西南大学
SOUTHWEST UNIVERSITY

GCCCE2024

第28届
全球华人计算机教育应用大会

The 28th Global Chinese Conference on Computers in Education

中国 重庆 | 西南大学(北碚校区)

Chongqing,China
Southwest University(Beibei Campus)

2024年6月1日—6月5日

June 01, 2024 —June 05, 2024



大会论文集(中文论文)

Main Conference Proceedings(Chinese Paper)



出版者：全球华人计算机教育应用学会
书名：第28届全球华人计算机教育 应用大会论文集 (中文论文)
作者：朱蕙君、孙丹儿、卢宇、孙之元、江绍祥、刘革平主编
出版年月：2024年 6月
版次：初版
ISSN：3005-3218

第 28 届全球华人计算机教育应用大会

The 28th Global Chinese Conference on Computers in Education

GCCCE 2024大会论文集

GCCCE 2024 Main Conference Proceedings

(Chinese Paper)

主编 Editors

朱蕙君 东吴大学
Hui-Chun Chu, Soochow University, Taiwan

孙丹儿 香港教育大学
Daner Sun, The Education University of Hong Kong, Hong Kong

卢宇 北京师范大学
Yu Lu, Beijing Normal University

孙之元 阳明交通大学
Jerry Chih-Yuan Sun, Yang Ming Chiao Tung University, Taiwan

江绍祥 香港教育大学
Siu Cheung Kong, The Education University of Hong Kong, Hong Kong

刘革平 西南大学
Geping Liu, Southwest University

联合编辑（依中文姓名筆畫順序排列）

Associate Editors (in Stroke orders)

朱志明 宜兰大学
Chih-Ming Chu, National Ilan University, Taiwan

朱高侠 南洋理工大学
Gaoxia Zhu, Nanyang Technological University

郑琨鸿 中兴大学
Kun-Hung Cheng, National Chung Hsing University, Taiwan

翟雪松 浙江大学
Xue-Song Zhai, Zhejiang University

吴俊育 阳明交通大学
Jiun-Yu Wu, National Yang Ming Chiao Tung University, Taiwan

涂芸芳 台湾科技大学
Yun-Fang Tu, National Taiwan University of Science and Technology

尚俊杰 北京大学
Jun-Jie Shang, Peking University

吕欣泽 政治大学
Owen H.T. Lu, National Chengchi University, Taiwan

陈真真 北京邮电大学
Zhenzhen Chen, Beijing University of Posts and Telecommunications, Mainland China

陈鹏鹤 北京师范大学
Penghe Chen, Beijing Normal University

责任编辑 Executive Editor

曾诗倩，东吴大学

Shih-Chien Tseng, Soochow University, Taiwan

目錄

一、 序言 MESSAGE FROM ORGANIZER XXI

二、 大会组织 CONFERENCE ORGANIZATION XXVI

三、 主旨演讲 KEYNOTES XXVII

主會議 Main Conference

C1 学习科学与计算机支持协作学习

1 使用基于渐进式提问的聊天机器人对大学生的艺术鉴赏能力、创造性思考倾向及学习满意度的影响

罗樾，张帝

9 生成式人工智能支持下的协作编程对学生批判性思维和编程群体自我效能的影响研究

严乙淼，陈创奇，叶新东

18 运用“HPS”+“PhET”实施科学教学的创新设计探索——以“地球圈层结构”教学活动设计为例

李冬，李悦飏

22 人工智能课程中协作脚本的设计与实践研究

列映湘，顾荣桢，黄桂闽，李绮华，钟柏昌

26 同伴审查策略和先验知识水平对学生计算思维和认知负荷的影响

梁晨，陈燊，林琳，赖文华，郭晓戈

30 促进大学生深度学习的在线同步同伴反馈研究

陈孝然，傅莉，朱敏捷，李栩晴

38 科学教育中生成性绘图的插图丰富度对注意力分配及学习效果的影响

章仪，何静，皮忠玲，杨九民

46 基于 AI 绘画的教学策略对学生写作成绩、写作想象力、学习动机的影响研究

郑益奇，杨刚，赵雨

56 以線上知識論壇學習對學生知識翻新環境與社群意識感知之影響

張芷瑄，薛智暉，蔡以寧，洪煌堯，洪國財

64 基于深度学习的在线异步讨论活动设计及实证研究
黄雅萱，彭炜明，邵文豪，汤名珍，杨时娇，周洋洋

69 智能技术赋能劳动教育评价创新的实施路径
梁心贤，陈晗，陈颖珊

73 大模型赋能高等教育的现状调查与对策分析
陆宇，张日悦

82 面向计算思维培养的学生生成问题策略构建及应用
张雪，牟歌，万海鹏

90 基於深度學習的瑜珈自動評量與回饋系統對大學生技能學習表現之影響
夏綠荷，黃國禎，黃展鵬，林彥男

98 深度學習增強藝術教育：運用人工智慧提升大學生藝術品鑑賞能力與繪畫技巧
邱敏棋，黃國禎，夏綠荷，李建輝

106 近三十年国内外学习空间研究进展——基于 1999-2022 的文献分析
刘思嘉，赵赫璇，付玉娜，马国威，毛晓龙，王妍莉

110 面向数字创造力培养的图形化编程学习活动研究
刘昺瑶，肖金爽，王伟

114 整合图示化支架的历史探究学习对学习者论证思维的影响研究
蔡慧英，孙佳悦

122 具身沉浸辨識的情境學習系統以提升學習者的學習成效
任家駿，吳建宇，王振漢，楊舒涵，趙伯堯，陳國棟

130 任务难度和元认知支架对学生协作编程学习成效的交互影响
李雯，李伟，曾秋蓉，刘城烨，黄佶艺，朱玉蝶，荣健康

138 “学生生成构想”教学理念应用于混合学习的历程分析
Lung Hsiang Wong, Thi Thu Ha Nguyen, Chew Lee Teo and Peter Seow

142 基于元认知支架的协作编程学习对不同性别学生学习成绩和元认知倾向的影响
周沫，李伟，曾秋蓉，黄佶艺，刘海英，季嘉炜，刘城烨

146 融合生成式 AI 的人机协同论证活动设计与应用
崔鑫，尤欣雅，马志强

154 教师与同伴组合式评价活动对评价质量与学习成绩的影响研究
张艺铧，陆雨菲，杨现民，李新

162 共享调节视角下计算机支持协作学习中的教师职能：基于系统性文献综述
顾佳博，孟柯颖，骆开燕，蓝敏

166 在 K-12 教育中 ChatGPT 应用的系统综述
叶碧滢，汪溢慧，王埕埕

172 不同动机水平的小学生对“人机双师”教学模式的技术接受度和满意度分析
孟晓培，贺丽媛，方建文

182 基于可视化协同辩论的反思性学习设计与效果研究
姚佳佳，闫千慧，刘明月

190 计算机支持的协作学习（CSCL）环境中的情感反馈：综述与展望
朱芳芳，杨九民，马志强

198 基于知识论坛写作的商务英语专业学生词汇复杂度与知识提升研究
雷春林，王玉，张俊杰

202 数智技术赋能 K-12 教育的研究综述
胡佩蓥，潘梦媛，林肖颖

206 学生协作学习过程中认知参与状态的自动分类方法
吴涵，张立山，王红晔，段腾飞，白志创，李雲帆

208 使用 GAFCC 游戏化模型提高在线协作学习背景下的学生参与度
Biyun Huang, Khe Foon Hew, Jun Peng, Yi Dai and Morris Jong

C2 移动、泛在与情境化学习

211 教中学策略中线上同伴在场方式对视频学习的影响
郭鑫，王婧然，徐珂，杨九民

219 基于场馆的 K-12 儿童非正式学习效果研究：系统性文献综述
刘倩，陈琦帆，罗恒

230 基于心流理论的虚拟现实学习模式会促进虚拟现实学习吗
梁友明，王陆

234 穿戴式虛擬實境結合合作概念圖即時互動回饋系統應用於中學英語閱讀教學：認知負荷對學習成效之影響
游師柔，盧瑩真，孫之元

238 虛擬實境下之全境式火場求生模擬演練學習系統
鄭樺嶸，區國良，唐文華

242 什么影响了一对一在线辅导的质量？基于在线辅导质量智能评估的分析视角
汪洋，赵梓明，陈玲

250 沉浸式虚拟现实如何影响学习体验与学习效果？——基于共情能力与虚拟同伴的视角
刘伊祉兰，兰翠玲，邢晓，宁王雪，余能保

258 B 站高传播力编程类课程的多组态路径分析
洪越洋，单俊豪

266 基于大语言模型的数字故事对学生批判性思维、认识论信念的影响研究
周雨欣，刘泽民，张文宇，单小份

271 基于增强现实技术的信息科技单元作业设计框架研究
张雪，魏雪峰，朱晓颖，朱梦雨，孙晓涵

275 支持學生於沉浸式虛擬實境社交互動並結合 STEM 與 SDGs 的學習環境開發
郭鋅恩，陳政煥

277 《氫朋好友》互動繪本在科學教育中的應用：以氫能源為例
湯梓辰，趙彥俞，林孟樺，周宏隆

279 基于 VR 技术的电磁波教学平台设计与交互式开发
张婉露，朱佳，潘昌藩，郭杭辉，陈张泽

C3 悦趣化学习、教育游戏与数字玩具

281 融入關鍵火災處理概念之教育遊戲—輔助國小學生火災教育議題設計探究
王曉璿，蔡承易，劉沛岑，張嘉益，楊逸捷

289 游戏化学习策略对大学生学习动机和学习投入的影响研究
张晓曼，李秀晗，刘巧伶，张颖

297 小学生 STEAM 项目化学习中科学探究能力培养的教育游戏设计
张蕊华，郭子霏，张新立，黄瑞婷

305 数学坐标游戏的设计与开发研究
相梦彤，张露，刘宇，尚俊杰

313 ARPG 與 Roguelite 遊戲對低學習成就國小學生的數學學習成效與學習投入之影響
謝榮益，游師柔，孫之元

317 游戏化学习理念下的小学数学教学活动设计与实施 —以“谁是卧底”之倍数与因数复习课为例

高佳慧，魏居林，张誉馨

321 游戏化学习在小学数学综合实践教学中的应用与开发研究—北京教育学院“教师培训”项目之一

刘月艳，王雪君，白永潇，曹慧，岳志刚，汤佳佩

329 理财素养实作模拟平台中操作行为与自我效能之关系研究

史安晴，李元萱，吴俊育

337 教育游戏嵌入发音对中职学生拼写及识别能力的影响研究

高志达，徐心远，乔建澍

345 学习动机如何影响国际中文教育游戏化学习投入？——基于滞后行为序列分析的实证研究

冯淼，方璐，杨国法，王抒诣，卫紫艳，张露，尚品言

353 紙筆遊戲不僅是遊戲：探討非數位國小數學教育的遊戲式學習

郭鋅恩，陳德懷

357 Exploring the Benefits of Incorporating Vital Energy-Saving Concepts into Online Game-Based for Elementary School Students Learning on Energy Issues 3

王曉璿，李思萱，邱光諺，陳俊霖，陳啟榮

361 整合實體展覽與遊戲式延展實境體驗於歷史課程之應用與研究

蔡遵弘，城菁汝

369 大腸直腸癌健康教育之 VR 教材開發與評估

劉怡姍，游君微，賴阿福

377 基于单元整体教学的编程游戏化学习活动设计 ——以高中《算法与问题解决》单元为例

林肖颖，赵晨畅

385 促进大学生高阶思维培养的协作推理游戏设计

张颖，李秀晗，刘巧伶，张晓曼

389 基于 Unity 2D 的小学数学教育游戏设计与实现 ——以《小数英雄》为例

张贤，赵源源

396 AR 技术赋能幼儿多角度高阶思维实证研究——基于北京军区机关幼儿园 AR 技术教学应用实践

段春梅，钱媛媛

404 面部表情分析方法在教育游戏中的应用
赵玥颖，张雨竹，张露，尚俊杰

412 融入環境議題概念之教育遊戲軟體輔助國小學生環境覺察課程學習效益探究
王曉璿，劉金英，楊淇鈞，林宥誼，李昀軒

416 指向计算思维培养的教育剧本游戏设计与实践研究 ——以浙教版五年级信息科技“算法的控制结构”单元为例
潘梦媛，黄璐，王运弘，章苏静

C4 高等教育与成人学习的技术应用、教师专业发展

420 面向成人学习者的医学知识学习系统的设计研究
陈守萍，陈梦洁

429 不同教学阶段中教师用语特征研究——LDA 主题模型和特征词共现网络分析
金宣萱，杨乐莹，许玮

437 核心素养视角下 AIGC 赋能的形成性评估模型的构建与展望
郝苗苗，彭俊

445 基于 TPB 的大学生专业认同感模型构建和检验
高欣羽，余俊，许玮

449 义务教育阶段教师智力流动意愿影响因素调查研究
曾颖红，林晓凡

457 趣創者理論在網頁設計課程上的應用
莊益瑞

465 中小学教师对教学视频使用意愿研究 ——基于整合的 TPB、TAM 和 UTAUT 模型
李妍，何静，章仪，刘彩霞，杨九民

473 基于多模态同伴反馈的同伴教学在师范生专业发展上的效果研究
何晓玲，徐刘杰，伍婧

480 生成式 AI 浪潮下教师发展的机遇与挑战：一种教师资格评价的新途径
张梦晨，陈彦君，周子骏

484 基于 ChatGPT 的人机协同教研中职前教师同伴反馈的社会网络分析
顾佳宁，李欣，蔡慧英

492 基于课堂视频的人机协同教研中教师交互模式的滞后序列分析研究
韩冰，蔡慧英

500 ChatGPT 与提问脚手架对师范生信息化教学设计的影响：一项混合研究
李荣娜，陈辉，陈琦帆，田金铭，杨九民

508 技术赋能教师职业幸福感提升：从技术观的视角出发
管乐，林易，文燕银，吴筱萌

512 沉浸式虚拟现实环境中存在感及认知情感因素对学习的影响机制研究 ——基于不同共情能力水平的比较
刘伊祉兰，兰翠玲，邢晓宁，王雪，余能保

520 多媒体学习中的积极情绪设计对学习效果的影响——基于生理数据的元分析
朱振凯，方苏，靳燕燕，王雅莉，彭晛

528 人工智能在教师教育中的应用：样态、挑战和优化策略
苗谦盈，黄禹，徐光涛

536 教育数字化转型背景下中小学多师协同教学方式的特征与思考
潘丹，杨明欢

544 生成式人工智能的个性化学习研究综述及展望——基于 Cite Space 可视化分析
刘小钰，代毅，庞何千慧，李可昕

552 数字原住民新手教师在个人学习环境中的专业发展模式： 基于扎根理论的一项探索性研究
林易，李姗姗，吴筱萌

556 融合技术的高校舞蹈多模态课堂教学行为的分析研究
张明，王海英，代毅

564 从海德格尔寻视理论看生成式人工智能对高等教育的影响 ——基于 ChatGPT 的实证研究
曾治，张誉月，卢晓东

572 人工智能时代教师数字素养评价研究—基于《教师数字素养》行业标准的思考
王旖，彭俊

580 移动教学对职前教师技术认知心理与 TPACK 发展的影响研究
吴莉霞，刘延申

582 基于 ITIAS 的小学数学课堂互动行为研究 ——以 2023 年度广东省“双融双创”示范课题为例

晏鑫，徐玲

584 人文关怀视角下开放教育教学管理的关怀阻滞及疏解路径

谭淇婧

586 数智赋能“做最好的自己”教师研修生态的构建研究

刘伟，钟芳，彭操红

588 澳门中学教师人工智能素养现状、分析及发展

凌智荣

C5 科技增强语言与人文学科学习

590 LLMs 在英语写作学习的应用：设计、实施和效果分析

郑伊彬，胡笑笑，叶新东

600 数字人文视域下国内外语言文学研究的对比分析

冀娇娇，任佩玲，武永

608 AI 協助英語對話機器人導入至國中——某國中英文課的課室觀察報告

余儘卿，洪霈羽，張立杰

616 虚拟现实技术应用于语言教育的系统性文献综述(2014 - 2023)

马萍，陈真真

624 跨越閱讀舒適圈：設計基於願望書單的聊書活動以培養學生的閱讀興趣

楊筱彤，廖長彥，陳德懷

628 5G 时代网络不良文化对儿童行为与认知发展的影响及对策研究

吴郑涵，杨清元，杨玉辉，赵亚萍，胡为民

632 大语言模型驱动的论证导学探索性研究

王红晔，张立山，吴涵

642 英语写作同伴反馈对学生反馈素养的影响研究

王殊雅，任佩玲，熊梦羽，郭紫钦

647 大语言模型在语言领域智能评估的探索——以外语对话场景为例

陈昂轩，贾积有

655 智能评分与修改：中文作文的提示框架开发与效果评测

刘怀亚，贾积有

663 基于生成式 AI 的小学英语跨学科项目化教学的实践研究

杨国超，刘小洁，姚晓芹，林秋纯，尹栩如，郑芳媛

667 基于 OBE 理念的现代教育技术课程混合式教学模式研究

张文馨，李西

674 開發 AI 英語自主學習課程規劃系統並探討其對教師態度之影響

劉奕帆，張心洧，陸以凡

682 当 ChatGPT 4.0 遇上医古文——以《黄帝内经·素问》英译为例

王子鹏，刘世英

690 基于语文核心素养的动画创作对小学生古诗学习成效的影响研究

李铭壹，蔡薇薇，王丹妮，张新立

696 演化博弈理论在教育研究中的应用分析

张盛，杨现民，李新

700 SPOC+产出导向法在国际商务谈判教学中的应用研究

蒋庆，李西，刘芮玲

C6 人工智能教育应用、智慧学习环境

704 国内教育界对生成式人工智能的研究热点与趋势——基于共词矩阵的知识图谱分析

任晨扬，陈婉琳，胡来林

712 ChatGPT 在教育中的研究热点分析

陈婉琳，任晨扬，胡来林

720 基于自然语言处理的深度学习课堂中教师话语特征分析与应用研究

曾钰琦，钟瑾，刘博

728 传统智慧与数字化转型：儒家教育思想在数字教育中的应用

刁均峰，连梅芳，韩锡斌

732 教师网络学习空间应用态度、应用行为、应用能力及应用效能关系的实证研究

张钰团，周洋洋，胡蓝心，邵文豪，任光杰

736 人工智能赋能课堂教学评价研究

易诗情，胡蓝心，周洋洋，王韵仪

740 基于 TAM 模型的中学生大语言模型平台使用意愿影响因素分析

许其鑫，梁燕，曾诗娅，李明蔚

749 基于 LSA 的小学人工智能课堂教学行为研究

宋玉婷，廖慧敏，张秀梅

757 作文说:大模型驱动的一对一作文智能导学系统

李雲帆，张立山，李万悦，廖俐丽，吴涵

765 数智技术赋能 5E 语文活动课生成性教学模式的构建与应用

陈良湾

769 基于微调大模型的教学视频辅助工具设计与研究

席少剑，吴子豪，杨烨枫，李锋

773 远程生理信号测量方法及其在教育领域应用综述

胡涛，魏艳涛，肖光润，潘东辉，徐琦，杨开

782 ChatPoet: 基于检索增强生成的诗词对话模型

崔咚咚，魏艳涛，肖光润，徐琦，董银锋

791 面向编程题解答的大语言模型性能强化研究

董银锋，魏艳涛，徐琦，肖光润，崔咚咚，潘东辉

800 基于 ARCS 模型的农村初中生 AI 课程学习活动设计研究

张乐佳

804 基于多模态数据融合的课堂学习情感识别研究现状及展望

杨开，魏艳涛，肖光润，潘东辉，徐琦，胡涛

809 生成式人工智能变革教育的实践逻辑：系统性综述与未来议程

陈小娇，胡哲冰，王成梁

817 结合教学属性的中小学古诗词图谱构建与分析

丁国柱，冯心茹

825 大学生人工智能技术使用意愿的影响因素实证研究

王浩名，陈小娇，戴坚，王成梁，徐显龙

833 数字技术赋能小学语文跨学科学习设计的探究

祝华轶

837 编程学习中基于大语言模型的 bug 整理与分析平台设计

刘发，高丹丹

845 教学视频中教师非言语行为的智能分析方法及应用研究

李妮辰，陈辉，何英杰，杨九民

853 课堂文本语言行为自动分析系统开发及应用

黄立新，卢卓豪，薛松

861 数字化学习资源关联网络研究综述：现状、问题及趋势

张惠影，李新，杨现民

869 基于大单元的特殊教育聋校自适应教学模式实践探究

米秀兰，邓太常

873 核心素养导向的个性化作业设计与应用研究

朱梦雨，魏雪峰，孙晓涵，张雪，朱晓颖

877 應用生成式人工智慧增強合作問題生成學習系統於國小社會課程之研究

洪丞威，洪暉鈞

885 自動問題生成之以知識追蹤為輔的適性程式設計學習初探

呂浚宏，洪暉鈞，楊叔卿

893 實體物件與數位雙生技術相結合的互動學習系統

游子諭，林孟衡，王振漢，楊舒涵，陳國棟

901 基于 Web of Science 的人工智能伦理研究关键词可视化分析

郭宝心，郑艺妍，廖林锟，张星语

909 深度学习视域下生成式人工智能支持小学生编程能力的实证研究

王丹妮，黄瑞婷，张新立，李铭壹

913 从错误中学习策略对高中生编程成绩和计算思维的影响研究

李琪，翁秋灵，方建文

918 数字化转型下基于分层学习资源的小学数学分层教学模式构建与应用

雷浩杰，郑光英，卜晓薇

926 生成式人工智能在职前教师跨学科教学设计中的应用研究

赵雨鑫，范志林，郑启航，叶乃方，卢宇

934 整合 AI 聊天機器人之學習管家數位平台培養學生恆毅力、成長型思維與提升學習成效

陳致諺，簡鈺笙，楊舒涵，王振漢，莊永裕，陳國棟

942 人工智能聊天机器人教育应用研究趋势与热点——基于 WoS（2014-2023）的文献计量分析

金哲宣，邹何林，王宇琳，唐颖

946 应用大型语言模型生成数学通用型提示的研究

张誉月，贾积有

950 生成式人工智能在教育中的应用—基于 Citespace 的可视化分析
白佳鹭，陈鹏，叶禹杉

952 支持教師基於個別學生閱讀歷程建構願望書單與聊書機制以延伸學生閱讀興趣
涂弘旻，廖長彥，陳德懷

954 国内外教育人工智能伦理可视化分析
贺玮，廖林锟，颜京莉，牛雯婷，赵东东

957 多模态大模型的教育应用价值、典型场景与发展建议
卜浩德，李新，杨现民

959 基于国际视野下的中小学 AI 课程内容设置建议
谭亚妮，王丽英

C7 学习分析与学习评估

961 从学习风格的角度探究在线协作学习中的批判性思维发展差异
李凤鸣，江鑫广，陈彦君，吴忭

969 設計閱讀成就目標系統輔助學生閱讀
葉彥呈，許喬珉，陳德懷

977 基于 iFIAS 的课堂师生互动行为研究——以初中信息科技优质课为例
庄凌，梁颖怡，徐玲

985 探討雙欄式步驟視覺化教學設計模式對不同先備知識學生之影響－以 AMA「幾何裝飾性風格」圖像創作為例
李佳蓉

989 基于改进 IEO 模型分析学生时间投入与数学成绩的影响机制：模型、算法与优化策略
谢晨，虞增艺，虞宸越

993 文學圈結合專題導向學習結合學習分析儀表板於書報討論創新教學研究
鄭博宇，洪暉鈞，楊叔卿

1001 信能數學：教師學習分析儀表板
周得揚，葉彥呈

1008 信息技术融合创新课堂教学互动现状研究——基于 2023 年广东省教育“双融双创”获奖视频的分析

徐玲，晏鑫，庄凌

1012 探索結合目標設定與視覺化學習分析之學習管理系統於 Python 教育之研究

柯炘德，洪暉鈞

1016 项目学习中协作问题解决能力的教学策略介入——基于 ENA 的循证分析

熊炀璐，金晨，邢倩雯，许玮

1024 生态学视角下学校疫情停课对学生综合素养的影响：揭示学校、家庭和个人支持的链式中介作用

虞增艺，陈诺言，金晟，徐刘杰

1032 基于数据挖掘的 MOOC 学习者行为特征与学习效果研究 ——以北京大学 H 课程为例

苏晗宇，贾积有

1036 大学生 MOOC 学习投入影响因素实证研究

马国威，赵赫璇，付玉娜，雅古，区文诗，王妍莉

1040 学习论坛中短文本语料的潜在语义信息挖掘

郑佳妮，王雯清，许玮

1048 多模态数据支持学习投入测评的系统性文献综述

曾佳尧，杨现民，李新

1056 基于深度学习的学生课堂行为识别方法研究综述

芪佩舒，魏艳涛，肖光润，潘东辉，徐琦，余游

1066 基于学习者反馈文本分析提升在线编程课程学习参与度研究

陈协玲，李冬梅

1070 ChatGPT 於計算機概論課程中的實際效果：測驗成效以及真實學習成效

吴嘉恩，邱嘉豪

1078 智能代理工具的关键技术与教育应用

刘倩，杨现民，李新

1080 科学虚拟实验中迷思概念的学习行为挖掘——以小学“简单电路”单元为例

李凤鸣，吴忭

1082 数学焦虑、自我效能感与数学素养的关系研究：考虑性别差异下自我效能感的中介作用及影响机制

张瑶，虞增艺，金泊西

C8 STEM 与创客教育

1084 以 STEAM 教育融入教育元宇宙之平台開發與研究
王政弘，吳聲毅

1092 不同模态叠加的教学视频对高中生编程内容学习的影响研究
徐珂，王瑞云，章仪，郭鑫，皮忠玲，杨九民

1100 元宇宙融入 STEAM 課程之科技接受模式探討
吳聲毅，王政弘，林在營，楊誠達

1104 小学 STEAM 课程设计的研究整合分析
张莹，段连鑫，谢利东，李志坚，杨家宽

1108 以 Scratch 軟體設計 SDGs 環境教育數位繪本之 STEAM 課程研究計畫
李曉旻，湯梓辰，陳瑞玲

1112 以虛擬實境教具運用在美術繪畫教學之認知負荷與心流經驗之研究
朱志明，林宜慶

1119 TPB 框架下中学生 STEM 行为投入的影响机制研究—— STEM 专业意向的中介作用
张欣欣，张屹，贾杰，王莹，赵苏楠，丁双婷

1127 6E 模式结合运算思维概念于 AIoT 学习活动对实作表现之研究
萧显胜，黄子芸，苏家莹，张玉山

1132 农村小学 STEM 教育实践的调查研究
杨翘楚，冯雨涵，李葆萍，朱立新

1136 国内外中小学 STEAM 教育研究文献分析
黄诘雅，李葆萍

1140 中學教師與 STEM 大屏教學
陳冠良，簡佑宏

1148 跨校分組的資訊教育課程對於國小學生運算思維能力及學習態度之影響
陳芷玄，張博裕，林秋斌，宋瑗玲，胡振猷

1156 PBL 中有效失败策略对小学生创造力影响研究
沈鑫明，张新立

1164 健康教育在线项目式学习课程设计——为社区居民设计个性化流感预防宣传册
陈柏因

1168 GPT/生成式人工智能与 STEAM 教育何以结合

沈柯，李锋

1172 戲劇學習對 STEM 機器人課程中建立負責任文化的影響：學生學習成效的探討

馮建中，王振漢，馮詩敏，馮詩婕，陳國棟

1180 中美 STEM 教育研究现状与热点的比较研究

陈博，唐瓷

1191 在线创客社区中的学习者学习情感研究 --基于文本情感分析的视角

叶龙俊，单俊豪

1199 指向跨学科思维的教学实践研究——以“制作 3D 校园建筑模型”项目为例

周千明，张涛，舒洁，秦琴，董连春

1203 基于 Mind+授课与智能菜园应用的教育创新研究

赵赫璇，马国威，付玉娜，孙伟华，毛云庆，王妍莉

1207 基于 STEAM 的项目式学习模式研究——以“创作同辈辅导视频”为例

姚瑶，伊亮亮

1209 培养职前教师跨科学教学专长的工作坊设计

付玉娜，马国威，赵赫璇，刘思嘉，毛晓龙，王妍莉

1211 基于 STEM 教育理念的“希罗喷泉”教学活动设计

高诗婕

1213 基于创客教育的跨学科主题学习设计研究

赵瑛，陈琦

C9 教育技术创新、政策与实践

1215 透过师范生生成视频中的隐喻表达揭示对教学中技术运用的认识
张祺晖，徐珂，章仪，杨九民

1223 教育数字化转型背景下网络学习空间促进学校变革的影响因素研究
王韵仪，黄雅萱，易诗情，胡蓝心，任光杰

1231 教育技术学公费师范生的困惑：职后为何走向分岔路
崔依冉，张书月

1239 后疫情时代大学生翻转学习需求对自我效能感的影响研究
叶禹杉，陈鹏，白佳鹭

1243 生成式人工智能辅助大学生自主学习的风险挑战及应对策略
路畅，田阳

1247 探討中小學教師使用平板數位教學行為意向相關因素之研究
鄭羽涵，林秋斌，楊肅健，高宏宣

1251 基于行为序列分析的双师课堂多元互动研究
王诗蓓，荆璐

1259 智能时代教育技术理论与实践新进展
罗文婧，邱艺

1263 数字化转型视角下教育评价的“变”与“不变”
宋琰玉，吕赐杰，赵笃庆，杨九妹，吴龙凯

1267 面向全纳教育的增强现实教学策略研究——以高中地理课程为例
刘嘉纯，张志兵，林晓凡

1271 “双减”政策的实施现状、问题成因及优化对策 -基于微博的数据分析
吴晓斐，梁琼月，万海鹏

1279 数字技术促进青少年心理健康的研究进展与启示 -基于 2013-2023 年文献的系统性综述和元分析
陈天娇，欧婧依，李格格，罗恒

1286 融入激励机制的小学编程类课程设计与实践研究
梁琼月，吴晓斐，万海鹏

1290 新技术赋能的爱国主义教育路径探析
章群璐，蓝靓

1294 共同体理论视角下研究生优质教学资源共享机制研究
曾佳尧，赵路瑶，李新，杨现民

1298 数字化何以赋能乡村教育高质量发展？ ——基于国内外经验与行动的综述
李永杰，张靖

1300 教师应用信息技术对数学课堂学习氛围的影响——基于不同信息技术参与度的序贯多元线性回归分析
宣申奥，虞增艺，景玉慧，徐刘杰

1302 课堂派支持下基于深度学习的单元作业设计模型构建
孙晓涵，魏雪峰，朱梦雨，朱晓颖，张雪

1304 基于 iFIAS 的课堂师生互动行为研究——以初中信息科技优质课为例
庄凌，梁颖怡，徐玲

1306 数字化转型背景下基于社会情绪能力培养的小学英语单元整体教学实践研究
丘秀文

一、 序言 Message from Organizer

全球华人计算机教育应用大会（Global Chinese Conference on Computers in Education, GCCCE）是由全球华人计算机教育应用学会（Global Chinese Society for Computers in Education, GCSCE）主办的双语（中文、英文）国际学术会议。该会议现已发展成为全球华人及非华人学者、教育工作者和政策制定者分享和交流计算机教育应用领域最新研究成果的年度盛会。第 28 届全球华人计算机教育应用大会（GCCCE 2024）于 2024 年 6 月 1 日至 6 月 5 日在西南大学举办。为了响应人工智能时代的到来及其对数字学习环境的影响，本届大会的主题定为“智能学习环境：重塑未来学习空间”（Reshaping the Future Learning Spaces with Intelligent Learning Environments）。

GCCCE 大会汇聚了来自世界各地的教育政策制定者、学者、教育工作者、校长及一线教师，互相交流与分享有关计算机教育应用的最新研究工作与成果。会议议程包括主旨演讲、论文报告、工作坊、博士生论坛、中小学教师论坛与企业参展等。同時，本次大會第五屆成立 EPT 英文分會，共收到 45 篇文章投稿。这显示了在过去五年中，已有许多非华裔研究者关注 GCCCE 研讨会，并吸引了众多英语系国家的作者前来交流与会。除此之外，本届 GCCCE 也延续上一届大会，安排两场英语主旨演讲，邀请两位顶尖国际学者主讲。

自 Covid-19 疫情以來，本届大会是首次取消线上参会，全体改为实体互动的模式。希望在经历了三、四年的疫情之后，通过本届大会的举办，能重新恢复学者之间的实体互动，增进全球参会者的情谊。本届 GCCCE 2024 大会包含 10 个会议主题：

C1 学习科学与计算机支持协作学习
(Learning Sciences & Computer-Supported Collaborative Learning)
C2 移动、泛在与情境化学习
(Mobile, Ubiquitous & Contextual Learning)
C3 悦趣化学习、教育游戏与数字玩具
(Joyful Learning, Educational Games & Digital Toys)
C4 高等教育与成人学习的技术应用、教师专业发展

(Technology in Higher Education & Adult Learning, and Teachers' Professional Development)

C5 科技增强语言与人文学科学习

(Technology-Enhanced Language and Humanities Learning)

C6 人工智能教育应用、智慧学习环境

(Artificial Intelligence in Education & Smart Learning Environments)

C7 学习分析与学习评估

(Learning Analytics & Assessments)

C8 STEM 与创客教育

(STEM & Maker Education)

C9 教育技术创新、政策与实践

(Educational Technology: Innovations, Policies & Practice)

EPT 英文论文分会

(English Paper Track)

各子会议均设有主席、副主席、委员，个别子会议另外设置了评审委员以及顾问。GCCCE 2024 中文论文向全球华人征稿，英文论文开放全球投稿。本届会议共收到 692 位作者的 375 篇投稿。这些论文投稿来自中国、台湾、香港、新加坡、澳门等 9 个国家和地区。投稿作者所在地区如表 1 所示。

表 1 GCCCE 2024 九个子会议与英文分会投稿作者所在地统计

地区	中国大陆	台湾	香港	新加坡	澳门	哈萨克	南韩	澳洲	智利
人数	631	116	22	20	11	4	2	1	1
投稿篇数	310	47	10	11	7	0.4	1	1	1
接受篇数	228	46	7	7	6	0.4	1	1	1

所有收到的论文，首先分配给至少 3 位委员进行初步审阅，初步审阅的结果再由相关子会议主席与副主席进行复议后才确认最终的论文审阅工作，经过这样严格的论 文评审过程，本届会议各子会议最终录取共 253 篇论文（见下表），录用率 67.4%。最后推荐了 6 篇最佳中文论文（限九个子会议录取的长论文）、5 篇最佳技术设计论文和 5 篇优秀学生论文（限九个子会议与英文分会录取的长论文或短论文）、4 篇最佳英文论文（限英文分会录取的长论文）。各子会议

投稿与录用状况如表 2 所示。

表 2. GCCCE 2024 各子会议及英文分会录取论文统计

	投稿	录用总数	长文	短文	海报論文	淘汰	錄取率
C1:学习科学与计算机支持协作学习	46	33	17	14	2	9	72%
C2:移动、泛在与情境化学习	20	13	5	5	3	7	65%
C3:悦趣化学习、教育游戏与数字玩具	31	22	14	8	0	9	71%
C4:高等教育与成人学习的技术应用、教师专业发展	37	28	16	6	6	9	76%
C5:科技增强语言与人文学科学习	22	17	11	5	1	5	77%
C6:人工智能教育应用、智慧学习环境	75	43	23	15	5	32	57%
C7:学习分析与学习评估	35	23	11	9	3	12	66%
C8: STEM 与创客教育	36	26	12	8	6	10	72%
C9:教育技术创新、政策与实践	28	21	7	9	5	7	75%
EPT 英文论文分会	45	26	12	10	4	19	58%
小计	375	252	128	89	35	123	67%

本届会议邀请了西南大学教育学部罗生全副部长、加拿大多伦多大学 Jim Slotta 教授、新加坡南洋理工大学陈文莉教授、台湾台中教育大学黄国祯教授等四位业界学界的专家学者进行四场主旨演讲，分别為：

主旨演讲 1：教学大数据应用与教学变革

演讲人: 罗生全，教授，博士生导师。西南大学教育教师学院院长

主旨演讲 2: The role of emerging technologies in 21st century education: Supporting multidisciplinary inquiry communities

演讲人: Jim Slotta, Professor, University of Toronto

主旨演讲 3：协作学习的多模态学习分析：希望与挑战

演讲人: 陈文莉，副教授，新加坡南洋理工大学

主旨演讲 4：人工智能在教育的应用议题与研究设计

演讲人：黄国禎，讲座教授兼副校长，台湾台中教育大学

这四场主旨演讲、九个子会议、英文分会，和三场多人座谈会形式的论坛（包括一场主会议特邀高端论坛与两场子论坛），组成了大会的主会议。主会议特邀高端论坛的主题为：「人工智能赋能教育创新」。子论坛一题目为「教育大模型

的研究与实践-生成式人工智能的创新应用」，子论坛二题目则为「AI 驱动的教育研究创新：思路启发与方法探索」，均为人工智能时代非常重要的讨论题目。

跟往届大会一样，本届大会还设有三项会前活动：中小学教师论坛、工作坊和博士生论坛。本届教师论坛共收录了来自中国大陆及港台地区的教师论文 39 篇，并推荐了 4 篇优秀教师论坛论文参与最佳中小学教师论文奖评选（限中小学教师论坛录取的长论文或短论文）。

工作坊部分也相当踊跃，本届大会共设立 9 个工作坊，最终收录论文 124 篇 ，这 9 个工作坊的主题分别叙述如下。

工作坊 1：迎接「元宇宙」的世代，如何融入新科技于教学工作坊

工作坊 2：教育大语言模型工作坊

工作坊 3：学习投入与学习行为建模工作坊

工作坊 4：数字游戏式学习(Game-based Learning)与游戏化(Gamification)教学策略运用于关键力教育工作坊

工作坊 5：电脑支援个人化与合作学习

工作坊 6：创新互动回馈科技提升学习动机

工作坊 7：人机协同视域下的协作学习设计、分析与评价

工作坊 8：中小学新一代人工智能教育:培育未来的创新者

工作坊 9：人工智能赋能知识建构创新研究与实践

另外，本届大会的博士生论坛，有 10 位博士生投稿，6 位导师专家的参与，分为三个部分，分别为短演讲、报告与讨论，以及一对一私人指导会议。

大会谨此向协作本届会议召开的所有人员致谢。我们衷心地感谢九个子会议、英文分会、工作坊、教师论坛、博士生论坛的主席、副主席、议程委员会委员，以及各组委会成员在会议筹备期间的帮助，并特别感谢大会主席的指导和支持会议统筹工作。

我们衷心希望大家能够享受 GCCCE 2024 的各项大会活动，从中获得丰富的启发和美好的体验。让我们共同努力，打造一个更具韧性、更具国际化的 GCCCE 学术社群，并将 GCCCE 的精神薪火相传，生生不息的传递下去。

江绍祥 香港教育大学(香港)
大会主席

朱蕙君 东吴大学(台湾)
国际议程协调主席

孙丹儿 香港教育大学(香港)
国际议程协调主席

卢宇 北京师范大学(大陆)
国际议程协调主席

孙之元 阳明交通大学(台湾)
国际议程协调主席

二、 大会组织 Conference Organization

主办单位 Organizer:

全球华人计算机教育应用学会

Global Chinese Society for Computers in Education (GCSCE)

承办单位 Host:

西南大学 Southwest University

西南大学教育学部 The Faculty of Education, Southwest University

大会主席 Conference Chair:

江紹祥 香港教育大学

Siu Cheung Kong, The Education University of Hong Kong

大会顾问 Consultants:

陳德懷 中央大学

Tak-Wai Chan, National Central University, Taiwan

施如齡 中央大学

Ju-Ling Shih, National Central University, Taiwan

国际议程协调主席 International Program Coordination Chair:

朱蕙君 東吳大学

Hui-Chun Chu, Soochow University, Taiwan

国际议程协调副主席 International Program Coordination Co-Chairs:

孫丹兒 香港教育大学

Daner Sun, The Education University of Hong Kong

盧 宇 北京师范大学

Yu Lu, Beijing Normal University

孫之元 陽明交通大学

Jerry Chih-Yuan Sun, National Yang Ming Chiao Tung University, Taiwan

组织委员会主席 Local Organizing Chair:

劉革平 西南大学

Geping Liu, Southwest University

组织委员会成员 Members of the Organizational Committee:

謝 濤 西南大学

Tao Xie, Southwest University

韩继宁 西南大学

Jining Han, Southwest University
梁　樱　西南大学
Ying Liang, Southwest University

中小学教师论坛主席 K-12 Teachers Forum Chairs:

宋燕捷 香港教育大学
吴　忭 华东师范大学
唐　颖 西南大学
吴骏杰 澳门理工大学
陈志鸿 台中教育大学
袁光姬 南洋理工大学

中小学教师论坛副主席 K-12 Teachers Forum Co-Chairs:

金伟明 香港计算机教育学会
文可为 东华三院郭一苇中学
张　思 华中师范大学
程妙婷 深圳大学
贾程媛 浙江大学
韦　卫 澳门理工大学
赖秋琳 台北教育大学
沈明勋 高雄师范大学
王其云 南洋理工大学

工作坊主席 Workshop Coordination Chair:

游师柔 阳明交通大学

工作坊副主席 Workshop Coordination Co-Chair:

刘子韬 暨南大学

博士生论坛主席 Doctoral Forum Chair:

孙丹儿 香港教育大学

博士生论坛副主席 Doctoral Forum Co-Chairs:

简淑芸 台北商业大学
杨玉芹 华中师范大学
詹　颖 香港教育大学

子会议议程委员会 Sub-Conference Program Committees

C1: 学习科学与计算机支持协作学习(Learning Sciences & Computer-Supported Collaborative Learning)

执行主席 Executive Chair

涂芸芳 台湾科技大学

副主席 Co-Chairs

余　亮 西南大学

马志强 江南大学

方建文 温州大学

C2: 移动、泛在与情境化学习(Mobile, Ubiquitous & Contextual Learning)

执行主席 Executive Chair

郑琨鸿 中兴大学

副主席 Co-Chairs

姜　强 东北师范大学

连晶晶 北京邮电大学

赖秋琳 台北教育大学

苏　友 北京邮电大学

C3: 悦趣化学习、教育游戏与数字玩具(Joyful Learning, Educational Games & Digital Toys)

执行主席 Executive Chair

尚俊杰 北京大学

副主席 Co-Chairs

庄绍勇 香港中文大学

张　露 北京邮电大学

C4: 高等教育与成人学习的技术应用、教师专业发展(Technology in Higher Education & Adult Learning, and Teachers' Professional Development)

执行主席 Executive Chair

翟雪松 浙江大学

副主席 Co-Chairs

代　毅 澳门城市大学

彭　睨 华中师范大学

刘　智 华中师范大学

C5: 技术增强语言与人文学科学习(Technology-Enhanced Language and Humanities Learning)

执行主席 Executive Chair

陈真真 北京邮电大学

副主席 Co-Chairs

贾积有 北京大学

黄德铭 香港都会大学

周秋惠 清华大学

陈维超 美国独立研究人员

刘世英 重庆邮电大学

C6: 人工智能教育应用、智慧学习环境(Artificial Intelligence in Education & Smart Learning Environments)

执行主席 Executive Chair

陈鹏鹤 北京师范大学

副主席 Co-Chairs

廖　剑 西南大学

朱　佳 浙江师范大学

刘泽民 新加皮国立大学

王　琦 北京外国语大学

C7: 学习分析与学习评估(Learning Analytics & Assessments)

执行主席 Executive Chair

吴俊育 阳明交通大学

副主席 Co-Chair

杨现民 江苏师范大学

范逸洲 北京大学

冯诗惠 香港大学

C8: STEM 与创客教育(STEM & Maker Education)

执行主席 Executive Chair

朱志明 宜兰大学

副主席 Co-Chairs

吴声毅 屏东大学

陈成志 新加坡国立教育学院

黄家伟 香港大学

傅　骞 北京师范大学

C9: 教育技术创新、政策与实践(Educational Technology: Innovations, Policies & Practice)

执行主席 Executive Chair

朱高侠 南洋理工大学

副主席 Co-Chairs

万海鹏 首都师范大学

林倍伊 清华大学

黄凌云 香港教育大学

英文分会议程委员会 (English Paper Track Program Committee)

执行主席 Executive Chair

吕欣泽 政治大学

副主席 Co-Chairs

Brendan John 京都大学

廖宜虹 屏东大学

汤凯喻 中兴大学

三、 主旨演讲 Keynotes

主旨演讲 1

- 2024 年 6 月 3 日（星期一）09:30-10:30 AM
- 会场：音乐厅

Speech title：教学大数据应用与教学变革



罗生全 教授，博士生导师
西南大学教師教育學院院長

Speech Abstract:
随着科技的不断发展，教育领域也在逐步融入大数据技术。教学大数据应用对于教育的变革起着至关重要的作用。罗生全教授将探讨教学大数据在教育领域中的应用，并阐述其对教学变革的积极影响。

Speaker Bio:
罗生全，教授，博士生导师。西南大学教育学部副部长。国家级人才称号获得者，全国课程学术委员会副秘书长，中国教育学会中小学综合实践分会常务理事，教育部基础教育教学指导委员会综合实践活动指导专委会委员，重庆市教育学会智慧教学专业委员会理事长，《Future in Educational Research》国际期刊主编全国“最具影响力青年学者”入围者。重庆市“巴渝学者”特聘教授。重庆市哲学社会科学领军人才。重庆市学术技术带头人。重庆市青年专家工作室领衔专家。美国密歇根州立大学和英国伦敦大学学院访问学者。主要研究教育学原理、教师教育和课程与教学论。在《光明日报》《教育研究》等报刊上发表学术论文 200 余篇，《教育研究》9 篇，《新华文摘》长文转载 4 篇；主持国家社科基金项目 3 项（其中重点 1 项），国际合作项目 1 项，省部级项目 20 余项；出版和编撰《新课程下的教学方式转变》《学业负担论纲》《学业负担问题解决：模型建构与治理机制》等学术著作十余部；获得教育部高等学校科学研究（人文社会科学）优秀成果奖一等奖、国家级教学成果奖、“明远教育奖”、钱学森城市学（教育）金奖提名奖、重庆市社会科学优秀成果奖、重庆市教学成果奖一等奖和二等奖等 10 余项。主要研究教育基本理论、基础教育课程改革、教材建设、教师专业发展等。

主旨演讲 2

- 2024 年 6 月 3 日（星期一）14:00-15:00 PM
- 会场：田家炳教育书院 802

Speech Title: The role of emerging technologies in 21st century education: Supporting multidisciplinary inquiry communities



Jim Slotta, Professor
University of Toronto

Speech Abstract:

This talk will explore recent trends in technology, media, and information, including a discussion of advances in artificial intelligence, physical computing, and classroom portal technologies. I will discuss the importance of fostering 21st-century competencies, and the need for education to prepare students as citizens, critical thinkers, problem, solver, and communicators. I will present recent advances in my own research of the Critical Action Learning Exchange (CALE), including a review of recent collaborations with teachers in South China. I will also discuss the importance of developing inquiry-oriented learning environments, including an important role for the physical and social classroom environments. I will argue that 21st century technologies are best suited for new forms of 21st century learning, and that we have new opportunities to care for all students, as they form identities as learners and future citizens.

Speaker Bio:

Jim Slotta is Professor and President's Chair in Knowledge Technologies and Education at the Ontario Institute for Studies in Education, University of Toronto. Since 2005, Dr. Slotta has led a team of students, designers and developers to investigate new models of collaborative and collective inquiry. These studies have advanced a pedagogical model known as Knowledge Community and Inquiry (KCI), in which students and teachers collaborate as a learning community to engage in STEM inquiry projects. Slotta currently directs the ENCORE lab (http://encorelab.org) in which KCI curriculum and technology environments are developed and researched. From 2006 - 2011, Slotta served as Canada Research Chair in Education and Technology, serving as

PI or co-I on more than 30 funded projects and supervising 20 doctoral and post-doctoral researchers. In 2019, Slotta launched the Critical Action Learning Exchange (CALE), where teachers develop, exchange and discuss competency-centered curriculum that empowers students as learners, providing meaning and purpose to their schooling experience and scaffolding their formation of learner and career identity.

主旨演讲 3

- 2024 年 6 月 4 日（星期二）08:30-09:10 AM
- 会场：田家炳教育书院 802

Speech Title: 协作学习的多模态学习分析：希望与挑战



陈文莉 副教授
南洋理工大学国立教育学院学习科学与评量学部主任

Speech Abstract:

多模态学习分析（MMLA）旨在通过测量、收集、分析和报告有关学习者及其环境的多种类型和模态的数据，全面理解和优化学习及其环境。陈文莉教授的研究团队借鉴学习科学和认知神经科学理论和方法，在协作学习环境中收集和分析学习者的多种模态数据，以更全面地理解和优化协作学习的设计与实施。这些多模态数据包括认知（近红外光谱脑功能成像）、视觉注意力（由眼动追踪仪）、以及学习行为（学习者参与协作学习活动时的口头、文本、手势等）等数据。这些多模态数据为科技支持的学习环境中的学习分析提供了新的视角与思路，通过测量与分析协作者之间的大脑同步性和共同注意力，用更为科学的方法研究协作学习中学习者之间的协同程度。另外，这些多模态数据不但帮助教育者与研究者更全面地了解协作学习的结果，而且对学习者的协作过程与轨迹提供更详细的解读，这种细粒度的学习分析对协作学习的优化设计很有价值。陈教授会通过实例具体介绍多模态学习分析的理论与方法，以及多模态学习分析在计算机支持的协作学习领域的应用。此外，本演讲还会说明多模态学习分析的前景与挑战。

Speaker Bio:

陈文莉博士现为新加坡南洋理工大学国立教育学院学习科学与评量学部主任，副教授。陈文莉多年来深耕于计算机支持的协作学习与学习分析领域的研究与实践，近 20 次在各国际会议获得最佳论文奖。她 2021 年荣获亚太计算机教育应用学会“杰出研究者奖”，是南洋理工大学最高教育奖项“南洋教育奖”得主。陈教授在 2016 至 2021 年担任国际学习科学学会计算机支持的协作学习委员会联合主席，她目前担任亚太教育计算机学会和全球华人教育计算机学会的执行委员。她曾担

任 2022 年国际计算机支持的协作学习会议、2017 年国际计算机教育应用会议与 2014 年全球华人计算机教育应用大会的议程委员会主席，以及 2016 年国际学习科学会议和 2012 年国际计算机教育应用会议的组织委员会主席。目前，陈教授担任 Journal of Computers in Education (Scopus, Q1; ESCI)和 Learning: Research and Practice (Scopus, Q2)两本期刊的主编，也是 Instructional Science (Scopus, Q1; SSCI)、Asia Pacific Journal of Education (Scopus, Q2; SSCI) ,和 Research and Practice in Technology Enhanced Learning (Scopus, Q1)三本期刊的副主编。此外，她还是 International Journal of Computer-Supported Collaborative Learning (Scopus, Q1; SSCI)的编委。

主旨演讲 4

- 2024 年 6 月 5 日（星期三）11:00-12:00 AM
- 会场：田家炳教育书院 802

Speech Title: 人工智能在教育的应用议题与研究设计



黄国禎 教授
台中教育大学 讲座教授兼副校长

Speech Abstract:

近年来人工智能快速的发展已经受到广泛的重视;与人工智能在教育应用的相关议题更是引起学者们的热烈讨论。然而，如何将人工智能应用在教学现场，并分析其带来的优势，仍然是一个具有挑战的议题。对于技术背景的学者而言，可能在实验设计及数据分析方面遭遇到困难；另一方面，教育背景的学者可能会感受到技术面的挑战。

本演讲主要介绍人工智能与数据分析技术在教育的应用、研究议题及发展方向。黄教授将透过实际在教学现场应用的实例，来说明不同背景的学者应如何进入人工智能教育应用的研究。除了说明人工智能及数据分析对于学习者、教学者，以及教育决策者带来的助益，黄教授也会针对可能的研究设计及评量方式提供具体的说明。本演讲对于有志于发展人工智能教育应用的研究人员极具参考价值。

Speaker Bio:

黄国禎教授是台中教育大学 讲座教授 兼 副校长。黄教授的学术专长包括行动与泛在学习、遊戏化学习、翻转学习及人工智慧教育应用。曾主持 150 多个专案计画，并多次获得研究奖励，包括 2007、2010 及 2013 年「科技部杰出研究奖」，更获得 2015 年「杰出资讯人才奖」及 2019 年「教育部师鐸奖」。在 2022 年，获得国科会杰出特约研究员的殊荣。

使用基于渐进式提问的聊天机器人对大学生的艺术鉴赏能力、创造性思考倾向及学习满意度的影响

Effects of using a progressive prompting-based chatbot on students' art appreciation, creative thinking and learning satisfaction

罗樾[1]，张帝[2]
[1]广东技术师范大学 美术学院
[2*]浙江师范大学 教育学院
[1]yueluo@gpnu.edu.cn
[2*] zhangdi0909@126.com

【摘要】聊天机器人可以通过自然语言处理为学生提供实时反馈，帮助他们进行认知建构。然而对于不同知识水平的学生，使用相同交互逻辑的聊天机器人无法在他们遇到困难时提供适性化的支持和反馈。因此本研究提出一种基于渐进式提示的聊天机器人应用在大学设计课程中提升教学效能。为了评估所提出方法的有效性，本实验进行了三周准实验研究。实验组学生(N = 29) 使用基于渐进式提示的聊天机器人进行学习。控制组学生(N = 29) 使用传统的聊天机器人进行学习。结果表明，实验组学生在艺术鉴赏能力和创造思考方面显著高于控制组学生。同时，两组学生对于学习策略的满意度无显著差异。

【关键词】 聊天机器人；渐进式提问；艺术鉴赏；创造性思考

Abstract:*Chatbots can provide real-time feedback to students through the natural language processing algorithm and help them with cognitive construction. However, chatbots with the same interaction logic cannot provide adaptive support and feedback to students with different knowledge levels when they encounter difficulties. Therefore, this study proposes a progressive prompting-based chatbot to enhance teaching effectiveness in university design courses. To assess the effectiveness of the proposed approach, a quasi-experiment was conducted for 3 weeks. The experimental group (N = 29) used a progressive prompting-based chatbot (PP-Chatbot), and the control group (N = 29) used a conventional chatbot (C-Chatbot). The results showed that students in the experimental group had significantly higher levels of art appreciation and creative thinking than students in the control group. There was no significant difference in the satisfaction of learning strategy between the two groups.*

Keywords: Chatbot, Progressive prompting, Art appreciation, creative thinking

1. 简介

设计是一种以服务人类的需求和赋予生活意义来改变对象或环境的方式(Yeh & Peng, 2019)。而艺术鉴赏引导着设计的进化趋势,帮助人们培养个人表达能力和发展世界意识(Hung & Young, 2017)。 研究表明通过创造性实践来理解和学习艺术概念能够让学习者更好地体验创作者的表达和情感(Palau-Pellicer et al., 2019)。而交流和协作能够刺激学习者激发审美体验，提高学生对视觉符号及文化差异的理解(Irwin, 2018)。 研究者指出，实现这一目标的一个潜

在解决方案是在学生认知建构过程中提供个性化学习指导和反馈(Schrader & Grassinger, 2021)。

随着大数据及人工智能的快速发展，聊天机器人(chatbot)与学习的结合逐渐成为教育者关注的课题之一。聊天机器人是一种通过自然语言处理算法模仿人类交谈方式为使用者提供对话体验的计算器程序(Deveci Topal et al., 2021)。与实时通讯软件相似，聊天机器人可以基于语音或文字识别，并遵循内置的规则和流程来与用户进行交互(Smutny & Schreiberova, 2020)。用户提供的信息能够让聊天机器人了解偏好、提出建议甚至预测需求(Deveci Topal et al., 2021)。然而技术增强型学习(需要搭配合适的学习模型，例如建模、辅导和鹰架 (Chen et al., 2021)。Schrader and Grassinger (2021)指出，根据学生当前的理解水平提供适性化的学习支持对于构建有效学习环境非常重要。大部分教育聊天机器人的行为和回答都是机械的，缺乏适性化的支持和反馈(Smutny & Schreiberova, 2020)。除此之外，如果缺乏适当的学习策略，无法最大化聊天机器人的教学效果(Lin & Mubarok, 2021; Yin et al., 2021)。

提示可以让学生反思学习进度，并引导他们在保持正确的方向并找到找到问题解决方案(Chen et al., 2014)。而渐进式提示策略则是根据学生的能力和认知状况给与不同程度的提示并达到适性化学习的目标(Chen et al., 2021; Schrader & Grassinger, 2021)。在本研究中，我们开发了一个基于渐进式提示的聊天机器人，通过聊天和互动提供不同程度的提示，并帮助学生提升艺术鉴赏能力。为了评估其有效性，本研究对以下研究问题进行了调查。

1. 使用基于渐进式提示的聊天机器人(PP-Chatbot)的学生是否比使用传统聊天机器人(C-Chatbot)的学生显示出更好的艺术鉴赏能力？
2. 使用 PP-Chatbot 的学生是否比使用 C-Chatbot 的学生具有更高的创造性思维倾向？
3. 使用 PP-Chatbot 的学生是否比使用 C-Chatbot 的学生具有更高的学习模式满意度？

2. 文献回顾

2.1. 聊天机器人

聊天机器人在 1960 年代被应用在教育领域作为数字学习环境中的教学代理(Laurillard, 2013)。如今，聊天机器人大多应用人工智能技术开发并被当做是辅助学习和增强交互的新技术(Fidan et al., 2022)。根据使用目的的不同，聊天机器人可以被分为任务式和非任务式(Hussain et a., 2019)。任务式聊天机器人为了完成特定任务而设计，用于为使用者提供服务或信息；非任务式聊天机器人为了模拟随意的对话而设计，用于提供创造性和娱乐性的闲聊。受限于学生人数的增加，传统课堂中教师无法为学生提供一对一的辅导(Vázquez-Cano et al., 2021)。而聊天机器人可以创造一对一的聊天环境，适合作为助教或是学习者在教育系统中的合作伙伴(Chen et al., 2020; Deveci Topal et al., 2021)。

然而聊天机器人也教育领域也面临许多挑战。例如学生的个体差异可能会影响聊天机器人的有效性(Yin et al., 2021)。设计者需要在建构教育聊天机器人时将个性化的学习支持融入到环境中来满足不同学习能力学生的需求(Chaves & Gerosa, 2021)。Luger and Sellen (2016)也指出大多数研究都在尝试提高聊天机器人的功能性和准确性，然而大部分开发的教育聊天机器人无法满足用户的交互期望。这些聊天机器人缺乏使用学习策略支持的交互或沟通模式，因此对不同学生的可访问性没有显著差异(Smutny & Schreiberova, 2020)。对于程度较差的学生来说，缺乏足够的外部支持可能会让他们在与聊天机器人的交互中感觉到沮丧和不满(Luger & Sellen, 2016)。如果能够增加聊天机器人的主动性并提供额外的信息来干预学习者的认知建构，不仅可以增加学生对于聊天机器人的兴趣，还能够带来更好的学习效果(Chaves & Gerosa, 2021)。因此设计者需要考虑如何使用合适的学习策略提高聊天机器人的教学技能(Chang et al., 2021)。

2.1. 基于渐进式提示的学习

渐进式提示策略是一种逐步向学习者提供帮助来提高他们的学习表现的方法(Yang et al., 2018)。如果学生缺乏必要先备知识或自我调节能力，就很难根据自己的学习弱点改善学习状

况(Sun et al., 2018)。渐进式提示策略的目标就是在学生在学习活动中遇到困难的时候提供适当的外部支持帮助他们思考如何走出困境(Chen et al., 2021)。当学习者在任务中受挫后，他们会得到抽象或简单的提示来引发思考。如果他们仍然无法解决问题，得到的提示将会变的越来越具体，直到像学习者示范完整的问题解决方案。而提示的数量或程度取决于学习者的学习表现或能力(Yang et al., 2018)。

规范渐进式提示的等级和内容会影响学生收到反馈的有效性。对能力较低的学生来说，抽象的提示无法帮助他们解决困境，导致他们产生挫折感和降低自我效能(ter Vrugte et al., 2017)。学生往往会认为这种提示是无关紧要的，并且倾向忽视他们(Chen et al., 2014)。另一方面，过于具体的提示可能会让学生对完整的问题解决方案产生依赖(Sun et al., 2018)。Wilson et al. (2012) 也强调了高层次的提示和高层次的回答之间具有较高的相关性。因此规范学生在遇到困难时获得不同类型的提示十分重要。

随着信息科技的进步，渐进式提示策略被应用在许多教育科技中并起到了积极的效果。例如 Chen et al. (2014) 开发了一种基于渐进式提示的情境感知学习方法 并应用在小学自然科学课程中。研究结果表明，与单阶段的提示相比，具有三个阶段提示的情景感知方法能够有效地提高学生的学习表现。同时也能够促进学生在具有挑战性的任务中对学习内容保持关注。Yang et al. (2018) 则提出了一种基于渐进式提示的教育游戏应用在数学课中。他们设计的游戏能够在学生第一次答错时给出文字提示，第二次答错时给出图形解释。结果表明实验组学生的数学成绩有显著提高。除此之外，渐进式提示策略也可以被应用于 VR 中来促进 EFL 学习者的口语发展、改善他们的互动和自我效能(Chen et al., 2021)。基于以上的研究结果，渐进式提示策略有望成为提升聊天机器人教学效果的鹰架。

3. 渐进式聊天机器人开发

在这项研究中，我们提出了一种基于渐进式提示的聊天机器人，通过逐步给学生提供相应的提示来帮助他们解决关于艺术鉴赏领域的问题。聊天机器人系统是使用 IBM Watson Assistant 开发，这是一种使用 AI 技术建制虚拟聊天机器人的平台。如 Fig. 1 所示，学习系统包括聊天机器人单元、提示单元以及数据库单元。聊天机器人单元包括交互逻辑以及对话文本。提示单元包括文本形容提示和图片形容提示。数据库则存储学习材料、题库、学习行为和提示内容。系统会记录学生在回答问题时的认知状况，并提供学习指导和问题提示。

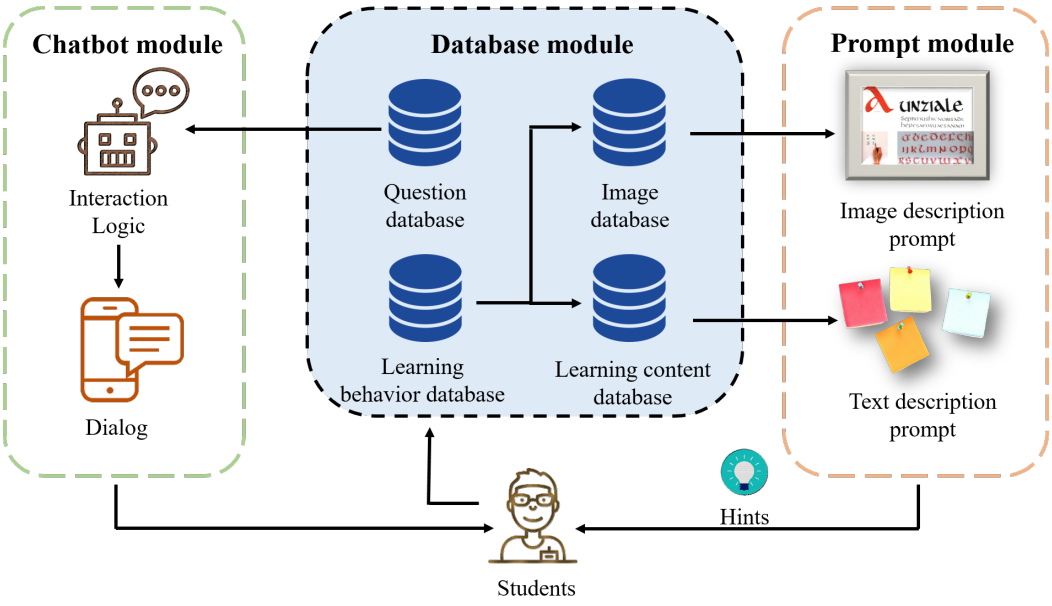


Fig.1 The structure of learning system

Fig.2 展示了聊天机器人制作平台以及学生使用聊天机器人进行学习的界面。学生可以使用移动设备打开网址来进入学习界面与聊天机器人对话。在交互过程中聊天机器人可以识别学生的回答并根据规则给出对应的回馈。学习文本以及问题基于大学课程中设计课程发展而来，目的希望让学生正确掌握中西文字体的分类、样式以及在艺术作品中的应用。在实验组中，学生第一次回答错误答案时，系统首先会提供正确的概念帮助学生纠正错误。当学生再次回答错误时，聊天机器人则会使用图像描述来进行提示，如 Fig.3 所示。在控制组中，当学生第一次回答错误时，聊天机器人不会给出任何提示。只有当学生第二次回答错误时，他们才会一次性得到文字描述和图像描述的提示。



Fig.2 The interface of Chatbot

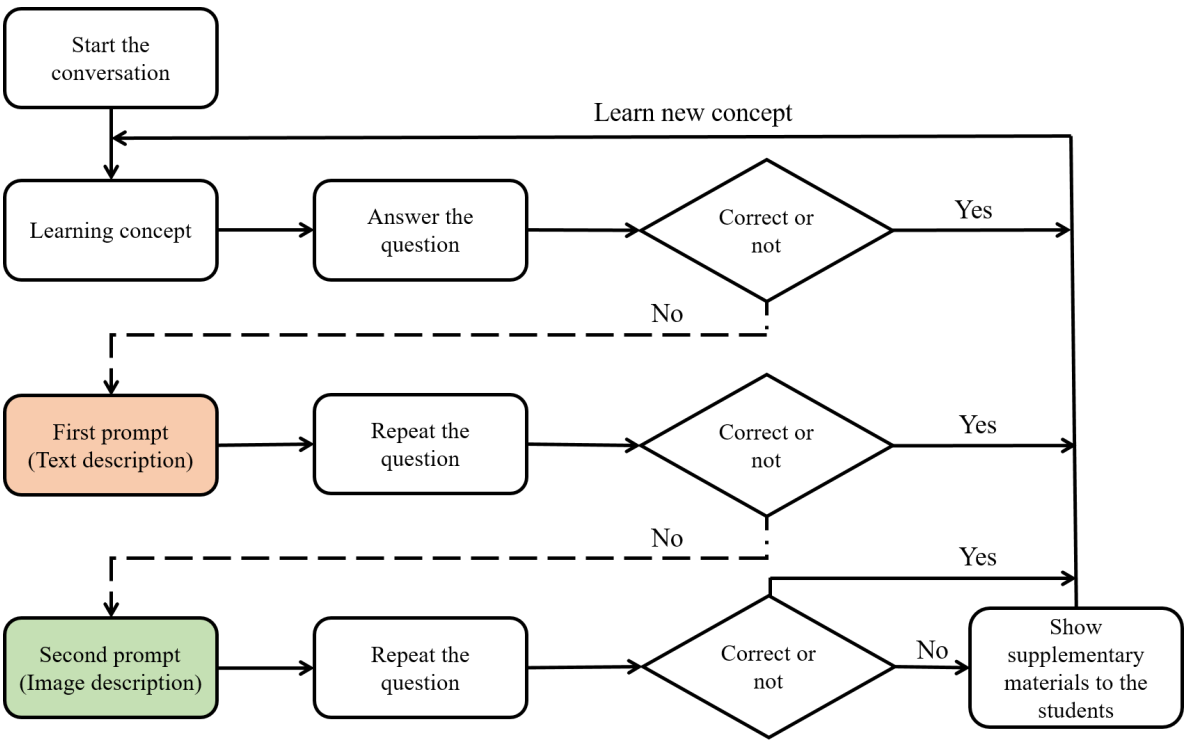


Fig.3 The mechanism of progressive prompting approach

4. 实验设计

4.1. 参与者

本研究采用准实验研究法，参与者为中国某大学设计系学生两个班共 58 名学生。学生的平均年龄为 19 岁。实验组共有 29 名学生，使用基于渐进式提示的聊天机器人进行学习。控制组共有 29 名学生，使用传统聊天机器人进行学习。两组学生均由同一位老师教授并完成相同的先备知识教学活动。

4.2. 实验流程

Fig. 4 为实验流程。两组学生在第一周会完成前测及前文卷，包含艺术鉴赏能力测验、创造性思考问卷。随后在 50 分钟的设计课程中，教师会明确教学目标，展示不同风格的设计字体并进行讲解。第二周学生在课堂中会是用手机与聊天机器人进行对话，并回答聊天机器人提出的问题。第三周学生会完成后测及后问卷，包含艺术鉴赏能力测验、创造性思考问卷以及学习策略满意度问卷。

4.3. 测量工具

本研究采用的测量工具包括艺术鉴赏能力测验试卷、创造性思考问卷以及学习策略满意度问卷。前后测中的艺术鉴赏能力测验试卷由两名经验丰富的老师进行设计，用来测量学生对于字体设计的鉴赏能力。试卷内容包括十道选择题(40%)、五道填空题(20%)以及一道设计题(40%)。满分为 100 分。

创造性思考问卷采用 Lai and Hwang (2014)提出的测量工具，包括 6 个题目，例如："我喜欢问一些别人想不到的问题"。采用李克特 5 点量表(1=极不同意；5=非常同意)。Cronbach's α 值为 0.83，具有良好的信效度。

学习策略满意度问卷采用 Chu et al.(2010)所提出的测量工具，包括 9 个题目，例如："这次学习活动的任务让我更好地理解了如何识别和分类目标学习对象的特征。"。采用李克特 5 点量表(1=极不同意；5=非常同意)。Cronbach's α 值为 0.80，具有良好的信效度。

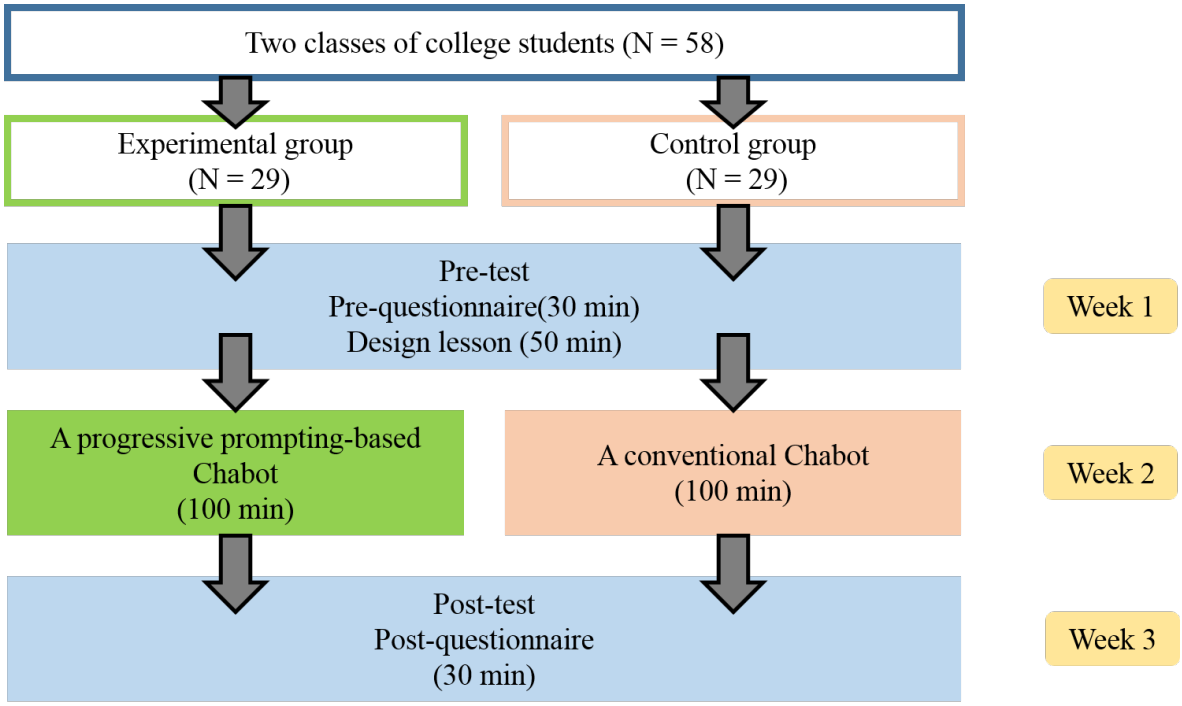


Fig.4 The experimental process

5. 结果

5.1. 艺术鉴赏能力

本研究使用单因子共变异数分析(ANCOVA)来评价实验组及控制组的艺术鉴赏能力。Levene's test 结果不拒绝变异数同构型假设($F(1, 58) = 1.75, p > 0.05$)。两组艺术鉴赏能力的 ANCOVA 结果如 Table 1 所示。在排除前测分数影响后，两组之间 $F(1, 55) = 5.10, p < 0.05$, 存在显著差异。调整后的均值实验组为 83.99，控制组为 81.90，两组调整后的变异数分别为 0.41 和 0.41。实验组分数显著高于控制组分数。这一结果意味着基于渐进式提示的聊天机器人比传统聊天机器人更能提升学生的艺术鉴赏能力。

Table 1. Analysis result of ANCOVA of the art appreciation ability

	N	Mean	SD	Adjusted M	Adjusted SD	*F*	η^2
Experimental group	29	84.76	12.95	83.99	0.41	5.10*	0.09
Control group	29	81.90	6.15	82.67	0.41		

*$p < 0.05$

5.2. 创造性思考

ANCOVA 被用来评量实验组及控制组学生的创造性思考。Levene's test 结果不拒绝变异数同构型假设($F(1, 58) = 0.93, p > 0.05$)。ANCOVA 结果如 Table 2 所示，在排除前测分数影响后，两组之间 $F(1, 55) = 4.22, p < 0.05$, 存在显著差异。调整后的均值实验组为 4.29，控制组为 4.02。调整后的变异数分别为 0.09 和 0.09。这一结果意味着使用基于渐进式提示的聊天机器人的学生比使用传统聊天机器人的学生具有更多的创造性思考。

Table 2. Analysis result of ANCOVA of the creative thinking

	N	Mean	SD	Adjusted M	Adjusted SD	*F*	η^2
Experimental group	29	4.26	0.55	4.29	0.09	4.22*	0.07
Control group	29	4.05	0.47	4.02	0.09		

*$p<0.05$

5.3. 学习策略满意度

为了了解学生对于学习策略的满意度，对实验组和控制组学生前后测结果分别进行独立样本 t 检定，结果如 Table 3 所示。实验组学生的平均分数为 4.14，控制组学生的平均分数为 4.33。两组学生在对学习策略的满意度上面无显著差异($t = 1.52, p > 0.05$)。结果显示学生对基于渐进式提示的聊天机器人和传统聊天机器人的满意度相同。

Table 3. Analysis result of ANCOVA of students' satisfaction with the learning approach

	N	Mean	SD	*t*	*p*
Experimental group	29	4.14	0.53	1.52	0.13
Control group	29	4.33	0.43		

6. 讨论及结论

本研究的目的是在聊天机器人中导入渐进式提示策略，帮助学生了解艺术作品中的设计思路，提升他们的鉴赏能力。本研究通过比较学生的艺术鉴赏能力、创造性思考以及对于学习策略满意度来评估新方法的有效性。结果表明基于渐进式提示策略的聊天机器人对促进学生的艺术鉴赏能力有很大的好处。本研究结果与之前的研究一致，表明渐进式提示能够鼓励

学生寻找线索并进行反思和调整(Chen et al., 2014; Yang et al., 2018)。相较于控制组学生一次性得到全部的提示，实验组学生在回答问题的过程中通过适性化的反馈有更多思考和认知建构的机会(Schrader & Grassinger, 2021)。这可能是实验组学生学习表现显著高于对照组的原因之一。

同时，基于渐进式提示策略的聊天机器人也能够促进学生的创造思考。Chen et al. (2021)指出，具有挑战性的学习场景有利于激发学生的高层次思考。学生可以将第一阶段的文字提示中的抽象概念具象化。而在第二阶段提示中呈现图像提示能够让学生产生比较和反思。适性化的认知建构方式能够帮助学生充分发展观念并形成具有发散思维过程的解决方案(Wen, 2021; Wilson et al., 2012)。

除此之外，研究还显示实验组和对照组学生在对学习策略满意度方面没有显著差异。Chang et al. (2021) 指出相较于传统的讲授式教学，基于知识的聊天机器人可以提供实时反馈，帮助学生有效地学习并提高满意度。本研究中两组都是通过聊天机器人来进行学习，因此这个结果意味着学生愿意参与到学习过程中。渐进式提示策略也不会影学生在聊天机器人环境下的满意度。

综上所述，基于渐进式的聊天机器人能够帮助提高学生的艺术鉴赏能力和促进他们的创造性思考，同时不会影响他们对于学习策略的满意度。未来我们计划根据更多个人因素，例如交流倾向、批判思维倾向等，应用这一机制来制定适性化的提问和提示策略。

7. 参考文献

Chang, C. Y., Kuo, S. Y., & Hwang, G. H. (2022). Chatbot-facilitated Nursing Education: Incorporating a Knowledge-based Chatbot System into a Nursing Training Program. *Educational Technology and Society*, *25*(1).

Chaves, A. P., & Gerosa, M. A. (2021). How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. *International Journal of Human–Computer Interaction*, *37*(8), 729-758.

Chen, C. H., Hwang, G. J., & Tsai, C. H. (2014). A progressive prompting approach to conducting context-aware learning activities for natural science courses. *Interacting with Computers*, *26*(4), 348-359.

Chen, C. Y., Chang, S. C., Hwang, G. J., & Zou, D. (2021). Facilitating EFL learners' active behaviors in speaking: a progressive question prompt-based peer-tutoring approach with VR contexts. *Interactive Learning Environments*, 1-20.

Chen, H. L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research*, *58*(6), 1161-1189.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education, 55*(4), 1618-1627.

Deveci Topal, A., Dilek Eren, C., & Kolburan Geçer, A. (2021). Chatbot application in a 5th grade science course. *Education and Information Technologies*, *26*(5), 6241-6265.

Fidan, M., & Gencel, N. (2022). Supporting the Instructional Videos With Chatbot and Peer Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation. *Journal of Educational Computing Research*, 07356331221077901.

Hung, H. C., & Young, S. S. C. (2017). Applying multi-touch technology to facilitate the learning of art appreciation: from the view of motivation and annotation. *Interactive Learning Environments*, *25*(6), 733-748.

Hussain, S., Ameri Sianaki, O., & Ababneh, N. (2019, March). A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the International Conference on Advanced Information Networking and Applications* (pp. 946-956). Springer, Cham.

Irwin, R. L., LeBlanc, N., Ryu, J. Y., & Belliveau, G. (2018). A/r/tography as Living Inquiry. *Handbook of arts-based research*, 37-53.

Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem-solving, meta-cognitive awareness and creativity. *International Journal of Mobile Learning and Organisation, 8*(3), 276-291.

Laurillard, D. (2013). *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies*. Routledge.

Lin, C. J., & Mubarok, H. (2021). Learning Analytics for Investigating the Mind Map-Guided AI Chatbot Approach in an EFL Flipped Speaking Classroom. *Educational Technology & Society*, *24*(4), 16-35.

Luger, E., & Sellen, A. (2016, May). " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5286-5297).

Palau-Pellicer, P., Mena, J., & Egas, O. (2019). Arts-based educational research in museums:'art for learning art', an a/r/tographic mediation. *International Journal of Art & Design Education*, *38*(3), 670-680.

Schrader, C., & Grassinger, R. (2021). Tell me that I can do it better. The effect of attributional feedback from a learning technology on achievement emotions and performance and the moderating role of individual adaptive reactions to errors. *Computers & Education*, *161*, 104028.

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151*, 103862.

Sun, C. T., Chen, L. X., & Chu, H. M. (2018). Associations among scaffold presentation, reward mechanisms and problem-solving behaviors in game play. *Computers & Education*, *119*, 95-111.

ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., & Elen, J. (2017). Computer game-based mathematics education: Embedded faded worked examples facilitate knowledge acquisition. *Learning and instruction*, *50*, 44-53.

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, *18*(1), 1-20.

Wen, M. (2021). Creative in future music teachers: Applying psychodiagnostic methods to assess the creative potential. *Thinking Skills and Creativity*, *41*, 100914.

Wilson, T., Perry, M., Anderson, C. J., & Grosshandler, D. (2012). Engaging young students in scientific investigations: prompting for meaningful reflection. *Instructional Science*, *40*(1), 19-46.

Yang, K. H., Chu, H. C., & Chiang, L. Y. (2018). Effects of a progressive prompting-based educational game on second graders' mathematics learning performance and behavioral patterns. *Journal of Educational Technology & Society*, *21*(2), 322-334.

Yeh, Y. C., & Peng, Y. Y. (2019). The Influences of aesthetic life experience and expertise on aesthetic judgement and emotion in mundane arts. *International Journal of Art & Design Education*, *38*(2), 492-507.

Yin, J., Goh, T. T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, *59*(1), 154-177.

生成式人工智能支持下的协作编程对学生批判性思维和编程群体自我效能的影响研究

A Study on the Impact of Collaborative Programming Supported by Generative Artificial Intelligence on Students' Critical Thinking and Programming Collective Self-Efficacy

严乙淼[1*]，陈创奇[1]，叶新东[1]
[1]温州大学 教师教育学院
[*] yym_1999@163.com

[摘要] 协作编程作为一种常见的编程教学方式可以极大地促进学生之间的合作与交流，能提高学生的团队协作能力。然而，学习者可能会因为在编程知识的不足影响了协作编程的有效进行。因此，本研究采用 GPT-4 介入常规协作编程教学为学生提供全面的学习内容来弥补学习者在知识方面的不足。本研究旨在评估在协作编程的环境下 GAI 对小学生批判性思维与编程群体自我效能感的有效性。研究结果表示：GAI 支持下的协作编程教学模式可以显著提升小学生的批判性思维与编程群体自我效能感。

[关键词] GAI；协作编程；批判性思维；编程群体自我效能感

Abstract: *Collaborative programming, as a common method in programming education, can significantly foster cooperation and communication among students, enhancing their team collaboration skills. However, learners may face challenges in effective collaborative programming due to deficiencies in programming knowledge. Therefore, this study incorporates GPT-4 into conventional collaborative programming teaching methods to provide comprehensive learning content and address learners' knowledge gaps. The research aims to assess the effectiveness of Generative Artificial Intelligence (GAI) in fostering critical thinking and group self-efficacy in programming among elementary school students within the collaborative programming environment. The study results indicate that the collaborative programming teaching model supported by GAI can significantly enhance elementary school students' critical thinking and group self-efficacy in programming.*

Keywords: GAI; collaborative programming; critical thinking; programming collective self-efficacy

基金项目:温州大学 2024 年硕士研究生创新基金学校资助基金项目“基于 LLM 的协作编程策略对中学生计算思维与自我效能感的应用研究”(课题编号: 3162024003019)。

1.引言

在这个信息充沛的时代，人们随时能获取大量信息(Halpern, 1999)，如何分辨与处理信息成为一种挑战。编程教育作为教授解决问题和批判性思维的绝佳工具，有助于发展学生的思维能力(Tasneem, 2012)，而协作编程(Cooperative Programming)作为编程教育中一种常见的方法，通过组员间共同构建代码的教学形式，为团队成员带来了诸多益处，可以培养学生的批判性思维与自我效能感(Denner et al., 2014; Scherer, 2016)。然而，在以往的协作编程过程中，

学习者往往面临一些挑战，其中一个显著的问题是团队成员在编程知识方面的存在的不足，对协作编程中学生的批判性思维与自我效能感产生负面影响(李新 et al., 2023)。

近年来，由 OpenAI 开发的 GPT-4 等大语言模型，是当下生成式人工智能(Generative AI)的代表性成就之一(Cao et al., 2023)。由 GPT-4 和类似 GAI 带来的教学变革产生的个性化教学的方式，可以在教学过程中给予实时反馈与支持，为学生提供更加全面、灵活的学习内容(Ahmad et al., 2023)，有效弥补了协作编程中成员知识方面存在的不足，解决了传统课堂在实施协作编程面临的许多限制和挑战。

尽管 GAI 在编程教学中有巨大潜力，但有关 GAI 如何融入 K-12 合作学习的实证研究仍相对有限。因此，本研究将 GAI 加入编程课堂中，探讨 GAI 与协作编程的教学环境对学生批判性思维与编程群体自我效能感的影响。研究问题如下：

问题 1：基于 GAI 的协作编程能否有效提升小学生的批判性思维？

问题 2：基于 GAI 的协作编程能否有效提升小学生的编程群体自我效能感？

2.文献综述

2.1. 协作编程中的批判性思维与编程群体自我效能感

在 21 世纪学习技能发展的背景下，协作解决问题的技能是至关重要的。协作编程是在协作学习的基础之上发展而来的，认知主义与建构主义认为，在协作学习的过程中，团队成员之间会积极互动，实现知识的共享共建(Kalaian & Kasim, 2015)从而培养学习者的批判性思维和解决问题能力(Susantini et al., 2018; Tang et al., 2020)。除此之外，协作编程的学习环境也为团队成员带来了诸多益处：节省了重复工作的时间，提高了代码质量和学习者的自我效能感(Williams & Kessler, 2003)。虽然课堂上的合作学习法已经被证明对编程学习者非常有益，但在今天的学习环境中固有的障碍往往会阻止合作学习方法的实施(Boyer et al., 2008)。除此之外，一项研究发现，在协作编程的过程中，只会提升编程能力较弱的成员的能力(Balijepally et al., 2009)，忽视了所有学生的共同发展。

2.2 生成式人工智能的发展及其在编程教育中的应用

近年来，先进人工智能技术的兴起对包括教育和研究在内的许多领域产生了深远的影响(Rahman & Watanobe, 2023)。ChatGPT、BERT、GPT-3、GPT-4 等类似 GAI 都是通过大量公开的文本数据进行训练，以在对话中生成类似人类的响应(Manning, 2022; Meyer et al., 2023)。GPT-4 作为一种 GAI，可以为学生提供个性化和适应性的学习体验，它可以分析学生对问题的回答，并使材料适应他们的个人学习模式(Ahmad et al., 2023)。有研究指出，在编程教育中使用 GAI 可以提升学习者的思维能力，增强其自我效能(Yilmaz & Yilmaz, 2023)。(MacNeil et al., 2022)等人最近的一项研究就是使用 GAI 生成代码解释，这项研究已经成功地证明了 GAI 通过解释给定代码片段的各个方面来支持学习的潜力。

在教育中使用 GAI 已经被认为是一个潜在的兴趣领域，然而在以往的编程教育中使用 GAI 的研究主要关注于高等教育。因此，本研究旨在研究在小学编程教育课堂中，借助 GAI 辅助的协作编程学习方法对学习者批判性思维能力发展的影响。对 GAI 在 K-12 阶段编程领域的影响进行了进一步的实证探讨。

3.研究设计

3.1 实验对象

本研究从浙江某所小学随机招募了 64 名小学六年级学生作为实验对象，年龄的平均数为 11.37，标准差为 0.49，共 30 位男生，34 位女生，实验参与者被随机分为两组。研究人员联系了孩子的父母或监护人，介绍并解释了研究方案，同意孩子参加的父母或监护人签署了知情同意书。研究结束后，对照组学生参加了三个精简的使用 GPT-4 的协作编程课程，重点关注 GPT-4 功能，确保平等获得干预的潜在好处。

3.2 课程设计

本研究的实验组将 GPT-4 融入常规协作编程课堂之中，采用 SpeechGPT 平台，接入 GPT-4 接口，平台界面如图 1 所示，Wang and Hwang (2017)设计的基于问题的练习(problem posing-based practicing，PPBP）策略为协作编程教学提供了一个完善的框架，包括授课、思考、练习、讨论、反思，课程具体实施步骤如下：

（1）教师授课阶段：教师通过集体授课的方式，讲授本节课的教学内容，要求学生全面理解并掌握，并对知识点进行提问，之后布置小组合作任务。在该阶段，GAI 可以充当辅助教师教学的角色。

（2）独立思考阶段：学生内化教师讲授的知识点，对问题进行深度思考，罗列解决问题的过程，提出对问题的初步看法，为小组合作打下基础。

（3）合作练习阶段：学生的主体地位得到进一步增大，根据独立思考阶段得出的结论，小组成员之间相互讨论，如图 2 所示。在该阶段，学生可以使用 SpeechGPT 平台，在用户提问区域请求 GAI 给予帮助。GAI 可以给学生提供实时的反馈，与小组成员交流以促进任务完成。每个小组成员需遵守以下规定：①所有小组成员掌握知识点之后才能进行小组合作编程阶段；②在请教老师或是 GAI 之前要先请求小组所有成员的帮助。



图 1 SpecchGPT 平台



图 2 合作讨论阶段

（4）初步结论与讨论反思阶段：通过小组之间相互讨论，指导者的帮助得出有关小组协作任务的初步答案。教师参考 GAI 的呈现内容对小组任务进行指导反馈，学习者根据教师的反馈与 GAI 的支持得出小组协作任务的最终答案。在该阶段，GAI 可以辅助教师审核小组任务的完成情况，辅助学生对答案进行进一步确认。

各阶段结束后，教师将最终小组合作成效进行记录，结合 GAI 对本次小组合作教学进行反思，为下次开展小组合作项目提出建议。

3.3 实验流程

实验在机房中进行，实验组与对照组都是四人一协作小组，每人都配备电脑，实验组四人一起使用一个 GPT-4。在实验准备与前测阶段，对照组与实验组的学生都会进行分组，对照组与实验组的教师都会对学生强调协作编程的注意事项，要求学生在课堂中通过分工的方式完成编程任务。实验组的教师会教授学生如何使用 GPT-4，知道如何操作系统之后实验组与对照组学生会签署知情同意书并完成批判性思维与编程群体自我效能感量表，实验流程如图 3 所示。在实验中，对照组学生在课堂中实施传统协作编程模式，通过分工合作、实时交流来高效地完成编程任务。实验组学生在 GPT-4 辅助下进行协作编程，每节编程课堂时间为 80 分钟，会有 30 分钟的协作编程讨论时间，实验持续四周。四周之后被试完成批判性思维与编程群体自我效能感量表，并在实验组随机抽取 5 名学生进行访谈。

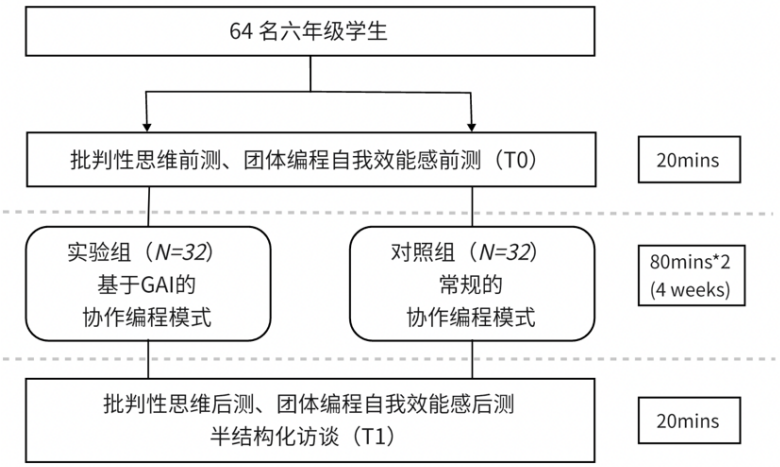


图 3 实验流程图

3.4 测量工具

本研究用到的测量工具包括批判性思维倾向量表、编程群体自我效能感量表和访谈提纲。

3.4.1 批判性思维倾向量表

关于批判性思维的倾向，使用的是 Chai 等人(Chai et al., 2015)的量表，由六个问题组成，（例如，“在学习的过程中，我会思考自己所学的是否正确。”），使用 5 点李克特量表，从 1= 完全不符合到 5= 完全符合。问卷具有较高的信度（Cronbachα=0.88～0.79），问卷主要用于评价学习者在课程中培养的批判性思维倾向。

3.4.2 编程群体自我效能量表

编程群体自我效能感前后测试题均改编自 Wang and Hwang (2012)，由八个项目组成，例如“我相信我的团队能理解这堂编程课上最难的部分”。本问卷采用 5 点李克特量表，从 1= 完全不符合到 5= 完全符合。原研究中该问卷的 Cronbach's alpha 为 0.92。

3.4.3 访谈提纲

访谈提纲改编自 Hwang 等人(Hwang et al., 2009)的访谈内容进行修改，包括五个半结构化问题（例如，“你喜欢使用 GPT-4 的哪些方面？有哪些方面你觉得困难？”），在实验组中随机选择 5 名受访者。访谈的目的是从学生的角度来了解此教学方法的优缺点，并为将来的研究

改进提供参考依据。

4.研究结果

在进行实验检测之前，对实验组与对照组的批判性思维与编程群体自我效能感前后测数据进行正态性检验与方差齐性检验，检验结果可知改组数据均符合正态且呈现方差齐性。进行 ANOVA 检验的结果发现不同组别的样本的批判性思维与编程群体自我效能感的前测分数均无显著差异（p>0.05）。

表 1 批判性思维、编程群体自我效能感前测 ANOVA 检验分析结果

Dimension	Group	N	Mean	SD	T	*p*
批判性思维	实验组	32	22.75	3.37	2.950	0.091
	对照组	32	21.19	3.89		
编程群体自我效能感	实验组	32	31.50	5.12	1.553	0.217
	对照组	32	29.56	7.15		

4.1 批判性思维

为检验在编程课堂中使用协作编程和使用融入 GPT-4 的协作编程对学生批判性思维倾向的影响，对实验组与对照组学生的前测和后测分数进行配对 T 检验（见表 2），结果发现实验组批判性思维前测和后测之间呈现出 0.001 水平的显著性(t=-4.383，p=0.000)，对照组批判性思维前测和后测之间呈现出 0.05 水平的显著性(t=-2.203，p=0.035)。经过四周的实验干预，研究结果显示，实验组和对照组学生在批判性思维方面都取得了显著的进步，其中实验组的效应量为中效果(d=0.577)，对照组的效应量为小效果(d=0.405)。

表 2 批判性思维配对 t 检验分析结果

Dimension	Group	Mean	SD	*t*	*p*
批判性思维	实验组前测	22.75	3.37	-4.383	0.000***
	实验组后测	24.69	3.35		
批判性思维	对照组前测	21.19	3.89	-2.203	0.035*
	对照组后测	22.69	3.50		

*p<0.05 ***p<0.001

为对比在协作编程课堂中使用 GPT-4 对学生批判性思维倾向的影响，对实验组与对照组的批判性思维后测进行 ANOVA 检验（见表 3），结果发现组间批判性思维后测呈现出 0.05 水平的显著性(F=5.452，p=0.023)，以及具体对比差异可知，实验组批判性思维后测的平均值(24.69)明显高于对照组后测平均值(22.69)。结果表明，基于 GPT-4 的协作编程课堂相较于常规的协作编程课堂能够显著提升学生的批判性思维倾向。

表 3 批判性思维实验组与对照组后测 ANOVA 检验分析结果

Dimension	Group	Mean	SD	F	*p*
批判性思维	实验组后测	24.69	3.35	5.452	0.023*
	对照组后测	22.69	3.55		

*p<0.05

4.2 编程群体自我效能感

为检验在编程课堂中使用协作编程和使用融入 GPT-4 的协作编程对学生编程群体自我效能感的影响，对实验组与对照组学生的前测和后测分数进行配对 T 检验（见表 4），结果发现实验组编程群体自我效能感前测和后测之间呈现出 0.005 水平的显著性(t=-3.602，p=0.001)，对照组编程群体自我效能感前测和后测之间呈现出 0.05 水平的显著性(t=-2.072，p=0.047)。经过四周的实验干预，研究结果显示，实验组和对照组学生在编程群体自我效能感方面都取得了显著的进步，其中实验组的效应量为中效果(d=0.587)，对照组的效应量为小效果(d=0.309)。

表 4 编程群体自我效能感能配对 t 检验分析结果

Dimension	Group	Mean	SD	*t*	*p*
编程群体自我效能感	实验组前测	31.50	5.12	-3.602	0.001**
	实验组后测	34.28	4.31		
编程群体自我效能感	对照组前测	29.69	7.32	-2.072	0.047*
	对照组后测	31.69	5.50		

*p<0.05 **p<0.005

为对比在协作编程课堂中使用 GPT-4 对学生编程群体自我效能感的影响，对实验组与对照组的编程群体自我效能后测进行 ANOVA 检验（见表 5），结果发现组间编程群体自我效能后测呈现出 0.05 水平的显著性(F=4.416，p=0.040)，以及具体对比差异可知，实验组编程群体自我效能后测的平均值(34.28)明显高于对照组后测平均值(31.69)。结果表明，基于 GPT-4 的协作编程课堂相较于常规的协作编程课堂能够显著提升学生的编程群体自我效能感。

表 5 编程群体自我效能感实验组与对照组后测 ANOVA 检验分析结果

Dimension	Group	Mean	SD	F	*p*
编程群体自我效能感	实验组后测	34.28	4.31	4.416	0.040*
	对照组后测	31.69	5.49		

*p<0.05

4.3 半结构化访谈

访谈题目一共有五道，可分为两类：1、在协作编程中使用 GPT-4 的优点，2、在协作编程中使用 GPT-4 的缺点。分析学生对使用 GPT-4 在协作编程学习过程的优点的看法，学生意见见表 6。

表 6 关于使用 GPT-4 的优点与好处

Sub-Themes	f
回答内容丰富易懂	2
回答迅速	4
帮助修改错误	1、3、5
帮助理解问题	4
提升对编程的兴趣	2、3、4、5
提升编程自信心	1、2、3、5

分析学生对使用 GPT-4 在协作编程学习过程的缺点与不足，学生意见见表 7。

表 7 关于使用 GPT-4 的缺点与不足

Sub-Themes	f
有时候会答非所问	3
有些内容看不懂	4
理解不了提问者的意思	5

根据表格所示的数据，清晰地展示了学生对使用 GPT-4 的积极态度，使用 GPT-4 的优点与好处远大于缺点与不足。这种明显的正面反馈表明 GPT-4 在协作编程教学环境中取得了显著的成功。除此之外，大多数学生都认为使用 GPT-4 可以提升编程的自信心，从另一方面证明融入 GPT-4 的协作编程可以显著提升学生的编程群体自我效能感。

5.讨论

本研究探讨了基于 GAI 的协作编程模式与常规的协作编程模式对小学生批判性思维与编程群体自我效能感的影响。研究结果证实了在进行四周干预实验之后，实验组与对照组学生的批判性思维与编程群体自我效能感均有了显著性提升，而实验组学生的提升显著高于对照组学生。

实验组学生批判性思维的显著提升可能归因于 GAI 的支持。在协作编程的过程中 GAI 可以为学习者提供更丰富、准确的语境和对困难代码的解释，促使学生在解决问题时更深入地思考(Yilmaz & Yilmaz, 2023)。通过这种方式，GAI 充当了一位积极的学习助手，为学生创造了更有利于深度思考和学科理解的学习环境。同时，团队编程的协作性质也为学生提供了相互学习和讨论的机会，激发了他们的批判性思维(Susantini et al., 2018; Tang et al., 2020)。这可以在一定程度上解决常规协作编程只提升编程能力较弱的成员的能力的情况(Balijepally et al., 2009)，重视的是所有学生的共同发展。

此外，先前的研究表明，学生对自己在编程任务中的能力和角色的认知与他们的学习体验和动机密切相关(Zimmerman, 2000)。通过基于 GAI 的协作编程课程，学生在团队合作中获得了更积极的反馈，感受到其在团体编程中的贡献，这种正面经验在一定程度上提升了学生的编程群体自我效能感。除此之外，从半结构化访谈的结果中可以看出，学生对 GPT 持有积极的态度，他们将其视为一个有益的工具，用以解决在编程学习中遇到的各种困难，以期更好地完成教师布置的协作讨论问题。这表明 GAI 的应用在提升学生学习体验和克服学习障碍方面具有显著潜力，为协作编程的未来发展提供了有力的支持。

6.结语

本研究发现，基于 GAI 的协作编程模式可以培养小学生批判性思维与编程群体自我效能感。在协作编程的学习环境中，学生通过向 GAI 提问，不仅解决了技术性难题，更激发了他们深入思考和批判性分析问题的能力，这种学术上的成长不仅为编程领域的学习提供了更全面的支持，也为他们日后面对复杂问题时培养了坚实的思考基础。尽管初步的研究结果证明了 GAI 在协作编程中的潜力，但研究仍处于初级阶段，还有许多不足之处，需要持续深入的实证研究作为更加有力的支撑。

参考文献

李新, 李艳燕, & 杨现民. (2023). 协作学习投入的结构模型构建与机理阐释. *现代教育技术*, *33*(12), 45-55.

Ahmad, N., Murugesan, S., & Kshetri, N. (2023). Generative Artificial Intelligence and the Education Sector. *Computer*, *56*(6), 72-76.

Balijepally, V., Mahapatra, R., Nerur, S., & Price, K. H. (2009). Are two heads better than one for software development? The productivity paradox of pair programming. *MIS quarterly*, 91-118.

Boyer, K. E., Dwight, A. A., Fondren, R. T., Vouk, M. A., & Lester, J. C. (2008). A development environment for distributed synchronous collaborative programming. Proceedings of the 13th annual conference on Innovation and technology in computer science education,

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *arXiv preprint arXiv:2303.04226*.

Chai, C. S., Deng, F., Tsai, P.-S., Koh, J. H. L., & Tsai, C.-C. (2015). Assessing multidimensional students' perceptions of twenty-first-century learning practices. *Asia Pacific Education Review*, *16*, 389-398.

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Pair programming: Under what conditions is it advantageous for middle school students? *Journal of Research on Technology in Education*, *46*(3), 277-296.

Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning*, *1999*(80), 69-74.

Hwang, G.-J., Yang, T.-C., Tsai, C.-C., & Yang, S. J. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & education*, *53*(2), 402-413.

Kalaian, S. A., & Kasim, R. M. (2015). Small-group vs. competitive learning in computer science classrooms: A meta-analytic review. In *Innovative teaching strategies and new learning paradigms in computer programming* (pp. 46-64). IGI Global.

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022). Generating diverse code explanations using the gpt-3 large language model. Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 2,

Manning, C. D. (2022). Human language understanding & reasoning. *Daedalus*, *151*(2), 127-138.

Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., & Gonzalez-Hernandez, G. (2023). ChatGPT and large language models in academia: opportunities and challenges. *BioData Mining*, *16*(1), 20.

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783.

Scherer, R. (2016). Learning from the past–the need for empirical evidence on the transfer effects of computer programming skills. *Frontiers in Psychology*, *7*, 1390.

Susantini, E., Prahani, B., & Mahdiannur, M. (2018). Improving collaborative critical thinking skills of physics education students through implementation of CinQASE learning model. Journal of Physics: Conference Series,

Tang, T., Vezzani, V., & Eriksson, V. (2020). Developing critical thinking, collective creativity skills and problem solving through playful design jams. *Thinking Skills and Creativity*, *37*, 100696.

Tasneem, S. (2012). Critical thinking in an introductory programming course. *Journal of computing sciences in colleges*, *27*(6), 81-83.

Wang, S.-L., & Hwang, G.-J. (2012). The role of collective efficacy, cognitive quality, and task cohesion in computer-supported collaborative learning (CSCL). *Computers & education*, *58*(2), 679-687.

Wang, X.-M., & Hwang, G.-J. (2017). A problem posing-based practicing strategy for facilitating students' computer programming skills in the team-based learning mode. *Educational Technology Research and Development*, *65*, 1655-1671.

Williams, L., & Kessler, R. R. (2003). *Pair programming illuminated*. Addison-Wesley Professional.

Yilmaz, R., & Yilmaz, F. G. K. (2023). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100005.

Zimmerman, B. J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, *25*(1), 82-91.

运用“HPS”+“PhET”实施科学教学的创新设计探索——以“地球圈层结构”教学活动设计为例

Innovative Design Exploration of Science Teaching Using "HPS" + "PhET" - A Case Study of "Earth's Layered Structure" Teaching Activity Design

李冬，李悦陁
上海市曹杨第二中学附属学校
上海师范大学生命科学学院
leedong110@163.com

【摘要】 以科学学科《地球圈层结构》一课的教学活动设计为例，应用科学史资源开展 HPS 模式的教学，引导学生基于科技史实的追本溯源探索，展现概念形成的过程；应用“PhET 虚拟实验技术“突破难点，实施了立足课堂的科学深度探究活动，给学生提供丰富的学习支架，帮助学生形成立体的知识结构；应用多种学习资源探索，感悟科学本质，树立学生良好的责任意识，从而促进科学学科核心素养的落实。
【关键词】 HPS；PhET；科学教学；创新设计

Abstract: *Taking the teaching activity design of the science subject "Earth's Layered Structure" as an example, this study applies the HPS (History and Philosophy of Science) model to explore the origins based on historical scientific resources, guiding students to understand the process of concept formation. Additionally, the "PhET Virtual Experiment Technology" is utilized to overcome difficulties and implement classroom-based scientific inquiry activities, providing students with a rich learning framework and helping them form a comprehensive knowledge structure. Through the exploration of various learning resources, students gain insights into the essence of science and develop a strong sense of responsibility, thereby promoting the implementation of core scientific literacy.*
Keywords: HPS,PhET,Science Teaching,Innovative Design

1. 前言

在实际科学教学中，教材中广泛分布、延伸复杂的知识概念可能加重学生学习负担，妨碍科学思维和核心素养的培养。为解决这一问题，“《义务教育科学课程标准 2022 版》建议在核心概念教学中融入科技史、模拟实验、虚拟技术等手段，引导学生理解科学本质”。（沈强，2023）在七年级科学课中，教师以“地球圈层结构”为例，采用 HPS 模式教学，利用科学史资源帮助学生重温科学家探索历程，展示概念形成过程；应用 PhET 虚拟实验技术突破难点，展开深度探究活动，建立立体知识结构；结合多种学习资源，培养学生责任意识，促进核心素养的实践。这种教学方法旨在帮助学生理解科学概念、培养科学思维，促进科学素养的全面发展。

2. 概念解析

HPS（History, Philosophy, & Sociology of Science）教育是近年来西方一些科学教育专家以建构主义为指导思想倡导的一种新的科学教育课程与教学模式。通过 HPS 模式学习课程（图 1），学生能够主动参与学习，在提出问题、获取信息、寻找证据、检验假设、发现规律等过程中多维度深入理解科学内涵，构建相应的科学观念，增强对科学学习的兴趣，培养终身学习和创新实践能力。



图 1 HPS 教学模式流程图

PhET 是（Physics Education Technology）的简称，即物理教育技术。该项目于 2002 年由诺贝尔物理学奖获得者卡尔·威曼发起，由科罗拉多大学进行制作和运营，其目标是通过自由互动的仿真实验程序，提高全球范围内人们的科学素养。

3. 基于教学内容的前期分析

3.1. 教学内容分析

本节课是牛津上海版《科学》七年级第二学期第十三单元《地球、矿物与材料》第一节“地球”中的第一课时（图 2）。教材结构明确，主要讲述地球的外形特点、内部圈层和外部圈层。通过“寻找地球是个球体的证据”等活动，培养学生科学思维，正确认识地球的结构。教师应结合“HPS 模式”和虚拟实验技术，选择科学史阅读材料和虚拟实验，帮助学生开展实践探究活动，创造适合的情境促进深度学习，培养学生科学思维，有效达成教学目标。

3.2. 教学目标分析

✧ 通过阅读人类认识地球形状的过程资料及帆船模拟实验，能够举例说出地球是球体的证据，体会科学家的勇于探索的科学精神，养成实事求是的严谨的科学态度。

✧ 通过测量、比较、归纳的方法，知道地球的内部圈层结构及各层主要特点，了解地球内部圈层不同圈层结构状态与其物质组成、温度、压力有关。

✧ 通过阅读、交流、观看视频等，了解地球的外部圈层结构及其空间分布和相互间的联系，意识到地球各圈层结构对人类生活的意义，关注人与自然的和谐相处。

3.3. 学生学情分析

七年级学生在小学《自然》和六年级《地理》学习中已具备一定知识储备和科学探究能力，对地球形状有基本了解，但在证据运用方面可能缺乏严谨的逻辑推理。本课涉及“帆船远去”实验，学生可能倾向于片面推断地球形状为球体。此外，地球圈层结构知识是六年级内容，时间间隔较长，可能存在遗忘。教学挑战在于帮助学生巩固地球形状知识的同时，简单认识地球圈层结构，尤其在科技手段无法直接观测地球内部的情况下，设计有效探究过程引导学生深入理解地球结构形成、物质组成、温度和压力之间的关系，以实现高效且有深度的学习。

4. 教学设计解析

在教学设计中，教师首先应用“HPS 模式”，引导学生通过提出观点、学习历史、模拟实验等方式了解人类认识地球形状的四个历程，让他们在探究过程中体会科学家勇于探索的科学精神和严谨的科学探究过程。随后，教师组织学生利用 PhET 虚拟仿真软件《板块构造论》进行虚拟实验，学生通过刻度尺、温度计等测量工具和学习单的任务设计自主探究大陆、海洋的地壳、地幔和地核的不同特点，然后收集证据、归纳分析，从而认识地球的内部结构。最后，教师让学生通过希沃互动与视频观看认识地球的外部圈层，并引导学生反思地球各圈层结构对人类生活的意义，关注地球生态环境，培养学生将知识与生活现象联系起来的思维习惯。(图 2)。

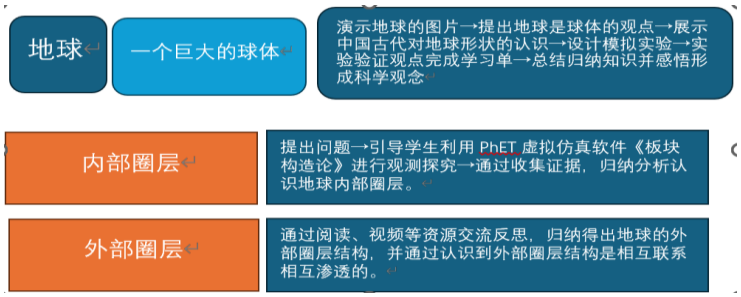


图 2 教学设计流程

4.1. 应用 HPS 模式探究地球形状，促进学生科学思维培养

在地球形状的教学中，教师通过科学史作为串联主线，对模拟实验的演示与日常现象进行对比类推，使学生理解各历史阶段人们对地球形状认知的依据根源，积极主动地逐步建构地球模型。通过科学史和现象判断地球的形状以及不同阶段下人类的认识历程，重现人类对地球形状从“平—不平—球体……”的多次认知飞跃（图 3）。这体现了模型在实验和证据的基础上不断修正与完善，学生对科学本质的理解也在认知过程中不断加深和渐进。

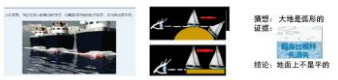
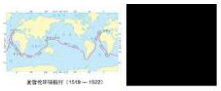
认识演变	时间	采用方式	材料呈现
平	公元前 4 世纪前	科学史主线：天圆地方说	
平-不平	公元前 4 世纪	模拟实验：模型活动比对现实现象	
不平—球体	公元 16-17 世纪	科学史主线：辨析麦哲伦的环球旅行过程	
巨大的球体	20 世纪-21 世纪初期	模型验证：观测技术的发展提升	

图 3“地球形状”认知学习结构表

在具体的教学设计中，教师首先提出问题，引导学生以实证的逻辑思维来逐步建构与完善模型，通过观察模拟实验的过程进而对比现实生活中的现象，在领略科学史历程和模拟实验在质疑和分析的基础上，探索一系列典型的历史事件，帮助学生不断经历思维碰撞，逐步建构出地球的整体模型，最后教师通过总结评价使学生感受到科学是动态发展的，激发学生不断探寻未知领域的热情，进而增强学生的科学责任感，使其更加热爱科学（图 4）。

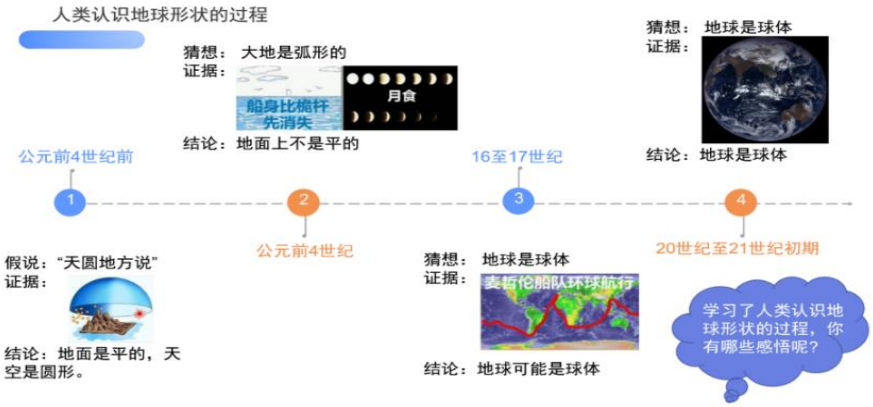
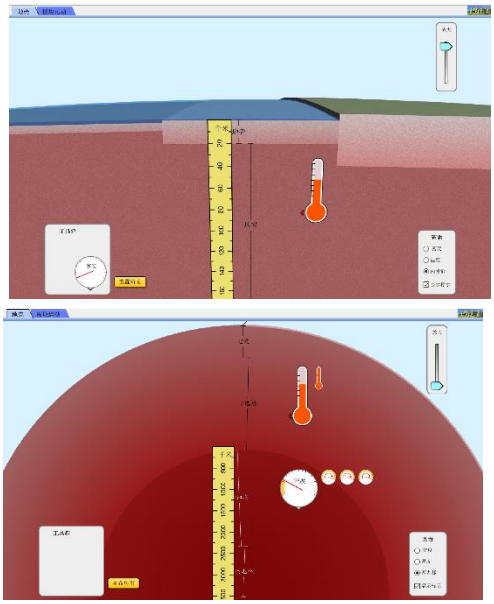


图 4 地球形状教学设计流程图

4.2. 应用 PhET 探究地球内部圈层，提升学生探究实践能力

针对地球的内部圈层结构的教学设计中，教师应用 PhET 虚拟实验——板块构造论项目，帮助学生应用软件，在虚拟场景中运用测量、分析、比较、归纳的科学学习方法来正确认识地球的圈层结构。具体设计如下

- ✧ 提问：地球的内部结构是什么样的呢?
- ✧ 出示 PhET 虚拟实验软件《板块构造论》，介绍软件使用方法。
- ✧ 组织学生利用 PhET 虚拟仿真软件《板块构造论》进行测量实验，并根据测量情况完成学习表格的填写。（图 5）
- ✧ 分析表格内容，组织学生对实验结果进行讨论分析，认知地球内部的圈层结构及特点。



活动二：认识地球内部圈层结构

1. 打开虚拟实验《板块构造论》，在最上面“地壳”标签下完成实验。在“查看”选项下，勾选“两者”和“显示标签”。使用左下角的工具，测量不同圈层结构的厚度和温度，并完成下面的表格。（对于温度，不需要使用单位，只需要记录刻度）

地壳类型	厚度（km）		温度		物质状态
地壳	海洋		海洋		固体岩石
	陆地		陆地		
地幔			100km		
			上部		熔融状物质
			中部		
			下部		
地核			上部		液体铁镍物质
			中部		固体铁镍物质
			下部		

提示：岩石的熔点为 3 刻度。

根据上表数据，请你思考不同圈层的物质状态为何不同?

图 5 地球内部圈层学习单

在明确探究目的的前提下，学生借助虚拟实验资源开展地球内部圈层结构探究实验。通过信息技术模拟真实地理环境，教师设计问题链和学习支架，引导学生在虚拟情境中操作、合作、思考，运用知识和方法获取、处理信息，形成科学探究能力。

4.3. 多角度探索外部圈层，促进学生社会责任的形成

教师在教学地球外部圈层结构时，利用多角度的教学资源如阅读材料、视频等，通过生动有趣的教学资料、游戏互动帮助学生理解地球外部圈层位置；引导学生归纳总结外部圈层结构；促进学生通过阅读交流理解外部圈层相互联系，丰富知识内容，提升认知水平。

5、结语

自 21 世纪以来，科学教育的目标已转向形成核心概念的进阶过程，而非简单获取事实和理论知识。通过观察分析课堂情况，可以发现创新地采用了“HPS”模型和“PhET”虚拟实验技术，这些尝试在地球圈层结构教学中取得了显著成效，学生们展现出高度的参与度和浓厚的兴趣，知识掌握情况也明显提升。由此也进一步阐释了科学课堂教学应满足时代发展需求，融入新教育理念，为学生树立正确价值观和提升综合素质奠定基础。当然也应该清醒地认识到，目前所使用的理论和教学资源仍处于试验阶段，还存在一些不足之处。同时，一节课的设计与实践并不能完全证明其普适性。因此，作者还将继续通过实践进一步推进教科研的实施，不断优化和完善教学方法，以更好地服务于学生的全面发展。

参考文献：

沈强（2023）。“科学史+模拟实验”促进概念理解的教学设计——以浙教版“地球和宇宙”为例[J]。**地理教学，2023(06)**，36-39+47。

人工智能课程中协作脚本的设计与实践研究

Design and Implementation Research of Collaborative Scripts in Artificial Intelligence Courses

列映湘，顾荣桢，黄桂闽*，李绮华，钟柏昌
华南师范大学教育信息技术学院
* 2023020788@m.scnu.edu.cn

【摘要】 人工智能课程的推进于我国推进人工智能教育战略具有重要意义。然而，当前对于人工智能课程内容的具体开展方式仍未有成熟明确的指导方案。人工智能课程因其课程内容具有跨学科、内容新等特点，应以学生小组合作为主要开展形式。协作脚本是协作学习中一种具有代表性的学习形式，其以脚本指导理论、调节学习相关理论、教育论证相关理论为基础，目前已有一些基于协作脚本设计的模型指导教学活动。为促进人工智能课程的开展，基于计算机协作学习环境的协作脚本模型设计有待探索，参考脚本指导理论第六原则（外部协作脚本引导原则）以及第七原则（外部脚本最优级原则），结合本科生课程《STEM 与创客教育》以“人工智能”为主题的项目式学习活动，设计以线下活动为开展形式、具有人工智能课程特色的协作脚本模型，能够为我国人工智能课程的建设与发展提供参考价值。

【关键词】 人工智能课程；人工智能教育；协作脚本

Abstract: *The promotion of AI curriculum is crucial for advancing our country's AI education strategy. However, there's currently a lack of mature guidance for developing specific course content. Due to AI's interdisciplinary nature and novelty, it should primarily involve student group cooperation. Collaborative scripting, rooted in script-guided, moderating learning, and educational argumentation theories, is a key learning format. Several models based on collaborative scripting are available to guide teaching activities. To advance AI courses, we need to explore collaborative script models in computer-based learning environments. By incorporating principles like external script guidance and optimal script levels, along with project-based learning in undergraduate STEM and Maker Education courses, designing a collaborative script model for AI curriculum in offline activities can offer valuable insights for China's AI curriculum construction and development.*

Keywords: artificial intelligence courses, AI education, collaborative scripts

1.引言

新课标改革背景下，我国部分地区试行单独开设中小学人工智能课程，其余大部分地区还是在中小学信息科技课程中加入人工智能相关内容。目前，大部分教材中关于人工智能的内容稀少，且仅限于基础性介绍，缺少对关键技术的深入探讨。这导致人工智能教学多停留在理论层面，缺乏实践活动，不能达到课程的学习目标。同时，教师普遍缺乏人工智能教学经验，无法深入探究问题。此外，软硬件等教学资源匮乏，许多学校缺少必要的计算机设备和软件资源，这都限制了人工智能课程的开展（银涛，2019）。传统的教学方法已难以满足学生对实践操作的需求。协作脚本作为一种新兴的教学工具，已经在许多领域得到了广泛的应用，在人工智能课程中，基于协作脚本的教学模式能够帮助学生获得更深层次的学习体验。

本文旨在探讨人工智能课程中协作脚本的设计与实践，通过对协作脚本和实践案例进行设计，从而为人工智能课程教学实践提供有益的启示与建议。

2. 协作脚本模型研究

协作脚本主要来源于脚本指导理论、调节学习相关理论以及教育论证相关理论。其中，脚本指导理论将协作脚本分为内部和外部两类，在组成结构上均分为剧本、场景、微脚本、角色四个部分（Fischer, Kollar, & Stegmann et al, 2013）。调节学习相关理论和教育论证相关理论由非教育领域迁移而来，有助于小组成员在交互中采取策略从而控制其行为、认知和信念，这两类理论可以支持我们对特定类型的协作脚本组件进行设计。

王小根等人（2019）对半结构化微协作脚本设计进行了探究，在小组讨论环节充分考虑到了组内和组间两个维度；王智颖（2021）为大学生的在线学习设计了角色轮换脚本，保证每个学生在学习过程中能体验到所有角色，但仍无法解决总结者与其他角色连接较弱的问题。王临登（2022）将协作脚本中学习者的角色定义进行了拓展，但仍有继续丰富的空间。

3. 人工智能课程协作脚本

本研究参考了脚本指导理论七大原则中原则 6（外部协作脚本引导原则）和原则 7（外部脚本最优级原则）、外部协作脚本的四种脚本类型以及学习活动支架具有的动态评估和渐隐等特点，针对人工智能课程设计了协作脚本，如图 1 所示。

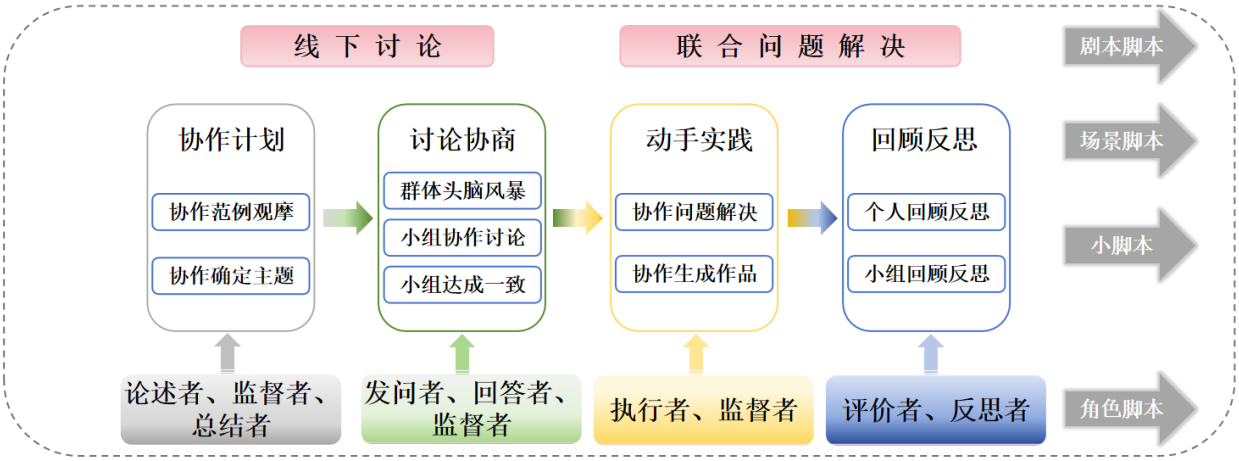


图 1 人工智能课程协作脚本设计

3.1. 剧本脚本设计

鉴于人工智能课程的形式和特点，我们将人工智能课程协作脚本的剧本脚本设置为线下讨论以及小组联合解决问题的形式，这会赋予各小组成员责任感，从而增加其学习投入度。

3.2. 场景脚本设计

结合人工智能课程情景体验式学习的特点，本研究设计了协作计划、讨论协商、动手实践、回顾反思这四个阶段的场景脚本。协作计划阶段，学生在组内讨论并确定作品主题；讨论协商阶段，学生轮流承担不同的角色，在讨论完成后总结观点；动手实践阶段，学生动手制作人工智能创意实践作品；回顾反思阶段，学生对整个活动过程进行总结反思。

3.3. 角色脚本设计

本研究共设计论述者等八个角色。通过独立思考提出观点，此为论述者；对讨论主题及节奏进行把控，此为监督者；对观念进行总结归纳，此为总结者；通过发问和回答的方式迸发思想火花，此为发问者和回答者；动手完成作品制作，此为执行者；对活动进行反思和总结，此为评价者和反思者。其中，每个学生在各阶段所承担的角色进行轮换，避免角色重复。

3.4. 小脚本设计

小脚本作为最底层的协作脚本，其组成内容必须细致且具体，它需要能够直接明确地告知学习者协作活动的操作内容。协作计划阶段，学习者自行组成小组，讨论并确定协作作品的主题；讨论协商阶段，学习者在组内进行讨论并总结；动手实践阶段，学生协作完成创意作品；回顾反思阶段，学生个人先进行反思总结，再共同讨论，最后反思总结小组协作过程。

4.案例设计

本研究选取广州市地方教材《人工智能》五年级上册第六课来阐释本研究脚本实现过程，该年级学生对人工智能概念有基础认识，具备一定的编程经验与硬件搭建能力。

4.1. 活动情景

本案例基于广州市无人驾驶出租车的现实场景，引导学生思考无人驾驶汽车在运行过程中可能出现的问题。课程内容分为协作计划、讨论协商、动手实践、回顾反思这四个阶段。

4.2. 活动目标

1.通过项目式活动，加深学生对无人车结构及功能的理解，感受人工智能对生活的影响。

2.通过人工智能课程协作脚本引导，帮助学生突破传统教学的限制，在亲身体验中获得深层次的人工智能技术感知。

4.3. 活动设计

学习者以小组为单位进行线下协作，通过虚拟仿真平台进行作品设计与制作。在协作计划阶段，学生观摩无人驾驶汽车案例，在组内分为论述者、监督者和总结者三种角色协作讨论确定主题。论述者需要提出作品主题的提议并作简要介绍，监督者把控讨论节奏，总结者对提议进行总结和记录。在讨论协商阶段，其余学生不明确设定发问者或回答者角色，学生自由讨论交流，对作品实现的细节（例如：无人驾驶汽车如何识别交通标志）进行协商。在动手实践阶段，设定一名学生作为监督者，其余学生通过虚拟仿真实验平台完成作品制作，在完成过程中互相帮助解决问题。在回顾反思阶段，组内分为评价者和总结者两种角色，对本组的合作过程及他组的仿真作品进行评价和反思。

5.总结

协作脚本作为计算机支持协作学习的一种重要学习模式，能够有效促进学生在小组协作学习形式中的学习投入程度。本研究以广州市地方教材五年级课程《无人驾驶汽车》为例实践了人工智能课程协作脚本模型。使学生能够深度参与到课程的活动内容中，充分保证其学习主体地位。诚然，本研究所设计的协作脚本模型较为初级，尚有完善的空间，今后可通过更多的课程内容进行验证及改进。此外，模型所设计的学生角色及角色轮换模式较为复杂，针对这一点，后续可进一步以课程内容为基础进行协作脚本模型中角色脚本的改进。

参考文献

银涛(2019)。浅谈人工智能在中小学教育中的应用。教育现代化,6(06):196-198。https://doi.org/10.16541/j.cnki.2095-8420.2019.06.069.

王临登(2022)。CSCL 环境下协作脚本对学习投入的影响研究(硕士学位论文,华东师范大学,上海市)。检索自中国知网数据库。

王智颖、翟芸和吴娟(2021).在线异步讨论中角色轮换脚本对大学生深度学习的影响。现代远程教育研究,33(03),100-112。

王小根和刘梦恒(2019)。在线知识建构中有效引发功能性冲突研究。 电化教育研究，(09),34-42. https://doi.org/10.13811/j.cnki.eer.2019.09.004

Beishuizen, J., & Steffens, K. (2011). *A conceptual framework for research on self-regulated learning.Sense Publishers*.

Dillenbourg, P. (1999). *What do you mean by collaborative learning?* UK: Elsevier Publishing.

Fischer, F., Kollar, I., & Stegmann, K. et al. (2013). Toward a script theory of guidance in computer-supported collaborative learning. Educational psychologist, 48(1), 56-66.

Gerlach, J.M. (1994). Is this collaboration. *New Directions for Teaching and Learning, 1994*, 5-14.

Järvelä, S., Kirschner, P.A., & Panadero, E. et al. Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Education Tech Research Dev 63*, 125 – 142 (2015).

Kobbe, L., Weinberger, A., &Dillenbourg, P. et al. Specifying computer-supported collaboration scripts. *Computer Supported Learning 2*, 211 – 224 (2007).

Kollar, I., Fischer, F., & Hesse, F.W.(2006).Collaboration Scripts–A conceptual analysis. *Educational Psychology Review, 2006*, 18, 159-185.

Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. *In Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community (pp. 72 – 81).* International Society of the Learning Sciences. https://doi.org/10.3115/1658616.1658627

Popov ,V., Biemans ,H.J.A., & Kuznetsov ,A.N. et al.(2014). Use of an interculturally enriched collaboration script in computer-supported collaborative learning in higher education. *Technology, Pedagogy and Education,23(3)*, 349-374.

同伴审查策略和先验知识水平对学生计算思维和认知负荷的影响

The Effects of Peer Review Strategies and Prior Knowledge Levels on Students' Computational Thinking and Cognitive Load

梁晨[1*]，陈燊[1]，林琳[2]，赖文华[1]，郭晓戈[1]
[1]温州大学教师教育学院
[2]瑞安市新纪元小学
[*]1835801545@qq.com

【摘要】 当前，在编程课程中使用协作学习来提高学生的计算思维，得到了许多学者的认可。然而，协作学习中的社会互动并不是自发的。如果没有提供适当的教学策略，学习效果可能会令人失望。基于此，不少研究人员使用同伴审查策略对学生计算思维的影响进行了探索，但都忽视了学习者先验知识对协作学习可能产生的影响。为此，本研究通过准实验的方式比较了在先验知识的调节下使用同伴审查策略与教师授课在计算思维和认知负荷方面的差异。研究发现，使用同伴审查策略对高低先验知识学生的计算思维提升和降低认知负荷方面具有显著的促进作用。
【关键词】 协作学习；同伴审查；先验知识；计算思维；认知负荷

Abstract： *At present, collaborative learning in programming courses is recognized by many scholars as enhancing students' computational thinking. social interaction in collaborative learning is not spontaneous. Without proper teaching strategies, learning outcomes may be disappointing. Many researchers have explored the impact of peer review strategies on computational thinking but have overlooked learners' prior knowledge in collaborative learning. Therefore, this study used a quasi-experimental approach to compare differences in computational thinking and cognitive load between peer review strategies mediated by prior knowledge and teacher instruction. The results show that peer review strategies have a positive impact on computational thinking and reduce cognitive load for students with different levels of prior knowledge.*
Keywords： Collaborative learning, Peer review, Prior knowledge, Computational thinking, Cognitive load

1. 问题提出

在过去数年里，研究人员对计算思维教育以及它在儿童获得思维技能和数字能力方面的作用越来越感兴趣（Angeli et al.,2020）。在计算思维培养上，编程教学被公认为是一种有效的途径，可以有效培养学生分析问题和解决问题的能力（Yang et al.,2020）。然而，在已有的信息科技课程中，编程学习还具有一定难度，需要学生进行长时间思考、时刻接受困惑和不确定性的考验（赵福君 等，2023）。特别是对初学者来讲，学习并独立完成编程任务还是一个困难的过程，这会加深学生的挫败感和畏难情绪，从而导致对学生计算思维的培养大打折扣。研究表明，在编程课堂中使用协作学习的方法可以提高学习学生的标称水平。但协作学习的社会互动并不是自发完成的，必须要有适当的指导策略进行支持。基于此，选取恰当的学习策略，以促进学生的社会互动是协作学习取得成功的关键。

同伴审查被认为一种有用的形成性评估方法，可以提供反馈，改善协作学习的效果，带来更好的协作学习体验（Mulder et al.,2014）。在同伴审查活动中，学生分别需要扮演评估者和被评估者的角色。作为评估者，他们需要审查同伴的工作并提供建设性的反馈；作为被评估者，他们可以选择接受反馈并做出相应的修改（Li et al.,2010）。在这个过程中，由于学生对于彼此的误解的理解，以及他们能够用熟悉的术语解释概念的能力，他们通常可以提供适当水平的帮助, 而不是超出新手水平的帮助（Brown et al.,2018）。借助同伴审查，学生可以对自己的表现进行深入反思，并在社会交互过程中提升自身的认知发展。然而，在这种情况下也并不是所有小组都能很好发挥作用，并促进个别学生的学习过程。因此，重要的是不仅要研究

同伴审查策略的效果，还要研究每个学习者头脑中发生的过程。考虑学习者在协作过程中面临的认知负荷可以为理解同伴审查有效和高效的条件提供更多的见解。

根据认知负荷理论，当群体成员经历积极的相互认知依赖时，他们就能够通过沟通和协调每个个体所拥有的知识来构建集体工作记忆资源，并获取比个人学习状况下更大的信息处理能力，这种分布优势使得协作个体相比单独学习的个体在执行任务时可以面临更少的认知负荷（Kirschner et al.,2011）。此外，先验知识作为学习者重要的认知基础，也会影响个体对于信息的加工方式和认知负荷，并最终表现在他们对于知识获取和技能提升上（Hathorn & Rawson, 2012）。然而当前鲜少有学者探讨关于同伴审查策略和先验知识如何共同作用于协作学习的过程，并研究对小学生计算思维和认知负荷的影响。因此，本研究旨在小学开展一项实证研究，并试图初步回答：（1）同伴审查策略和先验知识水平如何影响学生计算思维？（2）同伴审查策略和先验知识水平如何影响学生的认知负荷？

2. 研究设计

2.1. 研究对象

本实验选取以中国东部沿海地区浙江省某小学的学生作为被试，除去未全程参与实验以及数据存在异常的学生，本实验有效的被试人数共 92 人，其中男生和女生的人数占比相近，且学生均具有电脑使用经验。

2.2. 研究工具

本实验所作的准备包括计算思维量表、认知负荷量表。其中，计算思维量表改编自 Korkmaz（2017）量表， Cronbach's α 为 0.928，信度均为 0.95 以上；认知负荷量表改编自 Ryu Ji-heon（2010）量表，包括体力负担、精神努力、任务难度、自我评估和材料设计五个维度，Cronbach's α 为 0.82，信度均为 0.85 以上，以上问卷均具有良好的信效度。

2.3. 实验流程

本实验一共持续了 4 周，每周课堂实践为 40 分钟。在课前，教师对实验内容进行解释，并对学生的计算思维和先验知识水平进行测量。在前两周的课堂活动，学习内容以 python 基础知识为主，包括输入输出函数、while 循环、if 条件判断等。在后两周的课堂活动中，实验组采用了基于同伴审查的协作学习方法，在收到教师发布的任务单后，先进行独立的思考，自主完成任务单内容，然后和同伴进行分享比较, 在交流讨论的过程中，学生可以根据同伴提供的建议对自己的任务单进行完善修改。对照组采用传统的课堂讲授法，教师使用 ppt 来为学生讲解猜数字游戏的算法设计，并允许学生进行编程实践和独立填写任务单。如图 1 所示，在开展同伴审查活动过程，同伴针对任务单展开讨论、审查对方可能存在的问题并提出建设性的反馈。通过这种方式，学生能够对自己的表现进行深入的反思，并在交互过程中提升自身认知的发展。在课程全部结束后，教师回收所有学生的学习任务单，并要求学生填写计算思维后测问卷、认知负荷问卷。



图 1 课堂上同伴审查并讨论

3. 数据分析与结果

本研究采取双因素协方差分析，以计算思维前测为协变量，协变量与先验知识、组别交互作用的回归同质性被接受，F=2.493，F=2.608，p 值均大于 0.05，这表示数据符合回归同质性假定。协方差分析结果如表 1 所示。结果表明，先验知识和分组方式的主效应显著，即在计算思维上，实验组均显著高于控制组，高先验知识学生均显著高于低先验知识学生；在认知负荷上，实验组均显著低于控制组，高先验知识学生均显著低于低先验知识学生。此外，在计算思维上，是否使用同伴审查策略和先验知识之间不存在显著的交互作用，而在认知负荷上，两者交互作用显著。

表 1 计算思维、认知负荷的统计分析

	描述性统计 M（SD）				主效应		
	高先验知识		低先验知识		分组方式	先验知识	交互效应
	同伴审查	传统教学	同伴审查	传统教学			
计算思维	96.87（6.12）	91.70（13.09）	89.75（9.22）	78.05（9.73）	F (1,87) =15.567 p=0.000**	F(1，87)=11.135 *p*=0.001**	F (1,87) =3.512 *p*=0.064**
认知负荷	43.09 (9.40)	37.43 (10.33)	36.21 (11.34)	54.23 (11.65)	F (1,88) =7.657 *p*=0.007**	F (1,88) =4.921, *p*=0.029*	F (1,88) =28.052 *p*=0.000**

注：*表示 $p<0.05$, **表示 $p<0.01$

4. 总结与讨论

本研究探讨了同伴审查策略和先验知识水平对学生计算思维和认知负荷的影响，根据研究结果，本研究发现采用同伴审查策略能够有效提高学生的计算思维并减低他们的认知负荷。然而，从计算思维的角度来看，本研究并未发现同伴审查策略和学生先验知识水平之间存在显著的交互作用。这可能是被评估者有时并未真正理解评估者给出的建议、发布的任务较难造成的，从而使结果没有出现期望的“干预对低先验知识水平的学生更有效”的结论。因此，为防止低先验知识水平的学生在互动中被边缘化，教师必须要更细致地考虑教学干预的设计和个体学习需求之间的匹配程度，以及如何优化任务设置，以确保满足个体差异（Webb, 2009）。从认知负荷角度来看， 同伴审查策略和学生先验知识水平之间存在显著的交互作用。这与前人的研究结果保持一致，即异质分组会使具备较高先验知识水平的学生在协作学习活动中承担大部分的集体工作认知负荷。因此，在同伴审查活动中，较高先验知识水平的学生有可能会接收不到同伴的修改意见，从而导致外在认知负荷升高；相反，低先验知识水平的学生，他们能够从更有经验的学习者那里获取解决问题的指导，减少无指导的搜索过程。这一结论也再次印证了教学中存在着经验逆转效应，当一些教学指导与经验共同发挥作用时，对于低先验知识学习者有效的一些教学方法或指导原则，对高先验知识学习者可能并不能起到促进作用（Kalyuga et al.,2010）。

综上所述，本研究对于小学信息科技教学具有一定的借鉴意义。对于学生而言，同伴审查活动不仅可以促进学生积极参与和互动，避免在协作过程中出现“搭便车”行为，而且还可

以减轻学生在编程学习中的任务负担，培养计算思维。对于教师而言，评估者在一定程度上扮演了“小教师”的角色，帮助教师减轻了繁重的教学压力，使教师能够更集中精力处理更复杂、更需要专业指导的问题，这有助于提高教师的教学效率和质量，也使整个课堂更具有活力和互动性。但需要说明的是，此次研究的实践较短，短时间的数据难以捕捉学习者长期使用同伴审查策略后发生的变化，因此还需要更多的实践和探索，收集更全面的数据以进一步探讨同伴审查和先验知识的交互作用。

参考文献

赵福君, 刘妍君 & 李钰鹏.(2023).协同知识建构视角下高中生计算思维能力培养模式的设计研究. 兵团教育学院学报(03),10-16.

Angeli, C., & Giannakos, M. (2020). Computational thinking education: Issues and challenges. *Computers in human behavior*, *105*, 106185.

Brown, A. L., & Palincsar, A. S. (2018). Guided, cooperative learning and individual knowledge acquisition. In *Knowing, learning, and instruction* (pp. 393-451). Routledge.

Hathorn L G, Rawson K A. The roles of embedded monitoring requests and questions in improving mental models of computer-based scientific text[J]. Computers & Education, 2012,(3):1021-1031.

Kirschner, F., Paas, F., & Kirschner, P. A. (2011). Task complexity as a driver for collaborative learning efficiency: The collective working-memory effect. *Applied Cognitive Psychology*, *25*(4), 615-624.

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. Instructional Science, 38, 209-215.

Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). Computers in Human Behavior, 72, 558–569.

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. British journal of educational technology, 41(3), 525-536.

Mulder, R. A., Pearce, J. M., & Baik, C. (2014). Peer review in higher education: Student perceptions before and after participation. *Active Learning in Higher Education*, *15*(2), 157-171.

Ryu Ji-heon. The effect of prior knowledge level and computer screen division on task performance and cognitive load factors [J]. Educational Technology Research, 2010, 26(1): 87-109.

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. British Journal of Educational Psychology, 79(1), 1-28.

Yang, K., Liu, X., & Chen, G. (2020). The influence of robots on students’ computational thinking: A literature review. International Journal of Information and Education Technology, 10(8), 627-631.

促进大学生深度学习的在线同步同伴反馈研究

Online Synchronous Peer Feedback Research to Promote College Students' Deep Learning

陈孝然[1*]，傅莉[1]，朱敏捷[1]，李栩晴[1]
[1]华南师范大学 教育信息技术学院，广东广州 510631
[*] 1414639082@qq.com

【摘要】 数智时代，高校高度重视通过激发在线协作学习中生生深度互动来促进学生深度学习。基于在线协作平台实现同步同伴反馈有望成为激发学生高阶思维、加深学习投入的重要实现途径，但其视角效果和具体的机制路径尚不明确。在综合分析国内外面向深度学习的在线同伴反馈相关研究后，基于在线协作学习的四个交互机制和同伴反馈的五个基本流程，依托高校专业必修课《教育传播学》课程开展实证研究，通过相关性分析和行为序列分析等方法，分析和讨论在线同步同伴反馈对大学生深度学习的促进效果和其存在的行为表现特征与规律。

【关键词】 高校教育；在线协作学习；同步同伴反馈；深度学习

Abstract: *In the age of digital intelligence, universities attach great importance to promoting students' deep learning by stimulating deep interaction between students and students in online collaborative learning. Synchronous peer feedback based on online collaboration platform is expected to be an important way to stimulate students' higher order thinking and deepen learning engagement, but its perspective effect and specific mechanism are still unclear. After a comprehensive analysis of domestic and foreign research on online peer feedback for deep learning, based on the four interaction mechanisms of online collaborative learning and the five basic processes of peer feedback, empirical research is carried out relying on the compulsory course of Education Communication in colleges and universities, and through correlation analysis and behavioral sequence analysis, etc. This paper analyzes and discusses the effect of online synchronous peer feedback on deep learning of college students and its behavior characteristics and rules.*

Keywords: Higher education, Online collaborative learning, Synchronous peer feedback, Deep learning

1. 引言

随着全球科技的迅猛发展和全球化趋势，教育正朝数字化、智能化和高度个性化的方向发展（邵艳芬，2023），培养知识与思维相关联、具备协作学习能力和创新学习能力的高素质人才已经成为当今21世纪人才培养的主要要求。深度学习通过促使学生深度参与学习、采用高阶的学习与交流互动策略来促进学习者高阶知能的发展，能够满足对批判性思维、交流协作、问题解决等新时代创新型人才所需的能力培养需求（邢星，2018）。2015（L. 约翰逊等）到

2017（S·亚当斯贝克尔等）连续三年（L·约翰逊，2016）的《地平线报告（高等教育版）》中将深度学习（深层学习）方法作为推动学校教育技术应用的长期趋势之一。2018 年《教育部等六部门关于实施基础学科拔尖学生培养计划 2.0 的意见》中也指出“要创新学习方式，促进学生自主深度学习、建构知识体系、形成多维能力”（汪雅君等，2019）。可见促进学生深度学习已经成为国内外高校教育改革的一个重要方向。

同伴反馈作为一种课堂反馈形式（张军等，2020），也是协作学习的重要组成部分，要求学生能够对为同伴之间研讨话题、任务或存在的问题提供反馈。其中基于在线平台开展在线同步同伴反馈，能够协力促进自己和其他成员实现小组共同学习目标，有望实现培养学生的高阶认知与解决问题的能力，促进学生的深度学习。故本研究拟通过开展单组前后测准实验研究，深入研究与分析在线同步同伴反馈对大学生深度学习的影响。

2.相关研究综述

2.1. 学生深度学习的内涵与评价相关研究

学习科学领域将深度学习界定为学生主动获得多样化学习资源和结构化知识后，通过多维知识建构而形成的由浅入深的认知加工活动，从而实现对知识与技能全面完整、具体深入的理解本质并加以应用（祝智庭等，2017）。最直观表现在学生学习过程的质量，可以从学生在学习活动中的投入深度和对知识信息的建构与加工深度上加以判断。已有众多国内外专家分别从布鲁姆认知目标分类角度（Marton F，1076；Ramsden P，1981）、深度与浅表学习特征（Biggs J，1999；张浩等，2012）和课程（Warburton K，2003；袁金丽等，2022）等角度对深度学习的内涵进行阐释，不同研究视角各有侧重又有所关联。

较为权威且应用广泛的深度学习评价方式大致包括三大类：①面向过程的深度学习效果评价，主要是针对学习投入状态和学习方法展开的量化评价，如 Biggs（1999）面向中小学生深度学习学习过程测评的问卷和 Entwistle（1982）的学习方法问卷等。②面向结果的深度学习评价，主要包括基于权威评估模型的深层认知评估，以及基于问卷量表的高阶能力素养测评。其中较具权威性的几大评估模型包括，布鲁姆认知目标分类法、Web 的 DOK 知识深度模型和 Biggs 的 SOLO 学习结果分类框架（1982）。以及综合多个模型与理论构建的优化模型，如唐金娟等人（2011）基于布鲁姆目标分类、信息素养和认知心理学信息加工等理论所编制的数字化深度学习综合能力评估量表，张浩等人（2014）以布鲁姆的认知目标分类法等。③面向深度学习能力或素养的评价，国际上较权威的代表性评估工具主要包括 NRC、WFHF 和 AIR 共同提出的“深度学习领域与能力框架”（ Huberman，2020）、AIR（2018）开展的深度学习研究项目编制的“深度学习机会与结果”学生问卷等。

2.2.在线同步同伴反馈的相关研究现状

同伴反馈是教育教学中仅次于教师反馈的常见反馈类型，在高校课程混合学习和小组协作学习中均具有重要作用（盖淑华等，2013）。依据反馈依托的不同物理教学环境，同步反馈可分为在线同步同伴反馈、在线异步同伴反馈（姚佳佳等，2022）、面对面同伴反馈（Georgeta, Ion et al，2016）、线上线下混合式同伴反馈（王云等，2021）等。其中在线同步同伴反馈是指依托课程的线上协作互动平台进行同步的互动与评价，具体而言包括：①以在线直播、协作文档为代表的可多人实时同步互动、协作与反馈的平台，通过上麦互动、即时分析的线上讨论区互动和可视化同步协作等进行小组内部的观点争论与评价（胡天慧等，2022）；②以思维导图、三维概念图为代表的可视化在线协作工具，通过对已发布的协作探究任务和问题进行拆解分工，在小组内部同时段对问题解决方案展开协作共建，同时同伴间进行相互的补

充与反馈，从而优化可视化成果。

3.研究设计

3.1.研究问题

本研究旨在对在线协作学习活动中同步同伴反馈对大学生深度学习能力水平作用进行调查与分析，探究在线同步同伴反馈条件下学习者的深度学习能力层次的变化，以及所表现出来的群体行为特征。具体包括两个研究问题：①如何验证在线同步同伴反馈对大学生深度学习的提升效果？②大学生的在线同步同伴反馈行为序列转换具有哪些特征？

3.2.在线同步同伴反馈活动设计

本研究参考国际权威的PCF同伴纠正性反馈模型（Sato M，2017）和ICAP深度学习框架（Chi，2014），设计了面向深度学习能力提升的在线同步同伴反馈活动模型，如图1所示。第一，同步同伴反馈的5流程包括，①认知反馈：学生能够比较和对照他们自身对任务问题的认识和他人对其认知的差异，要求学生通过同伴之间的讨论，来比较、对照自己与教材或教师提供的教案材料中的示范样例之间的区别。②任务求解反馈：学生可以借助在线协作工具和在线会议平台来进行任务解决，要求小组内部进行充分的协商、质疑、批判和反思，用以推动任务的深化解决。③任务突破反馈：学生在与同伴交流过程中明确对任务价值的感知和参与任务的动机，形成反馈应答和任务深度理解。④纠正性反馈：在面对同伴的错误认知、错误问题解决或应用错误规则时，学生能够即时告知同伴存在错误或问题的内容、原因，并介绍改正的做法和途径。⑤成果检验反馈：学生通过协作实施反馈指导方案，并检验当前形成成果的有效性，对标任务目标优化和完善小组成果。第二，明晰学生在在线协作学习过程中四层深度学习互动机制，①组内交互：是小组内部在围绕协作学习任务的目标，开展的交互活动行为，即保证学生个体层面能够融入到小组任务中，也能够促进学生后续深度同伴反馈的建立。②知识建构：学生在知识建构层面进来回探讨，进行高效快速的知识建构。由于知识建构是在紧密的组内交互基础上进行，学生能够进行比较、分析、综合、批判、反思和得出结论，由此促进学生群体的深度学习。③纠错：针对在线协作学习任务求解与突破内容展开询问、评价和批判。④反思调控：学生需要反思小组主题任务的解决方法，确定问题，并制定行动计划和改进措施。

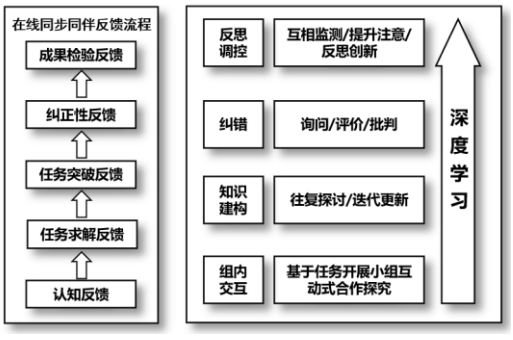


图1 在线同步同伴反馈活动模型

3.3.研究工具

1.深度学习能力测评问卷。本研究基于引用国际成熟的权威量表进行问卷设计的原则，综合参考NRC的深度学习领域与能力框架和AIR的深度学习机会与结果学生问卷，确定了深度学习能力测评的六个子维度，分别是小组协作能力、沟通表达能力、自主学习能力、学习毅

力、创造性思维能力、批判性思维能力。邀请 NX 省某大学的 21 位大二学生开展预调研，分析得到问卷整体信度为 0.902，各个子能力维度信度也均高于 0.8，说明问卷结构良好。

2.在线同步同伴反馈行为编码表。本研究参考 Burgess 等人(2021)提出的同伴反馈编码，如表 1 所示。其中反馈等级的校准包括，评审员（J1/J2/J3）最初进行了校准练习，每位评审员独立评估相同的学生反馈实例。在正式编码前进行预编码测试，得到评分者一致性系数为 0.819，即说明该编码表信度良好。

表 1 在线同步同伴反馈行为编码表

序号	反馈类型	行为指标	编码
1	认知反馈	倾听并尊重他人的意见	A
		鼓励、赞成、判断成员观点	B
2	任务求解反馈	与小组内的其他人密切协作	C
		针对任务中的某些材料/观点，反馈自己的困惑和疑问	D
3	任务突破反馈	清楚的表达支撑任务问题解决与突破的关键性意见和看法	E
		创造性地解决小组协作中的问题	F
4	纠正性反馈	对小组作品中存在的问题或不恰当之处进行纠正	G
		接受批评和反馈，并作出适当的回应	H
5	成果检验反馈	小组任务作品汇报后的反思性反馈（从优化作品视角出发）	I

3.4.研究设计与实施

本研究采用单组前后测准实验研究法，在 H 校教育技术学专业的《教育传播学（2022）》线上线下混合式本科课程的线上协作写环节进行教学实践，以验证在线同步同伴反馈对学生深度学习提升的可行性和有效性。研究对象为 35 位大二学生（其中男生 9 人，女生 26 人）。实验期间共开展 8 周活动实践。

实践后对 8 场的课在线同步小组协作活动实时录音频进行收集、清洗和分析。具体包括：①班级内的 11 个小组分别在每场在线协作活动中生成各自的研讨视频,每个视音频时长在 20 分钟左右，共计生成 88 个视音频，视频总时长大约 29 小时。②视音频借助“飞书”软件的自动转录文本功能，形成初步的文本数据文档。③编码者以事件为单位切片，依据同伴反馈行为编码表进行编码,得到按照时序排列的完整的在线同步同伴反馈行为序列样本,共计 6031 个行为编码。

4.数据分析结果

4.1.深度学习能力水平提升检验

经检验本研究的深度学习能力水平问卷前后测信效度均较好，数据呈现出整体的正态分布，进行配对样本 T 检验结果显示：实验对象的深度学习能力整体显著提高（t = 3.853，p < 0.001），且六个子维度均有不同程度的提升（如表 2 所示）提升呈现显著性的统计学特征。表明经过 8 周的在线同步同伴反馈活动的实施，学生的深度学习能力具有明显的提升。

表 2 深度学习能力前后测配对样本 T 检验结果

成对差分					t	df	Sig.（双侧）
均值	标准差	均值的标准误	差分的 95%置信区间 下限	差分的 95%置信区间 上限			

深度学习能力水平（后测-前测）	14.278	22.235	3.706	21.801	6.755	3.853	35	< .001
小组协作能力（后测-前测）	2.417	4.003	0.667	1.062	3.771	3.623	35	0.001
沟通表达能力（后测-前测）	1.889	3.446	0.574	0.723	3.055	3.289	35	0.002
自主学习能力（后测-前测）	2.944	4.459	0.743	1.436	4.453	3.962	35	< .001
学习毅力（后测-前测）	1.722	4.287	0.714	0.272	3.173	2.410	35	0.021
创造性思维能力（后测-前测）	2.028	4.088	0.681	0.645	3.411	2.976	35	0.005
批判性思维能力（后测-前测）	3.278	5.219	0.870	1.512	5.044	3.768	35	0.001

4.2.在线同步同伴反馈行为要素与深度学习的相关分析

为了进一步验证学生深度学习能力提升与本研究在线同步同伴反馈活动之间的联系，将在线协作活动互动过程中同步同伴反馈行为序列的总频次与学生深度学习能力测评分数在SPSS26.0软件中进行 Pearson 相关分析（Collins，1975），结果表明二者呈显著性的正相关($r=0.839, p<0.01$)。理论上学生样本在同步同伴反馈中可以产生81类同伴反馈行为序列。根据对编码数据分类统计，将真实出现的32个同伴反馈行为序列与深度学习能力测评得分进行Pearson相关分析，结果如表3所示，有13类同伴反馈行为序列与深度学习能力呈相关显著，即本研究在线同步同伴反馈多个行为转换均与学生深度学习能力呈现正向相关关系。

表3 与学生深度学习显著相关的同步同伴反馈行为投入序列

行为序列编码	Pearson 相关性	Sig.（双尾）	行为序列编码	Pearson 相关性	Sig.（双尾）
A→D	0.562**	.000	D→E	0.579**	.000
A→G	0.677**	.000	E→A	0.824**	.000
A→H	0.744**	.000	E→C	0.470**	.004
B→E	0.554**	.000	G→B	0.490**	.002
C→A	0.683**	.000	H→A	0.706**	.000
C→C	0.438**	.007	H→B	0.343*	.040
D→A	0.531**	.001			

注：*P＜0.05；**P＜0.01

4.3.在线同步同伴反馈行为序列特征分析

利用GSEQ5软件进行之后序列分析，得到行为转化频率表和调整后的残差表。提取Z分数大于1.96即统计学上认为达到显著性水平（Berk et al.，1992）的行为序列绘制“螺旋循环”在线同步同伴反馈行为转换特征图（如图2所示）。其中包括十个显著行为序列AH、AI、BD、BE、CC、DD、EH、FB、GG、IA。其中，AH、AI、BD、BE、CC、DD、EH、GG。

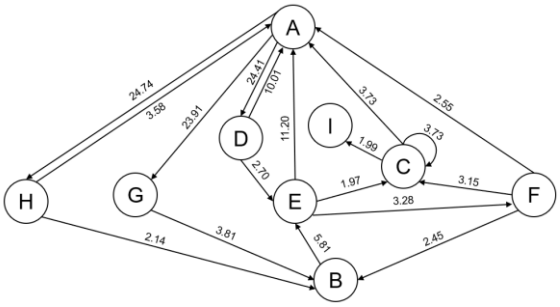


图2 “螺旋循环”在线同步同伴反馈行为转换

分析图 2 发现，研究对象形成了类似螺旋的循环体行为转换样态，其中代表性的路径有：①以 AD 为起点的 ADE(A)\(FA)\(CA)CI 的进阶类在线同步同伴反馈序列链，实现从“认知反馈→任务突破反馈→纠正性反馈→任务求解反馈→成果检验反馈”同伴反馈的行为路径，同时从该序列链上可以看出，进阶类在线同步同伴反馈在整个协作问题解决过程中，可以多次“任务突破反馈→认知反馈”“纠正性反馈→认知反馈”“任务求解反馈→认知反馈”往复循环，直至最终实现成果检验。②以 AG(H)为起点的 AG(H)B(EDA)\(EF)\(E)CI 的纠正类在线同步同伴反馈序列链，实现从“认知反馈→纠正性反馈→任务突破反馈→任务求解反馈→成果检验反馈”同伴反馈的行为路径，该路径反映了学生间的相互指正和纠错是小组后续深度协作的基础，同时从该序列链上可以看出，纠正性同伴反馈在任务解决过程中，可以多次在“纠正性反馈”“任务突破反馈→任务求解反馈→认知反馈”“任务突破反馈”往复循环，直至最终实现成果检验。

5. 结论与讨论

本研究设计了同步同伴反馈活动流程应用于高校本科课程的在线协作学习，并开展了单组前后测的准实验研究，在综合运用问卷调查、相关性分析、滞后行为序列分析等多种方法的基础上，从多角度探究了在线同步同伴反馈对大学生深度学习能力水平的影响。研究发现：

5.1. 实践标准化在线同步同伴反馈后，学生的深度学习能力水平显著提升

根据对 35 位研究对象的深度学习能力测评数据进行前后测数据的配对样本 t 检验结果（$t = 3.853$，$p < 0.001$），可见实施本研究构建的同伴反馈同策略后对学生的深度学习能力总体水平具有显著促进作用，即是实验对象的深度学习能力显著提高。结合深度学习能力后测问卷中最后一道“基于对 8 周在线协作任务解决活动的参与，谈谈您的收获与建议”开放性题目，学生们传达了众多正向的反馈，如“我感觉到了整个助教团队的用心，设计的活动和深度参与的同伴反馈除了提高了我的综合能力，最重要的是可以让我在协作交流中乐在其中，认真完成任务，自认为成为了一名优秀的小组成员！”，如“我觉得在同伴相互反馈的过程中学习到了很多同伴的优点和长处，对我来说非常有帮助”，又如“合作完成任务进行和同伴的交流中，我觉得能够集思广益，达到比独自学习更好的效果”，可见在 8 周高密度高质量的同伴反馈，收到了大多数学生的认可，大多数学生自觉自身的深度学习能力水平。

5.2. 在线同步同伴反馈多种行为转换与学生的深度学习能力呈现显著正向相关

根据相关性分析结果可见，在线同步同伴反馈策略总行为与学生的深度学习具有显著正相关($r=0.839$，$p<0.01$)，即说明在线协作学习中的同步同伴反馈策略整体而言能够促进学生的深度学习。另一方面，本研究的策略中有 13 类同步同伴反馈行为序列与深度学习能力呈显著正向相关。由此可见，在考虑到教育实验伦理未进行对照组实验的情况下，通过相关性分析验证了在线同步同伴反馈多种行为与学生深度学习结果的正向相关关系。例如“倾听并尊重他人意见→反馈自己的困惑和疑问”、“倾听并尊重他人意见→对任务解决/作品中存在的问题进行纠正”等同伴反馈行为转换，能够通过实现从“倾听中知识建构→批判性反思”的深层次学习生成。

5.3. 同步同伴反馈行为转换呈现“螺旋循环”的路径模式

通过对学生在线协作学习中同步同伴反馈的视音频数据进行编码和滞后序列分析，得到了“螺旋循环”同伴反馈行为转换模式：以 AD 为起点的 ADE(A)\(FA)\(CA)CI 的进阶类同伴反馈序列链和以 AG(H)为起点的 AG(H)B(EDA)\(EF)\(E)CI 的纠正类同伴反馈序列链，验证了本研究提出的同步同伴反馈在在线协作学习环境的应用下，学生能够在任务解决过程中实现循

环往复、交互逐渐深入的同伴反馈，其中的任务求解反馈、任务突破反馈、纠正反馈到认知反馈的逆向循环，表明了小组作品在多次同伴反馈中进行了认知的修正与重塑，无论是进阶类同伴反馈序列链还是纠正类类同伴反馈序列链，均最终均达成作品（成果）检验，促进学生的深度学习和知识建构。此外需要注意的是，同伴反馈的行为转换并不是越复杂越优，需要促进递进式、进阶式同伴反馈行为演进（姚佳佳等，2023）。

本研究虽然对促进大学生深度学习的在线同步同伴反馈机制与特征开展了具有一定创新性的研究，但也存在着研究遗憾和不足，一方面实证缺乏对照组，采用相关性分析验证同伴反馈行为和学生深度学习能力提升之间存在正向相关。另一方面，本研究重点探究了在线协作任务中组内的同步同伴反馈，后续可以进一步拓展探究组间竞合的异步同伴反馈内涵、方式、策略与行为特征。

参考文献

王云,李志霞,白清玉,姚海莹,汪存友（2021）.混合式教学中促进深度学习的同伴反馈研究.现代教育技术, 31(05):67-74.

邢星（2018）.教育信息化 2.0:深度学习、学校变革、智能治理.人民教育,(Z2):122-123.

张浩,吴秀娟（2003）. 深度学习的内涵及认知理论基础探析.中国电化教育, 2012, 10(11).

张军 程晓龙.(2020).我国近十年同伴反馈研究:回顾与展望.西安外国语大学学报(01),48-55.

邵艳芬(2023).高等教育数字化的缘由、理念与发展理路.现代教育管理,(08):31-39.

汪雅君,何晓萍（2019）.《地平线报告（2019 年高教版）》对高校信息化建设的启示.中国教育信息化,(23):6-12.

胡天慧,刘三女牙,粟柱,周子荷,余建伟 柴唤友.(2022).学习者同步和异步交互模式的比较与整合研究——基于社会网络和多层网络分析的方法.电化教育研究(05),53-60.

姚佳佳,李艳,潘金晶,程萌萌（2022）.同伴对话反馈对大学生在线深度学习的影响研究.华东师范大学学报(教育科学版),40(03):112-126.

祝智庭, 彭红超（2017）. 深度学习: 智慧教育的核心支柱. 中国教育学刊, (5): 36-45.

袁金丽,郭志涛（2022）.深度学习为核心的高校智慧教育实践路径研究.河北师范大学学报(教育科学版), 24(04):68-74.

盖淑华 周小春.(2013).基于动态系统理论的同伴反馈实证研究——社会文化观视角.外语与外语教学(02),36-40.

L.约翰逊,S.亚当斯贝克尔,V.埃斯特拉达,A.弗里曼,张铁道,殷丙山,吴莎莎,吴亚婕,杜贤,李玉（2015）.新媒体联盟地平线报告(2015 高等教育版).北京广播电视大学学报,(S1):4-52.

L·约翰逊,S·亚当斯·贝克尔,M·卡明斯,V·埃斯特拉达,A·弗里曼,C·霍尔,殷丙山,高茜,赵广元,李院春,李志强,冯振振（2016）.新媒体联盟地平线报告:2016 高等教育版.开放学习研究,(02):1-20.

S·亚当斯贝克尔,M·卡明斯,A·戴维斯,A·弗里曼,C·霍尔给辛格,V·安娜塔娜额亚婻,殷丙山,高茜,任直,刘鑫驰,曹红岩,王济军,赵广元,邵恒（2017）.新媒体联盟地平线报告:2017 高等教育版.开放学习研究,(02):1-20+62.

AIR （2018）. The study of deeper learning: Opportunities and outcomes *Student Survey Documentation*. Retrieved October 12, 2018, from https://air.prod.acquia-sites.com/sites/default/files/Deeper_Learning_Student_508-rev.pdf.

Biggs J（1999）. What the student does: Teaching for enhanced learning. *Higher education research*

& development, 18(1): 57-75.

Burgess A , Roberts C , Lane A S , et al（2021）. Peer review in team-based learning: influencing feedback literacy. *BMC Medical Education*, (1).

Chi M T H, Wylie R（2014）. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, 49(4): 219-243.

Collins, A. M., & Loftus, E. F（1975）. A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407-427

Entwistle, N., & Ramsden, P（1982）. Understanding student learning (ED244959). ERIC. 1982.

Huberman, M., Bitter, C., & Anthony, J. et al（2020）. The Shape of Deeper Learning: Strategies, *Structures, and Cultures in Deeper Learning Network High Schools*[DB/OL].[2020-08-04].

Ion, G., Barrera-Corominas, A., & Tomàs-Folch, M（2016）. Written peer-feedback to enhance students' current and future learning. *International Journal of Educational Technology in Higher Education*, 13(1).

Warburton K（2003）. Deep learning and education for sustainability. *International Journal of Sustainability in Higher Education*, 4(1): 44-56.

Marton F, Säljö R（1976）. On qualitative differences in learning: I—Outcome and process. *British journal of educational psychology*, 46(1): 4-11.

Ramsden P, Entwistle N J（1981）. Effects of academic departments on students' approaches to studying. *British journal of educational psychology*, 51(3): 368-383.

Sato M（2017）. Oral peer corrective feedback: Multiple theoretical perspectives.

姚佳佳,李艳,陈新亚, 杨智超.(2023).不同可视化实时互动界面对大学生同伴对话反馈的影响研究.远程教育杂志(01),102-112.

科学教育中生成性绘图的插图丰富度对注意力分配及学习效果的影响

Investigating the Impact of Illustration Richness in Generative Drawing on Attention Allocation and Learning Outcomes in Science Education

章 仪[1]，何静[2]，皮忠玲[3*]，杨九民[2]
[1]北京师范大学 未来教育学院
[2]华中师范大学 人工智能教育学部
[3]陕西师范大学
* pizl@snnu.edu.cn

【摘要】 本研究通过分析学习者在观看带有不同插图的科学教学视频时的行为，探索了是否使用插图和插图类型对注意力分配、绘图质量和测试成绩的影响。92 名大学生被随机分配到观看无插图、抽象插图和具象插图视频中，他们观看视频的注意分配模式、绘图的质量、测试成绩均被记录。研究结果显示，无插图情境下学习者更多地关注文本信息，而抽象插图有助于提高绘图质量，可能因为它们激发了更深层次的认知加工。此外，具象插图的直观性和具体性对即时测试成绩的提升有显著帮助。

【关键词】 科学学习；生成性绘图；图片丰富度；注意力分配

Abstract: *This study explored the impact of the use and type of illustrations on attention allocation, drawing quality, and test performance by analyzing learners' behavior while watching scientific educational videos with different types of illustrations. Ninety-two university students were randomly assigned to watch videos with no illustrations, abstract illustrations, or concrete illustrations. Their attention allocation patterns while watching the videos, the quality of their drawings, and their test scores were recorded. The results showed that learners in the no-illustration condition focused more on textual information, while abstract illustrations helped improve the quality of drawings, possibly because they stimulated deeper cognitive processing. Additionally, the intuitiveness and specificity of concrete illustrations significantly aided in enhancing performance on immediate tests.*

Keywords: Science Learning; Generative Drawing; Picture Richness; Attention Allocation

The work was support by the Special Funds for the Research on Teacher Development under Grant [2023JSY008] and the Philosophy and Social Sciences Planning Youth Project Of Guangdong Province under Grant [GD24YJY01].

1. 前言

自科学诞生之初，可视化就与之紧密相连。科学家们，如法拉第和麦克斯韦在研究磁感线，以及发现 DNA 双螺旋结构的科学家，都通过可视化的手段来构建新的关系、测试新理念和详细阐述知识（Ainsworth, Prain, & Tytler, 2011）。鉴于此，科学教育中的绘图受到研究者的广泛关注。一项权威的综述强调了科学绘图在课堂活动中的五个价值，包括：(a) 提高学生参与度；(b) 教授学生如何在科学中表达意义；(c) 促进学生参与科学推理；(d) 学习科学概念和过程；以及 (e) 传达科学主张。鉴于可视化在科学教育中的普遍应用，探究哪些可视化工具最能促进科学绘图，以及它们是否对所有学生同样有效，变得格外重要。可视化不仅支持科学绘图，而且可能成为一个特别有用的工具，尤其是在教授那些对学习者来说难以理解

或违反直觉的科学主题时。这些主题可能会对学生学习新概念和将知识应用于新情境构成挑战。在本研究中，我们探讨了图片的感知丰富性（即，图表包含的视觉特征数量）如何影响科学绘图的有效性。

绘图策略是指学习者在任何媒介上生成的二维描述性视觉表征，这包括了结构、关系或过程的表达。使用绘图策略时，学习者进行的绘图活动导致生成性加工发生，激发了学习者的主动学习倾向，因此被认为对学习有利。根据绘图效应的整合成分模型，绘图过程涉及以下几个关键步骤：**信息内容的细化（Elaboration）**，学习者首先需要理解和处理所学的信息内容，提取出关键要素和概念；选择合适的信息，在理解了信息之后，学习者需要选择哪些信息是重要的，哪些是可以被形象化的。这个过程要求学习者对信息进行筛选和抽象；**创建视觉图像（Visual Imagery）**，学习者在脑海中构建出与所学信息相关的视觉图像，这一步骤是将抽象信息转换成具体的视觉表征的过程；**通过手部动作完成绘图（Motor Action and Pictorial Representation）**，学习者通过手部动作将心中的视觉图像转化为实际的图画（绘图），这不仅是一个物理动作过程，也是对学习内容进行视觉化表达的过程。在整个绘图过程中，学习者不断地进行信息的选择、组织和整合。这种综合性的认知活动有助于加深对材料的理解和记忆，从而提高学习效果。通过将抽象的概念转换为具体的图像，绘图策略有效地增强了信息的加工深度，促进了学习者对知识的内化和应用。

一些研究显示绘图策略有助于学习，但也有研究揭示了它对学习的潜在消极影响。例如，Alesandrini（1981）的研究中要求本科生在阅读电化学概念的科学章节时进行绘图、写口头释义或反复阅读。结果显示，绘图组的学习者在学习测试中表现优于口头改写或反复阅读的学习者。Leopold 和 Leutner（2012）的研究也表明，在学习化学知识时，要求十年级学生进行绘图比总结能获得更好的学习成绩。然而，其他研究则提出相反的观点。Leutner, Leopold 和 Sumfleth（2009）发现，在学习化学知识（如水分子结构特征）时，要求十年级学生进行想象比绘图能带来更好的理解。Ploetzner 和 Fillisch（2017）的研究也显示，相比于采用反思策略的学习者，采用绘图策略的学习者在保持知识方面表现较差，绘图策略甚至可能阻碍学习内容的再认。

这些不一致的研究结果可能与学习材料的丰富性有关。例如，Fiorella 等人（2020）的研究发现，在学习具象且含有详细图片的人体肝脏系统时，采用绘图策略并不是最佳选择。这可能是因为绘图活动增加了学习者的认知负荷，导致对学习的损害。这表明，绘图策略的有效性可能受到学习材料具象性和学习者认知能力等因素的影响。在某些情境下，绘图可能是一种有益的学习工具，但在处理高度具象或详细的材料时，它可能反而成为认知上的负担。

图片的丰富度或逼真程度通常通过视觉表示与现实的匹配程度来衡量。在这方面，线条图等可视化形式被认为是最不逼真的类型，因为这些示意图通常会简化具象的形状，并省略某些对于特定任务可能不必要或过于详细的信息（Goldstone & Son, 2005）。因此，图片可以根据其表征的详细程度被区分为详细和抽象两类。图片的感知丰富性可能会影响信息的选择、组织和整合。根据诱人细节效应，如果给学习者提供过于丰富的细节，学习者可能会被这些细节吸引，从而将认知资源转移到处理这些细节上，而非专注于与课程相关的内容。Rey（2012）在其元分析研究中发现，提供丰富的细节会损害学习。然而，在他的元分析中，只有八项研究讨论了图片中的诱人细节，大部分研究关注的都是无关文本的细节问题。

总体而言，研究普遍认为，相比于含有具象细节的图片，学习者可以从简化的图像中学得更好。例如，在一项关于 6 至 8 岁儿童学习阅读抽象条形图的研究中，当课程涉及无额外细节的图表时，儿童学得更多，转移得更好（Kaminski & Sloutsky, 2013）。同样，在学习循环系统时，成人使用简化图表比使用详细、解剖学上正确的图表学得更多（Butcher, 2006）。

这些发现支持了简化图像在教育和学习过程中的优势，强调了在设计教育材料时考虑图像的逼真程度和细节水平的重要性。

然而，抽象图片的优越性也受到了一些研究的质疑。Siler 和 Willows（2014）的研究发现，在包含具体但相关细节的模块算术课程中，学习者的表现优于使用抽象表示的相似课程。此外，McNeil 和 Fyfe（2012）指出，从具体表示开始，逐步过渡到更抽象表示的模块算术课程（称为具体性逐渐减弱），比仅使用抽象表示的课程更有助于学习和知识转移。虽然高度现实的视觉化可能增加认知负荷，但 Skulmowski（2022）的研究表明，这种类型的视觉化在记忆保持方面可能有更好的效果。

有趣的是，Menendez 等人（2020, 2022）的研究发现，成人通常从简略（抽象）的表征中获得更多学习利益，而小学生则更容易从详细（逼真）的作品中学习。这种差异可能与学习者如何从图片中获益有关，这取决于他们对图片内容的加工方式。Butcher（2006）指出，学习者的图片加工方式对学习效果有重要影响。Jaeger 和 Fiorella（2023）进一步发现，只有当学习者能够精确地利用图片线索进行回忆加工时，他们才能从图片中获益；如果不能这样做，提供图片并不会带来任何优势。这些研究结果表明，图片的有效性不是绝对的，而是取决于多种因素，如学习者的年龄、认知发展水平、学习任务的性质以及学习者的信息加工方式等。

到目前为止，在绘图研究领域中，尚未有研究专门关注提供图片的具象性如何影响绘图活动，并进而影响学习结果。因此，本研究旨在探究在科学学习中，提供不同类型的图片如何差异性地影响学习效果。为了更深入地探索学习者对材料的加工过程，本研究不仅考察了学习效果和绘图质量，还将关注学习过程中学习者的眼动情况。以往关于绘图的研究表明，绘图策略要求学习者完成信息的整合，整个过程涉及语言和非语言信息的转化。因此，学习者需要对关键信息进行反复加工，以确保构建的图像表征准确无误。例如，Hellenbrand 等人（2019）的研究发现，采用绘图策略的学习者相比观看教师绘图组和总结组的学习者，更频繁地重复阅读词语，对重点信息有更高的注视时间比例，并在学习材料和绘图区域之间有更多的转换次数。本研究借鉴 Hellenbrand 等人（2019）的研究，采用了文本和图片之间的转换次数、注视文本和图片的时间比例作为衡量指标。基于以上考虑，本研究提出以下研究假设：

（1）相比于没有插图和具象插图，提供抽象插图的视频会让学习者更多地关注图片内容，更少地关注文字内容。

（2）相比于没有插图和抽象插图，提供具象插图的视频会让学习者表现出更好地成绩。

（3）相比于没有插图和抽象插图，提供具象插图的视频会让学习者画出更好地图片。

2. 研究方法

2.1. 被试

本研究共招募了 92 名来自湖北武汉某高校的本科及研究生，他们所学专业分别为经济学、教育技术学、电子信息、体育教育、法学、心理学、物理学、日语和西班牙语等。招募方式为通过 QQ、微信等社交平台发布通知，所有招募的被试均听力正常，无近视或矫正后视力正常，无精神疾病。在学习之前，所有被试均了解实验主题，且签署了知情同意书，完成实验后，被试根据表现得到得到 20-25 人民币的被试费和一包餐巾纸作为参与实验的酬劳，被试的具体基本信息如下：

表 1 被试基本信息统计表

性别		年龄（18-28）		学位		人数
男	女	M	SD	本科生	研究生	
13	79	21.85	2.21	64	28	92

2.2. 实验材料与仪器

2.2.1. 实验材料

结合以往研究结论，在对有关科学知识的传授上，使用绘图策略对学习有显著的积极影响，为此，本研究结合以往研究，选择科学知识--《心脏结构和功能》。内容主要包括以下四个方面：心脏的四个腔室，心脏的瓣膜，心脏的功能和血液循环。

为验证实验假设，本实验设计了三种视频学习材料，分别为无插图教学视频、含抽象插图视频和含具体插图视频（下分别简称为，无插图、抽象插图和具体插图；见图 1），具体见下图所示。

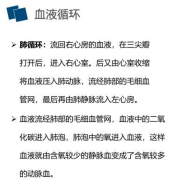


图 1a 无插图教学视频材料

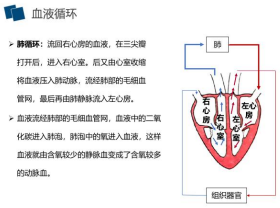


图 1b 抽象插图教学视频材料

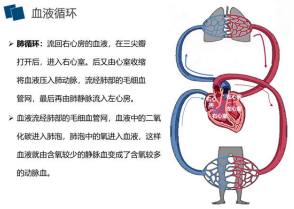


图 1c 具象插图教学视频材料

2.2.2. 实验仪器

本实验的眼动数据主要采用 Eyelink 1000 眼动仪。台式电脑是 Windows7 操作系统，视频分辨率为 1024×768，刷新率为 75 HZ。被试坐在距离显示器 60cm 的距离观看屏幕，眼动仪的采样率为 1000 HZ，采用九点矫正，双眼追踪。对被试进行眼动数据收集过程中，被试要求将头固定在头托上。

2.2.3. 测量工具

（1）人口统计问卷

该问卷用于统计被试的性别、年龄等人口学基本信息。

（2）先前知识经验测试

根据视频材料，本研究自行编制了先前知识测验题目。该测验包括 5 道单项选择题和 3 道填空题，一共 8 道题，其中选择题有 5 个选项，最后一个选项为“我不知道”，设置此选项的原因为避免学习者通过猜测而选择了正确的答案，具体见附录 2。题目主要考察心脏的结构和功能的基础知识，主要检测被试是否了解心脏的结构和功能。选择题和填空题每题答对得 2 分，答错或漏答不得分，得分越高，表明掌握得心脏的结构和功能知识越多，相反得分越低，则说明被试对血液循环相关知识越欠缺。该测验信度为 0.62。

（3）学习效果测验

测试问卷经过两名课程教师围绕教学视频中所有知识点进行认真地选题设计，过程严谨，科学客观。测试题目共包括 8 道题目，其中包括 4 道选择题和 3 道填空题（共 23 空）。选择题为单选题，共有 5 个选项，为了避免被试因猜测答案而答对问题，第五个选项设置为“我不知道”，每个单选题答对计 2 分，答错或漏答计 0 分，总分为 10 分。填空题每空分值为 2 分，答对计 2 分，答错或漏答计 0 分，总分为 56 分。因此，该测试的总分为 60 分。得分越高说明对学习内容的学习效果越好，相反则说明对学习内容掌握较差。学习效果测试信度分别为 0.66。

3. 研究结果

为了探讨使用绘图策略过程中，教学视频中是否提供插图及提供插图的具象性（抽象插图 vs.具象插图）对学习者注意力、即时后测和绘图测试的影响。所有变量的描述性统计分析

如下表 2 所示。其中，前测结果显示三组被试在先前知识经验水平上无显著差异（F（2，89）=0.97，p=.381，η=0.02）。

表 2 描述性统计分析表（n = 92）

	控制组（n = 30）		抽象插图组（n = 31）		具象插图组（n = 31）	
因变量	M	SD	M	SD	M	SD
前测成绩	5.93	3.04	5.65	4.17	6.90	3.82
文本-注视时间%	0.93	0.11	0.33	0.12	0.35	0.12
图片-注视时间%	/	/	0.62	0.14	0.58	0.14
图片→文本次数	/	/	28.59	9.31	22.52	8.24
绘图质量	24.63	10.64	31.62	14.85	23.67	11.20
知识测试	40.23	7.46	41.97	8.34	45.06	6.72
绘图测试	20.40	10.15	28.37	14.81	22.50	12.16

2.1. 注意力分配

为验证假设 1，为探究在绘图策略使用过程中，视频材料中是否插图及插图的具象程度对学习者的外部注意有何影响，本实验使用了单因素方差分别以对文字的注视时间比例和对图片注视时间比例进行分析。在对文字的注视时间比例上，三组被试在具有显著的差异（F（2，89）=253.47，p<.001，η=0.85）。事后分析发现，观看无插图组的学习者在文字上的注意停留时间显著高于观看抽象插图组和具象插图的学习者（MD = 0.60，p < .001；MD = 0.58，p < .001），但观看抽象插图组的学习者与观看无插图组的学习者在绘图质量上无显著差别（MD =0.02，p = .442）。在对图片的注视时间比例上，观看含有插图（抽象 vs.具象）组的学习者无显著差异（F（1，60）=0.97，p=.238，η=0.02）。在图片和文字的转换次数上，结果发现，相比于具象的图片，抽象的图片会导致学习者更多的从图片转移到文字（F（1，60）= 7.17，p=.010，η=0.110）。

2.2. 绘图质量

为探究在绘图策略使用过程中，视频材料中是否插图及插图的具象程度对学习者的绘图质量有何影响，本实验使用了单因素方差分析，实验结果表明三组被试在绘图质量表现上具有显著的差异（F（2，89）=3.69，p=.029，η=0.08;见图 2a）。事后分析发现，观看抽象插图组的学习者在绘图质量的表现上显著好于观看无插图组和具象插图的学习者（MD = 4.8，p = .032；MD = 7.95，p = .015），但观看具象插图组的学习者与观看无插图组的学习者在绘图质量上无显著差别（MD =0.97，p = .763）。

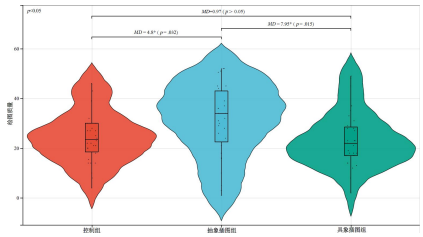


图 2a 绘图质量

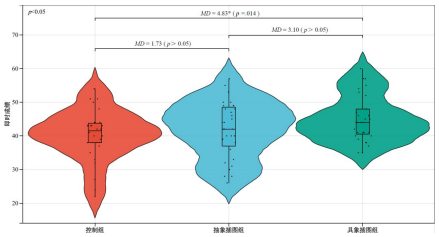


图 2b 即时成绩

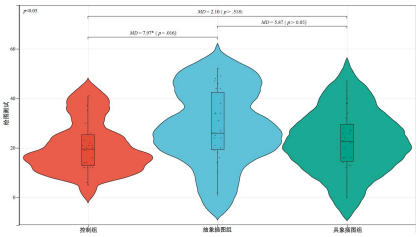


图 2c 绘图测试

2.3. 即时成绩

为验证假设 2，为探究在绘图策略使用过程中，视频材料中是否插图及插图的具象程度对即时成绩有何影响，本实验使用了单因素方差分析，实验结果表明三组被试在即时成绩表

现上具有显著的差异（F（2，89）=3.23，p=.044，η=0.07；见图 2b）。事后分析发现，观看具象插图组的学习者在即时测试的表现上显著好于观看无插图组的学习者（MD = 4.83，p = .014），但观看具象插图组的学习者与观看抽象插图组的学习者在即时测试上无显著差别（MD = 3.10，p = .109），观看抽象插图组的学习者与观看无插图组的学习者在即时测试上无显著差别（MD = 1.73，p = .371）。

2.4. 绘图测试

为验证假设 3，为探究在绘图策略使用过程中，视频材料中是否插图及插图的具象程度对学习者的绘图质量有何影响，本实验使用了单因素方差分析，实验结果表明三组被试在绘图质量表现上具有显著的差异（F（2，89）=3.26，p = .043，η=0.07；见图 2c）。事后分析发现，观看抽象插图组的学习者在绘图质量的表现上显著好于观看无插图组的学习者（MD = 7.97，p = .016），但观看具象插图组的学习者与观看无插图组的学习者在绘图质量上无显著差别（MD =2.10，p = .518），具象插图与抽象插图无显著区别（MD =5.87，p = .073）。

4. 研究讨论

本研究旨在探讨在科学学习中，提供不同类型的学习材料（无插图 vs.抽象插图 vs.具象插图）对学习者绘图策略使用过程中的影响。为此，本研究对学习者在观看带有不同类型插图的视频材料时的注意力分配、绘图质量、即时测试进行了详细的分析。总体而言，研究结果显示了插图类型对学习者的注意力分配、绘图质量和学习成效有显著影响。具体来说，观看无插图视频的学习者在文字上的注视时间比例显著高于其他两组，而观看含有抽象插图的学习者在绘图质量上表现更好。在成绩测试方面，观看具象插图的学习者表现优于无插图组，但与抽象插图组之间差异不显著。

在注意力分配方面，本研究的结果表明，在使用绘图策略的过程中，不同类型的插图对学习者的外部注意力有显著影响。具体来说，当视频材料中没有插图时，学习者更多地将注意力集中在文字上。相比之下，当材料中包含插图（无论是抽象还是具象）时，学习者在文字和图片上的注意力分配较为均衡。这种差异可能由几个因素导致。首先，无插图的情况可能迫使学习者更依赖文字信息来理解内容，因为他们没有其他视觉辅助来帮助解释或补充文本信息。这可能导致他们在文本上投入更多的时间和精力，以确保理解材料的每个细节。另一方面，当提供插图时，尤其是抽象图像，这可能促使学习者在视觉信息和文本信息之间进行更多的认知转换。这种转换可能是尝试将图像与文本内容联系起来的过程的一部分，从而形成对材料的更全面理解。抽象图像可能激发了学习者的内在认知过程，促使他们在图像与文本之间建立更深层次的联系。

在绘图质量方面，本研究的结果显示了视频材料中插图的具象程度对学习者绘图活动的显著影响。特别是，学习者在观看含有抽象插图的视频时，其绘图质量显著优于观看无插图或具象插图视频的学习者。这种结果可能是由于几个原因造成的。首先，抽象插图可能促使学习者进行更深入的思考和理解。由于抽象图像通常缺乏具体细节，学习者可能需要进行更多的内在认知加工，以解释和理解这些图像（Kaminski & Sloutsky, 2013）。这种加工可能包括将图像与他们已知的信息相结合，或者尝试在脑海中构建图像的更全面概念。这种深度加工可能导致更高质量的绘图输出，因为学习者不仅仅是在复制图像，而是在努力理解和内化所学内容。其次，抽象图像的使用可能减少了对具体视觉细节的依赖（Butcher, 2006），从而激发了学习者的创造性思维。学习者可能会被激励去创造性地表达他们对材料的理解，而不是仅仅模仿图像。这种创造性表达可能导致更高的绘图质量，因为它反映了学习者对概念的深入理解和个人化解释。最后，观看抽象插图的学习者可能更专注于理解图像背后的概念，

而不是被具体的细节所分散注意力。这种专注可能促进了更有效的信息整合和表达，从而提高了绘图的整体质量。

在测试成绩方面，本研究的结果揭示了视频材料中不同类型插图对学习者即时测试成绩的影响。研究发现，观看具象插图视频的学习者在即时测试中的表现优于观看无插图视频的学习者，而抽象插图和无插图之间的成绩差异并不显著。这一结果可能由以下几个因素造成：首先，具象插图由于其直观性和具体性，可能更易于学习者快速理解和记忆。具象图像通常提供了清晰的视觉表现，使得学习者能够迅速抓住关键信息，这对于短期记忆和即时测试表现尤为重要。相比之下，抽象图像可能需要更深层次的思考和内在认知加工，这在即时测试中可能不如具象图像直接有效。其次，具象插图可能促进了更好的记忆保持，因为它们提供了明确的视觉参照，这有助于学习者在测试中回忆信息。具象图像的具体细节可能作为记忆线索，帮助学习者更容易地在测试中回忆起学习内容。最后，虽然抽象图像可能促进深层次的信息加工，但在短时间内完成的即时测试可能无法充分体现这种深层次加工的优势。即时测试通常侧重于评估学习者对材料的直接记忆和理解，而不是他们对材料的深入分析或长期记忆。

研究存在如下局限，本研究使用的群体为大学生，以往有研究显示，相比于小学生，成年能够更好地从抽象插图中获益（Menendez et al., 2020, 2022）,因此本研究结果是否可以推广到小学生群体仍有待后续研究推进。此外，本研究侧重于即时测试成绩，而没有考虑长期记忆的影响。因此，研究可能没有完全揭示不同类型插图对学习者长期记忆和知识保持的影响。未来的研究可以包括随后的测试，以评估长期记忆效果。

本研究旨在探索不同类型插图（抽象与具象）对科学学习中绘图策略使用效果的影响，特别关注了注意力分配、绘图质量以及测试成绩。研究结果表明，插图类型对学习者的注意力分配有显著影响，尤其是在无插图情况下学习者更倾向于关注文字信息。在绘图质量方面，抽象插图似乎促进了更高质量的绘图，可能因为它们激发了更深层次的认知加工。而在即时测试成绩方面，具象插图的直观性和具体性似乎更有助于学习者的短期记忆和理解。这些发现强调了在教学设计中考虑视觉辅助类型的重要性。抽象图像可能更适用于促进深度理解和创造性思维，而具象图像则可能更适合于快速信息获取和即时记忆。然而，研究也存在局限性，包括样本代表性、研究设计、对长期记忆的考虑不足、插图类型定义的主观性，以及未直接测量认知负荷等问题。综合来看，这项研究为如何有效地利用视觉辅助工具提供了有价值的见解，并为未来在不同教学环境和学习者群体中应用绘图策略提供了指导。未来的研究应考虑这些局限性，并在更广泛的背景下探索不同类型插图对学习效果的影响。

参考文献

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science, 333*(6046), 1096-1097.

Alesandrini, K. L. (1981). Pictorial–verbal and analytic–holistic learning strategies in science learning. *Journal of Educational Psychology, 73*(3), 358–368.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*(1), 182–197.

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. Journal of Educational Psychology, 112(5), 895–906.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. Journal of the Learning Sciences, 69-110.

Hellenbrand, J., Mayer, R. E., Opfermann, M., Schmeck, A., & Leutner, D. (2019). How generative drawing affects the learning process: An eye-tracking analysis. *Applied Cognitive Psychology, 33*(6), 1147-1164.

Jaeger, A. J., & Fiorella, L. (2023). Metacognitive effects of instructional visuals: the role of cue use and judgment type. *Metacognition and Learning,* 1-43.

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology, 105*(2), 351–363.

Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction, 22*(1), 16-26.

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior, 25*(2), 284-289.

McNeil, N. M., & Fyfe, E. R. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction, 22*(6), 440-448.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2020). Do details bug you? Effects of perceptual richness in learning about biological change. *Applied Cognitive Psychology, 34*(5), 1101-1117.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2022). Detailed bugs or bugging details? The influence of perceptual richness across elementary school years. *Journal of Experimental Child Psychology, 213*, 105269.

Ploetzner, R., & Fillisch, B. (2017). Not the silver bullet: Learner-generated drawings make it difficult to understand broader spatiotemporal structures in complex animations. *Learning and Instruction, 47*, 13-24.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216-237.

Siler, S. A., & Willows, K. J. (2014). Individual differences in the effect of relevant concreteness on learning and transfer of a mathematical concept. *Learning and Instruction, 33*, 170-181.

Skulmowski, A. (2022). Is there an optimum of realism in computer-generated instructional visualizations?. *Education and Information Technologies, 27*(7), 10309-10326.

Skulmowski, A. (2022). Realistic visualizations can aid transfer performance: Do distinctive shapes and descriptive labels contribute towards learning?. *Journal of Computer Assisted Learning, 38*(3), 681-691.

Skulmowski, A. (2023). Realistic details impact learners independently of split-attention effects. *Cognitive Processing, 24*(2), 187-198.

Wang, C., Fang, T., & Gu, Y. (2020). Learning performance and behavioral patterns of online collaborative learning: Impact of cognitive load and affordances of different multimedia. *Computers & Education, 143*, 103683.

基于 AI 绘画的写作教学策略对学生写作成绩、写作想象力以及认知网络结构影响的实证研究

Research on the Effects of AI Painting-Based Writing Teaching Strategy on Students' Writing Achievement, Writing Imagination and Epistemic Network Structure

郑益奇*，杨刚，赵雨
温州大学教育学院
* 2328670001@qq.com

【摘要】 传统写作教学中存在学生想象力匮乏的问题，利用 AI 绘画所生成的创意图片，能够激发学生的想象力与写作表达。因此，本研究提出了一种基于 AI 绘画的写作教学策略，探讨该策略对学生写作成绩、写作想象力以及认知网络结构带来的影响。为测试策略的有效性，在小学写作课堂上进行了一项准实验研究。实验组 (N=38) 采用基于 AI 绘画技术的写作教学策略，对照组 (N=38) 采用传统讲授式写作教学策略。研究结果显示，与对照组相比，实验组在写作成绩、写作想象力方面表现出更大的改善，且认知网络结构更复杂。本研究结果表明，将 AI 绘画纳入中小学写作教学具有积极意义。

【关键词】 AI 绘画；写作教学；想象力；线索

Abstract: *Traditional writing teaching has the problems of students' lack of imagination, while the creative pictures generated by AI painting can stimulate students' imagination and writing expression. Therefore, this study proposes an AI painting based writing teaching strategy, focusing on the effects of this strategy on students' writing performance, writing imagination, and epistemic network structure. To test the effectiveness of the strategy, a quasi-experiment was conducted in an elementary school writing classroom. The experimental group (N=38) used the proposed strategy, and the control group (N=38) used the traditional lecture based writing teaching strategy. The results showed that the experimental group demonstrated greater improvements in writing achievement, writing imagination, and had a more complex epistemic network structure compared to the control group.*

Keywords: AI painting, writing teaching, imagination, clues

1. 前言

写作能力是 21 世纪对学生技能要求的关键素养之一 (Alismail & McGuire, 2015)。掌握写作能力是发展学生的高阶思维能力、培养学生的认知、行为、情感等综合素养的必要基础（杨启光和杨冰，2019）。然而，学生在写作学习中往往存在着许多问题。一是对于大多数学生来说，学习写作是一个具有挑战性和乏味的过程，需要非常复杂的解决问题的行为，比如记忆、计划、文本生成和修订 (Hayes, 2012)。二是传统的写作课堂缺少真情实感，缺乏想象力、创造力，从而使得学生写出的作文内容空洞。三是学生在写作学习过程中通常很少互动，从而难以解决他们所遇到的困难和问题 (Williams, 2018)。因此，如何提升学生在写作学习中的参与感，激发学生的想象力、促进学生抽象想法的具体表达是解决问题的重要突破口。

经过对文献的分析与深入讨论后，本研究深受启发：一是在写作教学过程中，可以将智能技术视为一种重要的教学要素。以智能技术为中介的学习，不仅可以调节学习者与写作对象、教师和同伴学习者之间的关系，而且也可以成为教学理念、模式和管理的干预力量，改变传统以讲授为主的教学范式 (Alexander et al., 2012)。二是高质量写作离不开学生丰富的想象力。Kolovou, Shen, & Smith (2021) 的研究强调写作是培养学生想象力的一种重要方式。此外，借助想象力，可以将学生的各种记忆表象和经验进行重新组合，进而针对写作主题创造出新的思路 (Ruusunen, Hallikainen, & Laukkanen, 2023)。因此，本研究将智能技术视为触发学生写作想象力的动力“引擎”。

基于上述的思考与启发，本研究拟提出一种基于 AI 绘画的写作教学策略。该策略的主要价值体现在两个方面：首先，借助 AI 绘画平台生成的图画，刺激学生的想象力，为他们解决写作上的问题提供思考线索，最终升华他们对写作主题的内在认知。其次，本研究希望改变传统讲授式写作教学模式。教师利用 AI 绘画技术整合到写作教学中，实现写作教学模式的创新，进而提高学生的写作质量。为了深入论证基于 AI 绘画的写作教学策略的有效性，本研究将开展一项准实验研究。本研究主要围绕以下几个问题进行探索。

RQ1: 与传统讲授式写作教学策略相比，基于 AI 绘画的写作教学策略是否更能提升学生的写作成绩？

RQ2: 与传统讲授式写作教学策略相比，基于 AI 绘画的写作教学策略是否更能提升学生的写作想象力？

RQ3: 与传统讲授式写作教学策略相比，基于 AI 绘画的写作教学策略是否更能优化学生的认知网络结构？

2. 关键概念

2.1. AI 绘画

基于人工智能的绘画是一项新兴技术，通过语义分析并模拟人类的绘画过程，能够在短时间内生成大量的图像和作品 (Min et al., 2022)。一方面，该技术使用庞大的文本和图像数据库，来训练机器学习模型，使其不断成熟与优化，进而能够基于用户给出的文本输入来生成图像 (Reed et al., 2016)。另一方面，AI 绘画融合了广泛而丰富的绘画技术和风格，并且利用机器学习逐次迭代升级，进而能够提高用户的绘画能力。因此，借助于 AI 绘画，无论使用者是否具有绘画基础与技巧，都可以轻松创作出令人印象深刻的作品。

AI 绘画在艺术创作中展现出了巨大的潜力，它依据强大的算法模型与动态性生成的图像的功能，可以创造出更加符合认知期望的艺术作品，从而在激发个体的情感表现、想象力发挥与创意创作具有重要的推动作用。如 Chen, Lin, & Chien (2022) 将具有 AI 辅助学习的数字艺术能力训练系统应用于 K-12 教育中的艺术学习活动，发现可以帮助提高学生的绘画表现与艺术学习活动的有效性，同时提升了学生的想象力和学生作品的质量。

因此，本研究拟定在中国小学课堂的写作学习活动中使用 AI 绘画平台，通过 AI 生成与写作主题相关的图像，促进学生对写作主题的深入理解，激发学生的联想与想象，进而提升学生的写作成绩。同时，也可以考察 AI 技术整合写作教学的有效性。

2.2. 写作中的想象力

想象力被定义为超越当前认识模式设想世界的认知能力，即人类通过心理意象想象事物和事件的能力 (Castoriadis & Curtis, 1997)。He & Shum (2005) 认为想象力对人的创造力发展有重要作用，它是产生想法的认知工具，是学生思维发散与提升的基础，也是写作和一般学习的关键要素。因此，激活学生的写作想象力对于写作任务的完成尤为重要（刘淼和曾洁，2011）。课堂是进行写作的重要场所，那么富有想象力的创造性教学课堂是激活学生想象力、解放写

作生命活力的关键（李文军，2014）。在课堂上，教师可以设计多感官的刺激来促进学生想象力的发展，从而补偿传统写作任务中的沉闷，使学生进行自由联想来生成更多的想法。比如，Cheong et al. (2022) 以连续性写作任务为背景，设计了一个探究想象力和写作表现之间关系的框架，证明了想象力在第二语言写作中具有促进作用。

因此，想象力的培养在学校教育和学生的未来发展中起着至关重要的作用 (Wu & Albanese, 2013)。如何在写作课堂中，充分发挥学生的想象力并培养学生的想象力是本研究关注的重点。故而，本研究将 AI 绘画整合到学生的写作学习过程之中，助力学生想象力的发挥，同时也改善他们的写作学习质量。

3. 基于 AI 绘画的写作教学策略设计

3.1. 基于 AI 绘画的写作教学策略的理论构建

本研究应用了一款名为“文心一格”的软件作为 AI 绘画平台。借助 AI 绘画平台，研究团队对写作教学设计进行优化，并提出了基于 AI 绘画的写作教学策略。这一策略主要将 AI 绘画平台应用到写作主题讨论环节之中，其目的是为学生提供视觉化学习支架。一方面帮助学生在讨论过程中，当遇到迷惑之处，他们能够透过生成的图片展开想象。具体见下图 1。另一方面以 AI 绘画生成图片为中介，为学生讨论营造了一种积极的“头脑风暴”氛围，帮助他们对写作主题进行深层次理解；同时也提升了学生合作学习的能力。



图 1 基于 AI 绘画的写作教学策略的理论框架

根据绘画平台的功能与特征，首先学生可以根据写作主题要求，输入写作线索词。学生通过输入不同的线索词可以生成不同类型的图片。我们将这些线索词从输入内容的复杂度分为三类，分别是单点式线索词，关联式线索词和拓展式线索词。线索词的复杂程度与生成图片的精细度相关，输入的线索词内容越具体，那么图片中所展现的细节越丰富。具体见下图 2。单点式线索词是学生根据写作中的问题，以单词的形式作为线索词输入一种模式来生成图片。关联式线索词，是学生在单词的基础上增加修饰词，比如形容词，从而以短语的形式作为线索词输入一种模式来生成图片，可以使原本常见的描述对象在细节呈现上变得更加丰富。而拓展式线索词，是学生将单词与短语进行组合形成了一个描述句子，以句子的形式作为线索词输入一种模式来生成图片，可以呈现具体情境从而激发学生展开想象。

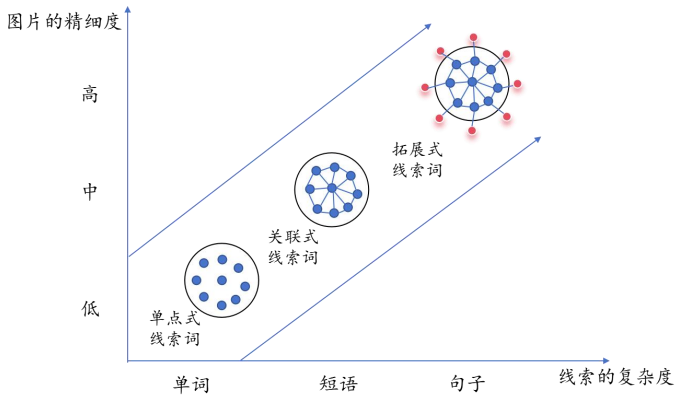


图 2 三种不同类型的线索词

3.2. 基于AI绘画的教学活动设计

根据上述的教学策略的设计，本研究将基于 AI 绘画的写作学习活动划分为五个阶段：（1）理解写作主题，奠定写作基础。（2）构思写作思路，搭建写作框架。（3）讨论写作主题，AI 绘画辅助认知。（4）撰写写作内容，接受写作指导。（5）评价写作质量，分享写作体验。具体见图 3。另外，需要说明的是，对照组在使用传统讲授式写作教学策略开展写作活动时，在第三阶段与实验组存在着差异，其他阶段基本一致。在这一阶段，对照组教师引导学生围绕写作主题展开讨论，积极分享个人观点，使用思维导图等工具来整理写作思路。

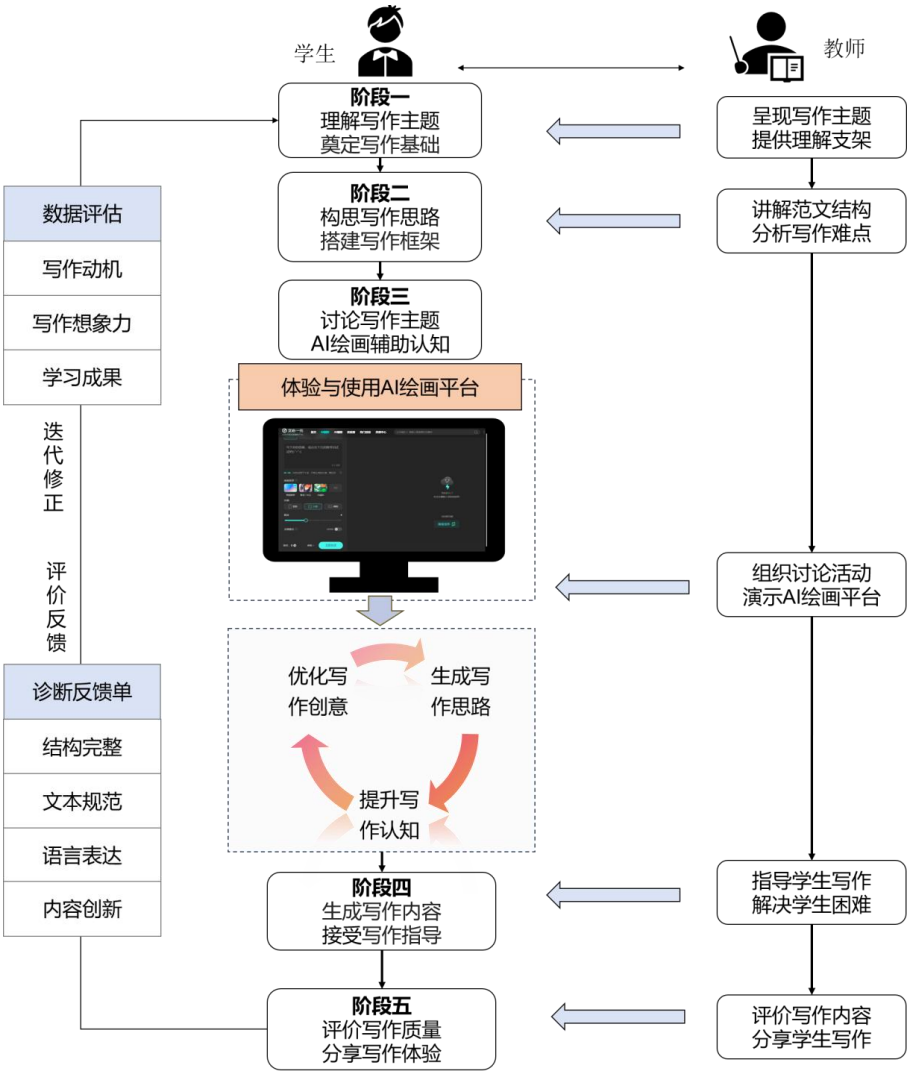


图 3 基于五个阶段的写作活动过程

4. 实验设计

4.1. 实验对象

本次实验的研究对象为某小学五年级的两个班级的学生。实验组采用基于 AI 绘画的写作教学策略，对照组采用传统讲授式写作教学策略。在写作学习过程中，本研究采用统一授课的方式，即由同一位老师讲授写作内容。

4.2. 实验过程

实验流程如图 4 所示。第一周是前测阶段，学生完成写作成绩的前测以及写作想象力和学习动机的问卷；第二周至第六周是实验活动阶段，实验组与对照组学生围绕着课内外一致的的写作主题，使用不同的教学策略开展写作活动，完成写作；第七周是后测阶段，实验组和对照组学生需进行写作成绩的后测，并完成写作想象力、学习动机的问卷；第八周是数据整理与分析阶段，同时在每个组随机抽取 10 名学生进行学习满意度的深度访谈。

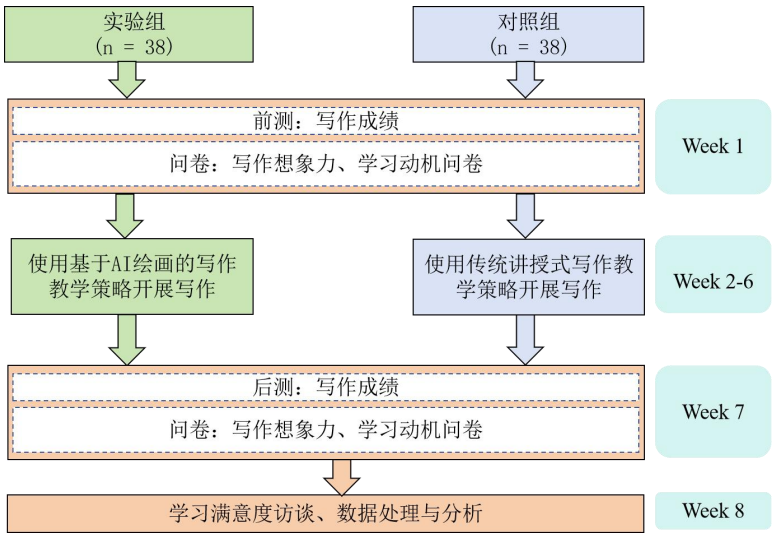


图 4 实验流程

4.3. 实验工具

本研究使用的测量工具包括作文评价量表、写作想象力问卷以及认知网络分析编码表。

4.3.1 作文量表

其中作文评价量表是在 Cheung et al.(2001) 提出的作文评价量表基础上开发的。经过实践应用后，形成了包括结构完整、文本规范、语言表达、内容创新的四维作文评价量表 (Yang et al., 2021)。

4.3.2 写作想象力问卷

写作想象力问卷改编自 Liang et al.(2012) 的想象力量表，一共由 9 个题项构成，包括创造性想象力与再现性想象力两个维度。问卷采用李克特 5 点量表，克朗巴赫系数α为 0.76 (Cronbach's alpha=0.76)，信度较高。

4.3.3 认知网络分析编码表

认知网络分析 (ENA) 是一种以个人认知框架为基础，以事实证据为中心的教育评价模式。主要基于对学习过程中产生的文本数据等进行分析，按照时间或是活动顺序划分为多个“节点”元素，并依据它们之间的关联进行网络建模，从而实现数据的可视化分析。因为 ENA 将复杂的认知结构通过网络图可视化表征，超越了传统基于频次的测量方法，被广泛用于分析和评价学生的学习过程。因此，本研究根据小学阶段学生写作特点，采用 Charmaz (2006) 的扎根理论方法对学生写作过程中生成的描述性好词进行编码，经两位教授对编码表进行审

查并确认，开发了一种 ENA 编码表。该编码表分为三个维度，包括心理、情节与环境。具体见表 1。

表 1 ENA 编码表

维度	编码	子维度	描述
心理 Psychology	PEM	情感 Emotion	描述对象的情绪体验，如喜怒哀乐等。注重刻画细腻，真挚感人的心理情感活动。
	PAS	联想 Association	在描述一个对象时，通过借助对其他事物的联想而产生的想象，如，我想起了、我灵光一现.....
情节 Episodes	ELI	直线性 Linear	在写作中，采用了直线式情节描述。主要体现衔接词的应用上。如，然后、接下来.....
	ELO	逻辑性 Logical	主要体现在写作情节上有强烈的故事感，描述的事情有新颖。主要体现以转折连接词作为段落的连接，如，突然，然而，但是.....
环境 Environment	EST	静态 Static	主要指对环境，景物进行客观、真实地描写，目的是静态描述事物景貌。体现在以形容词作为描述词，如，宽阔的小河.....
	EDY	动态 Dynamic	主要对环境、景物进行动态的描述。其目的是呈现环境具有动态、动作、行为等特征。主要体现在动词与形容词的应用上，如，水流湍急的小河，河面激起了水花。

5. 实验结果分析

5.1. 写作成绩分析

首先，本研究应用了描述性统计来探究两组学生的写作成绩，其结果如表 2 所示。由表可知，实验组与对照组在前测与后测的平均成绩均存在差异。如两组学生的后测写作成绩都高于前测，且实验组后测成绩高于对照组。为进一步检查这些差异的显著性及效应量，研究采用了单因素协方差分析 (ANCOVA) 比较两组学生实验后的写作成绩，排除前测成绩的影响。在此之前，为了证实 ANCOVA 模型的假设，我们进行了方差的同质性检验。Levene 检验表明实验组和对照组之间存在可比方差 (p =0.86 > .05) 和有效的一致数据变异性。因此，应用单因素 ANCOVA 分析实验组和对照组之间写作成绩的差异是合理的。

表 2 两组写作成绩的描述性统计

维度	组别	样本量	前/后测	均值	标准偏差	95%置信区间	
						上限	下限
写作成绩	实验组	38	前测	80.42	0.86	78.69	82.15
		38	后测	89.05	0.47	88.10	90.01
	对照组	38	前测	80.65	0.74	79.16	82.16
		38	后测	86.84	0.51	85.82	87.87

表 3 显示了单因素 ANCOVA 的结果。由表中数据可知，在排除前测的影响后，实验组和对照组学生在写作成绩 ($F_{(1,34)}$=14.17, p<0.05) 存在显著差异。效应大小即部分 eat 平方为 0.16，大于 0.14 被视为大效应量，因此两组写作成绩之间的差异是比较大的。同时进一步分析调整后的平均写作成绩发现，实验组学生的写作成绩 (M_{adj}=89.09, SD=0.429) 明显优于对照组 (M_{adj}=86.81, SD=0.429)。这一结果证明，相比于传统讲授式写作教学策略，基于 AI 绘画的写作教学策略能提升学生的写作成绩。

表 3 两组写作成绩的 ANCOVA 结果

来源	第III类平方和	*df*	均方	*F*	显著性	偏 Eta 方
组别	98.88	1	98.88	14.17	<.001***	.16
误差	509.24	73	6.98			

*** $p < .01$

5.2. 写作想象力分析

首先研究采用独立样本 t 检验，以验证两组学生在实验前后写作想象力是否有显著变化。数据结果如表 3 所示，发现实验组和对照组的整体想象力不存在显著差异 (t=1.358, p=0.180>0.05)。具体分维度来看，在创造性想象力方面，两组不存在显著差异 (t=0.437, p=0.663>0.05)；在再现性想象力方面，实验组与对照组存在显著差异 (t=2.023, p=0.048<0.05)。

为探讨基于 AI 绘画的写作教学策略对学生写作想象力的影响，需要排除对应的实验前测可能带来的干扰，因此研究针对实验组和对照组学生的写作想象力各维度进行协方差分析。Levene 的方差同质性检验表明实验组和对照组之间存在可比方差 ($p > .05$) ，两组学生想象力的 ANCOVA 结果如表 4 所示。在控制协变量后，两组在想象力总分($F_{(1,73)}$=7.98, p<0.05)、再现性想象力($F_{(1,73)}$=7.98, p<0.05) 上存在显著差异。这一结果证实了基于 AI 绘画的写作教学策略能提升学生的想象力整体水平与再现性想象力。

表 4 两组写作想象力的 ANCOVA 结果

维度	来源	第III类平方和	*df*	均方	*F*	显著性	偏 Eta 方
想象力总分	组别	158.69	1	158.69	8.07	0.006**	.099
	误差	1436.33	73	19.68			
创造性想象力	组别	35.23	1	35.23	2.830	0.097	.037
	误差	908.80	73	12.45			
再现性想象力	组别	45.88	1	45.77	7.980	0.006**	.099
	误差	419.74	73	5.75			

** $p < .01$

5.3. 认知网络结构分析

为进一步探究 AI 绘画对学生写作中认知能力造成的差异，本研究对实验组与对照组学生在写作过程中所使用的词汇进行认知网络分析。两组学生的写作词汇认知网络如图 5 所示，其中(a)为对照组，(b)为实验组，(c)为两组认知网络的叠加图。本研究将进行详细分析如下。

(a)中 PEM-EDY、PEM-ELO 的连接权重皆为 0.22，EDY-ELO、ELO-EST、EDY-EST、ELO-ELI 的连接权重皆为 0.15，这表明这些元素之间存在着连接关系。PEM-ELI、ELI-EST、PEM-EST 的连接权重皆为 0.43，这表明这些元素之间有相对较强的连接。值得注意的是，PAS 与任何元素都没有产生连接。由此可见，应用传统讲授式写作教学策略的学生能够从静态环境、直线情节、情感状态三方面写作，却忽视了对动态环境、逻辑情节的描写，同时他们很难产生积极的联想状态。

(b)中 PAS-EDY、PAS-ELO、PAS-ELI、EDY-EST、ELO-EST、PEM-EST 的连接权重皆为 0.15，这表明这些元素之间存在着连接关系；PAS-EST、ELI-EST 的连接权重皆为 0.05，表明这些元素之间有相对较弱的连接；EDY-PEM、PEM-ELO、ELO-EDY 的连接权重皆为 0.43，表明这些元素之间有相对强烈的连接关系。因此可知，应用基于 AI 绘画的写作教学策略的学生能够从心理状态、逻辑情节、动态环境三个方面着重写作，并能够积极地产生联想状态。

(c)中蓝色线条为对照组，主要强调 PEM、ELI、EST 元素之间的连接；红色线条为实验组，显示了 PAS 与 EDY、ELO、PEM、ELI 元素，以及 EDY、PEM、ELO 元素之间的连接。由此可见，对照组学生倾向于刻画写作对象在静态环境中的心理状态与直线情节的描述，而实验组学生则倾向于在作文中刻画写作对象在动态环境中的联想状态与逻辑情节的描述。

从上述认知网络结构的分析结果可知，相较于对照组，使用基于 AI 绘画的写作教学策略的实验组学生对事物的描写更加动态、全面和细致，并且能够表达自己的情感与心理活动，产生联想与想象，使作文更具丰富的想象力。

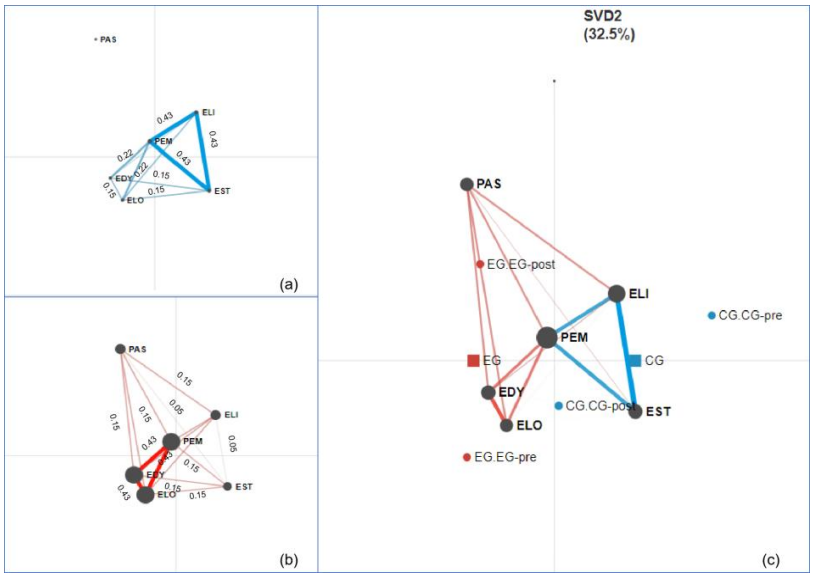


图 5 对照组（蓝色）和实验组（红色）后测写作中的认知网络结构

6. 结论

在这项研究中，我们探索了将 AI 绘画整合到写作学习活动中的潜力，旨在将图像生成式人工智能的艺术创造潜力与写作学习中的想象力需求结合起来。为此，我们设计并实施一项基于 AI 绘画的写作教学策略，让学生在使用 AI 绘画工具生成创意图片，以激发他们的想象力，并为他们解决写作学习中的困惑提供帮助。研究结果表明，基于 AI 绘画的写作教学策略对学生写作成绩、写作想象力、认知网络结构都具有积极性影响。

作为产生写作创意的认知工具，想象力是写作学习的一项关键要素 (Pavlik, 2002)。然而，培养想象力一直被置于中国学校教育体系的边缘。因此想象力的匮乏是导致学生写作表现不尽人意的一个重要原因之一。在制度化的写作教学课堂中，常见的是模仿式写作模式与单一化的思考方式，对于思维活跃的小学生来说，无疑是一种思想的枷锁。当学生使用 AI 生成的图片来解决写作中的问题时，他们就会暂时离开问题的漩涡，摆脱常有的惯性思考。由此，学生进入了新的思考空间，不再受到原有思考的局限与惯性思考的羁绊。这样新的思考避免了惯性思考所形成的写作套路，想象力便得到了唤醒。简言之，借助 AI 绘画工具激发学生的写作想象力，不仅让他们能够运用丰富修辞词汇来描述写作对象，而且也使得学生的认知思维结构的发展得到了促进。本研究中的认知网络结构分析结论就是一个很好的佐证。

当然，要想写出一篇优秀的且有创意的作文，光有想象力还是不够的。很多研究者认为，认知情感 (Vass, 2007)、同伴互动 (Marcos et al., 2020)、元认知技能与写作专业知识 (Barbot et al., 2012) 等因素都是影响写作的重要因素。这也是我们未来研究需要关注的重点。总而言之，这项研究对于技术整合于学科教学的研究具有实际意义，为在不断变化的技术环境中促进高质量学习提供进一步的指导。

参考文献

李文军. (2014). 在写作教学中培养学生的想象力. 中学语文教学 (08), 37-39.

刘淼 & 曾洁. (2011). 言语交际视域中的写作能力及写作教学. 课程·教材·教法 (09), 56-60.

杨启光 & 杨冰. (2019). 基于深度学习的小学语文写作:内涵、特征及教学策略. 教育导刊 (12), 65-69.

Alexander, P., Chabot, K., Cox, M., DeVoss, D. N., Gerber, B., Perryman-Clark, S., ... & Wendt, M. (2012). Teaching with technology: Remediating the teaching philosophy statement. *Computers and Composition, 29*(1),23-38.

Alismail, H. A., & McGuire, P. (2015). 21st century standards and curriculum: Current research and practice. *Journal of Education and Practice, 6*(6), 150-154.

Ammade, S., Mahmud, M., Jabu, B., & Tahmir, S. (2020). TPACK model based instruction in teaching writing: An analysis on TPACK literacy. *International Journal of Language Education, 4*(1), 129-140.

Barbot, B., Tan, M., Randi, J., Santa-Donato, G., & Grigorenko, E. L. (2012). Essential skills for creative writing: Integrating multiple domain-specific perspectives. *Thinking Skills and Creativity, 7*(3), 209-223.

Castoriadis, C., & Curtis, D. A. (1997). *World in fragments: Writings on politics, society, psychoanalysis, and the imagination*. Stanford University Press.

Chen, S. Y., Lin, P. H., & Chien, W. C. (2022). Children's Digital Art Ability Training System Based on AI-Assisted Learning: A Case Study of Drawing Color Perception. *Frontiers in Psychology, 13*, 823078.

Cheong, C. M., Zhang, J., Yao, Y., & Zhu, X. (2022). The role of gender differences in the effect of ideal L2 writing self and imagination on continuation writing task performance. *Thinking Skills and Creativity*, *46*, 101129.

English, A. R. (2019). John Dewey and the role of the teacher in a globalized world: Imagination, empathy, and 'third voice'. In *John Dewey's Democracy and Education in an Era of Globalization* (pp. 70-88). Routledge.

Finke, R. A. (1996). Imagery, creativity, and emergent structure. *Consciousness and cognition, 5*(3),381-393.

Galbraith, D., & Baaijen, V. M. (2018). The work of writing: Raiding the inarticulate. *Educational Psychologist, 53*(4), 238-257.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388.

He, K. K., & Shum, M. S. K. (2005). The study of psychological model and teaching approaches: Primary school composition. *Teaching Writing in Chinese Speaking Areas*, 109-130.

Jang, Y., & Lee, J. (2019). The effects of ideal and ought-to L2 selves on Korean EFL learners' writing strategy use and writing quality. *Reading and Writing, 32*, 1129-1148.

Klein, P. D., & Unsworth, L. (2014). The logogenesis of writing to learn: A systemic functional perspective. *Linguistics and Education, 26*, 1-17.

Kolovou, M., Shen, J., & Smith, B. E. (2021). Looping into hyperspace: Mechanisms of distributed imagination during a collaborative multimodal science fiction writing project. *Thinking Skills and Creativity, 40*, 100819.

Li, A. W. (2023). Using Peerceptiv to support AI-based online writing assessment across the

disciplines. *Assessing Writing, 57*, 100746.

Liang, C., Hsu, Y., Chang, C. C., & Lin, L. J. (2013). In search of an index of imagination for virtual experience designers. *International Journal of Technology and Design Education, 23*, 1037-1046.

Lothane, Z. (2007, September). Imagination as reciprocal process and its role in the psychoanalytic situation. In *International Forum of Psychoanalysis* (Vol. 16, No. 3, pp. 152-163). Taylor & Francis Group.

Marcos, R. I. S., Fernández, V. L., González, M. T. D., & Phillips-Silver, J. (2020). Promoting children's creative thinking through reading and writing in a cooperative learning classroom. *Thinking Skills and Creativity, 36*, 100663.

Min, S., Lyu, X., Holtzman, A., Artetxe, M., Lewis, M., Hajishirzi, H., & Zettlemoyer, L. (2022). Rethinking the role of demonstrations: What makes in-context learning work?. *arXiv preprint arXiv:2202.12837.*

Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Reading and Writing, 25*(3), 641-678.

Pavlik, L. (2002). Structured imagination and the writing of creative stories. *The Journal of Creative Behavior*, *36*(2), 117-137.

Rasteiro, I., & Limpo, T. (2023). Longitudinal and reciprocal links between writing motivation and writing quality in grades 4–5: A cross-lagged panel analysis. *Contemporary Educational Psychology, 75*, 102222.

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016, June). Generative adversarial text to image synthesis. In *International conference on machine learning* (pp. 1060-1069). PMLR.

Ruusunen, N., Hallikainen, H., & Laukkanen, T. (2023). Does imagination compensate for the need for touch in 360-virtual shopping?. I*nternational Journal of Information Management, 70,* 102622.

Sala-Bubaré, A., & Castelló, M. (2023). Exploring writing processes in authentic writing tasks: A multimodal mixed-method approach. *Journal of Second Language Writing, 61*, 101038.

Tao, W., Gao, S., & Yuan, Y. (2023). Boundary crossing: an experimental study of individual perceptions toward AIGC. Frontiers in Psychology, 14, 1185880.

Vass, E. (2007). Exploring processes of collaborative creativity—The role of emotions in children's joint creative writing. *Thinking skills and Creativity, 2*(2), 107-117.

Williams, C. (2018). Learning to write with interactive writing instruction. *The Reading Teacher, 71*(5), 523-532.

Wu, J. J., & Albanese, D. L. (2013). Imagination and creativity: Wellsprings and streams of education–the Taiwan experience. *Educational Psychology, 33*(5), 561-581.

Yang, G., Chen, Y. T., Zheng, X. L., & Hwang, G. J. (2021). From experiencing to expressing: A virtual reality approach to facilitating pupils′ descriptive paper writing performance and learning behavior engagement. *British Journal of Educational Technology, 52*(2), 807-823.

Yeung, P. S., Chung, K. K. H., Chan, D. W. O., & Chan, E. S. W. (2024). Relationships of writing self-efficacy, perceived value of writing, and writing apprehension to writing performance among Chinese children. *Reading and Writing*, 1-47.

以線上知識論壇學習對學生知識翻新環境與社群意識感知之影響

Influences of Knowledge Forum Online Learning on Students' Sense of Knowledge Building Environment and Community.

張芷瑄[1]、薛智暉[1]、蔡以寧[2]、洪煌堯[1*]、洪國財[3]
[1]國立政治大學教育學系
[2]國立東吳大學語言中心
[3]國立臺中科技大學應用日語系(科)暨日本市場暨商務策略碩士班
hyhong@nccu.edu.tw

【摘要】本研究旨在透過知識翻新教學與線上知識論壇平台促進班級學習社群的發展。研究對象為臺灣某國立大學修習「教學設計與科技研究」21 位學生。教學設計以知識翻新原則並輔以線上知識論壇平台，營造學生對話與交流的知識翻新協作環境。資料來源有二：第一、學生於學期初與學期末填寫「知識翻新環境量表」。第二、學生於學期初與學期末填寫「社群意識量表」。研究結果顯示：第一、經歷一學期知識翻新教學與知識論壇的學習環境，學生對於知識翻新環境感知有正向的提升。第二，學生到了學期末的階段，社群意識亦有顯著的提升。第三、學生在期末的知識翻新環境感知與社群意識的表現，有顯著正相關，代表當學生對知識翻新環境感知較高，也會有較高的社群意識表現。

【关键词】 知識論壇；知識翻新環境；社群意識

Abstract: *This study aimed to promote the development of a classroom learning community through knowledge building and an online Knowledge Forum. The participants in this research were 21 students from a national university in Taiwan who were enrolled in the course "Instructional Design and Technology Research." Instructional design was based on the principles of knowledge building and supplemented with an online Knowledge Forum to create a collaborative environment for knowledge building through student dialogues and interactions. Data were collected through two sources: first, students completed the "Knowledge Building Environment Scale" at the beginning and end of the semester, and second, they completed the "Classroom Community Scale" at the same intervals. The results of the study revealed the following: first, after experiencing one semester of knowledge building with Knowledge Forum online environment, students showed a positive improvement in their perception of the knowledge building environment. Second, by the end of the semester, students' sense of community also increased significantly. Third, there was a significant positive correlation between students' perception of the knowledge building environment and their sense of community at the end of the semester, indicating that when students had a higher perception of the knowledge building environment, they also had higher sense of community.*

Keywords: Knowledge Forum, knowledge building environment, sense of community

1. 前言

二十一世紀科技革新應用與跨域知識統整，增進了人類社會的福祉，也形成巨大而廣泛的變遷。美國的「21 世紀關鍵能力聯盟」（Partnership for 21st Century Skills，簡稱 P21），持續推動教育體系變革，以培養學生具備在 21 世紀成功所需的關鍵技能，應對快速變化的世界。P21 以批判性思維（critical thinking）、溝通能力（communication skills）、團隊協作

（collaboration）、創造與創新（creativity and innovation）為新世紀核心能力的學習技能框架，而學習者的想法是推動 4Cs 能力的關鍵核心。此外，國際學生能力評量計畫（Programme for International Student Assessment, PISA）自 2022 年起加入創意思考能力評量，期許各國教育可以幫助學生培養多元想法、創意想法、與改進想法的能力（OECD, 2019），此也為因應創新和創造力做為 21 世紀的關鍵能力養成與教育核心的重要性。於 2023 年誕生的 ChatGPT，以 AI 科技建構出龐大而驚人的知識量，對教育各界帶來無與倫比的衝擊。當人類以漫長時間記憶背誦而來的標準答案，不停面臨快速重整的挑戰時，培養人類創新想法與反思能力，也成為未來教學更不可忽視的課題。在這樣的世界潮流發展之下，過去傳統的教育方法已經無法符應現代社會快速變遷的需求，教育需要創新的思維與教學方法。

知識翻新理論（knowledge building）被學習科學視為是一種有效的教育創新途徑。知識翻新是指想法持續產出想法、改進想法的社群合作歷程。知識翻新打破傳統教育被動傳遞與單向接受的知識概念，將學生視為持續想法產出與創新的知識工作者；也不同於過去對學習單獨對個體的重視，在知識翻新的歷程中，社群協同合作以創新知識，才是社群的主要目的。知識翻新十二個原則，是教學設計重要的根基，透過原則性的教學，形成知識社群，而社群的合作不僅只是個人在小組之中的學習，更重要的是以小組達成集體學習目標的共同體（Gutiérrez-Braojos, et al. 2019; Scardamalia, & Bereiter, 2020）。值得關注的是，知識翻新翻轉了高等教育長期以來重視專業知識獲取、且將知識視為是個人認知的過程。知識翻新重視學習者對於想法的發想與改進，強調社群共同協作知識以增進福祉的目標，使高教學習者有機會透過社會互動與合作形成學習社群，讓學習從傳統競爭導向轉化為重視「公共利益」（common good）的共好環境，個體值價也於社群互動中昇華（Tan, et al.,2021; Scardamalia, & Bereiter, 2021）。

知識翻新以線上知識論壇（Knowledge Forum），營造無所不在的線上學習環境。知識論壇是一個基於知識構建原則設計的科技環境，以具體支持知識翻新以想法為中心的理念。知識論壇設計的工具功能都在於協助知識翻新社群中的個體與社群，以漸進的方式產生、改進、整合想法。學生可以使用知識論壇紀錄個體與社群在想法上所有的互動歷程由發想、反思與回饋中，呈現出想法如何被改進，以促進知識創新的過程。簡言之，在知識翻新與線上知識論壇科的學習環境中，學生與想法即是學習的主體，想法是無所不在且持續翻修的歷程（Scardamalia, & Bereiter, 2020; Tan, et al., 2021）。此外，知識翻新重視社群互動，學習社群的高度互動有助發展社群意識，對於學習效果與學生參與度具有重大影響（Rovia, 2002）。過去研究指出，社群意識在數位學習環境中，可提升學習者參與度與互動，支持彼此的情感需求進而提高學習成效（Balboni, et al., 2018; Ergün, & Avcı, 2018）。然而，知識翻新與社群意識的相關性研究，仍有發展空間。因此本研究將了解知識翻新歷程對社群意識的影響。此外，本研究以機遇式（opportunistic grouping）的方式，以全體班級為一個社群單位，打破既有「分組」的概念，以不同角度呈現知識翻新共同協作對社群意識影響性。

綜上所述，本研究以線上知識論壇讓學生沈浸在知識翻新的學習環境，促進學生社群意識的提升，目的在培養學生二十一世紀知識創新與社群合作之能力。本研究問題為：學生經歷知識翻新教學與線上知識論壇學習環境後，對知識翻新環境的感知情形為何？學生社群意識的感知情形為何？學生對知識翻新環境感知與社群意識感知是否有相關存在？

2. 研究方法

2.1.研究背景

本研究對象為台灣某國立大學學生共 21 位，修習「教學設計與科技研究」課程。修課學生來自 8 個系所的 16 位大學與 5 位研究生，課程時程為一個學期，共計十八週，課程目標為：學生以解決與自身相關的實際生活問題為前提，在學期初發想一個以解決問題為目標的方案，設計或改進一個教學科技或生活科技產品。在學期間的課程中，授課教師以機遇分組方式，使全班形成一個社群以充分進行討論。教師引導學生在課間與課後於知識論壇寫下自己的想法、提供建議與進行反思，並鼓勵學生隨時透過平台閱讀與紀錄。在學期中，學生報告初步成形的方案，同時彼此給予回饋。在學期末，學生報告完成的設計方案，並製作介紹影片上傳至 YouTube 影音頻道。本課程以全英文授課，修課學生同以英文進行討論與在知識論壇上留言。但受限於部分主題內容與特殊情況，學生會以中文溝通或留言。

2.2.教學設計與線上學習環境

2.2.1. 教學設計

本研究個案課程為「教學設計與科技研究」，教學設計以知識翻新原則與線上知識論壇平台，營造學生對話與交流的知識翻新協作環境。學生在知識翻新環境中，以解決教學科技問題為目標，改良與設計教學工具或設計。教師實際將核心原則融入課程運作時，重視：（1）想法發展原則：在課室中，本研究授課教師會肯定學生想法並鼓勵學生多產生想法。（2）知識自主原則：教師運用「少即是好」（less is more）的方式，不提出任何想法上的干涉，鼓勵學生產生自己的想法。（3）社群互動原則：本學期的每次分組討論，皆透過機遇分組方式，小組成員非固定，而是機遇形成的；每個人都有機會了解他人的想法，以及分享自己的想法，再透過互相回饋與反思，改進想法。

本研究個案課程每週進行一次，每次進行三堂課，共計 150 分鐘。學生每週上課，在課程知識之外，針對個人發想的教學設計方案進行改進。教師以知識翻新原則進行課程設計：教師於第一節課講述與引用課程相關文獻，刺激學生想法的產出；於第二節課與第三節課，學生以機遇分組方式，在教師營造的知識翻新環境中，進行討論。討論流程為：想法產出與想法與精進：學生向小組成員提出自己在設計上的想法，在每週的知識翻新過程中，學生會提出新的或改進的想法；在不斷的討論中，小組成員會彼此回饋與貢獻，成員互相使想法精進。接著，社群回饋與共同設計：成員個人在知識論壇留言，寫下對自己的產品設計的想法、對小組成員的設計提出建議、自己在接收建議後的自我回饋或反思。成員經由不同週次與不同成員的建議，透過知識翻新集體合作過程，共同改進自己與彼此的設計。最後階段為反思回饋：知識論壇不受時間與空間限制，每位學生在小組討論或閱讀其他小組成員的回饋後，可隨時隨地進行想法的回饋與反思，自由在知識論壇留下文字紀錄，想法得以再次發展或精進，知識翻新歷程可再向前推進。上述各個階段的互動，形成一個促進想法以解決問題的歷程，如圖 1 所示。

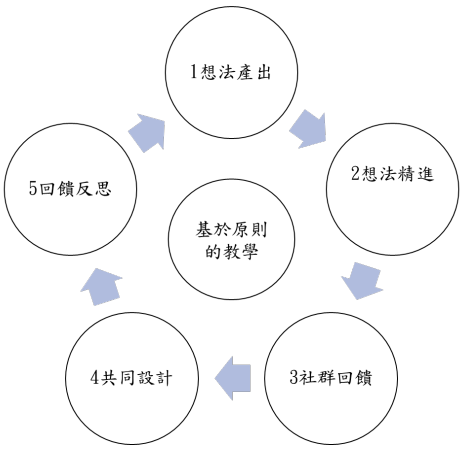


圖 1 想法中心於課程應用

2.2.2. 線上學習環境

知識論壇（knowledge forum 6, KF6）為本課程所使用之網路協作平台。知識論壇運用科技技術與網路環境，使使用者主動記錄與連結想法，想法得以互相建構而創新知識。虛擬的知識翻新社群得以成形，知識創造歷程，得以具體呈現。

本研究個案學生，在每週的機遇分組討論後，會先寫下個人對自己產品設計的想法，並對小組成員的設計提出建議；在課間與課後，再對其他成員對自己的建議，提出自我回饋或反思。知識翻新社群中的個人，透過知識論壇上每次不同的想法記錄，使想法產出與精進；經由社群共同回饋，形成知識翻新的集體合作過程，想法得以再次發展或精進。透過學生發表文、閱讀貼文、回文的行為，形成一個促進想法以解決問題的歷程，達成創新知識。圖 2 至圖 3 為知識論壇與紀錄想法之具體圖例。

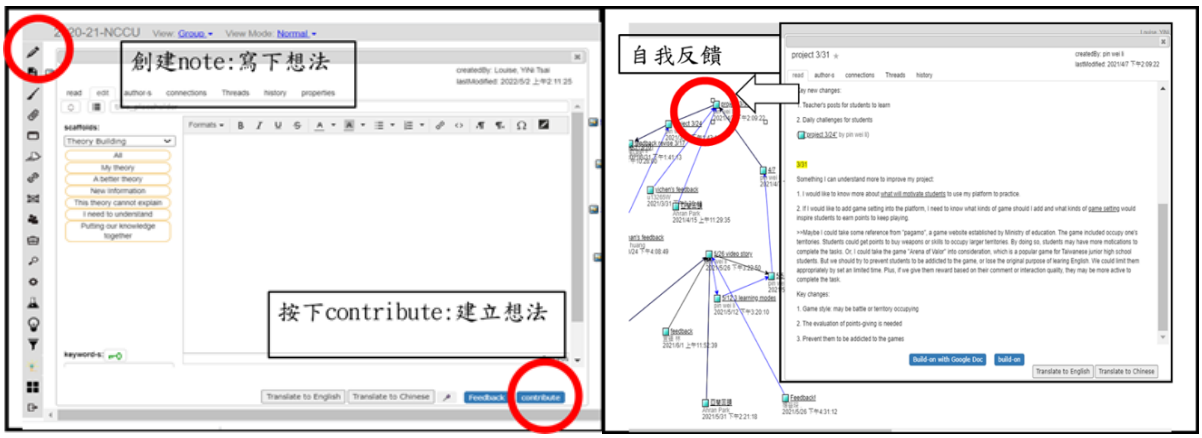


圖 2 在知識論壇貼文進行自我反饋

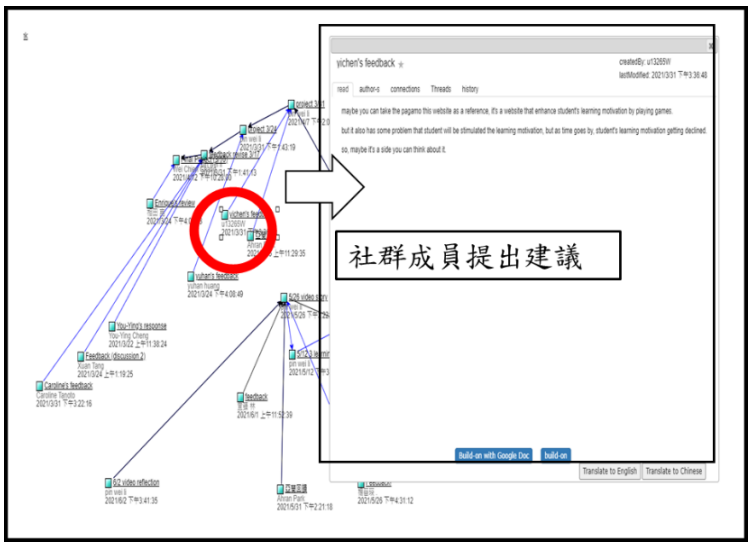


圖 3 知識翻新社群成員在知識論壇提出建議

2.3. 資料蒐集與分析

2.3.1. 知識環境量表

本研究以 Lin 等人（2014）編制的「知識環境量表」（Knowledge-Building Environment scale，簡稱 KBES） 探討研究個案在知識翻新的教學設計中，對知識翻新環境的感知。本量

表以知識翻新原則為基礎，以探索性因素分析將知識翻新環境分為「想法中心」(working with idea）、「自主學習者」（assuming agency）、「社群」（ fostering community）三個部份，透過相關分析，檢測量表具有良好之信、效度與穩定性與預測力（Lin et al., 2014）。

本量表共 24 題，「想法中心」部份有 7 題、「自主學習者」有 8 題、「社群」有 9 題。採四點量表（4-point Likert scale），受試者由勾選非常不同意、不同意、同意、非常同意，得分為 2 分至 5 分。在本研究中，研究者於學期初進行前測，並於學期末進行後測，由學生依自我對知識翻新環境的感知填答。

2.3.2. 社群意識量表

本研究採用 Rovai（2002）發展的「社群意識量表」（Classroom Community Scale）了解研究個案學生在知識翻新環境與發展知識翻新歷程後，成員對社群意識的感知情況。本量表分為二個子構面：連結項目（connectedness items）與學習項目（learning items）。連結項目用以測量社群成員間的情感性連結，例如:成員間相互分享與彼此信任的程度。學習項目用以測量社群成員在課堂中彼此連結的程度或對學習的滿意度等。本量表內部一致性信度（Cronbach's alpha）α 值為 0.93 到 0.96 的範圍內，具有穩定性。

此量表採五點量表（Likert scale），共二十個題目。受試者由勾選非常不同意、不同意、無意見（中立）、同意、非常同意，得分為 1 分至 5 分。本研究問卷調查施測時間為學期初與學期末，於課程中進行，以收集前測與後測資料。由學生依自我對知識翻新環境中的感知的社群意識填答。

3. 研究結果

研究結果分為三個部分，分析學生經過一學期知識翻新與知識論壇的學習環境後，第一、在知識翻新環境感知的改變情形；第二、社群意識感知的改變情形；第三、知識翻新環境與社群意識的關聯性為何？

3.1. 知識翻新環境感知

本研究分析學期前後學生對課堂知識翻新環境的感知情形。從表 1 的統計結果可以發現學生在知識翻新環境量表的分數後測高於前測，且達到統計顯著差異（t=-5.61, p< .001）。代表學生在經過一學期的課程，對知識翻新環境的感知有正向的提升。

表 1 知識翻新環境量表 t 檢定

	前測		後測		
	平均數	標準差	平均數	標準差	t 值
總分	3.66	0.62	4.50	0.47	-5.61***

****p< .001*

接著進一步針對，知識翻新環境量表包含三個子構面：想法中心、知識自主、社群互動進行前測與後測成對樣本 t 考驗。表 2 結果顯示學生在想法中心、知識自主、社群互動三個面向的平均數後測分數皆高於前測，且達到顯著差異（t=-5.91, p<0.01; t=-5.74, p<0.01; t=-4.98, p<0.01）。結果顯示學生經過一學期的知識翻新環境學習後，在知識翻新之想法中心、知識自主、社群互動此三個構面的感知程度皆有正向的提升。

表 2 知識翻新環境量表三項子構面成對樣本 t 檢定

變項	前後		後測		
	平均數	標準差	平均數	標準差	t 值
想法中心	3.69	0.55	4.50	0.50	-5.91***
知識自主	3.62	0.64	4.50	0.49	-5.74***
社群互動	3.67	0.73	4.51	0.53	-4.98***

****p< .001*

3.2.社群意識感知

本研究採用「社群意識量表」探究學生於學期前後社群意識感知的改變情形。表 3 結果顯示，學生在社群意識的整體表現後測分數高於前測分數，且差異達到顯著差異（t=-5.32, p< .001）。代表學生在經歷一學期知識翻新與線上論壇的學習，對社群意識有正向的提升。

表 3　社群意識量表前測與後測 t 檢定

變項	前測		後測		
	平均數	標準差	平均數	標準差	t 值
社群意識	3.24	0.52	3.90	0.47	-5.32**

***p< .001*

接下來，進一步分析社群意識二個子構面：連結項目與學習項目學生學期前後的感知情形。

由表 4 可知，學生在社群意識之連結項目與學習項目的感知程度從學期初到學期末的分數上升，且達到顯著差異（t=-5.13, p<0.01; t=-5.13, p<0.01），代表學生在社群意識的兩個子項度，經過一學期知識翻新與知識論壇的學習環境，有正向的提升。

表 4 社群意識量表兩項構面 t 檢定

變項	前測		後測		
	平均數	標準差	平均數	標準差	t 值
連結項目	3.20	0.62	3.89	0.55	-5.13**
學習項目	3.28	0.47	3.92	0.50	-5.13**

***p< .01*

4.3. 知識翻新環境與社群意識的相關性

由前述資料顯示，本研究學生知識翻新環境與社群意識的感知程度，到了學期末皆有正向且顯著的提升。本階段欲進一步了解學生對知識翻新環境與社群意識感知是否有相關存在。在此採用兩個量表的後測資料分數，先確認兩者間以及各構面間之間是否存在統計上的差異。再進一步進行皮爾森相關分析，了解兩個變項的關聯程度。

首先運用適於小樣本之魏克生符號等級檢定（Wilcoxon Signed Rank Test），進行知識翻新環境量表與社群意識量表後測的差異顯著檢定，分析結果如表 5 所示，Z 值為-3.91（p<0.001），達統計上的顯著。表示本研究社群對於知識翻新環境感知與對社群意識的感知，有統計上的顯著差異。

表 5 知識翻新環境量表與社群意識量表魏克生符號等級統計檢定

檢定	知識翻新環境量表後測 - 社群意識量表後測
Z 檢定	-3.91[b**]
漸進顯著性	<.001
精確顯著性	<.001

b. 基於正等級

表 6 顯示，本研究知識翻新環境量表後測與社群意識量表後測的皮爾森相關係數為 0.73（p<0.001），兩者屬於高度正相關，表示學生在學期末時，知識翻新環境感知與社群意識感知之間有相關性存在。由此結果可知，經過一學期知識翻新與知識論壇的學習環境後，當學生對知識翻新環境的感知程度較高時，也會有較高的社群意識感知。

表 6 知識翻新環境量表與社群意識相關係數摘要表

變項	社群意識量表
知識翻新環境量表	0.73***

****p< .001*

4. 結論

本研究透過知識翻新教學與知識論壇線上學習環境，並以機遇式分組的方式，以期能形塑知識翻新的學習環境，並促進社群的營造與學生社群意識提升。研究結果顯示，學生在「知識翻新環境感知」與「社群意識感知」，到了學期末有正向的提升。進一步了解子構面的表現：在知識翻新環境感知的三個構面想法中心、知識自主、社群互動，期末分數皆有顯著提升。社群意識感知的連結項目與學習項目，亦是在期末有顯著正向的成長。此外，本研究的學生在學期末的「知識翻新環境感知」與「社群意識感知」的分數，達到顯著正相關，顯示學生在知識翻新環境的知識翻新歷程中，社群在形成方案為共同目標下，學習與連結的能動性，使學生在發展想法與反思互動時，感受到彼此的支持，促進了社群意識的發展。

根據上述結果，本研究提出對未來教學現場的建議與應用：第一、在未來，教師在進行以知識創新或想法中心的課程時，可採用知識論壇做為增進學生合作學習的科技工具。本研究教師運用知識論壇做為線上工具，打破上課時間與課堂的時空限制，使學生於知識翻新歷程中，即時與想法互動，具體紀錄內容、閱讀建議、反思回饋。因此，知識論壇為學生知識翻新重要的協作平台，也促進了知識翻新環境的營造（Scardamalia & Bereiter, 2014）。第二、機遇式分組可使教師突破班級固定小組模式，可促進知識社群之經營和學習者社群意識的感知。鼓勵教師在教育實境或進行相關研究實，適時採用此分組方法，增進個體化的學習體驗與團體的輪動性。教師可藉由每次的隨機分組，使學生有機會與不同的成員互動，無論熟識或陌生；透過動態調整班級討論模式，使成員交流多元的觀點與想法，此合作方式有助增進創造力和解決問題的能力，達成知識翻新目標。同時，機遇式的互動模式與當今學習者主要接觸的虛擬或網路世界的合作方式，有很多共通性。教師運用此分組策略，可使學習者模擬

現實世界經驗，增加學習者動力；進一步使學生在參與各式團隊動態與合作溝通的經驗中，適應未來知識創造的社會。

參考文獻

Balboni, G., Perrucci, V., Cacciamani, S., & Zumbo, B. D. (2018). Development of a scale of Sense of Community in university online courses. *Distance Education*, *39*(3), 317-333.

Ergün, E., & Avcı, Ü. (2018). Knowledge sharing self-efficacy, motivation and sense of community as predictors of knowledge receiving and giving behaviors. *Journal of Educational Technology & Society*, *21*(3), 60-73.

Gutiérrez-Braojos, C., Montejo-Gámez, J., Ma, L., Chen, B., Muñoz de Escalona-Fernández, M., Scardamalia, M., & Bereiter, C. (2019). Exploring collective cognitive responsibility through the emergence and flow of forms of engagement in a knowledge building community. *Didactics of Smart Pedagogy: Smart Pedagogy for Technology Enhanced Learning*, 213-232.

Lin, K. Y., Hong, H.-Y., Chai, C. S. (2014). Development and Validation of the Knowledge-Building Environment Scale. Learning and Individual Differences, 30, 124-132.

Organisation for Economic Co-operation and Development (OECD) (2019). *PISA 2021 Creative Thinking Framework (Third Draft*). Retrieved on February 9, 2022, from https://www.oecd.org/pisa/publications/PISA-2021-creative-thinking-framework.pdf

Rovai, A. P. (2002). Development of an instrument to measure classroom community. *The Internet and higher education, 5*(3), 197-211.

Scardamalia, M., & Bereiter, C. (2014). Smart technology for self-organizing processes. *Smart Learning Environments, 1*, 1-13.

Scardamalia, M., & Bereiter, C. (2020). Will knowledge building remain uniquely human?. *Qwerty-Open and Interdisciplinary Journal of Technology, Culture and Education*, *15*(2), 12-26.

Scardamalia, M., & Bereiter, C. (2021). Knowledge building: Advancing the state of community knowledge. *International handbook of computer-supported collaborative learning*, 261-279.

Tan, S. C., Chan, C., Bielaczyc, K., Ma, L., Scardamalia, M., & Bereiter, C. (2021). Knowledge building: aligning education with needs for knowledge creation in the digital age. *Educational Technology Research and Development, 69*(4), 2243-2266.

基于深度学习的在线异步讨论活动设计及实证研究

Design and Empirical Research on Online Asynchronous Discussion Activities Based on Deep Learning

黄雅萱[1]，彭炜明[1*]，邵文豪[1]，汤名珍[1]，杨时娇[1]，周洋洋[1]
[1]华南师范大学教育信息技术学院
[*] 1390319329@qq.com

【摘要】 “互联网+教育”的时代背景下，在线学习成为了重要的学习方式，当代高质量教育改革提倡培养学习者深度学习能力，而在线学习对深度学习的促进效果取决于学习者之间是否能进行高质量、有深度的对话讨论。本文采用基于设计的研究范式，经过初步的设计与实践迭代，形成了基于深度学习的在线异步讨论活动设计框架，并通过量化数据验证其有效性，旨在通过活动设计的有效干预提升学生的深度学习能力，为提升在线讨论学习效果提供一条切实可行的路径。

【关键词】 深度学习；基于设计的研究；在线异步讨论

Abstract: *In the era of "Internet + Education," online learning has become a pivotal mode of education. Contemporary high-quality educational reforms advocate for the cultivation of learners' deep learning abilities. The effectiveness of online learning in promoting deep learning depends on whether learners can engage in high-quality and profound dialogue and discussions. This paper adopts a design-based research paradigm and, through preliminary design and iterative practice, establishes a framework for designing online asynchronous discussion activities based on deep learning. The validity of this framework is verified through quantitative data, with the aim of effectively intervening in student learning through activity design to enhance their deep learning abilities. This offers a practical pathway for improving the effectiveness of online discussion learning.*

Keywords: Deep Learning,Design-Based Research,Online Asynchronous Discussion

1. 绪论

随着科技的进步与时代的发展，“互联网+教育”已然成为当代教育发展的时势所趋。《关于促进在线教育健康发展的指导意见》中提出：“大力发展在线教育，将有助于构建网络化、数字化、个性化、终身化的教育体系，有助于构建一个人人都能学、随时随地学、每时每刻学的学习型社会(教育部，2019)。”美国国家研究委员会指出教育变革应侧重于教授深度学习能力，以培养 21 世纪技能錯誤！ 找不到參照來源。(James W. Pellegrino&Margaret L. Hilton，2012)。本研究通过在线异步讨论提高大学生深度学习能力，对深度学习的评价与实现模式发展具有积极影响，所构建的评价量表为相关研究提供参考。

2. 基于深度学习的在线异步讨论活动设计框架

2.1. 基于深度学习的在线异步讨论活动设计思路

本研究选取某大学大二本科生的一个教学班进行为期一个学期的实践研究，共分三轮进行。每轮实践都在上一轮的基础上进行修正和完善，以优化基于深度学习的在线异步讨论模型。

在实践开始前，向学生介绍了学习通讨论区的基本操作，并进行深度学习能力的前测。

实践过程中，利用学习通在线讨论区进行小组活动，48 名学生被随机分成 10 个小组，每组 4-6 人。通过三轮实践研究，不断调整活动框架和干预策略，具体设计思路如图 1 基于深度学习的在线异步讨论活动设计思路所示，以推动学生的在线讨论深度和深度学习能力的提升。

实践结束后，向学生发放后测调查问卷并进行分析，以验证框架的有效性。

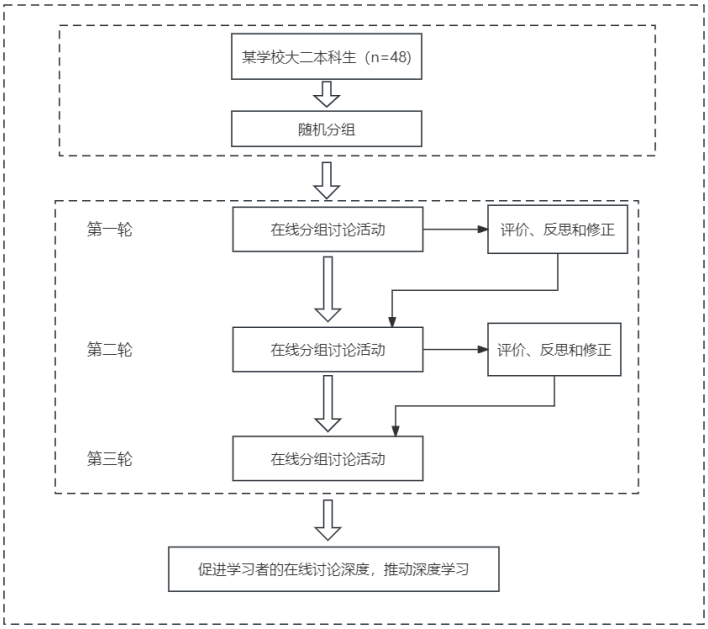


图 1 基于深度学习的在线异步讨论活动设计思路

2.2. 初步形成基于深度学习的在线异步讨论活动设计框架

2.2.1. 阶段策略应用于深度学习

在线讨论以学生为主导，阶段策略的干预为学生在缺乏教师指导的情况下有效规划小组讨论任务的推进及讨论的深度交互。王智颖发现阶段策略模式显著提升了学生的深度学习能力（王智颖，2023）。运用阶段策略能优化在线讨论，促进小组的深度讨论。

2.2.2. 形成基于深度学习的在线异步讨论活动设计框架初版

本研究以学生为主导，基于在线讨论区根据不同小组分别布置不同的讨论话题，设计了四个由浅入深的实施阶段。基于深度学习的在线异步讨论活动设计框架初版如图 2 基于深度学习的在线异步讨论活动设计框架初版所示。

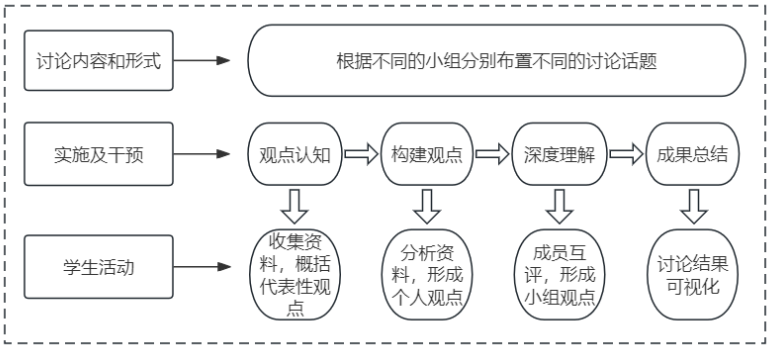


图 2 基于深度学习的在线异步讨论活动设计框架初版

2.3. 优化完善基于深度学习的在线异步讨论活动设计框架

2.3.1. 第一轮实施与调整

本研究首轮实践以了解学习者在线讨论初始水平为目的，通过小组讨论形式实施基于阶段策略的讨论活动。分析学生讨论记录后发现，学生讨论停留在浅层，缺乏深度思考与

互动反馈，未达到深度学习。讨论框架的局限性导致学习者被动参与，缺乏兴趣，讨论趋于浅层化。讨论活动数据见表 1。

针对上述问题，作出调整如下：首先，讨论主题设置更开放，允许学习者自选主题；其次，融入个人与小组反思支架到协作活动中，以促使学习者反思，深化讨论。

表 1 讨论活动数据表

	人数	发帖量	所占比例
深度	17	13	7.43%
中度	30	55	31.43%
浅度	42	105	60.00%

2.3.2. 第二轮实施与调整

第二轮实践采用开放性主题设计，融入反思支架，以激发学习者讨论积极性并引导其深入讨论。实践表明，此举提高了学生的在线讨论参与度，讨论内容更加深入，促进了深度学习。在讨论实施后，从讨论区文本中可以看到，学习者针对话题表达观点不再是复述资料，而是融入了更多自我想法，与第一次讨论相比促进了学生的深度学习，讨论活动数据见表 2。然而，小组内缺乏观点探讨和思维碰撞，问题式主题设计及不明确的小组分工导致学习者偏向回答而非质疑，影响了深层讨论。

表 2 讨论活动数据表

	人数	发帖量	所占比例
深度	20	29	15.43%
中度	38	88	46.81%
浅度	31	67	35.64%

针对过程中问题，活动调整包括：一，引入辩论主题增加观点冲突；二，明确小组分工，促进深度讨论；三，加入反思问题评估干预效果。

2.3.3. 第三轮实施与调整

第三轮讨论通过辩论和课堂成果展示激发观点冲突，并引入明确分工策略，明确学习者职责。同时，要求学习者在发帖时标明角色，确保发帖内容与职责匹配。在讨论实施后，从讨论区文本中可以看到，学习者针对话题深入地阐述了自己的想法，并且能根据小组分工完成相应职责，能对观点进行分析质疑，极大促进了深度学习。经过第二轮的设置开放问题与融入反思框架及第三轮的融入小组分工的干预策略，最终形成了基于深度学习的在线异步讨论活动设计框架，如图 3 基于深度学习的在线异步讨论活动设计框架所示。

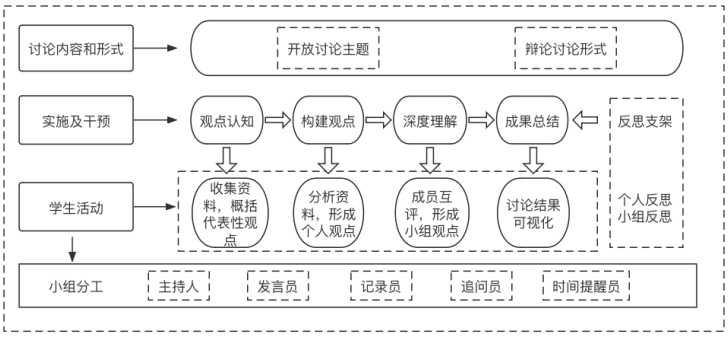


图 3 基于深度学习的在线异步讨论活动设计框架

3. 基于深度学习的在线异步讨论活动设计框架应用效果分析

3.1. 应用效果评价工具

研究团队基于现有深度学习评估工具，制定了深度学习能力问卷，包括三道基本信息题和六道矩阵量表题，共回收 42 份问卷，得到 38 份有效问卷。通过问卷调查以学习效果的前后测对比为依据，验证了学生深度学习能力的提升。

3.2.应用效果分析

研究通过单样本K-S检验发现后测数据非正态分布，故采用Wilcoxon符号秩检验分析学习者前后深度学习能力的差异。检验结果见表 3深度学习能力配对样本Wilcoxon符号秩检验所示。

表 3深度学习能力配对样本Wilcoxon符号秩检验

变量	测试	均值	标准差	Z	渐进显著性P
合作能力	前测	3.96	0.51	3.703	0.000
	后测	4.37	0.48		
有效沟通和表达	前测	3.92	0.64	2.395	0.017
	后测	4.24	0.59		
自主学习能力	前测	3.89	0.54	1.720	0.085
	后测	4.08	0.58		
学习毅力	前测	3.87	0.49	2.649	0.008
	后测	4.11	0.49		
元认知能力	前测	3.75	0.56	3.169	0.002
	后测	4.15	0.65		
自我效能感	前测	3.77	0.63	3.077	0.002
	后测	4.15	0.51		

从表3可知，表中六个维度的后测得分均高于前测得分，其中合作能力、有效沟通和表达、学习毅力、元认知能力、自我效能感五个维度的显著性均小于0.05，说明经过了三轮教学实践后，实验对学生的深度学习能力产生了积极影响，学生的整体深度学习能力得到了提升，其中，元认知能力与自我效能感这两个维度的显著性比较接近于 0，说明学生的元认知能力与自我效能感实施前后发生了较大的变化。

4.总结与展望

通过实验研究后，本研究发现在线异步讨论中开放的主题对学习者的深度学习具有正向的作用，不仅提高了学习者的深度讨论水平，同时显著激发了学习者参与讨论的积极性，同时恰当地实施干预策略，可以增强学习者参与在线讨论积极性，从而推动学习者深度学习能力的提高。在后续研究中，将会注重引导学习者对讨论时间的合理安排，并针对小组分工的干预策略进行多轮实践，分析其实际的干预效果，进一步验证本研究设计的有效性。

参考文献：

卜彩丽、李飒、王静、张思和董乐(2022)。为深度学习而思：反思日志促进大学生元认知发展的实证研究。**现代教育技术(09)**，73-81。

王智颖、翟芸和吴娟(2021)。在线异步讨论中角色轮换脚本对大学生深度学习的影响。**现代远程教育研究(03)**，100-112。

陶宇炜(2023)。具身认知视域下双线混融教学促进深度学习的行动研究。**现代教育技术 (01)**，66-73。

王京京(2019)。深度学习视域下的翻转课堂实证研究硕士学位论文,浙江工业大学。https://link.cnki.net/doi/10.27463/d.cnki.gzgyu.2019.000613doi:10.27463/d.cnki.gzgyu.2019.000613.

胡敏和叶映华(2022)。走向线上空间:基于异步在线讨论的批判性思维培养。**上海教育(24)**，43-44。

教育部(2019)。教育部等十一部门关于促进在线教育健康发展的指导意见。**中华人民共和国教育部公报(09)**。14-17。

James W. Pellegrino & Margaret L. Hilton,(2012).**Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century.[R]**。

MARTIN F,SUN T & WESTINE C D,(2020). A systematic review of research on online teaching and learning from 2009 to 2018 [J].**Computers & Education** ,159.

智能技术赋能劳动教育评价创新的实施路径

The Implementation Path of Smart Technology Enabled Labor Education Evaluation Innovation

梁心贤*，陈晗，陈颖珊
华南师范大学教育信息技术学院
*2023020795@m.scnu.edu.cn

【摘要】劳动教育作为五育的重要组成部分，近年来我国政府和教育部陆续发布各类文件，倡导鼓励各级各类学校重视劳动教育，确保劳动课程的质量。然而，当前劳动教育评价存在理念陈旧、缺乏系统的评价体系和科学的评价方式等问题，严重阻碍了劳动教育的发展。为此，本研究围绕新时代劳动教育改革新要求，提出以新时代先进理念为引领创新劳动教育评价理念；并在此基础上，将以人为本的先进教与学理论和劳动教育评价方法有机结合，以健全劳动教育评价体系；最后运用人工智能大模型等智能技术手段全方位、多角度收集数据，融入多元化教育评价，以丰富劳动教育评价方式，打造专业化的教育评价队伍。
【关键词】劳动教育；劳动教育评价；智能技术

Abstract: *As an important part of the five education, in recent years our government and the Ministry of Education have issued various documents to advocate and encourage schools at all levels and in all types to pay attention to labor education and to ensure the quality of labor courses. However, the current evaluation of labor education has problems such as outdated concepts, lack of a systematic evaluation system, and scientific evaluation methods, which seriously hinder the development of labor education. For this reason, this study focuses on the new requirements of labor education reform in the new era, and proposes to innovate the evaluation concept of labor education with the advanced concept of the new era as the leader; on this basis, organically combines the advanced human-oriented teaching and learning theory and the evaluation method of labor education to improve the evaluation system of labor education; and finally, uses the artificial intelligence big model and other intelligent technological means to collect data from all directions and angles, and integrates the diversified education evaluation to enrich the evaluation methods of labor education and create a professional education evaluation team.*
Keywords*:Labor Education, Labor Education Evaluation, Intelligent Technology*

1. 问题的提出

劳动教育一直以来都是老生常谈的话题。业界对于劳动教育的研究主要集中在劳动教育的内涵、目标、评价和实施方案这几个方面。严莹认为，劳动教育是每个青少年学习中的重要一环，是培养学生综合素质必不可缺的一部分（严莹, 2021）。檀传宝针对不同年龄段的孩子提供差异化的设计和说明，提供了劳动意识、劳动技能和劳动经验等方面的内容（檀传宝, 2019）。著名教育家莫尔曾提出，学生在坚持学习科学理论和文化知识的同时，每天至少应该劳动 6 小时（朱磊，2016）。《说文解字》云：“勤，劳也。”从某种意义来讲，勤和劳的意思是相通的。热爱劳动，勤奋劳作是植根于中华民族血脉深处的文化基因（张建伟，王光明，2022）。2020 年，中共中央国务院发布了《关于全面加强新时代大中小学劳动教育的意见》，明确提出了健全劳动素养评价制度，将劳动素养纳入学生综合素质评价体系，并制定了相应的评价标准（国务院, 2020）。同年，教育部印发了《大中小学劳动教育指导纲要（试行）》通知（以下简称“纲要”），纲要提供了劳动教育的评价依据和评价体系等（教育部，2020）。基于此，本研究立足本土新时代劳动教育评价改革方向，用好智能技术以赋能劳动教育评价创新，构建“智能技术+劳动教育”的评价指标体系，以期为新时代劳动教育评价的理论与实践做出积极贡献。

2. 劳动教育评价现状呼唤评价的创新

科学的评价对劳动教育的实施起着重要的导向和质量监控作用。然而，经过调查研究发

现，当前中小学劳动教育的评价体系存在一些问题，例如评价体系的不完善、评价目的不明确以及评价方式不够科学等。首先，劳动教育的评价体系过于片面。学校往往只关注学生的劳动成品，而忽视了对劳动意识、劳动习惯、劳动能力等核心素养的评价（程豪, 2021）。其次，劳动教育评价体系缺乏科学性。当前教师对劳动教育主要采用的是结果性评价，其评价依据主要源于教师的主观评价。这种评价方式缺乏多源数据的支持，无法对学生进行全方位的综合评价，从而无法实现通过评价驱动育人的目标（陆根书, 李运福, 2023）。最后，劳动教育评价目标不明确。由于教师素养不足，导致教师难以基于本节课内容对学生劳动教育目标做出系统的规划。为了解决上述问题，我们建议借助智能技术，建立系统科学的劳动教育评价指标体系。在开展劳动教育的过程中，可以借助物联网等技术实时监测与记实评价。此外，还需要探索多元的劳动教育评价方式，并着力解决分级分类的劳动教育评价问题。

3. 国内外智能技术赋能劳动教育评价的借鉴

国内对于智能技术赋能劳动教育评价的综述主要围绕理念策略、实践应用、评价体系等方面展开。具体而言，陈超等指出将劳动教育纳入高职教育的学分银行，利用互联网、大数据和人工智能等前沿技术构建学情数据库，提出构建“一核三阶五维”劳动素养评价体系（陈超, 欧彦麟, 2022）；栾慧敏等着眼于中小学劳动素养培养，提出利用电子档案袋评价方式，关注学生完整的劳动过程（栾慧敏, 田阳, 吉慧莹, 2023）；李燕玲等构建微认证和数字徽章赋能劳动教育评价方法并开展实践应用。国外方面对于劳动教育评价的研究主要集中在评价理论、模式、策略、方法工具及应用等方面，如芬兰利用电子档案袋赋能评价过程，注重对学生横贯能力的考查(European Commission,2023)；丹麦围绕课程目标、课程框架进行考核评估，鼓励将多种评价方式结合，运用技术赋能评价(Børne-og Undervisningsministeriet,2019)；德国将数字技术作为当前中小学劳动教育改革的重点，以学生学习、教学过程和产品成果为考核重点，构建了融合多模态、多维度数据且应用自然语言处理、大数据等数字技术支持的劳动教育评价框架(Senatsverwaltung für Bildung, Jugend und Familie Berlin,2019)。可见，国内研究在劳动教育评价理论上有较多探究，同时注重培养学生实践动手能力，且逐渐将“以人为本”作为劳动教育评价的内在要求，但缺乏在对劳动素养的深层评价实践，且劳动教育评价停留在技术支持数据收集的阶段。相比之下，国外关于劳动教育评价的研究从理论层面到实践层面已有较为完整的逻辑链了，但其整体会偏向于小场景教学为主，以未来就业为导向，更偏重脑力劳动的课程，缺少“自上而下”的整体设计。

从已有研究中得到如下启示：一是创新劳动教育评价内容，需要拓宽评价范围，丰富评价内容，运用智能技术使得评价结果可视化。二是革新劳动教育评价方式，鼓励实现个性化、综合性评价，开展伴随性评价，帮助教学者开展高效、个性化的教学。三是增值劳动教育评价主体权益，减少评价过程的主观偏见与不公，提供全面、精准、客观的评价结果。四是变革教育评价手段，摆脱“唯分数论”“技术至上”“算法偏见”的困境，将劳动教育主体从评价中解放出来。

4. 智能技术赋能劳动教育创新评价

劳动教育的培养目标和劳动教育核心素养，决定其评价在于促进学生的全面发展(阮瑜, 2022)。当前劳动教育评价存在的问题及未来改革方向呼吁智能技术的赋能劳动教育创新评价，即集成多种智能技术，利用人工智能大模型进行智能分析，通过收集多方数据，建立评价模型，充分利用实时反馈功能和人机协同，不断优化迭代评价模型，形成个性化评价和策略，供给教师诊断评价劳动教育情况。基于此，本研究提出了智能技术赋能劳动教育创新评价路径，如图 1 所示。

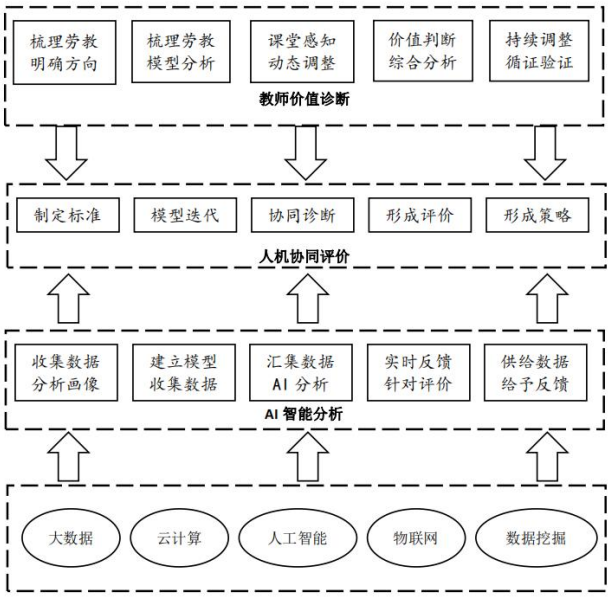


图 1 智能技术赋能劳动教育创新评价路径

劳动教育的实施从制定标准开始，明确评价目标和方向。随后，通过模型迭代，不断完善数据收集和处理机制。协同诊断阶段，AI 与教育者共同发挥作用，提高评价的准确性和精度。形成评价阶段，结合智能算法与人工评价，全面展示学生素养发展，为评价提供互证依据。最后，基于数据分析和最新教改要求，生成改进策略，实现课堂的持续优化。这一连贯过程，既提升了教师的教育教学素养，又为学生提供了更加精准有效的劳动教育。通过这一循环往复的机制，劳动教育的质量得以不断提升。

4.1 理念引领：树立新时代劳动教育评价观

在推进智能技术赋能劳动教育评价的过程中，首要任务是转变思想理念，确立符合新时代的劳动教育评价观。首先，我们要坚持立德树人的根本任务，明确评价的核心目标是促进学生的全面发展和教育质量的整体提升(刘邦奇等,2021)。在日常教育实践中，应积极引导各方主体形成正确的育人观和质量观，以适应新时代教育发展的需求。其次，我们需要树立多元评价的观念，注重实践劳动评价。劳动教育评价应运用多元主体和多元方式，全面、客观地反映学生的劳动表现与成果。同时，要特别关注学生的实践操作能力，激发他们的创新精神，致力于培养具备实践能力的劳动者。最后，我们应破除对技术的恐惧，树立创新性评价技术观。人工智能、大数据等智能技术能够为教育评价的各个环节提供有力支持，形成基于数据的客观性思维、动态变化的持续性思维和价值呈现的发展性思维等创新性技术思维。通过这些举措，我们将能够更好地利用智能技术推进劳动教育评价的改革与发展。

4.2 标准规范：构建劳动教育评价体系

评价标准是评价落地的重要梯子。当前，劳动教育缺乏完善的评价体系，导致劳动教育评价与教育两张皮(赵阳,2023)。为构建劳动教育评价体系，可从以下三方面完善。一是明确评价目标。在构建劳动教育评价体系之前，需要明确评价的目标和目的，基于学生劳动教育学习画像，分析学生劳动教育现状，结合课程标准，从而明确培养目标，确定评价方向。二是完善建立劳动教育专家组，加快推进专家组走进校园，与校内组成劳动教育小组，共同构建劳动教育评价体系，助力评价体系高质量建设。三是构建四位一体的评价体系，即即改进结果评价，强化过程评价，融入增值评价和健全综合评价。

4.3 专业支持：打造专业化教育评价队伍

在智能时代，教师评价数字素养的提升对于提高教育评价质量和效率至关重要(刘飞君,2022)。因此，建设专业化教育评价队伍是非常有必要的，可以从以下两个方面打造。一是组织开展专业培训，针对教师劳动教育评价数字素养的需求，成立专家教研团，组织专业

的培训课程和研讨会，帮助教师掌握数字化评价工具和技能，提升信息素养和数字技术应用能力。二是开展创新实践项目，鼓励教师将数字技术应用于劳动教育评价中，通过实践探索数字化评价的有效方法和模式，建设有效劳动教育评价模型。同时完善激励与评价机制，对教师在数字化评价方面的表现进行评估和奖励，激发他们提升评价数字素养的积极性和动力。

5. 总结

劳动教育评价作为劳动教育过程中必不可少的环节，是体现劳动教育情况最重要的着力点。本研究将培养学生劳动教育核心素养作为劳动教育发展阶段的核心，坚持以人为本，从评价理念、评价体系和评价方式等方面出发，对智能技术赋能劳动教育创新评价展开探索。为了响应国家深化改革新时代教育评价的号召，本研究围绕新时代劳动教育改革新要求，提出以新时代先进理念为引领创新劳动教育评价理念；并在此基础上，将以人为本的先进教与学理论和劳动教育评价方法有机结合，构建四位一体的评价框架，以健全劳动教育评价体系；最后运用人工智能大模型等智能技术手段全方位、多角度收集数据，融入多元化教育评价，以丰富劳动教育评价方式，打造专业化的教育评价队伍。我们希望，本研究能为新时代劳动教育评价的理论创新做出积极贡献。

参考文献

严莹. 基础教育阶段劳动教育的实施路径和保障措施[J]. 商丘师范学院学报, 2021(1): 103 — 105.

檀传宝. 劳动教育的概念理解——如何认识劳动教育概念的基本内涵与基本特征[J]. 中国教育学刊, 2019（2）

朱磊. (2016). 早期空想社会主义劳动教育思想及其当代价值. 广西师范大学学报(哲学社会科学版) (02), 163-167.

张建伟, 王光明. (2022). 新时代《小学生综合素质评价手册》改革：新要求及新路径. 教育理论与实践(17), 7-11.

中共中央、国务院. 关于全面加强新时代大中小学劳动教育的意见[EB/OL].(2020-3-20)[2024-01-10].http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202003/t20200326_435127.html.

教育部. 大中小学劳动教育指导纲要(试行)[EB/OL].(2020-07-09)[2024-01-10]. http://www.moe.gov.cn/srcsite/A26/jcj_kcjcgh/202007/t20200715_472808. html.

程豪. (2021). 从外在设计到内源发展：劳动教育变革的可能逻辑. 大学教育科学(03), 54-62.

陆根书, 李运福. (2023). 高校德体美劳过程性评价研究. 河北师范大学学报(教育科学版) (04), 6-13.

陈超, 欧彦麟. (2022). 高职院校“一核三阶五维”劳动教育评价体系构建研究. 教育与职业(20), 102-106.

栾慧敏, 田阳, 吉慧莹. (2023). 基于电子档案袋的中小学劳动教育评价：芬兰的经验与启示. 中国考试(10), 71-78+87.

European Commission.Single-structure primary and lower secondary education [EB/OL].[2023-05-31].

Børne-og Undervisningsministeriet.Folkeskolens formål[EB/OL].(2019-02-05)[2024-01-10]. https://emu.dk/grundskole/uddannelsens-formaal-og-historie/folkeskolens-formaal.

Senatsverwaltung für Bildung, Jugend und Familie Berlin. Entwurfsfassung Curriculare Vorgaben für die gymnasiale Oberstufe: Zusatzkurskurs Digitale Welten[EB/OL].2019-5-28.

阮瑜. (2022). 立德、笃行、培能的劳动教育评价体系建构与实施. 人民教育(12), 54-56.

刘邦奇, 袁婷婷, 纪玉超, 刘碧莹 李岭. (2021). 智能技术赋能教育评价：内涵、总体框架与实践路径. 中国电化教育(08), 16-24.

赵阳. (2023). 高职院校劳动教育评价的原则、内容体系及实施路径. 教育与职业(05), 13-20.

刘飞君. (2022). 智能时代大学生劳动教育的价值重塑及实施进路. 教育理论与实践(12), 8-12.

大模型赋能高等教育的现状调查与对策分析

Investigation and Countermeasure analysis of the Current situation of Large-scale model empowerment of Higher Education

陆宇，张日悦

北京邮电大学人文学院

2246039034@QQ.COM 305217556698QQ.COM

【摘要】 数智时代以 ChatGPT 为首的生成式人工智能在众多行业开展了广泛应用。由于其具有的强大语言理解和生成能力，使其适用于各种自然语言处理任务，在推动教育理念创新、探索教育模式新生态的过程中可以起到积极作用。本研究通过设立调查问卷，熵值赋权法等研究方法，开展大学生对大模型赋能高等教育的认知、预期、态度等现状调查，分析总结在大学生视角下大模型赋能高等教育中存在的问题，并针对这些问题提出了具体策略，以期为大模型赋能高等教育、促进高等教育数字化转型升级提供有益指导。

【关键词】 大模型；高等教育；数字化转型；熵值法

Abstract: *In the era of digital intelligence, generative artificial intelligence, mainly ChatGPT, has been widely used in many industries.Due to its strong language understanding and generation capabilities, it is suitable for various natural language processing tasks, and can play an active role in promoting the innovation of educational concepts and exploring the new ecology of education models.In this study, we set up a questionnaire and used entropy weighting methods to investigate the current situation of college students' cognition, expectation and attitude towards large model empowerment of higher education, analyzed and summarized the problems existing in large model empowerment of higher education from the perspective of college students, and put forward specific strategies for these problems, in order to provide useful guidance for large model empowerment of higher education and promote the digital transformation and upgrading of higher education.*

Keyword： *large models,Higher education, digital transformation, entropy method*

1. 引言

21 世纪随着信息技术迅猛发展，数字化、网络化、智能化的技术工具的不断演进，高等教育正面临着前所未有的转型机遇与挑战危机(尚智丛,2023)。自 2022 年底以 ChatGPT 为首的生成式人工智能（Generative Artificial Intelligence, Gen AI）广泛传播以来，凭借其具有的高效文本分析与数据处理能力，越来越多的国外教育工作者考虑将其与教育深度融合，打造个性化、智能化的新兴教学模式。

党中央、国务院也对生成式人工智能赋能教育事业投入了充分关注，在 2022 年 1 月印发的《“十四五”数字经济发展规划》中强调推进智慧教育，加快高等教

育数字化转型的实施与推进。在此背景下，大数据、人工智能以及机器学习等技术在教育领域得到广泛关注，尤其是生成式人工智能的引入，被认为是推动教育现代化、实现教育公平与质量提升的关键力量(吴兰岸，2023)。但由于大模型技术的诞生年限相对较短，在中国的应用根基相对薄弱，中文语料下大模型的实证效果有待考证，申丽萍以中学历史学科为例，开展了大语言模型赋能教育的实际效果研究，结果表明当前大语言模型虽然能够对自己输出的答案做出完整流畅的表达,但是在中学历史知识测验上仍然远远低于适龄儿童的平均水准，在知识储备、逻辑推理方面仍然有巨大提升空间（申丽萍，2023）。大模型赋能高等教育的实践之路仍然具有多重挑战，基于此为了更好地了解大语言模型在我国本土的适应状况，本文从大学生视角出发开展对大模型赋能高等教育的现状研究:探究大模型赋能高等教育存在哪些问题、大模型如何促进大学生学习增效、影响大模型赋能高等教育的可能因素等问题,以帮助研究者准确把握大模型与中国高等教育的契合程度并为后续研究打好基础。

2. 调查方法与实施过程

本研究采用文献分析法、问卷调查法、熵值赋权法等研究方法。研究分三个步骤展开。

步骤一:通过文献和问卷调查方法搜集并统计大学生对大模型的认知情况、大学生对大模型了解程度和态度、探究影响大模型赋能高等教育普及的社会因素。

步骤二:通过熵值赋权法分析获得大模型赋能高等教育各影响因素之间的重要程度，进行排序处理寻找重点要点。

步骤三:针对问卷调查与熵值法所进行的分析，对大模型赋能高等教育中存在的各项问题进行专项对策研究。

3. 问卷调查

本次调查通过问卷星平台进行设计并发布问卷，邀请北京 xx 大学在校生进行作答,共发放问卷 232 份,回收 221 份,其中有效问卷 200 份,有效回答率为 86.2%。且对所有有效问卷进行信度分析，Cronbach α 系数值为 0.89，证明问卷的信度较高。问卷各部分的结构效度均达到 0.73 以上满足探求研究目的的要求。

3.1 .问卷设计

本研究为探究不同大学生群体对大模型赋能高等教育的看法、大学生对大模型赋能高等教育的发展预期、影响大模型赋能高等教育的可能元素等问题。将问卷从三个部分进行展开 1.学生基本信息。2.大模型赋能高等教育的现状调查。3.影响大模型赋能高等教育的可能因素。问卷如表 1 所示,第一第二部分数据如表 2 所示。

表 1 问卷题目设计

结构	题目因素		题目设置
第一部分 受访者基本情况调查	基本情况	学习背景	Q1 性别 Q2 年龄 Q3 专业
		大模型的关注情况	Q4 关注度*

第二部分 大模型赋能高等教育现状情况调查	现状情况	大模型认知情况和态度	Q5 认知程度*Q6 认知渠道** Q7 期望态度*Q8 存在问题**
		大模型参与情况	Q9 参与程度*Q10 参与方式**
第三部分 影响大模型赋能高等教育影响因素调查	内部因素		Q11 硬件因素*
			Q12 软件因素*
			Q13 自我效能感*
	交互因素		Q14 学习沉浸感*
			Q15 人机交互性*
	外部因素		Q16 伦理与隐私*
			Q17 数据库的准确性*

注：问卷中*代表以分数（5 分制）进行答题，**代表为主观题，内容不少于三十字。

表 2 第一第二部分问卷数据表

Q1 性别	男			女	
	53%			47%	
Q3 专业	理			文	
	56%			44%	
Q4 关注度	1	2	3	4	5
	2.5%	15%	30%	29.5%	24%
Q5 认知程度	1	2	3	4	5
	3.5%	33%	33.5%	19.5%	10.5%
Q7 期望态度	1	2	3	4	5
	0%	5%	20%	43.5%	31.5%

Q9 参与程度	1	2	3	4	5
	5%	12.5%	35.5%	28%	19%

3.2. 问卷分析

从表 2 可以看出，在本次问卷调查的受访者信息中，男生比例要较大于女生，理科人数占比大于文科人数占比，但由于本研究的样本量相对较少，且实验地点是一所理工类院校，所以比例在正常范围，所产生的结果具有一定代表性水平。

问卷第二部分中关于大模型赋能高等教育的现状研究调查中，大学生普遍在关注度上呈现了较高水平，83.5%的学生对大模型赋能高等教育这件事的关注程度在 3 分以上；但是在认知程度上表现不佳，70%的大学生认为自己对大模型赋能高等教育的认知程度较低，并不能较好地将大模型与高等教育进行有机融合；在期待程度上绝大多数大学生希望未来的高等教育与大模型能够进行深度融合，促进高等教育的数字化转型；在参与程度上绝大多数用户在接受高等教育的过程中使用大模型作为辅助手段，但使用频率不高多在 3 分（偶尔使用）左右徘徊。

4. 熵值赋权

熵值法是基于各指标所含信息量大小的赋权方法，可避免权重受到人为因素过多的干扰。在信息论中，熵是衡量不确定性的一种度量，熵值法赋权通过计算每个决策因素的信息熵，来评估其对决策问题的重要性（沈光辉，2023）。

4.1. 熵值赋权计算

在综合评价中，若同一指标的数值差异较小，则意味着这一事件具有较大的不确定性，信息熵较大，因此赋予该指标较低的权重。反之，若同一指标的数值差异较大，则该指标的权重较高，具体步骤如下：

首先设决策举证为

$$D=\begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1n} \\ X_{21} & X_{22} & \ldots & X_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ X_{n1} & X_{n2} & \ldots & X_{mn} \end{bmatrix} \quad (1)$$

其中 xij 为第 i 个被评价对象的第 j 项评价指标的值，m 为被评价对象的个数，n 为评价指标的个数。运用熵值法确定指标权重的步骤如下。

一、计算第 j 项评价指标下，第 i 个被评价对象的归一化数值，得

$$P_{ij}=\frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \quad (2)$$

二、计算第 j 项指标的熵值，得

$$e_j=-\frac{1}{\ln m}\sum_{i=1}^{m} p_{ij}\ln p_{ij} \quad (3)$$

三、计算第 j 项指标的偏差度，得

$$d_j = 1 - e_j \tag{4}$$

四、将第 j 项指标的偏差度归一化，即权重为

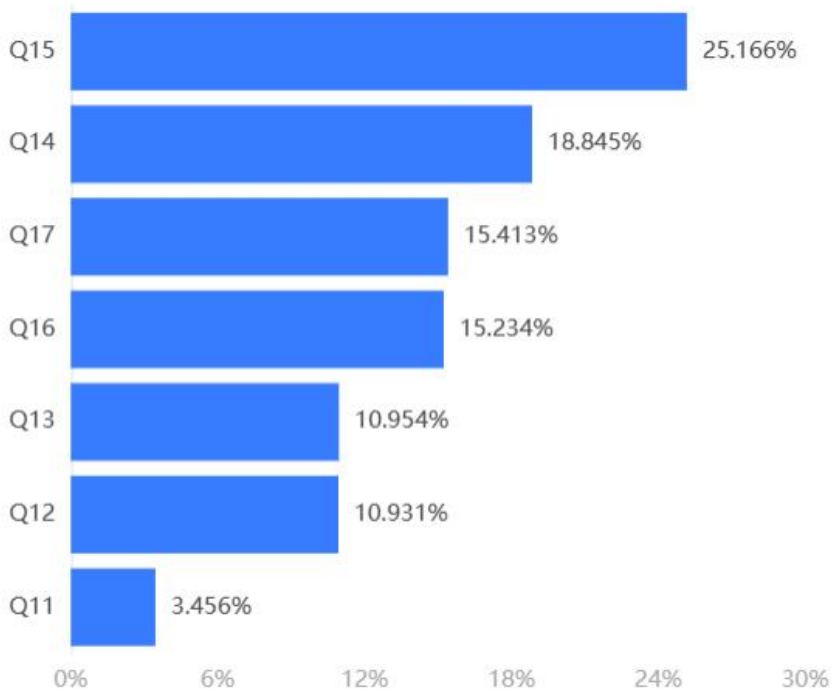
$$w_{cj} = \frac{d_j}{\sum_{k=1}^{n} d_k} \tag{5}$$

对 7 项可能的影响因素熵值权重如下表 3

表 3 影响因素权重表

因素	信息熵值 E	信息效用值 D	权重（%）
Q11	0.963	0.037	3.456
Q12	0.884	0.116	10.931
Q13	0.884	0.116	10.954
Q14	0.801	0.199	18.845
Q15	0.734	0.266	25.166
Q16	0.839	0.161	15.234
Q17	0.837	0.163	15.413

对上述影响因素权重进行排序处理可得到图一



图一 影响因素排序图

4.2. 熵值赋权分析

根据上述排序图我们可知，在大学生视角下大模型赋能高等教育的影响因素排名为 Q15>Q14>Q17>Q16>Q13>Q12>Q11，根据调查问卷中对影响因素的划分，我们可以总体上认为影响大模型赋能高等教育的主要因素为交互元素，其次为外部因素，最次为内部因素。其中权重最高为 Q15 人机交互性，说明绝大多数的学生都认为，在大模型赋能高等教育的学习过程中，提高人机交互性即增强学生主体与大模型工具之间的交互程度是提高大模型赋能高等教育效率的重要手段。而增强人机交互性也是有效提高学生学习沉浸感的有效手段，所以大多数的学生认为这两项因素所权重应当最高。其次相较于大模型本身所要求配套的软硬件设施，大学生认为大模型所受训练的数据库的准确程度与使用大模型所造成的科学伦理与隐私安全问题对高等教育的学习体验会产生更加重要的影响。排序最低为 Q11 硬件因素，说明大模型本身的低成本、易获得、易使用的特性使得绝大多数的大学生可以较为便利地使用大模型工具，硬件因素几乎不会对大模型赋能高等教育造成较大影响。

5. 存在问题的对策分析

5.1. 存在的问题

5.1.1. 大模型信息素养较低，文理差距较大

通过上述的数据分析，我们可以看出，在统计对大模型赋能高等教育的认知调查中，仅有 29.5%的大学生群体认为自己对大模型这一新型工具表示有充足认识，剩下 70.5%都对大模型的工作原理与基础概念表示并不理解，信息素养有待增强。且由于学科性质的限制，理科同学对大模型的认知程度要远高于文科同学，具体表现在认知程度水平三分以下的同学中文科同学占 68.5%。

5.1.2. 教学模式老旧，急需做出改变

在传统老旧的教学模式中往往强调学生将注意力完全集中在教师的课堂之中，对于使用辅助设备往往处于禁止的态度（苗勇胜，2013）。而使用大模型去赋能高

等教育则强调在学习的过程中将大模型作为学习帮手,老师使用大模型减少工作量,学生使用大模型增加知识面,大模型作为工具始终伴随着学习的始末。这与传统的教学模式所背道而驰。在 Q8 存在问题的问卷调查中多数同学表示目前存在的问题有:“老师不支持使用大模型工具”“上课老师会收手机”“老师强调课堂要注意听讲”等。

5.1.3.数据库质量难以保证,科学伦理受到挑战

数据库的质量问题是大多数学生所关注的重点所在,教育作为较为特殊的应用领域,其具有知识准确性、意识形态正确性、过程可解释性等要求(申丽萍,2023)对数据库的要求极为严格,而大模型的本身特征就是要求大量数据作为训练才能打造更加智能的新一代生成式人工智能,这两者的矛盾倒逼教育设计者必须做好顶层设计,从大模型的设计维度做好保障。其次大多数学生反映担心在高等教育领域使用大模型会造成学生成瘾、科研诚信、考试作弊等伦理问题。

5.2.解决的对策

5.2.1. “宣传+引导”:提高全体学生的大模型信息素养

宣传教育是提升学生大模型信息素养的基础工作,针对不同学科学生的不同性质应当开展统一的宣传活动、降低学习门槛、缩小不同学科间的大模型信息素养差距。如:增设相关公共课程,在高等教育阶段增设大模型及其应用的课程,内容覆盖大模型的基本概念、技术原理、应用案例等,确保学生能系统地学习和理解大模型。举办讲座和研讨会:定期邀请大模型领域的专家学者举办讲座、分享最新的研究成果和行业趋势、激发学生的学习兴趣和探索精神等。引导教育是提升学生大模型素养的进阶工作,在学生培养好正确的、完整的大模型科学素养之后,应当开展合适的实践引导促进学生大模型素养的知行合一,如从实验室和项目方面展开驱动:建立大模型实验室,提供必要的硬件和软件支持,鼓励学生参与到大模型的研究和开发中去。通过项目驱动的学习方式,让学生在实践中深化对大模型的理解和应用。

5.2.2. “内部+外部”:促进教学模式创新性改变

教学模式的创新性改变离不开外部学校机构的创新改变与内部教师的鼎力支持,所以本文从内外部两方面开展叙述。

内部改变:教师作为教学活动的主体,其态度和方法直接决定教学效果。因此,若项促进大模型+高等教育的模式创新则教师应当积极改变教学理念,与创新的教学模式保持步调一致,积极接受和掌握大模型技术,理解其基本原理和应用方法。采取阅读相关文献、参加培训课程、亲自动手实践等方式方法以全面理解和熟练使用大模型。其次教师应当创新教学方法,学会利用大模型技术来改变传统教学模式,让教学活动变得更具有交互性。例如可以使用大模型来创建个性化的教学计划,根据每个学生的学习速度和难度来优化学习方案,提高教学效率和质量。

外部改变:学校机构具有创建一个有利于大模型赋能高等教育的环境的责任。应当制定一系列鼓励和支持大模型应用的政策和机制。例如提供技术支持和设备硬件的投入,提供培训和专项基金,以及鼓励教师进行尝试和创新的激励机制等。为了让大模型更好地服务高等教育,学校应进行教学模式的变革,包括教师评估制度的改革,促进教师更多地采用创新教学模式,以及学习评价体系的改革,强调学生的素质教育和创新能力。其次应当打造良好的学校氛围,宣扬积极正确使用大模型的思想观念。

5.2.3. “支持+审查”：保障大模型在教育领域的健康使用

大模型的发展应用都是一柄双刃剑，为保障大模型在教育领域的合理开展，我们要做到既要鼓励支持又要严格审查。

首先是支持创作者制作适合中国教育国情的专属大模型这就要求政府部门做好相关配套措施如：政策支持，鼓励国内大模型生产者，通过优惠政策，提供研发资金支持，开展专属于中国教育国情的大模型制作，并且加强对模型制作过程的指导，保证其科学性和可靠性。其次政府主导加强培养合作，推动学术界、产业界、政府之间深入合作，共享资源，集聚全社会之力，共同打造教育大模型。解除硬件限制保障设备更新，提供先进的计算设备支持，以保证大模型的训练和测试能够顺利进行。

其次对做好对大模型的应用的严格监管其中包括内容审查、伦理审查、审查制度等，第一要对大模型生成的内容进行严格审查，任何不符合社会主义核心价值观的内容应禁止在教学中使用。为此可以设立专项审查机制，比如人工智能伦理审查委员会，对内容进行预审核和事后抽查。其次对使用大模型的科研过程进行监管，确保科研活动得到良性开展。所有研究活动必须遵守科学伦理如尊重原创、避免学术不端等，一旦发现存在违反科研伦理的行为，应依法依规进行处理，并纳入个人和单位的信用档案。最后要制订严格的审查制度，让大众了解大模型应用的执行标准，让执法有法可依、有法必依以实现社会各方对大模型的监管与理解。

6. 总结与展望

本文以北京 xx 大学为调查对象，开展了有关大模型赋能高等教育的现状调查与问题对策的研究，根据调查问卷的结果分析在大学生视角下目前影响大模型赋能高等教育的重点因素，并给出相应的应对措施，对帮助大模型赋能高等教育、促进高等教育数字化转型升级提供有益指导。

数智时代下，大数据、人工智能（AI）以及大模型等技术得到广泛应用，教育领域也应紧跟时代脚步，积极探寻“大模型+教育”的科学方案（荀渊，2023）。目前大学生对大模型这一新鲜事物的兴趣程度颇高，大模型本身的机制也符合高等教育行业的本质需求。所以本文认为将大模型赋能高等教育既符合时代进步的正确举措，也是顺应国家政策走好高等教育数字化转型的必要之路。但科技的进步不仅仅只有利好的一面，作为教育领域的研究者，我们也要看到科技对教育所带来的风险与挑战。在将大模型融入高等教育的过程之中做到进步与稳定并存，创新与反思共进。若想达成高等教育模式方法的科学进步，只有促进学生、教师、学校、政府等多方的共同努力，从多角度、多方面开展积极合作才能营造中国高等教育更加辉煌的未来。本研究也存在一些不足之处如样本量较少、实验对象缺乏广泛性等，这些不足之处将在后续研究中将不断逐步完善。本研究期望大模型作为当前最有望改变人类社会的强大工具，经过不断地完善发展最终会与教育深度融合打造人人向往的高数字化、高智能化教育，为推进实现教育平等、终身教育提供可行方案。

参考文献

吴兰岸,闫寒冰,黄发良等. 大型语言模型在高等教育中的应用分析与现实挑战[J]. 现代教育技术，2023，33 (08)：29-37.

申丽萍,何朝帆,曹东旭等. 大语言模型在中学历史学科中的应用测评分析 [J]. 现代教育技术，2024，34 (02)：62-71

沈光辉,叶彤,孙小坚等. 贝叶斯修正的AHP—熵值法在教育测评中的应用 [J]. 心理学探新，2023，43 (03)：269-277.

苗勇胜. 新型“移动学习”教学模式探究——以高等教育阶段教学为例[J]. 中国教育信息化，2013，(07)：21-24.

荀渊. ChatGPT/生成式人工智能与高等教育的价值和使命 [J]. 华东师范大学学报(教育科学版)，2023，41 (07)：56-63.

尚智丛,闫禹宏. ChatGPT 教育应用及其带来的变革与伦理挑战[J]. 东北师大学报(哲学社会科学版)，2023，(05)：44-54.

面向计算思维培养的学生生成问题策略构建及应用

Construction and Application of Student-Generated Questions Strategy for Cultivating Computational Thinking

张雪[1]，牟歌[1]，万海鹏[1*]
[1]首都师范大学教育学院
[*] dnvhp@163.com

【摘要】 计算思维是中小学信息科技课程核心素养的重要组成部分。学生生成问题策略作为内在认知活动，旨在引导学生提取知识以生成问题，其本质是一种深度学习。本研究构建了面向计算思维的学生生成问题策略，以 80 名初中生为被试进行了实证研究，实验组采用学生生成问题策略开展编程学习，对照组采用传统编程学习策略。结果表明，较之于传统的编程环境，采用学生生成问题策略能够显著提升学生的计算思维能力，助力自主探究学习与知识建构。这一研究发现将为中小学利用生成问题策略开展编程学习，培养计算思维能力提供有效的实践经验。

【关键词】 学生生成问题策略；核心素养；计算思维；编程教育

Abstract: *Computational thinking is the core literacy of information technology courses in education. As a cognitive activity, students-generated questions aim to guide students to extract knowledge to generate questions. In this study, students-generated questions oriented to computational thinking was constructed, and 80 students were participated in the research. The experimental group adopted students-generated questions strategy to carry out programming learning, while the control group adopted traditional programming strategy. The results show that, compared with the traditional programming, the strategy can significantly improve students' computational thinking ability and help independent inquiry learning and knowledge construction. The findings will provide effective practical experience for primary and secondary schools to develop programming learning and cultivate computational thinking ability by using student-generated questions strategy.*

Keywords: Student-Generated Questions Strategy, Core Literacy, Computational Thinking, Programming Education

基金项目：科大讯飞横向委托课题:基础教育智能教育产品评估与应用研究（课题编号：iflytek2021-002）。

1.引言

核心素养是课程育人价值的集中体现，是学生通过课程学习逐步形成的正确价值观、必备品格和关键能力。信息时代，构建高水准的义务教育课程体系，聚焦学生核心素养的发展是落实立德树人根本任务，发展素质教育，推动教育现代化进程的重要抓手。信息科技课程所要培养的核心素养，包括信息意识、计算思维、数字化学习与创新、信息社会责任，四个方面相互渗透，而计算思维作为核心素养的基石，逐年受到关注。当前，在 K12 阶段培养学生的计算思维能力，大多研究聚焦于技术介入和游戏化学习，为此，有研究指出虽然快节奏的游戏化活动有利于提升学生的学习动机和学习表现，但可能不利于培养学生的思维能力(Zhao & Shute, 2019)，故本研究试图从教学策略的内部结构出发，探究培养学生计算思维能

力的信息科技编程课。

学生生成问题策略(Student-Generated Questions Strategy)作为培养学生高阶认知思维能力的教学策略，能促进学生提取知识与理解知识以达到解决问题及创新问题的目的。该策略的核心是学生“生成问题”，外在表现是提出问题、生成答案，重点关注学生的自主探究能力和创造力。事实上，该策略的本质是一种深度学习，通过改变学习者的认知路径和思维方式，触发知识的联结，进而促进思维方式的转变与成绩的提升。目前，针对学生生成问题策略的研究尚未成熟，诸多研究者担忧学生设计的问题质量以及生成问题的能力，尚未对学生生成问题策略有明晰的认识，故未曾系统地投入教学活动。研究学生生成问题策略，不仅能帮助研究者及教育工作者系统地认识该策略，助力完善教学策略的制定，也能帮助学习者开展元认知监控，理解问题解决的过程，提高计算思维能力。基于此，本研究尝试构建面向计算思维的学生生成问题策略并投入应用，旨在鼓励学生积极主动地与学习资源互动，促进理解的发生与知识的联结，进而促进核心素养的提升。

2.文献综述

2.1. 学生生成问题策略

学生生成的问题(Student-Generated Questions, SGQ)是学习者围绕相关学习内容，监控和评估知识理解和掌握水平而提出的系列问题(刘清堂等, 2023)，是学生在认知顺应与同化过程中动态生成疑问、困惑的过程，强调主体在发现问题、提出问题和解决问题过程中的内在认知活动和深度思维活动(曹梅&马悦,2020)。学生生成问题策略源于“交互式教学(Reciprocal teaching)”，教师需要对所教授的认知过程进行建模，然后在学生尝试设计并完成任务时为他们提供认知支持和指导，即提供脚手架支持(Palinscar & Brown, 1984)。在问题生成的过程中，学习者回忆学习内容，感知学习内容重难点，监控知识理解状态，体现自主探究能力及问题解决能力。

学生生成问题策略强调赋予学生参与知识建构的能力(Chen & Yeh, 2019)，关键因素为学习者先验知识、生成问题的形式、借助的媒介技术以及迭代评估过程。学习者的先验知识影响提问能力，而提问能力是理解能力的映射。提出问题，需要学习者掌握相关的学科知识概念，开展元认知活动，明晰思维过程，期间可借助引导性的学习支架或元素，由此降低生成问题的难度，提高生成问题的质量。在可视化编程语言平台中嵌入学生问题生成策略，能够提升学习者的思维能力(Ramirez & Monterola, 2022)，学习动机(Cheng et al., 2023)和学习自信(Hsu & Wang, 2018)。此外，依托技术平台及同伴反馈机制，发挥迭代评估功能也有助于促进学习者对课程内容的参与及知识的联结(Yu, 2022)。但是，也有研究关注学习者生成问题的质量(Arruabarrena et al., 2019)，担忧生成问题的过程可能会伴随高认知负荷的产生(Chen & Yeh, 2019)。为了减轻学生的认知负荷，有研究表示教师可提供足够的学习支架(Hwang et al., 2020; 刘清堂等，2023)，并依托技术平台降低生成问题的难度。诚然，学生生成问题的质量最初可能会受到限制，但随着学生对知识的积累，生成问题的质量会有所提高。总之，现有研究较为关注问题生成策略的显性效果，较少关注该策略的学习支架设计以及学习者深层思维能力的激发。

2.2. 基于编程教育的计算思维能力培养

计算思维是参与制定问题及其解决方案的思维过程，涉及将解决方案以可处理可执行的方式表示出来。信息科技课程要培养的计算思维，涉及抽象、分解、建模、算法设计等思维活动。近年来，在编程教育阶段，学界大多以数字化平台，可视化编程语言，游戏化案例为依托，借助实体编程或人工智能应用，来探究学生在编程学习中的计算思维提升路径。张学

军等学者在 Python 课程中，利用游戏化案例为学生提供明确的问题情境，由此培养高中生的计算思维(张学军等，2021)。王佑镁等以具身认知理论为基础，发挥实体编程的效能，引导学生进行编程与实践，进而促进计算思维的提升(王佑镁等，2021)。岳彦龙等人以人工智能为背景，鼓励学生学习人工智能案例内化 Python 编程知识，由此改善高中生的计算思维态度(岳彦龙等，2022)。李琪等人以编程活动为支撑，发挥协作编程的优势，在协作学习探究中表征计算思维(李琪等，2023)。陈兴冶等人以问题驱动学习者开展深度学习，以可视化编程工具助力学习者进行思维表达(陈兴冶等，2021)。在编程教育中培养学生的计算思维，不仅可以从模式、平台、系统设计等宏观视角探究计算思维的提升路径，也可以从计算思维与编程教育本体微观视角出发，细化计算思维能力，将计算思维的组成部分与编程教学活动过程相融合，引导学习者在问题解决中开展抽象、分解、建模、算法设计等思维活动，系统全面地提升计算思维能力。

3.面向计算思维的学生生成问题策略构建

为了在教学活动中系统地表征计算思维的核心组成部分，需将其分解并整合到解决问题的具体步骤。本研究依据 Anderson 提出的计算思维五大核心能力，将计算思维分为分解、模式识别、抽象、算法设计和评估(Anderson, 2016)。其中，分解：将复杂的问题或系统分解为可管理的部分或子问题；模式识别：识别问题中的重复结构、模式或规则；抽象：基于数据的规则和模式构建模型；算法设计：创建有序步骤来设计问题的解决方案；评估：测试和修改错误，以确保算法或特定模型的正确性和可行性。根据本研究的目的和内容，分解是指学习者在确定问题主题后，将其分解为若干个子问题；模式识别指识别子问题中的编程语句模块，找到可用于有效解决子问题的规则；抽象指根据子问题的规则，绘制简单的算法流程图；算法设计指根据流程图系统地设计算法；评估指调试提交后进行互评交流完善。

该策略分为训练阶段和实战阶段,如图 1 所示，训练阶段主要让学生们熟悉该策略的使用方法，特别是主动利用学习支架（学习资源及问题提示清单）的意识，这一阶段的任务由教师提出，是项目式学习的映射。第二阶段为学生利用教师提供的问题提示清单及学习资源（问题提示清单如图 2 所示），产生灵感并提出任务主题，依次进行分解主题、识别编程模块、绘制流程图、设计算法及调试评估等环节。

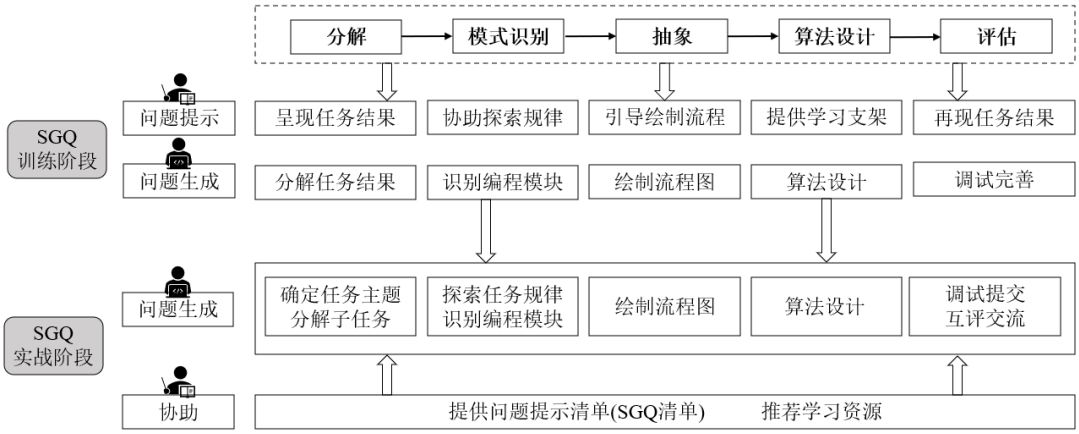


图 1 面向计算思维的学生生成问题策略

设计一个项目的基本步骤（小提示）
1.你提出的问题是什么？（你的主题）
2.为完成这个主题，你需要把它分解成几个小问题？分别是什么？
3.为实现每个小问题，你需要使用哪些编程语句？（赋值、顺序、选择、循环…）
4.这个主题的流程图大概是怎么样的？
5.恭喜你，可以开始编写！（提示：可以借助概念题、教材等学习资源）
6.调试成功后，思考可不可以再完善，使结果更有创意。

图 2 “问题生成”提示清单

4.研究设计

4.1. 研究对象

本实验的研究对象来自山东省济南市某中学的初一年级，采用准实验研究法，从四个班中随机挑选一个班为对照组，一个班为实验组。对照组共有 49 名学生，采用传统的教学策略，即教师讲解演示以及学生操练的形式；实验组共有 49 名学生，采用学生生成问题策略进行教学。实验组和对照组的学生均使用相同的教学方案。实验结束后，共回收有效数据 80 份，其中实验组回收 39 份，对照组回收 41 份，其余被试的数据因数据无效或未能完成测试而被舍弃。

4.2. 研究工具

计算思维试题。本实验利用试题的形式测试学生的计算思维能力水平，该试题为研究者及授课教师自编的计算思维试卷(A、B 卷)，试卷中的试题全部选自《Bebras 国际计算思维挑战赛试题》，因此该试题具有良好的区分度。两套试卷各包含 8 个情景题目，每个题 5 分，满分 40 分，均适用于初一学生，难度适中，所选的每个试题均覆盖分解、模式识别、算法设计和评估五个思维层面的两个及以上，题目示例如图 3 所示。

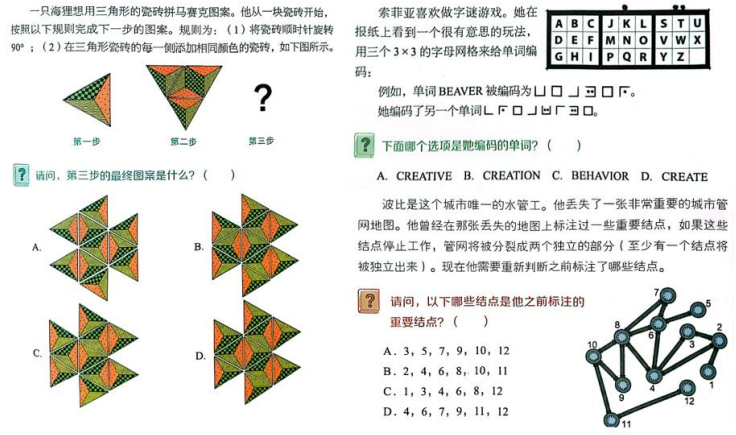


图 3 计算思维试卷试题示例

教学策略接受度问卷。本研究利用学习策略接受度量表评估学生对面向计算思维的学生生成问题策略的接受度，该量表是 Hwang, G 等人在提高学生学习表现的实验研究中使用的量表(Hwang et al., 2013)，分为感知有用性 6 题和感知易用性 7 题，本研究的笔者根据研究目的对该量表进行了微调，所有项目都使用 5 分李克特评分，很不同意(1)，不同意(2)，一般(3)，同意(4)和很同意(5)。

4.3. 实验流程

实验流程主要分为计算思维前测、学生生成问题策略 SGQ 训练阶段、学生生成问题 SGQ

实战阶段、计算思维后测四个阶段，如图 4 所示。

前测阶段：在正式课程开始之前，针对实验组，教师向学生介绍本次实验流程及“生成问题”提示清单的基本内容及使用规则，时间为 10 分钟。之后，实验组和对照组的学生分别独立作答计算思维测试卷 A 卷，时间为 35 分钟。该阶段的目的是测试两组学生的计算思维水平是否一致。

教学阶段：实验组和对照组的学生分别参与三个课时的学习，共 135 分钟，每个课时完成一个教学内容，分别是三大程序结构：顺序、选择和循环，由同一名教师授课。对于实验组而言，第一个课时是学生生成问题策略 SGQ 训练阶段，旨在训练学生熟悉“生成问题”的基本流程以及培养学生主动利用学习资源的意识；后两个课时为学生生成问题 SGQ 实战阶段，旨在激发学生借助教学支架设计问题并利用编程知识解决问题。对于对照组，三个课时均为教师讲解演示，学生练习反馈的模式。

后测阶段：在课程结束之后，实验组和对照组的学生作答计算思维测试卷 B 卷，实验组填写教学策略接受度问卷。

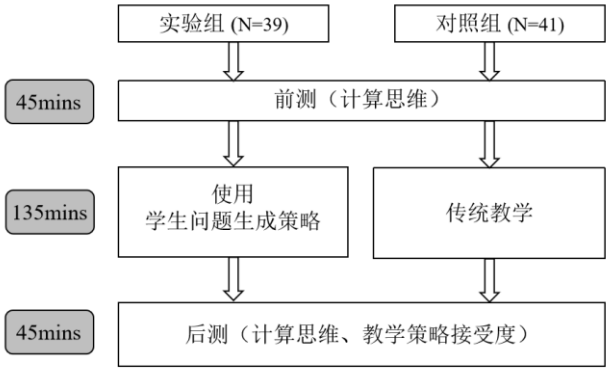


图 4 研究设计

5.研究结果

5.1. 实验组和对照组的计算思维对比

为探究实验组和对照组的计算思维水平差异，首先对两组分别进行正态性检验，显著性水平均低于 0.05，不满足正态分布，本研究利用 Mann-Whiteny 检验分析数据，结果如表 1 所示。在本实验前，实验组的秩平均值为 44.41，对照组的秩平均值为 36.78，两组的 Mann-Whiteny 检验显著性水平为 $p=0.125>0.05$，结果显示实验前实验组和对照组之间的计算思维并无明显差异，两组位于同一水平；在本实验后，实验组的秩平均值为 47.82，对照组的秩平均值为 33.54，两组的 Mann-Whiteny 检验显著性水平为 $p=0.005<0.05$，表明两组的计算思维水平存在显著差异，且实验组的计算思维水平显著高于对照组的计算思维水平，由此证明本研究构建的学生生成问题策略能显著提升学生的计算思维能力。

表 1 两组计算思维 Mann-Whitney 检验结果

		个案	秩平均值	Z	Sig.(双侧)
前测	实验组	39	44.41	-1.536	0.125
	对照组	41	36.78		
后测	实验组	39	47.82	-2.786	0.005**
	对照组	41	33.54		

5.2. 实验组和对照组的计算思维变化

为深入分析学生在实验前后计算思维的变化，本研究尝试利用 Wilcoxon 符号秩检验分析两组的计算思维数据，结果如表 2 所示。对于实验组，计算思维前后测 Wilcoxon 符号秩检验显著性水平 p=0.000<0.05，由此证明实验后实验组的计算思维水平有明显提升；对照组的计算思维前后测 Wilcoxon 符号秩检验显著性水平 p=0.004<0.05，证明对照组的计算思维水平也有一定地提升，但并无实验组显著。总之，无论是面向计算思维的学生生成问题策略还是传统的教学策略，都有利于培养学生的计算思维能力，但使用学生生成问题策略略胜一筹。

表 2 两组计算思维 Wilcoxon 检验结果

组别		个案	均值	标准误	Sig.(双侧)
实验组	前测	39	15.13	6.638	0.000**
	后测		23.08	9.004	
对照组	前测	41	13.66	5.003	0.004**
	后测		17.56	8.883	

注：*p<0.05, **p<0.01。

5.3. 实验组学生生成问题策略接受度

为了更好地了解学生对“生成问题策略”的感受，本研究收集了学生对“感知有用性”和“感知易用性”的反馈，“感知有用性”维度的克隆巴赫系数为 0.875，KMO 值为 0.891，“感知有用性”维度的克隆巴赫系数为 0.908，KMO 值为 0.868，故该量表有较好的信效度。根据表 3 可知，各个题项的均值都在 3.4 以上，故大多数学生对使用该策略的感受在两个维度都给出了积极的反馈，这意味着使用面向计算思维的学生生成问题策略让学生感受到了该策略确实能助力问题的提出与资源的利用，丰富自主探究学习活动。

表 3 实验组关于学生生成问题策略的感受(N=39)

	题项	均值	标准误
感知有用性	我觉得自己提出问题（主题）能让学习活动更丰富。	4	0.827
	我觉得自己提出问题对于我学习新知识很有帮助。	3.97	0.811
	使用学习资源和问题清单能让我提出问题的过程更为顺畅。	3.97	0.932
	课堂所提供的学习资源（灵感来源、可能用到的代码、提示等）可以帮助我在需要时获得有用的信息。	4.05	0.944
	自己提出问题可以让我学得更好。	3.95	0.857
	本次学习活动中，使用这样的学习方式（自己提出问题）比之前的计算机课更有效果。	3.49	0.97
感知易用性	使用问题清单，提出问题并不困难。	3.74	0.88
	使用问题清单提出问题时，并不需要花费过多的时间精力。	3.49	1.023
	我非常清楚，也非常理解提出问题的过程。	3.77	0.872
	我很快便了解了如何提出一个问题。	3.64	1.088
	本次学习活动中，使用“问题清单”及“学习资源”对我来说没什么困难。	3.77	0.902
	我觉得问题清单很容易使用。	3.9	0.821
	整体而言，借助问题清单提出问题是容易的。	3.95	0.826

6.讨论

培养学生计算思维能力的教学策略宏观层面分为三类，分别是基于游戏化理念的教学策略、基于问题解决理念的教学策略、基于可视化理念的教学策略(王国华&张立国，2018)。本研究基于问题解决的教学理念构建了学生生成问题策略，尊重学习者个体认知水平，鼓励学习者基于生活情境提出问题，其内在原理为促进学习者与自我认知及学习资源发生交互，触发知识的提取，进而解决自己提出的问题。本研究基于面向计算思维的学生生成问题策略，开展了实证研究，结果证明采用该策略确实能显著提升学生的计算思维能力，促进学生将学科知识与现实问题情境相联结，深度理解编程概念。相应的教学研究结果也表明，学生生成问题有助于鼓励学习者积极参与学习反思，产生知识联结，发挥学习者的主动性和主体性(张命华&刘电芝，2009)。此外，本研究的结果也显示了学生对于该策略的高接受度，从传统教学的限定问题到自主借助学习资源发现问题、提出问题、生成问题答案，学生逐渐接纳并熟悉自主探究学习模式，由此显示面对生成问题的高难度学习活动，学习者并非拒之千里。当然，如若欲将学生生成问题策略的教学效能最优化，在接下来的研究中还需斟酌该策略所用学习支架的内部结构与呈现序列，以期助力高阶问题生成，激发深度学习活动与知识联通。

7.结论与展望

本研究构建了面向计算思维的学生生成问题策略，并且以初一学生为被试进行了培养计算思维的实验研究，结果表明使用面向计算思维的学生生成问题策略能显著提升学生的计算思维能力。此外，对比传统“讲授练习”的编程学习，使用该策略更有利于培养学生的计算思维能力，帮助学生理解编程概念及提出问题解决问题的逻辑程序。当然，本研究也存在一定的局限性，一是统一为学生提供人工编制的问题提示 SGQ 清单，未能做到个性化；二是未能实时监测学生“生成问题”的障碍，及时为学生提供反馈。未来，还应继续完善学生生成问题策略，为学生提供自动生成的个性化情境化清单，采集学生使用学习资源及清单的过程数据，及时为学生提供适宜的脚手架，激发深度思维活动，全面提升学生的信息科技核心素养。

参考文献

曹梅 & 马悦.(2020).翻转课堂课前深层次学习的问题生成策略研究.**电化教育研究(11)**,101-107.doi:10.13811/j.cnki.eer.2020.11.014.

陈兴冶,马颖莹 & 杨伊.(2021).面向计算思维发展的深度学习模型建构——以可视化编程教学为例.**电化教育研究(05)**,94-100+121.doi:10.13811/j.cnki.eer.2021.05.013.

李琪,姜强 & 赵蔚.(2023).促进计算思维发展的编程投入机制研究：基于 ICAP 理论.**电化教育研究(08)**,88-95.doi:10.13811/j.cnki.eer.2023.08.012.

刘清堂,常瑀倍,吴林静 & 王春华.(2023).课前学习支架促进学生问题生成与认知水平提升.**电化教育研究(07)**,59-66.doi:10.13811/j.cnki.eer.2023.07.008.

王佑镁,宛平,南希烜 & 柳晨晨.(2021).实体编程促进计算思维发展：工具与策略.**中国电化教育(08)**,92-98.

岳彦龙,张学军 & 梁屿藩.(2022).人工智能教学如何培养高中生的计算思维?——基于人工智能案例驱动的 Python 编程教学的实证研究.**基础教育(01)**,74-84.

张立国 & 王国华.(2018).计算思维:信息技术学科核心素养培养的核心议题.**电化教育研究(05)**,115-121.doi:10.13811/j.cnki.eer.2018.05.017.

张命华 & 刘电芝. (2009). 培养远程学习者学习策略的教学思考. **中国远程教育, (01)**, pp.41-

44, 78-79.

张学军,岳彦龙 & 梁屿藩.(2021).Python 课程中数字化游戏教学培养高中生计算思维的实证研究.**电化教育研究(07)**,91-98.doi:10.13811/j.cnki.eer.2021.07.013.

Anderson, N. D. (2016). In Charlton S., Ryder K. and Taylor J.(Eds.), *A call for computational thinking in undergraduate psychology.* SAGE Publications. https://doi.org/10.1177/1475725716659252

Arruabarrena, R., Sánchez, A., Blanco, J. M., Vadillo, J. A., & Usandizaga, I. (2019). Integration of good practices of active methodologies with the reuse of student-generated content. *International Journal of Educational Technology in Higher Education, 16*(1), 1-20. https://doi.org/10.1186/s41239-019-0140-7

Cheng, Y., Lai, C., Chen, Y., Wang, W., Huang, Y., & Wu, T. (2023). Enhancing student's computational thinking skills with student-generated questions strategy in a game-based learning platform. *Computers and Education, 200*, 104794. https://doi.org/10.1016/j.compedu.2023.104794

Ching-Huei Chen & Hui-Chin Yeh (2019) Effects of integrating a questioning strategy with game-based learning on students' language learning performances in flipped classrooms, Technology, Pedagogy and Education, 28:3, 347-361, DOI: 10.1080/1475939X.2019.1618901

Fu-Yun Yu (2022) An online learning system supporting student-generated explanations for questions: design, development, and pedagogical potential, *Interactive Learning Environments, 30:5*, 782-802, DOI: 10.1080/10494820.2019.1686398

Hsu, C., & Wang, T. (2018). Applying game mechanics and student-generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Computers and Education, 121*, 73-88. https://doi.org/10.1016/j.compedu.2018.02.002

Hwang, G., Yang, L., & Wang, S. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers and Education, 69*, 121-130. https://doi.org/10.1016/j.compedu.2013.07.008

Hwang, G., Zou, D., & Lin, J. (2020). Effects of a multi-level concept mapping-based question-posing approach on students' ubiquitous learning performance and perceptions. *Computers and Education, 149*, 103815. https://doi.org/10.1016/j.compedu.2020.103815

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117-175. https://doi.org/10.1207/s1532690xci0102_1

Ramirez, H. J. M., & Monterola, S. L. C. (2022). Co-creating scripts in computer-supported collaborative learning and its effects on students' logical thinking in earth science. *Interactive Learning Environments, 30*(5), 908-921. https://doi.org/10.1080/10494820.2019.1702063

Zhao, W., & Shute, V. J. (2019). Can playing a video game foster computational thinking skills? *Computers and Education, 141*, 103633. https://doi.org/10.1016/j.compedu.2019.103633

基於深度學習的瑜珈自動評量與回饋系統對大學生技能學習表現之影響

Effect of a deep learning-based automated yoga assessment and feedback system on university students' skill acquisition performance

夏綠荷[1*]，黃國禎[2,3]，黃展鵬[4]，林彥男[5]
[1]勤益科技大學體育室
[2]台中教育大學教育資訊與測驗統計研究所
[3]台灣科技大學數位學習與教育研究所
[4]勤益科技大學資訊管理系
[5]嶺東科技大學體育室
[*] share.holder0130@gmail.com

【摘要】為提升學生的體育技能表現，提供個別化指導是非常重要的環節。然而學校體育普遍為大班制教學，每位學生能分配到的教師指導資源相當有限。為解決前述教學困境，學者嘗試應用人工智慧來提供更多專業且即時的個別化指導。但這個議題在實踐上，依然面臨著許多挑戰。因此，本研究基於反思實踐的教育理念，採用深度學習技術開發瑜珈自動評量與回饋系統，並分派 96 名學生使用此系統進行學習，另外 91 人採用一般線上學習系統。結果顯示，使用瑜珈自動評量與回饋系統，能顯著提升學生的瑜珈技能，並有促進反思、提升學習動機與獲得回饋等正面影響。
【關鍵字】 體育教育; 人工智慧; 反思實踐; 深度學習

Abstract: *To improve students' sports skills performance, it's important to individualized advice. However, in physical classes, it's almost impossible to provide detailed guidance to individual students. Scholars have trying to use Artificial Intelligence to provide personalized support to individual students; however, there's a lack of sufficient practices. This study develops a yoga automatic assessment and feedback system using artificial intelligence to provide personalized feedback to engage individual students in reflective practice, 96 students were adopted this system for learning, 91 students were adopted the general online learning system. The results showed that adopting the yoga automatic assessment and feedback system could significantly increase students' yoga skills performance, also have positive effects on promoting reflection, enhancing learning motivation, and obtaining feedback.*
Keywords: physical education, artificial intelligence, reflective practice, deep learning

1. 研究背景與動機

體育課程最常見的教學模式，是由教師進行運動技能的講解與演示，學生則透過模仿與反覆演練來習得技能 (Hung et al., 2018)。但每位學生的柔軟度、協調性、肌耐力等身體素質不盡相同；同時，對教師所講述內容的理解力，與接受能力也有所落差 (YanRu, 2021)。因此，為了確保落後的學生，可以跟上班級的群體學習進度；領先的學生，不會因為在課堂上感到無聊，而失去學習的興趣 (Chassignol et al., 2018)。體育教師普遍的挑戰是，需要盡自己最大的努力，提供學生個別化的指導 (Hung et al., 2018)。此外，在體育課程中，授課教師也會安排同儕回饋的學習活動。讓學生觀察同儕表現促進反思，並藉此獲取客觀的評價，以瞭解自己的技能水平與身體能力 (Trabelsi, 2020)。然而這樣的方式，學生需要在團體學習時才能執行。有學者即指出，相較於同儕回饋，使用人工智慧的體育課程，將能提供學生獨立學習時的輔助，例如：學生無需等待其他同學提供評價，即可透過人工智慧獲得客觀的回饋訊息，

更積極的進行體育活動，以提高自身的能力 (Lee & Lee, 2021)。因此，開發具有人工智慧的自動化回饋系統，除了能解決過去的教學困境，在課內、外提供更多專業且即時的個別化指導；也可以協助學生，在獨自練習時進行動作檢測，以了解自己的技能水平，進而展開精進表現的學習計畫。

目前已有不少學者關注，人工智慧應用於教育領域的實踐方式及理論基礎 (Ouyang & Jiao, 2021)。部分研究指出，AI 對學生的學習，產生了積極的助益 (Orucu & Selek, 2020; Wang, 2021; Yang et al., 2021)。然而，也有部分研究表明，相較於傳統教學，應用 AI 輔助的技能學習，並沒有顯著提升學生的學習表現 (Bonneton-Botté et al., 2020)。例如：相較於傳統的教學方法，提供學生 AI 分析練習結果的回饋訊息，並無法為各種程度的學童帶來學習優勢 (Bonneton-Botté et al., 2020)。顯見，AI 在提升學生的動作技能表現上，依然面臨著許多挑戰。當教育者無法確認 AI 的應用能帶來相對優勢時，AI 在教育領域的發展和應用都將受限 (Wang et al., 2020)。此外，目前多數的相關研究，並沒有指出可以應用哪些教育理論，來支持 AI 學習系統的可靠程度 (Deeva et al., 2021)。因此，需要更多納入教育觀點所設計的 AI 學習系統，以確保學生的學習效益，展現 AI 在教育領域上的有利價值 (Luckin & Cukurova, 2019)。

為了強化 AI 學習系統的可靠程度，本研究將以反思實踐的教育理念為基石，進行瑜珈自動評量與回饋系統的開發。反思實踐強調在實作練習時，引導學習者檢視並反思自己的表現，進而改善行動，從經驗中學習成長 (MacLean et al., 2019)。在過去已有不少研究支持，讓學生觀察自己或同儕的練習影像進行反思，並配合教師的指導，可以有效提升學生的運動技能表現 (Hung et al., 2018; Hsia & Hwang, 2020)。因此，本研究將引導學生觀察自己與同儕的練習，並應用 AI 提供與專家教師相仿的回饋指導，以引導學生進行反思。為檢驗本研究開發的瑜珈自動評量與回饋系統，對學生技能學習產生的影響，將針對下列研究問題進行探討：

(1) 使用 AI 瑜珈自動評量與回饋系統，或一般科技輔助學習模式，對學生的瑜珈技能表現是否產生差異？
(2) 使用 AI 瑜珈自動評量與回饋系統，或一般科技輔助學習模式，對學生的學習感受有何影響？

2. 文獻探討

2.1. 人工智慧在運動技能學習的應用

人工智慧技術的快速發展，為提供學生適性化學習，及改善學習體驗，增添了更多的可能 (Hsu et al., 2021; Salas-Pilco, 2020)。例如 Wang (2021) 建置了基於神經網絡的智能教學機器人，可以根據學生的學習狀態，制訂合適的教學策略。研究結果指出，學生的羽毛球技能表現、理論知識、學習興趣與學習熱情，都比傳統教學高出 20%左右。然而，也並非所有導入 AI 的智能教學模式，都能成功提升學生的動作技能表現。例如：Yang et al. (2020) 的研究結果即指出，AI 教育機器人可以提高學生對體育的學習興趣及態度，但無法提高學生的技能表現。Bonneton-Botté et al. (2020) 讓學齡前的幼兒在平板電腦上進行手寫運動，並透過 AI 分析以提供練習結果的回饋。該研究結果指出，相較於傳統的教學方法，AI 促進的學習優勢，僅在於起始點為中等水平的孩童身上。這也意謂著，僅提供學生練習結果的回饋訊息，並無法幫助不同程度的學生，反思自己的實踐。由此可知，AI 在動作技能學習領域中，尚有一些挑戰需要被克服。

為了強化學生的動作技能學習表現，建構更可靠的學習模式。有學者提倡，將反思引導策略導入課程設計，引導學生從自己的實踐中進行反思，進而達到反思能力與技能表現共同成長的雙重成果 (Lin et al., 2022; Samanian & Roohani, 2018)。而其中值得關注的是，要成功促進學生反思能力的成長，需要引導學生從客觀的角度審視自己，以及從多元的角度進行自我評估 (Hsia & Hwang, 2020)。在教學活動的規劃上，通常是透過教師提供學生個別化的回饋指導，及安排同儕觀摩與回饋的學習活動；以藉此幫助學生獲取客觀的評價，學習評估自己的

學習過程及表現，進而提升自我反思的效益 (Lin et al., 2020)。因此，本研究將納入反思實踐的教育觀點，進行 AI 瑜珈自動評量與回饋系統的開發，用以建構更可靠的學習效益。

2.2. 人工智慧在瑜珈的應用

瑜珈運動對動作的專業性與準確度要求極高，是需要應用人體姿勢辨識技術的領域之一。在近年的瑜伽姿勢辨識研究中，使用深度學習的方式是主流；主要是透過 OpenPose 來提取人體姿勢中的特徵，再將提取出的特徵，以不同的人工智慧技術，針對專業的瑜珈姿勢進行模型訓練，目前測試的準確程度可達 99%左右 (Yadav et al., 2019; Jain et al., 2021)。

然而，目前多數的相關研究，大多聚焦於驗證辨識技術效能的改善。但對運動技能學習領域而言，提供動作的正確度與專業回饋指導之建議，更是促進技能成長的關鍵 (Orucu & Selek, 2020)。另一方面，在 Covid-19 流行期間，教育當局並不鼓勵教師與學生進行接觸指導。因此，如何在師生沒有接觸的情況之下，提供學生正確的動作分析結果與專業回饋，以幫助他們在居家練習時，也能正確的學習瑜伽動作，亦是當前十分重要的教育議題。基於此需求，本研究將藉由 OpenPose 提取人體姿勢關鍵點，再透過知識擷取，建立出資深瑜珈教師，對不同瑜伽動作所制定的規則及回饋建議。最後，以此動作規則，開發出基於深度學習的瑜珈自動評量與回饋系統。用以提供學生在學習瑜珈技能時，可以主動獲得適性化的動作分析結果及回饋指導。

3. 瑜珈自動評量與回饋系統

本研究參考 Yadav et al. (2019) 及 Jain et al. (2021) 之建議，採用 OpenPose 作為主要的人體動作辨識之核心技術。它是以深度學習技術所建立的，人體姿勢辨識開源演算法程式庫。是由 CMU Perceptual Computing Lab (2017) 所提出一套以 C++撰寫，並使用 OpenCV 和 Caffe 的多人關鍵點檢測系統。它在近年已經取得很多學術上的應用，在人體姿勢辨識上，取得很高的準確率及良好效能 (Jain et al., 2021)。

系統採用響應式網頁設計，架構如圖 1 所示。使用者於前端，使用各式數位學習平台之功能進行學習。其中，最關鍵的獲取自動評量與回饋功能，係由學生上傳自己的瑜珈動作練習照片，系統於後端進行特徵擷取、CuDNN 分析運算與 OpenPose 標記關鍵點資訊。在將產出的 JSON 資料，透過專家知識庫規則進行分析。最後，再將分析結果回傳至使用者介面，呈現出該次練習的人體骨架視覺回饋、評分結果與分析指標等資訊。

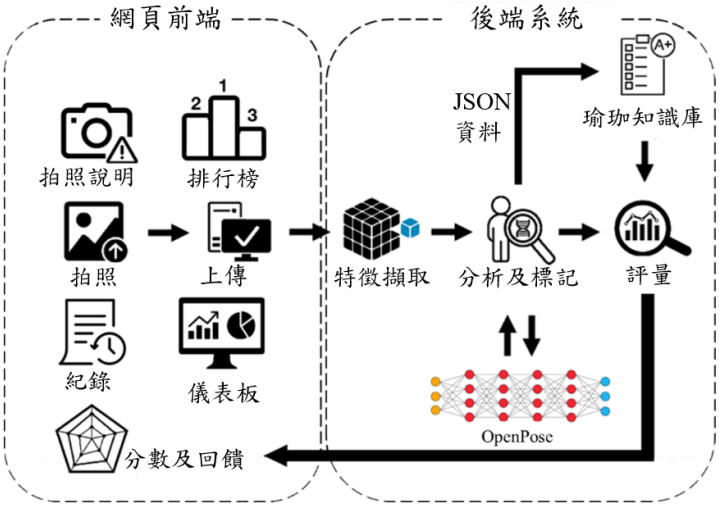


圖 1 系統架構圖

3.1. 專家知識庫規則

本研究邀請三位具備合格瑜珈教師證，及 10 年以上瑜珈教學經驗之教師組成專家小組。針對五個瑜珈姿勢，建構出各個姿勢正確性的分析指標。依專家知識庫之規則，在評量瑜珈姿勢的正確程度時，僅取用各動作所需的分析指標，與其對應的骨架關鍵點資料進行分析。

知識庫建立後，三位瑜珈教師之標示資料與系統標示之資料，Kendall's W 介於 0.707~0.881 之間。顯示系統評分與教師評分之間，具有高度的一致性。

3.2. 學習任務與系統介紹

學生在登入本研究所建立的 AI 瑜珈自動評量與回饋系統後，可先於「瑜珈姿勢」中，觀看各個姿勢的教師示範與重點說明。接著使用「上傳照片」功能，將練習的照片上傳至網頁伺服器，以獲得即時的分析結果。上傳照片經 OpenPose 分析標記，產生骨架資料以及 JSON 關鍵點檔案，並依專家知識所建立之評分規準，針對每一個分析指標進行評分。評分級距為最高 3 分至最低 1 分。隨後回傳給學生 OpenPose 所標記的人體骨架、評分結果、分析指標與綜合表現雷達圖。學生在接受到自己動作表現的分析結果後，可選擇再次觀看正確示範，或進入下一個動作的練習，如圖 2。



圖 2 學生接收分析結果介面

4. 實驗設計

4.1. 參與者

實驗的參與者為大專通識瑜珈課程之修課學生，共 4 個班級 187 人。年齡介於 18~19 歲之間，多數為首次接觸瑜珈之初學者。隨機分配 2 個班級 96 人為實驗組，使用本研究開發的瑜珈自動評量與回饋系統輔助學習；另外 2 個班級 91 人為控制組，採用一般科技輔助學習模式。由同一位具備 10 年以上教學經驗之教師進行授課。

4.2. 實驗流程

教學實驗為期 11 週，期間每週授課 1 次，每次進行 100 分鐘。前 4 週，教師會講解學習內容、任務與測驗方法。同時，採用一般傳統教學模式，教授一套入門的瑜珈動作組合。隨後，於第 5 週進行瑜珈技能表現之前測，以瞭解學生的起始點。

第 6-9 週，在教學活動開始之前，教師會利用 20 分鐘，講解學習系統的相關功能與使用方法。隨後，再進行第二套瑜珈動作組合之教學。在教學期間，每堂課的最後 30 分鐘，學生需進行自主練習，並拍下當週的練習影像回傳至學習系統。兩組學生之差異在於，採用不同的學習系統輔助學習。在實驗組的部分，學生可藉由觀看人體骨架所提供的視覺回饋，了解該瑜珈姿勢之正確性，並藉由評分結果、分析指標與綜合表現雷達圖，反思學習表現。控制組的部分，系統會提供教師的標準示範。在學生上傳實作練習影像後，告知是否上傳成功，並且於兩週內傳送教師批改的評分結果。相較於實驗組，控制組的線上學習系統，僅提供學生個人的學習空間；缺少同儕觀摩、同儕賞析等，支持學生從不同角度自我反思的群體學習

活動。第 10 週為瑜珈技能表現之後測，並於隔週進行半結構式的團體小組焦點訪談。

4.3. 研究工具

瑜珈技能表現之評分規準，由 2 位具備 7 年以上瑜珈教學經驗之教師，組成評分團隊。在考量本次的課程內容後，借鑑 Hsia and Hwang (2020) 提出的舞步流程、表現技巧與動作技巧等三個評分面向為基礎，再微調評分細則以符合本研究之情境。學生在參與前、後測之術科考試時，針對整套的動作組合，會分別獲得 1 個得分。在動作技巧的部分，則是每個單獨的瑜珈動作都會獲得 1 個評價。最所有得分的平均值為最終得分。2 位教師評分結果之 Pearson 相關係數前測為 0.83，後測為 0.77 (p< .000)，顯示評分者間具有高度的一致性。因此，以 2 位教師給分的平均數，作為學生的瑜珈技能表現得分。

訪談題目採用 Hwang, Yang, Tsai and Yang (2009) 針對資訊科技策略融入教學所編制的提問大綱，總計 7 題。再依據本研究導入的科技策略調整用詞。全體學生隨機分配為 5-6 人的小組，每組進行 10 分鐘左右的半結構式團體小組焦點訪談。訪談期間全程錄音，並打為逐字稿。實驗組代號"EG"、控制組代號"CG"，並依序紀錄訪談組別與發言出現的時間。訪談內容由兩位博士級研究人員進行分析，先整理出與學習有關之內容進行編碼，並且記錄下代碼出現的次數；最後，再將相似的代碼，歸納到同一個主題中 (Creswell, 2013)。當歸納意見出現分歧時，研究人員會透過會議從新審視內容並進行討論，以達成分類歸納之共識。

5. 結果

5.1. 瑜珈技能表現之分析

瑜珈技能表現之成績，經迴歸係數同質性考驗後，得出 F 值未達顯著差異 (F=2.94，p=.08>.05)。顯示 2 組學生的測驗成績，並未違反同質性之假定，可直接進行共變數分析，以排除先備技能對實驗結果產生影響。如表 1 所示，在排除前測成績的影響效果後，2 組學生的瑜珈技能表現達顯著差異 (F=4.60，p=.03< .05)。此外，η2=0.02，顯示為小效果量 (Cohen, 1988)。實驗組的調整後平均數為 3.12，而控制組則為 3.04。顯見，使用自動評量與回饋系統進行學習，能顯著提升學生的瑜珈技能表現。

表 1 瑜珈技能表現之共變數分析摘要表

組別	人數	平均數	標準差	調整後平均數	標準誤	F	η2
實驗組	96	3.12	0.31	3.13	0.03	4.60*	0.02
控制組	91	3.04	0.29	3.04	0.03		

*p <.05

5.2. 瑜珈自動評量與回饋系統帶來的學習優勢

表 2 分析學生訪談內容後，歸納出的學習優勢。其中，最常被學生提及的包含：促進反思、提升學習動機與獲得回饋。在促進反思的部分，兩組學生都有提出相關的歷程，兩組之間的差異則在於，實驗組的學生表示："透過系統學習，會比較標準。就是去看照片上的骨架圖，紅色、藍色的那些線條，可以知道動作的角度夠不夠延展等" (EG06037) ；" 系統它會提供不同面向的評分，你可以針對小區域的問題去做調整" (EG09426) ；可以去看前五名同學的動作，多一個參考，比較一下自己跟別人的差距 (EG09410)。顯見，實驗組學生可以經由瑜珈自動評量與回饋系統上的人體骨架標示、評分功能與前五名同學的實作照片，獲得更多的客觀訊息，以從不同的角度進行反思。

在提升學習動機的部分，兩組之間的差異則在於，實驗組學生表示：" 這個系統會讓我想做得更標準，就是打遊戲的感覺，有競爭的心態，會想把整體動作做得更好" (EG03235)。顯見，瑜珈自動評量與回饋系統提供的分數回饋、每次上傳的得分歷程記錄與排名機制，除了可以幫助實驗組學生了解自己進步的幅度，也能激發學生想挑戰更高的分數與排名，並從中

獲得成就感，進而促進學習動機的提升。

在獲得回饋的部分，兩組學生皆能從自己的實作練習影像，及教師的回饋指導中，獲取自己動作表現的訊息。實驗組學生還提報出更多，獲得立即回饋的歷程，例如：”很方便就可以知道自己的動作確不確實，可以立馬改善，不用等老師的輔助” (EG02348)。意謂著，瑜珈自動評量與回饋系統可以填補，學生在尚未獲得教師回饋時的空窗期，有助於即時修正動作表現。另一方面，有學生表示：”比較公正吧！有些老師比較喜歡某些學生，會影響評分，用這個機器去判斷標準比較客觀”(EG05138)。顯見，瑜珈自動評量與回饋系統的評分結果，讓實驗組學生覺得客觀、可信。

表 2 學習優勢分析摘要表

主題	代碼	提及次數	
		實驗組	控制組
促進反思	透過自我觀察反思自己的動作是否正確	12	17
	觀察同儕表現反思自己的動作是否標準	13	4
	從不同的角度進行反思	8	0
	透過系統的評分進行反思	21	0
	透過系統提供的人體骨架進行反思	5	0
提升學習動機	為了提升排名而激發學習的動力	30	0
	為了提升成績而增加練習的次數	25	2
	覺得新奇、特別、好玩	6	2
	為了提升動作品質而增加練習次數	18	13
獲得回饋	獲得即時的回饋	22	1
	知道自己的缺點	21	10
	獲得公平客觀的回饋	4	0
	知道多數人的問題	0	3

6. 結論與建議

為了提升體育課程的教學品質，本研究提出了基於反思實踐所開發的瑜珈自動評量與回饋系統，並實踐於大學的教學場域中。探討應用 AI 技術輔助的體育課程，對學生技能學習表現的影響。在經過 4 週的教學實驗介入後，使用瑜珈自動評量與回饋系統的實驗組，在瑜珈技能表現的部分，顯著優於使用一般學習平台輔助的控制組。意謂著，本研究引導學生觀察自己與同儕的練習影像，並藉由 AI 提供專業的即時評分，與人體骨架的視覺回饋，成功促進了學生的技能成長。這樣的發現，除了支持先前關於引導學生藉由自我觀察、反思，進而提升技能表現的研究成果 (Hung et al., 2018; Hsia & Hwang, 2020; Lin et al., 2020; MacLean et al., 2019)。更重要的是，證明了導入 AI 提供即時回饋的功能，能使學生在技能學習上獲得更好的學習成效 (Orucu & Selek, 2020; Yang et al., 2021)。從而為未來，想應用 AI 於技能學習上的相關研究做出貢獻。

其次，為了更深入的瞭解，學生是如何從學習系統中獲得幫助。本研究對所有參與的學生進行訪談，歸納出使用瑜珈自動評量與回饋系統能支持學習的三大優勢，包含：促進反思、提升學習動機與獲得回饋。在促進反思的部分，由於實驗組與控制組的學生，都有進行拍攝練習影像，與自我檢視的學習活動。因此，在兩組學生的訪談中，都有提報出關於反思的學習歷程。但實驗組的學生，提出了更多關於反思的體驗。同時，明確指出由 AI 技術支援的即時分析與回饋，能幫助他們從專家的評分角度，及人體骨架標示所提供的肢體角度等，不同的面向進行反思。誠如先前的研究結果顯示，學生僅透過觀察自己的練習影像，是很難有效反思教師所提出的概念性問題 (Nunes et al., 2020)。因此，導入 AI 輔助學生從更高層次的觀

點，來檢視自己的練習經驗，成功激起了更多的反思行為。

在提升學習動機的部分，兩組學生都表示，會為了提高動作品質或成績，更努力的進行練習。而其中值得關注的是，瑜珈自動評量與回饋系統中的分數回饋、得分紀錄與排行榜功能，除了有助於學生在自主學習時，瞭解自己的進步幅度，也促進了實驗組的學生，為了追求更高的成績表現與排名，進而更積極的投入練習。驗證了先前的研究結果，AI 的評測與回饋功能，可以促進學生持續與系統進行互動，產生重複練習的動力 (Lee & Lee, 2021; Yang et al., 2021)。

在獲得回饋的部分，兩組學生都談論到，能從觀察自己的練習影像與教師指導中，瞭解自己的表現。而實驗組的學生，對 AI 提供回饋的即時性、客觀程度與對學習產生的助益，更是讚譽有加。呼應了 Lee and Lee (2021) 提出的觀點，當學習者透過 AI 接收到客觀的回饋訊息後，可以立即展開積極的學習計畫以提高自身能力，而不必等待課堂上其他同儕的觀點或教師評價。此外，過去研究指出，多數學生認同在學習系統中導入 AI 技術，提供適性化的回饋建議，可以改善學習的質量，提供學生相似與教師指導的學習輔助 (Troussas et al.,2020)。綜觀本研究的質性與量化成果，可以發現這樣的論點，在本研究中也得到驗證。

最後，在延伸研究的部分。本研究展示了學生反思，與學習表現之間的關係。建議後續的研究，可以在學習系統後台記錄學生的線上學習行為，並進一步分析反思、學習表現及學習行為之間的關係。整理出更多具有教育意涵的訊息，以支持 AI 學習系統的設計，及探討 AI 應用於教育領域時，所帶來的學習效益。在系統開發的部分，過去在瑜珈姿勢辨識的相關研究中，有使用物聯網、感測設備，也有利用多位使用者的影片或相片，所訓練的深度學習模型。絕大多數的研究，皆是朝增加更多種姿勢的判斷，或是提升數據集判定的準確率為主。這為瑜珈姿勢辨識，提供很好的技術基礎，但並不全然符合，教育現場的學習者需求。為此，本研究藉由 OpenPose 辨識出肢體的關鍵點，開發一套以教育為目標的學習系統。從技術的角度來看，此系統是建立在專家的知識基礎上進行判定。其驗證方式，是採取和專家的判定，進行一致性比較。本研究所關注的焦點，並非改進 OpenPose 的準確度，而是以貼近教學需求為主軸，進行系統開發。以技術創新的角度來看，我們跳脫過往較單純的準確率提升框架。以合理的方式，在瑜珈姿勢判斷上，更多的融入人的因素。以符合讓學習者能收到，較貼近教師建議的回饋訊息，而不僅是動作的準確率或分數而已。

參考文獻

Bonneton-Botté, N., Fleury, S., Girard, N., Le Magadou, M., Cherbonnier, A., Renault, M., Anquetil, E., & Jamet, E. (2020). Can tablet apps support the learning of handwriting? An investigation of learning outcomes in kindergarten classroom. *Computers & Education, 151*, 103831.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science, 136*, 16-24.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerdt, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education, 162*, 104094.

Hsia, L. H. & Hwang, G. J. (2020). From reflective thinking to learning engagement awareness: A reflective thinking promoting approach to improve students' dance performance, self-efficacy and task load in flipped learning. *British Journal of Educational Technology, 51*(6), 2461-2477.

Hsu, T. C., Abelson, H., Lao, N., Tseng, Y. H., & Lin, Y. T. (2021). Behavioral-pattern exploration and development of an instructional tool for young children to learn AI. *Computers and Education: Artificial Intelligence, 2*, 100012.

Hung, H. C., Young, S. S.-C., & Lin, K. C. (2018). Exploring the effects of integrating the iPad to

improve students' motivation and badminton skills: a WISER model for physical education. *Technology, Pedagogy and Education, 27*(3), 265-278.
Hwang, G. J., Yang, T. C., Tsai, C. C., & Yang, Stephen J. H. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & Education, 53*(2), 402-413.
Jain, S., Rustagi, A., Saurav, S., Saini, R., & Singh, S. (2021). Three-dimensional CNN-inspired deep learning architecture for Yoga pose recognition in the real-world environment. *Neural Computing and Applications, 33*(12), 6427-6441.
Lee, H. S., & Lee, J. (2021). Applying Artificial Intelligence in Physical Education and Future Perspectives. *Sustainability, 13*(1), 351.
Lin, K.-C., Lee, I.-C., Cheng, C.-F., & Hung, H.-C. (2020). The Effects of Adopting Tablets and Facebook for Learning Badminton Skills: A Portfolio-Based WISER Model in Physical Education. *Educational Technology & Society, 23* (4), 89-105.
Lin, Y. N., Hsia, L. H., & Hwang, G. J. (2022). Fostering motor skills in physical education: A mobile technology-supported ICRA flipped learning model. *Computers & Education, 177*, 104380.
Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology, 50*(6), 2824-2838.
MacLean, S., Geddes, F., Kelly, M., & Della, P. (2019). Video Reflection in Discharge Communication Skills Training With Simulated Patients: A Qualitative Study of Nursing Students' Perceptions. *Clinical Simulation in Nursing, 28*, 15-24.
Nunes, M. E. D. S., Correa, U. C., de Souza, M. G. T. X., & Santos, S. (2020). Descriptive versus prescriptive feedback in the learning of golf putting by older persons. *International Journal of Sport and Exercise Psychology*. https://doi.org/10.1080/1612197X.2020.1717579
Orucu, S., & Selek, M. (2020). Design and validation of rule-based expert system by using kinect V2 for real-time athlete support. *Applied Sciences, 10*(2), 611.
Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2*, 100020.
Salas-Pilco, S. Z. (2020). The impact of AI and robotics on physical, social-emotional and intellectual learning outcomes: An integrated analytical framework. *British Journal of Educational Technology, 51*(5), 1808-1825.
Samanian, S., & Roohani, A. (2018). Effects of self-regulatory strategy development on EFL learners' descriptive writing and reflective thinking. Research in English Language Pedagogy, 6(1), 95-116.
Troussas, C., Krouska, A., & Sgouropoulou, C. (2020). Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Computers & Education, 144*, 103698.
Trabelsi, O., Gharbi, A., Masmoudi, L., & Mrayeh, M. (2020). Enhancing female adolescents' engagement in Physical Education classes through video-based peer feedback. *Acta Gymnica, 50*(3), 93-104.
Wang, P. (2021). Modeling of Badminton Intelligent Teaching System Based on Neural Network. *Wireless Communications and Mobile Computing*, 9933285.
Yadav, S. K., Singh, A., Gupta, A., & Raheja, J. L. (2019). Real-time Yoga recognition using deep learning. *Neural Computing and Applications, 31*(12), 9349-9361.
Yang, A. C. M., Chen, I. Y. L., Flanagan, B., & Ogata, H. (2021). Automatic Generation of Cloze Items for Repeated Testing to Improve Reading Comprehension. *Educational Technology & Society, 24* (3), 147–158.
YanRu, L. (2021) An artificial intelligence and machine vision based evaluation of physical education teaching. *Journal of Intelligent & Fuzzy Systems, 40*(2), 3559-3569.
Yang, D., Oh, E. S., & Wang, Y. (2020). Hybrid Physical Education Teaching and Curriculum Design Based on a Voice Interactive Artificial Intelligence Educational Robot. *Sustainability, 12*(19), 8000.

深度學習增強藝術教育：運用人工智慧提升大學生藝術品鑑賞能力與繪畫技巧

Enhanced Art Education through Deep Learning: Enhancing University Students' Ability to Appreciate Artworks and Painting Skills with Artificial Intelligence

邱敏棋 [1*]，黃國禎 [2]，夏綠荷 [3]，李建輝 [4]
[1] 嶺東科技大學資訊管理系
[2] 台中教育大學教育資訊與測驗統計研究所
及
台灣科技大學數位學習與教育研究所
[3] 勤益科技大學體育室
[4] 嶺東科技大學智慧製造科技系
[*] minky413@gmail.com

【摘要】在傳統的藝術課程中，教師根據學生的學習狀態為個別學生提供反饋和指導是非常重要的，因為誤解可能會影響他們在後續課程單元中的學習參與和表現。然而，一名教師通常需要指導數十名學生，因此對教師來說，在沒有任何輔助的情況下立即向學生提供反饋是具有挑戰性的。人工智能(AI)的進步為解決這一問題提供了可能的解決方案。在本研究中，我們開發了一個基於深度學習的藝術學習系統(DL-ALS)，通過使用微調的 ResNet50 模型來幫助學生識別和分類藝術作品。我們的目標是培養學生準確的欣賞知識和藝術創作能力，並利用 AI 技術提供即時反饋和個性化指導。為了探索這一系統的效果，我們在一所大學的藝術作品欣賞課程中實施了一個準實驗。共有 46 名在選修藝術課程的大學生參與了這項研究，他們來自兩個班級。其中一個班級是實驗組，採用 DL-ALS 學習，另一個則是控制組，採用傳統技術支持的藝術學習(CT-AL)。結果顯示，與 CT-AL 相比，通過 DL-ALS 進行學習可以促進學生的學習成就、學習態度、學習動機以及在藝術課程中的表現。

【關鍵词】 深度學習、人工智慧、藝術學習、藝術品欣賞、藝術品創作

Abstract: *In traditional art courses, it is very important for teachers to provide feedback and guidance to individual students based on their learning status, as misunderstandings can affect their participation and performance in subsequent course units. However, a teacher often needs to guide dozens of students, making it challenging to provide immediate feedback to students without any assistance. The advancement of Artificial Intelligence (AI) offers a potential solution to this problem. In this study, we developed a Deep Learning-based Art Learning System (DL-ALS) that assists students in identifying and classifying artworks using a fine-tuned ResNet50 model. Our goal is to cultivate students' accurate appreciation knowledge and artistic creation abilities and to provide real-time feedback and personalized guidance using AI technology. To explore the effectiveness of this system, we implemented a quasi-experiment in an art appreciation course at a university. A total of 46 university students participating in elective art courses from two classes were involved in this study. One class was the experimental group using DL-ALS for learning, while the other was the control group using traditional technology-supported art learning (CT-AL). The results show that learning*

through DL-ALS promotes students' learning achievements, learning attitudes, motivation, and performance in art courses, compared to CT-AL.

Keywords: Deep learning, artificial intelligence, art learning, art appreciation, art creation

1. 前言

傳統藝術教育中，學生需不斷練習並依賴課堂指導，但教師面對多位學生時，難以提供即時個別化回饋(Hayadi et al., 2018)，且學生學習能力差異增加教學負擔(Moubayed et al., 2020)。學生對課堂內容的理解程度不一致，導致學習困難(Emmers et al., 2020)。為解決這些問題，West(2012)開發線上學習系統，分析學生學習方法，協助教師提供適當指導，克服時間和距離障礙，滿足學生個別需求(Alenezi & Karim, 2010)。線上學習系統提高學生學習成就和動機(Chang et al., 2020)。隨著 AI 的發展，研究顯示其在提供個性化教學中具潛力(Popenici & Kerr, 2017)，已應用於醫學(Chang et al., 2021)、生物(Koć-Januchta et al., 2020)和語言教育(Nazari et al., 2021)。然而，AI 在藝術教育的應用尚待開發。本研究開發基於深度學習的藝術學習系統(DL-ALS)，學生可從專家作品學習並獲得專業反饋，用於藝術課程，提供動畫反饋激勵持續創作。本研究旨在探討 DL-ALS 對學生學習的影響，相關研究問題如下:

(1) 與傳統技術支持的藝術學習(conventional technology-supported art learning, CT-AL)相比，深度學習藝術學習系統(DL-ALS)能否提升學生的學習成就?
(2) 與 CT-AL 相比，DL-ALS 學習能否改善學生的學習態度?
(3) 與 CT-AL 相比，DL-ALS 學習能否增強學生的學習動機?
(4) 與 CT-AL 相比，DL-ALS 學習能否提高學生的繪畫表現?

2. 人工智慧在藝術及教育上的發展及應用

人工智慧(AI)賦予電腦類人的智能工作能力，包括推理、決策等(Lu, 2019)。AI 的核心技術，機器學習使電腦從數據中學習(Chen et al., 2020a) 。其中，深度學習顯著提升了 AI 的推理能力，成功應用於象棋、圖像識別、自然語言處理等領域(Zhang, 2019)。AI 在教育中日益重要，研究者開發了利用 AI 的學習環境，提供個性化學習支援(Stray, 2019)。然而，AI 在藝術教學領域的應用仍不普及(Kong, 2020)。藝術作品分析需專業知識，如理解作品的元素和特徵(Barnet, 1985)。儘管 AI 在藝術教育中具潛力，先前研究嘗試開發用於藝術品識別的 AI 系統(Tan et al., 2016)，如使用貝葉斯分類器(Keren, 2002)、卷積神經網絡(CNN)識別藝術品風格(Karayev, 2013)和 CNN 的藝術作品分類(Cetinic et al., 2018)。這些研究聚焦於 AI 識別藝術品，較少涉及輔助學習。本研究旨在開發深度學習藝術學習系統(DL-ALS)，探討在藝術教育的應用及學習成效。

3. 深度學習的藝術學習系統開發(DL-ALS)

基於深度學習的藝術學習系統包含三個步驟：首先，教師將藝術課程中的六個類別作品用於深度學習模型的離線訓練和測試。接著，學生使用電腦繪圖軟體創作藝術作品並上傳至模型進行分析。最後，模型分析結果將即時回傳給學生。若作品高度契合某藝術類別(達 90%以上相似度)，系統會展示煙火動畫以鼓勵學生。反之，若作品難以歸類，則出現雨天動畫，提示學生需要更多練習和改進。

3.1. 深度學習

深度學習是人工智慧的關鍵子領域，它模仿人腦運作模式，能從數據集自動提取有用的功能(Panch et al., 2018)。它主要應用類神經網絡來定義函數，其中卷積神經網絡(CNNs）在圖像分類等複雜機器視覺問題上表現優越，甚至超越人類(He et al., 2015)。CNNs 已被廣泛應用於物體檢測、人臉識別等領域(Guo et al., 2019)。ResNet 是一種 CNN 架構，以其多層次設計在圖像數據處理中表現卓越，勝過其他模型（Virnodkar et al., 2020)。它比 VGG 網絡具有更低複雜度和更高準確度，解決了退化問題(Theckedath & Sedamkar, 2020)。ResNet50 因其高可靠性和準確性，在醫學、農業及食品識別等領域得到廣泛應用(Farooq & Hafeez, 2020)。本研究通過 Fine-tune ResNet50 和遷移學習技術，使用 ImageNet 資料集的預訓練權重，增強模型泛化能力。本研究使用 WikiArt 開源資料集蒐集六個藝術類別共 323 張作品圖像。透過資料擴充技術，如平移轉換、縮放(0-40%)、亮度調整(10-30%)，擴充至 1292 張圖像，並將資料集分為訓練集(90%）和驗證集(10%)。DL-ALS 系統前端採用響應式網頁設計，適應不同裝置和螢幕尺寸，減少學生使用自備裝置時的操作負擔，提供最佳使用者體驗。

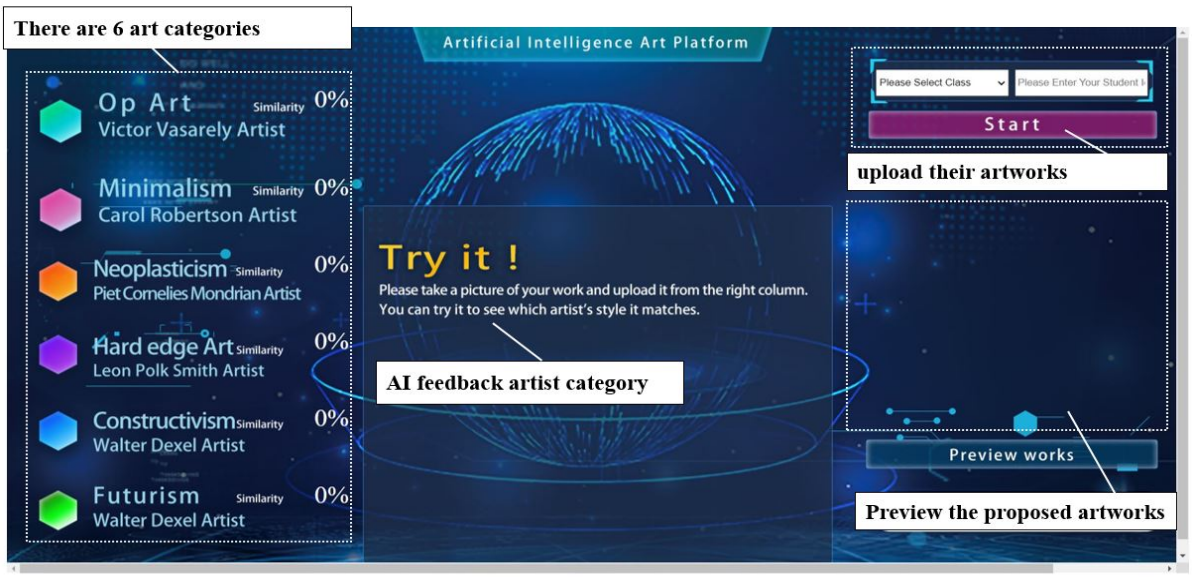


圖 1、系統剛開始畫面

4. 實驗設計

4 .1.參與者

參與實驗對象來自台灣中部同一所大學的 2 個班級，共 46 名選修藝術課程的大學生，以不具備藝術及設計背景的初學者為主。隨機分配 1 個班級 24 名學生為實驗組，使用基於深度學習之藝術學習系統進行學習; 另一個班級 22 名學生為控制組 ,採用一般科技輔助學習模式。這兩個班級皆由同一位具備 8 年教學經驗的授課教師進行教學。

4 .2.實驗過程

首先，教師在課堂上進行軟體操作及先備知識的課程教學，內容包含：藝術風格、歷史背景與藝術家作者的介紹。其次，進行 60 分鐘的學習成就前測及問卷前測。完成後，實驗組學生以電腦連結至 DL-ALS 系統進行學習，透過反覆練習及上傳繪畫創作，完成學習任'務。控制組學生以簡報影片搭配教師講述的方式進行學習，並完成繪畫創作練習。接續，兩組學生進行學習成就後測與問卷後測。最後，引導學生在 180 分鐘內完成繪畫創作成品，如圖 2。

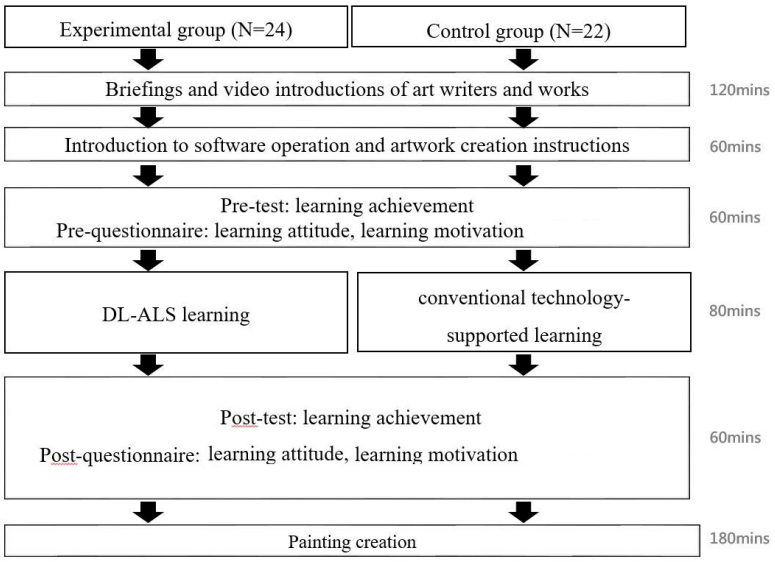


圖 2、實驗流程

4 .3.實驗工具

本研究中學習成就的測試由兩位具 10 年以上藝術設計教學經驗的教師出題，以評估學生對不同藝術風格的鑑別能力，涵蓋未來主義、歐普藝術等多種風格(前測 KR20 值為 0.82，後測為 0.88)。學習態度問卷引用 Hwang, Yang 和 Wang (2013)開發。它由 7 個項目組成，採用李克特五點量表，其中"5"代表"非常同意" ，"1"代表"非常不同意"。Cronbach's alpha 值為 0.89，顯示內部一致性信度良好。學習動機問卷參考自 Pintrich et al. (1991) 和 Wang & Chen (2010) ，包含內在動機和外在動機各 3 個項目的李克特五點量表(Cronbach's alpha 值為 0.88)。繪畫創作表現評分標準基於 Lai & Hwang (2015)，涵蓋圖案、空間、顏色和主題四個維度，每維度 1 至 3 分不等，由兩位教師評分，其 Kendall's ω 值為 0.76 ($p < 0.05$)，顯示評分者間一致性高。學生繪畫創作表現得分以兩位教師評分的平均數計算。

5. 研究結果

5 .1.學習成就

經誤差變異量的 Levene 檢定，驗證 F =0.661n.s. (p = 0.42> 0.05) 無違反回歸同質性假設，表示組內結果無顯著性差異，具有同質性，可使用 ANCOVA 分析學習成就。表 1 統計結果顯示，實驗組的學習成就顯著高於控制組（F =8.01，p =0.007< 0.01），實驗差異為大效果量 (η2 = 0.157) (Cohen, 1988)。兩組的調整後平均數分別為：實驗組 87.25 與控制組 70.03。表示使用 DL-ALS 系統進行學習，學生的學習成就會優於使用一般科技輔助學習模式。

表 1、兩組學習成就的共變數分析（ANCOVA）結果

Variance	Group	N	Mean	SD	Adjusted mean	*Std. error*	*F*	η^2
學習成就	實驗組	24	87.29	13.67	87.25	4.19	8.01**	0.157
	控制組	22	70.00	25.58	70.03	4.38		

**p<0.001

5 .2.學習態度

本研究以誤差變異量的 Levene 檢定，驗證學習態度 F =0.003n.s.（p = 0.96 > 0.05）無違反回歸同質性假設後，表示組內結果無顯著性差異，具有同質性，可使用 ANCOVA 進行分析。從表 2 統計結果顯示，實驗組的學習態度(Mean =4.73; SD = 0.47)高於控制組 (Mean = 4.40; SD = 0.49)，兩組之間具有顯著差異（F =5.11，p =0.02 <0.05）。η^2 值為 0.106，實驗差異為中等效應 (Cohen, 1988)。表示使用 DL-ALS 系統的實驗組學生，在學習態度上，優於使用一

般科技輔助學習模式的控制組學生；換言之，實驗組學生認為使用 DL-ALS 進行學習，是有趣且具重要性的。

表 2、兩組學習態度的共變數分析（ANCOVA）結果

Variance	Group	N	Mean	SD	Adjusted mean	Std. error	F	η^2
學習態度	實驗組	24	4.73	0.47	4.72	0.09	5.11*	0.106
	控制組	22	4.40	0.49	4.40	0.10		

*p<0.05

5.3.學習動機

經誤差變異量的 Levene 檢定，驗證內在動機 F =0.53 (p=0.46 > 0.05)及外在動機 F = 0.00 (p=0.99 > 0.05)皆沒有違反回歸同質性假設後，表示這兩組依變項的誤差變異量並沒有顯著差異，具有同質性，可使用 ANCOVA 進行分析。如表 3 所示，在內在動機的部分，實驗組表現(Mean =4.67; SD = 0.56)高於控制組 (Mean =3.95; SD = 0.78)，兩組之間具有顯著差異(F =8.98**，p =0.00 <0.01)。η^2 值為 0.173，實驗差異為大效果量 (Large) (η^2 = 0.157) (Cohen, 1988)。表示使用 DL-ALS 系統進行學習，學生的內在動機表現會顯著高於使用一般科技輔助學習模式。換言之，DL-ALS 系統可吸引學生學習及引起好奇心。在外在動機的部分，實驗組表現(Mean =4.58; SD = 0.65) 高於控制組 (Mean =4.00; SD = 0.87)，兩組之間達顯著差異(F = 8.73*，p =0.01 < 0.05)。η^2 值為 0.169，實驗差異為大效果量 (Large) (Cohen, 1988)。

表 3、兩組學習動機的共變數分析（ANCOVA）結果

Variance	Group	N	Mean	SD	Adjusted mean	Std. error	F	η^2
內在學習動機	實驗組	24	4.67	0.56	4.61	0.13	8.98**	0.173
	控制組	22	3.95	0.78	4.01	0.14		
外在學習動機	實驗組	24	4.58	0.65	4.59	0.14	8.73*	0.169
	控制組	22	4.00	0.87	3.99	0.14		

* p<0.05 , * * p<0.01

5.4. 繪畫創作表現

經獨立 t 檢定結果顯示，兩組學生的繪畫創作表現在圖案 (t=2.85**, p=0.007<0.01**)、空間 (t=2.57*, p=0.014<0.05*)、顏色(t=2.38*, p=0.021<0.05*) 及主題(t=2.50*, p=0.016<0.05*)皆有顯著差異。Cohen's d 值皆達 0.72 以上，代表具有較大的效果量(Cohen, 1988)，如表 4。顯見，實驗組學生在反覆練習繪畫創作，與上傳至 DL-ALS 系統獲得即時回饋的過程中，加強了對藝術作品的特徵瞭解，進而提升繪畫創作表現。

表 4、兩組繪畫創作表現的共變數分析（ANCOVA）結果

Variance	Group	N	Mean	SD	*t*	*d*
圖案	實驗組	24	2.66	0.48	2.85**	0.83
	控制組	22	2.18	0.66		
空間	實驗組	24	2.54	0.50	2.57*	0.77
	控制組	22	2.13	0.56		
顏色	實驗組	24	2.54	0.50	2.38*	0.72
	控制組	22	2.13	0.63		
主題	實驗組	24	2.62	0.49	2.50*	0.74
	控制組	22	2.27	0.45		

*p<0.05 , * *p<0.01

6. 結論

本研究貢獻在於開發基於 Fine-tune ResNet50 Model 架構的 DL-ALS 線上學習系統，採用響應式網頁設計，讓學生們不受無時間地點限制，可以隨時上傳作品獲得即時反饋。此系統加強學生們的記憶、專業知識，提升他們的學習成就和創作表現。動畫回饋增強學生互動、激勵練習提升學習態度和動機。系統介面簡單易用，提高學生們的滿意度。同時，系統評估學生作品可減輕教師負擔，提供學生學習成果數據。 然而，研究限制包括 DL-ALS 系統依賴 WikiArt 藝術作品，有限作品數量增加訓練模型困難。目前受測對象限於大學生，未來可擴及不同年齡層。系統後台可收集大數據，未來可供更廣泛學生使用，進行學習行為分析。目前系統主要用於學習繪畫創作與鑑賞，未來可擴展至雕刻、陶藝等其他藝術形式。

參考文獻

Alenezi, A. R., Karim, A., & Veloo, A. (2010). An empirical investigation into the role of enjoyment, computer anxiety, computer self-efficacy and internet experience in influencing the students' intention to use e-learning: A case study from Saudi Arabian governmental universities. *Turkish Online Journal of Educational Technology-TOJET, 9* (4), 22-34.

Barnet, S. (1985). *A short guide to writing about art*, Pearson Education.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cetinic, E., Lipic, T., & Grgic, S. J. E. S. w. A. (2018). Fine-tuning convolutional neural networks for fine art classification. *Expert Systems with Applications*, *114*, 107-118.

Chang, S. C., Hsu, T. C., & Jong, M. S. Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Computers & Education, 146*, 103758.

Chang, C. Y., Hwang, G. J., & Gau, M. L. (2021). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*.

Chen, M. Y., Mou-Te Chang, F., Chen, C. C., Huang, M. J., & Chen, J. W. (2012). Why do individuals use e-portfolios?. *Journal of Educational Technology Society, 15* (4), 114-125.

Emmers, E., Baeyens, D., & Petry, K. (2020). Attitudes and self-efficacy of teachers towards inclusion in higher education. *European Journal of Special Needs Education*, *35* (2), 139-153.

Farooq, M., & Hafeez, A. (2020). Covid-resnet: A deep learning framework for screening of covid19 from radiographs. *arXiv preprint arXiv:.14395*.

Guo, G., & Zhang, N. (2019). A survey on deep learning based face recognition. *Computer vision and image understanding*, *189*, 102805.

Hayadi, B. H., Bastian, A., Rukun, K., Jalius, N., Lizar, Y., & Guci, A. (2018). Expert system in the application of learning models with forward chaining method. *International Journal of Engineering Technology, 7* (2.29), 845-848.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013), A concept map-embedded educational computer game for improving students' learning performance in natural science courses, *Computers & Education*, *69*, 121-130.

Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A., & Winnemoeller, H. (2013). Recognizing image style. *arXiv preprint arXiv:1311.3715*.

Ke, C.H., Sun, H. M., Yang, Y. C., & Sun, H. M. (2012). Effects of User and System Characteristics on Perceived Usefulness and Perceived Ease of Use of the Web-Based Classroom Response System. *Turkish Online Journal of Educational Technology-TOJET, 11* (3), 128-143.

Keren, D. (2002). *Painter identification using local features and naive bayes.* Paper presented at the 16th international conference on pattern recognition, Quebec City, Canada, August 11-15.

Koć-Januchta, M. M., Schönborn, K. J., Tibell, L. A., Chaudhri, V. K., & Heller, H. C. (2020). Engaging with biology by asking questions: Investigating students' interaction and learning with an artificial intelligence-enriched textbook. *Journal of Educational Computing Research, 58* (6), 1190-1224.

Kong, F. (2020). Application of Artificial Intelligence in Modern Art Teaching. *International Journal of Emerging Technologies in Learning, 15* (13), 238-251.

Lai, C. L., & Hwang, G. J. (2015). An interactive peer-assessment criteria development approach to improving students' art design performance using handheld devices. *Computers & Education*, *85*, 149-159.

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics, 6* (1), 1-29.

Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2020). Student engagement level in an e-learning environment: Clustering using k-means. *American Journal of Distance Education*, *34* (2), 137-156.

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon, 7* (5), e07014.

Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of global health, 8* (2), 020303. https://doi.org/10.7189/jogh.08.020303

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research Practice in Technology Enhanced Learning, 12* (1), 1-13.

Stray, J. (2019). Making artificial intelligence work for investigative journalism. *Digital Journalism, 7* (8), 1076-1097. https://doi.org/10.1080/21670811.2019.1630289

Tan, W. R., Chan, C. S., Aguirre, H. E., & Tanaka, K. (2016). *Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification.* Paper presented at the 2016 IEEE international conference on image processing (ICIP), Phoenix, AZ, USA, September 25-28.

Theckedath, D., & Sedamkar, R. (2020). Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science, 1* (2), 1-7.

Virnodkar, S. S., Pachghare, V. K., Patil, V., & Jha, S. K. (2020). CaneSat dataset to leverage convolutional neural networks for sugarcane classification from Sentinel-2. *Journal of King Saud University-Computer Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.09.005

Wang, L. C., & Chen, M. P. (2010). The effects of game strategy and preference-matching on flow experience and programming performance in game-based learning. *Innovations in Education and Teaching International*, *47*(1), 39-52.

West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. *Governance studies at Brookings, 4* (1), 1-10.

Yang, X., Wang, Y., Byrne, R., Schneider, G., & Yang, S. (2019). Concepts of artificial intelligence for computer-assisted drug discovery. *119* (18), 10520-10594.

Zhang, C. (2019). Research on the Fluctuation and Factors of China TFP of IT Industry. *Journal of Industrial Integration Management, 4* (04), 1950013.

近三十年国内外学习空间研究进展

——基于 1999-2022 的文献分析

Progress in learning space research at home and abroad in the past 30 years: Literature analysis based on 1999-2022

刘思嘉*，赵赫璇，马国威，付玉娜，毛晓龙，王妍莉
*西北民族大学教育科学与技术学院

***1270325918@qq.com**

【摘要】 随着教育和技术的深度融合，学习空间的相关研究得到了越来越广泛的关注。通过 Cite Space 软件，对 CNKI 数据库和 WEB OF SCIENCE 数据库 1990-2022 年的 3619 篇文献进行可视化分析，在关键词共线分析与关键词聚类图谱分析基础上辨析了学习空间的相关概念，梳理了研究热点及相关案例、发展阶段及研究趋势。研究发现学习空间的相关研究呈现持续上升趋势；研究机构多为师范院校；研究热点可大体分为理论研究、框架研究、建设研究和教学应用研究四大类；在未来学习空间将进一步在支持智慧化的无缝学习方式、基于大数据的适应性学习空间和虚实融合的空间教学环境方面展开深入探索。

【关键词】 学习空间；文献计量；内容分析；文献综述

Abstract: *With the deep integration of education and technology, the study of learning space has been paid more and more attention. Using Cite Space software, this paper makes a visual analysis OF 3619 literatures in CNKI database and WEB OF SCIENCE database from 1990 to 2022, and distinguishes the relevant concepts of learning space on the basis of keyword collinear analysis and keyword cluster graph analysis. The research hotspots and related cases, development stages and research trends are reviewed. It is found that the related research of learning space presents a continuous upward trend. Most of the research institutions are normal colleges. The research focus can be divided into four categories: theoretical research, framework research, construction research and teaching application research. In the future, the learning space will be further explored in terms of supporting intelligent seamless learning methods, adaptive learning space based on big data and spatial teaching environment integrating virtual and real.*

Keywords: learning space, bibliometrics, content analysis, literature review

基金项目：本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学本科教学质量提升项目“教育技术学”课程思政示范专业（项目编号：2021KCSZSFZY-03）资助

1. 引言

任何形式的学习活动的开展都必须在一定的时空条件下进行(李秉德,1989)。学习空间作为信息技术发展并应用于教育领域的产物，它的优劣影响着教学活动开展的质量，而构建良

好的学习空间是实现学与教方式变革的基础（杨俊锋等,2013）。为了更好地了解学习空间国内外的研究现状、研究热点和研究方向。本文运用内容分析法，梳理了学习空间的研究历程，明晰了学习空间的相关概念与研究热点，并在厘清当前研究现状的基础上，对学习空间的未来的发展趋势进行了预测，丰富学习空间研究理论基础，为深入开展学习空间的研究提供理论支撑，为学习空间未来的研究提供参考建议。

2. 学习空间概念辨析

随着技术的发展，学习方式逐步多样化，混合式学习开始盛行。为了突出物理空间和虚拟空间的组合,“学习空间”也往往被叫做“混合学习空间”，二者无本质区别，均强调是学习发生的场所，是实体空间和虚拟空间的综合(王继新等,2010)；是线上、线下融合的无缝学习环境；是以建构主义、学习科学等理论为指导、通过运用信息技术为学习者提供最佳的、具有高度的灵活性的学习方式，以此来更好的促进和提升学习者的学习绩效(吴南中,2017)。

3. 学习空间研究变迁及因素分析

3.1 期刊载文量分析

本研究整理了国内外近 30 年的教育类相关核心期刊文章关于学习空间相关研究的年度发文量。整体而言 1990-2022 年间有关混合学习空间的期刊论文的发文篇数呈现持续上升的趋势，这表明学习空间的相关研究持续且越来越受到研究者们的关注。

3.2 关键词共线分析

通过分析可以看出学习空间的相关研究聚焦“智慧课堂”“智慧校园”“Collaborative Learning（协作学习）”“Higher Education（高等教育）”等相关领域，可以得出学习空间在发展和推进过程中聚焦领域的研究热点和变化趋势。

3.3 关键词聚类分析

通过对所选取文献进行关键词聚类分析，共得到国内 16 类聚类，国外研究 14 类聚类。技术、环境和执行过程是智慧环境建设的关键，根据这一理论，本文将聚类划分为四部分：学习空间理论研究、学习空间框架研究、学习空间建设研究和学习空间教学应用研究。

3.4 学习空间研究分析

根据智慧环境建设理论：技术、环境和执行过程 (Jordi et al.,2020)，结合国内外的关键词聚类分析图，本文将学习空间相关研究的聚类划分为四部分：学习空间理论研究、学习空间框架研究、学习空间建设研究和学习框架教学应用研究。

3.4.1 学习空间理论研究

与行为主义、认知主义和建构主义学习理论所不同的是，联通主义学习理论认为学习是发生在学习者个体外部得。因此，联通主义学习理论针对复杂数字环境提出了网络化的知识观和学习观，为学习空间之间的联系，尤其是网络学习空间存在的问题提供解决思路(曹璞,2021)。同时学习空间在面向学习者所呈现的技术支撑、环境保障、资源获取以及即时互动等方面，均会影响着学习者的个体学习行为的发生。社会认知理论是在班杜拉的社会认知学习理论基础上发展起来的，社会认知理论认为，个体、环境以及个体行为之间存在因果关系(Martin, J.,2004)。

3.4.2 学习空间框架研究

有关学习空间的框架研究，多运用 PST 框架或为基础根据实际需求进行修改运用。本文梳理了有关学习空间的建设框架，如表 1 所示包括 PSST 框架、FCDF 框架、PSE 框架等。

图 1 学习空间相关框架

研究团队	框架	主要内容
陈向东、吴平颐、张田力	PSST 框架	在原有框架的基础上增加了社会因素这一维度。
许亚锋、张际平	PST 改进框架	细化了学习空间设计、开发、评价的具体维度和过程，提供了可以借鉴实践取向的框架。
许亚锋、叶新东、王麒	FCDF 框架	遵循系统理论和系统方法，围绕以人为本的基本理念，使得教育、空间和技术三要素相互依赖、相互影响和相互作用。
蒋斌、郭庆华	PSE 框架	从工程建设角度来重新划分聚焦的维度，纠正实践中的负面倾向。

3.4.3 学习空间教学应用研究

以往的教育被设计成一个以工业为导向的系统，将学生视为在一条装配线待装的商品的工厂。但学习不仅是知识的传递，而是建构认知、分布式认知和情境认知的过程。这需要各方支持。学习空间的应用为学习者提供了基本的外部环境的保障，此外相应的教学应用为使用者提供了符合个性化发展的科学的方法论。如学习科学视域下构建的 E-Learning 环境下深度学习分析模型、O2O 教学模式的探索、探究创客空间和深度学习之间的关系实践等。也逐步形成了典型的学习空间建设应用案例。

4. 学习空间研究特征及趋势分析

4.1 学习空间研究发展阶段

本文参考已有研究的阶段划分，通过载文量和关键词的聚类分析结果，将近三十年学习空间的相关研究划分为四个阶段。

4.1.1 萌芽阶段（1990 年-1999 年）

这一阶段的研究主要以物理空间为主。学者们逐步重视教学空间的使用率以及对教学效果的影响作用，并开始关注教学空间的设计。但多数机构对于教学空间的有效使用率较低，并不利于学生的人际关系和学生个性化的发展(兰凤林,1993)，该阶段关于学习空间的研究刚刚开始并不全面。虽然已经注意到空间对于学生学习的重要影响，但并未深入探究。

4.1.2 兴起阶段（2000 年-2005 年）

随着网络技术的发展，各国分别颁发有关建设网络学习空间的文件，网络学习空间建设的重要性和必要性研究开始涌现。网络学习空间的资源应合理的使用，有利于促进教学和学习活动。这一阶段学习空间的相关研究以网络空间建设为主，且国内研究多以中小学学校、课堂为主，学者们开始大量探索网络学习空间的价值和意义，但受到技术的发展限制研究仍处于初步探索阶段。

4.1.3 发展阶段（2006 年-2015 年）

随着技术的不断更新网络学习空间进入快速发展期。这一阶段学习空间的研究多以网络学习空间建设以及探索个性化学习的网络学习空间教学模式与方法为主，并且开始从跨学科的角度，积极探索如何更好的将技术融入到学习空间的建设。如基于 2D 和 3D 技术的虚拟学习空间设计方案；跨学科的视角下的多显示学习空间开发和应用(陈向东等,2010)。

4.1.4 创新阶段（2015 年至今）

2015 年“第四次工业革命”开启，使得物联网、大数据、人工智能等技术逐步广泛的应用于教育领域。学习空间的研究开启了新阶段。这一阶段在推进学习空间理论、应用和建设的同时，也在探索如何建设更好空间以期为学生和教师提供更优质的服务，如何使用空间中的技术为学生提供更好的个性化学习支持，为教师提供更精准化的学情信息。

4.2 学习空间研究趋势分析

4.2.1 支持智慧化的无缝学习方式

无缝学习能够在移动设备的辅助支持不同情境下的学习，并且能够快速且容易地在不同学习情境间切换，融合了课内与课外、即时和非即时互动、正式和非正式学习、多种教学法和教学活动的结合，实现学习经验完整、连续的一种学习方式。学习空间具有高度的灵活性，方便学习者根据自身学习需求灵活地选择合适的学习方式，实现学习者不同学习需求之间的无缝连接，提高学习者的学习体验与对环境的感知。因此，混合学习空间将是开展无缝学习的新契机(肖君等,2021)。

4.2.2 基于大数据的适应性学习空间

创建适用性强的学习空间，在关注影响学习效果的不同变量和其关系的同时，也要创设理想的学习路径。通过描述和解释过去所发生的学习现象、预警和干预正在进行的学习活动、推断和预测未来的发展趋势，让学习者了解自己实际的学习需求和情况以及可能出现的结果，以此引导学习者向健康的方向发展，这是基于大数据进行的学习分析的优势也是未来学习空间的发展趋势。

4.2.3 虚实融合的空间教学环境

人工智能、人机交互、物联网等技术在教育教学中的广泛应用带来了对学习空间虚实融合的教学环境构建。虚实融合环境是通过多种传感设备识别、收集在真实环境下的学习活动所产生的学情信息，并依托互联网将课堂、真实学习环境和网络虚拟学习空间融为一体的一种新型学习环境。虚实融合的研究性学习环境的构建需要以融合学习情景和提升社会性互动为基准，构筑学习共同体、提供过程认知支架支持、学习资源的个性化聚合与推荐，构建个性化学习网络、虚实融合的研究性环境设计框架等几个方面的要素（杨进中和张剑平,2014)。

5. 总结与反思

随着技术的不断发展与革新，学习空间的发展将会越来越符合学习者的个性化需求，并且会受到越来越广泛的关注；因此必须以一种整体的方式来理解学习空间，以支持无所不在的学习。学习空间作为应对这种新范式而创建的混合空间，在未来的建设与发展过程中，应考虑到所有流程的个性化，以便为小组学习活动和个人任务的展开，提供更有效的方案。

参考文献

王继新，郑旭东，& 黄涛. (2010). 非线性学习空间的设计与创建. 中国电化教育(1)， 4.

兰凤林. (1993). 改变课堂教学的“秧田型”形态. 上海教育科研(04), 66.

李秉德. (1989). 对于教学论的回顾与前瞻. 华东师范大学学报(教育科学版)(03), 55-59.

杨进中 & 张剑平. (2014). 虚实融合的研究性学习环境设计. 电化教育研究(12), 74-80+85.

杨俊锋, 黄荣怀 刘斌. (2013). 国外学习空间研究述评. 中国电化教育(06), 15-20.

肖君, 梁晓彤, 黄龙翔 & 潘志敏. (2021). 无缝学习的焦点与趋势. 中国远程教育(02), 66-75.

吴南中. (2017). 混合学习空间:内涵、效用表征与形成机制. 电化教育研究(01), 21-27.

陈向东, 高山 金文. (2010). 多显示学习空间的开发与应用. 现代远距离教育(06), 59-63+74.

曹璞. (2021). 基于联通主义的新媒体课程教学改革. 青年记者(08), 98-99.

Jordi Mogas Recalde et al. (2020). Developments for Smart Classrooms: School Perspectives and Needs[J]. *International Journal of Mobile and Blended Learning (IJMBL)*, 12(4) : 34-50.

Martin, J. . (2004). Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, 39(2), 135-145.

面向数字创造力培养的图形化编程学习活动研究

Research on Graphical Programming Learning Activities for the Cultivation of Digital Creativity

刘昊瑶[1]，肖金爽[2]，王伟[3*]
[123]东北师范大学信息科学与技术学院
[*] dorrisww@126.com

【摘要】 创造力是创新型人才的特质，创造力培养需要把握创造力发展的儿童关键期，图形化编程能够助力儿童关键期的创造力发展。本研究首先从技术应用等方面厘清了图形化编程教学中数字创造力的内涵，从行为发生与产品产出方面对数字创造力的发生过程进行表征，并以此为依据设计了包括定义情境-问题转化等五个环节的学习活动。经过实践检验，所提出的图形化编程学习活动能够在数字背景下培养学生的创造力，为学生的数字创造力培养提供有效参考。

【关键词】 数字创造力；创造力；图形化编程；学习活动；学习活动设计

Abstract: *Creativity is the characteristic of innovative talents, and graphical programming can help children develop creativity in the critical period. This study first clarified the connotation of digital creativity in graphic programming teaching, and characterized the process of digital creativity, and designed five links of learning activities based on this. Through practice, the proposed graphical programming learning activity can cultivate students' creativity under the digital background, and provide effective reference for students' digital creativity training.*

Keywords: Digital Creativity; Creativity; Graphical Programming; Learning Activities; Learning Activity Design

1. 问题提出

创新是国家发展的不竭动力，培养具有创新精神的数字人才是智能时代国家发展的迫切需求。在《义务教育课程方案(2022 年版)》中也指出，要关注学生数字素养与技能的提升，进一步提高数字化合作与探究能力，发扬创新精神（教育部, 2022）。研究表明，2-6 岁是儿童个体创造性动力萌发、创造性人格和创造性思维培养的关键期，8 岁以后，他们早年的创造力似乎在一个并不重视不同做事方式的环境中丧失了（陈晓芳，2021）。作为具有“高天花板、低地板、宽墙”特点的图形化编程恰好契合了儿童创造力培养的需求，逐渐被各国政府与教育界重视，并逐步将其引入课堂教学中，成为儿童学习编程、提升数字技能、拓展创新思维的主要工具。

国务院《新一代人工智能发展规划》明确提出“在中小学阶段设置人工智能相关课程，逐步推广编程教育”（国务院, 2022）。以图形化为代表的儿童编程教育已经日益深入与普及，图形化编程教学为培养具有终身创造力的儿童提供了可行的实践场域。然而，有研究表明，对于不同年龄阶段的个体和不同的活动领域，创造力这种品质或能力有不同的表现（胡卫平和俞国良，2022）。因此，结合图形化编程这一活动领域的特征，界定与表征数字创造力，并探析适切的编程学习活动这一问题亟待解决。

为此，本研究旨在厘清图形化编程教学中数字创造力的内涵与表征，并据此进行图形化编程学习活动设计，以期为面向数字创造力培养的图形化编程课堂教学提供借鉴。

2. 图形化编程中的数字创造力解析及表征

结合图形化编程特点，可认为数字创造力发生过程的关键为环境支持下的行为发生与产品产出，具体如图 1 所示。行为发生包括信息搜集与创意启发、问题批判与独创生成、作品分享与智慧碰撞三个方面。信息搜集与创意启发是指学生根据需求对信息进行搜集处理并启发自己的创意灵感；问题批判与独创生成是指学生在使用技术解决问题、形成作品的过程中不断反思批判从而使作品在不断修补优化中具有独创性；作品分享与智慧碰撞是指学生愿意在技术支持下进行创意分享，达到相互学习、激发创意灵感的目的。产品产出包括学业成就与自我表达两个方面，其中学业成就主要指学生在学校环境内进行一系列活动如课堂学习、社团活动等的过程中产生的数字化知识与产品；自我表达是指除必要的学校任务外因学生的主观意愿而产生的数字创意作品。

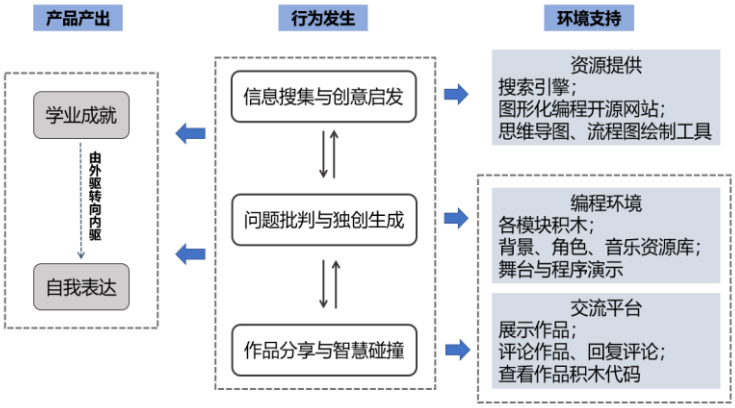


图 1 数字创造力表征

通过参考《托兰斯创造力测评量表》等测评量表，结合已有的创造力产品评价相关研究并综合考虑学生的思维特点，采用同感评估技术，确定作品评价的四个维度分别为：独创性、技术性、表现性、艺术性。利用 AHP 分析法得到四个维度的权重分别为 0.54、0.28、0.12、0.06，评分采用 5 分制进行计分，见表 1。对量表进行信度检验，本研究的量表克隆巴赫系数值为 0.836，即量表具有良好的内部一致性。

表 1 作品评价量表

评价维度	评价指向	具体内涵	权重
独创性	创新表达	作品具有独特性和其与众不同的程度	0.54
技术性	问题探究	作品使用技术的复杂程度	0.28
表现性	技术应用	作品表达的流畅程度与合理性	0.12
艺术性	创新表达	作品设计具备一定的艺术欣赏性，符合审美	0.06

3. 面向数字创造力培养的图形化编程学习活动设计

3.1. 定义情境，问题转化

编程教学以真实生活中的问题为基础，联系生活实际进行教学的具体情境创设。在教学中，教师通过故事等方法展开课程的具体情境，并进行问题描述、程序效果演示。学生根据教师的分析与自身体验对具体情境进行分析与定义，将真实问题转化为编程相关问题，在已有经验的基础上，建立编程语言与现实问题之间的逻辑关系，将问题转化为不同的子问题。

3.2. 头脑风暴，发散联结

确定待解决的问题链后，教师可以鼓励学生进行头脑风暴，通过思维发散提出观点。接着，教师可针对学生观点提供相应的资源，如作品案例，相关问题专题网站等，学生也可通过信息搜集、案例参考等方式，进行信息的交流分享。最后，通过联结现有知识与自身已有经验，厘清程序的主要逻辑以及关键算法。在此过程中，学生经历了思维发散-经验联结-明确思路的过程，为接下来自主设计与搭建程序打下了基础。

3.3. 设计实践，修补优化

明确问题解决思路后，教师可采用思维导图、流程图等方式引导学生进行搭建方案的设计，并在师生互动与生生互动的过程中，完善设计方案。接着，学生将方案应用于实践进行积木搭建，在搭建的过程中也会激发学生灵感，用全新的方式使用熟悉的材料等创造性地修补与创作（Resnikc, 2018）。同时，教师也需要关注学生的批判与反思，关注学生用辩证批判的眼光看待自己与他人的积木搭建过程，促进学生的积木优化与创意生成。

3.4. 聚合迁移，联系生活

编程教学中，根据学生学习能力的不同，提出解决方案的数量与质量也不相同。不同的作品会涉及不同的算法知识、形成不同的问题解决逻辑并用到不同的积木。教师可以抽取学生作品中涉及到的不同知识点，帮助学生选择问题解决的最佳方案。在此过程中，教师对知识点的抽取与总结过程也是学生聚合、联结所学知识的过程，有助于学生转变思维方式，将知识迁移至新的真实问题情境中，在已有方案的基础上进行作品的再创作。

3.5. 创造表达，智慧推动

经历设计、实践、优化、迁移等环节后，学生已对课程知识有所理解与掌握，学有余力的学习者更希望将创作的自由权利掌握在自己手中，结合已经学习的多个知识点进行独立自由的创作。无论是由教师启发的拓展任务还是学生的自主创作，其作品的产出过程都是学生的表达自身想法的过程。在这一环节中，学生需要首先完成自己的编程作品并录制展示，学生可以简述自己的作品创意、积木搭建过程并进行程序演示。同学之间可以互相查看对方作品，发表评论、提出优化想法等。教师也会对作品进行评价，选出优秀作品并邀请作者介绍自己的作品，包括积木搭建、视频剪辑创意等，推动学生的分享与智慧碰撞。

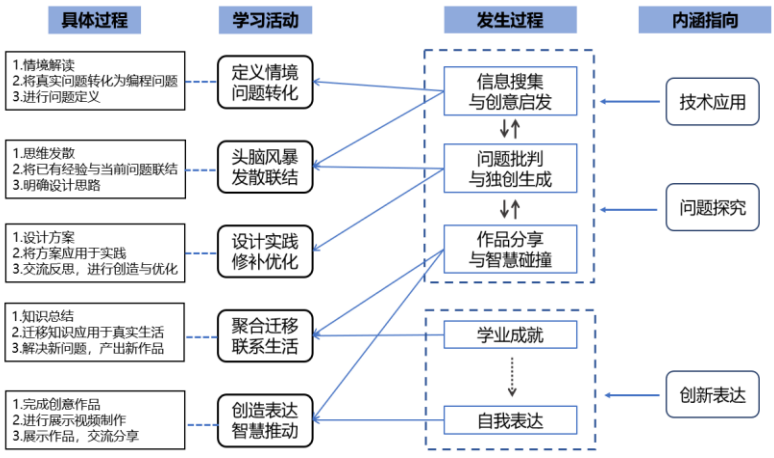


图 2 面向数字创造力培养的图形化编程学习活动

4. 教学实践与效果分析

4.1. 教学实施对象及流程

本次图形化编程教学实践的教学对象为公开招募的 50 位小学二年级学生，来自长春地区多所小学，平均年龄为 8 岁。50 位学生的计算机使用能力基本相同，图形化编程水平均为“未接触过编程”或“偶尔听说过”。学生需要进行为期两个半月共计 10 节编程课程的学习，教学内

容选取“讲故事，学编程”系列丛书中的图形化编程相关内容，教学活动的设计均以上述活动为基础。

4.2. 学生作品分析

本次效果分析采用“班级作品前后测对比”的方法，在课程实施过程中，分别在第一节课、第五节课以及第十节课结束后进行三轮随机作品收集，每轮收集 20 份学生作品共回收作品 60 份。对作品进行评分并统计加权，三轮作品的总分均值分别为 2.052、2.74、3.769，即随着课程教学的不断深入，学生的作品总质量不断提升。三轮样本数据均服从正态分布，对数据进行配对样本 T 检验，将第一轮与第三轮的作品数字创造力总体水平和各子维度的表现情况进行比较，学生的总体数字创造力水 p=0.000，效果十分显著；在独创性、技术性、表现性、艺术性四个维度，也都取得不同程度的显著。十次课程完成后，学生提交作品的独创性、技术性、表现性以及艺术性都有不同程度的提升，这表明基于数字创造力设计的教学活动与流程在保证了学生对于编程知识掌握的同时，也提升了学生运用数字技术进行创新创造的能力。

4.3. 教师访谈分析

课程结束后，从学生课堂表现与作品创意来源两个方面对任课教师进行访谈，以期对学生的学习情况有更全面的了解。经过十节课程的学习，学生不仅能够独立思考、探究交流从而为教师提出的问题提供解决方案，还能够在此基础上提出结合自身知识的新问题，创作新作品。作品创意分析方面，根据教师访谈可知，学生已经有意识对已学的积木进行迁移应用，并基于自身的兴趣来进行作品的创作，实现了对已内化知识的综合应用。

5. 结论与讨论

5.1. 结论

如前文所述，实施基于数字创造力培养的教学活动后，学生的数字创造力提升较为显著，学生作品的独创性、技术性、表现性以及艺术性都有不同程度的提升。可以说，在明晰图形化编程教学中的数字创造力内涵与表征的前提下，结合图形化编程教学特点与活动设计原则，让学生经历定义情境、问题转化等五个环节的学习活动，学生的创造力有所提升，学生的数字创造力能够得到培养。

5.2. 讨论与建议

在实施课程的过程中，仍有一些问题出现且未得到很好的解决，如课程教学更关注积木搭建，内涵不够丰富、学生容易受到外界信息的干扰、教师因能力局限无法为学生启发新创意等。因此，对于图形化编程课程中培养学生的数字创造力具有以下建议：（1）丰富编程教学内涵与外延，联结数字创意生活；（2） 设计润化创造力发展资源，营造创意实践环境 ；（3）关注设计思维与教学创新，契合数字时代发展。

参考文献

教育部. (2022).*教育部关于印发义务教育课程方案和课程标准（2022 年版）的通知*. http://www.moe.gov.cn/srcsite/A26/s8001/202204/W020220420582361024968.pdf

陈晓芳（2021）。儿童创造力培养的现实困境与实践突破。**人民教育，(Z3)**，29-32。

国务院(2022).*国务院关于印发新一代人工智能发展规划的通知*.http://www.gov.cn/zhengce/content/2017-07/20/content_52119 96.html.

胡卫平和俞国良（2022）。青少年的科学创造力研究。**教育研究，(01)**，44-48.

Resnikc, Mitchel（2018）。**终身幼儿园**。浙江：浙江教育出版社。

整合图示化支架的历史探究学习对学习者论证思维的影响研究

Research on the Impact of Graphical Scaffolding in Historical Inquiry Learning on Learners' Argumentative Thinking

蔡慧英[1*]，孙佳悦[1]
[1]江南大学人文学院
[*] caihy@jiangnan.edu.cn

【摘要】 论证思维是21世纪优秀人才所需的高阶认知思维。为了在历史教学中培养学习者的论证思维，基于探究的学习成为一种重要的教学模式。但是，在缺乏外在干预的情况下，基于探究的学习会对学习者产生较大的认知负荷，这会阻碍学习者论证思维的发展。因此，本研究提出了在基于探究的学习中整合图示化支架的教学策略。为了探究整合图示化支架的历史探究学习对学习者论证思维的影响，本研究按照基于探究的教学模式设计了整合图示化支架的历史教学单元，然后组织54名高中生进行了准实验。通过对学习者前后测数据的统计分析和认知网络分析，研究发现，整合图示化支架的历史探究学习显著提升了学习者对知识理解的深度，促进了其论证思维的发展。同时，本研究为教师优化历史教学质量提供了创新的教学方法和可行的实践方向。
【关键词】 图示化支架；历史论证思维；历史探究教学法

Abstract: *Argumentative thinking is a high-order cognitive skill required by outstanding talents in the 21st century. In order to cultivate learners' argumentative thinking in history education, inquiry-based learning has become an important instructional approach. However, without external intervention, inquiry-based learning can impose significant cognitive load on learners, hindering the development of their argumentative thinking. Therefore, this study proposes a teaching strategy that integrates visual scaffolds into inquiry-based learning. To investigate the impact of integrating visual scaffolds into historical inquiry learning on learners' argumentative thinking, the study designed a historical teaching unit with integrated visual scaffolds based on the inquiry-based teaching model. A quasi-experimental study was conducted with 54 high school students. Through statistical analysis of pre- and post-test data and cognitive network analysis, the research found that historical inquiry learning with integrated visual scaffolds significantly enhanced the depth of learners' understanding of knowledge and facilitated the development of their argumentative thinking. This study simultaneously provides innovative teaching methods and practical directions for teachers to optimize the quality of history education.*
Keywords: Visual scaffolding, historical argumentative thinking, history inquiry teaching methodology

1. 前言

论证思维是学习者日常生活中的重要组成部分（Osborne, 2010）。研究表明，论证思维不仅可以引导学习者进行知识的提升和创生（Scheuer, et al，2010），还可以引导学习者在新环境中做出有效决策并成功解决问题（Nussbaum, 2008）。因此，我们在中小学教学中需要培养学习者的论证思维。历史课程作为中小学教学中的重要学科，在教学中需要关注并培养学习者的论证思维。历史论证思维的培养能够使学习者对历史事件做出正确、合理的解释，并促使他们客观且理性地分析历史观点或历史事物（高明国，2022）。

为了培养学习者的论证思维，学习科学的研究者倡导教师采用基于探究的创新模式组织教学活动。该方法通常要求学习者通过围绕真实历史问题展开的任务，在探究过程中积极参

与到历史文献、文物等多种资源的阅读、分析和综合，从而形成对特定历史现象的理解（Van Boxtel，Voet & Stoel，2021）。但是，在缺乏外在干预的情况下，基于探究的学习会对学习者产生较大的认知负荷，这会阻碍学习者论证思维的发展。因此，本研究提出了在基于探究的学习中整合图示化支架的教学策略。但在历史教学中，图示化支架很少受到关注。因此，本研究将采用认知网络分析（Epistemic Network Analysis，简称 ENA），探究整合图示化支架的历史探究学习在促进学习者历史论证思维上的作用，为历史教师提升教育教学质量提供实践启示。

2. 文献综述

2.1. 历史论证思维能力

历史论证思维是一种认知过程，涉及到分析、解释和构建历史论证或历史观点的能力。在历史课程学习过程中，学习者需要在了解历史社会背景以及收集和分析历史证据的基础上，从多个角度进行证据分析和逻辑推理，从而形成观点和主张。与大多认知过程类似，论证思维也有一定结构。一般地，很多研究都是基于图尔敏论证模型而开展的（Toulmin，1958）。在图尔敏论证模型中，论证思维的认知结构包含 6 个要素，分别是：主张、证据、支撑、限定说明、反驳、背景性支持。另外，Jimenez-Aleixandre 和 Erduran（2008）将论证的认知结构简化为对真实世界的主张、以及用来支撑主张合理性的证据和理由等 3 个要素。“主张-证据-理由”这一框架不仅可以用来简化论证过程的复杂性，帮助学习者关注于相关任务的特征，还能表达论证科学家推崇支持主张的价值取向。

2.2. 支持历史教育中学习者论证思维能力的探究学习法

为了培养学习者的论证思维能力，我们需要变革传统的知识传递式的教学方式，使学习者能够运用历史知识对历史现象进行科学的分析和深度认知（Lee，2005）。研究表明，以探究为基础的学习（Inquiry-based Learning）是教授历史内容和历史学科思维技能最有希望的方法之一（Reisman，2012）。以探究为基础的学习主要包括五个步骤（Colby，2008）。第一步：建立情境。它要求教师建立相应的历史背景来吸引学习者的兴趣，并且明确历史探究的过程。第二步：提出问题。它要求教师提出或学习者集思广益，构建实践问题。第三步：分析史料。它要求教师给学习者提供史料或其他资料，帮助学习者探索自己的想法。第四步：分析历史现象。它要求学习者结合具体的历史现象以及二手和原始史料，对最初提出的实践问题进行分析。第五步：形成反思观点。它要求学习者将探究出的想法进行论证或表述，从而回答最初构建的实践问题。并且，在此基础上确定新的探索问题。

以探究为基础的学习的主要特点之一是需要给学习者设置结构不良的问题。与只有唯一正确答案的结构良好的问题相反，结构不良的问题往往有多种解决方案，且每种方案都有各自的优缺点。虽然结构不良的问题会引发学习者进行复杂的认知加工，促进学习者的知识建构和思维发展。但是，研究表明，结构不良的问题会给学习者带来沉重的认知负荷，导致以探究为基础的学习不能顺利开展（Hmelo-Silver, Duncan & Chinn，2007）。因此，为了降低学习者在参与以探究为基础的学习中的认知负荷，往往需要教师提供一定的支架帮助学习者学习（Kirschner，Sweller & Clark，2006）。

2.3. 支持历史探究学习的图示化支架

为了降低学习者的认知负荷，我们可以在学习者历史课程学习中整合图示化支架，引发学习者进行有效学习。图示化支架源自于社会文化理论中“学习支架”这一概念。它是指在复杂学习情境中为学习者提供图示化的学习支架，促使学习者进行有效地社会性交互，从而帮助学习者解决问题、建构知识并形成协作问题解决能力的学习干预手段。与一般的支架不

同，图示化支架具有显性表征知识结构和学习过程的特点。通过梳理发现，在历史教学中，我们可以整合三种类型的图示化支架，助力学习者论证思维的发展。

第一类是可视化信息的支架。这类支架是指运用图片、图表、视频等方式表征学习中的知识或者信息，帮助学习者理解知识。依据多媒体学习认知理论可知，以多种方式呈现信息对主动处理信息的学习者是有利的。第二类是可视化简单思维的支架。这类支架是指运用不同类型的图形组织器辅助学习者进行有序思考，用以表征思维过程的干预手段，有助于学习者解决复杂任务。在历史教学中，较多使用的是用来进行构建知识、发散思维、提高学习能力的思维导图工具。第三类是可视化知识结构的支架。这类支架是指可视化呈现学习过程中知识间结构关系。这类支架可以帮助学习者对历史知识的结构关系形成全局理解，并有助于引发学习者的创造性思维。

综上所述，为了探究整合图示化支架的历史探究学习对学习者论证思维的影响，以及了解学习者在这一教学模式中知识理解和论证思维的发展情况，本文提出以下研究问题：(1)在整合图示化支架的历史探究学习中，学习者的知识理解情况是否有显著提升？(2)在整合图示化支架的历史探究学习中，学习者的论证思维能力是否有显著提升？学习者的论证思维能力是如何发展的？

3. 研究设计

3.1. 参与者和研究情景

为了探究整合图示化支架的历史探究学习对学生历史论证思维的影响，本研究采用便利抽样的方式，选取了无锡市某中学高一年级一个班级的学生进行了实验。其中，男生 17 人，女生 37 人，共计 54 人，该班历史教师有 8 年的教学经验。在充分了解我们的研究期望后，该教师与研究团队围绕一个历史单元，共同设计并实施了整合图示化支架的历史探究项目。

3.2. 整合图示化支架的历史探究学习设计

为了顺利开展整合图示化支架的历史探究学习实验，研究团队在研读历史课程标准和高一历史第四单元教材内容的基础上，与教师协同完成了教学设计。根据历史探究教学法模型（Colby，2008），研究团队与教师就高一历史《中外历史纲要（上）》中第四单元“明清中国版图的奠定与面临的挑战”的教学内容进行协同交流，并按照如下思路在教学设计中整合了图示化支架。首先，明确教学目标和内容。确定教学的具体目标和要传达的核心内容，确保理解学生需要学习和掌握的重点。其次，选择合适的图示化支架。根据教学内容的性质和学生的学习需求，选择最适合的图示化支架类型，例如时间轴、流程图、概念地图或图表等。最后，设计简洁清晰的图示。突出教学内容的关键信息和逻辑关系，并配合文字说明帮助学生理解图示内容。

学习阶段	学习活动的描述	探究环节	图示化支架
学习活动1：课程导入	1. 教师呈现明朝版图、重要历史事件时间轴，引出学习主题。 2. 教师呈现二位历史学者对明朝的看法，形成驱动性问题：为什么大家会对明朝有不同看法？ 3. 学生思考并在学习单上填写思考结果。	建立情境； 提出问题； 明确初步观点	1. 图片 2. 时间轴 3. 学习单
学习活动2：理解明朝的政治制度	1. 学生阅读课本，总结明朝政治制度的三点变化。 2. 呈现秦朝、唐朝与明朝的政治制度的结构图和表格，引导学生理解内阁首辅和宰相的区别。 3. 教师引导学生阅读课本段落、分析史料等，引导学生填充因果复流图，帮助学生理解废宰相、设内阁、设司礼监的前因后果及特点。 4. 教师引导学生回顾所学知识，在学习单上总结明朝政治制度的特点。	分析史料； 分析历史现象； 形成观点	1. 括号结构图 2. 因果复流图 3. 表格 4. 学习单
学习活动3：理解明朝的对外关系	1. 教师引导学生阅读课本并浏览微视频，了解郑和下西洋的基本过程。在此基础上，教师引导学生分析史料和浏览地图，理解郑和下西洋的目的、特点、影响和意义。 2. 小组阅读材料，思助韦思国，引导学生讨论郑和下西洋与西方航海的不同与成因。 3. 教师引导学生回顾所学知识，在学习单上总结明朝对外关系的特点。	分析史料； 分析历史现象； 形成观点	1. 视频 2. 地图 3. 因果复流图 4. 韦思国 5. 学习单
学习活动4：理解明朝的领土变化	1. 运用表格，对比呈现明朝应对日本、葡萄牙和荷兰外来侵略势力的政策和结果；以及以地图的方式呈现戚继光抗倭成争的线路图，引导学生对明朝海疆局势形成整体认识。 2. 组织学生分工阅读书本关于明朝内陆边疆的内容，并合作完成明朝应对蒙古、西藏和东北等内陆民族势力的政策和结果，引导学生对明朝内陆边疆局势形成整体认识。 3. 通过回顾所学知识，在学习单上总结明朝对外关系的特点。	分析史料； 分析历史现象； 形成观点	1. 表格 2. 地图 3. 学习单
学习活动5：总结	1. 引导学生依据本节课所学的内容，思考为什么不同的人会对明朝有不同的看法，并把想法填写在学习单上。	形成反思观点	1. 学习单

图 1 第 13 课教学设计

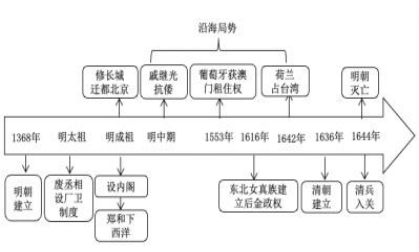


图 2 时间轴图示化支架

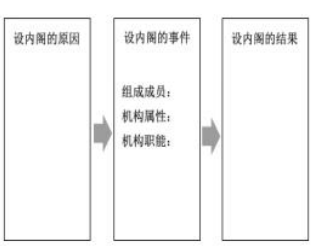


图 3 复流图图示化支架

最终，我们形成了第四单元的第 13、14 和 15 课的教学设计，其中，第 13 课的教学设计如图 1 所示。其中，在课程导入环节，我们会借用时间轴呈现明朝发展重要历史事件（如图 2），帮助学生对明朝形成全局理解。在学习明朝政治制度中“设内阁”的内容时，我们会引

导小组学生阅读史料，借用复流图分析“设内阁”历史事件发生的原因、内容和导致的结果，帮助学生形成逻辑推理能力（如图 3）。

3.3. 研究工具

为了判断历史单元教学中学生的知识理解水平以及论证思维的发展情况，研究者与教师共同设计了前测和后测的测试题。其中，为了评估学习者的知识理解水平，我们基于历史课程第四单元的学习内容，设计了 6 道单选题，用以测量学习者对该学习内容中简单知识的理解情况，如图 4 所示。另外，本研究还设计了 2 道简答题，用以测量学习者对该学习内容中复杂的知识理解情况，如图 5 所示。

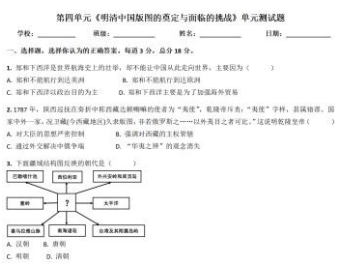


图 4 选择题示例



图 5 简答题示例

为了评估学习者的论证思维水平，我们设计了基于概念图的论述题，要求学生根据给定的概念绘制概念图，并用文字表达“对明清朝代的理解”。在绘制概念图时，详见图 6，学生需要根据已有的知识针对“概念列表”中的概念绘制概念图，展示学生对概念之间关系的理解。在写论述题时，学生需要根据所画的概念图，写出“对明清朝代的理解”，要求论述的文字需与概念图呈现出的理解一致、限制字数在 280 字左右，详见图 7。

3.4. 实验流程

教学设计和测试题设计好之后，研究团队与教师组织参与者开展了相应的课程学习实验。第一步，教师组织学生在 30 分钟内完成了前测。第二步，教师按照设计好的教学材料，运用三个课时，完成了第四单元中三节课的教学任务。每节课 45 分钟。第三步，教师组织学生运用 30 分钟完成了后测。在实验数据分析阶段，研究团队对收集的前测和后测数据进行了量化处理和分析。本研究对该高一班级学生发放测验卷并进行回收，在后测试卷回收中共收到 54 份，有效回收试卷 49 份，实现有效回收率为 90.74%。

3.5. 数据编码与分析

测试题中单选题每题 3 分，6 题共 18 分。其中，学生答对为 3 分，答错为 0 分。另外，根据 2 道简答题所包含的知识点和难度情况，我们规定第一题 8 分，第二题 14 分并分为两问（8 分和 6 分）。简答题评分时按照标准答案按“点”给分，并根据学生的回答理由是否充分、知识记忆是否准确、条理是否清楚等方面进行适当加或减分。

表 1 历史论证思维评价框架

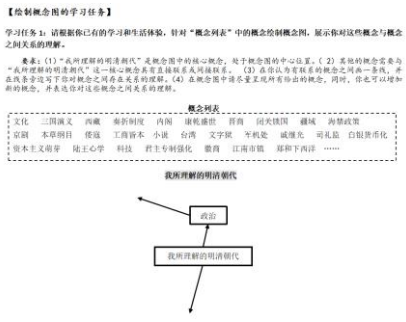


图 6 测验卷概念图题目



图 7 测验卷论述题题目

一级维度	二级维度	解释
主张	Level0	没有提出主张
	Level1	学生能够基于史料和历史情境等信息识别或提出历史主张
	Level2	学生能够根据历史情境，综合信息明确提出合理的历史主张
	Level3	学生能够根据历史情境，综合信息明确提出准确、凝练、严谨的历史主张
证据	Level0	没有提出证据
	Level1	学生能够引用一至两个具体史实证据支持主张，有简单描述
	Level2	学生能够引用两个以上合适且相关的史实证据支持主张，并进行简单描述
	Level3	学生能够引用两个以上具有可信性、相关性史实证据支持主张，并对证据进行合理解释
理由	Level0	没有提出理由
	Level1	学生直接叙述主张和史实证据之间存在的关系
	Level2	学生能够运用历史知识从原因、结果、意义、变化和多视角简单构建并解释主张和证据的关系，逻辑较严谨
	Level3	学生能够运用历史知识从原因、结果、意义、变化和多视角构建并解释主张和证据的关系，且论证充分，理由合理，逻辑严谨，表达通顺

根据简化的论证结构，本研究建立了历史论证思维的评价框架，见表 1。研究人员按照两个步骤对学生的论证文本数据进行了质性编码。第一步，研究人员提取了每个学生论证文

本数据中的话语，形成了最小的编码单元，确保每个编码单元包含完整的论证结构，即主张、证据、理由这三个要素。最终，共提取了 327 条单一话语，其中前测有 111 条话语，后测有 216 条话语。第二步，研究人员根据历史论证思维评价框架，识别和编码了每个最小编码单元中反映的论证结构及其水平。在编码过程中，研究人员要分析编码单元中包含哪些论证结构，并在此基础上，识别论证结构中该话语属于的水平。然后在该话语论证结构所属水平下编为 1，其余编为 0。

为了确保编码质量，我们对参与编码的 2 名人员进行了培训，确保他们对编码框架形成统一认识。编码完成后，经计算，编码人员之间的一致性系数均在 91%以上。对编码内容存在分歧的地方，进行协商讨论后达成统一意见。编码完成后，作者对编码数据进行系统地整理。首先，作者将前测和后测的数据放置在同一个工作表中。然后，作者根据每条数据的前后测来源，将前测编为 1，将后测编 2。根据每条数据所属学生，依次编码上相应的学生编号。最后，作者将整理好的数据导入 ENA 进行分析。

4. 研究发现与讨论

4.1. 整合图示化支架的教学是否可以促进学习者的知识理解

为了评估学习者在整合图示化支架后对知识的理解情况，本研究对学习者的前后测数据进行了配对样本 T 检验，结果如表 2 所示。我们可以发现，在总的知识理解上，学习者后测的分数与前测的分数存在显著性差异（T=-12.78，P<0.001），其中，后测选择题总分（M=17.3，SD=4.40）高于前测选择题总分（M=10.53，SD=3.09）。另外，学习者在简单知识理解和复杂知识理解上的后测得分均显著高于前测得分（简单知识理解：T=-12.65，P<0.001；复杂知识理解：T=-9.60，P<0.001）。具体而言，在简单知识理解中，后测选择题总分（M=5.47，SD=0.68）高于前测选择题总分（M=3.39，SD=1.30）。在复杂知识理解中，后测简答题总分（M=12.42，SD=3.86）同样高于前测简答题总分（M=7.29，SD=2.67）。这说明整合图示化支架的教学不仅显著促进了学习者对知识的整体理解，而且还提升了学习者对简单知识和复杂知识的深度理解。

表 2 学习者前后测中不同知识的均值和标准差

维度	前测		后测		t	df	p
	均值（M）	标准差（SD）	均值（M）	标准差（SD）			
简单知识的理解	3.39	1.30	5.47	0.68	-12.65	48	<0.001
复杂知识的理解	7.29	2.67	12.42	3.86	-9.60	47	<0.001
总的知识理解	10.53	3.09	17.3	4.40	-12.78	48	<0.001

4.2. 整合图示化支架的历史探究学习是否可以促进学习者的论证能力

为了了解在整合图示化支架的历史研究教学下学生历史论证思维变化情况，我们从主张、证据、理由三个方面对学生两次论证文本数据在水平上的频数进行了卡方检验。对学生论证文本数据的分析能够体现出学生在进行论证时的思维水平和严谨程度，从而明确学生论证思维的发展特点。由表 3 可知，从整体（X2=81.346，p<0.001）上来看，前测和后测学生在历史单元教学中在论证水平上存在着显著差异。具体地，我们也可以看出在主张维度（X2=25.714，p<0.001）、证据维度（X2=21.489，p<0.001）和理由维度（X2=44.217，p<0.001）上，前测和后测学生在历史单元教学中在论证水平均存在着显著差异。

表 3 学生两次论证文本数据编码卡方检验

维度		Level0	Level1	Level2	Level3	X^2	df	p
维度 1：主张	前测	16（14.4%）	92（82.9%）	3（2.7%）	0（0.0%）	25.71	3	<0.001
	后测	22（10.2%）	140（64.8%）	37（17.1%）	17（7.9%）			
维度 2：证据	前测	26（23.4%）	79（71.2%）	6（5.4%）	0（0.0%）	21.49	3	<0.001
	后测	32（14.8%）	129（59.7%）	37（17.1%）	18（8.3%）			

维度 3：理由	前测	80（72.1%）	30（27.0%）	1（0.9%）	0（0.0%）	44.22	3	<0.001
	后测	80（37.0%）	89（41.2%）	32（14.8%）	15（6.9%）			
整体	前测	122（36.6%）	201（60.4%）	10（3.0%）	0（0.0%）	81.35	3	<0.001
	后测	134（20.7%）	358（55.2%）	106（16.4%）	50（7.7%）			

为了了解在整合图示化支架的历史研究教学下学生历史论证思维变化情况，我们还从主张、证据、理由三个方面对学生两次论证文本数据在水平上的频数和占比进行了对比分析，结果如表 3 所示。通过分析表 3 中学生论证文本数据节点数量可以发现，在两次历史论证中，学生的历史论证思维在每个维度上水平的占比有一些变化。在前测正常历史课堂教学中，学生在 Level0 和 Level1 论证水平占大多数（均为 95%以上），在 Level2 和 Level3 论证水平上占极少数（均为 5%以下），其中在 Level3 论证水平上均为 0。然而，在后测整合图示化支架的历史探究学习中,学生在论证水平 Level0 和 Level1 论证水平上的占比有所下降(均为 78.2%以下）。在 Level2 和 Level3 论证水平上的占比有所增加（均为 21.7%以上），其中 Level2 论证水平均为 14.8%以上，Level3 论证水平均为 6.9%以上。这说明，在整合图示化支架的历史探究学习下，学生历史论证思维在水平上整体呈现向更高水平迁移的趋势，在结构上更加完善和严谨，学生的历史论证思维有所提升。

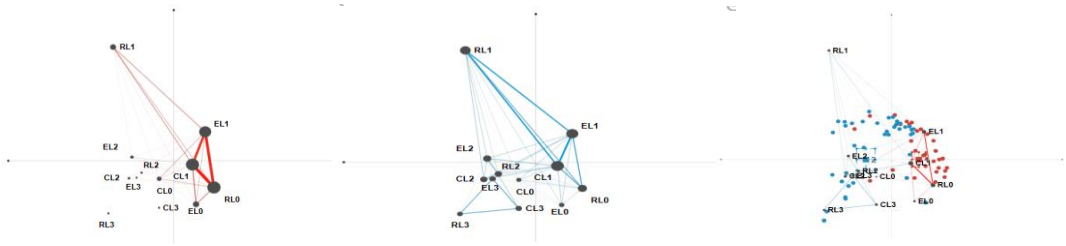


图 8 学生前测（左）、后测（中）以及两次对比（右）历史论证思维的认知网络结构

为了探究整合图示化支架的历史探究学习下学生在历史论证思维上的认知网络是否存在差异，对编码数据进行认知网络分析，结果如图 8 所示。学生在前后测历史论证思维所对应的认知网络结构有所差别。对认知网络分析结果进行双样本 t 检验后发现，学生前后测历史论证思维在 Y 轴上没有显著差异（P=1.00>0.05）；但在 X 轴上具有显著差异（P=0.00<0.05），如表 4 所示。在前后测历史论证思维对比图中可知，在前测的历史单元教学中，学生历史论证思维认知结构的质心靠近 X 轴右侧，侧重于“主张水平 1-证据水平 1-理由水平 0”这一论证水平及结构。然而，在后测的整合图示化支架的历史探究学习中，学生历史论证思维认知结构的质心靠近 X 轴左侧，侧重于“主张水平 1-证据水平 1-理由水平 1”这一论证水平及结构。

表 4 学生两次论证文本认知网络分析结果双样本 t 检验

	X 维度					**Y 维度**				
	Mean	**SD**	**N**	**t**	**p**	**Mean**	**SD**	**N**	**t**	**p**
前测	0.41	0.35	48	9.23	0.00	0.00	0.45	48	0.00	1.00
后测	-0.38	0.49	50			0.00	0.69	50		

另外，我们还可以发现，在后测中，学生历史论证思维认知结构在 X 轴左侧还很明显地呈现了“主张水平 1-证据水平 2-理由水平 2”以及“主张水平 3-证据水平 3-理由水平 3”这两种论证结构及水平。这表明，在历史单元教学中，图示化支架的介入对学生的历史论证思维的认知发展有着显著的正向作用。

表 5 学生两次论证文本认知网络共现关系示例表

共现关系	**前测**	**后测**	**共现关系**	**前测**	**后测**
证据水平 0-理由水平 0(EL0-RL0)	0.15	0.05	**主张水平 2-理由水平 2(CL2-RL2)**	0.00	0.07
主张水平 1-理由水平 0(CL1-RL0)	0.39	0.14	**主张水平 2-证据水平 2(CL2-EL2)**	0.00	0.06
证据水平 1-理由水平 0(EL1-RL0)	0.36	0.14	**证据水平 2-理由水平 2(EL2-RL2)**	0.01	0.10
主张水平 1-证据水平 0(CL1-EL0)	0.16	0.06	**主张水平 3-证据水平 3(CL3-EL3)**	0.00	0.12

主张水平 1-证据水平 1(CL1-EL1)	0.32	0.21	证据水平 3-理由水平 3(EL3-RL3)	0.00	0.12
主张水平 2-理由水平 1(CL2-RL1)	0.00	0.07	主张水平 3-理由水平 3(CL3-RL3)	0.00	0.12

根据认知网络分析结果的指标数据和表 5 中呈现的信息可知，在前测历史单元教学中，学生历史论证思维在主张水平 1 与理由水平 0 的关联度、证据水平 1 与理由水平 0 的关联度、主张水平 1 和证据水平 1 的关联度较高，但在后测中均呈现明显的下降趋势。另外，在前测历史单元教学中，学生历史论证思维在主张水平 3 与证据水平 3 的关联度、证据水平 3 与理由水平 3 的关联度、主张水平 3 与理由水平的关联度均为 0，但在后测整合图示化支架的历史单元教学中，学生历史论证思维在主张水平 3 与证据水平 3 的关联度、证据水平 3 与理由水平 3 的关联度、主张水平 3 与理由水平的关联度均为 0.12，呈现较大的上升趋势。这表明，在历史单元教学中，图示化支架的介入可以促进学习者历史论证思维的发展。一方面，学习者不再只简单罗列主张和证据，而是通过理由来建构主张和证据之间的桥梁，有效整合主张和证据，使论证结构更完善，更具有说服力，充分体现了学习者的逻辑性。另一方面，学习者在论证时能够根据历史情境，提出更具准确性和综合性的历史主张，并依托多个具有可信性和相关性的史实证据支持主张。还能从原因、结果等多视角构建并解释主张和证据的关系，使论证水平更高阶，论证更加充分。

5. 研究结论与启示

5.1. 研究结论

5.1.1. 整合图示化支架的历史探究学习可以促进学习者的知识理解

通过对整合图示化支架的历史探究学习的实证研究，我们得出结论：整合图示化支架的历史探究学习对学习者的知识理解起到了显著的促进作用。研究结果表明，在整合图示化支架的历史探究学习中，学习者对简单知识和复杂知识的理解都得到了显著提升。这体现在选择题和简答题的前后测得分差异上，后测得分均显著高于前测。值得注意的是，在复杂知识的理解方面，学习者在后测中的分数明显的提高。这表明整合图示化支架的历史探究学习在帮助学习者深化对历史复杂概念的理解上具有显著效果。图示化支架的引入为学习者提供了更直观、可视化的学习途径（Wood，Bruner & Ross，1976），帮助他们更好地理解历史事件之间的关联性。通过梳理和呈现主张、证据、理由的关系，学习者能够更系统地建构历史知识体系，提高对历史事件发展脉络的把握。

5.1.2. 整合图示化支架的历史探究学习可以提升学习者的论证思维能力

研究结果表明，整合图示化支架的历史探究学习对学习者的历史论证思维能力有显著的提升作用。通过卡方检验和认知网络分析，我们发现在整合图示化支架的历史单元教学后，学习者在主张、证据、理由三个维度上，前测和后测的论证水平均呈现显著差异。这表明，图示化支架的介入促进了学习者在历史探究中论证思维水平的提升。在论证思维的具体表现上，学习者在后测中呈现出更为完善和严谨的论证结构。认知网络分析则进一步揭示了学习者在历史论证思维认知结构上的差异。在整合图示化支架的历史单元教学后，学习者的认知结构呈现出更多元化、丰富化的关联，主张、证据、理由等要素之间的关联度明显增加，而且历史论证思维在水平上整体呈现向更高水平迁移的趋势。学习者在后测中更加注重“主张水平 3-证据水平 3-理由水平 3”这种高水平论证结构。一方面，这表明学习者能够更好地整合主张、证据和理由。学习者不再仅仅罗列主张和证据，而是通过理由来建构主张和证据之间的关系，形成更为有力和连贯的论证。另一方面，这表明在历史单元教学中，图示化支架的使用激发了学习者更高层次的论证思维，使他们能够更深刻地分析和解释历史事件，以及为自己的主张提供更有力的证据和理由。

5.2. 研究启示

在理论层面，本研究揭示了图示化支架在历史教学中的作用机制。图示化支架提供了一种直观、可视化的学习途径，通过帮助学习者建立主张、证据和理由之间的关系，加深对历史事件和概念的理解（马守信，2011）。这为教育理论提供了新的视角，使我们更好地理解学习者学习历史的过程，并有助于未来理论建构中更好地整合认知科学和教育心理学的成果。

在实践层面，研究强调了教师在历史教学中的关键角色。我们应倡导教师整合图示化支架进行教学，以更好地引导学习者构建论证结构、深化知识理解。教师可以通过设计和使用图示化支架，激发学习者的学习兴趣，提高他们对历史知识的领悟力（董海霞，2022）。因此，在今后的历史教育实践中，教师可以积极采用图示化支架，特别是在处理复杂历史概念和事件时，以更好地满足学习者的认知需求。

此外，本研究存在一定的局限性。未来的研究可以进一步深化对图示化支架在不同学科和不同年级的教学中的应用效果的探讨。通过深入挖掘图示化支架的特定元素和设计原则，研究者可以更加具体地了解哪些因素对于提升学习者的理解和论证思维能力最为关键。

参考文献

马守信. (2011). 历史教学中论证思维能力的培养. 历史教学问题 (03), 121-123+120.

高明国. (2022). 高中学生历史逻辑论证思维能力的现状分析及培养对策. 中学课程辅导 (04), 42-44.

董海霞. (2023). 协同设计视角下整合图示化支架提升教师信息技术应用能力的策略研究(硕士学位论文, 江南大学).

Colby, S. R. (2008). Energizing the history classroom: Historical narrative inquiry and historical empathy. Social Studies Research and Practice, 3(3), 60-79.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). Argumentation in science education. Perspectives from classroom-Based Research. Dordre-cht: Springer.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and. Educational psychologist, 42(2), 99-107.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational psychologist, 41(2), 75-86.

Lee, P. J. (2005). Putting principles into practice: Understanding history.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. Contemporary Educational Psychology, 33(3), 345-359.

Osborne, J. (2010). Arguing to learn in science: The role of collaborativecritical discourse. science, 328(5977), 463-466.

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. Cognition and instruction, 30(1), 86-112.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. International Journal of Computer-supported collaborative learning, 5, 43-102.

Toulmin S E. (1958). The use of argument. Cambridge University Press.

Van Boxtel, C., Voet, M., & Stoel, G. (2021). Inquiry learning in history. International handbook of inquiry and learning, 296-310.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of child psychology and psychiatry, 17(2), 89-100.

具身沉浸辨識的情境學習系統以提升學習者的學習成效

Embodied Immersion and Recognition-Based Situational Learning System to Enhance Learners' Learning Outcomes

任家駿[1]，吳建宇[2]，王振漢[3*]，楊舒涵[4]，趙伯堯[5]，陳國棟[6]
[126]中央大學資訊工程學系
[3]中央大學學習科技研究中心
[4]健行科技大學餐飲管理系
[5]元智大學資訊傳播學系
[*]harry@cl.ncu.edu.tw

【摘要】 情境認知強調在適當環境中學習，讓學生透過感官和運動行為獲得經驗。在過去的情境學習研究中常使用透過頭戴式裝置讓學習者進入到情境中，但這些方式缺乏讓使用者在過程中看到自己沉浸在數位情境中的結果。本研究在數位雙生劇場上，建立了具身沉浸與具身辨識的情境學習系統，提供學習者融入數位實境的學習。該系統會建立使用者的數位雙生，並根據身體感知和運動行為進行回應，顯示使用者的數位雙生與周圍環境的互動關係。研究結果表明，相較於傳統數位實境學習，本系統能有效地提升學習成效，在具身沉浸和認知感受也有正向的回饋。

【關鍵字】 情境認知；情境學習；具身認知；數位雙生；數位自我

Abstract: *Situated cognition emphasizes learning in environments, allowing students to gain experience through sensory and motor behaviors. Previous situated learning research often used head-mounted devices for immersion. However, these methods lacked the ability to allow users to see themselves immersed in digital environments. This study developed an Embodied Immersion and Recognition Digital Twin Theater learning system, providing learners with an immersive digital reality experience. It creates a digital twin for the user, responding to their bodily perceptions and motor behaviors, and displays the interactions between the user's digital twin and the surrounding environment. The results indicate that compared to traditional digital reality learning, this system effectively improves learning outcomes and receives positive feedback in terms of embodied immersion and cognitive perception.*

Keywords: Situated cognition, Situated Learning, Embodied Cognition, Digital Twin, Digital Self

1. 簡介

1.1. 研究背景

傳統的課堂學習中，大多數的教師會採用紙本教材工具，並透過口語闡述進行知識的傳授，讓學習者在被動的情況下接收知識，缺乏實際參與及應用的經驗。而情境認知主要目的是為了將課堂學習的知識內容場景移轉到現實場景中，並讓學生置身於情境之中，強調進行知識學習時，學習者應透過與情境之間的互動來獲得知識，該學習才是有意義(Brown et al., 1989)。

在人機介面互動方面上，傳統上許多研究是透過鍵盤、滑鼠、手持裝置和觸控的方式進行操控，並藉由虛擬化身代替學習者進入到虛擬世界進行互動，藉此代替使用者體驗數位情境。而具身認知強調學習者身體認知和動作行為的參與對認知形成是重要的(Wilson, 2002)，僅透過傳統的口語和紙本教學無法給予學習者的認知體驗，透過自身感官和運動行為等置身行為互動，才能讓學習者以直接參與的形式來與虛擬世界中進行聯繫。在過去研究也證實具身互

動可以提升使用者在數位實境中的參與度(Lindgren et al., 2016)，同時在部分研究中也強調其對於教育領域的重要意義(Shapiro & Stolz, 2019)，藉此我們可以觀察到相比過去一般的操作方式，透過使用者的具身互動可以提高使用者的學習認知感受，並改善使用者對於數位空間的體驗感。

由上述研究可知，當前的情境認知研究多依賴鍵盤、滑鼠等傳統控制器進行操作，這種方式使得學習者難以透過自身的感官和運動行為來進行互動，這種局限可能導致學習者在數位空間中難以建立自我認知。因此，本研究建立一套能夠讓學習者透過具身沉浸和具身辨識來進行情境學習的系統。該系統將學習者置身於數位情境中，透過自身的身體感官和運動行為參與互動，從而完全投入到數位情境中。系統透過具身辨識技術分析人體的運動和姿態，從而幫助理解和認知人體行為，實現具身認知的情境學習。在學習過程中，學習者還可以通過觀察數位情境中自身的肢體動作進行反思學習，進而提升學習效果和學習動機。

1.2. 研究目標

本研究提出的為使用者帶來具身沉浸與具身辨識的相關設計方法，並將此辨識系統應用於數位雙生學習劇場中，藉此驗證具備具身沉浸與具身辨識的情境學習系統是否能提升學習者的學習成效與學習動機，並探討該系統對學習者的影響。

1.3. 研究問題

本研究將探討使用具身沉浸與具身辨識的系統時，是否能提升學習者的學習成效與學習動機，及該系統對學習者的影響，因此本研究提出以下三個研究問題：「如何擷取使用者的全身資訊包含使用者認知和運動機制，以將使用者的完整數位雙生置入數位實境, 並且可以與數位實境中的景物互動, 達成置身認知的置身學習之情境學習?」、「如何讓使用者可以看到自己在數位實境中的學習表現並且如真實學習般的置身於真實情境中學習？」、「使用具身沉浸與具身辨識的情境學習系統對學習者的學習成效與學習動機的影響？」。

2. 相關研究

2.1. 情境認知

在 1989 年 Brown 等人提出情境認知的概念，其表示當抽象知識學習時，學習者應該是透過與情境之間進行互動來獲得知識，這樣的學習才是有意義的(Brown et al.,1989)。在現實研究中有許多方式可以達到情境認知的效果，有部分研究會採用真實學習的方式，藉由在實際環境中營造出學習內容相關的情境，並結合現實的議題，讓學習者置身於真實的環境當中進行學習，藉此提升學習成效(Wornyo et al., 2018)。或是透過戲劇式學習，將戲劇表演視為學習的一種方式，來豐富學習者的學習體驗(Hulse & Owens, 2019)，以及讓學習者以戲劇表演的方式來表達學習的內容，以提升學習者的口語表達能力及自信心，同時也幫助提升課堂的學習氣氛(Bora, 2022)。

但因為在現實世界中營造完全符合知識內容的情境並不是那麼容易，因此許多研究者開始將學習情境轉由透過數位實境世界進行模擬。部分研究利用數位學習劇場的方式，如 Wu 等人的數位劇場，把情境以劇場的方式進行建構，並讓學習者以不同的角色進入到劇場當中表演，並藉由完成劇場中的過程任務應用學習成果，完成學習目標，並將現實中人物影像和虛擬場景結合，讓使用者能以具身化的方式進入到情境中，提高使用者的情境沉浸體驗(Wu et al, 2015)。

2.2. 具身認知

根據 Wilson 提供六種具身認知的觀點，強調思維和認知過程是與身體感知和動作行為經驗密切相關的，表示我們的認知是通過身體的感知和動作進行實現(Wilson, 2002)。而具身學習是一種學習的方式，涉及讓身體的感官認知與感覺運動系統參與的學習過程，Lindgren 主

張認知不僅限於大腦，而與身體認知體驗和行為密切相關(Lindgren & Johnson-Glenberg, 2013)。部分研究開始透過手勢或者是使用者的身體動作進行直接的互動操作，如在虛擬世界利用手勢進行 UI 操控，用於建構 VR 環境的程式設計當中(Xiao et al., 2020)或利用 Kinect 進行動作捕捉，讓學習者透過動作來進行答題，以增加學習者的注意力(Kosmas et al.,2019)。

2.3. 數位雙生

數位雙生是指預期或實境的現實物理產品、系統或流程在數位平台中的雙胞胎，透過電腦計算來模擬現實中的環境，其主要包含真實環境、數位環境和資料連接(Tao et al., 2018)。而數位自我則是被定義為在數位環境中的虛擬分身和自我表現，看成一個獨立的個體，由在數位空間中的信息和活動構成，透過數位自我在數位空間中與他人進行交流(Zhao, 2005)，另外 Yee 等人則表示數位自我是指個人在網路世界中所創建的虛擬身分和數位自我形象，這個數位自我形象通常是由個人在網路平台的個人資訊等元素組成，透過這些元素，個人在網路上建構出代表自己的數位形象(Yee et al, 2009)。然而在工業或醫學等常見的數位雙生應用情境中較少提及數位自我的概念，在涉及使用者參與的數位雙生情境中，數位自我涉及到使用者的認知感知和互動行為，是數位雙生技術應用中不可忽視的一環。

2.4. 小結

綜合上述討論，情境認知和具身認知在教學領域的應用對於學習者在相關知識情境中的學習，以及透過身體感知和動作進行學習練習，具有顯著的幫助。然而現今許多數位學習系統雖然能夠在視覺上提供沉浸感，但仍缺少提供使用者具身沉浸的體驗，無法充分利用具身行為進行互動，也未能讓其他學習者進行觀察和學習。

因此本研究希望藉由使用者在數位雙生的情境進行互動，藉由建構使用者在情境中的數位自我沉浸於情境中，加強使用者的沉浸感和認知感受，並藉由辨識使用者多種不同的具身行為來達到具身學習的效果，來幫助學習者進行學習，並將其展現給教師和同學進行教學和觀摩。

3. 系統實作

3.1. 設計理念

本研究希望在數位雙生的劇場之中，設計一個具身沉浸與具身辨識的系統，提供學習者具身沉浸及具身辨識的學習情境，學習者將以自身感受和運動行為來體驗學習情境，加強其沉浸感和認知感受，同時讓在數位情境中進行合作學習，並藉由將學習者學習的相關知識透過螢幕畫面進行撥放，讓教師可以透過畫面進行說明教學，使學習者可以逐步修正錯誤的姿勢，且提供教室的其他同學進行觀摩來學習同儕之間的表現。

3.2. 環境設計

本研究在課堂的使用過程中，將教室規劃為展演區及觀眾區，如圖 1 所示。

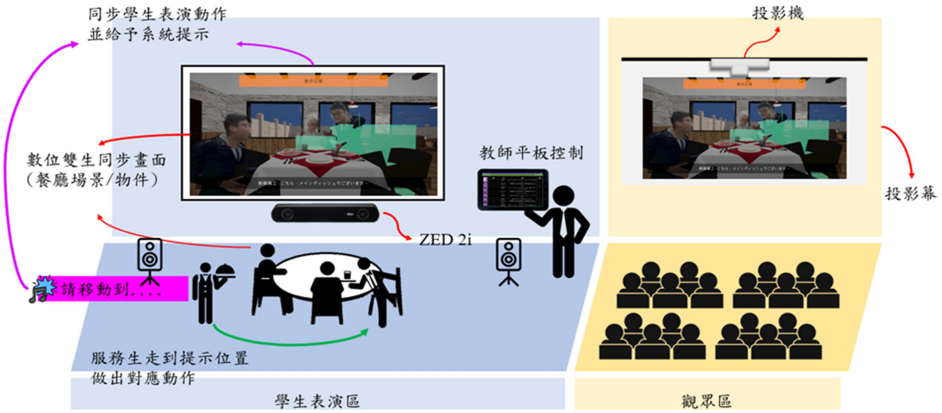


圖 1 課程現場環境架構

在表演進行時會將中間布幕拉上，將正在表演的學生分開，而觀眾區的學生可以透過同步畫面進行觀摩，而每一次演出後會由不同組別到展演區進行練習。

3.3. 系統架構

本研究系統架構如圖 2 所示，主要透過在 Unity 開發平台上建立的數位雙生劇場中，建立具身沉浸辨識系統。在學習者部分，系統會捕捉使用者的資訊，包含 3D 骨架資訊、3D 人物座標資訊、使用者去背影像及學習者的音訊資料，接著將其進行使用者的互動行為辨識，此部分包含使用者在情境空間中的走位、肢體行為的接觸範圍和動作角度及語音對話內容，並建立使用者在數位雙生劇場上的虛擬分身以還原使用者在現實中的表現，並將這些辨識結果透過數位雙生劇場的展示系統進行呈現；教師部分，在學習者使用系統前教師可以透過數位雙生劇本編輯系統編輯相關的資訊，並將課堂中會應用到的情境設定於資料庫當中給其他系統進行處理。且在學習者的表演的過程中，教師可以藉由劇場系統控制 APP 來控管表演的流程。

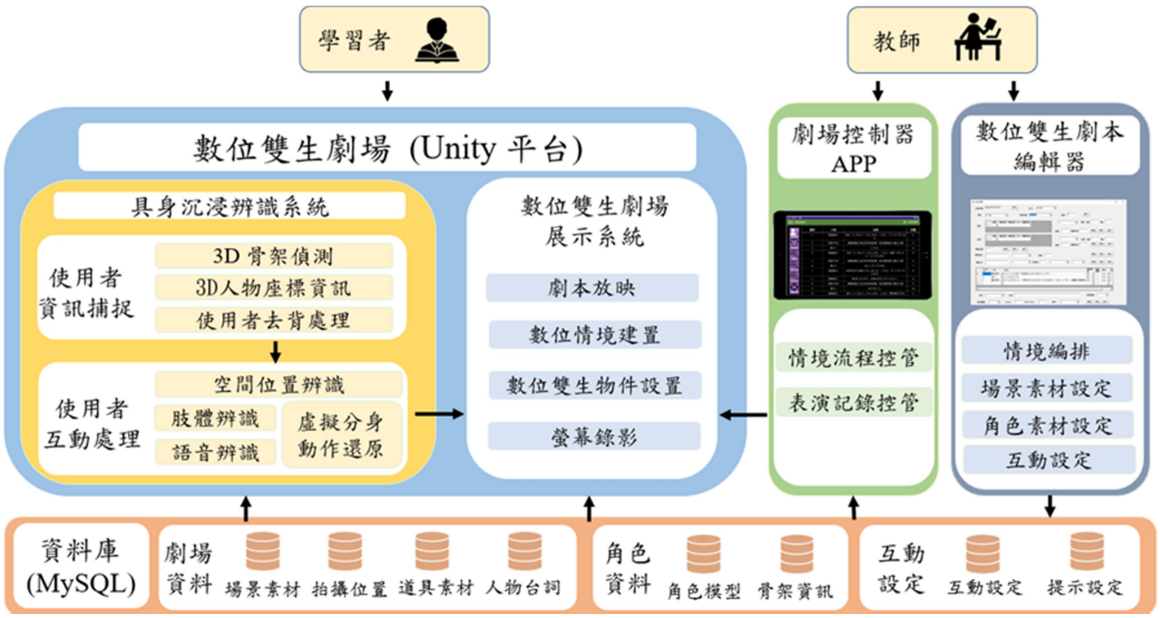


圖 2 系統架構圖

數位雙生劇場展示系統的部分主要會進行劇本放映、數位情境建置、數位雙生物件設置和螢幕錄影系統。在數位情境建置部分會依照劇本內容模擬使用者學習內容的情境，包含場景中建築物及並藉由物件辨識建構在情境中用到的物件道具，而在使用者行為處理部分，會接收 Zed 2i 拍攝的影像資訊及麥克風的聲音資訊，偵測使用者在現實中 3D 空間座標位置資訊、肢體骨架關節資訊及對話內容，並判斷使用者的行為是否符合當前情境的要求，在使用者行為不符合時會透過在虛擬數位情境中以視覺、聽覺及空間方向引導使用者，讓其進行即時的行為修正，最後回傳給展示系統的部分則會將畫面同步顯示給其他使用進行觀看及錄製進行反思學習。

數位雙生劇本編輯系統能夠讓教師進行數位情境的編輯，包括劇本內容設計、情境中的物件內容及使用者的動作要求，設計在不同的事件發生情況下會觸發的不同反應，並將其存入對應的資料庫。而資料庫內部儲存的教學內容及劇本皆參考自餐旅日語課本，並與餐旅管理系中餐旅日文課程的教授進行討論而成。教學內容包括服務生從接待客人、介紹菜單到結帳一系列的完整接待過程，幫助學生學習在接待過程中的相關日文及應對動作。

3.4. 互動設計

使用者動作捕捉過程，我們可以透過 Zed 2i 捕捉 3D 空間位置的辨識並擷取使用者肢體行為資訊做骨架辨識，以判斷使用者在數位實境中的互動行為。

在使用者的肢體偵測上，系統會在角色每一個關節部位設置碰撞檢測區域，這些區域可以檢測是否與其它的使用者發生碰撞事件，除了角色關節部位外，系統還可針對資料庫內的設定，在物體或使用者周遭產生碰撞區域。這樣的設計使系統可以捕捉到不同類型的互動事件，當使用者之間發生互動，或者使用者與虛擬物體進行互動時，系統能夠檢測將對應的事件，並且根據這些事件的發生情況，提供不同的回饋，如圖 3 所示。在語音辨識的處理流程則採用 Microsoft Azure Cognitive Service，將學生表達的音訊內容進行語音辨識處理並辨識其正確度，並將辨識到的結果呈現於畫面上，幫助學生進行修正。

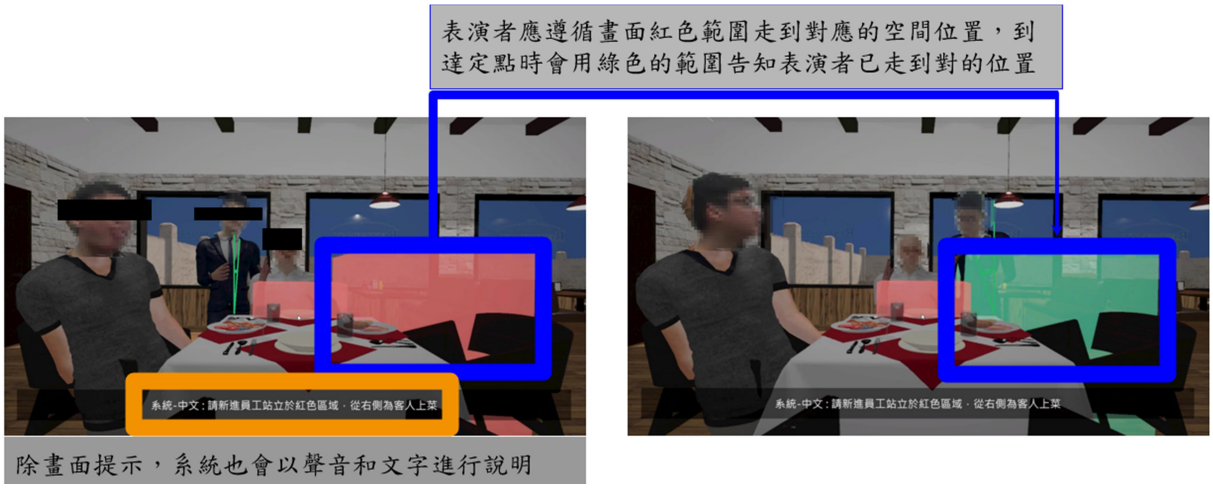


圖 3 系統給予學習者空間資訊及姿態等提示

4. 研究方法

本研究旨在確認學習者透過具身認知的學習情境對理解知識內容的學習表現影響，並設計實驗假設：學習者在具身沉浸辨識的學習情境中，透過攝影機即時辨識評量感官認知與運動行為，相較於僅使用紙本操作與知識說明的學習方式，以及透過教師情境畫面與同學互相評量建議的具身學習數位實境，能夠取得更好的學習成果。

4.1. 實驗對象

本研究與桃園某科技大學進行合作，挑選餐旅日文課程中的兩個班級總計 82 人作為受試對象，並隨機選取一班作為實驗組，另一班為對照組，其中實驗組包含 44 位學生，包含 11 位男同學及 33 位女同學;對照組包含 38 位學生，包含 16 位男同學及 22 位女同學，兩組受試對象完全獨立進行實驗，互不影響。

4.2. 測驗試題及問卷

本研究為評量參與實驗學生的學習成效，包括學生在餐廳接待時會用到的日文單字、基本用語及用於職場的對話內容和服務禮儀動作，作為教學的主要學習內容，其中包括客人入門的接待、點餐、上菜及結帳等內容，分別設計了一份前測與後測試卷，題目全部參考自《餐旅日語(下)》(黃招憲等, 2010)，並與餐旅日文課程的授課教授討論設計而成，題型包含是非題、選擇題、填充題，總分為 100 分。而為了解學生的感受，本研究使用 Likert 五點量表設計問卷題目，並以 Cronbach's alpha 來評比此問卷為高信度問卷。問卷內容方面，主要是參考 MSLQ 模型(Pintrich et al., 1991)進行設計而成，在此問卷設計模型中，包含多種不同的面性問卷設計範例，如評估學習者動機、學習策略和學習成果之間的關聯性，而在此研究中我們將針對「內在目標」、「任務價值」、「學習表現」及「同儕學習」，四個面向進行探討。

4.3. 實驗流程

本次實驗總共進行八周，實驗流程如圖 4 所示，在進行的實驗過程中，實驗組和對照組皆由課程專業的教師進行教學授課。在實驗組和對照組的實驗操控部分，兩組差異只在於學習者的使用系統部分，其餘的學習教材以及學習流程部分皆相同。進行情境學習系統時會以前測結果進行 S 型分組，將學生分為 6 人一組，並以小組的方式進行情境互動。

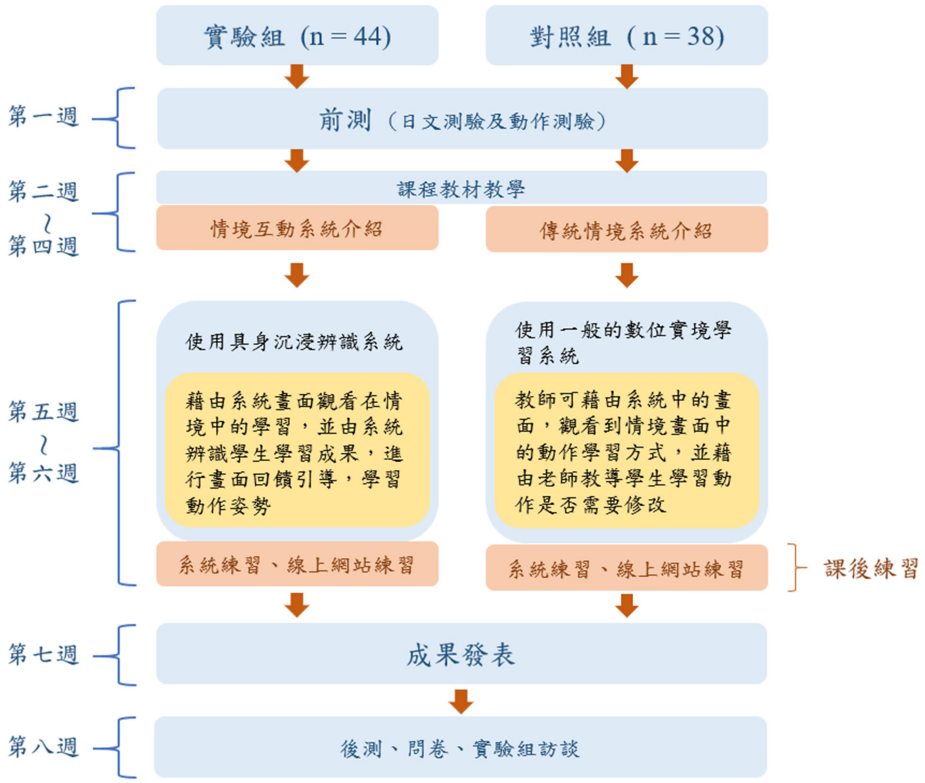


圖 4 實驗流程圖

5. 實驗結果與分析

5.1. 前後測學習表現結果分析

前後測試卷有效樣本共為 82 份，分為實驗組 44 份及對照組 38 份。使用 Shapiro-Wilk 進行常態分佈檢定，兩組前後測成績均符合常態分佈。ANCOVA 前進行組內迴歸同質性檢定結果符合基本假設。Levene 檢定結果顯示兩組後測的誤差變量並無顯著差異。

在排除前測成績的影響後，兩組在後測分數上具有顯著差異(F=5.309，p = 0.024 < 0.05)，由結果顯示在排除前測成績的影響後，實驗組和對照組之間在後測成績表現有顯著差異，且跟據 Cohen 所提出的效果量指標(Cohen, 1992)，效應值符合中等的效果量。故由該結果可驗證具身沉浸辨識的情境學習系統相較於使用一般情境學習系統之學生，在學習成效方面對學習者具有正面的影響。

最後針對題目兩個面向進行分析討論，在日文面向部分顯著性 p=0.098>0.05，兩組之間沒有顯著差異，在動作面向部分顯著性 p=0.046 < 0.05，兩組之間有顯著差異，透過 One-way ANOVA 分析結果顯示實驗組在動作分項部分相較對照組有明顯進步。

5.2. 問卷結果分析

此研究問卷是根據 MSLQ 模型進行設計，並以 Cronbach's Alpha 作為問卷信度分析之依據，可信度為 0.994 的高信度問卷。問卷內容主要探討該系統對學習者學習動機的影響，主要分為四個面向，包含內在目標、任務價值、學習表現、同儕學習，並進行獨立 t 檢定驗證各面向結果。

「內在目標導向」面向主要希望探討使用者在學習過程中對於自我增長的感受，探討在採用本研究的學習系統是否能讓受試者感到滿足，並加強受試者的學習意願，作答結果顯示實驗組的平均感受略高於對照組(M_{EG}=4.032, M_{CG}=3.973)，但未達顯著水準；「任務價值」面向則是針對使用者對於學習課程內容中的情境這項任務的理解情形進行探討，作答結果顯示實驗組的平均感受略高於對照組(M_{EG}=4.068, M_{CG}=4.026)；「學習表現」面向主要是針對使用系統對於學生自身的表現信心進行評量，作答結果顯示實驗組的平均感受略高於對照組(M_{EG}=4.091, M_{CG}=4.000)，但未達顯著水準，最後是針對「同儕學習」面向，本研究中的系統學習活動中，學生需要與其他的學生一同進行表演來完成，因此在此部份的問題希望探討是否有助於學生與其他同儕的互動學習作答結果顯示實驗組的平均感受略高於對照組(M_{EG}=4.091, M_{CG}=4.000)，但未達顯著水準。

結果表示對照組和實驗組在各面向並無顯著差異，因此兩組作答狀況屬於相同水平，均為「同意」區間。針對兩組在各面向均無顯著差異的結果，可能原因在於實驗時間只有 8 周，學習者無法充分體驗該系統，導致兩組無顯著差異，但實驗組的各面向的平均感受均略高於對照組， 可推測使用具身沉浸辨識的情境學習系統對學習者的學習動機仍具正面影響。

5.3. 訪談分析

除了測驗成績和問卷之外，本研究也針對實驗組的學生進行訪談。在問卷訪談部分，受試者表示與傳統的紙本學習相比，能有身歷情境的感覺，他們認為學習內容與現實餐廳情境的結合，幫助他們更好地理解和記憶學習內容。

而在系統使用意見訪談方面，受訪者對於將自己的動作和表情融入虛擬環境表示出高度的興趣和好奇，受訪者認為這種方式新奇有趣，但也希望系統能夠更細致地呈現個人特徵，以提高身份認同感。對於戲劇表演和語言學習的部分，部分學生表示一開始可能會感到緊張，但隨著熟悉程度的提升，表演會變得更自然，並建議增加語音輔助，以便於表演時更好地理解對話內容。整體而言，使用者對於具身沉浸辨識的情境學習系統的使用感受是抱持正面的態度，並且認同在表演的過程中同時進行語音和動作的不同模式學習是能有益處的。

6. 結論

本研究設計了一款具身沉浸與具身辨識的學習系統應用在數位雙生劇場中，為學習者提供了一種嶄新的情境學習方式，透過 ZED 2i 建置一個能夠反映學生具身現實表現的數位實境，並使用 3D 攝影機及影像擷取技術，建立使用者在數位實境中的數位雙生，並根據數位雙生的人物感測和行為進行互動回應，顯示與周遭使用者和環境的關係，以增強學習成效。以改善無法藉由與自身身體感知和運動機制下的方式接收學習知識、感受自身在空間和姿勢表情的呈現的問題。研究結果指出，使用具身沉浸與具身辨識的數位雙生劇場學習系統能夠顯著的提升學習成效，且從問卷結果可以得知在多個面向均給予了正向的回饋。綜合上述結果，該系統對於學習者有正面的影響。

在未來研究方面，可以改善數位雙生人物的精細度，新增更多自身細節讓畫面更符合自己的樣子。並嘗試將該系統應用於其他領域，利用不同的具身行為進行學習，不同認知的學習模式並將其應用在不同的學科當中，嘗試是否同樣有助於提升使用者在學習成效的表現。

7. 誌謝

本研究感謝國科會經費支持，計畫編號：MOST 110-2511-H-008 -004 -MY3；MOST 111-2410-H-008 -012 -MY3；NSTC 112-2811-H-008-006。

參考文獻

黃招憲、許惠端與王靖絜(2010)。餐旅日語(下)。致良

Bora, S. F. (2022). Drama pedagogy in foreign language learning: investigating the impact of theatre texts and theatre performance on L2 speaking accuracy and complexity. *Research in Drama Education: The Journal of Applied Theatre and Performance*, *27*(4), 507-523.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *1989*, *18*(1), 32-42.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98-101.

Hulse, B., & Owens, A. (2019). Process drama as a tool for teaching modern languages: supporting the development of creativity and innovation in early professional practice. *Innovation in Language Learning and Teaching*, *13*(1), 17-30.

Kosmas, P., Ioannou, A., & Zaphiris, P. (2019). Implementing embodied learning in the classroom: Effects on children's memory and language skills. *Educational Media International*, *56*(1), 59-74.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational researcher*, 42(8), 445-452.

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, *95*, 174-187.

Pintrich P., Smith D., Garcia T., & McKeachie W. (1991). A manual for the use of the motivated strategies for learning Questionnaire (MSLQ)

Shapiro, L., & Stolz, S. A. (2019). Embodied cognition and its significance for education. *Theory and Research in Education*, *17*(1), 19-39.

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. (2018). Digital twin in industry: State-of-the-art. *IEEE Transactions on industrial informatics*, *15*(4), 2405-2415.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, *9*, 625-636.

Wornyo, A. A., Klu, E. K., & Motlhaka, H. (2018). Authentic learning: Enhancing learners' academic literacy skills. *International Journal of Applied Linguistics and English Literature*, *7*(4), 56-62.

Wu, W. Y., Luo, Y. F., Huang, D. Y., Huang, C. W., Peng, Y. I., & Chen, G. D. (2015). A Self-Observable Learning Cinema in the Classroom. In *The 23rd International Conference on Computers in Education* (pp. 257-262).

Xiao, W. Y., Duan, J. Z., & Xue, C. Q. (2020, August). Research on visual programming system of gesture interaction interface in VR. In *Proceedings of the 2020 4th International Conference on Big Data and Internet of Things* (pp. 38-42).

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). The Proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, *36*(2), 285-312.

Zhao, S. (2005). The digital self: Through the looking glass of telecopresent others. *Symbolic interaction*, *28*(3), 387-405.

任务难度和元认知支架对学生协作编程学习成效的交互影响

The interactive effects of task difficulty and metacognitive scaffolding on students' collaborative programming learning outcomes

李雯[1]，李伟[1*]，曾秋蓉[2]，刘城烨[3]，黄佶艺[4]，朱玉蝶[1]，荣健康[1]
[1]温州大学 STEM 教育研究中心
[2]中华大学资讯工程学系
[3]香港中文大学课程与教学系
[4]温州市瓯海区瞿溪第一小学
[*]liweiwzu@wzu.edu.cn

【摘要】 本研究旨在探讨协作编程学习环境下，元认知支架与任务难度对学生编程能力的影响及其相互作用。台湾新竹县某初中一年级80名学生参与了为期13周的教学实验，被随机分配至实验组($n = 40$)和对照组($n = 40$)。实验组采用结合元认知支架的协作编程教学，对照组则采用传统协作编程方法。研究结果显示，元认知支架在简单及中等难度任务中对学习成绩的促进更为有效，且元认知支架和任务难度对编程成绩的交互影响不显著。此外，任务难度与元认知支架在影响学生元认知倾向上有显著的交互效应，随着任务难度的增加，元认知支架在激发学生元认知倾向方面的作用更加明显。

【关键词】 协作学习；编程学习；元认知支架；任务难度

Abstract: *This study explored the effects of metacognitive scaffolding and task difficulty on students' collaborative programming skills in a Taiwanese junior high school. Eighty first-year students participated in a 13-week experiment, randomly divided into experimental (n = 40) and control (n = 40) groups. The experimental group received programming instruction with metacognitive scaffolding, while the control group used traditional methods. The research results show that metacognitive scaffolding is more effective in promoting learning performance in simple and moderately difficult tasks, and the interaction between metacognitive scaffolding and task difficulty on programming performance is not significant. Furthermore, there was a significant interaction between task difficulty and metacognitive scaffolding in enhancing students' metacognitive tendencies, especially as task difficulty increased.*

Keywords: Collaborative learning, programming learning, metacognitive scaffolding, task difficulty

1.引言

近年来信息技术迅速发展，编程成为了一项必不可少的数字技能。然而，由于编程知识的复杂和抽象性的特点，使得大部分学生难以独立完成编程任务(Lin et al., 2018; Shadiev et al., 2014)。因此，有研究者提出使用协作学习的方式让学生一起完成编程任务 (Hanks et al., 2011)，有利于提高编程技能 (Lin et al., 2018)，培养学生的计算思维 (Wei et al., 2021)和解决问题的能力(Beck & Chizhik, 2013)。此外，协作编程使学生能够在学习编程的同时增加与同伴之间的互动和交流(Sandoval & Reiser, 2004)，帮助学生产生更多的解决问题的想法(Li et al., 2023a)。

然而，学生在协作编程的过程中面临一定的困难。一些研究指出，学生在协作编程的过程中，由于元认知能力不足(Cho, 2001)，不能有效的调节学习。在协作过程中缺乏对学习过程

的监控和反思能力(Hadwin & Oshige, 2011)，导致协作学习效果不佳。因此，在协作编程中为学习者提供适当的元认知支架是有必要的。元认知支架源于元认知概念，被定义为支持个体对自身的认知过程进行有效监控和控制的一种机制(Ouyang et al., 2022)。它可以帮助学生监控和调节学习目标、过程和结果(Kwon et al., 2013)。

元认知支架有利于学生在学习过程中认真的制定学习计划、监控自己的学习表现、根据所设定的目标进行自我评估以及根据设定目标不断的进行调整 (Huertas et al., 2017)。研究表明，将元认知支架整合到协作编程的学习中，可以帮助学生了解自己的学习状况，及时发现问题和解决问题，提高学生个人及协作学习的能力(Wang et al., 2023)。

此外，一些研究已经确定了任务难度是诱发元认知的一个重要因素 (Iiskala et al., 2004)。任务难度是指完成某项任务所需的认知和技能要求的程度(Robinson, 2001)，它被认为是影响人类行为表现的重要任务特征。不同的任务难度对激发学生的元认知水平也有所差异，学生的元认知在新的和相对困难的任务中表现得更为显著(Iiskala et al., 2011)。然而，在协作编程的研究中，较少的研究探讨不同的编程任务难度与元认知支架对学生编程学习的影响。因此，本研究旨在探讨任务难度和元认知支架对学生协作编程学习的交互影响。通过实验，本研究旨在探索以下研究问题：

RQ1: 任务难度与元认知支架对学生的编程学习成绩产生怎样的交互影响？

RQ2: 任务难度与元认知支架对学生的元认知倾向产生怎样的交互影响？

2.研究设计

2.1. 研究对象

本研究在台湾新竹县的一所初中进行，共有 80 名一年级学生参与。这些学生随机分配到实验组($n = 40$)和对照组($n = 40$)，两组均由同一位专业教师负责授课。该教师在编程教学领域拥有超过 5 年的经验。

2.2. 学习内容与环境

本研究的主要学习内容为程序设计，包括顺序结构、选择结构和循环结构。研究者并将其设计成三种难度（简单、中等和复杂）的任务。每种难度各出两个任务，共六个任务，进行协作学习任务，内容如下：

简单任务(顺序结构)：任务一，分配糖果。计算全班每人可分配到的糖果数。输入：糖果数及全班人数；输出：每人可得到的糖果。任务二，计算三角形面积。利用三角形的底边和高，计算出三角形面积。输入：使用者输入底边和高；输出：显示三角形面积。

中等任务(选择结构)：任务三，奇偶数判断。判断输入的正整数是奇数还是偶数。输入：使用者输入一个整数；输出：显示此数是奇数或偶数。任务四，比较两数大小。输入两正整数，判断这两正整数的大小。输入：输入两正整数；输出：两正整数的大小比较。

困难任务(循环结构)：任务五，找出 7 的倍数。找出 100 以内是所有 7 倍数的数字。输入：无；输出：100 以内所有 7 倍数的数字。任务六，计算奇数总和。计算输入整数的所有奇数总和。输入：一个正整数；输出：小于等于此输入正整数的所有奇数总和。

如图 1（左）所示，本案研究采用的结合元认知策略的编程学习环境。该协作编程学习系统由 Li et al. (2023a)开发，它为学生提供了一种基于块的编程环境，且为他们提供了元认知支架。在系统的引导下，学生在学习之前先设置学习目标，例如“我希望我们小组在这次任务取得几分”。确定学习目标之后，根据学习目标选择合适的学习方法，例如“为达成目标，我将采用以下哪些学习方法”。在完成学习任务之后，通过系统的自我反思功能对自己的学习成果进行反思，例如“在本次学习任务中，我积极参与了哪些合作任务”。同时，系统中

的监督机制，会提醒学生监控自己的协作学习过程，反思自己的学习是否达到目标，并及时根据目标进行调整，最终完成六个编程任务，如图 1（右）。

2.3. 研究工具

本研究采用的工具包括编程学习成绩的前测、低难度编程测试、中等难度编程测试、复杂难度编程测试，以及元认知倾向问卷。编程学习成绩的测试题由经验丰富的初中信息技术教师编制，经编程教育领域的专家共同审阅和修订。测试内容涵盖顺序结构、选择结构和循环结构的相关知识点。每套测试包含 10 道选择题，总分为 100 分。

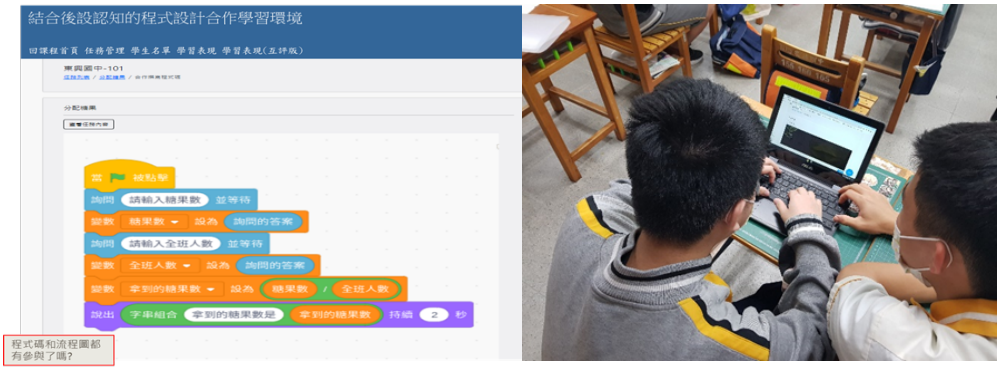


图 1 协作编程学习环境（左）和实验组协作解决编程问题的过程（右）

元认知倾向问卷改编自 Lai and Hwang (2014)所开发的量表，包含 5 个条目。例如“完成一项任务后，我会询问自己是否学到尽可能多的知识”。该问卷的 Cronbach's alpha 值为 0.87，以反映问卷的内部一致性。问卷采用 5 点李克特量表进行评分，分值从 1 至 5，依次代表“完全不同意”、“不同意”、“中立”、“同意”和“完全同意”。

2.4. 实验流程

本实验课程持续 13 周，每周安排 45 分钟的课时。在第一周，所有参与的学生完成了前置问卷调查和编程成绩的基线测试。接下来的三周（第 2 至 4 周），教师向学生们讲授实验课程的主要基础知识，包括顺序结构、选择结构和逻辑结构。在这期间，实验组学生采用融合元认知支架的协作学习模式，而对照组学生则应用传统的协作学习方式。第 5 至 7 周，学生们通过协作完成简单的编程任务，并进行第一次编程成绩测试。在第 8 至 10 周，学生们协作解决中等难度的编程任务，并完成第二次编程成绩测试。最后三周（第 11 至 13 周），学生们协作完成复杂的编程任务，并在课程结束时在教师指导下完成后置问卷调查和第三次编程成绩测试。

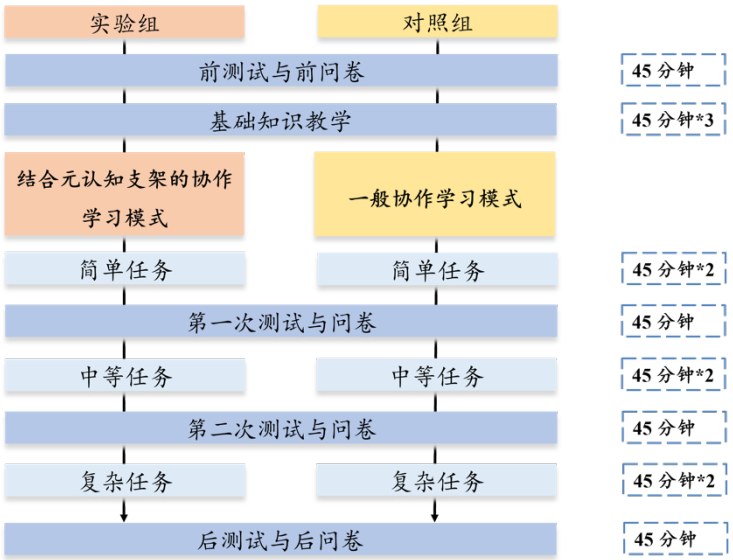


图 2 实验流程图

3.结果

3.1. 编程学习成绩

为了了解不同的任务难度和元认知支架对学生学习成就(对照组：$M = 55.5$, $SD = 16.58$; 实验组：$M = 58.5$, $SD = 19.81$)的影响，本研究进行了二因子重复测量方差分析(Two-way Repeated Measures ANOVA)分析。

表 1 编程学习成绩二因子重复测量方差分析结果

	平方和	自由度	平均值平方	F	η^2_p
任务难度	1517	197	771.5	4.167*	0.05
组别	3153	1	3153	9.39**	0.106
任务难度*组别	125	197	63.4	0.343	0.004
残差	28768	193	206		

*$p < .05$, **$p < .01$

编程学习成绩的二因子重复测量方差分析结果见表 1。结果显示，在编程学习成绩方面，任务难度与元认知支架不存在显著交互作用($F = 0.343$, $p = 0.710 > 0.05$)，说明任务难度和元认知支架并不会共同对学生编程学习成绩产生显著的影响。然而，任务难度（简单、中等和复杂）呈现出显著性($F = 4.167$，$p = 0.017 < 0.05$)，较小效应量($\eta^2_p = 0.05$)。这一结果说明任务难度的主效应存在，任务难度会对编程学习成绩产生显著影响；组别（对照组和实验组）呈现出显著性($F = 9.39$，$p = 0.003 < 0.01$)，较大效应量($\eta^2_p = 0.106$)。这一结果说明元认知支架的主效应存在，元认知支架会对学习成绩产生显著影响。

表 2 编程学习成绩事后多重比较结果

比较								
任务难度	组别		任务难度	组别	均值差	标准误	自由度	t
简单	对照组	−	简单	实验组	-8.591	3.01	79	-2.859*
		−	中等	对照组	-2.75	2.82	79	-0.977
		−	中等	实验组	-10.543	3.14	79	-3.355**
		−	复杂	对照组	-7.75	3.13	79	-2.478
		−	复杂	实验组	-12.982	3.45	79	-3.766**
	实验组	−	中等	对照组	5.841	3.15	79	1.857
		−	中等	实验组	-1.951	2.78	79	-0.702
		−	复杂	对照组	0.841	3.46	79	0.243
		−	复杂	实验组	-4.39	3.09	79	-1.421
中等	对照组	−	中等	实验组	-7.793	3.28	79	-2.378*
		−	复杂	对照组	-5	3.1	79	-1.613
		−	复杂	实验组	-10.232	3.57	79	-2.866**
	实验组	−	复杂	对照组	2.793	3.58	79	0.781
		−	复杂	实验组	-2.439	3.06	79	-0.797
复杂	对照组	−	复杂	实验组	-5.232	3.85	79	-1.36

*$p < .05$, **$p < .01$

如表 2 所示，对编程学习成绩进行事后多重比较分析发现，实验组学生的成绩均高于对照组，但是只有在解决简单($t = -2.859$，$p = 0.041 < 0.05$)和中等($t = -2.378$，$p = 0.023 < 0.05$)难度

任务时，实验组与对照组之间存在显著性差异。在解决复杂任务时，实验组与对照组无显著性差异(t = -1.36，p = 0.075 > 0.05)。这一结果表明在解决简单和中等难度的编程任务时，元认知支架对学生的编程学习成绩的影响更为显著。

3.2. 元认知倾向

为了了解不同的任务难度和元认知支架对学生元认知倾向的影响，本研究进行了二因子重复测量方差分析(Two-way Repeated Measures ANOVA)分析。如表 3 所示，在元认知倾向方面，任务难度与元认知支架有交互作用(F = 4.83, p = 0.03 < 0.05)，效应量为 0.02，表示效应量小。实验结果表示元认知支架会随着任务难度的不同对元认知倾向产生不同的影响。

表 3 元认知倾向二因子重复测量方差分析结果

	平方和	自由度	平均值平方	F	η^2_p
任务难度	0.25	1	0.25	1.92	0.01
组别	3.83	1	3.83	24.37***	0.112
任务难度*组别	0.64	1	0.64	4.83*	0.024
残差	25.52	193	0.13		

***p<.001, *p<.05

进一步分析元认知支架对元认知倾向的简单主效应，结果如表 4 所示。分析结果表明，任务难度对照组(F = 0.49，p = 0.49 > 0.05)及实验组(F = 1.420，p = 0.240 > 0.05)学生的元认知倾向均无统计学上的显著性差异。

表 4 元认知支架对元认知倾向的简单主效应分析结果

	任务难度	个数	平均数	标准差	调整后平均数	标准误	F
对照组	简单	40	3.32	0.38	3.32	0.03	0.49
	中等		3.43	0.32	3.43	0.03	
	复杂		3.42	0.41	3.42	0.04	
实验组	简单	40	3.39	0.3	3.39	0.03	1.42
	中等		3.55	0.3	3.55	0.03	
	复杂		3.66	0.49	3.66	0.05	

如表 5 所示，在简单难度任务 (F = 4.24, p = 0.04<0.05)、中等难度任务 (F = 9.98, p = 0.002 < 0.01)及复杂任务 (F = 16.47, p = 0.000 < 0.001) 中，元认知支架对学生元认知倾向具有显著性差异 (F = 4.24，p = 0.04 < 0.05)，且均为实验组优于对照组。

表 5 任务难度对元认知倾向的简单主效应分析结果

任务难度	组别	个数	平均数	标准差	调整后平均数	标准误	F	$\eta2$
简单	对照组	40	3.32	0.38	3.31	0.03	4.24*	0.021
	实验组	40	3.39	0.3	3.4	0.03		
中等	对照组	40	3.43	0.32	3.42	0.03	9.98**	0.049
	实验组	40	3.55	0.3	3.56	0.03		
复杂	对照组	40	3.42	0.41	3.41	0.04	16.47***	0.079
	实验组	40	3.66	0.49	3.67	0.05		

***p<.001, **p<.01, *p<.05

图 3 为元认知支架与任务难度对元认知倾向影响的交互作用图。图中显示，实验组学生的元认知倾向明显优于对照组。这说明随着任务难度的增加，元认知支架对学生的元认知倾向的影响更加显著。

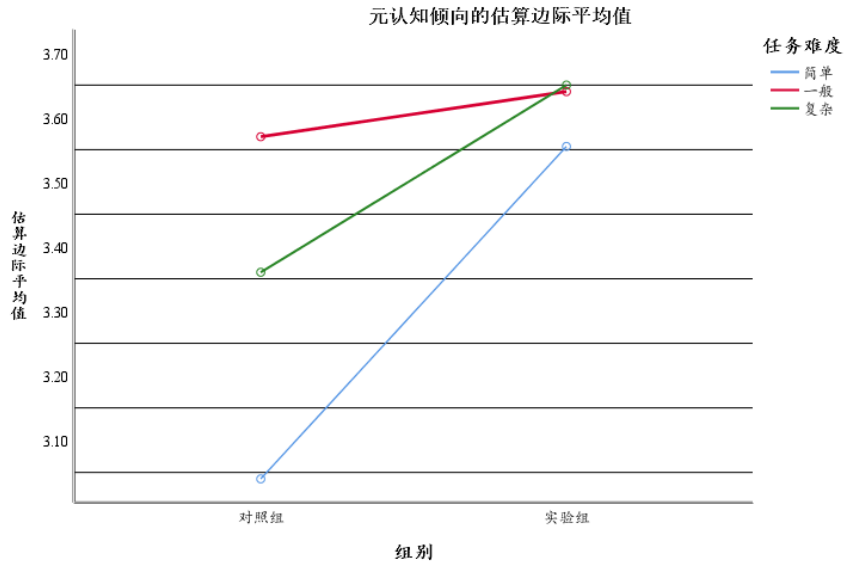


图 3 元认知支架与任务难度对元认知倾向影响的交互作用图

4.讨论与结论

本研究探讨了在协作编程学习情境下，元认知支架与任务难度对于学生协作编程表现的交互影响。研究结果表明，元认知支架在帮助学生解决简单和中等难度任务时对学习成绩的影响更加显著，且元认知支架与任务难度对编程学习成绩未产生显著的交互影响。此外，任务难度与元认知支架对学生的元认知倾向产生了显著的交互影响，且随着任务难度的增加，元认知支架越能够激发学生的元认知倾向。

4.1. 任务难度与元认知支架对学生的编程学习成绩产生怎样的交互影响？

结果表明，在简单及中等难度的任务中，实验组的编程学习成绩显著高于对照组。然而，在复杂任务中，实验组与对照组的学习成绩不存在统计学意义上的显著性差异。以往的研究认为学生的元认知在新的和相对困难的任务中表现得更为显著(Iiskala et al., 2004)，同时更高的元认知对于提升学生的协作编程表现具有显著的提升作用(Li et al., 2023b)。然而本研究的结论与该推论相反。任务难度与学习成就之间的最佳关系是在适度的挑战情境下，提供学习者更进一步思考、反思及评估自己的学习历程。因此，可能的原因是本研究所设计的复杂任务的难度太大。尽管学生非常专注于协作解决复杂的任务，但是由于任务难度太大，他们无法真正的解决问题（对照组：$M = 55.5$，$SD = 16.58$；实验组：$M = 58.5$, $SD = 19.81$）。所以，未来需要更多的研究探讨该问题。

4.2. 任务难度与元认知支架对学生的元认知倾向产生怎样的交互影响？

实验结果表明，任务难度与元认知支架对学生的元认知倾向产生了显著的交互影响，且随着任务难度的增加，元认知支架越能够激发学生的元认知倾向。研究指出当学习者通过元认知支架的引导，调整自我学习的思维模式，可以提升自我的元认知倾向(Goh, 2008)。而且任务难度增加时，学习者需要运用反思、评估等元认知策略来找出解决任务的方法(Mihalca et al., 2017)。任务的难度增加时，通常需要更高的认知能力和更复杂的策略才能成功完成。学习者往往会更加关注自己的思考过程，并更加审慎地评估自己的表现。这种察觉过程可以帮助他们理解自己的强项和弱点，并适应任务的要求(Zimmerman, 1990)。随着任务难度增加，学习

者会主动调整自己的认知历程及使用更复杂的策略来完成任务(Rivers, 2021)。因此，随着任务难度的提高，元认知支架更能够激发学生的元认知倾向。

此外，本研究也存在一些局限性。首先，本研究关注了元认知支架和任务难度对初中生的协作编程的影响，结论无法扩展到其他学段。未来的研究可以关注该方法对小学生或者大学生编程学习的影响。其次，本研究是一个量化的研究，没有收集和分析质性的数据。未来的研究可以采取混合的研究设计。

参考文献

Beck, L. L., & Chizhik, A. W. (2013). Cooperative learning instructional methods for CS1. *ACM Transactions on Computing Education*, *13*(3), 1–21. https://doi.org/10.1145/2492686

Cho, K., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, *50*(3), 5–22. https://doi.org/10.1007/bf02505022

Goh, C. C. M. (2008). Metacognitive instruction for second language listening development. *RELC Journal*, *39*(2), 188–213. https://doi.org/10.1177/0033688208092184

Hadwin, A. F., & Oshige, M. (2011). Self-Regulation, Coregulation, and Socially Shared Regulation: Exploring Perspectives of Social in Self-Regulated Learning Theory. *Teachers College Record*, *113*(2), 240–264. https://doi.org/10.1177/016146811111300204

Hanks, B., Fitzgerald, S., McCauley, R., Murphy, L., & Zander, C. (2011). Pair programming in education: a literature review. *Computer Science Education*, *21*(2), 135–173. https://doi.org/10.1080/08993408.2011.579808

Huertas, A., López, O., & Sanabria, L. (2016). Influence of a metacognitive scaffolding for information search in B-learning courses on learning achievement and its relationship with cognitive and learning style. *Journal of Educational Computing Research*, *55*(2), 147–171. https://doi.org/10.1177/0735633116656634

Iiskala, T., Vauras, M., & Lehtinen, E. (2004). Socially-shared metacognition in peer learning? *Hellenic Journal of Psychology*, *1*(2), 147-178.

Iiskala, T., Vauras, M., Lehtinen, E., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction*, *21*(3), 379–393. https://doi.org/10.1016/j.learninstruc.2010.05.002

Kwon, K., Hong, R., & Laffey, J. M. (2013). The educational impact of metacognitive group coordination in computer-supported collaborative learning. *Computers in Human Behavior*, *29*(4), 1271–1281. https://doi.org/10.1016/j.chb.2013.01.003

Lai, C.-L., & Hwang, G.-J. (2014). Effects of Mobile Learning Participation Time on High School Students' 21st Century Core Competences. 2014 International Conference of Educational Innovation through Technology.

Li, W., Liu, C.-Y., & Tseng, J. C. R. (2023a). Development of a metacognitive regulation-based collaborative programming system and its effects on students' learning achievements, computational thinking tendency and group metacognition. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13358

Li, W., Liu, C.-Y., & Tseng, J. C. R. (2023b). Effects of the interaction between metacognition teaching and students' learning achievement on students' computational thinking, critical

thinking, and metacognition in collaborative programming learning. *Education and Information Technologies*, *28*(10), 12919–12943. https://doi.org/10.1007/s10639-023-11671-2

Lin, Y., Wu, C., & Chiu, C. F. (2018). The use of Wiki in teaching programming. *International Journal of Distance Education Technologies*, *16*(3), 18–45. https://doi.org/10.4018/ijdet.2018070102

Mihalca, L., Mengelkamp, C., & Schnotz, W. (2017). Accuracy of metacognitive judgments as a moderator of learner control effectiveness in problem-solving tasks. *Metacognition and Learning*, *12*(3), 357–379. https://doi.org/10.1007/s11409-017-9173-2

Ouyang, F., Chen, S., Yang, Y., & Yun-Qing, C. (2021). Examining the effects of three Group-Level metacognitive scaffoldings on In-Service teachers' knowledge building. *Journal of Educational Computing Research*, *60*(2), 352–379. https://doi.org/10.1177/07356331211030847

Rivers, M. L. (2020). Metacognition About Practice Testing: a Review of Learners' Beliefs, Monitoring, and Control of Test-Enhanced Learning. *Educational Psychology Review*, *33*(3), 823–862. https://doi.org/10.1007/s10648-020-09578-2

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, *88*(3), 345–372. https://doi.org/10.1002/sce.10130

Shadiev, R., Hwang, W., Yeh, S., Yang, S. J. H., Wang, J. L., Lin, H., & Hsu, G. L. (2014). Effects of Unidirectional vs. Reciprocal Teaching Strategies on Web-Based Computer Programming Learning. *Journal of Educational Computing Research*, *50*(1), 67–95. https://doi.org/10.2190/ec.50.1.d

Wang, C., Gao, B., & Chen, S. (2023). The effects of metacognitive scaffolding of project-based learning environments on students' metacognitive ability and computational thinking. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12022-x

Wei, X., Lin, L., Meng, N., Tan, W., Kong, S., & Kinshuk, K. (2021). The effectiveness of partial pair programming on elementary school students' Computational Thinking skills and self-efficacy. *Computers & Education*, *160*, 104023. https://doi.org/10.1016/j.compedu.2020.104023

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

“学生生成构想”教学理念应用于混合学习的历程分析

Analysis of a Blended Learning Lesson Underpinned by the Pedagogical Approach of Student-Generated Ideas (SGI)

Lung-Hsiang Wong[1*], Thi Thu Ha Nguyen[1], Chew Lee Teo[1], Peter Seow[1]
[1]新加坡南洋理工大学国立教育学院教学法与实践研究中心
[*]lhwong.acad@gmail.com

【摘要】 本定性研究旨在通过“学生生成构想”（SGI）教学模式提升前线教师混合学习的实践。在新加坡小学五年级执行 SGI 混合学习课，通过媒介导向分析法解析学生学习历程，促成构想演化，实现深度学习及培养 21 世纪技能。研究发现混合学习中媒介与构想演化呈现多样性，并通过图示对比社群文化和认知主义视野下的学习历程。讨论中提出教学设计原则，如善用科技，强化课内外学习及线上与线下学习的联结。结论强调 SGI 模式与混合学习的理想契合，为教师提供培训-实践-反思路径，优化学生以构建知识为中心的学习历程。
【关键词】 “学生生成构想”模式；混合学习；媒介导向分析；定性分析；教师专业发展

Abstract: *This qualitative study aims to enhance teachers' practices in blended learning through the "Student-Generated Ideas" (SGI) teaching model. Implemented in a 5th-grade class in a Singaporean primary school, the research utilised the Artefact-Oriented Analysis (AOA) method to analyse students' learning processes, facilitating idea evolution, achieving deep learning, and fostering 21st-century skills. Findings indicate diverse interactions between media and idea evolution, illustrated through visual comparisons of socio-cultural and cognitive perspectives. Design principles, such as incorporating authentic tasks in out-of-class activities are subsequently proposed.*

Keywords: “Student-Generated Ideas” model; blended learning; artefact-oriented analysis; qualitative analysis

1. 前言

以“结合面对面教学和电脑中介教学”为核心的混合学习 (Cronje, 2020)，着重弹性切换线上线下教学模式，提供学生自步学习及善用科技学习的机会。新加坡自 2021 年起在所有中学实施混合学习。但教师在执行混合学习时，若缺乏建构或认知主义考量，恐流于形式，仍以教师为中心，在线学习仍是吸收数码内容或操练作业 (Bruggeman et al., 2021)。解决方案是从教育理论推导出教学法，协助教师提升教学。我们提出“学生生成构想”(Student-Generated Ideas; SGI) (Wong et al., 2021)，引导学生以假设、建议、提问、解释、制作作品等，贡献个人或小组构想，置于社群学习空间，产生思想碰撞，优化构想。

本研究在新加坡一个小学五年级班执行以 SGI 指引，为期三天的科学混合学习课，主题为“云朵的形成”。通过定性分析方法“媒介导向分析”(artefact-oriented analysis) (Wong et al., 2012)，我们通过解析贯穿课时内外、线上线下、个人与社群学习情境的五阶段学习历程，探究如何结合善用各种学习空间里的数码科技或非科技工具的赋能(affordances)，来执行多元化的学习任务，促成学生构想的演化(idea evolution)，达致深度学习。

2. “学生生成构想”（SGI）教学模式及媒介导向分析

SGI 教学模式由新加坡学习科学研究群（Lam et al., 2016)分析多种教学模式如知识建构、无缝学习、探究式学习、问题导向学习、创客、计算思维等，提炼出“最大公约数”，

即由教师设定真实的学习语境，激发学生生成并演化各种形式的“构想”。SGI 模式的理论依据是概念转变 (Hewson, 1981)。学生最初提出的构想称为初始构想，通过 SGI 学习活动，逐渐演化成最终构想。初始构想和最终构想之间可能生成多个中间构想。SGI 模式意在推动学生“构想生成及演化”历程。SGI 的另一个关键元素为元构想，即构思演化时的各种可能认知手段如解释、评估、具体化、抽象化、泛化、建模、改模、联结、整合、应用、调适等。

媒介导向分析法由 Wong et al. (2012)提出，解析学习者在学习历程中的认知行为与“媒介”(artefacts)的互动，如何影响学习任务轨迹。mediation by artefacts (Vygotsky, 1986)理念认为人类活动须通过特定媒介的“中介”(mediation)来达致。学习活动媒介可包括教师设计的任务、教师和同侪资源、学科内容及符号性工具如语言、电脑软件等 (Liang, 2009)。媒介也有在学习历程中临场生成的，如完成任务的好点子、学生作品的半成品、小组讨论内容等。学生可结合使用即有和现场生成的媒介，中介活动。但一些媒介在不同状况下，对学生可能有正面或负面影响，如滥用、造成误导或使学生分心等，称为 “局限性中介”。

媒介导向分析法（artefact-oriented analysis; AOA）以定形分析产生图示呈现个人或群体学习历程，解析学生在每一个学习任务中所调动或生成的媒介如何交互作用，“联合中介”此任务。媒介分四大类：学科媒介（含学科学习目标、教材、教师设计的教学任务等）、物质媒介（物质工具含科技工具、环境、人等）、社群认知媒介（如学生的临场构思、教师的临场支持等）和产出媒介（学生最终生成的物质或非物质媒介；可能在后续任务中成为输入媒介，中介下一个任务）。此分析法意在找出能提升学习任务执行的媒介组合，及哪些媒介有潜在的干扰性中介作用，让教师在后续教学设计改进时，有了兴利除弊的依据。

3. 研究设计及分析

新加坡一所小学五年级班 37 名学生参与研究。教师 TC（假名）接受有关 SGI 的培训，然后与团队合作设计两个 SGI 式混合学习教案。本文分析第一个教案“云朵的行成”的执行历程，第一天上午为一小时的实体课（活动 1A 和 1B）；放学后至第二晚为居家在线活动 2；第三天上午为另一个小时的实体课（活动 3A 和 3B）。以下是教学流程：**(实体课活动 1A)**学生五人一组，先独立思考，不得参考课文或讨论，个别在所分配的小组麻将纸上写下“什么是云朵？”和“云朵如何行成？”的答案，并发出一个新提问。**(实体课活动 1B)**学生小组讨论，回答同侪提问，并整合解答，利用平板电脑输入 Student Learning Space (SLS)的 Interactive Thinking Tool (ITT)上[1]。**(居家在线活动 2)**学生在家中上 SLS 参考各组电子便利贴，联结生活经验，进行反思，把至迄今个人对于“云的形成”的观念贴上 SLS。**(实体课活动 3A)**小组在之前所生成、演化的构想（云朵成份和如何形成的理解）指引，以 TC 提供的设备（烧杯、热水、冰块、保鲜纸）和基本指示进行在烧杯里“生成云朵”（悬浮的白色雾气）的实验；但指示刻意不完整，让小组试错调整实验设置。**(实体课活动 3B)**TC 带领全班讨论，联结学科知识和各个小组的实验成果。之后各个小组总结各自的新理解，输入 Padlet。

本研究所收集的学习历程数据包括：一、实体课录像及录音；二、各种学生初始构想、中间构想和最终构想（麻将纸上的观点和提问、贴文、实验设计等）；三、师生访谈。我们对数据进行 SGI 和 AOA 定性编码（SGI 编码找出初始构想、演化构想、最终构想，和联结构想的元构想的类型；AOA 编码则找出各阶段学生在生成及演化构时利用的促成或干扰此历程的媒介），整合成图示以显示整个学习历程及对之进行诠释。

[1] SLS 由新加坡教育部开发，全国所有公立中小学必须使用的线上学习平台。平台上的 ITT（互动思维工具）让教师能搭建引导学生思考历程的鹰架。

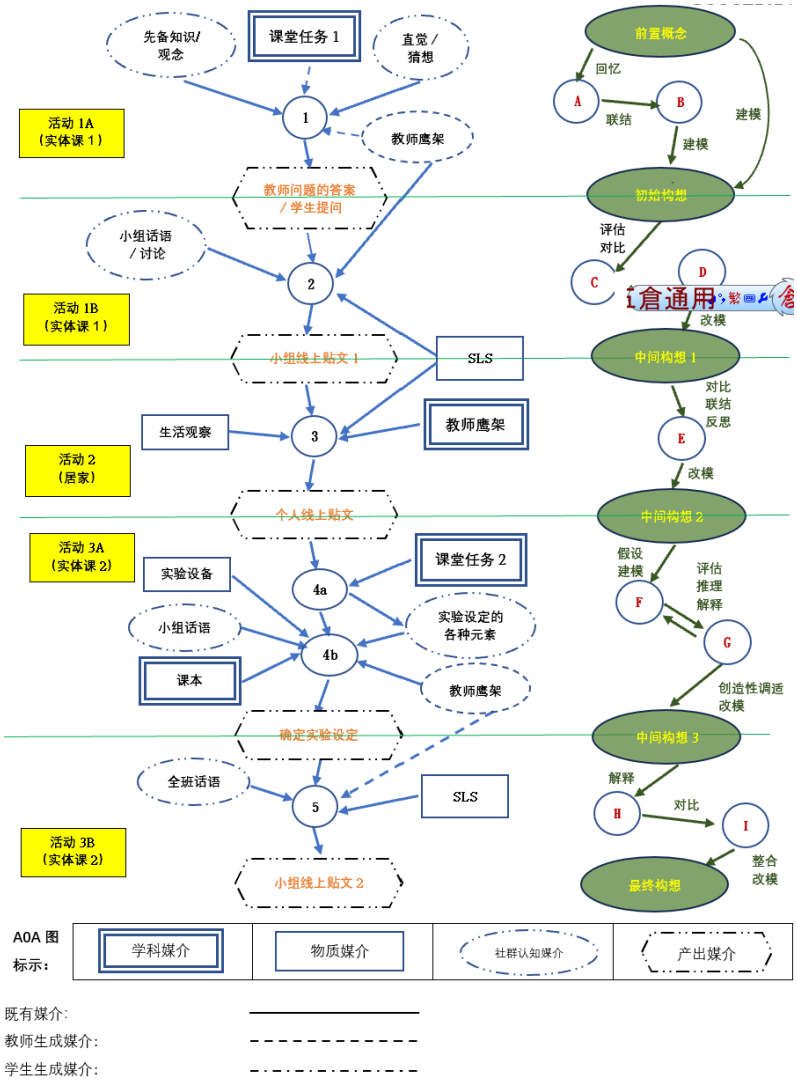


图 1 本次学习历程的 AOA 图和 SGI 图（左图的箭头表示中介作用；虚线箭头表示局限性中介）

4. 研究发现

图 1 并列我们整理而成的媒介导向分析图（AOA 图）和对应的学生构想生成及演化历程图（SGI 图）。图 1A 则显示 AOA 图的各种标示（不同媒介类型）。全图按上述五个活动划分，一个活动的产出媒介会输入下一个活动。这些产出媒介对应到 SGI 框架，便是初始、中间或最终构想。AOA 图里的圆形节点 1 至 8 代表联合中介的发生点，下称“JM (joint mediation)节点”。而 SGI 图里的节点 A 至 H 代表学生心智里的中间概念状态，下称“CS (conceptual state)节点”；联结 CS 状态和各种构想的箭头分别代表不同的元构想。篇幅所限，以下只挑选说明活动 1A 和活动 2。

根据图 1，活动 1A 显示 TC 预设的学习任务，及个别学生执行此学习任务时所调动的先备知识/观念，或他们的直觉/猜想，联合中介（JM 节点 1）了活动 1，在麻将纸上写下答案及新提问作为产出媒介。但任务对于 JM 节点 1 的作用是局限性中介，因为少数事先预习课文的学生在课后访谈中表示任务太简单，索然无味。此外，教师鹰架仍在活动的最后三分钟，多数学生仍未写下答案和提问时介入，因为 TC 为避免超时，赶着提供各种加快任务的指示，只给他们一分钟思考，可能干扰了学生的思路，被列为局限性媒介。

从 SGI 的角度来说，活动 1A 是学生从调动相关前置概念，到初始构想成形并以文本形式在麻将纸上表达的历程，含两个可能路径：（一）对于已有先备知识/观念的学生来说，他们先回忆前置概念（CS 节点 A），联结此任务课题（CS 节点 B），再建模形成初始构想（他们针对云的成份与如何形成的解释，可被视脑子里生成的科学模型）；（二）对于没有相关先备知识的学生来说，他们是临场在心智里生成直觉/猜想，并建模而形成初始构想。

居家活动 2，原先的小组贴文、教师鹰架（学科媒介）、学生生活环境（物质媒介）、SLS（物质媒介），联合中介 JM 节点 3 而产出个人线上贴文（社群认知媒介）。此教师鹰架是 TC 事先交待学生在贴文时应填入的句型结构（“我本来认为……现在我认为……”）（见图 2 的部分例子）——相对于在其他活动中的鹰架乃教师临场反应（社群文化媒介），此鹰架是事先设计的，属于学科媒介。少数学生受访时忆述在课余观察相关生活现象，联结到此居家学习任务，从而支持或修正他们的个人线上贴文。这是原本不在教师设想中的意外发现，因为学生身处在家中而对这些环境现象有更高敏感度。对应 SGI 图，学生对比和联结不同贴文、反思（CS 节点 E），促成部分学生再次改模，生成中间构想 2（个人贴文）。

5. 讨论与结论

本研究的创新意义在于整合了 AOA 分析和 SGI 分析这两种定性分析方法，对比社群文化视野和认知主义视野下的一个混合学习历程——前者着重分析学生与他人、环境的互动如何促成、推进或限制学习；后者聚焦学生心智中的思考历程。

我们在研究中发现的一个不在设计教案时预设的媒介应用，是居家活动 2 时少数学生观察生活现象，联结教学内容后演化构想，因而个人线上贴文更有深度。传统意义的混合学习的课外学习环节是纯在线的，但新加坡推的混合学习模式把课后学习扩展到线下学习，与实体环境互动，达致真实性学习。但一般的新加坡教师设计及执行混合学习，往往忽略这种可能性，课后学习仍把学生完全框在纯在线的任务。此发现可导出教学设计原则：“在课后学习环节中，设计与实体生活环境互动的相关任务”这就能确保所有学生都实现真实性学习。

此外，活动 1A 时 TC 所提供的教师鹰架被列为局限性媒介，这是我们的学术判断，因为它有悖 SGI 精神，即过早结束仍有可为的构想演化历程。但从实践角度来看，应与课时限制有关。这是 SGI 教学模式与教学资源（包括课时）之间的拉锯。我们推导出另一条教学原则“在学生讨论时，尽可保持开放性，避免过早提供‘权威’观点”，让教师有意识地延长讨论。但我们也理解现实挑战。科技和混合学习或可寻找平衡，如活动 1A 可改为课前居家任务。

以上例子也说明本研究一大目的：教师通过 SGI 教学模式的培训-实践-反思历程，提升保持教学开放性的意识，争取在有限资源下，将学生知识共建历程最大化。与 SGI 模式相得益彰的混合学习，用科技突破课时掣肘，让学生的知识共建开花结果，或许是理想的方案。

参考文献

Bruggeman, B., Tondeur, J., Struyven, K., Pynoo, B., Garone, A., & Vanslambrouck, S. (2021). Experts speaking: Crucial teacher attributes for implementing blended learning in higher education. *The Internet and Higher Education*, *48*, 100772.

Cronje, J. (2020). Towards a new definition of blended learning. *Electronic journal of e-Learning, 18*(2), 114-121.

Hewson, P. W. (1981). A conceptual change approach to learning science. *European Journal of Science Education*, *3*, 383-396.

Lam, R., Wong, L.-H., Gaydos, M., Huang, J. S., Kapur, M., Seah, L. H., Tan, M. B., Catherine, & Sandoval, W. A. (2016). Designing learning contexts using student-generated ideas. Proceedings of the International Conference of the Learning Sciences (ICLS) 2016, Singapore.

Liang, X. (2009). Investigating mediations in student activities in an English immersion context in mainland China. *English Language Teaching*, *2*(4), 38-50.

Vygotsky, L. (1986). *Thought and Language*. MIT Press.

Wong, L.-H., Chen, W., & Jan, M. (2012). How artefacts mediate small group co-creation activities in a mobile-assisted language learning environment? *Journal of Computer Assisted Learning*, *28*(5), 411-424.

Wong, L.-H. T., Chew Lee, Ogata, H., Song, Y., Wu, L., & Yu, F.-Y. (2021). Leveraging Student-Generated Ideas (SGI) to facilitate socio-constructivist learning and conceptual change: The roles of technology in SGI learning trajectories. *Proceedings of 29th International Conference on Computers in Education (ICCE 2021)* (pp.777-782), (online).

基于元认知支架的协作编程学习对不同性别学生学习成绩和元认知倾向的影响

Effects of metacognitive scaffold-based collaborative programming learning on academic achievement and metacognitive tendencies of students of different genders

周沫[1]，李伟[1*]，曾秋蓉[2]，黄佶艺[3]，刘海英[1]，季嘉炜[1]，刘城烨[4]
[1]温州大学 STEM 教育研究中心
[2]中华大学资讯工程学系
[3]温州市瓯海区瞿溪第一小学
[4]香港中文大学课程与教学系
[*]liweiwzu@wzu.edu.cn

【摘要】 协作编程学习旨在为学生提供与同龄人互动的机会，通过利用彼此资源和技能来促进编程技能的提升。研究者强调了引导学生进行反思和调节的重要性。因此，为提高学生的元认知能力，本研究提出了基于元认知支架的协作编程学习方法。此外，性别差异是影响协作学习的基本因素。不同性别的学生在认知、学习策略和解决问题方式上可能存在差异。因此，为了解元认知支架对不同性别学生协作编程的影响，在小学信息科技课堂中进行准实验研究。结果显示，基于元认知支架的协作编程学习方法显著提高了男生的学习成绩，且提高了所有学生的元认知倾向。

【关键词】 协作编程；元认知支架；性别差异

Abstract: *Collaborative programming learning provides students with the opportunity to interact with peers and promote programming skills by leveraging each other's resources and skills. The researchers stressed the importance of guiding students to reflect and adjust. Therefore, to improve students' metacognitive ability, this study proposes a collaborative programming learning method based on metacognitive scaffolds. In addition, gender differences are fundamental factors affecting collaborative learning. Therefore, to solve the effects of metacognitive scaffolds on the collaborative programming of students of different genders, a quasi-experimental study was conducted in the primary school information technology classroom. The results showed that the collaborative programming learning method based on metacognitive scaffolds significantly improved the academic performance of male students and the metacognitive tendency of all students.*

Keywords: *Collaborative programming, Metacognitive scaffold, Gender difference*

1.引言

近年来，K-12 年级计算机科学教育的重要性得到加强，计算思维的培养与重要性也成为教师与学者们关注的话题(Chen et al., 2017)。计算思维有助于培养学生的逻辑思维、算法能力，并提升他们在解决日常任务时的分析技能(Veenman et al., 2022)。研究表明，编程语言学习可以培养学生的计算思维(Tikva & Tambouris, 2021)，因为编程与学生语言表达方式接近，有助于降低他们的认知负荷，使其更容易获得计算思维和算法思维(Kelleher & Pausch, 2005)。然而，编程的复杂性使得许多学生难以实现高质量的编程(Hwang et al., 2012)。为解决这个问题，学者们发现协作编程作为一种计算机支持的协作学习模式，可以帮助学生共同解决编程问题，

促进学习交流和思想交流，提高学习效果和理解能力(Zheng, 2021)。然而，并非所有学生都能从协作编程中受益，有些学生可能缺乏元认知知识和能力，无法反思和监控自己的学习过程，导致协作学习表现不佳(Yao et al., 2020)。且计算思维是一个认知过程，受到元认知实践的监管(Allsop, 2019)，可以认为元认知对计算思维有影响(Sung, 2016)。因此，加入元认知支架的协作编程学习方法可以帮助解决这些问题，提升学生的计算技能、问题解决能力和自我调节学习能力(Kramarski & Mizrachi, 2006)。

此外，性别差异也是影响协作学习的因素(Ma et al., 2022)，不同性别的学生在认知、学习策略和问题解决方式上可能存在差异，这可能对协作学习的过程和结果产生影响。了解性别差异对学习表现的影响有助于教师和学者设计更有效的教学策略。

基于此，本研究提出了一种基于元认知支架的协作编程学习方法，并进行一项准实验，从不同性别学生的学习成就和元认知倾向方面来评估其有效性。研究问题如下：

1. 与传统的协作编程(CCC)学习方法相比，基于元认知支架的协作编程(MCC)学习方法如何影响不同性别学生的学习成绩？

2. 与 CCC 学习方法相比，MCC 学习方法如何影响不同性别学生的元认知倾向？

2.方法

2.1 实验对象

本研究以台湾 214 名小学一年级学生为实验对象。随机将一个班级分为实验组（男 53，女 51），采用基于 MCC 方法进行学习；另一个是对照组（男 50，女 60），采用 CCC 学习方法。

2.2 实验设计

本研究持续了 10 周，每周 1 课时，每次课程 45 分钟。本研究的学习系统参考了 Li et al. (2023)学者的研究。在实验组中，学生首先设置学习目标，然后进行小组讨论和协作学习，共同撰写流程图和代码。学习系统会监督学生的协作过程，并提醒他们监测自己的学习。任务结束后，学生进行自我反思并调整学习策略或目标。对照组在教师说明任务后，学生共同讨论并开始协作编写流程图和代码。第一周和第十周，学生完成学习成绩前测和后测以及元认知倾向前测和后测问卷。

2.3 测量工具

本研究采用的测量工具包括学习成绩前后测和元认知倾向的前后问卷。

学习成绩前测试题为国际运算思维能力测验(Bebras)，共 10 题选择，满分为 100 分。后测题目由具有丰富教学经验的国中编程教学老师所编写，共 10 题选择，满分为 100 分。后测旨在评估学生在不同的教学方法下，对编程方面知识的理解与应用能力。

元认知倾向问卷是根据 Lai and Hwang (2014)提出的测量工具应用而来。该问卷采用 5 点李克特量表（1=完全不同意；5=完全同意），包括五个问题，Cronbach's Alpha 系数为 0.815。

3.结果

3.1 学习成绩

为了解不同学习方法对不同性别学生学习成绩的影响，本研究采用双因素协方差分析方法。结果如表 1 所示，学习方法对学生的学习成绩有显著影响($F = 5.399$, $p = 0.021 < 0.050$)。此外，性别与学习方法间的交互作用对学生学习成绩也有显著影响($F = 4.070$, $p = 0.045 < 0.050$)。

表 1 学习成绩双因素协方差分析结果

变量	离均差平方和	自由度	平均值平方	F	p	η^2
学习成绩前测	112.514	1	112.514	0.598	0.440	0.003
性别	312.584	1	312.584	1.661	0.199	0.008

学习方法	1016.309	1	1016.309	5.399	0.021*	0.025
性别*学习方法	766.185	1	766.185	4.070	0.045*	0.019
误差	39340.934	209				

*$p < 0.05$

为了探讨不同性别对采用不同学习方法学习的学生学习成绩的影响，本研究进一步采用了简单主效应分析方法。结果发现，对照组中不同性别学生学习成绩具有显著的统计学差异($F = 5.591$, $p = 0.019 < 0.050$, $\eta^2 = 0.026$)。实验组中不同性别学生学习成绩之间没有显著性差异($F = 0.258$, $p = 0.612 > 0.050$, $\eta^2 = 0.001$)。这些结果表明，在采用 MCC 学习方法进行学习时，不同性别学生的学习成绩没有统计学意义上的显著差异。

本研究还探究了学习方法对不同性别学生学习成绩影响的简单主效应分析结果。结果发现，男生在不同学习方法下的学习成绩具有显著的统计学差异($F = 9.120$, $p = 0.003 < 0.010$, $\eta^2 = 0.042$)。而使用不同学习方法学习的女生之间不存在统计学上的差异($F = 0.049$, $p = 0.825 > 0.050$, $\eta^2 = 0.000$)。结果证明，在 MCC 学习方法指导下，男生比女生能更多的提高学习成绩。

图 1 显示了不同的学习方法和不同性别与学生学习成绩之间的交互作用。结果表明，采用 MCC 方法学习的学生学习成绩显著优于使用 CCC 方法学习的学生。此外，实验组男生的学习成绩显著优于对照组的男生，而女生使用两种方法下，学习成绩差异不大。也就意味着 MCC 方法对学生的学习成绩有一定的积极影响。

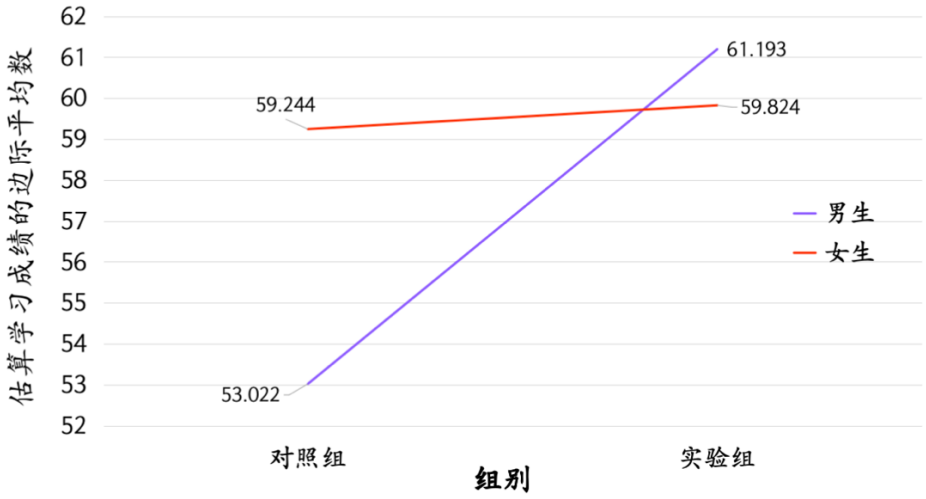


图 1 学习方法与性别对学习成绩的交互作用

3.2 元认知倾向

为了了解不同学习方法对不同性别学生元认知倾向的影响，本研究采用了双因素协方差分析方法。结果显示，性别与学习方法之间的交互作用对学生的元认知倾向没有显著影响($F = 0.000$, $p = 0.991 > 0.050$)。本研究进一步采用了主效应分析方法。结果显示，学习方法对学生的学习元认知倾向有显著影响($F = 4.370$, $p = 0.038 < 0.050$)。但性别对学生的元认知倾向没有显著影响($F = 3.745$, $p = 0.054 > 0.050$)。结合统计描述结果，可以知道采用 MCC 学习方法的学生，元认知倾向显著高于采用 CCC 方法学习的学生。

4.讨论和结论

本研究比较了 MCC 学习方法和 CCC 方法对学生学习效果的影响。结果显示，MCC 学习方法在提高学习成绩和元认知倾向方面取得了显著效果。这与以往研究的结果一致，表明采用元认知支架可以使学生在协作编程的学习环境中受益(Li et al., 2023)。

4.1 基于 MCC 学习方法如何影响不同性别学生的学习成绩？

本研究发现使用 MCC 方法可以显著提高学生的学习成绩。这与先前研究结果一致(Li et al., 2023)。该方法通过让学生在协作中互相讨论和分享编程思路，帮助他们发现问题并进行自我反思和修改，从而提高学习成绩。此外，研究还发现在 CCC 学习方法下，女生的学习成绩优于男生，但在 MCC 学习方法下，男女生之间没有显著差异。这可能是因为女生在协作和沟通方面更擅长，而男生在提供明确元认知支架的情况下能提升学习自觉性和自我反思能力。

4.2 基于 MCC 学习方法如何影响不同性别学生的元认知倾向？

本研究发现，采用 MCC 学习方式学习的学生，在元认知倾向方面显著高于采用 CCC 方式的学生。这种学习方式能够引导学生进行思考、监控学习进度、评估学习结果和反思学习过程，从而提高学生的分析思维、判断能力、学习动机和自主学习能力。此外，元认知学习还能帮助学生理解问题、调整学习策略、监控自我学习进展和进行自我评估，从而更好地应对问题解决过程中的挑战。

参考文献

Allsop, Y. (2019). Assessing computational thinking process using a multiple evaluation approach. *International Journal of Child-Computer Interaction*, *19*, 30-55.

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. *Computers & Education*, *109*(JUN.), 162-175.

Hwang, W. Y., Shadiev, R., Wang, C. Y., & Huang, Z. H. (2012). A pilot study of cooperative programming learning behavior and its relationship with students' learning performance. *Computers & Education*, *58*(4), 1267-1281.

Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers [Article]. *ACM Computing Surveys*, *37*(2), 83-137. https://doi.org/10.1145/1089733.1089734

Kramarski, B., & Mizrachi, N. (2006). Online Discussion and Self-Regulated Learning: Effects of Instructional Methods on Mathematical Literacy. *Journal of Educational Research*, *99*(4), 218-231.

Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem-solving, meta-cognitive awareness and creativity. *International Journal of Mobile Learning & Organisation*, *8*(3/4), 276.

Li, W., Liu, C.-Y., & Tseng, J. C. R. (2023). Development of a metacognitive regulation-based collaborative programming system and its effects on students' learning achievements, computational thinking tendency and group metacognition [Article; Early Access]. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13358

Ma, L., Luo, H., Liao, X., & Li, J. (2022). Impact of Gender on STEAM Education in Elementary School: From Individuals to Group Compositions [Article]. *Behavioral Sciences*, *12*(9), Article 308. https://doi.org/10.3390/bs12090308

Sung, W. (2016). Incorporating Touch-based Tablets into Classroom Activities: Fostering Children's Computational Thinking through iPad Integrated Instruction.

Tikva, C., & Tambouris, E. (2021). Mapping computational thinking through programming in K-12 education: A conceptual model based on a systematic literature Review [Review]. *Computers & Education*, *162*, Article 104083. https://doi.org/10.1016/j.compedu.2020.104083

Veenman, K., Tolboom, J. L. J., & van Beekum, O. (2022). The relation between computational thinking and logical thinking in the context of robotics education [Article]. *Frontiers in Education*, *7*, Article 956901. https://doi.org/10.3389/feduc.2022.956901

Yao, L. U., Mao, X., Wang, T., Yin, G., & Zude, L. I. (2020). Improving students' programming quality with the continuous inspection process: a social coding perspective. *Frontiers of Computer Science*, *14*(5).

Zheng, L. (2021). Improving Programming Skills Through an Innovative Collaborative Programming Model: A Case Study.

融合生成式 AI 的人机协同论证活动设计与应用

Design and Application of Human-Computer Collaborative Argumentation Activities Integrated Generative AI

崔鑫[1*]，尤欣雅[2]，马志强[2]
[1]江南大学设计学院
[2]江南大学教育学院 江苏省“互联网+教育”研究基地
[*] cuixin960220@163.com

【摘要】 传统协同论证情境中，小组成员难以从多个视角生成支持自己立场的观点，也缺乏批判性地分析、反驳论点的能力。而 ChatGPT 作为一种基于大语言模型的生成式聊天机器人，具备与学生进行人机协同论证的潜能，然而少有研究在真实的教学情境中开展人机协同论证实践并实证检验其效果。基于此，本研究设计了融合生成式 AI 的人机协同论证活动，并以准实验方法检验了该活动对学生论证对话的影响。研究发现，融合生成式 AI 的人机协同论证活动能够提升结构性论证要素的质量，完善论证结构，促进论证对话的社会性交互。
【关键词】 协同论证；生成式人工智能；人机协同论证；提示词；ChatGPT

Abstract: *In the traditional collaborative argumentation context, it is difficult for group members to generate ideas from multiple perspectives to support their own positions, and they also lack the ability to critically analyze and refute the arguments. ChatGPT, as a chatbot based on a large language model, has the potential to engage in collaborative human- computer argumentation with students, but few studies have practiced and empirically examined the effectiveness of collaborative human-computer argumentation in real teaching contexts. This study designed a human-computer collaborative argumentation activity of integrated generative AI and examined the impact of the activity on students' argumentation dialogues using a quasi-experimental approach. It was found that the human-computer collaborative argumentation activity of integrated generative AI can enhance the quality of structural argumentation elements, improve argumentation structure, and promote social interaction in argumentation dialogues.*
Keywords: collaborative argumentation, generative artificial intelligence, human-computer collaborative argumentation, prompt, ChatGPT

1. 引言

协同论证是一种有效的教学法，使学生在多方的对话中自然地从不同的角度考虑和权衡论点，澄清其不确定性，以实现对领域关键问题的理解(Noroozi et al., 2012)。然而在实际课堂中，小组成员由于先验知识、认知资源的限制，难以从多个视角生成支持自己立场的观点，也缺乏批判性地分析、反驳论点的能力(Chin & Osborne, 2010)。为了支持协同论证，现有研究开始将聊天机器人引入协同论证中，用于激发学生论点生成、考虑与整合反论点(Guo et al., 2023)。然而，现有的聊天机器人大多依赖于检索模型，这些模型通过预设的观点与学生互动，往往缺乏生成针对上下文需求的定制化、连贯响应的能力（Guo et al., 2022）。近期，以 ChatGPT 为代表的生成式人工智能技术取得了显著突破。ChatGPT 是一种基于大语言模型的生成式聊天机器人，其背后的大语言模型训练了大量公开可用的数字内容数据，能够理解并根据用户的提问生成响应，并在广泛的主题上保持连贯的对话，并近乎令人信服(Chan & Hu, 2023)，因而也具备了针对某一社会性议题进行论证的能力。因此，教育界鼓励学生与生成式 AI 开展人机协同论证，从而提高论证的质量与效率(Halaweh, 2023)。已有研究开始探讨人机协同论证的可能性，论述了生成式 AI 在支持学生生成论点、分析与批判反论点方面的潜能(Cress &

Kimmerle, 2023)，但少有研究在真实的教学情境中开展人机协同论证实践并实证检验其效果。基于此，本研究设计了融合生成式 AI 的人机协同论证活动，并在活动中聚焦于提示的设计，再以准实验方法检验了该活动对学生论证对话的影响。

2. 文献综述

2.1. 协同论证

协同论证的本质是基于理性的，对话性的、社会性的活动。面对其他论证主体提出的论点与依据，学生需要判断对方的观点，以确定是赞同还是反驳，从而将多方的论点整合起来。协同论证的理性本质通过论证结构来体现，社会性本质通过论证对话来体现。在论证结构上，研究者探索了协同论证过程的规范性指导与协同论证结果的结构性表征。在协同论证过程的规范性指导方面，代表性的教学法为知识整合教学法，学习者要经历析出观点、添加观点、辨分观点与反思观点四个阶段，这种教学法引导协作成员以对话的形式，对小组成员自己提出的观点和教师或其他代理提供的观点中众多相似或矛盾的论点进行辨分、协商，以建立对复杂问题的共同理解。在协同论证结果的结构性表征上，基于图尔敏论证模型的论证地图是一种有效方法(Noroozi et al., 2012)，能够直观清晰地将学生的论证结果外显为清晰的论证结构，包括主张(claim)、资料(data)、理由(warrant)、支持(backing)、限定(qualifier)、反驳(reservation)。目前，研究者已经开发了多种工具来支持学生绘制论证地图，如 iArgue 工具、AppleTree 工具、在线协作白板等(Zheng et al.,2023)。在论证对话上，观点生成与观点批判是论证对话的核心。观点生成是指考虑不同利益相关者的立场并与他们进行协商和对话，有可能生成更全面，多元化和包容性的观点，在决策中考虑到多种利益，并创建更具社会响应性的解决方案，从而更深入地理解有争议的话题和知识的构建 (Erduran & Jiménez-Aleixandre, 2008) 。生成多样化的观点是丰富论证结构，开展论证对话的基础。观点批判是指运用批判性问题进行论证，例如检查对方证据的可靠性，论点的完整性等，从而帮助学生判断其他论证主体的论点的可接受性，促进不同论点的整合(Nussbaum & Edwards , 2011)。

相应的，现有研究从结构性与社会性两个维度来分析与评价协同论证。在协同论证的结构性方面，常采用图尔敏模型来分析学生论证成果（如论证性文章、论证地图）中所体现的论证要素，但图尔敏模型难以分析学生论证性对话中所体现的结构(Nam & Chen, 2017) 。在协同论证的社会性方面，分析论证对话中的交互模式是必要的。论证性交互包括陈述自己的论点、指出他人论点的不足、支持他人的论点、修改自己的论点等等(Chin & Osborne, 2010)。其中，认知网络分析已被用于展示协作过程中话语属性的相互关联关系。但大部分研究运用 ENA 考察协同论证对话中的信息分享、提问、调节、或笼统的论证技巧(Ouyang et al., 2023; Zheng et al., 2023) ，少有研究综合分析协同论证话语中的结构与社会性要素之间的共现关系。

2.2. 融合生成式 AI 的人机协同论证

聊天机器人是人机协同论证领域的典型应用，它能够针对某一论证话题，使用自然语言与学生即时交互，不仅能降低学生与人类同伴论证时的焦虑，还能提升学生的论证能力与学业表现(Wu & Yu, 2023) 。目前聊天机器人的设计与实现基于两种模型，分别是检索模型与生成模型。基于检索模型的聊天机器人是有一个与特定论证话题相关的语料库，当学生选定论证主题与立场之后，聊天机器人会依据相似度算法从语料库中选择与学生立场相同或相反的论点来与学生互动 (Guo et al.,2023)。但是，基于检索模型的聊天机器人依赖于语料库中的预设数据，当学生的提问与互动超出语料库的范围时，聊天机器人的响应将受到限制。而基于生成模型的聊天机器人支持论证任何主题，并且能够根据人机对话的情境及时生成恰当的响应。但响应的质量取决于聊天机器人所训练的数据库、算法以及学生的提问。

ChatGPT 是一种基于大语言模型的聊天机器人，它使用了生成式的预训练模型 GPT。所使用的训练数据集规模涵盖几十亿的互联网文本数据，所使用的深度学习算法使其能够以更接近人类思考的方式、模拟多种人类情绪和语气，根据语境提供恰当的回答（Hirosawa et al.,2023）。研究发现，ChatGPT 能够通过提供多视角观点来支持学生的协同论证。Su 等人发现，学生在论证性文章写作时，ChatGPT 能够通过与学生的对话，提供多种视角与观点，从而激发学生的想法(Su 等, 2023)。然而，批判性地分析与判断 ChatGPT 的回复，实现人机论点的整合是必要的。但目前少有研究关注如何帮助学生批判性地整合 ChatGPT 的观点。

此外，ChatGPT 的观点生成质量取决于有效的提示（prompt）。提示是用户与 ChatGPT 交互时的输入过程，以便引出最准确、相关的响应。从 ChatGPT 的原理来看，ChatGPT 的回答质量不仅取决于底层算法和训练数据，还取决于用户输入提示的有效性。因此，提示工程受到广泛关注。从人机对话的本质来看，提示是人机交互的中介。在人机对话中，ChatGPT 鼓励学生通过提问问题和思考答案的方式与之进行互动。学生依据自己的思考提出问题，ChatGPT 根据学生的表述输出个性化的回答，通过多轮次的追问与回应，人机不断澄清、整合对话题的理解。这种对话式学习与传统的搜索式学习相比，对学生的提问能力提出了更高的要求。如何提示 ChatGPT 生成高质量的回答变得尤为重要(Liu 等, 2023)。因此，需要对提示进行精心的设计。目前，教育界的提示工程研究主要探索了提示语设计的原则。例如，Lo 提出了适用于教育情境中的 CLEAR 提示设计框架，包括五个核心原则:简洁性、逻辑一致性、明确性、适应性和反思性。应用该框架设计提示能够让 ChatGPT 输出更符合学生需求的回应(Lo, 2023)。然而现有教育界的提示工程研究更关注更广泛的教育情境下提示本身的语言与结构设计，未结合具体的人机协同论证情境考虑如何设计提示促进观点生成与观点批判，也未在教学实践中检验提示的有效性。

3. 融合生成式 AI 的人机协同论证活动设计

本研究基于 linn 提出的协作知识整合教学模型(Linn et al.,2006) 与 Halaweh 提出的 ChatGPT 教育应用策略(Halaweh, 2023)，构建了融合生成式 AI 的人机协同论证活动（如图 1 所示）。在人机协同论证活动开始之前，教师首先在课上明确说明鼓励学生使用 ChatGPT 进行论证，强调使用 ChatGPT 的目的是启发学生对论证话题进行多角度思考，而非直接提供答案，直接复制 ChatGPT 生成的文本将被视作抄袭。告知完毕后，教师讲授学科中的关键问题及其争议，并发布论证话题。课后，小组成员运用在线异步沟通工具企业微信小组通过文字交流进行协同论证。这种交流模式为所有学习者提供了足够的时间来构思论点和收集论据，避免即时交流中可能出现的错误和冲动回应，并且能够留下详细记录，便于后续的回顾与反思。在活动开始后的头脑风暴阶段，小组成员首先基于讨论主题独立思考，明确自己的立场和论点，并将其表达出来。在观点生成阶段，通过提示与 ChatGPT 进行对话，引入更多的思考角度。在观点批判阶段，成员通过提示对 ChatGPT 的观点进行评估、追问和解释，最终形成对论证主题的一致性理解并达成共识。人机协同论证活动结束后，小组成员使用在线协作白板工具，以论证地图的形式将协同论证活动的结果进行可视化展示。

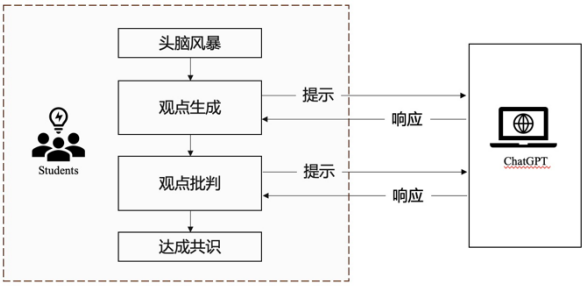


图 1 人机协同论证活动框架

观点生成阶段提示的设计依据来源于社会建构主义与教育界的提示工程研究。社会建构主义强调知识和社会现实是多元和相对的，与社会和文化环境紧密相关(Lock& Strong, 2010) 。不同角色和身份的个体和群体可能受到其文化、社会背景、利益的影响，具有不同的认知风格、经验和专业知识，对某一领域问题的有不同的理解。通过考虑不同利益相关者的立场，与其进行协商与对话，可以建立更全面、多元、包容性的论证，在决策中考虑多方利益，创造更符合社会需求的解决方案，从而深刻地理解争议性主题与建构知识(Erduran& Jiménez-Aleixandre,2008) 。教育领域的提示工程研究发现，当要求 ChatGPT 讨论一个话题时，指定特定的角色或视角可能会有所帮助(Atlas, 2023). ChatGPT 可以模拟不同的对话角色，生成的输出将与该特定角色对提出的问题的典型反应保持一致(Hwang & Chen, 2023)。因此，本研究设计的观点生成提示将引导 ChatGPT 模拟与论证话题相关的利益相关者，为学习者提供与不同角色与身份的人的交流体验，从而刺激学生者从多个视角考虑论点。针对论证话题“你认为 AI 绘画是不是艺术，为什么？”本研究所设计的观点生成提示如表 1 所示。

表 1 观点生成阶段的提示设计

	观点生成提示
1	假如你是一名原创画家，请用专业口吻向我解释 AI 绘画是艺术（划线处可以替换你的观点，如： AI 绘画不是艺术）
2	假如你是一名研发 AI 绘画技术公司的员工，请用专业口吻向我解释……
3	假如你是一名绘画收藏家，请用专业口吻向我解释……
4	假如你是一名消费者，请用专业口吻向我解释……
5	假如你是一名意识形态与伦理研究者，请用专业口吻向我解释……
6	假如你是一名学习绘画的学生，请用专业口吻向我解释……
7	假如你是一名美术老师，请用专业口吻向我解释……
8	假如你是一名（其他你能想到的想要 AI 扮演的角色），请用专业口吻向我解释……

观点批判阶段提示的设计依据为批判性问题模型(CQMAA) 。批判性问题模型发展于沃尔顿的对话理论，将其概括为支持论证的九个关键问题。学生运用这些关键问题思考其他论证主体的论点，可以判断其论证强度与说服力，从而帮助学生考虑是否接受他人立场并加强不同论点的整合。依据研究情境，本研究设计的观点批判提示如表 2 所示。

表 2 观点批判阶段的提示设计

	观点批判提示
1	我认为 AI 绘画是艺术（划线处可以替换你的观点）有哪些证据可以支持这一观点？
2	你提供的证据是否有参考来源？
3	你提供的证据是否有遗漏？
4	如果我不同意你的观点，你可以进一步解释你提供的证据吗？
5	请用具体例子来解释你的证据如何支持你的观点。
6	从以上证据到论点的推论过程是否正确？
7	我认为 AI 绘画是艺术（划线处可以替换你的观点），请反驳我的观点。
8	我认为 AI 绘画是艺术（划线处可以替换你的观点），请解释这一观点成立的限制条件是什么。

4. 融合生成式 AI 的人机协同论证活动效果检验

本研究运用准实验研究方法检验融合生成式 AI 的人机协同论证活动的效果。其中，对照组学生运用 ChatGPT 进行自由的协同论证，实验组学生在本研究提出的人机协同论证活动框架的指导下进行论证。实验结束后，比较两组论证对话在结构性与社会性维度上的差异。

4.1. 研究对象

研究对象是来自中国江苏某大学的教育学相关专业一年级研究生，共计 61 名。学生平均年龄为 23 岁，专业背景涉及多个学科，包括小学教育、心理健康教育、现代教育技术、语文教学、英语教学、音乐教学以及教育技术学。33 名学生参加了信息化教育理论与实践课程，28 名学生参加了教育技术学基本理论课程。两门课程均在 2023 年秋季学期开课，课时相同，由同一名教师授课。两门课程的教学内容有一部分是相同的，即要求学生通过协同论证感知与反思信息化教学实践中的争议性问题所有学生自愿组成 4-5 人的小组。其中，信息化教育理论与实践课程的 33 名学生被分配于对照组，共 7 个小组。教育技术学基本理论课程的 28 名学生被分配于实验组，共 6 个小组。

4.2. 实验流程

实验时间持续 4 周，每 2 周开展一轮完整的论证活动，共两轮。在活动开始之前，教师培训了如何使用企业微信、ChatGPT3.5 与在线协作白板工具进行协同论证。在第一轮活动中，教师在授课之后布置了论证话题，即：“你认为 AI 绘画是不是艺术？为什么？”。课后，对照组的小组进行自由论证，实验组的小组依据本研究的人机协同论证活动框架进行论证。第二轮活动中，教师在授课之后布置的论证话题为：“虚拟场景与现实场景哪一种对于学习更有价值？为什么？”，课后实验组与对照组的论证活动与第一轮活动相同。

4.3. 研究工具

本研究依据论证话语编码框架对学生的论证对话进行编码。该框架包含结构性与社会性两个维度，改编自(Darmawansah et al., 2022; Su et al., 2021; Chen et al., 2016) 。结构性维度包含主张（C）、证据（E）、推理（R）、反驳（CA）四个要素，每个要素下有三个由低到高的层级。如在主张要素中，低层级编码为 CC，即复述原材料或其他论证主体的主张；中层级编码为 CE，即简单提出自己的主张；高层级编码为 CD，即深入分析理由去支持自己的主张或在新证据的基础上对该主张进行改编。社会性维度包含 11 个要素，分别是提出自己的论点（P）、解释拓展自己的论点（E）、修改自己的观点（R）、询问他人的观点（A）、支持他人的观点（SP）、归纳总结他人的观点（SG）、指出他人论点的不足（C）、检查他人论据的有效性与可信度（CK）、说服他人接受自己的论点（D）、被他人激发出新的论点（SA）、将不同的论点整合为新的论点（I）。

4.4. 数据收集与分析

研究收集了每个小组企业微信的对话记录，共 1211 条。观察数据发现，每一条对话框中的文字内容都很丰富，最长的一条对话框包含 7 条句子。这可能是因为相比于面对面口头论证，文字通讯工具为所有学习者提供了充足的时间来构思论点、收集论据，一条对话框体现了学生完整的思考与表达，其中涉及多种论证要素，适合分析其共现关系。因此，分析者将一条对话框视作一个编码单元。依据 ENA 的编码表模板，按照编码框架标注其中出现的论证要素。由两名了解协同论证的研究者进行独立编码，当出现不确定的地方协商解决。编码结束后，首先统计对照组与实验组论证对话中出现的结构性与社会性要素的频率并比较其差异，其次，运用 ENA 分析两组结构性要素与社会性要素的认知网络及其差异。

4.5. 检验结果

4.5.1. 结构性维度

图 2 呈现了对照组与实验组论证对话中所体现的结构性论证要素频次的差异。结果显示，实验组的中高层次的证据频率高于对照组，中高层次的反驳频率高于对照组。

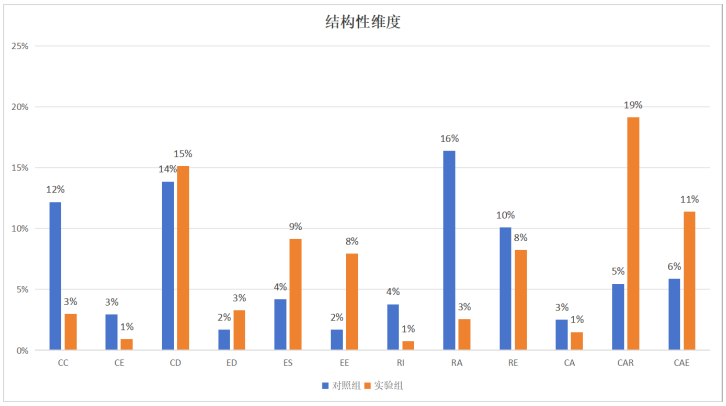


图 2 对照组与实验组的结构性论证要素频率

两组的认知网络分析进一步说明（如图 3），实验组的高层次主张、中层次证据、高层次推理、中高层次反驳的共现关系较强（CD-ES-RE-CAR-CAE），对照组中高层次主张、中高层次推理的共现关系较强（CC-CD-RA-RE）对两组认知网络的质心进行非参数检验后发现，两组在 X 轴上存在显著差异（p=0.00）。

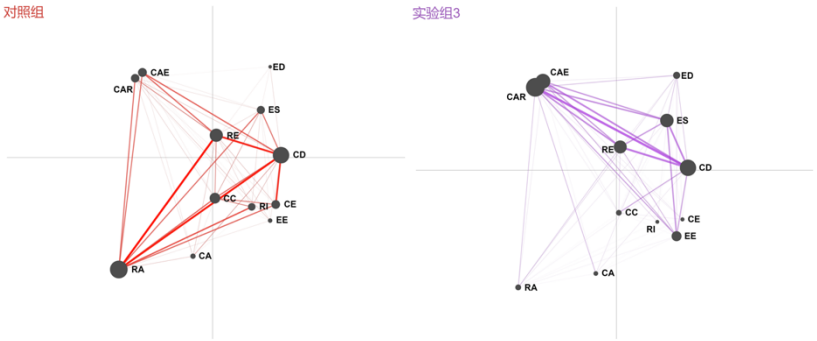


图 3 对照组与实验组的结构性论证要素认知网络

这说明，人机协同论证活动使得学生能够提出较高质量的主张、证据、推理与反驳，论证对话中体现的论证结构相较于对照组更为完善。在提出自己的主张时，从可靠来源中选择相关证据并对其进行细致描述，详细论证证据与主张之间的联系，有力地从证据推论到结论。并且能够以充足的证据反驳他人的复杂主张。

4.5.1. 社会性维度

对照组与实验组论证对话中所体现的社会性论证要素频次如图 4 所示。结果显示，对照组的提出论点（P）、解释拓展论点（E）、支持他人观点（SP）高于实验组；实验组的指出他人论点的不足（C）、检查他人论据的有效性与可信度（CK）、说服他人接受自己的论点（D）、修改自己的观点（R）、被他人激发出新的论点（SA）、将不同的论点整合为新的论点（I）高于对照组。

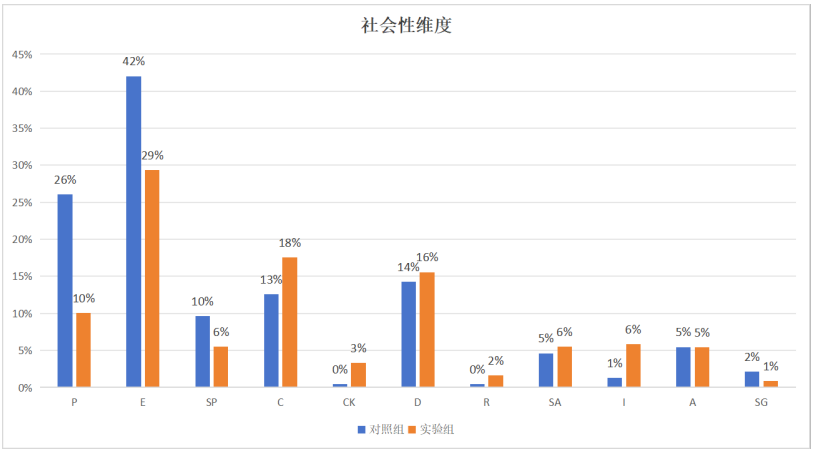


图 4 对照组与实验组的社会性论证要素频率

两组的认知网络分析进一步说明（如图 5），对照组的提出论点-解释论点（P-E）共现强度较高，而实验组在提出论点-解释论点-指出他人论点的不足-说服他人接受自己的论点（P-E-C-D）；提出论点-解释论点-检查证据有效性-支持(P-E-CK-SP)的共现强度较高，且两组在 X 轴上存在显著差异（p=0.05）。

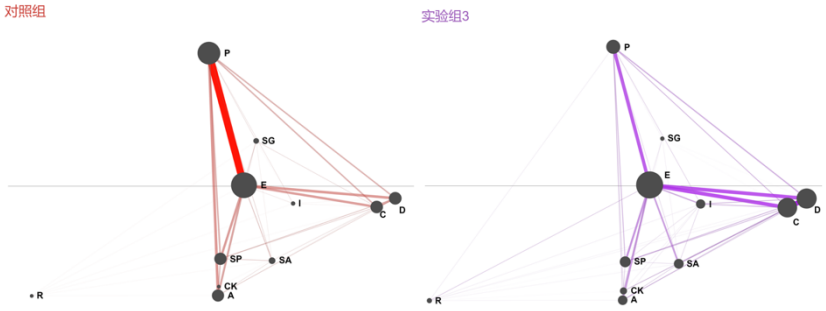


图 5 对照组与实验组的社会性论证要素认知网络

这说明，在融合生成式 AI 的人机协同论证活动中，学生除了论点论据的方式阐述自己的观点外，也会积极地反驳他人观点，来加强自身论证的合理性。此外，在论证过程中也会更加注重检查证据的有效性与可信度，来审查彼此提供的证据的真实性。

5. 结语

本研究聚焦协同论证这一重要的教育场景，所设计的人机协同论证活动能够提升结构性论证要素的质量，完善论证结构，促进论证对话的社会性交互。尽管得出了一些积极的发现，但本研究仍有一些值得注意的局限性。本研究设计的提示只引发了一问一答式的单向互动，学生更期待与生成式 AI 进行双向的、多轮次的沟通，实现真正意义上的人机协同论证，未来研究可以继续关注引发双向互动的提示设计，让提示真正成为人机交互的中介。此外，本研究的实验周期有限，因此建议未来的研究在更长的时间内验证融合生成式 AI 的协同论证活动的机制与作用。

参考文献

Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI.* Retrieved from https://digitalcommons.uri.edu/cba_facpubs/548.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 1-18.

Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of Research in Science Teaching, 47*(7), 883-908.

Cress, U., & Kimmerle, J. (2023). Co-constructing knowledge with generative AI tools: Reflections from a CSCL perspective. *International Journal of Computer-Supported Collaborative Learning*, 1-8.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in science education. Perspectives from classroom-Based Research.* Dordre-cht: Springer.

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education*, 104862.

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing, 54,* 100666.

Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology, 15*(2), 1-11.

Hirosawa, T., Harada, Y., Yokose, M., Sakamoto, T., Kawamura, R., & Shimizu, T. (2023). Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: A pilot study. *International journal of environmental research and public health, 20*(4), 3378.

Linn, M. C., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science, 313*(5790), 1049-1050.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys, 55*(9), 1–35.

Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship, 49*(4), 102720.

Nam, Y., & Chen, Y. C. (2017). Promoting argumentative practice in socio-scientific issues through a science inquiry activity. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(7), 3431–3461.

Noroozi, O., Weinberger, A., Biemans, H. J., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review, 7*(2), 79-106.

Nussbaum, E. M., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences, 20*(3), 443–488.

Ouyang, F., Tang, Z., Cheng, M., & Chen, Z. (2023). Using an integrated discourse analysis approach to analyze a group's collaborative argumentation. *Thinking Skills and Creativity, 47*, 101227.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, 100752.

Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, 1–24.

Zheng, X.-L., Huang, J., Xia, X.-H., Hwang, G.-J., Tu, Y.-F., Huang, Y.-P., & Wang, F. (2023). Effects of online whiteboard-based collaborative argumentation scaffolds on group-level cognitive regulations, written argument skills and regulation patterns. *Computers & Education, 207*, 104920.

教师与同伴组合式评价活动对评价质量与学习成绩的影响研究

A Study of the Effects of Teacher-Peer Portfolio Assessment Activities on Feedback Quality and Learning Achievement

张艺铧[1]，陆雨菲[2]，杨现民[3]，李新[4*]
[1234]江苏师范大学智慧教育学院
*lix2023@jsnu.edu.cn

【摘要】 同伴互评作为一种重要的评价方式，存在着评价质量不高、优等生受益较少等问题。因此，将同伴互评与教师评价相结合，形成一种新的评价方式，有利于优化同伴互评的效果，使所有学生从中受益。然而，这种方式对学生的评价质量及学习成绩的影响尚不清晰。基于此，本研究面向大学生开展了准实验研究，采用定量和定性相结合的方式对互评内容和学习成绩进行分析。研究发现，教师与同伴的组合式评价显著提升了反馈的质量，并有效促进了学生学习成绩的提高。
【关键词】 同伴互评；教师评价；反馈质量；学习成绩；影响研究

Abstract: *Peer assessment, as an important form of assessment, suffers from problems such as low quality of assessment and fewer benefits for top students. Therefore, combining peer-to-peer assessment with teacher assessment to form a new assessment method is conducive to optimizing the effect of peer-to-peer assessment and benefiting all students. However, the impact of this approach on students' assessment quality and academic performance is unclear. Based on this, this study conducted a quasi-experimental study for college students, using a combination of quantitative and qualitative methods to analyze the content of mutual assessment and learning achievement. It was found that the combined teacher-peer assessment significantly improved the quality of feedback and effectively contributed to the improvement of student's academic performance.*
Keywords: peer assessment, teacher evaluation, quality of feedback, academic achievement, impact study

基金项目：2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）；国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”(项目编号：62077030)

1.引言

当前我国教育实践中，学生评价仍存在以甄别与选拔为根本目的、以标准化测验与测定分数为主要评价方法的单一、线性评价不利于科学、有效地评价学生发展。因此，寻求有效的评价模式成为了增强学习成效的关键，特别是那些能够结合多种评价元素的评价方法。教师与同伴的组合式评价结合了教师评价的专业性和同伴评价的互动性，旨在提供更全面、更精确的反馈信息。尽管如此，对于教师评价与同伴互评的组合式评价的研究仍显不足。因此，本研究旨在探究此空白领域，综合分析教师与同伴组合式评价活动对学生反馈质量和学业成绩的影响，以期为教育实践提供有益的参考和启示。

2.文献综述

2.1. 同伴互评的研究现状

作为一种可以有效提升学习者学业绩效的教学方式，同伴互评受到学者的广泛关注（Kaufman & Schunn,2011)。Topping（1998）指出同伴互评是学习者对具有相同学习背景的其他同伴的学习作品或成果进行评价的一种方式，依据同伴作品完成的数量、水平、价值、作用、质量等进行评价。舒存叶（2017）进一步拓展同伴互评的应用场景，认为同伴互评不仅是一种评价方式，更是一种学习方式。近年来学者对同伴互评的研究多集中于互评评语本身和互评效果（柏宏权和李婷，2020)。针对互评评语本身，研究发现评价者撰写的问题或建议的质量能够预测自身完成作品的质量（Lu & Law，2012)；针对同伴互评反馈评语的效果，主要以分析反馈评语的数量、类型及其与学习效果的关联为主（马志强等，2016)，如同伴反馈的数量对作品质量有促进作用，同伴反馈的数量在第一轮到第二轮中会影响学习绩效，在第二轮到第三轮中的影响则不明显（Chen & Tsai，2009)。但也有学者发现同伴的评价质量往往无法达到专业评价的标准（Hwang et al.，2020)、学生的同伴评估结果可能受到多种因素的影响(Van et al.，2010）、优秀和中上水平学生从同伴评估中获益较少的问题仍然存在(Cheong et al.，2023）等等。

2.2. 教师评价的研究现状

教师评价通常被认为是权威和可靠的（Hu & Ren，2012；Huisman et al.，2019），它能引导学生实现学习目标（Wiese & Nortvedt，2023）。然而，在大学课程中，教师评价是一项繁琐的工作，尤其是在面对来自不同班级的学生时，由于时间和精力的限制，教师往往只能采取简单的反馈和评分方式，因而削弱评价的重要作用（Cui& Gai，2022 ）。还有学者指出，教师评价存在评价标准单一、评价方式简单、评价结果应用不足等问题（张钧和邵琳，2017)。

2.3. 研究问题

本研究将同伴互评与教师评价方式相结合，深入探讨这一组合方式对学习者评语质量和学习成绩的影响，旨在解决以下两个问题：

1.教师与同伴组合式评价活动对学生同伴互评的评语质量有什么影响？

2.教师与同伴组合式评价活动对学生的学习成绩的提高有什么影响？

3.研究设计

3.1. 研究对象与情境

本研究以江苏某高校的 2021 级教育技术学专业必修课“交互媒体设计”的 18 名本科生为研究对象，其中，男生 7 人，女生 11 人，年龄均为 20-21 岁。选取该课程中的“教育交互媒体设计”为活动内容，并要求参与者撰写设计方案。参与者根据自身兴趣和项目主题倾向自由组合，每组由 3 人组成。在评价活动中使用智慧教学平台，该平台支持开展同伴评价以及教师评价等多种评价方式，如图 1 所示。研究对象均未经历同伴互评学习活动，且各小组在实验前均已完成了教育交互媒体设计的项目需求分析，并熟练使用智慧教学平台。

3.2. 实验流程

本研究以撰写教育交互媒体设计方案为主题，以“教师与同伴组合式评价-方案修改-再次评价”为研究思路，探究学习者评语质量和学习成绩的发展，如图 2 所示。学生首先学习交互媒体设计的相关知识，合作撰写设计方案。接着，需要在课堂上完成两轮互评活动，依据教师提供的评价标准对其他小组进行匿名评价，教师对作品进行打分和撰写评语。第一轮互评活动结束后，各小组有一周的时间根据同伴和教师评价的意见进行方案修改，并使用修订模式追踪修改过程，修改后的作品作为第二轮互评活动的内容。



图 1 智慧教学平台评价页面

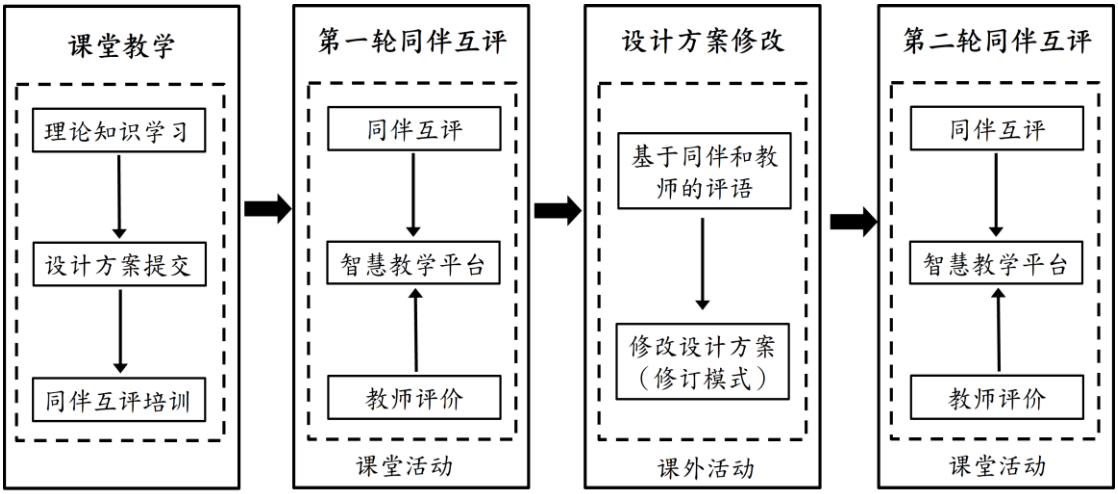


图 2 实验流程图

3.3. 研究工具

本研究旨在探讨教师与同伴组合式评价活动对学生学习者评语质量的影响，因此采用量规和评价标准作为研究工具。

反馈质量量规改编自 Hwang（Hwang & Chang，2021），包含四个维度：评语的清晰度、评语的准确性、评语的意见陈述和评语的可参考性，如表 1 所示。

表 1 反馈质量量规

维度	5	4	3	2	1
评语的清晰度	清晰解释所有主要观点	清晰解释大部分主要观点	清晰解释差不多一半的主要观点	清晰解释少数主要观点	没有清晰解释任何观点
评语的准确性	所有评语正确反映我的工作	大部分评语正确反映我的工作	差不多一半的评语正确的反映我的工作	只有少数评语正确反映我的工作	没有评语正确反映我的工作
评语的意	所有评语都提供了证据	大部分评语提供证据或参考	差不多一半的评语提供证据	只有少数评语提供了证据或	没有评语提供证据或参考材

见陈述	或参考材料	材料	或参考材料	参考材料	料
评语的可参考性	所有评语都富有启发性，可以为其他评估者提供很好的参考	大部分评语富有启发性，可以为其他评估者提供很好的参考	差不多一半的评语富有启发性，可以为其他评估者提供很好的参考	只有少数评语富有启发性，可以为其他评估者提供参考	没有评语富有启发性，可以为其他评估者提供参考

评价学生教育交互媒体设计作品的评价表由任课教师依据课程要求编制，从“教育价值”、“用户需求”、“界面设计”和“交互功能”四个维度展开。每个维度均采用 25 分制计分标准，总分为 100 分。

3.4. 数据收集与分析

本研究通过智慧教学平台收集了两轮同伴互评活动中的评语数据和设计方案。针对评语数据，由两名经过培训的编码员依据反馈质量量规对评语数据进行预编码，明确有异议的行为归属类别，形成一致的评判标准，分析单位为一条完整的同伴评语。为了确保编码规则的准确性，对两位编码员的编码情况进行 Cohen's Kappa 分析，以检查是否一致。若出现分歧（Cohen's Kappa 低于 0.7）的情况将再次进行讨论。达成一致意见后，由两位编码者进行独立编码。使用描述性统计分析评语在不同维度上的分布，采用配对样本 t 检验比较两轮同伴互评活动评语质量提升情况。针对设计方案，我们邀请了三名评分者依据评分标准对两轮设计方案进行评分，并进行 Kendall 分析，结果为 0.816 和 0.805，表明三位评分者的评分之间存在显著相关性。因此，以三名评分者的平均分作为学生两轮方案设计的成绩。采用描述性统计分析学习者两轮学习成绩的变化，为了进一步了解教师评语与同伴互评评语是如何提升学生学习成绩，对设计方案修订情况进行定性吸收分析。

4.研究结果

4.1. 教师与同伴组合式评价活动对评语质量的影响

根据反馈质量量规对同伴互评评语进行内容分析，以解决第一个研究问题。首先我们对编码情况进行了描述性统计分析，如图 3 所示。在评语清晰度维度上，第二轮互评评语在“清晰解释少数主要观点”、“清晰解释一半主要观点”的数量高于第一轮互评；在评语准确性维度上，第二轮互评评语在“所有评语正确反映我的工作”的数量高于第一轮；在评语的意见陈述维度上，第二轮互评评语在“一半的评语提供证据或参考材料”、“少数评语提供了证据或参考材料”以及“没有评语提供证据或参考材料”的数量高于第一轮。在评语的可参考性维度上，第二轮互评评语在“没有评语富有启发性，可以为其他评语者提高参考”的数量高于第一轮互评。

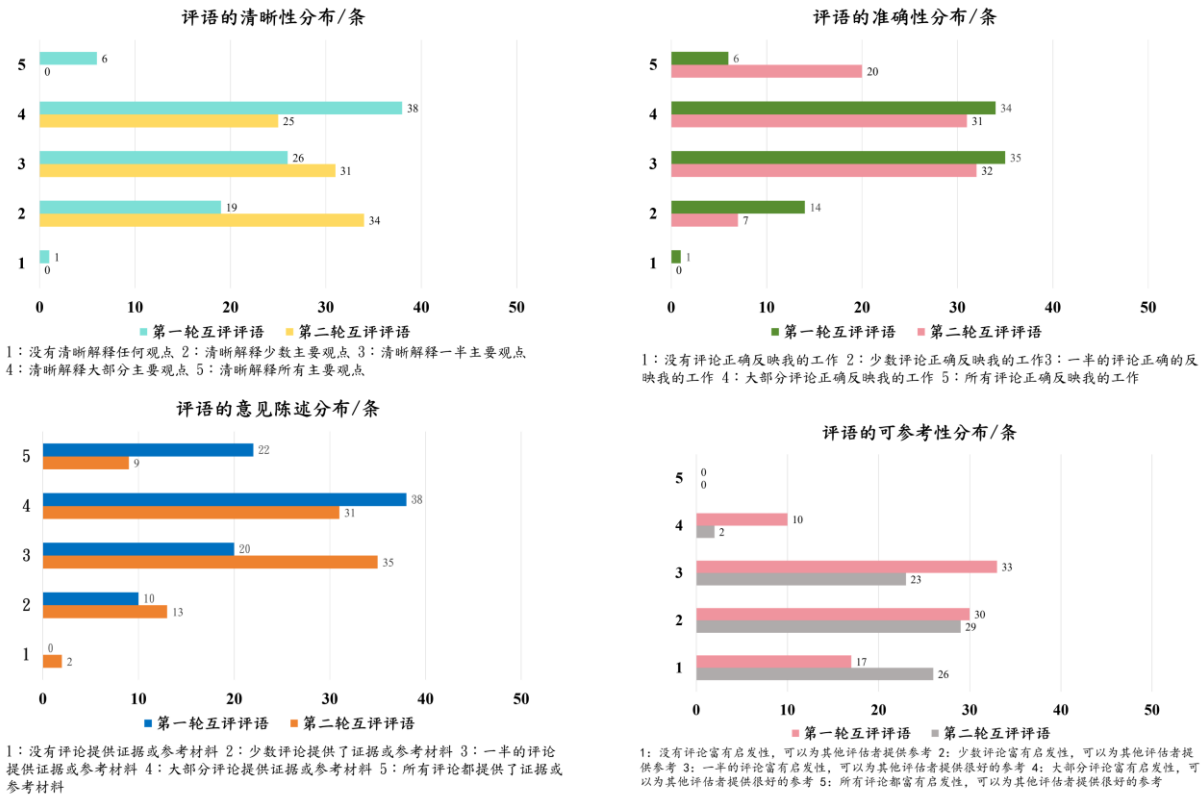


图 3 两轮互评评语分布图

为了分析两轮评语质量的差异，我们对数据进行了检验，发现数据满足配对样本 t 检验的假设。因此，可以对数据进行配对样本 t 检验，如表 2 所示。两轮同伴互评的评语质量存在显著差异，在第一轮同伴互评活动中，评语的清晰度、意见陈述以及可参考性显著高于第二轮，然而，第二轮评价活动中评语的准确性显著高于第一轮（t=-3.581，p<0.001）。

表 2 评语质量的配对样本 t 检验

第一轮-第二轮	M	SD	t	Sig
清晰度	.422	.114	3.571	<.001
准确性	-.378	-.041	-3.581	<.001
意见陈述	.444	.011	3.755	<.001
可参考性	.389	.121	3.084	.003

综上可知，教师与同伴组合式评价活动能够显著提高评语的准确性。这一结果可能得益于教师提专业性和针对性的反馈，凭借丰富的教育经验和深厚的专业知识，能够为学生提供精准而深入的点评，这种点评不仅有助于学生更全面的了解自己的作品优点和不足还能及激发他们深入反思和修改的积极性。通过同伴间的相互评价，学生也能从多角度审视自己的作品，进一步丰富评语的维度和深度，从而促使评语的质量提升。此外，我们还发现在清晰度、意见陈述以及可参考性等方面，第一轮的评语质量尤为突出。其有可能与作品本身的质量有着密切的关联，在高质量的作品中，由于作品本身已经充分展现了其卓越之处，评语无需过多展开详细评论、意见陈述及可参考性，仅需简明扼要地指出其优点，即可充分表达对作品的赞赏与肯定。例如，学生 A 在评价第四组时第一轮评语为“该学习平台针对教育技术本科生，对各个年级用户的需求进行分析，提供和制订个性化学习资源。核心功能齐全，界面设计美观简洁。首页课程类型繁杂，种类太多，根据类型可以分为不同板块。”第二轮评语则

为“该学习平台设计方案更加清晰，对目标用户特征分别进行分析，根据不同用户需求进行个性化学习资源推荐；核心功能设计合理，功能齐全；用户界面设计较为美观；市场前景广泛。”相反，当作品质量不尽如人意时，评语中则需要更多的详细评论、意见陈述及可参考性，具体指出作品中的不足之处，并为学生提供改进的方向与建议。例如，学生 B 在评价第五组时第一轮评语为“页面内容丰富，功能很多很完善，充分考虑了使用者的想法需求，但是比起一个教育软件更像看小说的软件。”第二轮评语为“笔记功能很完善，但是码字大厅有些不必要，或者说可以做的更完善一些，增加比如“草稿记录”功能。”

4.2. 教师与同伴组合式评价活动对学生学习成绩的影响

通过比较分析两轮教育交互媒体设计的作品成绩，探究教师与同伴组合式评价活动对学生学习成绩的影响。首先我们对两轮作品成绩进行了描述性统计分析，如图 4 所示。第二轮学生的教育交互媒体设计方案成绩相较于第一轮的成绩有显著提升。提升的程度通过两轮教师与同伴组合式评价活动学习成绩之间平均差(MD = 3.88)来说明。

为了进一步研究学生学习成绩是如何提升的，即在教师与同伴组合式评价活动中，教师评语和同伴评语如何有助于学习成绩的提升，因此我们进行了定性吸收分析，根据图 5，选取前后差异最大组进行定性吸收分析，即第四组（MD=6.7）。

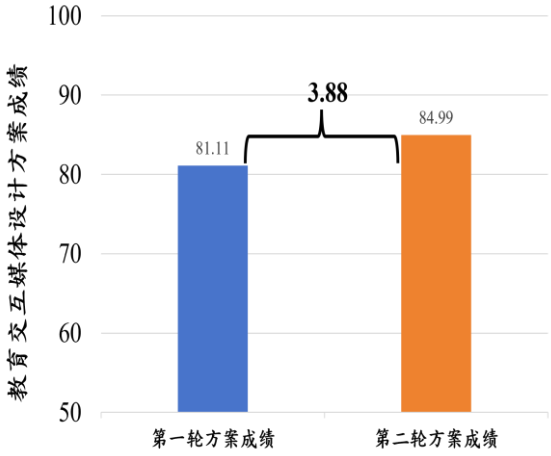


图 4 学生学习成绩

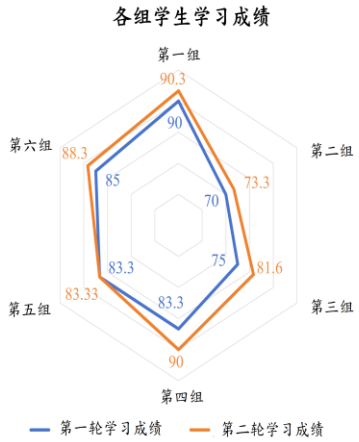


图 5 各小组学习成绩

图 6 显示了第四组在经历教师与同伴组合式评价活动后对评语的吸收图。顶端部分显示了在教师与同伴互评组合式评价前教育交互媒体设计的方案。底部部分展示了为回应同伴和教师的评语而改进的设计方案。中间部分展示了第四组所接受教师评语和其他小组的同伴互评评语，红色的文本表明该小组接受的修改建议。连接的箭头表明第四组是如何接受评语，并将其展现在设计方案中。分析可知教师所提供的评语是方向性的，例如“界面设计部分，可以进一步美化，色彩搭配”。而同伴互评更具有操作性，例如“课程类型繁杂”、“图片处于加载的状态”、“学术论文页面一眼看上去较乱”等。同时教师评语能提出在同伴互评过程中没有涉及的部分，例如“建议不要加入主观内容”、“市场前景部分建议从教育技术学学生培养这个角度进行”等，这些部分都是同伴互评过程中没有提及到的。

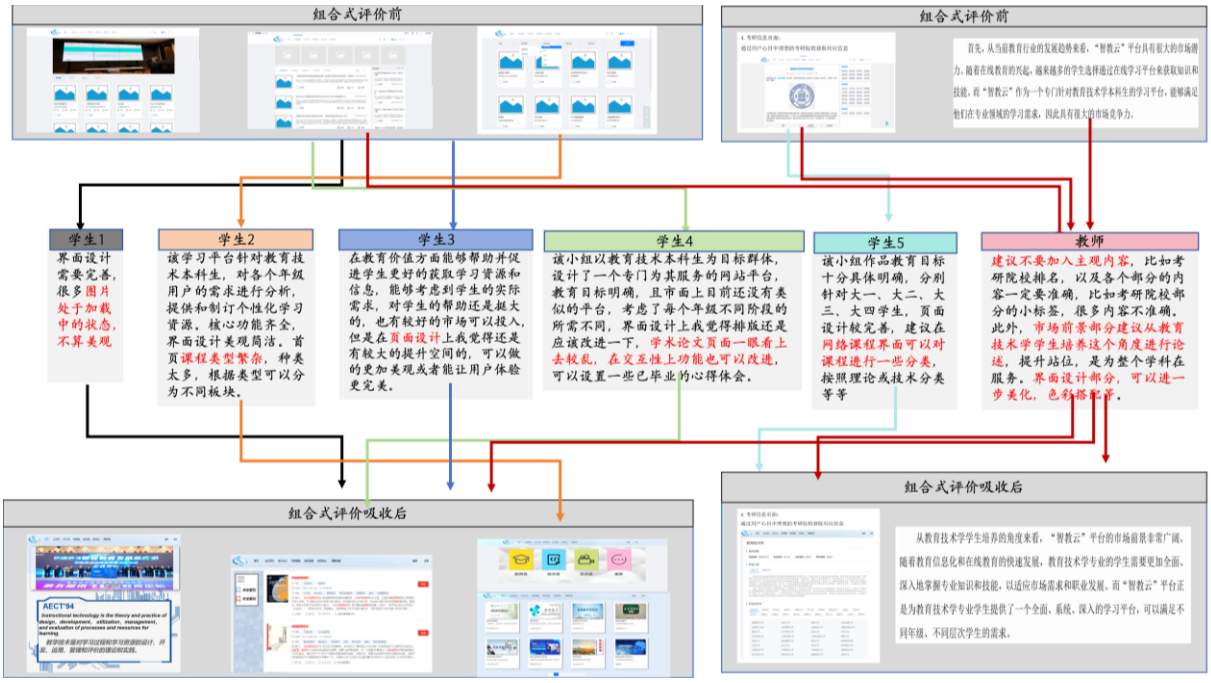


图 6 第四组评语吸收分析

可见，经过两轮教师与同伴组合式评价活动后，学生的学习成绩呈现出明显提升的趋势。教师与同伴组合式评价活动有助于学生掌握与教育交互媒体设计相关的知识，提高了教育交互媒体设计的质量，进而提升学生的学习成绩。同时，教师评语和同伴互评评语在学生学习成绩提升中都扮演了关键角色，教师评语提供了方向性的指导，为学生在学习过程中指明了前进的方向；而同伴互评评语则提供了具体的操作建议，帮助学生更好地理解和应用所学知识。

5. 结语

教师与同伴组合式评价能实现多方主体评价结果相融合，提升评价的准确性和全面性。然而，当前研究关于学生学习数据的评价方式简单、评价主体单一，弱化了评价在促进学生学习效果提升方面的作用。为此，本研究创新地采用教师与同伴组合式评价对学生的设计方案开展评价活动，接着对互评内容、学习成绩进行描述性统计分析、内容分析和配对样本 t 检验。研究发现，教师与同伴的组合式评价显著提升了反馈的质量，并有效促进了学生学习成绩的提高。但是，由于研究样本的规模相对较小，导致研究结果可能无法完全代表所有学生的情况。后续研究将进一步扩大研究样本，以获得更具有代表性的结果。未来研究可进一步挖掘教师与同伴评价对学生学习行为的影响机理，了解该评价方式如何影响学生的学习行为，如学习态度、学习策略、学习时间管理等，以更全面地了解评价活动对学生学习的综合影响。

参考文献

马志强, 王靖, 许晓群 龙琴琴. (2016). 网络同伴互评中反馈评语的类型与效果分析. 电化教育研究(01),66-71.

张钧 邵琳. (2017). 基于我国教师评价制度演进的思考. 东北师大学报(哲学社会科学版)(05),191-196.

柏宏权 李婷. (2020). 在线同伴互评的评语采纳倾向——基于情绪体验的视角. 开放教育研究(04),95-101.

舒存叶. (2017). 网络同伴互评系统的设计和优化. 电化教育研究(01),80-85.

Chen, Y. C., & Tsai, C. C. (2009). An educational research course facilitated by online peer assessment. Innovations in Education and Teaching International, 46(1), 105-117.

Cheong, C. M., Luo, N., Zhu, X., Lu, Q., & Wei, W. (2023). Self-assessment complements peer assessment for undergraduate students in an academic writing task. Assessment & Evaluation in Higher Education, 48(1), 135-148.

Cui, Y., C. D. Schunn, and X. Gai. 2022. "Peer Feedback and Teacher Feedback: A Comparative Study of Revision Effectiveness in Writing Instruction for EFL Learners." Higher Education Research & Development 41 (6): 1838–1854.

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. Computers & Education, 71, 133-152.

Hwang, G. J., & Chang, S. C. (2021). Facilitating knowledge construction in mobile learning contexts: A bi-directional peer-assessment approach. British Journal of Educational Technology, 52(1), 337-357.

Hu, G., and H. Ren. 2012. "The Impact of Experience and Beliefs on Chinese EFL Student Writers' Feedback Preferences." In Academic Writing in a Second or Foreign Language, edited by R. Tang, 67–87. London: Continuum.

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. Instructional Science, 39, 387-406.

Lu, J., & Law, N. (2012). Online peer assessment: Effects of cognitive and affective feedback. Instructional Science, 40, 257-275.

Topping, K. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249-276.

Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. Learning and instruction, 20(4), 270-279.

Wiese, E., and G. A. Nortvedt. 2023. "Teacher Assessment Literacy in Culturally and Linguistically Diverse Classrooms: A Norwegian Case Study." Teaching and Teacher Education 135: 104357.

共享调节视角下计算机支持协作学习中的教师职能：基于系统性文献综述

A Systematic Literature Review of Socially Shared Regulation Perspectives on Teacher Roles in Computer-supported Collaborative Learning

顾佳博，孟柯颖，骆开燕，蓝敏*
浙江师范大学教育学院教育技术学系
* lanmin@zjnu.edu.cn

【摘要】 信息技术时代下，作为一种重要的学习方式，计算机支持的协作学习已得到广泛的普及应用。共享调节作为一种新的协作学习研究与实践框架，于个人和集体在不同阶段以及不同维度的共同作用与调节发挥着重要的作用。但目前的研究多侧重调节机制理论的探究，忽视了教师角色及其在学生共享调节过程中所起到的职能效用。为此，本研究在 PRISMA 的数据筛选框架下，研究运用 Web of Science 等 6 个数据库实施数据检索与筛选，选取其中 22 篇作为研究文献来源，对其进行编码、统计和系统分析，阐述共享调节中教师干预内容、序列特征、支持工具，以期扩展共享调节理论。

【关键词】 计算机支持的协作学习；共享调节；教师职能；系统性文献综述

Abstract: *Computer-supported collaborative learning has become a widely used method of learning in the era of information technology. Shared regulation, a new framework for collaborative learning, plays an important role in regulating the interaction of individuals and groups at different stages and dimensions. However, current research focuses on exploring the theory of adjustment mechanism, while ignoring the role of teachers and their functional utility in the process of students' adjustment. Therefore, following the data screening framework of PRISMA, this study retrieved and screened data from six databases such as Web of Science. The study selected 22 research literature sources, encoded, counted, and systematically analyzes them. The study then explained the content, sequence characteristics and support tools of teacher intervention in shared regulation, in order to expand the theory of shared regulation.*

Keywords: computer-supported collaborative learning, shared regulation, teachers' function, systematic review

1.引言

学习者需要共同监控协作学习中的各种障碍与挑战，协商选择解决策略，调整认知、行为、情感动机等状态，以不断推进协作进程（Hadwin et al.,2018）。这一过程，被称为“共享调节学习”。但在此过程中，高质量的“协作”仅仅借助计算机是难以实现的，尤其是在小组学习中，若缺乏良好的教学设计和组织管理，计算机支持的优势将无从体现，甚至可能有消极的影响。团队的共享调节也与教师的引导与支持密切相关。在现有的研究中，教师在共享调节中充当的角色、教师干预序列的特征、共享调节中教师干预支持的工具等信息比较模糊，给教师实践带来的启示有限。因此，基于以上问题，本研究采用系统性文献综述法，对国内外共享调节的相关实证研究文献进行整理分析，综述共享调节中教师职能作用，以期为共享调节后续干预优化提供借鉴依据与创新基础。

2.文献综述

在协作学习过程中，学习者通过自我调节、合作调节和共享调节三个不同层次的方法调节学习（陈向东,罗淳 & 张江翔,2019）。自我调节是指学习者主动积极地参与学习过程，为实现学习目标而循环进行内部的元认知活动，涉及计划、执行与自我反思等周期性阶段与过程（Zimmerman,1989）。基于社会认知理论、社会文化理论、情境理论等的进一步发展以及协作学习的普及，Marja VAURAS 等人（2003）发现，学习者在真实协作学习情境中表现出的调节行为不仅限于自我调节的范畴，在部分情况下，学习者以平等的身份共同对任务进行调节，这种调节具有共享性。共享调节聚焦于集体层面的活动，强调群体成员通过有效的协调和沟通完成共同目标（Järvelä et al.,2013）。 Järvelä 等（2016）对计算机支持的协作学习（CSCL）课程中的协作小组在线聊天讨论和日志文件进行分析，发现协作计划的监管活动在实践中变得共享。使用多种监管流程以支持学习，并达成共同监管的协作小组取得良好学习成果。研究还表明个人社会共享调节对成功的协作学习起关键作用。

共享调节干预从群体感知、协作脚本和评价反思三个方面为协作学习提供支持（陈凯亮 et al.,2020）。根据其干预特征，共享调节支架可划分为静态支架和动态支架两种类型（Inge Molenaar et al.,2012）。根据其功能特征，共享调节支架可分为认知支架和社会支架（Miao et al.,2007）。Weinberger 等人（2005）还从学习目标、活动类型、序列特征、角色分配和团队状态五个方面，对共享调节的干预进行了归纳。尽管工具可以为学生共享调节提供过程支持，但教师的作用仍十分重要且不可代替（陈向东&刘冠群,2020）。CSCL 中教师的角色由“教授知识”转变为“引导学生”，由传统教学中以教师为中心的主体者转变为以学生为中心的主导者、启发者、帮助者和促进者（李艳燕 et al.,2008）。陈向东等人（2020）表明共享调节学习中的教师干预能有效促进集体认知与学生情感调节，推动社会参与。共享调节可以用于指导协作学习，在干预时需要支持学生协商任务共识，制定学习目标和计划，建立共享学习策略，感知集体进展和适应性调整等活动。然而，学生的共享调节不易自发产生，需要通过干预与引导，促使其发生（Panadero et al.,2015）。然而，对于共享调节学习中教师干预的作用的相关研究仍为少数，并且缺少 CSCL 下共享调节学习中教师干预的内容，特点，作用机制等的系统性描述。

本研究在对共享调节的概念、干预策略、CSCL 中的教师干预现状等研究整理的基础上，归纳出共享调节视角下计算机支持的协作学习中的教师干预内容、教师干预特征、教师干预的支持工具等，进一步为有效开展计算机支持的协作学习提供理论支持。

3.研究方法与过程

本研究采用系统性文献综述法，对国内外相关文献进行了深入检索与分析。我们确立了明确的文献筛选标准，并明确了文献来源和搜索策略，进一步进行了细致的文献筛选、数据整理与深入分析。为确保文献质量与研究目的的契合，我们设定了四个严格的筛选条件：首先，研究背景需限定在 CSCL 领域；其次，文献中必须包含对教师在 CSCL 共享调节中工作内容的详实研究与表述；第三，仅接受包含定量、定性或混合方法的实证研究或综述研究，排除纯理论性探讨；最后，文献来源限定为期刊文章，会议论文和报告等不在考虑之列。在文献来源方面，我们广泛涉猎了多个学术数据库，如中国知网、万方数据库、维普中文学术期刊、Web of Science、ScienceDirect 以及 Google Scholar 等，并辅以国内外知名教育学、心理学类学术期刊。搜索过程中，我们同时关注了中文和英文文献，搜索关键词涵盖了 CSCL、教师职能、共享调节等多个方面。在筛选过程中，我们遵循 PRISMA 流程图的四阶段方法，确保每一步都严谨有序。通过初步识别并去重后，我们根据设定的筛选标准对文章标题和摘要进

行了二次筛选，严格剔除了不符合要求的文献，如未涉及教师工作描述、非 CSCL 研究背景或缺乏实证研究支持的理论性探讨等。最终纳入研究的文献经过了详细的内容和数据提取，包括文献基本信息、研究结果与结论、相关数据和表格等。我们采用开放式编码、关联式编码、选择式编码的分析步骤，以期更全面地揭示教师在 CSCL 共享调节中的职能和作用。

4.结果与展望

本文对 2008 年-2023 年国内外有关共享调节中的教师职能研究成果，进行系统性的文献综述，分析了在现有共享调节实证研究中，教师角色类型、教师职能特性以及教师干预现状，主要得出以下结论与建议：

4.1. 教师干预的多维性与共享调节学习的深化理解

教师在共享调节学习中的多维角色及干预的四个核心维度：认知、元认知、情感动机和组织干预。这些维度构建了教师促进共享调节学习的综合框架，展现了教师干预的复杂性，并为全面理解 CSCL 中的教师职能提供了支撑。在认知干预方面，研究指出教师不仅传授知识，更引导协作学习中的认知过程，通过任务分析、深度理解引导和策略性思考机会提供促进学生认知发展和知识深层构建。教师需关注学生对知识的理解，并引导其在协作任务中进行有效分析和批判性思考。元认知干预强调教师引导学生监控、规划和调整自身学习过程。教师需具备引导学生自我反思的能力，并根据学生情况提供个性化指导。通过设定学习目标、制定学习计划和定期自我评估，强化学生元认知能力，培养其对学习过程的意识和调控能力。情感动机干预要求教师识别并应对学生在协作学习中的情感挑战和动机下降，采用多样化策略维持学生积极参与和学习热情。教师需创造积极学习氛围，及时鼓励，并深入参与学生情感状态和动机水平的调节。组织干预体现教师在学习活动设计、实施支持、团队管理和冲突调解等方面的关键作用。教师作为资源协调者和进程监督者，需有效管理学习资源、组织活动和协调团队合作，确保学习有序和目标实现。这些干预维度的多样性和复杂性要求教师能灵活适应学习情境和学生需求，切换不同干预策略，并持续评估和反思实践，确保有效性和适切性。

4.2. 教师干预过程的阶段特性与策略的适应性调整

教师在整个学习过程中扮演着举足轻重的角色，他们的干预不仅贯穿于任务的定义、目标的规划、计划的实施，还涉及到学习的监督与评价等各个阶段。在任务定义阶段，教师不仅需要清晰地阐述任务的要求，更重要的是要引导学习者对任务进行深入的理解，激发他们的学习兴趣和动力。进入目标规划阶段，教师需要与学习者共同设定具体可行的学习目标，并制定合理的学习计划，以确保学习活动的有序进行。而在计划的实施过程中，教师则需要密切关注学习者的学习进度和反馈，及时提供必要的支持和帮助，确保学习者能够顺利完成学习任务。最后，在监督评价阶段，教师不仅需要对学习者的学习成果进行客观的评价，更需要引导他们进行深入的反思和总结，以便更好地改进和提升学习效果。这一系列过程充分体现了教师干预策略的适应性和动态性，它们对于确保学习者的学习效果、促进其全面发展具有至关重要的意义。

4.3. 教师干预工具的集成应用与效果评估

本研究深入探讨了 KBS 学习分析工具、Standup、Visu 等教师干预工具在共享调节学习中的重要作用。这些工具不仅增强了教师对学习者行为模式、知识水平和协作交互情况的洞察能力，也提升了诊断和评估的精准性。然而，如何将这些工具有效融入日常教学，形成统一协调的系统，仍是当前面临的挑战。同时，本研究也强调了个性化和适应性在工具使用中的重要性，以及持续评估和反馈对于提升教师干预质量和效果的关键性。通过不断探索和创新，

教师可以更灵活地整合和应用这些工具，更有效地支持和促进学习者在共享调节学习环境中的全面发展和成长。

[基金项目]国家社会科学基金教育学青年课题《从调节理论视角优化教师在线教学设计的行动研究》[课题批准号：CCA220320]

参考文献

李艳燕,王晶,廖剑 & 黄荣怀.(2008).远程协作学习中教师角色研究.现代教育技术(06),53-56+12.

陈向东 & 刘冠群.(2020).共享调节学习中的教师干预：QCA 的视角.远程教育杂志(06),73-85.doi:10.15881/j.cnki.cn33-1304/g4.2020.06.008.

陈向东,罗淳 & 张江翔.(2019).共享调节:一种新的协作学习研究与实践框架.远程教育杂志(01),62-71.doi:10.15881/j.cnki.cn33-1304/g4.2019.01.006.

陈凯亮,包昊罡,李艳燕,彭禹 & 李新.(2020).协作情境下社会调节学习工具的设计与应用.现代教育技术(06),86-92.

Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), Handbook of self-regulation of learning and performance (2nd ed., pp. 83–106). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315697048-6

Inge Molenaar,Claudia Roda,Carla van Boxtel & Peter Sleegers.(2012).Dynamic scaffolding of socially regulated learning in a computer-based learning environment.Computers Education(2),515-523.

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. Educational Psychologist, 48(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Järvelä, S., Malmberg, J., & Koivuniemi, M. (2016). Recognizing socially shared regulation by using the temporal sequences of online chat and logs in CSCL. Learning and Instruction, 42, 1–11. https://doi.org/10.1016/j.learninstruc.2015.10.006

Miao, Y., Harrer, A., Hoeksema, K., & Hoppe, H. U.. (2007). Scripting computer-supported collaborative learning. Cognitive Computational & Educational Perspectives.

Panadero, E., Kirschner, P. A., Järvelä, S., Malmberg, J., & Järvenoja, H. (2015). How Individual Self-Regulation Affects Group Regulation and Performance: A Shared Regulation Intervention. Small Group Research, 46(4), 431-454. https://doi.org/10.1177/1046496415591219

Vauras, M., Iiskala, T., Kajamies, A., Kinnunen, R., & Lehtinen, E. (2003). Shared-regulation and motivation of collaborating peers: A case analysis. Psychologia: An International Journal of Psychology in the Orient, 46(1), 19–37. https://doi.org/10.2117/psysoc.2003.19

Weinberger, A., Ertl, B., Fischer, F. et al. Epistemic and social scripts in computer–supported collaborative learning. Instr Sci 33, 1–30 (2005). https://doi.org/10.1007/s11251-004-2322-4

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. Journal of Educational Psychology, 81(3), 329–339. https://doi.org/10.1037/0022-0663.81.3.329

在 K-12 教育中 ChatGPT 应用的系统综述

A Systematic Review of ChatGPT Applications in K-12 Education

叶碧滢[1]，汪溢慧[1*]，王埕埕[1]
[1]温州大学教师教育学院
*1460563383@qq.com

【摘要】随着 AI 技术在教育行业的日益普及，其对不同教育阶段的影响具有显著差异。尽管 ChatGPT 在高等教育中的应用已较为成熟和广泛，但在 K-12 教育领域，有关其应用的研究尚处于初级阶段。本文采用 PRISMA 框架进行了系统性回顾，分析了 19 篇关于 ChatGPT 在 K-12 教育中的应用研究，以深入探讨其在不同教育阶段和学科领域的实际应用情况，特别是在协作学习方面的作用。研究结果显示，ChatGPT 在 K-12 教育中扮演着多种角色，如思维工具、生成工具和学生学习的促进者等，但对于研究 ChatGPT 在不同学科的协作关系和成效仍较匮乏。教师和学生普遍认为 ChatGPT 对提升学习体验有积极影响，但他们也对其可能带来的负面效应表示关切。

【关键词】K-12 教育；ChatGPT；协作学习；应用

Abstract: *With the increasing popularity of AI technology in the education industry, its impact on different levels of education is significantly different. Although ChatGPT's application in higher education is mature and widespread, research on its application in K-12 education is still in its infancy. In this paper, PRISMA framework is used to conduct a systematic review of 19 studies on the application of ChatGPT in K-12 education, in order to explore its practical application in different education stages and subject areas, especially in collaborative learning. The results show that ChatGPT plays a variety of roles in K-12 education, such as thinking tools, generative tools and facilitators of student learning, but there is still a lack of research on the collaborative relationship and effectiveness of ChatGPT in different disciplines. Teachers and students generally agreed that ChatGPT had a positive impact on enhancing the learning experience, but they also expressed concerns about possible negative effects.*

Keywords: K-12 education，ChatGPT，Collaborative Learning，Application

1. 前言

近年来，OpenAI 的 GPT 系列成为人工智能研究的前沿，其中教育是一个重要应用场景(Hays, Jurkowski, & Sims, 2023)。教师借助 ChatGPT 可以轻松地创建定制课程、测验和其他学习材料，能显著提高教师教学的有效性和效率(Kasneci et al., 2023)。也有研究通过 ChatGPT 开发了一个评估任务来评估中学生的学习表现(Zhai, 2022)。这些研究大多关注教学方面的应用，对学生学习和协作类型的关注有所欠缺。本文旨在关注 ChatGPT 在不同学科领域、不同学段学生中的实施现状。Watrianthos et al.(2023)通过文献计量学分析，发现研究重点集中在高等教育应用上，指出了 ChatGPT 在 K-12 教育领域的研究存在缺口。因此，有必要对 K-12 教育中使用 ChatGPT 的研究现状进行全面审查。

ChatGPT 等大语言模型以人机协方式个性化辅导学(卢宇, 余京蕾, 陈鹏鹤, & 李沐云, 2023)。随着 ChatGPT 与师生进行协同，我们需要理解它在协作中扮演的具体角色以及协作学习的优化方向。这有助于教育者和研究者针对 ChatGPT 为进一步协作教师教学和学生学习提供更好的理论和实践指导。因此本研究提出以下研究问题：

- RQ1: ChatGPT 支持下的 K-12 教育有哪些协作类型？协作情况如何？
- RQ2: ChatGPT 应用在哪个教育阶段，是如何实施的？实施的结果如何？
- RQ3: 教师和学生对 ChatGPT 的应用态度如何？对未来发展有什么建议？

2. 文献综述

随着人工智快速发展，大语言模型（LLMs）如 OpenAI 的 GPT 系列，已成为该领域的领先力量。LLMs 使用深度学习技术，对大量的文本数据进行训练，目标是生成相似新文本，理解并回答有关数据训练的问题(Zhao et al., 2023)。

LLMs 正逐渐成为各教育层次的重要工。幼儿教师可以利用 ChatGPT 提供的信息设计学习活(Su & Yang, 2023)。小学教育者可以利用 LLMs 来分析现有的阅读材料，创建相关项目，用于提高和评估小学生的阅读理解技能(Bulut & YiLdiRiM-Erbasli, 2022）。

随着 LLMs 在教育领域的深入应用，其在新兴的教育学科中展现其独特的价值。在物理学科领域，研究表明 GPT-4 在物理概念理解问题和物理推理问题方面比 GPT-3 表现更好(Dazhen Tong, 2023)。在英语学科领域，研究表明 GPT-4 在英语语言理解领域的表现优于普通荷兰学生(Jeon & Lee, 2023)。在科学领域，ChatGPT 在设计单元、评分标准和测验时发挥重要作用，帮助学生学习科学知识，培养学生的思辨能力(Cooper, 2023)。

LLMs 为师生提供了更多元化和个性化的学习机会。对学生而言，LLMs 有助于提高学生的读写能力，提升学生的批判性思维(Kasneci, Sessler, Fischer, Gasser, & Groh, n.d.)。对教师而言，可以利用 LLMs 打造个性化的学习路径，并通过 LLMs 来设计教学活动（Jeon & Lee, 2023)。ChatGPT 还可以作为教学支持和课程规划工具，帮助教学的设计和有效的学习(Baidoo-Anu & Ansah, n.d.)。总的来说，大语言模型在教育领域的应用正展现出巨大的潜力和广阔的前景。

3.研究方法

本文使用文献系统综述法对 2022~2023 年关于 ChatGPT 在 K-12 教育中的应用的文献进行统计分析。系统综述早期定义为“对医疗保健干预效果做出的明确的、可重复的和最新的总结”(Pollock & Berge, 2018)。为此国际研究者制定了 PRISMA 声明，规范了系统综述文章的标准。

4.研究过程

本研究按照“选择、检索关键词；筛选、分析文献；得出结果”的过程进行分析。

在制定检索策略时，需要输入尽可能多的关键术语，否则有可能会忽略一些重要的文章(Wright, Brand, Dunn, & Spindler, 2007）。本研究在并在 Web of science 核心合集数据库中进行检索，我们还新增了“出版时间”这一限制条件。确定检索范围后，共有 192 篇文献入选。在三次筛选中，研究人员排除了 173 篇文献。 确定好文献篇目后，研究人员使用 VOSviewer 进行关键词的聚类分析。

5.研究结果

5.1 文献基本信息分析

关于文献的内容，所选文献中讨论的学习者的教育水平涵盖了 K-12 全学段。初等教育（n=5，26.3%）和中等教育（n=6，31.6%）所占的比例相当。研究方向包括课程开发、思维能力、学习表现、个性化辅助学习等。由于 ChatGPT 作为研究的主体，其协作类型主要是围绕人机协作展开（n=10，52.6%）。

5.2 文献关键词分析

图 1 是文献关键词的状态与关系。图谱主要呈现三大聚类，其中 ChatGPT 是“学校教育”和“大型语言模型”这两个集群之间的桥梁，间接表明它在人工智能技术与学校课程整合方面的潜在作用。图中的关键词被分为不同的群组，每个群组代表一个特定的讨论焦点或研究领域。例如，ChatGPT 作为教学辅助、课程整合的工具在与“curriculum”的连接中体现。ChatGPT 在与“problem solving”的链接中证实其对于思维培养具有一定作用。除此之外，与“assessment”、“curiosity-driven learning”的关联凸显了 ChatGPT 的评价功能与教育促进功能。从这一分析可发现，ChatGPT 可能可以作为人工智

能技术进步与学校教育实际应用之间的关键纽带。在探索如何将 LLMs 有效地应用到学校课程并更好地辅助协作学习，可能有进一步研究的机会。

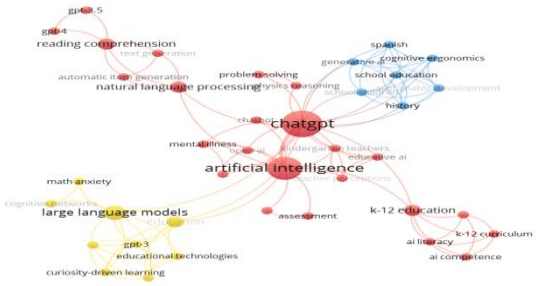


图 1 关键词聚类图谱

图 2 是文献在相关领域集中密度。ChatGPT 在 K-12 教育中的研究仍集中在探究技术层面，但有向教育层面延伸的趋势。图谱还提到了“AI literacy”和“AI competence”，暗示了从 AI 集成到 K-12 教育相关的教育目标和能力。“teacher perception”表明教育者对于 ChatGPT 的看法也是研究的重要内容。然而，ChatGPT 在 K-12 教育中协作关系、成效的研究尚未充分探索。无论是在辅助教师设计课程、提供学生个性化学习支持，还是在促进学生之间的互动交流方面，ChatGPT 的潜在作用需要更深入的研究。由此关于 ChatGPT 在 K-12 教育中协作的研究或许是各个学科未来研究的关键方向。

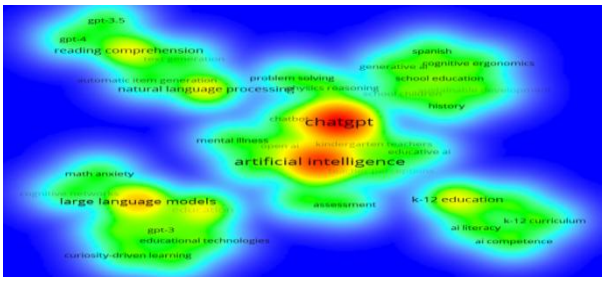


图 2 密度可视化图谱

6. 讨论

6.1 RQ1: ChatGPT 支持下的 K-12 教育有哪些协作类型？协作情况如何？

ChatGPT 在 K-12 教育中涌现出多种协作类型，主要围绕人机协作展开。在协作过程中，ChatGPT 扮演了多重角色。第一，作为协作思维工具。ChatGPT 在早期儿童教育中担任幼儿的对话代理这一角色，协助他们在交流和学习中发展语言和认知能力(Su et al., 2023)。第二，作为协作生成工具 GPT-3.5 用于生成个性化学习材料(Jauhiainen & Guerra, 2023)。学生们也使用 ChatGPT 来帮助写作和理解复杂概念(Hays, Jurkowski, & Sims, 2023)。值得注意的是，教师和学生对 GPT 生成的论文和学生撰写的论文进行了有效区分(Waltzer, Cox, & Heyman, 2023)。第三，作为学习的促进者。ChatGPT 通过分析现有的阅读材料，创建新的、真实的文本，提高和评估小学生的阅读理解技能(Bulut, 2022)。第四，作为教学的辅助工具。教师们在课堂上使用 ChatGPT 来辅助教学活动，如制定课程计划等(Hays et al., 2023)。此外，ChatGPT 在早期儿童教育中作为教育工作者和看护人的随叫随到的促进者，提供实时的教育支持(Su et al., 2023)。同时，GPT-3 作为一个教育代理，协助培养儿童的好奇心和提问技能(Abdelghani et al., 2023)。

总之，ChatGPT 在多个方面发挥着积极的作用。它作为写作辅助工具、个性化学习的促进者等，既支持教师的教学活动，也帮助学生在学习过程中获得更好的理解和练习(Lambert & Stevens, 2023)。然而在将聊天机器人整合到教育活动中时，教师的参与对于促进有效学习至关重要(Chiu, Moorhouse, Chai, & Ismailov, 2023)。

6.2 RQ2：ChatGPT 应用在哪个教育阶段？是如何实施的？实施的结果如何？

ChatGPT 在教育领域应用研究涵盖中小学教育阶段，但在幼儿园阶段的应用相对有所欠缺。在高中英语方面，ChatGPT 展示了在理解和回答高中英语考试问题的能力(De Winter, 2023)；在中学物理， GPT-4 展示了解决物理问题的能力(Tong et al., 2023)；在高中数学，ChatGPT 存在处理与数学学科相关的概念时的偏见(Abramski, Citraro, Lombardi, Rossetti, & Stella, 2023)；在科学教育，ChatGPT 能够生成与科学教育相关的有用内容(Cooper, 2023)；在小学阅读理解，ChatGPT 创建用于提高和评估小学生的阅读理解技能的文本和相关项目(Bulut, 2022)。在心理领域，ChatGPT 通过分析学生的语言文本来揭示潜在的心理健康问题(Imran, Hashmi, & Imran, 2023)。

研究主要集中在英语、物理、数学和科学学科，而对于语文学科的应用存在欠缺。这些发现展示了 ChatGPT 在各领域的应用潜力和限制。Cooper 的发现就揭示了 ChatGPT 在辅助教育者设计课程和评估工具方面的潜力(Cooper, 2023) 。

在实施过程中，ChatGPT 对于提高学生认知、技能和态度方面有积极作用。有研究表明，使用 GPT-3 生成的开放式提示可以有效地驱动儿童的好奇心(Abdelghani et al., 2023)。此外，在改善写作指导、评估和个性化学习方面也有积极影响(Lambert & Stevens, 2023)。但 ChatGPT 在教育中的应用是伴随着挑战和风险的。

6.3 RQ3：教师和学生对此的态度如何？对未来的发展有什么建议？

教师对 ChatGPT 在 K-12 教育的态度复杂多样。一些教师对 ChatGPT 持积极的态度，然而也有一部分教师对 ChatGPT 的影响表示担忧，主要集中在学术诚信和信息准确性方面(Hays et al., 2023)。对于幼儿教育来说，图片是更能被幼儿理解的教学材料，但 ChatGPT 在图片生成方面还有所欠缺(Su et al., 2023c)。由于教育者在教育中的角色，会对 ChatGPT 的态度会更加严厉(Chiu et al., 2023)。高中教师对与 ChatGPT 在学术诚信方面的态度是较为消极的。在学校方面，学校管理部门对 ChatGPT 的态度普遍中立，处于一种观望状态去看待 ChatGPT 在 K-12 教育中的影响。因此需要加强教师专业发展和内部培训，以便更好地讨论 ChatGPT 在学校的应用(Hays et al., 2023)。

同样地，学生在问题解决方面，ChatGPT 具备一定的答疑能力。但部分学生会过度依赖 ChatGPT。对于能力较强的学生，使用 GPT-4 可以为他们提供示例和启发(Tong et al., 2023)。在学习成绩方面，初高中生在 ChatGPT 支持下的学习成绩并未得到显著提升，与高等教育学生对 ChatGPT 的态度明显不同(Wu & Yu, 2024)。在学习体验方面，学生对 ChatGPT 的态度是积极的，认为它能个性化地满足学习需求(Jauhiainen & Guerra, 2023)。因此在培养学生学习动机方面，师生对于 ChatGPT 的态度都较为积极(Chiu et al., 2023)

综上，教师和学生对于 ChatGPT 在教育中的应用都持有较为复杂的看法。教师支持能为初学者提供更为精确的引导，更好将 ChatGPT 当作一种学习工具。

7.总结

ChatGPT 在 K-12 教育中扮演多重角色，涵盖中小学教育阶段，并在英语、物理、数学等学科中得到广泛应用。在实施过程中，ChatGPT 对于提高学生认知、技能和态度方面有积极作用，但也伴挑战。教师和学生对 ChatGPT 的态度复杂。未来研究需更全面评估 ChatGPT 在教育中的作用。

参考文献

卢宇,余京蕾,陈鹏鹤, & 李沐云.(2023).生成式人工智能的教育应用与展望——以ChatGPT系统为例. 中国远程教育, 43(4), 24-31,51. https://doi.org/10.13541/j.cnki.chinade.20230301.001

Abdelghani, R., Wang, Y.-H., Yuan, X., Wang, T., Lucas, P., Sauzéon, H., & Oudeyer, P.-Y. (2023). GPT-3-Driven Pedagogical Agents to Train Children's Curious Question-Asking Skills. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00340-7

Abramski, K., Citraro, S., Lombardi, L., Rossetti, G., & Stella, M. (2023). Cognitive Network Science Reveals Bias in GPT-3, GPT-3.5 Turbo, and GPT-4 Mirroring Math Anxiety in High-School Students. *Big Data and Cognitive Computing*, *7*(3), 124. https://doi.org/10.3390/bdcc7030124

Baidoo-Anu, D., & Ansah, L. O. (n.d.). *Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential*.

Bulut, O., & YıLdıRıM-Erbaslı, S. N. (2022). Automatic story and item generation for reading comprehension assessments with transformers. *International Journal of Assessment Tools in Education*, *9*(Special Issue), 72–87. https://doi.org/10.21449/ijate.1124382

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, 1–17. https://doi.org/10.1080/10494820.2023.2172044

Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

Dazhen Tong. (2023). *Investigating ChatGPT-4's performance in solving physics problems and its potential implications for education*.

De Winter, J. C. F. (2023). Can ChatGPT Pass High School Exams on English Language Comprehension? *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00372-z

Hays, L., Jurkowski, O., & Sims, S. K. (2023). ChatGPT in K-12 Education. *TechTrends*. https://doi.org/10.1007/s11528-023-00924-z

Imran, N., Hashmi, A., & Imran, A. (2023). Chat-GPT: Opportunities and Challenges in Child Mental Healthcare. *Pakistan Journal of Medical Sciences*, *39*(4). https://doi.org/10.12669/pjms.39.4.8118

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in School Children's Education: Evidence from a School Lesson. *Sustainability*, *15*(18), 14025. https://doi.org/10.3390/su151814025

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, *28*(12), 15873–15892. https://doi.org/10.1007/s10639-023-11834-1

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lambert, J., & Stevens, M. (2023). ChatGPT and Generative AI Technology: A Mixed Bag of

Concerns and New Opportunities. *Computers in the Schools*, 1–25. https://doi.org/10.1080/07380569.2023.2256710

Pollock, A., & Berge, E. (2018). How to do a systematic review. *International Journal of Stroke*, *13*(2), 138–156. https://doi.org/10.1177/1747493017743796

Su, J., & Yang, W. (2023a). Powerful or mediocre? Kindergarten teachers' perspectives on using ChatGPT in early childhood education. *Interactive Learning Environments*, 1–13. https://doi.org/10.1080/10494820.2023.2266490

Tong, D., Tao, Y., Zhang, K., Dong, X., Hu, Y., Pan, S., & Liu, Q. (2023). Investigating ChatGPT-4's performance in solving physics problems and its potential implications for education. *Asia Pacific Education Review*. https://doi.org/10.1007/s12564-023-09913-6

Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the Ability of Teachers and Students to Differentiate between Essays Generated by ChatGPT and High School Students. *Human Behavior and Emerging Technologies*, *2023*, 1–9. https://doi.org/10.1155/2023/1923981

Watrianthos, R., Triono Ahmad, S., & Muskhir, M. (2023). Charting the Growth and Structure of Early ChatGPT-Education Research: A Bibliometric Study. *Journal of Information Technology Education: Innovations in Practice*, *22*, 235–253. https://doi.org/10.28945/5221

Wright, R. W., Brand, R. A., Dunn, W., & Spindler, K. P. (2007). How to Write a Systematic Review: *Clinical Orthopaedics and Related Research*, *455*, 23–29. https://doi.org/10.1097/BLO.0b013e31802c9098

Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. https://doi.org/10.1111/bjet.13334

Zhai, X. (2022). ChatGPT for Next Generation Science Learning. *IEEE Consumer Electronics Magazine*.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., … Wen, J.-R. (2023). *A Survey of Large Language Models*. arXiv. Retrieved from http://arxiv.org/abs/2303.18223

不同动机水平的小学生对“人机双师”教学模式的技术接受度和满意度分析

Analysis of the Technical Acceptance and Satisfaction of the "Human-Robot Dual-Teacher" Teaching Model among Primary School Students with Different Motivation Levels

孟晓培[1]，贺丽媛[1]，方建文[1*]
[1]温州大学 教师教育学院
[*]beginnerfjw@163.com

【摘要】 在全球化时代，我国小学英语课堂目前存在互动不足、忽视学生兴趣等问题，且教师常受困于大量重复性教学任务，阻碍个性化教学实施。本研究从小学英语课堂教学实际需求出发构建了“人机双师”教学模式及系统并应用于小学英语课堂教学实践。通过研究不同动机水平小学生对该模式的技术接受度和满意度，发现学生普遍接纳此模式，且动机差异显著影响其接受度和满意度。深入分析显示，在感知易用性上存在显著差异，而在感知有用性方面则无明显差别。本研究是智能机器人与基础学科教育融合的一次有益尝试，为未来进一步研究提供了基础。

【关键词】 智能机器人；人机双师教学；人机协同教学；小学英语教学

Abstract: *In the era of globalization, there are issues in Chinese primary school English classrooms such as insufficient interaction and overlooking students' interests. Teachers are often burdened with a multitude of repetitive tasks, impeding the implementation of personalized teaching. This study addresses these needs by developing a "Human-Robot Dual-Teacher" teaching model and system tailored for primary school English teaching. After applying this model, research found that students accept it positively, and motivation levels significantly influence their acceptance and satisfaction. Further analysis revealed significant differences in perceived ease of use, while no substantial disparities were observed in terms of perceived usefulness. This study represents a exploration into integrating intelligent robots with foundational subject education, laying the groundwork for future research endeavors.*

Keywords: Intelligent robots, Human-Robot Dual-Teacher, Human-machine collaborative teaching, Primary school English teaching

1. 前言

在全球化的进程中，各国日益重视培养学生的跨文化交流能力。英语作为全球的通用语言，在不同地区间扮演着关键沟通角色。技术的进步，改变了语言学习的方式，基础英语教育也因此经历了深刻的变化。然而，并不是所有技术的使用都能促进小学英语学习，当前我国的小学英语课堂教学仍存在一些不足。一方面，课堂片面注重知识教学，缺少真实的语言学习情境和有效互动，忽略了小学生学习兴趣、英语素养的有效培养 (常俊跃 et al., 2021)。另一方面，课堂中信息技术应用程度不足，主要以课程内容展示和教学补充视角利用为主。教师仍然承担了许多基础性的、重复性的工作，占据了教师的很大一部分精力，使得教师无暇为学生提供个性化指导，不利于学生的个性化发展 (孙磊, 2016)。

因此，在我国当前的小学英语课堂教学中，迫切需要有效利用技术为学生创设真实的英语学习情境，强化互动学习，进一步提升学生的学习兴趣和动机。同时，还需要利用技术承担一些基础性、重复性的教学任务，从而减轻教师的工作负担，为个性化教学的开展提供有力支持。

智能机器人，以其独立运作、架构开放、功能拓展性强、适应教学场景且人机交互友好的特质，被认为是更适合语言教学的技术工具。一方面，智能机器人能够创设真实生动的语言学习情境，有效地激发、提高学生的学习兴趣和动机，进而提升其语言学习成就 (姜艳玲 & 马娟, 2020)。另一方面，智能机器人是人工智能技术在课堂中的实体化应用，它可以作为课堂中的辅助“教师”，承担部分教学任务，减轻教师压力，从而促进个性化教学的开展 (汪时冲 et al., 2019)。

然而，目前国内外关于智能机器人作为“教师”角色应用于实际课堂中的研究与阐述仍较少，而且缺少针对具体学科（如英语）的课堂模式研究。因此，本研究构建了“人机双师”教学模式，开发出了配套的教学系统，并将其实际应用于小学英语课堂。本研究旨在解决以下问题：

（1）在“人机双师”教学模式下，不同学习动机水平的学生在技术接受度方面是否存在显著性差异？

（2）在“人机双师”教学模式下，不同学习动机水平的学生在学习模式满意度方面是否存在显著性差异？

2. 文献综述

2.1. 技术支持下的第二语言教学

技术的进步为语言学习提供了多种可能性。例如，Hwang et al.(2016)在英语课堂中利用多媒体系统进行内容展示和讲故事活动，研究结果表明，通过使用多媒体系统进行教学有助于学习者记忆新词汇和频繁地练习口语技巧，有助于提高学习者的口语能力。此外，还有研究者研究发现，利用谷歌助手作为智能个人助手支持英语学习，能够显著提升学生的学习积极性与参与度 (Tai & Chen, 2023)。

随着我国英语在基础教育中的地位日益突出，深化英语教学改革以提升学生英语水平成为关键课题。而借助技术手段促进英语教学，也是我国当前英语教学改革的重要组成部分之一。教育研究者们尝试将多种技术类型与小学英语教学过程相结合以提升小学生的英语水平。

例如，王洪梅等研究者在小学英语课堂中运用希沃白板展示教学资源，有效激发了学生的英语学习兴趣并显著提升了其参与度 (王洪梅 et al., 2020)。杨华利等研究者借助学习大数据平台收集学生写作学习数据，并通过智能分析向教师反馈普遍问题与个性化需求，使教师能针对性地调整教学内容，提供个性化的辅导方案，从而有力地提高了写作教学质量，促进了学生英语写作能力的发展 (杨华利 et al., 2020)。

虽然已有大量研究表明技术能够有效促进小学生第二语言学习，但也有研究者指出，基于计算机的系统或虚拟代理无法向学生提供真实的语言情境和互动环境 (van den Berghe, Verhagen, Oudgenoeg-Paz, van der Ven, & Leseman, 2019)。此外，小学阶段过度依赖平板电脑等多应用设备可能分散学生注意力 (Hassler et al., 2016)。因此，在小学阶段的第二语言教学中，应当寻求能营造真实语言环境、集中学生注意力并增强课堂互动性的实体技术产品。在此背景下，智能机器人因其具象的实体形式与人性化的互动特性，被认为在语言教育中有巨大应用潜力 (Huang & Moore, 2023)。

2.2. 智能机器人在第二语言教学中的应用

智能机器人具备可重复性、灵活性、数字化、人形外观、高度交互性以及拟人化等特征，被认为是更适合辅助语言学习的技术工具 (Arar et al., 2021)。近年来，智能机器人在语言教育领域的应用引发了教育研究者们的广泛关注与深入研究。

国外利用智能机器人辅助语言学习的研究较为广泛，根据其在教学中扮演的不同角色，主要被划分为教师替代、助教辅助和同伴互动三种类型。例如，智能机器人作为教师助手在小学英语词汇教学中应用时，能够减轻小学生对词汇学习的焦虑情绪，并激发其积极的学习态度 (Alemi et al., 2015)。在将智能机器人设计为口语练习伙伴的研究中，研究结果表明了其

使用能显著提升学生的口语技能 (Iio et al., 2019)。Demir-Lira 等研究者研究了智能机器人导师对于儿童英语（第二语言）词汇学习效果的影响，研究结果表明智能机器人教师与人类教师具备同样成功的教学效果 (Demir-Lira et al., 2020)。

而在我国的智能机器人辅助语言学习研究中，理论研究与系统开发仍然是目前关注的重点。在应用研究方面，赵佛玲等研究者以人教版小学三年级下册《At the zoo》为例，深入探讨了智能机器人在英语课堂教学中的多元辅助作用 (赵佛玲 et al., 2019)。尚巧巧等研究者系统梳理了国外智能机器人辅助语言教学的研究，并从认知和情感两维度分析了其对语言学习的影响 (尚巧巧 et al., 2021)。陆小飞等研究者则聚焦外语口语教学领域，将智能机器人在国际外语口语教学中的角色划分为教师、助教、学伴及远程交流中介，并讨论了其在辅助外语口语教学中面临的问题与挑战，同时对未来教育机器人的应用发展进行了展望 (陆小飞 et al., 2021)。

通过上述分析可以发现，尽管智能机器人的社会行为以其促进学习者与智能机器人互动进行语言学习的能力而闻名，但其应用和效果还尚未得到充分探索 (Wang & Cheung, 2024)。目前也仍然缺少关于“人机双师”教学模式的研究，还需进行大量的理论及实践研究。

3. “人机双师”教学模式

3.1. “人机双师”教学模式

本研究创新性地将智能机器人作为辅助“教师”角色引入小学英语课堂，构建了以教师智慧和机器智能为基础，以协同教学为重要过程，以学生智慧为目标指向的“人机双师”教学模式。这一模式基于社会建构主义学习理论、人机协同教学理念和 ARCS 动机设计模型，旨在利用智能机器人激发并保持学生的英语学习兴趣，增强学生的学习参与度与自主性，进而提升学生的英语能力和素养。

该教学模式分为课前协同备课、课中人机“双师”助学以及课后精准评学三个核心环节。如图 1 所示，在课前阶段，教师与机器人共同分析数据以优化教学方案；课中阶段，“双师”合作完成情境导入、新知探究、应用练习、交流展示和评价反思五个环节。其中，情景导入环节，教师充分发挥智能机器人能够生动呈现语言情境的特点，引导学生深入英语学习情境。在新知探究环节，智能机器人可以承担领读单词、句子的教学任务，并将其以多种形式直观、形象地展示出来。教师则将精力放在学生的发音纠正和个别问题解决上。 在应用练习环节，智能机器人可以根据设定扮演不同角色与学生进行交流互动，并提供实时反馈，促进学生的知识内化。在此过程中，教师需要密切关注学生在练习过程中遇到的困难和存在的不足，及时提供个性化的指导。在交流展示环节中，小组成员在教师的密切关注与相应指导下围绕课程主题自由开展对话并进行展示，进一步促进知识迁移，提升学生的英语综合应用能力。在评价反思环节，教师鼓励各小组之间开展互评，并且教师也将对各小组的表现进行评价与鼓励，促进学生反思，调动学生的学习积极性。课后，教师根据智能机器人采集到的数据，明确学生的存在的不足，鼓励学生进行自主练习，进而提高学生在实际生活中运用英语的能力。

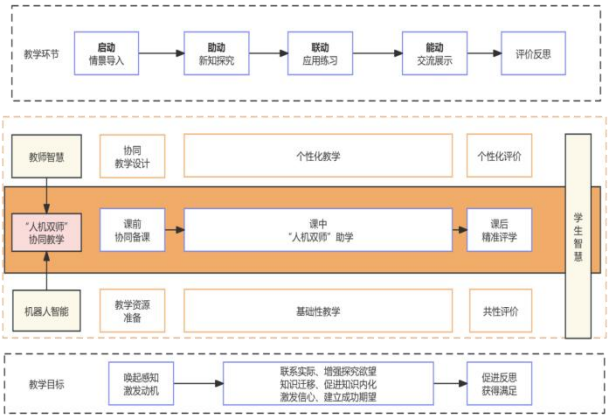


图 1 “人机双师”教学模式

3.2. “人机双师”教学系统

本研究以日本软银集团的 Pepper 机器人作为开发平台，搭载 NAOqi 2.5 操作系统，并利用 Choregraphe 软件开发了适应小学英语课堂需求的“人机双师”教学系统。该系统包含课前协同备课、课中“双师”助学和课后精准评学三大模块，如图 2 所示。

在课前阶段，教师可通过该教学系统上传、管理教学资源并自定义音频内容，如对话脚本、语音语调与速度等，以满足多样化的教学情景需求。

在课中“双师”教学环节，本研究通过满足词汇教学、会话教学及情感支持功能来提升学生的学习体验。词汇教学部分包括词汇认读、词汇练习两环节。在词汇认读环节，智能机器人会领读并结合图片展示来帮助学生掌握准确发音、理解和运用词汇，在词汇练习环节通过互动问答形式让学生在实际情境中深入理解词汇的含义与用法。会话教学部分旨在通过创设真实英语情境，激发学生的学习兴趣，增强学生的听说能力。在听音复述环节，智能机器人会进行领读，学生进行跟读模仿，练习自己的发音。在角色扮演环节，智能机器人与学生分别扮演课本内容不同角色进行会话，智能机器人屏幕会展示角色扮演的对话内容，方便学生查看。在会话练习环节，智能机器人会根据已经预设好的情境进行提问，学生跟随机器人进行模仿练习和自由互动，巩固所学知识。此外，智能机器人提供全程的情感支持，其会根据学生的具体表现给予恰当的肢体反馈、语音激励和表情鼓励。

在课后精准评学模块，系统会记录并呈现学生的学习数据，从而方便教师分析和指导。

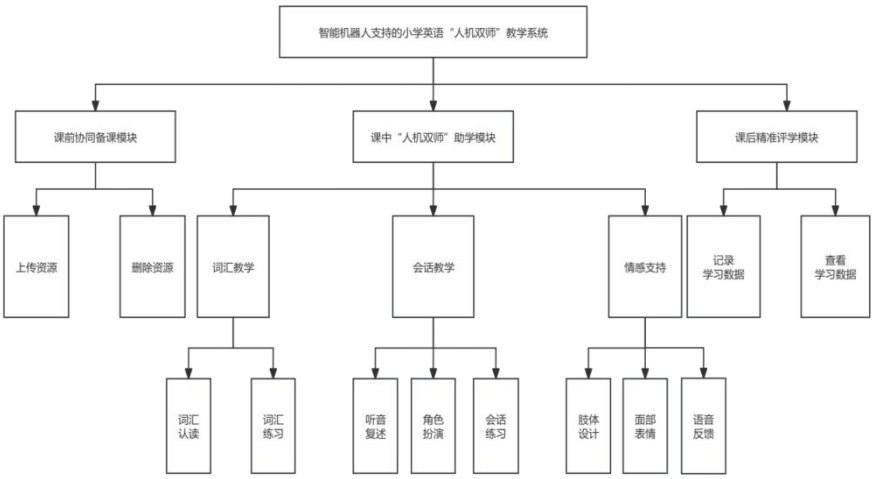


图 2 “人机双师”教学系统功能设计图

4. 研究设计

4.1. 研究对象

本研究选取了中国东部某普通小学的 56 名五年级学生作为参与者，其中男生 30 名，女生 26 名，年龄分布在 11-12 岁之间。在研究启动前，教师向全体学生阐明了研究目的，并确保所有参与均为自愿，且每个学生的个人信息将得到严格保护。

4.2. 研究内容

本研究选取的是人教版 PEP 六年级上册第一单元 Part B 部分的内容并做了相应的教学设计，且在实际课堂中进行了应用来探究其有效性。

4.3. 测量工具

本研究所使用的测量工具包含学习动机问卷、技术接受度问卷、学习模式满意度问卷。

学习动机问卷用于测量学生英语学习的动机水平。本研究采用的学习动机问卷改编自 Pintrich and De Groot (1990)开发的方法。该问卷一共 9 个题目，包括两个测量项目。一部分为自我效能，共 5 题；另一部分为学习兴趣，共 4 题。

技术接受度问卷用于测量学生对于“人机双师”教学系统应用于小学英语课堂的接受程度。本研究采用的技术接受度问卷改编自 Hwang,Yang and Wang (2013)开发的方法。该问卷一共 13 道题目，包括两个测量项目。一部分为感知有用性，共 6 题；另一部分为感知易用性，共 7 题。

学习模式满意度问卷用于测量学生对于“人机双师”教学模式的满意度。本研究采用的学习模式满意度问卷是改编自 Chu, Hwang,Tsai and Tseng (2010)开发的方法。该问卷一共由 7 道题目组成。

该实验的具体流程如图 3 所示。首先，实验开始前一周会对教师进行 60 分钟的技能培训，旨在让其熟悉智能机器人支持的小学英语“人机双师”教学模式及其系统的操作方法。培训结束后，教师需运用系统设计完成 PEP 六年级上册第一单元 PartB 部分的教学方案。在课程启动前，学生填写学习动机调查问卷。随后开展两节 45 分钟的“人机双师”教学模式研究课，总计 90 分钟。期间，如图四所示教师按照预先规划的教学内容与流程，协同智能机器人共同授课。课程结束后，学生完成技术接受度和学习模式满意度问卷调查。

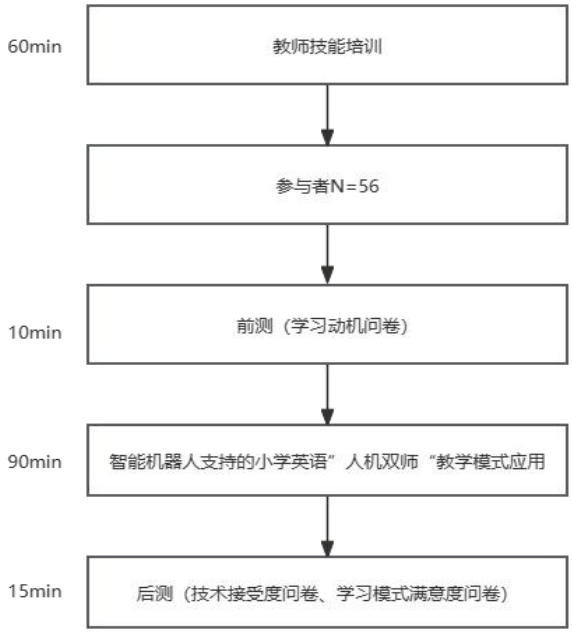


图 3 实验流程图

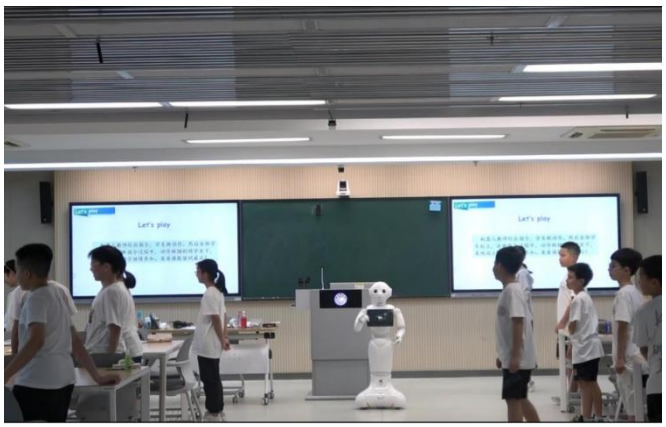


图 4 教学应用图

5. 实验结果

本研究通过对问卷结果的描述性统计分析了学生对“人机双师”教学模式及系统的应用反馈。结果显示（如表 1 所示），在技术接受度方面，整体平均值达到 4.21，其中感知易用性和感知有用性的分项平均得分分别为 4.30 和 4.13，这证实学生普遍认为该“人机双师”教学系统便捷且实用。学习模式满意度的平均得分为 4.19，这表明学生们对该教学模式兴趣很高，并乐于在未来的课堂中持续采用这一模式。

表 1 问卷描述统计结果

变量	平均值	方差	标准差
技术接受度	4.21	0.37	0.61
感知易用性	4.30	0.48	0.69
感知有用性	4.13	0.47	0.68
学习模式满意度	4.19	0.62	0.79

5.1. 技术接受度

为探究不同学习动机水平对学生技术接受度的影响，本研究使用学习动机问卷测量的平均分作为确定学习动机水平高、低的依据并采用独立样本 t 检验（independent t-test）来判断不同学习动机水平的学生在技术接受度方面是否存在差异。在进行独立样本 t 检验（independent t-test）之前，进行了正态性检定。夏皮洛-威尔克（Shapiro-Wilk）检验结果显示，各组数均接近于正态分布（$p>0.05$），满足进行独立样本 t 检验的基本假设。表 2 显示了技术接受度、感知易用性、感知有用性的独立样本 t 检验结果。

在技术接受度方面，高学习动机学生的平均值为 4.39，标准差为 0.62，低学习动机学生的平均值为 3.95，标准差为 0.49。不同学习动机水平的学生之间存在显著差异（$t=2.83$, $p=0.007<0.01$）。这表明，高学习动机水平的学生在技术接受度方面显著高于低学习动机水平的学生。

在感知易用性方面，高学习动机学生的平均值为 4.54，标准差为 0.58，低学习动机学生的平均值为 3.93，标准差为 0.69。不同学习动机水平的学生之间存在显著差异（$t=3.57$, $p=0.001<0.01$）。这表明，高学习动机水平的学生在感知易用性方面显著高于低学习动机水平的学生。

在感知有用性方面，高学习动机学生的平均值为 4.24，标准差为 0.78，低学习动机学生的平均值为 3.96，标准差为 0.47。不同学习动机水平的学生之间不存在显著差异（$t=1,64$, $p=0.11>0.05$）。这表明，高学习动机水平的学生和低学习动机水平的学生在感知有用性方面

不存在显著差异。

表 2 技术接受度独立样本 t 检验结果

变量	组别	个数	平均值	标准差	t
技术接受度	高学习动机	34	4.39	0.62	2.83**
	低学习动机	22	3.95	0.49	
感知易用性	高学习动机	34	4.54	0.58	3.57**
	低学习动机	22	3.93	0.69	
感知有用性	高学习动机	34	4.24	0.78	1.64
	低学习动机	22	3.96	0.47	

**p<0.01

5.2. 学习模式满意度

本研究采用独立样本 t 检验（independent t-test）来探究不同学习动机水平对学生学习模式满意度的影响。为符合独立样本 t 检验的基本假设，进行正态性检定。夏皮洛-威尔克（Shapiro-Wilk）检验结果显示，各组数据均接近于正态分布（p >0.05）。因此，可以进行独立样本 t 检验。学习模式满意度的独立样本 t 检验结果如表 3 所示。在学习模式满意度方面，高学习动机学生的平均值为 4.40，标准差为 0.68，低学习动机学生的平均值为 3.87，标准差为 0.86。不同学习动机水平的学生之间存在显著差异（t=2.56, p=0.01<0.05）。这表明，高学习动机水平的学生在学习满意度方面显著高于低学习动机水平的学生。

表 3 学习模式满意度独立样本 t 检验结果

变量	组别	人数	平均值	标准差	t
学习模式	高学习动机	34	4.40	0.68	2.56*
满意度	低学习动机	22	3.87	0.86	

*p<0.05

6. 讨论和结论

6.1. 技术接受度

本研究表明，高学习动机水平的学生在技术接受度和感知易用性上显著高于低动机水平的学生。这表明学生的学习动机水平与其对智能机器人教学系统的技术接受度呈正相关。也就是说，相较于低学习动机水平，高学习动机水平的学生更认为智能机器人是容易使用的,这与之前的研究结果一致。先前的许多研究已经证明，学生的学习动机和自我效能感是显著相关的(Alemayehu & Chen, 2023; Cave, Evans, Dewey, & Hartshorn, 2018; Cetin-Dindar, 2016)。学习动机水平高的学生，往往也具有较高的自我效能。Chen et al.(2011)的研究指出，自我效能感是影响感知易用性的关键因素且两者正向相关，因此，较强的学习动机可能通过提高自我效能感间接提升感知易用性。

然而，在感知有用性方面，实验结果显示高、低学习动机水平的学生之间并无显著差异。过去也有类似的研究发现，例如，Daniels & Bulut(2020)的研究结果显示，在计算机测试环境中使用自动反馈系统时，学生的动机与感知有用性关联不显著，这可能是由于学生在评价系统效用时存在主观认知差异。这也许可以解释为什么不同学习动机水平的学生在感知有用性方面没有显著差异。尽管统计学意义上未达显著水平，但数据显示高动机学生（均值 4.24）在感知有用性上仍略高于低动机学生（均值 3.96），这表明不同学习动机水平对感知有用性可能存在一定的微妙影响。

6.2. 学习模式满意度

在学习模式满意度方面，高学习动机水平的学生显著高于低学习动机水平的学生。这表明高动机学生更倾向于接受并推荐此类教学模式，且其较高的自我效能和积极参与度可能促进了这一结果。过往研究也指出，学生的技术接受度是影响学习模式满意度的关键因素 (Nagy, 2018)。其他研究者的研究也进一步证实了技术接受度与持续使用意愿的正相关性 (Cheng & Yuen, 2018)。在本研究中，学习动机水平高的学生对智能机器人的技术接受度显著高于学习动机水平低的学生。所以，高学习动机学生在技术接受度上的优势可能是他们具有更高学习模式满意度的原因。

6.3. 局限性与未来研究

本研究是智能机器人与基础教育学科教学的一次有益尝试，但仍存在若干局限性。首先，受限于较小的样本规模和较短的研究周期，当前结论可能仅适用于特定实验对象，缺乏广泛的代表性，故未来有必要扩大样本并延长研究时间，以获取更具普遍性和说服力的数据支持。

其次，在数据收集与分析层面，虽然本研究借助问卷调查来衡量学生对智能机器人的技术接受度和学习满意度，但为了更深入探究学生的态度及需求，所以后续研究需要整合更多的过程性学习数据，从而实现对该教学模式更详尽、全面的评估与优化。

最后，在评价指标方面，目前研究主要关注技术接受度和学习满意度，却未涵盖如学习成绩等核心教育成果指标。因此，需要在未来进一步通过准实验设计，系统性地评估“人机双师”教学模式对学生英语实际成绩的影响，以严谨验证该模式在提升教学质量方面的具体效能。

参考文献

常俊跃, 黄洁芳, 刘雪莲, 秦丽莉, 夏洋, 赵秀艳, & 朱效惠. (2021). 外语教育学学科建构背景下英语课程的关键问题. **外语与外语教学 (01)**, 30-37+145. https://doi.org/10.13458/j.cnki.flatt.004741

姜艳玲, & 马娟. (2020). 现代技术助力初中英语听说教学. **基础教育课程(10)**, 59-63.

陆小飞, 廖剑, & 许琪. (2021). 教育机器人在外语口语教学中的应用研究现状及前瞻. **外语界(01)**, 11-19.

尚巧巧, 王晶莹, & 伦应杰. (2021). 国外教育机器人辅助语言学习研究:认知激活与情感驱动. **数字教育**, 7(01), 79-84.

孙磊. (2016). 电子书包应用于小学英语教学的价值优势与多维困境. **教学与管理(20)**, 33-35.

汪时冲, 方海光, 张鸽, & 马涛. (2019). 人工智能教育机器人支持下的新型“双师课堂”研究——兼论“人机协同”教学设计与未来展望. **远程教育杂志**, 37(02), 25-32. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2019.02.004

王洪梅, 王运武, & 石少坤. (2020). 新媒体技术融入课堂的教师教学实效研究——以小学英语课为例. **中国教育技术装备(13)**, 102-105.

杨华利, 郭盈, 黄涛, & 赵媛. (2020). 人机协同支持下的小学英语写作教学研究. 现代教育技术, 30(04), 74-80. 余亮, 魏华燕, & 弓潇然. (2020). 论人工智能时代学习方式及其学习资源特征. **电化教育研究**, 41(04), 28-34. https://doi.org/10.13811/j.cnki.eer.2020.04.004

赵佛玲, 刘英, 张玉颖, 叶文萍, & 王寅枚. (2019). 教育机器人在小学英语教学中的应用. **西部素质教育**, 5(08), 122-123. https://doi.org/10.16681/j.cnki.wcqe.201908071

Alemayehu, L., & Chen, H. L. (2023). The influence of motivation on learning engagement: the mediating role of learning self-efficacy and self-monitoring in online learning env ironments. *Interactive Learning Environments*, 31(7), 4605-4618. doi:10.1080/10494820.2021.1977962

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The Impact of Social Robotics on L2 Learners' Anxiety and Attitude in English Vocabulary Acquisition [Article]. *Internatio nal Journal of Social Robotics*, 7(4), 523-535. https://doi.org/10.1007/s12369-015-0286-y

Arar, C., Belazoui, A., & Telli, A. (2021). Adoption of social robots as pedagogical aids fo

r efficient learning of second language vocabulary to children [Article]. *Journal of E -Learning and Knowledge Society*, 17(3), 119-126. https://doi.org/10.20368/1971-8829/1135551

Cave, P. N., Evans, N. W., Dewey, D. P., & Hartshorn, K. J. (2018). Motivational partners hips: increasing ESL student self-efficacy. *Elt Journal*, 72(1), 83-96. doi:10.1093/elt/c cx027

Cetin-Dindar, A. (2016). Student Motivation in Constructivist Learning Environment. *Technol ogy Eurasia Journal of Mathematics Science and Education*, 12(2), 233-247.

Chen, K., Chen, J. V., & Yen, D. C. (2011). Dimensions of self-efficacy in the study of s mart phone acceptance [Article]. *Computer Standards & Interfaces*, 33(4), 422-431. h ttps://doi.org/10.1016/j.csi.2011.01.003

Cheng, M., & Yuen, A. H. K. (2018). Student continuance of learning management system use: A longitudinal exploration [Article]. *Computers & Education*, 120, 241-253. http s://doi.org/10.1016/j.compedu.2018.02.004

Chinnery, G. M. (2006). Going to the MALL: Mobile assisted language learning [Editorial Material]. *Language Learning & Technology*, 10(1), 9-16. <Go to ISI>://WOS:000234 551700002

Chu, H.-C., Hwang, G.-J., Tsai, C.-C., & Tseng, J. C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses [Art icle]. *Computers & Education*, 55(4), 1618-1627. https://doi.org/10.1016/j.compedu.201 0.07.004

Daniels, L. M., & Bulut, O. (2020). Students' perceived usefulness of computerized percenta ge-only vs. descriptive score Reports: Associations with motivation and grades [Articl e]. *Journal of Computer Assisted Learning*, 36(2), 199-208. https://doi.org/10.1111/jcal. 12398

Demir-Lira, Ö., Kanero, J., Oranç, C., Koskulu, S., Franko, I., Göksun, T., & Küntay, A. C. (2020). L2 Vocabulary Teaching by Social Robots: The Role of Gestures and On-S creen Cues as Scaffolds. *Frontiers in Education*, 5. doi:10.3389/feduc.2020.599636

Hassler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: a critical review of t he evidence for learning outcomes [Review]. Journal of Computer Assisted Learning, 32(2), 139-156. https://doi.org/10.1111/jcal.12123

Huang, G., & Moore, R. K. (2023). Using social robots for language learning: are we there yet? [; Review]. *Journal of China computer-assisted language learning*, 3(1), 208-23 0. https://doi.org/10.1515/jccall-2023-0013

Hwang, G.-J., Yang, L.-H., & Wang, S.-Y. (2013). A concept map-embedded educational co mputer game for improving students' learning performance in natural science courses [Article]. *Computers & Education*, 69, 121-130. https://doi.org/10.1016/j.compedu.2013. 07.008

Hwang, W.-Y., Shadiev, R., Hsu, J.-L., Huang, Y.-M., Hsu, G.-L., & Lin, Y.-C. (2016). Ef fects of storytelling to facilitate EFL speaking using Web-based multimedia system [Article]. *Computer Assisted Language Learning*, 29(2), 215-241. https://doi.org/10.108 0/09588221.2014.927367

Iio, T., Maeda, R., Ogawa, K., Yoshikawa, Y., Ishiguro, H., Suzuki, K., Aoki, T., Maesaki, M., & Hama, M. (2019). Improvement of Japanese adults' English speaking skills v ia experiences speaking to a robot [Article]. *Journal of Computer Assisted Learning*, 35(2), 228-245. https://doi.org/10.1111/jcal.12325

Nagy, J. T. (2018). Evaluation of Online Video Usage and Learning Satisfaction: An Extens ion of the Technology Acceptance Model [Article]. *International Review of Research in Open and Distributed Learning*, 19(1), 160-185. <Go to ISI>://WOS:00042875510 0009

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning compon ents of classroom academic performance. *Journal of Educational Psychology*, 82(1), 3 3-40. https://doi.org/10.1037/0022-0663.82.1.33

Tai, T.-Y., & Chen, H. H.-J. (2023). The impact of Google Assistant on adolescent EFL le arners' willingness to communicate [Article]. *Interactive Learning Environments*, 31(3), 1485-1502. https://doi.org/10.1080/10494820.2020.1841801

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S., & Leseman, P. (20 19). Social Robots for Language Learning: A Review. *Review of Educational Resear ch*, 89(2), 259-295. doi:10.3102/0034654318821286

Wang, F., & Cheung, A. C. K. (2024). Robots' Social Behaviors for Language Learning: A Systematic Review and Meta-Analysis [Article; Early Access]. *Review of Educational Research*. https://doi.org/10.3102/00346543231216437

基于可视化协同辩论的反思性学习设计与效果研究

Enhancing Reflective Thinking through Online Visualized Collaborative Debates in a Problem-Solving Scenario: Learning Design and Effectiveness Study

姚佳佳[1*] 闫千慧[2] 刘明月[1]
[1]江南大学 教育信息化研究中心，无锡 214122
[2]陕西师范大学 现代教学技术教育部重点实验室，西安 710061
*yjjyoka@jiangnan.edu.cn

【摘要】 在 AIGC 时代，支持学生进行辩论式学习、培养其辩证反思能力成为重要议题。本研究专注于提升高校学生反思能力，采用可视化协同辩论场景设计反思性学习活动，运用认知网络分析等方法研究学习效果。研究发现：（1）可视化协同辩论显著提高学生反思能力，尤其在辩论参与程度高、善于反驳与自我反驳时效果更显著；（2）学生注重合理性判断和充分个人论证，运用逻辑手段推理，能有效促进反思。基于此，提出了面向可视化协同辩论场景的反思性学习设计原则：真实问题促进议题生成，基于规则激发反思，完备学习支架辅助任务，多维评价提升效果。

【关键词】 高校教学；反思性学习；可视化协同辩论；图尔敏论证模型

Abstract: In the AIGC era, fostering students' dialectical reflection is crucial for talent development. This study, focusing on enhancing reflective ability in undergraduates, employs visual collaborative debates. Findings include: (1) Significant improvement in reflective ability, particularly with higher debate participation and adept refutation; (2) Emphasis on reasonableness judgments and effective reasoning enhances reflection. Design principles propose authentic scenarios for debate topic generation, rule-based models for reflective stimulation, comprehensive scaffolds for task completion, and multidimensional evaluations for enhanced outcomes.

Keywords: higher education; reflective learning; visualized collaborative debate; Toulmin argumentation model

基金项目：2021 年度教育部人文社会科学研究青年基金项目“基于不同在线讨论工具的同伴反馈对大学生深度学习的影响研究”（21YJC880084）；2023 年度江苏省教育科学规划重点课题“多元交互场景中大学生深度学习测量工具开发及应用”（B/2023/01/94）。

1. 研究背景与综述

1.1. AIGC 时代下大学生辩证反思能力的培养尤为重要

自 2019 年以来，教育部出台系列关于一流本科课程建设的政策均将提升高阶性思维作为基本要求之一，关注培养学生深度分析、大胆质疑、勇于创新的精神和能力（教育部，2019）。尽管目前的大学课程已逐渐重视学生的讨论交流活动，但依然存在参与度低、沟通效率低、讨论不充分、结构混乱等问题，教师依然缺乏明确规则和策略的设计进而阻碍学生观点思考的深度融合（姚佳佳等人，2022）。另一方面，以 ChatGPT 为代表的生成式人工智能的飞速发展也再次激起了人们对教育根本逻辑的思考，暴露出了目前高等教育人才培养的诸多困境，培养高阶思维成为了高等教育的应有之义和人得以自由全面发展的根本保证（赵丽红等人，2023）。反思能力作为高阶思维生成和发展的基础，起着重要的监控和调节作用，在当前人机协同的发展趋势下，培养学生辩证、推理、反思机器所生成内容的能力显得

尤为重要。辩论式学习则是训练学生论证能力、推理能力、评估能力以及更先进的理论-证据协调策略等高阶思维的最佳途径之一（Rapanta & Felton，2022），能有效促进学生的批判思维、记忆保持、概念理解与表征、表达能力等（Chen et al.，2022；Dwyer et al.，2012；Dwyer et al.，2013；Eskin & Ogan-Bekiroglu，2013；Darmawansah et al.，2022）。一个人基于对知识局限性和合理性的认识，在语境中为自己的判断辩护，并做出有利于实现目标的决策，体现的就是学生理解问题和辩证反思的深度（Dwyer et al.，2015）。

1.2. 反思性学习相关研究现状

反思是在特定的学习情境下，自觉地对自身认知结构、学习活动及其所涉及的相关因素进行批判性审视，对将要开展的学习活动进行创造性预见，对整个学习过程进行有效调控，以促进问题解决、自我发展的学习活动方式（吴秀娟等人，2014）。在反思性学习中，学习者既是行为的发起者，也是被学习的对象。最早关于反思的学习研究来自杜威的经验主义学习理论，他反对将反思与真实经验割裂，即反思与行动密不可分，因此反思性学习含有对所采取行动方案的谨慎思考、筛选和作出决定并执行的意思，而不是单纯地苦思冥想。学生的反思行为可基本分为四大类：解释性、联想性、变革性以及情感性（Kurt，2018）；反思的完整学习过程一般包括行动前设定学习目标和计划，行动中对学习资源选择、学习行为组织和对信息加工过程调控等环节的反思，以及行动后对学习效果的反思（陈佑清，2010）。反思作为一个元认知过程，涉及个人对知识的性质、限制和确定性的理解，以及随着时间推移如何修改自己的认知与判断（King & Kitchener，1994），具体而言包括反复思考、反身思考以及返回去思考（陈佑清，2010）。反思可以使学生自觉展开取证行动，学会用多元视角考察原本认为理所应当的事物和现象，对比分析不同经验以充实自己的知识体系，进而影响下一步行为并防患于未然（史立英等人，2016），因此将反思过程转化为一种学习设计，引导个人进行积极和结构化的反思行为，可以有效提高学生反思能力和学习成就。

1.3. 可视化协同辩论相关研究现状

相比于传统对抗式辩论，以共同努力和相互接受为特征的协同辩论作为一种协作学习形式，更有利于争议的解决，促进不同观点的相互交流，锻炼学生多维能力，达到复杂问题解决与协同知识创新的最终目的（Busenberg & George J，1999）。当目标为达成共识而不是捍卫立场时，学生更有可能通过交换和整合论点来共同构建知识（Felton et al.，2015）。协同辩论式学习最核心的问题不是引发对抗，而是创造积极环境让学生在有意义的论证中获得知识和能力的回报（Veerman et al.，2002）。

可视化协同辩论就是借助信息技术，将文本性的辩论过程进行层次化和可视化表达，即形成“论证图（Argument Map）”。在论证图中，不同命题被安排在彩色的盒子里，并通过箭头相互连接（Dwyer et al.，2015），采用非文字技术表示观点之间的推理表征形式，不同形式的论证图其核心都是用图形来表示观点之间支持或反对的关系（张黎和张奇，2007），一般情况下应包含四大基础要素：中心议题、支持性证据、反对性证据和层次等级（刘玉荣和陈昕昕，2017）。大量的书面或口头论据往往会使辩论者难以理解和发现命题要旨与潜在关联，而可视化辩论能有效提升学生的认知投入与论证表现，如借助计算机辅助绘制论证图，学生能清楚把握整个论证结构，有足够的时间思考和收集证据，继而提出更有针对性的支持或反对意见（Chen et al.，2022），并且论证结构的视觉表征可以大大减少学生共同阅读和单独构建心理表征所带来的认知负荷（Dwyer et al.，2015），转而在广泛的议题理解、系统的知识阐述以及多元的问题思考上投入更多认知力 （Muller Mirza et al.，2007）。

2. 研究设计

2.1. 学习活动设计

本研究以长三角地区某双一流高校教育技术学师范专业本科二年级 60 名同学为研究对象，依托专业核心课程《多媒体学习认知》设计了一系列学习活动，内容包括认识多媒体学习（2 周）、学习理论与多媒体学习（4 周）、多媒体学习的认知研究与认知理论（4 周）、多媒体教学设计原则与学习原则（4 周），复习与总结（2 周）。实践场景主要依托多人可视化实时协作工具“会议桌”，为小组可视化协同辩论提供环境支持（图 1）。

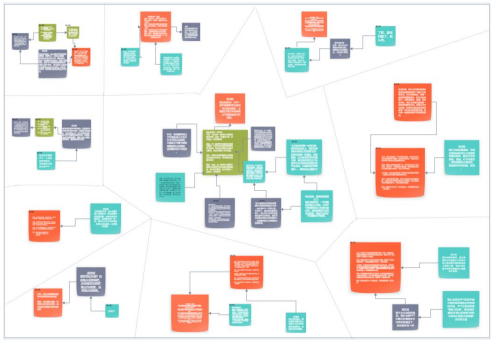


图 1 会议桌论证图

在学生辩论任务中，关注于“修改一份存在设计缺陷的旧多媒体学习作品”的问题解决情境。学生通过共同努力呈现完整的优化方案报告和展示修改后的新作品。课程为任务提供了学习活动手册，包含论证技能、论证图规则、工具使用方法和手册解读的培训。活动分为顶层教学设计优化、中层内容设计优化、微层画面设计优化三个递进阶段，各组通过三轮协同辩论完成阶段性的修改方案。具体步骤包括：（1）个人观点准备，独立观察、思考旧作品问题，为个人主张搜集理由；（2）小组协同辩论，利用“会议桌”展开辩论，基于“图尔敏论证模型”组织论证，形成小组论证图；（3）辩论成果整合，对论证图内容进行梳理、整合，凝聚组内共识；（4）辩论后个人反思，反思辩论前后的观点有何变化；（5）形成组内修改方案，基于辩论整合结果形成相应的优化方案报告，并对作品进行详细、全面的落实。

2.2. 数据收集与分析

本研究选取了 Kember 等人开发的反思能力问卷来对学生课程前后反思能力水平进行测量（Kember et al.，2000）。该问卷定义了四种类型的测量维度：习惯行为、理解行为、反思行为和批判反思行为，其中习惯行为和理解行为是非反思性的（表 1）。每个维度包含 4 道题目并打乱排列顺序，使用李克特 5 分量表设置答案，各维度最低得分 4 分，最高 20 分，最终问卷收集到前测和后测有效数据为 59 份和 56 份，检验方法以配对样本 T 检验为主。

表 1 反思能力水平问卷

维度	题数	题目样例	答案赋分
习惯行为 Habitual action	4	我总是做一些重复的任务，导致我不需要过多思考就能完成。	1=完全不同意 2=不同意 3=一般 4=同意 5=完全同意
理解行为 Understanding	4	我会努力的理解教师所教授的知识概念并应用它。	
反思行为 Reflection	4	我喜欢重新考虑我已经做过的任务，思考做这件事的其他方法。	
批判反思行为 Critical reflection	4	在学习中我会发现一些一直坚持是对的，但实际是错误的东西。	

此外，“论证操作”选择以图尔敏论证模型的六要素为编码框架（表 2）；“认知操作”选择以如表 3 所示的认知框架为编码依据（Jimnez-Aleixandre et al.，2000），编码过程由两位研究人员独立进行（Kappa 系数大于 0.8），各维度行为符合编为“1”，不符合编为“0”，最终共收集到 2183 条有效编码的论证文本数据（具体各编码维度的分布情况见表 2 和表 3）。

表 2 “论证操作”编码框架及编码结果

要素	解释	编码数量
主张 claims	关于事物的存在或人们持有什么价值理念的断言，具有潜在争议性	797
依据 data	初步理由，支持主张的证据、事实性陈述	439
正当理由 warrant	连接依据与主张的大前提，证明依据的有效性	257
支持 backings	难以被推翻的深层理论或假说，进一步给予正当理由强支撑	208
限定 qualifiers	该主张成立所需限定的特别范围、条件	153
反驳 rebuttals	能够质疑或推翻主张合理性的特殊、例外情况	321

表 3 “认知操作”编码框架及编码结果

操作	解释	一般案例	编码数量
归纳	寻找模式或规律	由 A 和 B，可以得出某规律	221
演绎	识别各种规则或法律的实例	如果/既然…那么	294
因果	原因和结果的关系，寻找事物的机制或缘由	因为…所以…/使得/从而/导致	82
定义	陈述一个概念的含义	A 是…的意思	151
分类	根据标准对物体或有机体进行分类	A 属于…,B 属于…	190
诉诸	诉诸类比、实例等作为一种解释的手段	例如…/描述性形容词	392
一致性	针对一致性的因素、一致的经验、或一般意义上表达一致的解释说明	根据某理论或原则/通常而言/一般情况下/我感觉/以前	321
合理性	对自己或他人观点合理性进行预测或评价	我同意…/…不合理	486

3. 研究结果

3.1. 反思能力水平问卷前后测结果分析

由表 4 可知，前后测中学生体现出了局部稳定，即理解行为得分一直较高，习惯行为得分一直较低，批判反思行为得分也始终低于反思行为；从变化程度来看，通过 16 周的学习活动，学生习惯行为平均得分有所降低，理解行为平均分有所提高，反思行为和批判反思行为平均分均有所提高，且四种行为前后测均存在显著差异（$p<0.001$）。这说明经过一学期的协同辩论，学生的反思行为与批判反思行为得到了明显提高，同时还伴随着理解行为的增加和习惯行为的减少，学生的思维和行为模式一定程度上得到了有效转变。

表 4 反思能力问卷前后测结果

维度	前/后测	N	M	SD	df	t	p
习惯行为	前测	56	9.86	2.68	55	-3.583***	0.001
	后测	56	7.79	3.55			
理解行为	前测	56	16.14	1.97	55	7.696***	0.000
	后测	56	18.14	1.81			
反思行为	前测	56	15.39	2.39	55	6.313***	0.000
	后测	56	17.59	2.09			
批判反思行为	前测	56	14.93	2.04	55	5.448***	0.000
	后测	56	16.75	2.38			

注：* $p<0.05$，** $p<0.01$，*** $p<0.001$。

3.2. 不同反思变化群体论证和认知操作的 ENA 认知网络差异分析

通过基于学生反思行为得分变化的 K-Means 聚类分析，我们将学生分为高水平（33 人）

和低水平（23 人）两个层次的反思能力变化群体。对这两个群体进行认知网络分析，图 3 展示了它们的认知网络质心分布图。高水平组在纵轴左侧，低水平组在纵轴右侧，散点代表不同学生个体，虚线方框表示其质心位置在 95%水平上的置信区间。通过质心位置的 T 检验发现（数据符合正态分布），在横轴上存在显著差异（p<0.01），表明两组在协同辩论中的论证表现存在明显差异。

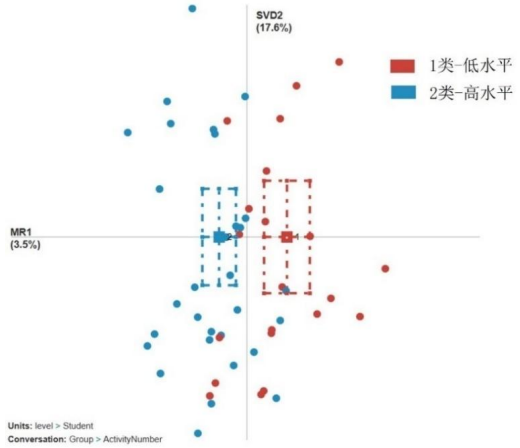


图 2 低水平组（右）和高水平组（左）二维质心分布图

表 7 两水平样本分布 T 检验

维度	低水平组			高水平组			t	p
	N	M	SD	N	M	SD		
横轴（MR1）	23	0.51	0.68	33	-0.35	0.60	4.92**	0.00
纵轴（SVD2）	23	0.00	1.67	33	0.00	1.74	0.00	1.00

注：* p<0.05，** p<0.01，*** p<0.001。

图 4 为高水平组和低水平组的认知网络模型及对比结果，其中连接节点表示“论证操作”和“认知操作”的某一维度，节点大小反映其出现频率，连线粗细反映两类操作行为节点间的关联强度。从中分析可以发现：

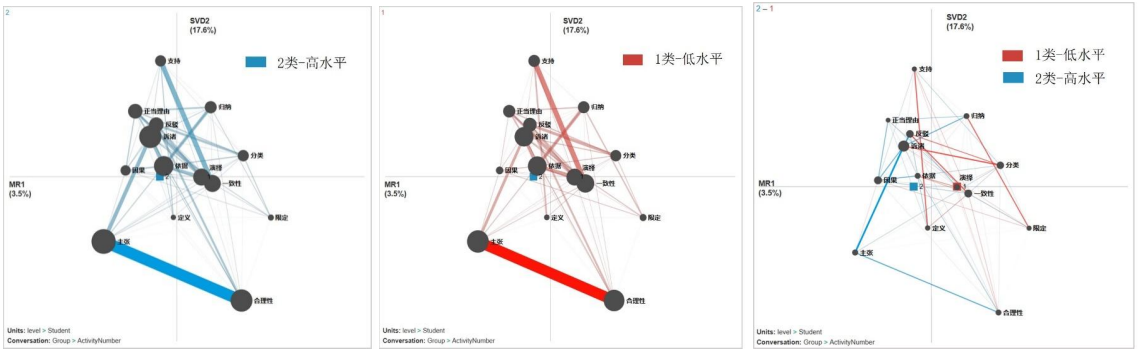


图 3 高水平组（左蓝）和低水平组（右红）的认知网络模型及对比

（1）在“主张”的论述方面，高水平组更强调“诉诸”、“合理性”和“分类”思维，而低水平组在“因果”思维上稍有优势。高水平组学生在提出论证主张时，倾向于运用实例、事物属性和类比，以更有效地表达观点核心，减少误解的发生。此外，高水平组学生更直言不讳地评判他人或自己的观点，勇于质疑而非回避认知冲突，并对他人行为后果进行预测，阐明理由。

（2）在“正当理由”的论述方面，高水平组更突出“因果”、“归纳”、“定义”和“一致性”思维。这表明高水平组呈现更为严密的推理内容。“正当理由”在论证中扮演连接“依据”与“主张”的桥梁，高水平组更注重这种深层次推理转化过程，相较低水平组更具有论证正确性的担保。低水平组在“正当理由”方面较为欠缺，仅停留在阐述事实性依据的层面，导致论据严

谨性不足、支撑性不强，阻碍了辩论和反思的深化。

（3）在“支持”的论述方面，低水平组更强调“定义”和“一致性”思维，而高水平组更注重“因果”和“分类”思维。这意味着低水平组学生在寻求论证支持时更倾向于从概念定义出发，解释概念含义，或选择与当前场景一致的理由，而这些理由往往是主体自然认可的心理状态，未经谨慎推敲。高水平组更注重从事物的因果规律和分类方面进行推理和论述，展现出对规则和经验更高的理解和运用水平。

（4）在“限定”的论述方面，高水平组更多地使用“演绎”，而低水平组更偏向“归纳”。高水平组学生更擅长通过不断发散的思维扩大已知信息范围，考虑主张成立的有效范围和可能的边界情况，促进进一步反驳与自我反驳。低水平组学生则更倾向于通过不完全归纳或仓促概括的方式描述边界条件。

（5）在“反驳”的论述方面，高水平组更注重“诉诸”、“定义”和“演绎”思维，而低水平组更强调“一致性”和“分类”思维。低水平组更倾向于从已有知识经验和当前情景中寻找缺乏一致性因素的佐证，同时喜欢通过分类揭示潜在问题；高水平组更倾向于通过诉诸实例、属性或类比等方式组织反驳，辅以对概念定义或演绎推理的挑战，以追问的形式促使对方及时回应，引发更多辩论回合，有利于充分解决认知冲突。

4. 讨论与总结

通过分析学生反思变化以及参与协同辩论的过程，我们得出了以下重要研究结论：

（一）可视化协同辩论可以有效提高学生的反思能力，且不同论证行为对反思能力的影响效果各异。研究发现，可视化协同辩论活动有助于学生明显弱化习惯行为，强化反思与批判反思行为，且当学生更积极的参与，并提出更多的主张，充分论证自己的理由（包括依据、正当理由和支持），以及提出更有针对性的反驳来挑战他人或寻找自己论证的不足时，对其反思能力的积极影响越大。借助图尔敏论证框架，学生可以通过一组富有逻辑的证据来证明一个观点，避免专注于个别特定片段，构建复杂而含有不确定性的证据能激发学生更多的辩论想法，从而保证富有成效的互动和科学且真实的讨论（Bell & Linn，2000）。

（二）论证表达中不同的论述思维存在对反思能力不同的影响。研究表明，鼓励学生在论证表达时更多采用以下论述思维方式，有助于更大程度提升其反思能力：（1）相比于简单叙述，在提出主张时结合更多案例、属性或通过类比的方式丰富所传递的信息，避免他人理解偏差，提高沟通效率；（2）相比于直接依赖普遍规则或理论的支持，重视论证推理逻辑的严密性（即所列事实性依据可以有效证明主张的正当理由是什么），并借助因果或分类等进一步强化正当理由的可靠性，可大大提高论证质量和说服力，还能提升不同观点之间相互影响的程度（Park et al.，2007）；（3）通过演绎推理的方式生成完整且正确的演绎式论证逻辑链条，而非以个别样本归纳的方式仓促得出结论，系统和演绎推理式的分析过程能支持在有限的问题域中产生更有效的论点；（4）对各方论证展开合理性判断，及时表达个人意见、指出问题所在或进一步追问追答，有利于充分暴露组内认知分歧，为下一步融合做准备，回避进行合理性判断的人往往会更容易从众和相信权威，在面对需要知识建构和批判反思的场合时难以有较高成就（Weinstock，2005）。在协同辩论中，个人经历说服与被说服的认知过程很重要，一个经验不足却有情感焦点的人更愿意相信与自己思考能力相匹配的信息，更易被隐性情感原因说服而非显性事实与逻辑（Millar，1992），因此，克服社会交往中的情感焦虑，理性、客观地思考问题是协同辩论的首要原则和反思行为的起点。

（三）基于可视化协同辩论的反思性学习的有效设计原则总结。针对基于可视化协同辩论的反思性学习设计，本研究提出了以下四个基本设计原则：（1）通过以真实问题为基础的学习环境，促进学生自主生成辩论议题。真实问题情景通常涉及复杂性和不确定性，鼓励多

人合作，使学生在动态环境中提出和解决问题。辩论议题的选择具有主观能动性，有助于学生协同合作，并在经验和兴趣指向的领域中加强反思结果的具体性和价值（Kunsch et al.，2014）。（2）通过基于规则的辩论模式激发反思。规则涵盖个人论证的规范化表达和多人协同辩论策略。使用图尔敏的论证模型有助于构建有说服力的论点，应用限定和反驳促使学生审慎思考，并确保对结论的反复验证。在协同辩论中，计算机协同绘制论证图有助于理清辩论结构，减轻认知负荷（李见恩等人，2019；Martin，2009）。（3）通过完备的学习支架辅助任务完成。协作问题解决和辩论涉及复杂认知过程，需要充分准备的学习支持（Veerman et al., 2002），包括论证规则知识、任务说明、学习资源和协同工作平台操作说明。有效的合作和辩论需要提前约定规则和可视化协同策略，编写宏观层面上的协作活动脚本提供过程性支持（Scheuer et al.，2010）。（4）通过多维的评价方案提升学习效果。评价方案应综合考虑小组辩论质量，包括逻辑强度、结果效度和论证图复杂度等。奖励论证丰富、分歧处理有效的小组，或以匿名、角色扮演等形式消除辩论人际压力。保留学生自我评价的空间，借助指示性话语引导学生自我反思。在评价中体现个人贡献与集体成果，确保小组成员获得合理分工和均衡参与（潘琳琳和宋毅，2016）。

参考文献

史立英、甘洪倩和曹洁.(2016).反思性思维的概念、价值与培养策略. **教学与管理,(33),**12-14. https://doi.org/10.3969/j.issn.1007-4260.2002.03.030

刘玉荣和陈昕昕.(2017).论证图——促进初学者学习科学论证的有效工具. **化学教育(中英文),(17),**10-15.https://doi.org/10.13884/j.1003-3807hxjy.2017020119

李见恩、肖玲和杨汉超.(2019).图尔敏论证理论中的科学创新本质. **自然辩证法研究,(07),**41-45. https://doi.org/10.19484/j.cnki.1000-8934.2019.07.007

吴秀娟、张浩和倪厂清.(2014).基于反思的深度学习:内涵与过程. **电化教育研究,(12),**23-28+33. https://doi.org/10.13811/j.cnki.eer.2014.12.004

陈佑清.(2010).反思学习:涵义、功能与过程. **教育学术月刊,(05),**5-9. https://doi.org/10.16477/j.cnki.issn1674-2311.2010.05.005

张黎和张奇.(2007).如何通过论证图提高学生的非形式推理能力. **现代教育技术,(06),**64-66+46+28. https://doi.org/10.3969/j.issn.1009-8097.2007.06.017

赵丽红、左敏和黄先开.(2023).人工智能时代高等教育教学的变革指向：培养高阶思维. **北京师范大学学报(社会科学版),(04),**40-48. https://doi.org/10.3969/j.issn.1002-0209.2023.04.005

姚佳佳、李艳、潘金晶和程萌萌.(2022).同伴对话反馈对大学生在线深度学习的影响研究. **华东师范大学学报(教育科学版),(03),**112-126. https://doi.org/10.16382/j.cnki.1000-5560.2022.03.010

教育部. (2019). 教育部关于一流本科课程建设的实施意见.Retrieved May 10, 2021,from http://www.moe.gov.cn/srcsite/A08/s7056/201910/t20191031_406269.html.

潘琳琳,宋毅.(2016).合作学习与思辨能力的培养.**外语与外语教学,(02),**97-105+147. https://doi.org/10.13458/j.cnki.flatt.004246

Busenberg, George J. (1999). Collaborative and adversarial analysis in environmental policy. Policy Sciences, 32(1), 1–11.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. International Journal of Science Education, 22(8), 797–817.

Chen, X., Wang, L., Zhai, X., & Li, Y. (2022). Exploring the Effects of Argument Map-Supported Online Group Debate Activities on College Students' Critical Thinking. Frontiers in Psychology, 13, 856462.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2015). The effects of argument mapping-infused critical thinking instruction

on reflective judgement performance. Thinking Skills and Creativity, 16, 11–26.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. Metacognition and Learning, 7(3), 219–244.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2013). An examination of the effects of argument mapping on students' memory and comprehension performance. Thinking Skills and Creativity, 8, 11–24.

Darmawansah, D., Lin, C.-J., & Hwang, G.-J. (2022). Empowering the collective reflection-based argumentation mapping strategy to enhance students' argumentative speaking. Computers & Education, 184, 104516.

Eskin, H., & Ogan-Bekiroglu, F. (2013). Argumentation as a Strategy for Conceptual Learning of Dynamics. Research in Science Education, 43(5), 1939–1956.

Felton, M., Garcia-Mila, M., Villarroel, C., & Gilabert, S. (2015). Arguing collaboratively: Argumentative discourse types and their potential for knowledge building. British Journal of Educational Psychology, 85(3), 372–386.

Jimnez-Aleixandre, M. P., Bugallo Rodrguez, A., & Duschl, R. A. (2000). Doing the lesson or doing science: Argument in high school genetics. Science Education, 84(6), 757–792.

Kurt, M. (2018). Quality in reflective thinking: Elicitation and classification of reflective acts. Quality & Quantity, 52(S1), 247–259.

King, P. M., & Kitchener, K. S. (1994). Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults. San Francisco: Jossey Bass.

Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K., Tse, H., Webb, C., Yuet Wong, F. K., Wong, M., & Yeung, E. (2000). Development of a Questionnaire to Measure the Level of Reflective Thinking. Assessment & Evaluation in Higher Education, 25(4), 381–395.

Kunsch, D. W., Schnarr, K., & Van Tyle, R. (2014). The Use of Argument Mapping to Enhance Critical Thinking Skills in Business Education. Journal of Education for Business, 89(8), 403–410.

Muller Mirza, N., Tartas, V., Perret-Clermont, A.-N., & de Pietro, J.-F. (2007). Using graphical tools in a phased activity for enhancing dialogical skills: An example with Digalo. International Journal of Computer-Supported Collaborative Learning, 2(2), 247–272.

Martin Davies, W. (2009). Computer-assisted argument mapping: A rationale approach. Higher Education, 58(6), 799–820.

Millar, M. (1992). Effects of experience on matched and mismatched arguments and attitudes. Social Behavior and Personality, 20, 47-56.

Park, H. S., Levine, T. R., Kingsley Westerman, C. Y., Orfgen, T., & Foregger, S. (2007). The Effects of Argument Quality and Involvement Type on Attitude Formation and Attitude Change: A Test of Dual-Process and Social Judgment Predictions. Human Communication Research, 33(1), 81–102.

Rapanta, C., & Felton, M. K. (2022). Learning to Argue Through Dialogue: A Review of Instructional Approaches. Educational Psychology Review, 34(2), 477–509.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. International Journal of Computer-Supported Collaborative Learning, 5(1), 43–102.

Veerman, A., Andriessen, J., & Kanselaar, G. (2002). Collaborative argumentation in academic education. Instructional Science, 30(3), 155–186.

Weinstock, M. P. (2005). Cognitive Bases for Effective Participation in Democratic Institutions: Argument Skill and Juror Reasoning. Theory & Research in Social Education, 33(1), 73–102.

计算机支持的协作学习（CSCL）环境中的情感反馈：综述与展望

Affective feedback in Computer-Supported Collaborative Learning Environments: A Review and Perspective

朱芳芳[1,2]，杨九民[2,*]，马志强[1,*]
[1]江南大学，江苏“互联网+教育”研究基地
[2]华中师范大学，人工智能教育学部
[*] yjm@mail.ccnu.edu.cn; mzq1213@jiangnan.edu.cn

【摘要】 计算机支持的协作学习是当下重要的学习方式之一，它不仅为学习者创建了不受时间和地点桎梏的虚拟互动空间，还能够为学习者提供多样化的学习服务，有助于培养和发展学习者的沟通能力、协作能力、批判性思维、创新能力等 4C 素养。情感反馈是计算机支持的协作学习环境中的重要设计要素，有助于提高学习者的动机和参与度，进而促进学习。本文通过回顾计算机支持的协作学习环境中的情感反馈相关研究成果，梳理情感反馈的发展与研究历程，剖析情感反馈的作用机制，并提炼计算机支持的协作学习环境中情感反馈的设计与优化策略，以期为构建更具人性化和有效性的计算机支持的协作学习环境提供启发与参考，推动个性化学习的发展。
【关键词】 计算机支持的协作学习；情感反馈；个性化学习；参与度

Abstract: *Computer-supported collaborative learning (CSCL) stands out as a significant contemporary educational paradigm, offering learners the opportunity to create a virtual interactive space free from temporal and spatial constraints. Beyond this liberation, it also provides learners with various educational services, promoting the development and improvement of 4C competencies—communication, collaboration, critical thinking, and innovation. Affective feedback emerges as a pivotal factor in the design of computer-supported collaborative learning environments. It contributes to heightened learner motivation and engagement, thereby facilitating the learning process and achievements. This paper systematically reviews relevant research findings on affective feedback within computer-supported collaborative learning environments. This article elucidates the developmental trajectory and research evolution of affective feedback, conducts a nuanced analysis of its operative mechanisms, and distills design and optimization strategies specific to affective feedback in computer-supported collaborative learning environments. The primary goal is to offer inspiration and guidance for developing more humanized and effective computer-supported collaborative learning environments, thereby accelerating the progress of personalized learning.*
Keywords: computer-supported collaborative learning; affective feedback; personalized learning; engagement

1.引言

近年来，信息技术发展迅猛，给人们的生活、工作和学习都带来了巨大的变化。现如今，计算机已成为人们进行社会互动和学习活动必不可少的媒介，计算机支持的协作学习（Computer Supported Collaborative Learning, CSCL）也已经发展为教育领域中一种重要的学习方式。CSCL 是指利用计算机技术和网络工具来促进和增强学习者之间、教师与学习者之间，甚至是计算机系统与学习者之间的协作和合作学习过程，在这一过程中，系统、教师和学习者会不断地进行认知互动、情感互动和社会互动（如图 1 所示）。这一概念融合了计算机科学、教育技术和协作学习理论，旨在利用数字化工具创造出更具开放性、互动性和协同性的学习环境。在 CSCL 环境中，学习者不受时间和物理地域的限制，可以通过在线平台、

协作工具或虚拟环境进行实时的学习、信息共享、交流和合作。这种学习方式的主要优点是让每一位学习者都能够突破时空限制，跨越地域差异，它允许所有具有计算机设备和网络环境的学习者共同参与一个研究课题、学习任务或学习活动，允许来自不同文化背景、不同学术背景的教师、学生建立学术联系。此外，教师可以利用计算机技术所提供的工具，更有效地监测参与协作学习活动的学习者的任务进展，与学习者进行同步/异步互动，为学习者提供认知和情感反馈（Cress, Rosé, Wise, & Oshima, 2021）。总的来说，CSCL 不仅将学习空间从物理教室拓展到虚拟空间，还能够为学习者创建个性化的学习体验，有助于培养学习者的沟通能力（Communication skills ）、协作精神（Collaboration）、批判性思维（Critical thinking）和创新意识（Creativity and innovation），推动教育教学方式的发展与创新。

CSCL 环境下，系统、教师和学习者之间的情感互动的重要形式之一就是情感反馈。情感反馈（emotional/ affective feedback）是指在学习过程中，由学习系统、教师或同伴向学习者传递的，与其情感状态、学习表现、活动参与度等有关的信息（Fongi & Schallert, 2023）。它是 CSCL 环境中的重要组成部分，对于全面提高学习参与，促进学习者学习具有重要意义。随着学习科学的不断发展，以及对个性化学习概念的探索不断深入，研究者们逐渐也认识到，学习者的情感状态在推动个性化学习发展的过程中具有重要作用。因此，通过整合情感计算技术、社会互动理论、认知加工理论和全球化教育技术发展与应用趋势，深入了解情感反馈在 CSCL 环境中的重要性成为研究者们关注的焦点之一。

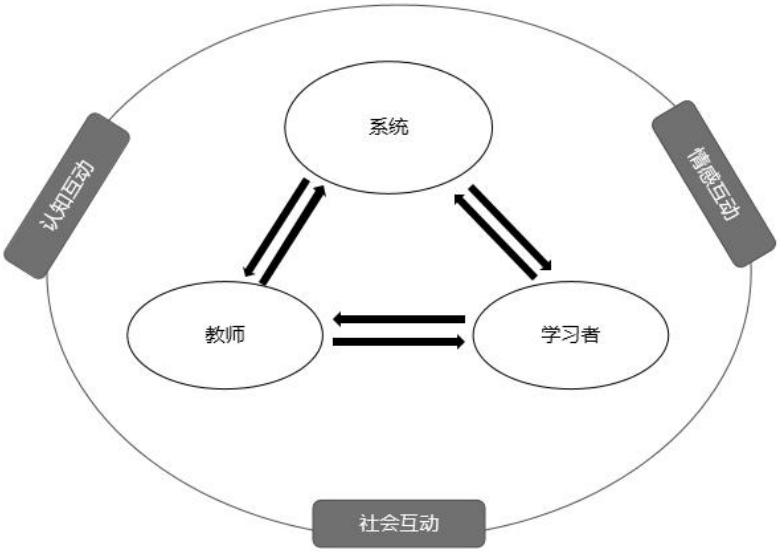


图 1 计算机支持的协作学习环境的关键要素及互动类型

基于此，本研究旨在梳理 CSCL 环境中的情感反馈相关研究成果，了解情感反馈的发展与研究历程，剖析情感反馈的作用机制，并总结 CSCL 环境中情感反馈的设计与优化策略，以期为构建更具人性化和有效性的 CSCL 环境提供启发与参考，推动个性化学习和协作学习研究的发展。

2.文献综述

2.1 情感反馈的内涵与分类

情感是个体对外界环境及环境中的事物的主观体验，它涉及心理、生理和行为方面一系列的变化。情感反馈，顾名思义，具有情感特征或情感导向的反馈。结合其应用场景来解读，情感反馈则是指在社会交流过程中或利用计算机系统学习的过程中，个体或者学习系统通过对学习者的面部表情、语音语调、肢体行为以及文本信息等多模态数据的记录，并基于机器学习、深度学习等算法对所收集的数据进行编码和分析，以了解其情感状态，之后通过相应的反馈机制，向学习者传达的情绪性信息以提供情感支持，与其进行情感互动，进而促进学

习投入和参与度，最终提高学习体验和学习效果（Coles, Larsen, & Lench, 2019; Terzis, Moridis, & Economides, 2012）。此外，个体在 CSCL 环境中对系统、学习活动以及自我表现的情感性评价，也是情感反馈的一部分。因此，在 CSCL 环境中加入情感反馈对于促进学习者学习有重要意义（Grawemeyer, Mavrikis, Holmes, et al., 2017; Molinari, Chanel, Bétrancourt, et al., 2013; Tan & Chen, 2022）。

情感反馈可以根据所传递信息的情感效价、反馈来源、反馈形式、反馈时间等方面划分为不同的类型（如表 1 所示）。

表 1. 情感反馈的类型

分类依据	分类结果
情感信息的效价 Valance	积极反馈：包含积极情感信息（如鼓励、表扬、认可、奖励、赞许等）的反馈
	中性反馈：所包含的信息不涉及明显的情感倾向，一般只包含对认知加工过程和表现的客观评价和总结
	消极反馈：包含消极情感信息（如批评、批判、纠错等）的反馈
反馈来源 Source	同伴反馈：由协作学习团队中的学习同伴所提供的情感反馈
	教师反馈：由教师或系统外的虚拟情感代理所提供的情感反馈
	系统反馈：由学习系统（包括系统中嵌入的人工智能角色）所提供的情感反馈
反馈形式 Model	静态反馈：以静态的文字、图片、表情等形式呈现的情感反馈
	语音反馈：以语音形式呈现的情感反馈
	视频反馈：以动态的视频、动画等形式呈现的情感反馈
反馈时间 Time	前馈：在学习活动开始前，用于鼓励、告知或警示学习者的情感反馈
	即时反馈：随着学习活动的开展而即时提供的情感反馈
	延迟反馈：一般是在学习活动的某个环节结束后，或整个学习活动结束后，延迟一段时间后提供的情感反馈
……	……

其中，根据所传递信息的情感效价对情感反馈进行分类是最常见的分类方式。根据罗素提出的情绪分类的环状模型（circumplex model of emotion），情感可以从效价（valance）和唤醒度（arousal）两方面来定义。从效价维度，可以将情感划分为积极情感（如开心、满意、激动、自豪、感激等）、中性情感或消极情感（如忧愁、悲伤、愤怒、紧张、焦虑、恐惧等）（Russell, 1980）。因此，根据所传递信息的情感效价，可以将情感反馈分为三种：积极反馈（positive feedback）、中性反馈（neutral feedback）和消极反馈（negative feedback）（Fongi & Schallert, 2023; Tan & Chen, 2022; Terzis, Moridis, & Economides, 2012）。其中，积极反馈通常会传达认可、鼓励、赞美、支持、奖励等积极信息，对学习者的学习表现进行肯定，可以唤醒他们的学习热情和积极性，促进参与（Cho, Schunn, & Charney, 2006; Grawemeyer, Mavrikis, Holmes, et al., 2017; Tan & Chen, 2022）；中性反馈一般不涉及明显的情感倾向，是对系统、学习任务和学习表现的客观态度和评价；消极反馈是相对于积极反馈而言的，一般

是对学习者进行批评、纠正、批判、惩罚，有时也会伴随意见和建议，这种反馈可以帮助学习者纠正错误的理解或调整不合适的学习方法,但是也可能会对学习产生消极影响（Kuklick & Lindner, 2023; Quintelier, Vanhoof, & De Maeyer, 2018; van Schie, Chiu, Serge et al., 2018）。

2.2 CSCL 环境中的情感反馈

CSCL 环境下的学习过程是一个由教师、学习者、学习材料和学习系统共同构成的认知与情感互动的过程。在这个过程中，学习者发出的情感反馈，以及收到的来自系统、教师和同伴的情感反馈，都可能会影响他们的情感状态、参与程度、认知加工过程和结果。这些影响，会因情感反馈的类型不同而不同。

通过对 CSCL 环境中的情感反馈相关研究的梳理，本文发现，积极反馈通常是对 CSCL 有益的（Arguedas, Daradoumis, & Xhafa, 2016; Cheng & Hou, 2015; Chien, Hwang, & Jong, 2020; Duijnhouwer, 2010; Lu & Law, 2012; Selcuk, Jones, & Vonkova, 2021）。例如，Arguedas, Daradoumis, & Xhafa（2016）研究发现，教师和一群高中生在一个 CSCL 环境中共同完成一项学习任务的过程中，教师基于学习者发表的文字性信息来了解他们的情感状态，并及时地在小组协作过程中提供积极的情感反馈。结果显示，教师所提供的积极情感反馈不仅显著促进了学习者的行为参与和认知参与，而且显著提高了他们在 CSCL 环境中的任务表现。之后，Chien, Hwang, & Jong (2020)设计开发了一个虚拟现实环境，在这个虚拟在线环境中学习者可以协同完成英语练习任务，并且互相评价，即给予同伴反馈。结果发现，来自同伴的积极情感反馈（即表扬）显著提高了学习者的英语口语表现，而消极情感反馈（即批评）则不利于学习者的英语口语表现。

然而，也有研究发现，相较于中性情感反馈，积极情感反馈可能无法显著促进学习者在 CSCL 环境中的学习效果（Chen & Tan, 2022; Kwon, Song, Sari, & Khikmatillaeva, 2019）。例如，Chen & Tan（2022）研究发现，由同伴所提供的积极情感反馈可能无法有效促进学习者的行为改善，相反，对学习者的任务表现进行客观总结和评价的中性情感反馈更有助于学习者发现问题，改善学习效果。这或许是因为，这种积极反馈大多是对学习者的鼓励和肯定，而没有直接有效地指出学习者在完成任务过程中的整体表现和所存在的问题，对于他们后续学习行为的完善和指导意义不强（De Sixte, Maná, Avila, & Sánchez, 2020）。

关于消极情感反馈，大多数研究聚焦于纠正性反馈（corrective feedback）对学习的影响，即向学习者提供有关学习活动和任务表现的批评性和纠正性信息，以帮助其改正错误，促进学习（韩晔 & 许悦婷，2020；Chen, Liu, & Lin, 2023）。例如，Kim (2023)研究发现，相较于无意义的鼓励，为学习者提供纠正性反馈更有其外语学习。然而，Guasch, Espasa, & Martinez-Melo（2019）却发现，相较于中性的总结性反馈，消极情感反馈无法有效促进学习者的认知参与，也没有提高其学业表现。

总的来说，近年来研究者们探索了不同类型的情感反馈对 CSCL 环境下学习者的学习参与和学习效果的影响。积极情感反馈能够鼓励学习者，满足其情感需求，从而促进学习投入，提高学习效果；消极情感反馈一方面能够帮助学习者发现错误，有助于调整认知加工的方法和策略，另一方面也会导致学习者产生被试、抑郁、焦虑等消极情绪，不利于学习。但是，由于研究设计和研究主题的差异，以往研究尚未达成一致性结论。因此，未来需要开展更多的实证研究来检验积极和消极情感反馈在 CSCL 中的作用。

3.CSCL 环境中情感反馈影响学习的作用机制

基于对以往 CSCL 环境中情感反馈对学习的影响的研究的梳理，可以发现，在 CSCL 环境下，情感反馈影响学习的作用机制涉及多个方面。如图 2 所示，首先，根据社会互动理论，情感互动能够满足学习者的情感互动需求，尤其是积极情感反馈，不仅能够加强学习者与反馈提供者（如教师、同伴）之间的社会联系，为学习者创造积极的学习氛围，从而激发学习者的学习兴趣和学习参与度，提高学习者在 CSCL 环境下的学习体验（Cheng & Hou, 2015; Lu & Law, 2012）。其次，从积极心理学视角来看，积极情感反馈通过提供鼓励和认可，增强了学

习者的自信心，激发了他们的内在动机，促使他们更积极地参与协作学习活动，提高学习效果（韩晔 & 许悦婷，2020）。消极反馈虽然含有批评等消极信息，但是能够提醒学习者进行元认知监控，进而学习者调整自己的认知加工策略。此外，根据认知加工理论，积极情感反馈和纠正性反馈都有助于学习者全面地了解和监测自己在 CSCL 中的角色以及在活动中的表现，能够为他们的学习后的提供一定的支持，从而达到促进学习的目的。

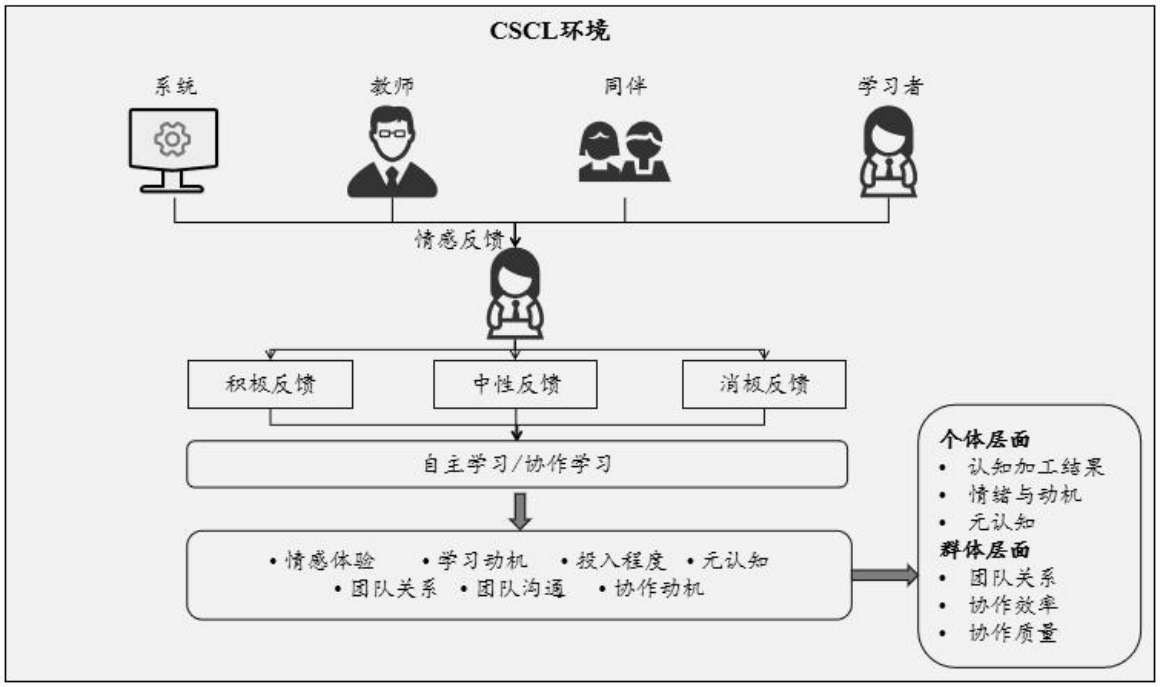


图 2. 计算机支持的协作学习环境下情感反馈影响学习的作用机制

由协作团队中同伴提供的积极情感反馈还可以增进团队成员之间的关系，提升团队协作氛围。通过及时的鼓励和认可，团队成员能够感受到对其贡献的肯定，建立了良好的协作基础。其次，积极情感反馈也可以视为一种积极的内部沟通，能够避免团队内部的误解和冲突。由积极情感反馈所诱发的积极情感体验会让其他成员更愿意分享观点、接受反馈，有助于在团队内部形成更为开放和有效的沟通模式。最终，积极情感反馈通过激发成员的协作动机，推动团队协作进程，提高团队协作质量。纠正性反馈则通过高效指出协作学习中出现的关键问题来提高协作效率，进而促进学习。

4.CSCL 环境下的情感反馈设计与优化策略

基于对 CSCL 环境中情感反馈的梳理，以及 CSCL 环境中学习活动的特点，本文为情感反馈策略的设计与优化提出了几点策略，主要包括：

（1）个性化定制与差异化反馈

在 CSCL 环境中，设计和优化情感反馈的首要原则是个性化定制与差异化反馈。学习者在学术背景、兴趣、学业能力、认知风格等方面存在差异，因此，情感反馈应当根据个体差异进行定制，以满足学习者的个性化发展需求（Chen & Wang, 2021）。差异化反馈不仅能够更有效地引导学习者的学习方向，还能够提高学习动机和积极性。个性化的反馈可以通过使用先进的技术工具和算法，如机器学习、大模型和情感计算，实现对学习者个体差异的精准感知，从而提供更为精准的反馈。

（2）积极性与建设性反馈

设计情感反馈时，强调积极情感并提供建设性反馈是至关重要的原则。积极的情感反馈，如赞美、认可、鼓励、表扬等，能够增强学习者的自信心和学习兴趣，促使他们更积极地参

与学术活动（韩晔 & 许悦婷，2020; Fongi & Schallert, 2023）。建设性的反馈（如纠正性反馈，corrective feedback）能够为学习者提供有建设性的建议和指导，注重具体性，结合具体的学习内容和活动表现来指导学习者如何改进，而非仅仅停留在表面性的表扬和认可。这种积极和建设性的情感反馈有助于创造支持性的学习氛围，促进协作团队的交流互动，进而提高学习者的学业表现和团队协作质量，培养学习者的沟通精神、协作精神、批判性思维和创新精神。

（3）及时性与即时反馈

及时性是设计和优化情感反馈的关键原则之一。及时提供反馈有助于学习者更好地理解和应对学术任务，促进他们的学习效果。在协作学习环境中，即时反馈不仅涉及学习者的个人表现，还包括团队协作的方方面面。及时性的反馈有助于纠正错误、鼓励正确行为，启发思考，并维持学习者参与协作学习活动的积极性。通过使用实时互动工具和即时通讯平台和软件，可以有效实现及时性反馈，提高学习效果。

（4）多模态反馈形式的整合

以往研究中，情感反馈多以静态文字、图片和表情为主。为了更全面地传达情感信息和提供更具体的指导，设计和优化情感反馈需要整合多模态反馈形式。多模态反馈形式包括文字、表情、声音、图像、视频等多种形式，可以更充分地表达情感内容（Chen & Tan, 2022; Fongi & Schallert, 2023）。例如，在语言表达不足以传达情感时，通过语音、图像或视频反馈来弥补。整合多模态反馈形式可以提高情感反馈的传达效果，使学习者更容易理解和接受信息。

（5）提供明确的方向性反馈

为了引导学习者朝着明确的学习目标前进，情感反馈应当结合具体学习内容和目标提供明确的方向性反馈。清晰的学习目标有助于学习者更好地理解他们的协作学习任务，并使得反馈更加有针对性。方向性反馈能够指导学习者朝着正确的方向发展，还能够启发其进行批判性思考，避免无意义的学习行为。因此，通过将情感反馈与学习目标直接关联，可以增强学习者对于学习任务的认知，提高学习效果，提升协作学习效率。

5.结论与展望

技术进步促使教育教学不断变革，人们从传统学习环境逐渐拓展到在线协作学习环境，以计算机为媒介的学习和交流成为当下重要的学习方式之一。CSCL 环境下学习者的自主性和个性化开始凸显，教育者也探究多种策略和方法来培养和提升学习者的自主学习能力和团队协作精神。在此过程中，学习者的情感需求也不容忽视。情感反馈是在 CSCL 环境中满足学习者的情感互动需求，促进学习参与的重要手段。研究发现，积极情感反馈和部分消极反馈（即纠正性反馈）有助于唤醒学习者的学习动机，促进认知投入，提高任务表现。然而，目前相关研究仍存在一定局限，需要更多实证研究来检验和探索不同情感反馈对学习的影响及其作用机制。

随着各种智能化技术与教育的深度融合，CSCL 环境也愈加智能化，因此，未来研究者们可以考虑在虚拟协作学习环境中不断开发和完善情感反馈系统的建设，以提升学习体验，提高学习质量。未来的研究可以从如下几方面展开：

（1）情感智能辅助工具的开发。设计具有情感智能的虚拟助手或工具，能够及时地识别和响应学习者的情感状态。这样的工具可以提供个性化的学习支持，根据学习者的情感需求来调整情感反馈和认知支持；

（2）虚拟角色和情感互动设计。引入能够模仿真人面部表情、语气、手势以及动作的虚拟角色或代理（Emotional Pedagogical Agent, EPA），使其能够以情感丰富的方式与学习者进行语音和虚拟面对面互动。这有助于模拟真实世界中的社会互动，提高学习者在学习过程中的情感投入和参与度；

（3） 情感分析和反馈系统的整合。结合情感分析技术，如面部表情识别技术（FACS、FaceReader)、基于及其学习和深度学习的文本情感分析技术（sentiment analysis），实时监测

学习者的情感状态，并将这些信息整合到反馈系统中。这有助于更准确地理解学习者的情感需求，提供更为精细化的情感支持；

（4）基于游戏化的情感反馈：利用游戏化元素，设计能够激发积极情感体验的学习环境。通过奖励机制（如徽章、金币、积分）、竞争性元素（如排行榜、梯度）或情感丰富的虚拟场景，提高学习者的学习动机和参与度；

（5）虚拟团队协作中的情感支持（emotional support）：着重于团队协作环境中的情感支持，包括团队建设、沟通技能培养、冲突解决等。情感反馈可以促进团队成员之间的有效合作，改善团队氛围，提高协作效果。

这些不仅有助于提高学习者在虚拟协作学习环境中的学习体验，促进学习效果，还有助于推动协作学习系统的建设与优化。通过整合情感反馈技术，不仅可以更好地满足学习者个性化的情感需求，还有助于培养学习者的协作学习精神和创新精神，推动教育数字化转型大背景下教与学方式的创新与变革。

致谢

本文系新文科研究与改革实践项目“‘人工智能+教育’背景下卓越中学教师培养的教学模式创新与实践——以现代教育技术应用》为例”（项目编号：2021180023）和国家自然科学基金面上项目“生成性学习策略影响视频学习的认知神经机制及智能干预研究”（项目编号：62177027）研究成果。

参考文献

韩晔 & 许悦婷. (2020). 积极心理学视角下二语写作学习的情绪体验及情绪调节策略研究——以书面纠正性反馈为例. *外语界*(01), 50–59.

Arguedas, M., Daradoumis, T., & Xhafa, F. (2016). Analyzing the effects of emotion management on time and self-management in computer-based learning. *Computers in Human Behavior, 63,* 517–529.

Cress, U., Rosé, C., Wise, A. F., & Oshima, J. (2021). *International Handbook of Computer-Supported Collaborative Learning.* Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-65291-3

Cheng, K. H., & Hou, H. T. (2015). Exploring students' behavioral patterns during online peer assessment from the affective, cognitive, and metacognitive perspectives: A progressive sequential analysis. *Technology, Pedagogy and Education, 24*(2), 171–188.

Chen, S. Y., & Wang, J. H. (2021). Individual differences and personalized learning: A review and appraisal. *Universal Access in the Information Society, 20,* 833–849.

Chen, W., Liu, D., & Lin, C. (2023). Collaborative peer feedback in L2 writing: Affective, behavioral, cognitive, and social engagement. *Frontiers in Psychology, 14*, 1078141.

Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-speaking performance and learning perceptions. *Computers & Education, 146,* 103751.

Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication, 23*(3), 260–294.

Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin, 145*(6), 610–651.

De Sixte, R., Maná, A., Avila, V., & Sánchez, E. (2020). Warm elaborated feedback. Exploring its benefits on post-feedback behavior. *Educational Psychology, 40*(9), 1094–1112.

Duijnhouwer, H. (2010). *Feedback effects on students' writing motivation, process, and performance.* Utrecht, Netherlands: Utrecht University.

Fongi, C. J., & Schallert, D. L. (2023). "Feedback to the future": Advancing motivational and emotional perspectives in feedback research. *Educational Psychologist, 58*(3), 146–161.

Guasch, T., Espasa, A., & Martinez-Melo, M. (2019). The art of questioning in online learning environments: the potentialities of feedback in writing. *Assessment & Evaluation in Higher Education, 44*(1), 111–123.

Lu, J., & Law, N. (2012). Online peer assessment: Effects of cognitive and affective feedback. *Instructional Science, 40*(2), 257–275.

Molinari, G., Chanel, G., Bétrancourt, M., Pun, T., & Christelle, B. G. (2013). Emotion Feedback During Computer-mediated Collaboration: Effects on Self-Reported Emotions and Perceived Interaction. In Rummel, N., Kapur, M., Nathan, M., & Puntambekar, S. (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013* Conference Proceedings Volume *1* — Full Papers & Symposia (pp. 336–343). Madison, WI: International Society of the Learning Sciences.

Kim, S. (2023). A comparative analysis of language choices in written peer feedback provided by high- and low-proficiency Japanese EFL students. *Language Teaching Research,* published online. https://doi.org/10.1177/13621688231193074

Kuklick, L., & Lindner, M. A. (2023). Affective-motivational effects of performance feedback in computer-based assessment: Does error message complexity matter?. *Contemporary Educational Psychology, 73,* 102146

Kwon, K., Song, D., Sari, A. R., & Khikmatillaeva, U. (2019). Different types of collaborative problem-solving processes in an online environment: Solution oriented versus problem oriented. *Journal of Educational Computing Research, 56*(8), 1277–1295.

Quintelier, A., Vanhoof, J. & De Maeyer, S. (2018). Understanding the influence of teachers' cognitive and affective responses upon school inspection feedback acceptance. *Educational Assessment, Evaluation and Accountability, 30,* 399–431.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.

Selcuk, H., Jones, J., & Vonkova, H. (2021). The emergence and influence of group leaders in web-based collaborative writing: Self-reported accounts of EFL learners. *Computer Assisted Language Learning, 34*(8), 1040–1060.

Tan, S. H., & Chen, W. (2022). Peer feedback to support collaborative knowledge improvement: What kind of feedback feed-forward?. *Computers & Education, 187,* 104467.

Terzis, V., Moridis, C. N., & Economides, A. A. (2012). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education, 59*(2), 710–721.

van Schie, C. C., Chiu, C. D., Rombouts, S., Heiser, W. J., & Elzinga, B. M. (2018). When compliments do not hit but critiques do: An fMRI study into self-esteem and self-knowledge in processing social feedback. *Social Cognitive and Affective Neuroscience, 3*(4), 404–417.

基于知识论坛探究话语的商务英语专业学生词汇与思维发展研究

雷春林[1]*，王玉[1]，张俊杰[2]
[1]上海对外经贸大学国际商务外语学院
[2]香港大学教育学院
*leichunlin@suibe.edu.cn

【摘要】本文在中国大陆语境下，研究商务英语学习者如何利用知识论坛平台，通过持续合作探究，促进产出性英语词汇的增长和思维水平的提升。参与者是某商科高校 30 名大一学生，依托混合式《商学导论》课程进行为期 16 周的教学实验。对论坛话语的分析表明，随着时间推移，学生使用词汇的复杂度、学术性和商务专业性不断增加。利用话语探究链分析表明，反映学生思维水平的集体知识也随着时间不断进步，且词汇增长和知识进步之间存在交织互惠的关系。该研究对如何设计学习环境以促进词汇学习和思维进步具有借鉴意义。
【关键词】 知识建构；词汇分析；知识论坛；知识提升；探究链

Abstract: *This study was conducted in Chinese mainland, examining Business English Majors' productive vocabulary growth and collective knowledge advancement through collaborative inquiry mediated by Knowledge Forum. Participants were 30 Year-1 university students, enrolled in a blended "Introductory Business Course" for 16 weeks. Analysis of the Forum discourse indicated students used greater number of sophisticated, academic and business domain words across time. Inquiry threads analysis also suggested students' thinking as reflected by collective knowledge advancement also moved from a low to a high level. There seemed an intertwined, reciprocal process of vocabulary growth and knowledge advancement. This study has implications on how to design learning environments to facilitate vocabulary learning and idea improvement.*

Keywords: knowledge building, lexical analysis, Knowledge Forum, knowledge advancement, inquiry thread

1. 研究背景

知识建构通过学习者的协作探究，实现观点的不断完善。这一教育模式依托知识论坛平台（knowledge Forum）和一系列知识建构的原则，使得学生的思想观点得以外化、质疑、补充、解释和凝练，这种有意义的协商，塑造和促成了学生的知识建构话语（Scardamalia & Bereiter，2021）。知识建构的研究近年来层出不穷，其裨益也被广泛证实（Tan et al., 2021）。例如，Sun、Zhang & Scardamalia（2010）曾在以英语为母语的国家（环境）中，研究了三年级小学生在知识建构过程中词汇的发展情况。但是，总体而言，学界对于知识建构过程中词汇发展的研究较少，特别是以英语为授课语言同时英语又是外语的情境（EFL）中，知识建构的模式如何影响学生的外语词汇发展以及思维水平，更是鲜有报告。

本研究在某商科大学商务英语专业一年级学生中，采用知识建构模式，进行为期 16 周的混合式《商学导论》课程的教学实验。学生在线下的课堂里以及线上的知识论坛平台上，使用英语进行交流，提出探究问题并寻找解决方案，在论坛上撰写笔记（Notes）记录这一过程。通过对论坛上探究话语的分析，回答两个问题（1）知识建构过程中，学生的英语词汇如何发展变化？（2）知识建构过程中，学生的探究思维水平如何发展变化？

2. 研究设计

一个自然班级的 30 名大一学生参与此项研究。课程中教师注重建立班级合作学习的文化，并强调了自主认知、共同担责、观点不断改进等一系列知识建构的原则。学生在课堂进行各种商务主题概念的讨论，在课后用知识论坛延续讨论，并撰写论坛笔记（见图 1）。教师鼓励学生尽量使用完整正确的语句和所学的商务词汇进行表达，并引导他们提问，解释，质疑、修正以及提出新问题，在相互协商、共同构建的过程中，推动集体知识的进步。

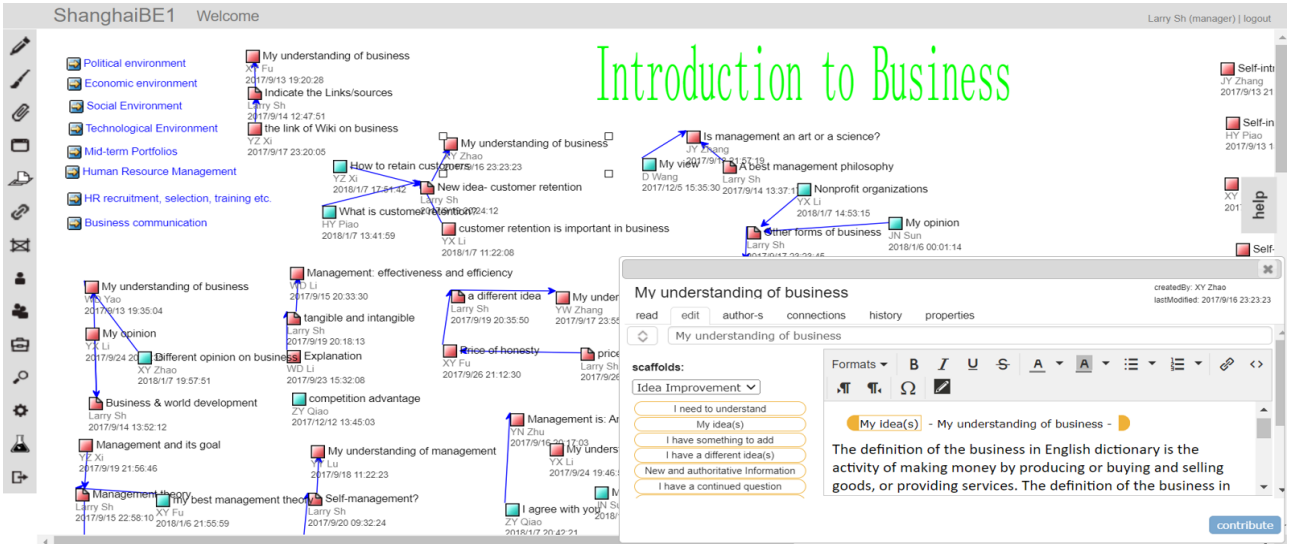


图 1 知识论坛协作探究的界面以及论坛笔记

本研究的数据包括学生在 16 周内撰写的所有论坛笔记/话语，一共 746 条。为了揭示词汇和思维水平的发展情况，数据被分成第一阶段（1-8 周）和第二阶段（9-16 周），分别包括 384 和 362 条笔记。分析之前对数据进行清洁和整理，撰写的词汇均为产出性词汇，词汇分析依托类符/形符比率（TTR）以及词汇频率配置的方法进行（Coxhead, 1998；Nation, 2001），该方法可以分析词汇的复杂度，学术性词汇使用，以及某一特定领域词汇（例如本文的商务词汇）的使用情况。思维水平的发展情况依托探究话语链进行，即根据论坛探究主题，将论坛笔记归类整理成探究链（inquiry thread，Zhang et al., 2007），探究链记录了一个主题从开始到结束的进程。学生作为一个学习共同体，在探究链里反射出一系列认知行为，比如提出问题、提供解释、提出质疑、完善或者修正观点、以及总结和凝练观点等等。这个集体认知过程通过内容分析，也可以编码为集体知识进步的高、中、低三个层次（Lei & Chan, 2018），反映思维的发展变化情况，如表 1 所示。

表 1 低、中、高三个层次的探究话语链编码方案

层次	对探究链的描述
低级	探究链通常较短，学生的讨论主要是对商务概念或主题的快速的、本能性想法，几乎没有参考权威信息或引用数据
中级	探究链呈现学生各种观点和思考，但观点之间频繁重复，讨论还是为了完成任务，学生缺乏群体意识
高级	探究链展示了一个“问题--解释“相互交织的过程，讨论中学生能识别知识欠缺、能意义协商，并对商务概念提出更深入的问题或观点

3. 结果与分析

3.1. 学生产出性词汇的发展

通过 RANGE32 和 SPSS 分析软件, 表 2 报告了两个阶段学生论坛笔记中词汇的形符和类符的平均值和标准差。T 检验后表明，随着时间的推移，学生使用的形符和类符比例显著增加。在第二阶段，学生撰写笔记的数量不如第一阶段，但他们倾向于撰写更长的笔记。

表 2 第一阶段和第二阶段学生笔记里词汇的形符与类符比较

	第一阶段 平均数（标准差）	第二阶段 平均数（标准差）	T-检验
全部形符	626.13 (244.30)	853.73 (278.35)	T= 4.95; p<0.001
全部类符	246.16 (80.03)	331.90 (90.70)	T= 6.47; p<0.001

表 3 报告了学生在两个阶段论坛笔记里的词汇频率配置, 显示了第一和第二个 1000 个词族、学术词汇、不在列表中的词汇、以及特定商务词汇类型的平均百分比。成对样本 T 检验显示，第一个 1000 词的比例在第二阶段显著下降，而第二个 1000 词、学术词汇、不在列表中的词汇、以及商务词汇的使用在第二个阶段均显著增加。这表明，在知识建构的过程中，学生越来越倾向于使用更多样化和更复杂的英语词汇，使用更多学术词汇以及特定的商务专业词汇。

表 3 基于词类数的词汇频度配置在第一阶段和第二阶段的比较

	第一阶段 平均数（标准差）	第二阶段 平均数（标准差）	T-检验
第一个 1000 词	73.69% (4.93%)	66.44% (4.69%)	T=7.75; p<0.001
第二个 1000 词	7.14% (2.00%)	8.24% (1.93%)	T=2.05; p<0.05
学术词汇	11.48% (3.65%)	14.10% (3.38%)	T=3.34; p<0.01
不在表中词汇	7.70% (3.83%)	11.21% (3.88%)	T=5.23: p<0.001
特定商务词汇	8.59%（1.55%）	10.44%（2.00%）	T=4.53: p<0.001

3.2. 学生思维水平的发展和集体知识的进步

根据学生在知识论坛上撰写的笔记，将相同主题的笔记加以识别并整理成了探究链。经过分类，一共有 29 个探究链从数据中浮现出来，主要围绕商务主题，例如“什么是商务”、“商务模式有哪些”、“创业需要哪些资源”、“如何增加内需”、“如何弥合贫富差距”等等。又根据讨论的时间轴进一步将这些探究链分类为两个阶段，在 1-8 周内结束的第一阶段探究链共有 16 条，在第 8 周之后结束的第二阶段探究链一共有 13 条。

探究链反映了一个思想观点或商务概念的发展轨迹。这一个观点和概念通过学习社区的共同努力，被共享、评估、丰富、质疑、修订或完善，这个认知思维过程就是集体知识进步的过程。根据表 1 中的编码方案对 29 个探究链进行分析，表 4 报告了第一阶段和第二阶段探究链在集体知识进步上体现的低、中、高级别的结果。在第一阶段，超过一半的探究链被编码为低层次知识进步探究链，这表明学生在很多情况下，讨论较为肤浅，依赖快速答案和信息分享，缺乏对思想观点的深度和连贯性处理。在第二阶段，学生更多意识到学习社区中的集体担责，更多地围绕“提问-解释-再提问-再解释”的过程展开讨论，能看到知识的欠缺并且

主动弥补并提出新的问题。在第二阶段，低层次知识进步的探究链百分比下降明显，而体现高层次知识进步探究链的百分比明显上升。

表 4 第一阶段和第二阶段包含低、中、高知识提升的探究链的分布情况

	第一阶段 数量（百分比）	第二阶段 数量（百分比）
低级别知识进步	9 (56.35%)	3 (23.1%)
中级别知识进步	4 (25.00%)	3 (23.1%)
高级别知识进步	3 (18.75%)	7 (53.8%)

4. 启示

这项知识建构的研究是在中国大陆英语作为外语（EFL）环境中进行的，主要考察了学生产出性词汇的发展和思维水平的进步。结果表明，随着时间的推移，学生能够使用更多复杂、学术和商务专业的词汇。同时，通过反映学生认知思维水平的探究链分析，发现学生的集体知识进步也经历了一个从较低水平向较高水平发展的轨迹，知识建构促进了学生词汇的发展和思维水平的提升，而且，词汇的增长和知识提升存在相互交织、相互促进的互惠过程。

Sun et al（2010 年）的研究表明，知识建构为学生提供了一个真实的情境，在这个情境中，学生会遇到、理解并反复使用学术和特定专业领域的词汇，以达到推进和完善思想观点的目标，本研究在高等教育的场景下进一步证明了这一结论。这项研究对如何设计学习环境，促进语言和内容整合式学习以及思维水平的提升，具有很强的实践意义。

注：本文是上海教育科学项目“知识建构模式促进商务英语专业学生的跨学科知识与思辨能力的实证研究”，编号 C2022316 的部分研究成果。

参考文献

Coxhead, A. (1998). *An Academic Word List*. Occasional Publication Number 18, LALS, Victoria University of Wellington, New Zealand

Lei, C., & Chan, C. K. (2018). Developing metadiscourse through reflective assessment in knowledge building environments. *Computers & Education*, 126, 153-169.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.

Scardamalia, M., & Bereiter, C. (2021). Knowledge Building: Advancing the State of Community Knowledge. In: Cress, U., Rosé, C., Wise, A.F., Oshima, J. (eds) *International Handbook of Computer-Supported Collaborative Learning*. Cham: Springer.

Sun, Y., Zhang, J., & Scardamalia, M. (2010). Knowledge building and vocabulary growth over two years, Grades 3 and 4. *Instructional Science*, 38(2), 147-171.

Tan, S. C., Chan, C., Bielaczyc, K., Ma, L., & Bereiter, C. (2021). Knowledge building: aligning education with needs for knowledge creation in the digital age. *Educational Technology Research and Development*, 69, 2243-2266.

Zhang, J. W., M. Scardamalia, et al. (2007). Socio-cognitive dynamics of knowledge building in the work of 9 and 10-year-olds. *Education Tech Research Dev*, 55, 117-145.

数智技术赋能 K-12 教育的研究综述

A Review of Research on Digital Intelligence Technology Enabling K-12 Education

胡佩莹[1*]，潘梦媛[2]，林肖颖[3]
[123]杭州师范大学经亨颐教育学院
*1066280374@qq.com

【摘要】 随着时代进步和技术变革，云计算、5G 通信技术、区块链、大数据等数字技术进入了一个迅速发展的黄金时期，这些技术被称为“数智技术”。在中国教育现代化背景下，数智技术的应用引发了许多学者研究，但人们对数智技术在教育中应用的知识和实践现状知之甚少。本文通过系统性文献综述对基于数智技术的教育应用的研究进行系统回顾。它侧重于三个主题：（1）为了教育目的，数智技术开发了哪些应用程序？（2）数智技术能给教育带来什么好处？（3）在教育中数智技术有哪些挑战？对每个主题进行了详细的结果分析，并对调查结果进行了深入讨论。

【关键词】 数智技术；K-12 教育；系统性文献综述

Abstract: *Digital technologies such as cloud computing, 5G communication technology, blockchain, and big data have entered a golden age of rapid development, and these technologies are known as "digital intelligence technologies". In the context of China's education modernisation, the application of digital technologies has triggered many studies by scholars, but little is known about the current state of knowledge and practice in the application of digital technologies in education. This paper provides a systematic review of research on educational applications based on digital intelligence technologies. It focuses on three themes:(1)What applications have been developed in digital intelligence technologies for educational purposes? (2)What benefits can digital intelligence technologies bring to education? (3)What are the challenges of digital intelligence in education? Detailed results are analysed for each theme and the findings are discussed in depth.*

Keywords: digital intelligence technologies, K-12 education, systematic literature review

1. 研究背景

以大数据、人工智能、移动互联等为代表的“数智”技术快速发展，推动了各类新产业、新业态、新平台蓬勃发展(金波 & 杨鹏, 2022)。相关部门强调中小学教师应充分利用现代信息技术特别是人工智能技术，提升教育管理工作效率，可见数智技术赋能 K-12 教育具有举足轻重的地位。近年来，关于数智技术在教育领域的应用的文献数量不断增加，但仍然比较零散。因此，为了揭示数智技术在教育层面的研究最新进展，本文对国内外数智技术在 K-12 领域的实证研究文献进行梳理与分析，以期为教育文献做出原创及时的贡献。

2. 研究设计

本研究选取 Web of Science、Science direct、Springer、ERIC 及 CNKI 五个主要的文献数据库，以字符串“数智技术与教育”、“数智技术与学习”和“数智技术与教学”以及(“cloud computing”OR“artificial intelligence”OR“digital technology”OR“big data”)AND (teach* OR education OR course OR learn*) and ("K-12" OR teen* OR school* OR “primary school”)进行检索，得到 4546 篇文献。将研究情境限定为“教育”，获得 873 篇文献，将学段设定为

“K-12 领域（幼儿园、小学、初中、高中）”获得 133 篇，其中 60 篇是针对于将数智技术作为一门课程的研究，例如人工智能课程、机器人课程等；26 篇是综述类、会议录论文。剩下 47 篇进行全文阅读，发现有 16 篇文章没有提供完整的审查数据，因此应该删除。最终，31 篇文章被认为适合收录。

3. 分析与讨论

全世界对在教育中使用数智技术越来越感兴趣。然而，鉴于已发现的研究数量相对较少，该领域需要更多的研究。总的来说，对 31 项研究的系统回顾发现现研究倾向于三个研究问题。

3.1 出于教育目的使用数智技术开发了哪些应用程序？

数智技术在 K-12 教育领域的应用主要体现在辅助教学、教育质量评估和教育资源共享三个方面。这些应用利用先进的算法和人工智能技术，旨在提高教育的质量和效率。

第一类侧重于辅助教学系统。这一类别的应用程序通过提供个性化的学习体验来促进教学和学习。例如，智能辅导系统和论文自动评分系统利用算法和学习数据分析来支持教学实践和提高学生的学习效率(Lechuga & Doroudi, 2022)。此外，对话聊天机器人的使用有助于知识整合，尤其对学习程度较低的学生有所帮助(Xia 等, 2022)。此类别涉及辅助教学过程、评价。第二类是教育质量评估系统。这些系统通过构建科学的评价指标体系，如小学教育 APP 和 STEM 课程的评价指标体系(X. Huang, 2022b)，以及采用熵权法和灰色聚类分析法对在线教学质量进行评价(Li & Su, 2020)，确保教育的质量和效率。这些评估方法有助于识别和提出提高教育质量的策略。第三类是教育资源共享应用。云计算技术在此类别中发挥重要作用，通过设计电子学习课程(Murphy, 2019)和提供易于获取的学习资源(Wang & Wong, 2019)，降低了教育成本并提高了资源的可及性。这些应用通过优化教育资源的分配和共享，使得教育质量和可承受性得到提升。

数智技术的应用在 K-12 教育中展现出其潜力，通过辅助教学、评估和资源共享，为提高教育质量和效率提供了有效的解决方案。随着技术的不断进步，我们可以期待未来教育领域将出现更多创新和高效的教学方法。

3.2 数智技术赋能 K-12 的教育有什么优势？

数智技术在 K-12 教育中的应用带来了显著的好处，这些好处主要体现在提升学生学习、促进教师教学和提高学校管理层的工作效率三个方面。

（1）提升学生学习：数智技术通过对话聊天机器人进行知识整合，提高低层次水平学生的自我调节学习能力(Xia 等, 2022)。个性化教育（PE）(Qushem 等, 2021)和人工智能增强的数字学习环境(Lashari 等, 2022)为学生提供个性化指导和支持。这些技术有助于培养学习者的元认知能力，进行个性化学习，监控学习过程，对 21 世纪技能的养成具有重要意义。

（2）促进教师教学：人工智能技术支持教师访问、改编和使用多语言内容的工具，提高教学效率(Chounta 等, 2022)。教师可以利用人工智能支持的课堂提供实时分析数据，帮助学生更好地学习(Holstein & Aleven, 2022)。机器学习和可解释的人工智能（XAI）技术预测学生表现(Jang 等, 2022)，帮助教师提前调整教学策略，提升学生的学习效果。

（3）提高学校管理层的工作效率：数智技术通过提供更好的数据分析和决策支持软件，提高学校管理层的工作效率(Kuleto 等, 2022)。此外，数智技术与教师的合作有助于缩小数字鸿沟，促进 K-12 学习空间的建设(Pelaez 等, 2022)。远程教育环境也支持残疾或非残疾学习(Starks, 2022)，使学校能够更高效地处理重复工作，更系统地存储学生学习表现数据。

3.3 数智技术赋能 K-12 教育有什么挑战？

数智技术在 K-12 教育中的应用虽然具有巨大潜力，但在实际应用过程中也面临着一系列挑战。这些挑战主要包括教师能力的提升、教学法的选择与应用、年级和学科的适应性，以及对技术应用的广泛性和潜在影响的考量。

（1）教师能力的影响：随着数智技术在教育中的广泛应用，教师需要提升自己的技术应用能力，以适应新的教学环境。这包括对人工智能(AI)和相关技能的理解，以及如何将这些技术有效地融入到教学实践中(Chiu 等, 2023)。教师的专业实践中增添了复杂性，他们需要在提供学习支持的同时，处理由技术引入的道德难题，如数据隐私和学生监控等问题。

（2）教学法的选择与应用：教师在采用数智技术时，需要选择合适的教学法，以促进学生的学习。这可能涉及到利用增强现实(AR)等技术来促进学习，或者灵活运用多种策略来提高教学效果。同时，教师需要意识到数智技术提供的实时交互情境学习方式，这对于学习者的练习和学习效果是有益的。

（3）年级和学科的适应性：数智技术在不同年级和学科中的应用需要特别考虑。目前，研究主要集中在城市学生和基础科学概念的教学中，而对乡村教育和非 STEM 学科如文学、艺术等领域的研究较少(贾桐, 2023)。这提示未来研究需要更多地关注这些领域，以确保数智技术的广泛应用和平等教育机会。

（4）技术应用的广泛性和潜在影响：数智技术的应用虽然带来了便利，但也存在一些潜在的不足，如可能分散学生的注意力(舒旭 & 胡超, 2020)、增加认知负荷，以及对学生视力和健康的影响。因此，教育者和研究人员需要对这些潜在问题保持警觉，并寻找相应的解决方案。

4. 结论

数智技术在 K-12 教育领域的应用正逐渐改变教育的面貌，其优势和挑战都需要未来的研究和实践给予充分关注。以下是对数智技术在教育中应用的精炼总结和建设性见解：

（1）优势聚焦：数智技术通过智能辅导系统和聊天机器人等工具，有助于提升低水平学生的自我调节学习能力，提供定制化的学习体验。教师通过技术辅助更高效地进行教学设计和实时学生学习数据分析，从而提前调整教学策略，提高教学质量。同时，云计算等技术降低了教育资源的成本，提高了资源的可及性和教育的整体效率。

（2）挑战应对：教师需要通过持续的专业发展，掌握必要的技术应用能力和教学法，以适应数智技术带来的教学变革。研究应关注如何创新教学法，使教师能够灵活运用数智技术，提供实时交互和情境学习，以增强学习效果。同时，数智技术的应用应考虑到不同年级、学科和地区的特点，特别是在乡村和非 STEM 领域，以实现教育公平和技术的广泛应用。

（3）未来研究方向：首先，未来的研究应全面评估数智技术在提升学习效果、教师教学和学校管理效率方面的长期影响。其次，研究应包括对数智技术项目的成本效益分析，以支持教育决策者做出明智的投资和实施决策。最后，研究应探讨如何减少数智技术可能带来的健康和认知负担，确保学生在享受技术便利的同时，也能保护其身心健康。

参考文献

贾桐. (2023). 互联网+时代乡村小规模学校教育困境现状、原因与对策. 中国教育技术装备, 1 – 3.

金波 & 杨鹏. (2022). “数智”赋能档案治理现代化：话语转向、范式变革与路径构筑. 档案学研究, *2*, 4 – 11. https://doi.org/10.16065/j.cnki.issn1002-1620.2022.02.001

舒旭 & 胡超. (2020). 人工智能视域下农村基础教育发展困境及突破策略研究. 教育导刊, *6*, 18－24. https://doi.org/10.16215/j.cnki.cn44-1371/g4.2020.06.003

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, *0*(0), 1－17. https://doi.org/10.1080/10494820.2023.2172044

Chounta, I.-A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, *32*(3), 725－755.

Holstein, K., & Aleven, V. (2022). Designing for human－AI complementarity in K-12 education. *AI Magazine*, *43*(2), 239－248. https://doi.org/10.1002/aaai.12058

Huang, X. (2022b). Application of artificial intelligence APP in quality evaluation of primary school science education. *Educational Studies*, *0*(0), 1－21. https://doi.org/10.1080/03055698.2022.2066462

Jang, Y., Choi, S., Jung, H., & Kim, H. (2022). Practical early prediction of students' performance using machine learning and eXplainable AI. *Education and Information Technologies*, 1－35.

Kuleto, V., Ilić, M. P., Bucea-Manea-Ţoniş, R., Ciocodeică, D.-F., Mihălcescu, H., & Mindrescu, V. (2022). The Attitudes of K－12 Schools' Teachers in Serbia towards the Potential of Artificial Intelligence. *Sustainability*, *14*(14), Article 14. https://doi.org/10.3390/su14148636

Lashari, T. A., Sajid, U., & Lashari, S. A. (2022). The Effective Use of Digital Storytelling and Flipped Classroom Instructional Approach to Improve Science Subjects. *International Journal of Instruction*, *15*(4).

Lechuga, C. G., & Doroudi, S. (2022). Three Algorithms for Grouping Students: A Bridge Between Personalized Tutoring System Data and Classroom Pedagogy. *International Journal of Artificial Intelligence in Education*, 1－42.

Murphy, R. F. (2019). Artificial intelligence applications to support K-12 teachers and teaching. *Rand Corporation*, *10*.

Pelaez, A., Jacobson, A., Trias, K., & Winston, E. (2022). The Turing Teacher: Identifying core attributes for AI learning in K-12. *Frontiers in Artificial Intelligence*, *5*. https://www.frontiersin.org/articles/10.3389/frai.2022.1031450

Qushem, U. B., Christopoulos, A., Oyelere, S. S., Ogata, H., & Laakso, M.-J. (2021). Multimodal Technologies in Precision Education: Providing New Opportunities or Adding More Challenges? *Education Sciences*, *11*(7), 338. https://doi.org/10.3390/educsci11070338

Starks, A. (2022). Serving students with disabilities in K-12 online learning: Daily practices of special educators during the COVID-19 pandemic. *Distance Education*, *43*(4), 596－620. https://doi.org/10.1080/01587919.2022.2144138

Wang, X., & Wong, B. (2019). Bridging Knowledge Divides Utilizing Cloud Computing Learning Resources in Underfunded Schools: Investigating the Determinants. *Journal of Educational Computing Research*, *57*(3), 591－617. https://doi.org/10.1177/0735633118759455

Xia, Q., Chiu, T. K. F., & Chai, C. S. (2022). The moderating effects of gender and need satisfaction on self-regulated learning through Artificial Intelligence (AI). *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11547-x

学生协作学习过程中认知参与状态的自动分类方法

Automatic Classification of Cognitive Engagement States in Student Collaborative Learning Processes

吴涵[1]，张立山[1*]，王红晔[1]，段腾飞[1]，白志创[1]，李雲帆[1]
[1]华中师范大学人工智能教育学部
[*] lishan.zhang@ccnu.edu.cn

【摘要】 在线协作学习已经成为现阶段一种至关重要的学习方式，然而很少有研究通过对学生协作学习中的认知参与状态进行自动分类来评估协作学习的质量。通过比较应用传统的集成学习模型、BERT 类预训练模型、GPT 类预训练模型这三种方法，对协作学习过程中收集到的文本数据进行自动分类，发现 GPT 凭借其少样本、结合上下文语序的特点分类结果最佳。将该方法应用到学生协作学习过程中的自动评估体系中，将有利于教师及时对学生的学习做出有效和准确的干预。

【关键词】 协作学习；自动分类；GPT

Abstract: *Online collaborative learning has become a crucial learning modality in the current stage. However, there is limited research that evaluates the quality of collaborative learning by automatically classifying students' cognitive engagement states during collaborative learning activities. By comparing three approaches, namely traditional ensemble learning models, BERT-based pre-trained models, and GPT-based pre-trained models, for automatic classification of textual data collected during collaborative learning processes, it was observed that GPT achieved the best classification results, leveraging its advantages in handling limited samples and considering contextual language sequencing. Integrating this method into an automated assessment system for student collaborative learning processes would facilitate timely and effective interventions by teachers to enhance student learning.*

Keywords: collaborative learning, automatic classification, GPT

1. 研究背景和分类方法

在“互联网+教育”时代，在线协作学习已经成为一种至关重要的学习方式，对协作学习过程的评估，大多是评估学生学习过程中的文本数据，例如，张立山等人使用连续词袋模型和 LSTM 分类器模型，对半开放式问题进行自动评分(Zhang, Huang, Yang, Yu, & Zhuang, 2022)。然而，鲜少有研究对学生协作学习过程中的认知参与状态进行自动评估。本研究使用的数据集来自华中地区某高校开设的一门诗词鉴赏类课程中收集到的 607 条批注数据和 404 条讨论数据，数据收集的工具是研究团队开发的协作批注平台。根据季清华教授等人提出的 ICAP 认知投入分类框架(Chi & Wylie, 2014)，研究首先对学生的批注数据和讨论数据进行分类编码。然后通过使用传统的集成学习模型、BERT 类预训练模型、GPT 类预训练模型这三种方法，对收集到的数据进行自动分类，最后对分类结果进行比较分析和总结。

在进行分类之前，研究首先将批注数据和讨论数据分别按照 8:2 的比例划分为训练集和测试集，得到批注训练集 484 条，测试集 123 条，讨论训练集 320 条，测试集 84 条，前两种方法均使用训练集训练模型，测试集评估模型。首先使用随机森林这种传统的集成学习模型对批注数据和讨论数据进行分类。模型的参数设置相同，决策树数量设置为 100，随机数为 42。然后采用 BERT 模型的改进版本 RoBERTa 模型对两个数据集进行分类，与其他预训练的大型语言模型相比，RoBERTa 模型在各种分类任务中具有更优越的性能(Shaheen, Wohlgenannt, & Filtz, 2020)。同样对两个数据集进行分类的模型参数设置相同，最大文本长度设置为 200，学习率为 2e-5，训练批次 batch_size 为 16，训练轮次 epoch 值是 5，在训练

过程中使用交叉熵损失和 Adam 优化器进行模型更新。第三种自动分类方法是使用最强大的生成式人工智能模型 ChatGPT 进行自动分类。首先使用 GPT4 的 playground 平台进行提示词的调试和修改，在对少量数据的分类达到一定准确度后再将提示词送入后端代码，通过调用 GPT4 的 API 进行自动分类。批注和讨论的提示词结构大致相同，主要由背景、分类指令和从训练集中挑选出的具有代表性的少量例子构成，唯一区别的地方在于批注数据的提示词还包含动态输入的对应高亮句子以及两个相似度较高的参考资料的句子。使用 GPT4 进行自动分类的温度值 temperature 设置为 0，文本处理最大长度设置为 4096。

2. 分类结果与讨论

评估三种分类方法的指标包括准确率、精确率、召回率、F1 分数和 Kappa 值。三种方法的准确率和 Kappa 值对比如表 1 和表 2 所示。从表中可以看出，随机森林的分类效果总体表现最低，讨论数据分类的准确率只有 0.48，并且 Kappa 值为 0.32。RoBERTa 分类效果有所改进，尤其是讨论数据，其准确率提高到 0.67，而 Kappa 值达到了 0.54。相比之下，GPT4 的分类效果最好，无论是批注数据还是讨论数据，分类准确率均高于 0.7，Kappa 值也都在 0.5 以上，模型分类的结果和实际结果具有较高的一致性。

表 1 三种自动分类方法应用于两个数据集的 Kappa 值

	批注分类的 Kappa	讨论分类的 Kappa
Random Forest	0.33	0.32
RoBERTa	0.39	0.54
GPT4	0.55	0.72

表 2 三种自动分类方法应用于两个数据集的准确率

	批注分类的准确率	讨论分类的准确率
Random Forest	0.59	0.48
RoBERTa	0.62	0.67
GPT4	0.70	0.77

本研究使用了三种自动分类方法来对学生协作学习过程中认知参与状态进行自动评估。从得到的结果可以看出，相比于传统的集成学习模型，BERT 类预训练模型利用其对语言的上下文理解，表现出更高的准确性，但是它对标记化输入的依赖以及对大量计算资源的需求限制了它的实用性。表现最出色的是 GPT 类预训练模型，即使在少量样本的情况下，也得到了较好的结果，这凸显了生成式大语言模型的潜力。在未来的研究中，我们将继续探索完善使用 GPT 类预训练模型进行自动分类的规范流程，将该方法应用到学生协作学习过程中的自动评估体系中，为教师带来更准确的学生认知学习状态的反馈情况，从而方便教师及时有效的对学生的学习做出干预。

致谢

本项目受国家自然科学基金面上项目支持，项目批准号：62377017。

参考文献

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. Educational Psychologist, 49(4), 219-243.

Shaheen, Z., Wohlgenannt, G., & Filtz, E. J. a. p. a. (2020). Large scale legal text classification using transformer models. arXiv preprint arXiv.

Zhang, L., Huang, Y., Yang, X., Yu, S., & Zhuang, F. (2022). An automatic short-answer grading model for semi-open-ended questions. Interactive learning environments, 30(1), 177-190.

使用 GAFCC 游戏化模型提高在线协作学习背景下的学生参与度

Using GAFCC Gamification Model to Enhance Student Engagement in Online Collaborative Learning Context

HUANG Biyun[1*], HEW Khe Foon [2], PENG Jun[1], DAI Yi[1], JONG Morris Siu-Yung [3]
[1]School of Education, City University of Macau
[2]Faculty of Education, The University of Hong Kong
[3]Department of Curriculum and Instruction, The Chinese University of Hong Kong
[*]byhuang@cityu.edu.mo

【摘要】 游戏化是在非游戏环境中使用类似游戏的元素，以吸引用户并鼓励他们做出期望的行为。本研究采用实证研究法来检验基于理论的游戏化设计模型在本科生课程中的应用，并研究游戏化在在线协作学习环境中能给学生习提供的支持。我们收集并分析了调查问卷、半结构式访谈数据。访谈结果显示，大多数学生更喜欢在游戏化环境中学习。他们表示，在游戏化的环境中，他们会更加关注自己提交材料的质量，与小组内外的同伴进行更多的互动，并从同伴身上学到更多的东西。

【关键词】在线参与；游戏化；激励；感知；高等教育

Abstract: *Gamification uses game-like elements in non-game environments to engage users and encourage desired behaviors. We conducted an empirical study to examine the application of a theory-based gamification design model in an undergraduate course to support student learning in collaborative online learning environment. We analyzed data from questionnaires and semi-structured interviews. The results of the interviews showed that most students preferred to learn in a gamified environment. They expressed that in the gamified environment, they pay more attention to the quality of their submissions, interact more with their peers, and learn more through interacting with their peers.*

Keywords: online engagement, gamification, motivation, perception, higher education

1. 前言

学生参与(student engagement)是指学生参与学习活动的程度，主要包括行为参与、情感参与和认知参与三个方面(Reeve，2012)。然而，在设计在线协作学习情境时，提升学习者认知参与度是一个挑战(Kyewski & Kraemer, 2018)。有证据表明，即使是在精心设计的合作过程学习中，学生们的互动状态也往往是在比較表层水平上的互动，他们往往缺乏与同学通过在线互动的方式，来进行更为深入的思维上的交流。有时在互动中，没有将以往的知识和经验融入到现有的讨论中，因此未能通过这样的活动提升高阶思维的能力(Hernández-Sellés et al., 2019)。针对这一问题，本文研究了游戏化是否能作为一种有效的策略，促进学生在协作学习情境中的互动，并在协作过程中生成高质量的内容。同时，本文还研究了在游戏化支持的环境中，学生是否会感到缺乏自主性，以及他们如何看待游戏化对他们的影响。

本文的研究问题是：

a. 在游戏化支持的在线协作学习情境中，学生的学习动机水平如何？学生是否会认为游戏化是一种额外的控制，从而降低学习动机？

b. 学生对游戏化在线协作学习的看法如何？他们是否认为游戏化有助于： (1) 加强同伴间的协作和互动；(2) 提高在线参与的数量；(3) 提高在线参与的质量？

2. 相关文献

参与度高的学生往往在学习上会更加努力，规范自己的行为以实现目标，勇于接受挑战以优化学习成效(Klem & Connell, 2004) 。学生的参与不仅受学习者个人特点的影响，也受学习环境的影响。因此，研究游戏化能否促进在线协作学习环境的营造，从而促進学生的学习成效的提高很有意义。为了应对让学生参与在线协作环境所面临的挑战，本研究探讨了游戏化设计能否在高等教育环境中提高学生的认知参与度和情感参与度。

为了让游戏化设计具备合理性，我们参考了基于理论的 GAFCC 游戏化设计模型(Huang & Hew, 2018)去设计游戏化体验。这个 model 由两个部分组成：(1) 理论框架部分和 (2) 五步设计流程。此模型介绍了如何针对课程学习目标引入游戏化元素，来激励学生并促进学生学习。GAFCC 游戏化设计模型(Huang & Hew, 2018; Huang et al., 2019)的理论基础有参考并整合了自我决定理论(Ryan & Deci, 2000)、行为强化理论(Skinner, 1953)、目标设定理论(Locke & Latham, 2002)、心流理论(Csikszentmihalyi, 1990)和社会比较理论(Festinger, 1954)等。GAFCC 所提出的设计流程分为五个步骤：检查、决定、匹配、启动和评估。

3. 研究背景和方法

参与者是来自本科一年级的一个班级的学生。实验课程为期 10 周。共有 50 名学生同意并参与研究。参考 GAFCC 游戏化设计模型，作者与课程教师合作对信息管理课程进行了游戏化设计。在整个课程中会发放交流者徽章、超级团队徽章和隐藏奖励徽章，以鼓励学生参与在线论坛，表彰高质量的帖子，并激发更高水平的团队合作。同时，老师会针对学生们每周提交的论坛帖子进行点评。这些帖子的讨论主题与每周的学习主题密切相关。本文使用了源于 Ryan(1982)的内在动机量表(IMI)调查问卷，并邀请部分同学进行了访谈。

4. 研究结果

4.1 问卷结果

配对样本 t 检验结果表明，在努力(effort)方面，尽管学生在论坛上提交了更多且更高质量的帖子，他们的后测数值(M = 9.5, SD = 2.15) 会低于前测数值(M= 10.67，SD= 2.26) 。在兴趣 (interest)方面，学生的前后测数值没有显着差异。 在自主度(autonomy)方面，学生前后测数值也没有显着差异。这些数据表明，虽然课程里融入了游戏化元素，学生并不认为游戏化给他们增加了更多的学习负担，也没有觉得自己学习的自主度方面有影响。学生们没有觉得游戏化是一种额外的控制，没有因此而降低学习动机。

4.2 访谈结果

一共有九名学生参加了半结构式访谈，分享了他们的看法。所有同学都提到游戏化方式促进了他们与同学之间的协作与互动。例如，有同学提到为了赢得超级团队徽章，他们的队友之间有更加努力地一起去完成论坛上的各种任务，“我们与组员的互动更多了，因为我们需要回复对方的帖子。此外，我和我的组员在完成 Moodle 论坛活动时更加配合，因为我们想赢得团队奖励徽章……”有同学也提到“这增加了我们的交流频率，使我们能够相互学习。”

在线参与的数量方面，大多数受访者表示，游戏化提高了他们在论坛的参与度。他们提到，在传统的非游戏化课程中，很少有人在论坛上发帖。学生提到，“我认为游戏化鼓励我们回答更多问题，并积极参与论坛。”在线参与的质量方面，大多数学生认为他们提交帖子的

质量更高。有同学分享到，为了能提供有质量的回复，他/她还另外花了很多心思在自学上。学生说，“由于本课程采用了游戏化，我决定不仅要用自己现有的知识在论坛上回答问题，还必须提供高质量的回复。这鼓励我学习了更多的相关知识。

5. 讨论与总结

总的来说，研究证实基于 GAFCC 设计模型理论的游戏化设计能有效提高参与度。提供明确的徽章规则、使用徽章对发帖质量进行反馈，以及在课堂上辅以教师反馈，可以进一步提高学生之间的互动和提升互动内容的品质。但本文的局限性在于目前只采用了单组实验设计。今后，我们可以引入对照组，收集更多类型的数据，以了解游戏化可能产生的效果。此外，将来 GAFCC 游戏化设计模型还可应用于更多课程的开发和学习情境中，并检视它在促进学生的互动和合作学习方面的效果。

参考文献

Czikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140.

Hernández-Sellés, N., Muñoz-Carril, P. C., & González-Sanmamed, M. (2019).Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers & Education, 138,* 1-12.

Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education*, 125, 254-272.

Huang, B., Hew, K. F., & Lo, C. K. (2019). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments*, 27(8), 1106-1126.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, 74, 262-273.

Kyewski, E., & Kramer, N. C. (2018). To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers & Education*, 118, 25-37.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705.

Reeve, J. (2012) A Self-determination Theory Perspective on Student Engagement. In: S. Christenson, A. Reschly, C. Wylie (eds) *Handbook of research on student engagement*. Springer

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450-461.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68.

Skinner, B. F. (1953). *Science and human behavior*. Simon and Schuster.

教中学策略中线上同伴在场方式对视频学习的影响

Effect of Online Peer Presence in Learning by Teaching Strategy on Video Learning

郭鑫[1]，王婧然[1]，徐珂[1]，杨九民[1*]
[1]华中师范大学 人工智能教育学部，湖北 武汉 430079
[*]yjm@mail.ccnu.edu.cn

【摘要】 随着教中学策略在视频学习中的逐渐深入，越来越多学者开始关注教中学策略对视频学习有效性影响的边界条件。研究从先前知识经验和线上同伴在场方式两方面出发，探究两者对学习者采用教中学策略进行视频学习产生的影响。研究共招募 54 名华中地区某高校大学生，结果表明，低经验学习者采用教中学策略进行视频学习时，相比于有同伴形象，无同伴形象时学习效果更好，学习效率更高，感知到的压力更小；而高经验学习者的学习效果无显著差异。研究结果对学习者采用教中学策略开展视频学习具有一定的启示意义。
【关键词】 视频学习；教中学；先前知识经验；同伴在场

Abstract: *With the gradual development of learning by teaching in video learning, there are more scholars focusing on the boundary conditions of the influence of learning by teaching on the effectiveness of video learning. The study explored the effects of prior knowledge and online peer presence on video learning by learners using learning by teaching. A total of 54 students from a university in the central China region were recruite, and results showed that low prior knowledge learners using learning by teaching were more effective, efficient, and perceived less stressful in video learning without peer image than with peer image; but there was no significant difference in the learning outcomes of the high prior knowledge learners. The findings have implications for learners using the learning by teaching in video learning.*

Keywords: video learning, learning by teaching, prior knowledge, presence of peer

基金项目：2023 年度优秀研究生教育创新资助项目“不同类型的虚拟教学代理对学习者视频学习的影响研究”（项目编号：2023CXZZ093）；教育部产学合作协同育人项目“智慧教学环境下的师范生教学能力培养实验教学体系研究”（项目编号：202101202009）。

1. 研究背景

随着教育信息化的不断发展，在线学习已成为正式和非正式教育环境中普遍存在的一种学习方式（Scagnoli et al., 2019）。视频作为在线学习的一种重要学习资源备受学习者喜爱。为提高视频学习的有效性，众多学者对视频中的关键元素进行探索与设计（例如，线索运用、教师形象呈现方式等），总结归纳了众多视频设计方法（Yang et al., 2021）。尽管这些方法可以有效管理学习者有限的认知能力，但不能保证学习者主动思考且进行有意义的视频学习。研究者逐渐开始意识到生成性学习策略的重要性，其中教中学是一种典型的生成性学习策略。

教中学（Learning by teaching）是指学习者通过将所学内容讲授给同伴以加深对知识的理解（王福兴，2023）。生成性学习理论认为，将学习内容讲授给同伴，可以触发一系列生成

性认知过程（Fiorella & Mayer, 2014）；有助于学习者监控自身对学习内容的理解水平，进而制定有效的认知策略以提升其学习效果（Roscoe, 2014）。关于教中学策略在视频学习中的应用，研究者进行了大量的实证研究。研究发现，相较于复习或总结策略，采用教中学策略可以提高学习者的学习效果。例如，Fiorella 等人（2020）的一项研究发现，在视频学习后，相比于重复学习，采用教中学策略的学习者保持和迁移测试成绩得分更高。Hoogerheide 等人（2019）在小学生群体中发现，相较于生成书面总结，采用教中学策略加深了学习者对学习材料的理解。然而，随着研究的深入，相比于单纯关注教中学策略是否对视频学习产生影响，越来越多的研究者开始关注影响教中学策略对视频学习有效性的边界条件。其中先前知识经验和线上同伴在场方式是非常重要的两个条件。线上同伴在场是指同伴与学习者处于不同物理空间，但处于同一虚拟或在线学习环境。线上同伴在场方式有两种，一种是呈现同伴形象的同伴在场，另一种无同伴形象的同伴在场。

社会促进理论认为，学习者在完成某项活动或任务时，他人在场或与他人一同活动会提高任务完成表现或效率（华红琴，2004）。基于社会促进理论，学习者在独立完成任务时，相较于无同伴形象，有同伴形象可能会激发学习者的表现欲，而表现欲的提高能够增强学习者的自我唤醒水平，促进学习者的学习表现。然而，社会抑制理论认为，学习者完成某项活动或任务时，由于有同伴的存在会导致学习者任务完成表现不如其单独完成（Laura et al., 2015）。基于此，学习者在运用教中学策略进行视频学习时，同伴形象的存在可能会使学习者心理压力增大，导致注意力无法集中，不利于学习者学习。因此，在运用教中学策略进行视频学习时，线上同伴在场方式对学习是否产生影响尚未可知。

不同先前知识经验的学习者采用教中学策略也会对学习产生一定的影响。基于经验反转理论，在视频学习时，使用教中学策略对于低经验学习者可能是一种有效的学习策略，但对于高经验学习者可能无效甚至产生消极影响。认知负荷理论认为，学习者有限的工作记忆容量会导致认知超负荷，但额外向同伴讲授的过程会降低认知负荷（Paas et al., 2016）。Pi 等人（2023）的一项研究也发现，不同先前知识经验的学习者在向同伴解释时确实存在经验反转效应。然而，Leonie 等人（2022）的研究并未发现在使用教中学策略时这一效应的存在。此外，同伴形象存在与否也会对不同先前知识经验的学习者采用教中学策略进行视频学习产生影响。Andreas 等人（2019）的研究发现，低经验学习者向虚拟听众讲授所学内容时，产生的担忧情绪对其学习效果和认知资源分配产生干扰。因此，向不同线上在场方式的同伴讲授学习内容对视频学习产生的影响也可能受学习者先前知识经验水平的影响。

综上所述，本研究主要探究学习者采用教中学策略进行视频学习时，线上同伴在场方式以及自身先前知识经验对学习产生的影响。

2. 研究过程与方法

2.1. 研究对象和设计

研究随机招募 54 名华中地区某高校大学生作为被试，年龄为 19 到 25 岁（M = 20.92，SD = 1.33），男生 14 名，女生 40 名。其中 23 名被试为生物科学专业，31 名被试为其他专业。所有被试在被告知实验流程后，签署一份知情同意书，完成实验后收到一定的被试费。

研究采用两因素混合实验设计，线上同伴在场方式为被试内因素（有同伴形象 vs.无同伴形象），被试的先前知识经验水平为被试间因素（高经验 vs.低经验）。被试先验知识高低依据其先验知识得分情况进行划分，且方差分析结果显示，两组先验知识存在显著差异（$t(52) = 4.99, p < .001$）。

2.2. 实验材料

实验材料包含两组教学视频，视频内容分别为《雌蕊的发育与结构》和《雄蕊的发育与结构》，视频画面示例如图 1 和图 2 所示。视频每隔两到三分钟便会出现一次教学解释提示，要求被试向线上同伴提供所学视频内容的教学解释，此过程中无法查看视频。每个视频被试需要向线上同伴提供三次教学解释。无同伴形象条件下的提示画面和有同伴形象条件下的提示画面分别如图 3 和图 4 所示。

三名生物专业教师对每组视频材料进行同质性评判，结果表明两组教学视频在内容难度与知识类型方面具有高度一致性。



图 1 雌蕊的发育与结构



图 2 雄蕊的发育与结构

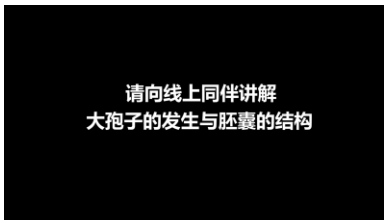


图 3 无同伴形象教学解释提示画面



图 4 有同伴形象教学解释提示画面

2.3. 测量工具

2.3.1. 基本信息问卷

基本信息问卷用于统计被试的人口统计学特征，包括被试的姓名、性别、年龄、专业等基本信息。

2.3.2. 先验知识水平测试

先验知识测试由教学视频讲授者编制，并经一名生物专业大学教师审核。该测试包含 3 道选择题和 2 道判断题，每题 10 分，总分 50 分。为避免预期对学习效果的影响，测试题目不包含与教学视频相关的内容。

2.3.3. 学习效果测试

学习效果测试包含即时测试和延时测试。被试在完成实验后立即填写即时测试问卷，在完成实验一周后填写延时测试问卷。每组实验条件下的即时测试由 10 道题组成，包括 5 道单选题和 5 道判断题。每题 10 分，满分为 100 分（问卷信度良好，Cronbach’s a=0.82）；延时测试包含 10 道单项选择题，涉及两组教学视频中讲解的知识，每题 10 分，满分 100 分（问卷信度良好，Cronbach’s a=0.85）。

2.3.4.学习效率

为了测量同伴在场方式对学习效率的影响，本研究利用心理努力和即时测试得分来计算被试的学习效率（Yang et al, 2021）：学习效率=（$Z_{即时测试}$-$Z_{心理努力}$）/$\sqrt{2}$。学习效率分数大于零，表示具有较高的学习效率，相反，学习效率分数小于零，表示具有较低的学习效率。

2.3.5. 学习动机问卷

学习动机问卷改编自 Lin 等人（2020）的研究。学习动机问卷共包括兴趣，能力，价值，压力四个维度。每个维度对应三道题目，其中兴趣维度 1 道题、压力维度 2 道题需反向计分。（问卷信度良好，兴趣维度：Cronbach’s a=0.88；能力维度：Cronbach’s a=0.86；价值维度：Cronbach’s a=0.90；压力维度：Cronbach’s a=0.88）

2.3.6. 认知负荷问卷

认知负荷问卷改编自 Pass 等人（2016）的认知负荷测量量表。该量表从心理努力和材料感知难度两个维度对被试的认知负荷进行测量，共七个题目，采用李克特 5 点量表计分（问卷信度良好，Cronbach’s a=0.90）。

3.3.7. 监控准确性

学习判断问卷（JOL）改编自 Lindner 等人（2021）的研究，共 3 道题目，评分范围由 0 到 100：用来评测被试学习内容掌握程度以及将所学知识应用到学习情境的信心。3 道题的平均分作为学习判断总分，然后通过观察学习判断得分与即时测试得分之差的绝对值来评估被试的监控准确性。得分越接近零，表明学习者的监控准确性越高（问卷信度良好，Cronbach’s a=0.92）。

3.3.8. 解释质量

解释质量的评分标准有四个维度：内容组织、原创性、流畅性和精细化（Fu et al., 2019）。内容组织指的是解释具有逻辑性和连贯性。原创性指的是解释教学内容时主动重组加工的程度。流畅性指的是能够产生更多的想法来解释知识点。精细化是指在讲解过程中细节描述的能力。研究采用录音的方式收集被试的教学解释。每个解释均由两位评审员从四个维度进行评分（最低 1 分，最高 5 分），最终得分为两评审员打分的平均值。

2.4. 实验流程

研究在实验室完成，大约持续 30 分钟。被试首先需要完成一份人口统计学问卷和先验知识水平测试。依据先验知识水平测试得分，被试被划分为低经验组或高经验组，每组被试均需要完成两轮相对独立的实验流程，被试随机分配到某一顺序中：①有同伴形象的教学视频→无同伴形象的教学视频；②无同伴形象的教学视频→有同伴形象的教学视频。每轮实验结束后，被试均需要填写即时测试问卷与其他后测问卷。并于七天后完成线上延时测试问卷。具体流程如图 5。

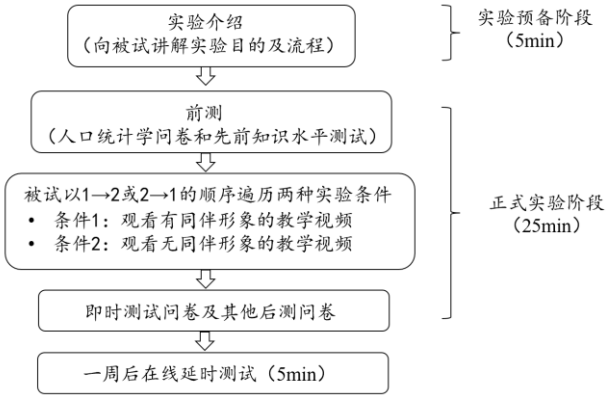


图 5 实验流程

3. 研究结果

本研究以先前知识经验（高经验 vs.低经验）*线上同伴在场方式（有同伴形象 vs.无同伴形象）为自变量，以学习效果（即时测试成绩、延时测试成绩）、学习效率、学习动机、认知负荷监控准确性以及解释质量为因变量进行重复测量方差分析。描述性统计结果如表 1 所示。

表 1 描述性统计结果

变量	低经验		高经验	
	无同伴 *M*(*SD*)	有同伴 *M*(*SD*)	无同伴 *M*(*SD*)	有同伴 *M*(*SD*)
即时测试成绩	64.44 (18.26)	45.19 (14.77)	81.48 (15.37)	77.41 (11.30)

延时测试成绩	12.59 (12.63)	10.74 (11.07)	14.07 (13.09)	11.48 (11.67)
学习效率	-10.21 (13.77)	-22.26 (12.96)	7.07 (19.80)	4.71 (15.69)
学习动机（兴趣）	3.48 (0.54)	3.51 (1.03)	3.53 (0.50)	3.61 (0.67)
学习动机（能力）	2.95 (1.10)	2.60 (1.15)	2.93 (0.77)	2.86 (0.61)
学习动机（价值）	3.75 (0.47)	3.50 (0.44)	3.99 (0.57)	4.10 (0.48)
学习动机（压力）	3.37 (0.58)	3.85 (0.50)	3.17 (0.35)	3.25 (0.50)
认知负荷	3.69 (0.45)	3.66 (0.47)	3.31 (0.52)	3.34 (0.59)
监控准确性	21.41 (15.27)	19.98 (16.02)	14.41 (12.71)	16.47 (11.72)
解释质量	12.76 (1.24)	12.17 (1.10)	13.83 (1.07)	13.65 (1.22)

3.1. 学习效果

3.1.1.即时测试成绩

重复测量方差分析结果显示，先前知识经验主效应显著（$F(1, 52) = 52.31$, $p < .001$, $\eta_p^2 = 0.50$），高经验学习者即时测试成绩显著高于低经验学习者（$MD = 24.63$, $p < .001$）；线上同伴在场方式主效应显著（$F(1, 52) = 25.42$, $p < .001$, $\eta_p^2 = 0.33$），学习者在无同伴形象条件下的即时测试成绩显著高于有同伴形象条件（$MD = 11.67$, $p < .001$）；先前知识经验和线上同伴在场方式交互效应显著（$F(1, 52) = 10.77$, $p = .002$, $\eta_p^2 = 0.17$；见图 6），简单效应分析结果发现，低经验学习者在无同伴形象条件下的即使测试得分显著高于有同伴形象（$MD = 19.26$, $p < .001$），而高经验学习者无显著差异（$MD = 4.07$, $p = .219$）。

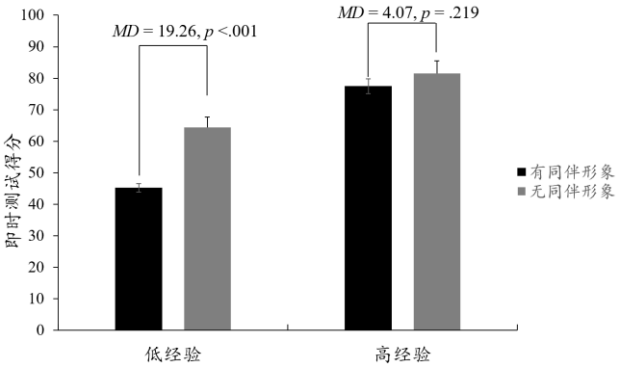


图 6 即时测试成绩结果

3.1.2.延时测试成绩

重复测量方差分析结果显示，先前知识经验主效应不显著（$F(1, 52) = 0.27$, $p = .606$, $\eta_p^2 = 0.05$），线上同伴在场方式主效应不显著（$F(1, 52) = 0.78$, $p = .381$, $\eta_p^2 = 0.02$），先前知识经验和线上同伴在场方式交互效应不显著（$F(1,52) = 0.22$, $p = .883$, $\eta_p^2 = 0.01$）。

3.2. 学习效率

重复测量方差分析结果显示，先前知识经验主效应显著（$F(1, 52) = 34.53$, $p < .001$, $\eta_p^2 = 0.40$），与低经验学习者相比，高经验学习者的学习效率更高（$MD = 22.13$, $p < .001$）；线上同伴在场方式主效应显著（$F(1, 52) = 12.16$, $p < .001$, $\eta_p^2 = 0.19$），与有同伴形相比，无同伴形象时学习者的学习效率更高（$MD = 7.20$, $p < .001$）；先前知识经验和线上同伴在场方式交互效应显著（$F(1, 52) = 5.50$, $p = 0.023$, $\eta_p^2 = 0.10$；见图 7），简单效应分析结果发现，低经验学习者在无同伴形象条件下的学习效率显著高于有同伴形象（$MD = 12.05$, $p < .001$），而高经验学习者无显著差异（$MD = 2.35$, $p = .423$）。

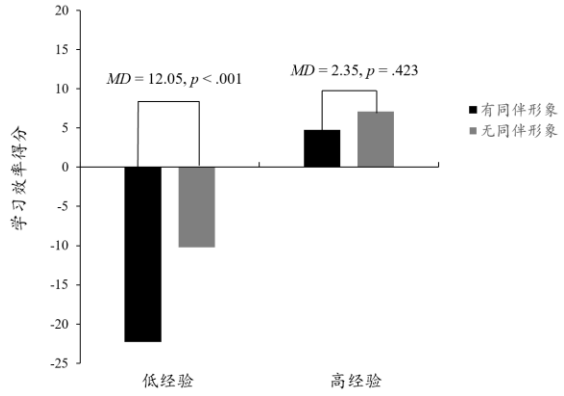


图 7 学习效率结果

3.3. 学习动机

重复测量方差结果显示，在兴趣维度，先前知识经验主效应不显著（$F(1, 52) = 0.90$, $p = .349$, $\eta_p^2 = 0.02$）；线上同伴在场方式主效应不显著（$F(1, 52) = 7.69$, $p = .078$, $\eta_p^2 = 0.07$）；先前知识经验和线上同伴在场方式交互效应不显著（$F(1, 52) = .36$, $p = .549$, $\eta_p^2 = 0.01$）。

在能力维度，先前知识经验主效应不显著（$F(1, 52) = .24$, $p = .624$, $\eta_p^2 = 0.01$）；线上同伴在场方式主效应不显著（$F(1, 52) = 5.32$, $p = .075$, $\eta_p^2 = 0.09$）；先前知识经验和线上同伴在场方式交互效应不显著（$F(1, 52) = 2.59$, $p = .114$, $\eta_p^2 = 0.05$）。

在价值维度，先前知识经验主效应显著（$F(1, 52) = 13.10$, $p = .001$, $\eta_p^2 = 0.20$），高经验学习者价值得分显著高于低经验学习者（$MD = 0.3$, $p = .001$）；线上同伴在场方式主效应不显著（$F(1, 52) = .98$, $p = .327$, $\eta_p^2 = 0.02$）；先前知识经验和线上同伴在场方式交互效应显著（$F(1, 52) = 6.80$, $p = .012$, $\eta_p^2 = 0.12$；见图 8），简单效应分析结果发现，低经验学习在有同伴形象条件下价值得分显著高于无同伴形象（$MD = 0.25$, $p = .014$），而高经验学习无显著差异（$MD = 0.1$, $p = .258$）。

在压力维度，先前知识经验主效应显著（$F(1, 52) = 14.37$, $p < .001$, $\eta_p^2 = 0.22$），低经验学习者压力得分显著高于高经验学习者（$MD = 0.40$, $p < .001$）；线上同伴在场方式主效应显著（$F(1, 52) = 11.65$, $p = .001$, $\eta_p^2 = 0.18$），学习者在有同伴形象条件下，压力得分显著高于无同伴形象（$MD = 0.28$, $p = .001$）；先前知识经验和线上同伴在场方式交互效应显著（$F(1, 52) = 6.27$, $p = .015$, $\eta_p^2 = 0.11$；见图 9），简单效应分析结果发现，低经验学习者在有同伴形象条件下的压力得分显著高于无同伴形象（$MD = 0.48$, $p < .001$），而高经验学习无显著差异（$MD = 0.07$, $p = .523$）。

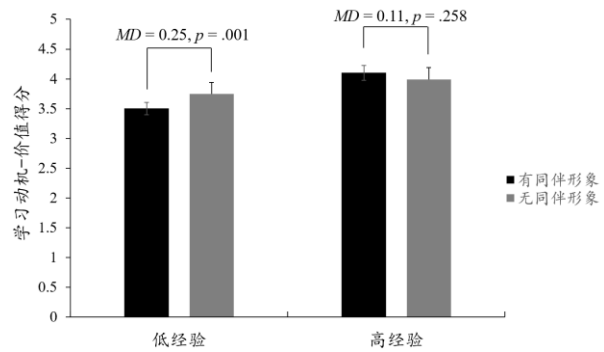


图 8 学习动机-价值结果

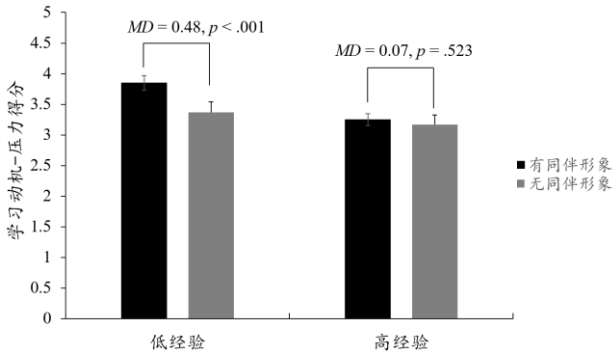


图 9 学习动机-压力结果

3.4. 认知负荷

重复测量方差分析结果显示，先前知识经验主效应显著（$F(1, 52) = 7.38$, $p = .009$, $\eta_p^2 = 0.12$），低经验学习者认知负荷显著高于高经验学习者（$MD = 0.35$，$p = .009$）；线上同伴在

场方式主效应不显著（$F(1, 52) = .00, p = 1.00, \eta_p^2 < 0.001$）；先前知识经验和线上同伴在场方式交互效应不显著（$F(1, 52) = 0.39, p = .535, \eta_p^2 = 0.01$）。

3.5. 监控准确性

重复测量方差分析结果显示，先前知识经验主效应显著（$F(1, 52) = 5.00, p = .030, \eta_p^2 = 0.09$），高经验学习者的监控准确性显著高于低经验学习者（$MD = 7.25, p < .001$）；线上同伴在场方式主效应不显著（$F(1, 52) = 0.70, p = .408, \eta_p^2 = 0.01$）；先前知识经验和线上同伴在场方式交互效应不显著（$F(1, 52) = 3.44, p = .069, \eta_p^2 = 0.06$）。

3.6. 解释质量

重复测量方差分析结果显示，先前知识经验主效应显著（$F(1, 52) = 18.37, p < .001, \eta_p^2 = 0.26$），高经验的学习者的解释质量显著高于低经验的学习者（$MD = 1.28, p < .001$）；线上同伴在场方式主效应显著（$F(1, 52) = 14.35, p < .001, \eta_p^2 = 0.22$），无同伴形象时学习者的解释质量显著高于有同伴形象（$MD = 0.39, p < .001$）；先前知识经验和线上同伴在场方式交互效应不显著（$F(1, 52) = 3.94, p = .053, \eta_p^2 = 0.07$）。

4. 讨论

本研究探究了学习者采用教中学策略进行视频学习时，自身先前知识经验（高经验 vs. 低经验）和线上同伴在场方式（有同伴形象 vs.无同伴形象）对视频学习产生的影响。

在学习效果方面，研究结果表明，低经验学习者在视频学习中向同伴讲授时，相较于呈现同伴形象，无同伴形象时即时测试成绩更好、学习效率更高；而高经验学习者在两种情况下学习效果均无显著差异。无同伴形象时学习者解释质量显著高于有同伴形象。这些结果与以往研究一致（Pi et al., 2023）。基于认知负荷理论，低经验学习者由于缺少所学知识相关的认知图式以及有限工作记忆容量容易导致认知超负荷，但向同伴讲授知识会促使一系列生成性学习过程的发生，改善学习效果。根据社会抑制理论与经验反转效应，低经验学习者在向他人讲授所学内容时，同伴形象的存在可能会使学习者感到紧张、焦虑，认知加工不充分，无法生成详细的解释，不利于学习者学习；高经验学习者本身拥有较高的认知图式，可以进行高效认知加工，教中学策略设计是多余的，因此，在讲授过程中，无论是否呈现同伴形象，对其学习效果均无显著影响。

在学习动机方面，研究结果表明，低经验学习者在视频学习中向同伴讲授时，相较于呈现同伴形象，无同伴形象时感知到的压力更低、感知到的价值更低；而高经验学习者均无显著差异。这与认知负荷的研究结果对应（Paas et al., 2016）。认知负荷相关的研究表明，同伴的存在往往能够提高学习者对于任务价值的感知，但也会使学习者感知到一定的压力，间接增加了感知任务难度，不利于学习者学习。由于低经验学习者有限的认知，这可能加剧了同伴形象存在对于压力的感知。

在认知负荷、监控准确性方面，研究结果表明，高经验学习者比低经验学习者有更少的认知负荷以及更准确的监控准确性。这与以往研究一致，低经验学习者有限的认知资源容量需要付出更多的心理努力对所学知识进行解释。元认知理论认为，学习者缺乏的元认知知识与技能可以通过学习策略来激发或培养（Backer et al., 2012）。研究结果证明了高经验学习者比低经验学习者更能从教中学策略中开展有效的元认知加工，促进学习。

5. 研究结论

本研究发现，对于低经验学习者，相较于有同伴形象，无同伴形象条件下视频学习效果更好、学习效率更高、压力感知更小以及价值感知更低，而高经验学习者在两种条件下各方面得分均无显著差异。基于此，本研究对于未来有效开展视频学习启示如下：在鼓励学习者

采用教中学策略进行视频学习，尤其是先前知识经验较低的学习者，最好不要呈现同伴的形象，以避免增加学习者学习的压力，并提高学习效果。

本研究还存有一定的不足之处，后续研究可以采用眼动技术或脑电技术进一步探索视频学习中教中学策略影响其学习效果的内在认知神经机制。

参考文献

华红琴(2004)。社会心理学原理和应用。上海：上海大学出版社.

王福兴、黄宇、冷晓雪和匡子翌(2023)。自我生成教学促进多媒体学习：绘图的作用. 苏州大学学报(教育科学版), 11(03):97-106.

De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional science*, 40, 559-588.

Fiorella L, Mayer R E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, 39(2): 75-85.

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology*, 112(5), 895.

Fred G w C Pass, Jeroen J G van Merrienboer. (1994). Measurement of Cognitive Load in Instructional Research. *Perceptual and Motor Skills*, 79(1): 419-430.

Fu, Q. K., Lin, C. J., Hwang, G. Z., & Zhang, L. X. (2019). Impacts of a mind mapping-based contextual gaming approach on EFL students' writing performance, learning perceptions and generative uses in an English course. *Computers & Education*, 137, 59-77.

Hoogerheide, V., Visee, J., Lachner, A., & van Gog, T. (2019). Generating an instructional video as homework activity is both effective and enjoyable. *Learning and Instruction*, 64, 101226.

Lachner, A., & Neuburg, C. (2019). Learning by writing explanations: Computer-based feedback about the explanatory cohesion enhances students' transfer. *Instructional Science*, 47(1), 19-37.

Lijia Lin, Paul Ginns, Tianhui Wang et al. (2020). Using a pedagogical agent to deliver conversational style instruction: What benefits can you obtain? *Computers &Education*, 143.

Lindner M A, Eitel A, Barenthien J, et al. (2021). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction*, 71:10110.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2016). Cognitive load measurement as a means to advance cognitive load theory. *In Cognitive Load Theory* (pp. 63-71).

Pi, Z., Guo, X., Liu, C., & Yang, J. (2023). Students with low prior knowledge learn more when explaining to a peer than to a teacher after viewing educational videos. *Active Learning in Higher Education*, 14697874231212260.

Ribosa, J., & Duran, D. (2023). Students' feelings of social presence when creating learning-by-teaching educational videos for a potential audience. *International Journal of Educational Research*, 117, 102128.

Scagnoli, N. I., Choo, J., & Tian, J. (2019). Students' insights on the use of video lectures in online classes. *British Journal of Educational Technology*, 50(1), 399-414.

Wolf, L. K., Bazargani, N., Kilford, E. J., Dumontheil, I., & Blakemore, S. J. (2015). The audience effect in adolescence depends on who's looking over your shoulder. *Journal of adolescence*, 43, 5-14.

Yang, J., Zhang, Y., Xu, K., & Pi, Z. (2021). The effects of learning strategies on video learning: Imagination, drawing and self-explanation strategies. *E-Education Research*, 42(10), 40-47.

基于场馆的 K-12 儿童非正式学习效果研究：系统性文献综述

A Museum-based Study on Informal Learning Outcomes of K-12 Children: a Systematic Literature Review

刘倩，陈琦帆，罗恒 [1*]
[1] 华中师范大学 人工智能教育学部，湖北武汉
[*] luoheng@mail.ccnu.edu.cn

【摘要】 作为非正式学习的主阵地，场馆建设趋于完善，场馆学习倍受关注。本文系统梳理了国内外 20 年来（2003-2023）有关 K-12 儿童场馆学习的高质量期刊文章，以 Scopus 与中国知网为数据库来源，共检索 65 篇高度相关的实证研究，开展细致的编码分析。研究结果表明，目前对于 K-12 儿童的场馆学习结果研究集中关注科学知识以及情感态度维度，多以信息技术的综合开展实证研究，如 AR 与教育游戏、移动设备的结合等。此外，目前主要研究场馆为科技馆，学习内容主要集中在物理学与生物学领域。本文还提出了场馆学习的教学干预对策与建议。

【关键词】 场馆学习；非正式学习；K-12；儿童；学习效果

Abstract: *As the main front of informal learning, museum has improved, and museum learning has attracted much attention. This paper systematically reviewed the high-quality domestic and foreign journal literature on K-12 children's museum learning in the past 20 years (2003-2023). Using Scopus and CNKI as database sources, a total of 65 highly relevant empirical studies were retrieved and detailed coding analysis was carried out. The results show that the current research on K-12 children's informal learning outcomes focuses on the dimensions of scientific knowledge and emotional attitude, and most of them carry out empirical studies based on the integration of information technology, such as the combination of AR and educational games, and the combination of AR and mobile devices. In addition, the main research sites are science museums, and the learning content mainly focuses on the field of physics and biology. This paper also puts forward some teaching intervention strategies and suggestions for museum learning.*

Keywords: museum learning, informal learning, K-12, children, learning outcomes

1. 前言

“教育要从娃娃抓起”，随着国家对科学教育的重视以及终身学习教育理念的影响，非正式学习相关研究越来越受到学者关注。作为正式学习的补充，非正式学习更强调使学生以一个主动建构者的身份进行文化性和社会性的活动，是一种社会知识的创造过程和意义建构过程（Anderson & Thomas, 2014），以科技馆、博物馆、水族馆、动物园与植物园为代表的非正式学习场馆在教育中发挥着越来越重要的作用。场馆学习也形成了一个交叉研究领域，吸引了不少来自教育技术学、博物馆学等专业学科领域的研究者。儿童作为场馆学习的主要参观群体之一（季娇等，2017； Piscitelli & Anderson，2001），通过参观场馆，在视觉、听觉乃至触觉等感知体验中进行知识的建构，塑造个人的认知结构，形成学习（张美霞，2017）。已有研究指出基于场馆的非正式学习能够为 K-12 学习者提供真实的学习情境，发展 K-12 学习者的科学技能、实践能力以及个人兴趣等多个层面（He et al., 2023）。但 K-12 学习者的学习效果如何，是学者们一直比较关心的话题。现阶段，随着人工智能时代的到来，以虚拟现

实、增强现实为代表的新兴技术融入场馆学习中，场馆学习呈现出新趋势与新特点。目前只有少量的非系统且古早的文献综述，本文系统梳理了新时代、新技术、新发展的场馆学习研究现状，以报告 K-12 儿童的场馆学习的最新研究成果。

2. 概念界定

2.1. 基于场馆的非正式学习

20 世纪 50 年代，美国成人教育之父 Malcolm Knowles 首次提出“非正式学习”（Informal Learning）的概念，但未被学界所重视。直到 20 世纪 90 年代，研究者对于非正式学习的关注度增加。非正式学习成为终身学习的不可分割的一部分（Park et al., 2021）。非正式学习与正式学习相对应，区别于基于传统教学理念的，以学校课堂学习、研讨会、任务等形式开展的系统性、连续性学习活动（王妍莉等，2011），非正式学习主要指学习地点不正式、学习时间不正式、学习方式不正式。具体而言，非正式学习可以发生在工作、生活、社交等多情境下，有研究表明非正式学习占个体所学知识的 75%以上，而该领域的关注度和投入度却不到 20%（余胜泉和毛芳，2005）。

基于场馆的非正式学习（Museum Learning）又叫做“场馆学习”，“非正式场馆学习”。作为一种典型的非正式学习模式，这种学习方式通常发生在自然博物馆、历史博物馆、科技馆、天文馆、美术馆、动物园等各类场馆中，是一种情境化的学习（叶贝和董雁，2022）。场馆学习基于其资源、情境的优势，能够给予参观者丰富的学习体验，但是由于其学习内容分散、部分知识距离日常生活较远、学习时间零散等特点在某种程度上也限制了参观者对科学知识的深入理解(李睿和刘颖芳, 2023)。儿童的场馆学习效果有待更多的探索(Silver et al., 2008)。针对场馆学习的学习效果分类，国内外均有不少学者进行讨论，广为认同的是将场馆学习的学习效果聚焦于知识概念、情感态度、动作技能与社会互动等方面（伍新春等，2009）在数智融合时代，以人工智能技术为代表的新技术的应用丰富了场馆的形式，如虚拟场馆（许玮和张剑平，2016）、智慧场馆（骆晓红，2016）、数字化场馆等（Barbieri et al.，2017），因此，本研究开展文献梳理时关注了学者研究情景下的场馆类型。

2.2. K-12 儿童的场馆学习

通常 K-12 儿童指幼儿园到 12 年级（4-18 岁）的学生，下文以“儿童”代称。早期的基础教育往往局限于学校的课堂，以传授知识、显性教育为主。随着建构主义学习观的兴起，研究者们逐渐认识到基础教育不能仅仅局限于学校，而应该与社会生活发生广泛而深入的联系。儿童时期处于具像思考阶段，场馆给予儿童非正式学习的场所与环境，儿童可以在场馆参与学习情境，通过亲身参与与展品互动，获得真实而深刻的学习体验。场馆的展品大多为承载科学知识的具体物体，有助于儿童在具体图形与抽象概念之间建立联系，有研究表明，基于场馆的非正式学习环境能够为儿童提供真实的情境和丰富的互动机会，激发儿童的学习兴趣，培养儿童的动手操作和沟通交流的技能，促进儿童在参观体验中主动建构科学知识（Falk et al.，2010；Han et al.，2020）。在我国，很多馆校合作的项目和课程，将博物馆教育对象细分到了小学、初中的各个年级，形成了系统的、递进式的教育项目和校本课程（王秀江和王瑛，2019）。美国等西方发达国家非常注重科技场馆与学校的合作，开展了工作坊、课程、材料制作、活动资源包、场馆宿营等各式各样的教育服务，使馆校合作成为了中小学生学习科学知识的“第二课堂”（Weinstein et al.，2014）。

总体来说，儿童作为场馆学习的主要受众，针对 K-12 儿童群体的场馆学习的文献梳理能够帮助研究者们关注儿童受众的不同特征，更具有针对性地探究场馆学习中儿童学习的痛点和难点。因此，本研究聚焦儿童的非正式学习，探讨场馆学习的特点与发展趋势，对 2003 年以来国内外高质量期刊文章开展系统性文献综述，尝试梳理并回答场馆学习研究中研究设

计、教学设计、学习效果等多方面的特点与发展趋势，旨在厘清目前国内外场馆学习的样态特征，总结国内外已有经验，期望能为研究者与实践者有效开展移动学习、馆校合作等的创新实践与研究提供思路与策略。

3. 研究过程及方法

本研究采用系统性文献综述法（Systematic Literature Review），是一种通过系统地收集、筛选、评价、分析和综合相关研究文献的方法，以回答特定研究问题或者验证某种假设。因其研究问题的明确性、研究过程的透明性、研究结果的创新性，系统性文献综述法能够有效规避传统文献综述法存在的描述性、主观性等不足。本文按照 A.P. Siddaway et al., (2019) 人提出的系统综述流程进行研究，如图 1 所示，研究过程分为五个阶段：提出研究问题，界定文献范围，确定文献编码方案，报告研究结果，总结研究发现。此处重点介绍文献检索、编码方案以及编码的信度分析。

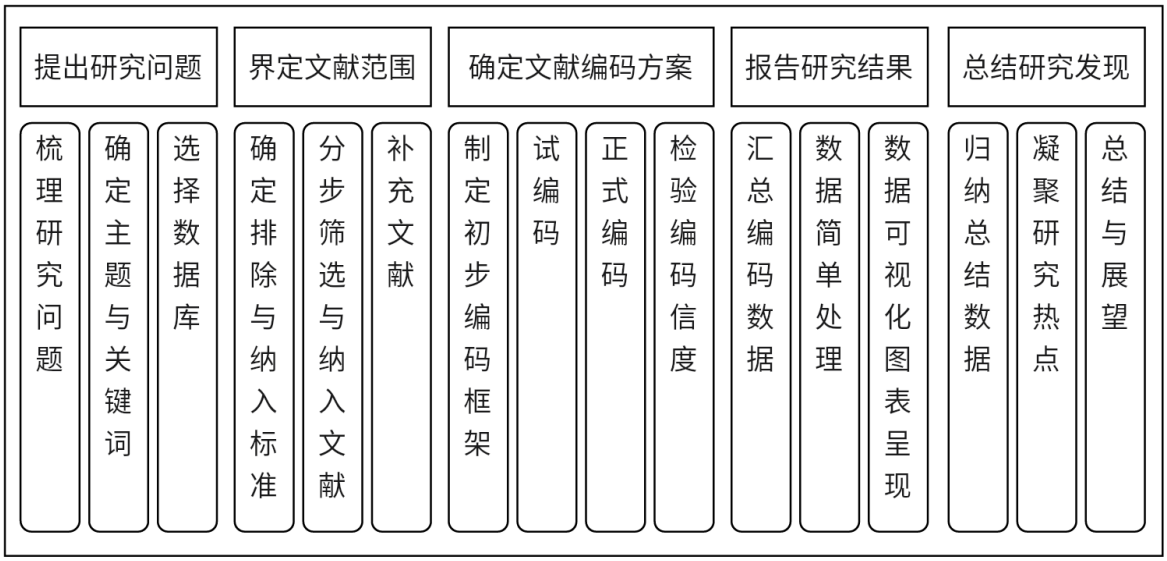


图 1 *系统性文献综述法的步骤流程图*

3.1. 文献检索

为保证样本文献的质量，本研究选取 scopus 数据库以及中国知网等数据库检索文献，以“非正式学习”（informal learning）、“场馆学习”（museum learning）、“K-12”、“学生”（student）、“儿童”（kid/children）、“科技馆”（science museum）、“水族馆”（aquarium）、“动物园”（zoo）、“植物园”（botanical garden）、“自然历史博物馆”（natural history museum）进行配对组合搜索文献，文献检索的时间为 2023 年 10 月到 12 月，将文献发表时间限定为 2003 年至 2023 年，限制研究领域为教育，以确保样本文献检索的全面性与准确性。文献类型为期刊论文，最终检索得到有效文献 65 篇（外文 57 篇、中文 8 篇）（如图 2）。

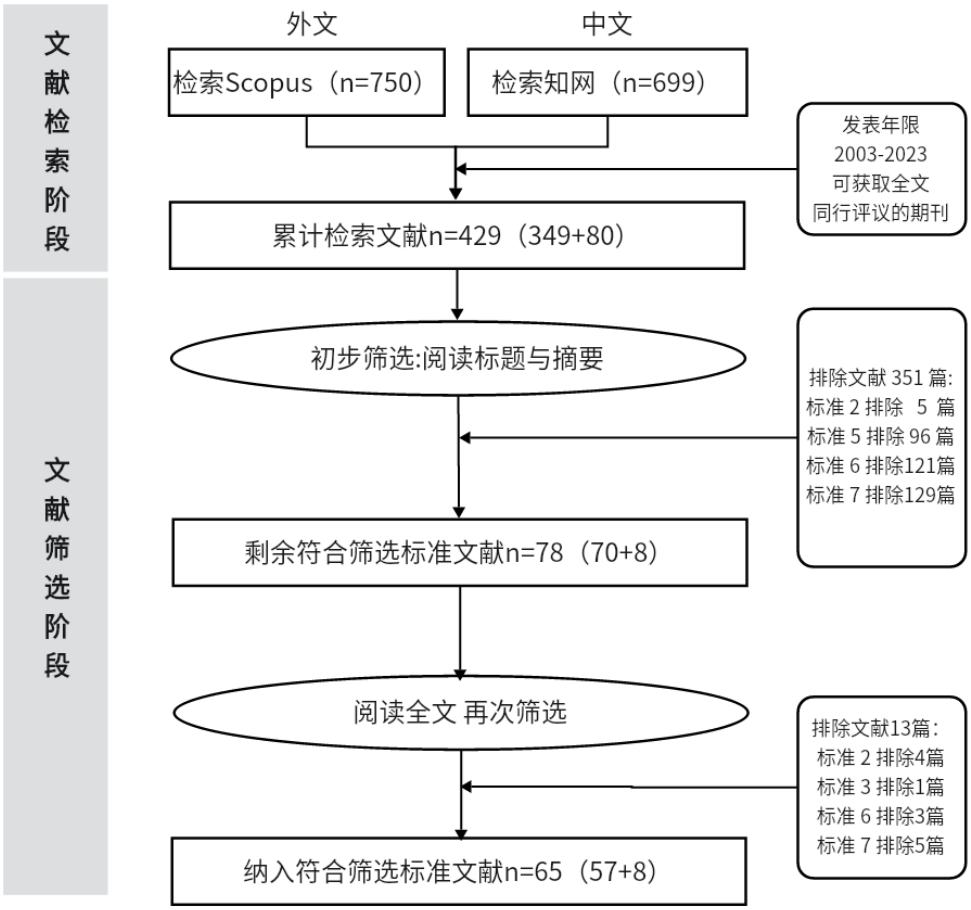


图 2 *文献筛选流程*

3.2. 文献筛选与编码

基于文献筛选标准（表 1），本研究对检索到的文献开展了三轮筛选，同时结合文献阅读以滚雪球方式补充新文献，最终纳入有效样本文献 65 篇（中文 8 篇，英文 57 篇）。研究者在编码前广泛浏览领域内文献，采取混合编码方式，结合季娇、伍新春（2021）学者的科技场馆三维学习效果模型，添加、扩展和提炼编码，最终从研究范式与取径、研究对象、教学设计以及学习效果等方面进行编码。具体的编码框架见图 3。

表 1 *文献筛选标准*

	筛选标准	纳入标准	排除标准
1	文献是否重复	否	是
2	文献是否可全文获取	是	否
3	文献篇幅是否在 3 页及以上	是	否
4	文献是否为期刊文献	是	否
5	文献是否采用实证研究方法	是	否
6	文献主题是否与本研究主题高度相关	是	否
7	文献关注的研究对象是否为 K-12 年龄阶段的学生	是	否

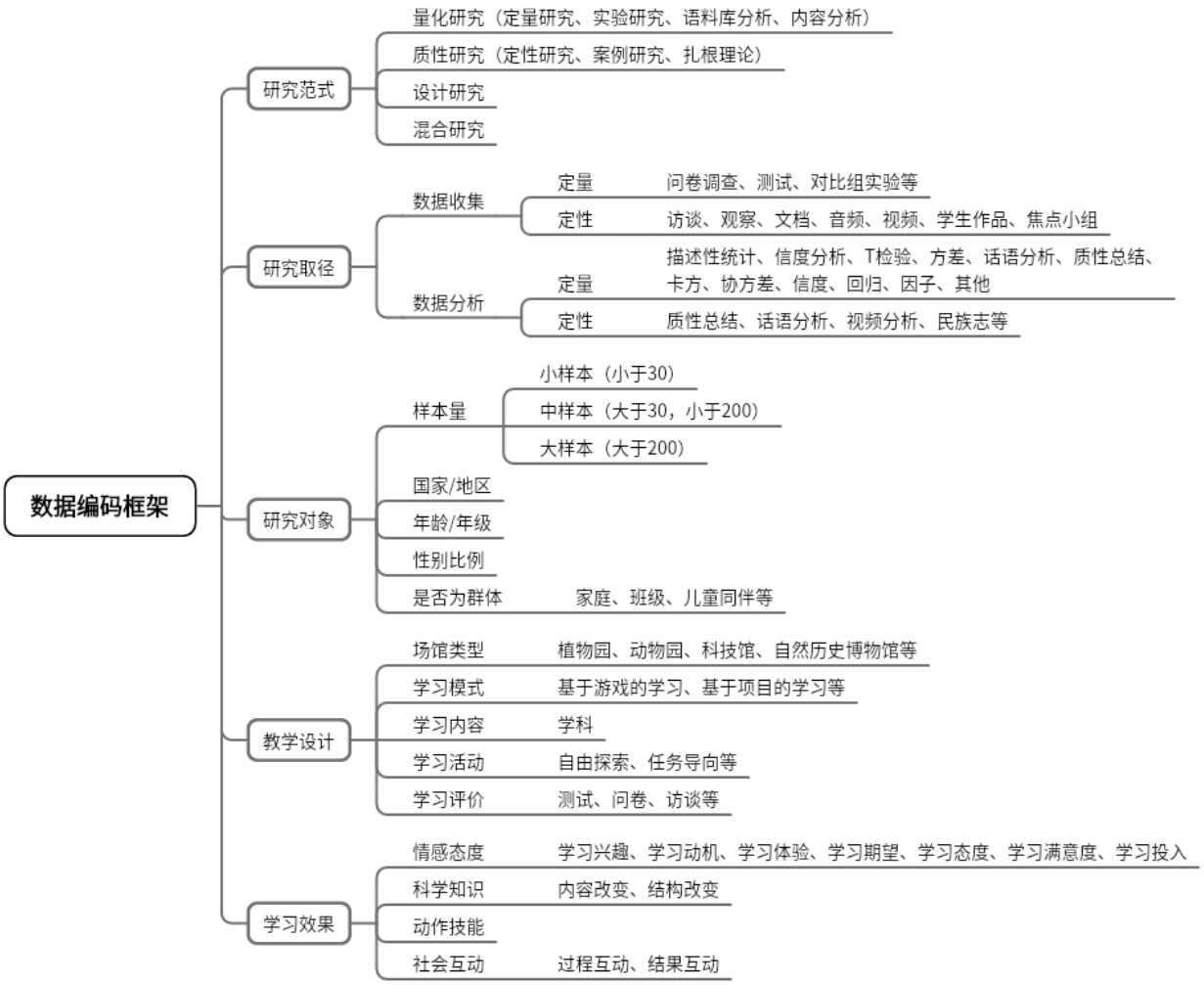


图 3 *数据编码框架*

3.3. 编码信度分析

在第一轮审查结束之后，两位编码人员开会讨论并协商确定纳入文章。接着研究小组在纳入/排除轮次后对具体研究内容进行了编码。为了支持评分者间的可靠性，研究小组开会并讨论了如何对前 10-15 篇文章进行编码。这使研究团队能够更好地理解编码，完善编码表，并进一步审查纳入/排除标准。选择纳入系统性文献综述的研究文章由两名研究小组成员独立编码。与纳入/排除轮次类似，最终讨论达成的协议信度为 100%。

4. 研究结果

4.1. 样本文献发文趋势

本文所选文献发表于 2006 年到 2023 年，其中从图 4 可以看出在 2009 年以前关于 K-12 儿童的非正式学习的实证研究较少（n=4），2010 年到 2012 年研究数量较多，2013 年高质量的实证研究较少；随后在 2016 年达到纳入数量峰值，这可能与《新媒体联盟地平线报告：2016 博物馆版》的发布有关。该报告由美国新媒体联盟与巴尔博亚公园在线协作完成，具有极大影响力。报告指出了最有可能影响今后五年(2015—2019)的新兴技术，具体包括六项主要趋势、六个关键挑战和六种重要技术（Freeman et al.，2016），这为研究者指出了具体多元的研究方向。随后文献数量波动增长，近些年又得到了较高的研究关注，侧重于创新研究，例如移动设备、教育游戏、增强现实等技术在场馆学习中的应用对儿童场馆学习效果的影响。

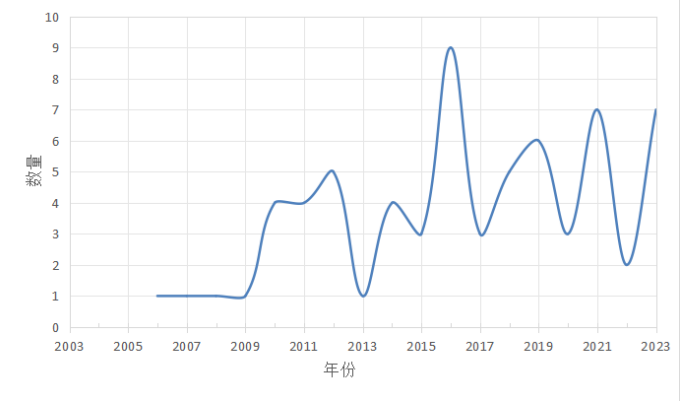


图 4 *纳入文章的发文趋势*

4.2. 场馆学习研究设计

4.2.1. 研究对象

研究地点和整合：纳入的研究来自 19 个国家，6 篇文章未汇报样本分布国家或地区。已发表的研究主要集中在亚洲地区（n=25，38%），其次是欧洲地区（n=16，25%），北美地区（n=11，17%）。样本主要分布在中国台湾地区（n=11），美国（n=10），中国（n=8）以及以色列（n=5），其中 6 篇未报告研究对象所在国家与地区。

样本来源与样本量选择：样本量根据刘清堂等学者的编码分类标准，将 30 以下认定为小样本，30-200 认定为中样本，200 以上认定为大样本（刘清堂等，2022）。纳入的研究小样本（n=18），中样本（n=39）以及大样本（n=9），其中谢娟、伍新春等（2018）学者的一项研究中，共有 1373 名参与者（覆盖学段为小学初中高中），似乎是一个异常值，因为其他审查过的研究一般都在 250 个样本量以下。原因可能是基于中国庞大的人口基数，该研究为国家政府支持的大规模的问卷调查，高样本量也增加了数据的说服力。目前审查的研究倾向于在一两个班级中进行（最小值=5，最大值=1373，中位数=64）。有 28 篇研究没有报告性别比例，大部分研究对象都是性别比例基本相同，个别研究专门研究男生/女生的场馆学习现状。例如在 Maher, Damian 人的研究中样本全为男生（Maher，2015）。在 Archer, Louise 等研究者研究中其中一个案例全为女生（Archer et al.，2021）。

研究情境：这些研究中的大多数是在科学博物馆、科学中心环境中进行的（n=26），其次是具有特色的专题场馆（如天文馆、历史博物馆、艺术馆、植物园与动物园等）进行的。

4.2.2. 数据收集

在分析的 65 项研究中，我们发现了量化研究（28 项）、质性研究（22 项）混合方法研究（9 项）和设计研究（6 项）。在定性研究中，主要是访谈（n=19）、视频（n=21）和观察（n=16）。23 篇研究使用混合方法收集数据，使用问卷（n=16）、观察（n=11）、访谈（n=10）、视频（n=8）或测验（n=6）等 2-5 种方法组合。在两项研究中分析了设备上的日志文件数据。在定量研究中，数据主要通过问卷收集（n=33）。总体而言，问卷调查是最为便捷的数据收集方法，但具有一定的主观性；测验可以直接检测学习效果，却忽略了学生内在学习体验；观察记录和访谈相较于前两种方式，能够深入了解学生的学习过程和学习体验，但需花费较多时间。因此，数据收集的方式应更加趋于多样化、多元化，将质性数据与量化数据相结合，提升研究结果的有效性及可靠性，从而深入探究儿童场馆学习效果。

4.2.3. 数据分析

纳入文献中大部分文献（n=10）运用混合数据分析方法，将定性与定量方法相结合，确保研究结论的正确性与可靠性。大部分文章采用定量的方法（n=33），定量分析方法除了描述性统计分析与信度分析以外，主要运用 T 检验（n=13）、方差分析（n=11）、协方差分析（n=8）。定性分析方法（n=22）包括内容分析、话语分析、视频分析、质性总结等。其中话

语分析（n=9）是运用最多的定性数据分析方法，其次是质性总结（n=7）。从纳入文章的整体研究思路来看，目前研究者以及开始重视到定性数据与定量分析相结合的方法，让非正式学习的效果得到更加全面与科学的解释。另外在研究方法上，Zhang, Jia 等学者运用滞后序列分析法识别学习行为模式，研究使用增强现实 worksheet 与使用传统纸质 worksheet 两种条件下学习者的学习效果比较，并得出使用增强现实 worksheet 比使用纸质练习题的学习效果好，并能提高学习者的社会互动（Zhang et al.，2023）。

4.3. K-12 学生场馆教学法

4.3.1. 教学设计

学习内容：本综述表明，纳入研究主要来自自然科学领域，尤其是生物（n=12）和物理学（n=9）。其他领域如地理、艺术、历史或音乐，研究较少。

学习活动：学习活动多以小组形式进行，访问时间平均为半小时至一个小时。研究对象往往是儿童群体（n=14），常见群体为家庭。例如，国内学者王婷、郑旭东等学者研究场馆学习中家庭群体行为特征与行为模式（郑旭东和王婷，2017；郑旭东，2015）；伍新春等学者研究场馆学习中亲子互动类型对学习的影响（伍新春、李长丽、曾筝和季姣，2012；谢娟、伍新春和季娇，2018）。

4.3.2. 技术策略

其中技术支持方面，目前已有研究的技术有 worksheet、概念图、移动设备、虚拟现实与增强现实技术、教育游戏等。目前的研究现状往往是多种信息技术的结合，例如 Liang, Hsin-Yi、Hsu, Tien-Yu 等学者采用基于移动概念图的提问方法（ARG-CMQ）的 ARG（Alternate reality game）来促进儿童在不同真实环境（校园和科学博物馆）中进行有意义的探究式学习（Liang et al., 2021）。此外，Hsu, Tien-Yu 学者团队开发一款 AR 教育游戏应用程序 “CoboChild”，开展一系列的研究探讨儿童在虚拟游戏世界和博物馆物理空间混合的环境中学习的影响因素，以促进儿童的持续学习和再访动机（Hsu et al.，2018）。

4.4. K-12 学生场馆学习效果

场馆学习效果的分类体系繁多，本文根据前人的研究，将纳入文献研究中的学习效果分成四类：科学知识、情感态度、动作技能与社会互动。编码结果表明评估最多的是科学知识与情感态度方面的学习效果，其次是社会互动（n=35），其中 45 项研究探究了科学知识层面的学习效果，42 项研究调查了情感态度层面的学习效果。

科学知识：科学知识维度是衡量场馆学习效果的基本且相对显性的维度。科学知识维度可以理解为认知层面，具体表现为以学科知识为导向和以科学概念学习为导向的学习。Storksdieck、Ellenbogen 和 Heimlich 提出，参观者在参观过程中也许并没有获得新的知识，而仅仅是对先前知识的再次确认(reaffirm)（Storksdieck et al. , 2005）。也就是说，通过参观场馆，参观者可能在具体的科学知识掌握量上没有增加，但他们能把先前散落的、具体的知识点进行联系和重组，形成知识网络体系。对于儿童而言，可以考察场馆参观是否促进了儿童对某一学科内或不同学科之间的知识建立联系，是否促进了儿童对科技场馆的情境性知识与个体生活中的知识联系，而这种知识的联结反映了博物馆参观对自己、家庭和社会等形成的更加复杂和完备的理解。

情感态度：学习效果除了认知层面，还应关注非认知层面。情感态度维度是场馆学习中教育的价值维度，也是研究者常关注的维度，具体包括学习体验、学习动机、学习态度、学习期望和自我效能感等多方面。其中，学习体验是研究最多的维度(n=18)。

动作技能：仅 4 篇研究评估了动作技能。尽管这些研究数量有限，但它们为动作技能的发展和评估提供了重要的见解。美国“21 世纪技能联盟”制定了《21 世纪技能框架》，强调培养学生的信息和通讯技能，思维和问题解决技能，人际和自我指导技能以及使用 21 世纪工具

(例如信息和通讯技术)的技能（2009）。纳入的四篇研究主要评估学生的问题解决能力、科学观察技能（Knight & Davies，2016；Land et al.，2020）以及动手实践能力（de Freitas & Bentley，2012）。科技场馆的展品设计是基于具体事件与实际问题的（伍新春等.2009），根据情境非正式学习提供了机会去解决现实生活中的问题，培养了个体的问题解决能力和创新思维。观察能力是场馆学习中的关键学习目标，因为参观场馆的过程中不仅是知识获取的过程也是只是构建的过程（Davies et al.，2008），观察能力可以用于产生问题与假设，捕捉信息与收集数据，同时观察还有助于回忆重要细节并支持解决问题（Horn et al.，2016；Oguz-Unver & Yurumezoglu, 2009）。

社会互动：社会互动通常被视为人与人之间的互动。在博物馆的背景下，这种互动主要发生在学生与学生之间、学生和讲解员以及家庭成员之间。然而，博物馆里的每件展品都包含着社会学的信息（Botelho & Morais，2006）。基于情境认知理论的观点，学习是学习者给予一定的社会和文化背景，在与环境的互动中建构自己的意义和身份的活动，因此社会互动应该包含人与展品的互动。展览传递的信息可能会被学生接收，也可能不会被学生接收。因此判断学生-展品之间的社会互动时，需要根据具体的研究情境与研究设计来考虑。

总之，2 项研究评估了 4 个方面的学习效果，10 项研究评估了 3 个方面的学习效果，34 项研究评估了 2 个方面的学习效果。

5. 讨论

本文重点从研究取径、教学设计、技术策略、学习效果四个方面对 65 篇高质量的高度相关的期刊文章进行了编码分析，并以系统的观点梳理了过去 20 年 K-12 儿童的场馆学习研究。研究表明，目前儿童的场馆学习研究主要集中在科学类场馆，主要学习内容为生物与物理。在研究设计上，研究者们大多采用小组的方式让儿童学习者开展基于展品/展览的学习、项目式学习或基于游戏的学习，常见的研究对象群体为家庭。其次研究者们逐渐意识到要采用定性与定量相结合的数据收集与分析方法，提高实验的准确性。在研究要素上，纳入的文章重点关注科学知识与情感态度维度的学习效果，在动作技能方面关注最少。

本系统文献综述的局限性在于以下两点。首先，检索仅限于 Scopus 和中国知网两个数据库，可能存在查找文献不全的问题。文献检索于 2023 年 12 月完成，因此较新的文章未纳入文献列表。另外，本研究在文献检索过程中并未发现幼儿园阶段的儿童的相关研究。其次，随着场馆建设的类型增加，例如画廊、大学科学研究中心等，增加了文献编码的难度。在这次搜索中，我们只包括一些非常突出的非正式学习场所，如博物馆、水族馆、植物园和动物园。未来的研究可使用更详细的关键词，包括与各种非正式学习场所相关的术语进行检索。尽管文献综述中收录的文章数量有限，但全文严格根据系统性文献综述步骤流程进行，根据信效度分析，结果可靠，为场馆工作者以及相关研究人员提供了重要的见解。

参考文献

季娇，伍新春和青紫馨(2017)。 非正式学习：学习科学研究的生长点。*北京师范大学学报(社会科学版)，(01), 74-82*。https://doi.org/10.3969/j.issn.1002-0209.2017.01.007

季娇，伍新春(2021)。何为“好”的博物馆教育——博物馆学习效果的多元维度及指标分析。*教育研究与实验, (01), 87-92*。

李睿和刘颖芳(2023)。 我国场馆教育近十五年研究热点和现状——基于 CiteSpace 的可视化分析。 *自然科学博物馆研究, 8(05), 50-58*。https://doi.org/10.19628/j.cnki.jnsmr.2023.05.006

刘清堂，马晶晶，余舒凡等(2022)。增强现实技术对场馆学习效果影响的研究及展望。*现代远距离教育, (03), 3-12*。https://doi.org/10.13927/j.cnki.yuan.20220428.001

骆晓红(2016)。智慧博物馆的发展路径探析。 *东南文化, (06), 107-112*。
王秀江，王瑛(2019)。儿童博物馆展项教育功能实现策略研究——以中国儿童中心老牛儿童探索馆为例。*中国博物馆, (03), 87-93*。
王妍莉，杨改学，王娟等。(2011). 基于内容分析法的非正式学习国内研究综述. *远程教育杂志, 29(04), 71-76.* https://doi.org/10.15881/j.cnki.cn33-1304/g4.2011.04.016.
伍新春，李长丽，曾筝 & 季娇. (2012). 科技场馆中的亲子互动类型及其对学习效果的影响. *教育研究与实验, (06), 88-92.*
谢娟，伍新春和季娇(2018)。科技场馆"第二课堂"育人价值实现路径——基于我国城市中小学生的实证调查。*中国教育学刊, (09), 68-73*。
许玮和张剑平 (2016)。 数字化场馆环境特征要素调查研究。*开放教育研究, 22(04), 95-102*。https://doi.org/10.13966/j.cnki.kfjyyj.2016.04.012
叶贝和董雁(2022)。国内外非正式学习研究回顾与启示——基于 CiteSpace 的可视化分析。*成人教育, 42(06), 1-7*。
余胜泉和毛芳(2005)。非正式学习——e-Learning 研究与实践的新领域。*电化教育研究, (10), 19-24*。
张美霞(2017)。新媒体技术支持下的场馆建设与场馆学习——以现代教育技术博物馆为例。*中国电化教育, (02), 20-24*。
郑旭东和王婷 (2017)。 场馆学习中的家庭行为模式：表征元素、形成机制与基本类型。 *中国电化教育, (09), 1-5*。
郑旭东(2015)。从博物馆教育到场馆学习的演进：历史与逻辑。*现代教育技术, 25(02), 5-11*。
Anderson, D., & Thomas, G.P. (2014). "Prospecting for Metacognition" in a Science Museum: A Metaphor Reflecting Hermeneutic Inquiry. *Issues in Educational Research, 24, 1-20.*
Archer, L., Godec, S., Calabrese Barton, A., Dawson, E., Mau, A., & Patel, U. (2021). Changing the field: A Bourdieusian analysis of educational practices that support equitable outcomes among minoritized youth on two informal science learning programs. *Science Education, 105(1), 166-203.* https://doi.org/10.1002/sce.21602
Barbieri, L., Bruno, F., & Muzzupappa, M. (2017). Virtual museum system evaluation through user studies. *Journal of Cultural Heritage, 26, 101-108.* https://doi.org/10.1016/j.culher.2017.02.005
Botelho, A., & Morais, A. M. (2006). Students-exhibits interaction at a science center. *Journal of Research in Science Teaching, 43(10), 987-1018.* https://doi.org/10.1002/tea.20135
Davies, R.S., Sprague, C. R., & New, C.M. (2008). Integrating technology into a science classroom. In Sunal, D.W., Wright, E. & Sundberg C., *The impact of the laboratory and technology on learning and teaching science K-16* (pp. 207 – 237). Tuscaloosa, AL: Information Age Publishing.
de Freitas, E., & Bentley, S. J. (2012). Material encounters with mathematics: The case for museum-based cross-curricular integration. *International Journal of Educational Research*, 55, 36-47. https://doi.org/10.1016/j.ijer.2012.08.003
Falk, J. H., Dierking, C. S., et al. (2010). Interactives and visitor learning. *Curator*, 47(2), 171-198.
Freeman, A., Adams Becker, S., Cummins, M., McKelroy, E., Giesinger, C. & Yuhnke, B. (2016). NMC Horizon Report: 2016 Museum Edition. Austin, Texas: The New Media Consortium. Retrieved April 14, 2023 from https://www.learntechlib.org/p/182007/.

Han, Y., Yang, F., Xiong, S., & Hu, M. (2020). Informal learning research based on mobile portable devices. *2020 International Conference on Information Science and Education (ICISE-IE),* 334-337. https://doi.org/10.1109/ICISE51755.2020.00080

He, X., Deng, Y., Liu, J., Sun, G., Xiong, J., & Xiao, Y. (2023). The chain mediating role of perceived family support for formal and informal science learning in the association between family socioeconomic status and informal science learning experiences. *Journal of Baltic Science Education, 22(2),* 232. https://doi.org/10.33225/jbse/23.22.232

Horn, M. S., Phillips, B. C., Evans, E. M., Block, F., & Shen, C. (2016). Visualizing biological data in museums: Visitor learning with an interactive tree of life exhibit. *Journal of Research in Science Teaching, 53(6),* 895-918. https://doi.org/10.1002/tea.21318

Hsu, T.-Y., Liang, H., Chiou, C.-K., & Tseng, J.C.R. (2018). CoboChild: A blended mobile game-based learning service for children in museum contexts. *Data Technologies and Applications, 52(3),* 294-312. https://doi.org/10.1108/DTA-05-2016-0043

Knight, K., & Davies, R. S. (2016). Using a mobile dichotomous key ipad application as a scaffolding tool in a museum setting. *Interactive Learning Environments, 24(1/4),* 814-828. https://doi.org/10.1080/10494820.2014.924532

Land, S. M., Zimmerman, H. T., Millet, C., & Choi, G. W. (2020). Supporting observing-on-the-move with proximity-based technology: Designing for children’s scientific observations outdoors. *Visitor Studies, 23(2), 182-204.* https://doi.org/10.1080/10645578.2020.1773720

Liang, H. Y. , Hsu, T. Y. , & Hwang, G. J. . (2021). Promoting children's inquiry performances in alternate reality games: a mobile concept mapping‐based questioning approach. *British Journal of Educational Technology, 52(5),* 2000-2019.

Oguz-Unver, A., & Yurumezoglu, K. (2009). A teaching strategy for developing the power of observation in science education. *Ondokuz Mayis University Journal of Education, 28, 105‐119.*

Park, K., Li, H., & Luo, N. (2021). Key issues on informal learning in the 21st century: A text mining-based literature review. *International Journal of Emerging Technologies in Learning, 16(17), 4‐18.* https://doi.org/10.3991/ijet.v16i17.23663

Piscitelli, B., & Anderson, D. (2001). Young children's perspectives of museum settings and experiences. *Museum Management and Curatorship, 19(3), 269-282*. https://doi.org/10.1080/09647770100401903

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747-770. https://doi.org/10.1146/annurev-psych-010418-102803

Silver, L., & Kisiel, J. (2008). Comparing visitors' conceptions of evolution: Examining understanding outside the United States. *Visitor Studies, 11(1),* 41-54.

Storksdieck, M., Ellenbogen, K., & Heimlich, J. E. (2005). Changing minds? Reassessing outcomes in free‐choice environmental education. *Environmental Education Research, 11(3),* 353‐369. https://doi.org/10.1080/13504620500081228

Weinstein, M., Whitesell, E. R., & Schwartz, A. E. (2014). Museums, zoos, and gardens: How formal-informal partnerships can impact urban students' performance in science. *Evaluation Review, 38(6), 514-545.*

Zhang, J., Huang, Y.-T., Liu, T.-C., Sung, Y.-T., & Chang, K.-E. (2023). Augmented reality worksheets in field trip learning. *Interactive Learning Environments, 31(1),* 4-21. https://doi.org/10.1080/10494820.2020.1758728

基于心流理论的虚拟现实学习模式会促进虚拟现实学习吗？

Will the Virtual Reality Learning Model based on Flow Theory Facilitate Virtual Reality Learning?

梁友明[1]，王陆[2]
[1]首都师范大学 教育学院
* liangyouming01@163.com

【摘要】 本研究旨在探讨基于心流理论的虚拟现实学习模式（VRFT）是否会影响学生在沉浸式虚拟现实环境中的学习。对开展过虚拟现实教学的 12 位中学一线教师进行深度访谈，使用扎根理论方法进行分析，构建了 VRFT 学习模式，并将该模式应用到实际的物理课堂教学中。基于 86 名学生的准实验研究结果表明：(1) VRFT 学习模式支持下，大部分学习者在虚拟现实课堂中顺利地进入到“心流”状态；(2) VRFT 学习模式相比起普通的沉浸式虚拟现实课堂，更易提高学生的学习成绩。研究为进一步优化虚拟现实技术应用于课堂教学有一定的参考价值。

【关键词】 虚拟现实；心流理论；学习模式；初中物理

Abstract: *The purpose of this study is to investigate whether a virtual reality learning model based on the theory of Flow (VRFT) affects students' learning in an immersive virtual reality environment. In-depth interviews were conducted with 12 front-line teachers in secondary schools who had conducted virtual reality teaching, analyzed using a rooted theory approach, constructed a VRFT learning model, and applied the model to actual middle school physical classroom teaching. The results of a quasi-experimental study based on 86 students show that: (1) most of the learners in the virtual reality classroom with the support of the VRFT learning model smoothly enter the state of "flow" states; (2) the VRFT learning model is more likely to improve the student's academic performance compared with the ordinary immersive virtual reality classroom. The research has a certain reference value for further optimizing the application of virtual reality technology in classroom teaching.*

Keywords: Virtual Reality; Flow Theory; Learning Models; Middle School Physics

1. 前言

如何巧用技术提升学生在课堂上的专注度和学习成绩，是教育领域关注的热点话题。近年来，随着计算机图形学、计算机视觉、计算机仿真技术等多种前沿技术的发展，以及国家层面“自上而下”的政策文件引领，使得以虚拟现实（Virtual Reality，简称“VR”）为代表的系列创新技术引入教学的探究逐渐增多。但众多研究表明，虚拟现实融入课堂仍面临诸多挑战：在虚拟环境中，课堂教学的难度增大，虚拟实验可能导致学习过程中的交流减少（Wu, 2018），师生间的情感交流易流失（祝士明，2019）；此外，虚拟环境下的教学活动设计、学习过程交互、学习效果评价等都是亟待解决的问题（李勋祥和游立雪，2017），这些问题若不能得到很好的解决，势必影响 VR 教学的实用性及普及性。归根结底，这是新技术与传统课堂融合所面临的不适切问题，迫切需要对虚拟现实技术融入课堂的学习模式进行探究。心流是虚拟学习环境中评估学习者学习体验的重要指标，心流能使学习者更专注于学习材料本身，全身心投入当前学习任务（刘哲雨、周继慧和周加仙，2022），进而促进目标的达成，因而学习者在学习时进入“心流状态”是重要的。本研究在心流理论指导下，结合对使用过沉浸式虚拟现实

系统的一线教师的访谈，采用“自上而下 + 自下而上”相结合——即“理论研究 + 访谈”的方式，构建了心流理论视域下的沉浸式虚拟现实课堂学习模式（An Immersive Virtual Reality Classroom Learning Model from the Perspective of Flow Theory，简称“VRFT”），并通过实证研究检验该模式应用于初中物理课堂的有效性，试图为 VR 技术融入课堂教学提供借鉴和参考价值。

2. 心流理论视域下的沉浸式虚拟现实学习模型构建

依据心流理论指导以及教学设计的原理与技巧，结合对一线教师的访谈分析，构建了 VRFT 学习模式，如图 1 所示。该学习模式分为心流前因阶段、心流体验阶段及心流结果阶段（Hoffman，1996），强调在虚拟现实介入课堂的基础上，充分尊重学生的主体性。该模式基本遵循“设疑—讲解—互动—合作—实验—汇报及点评—评估”的学习环节或流程进行。

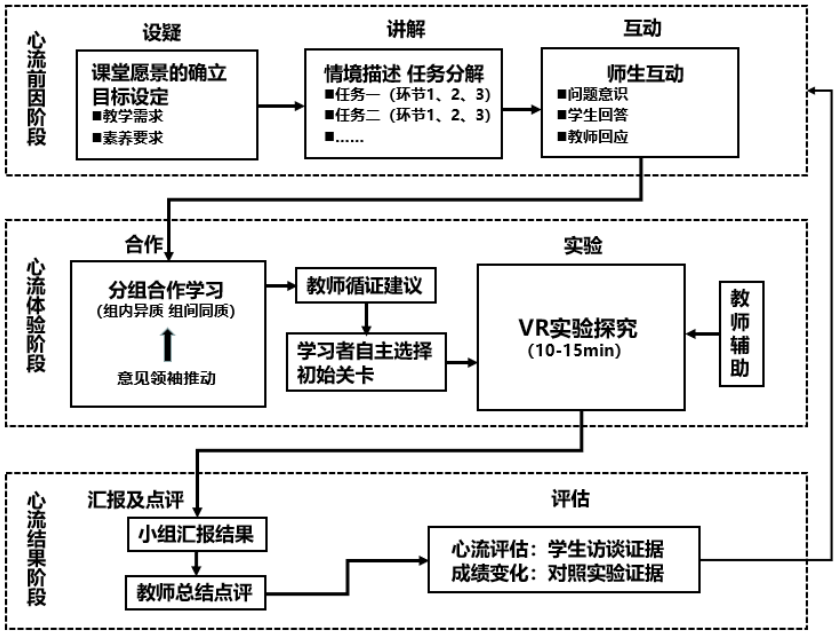


图 1　心流理论视域下的虚拟现实课堂学习模式构建

心流前因阶段以学习者产生心流体验的四个前提条件为导向：即清晰的目标、即时反馈、潜在控制感以及行动与意识的融合。这些因素共同确立了学习者能否进入心流体验。随着虚拟现实实验的介入，学习者进入心流体验阶段，表现为全神贯注、自我意识的丧失、时间扭曲感以及在场感。心流结果阶段包括自我意识丧失、时间体验失真和体验活动本身的价值等（刘哲雨，周继慧和周加仙，2022）。本研究首先通过对学生进行半结构化访谈，以评价学生在课堂中的心流体验；其次，在一段时间的准实验研究后，评估 VRFT 学习模式是否帮助学生提高成绩。

3. 心流理论指导下的 VR 教学实验设计与结果分析

3.1. 研究目的及问题

本研究的目的是验证 VRFT 学习模式介入课堂的有效性，具体的研究问题为：（1）介入 VRFT 学习模式后，与传统沉浸式虚拟现实课堂相比较，是否更有助于学习者进入心流状态？（2）介入 VRFT 学习模式后，与传统沉浸式虚拟现实课堂相比较，是否更有利于学生学习成绩的提升？

3.2. 实验设计及对象

本研究选取了两个平行班开展准实验研究。这两个班级有以下共同特征：（1）两班人数相等，均为 43 人，且男女生人数无明显差异；（2）两班的物理授课教师均为同一位老师。不

同的是，平行班采用传统的学习模式展开沉浸式虚拟现实课堂教学，即教学过程未严格设计，仅在需要时使用沉浸式虚拟现实设备。实验班则采取了心流理论视域下的沉浸式虚拟现实课堂模式进行教学。实验流程如图 2 所示。

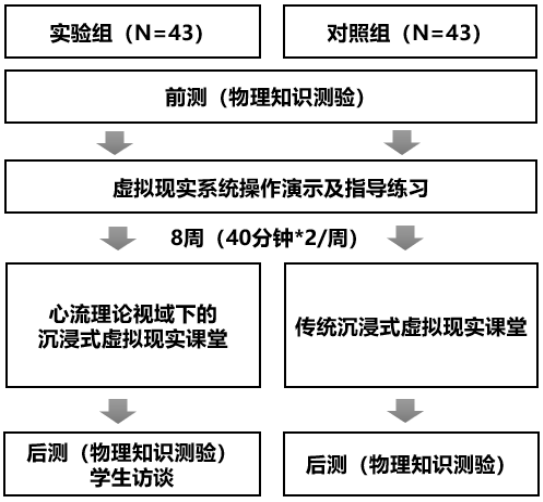


图 2　实验流程图

实验开始前，对两个班级进行考试前测，并对惠普虚拟现实系统进行说明，示范头盔的正确佩戴方法及手柄的连接和使用方式。每次课程结束后，随机抽取实验班的 1-2 名学生进行访谈，系列课程共抽取了 15 名学生，目的是验证实验班的学生在课堂中是否进入“心流状态”；两个月的准实验研究结束后，对两个班级进行考试后测。

3.3. 数据收集方法

心流是由内在动机（Intrinsic Motivation）决定的，当前已有众多测量心流的量表（Pearce，Ainley & Howard，2005；刘锦源和曹树金，2017），也有研究使用访谈法，研究心流理论如何影响学生进行虚拟游戏的过程（Akman & Çakır，2019）。本研究采用半结构化访谈的方式收集学生在学习过程中的心流体验证据。其次，通过实验前后分别对实验班及对照班进行成绩测试，以获得成绩证据。

3.4. 实验/教学内容设计

本研究共选择了八个实验，均为人教版物理教科书九年级第十五章《电流与电路》、第十六章《电压电阻》的内容，包括《摩擦起电》《电荷间的相互作用》《用滑动变阻器改变灯泡亮度》等，并开展了为期 2 个月的教学实践。这八个虚拟实验的开发，排除了真实环境下环境的不确定性、省去了实验器材的准备、使抽象的知识形象化，能以便捷的形式让实验顺利进行，教师及学生在课前只需熟练掌握 VR 设备的操作及使用，更专注于课堂内容本身。

3.5. 实验结果分析

3.5.1 学生访谈结果分析

实验结束后，通过随机抽样原则对实验班的 15 名学生进行访谈，得知本课创建的心流理论视域下的沉浸式虚拟现实课堂学习模式使大部分学生进入了心流状态，判断的依据是学生表达的感受与 Csikszentmihalyi 提出的学习者进入心流的八个主要特征是基本吻合的。例如有学生提到：“我分不清这是游戏还是 s 学习，感觉实验中的那只手就是我自己的手（自我意识的丧失）”。可以确定的是，VRFT 学习模式使绝大部分学生都进入了心流状态。

3.5.2 成绩分析

为期两个月的教学前后，对实验班和对照班分别进行了前测和后测。对实验班和对照班的成绩进行独立样本 T 检验分析，所得结果如下表 1 所示。

表 1　独立样本 T 检验前后测结果

前测	平均值	t 值	Sig（双尾）	后测	平均值	t 值	Sig（双尾）

实验班前测	68.62	-.553	0.717	实验班后测	74.58	3.403	0.002
对照班前测	65.77			对照班后测	68.25		

由表 2 可知，实验班与对照班在前测时，成绩不具有统计学意义上的差异（sig = 0.717 > 0.05），说明两个班级在实验前的初始条件是相似的。两章内容授课结束后，教师发放了后测问卷，并对成绩数据进行独立样本 T 检验，可知实验班的成绩远高于对照班，且 t = 3.403（sig = 0.002 < 0.01），表明通过一段时间的 VR 教学，学生的后测成绩呈现出了显著差异。

4. 研究结论与展望

本研究通过文献调研及对教师的访谈，构建了 VRFT 学习模式并开展准实验研究，以学生的访谈证据及成绩来验证学习模式的有效性，得出以下结论：(1) VRFT 学习模式能显著提升学习者的学习成绩。众多研究发现，心流理论指导下的虚拟现实教学在提升学生的学习成绩方面效果显著（Tan & Waugh，2013；李斐，2019；刘哲雨，刘宇晶和周继慧，2022），这与本研究的结论基本一致；(2) VRFT 学习模式能够提升学生的学习兴趣。对质性研究数据的分析结果表明，沉浸式虚拟课堂对于学生具有较大的吸引力，也有利于学习者进入心流状态。虚拟现实系统给学生营造了一种“游戏化学习”的意境，也为学生提供了更大的物理实验的想象空间。本研究也存在一定不足，目前研究的样本来自于特定的地区和学科（初中物理课堂），未来研究将扩大样本范围，包括不同年龄层、地理位置及学科领域，以增强研究结果的普遍性和适用性。

参考文献

李斐.（2019）.基于心流体验的 VR 历史学习软件设计与开发[D].陕西：陕西师范大学.

李勋祥和游立雪.（2017）.VR 时代开展实践教学的机遇、挑战及对策[J].现代教育术, 27(07):116-120.

刘哲雨、周继慧和周加仙.（2022）.教育神经科学视角下促进心流体验的智慧教学活动设计[J].现代教育技术, 32(07):14-21.

刘哲雨、刘宇晶和周继慧.（2022）.桌面虚拟现实环境中自我效能感如何影响学习结果——基于心流体验的中介作用[J].远程教育杂志,2022,40(04):55-64.

刘锦源和曹树金.（2017）.心流理论视角下信息检索体验测量与分析[J].图书情报工作, 61(08):67-73.

祝士明和陈静潇.(2019).虚拟现实学习环境的作用、挑战以及应对策略[J].现代教育技术, 29(02):39-45.

Akman, E., & Çakır R.(2019). Pupils' Opinions on an Educational Virtual Reality Game in Terms of Flow Experience[J]. International Journal of Emerging Technologies in Learning,2019,14(15), 121-137.

Hoffman, D.L.，& Novak, T.P.(1996). Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations[J]. Journal of Marketing, 60(3), 50–68.

PEARCE,J.M, AINLEY,M, & HOWARD,S. (2005). The ebb and flow of online learning[J].Computers in human behavior，21（5）：745-771.

Tan,S, &Waugh R.(2013). Use of virtual-reality in teaching and learning molecular biology. In: Cai Y. (eds) 3D Immersive and Interactive Learning[M]. Singapore: Springer.

Wu， J.(2018). A Space Design Teaching Model Using Virtual Simulation Technology[J]. International Journal of Emerging Technologies in Learning. 13(5), 163-175.

穿戴式虛擬實境結合合作概念圖即時互動回饋系統應用於中學英語閱讀教學：認知負荷對學習成效之影響

Wearable Virtual Reality Integrated with Collaborative Concept Maps IRS for English-Reading Instruction in Junior High School: Effect of Cognitive Load on Learning Performance

游師柔[1]，盧瑩真[2]，孫之元[3*]
[1, 2, 3]國立陽明交通大學
[*] jerrysun@nycu.edu.tw

【摘要】本研究探討穿戴式 VR 結合合作概念圖 IRS 對認知負荷預測英語閱讀學習成效的影響。本研究採用準實驗研究法，103 名國中生分為概念圖 IRS 組（*n* = 33）使用傳統講述式教學完成概念圖、合作概念圖 IRS 組（*n* = 34）利用合作學習完成概念圖，以及穿戴式 VR 合作概念圖 IRS 組（*n* = 36）利用穿戴式 VR 和合作學習完成概念圖，最後三組皆採用 IRS 自我檢核概念圖內容。研究工具包含認知負荷量表、先備知識測驗與學習成效測驗。研究結果顯示穿戴式 VR 合作概念圖 IRS 組由內在認知負荷負向預測學習成效。本研究建議教學者可依學習者的認知負荷狀態選擇教學設計提升學習成效。

【關鍵詞】 穿戴式虛擬實境；合作學習；即時回饋系統；認知負荷；英文閱讀教學

Abstract: *This study examined the effect of wearable VR integrated with collaborative concept maps IRS on the predictive power between cognitive load and learning achievement. The study employed a quasi-experimental design, dividing 103 high school students into three groups: the C-IRS (n = 33), the CC-IRS (n = 34), and the WVRCC-IRS (n = 36) groups. Traditional teaching, collaborative learning, and wearable VR with collaborative learning were employed to complete concept maps. The IRS was used to confirm the concept maps. Participants completed a cognitive load scale, a prior-knowledge test, and a learning achievement test. The findings indicated that the intrinsic cognitive load negatively predicted learning achievement in the WVRCC-IRS group. The study suggests instructors consider learners' cognitive load to enhance learning achievement.*

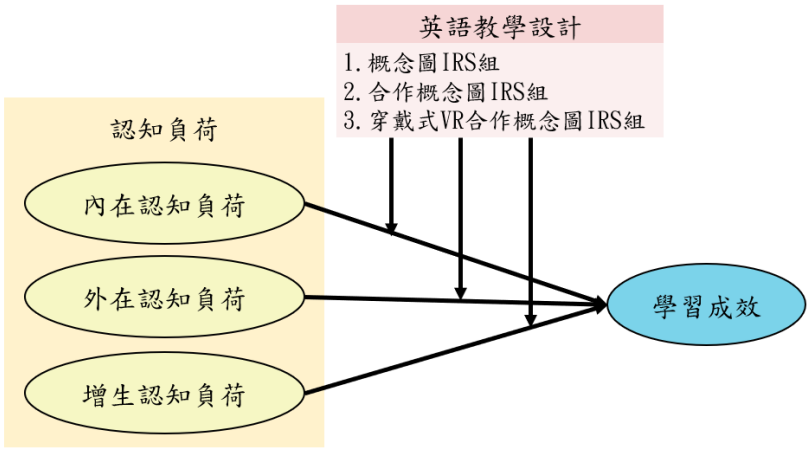
Keywords: Wearable virtual reality, collaborative learning, Interactive Response System, cognitive load, English-reading instruction

1. 前言

英語能力和認知發展會影響學習者對閱讀內容的選擇，過去研究針對中學階段探討不同英語閱讀教學策略，期望促進閱讀學習成效（Sun, 2023），尤其東亞地區中學生之學習動機和學習成效的平衡策略需進一步探討（Nishimura & Sakurai, 2017）。當教學設計不良時，認知處理過程超出工作記憶容量，便會產生外在認知負荷（Sweller et al., 2011），對學習造成負面的影響。經由過去研究已知概念圖有助於組織知識和聚焦學習問題（Novak & Cañas, 2007），即時回饋系統（Interactive Response System, IRS）提供學習者與教學者的互動對話機進而增進學習成效（Wang & Tahir, 2020），以及異質性分組能促進不同能力學習者互相協助提升學習效果（Zamani, 2016）。此外，虛擬實境（Virtual Reality, VR）與穿戴式設備結合的穿戴式 VR 能提供深度沉浸感，學習者能將知識遷移至真實世界（Yu et al., 2021）。雖然已知概念圖、IRS、合作學習與穿戴式 VR 由不同層面促進英語學習，但將其各別優點整合的教

學設計是否能改變學習者的認知處理進而影響學習成效需深入了解。因此，本研究目的為探討穿戴式虛擬實境結合合作概念圖即時互動回饋系統應用於中學英語閱讀教學，對於認知負荷預測學習成效之影響，本研究問題為「概念圖 IRS、合作概念圖 IRS 和穿戴式虛擬實境合作概念圖 IRS 之三種不同英語教學設計，是否為國中生的認知負荷（內在認知負荷、外在認知負荷、增生認知負荷）預測學習成效之調節變項？」，圖 1 為研究架構圖。

圖 1 研究架構



2. 研究方法

2.1. 研究對象

本研究採用準實驗研究法，研究對象為國民中學七年級學生，在國小階段已有英語學習經驗並具備國小基本單字量 300 字和基本英文短文閱讀能力，受試者人數共 113 人，有效樣本數為 103 人（91.15%），男生 52（50.49%）人，女生 51（49.51%）人，平均年齡為 12.89（*SD* = 0.39），由同一位英語教師授課，依照班級分為三組：概念圖 IRS 組 33 人，合作概念圖 IRS 組 34 人，穿戴式虛擬實境合作概念圖 IRS 組 36 人，其中合作概念圖 IRS 組以及穿戴虛擬實境合作概念圖 IRS 組採取分組教學，分組方式為 5-6 人為一組。

2.2. 研究流程

本研究參與者 103 位七年級學生在實驗開始前，需閱讀與簽署實驗同意書，待確認學生理解實驗內容並簽署同意書後，請學生完成先備知識測驗，約 10 分鐘，隨後依照班級分為概念圖 IRS 組、合作概念圖 IRS 組和穿戴式虛擬實境合作概念圖 IRS 組進行英語閱讀教學活動，三組在教師教學難度較高的字彙後，概念圖 IRS 組需完成學習單並獨自繪製文章結構概念圖。合作概念圖 IRS 組採用拼圖二代法，回答學習單問題與完成小組概念圖。穿戴式虛擬實境合作概念圖 IRS 組在分組前，先透過穿戴式 VR 觀看閱讀文章相關影片介紹，後續學習步驟如同合作概念圖 IRS 組共同完成小組概念圖。三組完成概念圖後，教學者皆會運用 IRS 檢視概念圖完成度並了解受試者的理解程度，教學活動約 40 分鐘。教學活動結束後，請學生填寫認知負荷量表與學習成效測驗，約 15 分鐘。

2.3. 研究工具

2.3.1. 認知負荷量表

本研究之認知負荷量表參考自 Lin 等人（2021）依據 Leppink 等人（2013）所改編的中文翻譯量表，僅配合本研究學習主題修改學習活動名稱，經兩位數位學習領域專家審查達專家效度。此量表分為三個子構面，內在認知負荷（3 題）、外在認知負荷（3 題）、增生認知負荷（4 題），共 10 題。本量表採用李克特氏六點量表，1 表示非常不同意至 6 表示非常同意，量表子構面信度依序為.94、.93、.93 和整體為.63（Nunnally et al., 1967）。

2.3.2. 先備知識測驗與學習成效測驗

先備知識測驗以及學習成就測驗為研究者根據英語閱讀教材內容結合 Google Arts & Culture（Google, 2022） VR 內容自行編製英語閱讀內容後，依照 Anderson 等人（2001）之新版 Bloom 認知歷程向度設計記憶、理解、分析向度各 5 題，共 15 題，試題經由二位 10 年以上的英語領域專業教師審查，達專家效度。試題為四選一選擇題，每題 6 分，共 90 分。依據 Aiken（1997）提出難度.3~.7 標準以及 Ebel 和 Frisbie（1972）提出鑑別度.2 標準篩選試題，共保留 10 題進行統計分析。先備知識測驗與學習成效測驗內容相同，僅調整答案順序。

2.3.3. 概念圖 IRS

本研究之概念圖由研究者參照英語閱讀文章內容設計，共有 11 道填空題須由學習者根據閱讀文章內容填寫對應答案。IRS 採用僅需教學者準備行動裝置即可運作的 Plicker（2022），教學者利用行動裝置掃描學習者的答案紙板，雲端系統會立即顯示作答情形。

2.3.4. 穿戴式虛擬實境

本研究使用的穿戴式 VR 為頭戴式穿戴設備結合智慧手機播放的 Google Arts & Culture 教材內容整合而成，頭戴式穿戴設備採用小宅 Z5 頭盔（小宅, 2018）搭配智慧型手機 Samsung Galaxy S8，頭盔配有具有環繞音效耳機，使原音重現提升沉浸感，教材內容為 Google Arts & Culture 的 The Fall of the Rebel Angels（Google, 2022）。

3. 分析結果

根據調節分析結果，概念圖 IRS 組中，內在認知負荷、外在認知負荷、增生認知負荷預測學習成效之解釋變異量為-2%，未具有統計顯著水準($F(3,29) = 0.80$, $p = .505$)，顯示認知負荷子構面對學習成效的解釋力不足。在合作概念圖 IRS 組，內在認知負荷、外在認知負荷、增生認知負荷預測學習成效之解釋變異量為 3%，未具有統計顯著水準($F(3,30) = 1.35$, $p = .278$)，代表使用合作概念圖 IRS 教學方式時，認知負荷子構面對於學習成效的解釋力不足。在穿戴式虛擬實境合作概念圖 IRS 組，內在認知負荷、外在認知負荷、增生認知負荷預測學習成效之解釋變異量為 17%，且具有統計顯著水準($F(3,32) = 3.35$, $p < .05$)，其中內在認知負荷顯著負向預測學習成效($\beta = -0.49$, $p < .05$)，代表使用穿戴式虛擬實境合作概念圖 IRS 教學方式時，學生的內在認知負荷越高，學習成效將會越低。

4. 討論與結論

本研究結果顯示教學設計為認知負荷預測學習成效之調節變項，在概念圖 IRS 組和合作概念圖 IRS 組中，認知負荷對學習成效解釋力不足，而穿戴式 VR 合作概念圖 IRS 組能經由內在認知負荷顯著負向預測學習成效。內在認知負荷與學習任務複雜度和學生的先備知識相關（Sweller et al., 2011）。如同 Lin 等人（2021）指出因學習者使用穿戴式 VR 的學習經驗較少，因此缺少先備知識作為學習與操作基礎，需付出較多的認知能力熟悉穿戴式 VR 的學習模式。因此，穿戴式 VR 合作概念圖 IRS 組較能引發內在認知負荷對學習成效的預測力。此外，在英語閱讀中，內在認知負荷與閱讀成績之間為顯著負相關（Jiang, 202），本研究結果可能因任務複雜度較高，同樣產生內在認知負荷顯著負向預測學習成效之結果。未來研究建議探討三組之間的認知負荷與學習成效差異、增加質性資料輔助量化資料解釋，以及增加穿戴式 VR 概念圖 IRS 組探討穿戴式 VR 有無加入合作學習是否會影響學習成效。

5. 誌謝

本研究感謝國科會(NSTC 112-2410-H-A49-019-MY3 和 MOST 111-2410-H-A49-018-MY4)之支持。

參考文獻

小宅（2018）。**小宅 Z5**。https://www.xiaozhai360.com/z5/

Aiken, L. R. (1997). *Pyschological testing and assessment* (5th ed.). Ally & Bacon.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Addison Wesley Longman, Inc.

Ebel, R. L., & Frisbie, D. A. (1972). *Essentials of educational measurement*. Pearson.

Google. (2022). *Google arts & culture*. https://artsandculture.google.com/

Jiang, C. (2023). Chinese undergraduates' English reading self-efficacy, intrinsic cognitive load, boredom, and performance: A moderated mediation model. *Frontiers in Psychology*, *14*, Article 1093044.

Leppink, J., Paas, F., Vleuten, C. P. M. V. d., Gog, T. V., & Merriënboer, J. J. G. V. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058-1072. https://doi.org/10.3758/s13428-013-0334-1

Lin, H. C.-S., Yu, S.-J., Sun, J. C.-Y., & Jong, M. S. Y. (2021). Engaging university students in a library guide through wearable spherical video-based virtual reality: Effects on situational interest and cognitive load. *Interactive Learning Environments*, *29*(8), 1272-1287. https://doi.org/10.1080/10494820.2019.1624579

Nishimura, T., & Sakurai, S. (2017). Longitudinal changes in academic motivation in Japan: Self-determination theory and East Asian cultures. *Journal of Applied Developmental Psychology*, *48*, 42-48.

Novak, J. D., & Cañas, A. J. (2007). Theoretical origins of concept maps, how to construct them, and uses in education. *Reflecting Education*, *3*(1), 29-42.

Nunnally, J. C., Knott, P. D., Duchnowski, A., & Parker, R. (1967). Pupillary response as a general measure of activation. *Perception & Psychophysics*, *2*(4), 149-155. https://doi.org/https://doi.org/10.3758/BF03210310

Plicker. (2022). *Plicker*. https://get.plickers.com/

Sun, X. (2023). Literature in secondary EFL class: Case studies of four experienced teachers' reading programmes in China. *The Language Learning Journal*, *51*(2), 145-160.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot! for learning–A literature review. *Computers & Education*, *149*, Article 103818. https://doi.org/https://doi.org/10.1016/j.compedu.2020.103818

Yu, S.-J., Hsueh, Y.-L., Sun, J. C.-Y., & Liu, H.-Z. (2021). Developing an intelligent virtual reality interactive system based on the ADDIE model for learning pour-over coffee brewing. *Computers and Education: Artificial Intelligence*, *2*, Article 100030. https://doi.org/10.1016/j.caeai.2021.100030

Zamani, M. (2016). Cooperative learning: Homogeneous and heterogeneous grouping of Iranian EFL learners in a writing context. *Cogent Education*, *3*(1), Article 1149959. https://doi.org/https://doi.org/10.1080/2331186X.2016.1149959

虛擬實境下之全境式火場求生模擬演練學習系統

A Full-Scale Fire Survival Simulation Learning System based on VR Technology

鄭樺嚎、區國良、唐文華

國立清華大學學習科學與科技研究所

klou@mx.nthu.edu.tw

【摘要】本研究旨在發展一套適合國小高年級學生使用的「VR 全境式火場求生模擬演練」學習系統，並探討其對於學習成效、記憶延續性與學習內動機的影響，以提供教師於防災教學時使用。本研究使用準實驗研究法，以自編「火災防災素養量表」與「動機學習策略問卷」為工具，選取新竹某國小高年級四個班的學生，得有效樣本數 83 人進行實驗，並利用描述性統計、獨立 t 檢定以及 ANCOVA 共變異數等統計方式進行分析。結果顯示，使用本系統進行全境式模擬演練對國小高年級學生在火災防災素養的學習成效及學習動機中的內在動機皆優於講述式桌上演練，代表利用虛擬實境技術所建構的全境式火場求生模擬演練系統適合作為國小防災教育使用，同時對於培養火災防災素養具有積極且正向的影響。

【關鍵詞】虛擬實境、防災教育、火災防災演練、火災防災素養、記憶延續性

Abstract: *This study aimed to develop a contextualized fire survival simulation learning system for upper elementary school students and to investigate the effects of this system on learning motivation, memory continuity, and learning motivation. There are 83 participants from four elementary classes of the 5th grade in a national elementary school in Hsinchu. Using a quasi-experimental research method, a self-developed "Fire Preparedness Competence Scale" and a motivational learning strategy questionnaire. The results indicated significant differences in the learning outcomes and intrinsic motivation of fire prevention literacy among the upper elementary school students using this system, which positively impacted the development of fire prevention literacy. This study will assist researchers, teachers, and disaster prevention education stakeholders use a quality approach to fire prevention* drills.

Keywords: *Virtual reality, Disaster prevention education, Fire prevention drills, Fire Preparedness Competence, Memory continuity*

1. 研究背景與動機

「火災」的發生，對人類的生命與財產容易造成重大的危害，而火災「機率低、影響高」的特性，容易導致民眾輕忽培養正確的防災知識與技能的重要性，民眾防災行動多仰賴自身對環境的認識與曾經歷的災害經驗，如何培養民眾在火場中正確的「初期應變行為」，則需建立在充足與良好的防災教育與演練基礎上;然而，防災教育與演練常面臨場地、設備、成本、師資與風險等問題，難以讓學生真正在災害情境中訓練如何判斷及應對災害的知識與技能；另外，若缺乏災害的學習情境，也導致學生不易理解防災知識與技能的重要性，使得在真正面臨災害時難以救己助人。

虛擬實境(Virtual Reality)為透過電腦建立的三維虛擬世界，提供包含視覺、聽覺、移動及肢體動作等模擬效果，讓使用者彷彿身歷真實情境的沉浸於虛擬世界中進行觀察與互動，讓學習者在安全且接近真實的情境脈絡中習得相關知識。許多研究開始嘗試導入 VR 科技進行防災教育，例如 Smith and Ericson (2009)研究顯示，虛擬實境中的遊戲式學習環境應用於火災防災教育中，能提升兒童對消防安全知識的興趣； Morélot et al. (2021)與 Mystakidis et al. (2022)的實驗中，將虛擬實境應用於火災消防安全培訓，發現有助於促進消防安全培訓的程序性知識與學習成效的表現，顯示出應用虛擬實境情境進行防災教育與訓練能有助於提升學習成效與學習動機。

2. 系統設計

本研究所開發的全境式火場求生模擬演練學習系統，開發工具為 Unity3D 引擎執行於結合於智慧型手機之 Cardboard VR 眼鏡，使用者戴上眼鏡後將可感受到全境式的空間情境，並

利用視野中央的小白點作為點擊資訊點，與環境物件互動。本研究提供受測者 4 種場景，分別是「聽到警報聲-樓上起火」、「聽到警報聲-樓下起火」、「聽到警報聲-門外起火」以及「無警報聲-房間內電器起火」，依照起火位置分為「火在門內」與「火在門外」兩路線進行。四個場景皆為臥室，但環境條件與逃生路徑情境條件則因不同情境而有所區別。第一場景與第二場景設定起火點在臥室外較遠處，分別在樓上及樓下；第三場景起火點較近，門外濃煙火焰由房間門縫竄入(圖 3)；第四場景則為房間內起火，室內具煙霧與火焰(圖 4)。

逃生路線中的「樓梯」場景則依照各關卡情境中不同火災發生位置，分為樓上起火而樓梯間「無煙與火」(圖 1) 與樓下起火而樓梯間「濃煙火焰」(圖 2)兩種場景條件，學習者進行火場求生演練時，如若逃生路線選擇進入「浴室」場景時，則代表該學生具有一般民眾常見的火場求生迷思概念，因此，本系統所設定之「浴室」場景進入後，立即顯示提示板警告，而其餘「浴室」場景各資訊點，將分別說明浴室為何不適合進行避難求生。



圖 1. 聽到警報聲-樓上起火



圖 2. 聽到警報聲-樓下起火



圖 3. 聽到警報聲-門外起火



圖 4. 無警報聲-房間內電器起火

學習者再進行求生演練過程中的不同選擇，將會影響其火場求生的成功與否，在各場景中，透過資訊點獲取情境資訊與火災防災知識，並藉此選擇正確的求生方式或路徑，即可逃生成功或順利求生避難獲救。順利求生後，也透過場景提示板提醒如何報案與提供哪些資訊協助消防人員進行後續救援任務。學習者進行火場求生演練過程中，如選擇錯誤方式或路線進行求生，則會進入失敗畫面之烈焰火場，透過火焰畫面與音效，讓學習者留下深刻的印象，並透過系統提示板了解其求生方式或路線選擇的錯誤，累積其失敗演練經驗，並藉此經驗重新挑戰該場景。

3. 研究結果

本研究採準實驗研究法，以教學方式為自變項，實驗組使用的是「VR 全境式火場求生模擬演練」學習系統，對照組則使用「講述式防災桌上演練」教學。本研究以新竹市某國小五年級學生為研究對象，研究選取了四個班級 83 名學生作為研究對象，實驗組 45 人，對照組 38 人，其中男性 46 人，女性 37 人，四個班級被隨機分配為兩個對照組和兩個實驗組，由同

一位教師負責進行測驗和演練教學。

實驗組與對照組的演練教學內容皆相同，實驗課程實施前一週，實驗組與控制組先進行「火災防災素養量表」之學習成效問卷量表測驗，取得前測分數。實驗當天將進行一節課 40 分鐘的實驗課程，實驗組使用「VR 全境式火場求生模擬演練」，而對照組則利用「消防宣導影片」進行的講述式防災桌上演練，實驗組與對照組的教學知識內容與授課教師皆相同。在教學活動結束後，參與者實施和前測相同內容之「火災防災素養量表」之學習成效問卷量表進行評量測驗，比較實驗組與對照組之教學模式對學生求生技能學習之學習成效。



圖 5. 實驗組操作系統情形



圖 6. 對照組上課情形

本研究使用之研究測量工具包括：「火災防災素養量表」與「動機學習策略問卷」，每個面向各十題，共二十題四選項之選擇題，經過三位警察大學消防領域專家評閱修正確保內容效度。另外再以基隆市某國小五年級的學生共 135 人作為樣本進行預測分析，使用檢測題目同質性的 KR20 信度（Kuder-Richardson Formula 20）值為.92，顯示本研究發展之火災防災素養量表試題內部一致性高。

(一) 學生在「火災防災素養」的學習成效表現

本研究以火災防災素養學習成效測驗的前測成績為共變項，教學方式為自變項，火災防災素養學習成效測驗的後測成績為依變項，進行獨立樣本單因子共變數分析(One-way ANCOVA)，探討實驗組與對照組學生在火災防災素養的學習成就是否有差異。結果顯示實驗組平均數為 79.11，標準差為 16.14；對照組平均數為 65.13，標準差為 20.18，實驗組防災素養學習成效的測驗平均數高於對照組，也就是實驗組火災防災素養的學習成效表現較佳，而標準差則顯示出實驗組的成績分布差異小於對照組，實驗組的後測成績較為集中，對照組的後測成績則較為分散。在後測結果部分，實驗組平均數為 77.71；對照組平均數為 67.26。由平均數顯示，實驗組的防災素養中火災防災素養學習成效測驗平均數較對照組高，顯示實驗組的學習成效優於對照組。再由共變數分析摘要中得知(表 1)，組別考驗效果達顯著，$F(1,79) = 4.498$， $p = .037 < .05$，淨相關 $\eta^2 = .054$，屬於中度效果量，顯示火災防災素養學習成效測驗因使用不同的火災求生演練方式而有顯著差異。

以上結果顯示，使用虛擬實境技術進行全境式火場求生模擬演練對學生學習火災防災素養時有正面的學習成效，代表本研究發展的模擬演練學習系統，利用虛擬實境技術讓學生在接近真實的虛擬環境中，實際體驗火場求生情境以引發真實生活的情感，並讓所學知識能立即實際應用在災害情境中，對於火災防災素養的學習與體驗有正面的影響，有助於加深對學習內容的印象與實際應用技能的能力。除此之外，本研究提供「安全」且接近真實的火場求生情境，讓學習者能避免空間、時間、成本與風險的限制，在接近真實的火場情境中進行逃生演練，累積成功與失敗的防災經驗，並透過系統全景畫面與情境音效，讓學習者透過沉浸在學習情境中，透過與學習情境進行互動，主動的建構知識的認知主體，實際將知識與技能

應用在災害情境，讓知識建構在實際應用知識的學習情境中，有助於提升學習者在火災防災素養的學習成效。

(二)「VR 全境式火場演練」對火災防災素養學習動機的影響

實驗組和對照組在內在動機前、後測成績部分，實驗組前測平均數為 11.71，標準差為 2.63，後測平均數為 12.38，標準差為 2.60；對照組前測平均數為 11.58，標準差為 2.49，後測平均數為 11.24，標準差為 2.60。內在動機的後測成績經 Levene's 的變異數同質性檢定未達顯著(F = .001，p = .972 > .05)，表示兩個數據樣本的離散情形並無明顯差異，而由假設變異數相等的 t 值與顯著性，發現檢定結果達顯著。由以上結果顯示，使用「VR 全境式火場求生模擬演練」學習系統對學生在學習火災防災素養相關知識時的內在動機優於使用「講述式桌上演練」，由標準差顯示，實驗組與對照組的成績分布都較為集中，此外，實驗組的內在動機明顯提升。在外在動機前、後測成績部分，實驗組前測平均數為 11.20，標準差為 2.15，後測平均數為 11.58，標準差為 2.41；對照組前測平均數為 11.34，標準差為 2.51，後測平均數為 11.71，標準差為 2.57。外在動機的後測成績經 Levene's 的變異數同質性檢定未達顯著(F = .008，p = .929 > .05)，表示兩個數據樣本的離散情形並無明顯差異。而由假設變異數相等的 t 值與顯著性，發現檢定結果未達顯著，表示實驗組和對照組在外在動機後測成績上並無明顯差異，由標準差顯示，實驗組的成績分布比對照組小，代表對照組的後測成績較為分散，實驗組的後測成績較為集中。

以上結果顯示利用虛擬實境技術提供學習者主動參與學習，促使學習者主動地去思考、探索與解決問題，對於探索學習及使用者體驗有正面的影響，而這些因素皆主要針對學習者的內在心理，實驗過程中較少有教師與同儕等外在因素影響，因此研究結果中主要提升的為學習者的「內在學習動機」，而 VR 虛擬實境演練能提供學習者體驗逼真的虛擬操作與高風險環境的現場體驗，也有助於提高學習者學習的積極性。

4. 結論

本研究之研究目的為以虛擬實境(VR)技術建構一個安全且具有學習成效的火場求生模擬演練系統，使用國小五年級自然與生活科技領域的學習主題為核心，結合火場求生原則與台灣地區 2016～2020 年統計的火災數據，設計火災知識內容與火場逃生演練場日口卜火，提供學生學習，再以「火災防災素養量表成效測驗」及「動機學習策略問卷」來評估學習者使用本系統後對學習成效與學習動機是否有顯著差異。研究結果顯示，兩種防災演練方式皆對火災防災素養的學習成效有所提升，其中以使用「VR 全境式火場求生模擬演練」學習系統的實驗組學習成效表現顯著優於使用「講述式桌上防災演練」的對照組；另外實驗組在火災防災素養的「內在動機」方面同樣也優於對照組，在顯著差異性方面達顯著差異；而兩組的「外在動機」則未達顯著差異。以上研究結果顯示使用「VR 全境式火場求生模擬演練」學習系統，相較於傳統講述式學習方式，對於學習者學習防災素養的學習成效及內在動機具有顯著的提升效果，且在安全及可重複操作的環境下，提供全境擬真的訓練環境，適合未來作為國小防災教育使用，提民眾面對火災時應變能力，降低災害風險、增加生存機率。

參考文獻

Morélot, S., Garrigou, A., Dedieu, J., & N'Kaoua, B. (2021). Virtual reality for fire safety training: Influence of immersion and sense of presence on conceptual and procedural acquisition. *Computers & Education*, *166*, 104145.

Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual reality*, *13*(2), 87-99.

什么影响了一对一在线辅导的质量？基于在线辅导质量智能评估的分析视角

What Affects the Quality of One-on-One Online Tutoring? An Analytical Perspective Based on Intelligent Assessment of Online Tutoring Quality

汪洋，赵梓明，陈玲*
北京师范大学未来教育高精尖创新中心
* chenling@bnu.edu.cn

【摘要】 双减背景下利用网络开展师生一对一在线辅导丰富了教育公共服务个性化供给，也成为学生个性化问题解决和扬长学习的有效途径，但如何保障辅导质量，利用技术实现对辅导过程的智能监管仍存在挑战。本研究基于“北京市中学教师开放型在线辅导计划”平台中 1329 条师生一对一在线辅导记录，从互动结构、话语表现、辅导情景、反馈评价四个维度共筛选出了 21 个特征，构建了一对一辅导质量自动评估模型，并进一步挖掘了影响辅导效果的关键因素。研究结果表明，静默占比、话轮交替频率、师生发言比例、学生积极反馈频率等是区分辅导层次的关键特征。该模型可为辅导平台及师生双方提供及时的信息反馈，从而有依据的施加干预和个性化支持以提升后续辅导质量。

【关键词】 在线教学服务；一对一辅导；在线学习；数据挖掘；质量分析

Abstract: *One-on-one online tutoring between teachers and students in the context of double-decrease has enriched the personalized supply of education public services and become an effective way for students to solve personalized problems and learn by their own strengths, but how to guarantee the quality of the tutoring and how to use technology to realize intelligent supervision of the tutoring process still remains a challenge. In this study, based on 1,329 records of one-on-one online tutoring between teachers and students on the platform of "Open Online Tutoring Program for Secondary School Teachers in Beijing", a total of 21 features are screened out from four dimensions: interaction structure, discourse performance, tutoring context, and feedback and evaluation, and a model for automatic assessment of the quality of one-on-one tutoring is constructed, and key factors affecting the effectiveness of the tutoring are further mined. The key factors affecting the tutoring effect were further explored. The results of the study show that the percentage of silence, the frequency of turn-taking, the ratio of teacher-to-student speeches, and the frequency of positive feedback from students are the key features that distinguish the tutoring levels. The model can provide timely feedback to the tutoring platform and both teachers and students, so that intervention and personalized support can be applied accordingly to improve the quality of subsequent tutoring.*

Keywords: online instructional services, one-on-one tutoring, online learning, data mining, quality analysis

1. 引言

目前，我国教育仍存在优质师资资源配置不均衡、个性化教育服务供给不充分等问题。中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》中提到，“各地积极创造条件，组织优秀教师开展免费在线互动交流答疑，做强做优免费线上学习服务，提升人民群众教育满意度”。在线辅导作为促进优质资源流转

共享的新教育形态，将师生交互从实体空间转移到虚拟空间，利用互联网提供了更加开放共享的学习资源，打破了传统教学时间与空间的束缚，为学生的课外个性化发展提供了有效的支持。

辅导过程的质量监管是在线辅导有效开展的重要保障，教育部等十一部门《关于促进在线教育健康发展的指导意见》明确指出需要持续提升在线辅导质量，完善其资源和服务标准体系，进一步健全治理体系。然而，目前对于线上教育的评估研究以师生一对多的大规模在线开放课程为主，缺少对一对一的在线辅导形式中高质量教学方式的特点和影响因素的挖掘与分析。具体而言，在线辅导质量评估的难点之一在于如何收集恰当且完整的数据刻画整个教学背景、实施与结果，并合理、及时的评估其成效(郭文革, 2020)，其中基于规模化数据的智能化质量监管和治理是在线辅导良性运转的关键。基于此，本文旨在构建一对一在线辅导质量的评估模型，并在广泛评估中挖掘教师和学生的有效辅导策略，通过数据驱动一对一在线辅导教学的完善。

2. 在线辅导质量评估研究现状综述

已有证据表明复杂的辅导过程能够被算法模型所捕获与解释(Boyer et al., 2011)。目前对于在线辅导质量的智能评估研究一般分为两个步骤，即先提取辅导过程的特征数据，再基于特征数据构建预测模型。例如，Lin (2022)等人结合人工与 Bert 模型对规模化辅导数据进行标记，并使用梯度提升决策树预测了学习者参与辅导的效果。Chen (2019)等人的研究也较为类似，这些研究结果都初步表明，不同质量一对一在线辅导中出现的师生对话行为存在差异。除了辅导过程的对话交流之外，一些非言语行为特征也能够对辅导质量产生影响，如 Cukurova (2022)等人总结出高效辅导中教师通常具有以适当的停顿行为来安排调控、积极检查屏幕或麦克风等辅导工具技术问题的行为特点。

综合来看，现有研究对在线辅导质量的评估依赖于人工标记的输入特征，人力成本较高，无法为师生或辅导平台提供及时的反馈与可执行的干预。在辅导质量的影响要素方面，多数研究结果集中于言语或非言语行为方面，对背景和反馈数据关注不足，这些数据与辅导参与度具有一定的相关性，进而可能对辅导质量产生影响(Shan et al., 2021)。基于此，考虑到本研究依托平台对辅导全过程数据记录的优势，我们尝试使用平台中易于提取的特征数据，并增加辅导情境和反馈评价两个特征维度，一方面以增强模型的数据输入和预测效果，另一方面旨在对在线辅导质量的影响要素进行更全面地挖掘。

3. 研究过程

3.1 研究对象

本研究依托“北京市中学教师开放型在线辅导计划”（简称为开放辅导）开展研究，其中一对一辅导是由学生需求所驱动，根据其学习需求实时向教师发起的辅导。图 1 为教师端辅导界面，教师与学生可以通过白板、语音或文字进行交流，平台支持双方上传分享图片、文件。此外，平台提供了一系列关于师生个体背景与辅导发生情境的数据，以及在辅导结束后师生对辅导的评价数据。开放辅导的学生用户为北京市初中在读学生，平台教师均为受过培训的北京市在职初中教师。辅导可以在周一至周五 18:00~22:00 与周末 8:00~22:00 间供学生免费使用，教师可通过辅导学生获得一定绩效激励。

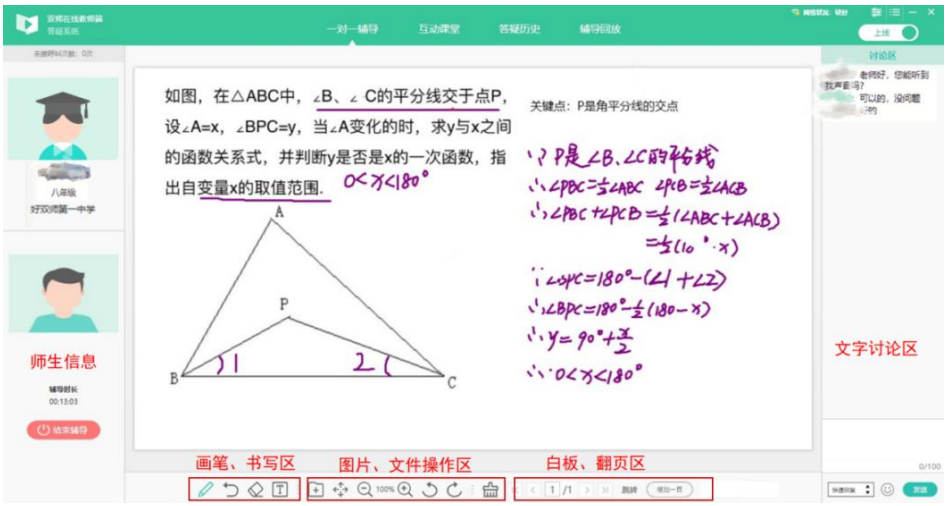


图 1 开放辅导一对一辅导教师端界面

经过随机抽取并去除存在数据缺失的辅导，本研究共获取 1329 条师生线上一对一辅导记录，所用数据覆盖多个学科及不同学校的师生。每条一对一辅导记录的教学质量由开放辅导项目监管人员手动标记为三个类别：优质辅导、普通辅导和低质辅导。优质辅导表现为师生之间具有富有成效的互动，如教师积极引导学生以及学生的问题得到解决，并在辅导课程中学到了东西。低质辅导表现为师生之间存在较多的非生产性互动，即教师没有提供有效的支架来帮助学生解决问题。普通辅导则是介于二者之间。本文所用数据包含优质辅导 220 条，普通辅导 982 条，低质辅导 127 条。

3.2 特征数据的初步选取与处理

鉴于多模态的非言语数据提取过程涉及相对复杂的算法，且受到提取准确性的限制，因此本研究选择从言语数据、辅导情景和反馈评价三个方面选取特征数据。其中，将言语数据分解为互动结构和话语表现两个维度，分别从师生双方的互动数据和教师或学生单一角色的言语数据进行考量，以更细致地评估辅导过程中的关键要素。此外，本研究还参考了前人相关研究的数据选取(Chen et al., 2019)，最终初步构建包含 4 个维度、58 个不依赖于人工标注且易于从平台提取的特征数据体系，待后续进一步筛选。为了便于模型处理，对二分类离散变量进行了哑编码处理，对多分类离散变量进行了独热编码处理。而后对数据进行标准化，以增强不同维度特征在数值上的可比性。20%的数据被随机抽取留作测试集未带入模型挑选，剩余的 80%数据用作训练集与验证集，并采用了 SMOTE 过采样方法应对数据不平衡问题，根据现有数据生成相似的新数据，将优质、低质辅导数据数量提升至与普通辅导相同的 782 条，训练与验证数据总计为 2346 条。

3.3 算法调优与特征数据筛选

在初步模型算法选取阶段，本研究将预处理过的训练与验证数据带入了 7 种分类器算法，并采用四折交叉验证方法评估分类器的差异。其中，随机森林(n_estimators=5000, max_depth = 15, random_state=42)被选中，最大深度为 20 时所带来的准确率提升只有 0.04%，因考虑到潜在的过拟合风险，并未选择更深的参数。各特征数据为随机森林模型带来的平均不纯度减少值被用作于筛选的依据。为提高模型运算效率并提升可解释性，本研究使用剔除策略处理多重共线性的特征数据，即在斯皮尔曼相关系数超过 0.5 的特征之间，仅保留重要性（平均不纯度减少）更高的特征数据。同时，还使用了递归特征消除法以规避部分特征数据所带来的平均不纯度减少极小，对准确度的提升效果不显著的情况。最终，从四个维度中筛选出了 21 个特征数据。表 1 展示了具体特征数据及其相对重要度排序。

表 1 衡量维度与筛选后的特征数据

衡量维度	特征数据	重要性排序	定义
互动结构	静默占比	1	对话音频未被识别出的文本时长占比
	话轮交替频率	2	去除过短发言后的师生话轮交替次数
	师生发言比例	3	教师和学生发言字数之间的比值
	辅导时长	5	辅导持续时长
	闲聊占比	9	与教学内容无关的对话时长占比
话语表现	学生积极反馈频率	4	学生平均每分钟对教师说出的积极肯定性语句次数
	学生语速	8	学生平均每分钟发言字数
	教师鼓励频率	10	教师平均每分钟对学生说出的鼓励性语句次数
	引导确认类问题频率	11	教师平均每分钟使用确认型问题次数
	教师语速	12	教师平均每分钟发言字数
	学生提问频率	13	学生平均每分钟对老师提问的次数
	教师提问频率	14	教师对学生提问次数
辅导情境	师生辅导计数	6	本次辅导是该教师与该名学生在平台内的第几次辅导，衡量师生熟悉程度
	教师教龄	7	教师从事教学工作的连续累计年份
	教师职称	15	教师中小学教师职称
	题目难度	16	学生对内容难度的评级
	发生在工作日或周末	17	辅导发生在日常教学后或是周末
	师生性别	18	教师与学生为同性或是异性
反馈评价	学生积极性	19	教师对学生积极性的评级
	教师对学生评分	20	教师对学生的整体评分
	是否解决了学生的问题	21	学生评价辅导是否解决了他的困惑

在选定特征数据后，再次测试了先前所用的 7 种分类器算法与其各自的超参数，以评估筛选后的 21 个特征数据在分类准确度方面的表现。随机森林(n_estimators=5000,max_depth=15, random_state=42)依然是分类准确率较高的算法。为评估模型的泛化能力，将预留的测试集带入所选的随机森林分类器，并得出在测试集中的判定结果并绘制混淆矩阵。同时，使用 SHAP 计算了每个特征数据的夏普利值，以判断特征数据对辅导所属质量分类的影响。最后，将 1329 条原数据带入可解释提升模型，刻画了每个特征数据与辅导质量之间的形函数，以揭示两者关系。

4. 研究结果

4.1 模型评估

图 2 展示了随机森林模型在测试集上的混淆矩阵，整体准确率为 78.20%。在各类别中，低质辅导的召回率和精确率分别为 86.96%和 90.91%，最高；普通辅导的召回率和精确率分别为 89.36%和 81.55%，居中；而优质辅导的召回率和精确率分别为 36.36%和 52.63%，最低。

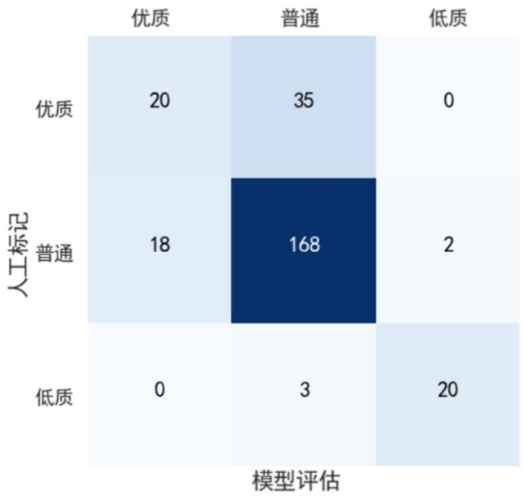


图 2 测试集预测结果混淆矩阵

4.2 在线辅导质量的影响要素分析

图 3 展示了所选随机森林模型在对三类辅导判定中各特征数据的 SHAP 值。横坐标表示特征数据对应的 SHAP 值(为负代表属于该分类的可能性更低,为正代表更可能属于该分类);纵坐标为不同的特征数据,自上而下对质量评估的影响逐步减弱;颜色越红表示特征的值越大,越蓝特征值越小。

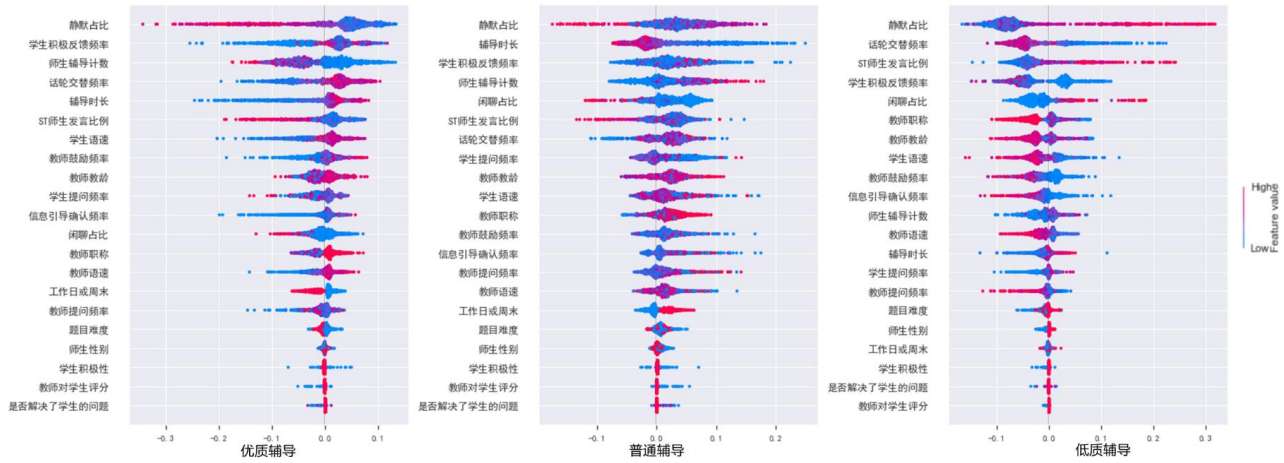


图 3 三类辅导质量评估中各特征数据 SHAP 值

综合来看,辅导的互动结构维度对质量的影响最为显著。结合图 4 进行分析,对比三类辅导的互动结构,其中影响质量最显著的共同特征数据是静默占比,其与辅导质量呈现负相关关系。当静默占比低于 23%时,辅导质量更倾向于优质;而静默占比超过 49%时,辅导质量较低的可能性显著提升。此外,话轮交替频率也是区分辅导质量的重要特征,当每分钟话轮交替频率低于 3 次时,低质概率较高;在 4 至 11 次之间,优质可能性略微提升。在判断低质辅导时,另一个重要特征是师生发言比例。当学生发言字数与教师发言字数比例超过 0.41 时,更可能出现低质量的辅导,这与闲聊时长相关,即低质辅导中师生闲聊较多,导致学生发言量占比较高。最后,辅导时长与辅导质量也有一定的关联,当时长在 30 分钟以内时,辅导质量倾向于为普通,而超过 55 分钟后,优质和低质的概率都会提高。

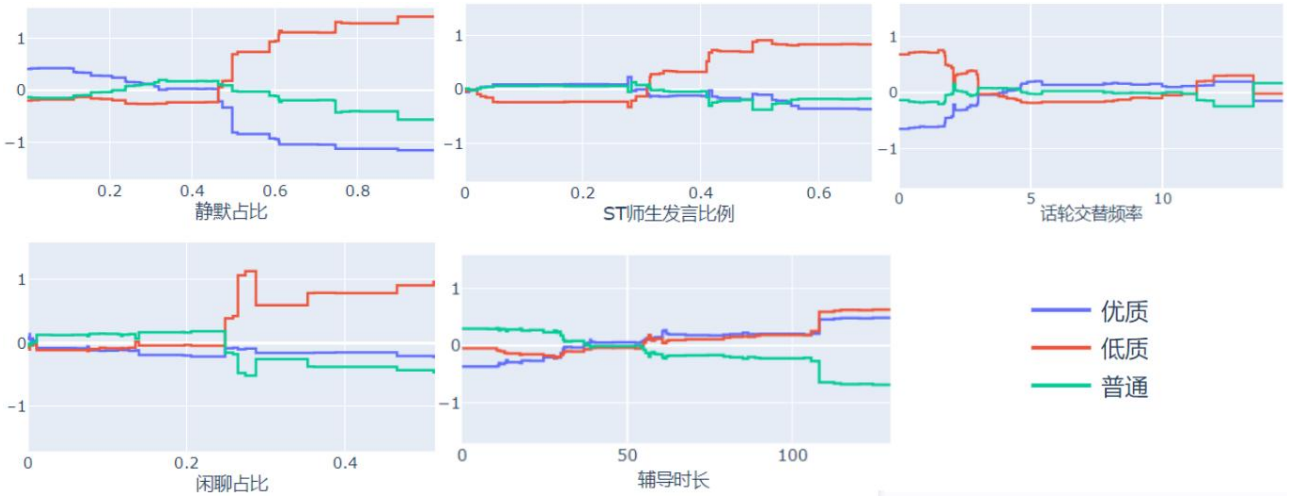


图 4 可解释提升模型中互动结构维度特征数据的贡献

此外，师生话语表现维度对辅导质量有重要的影响。结合图 5 进行分析，学生积极反馈频率对质量判定的影响最为显著，表现为若学生几乎没有积极反馈时辅导低质的可能性较高。反过来，教师对学生给予鼓励同样可以促进辅导质量，随着教师鼓励频率的增加辅导也会更倾向于优质，几乎没有鼓励时低质概率会略微增加。其次，教师与学生的提问表现对辅导质量的影响也较为显著。当教师平均每分钟提问过少时，辅导质量可能偏向普通水平，而高质量的可能性较低。优质辅导的可能性随着问题的增加而增加，在每分钟 1.5 次至 1.9 次的提问频率范围内表现最佳，而当教师提问过于频繁（超过每分钟 4 次）时，优质概率将降低。学生的提问频率同样具有一个最优区间，当学生每分钟提问次数在 0.4 至 1.2 次之间时，优质概率会提高，而在其他情况下，优质概率较低。在语速方面，师生语速与辅导质量具有相关性。学生语速存在着每分钟 130 字至 260 字的最佳区间，区间内略微倾向于优质，低于每分钟 130 字时倾向低质，高于 260 字时倾向普通。教师语速则在每分钟 250 字以内时辅导略微倾向普通或低质辅导，而在 250 字以上时略微倾向优质辅导。最后，引导确认类问题频率对辅导质量也有一定的影响，即优质辅导中往往存在较多的引导确认类问题。

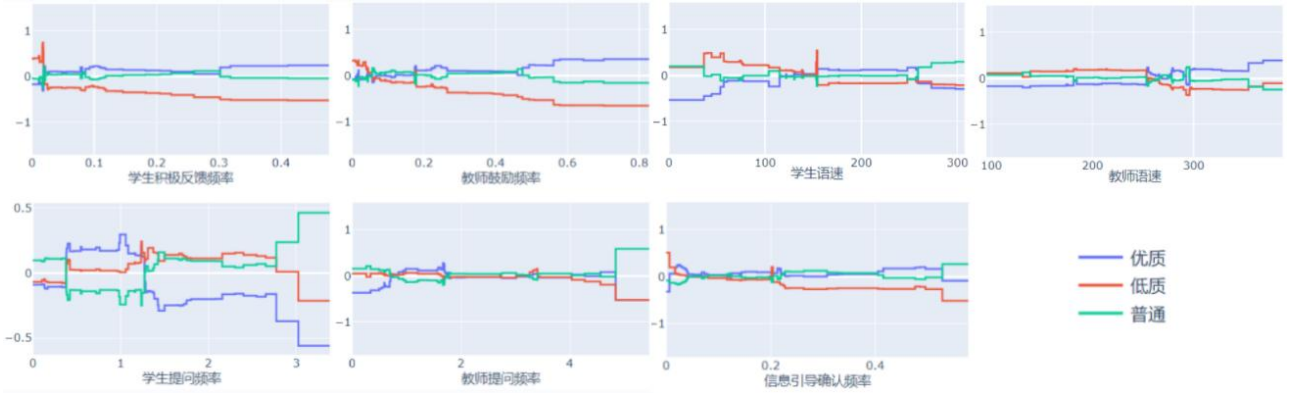


图 5 可解释提升模型中话语表现维度特征数据的贡献

结合图 6 进行分析，在辅导计数方面，师生前 15 次辅导中更可能为优质，而之后优质的可能性逐渐降低。此外，辅导优质的可能性会随着教师教龄与职称逐级提升，正高级教师能够使得辅导的优质概率显著提升。值得注意的是，若辅导在周末进行略微倾向于优质，周末师生双方时间相对充裕，辅导深度和广度更易落实。而在工作日，由于平台开放时间有限且师生晚间可能较为疲惫，这些因素对辅导质量产生一定潜在的负面影响。同时，师生性别之间的差异和题目难度并不会显著影响辅导质量。与其他特征数据相比，学生积极性、教师对学生评分、是否解决了学生问题等反馈评价维度的特征数据贡献相对较小。

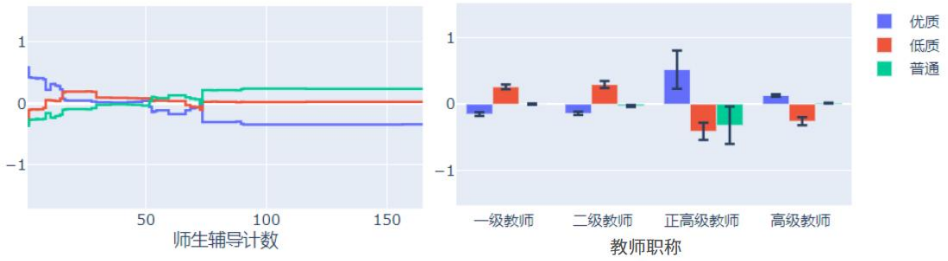


图 6 可解释提升模型中师生辅导计数与教师职称贡献

5. 总结与建议

本文结合开放辅导平台中 1329 条师生一对一辅导记录，挖掘了其中潜在的特征数据，最终构建了在线辅导质量智能评估模型，能够在每次辅导结束后提供即时的效果反馈，从而实现对辅导过程质量的实时监管。基于以上数据分析，对于后续如何提高师生辅导质量，本文从教师、学生、平台三个不同的主体提出如下建议：

5.1 教师：把控辅导节奏，促进学生反思

静默时长是影响一对一辅导质量的关键因素，它考验了教师的效率和掌控能力。由于辅导形式的实时性和随机性，教师需即兴应对学生的各种个性化问题，这要求教师能迅速理解学生的疑惑并采取适当的教学策略。有经验的教师更善于处理突发状况，减少学生等待时间。因此，教师需要在辅导前熟练掌握平台操作，并不断提高自身专业水平和教学策略的灵活性。

同时，教师需要关注学生在辅导过程中的情绪和学习进度，以减少学生注意力的流散。长时间的辅导会使其效果呈现两极分化的状态，即学生在辅导的过程中注意力会逐渐流失耗散，而教师能否引起学生兴趣使其积极参与到互动中以减缓这一过程则能够体现其辅导水平，进而决定辅导效果的好坏。本研究发现教师提供高频率鼓励时的辅导效果倾向优质，即教师的积极反馈能够对学生的情绪和学习状态产生有利的效果。此外，一些策略可以改善学生注意力的流失，包括合理使用确认类问题引导学生参与，并通过发展学生思维和认知的有效对话激发学生思考。然而，学生过高的参与也并不意味着导致更有效的辅导，当学生发言比例过高或提问过于频繁时，辅导效果也会倾向于低质，其原因除了其中较高的闲聊占比之外，还有教师权威性讲解不足、提供信息少，且对应的话题或问题对学生来说过于简单、能够快速回应，难以引起学生思考或带来启示。

最后值得注意的是，教师背景对辅导效果有正向影响，但非关键因素。有研究表明，专家辅导比非专家辅导更有效的原因在于特定的对话动作、互动模式和多话语轮换中(Lu et al., 2007)。这突显了教师需要专注于发展针对在线辅导中的话语互动策略和技能的必要性。

5.2 学生：主动表达需求，积极参与互动

在线上辅导中，教师缺乏学生的面部和肢体语言反馈，这对其判断学生的学习状态和调整教学造成了难度。而信息反馈可大幅度提升远程对象的社会存在感(Ye & Zhou, 2022)，学生的礼貌性和赞美性反馈能显著提高教师的辅导效果。其次，学生主动提问反映了他们的主动思考和知识链接，对有意义的学习至关重要。与面授教学形式不同，一对一辅导提供更自由的提问环境，优质辅导中学生的提问更加频繁。学生需要把握辅导带来的更为自由的提问机会，以获取对知识更深层次的理解。此外，学生的语速与辅导质量呈正相关，其流畅表达反映了清晰的需求和较高的先验知识水平，有助于教师实现针对性辅导。因此，需要鼓励学生在辅导前梳理自身需求和相关知识，主动与教师共同调节辅导节奏，实现更加高效的辅导交流。

5.3 平台：加强数据反馈，完善辅导机制

与线下一对多教学不同的是，线上一对一辅导特别强调教师为个别学生设定个性化目标、了解其能力，并评估自身对学生困难的贡献，是教师最难掌握的能力(Fuller, 1969)。在线辅导平台为教师提供了很好的练习机会，同时基于本研究提出的辅导质量智能评估方法，建议平台在辅导后提供完整数据报告，帮助教师了解效果，并给予改进建议，以提高教学效果和学生体验。同时，平台可以实时追踪学生的辅导情况，提供个性化学习资源。此外，相同师生间辅导效果会随着次数增加而降低，符合经济学中的边际效用递减规律。这意味着平台需要多样化师资库和引入轮换匹配机制，以提供多元教学风格，激发学生学习兴趣。最后在反馈评价模块中，师生之间彼此评分普遍较高，无法反映真实辅导质量。因此建议平台建立多维度的双向评价体系，转为结构化的过程评价，以提供更全面、客观的辅导质量评估。

参考文献

郭文革. (2020). 在线教育研究的真问题究竟是什么——“苏格拉底陷阱”及其超越. *教育研究, 41*(09), 146-155.

Boyer, K. E., Phillips, R., Ingram, A., Ha, E. Y., Wallis, M., Vouk, M., & Lester, J. (2011). Investigating the relationship between dialogue structure and tutoring effectiveness: a hidden Markov modeling approach. *International Journal of Artificial Intelligence in Education, 21*(1-2), 65-81.

Chen, G., Ferreira, R., Lang, D., & Gasevic, D. (2019). Predictors of Student Satisfaction: A Large-Scale Study of Human-Human Online Tutorial Dialogues. *International Educational Data Mining Society*.

Cukurova, M., Khan-Galaria, M., Millán, E., & Luckin, R. (2022). A learning analytics approach to monitoring the quality of online one-to-one tutoring. *Journal of Learning Analytics, 9*(2), 105-120.

Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American educational research journal, 6*(2), 207-226.

Lin, J., Singh, S., Sha, L., Tan, W., Lang, D., Gašević, D., & Chen, G. (2022). Is it a good move? Mining effective tutoring strategies from human–human tutorial dialogues. *Future Generation Computer Systems, 127*, 194-207.

Shan, D., Bhattacharya, P., Kao, B., & Phan, T. (2021). *Characterizing Amateur Tutoring Behavior on a Large Online Learning Platform.* Paper presented at the Proceedings of the Eighth ACM Conference on Learning@ Scale.

Ye, J.-m., & Zhou, J. (2022). Exploring the relationship between learning sentiments and cognitive processing in online collaborative learning: A network analytic approach. *The Internet and Higher Education, 55*, 100875.

Lu, X., Di Eugenio, B., Kershaw, T. C., Ohlsson, S., & Corrigan-Halpern, A. (2007). Expert vs. non-expert tutoring: Dialogue moves, interaction patterns and multi-utterance turns. In Computational Linguistics and Intelligent Text Processing: 8th International Conference, CICLing 2007, Mexico City, Mexico, February 18-24, 2007. Proceedings 8 (pp. 456-467). Springer Berlin Heidelberg.

沉浸式虚拟现实如何影响学习体验与学习效果？

——基于共情能力与虚拟同伴的视角

How does immersive virtual reality affect the learning experience and learning outcomes?

Perspectives Based on Empathic Competence and Virtual Peers

刘伊祉兰[1]，兰翠玲[1]，邢晓宁[1]，王雪*[1]，余能保[2]
1 天津师范大学，教育学部
2 天津工业大学，人文学院
* wangxuetjnu@qq.com

【摘要】 沉浸式虚拟现实环境能够影响学习者的学习体验与学习效果。同伴在学习过程中起到重要的作用，然而且共情能力对感知虚拟同伴及学习的影响尚不明确。基于此开展实验研究，通过不同共情能力的学习者在虚拟环境中观看虚拟同伴在场或不在场的学习材料，考察其对于学习体验及学习效果的影响。结果表明，共情能力对于沉浸式虚拟现实学习环境中学习者的学业情绪、学习动机、自我效能感存在显著影响，虚拟同伴对于学习者的社会存在感具有显著影响。本研究提供了一种新的看待虚拟同伴的视角，为深入探索沉浸式学习环境中的社会化教学提供了方向。

【关键词】 虚拟现实；虚拟同伴；共情能力；学习体验；学习效果

***Abstract*:** *Immersive virtual reality environment can affect learners' learning experience and learning effect. Peers play an important role in the learning process. However, the role of empathy in the perception of virtual peers and its impact on learning remains unclear. Based on this, this paper conducted an experimental study to investigate the influence of different empathic abilities on learning experience and learning effect by viewing learning materials with or without virtual peers in a virtual environment. The experimental results show that empathy ability has a significant impact on learners' academic emotion, learning motivation and self-efficacy in immersive virtual reality learning environment, and virtual peers have a significant impact on learners' sense of social presence. This study provides a new perspective on virtual reality learning environment and virtual peers, and provides a direction for in-depth exploration of social instructional design in immersive learning environment.*

Keywords: Virtual Reality, Virtual Peers, Empathy, Learning Experience, Learning Effect

1.引言

沉浸式虚拟现实以其独特的技术特征受到越来越多教育研究者的关注，并作为教育元宇宙的关键核心技术为教育元宇宙创造条件（马婧 等，2023）。然而，由于当前技术限制等因素，大部分沉浸式虚拟现实学习材料均为单机形式，学习者在沉浸式虚拟现实学习空间中与外部世界的事物、人物隔绝，可能会产生一定程度的孤独感，降低学业情绪，从而影响学习效果。因此，越来越多的研究开始关注社会性因素（周榕和郭佳瑞，2023）

[1].本文系国家自然科学基金青年项目（项目编号：62107030）的研究成果；本文系 2022 年天津市研究生科研创新项目（项目编号：2022SKY266）的研究成果；本文系 2021 年度天津市教育科学规划课题"虚拟现实(VR)教学资源多模态分析及设计优化研究"（HCE210337）的研究成果。

（韩继宁 等，2023）如协作（郑玲 等，2021）、共情（向维 等，2021）等在虚拟现实学习中的影响。VR环境中的虚拟同伴可以增强身临其境的体验（Huang et al., 2019）,起到陪伴、鼓励、监督等作用，可能促进学习者的学习。例如，Picard 设计具有情感计算能力的虚拟同伴，用其感知学习者在学习中的情感信息，并陪伴学习者来帮助他们进行更有效地学习（Picard et al., 2001）。然而作为虚拟环境中的人物形象，学习者能否像感知真实学伴一样感受到虚拟学伴的存在并影响学习，尚未有明确结论。共情是指在与他人接触时，能够识别他人的情绪、动机和身体信息，并理解他人情绪状态的倾向。具备高共情能力的学习者更能够在学习过程中受到学习同伴的认知与情绪等状态影响（Teding & Malouff, 2016）。但少有研究探究学习者的共情能力是否能够影响感知虚拟环境中的学习同伴，并进一步影响其学习。

基于此，本研究设计了两因素 2（共情能力：高共情/低共情）x2（虚拟同伴在场：无虚拟同伴/有虚拟同伴）的被试间实验，以探究虚拟同伴在场及共情能力对于沉浸式虚拟现实学习体验（学业情绪、学习动机、自我效能感、社会存在感）与学习效果（保持测试、迁移测试）的影响。

2. 文献综述

2.1. 沉浸式虚拟现实学习环境对学习的影响

作为新兴的教育形态，VR 能够提高学习者的学习效果。例如，高楠等人（2023）的研究认为，沉浸式 VR 能提高学习者的概念性知识的迁移效果、程序性知识的迁移效果与效率，以及学习过程中的情感体验。根据沉浸式学习认知情感模型（CAMIL）（Makransky & Petersen, 2021），在场和代理为沉浸式 VR 学习的一般心理支持，沉浸、控制因素和表征保真度能够促进这些支持。该模型描述了可以导致基于沉浸式 VR 的学习结果的六个情感和认知因素，包括兴趣、动机、自我效能、体现、认知负荷和自我调节。该模型还描述了这些因素如何导致事实性、概念性和程序性的知识获取和知识转移。综上所述，本研究选取其中的动机、自我效能等因素作为学习体验，知识获取和知识迁移作为学习效果。

2.2. 社会存在感及虚拟同伴对学习的影响

社会存在为学习者在在线学习中的社交和情感上自我投射的能力。更高的社会存在感会为群体的同一性或归属感提供更自主的动机（Shen et al., 2010），这不仅会带来更高质量的参与，还会带来更好的体验（Ryan & Deci, 2001）。社会存在感以学习者自身为主体，主要研究学习者的心理机制，与学习者的学习动机和学习效果有较大的联系，同时社会存在感还与同伴之间的感知互动联系密切（詹泽慧，2014）。王钰莹（2022）认为虚拟现实可引发同伴交互行为，学习者表现出对虚拟同伴的情感关切。Palloff 等（2007）认为社会存在感可以降低学习者的孤独感，帮助学习者获得个性化的学习体验。VR 环境能够增加学习者的社会存在感，例如，Biocca 和 Harms（2002）的研究表明，沉浸式 VR 可能比传统的环境更能唤起社会存在感。在 VR 环境下学习者可以与虚拟同伴进行互动，通过互动来进行知识的积累（刘蕾，2023），提高学习效果（余越凡 等，2022）。综上所述，社会存在感影响学习者的学习。因此，本研究选取社会存在感作为学习者在虚拟环境中感知到学习同伴的因素。

2.3. 共情能力对于学习效果的影响

共情一般是指对他人的感受与行为的的能力（Brydon et al., 2021）。共情是人类与生俱来的能力，不同的人具有不同的共情水平（Villalba et al., 2021），可分为共情能力高与

共情能力低两种。共情能力能够改善学习者之间的互动，提高学习动机，最终提升学生的学业成绩。共情能力更高的学习者无论是在学习上还是人际交往上均强于共情能力低的学习者（McQuiggan et al., 2010），为本研究奠定了基础。因共情本身所具有的换位感知性，其与虚拟现实密切联系。在教育领域中，教师可以通过 VR 技术营造的沉浸式体验来激发学习者的共情，在 VR 中学习者可以互相分享观点，与其他学习者建立联系进而产生共情（王浩和唐爱民，2023）。相关研究表明，沉浸感与互动性更强的虚拟现实环境可以激发学习者的共情（Barreda-Ángeles et al., 2000）。例如 Schutte 和 Stilinović（2017）的研究发现，在沉浸式 VR 中观看教学视频更能提高学习者的参与度与共情能力。Dyer 等人（2018）的研究通过在沉浸式 VR 环境下的训练，来提升医学学生的共情能力，进而增强其专业能力，最终提高学习效果。由此可见，虚拟现实的沉浸感可能影响共情能力，从而影响学习效果。进一步，共情能力高的学习者可能更能深入感知虚拟同伴的存在，并受到更加深入的影响。由于共情能力高的学习者其学业情绪波动可能较大，因此，本研究将学业情绪也纳入学习体验，考量情绪在其中起到的作用。综上所述，虚拟同伴对于学习的影响机制尚不明确，且少有研究将共情能力纳入虚拟同伴对学习的影响共同研究。为了解决上述研究需求，本研究通过随机实验探究该两因素对于（学业情绪、学习动机、自我效能感、社会存在感）与学习效果（保持测试、迁移测试）的影响。

3.研究方法

3.1. 研究对象

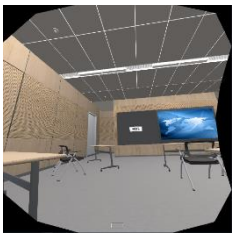
研究对象为 88 名在校大学生，被试年龄分布在 17~28 岁，平均年龄约为 20 岁，其中男性 25 名，女性 65 名。在实验开始前，被试被随机分布在无虚拟同伴组（n=48）和有虚拟同伴组（n=40）。其中，由于 2 名被试中途放弃实验，最终有效被试总数为 86 人。每位被试在实验前均签署了《实验知情同意书》，为自愿参与实验，并被告知如果正常完成实验可获得被试费作为报酬。

3.2. 实验设计

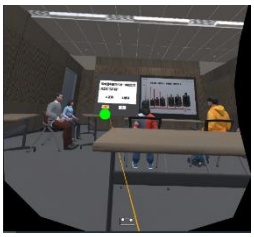
本实验为双因素 2（共情能力：高共情/低共情）x2（虚拟同伴在场：无虚拟同伴/有虚拟同伴）的被试间实验设计。其中，共情能力为后期根据人格类型量表（IRI-EC）测量数据取数据整体的中位数作为临界值，高于中位数的分类为高共情人格，不高于中位数的分类为低共情人格。由于被试被随机分布在无虚拟同伴与有虚拟同伴组，因此被试人格类型在两组之间的分布并不完全均匀，无同伴低共情 24 人、无同伴高共情 24 人、有同伴低共情 16 人、有同伴高共情 23 人。

3.2.1. 学习材料

本实验所用的学习材料参考借鉴国家虚拟仿真实验平台的教学资源《曾侯乙编钟演奏与创编实验》，基于 Unity 3D 开发曾侯乙编钟学习场景（如图 1 所示）。无虚拟人物组在空教室中学习相关知识并交互；有虚拟人物组在虚拟人物围坐的教室中学习相关知识并交互，交互过程中虚拟人物对被试做出相应点头、摇头、注视等形式的反馈。虚拟同伴组的人物基于 3dsMax 开发，共 11 人，具备点头、转头、注视转移、举手、招手等多种动作设计。各被试佩戴 HTC Vive Pro Eye 学习。由于具备交互环节，学习时长约 18-21 分钟。


a 无虚拟同伴-认知


b 无虚拟同伴-操作


c 有虚拟同伴-认知

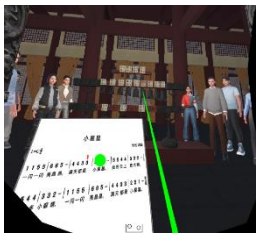

d 有虚拟同伴-操作

图 1 实验材料部分图示

3.2.2. 实验流程

随机招募被试单独开展实验。首先，被试需要阅读《实验知情同意书》并签字。其次完成人口统计学调查和先验知识测试，主要用于考察被试对于编钟及基本乐理知识的经验等。然后，主试向被试介绍整体实验流程、沉浸式虚拟现实设备使用方法及注意事项。之后，被试随机分布在不同实验组别中，学习实验材料。学习结束后，被试完成体验问卷和知识后测。

3.3. 测量工具

知识测试是参考国家虚拟仿真平台相关资源进行设计并扩展。保持测试包含 10 项选择题（每题 1 分），例如“甬钟的上部被称为（ ）”；迁移测试包含 10 项选择题（每题 1 分）（例如，“晓玉在一个编钟上看到了一些认不太出来的文字，晓玉认为首先这个钟可能是（ ）钟。”）和一项简答题（3 分）。

共情采用 Davis 的经典量表（1980）测量，Cronbach’s α=0.72。学业情绪采用经典的 PANAS 问卷测量，Cronbach’s α=0.925。社会存在感采用 Ning Shen 和 Khalifa（2008）开发的测量方法进行了修改，以适应本研究的学习活动，题项例如“我几乎没有注意到虚拟教室中的其他人”，Cronbach’s α=0.812。动机采用 Isen 和 Reeve（2005）编制的自我报告问卷并进行了修改，以适应本研究的学习活动，Cronbach’s α=0.942。题项如“它使我想进一步探索”。自我效能感采用 Artino 等人编制的自我报告问卷并进行了修改，以适应本研究的学习活动，Cronbach’s α=0.831。题项如“在刚刚的学习中，我有信心在没有他人的帮助下也能学好编钟。”

4.研究结果

4.1. 共情能力与虚拟同伴对于学习体验的影响

有无虚拟同伴对于学习者的积极学业情绪主效应不显著，F（1，83）=1.615，*p=.207*，η^2=.019。而共情能力对于积极学业情绪具有显著的主效应，F（1，83）=7.441，*p=.008*，η^2=.082，且低共情能力组的积极学业情绪（M=32.872，SD=.873）显著低于高共情能力组（M=36.439，SD=.973）。对于积极学业情绪，两者交互作用不显著，F（1，83）=2.152，*p=.146*，η^2=.025。

有无虚拟同伴对于学习者的内在动机主效应不显著，F（1，83）=.744，*p=.391*，η^2=.009。而共情能力对于内在动机具有显著的主效应，F（1，83）=10.813，*p=.001*，η^2=.115，且低共情能力组的内在动机（M=44.563，SD=1.016）显著低于高共情能力组（M=49.568，SD=1.133）。对于内在动机，两者交互作用不显著，F（1，83）=2.152，*p=.146*，η^2=.025。

有无虚拟同伴对于学习者的外在动机主效应不显著，F（1，83）=.718，*p=.399*，η^2=.009。而共情能力对于外在动机具有显著的主效应，F（1，83）=5.781，*p=.018*，η^2=.065，且低共情能力组的外在动机（M=15.344，SD=.420）显著低于高共情能力组

（M=16.857，SD=.468）。对于外在动机，两者交互作用不显著，F（1，83）=1.243，*p=.268*，η^2=.015。

有无虚拟同伴对于学习者的自我效能感主效应不显著，F（1，83）=.027，*p=.869*，η^2=.000。而共情能力对于自我效能感具有显著的主效应，F（1，83）=7.141，*p=.009*，η^2=.079，且低共情能力组的自我效能感（M=18.241，SD=.472）显著低于高共情能力组（M=20.131，SD=.526）。对于自我效能感，两者交互作用不显著，F（1，83）=.214，*p=.645*，η^2=.003。

有无虚拟同伴对于学习者的社会存在感主效应显著，F（1，83）=31.780，*p=.000*，η^2=.277，且无虚拟同伴组的社会存在感（M=5.161，SD=.377）显著低于有虚拟同伴组（M=8.346，SD=.421）。而共情能力对于社会存在感不具有显著的主效应，F（1，83）=.227，*p=.635*，η^2=.003。对于社会存在感，两者交互作用不显著，F（1，83）=2.375，*p=.127*，η^2=.028。

4.2. 共情能力与虚拟同伴对于学习效果的影响

有无虚拟同伴对于学习者的保持成绩主效应不显著，F（1，83）=.908，*p=.344*，η^2=.011。共情能力对于保持成绩同样不具有显著的主效应，F（1，83）=.037，*p=.849*，η^2=.000。对于保持成绩，两者交互作用不显著，F（1，83）=.960，*p=.330*，η^2=.011。

有无虚拟同伴对于学习者的迁移成绩主效应不显著，F（1，83）=.058，*p=.811*，η^2=.001。共情能力对于迁移成绩同样不具有显著的主效应，F（1，83）=.356，*p=.553*，η^2=.004。对于迁移成绩，两者交互作用不显著，F（1，83）=2.532，*p=.115*，η^2=.030。

5.研究结论与讨论

5.1. 共情能力与虚拟同伴对于学习体验的影响

研究结果表明，在沉浸式虚拟现实中共情能力不同的学习者的学业情绪存在显著差异。共情理论认为，当学习者周围出现外部信息时，学习者对其有所知觉，并激活大脑中相应的情感部位产生表征共享（刘洋和朱立芳，2023）。共情能力较高的学习者在沉浸式学习过程中的积极情绪显著高于共情能力较低的学习者，这与先前的研究一致（向维 等，2021），可能是共情能力较高的学习者在虚拟环境中更能深入感知虚拟环境并与其内部知识产生深度交互，调动长期记忆中的社会图式从而获得了更愉悦的情感体验。虚拟同伴对于学业情绪的影响无显著差异，与先前研究不符，原因可能在于实验设计部分，本研究中的虚拟同伴并未设计积极情绪状态如微笑、音色、音调、鼓励词等功能。

研究结果表明，无论是内在动机还是外在动机，共情能力高的学习者均表现出更高程度的学习动机，与低共情能力的学习者具有显著差异。表明在虚拟现实环境中学习者的共情能力能够影响学习动机。共情反映着动机相关的潜在因素，动机在通过建立同理心中起着重要作用（Weisz et al., 2021），高共情能力者能够准确识别他人的情绪并将环境中的有效因素主动整合到自己认知内容中，产生更高的兴趣和认同感从而提升学习动机（McQuiggan et al., 2010）。虚拟同伴对于学习动机的影响无显著差异，这与先前研究不符（Liao et al., 2019）。原因可能在于，实验设计中的虚拟同伴主要涉及对于学习者学习过程中问题以及练习的反馈，并未设计深层次的交互方式，比如让学习者进行观察与模仿，促进分享意愿的功能（Schwartz et al., 2016）。

研究结果表明，共情能力较高的学习者在沉浸式虚拟现实学习过程中的自我效能感明显高于共情能力较低的学习者。这与先前研究一致，共情能力与社会自我效能感显著正相关（韦丽炎，2021）。IVR 环境下共情能力高的学习者具有更好的情境理解与虚拟角色

产生情感共鸣，对于虚拟同伴提供的反馈，共情能力高的学习者能够充分理解自己的不足并改进从而提升自己的能力水平，提高学习者的自我效能感。在沉浸式虚拟现实学习中有无虚拟同伴对于自我效能感没有显著差异。

研究结果表明，虚拟同伴在场时，学习者的社会存在感明显高于无虚拟同伴组。这与先前的理论假设相吻合，同龄虚拟同伴在场时的真实体验能够促进学习者高水平吸收，增加学习者对虚拟身份的认同从而提高存在感以及对虚拟同伴的信念感。共情能力对于社会存在感的影响不显著，这可能是学习过程中学习者的焦点在于课程内容以及交互方式上，对虚拟同伴或情绪状态分配注意较少。

5.2. 共情能力与虚拟同伴对于学习效果的影响

研究结果表明，不同共情能力的学习者在沉浸式虚拟现实学习中的学习效果并无显著差异，即尽管其在虚拟环境中的沉浸式体验以及对环境的感知内容有所不同，但并不会直接影响保持记忆与迁移运用。即使共情能力较弱的学习者，在虚拟环境中也能获得与共情能力较强的学习者相当的知识掌握与运用能力。此外，虚拟现实学习环境中虚拟同伴的存在与否，对学习效果的影响较小。原因可能在于实验中虚拟同伴主要起到陪同的作用，仅限于单向的语音与动作反馈，无双向互动性。根据之前的研究，在 VR 环境中学习者的视野、活动空间以及与其互动的人数是有限制的，学习者可能无法从同伴那里吸收太多信息，甚至将同伴的反馈作为学习过程中的干扰，使对学习效果的影响具有不确定性（Chen, 2016）。同时，虚拟同学的数量在 VR 学习环境也影响（Liao et al., 2019），当虚拟同学人数较多时，学习者无法及时匹配虚拟同伴与反馈信息，从而无法促进学习。

6.研究总结及启示

本研究利用实证研究方法，应用两因素实验设计探索了不同共情能力的学习者在虚拟环境中对于虚拟同伴是否在场的学习体验及学习效果的不同影响。研究表明，共情能力对学业情绪、学习动机和自我效能感具有显著影响。学习者共情能力越高，其积极情绪、学习动机水平和自我效能感也越强。此外，虚拟同伴在场能够更有效地激发学习者的社会存在感。本研究还存在部分局限，如未能事先將學習者以高低共情能力分組以保持各組間人數的平衡。在未来的研究与教学中，教育工作者应根据学习者的不同特征（如共情能力），合理选择沉浸式虚拟现实的教学形式与内容；研究者需要关注沉浸式虚拟学习环境中社会化因素对学习者的长期影响（例如，设计具有不同功能的虚拟同伴），以增强学习者在学习过程中对社会性因素的感知，培养充满热情、活泼、善良、具备社会责任感与正义感的社会人，而非沉浸在单机环境下虚拟空间中孤独的个体。

参考文献

马婧, 金妍池, 杨阳 赵章红.(2023).教育元宇宙:互联网教育新形态的系统建构与实践应用.现代教育技术(07), 99-107.

王钰莹.(2022).虚拟学习环境下大学生批判性思维发展的影响因素及优化策略研究博士学位论文, 东北师范大学).

王浩，唐爱民.(2023).虚拟现实更能引发学生共情？——基于 19 项随机对照实验的元分析.开放教育研究(01), 60-69.

韦丽炎.(2021).大学生共情能力、社会自我效能感与人际交往能力的关系及干预研究硕士学位论文, 云南师范大学).

向维,李文昊,李亚芳 陈冬敏.(2021).虚拟现实环境中反馈对学习效果的影响机制——基于共情的中介效应分析.现代教育技术(06),88-95.

刘洋，朱立芳.(2023).国外信息行为研究中的共情理论：现状与展望.知识管理论坛(04),316-328.

刘蕾.(2023).从“身体在场”到“多重在场”——人工智能发展下的“教”与“学”变革探讨.国家教育行政学院学报(03),17-20.

余越凡,周晓云 杨现民.(2022).基于元宇宙的线上线下融合（OMO）学习空间构建与教学模式设计.远程教育杂志(04),14-22.

周榕 郭佳瑞.(2023).虚拟现实技术与社会-情感学习融合的框架设计研究.远程教育杂志(03),94-103.

郑玲,刘革平,谢涛,陈娟菲 张可.(2021).协作学习中虚拟现实技术对学习效果的影响——2007—2019 年国际实证论文的元分析研究.中国远程教育(04),56-64.

高楠,刘革平,谢涛 张领.(2023).沉浸式虚拟现实对学习者知识迁移效果及效率的影响研究.远程教育杂志(01),91-101.

韩继宁,刘革平 王思玉.(2023).元宇宙中社会化交互学习的特征与发展路向.现代远程教育研究(02),57-66.

詹泽慧.(2014).面授与远程学习者社会存在感之作用差异研究——以美国高校学生为样本.中国电化教育(02),35-39.

Barreda-Ángeles, M., Aleix-Guillaume, S., & Pereda-Baños, A. (2020). An “empathy machine” or a “just-for-the-fun-of-it” machine? Effects of immersion in nonfiction 360-video stories on empathy and enjoyment. Cyberpsychology, Behavior, and Social Networking, 23(10), 683-688.

Biocca, F., & Harms, C. (2002). Defining and measuring social presence: Contribution to the networked minds theory and measure. Proceedings of PRESENCE, 2002, 7-36.

Brydon, M., Kimber, J., Sponagle, M., MacLaine, J., Avery, J., Pyke, L., & Gilbert, R. (2021). Virtual reality as a tool for eliciting empathetic behaviour in carers: an integrative review. Journal of Medical Imaging and Radiation Sciences, 52(3), 466-477.

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. Computers & Education, 102, 152-171.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy.

Dyer, E., Swartzlander, B. J., & Gugliucci, M. R. (2018). Using virtual reality in medical education to teach empathy. Journal of the Medical Library Association: JMLA, 106(4), 498.

Huang, J. Y., Hung, W. H., Hsu, T. Y., Liao, Y. C., & Han, P. H. (2019). Pum** life: Embodied virtual companion for enhancing immersive experience with multisensory feedback. In SIGGRAPH Asia 2019 XR (pp. 34-35).

Isen, A. M., & Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behavior, and self-control. Motivation and emotion, 29, 295-323.

Liao, M. Y., Sung, C. Y., Wang, H. C., & Lin, W. C. (2019, March). Virtual classmates: Embodying historical learners' messages as learning companions in a VR classroom through comment map**. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 163-171). IEEE.

McQuiggan, S. W., Robison, J. L., & Lester, J. C. (2010). Affective transitions in narrative-centered learning environments. Journal of Educational Technology & Society, 13(1), 40-53.

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. Computers in Human Behavior, 72, 276-285.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. Educational Psychology Review, 1-22.

Ning Shen, K., & Khalifa, M. (2008). Exploring multidimensional conceptualization of social presence in the context of online communities. Intl. Journal of human–computer interaction, 24(7), 722-748.

Palloff, R. M., & Pratt, K. (2007). Building online learning communities: Effective strategies for the virtual classroom. John Wiley & Sons.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. IEEE transactions on pattern analysis and machine intelligence, 23(10), 1175-1191.

Ryan, R. M., & Deci, E. L. (2001). To be happy or to be self-fulfilled: A review of research on hedonic and eudaimonic well-being. Annual review of psychology, 52(1), 141-166.

Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. Handbook of motivation at school, 2, 34-54.

Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. Motivation and emotion, 41, 708-712.

Shen, K. N., Yu, A. Y., & Khalifa, M. (2010). Knowledge contribution in virtual communities: accounting for multiple dimensions of social presence through social identity. Behaviour & Information Technology, 29(4), 337-348.

Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. Journal of counseling psychology, 63(1), 32.

Villalba, É. E., Azócar, A. L. S. M., & Jacques-García, F. A. (2021). State of the art on immersive virtual reality and its use in develo** meaningful empathy. Computers & Electrical Engineering, 93, 107272.

Weisz, E., Ong, D. C., Carlson, R. W., & Zaki, J. (2021). Building empathy through motivation-based interventions. Emotion, 21(5), 990.

Yuan, Q., & Gao, Q. (2023). Being There, and Being Together: Avatar Appearance and Peer Interaction in VR Classrooms for Video-Based Learning. International Journal of Human–Computer Interaction, 1-21.

B 站高传播力编程类课程的多组态路径分析

Multi-configuration Path Analysis of High Transmissibility Programming Course in Station B

洪越洋[1]，单俊豪[2*]
[1,2]南京邮电大学教育科学与技术学院
[*] jhshan@njupt.edu.cn

【摘要】 对在线课程的质量影响因素和条件组态进行分析有助于在线课程的发展和优化。本研究从课程质量的视角出发，从教学媒体、课程内容和课程管理三个维度进行样本案例剖析，引入传播力将样本相关数据量化作为衡量在线课程优劣的指标，最后用定性比较分析法（QCA）进行因果条件探究，得到课程内容是影响课程质量的必要条件以及条件组态路径，依据这一结果，本研究提出了四点在线课程的发展和优化建议。
【关键词】 在线课程；传播力；层次分析法；定性比较分析法

Abstract: *The analysis of quality factors and condition configuration of online courses is helpful to the development and optimization of online courses. From the perspective of course quality, this study analyzes the sample cases from three dimensions: teaching media, course content and course management. It introduces the communication power to quantify the relevant data of the samples as an indicator to measure the quality of online courses. Finally, qualitative comparative analysis (QCA) is used to explore the causal conditions. It is concluded that the content of the course is the necessary condition and the condition configuration path that affect the quality of the course. According to the result, this study puts forward the development and optimization suggestions of four-point online course.*
Keywords: Online course, Transmissibility, Analytic hierarchy process, Qualitative comparative analysis

1. 引言

教育信息化变革推动了在线课程的持续发展，但同时在线课程的质量问题也逐渐显现出来（李爽、李梦蕾和赵宏，2020），如课程完成率较低（王淑娟、沈继红和朱磊，2019）和教学设计不合理（李爱霞，2017）等。在线课程的质量直接影响了在线学习者的学习效果，也是影响学习者在线学习粘性的核心要素（苗冬玲、吴昭和闫寒冰，2021），因此有必要探究在线课程质量的影响因素和课程优化途径。

用户选择是在线开放学习的重要方式，也是在线课程优胜劣汰的重要手段，新媒体技术的迭代升级更是加快了课程更新换代的速度。传播力是对一定覆盖范围内的目标受众造成潜在影响的能力（沈正赋，2016），在线课程作为内容特殊的信息传播形式同样拥有这一特性。传播力高意味着观看过这门课程的所有学习者中受到较大影响的学习者所占比例大，以此可以反映课程内容的受欢迎程度，因此课程传播力的高低可以作为衡量在线课程优劣的一个重要指标，结合这一指标就可以从在线学习者的体验视角来反馈在线课程质量。

本研究从课程质量的视角出发，以在线学习平台 B 站的编程类在线课程为例，选取合适的样本，运用定性比较分析的方法从课程质量评价量规出发归纳出高传播力在线课程的具体特征和组合路径，最终总结出在线课程设计和优化的普适性建议。

2. 文献综述

在线教育的发展过程现已从追求数量逐步转为追求质量，不少学者针对提升在线课程质量展开研究，通过对文献的阅读，发现已有研究多是基于MOOC平台或翻转课堂，关于如B站这种短视频平台上的知识类视频的研究较少。

2.1. 理论研究

提升质量的前提是评判质量，李青等根据国内外权威课程评价指标总结出课程质量影响因子，结合案例与实践从媒体技术、课程内容和课程管理三方面建立课程质量保证体系（李青和刘娜，2015）；滕霞等结合国内外分析和相关评价标准提出网络课程的初步指标体系之后用德尔菲法对制定的标准进行评议修正（滕霞、段崇江和张建伟等，2003）。除质量指标体系的建立外，还有学者以此为基础建立MOOC评价模型来更好地评价课程优劣，例如姚凯、李思志等以“MOOC学院”为数据源，基于文本挖掘技术、Vague集的相似度量法和AHP权重确定建立的评价模型（姚凯、李思志和李艳红等，2017）。由此可见，针对MOOC等在线开放平台课程的质量评价体系已经相对成熟。对在线课程传播机制的研究方面，有学者在研究中提出质量和内容是影响短视频传播力的关键因素（巫霞和马亮，2019），并且点赞、转发等行为能反映出用户对信息的接受和认可程度（李静，2016），可见传播力高低与在线课程质量成正比。因此本研究在之前研究的基础上结合B站在线课程的特点将质量评价体系进行修改，形成适用于B站在线课程的质量评价量规，并且将传播力作为结果变量，影响课程质量的诸多客观因素作为条件变量进行QCA分析得出高传播力在线课程的传播路径和内在逻辑。

2.2. 分析技术研究

目前的在线课程质量评价体系多是基于已有的指标或课程实例进行质性分析，缺少量化指标的支持。在传播学中，金心怡等运用熵权法量化抖音科普短视频的传播力（金心怡和王国燕，2021），张舒涵等通过构建短视频传播影响力指数，综合多个维度测算传播影响力，从而分析传播效果的影响因素（张舒涵、孔朝蓬和孔婧媛，2021）。本研究综合传播学的有关理论，将B站在线课程质量评价与传播力相结合，利用层次分析法（AHP）对B站视频传播核心影响因素设置权重，构建课程传播力指数计算公式来衡量课程传播水平。后在在线课程质量评价量规和视频传播力的基础上借助定性比较分析方法的中观理论建构方法进行研究，针对在线课程的高质量发展与改进提出建议。

3. 研究设计

3.1. 案例选取

B站提供了广泛的编程类课程，涵盖了多种编程语言和技术领域，为减少非研究变量对结果的干扰，本研究根据2023年12月TIOBE编程语言排行榜，选取了排名前三的面向对象编程语言对应的课程，分别是Python、C++和Java。本研究采用定性分析比较法进行高传播力视频路径探究，该方法适用于在传播力、影响力方面具有代表性的中小样本研究，且样本需具有异质性（武梦迪和冯锐，2022）。基于此共选取了28门编程类课程，其中Python语言10门，C++语言8门，Java语言10门，每一门语言都根据播放量和点赞量初步划分为正面案例和反面案例（Rihoux B & Ragin C C，2006）。

3.2. 变量赋值

本研究以在线课程质量量规（Quality Matter Rubric）为基础，结合B站编程类课程的特点和在线课程质量因子，将量规围绕B站课程进行个性化的补充优化，形成了围绕“教学媒体”、“课程内容”和“课程管理”三个维度，涵盖“教学媒体”、“课程概述与简介”、“教学设计”、“教学内容”、“交流与交互”、“项目化管理”、“学习支持”七个一级指标和29个二级指标的B站课程质量评价量规，见图1。

教学媒体（Teaching Media）		文字大小合适、图片清晰、排版美观，搭配协调
		视频材料应具有较高的清晰度和流畅性，画面构图美观
		提供同一视频的不同清晰度版本，以适应不同的带宽和设备访问
		视频内容应尽量使用图表代替大段的文字信息
		呈现教师头像或视频形象
		教师面对镜头表现自然，与镜头有眼神交流
		教学视频应该短小紧凑，长度以 10 分钟左右为佳
		为使用者提供一定限度的技术支持服务，例如编程环境配置等
课程内容（Course Content）	课程概述与简介（Course Introduction）	提供教师和教学团队的介绍
		简单介绍课程的主要内容及课程教学目标
		明确说明学习者应具备的基本知识和最低技术要求
	教学设计（Teaching Design）	面向广大学习者设计，考虑到学习者之间的差异
		提供清晰明了的教学大纲
		列出课程的学习目标与单元目标，并且目标是可测量的
		详细描述课程重点和难点，告知学习者遇到问题时的解决方法
		每周学习时长分配合理
	教学内容（Teaching Content）	资源和材料能够满足学习目标和教学要求
		教学内容具有足够的广度、深度，保证正确性和实时性
		教学内容的组织具有逻辑性，知识点安排合理
		课程中使用的所有材料和资源是合法、道德的
	交流与交互（Teaching Interactivity	规定课程的主要讨论方式，并提供其它补充渠道
		定期答疑，并和学生在线交流
		除教师外还聘请助教协助答疑解惑、参与讨论、批改作业
课程管理（Course Management）	项目化管理（Project	制定管理规定，规范在线课程整体规划与开发
		采用团队化的方式组织开发，并聘用专业的美术和制作人员
	学习支持（Learning Support）	支持任何学习者（ 包括残障人士） 无障碍地访问课程
		提供可替代视频和音频的文字内容
		为学生有效使用课程资源和工具提供帮助
		说明学习过程会遇到的常见问题及解决方案

图 1　B 站课程质量评价量规

围绕质量评价量规对所有案例课程按照一定标准进行打分，符合定义记为 1，反之记为 0，最终形成能够描述所有案例课程质量的二分赋值表。

3.3. 研究方法

3.3.1. 层次分析法确定传播力指标权重

在对传播力大小进行衡量的时候，涉及到点赞、收藏、转发、播放、投币等多个因素，但衡量时多依赖人的主观判断，降低了结果的可信度和科学性，因此需要对影响指标进行权重划分，将定性因素予以量化。本文选择层次分析法（Analytic hierarchy process，AHP）作为权重计算方法。

层次分析法是美国教授 T.L.Saaty 提出的一种把定性和定量相结合的多目标决策分析方法（廖红强、邱勇和杨侠等，2012），其优点在于将复杂系统的决策思维层次化（朱茵、孟志勇和阚叔愚，1999）。该方法将问题拆分成多个影响因素，然后将各因素的重要性两两相比，结合判断矩阵、排序计算和一致性检验等方法最终得到可靠的各影响因素的权重。本文参照 AHP 的要求，先将 B 站编程类在线课程传播力的影响因素划分成点赞、收藏、转发、播放、投币五个方面，然后结合 Saaty 提出的 1-9 标度法设计了 B 站编程类在线课程传播力指标权重打分表，并邀请五位相关领域的专家进行打分，最后将打分结果进行计算和一致性检验，最终得到五个因素对应的影响权重。

3.3.2. 以组态视角进行定性比较分析

定性比较分析（Qualitative Comparative Analysis，QCA）是一种以集合论和布尔代数为根本理论支持的因果关系探究方法，将研究对象视作条件变量不同组合情况下的组态，融合案例研究和变量研究，能用于回答较为复杂的因果问题（毛湛文，2016）。本研究探究的是 B 站高传播力编程类课程背后的影响因素，课程质量影响因素是条件变量，课程的传播力高低是结果变量，因此选择适用于二分赋值变量的清晰集定性比较分析方法（csQCA）来进行 B 站高传播力优质编程类课程背后质量因素多组态路径分析。依据传播力计算结果来划分所选择课程案例的质量，将计算所得传播力高于算术平均值的课程标记为高质量课程。具体步骤包括选取案例确定变量——案例编码——原始数据处理——必要性分析——构建真值表——

条件组态分析。

3.4 权重确定

3.4.1. 建立判断矩阵

将需要进行两两比较的传播力影响因素分别设为 u_i, u_j（i，j=1,2,3,4,5），相对权重值 u_{ij} 表示因素 u_i 对因素 u_j 的重要程度（申志东，2013），取值根据 1-9 标度法，具体定义如表 1 所示。

表 1 判断矩阵标度及其定义

标度	定义（重要性比较）
1	因素 u_i 和因素 u_j 相比具有同等重要性
3	因素 u_i 相较于因素 u_j 稍微重要
5	因素 u_i 相较于因素 u_j 比较重要
7	因素 u_i 相较于因素 u_j 非常重要
9	因素 u_i 相较于因素 u_j 绝对重要
2,4,6,8	介于相邻判断的两个标度之间，取中值
倒数 1/v（v 为 1-9）	若因素 u_i 和因素 u_j 的重要性之比为 v，那么因素 u_j 和因素 u_i 的重要性之比为 1/v

结合判断矩阵的标度及其定义设计专家打分表并对五位专家的打分结果进行收集，通过计算几何平均值得到最终判断矩阵如表 2 所示。

表 2 最终判断矩阵

	点赞量	转发量	收藏量	播放量	投币量
点赞量	1.000	0.334	0.225	4.425	0.164
转发量	2.993	1.000	0.748	5.376	0.485
收藏量	4.441	1.337	1.000	5.435	0.243
播放量	0.226	0.186	0.184	1.000	0.154
投币量	6.109	2.063	4.112	6.507	1.000

3.4.2. 指标权重计算

层次分析法计算各因素的权重的方法有几何平均法（方根法）、算术平均法（和积法）、特征根法、最小二乘法等（邓雪、李家铭和曾浩健等，2012），综合 B 站视频的特征和各计算方法的可信度、计算量，最终选择和积法对专家打分数据进行处理。对 3.1 中判断矩阵各列进行归一化处理，按行相加后得到的新向量重新进行归一化处理，最后得到 B 站编程类在线课程传播力各影响因素的权重如表 3 所示。

表 3 各影响因素的权重

指标	权重
点赞量	0.09
转发量	0.20
收藏量	0.22
播放量	0.04
投币量	0.45

3.4.3. 一致性检验

由于实际专家在对指标进行两两比较使会出现不一致的结论，因此需要对得到的权重结果进行一致性检验以保证指标权重的合理性。Saaty 提出将一致性指标（CI）与随机一致性指标值(RI)之比作为 AHP 一致性检验的衡量指标。本文利用特征向量计算得到 CI 值为 0.092，本研究构建五阶判断矩阵，对比随机一致性 RI 表格得到 RI 值为 1.12，随机一致性 RI 指标表格如表 4 所示。根据公示计算得到一致性指标 CR 值为 0.082<0.1，因此本次研究判断矩阵满足一致性检验，计算所得权重具有一致性。

$$CI = \frac{\lambda_{max} - N}{N-1} \quad ; \quad CR = \frac{CI}{RI}$$

其中 λmax 为判断矩阵的最大特征值 ，N 为判断矩阵的阶数

表 4 随机一致性 RI 指标

n 阶	3	4	5	6	7	8	9
RI 值	0.52	0.89	1.12	1.26	1.36	1.41	1.46

由于每个视频的点赞量等指标数值差异大，直接计算会产生较大的数据波动和较高的异方差性，因此参考已有传播力研究中利用自然对数将数据进行标准化的方法（王林、潘陈益和朱文静等，2018），结合权重得到 B 站编程类课程传播力 W 的计算方法如公式（1）所示：

$$W = a_1 Ln(X_1+1) + a_2 Ln(X_2+1) + a_3 Ln(X_3+1) + a_4 Ln(X_4+1) + a_5 Ln(X_5+1) \quad (1)$$

其中 X1、X2、X3、X4、X5 为该编程课程的点赞量、转发量、收藏量、播放量和投币量，a1、a2、a3、a4、a5 为对应的权重值。

4. 研究结果

4.1. 必要条件分析

将统计数据进行编码和基本处理之后依次对单个条件变量进行分析，判断其与结果变量之间的关系，利用一致性（Consistency）和覆盖度（Coverage）两个指标来探究现有条件变量中是否有导致传播力高这一结果存在的必要条件，本研究将一致性大于等于 0.9 的条件变量视作结果变量的必要条件（张明和杜运周，2019），覆盖度则代表成功案例中拥有这一条件变量所占有的比例，具体分析结果如表 5。

表 5 必要条件分析结果

	Consistency	Coverage
教学媒体	0.416667	0.357143
课程概述与简介	0.583333	0.583333
教学设计	0.416667	0.333333
教学内容	1.000000	0.444444
交流与交互	0.083333	0.333333
项目化管理	0.333333	0.444444
学习支持	0.500000	0.600000

结果显示只有教学内容这一条件变量的一致性大于 0.9，构成导致编程类课程高传播力

的必要条件，同时其覆盖度低于0.8，意味着它变量不能单独成为一个影响指标，还需要进一步分析。

4.2. 条件组态分析

通过构建真值表在fsQCA软件中进行多因素组合分析，得到高传播力B站编程类课程的组态路径如表6：

表 6 组态路径

条件	组态解				
	组态解1	组态解2	组态解3	组态解4	组态解5
教学媒体	⊗	⊗	●	●	●
课程概述与简介	●	●	●	●	⊗
教学设计	●	⊗		⊗	⊗
教学内容	●	●	●	●	●
交流与交互	⊗	⊗	⊗		⊗
项目化管理	⊗		●	●	⊗
学习支持		●	⊗	⊗	●
原始覆盖率	0.167	0.167	0.167	0.167	0.083
唯一覆盖率	0.167	0.167	0.083	0.083	0.083
一致性	1	1	1	1	1
总体覆盖率			0.667		
总体一致性			1		

注：●和●表示该条件存在；⊗和⊗表示该条件不存在；空白表示该条件可以存在也可以不存在

通过对比分析，可以得到B站编程类在线课程高传播力的四条组态路径：

（一）内容为王型：课程概述与简介+教学设计+教学内容

这一组态路径涵盖了课程的整体结构设计，通过简介提供对整个学习体系的总览，帮助学生明确导向和目标，结合新颖独特的教学方法和教学流程保障学生在课程中能够有所进步。这一路径旨在通过有序的组织和设计，使学生在课程中能够全面、深入地理解和掌握所学内容，达到教学目标。

（二）学习支持型：课程概述与简介+教学内容+学习支持

这一组态路径着眼于课程全局，课程概述与简介提供了整体框架，在此基础上专注于提供高质量、互动性强的学习材料支持，这些支撑材料不仅需要涵盖理论知识，还应包括实际应用案例和互动练习，以促进学生的深入理解和技能掌握，如项目化程序教学时，会提供完整源代码和关键代码注释。此外，利用B站这一在线平台的优势，教学内容可以多种形式呈现，此外还能通过有效的学习支持帮助学生克服在线学习中可能遇到的障碍，提高学习效率。

（三）系统化管理型：教学媒体+课程概述与简介+教学内容+项目化管理

这一组态路径注重教学设计整体框架和教学内容的综合考量，同时重视教学媒体的多样

化运用以及课程开发全过程的项目化管理。通过运用多样化的教学媒体，学生的学习体验得以丰富和多元化，有助于激发学生的学习兴趣和提高学习效果。在课程开发的全过程中，采用项目化管理的方法，即制定管理规定，聘请专业人员组建专业团队，并细化团队分工，以确保课程开发过程的高效运作和质量保障。这种综合运用教学媒体和项目化管理的方法，有助于提升编程类课程整体教学质量和学习效果，为学生的综合能力提升提供了有力支持。

（四）综合学习体验型：教学媒体+教学内容+学习支持

这一组态路径综合考虑了教学媒体、教学内容和学习支持，旨在为学生提供全面且有深度的学习体验。通过多样化的教学媒体激发学生的兴趣和参与度，同时深入的教学内容保障基础知识的讲解，学习支持关注学生可能面临的问题并提供解决方案。这一路径专注于打造一个综合性的学习环境，培养学生发现问题并解决问题的综合学科能力。

5. 研究结论

结合案例和理论研究可以得出B站学习视频的传播力高低受到多种因素影响，经过定性比较分析可以得到在线课程高传播力的四条影响因素典型组合路径。综合上述分析可以得出以下在线课程的发展和优化途径，从而更好地进行课程制作和传播。

（一）深化教学内容，优化教学设计

教学内容是课程类视频的核心，应从一定的理论基础出发，拥有足够的广度、深度。B站这一平台的特点势必会导致B站在线课程知识点的碎片化，因此在进行教学设计时应注重课程内容的完整性和逻辑性（顾晓薇、胥孝川和王青，2020）。

除此之外，研究中还发现教学结果评估是大多数课程未曾关注的部分，针对学习效果应采用多样化的评估方法，如项目式编程、在线测验和自评互评，多角度衡量学生的学科表现和能力发展。

（二）技术赋能，让教学媒体用得其所

在线课程本质上是课程录制，即将课本上晦涩的文字和代码经过讲师的讲解和计算机即时结果反馈，利用屏幕呈现给学生，在这一过程中技术就扮演着关键角色。例如可以结合AR和VR模拟真实问题情景从而提升学习者实际编程应用能力；可以通过学习分析工具实时追踪学生的学习进展，为教师和学习者提供有针对性的反馈和支持，从而提高学习效果；重视多媒体画面语言表征目标的原则和运用，促进深度学习（刘哲雨、侯岸泽和王志军，2017）。

（三）引入项目化管理与实践性学习

引入项目化管理与实践性学习，将理论知识与实际应用有机结合。通过组织学生开展实际项目，学生能够将所学编程理论结合真实场景进行应用和拓展，强化理论知识的同时还能培养学生解决问题的能力和团队协作技能，帮助学生全面发展，为未来在编程领域的职业发展奠定坚实基础。

（四）构建多元化学习支持服务团队

学习支持包括学习资源的提供、学习群体的分析、反馈信息的汇总分析和基本技术支持等，因此优质在线课程打造的背后应具备一个“跨功能团队”，包含教学总监、课程设计师、课程评估专家等成员以保证课程的管理、设计和分析优化，通过为学习者提供全面的服务来满足不同学习者不同层面的需求，促进更全面、个性化的在线学习体验。

高传播力的在线课程不仅需要高质量的教学内容，还要与规范化管理相结合，开发过程中务必警惕“重形式轻内容”，确保课程管理贯穿于课前、课中和课后的各个环节。在确保基础课程内容的前提下需要着眼于内容呈现形式，通过多种教学媒体的巧妙组合，推动项目学习和实践学习的多样化表达。这种全方位的管理和注重内容质量的策略，有助于在线课程的质的提升，提高学习效果和传播力。

致谢

*本文受2021年江苏省教育科学“十四五”规划青年重点课题“中职教师信息技术应用能力精准测评及提升微策略研究”与2022年江苏省高校哲学社会科学一般课题“微认证驱动的高校本科课程评价变革研究”支持。

参考文献

李爽,李梦蕾 赵宏.(2020).在线课程质量观和质量要素的质性研究——基于专家、实践者和学习者的视角.中国远程教育(03),42-50+81.
王淑娟,沈继红 朱磊.(2019).基于在线开放课程平台数据对MOOC完成率计算的思考.智库时代(19),216-217.
李爱霞.(2017).MOOC学习者粘性影响因素研究硕士学位论文,吉林大学)
苗冬玲,吴昭 闫寒冰.(2021).全视角学习理论视域下的在线学习粘性影响因素研究.中国远程教育(10),68-75.
沈正赋.(2016).新媒体时代新闻舆论传播力、引导力、影响力和公信力的重构.现代传播(中国传媒大学学报)(05),1-7.
李青 刘娜.(2015).MOOC质量保证体系研究.开放教育研究(05),66-73.
滕霞,段崇江,张建伟,王学优.(2003).网络课程评价标准的研制.现代教育技术(01),5-11.
姚凯,李思志,李艳红 邱静静.(2017).MOOC评价模型研究.复旦教育论坛(03),65-71.
巫霞 马亮.(2019).政务短视频的传播力及其影响因素:基于政务抖音号的实证研究.电子政务(07),22-30.
李静.(2016).微信点赞背后的情绪与文化表达.新媒体研究(16),12-13.
金心怡 王国燕.(2021).抖音热门科普短视频的传播力探析.科普研究(01),15-23+96.
张舒涵,孔朝蓬 孔婧媛.(2021).新媒体时代短视频信息传播影响力研究.情报科学(09),59-66.
武梦迪 冯锐.(2022).B站美食短视频走红的多组态路径研究.美食研究(04),37-46.
廖红强,邱勇,杨侠,王星刚 葛任伟.(2012).对应用层次分析法确定权重系数的探讨.机械工程师(06),22-25.
朱茵,孟志勇,阚叔愚.(1999).用层次分析法计算权重.北方交通大学学报(05),119-122.
毛湛文.(2016).定性比较分析(QCA)与新闻传播学研究.国际新闻界(04),6-25.
申志东.(2013).运用层次分析法构建国有企业绩效评价体系.审计研究(02),106-112.
邓雪,李家铭,曾浩健,陈俊羊 赵俊峰.(2012).层次分析法权重计算方法分析及其应用研究.数学的实践与认识(07),93-100.
王林,潘陈益,朱文静 邓胜利.(2018).机构微博传播力影响因素研究.现代情报(04),35-41.
张明 杜运周.(2019).组织与管理研究中QCA方法的应用:定位、策略和方向.管理学报(09),1312-1323.
顾晓薇,胥孝川 王青.(2020).国家精品在线开放课程建设研究.现代教育管理(06),77-83.
刘哲雨,侯岸泽 王志军.(2017).多媒体画面语言表征目标促进深度学习.电化教育研究(03),18-23.
Rihoux B , Ragin C C . (2006).Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques[M].
https://tiobe.p2hp.com/

基于大语言模型的数字故事对学生批判性思维、认识论信念的影响研究

A study of the impact of digital storytelling on students' critical thinking and epistemological beliefs based on the large language models

周雨欣[1*]，刘泽民[2]，张文宇[1]，单小份[1]
[1]温州大学教师教育学院
[2]华东师范大学教育信息技术学系
*1812484976@qq.com

【摘要】 信息过载时代培养学生批判性思维与认识论信念至关重要，数字故事是培养这些能力的有效策略，然而传统数字故事难以为学生提供即时和个性化的反馈。因此，本研究结合大语言模型强大的生成能力来弥补传统数字故事的不足，旨在探究基于大语言模型的数字故事对学生批判性思维、认识论信念的影响。同时借助移动学习的优势，学生通过移动设备参与课程。通过设计基于大语言模型的数字故事课程，并进行了一项随机对照实验，结果显示大语言模型组的学生批判性思维与认识论信念都取得了显著进步，并且批判性思维与认识论信念的提升都显著优于对照组。

【关键词】 大语言模型；数字故事；批判性思维；认识论信念；移动学习

Abstract: *Cultivating students' critical thinking and epistemological beliefs in the age of information overload is crucial, and digital storytelling is an effective strategy for developing these abilities; however, traditional digital storytelling is difficult to provide students with immediate and personalised feedback. Therefore, this study combines the powerful generative power of the Large Language Models to compensate for the shortcomings of traditional digital stories, aiming to investigate the impact of the Large Language Models-based digital stories on students' critical thinking and epistemological beliefs. It also takes advantage of mobile learning, where students participate in the course through mobile devices. By designing a digital storytelling course based on the Large Language Models and conducting a randomised controlled experiment, the results show that students in the Large Language Models group made significant progress in both critical thinking and epistemological beliefs, and that both critical thinking and epistemological beliefs improved significantly better than those in the control group.*

Keywords: Large Language Models, Digital storytelling, Critical thinking, Epistemological beliefs, mobile learning

1. 引言

随着数字技术的迅猛发展，我们迎来了信息过载的时代。在这个时代，学生在处理大量不同的信息时面临更多的挑战(刘永谋 et al., 2023)，这意味着学生需要形成对知识本质和认知本质的坚定不移的态度和观念，以便在面对信息时能进行独立、全面和深入的思考(林青青, 2022; 席静 et al., 2018)。因此，培养学生的批判性思维能力与认识论信念变得尤为重要。教学是培养学生批判性思维与认识论信念的主要途径(Valanides & Angeli, 2005)，然而传统教学模式缺乏使用技能的真实情景(谢晓宇, 2023)，难以真正促进这些能力的发展，需要在教学中加入有效的教学策略。数字故事（Digital Storytelling, DST）是一种有效的方法，有助于学习信息，提高学生的兴趣和动力(Yang & Wu, 2012)。但是许多学生在学习时会遇到疑惑和困难，而传统的教学模式往往难以为每个学生提供即时和个性化的反馈(Jeon & Lee, 2023)，导致学生学习思维陷入局限。大语言模型（Large Language Models，LLMs）强大的

语言生成能力可以为 DST 教学增添交互性和个性化元素，与学生进行实时对话，提供反馈(Wu et al., 2023)，并引导他们思考和探索更深层次的问题(Kasneci et al., 2023)。本研究设计了一项随机对照试验旨在探究基于 LLMs 的 DST 对学生批判性思维以及认识论信念的影响，研究问题如下：

问题 1：相比于传统 DST，基于 LLMs 的 DST 能否更加有效的提升学生批判性思维？

问题 2：相比于传统 DST，基于 LLMs 的 DST 能否更加有效的提升学生认识论信念？

2. 相关概念和研究现状

2.1. 数字故事（Digital storytelling）

Jing Wu (2020)提出教育数字故事（Education Digital Storytelling，EDS）的五个方向（专有导向、代理导向、反思导向、重建导向和反射导向），重建导向 DST 通常涉及重建给定概念含义的过程，高度注重培养参与者的批判性。但是目前关于重建导向 DST 的研究较少。并且在重建导向的数字故事中，学生在创作过程中得不到及时的反馈提示与反思评价，容易导致创作的故事情节平淡，缺乏批判性与反思性(池金章, 2019)。并且传统数字故事教学内容单一，无法给学生提供个性化支持(邱艳萍, 2014)。鉴于这些局限性，越来越多的学者开始关注使用技术特别是人工智能辅助的数字故事来应对这些挑战，LLMs 的应用或许可以为此提供一定的支持。

2.2. LLMs 在教育领域的应用

目前，LLMs 成了自然语言处理领域的重要技术，对文本生成、自动翻译、情感分析、问答系统等任务产生了巨大影响(焦建利 & 陈婷, 2023)。许多学生在学习时会遇到疑惑和困难，而传统的教育模式往往难以为每个学生提供即时和个性化的反馈(Jeon & Lee, 2023)。每个学生的学习需求和背景都是独特的，而传统的教材往往是固定的，不能满足所有学生的需求，LLMs 可以用于创建个性化的学习内容(Kasneci et al., 2023)。LLMs 融入 DST 能够引导学生参与更深层次的对话和探讨，在创作过程中给予学生及时反馈，但具体如何融入及其效果还需要进一步的实证探索。

3. 研究方法

3.1. 实验设计

本研究设计基于重建导向的 DST 课程，具体分为五部分，见表 1。

表 1 LLMs 在 DST 中的干预

步骤	主题	具体内容	LLMs 功能
一	了解 DST 及选题	了解 DST 并为 DST 选择一个科学概念现象作为故事主题	LLM 可以通过提供提示来指导学生进行研究来。例如，它可以建议要探索的问题或要检查的资源。
二	研究与批判性分析重建	使用教师提供的可靠来源研究该主题、识别与该主题相关的常见误解或刻板印象、根据他们的研究分析这些误解或刻板印象并且分析重建他们对主题的理解	LLM 可以提供鼓励批判性思维的提示。例如，它可以提出挑战现有信念或鼓励学生考虑不同观点的问题。
三	故事板的创作	计划 DST，决定故事事件的顺序、人物和叙述，完成故事板	LLMs 提供关于故事板的基本原理和设计原则的指导，帮助学生规划

			DST 的结构、确定故事事件的顺序以及选择适当的人物和情节。
四	DST 创作	根据故事板，利用 Puppet Pals 2 创建他们的 DST，用文本、图像、音频和视频来传达他们的故事	
五	分享与反思	与他人分享自己的 DST 并反思他们的学习过程	LLM 可以通过提供提示来指导学生进行反思来。例如，它可以询问他们学到了什么或者他们的理解发生了怎样的变化。

3.2. 实验流程

本实验在小学六年级科学社团课课堂上开展，实验以不同的教学策略（基于 LLMs 的 DST、传统 DST）为自变量，以学生的批判性思维和认识论信念为因变量，实验前通过批判性思维量表(Lin et al., 2019)和认识论信念量表(Conley et al., 2004)对学生进行前测。在 10 课时实验结束后进行后测，收集学生作品。

4. 结果

4.1. 批判性思维

为了比较干预对学生批判性思维的影响，对实验组和对照组的前后测批判性思维分别进行了配对样本 t 检验（见表 2），结果表明，在实验组使用基于 LLMs 的 DST 后，批判性思维水平得到显著提高，对照组使用传统的 DST 后，批判性思维水平也得到了一定提高。

表 2 批判性思维前、后测配对样本 *t* 检验分析

Dimension	Group	Mean（SD）		*t*	*d*
		Critical_pre	Critical_post		
Critical	实验组	3.69(0.563)	4.31(0.489)	-5.65**	0.983
	对照组	3.73(0.594)	3.96(0.436)	-1.83	0.314

*p<0.05 **p<0.001

为了对比 LLMs 对学生批判性思维的影响，以组别为自变量，后测分数为因变量对实验组和对照组进行独立样本 t 检验（见表 3）。结果证明，基于 LLMs 的 DST 在提高学生批判性思维方面比传统 DST 更有帮助。

表 3 批判性思维后测独立样本 *t* 检验分析结果

Dimension	Group	Mean	SD	*t*	*d*
Critical_post	实验组	4.31	0.489	3.113*	0.75
	对照组	3.96	0.436		

*p<0.05 **p<0.001

4.2. 认识论信念

为了对比干预前后学生认识论信念的变化，对实验组和对照组前后测认识论信念分别进行了配对样本 t 检验（见表 4），结果表明，在实验组使用基于 LLMs 的 DST 后，认识论信念总体水平得到显著提高，在对照组使用传统的 DST 后，认识论信念总体水平也得到了一定提高。

表 4 认识论信念前、后测配对样本 t 检验分析

Dimension	Group	Mean（SD）		*t*	*d*
		Pre	Post		

Source	实验组	3.53（0.740）	4.02（0.630）	-3.26*	0.567
	对照组	3.57（0.685）	3.66（0.750）	-0.730	0.125
Certainty	实验组	3.38（0.677）	3.92（0.600）	-3.53*	0.973
	对照组	3.35（0.689）	3.44（0.706）	-0.767	0.563
Development	实验组	3.73（0.637）	4.25（0.465）	-4.13**	0.212
	对照组	3.70（0.629）	3.82（0.635）	-0.944	0.45
Justification	实验组	3.87（0.586）	4.29（0.437）	-3.70**	0.644
	对照组	3.90（0.571）	4.19（0.421）	-3.214*	0.551

*p<0.05 **p<0.01

为了对比 LLMs 对学生认识论信念的影响，以组别为自变量，后测分数为因变量对实验组和对照组进行独立样本 t 检验（见表 5）。结果证明，基于 LLMs 的 DST 在提高学生认识论信念方面比传统 DST 更有帮助。

表 5 认识论信念后测独立样本 *t* 检验分析结果

Dimension	Group	Mean	SD	*t*	*d*
Source_post	实验组	4.02	0.630	2.160*	0.52
	对照组	3.66	0.750		
Certainty_post	实验组	3.92	0.600	3.030*	0.73
	对照组	3.44	0.706		
Development_post	实验组	4.25	0.465	3.137*	0.755
	对照组	3.82	0.635		
Justification_post	实验组	4.29	0.437	0.973	0.234
	对照组	4.19	0.421		

*p<.05 **p<.001

5. 讨论

本研究设计两种 DST 课程：一种为传统的 DST 课程，另一种为基于 LLMs 的 DST 课程，通过随机对照实验方法，旨在探讨基于 LLMs 的 DST 对小学生批判性思维和认识论信念的影响。

针对研究问题 1，相较于传统数字故事讲述（DST）方法，基于大型语言模型（LLMs）的 DST 能够更有效地提高小学生的批判性思维能力。这种提升可能源于 LLMs 的介入，它能更好地协助学生在学习过程中进行知识建构。LLMs 的语言生成能力不仅仅是一种工具，更是一种促进思考和学习的媒介(程思凡, 2023)，通过引导学生参与更深层次的对话和探讨，在激发学生独立思考和提出合理质疑方面发挥了积极的作用。并且 LLMs 生成内容的不确定性需要学生主动进行反思探究其科学正确性(Bitzenbauer, 2023)，这种主动性有助于培养学生对知识的质疑态度和深入思考的能力，从而促进了其批判性思维的发展。

针对研究问题 2，研究结果证实了基于 LLMs 的 DST 相较于传统的 DST 能够更有效的提高小学生认识论信念。这符合先前已有研究，专注于建构主义学习的学生具有更高的认识论信念(Liang & Tsai, 2010)。基于 LLMs 的 DST 课程，学生可以更频繁地接触到来自不同角度和观点的信息。这种多元性可能激发学生对于知识多样性和复杂性的认识。由此可能发生的认知冲突促使学生重新审视自己的认识论信念，并更倾向于构建更全面、综合的知识结构(Valanides & Angeli, 2005)，从而促进了其认识论信念的发展。

参考文献：

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, *15*(3), ep430.

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary educational psychology*, *29*(2), 186-204.

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 1-20.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

Liang, J. C., & Tsai, C. C. (2010). Relational analysis of college science-major students' epistemological beliefs toward science and conceptions of learning science. *International Journal of Science Education*, *32*(17), 2273-2289.

Lin, H.-C., Hwang, G.-J., & Hsu, Y.-D. (2019). Effects of ASQ-based flipped learning on nurse practitioner learners' nursing skills, learning achievement and learning perceptions. *Computers & education*, *139*, 207-221.

Valanides, N., & Angeli, C. (2005). Effects of instruction on changes in epistemological beliefs. *Contemporary educational psychology*, *30*(3), 314-330.

Wu, J., & Chen, D.-T. V. (2020). A systematic review of educational digital storytelling. *Computers & education*, *147*, 103786.

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122-1136.

Yang, Y.-T. C., & Wu, W.-C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & education*, *59*(2), 339-352.

程思凡. (2023). ChatGPT 知识生产框架、技术迷思与未来演进. *当代传播*(06), 60-64.

池金章. (2019). 数字故事在小学品德与社会课程中的应用研究. *教育界(基础教育)*(09), 39-40.

焦建利, & 陈婷. (2023). 大型语言模型赋能英语教学：四个场景. *外语电化教学*(02), 12-17+106.

林青青. (2022). *高中生生物学认识论信念对学业成就影响的路径分析* [硕士，江西师范大学].

刘永谋，马亮，孙宇，高政，龙宝新，袁靖华，罗理章，林玮，& 崔岐恩. (2023). ChatGPT 对公共治理的挑战与机遇（笔谈）. *东方论坛*(03), 1-24+165.

邱艳萍. (2014). 论数字故事在教学应用中存在的问题及对策. *电子测试*(03), 133-134+150. https://kns-cnki-net-443.webvpn.wzu.edu.cn/kcms/detail/11.3927.TN.20140218.1659.050.html

席静，李昕，& 武香萍. (2018). 在信息过载时代,如何保有批判性思维? *农家参谋*(17), 224.

谢晓宇. (2023). 西方批判性思维教学的审视与反思. *外国教育研究*, *50*(11), 60-69.

基于增强现实技术的信息科技单元作业设计框架研究

Research on the Design Framework of Information Technology Unit Assignment Based on Augmented Reality Technology

张雪，魏雪峰*，朱晓颖，朱梦雨，孙晓涵
鲁东大学 初等教育学院
*xuefengwei99@163.com

【摘要】 作业设计需要体现综合性、探究性和创新性，根据不同的学习内容及学生特点，设计形式多样、不同层次的作业是当前教育教学改革中的一个关键问题，即如何充分利用信息技术的特点和优势，提升作业的育人功能。文章梳理了增强现实的研究现状，提出了基于增强现实技术的信息科技单元作业设计框架。该框架以解决信息科技课堂中的作业问题为重要旨归，以信息社会责任素养和增强现实技术为关键要素，包含分析、设计、开发、实施和评价五个设计环节。文章旨在探索增强现实技术与单元作业设计耦合的实践途径，为小学信息科技课程作业模式与课堂教学创新提供理论支撑与实践指导。
【关键词】 作业设计；增强现实；信息科技；ADDIE 模型

Abstract: *Assignment design needs to reflect comprehensive, exploratory and innovative, the design of various forms and different levels of assignment is a key issue in the current education and teaching reform, that is, how to make full use of the characteristics and advantages of information technology, improve the educational function of assignment. This paper reviews the research status of augmented reality, and puts forward the design framework of information technology unit operation based on augmented reality technology and related cases. The purpose of this paper is to explore the practical way of coupling augmented reality technology with unit assignment design, and to provide theoretical support and practical guidance for the information technology coursework mode and classroom teaching innovation in primary schools.*
Keywords: Job Design, Augmented Reality, Information Technology, ADDIE Model

基金项目：本文系山东省人文社会科学课题（智库重点项目）“人工智能支持下的小学个性化作业设计与应用研究”（编号：2023-zkzd-014）的阶段性研究成果。

1. 前言

《义务教育信息科技课程标准（2022 年版）》强调学生在学习过程中的自我选择和自我设计，信息科技课程以提高学生的核心素养为宗旨（中华人民共和国教育部，2022），而信息科技学科在义务教育阶段却一度处于尴尬境地，学生和家长对信息科技学科比较忽视。此外，2021 年 7 月，中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》（简称“双减”政策）明确了“减轻学生过重作业负担，减少学生作业量及时长”的总要求，还对学生作业的数量、质量、评价等做出了具体要求。尤其在信息科技课程中，不能留家庭作业的现实问题亟待解决，那么巩固和复习所学知识、掌握技能技巧的重任就落在了课堂上。学生在有限的时间内不仅要切实掌握信息科技的基本知识和基本技

能，还要在学习活动过程中，学会应用信息科技知识解决问题，培养良好的信息素养。课堂作业作为课堂教学的重要环节之一，是老师教与学生学的交汇点，也是实现课程总目标的重要手段。本文通过梳理增强现实技术教育应用的相关研究，运用 Citespace 软件进行简单分析，旨在提出基于增强现实技术的单元作业设计框架。将增强现实技术与信息科技作业设计有机结合，以期为小学信息科技课程的作业模式和课堂教学创新提供理论支持和实践指导。

2. 文献综述

2.1. 培养信息科技核心素养的单元作业设计

如何借助信息技术特点和优势增进作业育人功能，是当前教育教学改革中需要关注和解决的现实问题，在最新发表的文章中，安雪慧等人（2024）梳理了目前各地探索借助信息技术推进作业改革的实践现状及遇到的新困顿，探讨了如何把握信息化背景下中小学作业的育人方式和新特征。武法提等人（2024）认为个性化作业是落实“双减”政策，实现减负提质的重要途径，基于人在回路机制的个性化作业设计理论模型，提出了个性化作业实践路径。李晓江（2023）分析了信息科技核心素养的组成要素并重点阐述了指向核心素养的初中信息科技作业设计有效策略。梁祥（2023）基于信息科技核心素养，对标学段目标，按照学生的认知特征和信息科技课程六条逻辑主线，运用可实时收集学生作业的软件，探索中小学信息科技作业特征下的设计理念与实践应用方式。关于培养信息科技核心素养的单元作业设计的研究多集中于义务教育阶段，多数学者从双减政策背景及《新课标》提出的核心素养出发思考如何进行作业设计，优化作业结构。

2.2. 增强现实在教育领域的应用现状分析

增强现实技术（Augmented Reality）在教育中的运用一直是备受关注的创新领域，为研究和实践提供了广阔的前景。其特点是将虚拟信息叠加到真实环境中，为学习创造了与传统方式截然不同的情境。Milgram 等（1995）提出了现实——虚拟连续体来定义增强现实（AR）、混合现实（MR）和虚拟现实（VR）。每个术语的简短定义表明，VR 使学习者沉浸在一个完全人工的数字环境中，AR 将虚拟对象覆盖在现实环境中，MR 以各种创新的方式将虚拟环境与现实元素结合在一起（Milman，2018）。过去的研究主要集中在截至 2012 年的时点，关注于 AR 学习环境的开发、发展和初步评估，而较少进行实验研究以深入探讨增强现实对学习的实际影响（Wu et al.，2013）。近十年来，随着对 AR 技术在学习中的潜在影响的认识不断加深，学者们开始广泛探索增强现实在教育领域的应用。本文在中国知网（CNKI）文献库中以“AR+教育”“增强现实+教育”“增强现实技术（AR）+教育”“增强现实应用+教育”等为主题对近十年（2014-2024）的 CSSCI 文献进行检索，共检索文献 422 篇，筛除与研究主题无关的文献，得到 380 篇有效文献。



图 1 作者合作图谱分析



图 2 机构合作图谱分析

借助 Citespace 软件对 380 条数据分别选择节点类型（Node Types）为作者（Author）和机构（Institution）的可视化图谱分析，得到图 1 与图 2，分别呈现了作者与机构在“增强现实+教育”领域的影响力。图中左上角的文字显示了相关数据，作者合作图谱中出现了 283 个节点，

连线为 167 条，从图中也可以直观地看出部分作者之间连线较多，联系较为紧密，出现了以李小平、黄业桃、刘清堂、刘炜及朱永海为核心的合作网络。从机构层面看，机构合作图谱中出现了 283 个节点，连线为 126 条，各机构之间联系比较紧密，发文量以华东师范大学教育信息技术系、南京大学信息管理学院及北京师范大学未来教育高精尖创新中心为主。

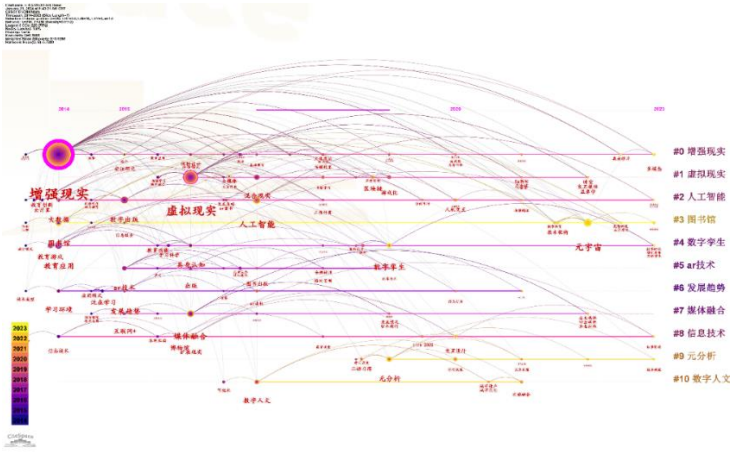


图 3 增强现实关键词分析时间线图

通过时间线图对文献关键词进行分析可以了解从 2014 至今十年间研究热点的变化趋势及未来方向，图中呈现出形成了“增强现实”“虚拟现实”“人工智能”等为主的聚类，其中“增强现实”时间跨度最大，该聚类持续时间最长；“虚拟现实”在 2016 年成为关注热点，至 2022 年热度降低；“人工智能”及“元宇宙”分别在 2017 及 2022 年兴起。因此，增强现实一直是教育领域的研究热点，是教育技术研究 CSSCI 期刊重要的选题方向。

2.3. ADDIE 模型应用于作业设计的研究

在作业应用领域，发现 ADDIE 模型主要集中于地理、语文、政治及化学等学科，如胡新雨（2021）采用 ADDIE 模型理论，按照分析、设计、开发、实施和评价步骤，阐述了地理大单元作业的操作模式；卢清菁（2023）认为该模型为高中政治课单元作业设计提供了可操作的思路，遵从模型的整体性和系统性，可有效把握单元作业设计的路径与结构；欧阳雪等人（2023）基于 ADDIE 模型，提炼了以“分析、设计、开发、实施、评价”为核心环节的单元作业设计与实施模型，以“自然界的水”单元为例，从单元主题分析、单元作业整体设计、单元作业案例开发、单元作业实施和单元作业评价五个方面探索了初中化学单元作业设计的具体路径。

3. 基于增强现实的“ADDIE+信息科技”单元作业设计框架

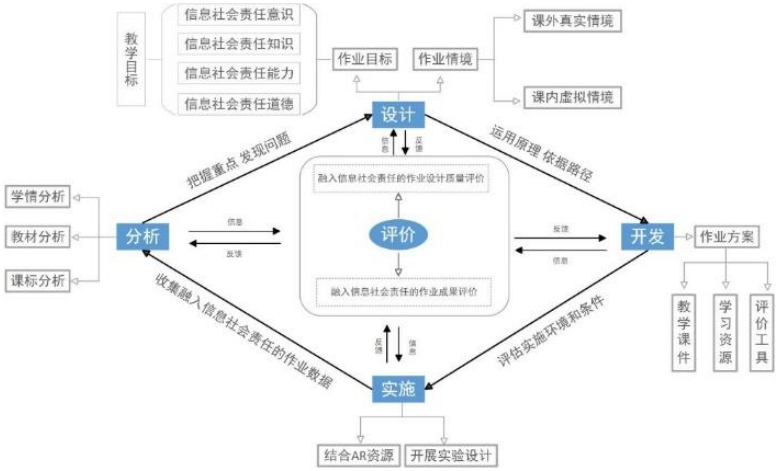


图 4 信息科技单元作业设计框架

ADDIE 模型包含分析（Analysis）、设计（Design）、开发（Develop）、实施（Implement）、评价（Evaluate）五个环节。基于 ADDIE 模型，融入增强现实技术及信息社会责任，结合小学信息科技学科特点，设计了小学信息科技单元作业设计的框架。分析是对学生情况、教材内容和课程标准进行分析，明确本单元的重点内容和学生在学习过程中可能存在的问题，为下一步的目标设计和情境设计奠定基础。设计是依据教学目标形成作业目标，结合课外真实情境与课内虚拟情境，通过增强现实技术对作业进行情境的设计。此后，在情境框架和目标框架的引领下，将作业方案内容开发与项目式学习原理、多元智能理论、系统论等理论进行有机结合，形成具备教学课件、学习资源与评价工具的作业方案，完成开发环节。还要依据学校和学生的情况，评估单元作业实施环境和条件，及时调整作业方案。教师确定好方案后，可以依据自己的实际教学情况，合理安排作业布置的时间，并在学生完成作业的过程中及时进行指导，同时收集学生的作业成果数据和对作业方案的评价数据，再次进行分析，并开展融入信息社会责任的作业成果评价与作业设计质量评价。五个环节构成了一个动态循环的过程，每个环节都详细规定了相应的内容，以确保在单元作业设计过程中的适应性和灵活性。

4. 总结

本文探讨了如何整合增强现实技术以提升信息科技学科的教育体验和实用性。通过设计一个创新的框架，成功地结合了增强现实的虚拟元素与真实世界的学科内容，为学生提供了更丰富、更互动的学习环境。我们期待这一作业设计框架能够在教育领域推动增强现实技术的应用，为学生创造更具有吸引力和实际意义的学习体验。在不断迭代和改进的过程中，增强现实与信息科技单元作业设计的结合将为学生提供更为丰富、创新的学习机会，助力其在未来科技领域取得更大的成就。

参考文献

中华人民共和国教育部.(2022).义务教育信息科技课程标准:北京师范大学出版社.

中共中央办公厅，国务院办公厅（2021-07-24）.关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见.

李晓江(2023).指向核心素养的初中信息科技作业设计.中小学电教(教学)(05),22-24.

安雪慧和张琨(2024).信息技术增效作业育人改革的困顿与突破.中国教育学刊(01),18-24.

卢清菁(2023).高中思想政治课单元作业设计研究（硕士学位论文,上海师范大学).

武法提,李坦,任伟祎和夏志文(2024).基于人在回路机制的个性化作业设计:理论模型及实践理路.现代远程教育研究(01),93-100.

欧阳雪和李远蓉.(2023).基于 ADDIE 模型的初中化学单元作业设计与实施研究——以“自然界的水”单元为例.化学教学(10),78-84+97.

胡新雨(2021).指向地理核心素养培养的大单元作业研究（硕士学位论文,南京师范大学).

梁祥,葛萱和查和胜(2023).基于素养导向的信息科技作业设计研究.中国信息技术教育(14),45-48.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995, December). Augmented reality: A class of displays on the reality-virtuality continuum. In Telemanipulator and telepresence technologies (Vol. 2351, pp. 282-292). Spie.

Milman, N. B. (2018). Defining and conceptualizing mixed reality, augmented reality, and virtual reality. Distance Learning, 15(2), 55-58.

Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. Computers & education, 62, 41-49.

支持學生於沉浸式虛擬實境社交互動並結合 STEM 與 SDGs 的學習環境開發

Development of Immersive Virtual Reality Learning Environment Supporting Students' Social Interaction and Integrating STEM and SDGs

郭鋅恩[1]，陳政煥[2*]
[1] 中央大學 網路學習科技研究所，臺灣
[2] 亞洲大學 行動商務與多媒體應用學系，臺灣
[*] chchen@asia.edu.tw

【摘要】 本研究針對沉浸式虛擬實境結合永續發展教育主題，開發一個能支持國中學生在元宇宙中社交互動的 STEM 專題學習環境。本研究以 VRChat 作為開發平台，參照國中八年級教材內容搭配永續發展目標，設計與真實世界相關的學習主題，讓學生能透過頭戴式裝置在元宇宙環境進行學習。此處虛擬環境聚焦於「氣候行動」，讓學生透過場景中的互動物件，探索海平面上升、溫室效應以及能源科技等議題，促進深入的互動學習與討論。
【關鍵字】 沉浸式虛擬實境；社交互動；專題式學習；STEM；SDGs

Abstract: *This study focuses on integration of immersive virtual reality with themes of sustainable development education, developing a STEM project-based learning environment that supports social interaction among junior high school students within the metaverse. The study used VRChat as the development platform, referencing eighth-grade instructional materials and sustainable development goals to design learning themes related to the real world. This enables students to learn within the metaverse environment using head-mounted display. For the 'Climate Action' theme, the virtual environment was designed to facilitate the exploration of issues such as rising sea levels, greenhouse effect, and energy technology using interactive objects in the scene, promoting in-depth interactive learning and discussion.*
Keywords: immersive virtual reality, social interaction, project-based learning, STEM, SDGs

1. 前言

元宇宙（metaverse）帶來虛擬實境技術的新應用，將沉浸式虛擬實境（immersive virtual reality, IVR）應用在教育是極具潛力的領域，IVR 能夠為 STEM 和永續發展目標（sustainable development goals, SDGs）教育創造多元且豐富的學習環境。目前已有一些文獻（e.g., Makransky & Petersen et al., 2021; Won et al., 2023; Wu et al., 2020）是關於 IVR 在教育領域的系統性回顧或後設分析，針對使用頭戴式裝置（head-mounted display, HMD）作為 IVR 呈現設備，另有一篇（Radianti et al., 2020）回顧 38 篇在高等教育中應用 IVR 的文獻，其中有 76% 採用了 HMD，都顯示 HMD 作為一種高度沉浸式媒介（Makransky & Petersen, 2021）的重要性，並且與桌面虛擬實境（desktop virtual reality）相比，HMD 在學習效果有更顯著的優勢（Wu et al., 2020）。IVR 在教育領域的應用中，STEM 領域有較多的相關研究（Radianti et al., 2020; Wu et al., 2020），有研究便讓學生使用 HMD 虛擬實地考察探索相關內容，增進他們 STEM 相關知識（Petersen et al., 2020; Pimentel & Kalyanaraman, 2023），IVR 能為 STEM 提供豐富多元的學習環境。然而，過往 IVR 環境中對於社交互動性的探討較少（Radianti et al., 2020; Won et al., 2023），以及如何將 SDGs 與 STEM 教育結合，有鑑於十二年國教課綱，跨領域教學是重點推動項目之一，本研究採用 VRChat 開發 IVR 環境，內容以 STEM 跨領域整合專題，結合 SDGs 學習內容，著重在具社交互動的 IVR 環境，讓學生在元宇宙的學習環境中進行討論等社交互動，解決與真實世界相關的任務與挑戰。

2. 研究方法與設計

本研究基於國中八年級的自然科學、科技和數學教材，結合聯合國的 SDGs 進行設計，採用 Won 等人（2023）在他們的回顧中，歸納出 IVR 的四個關鍵設計特徵，分別是感官性（包括視覺、聽覺和觸覺反饋）、行為性（涉及人機互動和身體動作）、敘事性（涵蓋角色定義、情境內容和挑戰性任務）以及社會性（指導性的社交互動），因此本研究選擇能支持多人互動和多種裝置的元宇宙平台 VRChat 作為開發平台，提供環境中自由移動、訊息交流、即時語音、以及豐富的互動物件設計，開發一個支援社交互動與 STEM 專題學習的 IVR 環境。

3. 研究結果

本研究針對 SDGs 第 13 項目標「氣候行動」設計場景，學生能在虛擬環境親身觀察海平面上升對小島國的影響，內容涵蓋了海平面變化的統計數據、溫室效應的成因及其後果以及能源科技的種類與優缺點，學生通過多媒體素材（如影片、照片、圖表等）和互動物件（如畫筆、噴漆、拼圖等）進行多元操作，從而在元宇宙深入反思、探索學習和互動討論，加深氣候變遷對環境的實際影響及對 SDGs 的理解。

4. 結論與未來展望

IVR 在科技教室及未來教育的應用極具潛力，本研究以強調社交互動元素的 VRChat 設計 STEM 結合 SDGs 專題的元宇宙學習環境，提供一個 IVR 在教育中應用的新途徑。

致謝

本研究在國家科學及技術委員會（NSTC 112-2628-H-468-001-）的資助下完成，謹此致謝。

參考文獻

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 1–22.

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology, 51*(6), 2099–2115.

Pimentel, D., & Kalyanaraman, S. (2023). Virtual climate scientist: A VR learning experience about paleoclimatology for underrepresented students. *Interactive Learning Environments, 31*(7), 4426–4439.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778.

Won, M., Ungu, D. A. K., Matovu, H., Treagust, D. F., Tsai, C.-C., Park, J., Mocerino, M., & Tasker, R. (2023). Diverse approaches to learning with immersive Virtual Reality identified from a systematic review. *Computers & Education, 195*, Article 104701.

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology, 51*(6), 1991–2005.

《氫朋好友》互動繪本在科學教育中的應用：以氫能源為例

Application of the Interactive Picture Book 'Hydrogen Friends' in Science Education: A Case Study on Hydrogen Energy

湯梓辰[1*]，趙彥俞[2]，林孟樺[3]，周宏隆[4]
[1*234]國立臺灣科技大學應用科技研究所
[1*]jttang0@mail.ntust.edu.tw

【摘要】 本研究旨在探討《氫朋好友》互動繪本在科學教育中的應用，我們開發了這部繪本，以擴增實境和虛擬實境豐富學生學習體驗。未來期待調查高中至大學學生，評估他們閱讀《氫朋好友》後在科學概念學習上的成效和對學習的反思。

【關鍵詞】 氫能源；互動繪本；擴增實境；虛擬實境

Abstract: *This study explores the application of the interactive picture book "Hydrogen Friends" in science education. We developed this picture book to enrich students' learning experiences with augmented and virtual reality. In the future, we look forward to surveying high school and college students to evaluate their effectiveness in learning scientific concepts and their reflections on learning after reading "Hydrogen Friends".*

Keywords: Hydrogen Energy, Interactive Picture Book, Augmented Reality (AR), Virtual Reality (VR).

1. 前言

氫能作為一種潔淨、高效率的能源形式，在世界各國面臨能源轉型的數十年來，引起了廣泛的關注與研究(Jain, 2009)。公眾對於氫能的認知相對較低，缺乏對於氫能的深入了解（劉世鈞，2014）。因此，本研究可以著手進行氫能源的科普教育，以提高大眾對於氫能的認識和理解。

綜此，研究者將開發《氫朋好友》並整合氫能源相關知識，應用擴增實境（AR）和虛擬實境（VR）提升學習體驗。並在未來邀請高中及大專學生參予實地實驗。

2. 繪本應用於科學議題之探討

綜合以上所述，為了使數位繪本更符合科學教育之議題，研究者針對文獻進行探討之，並將其應用於繪本之上，並期待日後有更完整的實驗結果。以下將針對繪本教育應用於科學議題探討之：

繪本在科學教育領域的研究中逐漸被看作有助於教學的輔助教材，有些研究者使用繪本來張著學生建構科學概念(Cho & Kim,1998) ,繪本中的科學故事不僅適用於原本即對科學有興趣的人，更提供給原本對科學不感興趣的人重新審視自己對科學與科學家印象的機會(Buxton & Austin, 2003)。

Down(1991)將科學事實融入到創作的故事中，可以提升學生對環境問題的了解，以及價值觀的建立。繪本的特性使得它在教育應用上具有價值，逐漸在各個領域的教學中得到應用。繪本教育在科學教育中展現出豐富的應用潛力，特別是在跨領域的研究和教學中。透過繪本，我們能夠以直觀的方式呈現科學知識，不僅引發對科學的興趣，還能促使學生深入思考和建構科學概念。

本研究透過融合繪本教育、科普閱讀課程，以及氫能源與科普教育的元素，構建了一個科學繪本教育架構。這種綜合性的教學模式不僅有助於拓展學生的知識視野，還能夠激發他

們對科學議題的興趣。從研究趨勢來看，將繪本納入正式的科學課程已經取得積極成果，證實了繪本在教學中的實際效益(Giasiranis & Sofos, 2016；林淑慧、吳靖國，2018；李文獻等人，2020)。同時，隨著數位科技的進步，數位繪本更為學習提供了新的可能性，不僅能夠豐富教學內容，還能夠激勵學生進一步探索科學知識。因此，在本研究中，我們提供了互動式繪本，並期待未來進行實驗研究進行探究之。

3. 互動式繪本

本研究主要透過繪本引導學生學習，以氫能源為主題，創作《氫朋好友》科學繪本，提升學生對氫能源的認知。互動式繪本成為主要教學工具，統一的圖像風格和引人入勝的故事情節，搭配擴增實境（AR）和虛擬實境（VR）技術，豐富學生知識與興趣，激發對能源的思考，強調人類與地球環境的相互關係，並鼓勵對環境永續發展採取實際行動。研究目標為提升對氫能源的理解、激發對能源的興趣，以及提高對永續議題的認識。繪本頁面設計如圖 1 所示。



圖 1 繪本的頁面設計

4. 結論

本研究透過創作《氫朋好友》科學繪本，結合教育和能源專業，以促進對氫能源的認知。研究者運用繪本和 AR/VR 技術，希望藉此提供學生豐富的學習體驗。未來本研究將其繪本進行實作研究並探討與發展。

參考文獻

Buxton, C. & Austin, P.(2003). Better books better teaching. Science and Children. 41(2), 28-32.

Cho, B. K. & Kim, J. J. (1998) Literature based science activities in kindergarten through children's picture book. Tampa, Florida. (ERIC Document Reproduction Service No.ED420526)

Dowd, F. S. (1991). Storybooks: Stimulating science starters. *School Library Media Quarterly*, 19, 105-108.

Giasiranis, S., & Sofos, L. (2016). Production and evaluation of educational material using augmented reality for teaching the module of "representation of the information on computers" in junior high school. *Creative Education*, 7(09), 1270- 1291

Jain, I. P. (2009). Hydrogen the fuel for 21st century. *International Journal of Hydrogen Energy*, *34*(17), 7368-7378. https://doi.org/https://doi.org/10.1016/j.ijhydene.2009.05.093

李文獻、林美君、張俊彥、賴信志、林慧敏（2020）。運用數位科技戶外教育場域之科學繪本教學設計－從教學者觀點探討。**科學教育月刊，第 428 期**，11-31。

林淑慧、吳靖國（2018）。**繪本教學提升學童海洋素養行動研究中的教師專業成長**。基隆市：國立臺灣海洋大學。

劉世鈞（2014）。大學生氫能認知與態度調查分析之研究。國立台南大學綠色能源科技學系碩士在職專班。

基于 VR 技术的电磁波教学平台设计与交互式开发

Design and Interactive Development of Electromagnetic Wave Teaching Platform Based on VR Technology

张婉露[1*]，朱佳[1]，潘昌藩[1]，郭杭辉[1]，陈张泽[1]
[1]浙江师范大学教育学院
[*] zhangwanlu@zjnu.edu.cn

【摘要】 本研究旨在将虚拟现实（VR）技术与课堂教学相结合，设计了面向“电与磁”物理课程的 VR 教学平台。研究通过对不同场景下的电磁波进行设计、建模与优化，以展示各应用场景下不同频率的电磁波。经过测试，该平台获得了 83.62%的学生满意率和 8.637 分（满分 10 分）的教师满意评分，表明该平台具备可用性。
【关键词】 虚拟现实；电磁波；教学展示

Abstract: *This study aims to combine virtual reality (VR) technology with classroom teaching and design a VR teaching platform for the "Electricity and Magnetism" physics course. In this paper, electromagnetic waves in different scenarios are designed, modeled, and optimized to show electromagnetic waves with different frequencies in various application scenarios. After testing, the platform achieved a student satisfaction rate of 83.62% and a teacher satisfaction score of 8.637 (out of 10), indicating that the platform is available.*

Keywords: Virtual Reality, Electromagnetic Waves, Teaching Display

基金项目：国家重点研发计划子课题“农村地区教师教学能力智能评测与教学精准辅助技术研究”（项目编号：YEC3303600）；国家自然科学基金“在线学习环境下基于事件知识图谱的可解释性学习评价研究”（项目编号：62077015）。

1. 引言

当前，虚拟现实技术在教育领域的应用潜力已受到广泛关注，特别是在优化学习体验、提升学习兴趣及促进知识内化方面（Jaziar,Tim&Jennifer,2020）。然而，在物理课程中，虚拟现实技术的研究尚处于探索阶段，特别是在电磁波这一具体课程领域中，尚缺乏深入的研究和实践应用（徐立君、潘嘉宁和范诗怡，2020）。鉴于此，本研究聚焦于电磁波教学的实际需求，致力于构建基于 VR 技术的电磁波教学平台，旨在解决传统电磁波教学中存在的抽象性、难理解和学生参与度不高的问题。

2. 平台的设计与交互式开发

本研究平台以经典的用户体验（User Experience, UX）设计理念为基础，深度融合三维仿真建模、虚拟现实（VR）技术及数据库技术，力求为用户带来直观且沉浸的学习体验。平台紧密围绕《电磁场与电磁波》教材的核心内容与教学要求，设计了均匀平面波、波的折射与反射以及矩形波导三大核心学习场景。

在电磁波场景建模的基础上，本研究进一步根据展示与交互需求，增添了能够调整视角及电磁波相关参数的组件。这一设计使得用户能够灵活地从不同角度观察具有不同参数的电磁

波，深入探索其内在特性。此外，平台还提供了在同一场景中不同实验台上电磁波的对比观察功能，有助于用户更加直观地理解电磁波的性质和变化规律。平台效果呈现如下图 1 所示。



图 1 电磁波教学平台使用效果图

3. 平台可用性测试

本研究对某中学九年级四个班级共 81 名学生（其中男 45 人，女 36 人）进行对照试验，同时，对四名物理老师进行问卷访谈，实验过程如图 2 所示。经前测，两组学生的学习水平相近；后测结果表明，实验组学生的成绩显著高于对照组（t=-2.46，P(Sig)=0.042），同时，学生对平台的认可度和流畅度的满意率分别达到 83.62%和 74.58%，学生对学习中电磁波呈现的直观性表现出高兴趣；教师对平台的流畅度、功能性、知识点覆盖度的满意度平均分均超过 8 分，超过半数老师高度肯定实验交互的设计。学生和教师的积极反馈表明电磁波教学平台在实际应用中具有可用性，平台为物理学科的学习提供了有力的支持。

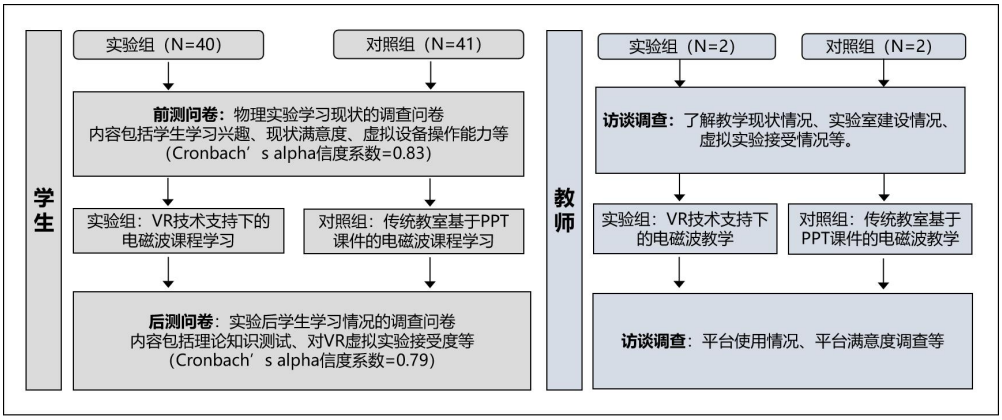


图 2 实验流程图

4. 研究结论

本研究将 VR 技术与电磁波教学相结合，旨在克服传统电磁波教学中的抽象性与理解难题。通过构建电磁波三维模型，平台创建了多频率电磁波的 VR 场景，并利用 Unity 实现用户交互，包括场景转换、操作装置、视角变换等。实际应用表明，该平台效果良好，得到师生积极反馈。展望未来，此平台可与更多教学资源结合，有一定的实践价值与推广空间。

参考文献

Jaziar Radianti, Tim A. Majchrzak&Jennifer Fromm.(2020).A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda.Computers & Education.Volume 147.

徐立君,潘嘉宁,&范诗怡(2020). 新工科背景下虚拟现实技术在大学物理教学中的应用研究. 中国农村教育, 4, 18-19.

融入關鍵火災處理概念之教育遊戲—輔助國小學生火災教育議題設計探究

Exploring the Educational Benefits of Integrating Educational Game-Assisted Concepts in Fire Incident Management to Enhance Fire Safety Education for Elementary School Students

王曉璿[1]，蔡承易[2]，劉沛岑[3]，張嘉益[4]，楊逸捷[5]

國立臺中教育大學數位內容科技學系

* hswang@mail.ntcu.edu.tw

【摘要】本研究主要聚焦於運用 3D 數位遊戲進行火災教育，同時評估其對國小學生的影響。現代教育越來越認識到傳統教學方式或許無法引起學生足夠的興趣和參與。數位遊戲作為一種富有趣味且具有高度互動性的學習方式，引起廣泛重視。此遊戲結合學習和娛樂，能有效激發學生的興趣和參與度。透過遊戲，我們得以建立一個虛擬世界，使學生在其中體驗和應對火災相關的挑戰，進而學習火災安全知識和應對技能。這種互動性有助於加深理解和學習，對於年輕學生而言更具吸引力和實用性。

【關鍵字】 數位遊戲式學習；火災教育；冒險類遊戲

Abstract: *Education nowadays is increasingly recognizing that traditional teaching methods may not be able to interest and engage students. Digital games have attracted a lot of attention as a fun and interactive way of learning. These games combine learning and entertainment and can stimulate students' interest and engagement. This study focuses on the use of educational games for fire education and evaluates the impact on elementary school students. Through digital games, we can create a virtual world in which students can experience and cope with fire-related challenges, learning fire safety knowledge and coping skills. This interactivity helps deepen understanding and learning, and is especially engaging and useful for young students.*

Keywords: Digital Game Based Learning, Fire Education, Adventure Games

1. 前言

隨著時代的演變，災害教育在現代社會中越來越受到關注和認可。「科技的進步，將數位遊戲式學習（digital game-based learning, DGBL） 運用在教育上，已成為新興趨勢」（戴于真、張懷綾、金凱儀，2020）。火災是一種嚴重的自然災害，每年都在世界各地造成巨大的破壞、人員傷亡以及財產損失。然而，只要我們提高社區成員的火災風險意識，並提供實用的火災安全知識，過多的火災事件實際上是可以預防的。目前國內防災缺乏教育與訓練，防災教育教材的不足與過時，導致學生學習意願低落；而在社區方面，專業訓練不足且演練流於形式，使得社區防災成效也明顯不足。不論是在學校或是社區，明顯可看出防災教育少

之又少(張玉連，2021)。因此我們希望能增加大眾對於火災逃生的學習意願。Prensky(2007)提出：二十一世紀是數位遊戲學習（digital game-based learning）的世代，他們與科技共同成長，從電腦輔助學習到數位學習，再輔以遊戲式融入教學，再進化到數位遊戲式學習。近年來，更有學術單位提倡「遊戲化學習」，欲結合教育與娛樂， 透過遊玩方式傳達教育理念。至此，玩遊戲已不再是讀書道路上的絆腳石，反倒成為學習的助力（Elizabeth Hanson-Smith,2020)。這些學習方法除了有各式各樣的科技工具輔助，亦有不同的效用，隨著遊戲科技的進步，數位遊戲式學習已受到矚目（張基成、林冠佑，2016）。

本研究主要是關注如何將火災逃生處理概念與教育遊戲結合，並評估其對國小生的影響。透過數位遊戲，我們可以營造出一個虛擬世界，讓學生透過遊戲中的任務來獲取並學習知識， 在遊玩過程中建立、培養與應用技巧(McFarlane, Sparrowhawk & Heald, 2002, 林友信、趙伯堯，2018)。這種互動性有助於深刻理解和學習，尤其對年輕的學生來說，更具吸引力和實用性。

2. 文獻探討

2.1. 火災教育

以目前台灣的學校教育來說，雖然將防災課程列為必些課目，但這些內容卻只注重學術知識缺乏防災態度、防災技能、心理建設及人文關懷，同時非常缺乏定期實質的操作與實練。另外「人為災害」的相關課程也幾乎沒有(洪志誠，2003)。而在社區防災這塊更是成效不佳，目前社區防災組織災前專業訓練不足，演練又流於形式，使得社區防災成效大打折扣。公部門領導者對社區防災教育的認知程度不一、學術資源與經費設備取得不穩定、地方派系的傾軋、社區人文屬性的區域與內部多元差異，使推動社區防災教育的問題更加棘手。(郭俊欽、莊翰華、康良宇，2007)。

2.2. 冒險類遊戲類型探討

Alessi & Trollip（1985）將遊戲分為冒險遊戲、棋盤遊戲（board games）、 紙牌或賭博遊戲、戰爭遊戲、邏輯遊戲、手腦並用的遊戲（psychomotor games） 以及角色扮演遊戲七類。根據 Adams 與 Rollings(2007)對冒險遊戲的定義:『冒險遊戲(Adventure Game，AVG)是電腦遊戲中的一個大類。玩家會以遊戲中的主視角進行探索，並根據線索來挖掘未知的劇情及線索、地圖，此遊戲種類著重於「探索」，探索未知的劇情、線索以及地圖等。透過感受未知所帶來的緊張感，因而做出不同的應對方式(迪迪那・里維, 2011)。根據 Hirameki International Group(2006)統計，2006 年在日本發行的電腦遊戲中，將近 70%屬於冒險遊戲。近年來這類遊戲更跨越文化藩籬，受到全世界玩家歡迎。例如:美國最大的遊戲製造商藝電(Electronic Arts)在 2010 年推出搭載在 ios 平台的冒現遊戲 cause of death(GameSpot,2012)。

2.3. 遊戲化學習

遊戲化學習(Learning through play)，又稱為學習遊戲化，就是採用遊戲化的方式進行學習。「遊戲化」並不是「玩遊戲」，而是把「遊戲的元素」應用到「非遊戲的現場」，遊戲化的核心元素總共有三種：積分（Points）、獎勵（Benefit）、排行榜（Leader board），以英文簡寫 PBL 來記憶。遊戲化教學，就是把這三種元素應用到教學的現場，它是目前比較流行的教學理論和教育實踐。將學生對遊戲的喜好轉化為對學習的動機，可以使老師的努力事半功倍。這正是遊戲化學習研究的核心。目前這種觀點，存在兩種傾向：(1)引用了評估中學習的理論，透過利用編訂教學遊戲軟體或電子課本，將教學的內容串聯遊戲或網上競技的方法引導學生進入學習狀態，而避免了傳統授課模式的死板和枯燥。這種方式目前在香港有小學正在試行，試點學校計有油蔴地天主教小學（海泓道)及將軍澳的佛教黃藻森小學。

(2)遊戲本身即具學習性。因此透過深入研究遊戲特性，以該等特性為基礎重塑學習機制，使學習過程本身具備遊戲性質，而非僅透過遊戲形式呈現學習內容。

那何謂數位遊戲化學習?數位遊戲化學習是一種結合遊戲元素與學習目標的教育方法，旨在提高學習者的參與度、動機和學習效果。遊戲式教學，可以加強學生參與度和掌握實踐經驗(Colarusso,1993)。Prensky（2001）認為數位遊戲式學習即任何教育內容與電腦遊戲的緊密結合，亦可把它定義為在電腦或線上（online）的任何教育性遊戲。而 Garris、Ahlers 及 Driskell（2002）的遊戲式的學習模式，也可詳實地描述何謂數位遊戲式學習，首先即需設計一個可以整合教學內容和遊戲特性的教育性遊戲，其次是此遊戲能讓人沈浸於面對挑戰而不斷地引發判斷、執行與系統回饋的循環，最後則藉由此種遊戲的投入而達成某一特定的學習目標。因此，簡單地說，數位遊戲式學習就利用內涵學習元素的數位遊戲，來達成某一特定的學習成果。

因此本研究經文獻探討，以火災處理概念為主，輔助國小學生火災教育議題，開發一款第一人稱防災教育的「冒險遊戲(AVG)」。相較於其他遊戲類型，在冒險遊戲中玩家以主視角進行探索，根據線索挖掘未知的劇情、地圖等，強調了「探索」的遊戲特性。通過感受未知所帶來的緊張感，玩家需根據情況做出不同的應對方式。這種設計有助於增加玩家的參與度和投入感，從而更有效地達到教育目的。因此，透過冒險遊戲的方式，我們能夠使學生在享受遊戲樂趣的同時，更深入地了解火災防災知識，提升其應對火災事件的能力，並期待本研究能為災害教育領域的發展提供有價值的參考，並為學生和社區的安全做出貢獻。

3. 研究方法

3.1. 研究對象

由上述可知，台灣目前學校教育存在弊處。現有的國小火災教育體系主要停留在理論知識的灌輸，缺乏實際應對火災的實務訓練，使學生難以在實際情境中有效應對火災風險。此外，教育方式單一，難以因應不同學生的差異性需求。在這樣的背景下，我們選擇以國小學生作為研究對象，將焦點放在他們的火災教育上，這具有重要的意義。考慮到遊戲的難易度，我們選擇了國小學生中的高年級作為研究的參考對象。因為高年級學生已經具備了一定的學習基礎，能夠更深入地理解和應用火災安全知識。相對地，國小學生對火災知識和應對技能的理解和實際應用能力較低，因此需要更多的教育和指導。此外，他們在面對火災時的行為模式可能與其他年齡段存在差異，因此專注於他們的火災教育更為重要。由上述文獻內容可知，將 3D 冒險遊戲技術融入國小火災教育中具有顯著的優勢。首先，3D 冒險遊戲提供了模擬真實火災情境的機會，讓學生能夠處於實際環境中進行應變練習，強化其實際操作能力。其次，遊戲的設計可以根據學生的年齡、認知水平調整難度，提供更貼近實際需求的教育體驗，提高學習效果。

3.2. 關卡設計

第一關：火勢嚴重的室外火災情境描述：在第一關中，起火點位於室外，火勢相當嚴重黑煙已經滾滾而來。玩家必須面對以下三個重要任務，以應對這一極端情境。任務一：門的關閉：玩家的首要任務是關上室內的門，以預防外頭的煙霧進入室內，確保室內空氣相對清新。任務二：口罩的尋找：玩家需要尋找口罩，以避免煙霧進入口鼻，防止嗆傷或昏迷。任務三：求助行動：玩家必須前往陽台，撥打 119 報火警並尋求協助。這是緊急狀況下的正確應對方式。獨立事件：若玩家試圖躲進廁所，將出現對話框，提醒玩家絕對不應該躲進廁所，因為這樣可能會增加危險。學習目標：玩家需要在有限時間內完成前往陽台求助的動作，這鼓勵玩家學習在火災緊急情況下的快速決策和應對技能。第二關：室內嚴重火勢的自

我逃生情境描述：在第二關中，起火點位於室內，火勢嚴重，但外頭的煙霧相對減少。玩家需要依次完成多個關鍵任務，以確保自己的安全。任務一：離開起始房間：玩家首先必須快速離開起始房間，以避免火勢擴散到該區域。任務二：關上房門：在離開房間後，立即關上房門，以隔離火源。任務三：樓梯前進：玩家需要迅速前進至樓梯方向，尋找下樓的出口。任務四：下樓逃生：玩家必須順著樓梯繼續往下逃生，確保離開高危險區域。任務五：出口尋找：玩家需要依循逃生指示燈的指引，找到安全出口，從而成功逃生。獨立事件：玩家在必要時刻需要學會使用滅火器，以應對火勢的控制。學習目標：本關的學習目標是使玩家學會在室內火災情況下自行逃生，包括適當的逃生程序、安全閉門、使用滅火器等技能。積分（Points）：積分可以根據玩家在逃生過程中的表現來獲得，例如快速逃生、有效使用逃生工具、正確回答火災相關問題等。這些積分可以直接在遊戲內顯示，並用於解鎖特殊獎勵或者在排行榜上展示。

3.3. 整體遊戲設計

本研究旨在創建一種以遊戲方式教育火災應對技能的教育工具，遊戲場景設計主要依照不同之應對狀況進行設計。使玩家能夠在虛擬情境中模擬現實中的火災逃生經驗。研究的關鍵要素包括火災勢態評估、適當的逃生方式選擇以及現實世界中的火災應對技能融入。(1)火勢嚴重程度評估：在遊戲情境中，參與者需精密審視虛擬火災場景，細致考察火勢擴散範圍、煙霧密度等諸多因素，以有效評估火勢的嚴重程度。此反映出玩家對於火災情境的深入觀察及分析能力，並迫使其根據所得觀察結果做出審慎而明智的抉擇。(2)逃生方式選擇：根據火勢的嚴重程度，遊戲提供了不同的逃生方式選擇。這涵蓋了窗外求救和自我逃生兩種主要方法。遊戲的挑戰在於，玩家必須在有限的時間內根據情況快速而明智地選擇最適合的逃生方式，以確保自身安全。這反映了火災應對策略的重要性，並提高了玩家的決策能力。(3)現實應對方式融入遊戲：模擬現實世界火災逃生情境，引入實際應對方法，包括低姿勢前進、煙霧回避、閉門、滅火器使用以及逃生指示的考量。透過虛擬情境的模擬，玩家得以實際體驗並學習根據火勢狀態作出適切反應，有助於培養其火災應對技巧和直覺。

表 1 關卡遊戲說明及場景與關鍵火災處理概念的設計

操作介面	關鍵學習概念設計	遊戲展示	說明
遊戲起始頁面	提高學習興趣		配色對比強烈，吸引玩家注意。
第一、二關遊戲前導動畫	熟悉遊戲背景		前導動畫引導玩家熟悉遊戲背景與操作，並繪製可愛風格的 2D 畫帶入玩家進入遊戲。

			
第一關	學習目標： 關上房門/帶口罩/撥打 119	 	第一關旨在於基本的火場逃生方式，透過對話指示學習逃生必要步驟。
獨立事件	根據火災勢態做出適當的反應		告知玩家遇到電梯不可搭乘原因。
任務介面			協助玩家了解任務需求與目標
第二關	學習目標： 離開房間/使用滅火器/跟隨逃生指示/低姿勢逃生	  	第二關擁有更多的學習目標與逃生技巧，需要玩家通過第一關後才可繼續參與，漸進式引導玩家掌握逃生觀念。

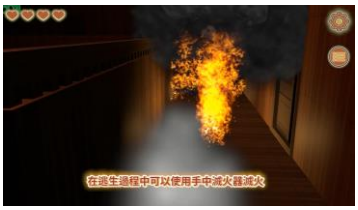
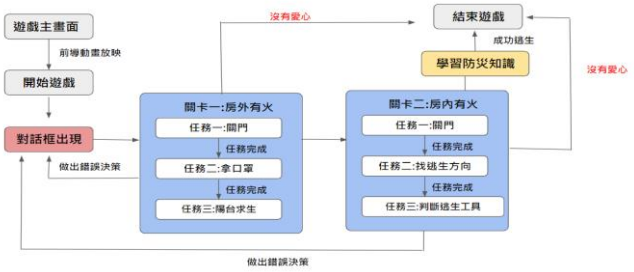
			
遊戲通關成功/失敗畫面	提供重新挑戰的選項，鼓勵玩家再次嘗試，並應用在前次失敗中所學到的知識和技能。	 	給予玩家通關與否的回饋，也代表玩家成功習得火災逃生知識。

圖 1 遊戲概念流程圖

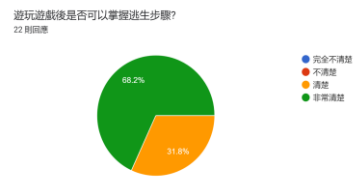
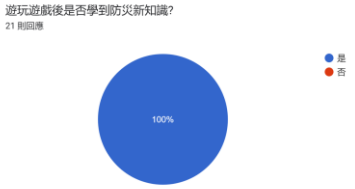
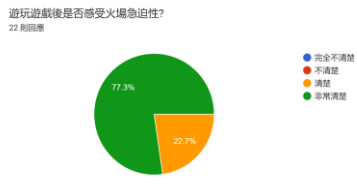
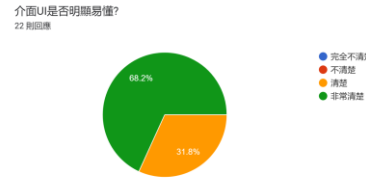
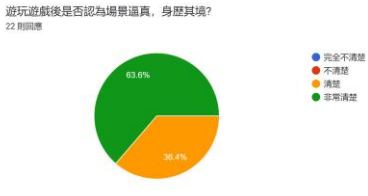
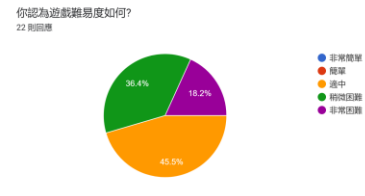


4. 結果與討論

我們原本計劃使用國小高年級年齡的玩家來驗證學習成效。然而由於時間限制，我們採取了一個臨時性的替代方法，將問卷分發給試玩遊戲過的 31 位大學生，以評估他們對於相關學習遊戲的觀感和理解。儘管這不是最理想的驗證方式，但我們相信這將提供一些初步的洞察和資訊，以支持我們對於該遊戲對於國小年齡學習者潛在影響的理解。

表 2 問卷回饋收集

遊玩遊戲後是否可以掌握逃生步驟?	遊玩遊戲後是否學到防災新知識?

	
遊玩遊戲後是否感覺火場急迫性?	介面 UI 是否明顯易懂?
	
遊玩遊戲後是否認為場景逼真?	你認為遊戲難易度如何?
	

4.1. 結合火災教育遊戲關卡設計分析

綜觀前述問卷分析，我們歸納出以下五項調整與修改建議：1.FIRE MENTOR 一款 3D 火災教育遊戲，因此在設計建模時，需不斷測試整體 3D 空間的規劃，包括尺寸、大小、形狀...等等，避免造成玩家在遊玩過程中受到空間的限制而分散對於遊戲的專注力，也會間接影響到玩家對於此款遊戲的興趣。2.我們希望能讓玩家有一種身歷其境的感覺，因此有設計音效等特效讓玩家有一場良好的遊戲體驗。3.在 UI 設計方面，我們希望能夠增加玩家對於遊戲的興趣，希望設計出更有活力的樣式。4.遊戲過程中適時給予玩家遊戲推動力，製造一些對話框情境，讓玩家身歷其境，並給予任務目標與危機感。並加入血條系統，在完成目標任務的同時讓玩家產生成就感，因而增加學習成效。5.對話框出現時遊戲中的火焰蔓延便會停止，避免在有時間壓迫的情況下吸收過多的資訊。

5. 結論與建議

本研究結果發現，所有研究對象均認為遊玩遊戲後可學習到防災新知且大多數人確實掌握了逃生的步驟，本遊戲提供相應於火災防災教育之相關知識，以提升學習者對於火災防災教育課程的興趣與學習效益。

FIRE MENTOR 於遊戲內進行多元情境模擬，即時培育學習者對於各種火災狀況的鑑別辨識與深刻理解之能力，同時使玩家習得精確的應對措施。透過此遊戲激發學習者積極主動地進行探索並參與互動的學習態度，同時提升其火災防災專業素養。此外，本研究計劃亦可投入教育部防災資訊網之資訊專區，以提供更多教育資源和信息。為了進一步評估本遊戲的

理解程度和學習效果，我們計劃參加 2024 年的「放視大賞」遊戲組比賽。這將為不同的年齡層或業界人才提供評估我們遊戲的機會，同時也希望通過比賽推廣和推動本遊戲的普及。

最終，透過遊戲的主動學習，深刻根植防災觀念於生活中，同時在此過程中反思並運用所學，期冀發展這類遊戲教材能夠為相同領域的研究者注入新的動力，以促進防災教育的實踐推廣，並激發更多不同的實踐方法。

參考文獻

洪志誠(2003)。**颱風與水災防災教育教材編撰(I)十二年一貫教育**。臺北市立教育大學地球環境暨生物資源學系

張玉連(2021)。**以情境模擬遊戲實施防災教育之教學與成效**。國立臺灣大學工學院土木工程學系博士論文。

張基成、林冠佑(2016)。**從傳統數位學習到遊戲式數位學習－學習成效、心流體驗與認知負荷**。

郭俊欽、莊翰華、康良宇(2007)。**土石流災區之防災教育初探**。環境與世界, 15 期 (2007/06) ，頁 99-128。

楊鎮豪(1998)。**個人電腦遊戲軟體消費者購買涉入與消費行為之研究**。國立交通大學管理科學研究所碩士論文。

楊瑞賓、李香儀、陳蓉蓉、陳怡艮、黃莞茗(2022)。**互動式冒險遊戲設計與開發**。幼兒運動遊戲年刊, 2 期, 頁 64-114。

趙慕芬、翁頂升(2021)。**評析數位遊戲式學習在高等教育中商管領域的應用與發展**。臺灣教育評論月刊，2021，10（2），頁 113-119。

劉以琳、劉婉儀(2013)。**冒險遊戲玩家心理特質與遊戲投入之研究**。南台學報, 38 卷 2 期, 頁 113-126。

戴于真、張懷綾、金凱儀(2020)。**結合多媒體技術與遊戲機制的數位學習系統－以淡水發展史為教學內容**。工程、技術與科技教育學術研討會, 第九屆(2020/06)，頁 175-186。

Alessi, S. M. & Trollip, S. R. (1985). Computer-based instruction:Methods and development.

Colarusso, C. A. (1993). Play in adulthood: A developmental consideration.

Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model.

Hanson-Smith, E. (2020). Pwyll Prince of Dyfed: the narrative structure.

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). Report on the educational use of games.

Nicola,W. (2007). Motivation and computer game based learning.

Prensky, M. (2003). Digital game-based learning.

游戏化学习策略对大学生学习动机和学习投入的影响研究

A Study on Effects of Gamified Learning Strategies on College Students' Motivation and Engagement in Learning

张晓旻，李秀晗*，刘巧伶，张颖
华中师范大学人工智能教育学部
1194765416@qq.com

【摘要】 游戏化学习策略的应用效果一直以来受到关注，它对大学生的影响需要实证证据支撑。本研究旨在调查游戏化学习策略的实施对大学生学习动机和学习投入的影响，选择 76 名本科生开展了半学期的准实验研究，结合线上工具平台进行游戏化教学活动，利用问卷和访谈调查他们的学习动机和学习投入。结果表明：（1）游戏化学习策略能有效提升大学生的学习投入；（2）高低绩效组之间的外部动机存在显著差异，低绩效组的外部动机更高。本研究结合访谈内容对以上结果进行解释，并针对实施中的问题给出高校游戏化学习策略实施的具体建议。

【关键词】 游戏化学习；学习动机；学习投入

Abstract: The application of gamified learning strategies has long been a concern, empirical evidence is needed to understand its influence on college students. This study aims to investigate the impact of gamified learning strategies on college students' motivation and engagement. A quasi-experimental design was conducted with 76 undergraduates over half a semester, combining gamified teaching activities with online platforms. Questionnaires and interviews were used to assess their motivation and engagement. The findings indicate that: (1) gamified learning strategies effectively enhance college students' engagement; (2) there is a significant difference in external motivation between high-achieving and low-achieving groups, with the latter exhibiting higher levels of external motivation. Based on interview analysis, this study provides specific recommendations for implementing gamification in higher education.

Keywords: gamified learning, learning motivation, learning engagement

1.引言

游戏化与教育的融合历史已久，早在 20 世纪 80 年代电子游戏就开始受到教学研究者的重视（曹琳琳，2022）。随着时代的推进与技术的革新，游戏进入教育领域的路径愈加明晰。然而，研究者们大多将目光放在学前教育和 K-12 教育上，且有部分研究集中在对游戏化理论和模式的探讨上,证实游戏化学习策略在高等教育真实课堂中应用效果的更是少之又少。与学前教育和 K-12 教育相比，高等教育中的学生更成熟，能够吸引他们的游戏元素与机制也发生了变化，游戏化学习策略在高校中的应用真的有效吗？贺宝勋（2021）等人证实了游戏化评价具有缓解大学生在线学习倦怠，进而提升在线学习成绩的潜质；韩飞飞（2022）探究了游戏化学习与学习投入度之间的关系及作用机理；鲜丹丹（2021）等人探究并验证游戏化理念下的翻转课堂教学模式对学生自我效能感及学习动机的影响。学习动机是学生学习质量评估的重要过程性指标（常静艳，2023），学习投入是指个体的一种与学习相关的积极、充实的精神状态,是学业倦怠的对立面（倪士光、伍新春，2011），这二者在一定程度上反映了学习

效果。因此，本研究在已有基础上进一步探索游戏化学习策略对大学生学习动机和学习投入的影响，主要研究问题为游戏化学习策略能否有效提高大学生的学习动机和学习投入？旨在为高校游戏化学习策略的有效性提供实证证据，并发现实施游戏化策略中的问题，凸显高校游戏化教学的重要性。研究强调游戏化教学是提升学习动机和投入度的有效形式，对于促进大学生社会化学习和协作学习有很深远的意义。

2.研究方法

本研究主要采用准实验设计来探究游戏化学习对高校教学课堂和大学生的影响。实验对象为选修“现代教育技术”课程的本科学生（年龄在 20 岁左右），人数为 76 人，开展为期半学期的实验研究。学生学习能力及心智成熟程度均相似，初始能力相当，且都基于同一平台（小雅）开展教学的学习活动，选择游戏化教学模式。我们根据课程内容设计了不同类型的游戏化活动：“趣味抢答”“猜猜他是谁”“超级词典”“智慧分享”“胡编乱造”“畅想未来”等十多种游戏，每堂课中组织学生开展与教学内容适配的游戏化活动，并利用积分、排行榜等常用游戏化元素辅助游戏评价。本研究采用问卷调查法和访谈了解大学生的学习动机和学习投入，还对实验班学生的游戏化学习活动偏好做了简单的调查。本次实验所需的问卷题目结合了被试者特征，并参考相关文献选择合适时间点发放给学生。因为有两份问卷因为缺少对应前测/后测数据，被归为无效数据，回收问卷 74 份，回收率为 97.4%。

利用 SPSS 软件对“大学生学习动机与学习投入量表”数据进行一致性检验和探索性因子分析，得出的结果如表 1 所示。前后测的量表的 Cronbach’s Alpha 和 KMO 值均大于 0.65，表明大学生学习动机与学习投入量表的信效度良好。在此基础上我们对调查结果做进一步分析。

表 1

大学生学习动机与学习投入量表的信效度检验

	实验班
前测 Cronbach’s Alpha 系数	0.684
后测 Cronbach’s Alpha 系数	0.862
前测 KMO	0.684
后测 KMO	0.744

具体的量表设计包含学习动机和学习投入，其中学习动机又分为外部动机和内部动机，具体题项设置如表 2 所示。学习动机量表参考了 Amabile（1994）和池丽萍（2006）等人的研究成果,从中挑选12题。学习投入量表结合公共课学习情况和方来坛(2008)修订的Schaufeli（2009）量表，制成 5 题。本研究采用 6 点量表（完全不符合，不符合，稍微不符合，稍微符合，符合，完全符合）调查大学生的学习动机和学习投入。其中 A1 为反向题项，得分越高表示外部动机越低，其余均为得分越高表示外部/内部动机越高。

表 2

量表题项具体描述

类别	题号	题项
外部动机	A1	在公共课上，我并不那么在乎别人对我的学习表现有什么看法
	A2	我比较喜欢在学习中有人替我设定清楚的目标
	A3	我十分清楚我的目标或目的是追求好成绩
	A4	我喜欢相对简单而直接的学习任务或作业
	A5	能赢得他人的肯定和赞赏是推动我去努力的主要动力

内部动机	A6	越困难的学习内容和问题，我越乐于尝试解决它
	A7	我希望课程学习能够提供我增加知识与技能的机会
	A8	我喜欢在公共课上独立思考解决疑难
	A9	我十分清楚我自己在学业成绩上要达到的目标
	A10	我做许多事情都是受好奇心驱使，能够学习我所感兴趣的内容是很重要的
	A11	当我能自己设定我的学习目标时，我会更为认真和投入
	A12	只要做的是我乐于做的事，我不那么在乎分数和奖励
学习投入	B1	我在公共课上能够积极参与，不开小差
	B2	在完成公共课课程作业时，我觉得全身心头投入和精力充沛
	B3	我对公共课的学习任务充满兴趣和激情
	B4	我经常针对公共课上的内容在课外进行扩展性学习
	B5	当完成公共课作业时，我有一种意犹未尽，还想继续做这些作业的感觉

3.研究结果

3.1.学习动机与学习投入描述性结果统计

为了进一步了解实验班学生在不同测量方面的积极性情况，对每个题目积极趋向（问卷选择“符合”，“完全符合”）和消极趋向（问卷选择“完全不符合”，“不符合”）的人数进行统计分析，结果如图 1 和图 2 所示。从两张图的纵来看，积极趋向选项占比大于消极趋向，说明整体上学生的学习动机和学习投入较好。

图 1 为消极趋向选项在各个题项上的分布情况，可以看出实验后整体上消极趋向选项占比有所下降，这说明游戏化学习策略在一定程度上改善了大学生在学习动机和学习投入上的消极态度。其中 A1-在公共课上，我并不那么在乎别人对我的学习表现有什么看法的不认同人数占比由原来的 21.62%下降至 9.46%，说明游戏化学习策略的实施降低了大学生对他人的外部期待，更关注自身。后测中，A4-我喜欢相对简单而直接的学习任务或作业、A7-我希望课程学习能够提供我增加知识与技能的机会、A9-我十分清楚我自己在学业成绩上要达到的目标、A10-我做许多事情都是受好奇心驱使，能够学习我所感兴趣的内容是很重要的、A11-当我能自己设定我的学习目标时，我会更为认真和投入、B3-我对公共课的学习任务充满兴趣和激情的消极趋向选项均为 0，说明实验班的学生在外部动机上都喜欢简单直接的学习任务，内部动机中关注知识与技能的提升、个人目标的达成和兴趣爱好，学习投入上对学习任务充满兴趣和激情。

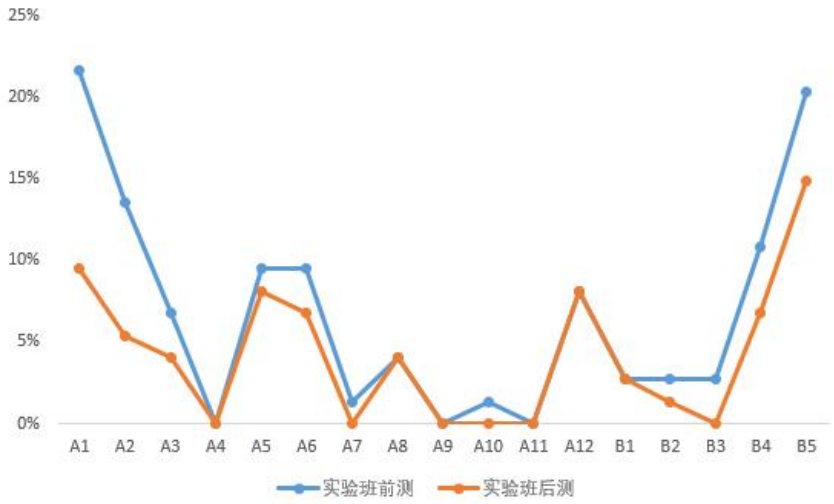


图 1 消极趋向选项分布情况

图 2 为积极趋向选项在各个题项上的分布情况，可以看出 B1-B5 的学习投入积极趋向占比均有所提升，这说明游戏化学习策略在提升大学生学习投入上产生效果。与消极趋向选项分布不同，学习动机后测占比并不都大于前测，其中 A2、A3、A6、A8、A12 相较于前测有了明显提升。可以看出，实施游戏化学习策略之后外部动机中的目标、成绩受到学生关注，内部动机中的个人兴趣和喜好有了提升。但是 A4-我喜欢相对简单而直接的学习任务或作业、A7-我希望课程学习能够提供我增加知识与技能的机会、A10-我做许多事情都是受好奇心驱使，能够学习我所感兴趣的内容是很重要的后测明显低于前测，这可能是因为游戏化的学习形式化解了学生对复杂任务的恐惧和逃避心理，有趣的教学形式能使学生不仅仅关注自己感兴趣的内容，而专注于教师设置的教学内容。

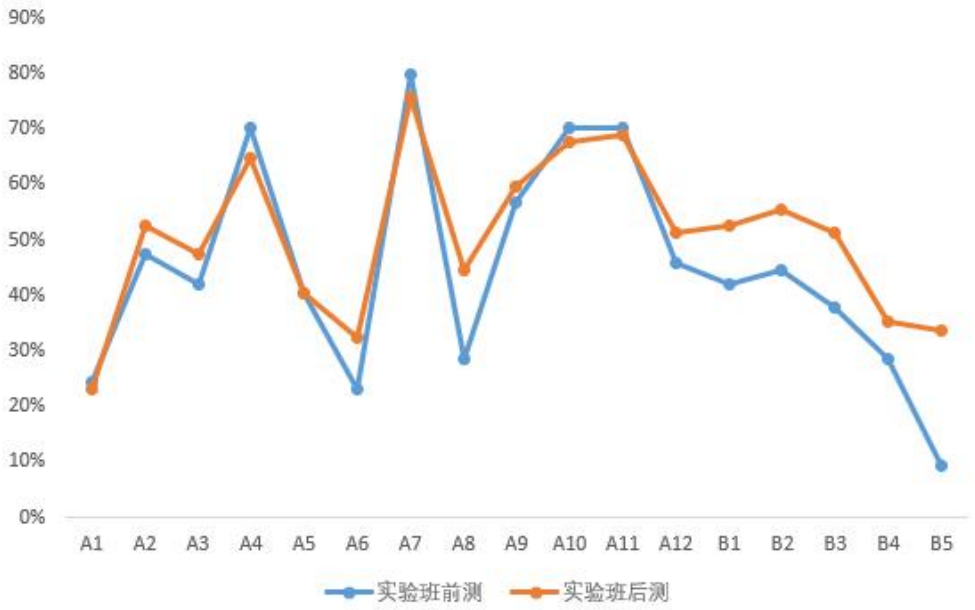


图 2 积极趋向选项分布情况

本研究将大学生在学习动机和学习投入上的题项得分进行统计，结果如表 3 所示：A7 得分最高，说明对于大学生来说增加知识与技能非常重要；B5 得分最低，表示在完成作业上学生的主动投入较低。从均值情况来看，实施游戏化策略后的学习动机和学习投入得分整体上高于前测，但是 A4、A7、A11、A12 的得分与前测相比有所下降，这个结果与积极趋向选项分布有相似之处。大学生门本能的喜欢简单而直接的学习任务和作业，因而在进行前测时得分较高，游戏化课堂的应用可能使他们体会到了不同层次游戏任务的乐趣，降低了对偏困难复杂任务和作业的逃避心理，因此得分有所下降。知识与技能的提升虽然是学生们最重视与最关注的点，但游戏化增加了学习过程中的学习体验，不再单纯的出于功能性目的进行学习活动。游戏实施前教师会设计安排目标，且过程中用到积分、奖励等游戏化元素，因此实施后 A11、A12 的均值有所下降。

表 3

学习动机和学习投入描述统计结果

题项	前测		后测	
	M	SD	M	SD
A1	3.58	1.29	3.81	1.02
A2	4.20	1.22	4.46	1.10
A3	4.18	1.13	4.42	0.97
A4	4.92	0.92	4.82	0.85

A5	4.15	1.16	4.26	1.07
A6	3.72	1.09	4.09	1.04
A7	5.12	0.88	4.97	0.76
A8	4.18	0.88	4.39	0.96
A9	4.66	0.86	4.72	0.88
A10	4.86	0.93	4.88	0.89
A11	4.92	0.82	4.88	0.84
A12	4.36	1.19	4.31	1.13
B1	4.38	0.84	4.51	0.91
B2	4.45	0.83	4.55	0.97
B3	4.28	0.79	4.41	0.92
B4	3.88	1.09	4.12	1.10
B5	3.32	1.05	3.80	1.28

3.2.学习动机与学习投入差异性分析

本研究对实验班的前后测得分进行配对样本统计及检验，结果如表 4 所示。可以看出，游戏化教学策略实施后学习动机和学习投入的均值都有所提升。“动机前测 - 动机后测”的 p 值>0.05，说明学习动机在前后测上不存在显著差异；“投入前测 - 投入后测”的 p 值为 0.04（<0.05），说明学习投入在前后测上有显著差异，换言之，游戏化学习策略显著提升了大学生的学习投入。

表 4

配对样本统计及检验

	M	SD	t	p
动机前测	4.40	0.42		
动机后测	4.50	0.53	/	/
投入前测	4.06	0.73		
投入后测	4.28	0.88		
动机前测 - 动机后测	-0.10	0.50	-1.68	0.10
投入前测 - 投入后测	-0.22	0.89	-2.09	0.04

为了检验高低绩效组在学习动机和学习投入上的差异，我们将期末测试中成绩排名前10%的学生定义为高绩效组，排名后 10%为低绩效组，对他们的学习动机和学习投入做独立样本 T 检验，结果如表 5 所示。其中 1 表示低绩效组，2 表示高绩效组。

表 5

高低绩效组在学习投入和学习动机上的差异

	分组	M	SD	t	p
学习投入	1	4.23	0.82	-0.14	0.89
	2	4.29	0.68		
外部动机	1	4.43	0.37	2.24	<0.05
	2	3.86	0.56		
内部动机	1	4.71	0.34	0.88	0.40
	2	4.53	0.43		

低绩效组-1，高绩效组-2

结果表明高低绩效组在外部动机上存在显著性差异，低绩效组（M=4.43）的外部动机显

著高于高绩效组（M=3.86）。在学习投入和内部动机上，结果未显示出显著性差异。为什么外部动机高的学生绩效反而越低呢？根据推测，不同的学生对游戏化教学策略的关注点不同，有些学生对其关注点停留在积分、奖励、竞争等外部元素和形式上，即“浅层游戏化”(Lieberoth, 2015)，他们在游戏过程中过分追求游戏目标而忽略对知识内容的深层理解，导致外部动机虽然有显著提升，但是并没有和教学内容相结合，因而在测试中成绩更低。此外，学习者本身的学习能力和风格也不同，高能力学习者不容易受到教师教学方式变化的影响，他们有自己的步调和习惯，学习成绩稳定性更强。

4.讨论

4.1.游戏化学习策略应用效果的原因分析

结合实证研究结果来看，游戏化教学策略的实施在提升大学生学习投入上有显著效果，这与 Ab Rahman（2019）等人的研究结果一致。游戏化元素的使用从外在形式上给学生提供持续的刺激，吸引他们持续参与到学习活动中来。游戏化学习的主要特点是利用不同方式的刺激（等级、积分、徽章、奖励等）来促进学习（Upadhyaya & Garg, 2019）。学生们的学习体验得到改善，从而能比较持续的投入到学习活动中去。但是学习动机并没有显著提升，结合访谈结果来看，可能有以下原因：①游戏化教学持续时间较长导致学生出现审美疲劳。本研究在每节课设置了不同的游戏化活动，持续时长为半学期，有同学反映“游戏化模式或别的，玩的多了会疲劳，丧失新鲜感”。②游戏化小组任务责任分散，游戏任务集中在团队中的某部分人身上，其他成员没有实际参与到游戏化教学活动中去。③实施的某些游戏本身不够有趣，吸引力不足。有同学表示趣味抢答的游戏化形式本质上并没有改变答题这种传统形式，动力不足。④期末考核压力导致后半程游戏化教学实施效果不佳。由于实验是在后半学期开展的，学生们会面临期末考试的压力而更关注成绩，无法全身心参与到游戏中去。与以往认知不同，研究结果表明高低绩效组的外部动机存在显著差异，低绩效组显著高于高绩效组。为什么会出现外部动机高的学生绩效反而低这种现象呢？首先学生绩效并不与学习动机呈线性相关，除了动机，绩效还会受到学生学习素养、学习行为、学习风格、教师投入等多方面的影响。其次，外部动机高一定程度上意味着学生关注外部因素如积分、奖励等，而忽略了知识内容的深度理解、与现实的融合，学生可能对学习内容并没有过多的印象或者只是零碎的知识点，无法建构起自己的知识体系。

4.2.高校游戏化学习活动实施建议

我们在学期结束后对实验班同学进行了访谈，了解本次游戏化学习活动实施的情况与不足，主要概括为以下几点：外在环境上，互动硬件设施有待提升。不少同学表示小雅平台的抢答功能不流畅，长期抢不到题目会挫伤积极性，例如在问到课程设计的不足时，有同学说“我觉得可能更多的还是偏向于技术方面或者说选用平台的不足，比如小雅平台的抢答，它的随机性太大了，机会不均等”。游戏设计上，部分游戏活动目标不明确、规则不清晰，学生在开展活动时找不到抓手，导致学生行为与教师预期不符，游戏化效果一般，有同学访谈时表示“没有完全理解，没有明确规则”；还有学生表示游戏化活动中的成绩直接与学分绩点挂钩，容易给学习者造成压力而无法享受游戏化互动。从学生本身来看，有些人的课前预习并不到位，导致某些与学生先验知识相关的游戏化活动效果不尽人意，学生体验感不强。

针对以上问题和实证研究结果，我们为高校游戏化学习活动的实施提出建议。首先，选择合适的游戏化开展场地和平台。线下游戏化教学活动可以选择更具灵活性的智慧教室，便于学习者展开游戏和讨论；提升在线游戏化平台的流畅性，教师充分发挥主观能动性调节由于网络等原因带来的机会不均等。其次，制定明确的游戏化活动规则，必要时给出游戏示例，引导学生有条不紊的展开游戏活动；还要注意游戏化任务分配的合理性，保证小组成员的充

分参与；游戏设计时明确学生需要掌握的先验知识，设置合理的课前学习活动和测验确保游戏化活动能顺利开展。此外，游戏活动的成绩与奖励需要充分考虑学生特征和课堂性质，增加游戏化活动结果与课程成绩的适配性。最后，实施游戏化教学时要想办法调动学生将知识与游戏联系起来，促进知识内容的深度理解。“游戏化学习”中“游戏化”是手段，是过程，最终目标是“学习”，在游戏化学习活动设计和实施时要牢牢把握这个理念，结合学生情况和教学情况灵活调整。

5.总结

总的来说，游戏化学习策略的实施能有效提升大学生的学习投入，对大学生学习产生了积极影响。但是，可能会出现提升外部动机而无法提升学业成绩的情况。这启示我们要注重利用游戏化培养学生的内外部动机，促进学习者对知识的深度理解和实践应用能力，而不仅仅停留在“浅层游戏化”上。本研究根据学生的访谈内容和课堂观察给出了高校游戏化活动实施的具体建议，为游戏化活动的开发提供一定的借鉴和参考。当然本研究也存在一定的局限性：游戏化活动设置在后半学期展开，期末的压力影响学生的状态，效果可能不如预期；利用量表和访谈来了解学生学习动机和学习投入，缺乏生理数据的支持；缺乏对照组的横向比较。但是这只是我们的初步研究，我们会在后期进一步完善实验方案，进一步探索游戏化学习策略在高校中的应用，构建出一个高校游戏化学习模型。

参考文献

方来坛、时勘 & 张风华.(2008).中文版学习投入量表的信效度研究.中国临床心理学杂志(06),618-620.

池丽萍 & 辛自强.(2006).大学生学习动机的测量及其与自我效能感的关系.心理发展与教育(02),64-70.

贺宝勋、张立国 & 庄科君.(2021).游戏化评价对大学生在线学习倦怠及学习成绩的影响研究.电化教育研究(03),62-68+119.doi:10.13811/j.cnki.eer.2021.03.009.

倪士光 & 伍新春.(2011).学习投入:概念、测量与相关变量.心理研究(01),81-87.

曹琳琳.(2022).面向计算思维培养的大学生在线游戏化编程学习活动设计研究硕士(学位论文,东北师范大学).硕士 https://link.cnki.net/doi/10.27011/d.cnki.gdbsu.2022.001487doi:10.27011/d.cnki.gdbsu.2022.001487.

常静艳.(2023).我国大学生学习动机的演化规律——基于“国家大学生学情调查”(NCSS)的分析.终身教育研究(05),22-30.doi:10.13425/j.cnki.jjou.2023.05.003.

韩飞飞.(2022).移动技术支持下游戏化学习对大学生学习投入度的影响研究.沈阳大学学报(社会科学版)(01),62-69.doi:10.16103/j.cnki.21-1582/c.2022.01.008.

鲜丹丹、黄冠 & 何江.(2021).游戏化理念下的翻转课堂对大学生自我效能感及学习动机的影响.中国教育信息化(14),6-12.

Ab Rahman, R., Ahmad, S., & Hashim, U. R. (2019). A study on gamification for higher education students' engagement towards education 4.0. In *Intelligent and Interactive Computing: Proceedings of IIC 2018* (pp. 491-502). Springer Singapore.

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. M. (1994). The Work Preference Inventory: assessing intrinsic and extrinsic motivational orientations. *Journal of personality and social psychology*, *66*(5), 950.

Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture*, *10*(3), 229-248.

Schaufeli, W. B., Bakker, A. B., & Van Rhenen, W. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *30*(7), 893-917.

Upadhyaya, D., & Garg, A. (2019, September). Leveraging on gamification and learning analytics for improved student learning. In *2019 International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 132-138). IEEE.

小学生 STEAM 项目化学习中科学探究能力培养的教育游戏设计

Educational Game Design for Cultivating Scientific Inquiry Skills in STEAM Project-based Learning for Elementary School Students

张蕊华[1]，郭子霏[2]，张新立[3*]，黄瑞婷[4]
[1234]温州大学教育学院
[*] zhxl@wzu.edu.cn;

【摘要】 时代的发展对人才的核心素养提出新要求，STEAM 教育以生活中的真实问题为导向，在科学探究中掌握基础知识点，在动手实践中强化知识的运用，在解决问题的过程中培养具有综合素养的科技创新人才。项目化学习是实施 STEAM 教育的主要途径，然而小学生在 STEAM 项目化学习中自主知识建构环节科学探究能力明显不足。教育游戏对提高学习兴趣、激发学习动机、提高知识理解与保持力等方面有显著作用—因此本研究在相关理论指导下，使用 RPG Maker MV 平台设计并开发了一款教育游戏以提高小学生科学探究能力，并在 STEAM 项目实践中检验了其有效性。

【关键词】 STEAM 教育；教育游戏；游戏开发；小学生；科学探究能力

Abstract: *The evolving landscape demands new skill sets from talents, with STEAM education addressing real-life challenges, fostering scientific inquiry, and promoting practical application of knowledge. It aims to nurture well-rounded innovators capable of problem-solving. While project-based learning is key, elementary students often lack proficiency in independently constructing knowledge during STEAM projects. Educational games significantly enhance learning engagement, motivation, and comprehension. Thus, guided by relevant theories, this study employs the RPG Maker MV platform to design an educational game. Its aim is to enhance elementary students' scientific inquiry skills and evaluate its efficacy within STEAM projects.*

Keywords: STEAM, educational games, game development, elementary school students, **Scientific inquiry skills**

1. 前言

STEAM 教育跨学科整合的教学模式颠覆了传统课堂的教学模式及理念，成为了培养学生适应未来关键能力的有效新路径。STEAM 教育以现实生活中的真实问题为导向，在科学探究中掌握基础知识点，在动手实践中强化知识的运用，在解决问题的过程中培养创新意识、工程思维等 21 世纪所需人才具备的核心素养。尤其在项目式的 STEAM 教学环节中的“知识探究环节”，强调让学生“像科学家工作般地学习”，注重的是正确运用科学探究方法，按照科学探究的基本步骤进行相关知识点的探究，以此逐步建立有序的科学探究意识，从而逐渐形成科学探究能力，最终为科学素养的养成打下良好的基础。但很多小学生在这一环节中并不知道如何进行探究，从而导致在这一关键环节中的学习效果并不是特别理想。

鉴于以上研究背景和问题，本研究针对 STEAM 项目化学习的“知识探究环节”，在建构主义学习理论、游戏化学习理论、需求层次理论、情境理论以及沉浸理论的指导下，探讨如何设计并开发以科学探究六个步骤为知识框架的教育游戏的过程，具体教育游戏开发的设计思路、设计原则以及应用策略，同时开发该游戏应用的实践案例，以检验该游戏在培养小学生科学探究能力方面的实际效果。

2. 文献综述

2.1. 概念界定

2.1.1. 科学探究能力

对于在科学探究概念的界定，不同学者或研究机构从不同的角度给出了不同的定义。美国颁布的《国家科学教育标准》倡导学生要像科学家一样思考，提出科学探究指的是学生们用以获取知识、领悟科学的思想观念、领悟科学家们研究自然界所用的方法而进行的各种活动，主要包含五个基本特征，即问题、证据、解释、评价和表达。因此，我们可以认为科学探究能力是在进行科学探究活动中必备的也是能逐步形成的一种特别的能力，也可以说是科学探究活动进行的根本保证。

美国奥勒冈州教育部门将科学探究能力概括为四个部分，即形成问题或假说、设计研究、收集并呈现数据和分析并解释结果。(刘东方 & 王磊, 2012)认为科学探究能力的构成要素包括提出问题能力、猜想与假设能力、制定计划能力、收集证据能力、解释与结论能力以及反思与评价能力和表达与交流能力。

2.1.2. 基于游戏的学习

基于游戏的学习（Game-based Learning）狭义上指将游戏尤其是电子游戏应用到学习中；广义上是指将教育游戏（Education Game）融入课程教学和自主学习活动，把游戏作为师生沟通平台，从传统单向说教模式转化为互动交流模式，让学生在轻松愉悦的学习体验中获得知识、技能，以培养正确情感态度的过程(张金磊 & 张宝辉, 2013)。基于游戏的学习可以归为游戏化学习领域中，所谓游戏化（Gamification）是将游戏理念、游戏元素和游戏机制应用到非游戏情境中的过程，游戏化学习即将游戏融入教学活动中达到“乐学”的结果。本研究所涉及到的基于游戏的学习就是指狭义上的概念，将开发完成的教育游戏依托平板这一媒介应用于课堂教学中，通过教育游戏进行知识点的学习与训练。

2.1.3 STEAM 项目化学习

STEAM 教育就是集科学（Science）、技术（Technology）、工程（Engineering）、艺术（Art）、数学（Mathematics）于一体的综合教育。STEAM 教育以现实生活中的真实问题为导向，在科学探究中掌握基础知识点，在动手实践中强化知识的运用，在解决问题的过程中培养创新意识、工程思维以及问题解决能力等 21 世纪所需人才具备的综合素养。

项目式学习，也可以称为基于项目的学习，英文为 Project-based learning，缩写为 PBL。对于项目式的学习概念的界定，国内学者李梅将它总结为是旨在使学生获取需要的知识以解决出现在项目中的问题，从而实现有意义的学习，而并不是始终接受二手知识的被动学习(李梅, 2017)。不难看出，这与 STEAM 所具有的跨学科、情境性等理念重叠，正如国内学者余胜泉所说：STEAM 存在着科学、技术、工程、艺术和数学之间的跨学科交叠，具有天然的跨学科性、情境性、实践性、协作性、实证性，尤其适于立足生活经验和社会实践取向进行 PBL 学习(余胜泉 & 胡翔, 2015)。因此，基于项目的 STEAM 学习或者 STEAM 项目化学习成为了当前 STEAM 学习模式中的主要模式之一。

2.2. 小学生科学探究能力培养国内外研究现状

针对科学探究能力培养的研究，国内外的学者大多将精力集中在了科学探究能力培养的模型建构以及实证研究方面。

在国内，(李春密, 2004)根据科学探究能力的定义及其特点，参照基尔福特的智力结构图，建构了科学探究能力的结构模型，并分析了模型的特征即其整体性、静态性、动态性和自调性，提出了科学探究能力的主要指标。在国外， (Hsu et al., 2016)进行了为期两年的研究，探讨在技术融合的学习模式中如何发展学生不同的探究能力。从国内外已有的关于学生探究能

力培养的研究来看，研究者多致力于培养科学探究能力的模型的建构，或应用于具体学科探究式课堂的教学模式的提出与改进，尤其我国大部分将其放在科学课程中，但很少有从综合的角度去研究小学生科学探究能力的培养效果。科学探究能力的培养也并非是单一学科的任务，因此，笔者认为科学探究能力的培养途径应当拓宽。

2.3. 教育游戏的国内外研究现状

近几年新媒体联盟《地平线报告》中也多次提及教育游戏以及将以游戏为基础的学习视为影响教育发展的关键技术。对于游戏融入教育对学习者自身或者学习成就等方面是否有影响，国内外已经有大量的实证研究为此提供了证明。

国内的相关研究表明基于教育游戏的游戏化学习能够支持个性化学习，符合学习者认知心理特征，帮助学生对抽象知识的理解，其趣味性、挑战性和高效性等特点可以大大提高学生学习的兴趣、动机、主动性和实践动手能力，强化学生的工程意识(李洋, 豆丹阳, & 王健, 2019; 李宜逊 et al., 2017; 于兰兰, 谭博学, 张厚升, & 苗飞, 2016)，把游戏作为师生沟通平台，从传统单向说教模式转化为互动交流模式，让学生在轻松愉悦的学习体验中获得知识、技能，以培养正确情感态度的过程(张金磊 & 张宝辉, 2013)。

在国外，学者们研究了在基于游戏的学习环境中，“心流”(可操作化为更高的挑战和技能)、参与感和沉浸感对学习的影响，发现游戏参与对学习有明显的正向影响，(Esther del Moral Perez, Guzman Duque, & Carlota Fernandez Garcia, 2018)研究发现在游戏化学习中儿童的数学逻辑、自然主义和语言学习水平存在显著差异，各项指标普遍提高。基于游戏的学习环境可以为教师提供一个理想的补充工具，丰富他们的方法，发展灵活性和适应性的问题解决能力。(Chen, Tsai, & Chang, 2019)探讨了情境模拟游戏与电子教材对小学生学习成效的影响，发现了与情景模拟游戏教学相比，电子教科书的使用对学习效果的影响较小。另外，(Wu, Tzeng, & Huang, 2020)认为教育游戏是影响学生情感体验的重要方式与工具，(Fu, Lin, Hwang, & Zhang, 2019)认为在支持个性化学习和优化教学环境方面极具潜力与前景。

国内外研究表明，游戏以其自身的趣味性、挑战性、沉浸性等特征可以提高学生的学习兴趣，激发学生的学习动机，以及对学生的知识保持效果、情感态度价值观的培养、创造力和问题解决能力等高阶思维能力的提升都有正方向的影响，大量的实证研究证明了游戏融入学习对学习者的学习有着正向的帮助，因此，笔者认为教育游戏也可以成为培养小学生科学探究能力的一个作用力。

3. 小学生 STEAM 项目化学习中科学探究能力培养的教育游戏设计与实现

3.1. 教育游戏《拯救米迪亚》的设计思路

3.1.1. 教育游戏《拯救米迪亚》的系统架构设计

依据本研究的问题，学习者在游戏中需要学习规范的科学探究过程。具体的学习目标是能够知道规范的科学探究过程要通过六个环节展开，理解每个探究环节的具体含义，可以按照正确的顺序排列科学探究的六个环节，最终能够迁移到实际探究学习活动中进行应用。因此，在建构主义学习理论、需要层次理论、游戏化学习理论、情境学习理论和沉浸理论的指导下，本研究将教育游戏设计为角色扮演类小游戏，选择了 RMMV 平台作为游戏开发环境。为了激励小学生参与游戏的兴趣和动机，增强游戏的趣味性、体验感、挑战性和成就感，将游戏情境设计为营救类冒险闯关游戏《拯救米迪亚》，游戏者需要在游戏中通过引导和调查经历六个冒险闯关过程，集齐六把钥匙并通过最后的排序测验，最终拿到宝剑战胜魔王，拯救米迪亚城主并解放米迪亚城。游戏运行环境为 Android 系统。基于角色扮演游戏的学习模式系统架构设计如图 1 所示。

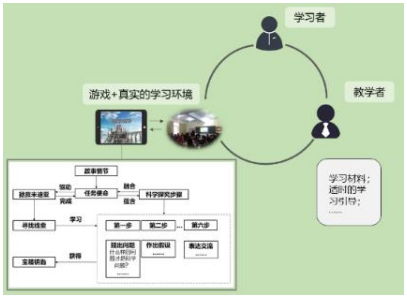


图 1 学习模式系统架构图

3.1.2. 教育游戏《拯救米迪亚》的策略

为了达到教育游戏的娱乐和教育性的相对平衡，笔者对本次教育游戏的设计提出以下策略。

动机的激发与维持

在教育游戏中由于教学目的的存在，玩家容易出现枯燥的感觉。针对此问题，研究提出以下三个策略。第一，合理的挑战难度。沉浸理论解释了在高挑战和高技能平衡的状态下，才能产生心流体验。因此，游戏的设计一定要具有合理的挑战难度，这样才能够让玩家不仅能够玩下去还能玩进去，也就是我们所说的沉浸感，如此才能达到我们所希望的在游戏中学习且快乐学习的状态。第二，及时的反馈激励机制。比如说玩家通过小关卡后，能够获得一定的奖励。再如，玩家寻找线索进行解谜时，超过一定次数或时长就要提供适当的帮助与说明。及时的反馈机制以及奖励机制能够一定程度的增强游戏的粘连性。第三，任务驱动策略。这与在教学中的“任务驱动法”的教学原理相似，运用在教育游戏中来讲，就是玩家需要通过获得的线索提示或者与 NPC（非玩家角色：Non-Player Character）的对话，获取一个接一个的任务，推动玩家寻找完成任务的办法。在教育游戏中，此任务往往是依附于故事情节的教学内容，以此来推进游戏的进程。

用户对游戏的适应性

适应性即游戏从元素的设计到触发机制的设计都要与玩家的认知相适应。本次开发的教育游戏的适应对象为高段小学生（五、六年级的学生），在设计时要考虑到学生的认知及偏好。例如，在界面设计方面，儿童往往更喜欢色彩艳丽或者对比度强的色调，画风倾向于卡通风格更容易受到儿童的喜爱，故事情节更偏于童话风，有明确的目的会让儿童更容易有使命感，更喜欢和怪兽战斗并且打败它从而获得成就感等。

游戏设计的合理性

游戏设计的合理性包括以下两个方面：第一，思想情感积极。除了要达到教学的目的，整个游戏带入的故事情节在思想上要带给孩子们正面的教育，如强调正义必胜等正能量主题，以寓教于乐的方式塑造孩子正向世界观。第二，操作设计简洁直观。避免过于复杂的动作要求，界面清晰突出重点，确保交互方式易于理解与执行，符合大众普遍认知习惯。

3.1.3. 教育游戏《拯救米迪亚》开发的整体框架设计

教育游戏开发的整体框架设计如图 2 所示。

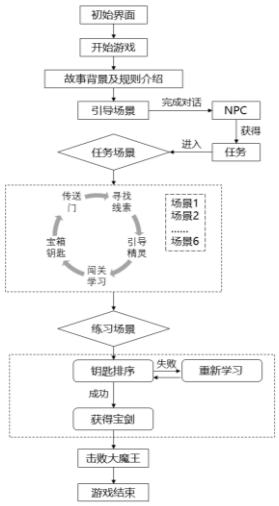


图 2 整体框架设计

3.2. 《拯救米迪亚》教育游戏的功能实现

3.2.1. 学习机制的功能实现

教学任务的设计与融入是本次教育游戏设计与开发的着重点，本游戏中融入的教育任务是依据科学探究的六个步骤而设定的。总体结构上分为两大部分，包括新知学习与测试，其中测试包括训练和最终测验。具体的教学任务框架设计如图 3 所示。

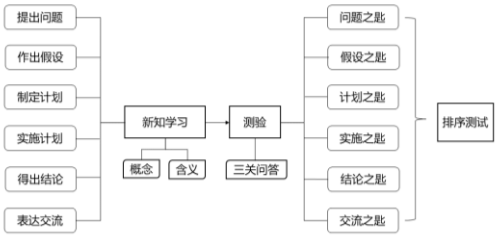


图 3 教学任务框架设计

在新知学习与训练部分，玩家通过调查和线索提示，找到关键 NPC，进入虚幻境地，也就是进行知识的学习与训练的场所，如图 4 所示和图 5 所示，以第一大关“提出问题”为例。在这里，玩家首先需要找到引导精灵（NPC）进行新知的学习。新知的学习分为两大部分，一部分为此步骤的排位顺序以及新知讲解，另一部分为实例呈现巩固知识。每一部分都会有结果的询问供学习者自主选择，以便学习者在这一环节里可以反复观看学习。



图 4 知识的学习



图 5 举例讲解巩固

待玩家认为自己已经学会之后，可以进入闯关环节。此环节设置三个小关卡，每答对一题便可进入下一小关，待三关全部闯过，此步骤的学习内容才算完成，玩家便可获得此步骤对应的相关钥匙。如图 6 所示为问答闯关部分截图和闯关成功获得钥匙截图。



图 6 闯关及闯关成功

每完成一关收集到相应的钥匙后，学习者会通过“传送门”进入下一关卡的学习情境。与科学探究六个步骤相对应，学习者需要经历六大关卡进行知识的学习和钥匙的收集。闯关完成后，学习者为了获得战胜大魔王的力量，必须进入试练关卡，如图 7 所示。在这里，学习者只需要将背包里的钥匙按正确的顺序放置在序号对应的台子上便可算成功通过测试。



图 7 拼钥匙测试

在整体的学习机制设定中，为了减缓学习过程的机械化，并且使其更加有趣且富有挑战性，学习者每到一个关卡，其情境、线索的调查以及关键人物寻找方式都不会与上一关卡相同，且关卡的情境以及遇到的各种 NPC 紧贴故事发展的主线，学习者在游戏中，由调查巫师的住所开启冒险之旅，到城堡的搜查再到圣坛的探寻，最终抵达魔王城，在魔王城中与怪兽比试，获得伙伴，搜到最后两把钥匙，试练功获得打败大魔王的力量。

3.2.2. 引导机制的功能实现

在游戏中，引导机制的存在是非常重要的，学习者可以通过引导来了解相应的游戏规则，也能够通过引导了解探索的方向。在本游戏中，引导可以分为两大类，一类是文字图形类的引导，一类是 NPC 的引导。

文字图形类的引导也可以称作提示，主要是为了使学习者明确总体任务，了解游戏规则以及获得线索提示。如图 8 所示，学习者进入游戏后会看到故事式的讲述，在简短的讲述中可以获得重要的任务信息，即“集齐六把钥匙才能获得击败魔王的力量”。再如图 9 所示，学习者在每个线索调查场景中都会看到不断闪耀的星星，这些便是提示学习者线索的所在处，且每找到一个线索，也会有下一个线索的指引。只有通过线索的调查和寻找，学习者才能被一步步引导着走向钥匙的所在处。



图 8 进入游戏的引导界面



图 9 城堡中的调查场景

NPC 引导的主要作用是对学习者学习任务获得的引导以及故事发展的推进，如图 10 所示，每一关卡都会有相应的引导精灵和学习者发生互动对话，推进学习者的学习过程。再如图 11 所示，学习者与 NPC 发生交谈，推进整体故事情节的发展。





图 10 引导精灵的引导场景 1　图 11 引导精灵的引导场景 2

3.2.3. 反馈机制的功能实现

反馈机制在游戏中是至关重要的，学习者只有通过及时的反馈才能获得正确的信息，激励的反馈机制也能增强学习者的成就感以及获得感，使得学习者在学习过程中能够有信心并且有兴趣挑战下去。如图 12，学习者在拼钥匙环节失败时，会出现错误反馈，并进入复习界面。学习者通过复习后可以再次进入挑战直到正确，最终会获得制胜宝剑，如图 13 所示。



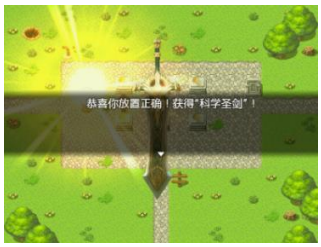


图 12 拼钥匙失败的反馈界面　图 13 拼钥匙成功的反馈界面

此外，学习者每成功闯下一关，除了会获得重要物品“钥匙”之外，还会获得能力的大幅度提升，等级的提升代表着能力值的提高，这对后边挑战 BOSS 会起到重要的作用。

4. 游戏的教学应用效果检验

4.1. 《拯救米迪亚》的教学应用案例

为了检验游戏开发的实际效果，研究将小学科学六年级下册的最后一个单元设计了以“环境”为主题的 STEAM 学习项目，项目由指导专家任课教师和笔者共同设计，包括《为垃圾找归处》和《给污水做个清洁》两个项目，每个项目均采用项目式学习（Project-based learning）教学模式，由六个教学环节组成，即项目引入、知识铺垫、方案制定、实施优化、评价总结和知识拓展。本游戏主要应用于“知识铺垫”学生自主探究学习过程中。

4.2. 《拯救米迪亚》的教学应用效果

研究随机抽取某小学 6 年级的一个班级为实验组（共 50 人，男生 32 人，女生 18 人），进行为期三周共 6 个课时的单组教学实验。参照张屹教授、赵亚萍学者等人于 2017 年编制的《STEM 教学成效调查问卷》(张屹, 赵亚萍, 何玲, & 白清玉, 2017)进行了改编，设计了小学生科学探究能力问卷，包括 15 道题，该问卷采用李克特式 5 级量表计分形式。

对实验组科学探究能力调查问卷前后测数据的配对样本 t 检验分析，实验组的前后测平均数分别为 58.22 和 62.26，且 Sig.=0.006，即 $p<0.05$，即实验组的前后测数据有显著性差异，且后测平均数明显的高于前测。这表明在 STEAM 项目化学习应用教育游戏能够显著提升学生的科学探究能力，本研究中教育游戏的开发达到了预期的效果。

5. 结论

结合小学生心理与教育特征，在 STEAM 项目化学习中加入了教育游戏，游戏自身的趣味性、挑战性、交互性等特点自然也融入到学习过程中，这就为学习者在学习情境中带来了更强的沉浸感以及更好的体验感。在游戏学习中获得的成就感、满足感等积极情绪会伴随着学

习活动的发生而产生，进而使学习者对此种学习模式能够获得较高的满意度。本研究中的教育游戏主要作用于“知识探究”环节，提升学习者在这一环节中的探究学习效果。

总之，教育游戏能够为 STEAM 项目化学习中培养小学生科学探究能力提供一条新路径，能够提高学生的学习动机，增强学生的学习兴趣，能够促使学生在课堂中达到心流状态，提高学生的专注度。此外，本研究也存在一定的不足之处，由于受到时间及实践条件的限制，对于项目的设计和开发尚不完整，在今后需要进一步解决。

参考文献

于兰兰, 谭博学, 张厚升, & 苗飞. (2016). 趣味化学习背景下“闯关游戏式”教学方法探究 %J 中国成人教育. (02), 112-114.

刘东方, & 王磊. (2012). 科学探究能力的构成要素——基于国外科学课程文件的分析 %J 化学教育. *33*(09), 44-49.

张屹, 赵亚萍, 何玲, & 白清玉. (2017). 基于 STEM 的跨学科教学设计与实践 %J 现代远程教育研究. (06), 75-84.

张金磊, & 张宝辉. (2013). 游戏化学习理念在翻转课堂教学中的应用研究 %J 远程教育杂志. *31*(01), 73-78. doi:10.15881/j.cnki.cn33-1304/g4.2013.01.007

李洋, 豆丹阳, & 王健. (2019). Kahoot!游戏化学习平台在生物学教学中的应用 %J 生物学教学. *44*(06), 41-42.

李春密, 梁., 蔡美洁. (2004). 中学生科学探究能力结构模型初探 %J 课程.教材.教法. (06), 86-90. doi:10.19877/j.cnki.kcjcjf.2004.06.021

李梅. (2017). 认知视角下的项目化学习解析 %J 电化教育研究. *38*(11), 102-107. doi:10.13811/j.cnki.eer.2017.11.016

余胜泉, & 胡翔. (2015). STEM 教育理念与跨学科整合模式 %J 开放教育研究. *21*(04), 13-22. doi:10.13966/j.cnki.kfjyyj.2015.04.002

Chen, M.-H. M., Tsai, S.-T., & Chang, C.-C. (2019). Effects of Game-Based Instruction on the Results of Primary School Children Taking a Natural Science Course. *Education Sciences, 9*(2). doi:10.3390/educsci9020079

Esther del Moral Perez, M., Guzman Duque, A. P., & Carlota Fernandez Garcia, L. (2018). Game-Based Learning: Increasing the Logical-Mathematical, Naturalistic, and Linguistic Learning Levels of Primary School Students. *Journal of New Approaches in Educational Research, 7*(1), 31-39. doi:10.7821/naer.2018.1.248

Fu, Q.-K., Lin, C.-J., Hwang, G.-J., & Zhang, L. (2019). Impacts of a mind mapping-based contextual gaming approach on EFL students' writing performance, learning perceptions and generative uses in an English course. *Computers & Education, 137*, 59-77. doi:10.1016/j.compedu.2019.04.005

Hsu, Y.-S., Fang, S.-C., Zhang, W.-X., Wu, H.-K., Wu, P.-H., & Hwang, F.-K. (2016). Identifying Effective Design Features of Technology-Infused Inquiry Learning Modules: A Two-Year Study of Students' Inquiry Abilities. *Educational Technology & Society, 19*(2), 228-244.

Wu, C.-H., Tzeng, Y.-L., & Huang, Y.-M. (2020). Measuring performance in leaning process of digital game-based learning and static E-learning. *Etr&D-Educational Technology Research and Development, 68*(5), 2215-2237. doi:10.1007/s11423-020-09765-6

数学坐标游戏的设计与开发研究

Design and development of a digital math coordinate game

相梦彤[1]，张露[2]，刘宇[3]，尚俊杰[4*]
[1, 4]北京大学学习科学实验室
[2]北京邮电大学教育技术研究所
[3]北京市第十九中学

【摘要】 平面坐标是数学学习的重要概念，但传统坐标教学面临诸多困境，教育游戏可作为有力工具支持学生的坐标学习。本研究在充分调研坐标学习和坐标教育游戏相关研究、课程标准、教材内容的基础上，尝试设计开发出一款支持坐标学习的游戏《坐标冲冲冲》,从教学内容设计和游戏机制设计两个方面对游戏设计展开探讨，以期创新坐标学习方式，增强学生坐标学习表现。

【关键词】 坐标学习；教育游戏；设计与开发；数学学习

Abstract: *Coordinate is a crucial concept in mathematics, but traditional coordinate teaching faces various challenges. Educational games can serve as powerful tools to support coordinate learning. In this study, we attempted to design and develop a digital game 'Go Coordinate' based on Chinese curriculum standards, textbook content and relevant studies on coordinate learning. Instructional design and game design are discussed in detail. The goal is to innovate the way coordinates are taught and enhance students' performance in coordinate learning.*

Keywords: coordinate learning, educational games, design and development, math learning

1. 引言

平面坐标是数学学习的重要概念，作为几何和代数的桥梁，是数形结合的有力工具。义务教育数学课程标准(2022)将“图形与坐标”作为初中阶段图形与几何领域三大主题之一，要求让学生掌握数形结合的数学思想，学会用代数方法研究图形。但是传统课堂中平面坐标系的学习面临诸多困难。教师只能通过黑板或课件的方式对平面坐标系的内容进行讲解和演示，学生通过听讲和练习习得知识，缺乏与坐标概念的直接交互和亲身体验；去情境化教学更使学生停留在对课本定义的机械记忆，无法做到问题解决中灵活的迁移应用(Saha et al., 2010)。

在教育数字化转型的今天，信息技术与教育教学的融合日益深入。教育游戏为解决学习和教学问题提供了有力的技术工具，不少研究和实践证明了教育游戏在激发学习动机(Yu et al., 2021)、理解和掌握知识、提升技能等方面的重要作用(Zeng et al., 2020)。

尽管如此，目前却很少有以坐标学习为主题的游戏开发研究与实践，尚未形成系统完整且科学的坐标游戏设计框架;且现有的坐标学习平台,多面向高等教育中大学生线性代数、STEM等课程学习，发表在非教育类期刊中，无法满足基础教育阶段初中生的学习需求。在游戏的趣味性和情境性上难以激发初中生的学习动机，在科学性上缺乏对学生认知规律和教育教学理论的探讨，无法贴合我国的课程标准和教学真实需求，更缺乏科学的学习理论指导。

因此，我们基于过往中小学数学教育游戏开发经验，结合我国义务教育课程标准和教材，尝试开发出《坐标冲冲冲》平面坐标学习游戏，帮助学生更好地学习和理解平面坐标知识，以坐标为切入点，打牢“数”与“形”结合思想的基础，促进其今后的理解迁移。

具体而言，本研究研究问题如下：

（1）如何设计一款教育游戏促进学生的坐标学习？

（2）数字教育游戏能否促进学生的坐标学习？

2. 文献综述

2.1 坐标学习与教学现状

坐标系是一个数学概念，有着数形结合双重内涵，用于确定几何对象在二维空间（平面）的精确位置。平面直角坐标系是连接数与形的桥梁，体现了代数和几何问题之间的相互转化，也是后续函数等知识学习的基础。但学生对坐标系的理解面临诸多困难。比如 Palupi(2013)指出学生无法理解坐标系具有唯一原点，构建和定位坐标。学生机械记忆规则，忽视对数学知识的理解。Sarama et al. (2003)则进一步指出，学生除了忽略原点外，还倾向于颠倒 x 轴和 y 轴。Akdeniz(2022)则报告称，学生对零的理解存在重大误解。引入负数后，概念学习更为困难。此外，Fuadiah & Suryadi(2017)发现，使用技术工具可以帮助学习者认识与坐标系相关的概念。技术环境对学生来说是一个有意义的环境，可以直观展示概念及其联系。

除了坐标学习本身的困难外，坐标教学也存在诸多障碍。目前我国课堂教学往往是以教师和课本为中心的静态教学观，教师把课本知识当作无需检验、只需理解和记忆的“绝对真理”教授给学生，把教材当作是教学顺序的“最优解”安排教学过程，缺乏知识再组织和再创造，缺乏对坐标教学形态的研究和开发，没有使坐标的学习走向“思维中的具体”。这种标准性和唯一性的教学观也轻视学生已有认知能力和知识经验，所以坐标教学更应该具有思维发展的层次性以及阶段性，让学生在自主的问题解决中体验探究的过程，建立坐标概念的感性和经验基础，并不断深化对坐标概念的理解(Kamaludin et al., 2015)。

2.2 教育游戏支持的坐标学习

教育游戏为解决坐标的学习和教学问题提供了有力的技术工具。游戏的视觉直观能够实现坐标“数”和“形”双重含义的不同表征，游戏中对概念对象的具体操作也能降低坐标概念理解的抽象性，游戏创设的真实情境有助于学生实现问题解决的灵活迁移。此外，一对一的游戏终端让每名学生都有体验参与的机会，有效地解决了传统一对多课堂教师无法兼顾每位学生，无法做到个性化自定步调学习的问题。

文献调研发现现有研究关于坐标学习的教育游戏较少。坐标学习上主要有 2D-CACSET(Carrasco-Navarro et al., 2022)、Electric Linear Actuator System(Konishi et al., 2002)、GeoGebra(Saha et al., 2010)等面向大学生线性代数学习的可视化技术平台。2D-CACSET（图 1 左）是一个实体的教育工具包，学生可以在坐标平面内自由移动游戏物体，x+、y+、x-、y-四个方向都配有传感器，能在电脑上实时地捕捉显示物体位置。GeoGebra（图 1 右）是一个面向五年级坐标学习的应用程序，分为三个级别，在坐标系中围成几何图形。实验发现，GeoGebra 能够帮助提升学生的学习兴趣，促进学生对于坐标知识的掌握。

但总得来说，目前关注坐标学习的教育游戏研究较少，几款教育游戏多是面向本科生，在国内，坐标学习是初中数学的重要内容，却少有适合初中学生教学的坐标教育游戏。同时，坐标教学更缺乏科学学习理论指导下，针对主要的学习困难点，展开有效的游戏机制设计。

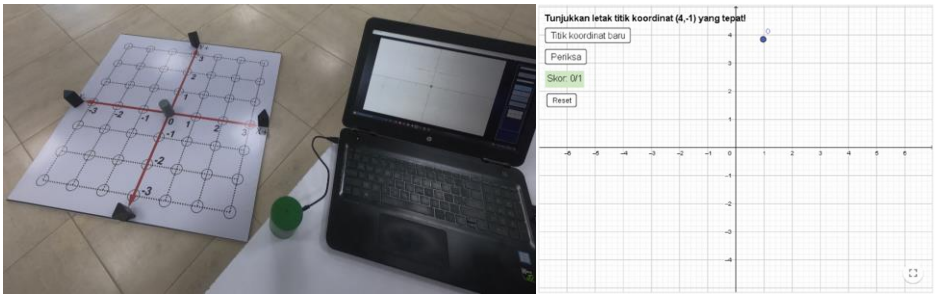


图 1：2D-CACSET（左）和 GeoGebra（右）

3. 坐标冲冲冲

3.1 教学内容设计

3.1.1 教学目标

课程标准中，“图形与坐标”属于初中阶段图形与几何领域三大主题之一，要求学生掌握平面上的点与数对表示的坐标一一对应，引导学生学会使用坐标表达图形的轴对称、旋转、平移变化过程，能够借助平面直角坐标解决现实问题，感悟数形结合的意义，发展推理能力和运算能力，增强应用意识。因此，初步把本游戏的教学目标定为【平面坐标系的学习】。

（1）理解平面上点与坐标之间的一一对应关系

（2）掌握用坐标表达图形的轴对称、平移等变化过程

（3）感悟数形结合的思想，会用数形结合的方法分析和解决问题

3.1.2 教学内容

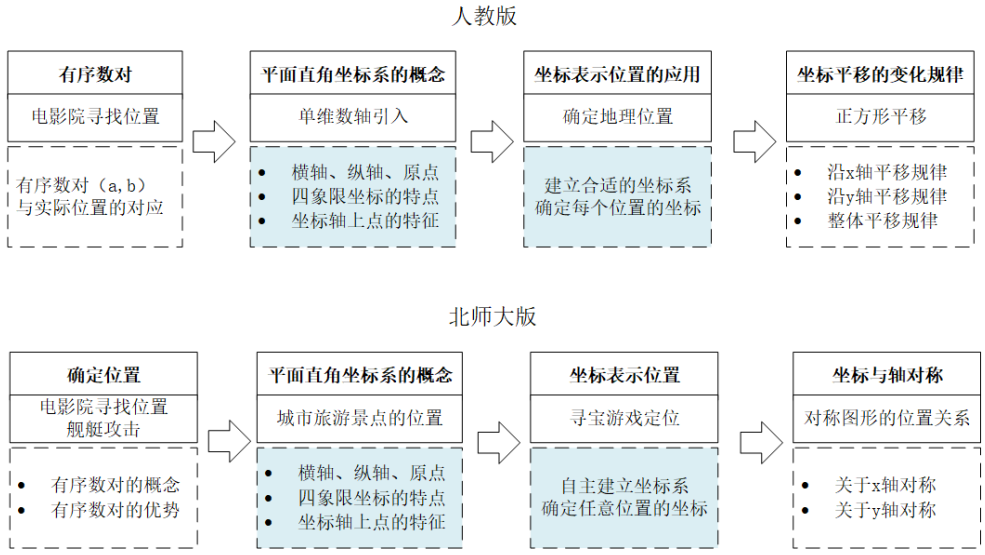


图 2：教学内容序列

教学内容上对人教版和北师大版初中教材“图形与坐标”部分展开分析，具体对比如图 2。

基于此，将本游戏的教学内容序列设置为“坐标确定点的位置-点与坐标轴的位置关系及特征-坐标的平移”。（1）坐标确定点的位置的学习目标为要求学生能根据坐标确定出点的位置，由点的位置写出坐标，实现平面中点与有序数对的一一对应。（2）点与坐标轴的位置关系学习目标为理解位于坐标轴上、与坐标轴平行、与坐标轴轴对称点的位置关系和特征。（3）坐标的平移学习目标为掌握坐标系中平移的变化规律。

3.1.3 知识网络

结合教材知识点和习题设计，进一步将三大知识模块细化为单个知识点，知识网络如图 3，比如第一个模块“坐标确定位置”包含“原点已知”和“原点未知”两种情况，原点已知又分为第一象限到第四象限坐标位置的特点。教材中原点已知主要以“写出已知点的坐标”或“根据坐标确定平面上的点”的题目出现，实现平面上点与数值坐标的一一对应；原点未知与实际问题情境结合，让学生自己建立坐标系，确定某一物体的位置。后两个模块同理。

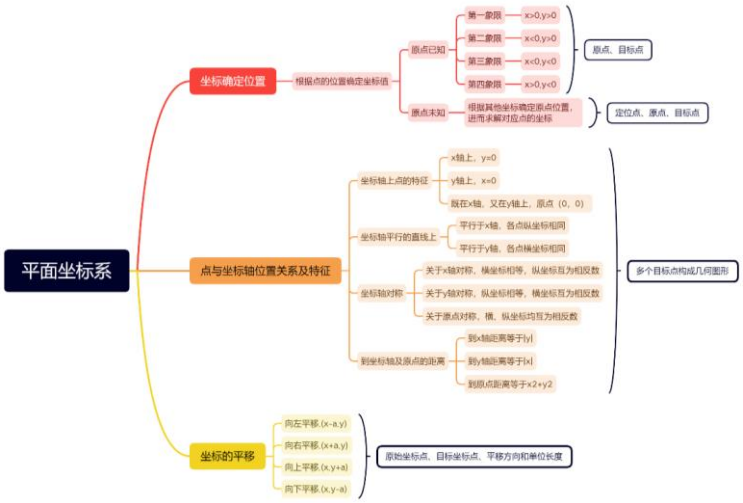


图 3：知识网络图

3.1.4 任务关卡

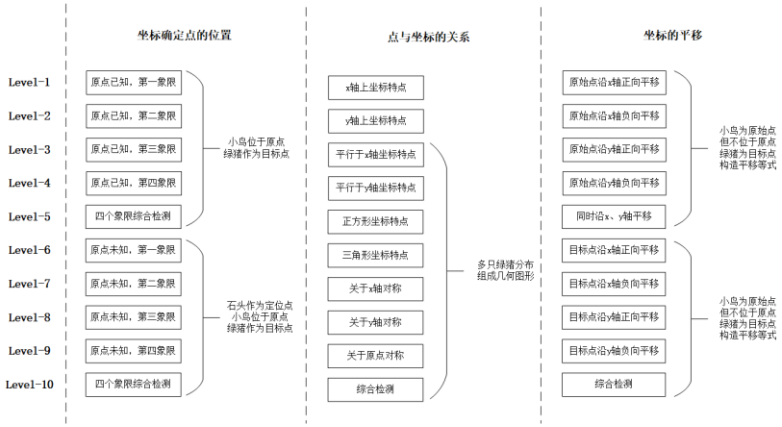


图 4：任务关卡序列

如图 4，本游戏共分三个模块，分别对应坐标确定点的位置、点与坐标的关系、坐标平移三大知识模块。每一模块下又分为 10 个小的关卡，对应模块下具体知识点的学习。比如关卡 1“坐标确定点的位置”的知识掌握包含原点已知和原点未知两种情况，原点已知又分为第一到第四象限点的位置特点。所以 Level-1 的教学内容为帮助学生掌握在原点已知情况下，确定位于第一象限点的坐标值。

3.1.5 核心辅助认知工具

本游戏的核心认知辅助工具有两个：

（1）平面直角坐标系，相当于二维平面数轴，是“数”和“形”结合的有力工具（坐标平面内的点与有序实数对一一对应，数轴上的点与实数一一对应）；每一个游戏物体（比如水果）可以在坐标系上找到一个对应位置点，每一个游戏物体具有图形（点在平面中的位置）与数值（点的坐标）两种表征方式。这样的设计可以更直观地展示坐标数与形的关系，帮助学生理解平面坐标系的实际意义。

（2）弓箭的方向：箭头指向方向与坐标的象限方向一致，可以帮助学生更直观地感受到四个象限坐标的特点。

3.2 游戏机制设计

3.2.1 游戏背景设计

游戏以小猴子的水果冒险之旅为背景（图 5），小猴子吉吉希望品尝到分布在平面坐标背景中的各种美味水果。为了帮助吉吉实现它的愿望，玩家需要运用学习到的坐标知识，选择合适的坐标弓箭，发射弓箭以精准地击中对应水果。



图 5：游戏背景设计

3.2.2 关卡地图设计

游戏设置了 3 个大的地图，每一地图包含从易到难 10 个关卡。为了便于教学，游戏设置闯关版和解锁版两个版本，闯关版需要学生逐步解锁，解锁版则在一开始开放全部关卡。



图 6：地图选择（左）；闯关版（中）；解锁版（右）

3.2.3 游戏玩法设计

游戏玩法类似射击游戏，玩家首先点击鼠标选择合适的弓箭；接着拖动鼠标，拉动弹弓，调整攻击方向，帮助小猴子顺利击中水果。游戏主界面是一个二维数轴的背景图片，每一关弹弓和水果的坐标点固定，水果会随机出现在平面的某一坐标点，玩家需要选择合适的弓箭，调整攻击的方向，使弓箭击中水果。

由于教学内容不同，三个地图的玩法也存在差异。地图 1 坐标确定位置，弹弓位于原点，玩家需要确定水果的坐标数值，选择正确的弓箭射中水果即为获胜；地图 2 坐标的位置关系，弹弓为原点，水果的分布体现坐标轴中点的特点（如平行与 x 轴、轴对称），弓箭击中多个水果，即可获得胜利；关卡 3 为坐标的平移，弹弓不位于原点，弓箭坐标等于水果坐标减去弹弓坐标。



图 7：游戏玩法设计

3.2.4 奖惩机制设计

在奖惩机制设计上，每个关卡 5 道题目，答对 3 道（正答率 60%）即可通关，并获得与之相应的能量星，在输赢界面显示，如果玩家对当前关卡获得的能量星数目不满意，可以重新尝试取得更高的星星奖励。同一关卡内每道题目共有两次试错机会（共 3 个选项，排除无效试错的可能），每错一次扣一颗能量星，游戏界面左上方会显示能量星的个数（共 5 颗），能量星耗完则游戏失败。

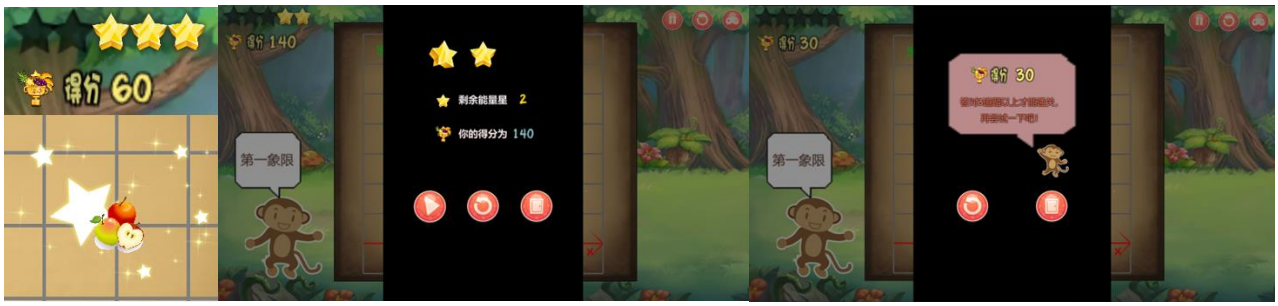


图 8：游戏奖惩机制设计

4. 实验结果

为检验游戏的学习效果，我们将《坐标冲冲冲》应用于课堂中，于北京市某中学开展实验研究。随机选取两个水平相当的班级，实验组 34 名学生（男生 14 人，女生 20 人），对照组 28 名学生（男生 11 人，女生 17 人）。实验前和实验后分别进行坐标知识的测试，测试参考课本习题编制，在授课教师的指导下修改。

项目	前测			后测			D-value	t	df	p	d
知识	N	Mean	SD	N	Mean	SD					
测试	34	13.26	54.37	34	7.79	44.23	-5.47	3.56	33	.000	1.69

表 1：前后测配对样本 T 检验

项目	实验组			对照组			D-value	t	df	p	d
知识	N	Mean	SD	N	Mean	SD					
测试	34	21.96	7.37	28	13.24	8.32	-8.72	-4.33	55	.000	1.67

表 2：实验组和对照组独立样本 T 检验

本研究采用配对样本 T 检验对学生前后测分数进行分析，探究游戏的教学应用效果。结果见表 1，发现学生知识后测总分均值（M=7.79，SD=44.23）显著高于前测（M=13.26，SD=54.37），t（34）=3.56，p=0.000，d=1.69。同时，对实验组和对照组进行独立样本 T 检验（表 2）发现，实验组的知识测试分数（M=21.96，SD=7.37）要显著地高于对照组（M=13.24，SD=8.32）。说明教育游戏《坐标冲冲冲》的应用能够促进学生的数学坐标知识学习。

5. 结语

本研究设计开发出一款支持坐标学习的数字教育游戏，从教学内容和游戏机制两个方面对教育游戏设计进行考虑。基于我国课程标准制定教学目标，综合人教版与北师大版教材确定教学内容，细化每一教学内容下的知识网络和与之相应的任务关卡序列，并选取平面坐标系和弓箭方向作为核心辅助认知工具，保证游戏设计贴合我国真实教学需求和学生认知规律。

在游戏机制上，以连贯有趣的故事情节为背景，设置三个地图，增强游戏代入感；不同地图由于教学内容差异在玩法上存在区别，针对易错点进行题目设计，并提供两种类型的脚手架作为学习支持。

教育游戏具有教育性和游戏性双重属性，其设计要考虑二者的平衡。游戏的教育性体现在以课程标准为指导，以内容知识为基础，以解决学生认知困难为目标，将游戏作为认知工具，符合科学认知规律，促进学生的概念理解。游戏性则需要考虑背景叙事、关卡地图、玩法、奖惩机制等元素，增添游戏的趣味性，让学生发现知识本身的乐趣，激发其内部动机。教育和趣味的交互展开，既能有效传递教育信息，又能激发学生的学习兴趣。这样的设计将为学生提供一种令人着迷且有益的学习体验，为教育领域注入新的活力和可能性。

本研究也存在一定局限：游戏设计为一次开发成果，主要是从文献和教材出发自上而下的设计思路，缺乏实践考量，也缺乏基于真实教学实践的多轮迭代和修改；其次，游戏的教学效果并未得到实证研究检验，其有效性有待考量。未来可以通过严谨的实验设计检验游戏教学效果，通过迭代优化实现螺旋上升的游戏设计，并落实到一线教育教学实践，助力学生的数学坐标学习。

致谢

本研究感谢北京市教育科学“十四五”规划 2021 年度延续课题“基于学习科学和游戏化学习的学习能力培养研究”（编号: CEFA21067）；田家炳基金会北大港大学习科学青年学人交流项目的支持！感谢北京市第十九中学王肖华老师、北京大学学习科学实验室李卓老师等提供的指导与帮助！

参考文献

Akdeniz, D. G., Topbaş, E. S. Y., & Argün, Z. (2022). Conceptions of Zero Among Students With and Without Learning Disabilities. *International Journal for Mathematics Teaching and Learning*, 23(2), 1-21.

Carrasco-Navarro, R., Luque-Vega, L. F., Nava-Pintor, J. A., Guerrero-Osuna, H. A., Carlos-Mancilla, M. A., & Castañeda-Miranda, C. L. (2022). MEIoT 2D-CACSET: IoT Two-Dimensional Cartesian Coordinate System Educational Toolkit Align with Educational Mechatronics Framework. *Sensors*, 22(13), 4802.

Fuadiah, N. F., & Suryadi, D. (2017). Some Difficulties in Understanding Negative Numbers Faced by Students: A Qualitative Study Applied at Secondary Schools in Indonesia. *International Education Studies*, 10(1), 24-38.

Kamaludin, A., Rusmin, P. H., & Harsoyo, A. (2015, 2015-1-1). *Design and implementation educational game of coordinate systems and least common multiple using educational games design model*. Paper presented at the.

Konishi, S., Ohno, K., & Munechika, M. (2002). Parallel linear actuator system with high accuracy and large stroke. *Sensors and Actuators a: Physical*, 97, 610-619.

Palupi, E. (2013). Understanding the Coordinate Systems. *Sriwijaya University*

Saha, R. A., Ayub, A. F. M., & Tarmizi, R. A. (2010). The effects of GeoGebra on mathematics achievement: enlightening coordinate geometry learning. *Procedia-Social and Behavioral Sciences*, 8, 686-693.

Sarama, J., Clements, D. H., Swaminathan, S., McMillen, S., & Gonzalez Gomez, R. M. (2003). Development of mathematical concepts of two-dimensional space in grid environments: An exploratory study. *Cognition and Instruction*, 21(3), 285-324.

Yu, Z., Gao, M., & Wang, L. (2021). The effect of educational games on learning outcomes, student motivation, engagement and satisfaction. *Journal of Educational Computing Research*, 59(3), 522-546.

Zeng, J., Parks, S., & Shang, J. (2020). To learn scientifically, effectively, and enjoyably: A review of educational games. *Human Behavior and Emerging Technologies*, 2(2), 186-195.

中华人民共和国教育部. (2022). *教育部关于印发义务教育课程方案和课程标准（2022年版）的通知* Retreved 2024-1-22 from http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html

ARPG 與 Roguelite 遊戲對低學習成就國小學生的數學學習成效與學習投入之影響

Effects of ARPG and Rouguelite Games on Lower-Achievement Elementary School Students' Learning Performance and Engagement in a Mathematics Course

謝榮益[1]，游師柔[2]，孫之元[3*]
[1, 2, 3] 國立陽明交通大學
[*] jerrysun@nycu.edu.tw

【摘要】根據 2022 年 PISA 調查報告指出台灣地區數學高低學習成就者之間的數學分數落差逐漸加大，本研究目標為探討創新的學習策略協助低學習成就學生克服數學學習困境。本研究透過文獻分析法回顧遊戲式學習應用於數學教育、ARPG 和 Roguelite 之教育發展現況與兩者之間差異，並探討 ARPG 和 Roguelite 對低學習成就學生的學習成效與學習投入之影響。研究結果為 Roguelite 有機會成為有效提升學習成效和學習投入之教學輔助策略，並且 ARPG 能促使學習成效與學習投入達到平均水準之上。本研究建議教學者可採用 Roguelite 遊戲穩固數學基礎能力，提升低學習成就者的學習成效與學習投入。

【关键词】 動作角色扮演遊戲；Roguelite；學習成效；學習投入；數學教育

Abstract: *In PISA 2022, the gap in mathematics scores increased between higher- and lower-achievement students in Taiwan. This study aims to examine innovative learning strategies to help lower-achievement students overcome the predicament of mathematics learning. The study employed a document analysis method to review the application of game-based learning for mathematics education and the differences between ARPG and Roguelite games. It also examines the effects of ARPG and Roguelite games on lower-achievement elementary school students' learning performance and engagement. The results indicated that the Roguelite games can potentially enhance learning achievement and engagement. The ARPG can raise learning achievement and engagement to above-average levels. This study suggested that instructors use the Roguelite games to improve lower-achievement students' mathematics learning achievement and engagement.*

Keywords: Action role-playing game, Roguelite, Learning performance, Engagement, Mathematics education

1. 前言

STEM 分別指科學(science)、科技(technology)、工程(engineering)與數學(mathematics)，發展 STEM 領域為提升人才全球競爭力的關鍵（Kayan-Fadlelmula et al., 2022），根據 OECD（2023）國際學生能力評估計劃（Programme for International Student Assessment, PISA）調查報告指出，台灣地區在 2022 年數學平均成績相較於 2018 年有所提升，然而高學習成就學生（前 10%）和低學習成就學生（後 10%）之間的成績差距卻擴大。因此，如何減緩高低成就之間的數學學習落差成為重要議題。Çakir 和 Ergül（2022）指出低數學成績學生的工作記憶表現顯著低於成績中等或高的同齡學生。根據認知負荷理論，教學設計可以調整工作記憶的認知處理需求（Sweller, Ayres, & Kalyuga, 2011），藉此提高學生的學習表現。Bodovski 和

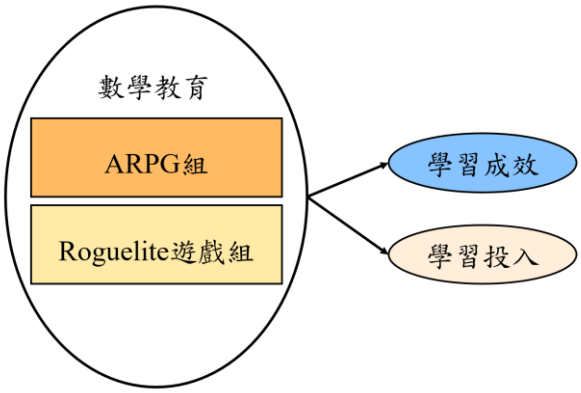
Farkas（2007）發現，雖然在校學習成就較低的學生所接受之教學時間最多，但投入教學的程度最低。因此，改善學習投入為促進低學習成就者之學習成效的方式之一。

藉由趣味遊戲流程學習知識是有助於提高學習成效和學習動機的因素之一（Weng et al., 2024），但是 Streicher 和 Smeddinck（2016）指出當活動過於冗贅、單調會讓使用者感到厭倦。反覆練習的任務關卡之遊戲機制可能會造成學習倦怠，Roguelite 加入隨機的遊戲機制，將有機會改善學習者對於遊戲重覆練習所產生的無趣學習經驗。本研究透過文獻回顧的方式，評估動作角色扮演遊戲（Action role-playing game, ARPG）和 Roguelite 的遊戲機制對數學低學習成就者的學習成效與學習動機之效益。因此，本研究目的為探討低數學學習成就國小學生，使用 ARPG 和 Roguelite 對公因倍數數學單元的學習成效與學習投入之影響，圖 1 為研究架構圖，研究問題如下：

研究問題一：低數學學習成就國小學生使用 ARPG 和 Roguelite 對學習成效是否具顯著差異？

研究問題二：低數學學習成就國小學生使用 ARPG 和 Roguelite 對學習投入是否具顯著差異？

圖 1 研究架構圖



2. 文獻探討

2.1. 遊戲式學習應用於數學教育

數學興趣與數學成績呈現顯著正相關（Brumm & Rathgeb-Schnierer, 2023），遊戲式學習指利用電子遊戲的激勵和趣味性建立教育或學習環境（Linek et al., 2009），並且遊戲本身相對於紙筆學習或是傳統面對面教學更具激勵性和吸引力（Girard, Ecalle, & Magnan, 2013）。遊戲的性質能夠促使學生對於學習的意願更加濃厚，利用遊戲機制帶領學生感受數學學習的樂趣，藉此提高學生的學習成效和學習投入。數學成績與學生對於數學之學習態度也具有顯著正相關，獲得好的成績能夠使學生以積極的態度面對數學，進而有強烈的動力學習數學（Evans & Field, 2020），形成正向循環。遊戲式學習有望減少學習者對數學的抗拒情緒，讓學習者主動投入學習過程，不同的遊戲設計類型帶來的學習成效與學習動機也不盡相同，故本研究將探討有助於數學學習的不同遊戲模式，期望能促進學習投入度和加強學習成效。

2.2. 動作角色扮演遊戲與 Roguelite 遊戲

角色扮演遊戲（Role-playing game, RPG）是常見的遊戲類型之一，易引起學生對於虛擬角色的共鳴，使學生投入遊戲情境，進而影響到現實世界的表現（Simpson & Elias, 2011）。RPG 可再藉由不同的遊戲性質與特色進一步分類，ARPG 強調允許玩家奔跑、行走等多種動作操作，並建立劇情探索關卡供玩家體驗，玩家可經由擊敗怪物和破解任務獲取經驗值，當經驗值累積達到升級標準時，玩家可以得到新的能力或魔法技能（Kao, 2020），然而 ARPG 尚未考量學生本身的學習成效差異。因低學習成就者需要較多時間學習，導致與同齡的高學習成就學生的學習程度產生落差，造成高低學習成就者之間的數學學習成效落差隨時間增加

（Scammacca et al., 2020），並且低程度學生的工作記憶所需花費的認知處理較多（Çakir & Ergül, 2022）。納入教學方法分析遊戲特性，根據呂錘卿（2022）指出問題解決教學法由教師引領學生探索教學內容，發掘問題並循序漸進的解決問題。ARPG 由遊戲機制引導學生發掘問題，透過達成各個關卡的目標歷程逐步解決問題，引起學生思考與運用學習概念，因其特性著重在思考和問題解決的過程，透過逐步解決學生問題達成學習目標，但對於鞏固記憶可能較缺乏。

為了幫助低學習成就學生鞏固知識，呂錘卿（2022）指出個別化取向的教學設計之精熟學習策略中，透過調整課程步調、較小的單元、多次形成性評量和充實活動，能夠讓大多數學生達到精熟的標準。Roguelite 是奇幻角色扮演遊戲的一個子類型，繼承 20 世紀 80 年代遊戲《Rogue》的核心遊戲玩法（Dias et al., 2022），其中「永久死亡」機制指若玩家在遊玩關卡中不慎失敗，將造成不可挽回的後果，需要再使用新角色挑戰（Szabados et al., 2022），程式生成的隨機地圖也能夠為每次遊戲提供不同的體驗（Lemoine, Laforcade, & George, 2023）。重複演練的過程，可達到多次形成性評量，並透過多個關卡設計小單元任務，以過關與否檢測學習者的精熟標準，整體遊戲機制較符合精熟學習策略，對於數學需要大量練習與精熟的科目，採用熟練遊戲之特色作為學習輔助策略，有助於達到精熟目標，結合 Roguelite 遊戲模式將能夠幫助低學習成就學生依照個人化需求進行學習，增進學習成效。Roguelite 獨特機制結合隨機環境的規則，使玩家重複遊玩時，可避免相同感受，豐富遊戲體驗，能達到重複練習、個人化的需求，建立不同以往的數學學習遊戲體驗，探討 Roguelite 應用於教育的可能性，也能拓展更深層的教育遊戲意涵，為遊戲式學習開闢一條新的道路，供未來莘莘學子使用。

3.結論

本研究探討 ARPG 和 Roguelite 遊戲在國小數學公因倍數單元的應用，對學習成效與學習投入之影響，透過回顧遊戲式學習、ARPG 和 Roguelite 遊戲對學習動機和學習投入相關文獻後，針對研究問題一，本研究推論 Roguelite 遊戲較能顯著提升學習成效，ARPG 組之學習成效能達到平均水準之上。由於 Roguelite 遊戲強調精熟學習策略，有助於鞏固學生的數學基礎能力，促進低成就學生吸收知識。ARPG 則傾向問題解決教學法，著重在思考探索層面，需要擁有相關的先備知識，學生才有能力完成遊戲活動，而非精熟學習。針對研究問題二，遊戲是激勵學生的有效學習工具，但依據 ARPG 的遊戲特性，有過場、移動等需要玩家執行與學習較無關的行為，可能導致學生注意力分散，影響後續的學習行為，Roguelite 則需要不斷面對數學學習任務，有助於學生透過行為投入於學習當中，因此本研究推論 Roguelite 組的學習投入顯著高於 ARPG 組，但兩者皆在平均水準之上。本研究建議教學者可嘗試採用 Roguelite 輔助教學，達到提升低學習成就者的學習成效與學習投入之目標，未來研究建議設計 ARPG 和 Roguelite 遊戲的實證性研究，將能夠提供未來數學教育實務應用的具體建議，以及探討其對提升學習動機與學習成效的差異，促進不同遊戲種類設計對數學教育的貢獻。

參考文獻

呂錘卿（2022）。**新教學設計與原理**。五南。

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal*, *108*(2), 115-130. https://doi.org/10.1086/525550

Brumm, L., & Rathgeb-Schnierer, E. (2023). The relationship between accuracy in numerosity estimation, math achievement, and math interest in primary school students. *Frontiers in Psychology*, *14*, Article 1146458. https://doi.org/10.3389/fpsyg.2023.1146458

Çakir, R., & Ergül, C. (2022). Investigation of working memory performances of children with low early mathematics achievement. *Education and Science*, *47*(209), 133-153. https://doi.org/10.15390/eb.2022.10389

Dias, D. M. P. L., Sousa, J. P. P., Barroso, B. C. V. B., & Magalhaes, I. M. B. (2022, September 16-18). A novel procedural content generation algorithm for tower defense games. International Conference on Algorithms, Computing and Systems, New York, NY, United States.

Evans, D., & Field, A. P. (2020). Maths attitudes, school affect and teacher characteristics as predictors of maths attainment trajectories in primary and secondary education. *Royal Society Open Science*, *7*(10), Article 200975. https://doi.org/10.1098/rsos.200975

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207-219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Kao, D. (2020). The effects of juiciness in an action RPG. *Entertainment Computing*, *34*, Article 100359. https://doi.org/10.1016/j.entcom.2020.100359

Kayan-Fadlelmula, F., Sellami, A., Abdelkader, N., & Umer, S. (2022). A systematic review of STEM education research in the GCC countries: Trends, gaps and barriers. *International Journal of STEM Education*, *9*(1), 1-24. https://doi.org/10.1186/s40594-021-00319-7

Lemoine, B., Laforcade, P., & George, S. (2023, April 21-23). An analysis framework for designing declarative knowledge training games using Roguelite genre. International Conference on Computer Supported Education, Lisbon, Portugal.

Linek, S. B., Schwarz, D., Bopp, M., & Albert, D. (2009, November 7-8). Game-based learning-conceptual methodology for creating educational games. International Conference on Web Information Systems and Technologies, Shanghai, China.

OECD. (2023). *PISA 2022 results: Factsheets – Chinese Taipei*. https://www.oecd.org/publication/pisa-2022-results/country-notes/chinese-taipei-ebda1f30/

Scammacca, N., Fall, A.-M., Capin, P., Roberts, G., & Swanson, E. (2020). Examining factors affecting reading and math growth and achievement gaps in grades 1–5: A cohort-sequential longitudinal approach. *Journal of Educational Psychology*, *112*(4), 718–734. https://doi.org/10.1037/edu0000400

Simpson, J. M., & Elias, V. L. (2011). Choices and chances: The sociology role-playing game—The sociological imagination in practice. *Teaching Sociology*, *39*(1), 42-56. https://doi.org/10.1177/0092055X10390646

Streicher, A., & Smeddinck, J. D. (2016, July 5-10). Personalized and adaptive serious games. Entertainment Computing and Serious Games, Dagstuhl Castle, Germany.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Elsevier.

Szabados, G. N., Bába, É. B., Fenyves, V., Bács, Z., Molnár, A., Ráthonyi, G., Rizwan, H., & Orbán, S. G. (2022). Roguelike games: The way we play. *International Journal of Engineering and Management Sciences*, *7*(4), 80-92. https://doi.org/10.21791/IJEMS.2022.4.7.

Weng, C., Tran, K. N. P., Yang, C.-C., Huang, H.-I., & Chen, H. (2024). Can an augmented reality-integrated gamification approach enhance vocational high school students' learning outcomes and motivation in an electronics course? *Education and Information Technologies*, *29*, 4025-4053. https://doi.org/10.1007/s10639-023-11966-4

游戏化学习理念下的小学数学教学活动设计与实施

——以“谁是卧底”之倍数与因数复习课为例

Design and implementation of primary school mathematics teaching activities under the concept of gamified learning

——Taking "Who is an Undercover" Multiplier and Factor Review Class as an Example

高佳慧[1] 魏居林[2] 张誉馨[3]
清华大学附属小学昌平学校
Gaojiahui77@163.com

【摘要】在新课标理念指导下，突破传统教学方式，采用游戏化学习理念、借助《谁是卧底》游戏程序设计北师大版五年级上册第三单元《倍数与因数》的单元复习课，总课时为 2 课时，并进行实践探究，最后对小学数学游戏化学习提出思考，优化游戏组织，游戏与课堂相结合，适时组织游戏复盘。

【关键词】游戏化学习；谁是卧底；活动设计

Abstract:*"Under the guidance of the new curriculum standards, we have broken through traditional teaching methods and adopted a gamified learning concept. With the help of the game program "Who is an Undercover", we have designed a unit review lesson for Unit 3 "Multipliers and Factors" in the first volume of the fifth grade of the Beijing Normal University edition. The total class time is 2 hours, and we have conducted practical exploration. Finally, we have proposed ideas for gamified learning in primary school mathematics, optimized game organization, combined games with the classroom, and timely organized game reviews.*

Keywords：*gamified learning,Who is the Spy,activity design*

1. 前言

《义务教育数学课程标准（2022 年）》明确指出“有效的教学活动是学生和教师共同参与，是学与教的统一”，教师组织、引导学生并与其合作，学生作为主体而主动学习。新课标还建议采用探究式、互动式等方式丰富教学方式，强化情景设计和问题提出，发挥其对学生主动参与学习活动的促进作用，实现“有效教学中师生教与学的统一”。

日常教学中的复习课常常为教师总结本单元学习内容或师生间低思维层级的一问一答，学生参与度低；最后教师直接组织单元质量检测，学生做完一张卷子就相当于复习了一个单元，学生兴趣度低，这些都不完全符合新课标理念，学生的能力素养得不到提升。

改变传统的课堂教学方法，发展高阶思维，游戏化学习早已应运而生，尚俊杰、庄绍勇等从学习动机、学习方式等角度阐述了游戏的教育应用价值，国内外的诸多学者也通过心流实验验证了游戏化学习的学习效果，在活动过程中实现素养目标。

综上，由于复习课学生参与度和兴趣度低的问题现状，本文在新课标“教与学统一”、“丰富教学方式”等理念的指导下，突破传统教学方式，基于游戏化学习理念，采用《谁是卧底》的游戏模式，与数学内容相结合，设计与实施北师大版五年级上第三单元《倍数与因数》的单元复习课。

2. 游戏化学习理念的概念界定

【基金项目】北京教育学院 2022 年课题：游戏化学习视角的小学数学作业设计研究（课题编号：YX2022-02）；2023 年北京教育学院协同创新学校计划：学习科学视角下的小学数学教与学的研究。

游戏化学习理念认为将游戏元素、游戏机制和游戏思想应用到非游戏的学习活动中以此设计教学活动，主要包括目标设定、规则设定、反馈机制、激励机制，将学习目标、教学目标内隐到游戏过程中，重点在于借助游戏这种有趣味、有挑战性、高参与度的特点来调动学生的学习主动性、提高学生的学习效果。在此理念下的师生关系是，游戏化学习理念明确教师与学生双主体的地位，教师作为创造学习环境、设计活动支架、引导过程发展的主导者，学生作为游戏中亲身体验、自觉思考、主动探究的主体，学生在教师基于游戏要素创设的最近发展区中，自然而然的学习并解决问题。

3. 在游戏化学习理念下的小学数学教学活动的设计与实施

笔者以小学数学教学活动“谁是卧底”为例，介绍设计流程以及基于此游戏设计的教学实践。

3.1. 学习者特征分析

本课程的学习对象为五年级学生，从智力因素上来看，学生在知识基础上已经认识了自然数、小数、倍数因数等有关数的知识；在认知能力上正处于由低段向高段的过渡阶段，具有比较强的自行探究的能力。学生在观察、思维、语言表达能力方面都有了较好的提高。从非智力因素上来看，学生们已经接触游戏，并且对游戏很感兴趣，同时也很喜欢提问和探索。

3.2. 学习目标

目标要以学生为中心并能够指导学生的学习活动和评价，教师要思考学生在此游戏化教学活动中能做什么，活动后能收获什么等问题，基于此设计了以下的学习目标：

1. 尝试运用关联、求异、排除的方法玩思维游戏，判断卧底的情况。
2. 体验完思维游戏带来的挑战，提升思维的敏捷性。
3. 通过游戏，使学生巩固有关数的概念知识，进行知识串联.

3.3. 游戏化教学实践

通过小学数学教学中《倍数与因数》单元复习课的游戏化教学案例进行实践，验证了游戏化教学模式能够提升学生思维，帮助他们建构知识间的联系，在此过程中也培养了学生的学习兴趣，提高学习效果的有效性。

“谁是卧底”之因数与倍数复习游戏规则：第一，参与游戏者均拿到写有一个数的卡片，其中有 1 人与他人所拿卡片的数不同，即为卧底。第二，不要直接说出卡片上的数，间接表达数的特征（不重复），要给同胞暗示，但不能让卧底发现。第三，每轮描述完毕，投票选出怀疑是卧底的人（小组），得票最多者出局，若两者票数相同，则进入 PK。第四，若卧底被投票出局，则游戏结束，反之继续；若三轮过后，卧底仍隐藏，则卧底获胜。

具体的教学过程见下图，包括两个课时的内容。

环节	设置目的	教师操作	学生活动
情境导入	激发学生的学习兴趣，给予学习支架	给出一个两位数	说出这个数的特征（例如数的组成）
		引出游戏，介绍规则	明确规则
知识共学	主要是熟悉规则，在游戏过程中串联相关知识，总结游戏经验，知道取胜的基本策略。	个人战：选取 5 个人上台进行演示，对其他人提出要求	台上人参与游戏，其他同学认真倾听，记录信息
		小组赛：以小组为单位进行谁是卧底游戏，教师作为游戏主持人推动游戏进程	小组内成员进行分工（记录，投票，智囊团，表达），所有人参与游戏
拓展提升	通过反思交流，巩固游戏的玩法和策略，更是总结相关数学知识。	引导学生复盘整局游戏，提出相应问题，例如：哪一组表述的更好?	认真倾听，并反思自己整局游戏中的表现
总结	回顾课堂内容，进行教学提升。	通过提问引导学生梳理本堂课内容	跟随教师回顾总结和提升

图 1 第一课时内容

环节	设置目的	教师操作	学生活动
情境导入	调动学生积极性	直接引入游戏	再次明确规则，学生参与其中
游戏共学	回忆游戏玩法，在已有经验的基础上不断巩固游戏的玩法和策略	小组赛：以小组为单位进行谁是卧底游戏，教师作为游戏主持人推动游戏进程	小组内成员进行分工（记录，投票，智囊团，表达），所有人参与游戏
拓展提升	了解一些制胜策略，并从策划者角度出发，寻找共性较多的两个数，综合运用所学数的知识。	引导学生转变角色，从玩家变成出题者，要怎样设置一组数?	学生认真思考，并尝试自己出一组数
		选取学生设计的数进行展示，引导学生思考哪组数出的更好?	学生认真思考并回答（这一组数共同特征得比较多）
总结	回顾课堂内容，进行总结提升	通过提问引导学生梳理本堂课内容	跟随教师回顾总结和提升

图 2 第二课时内容

本次教学实践是基于游戏化学习的理念下，是一次新的尝试，用谁是卧底这样的大型游戏和学生学习的内容进行串联进行游戏化教学。从此次教学实践来看，有以下几个特点：

①调动思维参与，促进同伴协作

数学游戏化学习活动改变传统教学中的学生通过教师讲授、做大量习题等方式同化、顺应数学知识概念，游戏规则、游戏机制等元素都可能使学生产生感官或思维上的刺激，有利于学生同化、顺应知识概念，构建图式体系、形成高层次的认知结构。

在游戏开始前，教师对学生进行引导，比如教师给出 72 这个数字，对学生进行提问，“你可以怎样描述这个数？”有的学生会说“十位上是 7，个位上是 2”，有的学生会说“这个数是一个合数”，还有的会说“十位和个位上的数字相加是 9”，总而言之，教师可以得到多种多样的答案，学生们在此过程中不断激活思维，积极地辩论与讨论，思考的越全面，在游戏中就越能够找到关键点进行判断。在游戏过程中，为了获胜，学生需要足够的专心和认真，要找到尽可能多的和数有关的信息，并且也要从别人的发言中去筛选信息，在提取信息的同时要推断自己的数与他人的数是否相同，也调动了他们思维的参与，学生在过程中经历猜测、推理、验证等过程，从而在推理能力方面能够有所提升。

②内容进阶助力思维进阶

根据布鲁姆认知领域的教育目标分类“分析、评价和创造”属于高阶思维，杜威及瑞斯尼克等认为高阶思维的发展是学生要在真实情景、真实问题中主动寻求得到的，“儿童在游戏中的行为水平略高于日常行为水平，游戏创造了儿童的最近发展区”。

在游戏设计中，难度要随着玩家的状态随时改变，使得玩家收获最佳的游戏体验。因此在本次教学实践中的两个课时就呈现了这个特点——有梯度的进行，第一课时学生作为玩家，能够利用所学知识积累游戏经验，努力制胜，第二课时学生作为出题者，揣摩出题人的想法，如何设计更好的题目，游戏角色的转变，也激发了学生的学习兴趣。

③增加游戏的互动性和竞争性，提升学生的兴趣度、参与度

游戏化学习理念明确教师和学生双主体的地位，强调学生是主动的学习，所以会倒逼教师基于游戏思想、游戏要素、游戏程序下合理设计在学生最近发展区的教学活动。通过谁是卧底游戏这样的方式增加了课堂的趣味性，丰富了课堂内容，提高了学生学习的积极性，通过几轮游戏自己能够总结出游戏经验。在游戏过程中进行小组 pk，在挑战中进行学习，并对表现突出的小组进行积分奖励，促进了他们的竞争意识。

④学生在数学游戏化学习活动中接受德育，发展道德品质

学生在不同的游戏种类中会获得不同方面的提升，在对决有输赢的游戏中，学生看待输赢的态度，输赢状态下的情绪心理的感受；在合作中，以小组为单位进行，这需要小组内分工明确，每个人找到自己在小组内的职责，帮助他们意识到如何分工，意见不同如何处理等；游戏结束时有积分排名，教师基于这样的反馈评价结果，适当引导学生积累解决非知识性问题的经验，获得方法，发挥游戏化学习的育人功能。

⑤引入现代科技，促进学生互动交流

在小组赛中，利用电脑演示，对小组进行积分，每加一分，就点亮一颗星星，学生们看的更加直观明了，并且在游戏化活动结束之后，排行榜前三名可以进行幸运大抽奖。



图 3 小组积分



图 4 刮刮乐抽奖

4. 小学数学游戏化教学的思考与改进

4.1.选取适当的游戏与课堂知识进行结合

数学游戏可以作为一种教学手段，与课堂教学相结合，更好地巩固和拓展课堂相关知识。但是要想保证游戏在小学数学教学过程中的有效性，作为教师要重点运用游戏元素，明确课堂教学的核心知识点。以本节课为例，教师要明确倍数与因数以及其他有关数的知识，进而在游戏过程中进行有针对性的数学学习，比如在课上引导学生对数的知识进行串联，从多方面去思考，学生也在此过程中对数越来越敏感，不断强化数的倍数特征等知识点的记忆。

4.2.合理安排游戏时机。

在本次教学实践中，游戏贯穿于整个课堂，但不是所有的都适合这样的安排，在常态的数学课中，还是以小游戏为主，教师在实施时应考虑实际情况。比如对于低年级学生而言，他们在课堂上的注意力时间比较短，当教学进行了一段时间后，可以导入课堂小游戏，创设游戏进行教学，吸引学生的注意力。

4.3.制定适当的游戏规则。

这包括班级公约以及具体到活动上的游戏规则，对于前者，更多的是强调参与游戏者的纪律方面的规则，大家要在游戏中遵守纪律，后者更多的是强调游戏规则，介绍时要简单明了，所有人都能看懂游戏规则。包括可以通过积分制等奖惩方法，促进他们形成竞争意识，激发他们的学习兴趣。

4.4.游戏细节再优化。

教师需要思考在游戏过程中哪个环节的游戏安排可以再优化，避免学生出现分歧。比如在此实践中的每一轮投票环节，由最开始的一个一个统计投某个组的人数，到大家同时投票，这样有效的避免的跟票的嫌疑。其他的还需根据教师的自身情况进行合理优化。

4.5.适时结束游戏并认真复盘，不宜太长。

游戏时间进行太长，学生的兴趣值下降，注意力涣散，甚至会出现其他不遵守规则的现象。游戏结束之后要注意对引导学生对数学知识或者是数学思维的总结，并结合游戏过程中学生们的表现进行评价，从育人的角度帮助学生进行成长。

5.结语

基于游戏化学习理念，在小学数学课堂中进行了游戏化教学的初步实践探索，能够发现游戏化教学能够使得小学数学课堂氛围更加的丰富和有趣，促进了教师在数学课堂教学上的改革和创新。希望此教学实践为大家提供实际案例，吸引更多的教师进行游戏化教学的实践。

参考文献

张金磊，张宝辉.游戏化学习理念在翻转课堂教学中的应用研究[J].远程教育杂志,2013,31(01):73-78.

尚俊杰,庄绍勇. 游戏的教育应用价值研究 [J]. 远程教育杂志, 2009, 17 (01): 63-68.

尚俊杰,庄绍勇,李芳乐,李浩文.教育游戏的动机、成效及若干问题之探讨[J].电化教育研究,2008(06):75.

张金磊，张宝辉.游戏化学习理念在翻转课堂教学中的应用研究[J].远程教育杂志,2013,31(01):73-78.

魏莹. 小学数学游戏化教学逻辑研究[D]. 福建师范大学, 2021.

Weisberg D S, Hirshpasek K, Golinkoff R M. Guided Play: Where Curricular Goals Meet a Playful Pedagogy[J]. Mind Brain and Education, 2013, 7(2):104-112.

Vygotsky, L, Trans. M. Cole.The Role of Play in Development. In Mind in Society. Cambridge[J]. MA: Harvard University Press.1978:92-10）

游戏化学习在小学数学综合实践教学中的应用与开发研究[1]

——北京教育学院“教师培训”项目之一

Application and Development of Gamification Learning in Comprehensive Practice Teaching of Mathematics in Primary School ——An Empirical Research Project from “Teaching Project”in Beijing Institute of Education

刘月艳[1*]，**王雪君**[2]，**白永潇**[1]，**曹慧**[1]，**岳志刚**[3]，**汤佳佩**[4]

[1]北京教育学院
[2]北京小学广内分校
[3]清华大学附属小学永丰学校
[4] 哈尔滨工程大学 智能科学与工程学院北京市朝阳区教育科学研究院
*liuyy1998@163.com

【摘要】本研究是以高等院校与一线学校联合展开的项目实证。本研究针对小学数学教材中的综合实践内容，传统和现代的益智器具，小学数学教材中的重要概念，有选择地进行价值分析和课程实践的分类研究。首先，确定具体的研究内容和相关的益智器具；其次，分析上述内容思维水平和适合的游戏方式，探讨实践活动的原生价值和教育价值；接下来，主体部分是对现有的游戏化学习的综合实践课程进行深度分析，探索小学生的思维特征，力求达到通过游戏化学习的实践课程，促成学生思维经验的积累和优化，学生思维品质的提升。
【关键字】益智器具；思维经验；实践课程；游戏化；小学数学

Abstract：*This is an empirical research project jointly carried out by universities and front-line schools. This research aims at the comprehensive practice content, traditional and modern educational devices, and the important concepts in the primary school mathematics textbook, and selectively carries out the value analysis and the classified research of curriculum practice. First of all, it determines the specific research content and related educational devices. Second, it analyzes the thinking level of the above-mentioned content and suitable game mode, and explores the original value and educational value of practical activities. Third, the main part is to conduct in-depth analysis of the existing comprehensive practical courses of game-based learning, explore the thinking characteristics of primary school students, and promote the accumulation and optimization of students' thinking experience, and improve students' thinking quality based on game-based learning practice curriculum.*

Keywords: educational devices, thinking experience, practical courses, gamification, primary mathematics

1.以游戏化学习为主题的教师培训项目背景分析

1.1.项目背景及分析

以游戏化学习为主题的教师培训项目主要包括2020-2022连续三年的专题培训和2023年北京教育学院第三期“协同创新学校计划”。这些教师培训项目试图带领项目校数学教师进

[1] 北京教育学院2022年课题：游戏化学习视角的小学数学作业设计研究（课题编号：YX2022-02）；2023年北京教育学院协同创新学校计划：学习科学视角下的小学数学教与学的研究。

行理论学习和实践研究，试图解决学校体现课标要求的课堂教学所需，解决“立德树人”目标的学科落实问题，培养学生的规则意识、问题解决能力和创新精神。“协同创新学校计划”是北京教育学院第一期“协同创新学校计划”之“游戏化学习在小学数学综合实践课中的运用”项目的延续，同时项目团队有已结题和在研的关于游戏化学习、高阶思维相关课题的研究成果。项目团队具备开展基于学习科学、游戏化学习而进行项目培训的实践基础和研究基础。

2015 年北京下半年北京教育学院“协同创新学校”计划的提出为项目实施提供了可能性。本着“把培训课堂建到学校，让校长教师研修真正发生”的理念，从培养学生核心素养的要求出发，围绕学校办学与教育教学目标和要解决的实际问题，学院与各区教师培训机构、中小学合作开展为期三年（2016—2018 年）的行动研究，真正促进学生发展、校长教师发展和培训者的发展，实现学校的持续改进，促进基础教育优质均衡发展。项目在三年中滚动实施，达到 200 所项目学校。项目学校以郊区学校为主，以普通学校为主。依据《北京教育学院“协同创新学校计划”实施方案》（〔2016〕1 号），有了本项目的出发点和落脚点。

2017 年 1 月，中小学幼儿园教师培训课程指导标准(义务教育数学学科教学)指出：积累数学活动经验

2019 年 6 月 3 日《中共中央 国务院关于深化教育教学改革全面提高义务教育质量的意见》中提出：好奇心、想象力、求知欲，激发学习兴趣，提高学习能力。促进信息技术与教育教学融合应用。推进“教育+互联网”发展，按照服务教师教学、服务学生学习、

2020 年 10 月 13 日中共中央 国务院印发《深化新时代教育评价改革总体方案》（四）改革学生评价，促进德智体美劳全面发展。

2022 年《义务教育课程方案（2022 年版）》中写到“合理设计小学一至二年级课程，注重活动化、游戏化、生活化的学习设计。”

1.2.项目背景分析结论

基于以上的背景分析，笔者认为：

第一要强调立德树人，以学生发展为本，希望能把老师们带到一个轨道上来。学生发展上存在什么问题？数学教学中，学生害怕数学，厌烦数学，而数学又很重要。如何来解决这个问题呢？提高孩子的学习兴趣很重要。知识是力量，能够把知识变成力量的桥梁是智慧，智慧的源泉在思维，思维的提升在于乐于思考的意识。

第二要解决教师和学校所急需，开设助推不少于 10%的综合实践课程的落实。

（1）落实《课标》中的四基目标，尤其新“两基”。课程标准的要求，要落实课程标准中的具体化内容。具体操作是基本活动经验获得的有效途径。

（2）解决学生的乐学问题：智慧是知识到力量的桥梁，智慧源是思维，思维核心点——乐于思考问题，而游戏化学习能促进学生乐于思考，动手操作能够积累和优化思维经验。

（3）填补“情”与“情”的关系。目前教学中更多的是教师的教授与学生的接受中的“授”与“受”的关系，而少了人与内容之间“情”的关系。

（4）综合实践课程需要一个有识有能有志的团队来开发。

当前小学数学教学的首要任务是激发小学生学习数学的兴趣，以及培养学生主动学习数学的习惯，游戏化学习凭借自身所具有的独特的魅力，能够吸引学生更多的注意力，使学生在短短的几十分钟内就可以有效的提高学习效率，为了达到这个效果，教师需要做的事不仅要选择与教学内容相适应的的教学方法，而且更重要的是教师还要能够把握整个教学过程。

2.以游戏化学习为主题的教师培训的实施

2.1.实施目标

（1）通过对学习科学的相关理论进行学习和研讨，了解学习科学的基本理论，理解游戏化学习的前沿理论，及其在改进和优化小学数学教学的意义和价值。

（2）通过案例分析和教学实践，理解游戏化学习应用于教学的思路与方法，提炼有效的教学策略，能够掌握游戏化学习在小学数学教学中运用的课例开发流程。

（3）形成基于学习科学的小学数学系列教学资源。

（4）经历课例研究过程，提升项目校数学教师教学能力。

（5）以“实践-提炼-辐射”的方式推进，发挥项目校的区域辐射和示范引领作用。

2.2.实施方式

2.2.1.培训方式

以校本研究的方式开展培训，相应的方式包括：集中讲座、自主阅读、工作坊、教学观摩、实践研讨、案例分析、学生调研、资源开发、网络学习等。

2.2.2.组织形式

本研究采用市区校三级管理的组织形式，包括项目负责人与项目助理、区学科指导专家、项目校负责人。项目研究人员共同体的构成：理论专家，实践专家，项目学校的参与者。项目校的具体研究主题由市级项目团队与项目校负责人共同确定，根据本校的已有研究基础和期待解决的问题有针对性开展研究与培训。

2.3.实施流程

2.3.1.课程结构图（见图 1）

左图为 2016-2018 年课程结构图；右图为 2023-2025 年课程结构图。

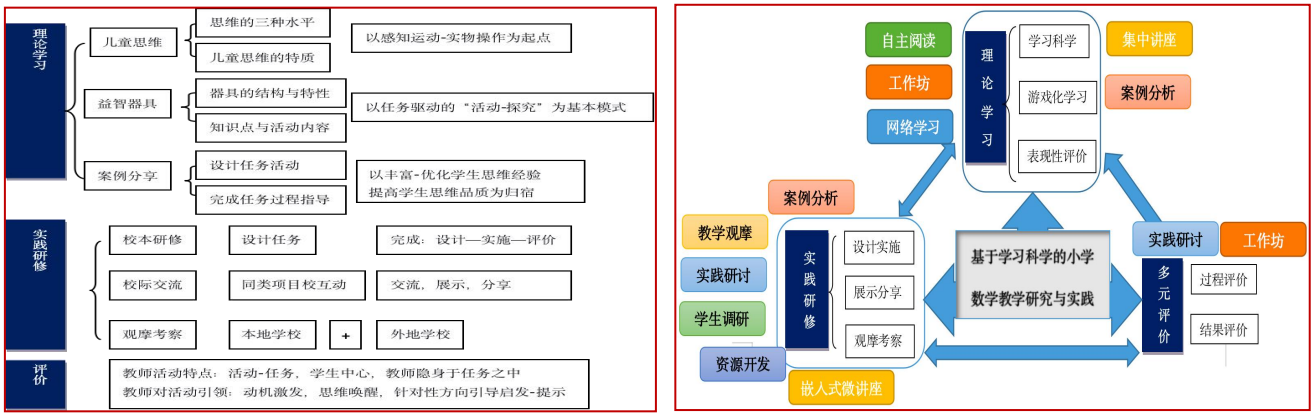


图 1：课程结构图

2.3.2.实施流程图（见图 2）

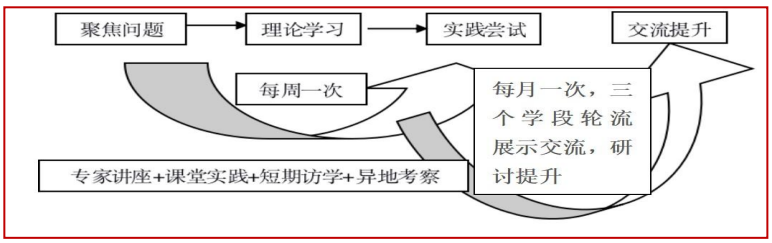


图 2：实施流程图

2.3.3.任务及评价流程图（见图 3）

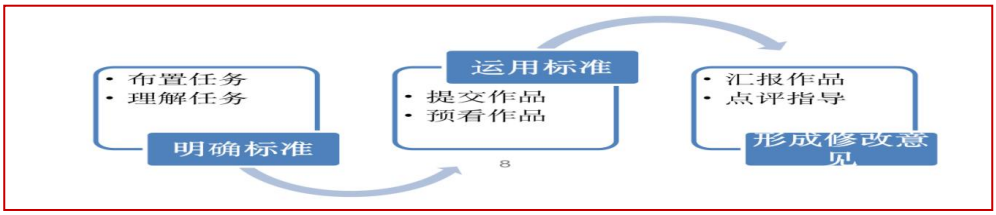


图 3：任务及评价流程图

3.以游戏化学习为主题的教师培训的初步成果

3.1.游戏化学习让课堂变乐园

数学学习好不好玩？不同的学生也许会给出不尽相同的答案。然而，随着游戏化学习方式进入数学课堂，越来越多的小学生开始体会到了数学的“好玩”。通过各种课堂游戏化操作活动，小学生发展了数学思维品质，更发掘出了“好玩的活动”中的数学价值。思维品质是核心素养的重要组成部分。在发展学生的核心素养的过程中，我们找到了一个抓手，通过设计好的学生活动，积累和优化学生的思维经验，进而提升学生思维品质，用游戏的方式来组织教学活动。我们强调思维教育，以问题解决为线索，培养学生发现问题、提出问题、解决问题、以及对问题解决过程的反思和回顾并形成初步的方法与策略的意识。强调操作是起点，思维提升是核心，问题解决能力的培养是目标。通过操作活动学生体会到数学好玩，更重要的是要引导学生在操作后的冷静思考，发现其中的数学价值。

游戏化学习在数学综合实践课中的运用项目基本的运行模式是：以课为载体，每个项目校在责任专家和指导教师的带领下，选择一个内容设计成综合实践课，校内经过三次研讨，第四次活动三个项目校在一起进行交流研讨，聚焦核心问题，澄清概念。

3.1.1.多种游戏资源进课程体现了的丰富供给

【案例 1】益智器具七巧板课程。《认识七巧板》ZXR 老师介绍过传统益智器具七巧板的来历后，一年级的孩子们两人一组旋转摆弄着手中的小木板，尝试用各种形状拼成一个三角形。过程中，孩子们或相互交流，或皱眉不语。渐渐地，活动记录单上出现了七八种拼法，一双双小手举了起来，想要急切地表达他们对这场空间探索之旅的认知。

【案例 2】挑战了不起——度量课程。《克与千克的认识》TJP 老师，要带领三年级学生接受度量的“挑战”——感受克与千克。三个苹果、一筐鸡蛋、一袋子棒冰、几大桶爆米花……感受了一千克的哑铃重量的学生，开始在日常生活用品里进行挑选，估量出一千克的重量。

【案例 3】24 点游戏课程。《24 点》DT 老师让三年级学生围坐在一起，24 点游戏取代了纸笔的计算，兴致勃勃的孩子们在欢笑中开启了颇为烧脑的运算活动，体会如何将运算技巧和算法多样化运用于解决问题之中。

“游戏不是学习目的，而是学习方式。”“游戏化的数学学习方式能够最大程度地调动学生的兴趣，在实践操作中学生由直观到抽象，一步步加深了对空间、量感、数感的认知，在不知不觉中积累和优化了他们的数学思维经验，提升了思维品质和解决问题的能力。”成为了大家的共识。

3.1.2.学生成为课堂学习主体实现教与学的转变

如何将游戏化学习贯通数学综合实践课？怎样的活动能积累孩子的思维经验，帮助孩子们优化思维经验？“学生的每一次犯错，都是他成为最好的自己的过程，教师只需在旁边给予适当的指导，给予他自我提升的时间与空间。”这是 ZHPY 老师写在反思中的话。

谁是课堂的主人？谁是游戏化学习的主体？在老师们的意识里渐渐有了清晰的答案。在北京明远教育书院实验小学郭勇老师的课堂上，学生对不规则物体的测量经历了“发散思维—动手实践—自主探索—合作交流”的主动学习过程。同校的张小晨老师执教《五子棋》一课，学校数学老师和三年级学生变成同学，共同研究五子棋。在和老师的小组合作中，学生积极尝试、敢于质疑，在多次的体验中找到了取胜的关键。

《实践园中的测量》，W 老师把学生带到真实场地中学习测量，教师团队有效合作，两人主讲，一个负责布置任务、组织活动，一个负责活动评价。新颖的课堂形式让学生兴奋的同时，及时反思和评价也得到了观摩的教师们的一致认可。

游戏化学习中，我们强调教师的第一目标不再是知识，而是关注学生的实际获得。同时，希望孩子关注游戏的规则。规则在社会中无处不在，在游戏中经历规则还可以培养孩子的担当精神和品质。

3.1.3.师生的获得感最大化促进共同发展

“在愉悦的环境下，完成挑战性的内容。”是很多教师对游戏化学习的评价。事实上，在游戏化学习和数学综合实践课结合的过程中，课堂有了很大变化，学生和老师成了共同的受益者。

在中国农业大学附属小学，针对一年级学生开设的数学综合实践课程《快乐的七巧板》是堂 60 分钟的大课。设课之初，大家都担心刚刚入学的一年级孩子能否适应得了。出乎所有人意料的是，因为加入了游戏化的体验，一年级的孩子们被有趣的形式和内容所吸引，求知欲被激发出来，兴高采烈地参与到做、拼、摆七巧板的过程中，积极地进行着关于空间感知的小学数学课程学习。

“在玩中学，在学中玩”，游戏化学习充分调动了学生的积极性。高举的双手，争相的发言，积极的尝试，勇敢的表达成为了数学实践活动课上的常态。充分思考、实验和交流时间，也使得学生深入思考的能力得到了有效锻炼；小组合作，汇报交流更培养了孩子的分享意识和沟通能力。

欣喜于孩子们的变化，项目校的教师开始在课堂上“放权”，俯下身子和孩子们对话。同时，越来越多的教师不仅仅满足于课堂教学经验的增长，开始将目光投入到了对游戏化学习的研究。项目组成员的研究成果也得以发表。

3.2.游戏化学习的几点共识

以感知运动-实物操作为起点；以思维教育为取向—问题解决为暗线；游戏中培养规则意识，实践中感悟契约精神。这就构成了整个项目的共同价值追求——立德树人精神在学科育人中的落实。



图 4：几种类型课举例

3.2.1.以感知运动—实物操作为起点

动手操作为抓手，操作的对象可以根据内容和主题有不同的选择。本文以三种类型课为例来说明，操作对象与活动内容和主题的关系。

【案例 1】《24 点游戏》三年级的数学综合实践课。课堂上教师设计了三次动手玩，玩扑克牌。第一次是用两张牌来找 24 点，根据两张牌上的数字通过四则运算（只有加和乘）得到 24。通过多次操作归纳总结得到只有五种组合可以得到 24。第二次活动是用三张牌来找 24 点，多了一些规定“A,J,Q,K 都看成 1”这样要求学生在活动中只在对三个 1-10 的数字通

过四则运算得到24。方法会多一些，这样把运算技巧和算法多样化能够主动地运用于解决问题之中。第三次活动是把刚刚得到的算式进行分类整理，帮助学生优化思维经验。

【案例2】《认识七巧板》一年级的数学综合实践课。本次课教师设计了三次动手操作七巧板的活动。第一次，初识七巧板，从形状上进行分类。1个正方形，1个平行四边形，5个三角形。第二次，再识七巧板，从大小关系上进行比较。两个大的三角形一样大，两个小的三角形一样大，两个小三角形还可以拼成1个正方形，也可以拼成一个平行四边形，还可以拼成那个中等的三角形。第三次，用七巧板拼三角形。用随便几块拼成三角形。在拼图的过程中能够积累图形的组拼的经验，在比较中能够优化学生怎么更快地拼成三角形的经验，对共同经验的总结再实践，进而达到提升学生的思维品质。

【案例3】《挑战了不起——千克与克》三年级的度量概念的综合实践课。本次课教师设计了三次动手体验和寻找1千克的活动。第一次，感受1千克有多重。为每组学生准备了重1千克的哑铃，每个学生都用手拿起来，通过掂一掂形成1千克有多重的初步感受。第二次，寻找1千克。以刚刚的哑铃为标准，寻找教师提供给小组的物品中大约1千克有多少。借助标准来找与标准相当的物品。第三次，寻找身边的1千克。再前两次活动积累和优化经验的基础上，再一次形成更多的经验。已经发表详见参考文献[2]。

3.2.2.以思维教育为取向，以问题解决为暗线

思维教育如何实现，我们进行了初步了探索，以问题解决为线索，培养学生发现问题、提出问题、解决问题、以及对问题解决过程的反思和回顾形成初步的方法与策略的意识。操作是起点，思维提升是核心，问题解决能力的培养是目标。通过操作活动学生体会到数学好玩，更重要的是要引导学生操作后的冷静思考，发现其中的数学价值。

在第一次三校交流研讨达成共识的基础上，每个项目校的团队成员又在各自责任专家的引导下，选择了新的内容，经过三次的研讨实践。在第二次活动中老师们分享了基于问题解决的数学综合实践课。游戏中做数学，实践中悟综合，成为了本次活动的最大特色。

明确问题——独立思考打算怎么做？——实践方案——聚焦解决中的真实困难——明确问题——修正思路——再实践——方法的回顾要贯穿始终——经验的分享——教师有结构的设计有利于发现规律——再尝试——转化为自己的思维经验。课堂中看到了老师和学生的变化。教师的目标不再知识是第一位，而更多地关注学生的实际获得。

【案例4】GY老师《有多大——不规则物体的测量》一课中，G老师将“问题”和“问题解决”有效贯穿于自己的课堂中。课堂上，学生们经历“发散思维—动手实践—自主探索—合作交流”的过程，改变了被动接受的学习方式。活动中，孩子们不仅体会到了数学就在身边，数学就在生活中，而且激发了学习的兴趣，获得了成功的体验，增进了学好数学的信心。在G老师的课中，教师成为了学生活动过程中的组织者、合作者和引导者，学生们互动交流、观点交锋、智慧碰撞，课堂变成了学生展示自我、发展智慧的舞台。本研究已经发表详细见参考文献[3].

【案例5】LJF老师带来的——《测量——到底有多长？》的课堂中。这节课中，L教师为学生们创设了一种和谐愉悦的学习氛围，积极调动学生参与，使学生各种感官全方位参与本节课活动。从“量小木条的长度”这一学生熟悉的生活实际入手，引导学生参与实际的测量活动，让学生经历用不同工具测量木条的过程，感受大小不同测量工具测量同一对象结果的不同，进而引导学生根据已有的测量经验探索测量方法。本节课中，L老师注重将主动权放给学生，引导学生操作-思考-交流，充分发挥了学生的积极性。

【案例6】YZHG老师呈现的一节三年级综合实践活动课——《制作七巧板》。这堂课中，学生们先是观察七巧板、构思制作七巧板的操作方案，然后自己独立动手制作七巧板。在实际操作与思考的基础上，经历了由关注图形到关注数的发现、提炼过程，学生们不仅体验了

动手的乐趣，也感觉到了做比想难，更重要的是感受到了七巧板的独特魅力——七巧板中的分数，也就是七巧板怎么玩都有数学。学生的动手、动脑能力在活动中得到发展与提升。已经发表，详见参考文献[4]。

3.2.3.游戏中培养规则意识，实践中感悟契约精神

在学科教学中培养规则意识是游戏化学习在小学数学教学中运用的突出价值之一。规则的培养分三个层面：理解规则，运用规则，创造规则。在项目推进过程中，多次讨论规则的问题如图 5。师生是规则意识是否有所提高，请看图 6 这是某小学每周一题，开始的规则是每个年级答每个年级的题，只有一个答题箱；师生的对话：老师，我可以答高年级的题吗？可以。那答高年级的答奖品应该不一样。老师接受你的建议。就多出了一个智慧箱。说明师生在这个过程中规则意识都自觉地应用于学习工作中。图 7 是其中一所学校在《快乐七巧板》课程开发过程中形成的流程图。图 8 是游戏还作业设计模版。图 9 是游戏化数学活动说明。

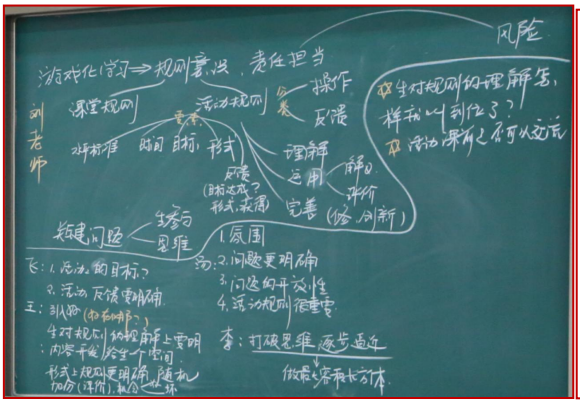


图 5：规则讨论



图 6：规则意识在师生行为中落实

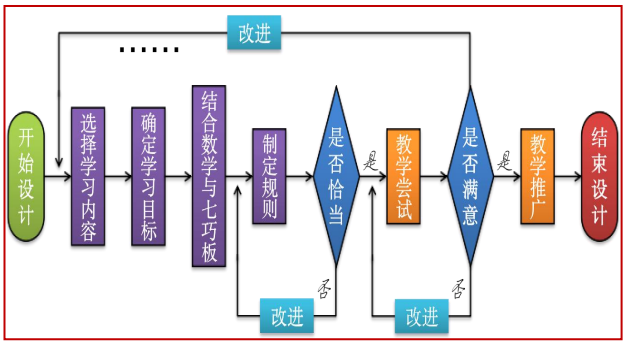


图 7：七巧板课程开发流程

设计内容		适用年级		设计教师	
设计	题目表述（包括完成作业用时）		目的：知识技能、思想方法、游戏元素		
	请用 3-5 个词概括本设计的特点：				
适用时段	课前	课中	课后		
作业解答参考答案					
作业反馈作业评价	作业反馈方式	作业评价框架			
其他关注点	我还关注了_____，我的思考是：				
疑问和困惑					

图 8：游戏化作业设计模版

开发者	游戏活动名称	适合范围			游戏类型			游戏参与者		游戏环境		游戏组织者要求	游戏图片或结构图	游戏规则简介	游戏价值分析
		游戏者几年级	课程内容	教学环节	电子游戏	器具游戏	卡牌游戏	单人	多人	网络环境	非网络环境				

图 9：游戏化数学活动说明

4. 结论

笔者主持过十多个不同发展阶段小学数学教师的培训，逐步形成了关于教师培训的主张：针对培训对象的特点，创设“社会建构”的环境，激发与支持教师的自我建构。在这个主张下， 开发并实施了游戏化的数学活动设计与实践为主题的专题培训和协同创新培训。

在培训课程的不断迭代中，我逐渐探索出了学术研究方向是：游戏化学习视角的小学数学的教与学，发表了二十多篇论文，这些工作主要围绕数学游戏化的课程资源开发与实践中的关键问题进行了探索，主要形成了以下的认识：

第一，游戏化学习视角下的数学活动是发展学生核心素养的可行路径之一。

从 2016 年就开始以课题研究和教师培训为抓手，项目团队带领一线教师共同开发游戏化学习视角的数学活动，通过设计恰当的学具和游戏活动，提出和探索凸显数学本质的问题，探索出了学生发展核心素养的路径。借助恰当的学具，组织具体操作，帮助学生积累和优化思维经验，进而提升思维品质。形成了系列研究文章，相关研究为新课标（2022 年版）的落地提供了实践经验。如“做个百数表”、“认识七巧板”、“小圆点”、“数学可以很美”等文章中，既关注了数学学科本质，又关注了学生是发展中的人，是有差异性的，为自主建构提供了良好的生态环境。文章被《人民日报-全国党媒》和《中国教育新闻网》转发推荐。

第二，高认知水平的任务的设计能够帮助教师提高数学专业素养；

高认知水平的任务设计需要教师对数学内容和对学生认知现状的深度理解，在培训中，为教师开发高水平的任务提供理论支持、工具支持以及策略支持，提高了他们的专业素养。高认知水平的任务设计基于对数学内容的理解和对学生认知现状的了解，重新构建高认知水平的任务。高认知水平的任务需要学生思维积极参与，典型活动包括说理、解释、推理、论证等。鉴于一线教师在教学实施中常把高认知水平的任务变成学生的低水平参与，提出在教学实施中运用恰当的方法保持学生高水平参与，如“先指向必要时指路”、“搭建脚手架”、“提供高水平示范”等具体建议。帮助学生由操作工变成解决问题的思维参与者。

第三，围绕游戏化和高认知水平任务设计的课程资源开发提升教师课程设计能力。

“游戏化学习在“七巧板课程开发”中的应用研究”和“快乐七巧板——《角的度量》第一课时教学与思考中。笔者将继续在游戏化的数学课程资源开发与实践方向继续深入研究。

参考文献

李铁安（2021）。**迷上数学：触动童心的数学文化课**。北京：教育科学出版社。

尚俊杰（2022）。**游戏的价值：教育游戏的设计与应用研究**。北京：北京大学出版社。

中华人民共和国教育部（2022）。**义务教育数学课程标准（2022 年版）**。北京：北京师范大学出版社。

汤佳佩、黄利华。游戏化学习在数学综合实践课程中的探索——“有多重”教学思考。**小学教学（数学）**，**2017（3）**，33-35。

郭勇。浅议游戏化学习的教学价值——以“不规则物体体积的测量”教学反思为例。**小学教学研究**，**2018（10）**，58-60。

刘月艳、杨小东、刘东旭。作业使学习持续发生——以北京市通州区东方小学《成长手册(数学)》为例考。**中小学数学（小学）**，**2022（04）**，6-8。

刘月艳、李友鹏。快乐七巧板——《角的度量》第一课时教学与思考。**小学教学设计（数学）**，**2020（Z4）**，80-83。

刘月艳。基于“高认知水平”教学案例分析——以“三角形的组拼”为例。**中小学数学（小学）**，**2020（09）**，52-56。

刘月艳、张雨濛。设计游戏化学习活动，促进学生数学思维发展——“做个百数表”教学实践与思考。**小学教学（数学）**，**2020（9）**，23-26。

刘月艳、裴杰。创编数独游戏，培养学生思维品质——以“有趣的圆点”一课教学为例。**小学教学（数学）**，**2022（9）**，28-31。

刘月艳、赵欣然。设计好玩的活动 提高学生思考能力——以《认识七巧板》教学为例。**小学教学（数学）**，**2021（Z2）**，81-82。

崔英梅、丁婷。从“先探后玩”到“玩中发现”——数学游戏的教学逻辑转向。**基础教育课程**，**2022（8）**，42-48。

理财素养实作模拟平台中操作行为与自我效能之关系研究

Investigating the Relationship between Junior High Students' Investment Simulation Behavior on their Financial Self-Efficacy in Hybrid financial literacy classrooms

史安晴[1]，李元萱[2]，吴俊育[1*]
[1] 台湾阳明交通大学
[2] 台湾清华大学
[*] jiunyu.rms@gmail.com

【摘要】本研究关注在理财素养课堂中，通过学生在模拟平台上的交易和浏览行为将其分组，并探索特定操作模式。分析了不同操作模式学生在交易金额上的差异，并比较了不同操作模式学生在金融操作自我效能和理财学习自我效能上的表现。研究结果将学生分为眼明手快组、谨慎多虑组和漠不关心组。虽然单笔操作金额没有显着差异，但眼明手快组在金融操作和理财学习自我效能方面表现较好。此外，随课程进行，学生的金融操作自我效能也有显着提升。建议老师在教学中可搭配模拟实作，鼓励学生多多实践，以提升学生理财相关的自我效能。

【关键词】理财素养教学；理财模拟平台；行为操作模式；理财自我效能；情境化学习

Abstract: *This study integrates practical teaching with a simulation platform in hybrid financial literacy classrooms. Within classroom instruction, junior high students are categorized by buying, selling, and browsing behaviors on the financial simulation platform to explore operational patterns. Analyzing variations in average transaction amounts among students with different patterns, the study compares investment behavior and financial self-efficacy performance at two time points. Three groups were identified: "Quick and Sharp Investors," "Information Seekers," and "Non-Interest Users." With no significant differences in transaction amounts, the "Quick and Sharp Investor" group consistently outperforms the "Non-Interest User" group in both financial literacy knowledge and confidence in financial literacy skills. Additionally, students' confidence in financial literacy skills significantly improves throughout the financial course. It is recommended that educators integrate simulated practical exercises into their instructional strategies to encourage students to engage in more hands-on practice, thereby fostering an increase in students' self-efficacy in learning financial literacy.*

Keywords: Financial Literacy Education, Financial Simulation Platform, Behavioral Operational Patterns, Confidence in Financial Literacy Knowledge and Skills, Contextualized Learning

1. 前言

金钱与我们的生活息息相关，对人影响足以改变一生。研究指出，当前消费者面临的金融环境相当危险(Boshara, 2010)，经济合作暨发展组织(OECD)也于 2005 年的一份报告强调了对于财务担忧。该报告指出金融知识匮乏在不同年龄层或各地理区域的族群中普遍存在。2012 年 OECD 对多个成员国进行理财素养的调查。结果显示，虽然在各国的共同努力下，大多数人已具备一定程度的财务知识，但仍有一定比例的人缺乏对更深入理财概念的理解 (Atkinson & Messy, 2012)。台湾金融研训院在 2022 年对 20 岁以上民众进行理财素养调查，结果显示超

过 78%的民众的理财素养仅有低或极低水平。其中拥有专科大学或以上学历者，金融素养为低和极低者，合计也都超过 60%。由上可见，尽管多国的理财素养教育正在进行，但仍有许多可加强的空间。因此本研究动机源于对当前金融环境的深刻关注和对理财素养教育的担忧，希望对中学校园中的理财教学效能进行了解，以異質學習分析手段(Authors, 2021)聚焦于了解理财模拟平台的操作模式对学生金融操作与理财学习自我效能的影响，提供具學理與實務價值之分析實證，協助建置優質理財素养教育系統，弭平財富落差，促進金融永續。

本研究目标在于透过对理财模拟平台操作行为的深入分析，揭示不同操作模式对学生理财学习的潜在影响因子。首先，分析中学生在理财模拟平台上的买卖操作和浏览行为，探索是否存在特定操作模式。其次，检验不同操作模式学生在理财模拟平台上单次操作金额的均值差异。最后，透过比较不同操作模式学生在这两种理财自我效能上的表现差异，期盼了解学生的平台操作行为模式是否对学生的自我效能产生影响。具体而言提出三个研究问题：

RQ1: 在理财模拟平台上，学生的买卖操作与浏览行为是否存在特定的模式？

RQ2: 不同操作模式学生在每笔操作金额均值上是否有差异？

RQ3: 不同操作模式学生在课程前后测的金融操作自我效能与理财素养学习自我效能是否有差异？

2. 文献回顾

2.1. 理财素养的定义

理财素养(Financial literacy)一词起源于美国非营利组织“Jump$tart Coalition for Personal Financial Literacy”，其为一种能做出明智且资讯充足的财务决策，并有效运用财务资源的能力(Hastings, 2013)。经济合作暨发展组织(Organization for Economic Co-Operation and Development, OECD, 2005)定义为：“金融消费者/投资者提高对金融产品和概念的理解。并透过资讯、指导或客观建议，发展技能和信心的过程更了解金融风险和机会，做出明智的选择行为，例如知道到哪里寻求帮助，并采取其他有效行动来改善他们的财务状况。”

2.2. 校园中的理财素养教育

从素养教育的“认知层面”而言，OECD(2015)指出，金融教育提供相关知识，培养消费者拥有判断与寻找资讯的技能，让他们能根据资讯做出符合自己利益最大化的决策。在 2005 年理财教育的国际研究报告中也建议：“理财教育应由学校开始，人们应在他们的人生中尽可能地提早被教育理财相关事物。”台湾为了巩固金融理财教育，教育部与金管会合作针对国中阶段发展三种不同教案，培养学生之理财素养。而尽管目前的国中社会领域或综合活动领域的教材中已然融入了金融知识，但对于金钱管理与规划教学内容仍然有加强空间。

除教学内容外，同样需要对不同族群针对性的教学(Goel & Khanna, 2014)。“钱”与“交易”可谓金融理财概念的基础，虽然不一定在课堂中学习，却要有十足的生活经验相伴。因此体验式教学在金融理财学习过程中十分合适(Collins & Odders-White, 2015)。体验式学习在行动、反思、概念化与新经验中不断反覆轮回以此来不断建构新的知识，让学生透过与环境的互动来建构他们的知识并探索正在与发展想法(Gregory, 2002)。因此本研究在理财素养教学与学习的过程中，透过严肃游戏的方式创造与生活相连的体验(Tang et al., 2019)，在兼顾学习与趣味性的同时达成教育目的(Susi, 2007; Gao, 2023)。

2.3. 操作行为对理财素养教育的影响

以下从“行为层面”切入思考。金融理财的行为所包含的面向十分广泛，Fernandes(2014)将金融行为被分为五个不同的面向进行评估，包括：为紧急基金储蓄、退休所需金额、银行和信用卡公司对信用评分的评估、信用和支票费用、以及储蓄/投资行为。若是依照 Amagir(2018)将理财素养定义所分为三个部分：知识和理解、技能和行为、态度和信心而言，行为是一个

十分适宜评量学习成效的面向。总体而言，理财教育的成功与否不仅取决于提供的知识量，更需要融入实际操作和行为培养。

2.4. 自我效能在理财教育中的角色

自我效能是由心理学家 Bandura 于 1977 年提出的概念，指一个个体对于自己能够完成特定任务的信心程度。自我效能信念决定了人们的情感、思考、自我激励和行为(Bandura,1978)在理财教育的背景下，自我效能扮演着关键的角色，对个体在面对金融决策和处理财务事务时的行为和态度产生深远的影响(Mottola, 2013)。

首先，自我效能对于个体参与理财活动的意愿和积极性具有重要的影响(Kurniasari, 2023)且自我效能与理财行为息息相关。其次，拥有较高自我效能的个体更倾向于对挑战采取积极的应对态度，他们相信自己具备克服困难的能力，因此更能应对金融不确定性和风险(Neymotin, 2010)。最后，理财教育中的自我效能培养是提高学生理财素养的重要途径。

3. 研究设计

3.1. 研究对象与实验设计

本研究使用自评问卷与行为资料探究金融理财教育搭配应用实作融入国中课程的影响因素。共征询 100 位有意愿参与研究之八年级学生国中生，研究结束后以约新台币 30 元价值的文具礼品作为研究酬金。其中受试者包括某校 50 位男生、50 位女生。

3.2. 实验流程与情境教学实施

本研究主要由该校 5 位教师进行授课。课程为期 20 周，每周一堂课(45 分钟)，并于后十五周搭配线上模拟理财系统。授课内容以 7 个不同金融理财产品为主题，主题分别为：活存、定存、外币、黄金、股票、房产与乐透。在课程中建立学生的理财知识与概念。金融商品与课堂所学单元对应，其中包括：定存、活存、股票等多项商品。模拟平台学习规范如下：学生于学期初时在线上模拟平台中获得资金共 300 万，可依判断自行买卖金融商品。学生可以于周间在线上模拟平台中任意操作资金投资以上商品。

3.3. 研究工具

理财学习及金融操作自我效能问卷皆以 PISA 于 2017 年编制的理财素养问卷做翻译。金融操作自我效能量表包含两个构念，分别为银行中的金融操作自我效能(共计 6 题)以及使用电子装置执行金融操作的自我效能(共计 5 题)，以李克特氏 4 点量表呈现，受试者依照自己在执行不同任务时所感受到的信心程度，给予 1 分(非常没有信心)到 4 分(非常有信心)的评分。最后将受试者各题得分加总并计算标准分数(*z*-score)指标。而理财学习自我效能量表共有 6 题，为李克特氏五点量表，受试者根据自己平常在学习理财技能与知识的情况，勾选一个最符合的状况，给予 1 分(非常不同意)到 5 分(非常同意)评分。最后将所有题数得分加总并计算标准分数(*z*-score)的理财学习自我效能指标。

3.4. 资料分析方法

3.4.1. k-means 群集分析

于研究问题一中，本研究者使用 *k*-means 群集分析，研究以买进、卖出及浏览之学期总次数作为分组变项将学生分群。再利用 ANOVA 检验各群组之间是否在买进、卖出及浏览次数之间存在显著差异，并根据各组差异为群组命名。

3.4.2. 单因子变异数分析

于研究问题二中，以使用者操作模式为自变项，单次操作金额均值为依变项，以 ANOVA 分析了解学生平均每次在平台中的操作金额是否受使用者因操作模式影响。若统计显著则意味着其中存在差异，则代表买卖与浏览次数的高低也同样对每次买卖金额的均值产生影响。以此确定不同操作模式的使用者在操作金额的差异。

3.4.3. 混合设计二因子变异数分析

为了进一步探讨理财课程实施前后，不同群组同学的理财学习与操作自我效能变化情况。研究问题三以“操作模式”及“时间点(理财课程前、后)”为自变项，并分别以“理财学习自我效能”与“金融操作自我效能”为依变项，进行两次混合设计二因子变异数分析。

4. 结果与讨论

4.1. 学生在理财模拟平台中操作模式的群集分布

为了解理财模拟平台上，学生是否在操作中建立了不同的行为模式，本研究利用整学期的操作次数分为买进、卖出、及浏览三种行为将其标准化后对学生进行分群，原数据描述性统计如表 1 所示。结果依序测试分群数二至五的四种分群方案，并针对不同数量的群集方案进行分析。在此四种分群方案中，分为三群的方案在 Dunn 指数中为最大值，表现最佳；在 SD 指数、PseudoT2 指数以及 Hubert 指数中皆为最小值，同样表现最佳。整体而言，代表群集内部距离近且相似性高。总体而言，如表 2 所示，本研究中许多指标皆表明,群集数为三的方案为一个能在“群集之间的分离性(Separation)”与“群集内部的紧凑性(Compactness)”之间取得良好平衡的方案，因此将使用者分为三群。

已註解 [J1]: 圖分開，放大，R^2

表 1 操作行为描述性统计表

	样本数	平均值	标准差	中位数	最小值	最大值	偏态	峰度
买进	100	805.51	927.36	492	1	5,240	1.82	4.41
卖出	100	15.30	20.07	7	0	112	1.97	4.60
浏览	100	136.23	179.82	66.5	2	799	2.01	3.40

表 2 *k*-means分群评量指标

分群数\指标	R^2	Pseudot2	Dunn	SDindex	Hubert	Silhouette	KL	Marriot
2群	50.79	5.5709	0.084	2.978	0.0061	0.5885	3.797	356276.7
3群	65.22	2.65	0.094	2.432	0.0061	0.5628	1.069	326151.1
4群	73.43	12.01	0.0438	2.608	0.0066	0.5340	3.781	223572.7
5群	69.91	-32.741	0.0434	3.312	0.0067	0.4015	0.196	174796.6

将 100 位受试者分别分到三组中，第一组(G1)、第二组(G2)、第三组(G3)分别包含 14、22、64 人。为验证分组的有效性，以 ANOVA 检验三组(G1、G2、G3)学生在买进、卖出以及浏览三种平台操作行为数是否存在显著差异。三种行为在 Levene 检定中皆拒绝同质性假设($F_{买进}$ = 6.21, p<.001; $F_{卖出}$ = 17.22, p<.001; $F_{浏览}$ = 61.00, p<.001)，因此进行 Welch's ANOVA。结果显示三种行为在三组间皆存在显著差异($F_{买进}$(2, 23.19) = 42.15, p<.001, η^2 =.784; $F_{卖出}$(2, 22.85) = 38.49 , p<.001, η^2 =.771; $F_{浏览}$(2, 23.89) = 43.53, p<.001, η^2 =.785)，且买进、卖出及浏览三种操作行为对组别的解释变异量皆超过 75%。再进一步使用 Games-Howell-Test 进行事后比较，结果显示在 G1 有最高的买进行为，显著高于 G2 与 G3，同时 G2 的买入行为也显著高于 G3(μ_{G1-G2} = 1.27, p<.05; μ_{G1-G3} = 2.22, p<.001; μ_{G2-G3} = 0.95, p<.001)。而 G2 有最高的卖出行为、G1 次之、G3 卖出行为则为最低(μ_{G1-G2} = -1.32, p<.001; μ_{G1-G3} =.64, p<.05; μ_{G2-G3} = 1.96, p<.001)。最后针

对浏览行为，G2 同样有最高的浏览行为、G1 次之、G3 最低(μ_{G1-G2} = -1.63, p<.001; μ_{G1-G3} =.41, p<.05; μ_{G2-G3} = 2.04, p<.001)。

若依照各组特性做总结，三组分别可见：G1 为高买进、中卖出、中浏览的使用者，将其命名为眼明手快组(N=14)；G2 为中买进、高卖出、高浏览的使用者，将其命名为谨慎多虑组(N=22)；G3 为低买进、低卖出、低浏览的使用者，将其命名为漠不关心组(N=64)。

4.2. 不同操作模式的使用者在平台的单次操作金额均值的差异

使用 ANOVA 进一步了解三种使用者在使用平台时，单次操作金额均值的差异。由描述性分析可见受试者的操作金额均值为每笔 86,862 新台币，且最小值为 25、最大值为 503,485 新台币。在 Levene 检定中不拒绝同质性假设(F(2,83) = 2.339, p =.10)。将数值标准化后，ANOVA 结果显示，三组在资产波动总额不存在显著差异(F(2,83) = 0.88, p =.417, η^2 =.02)。

由此可见，虽然不同操作模式的使用者在模拟平台中的操作买卖次数有显著的差异，但每位使用者平均每一笔买卖金额的大小并没有差异，并不会因着买卖与浏览多寡的操作模式差异造成操作金额大小的不同。

4.3. 不同操作模式的群组在理财课程前后自我效能的变化

本研究以三个组别(眼明手快组(G1)、谨慎多虑组(G2)、漠不关心组(G3))及两个时间点(理财课程学期前、后)为自变项，金融操作及理财学习自我效能分别为依变项，检验三个群组在不同时间点中，两种自我效能的变化情况。变项标准化后之描述性统计如表 3 所示。

表 3 自我效能标准化后描述性统计表

		金融操作自我效能前测		金融操作自我效能后测		理财学习自我效能前测		理财學習自我效能後測	
组别	样本数	平均值	标准差	平均值	标准差	平均值	标准差	平均值	标准差
眼明手快組(G1)	22	0.33	0.98	0.27	0.77	0.29	1.16	0.56	0.83
谨慎多虑组(G2)	14	0.24	1.13	0.24	1.10	0.03	1.14	-0.16	1.28
漠不关心组(G2)	64	-0.17	0.95	-0.38	1.03	-0.11	0.90	-0.16	0.93
总和	100	0	1	0	1	0	1	0	1

首先针对金融操作自我效能，变异数同质性假设在 Levene 检定无论在不同操作模式或不同时间点中，皆不拒绝虚无假设($F_{操作模式}$(2,197) = 1.27, p =.284; $F_{时间点}$(1,198) = 1.26, p =.264)，因此可进一步执行变异数分析。混合设计二因子变异数分析结果如表 4。因子间交互作用并未显著($F_{操作模式*时间点}$(2,97) = .10, p =.90)，表示操作模式与时间点两者之间不会相互影响，因此进行主要效果检验。結果顯示三种操作模式中学生的金融操作自我效能并不全相同($F_{操作模式}$(2,97) = 3.11, p<.05, η^2 =.04)，经事后比较得知，眼明手快组(G1)的金融操作自我效能显著高于漠不关心组(G3)，而 G1、G2 两组以及 G2、G3 两组在金融操作自我效能上并无显著差异。(μ_{G1-G3} = 3.82, p<.05; μ_{G1-G2} =.51, p =.96; μ_{G2-G3} = 3.31, p =.15)。而两个时间点的学生的金融操作自我效能恰好边缘显着($F_{时间点}$(1,97) = 3.83, p =.05, η^2 =.01)。依 Turkey 进行主要效果检验，结果显示以两个时间点的金融操作自我效能存在显著差异($\mu_{前测-後测}$ = 1.7, p<.05)，可见随课程进行，学生的金融操作自我效能也有显著提升。

表 4 金融操作自我效能变异数分析摘要表

变异来源	平方和 (*SS*)	自由度 (*df*)	均方 (*MS*)	*F*	*p*	η^2
组间	762	5				
平台操作模式	609	2	305	3.11	<.05*	0.04
时间点	145	1	145	3.83	0.05	
平台操作模式*时间点	8	2	4	0.10	0.90	
组内	13154	194	68			
受试者们	9493	97	98			
残差(误差)	3661	97	38			
全体(总和)	13916	199				

Note: *p<.05, **p < .01, *** p<.001

其次，理财学习自我效能变异数同质性检定结果显示，理财学习自我效能无论在不同操作模式或不同时间点，皆未达统计显著(*Levene's test* $F_{操作模式}(2,197) = 1.50, p = .22$; $F_{时间点}(1,198) = .02, p = .90$)，因此符合变异数同质假设可进一步执行变异数分析。混合设计二因子变异数分析结果显示，因子间交互作用并未显著($F_{操作模式*时间点}(2,97) = 1.65, p = .20$)，表示操作模式与时间点两者之间不会相互影响，因此进行主要效果检验，具体可见表 5 变异数分析摘要表。结果顯示，三种操作模式中学生的理财学习自我效能并不全相同($F_{操作模式}(2,97) = 3.40, p<.05, \eta^2 = .05$)，经事后比较得知，眼明手快组(G1)的理财学习自我效能显著高于漠不关心组(G3)，而眼明手快组(G1)与谨慎多虑组(G2)并无显著差异，同样谨慎多虑组(G2)与漠不关心组(G3)也未达统计显著($\mu_{G1-G3} = 4.22, p<.01$; $\mu_{G1-G2} = 3.70, p = .12$; $\mu_{G2-G3} = .51, p = .94$)。而两个时间点的学生的理财学习自我效能并未显著($F_{时间点}(1,97) = 1.71, p = .19, \eta^2 = .003$)，表示受试者的理财学习自我效能在课程学习前与学习后均值并无差异。

表 5 理财学习自我效能变异数分析摘要表

变异来源	平方和 (*SS*)	自由度 (*df*)	均方 (*MS*)	*F*	*p*	η^2
组间	696	5				
平台操作模式	590	2	295	3.40	<.05*	0.05
时间点	36	1	36	1.71	0.19	
平台操作模式*时间点	70	2	35	1.65	0.20	
组内	10469	194	54			
受试者们	8422	97	87			
残差(误差)	2047	97	21			
全体(总和)	11165	199				

Note: *p<.05, **p < .01, *** p<.001

由以上两种自我效能的分析结果可知，不同操作模式的使用者在理财学习或金融操作的自我效能都有所差异。于高买进中卖出中浏览的眼明手快组同学无论在理财学习或金融操作的

自我效能皆高于的漠不关心组的同学。而相对的，谨慎多虑组的同学拥有较高的浏览行为、在做购买决定时也较为谨慎，但其无论在理财学习或金融操作的自我效能上同时都与眼明手快组没有显著差异，也与漠不关心组没有显著差异。可见“多看多收集资讯”的浏览行为，在理财学习中并不会产生较高的金融操作自我效能，反倒是“多下单多尝试”的买进行为能对学生的金融操作自我效能产生正向影响。

就量测时间点而言，搭配金融模拟平台的理财课程情境画教学，能对学生的金融操作自我效能有显著提升。随着课程的进行与平台的运作训练，学生皆能在课程结束后有更强的金融操作自我效能。但理财学习自我效能则不同，结果显示课程前后学生的理财学习自我效能并没有显著差异。若是关注均值图，则发现虽然交互作用并不显著，但论均值不同组别的理财学习效能也有不同方向的改变。眼明手快组课程前测就拥有最高的自我效能，且在课程后有正向成长。但谨慎多虑组与漠不关心组前测时拥有较低的自我效能，且在后测均值中显示有所降低。另外分析金融操作自我效能时，时间在 ANOVA 中并不显著到主要效果显著的状况，研究者认为原因在于，ANOVA 主要关注因子的整体效应，而 Turkey 较关注具体组别之间的差异，因其更加敏感因而发现了在时间点上的显著差异。

5. 总结

本研究旨在理解搭配模拟平台教学的理财素养课堂中学生的行为操作模式对其金融操作自我效能和理财学习自我效能的影响。研究将学生于平台中的买进、卖出与浏览行为进行分群以了解操作模式。结果显示，学生可分为眼明手快组、谨慎多虑组、漠不关心组。尽管这三组在单笔操作金额上没有显著差异，显示虽然操作模式影响了操作次数，但并没有对单次操作金额产生显著影响。但眼明手快组在金融操作和理财学习两种自我效能上均显著高于漠不关心组，谨慎多虑组则与其他两组皆显著差异。另外，随着理财课程的进行，学生的金融操作自我效能也显著提升，但对理财学习自我效能则没有显著影响。在教学中，建议老师可多多搭配实践活动并鼓励学生「无论对错，多方尝试」，以此建立理财学习中的自我效能。

综上所述，尽管不同操作模式的学生在模拟平台中的行为和自我效能方面存在差异，但总体而言，搭配金融模拟平台的理财课程有助于提升学生的金融操作自我效能。未来也希望能更进一步分析使用者于模拟平台中的行为改变，并从学生在绩效上的表现、理财素养成绩等面向对此研究有更深入的分析。

6. 参考文献

已註解 [J2]: 我抓了一些錯，請檢查所有文獻 APA

Authors, 2021.

Amagir, A., Groot, W., Maassen van den Brink, H., & Wilschut, A. (2018). A review of financial-literacy education programs for children and adolescents. *Citizenship, Social and Economics Education, 17*(1)*,* 56-80.

Atkinson, A., & Messy, F. (2012). *Measuring financial literacy: Results of the OECD/International Network on Financial Education (INFE) Pilot Study* (OECD Working Papers on Finance, Insurance and Private Pensions, No. 15). Paris: OECD Publishing.

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy, 1*(4), 139–161.

Boshara R, Gannon J, Mandell L, Phillips JWR, Sass S (2010) *Consumer trends in the public, private, and nonprofit sector*. Denver: National Endowment for Financial Education.

Collins JM and Odders-White E (2015) A framework for developing and testing financial capability education programs targeted to elementary schools. *The Journal of Economic Education 46*(1):

105–120.

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science, 60(8)*, 1861–1883. doi:10.1287/mnsc.2013.1849

Gao, L. (2023, October). Literature Review for the Design Methods of Serious Games. In *Proceedings of the 17th European Conference on Game-Based Learning: ECGBL 2023*. Academic Conferences and publishing limited.

Goel, I., & Khanna, S. R. (2014). Financial education as tool to achieve financial literacy. *ZENITH International Journal of Multidisciplinary Research, 4(12),* 73-81.

Gregory, J. (2002). Principles of experiential education. In Jarvis, P. (Ed.), *The theory and practice of teaching (pp. 114-129).* Sterling, VA: Stylus Publishing, Inc.

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013).Financial literacy, financial education, and economic outcomes. *Annual Review of Economics, 5(1),* 347–373.

Kurniasari, I., & Ratnawati, K. (2023). The role of self-efficacy in mediating parental financial teaching on financial management behavior: a study on youths in Malang. *International Journal of Research in Business and Social Science (2147-4478), 12(4),* 422-433.

Mottola, G. R. (2013). In our best interest: Women, financial literacy, and credit card behavior. *Numeracy, 6(2)*, 1–15.

Neymotin, F. (2010). Linking self-esteem with the tendency to engage in financial planning. *Journal of Economic Psychology, 31(6),* 996–1007. https://doi.org/10.1016/j.joep.2010.08.006

OECD (2005). *Improving Financial Literacy: Analysis of Issues and Policies.* Paris: OECD

OECD (2015). National Strategies for Financial Education. OECD/INFE Policy Handbook. Paris: OECD .

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games—An overview* (Technical Report No. HS- IKI -TR-07-001). Skövde, Sweden: School of Humanities and Informatics, University of Skövde

Tang, J. S. Y., Falkmer, M., Chen, N. T. M., Bölte, S. & Girdler, S. (2019). Designing a Serious Game for Youth with ASD: Perspectives from End-Users and Professionals. *Journal of Autism and developmental disorders, 49*, 978-995.

教育游戏嵌入发音对中职学生拼写及识别能力的影响研究

A Study on the Impact of Pronunciation within Educational Games on the Competence of Spelling and Recognition for Secondary Vocational School Students

高志达[1*]，徐心远[2]，乔建澍[1]
[1]伦敦玛丽女王大学
[2]江阴中等专业学校
[*]December971230@hotmail.com

【摘要】 词汇学习是英语教学中的关键部分，中职院校的学生通常学习单词的动力不足。本研究旨在探索：（1）教育游戏是否可以促进学生的积极性；（2）教育游戏中的发音对拼写和识别能力的影响。本研究采用定量分析和问卷调查的形式进行对比试验，研究结果表明：（1）教育游戏可以促进学生拼写和识别能力；（2）教育游戏嵌入发音可以促进学生的拼写；（3）学生主观上更喜欢以教育游戏的方式学习单词，并认为这种方式较为高效。

【关键词】 英语词汇；中职教育；教育游戏

Abstract: *Vocabulary is a vital part of English education, however, vocational students usually lack the motivation to learn words. This research aims to explore (1) whether educational video games (EVGs) can promote students' motivation; and (2) the impact of pronunciation within EVGs on the competence of spelling and recognition. The difference between the experimental group and the control group is that the former has pronunciation. Utilizing the quantitative analysis and questionnaire method, the research reveals that (1) EVGs can facilitate the competence of spelling and recognition; (2) pronunciation within EVGs can promote students' spelling; and (3) students objectively like learn words in the EVGs fashion, and think it is an effectiveness way.*

Keywords: English vocabulary, vocational education, educational video games

1. 前言

随着科学技术的发展，人们学习英语的方式也随之变化。2022 年版的《义务教育英语课程标准》不仅对计算机和数字技术应用于教与学提出了要求，也强调教师要引导学生乐学善学。教育游戏凭借其通俗性、易玩性等特点，逐渐成为教育教学的工具之一。在英语教育中，教育游戏已经被证明能够有效提高学生的学习动力和教学效果。

当前，中职学生在英语学习方面面临兴趣不高和缺乏动力等挑战（李芝波，2022）。传统教学方法与学生的实际需求脱节，缺乏互动性和趣味性，导致学生难以有效吸收和应用所学知识（胡晓玲等人，2021）。教育游戏作为一种创新的教学工具，具有以下三点优势：（1）提供富有吸引力的学习环境：将英语知识与教育游戏融合，能够显著提高学生的学习兴趣和参与度；（2）模拟真实情境：让学生在游戏中实际运用英语进行交流，从而增强学习的实践性和应用性（王春艳等人，2019）；（3）提供即时反馈：在教育游戏中设置即时反馈，有助于加强学生的记忆力，同时提高视觉和听觉满足感，从而激发他们对英语学习的热情。长期来看，这种积极的学习态度和高涨的学习动力将进一步提高学生学习效率，形成良性循环，促进终身学习。

综上所述，教育游戏在中职英语教育中的应用具有重要的教育价值和社会意义。通过将教育游戏嵌入教室中，探索以教育游戏为主导的教学方法，有望解决当前中职英语教育难题。因此本研究以中职英语单词教育为切入点，从 5 个部分展开讨论，第一部分前言介绍教育游戏的优势；第二部分为文献综述，主要总结国内外教育游戏研究相关现状；第三部分为研究设计，主要介绍了研究问题、研究对象、研究工具、实验流程；第四部分为结果和讨论，主要对实验结果进行呈现和讨论；第五部分为结论，主要总结了本研究的结论，提出研究不足及对未来的展望。

2. 文献综述

教育游戏不断深入发展，并在多个学科领域展现出广泛的应用潜力。教育游戏的历史可追溯至 20 世纪 50 年代，主要以军事模拟和策略游戏形式出现（Demirbilek, 2009）。随着个人计算机和互联网的普及，教育游戏迅速发展并成为一种重要的教育工具（Herz, 2001）。Kim (2015) 探讨了将游戏作为学校课程活动的一部分的教育游戏开发方法。Ibrahim 等人（2019）提出了一种教育游戏质量框架，用于指导教育游戏的设计和评估，以确保游戏在教学中的有效性。 此外，Santanna Dos Santos 等人（2020）通过研究艺术历史教学中的棋盘游戏，探讨了教育游戏在增强学生学习动机方面的潜力。Hiriart（2020）通过开发基于历史的视频游戏，分析了教育游戏对历史教育的潜在贡献。Nugroho 等人（2020）的研究则侧重于利用教育游戏提升印尼历史课程的学习效果。

教育游戏在不同教育场景中的应用显示了其多样性和灵活性。Hyun 和 Wonhyong（2020）的研究强调了将游戏集成到历史教学中的重要性，指出教育游戏可以提升学生对历史的兴趣和理解。Goi（2021）回顾了商业模拟游戏的发展历程，指出其在教育和培训中的应用价值。Gampell 等人（2020）以新西兰四所学校为例，探讨了“严肃”灾难教育游戏作为地理学学生学习工具的潜力，发现教育游戏需要进一步考虑教学和学习过程。Sunarya（2022）总结了现阶段人工智能和机器学习嵌入各种不同类型的教育游戏产生全新教学内容的发展趋势，并对已经应用于 K-12 教育的部分教育游戏进行探索，结果表明，教育游戏对于 K-12 阶段教育拥有巨大的应用潜力。对于教育游戏在英语教学中的应用，Ageenko 等人（2021）通过改编后的 I.A. Zimbjaja 游戏对 10 名小学生进行英语教学，结果表明小学阶段使用教学游戏进行英语教学可以增强学生的积极性，并对学习、培训和修订教材产生积极影响。Ming 等（2021）通过快速证据评估回顾，探索了教育游戏在大学英语教学和学习过程中的整合，发现教育游戏在提高学生的大学英语能力和学习态度方面有效。他们的研究集中于教育游戏设计的重要性，但未对教育游戏在教室环境下英语发音学习的重要性进行探讨。Yichen 等人（2021）通过使用聊天机器人 AI 和语音转文字技术，创建了一个教育视频游戏，让玩家可以与 NPC 进行真实的口语对话，并在可能遇到的真实场景中与他们互动，从而提高英语口语能力。Tejedor-García 等人（2020）设计了一款新型学习游戏，用于第二语言发音训练，游戏中玩家可以相互挑战，通过移动应用进行多项发音活动，结果显示参与者通过大量活动和规律性练习显著提高了发音水平，证明了教育游戏在发音训练中的有效性。 Octaberlina 和 Rofiki（2021），通过对 22 名印尼学生完成 SpellingCity 游戏的反馈，得出教育游戏能够提高学生学习效率的结论，但是对于英语单词的发音对学生效率的影响，未做进一步研究。

综上所述，尽管教育游戏在多个领域得到了广泛应用，但发音在教室场景下的教育游戏中的效用尚未得到充分探索。目前的研究主要集中在游戏的设计、学习动机和学习成果上，而对于发音在教育游戏中的作用的研究相对较少。因此，本研究旨在填补这一研究空白，探索在教育游戏中融入发音对于提高中职学生英语词汇学习效率的影响。

3. 研究设计

本章对研究设计进行了全面介绍，包括研究问题、研究对象、研究工具、实验流程四方面。

3.1. 研究问题

为了量化研究蛇游戏中添加发音对于学生识记单词的效果的作用，将蛇游戏应用于中职英语课堂教育中，研究其对中职学生学习单词的影响，提出了以下问题：

1.蛇游戏在拼写和识别方面能否帮助学生记忆单词？

2.蛇游戏中嵌入发音能否提升识别和拼写的效果？

3.蛇游戏能否提高学生的学习兴趣？

基于此研究问题，本研究做出以下假设：

假设 1：蛇游戏可以帮助学生提高拼写和识别能力。

假设 2：蛇游戏中添加发音能提高学生的识别和拼写能力。

假设 3：蛇游戏可以提高学生的学习兴趣。

3.2. 研究对象

参与实验的学生为某中等专业学校某个班级的学生。实验在学校电脑教室中进行。为保证所选实验组及对照组学生学习水平较为相似，本研究对一个班的学生按照成绩排名，采取间隔分组。选取样本总量为 34 名学生，其中实验组 17 名，在蛇游戏中提供强制和选择发音两种方式；对照组 17 名，在蛇游戏中不提供发音。

为保证所学习单词均为新单词，本研究首先挑选 30 个单词，并让学生按照中文写出英文拼写，最后选择这 34 名学生均不会的 10 个单词进行实验。

3.3. 研究工具

3.3.1. 蛇游戏

蛇游戏是此次实验的载体游戏，使用 python 语言编写。游戏的基本信息是，开始游戏后，区域 1 包括当前任务、音标、拼写追踪；区域 2 包括得分和剩余任务信息；区域 3 为基本操作的指引；区域 4 是主游戏区域，两个黄色的圆圈是蛇头，区域空间的其他两个字母分别为正确字母及迷惑字母。每次吃到正确字母加 10 分并提供积极的音效反馈，吃到错误字母减 10 分并提供消极的音效反馈。区域 5 为提示区域。如果学生无法完成任务，可以选择吃红色十字架来获得拼写提示，提示展示 5 秒后消失。

蛇游戏的基本规则是：学生首先根据区域 1 明确任务，在区域 4 通过上下左右操作拼写单词，每吃一个字母就会加入到蛇的尾部，吃到错误字母不会添加，每次完成一个任务以后，蛇头后面的字母会消失。当完成所有的任务之后，学生可以根据图 2 的反馈加强记忆并反思自己错误，并且可以返回主菜单来继续游戏。拼写正确的单词会移出任务组，拼写错的单词会保留在任务组，游戏完全结束的条件是学生可以正确拼写所有的单词，即任务组中再也没有单词，整个游戏时长上限为 20 分钟。

蛇游戏设计的特点是：首先，蛇游戏通过逐一吃掉字母以组成正确单词的游戏形式，使学生能够将单词在脑海中主动复现，加强对新单词的记忆；其次，蛇游戏通过提示单词发音的方式加强了学生“音、形”结合的能力；再次，蛇游戏能够对学习过程中拼写错误的单词进行重复训练，增强单词学习的针对性；最后，由于蛇游戏的轻量级设计，可以离线使用，对学校的硬件设施要求较低，可以广泛适配教室环境。

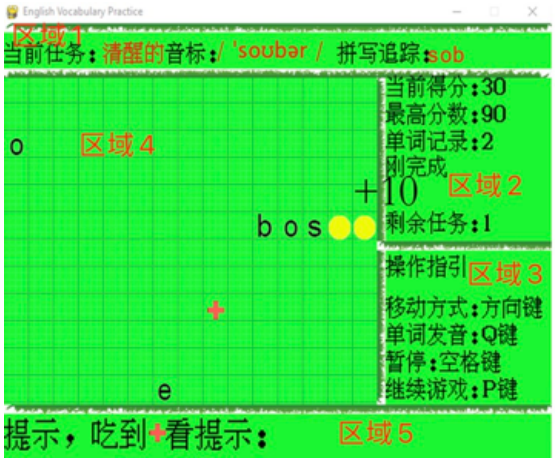


图 1 游戏界面



图 2 反馈界面

3.3.2. 拼写及识别测试

拼写测试（见附录）通过提供中文和音标让学生拼写，学生需要根据所给信息拼写完整的单词，评分标准均为正确得 1 分错误为 0 分。该测试的目的是为了研究蛇游戏是否能够提高学生的拼写能力。

识别测试（见附录）通过单项选择题，让学生根据汉语从四个选项中选择正确的英文单词，评分标准均为正确得 1 分错误为 0 分。该测试的目的是研究蛇游戏是否能够提高学生的识别能力。

3.3.3. 问卷调查

在实验结束后对所有参与实验的学生进行问卷调查（见附录），旨在了解教育游戏对学生英语词汇学习动机及兴趣的影响。

3.4. 实验流程

2023 年 12 月 14 日在学校中进行实验。研究过程分为四个阶段，分别为：实验准备、实验实施、能力测试、数据收集。

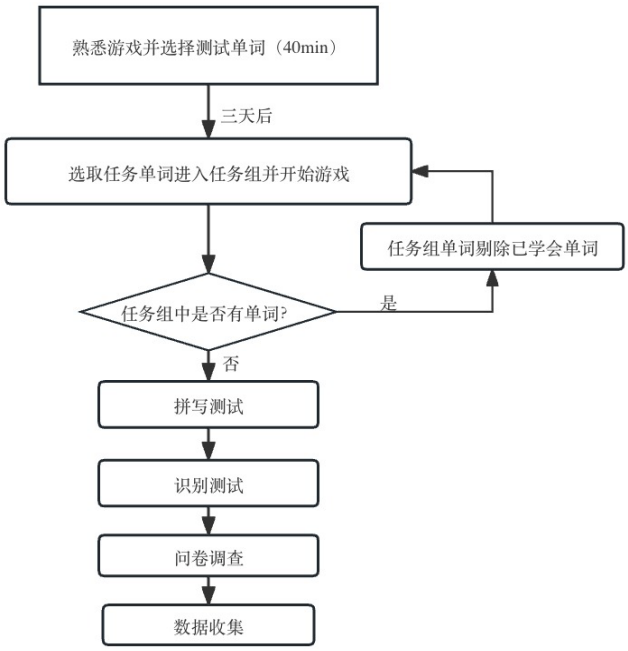


图 3 实验流程图

首先，研究者通过一节课的时间（40 分钟）让学生熟悉整个游戏的操作，此目的是避免学生在实验过程中因为操作不熟练导致已经记住的单词无法有效操作蛇拼写正确，进而影响实验数据统计；随后，通过学生对单词的了解程度寻找实验组与对照组学生均不会的 10 个单词进行蛇游戏实验。

为避免学生在单词选择过程中形成的短期记忆对实验结果的影响，研究者在 3 天后进行正式实验：学生首先使用蛇游戏记忆单词，时长为 20 分钟，待实验结束后先进行拼写测试（8 分钟），待拼写测试完成后，进行识别测试（8 分钟）；最后，向所有参与实验的学生发放问卷调查，并收集结果进行分析。正式实验及测试的总时间为一节课（40 分钟）。

4. 结果与讨论

本章主要通过对测试结果及问卷调查数据进行结果呈现。并对统计结果使用 SPSS 进行单因素 ANOVA 分析，旨在探索蛇游戏对学生在拼写和识别方面的影响，以及蛇游戏中的发音能否促进拼写和识别的效果。

4.1. 拼写测试

参与者在使用蛇游戏记忆单词的平均数为 9.56，而在实验之前学生不认识所有的单词，可以证明蛇游戏可以帮助学生记忆单词拼写。通过对实验组和对照组的统计量分析，可以发现，实验组的平均数 9.77 高于对照组 9.35，实验组的标准差 0.44 小于对照组的 0.70，说明实验组的平均成绩要高于对照组，且个体差异较少。为了确认实验后两组学生的均值有无显著差异，对两组学生的数据进行了单因素 ANOVA 检验。两组学生的 Sig=0.048<0.05，在 95%的置信水平上有显著的统计意义。即两组学生的拼写效果出现显著差异。由此可见，在教育游戏中添加发音能够促进学生的拼写能力。

表 1: 拼写表格

	个案数	平均分	标准差	最小值	最大值	F	Sig.
无发音	17	9.35	0.70	8	10	4.220	0.048
有发音	17	9.77	0.44	9	10		
总数	34	9.56	0.61	8	10		

4.2. 识别测试

参与者在使用蛇游戏记忆单词的平均数为 9.79，而在实验之前学生不认识所有的单词，可以证明蛇游戏可以帮助学生提高识别能力。通过对实验组和对照组的统计量分析，可以发现，实验组的平均数 9.94 高于对照组 9.65，实验组的标准差 0.24 小于对照组的 1.22，说明实验组的平均成绩要高于对照组，且个体差异较少。为了确认实验前后两组学生的均值有无显著差异，对两组学生的数据进行了单因素 ANOVA 检验。两组学生的 Sig=0.338>0.05，在 95%的置信水平上没有显著的差异。即由此可见，在教育游戏中添加发音对学生的识别能力没有太大的影响。

表 2: 识别表格

	个案数	平均分	标准差	最小值	最大值	F	Sig.
无发音	17	9.65	1.22	5	10	0.948	0.338
有发音	17	9.94	0.24	9	10		
总数	34	9.79	0.88	5	10		

根据假设 1：蛇游戏可以帮助学生提高拼写和识别能力。通过拼写测试和识别测试发现，学生在不认识任何单词的情况下，在 20 分钟的学习之后的识别和拼写测试的平均分分别为 9.56 和 9.79，假设 1 得到验证。蛇游戏可以提高学生的拼写和识别能力的主要原因是，通过吃字母的方式来记忆单词比较有趣，这可以通过参与者的反馈得到验证，例如，“喜欢因为比较有趣”、“形式新，可以边玩游戏边学习”、“比日常背单词软件好玩”，所以蛇游戏较为容易抓住学生的注意力。学生在游戏过程中通过操作蛇头吃到正确或错误字母所提供的实时音效反馈及区域 4 的加减分反馈，可以提升学生的成绩，这也符合 Hanson-smith（2016）对于教育游戏中添加得分和点数等反馈机制有助于学生获得更高成绩的结论。因此，通过提供奖励及视觉和听觉反馈可以促进学生的学习。在一轮游戏结束以后，游戏会展示正确拼写和学生拼写，可以帮助学生修正自己的记忆。这种学习-测试的组合与 Batsell（2017）等人对于测试效应在课文阅读情景下的研究结论相符。所以蛇游戏可以提高学生拼写和识别能力。

根据假设 2：蛇游戏中添加发音能提高学生的识别和拼写能力。通过表 1 和表 2 可以发现，实验组和对照组在拼写能力上存在显著差异，在识别能力上不存在显著差异。实验组和对照组的差别为游戏有无发音。在拼写和识别上表达出不同的效果，原因可能是因为拼写测试没有提供正确的英文单词，所以大脑需要检索更多的信息来拼写出正确的单词。在有发音的实验组接受发音，所以对照组的拼写成绩高于实验组。Futami（2022）等人的实验证明，通过提供更多的信息可以提高学生记忆单词的能力得到验证。这也可以通过学生的建议得到验证，例如，“最好有由来和句子示例”、“有图片最好”，而识别测试提供了正确的单词，所以学生主要是匹配汉语和英语的信息，大脑不需要检索发音信息，所以在识别能力方面没有表现出显著性差异，尽管如此，实验组的学生的识别平均成绩还是高于对照组。

表 3: 问卷调查结果

	非常不喜欢	不喜欢	没感觉	喜欢	非常喜欢
你喜欢这种背单词的方式吗？	1	2	7	18	6
	非常低	低效率	没感觉	高效率	非常高效率
你觉得这样背单词的效率会更高吗？	0	3	8	20	3

根据假设 3：蛇游戏可以提高学生的学习兴趣。表 3 显示大多数学生对蛇游戏持正面态度，这反映了蛇游戏在设计上的吸引力和互动性。游戏作为一种动态和参与性的学习媒介，能够激发学生的学习兴趣，成为促进学习动机的重要因素，假设 3 得到验证。

具体来说，“非常喜欢”与“喜欢”的比例高（18 和 6），与蛇游戏满足了学生在视觉、听觉和互动体验上的需求有关。这样的设计不但增加了学习的乐趣，而且提高了信息的记忆和处理能力，因为它们通过多感官的方式呈现，与传统的教学方法相比，更容易被学生接受和记忆。

在效率方面，高比例的“非常高效率”与“高效率”反应（3 和 20）表明学生主观上认为蛇游戏促进了学习效果。当学生对学习内容感到兴奋和参与时，他们更有可能在认知上投入更多精力，从而提高了学习效率。此外，蛇游戏提供的即时反馈和奖励系统增强了学习的强化机制，这在教育心理学中是促进学习的已知因素。综上所述，蛇游戏能够提高学生的学习兴趣。

然而，对于那些“不喜欢”和“没效果”的学生（7 和 3），这可能表明蛇游戏并不是所有学生都能接受的学习工具。对这部分学生来说，蛇游戏可能无法引起他们的兴趣或不符合他们的学习风格。例如，一些学生可能更喜欢传统的、结构化的学习环境，或者他们可能觉得游戏分散了他们对学习目标的注意力。

5. 结论和启示

本章总结研究结论并提出研究的不足与展望。

5.1. 研究结论

本研究以蛇游戏为载体，采用了拼写测试和识别测试的定量分析方法和问卷调查的定性分析方法，旨在讨论教育游戏对学生拼写和识别能力的影响以及游戏中发音对学习效果的影响。

首先，通过表 1 和表 2 可以得出，实验后学生的拼写测试和识别测试的平均数分别为 9.56 和 9.79，接近于满分（10），所以教育游戏可以帮助学生提高其的拼写和识别能力。其次，在教育游戏中提供发音可以提高学生的拼写能力。在拼写方面，实验组的平均成绩要高于对照组，且个体差异较少。单因素 ANOVA 检验得出两组学生的 Sig=0.048<0.05，在 95%的置信水平上有显著的统计意义，即在拼写方面实验组要大于对照组。在识别方面，实验组的平均成绩要高于对照组，且个体差异较少。单因素 ANOVA 检验表明两组学生的 Sig=0.338>0.05，在 95%的置信水平上没有显著的差异，由此可见，在教育游戏中添加发音对学生的识别能力影响较小。通过表 3 可以得出，学生对教育游戏的主观喜爱程度和教育游戏的效率存在一定关联，即学生喜欢教育游戏（24）主观上也会认为教育游戏能够提高学习效率（23）。

5.2. 不足与展望

本研究的对象是某个地区的某职业中学的一个班级，并没有对不同地区、不同年级的学生进行比较。因此学生之间的个体差异和教育水平等变量会影响实验结果的准确度。本研究对学生进行短期记忆测试，没有对长期记忆的效果进行探索。本研究测试了学生的拼写和识别能力，没有测试学生的发音能力。因此未来的研究希望可以对不同地区的不同学生的进行探索，设计更多的测试方式并从长期记忆和短期记忆多种评价维度探索教育游戏在课堂教育中运用的可行性。虽然科学技术和硬件设备在不断发展，但是关于教育游戏在教室中的运用受到各种方面的限制，比如说教室硬件水平、学生和老师的计算机能力、课程设计、每节课的时间及课程安排等。因此希望未来的研究能够真正开发出能够完全嵌入到学校的教学规划中，便于学生及老师操作与应用，最终在全国各个地区实现寓教于乐，促进学生的终生化学习，实现教育高质量发展的目标。

参考文献

王春艳, 甘甜, 吴倩莲, 王昱霖 & 高伟 (2019) 。基于 Unity 3D 的 VR 英语教育游戏的设计与开发。*计算机时代*, 10, 74-77。

李芝波 (2022) 。中职英语教学单词记忆法探究。*职业教育*, 9, 67-69。

胡晓玲, 赵凌霞, 李丹 & 范博 (2021) 。游戏化教学有效性的系统评价与元分析。*开放教育研究*, 2, 69-79。

Ageenko, N., Rybkina, A., & Startseva, N. (2021). Motivation fostering through gaming technologies. *SHS Web of Conferences, 113,* 00025.

Batsell, W., Perry, J., Hanley, E., & Hostetter, A. (2017). Ecological Validity of the Testing Effect. *Teaching of Psychology, 44,* 18-23.

Demirbilek, M. (2009). The Use of Electronic Games in Distance Learning as a Tool for Teaching and Learning. *In Encyclopedia of Distance Learning*, 2209-2223.

Futami, K., Kawahigashi, D., & Murao, K. (2022). Mindless Memorization Booster: A Method to Influence Memorization Power Using Attention Induction Phenomena Caused by Visual Interface Modulation and Its Application to Memorization Support for English Vocabulary Learning. *Electronics, 11*(14), 2276.

Gampell, A., Gaillard, J., Parsons, M., & Dé, L. (2020). 'Serious' Disaster Video Games: An Innovative Approach to Teaching and Learning about Disasters and Disaster Risk Reduction. *Journal of Geography, 119,* 159-170.

Hanson-Smith, E. (2016). Games, Gaming, and Gamification: Some Aspects of Motivation. *TESOL Journal, 7,* 227-232.

Herz, J. C. (2001). *The history of video games.*

Hiriart, J. (2020). How to be a 'Good' Anglo-Saxon: Designing and Using Historical Video Games in Primary Schools.

Hyun, L., & Wonhyong, L. (2020). A Study on the educational application of historical game.

Ibrahim, R., Yusoff, R., Zainudin, N. M. M., Sjarif, N. N., & Yahya, Y. (2019). Educational Games Quality Framework for Learning Islamic History in Primary School.

Kim, N. (2015). Games Development Methodology as a Teaching Tool for Elementary School: Case Study of Developing History Learning Game. *Journal of Korea Game Society, 15*(2), 53-62.

Li, M., Zhang, Q., & Yu, Z. (2021). Integrating Educational Games into College English Learning. *Education Research International,* 1-9.

Liu, Y., Sahagun, J., & Sun, Y. (2021). An Adaptive and Interactive Educational Game Platform for English Learning Enhancement using AI and Chatbot Techniques. *Natural Language Processing*.

Nugroho, S. M. S., Rahman, M. D. A., Hariadi, M., & Arief, M. (2020). Development of Educational Game for Indonesian History Lesson Focusing on Indonesia Pre-Literacy Relics "Prehistoria" for Grade X High School Student. *In 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia*, 149-154.

Octaberlina, L., & Rofiki, I. (2021). Using Online Game for Indonesian EFL Learners to Enrich Vocabulary. *Int. J. Interact. Mob. Technol., 15,* 168-183.

Santos, M. S. D., Schmitt, M. A. R., Peres, A., & Reategui, E. (2020). Conquer The Artwork: A Board Game for Teaching History of Art. *Educational Review.*

Sunarya, P. (2022). Machine Learning and Artificial Intelligence as Educational Games. *International Transactions on Artificial Intelligence, 1*(1), 129-138.

Tejedor-García, C., Escudero-Mancebo, D., Cardeñoso-Payo, V., & González-Ferreras, C. (2020) Using Challenges to Enhance a Learning Game for Pronunciation Training of English as a Second Language, *IEEE Access, 8,* 74250-74266.

Ya, L. (2020). An Analysis of the Application of Game Teaching Method in Oral English Class of Preschool Education. *Frontiers in Educational Research, 3*(9), 102-104.

附录

网盘链接: https://pan.baidu.com/s/1LBF8uGgcncE5PBsk_Xrz1A　　提取码: r62f

学习动机如何影响国际中文教育游戏化学习投入？——基于滞后行为序列分析的实证研究

How does motivation affect gamified learning engagement in international Chinese Education? -- Empirical research based on lag sequential analysis

冯淼[1]，方璐[2*]，杨国法[3]，王抒诣[4]，卫紫艳[5]，尚品言[6]，张露[7]
1.福建师范大学 海外教育学院；
2.浙江大学 教育学院；
3.内蒙古师范大学 国际交流学院；
4.宁波职业技术学院 国际商旅学院；
5. 西京学院 人文与教育学院
6.云南大学 昌新国际艺术学院
7.北京邮电大学 人文学院；
[*] fanglu_0707@163.com

【摘要】 学习投入作为影响教育游戏学习表现的关键因素，成为当前的研究热点之一。然而，多数研究仅对结果性数据进行分析，而对过程性数据缺乏关注。为深入了解学习动机对学习投入的影响，本研究运用滞后行为序列分析法，对国际中文教育游戏中不同学习动机强度学习者的行为特征与模式进行分析。研究发现，高学习动机学习者表现出更丰富、更高学习投入度的行为，而低学习动机学习者在学习过程中行为模式相对单一。本研究能为理解学习动机与学习投入之间的关系提供思路，并为后续教育游戏的改进提供理论支持。
【关键词】 学习动机；学习投入；游戏化学习；滞后行为序列分析；

Abstract: *Engagement is the key factor affecting learning performance in educational games and has become one of the research hotspots. However, most studies only analyze the outcome data and pay little attention to the process data. To deeply understand the influence of learning motivation on learning engagement, this study used lag sequential analysis to analyze the behavioral characteristics and patterns of learners with different learning motivation intensities. It is found that students with high motivation will show richer and more engaged behaviors, while students with low motivation have relatively simple behavior patterns. This study can provide ideas for understanding the relationship between motivation and engagement and provide theoretical support for the improvement of educational games.*
Keywords: Learning motivation, Learning engagement, game-based learning, lag sequential analysis

1.引言

科学技术的发展推动了学习形式的变革，教育游戏作为一类新兴的技术支持学习形式，受到学界的广泛关注。多项实证研究显示，使用教育游戏有助于提高学习效率，其“最主要的价值在于激发和维持学习者的学习动机”(尚俊杰 & 裴蕾丝, 2015)。自我决定理论指出，学习投入与学习动机的激发相关，“增强学习动机有助于实现高质量学习投入”(Ryan & Deci, 2000)。相关研究证明，“学习动机可以有效预测学习过程中的学习投入程度”(Ninaus, Moeller,

McMullen, & Kiili, 2017)及“问题解决效果”(Eseryel, Law, Ifenthaler, Ge1, & Miller, 2014)。认知投入作为学习投入的重要组成部分，指的是“学生在学习过程中的认知努力，表现在过程性行为的正确率和结果性认知结果中”(张露，胡若楠，雷悦,& 尚俊杰,2023)。然而，当前研究主要对问卷、量表等结果性学习数据进行分析，而对过程性学习行为缺少关注。

故此，本研究运用滞后行为序列分析法，对国际中文教育游戏过程性学习行为数据进行分析，旨在探究学习动机对游戏化认知层面学习投入的影响，包括以下两个研究问题：不同学习动机强度的学习者在游戏学习行为频次上有何差异？不同学习动机强度的学习者在游戏学习行为模式上有何差异？

2.理论基础

2.1. 具身认知理论

具身认知理论强调身体认知过程中发挥的重要作用。狭义的具身认知理论主张身体的生理结构、活动方式、感觉和运动体验等均会影响人们的认知过程，认为身体及感觉在认知过程中具有核心作用；而广义的具身认知理论则主张认知来源于身体与环境的相互作用，不仅强调了身体在认知活动中的作用，还强调其“与环境的关系”(Cowart, 2015)。本研究所涉及的游戏遵循具身认知理论，确保学习者在线虚拟情境中学习交互的多样化，调动学生多感官参与学习，使其在广阔的游戏场景中自由探索。

2.2. 情境学习理论

Falk 和 Dierking 指出，学习是一个高度个人化的过程。学习不但依赖于先前的经验，还依赖于特定的社会文化和物理环境。同时，学习涉及多种经验和信息来源，是“多类情境共同促进知识构建的结果”(Falk & Dierking, 2000)。情境学习理论认为，学习是“合法的边缘性参与”的过程，学习者在学习共同体内通过对专家技能的模仿、互动和再创造，逐渐从一个新手变成熟手，最终掌握技能来获得知识(Lave & Wenger, 1991)。建构真实的语言运用情境，能帮助学习者“提高语言交流与应用能力，发展高阶思维”(Ozverir, Osam, & Herrington, 2017)。以情境理论为指导，本研究涉及的教育游戏高度调情境的真实性，通过在虚拟环境中模拟真实的语言使用场景，学习者可以更好地理解和运用汉语。例如，当学习者置身虚拟购物商场时需要完成“购买野餐清单上的物品”这一任务。为此，学生首先需要根据个人兴趣和中文标注的商品信息选择商品，随后与虚拟商家交流价格、砍价，最终完成指定物品的购买。

2.3. 多媒体认知理论

多媒体学习原则表明，学习者在游戏中通过“独立的双通道处理视觉与听觉信息”(Mayer, 2002)。除却视觉内容的设计外，还需关注游戏内声音信息的设计。本游戏采用了 Microsoft Speech API 的智能语音合成技术，为每一位不同的游戏角色生成真实生动的机器配音。同时，每个游戏场景内都设置了适合的背景音乐，以增强游戏的沉浸感。例如，商场场景以人群走动声和讲价声为背景音乐，以模拟真实的商场环境。此外，学习者在进行移动和交谈时拥有对应的动作音效，作为游戏对学习者行为的反馈。

3.研究方法

3.1. 研究对象

本研究采用自主报名的方式，对北京市某大学的非汉语母语学生进行招募。总计招募 9 位学习者，其中男性 8 名，女性 1 名。学习者的平均年龄为 29.4 岁，汉语平均学习时长为 1.9 年，汉语平均等级为 3.7。所有学习者均具备基本的电脑操作能力，在参与本研究前均未接触过本研究涉及的汉语学习游戏。在实验开始前，学习者均充分了解实验的目的和流程。

3.2. 游戏选择

基于“游戏化学习相关理论框架”(张露 & 尚俊杰, 2018)，本研究设计开发了国际汉语学科教育游戏——《Ethan 的北京生活》。该游戏旨在辅助国际中文初级学习者在课后进行语篇练习。在基于语境的认知体验方面，游戏内容的设计遵循了《国际中文教育中文水平等级标准》(GF0025-2021）对语篇理解的要求，并参考了《HSK3 标准教程》的内容和形式进行了设计。游戏共包含 5 个关卡，每关卡有 3-4 个语篇对话，每个语篇设置了 1-2 个问题，平均每关有 10 题。使用 Unity 3D 三维数字动画技术创设了语篇对应的情境，包括教室、咖啡厅、北京胡同等虚拟场景。本游戏中，学习者需要扮演在北京的美国留学生易森，游戏内容则着眼于易森在北京城内的所见所闻。为增加游戏的趣味性和新鲜感，除主角外，本游戏还设计了性格各异的其他游戏角色，如同学、老师，导游等。每个关卡分为“背景介绍，任务接取，进行对话”三个主要部分，学习者在进入关卡之前会接收到与场景相关的背景信息，逐步融入游戏情境。下图 1 展示了游戏内的对话界面。



图 1 游戏内对话截图

3.3. 研究流程

学习者将使用游戏化虚拟情境进行语言学习。在实验开始前，所有参与者需要填写 HSK-3 级的汉语交流意愿前测。在实验过程中，实验组的参与者在两周内共玩 6 次游戏，每次玩 20 分钟左右。所有参与者使用自己的设备独立完成实验任务。在实验结束后，参与者填写 HSK-3 级的语言知识后测，两组问卷的前后测题目一致。

3.4. 研究工具与技术

3.4.1. 汉语交流意愿量表

本研究使用语言交流意愿量表修正版(Willingness To Communicate Revised Scale)测量学习者在游玩过程中的语言交流兴趣与意愿的变化。语言交流意愿量表修正版对原语言交流意愿量表(WTC-S)进行了扩充。原量表包含 8 个题目，主要对外语学习者日常生活环境中(如医院、警察局)的外语交流意愿进行测量。然而，大多数外语学习者并不倾向于在日常生活中使用外语，其外语交流环境通常位于外语课堂以及网络环境(如外语论坛)中。为此，修正版量表在原量表基础上增加“在网络环境中的交流意愿”这一维度，同时保留原量表中的“在课堂中的交流意愿”以及“在课后的交流意愿”两维度内的题目，共包括 11 题，一致性系数(α)为 0.75-0.79，类型为李克特五点量表(Lee & Drajati, 2020)。修正版量表的有效性已在技术支持二语学习领域相关研究中得到验证(Lee & Lu, 2023)。量表选项从“绝对愿意(1)”到“绝对不愿意(5)”五点分布。得分越高，表示学习者的汉语交流意愿越低。本研究中，该量表平均得分为 3.51，一致性系数(α)为 0.91。

3.4.2. 游戏行为编码方案

本游戏连接了 MySQL 数据库，当学习者在游戏内进行操作时，其游戏行为会被记录到 UserPath 数据表中。UserPath 数据表记录了学习者的答题行为，当学习者进行一次选择后，就会增加一条记录，记录的内容包括当前提交的选项 ID、题目 ID、正误情况、提交时间等。当学习者出现错误时，其下一次提交将存在三种可能：①学习者继续提交同一个错误回答；②学习者提交了另一个错误回答，产生新的错误；③学习者提交了正确回答，通过测试。研究表明，体现学习者认知投入的行为主要包括“收集信息、测试回答和提出改进策略”等(Moon & Ke, 2020)。因此，学习者对原有错误的改进可以体现其较高的认知参与，而出现新错误或重复已有的错误则反映出较低的认知参与。为进一步形成对学习者行为的评价框架，本研究设计了下表 1 所示的编码方案，将学习数据为五类行为，包括成功(Success, S)、失败(Failure, F)、重复(Repeat, R)、修正(Modify, M)与新错(New Mistake, N)。“成功”表示学习者回答正确，“失败”表示学习者回答错误，“重复”表示学习者在答错的基础上再次犯相同错误，“修正”表示学习者对上一次错误进行修正，“新错”表示学习者出现了与上一次错误不同的错误。

表 1 学习行为编码方案表

编码	行为	行为描述
S	答对题目	当前提交回答正确
F	答错题目	当前提交回答错误
R	重复答错题目	对当前错误题目，与上一次提交相比，提交相同的错误选项
N	出现新错误	对当前错误题目，与上一次提交相比，提交不同的错误选项

M	改正错误	对当前错误题目，提交正确选项

3.4.3. 滞后序列分析

基于量表结果，将个体平均分低于总体平均分的学习者归为高学习动机组(N=5)，而将个体平均分高于总体平均分的学习者归为低学习动机组(N=4)。本研究将对不同类型学习者的学习行为数据进行滞后序列分析，以“探究不同类型学习者的典型学习行为模式”(Bakeman & Gottman, 1997)。滞后序列分析法是“通过分析某种行为在另一种行为后出现概率的显著性来探究其行为模式”的方法，可分析发起行为之后另一种行为伴随的可能性，及发起行为与伴随行为形成的行为序列是否存在显著性(Bakeman & Quera, 1995)。首先使用 pyseqlsa 工具将编好码的每一类学习者的行为数据表格进行滞后序列分析，随后根据程序生成的调整后残差表绘制行为转换图。

4.研究结果

4.1. 不同学习动机强度学习者的游戏化学习行为频次分析

下图 2 展示了不同类型学习者学习行为的平均频次。高学习动机组学习者的“成功-失败(S/F)”比为 1.4，而低学习动机组的“成功-失败”比为 2.1，说明低学习动机组的学习者汉语知识水平更高，拥有更好的学习表现，在游戏中犯错的次数更少。然而，由于学习动机的不同,二者在的学习行为具有显著差异。具体表现为,高学习动机组的学习者成功通过测试(S)、修正错误（M）以及未通过测试（F）的频次均显著高于低学习动机组，产生新错误（N）的频次略高于低学习动机组。低学习动机组的学习者明显比高学习动机组更容易出现重复错误（R）。以上数据表明，高学习动机组的学习者倾向于积极投入游戏任务，包括更主动地探索游戏内容、尝试新的学习策略，对学习任务更加专注和投入等；相反，低学习动机组的学习者缺乏内在动机去积极参与游戏学习任务，虽然其答题正确率更高，但对于未掌握的知识点往往采取消极的态度，不主动纠正已有的错误，而对其进行不断重复。

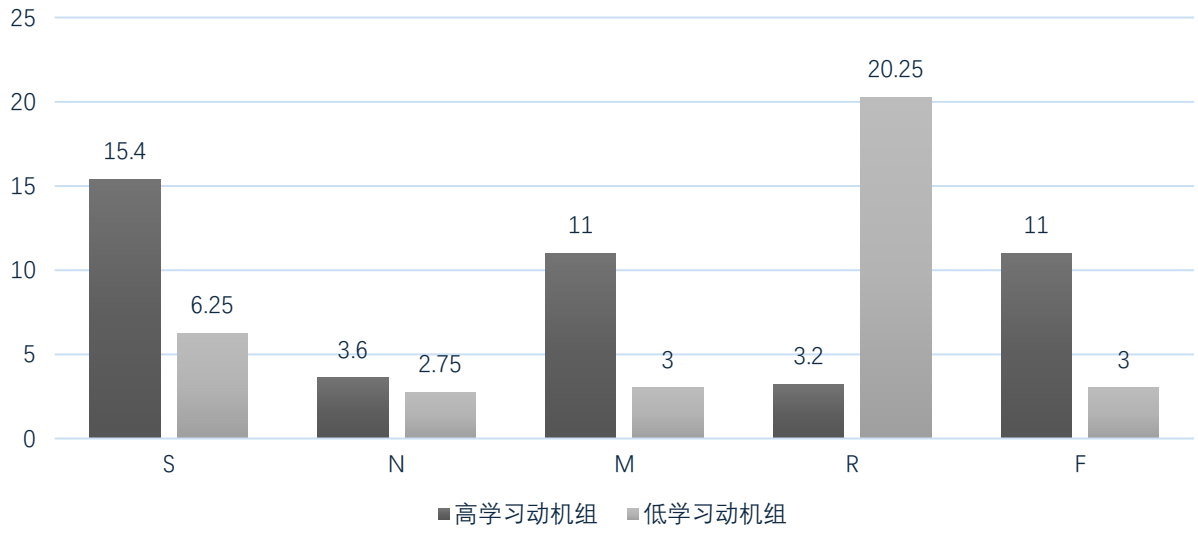


图 2 不同学习动机强度游戏学习行为平均频次统计图

4.2. 不同学习动机强度学习者的游戏化学习行为序列分析

为进一步探究不同类型学习者的学习模式，使用滞后序列分析法对二者的游戏行为数据进行进一步分析。滞后序列分析可以“计算某一时间段内各游戏行为的过渡频率”，并得到如下表 2 所示的调整后的残差表(Hou, 2010)。表中的行表示开始行为，列表示后继行为，若“a 行 b 列对应的残差值(z-score)大于 1.96，则表示行为序列 a→b 呈显著($p<0.05$)”(Bakeman & Gottman, 1997)。为了更直观地展示行为间的转换关系，根据调整后的残差表中的显著序列，绘制出如图 3、4 所示的学习行为转换图。图中节点表示各种用户行为，连线表示行为与行为之间的连接具有显著意义，箭头代表行为转换的方向，线条的粗细表示行为连接的显著水平，线条上的数据则是调整后的残差。

和行为频率分析的结果一致，低学习动机组学习者连续成功(S→S,M→S)的显著性更高，在游戏中更容易连续做对题目，说明其拥有更高的汉语水平。当学习者在游戏中回答错误时，界面内会显示题目以及所有选项的英文翻译，以降低学习者的答题难度。然而，低学习动机组学习者对错误的首次修正(F→M)显著性低于高参与组学习者，说明他们无法对以上认知辅助工具进行有效利用。同时，低学习动机组学习者较少出现对新错误和重复错误的修正，当他们出现以上两类错误时，往往对其进行不断重复(N→N，R→R)。此类学习者很可能不会对错误进行反思，而是通过不断尝试选项来通过测试。由于每道题的选项数量有限，他们才得以在排除某些选项的情况下将题目做对。上述学习策略说明，此类学习者对错误知识点的认知并不清晰，较低的学习动机水平使他们无法有效对学习过程进行监控。同时，由于学习动机的不足，低学习动机组学习者运用认知辅助工具修正错误的能力较弱。对于不理解的知识点，即使在阅读了英文翻译的前提下，此类学习者也无法将知识点进行内化，并在情境中进行运用。

相比之下，高学习动机组展现出更为复杂的学习行为模式，具有 12 种显著序列。他们在遇到错误时，通常会立即进行修正，也可能犯新的错误（F→N）或重复原有的错误（F→R）。与低学习动机组不同的是，高学习动机组最显著的序列是对错误回答进行修正（F→M），以及不同指向错误改进行为的序列（N→M，R→M），表明他们在游戏过程中的专注度较高，付出了较多的认知努力。此外，他们重复错误的显著性较低，说明此类学习者能根据认知辅助工具对知识点进行掌握和理解，从而对原有回答进行修改，其修改的正确率通常较高。以上信息说明，高学习动机组学习者在游戏参与度较高的情况下，遇到错误时会积极采取各种策略进行修正，体现了其较强的元认知策略和能力。

表 2 不同类型学生学习行为调整后残差表(* $p < 0.05$)

Z	S	N	M	R	F
高学习动机组					
S	6.306*	-3.133	-6.074	-2.939	3.000*
N	-3.166	3.118*	2.496*	2.507*	-2.558
M	3.108*	-2.590	-5.021	-2.430	4.779*
R	-2.970	1.567	2.938*	2.792*	-2.400
F	-6.137	3.061*	8.244*	2.341*	-4.959
低学习动机组					
S	7.755*	-1.471	-1.543	-5.989	3.490*

N	-1.471	5.920*	-1.072	-0.962	-1.072
M	3.490*	-1.072	-1.124	-4.362	5.291*
R	-5.989	-1.601	-1.903	8.877*	-4.362
F	-1.543	0.041	7.429*	-2.517	-1.124

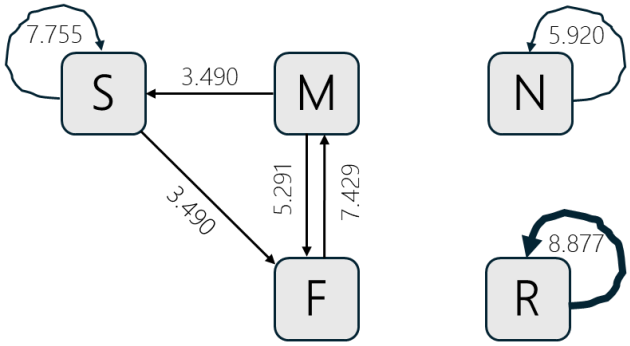


图 3 高学习动机组行为序列转化图

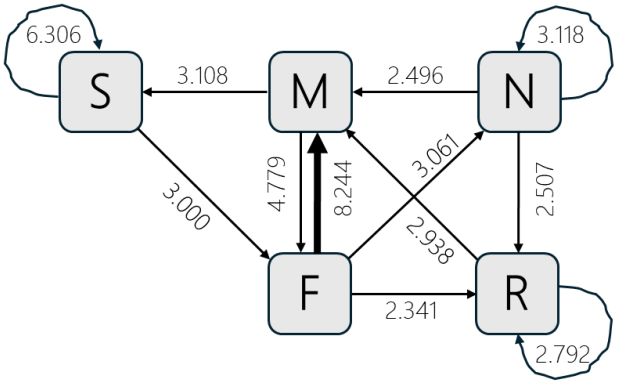


图 4 低学习动机组行为序列转化图

5.讨论与分析

本研究表明,学习动机对学习者在教育游戏中的学习行为和表现产生重要影响。具体而言,低学习动机组学习者在学习过程中表现出较低的学习参与、对错误的修正能力不足以及较低的元认知监控能力。相反,高学习动机组学习者则展现出更积极的学习行为和更高的元认知能力。为了帮助低学习动机的学习者提高学习效果,设计学习辅助工具时可以考虑以下几个方面:第一,提供个性化的学习支持。研究者应当通过系统分析学习者的学习行为和错误模式,设计个性化的学习辅助工具,针对其弱点提供针对性的帮助和指导。第二,研究者应增加认知辅助功能,为此类学习者设计更加直观易懂的界面,提供更多的认知辅助功能,例如错误提示、解释说明或者答题建议,以帮助学习者更好地理解和纠正错误。第三,提供积极的反馈机制。游戏内应包含积极鼓励的反馈机制,及时表扬学习者的努力和进步,激发其学习的积极性和自信心。通过以上方式,可以为低学习动机的学习者设计更加有效的学习辅助工具,提高其学习参与度和学习效果,促进其学习动机和兴趣的提升。

参考文献

尚俊杰 & 裴蕾丝. (2015). 重塑学习方式:游戏的核心教育价值及应用前景. 中国电化教育, 41–49.

张露, 胡若楠, 雷悦, & 尚俊杰. (2023). 基于人工智能表情识别与差异分析的游戏化学习投入状态研究. 现代教育技术, 89–99.

张露, & 尚俊杰. (2018). 基于学习体验视角的游戏化学习理论研究. 电化教育研究, 39, 11-20+26.

Bakeman, R., & Gottman, J. M. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. Cambridge University Press.

Bakeman, R., & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS & GSEQ*. Cambridge University Press.

Cowart, M. (2015). Embodied cognition, PTSD and trauma theory. *The philosophers' magazine*, 88–95.

Eseryel, D., Law, V., Ifenthaler, D., Ge1, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Journal of educational technology & society*, 17, 42–53.

Falk, J., & Dierking, L. (2000). *Learning from museums: Visitor experiences and the making of meaning.* Rowman & Littlefield.

Hou, H.-T. (2010). Exploring the behavioural patterns in project-based learning with online discussion: Quantitative content analysis and progressive sequential analysis. *The Turkish online journal of educational technology*, 9.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Lee, J. S., & Drajati, N. A. (2020). Willingness to communicate in digital and non-digital EFL contexts: Scale development and psychometric testing. *Computer assisted language learning*, 33, 688–707.

Lee, J. S., & Lu, Y. (2023). L2 motivational self system and willingness to communicate in the classroom and extramural digital contexts. *Computer assisted language learning*, 36, 126–148..

Mayer, R. E. (2002). *Multimedia learning*. Academic Press.

Moon, J., & Ke, F. (2020). In-Game actions to promote game-based math learning engagement. *Journal of educational computing research*, 58, 863–885.

Ninaus, M., Moeller, K., McMullen, J., & Kiili, K. (2017). Acceptance of game-based learning and intrinsic motivation as predictors for learning success and flow experience. *International journal of serious games*, 4.

Ozverir, I., Osam, U. V., & Herrington, J. (2017). Investigating the effects of authentic activities on foreign language learning. *Educational technology & society*, 20, 261–274.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist*, 55, 68–78.

紙筆遊戲不僅是遊戲：探討非數位國小數學教育的遊戲式學習

Paper-and-Pencil Games are More Than Games: Exploring Non-digital Game-Based Learning in Elementary Mathematics Education

郭鋅恩[1*]，陳德懷[1]
[1]中央大學 網路學習科技研究所，臺灣
[*] wesley.ncu@gmail.com

【摘要】 本研究探討紙筆遊戲在國小數學教學上的應用，擬針對台灣某小學的三、四年級學生為對象，分為實驗組和對照組，實驗組在相關的數學單元中，除了傳統教學外，還加入三款紙筆遊戲，對照組則只使用傳統教學。透過數學學習興趣量表和自我效能量表進行評估，並使用 ANCOVA 分析數據，同時通過開放式問卷與訪談，深入了解遊戲對學習的影響，以及對教學策略的改進建議。本研究的優勢在於紙筆遊戲的易用性與低成本，希望透過遊戲式學習，讓學生體驗數學的樂趣和價值，培養他們的數學素養和能力。
【關鍵字】 遊戲式學習；非數位的；紙筆遊戲

Abstract: *This study explores the application of paper-and-pencil games in elementary mathematics education. It is planned to target third and fourth-grade students at a certain elementary school in Taiwan, which will be divided into experimental and control groups. In addition to traditional teaching methods, the experimental group will incorporate three paper-and-pencil games within relevant mathematical units, while the control group uses only traditional teaching. The assessment will be conducted using a mathematics learning interest scale and a mathematics self-efficacy scale, with data analysis performed through ANCOVA. Open-ended questionnaires and interviews will further assess the games' educational impact and inform teaching strategies. The advantage of this research lies in the ease of use and low cost of paper-and-pencil games. Through game-based learning, it aims to let students experience the fun and value of mathematics, fostering their mathematical literacy and abilities.*
Keywords: game-based learning, non-digital, paper-and-pencil game

1. 前言

在數位化時代下，數位遊戲式學習（digital game-based learning）利用數位科技和遊戲元素，創造有趣的學習環境，讓學習者在遊戲中獲得知識和技能，長久以來是教育界的熱門話題。然而，數位遊戲式學習並不能完全取代傳統遊戲的教育價值，其中傳統遊戲，如大富翁、踢毽子、打陀螺等，不僅可以培養學習者的基本素養，也可以促進學習者的社交、情感、品格等發展。此外，許多遊戲與數學思維有密切關係，可以訓練學習者的計算、推理、策略等能力，例如賓果（bingo）、井字遊戲（tic-tac-toe）、四子棋（connect four）、麻將（mahjong）、拉密（rummikub）等，甚至在印尼 State University of Surabaya（UNESA）每年還會針對中學舉辦 Rumah Perkalian 桌遊比賽（Garry et al., 2019）。

聯合國教科文組織（United Nations Educational, Scientific and Cultural Organization, UNESCO）在報告中指出，數位科技在教育中的應用逐漸普及，但在非先進地區，電腦和其他數位設備在課堂上的使用仍然不普遍，此情況短期內無法改變，因為數位科技的應用需要逐步發展（2023）。此外，對教育科技的關注往往伴隨著高昂成本，導致弱勢群體經常無法從這些技術中獲益，使得邁向全球教育目標（quality education, SDG 4）的實現更加困難，還

強調科技在教育中的使用需要有明確的目標和原則，以確保其帶來的是利益而非傷害，例如需要避免科技使用時的分心和缺乏人際關係等風險（UNESCO, 2023）。

目前有關遊戲式學習的研究，過於關注數位遊戲的研究（Balladares et al., 2023; Talan et al., 2020），忽略非數位遊戲的效果，也可以看到世界組織呼籲只有在數位設備真正有助於支持學習時才應該使用它們，以及一些國家已經制定關於兒童和青少年使用數位設備的明確法規（UNESCO, 2023）。因此本研究關注非數位的遊戲式學習，尤其是紙筆（paper-and-pencil）遊戲，這是一種方便且經濟的教學工具，只需紙筆即可進行，無需任何數位設備或桌遊材料，但卻鮮少被研究探討，本研究目的是探討將紙筆遊戲運用在國小數學教學上，並分析其對學習者學習成效的影響。

2. 文獻探討

2.1. 遊戲式學習

遊戲式學習是一種結合遊戲特性與機制，以促進學習者獲得知識、技能、態度和動機的教學策略。遊戲式學習可以分為數位遊戲式學習和非數位遊戲式學習，前者是利用電腦、平板、手機等數位設備進行遊戲，後者則是使用紙筆、桌遊、卡牌等實體材料進行遊戲。然而，近年來的研究似乎過於集中在數位遊戲式學習，而忽略了非數位的遊戲式學習效果和價值（Balladares et al., 2023; Talan et al., 2020）。事實上，非數位遊戲式學習也有許多優勢，例如提供更多的同儕互動機會、更廣泛的活動選擇、更大的設計和內容靈活性，以及多樣化的遊戲平台，還有發現非數位遊戲中學生的表現效果大於數位遊戲，儘管差異不顯著（Talan et al., 2020）。此外，Balladares 等人（2023）針對多篇學齡前桌遊研究進行回顧，發現這些桌遊對兒童的數學技能和興趣有正向影響。

在台灣有個非數學且非數位遊戲的實證研究，Shih 等人(2022)設計了一款名叫「Fragrance Channel v2.0」的策略性桌上遊戲，該遊戲結合一些科技輔助，透過角色扮演和模擬歐洲探險的情境，讓中學生在遊戲中探索 1400 年代至 1700 年代的地理及歷史知識，探討他們的利社會行為及反社會行為的互動策略。

2.2. 遊戲式學習在數學的應用

兩篇回顧型文獻指出，數位遊戲式學習的研究過於集中，而忽略了非數位遊戲式學習的效果和價值（Balladares et al., 2023; Talan et al., 2020），加上本研究的重點是非數位的遊戲式學習，因此文獻回顧以實體遊戲教材在數學教學上的應用為主。

在實體遊戲教材於數學的應用上，有些研究探討針對不同文化的傳統遊戲在數學教學上的潛力和效益。例如，Fouze 和 Amit（2017）認為貝都因人的傳統遊戲，可以增進學生的數學興趣、動機和技能，同時也讓學生感受到數學與他們的文化和日常生活的關聯，有助於數學教育的多元化、民主化和人性化；Bayeck（2018）回顧了非洲地區流行的五種傳統桌遊，發現它們需要數學技能和空間思維，同時也可以促進社交技能、文化理解和和諧。另外，也有些研究將桌遊數位化，Garry 等人（2019）參考印尼的「Rumah Perkalian」桌遊，將其數位化，並發現使用者對這款井字遊戲有正向的反應，該遊戲主要是提升 1 至 9 的乘法應用。

3. 研究方法與設計

3.1. 遊戲內容

一學期的課程內容會針對單元搭配遊戲進行教學，已參考當前主流的教科書內容，針對三、四年級運算符號相關單元，第一款是由 David Johnson-Davies 所設計的，名叫「Number Countdown」（http://www.papg.com/show?2XLM ）的紙筆遊戲，可以參考圖 1，需要兩名玩家，一名玩家秘密選擇四個較小的數字（1 至 10）以及 2 個較大的數字（10 至 100），另一

名玩家則秘密寫下一個三位數（100 至 1000）的目標數字，雙方互相公開數字後，僅能使用這六個數字嘗試用加減乘除進行整數運算，六個數字不一定要全部使用，但每個數字只得使用一次，雙方盡可能接近目標數字，誰愈接近該目標數字或者用最少運算步驟便獲勝；第二款遊戲是由 Bonnie Tank 和 Lynne Zolli（2001）所設計的「Four Strikes and You're Out」，需要兩名玩家，一名玩家秘密寫下 2 位數加法的目標算式，另一名玩家需要猜測當中有 0 至 9 哪些數字，若猜中目標數字，寫下數字的玩家要在相應的空格填上這個數字，若猜數字的玩家在失誤 4 次前得到正確的目標加法算式，便是猜數字玩家勝利，反則寫下數字的玩家勝利，參考圖 2；第三款遊戲是「點線遊戲」，規則是紙上畫上 15 個點的大三角形，兩個人遊戲，每個人輪流只能用直線畫上最多 3 個點，最少畫 1 個點，畫到最後 1 個點的玩家便輸了遊戲，

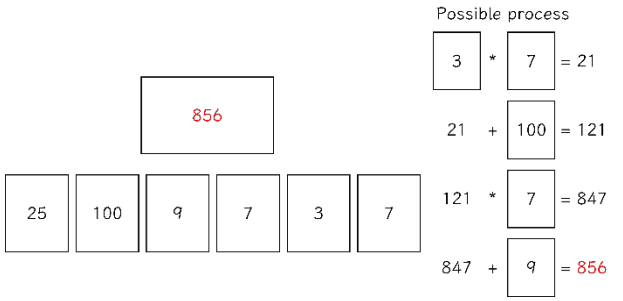


圖 1、Number Countdown

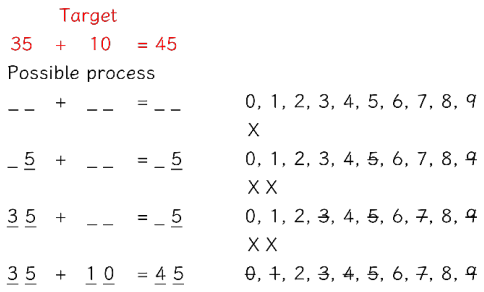


圖 2、Four Strikes and You're Out

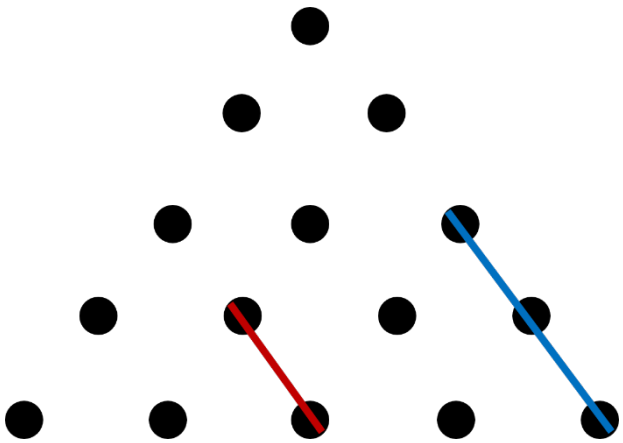


圖 3、點線遊戲

這遊戲將運算邏輯結合了點、線與面，玩家要想辦法給對方留下最後一個點，可以參考圖 3。

3.2. 教學流程與方法

在完成相關知識單元的教學後，向全班介紹遊戲規則，並引導學生們一同嘗試遊戲，同時鼓勵學生們在休閒時間也可以遊玩，因為它只需要紙筆就可以進行，不需要其他的數位設備或其它材料；接著，教師可以讓兩人一組，互相配合進行遊戲練習，教師可以觀察與指導練習狀況，並邀請其他教師或助教也參與遊戲，讓學生發現或分享可以連續獲勝的特殊技巧與策略，激發學生的好奇心和學習意願，教師也可以視情況設置不同的難度等級，讓學生有挑戰性和成就感；最後由教師舉辦正式的淘汰賽，讓學生之間進行競爭，選出前三名的優勝者，讓他們有機會與其他強者對戰，檢測學生的理解和運用，並給予適當的獎勵和回饋。

4. 研究結果

預計採用準實驗研究，以台灣某小學的三年級和四年級各兩班共 80 位學生為研究對象，分為實驗組和對照組，每組各 40 位。實驗組在數學教學中，除了使用傳統的教學方式外，還融入紙筆遊戲的數學教學活動，而對照組則只使用傳統的教學方式。兩組學生將接受相同時

數的數學教學，以確保研究的公平性。研究以數學學習興趣與自我效能量表進行前後測，以兩組學生的前測成績作為共變數，並透過 ANCOVA 分析兩組學生的學習成效差異，預期實驗組學生在數學學習興趣與自我效能上將顯著優於對照組。此外，也會針對實驗組進行開放式問卷與訪談，深入了解紙筆遊戲對學習情況的影響及需改進之處，這將是未來相關教學策略提供寶貴的見解與建議。

5. 結論與未來展望

本研究的優勢在於紙筆遊戲的易用性與低成本，使得教師和學生都能夠輕鬆地進行遊戲式學習，希望透過紙筆遊戲，能夠讓學生體驗數學的樂趣和價值以及培養他們的數學素養和能力，學生若有興趣可以隨時找人一起遊玩，不必像數位遊戲需要科技裝置才能進行，在未來甚至可以創造屬於自己的遊戲。以下是未來可能的研究方向，首先，其一可以將紙筆遊戲的內容數位化，了解實體遊戲與數位遊戲在教育上的差異；其二將紙筆遊戲運用在不同年齡層或不同遊戲，探討不同年齡或遊戲對學習的影響和效果；其三嘗試將紙筆遊戲與第二外語教學結合，探討遊戲是否能夠促進學生在數學和第二外語的習得。

致謝

本研究在國家科學及技術委員會 (NSTC 112-2410-H-008-020-MY3) 與「中央大學學習科技研究中心」的資助下完成，謹此致謝。

參考文獻

Balladares, J., Miranda, M., & Cordova, K. (2023). The effects of board games on math skills in children attending prekindergarten and kindergarten: A systematic review. *Early Years,* 1-25.

Bayeck, R. Y. (2018). A review of five African board games: Is there any educational potential?. *Cambridge Journal of Education, 48*(5), 533-552.

Fouze, A. Q., & Amit, M. (2017). Development of mathematical thinking through integration of ethnomathematic folklore game in math instruction. *Eurasia Journal of Mathematics, Science and Technology Education, 14*(2), 617-630.

Garry, M. H., Yamasari, Y., Nugroho, S. M. S., & Purnomo, M. H. (2019, November). Design and Implementation Serious Game “Tic Tac Toe Math”. In *2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)* (pp. 1-6). IEEE.

Shih, J. L., Chiu, M. M., & Lin, C. H. (2022). Personalities, sequences of strategies and actions, and game attacks: A statistical discourse analysis of strategic board game play. *Computers in Human Behavior, 133*, 107271.

Talan, T., Doğan, Y., & Batdı, V. (2020). Efficiency of digital and non-digital educational games: A comparative meta-analysis and a meta-thematic analysis. *Journal of Research on Technology in Education, 52*(4), 474-514.

Tank, B., & Zolli, L. (2001). *Teaching Arithmetic: Lessons for Addition and Subtraction Grades 2-3*. Math Solutions Publications, Marilyn Burns Education Associates.

United Nations Educational, Scientific and Cultural Organization. (2023). *Global education monitoring report summary, 2023: technology in education: a tool on whose terms?*. United Nations Educational, Scientific and Cultural Organization.

融入關鍵節能概念線上遊戲輔助國小學生能源議題設計研究

Exploring the Benefits of Incorporating Vital Energy-Saving Concepts into Online Game-Based for Elementary School Students Learning on Energy Issues

王曉璿 1，李思萱 2*，邱光諺 3，陳俊霖 4，陳啟榮 5
1 國立臺中教育大學
2 國立臺中教育大學
3 國立臺中教育大學
4 國立臺中教育大學
5 國立臺中教育大學

* xunym0@gmail.com

【摘要】 能源議題日益受到關注，而遊戲在日常生活中的普及度也日漸提升，年齡層更是越來越降低。因此，我們提出一項結合能源知識和遊戲互動的專題，主旨在透過遊戲方式培養兒童的節能意識。我們將開發一款結合教育元素的遊戲，通過虛擬場景模擬現實能源使用情境，引導玩家深入了解節能的必要性。建立孩童正確的節能價值觀，為未來社會培養更為節約能源的新一代。

【關鍵字】 環境覺察；數位遊戲式學習;趣味化學習

Abstract: *The issue of energy conservation is increasingly gaining attention, and the popularity of gaming in daily life is also on the rise, especially among younger age groups. Therefore, we propose a project that combines energy knowledge with interactive gaming, aimed at cultivating energy-saving awareness in children. We will develop an educational game that simulates real-world energy usage scenarios through virtual environments, guiding players to gain a deeper understanding of the importance of energy conservation. Establishing the correct values of energy conservation in children aims to foster a new generation committed to saving energy for the future society.*

Keywords: *Environmental Awareness, Digital Game-Based Learning, Fun Learming*

1. 前言

能源議題在現代社會備受關注，隨著全球人口不斷增長和能源需求急速擴大，有效利用有限的能源資源成為當務之急。節能不僅減緩氣候變遷，也減輕能源壓力，並促進可持續發展。本研究旨在探討如何結合能源知識和遊戲互動，引導孩童參與節能活動。透過此方法，不僅提升孩童對能源議題的認識，也培養其節能意識和行為，為未來打下堅實基礎。

2. 文獻探討

2.1. 能源議題探討

世界上有大約 30 億人用簡單的爐灶或燃燒木材、動物糞便或煤炭來烹飪和為房屋供暖。而隨著發展中國家工業化程度的加深，將需要獲得可靠的電力供應，此外，已經進入快速發展期的國家，中產階級的擴大以及相應地生活方式使能源需求也有所增加。全球的能源消耗的增加，讓節能和能源轉型成為很重要的一環。世界能源理事會的一份報告認為，能源行業

正面臨長達數十年的轉型。「菲律賓人權委員會於 2015 年宣布針對 47 家全球化石燃料的主要生產者展開調查」、「世界氣象組織曾發出警告，全球年均溫最早可能在 2026 年升溫超過攝氏 1.5 度，屆時水資源和糧食短缺等風險將倍數成長，威脅到全世界的安全。現在已無法再拖延，必須立即淘汰化石燃料！」從文獻可看到，能源問題是被全球共同重視的議題(Gray，2017）。

2.2. 遊戲的影響力

史密蘭斯基認為「遊戲類似鷹架，可以讓兒童在現有能力上沿此發展」、皮亞傑認為「遊戲反映出兒童發展的層次」，從這些理論可以知道，遊戲在兒童發展扮演很重要的角色。此外兒童能從遊戲中揭露本身的經驗、對經驗的反應、對經驗的感覺、願望及需求、對自我的概念，是瞭解兒童及與兒童溝通最適當、最自然的方式。這些反饋可以讓我們更好去得知孩童遊戲後的感受與效果（Smilansky，Piaget，1919）。

2.3. 線上學習與孩童

在高達一至二百萬的遊戲人口中調查發現，未成年孩子曝露在網路社群的時間愈來愈早，年齡也愈來愈小。這些年輕玩家在上網時，多數是在遊戲中停留。創市際公司「ARO 網路測量研究」分析最近的統計數據發現，十到十四歲的年齡族群造訪線上遊戲網站，是一般上網的二倍以上。巴哈姆特也調查發現，國中與國小教育程度的玩家比例，合計將近 24%。數據顯示年輕玩家對遊戲的喜好度（林照真，2004）。

2.4. 教育與遊戲

米哈里·齊克森提出的心流理論中，提到如此全神貫注的狀態甚至能於其中獲得自由與幸福感。一款有吸引力、有趣味性的教育遊戲，才能讓學習者持之以恆並專心學習，進而大幅提升學習者的心流狀態，提升學習效果（Mihaly，1975）。

我們以「節能概念」為主軸，輔導國小兒童加強能源觀念，透過遊戲引發其思考，發現日常生活中的能源危機。希望透過研究，讓孩童了解能源知識，並從購買產品時開始實踐節能。讓我們共同為地球盡一份微薄心力。

3.研究方法

本研究主要透過文獻探討關鍵能源與教育遊戲議題，然後設計對應 3D 教育遊戲，並進行初步實驗以了解學生學習成效，研究對象是小學中年級的學生。我們從台灣展翅協會(2015)的調查看到兒童擁有自己專屬的上網工具高達 84.7%，且從愛傳資訊媒體發展協會對孩子使用手機的用途調查看到 64.5%的孩子使用手機的用途在於玩遊戲，因此我們希望透過日常結合教育，引發孩童學習興趣。而從「運用手機遊戲專題提升學生於視窗程式設計課程之興趣」研究報告的結果我們看到手機遊戲融入教學對於學習成效是有所提升的。此研究同時是基於 ARCS 動機教學模式之理論研發遊戲，藉由線上遊戲引起孩童注意，而主題是與日常切身相關的居家節能，遊戲過程以獎勵為主，建立孩童自信心、獲得滿足。希望藉此讓受眾主動運用所學於日常，落實節能。

3.1. 遊戲文本設計

未來，哈基米星球面臨著快速上升的海平面挑戰，成為一個以火力發電為主要能源的世界。小咪決定行動起來，立志成為節能和再生能源專家，保護家園。她深入研究節約用電技巧，掌握太陽能和風能等可再生能源的運作方式。小咪堅信，節能是每個人都能參與的解決方案，要改變能源使用模式，從日常生活做起。她開啟了一場居家節能的探索之旅。

3.2. 場景設計

3.2.1. 客廳

清潔冷氣濾網很重要，因髒的濾網會降低冷氣效能。選擇淺色系窗簾可以降低室內溫度，減少冷氣使用。將冷氣溫度設定在 23 度以下有助於節省能源。同時使用電扇和冷氣可以提高空氣流通，減少對冷氣的依賴，節省能源。確保門窗緊閉有助防止冷氣流失，進一步減少能源浪費。另外，定期清潔灰塵和燈管可以提高照明效率，減少能源消耗。

3.2.2. 房間

冷氣噸數的選擇至關重要，過小的噸數會增加能源浪費。其次，未關閉照明設備和電腦待機模式也會浪費能源。窗戶未完全關閉會影響冷氣效果。最後，未拔除不在使用的吹風機插頭也是能源浪費的行為。調整這些細節可實現更節能的生活方式。

3.2.3 廚房

冰箱的過度擠滿影響空氣流通，增加冷氣機工作負荷，導致能源浪費。其次，老舊電冰箱能效較低，考慮升級為能源效率更高的 LED 燈泡可節省能源開支。直接將熱食放入冰箱會增加冷卻負荷，應等食物冷卻後再存放。此外，未拔除未使用的熱水壺插頭和電鍋外鍋卡污漬也是潛在的能源浪費點，應當注意清潔和插頭管理。

3.3 角色設計

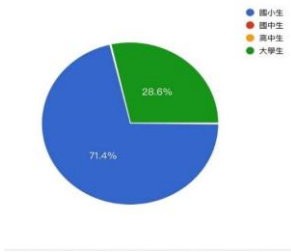
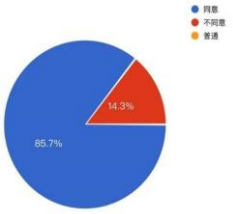
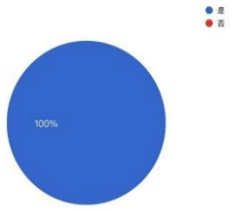
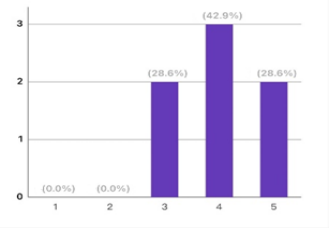
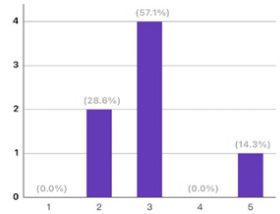
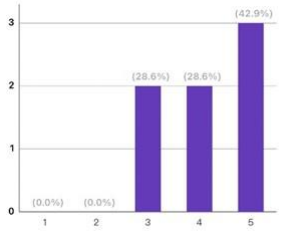
遊戲設計配合小學高年級的受眾製作。角色形象以可愛的女孩呈現，有助於玩家更容易投入遊戲中、感到親近。畫風以手繪風格為主，較受兒童歡迎、能夠引起孩子的注意，激發他們的好奇心。

3.4 遊戲機制

在這個能源節約遊戲中，玩家需要保護自己的能源條，避免它消失。玩家將在不同場景進行空間探索，找出可能造成能源浪費的地方。發現浪費源後，將觸發生動的能源知識影片，提供深入解釋並回答問題以獲得獎勵。整個遊戲系統將記錄玩家的正確回答率，讓玩家了解學習成效。這個能源節約遊戲旨在娛樂的同時，提高玩家對能源議題的認識，鼓勵節約能源的行為。

類型	畫面示意圖	設計說明
封面		進入遊戲後可見設定鍵即開始鍵，開始後即選擇關卡，並可在遊玩後掃描 QRcode 進行回饋
設定畫面		可自定義遊戲音量 有中英文語言可供切換
第一關		在場景中找尋耗能物品，也可在下方列表看到物件
題目畫面		每找出一個一個物件便會產生與其相關的對應問題
遊戲成功		過關後顯示該關卡通關時間與答題正確率

4. 回饋建議整理和分析

遊戲適合年齡層	是否習得新知	是否了解遊戲意義
		
遊戲流暢度	**遊戲問題難易度**	**遊戲畫面風格**
		

4.1 課綱

針對我們的目標客群，將遊戲設計貼近日常的常見情境，以間接達成教育目標。我們將加入我國十二年國民基本教育課程綱要中的相關知識，結合日常，以增強教育效果和連結性。

4.2 UI 設計

玩家指出遊戲互動感較不足，因此我們增加了按鈕變化和音效提示，提升與玩家的互動感

4.3 挑戰性

遊戲挑戰性相對缺乏，使"SAVE" the word 留住玩家的機率降低。

我們嘗試再增加紀錄功能的資料庫，其中像是排行榜功能使玩家能相互競爭。

玩家對"SAVE the word"持樂觀態度，認為這類教育遊戲能有效提升保育環境和節約能源的知識。未來我們希望解決負面問題，持續運作"SAVE the word"，達成我們的目標。

5. 結論與建議

遊戲具有教育用途的高度潛力，使過程更加趣味，並與能源議題連結更加重要。在創作遊戲時，必須考慮故事性、吸引力、互動感和環境營造等因素，以展現我們想表達的意圖。

參考文獻

世界能源理事會(2017)。人類面臨的最大能源挑戰是什麼？

世界氣象組織(2022)3 分鐘看懂全球環境焦點：2022 年 5 月：立即淘汰化石燃料

張高賓（2009）遊戲治療能力評量表之編製研究。國立嘉義大學輔導與諮商學系論文

林照真(2004)線上遊戲伴孩子長大。台北:天下雜誌 301 期。

米哈里·齊克森(1975)心流理論。蜂窩兒童大學

Franzoi, Stephen.(2014) Essentials of Psychology

Csíkszentmihályi, Mihály, Creativity(1996) Flow and the Psychology of Discovery and Invention,New York: Harper Perennial

整合實體展覽與遊戲式延展實境體驗於歷史課程之應用與研究

Integration of Physical Exhibitions and Gamified Extended Reality Experiences in the Application and Study of Historical Courses

蔡遵弘[1*]，城菁汝[2]
[1]國立清華大學科技藝術研究所
[2]國家圖書館知識服務組
[*] otaru@gapp.nthu.edu.tw

【摘要】文本研究探討利用延展實境（XR）技術在歷史教育中的應用，特別是結合實體展場與遊戲式實境體驗以提升學習動機和參與度。透過台灣記憶典藏系統的校園資源，本研究創建了跨越四代的教育情境。透過遊戲化設計和情境式學習，學生能體驗不同時期的教育環境和社會氛圍，從而深入理解歷史和文化的演變。本研究強調 XR 技術在教育領域的應用可能，目標為提升對歷史內容的興趣和加深對過去社會和教育環境理解方面的效果。透過這種創新的教學設計，本研究不僅加深了學生對歷史的瞭解，也為 XR 應用於歷史教育提供了相關研究參考。

【關鍵字】 延展實境；遊戲化；情境式學習；歷史教育；沉浸式體驗

Abstract: *This study examines the application of Extended Reality (XR) technology in historical education, focusing on integrating physical exhibitions with gamified virtual experiences to enhance student motivation and participation. Utilizing Taiwan's Memory Collection System's campus resources, it constructs educational scenarios spanning four generations. Through gamified design and situated learning, students experience varied educational environments and societal atmospheres across periods, deepening their understanding of historical and cultural evolution. The study highlights the potential of XR technology in education, aiming to increase interest in historical content and enhance comprehension of past societal and educational contexts. It provides valuable insights into the use of XR in historical education.*

Keywords: Extended Reality, Gamification, Situated Learning, Historical Education, Immersive Experience

1. 前言

歷史教育不僅是知識傳授的過程，更是一種鍛鍊批判性思考和反思性理解能力的重要途徑。通過歷史視角的思考，我們不僅能培養思維能力，還能深入理解文化脈絡，從而提高對本土歷史和文化的認同感 (Carretero, Rodríguez-Moneo & Asensio, 2012)。這種教育方法在全球範圍內被廣泛認可，並被視為學術和文化教育中不可或缺的一部分 (Keleşzade, Güneyli, & Özkul, Citation2018)。

傳統上，歷史教育主要依靠文字記載、圖片資料以及博物館中的文物展示來重現過去的生活和歷史事件。雖然這些方法能提供豐富的資料，幫助我們構建對過去的理解，但在某種程度上，這些方式仍受限於媒介的局限性，對於完整重現歷史場景的體驗，往往依賴於觀眾的想像力 (Greipl, Moeller & Ninaus, 2020)。

然而，隨著科技的進步，尤其是延展實境（Extended Reality, XR）技術的發展，我們現在有機會以傳遞體驗的方式來輔助歷史教育。XR 技術涵蓋了擴增實境（Augmented Reality,

AR）、虛擬實境（Virtual reality, VR）和混合實境（Mixed Reality, MR），使得虛擬與現實世界的結合成為可能。這些技術能夠將過去的人物和物件帶入當前的實體環境中，創造出一個既互動又沉浸的學習空間 (Harley, Lajoie , Haldane, McLaughlin & Poitras, 2020 ; Martin, 2019)。

本研究旨在探討如何通過遊戲化設計和 XR 技術，將歷史課程內容以劇情化的形式呈現，從而提高學習者的動機和參與度。本研究將聚焦於如何在實體展覽中整合 XR 技術，以體驗的形式傳遞知識，增強學生對歷史課綱中對兩次國語運動章節的理解和學習成效。

2. 文獻探討

本研究目標在借助 XR 技術能整合虛實的特質，將歷史上的過去所存在的人設與環境物件與真實展覽結合，並透過遊戲式的情節設定促使學生進行探索，進而獲得課程中想傳遞的知識，意即透過 XR 技術來創造情境式學習 (Situated Learning) 的效果，在經遊戲式體驗的過程來提供探索式學習的結構，本章節將就 XR 技術在教育中的應用，與結合遊戲式體驗於情境式學習的相關文獻，來建構本研究之基礎。

2.1. XR 技術在教育中的應用

近年來，XR 相關技術在教育領域的應用日益增多，為學生提供了更加豐富和具互動性的學習體驗。相關技術通過結合現實世界與虛擬元素，創造出一種新的學習情境，使學習內容得以轉化的更加直觀具吸引力。澳洲學者 Paola 的研究中展示了如何利用沉浸式虛擬實境來提高學生學習解決問題的技巧。在這個研究中，學生通過 VR 環境來進行問題解決的練習，成果除了有效提高學生的學習興趣，也幫助他們更好地理解了複雜概念 (Araiza-Alba, 2021)。學者 Bogusevschi 等人的研究則聚焦於使用 3D 虛擬學習環境來教授物理學。通過這種方法，學生能夠在虛擬實驗室中進行物理實驗，這不僅使學習過程更加有趣，還幫助學生更好地理解物理學原理(Bogusevschi, Muntean & Muntean, 2020)。此外，亦有研究探討了 AR 和 VR 在教授焊接技巧的應用，研究成果顯示 VR 和 AR 技術可以有效地提高學習者在實際操作中的技能水準 (Chan, Haron, Isham & Mohamed, 2022)。學者 Checa 等人則研究了沉浸式 VR 嚴肅遊戲對學習和訓練的提升作用。研究結果顯示：這種遊戲化的學習方法可以提高學生的參與度和學習成效 (Checa & Bustillo, 2020)。學者 David 與 Andres 的研究透過分析 135 個沉浸式 VR 環境中打造嚴肅遊戲的研究，探討了沉浸式技術對使用者存在感的影響。這項研究表明，沉浸式技術能夠顯著提高學生的參與度和學習體驗 (Cummings & Bailenson, 2016)。

綜合上述研究，可以看出 XR 技術在提高學生學習動機、參與度以及學習效果方面具有顯著優勢。特別是在歷史教育中，這種技術能夠通過創建沉浸式的歷史環境，使學生能夠更加生動地體驗和理解歷史事件和文化。這種新穎的學習方式有望為傳統學習方式為歷史教育帶來突破性的改變，也是本研究之所以採用相關技術的原因。

2.2. 結合遊戲式體驗於情境式學習

遊戲化學習的概念在教育領域已展現其對學習動機、參與度和學習成效的顯著影響。特別是在歷史教育中，遊戲化學習結合情境式學習可以創造更吸引人和互動的學習環境，遊戲化學習元素的應用可以根據學生的個性特質以不同方式影響他們的學習和行為 (Smiderle, Rigo, Marques, Peçanha & Jaques, 2020)。在歷史教育領域，遊戲化可以透過多種方式實現。例如，透過基於歷史事件的模擬遊戲，學生能夠在虛擬環境中體驗歷史情境，從而更深入地理解歷史事件的背景和影響 (Martínez-Hita, Gómez-Carrasco & Miralles-Martínez, 2021)。此外，設計解謎遊戲或尋寶遊戲鼓勵學生探索歷史知識，同時提升他們的批判性思維和問題解決能力 (Shavab, Yulifar, Supriatna & Mulyana, 2021)。

承前一章節所述，XR 技術已被廣泛應用於教育場景。這些技術不僅增加了教育過程中的互動性和吸引力，還能夠創建一個更加沉浸式的學習環境。特別是在歷史教育中，通過利用這些技術來重現歷史場景，學生可以進入這些虛擬重現的歷史情境，這不僅提高了學習的趣味性，亦幫助學生理解和記憶 (Guo, Guo & Liu, 2021)。相關技術的應用延伸了許多新的教學方法，可從根本上改變學生對歷史學科的學習動機與方式，使之成為一個更加生動和注重情境體驗的學習方法。

2.3. XR 結合遊戲式學習之優勢

結合遊戲式體驗於情境式學習對於提升學生的學習效果和參與度具有顯著優勢 (Moseikina, Toktamysov & Danshina, 2022)。結合 XR 體驗與遊戲式學習的方法，不僅僅只是提升體驗或動機，更能透過設計幫助學習者從 XR 體驗中獲得隱性知識 (Majgaard & Weitze, 2020)。XR 的特質讓傳統學習內容以延展的方式構成體驗，讓學習從平面的閱讀轉化成體驗，而整合了遊戲式學習的優勢，使遊戲情境提升學習成效與動機的特質，能夠與 XR 相得益彰，在兼顧體驗式學習的優勢之下，發揮遊戲式學習的特質，讓學習內容讓學習者更加深刻。

無論是在體驗過程中，因為決策或情境的驅使、抑或是在遊戲的機制下做出選擇與判斷，無不是在累積學習者重複咀嚼其中設計好的學習內容，本研究認為這樣的整合優勢，尤其在歷史教育中有更充分的發揮空間，因為其不僅增加了學習的趣味性，還有助於學生更全面和深入地理解歷史內容，並在體驗與遊戲的過程中加深印象，這亦是本研究希望整合遊戲式體驗於情境式學習的重要因素。

3. 研究方法

本研究具體目標，為結合國家圖書館的實體展覽，融合高中課程綱要中歷史科目中針對兩次國語運動的內容而設計，為由於課程內容涵蓋了近一百年教育現場的嚴格，不同的時代背景下學校的氛圍與文化皆有顯著的差異，然而在教室物理空間的安排上，桌、椅、講台與講桌…等配置，大致維持了相似的格局與風格，因此本研究透過模仿教室的格局打造了實體展場，並透過 XR 技術將不同時間點的教育情境融合成遊戲式情境體驗，讓體驗的學生可以在空間中遊走探索，並獲得課程中的相關知識。為了達成此目標，本研究從內容考據乃至於 XR 的內容開發製作都有相當的細節處理，以下分六個小節做說明。

3.1. 內容考據與劇情設計

研究初期，我們選擇國家圖書館提供的台灣記憶典藏系統中的校園資源作為考據基礎，這些資料涵蓋了豐富且多元的歷史文獻，如台灣 1895 年至 1945 年間的畢業紀念冊、1945 年至 1987 年的教科書、考試題本，以及校園的老照片和百年教育文獻……等。這些珍貴的歷史資料不僅提供了深入了解台灣教育史的機會，也為本研究創造了一個獨特的歷史敘事背景。

在研究過程中，我們將這些史料進行細致的整理與分析，並將故事情境劃分為四個不同的時代。為了讓體驗可以涵蓋將近一個世紀的時間線，我們透過考據創作了一劇本，故事講述一位現代高中少女，從一個因方言造成的誤會引發的神秘事件開始，少女意外穿越時空，回到過去，相繼遇到了她的曾祖父、祖父和父親還在中學時的情景。這一系列的時空旅行不僅讓她親身體驗到不同年代的教育環境和社會氛圍，也讓她深入理解了台灣當前多語言使用環境的形成原因，並對最初的誤會有了新的認識，也呼應了課綱中兩次國語運動的整個脈絡。

為了讓學生在遊戲式的情境體驗中能夠身歷其境的感受到不同時代下的社會氛圍，並了解當時曾發生過的重要事件，穿越的教室中會遇到曾祖父、祖父和父親時期的同學和老師，透過遊戲化的情境設計，學生會經歷一連串任務，並在與各角色的互動中，以體驗的方式去理解課本中對於這些時代背景介紹的內容，文本中整個體驗內的角色總計有 19 人，彼此之間的關係如下圖。

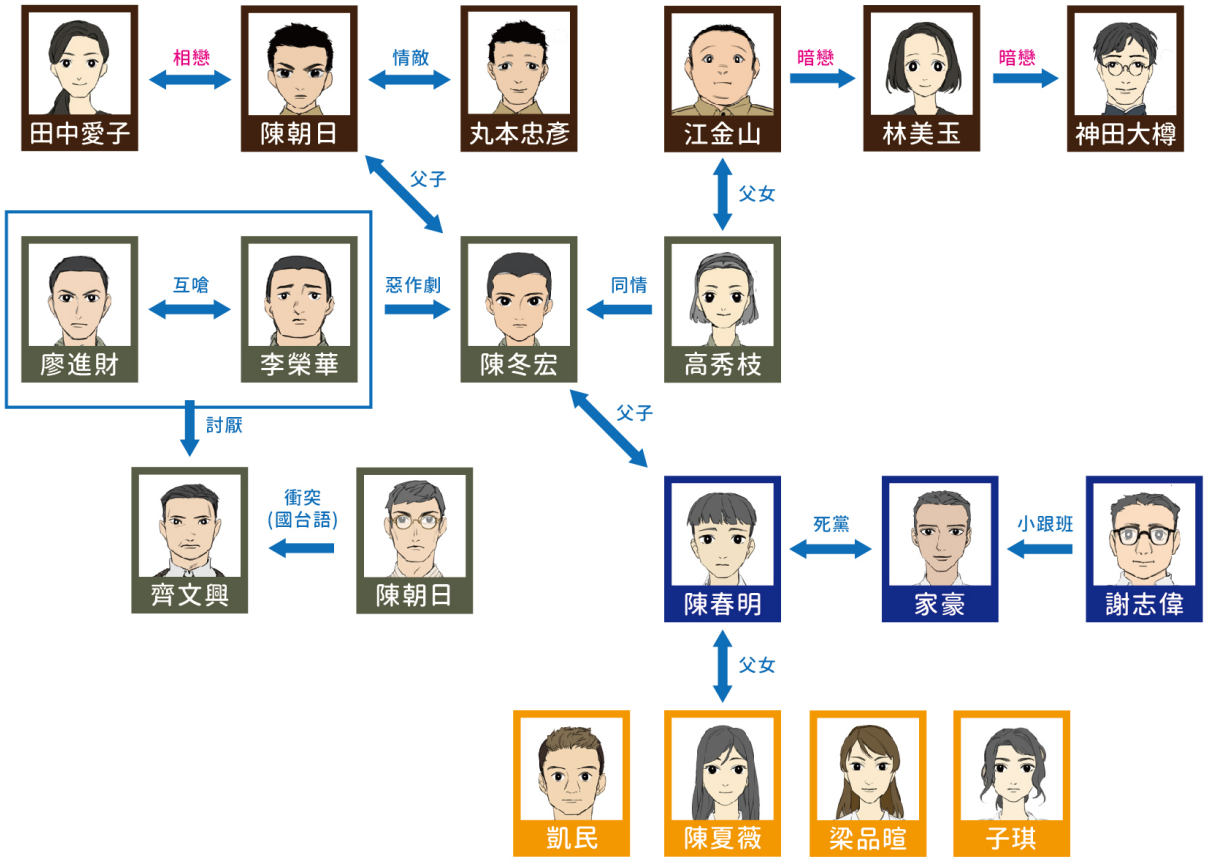


圖 1 憶起上學 XR 體驗內容人物關係圖（本研究提供）

這種獨特的敘事方式使體驗者能夠通過互動過程深入了解台灣不同時代的教育特色，並從中獲得對歷史和文化的全新認識。

3.2. 人物設定與場景建置

確認劇本內容後，本研究考據資料中的老照片，詳細對照不同時期對制服、髮禁等細節的規範進行美術設定，爾後由於 XR 的技術體驗需求上，將允許使用者字由於真實空間中走動與角色互動，因此所有的人物與虛擬場景物件都需要以 3D 建模的方式構成，方能以遊戲引擎透過程式做串接，如下圖 2

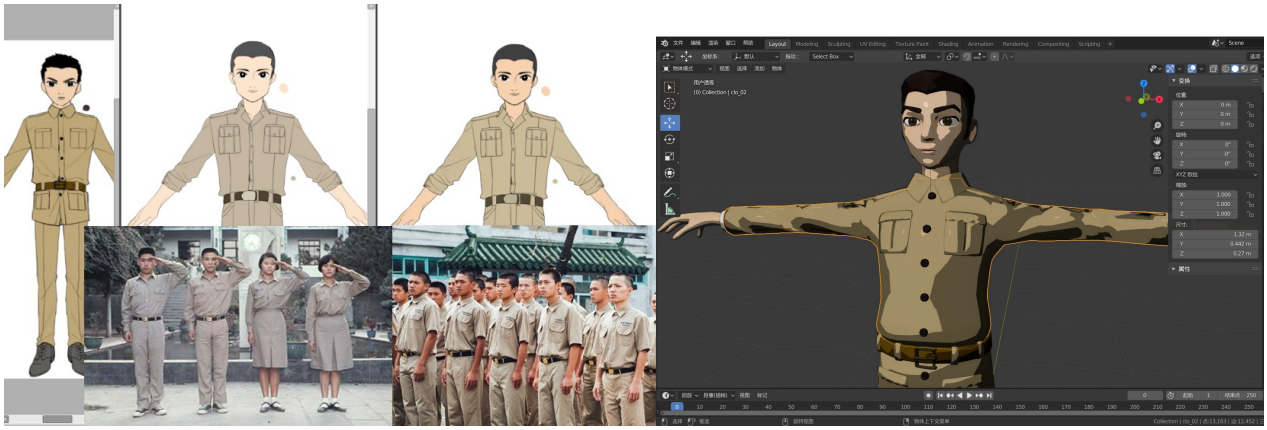


圖 2 透過考據建立人設與模型（本研究提供）

實體展場雖具備基本的模仿教室擺設的桌椅格局，然而為了能讓虛擬和真實場景更加貼合，並能呈現不同時代的環境細節，部分桌椅與教室周遭的擺設物件，亦透過考據當時時代背景下的產物加以添置，並同樣製作 3D 模型供 XR 內容開發使用，如下圖 3



圖 3 場景 3D 模型（本研究提供）

3.3. 動畫製作

由於本研究中的動畫多來自角色，而為了讓不同時代的角色其行為與表現更加自然生動，並考量製作時間之因素，本研究使用了穿戴式的動態捕捉〈Motion Capture, MOCAP〉技術，以不同的演員詮釋不同角色的個性細節，並在不同對話時會有不同的肢體表演，除了提高遊戲的品質，也讓情境能更加沉浸。並使用穿戴式的動態捕捉技術可機動性的轉移場地，讓眾多演員可以在不同時間不同地點進行錄製。

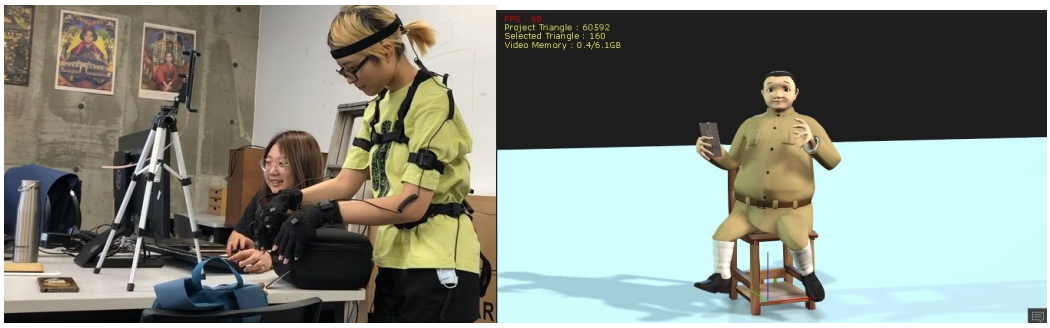


圖 4 透過動態捕捉製作動畫（本研究提供）

3.4. 配音配樂

在故事設定內，不同時期的學生組合不同，亦有不同的語言與腔調，為了能真實還原當時可能的情境，本研究透過專業配音員的培訓，讓參與計畫的成員均接受配音訓練，爾後按照不同情境與角色的個性設定為所有角色配音，其中例如有以日語或原住民腔調發音的角色，便尋找日本籍同學與原住民同學協助，透過此方法盡可能讓情境得以貼近真實。

在配樂的部分，本研究以時常聽到的鐘響旋律作為主旋律，以春夏秋冬四季來形容四個時期的社會氛圍，進而產生四種不同的變奏曲，使學生在體驗 XR 內容時，遊戲感的部分能被提煉出來，增加趣味性也增加情境氛圍。

3.5. 遊戲設計與程式開發

本研究使用遊戲引擎 Unity 做為開發平台，並由於最終以 Apple 的平板電腦 iPad 作為 XR 體驗之載體，因此透過 Arkit 的 SDK 來進行開發，此外，為了讓每位學生在起始遊戲時能和真實場景準確地疊合，本研究在實體展場布置的同時，運用光達(LiDAR)掃描器將展場空間進行 3D 掃描，獲取點雲 (point cloud) 資訊後用以做 XR 虛擬與真實空間的對位基礎，爾後，透過 Arkit 的同步定位與地圖構建 (Simultaneous localization and mapping, SLAM)技術結合物件偵測的算法，並以遊戲機制的引導，在劇情處引導使用者將平板對上指定的實體桌子，待程式辨識到正確的特徵點，便會起始擴增，進入遊戲情境，如下圖 5



圖 5 以遊戲劇情引導使用者正確觸發 XR 體驗（本研究提供）

在遊戲機制的設計上，在四個不同的時期的體驗當下，作為玩家的學生可以手持平板在實體空間中穿梭，在故事中遇到的角色會引導玩家進行不同的任務與對話，其中包含協助角色遞送物品，找到散落在桌上的報紙，或者放在椅子或抽屜裡的物品，因此玩家會在實體空間中來回走動，並透過點擊畫面來進行對話或拾取的動作，此外，本遊戲式實境體驗也設計了彩蛋，將對考據中找到、時期生活上的一些物品，放置在各角落，如果玩家不小心遇到，會發現是可以進行收藏的成就，並且一旦發現的物品，在選單中可以看到詳細介紹，以探索遊戲的情境設計來引發同學的好奇心，並增強其對知識內容的記憶。

3.6. 實體展覽規劃與設計

於本研究中，我們以「憶起上學」為主題，在國家圖書館的穿堂處設計了互動展覽區，為了提升體驗過程的沉浸感，展場以台灣常見的教室的格局為基礎，打造出具有 9 張木製桌椅、一個講台、一個黑板與後方有以互動投影牆構成的佈告欄的展場空間，如下圖 6



圖 6 實體展場空間（本研究提供）

於此外，有鑑於 XR 體驗需要透過裝置來進行，但裝置有數量限制，因希望讓等待體驗的學生也能在實體空間中探索，展場中透過互動科技設計了遊戲式的互動內容，包含互動便當、互動書包、互動抽屜和互動公佈欄。

互動便當以電容感測器與真實的鐵製便當作結合，當使用者接觸到便當時便會發出旋律，而音樂內容是對應過去不同年代的民歌，塑造出符合該時期的氛圍；互動書包則透過結合光敏電阻的設計，當學生打開書包時，便會觸發互動機制，與便當相同是以撥放不同時期的民歌為回饋效果。

互動抽屜使用了樹梅派與霍爾感測器以及電容式觸控螢幕，當學生將抽屜拉出時，抽屜內的觸控螢幕會亮起，內容則是不同時期的課本，並以混合塗鴉的方式，參考網路上許多民眾童年都有塗鴉課本的共同回憶為靈感，雖為塗鴉課本的趣味互動，實而讓學生翻閱不同時期的課本。

互動布告欄則是將圖書館提供，但並沒有情境可以融入 XR 體驗的圖文內容，按照不同的年代與類別分類，使用光達的技術搭配投影製作成互動牆，當學生觸碰有興趣了解的年代或

物件，就會出現對應的知識內容，使等待體驗 XR 的過程也一樣可以進行探索與學習。如下圖 7



圖 7 互動便當、互動書包、互動抽屜與互動牆（本研究提供）

4. 研究結果

展場與 XR 內容設計完成後，由台灣中正高中的歷史老師於課堂中相應的章節進行實，先進行了一場分組討論，引導同學對該段歷史進行省思與辯論，如下圖 8 左，爾後便帶領同學到國家圖書館的展場進行遊戲式實境體驗學習，透過本研究設計與發展的應用進行情境式學習，如下圖 8 右



圖 8 課程引導與 XR 遊戲式實境體驗（本研究提供）

課程結束後完成後，本研究進行了不同主題的訪談，而針對透過 XR 技術結合遊戲式 XR 體驗的設計，獲得同學與老師們的一致肯定，尤其對於實體展場與 XR 內容結合的沉浸感與悅趣感給予高度的評價，而現場空間中的互動設計，也刺激了同學許多想法，表達希望能在其他課程中也能體會到整合 XR 與互動技術的學習方式，肯定其提升學習動機與成效的效果，未來期望有機會擴展到更多類別的課程中。

5. 結論與未來展望

本研究強調了在 XR 遊戲式體驗中建構歷史情境以加強學習成效的可能性，讓體驗者能夠更沉浸的感受並理解歷史上的教育變革。透過 XR 的遊戲式體驗設計曾提升了學生對歷史內容的興趣，並加深了他們對過去社會和教育環境的理解。通過這種可互動和沉浸式的學習體驗，我們希望能夠激發學生對歷史和文化的新認知，也希望進一步因此從而更加珍惜和理解我們的多元文化遺產。此外，基於結合展場的 XR 體驗獲得了師生一致的肯定，為了能讓更多的學生有機會接觸到開發的內容，實體展覽結束後，希望能調整體驗內容的互動的機制，讓其可以於任何空間中自由體驗，造福更多學生，並在此後透過質化觀察法、量化問卷與學習評分進行更進一步的實驗評估。

參考文獻（请参阅 APA 7th 格式）

Carretero, M., Rodríguez-Moneo, M., & Asensio, M. (2012). History education and the construction of a national identity. *History education and the construction of national identities*.

Keleşzade, G., Güneyli, A., & Özkul, A. E. (2018). Effectiveness of history teaching based on social constructivist learning and development of historical thinking skills. *Egitim ve Bilim*, *43*(195).

Greipl, S., Moeller, K., & Ninaus, M. (2020). Potential and limits of game-based learning. *International Journal of Technology Enhanced Learning*, *12*(4), 363-389.

Harley, J. M., Lajoie, S. P., Haldane, C., McLaughlin, B., & Poitras, E. G. (2020). Beyond Historical Books, Names and Dates: Leveraging Augmented Reality to Promote Knowledge, Reasoning, and Emotional Engagement. *Augmented Reality in Education: A New Technology for Teaching and Learning*, 199-216.

Majgaard, G., & Weitze, C. (2020,). Virtual experiential learning, learning design and interaction in extended reality simulations. In *I 14th European Conference on Games Based Learning, ECGBL* (pp. 372-379).

Martin, S. M. (2019). Artificial intelligence, mixed reality, and the redefinition of the classroom. Rowman & Littlefield.

Araiza-Alba, P., Keane, T., Chen, W. S., & Kaufman, J. (2021). Immersive virtual reality as a tool to learn problem-solving skills. *Computers & Education*, *164*, 104121.

Bogusevschi, D., Muntean, C., & Muntean, G. M. (2020). Teaching and learning physics using 3D virtual learning environment: A case study of combined virtual reality and virtual laboratory in secondary school. *Journal of Computers in Mathematics and Science Teaching*, *39*(1), 5-18.

Chan, V. S., Haron, H. N. H., Isham, M. I. B. M., & Mohamed, F. B. (2022). VR and AR virtual welding for psychomotor skills: a systematic review. *Multimedia Tools and Applications*, *81*(9), 12459-12493.

Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, *79*, 5501-5527.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media psychology*, *19*(2), 272-309.

Smiderle, R., Rigo, S. J., Marques, L. B., Peçanha de Miranda Coelho, J. A., & Jaques, P. A. (2020). The impact of gamification on students' learning, engagement and behavior based on their personality traits. *Smart Learning Environments*, *7*(1), 1-11.

Martínez-Hita, M., Gómez-Carrasco, C. J., & Miralles-Martínez, P. (2021). The effects of a gamified project based on historical thinking on the academic performance of primary school children. *Humanities and Social Sciences Communications*, *8*(1), 1-10.

Shavab, O. A. K., Yulifar, L., Supriatna, N., & Mulyana, A. (2021). Gamification in History Learning: A Literature Review. In *6th International Conference on Education & Social Sciences (ICESS 2021)* (pp. 254-258). Atlantis Press.

Guo, X., Guo, Y., & Liu, Y. (2021). The development of extended reality in education: Inspiration from the research literature. *Sustainability*, *13*(24), 13776.

Moseikina, M., Toktamysov, S., & Danshina, S. (2022). Modern technologies and gamification in historical education. *Simulation & Gaming*, *53*(2), 135-156.

大腸直腸癌健康教育之 VR 教材開發與評估

A Study on the Development and Evaluation of VR Learning Materials in Colorectal Cancer

劉怡姍[1*]，游君微[1*]，賴阿福[2]
[1]臺北市立大學 資訊科學系 碩士班 研究生
[2]臺北市立大學 資訊科學系 教授
[*] G11116005@go.utaipei.edu.tw
[*] G11116007@go.utaipei.edu.tw
[*] laiahfur@gmail.com

【摘要】 本研究運用虛擬實境技術開發大腸直腸癌健康教育 VR 教材，以教導學生關注健康議題。教材包含三個遊戲關卡，第一關「飲食偵查」玩家需區分健康和不健康食物；第二關「消滅病毒」需要清除腸道中的病毒；第三關「驅逐之戰」要清除腸道糞便。完成之 VR 教材邀請專家試用及評估，結果顯示教材操作介面流暢，能提高學習趣味性，有助學生獲得大腸直腸癌知識。本研究遵循聯合國永續發展目標，透過科技強化學生健康意識，實現健康與教育的普及。期待教材未來能應用於校園教學，進一步探討虛擬實境在健康教育的成效。

【關鍵詞】 大腸直腸癌；健康教育；數位遊戲式學習；虛擬實境教材

Abstract: *This study utilized virtual reality technology to develop a VR educational material for colorectal cancer health education, aimed at teaching students to pay attention to health issues. The material includes three game levels: the first level, "Diet Investigation," requires players to distinguish between healthy and unhealthy foods; the second level, "Virus Elimination," involves clearing viruses in the intestines; the third level, "Expulsion Battle," involves clearing fecal matter from the intestines. Experts were invited to test and evaluate the VR material. The results showed that the material has a smooth user interface, can enhance learning interest, and helps students acquire knowledge about colorectal cancer. This study adheres to the United Nations Sustainable Development Goals, enhancing students' health awareness through digital technology, and achieving the popularization of health education. The material will be applied to campus teaching in the future, and further explore the effectiveness of virtual reality in health education.*

Keywords: Colorectal Cancer, Health Education, Digital Game-based Learning, Virtual Reality Educational Materials

1. 前言

2015 年聯合國提出永續發展目標（Sustainable Development Goals, SDGs）發表了 17 項全球邁向永續發展的核心目標及 169 項相關具體目標，藉此引領政府、地方政府、企業、公民團體等行動者，主要致力於實現經濟、社會、環境三個層面的永續發展，這些目標是不可分割的、全球性且普遍適用的，以尊重、包容角度考慮各國家具體情況，目標願景將持續至 2030 年（United Nations [UN], 2015）。可見，反映了在全球化時代面臨的共同挑戰，為了實現永續發展目標，人類須積極採取行動並進行改變。

隨著外送平台興起，讓外食文化再升級，然而，每一次的美食饗宴背後卻藏著莫大的罹癌危機。世界衛生組織（World Health Organization）於 2023 年報導，大腸直腸癌是全世界第三常

見的癌症，約占所有癌症病例的 10%，它也是全世界癌症相關死亡的第二大原因；根據 2020 年的統計數據顯示，全球大腸直腸癌的新病例數估計超過 190 萬例以及死亡病例超過 93 萬例。預估到 2040 年，每年的大腸直腸癌負擔將增加至 320 萬例新病例，死亡病例則預計增至 160 萬例（World Health Organization，2023）。而大腸直腸癌有許多的危險因子，多半和生活作息有關，包含肥胖、高脂肪食物、紅肉、低纖維飲食、抽菸及飲酒等（健康世界編輯部，2020）。因此，在這個忙碌的城市裡，無論老幼都常常為了生活忽略了自己的健康，讓我們在熱愛美食的同時，也不禁對所見所聞的警訊深感擔憂「大腸癌」這個可怕的疾病正逐漸成為臺灣人最大的健康威脅之一。

為因應罹癌率上升，本研究運用 CoSpaces EDU 平臺建置虛擬實境，以熱門的 VR 遊戲方式呈現大腸癌衛教宣導。透過了解腸道狀態、選擇健康飲食，旨在創新有趣的方式中強化對大腸癌的認識，同時打破健康與美食的不平衡。

2. 文獻探討

2.1. 虛擬實境（Virtual Reality，簡稱 VR）

隨著科技的快速發展，擴增實境（AR）和虛擬實境（VR）已成為當前熱門的話題。關於元宇宙在教育上的應用，Mystakidis（2022）提到 STEM 教育是最早利用 AR 和 VR 進行教學，並取得顯著成效的領域之一，研究結果顯示，元宇宙技術能夠提高學習速度、增強學習表現，並能促進記憶力，為教育帶來了新的可能性。

VR 是一個高度互動性的 3D 虛擬空間環境，可以不受地域及時間的限制（Petersen, Petkakis, & Makransky, 2022），學生可以透過 3D 顯示頭盔及感測器以更有趣的方式與虛擬世界中的物件進行互動（Elmqaddem, 2019），這樣的學習環境與傳統學習環境相比之下，更具有激勵性且增加參與度，可以將抽象內容可視化，靈活的去做學習（Meharkhe, Schutte, & Kamphorst, 2022），故本研究透過 VR 讓學生可以與平常看不見的大腸腸道進行互動，探索各種不同的場景和學習物件，了解大腸直腸癌發生的原因及有效的預防這項疾病。

2.2. 數位遊戲式學習（Digital game based learning）

由於資料通訊快速進步，電腦、平板和手持行動裝置等所帶來的便利，已經與我們密不可分，而學習也不再像過去一樣，呆板地於室內講課，各種新型態的教育方式也隨著科技有了諸多變化，傳統的教學與學習已不能發揮最大的學習成效，我們必須放棄侷限在以前的大班講述教學，邁向多元不同的教學法選擇（楊雅雯，2017）。而數位遊戲式學習提供互動式體驗和學習環境，逐步的傳達學習概念並引導學生朝往目標前進，因此數位遊戲式學習被視為一種適當的教育工具（Anastasiadis, Lampropoulos, & Siakas, 2018）。

Prensky（2001）提及遊戲式學習應該結合學習內容和遊戲動機，激發學習動機，提高學習成效。成功的遊戲式學習策略蔡仕筌等人（2018）列出了三個觀點：（1）學習者必須沉浸於數位學習遊戲中，且要與場景遊戲有互動性，讓學習者有動力去完成關卡（2）設計出簡易適中的遊戲關卡，引發學習者學習興趣（3）學習目標為遊戲式學習真正的價值，提供學習者修正行為及強化個人技能。故本研究以這三個觀點的角度出發，透過 VR 的輔助，達到沉浸式學習，並且設計不同的遊戲關卡，引導學生吸收相關大腸直腸癌健康知識，提高學習成就及興趣。

2.3. 永續發展目標（Sustainable Development Goals, 簡稱 SDGs）

2015 年聯合國提出 SDGs 的重要性在於推動全球永續發展，並為實現人類共同目標提供了指導與詮釋架構，其涵蓋了經濟、社會和環境三個層面，旨在消除貧困、保障食品安全、確保健康及促進教育等方面的發展（曾文鑑，2023）。本研究參照 SDGs 永續發展第三點–良好

健康與福祉，致力於將科技融入至大腸癌衛教宣導、正確了解飲食習慣對健康的重要性，如下表 1 所述。

表 1 本研究參照 SDGs 永續發展之規劃

SDGs 指標	SDGs 細項	本研究目標
社會生活	3.4「減少非傳染性疾病造成的死亡」	虛擬實境遊戲方式，輔助大眾的健康知識，減少大腸直腸癌的好發速率。
經濟發展	3.D「完善全球健康風險預警」	透過科技技術，讓全民健康進行掌握和管理，實行超前部屬並深入認識。

3. 研究方法

3.1. ADDIE 教學模式

本研究透過「ADDIE 教學設計模式」進行大腸直腸癌教材設計規劃，另外邀請專家實際操作教材並做問卷分析，了解實際操作感受及遇到的問題。ADDIE 教學設計模式主要分為五個階段，分析（Analysis）、設計（Design）、發展（Development）、實施（Implementation）、評鑑（Evaluation），以下分別做課程規劃及分析。

3.1.1. 分析階段

針對健康教育擬定教師訪談內容，蒐集資料後進行整理及分析。據前言說明，世界衛生組織（World Health Organization）研究數據顯示，全球大腸直腸癌罹患率甚高；查找相關文獻時發現鮮少有以大腸腸道為出發點的健康教育教材，有關大腸直腸癌預防方式的研究，幾乎都聚焦於手術實踐或減輕檢查疼痛的引導等方面。為了確認學生對於預防大腸直腸癌的健康知識，會透過課程前測問卷進行評估，用以確保教學內容是否符合學習成效。因此，透過文獻探討 2.2 數位遊戲式學習，結合學習內容和遊戲設計關卡，使學生激發學習動機。

3.1.2. 設計階段

製定國小健康教育課程目標及教學架構,透過教師引導學生了解目前大腸直腸癌對於人類的威脅狀況並且開發主要學習情境。使用 VR 教材之前，教師先與學生討論飲食習慣及生活習慣，透過新聞時事和醫療保健知識，與學生一起探討腸道健康。課程中後期，適時加入 VR 教材，透過趣味性互動和挑戰性學習任務，加深學生對大腸直腸癌疾病的認識，最後課程結束前，對整體課程流程進行總結，回顧學習成果並進行成效評估，同時，提醒學生積極攝取健康飲食、養成運動習慣，對於自我監測、早期篩檢和風險預防意識的傳播是必要性的，才能有效降低致癌機率。

3.1.3. 發展階段

利用 CoSpaces EDU 網頁平臺開發大腸直腸癌健康教育之 VR 教材，輔助學生可以在虛擬環境中建立大腸直腸癌衛教知識，根據陳哲民等（2006）研究發現，大腸直腸癌的關鍵在於大腸瘜肉，其中 75%的大腸直腸癌是偶發性，飲食習慣上的風險因素，如果不及時處理，容易導致大腸黏膜病變並進而形成瘜肉對患者的健康造成極大的威脅，且大腸癌有越漸年輕化趨勢，所以早期篩檢及早期治療顯得格外重要。本 VR 教材依據上述研究進行開發，主要分為三大關卡，詳細於 4.1 研究結果展示。學生透過遊戲獲取大腸直腸癌防治與認知度提升，生動的視覺和互動體驗，進一步加深學生觀念。

3.1.4. 實施階段

未來將教材融入至國小健康教育課程當中，目前由專家試用與評估，收集相關意見及問題，並分析遊戲內容，及時做優化與修正，將來進行教學實驗，以驗證教材之學習成效。

3.1.5. 評鑑階段

本研究教材設計完成後，根據教學目標和學生學習需求，並且依專家實際使用的反饋和問卷評估，進行修正，以提供更好的教學體驗和學習效果。依單一樣本 t 檢定去做分析，了解教材問題點，進一步明確改進和提升教材內容。未來方面，在課程結束前，給予學生課程問卷後測，以評估學生對於課程了解的程度及學習狀況。茲將此研究 ADDIE 教學設計模式之流程圖整理如圖 1。

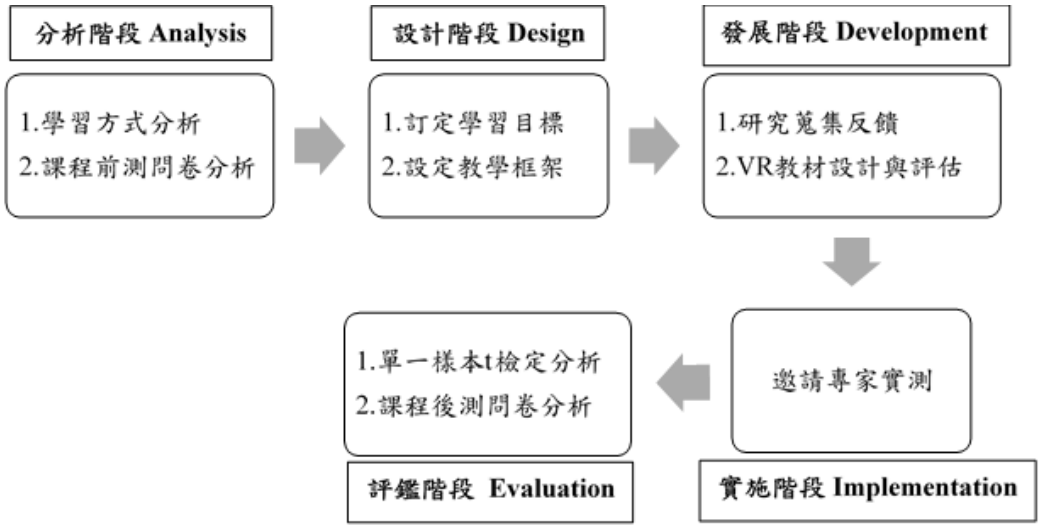


圖 1　ADDIE 教學設計模式之流程圖

3.2. 研究工具

本研究透過 Google Forms 表單編製評估問卷，主要分為六個向度「使用本 VR 教材前」、「本 VR 教材之操作介面設計」、「本 VR 教材之幫助性」、「本 VR 教材之學習效益」、「本 VR 教材之學習價值」、「使用本 VR 教材後」，並使用李克特量表（Likert Scales）分為五個等級的選項，從強度 5（非常同意）、強度 4（同意）、強度 3（普通）、強度 2（不同意）、強度 1（非常不同意）衡量使用者認同程度和客觀評價。

依據 Nunnally（1978）的觀點來看，α 大於 0.7 表示此問卷有高信度。從表 2 來看，各向度的 Cronbach's Alpha 值皆大於 0.7，整體的 Cronbach's Alpha 值為.961，顯示出此問卷的信度良好，適合給專家做評估之用。

表 2 VR 教材專家評估結果之 Cronbach's Alpha 統計

五大向度	題目總數	Cronbach's Alpha值
使用本VR教材前	4	.745
本VR教材之操作介面設計	4	.902
本VR教材之幫助性	5	.917
本VR教材之學習效益	4	.916
本VR教材之學習價值	5	.913
使用本VR教材後	7	.945
整體	29	.961

4. 研究結果

4.1. 大腸直腸癌健康教育 VR 教材成果

本研究在虛擬實境環境採用 CoSpaces EDU 網頁平台建置，透過探索大腸癌的嚴重性及攝取健康的飲食，利用科技的輔助，吸引學生遊玩，達成宣導的效果。第一關「飲食偵查」腸道上會遇到各種食物（如下圖 2 所示），有些食物可能增加罹患大腸癌的風險，透過吸收健康食物而得分，若遇到較不健康的食物時，遊戲會挑戰你的健康知識，帶出相關的衛教問題（如下圖 3 所示），需要回答正確才能繼續前進到下一個階段；進入第二關「消滅病毒」，將目睹腸道中出現的各種病毒和瘜肉（如下圖 4 所示），它們是潛在的健康威脅，此關的任務是利用手術刀，一個一個地將這些有害元素割除，清理腸道，並確保身體免受這些潛在威脅的侵害，每一個成功清理的病毒或瘜肉都將獲得額外的健康積分；最後一關「驅逐之戰」，根據你所吃下的食物和消除病毒後，這裡是大腸中無法吸收而產生糞便的地方(如下圖 5 所示)，須將這些糞便清除，同時隨著身體的自然運作，虛擬腸道將與這些病毒一同排除體外（如下圖 6 所示），象徵著身體的免疫力得以戰勝潛在威脅，最後，順利離開腸道(如下圖 7 所示)。

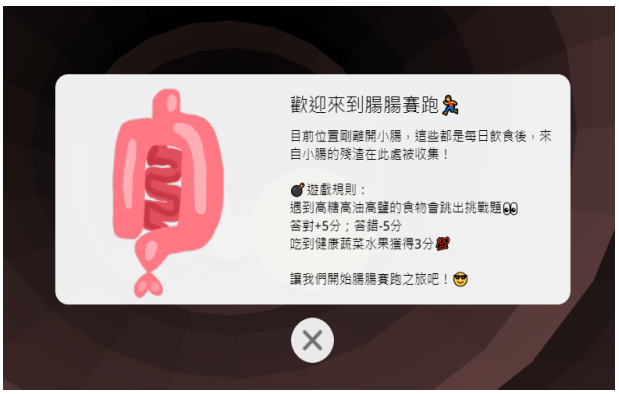


圖 2　VR 教材–第一關「飲食偵查」



圖 3　VR 教材–相關的衛教問題



圖 4　VR 教材–第二關「消滅病毒」



圖 5　VR 教材–第三關「驅逐之戰」



圖 6　VR 教材–腸道衛教宣導



圖 7　VR 教材–破關成功畫面

4.2. VR 教材評估結果

本研究完成大腸癌衛教宣導之 VR 教材後，邀請了共 25 位專家進行試用及評估，評估資料採用統計軟體 SPSS 作為分析工具，透過描述性統計及單一樣本 t 檢定（檢定值為 3）分析問卷數據，並運用 Cronbach's Alpha 檢測問卷信度，專家評審結果總表如下表 3。

「使用本 VR 教材前之評估分析」共有四題，平均值介於 3.36~4.40，向度平均值為 3.86，標準差為.746，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向度平均 t 檢定統計量為 5.758，p 值<.001，達顯著水準，表示專家對於本 VR 教材「使用前之評估分析」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

「本 VR 教材之操作介面設計」共有四題，平均值介於 3.72~4.36，向度平均值為 4.00，標準差為.869，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向度平均 t 檢定統計量為 5.754，p 值<.001，達顯著水準，表示專家對於「本 VR 教材操作介面設計」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

「本 VR 教材之幫助性」共有五題，平均值介於 4.28~4.40，向度平均值為 4.34，標準差為.662，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向

度平均 t 檢定統計量為 10.149，p 值<.001，達顯著水準，表示專家對於「本 VR 教材之幫助性」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

「本 VR 教材之學習效益」共有四題，平均值介於 4.12~4.28，向度平均值為 4.21，標準差為.846，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向度平均 t 檢定統計量為 7.144，p 值<.001，達顯著水準，表示專家對於「本 VR 教材之學習效益」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

「本 VR 教材之學習價值」共有五題，平均值介於 4.08~4.28，向度平均值為 4.16，標準差為.810，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向度平均 t 檢定統計量為 7.157，p 值<.001，達顯著水準，表示專家對於「本 VR 教材之學習價值」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

「本 VR 教材使用後之評價分析」共有七題，平均值介於 4.16~4.48，向度平均值為 4.37，標準差為.698，顯示專家認同此大腸癌衛教宣導之 VR 教材。從單一樣本 t 檢定（檢定值為 3）結果顯示，向度平均 t 檢定統計量為 9.815，p 值<.001，達顯著水準，表示專家對於「本 VR 教材使用後之評價分析」同意者與不同意者有顯著差異，並且同意者顯著高於不同意者。

表 3 專家評審結果的描述性統計及單樣本 t 檢定總結

	向度平均值	標準差	t(24)
使用本VR教材前	3.86	.746	5.758***
本VR教材之操作介面設計	4.00	.869	5.754***
本VR教材之幫助性	4.34	.662	10.149***
本VR教材之學習效益	4.21	.846	7.144***
本VR教材之學習價值	4.16	.810	7.157***
使用本VR教材後	4.37	.698	9.815***

****p<.001*

綜合以上，專家們認為本大腸癌衛教宣導 VR 教材具有高度價值，且顯示 VR 教材得到一致好評，特別強調互動性、衛教宣導的流暢度和學習場景增加吸引力，透過遊戲關卡，可以提高學習的趣味性及學習成效。

5. 結論

本研究在評估教材前，評估者對於大腸癌有一定的知識與了解占大多數，體驗遊戲過後，由專家評估結果，獲得肯定，皆認為此教材能適合做為學習者之用，而本 VR 遊戲教材體驗感受能有效地提升大腸直腸癌知識學習成效，且 VR 技術能讓學習者再接近真實性環境，提升學生學習相關大腸癌保健的意願及態度，適合融入教學現場，做為輔助課堂學習教材之使用。

在執行本研究時，我們利用長時間研究腸道的健康知識，和專業人士討論，得到最正確和精準的知識運用，後續將 VR 教材設計遊戲關卡，思考如何讓遊戲體驗融入學習內容，將遊戲設計符合學生學習興趣，讓原本照本宣科的健康課程變得生動活潑，加強學生的健康認知及意識。

未來期望可以進一步進行校園教學實驗，打造具教育意義的腸胃健康指南，透過生動的視覺和互動體驗，探討 VR 技術在大腸直腸癌的校園衛教宣導之教學成效，也能探討教師在教學中運用此 VR 教材之情況，了解教師對於 VR 技術的應用意見和困難，並提供相應支持和培訓。

最後，本研究期望藉此完成 VR 教材，強化大腸癌議題的關注度與實際認知的教育性，以真實互動性、遊戲式學習，並遵循 SDGs 永續發展目標帶給學生正確的腸胃健康知識。

參考文獻

曾文鑑（2023）。SDGs 與學校課程和教學的相遇。**師友雙月刊，（639）**，25-32。

陳哲民、李熹昌、林光洋、陳建志、郭冠良、黃惠娟、吳岱穎（2006）。大腸直腸癌篩檢。**北市醫學雜誌，3（1）**，1-10。doi:10.6200/TCMJ.2006.3.1.01

健康世界編輯部（2020）。大腸癌是最可預防的癌症－多運動、多蔬果、少紅肉創造健康生活。**健康世界，（519）**，49-50。

楊雅雯（2017）。玩中學－數位遊戲式學習。**臺灣教育評論月刊，6（9）**，300-302。

蔡士荃、蔡銘修（2018）。淺談數位遊戲式學習與 APP 設計應用實施。**台灣教育評論月刊，7（5）**，229-233。

Anastasiadis, T., Lampropoulos, G., & Siakas, K. (2018). Digital game-based learning and serious games in education. *International Journal of Advances in Scientific Research and Engineering, 4(12),* 139-144.

Elmqaddem, N. (2019). Augmented Reality and Virtual Reality in Education. Myth or Reality?. *International Journal of Emerging Technologies in Learning, 14(3),* 234-242.

Meharkhe, J., Schutte, P., & Kamphorst, B. (2022). Learning in Virtual Reality: Bridging the Motivation Gap by Adding Annotations. *In Proceedings of Conference on Human Factors in Computing Systems (CHI '22)*. New York, NY: Association for Computing Machinery.

Nunnally, J. C.（1978）. *Psychometric Theory*. New York: McGraw-Hill.

Petersen, G. B., Petkakis, G., & Makransk, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education, 179,* 104429.

Prensky, M. (2003). Digital Game-Based Learning- Book02. *ACM Computers in Entertainment,1*(1), 1-4.

基于单元整体教学的编程游戏化学习活动设计

——以高中《算法与问题解决》单元为例

Programming gamification learning activity design based on the holistic perspective of large units

-- Take "Algorithm and Problem Solving" unit in high school as an example

林肖颖[1]，赵晨畅[2*]
[1]杭州师范大学经亨颐教育学院
[2]杭州师范大学经亨颐教育学院
[*] 2022112004125@stu.hznu.edu.cn

【摘要】 通过问卷调查和访谈表明，传统的高中编程教学存在兴趣不足等问题，不能较好地培养学生的核心素养。而游戏化学习结合大单元教学，能够激活学生思维，落实核心素养。本研究梳理了游戏化学习、大单元整体教学的发展，拟将高中编程内容、游戏化设计与大单元教学相融合，为高中生引入游戏化学习方式。
【关键词】 信息技术；游戏化学习；游戏化学习活动；算法；大单元教学

Abstract: *Through the questionnaire survey and interview, the traditional high school programming teaching has some problems, such as lack of interest, and can not cultivate the core quality of students. The combination of gamified learning and large unit teaching can activate students' thinking and implement their core qualities. This study combs the development of gamified learning and large unit integrated teaching, and intends to integrate high school programming content and gamified design with large unit teaching to introduce gamified learning methods for high school students.*

Keywords: Information technology, Gamified learning, Gamified learning activities, Algorithm, Large unit teaching

1. 引言

2017 年，国务院印发《新一代人工智能发展规划》，提出“实施全民智能教育项目，在中小学阶段设置人工智能相关课程，逐步推广编程教育”。通过问卷调查和访谈表明，传统的高中编程教学存在兴趣不足、动机较低、课程乏味等问题，使得教学效率低，不能较好地培养学生的核心素养。游戏化学习作为落实核心素养的有效学习策略，更容易激发学习者的学习兴趣和学习动机，并有助于培养学习者计算思维等高阶思维能力。(张学军　等，2021；张屹　等，2022)且游戏化学习强调团队合作，注重师生的双向交互作用，能够深化学生与学习内容的联系，培养学生的核心素养。但是，研究表明游戏化教学应用主要集中于初中和小学阶段，高中阶段应用较少。(屈志伟　等，2023)大单元教学创设情境，能够激活学生思维，促进学生的参与，落实核心素养。(任明满，2022)《普通高中信息技术课程标准(2017 年版)》中也提出重视以学科大概念为核心，使课程内容情境化。因此，本研究拟将高中编程课程内容、游戏化设计与大单元教学相融合，为高中生引入游戏化学习方式。

2. 文献综述

2.1. 游戏化学习

游戏化(Gamification)是在非游戏环境中使用游戏设计的元素，目的是提升学生的兴趣和参与度。游戏化学习是游戏化的概念在教学领域的实践，前身是基于游戏的学习(Game-Based Learning,GBL)、教育游戏(Education Game)和严肃游戏(Serious Game)(鲍雪莹等，2015)。Deterding 等人明确指出，游戏化学习的特点包括:倾向于使用游戏的核心设计元素；更关注游戏元素的应用；其目的是利用游戏的教育属性；不局限于特定的目标和场景(DETERDING，et al, 2011)。这对理解和实践游戏化学习有着重大的启示。现阶段，高中信息技术教学中的游戏化学习还处于发展阶段，研究者看重的是如何在教学环境中巧妙地利用游戏精神，吸引学生的注意力，带来优质的教学效果。

从 1970 年代至今，游戏化学习的设计理念经历了从行为主义和认知主义，再到建构主义的发展过程。这一过程本质上也贴近百年来教育目标的变化过程——从机械学习陈述性知识，到主动理解概念性知识，再到运用已有知识开展社会与科学实践(曹鹭，2022)。游戏化学习理论主张在游戏的场景和任务中促进探索性学习，利用游戏的趣味性和竞争性，以增强他们的学习动力以深入学习。游戏化学习的相关理念如表 1 所示。

表 1 游戏化学习相关理念

相关理念	含义
学习者中心	游戏化学习让学习者成为活动的中心，鼓励他们主动参与和探索，以提高学习效果。
动机驱动	游戏化学习活动通常利用激励机制（例如：积分、徽章、排行榜等）来提高学习者的参与度和学习动力。
持续反馈	即时反馈是游戏中的重要机制，它让玩家了解自己的表现，调整策略，学习效果变得明显。这一原则在游戏化学习中也占据重要地位。
安全的失败	在游戏中，失败通常被看作是学习的机会，可以通过重试、反思和探索新的策略来改善。这个原则也适用于游戏化学习，让学生在失败中发现问题、提出解决方案并自我提升。
自主学习	游戏化学习可以增加学生的自主性，他们可以自选学习路径，掌握自己的学习节奏，从而满足他们的个体化需要。
合作竞赛	游戏化学习可以创设充满竞争和合作的学习环境，从而促进学生的互动交流和团队合作。
非线性学习	像游戏一样，知识点可以是非线性和模块化的，学生可以根据自己的兴趣和需要，选择适合的模块进行学习。
多元智能	游戏化教学能够迎合学习者的多元智能，例如逻辑智能、空间智能、人际智能等，让所有学生在学习过程中找到属于自己的优势和乐趣。

2.2. 信息技术学科中的游戏化教学

国外的游戏化教学方面的研究相对较为成熟。Yallihep 等人通过实证表明严肃游戏能更好地提升学生的编程能力(YALLIHEP, et al, 2020)。Koupritzioti 等人以迷宫教育游戏为基础展开实践，发现该游戏可供学生完成编程基础知识的学习(KOUPRITZIOTI，et al, 2020)。相比之下，

我国在游戏化教学研究方面起步较晚。蒋宇等人以真实情境和游戏学习为导向，提出了包括自主学习、合作探究学习、总结分享三个阶段的游戏化探究学习模式(蒋宇 等，2011)。程佳铭等人在需求满足游戏体验设计框架的基础上对异步在线课程内容进行了游戏化教学设计发现该设计能够对中学生的互动质量产生积极影响(程佳铭 等，2022)。

2.3. 大单元整体教学

单元整体教学最初起源于欧美“新教育运动”，主张知识学习应当是完整的，并与实际活动相结合，形成较为完整的学习单元。而后，杜威提出学生应当明确学习内容与实际情境的对应关系，自主构建具有综合性的学习单元。克伯屈在此基础上提出了“设计教学法”，倡导大型学习单元，避免了教学碎片化。我国对单元整体教学的研究始于对语文学科的探讨，并逐渐扩展到其他学科的研究领域。钟启泉认为单元教学设计是促使课堂教学转型的支点，是提升学生学科素养和综合能力的重要路径。

3. 基于单元整体教学的编程游戏化学习活动设计

3.1. 设定学习目标，指向核心素养

《普通高中信息技术课程标准(2017 年版)》中提出高中信息技术学科核心素养由信息意识、计算思维、数字化学习与创新、信息社会责任四个核心要素组成。

3.2. 分析学习内容，创设情境任务

学习内容的设计明确了学生在学习之后必须掌握的知识范围与能力要求，按照学习者的认知发展水平呈现游戏任务。因此在设计游戏化学习活动之前，要深入理解与剖析信息技术课程，以便科学的选择最合适的学习内容。

3.3. 设计教学活动，融入游戏化元素

Garris 等认为游戏要素包括反馈、幻想性、规则/目标、感官刺激、挑战、神秘感、控制、社会性、故事情节等。本文设计了教学活动，融入了各种游戏化元素。

4. 基于单元整体教学的编程游戏化学习活动设计案例

本研究以浙教版高中信息技术《算法与问题解决》为例设计游戏化学习活动。

4.1. 设定学习目标，指向核心素养

本单元旨在培养学生具备以下能力：首先，能够在日常生活和学习中主动发现算法的存在，深刻理解算法的内涵和外延。其次，学生应具备根据实际问题进行抽象与建模的能力，进而能够完成算法的设计与描述，将理论知识应用于实际情境中。此外，学生将初步认识算法的多样性，理解不同算法在解决问题时的灵活性和适用性。通过这一学习单元，学生将不仅仅获得对算法的深刻理解，还能够灵活运用算法解决各类实际问题，为未来深入学习信息技术领域打下坚实基础。

信息意识	学生将培养对信息的敏感性，能够主动从日常生活和学习中发现信息，并理解信息对问题解决的关键性作用。通过游戏化教学，学生将参与信息获取、分析和利用的过程，提高信息意识，使其在信息丰富的环境中更加自信和能动。
计算思维	学生将发展计算思维，能够将实际问题抽象为计算机可处理的形式，培养解决问题的逻辑和算法思维。游戏化教学将提供虚拟场景，要求学生在游戏中运用计算思维解决挑战，促进学生在解决实际问题时更具计算思维的敏感性。
数字化学习与创新	学生将学会根据实际问题进行抽象与建模，完成算法的设计与描述，培养数字创新的能力。通过游戏中的创新性任务和挑战，学生将在数字化学习的过程中培养对新思路和解决方案的开放性态度，促使其在数字领域展现创新精神。

信息社会责任	学生将了解信息社会的背景和发展，认识信息技术对社会产生的影响，并培养信息社会责任感。游戏化教学将引导学生在虚拟社会中扮演角色，体验信息决策的社会责任，使其在信息使用中更具责任心和道德意识。

4.2. 分析学习内容，创设情境任务

本单元知识点如图所示：

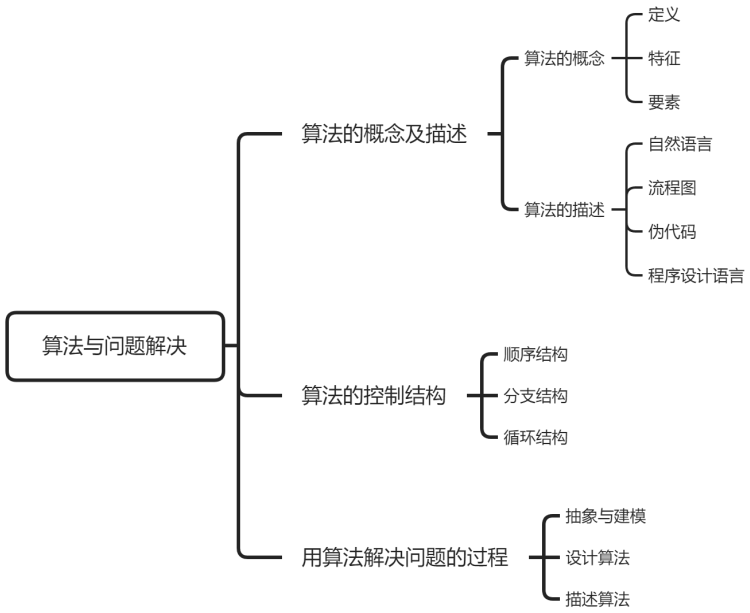


图 1 《算法与问题解决》单元知识点

本单元情境：环游宇宙之蒜蒜星球历险记。“我”和“我”的电子宠物小 Q 在星球旅行的途中，飞行器故障了，掉落在蒜蒜星球，“我”需要学习相关的算法知识修好飞行器，回到地球。

4.3. 设计教学活动，融入游戏化元素

本教学活动以“蒜蒜星球”为主要地点，包括四个游戏场景，分别是陌生星球、迷宫、铁匠铺和飞船，各个场景环环相扣，学生和 NPC 互动对话，在场景中完成相应的游戏活动和学习任务，进而掌握知识点，如表 2 所示。

表 2 游戏场景对应的活动、游戏化元素、知识点

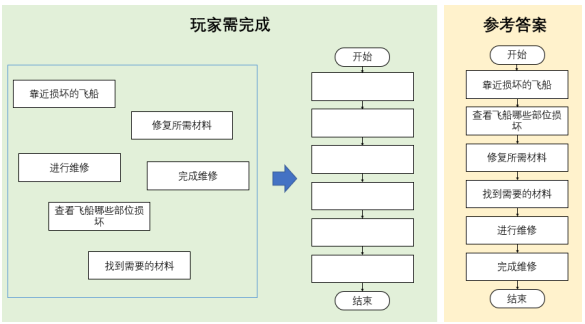
游戏场景	游戏活动	游戏化元素	知识点
场景一：掉落陌生星球	活动一：制定维修飞船的步骤，学习自然语言 活动二：解锁飞船检测仪，学习流程图	幻想性：包含虚构或幻想的情境、主题或人物。 故事情节：通过叙事的方式，将游戏的相关信息以情景化的形式呈现给学习者。	算法的定义 算法的特征 算法的要素 流程图
场景二：进入迷宫，收集元素	活动三：寻找缺失的元素 第一关：元素拐个弯就能拿到——学习顺序结构 第二关：小心分叉路口——学习分支结构	挑战：有明确的目标，采用渐进的难度，游戏过程中持续跟踪的反馈和记分。	算法的控制结构

	第三关：妈妈总是和我说循环的重要性——学习循环结构 终极关卡：夺取元素，逃出迷宫		
场景三：铁匠铺	活动四：制作“棋怪”机器人，学习“IPO 模型” 活动五：购买材料，巩固“IPO 模型”	规则与目标：游戏由清晰的规则、目标以及通往目标过程中的一系列反馈组成。	用算法解决问题的过程
场景四：回到飞船	活动六：闯关答题，总结知识点	成就：对已完成任务的虚拟的表征，例如，徽章、等级等。	

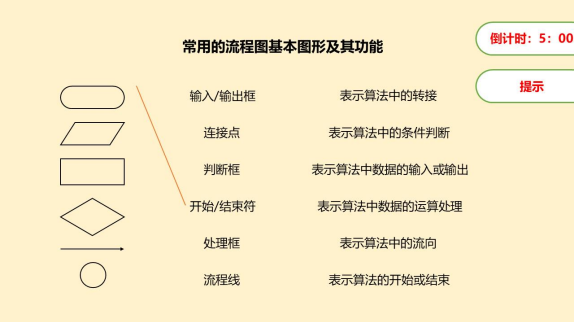
4.4. 游戏化学习活动设计案例

基于上述设计框架，对浙教版高中信息技术《算法与问题解决》进行了游戏化学习活动的整体设计，具体如表 3 所示。

表 3 具体游戏化学习活动设计

游戏脚本	游戏活动	知识点
场景一：掉落陌生星球		
我：这是哪呀？ 小 Q：我们好像掉落在蒜蒜星球。 我的眼前出现了一幕幕在网上看来的关于蒜蒜星球的记忆，这个星球因“算法”著名。 小 Q 打断了我的回忆，说：我们的飞船好像故障了，飞不起来了，怎么办呢？ 我：那就让我们修好飞船，重返地球吧。	【活动一：制定维修飞船的步骤】 根据自然语言，拖动方块，对维修飞船的步骤进行排序。 自然语言：先靠近损坏的飞船，之后查看飞船是那些部位损坏了，需要什么元素材料来进行修复，找到需要的材料，放入维修，维修完成。 	算法的定义 算法的特征 算法的要素
小 Q：那就让我们看看飞船缺什么元素吧。 我打开背包，拿出了飞船元素检测仪。 小 Q：这个飞船元素检测仪怎么被锁住了？（飞船检测仪的屏幕上出现：请完成以下命令） 解锁完毕，开始检测飞船缺失的元素，屏幕上显示	【活动二：解锁飞船检测仪，学习流程图】 自主学习常用流程图基本图形及其功能，完成连线题，解锁飞船检测仪。	流程图

缺少的元素 100 个铁块、十个工具和武器。
小 Q：让我们一起来寻找这些材料吧！

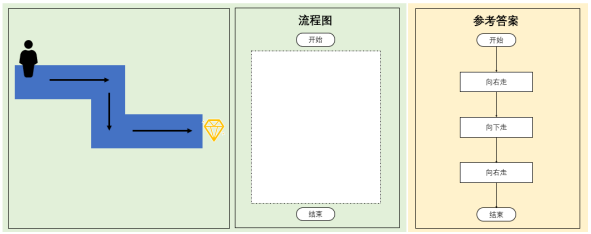


场景二：进入迷宫，收集元素

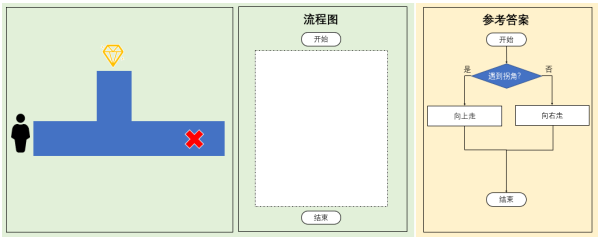
我和小 Q 出发寻找需要的修复元素。
小 Q：飞船检测仪响了，附近一定有我们需要的元素吧！你看，前面出现了一个迷宫！
我们跑向了迷宫入口，飞船检测仪越来越响。
小 Q：看来就是这个迷宫了，我们想找的东西就在里面！
我们走进迷宫。
阿守：你们是谁，好像不是我们星球的？
小 Q：我们需要一些元素来修好飞船。
阿守：想要拿到元素可不是简单的事，必须通过我的考验。
眼前出现了解密界面。

【活动三：寻找缺失的元素】
画出流程图，使得小人能拿到元素。

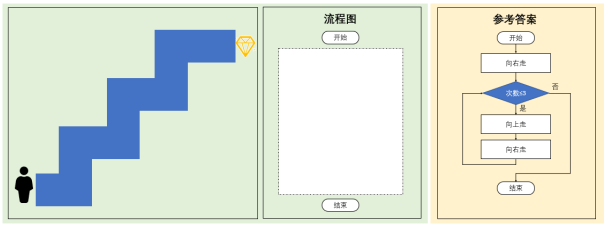
- 第一关：元素拐个弯就能拿到——学习顺序结构



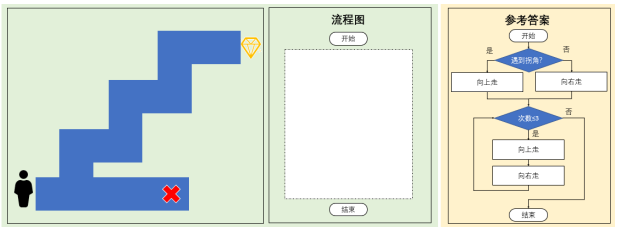
- 第二关：小心分叉路口——学习分支结构



- 第三关：妈妈总是和我说循环的重要性——学习循环结构



- 终极关卡：夺取元素，逃出迷宫



场景三：铁匠铺

阿巧：嗨，新面孔呀。
我做了个下棋的机器人，名叫“棋怪”，我想知道

【活动四：制作“棋怪”机器人，学习“IPO 模型”】

- 填空

算法的控制结构

用算法解决问题的过程

他这几次下棋的胜负如何，不知该如何操作。

小 Q：好神奇的机器人，阿巧你的手可真巧呢！

我：下棋结束后，想让棋怪能记录什么信息呢？

阿巧：我是想如果棋怪赢得了对方则输出：“zhuangshenme”

如果对方赢了输出：“G”

如果平局则输出：“henfan”

小 Q:这个棋怪确实奇怪。我们该如何解决呢？

我：若是用算法的思维来想，输入的是棋怪和对方棋手的胜负，输出的是棋怪的胜负记录，中间过程我们需要判断胜负。

小 Q：我知道，这是大名鼎鼎的“IPO 模型”吧——“输入（input）-处理（process）-输出（output）”。

阿巧：我这里还有一本算法秘籍，讲的就是这些。

我接过算法秘籍并打开，在算法秘籍的帮助下完成活动。

小 Q：为了修好飞船顺利回地球，必须购买 10 样工具和武器。我们有多少钱呀？

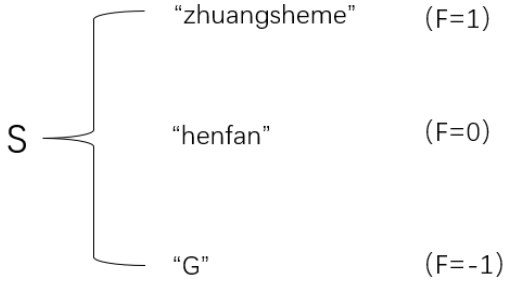
我：我们已经获得了 2000 星特币。

小 Q：那么，最终每样每个工具将购得多少件呢，还有剩余的星特币吗？

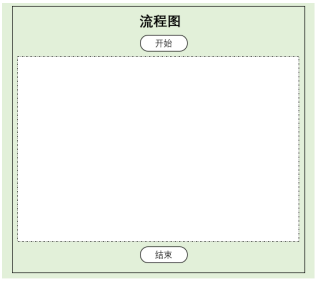
这又回到了用算法解决问题的过程了！你还记得是哪几个步骤吗？

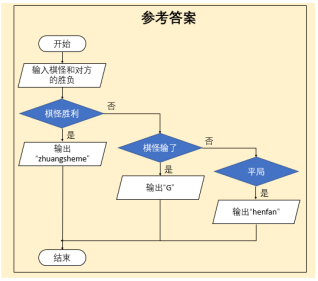
我：抽象与建模，设计算法和描述算法。

已知 n（1≤n≤30）组数据，输出棋怪的胜负话语 S。

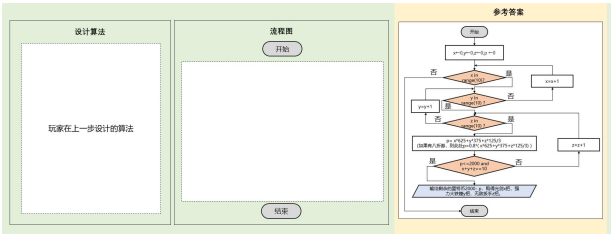


- 完成流程图





【活动五：购买材料，巩固“IPO 模型”】



小 Q：没错！这下我们可熟练啦。	
场景四：回到飞船	
小 Q：这一段旅程经历了不少，收获满满呢。你还记得我们都学了哪些知识吗？ 最后，我们获得了工具和燃料，乘坐着飞船，踏上了返回地球的旅程。	【活动六：闯关答题，总结知识点】

5. 结语

本研究梳理了游戏化学习的发展、相关理论、信息技术学科中的游戏化教学以及大单元教学，将高中编程课程内容、游戏化设计与大单元教学相融合，设计了学习流程，并以《算法与问题解决》为例，展示了该模式的应用过程，希望能为游戏化教学研究、高中信息技术课程教育提供借鉴和参考。

参考文献

鲍雪莹, 赵宇翔, 2015. 游戏化学习的研究进展及展望[J/OL]. 电化教育研究, 36(8): 45-52. DOI:10.13811/j.cnki.eer.2015.08.008.

曹鹭, 2022. 三种基本学习理论与教育游戏的设计[J/OL]. 开放教育研究, 28(05): 29-38+92. DOI:10.13966/j.cnki.kfjyyj.2022.05.004.

程佳铭, 张力力, 2022. 游戏化设计对在线异步讨论生生互动的影响研究[J/OL]. 现代远距离教育(2): 12-23. DOI:10.13927/j.cnki.yuan.20220221.001.

蒋宇, 尚俊杰, 庄绍勇, 2011. 游戏化探究学习模式的设计与应用研究[J]. 中国电化教育(5): 84-91.

屈志伟, 车璐, 刘佳明, 等, 2023. 基于 CiteSpace 的中小学信息技术课程游戏化教学现状研究[J]. 内江科技, 44(9): 117-119.

任明满, 2022. 大单元教学:历史脉络、研究现状及路径选择[J/OL]. 课程.教材.教法, 42(4): 97-105. DOI:10.19877/j.cnki.kcjcjf.2022.04.018.

张屹, 马静思, 周平红, 等, 2022. 人工智能课程中游戏化学习培养高中生计算思维实践的研究——以“挑战 Alpha 井字棋”为例[J/OL]. 电化教育研究, 43(9): 63-72. DOI:10.13811/j.cnki.eer.2022.09.008.

DETERDING S, DIXON D, KHALED R, 等, 2011. From game design elements to gamefulness: defining “gamification”[C/OL]//Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. Tampere Finland: ACM: 9-15[2023-07-07]. https://dl.acm.org/doi/10.1145/2181037.2181040. DOI:10.1145/2181037.2181040.

KOUPRITZIOTI D, XINOGALOS S, 2020. PyDiophantus maze game: Play it to learn mathematics or implement it to learn game programming in Python[J/OL]. Education and Information Technologies, 25(4): 2747-2764. DOI:10.1007/s10639-019-10087-1.

YALLIHEP M, KUTLU B, 2020. Mobile serious games: Effects on students’ understanding of programming concepts and attitudes towards information technology[J/OL]. Education and Information Technologies, 25(2): 1237-1254. DOI:10.1007/s10639-019-10008-2.

促进大学生高阶思维培养的协作推理游戏设计

Collaborative Reasoning Game Design for Fostering Higher-order Thinking in University Students

张颖[1]，李秀晗[2*]，刘巧伶[3]，张晓曼[4]
华中师范大学人工智能教育学部
2071558095@qq.com

【摘要】 本研究对能有效促进大学生高阶思维培养的协作推理游戏设计进行探索。首先，文章明确了高阶思维的概念，并分析大学生在协作学习中思维能力方面存在的问题和挑战。其次，通过文献综述，总结了已有的推理游戏在教育应用中的研究，为游戏设计提供基础。在此基础上，本文提出了一种协作推理游戏的设计框架，并以课程为例实现协作推理游戏的设计，该游戏的设计旨在引导大学生在游戏过程中进行深入思考、推理、分析和判断，从而培养其问题解决、批判性思维和创新能力等高阶思维能力。
【关键词】 推理游戏；协作学习；大学生；高阶思维；游戏设计

Abstract: *This study explores the effective design of collaborative reasoning games to enhance higher-order thinking of university students. Firstly, the article defines the concept of higher-order thinking and analyzes the issue of college students' thinking ability in collaborative learning. Furthermore, through a literature review, existing research on the educational applications of reasoning games is summarized. Building upon this foundation, the paper proposes a design framework for collaborative reasoning games and takes a course as an example to realize it, which is designed to guide college students in engaging in profound reflection, reasoning, analysis, and judgment. The objective is to nurture their higher-order thinking abilities, including problem-solving, critical thinking, and innovation skills.*
Keywords: Reasoning Games, Collaborative Learning, University Students, Higher-order Thinking, Game Design

基金项目：2023 年湖北省本科高校省级教学改革项目，智慧教学环境下自主设问促进大学生高阶思维发展的研究与实践；2023 年华中师范大学校级教学改革项目，智慧教学环境下自主设问促进大学生高阶思维发展的研究与实践。

1. 前言

1.1. 研究背景

在当今迅速变化的学习环境中，高阶思维对学生适应这样的学习方式和实现认知平衡具有积极影响，是当代教育和教学改革中备受关注的焦点。但目前大学生在协作学习中的思维能力面临很多的问题和挑战，如独立思考的能力、批判性思维能力、创新能力、协作能力不足等。而教育游戏能很好改善这个现状。教育游戏思想的实质就是尊重学习者的主体性，尊重人发展的自然规律，实现真正的“寓教于乐”。且相关研究也表明将思维能力的培养与游戏化的学习情境进行结合可以收到事半功倍之效。基于此，本研究设计了一个协作推理游戏，以《教学系统设计理论与实践》课程为例， 探究推理游戏在协作学习中的设计方法，从而能够引导大学生在游戏过程中进行深入思考、推理、分析和判断，培养高阶思维。

1.2. 高阶思维

高阶思维最早起源于 1956 年布鲁姆提出的认知目标分类法，包含知道、领会、应用、分

析、综合和评价六个层次，其中分析、综合和评价是高阶思维的表现形式。目前钟志贤教授对高阶思维的定义在国内认可度较高，他将高阶思维看作一种能力，把高阶思维定义为发生在评价、分析、创造等高水平认知活动中的，以创造性思维能力、批判性思维能力、问题求解能力和决策能力为主的高层次思维(钟志贤, 2004)。

2. 教育游戏对高阶思维培养的研究现状

高阶思维培养的研究一直是教育领域的一个重要议题，已经有一些研究将教育游戏作为一种途径来培养学生的高阶思维。Yang 开发了虚拟 CEO 游戏，并将其应用于“商店规划与管理”课程中，并表明该游戏可以用于提高职业高中学生的问题解决能力、批判思维和创造思维能力(Yang, 2015)。Huang 等人的研究也得出相似的结论，他们开发一个商业模拟游戏并在商业大学中开展研究（Huang, Silitonga, & Wu, 2022）。除了商业专业，有研究以工程为主体，设计了能促进工程专业的学生更高阶的学习的计算机游戏(Siddique et al., 2013)。以上可观，目前针对高阶思维培养的教育游戏是为特定专业而设计的，对于可推广性是受限的，因此本研究设计的不限学科的协作推理游戏框架是有必要的。

3. 协作推理游戏的设计与实现

3.1. 游戏设计

协作推理游戏的重要特征是将推理机制应用于协作学习中。该游戏是在建构主义理论、情境认知理论基础上设计的，利用游戏化学习情境来实现学习者轻松学习的目的。本研究的协作推理游戏的设计框架如图 1 所示。该游戏设计框架将设计流程分为学习目标分析、学习者分析、游戏元素设计、任务设计和评价设计五个环节。在高阶思维的培养上，游戏元素设计、任务设计、评价设计与其息息相关。

（1）学习目标分析

学习目标是对学习者在教学后应该表现出来的可见行为的具体的、明确的表述，是学习活动的出发点、依据和归宿。该协作推理游戏的游戏目标就是促进培养学习者的高阶思维。

（2）学习者分析

学习者是学习活动的主体，学习者具有的认知的、情感的、社会的等特征都将对学习的信息加工过程产生影响。教育游戏是为学习者服务的，所以对学习者进行分析是很有必要的。该协作推理游戏的学习者是大学生群体。

（3）游戏元素设计

游戏元素是教育游戏的核心。通过对情境、情感、学习策略和资源进行设计完成对学习者进行学习活动的支持。元素的设计要为高阶思维的培养所服务。

①情境设计。游戏的情境设计包括游戏主题和游戏角色。游戏主题的设计要符合学习者的认知水平。游戏角色的设计要可以引起学习者的学习兴趣。

②情感设计。游戏场景设计要符合学习者的心理特征，设计与真实接近的情感事实，这有助于学习者沉浸式进行游戏活动。

③策略设计。学习策略的设计是为了提高学习者的学习效果与效率。在协作推理游戏的策略设计中，主要有协作学习、自主学习，学习者通过不断探索以推进游戏进程。

④资源设计。除了学习者参加游戏的自身知识，相关资源和背景的指导也是有必要的，有助于玩家顺利完成游戏活动。

（4）任务设计

任务设计与高阶思维的培养密不可分，它包括教师任务设计和学生任务设计。游戏任务是指学生任务，将学习内容以问题的形式融入到游戏情境中，使学习者在游戏化学习、探索中

与原有经验实现一定的关联，从而达到认知的重新建构。在推理游戏设计中教师任务是学习内容的转化，即游戏库的扩充。

（5）评价设计

评价是游戏化学习情境中互动要素的体现。在游戏中主要采用的是自我评价、教师评价、其他学习者评价和游戏的反馈评价。在自我评价的过程中，学习者通过不断反思和认知重建实现认知结构的整合。通过其他学习者评价和教师评价，学习者了解知识掌握程度，通过他人的反馈不断理清知识点，并实现知识的迁移和创新。游戏反馈提供给学习者不断复盘，推陈出新的机会。

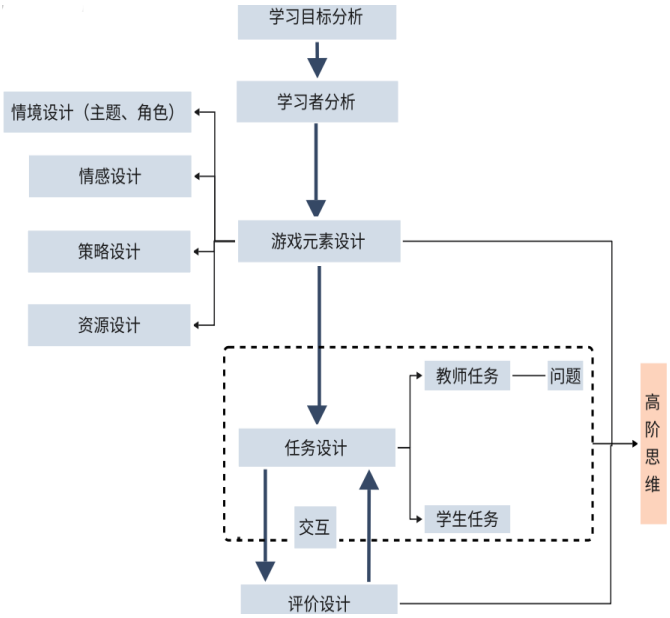


图 1 协作推理游戏设计框架

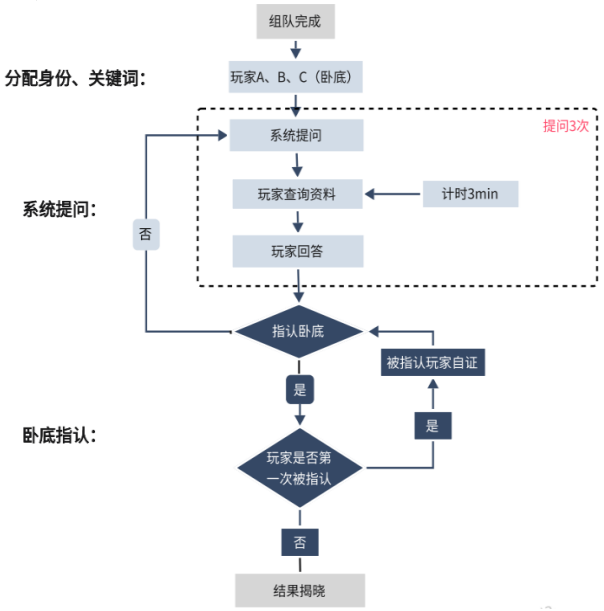


图 2 游戏流程

3.2. 游戏实现

该环节以《教学系统设计理论与实践》课程为例，完成协作推理游戏的实现。

3.2.1. 学习目标

在课程内容上，将课程知识与协作推理游戏的关卡结合，使学生在团队协作的过程中理解和掌握课程内容，提高学习效果。

在团队协作上，玩家在游戏过程与队友紧密协作，在清晰、准确地表达自己观点的同时倾听和理解他人观点，提高其协作能力和沟通能力。

在高阶思维培养上，学生在整个游戏过程中，能用创造性思维回答提出的开放性问题，能用批判性思维排除模糊信息的干扰，能用逻辑推理能力分析陈述进行推理，能用问题解决能力打消他人的怀疑，能运用决策能力筛选出有用信息，做出正确决策。

3.2.2. 学习者分析

该游戏对象是大学生，他们已经具备一定的思维能力。为了便于游戏的开展，游戏对象需具备该课程的基础知识并有初步的问题解决的能力，以便能在游戏中应用。

3.2.3. 游戏元素分析

（1）情境设计

该游戏在卧底追踪的情境下进行，主题围绕课程知识。该游戏角色有“平民”和“卧底”，由系统随机分配。“卧底”任务除了隐藏自己还需要猜测平民的关键词。“平民”任务则是根据游戏过程中对问题答案的陈述来寻找线索，指认卧底。

（2）情感设计

该游戏的情感主基调是期待和悬念。通过对回答语句的不断斟酌和分析，对卧底指认，直到游戏结束才会揭晓悬念。在游戏过程中平民怀着期待的情绪找出真正的卧底，卧底则在焦虑中一边回答问题，一边隐藏自己。

（3）策略设计

在回答问题环节，玩家有 3 分钟的资料查询时间，主要进行自主学习。在指认及自证环节，学习者之间需要进行协作交流学习，进行决策。

（4）资源设计

该游戏的资源主要是游戏库以及玩法指南。游戏库是该游戏的核心，通过游戏库中的关键词和问题，玩家能更好地开展游戏活动。

3.2.4. 任务设计

（1）游戏任务

游戏开始，由系统分配玩家关键词和身份，完整的游戏流程如图 2 所示。C 被分配“卧底”身份和对应关键词，AB 被分配“平民”身份和对应关键词。玩家的第一个任务是依次在对话框回答系统提出的问题，提问持续三次。第二个任务是玩家选择“指认卧底”或者“再答题”。被指认为卧底的玩家有第三个任务即进行自证。

（2）教师任务

教师在该游戏中的任务是游戏库的设计，游戏中的关键词应该是两个易混淆的知识点，针对关键词的提问应该是开放的、具有创造性的，这些问题应该依据布鲁姆的目标分类中的高阶思维层次来进行设计。

3.2.5. 评价设计

在游戏过程中的他人评价是以指认环节来进行的，玩家针对问题的回答不断推理、分析，对他人的指认不断反驳并举证，经过几回合的碰撞摩擦，直至最后游戏的结束。游戏的反馈评价是以游戏结果呈现的。在教师端，更加详细的游戏记录将会显示，包括游戏的关键词、学生对于每个问题的回答、指认的结果等信息，以便教师和学生对知识掌握程度进行复盘和分析，有助于教师改善教学，学生提高其高阶思维。

4. 总结与展望

本研究对协作推理游戏的设计进行探究，并以一门课程为例实现对协作推理游戏的设计，该游戏旨在有效促进大学生高阶思维的培养。该游戏的对象以大学生为主，希望通过这种娱乐化的学习体验，为大学生高阶思维水平的培养提供新颖而有趣的途径，但是关于游戏效果，我们还没有进行一个量化的研究，当然这将是我们未来工作的重点，接下来的研究我们将会在教学实践中对其进行验证和完善。

参考文献

钟志贤 (2004)。促进学习者高阶思维发展的教学设计假设。**电化教育研究(12)**，21-28。

Huang, Y. M., Silitonga, L. M., & Wu, T. T. (2022). *Applying a business simulation game in a flipped classroom to enhance engagement, learning achievement, and higher-order thinking skills.* Computers & Education, 183, 104494.

Siddique, Z., Ling, C., Roberson, P., Xu, Y., & Geng, X. (2013). *Facilitating higher-order learning through computer games.* Journal of Mechanical Design, 135(12), 121004.

Yang, Y. T. C. (2015). *Virtual CEOs: A blended approach to digital gaming for enhancing higher order thinking and academic achievement among vocational high school students.* Computers & Education, 81, 281-295.

基于 Unity 2D 的小学数学教育游戏设计与实现

——以《小数英雄》为例

Design and Implementation of Primary School Math Education Game Based on Unity 2D

--Take "Decimal Hero" as an example

张贤[1*]，赵源源[2]
北京邮电大学 人文学院
*zhangxian@bupt.edu.cn

【摘要】 在教学方式创新丰富的当下，教育游戏展现出了不可忽视的研究价值。在小学教育阶段，趣味性和实践性的结合备受推崇，这就要求把课堂教学和游戏有机结合起来，让学习变得更加生动。本研究以《小数英雄》这款基于 Unity 2D 平台开发的小学数学教育游戏为例，着重研究如何将小数相关知识以游戏的形式呈现出来，详细分析了游戏设计中涉及到的教学理论，尝试揭示这些理论在实际设计过程中运用的具体方案和策略，该游戏使学生在轻松愉快的游戏中，有效学习并掌握数学内容。最后，本研究藉由 AI 技术，并结合智能元素提出改进方式，并对其优势进行剖析。

【关键词】 小学数学；教育游戏；Unity 2D；小数学习

Abstract: *At a time when teaching methods are rich in innovation, educational games show a research value that cannot be ignored. In the primary education stage, the combination of fun and practice is highly respected, which requires the organic combination of classroom teaching and games to make learning more vivid. This study takes "Decimal Hero", an elementary school mathematics educational game developed on Unity 2D platform, as an example, focuses on how to present the knowledge related to decimals in the form of a game, analyzes in detail the teaching theories involved in the design of the game, and tries to reveal the specific solutions and strategies used in the actual design process of these theories, which enables students to effectively learn and master the The game enables students to learn and master math content effectively in a relaxed and enjoyable game. Finally, this study proposes ways to improve the game by AI technology and combines intelligent elements, and analyzes its advantages.*

Keywords: primary school mathematics, educational games, Unity 2D, fractional learning

1. 前言

随着信息技术的飞速发展和义务教育数学课程标准改革的实施推进，小学数学课堂的教学环境、教学内容和教学手段都经历了显著的转变。理想的小学数学课堂应该是活泼、丰富且有趣的学习环境，在此环境下，学生不仅乐于学习，而且学习成效显著。新课程标准也提出“要改变以接受学习、死记硬背和机械训练为重的教学方式，强调培养学生的积极学习态度，鼓励学生积极参与、乐于探究和享受学习过程”。传统的教学形式已经不能满足学习者个性化发展的需求，需要通过多种模态的教学方法来促进学习者的个人认知成长。

近年来，游戏化学习引起了广泛关注，注重“寓教于乐”的教育游戏不仅可以激发学生的学习兴趣，还能促进他们的认知发展(尚俊杰 & 张露,2017)。教育游戏是把学习内容与游戏相结合，游戏者可以在游戏中获取知识、技能、能力和情感等的提升，从而完成设定的教学

目标。有研究表明，教育游戏能帮助学生促进大脑发展(Fissler et al., 2015)、改善学习成绩、提升学习者学习动机等(尚俊杰 & 裴蕾丝, 2015)。教育游戏正向影响学生的学习效果，给学生提供轻松、愉悦的学习情境(和文斌 & 董永权, 2021)。学生在日常学习生活中，承受了作业、考试等一系列负担，而游戏可以作为宣泄工具，摆脱巨大的学业负担，消除枯燥学习的消极情绪，抛开现实因素，将人从被动、不良经验的角色中转移出来，发泄情绪(陶漪 & 恽如伟, 2008)。

本研究将教育游戏的理念引入到小学数学教学设计中，通过分析小学数学课程目标、课程内容及学科特点，研读数学教学法的设计要求，结合教学对象，探索符合小学生认知特点的游戏化学习教案，最终选定三年级《小数的认识》与《小数的大小比较》两个知识节点设计了“小数英雄”游戏案例。

2. 基于 unity 的小数游戏设计与开发

2.1. 游戏设计理念

本游戏旨在通过游戏化学习，让用户能够在有趣的互动中掌握小数相关知识，打破传统教学方式的枯燥与单一。游戏的设计首先依照教学目标设定关卡，通过植入故事情节，赋予学生更强的参与感和使命感，使他们在完成游戏任务的过程中自然而然地学习和领会数学知识。

2.2. 游戏概述

《小数英雄》是一款基于 Unity2D 平台开发的数学教育游戏，针对小学三年级的学生，该内容对应小学三年级下册第七单元的认识小数以及四年级下册第四单元的大小比较的教学知识点。游戏以情景式的设计策略，把小数相关的知识嵌入到一段架构紧张刺激的冒险故事中，使得学生能在富有情感投入和互动体验的游戏活动中理解和掌握小数概念。

2.3. 游戏背景

《小数英雄》设定在一个面临巨大危机的虚构城市，学生在游戏中扮演一名英雄角色，他们需要运用小数知识，解决各种挑战，如解救人质和消灭怪物，以实现拯救城市的目标。这种创新的情境设定把学生们的想象力和游戏中的教学内容有机地结合在一起，以冒险故事作为情境载体，激发学生的学习兴趣。

2.4. 游戏交互

2.4.1. 卡牌游戏——认识小数（如图 1）

在《小数英雄》游戏的设计中，为了将学习与游戏有机地结合起来，使学生在寓教于乐中感受到学习的乐趣，游戏的“小数认识”部分采取了卡牌式的游戏形式。根据教育部颁布的《小学数学课程标准》，并结合《人教版三年级下册第七章——小数的初步认识》的教学要点，运用了分数、图像以及实际生活应用这三种形式的题目，让学生在多维度与多情境下进行小数认知的锤炼。此外，为了增加游戏的紧张感和挑战性以提高学生的学习兴趣和动力，本研究设定了有限的翻牌次数作为答题机会。每正确回答一个问题，象征着成功解救一个人质，即在游戏人质区域的一张卡牌从暗淡变为明亮，表明他/她已经被解救。值得注意的是，正确的回答将推动城市居民的安全状况朝着积极的方向发展，而错误的回答则会使城市的危机加剧。这一设计旨在鼓励学生在游戏中养成正确的数学思维方式并体验学习的幸福感，采用情景化教学的设计理念可以在寓教于乐的同时，提高学生对小数知识理解的深度和广度。



图 1 认识小数关卡

2.4.2. 射击游戏——小数比大小（如图 2）

在“小数比大小”部分，本研究引用《四年级下册第四章 小数的意义和性质》作为教学参考，采用射击游戏的方式实现小数大小比较的学习目标。具体操作方式包括：设置两个携带小数的怪兽图像，学生需根据怪兽下方的小数进行比较，并选择标志为“大”或“小”的火箭进行打击。答题正确意味着怪兽被消灭，错误则火焰会下移一格，火焰移至装备栏时，游戏失败。通过结合射击游戏的模式与小数的学习，该设计旨在提供一种兼具挑战性与趣味性的学习体验。该设计让学生在面对怪兽以及怪兽下方不断变化的小数时，需迅速判断并作出相应的动作。与此同时，子弹的正确发射表示了学生对小数比较这一知识点的理解与掌握，而错误的发射代表学生需要在此知识点上进行深入的学习与反思。因此，通过该方式，学生得以以更高的积极性参与到这一充满娱乐性以及学术性的学习过程中。这一设计将寓教于乐的理念与学术研究相结合，通过以游戏为载体提供学习体验，让学生在乐趣中获得知识，让学习成为真正的享受。



图 2 小数比大小关卡

2.5. 游戏元素

2.5.1. 数轴

在认识小数部分，借助数轴这个认知辅助工具，帮助初学者了解小数在数轴中的位置，建构小数概念。数轴作为学习者已经习得的工具，能帮助学习者理解小数的数学意义。设置数轴的核心目的是让学生掌握数量大小的方向感，右边的数字大于左边，越往右延伸数字越大（张露，胡若楠，曾嘉灵，& 尚俊杰，2021）。

2.5.2. 卡牌元素

卡牌的两面设计考虑到了学习者的游戏体验与动态反馈，背面是怪兽形象，正面是人质以及绿色怪兽。这一独特设计旨在立即呈现出学习者答题的结果（答对为人质，答错为绿色怪兽），从而增强学习者的游戏兴趣，提高其参与度和情绪投入。

2.5.3. 人质设定

为了扩增学习者的答题激励和成就感，设计了解救人质环节，当学习者在认识小数环节的答题正确后，可以获得解救人质的奖励，表现为人质牌的由暗变亮。该设计能让学习者从成功的经验中获取成就感，此种正向反馈能提升学习者的学习积极性。

2.5.4. 怪兽角色

在角色设定上，怪兽形象借鉴了人气动物形象，旨在迎合三年级学生的兴趣取向，增加其对游戏的认同感和参与度。在怪兽的设计上，主要呈现可爱风格，以应对学生可能对凶神恶煞形象所产生的负面情绪。此外，人质角色设定覆盖了男女老少各年龄段，增加了游戏元素的多样性。

2.5.5. 火焰障碍

为了创造紧张刺激的游戏氛围，怪兽下方的火焰元素提供了视觉上的压迫感，有效刺激了学习者加速思考和做出反应，从而提升其问题解决的效率。

2.5.6. 带有符号的火箭道具

在火箭上显著地标明“大于”“小于”符号这一设计创新，有助于学生在玩耍中自然学习和掌握这些重要的数学符号，进一步提升他们对游戏的投入度。

2.5.7. 游戏音乐环境

《小数英雄》通过设计精细的音效，如各场景特有的背景音乐，以及不同答题结果对应的正确、错误音效，关卡结果的胜利、失败音效等，为学习者的游戏体验增添了层次感。这些积极的听觉刺激不仅促进学习者的感官投入，也能引导其积极情绪，从而提升其学习的参与度和心流体验，进而提高数字游戏化学习的效果(詹杰妮, 王智铿, 桂阳, & 蔡志慧, 2023)。

2.5.8. UI 设计

在总体的 UI 制作中，本研究采取了积极视觉情绪设计，各元素均采用了明亮的颜色。明亮的颜色可以明显提高学习者的积极情绪、增加相关认知负荷，并且巩固学习者在游戏中的学习成绩(王智铿, 2023)。

3. 该教育游戏的学习理论以及教学策略部分

3.1 建构主义理论

教育游戏的建构主义主要强调交互丰富的学习环境（该游戏中的人质、怪兽以及火焰），学习者在与各类对象进行交互的过程中开展学习(曹鹭, 2022)。《小数英雄》为初学者专门定制，教育与游戏的结合符合建构主义理论。学习者通过与游戏情境的互动，在积极参与问题解决和决策制定的过程中，学习者通过与游戏内各类元素的频繁交互以及反复的实践，逐渐构建和完善自己的知识体系，实现对小数知识的建构。

3.2 情感认知理论

情感认知理论强调学习是并行的情感和认知两个方向的互动过程。根据这一理论，《小数英雄》创造了一个让学习者在乐趣中学习的环境，能激发学习者的积极情绪反应，如兴奋、好奇、满足感等，从而吸引他们积极参与到学习过程中。同时，学习者在游戏中扮演英雄的角色，接受并完成拯救城市的挑战任务，这一过程也会产生强烈的成就感和自我价值的实现。而且，该游戏通过各种趣味性设计（翻牌、射击）、挑战性任务（解救人质），激发学生的兴趣，使学习过程变得更加积极主动。情感投入有助于加深对数学的情感认知，促使学生更

深层次地记忆数学知识。显然，这种情感上的投入和满足感反过来又能强化学习者对数学知识的学习印象，使学习者产生更深层次的数学知识理解和记忆。

3.3 认知负荷理论

认知负荷理论主张学习者的认知处理能力是有限的，过多的信息将会对学习者的认知过程产生干扰，从而影响学习效果。在《小数英雄》的设计中，为了使学习者在愉快的环境中更好地吸收知识，开发者采用了逐层递进的高级设计策略，通过设计不同种类和难度的游戏关卡，逐步引导学习者面对并解决游戏中的数学问题。针对小数入门阶段的学习内容设计，存在由简单到复杂的递进关系，辅以不同的游戏形式来适应学习者的学习节奏进程。通过这种方式，该教育游戏可以在控制学习者的认知负荷的同时，不断激发和满足他们的学习兴趣，从而达成顺利推进知识学习的目标。

3.4 先行组织者教学策略

大卫·奥苏贝尔的先行组织者理论强调通过使用某些引导性的教材使所学新知识与既有认知构造之间建立链接，以促成意义性的学习。在《小数英雄》这款游戏中，游戏将数轴和分数等学习者已有的知识作为桥梁，将新知识（例如小数）与旧知识有机地连结在一起。学习者借助已有的知识范畴，更易接受和理解新学的小数相关知识，将新旧知识进行整合，从而推进新知识的吸收和掌握。如，游戏中小数的出题区会从基础的分数开始，通过连接数轴，帮助学习者以分数为桥梁，引领他们探索和理解小数与分数之间的关系，进一步推动他们了解并掌握小数知识。

4. 《小数英雄》的教学应用探讨

4.1 提升学生学习兴趣

《小数英雄》将数学概念和游戏元素充分结合，利用生动有趣的画面和音效，以及吸引人的故事情节，拉近了学生对学习小数知识的距离，使他们在无压力的环境中，积极主动地去探索学习。游戏中的数轴和符号火箭等辅助道具，以形象具体的方式引导学生从感性认识了解到小数的抽象概念，通过具体形象的感知延伸到抽象的概念，循序渐进，辅助学生获得良好的游戏与教学体验(李新兰 & 韩瑛, 2009)。此外，游戏的关卡设计强调趣味性，涵盖了小数相关的多种概念，让学习过程充满乐趣，不再枯燥。通过参与游戏互动，学生能够更深入地理解小数概念，养成自主学习的好习惯。当学生进入充满乐趣的游戏活动时，他们可能会经历一种被称为“心流”状态的体验，即完全投入并乐在其中，这种状态的体验能极大地提升他们的学习效果和体验。

4.2 强化学生实践应用

在数字学习领域，应用和实践是推动认知发展的重要方式。对此，《小数英雄》游戏进行了深思熟虑的设计，以使得学习者能够在理解和掌握抽象概念的同时，看到其在生活中的具体应用。“认识小数”部分从三年级下册教材中选取适当的题目，将理论和实践相结合，例如“3 米 6 分米=多少米？”这样的题目，既涵盖了小数的计算，又展示了这种计算在现实生活中的具体应用。这类题目能激发学生对学习的兴趣，使他们在学习过程中能够理解抽象的小数概念，并且能在游戏中体验小数的实际应用，从而更直观地理解和掌握小数知识。而这种教学方法，也符合现代教育理论中强调的“应用性”和“实践性”原则，即在掌握知识的基础上，还需学会应用知识解决实际问题，提升学生的实践能力。学生们不仅是知识的学习者，也是知识的实际运用者。同时，这种方法也善于激发学生的创新思维，并提高他们将所学应用于实践的能力。

4.3 揭示适应性学习有效性

《小数英雄》利用关卡设计及其难度设置，实现了真正的适应性教学。游戏包含“认识小数”和“小数比大小”两大部分，并且在每一部分中都设有层次分明的关卡，难度由易到难。这种设计允许和鼓励学生根据自身的学习需求和能力情况选择适合自己的游戏环节及难度，学生可以先从最基础的关卡开始，然后通过学习和反馈，逐渐提升自己解决问题的能力，挑战更高级的关卡。这样的设计给予学生进一步提升小数知识认知水平的可能性，让游戏变成了适应个体差异的有力工具。此外，该游戏的适应性教学模式还考虑到了认知负荷理论。当学生的认知能力和关卡难度配合得当，才有助于提升学习效果，避免出现头脑混乱甚至厌烦学习的情况。只有将认知负荷控制在一个合适的范围内，才能提高学习的积极性和效率。

4.4 帮助家长与教师进行针对辅导

《小数英雄》游戏采用自动记录玩家游戏数据的方式，实时收集和存储学生在游戏过程中产生的学习数据，以照应其学习情况，从而在一定程度上降低了家长和教师追踪学生学习情况的时间和精力成本。通过对游戏数据的分析，他们可以清楚地了解学生的学习进度和存在的问题，从而提供更有针对性的支持和指导。例如，如果数据显示学生在“小数比大小”部分的表现较弱，那么教师和家长就可以着重对这一部分进行再次教学和训练，这种定向性的教学辅导大大提高了教学效率和效果。通过对学生学习过程的系统化记录和反馈，教师和家长可以更科学地进行教学和辅导，从整体上提高学生的学习质量和效果。

5. 结论与展望

本研究以小学数学三年级“小数的认识”与“小数的大小比较”两个知识点为例，进行了基于 Unity2D 小学数学教育游戏的游戏设计和开发，通过应用科学的学习理念来最大化教育游戏的教学效果。总之，《小数英雄》在界面设计，游戏元素设计，音效设计等方面以儿童的心理需求为根本，力求为儿童的数学学习提供轻松、愉快的学习环境。

随着信息技术的发展，在教学方法与技术的对接上，应用越来越成熟的人工智能技术有望实现更加个性化、精细化的教学辅导。例如，以智能算法帮助个性化教学诊断和策略制定，根据学生的学习习惯和理解速度，生成适应他们的个性化游戏关卡，从而进一步提升学生的学习效率和兴趣。其次，借助人工智能技术，可以对学生的学习历史进行深度分析，并据此为学生定制个性化的学习路径，以便精确定位学生在小数学习中的薄弱环节并进行有效提升。在今后，本研究将对教育游戏与其教学应用进行更深入的研究，通过开发引擎设计开发出更多种类的教育游戏案例，使得教育游戏能够在数学课堂教学中发挥更大的作用。

参考文献

曹鹭. (2022). 三种基本学习理论与教育游戏的设计. **开放教育研究**, *28*(5), 29-38+92. https://doi.org/10.13966/j.cnki.kfjyyj.2022.05.004

和文斌 & 董永权. (2021). 教育游戏对学生学习效果的影响研究——基于 41 项实验和准实验的元分析. **现代教育技术**, *31*(4), 44－50.

李新兰 & 韩瑛. (2009). 从心理学角度谈小学数学教育游戏的设计. **中国教育技术装备**, (33), 22－23.

尚俊杰 & 张露.(2017).基于认知神经科学的游戏化学习研究综述.**电化教育研究** (02),104-111.doi:10.13811/j.cnki.eer.2017.02.014.

尚俊杰 & 裴蕾丝.(2015).重塑学习方式:游戏的核心教育价值及应用前景.**中国电化教育** (05),41-49.

陶漪 & 恽如伟. (2008). 教育游戏设计理论的分析与运用. **远程教育杂志**, (2), 75 - 78.

王智铿. (2023). 数字游戏化学习中视听情绪设计对学习效果的影响 (**硕士, 华中师范大学**). 华中师范大学. https://doi.org/10.27159/d.cnki.ghzsu.2023.000742

詹杰妮, 王智铿, 桂阳, & 蔡志慧. (2023). 数字游戏化学习中听觉情绪设计对学习效果的影响. (eds.)**第二十五届全国心理学学术会议摘要集——分组展贴报告**(pp.23-24). 华中师范大学心理学院. https://doi.org/10.26914/c.cnkihy.2023.057354

张露, 胡若楠, 曾嘉灵, & 尚俊杰. (2021). 如何设计科学、有效、有趣的教育游戏——学习科学跨学科视角下的数学游戏设计研究. **电化教育研究**, *42*(10), 70 - 76. https://doi.org/10.13811/j.cnki.eer.2021.10.010

Fissler, P., Kolassa, I. T., & Schrader, C. (2015). Educational games for brain health: revealing their unexplored potential through a neurocognitive approach. Frontiers in psychology, 6, 140709.

AR 技术赋能幼儿多角度高阶思维实证研究

------基于北京军区机关幼儿园 AR 技术教学应用实践

An Empirical Study on Multi-angle Higher-order Thinking of Children Enabled by AR Technology——Based on The Application of AR Technology in Kindergarten of Beijing Military Region

段春梅，钱媛媛
北京军区机关幼儿园
59817266@qq.com

【摘要】基于 AR 技术赋能的互动手段为教育者提供了一种新的表达方式给学习对象，同时也以最贴近自然的互动方式为学习者构建了一个自我探索的空间。北京军区机关幼儿园连续 5 届参加 COR 汉语在线快速阅读饶宗颐著作读后随笔国际邀请活动，积极探索 AR 技术赋能幼儿教育应用，关注幼儿多角度高阶思维能力培养质效。本文通过实证研究方法，分享幼儿教育中 AR 技术实践应用过程。结果表明，AR 技术的应用能有效助力提升幼儿认知能力、创造力和问题解决能力，为幼儿教育提供了一种新的有效工具。

【关键词】 AR 技术赋能；幼儿多角度高阶思维；实证研究

***Abstract*:** *The interactive method based on AR technology provides a new way of expression for the learners, at the same time, it also constructs a space of self-exploration for the learners. For the fifth consecutive year, the Beijing Military Region Government Kindergarten participated in an international invitation activity for Chinese online speed reading of Jao Tsung-I's works, actively exploring the preschool application of AR technology, to pay attention to the quality and effect of cultivating children's multi-angle higher-order thinking ability. In this paper, we share the application of AR technology in preschool education. The results showed that the application of AR technology can effectively improve children's cognitive ability, creativity and problem-solving ability, providing a new and effective tool for preschool education.*

***Keywords*:** *AR technology empowerment, multi-angle higher-order thinking of children, empirical study*

1. 引言

当今世界日新月异，教育现代化已成为大势所趋、发展所需。世界各国纷纷将教育现代化发展、人的现代化发展，作为推动教育教学改革和社会进步，以及国家综合实力提升的重要举措。教育部科学技术与信息化司司长周大旺在教育部新闻发布会上讲到，“发展数字教育、推进教育数字化，推进教育现代化是大势所趋、发展所需，也是改革所向。”学前教育是教育活动的第一个阶段，作为基础教育的重要组成部分，对人一生的发展起着不可替代的作用。

《中国儿童发展纲要（2021-2030）》中明确指出：“加强儿童媒介素养教育。保障儿童利用和参与媒介的权利。丰富儿童数字生活体验，提高数字生活质量。”（国务院.《中国儿童发展纲要(2021～2030 年)》[EB/OL]．https://www.gov.cn/gongbao/content/2021/content_5643262.htm）

有关报告显示，增强现实 AR 技术对于幼儿教育教学具有特殊性、必要性和普适性，已成为未来社会发展中最具有影响力的技术之一。第一，特殊性。虚实结合与过程性体验是 AR 技术的主要特征，对幼儿教育教学来说，最特殊的就是能呈现不同的情境，让幼儿可以亲身体验和直接感知，帮助其提升经验。第二，必要性。AR 技术能够通过技术手段以三维场景，将幼儿园里无法直接呈现给幼儿的教学相关知识真实地呈现给幼儿，既能启发幼儿的心智又能开拓幼儿的视野。第三，普适性。AR 技术应用于幼儿园教学的核心价值在于提升教学活动的质量，促进幼儿多角度高阶思维的发展，在此意义上是普适的。因此，探索将 AR 技术应用于幼儿园教育教学,不仅顺应教育信息化的发展潮流,而且更有利于提高幼儿园教育整体质量。

2.研究依据

2.1. AR 技术赋能教育理论

2019 年中国教育三十人论坛第六届年会上，顾明远教授说:“关于未来教育，我们要充分利用人工智能、大数据、数字化等信息技术改变人才培养的模式，关键是要充分认识学生的内在潜力”。斯坦福大学人工智能、机器人与未来教育项目主任、斯坦福大学全球创新设计联盟联席会主席蒋里教授认为：人工智能思维是以后每一个人都需要具备的思维方式。

AR（Augmented Reality），是在现实世界中迭加虚拟信息，也即给现实做“增强”，这种增强可以是来自视觉、听觉乃至触觉，主要目的均是在感官上让现实的世界和虚拟的世界融合在一起。AR 流程的三个关键点。首先是 3D 环境理解。主要依靠物体/场景的识别和定位技术。第二个关键技术：Research on AR Technology and Appreciation of Literary Works 显示技术。这其中又分为视频透视和光学透视，其他的代表有光场技术（主要因 Magic Leap 而显名）、全息投影。AR 中的第三个关键技术在于人机交互，用以让人机迭加后的虚拟信息互动，AR 追求在触摸按键之外自然的人机交互方式，比如语音、手势、姿态、人脸等，用的比较多的语音跟手势。（选自网络数据）虚拟现实和现实增强技术（VR/AR）技术是现代人工智能高度发展的结果。AR 技术作为新兴的现代教育技术，具有趣味性、沉浸性、交互性、可视性等特点，同时，提供给读者良好的阅读体验。（林小苹等.AR 技术与文学作品鉴赏应用研究[EB/OL].北京：中国科技论文在线 [2020-02-20].http://www.paper.edu.cn/releasepaper/content/202002-95.）

2.2. AR 技术赋能心理学理论

梅斯基（Mey-sky）和纽曼（Newman）对儿童创造力下了如下定义：创造力是一种能力，它能使一个人以一种让别人听取和欣赏他们讲话的方式来表达自己的意见；创造力是一种素质，它能使人发现别人以前未能理解的定义。

好奇心是人对新异事物产生诧异并进行探究的一种心理倾向。求知欲又称认知兴趣，它是好奇心的升华，是人渴望获得知识的一种心理状态。好奇心和求知欲是推动人们主动积极地观察世界，进行创造性思维的内容动因，具有好奇心和求知欲的人。对事物有着执着的追求和迷恋,不会感到学习和创造是一种负担,而是在活动中获得极大的精神鼓舞和情感满足。

高阶思维是发生在较高认知水平层次上的认知能力，它主要由信息整合能力、新知识建构能力、逻辑思维能力、批判思维能力、创造性思维、评价反思能力等内容构成。当前，互联网推动人类文明迈上新台阶，改变了人们获取信息和知识的途径、交流方式。AR 助力文学作品鉴赏以“品、赏、析、联、创”为引导的教学原型，讨论创新科技与教育科学理论给中文教育教学实践带来的新挑战。（林小苹等.AR 技术与文学作品鉴赏应用研究[EB/OL].北京：中国科技论文在线 [2020-02-20].http://www.paper.edu.cn/releasepaper/content/202002-95.）

分析AR技术对文学作品鉴赏的突出作用，突出“以幼儿为主体”与“培养幼儿多角度高阶思维”的互动引导，养成善于“发现”，着眼于人才好奇心和创造力培养。

3. 研究内容

3.1. AR技术赋能文学作品鉴赏样品选择

AR技术以增强现实（AR）等人工智能技术作为媒介手段，包含多媒体、三维建模、实时视频显示及控制、多传感器融合、实时跟踪及注册、场景融合等新技术与新手段，研究设计并评估基于人工智能与视觉元素交互作用，对增强诗词的现实体感效果。

这个项目AR的样品来源：由林小苹总干事主持向香港特别行政区创新科技署科技券计划申报获《建立设有扩增实境互动系统的手机应用程序》TVP/0507/21完成项目，AR样品实验群体获得应用体验，如下图：



图1 AR样品界面

其中，以中华古诗词白居易名篇“钱塘湖春行”为体验实例，本实证研究采用的AR样品是“钱塘湖春行”普通话版本。

3.2. AR技术赋能幼儿的学习动力

幼儿能够与AR虚拟物体产生交互，通过游戏化学习，达到沉浸式的体验，或者是使他们主动地去探索知识，增强幼儿的学习动力。当幼儿使用AR进行学习时，往往能有更高的满意度，体会到更多的乐趣以及会很愿意进行重复体验。

（1）信息的选择性建构和互动性呈现，引导注意力投入。AR呈现的虚拟教育信息，主要以形式和内容的良好设计来吸引个体的注意力，进而引导和保持幼儿的注意力。内容上，AR一般以有趣的互动设计激发幼儿的学习兴趣，使其持续投入注意力。在实际学习过程中，AR学习内容是在幼儿与学习对象的互动中生成的，对此幼儿不仅可看，还能触摸和操作，在此互动过程中幼儿的注意力易被吸引并维持一段时间。

（2）3D虚拟空间的创建，增进行为参与和个体体验。幼儿注意力得到引导的同时，AR逼真的虚拟信息能使幼儿产生较强的近体空间效应，促进身体姿势、行为动作的参与，并丰富个体体验。

（3）多模态信息的刺激，丰富有关客体的经验。通过多模态信息对个体多种感官的刺激，以及多种信息互相补充、共同诱发和引导，有助于唤醒个体积极的情绪、情感，增强个体的情境感，进而丰富个体有关客体的信息和经验。

4. 研究方法

4.1. AR赋能，聚焦发现，提升思维能力

学习知识的认知过程，区分为低阶思维和高阶思维。相较于低阶思维，高阶思维是发生在较高认知水平层次上的心智活动或较高层次的认知能力，主要由问题求解、决策、批判性思维、创造性思维这些能力构成。

尤其是发现式学习，能比较有效地促进学习者高阶思维能力发展。老师要成为幼儿的助学者，幼儿要养成良好的学习习惯，根据幼童的特点，而不仅仅限于要求学习者回忆事实性

信息的活动。开发想象力，启迪智慧。发挥再造想象和创造性想象。读、写、说、画、游戏等等，个性化创意特点。问题发现与求解、在资料查找中去发现新创意。模式设计如下图：



图 2 AR 赋能幼儿活动模式设计图

（林小苹等.AR 技术与文学作品鉴赏应用研究[EB/OL].北京：中国科技论文在线 [2020-02-20].http://www.paper.edu.cn/releasepaper/content/202002-95.）

4.2. AR 技术助力文学作品鉴赏应用体验

2024 年初，北京军区机关幼儿园段春梅园长主持，以幼儿园 10 个大班和 1 个中班为主，有 13 组共 39 位老师自费参加的 COR8·AR 助力文学作品鉴赏教学设计文本选评，老师们不仅积极体验创新科技给教学实践带来的创新机遇，更是想方设法，将设计通过教学实践，在宝宝的感受和体验中，逐步提升自己的经验与水平。以中十班钱媛媛老师率先尝试的将 AR 情景引入小朋友游戏活动为例。如下图：



图 3 钱媛媛老师将 AR 情景教学引入课堂实践

4.2.1. 活动对象

北京军区机关幼儿园中十班共 8 名 5 岁左右幼儿参加，设计包含问卷调查、课堂观察、个案访谈和幼儿作品分析等。

4.2.2. 活动材料

《钱塘湖春行》AR 样品，来源于林小苹总干事向香港特别行政区创新科技署申报获批科技券计划：TVP/0507/21，熊猫 AR 样品，《春晓》音乐，古筝乐曲，安卓手机，多媒体一体机，电子图谱，非洲鼓（每名幼儿一个）。

4.2.3. 活动过程

（一）前侧环节。谈话导入，用熊猫 AR 技术，引出话题。

（二）中间环节。①教师通过安卓手机，将《钱塘湖春行》AR 样品投放到一体机上，引导幼儿初步品赏古诗中的寓意，感受古诗的意境美。②结合 AR 视频，教师鼓励幼儿用自己的语言大胆地表达对春天景色的感受，体会诗歌的情趣。③根据 AR 画面，教师配乐幼儿抑扬顿挫朗诵古诗《春晓》，表达出对春天景色的喜悦之情。④结合《钱塘湖春行》AR 画面，教师介绍古诗《春晓》作者，扩大认知。由幼儿根据 AR 画面背景多角度联想。⑤出示一体机电子图谱，鼓励幼儿结合古诗《春晓》看图谱用声势、动作边打节奏边唱古诗。⑥出示结合一体机 AR 样品，分组操作后讨论小结。

（三）后侧环节。①出示一体机电子图谱，播放古诗《春晓》音乐，鼓励幼儿结合非洲鼓进行古诗表演。②研究者将有关《钱塘湖春行》AR 样品投放到美工区，鼓励幼儿将古诗的春天美景进行绘画、手工和泥塑，帮助幼儿更深入体会诗词的意境。

（四）环节延伸。①细读赏景，品味语言：提出问题“你最喜欢 AR 画面哪些春天景象？”“你发现哪些景物有何特点？”（突出词组如“绿绿的、白白的、清清的，粉粉的等”）感受诗优美的语言和意境。②创诵诗歌，拓展提升：（1）打开手机平板中的 AR 样品，分小组找出诗中春天的景物，可添加图画。（2）小组讨论，分享小组的理解（以问题为导向，促进幼儿思维能力的发展和提升）。③结合本活动收获引发幼儿环保意识，鼓励幼儿设计一张环保宣传画，并讲述画面内容，教师一对一倾听记录，争当“优秀环境环保宣传员”，完成后发布至 AR 互动区，师幼、幼幼共同点评。

4.3. AR 技术助力文学作品鉴赏实证研究

4.3.1. 个案访谈

他怎么会在里头呢？



图 4 小女孩的自主发现回答

个案访谈：我发现宋庆平很可爱，我看见上面的屏幕，感觉他在吃草，我想进去想看看，和平时游戏不一样，我觉的很开心，我喜欢里面有鸟叫，我喜欢 AR 体验游戏，下次还想看，还想去探索一下。

4.3.2. 问卷统计分析

北京军区机关幼儿园中十班共 8 名 5 岁左右幼儿参加 AR 体验，设计包含问卷调查，对小朋友和家长邀请在线调查问卷；邀请另外 8 名未参加 AR 体验的小朋友和家长参加在线调查问卷。由平台 https://www.chxckc.com/COR8 在线问卷栏目发布和自动数据统计分析。

基本情况

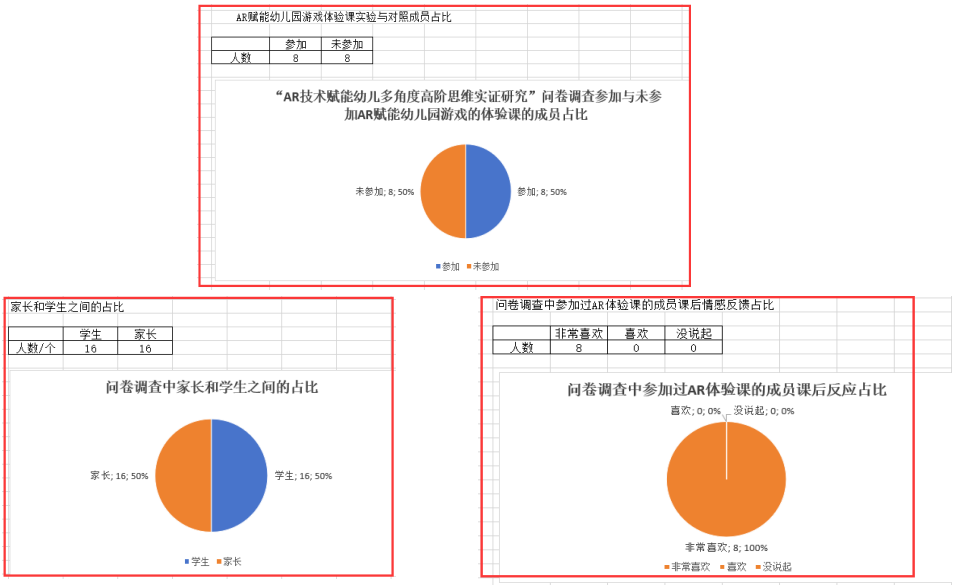


图 5 参与 AR 体验在线调查问卷数据统计分析

（1）学生

序号	请简单表述小朋友参加“钱塘湖春行”AR 体验课后的最大变化：
1	我喜欢上了古诗，我每天晚上都喜欢朗诵古诗。
2	我喜欢这个游戏，以后还要上。
3	喜欢老师，喜欢古诗。
4	喜欢唐诗。
5	我感觉很好玩，还想玩这个 AR 游戏课，我还喜欢探索，还喜欢古诗。
6	爱上了唐诗，还有古诗。
7	喜欢古诗，我每天晚上都要学一首才睡觉。

（2）家长

序号	请简单表述小朋友参加“钱塘湖春行”AR 体验课后的最大变化：
1	孩子特别喜欢，特别开心，还喜欢朗诵古诗了。

（1）学生

序号	请参加的小朋友表述参加的最新发现?
1	我发现有小河、小鱼、大山、小亭子，还有小草，小马而且都会动。
2	有小船、有小鱼、还有云，比平时的游戏活动好玩，平时没有这种感觉，那天我非常高兴所以看看能不能咬到草地。
3	我最喜欢的是小船，大山，还有云朵，还可以变化，还有荷叶，有水和水上的那些花，树上全是小鸟的叫声。
4	我特别特别喜欢，感觉特别好，比别的游戏都好，还想上一次，特别好玩。
5	我发现宋庆平很可爱，我看见上面的屏幕，感觉他在吃草，我想进去想看看，和平时游戏不一样，我觉的很开心，我喜欢里面有鸟叫，我喜欢 AR 体验游戏，下次还想看，还想去探索一下。

（2）家长

序号	请参加的小朋友表述参加的最新发现?
1	平时的游戏不能投到屏幕上，这个可以用手机投影，里面的小船、水、石头、花和平时的不一样，还可以触摸。
2	和我们平时玩的游戏不一样，而且是手机弄出来的。
3	孩子发现了山、还有小鱼、还有小船、还发现了一座大山。

序号	请参加的小朋友表述参加的最新发现?
1	平时的游戏不能投到屏幕上，这个可以用手机投影，里面的小船、水、石头、花和平时的不一样，还可以触摸。
2	和我们平时玩的游戏不一样，而且是手机弄出来的。
3	孩子发现了山、还有小鱼、还有小船、还发现了一座大山。

题目：对于这一先导探索，你有哪些宝贵建议?

（1）学生

序号	对于这一先导探索，你有哪些宝贵建议?
1	特别喜欢这个游戏，下次还想玩。
2	希望下一次还要玩这个游戏。
3	我感觉很好玩，还想玩这个 AR 游戏课。
4	下次还想探索 AR。

（2）家长

序号	对于这一先导探索，你有哪些宝贵建议?
1	很期待。
2	特别喜欢。
3	希望下次还有这个游戏。
4	感觉那个游戏特别好。

图 6 AR 体验在线调查问卷汇总

以上的图表提取部分了问卷内容，小朋友和家长的反馈是具体而真实，效果一目了然。

5．研究分析

运用 AR 技术将真实世界信息和虚拟世界信息“无缝”集成，把原本在现实世界的一定时间空间范围内很难体验到的实体信息（视觉信息、声音、味道、触觉等）通过计算机等科学技术，仿真模拟后再迭加，将虚拟的信息应用到真实世界，被人类感官所感知，从而达到超越现实的感官体验，使真实的环境和虚拟的物体实时地迭加到了同一个画面或空间同时存在。体验中，重点让幼儿在欣赏 AR《钱塘湖春行》的春景中活动。诗，美在意境，欣赏古诗的目的就是提升思维能力为主线的品鉴、欣赏、解析、联想、创造性的阅读活动，将孩子带入诗境之中。因此，借助 AR 信息技术与《钱塘湖春行》AR 样品相融合，以深度融合赋能生动活泼的 AR 互动教学情境、构建以 AR 动画为中介的活动，以合作探究模式、注重持续动态的延伸拓展与交流评价三大实施路径，将抽象概念可视化，将教学时间多元化，促进幼儿习惯迁移内化。而从心理学的角度来说，人的右脑主于综合、整体的听辨，包括想象、感知、直觉等等，而左脑主于语言及与语言相关的功能。因而，意境的感悟关键在于激活幼儿的右脑，使幼儿在音、图、声等多种显性手段结合的形象环境中，加强对诗句语言的咀嚼体味，利用幼儿与诗句相通或相关的生活经验，唤起幼儿的亲切感，缩短诗与幼儿的心理距离，让幼儿带着一种朦胧的抒情的情绪，进入诗的意境，感悟诗的意境。因而，活动主要是靠 AR 信息技术来引领，让幼儿通过研究性学习，培养幼儿的创意思维能力，问题发现与求解，在探索中发现新创意，多感官、多角度、多元素，全方位进行联想，让每一名幼儿都能开动小宇宙，创造出充满智慧火花的杰作。

5.1. 个体与环境的交互

幼儿的年龄特点决定了幼儿是运用触觉、嗅觉、视觉、听觉等多感官进行学习和游戏，这既能满足幼儿的好奇心，还能促进幼儿认知的发展。增强现实技术应用于幼儿园教学活动，恰好能实现幼儿视觉、听觉、触觉、运动觉等多感官通道与环境进行交互，是所有幼儿学习与发展的需要，也有利于支持不同风格偏好的学习。在传统教学活动中，幼儿在学习中主要采用视觉和听觉，不能全身心地投入到活动中。而在基于 AR 技术的教学活动中，幼儿能运用视觉影像进行游戏和学习，实现各感官的结合，达到幼儿最大化的学习和发展。在增强现实技术的支持下，能够充分调动幼儿多感官通道参与学习，提供多样化的交互方式，从而提升幼儿的学习兴趣，促进幼儿的主动学习。

5.2. 虚实结合构建情境

当 AR 技术应用于幼儿园教学活动时，幼儿在真实的环境学习过程中，还能感知虚拟技术创造的信息世界。此外，虚拟世界是幼儿学习情境的重要组成部分，既拓展了真实的物理空间，也实现了信息空间与真实物理空间的有机融合。在这种技术场景下，幼儿获得的是一体化的空间体验。比起传统的借助平面图像，幼儿在三维环境里去认识虚拟化的信息和事物，更容易被理解和认识，这不仅有利于幼儿对抽象知识的理解，还有利于幼儿空间思维能力的训练和发展。此外，幼儿将在沉浸式媒体（以 AR 技术为主）中，实现一种身临其境、对虚拟世界和真实世界的主观感受。

5.3. 我者和他者的协作

在传统的幼儿园活动中，由于多种因素影响，往往无法考虑每一个幼儿的需求和想法，教师与幼儿之间仅仅是一种简单的交流、对话，从而影响了师幼互动的质量，幼儿没有得到足够的刺激和激励，于是游戏和学习活动质量偏低，进而阻碍幼儿的健康可持续发展。而虚拟现实技术的使用则是另一种极端的表现。AR 技术不会完全让使用者封闭在信息世界里，绝缘于真实世界中，教师可以和幼儿进行互动和协作。基于 AR 技术下的教学活动，师幼之间、

幼幼之间相互合作、相互学习，共同发现问题、提出假设、解答问题，更有助于幼儿新经验的获得、自身意义世界的建构和社会性的发展。

6. 结论及展望

通过实证研究证明，AR 技术在幼儿教育中能有效促进多角度高阶思维的发展。AR 技术为幼儿提供了沉浸式的学习体验，激发了其探索欲望和创新思维。未来，我们将在更多场景下利用 AR 技术促进幼儿多角度高阶思维能力培养，并进一步探索其在不同领域的应用价值。要升级功能，对教学资源进行系统化管理和开发，并且积极落实教学工作，在满足技术发展需求的同时，将沉浸式、交互式、构想式教学等模块融合在一起，提高教学水平并为教学应用的系统化升级提供保障。

未来，还有很多需要进一步深入探索，比如：如何进一步优化 AR 技术在幼儿教育中的应用、如何保障孩子们在使用 AR 技术过程中的身心健康等。相信在不久的将来，随着技术的不断进步和研究的深入开展，AR 技术将在幼儿教育中发挥出更大的潜力。

习近平总书记在中共中央政治局第五次集体学习时强调:“教育数字化是我国开辟教育发展新赛道和塑造教育发展新优势的重要突破口。进一步推进数字教育，为个性化学习、终身学习、扩大优质教育资源覆盖面和教育现代化提供有效支撑。”教育数字化是推动教育高质量发展，建设教育强国的重要引擎。面对新一轮科技革命浪潮，教育要主动变革、积极作为，用数字化转型助推教育强国建设，努力办好人民满意的教育。

参考文献

国务院.《中国儿童发展纲要(2021～2030 年)》［EB/OL］.https://www.gov.cn/gongbao/content/2021/content_5643262.htm.

林小苹、谢维平、曾金金、王暑胜.AR 技术与文学作品鉴赏应用研究[EB/OL].北京：中国科技论文在线 [2020-02-20].http://www.paper.edu.cn/releasepaper/content/202002-95.

钟志贤.信息化教学模式:理论建构与实践例说［M］.北京:教育科学出版社,2005.

香港中文电脑应用学会：纵横在线和移动创新模式汉语学习资源平台 https://www.chxckc.com/

Dede C. Immersive interfaces for engagement and learning ［J］.Science,2009,323(5910):66-69.

面部表情分析方法在教育游戏中的应用

Application of Facial Expression Analysis Methods in Educational Games

赵玥颖[1*]，张雨竹[2]，张露[3]，尚俊杰[1]
[1]北京大学教育学院学习科学实验室
[2]北京市朝阳区芳草地国际学校双花园校区
[3]北京邮电大学人文学院
[*] yueying_zhao@126.com

【摘要】 学习情感计算技术的出现与发展，为监测学习者在学习过程中的情绪情感产生与变化、适时提供个性化教学反馈与情感支持提供了新的方向。面部表情分析是学习情感计算的重要组成部分，通过对学习者面部关键组件的肌肉变化数据进行采集与分类，能够判断与揭示其行为背后的内在情绪情感体验。本文旨在探索面部表情识别方法在教育游戏中的应用，重点阐述面部表情识别的含义与发展、面部表情识别在教育游戏中面临的挑战与应用策略，以期为将面部表情识别方法应用于学习中的情感计算领域提供参考与方向。
【关键词】 情感计算；面部表情识别；FER；面部表情分析；教育游戏

Abstract: *The emergence and development of learning affective computing technology provides a new direction for monitoring the learners' emotions in the learning process as well as providing personalized teaching feedback and emotional support. Facial expression analysis is an important part of learning affective computing, which evaluates and reveals the inner emotional experience behind learners' behaviors by collecting and classifying the data of learners' facial muscle changes. The purpose of this paper is to explore the development of facial expression recognition, the challenges and application strategies of facial expression recognition in educational games. As a result, this paper provides reference and direction for the application of facial expression recognition methods in the field of affective computing in learning by educational games.*

Keywords: affective computing, facial expression recognition, FER, game-based learning

1. 前言

认知过程通常伴随着情绪情感的产生，这些情绪情感与所处环境密切相关，能够影响人们的动机与状态、行为与表现、心智与思维（徐晓青 等, 2018），使得探究情感这一因素逐渐受到各个领域的广泛关注。探究与分析人们在认知过程中的情感产生及其变化，有助于理解其内在认知过程，揭示其心理变化，了解其情感体验，为调节认知过程、引导积极情绪产生提供方向。

随着信息技术的发展，情感计算为情感状态挖掘与识别提供了方法与技术支持（Picard, 2003）。通常来说，可以通过语言形式与非语言形式来识别人的情绪情感（Friend & Bryant, 2000），其中非语言形式主要表现为面部状态、身体姿势与肢体动作等。相比于身体动作与姿势，面孔承载着大量的信息线索（Bijlstra et al., 2010），是人类情感表达最重要的形式（De la Torre & Cohn, 2011）。有研究表明，人们在交流时，约有 55%的有用情感信息通过面部表情来传递，而来自身体姿势和语言的分别约占 38%和 7%（Hale et al., 2017）。由此

可知，作为人们在日常生活中一种常见的非语言交流形式，面部表情是人类情绪输出的主要媒介，也是人类最重要、最直观的情感表达方式与推测指标之一（Mehrabian, 1971）。因此，越来越多情感计算相关研究将目光聚焦于面部表情的识别与分类（Adolphs, 2002）。

随着人工智能、人机交互、计算机视觉等领域的发展，面部表情识别（facial expression recognition, FER）研究逐渐深入。FER 指通过判断不同小肌肉在面部的运动来判断个体面部表情、推断其情绪情感状态。这一方法为测量与理解人类情绪情感变化提供了更加准确高效的方式（Faria et al., 2017），被广泛应用于用户研究、人机交互、医疗诊断、抑郁或情绪障碍监测、刑侦测谎、安全检测、疲劳驾驶、生物特征识别、市场营销、广告定制、多媒体学习与在线教育等领域中（王欢欢 等, 2022）。

2. 情感计算与面部表情分析概述

2.1. 面部表情分析概况

通过对面部器官和肌肉的控制，人脸可以借助简单表情来传达复杂的情绪情感线索。其中，面部表情运动变化可视作眉毛、眼睛、鼻子、嘴巴等关键部位的动态变化组合，而脸、眼、口等三个部分的肌肉变化最丰富、轮廓相对更易识别（徐琳琳 等, 2017），因而常作为面部表情识别的重要指标。基于计算机视觉、人机交互等多学科的发展基础。国际上关于面部表情分析与识别的研究可以划分为基于心理学的研究与基于计算机识别的研究两类。

基于心理学的研究出现较早，最早可以追溯到 19 世纪，达尔文（1982）在其著作《人和动物的表情》（《The Expression of the Emotions in Animals and Man》）中，指出人与动物的面部表情之间存在一定的相似性，并由此推断人类的面部表情是进化的产物（程静与刘光远, 2012）。这一著作引发了学界对面部表情的关注，吸引了一系列学者围绕面部表情的跨文化性和分类标准展开讨论与研究。

1971 年，Ekman 与 Friesen 指出不同文化背景下的人类具有 6 种基本情绪，分别为快乐（happy）、悲伤（sad）、愤怒（angry）、惊讶（surprise）、厌恶（disgust）和恐惧（fear）。在此基础上，他们系统地研究了识别对象的类型，于 1978 年提出面部动作编码系统（Facial Action Coding System, FACS）。该系统根据人脸面部肌肉的生理解剖结构特点，将每一种表情分解为一些相关的运动单元（Action Unit, AU），通过眉毛、眼睛、眼睑、嘴巴等面部结构方面的变化来描述面部肌肉运动。2022 年，相关的 AU 数量由 44 个更新为 98 个。通过分析 AU 的运动特征和相关区域，可以进一步描述人脸表情的组成与变化，判断个体的情感类型与强度。作为面部表情识别最普遍使用的系统（Ekman & Friesen, 1978），FACS 目前已被广泛用于提供面部表情肌肉运动的权威参照标准、指导情绪能力提升的干预方案设计（彭小江与乔宇, 2020），为近现代面部表情识别研究的开展奠定了坚实基础。

许多研究最初采用手工方式开展面部表情识别，对时间投入要求较高，尤其动态视频处理耗时较长，于是基于计算机的自动表情识别逐渐兴起（程静与刘光远, 2012）。1978 年，Suwa 等对一段连续图片序列中选定的 20 个脸部固定点进行跟踪，并与典型模式（prototype patterns）进行比较以实现面部表情识别，实现了最早的基于计算机的面部表情自动识别。目前所说“面部表情分析”，通常指通过计算机技术开展的面部表情自动识别与分析。

面部表情分析通常只需一个摄像头或一台带头摄像头的计算机，且可以在无接触的情况下远程操纵，同时能够结合其他方法一同使用，具有设备要求低、操作方式易、应用范围广、融合度高的优势，因此被广泛用于用户情感测量领域。

2.2. 面部表情分析及相关情绪理论模型

目前，用于指导面部表情识别中情感分类的理论大概分为两种：离散情绪模型（也称分类情绪模型）和连续情绪模型（也称维度情绪模型）。

离散情绪模型认为，个体的情感可以划分为有限数量的基本情绪，这些基本情绪在日常生活中出现频率最高，它们彼此离散独立，具有生理特异性，而其它更复杂的情绪都是基本情绪的组合。离散情绪理论最有代表性的为艾克曼（Ekman）提出的 6 种基本情绪。

连续情绪模型认为情绪是一个连续体，不能用非此即彼的离散状态来表征，而应在更广泛宏大的多维坐标或多维空间中来描述。目前受到最广泛认可的模型为罗素（Russell, 1980）提出的效价-唤醒环形模型（Valen-Arousal circumplex model）。这一模型通过效价和唤醒两个维度来表征情绪，效价描述情绪的负向（消极）或正向（积极），唤醒描述情绪的高低程度。此外，普拉奇克（Plutchik）提出了情绪轮模型，这一模型采用倒锥体形状，通过强度、两极性、相似性三个维度来表征情绪。最内层由快乐（joy）、信任（trust）、恐惧（fear）、惊喜（surprise）、悲伤（sadness）、厌恶（disgust）、愤怒（anger）和期待（anticipation）8 种核心情绪组成，向外依次扩展为强度较弱的情绪，相邻处的情绪相似，对角处的情绪相反，间隔角度越小，情绪越相关，由此可以构成多组情绪之间的配对。

虽然连续情感模型能够表征无限种类的情感类型，但关于情绪类型与维度空间之间的转换是较难突破的一个问题（陈子健与朱晓亮, 2019），学界目前尚未对应采用的情绪表征维度这一问题达成共识。相比之下，离散情绪模型采用简洁的方式区分不同情感类型，更便于利用计算机进行检测与识别，因此，目前关于面部表情识别的研究大多基于离散情绪模型。

2.3. 基于 FER 的学习情感计算分析流程

面部表情识别通常包括采集图片/视频、预处理（人脸检测与人脸配准）、特征提取、表情分类等步骤（见图 1）。预处理具体包括人脸检测与人脸配准，其目的是去除噪声，剔除与表情无关的区域。其中，人脸检测具体体现为检测图像中是否包含人脸，然后准确地找到并定位人脸；人脸配准则在人脸检测的基础上去除旋转、遮挡等因素的影响，准确地定位人脸的特征点。目前，人脸检测与匹配算法比较成熟，许多市面上的产品都能实现较精准的人脸检测与配准。



图 1 面部表情识别与分析流程示意图

在人脸检测与配准后，便可以进行特征提取与表情分类。其中，特征提取的具体操作为提取图像上的特征点，判断其是否属于某类特征；表情分类则是根据提取的特征来判断该表情所属类别。特征提取与表情分类是面部表情识别的核心步骤，也是面部表情识别与人脸识别最大差异之所在。特征点的判断与提取的准确度直接影响面部表情识别的性能和准确度。

目前，基于采集的数据类型，特征提取与表情识别主要分为基于静态图片的特征提取与基于视频的特征提取两类，也即静态面部表情识别 （Static FER, SFER）和动态面部表情识别（Dynamic FER, DFER）。SFER 借由单张图片开展特征提取，实现表情辨别和情绪判断，主要的特征提取方法为全局人脸特征提取和局部人脸特征提取，前者将人脸视作一个整体，通过识别整体面部人脸的形变来提取特征。FFER 考虑到面部表情动态变化的时空特性，试图通过对多张图片的动态序列进行特征提取，以实现表情与情绪状态的理解与判断，通常能够实现更有效、更精确的面部表情识别。

3. 基于 FER 的教育游戏学习情感计算研究进展

在教育领域研究中，情感计算主要用于精准识别与判断学习者的情绪，也即学习情感识别与计算（翟雪松 等, 2022）。当前国内外基于面部表情识别在教育游戏学习情感计算中的研究主要集中于：（1）真实学习情境中的 FER 与数据库构建；（2）学习者情绪情感监测与评估系统构建；（3）基于多模态的学习情感计算。

3.1. 真实学习情境中的面部表情识别或数据库构建

以往研究通常使用现成的面部表情数据库，这些数据库大多基于实验室环境或某种特定场合采集，仅有适合该特定场合的几种情绪分类。然而，学习者通常在传统课堂、网络学堂或其他学习情境中开展学习，且在此过程中产生的表情远不只限于 6 种基本表情。随着情感计算技术的发展，FER 的关注点逐渐转移到利用深度学习技术来识别和构建真实学习场景中常见的面部表情。

有研究表明，学习者在学习过程中常出现的情感更多表现为困惑、无聊、中性、好奇、喜悦、焦虑和沮丧（Kolog et al., 2019），其中尤其是困惑的表达能够作为学习过程成功一个良好指标。在学习过程中，学习者可能常出现“疑惑”这一情绪。倘若未能及时解决学习者的困惑，可能造成其动机兴趣不足、知识能力学习不到位等问题。然而，学习者在在线学习中常出现困惑情绪，而长时间的困惑可能引发学习僵局。以往研究对困惑情绪的关注较少，且大多主要基于文本信息、学习行为、生理信号等方式开展识别，在准确率、推广性等方面存在诸多局限。

基于 FER 方法与技术，江波等（2018）提出了一组基于面部表情的困惑情绪自动识别算法。这一研究在实验室环境下开展，招募某大学 20 名在校研究生为被试（男女比 1:1）。被试在电脑上完成 30 道英语测试题（学生选择容易、中等、困难三种难度），同时利用 Intel VF8010 摄像头采集学生的面部表情。每道题采集 5 张照片，每位学生共采集 150 张照片，实验总计采集照片 3000 张。在定义数据标签方面，采用“困惑”与“不困惑”两种标签，只有当被试对题目含义及考察知识点完全了解的情况下才被定义为“不困惑”。然而，该研究建立的困惑情绪数据库虽然能够作为基本情绪数据库的补充，但也是基于实验室环境开展的。这一研究只研究了“困惑”情绪，并未关注其它学习过程中经常出现的情绪。此外，研究只判断学生是否产生困惑情绪，并未对其情绪强度进行区分。

3.2. 学习者情绪情感监测与评估系统构建

学习过程中常伴随着情绪情感的产生，尤其像计算机界面、智能辅导系统、模拟环境、动画、沉浸式游戏等多媒体学习环境，不仅能够为学习者提供学习平台与互动机制，也可能在交互过程中促成更丰富、更强烈的情感产生。比如，有研究表明，学习者可能在多媒体学习环境中更频繁地出现投入、挫败、厌倦等情绪；此外，师生面对面交流的匮乏可能导致“情感缺失”现象的产生与凸显，造成学习者的孤独感和倦怠感（De Baker et al., 2010）。随着信息技术的发展，研究者开始关注实时监测学习者在学习过程中的情感状态，以及构建相关的评估或预测模型。教育游戏能够为学习情感计算提供天然的隐性评估平台，由此受到广泛关注。

Verma 等（2022）构建了化学游戏《Chemo-o-crypt》，基于 Ekman 提出的 7 种基本情绪构建情感标签。游戏使用 Affectiva 的 Affdex SDK 自动检测和评估学生在游戏化学习过程中产生的无聊、心流、挫败感等情感状态，实现游戏化学习环境中的隐性评估（stealth assessment）。研究区分了采用隐性评估和适应性调整机制的实验组，以及无控制条件的对照组，通过比较两组学习者的情感体验与学习效果，验证了监测情绪和调整游戏可以改善低领域知识参与者的学习，且有助于缓解无聊和沮丧等消极情绪。该游戏系统实现了对学习者情感的隐性评估，并在此基础上尝试自适应地调整游戏玩法，匹配学习者的化学知识水平和学习体验，以激发学习者兴趣，促使其愿意停留在学习过程中，并尽可能最大化地学习。

3.3. 基于多模态的学习情感计算

一般来说，仅依靠单模态数据的分析存在一定的数据丢失率和时空局限。当前关于面部表情识别的研究通常针对整个学习过程，难以揭示不同微小时刻的情绪体验，且对每种情绪识别的精确度也不同。此外，人脸是一个多重信息系统（Ekman & Friesen, 2003），不同的关键部位对面部表情识别的贡献度也不同。比如有研究表明，人脸中眼睛和嘴巴的重要程度大于鼻子，而眼球转动则承载了较为快速的面部信号。将多种数据模态结合，可能有助于规避数据损失，提高数据利用率与识别精确度。结合具有较高准确度的生理测量形式（雷婕与丁亚平, 2013），如眼球数据的追踪与分析和大脑成像技术，有助于从“眼睛”的视角出发探索学习过程中的即时加工过程，从大脑结构的角度出发挖掘状态变化（隋雪与任延涛, 2007）（程萌萌 等, 2013）……

学习分析与基于传感器技术的发展提供了多模态学习情感计算的条件。结合学习者的基础数据、外显行为（如文本、对话、语音语调、身体姿势、肢体动作、面部表情等）、生理指标（心率、皮肤电、眼动、脑电、肌电、体温等）、日志与键鼠等多模态数据，能够指导研究者从不同的时间特性和维度来进行学习评价（张家华 等, 2022）、开展情感识别与分析（黎孟雄 等, 2015）。

Emerson 等（2020）使用游戏《水晶岛（Crystal Island）》，招募了具有不同先验知识水平的学生。在游戏过程中，研究者采集了学生的游戏玩法、面部表情、眼神注视、后测成绩和自我报告的兴趣等数据，并对游戏玩法、眼神凝视和面部表情等不同模态数据进行组合，构建单模态、双模态或多模态数据学习模型，用于预测学生的后测成绩或兴趣程度。结果表明，在预测后测成绩方面，两种模态相结合构建的模型，表现优于仅用游戏玩法构建的模型；在预测兴趣程度方面，两种模态相结合的模型，表现优于仅用游戏玩法构建的模型。研究者还指出，由于许多游戏玩法中的功能是基于单个游戏内对象参数化的，由此导致眼神凝视数据的加入造成数据冗余，使得基于游戏玩法和面部表情双模态的预测模型表现优于仅基于眼神凝视数据的模型，而基于三种数据模态的预测模型表现差于双模态预测模型。

4. 基于 FER 的教育游戏应用案例

4.1. 实验工具

本研究的实验工具包括：（1）基本信息问卷，包括实验对象的姓名、性别、年龄、班级；（2）电脑自带摄像头，实现对学习者面部表情的实时记录；（3）分数教育游戏《分数跑酷》（Run Fraction）；（4）情感分析软件 FaceReader。

其中，《分数跑酷》为研究者所在实验室于 2018 年开发的教育游戏，主要侧重于帮助小学阶段的学生理解分数数值与“数轴”概念。游戏包括认知体验与主观性体验两方面的设计体系。其中，认知体验通过认知辅助、分数表征、内容框架和教学支架等四个主要模块的设计，实现促进学习者对分数的理解，将“数轴”这一抽象概念转化为“跑酷”运动中的道路，通过闯关形式理解与掌握分数知识。主观性体验方面，游戏的设计充分使用了游戏化元素，以试图促进学习者的情感体验、增强其学习投入。具体体现为：1）3D 场景设计，2）可选择的游戏化身，3）声效设计与动效设计， 4）简单明晰的故事线设计， 5）奖惩机制引发不确定性，游戏中包含奖励、积分、排行榜 3 种游戏化元素，其中数轴墙和能量星两种奖励取决于玩家的认知表现，与其努力程度有关，而随机出现的额外奖励则与其运气有关，这两种类型的奖励保证了游戏化学习过程中的不确定性，更有助于调动学习者的不同情绪。

4.2. 数据来源及实验流程

本文所用视频数据来自于研究者所在实验室先前开展实验时采集的数据，该实验选取北京市顺义区某所小学的四至五年级学生，被试年龄范围在 9-12 岁之间，均为自愿参加本次

实验，同意并知情实验内容。所有受试者均能独立顺畅地操作计算机，具有较高的信息素养，且在本实验前从未观看或使用过实验所用教育游戏。

考虑到学习者的真实学习环境，在学校教室中开展实验，实验场所光线充足，对比度充足，使用电脑自带摄影头完成对受试者面部表情的实时记录。在实验过程中，让每位受试者玩分数教育游戏《分数跑酷》的某个模块。通过镜头实时采集学生的面部表情视频，视频记录为 WMV 格式；同时通过屏幕录制采集游戏过程视频，视频记录为 MP4 格式。考虑到学校课程安排时间与学习者注意力集中时长，每次数据采集时间不超过 50min，最终所得每段视频时长均在 20 分钟至 30 分钟之间。

4.3. 数据分析及主要发现

本研究首先对采集到的数据进行了处理，包括视频筛选、剪切、对齐和标注。1）筛选：对拍摄与录制的视频进行筛选，将受试者未完整露出脸部、脸部有大幅度遮挡及长期移动身体的视频排除，确保视频中只有一个清晰可见的人脸，以便剩余视频能够使用软件开展后续面部表情识别与分析。2）剪切与对齐：将面部表情录制视频与屏幕录制视频进行剪辑与对齐，确保视频时长与录屏时长保持一致视频，然后剪切为 20-30 分钟之间的视频片段，本研究中展示 10 段视频分析的结果。3）标注：为学习视频打上标签。

由于软件的模型质量并不完全等同于情感识别的精度，在对视频进行分析之前，先对每个参与者的面部表情进行校准，以减少由于参与者本身表情或长相特征带来的“中性”表情识别误差。比如，有的学习者天生嘴角略微下垂或习惯性做撇嘴动作，可能被软件识别为主要呈现悲伤情绪；有的孩子天生喜欢皱眉，可能被软件识别为主要呈现愤怒情绪；甚至可能因为有的孩子天生眉毛上扬，可能被软件识别为主要呈现厌恶或轻蔑情绪。校准结果良好。

基于以往研究，将视频中的情感标签定义为高兴、悲伤、惊讶、困惑、愤怒、中立、轻蔑、无聊、恐惧、厌恶 10 种状态。由于东亚人群与欧洲人群的面部特征具有差异，选择东亚人群的面部表情分析模型（准确率达 96%）。将剪切完成的视频导入 FaceReader，按照动作单元分类，识别受试者的面部状态。

对所有参与者的所有表情信息进行识别与分析，发现中性情绪出现最多，随后较常出现的表情依次为愤怒、悲伤、愉快、惊奇。厌恶、恐惧、轻蔑、困惑等情绪出现较少，无聊情绪几乎没有出现。总的来说，该教育游戏能够调动学习者的情绪，参与者在此过程中总体感受较为平静，较少走神或感到无趣，在一定程度上说明游戏设计符合学习者的认知水平与兴趣偏好。

对每位参与者的面部表情进行分析，首先分析每位参与者的总体表情比例。根据参与者的表现与表情比例，大概可将其分为三种类型：平静型、激动型和悲伤型。

其中，**平静型学习者**在学习过程中大多时候情绪平稳，表现为“中性”情绪最常见，其他诸如愉快、悲伤、愤怒、惊奇等情绪出现比例均低于中性情绪的 50%，且总体比例平均。**激动型学习者**在学习过程中出现“愤怒”情绪显著高于除中性以外的其他情绪，面部肌肉常表现为眉毛降下、上眼睑提起、嘴角下撇、嘴唇拉伸，看起来神情紧张。**复合型学习者**在学习过程中，除中性情绪之外，有 2 种或 2 种以上其他情绪出现比例较高（超过中性情绪的 50%）或比例均衡。比如图 9 中，学习者除中性情绪以外，悲伤、愉快、愤怒三种情绪出现比例均衡，且均超过中性情绪的 50%。再比如图 10 中，学习者愤怒情绪出现最高，但同时中性和悲伤情绪出现比例均衡，且均超过愤怒情绪的 50%。总的来说，复合型学习者大多数时间限制表现出不止一种情绪主导。

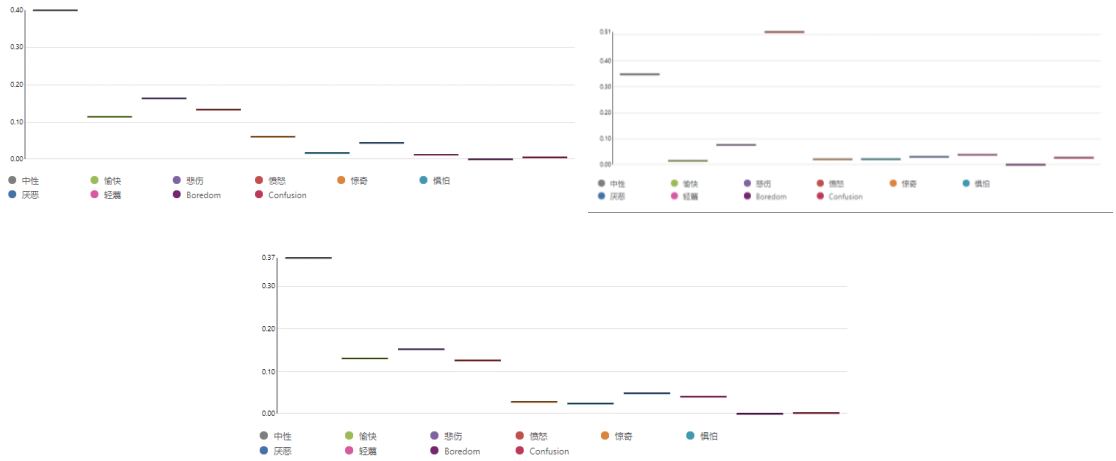


图 2 三种类型学习者情绪比例示意

5. 基于 FER 的教育游戏学习情感计算面临挑战与应用策略

虽然 FER 在学习情感计算中的应用受到广泛关注与认可，但同时也面临一些挑战：（1）目前受到普遍认可和广泛使用的数据库均基于离散情绪，公认的基于连续情绪模型的数据库数量较少；（2）面部表情识别依赖于算法的准确度和鲁棒性，对研究者的技术要求较高；（3）对受试者要求严格，面部遮挡或大幅度转动都会影响识别的准确度；（4）现有面部表情数据库通常在实验室环境下诱发形成，受试者面部表情幅度大（Wang et al., 2015），与真实学习环境中产生的幅度更小（魏艳涛 等, 2022）、持续时间更短的表情存在一定差异（翟雪松 等, 2022）；（5）仅对面部表情这一单模态数据分析，难以避免信息缺失现象，且情感标注过程易受被试和主试的主观性影响，导致基于 FER 的学习情感分析相较于多模态学习情感分析而言可解释性较弱。

6. 总结

情感计算提供了基于计算机的情感测量与评估方式，FER 具有操作简便、准确度高等特征，能够通过外显的数据采集与识别，实现对受试者内隐的情绪情感状态的发现与揭示。将 FER 应用于教育游戏，既为解释学习者内在学习过程与变化提供了新的视角，有利于从更加科学的角度回答学生在学习过程中心理变化的追问；同时也为教学者提供及时合适的隐性评估和个性化学习与情感支持提供了依据，有助于规避学习者在学习过程中的消极情绪，丰富其积极的学习体验与情感体验。未来研究可以关注与捕捉学习者“瞬时”出现的真实表情；结合键盘鼠标、学习日志、眼动、脑电、皮肤电等设备采集数据，从多模态学习情感计算的角度，更精确地还原与探索学习者的情绪变化、情感体验及其发生机理。

参考文献

王欢欢, 吕紫藤, & 李现昆. (2022). 基于表情识别技术的用户研究方法. *包装工程艺术版*, *43*(2), 116-121.

江波, 李万健, 李芷璇, & 叶韵. (2018). 基于面部表情的学习困惑自动识别法. *开放教育研究*, *24*(4), 101-108.

陈子健, & 朱晓亮. (2019). 基于面部表情的学习者情绪自动识别研究——适切性, 现状, 现存问题和提升路径. *远程教育杂志*, *37*(4), 64-72.

徐晓青, 赵蔚, & 刘红霞(2018). 混合式学习环境下情绪分析应用与模型研究——基于元分析的视角. *电化教育研究,39(08)*,70-77.

徐琳琳, 张树美, & 赵俊莉. (2017). 基于图像的面部表情识别方法综述. *计算机应用*, *37*(12), 3509-3516.

隋雪, & 任延涛. (2007). 面部表情识别的即时加工过程. *心理学报*, *39*(1), 64-70.

程萌萌, 林茂松, & 王中飞. (2013). 应用表情识别与视线跟踪的智能教学系统研究. *中国远程教育*, (5), 59-64.

程静, & 刘光远. (2012). 学科交叉视角下的情感识别研究进展. *计算机科学*, *39*(5), 19-24.

雷婕, & 丁亚平. (2013). 面部表情: 一些争论. *心理科学进展*, *21*(10), 1749.

翟雪松, 许家奇, & 王永固. (2022). 在线教育中的学习情感计算研究——基于多源数据融合视角. *华东师范大学学报(教育科学版)*, *40*(9), 32.

黎孟雄, 郭鹏飞, & 黎知秋. (2015). 基于情绪识别的远程教学自适应调节策略研究. *中国远程教育*, (11), 18-24.

魏艳涛, 罗洁琳, 胡美佳, 李文昊, & 姚璜(2022). 基于视频的在线学习情感识别研究. 计算机科学, 49(S02), 6.

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, *1*(1), 21-62.

De Baker, R. S. J., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: the incidence and prevalence of affect during interactions with three computer-based learning environments. *Int. J. Hum. Comput. Stud*, *68*, 223-241.

Bijlstra, G., Holland, R. W., & Wigboldus, D. H. J. (2010). The social face of emotion recognition: evaluations versus stereotypes. *Journal of Experimental Social Psychology, 46(4)*, 657-663.

De la Torre, F., & Cohn, J. F. (2011). Facial expression analysis. *Visual analysis of humans: Looking at people*, 377-409.

Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. (2020). Multimodal learning analytics for game-based learning. *British Journal of Educational Technology*, 51(5), 1505-1526.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues* (Vol. 10). Ishk.

Friend, M., & Bryant, J. B. (2000). A Developmental Lexical Bias in the Interpretation of Discrepant Messages. *Merrill-Palmer Quarterly*, *46*(2), 342-69.

Hale, A. J., Freed, J., Ricotta, D., Farris, G., & Smith, C. C. (2017). Twelve tips for effective body language for medical educators. *Medical Teacher*, *39*(9), 914-919.

Kolog, E. A., Devine, S. N. O., Ansong-Gyimah, K., & Agjei, R. O. (2019). Fine-grained affect detection in learners' generated content using machine learning. *Education and Information Technologies*, *24*(6), 3767-3783.

Mehrabian, A. (1971). *Silent messages* (Vol. 8, No. 152, p. 30). Belmont, CA: Wadsworth.

Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies*, *59*(1-2), 55-64.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of Emotion*, 3-33.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Suwa, M. (1978). A preliminary note on pattern recognition of human emotional expression. *Proc. IJPR*, 408-410.

Verma, V., Craig, S. D., Levy, R., Bansal, A., & Amresh, A. (2022). Domain knowledge and adaptive serious games: Exploring the relationship of learner ability and affect adaptability. *Journal of Educational Computing Research*, *60*(2), 406-432.

Wang, S., Wu, C., He, M., Wang, J., & Ji, Q. (2015). Posed and spontaneous expression recognition through modeling their spatial patterns. *Machine Vision and Applications*, *26*, 219-231.

融入環境議題概念之教育遊戲軟體輔助國小學生環境覺察課程學習效益探究

Research on the effectiveness of educational software to assist primary school students in learning environmental awareness courses on key environmental issues learning concepts

王曉璿[1]、劉金英[2]、楊淇鈞[3]、林宥誼[4]、李昀軒[5]
[12345]國立臺中教育大學數位內容科技學系
* adt109105@gm.ntcu.edu.tw

【摘要】 本研究主要以 2D 卷軸遊戲之第三視角，綜觀故事劇情與角色經歷，並加入道具解謎、冒險戰鬥等要素，進一步認識臺灣主要發生的各種環境災害，提高玩家對於保護環境的意識、了解災害的發生與預防，並深化學習者愛護地球的心。

【關鍵字】 環境覺察；數位遊戲式學習；2D 卷軸環境教育遊戲

Abstract: *This study adopts a third-person perspective in a 2D side-scrolling game. It includes puzzle-solving with items and combat, aiming to raise awareness about environmental disasters in Taiwan. The goal is to deepen understanding of disaster occurrence and prevention. Through gameplay, players are urged to become more conscious of environmental protection and develop a strong sense of care for the Earth.*

Keywords: Environmental Awareness, Digital Game-Based Learning, 2D Scroll Environmental Education Game

1. 前言

臺灣常因地震、土石流等天然災害，對社會帶來負面影響。臺灣教育制度長期偏重學科教學，學生為應付考試而消極面對相關之防災教育，導致防災課程結束後，便忘記防災相關知識（張全亨，2009）。加上防災課程多著重於知識、學習層面，較忽略情感與技能，使學生對防災教育抱持輕忽的態度，以至防災教育的成效極為有限(趙家民、張宏仁，2007)。

為因應現今的教育制度，強化國家災害防救中心對中小學學生防災意識的宣導，本研究嘗試以環境災害融入多元領域素養的遊戲，傳授知識，且結合道德和價值觀，期望數位教材讓學習者跳脫課本框架，達成「虛實合一」的學習效果，加深對環境防災的基礎知識和技能。

2. 文獻探討

本章將進行與研究內容有關之主題的文獻整理。以下分為三個小節來做探討：

2.1. 環境災害的定義

王秋原、徐美玲（1990）指出：『環境』是指人類活動領域的總稱，包括許多環境因素，這些因素與人類活動互相作用。當環境因素的變化導致人類生命和財產受損時，可稱之為『環境災害』。由於災害的成因多種多變，必須從自然因素和人為因素分別追究其原因，才能更有效地進行防災和減災工作（黃朝恩，2000）。

2.2. 環境災害教育遊戲設計

陳若函（2014）提出，數位遊戲式學習（digital game-based learning）能提升學習動機、

專注力及興趣，使學習變得有趣，幫助學習者深入學習。Prensky（2006）指出，21 世紀的孩子們是數位科技的「原住民」（natives），而我們則是「移民者」（immigrants）。孩子的接觸經驗影響他們的思考方式，使其偏好遊戲化學習，這種學習方式是主動的、互動性、挑戰性、有決策權、有目標性的。教師可利用教育遊戲元素來豐富教學方法。臺灣教育偏重學科教學，學生對防災課程常感被動，導致學習成效有限（張全亨，2009）。相較於傳統教育方法注重理論和知識傳授，數位遊戲式學習可以激發學生的學習興趣和參與度（Fujiwara et al., 2012）。因此本研究透過數位遊戲進行環境教育，以提升學生的防災知識和素養。

2.3. 融入數位教學的成效

數位遊戲式學習具備多種特質，如目標與任務、挑戰性、真實性、探索、互動、回饋、競爭、合作及行動學習等，可將抽象的學習概念轉化為具體的實例呈現在學習者面前（Shih, J. L., Shih, B. J., Shih, C. C., Su, H. Y.,& Chuang, C. W.2010）。因此遊戲是學習者一自然且必要的學習元素，鼓勵融入教學策略中，降低學習的乏味感（Prensky, 2001）。Ferdig 提到將建構主義或社會建構理論架構引入遊戲中，例如：探究式教學，可促進學習者的學習成效（Ferdig, 2007）。故本研究以融入環境議題概念之教育遊戲軟體探究國小學生在環境覺察課程的學習效益，透過數位遊戲式學習進行環境教育，使其具備防災相關知識與環境覺察能力。

3. 研究方法

3.1. 遊戲設定族群

本研究教育遊戲軟體之使用者設定上，以國小高年級學生為主，呼應譚至晢所提：「國小教育是各階段教育的基礎」，且十二年國民基本教育課程綱要中已將環境教育納入各教育階段之議題，期望學生認知環境倫理、永續發展、氣候變遷、災害防救、能源資源永續利用五大實質內涵（國家教育研究院，2020），強化大眾防救災素養能力，降低災害發生的損傷及風險。

3.2. 遊戲設計

本研究使用的教育遊戲「尋境 NESOI」，描述少年和小精靈揭開環境災害秘密的故事，藉道具解謎和探索活動，認識臺灣環境災害，提升保護環境和災害預防的意識。

3.2.1. 文本設計

玩家藉劇情推演與對話、觀察場景，獲取災難特徵、發生因素與防範等資訊。故事描述大地女神守護土地，卻因人類破壞生態平衡，引發災難。主角們先到受土石流摧毀的村莊，經對話了解災害前兆，為應對颱風和土石流做好準備。下一個目的地他們收集藥材製成藥劑救治村民。最後抵達根源之地，面對廢棄工廠和淨化神像的挑戰，成功喚醒守護神。

3.2.2. 場景設計

本研究各關卡場景呼應災難主題與背景故事，第一關設計在山腳下靠海村莊，以泛紅的天空、枯萎植物，呈現颱風及土石流發生前的景象；第二關為偏遠山區村莊，玩家需幫助生病的村民和採集草藥；第三關為沙漠中的城市展示全球暖化的景象。

3.3. 遊戲機制

本研究使用環境感知教育遊戲進行互動。通過控制主角移動，觸發與 NPC 的對話和物件，完成關卡任務，推動劇情發展並達成最終结局。劇情對話、文字描述和物件說明使用醒目顏色標記，讓玩家快速獲取關卡相關知識、解謎提示和重要劇情信息。

3.4. 遊戲整體設計呈現與修正歷程

針對遊戲設計與其修正，以下分為遊戲整體設計呈現、遊戲修正歷程兩部分呈現說明：

3.4.1. 遊戲整體設計呈現

表 1 遊戲整體設計呈現表格

類型	畫面示意圖	設計說明
主角		約 11、12 歲之少年，並將服裝、配件繪製成旅行者相關元素。
小精靈		角色設計為圓潤蘿蔔形狀，頭頂有綠芽，展現與自然的連結與活潑特質。
NPC		居住於不同關卡中的角色，其外型以老人、青年、少女的外型進行繪製，藉由對話推動關卡劇情故事。
主畫面		UI 設計以藍綠色調和綠葉元素，體現環境覺察主旨，與自然呼應。
遊戲場景		第一關以颱風與土石流為主題，畫面呈現昏暗的橘褐色系，加上枯萎植物和泛紅天空的景象。

3.4.2. 遊戲修正歷程

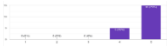
表 2 遊戲修正歷程表格

類型	第一版	第二版	修正說明
主畫面			原 UI 介面為簡約風格，後期調整為綠色系並加入綠葉元素，以符合環境主題並豐富畫面。
遊戲說明			遊戲說明畫面由初期的灰白紅簡約風格，調整為更貼近自然的藍綠色系。
主角設計			初版主角身材高挑、臉型削瘦；根據建議，調整為 Q 版、臉型圓潤，更適合 11 至 12 歲男孩形象。
場景			初版場景空曠、色彩鮮明，經回饋後調為昏暗橘褐色，符合颱風與土石流主題。

4. 結果與討論

本研究以問卷蒐集使用者建議，針對介紹與操作、吸引遊玩因素、整體體驗等面向，探究遊戲是否適合國小學童操作學習環境議題，以及遊戲修改之建議。以下為回饋之結果：

表 3 問卷回饋蒐集結果表格

遊戲介面與操作說明是否清楚	是否適合教育國小學童環境覺察和環境教育之議題學習	遊玩難易度是否適合國小高年級學童遊玩
		
有哪些內容吸引你前來遊玩	遊戲體驗滿意度	體驗後是否發現以下缺點
		

4.1. 遊戲劇情設計修改

本研究經問卷調查與建議後，針對遊戲劇情進行修正：1.針對國小高年級玩家，將遊戲中艱深的詞彙簡化成易懂的文字 2.縮減文本長度，避免世界觀過於複雜，玩家無法理解 3.增加故事或遊戲彩蛋，藉此增加玩家探索遊戲之慾望。

4.2. 學習內容修正與調整

本研究經調查與建議後，針對遊學習內容進行修正：1.以環境教育為核心，將臺灣常見災害作為各關卡主題呈現，以因應時間、技術等考量而進行討論後確定 2.針對國小高年級學

童，修改劇情篇幅與遊戲玩法 3.透過各關卡遊戲操作檢核玩家的學習成效。

4.3. 整體美術與程式修正歷程

本研究經調查與建議後，針對美術與程式進行修正：1.角色調整為Q版，使外型圓潤可愛 2.場景整體色系改為昏暗的橘褐色，增加素材以豐富場景外觀 3.UI 設計中加入綠色、藍色、綠葉等元素以符合「環境」概念 4.統一角色立繪在對話框相同位置，避免視覺疲勞。

5. 結論與建議

綜上所述，本遊戲《尋境 NESOI》，期望透過角色扮演、對話和劇情設計，讓玩家代入角色情緒和思考，以吸引玩家並達到共鳴，同時傳達正面積極意象，提升環境意識。

有鑑於近年綠色意識抬頭，政府推出節能減碳與永續發展相關政策，在教育上將環境教育納入課綱，期許學生將知識與技能落實至日常生活，普及「環境教育」至家庭、學校、社區，增強社會大眾防救災素養，且呼應譚至晢所論:「國小教育是各階段教育的基礎（譚至晢，2002）。」，故本研究以環境覺察為主題之教育遊戲，針對國小高年級學童，並因應新冠肺炎（Covid-19）對教育帶來的挑戰，期望該遊戲成為教師教授環境知識的輔助工具和學生的自學平台，讓師生可以不限時間、地點使用，透過遊戲引發學生好奇心，培養其正確的環境素養。

參考文獻

王秋原、徐美玲（1990）。**影響基隆河流域居民對其住家環境之識覺及其因應行為因子之研究**。中國地理學會會刊，18，23-39。

國家教育研究院（2020）。**十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校議題融入手冊**。

張全亨（2009）。**高中生颱風和水災概念理解與防災素養之現況調查**。國立臺灣師範大學地球科學研究所，臺北市。

陳若函（2014）。**數位教育遊戲之開發與評估：以「Mr.道耳頓的奇幻歷險」為例**。國立中央大學網路學習科技研究所。

黃朝恩（2000）。人類活動所激化的自然災害。**環境教育季刊，41**，49-56。

趙家民、張宏仁（2007）。**國民小學實施防災教育成效之研究--以南投縣內湖國小學童為例**。臺北市立教育大學環境教育與資源研究所。

譚至晢（2002）。**臺灣中部地區國小教師對自然災害的防備態度及因應行為之研究**。國立臺中教育大學環境教育研究所。

Prensky, M. (2006). Don’t bother me, Mom, I’m learning!: How computer and video games are preparing your kids for 21st century success and how you can help! St. Paul, MN: Paragon House.

Ferdig, R. (2007). Preface: Learning and teaching with electronic games. Journal of Educational Multimedia and Hypermedia, 16(3), 217-223.

Fujiwara, T., Yamada, F., Okuyama, M., Kamimaki, I., Shikoro, N., & Barr, R. G. (2012). Effectiveness of educational materials designed to change knowledge and behavior about crying and shaken baby syndrome: A replication of a randomized controlled trial in Japan. Child Abuse & Neglect, 36(9), 613-620.

Prensky, M. (2001). Digital game-based learning. New York: McGraw-Hill.

Shih, J. L., Shih, B. J., Shih, C. C., Su, H. Y.,& Chuang, C. W.(2010). The influence of collaboration styles to children’s cognitive performance in digital problem-solving game “William Adventure”: A comparative case study. Computers & Education, 55(3), 982-993.

指向计算思维培养的教育剧本游戏设计与实践研究——以浙教版五年级信息科技“算法的控制结构”单元为例

Research on the Design and Practice of Educational Script Games Targeting the Cultivation of Computational Thinking——Taking Algorithm Control Structure Zhejiang Education Press Fifth Grade Information Technology as an Example

潘梦媛[1*]，黄璐[2]，王运弘[3]，章苏静[4]
杭州师范大学经亨颐教育学院
15837683638@126.com

【摘要】 计算思维作为信息科技课程学科核心素养之一，是数字化时代学生必备的基本素养与能力，如何培养学生的计算思维已成为信息科技课程教学研究的重要研究课题。教育剧本游戏具有教育性、文学性、游戏性的特征，有助于激发学生学习动机，提升问题解决能力，为培养小学生计算思维提供了新途径。本研究以小学生为研究对象，浙教版小学信息科技课程五年级“算法的控制结构”单元为研究内容，设计小学信息科技课程教育剧本游戏学习活动。通过对比实验发现，教育剧本游戏学习活动能够较好地提升小学生计算思维能力，为信息科技教学提供了创新教学案例。

【关键词】 计算思维；教育剧本游戏；信息科技

Abstract: *Computational thinking, as one of the core competencies of information technology courses, is an essential basic literacy and ability for students in the digital age. How to cultivate students' computational thinking has become an important research topic in information technology course teaching and research. Educational script games have the characteristics of education, literature, and gameplay, which help stimulate students' learning motivation, improve problem-solving abilities, and provide a new way to cultivate primary school students' computational thinking. This study focuses on primary school students, with the fifth grade unit of "Algorithm Control Structure" in the Zhejiang Education Press Primary School Information Technology Curriculum as the research content, and designs educational script game learning activities for primary school information technology curriculum. The experimental results show that educational script game learning activities can effectively enhance the computational thinking ability of primary school students, providing innovative teaching cases for information technology teaching.*

Keywords: computational thinking, educational script game, information technology

1. 问题提出

《义务教育信息科技课程标准（2022 年版）》（以下简称新课标）明确指出，计算思维是信息科技课程学科核心素养之一，对学生的全面发展具有重要意义(熊璋等, 2022)。信息科技课程教学应顺应时代发展的需要，着力培养学生的计算思维，提高学生利用计算思维解决实际问题的能力。然而，在当前的信息科技课程教学中，存在着教学方法陈旧、学生学习兴趣低等问题，学生较少有参与和探究的机会，不利于学生计算思维的培养(彭少霞, 2023)。近年来，剧本游戏被引入教育领域，其优势逐渐凸显，有研究表明，剧本游戏学习活动能提高学生的学习动机和参与度，学生在游戏活动中轻松愉悦的思考和行动，会发挥更大的潜力

(廖紫兰 & 张曼, 2022)。此外，新课标提出“注重培养学生在真实情境中综合运用知识解决问题的能力，学会在真实情境中发现问题、解决问题”，要求强化课程的综合性和实践性，注重活动化、游戏化、生活化的学习设计。基于此，本研究以小学生为研究对象，针对浙教版小学信息科技课程五年级“算法的控制结构”单元，设计教育剧本游戏学习活动，并分析其教学效果，以期为小学生计算思维的培养提供新途径、新思路和新方法。

2. 文献综述

教育剧本游戏借鉴了教育戏剧的形式(武怡帆 & 姚春敏, 2023)，通过创设情境、提供场景和情节，让学生进行角色扮演等方式来展开学习活动，同时融入剧本杀的“推理”元素(杨紫馨 & 王艳, 2022)，借助环境、推理与自我演绎、DM 引导、互动社交等方式，赋予参与者极强的沉浸感、参与感。基于此，本研究认为教育剧本游戏是以核心素养为导向，学生体验为中心，通过情境推理、角色扮演、游戏活动学习学科知识和跨学科知识的综合学习活动，具有教育性、文学性和游戏性三个特征。

计算思维由周以真教授于 2006 年首次提出，她认为计算思维是一种解决问题的思维过程并有助于提升解决问题的能力(Wing, 2006)。新课标将计算思维界定为个体运用计算机科学领域的思想方法，在解决问题过程中，涉及的抽象、分解、建模、算法设计等思维活动。计算思维培养的途径主要有编程、机器人教育、游戏、基于设计的学习活动等(Shute 等, 2017)。张学军等人认为游戏提供了一个“有意义的框架”，通过游戏的方式可以解决编程教学枯燥乏味、难以激发学生兴趣的难题，为计算思维的培养创造了良好的氛围。

3. 基于教育剧本游戏的“算法的控制结构”单元学习活动设计

3.1. 教育剧本游戏设计原则

教育剧本游戏设计遵循以学生为中心、融合学科知识、注重互动合作和强调问题解决四个原则，以确保学习活动的有效性。首先是以学生为中心，应关注学生认知发展阶段，根据其学习能力，将抽象概念具象化，激发学习热情。其次是融合学科知识，在设计游戏学习活动时，应该根据教学目标要求，将教学内容与游戏活动相结合，以达到教育性和游戏性的平衡。再次是注重互动合作，鼓励小组探究性学习，利用学生间的多样性，体验团队合作的乐趣，提高解决问题的能力。最后是强调问题解决，学生需根据不同的问题情境，学会对复杂的问题进行抽象和分解，会用流程图表达算法，从而学会解决问题。

3.2. 教育剧本游戏学习活动设计

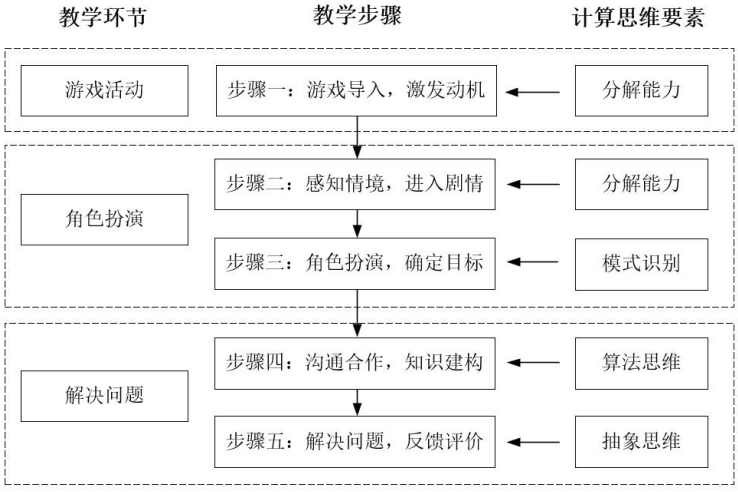


图 1 基于教育剧本游戏的单元学习活动

依据设计原则，结合浙教版信息科技五年级上册“算法的控制结构”单元内容，本研究设计了基于教育剧本游戏的单元学习活动，包括三个环节五个步骤：环节一，在游戏活动中了解算法控制结构的基本类型；环节二，在角色扮演中分析问题，提炼解决问题的办法；环节三，在解决真实性问题中掌握算法的基本结构。如图 1 所示。

3.2.1.环节一：游戏活动

在“游戏活动”环节，教师导入活动主题，描述故事背景：穿越到南宋临安城破解谜题。学生感知情境投入其中并在游戏活动中了解顺序结构。该环节从具体的故事情境出发，吸引学生注意力，激发学习兴趣和动机，进而投入课堂引发情感体验，培养学生发现问题与分析问题的能力。

3.2.2.环节二：角色扮演

在“角色扮演”环节，教师带领学生穿越到南宋临安城，在案发现场搜集线索，调查案件并找出间谍。学生根据剧本要求，扮演古代相应角色，运用顺序结构通过小组合作画出调查案件及寻找间谍的流程图，并根据流程图的步骤完成任务。该环节让学生体验把解决问题的步骤用流程图进行表述，培养学生模式识别与分解问题的能力。

3.2.3.环节三：解决问题

在“解决问题”环节，教师提供任务线索，有语音信息和地图等。学生首先从语音信息中提取线索，获得文案藏匿地点，运用分支结构验证地点是否正确。接着小组成员在现场搜索证据，通过组员的相互合作分析推理关键信息，整合各种线索破解难题，利用单分支和双分支结构的知识，合并寻找文件流程图，最终通过挑战，找到三份文件的下落，成功完成此次临安城探秘任务并找出间谍。最后，教师带领学生复盘反思，针对整个学习活动进行点评与总结，学生感悟体验与收获。该环节将算法知识融入游戏活动和情境学习任务中，使抽象思维具象化，让学生灵活运用算法的基本结构解决问题，进一步强化知识的应用，培养学生算法思维及解决问题的能力。

4. 基于教育剧本游戏的“算法的控制结构”单元学习效果分析

4.1.研究对象

本研究选取杭州市某小学五年级两个平行班为研究对象，共 80 名学生，男女性别相当，实验组 40 名，对照组 40 名，教学内容为五年级上册“算法的控制结构”单元，实验组采用教育剧本游戏教学方式，对照组采用传统教学方式。

4.2.研究工具与数据收集

本研究的研究工具采用了小学生计算思维能力量表(张屹等, 2020)，该量表一共有五个维度，分别是创造力、批判思维、问题解决、算法思维及合作能力，每个维度下有一定数量的题目，一共 23 个题项，问卷信度为 0.902，效度为 0.769，信效度良好。

4.3.结果分析

本研究在实验前后分别测量了实验组和对照组学生的计算思维，如表 1 所示。实验前，实验组平均得分为 3.71，对照组平均得分为 3.61，可见两组学生的计算思维水平无显著性差异（t=0.781,p=0.437>0.05）。实验后，实验组的平均得分为 4.07，对照组平均得分为 3.77，可见两组学生的计算思维水平存在显著性差异（t=2.356,p=0.021<0.05）。

表 1 计算思维统计结果

	组别	人数	均值	标准差	t
计算思维前测	实验组	40	3.71	0.58	0.781
	对照组	40	3.61	0.52	

计算思维后测	实验组	40	4.07	0.57	2.356*
	对照组	40	3.77	0.54	

注，*$p<0.05$

通过数据结果可以看出，实验前两组学生的计算思维处于相同水平，实验结束后，两组学生的计算思维水平都有提升，但存在显著性差异，实验组明显优于对照组，说明通过教育剧本游戏教学方式，能够提升小学生的计算思维水平。

5. 结论与展望

本研究设计了基于教育剧本游戏的信息科技单元学习活动，并将其应用于教学实践中。通过实证数据发现，本研究所设计的教育剧本游戏学习活动能够较好地提升小学生计算思维能力。随着新课程改革的深入，信息科技课程内容趋向思维训练，难度相应提高，导致学生学习兴趣降低，本研究对如何有效提升学生计算思维具有重要的现实意义。我们认为，在信息科技课程中融入兼备挑战性和趣味性的教育剧本游戏学习活动，不仅创设了吸引学生学习的环境，激发了学生学习的热情，而且让学生在有趣的游戏活动中进行团队合作，提升了问题解决的能力，促进了计算思维能力的提升。教育剧本游戏作为一种创新性学习方式，有潜力深度整合到更广泛的学习活动中，为学生提供动手操作、意义学习和团队协作的机会。

参考文献

廖紫兰 & 张曼. (2022). 学生兴趣视角下剧本游戏对研学教育活动的优化. 科学咨询(教育科研), *3*, 54–56.

彭少霞. (2023). 新课标背景下初中信息科技计算思维的培养策略. 中国现代教育装备, *6*, 37–39. https://doi.org/10.13492/j.cnki.cmee.2023.06.012

武怡帆 & 姚春敏. (2023). 教育戏剧及其在小学教育中的育人价值与实践. 教育理论与实践, *43*(2), 62–64.

熊璋, 赵健, 陆海丰, & 欧阳元新. (2022). 义务教育阶段信息科技课程的时代性与科学性——《义务教育信息科技课程标准(2022 年版)》解读. 教师教育学报, *9*(4), 63–69. https://doi.org/10.13718/j.cnki.jsjy.2022.04.008

杨紫馨 & 王艳. (2022). 还原·沉浸·交互：论剧本杀剧本的三重创新. 电影文学, *2022*(12), 46–50.

义务教育课程方案（2022 年版）. (2022). 基础教育课程, *2022*(9), 72–80.

张立国, & 王国华. (2018). 计算思维:信息技术学科核心素养培养的核心议题. 电化教育研究, *39*(5), 115–121. https://doi.org/10.13811/j.cnki.eer.2018.05.017

张学军, 岳彦龙, & 梁屿藩. (2021). Python 课程中数字化游戏教学培养高中生计算思维的实证研究. 电化教育研究, *42*(7), 91–98. https://doi.org/10.13811/j.cnki.eer.2021.07.013

张屹, 莫尉, 张岩, 高晗蕊, 李妞, & 林郁菲. (2020). 我国小学生计算思维量表研发与应用. 中国电化教育, *10*, 49–57.

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, *22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

面向成人学习者的医学知识学习系统的设计研究

A Study on the Design of Medical Knowledge Learning System for Adult Learners

陈宇萍[1*]，陈梦洁[2]
[1 2]华东师范大学教育学部教育信息技术学系
[*]51254108019@stu.ecnu.edu.cn

【摘要】自适应学习系统能够根据学生的个体情况和学习需求，个性化地提供各类学习内容与活动形式。目前，医学知识学习系统主要面向医学生和医生群体，针对成人学习者的医学知识学习系统有待进一步研发。本研究先基于访谈了解成人学习者对医学知识学习系统的功能需求，再分析系统的设计原则，以构建系统模型并设计系统原型（系统又称为“医教平台”）。后改编用户体验问卷（User Experience Questionnaire，UEQ）调查成人学习者对系统原型的使用体验，结果显示“医教平台”满足成人学习者对医学知识学习系统的期待。这为进一步开发和推广自适应学习系统在医学教育领域的应用提供了有力的支持。
【关键词】成人学习者；医学知识；自适应学习系统

Abstract: *The adaptive learning system possesses the capability to tailor various learning content and activity formats according to individual student profiles and learning requirements. At present, the medical knowledge learning system is mainly for medical students and doctors, and the system for adult learners needs to be further developed. Based on interviews, this study first understood the functional requirements of adult learners for the medical knowledge learning system, and then analyzed the design principles of the system to build a system model and design a prototype of the system (the system is also called "medical education platform"). After that, a user experience questionnaire (UEQ) was compiled to investigate adult learners' experience of using the system prototype, and the results showed that the "medical education platform" met adult learners' expectations for the medical knowledge learning system. This provides strong support for further developing and popularizing the application of adaptive learning system in the field of medical education.*
Keywords: Adult learner, Medical knowledge, Adaptive learning system

1. 引言

中国智慧教育蓝皮书（2022）强调智慧教育将构建人人皆学、处处能学、时时可学的高质量个性化终身学习体系(中国科技网, 2023)。在学习基本的数理知识外，成人的身体健康也受到国家关注，国务院颁发的《“健康中国”2030 规划纲要》，强调以人民健康为中心，促进个性化健康管理服务发展，强化面向全员的继续医学教育制度(国务院, 2016)。但第八次全国学生体质与健康调研的结果显示，我国大学生的身心健康状况不容乐观，呈逐年下降趋势(红星新闻, 2019)。近年来，期望学习医学知识以改善自身健康状况的成人学习者与日俱增，但实现成人学习者个性化终身学习医学知识的平台较少。

自适应学习作为推动个性化教育实践落地的抓手(陈昂轩等, 2023)，是指根据学习者已有信息，建立用户适应模型，实现个性化内容呈现、路径推荐和学习支持(吴南中, 2017)。与学校教育的学生不同，成人学习者的工学矛盾突出，难以有整块时间进行系统学习(杜秀芳 & 任淑红, 2007)，尤其是针对繁多且复杂的医学知识方面的学习，成人学习者不仅要对理论知识熟练于心，更需要对实际操作得心应手，在学习上需要花费更多的时间和精力。此外，成

人学习者涵盖了医学生、医生以及其他期望了解医学知识的成人等不同群体。这些群体对医学知识的掌握程度要求存在较大差异，从而使得医学知识的学习呈现出高度个性化的特点。

自适应学习系统迎合了因材施教、个性化教学的教学理念，针对性的提出适合每一个学生自身特点的学习策略、学习方案、学习路径以及学习资源，可以适应学习者多样性的学习需求(张宇新, 2021)。为了克服成人学习者学习时间、专注力、先验知识水平参差不齐等障碍，本文提出了面向成人学习者的医学知识学习系统（又称“医教平台”），旨在基于机器学习算法和数据分析技术，根据学习者的个性化学习需求、学习特征和风格等提供针对性的医学知识学习资源，并为学习者提供科学指导和反馈。

2. 自适应学习系统在成人医学知识学习中的应用现状

自适应学习系统的本质是一类支持个别化学习的学习环境，可以针对个体在学习过程中的差异性（因人、因时）提供适合个体特征的学习支持，包括个性化的学习资源、教学进程和学习策略等(李斯萌, 2014)。当前，自适应学习系统在成人医学知识学习上的应用，主要体现在辅助医学生或医生的知识学习和虚拟仿真实践。

在知识学习方面。ESP（electronic standardized patient）将复杂知识点以实时动画和互动的形式展示，让数字人扮演标准化病人的角色，创造出几种标准化病例，供学生巩固练习，从而强化学生的知识运用能力(袁艺标等, 2020)。Crystal Island 是以故事为中心的虚拟沉浸式学习环境，学习者扮演生物学家的角色，在游戏中学习，让学习者有身临其境的感觉，提高学习体验(韩建华等, 2016)。Gupta 等人通过创设虚拟病人场景，让学习者进行实践操作，有效地促进了学生关于内分泌生理学的知识学习(Gupta et al., 2018)。

在虚拟仿真实践方面。Sadeghnejad 等(2023)开发了基于 VR 的触觉虚拟训练系统，用于培训医生进行内镜鼻窦颅底手术，实验结果发现系统的性能和稳定性均较好。Whitaker 等(2019)将智能助教作为个性化病人，成人学习者与病人的交互、实践操作过程中，实现了其对专业知识的理解，提高了医疗护理的技能。临床教学中也会运用此类方法，将智能助教作为虚拟导师为学习者提供个性化的训练任务，并在训练的过程中进行针对性的互动指导(Yu et al., 2020)。如 Leng 等(2023)开发的将视觉、触觉和听觉相结合的颧骨植入手术训练系统，并通过实验表明系统可以提高外科医生对颧骨植入手术的熟练程度。

医学教育的目的是满足普通民众的保健需要(Lomis et al., 2021)。医学知识的学习系统不应该仅围绕医学生或医生等专业人士来设计和开发，普通民众也需要学习相关知识，而不是每次遇到身体问题时，在网上随意检索，检索结果却是“癌症起步”，不免引起恐慌。为满足普通民众和医学生等的学习和交流需要，本文设计了面向成人学习者的医学知识学习的自适应学习系统，即“医教平台”。

3. 医教平台的设计

3.1. 成人学习者对医教平台功能的需求

本研究首先通过线上半结构化访谈，系统梳理了 15 名受访者（其中 6 名男性，9 名女性）对医教平台功能的需求，并深入探讨他们对系统中运用数字人和严肃游戏两种形式的态度。根据访谈结果，受访者对平台功能的需求主要集中在论坛、多媒体资源和师生互动方面。

在论坛功能方面，近 60%的受访者认为论坛虽然有益，但并非主要功能。他们认为，首先，论坛作为一个自由发言的空间，需要管理员来确保内容的质量；其次，论坛内容按照知识点分类，以便学习者讨论和参与；最后，如果病人能够在论坛中描述症状，可以为学习者提供诊断练习和实践的机会，但在这一过程中专家参与必不可少，以保证诊断结果的正确性。在

多媒体资源方面，近 70%的受访者表示希望能够提供丰富的案例，通过视频呈现案例分析过程会更加生动；例题和习题也是课程学习后必不可少的；资源中可以涉及人体结构图，并通过超链接的方式连接到相应的知识点或医学权威网站。在师生互动方面，近 34%的受访者希望课程学习后能够与教师进行互动，期望教师能够为学习者解疑答惑。

关于受访者对系统中运用数字人和严肃游戏形式的态度，调查结果显示，（1）数字人：约 80%的受访者表示希望系统引入数字人，以具备搜索答疑、互动交流、语音播读和自适应提示等功能。具体而言，受访者期望数字人在答题时能够提供自适应提示功能，如根据个体停留时间的个性化调整提示;提示内容更侧重于提供解题思路而非直接给出答案等。超过 90%的受访者认为在实操练习中引入标准化病人是有益的。（2）严肃游戏：近 50%的受访者认为成年人具备足够的自控力，认为游戏功能可能有些花里胡哨，使用起来会显得繁琐且容易耽误时间。这表明部分受访者对于成年人使用严肃游戏进行学习持保留态度。

3.2. 医教平台设计的原则

3.2.1. 符合成人学习者的学习特征

诺尔斯认为成人学习者具有以解决问题为目标导向、独立自主学习、善于运用已有经验等特点(王海东, 2007)。基于成人学习者的学习特点，平台设计应首先支持个性化学习路径和内容，让学习者能够根据自身的兴趣、水平和学习需求进行自主学习，帮助学习者高效地达到学习目标。其次，成人学习者工学矛盾突出、学习时间较少，因而平台不应加入复杂且费时的游戏活动；成人学习者能够依靠其强烈的内驱力克服学习内容的专业但相对枯燥的问题。最后，平台应该充分利用学习者已有的经验，通过案例分析、问题解决等实践性学习活动，促进学习者将理论知识应用到实际问题中。

3.2.2. 强调学习的协作与意义建构

建构主义学习理论认为,知识是在一定的学习环境中,在他人的帮助下,利用现有的资源,通过有意义的建构方式获得的(黄凌梅 & 钟秉林, 2020)。协作和意义建构是学习环境中至关重要的两个方面。协作强调学生之间的合作与互动，学生在合作中能够共同探讨问题、分享观点，从而促进彼此的学习。因此，医教平台的设计应当提供多种交互方式，如在线讨论、实时问答等，促进学生与教师和学生之间的互动和合作。意义建构强调学习者将个人经验与新知识相结合，对所学知识有更深层次的理解。因此，医教平台的设计应提供丰富的学习资源，并利用数字人辅助设计实操活动，让学生学以致用，加深对知识的掌握。

3.2.3. 注重模拟实践的情境性

情境认知理论认为学习和思维的塑造受特定情境和社会环境的影响(刘革 & 吴庆麟, 2012)。因此，在医教平台的设计中，除了注重学习者之间的互动性，还应重视模拟实践活动的设计，这些活动应真实地模拟医学场景，使学生能够身临其境地进行实践，帮助学习者将抽象的理论知识转化为实际操作能力，从而更好地应对现实生活中的挑战。此外，医教平台还应当呈现学习者的学情分析，帮助他们进行反思和总结。通过模拟实践活动和个性化的学情分析,为学习者提供一个真实情境和个性化学习支持的学习环境,促进他们的学习和成长。

3.3. 医教平台的模型框架

本文参考 De Bra 将 DHRM 模型的存储层进行扩展,形成了自适应超媒体应用模型（AHAM 模型），主要由学习者模型（User Model）、领域模型（Domain Model）和呈现模型（Adaptation Model）组成(Halasz & Schwartz, 1994)。学习者模型（又称用户模型）通过对学习者个性化特征信息进行描述,反映学习者的个性化差异(姜强 & 赵蔚, 2011)。领域模型（又称内容模型），是对应用领域的各组成元素及其结构的描述，表示领域知识内部各组成元素及元素之间的相互关系(马相春等, 2017)。呈现模型（又称自适应模型）是指系统将学习者的学习内容和交互

方式进行可视化的呈现。本文在姜强等(2012)提出的“智能助教下的自适应学习系统通用参考规”基础上，参考了刘博文等(2017)的模型，并依据用户对平台功能的实际需求与平台设计的原则，构建了医教平台系统模型（见图 1）。

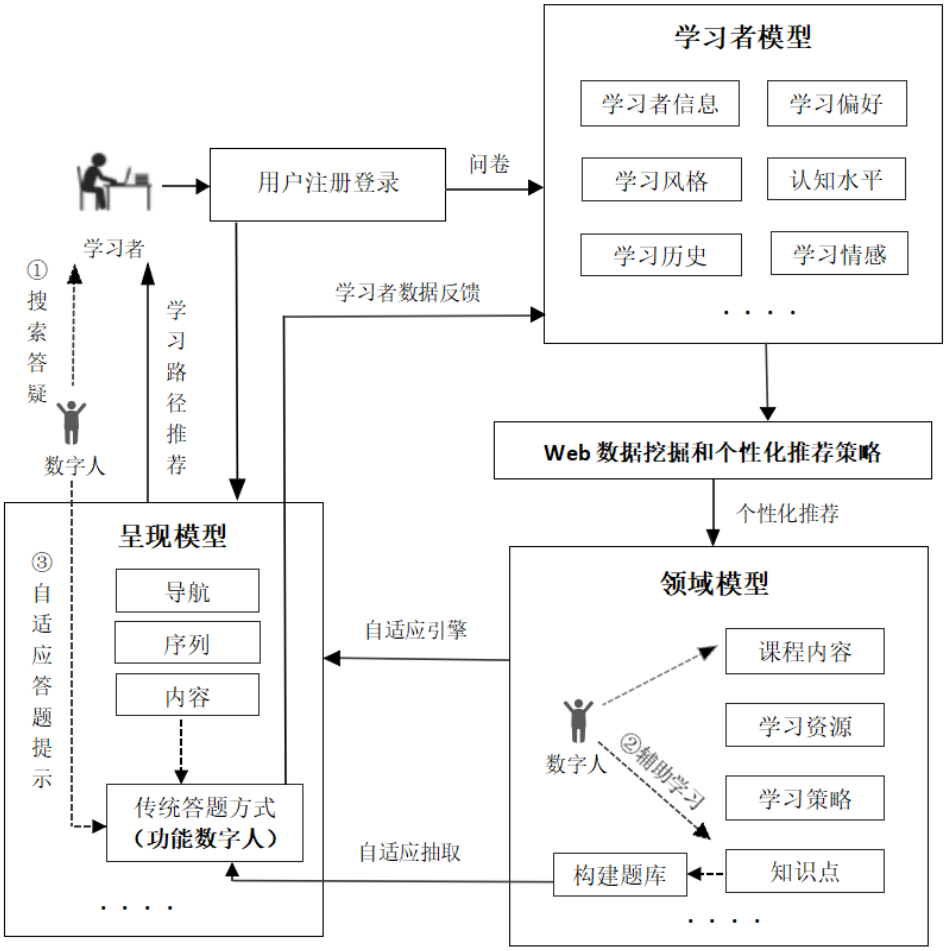


图 1 医教平台模型框架

首先，在已有的研究中，学习者模型的构建主要参考学习者的学习目的、认知风格和知识水平维度(Abu-Alsaad, 2019)。例如，在 MASPLANG 系统中，研究者根据 Felder-Silverman 模型编制了问卷，以调查学习者的知识水平和学习风格，从而建立初始学习者模型(Pena, et al., 2002)。赵学孔等(2015)从基本特征、学习风格和认知水平三个方面来构建学习者模型。鉴于已有研究成果，本文构建的学习者模型，先通过问卷调查学习者学习风格、认知水平和学习偏好，再通过数据挖掘得到学习历史、学习情感等学习者特征。

其次，在个性化推荐策略中，参考 Shahbazi 和 Byun(2022)基于学习者特征和资源内容，利用相关算法，向学习者推荐个性化的课程资源。Muangprathub 等(2020)根据学习者学习特征，在知识结构中提供自适应表示和个性化学习的自适应算法，为其检索相关知识点提供辅助支持。本文采用基于内容的推荐和基于知识的推荐，为学习者推荐个性化的课程内容、学习资源，同时数字人能够为学习者学习知识提供个性化支持。

再次，领域模型的构建必须考虑学习者模型中学习者的偏好和特征，这是自适应系统能够进行精确的个性化推荐的关系。如在一个面向 BOP 人群的在线自适应学习的模型框架中，研究者考虑到学习者从事不同的职业，如农业、渔业、木工等，他们深受社会、经济和天气变化的影响，因而在领域模型中增加对实时气候数据的采集(Irugalbandara, et al., 2019)。医教平台是为成人学习者学习医学知识而准备，在领域模型中，除了课程内容、学习资源、学习策

略外，数字人也能够根据学习者的需要辅助其学习；此外，系统根据学习者学习的知识点构建题库，为每位学习者个性化呈现测试题内容。

最后，呈现模型的目的是运用文本、图像、音频、视频等方式将领域模型中的知识和概念呈现给学习者。董晓辉等(2017)认为自适应学习环境应依据学习者的差异和学习需求，为他们提供需要的学习资源和学习导航支持。举例来说，Shahbazi 和 Byun(2022)基于学习者特征和资源内容，利用相关算法，向学习者推荐个性化的课程资源。Muangprathub 等(2020)根据学习者学习特征，在知识结构中提供自适应表示和个性化学习的自适应算法，为其检索相关知识点提供辅助支持。本文构建的呈现模型能够在呈现内容资源之外，数字人还提供搜索答疑、重难点内容提示、呈现已学习内容的思维导图等功能。此外，在学习者答题时，数字人还能够模拟成功能数字人（即标准化病人），增加学习者在模拟实践中的沉浸感。

3.4. 医教平台的主要功能

3.4.1. 自适应内容推荐

医教平台首先通过问卷了解学习者的学习风格和学习偏好，为学习者推荐相关课程和资源。再根据学习者使用系统的行为和表现，如访问频次、访问时间段、收藏与否等，为学习者提供与其知识水平和学习风格相匹配的学习内容以及多媒体资源。平台还能根据学习者的知识掌握程度、学习速度和学科兴趣变化时灵活地调整内容推荐，确保学习者始终处于一个有挑战性地学习环境中，提高学习效果，并激发学习者的学习兴趣。

3.4.2. 实时自我监控与反馈

医教平台能够实时记录学习者的课程进度、在线活动和练习成绩。学习者随时可通过“我的学习”中详尽了解在不同主题和任务上的表现，包括已完成的课程模块和学习时长等信息。此外，平台提供了具有交互性的“论坛”和专注于课程学习的“主题讨论”区，促进学习者之间、学习者与教师之间的异步交流；数字人与学习者之间的互动机制能够让学习者获得即时反馈。这种交流不仅能够帮助医学生回顾知识，还能让普通民众得到较权威的问题解答。通过综合性的学习监测和反馈机制，有助于塑造积极的学习环境。

3.4.3. 数字人辅助学习

数字人在医教平台中担任三种关键角色。其一，作为课程内容学习的辅助者，数字人提供一键检索的功能，辅助学习者更好的理解课程内容；其二、章节巩固测试的提示者，数字人能够监测学习者的答题情况，当学习者长时间未答题时，主动询问是否需要提示，提供有针对性的帮助以促进学习巩固；其三、作为学习者学习的伴学者，当学习者做模拟实践题时，数字人充当标准化病人的角色，陪伴学习者学习，为学习者提供情境化的学习体验。通过多样化的表现形式，数字人为学生提供了高效学习和巩固练习的支持，实现了去中心化、个性化的学习服务。

4. 医教平台原型实现

4.1. 医教平台原型设计

根据访谈结果与医教平台的模型框架，本研究使用 Axure RP9 软件设计了医教平台原型。医教平台原型系统主要包括“课程”、“论坛”、“我的学习”、“小医数字人”四个板块。

在“课程”板块，如图 2。登录网站后，首页呈现最新课程、热门课程的推荐；可以选择需要学习的多种课程。当点击“口腔医学”板块，会呈现口腔医学系列课程。以“口腔修复学”为例，可以看到课程简介、课程大纲、课程评价。



图 2 “课程”板块介绍

在“论坛”板块，如图 3。一方面，可以查找到不同科目的论坛专区；另一方面，学习者之间可以相互讨论。以“心血管专区”为例，在病例帖中，学习者可以对实际病例提出诊断意见并讨论，也可以自主发帖、查看各类课程的发帖回帖情况。



图 3 “论坛”板块介绍

在“我的学习”板块，如图 4。可以在此查看全部课程、正在进行的课程等；每门课程可以查看作业、成绩分析等。以“医学遗传学”课程为例，可查看课程公告、课程学习、学习测试、主题讨论、成绩分析（包括学习进度、学习习惯、互动与测试成绩）、学习笔记。



图 4 “我的学习”板块介绍

在“小医数字人”板块，如图 5。此板块具有搜索、查看人体结构图、学习总结与学习检测功能。在学习总结页面中可以查看学习时长、学习正确率等学情情况。在学习检测页面的小医数字人中，学习者可以自主选择是否“获取提示”，同时可以看到相关模拟实操题。

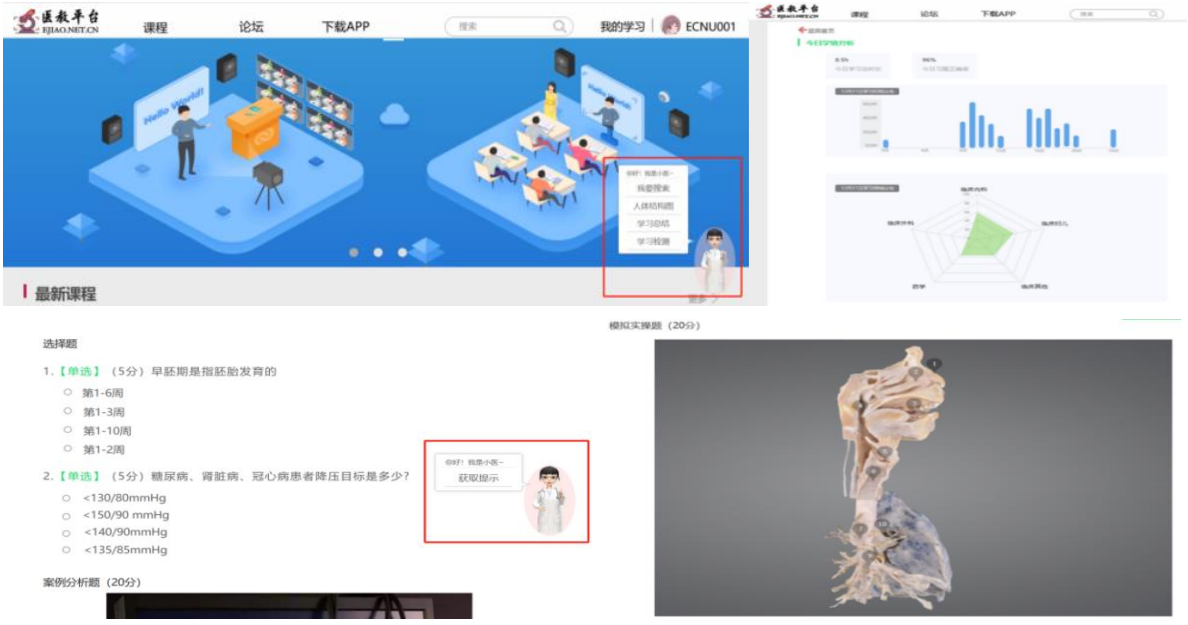


图 5 “小医数字人”板块介绍

4.2. 医教平台系统原型的用户体验分析

调研问卷是基于用户体验问卷（User Experience Questionnaire，UEQ）进行改编的，该问卷涉及评价软件的吸引力、实用性和享乐方面。实用性方面主要关注有效性、易用性和可靠性；而享乐方面则关注激励性与新颖性(Hinderks et al., 2019）。UEQ 在多个平台或系统的用户体验调查中被广泛使用(Macis et al., 2020; Maslov et al., 2021; P é rez-Rodr í guez et al., 2021)，但 UEQ 的整体问卷关于软件的具体内容较少，因此在本研究中，我们对其进行改编，增加了对软件设计、交互设计、内容资源、“论坛”、“小医数字人”和成绩分析等具体方面的评估，以便让填写者更清晰地了解题目意图。

本研究通过线上发放问卷，共收集 20 份有效问卷样本。在对数据进行计算分析后，得到系统原型体验的平均得分为：吸引力（M=3.83）、有效性（M=3.98）、易用性（M=3.92）、可靠性（M=3.73）、激励性（M=3.825）、新颖性（M=3.39）。在系统原型设计与功能体验方面，平均得分为：页面设计（M=3.95）、交互设计（M=3.85）、内容资源（M=3.925）、“论坛”（M=4.03）、“小医数字人”（M=3.64）、成绩分析（M=3.74）。通过对比用户体验问卷各维度的平均得分，我们发现系统有效性的平均得分最高，其次是易用性和激励性。对比系统具体内容方面的平均得分，系统的“论坛”板块得分最高，页面设计和内容资源次之。这表明系统的功能整合较好，功能结构组织清晰，尤其是“论坛”板块，易于学习者操作，并能有效激励学习者学习。

5. 结论

个性化自适应学习系统可以对学生的学习过程做出个别化的适应性智能指导，有效地收集、测量、分析、报告与自主学习、自主探究过程中行为有关的各种信息，从中提取隐含的、对自主学习和自主探究有指导意义的信息、知识与模式(陆玉书, 2021)。本文为医学生继续学习、普通民众学习医学知识设计的自适应学习系统（医教平台），首先通过访谈成人学习者对医学知识学习的需求，梳理出医教平台的主要功能。其次，基于用户需求与系统设计的原则，构建医教平台的系统原型。最后，在用户体验后发放系统原型的用户体验问卷（UEQ），了解成人学习者对系统的使用体验。问卷调查结果显示，系统的功能组织结构清晰，有益于学习者之间的交流，能够对学习者学习起到激励作用。但系统新颖性得分较低，表明系统在创新性和独特性方面存在不足，需要更多的创新元素来吸引用户；系统中“小医数字人”得分最低，表明数字人目前的功能未能满足学习者的学习需求，数字人需要开发更多的功能。因此，医教平台设计可以从三个方面进行改进，一是增加创新元素和新颖功能，例如在领域模型的学习资源中增加更多实例，开发与虚拟现实（VR）穿戴设备结合的实践操作场景，提高学习者的沉浸感；二是丰富“小医数字人”的功能，包括接入生成式人工智能工具以提供互动式问题解答、学习建议、个性化学习路径推荐等功能，满足学习者更广泛的学习需求；三是梳理国内外的自适应学习系统的界面设计，总结其独特之处，进一步优化本系统中“导航”和“序列”的设计。未来的研究重点将集中在医教平台的模型框架迭代和系统开发上，通过让更多的学习者体验该平台，获得的问卷与访谈结果持续修正系统模型并丰富平台功能，逐步实现医教平台在成人学习者学习医学知识方面的通用性。

参考文献

陈昂轩, 刘怀亚, 贾积有, 张君, & 徐玄冲. (2023). 双重身份学生自适应学习模型的构建与实证研究. *现代教育技术*(10), 108-117.

董晓辉, 杨晓宏, & 张学军. (2017). 自适应学习技术研究现状与展望. *电化教育研究*(02), 91-97+121.
杜秀芳 & 任淑红. (2007). 成人学习的特点分析. *中国成人教育*(15), 121-122.
国务院. (2016, 10 月 25 日). “健康中国 2030”规划纲要》. Retrieved from https://www.gov.cn/zhengce/2016-10/25/content_5124174.htm
韩建华, 姜强, & 赵蔚. (2016). 基于元认知能力发展的智能导学系统研究. *现代教育技术*(03), 107-113.
红星新闻. (2021,9 月 3 日). 第八次全国学生体质与健康调研:中小学生超重肥胖率上升,大学生整体身体素质下降. Retrieved from https://baijiahao.baidu.com/s?id=1709853987560916984&wfr=spider&for=pc.
黄凌梅 & 钟秉林. (2020). 建构主义视域下实习辅助课的启示. *中国大学教学(06)*, 64-69+81.
姜强 & 赵蔚. (2011). 面向“服务”视角的自适应学习系统设计与实现. *中国电化教育*(02), 119-124.
姜强. (2012). 自适应学习系统支持模型与实现机制研究(博士). 东北师范大学, Available from Cnki
李斯萌. (2014). 自适应学习系统设计模型相关研究(硕士). 吉林大学, Available from Cnki
刘博文. (2017). 自适应学习平台的设计与原型实现研究(硕士). 湖南师范大学, Available from Cnki.
刘革 & 吴庆麟. (2012). 情境认知理论的三大流派及争论. *上海教育科研(01)*, 37-41.
陆玉书. (2021). 基于学习技术的个性化自适应在线学习分析系统设计——以初中语文个性化自适应学习系统为例. *中国教育技术装备*(17), 29-31.
马相春, 钟绍春, & 徐妲. (2017). 大数据视角下个性化自适应学习系统支撑模型及实现机制研究. *中国电化教育*(04), 97-102.
王海东. (2007). 美国当代成人学习理论述评. *中国成人教育(01)*, 126-128.
吴南中. (2017). 自适应学习模型的构建及其实现策略. *现代教育技术*(09), 12-18.
袁艺标, 王觉进, 林茂辉, & 高兴亚. (2020). 基于功能数字人内核的虚拟仿真教学项目建设. *生理学报*(06), 730-736.
张宇新. (2021). 基于自适应学习产品的学习体验研究(硕士). 山西师范大学, Available from Cnki
赵学孔, 徐晓东, & 龙世荣. (2015). B/S 模式下自适应学习系统个性化推荐服务研究. *中国远程教育*(10), 71-78+80.
中国科技网. (2023, 2 月 13 日). 《中国智慧教育蓝皮书》：智慧教育是数字时代的教育新形态. Retrieved from http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2023/2023_zt01/mtbd/202302/t20230214_1044497.html
Abu-Alsaad, H. A. (2019, June). Agent applications in e-learning systems and current development and challenges of adaptive E-learning systems. In 2019 11th international conference on electronics, computers and artificial intelligence (ECAI), 1-6. IEEE.
Gupta, A., Singh, S., Khaliq, F., Dhaliwal, U., & Madhu, S. V. (2018). Development and validation of simulated virtual patients to impart early clinical exposure in endocrine physiology. Advances in Physiology Education, 42(1), 15-20.
Halasz, F.G., & Schwartz, M. D. (1994). The Dexter Hypertext Reference Model.Communications

of the ACM, 37(5), 30-39.

Hinderks, A., Schrepp, M., Mayo, F.J., Cuaresma, M.J., & Thomaschewski, J. (2019). Developing a UX KPI based on the user experience questionnaire. Comput. Stand. Interfaces, 65, 38-44.

Huang, Z. H., Zhang, J. W., Tian, C. Q., et al. (2016). Survey on Learning-to-Rank Based Recommendation Algorithms. Journal of Software, 27(3), 691-713.

Irugalbandara, T. C., & Fernando, M. S. D. (2019). Context aware adaptive mobile learning framework for bottom of pyramid people (BOP). International Journal of Advanced Applied Sciences, 6(12), 27-40.

Leng, A., Zeng, B., Chen, Y., Tu, P., Tao, B., & Chen, X. (2023). Development of a virtual reality-based zygomatic implant surgery training system with global collision detection and optimized finite element method model. Computer methods and programs in biomedicine, 243, 107940 .

Lomis, K.D., Mejicano, G.C., Caverzagie, K.J., Monrad, S.U., Pusic, M.V., & Hauer, K.E. (2021). The critical role of infrastructure and organizational culture in implementing competency-based education and individualized pathways in undergraduate medical education. Medical Teacher, 43, S7-S16.

Macis, S., Loi, D., Ulgheri, A., Pani, D., Solinas, G., Manna, S.L., Cestone, V., Guerri, D., & Raffo, L. (2020). Design and Usability Assessment of a Multi-Device SOA-Based Telecare Framework for the Elderly. IEEE Journal of Biomedical and Health Informatics, 24, 268-279.

Maslov, I., Nikou, S., & Hansen, P. (2021). Exploring user experience of learning management system. The International Journal of Information and Learning Technology.

Muangprathub, J., Boonjing, V., & Chamnongthai, K. (2020). Learning recommendation with formal concept analysis for intelligent tutoring system. Heliyon, 6(10), e05227.

Pena, J. L. Marzo, & J. De Rosa. (2002). Intelligent Agents in a Teaching and Learning Environment on the Web. Proceedings of the International Conference on Advanced Learning Technologies, 21–27.

Pérez-Rodríguez, R., Villalba-Mora, E., Valdés-Aragonés, M., Moral-Martos, C., Mas-Romero, M., Abizanda-Soler, P., & Rodríguez-Mañas, L. (2021). Usability, user experience and acceptance evaluation of CAPACITY: a technological ecosystem for remote follow-up of frailty (Preprint).

Sadeghnejad, S., Khadivar, F., Esfandiari, M., et al. (2023). Using an Improved Output Feedback MPC Approach for Developing a Haptic Virtual Training System. Journal of Optimization Theory and Applications, 198, 745–766.

Shahbazi, Z., & Byun, Y. C. (2022). Agent-Based Recommendation in E-Learning Environment Using Knowledge Discovery and Machine Learning Approaches. Mathematics, 10(7), 1192.

Whitaker, E., Trewhitt, E., & Veinott, E. S. (2019). Intelligent Tutoring Design Alternatives in a Serious Game. In R. Sottilare & J. Schwarz (Eds.), Adaptive Instructional Systems (pp. 745–766). Lecture Notes in Computer Science, 11597. Springer.

Yu, M., Zhang, P., & Zhang, W. (2020). Design and Implementation of Intelligent Tutor System for Training of Scalpel Operation. Microelectronics Systems Education. IOP Publishing.

不同教学阶段中教师用语特征研究——LDA 主题模型和特征词共现网络分析

A Study of Teacher Language Characteristics in Different Educational Stages:LDA Topic Modeling and Feature Word Co-occurrence Network Analysis

金宣萱[1]，杨乐莹[2]，许玮[3]
[1]浙江工业大学教育科学与技术学院
[2]浙江工业大学教育科学与技术学院
[3]浙江工业大学教育科学与技术学院
792309690@qq.com

【摘要】 在当前注重教学高质量发展的背景下，提高教师的教学水平显得尤为重要。本研究针对教师用语这一关键领域，通过分析在省、校级教师技能大赛中获奖的案例视频，采用 LDA 主题模型通过文本聚类揭示视频文本数据的深层结构和内在关联。此外，利用共现网络分析方法构建教师用语中高频关键词的共现网络图，以探究教师用语的语义关系、主题结构和关键词之间的联系。研究旨在从大规模文本数据中挖掘出有意义的信息，为教育工作者提供关于教师用语改进的具体建议。本研究不仅丰富了教师用语的理论研究，也为实践中的教学改进提供了科学依据。

【关键词】 教师用语；LDA 主题模型；共现网络分析

Abstract: *In the context of emphasizing high-quality educational development, enhancing teachers' instructional proficiency is of paramount importance. This study delves into the critical area of teacher linguistic practices. It analyzes case videos from award-winning entries in provincial and institutional teaching skill competitions, utilizing LDA topic modeling to uncover the deep structure and inherent associations within video text data. Moreover, by employing co-occurrence network analysis, the study constructs a network graph depicting the high-frequency keywords in teacher linguistic practices. This graph aids in examining the semantic relationships, thematic frameworks, and interconnectivity of keywords. Aimed at extracting meaningful insights from extensive text data, the research provides specific recommendations for refining teacher language use. This study not only enriches theoretical research in teacher linguistic practices but also offers a scientific foundation for practical improvements in teaching.*

Keywords: *Teacher language use，LDA Topic Modeling，Co-occurrence Network Analysis*

1. 引言

在当前教育领域，追求教学质量的高效发展是核心任务，其中，提升教师的教学水平显得尤为关键。教师教学能力的提升不仅关系到学生学业成就的提高，也是高校和教育机构工作的重点。本研究关注的是教师用语在教学质量提升中的作用，特别是在高校教学环境下，教师用语的有效运用对提高教学效果具有重要意义。

近年来，随着教育技术的发展，越来越多的研究开始关注如何通过科技手段来辅助教育质量的提升。在这一背景下，本研究采用了先进的文本分析技术——潜在狄利克雷分配（Latent Dirichlet Allocation, LDA）主题模型，对校级及省级教师技能大赛中获奖案例的视频文本进行深入分析。通过这一方法，我们不仅能够揭示视频文本数据中的潜在结构和关联性，而且

能在大规模文本数据中挖掘出有意义的信息。此外，通过构建共现网络图，本研究深入挖掘了教师用语的语义关系、主题结构以及关键词之间的联系，旨在为教学者在教师用语方面提供科学、有效的建议。

本研究的目的在于通过对教师用语的深度分析，为提升教师的教学水平提供理论支持和实践指导。通过结合现代文本分析技术和教育实践，希望为当前教育领域中的教学质量提升贡献新的见解和方法论。

2. 文献综述

2.1. 教师用语研究现状

当前教师用语研究面临多重挑战。由于教师教育体系中教师语言能力的培育及研究的不足，教师的语言素养普遍缺失（崔勇，2018），目前的研究主要聚焦于课堂语言的结构和功能，包括纠错、简化和解释等方面（罗书全，2006）。这些研究表明，教师与学生的互动和交流模式对学生的学习动机和绩效具有显著影响（Boden Kelly K.，2019）。然而，当前研究仍存在局限性，如缺乏对教师用语在实际教学情境中如何运用以及这种运用对教学过程影响的深入探讨。

尽管当前诸多研究通过现状分析阐述了教师用语对教育水平的影响，并提出了高质量、多样化、饱含情感的教师用语建议（吴春秀，2009），另外有研究将教师用语分为寒暄语、教学语言与评价用语，将教学阶段划分为寒暄、教学、评价三个阶段，并通过问卷调查针对不同阶段提出了建议（杨渺，2019）。但多数建议主要局限于外语课堂，未能为其他科目课程提供具体的可行性意见。

2.2. LDA 主题模型在教育领域的应用

潜在狄利克雷分配（LDA）主题模型在教育领域的应用逐渐受到重视，其在文本分析方面展现出独特的优势。通过应用 LDA 模型，研究者能够有效地识别教学内容中的主要主题和趋势，从而为教育策略的制定提供数据支持（Saura, J.R，2019）。许多研究进一步展示了 LDA 模型在互联网教育课程识别中的应用，他们利用 LDA 主题模型对课程进行层次聚类，揭示了不同抽象层次主题之间的关联，为课程设计和优化提供了有益的参考（李梦杰，2018）。

此外，有诸多研究结合了 LDA 主题建模和隐马尔可夫模型，对学习者会话行为进行实体建模，成功地将主题建模结果与学习行为、学习绩效相联系（张思，2021）。这种跨模型的整合分析不仅深化了我们对学习者行为的理解，也为教育评估和个性化学习提供了新的视角。

2.3. 共现网络在教育领域的应用

共现网络分析法，作为一种以文章关键词及其共现关系为基础的特殊知识网络构建方法，近年来在教育研究中逐渐受到关注。这种方法不仅能够在微观层面展示关键词之间的关系，还能够通过其演化过程反映科学概念和命题的增长规律。

在教育领域，刘萍和周梦欢等学者进一步将共现网络分析法应用于实际研究中。他们通过从相关文档集中抽取特征词，形成共现网络，并运用词频统计、同义词合并、通用词去除等方法，选取高频词汇作为共词网络的节点。借助社区划分的方法对词汇进行聚类，并根据点度中心度确定社区主题，实现了对主题的自动识别（刘萍，2012）。这种方法的精确度得到了验证，为教育领域的关键词分析提供了新的视角。

共现网络分析法用于识别和可视化教育研究中关键词的关联性。通过分析词汇的共现频率，这种方法揭示了教育领域的核心主题和发展趋势（吉添衡，2020）。此法不仅从微观层面展示了教育研究的关键词关系，而且通过其演化过程反映了科学概念和命题的发展规律。

2.4. 针对当前研究现状的总结

当前研究虽揭示了教师用语的特征，但在实际教学中仍显示出一定的局限性，特别是在课堂语言情境与意义的复杂性方面。现有研究提供了客观特征，但对语言运用及其对教学过程的影响缺乏深入探讨。此外，虽然 LDA 模型和共现网络分析为教育研究提供了新视角，但在实际的教学应用中仍需进一步探索。未来研究应关注不同教学情境下的教师用语，结合量化和质性研究，以全面探究教师间接语言策略，提升教育质量。

3. 研究设计

3.1. 研究问题

本研究中，LDA 主题模型与共现网络搭建主要用于针对不同阶段教师语言的主题建模与高频关键词的网络图。教师用语对于提升教学水平至关重要。本研究将着重研究不同教学阶段的主题建模结果与高频关键词的权重及词与词之间的关联性，进一步总结教师用语的使用技巧方法。具体来说，主要研究以下问题：

Q1.在不同教学阶段，教师应如何选择和利用特定的语言结构？

Q2.在整体课堂教学中，有哪些普遍适用的教师语言技巧可被推荐以提升教学效果？

3.2. 数据来源

为确保研究的时效性，本研究聚焦于 2018 年至今的比赛获奖视频，并根据本研究的目标进行了筛选，因此过滤掉需要较强表现力用语的学前教育与文化差异较大的英语外语教育，以提高研究结论的普适性。

最终，研究成功收集了共计 272 个获奖视频。收集视频中文理分布、年龄阶段分布较为均匀，这对于研究展开及研究结论普适性提供了重要保障。这些收集到的案例视频为研究提供了丰富的素材，涵盖了多个年份和省份的教师技能大赛获奖案例。这些案例将为深入研究教师的优秀实践和创新教学方法提供有力支持，并为研究提供了较高水平的数据基础。

表 1 素材收集具体分类数目

比赛名	视频总数	理科组数目	文科组数目	综合组数目	小学教育数目	中学教育数目	职业教育数目
2019 浙江省师范生教学技能竞赛	52	/	12	24	11	12	5
2023 浙江省师范生教学技能竞赛	26	4	4	12	4	8	2
2018 年湖南省师范生技能竞赛	20	8	8	/	4	/	/
2019 年湖南省师范生技能竞赛	18	8	8	/	4	14	/
2020 年湖南省师范生技能竞赛	76	31	32	/	15	/	/
山东省第六届师范生技能大赛	53	6	26	/	/	53	/
山东省第七届师范生技能大赛	27	8	19	/	12	15	/
总数	272	65	109	36	50	15	7
注：理科数目、文科数目指明确区分文理组的比赛；小学、中学、职业教育也仅针对有明确赛道区别的比赛所对应个数							

3.3. 数据处理

获取视频源后本研究利用“讯飞听见”软件进行视频转译文本，转译后由一位教育学专业学生对于转译文本抽样检测，抽取了其中任意 50 个视频的五分钟片段以正确转译字数为标准进行校验，最终正确率为 82.33%，可见该软件转译效果较好，得到的教学文本可信度高。

由于本研究聚焦不同教学阶段的教师语言使用，因此还需对教学过程进行人为划分。本研究基于赫尔巴特五步教学模式（Guoliang Wang, 2019）结合划分阶段的可分辨性，最终将教

学阶段按照如下标准：话题导入，知识学习与深化，回顾总结与复习分为三阶段。为使划分结果更可信，划分教学阶段过程由两位教学系统设计课程学业成绩较好的大学在校生进行，划分过程双方独立操作、禁止交流采用双盲独立评估，并对两者划分结果使用一致性检验方法，最终 Kappa 指数为 0.728，可信度高。

其次，对已划分好教学阶段的课程文本分别进行 LDA 主题建模，主要步骤为（1）文本预处理：通过过滤停用词、替换同义词等方法清洗无效数据；（2）分词文本向量化；（3）LDA 分析，提取主题词，并根据困惑度确定最佳主题数；（4）进一步绘制主题特征划分结果的共词网络图，挖掘不同教学阶段教师用语的内在联系与特点。

4. 数据分析

4.1. LDA 主题建模

在 LDA 主题建模中，研究通过评估不同主题数下模型的性能，使用困惑度（Perplexity）、相干性分数（Coherence Score）确定不同阶段最合适的主题数。最终确定（主题分别对应图中序号及主题权重）：（1）“话题导入”阶段四个主题：探索式学习与知识建构、文学欣赏与情感教育、视觉元素与课堂互动、生活化场景与思维激发；（2）“知识学习与深化”阶段五个主题：实验与现象分析、阅读与情感理解、逻辑与问题探究、结论与验证、观察与创造性思考；（3）“回顾总结与复习”阶段四个主题：课堂回顾与知识综合、情感交流与故事分享、问题解决与实际应用、成果总结与作业布置。

4.1.1. 课程导入阶段

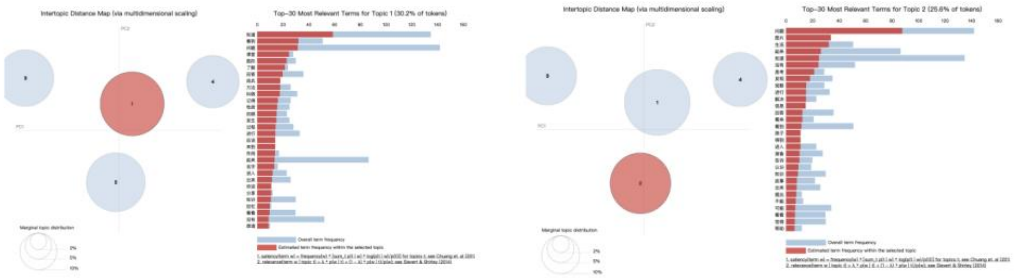


图 1 课程导入阶段主题一、主题二主题建模图

课程导入阶段，主题一“探索式学习与知识建构”在总文档中主题使用占比 30.2%，在该阶段占比最高，该主题中例如“问题”、“回忆”、“探究”等问题研究、知识回顾与建构类型词汇出现的频次较高、文本权重较大；主题二“文学欣赏与情感教育”在总文档中主题使用占比 25.6%，该主题中例如“情感”、“感受”、“故事”等情感词汇出现的频次较高、文本权重较大；

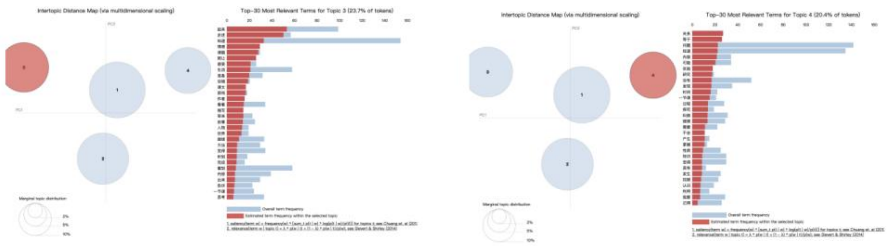


图 2 课程导入阶段主题三、主题四主题建模图

主题三“视觉元素与课堂互动”在总文档中主题使用占比 23.7%，该主题中例如“图形”、“图片”、“分享”等词汇出现的频次较高、文本权重较大；主题四“生活化场景与思维激发”在总文档中使用占比 20.4%，该主题中例如“生活”“观察”“解决”“思考”等有意义词汇出现的频次较高、文本权重较大。

4.1.2. 知识学习与深化阶段

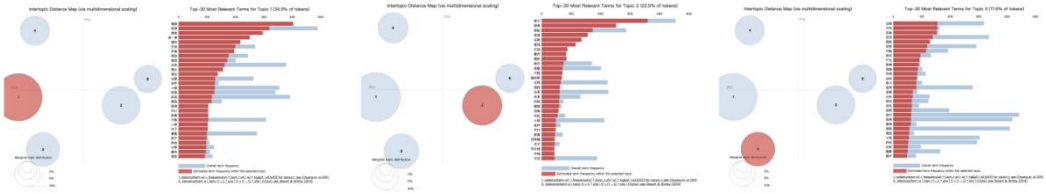


图 3 知识学习与深化阶段主题一、主题二、主题三主题建模图

知识学习与深化阶段，主题一“实验与现象分析”在总文档中主题使用占比 34.5%，该主题中例如“实验”、“现象”、“影响”、“过程”等实验总结性词汇出现的频次较高、文本权重较大；主题二“阅读与情感理解”在总文档中主题使用占比 22.5%，该主题中例如“感受”、“故事”、“描写”等情感解读类词汇出现的频次较高、文本权重较大；主题三“逻辑与问题探究”在总文档中主题使用占比 17.6%，该主题中例如“计算”、“思考”、“定义”等逻辑探究类词汇出现的频次较高、文本权重较大；

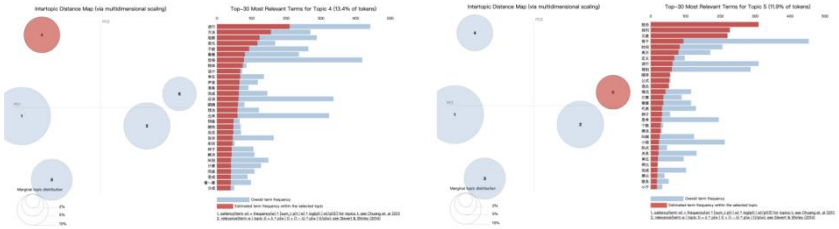


图 4 知识学习与深化阶段主题四、主题五主题建模图

主题四“结论与验证”在总文档中主题使用占比 13.4%，该主题中例如“证明”、“猜想”、“关系”等词汇出现的频次较高、文本权重较大；主题五“观察与创造性思考”在总文档中主题使用占比 11.9%，该主题中例如“观察”、“解决”、“变化”、“想法”等涉及思维转换与创造性思维词汇出现的频次较高、文本权重较大。

4.1.3. 回顾总结与复习

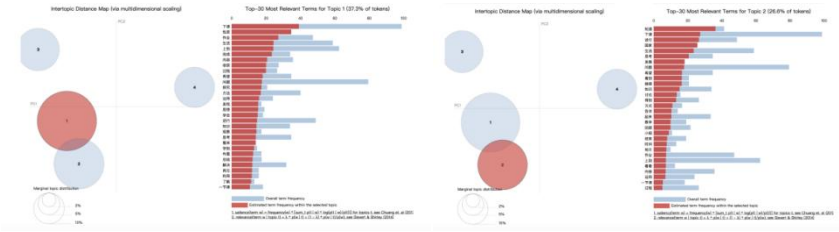


图 5 回顾总结与复习阶段主题一、主题二主题建模图

回顾总结与复习阶段，主题一“课堂回顾与知识综合”在总文档中主题使用占比 37.3%，该主题中例如“思考”、“发展”、“回顾”等回忆课程内容与知识再运用相关词汇出现的频次较高、文本权重较大；主题二“情感交流与故事分享”在总文档中主题使用占比 26.6%，该主题中例如“感受”、“分享”、“收获”等词汇出现的频次较高、文本权重较大；

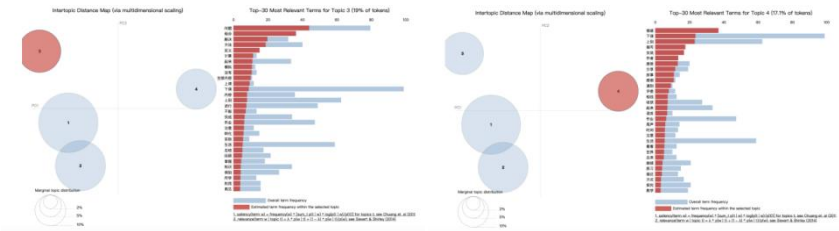


图 5 回顾总结与复习阶段主题三、主题四主题建模图

主题三“问题解决与实际应用”在总文档中主题使用占比 19%，该主题中例如“解决”、“方法”、“转化”、“实际”等问题解决与课程知识使用相关词汇出现的频次较高、文本权重较大；主题四“成果总结与作业布置”在总文档中主题使用占比 17.1%，该主题中例如“总结”、“布置”、“知识”等词汇出现的频次较高、文本权重较大。

4.2. Gephi 共现网络分析

共现网络搭建分析的过程按照话题导入，知识的学习与深化，回顾总结与复习三阶段绘制共现网络图。分别对三阶段网络图进行平均加权度与社区模块化（即进行主题分类模块）。

共现网络图中，不同节点颜色对应不同模块（即不同主题）；节点或字体越大，节点词汇在总体文本中的权重越大；边的厚度越大，两节点的平均加权度越大。例如，图 6 中话题导入阶段，该节点颜色为紫色、节点中心大、与其他词汇连接边厚度大，因此“问题”词汇为主题 1、权重大，与其他词汇平均加权度大。

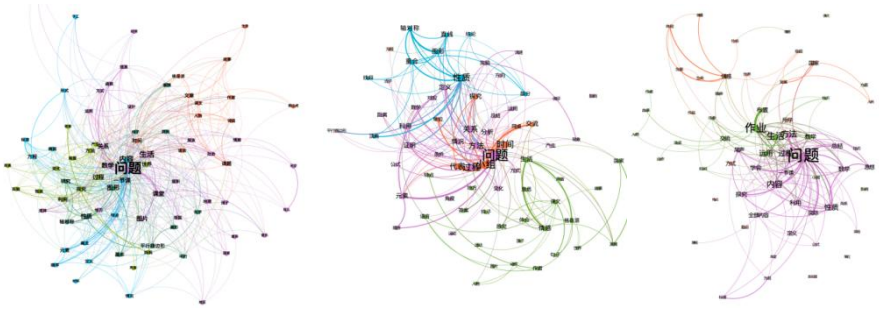


图 6 导入、知识学习、回顾三阶段共现网络图

在 Gephi 数据表格中可以进一步研究节点占文本权重与节点间的关联性。研究中，根据边的 PMI 值计算节点与节点间的关联性。共现网络中的 PMI（Pointwise Mutual Information）是一种衡量两个事项同时发生的概率与这两个事项各自独立发生概率之积的比值。在共现网络的背景下，PMI 可以用来衡量两个词或概念在文本中共同出现的程度超过了它们各自独立出现的程度多少。如果两个词经常一起出现，而且比它们单独出现的频率要高很多，那么它们的 PMI 值就会很高，这表明这两个词之间可能有强烈的关联性。PMI 特别适用于发现文本数据中的意义相关性，比如在自然语言处理和文本挖掘中发现词语间的语义关联。

表 2 不同阶段节点加权度与边的 PMI

Period.话题导入		Period.知识学习与深化		Period.回顾总结与复习		Period.话题导入			Period.知识学习与深化			Period.回顾总结与复习		
Lable	加权度	Lable	加权度	Lable	加权度	节点 1	节点 2	PMI	节点 1	节点 2	PMI	节点 1	节点 2	PMI
问题	159	问题	649	问题	244	问题	数学	8.1	小组	代表	9.1	作业	布置	8.3
生活	53	小组	433	作业	173	问题	课堂	8.1	问题	元素	8.2	问题	方法	7.7
内容	45	过程	386	生活	136	问题	关系	8.1	小组	交流	7.6	问题	内容	7.7
数学	38	性质	357	内容	127	内容	一节课	8.1	图形	轴对称	7.5	问题	性质	7.7
图形	37	时间	351	方法	113	问题	元素	8.1	小组	时间	7.4	问题	实际	7.7
课堂	33	关系	289	性质	108	数学	信息	8.1	情感	生活	7.4	分母	分数	7.7
一节课	32	代表	279	运用	106	问题	生活	7.8	问题	过程	7.3	问题	作业	7.3
性质	28	方法	273	过程	101	生活	图片	7.7	问题	小组	7.3	问题	探究	7.3
关系	28	证明	250	数学	88	问题	顺序	7.7	问题	时间	7.2	问题	学会	7.3
研究	26	生活	244	教学	82	数学	课堂	7.4	问题	定义	7.2	问题	数学	7.3
轴对称	24	分析	212	利用	78	图形	轴对称	7.3	性质	重合	7.1	生活	作业	7.3
图片	23	元素	207	探究	77	图形	基本	7.3	小组	单位	7.1	生活	运用	7.3

方程	23	定义	206	学会	75	关系	流体	7.3	生活	桃花源	7.1	问题	生活	7.0
平行四边形	22	情况	198	总结	73	方程	标准	7.2	性质	轴对称	7.1	问题	过程	7.0
基本	20	图形	198	布置	73	天空	鸟儿	7.2	过程	证明	7.1	问题	利用	7.0
标准	20	探究	197	情感	71	问题	内容	7.1	问题	关系	7.0	作业	内容	7.0
文章	19	交流	192	交流	68	问题	过程	7.1	问题	数学	7.0	作业	性质	6.7
时间	19	直线	190	尾声	64	问题	运用	7.1	性质	证明	7.0	内容	总结	6.7
元素	18	利用	188	实际	62	问题	地方	7.1	性质	顶角	7.0	性质	数学	6.7
利用	18	轴对称	186	全部内容	60	平行四边形	性质	7.1	重合	轴对称	6.9	数学	思想	6.7
课题	18	重合	181	国家	59	性质	利用	7.1	问题	性质	6.9	情感	负面	6.7
过程	17	情感	180	定义	57	问题	性质	7.0	问题	分析	6.9	情感	体验	6.7
社会	16	变化	169	一节课	56	问题	时间	7.0	问题	顺序	6.9	问题	国家	6.7
顺序	15	课文	168	方式	55	问题	研究	7.0	元素	定义	6.8	问题	运用	6.7
形式	14	运用	157	思想	52	问题	流体	7.0	元素	顺序	6.8	问题	全部内容	6.7
信息	14	对应	154	所学	52	问题	社会	7.0	性质	利用	6.8	生活	数学	6.7
椭圆	14	实验	153	小组	48	问题	函数	7.0	小组	探究	6.7	性质	利用	6.7

具体到表格数据，在话题导入阶段，“问题”、“生活”节点的加权度最高，表示这类词汇在该阶段经常被使用。而在知识学习与深化阶段，“问题”节点相较话题导入阶段的加权度大幅增加，“小组”、“过程”等节点加权度较高，表明问题的探讨和解决、小组合作是这个阶段的核心活动。在回顾总结与复习阶段，“问题”、“作业”、“生活”节点的加权度高，表明解决问题、布置作业、生活化总结为该阶段核心活动。

从 PMI 值来看，在话题导入阶段，“问题”与其他节点的边的 PMI 值位居前列，可见问题化引入课堂针对不同主题都很适用，各种课程都倾向以问题式话题引导开始课堂。例如“问题”节点与“数学”节点的 PMI 值为 8.12，表明这两个节点之间有很强的关联性。同样地，在知识学习与深化阶段，“小组”节点与“代表”节点 PMI 值为 9.06，可见两节点关联性强，即在小组讨论进行过程中倾向于代表回答的方式进行课程组织；在回顾总结与复习阶段，“作业”与“布置”节点 PMI 值为 8.31，即两节点在该阶段关联使用的程度较高。除此之外，数据也可观测到不同阶段的 PMI 值有所变化，这反映了不同教学阶段中概念之间关联性的变化。

5. 结论与建议

5.1. 各阶段用语建议

在话题导入阶段，需要重视知识回顾与建构、生活化场景问题的搭建。根据 LDA 主题建模与共现网络图，在获奖的优秀案例中教师用语多集中于探索式学习与知识构建、生活化场景与思维激发。在该阶段，例如“回忆”、“回顾”、“上节课”等词汇出现频率较高，加之知识的回顾与构建之于教学的完整性、及学生建构知识框架的重要性，因此教师应重视在这一阶段通过引导学生回顾先前知识。同时，“问题”一词与其他各类词汇的 PMI 值均较高，可见任何科目的优秀教学案例都倾向于将问题化话题引导与课程内容结合，以实现在生活化场景问题搭建的过程将学生引入到课程知识学习过程。

在知识学习与深化阶段，注重小组合作。除学科性词汇聚类显示之外，在该阶段中不论是结论与验证、观察与创造性、实验与现象分析主题中均有“小组合作”词汇，且该词的权重较大、与其他节点 PMI 值较高。因此，在知识学习与深化阶段，教师应积极倡导并支持学生之间的小组合作和有深度的讨论。这可以通过强调和鼓励学习社群的形成，利用类似于“小组”、“讨论”、“合作”等关键词汇来引导学生的学习体验。此外，教师还应运用实验设

计和案例分析等教育方法，以促进学生对所学知识的深入理解与掌握。这种教育模式有助于培养学生的协作能力，提高他们的综合思考和问题解决能力，从而更好地应对复杂的学术挑战。

在回顾复习阶段，关注知识回顾、实际应用与任务布置。该阶段，研究表明教师多使用“回顾”、“回想”、“回忆”等词汇引导学生回顾课程内容，并结合故事分享、回忆整体生活化场景的方式将课程内容再次转化至实际生活应用，以“作业布置”等词汇提醒方式引导学生完成课后任务与拓展。实际应用中，教师应着重于知识的整合和应用，使用如“回顾”、“总结”、“应用”等词汇来帮助学生现固知识。此外，布置与生活相关的作业可以帮助学生将课堂知识与实际情境相结合。

5.2. 总体教学阶段：问题贯穿始终，合理规划用语策略

不论是哪个教学阶段，由共现网络可发现，“问题”节点的加权度数值均较大。由此可见，问题化地组织课程（可以是整体的问题化场景构建，也可以根据知识内容设计不同的子问题）对于优秀课程都至关重要。因此在实际应用中，教师应思考如何结合课程内容构建学习场景，随着知识的学习逐步解决场景问题，以实现学生在现实应用中构建知识框架的总体过程。

由共现网络的节点 PMI 值可以发现，在广泛的教师用语文本中，教师总有意或无意地将一些词汇合并使用，因此在表 2 中可以发现，许多词汇的关联性较强。因此，教师在备课和授课时，应兰考虑到词汇的网络关联性，合理规划用语策略，以便更有效地构建知识框架，加强词汇间的逻辑联系，促进学生深层次的学习。

致谢

本项目受“国家级大学生创新创业大赛（项目编号：202310337027）”和“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”共同资助

参考文献

崔勇. (2018). 教师语言品质背后的人文情怀与思维方式. 教育科学论坛, (31), 59-64.

罗书全. (2006). 教师用语对学习者语言能力发展的影响. 中国成人教育, (03), 147-148.

吉添衛, 服部宏充, 江間有沙, 大澤博隆, & 神崎宣次. (2020). 多様な価値観への気づき支援. 科学技術社会論研究, 16(0), 120-133.

吴春秀. (2009). 英语课堂教师用语研究. 教学与管理, (24), 127-128.

杨渺. (2021). 教师课堂用语与大学生英语学习的关系探析. 海外英语, (22), 107-108.

李梦杰,刘建国,郭强. (2018). 基于文本挖掘的互联网教育课程主题发现与聚类研究. 上海理工大学学报, 40(3), 259-266.

张思,高倩倩,马鑫倩,魏艳涛,杨海茹. (2021). 私播课论坛中学习者会话行为建模研究. 电化教育研究, 42(11), 63-68+106.

刘萍,周梦欢. (2012). 基于共词网络的专家专长挖掘. 情报科学, 30(12), 1815-1819.

Boden, K. K., Zepeda, C. D., & Nokes Malach, T. J. (2019). Achievement goals and conceptual learning: An examination of teacher talk. Journal of Educational Psychology.

Saura, J. R., Reyes-Menendez, A., & Bennett, D. R. (2019). How to Extract Meaningful Insights from UGC: A Knowledge-Based Method Applied to Education. Applied Sciences, 9(21), 4603.

Wang, G., et al. (2019). Exploration of Programming Teaching Based on Herbart Model. IEEE.

核心素养视角下 AIGC 赋能的形成性评估模型的构建与展望

The construction and prospect of formative evaluation model empowered by AIGC from the perspective of core competencies

彭俊[1]，郝苗苗[1]
[1]澳门城市大学教育学院
[1]澳门城市大学教育学院
[*] 741575238@qq.com

【摘要】 以生成式人工智能技术为支撑的课堂评价已经成为教育变革的新动向。研究呼应国家建设技术赋能教育评价变革的时代诉求，立足核心素养创新人才培养方向，围绕人机协同深度探究学习的实践逻辑， 提出 AIGC 赋能的形成性课堂评价创新概念模型，其促进学生核心素养培养和知识创造的特征契合新时代育人质量的要求；后结合技术支持的形成性评估系统的四大功能以及四大应用策略分别论述了其应用场景和实施路径，以期为教育信息化发展和教育评价改革的提供参考。

【关键词】 生成式人工智能；核心素养；形成性评估；

Abstract: *Classroom evaluation supported by generative artificial intelligence technology has become a new trend in educational reform. The study responds to the national call for technology-empowered educational evaluation reform, focuses on the cultivation of core literacy and innovative talents, and explores the practical logic of human-machine collaborative deep learning. It proposes an innovative conceptual model of formative classroom evaluation empowered by AIGC, which meets the requirements of nurturing quality in the new era with its characteristics of promoting students' core literacy and knowledge creation; it then discusses its application scenarios and implementation paths respectively with the four major functions and strategies of technology-supported formative assessment system, aiming to provide references for the development of educational informatization and education evaluation reform.*

Keywords: AIGC, core literacy, formative assessment

1. 前言

课堂是教学过程的主阵地，课堂质量决定了教与学的效果，因此教育评估成为了课程改革的重要一环，也引起了国家的高度重视，2020 年，中共中央、国务院印发的《深化新时代教育评价改革总体方案》（简称《总体方案》）明确提出，发挥教育评价指挥棒作用，改进结果评价、强化过程评价、探索增值评价、健全综合评价(中华人民共和国教育部, 2020)。过程性评价是教学评价的重要一环，也是培养学生能力的主要根据地，北京师范大学的研究团队提出 21 世纪 5C 核心素养（简称 P21）：创新、合作、沟通、审辩思维、文化理解和传承(大卫·罗斯, 2020)，对学生发展社会的必备能力提出新的要求，强化过程性评价则可以从提升学生能力的维度出发。近年来我国不断加大人工智能政策的供给， 2023 年习近平总书记就曾指出，要深化新时代教育评价改革，构建多元主体参与、符合中国实际、具有世界水平的教育评价体系(教育部, 2023)。2022 年 ChatGPT 的出现将生成式人工智能技术（Artificial Intelligence-Generated Content, 简称 AIGC）推入大家的视野。通用人工智能技术模型在教育领域的应用，实现了个性化学习和学生伴随式数据的收集(杨宗凯等, 2022)，教育评价也此过程中不断的发生变革，由此可见，新时代具有世界水平的教育评价体系必然离不开 AIGC 的赋能，如何以

发展学生的核心素养为指挥棒，构建智能技术赋能的课堂形成性评估模型则是本研究探讨的问题。

2. 技术赋能的形成性评估助力课堂质量改革

AIGC 赋能的形成性评估是指利用数智化学习特征来评估学生的成长路径的方法，这里的数智化学习是指生成式人工智能技术支持的学习。研究表明，技术赋能的形成性评估用于课堂教学具有不同的优势。如促进学生复杂的认知过程的形成，包括增强的自我调节能力；通过提供更加即时有效和个性化的反馈帮助学生获取知识和成就感(McLaughlin & Yan, 2017)。此外，技术支持的形成性评估使学生能够提高学习的动力(Bhagat & Spector, 2017)，并支持学生不受任何时间和地点的限制的学习(Barana & Marchisio, 2016)。目前教育工作者缺乏关于如何将技术支持的形成性评估工具整合到教学过程中的能力(Zhan 等, 2021)。随着大数据挖掘技术、语义分析技术以及人工智能技术的发展，越来越多的智慧教育评价系统进入学校为课堂赋能，但是由于技术的限制，系统仍在数据反馈的智能性和及时性，评价的趣味性等方面存在局限性。生成式人工智能是基于算法、 模型、 规则自动化生成文本、 图片、 声音、 视频、 代码等多模态内容的技术 (国家互联网信息办公室, 2023)， 主要依托大语言模型（Large Language Models, LLM）进行强化训练， 获得了类似人类的自然语言理解和生成能力。AIGC 技术下的用户与技术间的常态化多模态交互使得“数字化教育”向“数智化教育”转变。传统教学模式下教学评估存在以下问题：教学资源不个性不全面、教学评估负担过重、课堂教学反馈不及时等问题，因此在传统的课堂上教师传授知识与课堂反馈占据了大部分时间。将生成性人工智能应用于形成性评估过程中，评估模型变得更为即时、智能化和人性化，可以成为学生的智能学习小助手，使教师更聚焦于探究从哪些方面培养学生的核心素养能力(黄荣怀, 2023)。

综上所述，技术赋能的形成性评估系统在教学中的应用效果是正向的，如何通过更加精细化的系统设计以及教学策略设计强化教师运用评估系统的能力成为了重要的研究方向。通过对主题词的知识图谱可视化分析发现，关于 AIGC 赋能形成性评估系统的研究目前仍处于展望、设想等初期阶段，因此结合生成性人工智能的研究热点建构形成性评估模型的研究想法为本文之创新点。

在 AIGC 赋能新的学习范式中，形成性评估系统以“数智人”的身份存在在课堂中，数智人、教师和学生构成课堂教学三大关键元素，与人机协同深度探究理论共同构成“人人共生”智能化课堂生态。教师将更多时间投入至探究式课堂的设计与组织中，评估系统在技术支持下以更加人性化的形式与学生共处，教师成为“人人共生”课堂的组织者、管理者和建构者，学生以“主人翁”的角色出现在“人人共生”环境中，“智能人”成为知识的传授者，学生成为知识的创造者。

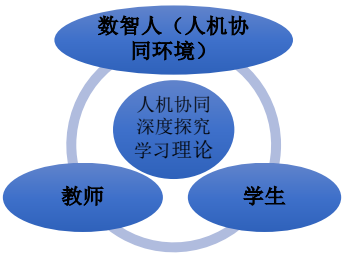


图 1“人人共生”智能化课堂生态

为保证“人人共生”环境中高效学习的发生，人机协同深度探究性学习理论被运用，其基于人机协同学习理论、社会主义建构理论和涟漪拓展探究法建构而来(李海峰 & 王炜, 2023)，该理论强调 1.指向人机协同素养的发展目标；2.智能系统的生成性；3.驱动人机协同的认知冲突；4.人机协同环境的有效支持。智能形成性评估系统在人机协同深度探究性学习理论的支持下被有效设计，从而达到提升学生的核心素养，为教学增效减负的目的。

3. AIGC 赋能的形成性评估系统建构

3.1 核心素养诊断模型构建

核心素养诊断模型依据 5C 核心素养对应的 16 个子要素的描述以及行为表现进行训练。(甘秋玲等, 2020; 康翠萍等, 2020; 刘妍等, 2020; 马利红等, 2020; 徐冠兴等, 2020)P21 提出需要进行持续的形成性评估来培养这些技能，这在其他一些文献中也得到了强调(Bransford 等, 2000; Johnson 等, 2016; Woolf, 2010) ，因此学生的核心素养能力需在形成性评估过程中不断得到培养和提升。核心素养的诊断系统通过教师在课堂上布置的情景任务，采集学生行为数据，经过数据清洗，基于 5C 核心素养行为指标，进行人机协同的机器模型训练，用到如 KNN、决策树等方法，达到深度的机器学习，有效剔除冗余数据，优化系统的数据分析的能力，建构大数据模型，最后形成核心素养诊断模型可以对课堂上学生行为进行预测、诊断与反馈。

3.2. AIGC 赋能的形成性评估系统理论基础及功能

Ecclestone 认为技术支持的形成性评价系统应该包含 5 个要素来保证学生的个性化学习以及教师的精准教学(Ecclestone, 2010)，本研究基于 5 个要素以及生成式人工智能技术的特征总结了 7 条适用于基础教育阶段形成性评估模型需要遵循的要素：1.能够预估学习者学习的起点以及当前的知识水平；2.针对学习数据和目标，能够分析出学习者个人画像，优势和薄弱点一目了然；3.能够生成学生的个性化学习路径；4.支持自我评估和同伴评估；5.根据学生的个性化学习数据可以生成下阶段的学习目标；6.根据个性化学习数据生成个性化学习建议以及资源。7.反馈的及时性和智能性。

生成式人工智能技术（AIGC）的出现使得学生的学习更加个性化和沉浸式，形成性评价可以从原来的机器式伴随转变为资源生成式的伴随，本研究依据以上 7 个功能将 AIGC 赋能的形成性评估系统分为四个子系统：发送/展示系统、收集/分析系统、生成/反馈系统，自评/互评系统。为更好保证形成性评估系统的教学应用效果，Wiliam 和 Thompson 提出的五种形成性评估策略(Wiliam & Thompson, 2017)在过往研究设计形成性评估模型中被广泛使用。基于此以及 AIGC 技术的特征总结以下 4 条形成性评估系统应用策略：1.明确个性化的学习路径与学习标准；2.设计诱发认知冲突的课堂活动与学习任务；3.提供有助于学生进步的反馈；4.让学生成为彼此的学习资源和自己的主人。

AIGC 赋能的形成性评估系统主要是指教育评价在生成式人工智能大模型的支撑下，以人机协同深度探究性学习理论为系统在教育场景应用的理论基础，以学生核心素养诊断系统评估为评估系统功能的指挥棒，以形成性评估四大策略为系统课堂应用的实践指南，以生成式

人工智能技术、大数据技术、语义分析技术等作为技术支撑，构建的发送/展示、收集/分析、生成/反馈，自评/互评四大系统组成的形成性评估系统，如下图所示。

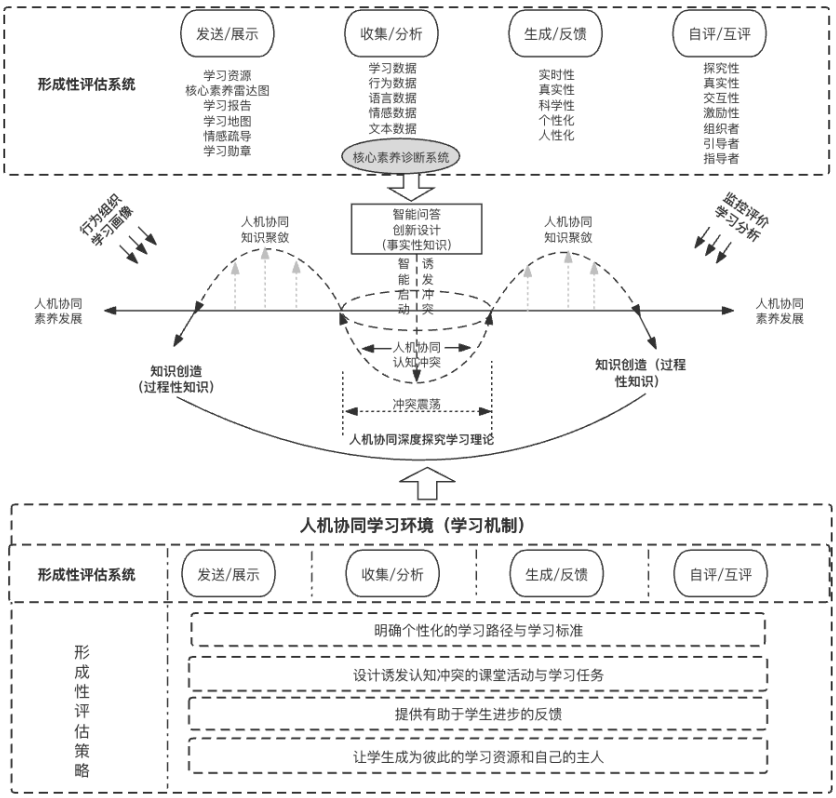


图 2 AIGC 赋能的形成性评估系统

下面结合形成评估的四大策略对四大系统的功能进行简要描述：

3.1.1.发送/展示系统

如模型图所示,此系统包含发送学习资源、生成核心素养可视化雷达图、个性化学习报告、学习地图、学习勋章以及情感陪疏导的功能。向学生发送数字化的学习资源，以保证学生在建构课堂知识的时候拥有足够多的背景知识，同时学生在对数字化学习资源学习过程中，通过不同的渠道和方式建立学生、同伴、教师以及数智人之间的沟通数据，以此促进学生记忆的形成和知识的建构，同时帮助数智人获取足够多的行为数据，给出精准的反馈，数智人通过数据分析，生成每个学生各个阶段的核心素养数据雷达图。在发送展示系统中，数智人在每位同学完成学习资源的学习之后告知学生的学习进度以及学习效果。并告知本节课的学习路径、每一步的目标和评价标准。课堂活动和学习任务、系统反馈、自评和同伴互评的结果须通过该版块传达给学生。

3.1.2.收集/分析系统

此系统收集和分析学生的学习数据、行为数据、语言数据、情感数据、文本数据等，是形成性评估系统的核心系统。通过收集整个教学过程中的数据，使用人工智能、大数据和深度分析等技术对这些学习和教学过程进行分析,是其他三个分支系统的技术基础。使用此功能，

学生和教师可以充分了解他们的学习和教学情况。此功能有助于选择性地提取或总结相关数据，并通过基于学习目标处理或分析学习者的表现，用数据支持形成性评估的结果。核心素养诊断系统也将基于收集的数据，分析学生的核心素养能力，同时在对应学习环节中给出反馈。

3.1.3.生成/反馈系统

该系统强调反馈的智能性、及时性以及情感陪伴性，因此需要系统具备技术支持下的实时性、真实性、科学性和人性化，是保证评估有效性的关键系统。数智人是该系统的主人公，通过收集/分析每位同学的行为数据，数智人已经建构出每位学生的个人学习画像以及核心素养画像，利用数智人的生成性技术，提供给每位学生关于核心素养提升的激励策略以及资源，以满足他们的个性化学习路径和个性化需求；同时数智人可以做到反馈的智能性，与传统的纯数据反馈不同，该系统可以根据学生的路径自动生成学习资源，同时在每一个路径中提供情感维护，激励、陪伴和与学生交互。

3.1.4.自评/互评系统

该系统主要为了营造一个由数智人、学生、同伴和教师构成的充满活力的社群空间，数智人在这个社群空间充当主持人的角色，而教师充当指导员的角色，学生和同伴是社群的主要活动者。数智人充当组织者、引导者和指导者，要求其在技术支持下具备探究性、交互性、真实性、激励性。该版块可以帮助学生更好的通过智能工具获得过程性知识，也成为形成性评估的主要组成部分，自我评估帮助学生发展了自我认知和自律的技能（自主发展），同伴评估帮助学生发展了批判质疑和理性思维、其中问题的解决，技术的运动，以及人文底蕴可以通过教学活动的设计融合相应的元素(Bransford 等, 2000)。在数智人的主持环境中，学生和同伴可以完成知识的学习、自评和互评等活动。通过此功能也锻炼了学生的批判性思维以及高阶思维。

4. 核心素养视角下 AIGC 赋能的形成性评估系统课堂实施路径和展望

AIGC 赋能的形成性评估系统是一个适用性极强的教学评估工具，在人机协同深度探究性学习理论的支持下，它主要可以在情景探究式课堂、任务驱动式课堂以及游戏化教学课堂等教学情景中发挥提升教学质量的作用。原因是诸如此类课堂，教师有更多空间组织和牵引学生和数智人之间发生认知冲突，从而产生知识的聚敛和创造。正如前文所述，教师应用智能形成性评估系统的能力也是影响教学质量改革的重要原因，形成性评估策略为 AIGC 赋能的评估系统在课堂中的实施路径指明了方向。

4.1．明确个性化的学习路径与学习标准

个性化的学习路径是学生自主学习质量的重要条件，为了保障学生的学习质量，评估系统需在课前、课中、课后整个教学流程根据学生的行为数据，实时或者阶段性发送个性化的学习路径以及学习的标准，并通过评估反馈的形式确保学生对于学习标准的认知水平。从教师角度出发，该策略可以保证课前教师清晰了解学生的认知水平，触发教师的精准备课；课中，教师可根据不同学生的人物画像，推送个性化学习资源和智能问答；课后，教师亦可根据学

生画像以及核心素养画像，实现精准辅导。对于数智人而言，可根据不同的教学场景触发学生不同的学习路径和学习策略，从而提供针对性的技术辅助，如情感分析技术可通过学生与数智人的对话内容分析其情感变化，调节学习氛围；学习分析技术可以通过监控学生与数智人之间的会话关键词、会话频次、会话内容结合核心素养模型进行主动引导和深度探究，构成每个人不同的学习路径，从而保证学生核心素养培养的全面性。

4.2．设计诱发认知冲突的课堂活动与学习任务

有效的课堂活动以及学习任务设计是保障课堂质量的关键，在“人人共生”的学习环境中，教师的角色发生了转移，课堂活动与学习任务不再成为教师和学生之间的互动形式，而是学生和数智人之间产生认知冲突的桥梁，因此有效教学设计成为教师的重要任务。认清数智人的角色定位以及功能对于教学设计来说十分关键，根据人机协同深度探究学习理论，学生的过程性知识需在智能问答功能的基础上产生持续的认知冲突震荡(李海峰 & 王炜, 2023)，数智人需要具备在教师设计的课堂活动中充分发挥主动参与、解答疑问、提出质疑以及生成策略的能力。而在学习任务中，需要根据知识图谱的底层逻辑设计，与学生产生以知识梯度为核心的任务交互，同时辅之以情感交互等学习支持服务。因此，教师需结合数智人功能设计有效的课堂活动，同时需在课堂活动中充当脚手架的作用，线下充当学生知识和技术的答疑解惑辅导者，线上根据学生行为数据触发评估系统不同的功能以及提供个性化学习资源。

4.3．提供有助于学生进步的反馈

作为伴随式评估的一种方式，形成性评估最重要的一点是反馈的及时性，在生成式人工智能技术的支持下，反馈又可以变得智能化和个性化。教师可以在课前预习活动，课中学习活动以及课后复习活动中通过线上或者线下的方式给予学生反馈。人机协同深度探究性教学模式能显著提高学习者的批判性思维能力(Chang 等, 2022; Guo 等, 2023)，数智人可以通过数据呈现、疑问解答、鼓励参与、及时反馈等形式引导学生明确问题、分析问题、形成解决问题的方案，实现人人共生环境中的人机协同深度探究、问题解决和知识创生；教师可利用距离优势，及时关注课堂上学生的状态，通过语言互动、及时指导的形式给学生提供正向的反馈。

4.4．让学生成为彼此的学习资源和自己的主人

在学生、教师和数智人共建的“人人共生”课堂中，学生是课堂的主人，也是知识的创造者，形成性评估可通过系统反馈和教师反馈实现，也可以在学生相互评价之中产生。同伴评估是个人或群体对同伴的学习成果进行评价的过程。这种评估可以采取授予分数的形式，也可以采取书面或口头评论的形式，但是在许多情况下，通常是采取这两者的组合的方式呈现，同伴互评除了减少教师的工作量外，也被认为为学生学习带来许多潜在的好处，包括增加学生的归属感和自主性(Rust, 1995)，增加学习动机(Bostock, 2000; Vu & Dall’Alba, 2007)，增强了社会存在感(Strijbos & Sluijsmans, 2010; Topping 等, 2000)，以及促进高阶思维和元认知技能的发展(Lydia Wen 等, 2006; Mok, 2011)。在技术融合课堂的环境中，同伴互评以及自我评估强调了学生的主体地位，弱化了技术进入课堂的认知冲突，有助于打造人机和谐共生的学习环境。因此在不同的教学场景中，数智人可以通过打造论坛空间、问答互动、在线评价等形式实现学生对于自己或者他人的评价，共建和共创知识，成为掌控自己学习路径的主人。

5. 结语

以 ChatGPT 为代表的生成式人工智能技术的出现，将教育推向了百年未有的大变局。技术赋能教育又将面临新的高度，形成性评估作为课堂教育应用的主阵地，势必需要扛起技术变革的大旗，为教育创新添砖加瓦。教育评价改革需要创新人才培养理念-核心素养作为支撑，智能化测评通过引入生成式人工智能新技术，以更可靠、更高效、更智能的手段整合多维度、多层次的信息，为教师提供更加精准有效的教学支撑，为课堂提供“人人共生”的和谐生态。随着生成式人工智能和教育教学的深度融合，形成性评估既吸取传统测评的优势，提高评估的精细化和针对性，又结合技术革新的力量，更加智能化和生成式，并结合形成性评估策略，增加系统的实用性、适用性和易用性，为教师增效减负、为课堂学习注入乐趣和活力，为教育质量的提升增添新的生命力。

参考文献

中文部分

中华人民共和国教育部 （2020）。深化新时代教育评价改革总体方案。 取读于 2024 年 1 月 13 日，从 http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202010/t20201013_494381.html

刘妍、马晓英、刘坚、魏锐、马利红、徐冠兴、… 甘秋玲 （2020）。文化理解与传承素养:21 世纪核心素养 5C 模型之一。华东师范大学学报(教育科学版), (2), 29–44。

国家互联网信息办公室 （2023）。生成式人工智能服务管理办法（征求意见稿）。 取读于 2024 年 1 月 13 日，从 http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm

大卫·罗斯 （2020）。致辞:从《4C》到《5C》——祝贺《21 世纪核心素养 5C 模型》发布。华东师范大学学报(教育科学版), (2), 19。

康翠萍、徐冠兴、魏锐、刘坚、郑琰、刘妍、… 马利红 （2020）。沟通素养:21 世纪核心素养 5C 模型之四。华东师范大学学报(教育科学版), (2), 71–82。

徐冠兴、魏锐、刘坚、李静懿、康翠萍、马利红、… 刘妍 （2020）。合作素养:21 世纪核心素养 5C 模型之五。华东师范大学学报(教育科学版), (2), 83–96。

教育部 （2023）。习近平在中共中央政治局第五次集体学习时强调 加快建设教育强国 为中华民族伟大复兴提供有力支撑。 取读于 2024 年 1 月 13 日，从 http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/202305/t20230529_1061907.html

李海峰和王炜 （2023）。人机协同深度探究性教学模式——以基于 ChatGPT 和 QQ 开发的人机协同探究性学习系统为例。开放教育研究, (6), 69–81。

杨宗凯、王俊、吴砥和王美倩 （2022）。發展智能教育學 推動教育可持續發展。電化教育研究, (12), 5-10,17。

甘秋玲、白新文、刘坚、魏锐、马利红、徐冠兴、… 康翠萍 （2020）。创新素养:21 世纪核心素养 5C 模型之三。华东师范大学学报(教育科学版), (2), 57–70。

马利红、魏锐、刘坚、马鸣燕、刘妍、甘秋玲、… 徐冠兴 （2020）。审辨思维:21 世纪核心素养 5C 模型之二。华东师范大学学报(教育科学版), (2), 45–56。

黄荣怀 （2023）。人工智能正加速教育变革：现实挑战与应对举措。中国教育学刊, (06), 26–33。

英文部分

Bhagat, K. K., & Spector, J. M. (2017). Formative Assessment in Complex Problem-Solving Domains: The Emerging Role of Assessment Technologies. Journal of Educational Tech nology & Society, (4), 312–317.

Bostock, S. (2000). Student peer assessment. Keele, Staffordshire: Centre for Learning Techn ology, Keele University.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). How people learn (Vol. 11). Was hington, DC: National academy press.

Chang, C.-Y., Kuo, S.-Y., & Hwang, G.-H. (2022). Chatbot-facilitated nursing education. Ed ucational Technology & Society, (1), 15–27.

Ecclestone, K. (2010). EBOOK: Transforming Formative Assessment in Lifelong Learning. McGraw-Hill Education (UK).

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class de bates on students' argumentation skills and task motivation. Computers & Education, 104 862.

Johnson, L., Becker, S. A., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). N MC horizon report: 2016 higher education edition. The New Media Consortium.

Lydia Wen, M., Tsai, C., & Chang, C. (2006). Attitudes towards peer assessment: A compa rison of the perspectives of pre-service and in-service teachers. Innovations in Education and Teaching International, (1), 83–92.

McLaughlin, T., & Yan, Z. (2017). Diverse delivery methods and strong psychological benef its: A review of online formative assessment. Journal of Computer Assisted Learning, (6), 562–574.

Mok, J. (2011). A case study of students' perceptions of peer assessment in Hong Kong. E LT Journal, (3), 230–239.

Rust, S. B., Phil Race, Chris. (1995). Using and Experiencing Assessment. In Assessment f or Learning in Higher Education. Routledge.

Strijbos, J.-W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functi onal, and conceptual developments. Elsevier.

Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative Peer Assessment of Academic Writing Between Postgraduate Students. Assessment & Evaluation in Highe r Education, (2), 149–169.

Vu, T. T., & Dall'Alba, G. (2007). Students' experience of peer assessment in a professiona l course. Assessment & Evaluation in Higher Education, (5), 541–556.

Wiliam, D., & Thompson, M. (2017). Integrating assessment with learning: What will it tak e to make it work? In The future of assessment (pp. 53–82). Routledge.

Woolf, B. P. (2010). A roadmap for education technology. Washington, DC: National Scienc e Foundation.

Zhan, Y., Sun, D., Chan, N. C., Chan, K. W., Lam, T. S., & Lee, T. H. (2021). Enhancing Learning Engagement Through Formative E-Assessment in General Education Foundatio n Course Tutorials. In C. P. Lim & C. R. Graham (Eds.), Blended Learning for Inclusiv e and Quality Higher Education in Asia (pp. 281–299). Singapore: Springer Nature.

Barana, A., & Marchisio, M. (2016). Ten good reasons to adopt an automated formative ass essment model for learning and teaching Mathematics and scientific disciplines. Procedia-Social and Behavioral Sciences, 608–613.

基于TPB的大学生专业认同感模型构建和检验

Construction and Testing of a Model of College Students' Professional Identity Based on TPB

高欣羽 余俊 许玮

浙江工业大学

1291723049@qq.com

【摘要】 深化对大学生专业认同感的研究将进一步推动“高质量教育体系”的建设。本研究基于计划行为理论（TPB）构建了六维度的专业认同影响机制模型，并对浙江高校445名大学生的问卷数据进行了分析。结果表明，专业环境感知对行为倾向具有显著的正向影响，专业角色对行为倾向有显著的负向影响。行为倾向在六维度的专业认同影响机制中起到部分中介作用。由此为学校和教育部门调整教育工作及相关政策提供了切实的参考。

【关键词】专业认同感；计划行为理论；结构方程模型

Abstract: *Based on the Theory of Planned Behavior,this study built a six-dimensional model of professional identity influence and analyzed data from 445 students in Zhejiang universities. Findings show that perceived professional environment positively influences behavioral tendency, while professional role has a negative impact. Behavioral tendency partially mediates the six-dimensional professional identity influence. These results offer practical insights for schools and educational authorities to adjust educational practices and policies effectively.*

Keywords: *Professional Recognition,Theory of Planned Behavior,SEM*

1. 引言

在“课程思政”背景下，高等学校教育对于学习者情感态度价值观层面的培养也逐步提升[1]，这表明学习者专业认同等非认知层面的因素非常值得关注。学生进入大学，从应试化教育转为更开放的教育模式，而专业认同感是决定其能否适应新的教学模式的重要因素。本研究在计划行为理论模型(TPB)的创新基础上，编写专业认同感影响因素问卷，使用结构方程模型验证理论模型，为推进政府教育部门对高等教育工作的调整、提升教育政策的实施效率提供可行的参考依据。

2. 研究综述和理论基础

2.1. 专业认同感

心理学家弗洛伊德最早提出来“认同”一词，他认为认同是个人与他人、群体或模仿人物在感情上、心理上趋同的过程。专业认同这一概念由“认同”理念延伸开来，现有的研究为我们提供了大学生专业认同感的定义和多维度衡量的理论指标，李彦花[2]将教师专业认同分成六个因素：专业自我、专业角色、专业知

本项目受“浙江工业大学2024年院级运河杯科技项目”和“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”共同资助

能、专业态度和动机、专业行为倾向和专业环境认同，本研究正是基于此，与TPB结合为待验证的影响因素模型。并采用结构方程模型对问卷数据进行分析，不断优化专业认同感CFA模型，得到一个较为完善且走在前沿的大学生专业认同感影响机制模型，为学校和教育部门调整工作及相关教育政策的制定提供切实的参考。

2.2. 行为主义理论

计划行为理论(TPB)是由学者Ajzen(1985)根据多维属性态度理论与理性行为理论所提出的[3]，是社会心理学中最著名的态度行为关系理论。该理论认为行为意向是影响行为最直接的因素，行为意向反过来受态度、主观规范和知觉行为控制的影响，并被证实能提高研究对行为的解释力和预测力。[4]“计划行为理论”的基础理论模型如图1所示。

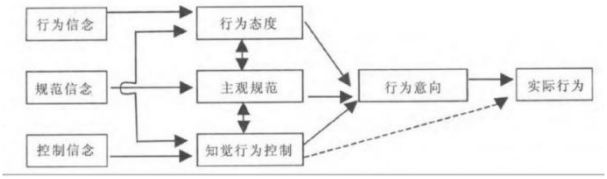


图1 “计划行为理论”基础理论模型

2.3. 专业认同影响机制假设模型

本研究将李彦花验证的态度动机分解为专业态度和专业动机两个维度，基于“专业自我、专业角色、专业态度、专业动机、专业环境感知、专业知能感知”六个维度对专业认同的影响机制模型，考察行为倾向的中介效应，具体结合李彦花的研究成果，基于六维度要素做出以下假设：

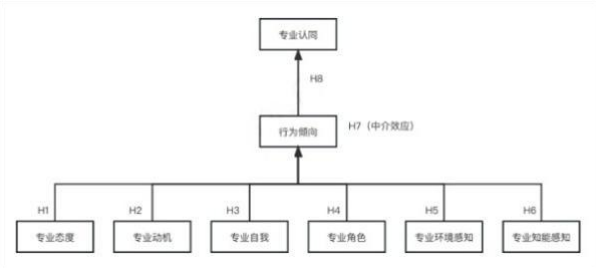


图 2 专业认同感影响机制假设模型

3.研究方法

3.1. 研究工具

研究工具包含有 1 个量表，各维度均采用李克特五点计分设计。涉及专业自我（5 题项）、专业角色（4 题项）、专业环境感知（3 题项）、专业知能感知（5 题项）、专业态度（3 题项）、专业动机（2 题项）、专业行为倾向（4 题项）及专业认同感（3 题项）共 8 个维度来验证理论模型。共回收有效问卷 445 份。采用SPSS 27.0+Amos27 进行数据分析，通过结构方程模型检验模型拟合度。

3.2. 研究对象

本研究以浙江省数所高校的在读学生中进行广发放，其中有效问卷 445 份，有效回收率为 95.9%。有效被试中，男、女学生所占的比重分别为 73.03%、26.97%，其中 81.60%的学生所处年龄段在 22 岁及以下，且 90%以上的学生都为在读本科生。通过以上样本特征描述可以看出，本文涉及的被试对象具有较强代表性。

本项目受“浙江工业大学2024年院级运河杯科技项目”和“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”共同资助

4.数据分析和结果

4.1. 问卷信度分析

在本次研究中，专业认同感影响因素均通过量表的形式进行测量，因此对测量结果的数据质量进行检验是保证后续分析具有意义的重要前提。本次分析中，“专业认同感影响因素问卷量表”的总体及各个维度题项的信度系数均在0.839-1的范围内，说明本次研究所使用的量表具有很好的内部一致性，各维度的题目具有较好的可靠性。

4.2. 量表效度分析

4.2.1. 验证性因子分析（CFA模型适配度检验）

根据模型适配检验结果可以看出，CMIN/DF（卡方自由度比）=3.203，在1-5的范围内，RMSEA（误差均方根）=0.070，在<0.08的良好范围内。另外的IFI、TLI以及CFI的检验结果均达到了0.9以上的优秀水平。因此，综合本次的研究结果可以说明，专业认同感CFA模型具有良好的适配度。

4.2.2. 收敛效度与组合信度分析

在专业认同感量表CFA模型具有良好适配度的前提条件下，将进一步检验量表各个维度的收敛效度（AVE）和组合信度（CR）。根据分析结果可以看出，在本次专业认同感量表效度检验中，各个维度的AVE值达到了0.5以上，CR值达到了0.7以上，综合可以说明各个维度均具有良好的收敛效度和组合信度。

4.3.结构方程模型

4.3.1. 专业认同感影响因素SEM模型适配度检验

在对变量的拟合指标进行分析可知，本研究初始模型的CMIN/DF值为6.658，其余适配指标RMSEA的值为0.113，NFI的值为0.828，RFI的值为0.805，CFI的值为0.823，其中CMIN/DF和RMSEA不合格，其他的指标处于可以接受的范围。因此可以对模型进行MI修正再次测试模型关系。

4.3.2 对专业认同感影响因素SEM模型进行修正后适配度检验

利用AMOS软件计算出本次测量模型模型修正指标（modification indices），针对MI值相对较大的变量对其进行连接。

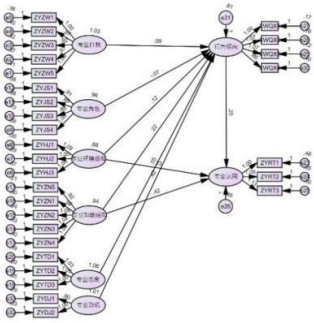


图5 修正前的模型路径关系图

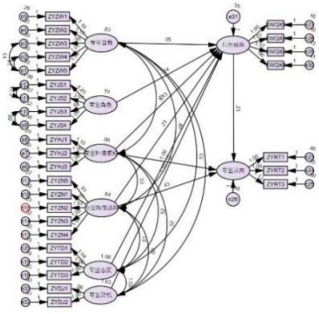


图6 修正后的模型路径关系图

在对变量的拟合指标进行分析可知，本研究修正后的模型的CMIN/DF值为2.744，RMSEA的值为0.063，NFI的值为0.948，RFI的值为0.940，CFI的值为0.948，各项指标均处于良好或者优秀的水准。因此修正后的模型关系是合理的。

4.3.3 专业认同感影响因素SEM模型路径假设检验结果

根据分析结果可以看出，专业自我对行为倾向的影响系数为0.04（Z=0.538，p＞0.05），因此专业自我对行为倾向不具有显著性的正向影响。同理，专业角色对行为倾向具有显著性的反向影响。专业环境感知对行为倾向不具有显著性的

本项目受“浙江工业大学2024年院级运河杯科技项目”和“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”共同资助

正向影响。专业知能感知对行为倾向具有显著性的正向影响。专业动机对行为倾向不具有显著性的正向影响。专业态度对行为倾向不具有显著性的反向影响。专业环境感知对专业认同具有显著性的正向影响。专业知能感知对专业认同具有显著性的正向影响。行为倾向对专业认同具有显著性的正向影响。

4.3.4 中介效应分析

调查行为倾向在专业态度、专业动机、专业自我、专业角色、专业环境和专业知能对专业认同的中介效应作用中，所有维度都符合要求（$p<0.05$）。行为倾向在专业自我对专业认同的影响中Bootstrap95%CI不包含0([0.110,0.253])，因此行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为31%。同理得出，行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为42%。行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为32%。行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为25%。行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为21%。行为倾向在专业自我对专业认同的影响中呈现部分中介作用，中介效应占总效应的比例为26%。

5.结论与讨论

5.1. 知觉行为层面对专业认同的影响：专业环境感知的重要性

从数据可知，专业环境感知对行为倾向具有显著性的正向影响，专业角色对行为倾向有显著反向影响；行为倾向在六维度对专业认同的影响机制中起部分中介作用。知觉与情感层面的要素对行为要素具有预测影响，对于专业认同的发展具有重要作用。从这个角度，如何分阶段给专业学习者构建合适的专业环境，为其产生正向专业认同认知具有重要作用；同时，合理的专业学习资源的投入与分配，对于学习者提升个人的知能认知，优化专业认可度也有重要的正向预测作用。

5.2. 大学生专业认同感的影响要素间存在相互影响：中介效应的作用

在对模型进行修订后可以发现，各要素之间适当的两两关联对于模型拟合度的提升有明显的优化作用，这是因为专业认同的形成需要经过个体与他人、与环境、与个人情感建构的合理互动、优化协调才能逐步形成，这种影响与优化的过程具有动态性、流动性和社会建构性[5]。但无论要素之间如何影响都可以看出，作为专业学习者的主体性在其中的作用都是不容忽视的。需要在以个体主体性为基础的前提下，探究各要素之间带有关联性或相互制约性的影响，并基于此，重点考虑在实际专业认同提升实践中的体系构建。

参考文献

[1] 教育部关于做好2022届全国普通高校毕业生就业创业工作的通知[J].中华人民共和国教育部公报,2022(03):74-78.

[2] 李彦花. 中学教师专业认同研究[D].西南大学,2009.

[3] AJZEN I. From intentions to actions:A theory of plannedbehavior [J] . Action control. Springer, 1985:11－39.

[4] 段文婷,江光荣.计划行为理论述评[J].心理科学进展,2008(02):315-320.

[5] 袁朝阳.学前教育专业认同感现状及对策研究[J].兰州教育学院学报,2011,27(05):124-126

本项目受“浙江工业大学2024年院级运河杯科技项目”和“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”共同资助

义务教育阶段教师智力流动意愿影响因素调查研究

A Study on the Factors Influencing Teachers' Intellectual Mobility Willingness in Compulsory Education Stage

曾颖红[1*]，林晓凡[1]
华南师范大学教育信息技术学院
*1229812519@qq.com；linxiaofan@m.scnu.edu.cn

【摘要】近年来教育部发布了多项政策，旨在推动教师交流轮岗，以实现义务教育优质均衡发展。然而传统的交流轮岗方式并未达到预期效果，因此引出了“智力流动”概念。一些学者指出，智力流动是解决教师实体流动问题的有效方式之一，但对于影响教师智力流动的因素却鲜有关注。本研究选择了 161 名义务教育阶段教师作为研究对象，通过调查人口特征、学校条件、对政策措施的态度以及数字能力四个方面来探究影响他们智力流动意愿的因素。研究结果显示，教师的职称、对政策措施的态度以及数字能力等五个方面对义务教育阶段教师智力流动意愿具有显著的影响。

【关键词】智力流动；义务教育阶段教师；流动意愿

Abstract : *In recent years, the Ministry of Education has issued a number of policies to promote the exchange and rotation of teachers in order to realize the quality and balanced development of compulsory education, but the traditional exchange and rotation has not achieved the expected results, which has led to the concept of "intellectual exchange", which has been pointed out by many scholars to be one of the effective ways to solve the problem of the physical mobility of teachers, but little attention has been paid to factors affecting the intellectual mobility of teachers. However, little attention has been paid to the factors affecting teachers' intellectual mobility. In this study, 161 teachers in compulsory education were selected as research subjects to investigate the factors influencing teachers' willingness to move intellectually from four aspects: demographics, school conditions, attitudes toward policy measures, and numerical ability. The results of the study showed that five aspects, namely, job title, teachers' attitudes toward policy measures, and digital competence, had a significant effect on teachers' willingness to be intellectually mobile at the compulsory education level.*

Keywords: Intellectual mobility; teachers in compulsory education; willingness to move

1. 引言

近年来我国发布了大量政策将教师交流作为一项重要举措来促进义务教育质量优质均衡发展，在义务教育阶段，我们追求的是教育的公平性，而为了实现这种公平，教师队伍的均衡配置就是关键之一。目前社会对教师交流的关注点聚焦于教师实体流动,《关于推进县（市、区）域内义务教育学校校长教师交流轮岗的意见》（教师〔2014〕4 号）明确提出“要实现县（区）域内校长教师交流轮岗的制度化、常态化，实现县（区）域内校长教师资源均衡配置”，但目前教师实体流动参与度低，执行过程并不顺利。因此，利用信息技术来改变传统的教师流动已经成为促进教育资源均衡配置的一种创新方式（安富海, 2017)，但教师智力流动能否持续开展，会受到哪些因素的影响和限制，为了提升教师智力流动的效率，这也是我们值得

*基金项目：本文为国家自然科学基金项目“智能增强现实学习环境中师生双向支架匹配机理与评测研究”（项目编号： 62007010）、2024 年广东省科技创新战略专项资金重点项目 “数字化转型赋能义务教育公平的机理研究”（项目编号：pdjh2024a114）的阶段性研究成果。

思考的问题。鉴于此，本研究从影响义务教育阶段城乡一线教师智力交流实施的意愿和行为视角出发，探究影响义务教育阶段教师进行智力交流的关键因素，并提出相应策略，以期促进教师智力流动的常态化实施与发展。

2. 文献综述

2.1. 教师智力流动

教师智力流动当前有四个核心要素，主要为：人岗不动、优势师资、网络学习空间以及机制保障，其中网络学习空间为必要条件（安富海, 2017)。2009 年，欧盟为了解决高等教育领域教师或研究人员因“实体流动”而带来的科研困难，率先提出了“虚拟流动”这一概念，即通过 ICT 构建的学习环境，教师和学生无需离开自己所在的大学即可实现和其他学校的交流(Daukšien et al., 2009)。Sadowski-Smith et al.（2015）等人认为智力流动是指高技能、高学历专业人员跨国网络人才流动的一种趋势，他们比起普通的低技能劳动力移民更具有贡献值，可以概念化为一种人才循环形式。张立国和贺宝勋(2023)基于欧盟所提出的“虚拟流动”这一概念做了进一步的扩展，他认为教师的智力流转是一种虚拟流动，它是指通过网络空间来对优秀教师的无形教育智力资源进行共享，以实现优质教师智力资源优化配置的一种教师流动方式，现阶段，“教师虚拟流动” 亦可被称为“教师在线流动”“教师网络流动”或“虚拟教师流动”。

总体而言，教师智力流动是教师流动的特殊流动形式，他是一种可以通过网络环境，将教师智力输送到农村和薄弱学校地区，从而帮助当地的师生获得精准服务以实现教育资源优质均衡发展的独特流动方式。

2.2. 教师智力流动影响因素

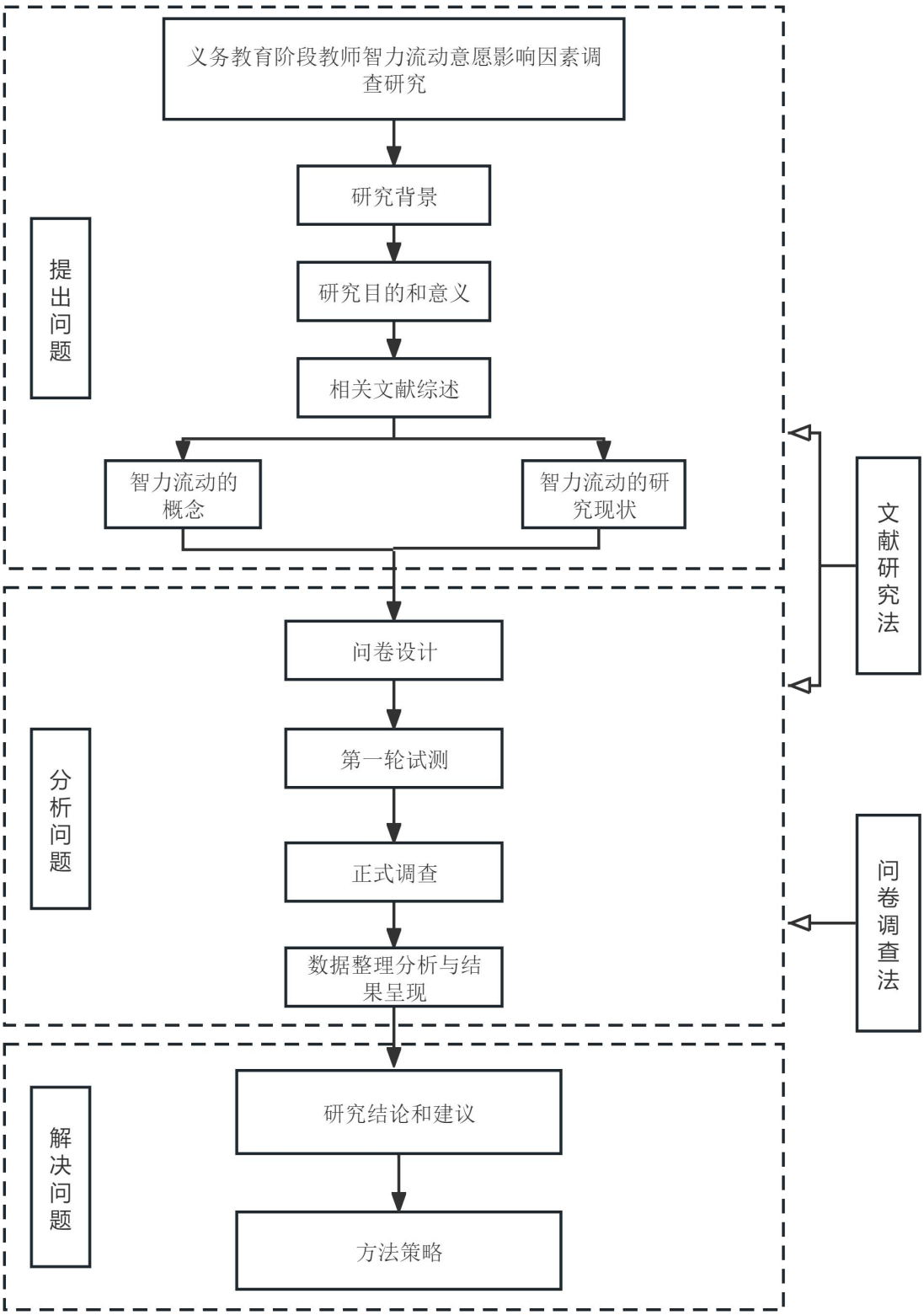
教师决定任教地点往往会受到多方面因素的影响，其中个人特征就是一个重要因素，教师的性别、年龄、职称以及教学经验等个人特征在教师工作这一研究领域中受到了重视(Forsey, 2010；安晓敏 & 佟艳杰, 2019; Vagi & Pivovarova,2017）。已有研究表明，科学配置教育空间资源是保障城乡教师流动空间公平的重要保障，当前有些地方通过“网络走教”模式来促进教育公平，但在一些偏远地区，信息设备工具缺乏，教学设施不完善，这一模式也未能得到很好的实施，因此学校条件也是智力流动的影响因素之一（尹建锋, 2020）。为了实现更高效的智力流动，安富海（2017)和徐刘杰(2021)表明应通过一定的政策以及机制保障来引导优质教师在规定的时间内从实体学校流动到网络学校空间，以实现教师的“虚拟流转”。当前传统的校长教师交流轮岗在政策层面得到了充分的保障机制，但教师智力流动的政策目前还比较缺乏（张立国 & 贺宝勋，2023）。除此之外，美国加纳地区一项相关研究也表明学校教师的数字化应用能力是学校之间数字鸿沟的重要因素，在教育薄弱地区的学校，教师的数字技术应用水平明显较低，同时开展教师网络空间交流活动的频率也较低(Quaicoe & Pata, 2023)，教师智力流转涉及到信息技术设备的使用、数字技术知识与技能等，要实现高效的智力流转，教师需要具备相应的数字能力，因此教师的数字能力这一因素也在本研究的考量之中。

综上所述，教师的个人特征、学校条件、政策措施以及数字能力四个方面在教师智力流动四个方面有着重要影响，但对于他们是如何影响教师智力流动的，以往的研究存在着不同的观点和看法。

2.3. 研究目的

本研究在以往教师传统流动的基础上，从影响义务教育阶段城乡一线教师智力交流实施的意愿和行为视角出发，基于人口特征、学校条件、政策措施以及教师数字能力三个方面来探究影响义务教育阶段教师进行智力交流的关键因素。

2.4. 技术路线



3. 调查研究设计

3.1. 调查对象

本研究选取了义务教育阶段的教师进行调查，涵盖小学和初中两个学段，在全国范围内对义务教育阶段的中小学老师采用了随机抽样的方法。本调查共收集问卷 180 份，根据作答

时间过短以及随意填写等因素，剔除无效问卷 19 份，回收有效问卷 161 份。在调查对象中女性教师为 96 名，男性教师 65 名，女性教师占被试的多数。问卷填写大约需要 3 分钟，问卷回收完成后，将所有样本数据放入 SPSS 27.0 中进行统计分析。

3.2. 调查工具

本研究的问卷主要由两部分组成，包括教师基本信息和量表两部分。基本信息共 7 个题项，主要涉及教师的性别、年龄、职称、流动意愿等。量表有三个维度，分别是教师对学校条件、政策措施的态度以及教师的数字能力。本量表中对学校条件的态度参考了王艳玲（2023）用于调查乡村教师离职意愿影响因素的问卷。对政策措施的态度改编自 Wang（2023）的量表，通过对学校一线教师的实际访谈，部分教师认为参与智力流动更倾向于是为了获得精神满足，因此在原量表中新增一条针对教师职业满足感的项目，同时将原量表的描述中的教师流动改为教师智力流动。教师数字能力自评部分采用了由 Hatos 等人（2022）编制的用于衡量教师自我数字能力的量表，并将量表题项中的相关主体改编为教师智力流动的相关内容。量表中每个题项均采用李克特五点量表评分，1～5 分表示从“完全不同意”到“完全同意”。

表 1 各维度基本信息

维度	题数	含义	参考来源
学校条件	5	测量目前教师对条件的态度，以此来了解学校条件对智力流动的影响程度	王艳玲, 2023
政策措施	6	测量教师对各项教师智力流动有关措施的态度，以此来了解每项措施对教师智力流动意愿的影响程度	Wang et al., 2023
数字能力	4	测量教师使用数字化工具进行智力资源制作及开展智力交流活动的能力程度	Hatos et al., 2022

3.3. 数据分析

本研究首先采用主成分分析法来确认因子的结构，并计算每个因素的 alpha 系数来确保每个因素的可靠性。随后使用卡方检验来检验义务教育阶段教师智力流动意愿在人口统计学变量方面的差异，同时通过独立样本 *t* 检验来揭示义务教育阶段教师智力流动意愿在学校条件、教师对政策措施的态度以及教师数字能力之间的差异，并使用描述性分析来对这一方面的数据进行说明。此外，针对研究问题，本研究还采用皮尔逊相关法来探讨教师流动意愿与学校条件、教师对政策措施的态度以及教师数字能力之间的相关关系。

4. 调查结果

4.1. 探索性因子分析

对量表采用的提取方法是主成分分析法，其结果（表 2）显示 KMO 值为 0.87，大于 0.7，同时 p 值小于 0.001，说明数据可以进行主成分提取。基于特征根大于 1 的标准，本次共提取了 3 个因子。探索性因子分析结果如表 3 所示，本量表中最终共 15 个项目被保留了下来，所有研究项对应的共同度值均高于 0.4，说明研究项信息可以被有效的提取，旋转后累积方差解释率为 62.308%>50%。意味着研究项的信息量可以有效的提取出来。

表 2 KMO 和 Bartlett 的检验

KMO 值		0.87
Bartlett 球形度检验	近似卡方	1029.565
	df	105

	p 值	0.00

表 3 旋转后的成分矩阵 a

	成分		
	因子 1	因子 2	因子 3
数字能力 1			0.768
数字能力 2			0.688
数字能力 3			0.763
数字能力 4			0.687
学校条件 1		0.718	
学校条件 2		0.73	
学校条件 3		0.811	
学校条件 4		0.793	
学校条件 5		0.688	
政策措施 1	0.826		
政策措施 2	0.791		
政策措施 3	0.753		
政策措施 4	0.722		
政策措施 5	0.843		
政策措施 6	0.744		

旋转后累积方差解释率：62.308%

4.2. 信度分析

运用 SPSS 27.0 统计分析软件对问卷信效度进行测验，结果如表 4 显示，问卷各部分信度为(Cronbach α)为 0.762~0.887，表明问卷的信度较好，符合测量要求。

表 4 Cronbach 信度分析

维度	项数	样本量	Cronbach α系数
学校条件	5	161	0.833
政策措施	6	161	0.887
数字能力	4	161	0.762

4.3. 义务教育阶段教师智力流动意愿影响因素分析

首先是人口统计学因素，为了确定人口统计学因素对教师智力流动意愿的影响，本次调研使用卡方检验，通过表 5，我们可以发现，职称、学历以及教学经验三个因素对教师智力流动意愿有显著影响。在职称方面，高级职称愿意流动的比例明显低于其他三个职称的老师。其次是学历方面，本科学历的老师更倾向于接受智力流动。最后是教学经验，教学经验为 0-5 年的教师接受教师流动的意愿最高。

表 5 人口统计学特征的卡方检验结果

个人特征	分类	智力流动意愿		总计	X^2	*p*	*eta*
		愿意	不愿意				
性别	男	37	28	65	1.574	0.21	0.099
	女	64	32	96			
职称	无职称	35	16	51	11.42	0.010**	0.213
	初级职称	41	14	55			
	中级职称	20	22	42			
	高级职称	5	8	13			
学历	专科	13	20	33	9.749	0.008**	0.181

	本科	53	23	76			
	研究生及以上	35	17	52			
学校类型	农村	49	27	76	0.187	0.666	0.034
	城镇	52	33	85			
	0-5 年	59	21	80			
	6-10 年	12	7	19			
教学经验	11-15 年	12	14	26	10.968	0.027*	0.205
	16-20 年	4	7	11			
	20 年以上	14	11	25			
	25 岁以下	37	13	50			
	26-30 岁	21	11	32			
年龄	31-35 岁	13	9	22	5.443	0.245	0.178
	36-40 岁	11	10	21			
	40 岁以上	19	17	36			

注：* p<0.05， ** p<0.01， *** p<0.001

其次调查从教师对学校条件、政策措施的态度以及教师数字能力入手，利用 *t* 检验去研究智力流动意愿对于学校条件、政策措施以及数字能力之间的差异性。由表 6 可知，不同智力流动意愿样本对于学校条件并不会表现出显著性($p > 0.05$)，意味着不同智力流动意愿样本对于学校条件均表现出一致性，不存在差异性。另外政策措施、数字能力呈现出显著性($p < 0.05$)，意味着不同智力流动意愿样本对于在政策措施和数字能力之间存在着差异性。

表 6 学校条件、政策措施以及数字能力 *t* 检验分析结果

	智力流动意愿(平均值±标准差)		*t*	*p*
	愿意(n=101)	不愿意(n=60)		
学校条件	3.94±0.68	3.94±0.80	0.012	0.991
政策措施	3.95±0.65	3.59±1.07	-2.337	0.022*
数字能力	4.05±0.60	3.76±0.85	-2.304	0.023*

注：* $p<0.05$ ， ** $p<0.01$ ， *** $p<0.001$

4.4. 教师智力流动意愿影响因素的相关性分析

本研究利用相关性分析近一步探讨了教师智力流动意愿影响因素，由表 6 可知，智力流动意愿和学校条件之间的相关系数值为-0.001，接近于 0，并且 *p* 值为 0.991 > 0.05，说明智力流动意愿和学校条件之间并没有相关关系。智力流动意愿和政策措施之间的相关系数值为 0.204，并且呈现出 0.01 水平的显著性，因而说明智力流动意愿和政策措施之间有着显著的正相关关系。智力流动意愿和数字能力之间的相关系数值为 0.195，并且呈现出 0.05 水平的显著性，因而说明智力流动意愿和数字能力之间有着显著的正相关关系。

表 7 教师智力流动意愿影响因素的相关性分析

	智力流动意愿	学校条件	政策措施	数字能力
智力流动意愿	1			
学校条件	-0.001	1		
政策措施	0.204**	0.321***	1	
数字能力	0.195*	0.431***	0.361***	1

注：* $p<0.05$ ， ** $p<0.01$， *** $p<0.001$

5. 研究结论

由上文的统计分析结果可以得出，职称、学历、教学经验、政策措施和教师数字能力对教师智力流动的意愿有显著影响。

首先是在职称和教学经验这一部分，本研究印证了 Wang（2023）和 van Geffen（2014）的研究，高职称以及教学经验较丰富的教师流动意愿显著低于年轻或低职称教师。在学历方面，本科学历的老师流动意愿也较高，受到学历的影响，他们更希望从提升教学能力这一方面来寻求自己的晋升机会（Cañado, 2020）。

其次是政策措施和教师数字能力方面，本研究结果表明，教师对政策措施的看法表明了政策措施制定得越完善，教师的流动意愿越高，政府需要提供相应的支持和资源，如教师培训计划、交流机会和奖励机制。而对于数字能力，在本研究中，智力流动意愿和数字能力之间有着显著的正相关关系，这说明如果教师具备较高的数字能力，他们会更加自信地应对数字化教学环境，并且更愿意进行智力交流。此外，数字技术也可以在教师之间实现更加便捷的智力交流，如在线会议、协作平台和共享资源库等。因此，教师的数字能力是非常重要的，这将有助于促进教师之间的智力交流和创新，提高教育质量和教学效果。

最后，和以往传统的教师流动研究相比，传统教师流动在互联网这一中介的支持下，学校条件不再成为限制教师流动的条件之一，以往研究表明教师流动会受到学校环境的限制（Vagi & Pivovarova, 2017），而在本研究中教师智力流动意愿与学校条件并未呈现出显著的相关性。

6. 方法策略

为了实现教师轮岗的目标，国家应加强相关政策的制定与实施，例如提供相应的支持和资源，如教师培训计划、交流机会和奖励机制等，并确保奖励政策落实。同时教师个体也应意识到数字能力的重要性，并主动增强自己的数字能力。一方面，学校应提供给教师提升的机会，例如优秀教师经验分享，或者定期的培训计划，除此之外，还应采购一些必要的信息化设备，例如希沃白板等。另一方面，教师应不断学习和掌握最新的数字工具和技术，如国家中小学智慧教育平台、在线协作平台等，以便更好地运用这些工具和资源来支持教学和智力交流。其次，在教学方面，教师应学会使用技术和工具来辅助自己的教学，通过对学生的课堂行为数据进行分析和挖掘，了解到学生的个性和需求，从而针对实际情况实施个性化和差异化的教学。此外，教师还应积极参与智力交流活动，在网络教研共同体中与优秀同行分享交流与学习，以此提升自己的数字素养和创新能力。

参考文献

中国人民共和国教育部.教育部 财政部 人力资源和社会保障部关于推进县（区）域内义务教育学校.校长教师交流轮岗的意见[EB/OL].[2014-08-15].http://www.moe.gov.cn/srcsite/A10/s7151/201408/t20140815_174493.html

安富海.(2017).学习空间支持的智力流动:破解民族地区教师交流困境的有效途径.**电化教育研究(09),**102-107.doi:10.13811/j.cnki.eer.2017.09.015.

张立国 贺宝勋.(2023).教师虚拟流动的概念特征、基本业态与实践应用.**现代教育技术(05)**,41-49.

安晓敏 佟艳杰.(2019).教师轮岗交流意愿影响因素研究.**教育科学(03)**,43-50.

尹建锋.(2020).城乡教师流动的空间正义及其实现.**教育研究(01)**,136-147.

徐刘杰.(2021).“互联网+”教师智力流动的影响因素研究.**数字教育(02)**,27-32.

王艳玲.(2023).乡村教师离职意愿影响因素的变与不变——基于《云南省乡村教师支持计划（2015—2020 年）》实施前后的比较.**华东师范大学学报(教育科学版)(09)**,85-99.doi:10.1638

2/j.cnki.1000-5560.2023.09.006.
Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113–132. https://doi.org/10.1002/pam.20072
Cañado, M. L. P. (2020). Addressing the research gap in teacher training for EMI: An evidence-based teacher education proposal in monolingual contexts. *Journal of English for Academic Purposes, 48*, 100927. https://doi.org/10.1016/j.jeap.2020.100927
Daukšien, E., Teresevičien, M., & Volungevičien, A. (2009). Teacher virtual mobility for research and practice. In Innovation and creativity in e-learning: international conference proceedings (pp. 1-6). https://hdl.handle.net/20.500.12259/44152
Forsey, M.G. (2010). Publicly minded, privately focused: Western Australian teachers and sc hool choice. *Teaching and Teacher Education, 26*(1), 53–60. https://doi.org/10.1016/j.tate. 2009.06.011
Hatos A, Cosma M L, Clipa O. Self-assessed digital competences of Romanian teachers dur ing the COVID-19 pandemic[J]. *Frontiers in Psychology,* 2022, 13: 810359. https://doi.o rg/10.3389/fpsyg.2022.810359
Sadowski-Smith, C., Yu, W., Wang, H., & Li, W. (2015). Intellectual migration and brain circulation: Conceptual framework and empirical evidence. Journal of Chinese Overseas, 11(1), 43-58. https://doi.org/10.1163/17932548-12341300
Quaicoe, J. S., & Pata, K. (2020). Teachers' digital literacy and digital activity as digital divide components among basic schools in Ghana. *Education and Information Technologies, 25,* 4077-4095. https://doi.org/10.1007/s10639-020-10158-8
Vagi, R., & Pivovarova, M. (2017). "Theorizing teacher mobility": A critical review of literature. *Teachers and Teaching, 23* (7), 781–793. https://doi.org/10.1080/13540602.2016.1219714
van Geffen, R. E., & Poell, R. F. (2014). Responding to teacher shortages: relationships among mobility experiences, attitudes, and intentions of Dutch teachers. *Asia-Pacific Journal of Teacher Education, 42*(3), 275-290. https://doi.org/10.1080/1359866X.2014.902424
Wang, F., Tang, H., Shen, W., Wang, Y., & Lo, L. N. (2023). Factors influencing teachers' willingness to move in the Chinese school system. *Asia Pacific Journal of Education, 43*(1), 203-218. https://doi.org/10.1080/02188791.2021.1923457

趣創者理論在網頁設計課程上的應用

Appling IDC Theory in Web Design Courses

莊益瑞
景文科技大學 行動商務與多媒體應用系
*yrjuang@just.edu.tw

【摘要】 設計類課程重點在培養學生創意思考，進而將之實踐在作品的創新上，而促成學習的動力來自興趣，培養學生學習興趣則取決於學習活動的設計，如何設計能引起學生產生興趣、持續學習、持續創作的學習活動，是累積趣創者理論教學實踐經驗的重要實徵研究工作。學習科技學者陳德懷等人在 2018 年正式發表趣創者理論，將學習視為一個創造過程，反之亦然。本研究基於趣創者理論的興趣環嵌入創作環的概念，設計由模仿、結合和展演三個學習活動建構的教學模式，並在每個學習活動中設計了引趣、入趣和延趣組成的學習循環步驟，應用在以技術實作為主軸的網頁設計課程上，藉以觀察學生的學習興趣、學習成效和創造力是否有所提升。經過單組前後測實驗設計，透過學習興趣量表、期中期末考和創造力測驗 ATTA 進行量化研究分析，研究結果顯示學生的學習興趣、學習成效與創造力均有顯著提升。焦點團體訪談的質化分析也顯示學生的正向回饋，同時提出許多待改善的缺點，以及未來的研究建議。

【關鍵詞】 趣創者理論；學習興趣；創造力；網頁設計

Abstract: *In technical and vocational education, the curriculum of the design cluster focuses on cultivating students' creative thinking and practicing it in the innovation of works. The motivation to promote learning comes from interest, and cultivating students' learning interests depends on the design of learning activities. Designing learning activities that can arouse students' interest, continuous learning, and continuous creation is important research work to accumulate practical experience. In 2018, Prof. Chan, an outstanding scholar of learning technology, and his colleagues officially published the IDC theory, which views learning as a creation process and vice versa. This study is based on the concept of the interest loop embedded in the creation loop of IDC theory and designs a teaching model constructed from three learning activities: imitating, combining, and staging. Each learning activity is designed with cyclic learning steps consisting of triggering, immersing, and extending. This study applies this teaching model to a web design course that focuses on technical implementation to observe whether students' learning interests, learning effects, and creativity have improved. After a single-group pre-test and post-test experimental design, quantitative research and analysis were conducted through the learning interest scale, midterm and final-term, and creativity scale ATTA. The research results show that students' learning interests and creativity have been significantly improved. Qualitative analysis of focus group interviews also revealed positive feedback from students, and we also identified some shortcomings that need to be improved, as well as suggestions for future research.*

Keywords: IDC Theory, Learning Interest, Creativity, Web Design

1. 前言

設計類課程的主要教育目標是培養學生具備設計產業所需之知識與實作技能、系統思考與美學素養，強化透過設計解決問題與創新能力，務求學習內容與設計產業技術接軌（台灣教育部，2022）。換言之，學生進入大學階段後，除了深化技術能力之外，更要注重創造力（creativity）的培養，才能提升未來進入職場的價值與優勢，更是職涯發展中自主學習和不斷精進自我效能的基石。然而過往設計類專業能力檢定或證照考試容易讓學生流於背考古題，

限制了知識與技術的廣泛應用，也抹煞了學生的創造力，即便考取了證照，實作能力仍顯薄弱，難以應付未來職場的挑戰。

學習科技學者陳德懷教授在台灣所推行的「身教式持續安靜閱讀（Modeled Sustained Silent Reading; MSSR）」計畫（陳德懷和明日閱讀研究團隊，2016）強調以興趣為本的閱讀，培養學生終身閱讀習慣與能力。此計畫的背後是基於一個學習設計的理論，陳德懷教授帶領一批亞洲學者在 2015 年的「國際電腦在教育應用研討會（ICCE）」上首次發表了「趣創者理論（Interest-Driven Creator Theory）」（Chan et al., 2015; Wong et al., 2015; Chen et al., 2015）。趣創者理論試圖建議如何引導學生有興趣地、習慣性地進行創造，以期培育年輕的下一代成為終身的創造者。當中提出的興趣、創造和習慣等三個循環框架可以各自執行，也可以組合成相輔相成的正向循環，依學習內容設計不同的教學活動。後續陳教授等人持續修正、歸納與統整趣創者理論的論述，於 2018 年在國際期刊 Journal of Computers in Education 提出 IDC 理論（Chan et al., 2018），尋求學術社群更多的討論和協作，以加深及進一步發展該理論。本研究試圖基於此宏觀理論設計一個以興趣為驅動的教學模式，應用在網頁設計課程上，並透過實際的教學實驗，評鑑此模式是否能有效提升學生的學習興趣、學習成效及創造力。

2. 文獻探討

趣創者理論包含三個錨定概念：興趣、創造和習慣，探討如何引導學生有興趣地、習慣性地進行創造。每一個概念都會經歷一個連貫的學習過程，均包含三個「元件（components）」，形成一個循環，且這三個環可依教材內容、教學方法、學生特性等以多元的方式連結在一起，形成一個教學模式（Chan et al., 2018）。

興趣環（Interest Loop）由引發興趣（引趣）、沈浸興趣（入趣）和延伸興趣（延趣）三個元件組成，可以粗略的對應到 Hidi 和 Renninger（2006）興趣發展模型的前三個階段（即觸發情境興趣、維持情境興趣和發展個人興趣），但在發展學習設計時，最後將和諧地融入 IDC 整個學習歷程，即將興趣環與與創造環和習慣環連結起來，其目的是將情境興趣轉化為個人興趣，才能讓興趣延續下去（Wong et al., 2020）。

在趣創者理論中認為學習就是創造，創造就是學習，學習過程就是一種創造過程，創造讓學習者有產出、有創意、有成果。從遠古時代的文字描述到現代電子媒體的呈現，讓模仿演變為創造，觀察演變為吸收，複製演變為結合，並建構舞台展演創造作品。因此「創造環（Creation Loop）」由模仿（創造的基礎）、結合（創造的核心）和展演（改進創造品）等三個元件組成，其中模仿又可稱為吸收（Chan et al., 2019）。

心理學家研究發現，習慣是有規律地重複的行為，這種行為通常是無意識的。習慣是由重複發展的、自發認知過程的結果，它並不需要「有意識的」努力。趣創者理論的習慣環（Habit Loop）由三個部分組成：提示環境（地點、時間、人員或事件的安排）、例行活動（重複模式）和滿足（作為學習後正面的心理獎勵）。學習者要在一個能夠引發其動作的情境，接著規律地重複行為而形成習慣，最後透過成就感或獎勵增強學習者的滿意度，以達到持之以恆的目的（Chen, 2020）。

應用趣創者理論在教學上的研究文獻中，對象以國小、大學及研究生為多數，科目包含了語文（含閱讀與寫作）、數學、運算思維（程式設計）、教育科技等，以語文學科領域居多，所發表的文章集中在「全球華人計算機教育應用大會（GCCCE）」和「International Conference on Computer in Education（ICCE）」兩個以數位學習科技領域為主的國際學術研討會。多數結果顯示運用趣創者理論來設計教學活動，可以提升學習興趣或學習動機，以及更主動的學習態度。同時也提出了許多對本研究有用的建議，例如：Kong（2017）建議教師可以提供學生選擇教學內容，決定哪些內容是學生感興趣的；Khambari（2019）建議興趣環

可以應用在微觀的教學活動設計上，創造環可以應用在宏觀的教學設計層面上；Cheng & Zhang（2019）建議在學習活動的設計上可以考量具不同自我調節（self-regulated）能力的學生需求，促使學習習慣的養成；Wong 等人（2020）建議教師對學生感興趣的內容要做出回應很重要，這是促使學生持續保持學習動機的催化劑；劉偉瑄（2018）建議設計一個可以自我調節學習進度的、循序漸進的闖關活動，給予不同的獎勵或稱號，有助於學習興趣的保持和推動自我進度；黃政理（2019）建議可以透過創作教學影片來促成同儕相互教學、合作學習，提升學習動機；劉宸穎（2020）建議增加興趣環中的促進活動或循環次數，可以更加提升學習興趣，亦能增進創造品的品質。

以上文獻雖然研究對象、學科、環境、輔助工具等有所不同，但在掌握趣創者理論的基本概念、實務應用、進行方式與彈性上，提供了建設性的建議。此外，由於培養學生持續創作的習慣與興趣環和創造環的運作模式有很大的關聯，一旦興趣環和創造環能有效正向帶動學生的學習動力，此時配合習慣環的提示環境、例行活動和滿足等三項策略，便能水到渠成。為降低教學設計與評鑑的複雜度，本研究企圖結合興趣環和創造環初探在設計類課程（網頁設計為例）應用的可行性。

3. 教學設計

本研究參考趣創者理論，將興趣環融入在創造環中的每一個元件中，亦即創造環中的模仿、結合和展演三個元件均包含了引趣、入趣和延趣三個元件，使創造過程中都能以興趣為導向來設計教學活動，形成「趣創者式教學（Interest-Driven Creator Teaching）」（圖 1）。以下先說明此教學模式每個元件的教學活動設計構想。

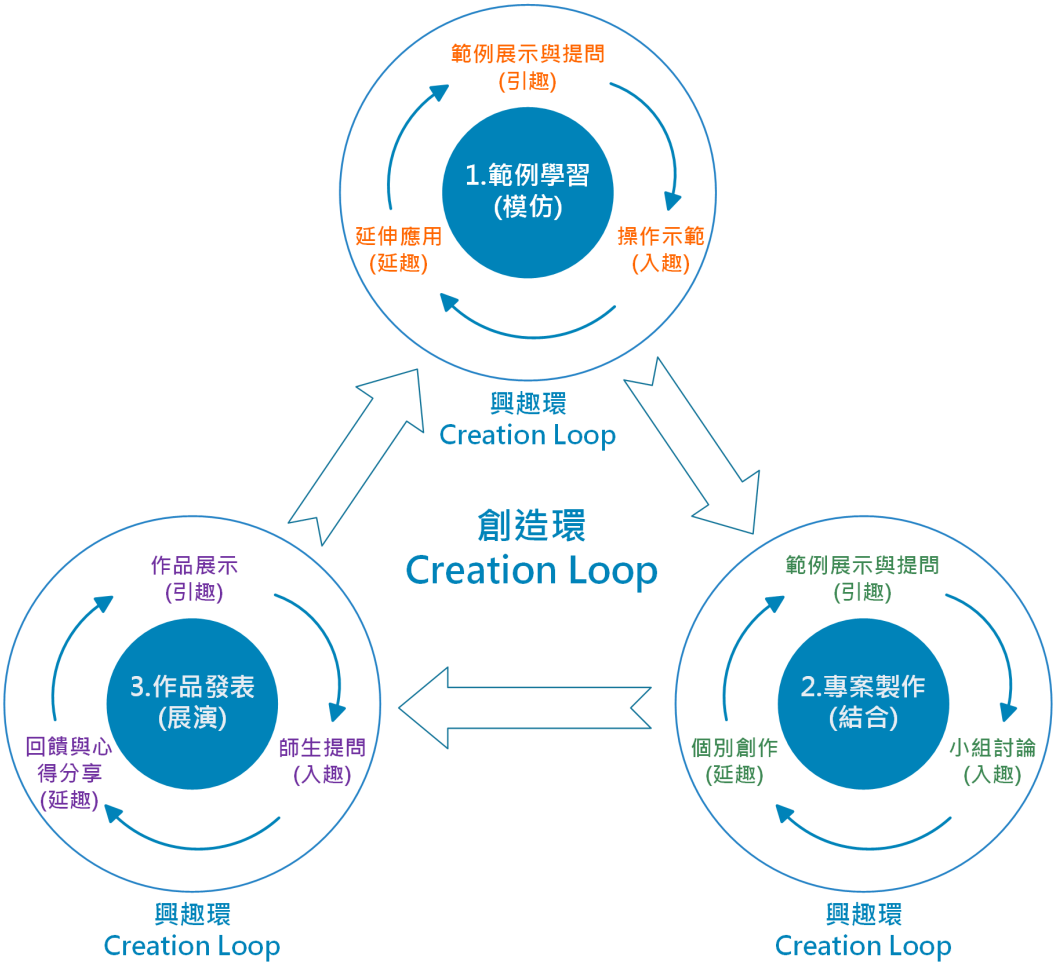


圖 1 趣創者式教學（Interest-Driven Creator Teaching）

模仿：此階段是大量吸收知識與技術的階段，除了閱讀相關教材外，也要多練習範例，從範例裡學習、吸收實作技巧。一開始教師先透過範例作品展示，並在過程中提出可以引起學生好奇心的問題，達到「引趣」的目的。接著實際操作範例的做法，一步步引導學生完成範例，滿足學生的成就感，進入到心流狀態，沈浸在吸收實作技巧的成功經驗中，達到「入趣」的目的。最後，引導學生思考該範例每個操作步驟的意義，是否有其他的應用，或者可以修改成為自己覺得滿意的作品。從該範例延伸嘗試不同的操作和使用不同的工具，使學生更明白範例中所傳達的知識與意義，讓情境興趣轉化為個人興趣，達到「延趣」的目的。

結合：此階段將進入到小組合作學習的階段，共同討論及和實作完成教師指定的專案（Project）。教師先展示該專案可能的完成品，並提示可以運用前面所學的哪些設計技巧來完成，也可以用提問的方式來引起學生的好奇心，達到「引趣」的目的。接著進入到小組討論階段，教師提供具體目標和即時清楚的回饋，給予學生自行選擇和掌控的彈性，以便進入心流狀態。為了讓每位同學都必須有所貢獻，可先讓學生各自寫下對專案的作法，再開始組內彼此分享與回饋。過程中教師巡迴各組，可提示線索或讓學生提問，透過教師的即時回饋，讓討論可以順利進行，甚至熱烈討論，達到「入趣」的目的。最後，學生融合了自己和其他小組成員的建議後各自完成專案，而不是小組共同完成，目的是讓學生能將來自小組討論的收穫整合成自己的創作想法，令他們反思自己到底懂得多少，達到「延趣」的目的。

展演：此階段為接續前一階段完成個人創作後的展示階段，所有同學均將自己的創作上傳至學習管理系統（如 Moodle），教師可依據學生人數和時間評估選擇多少位同學上台分享作品，選擇的方式可以是徵求自願者、經由老師巡堂觀察到不錯的作品、由小組遴選推薦、隨機抽點等。透過作品展示與分享創作過程，除了讓學生產生好奇心，也可以學習他人的設計技巧，達到「引趣」的目的。在展示期間，鼓勵同學向展示者發問，但通常學生比較保守，可以先由老師提問（可向展示者或聆聽者提問），或者給予提問者獎勵（例如加分），促成同學之間的熱烈討論，達到「入趣」的目的。在分享者結束分享後，可以請同學在學習管理系統給予分享者評價與回饋（同儕互評），同時也分享自己的心得，藉此將展示階段所學習到的知識與技能留下紀錄，作為日後設計更大型的專案時參考，達到「延趣」的目的。

4. 研究設計

本研究以台灣一所科技大學的多媒體設計系在大一開設的「網頁設計」課程為例進行教學實驗，由於該系大一只有一個班級，在無法找到相同背景及無法同時開設兩班的情況下，將採用「前實驗設計（pre-experimental design）」之「單組前後測設計」，在沒有控制組的情形下，儘量控制教學實驗過程均在同一個場域、同一位授課教師、同一份教材下進行。整學期的前八週將採用一般翻轉教學模式，每一個教學單元包括課前自學、課堂核心觀念教學、教師帶領範例練習、類似題實作、以及教師評量等五階段。期中考後的八週則採用上述之「趣創者式教學」。研究架構將以學生在不同教學方法（趣創者式教學和一般翻轉教學）為自變項，研究目的中欲提升或改善的學習興趣、學習成效和創造力等為依變項，進行量化與質化的研究。研究假說如下：

H1：採用趣創者式教學比一般翻轉教學更能引起學生的「學習興趣」。

H2：採用趣創者式教學比一般翻轉教學更能提升學生的「學習成效」。

H3：採用趣創者式教學比一般翻轉教學更能增進學生的「創造力」。

在「學習興趣」方面的研究工具上，本研究將參考 Pintrich & De Groot（1990）的動機與自我調節學習量表（Motivational and self-regulated learning scale），採用其中的動機信念的題目改編為適合本研究情境之學習興趣量表。此量表分為自我效能（self-efficacy）和內在價值（intrinsic value）等二個向度。自我效能是指對自己是否學會某技巧、是否有能力完成特定

目標等抱持的信念。內在價值是指對特定教材或課程所抱持的信念，包含認為某課程重要與否、實用與否、有無興趣等。此動機信念量表採用 Likert-5 point scale 編製，這二個向度的信度係數 Cronbach α 分別為：自我效能 α=.89，內在價值 α=.87。此「學習興趣量表」於學期初、期中評量和期末評量各填寫一次，用來評估學生在這三個時間點學習動機的變化，做為學生學習興趣改變的參考。

在「學習成效」方面的研究工具上，將由研究者依照授課內容編製一份學習成就測驗，內容包含觀念題和實作題，觀念題以選擇題為主，實作題以實際操作網頁設計軟體可完成的應用問題為主。分別在期中考和期末考實施，以比較兩次成績的差異。

在「創造力」方面的研究工具上，本研究將參考學者陳長益（2006）所修訂的「陶倫斯創造力測驗成人適用精簡版」，它是根據 Kathy Goff 與 E. Paul Torrance 編製之 Abbreviated Torrance Test for Adults（ATTA），建立台灣地區的常模所修訂而成。該量表結合原 TTCT 的圖形版與語文版測驗，簡化為成人適用之精簡版本。要求受試者分別以書寫與畫圖的方式，表現對刺激圖文的反應，藉此評估受試者的創造力。在信度和效度的評估上，該量表從台灣地區成人中，考量地區性、學系、職種等因素，取樣 627 名成人樣本，年齡層涵蓋 18~57 歲。信度採重測信度與評分者間信度，各項創造力指標的重測信度介於.34~.68，評分者間信度則介於.311~.975，皆達.01 顯著水準，顯示本測驗具有不錯之穩定性。在效度方面，以「問題解決創造力測驗」（朱錦鳳，2005）為效標分析，相關係數為.37 與.46，均達.05 之信賴水準，顯示本測驗具有不錯之效度（陳長益，2006）。

本研究將依照 ATTA 作答評分標準取得受測者四種創造力之反應分數，分別為流暢性、獨創性、變通性、精密性四個層面，再依據指導手冊內所指示的評分標準，分別算出各層面的原始分數，對照等級表決定受試者在這四個層面的等級分數，加上創造力指標的分數，即為本研究之創造力指數，藉以評估受測者的創造力。

此外，為補充量化研究的不足，同時採用焦點團體訪談，以學生立場來看趣創者教學進行過程中的優缺點及遇到的問題。受訪對象選擇學習成就評量低、中、高分的學生各 2 名，以及在上述量化評鑑過程中最高分與最低分的同學各 1 名，進行訪談與釐清學生實際的想法。

5. 研究結果與討論

5.1. 學習興趣（H1）

期初到期中採用的是一般翻轉教學，期中到期末採用的是趣創者教學，因此學生在期初、期中和期末三個時間填寫「學習興趣量表」，分數依李克特 5 點量表換算（5 分-非常同意 ~ 1 分-非常不同意）。由於三個測驗時間的受測人數不同，且研究者使用的是單組前後測實驗設計，故採用三次皆有填寫量表的受測者數據（共 23 人），且期中測驗數據可同時當作一般翻轉教學的後測及趣創者教學的前測，藉以進行成對樣本 *t* 檢定。量表中所包含的二個向度及總平均在不同教學法的前後測所得之平均數及 *t* 檢定結果如表 1。

從表 1 的數據可觀察到一般翻轉教學的前後測並未達到顯著改變，顯示受測者在這個教學法期間的學習興趣沒有太大的改變，甚至有少許的降低（t<0）。反之，在趣創者教學期間，不論在自我效能、內在價值和總平均，都有顯著的正向改變（p<.01），顯示此教學法能提升受測者的學習興趣。當中幾個有達到顯著改變的題目包括：「與班上其他人相比，我認為我是一個好學生」、「對於老師指定的作業和任務，我確信我可以做得很好」、「與班上其他人相比，我的學習能力很強」、「我喜歡具有挑戰性的課堂作業，這樣我就可以學習新事物」、「我認為我們在這堂課中學到的東西很有趣」等五項。

表 1 「學習興趣量表」各向度在不同教學法的 t 檢定結果

自變項	測驗向度	測別	人數	平均數	變異數	F	P(F<=f)	t	P(F<=f)
一般翻轉教學(成對樣本t檢定)	自我效能	前測	23	3.48	0.4002	0.8811	0.3846	-0.1043	0.9179
		後測	23	3.47	0.3527				
	內在價值	前測	23	3.69	0.5488	0.7042	0.2086	-0.0858	0.9324
		後測	23	3.68	0.3864				
	總平均	前測	23	3.59	0.4166	0.7966	0.2991	-0.1073	0.9155
		後測	23	3.58	0.3319				
趣創者教學(成對樣本t檢定)	自我效能	前測	23	3.47	0.3527	1.4159	0.2106	3.1744	0.0044**
		後測	23	3.76	0.4993				
	內在價值	前測	23	3.68	0.3864	1.2879	0.2790	3.2293	0.0039**
		後測	23	3.84	0.4977				
	總平均	前測	23	3.58	0.3319	1.4559	0.1926	3.5962	0.0016**
		後測	23	3.85	0.4832				

(**<.01)

5.2. 學習成效（H2）

實作能力透過「學習成就測驗」分別在期中和期末施測，將測驗成績進行「相依樣本 t 檢定」，結果如表 2。觀察檢定結果發現期末考平均成績顯著高於期中考成績（p<.05），顯示學生期末考比期中考有顯著進步。

表 1 學習成就評量之相依樣本 t 檢定結果

	期末考	期中考
平均數	70.3313	56.0681
變異數	538.8657	1056.3153
觀察值個數	47	47
自由度	46	46
F	0.5101	
P(F<=f) 單尾	**0.0122**	
臨界值：單尾	0.6126	

	期末考	期中考
平均數	70.3313	56.0681
變異數	538.8657	1056.3153
觀察值個數	47	47
假設的均數差	0	
自由度	46	
t 統計	2.6461	
P(T<=t) 單尾	**0.0056**	
臨界值：單尾	1.6787	
P(T<=t) 雙尾	**0.0111**	
臨界值：雙尾	2.0129	

5.3. 創造力（H3）

受測者在期中和期末進行陳長益（2006）所修訂的「陶倫斯創造力測驗成人適用精簡版（ATTA）」，參加期中測驗（前測）的人數有 35 位，參加期末測驗（後測）有 26 位，同時參加期中與期末測驗則有 21 位。研究者針對不同測驗人數的前後測進行獨立樣本 t 檢定，為了更精確的比較同一學生前後測的創造力表現，亦針對同時參加前後測的受測者進行成對樣本 t 檢定，檢定總結果如表 3。

ATTA 的計分內容包含四項常模參照評分及十五項效標參照創造力指標，兩者加總後為「創造力指數（Creativity Index, 簡稱 CI）」，表 3 呈現常模參照總級分、效標參照創造力總指標及創造力指數等三種創造力檢定項目在不同 t 檢定下的結果，從檢定結果可知不論是全部樣本或前後均測者樣本的 t 檢定, 這三種檢定項目均達顯著水準 (p<.05 or p<.01 or p<.001)，尤其在常模參照創造力和創造力指數上，讓我們更有信心推斷本研究提出的趣創者教學法可以有效提升受測者的創造力。

表 3 ATTA 創造力測驗 t 檢定結果

t檢定方式	創造力項目	測別	人數	平均數	變異數	F	P(F<=f)	t	P(T<=t)雙尾
獨立樣本t檢定(全部樣本)	常模參照創造力總級分	前測	35	52.63	31.01	1.0448	0.4460	3.9159	0.0003***
		後測	26	58.35	32.40				
	效標參照創造力總指標	前測	35	3.29	9.39	2.1138	0.0214	2.4834	0.0171*
		後測	26	5.81	19.84				
	創造力指數(總級分+總指標)	前測	35	55.91	68.79	1.1510	0.3466	3.7200	0.0004***
		後測	26	64.15	79.18				
成對樣本t檢定(前後均測者樣本)	常模參照創造力總級分	前測	21	53.57	31.66	0.9835	0.4853	4.8486	0.0001***
		後測	21	58.67	31.13				
	效標參照創造力總指標	前測	21	3.95	12.65	1.4744	0.1963	2.3205	0.0310*
		後測	21	5.62	18.65				
	創意指數(總級分+總指標)	前測	21	57.52	78.56	1.0567	0.4516	4.2368	0.0004***
		後測	21	64.29	83.01				

(*<0.05 · **<0.01 · ***<.001)

5.4. 學習成果的質性回饋

透過焦點團體訪談與學生討論趣創者教學的優點，受訪者表示原本以為網頁設計像是程式設計那般難學與無趣的課程，但透過適當的範例展示、方便好用的設計工具、教師循序漸進的示範與練習、小組合作學習、以及公開展示學習成果等以學生為中心的教學模式，提升了學生的學習動機，減少了挫折感，激勵了想要學更多、創作更多的進取心。有受訪者特別提及在專案製作時表面上是小組必須合作完成實作，但教師卻要求每位組員都要做，這樣讓組員間減少搭便車、裝笨或專斷等社會性懈怠的現象，改變了過去分組活動惱人的溝通協調困境。而透過同儕互評也能讓學生嘗試以後設認知的角度，學習評價他人的客觀思維，進而反饋到自己的作品改善上。

在缺點方面，有受訪者表示在「範例學習」階段讓學生思考不同的做法或進階的範例時，在沒有老師的引導下，往往會不知所措，即便與同學討論後仍有困難。此問題牽涉到範例的難易度及是否提供足夠的提示鷹架有關，這是未來設計教材時要評估的重點，教師教學時亦需臨場應變。因為沒有過去的經驗可參考，教師需要花費較多的時間與精力，給學生較多的提示與指導，學生的創造力才會慢慢被啟發，之後的專題就比較順利。

6. 結論與建議

從教學與研究的成果來看，趣創者教學基於學習興趣為導向的概念下，教師改變了教學流程，從引趣到入趣是很重要的啟動階段，從陌生到有動機想一探究竟或動手做，若能妥善的引導與鼓勵，讓學生產生興趣後，再從入趣到延趣，延續其探索更複雜、深入的相關知識與技能，便可順利進入到創造階段。研究者在實施的過程中，所採用的應用範例必須有足夠的誘因與難易度在學生能力所及，方有機會啟發學生學習動機。當學生的學習動力來自學習興趣，一旦進入學習興趣的心流，就能愈熟練與深入知識的核心與關鍵技術，自然能夠在創新與創意上舉一反三，也能拓展更寬廣的創作思維，進而反饋到新知識或技術的學習，形成一個良性的學習循環。

因此，若要採用趣創者教學模式在網頁設計的課程，建議教師需重整教材，多以學生可能有興趣的方向尋找或製作應用範例，此範例除了要能達成的知識與技能目標外，也要具有擴展性，能讓學生舉一反三的仿作或加值的空間。透過不斷練習舉一反三的思維與實作，並養成課前預習的習慣，便可培養學生主動學習與創造思考的學習態度與能力。

参考文獻

台灣教育部（2022）。技術型高級中等學校群科課程手冊。2023 年 1 月 18 日從: https://www.naer.edu.tw/upload/1/16/doc/2035/設計群課程手冊（111 年 7 月更新）.pdf。

朱錦鳳（2005）。問題解決創造力測驗。台北：心理出版社。

陳長益 (2006)。陶倫斯創造力測驗成人適用精簡版(ATTA)(Abbreviated Torrance Test for Adults)。台北市：心理出版社。

陳德懷、明日閱讀研究團隊 (2016)。明日閱讀：明日主題學習的基礎。台北市：天下雜誌。

陳德懷、廖長彥、呂賜杰、張立杰、簡子超、陳志洪和周志岳 (2016b)。趣創者理論: 創造與創造環（第一部分）。第 20 屆全球華人計算機教育應用大會（GCCCE 2016），中國香港。

陳德懷、廖長彥、鄭年亨、陳志洪、張立杰、黃龍翔、陳文莉、呂賜杰和趣創者任務小組 (2016a)。趣創者理論初探：二十一世紀學習設計理論。第 20 屆全球華人計算機教育應用大會（GCCCE 2016），中國香港。

黃政理 (2019)。透過學生影片創作進行國小數學學習:趣創者理論之應用。國立中央大學博士論文。引自: https://hdl.handle.net/11296/5t4ng6

劉宸穎 (2020)。設計與實踐「提升式寫作」活動以提升國小學生寫作品質與寫作興趣。國立中央大學碩士論文。引自：https://hdl.handle.net/11296/7g5t9q

劉偉瑄 (2018)。探討國小低年級學童在數學自我進度學習下的發展與影響。國立中央大學碩士論文。引自: https://hdl.handle.net/11296/7kd3b7

Chan, T. W., Looi, C. K., Chen, W., Wong, L.-H., Chang, B., Liao, C. C. Y., et al. (2018). Interest-driven creator theory: Towards a theory of learning design for Asia in the twenty-first century. *Journal of Computers in Education*, 5, 435–461.

Chan, T.-W., Looi, C.-K., Chang, B., Chen, W., Wong, L.-H., Wong, S. L., et al. (2019). IDC theory: creation and the creation loop. *Research and Practice in Technology Enhanced Learning, 14*(26).

Chen, W., Chan, T.W., Wong, L.H., Looi, C. K., Liao, C. Y., Cheng, H. N. H., et al. (2020). IDC theory: habit and the habit loop. *Research and Practice in Technology Enhanced Learning, 15*(10).

Cheng, H. N. H. & Zhang X. (2019). How the habits of self-regulated academic writers are behaviorally facilitated? In M. Chang, H. J. So, L. H. Wong, F. Y. Yu, & J. L. Shih (Eds.), *Proceedings Volume 2 of ICCE 2019 – 27th International Conference on Computers in Education* (pp. 164-171).

Csikszentmihályi, M. (1990). Literacy and intrinsic motivation. *Daedalus, 119*(2), 115–140. Csíkszentmihályi, M. (1991). *Flow: The psychology of optimal experience*. New York: Harper Perennial.

Goff, K., & Torrance, E. P. (No year specified.). Abbreviated Torrance Test for Adults (ATTA) [Database record]. APA PsycTests. https://doi.org/10.1037/t12015-000

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Khambari, M. N. (2019). Designing an interest-driven challenge-based learning and alternative assessment method for an educational technology undergraduate course. In M. Chang, H. J. So, L. H. Wong, F. Y. Yu, & J. L. Shih (Eds.), *Proceedings Volume 2 of ICCE 2019 – 27th International Conference on Computers in Education* (pp. 153-158).

Kong, S. C. (2017). A survey study for understanding interest of primary school learners in programming mobile apps: Designing an interest-driven curriculum. In M. Chang, M. Jong, T.-W. Chan, S. Yu, F. Wu, B. Li, et al. (Eds.), *Conference proceedings of the 21st Global Chinese Conference on Computers in Education 2017* (pp. 783-786). Beijing: Beijing Normal University.

Wong, L.-H., Chan, T.-W., Chen, W., Looi, C.-K., Chen, Z.-H., Liao, C. C. Y., King, R. B., Wong, S. L. (2020). IDC theory: Interest and the interest loop. *Research and Practice in Technology Enhanced Learning*, *15*(3),

Wong, S. L., Khambari, M. N., Wong, S. L., Voond, X. P., & Wong, L. H. (2020). Igniting student interest towards educational technology through interest driven creator theory: A case study at Universiti Putra Malaysia. In H. J. SO, Ma. M. Rodrigo, J. Mason, & A. Mitrovic (Eds.), *Proceedings of ICCE 2020 – 28th International Conference on Computers in Education* (pp. 279-285).

中小学教师对教学视频使用意愿研究

——基于整合的 TPB、TAM 和 UTAUT 模型

A Study of Primary and Secondary School Teachers' Willingness to Use Instructional Videos ——The Integration-Based TPB, TAM, and UTAUT Models

李妍[1]，何静[1]，章仪[2]，刘彩霞[1]，杨九民[1*]
[1]华中师范大学人工智能教育学部
[2]北京师范大学未来教育学院
[*]yjm@mail.ccnu.edu.cn

【摘要】 为探究中小学教师对教学视频使用意愿的影响机制，本文整合技术接受模型、计划行为理论和技术接受与使用统一模型建立概念模型。使用结构方程模型对收集到的 464 份问卷进行分析，研究结果发现：①感知有用性、感知易用性和主观规范均对使用态度产生显著正向影响，且均通过使用态度对使用意愿产生显著正向影响，但便利条件对使用态度产生显著负向影响。②主观规范和使用态度均对使用意愿产生显著正向影响，但便利条件对使用意愿并无显著的直接影响。研究结果有助于为中小学教育信息化和教师培训的设计和实施提供建议和参考。

【关键词】 中小学教师；教学视频；技术接受模型；计划行为理论；技术接受与使用统一模型

Abstract: *To explore the influence mechanism of primary and secondary school teachers' willingness to use instructional videos, this paper integrates technology acceptance model, theory of planned behavior, and unified theory of acceptance and use of technology model to establish a conceptual model. Analyzing 464 questionnaires collected using structural equation model, the research results found that: ① Perceived usefulness, perceived ease of use, and subjective norm all have significant positive effects on attitudes toward usage, and all have significant positive effects on behavioral intention to use through attitudes toward usage, but facilitating conditions have significant negative effects on attitudes toward usage. ② Subjective norms and usage attitudes have significant positive effects on behavioral intention to use, but facilitating conditions have no significant direct impact on behavioral intention to use. The research results provide valuable suggestions and references for the design and implementation of educational informatization and teacher training in primary and secondary schools.*

Keywords: Primary and secondary school teachers, Instructional video, TAM; TPB, UTAUT

1. 引言

党的二十大报告中明确提出“推进教育数字化，建设全民终身学习的学习型社会、学习型大国”。习近平总书记也指出：“教育数字化是我国开辟教育发展新赛道和塑造教育发展新优势的重要突破口”，深刻揭示了教育数字化的关键作用。如何落实国家教育数字化战略行动，优质数字技术资源的建设与使用是关键。教学视频作为一种核心的数字技术资源，因其能够结合视听双通道更生动、直观、形象地呈现教学材料，并可打破空间、时间的限制，极大地扩展了教学的灵活性而备受人们的关注（杨九民等, 2022; 尤洋等, 2020）。目前的教学视频研

究主要关注于教学视频的设计和应用方面，例如教学视频的设计、实践中的教学效果以及影响教学有效性的因素等（杨九民等, 2023; 王雪等, 2022），而从教师角度探究教学视频使用意愿及其影响因素的实证研究较少，忽略了教师作为教育教学的主导者对教学视频的使用意愿的重要性。

后疫情时代，使用教学视频进行教学已经成为教育形式的新常态（郑茜茹等, 2023），但在实际教学中，教师在决定课堂上是否使用技术方面发挥着重要作用（Scherer et al., 2019），且仍存在教师抵触使用技术的现象，甚至出现“计算机恐惧症”（陈奕延等, 2022）。教师对待技术的态度、外在阻碍等均会影响到技术的使用效果（伏蓉等, 2020）。例如教师在课堂上使用技术的态度会影响成功使用技术的程度（Thapaliya et al., 2023），难以检索到合适的视频（boclips, 2019）、视频搜索功能的局限性(Ekstrand et al., 2020）会影响教师使用教学视频的积极性。且根据技术接受模型（TAM），教师直接使用技术进行教学的行为意图会决定其在日常教学中使用技术的程度（Davis, 1989）。因此，教师是否愿意接受并运用教学视频，以及哪些因素会影响教师对教学视频的接受度和使用行为，仍然是一个值得深入探讨的问题。然而，TAM 模型主要侧重于外部动机和心理因素的探究，却忽略了内在动机的重要性（Taherdoost, 2018）。因此，在本文中将引入计划行为理论（TPB）和技术接受与使用统一理论（UTAUT），以填补 TAM 模型的不足之处，并更全面地探究内外部因素对教学视频使用意愿的影响。此外，基于对教师教学视频的使用意愿和影响因素的探索，我们能够更好地理解教师的行为意向和决策过程，从而为教育行政部门和学校制定相关的教师培训和支持策略提供指导，具有重要的现实意义。

本文拟通过问卷调查的方式，以技术接受理论、计划行为理论和技术接受与使用统一理论为基础，构建影响中小学教师对教学视频使用意愿的概念模型，验证感知易用性、感知有用性、主观规范等五个解释变量对中小学教师教学视频使用意愿的影响假设，以教师角度探索提高教学视频使用意愿的关键影响因素。

2. 理论基础

2.1 技术接受模型（TAM）

技术接受模型（Technology Acceptance Model, TAM）是 Davis 等人（1989）基于理性行为理论和计划行为理论提出的，该模型的两个主要前因变量为感知易用性（PEU）和感知有用性（PU）（Marangunić et al., 2015）。行为态度是指使用者对使用技术的积极或消极感受的程度（Li et al., 2019）。在本研究中，感知易用性定义为教师感知使用教学视频的难易程度，感知有用性指教师认为使用教学视频能够提高其工作绩效的程度，行为态度指教师对使用教学视频的积极或消极态度。TAM 模型主要观点为使用者对技术的感知易用性和感知有用性共同影响使用者使用技术的态度，感知有用性和态度共同决定使用者的行为意向，而感知有用性受到感知易用性和外部变量影响（Davis, 1989）。在教育研究领域，关于 TAM 模型的已有研究多集中于对教育游戏、在线教学、在线学习、信息化教学等信息化技术的使用意愿研究，且研究结果较为统一，即技术使用态度会受到感知易用性和感知有用性的影响，且感知有用性是最主要的决定性因素（Yeo et al., 2022; 孙爱晶等, 2023）。但以往研究的研究对象多集中于高校师生，对中小学教师的关注较少。基于此，本文提出以下假设：

H1：感知易用性与对教学视频的感知有用性呈正相关。

H2：感知易用性与对教学视频的使用态度呈正相关。

H3：感知有用性与对教学视频的使用态度呈正相关。

H7：使用态度与对教学视频的使用行为意向正相关。

2.2 计划行为理论（TPB）

计划行为理论（Theory of Planned Behavior, TPB）是 Ajzen（1991）在理性行为理论的基础提出的，该理论认为主体的行为意愿会受到主观规范、行为态度和感知行为控制的共同影响。主观规范是指在某一特定情境下，个体对于自己在意的相关群体对其的期望所产生的感知程度，在本研究中主观规范定义为教师感知到的重要他人对其使用教学视频的期望程度。在 TAM 模型中，其重点强调个体的认知信念对技术使用态度和行为意向的内在影响，却忽略了主观规范的影响（宋昊泽等, 2023）。因此在本研究中，将主观规范因素纳入概念模型中。在关于主观规范的研究中，Teo 等人（2008）的研究结果发现主观规范是决定行为态度的重要决定因素之一，也是感知有用性的积极预测因素。魏叶美和范国睿（2021）基于 TPB 框架进行分析发现主观规范能显著影响教师的行为意向。基于已有研究，本文提出如下假设：

H4：主观规范与对教学视频的感知有用性呈正相关。

H5：主观规范与对教学视频的使用态度呈正相关。

H6：主观规范与对教学视频的使用行为意向呈正相关。

2.3 技术接受与使用统一模型（UTAUT）

技术接受与统一使用模型（Unified Theory of Acceptance and Use of Technology Model, UTAUT）是 Venkatesh 等人（2003）在技术接受理论、计划行为理论等理论模型基础上提出的，该模型常用于信息技术接受度评测，由四个核心变量和四个调节因子组成。便利条件是其中的一个核心变量，指使用者认为现有条件支撑技术使用的的程度（van Twillert et al., 2020）。在本研究中，便利条件指教师认为现有条件支撑使用教学视频的程度。

由于 TAM 模型侧重于个体对技术的主观能动感知，忽略了外部环境和使用者的个体特征对技术接受的影响（Dishaw&Strong, 1999）。因此在本研究中将便利条件纳入概念模型。以往研究指出，便利条件和基础设施会影响教师对技术使用的态度（van Twillertet al., 2020），便利条件对行为意向具有显著的正向影响（黄玫, 2022），同时对感知易用性有显著影响（Ranellucci, 2020）。基于此，提出如下假设：

H8：便利条件与对教学视频的感知易用性呈正相关。

H9：便利条件与对教学视频的使用态度呈正相关。

H10：便利条件与对教学视频的使用行为意向呈正相关。

基于以上假设，本研究整合技术接受模型、计划行为理论和技术接受与使用统一模型，构建影响中小学教师对教学视频使用意愿的概念模型，如图 1 所示。

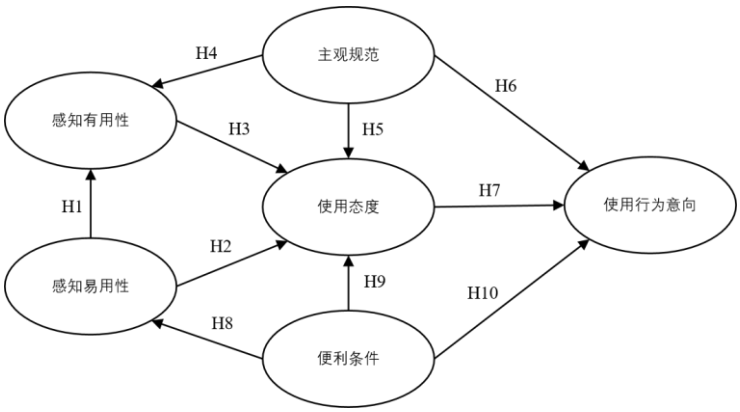


图 1 中小学教师对教学视频使用意愿概念模型

3. 研究设计

3.1 研究工具

本研究使用调查问卷收集数据。问卷由两个部分组成，第一部分为人口统计学信息收集，收集信息包括性别、教学阶段、年龄和教学经验。第二部分是对 6 个潜变量的测量，均改编自前人量表，包括感知有用性（Daviset al., 1989）、感知易用性（Daviset al., 1989）、主观规范（Ajzen, 1985）、便利条件（Thompson et al., 1991）、使用态度（Compeau & Higgins, 1995）和使用行为意向（Daviset al., 1989），采用五点李克特量表，从 1 分（“非常不同意”）到 5 分（“非常同意”）。

3.2 样本数据情况

样本选取来自全国 500 名中小学教师，调查时间为 2022 年 8 月 3 日至 11 月 2 日，历时近三个半月。在排除不可靠的问卷后，本文共使用 464 份有效问卷进行数据分析。从样本的描述性统计可知，教师的平均年龄为 M=37.55（SD=7.081），平均教龄为 M=14.55（SD=8.058），54%为女性。在教学阶段方面，共计小学教师 114 人（24.6%），初中教师 230 人（49.6%），高中教师 120 人（25.9%）。在教授学科方面，共计信息技术 105 人（22.6%）、语文 95 人（20.5%）、数学 82 人（17.7%）、英语 42 人（9.1%）、化学 35 人（7.5%）、物理 30 人（6.5%）、地理 28 人（6.0%）、生物 24 人（5.2%）、道德与法治 14 人（3.0%）、美术 9 人（1.9%）。

3.3 数据分析方法

本文采用偏最小二乘法构建结构方程模型（PLS-SEM），并使用 SmartPLS3.0 分析影响中小学教师对教学视频使用意愿的概念模型，以验证教师感知有用性、感知易用性、主观规范、便利条件和使用态度之间的关系，以及这些关键解释变量与教师使用教学视频的行为意向之间的关系。PLS-SEM 在探索性研究中具有显著的优势，其特点体现在能够处理小样本量数据，对样本分布要求宽松，同时能够求解更为复杂的结构方程模型（Hair et al., 2011）。

4. 研究结果

4.1 信效度分析

为保证概念模型的信效度，本研究采用 Cronbach's α 系数和组合信度 CR 值测量问卷量表的内在信度。如表 1 所示，Cronbach's α（范围：0.791-0.956）和组合信度（范围：0.878-0.972）均高于标准值（0.7），表明测量模型的各量表信度较高。效度指测量题项有效衡量目标对象的程度，一般包括收敛效度和区别效度。本研究主要采用因子载荷系数和平均方差萃取量（AVE）来评价收敛效度，使用 Fornell-Larcker 标准检验区别效度。所有的指标项标准化因子载荷系数（范围：0.777-0.971）均高于标准值（0.7），AVE（范围：0.707-0.919）高于标准值（0.5），即所有变量都具有较好的收敛效度。关于区分效度，通过对比各潜变量的平均方差萃取量（AVE）的平方根和该潜变量本身与其他潜变量相关系数可知（见表 2），AVE 的平方根均大于潜变量本身和其他潜变量相关系数，即本概念模型各变量之间具有较好的区分效度。

表 1 测量模型的信度和效度检验结果

潜变量	代码	因子载荷	AVE 值	CR 值	Cronbach's α 值
感知有用性（PU）	PU01	0.942	0.919	0.972	0.956

	PU02	0.971			
	PU03	0.964			
	PEU01	0.777			
感知易用性（PEU）	PEU02	0.899	0.707	0.878	0.791
	PEU03	0.842			
主观规范（SN）	SN01	0.954	0.914	0.955	0.906
	SN02	0.958			
便利条件（FC）	FC01	0.961	0.908	0.952	0.899
	FC02	0.944			
使用态度（ATU）	ATU01	0.93	0.864	0.927	0.842
	ATU02	0.929			
使用意愿（BIU）	BIU01	0.957	0.914	0.955	0.905
	BIU02	0.955			

表 2 区分效度的检验结果

变量	ATU	BIU	FC	PEU	PU	SN
ATU	0.929					
BIU	0.770	0.956				
FC	0.251	0.269	0.953			
PEU	0.325	0.323	0.399	0.841		
PU	0.342	0.327	0.131	0.272	0.959	
SN	0.600	0.597	0.514	0.366	0.343	0.956

注：对角线数值是平均方差萃取量(AVE)的平方根；对角线左下方数值是相应潜变量之间的相关系数

4.2 结构模型的假设检验

对中小学教师在教学中使用视频意愿的概念模型进行检验分析，得出各假设路径的标准化系数、*p* 值（见表 3）和模型图（如图 2）。根据检验结果可知：感知易用性对教学视频的感知有用性和使用态度均有显著的预测作用。感知有用性对使用态度均有显著的积极影响。因此，假设 H1、H2 和 H3 均得到支持。主观规范对感知有用性、使用态度和使用行为意向有显著的正向影响。此外，教师的使用态度对教学视频的使用行为意向也有显著影响。因此，结果支持假设 H4、H5、H6 和 H7。便利条件对感知易用性有显著影响，但便利条件对使用态度呈显著负向影响，对使用行为意向无显著影响，支持假设 H8，不支持假设 H9、H10。

表 3 模型参数检验值及研究假设检验

研究假设	假设路径	路径系数	t 值	检验结果
H1	PEU→PU	0.169*	2.510	支持假设
H2	PEU→ATU	0.126**	2.562	支持假设
H3	PU→ATU	0.127**	2.685	支持假设
H4	SN→PU	0.281***	4.203	支持假设
H5	SN→ATU	0.565***	11.787	支持假设
H6	SN→BIU	0.211***	4.740	支持假设

H7	ATU→BIU	0.644***	16.377	支持假设
H8	FC→PEU	0.399***	8.288	支持假设
H9	FC→ATU	-0.106*	2.221	拒绝假设
H10	FC→BIU	-0.001	0.026	拒绝假设

注：*$p < 0.05$, **p < 0.01, ***$p < 0.001$.

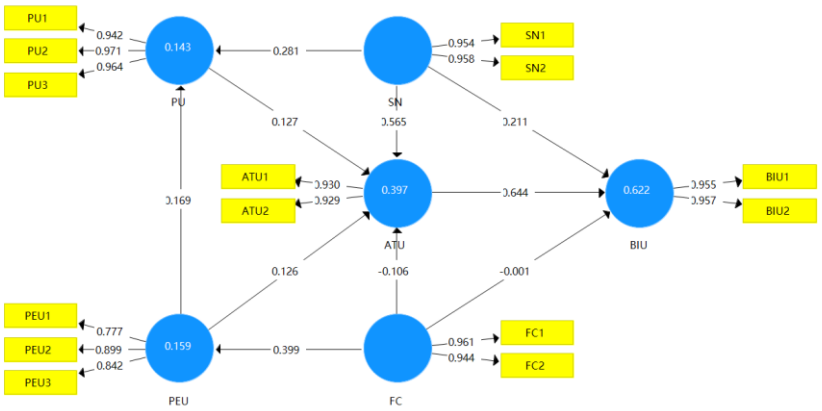


图 2 结构模型的运行结果

5. 讨论与结论

5.1 感知有用性、感知易用性和使用态度与教学视频使用意愿之间的关系

研究结果表明，感知有用性和感知易用性均与使用态度显著正相关，感知易用性对感知有用性具有正向的预测作用，且感知有用性和感知易用性通过使用态度间接正向影响教学视频使用意愿，验证假设 H1、H2、H3。使用态度对教学视频使用意愿有显著的正向影响，验证假设 H7。以上研究结果与 Adanır 等人（2022）和 Bajaj 等人(2021)的研究一致，当使用者了解使用教学视频的好处并主动想要使用教学视频时，其对教学视频的利用率会进一步提升。而感知有用性侧重于通过技术提高工作绩效和效率（Venkatesh et al., 2003）。这说明中小学教师越容易得到所需的教学视频，便越认为教学视频能够提高其工作效率，进一步提升对使用教学视频的积极态度，使其对教学视频的使用意愿越强。

5.2 主观规范与教学视频使用意愿之间的关系

主观规范对感知有用性、使用态度和使用行为意向均具有显著的直接正向影响，验证假设 H4、H5、H6。这一研究结果与 Venkatesh 等人（2003）和 Teo 等人（2008）的研究观点相一致，说明中小学教师认为教学视频的有用性、对教学视频的使用态度和使用意愿均受到权威人士和其他重要群体的影响。可能是因为目前教学视频的应用仍是自上而下地推广，中小学教师更多的是在学校和政策监督下使用教学视频，教师个人对教学视频使用的主观意愿不强，受外界影响较大。因此在推广教学视频使用过程中，可以通过完善政策要求和考评措施的方式，促使中小学教师使用教学视频。

5.3 便利条件与教学视频使用意愿之间的关系

研究结果表明，便利条件对感知易用性具有显著的直接正向影响，对使用态度有显著的负向影响，与使用行为意向无显著相关性，验证假设 H8，不符合假设 H9、H10。该研究结果与 Rahiman 和 Kodikal（2024）不同，而与 Teo 学者（2008）的研究一致，便利条件并不直接影响使用技术的态度，而是通过感知的易用性间接影响使用技术的态度。本研究结果同时发

现便利条件通过感知易用性间接正向影响使用态度，说明对于中小学教师而言，使用教学视频的态度并不直接受到现有条件的影响，教师的主观能动性较强。当具备足够的外部条件促使教师意识到教学视频的有用性时，教师会主动增强对教学视频使用的积极性并增强对教学视频的使用意向。

因此，为促进中小学教师教学视频的使用意愿，应在保证外部支持充足的情况下，让教师能够较为便利地获取所需要的教学视频。外部支持可以是教学视频资源平台的建设（例如我国已建立的中小学智慧教育平台），和对教学视频资源搜寻和制作培训的外部支持。通过技术的支持，提高教师对教学视频使用的积极性，激发教师使用教学视频的内生动力，从而达到提升教师对教学视频的使用意愿的目的。

参考文献

陈奕延 李晔.(2022).人工智能技术恐惧症的定义、诱因、衡量及克服路径研究.计算机应用与软件(12),23-33+63.

冯吉兵,张国良,靳帅贞,王俊,钟正,谭婷 邓雅心.(2022).中小学教师虚拟实验教学能力提升的对策研究——基于技术接受度测评.中国电化教育(07),120-126+133.

伏蓉,包磊,孙淑艳,菅光宾,武咏梅 徐万东.(2020).基础教育信息化应用研究与案例.中国电化教育(10),121-133.

黄玫.(2022).中小学教师信息技术接受度影响因素研究——以陕西省为例.教育信息技术(03),18-21.

宋昊泽,赵婧 洪阳.(2023).高校师范生 MOOC 学习意愿研究——基于整合的 TPB 模型和 TAM 模型.天津师范大学学报(社会科学版)(01),74-80.

孙爱晶,范九伦 张二锋.(2023).教育数字化背景下高校教师在线教学持续使用意愿的实证研究——以 S 省为例.高教学刊(27),97-100.

王雪,张蕾,王鉴羽,乔玉飞 李慧洋.(2022).弹幕教学视频中学习者的眼动行为模式及其作用机制研究.远程教育杂志(05),103-112.

魏叶美 范国睿.(2021).教师参与学校治理意愿影响因素的实证研究——计划行为理论框架下的分析.华东师范大学学报(教育科学版)(04),73-82.

杨九民,陈辉,杨文蝶,李妍,喻邱晨 皮忠玲.(2022).教学视频中的字幕应该如何设计？——基于 2000—2021 年 44 篇实验和准实验研究文献的元分析.电化教育研究(08),59-67.

杨九民,刘彩霞,陈楼琪 皮忠玲.(2023).教学视频中的描述性手势何以有效——手势学习策略及学习者空间能力的影响.现代远程教育研究(04),92-101.

尤洋,王以宁 张海.(2020).智慧课堂环境下教学视频复杂度与学习者认知负荷关系研究.现代远距离教育(02),91-96.

郑茜茹,胡萍,冯小燕 张钰川.(2023).“互联网+”时代下教学视频的研究现状与趋势.电脑知识与技术(23),122-125.

Adanır, G. A., Muhametjanova, G., Borkoev, B., & Salieva, K. (2022). University learners' utilisation of online videos in a general chemistry course. *Research in Learning Technology, 30*.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179-211.

Bajaj, P., Khan, A., Tabash, M., & Anagreh, S. (2021). Teachers' intention to continue the use of

online teaching tools post Covid-19, *Cogent Education, 8*(1).

Boclips (2019). Access Denied: Teachers Struggle to Access Safe Video Content. Link: www.boclips.com/teachers-struggle-to-access-video-whitepaper

Compeau, D.R., & Higgins, C.A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, *19*(2), 189–211.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340.

Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task–technology fit constructs. *Information & management*, *36*(1), 9-21.

Ekstrand, M. D., Wright, K. L., & Pera, M. S. (2020). Enhancing classroom instruction with online news. *Aslib Journal of Information Management*, *72*(5), 725-744.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139-152.

Li, Y., Wang, Q., & Lei, J. (2019). Modeling Chinese Teachers' attitudes toward using Technology for Teaching with a SEM approach. *Computers in the Schools*, *36*(2), 122-141.

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal access in the information society*, *14*(1), 81-95.

Rahiman, H. U., & Kodikal, R. (2024). Revolutionizing education: Artificial intelligence empowered learning in higher education. *Cogent Education*, *11*(1), 2293431.

Ranellucci, J., Rosenberg, J. M., & Poitras, E. G. (2020). Exploring pre-service teachers' use of technology: The technology acceptance model and expectancy-value theory. *Journal of Computer Assisted Learning*, *36*(6), 810-824.

Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, *128*, 13–35.

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia manufacturing*, *22*, 960-967.

Teo, T., Lee, C. B., & Chai, C. S. (2008). Understanding pre-service teachers' computer attitudes: applying and extending the technology acceptance model. *Journal of computer assisted learning*, *24*(2), 128-143.

Thapaliya, M., Adhikari, S., & Rana, L. (2024). Opportunity in COVID-19 crisis: Moving away from chalk and talk to technology-integrated teaching in Nepalese higher education institutions. *E-Learning and Digital Media, 21*(1), 87-105.

Thompson, R.L., Higgins, C.A., & Howell, J.M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, *15*(1), 124–143.

van Twillert, A., Kreijns, K., Vermeulen, M., & Evers, A. (2020). Teachers' beliefs to integrate Web 2.0 technology in their pedagogy and their influence on attitude, perceived norms, and perceived behavior control. *International Journal of Educational Research Open*, *1*, 100014.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, *27*, 425-478.

Yeo, S., Rutherford, T., & Campbell, T. (2022). Understanding elementary mathematics teachers' intention to use a digital game through the technology acceptance model. *Education and Information Technologies*, *27*(8), 11515-11536.

基于多模态同伴反馈的同伴教学在师范生专业发展上的效果研究

Effects of multi-peer feedback-based peer instruction on pre-service teachers' professional development

何晓玲[1]，徐刘杰[2]，伍婧[1]

[1]中国杭州浙江工业大学教育科学与技术学院

[2]中国芜湖安徽师范大学教育科学学院

xulj2004@126.com

【摘要】 促进师范生专业发展是每个国家的重中之重，因此有大量文章研究有哪些因素能够提高师范生专业能力。本文采用准实验研究方法，对比传统同伴教学，研究基于多模态同伴反馈的同伴教学在师范生专业发展上的效果。结果表明多模态同伴反馈支持下的同伴教学不仅能够提高师范生的 TPACK 和批判性能力，还能够缩小不同水平师范生之间在专业发展方面的差距。本文为国家如何培养师范生能力提供了一定的参考，也为研究者后续研究师范生专业发展提供了思路。

【关键词】 师范生；专业发展；同伴教学；同伴反馈；多模态

Abstract: *Promoting the professional development of teacher trainees is a top priority in every country, so there are a large number of articles examining what factors can improve the professional competence of teacher trainees. This paper adopts a quasi-experimental research method to investigate the effect of peer teaching based on multimodal peer feedback on the professional development of teacher trainees in comparison with traditional peer teaching. The results show that peer teaching supported by multimodal peer feedback not only improves teacher trainees' TPACK and critical competence, but also reduces the gap in professional development between teacher trainees of different levels.*

Keywords: Teacher educators,professional development,peer teaching,peer feedback,multimodality

1.前言

教师是培养未来人才的关键，师范生是建设优质师资队伍的主力，是实现教育高质量发展的必要力量。无论是国外还是国内都对师范生专业发展十分重视。例如：中国教育部等八部门在 2022 年印发了《新时代基础教育强师计划》的通知中提出要开展国家教师队伍建设改革试点来培养师范生和促进教师专业发展。印度在《2020 年新教育计划》中强调了教师专业发展的重要性，要努力将理论知识转向实践过程。美国在 2020 年 1 月 30 日的《REL Pacific》中提出要支持帕劳教师的成长和发展，帕劳教育者集体审查教育系统来确保教师成长的成功。美国教育部 SSED 在 2022 年为支持教师专业学习拨款 6500 万美元。法国教育部发布《2023—2027 年教育数字化战略》的报告中提出要加强教师培训，开发 Pix 来支持教师技能发展。

国家、组织和地方政策以及反思和制定是促使师范生专业持续发展的动力源泉，但目前师范生并不能制定出促进所有学生学习和发展的实践和策略，归其原因是因为师范生学习的理论知识与缺少课堂实践之间的矛盾(Chaaban，Du，Lundberg，& Abu—Tineh，2023；Romijn，Slot，& Leseman，2021)。实践知识能够促使教师专业有效发展，师范生大多利用同伴教学的实践过程

来反思获得相关知识。提供同伴反馈的能力是师范生成为教师角色的要求之一，该能力能够使他们的专业实践知识得到持续发展。但目前的大多研究要么只研究同伴教学在教和学过程中的作用，要么只研究同伴反馈在实践过程中发挥的作用，又或者虽研究了同伴教学和同伴反馈对教师的促进作用，但促进方面是体现在教师的自我效能感上，并没有明确研究同伴教学和同伴反馈对师范生专业发展的促进作用。

多模态数据能够提高评价数据的价值，因此本文对基于多模态同伴反馈的同伴教学对师范生专业发展的效果通过准实验和对比研究的方式展开研究。本研究的目的是探究在实施同伴教学的基础上，引入多模态的同伴反馈对师范生专业发展的影响。本研究包括以下三个研究问题。

问题 1：与传统的同伴教学相比，基于多模态同伴评价的同伴教学是否能够促进师范生的 TPACK 能力？

问题 2：与传统的同伴教学相比，基于多模态同伴评价的同伴教学是否能够促进师范生的批判性思维能力？

问题 3：与传统的同伴教学相比，基于多模态同伴评价的同伴教学是否能够缩小不同水平师范生之间在专业发展（TPACK 和批判性思维）方面的差距？

2.综述

2.1.师范生专业发展

国家发展，根本在教师，教师专业发展被认为是各种教育改革的基石，其重要性不言而喻(Chai，2019)。衡量教师专业发展框架有 TPACK，其包括教学内容知识、技术教学知识和技术教学内容知识三部分，该框架对教师教育的专业发展具有重大影响(Xie，Kim，Cheng，& Luthy，2017)。批判性思维是高阶思维能力的一部分，也是衡量学生学习质量的标准之一，教师的职责是教书育人，因此教师需要有批判性思维，来帮助学生和促进社会发展。提高教师专业发展是刻不容缓的事，师范生会成为未来的大部分教师，现在培养师范生的专业能力会改善一些未来教师在专业发展上出现的问题。但目前在培养师范生上还存在一系列的问题，比如只注重实践培训的静态发展、培训方式采用千篇一律的方法进行等。

那么如何更好的培养师范生？一个人的力量是有限的，有时并不能发现自身教学存在的问题，若每位师范生都能够得到专家或同伴持续的指点，师范生的专业能力将会得到稳步提升(徐晓东，李王伟 & 曹鑫，2021）。同伴教学可以实现师范生以互教和互评等方式的实践活动来充分展示自己的教学方式，让自己的理论知识得到实践，且通过与同伴之间的有效互动，能够实现有意义的建构知识(Hamlen Mansour，Jackson，Bievenue，Voight，& Sridhar，2023)。

2.2.同伴教学

同伴教学是指学习同伴之间相互教学的过程，利用这种教学模式可以帮助学习者更好地理解和掌握新知识。实施同伴教学主要包括以下四个阶段，第一，为学习创造激励条件；第二，通过创建与学生相关的学习材料来引入初始激励技巧；第三，提供刺激性活动和培养学生之间的自尊、自信和合作来保持学生的积极性；第四，将评估和反馈转化为积极的体验(Rodis & Locsin，2019)。同伴教学有助于学习者在认知、情感和社会领域的发展，以他们的视角来进行教学，提升他们的学习和领导能力，在这种更具有包容性的学习环境下，充分尊重学习者的能动性，培养学习者批判性思维(Maxson，Neely，Roberts，Stone，Lowe，Macy，& Miller，2019)。

同时同伴教学也能够促进教师专业发展，它能够帮助师范生积极理解概念知识，在师范生进行同伴教学时，积极的反馈和评论不仅能够改善他们的技能，还能够提高他们的自信心和激励他们的教学热情(Alsaleh，Alabdulhadi，& Alrwaished，2017；Al—Hebaishi，2017)。但目前同伴教学在应用方面存在一定的问题，大多数教师没有进行培训同伴教学的相关知识，在课堂上教

学时没有使用同伴教学的意识；且课堂教学时间有限，进而使得使用同伴教学的时间非常有限(Alsaleh，Alabdulhadi，& Alrwaished，2017)。且就算进行同伴教学实践，也缺乏立即的质疑与分析，会影响到同伴教学的效果，因而需要在观察之后立即提出反馈，而同伴指导具有时效性的作用(Lowman，2016)。

2.3.同伴反馈

同伴反馈是指学习者对同伴学习状态的一种评价模式，主要从同伴的文章内容、组织架构和语言层面来提供反馈（黄爱琼，2020）。同伴反馈的实施过程主要包括三个阶段，第一：观察同伴表现；第二：对同伴表现的优点或缺点进行反馈；第三：被反馈的同伴反思提供的信息(Prilop & Weber，2023)。同伴反馈被认为是教师获得教学专业知识的关键因素，教师可以通过同伴者在对自己教学实践过程中提供的反馈的基础上进行反思去促进成长(Prilop，Weber，& Kleinknecht，2020)。反馈是教师专业能力的其中之一，也被认为是促进各领域发展的重要因素之一，反馈能够影响师范生的专业知识、实践和自我效能感，高质量的同伴反馈是教师专业发展所需要的(Prilop，Weber，Prins，& Kleinknecht，2021)。

数据应用于教育是必然走向，不过在教学应用中会出现单一数据的问题，从而容易给教与学过程带来误导作用。而多模态数据能够很好的解决这个问题，对于学生而言，多模态支持的同伴反馈能够直接影响到学生对自己当前不足修改和吸收当前学习的知识的情况(田丽丽& 周羽，2022)。对于教师来说，教师可以在课堂上利用同伴反馈过程中产生的数据，了解学生掌握当前课堂教学的哪些知识，以及哪些是未掌握的知识，从而更好进行当前教学 (Visscher，2021)；在课后对同伴反馈产生的数据进行分析，能够得到反思与教学问题的改进，进而优化教学实践(王阿习，余胜泉，湛志强 & 刘秀英，2022)。

3.研究设计

3.1.基于多模态同伴反馈的同伴教学环境

本研究借助智慧教研和学习元平台实现多模态同伴反馈的同伴教学环境。在课前，授课师范生根据教师布置的主题利用学习元平台上的微视频和资料来进行自学，设计出与主题相应的教学设计并上传到智慧教研平台，其他师范生和教师查看授课师范生的教学设计并提出自己的疑问，授课师范生与他们积极讨论，在这个阶段，平台会根据评论和讨论数据来分析每位师范生的学习和参与情况。师范生参与度和贡献越高，以及有效评论和讨论越多，则越靠近内圈，如下图 1 参与水平图所示。

在课中，录制授课师范生在课堂上的讲课视频并上传平台，其他师范生听授课师范生讲课并进行评价，在此阶段，平台不仅会对授课师范生和其他听课师范生进行情绪识别，还会分析同伴反馈的文本、语言、视频和面部表情的多模态数据，并将它们以可视化的结果反馈给授课师范生。如下图 2 情感评价图所示，该图明确表达了听课师范生的听课情感状态，在哪个时间段高兴、质疑、赞同和失望，以及听课师范生情感状态的占比。

在课后，授课师范生根据平台反馈的结果进行反思，修改本次上课的教学设计并再次上传。



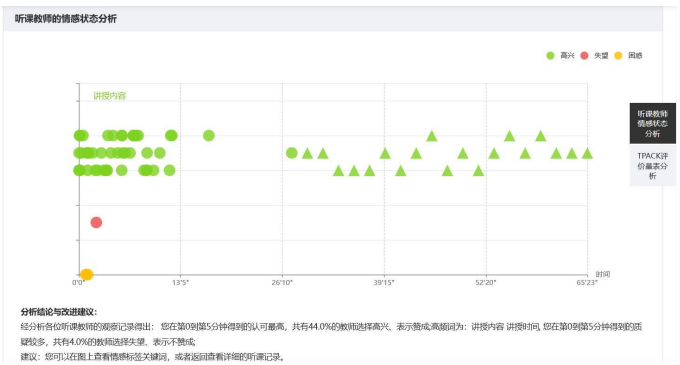


图 1 参与水平图　　图 2 情感评价图

3.2.研究对象

本实验研究对象来自某高校教育技术学专业 21 级班级的 68 名本科师范生，平均年龄 21 岁，其中男生 46 名，女生 22 名，随机将他们分为实验组和控制组。其中实验组有 41 名学生（男生 24 名，女生 17 名），采用多模态同伴评价的同伴教学方式，控制组有 27 名学生（男生 22 名，女生 5 名），使用传统同伴教学方式。

3.3.研究过程

整个实验流程图如下图 3 所示。其中在采用基于多模态同伴评价的同伴教学方式学习过程中主要包括授课者开展同伴教学、师生进行同伴反馈、评价者利用技术工具进行评价、授课者查看多模态可视化反馈结果、师生基于反馈结果进行讨论交流和反思以及教师总结这六个部分。采用传统同伴教学包括授课者开展同伴教学、评价者通过授课者的表现进行打分、师生基于打分结果开展讨论交流和反思以及教师进行总结这四部分。

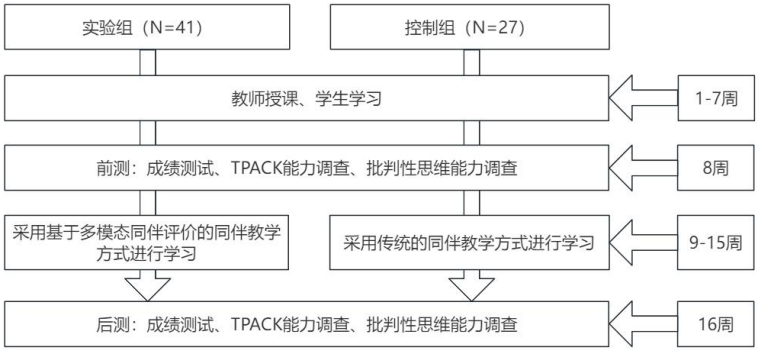


图 3 实验流程图

3.4.研究工具

本实验研究工具有三个，分别是 TPACK skills 问卷、critical thinking 问卷和学业成绩测试题。前两个问卷分别改编自 Schmid 等人的 TPACK skills 问卷(Schmid，Brianza，& Petko，2020) 和 Sosu 等人的 critical thinking 问卷(Sosu，2013)。测得 TPACK skills 问卷和 critical thinking 问卷的克朗巴赫 α 值分别为 0.958 和 0.889，问卷均采用采用五点李克特量表计分法（1 =强烈不同意；5=强烈同意）。

3.5.数据采集与分析

研究数据包括三方面的数据，第一部分是利用 TPACK skills 问卷测试师范生 TPACK 能力的数据，第二部分是利用 critical thinking 问卷测试师范生批判性思维能力的数据，第三部分是利用学业成绩测试题去测试师范生的成绩。根据所得到的数据，利用多次 t 检验的方式来分析在不同条件下师范生的 TPACK 能力、批判性思维能力以及学业成绩之间的差异。

这项研究使用定量研究和对比研究方法来回答我们提出的研究问题。对于第一个问题，本文研究的重点是对比实验前后师范生的 TPACK 的能力是否得到了提升，即师范生采用多模态同伴教学的方式能不能促进 TPACK 能力，利用 t 检验来分析第一个问题中实验组和控制组的数据。问题二采用独立样本 t 检验和单因素协方差方法进行分析处理。问题三使用 t 检验分析不同水平师范生专业发展之间的差距。

4.结果与讨论

4.1.职前教师 TPACK 能力

首先对实验组和控制组的前测数据进行独立样本 t 检验，结果显示实验组（M=14.54，SD=2.560）和控制组（M=14.74，SD=2.474）没有显著差异（p=0.886>0.05），表明实验组

和控制组的起始 TPACK 能力几乎没有太大差异。再次使用独立样本 t 检验，对两组的后测数据进行分析，结果如下表 1 所示，根据结果 p=0.049<0.05 可知基于多模态同伴评价的同伴教学与传统同伴教学相比更能够促进师范生 TPACK 能力。

TPACK 由 TK、CK、PK、TCK、TPK 和 PCK 组成，是一个教师专业知识框架。TPACK 提出的目的是为了帮助教师把信息技术与教学实践有机结合，进而改善教学效果，从而提高学习者的成绩。因此研究如何提高师范生的 TPACK 能力是有价值的，本研究在对比实验中发现，师范生的 TPACK 能力是可以通过多模态同伴评价的同伴教学方式来提高。

表 1 职前教师 TPACK 能力的独立样本 t 检验分析结果

分组	样本	平均值	标准差	*SE*	*F*	*p*	*T*
实验组	41	113.02	12.02	3.66	0.097	0.049	2.00
控制组	27	105.66	18.21	3.97			

4.2.职前教师的批判性思维能力

分别对实验组和控制组的评价思维和反思思维的前测数据使用独立样本 t 检验，结果如表 2 所示，实验组和控制组的评价思维的前测数据没有显著差异（p=0.811>0.05），不过反思思维的前测数据有显著差异（p=0.015<0.05）。为了进一步研究实验组与控制组后测数据之间的差异，在验证满足回归齐性假设后（F=2.55，p=0.086>0.05）后，应用单因素协方差方法，将前测数据作为协变量，后测数据作为因变量进一步验证两组数据之间的区别。结果如下表 3 所示，由 F=5.41，p=0.023<0.05 数据可知，师范生基于多模态同伴评价的同伴教学的批判性思维能力显著高于采用传统的同伴教学方式。

批判性思维能力包括评价思维和反思思维两部分，是决定怎么想、怎么做和怎么样的过程。孔子曾说学而不思则罔，思而不学则殆。曾子也曾说吾日三省吾身，充分证明了人进行批判性思考的重要性。反思和评价一直贯穿教师的整个教学过程，知识是学无止境的，教师在备课和上课过程中会遇到新的知识，这时候教师就需要反思这个知识是否正确、是否与现阶段教学相适以及该怎样将该知识和谐的应用，且对学习者的诊断性评价、形成性评价和总结性评价是个动态过程，是需要教师时时刻刻进行评判的。在师范生进行同伴教学的实践过程中，加入同伴反馈过程，充分培养师范生的反思和评价思维能力，进而提高他们的批判性思维能力。

表 2 职前教师批判性思维能力的描述性数据

分组	维度	样本	前测平均值	前测标准差	后测平均值	后测标准差
实验组	CO	41	26.39	4.352	29.43	2.88
	RS	41	15.63	2.88	16.43	1.74
控制组	CO	27	25.89	4.108	27.07	4.99
	RS	27	15.00	2.68	15.48	2.79

表 3 职前教师批判性思维能力的单因素协方差分析结果

变异来源	平方和	*df*	均方	*F*	*p*	η^2
协变量1 (PreCO)	166.45	1	166.45	6.39	0.014	0.091
协变量2 (PreRS)	2.757	1	2.75	0.10	0.746	0.002
主效应	140.86	1	140.86	5.41	0.023	0.078

4.3.职前教师的学业成绩

使用独立样本 t 检验来解决问题三，比较职前教师的前测、后测和进步分数（后测−前测分数），结果如下表 4 所示，根据前测分数（t=−1.47，p=0.144>0.05）可知师范生的学习成绩没有显著差异，前测数据不会影响后测数据；后测分数（t=2.20，p=0.031<0.05）可知实验组与控制组之间存在显著性差异；以及进步分数（t=3.55，p=0.001）可知实验组和控制

组之间的分数有明显的显著差异。因此基于多模态同伴评价的同伴教学是否能够明显缩小不同水平师范生之间在 TPACK 和批判性思维方面的差距。

从古至今学业成绩都有一定的重要性，古代利用科举考试的成绩选拔人才，在现代学业成绩是进入高中、大学和研究生的主要依据。学业成绩是用来测量人的学识和能力的手段，因此在本研究中将测量不同水平师范生专业发展差距转换成可以测量的学业成绩差距。本研究的实验结果证实了利用多模态同伴反馈的同伴教学能够减小不同水平师范生专业发展之间的差异。

表 4 职前教师学业成绩的独立样本 t 检验分析结果

变量	分组	样本	平均值	标准差	*SE*	*F*	*p*	*T*
前测	实验组	41	78.12	24.58	3.84	2.718	0.144	−1.47
	控制组	27	85.59	11.11	2.13			
后测	实验组	41	85.83	12.68	1.98	0.025	0.031	2.20
	控制组	27	78.74	13.41	2.58			
进步分数	实验组	41	7.70	19.20	2.99	0.058	0.001	3.55
	控制组	27	−6.85	11.20	2.15			

4.4.研究不足与未来研究趋势

本研究有一些局限性，第一，本研究参与者同伴教学的科目固定，不同科目实施同伴教学的效果是不一样的，因此未来可以设计多种科目下的基于多模态同伴反馈的同伴教学的效果差异，从而扩大这种方式的应用范围，进而证明该方式的普适性。第二，没有统计师范生为同伴教学和同伴反馈所花费的平均时间，在时间计划上不能给教师选择这种教学方式提供参考。第三，现有数据仅来自一个班，数据量不够大和广，在未来可以扩大样本量。

5.总结

本研究表明利用基于多模态同伴反馈的同伴教学培养师范生是有效的、可操作性的教学策略，具体体现在 TPACK 能力和批判性能力的提高，以及缩短不同水平师范生之间的差距。未来师范生的培养应多注重实践，在实践过程中使用一定的技术手段去全面剖析师范生各方面能力，比如本研究利用的智慧教研和学习元平台两个平台，能够记录师范生进行同伴教学和同伴反馈的全过程多模态数据，使教师能够有依据的认真剖析，且该平台能够自动分析文本、语言、面部表情等数据，并以可视化的结果呈现给教师，不仅能节省教师分析的时间，还能避免教师因数据分析能力不足而导致分析出现偏差的问题。

参考文献

王阿习，余胜泉，湛志强 & 刘秀英. (2022). 在线同伴反馈情感识别与可视化：促进教师深度反思与教学问题改进. 中国远程教育 (08), 18−25+34+78.

田丽丽 & 周羽. (2022). 在线英语写作课程多元反馈吸收研究. 外语与外语教学 (01), 95−104+150.

徐晓东，李王伟 & 曹鑫. (2021). 如何在教师专业发展项目中指导教师转变?——来自质性研究的发现. 电化教育研究 (07), 99−105.

黄爱琼. (2020). 以反馈促学：构建基于写作任务的多重反馈模式. 外语教学 (06), 67−71.

Alsaleh, A., Alabdulhadi, M., & Alrwaished, N. (2017). Impact of peer coaching strategy on pre-service teachers' professional development growth in Kuwait. International Journal of Educational Research, 86, 36-49.

Al-Hebaishi, S. M. (2017). "The Effect of Peer Instruction Method on Pre-Service Teachers' Conceptual Comprehension of Methodology Course." Journal of Education and Learning, 6(3), 70–82.

Chaaban, Y., Du, X., Lundberg, A., & Abu-Tineh, A. (2023). Education Stakeholders' Viewpoints about an ESD Competency Framework: Q Methodology Research. Sustainability, 15(3).

Chai, C. S. (2019). "Teacher Professional Development for Science, Technology, Engineering and Mathematics (STEM) Education: A Review from the Perspectives of Technological Pedagogical Content (TPACK)." Asia-Pacific Education Researcher, 28(1), 5-13.

Lowman, J. J. (2016). "A Comparison of Three Professional Development Mechanisms for Improving the Quality of Standards-Based IEP Objectives." Communication Disorders Quarterly, 37(4), 211-224.

Hamlen Mansour, K., Jackson, D. K., Bievenue, L., Voight, A., & Sridhar, N. (2023). Understanding the Impact of Peer Instruction in CS Principles Teacher Professional Development. ACM Transactions on Computing Education, 23(2), 1-21.

Maxson, B. K., Neely, M. E., Roberts, L. M., Stone, S. M., Lowe, M. S., Macy, K. V., & Miller, W. (2019). The power of peers: approaches from writing and libraries. Reference Services Review, 47(3), 314-330.

Prilop, C. N., & Weber, K. E. (2023). Digital video-based peer feedback training: The effect of expert feedback on pre-service teachers' peer feedback beliefs and peer feedback quality. Teaching and Teacher Education, 127.

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2020). Effects of digital video-based feedback environments on pre-service teachers' feedback competence. Computers in Human Behavior, 102, 120-131.

Prilop, C. N., Weber, K. E., Prins, F. J., & Kleinknecht, M. (2021). Connecting feedback to self-efficacy: Receiving and providing peer feedback in teacher education. Studies in Educational Evaluation, 70.

Rodis, O. M., & Locsin, R. C. (2019). The implementation of the Japanese Dental English core curriculum: active learning based on peer-teaching and learning activities. BMC medical education, 19(1), 1-7.

Romijn, B. R., Slot, P. L., & Leseman, P. P. (2021). Increasing teachers' intercultural competences in teacher preparation programs and through professional development: A review. Teaching and Teacher Education, 98.

Schmid, M., Brianza, E. & Petko, D. (2020). Developing a short assessment instrument for Technological Pedagogical Content Knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model. Computers & Education, 157.

Sosu, E. M. (2013). The development and psychometric validation of a Critical Thinking Disposition Scale. Thinking Skills and Creativity, 9, 107–119.

Visscher, A. J. (2021). On the value of data-based decision making in education: The evidence from six intervention studies. Studies in educational evaluation, 69.

Xie, K., Kim, M. K., Cheng, S. L., & Luthy, N. C. (2017). "Teacher professional development through digital content evaluation." Etr&D-Educational Technology Research and Development, 65(4), 1067-1103.

生成式 AI 浪潮下教师发展的机遇与挑战：一种教师评价的新途径

Opportunities and Challenges in Teacher Development Amid the Wave of Generative AI: A New Approach to Teacher Qualification Assessment

张梦晨[1*]，陈彦君[1]，周子骏[1]
[1]华东师范大学教育信息技术学系
[*]51254108017@stu.ecnu.edu.cn

【摘要】 以 ChatGPT 为代表的生成式人工智能技术为教育带来变革的同时，也使教育的各环节面临一系列挑战。在教育数字化转型的背景下，新技术如何赋能教师专业发展是一个亟待解决的问题。本文剖析了 ChatGPT 及其背后的大语言模型技术应用到教师发展及教师评价中的作用与挑战，挖掘教师评价环节面临的问题与需求，通过设计嵌入生成式 AI 的教师资格考核试讲评价系统、结合大语言模型 API 和人工智能 API 等开发 App 并验证其有效性，为 ChatGPT 等生成式 AI 在教师评价过程中的广泛应用提供了范例。
【关键词】 生成式人工智能; ChatGPT; 教师评价; 教师发展

Abstract: *The Transformative Impact of Generetive AI, represented by ChatGPT, brings both opportunities and challenges to various aspects of the educational landscape. Against the backdrop of the digital transformation of education, addressing how new technologies can empower teachers' professional development is an urgent issue. This article analyzes the role and challenges of ChatGPT and its underlying large language model(LLM) to teacher development and assessment. It outlines the issues and requirements in teacher qualification assessment, proposing a solution through the design of an generative AI-based teacher qualification assessment system. The development of application integrating LLM APIs and AI APIs is discussed, along with the validation of its effectiveness. This study provides an example for the widespread application of ChatGPT in the teacher assessment process.*
Keywords: Generative AI, ChatGPT, Teacher assessment, Teacher development

1. 引言

人工智能技术的进步正持续推动着教育教学的改革与创新发展。把握人工智能技术发展态势，积极应对其在教育领域应用中面临的风险和挑战，培养具备智能素养和高阶能力的“AI+教育”教师队伍，是教育领域发展的主要方向(孟青泉，贾积有, 2019)。2022 年底，以 ChatGPT 为代表的生成式人工智能(Generative Artificial Intelligence, GAI)取得重大技术突破，越来越强大的生成式 AI 将拓展更广泛的应用场景，变革人机协同关系，对“AI+教育”理论与实践产生巨大影响。教师评价是对教师工作表现进行全面评估的过程，其目的在于客观、准确地了解教师的教学质量，应用 AI 技术能够提供客观数据分析、个性化反馈、自动化流程等，是生成式 AI 技术和教师发展的一个良好的结合点。

本文将深入分析人工智能驱动教育变革的发展现状，剖析 ChatGPT 应用在教师发展中所面临的问题，以相关生成式 AI 技术与教师评价的结合作为切入点，重点以生成式 AI 赋能教师评价的新途径为例开发一个应用案例，最后探讨人工智能赋能教育的发展之道。

2. 问题背景概述

教师专业发展与教师评价之间存在紧密的相互关系，要促进教师专业发展，就要建立完善的教师评价体系。课堂教学能力是教师评价中考察的基本能力。教师上课涉及到五个关键环节，包括导入新课、讲授新课和巩固练习、归纳小结、布置任务(姚声远，于越, 2002)。在中

学教师资格试讲的具体要求中，就提到了教学过程要涉及到以上五个环节。对于未通过教师资格面试的职前教师来说，学习好五环节课堂教学是后续专业发展的基石，只有深入学习基本的教学知识，在了解、掌握的基础上才可以应用、迁移。

然而，传统的教师资格面试试讲评价过程常常依赖于主观评价，容易受到个人主观偏好和经验的影响。教师资格评价中呼吁着更加客观公平、能为教师提供个性化评价与指导以及实时反馈的技术途径，这种技术赋能是教育领域现代化、高效化的必然趋势。而人工智能技术，以及以 ChatGPT 为代表的生成式 AI 技术，可以作为一种创新的工具，为教师资格面试评价带来更科学的依据，推动教师资格评价效率的提升。

3. 生成式 AI 对教师发展的影响与挑战分析

为明确 ChatGPT 对教师发展的影响，本研究针对其在教育教学中的应用作 SWOT 分析如表 1。针对 ChatGPT 应用于教师评价的伦理和安全问题，开展利益相关者分析如表 2。

表 1 ChatGPT 应用于教师发展的 SWOT 分析

S-strengths（优势）	W-weakness(劣势)
1.ChatGPT 能够加速专业知识的传递，提高知识获取效率，降低教师专业发展的成本。 2.技术开放性高，可以整合到智能辅导系统、辅助教师教学、协助管理决策等多个教育场景，促进教师专业发展途径多样化。 3.促进自主学习和个性化研习，多轮次对话能够为教师提供良好的互动研习体验。 4.解放教师生产力，使教师能够更多地投入到创新教学设计等创造性活动中。	1.ChatGPT 的对话质量会出现波动，影响教师研习效果。 2.ChatGPT 的性能受到数据影响，如果教学数据质量不好，会影响准确性。 3.由于 ChatGPT 基于深度学习技术训练，可能会出现对话前后不一致的情况，影响对话效果，制约教师专业发展的连贯性。 4.由于语言的多义性和复杂性，ChatGPT 可能会误解教师的意图，给出错误的建议。
O-opportunity(机遇)	T-threat(挑战)
1.促进教师发散性思维、创造性思维、批判性思维的培养，促使教师的专业水平提升。 2.促进教育范式重塑，促使大规模因材施教和全流程个性化教育的实现。 3.推动教师培养模式的重构和评价体系的转变，人机协同素养将成为重要的发展方向。 4.具有广泛的下游业务群，可以基于此技术设计其在教育教学领域的应用方案，促使教师专业发展的协同素养提升。	1.可能导致教师过度依赖 ChatGPT，缺少独立思考，对专业发展产生负面影响。 2.对教师的信息甄别能力要求更高，部分错误信息难以识别，需要教师反复检查与确认。 3.对监管的要求更高，迫切需要划定责任、明确权益，影响教师专业发展的规范性。 4.可能引发新的隐私保护问题，ChatGPT 数据处理过程的“黑箱”会引发教师对个人隐私、数据安全及伦理问题的担忧。

表 2 ChatGPT 教育伦理安全的利益相关分析

利益相关者	学生	教师	机构/部门
在教师评价的相关数据、算法及其交叉因素等方面	获得个人教学数据的归属和支配权; 意识到数据的商业价值； 算法隐私权(如：拒绝 ChatGPT 评价情绪状态)； 能够自主选择加入/退出评价算法。	获得 ChatGPT 培训,了解其数据如何被用于评价； 反思 ChatGPT 决策路径； 了解如何解释 ChatGPT 算法分析结果； 深刻理解 ChatGPT 的潜力和影响。	尊重法律法规和数据所有权，防止数据泄露和欺诈； 教育数据准确性和更新； 允许个人决定如何分析其数据； 确保引入的 ChatGPT 和算法准确、无偏向性。

4. 案例研究：开发嵌入生成式 AI 技术的教师资格评价 App

本研究致力于开发一款教师资格面试试讲视频分析系统，旨在通过应用 AI 技术，为教师试讲提供客观、科学且全面的评价工具。该系统结合语音识别 API、OCR 文字识别 API 以及大语言模型 API 等技术，以分析和评估教师试讲在导入、授课和复习环节中的表现。

4.1. 需求调研

本研究对备考教师资格面试的情况做调查分析，主要分为问卷调查和访谈两个部分。调查发现教师在备考教资面试时存在许多困难，使用人工智能技术来辅助学生教资面试备考的市场需求较大，如图 1 所示。



图 1 备考教师资格面试的需求调查

结果显示，有 82%的教师采用自学的方法，14%的教师采用上网课的方式，仅有 4%的教师会找人进行专门辅导。当被问及教师如何评价自己的教学过程时，有 87%的教师表示，通过试讲给其他人并获取他们的评价，以此改进自己的教学方法、教学习惯等。有 11%的教师采用微格教学的方法，即把自己上课的视频录制下来，反复打磨各个教学环节。教师自我评估的方法是偏主观的，这种局限于和教师同等水平的自评和他评方式难以得到较好的效果。

4.2. 功能设计

为了满足上述需求，本研究设计的教师备考与资格评价 app 主要有以下功能：

（1）账号注册与登录。教师通过账号注册，建立一个唯一标识主键，即 id。不同的 id 将存储不同的视频数据和分析反馈记录数据。

（2）视频上传与分析。教师登录后，进入视频分析模块，可以得出试讲主题、心理素质、板书检测、识别试讲文本等多维度的数据。试讲主题会将教师提及的关键词按照频次从高到低排列。心理素质显示的是教师的情感组成，如教师在讲课过程中积极、消极情绪的占比情况。板书检测则是识别并呈现教师的板书内容以及排版。试讲文本识别可以将教师试讲过程中的全部文本记录下来。

（3）教案识别和生成。在视频分析模块的下方，有教案识别模块，用户可以上传图片识别。此外，可以输入学科、教学主题和学段，输出大模型生成的教案。

4.3. 系统技术架构与开发实现

本研究开发的系统的技术架构如图 2 所示。

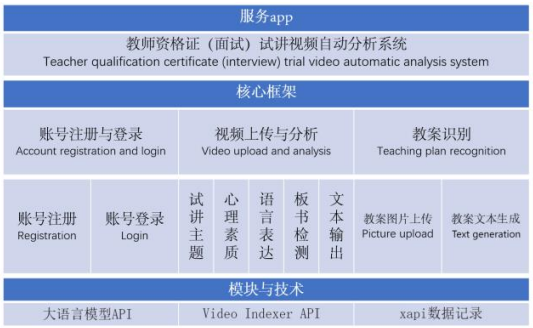


图 2 系统架构图

在关键技术方面，主要涉及大语言模型 API 调用、Video Indexer 视频分析 API 调用和 XAPI(Experience API)数据追踪技术。

在开发部分，本研究使用 App Inventor2 快速开发验证原型。最终实现的移动端 App 如图 3 所示。

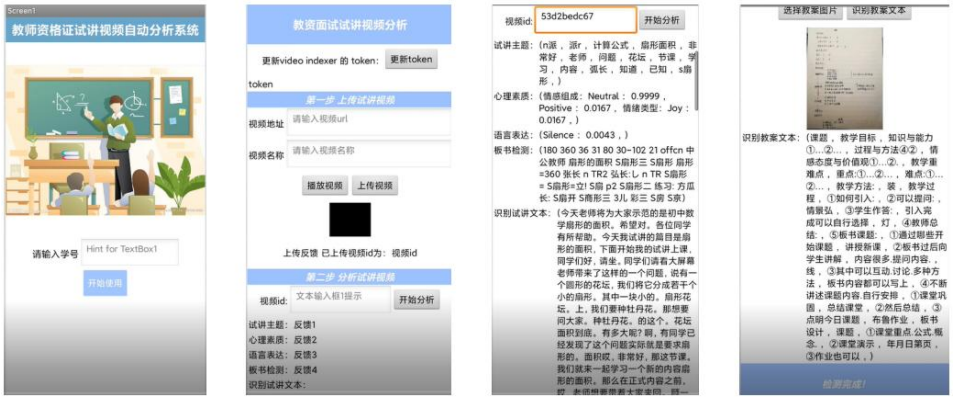


图 3 教师资格考试试讲视频分析 App 界面

4.4. 用户使用体验调查与分析

针对平台的使用体验评估，我们让用户体验手机端 App 生成体验感受，并用改进的 UEQ 问卷进行调查。问卷分为基本信息、用户体验打分、开放性问题三大部分，其中基本信息包含用户性别和年龄，用户体验打分从可靠性、易用性、高效性、激活性、吸引力、创新性六个维度开展，采用李克特五点量表收集用户打分，开放性问题则询问了用户的意见和建议。问卷发放后收集了 52 份答卷，其中有效答卷 51 份，统计性分析结果如图 4 所示。可见用户体验感在六个维度均超过 4.0 分（满分 5.0），访谈中用户也显示了较为积极的使用倾向。

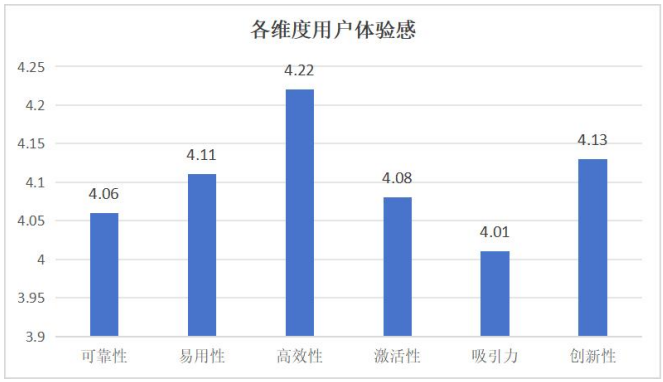


图 4 用户体验分析结果

5. 结论与展望

运用 AI 技术以及其中的大模型技术进行教师评价，能够准确捕捉教师教学过程中的关键信息，为教师提供客观全面的多维度反馈，提高评价的效率和准确性，为大规模的评价提供了可能。实时评价反馈使教师能够迅速了解自己的教学效果，指导教师制定针对性、个性化的发展计划，促进教师发展。

参考文献

孟青泉，贾积有.(2019). 人工智能教育研究及应用中的问题剖析与发展建议. 人工智能(第 3 期), 110-118. doi:10.16453/j.cnki.issn2096-5036.2019.03.015

姚声远，于越.（2002）. “五环节教学模式”释放—“研究体验式创新教学法”课堂教学模式探究. 教育探索(第 6 期), 12-13.

基于 ChatGPT 的人机协同教研中职前教师同伴反馈的社会网络分析

Social network analysis of pre-service teacher peer feedback in human-computer collaborative teaching and research based on ChatGPT

顾佳宁[1]，李欣[1]，蔡慧英[1*]
[1]江南大学人文学院 教育技术系
*caihy@jiangnan.edu.cn

【摘要】 职前教师同伴反馈是教研的核心。随着人工智能技术的发展，智能技术的引用也成为促进职前教师同伴反馈的创新手段。在同伴反馈活动过程中，社会交互作为同伴反馈的内核成为研究职前教师同伴反馈的关键。因此，本研究开展基于 ChatGPT 的教研活动，探究智能工具对于职前教师同伴反馈的影响。本研究共招募 60 名职前教师进行情境实验，收集职前教师同伴反馈的社会交互数据。对实验过程数据采用社会网络分析与独立样本 t 检验方法进行分析，基于结果提出促进职前教师同伴反馈的策略。
【关键词】 ChatGPT；教师同伴反馈；社会网络分析

***Abstract*:** *Pre-service teacher peer feedback is the core of teaching research. With the development of artificial intelligence technology, the reference of intelligent technology has also become an innovative means to promote peer feedback of pre-service teachers. In the process of peer feedback activities, social interaction as the core of peer feedback has become the key to studying peer feedback of pre-service teachers. Therefore, this study carried out teaching and research activities based on ChatGPT to explore the impact of intelligent tools on pre-service teachers' peer feedback. This study recruited a total of 60 pre-service teachers to conduct a situational experiment and collect social interaction data of pre-service teachers peer feedback. The experimental process data were analyzed using social network analysis and independent sample t-test methods, and based on the results, strategies to promote peer feedback of pre-service teachers were proposed.*
Keywords: ChatGPT, teacher peer feedback, social network analysis

1. 前言

职前教师的同伴反馈是一次社会交互的体验，也是教学提升与发展的机会。教师同伴反馈（Teacher Peer Feedback）是指基于教学实践、教研活动，教师之间共享关于教学表现、教学计划和实际问题等方面的信息，并进行相互评价和建议的过程，强调教师同伴之间的相互支持和合作(Jin，2021)。教师的社会交互作为同伴反馈的核心，逐渐成为研究教师同伴反馈的关键。

同时，当前人工智能的高速发展为教师教研提供了创新思路。尤其是应用生成式人工智能 ChatGPT 被认为能向教育者提供更多资源并促进教师的同伴反馈，以更好地应对教育挑战。但是，目前职前教师同伴反馈较为浅显、难以把握并分析教学中存在重点问题，所以，创新生成式人工智能的出现与使用能够支持教师深入思考教学问题，促进职前教师同伴之间更密切的话语讨论、激发职前教师更有效地分享经验（吴军其，吴飞燕等人，2023）。因此，本研究欲通过探究基于人工智能 ChatGPT 支持的人机协同教研中职前教师同伴反馈社会交互情况，分析 ChatGPT 对于职前教师同伴反馈的作用。以期挖掘出基于人工智能 ChatGPT 促进教师同伴反馈的方案，为教育改革和提高教学质量做出更大的贡献。

2. 文献综述

2.1. 教师教研中的同伴反馈

同伴反馈（Peer Feedback）是指至少两个同伴之间的协作活动（Kollar & Fischer, 2010），同伴之间相互提供信息与资源。教师同伴反馈是指，在教育或工作环境中，基于某一节课堂或某一个教学任务，两名或两名以上教师同伴就教学表现、教学计划、实践问题等环节互相观察进行信息共享、协商共议、点评指导等反馈活动以促进优化教学和提升个体专业能力的协作过程(Kollar & Fischer, 2010; Jin, 2021)。教师同伴反馈可涵盖多种类型的教师学习活动，如教学反馈、同伴评估、教学示范和同伴辅导与经验传授等（Iacono, Pierri, &Taranto, 2019）。

已有研究表明，同伴反馈是一个能为教师提供学习资源、提升教师知识储备、刺激教师反思和行为改变的关键活动(Pearce et al., 2019)。另外，部分研究者认为，培养职前教师同伴反馈是职前教师未来发展的关键（Prilop & Weber，2023）。例如，Prilop 和 Weber（2023）的研究中，同伴反馈能够促进职前教师专业知识的提升。Axi、Shengquan 等人 (2023)的研究中，同伴反馈能够有效提高各阶段教师的技术教学内容知识（TPACK）。

2.2. 教师同伴反馈中的社会交互与社会网络分析

同伴反馈其本质在于同伴之间的社会交互。同伴间的社会交互可以理解为教师或学生之间建立关系并进行交流的行为（刘权纬，王兴辉，蒋红星，2019）。部分研究者推测高质量的社会交互为个体和组织提供了宝贵的反馈资源。例如，Tan 和 Chen（2022）认为社会交互行为可以帮助教师接收来自同伴的直接观察和评价，了解教学中可能存在的问题，并寻找针对性的建议和解决方案。刘权纬、王兴辉和蒋红星（2019）认为高质高效的社会交互能够加强团队的协作，影响着同伴反馈质量与互动多方的工作效能。

文献梳理发现，贯穿教师同伴反馈全过程的社会交互是多维度的（Falchikov，2003）。社会交互包括寻求、解释、分析和总结信息，通过各种观点和论点进行批评和推理，以及做出决定与指导（Zhu，2006）。不同研究者对于社会交互过程的分析角度不同。基于学习过程的内容分析，Henri（1992）从社交、互动、元认知、认知等四个维度提出了基于计算机会话内容的分析框架。在 Henri 所构建的模型的基础上，朱尔平（Erping Zhu）（2006）整合了布鲁姆认知分类目标理论，从提问、回答、反映、指导、脚手架五个维度对参与者交互参与程度及其认知水平的变化进行归类。

社会网络分析（SNA）是探究行动者社会交互的主要方法。该方法关注行动者及其关系形成的社交网络结构，并通过图形化表征和量性表征方式分析学习者交互的网络结构属性和规律（胡天慧,刘三女牙,粟柱等人，2022）。已有研究者采用社会网络分析对学习者的社会交互特点进行分析。例如，刘权纬和王兴辉等人（2019）采用社会网络分析方法聚焦教师网络研修过程中的学习交互行为，分析教师的交互内容，为改革教师培训模式、解决研修问题提供借鉴。陈玲等学者（2013）从网络结构图、中心性、凝聚子群、派系四个方面分析了教师区域网络协同备课互动网络中教师群体和个人层次的参与特点及协同知识构建情况。综上，社会网络分析能够用于分析职前教师同伴反馈中社会交互的特征。

2.3. 人工智能 ChatGPT 助力教师同伴反馈

近年来，随着人工智能技术的不断发展，GPT-4、BERT 等新一代大规模语言模型（Large language models，LLMs）成为了引领潮流的先进技术。在自然语言处理（NLP）技术支持下，OpenAI 的 ChatGPT、Microsoft 的 Newbing、百度的文心一言等搭载大规模语言模型的人工智能聊天工具应运而生，为人们提供了更智能、高效、个性化的交流方式。ChatGPT 是由 OpenAI 推出的一种大规模语言模型，本质上，它是一种从人类反馈中强化学习的技术开发程序，专门设计用于进行对话式交互的人工智能交互工具（李会春，2023）。ChatGPT 使用基于 GPT-3.5（GPT-4.0）架构的大型语言模型并通过强化学习进行训练。它能根据已有信息预测下一段文本，能自主从数据中学习，并在对大量文本数据集进行训练后，生成复杂且看似智能的文字。

ChatGPT 可以作为辅助工具支持教师的学习与发展（吴军其，吴飞燕等人，2023；夏琪，程妙婷等人，2023）。首先，ChatGPT 语料库中存储的大量信息和专业知识。教师可以向

ChatGPT 咨询特定学科的概念、原理、规律等基础知识，并获取相关的参考资料，在与 ChatGPT 的互动中不断吸纳和借鉴全新的知识视角。教师也可以向 ChatGPT 咨询定制化教学策略、课程计划、活动设计等案例信息，快速累积教学经验。其次，ChatGPT 能够对教师输入的教学实践、作品等内容进行快速分析并提出针对性建议，辅助教师进行反思。最后，ChatGPT 可以模仿专家，支持教师以对话的形式与“虚拟专家”进行互动。ChatGPT 支持教师提出教学问题，助力教师分析、整合已有知识，立足问题、从理论层面分析问题成因，提供个性化的“虚拟专家指导”。

因此，部分研究者认为，在教师教研过程中引入 ChatGPT 等人工智能聊天工具对促进教师同伴反馈，引导教师社会交互具有一定意义（Bauer，Greisel，2023）。本研究旨在探究基于 ChatGPT 的人机协同教研对职前教师同伴反馈的社会交互存在何种作用。因此，基于上述的梳理，本研究的研究问题如下所示：

1.基于 ChatGPT 的人机协同教研中职前教师同伴反馈的社会网络分析结果是否存在差异？

2.基于 ChatGPT 的人机协同教研中职前教师同伴反馈的社会网络分析结果存在什么差异？

3. 实验设计

3.1. 参与者与实验情境

为了从实证角度探究基于 ChatGPT 的人机协同教研中的职前教师同伴反馈情况，研究团队运用随机招募的方式，面向江苏某市共招募了 60 名职前教师，并组织他们参与情境实验。所招募的职前教师中，男性有 18 人，女性有 42 人。在本研究中，研究团队将 60 名职前教师分为控制组与实验组两组。其中，控制组职前教师 30 名，需要参与传统条件下的同伴反馈活动实验。实验组职前教师 30 名，需要参与 ChatGPT 这一人工智能工具支持条件下的同伴反馈活动实验。此外，研究团队对控制组和实验组的职前教师进行随机分组，3 人一组共 20 组。其中，1-10 组为控制组，11-20 组为实验组。

本研究于 2023 年 10 月组织并开展实验，地点在江苏省某高校机房展开，实验时间为 3 小时左右。参与的职前教师需要根据实验要求同伴协作完成教案的修改与完善，在协作过程中进行同伴反馈的交互。职前教师使用腾讯 QQ 平台进行同伴反馈的交流与讨论，使用金山文档进行教案的修改与完善。

3.2. 研究工具：ChatGPT

为了研究基于 ChatGPT 的人机协同教研中职前教师的同伴反馈情况，研究团队选取了 Esaychat 聊天机器人作为研究工具。Esaychat 作为智能聊天机器人可以帮助职前教师更好的理解、处理同伴提出的问题，起草、优化反馈给同伴的问题回答。具体来看其功能包括，第一，EasyChat 能够为职前教师提供信息。例如，EasyChat 能够为用户提供各种基本信息和信息检索的方法，或根据用户的指令和要求模仿或生成相关文本信息。第二，EasyChat 可以实现一些基本的文本匹配、相似度分析、关系判断和情感分析等文本分析功能。第三，EasyChat 能够与职前教师进行交互式问答对话，帮助职前教师找到特定问题的解决方案。第四，EasyChat 可以进行角色扮演，扮演不同的角色能够使其表现出具有特定身份特色的话语表达，以更加个性化或专业化的方式来处理问题或进行对话。

3.3. 实验过程

本次实验主要包括三个阶段。第一阶段为实验准备阶段。研究团队准备职前教师开展实验所需的实验材料，包括教案、实验任务单和问题提示支架。教案是职前教师同伴反馈活动开展的载体，职前教师需要通过对教案的修改与完善进行同伴反馈对话。实验任务单用于支持职前教师更好地了解实验进程和所需完成的同伴反馈任务。问题提示支架辅助职前教师针对教案提出疑问，可以支持职前教师同伴反馈活动的开展。第二阶段为实验开展阶段。研究团队组织参与研究的职前教师在高校机房集中实验。在实验开始前，团队成员下发实验过程中所需的全部材料，并向职前教师讲解教案、实验任务单和问题提示支架等材料的使用方式。其中，研究团队还需要向实验组职前教师介绍 ChatGPT 的登陆与使用方式。在实验过程中，

职前教师需要按照实验任务单的流程，首先对金山文档中的教案进行问题批注，提出自己认为的不足之处。然后，职前教师同伴们相互回复批注，同伴反馈全程可在 QQ 群聊中进行对话讨论。最后，职前教师基于讨论总结并填写教案的修订方案，结束实验。在整个实验过程中，实验组的职前教师可以根据自身需要使用 ChatGPT 工具。第三阶段为数据收集与分析阶段。研究者在同伴反馈活动完成后，收集和摘取职前教师在 QQ 平台上的对话记录和金山文档上的批注记录。此后，对所收集的职前教师同伴反馈对话数据进行社会网络分析。

3.4. 数据收集与分析

为了探究基于 ChatGPT 的人机协同教研中职前教师同伴反馈的社会交互情况，研究团队下载并分析职前教师在 QQ 群聊中的聊天记录以及金山文档中的教案讨论与修改批注。而后，研究团队中的两位研究助理分别对下载的职前教师同伴反馈社会交互数据进行质性编码。

表 1 社会交互内容分析框架（朱尔平，2006）

层次	子维度	内涵	编码
提问	信息寻求	有直接和正确答案的问题	Q1
	探究式讨论	没有直接和正确答案的问题	Q2
陈述	回应	直接回应之前的信息，提供反馈、意见等的陈述	S1
	信息提供	提供与所讨论的主题相关的（传闻或个人）信息的陈述	S2
	说明	以有限的个人意见来解释相关阅读或信息的事实性陈述	S3
	分析	对回应信息或相关阅读材料提供分析性意见的陈述	S4
	总结	总结或试图提供讨论信息和相关阅读材料的摘要的声明	S5
	评价	对讨论或相关阅读中的关键点提供评价性或判断性意见的陈述	S6
反映	变化的反映	反映出个人观点和行为变化的陈述	R1
	使用认知策略的反映	解释或反映一个人在完成某些学习任务中使用的认知策略/技能的陈述	R2
指导	指导	解释或展示如何理解特定概念（想法等）的陈述	M1

首先，研究团队选取社会交互内容分析框架（朱尔平，2006）中的 11 个子维度作为本研究的编码依据，对职前教师同伴反馈中成员交互行为进行分析，如表 1 所示。其次，研究助理对每位教师的话语数据进行了单一语义的切割处理，将其作为后续的编码单元。最后，两位研究助理对数据进行编码。编码者一致性信度为 93.6%，在 90%以上，说明编码结果可靠性较高。基于编码结果，研究团队选用了社会网络分析方法进一步分析与探究。

4. 数据分析结果

为了探究基于 ChatGPT 支持的人机协同教研中职前教师同伴反馈的社会网络结构差异，研究者选用 UCINET 6.186 工具，对控制组与实验组职前教师的同伴反馈对话编码结果进行了社会网络分析，构建职前教师社会交互的社会网络结构图，并得出社会网络结构基本属性及中心性指标。此外，为了进一步比较有无 ChatGPT 支持的同伴反馈活动参与对于职前教师同伴反馈感知过程的影响差异，研究者选用 SPSS26.0 工具，对控制组与实验组职前教师的社会网络结构图及相关指标进行了独立样本 T 检验。

4.1. 基于同伴反馈社会交互数据编码结果的职前教师认知交互社会网络结构基本属性

针对控制组与实验组的职前教师社会交互网络结构的基本属性进行分析可以得到如下表 2 的结果。其中，密度是用来测量社会网络中节点间的联结程度，它主要表示节点整体的关系是否紧密,。密度值介于 0 和 1 之间，值越接近 1 则代表彼此间关系越紧密。聚类系数是用于衡量节点与邻居之间连接紧密程度的指标，反映了节点的邻居是否倾向于相互连接，形成紧密的群体或社团。聚类系数的取值范围也在 0 到 1 之间，越接近 1 的聚类系数表明节点邻居之间连接越密集，易形成社团或群体；而接近 0 则表示节点邻居之间连接相对稀疏。互惠性是指节点之间的双向连接关系，具体而言，如果节点 A 与节点 B 有连接，而且节点 B 也与节点 A 有连接，那么存在互惠性。

表 2 职前教师社会交互网络结构基本属性

组别	组号	网络基本属性			组别	组号	网络基本属性		
		密度	聚类系数	互惠性			密度	聚类系数	互惠性
控制组	1	0.07	0.38	0	实验组	11	0.10	0.19	0
	2	0.09	0.12	0		12	0.07	0.52	0.14
	3	0.05	0	0		13	0.26	0.63	0.38
	4	0.07	0.61	0.14		14	0.30	0.39	0.10
	5	0.14	0.6	0.15		15	0.10	0.22	0
	6	0.14	0.27	0.07		16	0.15	0.58	0.14
	7	0.55	0.35	0		17	0.26	0.61	0.32
	8	0.05	0	0		18	0.23	0.50	0.09
	9	0.04	0	0		19	0.23	0.39	0.09
	10	0.08	0.35	0		20	0.35	0.59	0.26

对于控制组而言，从密度数值分析可知，网络的密度较小，即职前教师社会交互的联系较小。聚类系数的数值结果显示控制组中多数组别的网络社交节点与邻居较为分散，凝聚力较低。互惠性结果显示，组内的交互内容回应频率并不高，只有个别组数值较为突出。对于实验组而言，从密度数值分析可知，网络的密度同样较小，即教师社会交互的关联性也较弱。聚类系数的数值结果显示控制组中多数组别的网络社交节点较为密集，数值超过 0.5，网络凝聚力较高。互惠性结果显示，组内的社会交互的往来适中，只有个别组数值较低。同时链接数也说明，实验组教师在交互过程中的对话互动较多。

表 3 职前教师网络基本属性数据独立样本 t 检验分析结果

维度		控制组		实验组		t	P
		M	SD	M	SD		
网络基本属性	密度	0.13	0.15	0.21	0.09	-1.36	0.191
	聚类系数	0.27	0.23	0.46	0.16	-2.17	0.044
	互惠性	0.04	0.06	0.15	0.13	-2.57	0.019

将实验组与控制组的网络基本属性数据进行独立样本 t 检验，得到结果如表 3 所示。控制组与实验组的社会网络基本属性在密度（t=-1.36，p=0.191＞0.05）方面不存在显著差异。但控制组与实验组的社会网络基本属性在聚类系数（t=-2.17，p=0.044＜0.05）、互惠性（t=-2.57，p=0.019＜0.05）方面存在显著差异。对比发现，实验组与控制组的密度数值相差无几，即整体交互紧密程度不高；实验组的聚类系数与互惠性数值高于控制组，即实验组职前教师同伴反馈社会交互凝聚力与双向关系显著高于控制组。总体而言，虽然实验组与控制组职前教师同伴反馈社会交互的整体交互结构较为稀疏，即基于 ChatGPT 的人机协同教研对于职前教师同伴反馈整体交互程度作用较弱，但 ChatGPT 能够促进职前教师的局部社会交互、反馈话语的密切联系与凝聚程度。

4.2. 基于同伴反馈社会交互数据编码结果的职前教师认知参与社会网络结构中心性分析

针对控制组与实验组的职前教师社会交互网络结构的网络中心性进行分析可以得到如下表 4 的结果。点度中心性指节点在某一网络中所拥连接的数目，刻画节点的局部中心指数，测量网络中节点本身的交互能力。中介中心性指的是一个节点担任其他两个节点之间最短桥梁的次数，充当“中介”的次数越高，节点的中介中心性就越高。

表 4 实验组与控制组网络中心性指标分析结果

组别	组号	网络中心性		组别	组号	网络中心性	
		点度中心性	中介中心性			点度中心性	中介中心性
控制组	1	2.82	0	实验组	11	5.82	0.36
	2	3.64	0		12	2.00	0.18
	3	1.09	0		13	7.36	2.27
	4	2.36	0.46		14	8.82	2.27
	5	9.73	0.64		15	2.55	0.18
	6	1.36	1.18		16	1.46	2.36

7	2.82	0	17	8.18	3.91
8	0.82	0	18	11.18	0.55
9	1.18	0.09	19	9.72	1.18
10	2.73	0	20	15.27	2.36

对于控制组而言，从点度中心性的数值可知，职前教师在教研过程中同伴反馈社会交互处于较低水平，职前教师社会交互较为分散，网络结构较不稳定。中介中心性均值为 0.237，从整体数据来看，中介中心性较低，说明拥有重要影响的节点较少，在教师社会交互的过程中需要注意具有关键性节点在教师学习中发挥的重要作用。对于实验组而言，从点度中心性的数值可知，职前教师在基于 ChatGPT 的教研过程中同伴反馈社会交互水平较高，社会交互行为更为密集。中介中心性均值为 1.562，从整体数据来看，中介中心性处于较高的水平，说明拥有重要影响的中介节点较多。

表 5 实验组与控制组网络的网络中心性独立样本 t 检验分析结果

维度		控制组		实验组		t	P
		M	SD	M	SD		
网络中心性	点度中心性	2.86	2.59	7.24	4.40	-2.71	0.014
	中介中心性	0.24	0.40	1.56	1.26	-3.17	0.009

将实验组与控制组的社会交互网络结构的网络中心性数据进行独立样本 t 检验，得到结果如表 5 所示。控制组与实验组的社会网络中心性在点度中心性（t=-2.71，p=0.014＜0.05）、中介中心性（t=-3.17，p=0.009＜0.05）方面存在显著差异。对比发现，实验组分数均高于控制组分数。总体而言，基于 ChatGPT 的人机协同教研中职前教师同伴反馈能够增强职前教师重要社会交互数量，突出了中介节点对于整体同伴反馈社会交互的影响，也提升职前教师的社会交互质量与效率。

4.3. 基于同伴反馈社会交互数据编码结果的职前教师社会交互社会网络结构图

利用社会网络结构图对社会网络中的关系进行描述，可以帮助研究者直观清晰地了解并分析不同类型知识的交互关系以及强度。研究团队采用 UCINET 工具分析 20 个组的职前教师社会交互社会网络结构，并从控制组与实验组两个群体角度分别进行组内对比与组间对比。首先，控制组社会网络结构图如下图 1-3 所示。基于社群结构图可知，未使用 ChatGPT 的职前教师在同伴反馈过程中其社会交互结构较为分散，节点与节点间的链接较少。同时，链接节点的线段粗细代表了交互的强度，可以看出控制组职前教师整体的社会交互情况较弱，更有第 3、8 等组别结果显示社会交互与思考的维度较局限。

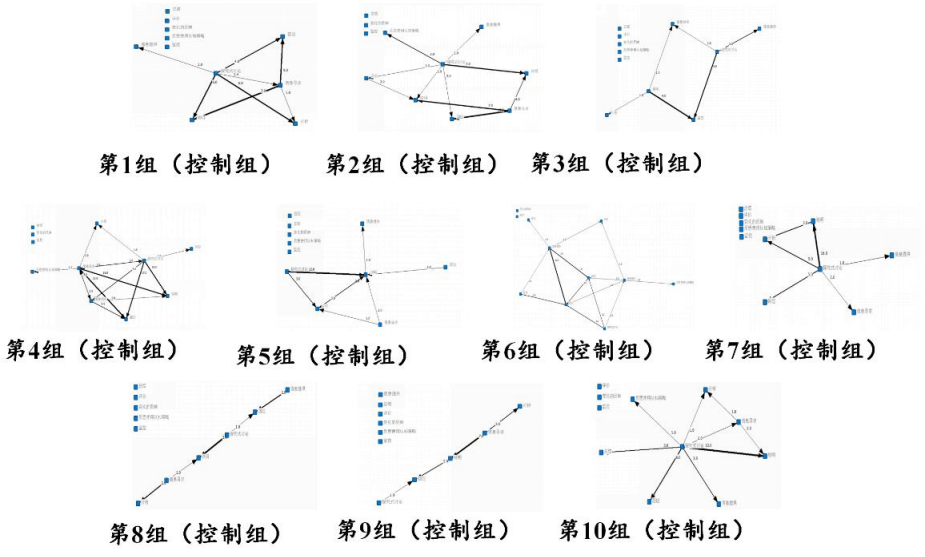


图 2 第 1-10 组（控制组）社会交互网络图

其次，实验组社会网络结构图如下图 4-6 所示。结合社群结构图可知，使用 ChatGPT 的职前教师在同伴反馈过程中其社会交互结构较为密集，节点与节点间的链接较多。同时链接线段也更粗，可以看出实验组职前教师整体的社会交互情况较强，其中第 18、19、20 组尤为显著。

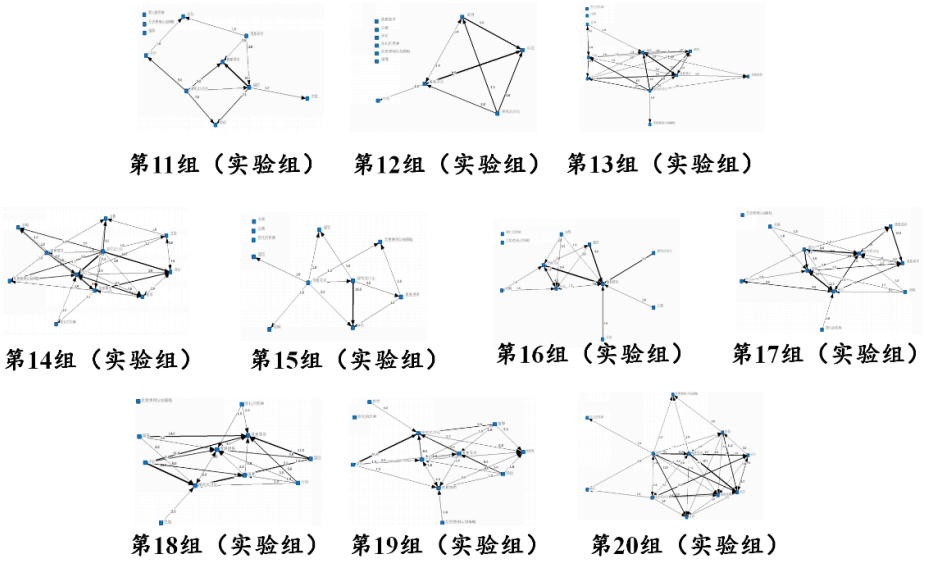


图 3 第 11-20 组（实验组）网络图

对比控制组与实验组的社会交互社会网络结构图可以发现，不同情境下的教师社会交互情况存在差异。第一，ChatGPT 支持下的实验组职前教师，其社会交互强度整体高于控制组。第二，与实验组使用 ChatGPT 的情景对比而言，控制组的所有社群图中显示均存在个别孤立的节点，这说明控制组职前教师的社会交互行为较为居中也较为局限。第三，控制组中的孤立节点均为高维度的交互行为，即评价、总结、指导等维度，但这些维度的节点在实验组的结构图中并不全是孤立的，这说明未使用 ChatGPT 进行人机协同教研的职前教师在同伴反馈社会交互过程中表现较为浅显的认知与行为，ChatGPT 的支持有助于职前教师深度思考教学问题。

5. 结论与启示

整体分析的结果进一步证实了基于 ChatGPT 的人机协同教研中职前教师同伴反馈情况更为良好。首先，ChatGPT 工具的使用显著增强了职前教师同伴反馈的社会交互行为。在使用 ChatGPT 作为辅助工具的实验组中，其社会结构网络图的多个关键指标表现出显著优势。这意味着实验组中职前教师同伴反馈的互动更为密切和频繁，以及他们在教学和反馈过程中的更深入的交流和讨论。其次，通过对比控制组与实验组的社会网络结构图，研究者明显地观察到了两者之间的差异。在控制组中，职前教师的社会交互主要停留在较为基础的层次，如简单的提问和回答。这些互动主要集中在信息寻求、探究性讨论、简单回应以及提供基本信息等方面。然而，在使用 ChatGPT 的实验组中，研究者观察到职前教师展现出了更加高维和深入的社会互动行为。他们不仅在同伴反馈中进行了更为全面的总结和评论，还展现出了对同伴提供反思和指导的能力。究其原因，一方面，ChatGPT 作为一个强大的辅助工具，为职前教师提供了一个更为丰富和多元的资源和信息源来，使他们能够从多个维度和角度思考问题，并产生更深入的交互和反馈（Bauer，Greisel，2023）。另一方面，这种工具可能也激发了教师之间的创新和合作精神，鼓励他们更加主动地参与同伴反馈过程，分享经验和知识。

同时，本研究也存在一些局限。第一，需要扩大研究样本规模。本研究选取的样本量较少，也存在男女比例不平衡等问题，后续可以扩大研究样本规模。第二，本研究只是进行了简单的社会网络分析，并对结果进行表述，后续也将继续采用社会认知网络分析等方法对于

研究数据进行更深度的挖掘与分析。

未来，领导者可以在教研过程中引入 ChatGPT 等生成式人工智能辅助教师展开同伴反馈与协作的社会交互活动。首先，相关教育研究者应倡导教师们利用 ChatGPT 提出问题，并得到智能化的建议，以促进同伴之间的深度交流。其次，研究者应鼓励教师借助 ChatGPT 进行教学分析与评价，例如向 ChatGPT 提交教学设计获取初步反馈，然后再与同伴进行讨论。最后，通过与 ChatGPT 的对话，帮助教师进行教学反思，促使他们更深入地思考自己的教学方法和策略，找到改进的空间。总体而言，这些发现为教育研究提供了宝贵的见解，揭示了在现代教育环境中，如何有效地利用技术工具来促进职前教师之间的同伴反馈。

参考文献

陈玲,王冰洁 刘禹.(2013).教师区域网络协同备课中的社会网络分析.远程教育杂志(03),17-23.

胡天慧,刘三女牙,粟柱,周子荷,余建伟 & 柴唤友.(2022).学习者同步和异步交互模式的比较与整合研究——基于社会网络和多层网络分析的方法. 电化教育研究(05),53-60.

李会春.(2023).ChatGPT 技术热潮下教育变革的挑战和对策. 复旦教育论坛(02),13-23.

刘权纬,王兴辉 蒋红星.(2019).教师工作坊成员学习交互行为的社会网络分析.现代远距离教育(03),22-29.

吴军其,吴飞燕,文思娇,张萌萌 王嘉桐.(2023).ChatGPT 赋能教师专业发展：机遇、挑战和路径. 中国电化教育(05),15-23+33.

Axi Wang, Shengquan Yu, Minhong Wang & Ling Chen (2023): Effects of the visual analytics of peer feedback on teachers' TPACK development in a lesson study, Interactive Learning Environments,

Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., ... & Fischer, F. (2023). Using natural language processing to support peer‐feedback in the age of artificial intelligence: A cross‐disciplinary framework and a research agenda. British Journal of Educational Technology.

Falchikov, N. (2003). Learning together: Peer tutoring in higher education. Routledge.

Henri, F. (1992). Computer conferencing and content analysis. In Collaborative learning through computer conferencing: The Najaden papers (pp. 117-136). Springer Berlin Heidelberg.

Iacono, U. D., Pierri, A., & Taranto, E. (2019). Peer Review Methodology in a Blended Course for Mathematics Teacher Education. International Journal for Technology in Mathematics Education, 26(2).

Jin, X. (2021). Peer feedback in teacher professional development.

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. Learning and instruction, 20(4), 344-348.

Pearce, E., de la Fuente, Y., Hartweg, B., & Weinburgh, M. (2019). Peer‐coaching as a component of a professional development model for biology teachers. School Science and Mathematics, 119(3), 117-126.

Prilop, C. N., & Weber, K. E. (2023). Digital video-based peer feedback training: The effect of expert feedback on pre-service teachers' peer feedback beliefs and peer feedback quality. Teaching and Teacher Education, 127, 104099.

Tan, J. S., & Chen, W. (2022). Peer feedback to support collaborative knowledge improvement: What kind of feedback feed-forward?. Computers & Education, 187, 104467.

Zhu, E. (2006). Interaction and cognitive engagement: An analysis of four asynchronous online discussions. Instructional Science, 34, 451-480.

基于课堂视频的人机协同教研中教师交互模式的滞后序列分析研究

Lag Sequence Analysis of Teacher Interaction Patterns in Human-Computer Collaborative Teaching and Research Based on Classroom Videos

韩冰[1]，蔡慧英[1*]
[1]江南大学教育技术系
[*] caihy@jiangnan.edu.cn

【摘要】 在教师专业发展领域，基于课堂视频的教研是支持教师协作知识建构，助力教师专业成长的重要方式。但是，在缺乏外在干预的情况下，教师之间基于课堂视频的教学研讨存在交互质量不高的问题。为此，本研究提出人工智能赋能的基于课堂视频教研的技术解决方案。为了探究这一技术解决方案的有效性，本研究招募 48 名职前教师进行了准实验研究，采用滞后序列分析方法分析了实验组和控制组中教师在教研过程的交互模式的差异情况。研究发现，在基于课堂视频的人机协同教研中教师交互模式具有任务衔接度好、投入度高、灵活性强的特点。本研究不仅验证了基于课堂视频的人机协同教研对教师交互模式的影响，而且为人工智能技术赋能教师教研方式提供了指导。

【关键词】 教师交互模式；课堂视频；人机协同教研；滞后序列分析

Abstract: *In the field of teachers' professional development, teaching and research based on classroom videos is an important way to support teachers' collaborative knowledge construction and promote teachers' professional growth. However, in the absence of external intervention, teaching seminars merely based on classroom videos have the problem of low interaction quality. This study proposes AI-empowered classroom videos as a solution. To examine its effectiveness, this study recruited 48 pre-service teachers to conduct a quasi-experimental study and used LSA to analyze the differences between the experimental group and control group. The study found the teacher interaction model in human-computer collaborative teaching and research based on classroom videos has the characteristics of good task connection, high investment and strong flexibility. This study not only verifies the impact of human-computer collaborative teaching and research based on classroom videos on teachers' interaction patterns, but also provides guidance for AI-empowered teaching and research.*

Keywords: Teacher interaction mode, Classroom Video, Human-Computer Collaborative Teaching and Research, LAS

1.前言

教师在教研过程中的协作交互质量对教师专业发展发挥着重要作用（Chen et al, 2016; Yang et al, 2015）。知识建构理论认为，学习不仅仅是认知加工的过程，还是在一定的环境中表现出来的有意义的社会性协商的过程(Scardamalia & Bereiter, 2006)。其中，社会性互动被看成是促进学习者有效的主要驱动力（Vigotsky, 1939）。在教师教研过程中，只有当教师个人的能力与其他从业者的集体想法相结合时，才能实现有意义的知识建构（Jina, 2017）。因此，我们需要关注教师教研的交互质量。

目前，基于课堂视频的教研已成为许多教育环境中教师教研的强大工具（Kang & van, 2019）。基于视频的教研是指教师协作观看、分析和讨论课堂视频，促进教师深入反思教学实践，从而提高专业能力的方法（Hamel & Viau-Guay, 2019）。但基于课堂视频的教研还存在一定的优化空间，例如，在缺乏外在支持的情况下，教师可能无法聚焦于课堂视频提供信息中的要点、缺乏与其他教师的积极互动交流、忽略了其他教师在同样问题上的看法（Hu et al, 2023）等，这些情况都会导致教研交互质量不高。为了解决这些问题，基于课堂视频

的人机协同教研是一种有效解决方案。

基于课堂视频的人机协同教研是指运用智能技术对课堂视频数据进行智能提取与分析，支持教师进行数据驱动的教学研讨。从理论上看，基于课堂视频的人机协同教研为更好的促进教师间的交流合作提供了可能性。为了探究基于课堂视频的人机协同教研对于提升教师教研的交互质量的有效性，本研究将创建基于课堂视频的人机协同的教研场景，运用滞后序列分析法挖掘教师在教研过程中的交互模式，为优化人机协同教研实践模式提供研究启示。

2.文献综述

2.1. 教师教研中的协作知识建构及其交互模式

从情境学习理论的角度看，在特定的社会情境中，小组教师在教研过程中的协同知识建构是促进教师专业发展的重要途径。情境学习理论认为学习发生在社会参与和协作的环境中(赵健, 2006)，也就是说，教师需要通过与其他教师交流与共享经验、观点和想法，才能真正实现自身的专业发展。协作知识建构正是指具有共同学习目的的教师间通过社会互动中的想法、观点和认知持续共享、协商和改进过程（Driver et al, 2013）。因此，教师之间协作知识建构的质量可以反映和影响教师专业发展的好坏（魏宁，2006）。

通过文献梳理发现，我们可以运用不同的方法评估教师协作知识建构。第一，我们可以从教师在不同类型的学习活动中时间投入的角度评估教师协作知识建构的情况。例如，So等人通过对话片段中各个想法单元的实际投入时间进行编码，发现视频注释工具的介入极大程度缩减了教师观看与讨论视频内容的时间，有效提高了教师的参与水平（So et al, 2016）。第二，我们可以从教师在不同类型的学习活动中交互模式的角度评估教师协作知识建构的情况。例如，Van 等人通过对教师交互模式的密度、互惠性和集中性进行分析，发现随着讨论、解释、诊断和行动四个阶段的推进，教师间的协作知识建构变得更加激烈(Van et al, 2017)。

目前，已有研究者对协作知识建构的一般过程进行了系统化描述,形成了对应模型。如，Henri的五维分析框架包括参与、讨论、认知、元认知、社会（Henri,1992）。Gunawardena等人（1997）则在 Henri 模型的基础上提出了互动模型框架,具体包括以下五个阶段：（1）共享或比较问题的信息，指教师介绍或者反对观点；（2）发现和探索想法、概念或陈述之间的不协调或不一致，指教师确定分歧；（3）意义协商或知识共建，指教师开展协商或澄清,确定冲突概念之间的一致；（4）拟议合成或共建的测试和修改，指教师根据现有的认知模式或个人经验测试提出新知识；（5）协议声明或新构造意义的应用，指教师陈述总结。

2.2. 人工智能赋能的基于课堂视频的教师教研

基于课堂视频的教研（Classroom Video-based Professional Development）是指教师对录制的课堂视频进行细致地分析和剖析，以促进教师对教学实践进行反思，从而提升教师的专业能力（Hamel&Viau-Guay,2019）。课堂视频是一种具有情境性和真实性的天然媒介，可以为教师提供反复观看和学习的教学实例，能够有效促进教师专业发展。但是，单纯看课堂视频并不能有效促进教师的协作知识建构。研究发现，在缺乏外部干预的条件下，教师无法聚焦到繁多课堂视频信息中的要点内容，难以实现高质量交互。因此，为了充分发挥基于课堂视频的教研对教师协作知识建构的促进作用，我们需要为教师提供一些外在的支持和干预。

目前，人工智能技术正在不断进入教育应用场景，对学习者的学习和教师的教学等提供智能分析、诊断、推荐和干预等支持，为提升教育教学效率和质量提供了发展前景。在教育教学场景中，人工智能技术可以对教育教学过程中产生的多模态数据进行智能分析，从而生成刻画教学过程的诊断结果和分析报告。它们可以作为一种“证据性”资源，支持教师进行人机协同的精准教研。这种人机协同的教研方式不仅可以帮助教师能更加精准地知晓课堂教学过程，还能帮助教师明确协作目标与协作任务，助力协作知识建构的发生。

通过文献梳理发现，我们可以从师生教学行为、师生教学话语和学习者面部情绪生理特征等角度对课堂视频数据进行智能分析（蔡慧英,卢琳萌和顾小清，2023）。(1) 针对师生教学行为的智能分析技术。依据图神经网络的深度学习算法（如图卷积网络、图注意力网络、

图自编码器、图生成网络和图时空网络等），我们可以对视频中的教学行为进行特征提取和计算分析，从而量化表征和分析教师教学行为特点。（2）针对师生话语的智能分析技术。例如，语音识别中的基于动态时间规整的算法可以对连续语音识别分析，采用梅尔频率倒谱系数已修改可以实现对话语中的音调等时域特征的表征。（3）针对面部情绪生理特征的智能分析技术。借助卷积神经网络、多任务卷积神经网络等算法，我们可以通过人脸检测、面部表情检测、依据面部表情识别情绪状态等三个关键步骤，对视频中的人脸信息进行智能识别。

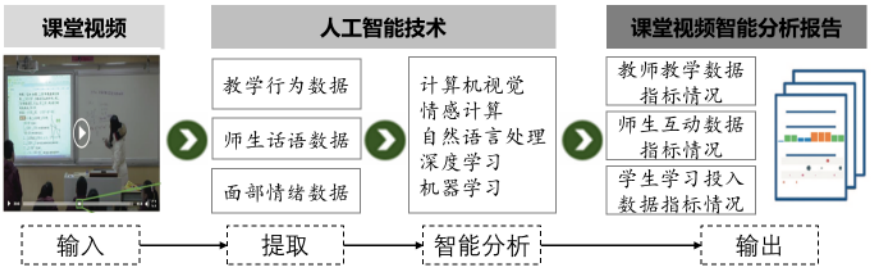


图 1 基于课堂视频智能分析技术的技术方案

因此，为了提升教师在基于课堂视频的教研过程中的交互质量，研究团队对课堂视频进行分析的具体技术方案如图 1 所示。首先，对课堂视频中的教学行为、师生话语、面部情绪等数据进行提取。接下来，借助平台的自然语言处理、深度学习等算法对视频进行智能分析。最后，将结果以课堂教学分析报告的形式呈现出来。

3. 研究问题

为了探究基于课堂视频人机协同教研教师交互模式，本研究开展智能技术辅助下的基于课堂视频的教研，通过收集教研过程中各种量化和质性数据，运用滞后序列分析方法，探究人机协同教研过程中教师交互行为模式特点。本研究意图探究的问题是：（1）在基于课堂视频的人机协同教研中，小组教师在不同类型活动中交互行为投入时间存在何种特点？（2）在基于课堂视频的人机协同教研中，小组教师在不同类型活动中交互模式存在何种特点？

4. 研究方法

4.1. 参与者与研究情境

为了探索基于课堂视频的人机协同教研过程中职前教师交互模式的质量，研究团队招募并组织了 48 名职前教师（其中 27 名来自教育技术学专业，21 名来自教育学专业）参加基于视频的专业发展对比实验。根据职前教师前测结果，研究团队将 48 名职前教师分为 16 组，每组有 3 名职前教师。其中，实验组情境下包括 8 个小组，控制组情境下包括 8 个小组。

4.2. 人工智能平台与实验实施过程

本研究基于人工智能网站“全景教研平台”分析课堂视频。全景教研平台的人工智能算法从教学行为、师生互动和学习投入等三个角度对课堂分析报告中的数据指标进行归类，还对数据指标的含义进行了详细说明，其界面示例如图 2 所示。报告对各数据指标与教学效果对学生学习的关系进行了一定的描述。这样，教师可以了解教学行为以及这些行为将如何影响学习者的学习，为教师与共同体内其他成员之间的交互提供了明确的切入点。



图 2 课堂视频智能分析报告部分示例

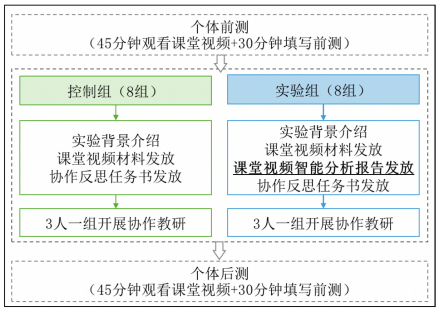


图 3 实验步骤流程图

在实验准备阶段，研究团队运用全景教研平台对苏教版七年级《信息技术》第三章第 1 节《收集与编辑数据》的课堂实录进行了智能分析，并形成相应智能分析报告。在小组协作教研阶段共包含三个主要环节。第一环节是个体前测。小组教师登录慕课平台观看 45 分钟的课堂视频，用约 30 分钟时间独立完成前测。第二环节是分别开展小组协作教研活动，如图 3 所示。在组织实验组教师开展协作教研活动中，研究团队向参与者介绍本次活动的背景。随后，小组教师登录金山文档查看协作研讨所需的课堂视频、实验准备阶段形成的课堂视频分析报告、教案、教研任务书材料，并在 QQ 上完成教研任务书中的问题讨论。最后，小组教师根据讨论结果生成一份优化教案。在组织控制组教师开展协作教研活动时，流程与实验组相同，但在发放教研相关材料时没有课堂视频分析报告。第三环节是个体后测。一周后，小组教师登录慕课平台完成课堂视频的观看，用约 30 分钟独立完成后测。

4.3. 数据收集和分析方法

本研究依据交互分析模型对收集到的小组的对话数据进行编码分析，并在编码结果的基础上进行滞后序列分析，以深入探究控制组和实验组实验情境下小组教师教研的交互模式。具体来说，本研究的数据收集与分析过程具体包括以下三个阶段。

阶段一，确定教师交互模式的编码框架。本研究依据 Gunawardena 等人提出的交互分析模型，在原有的五个阶段的基础上，对交互分析模型进行了情境化改编，通过审阅各个小组的对话数据，增加了“与学习无关的活动”、“团体计划”和“团体监控”三个交互阶段。最终改编后的交互模式编码框架共包含 8 个维度,如表 1 所示。

表 1 交互模式编码框架

交互分析维度	内涵	编码
与学习无关的活动	教师表达关心、鼓励，以及讨论与任务无关的话题。	P1
团体计划	教师关注协作教研任务如何分工，开展任务的方式等。	P2
团体监控	教师关注协作教研任务完成的进度和任务进展的快慢，当前任务是否完成，是否可以进入下一个任务。	P3
陈述教学信息	教师以观察者的角度描述教学设计中的教学元素、课堂视频中体现的教学事件。	P4
分享或比较对教学信息的理解	教师发表对课堂视频中的教学事件发表自己的感受和理解，表达赞同或者质疑。	P5
协商与知识构建	教师从教学效果的角度对各自的观点进行互动讨论。	P6
达成共识	教师达成对观点的共同认识。	P7
得出方案	教师得出改善课堂教学的方案和计划。	P8

阶段二，基于框架编码数据。本研究依据交互模式编码框架对教师对话数据进行编码，具体的编码过程分为以下三步。第一步，阅读熟悉小组教师对话数据，识别具有单一语义的对话片段。第二步，运用交互模式编码框架，对识别出来的对话片段进行逐段编码。第三步，统计教师在每个语义片段所投入的时间。根据一个对话片段的起始时刻和结束时刻，计算小组成员在每个行为所投入的时间。以上环节是在 3 位编码者经过编码培训之后进行，且每份数据均至少有两位编码者进行编码。编码者一致性系数超过 83%，编码结果具有可靠性。

阶段三，对编码结果进行滞后序列分析。滞后序列分析主要用于检验一种行为之后另外一种行为出现的概率及其是否存在统计学意义上的显著性。滞后序列分析会产生两个重要的表格以及一个重要的图示，分别是行为转换频率表、行为序列残差表与行为转换图。以上表格与图示能够实现行为因果关系分析与行为结构模式的可视化，有利于探索隐性的关键行为序列与结构（杨现民, 王怀波和李冀红, 2016）。因此，研究团队将编码结果导入到滞后序列分析的代表性工具 GSEQ5.0 软件中进行分析，进而生成相应的行为转换频率表、行为序列残差表以及行为转换图，以更清晰地探究控制组和实验组实验情境下小组的交互模式。

5.研究发现

5.1. 基于课堂视频的人机协同教研的交互行为投入时间差异分析

为了深入探究控制组和实验组小组教师在协作教研过程中交互行为投入时间的差异，研

究团队对控制组和实验组小组教师在协作教研过程中交互行为投入时间进行了统计与分析，其卡方分析结果如表 2 所示。控制组和实验组小组教师在交互行为中投入的时间存在着显著差异，X^2=2675.00，p=0.00<0.05。

表2 控制组和实验组小组教师在协作教研过程中交互行为投入时间的卡方分析结果

	P1	P2	P3	P4	P5	P6	P7	P8	X^2	df	p
控制组	5808	5154	1950	3720	16667	6303	2706	6440	2675.00	7	0.00
实验组	3549	1433	2851	1691	11822	6058	859	4246			

本研究对教师交互行为投入时间百分比情况进行了可视化，观察图 4 可以发现，首先，控制组和实验组小组教师均在分享或比较对教学信息的理解（控制组 p5，34.19%；实验组 p5，36.37%）上投入了大量时间，且所投入的时间均为控制组和实验组小组教师中占比最高的。这表明，教师在开展基于课堂视频的协作教研时，侧重分享或比较对教学信息的理解。其次，实验组小组教师在团体监控（p3）和协商与知识构建（p6）所投入的时间明显高于控制组小组教师。在协商与知识构建（p6）中，实验组小组教师所投入时间的占比（p6，8.77%）高于控制组小组教师（p6，4.00%）。在团队监控（p3）中，实验组小组教师所投入时间的占比（p6，18.63%）高于控制组小组教师（p6，12.93%）。这表明，当课堂视频分析报告介入时，教师的高层次交互行为总体频率占比会有所增加。最后，在团体计划（p2）中，实验组小组教师（p2，4.41%）投入时间占比明显低于控制组小组教师（p2，10.57%）。这表明，当课堂视频分析报告介入时，低层次的交互行为总体会有所降低。

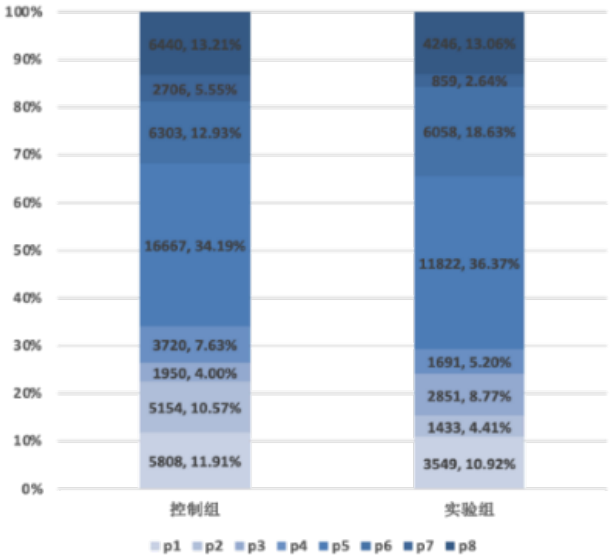


图 4 控制组和实验组小组教师在协作教研过程中交互行为投入时间百分比情况

综上，实验组和控制组基于课堂视频的协作教研的相同点在于开展分享或比较对教学信息的理解这一活动，且时长较长。控制组和实验组小组教师的交互行为差异主要体现在两个方面，一是在课堂视频分析报告的干预下，教师的高层次交互行为总体频率和总时长占比会有所增加；二是教师的低层次交互行为频率和总体时长会有所降低。

5.2. 基于课堂视频的人机协同教研的交互行为序列模式差异分析

为深入探究控制组和实验组小组教师在协作教研过程中交互模式，本研究借助 GSEQ 软件分析小组教师的交互行为序列模式，得到调整后残差表，见表 3。根据滞后序列分析理论，残差值 Z-score 大于 1.96 表示行为路径具有显著意义（即 p<0.05）。具体来说，Z-score＜1.96 称“不显著”；而 Z-score≥1.96 称“显著”，以*标记。

表 3 控制组和实验组小组教师交互行为序列残差表(Z-scores)

		p1	p2	p3	p4	p5	p6	p7	p8
控	**p1**	-0.49	2.06*	-0.66	-0.59	-1.12	-0.63	1.06	1.26
制	**p2**	0.1	-0.9	-0.8	0.98	1.18	-1.37	0.46	0.12
组	**p3**	-0.16	-1.65	-0.79	0.33	1.38	0.95	-1.06	0.22

	p4	-1.1	-0.69	-0.62	-1.33	4.31*	-1.43	-0.98	-0.47
	p5	-0.65	0.99	0.61	-0.04	-1.9	0.8	-0.26	1
	p6	-1.19	-0.86	1.49	1.82	-0.37	-1.54	-1.06	0.22
	p7	2*	2.12*	-1.09	-0.98	-1.06	1.22	-0.72	-1.02
	p8	2.02*	2.49*	0.13	-1.38	-2.62	1.93	0.85	-1.43
实验组	**p1**	-0.8	-0.67	-0.16	-0.67	0.65	1.38	-0.47	-0.11
	p2	0.1	-0.88	-1.61	0.44	0.59	-0.39	-0.61	1.85
	p3	-0.23	-0.6	-2.74	2.1*	0.74	0.94	0.21	0.24
	p4	0.21	-0.83	-0.64	-0.83	0.81	0.72	1.35	-0.58
	p5	-1.17	0.12	1.44	0.12	-0.67	1.98*	-1.12	0.48
	p6	-0.51	-0.14	0.49	-1.19	0.74	-1.06	2.07*	0.6
	p7	-0.65	-0.54	-0.99	-0.54	1.34	0.63	-0.38	0.35
	p8	1.56	2.5*	3.04*	-0.39	-2.72	-1.36	0.4	-2.39

依据表 3 的结果，研究团队绘制出了控制组小组教师和实验组小组教师交互行为转换图，如图 5 所示。所有重要的行为序列在行为转换图中都用箭头表示。圆圈代表行为类别，数值代表调整后的残差值（Z-score），箭头代表过渡方向，粗细代表行为连接的显著性。从行为过渡图中可以看出，在控制组和实验组小组中，存在着不同的重要行为序列。

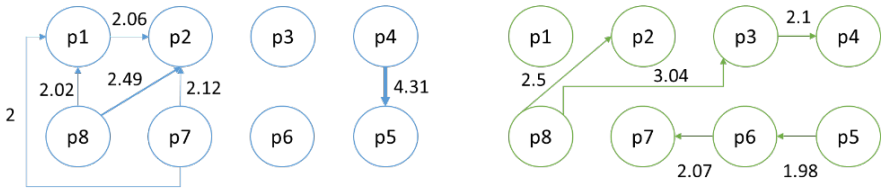


图 5 控制组小组教师和实验组小组教师交互行为转换图（左图为控制组、右图为实验组）

分析发现，实验组小组教师和控制组小组教师在协作教研的交互模式上存在的相同点主要有以下两点。第一，控制组小组教师和实验组教师都包含同一条交互路径，即“得出方案→团体计划”（p8→p2）。两组情境下的小组教师在讨论得出某一问题的解决方案后，均会开展团体计划行动，关注后续的协作教研任务如何分工，以推动进展。第二，控制组小组教师和实验组教师都没有以下两条交互路径，分别是“团体计划→团体监控”（p2→p3）路径，与达成共识→得出方案“（p7→p8）路径。两组的教师都不会在明确教研任务分工、开展教研的方式后，即刻开展任务监控，也都不会在达成观点的共同认识后，立刻得出改善课堂教学的方案和计划。

控制组和实验组小组教师在协作教研的交互模式上的差异主要有以下四点。第一，“陈述教学信息→分享或比较对教学信息的理解”（p4→p5，Z-score=4. 31）是控制组小组教师特有的具有显著意义的行为路径。这表明，控制组小组教师会在陈述教学信息后，发表对课堂视频中教学事件的感受和理解。而实验组小组教师则会基于课堂视频智能分析报告，直接发表感受并相互赞同或质疑。无需经历陈述教学设计中的教学元素这一环节。第二，“得出方案→团体监控”（p8→p3，Z-score=3. 04）是实验组小组教师特有的具有显著意义的行为路径。且“得出方案→团体监控”（p8→p3，Z-score=3. 04）是实验组小组教师特有的具有显著意义的行为路径。这表明，实验组小组教师会在得出方案后进一步关注协作任务的进度与进展，关心任务的推进。而控制组小组教师则缺少对任务的持续性监控。第三，控制组小组教师涉及到推理过程的交互路径只有一条是有显著意义的，即“陈述教学信息→分享或比较对教学信息的理解”（p4→p5，Z-score=4. 31）。教师在陈述教学事件或者教学元素相关信息之后，教师之间会就这些信息发表自己的感受和理解。而实验组小组教师具有显著意义的推理过程的交互路径有 3 条，分别是“团体监控→陈述教学信息”（p3→p4，Z-score=2. 1）、“分享或比较对教学信息的理解→协商与知识构建”（p5→p6，Z-score=1. 98）、“协商与知识构建→达成共识”（p6→p7，Z-score=2. 07）。第四，实验组小组教师则没有出现有显著意义的与“与学习无关的活动”相关的行为路径。而控制组小组教师有 4 条与“与学习无关的活动”的行为路径具有显著意义，分别是“与学习无关的活动→团体计划”（p1→p2，Z-score=2. 06）、“达成共识→与学习无关的活动”（p7→p1，Z-score=2）、“得出方案→与学习

无关的活动”（p8→p1，Z-score=2. 02）。

6.研究结论与启示

6.1.研究结论

对于“研究问题（1）：在基于课堂视频的人机协同教研中，小组教师在不同类型活动中交互行为投入时间存在何种特点？”，本研究发现教师前后两次所体现的交互行为投入时间表现出显著差异。在有课堂视频智能分析报告的协作教研中，教师会投入更多的时间从教学效果的角度对各自的观点进行互动讨论,且教师在团体监控方面投入更多的时间，注重协作教研任务完成的进度和任务进展的快慢，当前任务是否完成，是否可以进入下一个任务。相反地，控制组教师在团体计划方面投入更多的时间，关注协作教研任务如何分工，开展任务的方式等。这表明，在课堂视频分析报告这一干预条件下，教师的高层次交互行为总体频率和总时长占比会有所增加，低层次交互行为频率和总体时长会有所降低。

对于“研究问题（2）：在基于课堂视频的人机协同教研中，小组教师在不同类型活动中交互模式存在何种特点？”，本研究发现基于课堂视频的人机协同教研中，职前教师交互模式具有任务衔接度好、任务投入度高、任务灵活性强的特点。第一，课堂视频智能技术赋能的小组教师交互模式具有任务间衔接度好的特点。控制组会出现小组教师进行团体计划，将教师从无关活动中拉到本次教研任务当中的情况。对比之下，在课堂视频分析报告这一干预条件下，实验组小组教师在一个任务得出方案之后会直接积极投入下一个任务的安排和规划当中。这表明，课堂视频智能分析报告上的内容能够促进教师主动思考并投入到接下来的任务。第二，课堂视频智能技术赋能的小组教师交互模式具有任务投入度高的特点。控制组小组教师达成共识或者得出方案之后，有的教师会注意力不集中，出现倦怠情绪，出现与学习无关的活动。对比之下，在课堂视频分析报告这一干预条件下，实验组小组教师在得出方案之后会关注协作教研任务完成的进度，对当前任务是否完成以及是否可以进入下一个任务进行流程上的把控，这表明课堂视频智能分析报告上的内容能够提高教师协作教研活动的投入度。第三，课堂视频智能技术赋能的小组教师交互模式具有任务灵活性强的特点。控制组小组教师会基于经验按照陈述教学设计中的教学元素、表达赞同或者质疑的顺序展开教研。对比之下，在课堂视频分析报告这一干预条件下，实验组小组教师可以直接基于报告中呈现的信息发表感受，进而从教学效果的角度对观点展开互动。这表明课堂视频智能分析报告上的内容能够促进教师灵活开展教研，并提高交互的效率。

6.2.研究启示

从教师交互模式干预的角度出发，相关人员在组织教师开展智能技术辅助的教研时需要关注教师交互的数据使用、推理过程及具体路径。

第一，关注教师交互的数据使用过程，给予教师个性化的精准帮扶与指导，促使教师将智能分析结果内化于心。虽然智能技术可以帮助教师收集和分析课堂教学数据，并提供课堂数据智能分析报告，但这些技术仅仅是工具而已。若想激发智能技术的最大潜能，研究人员则需要关注教师智能分析数据使用的过程，从理论上提出干预措施，为教师提供例如同伴合作等实时指导。若教师把课堂视频智能分析报告中的数据转化为有教育价值的信息和知识，并最终内化为头脑中的智慧。那么，教师协作知识建构的成效可以得到最大化发展。

第二，关注教师交互的推理过程，引导教师开展推理性和论证性交流，促使教师展开深度思考。在基于课堂视频的人机协同教研的交互模式中，教师并非单纯倾听和接受彼此的观点想法，而是通过共同探讨、推理和论证来支持或者反对对方的观点。这种推理论证的交互模式可以帮助教师思考和评估事物表面下的意义和真相，促进教师的协作知识建构。为了进一步激发教师的深度思考，相关人员需要关注教师交互的推理过程，将教学干预措施上升到实践层面，为教师提供例如引导性问题等支架。

第三，关注教师交互的具体路径，提供对教师教研过程的动态监管，促使教师实现持续性自我监控。在基于课堂视频的人机协同教研的交互模式中，教师仍缺乏计划与监控意识。

从控制论的视角来看，获取教研交互模式的根本目的在于调控和优化教研的组织和过程。因此，教师交互质量反馈的及时性、准确度将决定决策的科学程度。在智能精准教研中，相关人员需要基于对教研数据的实时掌握和智能分析，对教师教研全过程进行实时监控，并借助监测数据做出科学决策。

6.3.研究不足与展望

本研究创设并实施了人工智能辅助下教师基于课堂视频的教研实验，这从教研效果的角度验证了课堂视频智能技术对于教研交互质量提升具有积极作用，并提出了激发人工智能赋能教师基于课堂视频教研过程的实施策略。但本研究也存在一定局限，如参与实验的样本数量较少、性别比例不够均衡、实验周期较短等。为此，研究团队将在现有研究的基础上，扩大研究样本和研究周期，进一步揭示基于课堂视频的人机协同教研提升教师交互质量的机制。

参考文献

蔡慧英,卢琳萌,顾小清.人机协同教研会促进教师教学反思能力的发展吗?——基于课堂视频智能分析技术的实证研究[J].现代远距离教育,2023,(01):40-49.DOI:10.13927/j.cnki.yuan.20230327.001

魏宁. (2006). 协作知识建构: 教师博客群发展的新视角. 中小学信息技术教育, 11, 52-54.

杨现民,王怀波,李冀红.(2016).滞后序列分析法在学习行为分析中的应用.中国电化教育(02),17-23+32.

赵健. (2006). 学习共同体: 关于学习的社会文化分析. 华东师范大学出版社.

Chen, P., Lee, C. D., Lin, H., & Zhang, C. X. (2018). Factors that develop effective professional learning communities in Taiwan. In Global Perspectives on Developing Professional Learning Communities (pp. 56-73). Routledge.

Driver, R. (1994). Making sense of secondary science: Research into children's ideas. Routledge.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. Journal of educational computing research, 17(4), 397-431.

Hamel, C., & Viau-Guay, A. (2019). Using video to support teachers' reflective practice: A literature review. Cogent Education, 6(1), 1673689.

Henri F. Computer conferencing and content analysis[C]//Collaborative learning through computer conferencing: The Najaden papers. Springer Berlin Heidelberg, 1992: 117-136.

Hu, B. Y., Guan, L., LoCasale-Crouch, J., Song, Z., Dou, L., Li, S., ... & Zhang, X. (2023). Effects of using video-based coaching to promote preservice teachers' interactional skills in Chinese preschool classrooms. Early Childhood Research Quarterly, 65, 284-294.

Kang, H., & van Es, E. A. (2019). Articulating design principles for productive use of video in preservice education. Journal of Teacher Education, 70(3), 237-250.

Ro, J. (2017). A journey from the classroom to the world of educational reform: a study of three Korean teachers' practitioner inquiry. Asia Pacific Journal of Education, 37(1), 28-41.

Scardamalia M, Bereiter C. Knowledge building[J]. The Cambridge, 2006.

So, H. J., Lim, W., & Xiong, Y. (2016). Designing video-based teacher professional development: Teachers' meaning making with a video annotation tool. Educational Technology International, 17(1), 87-116.

Van Gasse, R., Vanlommel, K., Vanhoof, J., & Van Petegem, P. (2017). Unravelling data use in teacher teams: How network patterns and interactive learning activities change across different data use phases. Teaching and Teacher Education, 67, 550-560.

Vigotsky, L. S. (1939). Thought and speech. Psychiatry, 2(1), 29.

Yang, X., Li, J., Guo, X., & Li, X. (2015). Group interactive network and behavioral patterns in online English-to-Chinese cooperative translation activity. The Internet and Higher Education, 25, 28-36.

ChatGPT 与提问脚手架对师范生信息化教学设计的影响：一项混合研究

Impact of ChatGPT and Prompt on Pre-service Teachers' Technology-Enhanced Learning Design: A Mixed Method Study

李荣娜[1]，陈辉[1]，陈琦帆[1]，田金铭[1]，杨九民[1*]
[1]华中师范大学 人工智能教育学部，湖北 武汉 430079
[*]yjm@mail.ccnu.edu.cn

【摘要】本研究探究了 ChatGPT 和提问脚手架对师范生开展信息化教学设计和学习体验的影响。研究选取了中国中部某师范大学选修《现代教育技术应用》课程的 57 名师范生，采用混合研究设计，被试随机分配至两种实验条件之一：（1）ChatGPT 与提问脚手架结合组（n=37）；（2）仅 ChatGPT 支持组（n=21）。通过对师范生的信息化教学设计方案的编码评分、以及后测问卷和访谈数据的分析，结果表明，相较于仅 ChatGPT 支持，ChatGPT 与提问脚手架结合能够显著提升师范生的信息化教学设计作品质量，并降低其在设计过程中的外在认知负荷。访谈结果也说明，提问脚手架能够提升师范生与 ChatGPT 的对话效率和对话质量。本研究对于 ChatGPT 应用于师范生教育提供了实践启示：教师应设计良好的提问脚手架，帮助师范生与 ChatGPT 互动中取得更好的学习体验和学习效果。

【关键词】ChatGPT；提问脚手架；信息化教学设计；教师专业发展

Abstract: This study explores the effects of ChatGPT and prompt on improving the pre-service teachers' technology-enhanced learning design ability and learning experiences. Fifty-seven pre-service teachers taking a course named Modern Educational Technology at a normal university in central China were selected for the study, and a mixed research design was used in which participants were randomly assigned to one of two conditions: (1) a combination of ChatGPT and prompt group; and (2) a ChatGPT-only group. By analyzing the coded scores of the pre-service teachers' technology-enhanced learning design artifacts, as well as the data from the posttest questionnaire and interviews, the results indicated that, compare with the ChatGPT support only, the combination of ChatGPT and prompt significantly improved the quality of the pre-service teachers' technology-enhanced learning design artifacts and reduced their extrinsic cognitive loads during the design process. Moreover, the interview results also illustrated that prompt can enhance the efficiency and quality of pre-service teachers' interactions with ChatGPT. This study provides practical implications for the application of ChatGPT to teacher trainee education: teachers should design good prompts to help pre-service teachers achieve better learning experiences and learning outcomes in their interactions with ChatGPT.

Keywords: ChatGPT, prompt, technology-enhanced learning design, teacher professional development

1. 引言

2022 年 11 月，OpenAI 推出了 ChatGPT，它能够智能响应提示和问题，并以类人对话的方式与用户进行交互。2023 年 9 月，联合国教科文组织发布的《生成式人工智能教育和研究应用全球指南》中提到，要培养教师和研究者利用生成式人工智能的能力，同时可以设计特定的基于生成式人工智能的辅助工具，促进学生课堂学习和教师专业发展（UNESCO, 2023）。

在生成式人工智能技术应用于教师专业发展方面，有研究指出，教师可以在教学活动的特定阶段有针对性、有策略性地使用 ChatGPT 辅助其更好地开展教学实践（Barrot, 2023）。从“课前-课中-课后教学评一体化”的视角来看，在课前阶段，教师可以使用 ChatGPT 辅助其进行备课工作，包括辅助生成教学设计方案，制作课堂所需的评价量规等教学材料；在课中阶段，教师可以使用 ChatGPT 作为课堂智能助教，构建 AI 双师课堂，对学生在学习活动中的问题予以即时反馈和帮助；在课后阶段，教师可以使用辅助其对课堂教学效果进行评价（吴军其等，2023）。本研究则聚焦于 ChatGPT 辅助教师的课前备课，具体而言，旨在探究 ChatGPT 对于师范生的信息化教学设计的影响。

研究表明，为了更高效的使用 ChatGPT 辅助开展教学和学习，同时进一步提升 ChatGPT 生成的反馈信息的质量，使用者应当扮演好“提问设计者”角色（王丽等，2023）。具体而言，教师和学习者需要结合实际情况和需求提前做好提问设计，通过指明所问对象、不断追问、设定角色等优化策略对提问设计进行迭代优化，以确保提问设计的针对性和实用性。例如，可以通过设计良好的提问脚手架、提示语的方式来充分发 ChatGPT 工具的教育潜能，使其更为准确地向使用者生成反馈信息（赵晓伟等，2023）。然而，目前还未有研究针对课前环节，探索 ChatGPT 辅助教师进行备课的效果，并通过开展实证研究探究相关的提示词设计的应用效果。因此，本研究旨在针对“教师利用 ChatGPT 生成信息化教学设计”这一人机交互活动，设计提问脚手架，并通过开展一项混合研究，探究该提问脚手架对于师范生的信息化教学设计作品质量和学习体验的影响。基于此，本研究研究问题如下：

（1）相较于仅 ChatGPT 组的师范生，ChatGPT 与提问脚手架结合组的师范生制作的信息化教学设计作品的质量是否更高？

（2）相较于仅 ChatGPT 组的师范生，ChatGPT 与提问脚手架结合组的师范生在学习过程中是否有更好的学习体验？

2. 研究设计

2.1. 研究对象

本研究在中国中部某师范大学的一门教育教师类课程——《现代教育技术应用》课程开展。该课程的目标之一是提升师范生的信息化教学设计能力。研究采用方便抽样法，招募 2023 年秋季学期《现代教育技术应用》某课堂的 58 名历史学专业大三年级的师范生作为研究对象，其年龄在 19~24 岁之间（$M = 20.31$）。研究预调查显示，所有被试在大二年级均选修过教育学、心理学和学科教学论等先行课程，具备基础的教学设计能力，然而，在信息化教学设计能力方面有所欠缺。在研究开始前，被试均签署了知情同意书，知晓是否参与本研究与课程成绩无关。

2.2. 实验设计

研究采用随机对照组间设计，被试被随机分配至两组：（1）实验组（n=37）：ChatGPT 与提问脚手架结合辅助师范生进行信息化教学设计；（2）对照组（n=21）：仅 ChatGPT 辅助师范生进行信息化教学设计。采用不等组设计的原因是受限于实验环境的限制，对照组所在的实验机房仅有 24 台计算机，因此仅能支持 24 名师范生同时开展信息化教学设计。研究因变量包括：师范生设计的信息化教学设计作品质量、信息化教学设计能力测验成绩和学习体验。学习体验包括认知负荷和师范生对 ChatGPT 的技术接受度。

两组的师范生均被要求在 2 个课时（90min）内独立完成一个信息化教学设计方案的学习任务。在学习任务开始前，助教向所有师范生提供了完全一致的辅助材料，包括信息化教学设计模板、两个优秀的信息化教学设计案例、以及信息化教学设计评价量规。并且，每个被

试随机分配了一个 ChatGPT3.5 独立账号。两组唯一的区别是，助教为实验组的师范生提供了基于 ChatGPT 开展信息化教学设计的提问脚手架，而对照组则没有提问脚手架的支持。

2.3. 研究流程

本研究依托《现代教育技术应用》课程第十章节“信息化环境下的教学设计”开展，实验环境为两个实验教学专用机房，实验组和对照组随机分配在两个不同的机房内完成学习任务。实验流程如下：实验开始的前一周，教师在课堂 QQ 群中通过问卷星发放实验知情同意书和前测问卷，并提醒师范生自愿填写。第二周的课中环节，教师向师范生们介绍本次实验的学习任务：自选教学主题，在 ChatGPT 以及助教提供的辅助材料支持下独立地完成一份信息化教学设计方案，限时 90 分钟。在课堂结束后，学生通过机房局域网向教师提交信息化教学设计作品，并填写后测问卷。实验结束后，研究者在两组中分别随机抽选了 10 名师范生开展线上访谈。实验流程如图 1 所示。

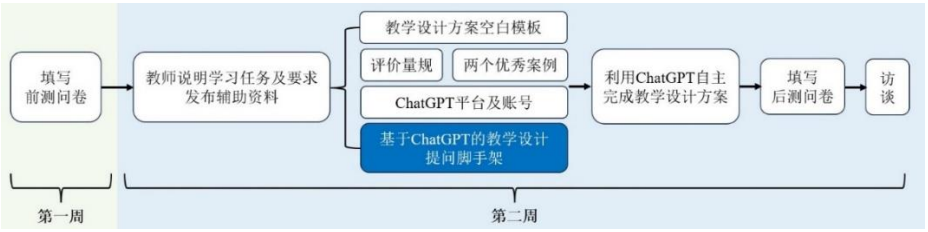


图 1 研究流程图

2.4. 测量工具

2.4.1. 前测问卷

包括人口统计学信息（性别、年龄等）以及对 ChatGPT 工具的了解程度和使用意愿。问卷采用五点李克特量表的形式（1=完全不符合，5=完全符合），其中，师范生对 ChatGPT 工具的了解程度包含一个题项，“你对 ChatGPT 以及相关的生成式人工智能产品的了解程度”；使用意愿包含三个题项，例如：“作为一名未来教师，我愿意使用 ChatGPT 类生成式人工智能产品来辅助我备课”。

2.4.2. 基于 ChatGPT 的信息化教学设计提问脚手架

本研究在课堂任务过程中为师范生提供了教学设计模板，包括学习者分析、教学内容分析、教学目标设计、教学环境和媒体选择、教学模式和教学策略、教学过程、教学评价设计和教学反思八个部分。按照信息化教学设计的基本流程，本研究为实验组的被试设计并提供基于 ChatGPT 的信息化教学设计提问脚手架，如图 2 所示。

依据提问脚手架的提示，师范生可以通过向 ChatGPT 提出问题并逐步完善教学设计，在每一步提问中都包括三个部分：提问模板、提问示例和模拟演示。提问模板是研究者提前设计好的提问结构，并为师范生提供可能得提问关键词，提问示例则是一个具体的提问问题。例如，在选择教学模式和策略环节，师范生需要依据确定的教学内容和教学对象，向 ChatGPT 提出选取某一种教学模式和策略，并以此为依据进一步详细优化教学过程。通过基于 ChatGPT 的教学设计提问脚手架的引导，师范生可以在每一步针对性的向 ChatGPT 提出问题，调整提问方式，整合信息并优化整体信息化教学设计。



图 2 基于 ChatGPT 的教学设计提问脚手架

2.4.3. 信息化教学设计评价量规

信息化教学设计的评价量规由研究人员和课程讲师共同开发。根据 Tan & Chen（2022）的研究中包含的评价标准，信息化教学设计方案的评价量规应包括针对所有三个部分的评价指标（即，内容，教学和技术知识），并强调这三个组成部分之间的相互作用。因此，针对信息化教学设计方案的评价量规由四个维度组成：（1）前端分析（针对内容知识的三个项目；例如："学习需求分析有可靠依据；教学内容分析能够充分揭示各部分之间的逻辑关系，体现重难点；"）；（2）学习目标与学习评价设计（针对教学知识的四个项目，例如："学习目标分类合理;学习目标的描述具体、可观察和可测量;学习评价体现学习过程和学习结果;"）；（3）信息化教学过程设计（针对内容、教学和技术知识综合设计的四个项目，例如："采用的教学方法/策略能够支持教学内容的呈现，能够促进学生对教学内容的理解；"）（4）信息化教学环境设计（针对技术知识的四个项目，例如："对环境中的技术工具的教学功能和特性描述清晰、理解到位；技术工具的使用能够支持和促进教学活动的开展等"）。

2.4.4. 后测问卷

后测问卷针对师范生在 ChatGPT 支持的条件下开展信息化教学设计过程中的学习体验进行调查，包括技术接受度和认知负荷两个部分，依据 Zhang et al.,（2023）较为成熟的问卷，采用五点李克特量表，分别从绩效期望、感知易用性、感知风险、外在认知负荷、内在认知负荷和关联认知负荷等六个维度，共 18 个题项。问卷中六个维度的信度（Cronbach' α 系数）分别为 0.784、0.732、0.749、0.855、0.843、0.841，表明量表信度较高，通过该量表能较真实可靠地反映师范生对 ChatGPT 的技术接受度和开展信息化教学设计过程中的认知负荷。量表选项从 1 到 5 依次代表"非常不赞同"到"非常赞同"。其中，绩效期望包含 5 个题项，是指师范生认为通过 ChatGPT 对自身提升信息化教学设计能力的影响程度，如"使用 ChatGPT 可以丰富和补充我的观点和想法"等，当个体对该方式的绩效期望越高时，其对 ChatGPT 技术的接纳意愿也将越明显；感知易用性包含 5 个题项，是指师范生利用 ChatGPT 辅助完成信息化教学设计方案的过程中感知的便利条件，如"我能够掌握高效使用 ChatGPT 的知识和技巧"等，当个体感知易用性越高，越会对该技术产生较高接受度；感知风险包含 5 个题项，是指师范生在使用 ChatGPT 时可能会产生的担忧和疑惑，如"经常使用 ChatGPT 会降低我的批判性思维能力"，感知风险越高，师范生对持续性使用该项技术的意愿会越低。认知负荷维度包含 3 个题项，依据 Leppink et al.,（2014）的研究，分别从内在认知负荷、外在认知负荷和关联认知负荷三个方面测量师范生在开展信息化教学设计任务时的认知负荷。

2.4.5. 访谈

为综合分析基于 ChatGPT 的教学设计提问脚手架对师范生信息化教学设计能力的影响，本研究采用访谈法了解学生使用 ChatGPT 的详细过程、提问方式以及体验感受，包括对

ChatGPT 辅助信息化教学设计的态度层面、使用层面、困难问题层面、建议层面等 4 个维度（见表 1）。

表 1 访谈提纲

访谈维度	访谈内容
态度层面	1.你觉得在你做这个信息化教学设计的过程中，ChatGPT 扮演什么样的角色？ 2.你怎样看待使用 ChatGPT 这个工具来辅助师范生进行信息化教学设计这件事情？ 3.你觉得 ChatGPT 能帮助/辅助到你进行信息化教学设计吗？如果能，请说说你觉得它帮助了你哪些方面？
使用层面	1.在这堂习题课中，你是如何使用 ChatGPT 来辅助你进行信息化教学设计的？ 2.除了使用 ChatGPT,你还有没有参考了哪些其他的资料或资源来帮助自己更好地进行信息化教学设计呢？请你展开说说。
困难问题层面	1.你在使用 ChatGPT 来辅助你进行信息化教学设计的过程中，你有遇到什么问题或者困难或者困惑的地方吗？
建议层面	1.你有什么建议或者想法,能够更好地使用 ChatGPT 帮助自己进行信息化教学设计？ 2.你觉得老师和助教学姐/学长还需要给你提供什么样的帮助或支持来帮助你更好、更高效地利用 ChatGPT 进行信息化教学设计？

2.4.6. 信息化教学设计编码框架

本研究采用已被证实有效的 TPACK 方案来评估师范生的信息化教学设计质量（Zhang et al., 2022）。TPACK 方案包括七个维度：内容知识（CK），教学知识（PK），技术知识（TK），技术教学知识（TPK），教学内容知识（PCK），技术内容知识（TCK）和技术教学内容知识（TPACK）进行分级。与以往研究一致，信息化教学设计方案中 TPACK 维度均采用 5 分制测量，范围从 1（最低质量）到 5（最高质量），信息化教学设计方案的最高得分为 35 分。由两名接受过培训的评分员分别使用编码方案对所有信息化教学设计方案进行独立评分。7 个 TPACK 指标的评分者间信度（Cronbachα 系数）分别为 CK：0.946、PK：0.874、TK：0.750、TPK：0.819、PCK：0.835、TCK：0.704 和 TPACK：0.838。考虑到评分者间的高可靠性，两名评分者的平均得分被视为每个师范生信息化教学设计方案的最终得分。

2.5. 数据分析

本研究的数据包括（1）依据 TPACK 评价框架，师范生在 ChatGPT 支持下自主完成信息化教学设计方案的作品质量；（2）依据 TPACK 评价框架，师范生在期末测试中自主完成信息化教学设计方案的作品质量；以及（3）师范生在 ChatGPT 支持的条件下开展信息化教学设计过程中的技术接受度和认知负荷表现。首先，为排除组间差异对实验效果的影响，对实验组和对照组的师范生对 ChatGPT 技术的以往使用经验和使用意愿分别进行独立样本 t 检验。其次,为了比较实验组和对照组在 ChatGPT 支持下自主完成信息化教学设计方案的作品质量，基于 TPACK 评价的总分和 7 个子维度的得分进行了一组独立样本 t 检验。两组数据符合正态分布和等方差假设(正态性假设的 Shapiro-Wilk 检验:$ps > .001$;等方差假设的 Levene 检验:$ps > .05$），满足开展独立样本 t 检验的条件。

3. 研究结果

为了解决前面提到的研究问题，本研究主要采用独立样本 t 检验的方法对于收集到的数据进行分析。在正式分析之前，本研究使用 Shapiro-Wilk 检验分别评估被试个体特征、信息化教学作品设计分数、认知负荷和 ChatGPT 接受度情况等数据的正态性，数据均符合满足正态性和方差齐性检验，满足使用独立样本 t 检验进行组间比较的前提条件。

3.1. 个体特征

为了排除一些无关变量的干扰，首先对师范生对 ChatGPT 类生成式人工智能产品的了解程度和使用意愿进行了独立性检验。结果表明（见表 2），实验组和对照组在两个方面没有显著差异。

3.2. 信息化教学设计作品的质量

为了探究 ChatGPT 与提问脚手架对师范生的信息化教学设计能力的影响，本研究对师范生的信息化教学设计作品的整体得分和 7 个子维度（即 CK，PK，TK，TPK，TCK，PCK 和 TPACK）分别进行了独立样本 t 检验。结果显示（见表 2），实验组的师范生在教学设计作品的整体得分和 3 个子维度（即 PK，TPK，TPACK）的得分显著高于对照组的师范生。更重要的是，TPACK 的效应量中等，PK，TPK 和 Total 的效应量较大。这些结果均表明，与对照组相比，使用提问脚手架有助于提升师范生信息化教学设计作品的质量。

表 2 描述性统计结果

因变量	实验组 *Mean (SD)*	对照组 *Mean (SD)*	*t*	*p*	*Cohen's d*
前测变量					
对 ChatGPT 的了解程度	5.49（1.05）	5.08（1.34）	1.29	0.204	0.34
对 ChatGPT 的使用意愿	4.46（1.24）	3.90（0.94）	1.78	0.081	0.51
信息化教学设计能力					
CK（内容知识）	4.11（0.72）	3.81（0.62）	1.60	0.116	0.45
PK（技术知识）	4.19（0.66）	3.57（0.80）	3.18**	0.002	0.85
TK（教学知识）	3.84（0.60）	3.64（0.76）	1.08	0.286	0.29
TPK（技术教学知识）	4.16（0.50）	3.67（0.70）	2.87**	0.003	0.81
TCK（教学内容知识）	3.85（0.53）	3.71（0.72）	0.84	0.407	0.22
PCK（技术内容知识）	4.76（0.40）	4.60（0.60）	1.22	0.227	0.31
TPACK（技术教学内容知识）	4.26（0.36）	3.99（0.54）	2.23*	0.03	0.59
总分	29.16（2.49）	26.99（3.26）	2.85**	0.006	0.75
认知负荷					
外在认知负荷	5.76（1.92）	6.65（1.94）	-1.45	0.153	-0.46
内在认知负荷	2.76（1.48）	4.85（1.79）	-4.78***	<0.001	-1.27
关联认知负荷	7.28（1.44）	7.16（1.49）	0.29	0.114	0.08
对 ChatGPT 技术的接受度					
绩效期望	4.1（0.51）	4.05（0.49）	0.71	0.479	0.20
感知易用性	3.47（0.61）	3.32（0.72）	0.82	0.414	0.22
感知风险	3.22（0.12）	3.50（0.14）	-1.46	0.151	1.25

3.3. 信息化教学设计过程的学习体验

3.3.1. 认知负荷

在 ChatGPT 与提问脚手架对师范生认知负荷的影响方面，本研究对师范生内在认知负荷，外在认知负荷和关联认知负荷分别进行了独立样本 t 检验。结果显示，实验组的外在认知负荷得分显著小于对照组的外在认知负荷得分，而实验组和对照组在内在认知负荷、关联认知负荷两个方面没有显著差异。结果表明，和对照组相比，使用提问脚手架能够帮助师范生减少外在的认知负荷。

3.3.2. 技术接受度

为了进一步了解加入提问脚手架对师范生使用 ChatGPT 的技术接受度的影响，首先，我们对绩效期望、感知易用性和感知风险三个维度分别进行了独立样本 t 检验。结果显示，实验组和对照组在绩效期望、感知易用性和感知风险三个方面均不存在显著差异。结合对师范生的后期访谈，虽然实验组被提供了提问脚手架，但是他们对于除了模板示例以外的提问感到困惑。我们猜测，这可能也是影响提问脚手架在 ChatGPT 辅助的信息化教学设计的效果的因素之一。实验组#6 师范生表示：“在习题课中的教学设计中，我第一次接触并利用 ChatGPT 进行教学设计，有点依赖 ChatGPT，有被它牵着走的感觉，没有自己连贯的逻辑思路。”

在感知风险方面，实验者的师范生表示：虽然使用 ChatGPT 提问脚手架的模板进行提问可以辅助信息化教学设计提供设计思路并提高设计效率，但是他们依旧对提问的有效性和正确性感到担忧。实验组#1 师范生表示：“在我进行信息化教学设计的过程中，ChatGPT 充当一个辅助者的身份。它能根据我提出的问题进行相关设计并回复，但是因为他的回答比较客观，属于任何课题都可以参考的类型，所以还需要根据自己的实际情况对回答进行修改完善，并不是完全可靠的。”

4. 讨论

为了探究提问脚手架对师范生开展 ChatGPT 辅助的信息化教学设计的影响，本研究对两组师范生的信息化教学设计方案和后测问卷数据进行了描述性统计，深入比较有无提问脚手架的条件下师范生的信息化教学设计方案和学习过程体验。

对于研究问题一，相较于无提问脚手架组的对照组，通过完成信息化教学设计的学习任务，有提问脚手架的实验组生成的信息化教学设计作品质量更高。这意味着，相比于对照组，实验组的信息化教学设计能力提升更多。具体表现在信息化教学设计方案的整体得分和 3 个子维度（PK，TPK，TPACK）上。这一结果与以往研究一致（Kumar., 2021），师范生在利用 ChatGPT 开展教学设计的过程中能够通过向 ChatGPT 探寻观点和发表观点，依据自身需求和想法开展个性化学习（马卓等，2023）。对于实验组在信息化教学设计的技术知识方面的突出表现，可能与本研究的被试的专业学科背景（均来自历史学科）有关，他们在以往学习中并未具备较为丰富的结合现代教育技术开展教学设计和实施的经验，因此对于如何将技术有效融入课堂教学策略选择和教学环节实施中缺乏想法，这一点也在后续的访谈中得到证实。而在教学设计中 ChatGPT 与提问脚手架的使用恰好能弥补这一点：ChatGPT 能够结合具体学科和已有社会实践经历，搜集信息和梳理想法，并介绍数字化学习工具的特点、功能和用途，进而帮助设计者更好的筛选和选择可用于教学过程中的技术和工具。

对于研究问题二，相较于无提问脚手架组的对照组，有提问脚手架的实验组在信息化教学设计过程中的认知负荷更低，在技术接受度的绩效期望、感知易用性和感知风险等方面没有显著差异。首先，在认知负荷方面，实验结果表明向师范生提供提问脚手架在一定程度上能够减轻师范生独立完成信息化教学设计的压力，他们能够更多的关注 ChatGPT 生成的内容质量。在后续访谈中，对照组中部分师范生表示由于 ChatGPT 并不总是能够理解给出的指令，因此在与 ChatGPT 交流的过程中存在交流与理解困难，需要设计者多次提出设计要求和改进意见。除此之外，也有实验组的师范生表示除了被给出关于信息化教学设计的特定问题脚手架外，他们也希望能够学习更多关于与 ChatGPT 有效沟通的提问技巧。其次，在技术接受度方面，值得注意的是，虽然实验结果并未呈现显著差异，但是从平均值来看，相比于对照组，实验组在绩效期望和感知易用性的表现较高，但同时也会感知更多的使用风险。结合后期访谈内容分析，部分实验组的师范生表示在有提问模板的情况下，他们会更相信 ChatGPT 响应的正确性，然而对于提问脚手架没有提及到的部分，他们并不是总能自行开展有效提问。

基于以上研究结果，本研究提出开展 ChatGPT 辅助的信息化教学设计的几点应用启示：首先，教师可以通过向学习者提供提问脚手架来提升 ChatGPT 辅助学习的应用效果。其次，为了更好的提升 ChatGPT 辅助学习的学习体验，可以预先向学习者提供一些有效沟通的技巧和方法，更有助于学习者了解该技术，进而能够有针对性和目的性的使用。

本研究也有几个局限性需要在未来研究中深入考虑和解决。首先，本研究的被试是从单一课程中采用方便抽样的方式选取的，因此，实验结果可能不能够代表全学科的师范生群体。其次，本研究的样本量相对较小，因此本研究的结果需要谨慎解读，未来的研究可以采用更大的样本量和从 ChatGPT 辅助学习的行为视角讨论，以增强研究结果的可信度和可推广性。

5. 参考文献

王卓，马洋珍，杨现民等（2023）。ChatGPT 类阅读平台对研究生学术阅读能力的影响[J]。开放教育研究， 29(06)：60-68。

吴军其，吴飞燕，文思娇等（2023）。ChatGPT 赋能教师专业发展：机遇、挑战和路径[J]。中国电化教育， (05)：15-23+33。

赵晓伟，祝智庭，沈书生（2023）。教育提示语工程：构建数智时代的认识论新话语[J]。中国远程教育，43(11)：22-31。

Agustini, N. P. O. (2023). Examining the role of ChatGPT as a learning tool in promoting students' English language learning autonomy relevant to Kurikulum Merdeka Belajar. Edukasia: Jurnal Pendidikan Dan Pembelajaran, 4(2), 921–934.

Barrot, J. S. (2023). ChatGPT as a language learning tool: An emerging technology report. Technology, Knowledge and Learning, 1-6.

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. Library Hi Tech, (ahead-of-print).

Leppink, J., Paas, F., Van Gog, T., van Der Vleuten, C. P., & Van Merrienboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. Learning and instruction, 30, 32-42.

Kumar, J. A. (2021). Educational Chatbots for Project-Based Learning: Investigating Learning Outcomes for a Team-Based Design Course. International Journal of Educational Technology in Higher Education, 18.

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. Education and Information Technologies, 1-23.

Tan, J. S., & Chen, W. (2022). Peer feedback to support collaborative knowledge improvement: What kind of feedback feed-forward?. Computers & Education, 187, 104467.

UNESCO（2023）. Guidance for generative AI in education and research [EB/OL]. [2023-10-03]. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research.

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: a multigroup analysis. International Journal of Educational Technology in Higher Education, 20(1), 49.

Zhang, S., Gao, Q., Sun, M., Cai, Z., Li, H., Tang, Y., & Liu, Q. (2022). Understanding student teachers' collaborative problem solving: Insights from an epistemic network analysis (ENA). Computers & Education, 183, 104485.

技术赋能教师职业幸福感提升：从技术观的视角出发

Enhancing Teachers' Occupational Well-being through Technological Empowerment: A Perspective from the Technological View

管乐[1]，林易[1]，文燕银[1]，吴筱萌[1]
[1]北京大学教育学院教育技术系
2201213193@stu.pku.edu.cn

【摘要】 中国已将教师职业幸福感提升纳入教师队伍建设改革目标。如今技术正越来越多地融入提升教师职业幸福感的干预措施中。鉴于此，论证是否应该使用技术、如何使用技术显得十分必要。这些问题直接涉及哲学取向的技术观，技术观通过影响实践取向进而影响技术所发挥的效应。本研究分析国内外学者的干预实践中的技术应用，发现了其提升教学成就、便利沟通和扩大干预影响的优势效应。随后运用卡尔·米切姆的技术哲学理论，分析技术作为人工物、知识、过程和意志，如何促进其优势发挥，以及存在何种需要警惕的潜在危害，据此提出实践建议。

【关键词】 教师专业发展；职业幸福感；技术干预；技术观

Abstract: *China has incorporated the enhancement of teachers' occupational well-being into the educational reform goal. In the digital age, technology is increasingly integrated into interventions aimed at improving teachers' occupational well-being. Therefore, it becomes crucial to demonstrate whether technology should be used, and how it should be used. These questions directly involve the philosophical perspective of technological view. This study analyzed the technological applications in interventions and identified the positive effect of technology in enhancing teaching achievement, facilitating communication, and expanding impact. Subsequently, applying Carl Mitcham's theory, the study demonstrated how technology, as an artifact, knowledge, process, and will, promotes its positive effect and identified potential hazards that need caution. Based on these analyses, the study presents technological practice recommendations.*

Keywords: teacher professional development, occupational well-being, technological intervention, technological view

1. 问题缘起：教师职业幸福感的多元"失衡"

如今，教师的职业幸福感受到多方面因素的威胁，导致教师职业幸福感在情感、认知、社交、身体等维度上存在不同程度的缺失。这一问题源于"失衡"：自我调节能力与负面情绪生成速率的失衡，扩大了负面事件的影响。学习特征与专业发展模式之间的不完全适配，降低了工作成就、提高了职业倦怠。高社会期待与低社会支持之间的发展悖论，进一步激化了教师个体幸福和社会大环境之间的矛盾。本文从技术视角出发，为教师个体、教师所属组织找到更有效的职业幸福感改善路径，阻止多元"失衡"加剧，重构新平衡模式。

2. 聚焦技术：以多元形式为提升教师职业幸福感赋能

目前已有融入技术的干预措施被投入实践场域，旨在扩大受益群体、提高干预方案的便捷性与灵活性等。梳理相关文献发现，技术的赋能手段主要体现为以下五种。

第一，提供电子资源。自主跟练瑜伽视频资源（Tamilselvi, 2011）能使教师减轻负面情绪。Rahayu等（2022）为幼儿教师研发帮助教师鼓励学生的电子课本，提高教师与学生社交的能力，间接提升幸福感。电子资源实现了远程教学，扩大了教师接收教育资源的时空便利性。

第二，利用在线课程。比起电子资源，在线课程则能够将更为详实、互动性更强的信息呈现给教师做理论解释与实践展示。例如情绪管理和教学策略两方面的课程能够直接或间接提升教师幸福感。

第三，建立在线社区。教师需要与内容研发者、学习同伴产生交互，以获得更良好的情绪和更有效的实践模式。例如本就有助于减轻教师倦怠和压力感知的阅读疗法，若配合在线学习社区，让教师在线上互助小组内进行阅读分享，效果将更显著（Eddy et al., 2022）。

第四，采用线上线下混合方式。将具备强交互性的在线干预手段与线下培训活动相结合，会产生更好的干预效果。研究发现，促进师生互动的在线课程仅对教师进行线上技能培训效果并不理想，但与线下研讨会、反思性写作结合后，教师的自我效能感、职业倦怠、压力感知的改善效果更好（Roberts et al., 2020）。除了信息、教师还需要与教育过程中的各类主体进行真实交互，才能将外部环境的指导落成于现实工作环境。

第五，采集和显示数据。此时技术更多辅助研究者提高数据的采集效率和效度，而非直接作用于教师。例如，采集参与者的皮质醇浓度，以判断参与者处于愉悦或高压状态。此外，干预措施实施过程中的前后测问卷也可通过电子邮件、网络平台的方式实现点对点的精准发送与回收。在未来，易于携带的可穿戴式设备将被更多使用，采集教师真实工作状态下的过程性数据，产出更具实践意义的研究结论。

综上所述，在线技术在提升教师职业幸福感方面具有重要意义。它既能低成本拓展课程受众，又能支持学习小组和教师交流，促进教师转型为社区型的开放教室。技术的优势如交互性、即时性和多样性，有助于教师短期内提升幸福感。因此，未来还需开发简短有效的技术干预措施和培训项目。而在使用技术时应明确其角色和作用，才能避免盲目使用和为用而用，以确保技术能够充分发挥其潜力。

3. 理论切入：以米切姆的技术哲学框架审视技术的使用

米切姆综合阐发了四种技术类型：作为物体的技术、作为知识的技术、作为活动的技术和作为意志的技术。采纳这一框架是因为技术和人对技术的观点都处于发展过程中，该理论比较完整地呈现出了变化中蕴含的多样化理解。这种多样化能够与干预目的及实施条件、教师群体特征及其需求的多样化相适应。本文并非比较四种技术类型的优劣，而是讨论它们分别适用于何种类型的情境。

3.1. 作为物体的技术：技术手段的应用

作为物体的技术包括工具、机器等所有人类制造出的物质人工物，在前述的技术使用方式中体现为：第一，提供电子资源，此时技术作为交流工具，将信息传递给更广泛的教师群体；第二，以技术辅助研究者采集和显示数据，经由人力操控相应的工具，将信息以可理解、可分析的形式快速、高效地传递给研究者。第三，在线课程符合“作为过程容器的机器”（2008,p.225-227），在线课程的实施需要人的指导、也承载着整个信息传递和接受的过程，线上实施营造了独立于真实世界的空间和环境，教师于该环境中受干预影响，发生定向改变。

上述应用中，研究者或领域专家通过技术物体对教师产生影响，即技术物的主动权掌握在作为他者的人手中。这类干预措施本质上仍是一群人（研究者、专家）对另一群人（教师）的影响。因此也需要在干预的设计阶段保持警惕——技术物体作为影响的载体或辅助工具，在发挥作用时有可能出现偏离预期的情况，即预期中产生的影响被扩大、缩小或转向。

然而这一反思本质上还是“作为物体的技术的价值取决于人的使用方式”，是“技术双刃剑”

的论调。虽有合理性，但不具备在实践方面的指导作用，所以仍需深入对技术进行价值判断。米切姆（2008,p.242）赞同曾悦刘的观点：“机器可以提供速度、力量、统一的规格以及准确性，但却无法提供创造力、适应性、自由和异质性。”作为过程的机器在具体实践中最典型的表现方式就是工业生产线，高效、统一、高合格率是其巨大优势。但教育关注于人和个性发展，对这种优势需要进行更细致地辨析和更谨慎地处理。对于已被证明有效的呼吸疗法等需以标准方式执行的干预措施，运用固定的传授方式能让教师最大化地接近标准，确保疗法作用。但压力管理模式、教学策略调整等需要因时因地制宜，对灵活性与变通性有高要求的内容，需要避免工业流水线模式的教学培训所带来的僵化风险。因此，在干预设计中要有相应模块来回应教师的个性化问题，例如提供跟踪辅导、讨论交流机会或线下答疑研讨活动等。

上述分析提供了以下实践启示：首先要重视人与技术物体的交互。每当技术带来了更为实用的问题解决工具时，原有的解决方式就会被逐渐抛弃，人所具备的相关能力也可能随之退化。例如当课堂讲授越来越依赖于多媒体课件时，教师设计并美观呈现板书的能力就减少了锻炼机会，进而导致板书技能的退化。所以，既不能回避技术带给人的便利性，也要关注如何通过技术为使用者带来长足进步。此外，在技术公平角度米切姆（2008, p.254）引用伊里奇对新型工具的倡议：工具应当可以被自由选择；工具应是个人生活的积极表达；工具不能被部分职业精英垄断。技术物体应该让教师“想用就有能力获取、同时掌握使用方法”，而不是“必须用”但又不确保“能用好”。所以技术产品要具备较高的可获取性和易用性。

3.2. 作为知识的技术：理论与实际的结合

提供电子资源、组织在线课程、建立在线社区、实施混合教学都可归入作为知识的技术类型。米切姆（2008, p. 262）认为，知识是对真实信仰的验证，而关于人工物的制作和使用的真实信仰，可以通过对技能、格言、法则、规则或理论的诉求来得以验证，并产生了各种不同种类的作为知识的技术。这类干预措施可以理解为通过实验来评估研究者们所“信仰”的某种方案及其所依托的理论的正确性，因而可以理解为作为知识的技术类型。

此类技术兼具认知和实践两个维度。即技术的使用者需要认知到目前情况如何，需要什么问题，需要采取什么行动。不同于以科学为主导的对理论的应用与检验，这类技术在发挥作用的过程中体现出一定的主动性：一方面，它不会完全遵照理论框架，会将理论模型中出于种种考虑而忽视、但在实践中体现出重要性的领域重新强调出来；另一方面，这也提示研究者在实施过程中不断调整和改进，既要及时对理论做出贡献，也要把这些贡献立即运用于实践中，这与基于设计的研究相似。

马丁·海德格尔进一步指出技术是一种特殊的真理，现代技术作为座架展现了这个世界（米切姆, 2008, p. 281）。技术应用的目标在于“使自然起作用”。这其实也提醒研究者在设计在线课程等干预项目时，要真正贴合教师们的“自然”，即考虑到教师作为活生生的个体的天性、喜好、观念等因素，避免强硬地灌输新的方法和理念，或强行要求教师放弃早已习惯的行为方式和思维模式。要根据即时掌握的教师动向，做出适时温和的引导。这也是小规模、有即时跟踪反馈和系统指导、交互性强的干预课程普遍效果更显著的原因之一。

3.3. 作为活动的技术：研究设计的细化

“活动”是“知识和意志联合起来使人工物得以存在或让人使用的关键事件”（2008, p. 283）。比之作为物体的技术类型中承载了整个过程的“机器”，此类技术还涵盖了人的认知和意志与行为相统一的维度。即促成某一结果达成的意志、知道如何达成这一结果的知识，和促成这一结果的行为相统一。从这个意义上看，作为活动的技术更多指导着大规模干预项目的宏观设计与实施过程的管理环节。

米切姆（2008, p. 302）分析工程设计时指出，绘图可能是完成更为普遍的模仿工程活动的唯一方式。“模仿”不等于“复制”，按部就班地接收知识和理论、模仿相应的行为只是手段，

逐渐形成属于教师自身的情绪调节能力、让教师自己成为自身幸福感的来源才是最终的目的。因此，对于干预措施的设计需要更加精细化、系统化，必要时引入教学理论和专家经验，不仅关注实施了什么，更要重视如何实施。要将抽象地存在于专家经验中相关知识与能力，具象化为教师可理解、可模仿的行为和观念，内化为教师内在的能力。评估干预项目的效果时，不能武断地归咎于某种方法效果不理想，更应该看到实施方法是否对最大化发挥起作用造成阻碍。另一方面，设计本质上是为省力而做出系统努力、以效率为目标的行为，因此也需要警惕过度追求“最大化效益”导致的对精细化取向和个体差异的湮灭。内在于“设计”一词中对效率的追求，也在一定程度上能够解释大规模干预措施的实际效果为何不能完全实现预期。

管理的目的是形成一套正确使用工具和机器的方法并按此实施。技术评估是在尝试阐明产品制作和使用决策的形成基础，也属于管理。对于管理环节的强调启示研究者：提升教师职业幸福感的干预措施是否按照设计方案中完整实施、实施过程中的推力与阻力，以及最终效果与预期的差距，还与学校领导、教育行政部门的决策息息相关，这些部门和群体之间的和谐联动是更为宏大的问题，需要借助管理学等学科的理论进行更为深入的讨论。本文主要关注于教育领域和技术应用，不做详细展开。此外，管理必然涉及到权力关系，可能会引发对劳动的异化。因此，米切姆（2008, p.334-338）倡导“进入自然的行动”，即任何干预手段都不应成为教师新增的必须完成的工作，而是在有需求时有所依靠。干预何时启用、如何使用的选择权应归于教师，以免其成为外在于教师、反过来对教师形成支配与异化的一股力量。

3.4. 作为意志的技术：技术实践的整体反思

作为意志的技术反对技术中立的观点，认为技术独立于人的意志之外，既能影响意志，也能被人的意志改变其使用方式。这一类技术主要指征技术与人的复杂关系，更多属于技术哲学与伦理学的讨论范畴，因此本文不做详细论述，只围绕对于技术使用的启示做简要讨论。

海德格尔提倡恰当地使用技术，这里的“恰当”意味着与技术保持合适的距离：在任何需要的时候，人都可以脱离技术而完成自足的生活，这也是人对于技术的自由（2008, p. 354-355）。思考干预措施传递给了教师怎样的观念和意志是有必要的：情绪管理和教学策略等内容会引导教师把不幸福的因素更多归结于外部因素还是个人能力等内部因素？长期的干预和类似于“应急”的危机干预项目是否暗示着不幸福是需要长期与之相处的常态、还是可以被一次性解决的问题？技术手段的介入与教师工作和职业幸福感又是怎样的复杂关系？这些问题的答案体现在干预的设计与实施中，教师们对此的答案则隐藏在他们对于干预的反馈之中。因此，研究者还需要了解技术活动产生的结果，进一步进行试验和创新，以求更为简洁、高效地实现教师职业幸福感的提升。

参考文献

[1] 卡尔·米切姆(2008)。陈凡、朱春艳译。通过技术思考。沈阳：辽宁人民出版社。

[2] Eddy, C. L., Herman, K. C., Huang, F., & Reinke, W. M. (2022). Evaluation of a bibliotherapy-based stress intervention for teachers. Teaching and Teacher Education, 109, 103543.

[3] Rahayu, S., Usman, H., Sugito, S., & Herwin, H. (2022). The Digital Module Encourages Expression to Develop the Social Competence of Early Childhood Education Teachers. World Journal on Educational Technology: Current Issues, 14(3), 682-691.

[4] Roberts, A. M., LoCasale-Crouch, J., Hamre, B. K., & Jamil, F. M. (2020). Preschool teachers' self-efficacy, burnout, and stress in online professional development: A mixed methods approach to understand change. Journal of Early Childhood Teacher Education, 41(3), 262-283.

[5] Tamilselvi, B. (2011). Development and validation of yoga video package and its effectiveness on depression, anxiety and stress of school teachers.

沉浸式虚拟现实环境中存在感及认知情感因素对学习的影响机制研究*

——基于不同共情能力水平的比较

A Study on the Mechanism of Influence of Presence and Affective & Cognitive Factors on Learning in IVR Environments: A Comparative Study Based on Different Levels of Empathy

刘伊祉兰[1]，兰翠玲[1]，邢晓宁[1]，王雪[1*]，余能保[2]
[1]天津师范大学，教育学部
[2]天津工业大学，人文学院
* wangxuetjnu@qq.com

【摘要】 沉浸式特征对认知情感因素具有影响作用，但较少有研究关注不同共情能力下的影响机制。基于此开展验研究，通过不同共情能力的学习者在虚拟环境中学习，探索存在感对认知与情感因素及学习的影响。通过收集实验数据，利用最小二乘结构方程模型（PLS-SEM）探索变量间关系，研究发现：1.存在感对积极情绪具有显著影响，高共情条件下影响更显著；2.存在感对具身认知、内部动机、自我效能感、自我调节具有显著影响，高共情条件下影响更显著。3.积极情绪对满意度有显著影响，自我效能感对保持成绩有显著影响，内部动机与认知负荷对迁移成绩有显著影响。

【关键词】 沉浸式虚拟现实；CAMIL（CAMIL）；存在感；共情；结构方程模型

Abstract: *Immersive characteristics have an effect on cognitive affective factors, but few studies have focused on the mechanism of influence under empathy. Based on this, an experiment was carried out to explore the influence of presence on cognitive and affective factors and learning through learners with different level of empathy learning in an IVR environment. Through the collection of experimental data, the structural equation modeling (PLS-SEM) was used to explore the relationship between variables, and the research found that: 1. Presence has a significant effect on positive emotion, especially in the condition of high empathy. 2. Presence has a significant impact on embodiment, internal motivation, self-efficacy and self-regulation, especially in the case of high empathy. 3. Positive emotions have a significant impact on satisfaction, self-efficacy has a significant impact on retention, and internal motivation and cognitive load have a significant impact on transfer.*

Keywords: Virtual Reality(VR), The Cognitive Affective Model of Immersive Learning(CAMIL), Presence, Empathy, Structural Equation Modeling

1. 前言

沉浸式虚拟现实（Immersive Virtual Reality，IVR）作为一种新兴的学习方式逐渐在教育领域受到关注。虚拟学习环境的沉浸感和交互性有助于发展更复杂、更高层次的思维（Chen, 2016），增强学习者的存在感，促进学习者的学习（李文昊 等，2023）。存在感作为 IVR 环境中核心特征，是指学习者在 IVR 环境中感到身临其境、真实存在的程度。存在感的高低

* 本文系国家自然科学基金青年项目（项目编号：62107030）的研究成果；本文系 2022 年天津市研究生科研创新项目（项目编号：2022SKY266）的研究成果；本文系 2021 年度天津市教育科学规划课题"虚拟现实(VR)教学资源多模态分析及设计优化研究"（HCE210337）的研究成果。

不仅会影响学习者对学习任务的投入程度，与学习者的主观体验、认知情感因素也密切相关（Makransky & Lilleholt, 2018）。CAMIL（The Cognitive Affective Model of Immersive Learning）认为存在感可以通过认知情感因素影响学习效果（Makransky & Petersen, 2021），但少有研究对该模型进行验证（Makransky & Mayer, 2022），同时 IVR 环境中学习者的存在感及认知情感因素对学习的影响机制尚不明确。

共情能力是指个体对他人情感状态的感知和理解能力（Blair, 2005）。高共情力的人可以准确地从周围获取有价值的信息，更高从而促进学习，高共情力者在学业表现、自我效能感等方面都优于共情能力低的学生（McQuiggan, Robison, & Lester, 2010），VR 环境的沉浸感带给学习者较强的存在感，促使学习者展现出更高的共情能力（向维 等，2021）。然而，研究大多关注 IVR 环境能否引发共情，基于不同共情能力水平比较 IVR 环境中的学习机制是否有差异以及有何差异尚未明确，需进一步探讨。

因此，本研究通过最小二乘结构方程模型（PLS-SEM）分析存在感与认知情感因素的相互作用及其与学习效果间的关系，基于不同共情能力水平进行比较研究，进一步探讨共情能力在其中的影响机制；揭示 IVR 学习的内在机制和影响因素，验证并深化 CAMIL，推动 IVR 学习向更智能、更个性化的方向发展，为教育领域的实践提供有力的理论支撑。

2. 理论模型与研究假设

2.1. 理论基础

沉浸式学习认知情感模型（CAMIL）指出，在 IVR 中，认知与情感因素占据核心地位，而存在感和能动性是影响这两者的关键因素(Makransk & Petersen, 2021)，该模型详细阐述了存在感和能动性如何由技术特征产生，以及其如何通过认知与情感过程预测学习效果。此外，该理论还强调了认知与情感因素在获取和转移事实、概念和程序知识过程中的作用。沉浸式学习认知情感模型为本研究揭示沉浸式虚拟现实环境中变量之间的关系提供了重要的理论支撑。

2.2 变量选取

根据 CAMIL，认知与情感因素在虚拟现实环境中的学习过程中起着重要作用，其包括兴趣、内在动机、自我效能、具身认知、认知负荷和自我调节，认知与情感因素又与学习满意度以及学习成绩有关(Makransk & Petersen, 2021)。因此，本文选取存在感、具身认知、内部动机、积极情绪、认知负荷、自我效能感、自我调节以及学习满意度和学习效果 9 个指标，基于 CAMIL 框架探讨沉浸式特征如何影响学习者的认知与情感体验，具体如下：

存在感。存在感可以理解为“身临其境”的感觉(Riva, David, & IJsselsteijn, 2003)，代表着个体对于虚拟环境真实感的感知程度。本研究将深入探讨存在感如何影响学习者的认知与情感因素，以揭示沉浸式特征与学习者体验之间的关联。具身认知。具身认知强调身体、心智与环境的交互融合，指虚拟身体的体验。IVR 中，具身认知水平的高低会影响学习者的学习(Yu, Liu, Johnson-Glenberg, Han, & Wu2023)。本研究采用具身认知反映认知情感因素对于学习效果的影响。内在动机。内在动机是指为了与活动本身相关的内在满足感而从事活动(Ryan & Deci, 2000)。在基于 IVR 的课程中，先前研究确定了与 IVR 学习相关的情感路径，其中存在感与动机相关(Makransky & Lilleholt, 2018)。本研究中动机主要作为可能的认知情感因素，中介存在感对于学习效果的影响。积极情绪。在 IVR 情境下，存在感可以为激发学习者的情绪创造必要的条件(Makransky & Petersen, 2021)，学习者处于积极情绪状态在理解和迁移测试中具有更好的学习成果 (Park, Knörzer, Plass, & Brünken, 2015)。本研究中积极情绪主要作为可能的认知情感因素，中介存在感对于学习效果的影响。认知负荷。根据认知负荷理论， IVR 旨在带来模拟学习体验，但可能会增加无关的认知负荷(Zhong, Zhang, Wang, & Feng, 2022)。

同时 IVR 环境能够建立更强的存在感，从而对学习材料进行了更深入的认知处理(Lee, Wong, & Fung, 2010)。因此，本研究通过考虑学习者在 IVR 环境中的认知负荷来评估沉浸式特征对学习效果的影响。自我效能感。自我效能感是指学生对学习或表现的感知能力(Bandura, Freeman, & Lightsey, 1999)，沉浸式虚拟现实环境对自我效能感具有积极影响，并且存在感和自我效能之间存在着积极的关系(Makransk & Petersen, 2019)。因此，本研究中自我效能感主要作为可能的认知情感因素，中介存在感对于学习效果的影响。自我调节。自我调节是指个体管理自己行为的能力。根据 CAMIL，IVR 环境可以通过高水平的社会存在来潜在地促进自我调节的过程，已有研究证明基于自我调节设计的虚拟现实课程提高了学生学习成绩。本研究中自我调节主要作为可能的认知情感因素，中介存在感对于学习效果的影响。学习满意度。高度沉浸式的学习环境更有可能帮助学习者提高学习准确性，并增加他们在学习过程中的满意度和成就感(Hung, Lin, Yu, & Sun, 2023)。本研究将学习满意度纳入 CAMIL，试图探索学习满意度在沉浸式学习中所起到的作用。学习效果。本研究主要用保持成绩与迁移成绩来反映学习效果。

2.3. 研究假设

综上所述，基于 CAMIL，探索存在感对于情感（积极情绪）与认知因素（具身认知、内部动机、积极情绪、认知负荷、自我效能感、自我调节）之间的关系，以及认知与情感因素对学习效果的调节作用。相关假设如下：

2.3.1. 存在感对情绪的影响

根据 CAMIL，IVR 环境中存在感可以激发学习者的积极情绪产生(Makransky & Petersen, 2021)，同时不同共情能力的学习者在 IVR 环境中所产生的存在感有所不同，为了更好理解不同共情能力水平的学习者在 IVR 环境中存在感与情绪之间的关系，本研究提出假设 H1：存在感对情绪有积极影响。

2.3.2. 存在感对认知因素的影响

多媒体学习的认知情感模型(Moreno & Mayer, 2007)表明，IVR 可以通过提供更逼真的体验来促进生成处理，这将导致更高的存在感，根据学业情绪的控制价值理论，存在感可以培养学习者的积极认知价值(Makransk & Lilleholt, 2018)。因此本研究提出假设 H2：存在感对认知因素(具身认知、认知负荷、内部动机、自我效能感、自我调节)有积极影响。

2.3.3. 情绪与认知因素对学习效果的影响

根据 CAMIL，情感和认知因素在 IVR 学习中起着重要作用(Makransk & Petersen, 2019)，内在动机、自我效能感、具身认知感和自我调节水平的提高以及认知负荷水平的降低可以对学习效果（保持测试、迁移测试）产生积极影响。因此，本研究提出假设 H3：认知与情感因素对于学习效果有积极影响。

3. 实验设计

3.1. 研究对象及实验材料

研究对象为 88 名在校大学生，被试年龄分布在 17~28 岁，平均年龄约为 20 岁。其中，由于 2 名被试中途放弃实验，最终有效被试总数为 86 人。每位被试在实验前均签署了《实验知情同意书》，为自愿参与实验，并被告知如果正常完成实验可获得被试费作为报酬。本实验所用的学习材料参考借鉴国家虚拟仿真实验平台的教学资源《曾侯乙编钟演奏与创编实验》，基于 Unity 3D 开发曾侯乙编钟学习场景。虚拟环境对真实编钟进行了复原，使学习者能够在真实环境不允许敲钟的情况下沉浸体验虚拟环境中的编钟演奏。在原有教学环节基础上，每个环节都增设 3 道测试题交互环节。各被试佩戴 HTC Vive Pro Eye 学习。由于具备交互环节，学习时长并不完全一致，约 18-21 分钟。流程为：首先，被试阅读《实验知情同意书》并签

字，然后完成人口统计学调查和先验知识测试。然后，介绍整体实验流程、IVR 设备使用方法及注意事项。之后，被试学习实验材料。学习结束后，被试完成体验问卷和知识后测。

3.3. 研究工具

测量的变量主要包括存在感、认知情感因素（积极情绪、具身认知、认知负荷、内部动机、自我效能感、自我调节、满意度）及学习效果（保持测试和迁移测试）。此外，还收集了被试的年龄、性别、年级、专业、编钟及基本乐理知识的经验等人口统计学特征信息。

3.3.1 存在感

存在感采用 Kim（1997）编制的量表，Cronbach's α=0.777。

3.3.2. 认知与情感因素

学业情绪采用经典的 PANAS 问卷测量，Cronbach's α=0.925。动机采用 Isen 和 Reeve（2005）编制的自我报告问卷并进行了修改，以适应本研究的学习活动，Cronbach's α=0.942。题项如“它使我想进一步探索”。自我效能感采用 Artino 等人(2008)编制的自我报告问卷并进行了修改，以适应本研究的学习活动，Cronbach's α=0.831。题项如“在刚刚的学习中，我有信心在没有他人的帮助下也能学好编钟。”认知负荷采用 Pass（1992）编制的自我报告问卷（Cronbach's α=0.74）。具身认知采用 Shin 和 Biocca（2018）改编的自我报告问卷（Cronbach's α>0.7）。自我调节采用 Broadbent 等人（2023）使用的量表（Cronbach's α>0.7）。满意度采用 Shin 和 Biocca（2018）改编的自我报告问卷（Cronbach's α>0.7）。

3.3.3. 学习效果

学习效果是参考国家虚拟仿真平台（*https://www.ilab-x.com/*）相关资源进行设计并扩展。保持测试包含 10 项选择题（每题 1 分），例如“甬钟的上部被称为（ ）”；迁移测试包含 10 项选择题（每题 1 分）（例如，“晓玉在一个编钟上看到了一些认不太出来的文字，晓玉认为首先这个钟可能是（ ）钟。”）和一项简答题（3 分）。

3.4. 分析方法

本研究通过 SmartPLS 3 采用偏最小二乘结构方程模型（PLS-SEM）来检验测量模型中各变量的关系(Hair, Hult, Ringle, Sarstedt, Danks, & Ray, 2021)。与基于协方差的结构方程模型（CB-SEM）相比，PLS 分析更适用于理论的探索和发展，它可以接受较小的样本量，允许偏态数据，还支持模型中存在形成性指标。在本研究中，基于 CAMIL 对模型进行扩展，因素关系较为复杂，新增社会存在感、学习满意度、积极情绪等变量对模型开展理论性探索，且样本量为 82，属于较小样本量。因此，本研究更适合采用 PLS-SEM。

4. 数据分析

4.1. 结构模型评估

本研究采用包含 5000 个子样本和 95%置信区间的偏差校正自举程序，进行全样本（N=82）结构模型中假设关系的评估，结果如图 1 所示。内部动机（β=0.357，p<0.01）对迁移成绩有积极影响，存在感（β=0.417，p<0.01）对具身认知体验存在积极影响，存在感（β=0.480，p<0.01）对内部动机具有具有显著影响，存在感对积极情绪具有显著影响（β=0.477，p<0.01），存在感对自我效能感有显著影响（β=0.422，p<0.01），存在感（β=0.376，p<0.01）对自我调节有显著积极影响，社会存在感（β=-0.251，p<0.01）对存在感有显著的负向影响，积极情绪（β=0.502，p<0.01）对满意度有显著的正向影响，自我效能感（β=0.014，p<0.01）对保持成绩有显著影响，认知负荷（β=0.187，p<0.01）对迁移成绩有显著的正向影响作用。

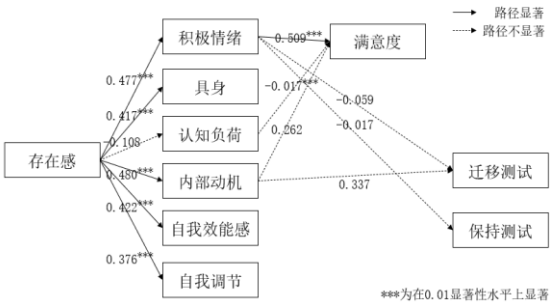


图 1 全样本模型检验结果

4.2. 不同共情能力下模型路径对比分析

为探究不同共情能力的学习者在 IVR 中的认知情感机制，以共情能力统计量中位数为分界，将高于总共情能力中位数的学习者列为高共情，其余学习者为低共情，检测共情能力的调节效应。首先，采用偏最小二乘多组分析（PLS-MGA）检验两种共情能力水平的学习者在路径系数上是否存在显著差异(Hair, Hult, Ringle, Sarstedt, Danks, & Ray, 2021)。表 2 为基于 t 检验的多组比较分析结果。结果显示，在模型的路径钟，多条路径均出现了不同共情能力下的显著差异。高共情能力水平下与低共情能力水平下的学习者在 IVR 学习环境中，内部动机（|D|=0.718，p<0.01）对学习满意度的影响、内部动机（|D|=0.846，p<0.01）对迁移成绩的影响、积极情绪（|D|=-0.578，p<0.01）对满意度的影响，积极情绪（|D|=-1.076，p<0.01）对迁移成绩的影响均存在显著差异。

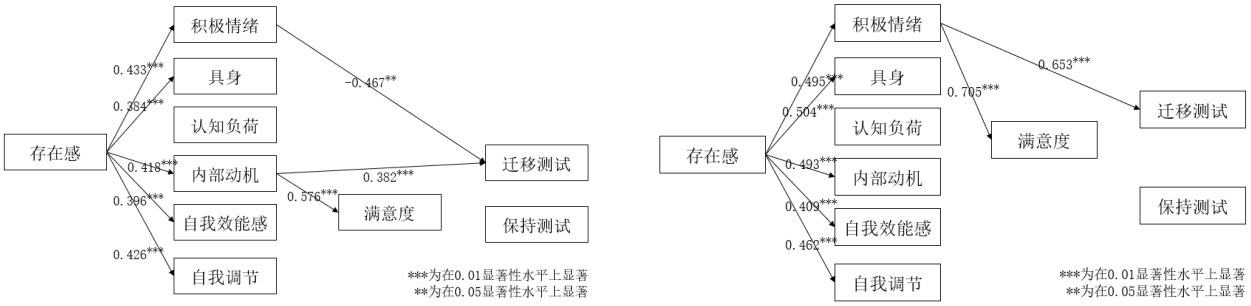
表 2 共情能力不同水平 t 检验及其比较

路径	路径系数差|D|	*p* 值	低共情		高共情	
			路径系数	*P 值*	路径系数	*p* 值
内部动机 -> 满意度	0.655	0.004	0.576	0.001	0.062	0.749
内部动机 -> 迁移成绩	0.888	0.019	0.750	0.000	-0.367	0.225
存在感 -> 具身认知	-0.283	0.157			/	
存在感 -> 内部动机	-0.121	0.453			/	
存在感 -> 积极情绪	-0.177	0.316			/	
存在感 -> 自我效能感	-0.219	0.274			/	
存在感 -> 自我调节	-0.192	0.332			/	
存在感 -> 认知负荷	0.374	0.157			/	
社会存在感 -> 存在感	0.101	0.607			/	
积极情绪 -> 保持成绩	-0.068	0.742			/	
积极情绪 -> 满意度	-0.596	0.011	0.206	0.199	0.705	0.000
积极情绪 -> 迁移成绩	-1.041	0.005	-0.467	0.025	0.653	0.004
认知负荷 -> 满意度	0.134	0.327			/	

其次，为了探究不同共情能力水平的学习者在 IVR 中的认知情感机制模型具体存在哪些不同，将高共情能力水平（N=48）与低共情能力水平（N=39）的学习者认知情感结构模型进行对比。高共情水平与低共情水平学习者的结构模型分别如图 2 所示，两种共情水平对比的假设检验结果见表 2。

高共情能力与低共情能力两种水平的学习者 CAMIL 有所不同。低共情能力水平学习者其内部动机对迁移测试成绩影响显著（β=0.382，p<0.01）（图 2-a），而在高共情能力学习者处，这一路径并不显著（β=-0.367，p>0.05）。低共情能力水平学习者在沉浸式学习中的内

部动机能够对满意度产生显著影响（β=0.750，p<0.01），而在高共情能力水平学习者处并无显著效果（β=-0.367，p<0.01）。高共情能力水平学习者其积极情绪对迁移测试成绩影响显著（β=0.653，p<0.01）（图 2-b），而在低共情能力学习者处，这一路径并不显著（β=-0.467，p>0.05）。高共情能力水平学习者在沉浸式学习中的积极情绪能够对满意度产生显著影响（β=0.705，p<0.01），而在高共情能力水平学习者处并无显著效果（β=0.206，p<0.01）。



a 低共情能力水平　　b 高共情能力水平

图 2 不同共情能力水平下学习者的 CAMIL

5. 结果讨论

本研究旨在探究不同存在感的沉浸式虚拟现实环境，如何影响学习者的情绪与认知因素（具身认知、认知负荷、内部动机、自我效能感、自我调节），同时探究不同共情能力水平的学习者在 IVR 中的认知情感机制模型具体存在哪些不同。通过结构方程的检验，本研究发现，存在感对情绪有显著的影响，对具身认知、内部动机、自我效能感、自我调节有显著的影响，同时发现，不同的共情水平下的认知情感模型路径不同。

5.1. IVR 特征对情绪的影响

总体上，存在感对积极情绪具有显著影响，假设 1 成立。进一步对比分析不同共情水平下，存在感对情绪的影响发现。在高共情条件下，存在感对情绪的影响相比于低共情条件下更显著。也就是说，共情能力高的学习者可能更能深入感知虚拟同伴的存在，并受到更加深入的影响。这与 Teding van Berkhou & Malouff 的研究(2016)一致，即具备高共情能力的学习者更能够在学习过程中受到学习同伴的认知与情绪等状态影响。

5.2. IVR 特征对认知因素的影响

总体上，存在感对具身认知、内部动机、自我效能感、自我调节具有显著的影响，但对认知负荷的影响并不显著，假设 2 部分成立。进一步对比分析不同共情水平下，存在感对认知因素的影响发现。在高共情的条件下，存在感对具身认知、内部动机、自我效能感、自我调节的影响更显著。这符合高共情学习者的特点，即共情一般是指对他人的感受与行为的的能力(Brydon, Kimber, Sponagle, MacLaine, Avery, Pyke, & Gilbert, 2021)，同时共情能力会影响学习，在学习的过程中会改善学习者之间的互动，提高学习者的动机。

5.3. 情绪与认知因素对学习效果的影响

总体上，积极情绪对满意度有显著的正向影响，自我效能感对保持成绩有显著影响，内部动机对迁移成绩有积极影响，认知负荷对迁移成绩有显著的正向影响作用，部分支持假设 3。进一步对比分析不同共情水平下，存在感对情绪的影响发现，低共情学习者的内部动机对满意度和迁移成绩影响显著，但积极情绪对迁移成绩具有负向影响。这可能是因为，低共情的学习者可能更难在学习过程中受到学习同伴的认知与情绪的影响，即若想提高学习满意度与学习效果，可通过激发内在动机来驱动学习(Teding van Berkhout & Malouff, 2016)。高共情学习者的积极情绪对满意度与迁移成绩影响显著，这可能是因为高共情学习者在学习的过程中更好地与虚拟同伴进行互动，积极情绪越高越能提升动机与满意度，最终提升学业成绩。

6. 结语

本研究通过实验采集学习者在不同沉浸式特征的环境下学习者的相关数据，利用最小二乘结构方程模型（PLS-SEM）探索了存在感及认知情感因素对学习的影响机制，同时比较了不同共情能力水平下的认知情感模型的异同。研究发现 CAMIL 的大部分路径均被验证，且不同共情能力水平的学习者在 IVR 环境中的学习机制模型具有一定的差异。存在感对积极情绪、具身认知、内部动机、自我效能感、自我调节有显著影响。说明 IVR 中，可通过增加学习者的存在感促进学习者的学习体验与学习效果。本研究还发现，低共情能力水平的学习者主要通过内部动机来影响满意度与迁移成绩；而高共情者通过积极情绪。此外，本研究还存在不足，如被试样本分布上性别存在不均衡情况等。此外，未来的研究可以从调节作用的角度来进一步检测不同共情水平在存在感及认知情感因素对学习的影响的调节水平。

根据本研究结论提出两点建议：第一，因教育才。根据结构模型总体评估，沉浸式虚拟现实学习环境并不适用于所有类型的学习内容，需要充分调动学习者内在动机、积极情绪等情感因素的学习材料更适合于 IVR 中呈现。第二，因材施教。一方面，对于低共情能力水平的学习者而言，相比于同伴的陪伴，在 IVR 环境中更应该设置能够引起学习者兴趣、增强学习者内部动机的元素如闯关任务、游戏等以促进学习；对于高共情能力水平的学习者，除了 IVR 本身的教学设计之外，需要在虚拟环境中充分提供认知与情感的社会化线索如人物的交流对白、价值诱导等等，以促进该类学习者在沉浸式虚拟现实中的学习。

参考文献

李文昊,钱莉,冯秦娜 & 黄静.(2023).增强沉浸感能提高学习成效吗？——沉浸感对学习结果的影响及作用机制. 电化教育研究(12),55-63.

向维,李文昊,李亚芳 陈冬敏.(2021).虚拟现实环境中反馈对学习效果的影响机制——基于共情的中介效应分析.现代教育技术(06),88-95.

Artino Jr, A. R., & McCoach, D. B. (2008). Development and initial validation of the online learning value and self-efficacy scale. Journal of Educational Computing Research, 38(3), 279-303.

Bandura, A., Freeman, W. H., & Lightsey, R. (1999). Self-efficacy: The exercise of control.

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. Consciousness and Cognition, 14(4), 698–718.

Broadbent, J., Panadero, E., Lodge, J. M., & Fuller-Tyszkiewicz, M. (2023). The self-regulation for learning online (SRL-O) questionnaire. Metacognition and Learning, 18(1), 135-163.

Brydon, M., Kimber, J., Sponagle, M., MacLaine, J., Avery, J., Pyke, L., & Gilbert, R. (2021). Virtual reality as a tool for eliciting empathetic behaviour in carers: an integrative review. Journal of Medical Imaging and Radiation Sciences, 52(3), 466-477.

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. Computers & Education, 102, 152-171.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial least squares structural equation modeling (PLS-SEM) using R: A workbook (p. 197). Springer Nature.

Hung, C. Y., Lin, Y. T., Yu, S. J., & Sun, J. C. Y. (2023). Effects of AR–and VR–based wearables in teaching English: The application of an ARCS model–based learning design to improve elementary school students' learning motivation and performance. Journal of Computer Assisted Learning.

Isen, A. M., & Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behavior, and self-control. Motivation and emotion, 29, 295-323.

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. Journal of computer-mediated communication, 3(2), JCMC325.

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. Computers & Education, 55(4), 1424-1442.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educational Technology Research and Development, 66(5), 1141-1164.

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. Computers & Education, 134, 15-30.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. Educational Psychology Review, 1-22.

Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. Educational Psychology Review, 34(3), 1771-1798.

McQuiggan, S. W., Robison, J. L., & Lester, J. C. (2010). Affective transitions in narrative-centered learning environments. Journal of Educational Technology & Society, 13(1), 40-53.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. Educational psychology review, 19, 309-326.

Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. Computers & Education, 86, 30-42.

Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). Being there: The experience of presence in mediated environments. Being there: Concepts, effects and measurement of user presence in synthetic environments, 5.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary educational psychology, 25(1), 54-67.

Shin, D., & Biocca, F. (2018). Exploring immersive experience in journalism. New media & society, 20(8), 2800-2823.

Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. Journal of counseling psychology, 63(1), 32.

Yu, S., Liu, Q., Johnson-Glenberg, M. C., Han, M., Ma, J., Ba, S., & Wu, L. (2023). Promoting musical instrument learning in virtual reality environment: Effects of embodiment and visual cues. Computers & Education, 198, 104764.

Zhong, Z., Zhang, G., Jin, S., Wang, J., Ma, N., & Feng, S. (2022). Investigating the effect of peer instruction on learners with different cognitive styles in VR-based learning environment. Education and Information Technologies, 27(8), 11875-11899.

多媒体学习中的积极情绪设计对学习效果的影响——基于生理数据的元分析

The Influence of Positive Emotional Design on Learning Effects in Multimedia Learning: A Meta-Analysis Based on Physiological Data

朱振凯[1]，方苏[1]，靳燕燕[1]，王雅莉[1]，彭晛[1*]

[1]华中师范大学 人工智能教育学部，湖北 武汉 430079

*pengxian@ccnu.edu.cn

【摘要】 多媒体教学中融入积极情绪设计已成趋势，尽管利用生理设备的实验对其学习效果的评估存在不一致。本研究通过元分析方法研究了积极情绪设计对学习者情绪、生理数据（即总注视时间）、认知负荷、内在动机以及学习成绩的影响。结果显示积极情绪设计显著延长了学习者的总注视时间（$g = 0.44$）和提高了内在动机（$g = 0.31$），同时有效提升了学习保持成绩（$g = 0.26$）和迁移成绩（$g = 0.34$）。此外，学习者的语言文化和学习内容主题对其效果有调节作用。研究表明积极情绪设计在多媒体学习中的积极作用，因此，在设计教学材料时，教学者应主动加入积极情绪设计以提升教学质量。

【关键词】 积极情绪设计；多媒体学习；生理数据；学习效果

Abstract: *The integration of positive emotional design into multimedia teaching has become a trend, despite inconsistent findings in assessing its learning effects using physiological devices. This study employed a meta-analysis approach to investigate the impact of positive emotional design on learners' emotions, physiological data (i.e., total gaze time), cognitive load, intrinsic motivation, and learning performance. The results indicated that positive emotional design significantly extended learners' total gaze time (g = 0.44) and enhanced intrinsic motivation (g = 0.31), while effectively improving both retention scores (g = 0.26) and transfer scores (g = 0.34). Additionally, learners' language and cultural background, as well as the thematic content of the learning materials, had moderating effects on its effectiveness. The study highlights the positive role of emotional design in multimedia learning, suggesting that educators should proactively incorporate positive emotional design into instructional materials to enhance teaching quality.*

Keywords: *Positive Emotional Design, Multimedia Learning, Physiological Data, Learning Outcomes*

1. 引言

当前，多媒体技术已广泛应用于教育领域，并为不同教学模式提供了诸多资源，为提升教学效果创造了机遇，教学者也逐渐意识到多媒体在教学过程中的重要性，所以在教学过程中积极采用多媒体进行教学，而多媒体学习也因此成为教育创新的重要一环。

多媒体学习是指在教学中同时呈现言语和画面学习材料的一种学习形式(Mayer, 2006)。而情绪设计在其中扮演着重要的角色，对此，各国学者展开了研究。Um 等人的研究表明，情绪设计能够诱发学习者的积极情绪，并导致更高的理解成绩和迁移成绩(Um, Plass, Hayward, & Homer, 2012)。之后，有学者在进一步研究后发现“拟人化+暖色”结合能促进多媒体学习的保持成绩与迁移成绩(Mayer & Estrella, 2014)。Uzun 等人研究则表明多媒体学习中不同层次和复杂程度的情感设计对学生的积极情绪有显著影响，即随着情感设计元素数量的增加，积极情绪也会增加(Uzun & Yildirim, 2018)。龚少英等人也做了实验进行验证，结果

则表明暖色因素可以激起学习者的积极情绪，拟人化只有与暖色结合才能诱发学习者的积极情绪，并且暖色和拟人化都可以促进迁移成绩，但都没有促进保持成绩(龚少英，上官晨雨，翟奎虎，& 郭雅薇，2017)。王雪等人使用“婴儿脸+圆形+暖色”的情绪设计方法进行实验，结果表明积极情绪设计不仅能有效激发学习者的积极情绪，还能增加学习者的相关认知负荷，使其投入更多认知资源(王雪，高泽红，徐文文，& 张蕾，2021)。由此可以看出，在有关“积极情绪设计是否促进多媒体学习的学习效果”的研究中，其结果与解释仍存在分歧。

在多媒体学习过程中，学习者的实时生理数据对于研究有较大影响。可以通过采集人体的各种生理信号，如眼动、脑电、皮肤电、肌电、心电等，来推测个体当前所处学习状态。有研究通过眼动测量，发现积极情绪设计会增加学习者对情绪设计元素的注视时间(X. Wang, Mayer, Han, & Zhang, 2023)。有学者根据眼–心假说(Just & Carpenter, 1976)，认为在多媒体学习环境下，学习者对文本部分的总注视时间越长、注视次数越多，学习效果会越好(X. Peng et al., 2021)。也有学者将采集心电数据与对照组的迁移成绩进行相关性分析，结果显示，对照组的迁移成绩和频谱功率这两个变量呈弱相关(X. Peng et al., 2022)。因此可以看出，生理数据与学习者的情绪状态和学习成效之间存在一定的联系。

由此可见，目前研究集中于生理数据应用和分析趋势，并在探讨积极情绪设计对多媒体学习效果的影响上存在分歧。本研究整理了相关文献，通过进一步整理、分析、统计、总结，从而得到积极情绪设计对于学习效果的影响。同时我们发现文献中大多为包含眼动追踪的数据，所以我们侧重分析积极情绪设计对总注视时间的影响。基此，研究将针对以下问题展开：

问题一：多媒体学习中的积极情绪设计能否唤醒学习者的积极情绪？

问题二：多媒体学习中的积极情绪设计是否对生理数据（如总注视时间）有显著影响？

问题三：多媒体学习中的积极情绪设计能否影响学习者的主观感知变化（如认知负荷、内在动机等）？

问题四：多媒体学习中的积极情绪设计是否可以提高学习者的保持、迁移成绩？

问题五：多媒体学习中的积极情绪设计效果是否存在潜在调节变量？

2. 研究过程

2.1. 文献来源与检索方式

本研究所引用的文献来源于 ERIC、PubMed、Web of science、ScienceDirect、PsycINFO，根据关键词或主题词及其组合词从所选数据库中进行检索。主题词为(“Multimedia learning” or multimedia or learn) AND (“emotional design”)。检索时间范围为 2010-2023 年，最终得到 17 篇文献（见图 1），共 25 个独立样本。

2.2. 文献纳入与排除

文献纳入标准如下：(1）文献须为实证研究。(2）文献中需要出现积极情绪设计，即需要包含颜色、形状、拟人化、图片、情感教学代理等多个因素中的至少一种。(3）文献中须包含生理数据（如脑电、心率、眼动、皮肤电等数据）作为因素指标。(4）实验中的情绪设计须在多媒体学习情境中有所体现。

文献排除标准如下：(1）非实验研究予以排除。(2）无情绪设计。(3）数据中无任何生理数据作为因素指标。(4）实验材料中为纯文字（无图片、画面等）的情况须予以排除。

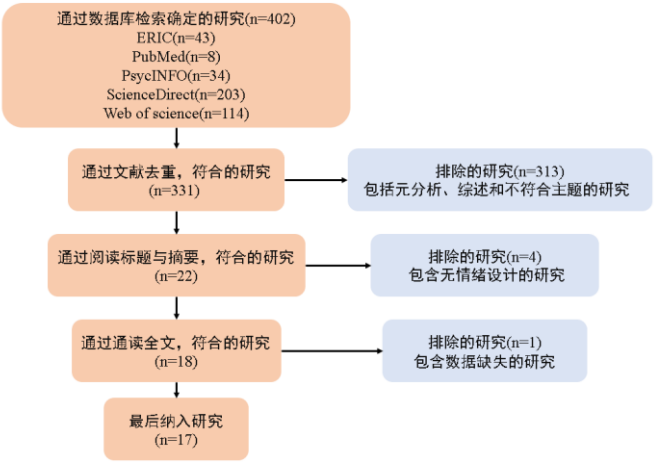


图 1：文献检索流程图

2.3. 研究方法

本研究采用元分析方法，使用 Stata MP(17)和 CMA 软件，并选取 Hedges's g 作为效应量，以校正 Cohen's d 的估计偏差。文献包含多个实验时，根据是否为研究调节变量分拆或合并效应量。另外，异质性通过 Q 检验法判断，若 Q 检验 p 值小于 0.05，表示存在异质性。同时用 I^2 检验结果作为辅助检验，采用 Higgins 等人对异质性高低的划定标准(Higgins, Thompson, Deeks, & Altman, 2003)， 25%、50%、75%分别表示异质性为低、中、高程度。如若两者数据（Q 检验与 I^2 检验）矛盾，以 Q 检验数据为准进行研究。若存在异质性，需探讨潜在变量和边界条件，并对研究进行发表偏倚检验。

2.4. 数据编码

编码的标准以 Brom 等人(Brom, Starkova, & D'Mello, 2018)对情绪设计元分析的编码为基础，进行改进整合，最终得到的编码内容则包含以下几点：情绪设计元素数量、主题、学习者学段、语言或文化背景、材料学习时长、呈现通道。

情绪设计元素数量：我们将这个变量编码为情绪设计元素数量，一共分为两种，分别为单一(Park, Knoerzer, Plass, & Bruenken, 2015)和混合(X. Wang et al., 2023)。其中，“单一”指的是独立样本中实验组与对照组间仅有一个情绪线索；而“混合”则代表存在多个情绪线索。

主题：经过筛选，我们将主题分为三大类：生物（主要包括免疫系统、ATP 合成酶等内容）、自然科学（主要为闪电的形成、天文学等内容）和其他。

学习者学段：学段主要分为两类：大学学段与其他，其中大学的学习者占大多数（85.71%），其他学段包括混合学科、初中学段（7 年级）、小学学段（3、4 年级）。

语言或文化背景：实验将学习者根据地区和语言分为四类，即中文（中国学生）、德语（德语母语者）、土耳其语（土耳其语母语者）和其他（即包括 2 个以“捷克语”为母语和 1 个以“西班牙语”为母语的 3 个样本）。

材料学习时长：我们将其编码为短（材料学习的时间小于或等于 5min）、长（材料学习的时间大于 5min）、其他（未标明材料学习时长或者实验时长超过一天）。

呈现通道：与 Brom 等人(Brom et al., 2018)一致，我们将文本、图片和幻灯片编码为静态，而将所有其他情况(如视频、动画)编码为动态。

3. 研究结果

3.1. 主效应检测

本研究采用随机效应模型进行主效应检验，研究分别对积极情绪、生理数据（即总注视时间）、主观感知（即认知负荷、内在动机）、保持成绩以及迁移成绩的主效应进行随机效应

模型分析，结果如表 1 所示。其中，对于总注视时间（$g=0.44$，$p<0.001$，见图 2）、内在动机（$g = 0.31$，$p = 0.003$）、迁移成绩（$g = 0.34$，$p < 0.001$）以及保持成绩（$g = 0.26$，$p = 0.035$），积极情绪设计在多媒体教学中展现出中等平均效应，能延长学习者对材料的注视时长，提升内在动机，进而加深记忆与理解；然而，其对积极情绪和认知负荷的影响并不显著。

表 1：主效应检验结果

结果变量	k	Hedges's g	95%CI	*p*
积极情绪	20	0.13	[0.00, 0.27]	0.053
总注视时间	13	0.44	[0.20, 0.68]	< 0.001
内部认知负荷	15	0.00	[-0.14, 0.15]	0.753
外部认知负荷	10	-0.02	[-0.30, 0.27]	0.908
内在动机	10	0.31	[0.10, 0.51]	0.003
保持成绩	14	0.26	[0.02, 0.50]	0.035
迁移成绩	18	0.34	[0.17, 0.51]	< 0.001

（注：k 代表独立效应量个数，此处 95%CI 指各个结果变量对应的效应量 *g* 的 95%置信区间）

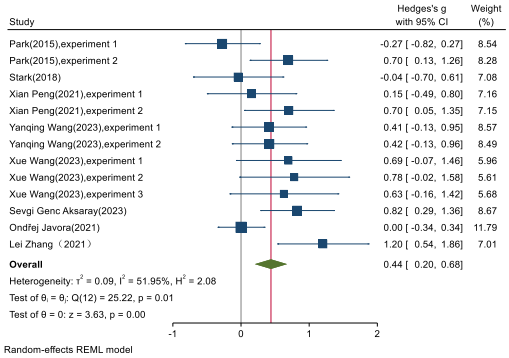


图 2：生理数据（即总注视时间）的森林图

3.2. 异质性检测

本研究分别对积极情绪、生理数据（即总注视时间）、主观感知（即认知负荷、内在动机）、保持成绩以及迁移成绩的异质性进行分析（结果如表 2 所示），发现除内在动机以外的其他 6 个结果变量 Q 检验均显著（$p<0.05$），揭示了这些变量间的异质性。这表明在多媒体教学背景下，积极情绪设计可能受到潜在调节变量的影响。为进一步明确积极情绪设计的效果边界，有必要对这些结果变量进行调节效应检验。

表 2：异质性检验结果

结果变量	Q	df	*p*	I^2
积极情绪	31.06	19	0.040	14.95%
总注视时间	25.22	12	0.014	51.95%
内部认知负荷	31.56	14	0.005	55.64%
外部认知负荷	25.50	9	0.003	64.50%
内在动机	13.56	9	0.139	33.17%
保持成绩	29.48	13	0.006	56.65%
迁移成绩	33.76	17	0.009	48.69%

3.3. 发表偏倚检测

本研究采用漏斗图、失安全系数及 Egger 线性回归检验来评估发表偏差（见表 3）。Sterne 等人指出使用漏斗图时研究样本量需要大于 10（Sterne et al., 2011），故仅对积极情绪、生理数据（即总注视时间）、内部认知负荷、保持成绩以及迁移成绩进行漏斗图的发表偏倚检验。根据各项研究在漏斗图中的大致分布是否对称，来比较直观地看出是否存在发表偏倚，结果发现保持成绩（见图 3）在漏斗图两侧的分布不均匀对称，可能存在发表偏倚；而其他变量的漏斗图分布较为均匀对称，存在发表偏倚可能性较小。从失安全系数（Rosenthal's N_{fs}）来看，在总注视时间和迁移成绩上，该指标值均大于“5k＋10”（k 指纳入元分析的独立效应

量个数），表明存在发表偏差的可能性较小；但在其他结果变量上，该指标值均小于“5k+10”，表明可能存在发表偏差。另外，我们还通过 Egger 线性回归检验发表偏倚，当 $p < 0.05$，表示存在发表偏倚，$p > 0.05$，表示无明显发表偏倚。结果发现总注视时间、内在动机、保持成绩、迁移成绩的 p 值均大于 0.05，故不存在明显的发表偏倚，而积极情绪、内部认知负荷、外部认知负荷的 p 值均小于 0.05，可能存在发表偏倚。最终，针对三种方法的结果差异，我们采用对合成效应量两侧的文献进行剪补，结果发现，剪补过后总注视时间、内在动机、保持成绩、迁移成绩的效应量仍显著，积极情绪、内部认知负荷和外部认知负荷的效应量仍不显著。综上所述，我们可以认为本研究不存在较大的发表偏差。

表 3：发表偏倚检验

结果变量	Rosenthal's N_{fs}	Egger's intercept	p
积极情绪	101	2.162	0.007
总注视时间	78	3.029	0.058
内部认知负荷	63	1.703	0.366
外部认知负荷	39	-14.354	0.106
内在动机	25	1.789	0.401
保持成绩	24	-1.907	0.496
迁移成绩	125	-0.228	0.822

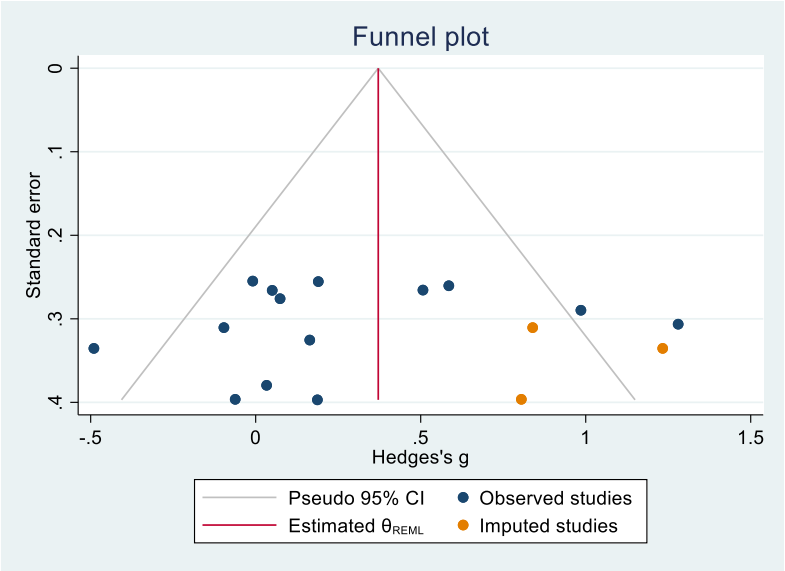


图 3：保持成绩的漏斗图（Imputed studies 是插补部分）

3.4. 调节效应检测

本研究根据文献之间的差异性，针对存在异质性的 6 个结果变量，选取了数据编码的 6 个编码结果作为调节变量进行调节效应检验。结果发现，在积极情绪（见表 4，Q(1) = 6.61，$p < 0.05$）、外部认知负荷（见表 4，Q(1) = 7.15，$p < 0.05$）、内部认知负荷（见表 4，Q(1) = 11.25，$p < 0.05$）上，“主题”均有显著调节作用，针对主题为“自然科学”的亚组，其在积极情绪的激发方面表现显著优于主题为“生物”的亚组。此外，主题为“自然科学”的亚组在外部认知负荷方面产生显著影响，远超过主题为“生物”的亚组。同样，主题为“自然科学”的亚组对内部认知负荷的影响也显著大于主题为“生物”的亚组。而在生理数据（即总注视时间）上，语言或文化背景有显著调节作用（见表 4，Q(3) = 11.11，$p < 0.05$）。在迁移成绩上，语言或文化背景有显著调节作用（见表 4，Q(3) = 10.55，$p < 0.05$），其中语言或文化背景为“中文”的亚组对于迁移成绩的促进效果显著大于其他三组，另外 5 个调节变量作用不显著（$p > 0.05$）；而在保持成绩上，6 个调节变量作用均不显著（$p > 0.05$）。

表 4：调节效应检验结果

结果变量	调节变量	k	效应量		亚组间同质性检验		
			Hedges's g	95%CI	Q(B)	df	*p*
积极情绪	主题						
	生物	11	0.02	[-0.116，0.156]			
	自然科学	9	0.436	[0.15，0.722]			
	Between levels (QB)				6.61	1	0.01
总注视时间	语言或文化背景						
	中文	8	0.588	[0.356,0.82]			
	德语	3	0.129	[-0.461,0.719]			
	土耳其语	1	0	[-0.337,0.337]			
	其他	1	0.824	[0.29,1.358]			
	Between levels (QB)				11.11	3	0.011
外部认知负荷	主题						
	生物	7	-0.206	[-0.477，0.066]			
	自然科学	3	0.426	[0.051，0.8]			
	Between levels (QB)				7.15	1	0.007
内部认知负荷	主题						
	生物	8	-0.214	[-0.434，0.006]			
	自然科学	7	0.372	[0.109，0.634]			
	Between levels (QB)				11.25	1	0.001
迁移成绩	语言或文化背景						
	中文	13	0.419	[0.236，0.602]			
	德语	2	0.079	[-0.328，0.486]			
	土耳其语	1	0.79	[0.258，1.322]			
	其他	2	-0.034	[-0.342，0.273]			
	Between levels (QB)				10.55	3	0.014

4. 讨论

本研究通过元分析探究了积极情绪设计在多媒体学习中对学生情绪、生理反应（总注视时间）、主观感受（如认知负荷和内在动机）以及学习成绩和迁移成绩的影响，并考察其中的潜在调节变量。具体分析如下：

4.1. 多媒体学习中的积极情绪设计对学习者积极情绪的影响

本研究结果表明，积极情绪设计对情绪唤醒的影响不显著，即积极情绪设计不足以唤醒学习者的积极情绪，与先前部分研究结果一致(Heidig, Muller, & Reichelt, 2015; Li, Luo, Zhang, & Shadiev, 2020)。本研究认为主要原因可能是主观量表测量的局限性和部分实验时间过长(Uzun & Yildirim, 2018)导致的情绪不稳定。建议后续研究应用先进仪器，在控制的环境中精确测量情绪状态。

4.2. 多媒体学习中的积极情绪设计对生理数据（即总注视时间）的影响

本研究发现积极情绪设计能够显著增加学习者的生理数据（即总注视时间），这与先前的研究结果保持一致(X. Peng et al., 2021)。同时本研究也支持眼-心假说(Just & Carpenter, 1976)，即学习者对于学习材料的注视持续时间与他的学习效果有关。本研究与先前研究一致(X. Wang et al., 2023)，认为积极情绪设计更能够引导学习者关注学习内容。

4.3. 多媒体学习中的积极情绪设计对学习者主观感知变化（如认知负荷、内在动机等）的影响

结果显示，积极情绪设计组相比于对照组，学习者的内在动机更高（g=0.31），这个结果与以往元分析(Brom et al., 2018)和部分研究结果(X. Peng et al., 2022)一致。然而，在内部认知负荷和外部认知负荷的评估上，积极情绪设计组与对照组之间没有显著差异，原因可能是以往研究多采用自我报告问卷来测试认知负荷，受到多方面因素影响，数据可能不够准确，

未来可以采用更加先进的生理信号测量仪器去检测学习者的实时认知负荷。另外根据认知负荷理论(Sweller, van Merriënboer, & Paas, 2019)，认知负荷可以分为内在认知负荷、外在认知负荷和相关认知负荷，三种认知负荷是相互影响的关系，共同构成学习者总的认知负荷，所以分析不同认知负荷可能会造成结果不显著。

4.4. 多媒体学习中的积极情绪设计对学习者学习成绩的影响

我们通过研究发现在多媒体学习中积极情绪设计组的材料相对于对照组，能显著提高学习者的保持成绩($p = 0.26$)、迁移成绩($p = 0.34$)，这个结果也与以往部分研究结果保持一致(Y. Wang, Gong, Cao, & Fan, 2023)。即积极情绪设计能够帮助学习者对于材料有更好地记忆与理解，提高对细节内容的记忆程度和学习者的迁移能力。

4.5. 多媒体学习中的积极情绪设计效果的潜在调节变量

从调节效应上看，积极情绪设计对于积极情绪、生理数据（即总注视时间)、内部认知负荷、外部认知负荷、迁移成绩、的影响都有一定的调节作用，而对于保持成绩并未发现显著的调节作用。其中语言/文化背景对于生理数据（即总注视时间)、迁移成绩有显著调节作用，另外 5 个调节变量作用不显著；学习材料的“主题”对于积极情绪、内部认知负荷、外部认知负荷的影响具有显著调节作用，其他调节变量无明显调节作用。

5. 结论与展望

本研究探讨了多媒体学习中积极情绪设计对于学生积极情绪、生理数据（即总注视时间)、主观感知（即认知负荷、内在动机)、保持成绩以及迁移成绩的影响，采用元分析方法，依次经过主效应检验、异质性检验、发表偏倚检验等过程，得出以下结论：首先，基于有生理数据的实验中，在多媒体学习中加入积极情绪设计能够有利于引导学生的注意分配，增加对学习材料内容的总注视时间，提高他们学习时的内在动机。同时积极情绪设计提高学习者的保持、迁移成绩。其次，积极情绪设计对学习者积极情绪、认知负荷的影响不显著。最后，积极情绪设计的效果受到学习材料主题以及语言/文化背景等因素的调节。这些研究结论为我们深入研究情感设计原理在教育和学习中的应用提供了关键线索。

当然本研究也存在一定的不足之处：第一，文献的总数量和样本太少，因此效应量过少；另外调节效应检验中部分样本量分布不均或太少，导致部分调节变量不能纳入研究进行调节效应检验；第二，本研究中大部分为加入眼动追踪的研究，故没有对其他生理数据进行元分析处理，在之后的研究中可以针对不同生理指标（如脑电图(EEG)、心率变异性(HRV)、皮肤电(GSR)等）进行分析；第三，少数研究中，在学习之前有一定的情绪诱导，这也可能对学习效果的影响有一定的调节作用，值得在未来研究中深入探讨。因此，未来研究可深入检视积极情绪设计通过生理数据分析学习影响，审视其对认知负荷的作用，以及扩展到中老年群体，评估跨年龄层次的适用性，进一步丰富多媒体学习的理论基础。

感谢

在文章即将完成之际，我向所有给予我支持和帮助的人表示衷心的感谢。首先，特别感谢我的实验室的同学们，他们在研究过程中与我并肩作战，共同克服了许多困难。此外，我还要感谢项目 The Research Funds from National NaturalScience Foundation of China (Grant No.62107016，62293555,62293550)，以及 Humanities and Social Sciences Foundationof the Ministry of Education (Grant No.21YJC880057)的支持，使我的研究得以顺利进行。再次感谢！

参考文献

王雪，高泽红，徐文文，& 张蕾．(2021)．反馈的情绪设计对视频学习的影响机制研究．*电化教育研究,*

42(03), 69-74. doi:10.13811/j.cnki.eer.2021.03.010

龚少英，上官晨雨，翟奎虎，& 郭雅薇. (2017). 情绪设计对多媒体学习的影响. *心理学报，49*(06), 771-782.

Brom, C., Starkova, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review, 25*, 100-119. doi:10.1016/j.edurev.2018.09.004

Heidig, S., Muller, J., & Reichelt, M. (2015). Emotional design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior, 44*, 81-95. doi:10.1016/j.chb.2014.11.009

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal, 327*.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441-480. doi:10.1016/0010-0285(76)90015-3

Li, J., Luo, C., Zhang, Q., & Shadiev, R. (2020). Can emotional design really evoke emotion in multimedia learning? *INTERNATIONAL JOURNAL OF EDUCATIONAL TECHNOLOGY IN HIGHER EDUCATION, 17*(1). doi:10.1186/s41239-020-00198-y

Mayer, R. E. (2006). 多媒体学习.

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning & Instruction, 33*, 12-18.

Park, B., Knoerzer, L., Plass, J. L., & Bruenken, R. (2015). Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *COMPUTERS & EDUCATION, 86*, 30-42. doi:10.1016/j.compedu.2015.02.016

Peng, X., Jin, Y., Zhao, L., Yang, B., Liu, S., & Su, Z. (2022). *Under the Condition of Stress Arousal: The Influence of Positive Emotion Design on Multimedia Learning* (Vol. 2022 Eleventh International Conference of Educational Innovation through Technology (EITT)).

Peng, X., Xu, Q., Chen, Y., Zhou, C., Ge, Y., & Li, N. (2021). An eye tracking study: positive emotional interface design facilitates learning outcomes in multimedia learning? *INTERNATIONAL JOURNAL OF EDUCATIONAL TECHNOLOGY IN HIGHER EDUCATION, 18*(1). doi:10.1186/s41239-021-00274-x

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Schmid, C. H. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj, 343*(5), d4002.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20Years Later. *Educational Psychology Review, 31*(2), 261-292. doi:10.1007/s10648-019-09465-5

Um, E. R., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology, 104*(2), 485-498. doi:10.1037/a0026609

Uzun, A. M., & Yildirim, Z. (2018). Exploring the effect of using different levels of emotional design features in multimedia science learning. *COMPUTERS & EDUCATION, 119*, 112-128. doi:10.1016/j.compedu.2018.01.002

Wang, X., Mayer, R. E., Han, M., & Zhang, L. (2023). Two Emotional Design Features Are More Effective Than One in Multimedia Learning. *JOURNAL OF EDUCATIONAL COMPUTING RESEARCH, 60*(8), 1991-2014. doi:10.1177/07356331221090845

Wang, Y., Gong, S., Cao, Y., & Fan, W. (2023). The power of affective pedagogical agent and self-explanation in computer-based learning. *COMPUTERS & EDUCATION, 195*. doi:10.1016/j.compedu.2022.104723

人工智能在教师教育中的应用：样态、挑战和优化策略

Application of Artificial Intelligence in Teacher Education: Modalities, Challenges, and Optimization Strategies

苗谦盈[1]，黄禹[1]，徐光涛[1*]
[1]杭州师范大学经亨颐教育学院
[*] xuguangtao@hznu.edu.cn

【摘要】 近年来，人工智能正逐渐被应用到教育领域，有潜力成为支持教师教育的重要工具。人工智能在教师教育中的应用样态和方式成为研究人员和教师关注的话题。在此背景下，文章采用系统性文献综述的方法，梳理了 2016-2023 年的 41 项国际研究成果。首先从文献发表情况、研究对象、国别等维度，分析了人工智能应用于教师教育研究的现状；随后，归纳出人工智能在教师教育中的 5 种应用样态和应用实践中存在的诸多挑战，并提出应对策略，以优化人工智能在教师教育的应用效果。最后，本文讨论了研究的局限性，并根据当前的研究空白提出了未来研究的建议。

【关键词】 人工智能；教师教育；应用样态；挑战；优化策略

Abstract: *In recent years, artificial intelligence (AI) has gradually been applied to the field of education, showing potential to become a significant tool in supporting teacher education. The application modes and methods of AI in teacher education have become a focal point for researchers and educators. Against this backdrop, this paper employs a systematic literature review approach to examine 41 international research outcomes spanning the years 2016 to 2023. Initially, an analysis is conducted across various dimensions, such as publication trends, research subjects, and geographic distribution, to assess the current state of AI application in teacher education research. Subsequently, the paper identifies five application modes of AI in teacher education and highlights numerous challenges encountered in practical implementation. Strategies to address these challenges are proposed to optimize the effectiveness of AI in teacher education. Finally, the paper discusses the limitations of the study and puts forth recommendations for future research based on current research gaps.*

Keywords: artificial intelligence, teacher education, modalities of application, challenges, optimization strategies

1. 前言

随着人工智能技术的迅猛发展，教育领域正迎来一场深刻而革命性的变革。人工智能已经开始在教育中扮演着日益重要的角色，为学习者提供了前所未有的个性化和高效性的学习体验。这种转变也为教师教育带来了全新的挑战与机遇。

在国际与国内的教育领域，人工智能与教育的结合逐渐成为关注的焦点。联合国教科文组织、美国教育部等权威机构纷纷发布相关政策文件，提出人工智能在教育中的潜在作用和发展方向，为全球范围内的教育改革和教师培训提供了指导和倡议。2019 年，联合国教科文组织(UNESCO)在首届人工智能与教育国际会议上提出《北京共识——人工智能与教育》，呼吁高等教育机构加强人工智能相关研究，培养具备系统设计和开发能力的人才队伍(UNESCO,

2019)。2021 年，联合国教科文组织发布《人工智能与教育：政策制定者指南》，旨在为国家领导人和政策制定者提供更好的理解，促使提供包容、公平的高质量教育(UNESCO, 2021)。2023 年，美国教育部发布《人工智能与教学的未来：见解与建议》，强调人工智能助手在减轻教学负担、提升个性化学习等方面的潜在作用(ED, 2023)。而在国内，国务院(国务院, 2018)以及教育部(中华人民共和国教育部, 2021)也通过政策文件明确了教师队伍适应信息化、人工智能等新技术的重要性，并提倡新技术融合教师队伍建设的新路径和新模式。这些政策文件共同构建了一个促进人工智能与教育融合的全球化框架，为各国在教育改革和教师教育方面提供了有益的指导和经验。然而，实施过程中需要注意在技术推广和应用中确保教育的公平和包容性，以充分发挥人工智能在教育中的潜力。

由此可见，中国和世界都将人工智能技术视为提高教师队伍质量、推进教育发展改革的重要支撑。教育界的前沿领域将人工智能视为促进教师教育发展的至关重要的力量，然而，为了充分发挥人工智能在教师教育中的潜力，我们必须以审慎的态度直面人工智能应用于教育领域所带来的种种挑战，以期提出有针对性的优化策略。

2. 研究方法

2.1. 研究工具

本研究采用的是系统性文献综述。系统性文献综述是一种可以识别、评估、解释与某一特定研究问题、主题领域或感兴趣的现象相关的所有可获得研究的方法(Kitchenham & Charters，2007)。

2.2. 数据来源

本研究的系统性文献综述遵循了 PRISMA 指南(Moher, Liberati, & Tetzlaff，2009)。为保证样本文献的质量，本研究选取比较知名且成熟的数据库进行搜索：Web of Science、EBSCO、SCIENCE Direct、Google scholar、CNKI。在前 4 个数据库中，搜索词的布尔逻辑式为((artificial intelligence OR AI)AND(teacher education OR teacher training OR teacher preparation))，在 CNKI 中以“人工智能”“AI”和“教师专业发展”为中文关键词，检索时间设为 2016 年 1 月 1 日～2023 年 12 月 31 日，文献类型为期刊论文，最终检索得到有效文献 764 篇（英文 641 篇、中文 123 篇）。经过筛选，本研究最终获得的样本文献为 46 篇（英文 27 篇，中文 19 篇）。为便于分析，本研究从题名、国籍、作者、发表年份、期刊来源、机遇、挑战与困难、技术方法等维度对 46 篇样本文献进行了编码分析。

3. 研究结果与分析

3.1. 国家/地区分布情况

本文以第一作者为准对作者分布情况展开分析。数据统计显示，41 篇样本文献的作者来自 12 个国家，其中中国作者是主要的研究群体（N = 24，59%），美国作者发文量排名第二（N = 4，10%），德国、印度尼西亚、以色列作者发文量排名并列第三（N = 2，5%），具体的按国家列出的发文量如图 1 所示；区域涵盖欧洲（N = 5，12%）、美洲（N = 4，10%）、亚洲（N = 31，75%）、非洲（N = 1，3%）。

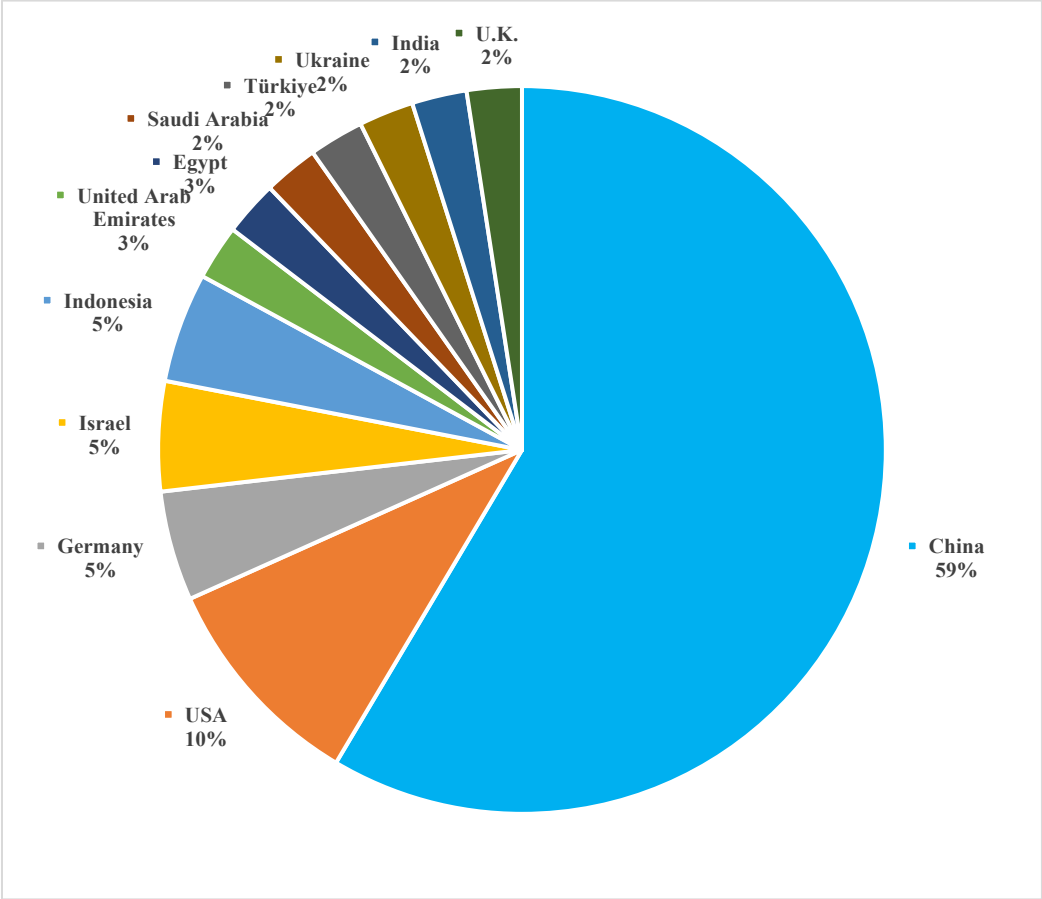


图 1 按国家列出的发文量

3.2. 发表年份分布情况

从文献发表年份的情况来看，2021、2022、2023 年的发文量增幅比较显著，具体发文情况如图 2 所示。2021～2023 年相关主体的发文量增加，可能与人工智能技术的快速兴起，各国、各类组织机构对于教师教育的重视有一定关联。从期刊来源来看，41 篇样本文献发表于 26 种期刊，发表量较多的期刊为 Education and Information Technologies、Computers and Education: Artificial Intelligence、中国电化教育、现代教育技术。从学科背景来看，人工智能技术在教师教育中的应用设计计算机学科、教育学科等多学科领域。



图 2 2016～2023 年样本文献的数量（单位：篇）

3.3. 教师教授学科分布情况

在被筛选出的文献中，只有一部分表明了接受教师教育的教师所教授的科目。教授的科目包含计算机、科学、数学、音乐、生物、英语、语文。其中研究对象为教授数学的教师的研究最多（n＝3，27%）。研究对象教授科目呈现多元化的特征。

3.4. 职前教师与在职教师数量分布情况

教师教育的对象包含在职教师和职前教师，因此筛选到的研究中也有这两种研究对象。在有报告教师类型的研究中，研究对象为在职教师的研究有 11 个，研究对象为职前教师的研究有 6 个，研究对象为在职教师的研究更多。

4. 人工智能在教师教育中的应用形态

人工智能技术主要包含机器学习技术、深度学习技术、自然语言处理技术、计算机视觉、专家系统等。这些技术应用方式不同就产生了不同的应用形态。根据纳入综述的研究，主要可以总结出 5 种应用形态：多样化的数据收集、自动评估反馈、虚拟学生代理、算法预测、基于人工智能技术的应用平台。

4.1. 多样化的数据收集

人工智能技术可以实现多样化数据的实时收集，为教师自我反思、提高教学能力提供数据支撑。传统的数据收集以学生作业情况、考试情况、教师备课情况等为主，而基于人工智能技术的数据收集，得益于深度学习算法的支持，可以实现更多样的数据收集。

在真实课堂中，可以通过语音识别技术收集学生和教师的发言(al-Zyoud, 2020)，不仅如此，还可以收集到他们语气调节数据、学生参与水平、内容教授节奏等较为细节的数据(Barrett, Ke, Dai, West, Bhowmik, & Zhang, 2023)。同时，基于图像识别技术对教师和学生的捕捉，人工智能可以通过肌肉细节识别出学生的认知情感状态(Barrett, Ke, Dai, West, Bhowmik, & Zhang, 2023)。在虚拟课堂中（如微格教室），职前教师也可以通过人工智能技术获得自己的讲授数据，提升自己的教学能力。例如，有研究者利用阿凡达技术创建模拟课堂场景，在这样的环境中让职前教师锻炼自己的教学技能，同时进行收集、分析职前教师的教学数据，使得反馈建议能够量身定制。此外，该系统中的机器学习模型的集成意味着，随着时间的推移，摄入的数据越来越多，洞察力会变得更加敏锐，对于教学实践的指导更加精细(Kusmawan, 2023)。总的来说，人工智能驱动的数据收集更加细微，使用的深度学习算法和机器学习模型更加先进，为更加精准的教学指导提供数据支撑。

4.2. 自动评估反馈

人工智能在对教师的自动评估反馈方面也有较多应用。如在美国应用较多的电子评估系统 Assisments，就受到人工智能系统“认知导师”的启发，用到认知导师项目中的模型和示例追踪范式(Porayska-Pomsta, 2016)。除此之外，还有人工智能代理可以分析课堂上的实时数据，例如提问技巧、行为、对于教学内容知识诊断的反应，基于此提供对教学的评估和建议，且人工智能的评估反馈足够客观，因为教师不会因为批评而感到冒犯(Chiu, Xia, Zhou, Chai, & Cheng, 2023)。还有研究通过教研平台的人工智能分析算法，智能化分析教师课堂教学基本情况，从而为教师形成个性化的课堂教学分析报告(蔡慧英、卢琳萌、顾小清, 2023)。人工智能基于课堂数据进行实时评估与反馈，提供实时精准的改进建议，可以帮助教师反思教学实践，提升教学能力。

4.3. 虚拟学生代理

得益于自然语言处理技术的发展，以 ChatGPT 为代表的人工智能技术迅速兴起并获得了广泛的使用，虚拟学生代理也应运而生。在教师教育中，ChatGPT 主要被用于话语训练器，帮助职前教师掌握真实课堂中的对话技能(Kortenkamp & Dohrmann, 2023)。早在 ChatGPT2.0 发布时，就有研究者使用基于生成式预训练变压器的深度神经网络模型，利用真实课堂对话

中的语言资源数据创建了一个人工智能聊天机器人(Bhowmik 等, 不详)。随着 ChatGPT3.5 版本的发布，其自然对话功能大大增强。同时有研究发现 ChatGPT 在表现出某些错误、或者纠错的时候与真实学生的表现有相像之处(Kortenkamp & Dohrmann, 2023)，因此，职前教师在初期训练时可以与 ChatGPT 进行对话，学习与学生的沟通方式，通过与 ChatGPT 的不断对话和对其纠错，可以帮助师范生获得诊断和沟通技能。目前，已有职前教师通过虚拟教室中的虚拟对话培养沟通和教学技能，并取得较好的效果(Asst. Prof., 2022)。

在具有自然对话能力的基础上，人工智能可以渲染基于模拟的虚拟代理，反映真实的认知情感状态。有研究就构建了在 3D 虚拟世界 OpenSimulator 中集成 AI 的虚拟学生代理，它使用基于 Transformer 的生成式预训练深度神经网络模型进行编程,模型经过真实的 STEM 课堂对话训练，能够根据四种典型的功能类型：陈述式、疑问式、祈使式和感叹式，通过按语句功能的频率对转录本进行编码来分析数据(Barrett, Ke, Dai, West, Bhowmik, & Zhang, 2023)，服务于教师教学能力的成长。在未来，教师可以在实验室的环境中获取诊断和交流技能，而无需学生的真正参与。

4.4. 算法预测

有研究利用人工智能的三种算法对高等师范教育的绩效数据进行了分析，并且在宏观上对其进行了预测，显示了人工智能算法在职前教师教育绩效数据的预测分析作用。该研究主要利用三种不同的人工智能技术：基于免疫算法的支持向量机、基于粒子群优化的支持向量机、纯支持向量机。利用上述三种人工智能技术，对高等师范教育培训涉及的组织、任务、人员三个主要模块对应的绩效数据进行了预测和分析(Deng, Jia, & Chai, 2022)。多种人工智能技术的预测结果表明，基于免疫算法优化的支持向量机方法对应的判定系数最大，均方根误差最小。人工智能技术对三个模块指标都有较好的预测效果(Deng, Jia, & Chai, 2022)。可见，人工智能的预测算法可以服务于教师教育相关数据的预测与分析。

4.5. 基于人工智能技术的应用平台

基于人工智能技术的应用平台实质上是对上述部分功能的整合优化，可以实现一些综合性的功能。例如，近些年来出现的专家学习系统、智能教学系统等能为教师教学和反思提供了个性化的分析报告(Salas-Pilco, Xiao, & Hu, 2022)。还有的人工智能应用系统可以为教师提供诊断测试、综合科学主题、综合科学课程、在线论坛、综合科学学习中的课程建议。通过人工智能系统辅助培训，他们可以根据自己的初步知识参加课程和在线论坛，系统给出的指导可以作为参与者更新知识的参考(Gunawan, Liliasari, Kaniawati, & Setiawan, 2021)。此外，Waston 智能备课系统可以实时记录学生的学习状况数据。通过大数据智能分析，结合学习情况分析,根据具体章节的知识点向教师提出相应的教学建议,能够帮助新手教师快速成长(Liu, Chen , & Yao, 2022)。

5. 讨论

本次综述的目的是基于文献证据，系统地审查人工智能在教师教育中应用的现状。结果表明，人工智能在教师教育中的应用形态包括多样化的数据收集、自动评估反馈、虚拟学生代理、算法预测和基于人工智能技术的应用平台。这些应用的目的要是通过数据分析、虚拟对话和个性化指导促进教师对教学的反思和改进，帮助教师成长。

然而，将人工智能应用于教师教育中，会带来许多不同的挑战。人工智能应用于教师教育的挑战包括对教师身份认同的弱化、对“双师”协同融合理念的冲击、伦理问题以及不完善的教师教育课程体系。在人工智能时代，智能教育环境下学生成为知识创造的核心，人工智能技术如 ChatGPT 提供个性化学习指导，挑战传统以教师为中心的教学模式，削弱了教师地位，迫使其重新思考教学方法和评价(张一春、汤玲和马春兰，2023)。过度依赖人工智能工

具可能对教师的专业情感和自我认同产生负面影响，威胁教师职业发展的稳定性(吴军其、吴飞燕、文思娇、张萌萌和王嘉桐，2023)。ChatGPT 的介入构成“双师”协同模式，与人类教师协作，但当其超越态势时可能削弱教师的“立德树人”教育观，限制教师创新精神和自主能动性，固化教学改革(吴军其、吴飞燕、文思娇、张萌萌和王嘉桐，2023)。此外，人工智能技术的出现引发伦理问题，包括潜在偏见加剧社会不平等、对学生和教师数据隐私的担忧以及与人性特质相关的风险(Afiya，2023)。在教师教育课程方面，存在不完善的体系，缺乏统一的课程标准，教育者和师资在技术接受与整合方面存在不足，制约了教师教育课程充分发挥智能技术赋能的潜力(赵磊磊、张黎、代蕊华和王晓茜，2022)。此外，ChatGPT 产生的知识可能存在碎片化，使教师专业知识发展面临系统化构建的挑战(柳立言、张会庆和闫寒冰，2021)。这些挑战需要综合考虑，以更好地应对人工智能在教师教育中的应用。

本研究认为可以通过如下手段面对这些挑战，优化人工智能在教师教育中的应用。在人工智能优化教师教育的策略中，首先强调了打造智能环境，通过在城区学校进行试点，建设学习平台和智慧教育环境，利用人工智能为教师精准画像，助推教师专业发展(罗强，2022)。通过赋能、超越和共生的方式，实现智能研修平台与教研机制深度融合，形成共生效应(李阳和曾祥翊，2022)。其次，强调了重塑教师角色的重要性。鉴于人工智能只能发挥辅助作用，教师在“AI”时代必须成为学生的“灵魂的工程师”和情感的呵护者，加强智能教育培训，提升信息理念和意识，理性自觉强化教育洞察力(宋灵青和许林，2022)。最后，强调了在“AI”时代需要变革教学方式。教师的工作重点应放在推动以自主、合作、探究为主要特征的教学方式变革上，以培养学生一系列机器无法具备的核心素养和能力(宋灵青和许林，2022)。人工智能技术在促进个性化学习方面发挥作用，帮助克服人脑在处理注意力、知觉、知识获取与知识转换过程中的局限性(罗强，2022)。这使得教师能够专注于辅助“学弱生”和“学优生”，对教师的专业发展有裨益(王宇和汪琼，2022)。

这项系统评价有一些局限性。首先，它侧重于教师教育，包括教育培训项目中的在职教师和职前教师。未来的研究可以扩大范围，并将未参加教育培训项目的教师纳入其中。其次，本综述重点关注人工智能的部分技术，主要包括机器学习、深度学习、计算机视觉、自然语言处理在教师教育中的应用。未来的研究可能包括其他人工智能技术和工具。最后，不能保证找到每一篇相关文献。尽管如此，本研究有助于分析人工智能在教师教育中的应用现状。

6. 总结

本次系统性文献综述重点介绍了人工智能在教师教育中的应用。人工智能正在逐渐被用来实现更高质量的教师教育。通过审查，本研究归纳出五种人工智能技术应用于教师教育的形态，这讲帮助教师在数据收集分析、虚拟对话和个性化指导方面获得帮助。然而，在具体应用实施的过程中，仍在存在对教师身份认同的弱化、对“双师”协同融合理念的冲击、人工智能应用于教师教育的伦理问题、不完善的教师教育课程体系的问题。本研究认为，通过打造智能环境，以点促面引领教师成长、重塑教师角色、变革教学方式，可以应对这些挑战，优化人工智能在教师教育中的应用。

参考文献

蔡慧英,卢琳萌 顾小清. (2023). 人机协同教研会促进教师教学反思能力的发展吗?——基于课堂视频智能分析技术的实证研究. 现代远距离教育, 1, 40－49.

李阳 曾祥翊.(2022).人工智能赋能教研高质量发展：智能精准教研的理论框架、实践蓝图与发展脉络.中国电化教育(11),99-107+122.

柳立言,张会庆 闫寒冰.(2021).智能时代乡村教师专业发展的困境、机遇和实践路径.中国电化教育(10),105-112.
罗强.(2022).智能时代教师知识结构的发展框架及其实现路径.现代教育技术(07),31-39.
宋灵青 许林.(2018).“AI”时代未来教师专业发展途径探究.中国电化教育(07),73-80.
王宇 汪琼.(2022).人工智能助推教师专业发展的若干思考.中国远程教育(01),12-19+92.doi:10.13541/j.cnki.chinade.2022.01.002.
吴军其,吴飞燕,文思娇,张萌萌 王嘉桐.(2023).ChatGPT 赋能教师专业发展：机遇、挑战和路径.中国电化教育(05),15-23+33.
张一春,汤玲 马春兰.(2023).人工智能助推教师发展的路径与对策研究.电化教育研究(10),104-111.doi:10.13811/j.cnki.eer.2023.10.014.
赵磊磊,张黎,代蕊华 王晓茜.(2022).人工智能赋能教师教育:基本逻辑与实践路向.中国教育学刊(06),14-21.
中共中央国务院.(2018).中共中央国务院关于全面深化新时代教师队伍建设改革的意见.http://www.gov.cn/zhengce/2018 -01/31/content_5262659.html.
中华人民共和国教育部.(2021).教育部关于实施第二批人工智能助推教师队伍建设行动试点工作的通知.http://www.moe.gov.cn/srcsite/A10/s7034/202109/t20210915_563278.html.
Asst. Prof., Curriculum and Instruction in Computer and Educational Technology, Prince Sattam bin Abdulaziz University, Saudi Arabia, h.elgohary@psau.edu.sa, Elgohary, H. K. A., Al-Dossary, H. K., & Lecturer of curricula and methods of teaching mathematics, Prince Sattam bin Abdulaziz University, Saudi Arabia, mmrcomputer2013@gmail.com. (2022). The effectiveness of an educational environment based on artificial intelligence techniques using virtual classrooms on training development. *International Journal of Instruction*, 15(4), 1133–1150.
Afiya Jamal. (2023). The role of artificial intelligence (AI) in teacher education: Opportunities & challenges
al-Zyoud, H. M. M. (2020). The role of artificial intelligence in teacher professional development. *Universal Journal of Educational Research*, 8(11B), 6263–6272.
Barrett A., Ke F., Dai C.-P., West L., Bhowmik S., & Zhang N. (2023). AI-integrated virtual students for teacher training: Comparing simulation-based classroom dialogue with the real thing.
Chiu T. K. F., Xia Q., Zhou X., Chai C. S., & Cheng M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118.
Deng, H., Jia, W., & Chai, D. (2022). Discussion on innovative methods of higher teacher education and training based on new artificial intelligence. *Security and Communication Networks*, 2022, 1–10.
Gunawan, K. D. H., Liliasari, L., Kaniawati, I., & Setiawan, W. (2021). Implementation of competency enhancement program for science teachers assisted by artificial intelligence in designing HOTS-based integrated science learning. *Jurnal Penelitian Dan Pembelajaran IPA*, 7(1), 55.
Kitchenham B A, & Charters S. (2007) . Guidelines for performing systematic literature reviews in software engineering. *UK: University of Durham: Technical report*, EBSE Technical Report EBSE,3

Kortenkamp U., & Dohrmann C. (2023). Pre-service teacher trainging with AI: Using ChatGPT discussions to practice teacher-student discourse.

Kusmawan, U. (2023). Redefining teacher training: The promise of AI-supported teaching practices. *Journal of Advances in Education and Philosophy*, 7(09), 332–335.

Liu, Y., Chen, L., & Yao, Z. (2022). The application of artificial intelligence assistant to deep learning in teachers' teaching and students' learning processes. *Frontiers in Psychology*, 13, 929175.

Miao F., Holmes W., & Huang R.(2021). *AI and education: guidance for policy-makers*. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000376709

Moher D., Liberati A., & Tetzlaff J. (2009). Preferred reporting items for systematic reviews and meta-analyses:the PRISMA statement. *BMJ*, 2009, 339 (undefined) :b2535.

Porayska-Pomsta K. (2016). AI as a methodology for supporting educational praxis and teacher metacognition. *International Journal of Artificial Intelligence in Education*, 26(2), 679–700.

Salas-Pilco, S., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, 12(8), 569.

UNESCO.(2019). *Beijing consensus on artificial intelligence and education*. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000368303

U.S. Department of Education's Office of Educational Technology.(2023).*Artificial Intelligence and the Future of Teaching and Learning*. Retrieved from https://tech.ed.gov/

教育数字化转型背景下中小学多师协同教学方式的特征与思考

Characteristics and Thinking of Multi-teacher Collaborative Teaching Methods in Primary and Secondary Schools in the Context of Digital Transformation of Education

潘丹[1*]，杨明欢[2]
[1]华南师范大学教育信息技术学院
[2]广东省电化教育馆
[*] 2325349427@qq.com

【摘要】 随着数智技术的不断变革，多师协同教学形式越来越多样。本研究以协同学、系统论、教学设计理论为理论依据，结合多师协同教学方式的内涵与影响因素，构建了教育数字化转型背景下中小学多师协同教学方式分析模型，对来自广东省中小学的多个多师协同教学案例进行系统分析，总结形成了三类教育数字化转型背景下典型中小学多师协同教学方式，并分析了教育数字化转型背景下中小学多师协同教学的典型特征、存在的问题，并基于此提出解决中小学多师协同教学问题的建议，以期满足当前教育数字化转型背景下中小学多师协同教学创新的需要。

【关键词】 多师协同教学；教学方式；教育数字化转型；中小学；特征

Abstract: *With continuous change of digital intelligence technology, the forms of multi-teacher collaborative teaching are becoming more and more diverse. Based on collaborative theory, system theory and instructional design theory, combined with connotation and influencing factors of the forms of multi-teacher collaborative teaching, this study builds an analysis model of the forms of multi-teacher collaborative teaching in primary and secondary schools under education digital transformation's background, and systematically analyzes multiple multi-teacher collaborative teaching cases from primary and secondary schools in Guangdong Province. This paper summarizes and forms three typical multi-teacher collaborative teaching methods, analyzes the typical characteristics and existing problems of multi-teacher collaborative teaching, and puts forward suggestions to solve the problems of multi-teacher collaborative teaching based on this, in order to meet the needs of multi-teacher collaborative teaching innovation.*

Keywords: multi-teacher collaborative teaching; teaching methods; digital transformation of education; primary and secondary schools; characteristics

1. 问题的提出

为贯彻落实教育数字化战略行动和《教师数字素养》等有关文件精神，广东省教育厅组织开展了广东省教育“双融双创”教师信息素养提升实践活动融合创新应用教学案例优秀作品提炼培育及示范辐射工作（广东省教育厅，2023）。其中，多师协同教学案例由于其复杂性及涉及范围广而成为了热点话题。教育数字化转型为多师协同教学提供了更广阔的应用场景和工具，增强了教学的灵活性和效果，提供了更加丰富、个性化的教学资源和支持，促进了教师之间的合作与专业发展。但在实施过程中，由于存在对多师协同教学的理论认识不足、教师团体分工及协作标准不够明晰、多师协同互动质量和频率不高等问题，导致教师之间缺乏默契，协同教学效果不佳，无法充分发挥多师协同教学的优势。基于此，本研究以协同学、系统论、教学设计理论为理论基础，结合多师协同教学方式的内涵与影响因素，构建数字化

转型背景下多师协同教学方式分析模型，分析教育数字化转型背景下多师协同教学方式的特征和多师协同教学亟待解决的问题，并提出解决多师协同教学问题的建议，以期解决当前教育数字化转型背景下中小学多师协同教学理论指导和实践创新的迫切需要，同时也为新时代我国多师协同教学可持续发展提供有益借鉴。

2. 多师协同教学方式的理论基础及影响因素

2.1. 多师协同教学方式的理论基础

教学方式是教学主体为达成教学目标、实现教育价值而采用的动态化教学行为序列及系统化教学活动的总和（谢幼如等，2020）。协同学倡导通过各要素的协同保持事物整体性的状态和趋势，系统论强调系统各部分之间相互依赖、互为资源时才能发挥整体的作用（周慧敏等，2023），教学设计是指应用系统方法分析教学问题，确定解决教学问题的方法、策略和步骤，并对结果作出评价的计划过程和操作程序（谢幼如，2016）。其为教育数字化转型背景下课堂教学创新带来了重要启示，为多师协同教学奠定了理论基础。协同思维及系统思维的融入革新了原来单一教师教学的教学方式，而数智技术的变革打破了物理空间壁垒，转变了原有同一场域协同教学的教学方式，为实现协同教学方式的创新提供了技术支持。多师协同教学方式不是简单的教师之间的协同合作，是教师、学生、现实环境、虚拟环境的多方协同融合，也代表着以数智技术为支持的课堂教学要素的革新。

2.2. 多师协同教学方式的影响因素

李贺（2007）认为影响师生协同合作的因素为师生协同合作本身的特点、教师、学生、课程，唐静静等（2023）认为影响协作学习行为的因素包括学生个体、小组群体、学习任务、学习环境，彭梓涵等（2019）认为影响影响小组协作学习的因素有协作学习自我效能感、组织有效性、心理信任、学习支持服务；张靖等（2022）提出了影响双师课堂教学的因素包括双师教学职能划分、课堂教学关系结构、辅助教师职能、硬件环境、共进机制。因此，结合教学方式的定义，本研究认为协同教学目标、教学环境、教师、学生、教学规则、教学资源、教学活动、教学评价等是影响多师协同教学方式的主要因素。

3. 教育数字化转型背景下中小学多师协同教学方式的分析

3.1. 多师协同教学的模型构建

3.1.1. 主要依据

教学方式是教学活动的表现形式。影响多师协同教学的因素主要包括协同教学目标、 教学环境、教师、学生、协同规则、教学资源、教学活动、教学评价等八个方面。多师协同教学方式不仅代表数智技术支持的教学方式多方协同化，更展现出以数智技术作为创新要素对多师协同教学八个基本要素的重构和变革。面对教育优质均衡创建，教育部号召通过“双师课堂”改革，引进教育发达地区或本地区高水平师资与课程资源，与乡村学校及薄弱学校教师协同发展（教育部，2018）。同时教育部要求加大国家智慧教育平台在双师课堂中的应用，促进优质教育资源广泛共享（教育部，2023），由此可见，教育数字化转型背景下多师协同教学方式不仅需要有效支持学生个性化学习，也应达成教育资源共建共享及教研联动的目标。

3.1.2. 分析模型构建

综合上述分析，本研究构建了教育数字化转型背景下中小学多师协同教学方式分析模型，如图 1 所示。该模型在多师协同教学理论指导下和多师协同教学环境的支持下，制定多师协同教学规则、应用和开发多师协同教学资源、重构多师协同行为活动、开展协同多元教学评价，最终达成教育数字化转型背景下多师协同教学的目的。

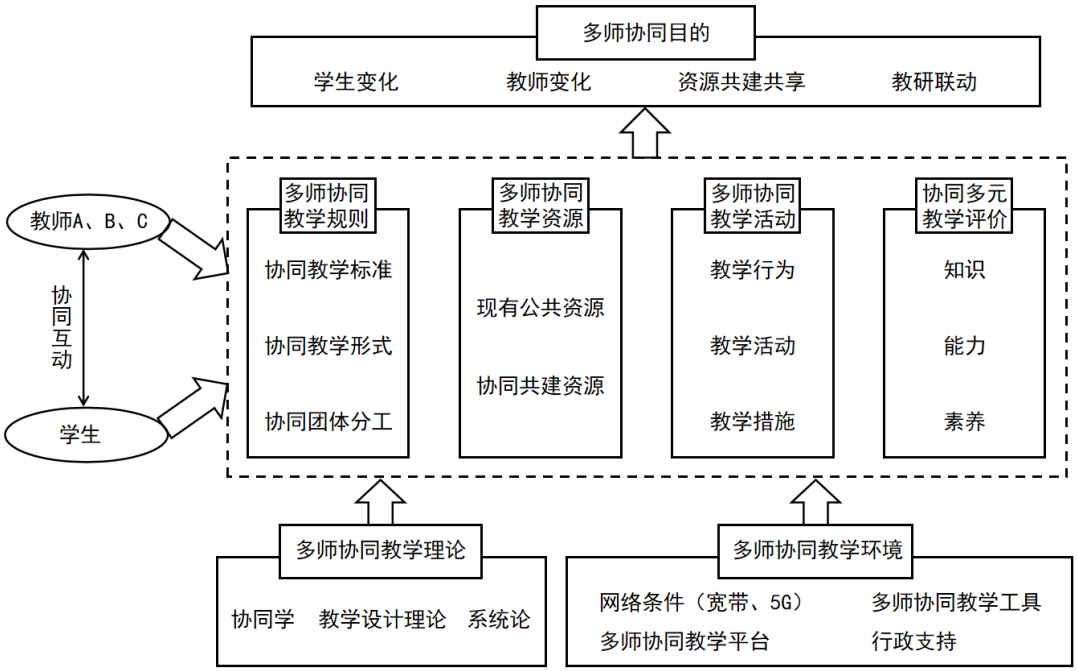


图 1 教育数字化转型背景下中小学多师协同教学方式分析模型

3.2. 分析框架构建

本研究根据图 1 所示的教育数字化转型背景下中小学多师协同教学方式分析模型，转化为表 1 所见的分析维度及其内涵，从多师协同教学理论基础、多师协同教学环境、教师、学生、多师协同教学规则、多师协同教学资源、多师协同教学活动、协同多元教学评价 8 个维度，对教育数字化转型背景下中小学多师协同教学方式进行分析。

表 1 数字化转型背景下中小学多师协同教学方式分析框架

一级维度	二级维度	描述性指标
1. 多师协同教学理论基础	协同教学理念	融合协同学、系统论思维，立德树人，培养全面发展的人，开展协作探究教学、跨学科教学等，促进学生协作、交流、跨文化理解等能力的提升
	协同教学原则	遵循合作性、互动性、系统性等教学原则
	协同教学设计	在协同学、系统论等理论指导下，确定教学目标、重塑教学内容、重构教学结构、再造教学流程、重整教学评价
2. 多师协同教学环境	网络条件	具备多师协同教学的网络环境，如宽带、5G 网等
	多师协同教学平台	具备支持多师协同教学的平台，如智能化教学互动平台、直播平台、协作教研平台等
	多师协同教学工具	具备支持不同教学场景的教学工具，如直播工具、协作工具、学科教学工具等；具备学情分析工具、问卷调查工具等
	行政支持	学校行政管理层支持多师协同教学
3. 教师	多师协同教学认识	了解协同学和系统论等理论基础，认识到开展多师协同教学的价值和挑战
	学科教学水平	具备较强的学科教学水平，具有教育教学理论、教学设计、学科教学等方面的专业知识和技能
	协作交流能力	具有较强的合作能力、交流表达能力和问题解决能力
	数字素养	具备良好的数字素养，能保证多师协同教学的顺利开展

4. 学生	协作交流能力	具有较强的合作能力、交流表达能力
	数字素养	具备良好的数字素养，能利用数字学习工具、平台、资源进行学习
5. 多师协同教学规则	协同教学标准	明确多师协同教学的技术规范、质量要求、协同频度
	协同教学形式	明确教学形式是多师同步直播教学或多师同堂线下教学
	协同团体分工	明确教师之间的关系、师生关系、生生关系
6. 多师协同教学资源	现有公共资源	应用公共教育平台的微课、网络视频、音频等
	协同共建资源	协同开发多师协同教学的课件、微课、学历案、练习题等
7. 多师协同行为活动	教学行为	授课、推送学习资源、设疑提问、互动交流、个性化指导、组织管理、发布评价表、学情反馈等
	教学活动	组织同步直播教学、合作探究教学、项目式教学、大单元教学、跨学科教学、校企合作教学等
	教学措施	明确学生协同互动规则、学情反馈、提供学习平台、工具及资源、维护教学秩序、执行教学计划、开展个性化指导等
8. 多师多元教学评价	知识	学生知识的内化、技能的获得等
	能力	知识生成能力、互动交流能力、问题解决能力、跨文化理解和跨地区的协作学习能力等
	素养	学生形成正确的世界观、人生观、价值观等

3.3. 案例选取

本研究结合已有关于多师协同教学方式的文献梳理和2023年数字化转型赋能高质量课堂的广东实践，从广东省教育“双融双创”教师信息素养提升实践活动部分优秀作品提炼培育及示范辐射活动（来自广东省教育双融双创智慧共享社区平台）中选取了小学、初中、高中多个学段，语文、数学、英语、化学、人工智能、信息技术、美术等多个学科的10个多师协同教学典型案例。

3.4. 教育数字化转型背景下典型中小学多师协同教学方式分析

根据教师协同团体分工的不同，分析总结出了以下三种协同教学方式，包括主讲制多师协同教学、平行式多师协同教学、校企合作多师协同教学。

3.4.1. 主讲制多师协同教学

主讲制多师协同教学是指通过数智技术的赋能作用，主讲教师与其他教师协同合作，主讲教师负责讲授教学内容及与学生协同互动，其他辅助教师负责课堂组织管理、个性化指导学生的教学方式。主讲制多师协同教学中，主讲教师与辅助教师协同制定教学计划，协同备课，协同互动教学，协同评价与反思，实现个性化学习和优质资源共享，其中主讲老师承担课程的总体框架建设和主要讲解。这就要求主讲教师具有丰富的教学经验，能掌控教学节奏与进度，引领、指导辅助教师参与备课，了解教学理念，学会教学方法，并且丰富学科知识。对于直播教学而言，远端学生受益更大，远端学生借助数智教学平台和工具参与前端教师（主讲教师）组织的教学活动（郭炯等，2017）。《双师课堂——等腰三角形的分类讨论》多师协同教学案例中，主讲教师与辅助教师协同进行教学设计，共同开发及共享教学资源，按照教学设计环节逐步推进教学，并实时反馈两地学生的学情。辅助教师根据各教学环节的要求，维护课堂秩序，组织学生参与学习活动，并对学生进行个性化指导。最后，主讲教师和辅助教师协同完成教学评价并进行总结和反思。

3.4.2. 平行式多师协同教学

平行式多师协同教学是指多位教师平等协同合作，每位教师负责讲授部分教学内容或部分课时的教学方式。在平行式多师协同教学中，每个教师负责讲解课程的一个部分或一个课时，其特点在于学生可以获得多个教师的授课，学习不同的思路和方法，拓宽知识面；教师之间、师生之间的互动更加频繁和多样化，有助于激发学生的学习兴趣和提升学习参与度；多个教师可以共享彼此的资源和经验，提高教学质量和效果。《“三师四同步智慧课堂”大单元案例—以听说课为例》多师协同教学案例中，三位授课老师基于学情，通过远程互联网进行线上深度教研，共研共商，反复磨课，当主讲老师授课时，学生与主讲老师进行互动。《华夏艺匠——中国传统建筑（斗拱）》多师协同教学案例采取同课同构的课例研究方式，寻找高中语文、高中美术、高中通用技术共同的中国建筑作为切入点，共同研究主题、设计教学目标、改进教学设计、融合数智技术辅助教学，进行反思再加以改进。

3.4.3. 校企合作多师协同教学

校企合作多师协同教学是指学校教师与企业教师通过数智技术的赋能作用协同合作，学校教师讲授教学理论知识，并组织同步直播教学活动或实地考察与实践操作活动，企业教师讲授实践知识，展示实践过程，最终实现产教融合协同育人的教学方式。学校邀请企业专业人员进入直播课堂，为学生分享实践经验和行业知识。此外，学校还可以利用企业提供的设备和场地，组织学生进行实地考察和实践操作，以提升学生的实际操作能力和职业素养，为学生提供更加贴近实际的学习体验，了解企业的实际运作和行业发展趋势，学校教师也可以不断提升自己的专业素养和实践能力。如《匠心炼石 钢铁强国—铁的冶炼》多师协同教学案例以校企协同育人为基础，将企业和学校的资源进行有机整合，使课堂教学更加贴近生产生活实际，打破了传统教学中学科知识割裂、与社会脱节的壁垒。

4. 教育数字化转型背景下中小学多师协同课堂教学方式的典型特征

本研究以分析广东省中小学多师协同教学情况为目的，采用内容分析方法对中小学多师协同教学案例的指标项与描述内容类属编码后定性比较分析，以提炼中小学多师协同教学的典型特征与发展趋势，找出中小学多师协同教学中存在的问题。具体分析结果如图 2 所示。

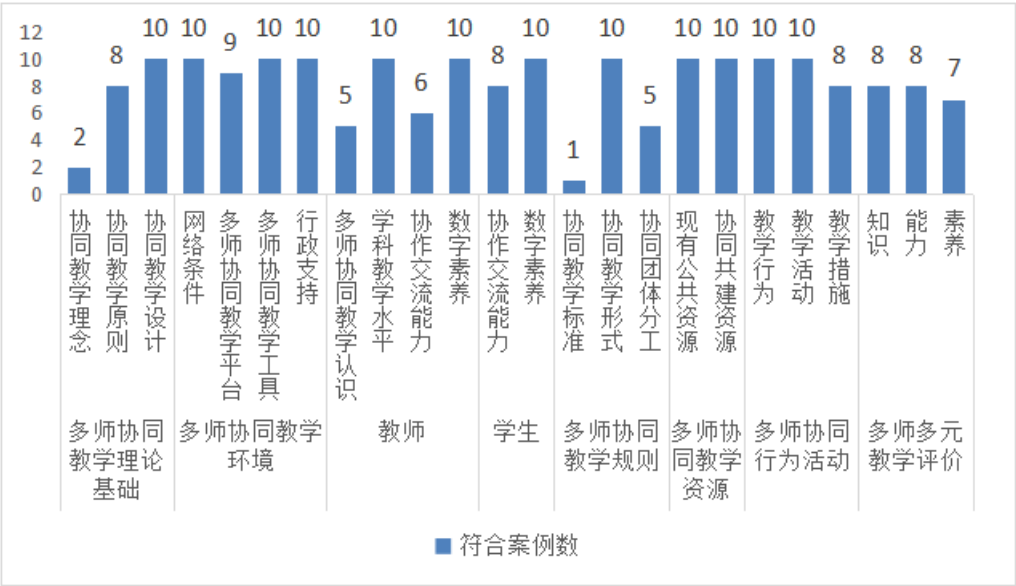


图 2 数字化转型背景下中小学多师协同教学方式分析结果

4.1. 多师协同教学环境完备且走向数智化

案例分析发现，在教育数字化转型背景下，中小学在教育部门和相关企业的支持下，均可保障网络环境畅通运行，多师协同教学平台占比 90%，说明大部分案例运用了智能化教学互动平台、直播平台、协作教研平台。多师协同教学工具及行政支持所占比例均为 100%，说

明教师均使用了学科教学工具、问卷调查工具等数字化教学工具以支持多师协同教学活动顺利多样地实施，且教育行政管理层大力支持教师开展多师协同教学实践。如《“三师四同步智慧课堂”大单元案例——以听说课为例》案例中，三位教师借助数智技术对英语教与学的支持和服务功能，创新使用数智化工具、资源，如希沃一体机、广州共享课堂、ClassIn 直播教学平台、翼课网等开展同步教学，用词云图直观展示高频词，用语音合成工具制作教学素材，并用剪辑工具制作视频创设真实情境。

4.2. 多师协同教学资源丰富且呈现形式数智化

案例分析发现，现有公共资源及协同自建资源均为 100%，说明现有多师协同教学资源较为丰富，中小学教师基本能够结合自身实际情况，恰当选用国家公共教育资源、网络开放资源、区域教学资源、校本特色资源，为本校开展网络教学工作提供多元资源体系。但由于其需要满足多个班级学生的学习需要，大部分教师协同开发了符合自身教学需要、学习者特征和思政育人目标的多师协同教学资源，通过教师间的协同合作，实现教学资源的共建共享。如《双师互动课堂让每个孩子获得更优质的艺术教育》案例中，教师们根据小学美术教育目标和教学计划，协作开发了适合双师教学的教学内容和数字化教学资源，并利用网络化平台实现了优质教育资源的共享共建。

4.3. 多师协同教学活动多样且融入大单元、跨学科等特色

案例分析发现，教学活动及教学行为均为 100%，说明教师能够结合自身实际情况开展灵活多样的教学活动。且 4 个案例融入跨学科教学特色，2 个案例融入大单元教学特色，1 个案例融入项目式教学特色。如《华夏艺匠-中国传统建筑（斗拱）》案例将语文、美术、通用技术 3 个学科融合，两位教师开展了练习、测评、小组合作探究、实物教具演示、自制微课讲解、视频播放等多样化教学活动以实现课堂教学多样化、趣味化。《基于腾讯会议开展两校双师混合式教学-机器诗人》案例中，学生结合学习内容，编写机器生成近体诗的程序，同时运用语文学科知识，对机器生成的近体诗进行修改，运用跨学科知识解决实际问题。

4.4. 协同教学评价多元化且注重对学生协作能力和问题解决能力的评价

案例分析发现，知识与能力均为 80%，素养为 70%，说明大部分教师注重综合运用在线测试、问卷、项目作品自评、互评、师评等方式对学生的知识、能力、素养进行多元多维评价，注重提升学生知识素养，培养学生知识生成能力、互动交流能力、问题解决能力、跨文化理解和跨地区的协作学习能力等，从而促进协同教学目的达成。如《基于 ClassIn 三校三师联合教学—电子门牌的设计与制作》案例中，教师将学生协作学习能力作为学历案自评表的一个维度。《匠心炼石 钢铁强国—铁的冶炼》案例中，教师将化学学科知识、批判性思维和自主创新能力、责任担当、工匠精神纳入评价标准，以多方位评价学生的学习效果。

5. 思考

5.1. 协同教学的理论认识不足，需强化多师协同教学理论认识

案例分析发现，只有 2 个案例提及了协同学、系统论、教学设计理论，多师协同教学认识占比为 50%，说明有一半教师对多师协同教学的定义和内涵理解不够深入，尤其需要深度学习数字化教学设计理论。因此，教师应充分考虑多师协同教学的协作性，在协同学和系统论指导下，协同制定多师协同教学规则，协同开发教学资源、再造教学流程、制定多元多维协同教学评价量表，注重课程思政融入，融合大单元、项目式、跨学科等特色，创新教学方式，以达成多师协同教学目标。

强化多师协同教学理论认识。首先要深入理解多师协同教学的含义与形式，这种教学方式强调教师之间的协作与配合，共同完成教学任务，实现教学目标。其次要掌握多师协同教学的理论基础，主要包括协同理论、系统理论和教学设计理论等。这些理论为多师协同教学

提供了重要的理论支撑和实践指导，有助于深入理解其内在机制和优势。再次，要明确多师协同教学的价值与意义，认识到多师协同教学有助于提高教学质量和效果，培养学生的综合素质和创新能力。

5.2. 教师团体分工不够明晰，需明确教师团体分工和协作标准

案例分析发现，协同教学标准占比仅为 10%，说明教师们不注重多师协同教学标准的制定，没有明确协同教学的技术规范、质量要求、协同频度。50%的案例中，教师没有明确教师的角色及分工，部分教师没有厘清教师间关系及教师与学生的协同交互方式。因此，教师应明晰自己及协同教师的角色定位，包括讲授者、讲授时长、辅助者、以及与自己学生及其他班级学生的关系，明确知识讲授、组织管理、学习支持服务、习题讲解、实验演示、程序操作等教学活动由谁负责，同时也要制定相应的协同教学标准，明确协同教学的要求。

明确教师间关系。首先，协同教师要明确教学目标和计划，确保每个教师都清楚自己的任务和角色，避免教学过程中的混乱和重复。其次，要建立科研引领-教师发展中心培训-区域组织-学校实施-教研指导-技术支持的联动机制，通过科研、教发中心、区域、学校、教研、技术公司的协同联动，制定出争对不同角色分工的多师协同教学的标准，并实现区域学校的优秀经验交流共享。最后，教师可以根据教师的专业特长进行分工，擅长讲解理论知识的教师可以负责理论部分的授课，而擅长实践教学的教师可以负责实验或实践环节的教学；也可以根据教学阶段进行分工，如准备阶段、授课阶段、辅导阶段和评估阶段等。

5.3. 协同互动质量和频率有待提高，需全方位全过程加强协同互动

案例分析发现，主讲制多师协同教学中，部分辅助教师参与多师协同教学的积极性不高，主讲教师与辅助教师之间缺乏及时的互动交流，很难有效协作，教学效果受到较大影响。在对案例的教学设计进行分析时，发现大部分教师未分开分析多个班级的学习者特征，未统筹兼顾各班级学生的学习起点能力和学习风格，此外，部分学生的参与意愿不强，参与质量有待提高，也会对多师协同教学的效果产生影响。教师应在备课时协同调研并分析各班级学习者特征，运用数智技术工具协同设计适应多个班级学生学习风格的教学活动。

全过程多方协同合作是多师协同教学的关键。首先，应建立有效的沟通机制，多师协同教学需要教师之间进行频繁的沟通与协作，如定期召开教学协调会、使用实时通讯工具等，以便教师们及时交流教学情况，解决教学中遇到的问题。其次，应协同分析各班级学习者特征，比如可以通过问卷星或测验平台诊断两地学生学情，并根据学情诊断结果协同设计教学活动。最后，应关注学生的学情反馈，根据学情反馈结果及时调整教学方法和教学策略，提升学生的学习体验和效果。

6. 结语

本研究积极响应教育数字化转型课堂教学高质量发展的需求，以协同学、系统论、教学设计理论为支撑，结合多师协同教学方式的内涵与影响因素，根据教育数字化转型背景下教育部有关文件精神，构建了数字化转型背景下多师协同教学方式分析模型，对来自广东省中小学的多个多师协同教学案例进行系统分析，总结形成了三类典型数字化转型背景下中小学多师协同教学方式，包括主讲制多师协同教学、平行式多师协同教学、校企合作多师协同教学。同时指出其主要特征，即多师协同教学环境完备且走向数智化，多师协同教学资源丰富且呈现形式数智化，多师协同教学活动多样且融入大单元、跨学科等特色，协同教学评价多元化且注重对学生协作能力和问题解决能力的评价。研究还发现数字化转型背景下中小学多师协同教学存在对多师协同教学的理论认识不足、教师团体分工不够明晰、多师协同互动质量和频率有待提高等问题，并基于此提出解决多师协同教学问题的建议，即强化多师协同教学理论认识、明确教师团体分工及协作标准、全方位全过程加强协同互动，以期解决当前教育

数字化转型背景下中小学多师协同教学创新的迫切需要，同时也为新时代我国多师协同教学可持续发展提供有益借鉴。

参考文献

广东省教育厅（2023）。关于组织开展广东省教育“双融双创”教师信息素养提升实践活动部分优秀作品提炼培育及示范辐射工作的通知。检索自 https://srsc.gdedu.gov.cn/srsc/home/noticeDetails?id=1732704252307730519。

谢幼如,邱艺和黄瑜玲等（2020）。疫情防控期间“停课不停学”在线教学方式的特征、问题与创新。电化教育研究，03，20-28。

周慧敏,张远索和董朋（2023）。多师协同教学研究进展与展望——基于 CiteSpace 分析。改革与开放，14，64-72。

谢幼如（2016）。教学设计原理与方法。北京，高等教育出版社，23-24。

李贺（2007）。远程合作学习中的师生关系特点及影响因素分析。中国电化教育，08，54-56。

唐静静,卢娜和朱红瑜（2023）。智慧教育环境下协作学习影响因素及效果提升策略研究。河南教育(高等教育),12，45-47。

彭梓涵和王运武（2019）。基于混合学习模式的小组协作学习行为影响因素研究。黑龙江高教研究,06，141-147。

张靖和郑新（2022）。教学关系视域下的双师课堂：特征、问题与优化策略。电化教育研究，10，19-25。

中华人民共和国教育部（2018）。教育部办公厅关于开展第二批人工智能助推教师队伍建设试点推荐遴选工作的通知。检索自 http://www.moe.gov.cn/srcsite/A10/s7034/202104/t20210423_527853.html。

教育部和国家发展改革委（2023）。财政部关于实施新时代基础教育扩优提质行动计划的意见。中华人民共和国国务院公报，27，44-48。

郭炯,郑晓俊和黄彬（2017）。网络学习空间支持的协同教学模式与应用案例研究——网络学习空间内涵与学校教育发展研究之八。电化教育研究,10，23-29。

梁馨月（2023）。基于“在线互动课堂”的双师型教学模式构建与应用研究（硕士论文）。宁夏大学。

生成式人工智能的个性化学习研究综述及展望——基于 Cite Space 可视化分析

A Review and Outlook for Personalized Learning with AIGC:

A Scientometric Analysis Using Cite Space

刘小钰[1]代毅[*]庞何千慧[2]李可昕[3]
[1*23]澳门城市大学 教育学院
yidai@cityu.edu.mo

摘要：教育信息化的推进，再次推动了人工智能与个性化教育的融合。基于 2008 年至 2023 年 15 年间 388 篇 CSSCI 来源期刊和北大中文核心期刊的文献数据分析，并借助信息科学知识图谱软件 CiteSpace 可视化呈现，探讨人工智能在个性化学习领域的研究趋势。研究发现，目前生成式人工智能在个性化学习方面具有政策导向性并备受师范类院校的关注，且研究内容主要集中在人工智能平台搭建和信息技术应用方面。研究有助于进一步挖掘生成式人工智能在个性化学习领域的深层次研究和探索未来发展方向，为教育领域的决策者、研究人员和师生等提供了有益的参考。

关键词：生成式人工智能；个性化学习；发展趋势分析；CiteSpace

Abstract: *Advancing educational informatization has once again spurred the integration of AI and personalized education. This study examines 388 CSSCI source journals, as well as Peking University's Chinese core journal literature data, spanning a period from 2008 to 2023. The research employs the visualization software CiteSpace for scientometric analysis, aiming to explore the trends in AI research within the context of personalized learning. The results indicate that AIGC is currently characterized by its policy-oriented approach to personalized learning and has attracted the attention of mainstream universities. The primary focus of current research is on the development of AI platforms and the application of information technology. This analysis provides insights into the ongoing research and potential future trajectory of AIGC in personalized learning, offering valuable perspectives for decision-makers, researchers, educators, and students in the field of education.*

Keywords: AIGC; Personalized Learning; Trends; CiteSpace

1. 研究背景

近年来，人工智能、大数据和学习分析技术在教育领域的应用不断深化，取得了许多创新性成果，为探索和解决个性化学习问题提供了新的思路。在中国，教育部门也积极推动教育信息化发展，致力于将人工智能技术与个性化教育相结合。首先，中共中央、国务院于 2019 年印发的《中国教育现代化 2035》强调了“注重全面发展、面向人人、因材施教”的基本理念，提出了实现规模化教育与个性化培养有机结合的目标。自 2021 年教育部联合多个部门发布的《关于推进教育新型基础设施建设，构建高质量教育支撑体系的指导意见》再次强调了人工智能在教育领域的重要性。此外，习近平总书记在中央政治局第五次集体学习时强调，“教育数字化是我国开辟教育发展新赛道和塑造教育发展新优势的重要突破口。”教育部高度关注教育数字化发展，推进国家教育数字化战略行动。而今 2024 年 1 月 30 日，教育部长怀进鹏在今年召开的世界数字教育大会上发表了主旨演讲，强调要深化实施教育数字化战略

行动，推进资源数字化、管理智能化、成长个性化和学习社会化，将大规模个性化教育变为现实。在政策的引领下，人工智能与个性化教育的融合发展成为当前教育研究的重要发展方向。本研究旨在深入探讨我国人工智能实现个性化学习的发展，通过探讨前沿热点发展，以期促进教育领域的持续发展和进步。通过充分利用人工智能技术的优势，推动教育信息化的创新应用，可以更好地满足学生的个性化学习需求，提高教育的质量和效果。

2. 文献综述

2.1. 生成式人工智能的概念界定

生成式人工智能（Artificial Intelligence Generated Content，AIGC））是指一类人工智能系统，能够自主地生成新的内容、图像、音频或其他形式的数据，而不仅仅是对已有数据的分析和处理 (顾小清, 2019)。生成式人工智能系统具备创造性和创新性，能够以某种方式模拟人类的创作过程。生成式人工智能主要基于深度学习和生成模型等技术，包括生成对抗网络（GANs）、变分自编码器（VAEs）等。这些模型通过学习大量的数据，能够生成具有相似特征的新数据，并且能够在生成过程中融入一定程度的创造性和变化性 (顾小清等, 2024)。生成式人工智能在多个领域有广泛的应用，如自然语言处理、计算机视觉、音乐创作、艺术生成等。通过生成式人工智能系统，可以自动生成文本、图像、音乐等内容，为人类创作者提供灵感和辅助工具。

2.2. 研究问题和研究方法

生成式人工智能在个性化学习领域的研究现状是怎样的？通过对2008-2023年的文献进行综述和分析，可以了解生成式人工智能在个性化学习中的应用情况、研究热点和趋势。生成式人工智能在个性化学习中的主要技术和方法有哪些？通过对文献的分析，可以了解生成式人工智能在个性化学习中所采用的具体技术和方法，如生成对抗网络（GANs）、变分自编码器（VAEs）等。生成式人工智能在个性化学习中的优势和挑战是什么？通过对文献的综述，可以分析生成式人工智能在个性化学习中的优势，如创造性和创新性的输出，以及面临的挑战，如数据偏见和可解释性等问题。未来生成式人工智能在个性化学习中的发展趋势和前景如何？通过对文献的分析，可展望生成式人工智能在个性化学习中的未来发展趋势，如更加智能化和个性化的生成模型，以及与其他技术的融合等。通过对以上研究问题的探索，可以全面了解生成式人工智能在个性化学习领域的研究现状、技术方法、应用案例、优势挑战以及未来的发展趋势，为相关研究提供有益的参考和指导。

CiteSpace 是美国德雷赛尔大学教授陈超美及其团队应用Java语言开发的一种文献分析的信息可视化软件，主要通过共引分析理论（co-ciation）和寻径网络算法（path-Finder）等进行分析研究（陈悦&陈超美,2015）。数据处理方面，本研究采用学术界较为广泛使用的 Cite Space 知识图谱分析软件作为研究工具，对文献的统计分析进行可视化呈现。

2.3. 数据来源与研究方法

本研究以中国知网 CNKI 数据库为数据来源，以“人工智能”，“个性化学习”等为关键词进行检索，共检索论文 540 篇（检索时间截至 2024 年 1 月 29 日）；依据文章类型及学科来源等对检索结果进行精炼，共提取论文 388 篇。本研究利用 CiteSpace 对所检索的文献进行文献共被引分析，以了解人工智能背景下个性化学习策略研究领域的前沿；通过关键词共现分析探索人工智能背景下个性化学习策略研究领域的主题和热点，通过关键词共现分析中突现关键词的演变，追踪人工智能背景下个性化学习策略研究的发展趋势等。

3. 生成式人工智能的个性化学习研究总体趋势

3.1. 发文趋势

在之前的研究中，来自华东师范大学信息技术学系的学者李世瑾等 (李世瑾&顾小清, 2023) 已经将我国近十年以来的人工智能个性化发展的研究文献进行了系统梳理和回顾。因此，本研究在该研究基础上，选择时间跨度为 2008 年 1 月 1 日到 2023 年 12 月 31 日，总结人工智能的个性化学习策略”的最新研究进展，洞察未来的研究方向。图 1 是根据本研究检索出的 153 篇科学教育相关文献构建的年际与累计发文量。前面学者将 2017 年后划分为科学教育研究的快速发展时期 (Cai, 2020)。本研究具体细分为三个发展阶段：平稳增长、快速增长和持续增长。作为教育的基础，个性化学习问题一直是教育界、学界研究和关注的重要话题。为了解我国“人工智能的个性化学习策略”研究成果的年产出情况，将文献按照发表年度进行统计整理如图 1 所示：2008 年至 2023 年间，大致可以分为三个阶段：2008 年至 2016 年起步阶段，稳步发展；2016-2019 年迅速发展阶段，在《教育信息化“十三五”发展规划》明确提出，要深化教育信息化改革，推动数字教育资源的集约化开发和应用，推进数字教材的普及和应用 (教育部, 2016)；2019 年至 2023 年平稳上升阶段，总体呈现稳步上升的趋势，快速发展。当前，全球正在进入人工智能赋能的新时代，大大改变了我们的生产生活和学习、科研的方式，推动着社会向智能化、个性化方向发展。在此背景下，教育部发布了《教育信息化 2.0 行动计划》 (教育部, 2018)，要求以人工智能、大数据、物联网等新兴技术为基础，依托各类智能设备及网络，积极开展智慧教育创新研究和示范。

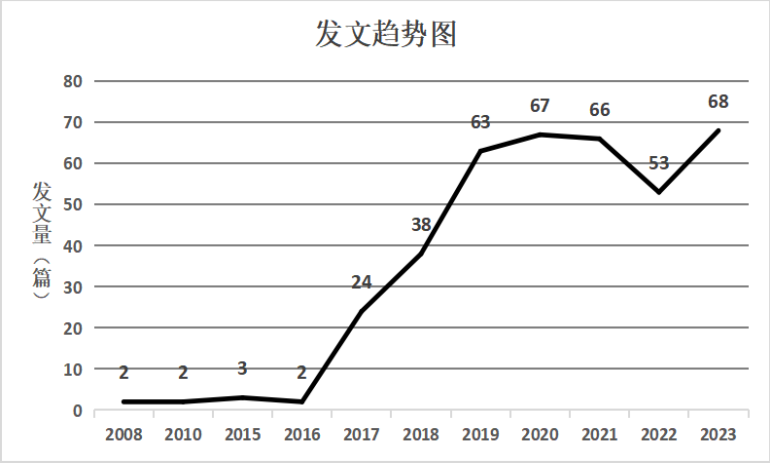


图 1 “人工智能”+“个性化学习”文献发表趋势图

3.2. 核心作者与研究机构

核心作者是指对某一研究领域的发展具有较大贡献的主要研究者，分析这一群体有利于把握该领域的主要研究着和整体研究态势。对核心作者进行分析有助于了解该领域的研究现状及进程。根据图谱中的数据，发现有 195 个节点（作者），71 条连线（合作关系），密度仅为 0.0038，关系较为稀薄。依据普莱斯定理，其计算公式为 M= 0.749（AMAX）（M 是核心作者发表的论文数量；AMAX 是最高产作者的发表论文量），由于已知 AMAX =3，代入公式计算得出 M =2.236。

进一步利用 CiteSpace 软件对作者发文量进行统计发现（表 1 所示），根据普莱斯公式计算该领域的核心作者发文数量。在 388 篇文献中，发文 3 篇以上的作者可视为该研究领域的核心作者。通过计算发现核心作者共有 9 人，共发表了 30 篇学术论文，占相关文献总量的 7.73%。其中，顾小清发文量最多达到 5 篇 ，其次是唐烨伟 4 篇。为了进一步观察作者之间

合作情况，利用 CiteSpace 软件生成作者合作图谱进行可视化分析。合作图谱中的每个源泉节点代表以为作者，节点大小与作者的发文量成正比，节点之间连线代表合作情况。如图 2 所示核心作者分布较为离散，部分作者之间进行了多次合作，开始初步形成多个研究团队。

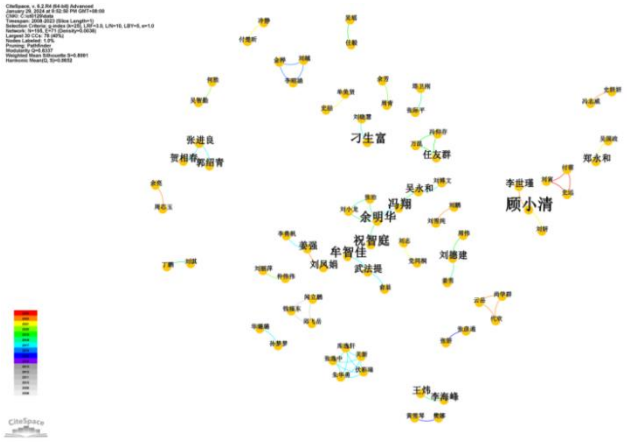


图 2　“人工智能”+“个性化学习”作者合作图谱

表 1 “人工智能”+“个性化学习”发文作者图（节选前十位）

排序	发文数	年份	作者
1	5	2021	顾小清
2	4	2018	唐烨伟
3	3	2017	冯翔
4	3	2018	祝智庭
5	3	2017	刁生富
6	3	2017	余明华
7	3	2019	周东岱
8	3	2017	牟智佳
9	3	2019	于泽元
10	2	2017	武法提

对机构进行定量分析能在一定程度上反映出“人工智能”+“个性化学习”领域的研究布局，运行 CiteSpace，选择 Node Types 为 Institution，绘制出生成式人工智能研究机构可视化图谱（详见图 3、表 2），有 172 个节点，100 条连线，密度 0.0068。同时，利用 Citespace 软件生成主要研究机构的合作图（图 3），“人工智能”+“个性化学习”研究机构主要包括综合类院校，教育研究机构，师范类高等院校与高职类院校。据图谱研究可得出，高校是我国目前研究人工智能个性化学习问题的主力军（表 2），特别是师范类的高校在图谱中出现次数最多。研究机构众多，形成了以华东师范大学教育信息技术学系、北京师范大学教育学部，东北师范大学信息科学与技术学院为代表三个合作较多研究机构组织。不难发现高等师范类院校独占鳌头，高校研究集团正逐渐形成。

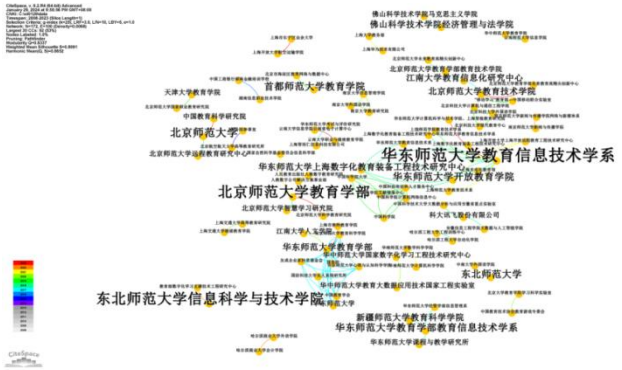


图 3　“人工智能”+“个性化学习”主要研究机构合作图谱

表 2“人工智能”+“个性化学习”机构发文表（节选前十所）

排序	发文量	年份	机构
1	14	2017	华东师范大学教育信息技术学系
2	10	2018	北京师范大学教育学部
3	10	2018	东北师范大学信息科学与技术学院
4	6	2017	西北师范大学教育技术学院
5	5	2018	北京师范大学
6	4	2021	无锡城市职业技术学院
7	4	2020	东北师范大学
8	4	2017	华东师范大学教育学部教育信息技术学系
9	4	2019	首都师范大学教育学院
10	4	2017	华东师范大学开放教育学院

4. 研究热点和趋势分析

4.1. 关键词共现分析

词频分析法是一种文献计量方法，利用关键词或主题词在某一领域文献中出现频次的高低来揭示该领域研究热点和发展动向 (张勤&马费成, 2007)。运用 Citespace 软件，通过关键词共现分析可以帮助研究者追踪热点。关键词用年轮状节点显示，年轮大小表示关键词出现频次高低，频率越高，就越可能代表研究的核心问题。网络节点数量 N=241，连线数 E=273，网络密度 D=0.0094。表示该图共有 241 个节点，273 条连线，关键词之间存在较强的相互关联。结合表 3 和图 4 可知，关键词“人工智能”出现的频次最多达 218 次，反映该主题在学术界有较高的研究热度，其次“大数据”、“智能教育”和“个性化”分别达到 31 次、22 次和 16 次，与当前数字教材研究中的热点相一致，且高频词与其它关键词均有连线且有共线关系，聚集程度较高。其中，“人工智能”“大数据”“学习”等 8 个关键词节点的最外围呈现紫红色光环并有向外扩张的趋势，说明该类关键词在现阶段发展迅猛，符合当前数字教材研究中的热点与焦点（图 4、表 3）。

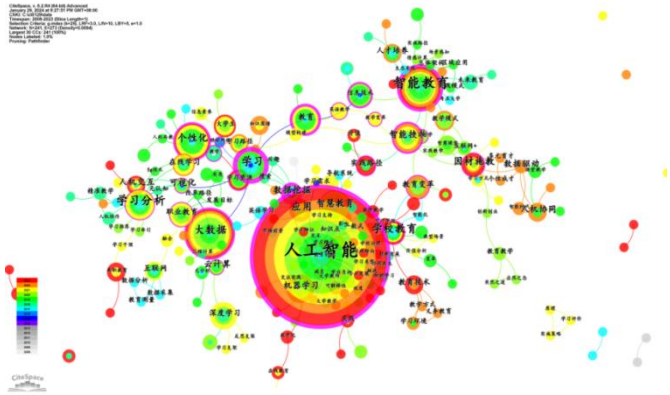


图 4　“人工智能”+“个性化学习”研究关键词共现图谱

表 3“人工智能”+“个性化学习”研究关键词共现表（节选前十位）

排序	频次	中心性	年份	关键词
1	218	0.69	2015	人工智能
2	31	0.27	2015	大数据
3	22	0.5	2017	智能教育
4	16	0.23	2008	个性化
5	12	0.14	2017	智慧教育
6	12	0.9	2008	学习
7	12	0.18	2017	学习分析
8	10	0.07	2017	机器学习
9	10	0.37	2018	智能技术
10	9	0.47	2015	教育

4.2. 关键词聚类图谱分析

通过对关键词共现聚类与突出主题词聚类叠加进行分析，进一步挖掘国际跨学科教育研究的热点与前沿 (陈悦&陈超美, 2015)。聚类分析是指通过对所收集的数据进行分类，将类似的研究对象进行集合分组，以此呈现出某一研究领域内的主题结构，推断出研究热点。模块值 Q 和平均轮廓值 S 是判断一个图谱的网络结构和聚类清晰与否的标准，如果 Q 值大于 0.3，图谱结构显著，如果 S 值大于 0.5，说明聚类较为合理。如表 4 所示，在关键词聚类表中“人工智能“+”个性化学习”共形成 13 个有效聚类标签。除检索关键词人工智能和个性化学习外，聚类标签的节点数量较多的：学校教育，智能教育，学习分析，互联网+，教育教学等；聚类标签轮廓值大部分都在 0.9 以上，说明聚类效果较好。如图 5 所示，其 Q 值为 0,8337，S 值为 0.8991，图谱具有科学性有效性。观察图 5 可知，形成了 11 个聚类，分别为#0 人工智能、#1 学校教育、#2 智能教育、#3 学习等 11 个聚类，根据图谱可以看出，各个聚类之间无重合。

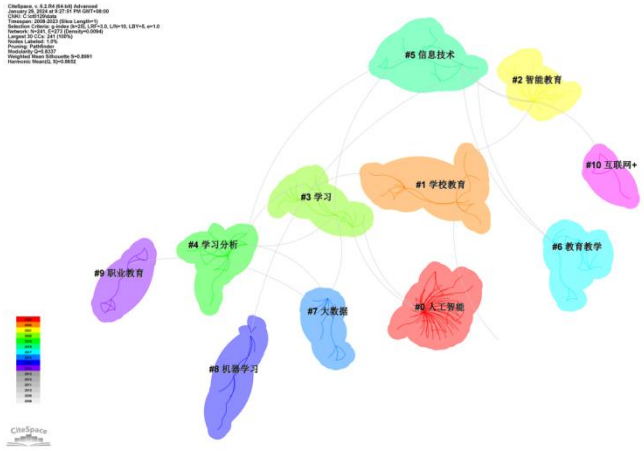


图 5　“人工智能”+“个性化学习”研究关键词聚类图谱

结合关键词聚类知识图谱和共线网络聚类表可以发现，大部分聚类都有重叠之处，同一关键词可能出现在多个聚类中，其中，“人工智能”“学校教育”“智能教育”等标签词出现聚类较强趋势，说明该类关键词符合当前生成式人工智能个性化学习研究中的热点与焦点，也反映出学者们针对某一关键词会从不同角度进行研究。基于此得出主要观点：一是从人工智能时代的人才培养战略出发，探索人工智能技术在引发社会变革的同时，个性化学习将满足学习新需求，以及教育可为人工智能推动的社会发展提前做好知识蓝图绘制；二是从人工智能使能的大规模教育变革、人工智能重塑的知识观，以及人工智能赋能未来个性化学习的发展出发，深度解析了在人工智能创造的技术条件下，解决学习、创造与传递知识和教学等核心问题。把握发展趋势，站在研究前沿，能够更加有效地勾勒教育生态的未来图景。

表 4　“人工智能”+“个性化学习”研究关键词聚类表（节选前十个）

标签	节点数	轮廓值	年份	关键词
0	55	1	2020	人工智能；学习；成人教育；学习分析；机器学习
1	20	0.889	2018	学校教育；教育变革；应用；英语学习；智能化
2	18	1	2019	智能教育；未来教育；人才培养；协同共存；区域应用
3	18	0.931	2015	学习；个性化；搜索；知识；课堂转型
4	18	0.985	2019	学习分析；精准教学；元认知；　演进趋势；因学定教
5	15	0.806	2019	信息技术；　能技术；模型构建；　智慧课堂；教学变革
6	15	1	2020	教育教学；数据驱动；因材施教；　评价指标；多元育才
7	13	0.958	2019	大数据；深度学习；云计算；元分析；人工智能技术
8	11	1	2020	机器学习；智慧教育；教育创新；虚拟仿真技术；医学教育
9	9	1	2018	职业教育；数据分析；数据采集；　教育测量；自适应测评
10	5	0.976	2018	互联网+；实践教学；学习方式；精准就业；网络学习空间

4.3. 关键词突变分析

一组突现的动态概念和潜在的研究问题可以代表某一领域的研究热点 (Chen, 2006)。根据 Cite Space 软件进行相关分析得到突现主题以及对应的突现率（图 6）。可将我国 2008 年至 2023 年“人工智能”+“个性化学习”大致划分为三个时期：2008 年至 2017 年，2017 年至 2021 年以及 2021 年至 2023 年。在 2017 年，学者们主要关注“学习”，“数据挖掘”以及“教学”。在知识生产主体层面，生成式人工智能向知识生产主体靠近，以其自身知识生产的逻辑弥补人类不足。人类不再是唯一的知识生产和拥有者 (李冀红等, 2021)，“自动化智能助手”“机器科学家”“人机共同体”等人工智能重塑的知识生产主体也具有同样的能力。知识生产者不再局限于科学界精英所组成的科学共同体，也不限于为了解决真实问题而吸纳的更广泛的社会参与 (顾小清, 2019)。同时，机器“智能”的参与，延展了知识生产主体的视野，未来知识生产主体将出现人、机器和人机协同三种共生，进而推进个性化学习的深层次发展。

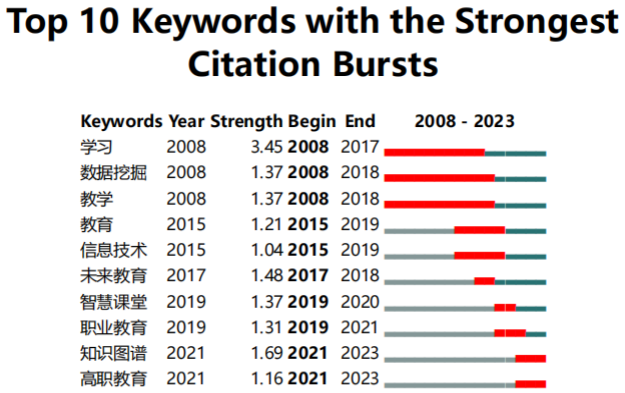


图 6　人工智能”+“个性化学习”研究关键词突变图谱（前十位）

自 2021 年始，“知识图谱”和“高职教育”也慢慢走进视野之中，热度不减。随着人工智能等信息技术与教育教学的深度融合，这些技术不断对教育赋能，增强了教育教学力量，提高了教学环境的智慧，使教育从规模化走向个性化成为现实 (褚宏启等, 2019)。基于生成式人工智能个性化教育依托三类人工智能关键技术的支持：准确构建学习者画像模型的智能技术、建构各类学科“知识图谱”的智能技术和基于个体学习干预的智能技术 (顾小清&李世瑾, 2022)。赋能并构建由人类教师与机器人教师共同提供支持服务的个性化教育模式，为学生定制个性化学习方案，智能推送实时匹配课程资源，全程追踪学生学习行为，数据化立体多维评价学生的学习效果。

5. 研究结论与未来发展建议

综上所述，本文基于 CNKI 数据库中的可视化功能以及信息可视化软件 CiteSpace 对近 15 年来北大中文核心和 CSSCI 收录的与"人工智能个性化学习"相关的研究文献进行了深入分析，得出以下结论。首先，从发文量的趋势来看，生成式人工智能的个性化学习研究呈现出总体增长的态势，经历了平稳增长、快速增长和持续增长的不同阶段。这说明该领域的研究具有一定的政策导向性特点，反映了相关政策对于个性化学习研究的重视程度。其次，研究团体和机构的分布呈现一定的离散性。值得关注的是，在师范类院校中，相关研究团体和机构已经初具规模，这体现了教育领域对个性化学习研究的关注和投入。最后，研究的重点主要集中在人工智能平台搭建、信息技术、数字化等方面。这表明个性化学习研究的核心关注点在于如何应用人工智能技术支持教育教学的改进，提升教育教学的水平和效果。因此，生成式

人工智能的个性化学习研究具有一定的政策导向性特点，并得到师范类院校的关注。此外，研究内容主要集中在人工智能平台搭建和信息技术应用方面。这些发现为进一步推动个性化学习的改进和提升教育教学质量提供了重要的参考依据。

教育事业是促进社会发展的人力资本发动机 (曹培杰, 2020)。伴随人工智能新一轮技术革命的热潮，教育工作者有必要重新思考人工智能时代教育的走向，以及如何充分利用人工智能技术优势，重塑未来教育生态等关键问题。同时，个性化教育是人工智能时代学习者的需求，也是教育在信息 2.0 时代下教学生态转型的新方向。最后，我们期待在智能化教育环境、开放式教育理念中，教学通过顶层设计、智慧系统构建、师资队伍建设等措施，促进个性化教育生态落地，满足每一位学习者的个性化学习需求。

参考文献

张勤, 马费成. (2007). 国外知识管理研究范式——以共词分析为方法. *管理科学学报*, (06), 65-75.

李冀红，庄榕霞，年智英，刘德建&黄荣怀.(2021).面向人机协同的创新能力培养—兼论面向智能时代的创造性人才诉求. *中国电化教育,(07)*, 36-42+61.

李世瑾,顾小清,李睿&王欣苗.(2023).人工智能教育发展指数：内涵、框架与测度.*开放教育研究 29*(04),21-33.

陈悦,陈超美, 刘则渊, 胡志刚&王贤文. (2015). CiteSpace 知识图谱的方法论功能. *科学学研究 33*(02), 242-253.

顾小清(2019). 数字技术带来教育生态变革. *光明日报*,08-06, 13.

顾小清，李世瑾. (2022). 人工智能促进未来教育发展：本质内涵与应然路向. *华东师范大学学报(教育科学版), 40*(09), 1-9.

顾小清,卢琳萌&宛平.(2024).教育数字化转型下的教育研究范式变革.*中国远程教育 44*(02), 36-46.

曹培杰. (2020). 人工智能教育变革的三重境界. *教育研究 41*(02), 143–150.

教育部关于印发《教育信息化“十三五”规划》的通知[EB/OL].(2016-06-07)[2024-02-03]. http://www.moe.gov.cn/srcsite/A16/s3342/201606/t20160622_269367.html

教育部关于印发《教育信息化 2.0 行动计划》的通知[EB/OL]. (2018-04-18)[2024-02-03]. http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html

褚宏启, 瞿志印, 丁新, 张震&王帅国. (2019). 从规模化到个性化：走向技术与教育的深度融合之路. *开放学习研究, 24*(05), 24-28.

Cai, Y. (2020). Progress and trends of Science Education Research in primary schools in China—visual analysis based on CNKI database. *Creative Education Studies, 08*(03), 380–387.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology, 57*(03), 359–377.

数字原住民新手教师在个人学习环境中的专业发展模式：基于扎根理论的一项探索性研究

A Professional Development Model for Novice Digital Native Teachers in Their Personal Learning Environment：An Exploratory Study Based on Grounded Theory

林易[1]，李姗姗[2]，吴筱萌[3*]
[1-3]北京大学教育学院教育技术学专业
*wuxm@pku.edu.cn

【摘要】 数字化时代对劳动者的科学文化素养、教师的能力素养提出了进阶要求。因此，本研究对18位数字原住民新手教师进行个人深度访谈与焦点小组访谈，基于扎根理论构建了数字原住民新手教师在个人学习环境中的专业发展模式。我们以“动因—诱因—行为—产出”来归纳这一专业发展模式，凝练了个人学习环境中的教师后现代专业主义意涵，列举了影响教师基于个人学习环境进行专业发展的效率的六大桎梏，将其与诞生于数字时代的联通主义学习观相比对，提出要以联通视域看待未来教师教育的发展方向。

【关键词】教师专业主义；数字原住民；个人学习环境；扎根理论；联通主义

Abstract: *Based on the background of the digital era and the advanced demand for the scientific and cultural literacy of workers and teachers, this study conducted individual in-depth interviews and focus group with 18 novice digital native teachers. We developed a professional development model for novice digital native teachers in their personal learning environment based on grounded theory, using the "motivation—incentive—behavior—output" framework. The study distilled the implication of teachers' postmodern professionalism in personal learning environment and enumerated six major constraints affecting the efficiency of teachers' professional development based on their individual learning environment. By comparing these constraints with the connectivism, the study suggests incorporating connectivism when considering the future direction of teacher education.*

Keywords: teachers' professionalism, digital native, personal learning environment, Grounded Theory, connectivism

1. 研究问题

在数字化时代背景下，社会对劳动者的科学文化素养提出了进一步的要求。与劳动者的市场准入标准提升相伴随的，是人民群众对新时代教师专业化发展成果的进阶要求。纵观历史，当社会变革发生时，教师的专业化发展方向便会发生转变，继而演变出新的专业主义内涵（Hargreaves，2000），即教师在教育过程中新增的知识、技能与素养要求。

依照Hargreaves对教师专业化进程的划分，我们已迈过了前专业时代、自主专业时代、团体专业时代，正步入更灵活、开放、民主的后现代专业时代（Hargreaves，2000）。在以数字化为标志的第四次工业革命背景下，我们观察到身边的新晋中小学教师正在校内集中管理的专业发展模式以外，通过现代数字工具自主寻找各型各类的教学资源，例如基于“小红书”APP获取与时俱进的教案资源，基于“空中课堂”授课平台梳理教学流程，基于AIGC大语言模型预演课堂中的师生问答等。这批年龄特征隶属于数字原住民的新手教师（Prensky，2001），正通过搭建一个个性化定制的个人学习环境来扩展自身的专业发展路径。因此我们将数字原住民新手教师作为研究对象，通过个人深度访谈法与焦点小组法收集质性资料，基于扎根理论开展了一项探索性研究，旨在回答下述两个问题：

（1）数字原住民新手教师搭建个人学习环境进行专业发展的原因与产出是什么？

（2）在搭建个人学习环境进行专业发展的过程中，涌现了哪些后现代专业主义新要求？

2. 文献述评

个人学习环境是一个工具、资源与人进行组合连接的基于网络力量的凝聚体（Siemens，2008），其构建过程即学习者、学习同伴、学习工具、学习资源之间形成动态链接的过程（张丽霞和王丽川，2014），目的是支持学习者的个人学习需求（尹睿、何丽珍和彭丽丽，2013）。

王利等基于中小学教师可获取信息的五大来源，将其个人学习环境拆解为人际交往、广播媒体、网络媒体、书籍报刊和实践活动五个子环境（王利、董杰和张景生，2014）。其中，书籍报刊这类传统纸媒在主流观点中，并不归属于强调网络力量的个人学习环境，但它们又确实服务于数字移民教师的专业发展过程中。对于数字移民与数字原住民教师而言，他们用于构建个人学习环境的材料与工具存在先天差异。所以，我们重点关注于中小学新手教师中的数字原住民，探究其独特的个人学习环境构建模式，以及这些构建行为的动因与产出。

3. 研究方法和数据来源

多项研究表明，在探究教师教育问题时，一线教师的切身经历、真实想法等质性资料具有很强的研究必要性（申继亮和王凯荣，2000；林崇德、申继亮和辛涛，1996）。所以，本研究基于扎根理论开展质性研究。我们选取了18位中小学新手教师作为研究对象，于2024年1月开展访谈，共开展一对一深度访谈9人次、焦点小组9人次。所有受访者的职称均为二级教师，教龄均在5年以内，16位来自公办学校，2位来自民办学校，学校类型包括乡村学校、重点学校与普通学校。访谈结束后，我们使用NVivo11工具进行开放性编码（一级）、主轴编码（二级）与选择性编码（三级）。资料分析过程采用连续比较法不断凝练、修正理论模式。

4. 范畴提炼和模式建构

4.1. 开放性编码

开放性编码为资料分析的第一阶段，需要对访谈原文逐句进行编码，从研究对象的经验事实中指认现象、发掘概念、提炼范畴（米高·奎因·巴顿，1990）。本研究从资料中抽象出99个初始概念，并将其划归为29个范畴。碍于篇幅仅展示部分，如表1所示。

表 1 开放性编码形成的范畴（部分）

范畴	初始概念
数字工具的教学依赖	基于纸质资料、惯用板书、必用 PPT
协同创生氛围不足	任务分配制、分工大于协作、个人完成、内容整合
基于网络社区的实践经验传递	高效避错、对知识的理解加深、游刃有余的教学心态
教育资源的资产化取向	资产归属的模糊界限、制度约束、领导抽查、禁止分享

4.2. 主轴编码

主轴编码可以理解为范畴的聚类分析，旨在梳理范畴间的逻辑关联，确定范畴间的类属关系，继而发展出主范畴与副范畴。我们共归纳出6个主范畴，分别是专业发展特征的差异、现行专业发展模式的桎梏、个人学习环境的使用优势、个人学习环境中的后现代专业主义、基于个人学习环境的高效专业发展、各方桎梏下的局部资源联通。

4.3. 选择性编码

选择性编码即从主范畴中挖掘核心范畴，围绕核心范畴构建理论模式。我们确定将“数字原住民新手教师在个人学习环境中的专业发展”作为核心范畴后，构建了如图1所示的“动因—诱因—行为—产出”理论模式。梳理的“故事线”所下：数字原住民新手教师与现有专业发展模式间存在差异，导致数字原住民教师在专业发展过程中遭遇资源局限、创生环境缺失

等问题，这构成了他们搭建个人学习环境的根本动因；数字原住民教师拥有较好的数字素养，同时对个人学习环境带来的效率提升有深刻的认识，这种先天使用优势构成了他们自主搭建个人学习环境的行动诱因；在搭建个人学习环境的过程中，又分化出了聚焦于教学资源、数字工具、实践伦理三个主体的行为规范；最终这种搭建行为产出了更为高效的专业发展结果，但这一结果也明显受到了教育系统中各级各类主体的观念束缚、制度制约与数据壁垒限制。

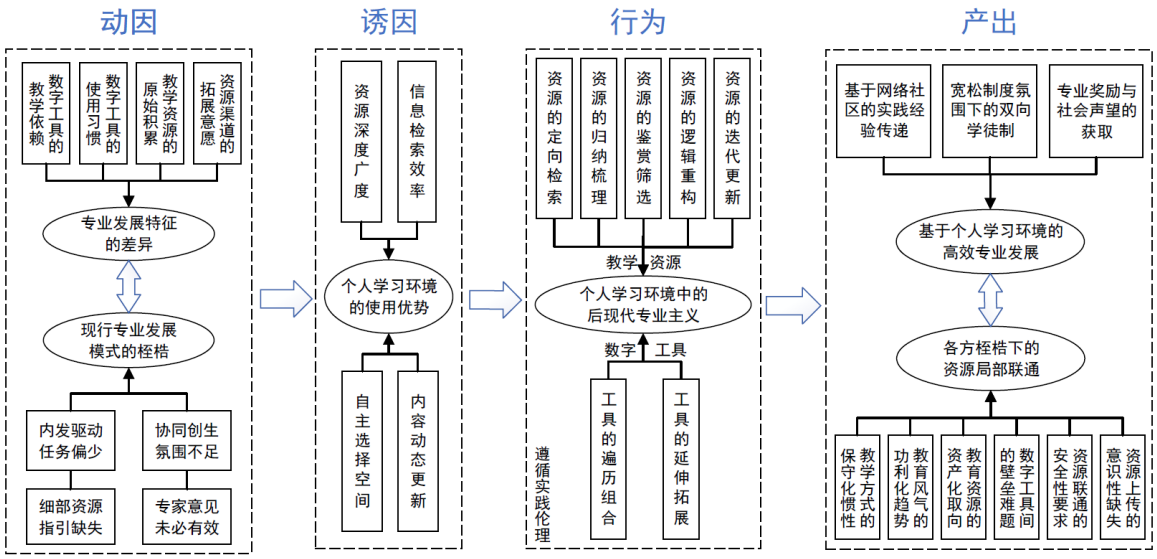


图 1 数字原住民新手教师在个人学习环境中的专业发展模式

5. 数字原住民新手教师在个人学习环境中的专业发展模式的阐述

5.1. 动因：“专业发展特征的差异”与“现行专业发展模式的桎梏”

由于数字原住民新手教师所处的数字社会的时代背景，他们的行为模式中先天就刻入了数字工具的使用惯习，将其视为优先级更高、效率更高的专业发展手段。与此同时，他们所接受的教师职前教育同样默许了教学过程中数字工具对传统媒介的替代性参与，这就加剧了他们对数字工具的教学依赖。此外，数字移民教师往往都已步入职业生命周期的中后段，他们已经实现了大量可用教学资源的原始积累，所以对拓展资源获取渠道的意愿不高。

现行的专业发展模式，例如师徒带教制、集体备课、专家理论讲座等，大都诞生于团体专业时代，呈现出强调集体性的管理特征，与后现代专业时代强调的灵活、民主等特征不尽契合，与数字原住民基于网络解决问题的习惯有所违背（McLoughlin & Lee, 2007）。所以，数字原住民新手教师易于在专业发展初期遭遇认知桎梏。这种专业发展特征与现行专业发展模式的碰撞与失衡成为了他们构建基于网络的、数字化的个人学习环境的关键驱动力。

5.2. 诱因：“个人学习环境的使用优势”

为了应对这种发展失调以满足职业生存的基本需求，数字原住民新手教师将目光转向了那些熟悉的数字工具，例如小红书、B站、知乎等。他们发现，基于这些工具搭建的个人学习环境在获取教学资源时具有极强的使用优势，包括资源的深度广度大、信息检索效率高、自主选择空间广、内容动态更新，能够很好填补现行专业发展模式与教学实践所需之间的沟壑。

5.3. 行为：“个人学习环境中的后现代专业主义”

搭建个人学习环境的过程是做动态加减法的过程，教师会从熟悉的数字工具逐步延展到AIGC大模型、各类资源网站等陌生工具中。这一过程中，后现代专业主义的一部分细部行为表征逐渐明晰。在教学资源的内化方面，包括定向检索、归纳梳理、鉴赏筛选等初级行为，以及依照自身教学逻辑对资源进行重构、周期性地对现有资源进行迭代更新等进阶行为；在数字工具的使用方面，包括基于问题情景遍历寻找最适切数字工具的初级行为，基于数字工

具的特性进行组合求解的进阶行为，以及处于高阶的面向终身学习的行为模式，即依据实际需求不断拓展延伸、淘汰更迭个人学习环境中的数字工具的行为；最后是与资源内化、数字工具使用这两类行为相对应的实践伦理要求。

5.4. 产出：“基于个人学习环境的高效专业发展”与“各方桎梏下的局部资源联通”

个人学习环境的搭建不仅解决了新手教师的生存危机，同时还与传统专业发展模式之间形成了相互补充的并行双轨关系。数字原住民新手教师可以在网络社区等实践共同体中完成实践经验的有效扩容、实现自身教育实践过程中的有效避错，以高效的专业发展进程产出专业奖励与社会声望等良性反馈。在宽松平等的制度氛围下，新手教师甚至能与带教师傅在数字工具范畴上形成逆向的学徒制，这就近似于后现代专业时代所期望达成的具备多向度信息传输能力、具备高自由度与去中心化特征的专业发展模式。

理想情况下，个人学习环境应当是充分联通的，教师可随时改写、淘汰或扩展环境中的资源与工具节点，可随时与任何学习同伴进行交互与创生（田富鹏和田斐予，2014）。但数字原住民新手教师很快就能觉察到强压在个人学习环境搭建行为上的多方压力，它受到教育系统中其他主体在观念、制度、氛围、环境等要素上的桎梏，并不能与外界达成完全公开的联通。具体包括上位者所传递的教育方式的保守化惯性、整体教育风气的功利化趋势、教育资源的资产化取向、数字工具间的壁垒难题、紧张制度氛围下资源联通的安全性要求、教师自身资源上传意识的缺失。这些局限性发展出了教师搭建个人学习环境时的独特行为模式，即公开场合下的缄默性、安全环境下的隐秘联通性。

6. 结语

传统的专业发展模式已经无法完全满足当下数字原住民新手教师的专业发展需求，故而他们在职业生存的根本需要下启用了个人学习环境作为发展副轨，完成了后现代专业主义的发展跃进。在未来，应当推动“面向联通”的教师专业发展模式建设，使得教师更从容得应对当下开放包容、互联互通的世界。在未来，本研究也将继续扩大研究样本，以便于在数字原住民新手教师的专业发展主轨与副轨探究上得到更深入、更可靠的结论。

参考文献

王利、董杰和张景生 (2014)。个人学习环境促进教师学科教学技能发展的实证研究。中国电化教育, (5), 89-96。

尹睿、何丽珍和彭丽丽 (2013)。个人学习环境构建方式与实现技术的对比分析。*中国电化教育*, (9), 10-15。

申继亮和王凯荣 (2000)。论教师的教学能力。*北京师范大学学报: 社会科学版*, (1), 64-71。

田富鹏和田斐予 (2014)。基于联通主义的个人学习环境 (PLE) 研究综述。*电化教育研究*, 35(7), 72-78.

米高·奎因·巴顿（1990）。*质的评鉴与研究*。台北：桂冠图书公司。

张丽霞和王丽川(2014)。论连通主义视域下的个人学习环境构建。*电化教育研究*, 35(12), 63-67。

林崇德、申继亮和辛涛(1996)。教师素质的构成及其培养途径。*中国教育学刊*, (6), 16-22。

Hargreaves, A. (2000). Four ages of professionalism and professional learning. *Teachers and teaching*, 6(2), 151-182.

McLoughlin, C., & Lee, M. J. (2007). Listen and learn: A systematic review of the evidence that podcasting supports learning in higher education. *EdMedia+ Innovate Learning*, 1669-1677.

Prensky, M. (2001). Digital natives, digital immigrants part 2: Do they really think differently?. *On the horizon*, 9(6), 1-6.

Siemens, G. (2008). Systematization of education: Room for PLEs. *Learning Technologies Centre Research Blog*, 2008.

融合技术的高校舞蹈多模态课堂教学行为的分析研究*

Analysis and Research on Multimodal Classroom Teaching Behaviors in Dance at Higher Education Institutions with Fusion Technology

张明[1,3]，王海英[2**]，代毅[3]
1 广东技术师范大学 音乐学院
2 华南师范大学 音乐学院
3 澳门城市大学 教育学院
**mzhang17@foxmail.com

【摘要】舞蹈教育中传统的教与学形态以“口传身授，教为中心”，课堂教学的方法较为单一。信息技术为高校的舞蹈课堂教学开创了新模式，成为舞蹈教育发展的新趋势。针对如何理解技术支持下的高校舞蹈课堂教学模式的有效性这一问题，本文根据舞蹈课堂的特点构建了包含 23 种模态的课堂多模态教学行为分析编码系统，以某高校构建的舞蹈教学视频录像作为语料库，设计了学习分析模型分别对融合技术的、传统舞蹈课堂的教学行为进行了多模态的标注和分析。在滞后序列分析方法和行为特征量化分析的辅助下发现教学模式的行为状态转移规律，和两类模式间存在的同质性和差异性。研究结果发现了“知识传授”“文化理解”“创编实践”“师生互动”等高频模态协同，反映了舞蹈教学过程中的行为规律与特征，从而推动舞蹈课堂中的技术应用进程。
【关键词】舞蹈教学；多模态；行为编码；滞后序列分析；差异性

Abstract: *The traditional form of teaching and learning in dance education, characterized by "oral transmission and physical demonstration with teaching-center," tends to have a relatively singular approach in classroom instruction. Information technology has introduced a new paradigm for dance classroom teaching in higher education, becoming a novel trend in dance education development. To address the question of understanding the effectiveness of the technology-supported dance classroom teaching model in higher education, this paper constructs a Classroom Multimodal Teaching Behavior Analysis Coding System comprising 23 modalities based on the characteristics of dance classes. Using dance education videos recorded at a university as a corpus, a learning analysis model is designed to annotate and analyze the multimodal behaviors of both technology-integrated and traditional dance class teaching. Through lag sequential analysis and quantified analysis of behavioral features, the study reveals the behavioral state transition patterns in teaching models, as well as the homogeneity and differences between the two modes. The research findings provide crucial behavioral evidence for the development of higher education dance teaching models, thereby advancing the application of technology in dance classes.*
Keywords: Dance education, Multimodal; Behavioral coding, Lag sequential analysis, Differences

1. 问题的提出

舞蹈教育中传统的教与学形态以“口传身授，教为中心”，授课的方式往往较为单一。鉴于传统舞蹈课堂有限的教学时空，“示范——模仿——练习”的常规三角循环教学模式，舞蹈教师对多模态意义建构、多元话语能力培养、多模态教学方法的认知不足，“教”与“学”过程中各种要素间“化学反应”式的深度融合不充分（季明明，2019）。高校舞蹈课堂运用多模态教学模式能够搭建完整场景，补足教学环节，促进知识整合，推进多元智慧，克服传统舞蹈教学资源不足、教学条件不具备或实际运行困难等问题（张翔，2020）。然而，在过去若干年的探索中，高校舞蹈教师的多模态模式以及该模式对优化教学的作用尚未得到足够的关注。

事实上，理解舞蹈教师的技术融合课堂教学模式，一方面可以促进教师充分反思教学过程，干预教师对技术的使用，达到更好的智慧育人效果；另一方面，它可以为舞蹈课堂的教学评价提供解释性支撑，并与特定的教学环节和教学模型相关联。因此，伴随着教师 ICT 技术的逐步完善和学习分析技术的发展，本文通过采集教师行为与表现数据，将其转变为结构

*本文受 2022 年教育部人文社会科学重点研究基地重大项目“审美过程与中小学生审美素养培养研究”(22JJD190005)和 2023 年广东省本科高校教学质量与教学改革建设项目“创新型人才视域下高校舞蹈多模态教学模式构建与实践研究”(802)的支持

化知识，为推动高校舞蹈教师专业发展和舞蹈课堂的创新型人才导向的育人目标提供行为层面的证据，从而推动舞蹈课堂中的技术应用进程。

2. 研究基础

2.1. 融合技术的高校舞蹈课堂教学模式

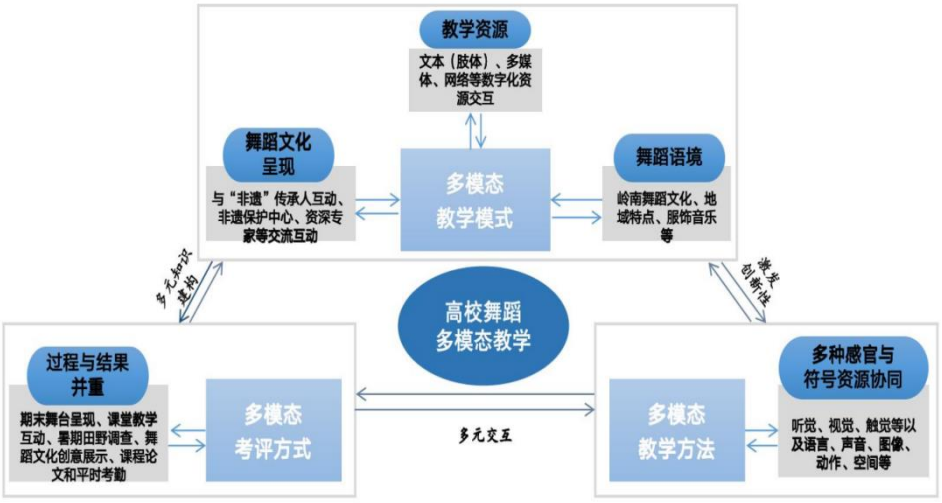
从课堂教学的角度，多模态是增强教学符号系统、增强教学表现力的教学方法，从评价角度多模态是一种对课堂教学行为合理性分析和评判的方法（胡炬 等，2022）。新伦敦小组（New London Group）指出，多模态教学（Pedagogy of Multiliteracies）是通过融合文本和数字化资源，运用不同的模态进行意义构建的一种教学方法（New London Group, 1996），强调身体和大脑通过多模态、多感官协同参与交际的不可分割性。NORRIS S（2017）研究指出模态密度与活动者对传播知识、转移注意、协商分歧等高层次活动的注意程度呈正相关。Ganapathy（2016）的研究表明，多模态教学方法能够促进学生的自主学习和个性化学习，并提高他们的学习动机。

舞蹈专业教学以集体授课为主，教师通过身体传达知识，其教学内容与成果均以身体动态呈现为主（李琼，2022）。针对舞蹈教学的复杂性和特殊性，刘春（2005）从创作、传播、教育、娱乐四个方面提供了舞蹈教学与新媒体技术结合的理论构架。田培培（2020）通过教育空间、教育技术、教师与学生等方面的分析和实际教学案例，展示了在线上舞蹈教学中具体的流程、教学方法以及面临的挑战和解决方案，对于反思传统教学模式，探索舞蹈教学多元化发展具有较好地启示。蒙曦（2023）从舞蹈教学实践与信息技术融合，通过 3D 全息沉浸技术，提出“全景体验”多模态框架下的“口传身授+视听感受”教学模式，将视觉、听觉、触觉舞蹈虚拟空间与舞蹈生态、文化体验、地理人文等方面相互渗透，促进学生对舞蹈文化的全方位感知与理解。张铭（2020）通过舞蹈教学资源、民间舞蹈文化呈现、舞蹈语境以及考核方式的多模态教学模式，以验证舞蹈教学的有效性和适用性。可见，对技术支持的舞蹈多模态教学模式有传统教学无可比拟的优势，在提高高校舞蹈教学效率、培养学习者的多元知识理解、创新能力、优化课堂教学效果和推进高校舞蹈教学改革进程等方面都具有可行性和突破性。

如图 1 所示，本文认为高校舞蹈多模态教学是通过多模态的教学模式（教学资源、 舞蹈文化呈现、舞蹈语境）、多模态的教学方法和多模态的考评方式，在这一过程中利用现代多媒体、立体化教学资源，通过听觉、视觉、触觉等多种感官协同运作，以及语言、声音、图像、动作、空间等多种手段和符号资源，在多元交互与“合力”的舞蹈多模态学习中，培养学生的多元知识建构。

图 1

高校舞蹈多模态教学模式的架构图



2.2. 高校舞蹈教学行为分析的学理认识

从教师教学行为的定义可以得知，教师教学行为会受到教师教育思想、教学理念、教学技能等因素的影响；同时，教师教学行为也反映教师的教学意图。李琼（2022）以高校舞蹈课堂观察为个案研究对象，从教学本质、目标、内容、师生角色、教学方式对舞蹈课堂教学行为关系的因素分析表明，高校舞蹈教师教学观与教学行为呈现差异性，主要体现在目标落实和活动实施两个维度。彭林华等（2021）采用滞后序列分析探索了教师智慧教学行为模式，结果表明不同层级教师具有不同的信息技术应用能力。高校舞蹈教师的多模态教学行为是教学行为的一部分，是教师根据教学目标与教学内容专门设计与安排的活动，受教师教育思想与教学理念等影响。

教师教学行为的分析和改进，可以促进课堂教学效果提升及创新人才培养的素养目标的达成。胡炬等（2022）通过音乐课堂教师教学行为分析的多模态分析，提出课堂中教师以有效的多模态表现为中介强化人际互动，使教师的表达更加多元化的同时，可以使学生更加直观地了解教师所表达的意思，激发学生的学习兴趣，发展认识美、体验美、感受美、欣赏美和创造美的能力。尹睿等（2022）通过多模态数据采集、数据融合、建立学习者模型、强调情景感知实现多模态教学资源的构建，以优化学习者的学习体验，助力学习者更加投入。Kress G（1996）认为除了使用语言资源外，人类学习还须借助其他一种或几种模态符号参与才能达到有效学习的目的。简言之，课堂的多模态话语形式共同作用从而构建学习意义的实现。以上研究，说明技术是否能提高教学效果，还取决于技术是否能与课程合理的融合。

3. 研究设计

3.1. 问题定义

多模态教学成为高校舞蹈教学模式发展的趋势。但由于舞蹈有着丰富的空间动作变化，在具体的舞蹈课程实施过程中，是主要基于肢体动作的舞蹈教学，与传统的课堂教学以语言教育为主的形式有很大的差异，对“在场”的需求更为强烈。但是以肢体动作的行为研究目前非常缺乏，因此，仅参考其他学科的研究成果，不能充分表达舞蹈学科教学的行为意义，也无法证明多模态教学的有效性。

基于以上现象，本文通过选取的八节教学视频课例（传统教学和多模态教学各四节），分析高校舞蹈的多模态教学情况，从而提出改善建议。具体研究问题如下：

（1）高校舞蹈教学中采用的多模态主要有哪些，模态分布特征如何？

（2）在教学过程中不同教学阶段，行为状态转移有何规律？

（3）两类舞蹈教学模式（传统、多模态）间存在何种同质性和差异性？

3.2. 研究对象

本文以 G 大学开设的《岭南舞蹈》中“排瑶长鼓舞”一课为例作为语料来源，该课程为舞蹈学（师范）本科二年级的专业选修课，教学形式为实训与理论相结合。教学对象选择该年级两个平行班级，共 72 人，设定 A 班作为实验班，采用多模态教学方式，B 班作为对照班，采用传统舞蹈教学方式。课程内容主要包括：培养学生掌握岭南舞蹈（排瑶长鼓舞）基础理论知识、基本动作技能；践行“五育并举、美育并行”为主旨的教育指导思想，培养学生的创新精神与实践能力；坚守专业本位，融入课程思政理念，引领学生形成“真、善、美”的正向价值观，了解理论、创演、教学间的关系，解读动态形象背后的传统文化积淀，使学生形成对中华优秀传统文化中大美、善美、和美精神的自觉接受与传承。

3.3. 研究方法及研究工具

研究根据《岭南舞蹈》课程的特点，通过课堂视频分析的形式，结合弗兰德斯互动分析系统 FIAS(Flanders Interaction Analysis System)，将课堂教学中基本的师生交互行为加以分类，将课堂教学中的模态类型进行细化，以及结合张德禄（2009）的多模态话语分析综合框架和模态选择主导因素的划分框架，形成了包含 19 种模态的舞蹈课堂教学行为分析框架，见表 1。其中模态类型总体上可分为“语言”模态和“非语言”模态，“语言”模态包括教师的口头语

言、伴唱语言、文字和伴语言；伴语言包括音响、道具、音调、队形和服饰。“非语言”模态包括身体和非身体两部分，身体包括动作、比拟、面容和肢体；非身体包括 PPT、网络、图片、视频和环境。

表 1

高校舞蹈课堂教学多模态行为分析框架

模态类型 \ 教学阶段				协同及作用	话语意义
语言	纯语言	口头语言		主模态：主导课程进度	概念、人际和话语意义
		伴唱语言		强化关系：突出排瑶长鼓舞文化特点	概念、人际和话语意义
		文字		补充、优化关系：加深对排瑶长鼓舞文化特点理解	概念、人际和话语意义
	伴语言	音响		辅助、强化关系：提供音乐文化背景，吸引注意力	话语意义
		道具		主模态：辅助教学	话语意义
		音调		补充、强化关系：表达情感、提问、强调重难点	人际和话语意义
		队形		补充关系：辅助教学、加强学生对瑶族民族气质的理解	话语意义
		服饰		优化关系：突出、扩充身体语言表达，发现服饰美、培养兴趣，达到审美能力的提高	人际和话语意义
非语言	身体	动作		主模态：由基础认知到直观感受，加深对瑶族广博的文化底蕴的知识建构	概念、人际和话语意义
		比拟		补充、优化关系：从现实生活与民族文化的区别之处，对瑶族文化产生认知，激发创新意识	人际和话语意义
		面容	表情	补充、优化关系：给予鼓励、肯定	人际意义
		肢体	眼神	辅助关系：获取信息、释放信息、控制课堂	人际意义
			手势	补充、优化关系：吸引注意力	人际意义
			姿势	补充、优化关系：吸引注意力	人际意义
	非身体	工具	PPT	辅助、强化关系：感受排瑶精神的薪火相传，使民族气质与自身产生联系	概念、语篇意义
			网络	辅助、强化关系：打破场域、语境交互，认知老艺人舞蹈中自然流露的“生命性”原点	概念、人际和语篇意义
			图片	辅助、强化关系：感受排瑶精神的薪火相传，使民族气质与自身产生联系	概念、语篇意义
			视频	辅助、强化关系：感受排瑶精神的薪火相传，使民族气质与自身产生联系	概念、语篇意义
			环境	补充、优化关系：吸引注意力、提高教学效果	人际、语篇意义

统计工具采用采用多模态语料标注软件 ELAN6.3 进行语料的标注。使用 ELAN 软件依据表 1 行为分析框架，以一节课为单位，每隔 3 秒对教师的语言、眼神、手势、姿势及视频等多媒体工具进行多层级标注，形成如图 3 所示的编码表。以 FIAS 为代表的编码系统并没有对同时发生的 多类教学互动行为进行编码，而是选择一类主要行为进行编码与分析，这种单一维度的编码方式，无法适应舞蹈教学中语言和肢体动作同时发生的情境。因此，本文在编码基础上使用颜色标注，增加了对同时发生的行为进行标注的方式。

表 1

行为编码

行为	编码	行为	编码
讲授	IT	学生分队形创编练习	SD
指令	GD	教师舞蹈动作示范	TDD
提出问题	AQ	学生舞蹈动作练习	SDP
鼓励表扬	PE	舞蹈动作对比	CA
主动应答	TR	学生实践创作展示	SPD
主动提问	TI	微笑，给予肯定	SA
分组讨论	SGD	扫视全班	SC
教师哼唱音乐伴奏	MA	注视个人	LI
教师播放舞蹈音乐	DM	鼓励性手势表示肯定	EG
教师道具训练	TRT	伴随性手势提示学生动作	AG
学生道具训练	SRT	舞蹈视频演示	VP
教师语调	TUI	师生距离	CR

分析工具采用滞后序列分析方法（Lag Sequential Analysis，LSA），该方法主要用于检验发生一种行为后发生另一种行为在统计学上是否存在显著性，进而探索人类的行为模式。LSA 作为学习分析技术之一，逐渐成为了分析学习和教学行为序列的有效方法。滞后序列分析常用的工具有 Enterprise Miner、IntelligentMiner 和 GSEQ。相比于 Enterprise Miner 和 IntelligentMiner，GSEQ 简化了行为分析流程，提升了数据处理效率（杨现民 等，2016），因此本文使用 GSEQ5.1 进行滞后序列分析。应用 GSEQ 分析高校舞蹈教学中行为序列的主要过程，包括：（1）将编码后的课堂教学行为数据以课时为单位，把教师的行为整理为{序列编号，行为序列}的形式；（2）将整理后的数据导入 GSEQ 中，生成行为状态转移频率表和调整后的残差表；（3）根据调整后的残差表筛选出具有显著意义的行为序列，并利用 Gephi 绘制行为状态转移图。

图 2

教学录像量化编码表（截选）

	0秒	3秒	6秒	9秒	12秒	15秒	18秒	21秒	24秒	27秒	30秒	33秒	36秒	39秒
0分钟	GD	IT	IT	IT	IT	IT	IT	IT	IT	IT	IT	IT	IT	IT
1分钟	GD	SDP	SDP	GD	GD	GD	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP
2分钟	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	GD	TUI	SA
3分钟	IT	IT	IT	IT	GD	GD	GD	GD	IT	IT	IT	IT	GD	GD
4分钟	SDP	SDP	SDP	SDP	SDP	GD	SC	SC	GD	GD	GD	GD	GD	GD
5分钟	SDP	SDP	LI	GD	GD	LI	IT	IT	IT	GD	IT	IT	AQ	
6分钟	SRT	SRT	PE	GD	GD	SRT	SRT	SRT	IT	IT	IT	TRT	TRT	
7分钟	GD	GD	GD	TRT	TRT	TRT	SRT	SRT	SRT	SRT	GD	PE	GD	
8分钟	GD	GD	GD	GD	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	
9分钟	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	SDP	

教师提示　　个人距离

表 3

舞蹈教师多模态行为状态转移频率（截选）

	IT	GD	AQ	PE	TR	TI	SGD	MA	DM	TRT	SRT	TUI	SD	TDD	SDP
IT	0	14	1	0	0	0	0	0	0	1	0	1	0	0	0
GD	8	0	3	4	1	0	0	0	0	1	1	1	0	0	10
AQ	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
PE	3	4	1	0	0	0	0	0	0	0	0	0	0	0	2
TR	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0
TI	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
SGD	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1
MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TRT	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0
SRT	1	2	0	1	0	0	0	0	0	0	0	0	0	0	0
TUI	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0
SD	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
TDD	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SDP	0	7	0	4	0	0	0	0	0	0	0	0	0	0	2
CA	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0
SPD	0	2	0	1	0	0	1	0	0	0	0	0	0	0	0
SA	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0
SC	6	3	0	0	0	0	2	0	0	0	0	1	0	0	0
LI	1	2	3	0	1	0	0	0	0	0	0	0	0	1	0

4. 数据分析

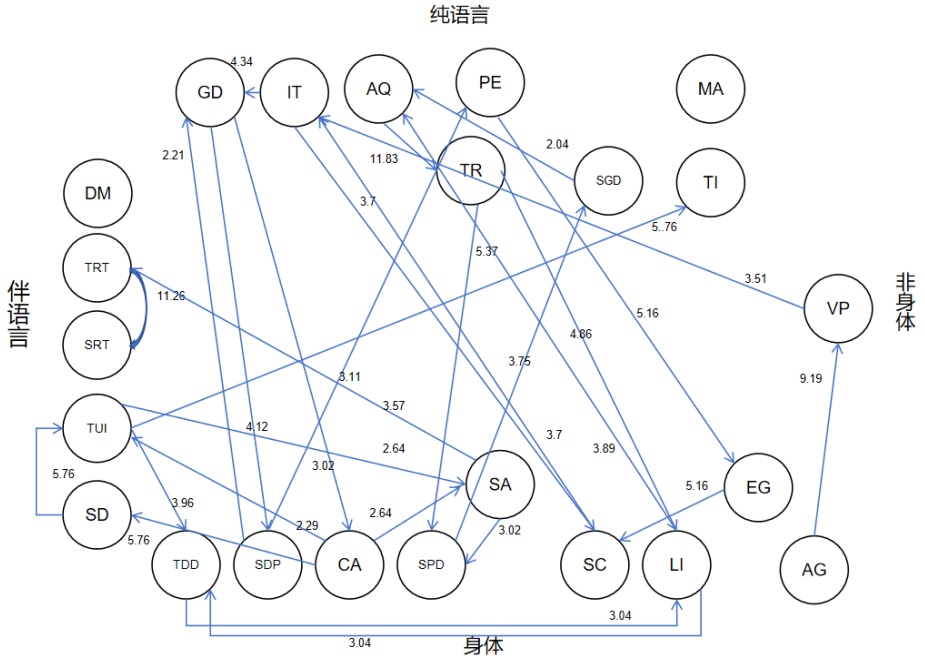
4.1. 滞后序列分析方法

GSEQ 生成的一节舞蹈教师多模态行为状态转移频率（截选），见表 3，表中的行表示起始行为，列表示紧随发生的目标行为，编码表示发生起始行为后又发生目标行为的频率。例如表 2 中第三行第五列对应的数字为 10，表示教师在提出问题（AQ）后学生回答问题（TR）发生的频率为 10。

GSEQ 对频率表中的数据进行转换，得到了对应的残差表，表中的数据是行为序列发生频次调整后的残差值（Z -score）。根据滞后序列处理原则，Z-score>1.96 即表示该行为序列具有显著意义，筛选并整理舞蹈教学行为序列残差表中具有显著性的行为序列，大于 1.96 的有 30 个。以边表格的形式{soure,target,weight}绘制舞蹈教师的行为状态转移图，如图 3 所示。图中各节点表示舞蹈教学的多模态行为，边连接连续发生的两个行为，箭头表示行为转换的方向。节点上的数字越大，表示该行为出现的频率越高，即发生上一节点所代表的行为后发生下一节点所代表的行为的可能性越大。

图 3

舞蹈多模态教学行为状态转移图



4.2.高频模态协同的启示

4.2.1. 促进知识讲授

“口头语言+动作（移动）+道具+比拟”是一组高频模态组合，帮助学生理解并掌握舞蹈动作。教师主要通过口述舞蹈动作要领、亲身示范舞蹈动作，与规范的动作姿态进行比对，及时纠正动作的规范性。例如教师在讲授瑶族长鼓舞的传统动作时，手持传统竹片，横架长鼓道具，向学生展示排瑶长鼓舞传统动作短句、传统节奏型。教师通过直观展示，随后学生进行舞蹈动作练习，过程中教师进行口头语言提示讲解，并带领学生一起跳传统动作短句。通过教师语言模态、身体动作示范、比拟以及眼神手势等多种模态交叉教学，使学生能较好地掌握并灵活运用广东排瑶长鼓舞手持竹片击鼓的传统动作，感受“鼓舞一体”的动作特点，多种模态的交互体现了概念、人际和语篇意义。

4.2.2. 加强文化理解

“口头语言+队形+视频+比拟”是一组高频模态组合，基于声音、视频与队形演示和对比，探讨排瑶长鼓舞的民间传统形态、衍生形态与舞台表演性作品形态。使学生由基础认知到直观感受，从现实生活与民族文化的区别之处，感受排瑶精神的薪火相传，使触不可及的民族气质与自身产生联系。例如在情景教学练习中，教师播放排瑶长鼓舞的民间传统形态、衍生形态、与舞台表演性作品视频，组织学生探讨，通过对地域特点、民俗文化的探讨，引出长鼓舞沉稳朴实、粗旷洒脱的风格特点；接着带领学生围圈队形集中，仿佛置身于瑶族民间传统节日“耍歌堂”，回顾所整理的瑶族文化，从而调动学生多种感官的协同参与；采用学生集体讨论的方式，对梳理出的瑶族文化线索进行分析补充。

4.2.3. 培养创新实践

“队形+动作（移动）+比拟”是一组高频模态组合，培养学生在动作的对比与小组合作创编中激发创新意识。例如教师通过概括传统排瑶长鼓舞动作要领，以及“鼓舞一体”的动作特点，接着引导学生将排瑶长鼓舞沉稳朴实、粗旷洒脱的风格特点，以小组为单位进行动作元素的提炼发展与创编。学生通过小组合作创编练习的形式，组内与组间互相探讨、评价，深化对于排瑶文化的理解与创新转化。发挥学生学习的主动性同时，充分实现学生学习的主体地位，培养学生善于思考、发现、提出问题和创新思维能力。

4.2.4. 深化师生互动

“口头语言+音调+表情+肢体”是一组高频模态组合，主要通过模态协同吸引学生注意力，使学生深入参与课堂互动，提高教学效果。教师通过提出问题，对易混淆的横摆鼓、提压和绕圆动律中的动作与学生进行开放式交流问答，并总结排瑶长鼓舞价值体现。其次，鼓励与引导学生打破传统课堂空间，走进民间田野、传统节日现场、传承人屋舍等，建构非常态化“教学场域”的对话平台，对岭南传统舞蹈的动作元素，表演套路、构图调度、服饰道具等原生态素材进行收集和整理。认知老艺人舞蹈中那种自然流露的“生命性”原点，使获得语境知识更加具体。在“采于民间”的基础上，以民族精神、内涵为基础，引导学生对舞蹈形态中所具有的典型风格、文化内涵展开纵深性学习探索。感受中国传统文化所孕育的神韵与气质，形成对中国传统文化的热爱与自信，同时在互动中培养学生的多元知识理解。

4.3.行为比例统计

表 4

传统与多模态课堂行为比率统计表

模态类型 \ 比率					传统课堂		多模态课堂	
					标注次数	时间占比	标注次数	时间占比
语言	纯语言	口头语言	教师	讲授	72	11.25%	47	10.20%
				指令	91	14.22%	68	14.75%
				提出问题	6	0.94%	8	1.74%
				鼓励表扬	14	2.19%	3	0.65%
			学生	主动应答	23	3.59%	4	0.87%
				主动提问	2	0.31%	0	0.00%
				分组讨论	18	2.81%	0	0.00%
		伴唱语言	教师哼唱音乐伴奏		0	0.00%	0	0.00%
	伴语言	音响	教师播放舞蹈音乐		0	0.00%	28	6.07%
		道具	教师道具训练		9	1.41%	2	0.43%
			学生道具训练		14	2.19%	0	0.00%
		音调	教师语调		7	1.09%	17	3.69%
		队形	学生分队形创编练习		18	2.81%	0	0.00%
非语言	身体	动作（移动）	教师舞蹈动作示范		6	0.94%	154	33.41%
			学生舞蹈动作练习		107	16.72%	105	22.78%
		比拟	舞蹈动作对比		19	2.97%	7	1.52%
			学生实践创作展示		18	2.81%	0	0.00%
		面容	表情	微笑，给予肯定	5	0.78%	1	0.22%
		肢体	眼神	扫视全班	14	2.19%	14	3.04%
				注视个人	8	1.25%	3	0.65%
			手势	鼓励性手势表示肯定	3	0.47%	0	0.00%
				伴随性手势提示学生动作	1	0.16%	0	0.00%
	非身体	视频	舞蹈视频演示		185	28.91%	0	0.00%

本文对两节对比课（传统课堂与多模态课堂）的教学视频进行编码分析，从表 3 中可以了解两类课堂环境下教学的基本结构。从“口头语言”角度分析，讲授和指令行为都是高频发生的行为，但是多模态课堂鼓励表扬（2.19%）、主动应答（3.59%）和分组讨论（2.81%）均占据更大的比重，这说明多模态课堂的更加面向学生的高阶思维培养，体现以学生为中心的理念。从“身体”角度分析，传统教学在教师舞蹈动作示范（33.41%）、学生舞蹈动作练习（22.78%）占据了课堂一半以上的时间，多模态课堂具有更多的舞蹈动作对比（2.97%）、学生实践创作展示（2.81%），更多地鼓励学生的创新思维，已经关注技术在教学中的深层次应用。从“非身体”角度分析，以舞蹈视频演示为代表的多媒体演示功能占据 28.91%，而传统课堂一般不具备该环节。在教学前段多模态课堂一般使用视频演示使学生理解舞蹈文化，在后段视频演示主要是帮助学生了解舞蹈动作的舞台呈现。

5. 结论

针对如何理解技术支持下的高校舞蹈课堂教学模式的有效性这一问题，高校舞蹈课堂教学的行为数据反映了舞蹈教学过程中的行为规律与特征。本文基于行为序列分析视角，通过滞后序列分析法和行为特征量化分析，探索了高校舞蹈课堂多模态行为的特征与规律。主要

探讨了高校舞蹈教学中有何多模态特征；比较了在教学过程中不同教学阶段，行为状态转移规律；传统模式与多模态模式间存在的同质性和差异性。研究表明，高校舞蹈多模态课堂教学包括口头语言的纯语言模态和身体的非语言模态占课堂的主要成分，模态运用合理，特别是比拟、音调、队形和视频课件的模态协同达到了舞蹈多模态课堂教学的预期目标，但教师舞蹈动作示范时学生主动提问的口头语言模态频率较少，以教师提出问题、指令与鼓励模态为中介时增强人际互动。研究结果为进一步分析教师教学的行为模式提供了新思路，能够为高校舞蹈教学模式的发展提供行为层面的关键证据，从而推动舞蹈课中的技术应用进程。需要提及是，因目前国内高校舞蹈课堂教学视频录像的数量尚不充足，未来还需研究者在教学实践中进一步丰富，以期为更多的研究者提供研究样本，得到实践性更为充分的证据。

参考文献

Ganapathy, M. (2016). Qualitative Data Analysis: Making it easy for nurse researcher. *International Journal of Nursing Education, 8*(2), 106. https://doi.org/10.5958/0974-9357.2016.00057.x

Laba, N. (2022). Book review: Reading images: The grammar of visual design. *Visual Communication*, 147035722110696. https://doi.org/10.1177/14703572211069606

Nelson, L. N. (1966). Teacher leadership: An empirical approach to analyzing teacher behavior in the classroom. *Journal of Teacher Education, 17*(4), 417–425. https://doi.org/10.1177/002248716601700404

Norris, S., & Pirini, J. P. (2017). Communicating knowledge, getting attention, and negotiating disagreement via videoconferencing technology: A multimodal analysis. *Journal of Organizational Knowledge Communication, 3*(1), 23. https://doi.org/10.7146/jookc.v3i1.23876

The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*(1), 60–93. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

胡炬、卢甜和赵健(2022)。师范生实习支教云端音乐课堂多模态教学行为分析。**电化教育研究**，**09**，114-128。https://doi:10.13811/j.cnki.eer.2022.09.014.

季明明(2019)。一流大学要引领终身学习和继续教育发展。**开放学习研究**，**01**，1-7。https://doi:10.19605/j.cnki.kfxxyj.2019.01.01.

刘春(2005)。舞蹈新媒体初探。**北京舞蹈学院学报**，**04**，90-97。https://doi:CNKI:SUN:BJWD.0.2005-04-021.

李琼(2022)。高校舞蹈专业教师教学观与教学行为的关系初探。**北京舞蹈学院学报**，**03**，112-120。https://doi:CNKI:SUN:BJWD.0.2022-03-017.

蒙曦(2023)。3D 全息沉浸式舞蹈教学实践探究。**大数据时代**，**06**，28-38。https://doi:CNKI:SUN:DSJD.0.2023-06-003.

彭林华、林晓凡(2021)。基于滞后序列分析的智慧课堂教学行为研究。**现代教育技术**，**07**，55-61。https://doi:CNKI:SUN:XJJS.0.2021-07-008.

田培培(2020)。“停课不停学”的挑战与机遇——基于“互联网+”舞蹈实践课程的案例研究。**舞蹈**，**03**，13-17。https://doi:CNKI:SUN:WDWD.0.2020-03-003.

尹睿、何淑茵(2022)。基于系统性文献综述的多模态学习分析研究进展与前瞻。**现代远程教育研究**，**06**，54-63。https://doi:CNKI:SUN:XDYC.0.2022-06-006.

杨现民，王怀波，& 李冀红. (2016). 滞后序列分析法在学习行为分析中的应用。**中国电化教育**，**02**， 17-23.

张德禄(2009)。多模态话语分析综合理论框架探索。**中国外语**，**01**，24-30。https://doi:CNKI:SUN:ZGWE.0.2009-01-010.

张铭(2020)。新疆艺术学院《中国民间舞蹈文化》课堂教学与考核方法改革与实践。**新疆艺术学院学报**，**01**，43-49。https://doi:CNKI:SUN:XJYA.0.2020-01-007.

张翔(2020)。**当代高等舞蹈教育改革及人才培养研究**。中国书籍出版社。

从海德格尔寻视理论看生成式人工智能对高等教育变革的启示

——基于 ChatGPT 的实证研究

The Insights of Generative Artificial Intelligence on Higher Education Reform from the Perspective of Heidegger's Umsicht Theory: An Empirical Research Based on ChatGPT

曾治* 张誉月 卢晓东
北京大学教育学院，北京 100171
*2201213181@pku.edu.cn，persistbetter@stu.pku.edu.cn，luxiaodong@pku.edu.cn

【摘要】大语言模型的技术原理和海德格尔的寻视理论，预示了人工智能与人类本质之间的差异。对大语言模型和人类操作者进行的实证研究发现，二者围绕实验操作的回答具有显著区别：大语言模型基于语言世界，其回答具有全面性、流程化、客观性与限定性的工具性优势；而人类基于寻视世界，他们的回答展现出针对性、碎片化、感官性与延展性的具身性特点。因此，生成式人工智能为高等教育变革既带来了挑战，也带来了机遇。一方面，未来的教育者应当充分重视高等教育的人本底色，强化学生的劳动教育与实践教育，完善他们的生命体验与人格素养；另一方面，也应当充分发挥信息技术的优势，助力学生应用人工智能拓展寻视世界。

【关键词】生成式人工智能；海德格尔；寻视理论；高等教育变革；大语言模型

Abstract: *The technical mechanism of the large language model (LLM) and Heidegger's Umsicht theory implies a difference between artificial intelligence (AI) and human nature. The empirical research on LLM and human operators finds that their answers to questions about experimental operations are significantly different: LLM is rooted in the language world and its answer is comprehensive, process-oriented, objective and limited, with instrumental strength; while human operators are based on the Umsicht world and their answer is targeted, fragmented, sense-based and extensive, demonstrating their embodied characteristics. Therefore, generative AI brings not only challenges but also opportunities for education reform. On the one hand, educators should focus more on the humanistic elements in education and enhance labor education and practice education to enrich their experience and personality. On the other, educators in the future should make better use of the technological instruments and empower students to extend their Umsicht world with AI.*

Keywords: generative artificial intelligence, Heidegger, Umsicht theory, higher education reform, large language model (LLM)

1. 引言

ChatGPT 凭借其流畅的对话能力、出色的逻辑能力，以及在问题解决与文字写作方面的优秀表现，为知识的生产、传授、应用带来了巨大的挑战与机遇。尽管 ChatGPT 容易带来学习者主体异化、学术不端、技术依赖、信息错误等问题（吴青和刘毓文，2023；张峰和陈玮，2023），但也不可否认其在辅助教案优化、个性化学习和评估、打破信息茧房、创新学习模式等领域的潜在价值（荀渊，2023）。因此，生成式人工智能不仅对高等教育领域造成了巨大的冲击，更为高等教育发展带来了前所未有的创新契机。如何顺应生成式人工智能的时代潮流开展教育变革，成为了国内外学者普遍关注的问题。

国内学者对 ChatGPT 的研究大多停留在教育的技术层面，关注人工智能如何融入教学过程之中。但也有部分学者开始思考 ChatGPT 对整个高等教育发展方向的影响，尤其是对高校

学生培养的影响。荀渊（2023）提醒高等教育研究者应当关注人工智能所不能及的领域，如道德、情感、文化领域等。赵丽红等（2023）则认为，应当重视培养学生的高阶思维能力，从而更加有效地超越人工智能。钟秉林等（2023）认为，尽管教师应当帮助学生应用人工智能更好地获取信息，但更应该关怀学生的人格发展，肩负起“引路人”的角色。相关研究大多停留在纯粹理论的层面，很少有针对 ChatGPT 的实证研究（王佑镁等，2023）。

早在 ChatGPT 问世之前，李丽萍与卢晓东（2023）便已通过对实证访谈材料的分析，引入海德格尔寻视理论，对传统计算主义的离身认知理论取向进行反思，证明了实验教育中“人”的不可替代性。但是，李与卢没能实现人与计算机之间的对比分析，而 ChatGPT 作为生成式人工智能所具有的交互性，为理解并呈现人与计算机之间的差异提供了实证研究的可能。

因此，本文将以实验活动为主题，分别对 ChatGPT 与人类实验操作者进行提问访谈，了解 ChatGPT 在面对需要身体参与的活动之时，会有怎样异于人类学习者的表现。此外，本文还将借助海德格尔的寻视理论，对该差异进行深入的思考与讨论，揭示 ChatGPT 的离身认知本质及其局限。在此基础之上，本文将对 ChatGPT 时代的高等教育发展方向进行再思考，以期为高等教育变革提供理论与实证参考。

2. 从技术原理看 ChatGPT 的局限性

2.1. ChatGPT 的两大核心技术

早在上世纪 60 年代，麻省理工学院就运用自然语言处理技术发明了第一个能够按照固定模式聊天的机器人 ELIZA，用于模拟心理医生的对话（Weizenbaum，1983）。经过数年大规模的语料训练和不断的参数优化，上游产业的数据查询、转换、标注、管理等服务产业发展渐渐壮大。整条产业链的全面建设完善，产生了从量变到质变的效果。ChatGPT 能够作为最前沿、最出色的自然语言处理模型脱颖而出，主要依赖本世纪两大突破性的技术：Transformer 模型和基于人类反馈的强化学习。

2.1.1. Transformer 和注意力机制 早期的文本生成人工智能通常使用基于概率的判别模型，如 N-gram 统计语言模型（Brown et al.，1992）。随着强化学习技术的发展，人们提出了通过神经网络学习来训练语言模型（Bengio et al., 2000），但只能串行处理问题而不能并行处理，常常“忘记”前面输入的话。直到 2017 年，谷歌提出了全新的 Transformer 模型，引入了注意力机制（Vaswani et al., 2017），通过计算词向量之间的权重系数，将句子中每个词与其他词汇并行计算，解决了深度神经网络“记性差”的问题，成功突破了时序屏障。由此， ChatGPT 开始能够回答假设性问题、主动承认错误、在受到质疑时优化回答等。

2.1.2. 基于人类反馈的强化学习 由于神经网络只能根据概率进行语言生成，其中可能会产生错误的知识，因此需要通过人工随机抽取问题并给出准确答案来进行人工标注。但是，人工标注只能在一定程度上减少语法和常识错误，仍然不能够保证所生成的内容符合人类的认知倾向。于是， DeepMind 和 OpenAI 共同提出了“基于人类反馈的强化学习”（Christiano et al,. 2017），将人工标注纳入模型的反馈体系，使其能够根据人类反馈进行自动调整，从而保证模型的输出结果与人类的认知、态度、情感保持一致。

2.2. ChatGPT 技术的局限性

ChatGPT 的底层原理为输出“概率出现次数最多”的文本，并不具备识别情境的能力。例如，安排时间规划表时，它只能简单地对事件加以排列，而不能根据执行者的实际情况给予有效的帮助。因此，ChatGPT 仍然需要依靠大量的人工处理来保障其回答质量。此外，在算力不足、“喂食”不够等客观条件等限制下，生成式人工智能还远未到“通用人工智能（AGI）”或“强人工智能”的程度。

ChatGPT 更本质的局限性，则是它与人类的认知思维上的鸿沟。生成式人工智能依靠足够多的节点和数据、足够精细的算法、足够强大的算力支撑起来的逻辑能力和认知模式，与人类的理性并不相同。如“中文房间”理论所揭示的，要想真正学通一门语言，除了学习语言本身的规则以外，还要学习语言之外更加广阔的“人类世界”（Searle et al., 1982）。

3. 海德格尔的寻视理论与大语言模型的局限性

在海德格尔的哲学理论体系中，寻视世界就是通过操劳上手所揭示的用具世界，其核心在于上手使用。海德格尔将“寻视”定义为一种“顺应于事的视”（海德格尔，2014），即通过上手去“用”某种用具从而“看”到它的过程。因此，海德格尔寻视理论中的“看”并不是简单地用眼睛“看”，而是用整套综合的感官系统去把握、体悟和理解，本质上是一种具身认知（陈巍，2011）。例如，当仅仅用眼看到锤子时，并没有真正地认知到这个锤子，只有当用锤子来锤的时候，才通过寻视“看”到了这个锤子。

与寻视世界相对，海德格尔认为，语言世界只是由抽象的命题所组成的世界，只能规定并揭示“现成状态”的事物。在海德格尔的寻视理论当中，现成状态与上手状态相对，指的是在没有操劳上手的前提下对某件事物获得的模糊理解，往往表现为通过单一视觉或听觉获得的片面认知，大致对应日常语境下的形象和概念，是一种抽象的、无实际意义的存在状态（郁振华，2020）。例如，在真正使用锤子之前，即使充分掌握锤子的材质、构造、用途与操作方式，只要没有真正上手使用，对锤子的所有认知都只能停留在现成状态层面。

因此，寻视世界与语言世界最大的区别在于，寻视不仅要求大脑的参与，更要求眼、耳、手的介入；而语言世界则只是抽象思维的音节化或文字化的呈现，可以脱离各种感官。也就是说，寻视世界是由具身认知所展开的丰富世界，其中每一次上手操劳所带来的视觉、听觉、触觉等感官反馈都是必不可少的，这是一个由海量信息组成的经验世界。而语言世界则是由离身认知所展开的抽象世界，其中的音节或文字在脱离感官经验的条件下都是没有意义的；好比对于婴儿而言，任何的音节都与噪声无异，任何的文字都与涂画相同。

通过寻视或具身认知所获得的知识，可以对应波兰尼所提出的“默会知识”（tacit knowledge）；而语言世界当中的知识，则对应“明言知识”（郁振华，2006）。默会知识是在实践中获得的知识，蕴含了实践过程中无穷无尽的细节，而对这些细节的认知往往是无法用语言来表达的，但它们却是知识最完整的形态（波兰尼，2017）。与此相对，明言知识则是以语言或其他符号为载体所呈现出的准则、指导、关系等，只是知识的很小一部分（Polanyi, 2013）。例如，即使掌握了骑自行车的所有书面技巧，如果从来没有骑过，那依然是不会骑。因此，所有的知识从根本上来说都是默会知识，这种真正的知识是不可能被语言穷尽的。

但是，波兰尼并没有将默会知识与明言知识完全割裂，而是从获取知识的角度对二者的关系进行了更加细致的说明。明言知识并非毫无意义，它能够帮助学习者限定一条获得默会知识的道路、确定默会的先后次序，从而尽可能减少不必要的试错成本（郁振华，2001）。例如，拿到一份汽车驾驶手册时，如果能够记住手册上的内容，那真正开始学车的时候，就能够将现实经验的每个步骤与手册的指导建立起关联，从而提高默会认知的效率。

由此可见，无论是海德格尔的寻视理论，还是波兰尼的默会认知理论，均预示出 ChatGPT 作为大语言模型的根本局限：无法通过上手操劳进入寻视世界，无法通过具身认知获得默会知识。

4. 研究设计与研究结论

为了进一步说明大语言模型所处的语言世界与人类所处的寻视世界之间的关系，更加清晰地呈现 ChatGPT 的局限性，本研究以 ChatGPT 为访谈对象，创新性地展开了一场实证研究。

4.1. 研究设计

本研究选取化学与工程学这两门一级学科，围绕各自的基础实验操作，对 ChatGPT3.5 和相关学科毕业生进行半结构化访谈，并在受访者知情同意的前提下进行记录。受访者共 2 名，均为该学科的本科毕业生，具有丰富的实验操作经历。本研究的访谈人员几乎从未进行过相关的实验操作，在访谈过程当中会以入门学习者的身份对实验过程中涉及到的某些术语、细节进行适时且充分的追问。为了便于理解，访谈围绕两门学科最基础的实验操作——“液体转移操作”和“电路焊接操作”展开，并分别设计了一组访谈提纲。此外，为了保证 ChatGPT 与人类受访者之间的可比性，访谈全部通过线上文字的方式进行。校对主要针对标点符号与段落格式进行修改，对内容本身没有实质性的影响。受访者信息见表 1。

表 1 受访者信息

编号	性别	年龄	就读院校	入学年级	学科专业
Chemi-Human01	男	23 岁	浙江大学	2017 级（已毕业）	化学与高分子材料
Engi-Human01	女	21 岁	西安交通大学	2018 级（已毕业）	电气工程与自动化

4.2. 研究结果

4.2.1. ChatGPT 的全面性、流程化与人类的针对性、碎片化 首先，ChatGPT 的回答相当全面，覆盖了实验前、中、后整个过程，还提供了安全注意事项；而人类受访者的回答往往针对某一重难点步骤，不会拓展到实验操作的整个流程。例如，在回答“如何将烧杯当中的液体倒入试管”时，ChatGPT 不仅描述了液体转移操作的详细步骤，还增加了“准备工作”和“清洁操作”两个环节。人类受访者（Chemi-Human01）则紧紧围绕问题本身展开说明，没有额外增加任何的环节。

其次，ChatGPT 的回答具有流程化的特点，而人类受访者的回答则比较碎片化，需要进一步的梳理才能够明确前后关系。在所有实验操作的回答中，ChatGPT 都会提供流程化的步骤，总是使用“以下步骤”“以下方法”等词组作为引入，随后分点回答。而人类受访者往往先聚焦某一个方面，再基于这一方面展开补充，没有明确的时间先后关系。例如，在提问如何焊接电路板时，人类受访者直接聚焦“焊接的接口”，说明焊接的最佳标准就是焊接接口的“牢固”与“均匀”，然后再继续围绕“牢固”与“均匀”进行补充。并且，人类受访者（Engi-Human01）倾向于使用“然后”“还有就是”等连接词，具有自我补充的特点。

全面性和流程化是语言世界的一大优势，也是学习过程当中的效率保障。语言世界有可能让人陷入空洞的概念漩涡与理论迷雾，但语言作为高效的信息传播载体，也具有着不可替代的作用。ChatGPT 的回答充分展现了语言的这一价值，即能够提供全面的操作说明，为实验者提供寻视的方向。这意味着，只要能够对 ChatGPT 的语料库进行严格的筛选、把关，保证其输出的权威性与可信度，便能充分发挥其指导寻视方向的功能，大大提高学习效率。

4.2.2. ChatGPT 的客观性与人类的感官性 ChatGPT 所提供的标准具有较强的客观性，其精确度有时甚至超出常识范围；而人类受访者则倾向于将标准与感官反馈相关联，尤其是借助视觉经验来呈现规范的操作。例如，在回答“如何判断流速快慢”时，ChatGPT 提出“可以用计时器或秒表等工具，记录液体从烧杯到试管中所需要的时间，并根据时间来判断流速的快慢”。而人类受访者（Chemi-Human01）则直接使用了“层流”与“湍流”的对比来展现流速的不同：“液体在一定的流速内呈现层状，流速更大的时候，层流出现表面的波动就叫湍流”。

人类感官性与 ChatGPT 客观性的区分，体现的正是寻视世界与语言世界的差别，即上手与否的差别。例如，在说明焊接电路的基本操作时，人类受访者（Engi-Human01）直接就从双手开始：“一手拿着焊丝，另一手拿着一个已经加热好的高温的枪”；而在描述微调操作

的时候，其回答也展现出较强的上手动作：“把焊锡往那边稍微带一下，就是围绕着那个圆点转一下”。而 ChatGPT 的回答则基本围绕实验操作的相关部件展开，如焊笔温度、焊锡丝长度、焊点牢固性等，基本感受不到手部动作的存在。

语言世界的局限在 ChatGPT 回答的客观性当中得到了明显的展现。例如，当问到焊台的温度怎样才算合适时，ChatGPT 直接给出了一串精确的温度：“大部分电子焊锡的熔点大约在 183°C（361°F）”。的确，在语言世界当中，ChatGPT 的回答作为标准答案无懈可击，但对于需要上手实验操作的新手而言，如此客观、精确的数字毫无助益，反倒是人类受访者的感官经验更具参考价值。

4.2.3. ChatGPT 的限定性与人类的延展性 ChatGPT 的回答能够始终切题，基本不会脱离设定的实验场景，具有较强的限定性。例如，在烧杯液体转移的场景中，烧杯、试管、液体均为强关联的组分，而工作台（用于放置烧杯和试管）、纸巾（用于擦拭烧杯和试管的外壁）、秒表（用于判断流速的快慢）则可视为弱关联的组分，但都不脱离实验场景。

与之相对应，人类受访者倾向于援引日常生活中的例子，用实验外的现象来说明实验内的操作，具有显著的延展性。例如，在说明如何通过微调实现均匀焊接时，人类受访者（Engi-Human01）的场景拓展到了“油漆刷墙”的场景：“当你在加热过程中，你发现那些融化成液体状态的焊锡都跑到了焊点的某一侧，但是另一侧没有的话，这时候你可能就需要把焊锡往那边稍微带一下……让它继续成一个圆的状态，就好像用油漆刷墙一样，看到这一侧的墙上没有油漆了，就把那个刷子往那边带一下就可以了。”

此外，延展的程度实际上体现了对话双方的经验重合程度。例如，在回答液体转移的流速“怎么才叫缓慢”时，受访者（Chemi-Human01）引用了河流的类比：“保证液体只沿着一侧的管壁呈现液流状，不出现那种湍流。”对于从小生活在河流边上的提问者来说，这一简短的回答就已经足够清晰了；但是对于较少接触河流的另一研究者来说，这一回答便显得太过模糊。这意味着，通过上手展开的寻视世界构成了人类沟通的基础，而 ChatGPT 的限定性则表明它完全缺失了这进入共同寻视世界的能力。

综上所述，在实验操作的回答中，ChatGPT 和人类的表现存在本质的差异：ChatGPT 的回答是一种展示（presentation），而人类的回答则是一种交流（communication）。ChatGPT 能够以教科书式的全面性、流程化、客观性与限定性来展示出实验操作的步骤；而人类受访者则事先预设受访者与自己有相似的经历与感受，并在此基础上作出具有针对性、碎片化、感官性与延展性的回答。

5. ChatGPT 对高等教育变革的启示：面向真实的世界

随着计算机科学的不断发展，ChatGPT 的语言能力终将超越所有个体，这将为未来的学习者提供巨大的挑战和机遇，而克服挑战、抓住机遇的关键则在于掌握正确应用明言知识的方法。但受限于语言模型的技术逻辑，我们也应当意识到，无论如何提升 ChatGPT 的算力，都无法跨越语言世界和寻视世界的鸿沟。寻视世界始终是人类独有的世界，寻视是人类区别于计算机的根本所在。

仅有语言世界的大语言模型擅长具有明确规则的知识，如程序代码的编写、学习路径的规划（如一张先学什么、再学什么的知识列表）、法律条文的输出等。而具有寻视世界的人类更擅长应用代码去解决实际问题、对学生进行个性化学习指导和评估、在教学过程中实现情感支持与道德引导等。

表 2 展示了两者在高等教育领域的五个维度的差异：对于仅有语言世界的 ChatGPT 和具有寻视世界的人类而言，前者只有理论、后者具有经验；前者进行从信息到知识方面的提高，

后者进行从知识到智慧层面的提高；前者借助强化主义的学习理论，后者偏向建构主义的学习方式；前者只是一个工具，而后者才培养全人。这也为高等教育变革带来五点启示。

表 2 大语言模型与人类导师的差异维度

划分维度	大语言模型-语言世界	人类导师-寻视世界	教师/学生关注的方向
理论和经验	有明确的明言知识	实践性的默会知识	实践教育、劳动教育
知识和智慧	从信息到知识方面的提高	从知识到智慧层面的提高	高阶能力和德性的培养
强化和构建	适配强化主义学习方式	适配建构主义学习方式	自我调节学习和元认知
工具和全人	工具性	情感陪伴、持续性、榜样	加强教师榜样引领作用
共性和个性	思路扩展、打破信息茧房	场景聚焦、个性化	语言和寻视世界的迁移

5.1. 理论和经验导向的实践教育

《大中小学劳动教育纲要》中提到：“必须面向真实的生活世界和职业世界”（教育部，2020），展现出对寻视世界的重视。然而，受到大学传统教学模式的惯性影响，我国的高等教育仍然具有重理论轻实践的倾向，“实践实习活动不过是一种点缀”（王洪才，2023）。当下中国的劳动教育也往往只关注到其体育锻炼功能，甚至只将其当作政治任务（刘向兵等，2021），而未能充分发挥其帮助学生上手操劳的价值。正如汽车等交通工具虽便利出行，却无法使人强身健体一样，在生成式人工智能的背景下，学生更有可能忽视实践与劳动的重要地位，从而不利于他们寻视世界的拓展（贾积有和张誉月，2023）。因此，未来的高等教育，必须加强对劳动教育的重视，将劳动教育提高至与专业教育、政治教育相当的地位。

5.2. 知识和智慧导向的高阶能力

人类的思想内容可分为数据、信息、知识和智慧（荆宁宁和程俊瑜，2005）。数据是原始的，信息由一系列特定的数据组成，知识为数据的应用，智慧则是对知识本质的理解。ChatGPT 带来的最直接的改变在于，明言知识的获得效率得到了前所未有的提升，以语言为载体的信息能够由繁杂走向精简、由碎片走向整合，并由此能够带动学习者进行从信息到知识方面的提高。但是，从知识到智慧层面的提高，如高阶能力、德性等，依然需要人类导师指引。因此，高等教育需要培养学习者超越传统的获取信息、整合信息的学习方式，着重培养目标意识、模型构建、创造创新、学科融合等高阶能力（赵丽红、左敏和黄先开，2023）。

5.3. 强化和构建导向的自我调节

微软研究人员测试了 GPT 在标准化考试中的表现，发现 ChatGPT 在外语类、事实类、程序类上的得分高于人类的最高分。这些知识依赖强化主义学习理论，通过不断重复便能习得。而 ChatGPT 在重视逻辑推理的数学、具有创新性的语文学科上还没有超过人类平均分（Zhong et al., 2023）。因此未来的高等教育更要注重培养学生通过认知主义、建构主义、连接主义学习理论知识的能力，深入掌握知识的原理层面，并帮助他们将所学知识与其他概念、知识相互关联（王竹立，2023），从而增强学生的自我调节能力和元认知策略，在学习的过程中积极进行自我计划、监控和评估。

5.4. 工具和全人导向的人文关怀

从技术原理上看，ChatGPT 的本质是更加高效、更加实用的搜索引擎。而高等教育要培养的不是“工具人”，而是能够有效使用工具的“人” （本刊编辑和张绒，2023）。因此，未来的教学不再是一种传播知识的手段，更是一种基于师生关系的实践。关怀学生是教学实践的基础，也是教学实践独特的内在卓越标准。这要求高等教育的教师进行自身角色定位的转变，从知识传播转为提供持续的情感陪伴和榜样作用，使高等教育回归“育人”的使命。

5.5. 共性和个性导向的寻视拓展

ChatGPT 的回答具有发散性，可以帮助学生扩展思路、打破信息茧房。然而，由于 ChatGPT 不具备寻视世界，只能根据提供的数据进行学习建模，无法像人类一样从性格、肢体语言、神态等真正地“认识”这个学生，无法对学生未来发展、职业选择或人生困惑提供细致入微的解决方案。因材施教需要人类导师聚焦特定问题和场景，帮助学生打破、扩展寻视世界，真正提供个性化的辅助手段和建议（李芒和张华阳，2022）。此外，由于 ChatGPT 具体功能依赖于提示词（prompt）的设计，个人提出的 prompt 也能体现个人的“寻视世界”，这也提醒我们要训练学生的文字达意能力，将个人的寻视世界体现在提示词中，将无寻视世界的语言世界迁移到寻视世界中，让 ChatGPT 提供更有效的指导。

6. 结论

本研究通过对 ChatGPT 与人类实验操作者分别进行文字访谈，发现 ChatGPT 和人类操作者的回答呈现出显著的差异：首先，ChatGPT 的答案倾向于具备全面性、流程化、客观性与限定性的特征，人类的答案更注重于针对性、碎片化、感官性与延展性。其次，ChatGPT 的回答是在纯粹的语言环境中生成的，而人类的回答则反映了他们对实际上手的寻视经验。最后，ChatGPT 在提供全面、详尽的工具性指导方面具有明显的优势，然而人类操作者在传授默会知识和引导学生全人发展方面拥有不可替代的价值。该结果不仅符合 ChatGPT 的技术局限，也与寻视理论的预测保持一致，由此证明了大语言模型与人类之间的确具有不可逾越的鸿沟。基于这些根本差异，本研究认为人工智能时代下的高等教育不仅应该充分利用信息技术的优势，还应该重视以人为本的发展方向，并提出以下 5 点建议：1）强化学生的实践教育、劳动教育；2）培养学生高阶学习能力；3）掌握建构主义学习方式；4）突破教育的工具性，回归教育的人文性；5）开发通过生成式人工智能拓展寻视世界的方式。

本研究为深入认识人工智能与人的本质均提供了有力的理论与实证支持，为利用人工智能、回归人的本质提供了有力的参考。此外，本研究是国内外首个以 ChatGPT 为访谈对象并将其与人类受访者进行平等对比分析的研究，拓宽了对“实证”的理解，也为之后的研究者提供了一条全新的实证研究路径，在方法论层面具有重要的创新意义。

参考文献 *（请参阅 APA 6th 格式）*

本刊编辑和张绒. (2023). 生成式人工智能技术对教育领域的影响——关于 ChatGPT 的专访. 电化教育研究, 44(02), 5-14. https://doi.org/10.13811/j.cnki.eer.2023.02.001

波兰尼. (2017). 认知与存在. 南京大学出版社.

陈巍. (2011). 交互心灵观与具身认知科学的现象学始基. 科技导报, 29(24), 80.

海德格尔. (2014). 存在与时间 (修订译本). 生活·读书·新知三联书店.

贾积有和张誉月. (2023). 人工智能与教育:机遇、挑战与对策. 北京大学教育评论, 21(01), 49-61+188-189.

教育部. (2020). 教育部关于印发《大中小学劳动教育指导纲要（试行）》的通知. 中华人民共和国教育部公报(Z2), 2-11.

荆宁宁和程俊瑜. (2005). 数据、信息、知识与智慧. 情报科学(12), 1786-1790.

李丽萍和卢晓东. (2023). 实验教育中面向创新的具身认知发展价值——以化学学科为例. 现代大学教育, 39(02), 103-110.

刘向兵、柳友荣、周光礼、卢晓东、孙振东、张有奎、廖辉、王岩和蔡文浩. (2021). 全面加强新时代高校劳动教育（笔谈）. 中国高教研究(04), 9-13.

王洪才. (2023). 对 ChatGPT 的高等教育哲学审视. 重庆高教研究, 11(04), 3-13.

王佑镁、王旦、梁炜怡和柳晨晨. (2023). “阿拉丁神灯”还是“潘多拉魔盒”：ChatGPT 教育应用的潜能与风险. 现代远程教育研究, 35(02), 48-56.

王竹立. (2023). 回归与重构:智能时代的新知识观——再与陈丽教授等商榷. 电化教育研究, 44(07), 13-20. https://doi.org/10.13811/j.cnki.eer.2023.07.002

吴青和刘毓文. (2023). ChatGPT 时代的高等教育应对：禁止还是变革. 高校教育管理, 17(03), 32-41. https://doi.org/10.13316/j.cnki.jhem.20230504.004

荀渊. (2023). ChatGPT/生成式人工智能与高等教育的价值和使命. 华东师范大学学报(教育科学版), 41(07), 56-63.

郁振华. (2001). 波兰尼的默会认识论. 自然辩证法研究(08), 5-10.

郁振华. (2006). 默会认识论视野中的“在世”观念. 学术月刊(07), 53-60.

郁振华. (2020). 理论之发生——兼论海德格尔和实用主义的关系. 天津社会科学(01), 21-35. https://doi.org/10.16240/j.cnki.1002-3976.2020.01.003

张峰和陈玮. (2023). ChatGPT 与高等教育：人工智能如何驱动学习变革. 重庆理工大学学报(社会科学), 37(05), 26-33.

赵丽红、左敏和黄先开. (2023). 人工智能时代高等教育教学的变革指向：培养高阶思维. 北京师范大学学报(社会科学版)(04), 40-48.

钟秉林、尚俊杰、王建华、韩云波、刘进、邹红军和王争录. (2023). ChatGPT 对教育的挑战(笔谈). 重庆高教研究, 11(03), 3-25.

Bengio, Y., Ducharme, R., & Vincent, P. (2000). A neural probabilistic language model. Advances in neural information processing systems, 13.

Brown, P. F., deSouza, P. V., Mercer, R. L., Pietra, V. J. D., & Lai, J. C. (1992). Class-based n-gram models of natural language. Comput. Linguist., 18(4), 467–479.

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. Advances in neural information processing systems, 30.

Polanyi, M. (2013). The Study of Man: the Lindsay memorial lectures 1958. Routledge.

Searle, J. R. (1982). The Chinese room revisited. Behavioral and Brain Sciences, 5(2), 345-348.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Weizenbaum, J. (1983). ELIZA — a computer program for the study of natural language communication between man and machine. Commun. ACM, 26(1), 23–28.

Zhong, W., Cui, R., Guo, Y., Liang, Y., Lu, S., Wang, Y., Saied, A., Chen, W., & Duan, N. (2023). Agieval: A human-centric benchmark for evaluating foundation models. arXiv preprint arXiv:2304.06364.

人工智能时代教师数字素养评价研究

——基于《教师数字素养》行业标准的思考

Research on the Evaluation of Teacher Digital Literacy in the Era of Artificial Intelligence

——Based on the Industry Standards of Teacher Digital Literacy

王旖[1 2]，彭俊[2*]
[1]长沙商贸旅游职业技术学院
[2]澳门城市大学 教育学院
[*] junpeng@cityu.mo

【摘要】在人工智能的时代背景下，数字素养已然成为教师开展教育教学活动的核心基本素养。随着《教师数字素养》标准的正式发布，如何精准地评价教师的数字素养引发了广泛关注。本文通过深入剖析教师数字素养评价的现实意义，系统梳理了国内外关于此领域的研究现状，并揭示当前评价存在的问题。结合人工智能前沿技术所带来的变革，本文从评价主体、评价内容、评价方法、评价标准和评价模式五大维度，全面探讨了人工智能时代下教师数字素养评价体系的构建。文章最后提出了人工智能时代教师数字素养评价实施建议：一是强化顶层设计，协同推进评价改革；二是完善数字底座，突破关键评价技术；三是构建多元主体评价，强化数据共享与应用；四是提供智能化的辅助，强化评价反馈功能。

【关键词】教师数字素养；评价研究；人工智能时代；评价体系的构建；实施建议

Abstract: *In the era of artificial intelligence, digital literacy has become an essential basic literacy for teachers to carry out educational and teaching activities. With the release of the Teacher Digital Literacy standard, the issue of accurately evaluating teacher digital literacy has received widespread attention. The article summarizes the practical significance of teacher digital literacy evaluation, summarizes the current research status at home and abroad, and identifies existing problems. Based on the changes brought about by cutting-edge artificial intelligence technology, the article analyzes the construction of a teacher digital literacy evaluation system in the era of artificial intelligence from five aspects: evaluation subject, evaluation content, evaluation method, evaluation standard, and evaluation mode. Finally, suggestions were put forward for the implementation of digital literacy evaluation for teachers in the era of artificial intelligence: firstly, to strengthen top-level design and promote evaluation reform in a coordinated manner; The second is to improve the digital base and break through key evaluation technologies; The third is to construct a diversified subject evaluation, strengthen data sharing and application; The fourth is to provide intelligent assistance and strengthen the evaluation and feedback function.*

Keywords: Teacher's digital literacy; Evaluation research; The era of Artificial Intelligence; Construction of evaluation system; Implementation suggestions

1. 人工智能时代教师数字素养评价的现实意义

随着信息技术的蓬勃发展，人工智能技术的出现标志着人类已经进入智能时代。以人工智能技术、大数据为代表的新一代信息技术正趋向于智能化，赋予了智能时代独有的特征，例如人机协同、万物智联等。这些技术给教育领域带来了深远影响，极大地改变了教育环境、教育模式、教学组织形式以及教育评价。作为教育管理的重要环节，教育评价对于了解教育的适宜性和有效性，调整和改进教育工作，以及提升教育质量具有不可替代的作用。在人工智能的助力下，通过自然语言处理、机器学习、数据挖掘、人脸识别等先进技术和方法，我

们可以为评估主体提供更为客观、及时、精准的评估结果，帮助教师获得更具针对性和个性化的评估与指导，同时为教学改进提供宝贵的反馈。

2022 年 11 月 30 日，教育部颁布了《教师数字素养》行业标准（教科信函〔2022〕58 号），这是我国首份关于教师数字素养的教育行业标准，对于推动教师数字素养的高效发展具有举足轻重的意义。该标准清晰地构建了教师数字素养的五维三级框架，涵盖数字化意识、数字技术知识与技能、数字化应用、数字社会责任以及专业发展这五个核心维度。在每个一级维度下，又进一步细分了多个二级维度、三级维度和详细描述。此标准主要用作对教师数字素养的评估和培训工作。其中，教师数字素养被定义为教师应具备的，利用数字技术获取、加工、使用、管理和评价数字信息与资源的能力，以及发现、分析和解决教育教学问题，优化、创新和变革教育教学活动的意识、能力和责任感。《教师数字素养》行业标准的出台，为教师数字素养的评估工作奠定了有力的技术基础。在人工智能信息时代，如何有效依据该行业标准，并整合科技前沿手段，有效促进了教师数字素养的评估工作，已成为新时期教师发展中亟待解决的关键问题。

2. 教师数字素养评价的国内外研究现状

2.1 国外教师数字素养评价相关研究

随着人工智能与教育教学的融合发展，数字技术已逐渐成为教育变革的重要驱动力量，数字素养，作为未来教师的核心素养，已经引起了国外的深入研究。数字素养是学术课程中整合技术的挑战之一（Blau et al., 2020），它代表了处在一个复杂且零散的信息生态系统所需的能力和技能（Eshet,2004）。为了推动数字素养的发展，各国研究机构和国际组织已相继发布了一系列教师数字素养相关的评估框架或标准。其中“数字素养框架”被定义为包含六个类别的思维，分别是照片视觉思维、实时思维、信息思维、分支思维、复制思维和社会情感思维（Eshet,2012）。此外，数字素养还可以被划分为三个集群：识别和使用数字内容、创建数字内容，以及交流或共享数字内容（Pangrazio & Luci,2021）。

国外学者在数字素养评估方面的起步相对较早，通过发展了自评工具和他评价工具来促进了数字素养评估实践的开展。联合国教科文组织于 2018 年推出了《数字素养全球框架》，该架构共包含了七个数字素养范畴，其评价方式主要以“自我报告量表”为主，辅以“在线知识测试”，并融入了具有诊断功能的电子画像技术（刘宝存&岑宇，2022）。欧盟开发了多个版本的教师数字素养自我评估工具，《欧洲教育工作者数字素养框架》将教育工作者的数字胜任力细分为专业化参与、数字资源、教学与学习、评价、赋权学习者六个方面，并据此研发了相应的教育工作者在线测评工具，供全球教育工作者使用以实现在线自我评估（闫广芬&刘丽，2022）。西班牙教育技术与教师培训学院依据本国的《教师通用数字素养框架》，打造了“教师数字素养档案包”，教师自行在网络上传自己在数字素养方面取得的主要成果，从而获取数字素养水平的认证（Digital Docente，2021）。瑞士的学者开发了一套职业教育与培训数字能力量表来测试教师的教师数字化能力（TDC），并对瑞士 1692 名 VET 教师进行了实证研究，发现教师的数字能力总体程度类似，但在一些子能力方面存在一些差异，同时教师对技术的态度和数字工具使用频率对教师的 TDC 有影响（ Cattaneo，Antonietti & Rauseo，2022）。

通过梳理总结国外关于教师数字素养评价的文献资料，我们发现科学合理的指标体系是开展教师数字素养评价实践的基础条件。随着互联网技术的快速发展，在线评估已经成为目前一种较为普遍的方式。此外，表现性评价与成长记录袋评价也开始被应用于教师数字素养评价中，这些评价方式有助于减少评价结果对教师主观判断的依赖，使评价结果更加客观有效。

2.2 国内教师数字素养评价相关研究

当前国内教师数字素养评价研究主要可分为关于教师数字素养的内涵探究和关于教师数字素养评价的实践研究两方面。

在教师信息素养相关评价内涵探究方面，有研究者通过研究《教师数字素养》标准，对教师数字素养概念作了阐释（郭绍青&林丰民，2023）。还有研究者基于《教师数字素养》标准，提出了开展分层分类的精准评价、数智驱动的全面评价等实现路径促进教师数字素养发展(吴砥&桂徐君,2023)。在评价改革未来趋势的基础上，有研究者指出随着智能技术自身的发展进步及其在评价中日益广泛深入的应用，将助推结果评价更加科学、过程评价更加智慧、增值评价更加便捷、综合评价更加完善（吴砥，郭庆，吴龙凯&程浩，2023）。实践方面，研究者通过设计问卷，运用问卷调查研究方法开展评价研究（方紫帆&徐娟，2023）。有学者参考教育部发布的《教师数字素养》以及《信息科技课标》，确定了中小学教师数字素养的五个维度，编制了《中小学教师数字素养测试题》，发现中小学教师数字素养整体水平较高，在数字社会责任维度上表现最佳，在数字技术知识与技能、数字化应用维度上表现最差，同时关注了教师地区和性别差异（宋灵青，2023）。

综上所述，当前有关教师数字素养评价方面的研究仍然较为薄弱，缺乏统一的数字素养评价标准，评价手段和方法相对单一，评价实践覆盖面不广等问题较为突出。教师数字素养已成为当今教育工作者进行有效教育教学的必要素质之一，因此，进行教师数字素养评价，并依据评价结果有针对性地提高教师数字素养水平，是当前教师队伍建设的重要任务。因此，本研究深入分析了我国教师数字素养评价的现状，构建了教师数字素养评价的体系并提出实施路径，以期能够为相关研究和实践提供有益的参考思路。

3. 人工智能时代教师数字素养评价存在的问题

当前我国教育信息化发展已进入 2.0 阶段，教师应用数字技术优化教育教学活动的意识和能力已有较大提升(方紫帆&徐娟,2023)。教育部出台《教师数字素养》标准后，部分地区也开展了一些评价实证研究。目前来看，我国教师数字素养评价尚处于起步阶段，在顶层设计、评价方式、评价运用等方面有待进一步完善与提高。

3.1 评价顶层设计缺乏，无法满足技术发展的需求

评价教师数字素养需要完善的机制来支撑，否则即使制定了标准，在实践过程中也难以有效实施。然而，现有的评价机制还存在诸多不足，具体表现在以下几个方面：一是缺乏系统的培训机制，政府和学校未能提供全面且系统的教师数字素养培训计划，或培训资源有限，难以满足所有教师的需求(相巨虎&陈鹏,2023)；二是缺乏专业的评价指导，随着人工智能等科学技术和产品更新迭代的速度远超现有数据平台的更新速度，现有教师数字素养培训的内容过于侧重基础技术操作，缺乏对人工智能在教学设计、学生指导和数据分析等更深层次的应用指导(周刘波&张梦瑶,2023)。虽然一些城市已经建立了评价平台和系统，但预期用途与实际实施之间存在较大差距(张硕,2023)；三是缺乏持续的支持和反馈机制，教师数字素养的培训往往仅限于短期课程，缺乏长期的跟踪和反馈机制(何文涛&庞兴会,2023)，从而影响了评价的实际效果。

3.2 评价方法较为单一，数据客观性和真实性难以核实

教师的数字素养涉及面广，涉及数字化意识、数字技能知识与技能、数字化应用等方面，对其进行评估需要采集大量的定量和定性的数据来得出评估结论(冯修猛&范笑妤,2022)，多样化的评价方法可以保证数据来源的广泛性和真实性。但目前教师数字素养评价方法注重结果评价，过程性评价较少。评价方式主要依靠表现或成果评价，或是采取问卷、知识测试等方式，现代教育评价理论与方法很少被利用。同时教师数字素养评价中的主观性也可能导致

数据真实性难以核实，尤其是数字化意识、数字社会责任一般通过问卷和访谈来的研究，教师的情感、心态、价值观等指数，因为无法外化，在评估过程中很难以定量的方式得到真实的结果，导致了此类指数的评估主观性很强，在一定意义上削弱了数字素质评估的公正性。

3.3 数据孤岛现象存在，难以发挥评价数据价值

在人工智能时代，教师数据平台应该是开放和共享的，使不同的组织和研究者可以共享数据资源，同时数据应该具备互操作性，能够被不同的人工智能系统和算法所使用和分析。由于技术限制、安全考量和层级管理等原因，我国教师数据数量庞大，教师们在不同的平台上申报信息、教学资料信息，但由于这些资源库之间相互联系彼此独立，丰富的信息资源不能得到合理发掘与广泛共享，从而必然地产生了信息“信息孤岛”、“信息壁垒”的问题。教育信息平台数据的缺乏与共享体系，使得教育评价主体在对教师数字素质的评估和对接与管理等环节中，需要耗费大量的人力、物力和财力，也易造成数据资源的浪费，影响了评价效果和效率。

3.4 评价结果使用存在偏差，缺乏有效的反馈机制

现有评价制度功能之一是帮助管理者对管理对象进行绩效考核、职称晋级或是奖金发放等任务。在教师评价层面，其功能往往定位于对教师的教学成绩和科研成果进行评定，一般在特定的时段开展，通常是一个学期期末或一个年度年末，教师需要提供各类表格、在线课程资源平台、在线培训证书截图等其他佐证材料来开展测评，评价主体需要耗费大量精力从繁杂数据中提取出有价值的内容。在经过相关部门的核验后，教师会收到评价结果，一般为等级标定，如优良，良好，一般，不合格等，但缺乏有针对性的意见和反馈建议，且评价结论一旦公布，评价工作也宣告完成，缺乏后续的反馈，教师在决定今后究竟应朝哪方面努力的问题上变得无所适从。

4. 人工智能时代教师数字素养评价体系的构建

教师数字素养 Gilster 于 1997 年首次提出“数字素养”（Digital Literacy）后，在新媒体的帮助下，其概念在世界范围内得到广泛关注和传播（Gilster P.1997）。新概念的提出对教师传统的评价提出了新的挑战。构建教师数字素养评价体系，本文从评价主体、评价内容、评价方法、评价标准、评价模式五个方面来进行关系梳理(见图 1）。

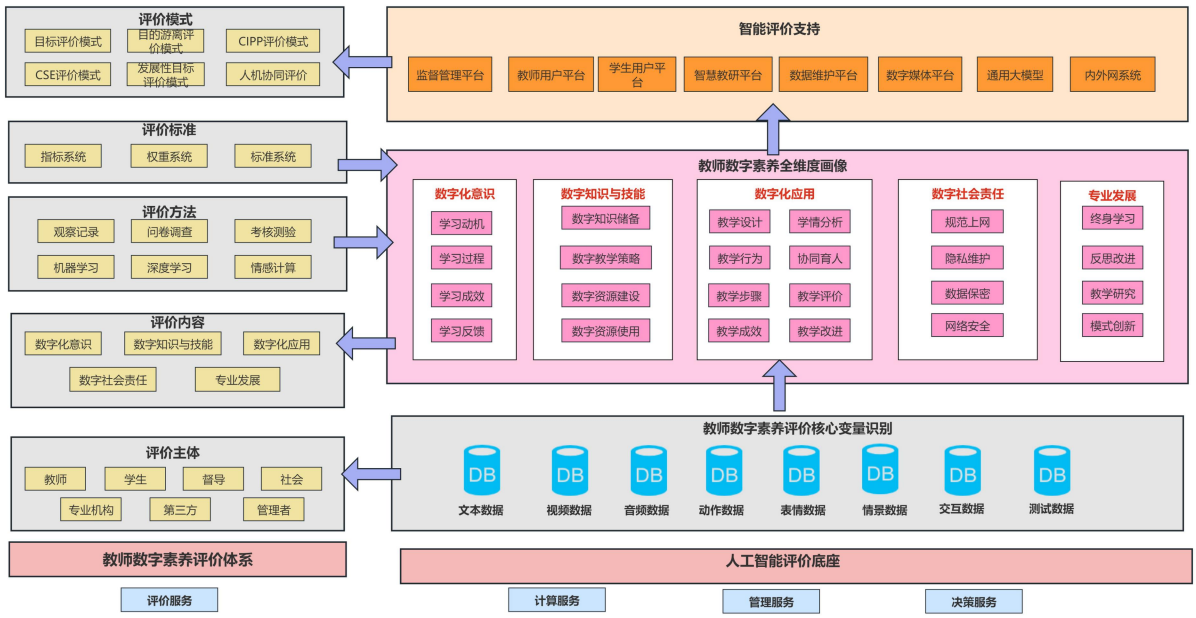


图 1 教师数字素养评价体系

4.1 评价主体：人工智能驱动教育评价主体多元化

教育评价作为一种兼具繁杂性与丰富性的实践活动，其生命力在于多方主体的共同投入与协商（陆启越&余小波，2018）。“多元主体参与”最早由美国教育评价学专家派特（Patton，M.Q）于1978年在《采用聚焦评价》一书中正式提出（史晓燕，2020），后来进一步发展出了多元主体“全面参与”“协商对话”“共同建构”以及“价值多元化”评价理念（朱德全&吴虑，2019）。智能技术赋能教育评价，可以通过利用云计算的存储计算、资源共享、服务提供等技术，对教育评价的形式、渠道、主体等进行拓展和丰富，从而提高教育评价的多元性和包容性。例如，智能技术可以通过构建云端的教育评价平台，实现对多种类型的教育评价数据（如文本、视频、音频、动作、表情、交互等）的快速处理和呈现，同时支持多种角色（如教师、学生、督导、专业机构、社会等）的参与和协作。

4.2 评价内容：依托行业标准确认和改进

在《教师数字素养》教育行业标准出台后，对于教师数字素养进行了定义和明晰，教师数字素养框架包括5个一级维度、13个二级维度和33个三级维度，为教师数字素养评价提供全面的导向和支撑。数字化意识指的是教师在开展教育教学工作应具备的基本意识，包括数字化认识，数字化意愿，以及数字化意志；数字技术与技能包括数字技术知识和数字技术技能；数字化应用体现的是教师开展各类数字化教学活动的综合能力，也是核心内容，包括数字化教学设计、数字化教学实施、数字化学业评价和数字化协同育人；数字社会责任是对教师在数字社会的基本道德品质要求的规定，包括法治道德规范和数字安全保护；专业发展主要是指教师的数字化学习与研修、数字化教学研究与创新。

4.3 评价方法：全面科学的大数据评价

评价方法是指通过不同的途径、手段对被评价对象的工作、行为和绩效等进行量化或非量化的分析和判断。常见的教育评价方法包括了观察记录、问卷调查和测验考核等，人工智能可以创新评价方法，进行全面科学的大数据评价：一方面可以利用人工智能的数据挖掘、机器学习、自然语言处理等技术，对海量的教育数据进行有效地收集、整理、分析和应用，从而提高教育评价的科学性和准确性；另一方面通过数据验证、去重和重复检验，可以增加信息资料的真实性，使每项信息都有迹可循，全面分析教师数字素养相关信息，从而深度挖掘潜藏在数据之后的深层关系与价值，为政府制定教师发展策略及教育政策提供证据支持。

4.4 评价标准：教师数字素养评价指标体系

教师数字素养评价指标体系构建是教师数字素养评价研究的核心问题，一般包括指标系统、权重系统和标准系统。教师数字素养评价相关的指标体系主要考察教师数字化意识、应用数字化技术开展教学活动的知识和能力，还要考虑到社会责任和专业发展等综合方面。可以利用大数据的关联分析、预测分析、优化分析等技术，对评价的标准和权重进行动态的调整和优化，从而提高教育评价的精准性和有效性。例如，智能技术可以根据教师的专业背景、职业规划、潜能发展等因素，为教师匹配合适的数字素养评价指标体系，实现个性化的教育评价。

4.5 评价模式：过程与结果相结合的智能协同评价

人工智能时代下对教师数字素质的评价内容不但重视其数字知识掌握、信息技术运用情况的外在表征，同时重视老师的数字意识、法治道德规范、数字安全保护意识等内在思维，除了传统的目标评价模式、目标游离评价模式、CIPP模式外，还需要不断创新评价模式，提高评价的准确性和科学性。一方面对教师课堂教学全程实施监控评价，对教学过程实施事前控制、过程控制与事后控制,综合利用现实的课堂教学全过程数据刻画教师数字素养画像，帮助精准评价教师数字素养水平；另一方面可以通过各类数据平台进行教学实时分析，科学调

整教师的教学目标和计划，为教师提供适合的教学资源和方法，并根据学生的反馈和表现，为教师提供及时的评价信息和建议，从而实现人机协同的教学评价。

5. 人工智能时代教师数字素养评价实施建议

5.1 强化顶层设计，协同推进评价改革

教师数字素养的评价离不开政府部门的政策引导，目前我国教师数字素养行业标准的文件虽已出台，但考核和评价机制仍需结合实践不断完善与细化。一是区域层面需要建立新的政策机制，大力推动教师数字素养评价项目，通过教师数字素养评价了解现状，从而加快教师数字素养培育的创新变革、推动人才培养高质量发展；二是建立起关于教师数字素养的培训机制，由学科专家、教研员与教育技术专家共同组成专家团队，打破学科教学与教育技术各自为政的壁垒，让教师们了解并掌握数字素养评价背景、评价目的、评价方法等，真正实现数字技术与学科教学融合、评价与教学改革融合；三是搭建教师数字素养评价的智慧平台系统，充分发挥教育大数据的采集和评价作用，利用教师课堂教学、专业学习的伴随性数据采集、教师画像等功能，实现评价由“静态粗放式”支持向动态精细化”支持转变，帮助教师通过 APP 或微信小程序在移动端完成评价以及查看反馈。

5.2 完善数字底座，突破关键评价技术

数字底座是指在教育领域中，用于支持和促进教育信息化、数字化发展的基础设施和平台，是构建数字化学习环境和教育生态系统的基础，能为教育工作者、学生和家长提供良好的数字化教育资源、工具和服务。相关部门应积极完善网络环境健全数据中心建设、推动平台开放协同助力教师数字素养评价（史陈新&李艳燕，2022）。一是打造评价数据大脑，建立数字素养数据库，借助数字设备如移动智能设备、教学工具等智能数据采集终端，实现多空间活动的数据采集；二是构建教师素养模型，通过智能化分析处理多模态数据，实现对教师的全方位追踪和变化分析，形成教师数字画像，帮助教师做好职业生涯规划；三是围绕教师数字素养标准，形成按学期和学年等期的动态发展跟踪监控制度，通过对比研究教师的数字素质各个层面呈现的不同变化趋势，描绘教师数字素质的个体增长曲线和群体增长态势，达到对教师数字素质持续跟踪、深入全面的研究，为教学管理人员打造优秀教师提供政策保障。

5.3 构建多元主体评价，强化数据共享与应用

多元化的评价主体可以扩大信息交流范围，提高问题分析的全面性，也有利于被评教师多角度认识问题和反思自己的数字素养水平和成效。一是积极吸纳各方评价主体参与评价标准的制定和评价的实施中，除了教师自评和互评，还要争取专业的评价机构、教育决策机构、教学督导人员、学生、家长和个体以及学校以外的其他有关人员，通过不同方式参与评价，进一步扩大评价数据的来源，充分发挥多元评价主体的作用，使评价结果更为客观；二是要实现资源共享，打通不同数据平台的壁垒，建立数据资源目录的动态更新机制，规范数据收集与使用，建立数据共享交换机制，减少重复收集；三是建立安全评价制度，提高安全保障水平，从制度、技术手段、操作等多维度确保安全，避免数据风险，实现数据的最高价值。

5.4 提供智能化的辅助，强化评价反馈功能

评估的直接目的有两个方面，一方面是帮助教育管理部门了解现实情况，分析政策实施效果及为作出进一步的决策提供依据，另一方面是为被评对象改进工作提供依据。教师评估作为教育管理的重要环节，过去比较注重鉴定与分类，现在应该是注重对教师的提升与发展。教师数字素养的评价需要借助智能化手段，强化评价反馈机制，一方面建立教师数字素养反馈机制，并使得其贯穿于整个评估过程，通过人工智能大数据技术的实时分析功能，在教学的全过程实时调整教学内容和教学策略；另一方面利用人工智能分析技术对教师课堂行为、课堂互动话语、情绪变化进行智能分析，形成总体分析报告，把教师数字素能评价结果纳入

教师的日常测评和考核过程中，完善相应的管理的激励机制，树立典型榜样，为打造数字素养和专业学术相结合的未来教师提供方向。

参考文献：

[1] Ina, B., Tamar, S. I., Orit1, A. (2020). How does the pedagogical design of a technology-enhanced collaborative academic course promote digital literacies, self-regulation, and perceived learning of students? *Internet and Higher Education,* 25(0), 1096-7516. https://doi.org/ 10.1016/j.iheduc.2019.100722

[2] Eshet, Y.（2004）Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, 2004.13(1),93-106.

[3] Eshet, Y. (2012) Thinking in the digital era: A revised model for digital literacy. *Issues in Informing Science & Information Technology,* 9(2), 267-276. https://doi.org/ 10.28945/1621

[4]Pangrazio,Luci.Sefton-Green,Julian(2021).Digital Rights, Digital Citizenship and Digital Literacy: What's the Difference? *Journal of New Approaches in Educational Research*.(10) 15-27

[5]刘宝存&岑宇.(2022).世界教育数字化转型的动因、趋势及镜鉴.现代远程教育研究 (06), 12-23.

[6]闫广芬&刘丽.(2022).教师数字素养及其培育路径研究——基于欧盟七个教师数字素养框架的比较分析.比较教育研究(03),10-18.doi:10.20013/j.cnki.ICE.2022.03.02.

[7]Marco de Competencia Digital Docente [EB/OL].(2020-07-13)[2021-09-13]. http://aprende.intef.es/mccdd/.

[8]Cattaneo,A.P., Antonietti,C.,& Rauseo,M.(2022).How digitalised are vocational teachers? Assessing digital competence in vocational education and looking at its underlying factors.*Computers & Education.176,* Article 104358.

[9]郭绍青,林丰民,于青青,张治,张云峰,段元丽& 管延端.(2023).数字化赋能教师专业发展实践探索.电化教育研究 (07), 96-106. doi:10.13811/j.cnki.eer.2023.07.013.

[10]吴砥,桂徐君,周驰&陈敏.(2023).教师数字素养：内涵、标准与评价.电化教育研究(08), 108-114+128. doi:10.13811/j.cnki.eer.2023.08.015.

[11]吴砥,郭庆,吴龙凯&程浩.(2023).智能技术赋能教育评价改革.开放教育研究(04), 4-10. doi:10.13966/j.cnki.kfjyyj.2023.04.001.

[12]方紫帆&徐娟.(2023).国际中文教师数字素养指标体系建构研究.天津师范大学学报(社会科学版)(06),25-33.

[13]宋灵青.(2023).我国中小学教师数字素养的实然状态与突破路径——基于全国9405名中小学教师的测评.中国电化教育(12),113-120.

[14]方紫帆&徐娟.(2023).国际中文教师数字素养指标体系建构研究.天津师范大学学报(社会科学版)(06),25-33.

[15]相巨虎&陈鹏.(2023).迈向高质量：教师队伍数字化转型的洞见症结与破解之道.中国电化教育(09),59-68+75.

[16]周刘波,张梦瑶&张成豪.(2023).数字化转型背景下教师数字素养培育：时代价值、现实困境与突破路径.中国电化教育(10),98-105.

[17]张硕.(2023).教育数字化转型背景下教师数字素养提升研究.教书育人(28),44-46.

[18]何文涛,庞兴会,朱悦,阮高峰&石倩.(2022).人工智能时代中小学教师信息化教学能力发展现状与提升策略.现代教育技术(03),92-101.

[19]冯修猛&范笑妤.(2022).区块链技术驱动教育评估现代化：现实图景、适用价值与实施进路.教育发展研究(19),69-74. doi:10.14121/j.cnki.1008-3855.2022.19.011.
[20]GILSTER P（1997）. Digital literacy.*New York:Wiley Computer Publications*
[21]陆启越&余小波.(2018).高校德育评价范式及其转变.湖南师范大学教育科学学报(02),103-108.doi:10.19503/j.cnki.1671-6124.2018.02.016.
[22]史晓燕（2020）.教师教学评价：主体·标准·模式·方法.北京师范大学出版社.
[23]朱德全&吴虑.(2019).大数据时代教育评价专业化何以可能：第四范式视角.现代远程教育研究(06),14-21.
[24]史陈新,李艳燕,方海光,王红& 刘雪娇. (2022).迈向数字教育：北京市教育数字化转型的现实意涵与实践策略.中国教育信息化(10),58-66.

移动教学对职前教师技术认知心理与 TPACK 发展的影响研究

Study on the Impact of Mobile Teaching on Pre-service Teachers' Technological Cognition and TPACK Development

吴莉[12]，刘延申[1]
[1]华中师范大学国家数字化学习技术工程研究中心， 湖北 武汉 430000；
[2] 南师范大学教育与心理学 ， 福建 漳州 363000
Lixia_wu@163.com

【摘要】 信息技术与教育融合发展的当下，移动教学在教师教育中的普 用为职前教师技术认知及 TPACK 发展培养带来机 。为探究移动教学对职前教师技术认知及 TPACK 发展的影响，以技术接受模型和 TPACK 框架为依据，设计相应 卷，对 149 个信息技术专业师范生展开了调查。结果表明：(1)移动教学对信息技术职前教师技术认知心理有正向促 作用。(2) 移动教学对信息技术职前教师 TPACK 七个单元均有不同程度的促 作用。本研究的探讨为职前教师培养提供一定的参考和借 。

【关 词】 移动教学；职前教师；技术认知；TPACK

Abstract: *When information technology and education are integrating, the widespread adoption of mobile teaching in teacher education presents opportunities for the development of pre-service teachers' technological cognition and TPACK development. To investigate the impact of mobile teaching on the technological cognition and TPACK development of pre-service teachers, a survey based on the Technology Acceptance Model and the TPACK framework was designed and conducted with 149 information technology education students. The analysis results show: (1) Mobile teaching has a positive effect on the psychological technological cognition of pre-service information technology teachers. (2) Mobile teaching promotes the development of the seven components of TPACK for pre-service information technology teachers to varying degrees. This research provides certain references for the cultivation of pre-service teachers in the information age.*

Keywords: Mobile teaching, pre-service teacher , Technology Perception ,TPACK

1. 引言

研究表明，TPACK 是教育信息化时代教师所必备的知识框架。移动教学作为一种整合了移动技术的创新教学模式，已被广泛应用在正式学习环境中。移动教学实践能否对职前教师相关能力发展起到促 作用呢？ Matthew L. Wilson 等人（2020） 过对相关研究（num=3271） 行元分析发现，注 技术整合的教师教育课程对职前教师知识增 有 常显著的正向影响（ES 值为 1.057）。然而，该研究没有明确指出是何种形式的技术整合课程。另外，该研究并没有对技术整合知识各个组成 分分别展开探讨。因此，本研究拟深入探讨移动教学对职前教师技术认知心理及 TPACK 各个组成 分的影响，为职前教师 TPACK 发展提供思路。

2. 相关研究基础

（1）技术认知。技术认知是一个认识层 的概念，是一种典型内 的心理活动。教师的技术

感知，一般用来衡　教师对于信息技术的感知程度。**（2）教师 TPACK 结构研究。**近年来，关于职前教师 TPACK 结构的研究主要　过调查　卷、结构方程模型等方法展开。我国学者张哲等人（2015）　过对某师范　校 384 个职前教师展开调研，并　用结构方程模型方法。结果显示该师范大学职前教师 TPACK 结构中的七类知识具备相对较　的水平。**（3）移动教学模式的本质特征。**Michael（2019）构建的移动教学特征框架提出如下观点：“学习者是移动的”、“设备是移动的”、“教学/学习内容是移动的”、“数据服务是持续不断的”以及“教师提供的指导是易得的”是移动教学构成的必　特质。

3. 研究方法

（1）　卷设计。　用　卷调查的方法。共有 3 个　分：对象人口学信息、移动教学实施情况和移动教学对职前教师技术认知和 TPACK 发展的影响。**（2）　卷　目、信效度分析。**　取某　校教育技术学专业大四学生（n=41）做　测试。　**目分析结果：**所有　　t 值　大于 3.000，表明　　独立性好；**信度分析结果：**各维度及总体 α 值分别为 0.780−0.879、0.895，说明　卷内和整体的一致性好；**效度分析。**探索性因子分析 KOM 值为 0.798，Bartlett 球形检测值为 0.000（小于 0.05），说明变　　相互依存。主成分分析显示：所有　　共同性萃取值均大于 0.2，因此保留全　　。**（3）确定研究数据样本。**回收　卷 320 份，筛　得到有效样本 149 份。

4. 结果分析

(1)技术认知（TP）结果分析。TP 维度平均分为 3.79，各　指标的得分均在 3 以上，说明研究对象认为基于移动教学环境的技术课程有助于他们技术认知水平的提升。**(2)TPACK 七个维度结果分析。**本节分别对研究对象 TPACK 七个维度对应　　的均值和标准差　行统计分析，探究移动教学对职前教师 TPACK 七个维度的影响。①技术知识 TK 维度平均分 3.73，各　指标得分均大于 3。②学科内容知识 CK 维度均值　在 3 以上，该维度平均分为 3.86。③教学法知识 PK 维度平均得分为 3.86，各　指标得分均在 3 以上。④学科教学知识 PCK 维度平均得分为 3.83，各　指标得分均在 3 以上。⑤整合技术的学科内容知识 TCK 维度平均得分为 3.80，各　指标得分均在 3 以上。⑥整合技术的教学法知识 TPK 维度平均得分为 3.76，各　指标得分均在 3 以上。⑦整合技术的学科教学知识 TPACK 维度平均得分为 3.77，各　指标得分均在 3 以上。

5. 结论

调查结果显示，移动教学有助于职前教师技术认知及 TPACK 七个维度的发展。其中，较显著的是学科内容知识 CK 和与之相关的学科教学法知识 PCK 及整合技术的学科内容知识 TCK 维度。本研究也存在不足。第一，数据样本　择的规模不大，使得研究结果的代表性不突出。第二，没有对职前教师的技术认知与 TPACK 二者之　的关系做探讨。

参考文献

张哲，张海，刘绩宏，王以宁（2015）。 职前教师整合技术的学科教学知识结构研究——基于结构方程模型的实证分析。中国电化教育，8，103-109。

Matthew L. Wilson, Ph.D. a,* , Albert D. Ritzhaupt, Ph.D. b , Li Cheng, Ph.D. c(2020).The impact of teacher education courses for technology integration on pre-service teacher knowledge: A meta-analysis study. Computers and Education,（7）：21-32.-

Michael M. Grant(2019).　Difficulties in defining mobile learning: analysis, design characteristics, and implications. Educational technology research and development ,67:361–388.

基于 ITIAS 的小学数学课堂互动行为研究

——以 2023 年度广东省“双融双创”示范课题为例

A Study of Interactive Behaviour in Primary Mathematics Classrooms Based on ITIAS

--Taking 2023 Guangdong Province "Double Integration and Double Creation"

Demonstration Project as an Example

晏鑫[1]，徐玲[2]
[12]华南师范大学教育信息技术学院
1736966113@qq.com

【摘要】 在当今的教育环境中，课堂互动被认为是提高教学质量和促进学生发展的重要因素。然而，目前课堂互动的分析多依赖于主观经验和观察，缺乏数据支持的科学分析，这使得课堂互动的效果难以准确评估和量化。本研究基于 ITIAS，采用视频分析法和个案研究法，以 2023 年广东省“双融双创”示范课题《分数的加法和减法——喝果汁问题》为研究对象，深入探索课堂互动行为的特征和影响。研究结果显示该课堂师生互动频繁，学生主动参与度高，技术应用有效促进了教学互动。

【关键词】 ITIAS；FIAS；课堂互动行为

Abstract: *In today's educational environment, classroom interaction is recognized as an important factor in improving the quality of teaching and promoting student development. However, the current analysis of classroom interactions mostly relies on subjective experiences and observations, and lacks data-supported scientific analysis, which makes it difficult to accurately assess and quantify the effects of classroom interactions. Based on ITIAS, this study adopts the video analysis method and case study method, and takes "Addition and Subtraction of Fractions - Drinking Juice Problems", a demonstration project of Guangdong Province's "Double Integration and Double Creation" in 2023, as the object of the study to explore the characteristics and impacts of classroom interaction behaviors. influence. The results of the study show that the classroom has frequent teacher-student interactions, students' active participation is high, and the application of technology effectively promotes teaching and learning interactions.*

Keywords: ITIAS; FIAS; Classroom Interactive Behavior

1. 引言

弗兰德斯互动分析系统（FIAS）是美国学者弗兰德斯提出的课堂教学行为分析工具，通过编码系统记录和分析教学行为及师生互动。然而，它忽视了学生的语言表现，无法全面反映无语言行为和信息技术在课堂中的作用。华东师大顾小清教授针对这些问题，改进了 FIAS，提出 ITIAS，增加了反映学生参与、有意义沉寂及信息技术与师生互动的编码，从而更全面地分析课堂互动行为。

2. 研究过程

本研究旨在探讨《分数的加法和减法之喝果汁问题》课堂中师生互动行为特征及其教学效果。为此，我们借助首都师范大学数字化学习实验室开发的编程助手和分析程序，深入剖析了课堂互动过程。通过梳理教学时序列，结合 ITIAS 编码系统的量化分析，我们得出了各类互动行为的比例，并绘制了各参数比率动态特征曲线，直观展示了课堂互动的阶段性特征，为提升教学效果提供了科学依据。

3. 研究结果

课堂结构方面，学生语言行为比例较高，达到 43.20%，显示出学生较高的课堂参与度。教师语言比例相对较低，为 33.46%，但仍保持了必要的引导作用。同时，技术的使用比例较低，仅为 7.49%，表明技术在该课堂中的应用有待加强。教师语言教学风格方面，教师采用了多种影响方式，其中间接影响与直接影响比率为 32.83%，显示出教师在课堂中的多样化教学策略。然而，积极强化与消极强化的比率为 51.61%，说明教师在激励学生的过程中可能过于依赖消极强化手段，需要进一步提高积极强化的使用频率。课堂情感气氛方面，积极整合格的记录次数占总次数的比率高于缺陷格的比率，反映出教师与学生之间情感氛围较为融洽，有助于形成良好的师生互动环境。在学生言语行为特征方面，学生被动应答较多，主动说话和同伴讨论的频次相对较低，这可能与学生参与课堂讨论的方式和教师的提问策略有关。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	合计
1	3	1	0	0	1	0	0	0	0	0	0	0	0	0	5
2	0	8	0	2	5	0	0	0	0	0	0	0	1	0	16
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	12	7	4	0	22	3	0	0	1	0	0	49
5	1	3	0	22	120	6	0	5	3	0	0	1	6	0	167
6	0	0	0	0	1	16	0	8	0	0	0	3	3	0	31
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	4	0	7	22	1	0	171	0	0	0	2	3	1	211
9	1	0	0	1	4	0	0	0	28	0	0	0	0	0	34
10	0	0	0	0	1	0	0	0	0	99	0	1	0	0	101
11	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
12	0	0	0	2	2	2	0	0	0	3	0	116	1	0	126
13	0	0	0	2	4	2	0	5	0	0	0	2	44	0	59
14	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
合计	5	16	0	49	167	31	0	211	34	101	1	126	59	1	801

图 1 课堂互动分析矩阵

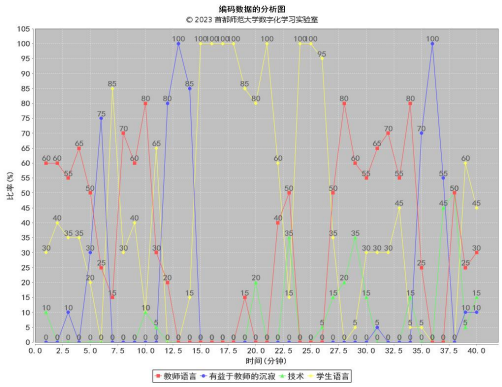


图 2 各参数比率动态特征曲线

4. 结论

首先，本节课师生互动性强，学生参与度高。教师利用青鹿平台实现了多维度的互动，使学生能充分表达观点并获得反馈，有助于深化理解、提升学习效果，同时培养学生的沟通能力和合作精神。其次，技术应用恰当。虽然技术应用比率不高，但教师将其用于实时监测学情和引导学生选择表征方式，有效提升了学生的理解和思考能力。最后，课堂注重学生主体性。教师减少讲授，通过追问和组织汇报交流等方式引导学生主动思考和探究，激发了学生的学习兴趣和探究精神，培养了学生的自主学习能力和创新思维。

参考文献

方海光,孔新梅,郑志宏,等.一种智慧课堂交互行为数据的标准分数常模计算方法研究——基于改进型弗兰德斯互动分析系统 iFIAS[J].远程教育杂志,2023,41(05):67-75.

朱涵,汪甜甜,邓猛,等.培智学校数学课堂教学师生言语行为互动研究——基于改进型弗兰德斯互动分析系统(iFIAS)[J].中国特殊教育,2022(01):39-46.

人文关怀视角下开放教育教学管理的关怀阻滞及疏解路径

Path of care and relief paths of open education and teaching management

under the perspective of humanistic care

谭淇婧
四川开放大学
416293477@qq.com

【摘要】 开放大学是以建设以现代信息技术为支撑、以“互联网＋”为特征的新型高校，然而开放大学教师由于课程管理工作繁杂，缺乏与学生的有效沟通，社会认同较低而在教育工作中出现需求关怀缺失、成长关怀缺失和价值关怀缺失。数智时代的开放教育，需要加强数字赋能与人文关怀的协同作用，开放教育教师需要由多项业务管理向专项能力提升转变，由人际交流能力向数字交流能力转变，由学习主导本位向学习陪伴本位转变。
【关键词】 人文关怀；开放教育；教学管理

Abstract: *Open University is a new type of university characterized by the construction of modern information technology and "Internet +". However, due to the complexity of the course management work and lack of effective communication with students, the low social identity appears in education. Lack of demand care, lack of growth care, and lack of value care. The open education in the era of wisdom needs to strengthen the synergistic role of digital empowerment and humanistic care. Open education teachers need to change from multiple business management to the improvement of special capabilities. Companion with the transformation.*
Key words: Humanistic care，open education，teaching management

1.问题的提出

开放大学作为开放教育的核心载体，立足于建设以现代信息技术为支撑、以“互联网＋”为特征的新型高校，但其发展之初的主要功能还是教育资源“满堂灌式”的传递，学习交互及情景体验感不足（熊军，2023）。开放教育数字化转型的重点是将开放教育从“技术导向”推至“学习者导向”，从过去反向求解技术以适应教育，到未来关注学习者生命周期和学习需求（祝智庭，2023）。数智时代的开放教育，需要利用智能技术不断提升综合能力，加强数字赋能与人文关怀的协同作用，在多样化教学实践中总结经验，促进开放教育向更高层次发展。

2.开放大学教学管理“关怀”阻滞

2.1. 课程管理能力繁杂与需求关怀阻滞

开放教育的主要教学通过“一网一平台”开展，包含学生选课、资源建设、直播教学、课程辅导、论文指导等。由于教育高质量的发展要求，教师工作覆盖程度与学生学习需求精度不平衡的问题日益明显。一是教师课程管理工作内容繁杂，难以针对学生的个性化需求展开有效推进。二是任务式的学习形式不利于激发师生双方营造积极的学习氛围。三是课程资源建设缺乏深度调研，学习内容设计未能适应市场多元竞争。

2.2. 数字交往能力乏力与成长关怀阻滞

开放教育课程教师每月在教学平台上开设一次的教学活动周，根据课程进度发布四个主题贴引导学生交流讨论，然而，近年来，教师与学生之间形成的关怀与沟通联结却日益薄弱。

一是公共服务在教学上的体现逐渐弱化。教学服务重点在于如何顺利的帮助学生完成学业，个体成长服务方面则显得驱动不足。二是开放教育学生功利性的入学目的不利于自身成长。部分学生将获得文凭作为入学的唯一目的，对于与教师的沟通漠不关心，学习的获得感较低。

2.3. 人格塑造能力不足与价值关怀阻滞

开放大学的品牌效应、师生的社会认同与普通高校相比，都处于较低水平，开放大学的学生大部分来自于当前或者曾经在学习当中发展不利的群体，其内心对自我价值实现或多或少存在怀疑或否定。一是开放大学教师长期以来缺乏足够的价值定位自信。二是地域差异的不均衡发展未得到有效缓解。例如，西南地区对开放大学的社会认可跟沿海地区相比，明显处于劣势，学生一方面希望通过开放学习为自己的成功提供助力，一方面又因为社会认可度不高而在学习中产生自我怀疑。

3.开放教育数字化转型中的

3.1. 由多项业务管理向专项能力提升转变，定位需求关怀

开放教育数字化转型要求教师团队优化内部结构，按照课程管理的不同功能进行重组，实现个性需求的最大化满足。根据开放教育的数字化教学特征，教师团队可以划分为课程设计团队、学习辅导团队、课程服务团队及技术研发团队。课程设计团队致力于根据市场趋势和学生需求确定课程的框架及核心内容；学习辅导团队重点在于提升在线授课技术及相关课业辅导；课程服务团队的主要职责在于加强与学生的沟通联系，确保学习过程中的良好学习体验；技术研发团队则依托人工人工智能技术构建虚拟学习社区、虚拟实验室辅助在线教学，同时，加强数字化研究和分析功能，以数字分析赋能其他团队的实施策略。

3.2. 由人际交流能力向数字交流能力转变，重视成长关怀

在教师团队结构优化，专项能力达到数字化转型发展要求时，各团队教师都要具备与数字交流的能力，从各项数据分析中读懂学生的成长状态，明确教学对策方向。例如课程设计教师要从调研数据中明确课程目标并设计学习活动；学习辅导团队要从学生课堂分析数据追踪学习体验，优化教学方式；课程服务团队要根据学生行为数据强化学习支持服务，对行为异常的学生采取干预措施；数字分析团队更需要提升教育技术来赋能智慧化学习。尽管各团队教师分工不同，但都要具备跟数字交流的能力，使学生更加关注学习本身，提升内在学习动力。除此之外，开放教育教师还要掌握技能迭代更新的能力，保持自身在数字化发展中的先进性。

3.3. 由学习主导本位向学习陪伴本位转变，强调价值关怀

多元智力理论阐述了人的发展的不平衡性，为人的多元发展提供了理论基础，教师要转变过去主导学习的思想，并充分尊重个体的自主学习，让学生根据自身情况选择符合实际发展需要的课程，培养学习的自信心和主动发展的能力。一是通过学习者画像来精准定位学习需求，教师团队不再是学习的主导者，而是作为学习的陪伴着跟随学生共同发展，学生在获得自我肯定和收获的同时，教师也体会到自我价值的实现。二是采取基于平台的合作学习模式，在一体化平台的数字生态中，教师和学生均作为系统的用户开展多元交互，在解决学生问题的同时，也提高了教师教学水平。

参考文献

熊军（2023）。数字化转型和智能化升级：“互联网 +”新型大学建设研究。**中国教育信息化，29（02），**64-78。

祝智庭（2023）。开放教育数字化转型与融创发展：新议题、新思维、新行动。**中国远程教育，（6），**19-28。

数智赋能“做最好的自己”教师研修生态的构建研究

Research on the Construction of a “Be Your Best Self” Teacher Training Ecosystem Empowered by Digital Intelligence

刘伟[1*], 钟芳[1], 彭操红[1]
[1]广州市天河区体育东路小学兴国学校
[*]105172120@qq.com

【摘要】 教师是教育发展的第一资源。本研究针对集团校教师队伍流动性大、年轻化、专业水平参差不齐等问题，以教育生态学理论等为指导，构建数智赋能“做最好的自己”教师研修生态，并以体育东路小学兴国学校为例进行实践。结果表明研修生态能有效提高教师专业化发展水平，助推集团化办学高质量发展。
【关键词】 数智赋能；教师研修生态；做最好的自己；多元协同

Abstract: *Teachers are the primary resource for educational development. Addressing issues of high turnover, youth and uneven professional levels in group school, this study, guided by educational ecology theory, constructs a “Be the Best Self” teacher training ecology empowered by digital intelligence. Practice has shown that it effectively enhances teachers’ professional development and promotes high-quality, group-based educational development.*
Keywords: digital intelligence empowerment, Teacher training ecology, Be the best self, Multivariate collaboration

1. 问题的提出

教师是教育发展的第一资源。集团校面临着教师队伍流动性大、年轻化、专业水平不一等问题，急需融合数字化、网络化、智能化技术，构建全场域教师研修生态，加快培养高素质教师，成就教师“做最好的自己”。基于此，研究以教育生态学等理论为指导，构建数智赋能“做最好的自己”教师研修生态，以期为新时代教师研修提供理论参考与实践借鉴。

2. 相关研究述评

现有教育生态研究聚焦教育系统中各因子相互作用路径及效果，如胡洁雯(2013)构建了教育生态学视角下的教师发展机制。技术的发展使教育要素乃至结构发生变革，教师亟需提升专业能力。数智赋能教师研修研究聚焦于资源供给和模式构建，如吴南中等(2024)构建了人工智能环境下有组织教研模式。因此，需探索如何利用数智技术构建良好的教师研修生态。

3. 数智赋能“做最好的自己”教师研修生态构建

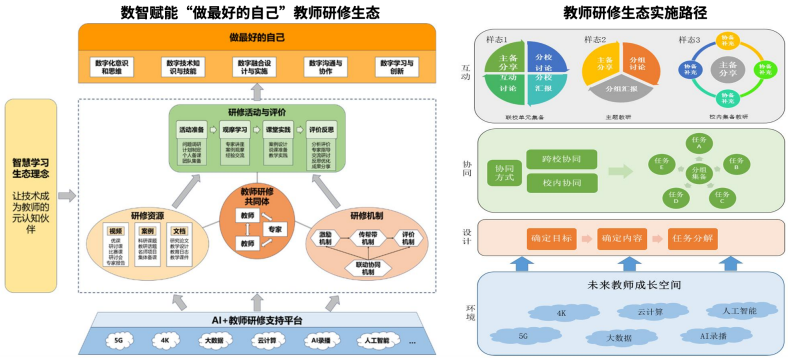
3.1. 理论基础与核心理念

教育生态学理论关注教育中各生态因子的相互作用，以促进生态平衡发展；活动理论认为个体与环境、共同体等交流互动是发展的关键，二者为研修生态的构建提供了理论指导。基于此，本研究提出数智赋能“做最好的自己”教师研修生态核心理念，即基于互联网、人工智能等技术支撑的移动互联环境和“信任、融洽、互动、开放”新型师生关系的人文环境，让技术成为教师发展的元认知伙伴，创新教学与教研，促进教师自我认知与提升。

3.2. 数智赋能“做最好的自己”教师研修生态构建

本研究基于核心理念，构建出数智赋能“做最好的自己”教师研修生态和实施路径，如图 1 所示。该生态依托 AI+教师研修支持平台，基于研修资源与研修机制的支持，让教师在不同样态的研修共同体中，按照“活动准备-观摩学习-课堂实践-评价反思”的流程参与研修活动与评价，提升教师数字素养，数智赋能教师“做最好的自己”。

图 1 数智赋能“做最好的自己”教师研修生态与实施路径



4. 数智赋能“做最好的自己”教师研修生态实践

本研究以体育东路小学兴国学校为例，从空间、资源、活动、评价和机制等方面开展生态实践。学校打造未来教师成长空间，建设共享优质研修资源；形成跨校教师研修共同体，遵循“设计-协同-互动”方式实施多样态网络教研；构建数智赋能以学为中心的共同体深度教研模式，通过“确定目标-深度调研-拟定方案-深度观察-深度反思”流程深化教研实践，如图 2 所示；基于大数据看板，引入第三方 AI 分析工具形成教学分析报告，实现精准教学反思、指导与帮扶；建立“校-企-研-院”多方协同机制、集团联动机制等，助力研修高质量开展。

图 2 数智赋能以学为中心的共同体深度教研模式



5. 数智赋能“做最好的自己”教师研修生态效果

通过生态的实践，有效解决了集团校教师队伍发展难题，培育了大量优秀骨干教师。共培育特级教师 2 名、正高级教师 2 名、高级教师 19 名，其中市级以上名师 7 人、市级以上骨干教师 14 名，并建设有市级以上名师工作室 4 个，打造各级各类优课百余节。学校还在教学竞赛、课题研究等方面取得丰硕成果，推动了教师队伍建设和集团化办学高质量发展。本研究针对集团校教师队伍建设问题，以教育生态学理论等为指导，构建数智赋能“做最好的自己”教师研修生态，并依托体育东路小学兴国学校开展实践。结果表明，该生态能够有效提高教师专业化发展水平，助推集团化办学高质量发展。

参考文献

吴南中、陈恩伦和吴云(2024)。有组织教研：人工智能环境下教师研修的有序进化。电化教育研究，(01)，122-128。

胡洁雯(2013)。教育生态学视域下的教师发展研究。教学与管理，(03)，52-54。

澳门中学教师人工智能素养现状、分析及发展

The Current Situation, Analysis and Development of Artificial Intelligence Literacy among Secondary School Teachers in Macau

凌智荣 [1]
* chiwengamos@qq.com

【摘要】本研究探讨了信息化环境下澳门人工智能教育的改革和发展。通过 212 名澳门中学教师的调查研究来分析和理解他们的人工智能素养和应用能力。

【关键词】澳门教师 人工智能 信息素养 影响因素

Abstract: *This research investigates the transformation and progress of artificial intelligence (AI) education in Macao within the framework of an informationized environment. Through an investigation involving 212 middle school teachers from Macao, the study seeks to dissect and comprehend their AI-related literacy and application skills.*

Keywords: Macau teachers, Artificial intelligence, Information literacy, Influencing factors

1. 研究背景与问题提出

本研究分析了澳门中学教师人工智能信息素养现状，探讨了人工智能技术在澳门中学教师中的积极作用以及所面临的挑战。

1.1. 研究问题

研究旨在回答两个核心问题：（1）澳门教师人工智能素养现状如何；（2）如何培养和提升教师的生成式人工智能技术能力。通过探讨这些关键问题，希望普及澳门人工智能教育发展，提升澳门教师使用人工智能教学能力。

2. 研究设计

本研究采用问卷调查法对澳门中学教师人工智能素养现状进行调查，信效度较高 (Cronbach α-0.963; KMO-0.928)。

澳门教师人工智能素养现状信度分析

变量	Cronbach's Alpha	基于标准化项的 Cronbach's Alpha	项数	总体 Cronbach's Alpha
智能意识	0.905	0.905	5	0.963
智能知识	0.942	0.942	4	
智能能力	0.948	0.948	9	
智能道德	0.863	0.863	3	

3. 结果与分析

3.1. 澳门中学教师人工智能素养现状

由图 1 得知，澳门中学教师人工智能素养得分为 3.74，可见目前澳门中学教师的人工智能素养处于中等水平。在各维度中，澳门教师在智能意识得分为 4.02，智能道德为 3.98，得分相对较高。智能知识得分为 3.46 和智能能力为 3.90，反映出部分教师在人工智能知识方面有所欠缺。性别和学历间不存在显著差异，但

是从年龄来看年轻教师在智能能力得分较高，51 及以上的教师得分较低。在智能道德方面，51 及以上的教师具有良好的道德水平。

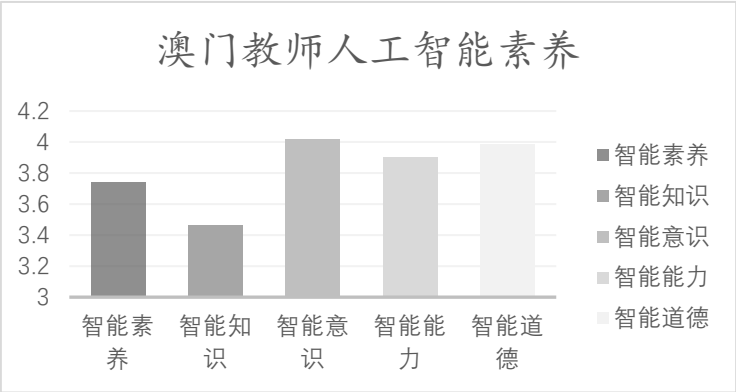


图 1 澳门教师人工智能信息素养的影响因素及分析

3.2. 澳门不同中学的教师智能素养存在差异

表 1 不同性别的受访者在智能素养上的差异性分析(M±SD)

变量	男性	女性	t	P
智能素养	3.76±0.77	3.72±0.78	0.360	0.749
智能意识	4.04±0.97	4.00±0.92	0.549	0.922
智能知识	3.45±0.99	3.47±0.96	1.246	0.765
智能能力	3.89±0.96	3.91±0.99	0.965	0.433
智能道德	4.00±1.07	3.96±0.99	0.662	0.831

表 2 不同年龄的受访者在智能素养上的差异性分析(M±SD)

变量	18-25	26-30	31-40	41-50	51 及以上	*F*	*P*
智能素养	3.74±0.70	3.60±0.76	3.68±0.86	3.98±0.74	3.70±0.73	0.288	0.260
智能意识	4.20±0.94	3.99±0.90	4.06±1.00	3.83±0.89	4.02±0.93	0.627	0.821
智能知识	3.39±1.00	3.09±0.91	3.26±1.03	3.41±0.93	3.33±0.96	0.359	0.345
智能能力	3.60±0.85	3.55±0.90	3.48±1.02	3.50±0.99	3.17±0.94	5.323	0.031
智能道德	3.90±1.08	3.85±1.11	3.88±1.22	3.95±1.18	4.32±1.16	3.569	0.002

表 3 不同学历的受访者在智能素养上的差异性分析(M±SD)

变量	学士	硕士	博士	*F*	*P*
智能素养	3.72±0.79	3.74±0.77	3.76±0.75	0.735	0.446
智能意识	4.00±0.92	4.02±0.96	4.04±0.94	1.487	0.753
智能知识	3.42±0.99	3.46±0.94	3.50±1.03	1.309	0.439
智能能力	3.90±0.95	3.88±1.01	3.92±1.00	0.574	0.858
智能道德	3.94±0.95	3.98±1.01	4.02±1.00	0.896	0.321

4. 建议

澳门教青局可通过举办研讨培训、与技术公司合作开发定制 AI 教学工具、实施分阶段专业发展课程等措施，提高教师对信息化政策的理解，增强人工智能应用能力，并促进教师间的互助学习，以培养适应未来教育挑战的师资队伍。

LLMs 在英语写作学习的应用：设计、实施和效果分析

The Application of LLMs in English Writing Learning: Design, Implementation, and Effect Analysis

郑伊彬[1*]，胡笑笑[2]，叶新东[1]
[1]温州大学 教师教育学院
[2]温州肯恩大学 人文学院
[*] yibinzyb0806@gmail.com

【摘要】 本研究旨在调查整合大语言模型（LLMs）的英语写作课程在以英语为外语写作中的教学效果。通过设计整合大语言模型的英语写作课程，并进行了随机对照实验，将 65 名小学生分为两组，一组接受传统的 SBI 英语写作课程，一组接受整合 LLMs 的 SBI 课程。经过了为期 5 周，共 10 节课的干预，结果显示 LLMs 组在英语写作能力和 SRL 策略使用频率上从干预前到干预后都取得了显著进步，并且在写作能力、文本生成策略和自我监控策略方面的提升显著优于对照组。以上结果证明了大语言模型在英语写作领域的应用潜力。
【关键词】 大语言模型；英语写作；基于策略的教学；EFL 小学生

Abstract: *This study aimed to investigate the instructional impact of integrating Large Language Models (LLMs) into an English writing course for English as a Foreign Language (EFL) learners. Through the design of an English writing program incorporating LLMs, a randomized controlled trial involved 65 elementary students divided into two groups: one receiving traditional SBI English writing classes and the other receiving SBI classes integrating LLMs. Following a 5-week, 10-session intervention, results indicated significant improvements in EFL writing proficiency and SRL strategy utilization within the LLMs-integrated group compared to the control, specifically in writing proficiency, text generation, and self-monitoring strategies. These outcomes signify the promising potential of Large Language Models in the realm of English writing.*

Keywords: Large Language Models; English Writing; Strategy-Based Instruction; EFL Elementary Students

基金项目：温州大学 2024 年硕士研究生创新基金自筹基金项目“自我调节策略支持下的大语言模型对小学生英语写作能力的应用研究”（课题编号：3162024004024）。

1. 引言

学习英语通常分为四个部分：听、说、读、写，而对于英语作为外语（EFL）的学习者来说，写作往往被是视为是所有语言学习中最难的一项(Zhang & Guo, 2012)。这是由于写作是一个复杂而多维的过程，EFL 学习者需要应对语言和语句的复杂性，同时应对英语写作的修辞和组织挑战(Sun & Wang, 2020)。基于策略的教学干预（SBI）是解决这些问题的一个关注点。虽然基于策略的教学干预能在一定程度上提高 EFL 学习者的英语写作能力，但存在个性化需求难以满足和反馈不及时等问题。近年来，随着人工智能（AI）的发展，尤其是像 ChatGPT、GPT-4 等大语言模型（Larger Language Models, LLMs）的问世，为解决这些问题提供了新途径。LLMs 能够根据用户的个性化需求生成高质量内容，并提供实时的上下文反馈(Wu et al., 2023)。这使得它们能更有效地适应学生的个性差异和学习方式，推动个性化教育的发展。

在 SBI 英语写作课程中，LLMs 可以充当个性化支架，与传统的教学方式相辅相成。LLMs 能够根据学习者的差异和写作需求提供创造性的反馈和建议，有助于学习者组织思维、丰富修辞表达和提供词汇选择，并及时给予反馈。将 LLMs 整合于 SBI 课程为克服传统评价、反馈和个性化教学所面临的挑战提供了新的解决途径，为 EFL 英语写作教育提供更高效和个性化的学习体验。虽然 LLMs 应用于教学中具有理论支持，但将其融入课堂的实证研究非常有限。因此，本研究设计了基于 SBI 的两种课程：一种是传统的基于 SBI 英语写作课程，另一种是整合 LLMs 的 SBI 英语写作课程。通过随机对照实验，旨在评估整合 LLMs 的 SBI 英语写作课程对小学生英语写作能力和自我调节策略使用频率的影响。

2. 文献综述

2.1. 自我调节策略（SRL）在 EFL 写作中的作用

自我调节策略（Self-Regulated Learning, SRL）是指学习者在学习过程中进行计划、设定目标、组织、自我监控以及自我评价(Zimmerman, 2000)。目前 SRL 已成为提高 EFL 英语写作能力的重要方法。写作的过程主要包括三个主要阶段：计划、写作和回顾(Flower & Hayes, 1981)，SRL 策略表现为贯穿于英语写作的整个过程，这些策略包括但不限于:计划、组织、监控、寻求帮助和资源、修订(Bai, 2015)。在写作前的计划阶段，学习者可以使用 SRL 策略设定具体的目标；在写作阶段，学习者可以使用 SRL 策略积极调整写作过程，确保与初始目标一致；在写作后的回顾阶段，学习者使用 SRL 策略对写作内容进行修改和寻求反馈。二语研究者指出，EFL 学习者的英语写作能力和自我效能感与运用 SRL 策略部署写作过程之间存在显著相关性(Sun & Wang, 2020)。即使用 SRL 策略频率越高，其英语写作能力以及自我效能感增长越显著。这些研究表明了 SRL 策略教学在提升 ESL/EFL 学习者写作能力方面的重要性。而基于策略的教学（SBI）是 SRL 写作策略教学的典型方法。

2.2. 基于策略的英语写作教学模式

语言学习策略是语言学习者为了提高学习效率，增强语言能力，达到学习目标，导向独立自主学习所采取的有意识活动(Oxford, 1990)。在基于策略的教学（Strategy-based instruction, SBI）中，教师和学习者共同努力，将策略整合到常规的语言课程中(Gu, 2007)。已有研究证明学习者的策略行为和他们的语言学习结果之间存在积极的关系(Grenfell et al., 2007)。目前已开发和评估了多种权威的基于策略的教学框架，如认知学术语言学习法（CALLA）、自主学习策略发展（SRSD）模型和战略教学模型（SIM）(Chamot, 2005; Deshler & Lenz, 1989; Graham & Harris, 2003)。

然而，传统的 SBI 课程在不同课堂的实施中面临着一些挑战。SBI 课程难以充分满足不同学习者的个性化需求，会受到教师自身专业知识的影响，并且传统的 SBI 课堂难以提供语言学习中至关重要的及时有效的反馈(Lim et al., 2021)，无法满足语言学习过程中学习者对于迅速纠正和改进的需求。鉴于这些局限性，越来越多的学者开始关注使用技术，特别是人工智能和大语言模型，来应对这些挑战。LLMs 的个性化实施反馈能力为未来的 SBI 课程研究和应用提供了新的途径。

2.3. 大语言模型（LLMs）以及在教育领域的作用

大语言模型（Large Language Models, LLMs）是近年来在自然语言处理（NLP）方面取得重大突破的技术。这些模型通过海量的文本数据训练，能够生成类似人类的文本、回答问题并高精度地完成其他与语言相关的任务(Kasneci et al., 2023)。LLMs 的代表性模型包括 GPT-3、GPT-4。GPT-4 在 NLP 的准确性和效率方面取得了进一步的提高，展现出了非凡的语言能力为教育领域提供了前景(Kohnke et al., 2023b)。它可以用作课堂辅助工具，提供问题的快速解答，也可以用作辅助文献综述和文章大纲的学习工具(Kasneci et al., 2023)。Kohnke et al. (2023a)

强调了 LLMs 能够调整对话复杂性并生成扩展资源的能力，从而为学习提供了适应性支持，它们提供实时反馈并进行有意义的个性化互动能力，能够克服传统课堂环境中存在的许多限制。

类似于维果斯基的脚手架概念(Vygotskij, 1979)，LLMs 能够在学习过程中充当动态的脚手架，为不同需求和水平的学习者提供个性化的写作辅助和反馈。LLMs 可以通过提供个性化的建模、实践和反馈优化 SBI 英语写作课程。然而，LLMs 的实际影响、它们如何融入到传统的教学法以及对教育环境的影响仍然需要进一步的实证探索。

3. 研究问题与假设

本研究设计基于一种整合 LLMs 的 SBI 英语写作课程，并进行了一项随机对照实验，旨在评估整合 LLMs 的 SBI 英语写作课程对小学生英语写作能力以及自我调节策略使用频率的影响。研究问题与假设如下：

问题 1：整合 LLMs 的 SBI 英语写作课程相较于传统的 SBI 英语写作课程能否更有效的提高小学生英语写作能力？

问题 2：整合 LLMs 的 SBI 英语写作课程相较于传统的 SBI 英语写作课程能否更有效的提高小学生自我调节策略使用频率？

假设 1：整合 LLMs 的 SBI 英语写作课程相较于传统的 SBI 英语写作课程能够更有效的提高小学生英语写作能力。

假设 2：整合 LLMs 的 SBI 英语写作课程相较于传统的 SBI 英语写作课程能够更有效的提高小学生自我调节策略使用频率。

4. 研究方法

4.1. 被试与实验设计

参与者是来自温州某小学六年级的 65 名学生，年龄分布在 11-13 岁之间。实验参与者被随机分为两组。实验组 32 名（18 名女孩，平均年龄=11.5 ± 0.57）学生接受整合 LLMs 的 SBI 英语写作课程，对照组 33 名（19 名女孩，平均年龄=11.3 ± 0.48）学生接受传统的 SBI 英语写作课程。研究方案的制定经过了学校校长和所在年级多位教师的参与和批准。纳入标准是视力正常或矫正为正常，认知和运动功能正常。研究人员亲自联系了孩子的父母或监护人，介绍并解释了研究方案，同意孩子参加的父母或监护人签署了知情同意书。

4.2. SBI-LLMs 教学设计

本研究中的 SBI 课程是基于面向过程的写作指导框架开发的，包括规划、起草、修改的递归阶段(Flower & Hayes, 1981)。课程改编自 Bai (2015)为新加坡公立学校小学五年级学生设计的 9 节 SBI 课程，在实验组集成了 LLMs 的功能以提供个性化支架。

每一节课的支架改编自 Chamot and O'malley (1994)的模型，该模型包括五个阶段：准备阶段、呈现阶段、实践阶段、评价阶段以及扩展阶段，LLMs 的功能主要集成实践阶段、评价阶段以及扩展阶段。这种集成了 LLMs 的 SBI 课程被称为 SBI-LLMs 课程。本课程计划包括为期 5 周、共 10 节课的课程，每周两节课，每节课都与写作课程的一个阶段相关。在整干预过程中教授四组 SRL 写作策略，即规划、文本生成、修订以及自我监控策略，四组写作策略被嵌入到课程中（见表 1）。在对照组每节课的教学步骤遵循 Chamot (1995)的 CALLA 模型，课程包含了同样 SRL 写作策略。

表 1 SBI-LLMs 课程

课	主题	LLMs 的功能	目标策略

程		实践阶段	评价阶段	扩展阶段	
1	设定目标	作为学生获取灵感和想法的来源，帮助学生在设定写作目标时扩展视野和引发思考，从而更有针对性地确定个人的写作目标	利用 LLMs 检索相关内容，评估其与所设定的写作目标的相关性和可行性，从而调整和完善目标	利用 LLMs 拓展思路，以丰富和扩展写作目标的内容	规划 自我监控
2	规划内容	提供内容引导和创作建议；帮助学生理清思路，初步规划内容组织和结构	LLMs 评估和审查学生规划内容的质量和相关性	利用 LLMs 获取更多信息、观点和细节，以丰富写作内容	规划 自我监控
3	评估受众	利用 LLMs 搜寻相关信息，初步了解不同受众群体的需求，以便更好地针对受众撰写内容	LLMs 提供的信息和反馈，调整和优化写作内容，使其更符合目标受众的需求	根据 LLMs 提供的内容，更精准地为不同受众定制内容	规划 自我监控
4	获取灵感——阅读相关内容	搜索不同主题的相关资料和观点，初步了解相关领域的内容和见解	作为学生信息质量评估的工具	通过 LLMs 获取的信息和语言参考，能够在其他写作中灵活应用	规划 自我监控
5	词汇的重要性——关注词汇	提供词汇积累和实践支持。学生可以通过 LLMs 搜索、学习和应用新的词汇，积累丰富的词汇库，并在写作过程中尝试使用这些词汇	可用于评估词汇的准确性并提供建议和修正	促进词汇的迁移和应用	文本生成 自我监控
6	结构的重要性——关注语法结构	引入不同的语法结构，解释其在写作中的重要性和作用	学生利用 LLMs 进行语法错误的检查和修正，获得即时的语法反馈和建议	学生利用 LLMs 进行自主学习和探索，加深对语法结构的理解，提高语言表达水平	文本生成 自我监控
7	组合——编排文章	帮助学生理解和应用不同的篇章结构；学生利用 LLMs 进行文章组织和排列的练习	学生从 LLMs 获得反馈和建议，指导他们改进篇章结构，提高文章的连贯性和逻辑性	LLMs 作为参考工具，帮助他们在不同的写作项目中灵活运用不同的篇章结构	文本生成 自我监控
8	修改文章	学生通过 LLMs 进行修订和编辑的练习，学习如何修改句子、改进表达、提升文章逻辑性和连贯性	LLMs 用于提供修改建议和反馈。学生利用 LLMs 的反馈，了解文章的问题所在，并获取改进建议，指导他们进行文章修订	LLMs 作为自主学习的工具，鼓励学生自主进行文章修订和改进	修订 自我监控

9	获取并回应反馈	通过LLMs提供的反馈指导，对自己的写作进行调整和改进。	LLMs帮助学生比较修改前后的文章，分析改进之处。	从LLMs反馈中吸取经验，不断改进自己的写作技巧	修订 自我监控
10	掌握所学策略的应用	通过LLMs的支持和示范，将学到的策略融入到具体的写作实践中	学生评估LLMs帮助他们应用策略的效果。他们通过比较应用策略前后的写作成果，分析策略应用的实际效果	学生利用LLMs在不同的写作场景中灵活运用所掌握的策略	规划 文本生成 修订 自我监控

4.3. 方法

4.3.1. 大语言模型平台

本研究开发了一个专门的 Web 应用程序，利用开源代码，通过 OpenAI 的 API 实现了与 LLMs 的互动。在 API 中，使用“系统角色”来限制 LLMs 的行为。教师有权输入或切换提示语，以确保 LLMs 的反馈和指导与当前课程阶段相关。同时，学生在特定课堂阶段能够自由与 LLMs 互动。

4.3.2. 写作能力

学生的写作能力是通过其作文成绩来衡量的。对作文成绩的评分采用了 Hyland (2003)提出的分析方法，这种方法因其对写作要素的全面描述而被广泛采用。评分主要包括三个方面：句子结构和词汇（40 分）；组织和逻辑连贯性（20 分）；以及格式和内容（40 分）。

为了确保可靠和客观的评估，两位具有丰富英语教学经验的高级小学教育工作者被委托负责评分工作。在实际评估之前，组织了一次培训会议，以使他们对评分标准有一致的理解，并校准其评分技巧。每篇作文由两位评分人独立评分，每篇作文的最终得分代表了两位评分人给出的分数的平均值。评分者间的一致性很强，Spearman's rho = 0.88, $p < .001$。”

4.3.3. SRL 写作策略使用量表

使用了 Bai et al. (2022)提出的写作策略量表，用于评估参与者在 EFL 写作过程中使用的 SRL 策略。该量表包含 17 个陈述项目，采用了李克特 5 分评分方案，从“从不”到“总是”进行评分，分数越高表示使用相应的 SRL 写作策略越频繁。这些项目涵盖了四个方面的策略：规划策略，文本生成策略，修订策略以及自我监控策略。该量表在四个维度上的 Cronbach's alpha 值分别为 0.85、0.88、0.79 和 0.82。

4.4. 实验流程

本研究在英语写作课中开展，参与者来自同年级的多个班级，通过随机分组被分配到 LLMs 组或对照组。为了确保客观性，在整个干预期间，同一受过培训的教师对 LLMs 组和对照组提供指导。实验流程如图 2 所示。在实验准备与前测阶段，被试指导如何操作系统之后签署知情同意书并完成人口统计学问卷，随后完成英语写作能力前测以及 SRL 使用频率量表。在接下去的五周时间里，所有被试都参与了 10 节 SBI 课程。LLMs 组接受整合 LLMs 的 SBI 英语写作课程指导，而对照组则接受传统 SBI 英语写作课程指导。在五周干预之后，被试完成英语写作能力后测以及 SRL 使用频率量表。

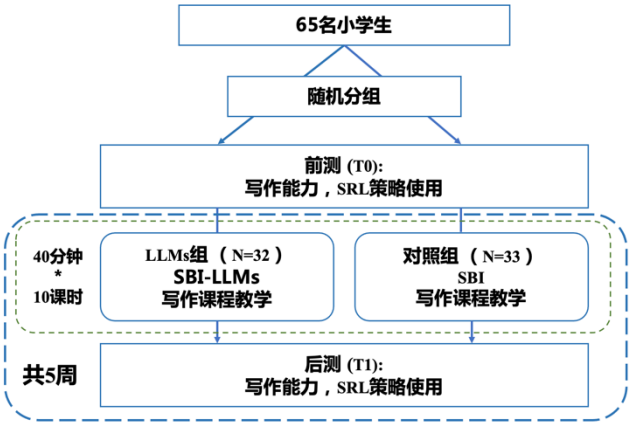


图 1 实验流程图

5. 结果分析

为比较实验组和对照组干预前的英语写作能力、SRL 策略使用频率是否存在差异，以组别为因变量，前测分数为自变量进行独立样本 *t* 检验（见表 2）。结果发现不同组别样本对于英语写作能力、规划策略、文本生成策略、修订策略以及自我监控策略前测分数均没有显著差异，因此可以排除组别对研究结果的影响。

表 2 写作能力、SRL 策略使用频率前测独立样本 t 检验分析

Dimension	Group	N	Mean	SD	t	p
写作能力	实验组	32	63.05	9.78	-0.36	0.723
	对照组	33	62.12	11.12		
规划策略	实验组	32	3.06	0.51	0.39	0.698
	对照组	33	3.11	0.55		
文本生成策略	实验组	32	3.08	0.65	-0.44	0.661
	对照组	33	3.01	0.64		
修订策略	实验组	32	3.19	0.55	0.45	0.655
	对照组	33	3.25	0.57		
自我监控策略	实验组	32	3.28	0.45	0.48	0.633
	对照组	33	3.34	0.53		

*p<0.05 **p<0.01

5.1. 英语写作能力

在干预后，对实验组和对照组的后测英语写作成绩分别进行了配对样本 t 检验（见表 5），结果显示实验组的后测英语写作成绩（M=72.44）显著高于实验组前测（M=63.05），p=0.000，对照组的后测英语写作成绩（M=67.35）也显著高于前测成绩（M=62.12），p=0.001。这一结果表明，无论是传统的 SBI 课程还是整合 LLMs 的 SBI 课程能够提高学生的写作能力。

为了对比 LLMs 对学生英语写作能力的影响，对实验组和对照组的英语写作成绩后测结果进行分析。ANOVA 分析结果显示（见表 3），实验组（M=72.44）的成绩显著优于和对照组（M=67.35），F（1,63）=4.57，p=0.036，p<0.05。这一结果证明，整合 LLMs 的 SBI 写作课程在提高学生写作能力比传统 SBI 写作课程更有帮助。

表 3 英语写作成绩后测 ANOVA 分析

Dimension	Group	Mean	SD	F	p
写作成绩	实验组	72.44	10.74	4.57	0.036*
	对照组	67.35	8.33		

*p<0.05 **p<0.01

5.2. SRL 策略使用频率

为了对比 LLMs 对学生 SRL 策略频率的影响，对实验组和对照组的 SRL 使用频率的四个维度结果进行分析。ANOVA 分析结果显示（见表 4），实验组在文本生成策略使用频率（M=3.49）显著优于对照组（M=3.14），F（1,63）=5.93，p=0.018；自我监控策略使用频率（M=3.75）显著优于对照组（M=3.5），F（1,63）=5.93，p=0.047。这一结果证明，整合 LLMs 的 SBI 课程在提高学生文本生成策略以及自我监控策略使用频率比传统 SBI 写作课程更有帮助。

表 4 SRL 使用频率后测 ANOVA 分析

Dimension	Group	Mean	SD	F	p
规划策略	实验组	3.57	0.49	3.83	0.055
	对照组	3.32	0.52		
文本生成策略	实验组	3.49	0.59	5.93	0.018*
	对照组	3.14	0.56		
修订策略	实验组	3.66	0.48	2.73	0.104
	对照组	3.45	0.54		
自我监控策略	实验组	3.75	0.48	4.10	0.047*
	对照组	3.51	0.48		

*p<0.05 **p<0.01

规划策略和修订策略在实验组和对照组后测之间无显著差异，但是使用配对样本 t 检验对两组分别前后测进行比较（见表 5），结果发现实验组在规划策略（p=0.000）、文本生成策略（p=0.000）、修订策略（p=0.000）以及自我监控策略（p=0.000）后测成绩都显著高于前测成绩。对照组在规划策略（p=0.021）、修订策略（p=0.005）以及自我监控策略（p=0.046）后测都显著高于前测成绩。这一结果证明，无论是传统的 SBI 课程还是整合 LLMs 的 SBI 课程在提高学生 SRL 使用频率上都有帮助。

表 5 写作成绩和 SRL 使用频率前、后测配对样本 t 检验分析

Dimension	Group	Mean（SD）		t	p
		Pre-test	Post-test		
写作成绩	实验组	63.05（9.78）	72.44（10.74）	-6.08	0.000**
	对照组	62.12（11.12）	67.35（8.33）	-3.76	0.001**
规划策略	实验组	3.06（0.51）	3.57（0.49）	-6.96	0.000**
	对照组	3.11（0.55）	3.33（0.52）	-2.44	0.021*
文本生成策略	实验组	3.08（0.65）	3.49（0.59）	-5.40	0.000**
	对照组	3.01（0.64）	3.14（0.56）	-1.74	0.092
修订策略	实验组	3.19（0.55）	3.66（0.48）	-6.62	0.000**
	对照组	3.25（0.54）	3.45（0.54）	-2.99	0.005**
自我监控策略	实验组	3.28（0.45）	3.75（0.48）	-7.00	0.000**

对照组	3.34（0.53）	3.51（0.48）	-2.08	0.046*

*p<0.05 **p<0.01

6. 结果讨论

6.1. 主要发现

本研究设计基于 SBI 的两种课程：一种为传统的基于 SBI 英语写作课程，另一种为整合 LLMs 的 SBI 英语写作课程。通过随机对照实验方法，旨在探讨整合 LLMs 的 SBI 英语写作课程对小学生英语写作能力、SRL 策略使用频率的影响。研究结果证实了整合 LLMs 的 SBI 英语写作课程相较于传统的 SBI 英语写作课程能够更有效的提高小学生英语写作能力。这可能是由于整合 LLMs 的 SBI 课程中所提供的个性化脚手架与维果斯基的最近发展区（ZPD）理论非常吻合。Vygotskij (1979)认为 ZPD 是学习者学习者可以独立完成的事情与他们在帮助下可以完成的事情之间的差距。在 SBI-LLMs 课程中 LLMs 充当一个“知识渊博的指导者”，为每个学生提供个性化的支持和指导，填补了学习者自主能力和有指导帮助时能够完成任务的差距。这种个性化支架和指导有助于学习者更有效地提高英语写作能力。此外，LLMs 提供的及时反馈与 Hattie and Timperley (2007)提出的反馈作用吻合。Hattie and Timperley (2007)认为反馈是语言学习的关键因素。它可以促进最低程度的学习，也可以促进深度学习。SBI-LLMs 课程中 LLMs 提供及时的反馈促进了学习者的深度学习,从而更有效地提高了学习者的英语写作能力。

此外，SRL 策略使用频率的提高符合自我调节学习理论的原则。通过写作策略的应用，学生在写作过程中的自我调节能力得到提升(Zimmerman, 2000)。在本研究中，LLMs 对写作过程中策略的建模，增强了学生在写作过程中自我调节策略的使用，提高了写作能力。

6.2. 局限和未来方向

虽然这项研究带来了有价值的初步成果，但其存在一些限制需要我们注意。首先，研究的样本范围仅覆盖了单一年级的学生，这可能对研究结果的普适性产生一定的限制。同时，干预的时间相对较短（仅为五周），可能无法充分展示基于 LLMs 的教学方法在长期和持续性方面的效果。考虑到这些限制，未来的研究可以考虑使用更大、更多样化的样本，涵盖不同年龄组、学习背景，甚至扩展到英语教育之外的其他学科，以更全面地了解 LLMs 教学方法的影响。另外，本研究的重点仅局限于写作策略指导方面。未来的研究可以拓展视野，探索将 LLMs 支持整合到其他语言学习领域，例如阅读、听力和口语等方面。这种扩展或许能够提供更广泛和全面的认识，揭示 LLMs 在语言教育中的多方面潜力。

更进一步，未来的研究或许可以采取纵向研究设计，以深入探究 LLMs 教学对学生写作表现和自我调节能力的长期影响。这种长期跟踪或许能够更清晰地展现教学方法在学生学习过程中的长期影响和变化。这样的研究设计将有助于全面理解 LLMs 教学方法对学习者的影响，并为未来的教学提供更为深入和持久的指导。

7. 结语

本研究旨在探究 LLMs 在英语写作课程的应用，并揭示了与传统的 SBI 课程相比，整合 LLMs 的 SBI 课程在提高学生写作能力和 SRL 策略使用上表现为更积极效果。LLMs 作为个性化的学习支持工具，不仅填补了学生自主学习能力和有指导帮助下完成任务之间的差距，而且为学生提供了个性化的、及时的支持和指导。尽管初步的研究结果证明了 LLMs 在语言教育方面的潜力，但研究仍处于初级阶段，需要持续深入的实证研究作为更加有力的支撑。

参考文献

Bai, B., Wang, J., & Zhou, H. (2022). An intervention study to improve primary school students' self-regulated strategy use in English writing through e-learning in Hong Kong. *Computer Assisted Language Learning*, *35*(9), 2265-2290. https://doi.org/10.1080/09588221.2020.1871030

Chamot, A. U. (1995). Implementing the cognitive academic language learning approach: CALLA in Arlington, Virginia. *Bilingual Research Journal*, *19*, 379-394.

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual review of applied linguistics*, *25*, 112-130. https://doi.org/10.1017/S0267190505000061

Chamot, A. U., & O'malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Addison-Wesley Publishing Company Reading, MA.

Deshler, D. D., & Lenz, B. K. (1989). The strategies instructional approach. *International Journal of Disability, Development and Education*, *36*(3), 203-224. https://doi.org/10.1080/0156655893603004

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, *32*(4), 365-387. https://doi.org/10.2307/356600

Graham, S., & Harris, K. R. (2003). *Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies*. The Guilford Press.

Grenfell, M., Macaro, E., & Cohen, A. (2007). Language learner strategies: 30 years of research and practice. In.

Gu, Y. (2007). Strategy-based instruction.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hyland, K. (2003). *Second Language Writing Book*.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for language teaching and learning. *RELC Journal*, 00336882231162868. https://doi.org/10.1177/00336882231162868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, *5*, 100156. https://doi.org/10.1016/j.caeai.2023.100156

Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2021). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, *72*, 101202. https://doi.org/10.1016/j.learninstruc.2019.04.003

Oxford, R. L. (1990). Language learning strategies: What every teacher should know.

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, *90*, 102221. https://doi.org/10.1016/j.system.2020.102221

Vygotskij. (1979). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122-1136.

Zhang, Y., & Guo, H. (2012). A study of English writing and domain-specific motivation and self-efficacy of Chinese EFL learners. *Journal of Pan-Pacific Association of Applied Linguistics*, *16*(2).

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation* (pp. 13-39). Elsevier. https://doi.org/10.1016/B978-012109890-2/50031-7

数字人文视域下国内外语言文学研究的对比分析

Comparative Analysis of the Language and Literature Research from the Perspective of Digital Humanities at Home and Abroad

冀娇娇[1]，任佩玲[2]，武永[1*]
[12]北京邮电大学 人文学院
[*] wuyong@bupt.edu.cn

【摘要】 数字人文是计算机科学与人文学科交叉研究的一个新兴领域。随着大数据时代的到来，语言文学研究在技术的加持下正不断呈现出新的发展态势。本研究运用文献计量方法，基于 Web of Science 和 CNKI 数据库中数字人文视域下的语言文学研究，借助 CiteSpace 对相关文献进行知识图谱绘制，以对比国内外研究成果，总结宏观规律，发现研究不足，为数字人文视域下的语言文学研究提供建议和参考。
【关键词】 数字人文、语言文学、CiteSpace、研究热点

ABSTRACT: *Digital humanities is an emerging interdisciplinary field of computer science and humanities research. With the advent of the era of big data, language and literature research is presenting new trends with the support of new technologies. Using the bibliometric method, this study is based on the language and literature research in terms of digital humanities from Web of Science and CNKI database and use visual analysis software CiteSpace to draw knowledge map of the literature data. By analyzing the knowledge map of related literature, this study compares the domestic and foreign research results, summarize the macro law and find out the research shortcomings, thus offering new advice and references for researchers of language and literature in the field of digital humanities.*
Keywords: digital humanities, language and literature, CiteSpace, hot spots

1. 研究背景与意义

学界认为“数字人文”是从“人文计算”发展而来，20 世纪 40 年代“人文计算之父”布萨将文学文本与计算机相结合的成功实践开启了人文计算这一新领域。自 2009 年王晓光(2010)将“数字人文”概念引入中国以来，数字人文领域研究的热度逐年上升。随着信息技术、互联网技术的不断发展，各类技术方法在人们的生活中呈现出全方位的渗透，这正在改变传统意义上的以纸质媒体为主要对象、细读为主要方法的人文学科研究，“数字人文”概念成为最近一二十年以来人文社会学术发展的一股新潮流。数字人文作为典型的跨学科研究领域，涉及范围广泛。语言文学作为人文研究的重要方面，包含语言、文学、翻译、语言教学等多个研究方向，但目前数字人文在语言文学领域的研究和应用较少，应用层面尚浅，有必要对现有数字人文视域下的语言文学研究进行文献综述分析，厘清研究趋势，归纳研究热点，为相关领域的研究者提供参考和建议。

2. 研究方法

2.1. 文献检索方法

本研究选取 Web of Science 数据库的 SSCI(Social Science Citation Index，社会科学引文索引)和 A&HCI(Arts and Humanities Citation Index，艺术与人文科学引文索引)引文数据库，以 digital humanities, humanities computing, computing in humanities 为关键词进行主题检索，设置

时间跨度为 2010 年 1 月 1 日—2022 年 12 月 31 日，文献类型为论文，语言为英语。中文文献选取中国知网数据库(CNKI)的北大核心和中国社会科学引文数据库 (CSSCI)，以“数字人文”和“人文计算”为关键词进行主题检索，时间跨度为 2010 年 1 月 1 日—2022 年 12 月 31 日，文献类型为论文。此外，国内两本数字人文期刊《数字人文》和《数字人文研究》分别于 2020 年和 2021 年创刊，自其创刊至 2022 年底所刊载的文献被纳入本研究。

2.2. 文献筛选标准

本研究对数据进行了多轮筛选。首先，删除书评、会议摘要、新闻、报告等不符合本研究要求的文献。第二步，删除与语言文学无关的文献，如讨论数字人文研究方法的文献，仅保留语言文学领域的数字人文实证研究。逐一验证检索文献结果后，最终得到英文文献记录 156 条，中文文献记录 117 条。本研究将筛选完成的文献数据经过 CiteSpace(6. 2. R6 版本）进行可视化呈现与分析。

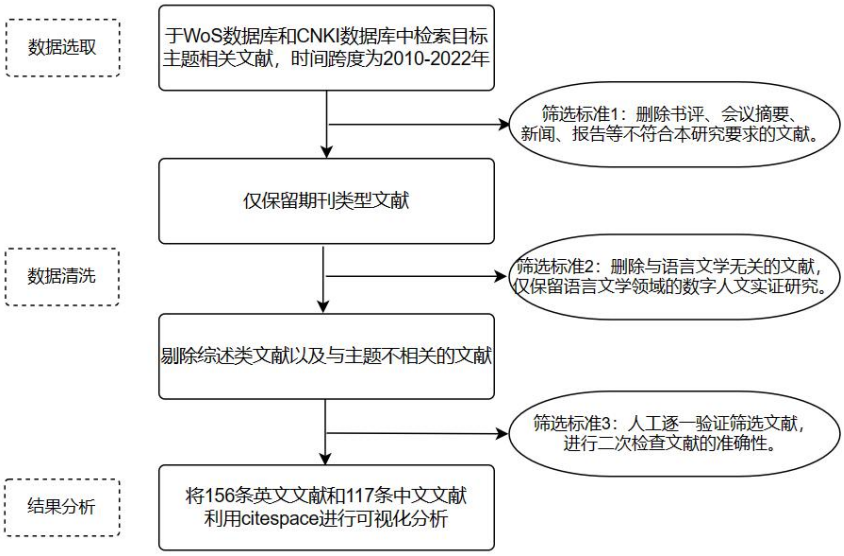


图 1 数字人文视域下国内外语言文学研究的对比分析研究框架

2.3. 研究问题

本研究旨在探讨数字人文视域下语言文学研究的 4 个问题：

1） 该研究领域国内外发文趋势有何特点？
2） 该研究领域国内外发文国家、机构有何特点？
3） 该研究领域国内外高被引文献与高影响力作者有何特点？
4） 该研究领域国内外文献的关键词及关键词聚类结果有何特点？

3. 结果与讨论

3.1. 国内外发文趋势分析

在发文量(见表 1 与表 2）方面，国外在数字人文领域的语言文学研究从 2011 年开始进入探索阶段，到 2015 年左右已具备一定成果，并在近年来保持稳定发展。而国内相关研究始于 2015 年，到 2018 年才展现出快速发展的态势，此后的 2020 年至 2022 年发文量较之前显著增长，但在总量上仍低于国外。从 2020 年开始，国内数字人文的语言文学研究发文量赶超国外，并在 2021 年将差距明显拉大。出现这一情况的主要原因在于“数字人文”起源于欧美国家，从 1949 年的“托马斯著作索引”发展至今已有 70 多年历史，而“数字人文”作为舶来词汇引入中国时间较晚。其次，随着大数据时代的到来和新文科建设的开展，“数字人文”在国内语言文学研究领域的地位和关注度高涨，成为学科发展的契机，使得数字人文领域的

语言文学研究不断涌现。

表 1 国外论文发文趋势(2011 年-2022 年)

年份	2011	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
发文量	1	6	4	12	10	14	15	22	21	21	30

表 2 国内论文发文趋势(2015 年-2022 年)

年份	2015	2016	2017	2018	2019	2020	2021	2022
发文量	1	2	0	10	9	25	37	33

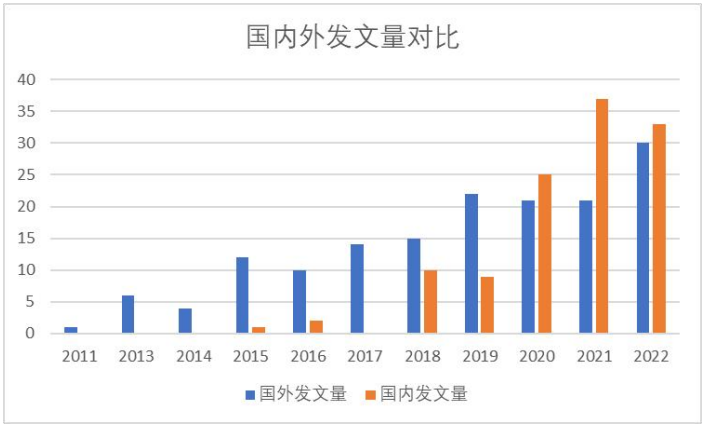


图 2 国内外数字人文语言文学研究的发文量对比

3.2. 国内外发文地区及机构分析

在发文地区以及机构(见图 3)方面，国际方面的数字人文研究主要来自美国和德国(美国有 33 篇，德国有 22 篇)。其次是英国、西班牙和荷兰，分别有 19 篇、9 篇和 9 篇，高发文量研究机构主要是大学的研究团队，如伦敦国王学院、斯坦福大学、悉尼大学、伦敦大学等。国外数字人文研究起步较早，各研究团队在不同研究方向上有深厚积淀和系统深耕。例如，伦敦国王学院早在 1973 年就成立了文字与语言学计算学会。由图 4 可以分析出，国内数字人文领域的语言文学研究主要集中在各高等院校进行，且其中南方高等院校机构占比较高。进一步分析可知，由南京农业大学信息管理学院和南京大学信息管理学院牵头的研究集中在古汉语典籍方面，其中针对古籍的自动标注、实体识别、自动分词和标点及其关键技术的应用为这些机构的主要研究内容，为国内数字人文视域下的语言文学研究提供了重要借鉴。



图 3 国外主要研究国家共现



图 4 国内主要研究机构共现

3.3. 国内外高被引文献与高影响力作者分析

表 3 列出了数字人文视域下的语言文学研究中，总被引频次排名前十位的中英文文献及其作者。从英文文献来看，学者发文所涉方向、学科分布更为广泛，如 Kong, Siu Cheung(2014)通过定性研究证实了数字课堂教学设计在支持信息素养能力和批判性思维技能发展方面的有效性。Ross, C.和 Terras Melissa 等学者(2011)对推特社交平台的文本进行挖掘，以分析此类平台如何能够被学者使用。Lansdall-Welfare(2017)收集了英国地区 150 年以来的报纸并对其进行文本挖掘，揭示出英国一个多世纪的文化变革。从英文高被引文献情况可以看出，英文高被引文献有主题较为分散、核心较为相似的特点。

通过分析中文高被引文献，可以发现近年来古汉语典籍文本研究这一具有鲜明中国特点的代表性研究呈现出热点趋势，研究者们利用语料库、机器学习和可视化等方法分析出新的数据，在古籍文本研究上寻找新的突破。在高被引学者中，黄水清等(2015)基于《汉学引得丛刊》制定词汇表，结合使用统计和人工内省方法确定的特征模板，进行了先秦典籍进行自动分词的探究。王东波等(2018)将先秦时期的《公羊传》《谷梁传》等典籍作为处理语料，对实体进行内部数量统计和外部特征分析，构建了历史事件基本实体构件自动识别的特征模板。欧阳剑等(2021)通过人机协作进行典籍数据抽取与整理，对全球目前存世的主要中国历代典籍书目信息进行了较全面的描述。从单一典籍语料到多本典籍语料，再到全球现存的主要中国典籍，国内对于古籍的研究不断朝着纵向深入发展，在数字人文兴起的背景下，数字人文方法显示出广阔的应用前景，为古籍深层次利用与开发带来了全新模式。

表 3 国内外高影响力文献与高被引作者统计

中文高被引文献	被引用次数	英文高被引文献	被引用次数
面向数字人文研究的大规模古籍文本可视化分析与挖掘	欧阳剑 (2016) 被引 143 次	Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy	Kong, Siu Cheung (2014） 214 citations
唐诗知识图谱的构建及其智能知识服务设计	周莉娜 (2019) 被引 83 次	Enabled backchannel: conference Twitter use by digital humanists	Terras Melissa (2011） 116 citations
SikuBERT 与 SikuRoBERTa：面向数字人文的《四库全书》预训练模型构建及应用研究	王东波 (2022) 被引 37 次	Content analysis of 150 years of British periodicals	Lansdall-Welfare, Thomas (2017） 51 citations
数字人文视域下的古文献文本标注与可视化研究——以《左传》知识库为例	李斌 (2020) 被引 33 次	On the Meanings of Self-Regulation: Digital Humanities in Service of Conceptual Clarity	Burman, Jeremy T. (2015) 43 citations
以《汉学引得丛刊》为领域词表的先秦典籍自动分词探讨	黄水清 (2015) 被引 32 次	Framing Psychology as a Discipline (1950-1999): A Large-Scale Term Co-Occurrence Analysis of Scientific Literature in Psychology	Flis, Ivan (2018) 42 citations

中文高被引文献	被引用次数	英文高被引文献	被引用次数
基于深度学习的方志物产资料实体自动识别模型构建研究	徐晨飞 (2020) 被引 31 次	ANNIS3: A new architecture for generic corpus query and visualization	Krause, Thomas (2016) 38 citations
面向先秦典籍的历史事件基本实体构件自动识别研究	王东波 (2018) 被引 30 次	The impact of lacking metadata for the measurement of cultural and linguistic change using the Google Ngram data sets-Reconstructing the composition of the German corpus in times of WWII	Koplenig, Alexander (2017) 30 citations
大规模中国历代存世典籍知识图谱构建研究	欧阳剑 (2021) 被引 27 次	Quantifying the impact of dirty OCR on historical text analysis: Eighteenth Century Collections Online as a case study	Hill, Mark J. (2019) 28 citations
数字人文框架下《德意志人物志》的群像描绘与类型分析	王涛 (2018) 被引 24 次	Anti-Asian Sentiments During the COVID-19 Pandemic Across 20 Countries:Analysis of a 12-Billion-Word News Media Database	Ng, Reuben (2021) 24 citations
社会网络分析与“《大波》三部曲”的人物功能	赵薇 (2018) 被引 24 次	Reading videogames as (authorless) literature	Berger, Richard (2013) 22 citations

3.4. 国内外数字人文领域语言文学研究的关键词统计

表 4 国外文献关键词统计(词频大于 3)

关键词	频次
digital humanities	71
distant reading	9
corpus linguistics	6
digital history	4

表 5 国内文献关键词统计(词频大于 3)

关键词	频次
数字人文	78
深度学习	11
可视化	6
先秦典籍	4
古籍	4
知识组织	4

本研究首先对中英文文献的高频关键词进行了统计，表 4 和 5 分别列出了中英文文献中频次大于 3 的所有关键词，英文文献主要关键词有 digital humanities、distant reading、corpus linguistics 和 digital history 等，中文文献主要关键词有数字人文、深度学习、可视化、先秦典籍和古籍等。

3.5. 国内外数字人文领域语言文学研究的关键词聚类统计

利用 CiteSpace 对国内外文献的关键词分别进行聚类分析，156 篇英文文献形成聚类共 12

个，117 篇中文文献形成聚类共 14 个，分别通过核心聚类筛选功能使其显示出 6 个具有统计意义的聚类，如图 5 和 6 所示，其中英文文献聚类标签为 digital humanities、data visualization、distant reading、teaching / learning strategies、data science 以及 corpus linguistics。中文文献聚类标签分别为数字人文、时空分布、语义、深度学习、大数据和机器学习。

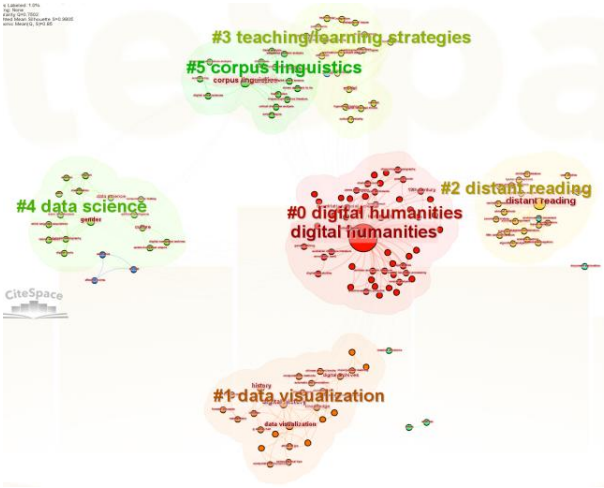


图 5 国外文献关键词聚类



图 6 国内文献关键词聚类

根据中英文文献的关键词词频统计以及聚类结果，可归纳出国内外数字人文视域下的语言文学研究主要有如下四个相似的热点领域：

1）文本远读：“远读”一词最早在 2013 年由斯坦福大学英语和比较文学的讲席教授弗朗哥·莫莱蒂 (Franco Moretti) 提出，“远读”概念是一个以计算机技术介入文本分析的重要方法和理论 (向帆、何依朗，2018)。此概念一经提出便引起国内外众多学者的关注。例如，Henrichs (2019) 和吉云飞 (2020)都探讨了“远读”这一定量与计算方法在文学研究中的适用性问题；Pikkanen (2019)和郭佳欣等(2021)都从数字人文技术出发对文学作品进行解析与挖掘，以呈现远读的特点和方法。

2）深度学习模型的应用：近年来随着自然语言处理技术的发展，国外学者开始进行研发深度学习监督算法的探索(Kestemont et al., 2016; Suissa et al., 2022)。而国内学者更多利用 BERT 等深度学习模型处理古籍文本(杜悦等，2021；李娜，2022；袁义国等，2022)。深度学习技术为语言文学研究带来了深刻革新与影响，鉴于技术的滞后性，更加广泛的研究领域尚未被涉及，研究特点与潜力并未完全显露，但根据已有研究成果，已经可以一窥深度学习技术驱动下语言文学研究的未来趋势。将有更多人文学者投身于相关研究领域，利用智能化的深度学习以全新角度看待人文研究。

3）大规模语料库的应用：自托马斯索引之后，数字人文随着语料库语言学的进步而不断发展(Sula & Hill, 2019)。语料库语言学作为计算语言学的分支学科，与语言学有着诸多联系。国外已有学者作出了将语料库语言学方法置于数字人文背景下文学文本研究的尝试，捷克国家语料库研究员切尔马科娃和伯明翰大学语料库语言学教授马尔伯格(Čermáková, & Mahlberg, 2021)通过研究两个不同时期的儿童文学语料库证实了随着时间的推移，肢体语言的使用的变化趋势。数字人文研究需要大规模语料库和高性能古文自然语言处理工具的支持(王东波, 2022)。国内多数古籍研究都需建立在大规模古籍文本语料库的基础上方可进行后续的应用研究。例如，杜悦(2021)等基于 25 本先秦典籍构建了古籍语料库，又分别基于不同规模的语料库训练深度学习模型，从而证实了深度学习应用于大规模文本数据集的可行性。

4）可视化分析技术的应用：目前在数字人文的语言文学领域所应用的可视化分析技术主要有地理信息系统(GIS)以及社会网络分析(SNA)两种。例如，韩国延世大学学者柳俊杰(Yoo, J.

J.,2021)通过以上两种可视化方法重建了18世纪首尔诗人群体的结构。地理信息可视化相关研究是在对历史地理信息研究的基础上，实现与文学领域研究相结合，为其研究提供多样的展现形式与有价值的参考。宋雪雁等(2022)使用计量统计方法与QGIS地理信息系统，对《全唐诗》贬谪诗人时空信息进行可视化展示，挖掘贬谪诗人在时间与地点上的分布特点。此外，在语言文学中，社会网络分析是很多关系型数据库的方法论基础(赵薇，2021)。学者通过将文学作品中的人物社会关系进行处理，编译为计算机可读编码，并利用已有数据库内的计算机算法来计算人物的一些基本信息权重和社会关系权重，从而对其社会关系进行清晰的分析以及可视化呈现。

4. 总结与展望

本文以WoS和CNKI的相关主题文献作为数据源，利用可视化分析工具CiteSpace呈现数字人文视域下国内外语言文学研究的发文趋势、发文国家与机构分布、关键词聚类等特征，厘清了国内外相关研究的发展现状与研究热点。依据分析结果，数字人文视域下国内外语言文学研究的特点可总结如下：

第一，从发文趋势来看，国内发文量虽然低于国际水平，但也呈历年增长态势，这表明国内学者对该领域关注度逐年提升。通过对国外文献进行发文地区的统计分析，可以得出美国是数字人文领域的语言文学文献发表数量最多的国家，学术影响力最大。通过对国内文献进行研究机构的统计分析，发现南京的高等院校是目前在数字人文方面语言文学研究的领军机构。

第二，从高被引文献与高影响力作者来看，国外高被引文献的主题与国内相比更为多样化、所涉及的学科分布也更为广泛。国内高被引文献则是聚焦于古籍文本的研究，例如古籍文本挖掘、典籍实体标注技术的研发等。由此可见，国内语言文学领域的数字人文研究在与国际相关研究接轨的基础上，向着凸显中国历史文化底蕴与特色的方向推进。

第三，从研究主题来看，根据关键词词频统计及聚类结果，国内外数字人文视域下的语言文学研究有四个相似的热点：1）文本远读的研究，2）深度学习模型的应用，3）大规模语料库的应用，以及4）可视化分析技术的应用。

从早期以量化研究作为主要方法的人文计算，到跨多学科多领域的数字人文，其已在科学、技术和人文之间建立了重要的桥梁，为探索文化、语言、文学和历史等领域提供了新的研究范式，数字人文在语言文学研究领域中正展现出广阔的前景和潜力。未来，数字人文技术将继续推动语言文学研究的创新和发展，助力人文学者挖掘新问题，得到新启发。

致谢

本研究受下列项目资助：2023年北京邮电大学大学生创新创业训练计划项目“儒家思想传承之数字人文研究”（项目编号:202309017）、2024年北京邮电大学研究生创新创业项目“数字人文视角下朱熹理学格局的形成过程研究”（项目编号:2024-YC-A192）、2024年北京邮电大学研究生教改项目：基于“项目式教学”的“数字外文”人才培养模式探索——以《数字人文概论》为例(项目号：2024Y016)。

参考文献

王东波，刘畅，朱子赫，刘江峰，胡昊天，沈思，& 李斌. (2022). SikuBERT与SikuRoBERTa：面向数字人文的《四库全书》预训练模型构建及应用研究. **图书馆论坛，31(06)**，43。

王东波，高瑞卿，沈思，& 李斌. (2018). 面向先秦典籍的历史事件基本实体构件自动识别研究. **国家图书馆学刊，65(01)**，77。

王晓光. (2010). 《“数字人文”的产生、发展与前沿》. **《方法创新与哲学社会科学发展》。**

吉云飞. (2020). 作为“计算批评”的“远读”——以网络小说“升级文”中的节奏与情绪为例. **中国现代文学研究丛刊 (08)**, 177-187。

向帆 & 何依朗. (2018). “远读”的原意:基于《远读》的引文和原文的观察. **图书馆论坛 (11)**, 44-48+43。

杜悦, 王东波, 江川, 徐润华, 李斌, 许超 & 徐晨飞. (2021). 数字人文下的典籍深度学习实体自动识别模型构建及应用研究. **图书情报工作 (03)**, 100-108。

李娜. (2022). 基于深度学习的《方志物产》用途实体自动识别模型构建与应用. **数字图书馆论坛 (12)**, 19-28。

宋雪雁, 霍晓楠, 刘寅鹏 & 邓君. (2022). 数字人文视角下《全唐诗》贬谪诗人的时空轨迹分析. **图书情报工作 (07)**, 26-34。

欧阳剑, 梁珠芳, & 任树怀. (2021). 大规模中国历代存世典籍知识图谱构建研究. **图书情报工作, 126(05)**, 135。

赵薇. (2021). 数字时代人文学研究的变革与超越——数字人文在中国. **探索与争鸣, 2021(06)**, 191-206+232-233。

袁义国, 李斌, 冯敏萱, 贺胜 & 王东波. (2022). 基于深度学习的古籍文本自动断句与标点一体化研究. **图书情报工作 (22)**, 134-141。

郭佳欣, 马昭仪, 肖天意 & 何捷. (2021). 《长安十二时辰》对唐长安城市空间的当代重构——一种文学制图的视角. **数字人文研究 (02)**, 9-20+113-116.

黄水清, 王东波, & 何琳. (2015). 以《汉学引得丛刊》为领域词表的先秦典籍自动分词探讨. **图书情报工作, 127(11)**, 133。

Čermáková, A., & Mahlberg, M. (2021). Gender inequality and female body language in children's literature. *Digital Scholarship in the Humanities, 36*(Supplement_2), ii72-ii77.

Henrichs, A. (2019). Deforming Shakespeare's sonnets: Topic models as poems. *Criticism, 61*(3), 387-412.

Kestemont, M., De Pauw, G., van Nie, R., & Daelemans, W. (2017). Lemmatization for variation-rich languages using deep learning. *Digital Scholarship in the Humanities, 32*(4), 797-815.

Kong, S. C. (2014). Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy. *Computers & Education, 78,* 160-173.

Lansdall-Welfare, T., Sudhahar, S., Thompson, J., Lewis, J., FindMyPast Newspaper Team, & Cristianini, N. (2017). Content analysis of 150 years of British periodicals. *Proceedings of the National Academy of Sciences, 114*(4), E457-E465.

Pikkanen, I. (2019). The metrics and poetics of historical drama: The dramatis personae of a premodern revolt in early nineteenth-century Finland. *Orbis Litterarum, 74*(5), 311-339.

Ross, C., Terras, M., Warwick, C., & Welsh, A. (2011). Enabled backchannel: Conference Twitter use by digital humanists. *Journal of documentation, 67*(2), 214-237.

Suissa, O., Elmalech, A., & Zhitomirsky-Geffet, M. (2022). Text analysis using deep neural networks in digital humanities and information science. *Journal of the Association for Information Science and Technology, 73*(2), 268-287.

Sula, C. A., & Hill, H. V. (2019). The early history of digital humanities: An analysis of Computers and the Humanitites (1966-2004) and Literary and Linguistic Computing (1986-2004). *Digital Scholarship in the Humanities, 78*, 160-173.

Yoo, J. J. (2021). Poetry in action: Networks of literary communication and the cultural leverage in the eighteenth-century Seoul. *Digital Scholarship in the Humanities, 36*(3), 766-781.

AI 協助英語對話機器人導入至國中——某國中英文課的課室觀察報告

Integrating AI-based English Conversation Chatbot into Junior High School: An English Course Observation Report

余儘卿*，洪霈羽，張立杰
中央大學學習與教學研究所
chinching@ms.tyc.edu.tw

【摘要】生成式 AI 對話機器人有機會進入課堂提供學生英語對話機會。過往 AI 在語言學習研究多重於學生學習成效，但較少對於高低成就學生參與度及教師接受度和觀感的研究。為了解 AI 英語對話機器人成效，本研究團隊開發名為 ALICE 的 AI 英語對話機器人，探討如何將此導入國中英語課堂，其方法採用課室觀察和教師的問卷回饋，探討 ALICE 對高低成就學生英語對話機會是否有實質幫助，以及師生對 ALICE 的反應為何。研究結果顯示學生對 ALICE 的態度和參與度有所提升，高低成就學生在 5 至 10 分鐘至少有一次英語對話機會，老師接受度及觀感有提升，此結果為教育應用提供有益的參考。

【关键词】 生成式 AI；英語對話；TELL

Abstract: *Generative AI-powered chatbots present an opportunity to enhance students' English conversational skills in the classroom. Previous research on AI in language learning has predominantly focused on the learning outcomes, lacking in-depth exploration of engagement and teacher perceptions, particularly among high and low-achieving students. To address the efficacy of an AI chatbot, the research team developed an English conversation chatbot named ALICE, investigating its integration into junior high school English classes. Through classroom observations and teacher surveys, the study explores whether ALICE provides substantial benefits for English conversation opportunities among high and low-achieving students and assesses teacher and student responses. The results indicate the attitudes and engagement were improved with both high and low-achieving students having meaningful English conversation opportunities. Teacher's acceptance and perceptions also showed enhancement. These findings offer valuable insights for educational applications.*

Keywords: Generative AI, English Conversation, Technology enhanced language learning

1. 前言

國中學生英語能力在聽說讀寫四部分中，以說的能力在英文課堂上較少有機會被加強，在一堂課程 45 分鐘的情況下，一位英文老師對約三十位學生，能協助到每個學生的英文對話時間，明顯過低。這樣的時間分配，老師很難照顧到每位學生的英語對話練習。此外，教師的人力成本昂貴，教學上很難再增加額外的英語師資。

目前科技技術進步，包含生成式 AI，語音合成及語音辨識的技術均有很大幅度的突破，使得使用生成式 AI、語音合成及語音辨識在學生進行課堂中的英語對話練習變得可能。如果導入 AI 的對話機器人，除了可以提升學生對話的機會跟能力外，更可以降低老師在教室內一對多位學生，無法兼顧學生對話的困難（阮家慶，2021）。此外，過往，國中學生和英文老師進行口說英文對話時，通常較為緊張，但透過 AI 對話機器人，或許有機會降低進行英文對話時的緊張程（楊書佩，2020)，也可以比較進行多次的練習，來達到較好的英文對話練習，提升英文聽和說的能力。

因此，這個研究中，研究團隊開發了一個名為 ALICE(Advanced Learning through Interactive and Communicative Environment) 的「生成式 AI 對話機器人」。ALICE 可以協助學生英語口說對話練習，每位學生可使用平板進行對話練習， 像 ALICE 這類的「生成式 AI 對話機器人」進入到教室後，會產生很多不同面向的學習改變，這些面向包含：老師、學生、甚至學校的教學環境。本篇研究，針對課堂上使用 ALICE 提升學生口說英文的機會、高低成就學生英語對話的參與度、老師對 ALICE 的接受度及觀感，採課室觀察及問卷回饋的研究方法實施。

2. 文獻探討

針對上面提到的研究問題，涵蓋的面向包含英語教學、科技輔助語言學習等二個面向，因此，我們針對這二個面向進行文獻探討並報告如下：

2.1. 英語教學

過去已有許多研究強調英語對話對於語言學習的重要性，需要考慮的點除了在課室內外應有不同的學習環境外，也應該整合英語學習的資源設計出適合學習的雙語環境（陳浩然，2021）而在科技進步的今日，語言學習者更可以將這些技術利用於第二語言或是外語學習。數位科技在語言學習中雖然帶來了便利性及提高了學生的參與度，但同時也存在著溝通和互動方面的限制（Lo，2023）在關注學生學習狀態的同時，對於教師也應探討其資訊通信技術之接受度，Bai 與 Chai（2019）就曾以科技接受模型（TAM）、期望價值理論及學習視角開發面向教師在課堂中使用科技的模型，發現教師們在是否有持續使用其技術，興趣是一大影響。而機器人在雙語教育及師資培育能夠進行的協助應是在學生、語言教師及學科教師之間進行協調，透過學生模型的記錄，老師也能根據其學習歷程協助學生進行學習（阮家慶，2021）。

2.2. 科技輔助語言學習

對於科技融入教育的過往研究中，線上教學影片、各式網路資源和數位遊戲等電子學習活動，常在老師的鼓勵之下提高學生對科技融入的接受度，並且有效地降低學習焦慮（Lau 與 Keung，2021）。關於語音辨識技術，夏崇舜等人（2005）的研究中指出，如果能夠在生動的多媒體數位教材中融入語音辨識技術，有機會提高學生學習英語的興趣、鼓勵他們開口說英語，甚至使學生自動自發進行有效學習。而在 2022 年，由 Open AI 開發的人工智慧聊天機器人程式 ChatGPT 發行，將生成式 AI 這個技術帶來語言學習領域，前期的聊天機器人對於語言學習較無明確的概念，忽略不同模組之間的特徵，但是在對話系統的改進，聊天機器人可以提供較為自然及豐富的對話環境，提供更好的交流體驗（Jeon 等人，2023）。

依據上面探討的文獻，我們發現在人工智慧應用在語言學習上，多半探究學生學習的成效，但是對於提升學生口說英語的機會、高低成就學生英語對話的參與度、老師的接受度及觀感，則尚未有比較深入的研究。

3. 研究方法

針對上面提到的目標，包含提升學生口說英文的機會、以及降低老師的教學負擔，我們設計開發了一個名為 ALICE 的系統。此外，更重要的部分在於這個系統實施後，我們預計實施提升學生口說英文的機會、高低成就學生的英文對話參與度、老師的接受度及觀感的觀察。

底下我們針對這個系統的設計與實施過程進行說明，並闡述我們預計實施的研究者課室觀察及英文任課教師回饋問卷內容。

3.1. ALICE 系統設計說明

ALICE 為本研究所使用之英語學習平台，利用語音辨識與合成、OpenAI 及學生模型之技術，提供適於學生的英語對話環境，結合系統紀錄及任務導向之活動設計，可使在無縫的網路環境下按照學習進度，隨時隨地進行英語對話學習。

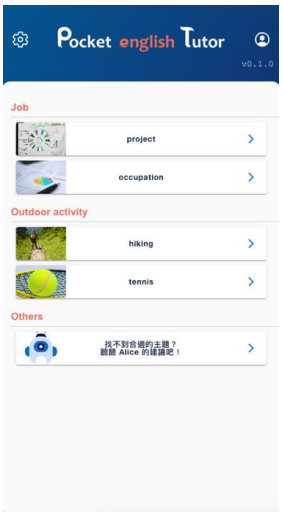
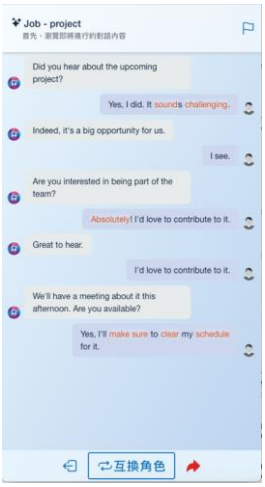
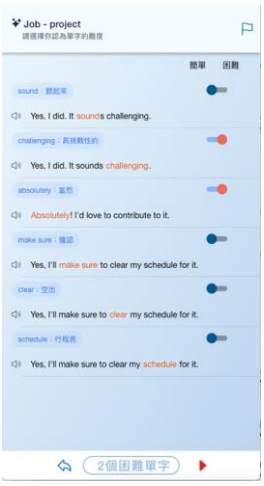
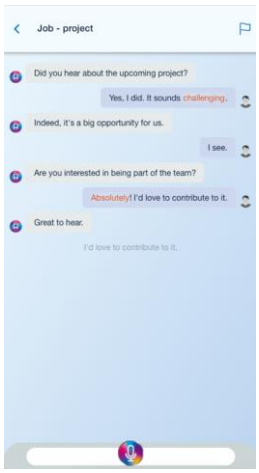
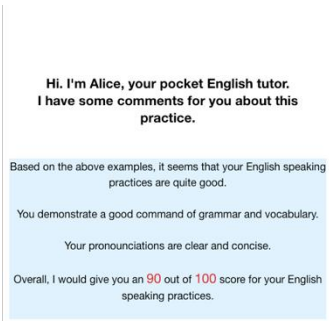
<table>
<tr><td> 
</td><td> 
</td><td> 
</td></tr>
<tr><td>步驟一：選擇對話</td><td>步驟二：對話內容確認</td><td>步驟三：生詞理解</td></tr>
<tr><td> 
</td><td> 
</td><td></td></tr>
<tr><td>步驟四：對話練習</td><td>步驟五：結果評估</td><td></td></tr>
</table>

圖 1 英語對話互動示意圖

此軟體的運作是以學生模型中的記錄，從課程樹中獲取符合學生現階段能力之教材，而整體對話練習共會經過五個步驟：選擇對話 、對話內容確認、生詞理解、對話練習、結果評估。

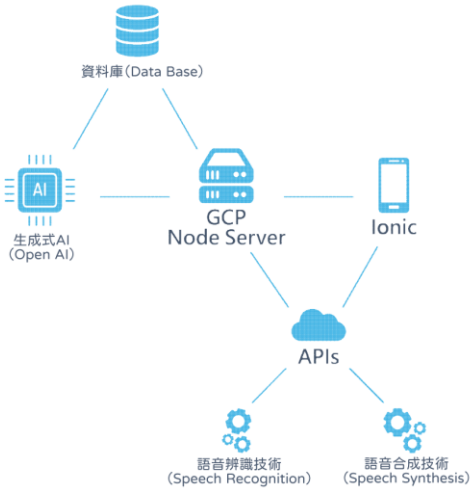


圖 2 ALICE 技術結構圖

3.2. ALICE 系統進入教室內的課堂學生實施過程

本研究安排在台灣桃園市內某一所市區型國中某一八年級班級的學生做實施，研究者預先安排一場生成式 AI 對話機器人的計畫說明會，由計劃指導教授主講，邀請該班級的導師、英文任課老師、資訊組長、資訊老師、教務主任及校長進行本計畫內容的了解，取得學校合作計畫意願書後，發下知情同意書給該班全部學生，學生共計 26 人，先做該班環境測試。

每次課堂一開始，學生使用平板上網連結 ALICE 英語學習系統(Pocketor.CoCoing.info)，並依說明進入單字練習畫面，然後再進入四個單元主題，依次選一個單元主題 8 至 12 句的英文對話內容，和 ALICE 進行 5-10 分鐘英語對話練習，完成後平台會依據學生的回答給予一個評分，學生可以在時間內再重複對話練習，來取得一個較滿意的分數。

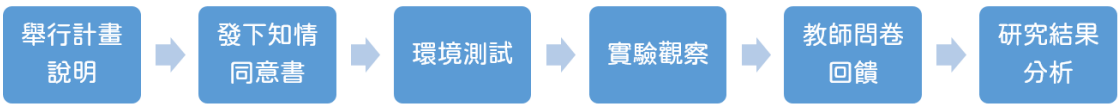


圖 3 研究實施流程圖

3.3. ALICE 系統課室觀察及問卷回饋實施過程

3.3.1. 研究者的課室觀察

針對研究者的課室觀察，採用課堂上學生使用平板操作 ALICE 的過程中，隨班巡視觀察學生的使用情況，隨堂紀錄高低程度學生和 ALICE 進行英語對話的反應，並觀察及詢問英文任課教師在每次學生和 ALICE 對話後的使用度及教師觀感的了解。

3.3.2. 英文任課教師的問卷回饋

本研究自行設計任課教師問卷回饋的內容(如表一)，並在學生完成四個 ALICE 主題對話練習後，進行英文任課教師問卷回饋的填寫。

表一、問卷內容

「AI 協助英語對話導入至國中—某國中英文課的課室觀察報告」問卷回饋

填寫人：_____ 學校：_____ 職務：_____
任教班級：_____，學生人數：_____ 日期：____年_____月_____日

1. 請問您任課的班級學生中，原先勾選不參加生成式 AI 協助英語對話研究計畫，後來參加的學生人數有多少？_____
2. 經您了解原因為何呢？_____
3. 請問您任課的班級學生中，在 AI 協助英語對話導入至國中課堂第一次主題內容：職業(Occupation)的過程中，您觀察到全班學生的使用興趣如何？_____
   英文高學習成就的學生使用後的反應如何？_____
   英文低學習成就的學生使用後的反應如何？_____
4. 請問您任課的班級學生中，在使用 AI 協助英語對話導入至國中課堂第二次主題內容：網球(Tennis)的過程中，您觀察到全班學生的使用興趣如何？_____
   英文高學習成就的學生使用後的反應如何？_____英文低學習成就的學生使用後的反應如何？_____
5. 請問您任課的班級學生中，在使用 AI 協助英語對話導入至國中課堂第三次主題內容：專案(Project)的過程中, 您觀察到全班學生的使用興趣如何？_____
   英文高學習成就的學生使用後的反應如何？_____
   英文低學習成就的學生使用後的反應如何？_____
6. 請問您任課的班級學生中，在使用 AI 協助英語對話導入至國中課堂第四次主題內容：海邊(Beach)的過程中，您觀察到全班學生的使用興趣如何？_____
   英文高學習成就的學生使用後的反應如何？_____
   英文低學習成就的學生使用後的反應如何？_____
7. 請問下次邀請您的任課班級加入 AI 協助英語對話導入至國中課堂, 您會參加嗎？_____
8. 經過本次使用後，如果下一個英文主題對話內容，您會建議哪一個主題呢？_____
9. 有關學生使用 AI 協助英語對話練習，對英文高學習成就的學生使用，您會建議如何實施呢？_____
   對英文低學習成就的學生使用，您會建議如何實施呢？_____
10. 有關學生使用 AI 協助英語對話練習，想推廣到更多國中生來做使用，請問您建議如何讓英文任課教師願意加入的做法為？_____
11. 有沒有什麼意見，是我們沒注意到，而想提供給我們知道的呢？_____

4. 研究結果

依據上述 3.2 和 3.3 節的實施結果，我們將研究結果分研究者的課室觀察、任課教師的問卷回饋及 ALICE 系統對於學生提升對話機會的觀察進行報告。

4.1. 研究者的課室觀察

4.1.1. AI 自動評分對學生產生的反應

研究者課室觀察到，第一次英文主題職業(Occupation) 和 ALICE 來進行英文對話練習後，有一位學生獲得 100 分，馬上詢問老師可以進行下一主題嗎？顯示英文高學習成就學生對 ALICE 來進行英語對話練習是很有參與度的。其他學生在第一次使用 ALICE 英語對話完，獲得的分數為 40 多分、或 60 多分、或 80 多分，練習完第一次獲得的分數，學生會詢問旁邊的同學：「你得幾分?」然後再次和 ALICE 對話練習挑戰得高分，研究觀察到大部分的學生會持續和 ALICE 進行英語對話練習，直到實施時間截止。

4.1.2. 高低成就學生使用 ALICE 的反應

研究觀察到高成就學生對使用 ALICE 的參與度很高，有一位學生獲得 100 分，馬上詢問老師可以進行下一主題嗎？顯示高成就學生對 ALICE 來進行英語對話練習是很有興趣的。但第二次至第四次測驗，因為評分系統不穩定，所以高成就學生因為評分成績為 0 分，感覺困惑，會馬上舉手詢問老師為何未如此，老師請學生重複再次測試並將口說聲量放大，仍獲得 0 分，老師告知會反應系統的評分異常狀況，高成就學生對於自己英文能力的認知無明顯影響。研究觀察到低學習成就學生會很慢才開始和 ALICE 對話練習，他們會先將不懂的單字逐一聽完發音後才開始，或需要老師一旁鼓勵即便無法說出完整的一個句子，但能說多少單字就儘量說，才能完成一個單元主題的對話練習。

4.1.3. 學生其他的反應

研究觀察到學生第一次參與度的變化，原先不參加使用 ALICE 英語對話平台的 2 位學生因看到同學操作體驗的活動，發現不是考試，現場表示想參加。另研究觀察到學生在使用 ALICE 最後一次時，有學生表示希望還有再次使用的機會，及大部分學生可以馬上操作平板和 ALICE 進行英語對話練習，學生在使用過程中專注參與，並且較無觀察到學生和 ALICE 對話的緊張感。

4.1.4. 任課教師的使用度及觀感了解

研究者在學生第一次至第四次使用 ALICE 英語對話練習後，立即訪談任課教師的使用情況及觀感，英文任課教師表示學生反應很喜歡，並會詢問下次何時再使用，老師也覺得對學生英語對話練習很有幫助，但對於 AI 自動評分出現 0 分的現象則反應希望能儘快解決，讓學生在練習英語對話時，能有評分回饋，若評分系統穩定度提高可讓任課教師多一項學生英語口說的評分依據。

4.2. 英文任課教師的問卷回饋

4.2.1. 高低成就學生對 ALICE 四個英語對話主題的反應

- 請問您任課的班級學生中，在 AI 協助英語對話導入至國中課堂第一次主題內容：職業(Occupation)的過程中，您觀察到全班學生的使用興趣如何？
  - 相當感興趣，因為是第一次使用 iPad 做口說測驗，會主動問甚麼時候要再做實驗。英文高學習成就的學生使用後的反應如何？成績皆相當理想，學生很開心。英文低學習成就的學生使用後的反應如何？大部分低成就學生因為對自己的英文口說能力本來就較無信心，所以對於成績較沒有明顯的情緒表現。

4.2.2.高低成就學生對 ALICE 四個英語對話主題的反應

- 請問您任課的班級學生中，在 AI 協助英語對話導入至國中課堂第一次主題內容：職業(Occupation)的過程中，您觀察到全班學生的使用興趣如何？
  - 教師反應如下：相當感興趣，因為是第一次使用 iPad 做口說測驗，會主動問甚麼時候要再做實驗。英文高學習成就的學生使用後的反應如何？成績皆相當理想，學生很開心。英文低學習成就的學生使用後的反應如何？大部分低成就學生因為對自己的英文口說能力本來就較無信心，所以對於成績較沒有明顯的情緒表現。

- 請問您任課的班級學生中，在使用 AI 協助英語對話導入至國中課堂第二次至第四次主題內容：網球(Tennis)、專案(Project)及海邊(Beach)的過程中，您觀察到全班學生的使用興趣如何？英文高學習成就的學生使用後的反應如何？英文低學習成就的學生使用後的反應如何？

➢ 教師反應如下：學生對於用 iPad 做口說測驗感興趣。英文高學習成就的學生因為評分成績為 0 分，感覺困惑，但對於自己英文能力的認知無明顯影響。對於是最後一次做測驗感到不捨。大部分低成就學生因為對自己的英文口說能力本來就較無信心，所以對於成績較沒有明顯的情緒表現。

- 有關學生使用 AI 協助英語對話練習，對英文高學習成就的學生使用，您會建議如何實施呢？對英文低學習成就的學生使用，您會建議如何實施呢？

➢ 教師反應如下：高成就學生對於覆誦題目較易上手，所以如果是開放性問題，讓學生可以自由回答，更能測驗出高成就學生對語言使用的掌握度，此外，AI 回饋評語如果可以更精細，例如精確指出該修正的單字，學生可以修正及進步的更快。對於低成就學生必須反覆聽相同字詞才能記住該字詞發音，所以回答問題時如果可以在同一介面反覆點選該字詞或句子，學生才能更快記住字詞或句子念法。

4.2.3. 教師的使用度和觀感及未來 ALICE 應用建議

- 請問下次邀請您的任課班級加入 AI 協助英語對話導入至國中課堂，您會參加嗎？如果會，您會建議哪些主題呢？

➢ 教師反應如下：會。建議旅遊，購物等日常生活對話。

- 有關學生使用 AI 協助英語對話練習，想推廣到更多國中生來做使用，請問您建議如何讓英文任課教師願意加入的做法為？

➢ 教師反應如下：AI 可以直接輸出成績及評語，減輕老師分析及評量各個學生的口說能力的負擔，所以如果可以結合現有的教課書內容設計題目，或者可以讓老師自行輸入相關題目，作為口說練習或測驗的方式，不但可以節省老師上課點選不同學生朗讀的時間，英文任課教師也會有更高願意加入。

- 有沒有什麼意見，是我們沒注意到，而想提供給我們知道的呢？

➢ 教師反應如下：主要問題是 AI 評分部分，第一次使用時有低成就學生未回答但得到高分，第二次之後有些高成就學生回答後但得到 0 分，所以 AI 評分規準可能需要再修正的更精細及精準，承上題，如果 AI 評分結果能更精確呈現學生的英文口說程度，英文任課教師加入的意願也會提高。

5. 討論與結論

本研究經由學生四次使用 ALICE 系統進行英語對話練習的研究結果討論與結論。

5.1. ALICE 系統對於學生提升對話機會的觀察

研究觀察到學生參與度變化第一次是有提升的，原先不願意參加的學生現場表示想參加。AI 自動評分系統對學生產生的反應會讓高成就學生持續想使用 ALICE 進行其他主題測驗或持續對話練習，以獲取較高分數的回饋，直至測驗時間截止，也讓低成就學生在教師鼓

勵下完成一次英語對話練習，顯示 ALICE 對學生在課堂上的使用興趣及參與度提升。另研究觀察到學生在與 ALICE 進行對話練習時，較無緊張感。研究觀察到學生在四次主題中使用平板進入 ALICE 系統五至十分鐘至少都有一次口說英語的機會，其中高成就學生觀察到可以使用二至三次次的英語對話機會，而低成就學生可以使用一次的英語對話機會，顯示 ALICE 系統能提升學生口說英語的機會。

5.2. 高低學習成就學生對 ALICE 的反應

研究觀察到學生使用 ALICE 口說對話練習的參與度提高，高成就學生對 ALICE 主題式口說英語對話練習的機會有二至三次，對 ALICE 英語對話練習積極參與；而低成就學生也有一次口說對話機會，但需教師鼓勵，才能將口說對話練習完成。研究發現學生在使用 ALICE 完成四次英語口說對話練習後，表示希望有下次使用的機會，表示學生對 ALICE 使用興趣提升。因觀察學生在英語課堂上使用 ALICE 口說對話練習，高成就學生積極參與想獲得高分，且在 5 至 10 分鐘即可完成二至三次英語對話練習，而低成就學生則有一次主題式口說英語對話練習，顯示使用度及觀感有提升。

5.3. 教師對 ALICE 的使用度及觀感

研究觀察到英文任課教師對英語課堂加入 AI 協助英語對話導入至課堂願意參加，表示教師使用過 ALICE 後對 AI 英用於英文課堂的接受度及觀感提升。另英文任課教師表示 AI 自動評分系統的準確度若能提高及穩定，也會提升英文任課教師使用的意願，透過 ALICE 系統對於學生在課堂中使用 5 至 10 分鐘能提升英文對話的機會是有幫助的，顯示教師觀感提升。

研究未來展望：未來研究設計將會加入多元英文主題及國中七至九年級的英文對話練習課程設計，透過生成式 AI 的快速演算，推進適合學生程度的英文主題對話練習課程。AI 評分系統能更精確呈現學生的英文口說程度並能直接輸出成績及評語，減輕老師分析及評量各個學生口說能力的負擔。此外，後續研究將增加系統的學生及高低學習成就學生的人數，以期達到統計數據上的提升學習成效等多元面向的研究分析。

6. 參考文獻

阮家慶（2021）。人工智慧在雙語教育的重要性。**台灣教育**。**727**。3-64。

夏崇舜和王定欽（2005）。語音辨識系統對於英語文學習的影響。**元培學報**。**12**。21-32。

楊書佩（2020）。探討沉浸式虛擬實境英語對話練習系統對學習動機口說焦慮學習行為與學習成效之影像〔未出版之碩士論文〕。國立雲林科技大學。

陳浩然（2021）。有效營造雙語環境，提升中小學學生英語力。**台灣經濟論衡**。19(3)。26-35.

Bai, B., Wang, J., & Chai, C. S. (2021). Understanding Hong Kong primary school English teachers' continuance intention to teach with ICT. *Computer Assisted Language Learning, 34*(4), 528-551.

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 104898.

Lo, N. P. K. (2023). Digital learning and the ESL online classroom in higher education: Teachers' perspectives. *Asian-Pacific Journal of Second and Foreign Language Education, 8*(1), 24.

Lau, K.L., & Keung, H. M. (2021). Effectiveness and user acceptance of using out-of-class e-learning activities to support classical Chinese learning. *Journal of Chinese Language Teaching, 18*(3), 37-71.

虚拟现实技术应用于语言教育的系统性文献综述(2014 - 2023)

A Systematic Literature Review of Application of Virtual Reality in Language Education (2014--2023)

马萍[1]， 陈真真[*]
[1]北京邮电大学人文学院
[*]北京邮电大学人文学院
czz@bupt.edu.cn

【摘要】 近年来，虚拟现实（VR）技术在语言教育领域的应用日益广泛，它能够支持互动式、情境化和协作式的语言学习。本研究基于 4 本教育技术以及计算机辅助语言学习（CALL）领域的高影响力国际期刊，筛选出近 10 年（2014-2023）虚拟现实技术应用于语言教育领域的 32 项实证研究，从整体研究趋势、研究方法、研究主题三方面构建编码框架，采用内容分析法对虚拟现实技术在语言教育中的应用展开系统性文献综述。本研究能够为虚拟现实技术应用于语言教学研究与实践提供有益的参考。

【关键词】 虚拟现实；语言教育；系统性文献综述；内容分析

Abstract: *In recent years, the application of Virtual Reality (VR) technology in the field of language education has received increased attention, as it can support interactive, contextualized, and collaborative language learning. Drawing from four high-impact international journals in educational technology and computer-assisted language learning (CALL), this paper offers a systematic review of existing research on VR-based language education, encompassing 32 peer-reviewed empirical studies between 2014 and 2023. These studies were coded on a basis of the general research trends, research methodologies and major research themes, and then were systematically analyzed by content analysis method. This study can provide useful insights into the application of virtual reality in language teaching research and practice.*

Keywords: virtual reality, language education, systematic literature review, content analysis

1. 前言

虚拟现实，是指利用计算机技术模拟产生一个三维立体的虚拟空间，为用户提供视觉、听觉等感官模拟，令其获得身临其境的幻觉（Huang et al., 2021）。自上世纪 60 年代问世以来，虚拟现实技术（Virtual Reality，以下简称 VR）经历了飞速的发展和演变。相比于早期基于文本的二维（2D）在线虚拟环境，如今的 VR 工具已变得更加复杂和更具互动性，它可以提供三维（3D）虚拟空间以及可定制的虚拟化身（Lin & Lan, 2015）。VR 技术逐渐在语言教育领域展现其极大的潜力，它可以通过创造出强烈的身临其境感为学习者提供一种情境化和交互式的学习体验。虚拟现实环境中外语学习者的二语习得过程十分复杂，然而，相关实证研究的数量还非常有限。本研究采用内容分析法，针对近十年国际上相关的实证研究展开系统性文献综述，旨在梳理虚拟现实技术应用于语言教育领域的发展趋势、研究方法与研究主题，总结现有研究的不足之处，为未来虚拟现实技术应用于语言教学研究与实践提供有益的参考。

2. 研究方法

本研究基于近十年(2014-2023) 的国外文献，对虚拟现实技术应用于语言教育领域的实证研究进行筛查，对筛选的文献开展系统性文献综述(Systematic Literature Review)。作为一种新兴且快速发展的综合性研究方法，系统性文献综述法以其特有的客观性、标准性、创新性和亲和性，在研究主题、研究主体和研究方法学上展现极大的发展潜力(黄甫全等, 2017)。

2.1 文献筛选

为有效获取国外近十年虚拟现实技术应用在语言教育领域的高质量实证研究文献，本研究选择 4 本国际期刊，分别为 *Computer Assisted Language Learning (*以下缩写为 *CALL), ReCALL, Computers & Education (*以下缩写为 *C & E)* 和 *Language Learning & Technology (*以下缩写为 *LLT)*。这 4 本学术期刊发表的文献均经过严格评审，是国际语言学习与教育技术领域具有代表性、权威性以及学术影响力的学术期刊，期刊内容均聚焦于技术支持下的语言学习、教学与测试。

本研究采用逐刊检索的方式，结合标题、摘要和关键词，进行人工筛查。在剔除编辑部致辞、书评、软件简评以及期刊中未涉及虚拟技术语言教育应用的文献之后，以“virtual reality”, “virtual environment”，并含“language learning”，“language teaching”等关键词进行主题精确检索。最终经过人工筛查，剔除无关文献 23 篇，获得有效英文文献 32 篇。

2.2 文献编码

如图 1 所示，笔者分别从研究趋势、研究方法、研究主题三个方面对文献进行编码。



图 1　编码框架

3. 研究结果和讨论

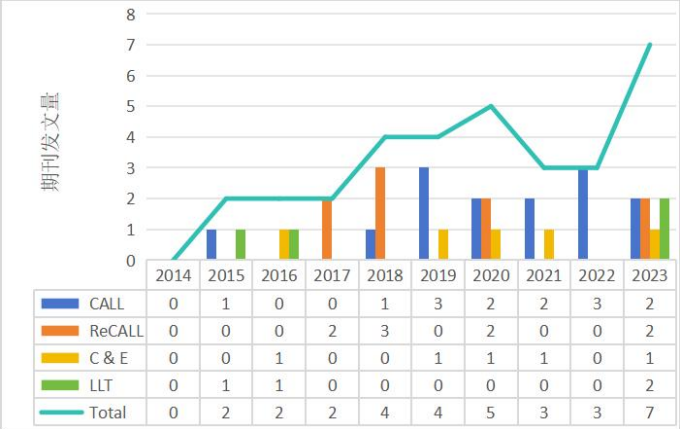
3.1 整体研究趋势

3.1.1 发文数量

近十年来，四本国际期刊共刊载虚拟现实应用于语言教育的实证研究合计 32 篇。从发表时间来看，当前国外将 VR 技术应用在语言教育领域的实证研究虽然发文量不大，但整体呈现稳步发展的趋势，可以窥见新兴智能技术在语言教育领域的应用正在逐渐受到研究者的关注。

从期刊分布来看，在过去十年中发表文章数量最多的是 *CALL*（14/32），其次是 *ReCALL*（9/32）、*C & E*（5/32）和 *LLT*（4/32）。与 Recall、C & E 和 LLT 期刊相比，CALL 发表了更多与本研究主题相关的文章，可能是因为 CALL 每年出版 8 期，C & E 每年出版 6 期，而 ReCALL 和 LLT 每年只出版 3 期, 其他相关研究也报告了类似的发现 (Shadiev & Yang, 2020;

Shadiev et al., 2017)。



	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
CALL	0	1	0	0	1	3	2	2	3	2
ReCALL	0	0	0	2	3	0	2	0	0	2
C & E	0	0	1	0	0	1	1	1	0	1
LLT	0	1	1	0	0	0	0	0	0	2
Total	0	2	2	2	4	4	5	3	3	7

图 2　2014-2023 年 4 本期刊实证研究年度发文量

3.1.2 目标语言

在这 32 项研究中，所涵盖的语言范围仅限于第二语言或外语（L2）。绝大部分文献关注的目标语为英语(23/32) ，只有小部分文献分别讨论了虚拟现实技术在汉语(5/32) 、西班牙语（2/32）、法语(1/32) 和日语（1/32）教学中的应用，这一结果与众多过往技术增强环境下的语言教育综述研究结论相一致（例如 Chen et al., 2020; Shadiev et al., 2017; Ghanizadeh et al., 2015 等）。英语教学在全球仍位居主流，这与英语全球“通用语(lingual franca)”的地位相吻合，汉语也因为在全球的使用人数广排在第二位。而关于 VR 技术应用于其他小语种教学的实证研究仍相对匮乏，需要在今后的研究中予以重视。

3.1.3 研究对象

这 32 项研究在参与对象的规模上存在较大差异，参与者人数从最少的 2 人（Kruk，2022）到最多 274 人（Chen & Hsu, 2020）。大多数研究（22/32）的参与者人数在 10 到 100 人之间，一些研究（7/32）少于 10 人，少数研究（3/32）超过 100 人。就参与者类型而言，1 项研究的参与者背景多元化（Chen, 2022），仅 2 项研究调查了小学学段，6 项研究调查中学学段，5 项研究调查了不同年龄的成人学习者，其余 18 项研究的参与对象均为大学生。通过分析研究对象所在的学段，超过 2 /3 的实证研究针对高等教育或继续教育等成人学习者展开(22/32)，只有少量研究以中学生(6/32)或小学生(2/32)为研究对象。与中小学生相比，成人学习者更容易在虚拟现实环境中开展相对复杂的操作，适应多样化的任务型教学设计，因而，针对成人学习者的实证研究也更容易开展。此外，2014 年-2018 年的研究多以高等教育学生或成人学习者为目标对象，针对中小学生展开的实证研究基本集中在 2019 年-2023 年，尽管数量有限，但显而易见的是研究者已经开始关注 VR 技术在基础教育中的应用，有不少学者认为 VR 技术有利于提升中小学生外语词汇能力与阅读能力(如 Alfadil，2020) 。

3.2 研究方法

3.2.1 研究类型

文秋芳和林琳(2016) 将研究方法分为量化、质性和混合研究，其中混合研究是指量化与质性相结合的研究方法。本研究参考此分类方法，将综述的 32 篇实证研究分为量化、质性和混合研究 3 种类型。研究结果显示，量化研究占比 28.1%（9/32），定性研究占比 9.4%（3/32），混合研究占比 62.5%（20/32），相对更大比例的混合研究方法反映了教育领域研究方法的发展新趋势。混合研究可以修正和弥补纯定性或定量研究的不足，并帮助研究人员更全面、深入地回答研究问题（Flick, 2018），这更有利于研究人员分析学习者在虚拟现实环境中动态复杂的二语习得过程。在 VR 辅助语言教育领域，定量研究和混合研究呈压倒性趋势，这与国际应用语言学研究现状基本吻合（肖俊敏和王春辉, 2023）。

3.2.2 数据收集

这 32 项研究采用了多种数据收集方法（如图 3 所示），包括问卷（22 项）、访谈（14 项）、测试（14 项）、语料分析（11 项）、音频材料（7 项）、视频材料（4 项）、对比组实验（4 项）、观察（4 项）。问卷、访谈和测试长久以来在实证研究中都是常用数据采集方式，在 VR 研究领域同样也不例外，此外，基于视频资料、对比组实验和观察的深度数据挖掘和分析的实证研究还比较缺乏。

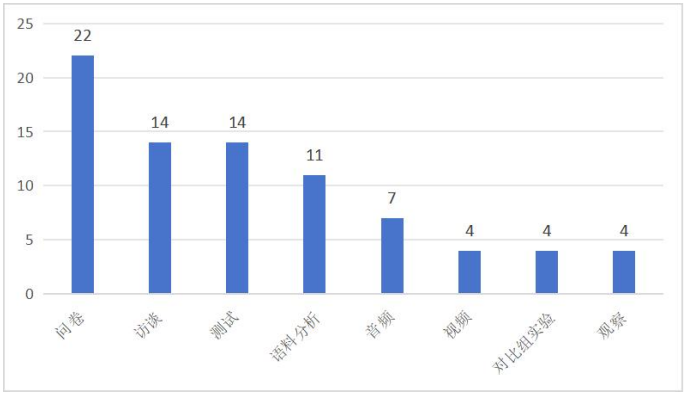


图 3　2014-2023 年国外实证研究数据收集方法频次统计

3.3 研究主题

基于内容分析法，本研究总结以下四个研究主题(如表 1 所示)，包括学习者特征(21 篇)，语言技能学习(19 篇)和非语言技能学习(11 篇)，以及 VR 的应用效果(1 篇)，其中有多篇论文同时涵盖多个研究主题(如 Saman & Marjan, 2022 等)。

3.3.1 学习者特征

关于 VR 环境中学习者特征的实证研究在发表数量上位居榜首（21 篇），由此可推测，学习者特征是国外近十年实证研究关注的热点，其中学习者动机(8 篇)占比最大，其次是学习者体验与态度(7 篇)，学习者焦虑(7 篇)，学习者投入度（6 篇）。多项研究结果表明 VR 技术对于增强外语学习者的学习动机具有显著成效（如 Kruk, 2022; Chen et al., 2020; Chen & Yuan, 2023 等）。因此，第二语言/外语教师有必要思考如何调整他们的教学方法，以更好地激发学生参与语言学习的学习动机和积极性，从而促进语言学习效果。此外，VR 技术因其特有的虚拟化身可用于减轻语言学习者的焦虑感，例如 Chen(2022)通过比较基于讲座、移动辅助和 VR 辅助的 4 周教学效果，调查了智能教育技术对降低英语学习者公共演讲焦虑的影响，其中只有 VR 辅助组达到了统计学意义。尽管绝大部分研究都集中在探究 VR 辅助语言学习的有效性，但仍然有相当一部分研究转向关注语言学习者在 VR 环境中的学习体验（7/32），这也体现未来研究可能会逐渐趋向于以学习者为中心。

3.3.2 语言技能学习

VR 环境中的语言技能提升是另一研究焦点（19 篇）。VR 技术如今已普遍应用于促进各项语言知识和技能的提升，但深广程度存在一定差异。如下表所示，AR 技术最常用于口语（7/19）和词汇学习（6/19），其次分别是针对写作（5/19）、听力（4/19）、综合语言能力（4/19）和阅读（1/19）技能的提升，此结果与过往相关综述研究（Huang et al., 2021; Chen et al., 2022; Qiu et al., 2023）一致。这也侧面证明国外 VR 辅助语言学习更注重语言的实际应用，且 VR 可以在短时间内增强学习者对语言表征和理解的学习表现，如口语学习（Huang et al., 2021）。但是针对语法技能的相关研究在本研究综述文献中并未涉及，这可能是由于 VR 环境更有助于通过创设真实情境促进听说读写等技能的习得，而语法练习对于沉浸式的环境要求不高。此外，前五年中针对语言技能的研究占比 50%（5/10），而 2019 年-2023 年中的占比增长到了 82%（18/22），这一增长进一步表明表明，VR 技术的应用越来越关注对于语言

技能提升的作用效果。

3.3.3 非语言技能学习

共计 11 篇文章探讨了 VR 环境中非语言技能的提升，这些文章主要讨论了 VR 环境在培养学习者交际能力（8/11）、认知能力（3/11）和批判性思维（1/11）方面存在的的潜在优势。交际语言教学法常被认为是第二语言/外语教学最有效的方法，其核心在于学生的参与。Saman 和 Marjan（2022）通过对伊朗英语学习者的交际意愿和口语熟练度展开了对照实验，研究 VR 工具对其产生的影响，结果显示该 VR 工具提高了参与者的交际意愿、动力和自信心。Barrett 等人（2021）调查了用户对高度沉浸式 VR 学习环境中学习英语段落写作的接受程度。他们提出，高沉浸度的 VR 学习环境可以激发学习者的想象力，从而提高他们解决问题的能力。批判性思维是一种高阶思维能力，学习者从不同的角度对作品进行判断，并做出合理的陈述或评价，过程涉及逻辑思维等技巧(Hovardas et al., 2014)。Chien 等人（2020）研究开发了一个基于球形视频的虚拟现实(SVVR)环境，研究结果证明了采用基于同伴互评的 SVVR 教学法的学生的批判性思维能力显著高于采用非基于同伴互评的 SVVR 教学法的学生。

3.3.4 应用效果

VR 技术在语言教育中的应用效果研究聚焦使用态度和学习成效两个方面。在使用态度方面，已有研究几乎一致性地表明 VR 技术为学生提供了更好的学习体验，从而提高了学生的学习成绩和学习积极性。Chen(2016）考察了 EFL 学习者在任务型虚拟课堂在二语学习中的语言学习体验，研究发现既有学生提出 VR 高交互、高沉浸的学习环境和多文化/多语言和协作环境的优点，同时也有少数同学指出 VR 仍具备一些技术问题以及缺乏副语言特征的不足。除此之外的研究中，学生对 VR 辅助语言教学的方式和效果均体现出较高的满意度(Wang et al., 2021; Lai & Chen, 2023; York et al., 2021) ，表明 VR 辅助语言教学是一种较受欢迎的技术增强型学习方式。

在学习成效方面，大多数学习者通过使用 VR 技术获得了语言或情感上的收获。Lan(2015)开展了为期一学期的 2 次迭代行动研究以评估虚拟语境对学习者英语学习的影响，结果显示 EFL 学习中使用虚拟语境可以有效提高学习者的外语成绩。Chen 和 Yuan(2023)探讨了 VR 对留学生在汉语词汇学习中的有效性，研究结果表明虚拟现实教学方法可以有效促进学生词汇的学习和记忆，并提供积极的学习体验。此外，针对听(Levak & Son, 2017)、说（Chen, 2022)、读(Guo & Lan, 2023)、写(Chen et al., 2019)等外语专项技能，VR 技术均表现出显著的积极影响。VR 技术在降低学习者焦虑感方面也有显著成效，Chen(2022)开展了为期 4 周的基于讲座、移动辅助或 VR 辅助的对比教学实验，研究证明了使用 VR 可以有效降低外语学习者公共演讲的焦虑。

表 1 2014-2023 年国外实证研究主题分类及数量

主题	子主题	研究数量
学习者特征（21 篇）	学习者动机	8
	学习者体验与态度	7
	学习者焦虑	7
	学习者投入度	6
	学习者自我效能感	2
	交际意愿	2
	学习者自主性	1
语言技能学习（19 篇）	口语	7
	词汇	6

	写作	5
	听力	4
	语言综合能力	4
	阅读	1
非语言技能学习（11 篇）	交际能力	8
	认知能力	3
	批判性思维	1

4. 总结与展望

基于以上文献分析结果，笔者对国外 VR 技术在语言教育领域的应用现状进行总结，指出研究的有待改进之处，展望研究发展方向，供相关研究者借鉴参考。

在研究对象方面，考虑到中小学学生的特点，未来的研究应更多关注于中小学教育，通过基于 VR 的学习增强学生的学习动机、满足情感需求和提高学习成绩。对于学前教育阶段的学习者来说，虚拟环境的学习也具有很大的潜力，但在虚拟环境中进行语言学习具有一定的技术难度和操作难度，未来研究者应根据研究参与者的学习需求和对技术的接受程度，改进系统功能和学习方式，以满足不同类型学习者的需求。此外，VR 应用中的教师因素也值得进一步进行探讨。

虚拟环境下语言学习的主要目标包括语法、词汇、语音/口语、阅读、写作和听力。目前国际实证研究主要集中在与发音/口语和词汇提高相关的活动上，通常是在虚拟环境中创设接近真实情境的学习情景，运用任务型学习和游戏学习等策略为学习者组织多样化的交互活动，很好地体现了 VR 的交互性和沉浸感。未来的研究应将如何保证虚拟环境下语言技能的全面发展作为研究的重点，逐步提升对于写作、语法等技能以及学生综合语言能力的关注。现有实证研究的目标语言主要以英语为主，很少有研究将其他语言作为第二语言或外语来学习，针对小语种的研究成果相对比较匮乏，这一现象同样存在于其他技术增强语言学习的研究当中，例如 AR 技术（Huang et al., 2021）。因此，未来的研究应该考虑语种类型和语言技能的多元化，以扩大 VR 辅助语言学习的多样性。

在数据收集方面，从本研究数据可知，目前国际实证研究主要采用问卷和测试方法，注重对学习者学习效果的考察，但对学习过程的关注不足，今后可多采用访谈、观察等方法收集数据，使用跟踪研究、案例分析等更深入地探究学生的学习过程。此外，虚拟场景的创建不仅基于现实生活中的场景，还基于更多可扩展的功能。研究表明，人的复杂行为和活动可以在虚拟环境中进行模拟，便于对其进行深入的记录和分析，但很少有研究者使用 VR 技术来研究学习者的学习行为，尤其是在语言学习领域（Qiu et al., 2023)。通过记录学习者的语言学习行为，研究者可以分析学习者在虚拟环境中的行为，如学习时间、交互信息、动作、姿势等(Tan et al , 2016)。因此，VR 技术创设的多模态环境在语言学习中的功能需要得到更好的延伸。

5. 结语

虽然 VR 技术在语言教育领域中的应用还处于起步阶段，但不可否认的是，VR 技术的应用为语言教育的发展注入了新的活力。本研究回顾了 2014 - 2023 年间 VR 技术在语言教育中应用的 32 项实证研究，从研究趋势、研究方法、研究主题三个方面进行编码并梳理，但由于仅基于 4 本国际期刊的相关论文展开分析，未能涵盖其他期刊论文或会议论文，可能不足以反映国际上 VR 技术应用于语言教育领域研究的全貌，因此今后可以开展更大范围的研究。

参考文献

黄甫全,游景如,涂丽娜 & 曾文婕.(2017).系统性文献综述法:案例、步骤与价值. *电化教育研究*(11),11-18+25.

马武林 & 徐梦云.(2023).21 世纪以来我国智能语言教学研究述评. *外国语文*(06),140-150.

文秋芳 & 林琳.(2016).2001-2015 年应用语言学研究方法的使用趋势. *现代外语*(06), 842-852+874.

肖俊敏 & 王春辉.(2023).虚拟现实技术在语言教育中的应用——研究现状、作用机制与发展愿景. *首都师范大学学报(社会科学版)*(05),91-105.

Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education*, 153, 103893.

Barrett, A. J., Pack, A., & Quaid, E. D. (2021). Understanding learners' acceptance of high-immersion virtual reality systems: Insights from confirmatory and exploratory PLS-SEM analyses. *Computers & Education*, 169, 104214.

Cerezo, R., Calderón, V., & Romero, C. (2019). A holographic mobile-based application for practicing pronunciation of basic English vocabulary for Spanish speaking children. *International Journal of Human-Computer Studies*, 124, 13-25.

Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, 14(6), 3147.

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education*, 102, 152-171.

Chen, Y. C. (2022). Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning*, 1-25.

Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. *Educational Technology Research and Development*, 68(4), 1769-1789.

Chen Chen & Yifeng Yuan (2023) Effectiveness of Virtual Reality on Chinese as a second language vocabulary learning: perceptions from international students, *Computer Assisted Language Learning.*

Chen, Y. L., & Hsu, C. C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers & Education*, 154, 103910.

Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education*, 146, 103751.

Flick, U. (2018). *Doing triangulation and mixed methods* (Vol. 8). Sage.

Ghanizadeh, A., Razavi, A., & Jahedizadeh, S. (2015). Technology-enhanced language learning (TELL): A review of resourses and upshots. *International Letters of Chemistry, Physics and Astronomy,* 54, 73-87.

Gruber, A., Canto, S., & Jauregi-Ondarra, K. (2023). Exploring the use of social virtual reality for virtual exchange. *ReCALL*, 35(3), 258–273.

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education*, 71, 133-152.

Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A systematic review of AR and VR enhanced language learning. *Sustainability*, 13(9), 4639.

Kruk, M. (2022). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in Second Life: The case of two advanced learners of English. *Computer Assisted Language Learning*, 35(1-2), 190-216.

Lai, K. W. K., & Chen, H. J. H. (2023). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Computer Assisted Language Learning*, 36(3), 312-345.

Levak, N., & Son, J. B. (2017). Facilitating second language learners' listening comprehension with Second Life and Skype. *ReCALL*, 29(2), 200-218.

Liang, J. C., Hwang, G. J., Chen, M. R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: an integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 31(7), 4270-4296.

Lin, T. J., and Lan, Y. J. (2015). Language learning in virtual reality environments: past, present, and future. *Educ. Technol. Soc*. 18, 486–497.

Qiu, X. Y., Chiu, C. K., Zhao, L. L., Sun, C. F., & Chen, S. J. (2023). Trends in VR/AR technology-supporting language learning from 2008 to 2019: A research perspective. *Interactive Learning Environments*, 31(4), 2090-2113.

Saman Ebadi & Marjan Ebadijalal (2022) The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency, *Computer Assisted Language Learning*, 35:8, 1975-2000.

Shadiev, R., Hwang, W. Y., & Huang, Y. M. (2017). Review of research on mobile language learning in authentic environments. *Computer Assisted Language Learning*, 30(3-4), 284-303.

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, 12(2), 524.

Tan, S., O'Halloran, K. L., & Wignell, P. (2016). Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL. *ReCALL*, 28(3), 253-273.

T. J. Lin, Y. J. Lan, and Y. H. Kan. (2013). "*Research trends of language learning in virtual reality from 2003 to 2012: A*GCCCE2024 *content analysis," in International Conference on Web-Based Learning*, pp. 73-81.

Wang, Y., Grant, S., & Grist, M. (2021). Enhancing the learning of multi-level undergraduate Chinese language with a 3D immersive experience-an exploratory study. *Computer Assisted Language Learning*, 34(1-2), 114-132.

Yan Chen, Thomas J. Smith, Cindy S. York & Hayley J. Mayall (2020) Google Earth Virtual Reality and expository writing for young English Learners from a Funds of Knowledge perspective, *Computer Assisted Language Learning*, 33:1-2, 1-25.

Yi-chen Chen (2022) Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety, *Computer Assisted Language Learning*.

York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, 33(1), 49-70.

跨越閱讀舒適圈：設計願望書單活動以培養學生的閱讀興趣

Stepping Beyond the Comfort Zone of Reading: Designing Wish Booklists Activities to Cultivate Students' Reading Interest

楊筱彤[1*]，廖長彥[2]，陳德懷[1]
[1]台灣中央大學 網路學習科技研究所
[2]台灣中央大學 客家語文暨社會科學學系
[*] jessica4625@gmail.com

【摘要】閱讀是人們獲得知識最直接而重要的管道。台灣推行「身教式持續安靜閱讀」(MSSR)已行之有年，幫助學生培養閱讀興趣與習慣，然而，多數學生容易停留在「閱讀舒適圈」，甚至形成「閱讀偏食」。因此，研究團隊曾提出「願望書單」概念，以期能解決先前研究之問題，但仍未臻完善。本研究將基於前導研究和師生進行訪談的結果，發展願望書單系統，設計輔助學生願望書單的選書機制與工具，並且在系統中設計聊書同伴機器人，提供學生自主練習聊書的機會，希望能夠為學生的閱讀提供更多可能性，促進他們跨越閱讀舒適圈。
【關鍵字】願望書單；選書機制；聊書；閱讀目標

Abstract: *Reading is the most direct and crucial channel for acquiring knowledge. In Taiwan, the longstanding practice of 'Modeled Sustained Silent Reading' (MSSR) has played a key role in cultivating students' reading interests and habits. However, many students tend to remain within their 'reading comfort zones' and may even develop 'narrow reading.' Therefore, the research team proposed the concept of "Wish Booklists" in an attempt to address previous research issues, but it has not been fully perfected. Building upon preliminary research and interviews with teachers and students, this study will develop a Wish Booklists System, design mechanisms and tools to assist students in creating wish booklists, along with a Book Talk Companion Chatbot to provide students with opportunities for autonomous practice in discussing books. It is hoped that this research can promote them to step beyond their reading comfort zone.*

Keywords: wish booklist, book selection mechanism, book talk, reading goal

1. 前言

閱讀為所有領域的學習基礎，透過吸收不同種類書籍的知識，使我們在遇到問題時能將各領域知識串聯，思考全面性的解決方案(Madhu Nair, 2019)。台灣自 2009 年起，於全台多所中小學推動「身教式持續安靜閱讀」(Modeled Sustained Silent Reading, MSSR)，以成人做為孩子們的閱讀楷模，帶動孩子持續進行安靜的閱讀活動，讓閱讀自然而然地成為生活的一部分。然而，據研究結果發現，多數學生雖然建立閱讀的基本興趣與習慣，卻存在閱讀偏食、缺乏閱讀目標的制定與規劃問題，即學生養成閱讀習慣後，在閱讀時，更傾向於選擇熟悉的主題或類型的書籍來閱讀，停留於閱讀舒適圈，久而久之，逐漸形成閱讀偏食的狀況，而由於 MSSR 強調「以興趣為本」，讓學生可以依個人喜好來挑選書籍，但若缺乏確切的閱讀目標，可能使得學生在閱讀中缺乏方向性，對於自己的閱讀歷程缺乏反思，無法滿足個人的學習需求(Victoria Wood, 2021)。

為了改善前述問題，先前研究團隊曾提出「願望書單」的概念（賴瑞霖，2021），其設計為讓教師與學生定期規劃並共同協商篩選出書單，這樣不僅能幫助學生規律地訂定閱讀目

標，選中的書籍還能符合學生的興趣，也能讓教師協助學生定期關注其閱讀狀況並適時地進行調整。該模式目前仍在桃園市某實驗學校中持續使用。然而，研究者先前針對桃園市某實驗學校的二至六年級學生以及教師進行半結構式訪談發現，現行模式仍存在著許多問題，特別是於「選書」方面缺乏明確方向與機制，以及「聊書」方面學生對於整理並表達自己想法的能力與經驗存在一定的不足。本研究將基於訪談結果發展一個電子願望書單系統，發展學生願望書單的選書機制與工具，並於系統中設計聊書同伴，讓學生得以在進行較無壓力的自發性練習。希望透過這樣的系統提升師生之間聊書的互動效果，幫助學生培養建立閱讀目標的能力，改善學生閱讀偏食的狀況。

2. 文獻探討

2.1. 聊書定義與策略

聊書為一個閱讀後的延續活動，即對自己讀過的書籍進行反思、歸納、整理過後，與他人分享的活動。透過與他人分享閱讀經驗，我們可以交流不同的觀點，學習從不同面向來看事物，並了解更多類別和主題的書籍。對於聊書較不熟悉的學生，我們認為有以下幾個常見提問或分析法可以做為鷹架，輔助學生進行聊書：

首先是 4F 引導思考法(Roger Greenway, 1992)，當中有 4 個提問重點，分別為事實(Facts)、感受(Feelings)、發現(Findings)和未來(Future)，為透過問題引導學生回顧書中的事實、感受、發現和未來展望，促進深入思考和討論；PMI 分析法(Edward de Bono, 2007)表示 Plus（正向）、Minus（負向）、Interesting（有趣），讓學生能從正、負、有趣三方面探討書籍，有助於他們對書本內容更全面的理解。最後是 ORID 焦點討論法，其包含四個層次：分別是 Objective（客觀性層次）、Reflective（反應性層次）、Interpretive（詮釋性層次）以及 Decisional（決定性層次），引導學生從不同角度討論書籍，從中獲得啟示和決定行動。

2.2. 制定閱讀目標

訂定閱讀目標對於個人發展和學習是至關重要的。透過明確的閱讀目標，個人能夠更有方向地進行學習，提升學習效果(Natalie Förster & Elmar Souvignier, 2014)。根據 Bandura(1986)的社會學習理論，訂定目標是一個主動的過程，能夠激發個人內在的學習動機，進而影響行為。在閱讀領域，Locke 和 Latham 的目標設定理論也強調，明確的目標設定能夠提高工作效率，而這個理論同樣適用於學習活動。有效的閱讀目標需要具備一定的特徵，包括具體性、可衡量性、可達成性、相關性，以及有時限性(Locke & Latham, 2002)。這表示閱讀目標應該清晰、具體，使個人能夠明確知道他們努力的方向，同時目標應該是可以量化和達成的，進而提升學習的動機。此外，閱讀目標應該與個人的興趣和長期學習目標相關，才能提高目標的吸引力和可持續性。

3. 系統介紹

本研究將學生選書介面設計如圖 1，上方提供篩選器，讓學生可以依書籍類別、書籍語言以及書籍深度，學生可以透過點擊不同的篩選條件，快速地縮小選書範圍，幫助學生更準確地找到符合其閱讀需求的書籍。下方可以查看各個書籍的簡易資訊，包含書籍封面、書名、書籍種類、書籍深度等資訊。其中，書籍種類有 3 個領域（故事小說、人文社會、科學技術），每個領域分別包含 4 小類（溫馨勵志類、冒險推理類、神話寓言類、科幻武俠類、生活類、藝術類、社會類、文學類、工程技術類、自然科學類、生命科學類、永續發展類）。書籍深度則分成了 5 個等級，以星星符號表示，1 顆星代表繪本等級，2 顆星是橋梁書的等級，3 顆

星是初階文字書，4 顆星是中階文字書，5 顆星是高階文字書。若閱讀過則會於封面圖片的右上方顯示「於 202X-XX-XX 借閱過」。

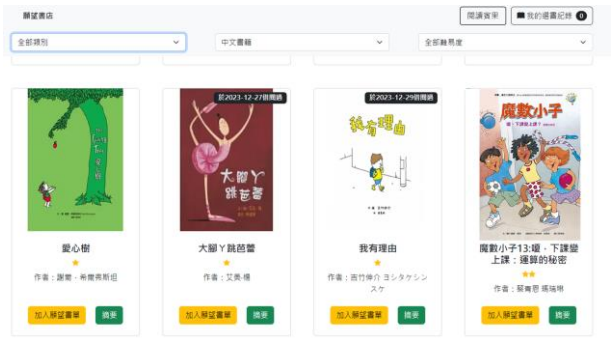


圖 1 學生選書介面

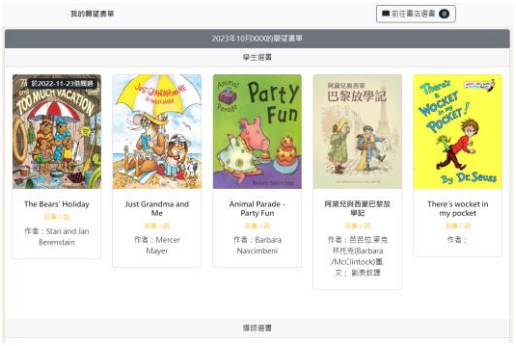


圖 2 學生的願望書單介面

學生亦可點擊右上方的「閱讀實果」按鈕，此閱讀實果根據學生的閱讀歷程動態生成，同時考量各年級的學習主題，將學生較少閱讀種類或深度，以及隨機從各年級的學習主題挑選，做為實果內容。學生可透過挑戰實果遊戲，同時也可以從多元角度思考書籍的選擇，做為願望書單選書的參考依據。點擊畫面右上方的「我的選書紀錄」，會跳至學生的願望書單介面（圖 2），可於此處查看已被選中的書籍，並且提供歷史願望書單內容，依月份排列供學生查看。

針對學生聊書能力的強化，我們設計了一個聊書同伴（圖 5），我們採用前面文獻提及的 4F 引導思考法、PMI 分析法、ORID 焦點討論法三種提問策略，將策略定義與書籍摘要做為 Prompt，使用 GPT 輔助題目的設計。學生可自由選擇想聊的書和想練習的策略，書本的選擇包含它願望書單中已借閱過的書籍以及學生該月份已借閱完畢之其他書籍。選擇好策略與書籍之後，便可開始與系統進行聊書。以圖 5 為例，若該學生選擇 4F 引導思考法以及《聰明的波麗和大野狼》做練習，聊書介面將會顯示如圖 6，該題目運用 4F 的第一個提問重點「Facts（事實）」命題，第二題則會運用「Feelings（感受）」出題，以此類推。題目除了以文字敘述，同時也播放內容，並且提供重播功能，增加擬真度。學生可於下方回答區塊進行錄音、回放、重錄等功能，透過儲存學生的聊書語音，使用 whisper AI 將聊書語音轉文字，紀錄並使用 GPT AI 將聊書狀況進行摘要，以及初步分析聊書結果。



圖 3 聊書同伴機器人起始介面

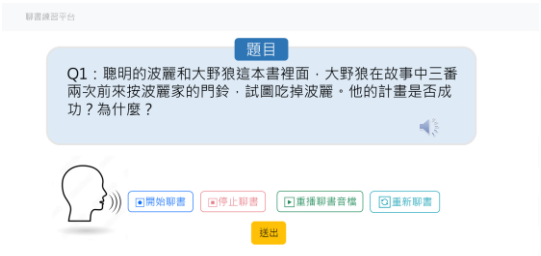


圖 4 聊書介面

4. 研究方法

本研究之研究對象為桃園市某實驗學校的國小二到六年級，共計 89 名學生。研究參與者皆已實施課堂 MSSR 活動以及紙本願望書單活動多年，並且曾有聊書的經驗。本研究目前已對研究對象之學生和教師進行訪談，了解師生於目前模式進行下遭遇的困難以及需要幫助的地方。以蒐集到的資料將願望書單執行流程歸納為四個環節（圖 5），依各環節進行資料的整理與分析，發現學生對於選書沒有明確方向、教師在書籍推薦環節中較無影響力以及與教

師聊書時因為不熟悉或容易緊張而沒能將自己的想法完整的表達出來。因此，本研究將就上述問題改善願望書單活動流程與機制，發展電子願望書單系統。使用願望書單系統輔助願望書單活動的執行，觀察並評估其對學生選書、聊書能力以及閱讀偏食改善之影響。



圖 5 願望書單執行流程圖

5. 預期結果

本研究旨在探討與評估使用電子願望書統對學生選書、聊書能力以及改善閱讀偏食情形的影響。研究者使用多位語文教師們共同推薦的書籍做為系統的候選書單，書單中涵蓋前文提及的 3 大領域與 12 小類，共計 4427 本中文書以及 2300 本英文書，提供學生更多元的閱讀選擇，滿足其各種興趣和需求，並將學生閱讀歷程依類別、深度視覺化，讓教師可以根據個別學生的閱讀歷程表現覺察學生的閱讀偏食情形，再進行書籍的推薦以及給予學生調整的方向與建議；系統提供篩選器以及一目瞭然的書籍資訊，以期能幫助學生於大量書籍中能更快速地找到符合自身需求的書，並運用個人化的閱讀賓果促進學生從多個需求面向思考閱讀的選擇，期望能協助學生規劃閱讀目標，進行足量書籍的閱讀；透過使用聊書同伴讓學生無須受時間限制，可以於任何時候進行自發性的聊書練習，使他們能更自在地表達自己的想法，克服可能因為與他人交流而產生的壓力或害羞感，進而提升他們的自信心。透過平台的初步分析資料，讓學生能夠快速瞭解自己的聊書表現和結果，從而調整和改進自己的表達方式。

致謝

本研究在臺灣國科會人文處（NSTC 111-2410-H-008-013-MY3）與「中央大學學習科技研究中心」的資助下完成，僅此致謝。

參考文獻

賴瑞霖（2021）。閱讀島：應用數位閱讀歷程擴大學生閱讀舒適圈及提升閱讀動機之系統設計。〔未出版之碩士論文〕。國立中央大學網路學習科技研究所。

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ, 1986*(23-28).

Bengel, J. (2014). *Speaking and Listening Skills With a Book Talk.* Retrieved from https://www.outofthisworldliteracy.com/speaking-listening-skills-with-book-talk/

De Bono, E. (2007). *Tactics: The art and science of success.* Profile Books.

Greenaway, R. (1992). Reviewing by doing. *Journal of Adventure Education and Outdoor Leadership*, 9(1), 21-25.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9), 705.

Nair, M. (2019). *Why Is Reading Important for Your Growth?* Retrieved from https://www.uopeople.edu/blog/why-its-important-to-read/

Wood, V. (2021). *How To Set and Achieve Your Reading Goals.* Retrieved from https://www.uopeople.edu/blog/why-its-important-to-read/

5G 时代网络不良文化对儿童行为与认知发展的影响及对策研究

The Research on the Impact of Unhealthy Internet Culture on the Development of Children's Behavior and Cognition in 5G Era

吴郑涵[1]，杨清元[2*]，杨玉辉[2]，赵亚萍[2]，胡为民[2]
[1]杭州师范大学 人文学院
[2]浙江大学 信息技术中心
[*]yangqingyuan@zju.edu.cn

【摘要】 5G 时代“网络梗”在儿童成长的过程中，潜藏着一系列问题与挑战，如孤立霸凌、价值观畸形、群体行为的极端化等。随着网络梗的广泛传播，儿童认知发展所面临的潜在影响引起了各界的密切关注。研究通过探究 5G 时代网络梗对儿童行为与认知发展的影响机制，重点聚焦于梗的传播模式、群体效应以及信息循环强化等方面，系统分析网络不良文化导致儿童群体极端化行为以及负面思维模式的成因。同时，研究从学习共同体的视角出发，围绕学校、家庭和社会三方，讨论相关方面的引导和监管的策略与措施，并提出相关干预方案，以期为该领域的研究提供理论基础和实践参考。

【关键词】 网络文化；数字土著；认知发展；群体行为

Abstract: *In the 5G era, unhealthy internet culture poses a series of problems within the growth of children, such as social isolation, bullying, values distorting, extreme group behavior, and so on. The widespread dissemination of the bad culture has raised significant concern among various sectors about the potential impact on the development of children's behavior and cognition. The study focuses on the mechanisms of culture transmission, group effects, reinforcement of information cycles, and other relevant aspects, which systematically analyzes the causes of extreme behaviors and negative thought patterns among children. Additionally, the study adopts a perspective of the learning community to discuss the regulatory strategies, which also proposes some intervention plans that aim to provide a theoretical foundation and practical reference in this field.*

Keywords: internet culture, digital natives, cognitive development, group behavior

基金项目：教育部产教融合协同育人项目（NO.220901214230904）；浙江省“十四五”研究生教育改革项目“数智驱动跨学科创新人才多模态融合培养支撑体系构建”。

1.引言

根据《说文解字》的记载，梗被定义为“山枌榆。有朿，荚可为芜夷者。从木，更声”，其本义为一种带刺的植物(许慎, 2004)。如今互联网中普遍存在的“梗”则被视为“哏”的误读、误解(徐娜娜, 2023)。而“烂梗”，则通常被认为是缺少内涵与品味、被不合时宜滥用的网络表达，通过泛娱乐化和情绪化的方式在网络上传播(田萍, 2020)。随着 5G 时代的来临，数字原生代儿童在互联网文化中的参与越发显著。然而，伴随着网络梗的广泛传播，儿童认知发展所面临的潜在影响，如孤立霸凌、价值观畸形、群体行为的极端化等，引起各界的关注。因此，梳理清网络梗对儿童认知发展的影响机制、梗的传播模式、群体效应以及信息循环强化等内容，对全面理解并有效干预消极影响具有重要意义。

2.时代背景

5G 时代线上教育和网络娱乐蓬勃发展，儿童可以在线获取学习资料，通过社交媒体深度融入网络空间，成为数字土著，即深度融入互联网的一代人。如何充分利用网络，以促进儿童进行良性社交，是一个备受关注的议题。另外，儿童通过 5G 网络接触到更为丰富的信息内容，对其认知和行为方式产生更复杂的影响。在娱乐层面，互动媒体为儿童提供了更具参与

性的体验，同时也引发了对虚拟世界的过度沉迷。此外，家庭与社会关系的变化也不可忽视，5G 网络改变了家庭和社会交往的方式，对儿童的社交关系和家庭的良性互动产生了新的挑战。

3.形成原因

5G 时代网络不良文化正对儿童本身、教育共同体的引导和监管三方面造成挑战。由于儿童容易受到外部环境的深刻影响，致使信息隔离效应、排斥孤立和群体价值取向极端化等现象进一步强化了儿童趣缘群体的异质集聚。在教育引导方面，学习共同体理念尚未充分实践，学校和家庭教育存在引导不足、倾听不足以及榜样示范的缺陷。此外，学校和家庭的监管存在空白，网络监管在细致分类管理上存在缺陷。这些问题综合影响着儿童的认知能力发展和整体成长，需要采取综合而有力的措施，以促进儿童在良好成长环境中全面发展认知能力。

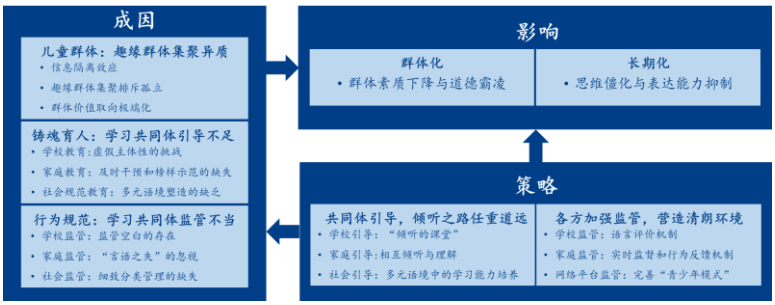


图 1 5G 时代网络不良文化对儿童行为与认知发展的影响及对策

3.1. 儿童群体：趣缘群体集聚异质

在儿童的世界观、人生观和价值观形成时期，个体的自我调节能力相对有限，同时自我效能感正处于建构的阶段(余俊渠, 2020)。这意味着儿童在应对外部刺激和情境变化时，更依赖外部环境的引导和规范。缺乏自我调节能力可能使得儿童更容易受到周围环境的直接影响。同时，网络不良文化的影响下的恶性集聚是儿童认知能力发展阻碍的主要原因之一。

信息隔离效应。当某一“烂梗”在社交媒体上持续传播时，容易形成泛滥现象。这种信息在特定群体内形成了封闭的循环，导致群体内观点和信息的多元化和平衡受到阻碍。这种信息隔离效应使得群体内的认知逐渐变得单一，阻碍了多种观点和信息的有效交融。

趣缘群体集聚排斥孤立。为获得认同，个体愿意抛弃是非，换取让人倍感安全的归属感(庞勒，2015)。当一个烂梗在儿童群体中传播甚广，将会不断产生对群体之外儿童个体的文化孤立和爽感吸引。即使儿童个体自身对此梗没有过多的兴趣，也会对此进行学习和模仿，以换取趣缘群体的认可和归属感。在不断的排斥孤立中，烂梗的传播速度和强度会逐渐升高。

群体价值取向极端化。首先，儿童群体认知水平和辨别能力不成熟，易受单一信息影响，导致价值极端化。其次，缺乏正向规范引导，庸俗价值可能消解高尚价值。信息传播偏向哗众取宠，负面信息更易吸引儿童，加剧价值极端化趋势，影响认知发展。

3.2. 铸魂育人：学习共同体引导不足

教师、家长和社会联合参与学习共同体正促进为学生提供积极的发展环境的教育生态形成（印贤文，2020）。然而，学校虚假主体性的课堂氛围导致教师缺乏倾听和引导学生思考的能力，阻碍了学生真实状态的表达；家长的参与不足和缺乏良好的榜样示范也限制了学生对不当行为的认知；社会规范不足和网络监管漏洞使学生在多元语境中难以获得正向的指引。

学校教育中虚假主体性的挑战。以心理辅导班会课为例，传统的班会课主要通过教师对某一问题，如烂梗在全班风靡，进行心理辅导和正向引导。然而，这种模式通常处于“虚假的主体性”之中（陈静静，2018），即课堂氛围充满热闹的虚假快乐，缺乏思维上的挑战以及真正的倾听和情感交流。教师更侧重于“教”而不够注重“听”，导致学生缺乏理性思考、情感交流与批判性思维。根据佐藤学（2014）的思想，要在课堂中实现真正的学习共同体转型，教师需对儿童学习的深入理解，锤炼自身倾听的方法，并对先前教学方法的深刻反思。

家庭教育中及时干预和榜样示范的缺失。家长或许未重视及时改正孩子消极行为，造成负

面影响。同时，批评方式可能过于极端，缺乏耐心疏导与理解，妨碍有效沟通与教育。甚至一些家长未能做出良好示范，反而成为孩子负面行为的模仿对象。另外，家长未积极参与学校教育，未能深入了解孩子语言表达或其他问题的触发因素。

社会规范教育中多元语境塑造的缺乏。低年龄段儿童教育常限于“普通话”和“规范字”，未充分关注语言多样性，需更深入融入社会实践、文化活动，全面激发儿童学习潜力。社会中不良示范可能影响儿童，所以应全面培养儿童社会规范素养，注重“相互倾听和理解”。

3.3. 行为规范：学习共同体监管不当

在学校监管中，教师直接监管和施行班干部的间接监管存在一些不足。在直接监管时，可能缺乏有效的监管手段和方法，导致在学生的某些问题行为较难得到及时纠正。因此，为提升监管的效力，要深入剖析教师直接和间接监管的空白之处，更好地引导和管理学生的行为。

部分家庭在孩子有意或无意表达烂梗时，缺乏及时的纠正和批评，这种对孩子不当行为的忽视可能会强化该行为，并在其未来成长中造成持续性负面影响。因此，强调家长在及时发现、纠正和批评孩子不良行为方面的作用，对于孩子的良好成长至关重要。

社会监管细致分类管理的缺失。网络主体的虚拟性和自由化致使在针对低龄儿童的网络监管方面存在明显不足。如网络信息缺乏对这一群体进行分级、分类和分层管理的有效手段。这使得儿童在网络空间中容易接触到与年龄阶段不相适宜的内容，缺乏足够的保护和引导，从而增加了不良信息对其产生负面影响的可能性。

4.负面影响

三言两语怎会成“毒”？“毒”在何处？烂梗使用者的集群化严重影响儿童认知发展。第一，儿童容易受到庸俗影响，削弱了理性和批判性思维能力。第二，烂梗群体化减弱了个体的自责意识。第三，群体排他性导致狂欢趋势不断升级，可能最终演变为群体性霸凌行为。

长期来看，思维模式固化、表达能力抑制。通常，一个烂梗从创造到消亡的周期仅数天或数月。然而，海量的新梗不断涌现，延续梗的生命力至下一个烂梗之上。这种惯性思维和新热词暗示着避免思考的倾向，这可能导致儿童思维模式僵化、语言表达抑制、价值观畸形发展。儿童快速学习能力加速同质化行为。

5.策略与措施

不管是如何偏执而狂热的烂梗，即使已经得到了儿童群体普遍崇拜，却仍摆脱不了其虚无主义的本质，而唯一能让它消亡方法就是使其接受理性的思辨和长期的健康引导。

5.1. 教育教学引导，倾听之路任重道远

学校、家庭和社会的教育引导共同构建了一个以“倾听”为核心理念的教育体系，为儿童的全面发展奠定了积极而深远的基础。学校采用学习共同体式的班会，发挥教师的关键角色在培养儿童对网络文化理性认知方面起到了关键作用(陈静静, 2018)。同时，家庭教育通过家长的倾听、悉心引导以及自身示范作用，为儿童形成理性的网络认知和良好的网络行为习惯提供支持(孙海燕，2019)。社会教育引导强调对儿童多样性和潜能的挖掘，通过社会实践和文化活动培养儿童积极素养，为其认知水平和社会适应能力的提升奠定基础(王晓丹, 2020)。

学校通过引入学习共同体式的班会教学，能够创设一个有利于学生集体思考和合作交流的环境，不仅有助于培养巩固学生的道德观念，也提供平台让学生审视网络潮流并形成对烂梗的审辨性思维。此外，教师应该秉持学习共同体的思想，为学生创设安心、安静的课堂环境，改变忽略倾听学生的习惯，成为学生倾听的示范者，并能长时间地耐心支持(张鲁宁，2018）。通过言行一致，教师能够向学生传递正确的价值观，成为学生认知引导的榜样，引导其形成积极健康的网络行为习惯,建构学生主体性心理课堂(黄洪霖，2017)。

家庭教育应以循循善诱的方式探寻孩子使用烂梗的原因，理解其背后动机，并帮助他们认识烂梗的危害，提高语言修养和鉴别能力。其次，家长需以身作则，示范良好的语言表达和道德习惯，成为儿童认知发展的引导者。最后，家长应积极参与学校教育，深入了解孩子语言表达及其他问题的触发因素，将支持融入学校发展，促进双方建立互相理解的关系。

社会教育应重视儿童多样性和潜能的挖掘，通过社会实践和文化活动激发其多元学习能力。另一方面，社会规范教育需培养儿童积极素养，通过正向宣传传递积极价值观，提升儿童认知和社会适应能力，奠定道德基础，促使其全面发展。而相互倾听和理解的普及有望促进社会教育规范，推动人际关系和社会互动注重合作与理解，促进社会进步。

5.2. 加强各方监管，营造清朗环境

学校、家庭和网络平台的监管需考虑儿童特点，通过学校评价机制、家庭监督反馈和网络平台青少年模式，共同营造积极网络环境，促进儿童健康成长。

学校可以建立学生间相互监督和奖惩分明的机制，将言语习惯纳入评价体系。这一评价机制应以互相监督机制为重点，使学生在校内相互提醒和纠正，并将言语表现及时与家长分享。

家庭监管有助于培养儿童良好语言习惯和社会价值观。家长应及时发现并纠正孩子不良行为，尤其是频繁使用烂梗和污言秽语，进行有效监督和惩戒。家长可采用开放沟通，了解原因和动机，表达反感态度，并给予孩子批评和教育，引导其正确语言表达。

网络平台需加强审核机制，完善青少年模式以精准管理儿童信息；推送适宜内容，避免不适信息影响儿童身心发展。同时整顿劣迹网红，保障信息质量；建立智能审核系统，培养专业审核团队，构建安全网络环境，提供有益信息资源。

6.总结

在儿童认知发展的关键时期，个体受到趣缘群体集聚异质的负面影响，主要表现为烂梗的恶性集群泛滥和儿童思维模式的长期侵害，导致其语言表达能力下降、思维方式固化，丧失对深度理解能力。分析结果显示，儿童自我调节能力的有限、认知水平和辨别能力的不成熟，以及群体信息隔离效应、趣缘群体排斥孤立和群体价值取向极端化促使这一现象的发生。此外，学习共同体引导不足、监管疏漏等因素也综合影响了儿童的行为判断。

对此，本文提出一系列策略。在教育引导方面，学校应创设倾听氛围，教师示范正确价值观，家长以循循善诱了解孩子使用烂梗原因，积极参与学校活动建立相互倾听关系。社会教育强调多样性挖掘，促进学习能力。在加强监管方面，学校建立语言评价机制，家长实施实时监督反馈机制，社会加强网络平台审核，完善青少年模式，保障网络信息质量和健康。

通过对形成原因的深入分析、负面影响的明确阐述以及策略的提出，研究为引导儿童在网络文化中培养良好的认知发展能力和语言行为规范提供了一些积极有效的建议和方法。但这些策略与措施尚未经过实证研究的检验。未来的研究方向可以更加深入地探讨提出的策略在实践中的效果，以及进一步挖掘其他可能的因素和解决方案。

参考文献

许慎. (2004). *说文解字*. 北京: 中华书局.

徐娜娜. (2023). 流量时代“梗”的多维呈现及引导策略. 南京邮电大学学报(社会科学版), 25(01), 66-75.

田萍. (2020). 由误用到走红的“梗”族. 语文建设, (20), 76-78.

余俊渠. (2020). 网络违法不良信息与青少年越轨行为探析. 广西民族大学学报(哲学社会科学版), 42(03), 155-160.

庞勒，G. (2015). *乌合之众* (高山译). 北京：新世界出版社.(原著出版于 1895)

印贤文. (2020).为成长赋能：在选课走班中构建师生学习共同体. 中小学管理(10),53-54.

陈静静. (2018).佐藤学“学习共同体”教育改革方案与启示. 全球教育展望(06),78-88.

佐藤学. (2014). *学校见闻录——学习共同体的实践* (钟启泉译). 上海：华东师范大学出版.

孙海燕. (2019). 班级学习共同体及其构建策略研究（硕士学位论文, 曲阜师范大学).

王晓丹. (2020). 佐藤学“学习共同体”思想及其实践研究（硕士学位论文, 福建师范大学).

张鲁宁. (2018). 佐藤学“相互倾听课堂”理论及启示. 当代教育科学, (06), 76-79.

黄洪霖. (2017). 佐藤学的合作学习理论及其镜鉴. 福建基础教育研究, (05), 4-7.

大语言模型驱动的论证导学探索性研究

Exploratory research on argumentation guidance driven by large language model

王红晔[1]，张立山[2*]，吴涵[3]
[1]华中师范大学
[2]华中师范大学
[3]华中师范大学
[*] 1196414083@qq.com

【摘要】 近年来，向学生教授元认知技能，如批判性思维，已成为教育工作者的核心兴趣。一篇具有说服力的议论文要求作者逻辑清晰合理、单词使用正确，因此议论文写作被认为是所有写作体裁中要求最高的写作活动。然而，有研究表明，大多数青少年并不擅长论证，他们经常在撰写议论文、理解书面论点或论据等方面遇到困难。因此，本文以大学生为研究对象，以议论文为研究载体，对比预训练语言模型和新兴的生成式大模型在论证结构解析方面的性能指标，结果表明，与预训练语言模型相比，生成式大模型技术更适合应用于学生议论文写作辅导系统中。

【关键词】 **议论文写作辅导；预训练语言模型；生成式大模型；论证结构解析**

Abstract: *Teaching metacognitive skills, such as critical thinking, to students has become a core interest for educators. A persuasive argumentative essay requires the author to be reasonably logical and use correct words, so argumentative essay writing is considered the most demanding writing activity among all writing genres. However, research shows that most teenagers are not good at argumentation. They often encounter difficulties in writing argumentative essays and understanding written claims and premises. Therefore, this paper takes college students as the research object and argumentative papers as the research carrier to compare the performance indicators of pre-training model and emerging generative large model in argument structure analysis. The results show that generative large model technology can be applied to In the student argumentative essay writing tutoring system.*

Keywords: Argumentative essay writing assisting, Large-scale pre-trained models, *generative large model, argument structure analysis*

1. 引言

论证是我们理性地解决问题、处理争端的手段，在学习环境中嵌入和培养论证思维活动可以促进富有成效的批判性思维。教会学生批判性思维，其中一个关键的方法是教学生写议论文，一篇好的议论文能够试图说服读者在有争议的问题上同意作者的观点。

使用自然语言处理技术自动检测论文中的论点、论据对于教师和学生非常有价值(Stede, Schneider, & Hirst, 2019)。当将其纳入论文中的校正或自动检测算法时，它可以帮助教师改进修正过程，减轻老师修改论文的负担；对于学生来说，当将其纳入到议论文写作反馈辅导系统中时，它即可以向学生清晰地呈现文本的论证结构，又能给学生提供论证参考。此外，它使得议论文评估程序能够更大规模地应用，为许多学科的教学工作提供指导。

目前提取论证成分来辅助学生进行论证写作的自适应反馈系统(Wambsganss 等, 2020)，以学生写出一篇初步的论证文稿为前提，在此前提下给学生提供一些自适应的反馈信息

(Wambsganss 等, 2020)。虽然自适应反馈信息一定程度上能帮助学生提升论证文本写作技能，但忽视了学生自身的写作水平。对于写作水平低的学生仍然很难写出一篇较完整的议论文，因为他们缺乏预先的知识储备、缺乏论点论据参考资料的储备。因此，这些系统仍然难以全面评估计算机支持的论证教学对学生议论文写作的具体影响，尤其是对于写作水平低的学生。

除了提供学生写作完成后的反馈信息外，我们的研究致力于将特定主题的论证成分作为预先参考资料储备在大学生论证写作辅导系统中，这样学生在写作过程中能够自由选择是否需要参考资料的帮助。开发该系统之前，本文先对预训练模型和 GPT4 在抽取论点论据上的性能做了对比和评估，为将人工智能技术加入到论证写作辅导系统中做一个选择参考，本文的研究工作主要聚焦于探讨哪项人工智能技术提取的论点论据更能有效应用于实际的论证教学中。

2. 文献综述

2.1. 论证教学的相关研究

虽然一些研究人员认为，论证的技能应该在儿童和青少年（6 至 9 年级）之间培养，大学生已经获得了很多论证技能(Felton & Kuhn, 2001)，但大多数其他研究人员却认为，即使是大多数年龄较大的学生也不擅长辩论。例如，(Reznitskaya 等, 2001)表明，青少年的在撰写议论文时的问题大多数青少年学生不理解争论性话语，他们在撰写有说服力的文章、理解书面论证方面常常面临困难，主要表现在难以区分理论和证据；难以写出真实的证据或替代理论。在传统的课堂教学中，大多数正式课程都缺乏对论证学习的支持，学生只是在学习过程中通过与同学和老师进行简单地互动来学习论证。此外，教师教授论证的能力也很有限，受到时间和可用性的限制(Scheuer, Loll, Pinkwart, & McLaren, 2010)。因此，议论文教学成为传统课堂教学中的一大难题。

(Jonassen & Kim, 2010)等人认为当学生参与到基于项目或基于问题的学习环境时，并且这个学习环境提供论证的合法替代方案，那么学生将会更加成功地、更加有效地学习到论证技巧。相反，被要求记住现存论证方案的学生没有理由进行论证，因此不太可能取得学习成果。换句话说，需要给学生提出一个有争议性的问题，以供学生思考，明确自己的立场；然后在该学习环境下为学生提供一些可以参考的论证方案，以供学生借鉴。没有提供论证参考的学生群体，他们必须直接在脑海中解决问题，这可能涉及相当大的工作记忆负荷，因此不太可能取得学习成果。然而，论证教学的方法是有限的。(Jonassen & Kim, 2010)也指出了论证教学中的三个主要挑战：“教师缺乏在课堂上促进论证的教学技能，因此缺乏练习论证的机会；外加教学材料的压力，没有时间发展技能；以及缺乏关于论证的先验知识。”

教育技术领域中的一些研究人员将论辩挖掘算法应用于辅导学生论证写作中，他们使用论辩挖掘算法从无结构的文本中抽取论点论据。这些信息可用于对文本的论证水平进行评分，并提供关于文本说服力的适应性辅导和反馈 例如 ArgueTutor(Wambsganss, Kueng, Soellner, & Leimeister, 2021),作为一个提供自适应反馈的论证写作系统，致力于辅导大学生写出一篇具有说服力的评审文本。以同行评审为学习场景 ，使用论辩挖掘算法自动识别大学生文本中的论点和论据，在写作过程中根据学生需求提供论证理论解释的帮助，以文章中论点论据的数量、文章的可读性、自适应反馈话语作为文本评价信息反馈给学生。

2.2. 论证挖掘的相关方法&算法

大多数研究人员为了克服论证教学的困难，探索了很多方法，早期的研究主要集中在人身上。随着计算机技术和人工智能技术的发展，人们逐渐尝试使用特征工程、机器学习、深度学习、大规模预训练模型等方法来探索论证教学。其中，手动提取论证结构是一个熟练且耗时的过程。例如，罗伯特·霍恩在谈到他在“计算机是否可以思考？”的论证中制作的论证图

时，引用了一名学生的话说“这些图可以为我研究生第一年节省 500 个小时的时间”；然而，Metzinger 指出，Horn 和他的团队需要 7,000 多个小时的工作才能创建该论证图(Lawrence & Reed, 2020)。

随着自然语言处理技术的发展，研究人员逐渐使用论辩挖掘算法——自然语言处理领域一个新颖的研究方向，来探索论证教学。论辩挖掘处理的是从无结构的文本中自动地识别论证结构。(Palau & Moens, 2009)认为，论证挖掘需要检测论点和论据以及他们之间关系。(Stab & Gurevych, 2014)创建了 PE 数据集，这是一个用包括论点、论据以及攻击或支持关系的方案注释的语料库。(Eger, Daxenberger, & Gurevych, 2017)通过将该问题视为序列标记问题，在 PE 数据集上实现了最先进的性能，他们通过使用(Miwa & Bansal, 2016)引入的 LSTM-ER 模型的修改版本，使得该任务具有 61.67%的最佳准确度，该模型使用了基于序列和树型 LSTM 的堆叠架构。

随着深度学习和预训练模型的发展，(Wang, Huang, Dou, & Hong, 2020)从议论文的结构特征出发，提出了一个多尺度、多层次论证挖掘模型，使用 BERT 编码器对文本进行编码，并对三种成分的识别(主要论点、论点、论据)提出了相应的解码策略，使用指针网络的方法抽取主要论点和论点，同时将抽取论据视为序列标注任务。他们在 PE 数据集上获得了最高 F_mico 得分为 66.93%的最佳性能。(Sazid & Mercer, 2022)等人通过将论证挖掘任务定义为一个多标签问题，以统一和解耦的方式处理论证挖掘问题，使用神经网络构建了 Unified-AM 和 Decoupled-Solution 两个深度学习模型(核心使用了 BiLSTM 和多头注意力机制)。为了将所有子问题（议论文和非议论文单元分类；主要主张、论点和论据成分抽取和分类；关系识别）整合到一个问题中，他们为目标标签构建了一个大小为 33 的二元向量，在 PE 数据集上实现了 68.02%的最高 token 级准确度。

论证成分的识别和分类是议论文写作辅导系统的一个最基础的技术问题，当需求量非常大时，目前用于提取论证成分的技术仍在冷启动、零样本等问题上仍然面临着挑战。ChatGPT、GPT4 的发布给论证教学带来了新的机会，该技术在许多工作流程和日常生活中的应用引起了人们的广泛兴趣，在语言、沟通和信息至关重要的教学中也是如此。因此，本文的研究重点在于对比和评估了预训练语言模型和生成式大模型—GPT4 在抽取论点论据上的性能，为将人工智能技术加入到论证写作辅导系统中做一个选择参考。

3. 实验设计

3.1. 数据集描述

本文沿用了 2.2 节中提到的 PE 数据集来验证 GPT4 在论证成分的识别和分类上的性能，PE 数据集由(Stab & Gurevych, 2014)创建，在 2017 年对数据集进行了更新(Stab & Gurevych, 2017)。他们从 essayforum.com 上随机选择了 402 篇带有写作问题描述的英文文章。essayforum.com 在线论坛是一个活跃的社区，提供对研究论文、散文或诗歌等不同文本的更正和反馈。例如，学生在准备标准化语言考试时，可以在该论坛发布他们的文章，以获得关于写作技能的反馈。这些文章都是关于有争议的话题写的，这样作者就可以发表自己的观点和立场，该语料库已标记有成分类型、立场以及从论据到论点或从论据到论据的距离，该数据集包含 1587 个段落，其中训练集中有 105988 个标记；测试集中有 449 个段落，29537 个标记；该验证集有 12657 个标记，分为 199 个段落。

论证结构可以被视为一片森林，每棵树都植根于作者的一个主要主张(Major Claim)，论点(Claim)与所有具有“赞成(for)”或“反对(against)”关系的主要主张有关。论据(Premise)仅与一项论点或论据相关，论据“支持(support)”或“攻击(attack)”某个论点或论据。值

得注意的是，除了论点和主要主张之间的关系不在一个段落之外，其他成分的论证结构完全包含在一个段落中。正如(Eger 等, 2017)所提到的，该语料库是不平衡的。

3.2. 方法介绍

本文以 BERT 作为基线模型，在此基础上使用自然语言处理的方法，来提升论证抽取的整体性能。同时使用提示词工程探索 GPT4 在论证抽取方面的整体性能。具体方法流程如下图所示：

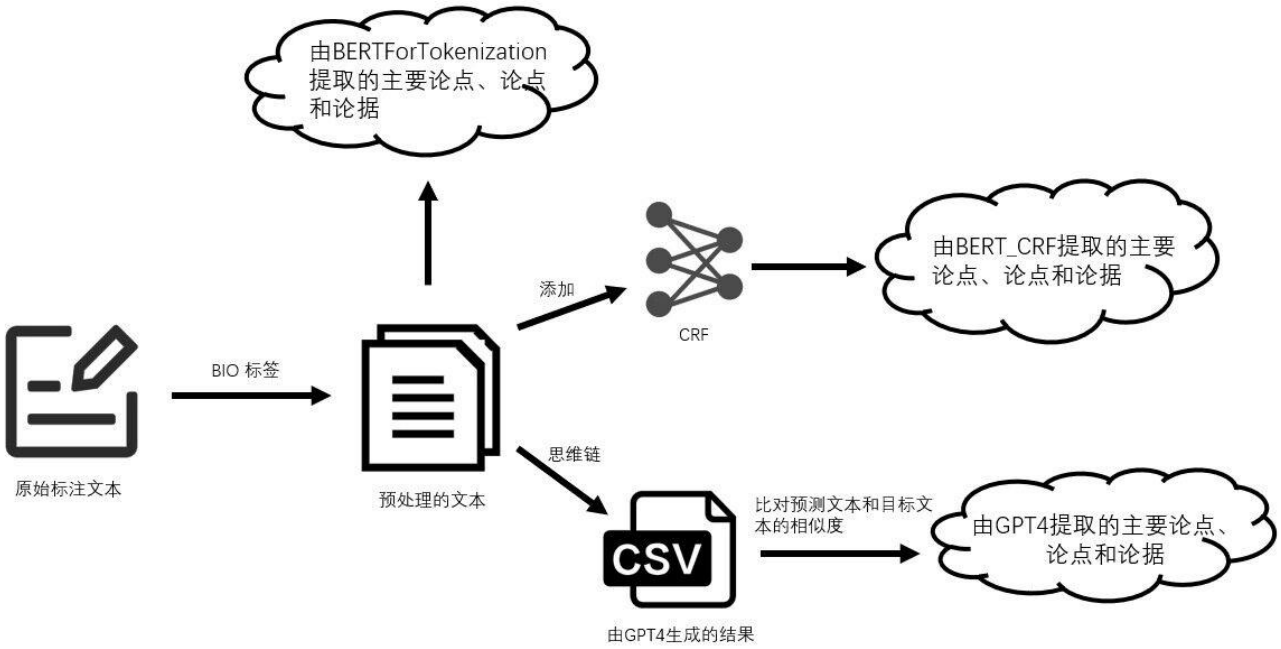


图 1 方法流程图

3.2.1. BERT

使用 BERT 语言模型得到文章的语义向量，能够捕捉更多的内在信息，提高了预测的精度。BERT 模型的全称是 Bidirectional Encoder Representations from Transformers，BERT 是通过随机遮蔽和下一句预测的训练方式得到的语言模型，BERT 旨在通过联合调节所有层中的上下文来预先训练深度双向表示。BERT 的基础集成单元是 Transformer，但其语义理解能力要优于 Transformer 等其他模型，预训练的 BERT 表示可以通过一个额外的输出层进行微调，适用于广泛任务的最先进模型的构建，比如文本分类、文本匹配、问答任务和语言推理等，无需针对具体任务做大幅架构修改。其中 BertForTokenClassification 模型用于序列标注任务，对文本序列的每个单词做分类。通常输入是一段文本，输出是每一个 token 的标签属性。BertForTokenClassification 是在 BertModel 的输出上增加一个分类层，即依次经过一个 Dropout、全连接操作和非线性激活输出类别。由于每个单词都要输出分类，因此不需要 BertModel 的 pooling 层。

3.2.1. Generative AI - GPT4

GPT-4（Generative Pre-trained Transformer 4）是 OpenAI 发布的最新 GPT 系列模型。它是一个大规模的多模态模型，可以接受图像和文本输入，产生文本输出，输出任务依旧是一个自回归的单词预测任务。整体来说，GPT-4 的能力已在各种专业和学术基准上表现出了人类的水平，包括以大约前 10% 的成绩通过模拟律师资格考试。相比其他 GPT 模型，GPT-4 毫无疑问是目前最强的文本生成模型。模型参数量估计约为 10 万到 100 万亿量级。

本文首先采用“chain of thoughts”的提示方法给 GPT4 输入提示词，先后使用了两个提示，分别如下：

表 1 提示词

Chain of thoughts	Prompt text
Prompt1	'''You now need to extract claims and premises from the argumentative essays written by students. The specific steps are as follows: 1. Read the article thoroughly to understand the author's main position on a controversial topic; 2. Output all claims. The purpose of using claims is to provide reasons for the author's main position. Before the author expresses his claims, there will be some precursor words, such as "I think", "accordingly", "as a result", "hence", "I believe that", etc. 3. Output all premises. The purpose of using premises is to explain reasons. Before the author expresses his premises, there will be some precursor words, such as "for example", "the reason is...","because", "since", etc. 4. Remember: You only need to output the original text fragment containing the above two components, no need to summarize or modify them by yourself. The output format needs to be the format of the dictionary in python, the original text fragments output are enclosed in double quotes.:{{'Claim':["text fragment", "text fragment",...],'Premise':["text fragment", "text fragment",...]}} The input: ```{}``` The output: '''.format(article)
Prompt2	'''These are the claims list ```{}``` you just extracted. Now you need to filter out MajorClaim from the claims list. You can filter by referring to the article ```{}```. In the process of performing the task, there are some points that need to be kept in mind at all times: (1) MajorClaim only appears in induction and conclusion. (2) MajorClaim's function is the author's position or assertion on a controversial topic. The output format needs to be the format of the dictionary in python, the original text fragments output are enclosed in double quotes.:{{' MajorClaim ':["text fragment", "text fragment",...]}} '''.format(claim_list, article)

由于以上两个提示包括两个论证挖掘子任务，分别是论证成分提取和分类。因此，在得到了 GPT4 生成的文本后，使用自然语言处理技术对其进行后处理，通过比对预测文本和目标文本的相似度，阈值设 0.8(即>=0.8 表示提取正确)，来判断是否提取正确，进一步判断论证分类是否正确。

4. 实施和结果

本文用 pytorch1.9.0 实现了以上模型，并用 NIVIDIA RTX3090 GPU 在计算节点上进行了实验。在实验中，采用预训练的 BERT-uncased 作为编码器。我们使用初始学习率为 3e-5 的

Adam 优化器，并选择 16 的批量大小来避免显存不足的问题。我们还采用了超参数优化，其丢弃概率为 0.5。在每种情况下，我们训练 15 个时期，并在测试集上选择性能最好的模型参数。

4.1. 评价指标

为与以往的论辩挖掘方法进行对比，本文选用文献(Wang 等, 2020)中所使用的评价指标。我们计算了三种方法在测试集上的平均精确度 P、平均召回率 R 以及平均 F1 分数。此外，我们使用F_{micro}、F_{macro}作为总体评估指标。

$$F_{marco}=\frac{1}{n}\sum_{i=1}^{n}\frac{F_{MC_i}+F_{C_i}+F_{P_i}}{3}$$

其中 n 是测试集上的文章数量。

4.2. 实验结果

表 2 实验结果

Model	P_{MC}	R_{MC}	F_{MC}	P_C	R_C	F_C	P_P	R_P	F_P	F_{micro}	F_{macro}
BERT	0.08	0.12	0.10	0.07	0.17	0.10	0.36	0.62	0.46	0.32	0.22
BERT_CRF	0.58	0.43	0.49	0.41	0.30	0.35	0.68	0.59	0.63	0.54	0.49
GPT4	0.77	0.78	0.78	0.46	0.67	0.54	0.88	0.71	0.79	0.70	0.73

表 2 总结了预训练语言模型与生成式大模型识别不同论证类型的实验结果，显然，生成式大模型在识别论证类型任务上的整体性能优于预训练语言模型，F_{micro}、F_{macro}得分最高。相比 BERT、CRF，GPT4 在三种论证成分的识别和分类任务中表现出最佳性能，F1、精度率和召回率都达到了最高分数。

相比 BERT，由于 CRF 模型中概率转移矩阵对 BERT 输出序列的约束，使得 BERT_CRF 在序列标注任务中表现更好，因为其最终输出序列通过逐字逐句地推理得出。另外，BERT_CRF 在提取论据子任务的结果中，F1 分数提高，性能明显提高，这源于精度的显著提高和召回率的相对轻微下降。然而，GPT4 强大的文本理解能力和“思维链”的设计，使得 GPT4 在论证成分的识别和分类上比 BERT、BERT_CRF 好。此外，GPT4 在抽取论据子模块任务上性能最好，F1 得分和精确率最高。

值得一提的是，与提取主要主张和论据子任务相比，GPT4 在提取论点时呈现明显的高召回率低精确率的现象。由图 2 可知，这是因为 GPT4 在抽取论点的过程中，混淆了论点论据，错误地将 28%的论据识别成了论点。

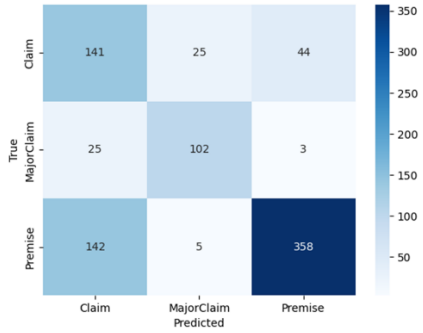


图 2 GPT4 结果混淆矩阵

5. 讨论

本研究表明，与大规模预训练模型相比，在论证成分抽取和分类的子任务上，生成式人工智能—GPT4 的效果明显更好，相比 BERT_CRF 下 F_micro 提升了 16%。与(Sazid & Mercer, 2022)和(Wang 等, 2020)提出的方法相比，分别提升了 1.98%和 3.07%。该研究结果也回答了前文提到的研究问题，GPT4 的强大功能附以 chain of thoughts 的提示词撰写方法使得 GPT4 在提取论点论据的任务上表现较好，相比训练模型，GPT4 更适合应用于真实的论证教学中。GPT4 在论证成分识别和分类上较高的性能一定程度上可以帮助教师改进修正过程，减轻老师修改论文的负担；对于学生来说，当将该技术嵌入到议论文写作反馈辅导系统中后，除了可以向学生清晰地呈现文本的论证结构外，使用 GPT4 技术将提取到的特定主题的论证成分作为预先参考资料储备在论证写作辅导系统中，这样学生在写作过程中能够自由选择是否需要参考资料的帮助，这个做法即满足了写作基础好的学生对写更具说服力的论证的追求，又照顾到了写作基础差的学生参考论证的需求。

本文的工作对当前论辩挖掘在教育领域中的研究做出了一些贡献，这项研究探索了 GPT4 在抽取论证成分上的能力，它为致力于使用将生成式大模型技术辅助学生议论文写作的研究人员提供了一个基础。采用“chain of thoughts”的方式为 GPT4 在具体任务上设计了提示词，为相关领域研究人员提供了参考。

6. 结论

本研究构建和评估了大规模预训练模型和生成式人工智能技术在提取论证成分上的性能，并采用“chain of thoughts”的方式为 GPT4 在具体任务上设计了提示词。本研究发现，与大规模预训练模型相比，生成式人工智能在论证提取子任务上的效果明显更好。同时，本研究也有一些局限性，对于论辩挖掘的四个子任务(议论文和非议论文单元分类；论证抽取；主要主张、主张和前提成分分类；关系识别),只关注了前 3 个子任务，缺乏对论证成分之间关系的研究，关系是构成论证结构的基础。这一局限性将在下一步的研究中补足。其次，对于 GPT4 在提取论点上较低精确度的情况，也将会在下一步的研究中更新。下一步的工作将致力于将 GPT 技术融入真实的论证教学情景中去。

参考文献

Felton, M., & Kuhn, D. (2001). The Development of Argumentive Discourse Skill. *Discourse Processes*, *32*(2–3), 135–153. https://doi.org/10.1080/0163853X.2001.9651595

Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development*, *58*(4), 439–457. https://doi.org/10.1007/s11423-009-9143-8

Lawrence, J., & Reed, C. (2020). Argument Mining: A Survey. *Computational Linguistics*, *45*(4), 765–818. https://doi.org/10.1162/coli_a_00364

Miwa, M., & Bansal, M. (2016, 六月 7). *End-to-End Relation Extraction using LSTMs on Sequences and Tree Structures*. arXiv. 取读于 从 http://arxiv.org/abs/1601.00770

Palau, R. M., & Moens, M.-F. (2009). Argumentation mining: The detection, classification and structure of arguments in text. *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 98–107. Barcelona Spain: ACM. https://doi.org/10.1145/1568234.1568246

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of Oral Discussion on Written Argument. *Discourse Processes*, *32*(2–3), 155–175. https://doi.org/10.1080/0163853X.2001.9651596

Sazid, M. T., & Mercer, R. E. (2022). A Unified Representation and a Decoupled Deep Learning Architecture for Argumentation Mining of Students' Persuasive Essays.

Proceedings of the 9th Workshop on Argument Mining, 74–83. 取读于 从 https://aclanthology.org/2022.argmining-1.6/

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102. https://doi.org/10.1007/s11412-009-9080-x

Stab, C., & Gurevych, I. (2014). Annotating argument components and relations in persuasive essays. *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers*, 1501–1510. 取读于 从 https://aclanthology.org/C14-1142.pdf

Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, *43*(3), 619–659.

Stede, M., Schneider, J., & Hirst, G. (2019). *Argumentation mining*. Springer.

Wambsganss, T., Kueng, T., Soellner, M., & Leimeister, J. M. (2021). ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3411764.3445781

Wambsganss, T., Niklaus, C., Cetto, M., Söllner, M., Handschuh, S., & Leimeister, J. M. (2020). AL: An Adaptive Learning Support System for Argumentation Skills.

Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–14. Honolulu HI USA: ACM. https://doi.org/10.1145/3313831.3376732

Wang, H., Huang, Z., Dou, Y., & Hong, Y. (2020). Argumentation mining on essays at multi scales. *Proceedings of the 28th International Conference on Computational Linguistics*, 5480–5493. 取读于 从 https://aclanthology.org/2020.coling-main.478/

致谢

本研究受国家自然科学基金面上项目支持，项目批准号：62377017

英语写作同伴反　对学生反　素养的影响研究

The Effect of Peer Feedback on Student Feedback Literacy in EFL Writing

王殊　[1]*，任佩玲[2]，熊梦羽[3]，　紫　[4]
[1234]北京　电大学人文学
* Gloria_Wang@bupt.edu.cn

【摘要】学生反　素养是学生利用反　提　写作能力的　要影响因素，提　学生反　素养是提　写作教学效果的　要手段。本文　用　卷调查和文本分析的方法，研究了英语写作同伴反　对提　学生反　素养的影响。过对数据　行对比分析，本研究发现：1）学生反　素养整体处于中等偏上水平；2）经过一个学期的同伴互评教学，学生在　取行动方　的素养有显著提升；3）在学期末的写作任务中，学生能为同伴作文提供更多具体的批评建议。以上研究结果说明英语写作同伴反　对提　学生反　素养有积极的影响。
【关　词】学生反　素养；同伴互评；提供反

Abstract: *Student feedback literacy is crucial for enhancing writing abilities, and improving it is essential for effective writing instruction. This study employs a questionnaire survey and textual analysis to explore peer feedback's effect on student feedback literacy. The findings are as follows: 1) The overall level of student feedback literacy is moderately high; 2) After a semester of peer feedback, students show a significant improvement in action-oriented literacy; 3) In the end-of-semester task, students offer more specific critical suggestions for peers' compositions. Results show peer feedback has a positive effect on student feedback literacy in English writing.*
Keywords: student feedback literacy，peer review，providing feedback

1. 引言

学生反　素养是影响学生学习行为和学业成就的　要因素（Yu & Liu, 2021）。Han 和 Xu（2020）认为，学生可以　过参与反　活动提　反　素养，从而充分利用反　提　写作能力。现有学生反　素养研究主要　中于测评工具的研发（e.g., 董哲和　瑛，2023；Yu et al., 2022; Zhan, 2022）或从收到反　的角度研究学生反　素养（e.g., Zhang & Mao，2023；Zhang et al.，2023），从提供反　角度研究学生反　素养的研究较少。　于此，本文以北京某理工类大学英语专业本科二年级学生为研究对象，　过学生反　素养　卷调查和分析学生提供的反　，探讨英语写作同伴互评对学生反　素养及对学生提供反　能力的影响。

2.文献综

2.1. 学生反　素养研究

在英语写作研究　域，学生反　素养研究仍处于早期　段。Yu 和 Liu（2021）认为英语写作中的学生反　素养是学生感知、评估和使用反　信息的能力。在此基础上，Yu 等（2022）将学生英语写作反　素养定义为学生欣赏、评估和利用反　的能力以及管理情绪的能力。学生反　素养的概念化推动了　一　域的实证研究。Zhang 和 Mao (2023) 研究了学生在英语写作课堂上反　素养的发展轨　。研究发现学生愿意利用反　来改　自己的写作，但不同维度　步不同：从不同来源寻求反　、做出判断和欣赏反　的能力显著提　；而　取行动和管理情绪的能力则提　较少。Zhang 等（2023）则探讨了三位学生在理解、评价、反思和使用

同伴反 及情绪反应方 的反 素养发展情况。

2.2. 学生提供反 的研究

大多数学者致力于研究学生对反 的 纳情况,但提供反 的能力也是英语写作反 素养的 要组成 分。Han 和 Xu（2020）分析了 3 位学生的反 素养发展轨 ，发现学生有的英语能力和对同伴反 的质疑会 碍其反 素养的发展，而积极主动的英语学习者则能及时、有效地调整自己提供反 的策略,反 素养的提升较为明显。该发现与 Patchan 和 Schunn（2015）的研究结果相一致，提供反 的质 受评 者英语能力的影响。

过综 现有研究，我们发现虽然有文献对学生自我报告的反 素养水平 行了探讨，但对反 素养的具体表现缺乏深入了解，此外，在提供反 方 的学生反 素养研究仍然较少。了解 些 有助于我们深入了解学生反 素养，为英语写作教学提供参考建议。本研究将回答以下两个 ：1) 英语写作同伴互评对学生反 素养的影响如何？2) 英语写作同伴互评对学生提供反 能力的影响如何?

3.研究方法

3.1. 研究对象和教学情境

研究对象为北京某理工类 校英语专业二年级学生（女生 20 人，男生 4 人），学生 要在 16 周内完成 6 次不同主 的写作任务，每次任务 定在两周内。

3.2. 研究工具

3.3.1. Peerceptiv 同伴互评系统

在本研究中，学生使用 Peerceptiv 同伴互评系统 行互评，该系统具有匿名同伴反 与回评反 质 等特征（Schunn, 2016），能提供同伴反 的质 。学期初，教师演示了如何使用 Peerceptiv 同伴互评系统 行互评，并讲解了 Paulus（1999）的写作评价标准。学生在系统上提交初稿后， 要批 三名同学的作文，并在六个维度为同伴作文提供反 ：论证、衔接贯、结构、句子、词汇和写作规范。作者收到三名同学的反 后， 要对其反 的有效性行评分，并根据收到的反 修改初稿，修改完成后在系统上提交修改稿。

3.3.2. 学生反 素养 表

我们在学期初和学期末发放了两次学生反 素养 表,第二份 表在第一份 表的基础上调整了 序。研究 用 Zhan（2022）的学生反 素养 卷并 行了细微调整，让 表更切合英语写作情境。该 表共包含 6 个维度：寻求反 、加工反 、 取行动、认可反 、接纳反 和做出改变，每个维度下设置 4 单 ，其中，前三个维度与学生的能力有关，后三个维度与学生的情感有关。寻求反 是指学生从不同来源接收反 的能力；加工反 是指学生理解和评估收到的反 的能力； 取行动是指学生收到反 后 取的行动（如计划、监控）；认可反 用以衡 学生是否认识到反 在学习过程中的 要性；接纳反 关注学生在 对批评性反 时是否能够积极调节情绪；做出改变则用来评价学生是否愿意投入时 和精力参与反 以提 写作能力。研究 用了 6 点 表（1 表示完全不同意, 6 表示完全同意），并设置了 5 个反 ， 样的设计有助于获得更真实的学生反 ， 免回答惯性。学生在卷星上匿名答卷。我们对前后两次 表 行了独立样本 t 检 ，以了解学生反 素养水平变化。

3.3.3. 提供反

本研究对比学生在第一个和最后一个写作任务中提供的反 。 先，切分学生提供的所有反 ，确保每条反 只指出作文中的一个 。第二，标注反 类型（批评、表扬或总结性反 ）和批判性反 类型（具体批评或宽泛批评），具体批评是指 对作文中的具体 提出解 性信息和修改建议（如：英语讲究不 复，原文中 stimulate 使用多处，可以考虑替换

成 boost，advance 等多样表达），宽泛批评是指只指出有　　，但是未　对具体　　（如：多为基础词汇，可以将其中一些替换为　级表达）。为保证标注的质　，每条反　至少由两人背对背标注，如标注有异，两人讨论后达成一致。

4.研究结果与讨论

4.1. 学生反　素养水平

学期初和学期末，分别有 24 和 22 位学生填写了　卷。　卷分析显示，学期初，学生在 6 个维度上的反　素养水平均值　在 4.1 以上，学期末则升至 4.6 以上，学生反　素养处于中等偏上水平。　过对比学期初和学期末的　卷（见表 1），我们发现学生只在　取行动　一维度有显著性提　（$t(44) = -2.67, p < .05$）。　过对两次　卷的所有　　行了对比发现，有 5 个　　有显著差异（见表 1）：加工反　维度下的加工改　，　取行动维度下的制定计划、管理时　和监控　程，认可反　维度下的学习他人。

学生从他人反　中获取有益信息的能力有提　，并且能学习他人优点。另一方　，　过评　他人作文，不仅能提　读者意识，了解读者的期待，还能看到他人的优点，两种经　的交织能帮助学生提　判断反　质　和吸取他人优点的能力（Wu & Schunn, 2020）。学生　取行动　一维度上的素养显著提　，可能与本研究　用的互评模式有关。学生　在规定时　内完成一系列互评任务，因此对于学生的执行力、时　管理能力和自我调控能力有较　要求。

表 1 学生在学期初和学期末的反　素养水平

维度	学期初（$N = 24$）	学期末（$N = 22$）	t 值
	均值 (标准差)	均值 (标准差)	
维度 1: 寻求反	4.33 (.75)	4.67 (.78)	-1.49
维度 2: 加工反	4.58 (.56)	4.83 (.49)	-1.57
加工反　-加工改	4.71 (.55)	5.14 (.56)	-2.61*
维度 3:　取行动	4.13 (.75)	4.68 (.66)	-2.67*
取行动-制定计划	4.13 (.90)	4.68 (.89)	-2.10*
取行动-管理时	4.00 (1.02)	4.64 (.79)	-2.34*
取行动-监控　程	3.88 (.85)	4.55 (.85)	-2.66*
维度 4: 认可反	4.89 (.68)	4.91 (.59)	-.13
认可反　-学习他人	4.67 (.82)	5.18 (.58)	-2.43*
维度 5: 接纳反	5.02 (.62)	5.07 (.58)	-.27
维度 6: 做出改变	4.76 (.57)	4.69 (.58)	.40

注：*$p < .05$

4.2. 提供反

提供反　的能力是反　素养的　要组成　分（董哲、　瑛，2023），现有研究多从收到反　的角度分析学生反　素养，学生提供反　能力的研究较少。为了解学生提供反　的能力，我们分析了学生在学期初和学期末的两个写作任务中提供的反　。我们发现学生提供的批评性反　和表扬性反　占比在两次任务中没有显著差异，但是学生提供的具体批评占比从学期初的 74%升至 87%（$t(23) = -4.07, p < .001$）（见表 3），　一显著　步说明经过一个学期多次同伴互评活动，学生能提供更多具体、有帮助的反　。每次互评时，学生均　根据评价标准为同伴作文提供反　，作者收到反　后，还会回评收到反　的质　，具体的批评性反　能

获得作者的认可，也更可能被 纳（Wu & Schunn, 2020），因此学生会提供具体批评。例如，该建议指出了作文中的 ，并提出明确的修改建议（“原句感觉有些地方不必要，作者后 的论证核心就是利于填补人才缺口，建议 一段中心句就改成：利于培养专业人才，从而填补社会人才缺口，作者 一段既在讲对于个人，又在讲对于社会就业，不如直接落在填补社会人才缺口 一个点”）。

表 3 学生学期初和学期末提供反 的对比

维度	学期初（$N = 24$）	学期末（$N = 24$）	t 值
	均值 (标准差)	均值 (标准差)	
批评性反 占比	.49 (.16)	.49 (.17)	.11
表扬性反 占比	.41 (.14)	.39 (.17)	.75
具体批评占比	.74 (.14)	.87 (.09)	-4.07***

注： ***p < .001

5.结论

本文 过 卷调查和文本分析的方法，研究了英语写作同伴反 对学生反 素养的影响，发现经过一个学期同伴互评教学，学生在 取行动方 的素养有显著提升，并且提供反 的能力有 步。但是要注意的是，本研究 用的同伴互评有以下特点：1）小组匿名互评，有利于让评 人有更多机会评 他人作文，学习他人优点，同时培养学生发现 和解决 的能力；2）学期初对学生 行互评培训，讲解了评价标准，有助于学生提供具体的反 。

本研究存在一定的局 性。第一，未设置参照组。由于教学条件等的 制，本研究未加入参照组，因此对于同伴互评、学生反 素养和提供反 能力的分析只是相关分析，而 因果结论。将来可 用实 设计，设置参照组，探讨同伴反 对反 素养的影响。第二，未分析质性数据。将来可收 访谈、反思报告等质性数据，从而对定 分析结果 行深入解 。

致谢

本研究受北京 电大学大学生研究创新基 资助：2023 年北京 电大学大学生创新创业训练 目“教师反 和同伴反 对学生二语写作能力的影响研究”（ 目号：202309013）及“ 校英语学习者反 素养测评框架研究”（ 目号：202309008）。

参考文献

董哲和 瑛 (2023)。英语写作同伴反 素养 表的编制与应用。解放军外国语学 , 3, 18-25。

Han, Y., & Xu, Y. (2020). The development of student feedback literacy: the influences of teacher feedback on peer feedback. *Assessment & Evaluation in Higher Education, 45*(5), 680-696.

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: how students respond to peers' texts of varying quality. *Instructional Science, 43,* 591-614.

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of second language writing, 8*(3), 265-289.

Schunn, C. (2016). Writing to learn and learning to write through SWoRD. In *Adaptive educational technologies for literacy instruction* (pp. 243-260). Routledge.

Wu, Y., & Schunn, C. D. (2020). From feedback to revisions: Effects of feedback features and perceptions. *Contemporary Educational Psychology, 60,* 101826.

Yu, S., & Liu, C. (2021). Improving student feedback literacy in academic writing: An evidence-based framework. *Assessing Writing, 48,* 100525.

Yu, S., Di Zhang, E., & Liu, C. (2022). Assessing L2 student writing feedback literacy: A scale development and validation study. *Assessing Writing, 53*, 100643.

Zhan, Y. (2022). Developing and validating a student feedback literacy scale. *Assessment & Evaluation in Higher Education, 47*(7), 1087-1100.

Zhang, T., & Mao, Z. (2023). Exploring the development of student feedback literacy in the second language writing classroom. *Assessing Writing, 55,* 100697.

Zhang, F., Min, H. T., He, P., Chen, S., & Ren, S. (2023). Understanding EFL students' feedback literacy development in academic writing: A longitudinal case study. *Assessing Writing, 58,* 100770.

大语言模型在语言领域智能评估的探索——以外语对话场景为例

Exploration of Large Language Models in Intelligent Language Learning Assessment: A Case Study in Foreign Language Dialogue

陈昂轩[1]，贾积有[1*]
[1]北京大学教育学院教育技术系

angxuan.chen@stu.pku.edu.cn; * jjy@pku.edu.cn

【摘要】 大语言模型作为当前最具突破性的前沿人工智能技术，展现出了自然语言提示的控制性、语言判断的智能性和生成反馈的可能性。这使其在语言领域的评估之中具有重要的发展潜力。本研究针对外语学习的语言对话场景为例，针对外语对话评估的具体需求和大语言模型的能力特点，构建了基于大语言模型的 REACH 外语评估设计框架，并针对当前众多开放使用的大语言模型进行了对比实验。结果表明，大语言模型具有作为语言对话评估工具的可行性，其技术特点能够满足语言对话这一情境的现实评估需求；不同的大语言模型在 REACH 框架下的表现不同，国产模型在反馈设计和支持设计上具有优势，而国外模型在评估设计中的精准性更高。未来的研究有望基于 REACH 框架来设计符合语言评估情境的大语言模型应用。

【关键词】 大语言模型；语言对话；智能评估；场景对话

Abstract: *Large language models, as the most breakthrough frontier artificial intelligence technologies currently, have demonstrated control in natural language prompting, intelligence in mistake judgement, and the potential for generating feedback. This makes them with significant developmental potential in the field of language assessment. This study uses foreign language learning dialogues as an example, addressing specific needs for foreign language dialogue assessment and the capabilities of large language models. We built a REACH foreign language assessment framework based on large language models and compared many currently available large language models. The results show that large language models are feasible as language dialogue assessment tools, meeting the real assessment needs of this scenario. Different large language models perform differently under the REACH framework. Domestic models have advantages in feedback and support design, while foreign models are more precise in assessment design. Future research is expected to design large language model applications for language assessment scenarios based on the REACH framework.*

Keywords: large language model, language dialogue, intelligent assessment; situational dialogue;

1. 引言

语言对话是外语教学领域至关重要的部分，在对话中，语言知识能够被拓展迁移到丰富的情境，拓宽学习者对于语言本身的使用理解，发展学习者的语言应用素养。从理论上看，社会建构主义者认为认知发展很大程度上与语言发展的程度相关。维果斯基提出儿童在认知发展过程中，需要通过语言与成人、其他儿童和外界文化交流，儿童是在社会交互的过程中建构意义，实现其学习和发展(Vygotsky, 1967)。在维果斯基的理论基础上，后续的研究进一步探讨了对话在外语教学中的实际应用及其效果。例如, Swain 和 Lapkin 的“输出假设”(Output Hypothesis）强调学习者在语言产出过程中的认知活动。他们认为，当学习者试图用外语表达思想时，他们不仅在重现已有的语言知识，而且还在通过输出活动来测试和扩展这些知识。这种产出过程促进了学习者的语言意识和问题解决能力的提升(Swain, 2005)。

鉴于语言对话在外语教学中的重要地位，如何对学习者的语言对话结果进行评估和答疑，从而提升学习者的语言对话水平，受到了领域内研究者的重点关注。过去的评估方式注重以教师为主的主观判断，如共同欧洲语言参考框架（CEFR）(Jones & Saville, 2009)、ACTFL 口语能力指南(Isbell & Winke, 2019)等等。这种基于人工的主观评估方式难以覆盖广大的学习者人群，也难以给予学习者即时的评估答疑反馈。随着人工智能技术的逐渐兴起，一些研究者也开始探究在语言学习中融入聊天机器人的形式来支持学习者的对话评估。例如，一些研究者构建的 Computer Simulator in Educational Communication（CSIEC）系统，其专注于提供一个虚拟聊天伙伴，可以随时随地与英语学习者进行英语聊天。它根据用户输入、对话上下文、用户及其自身的个性知识、常识知识和推理知识产生反应和反馈(Jia, 2009)。一系列的后续研究也表明了 CSIEC 系统在辅助语言对话方面的有效性(Ebadi & Amini, 2022;丁竹卉等, 2012)。一些研究者采用定制的聊天机器人来进行对话词汇学习的动态评估，表明了以自然语言交流的形式给予学习者语言评估和反馈能够有效促进其对话词汇习得(Jeon, 2023)

然而，之前的研究为了保证对话评估的质量，采用的是基于规则的人工智能方式构建聊天机器人，需要领域专家能够尽可能多的编写规则，来尽可能覆盖学习者对话过程中产生的错误。这使得聊天机器人的开发和维护变得相当繁重，因为它需要不断更新和扩展规则库以应对新的语言使用情况。此外，基于规则的系统往往难以处理语言的多样性和复杂性，特别是涉及到了不同文化背景或非标准语言变体的时候。这种产生式规则驱动的评估方式往往限制了学习者的语言对话创造性，在处理未预见的语言模式或创新性表达时可能显得僵硬和有限。随着以 ChatGPT 为代表的生成式人工智能的崛起，过往以定制和规则为主要范式的聊天机器人正在逐渐转向更加自然和灵活的交流模式(贾积有等, 2023)。借助深度学习和自然语言处理技术，大模型能够学习和模拟复杂的语言模式，并提供更自然和人性化的交流体验。同时，生成式人工智能由于其极大的神经元参数引发了智能的“涌现”现象，使得生成式人工智能能够具备一定的判断能力。因此，本研究探索如何设计符合语言对话评估情境的大语言模型产出，并通过模拟案例评估大语言模型在对话评估之中的可行性。

2. REACH 框架：大语言赋能语言对话评估的设计模式

生成式模型基于人类反馈强化学习（RLHF）范式进行人类价值观的对齐，使得其本身能够通过人类的自然语言的范式来控制其自身的输出，因此，基于大语言模型的对话评估设计可以通过自然语言的处理方式（提示词）进行展开设计。当前对于提示词的设计研究主要关注如何能够更好的发挥大语言模型的能力，而非关注大语言模型的输出是否能够被学习者认可和理解、满足学习者的需求，这使得已有的提示词研究设计未能充分有效的应用于教育场景之中(Zhang & Tur, 2023)。

针对以上的研究现状，本研究拟基于生成式大语言模型，构建对学习者在计算机支持的学习环境中进行外语对话评估的大语言模型框架（REACH 框架）。综合来说，该框架由五个方面的设计组成：

2.1. 角色设计

角色设计重点关注在语言对话评估中大语言模型所扮演的角色。现有的大语言模型提示词设计都重点强调了扮演角色设计在大语言模型性能上的重要性(赵晓伟, 祝智庭, & 沈书生, 2023)。在语言对话评估任务中，大语言模型扮演的不同角色会改变其评估认识的角度及输出的话语语气，从而影响学习者对于话语评估结果的吸收与采纳。除了传统的以教师角色提供对话评估支持外，一些研究也证明了同伴评估对学习者语言对话能力上的提升(Ashraf & Mahdinezhad, 2015)；另一方面，家长的家庭互动方式也能使得学生的语言对话能力受益(Kwok,

Jane Cunningham, & Oram Cardy, 2020)。因此，对于基于大语言模型构建的对话评估设计而言，可以基于学习者的认可差异来进行差异化的评估角色选择。

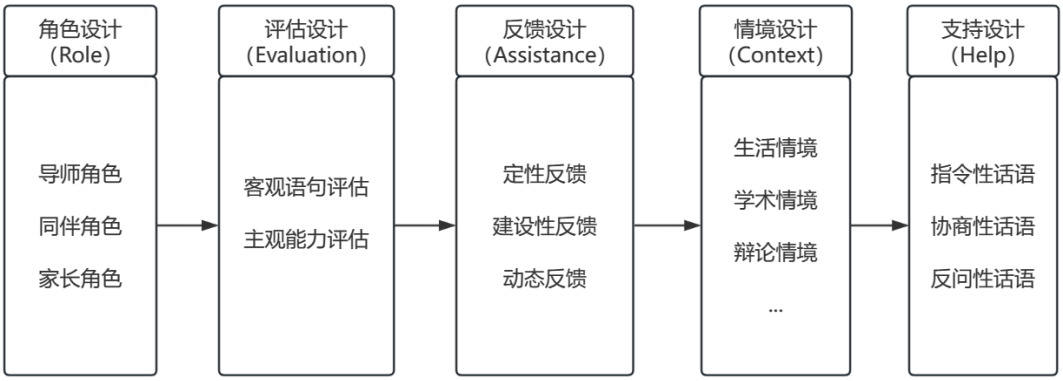


图 1 支持对话评估的大语言模型设计框架 REACH

2.2. 评估设计

评估设计是过往语言评估框架中的关键组成部分，注重如何让大语言模型的回应生成有价值的信息。在语言对话情境中，大语言模型的评估方式可以从客观的语句问题中进行评估。例如，可以通过提示词的具体要求（如“请注重分析用户输入的对话句子的语法错误”）来让大语言模型分析对话的语法结构、语句时态、语义连贯性以及情感表达等客观语句错误。这种客观语句评估可以帮助确定学习对话中的客观错误。此外，大语言模型的评估方式还可以从学习者的话语错误类型中挖掘学习者在语法、时态、句式结构的高维能力差异（如“请评估用户所犯时态错误对应的时态知识点问题”），这种对于学习者语言使用能力的高阶评估可以为其提供进一步改进的方向。

2.3. 反馈设计

反馈设计重点在评估之后如何量化和解释这些信息。在以学习者为中心的评估中，不仅要确定评估的指标和方法，还要确保这些反馈对于学习者是透明和可理解的，才能达到“以评促学”的效果(曾文婕, 周子仪 & 赖静, 2019)。在反馈设计里，需要考虑如何结合量化的评估和定性的建议来提供有效的学习指导。定性的反馈指让大语言模型在反馈之时提供具体的例子（如“请在回应时提供正确的语句”），以促进学习者的深入反思和语言技能的进一步发展；而建设性反馈重点在于提供可操作性的修改意见，使得学习者能够充分操作理解后的反馈信息（如“请给出语法错误应该如何修改”）；而动态反馈重点是提供步进式的反馈建议，使得学习者能够以和大语言模型交互的形式逐步修改提升语言对话的问题（如“请在修改过程中一步一步地向学习者提问”）。步进式的反馈能够满足学习者的语言对话水平过程正在逐步上升的变化现状。反馈设计中不同的设计方式并不是不可兼容的，在基于大语言模型的设计中往往可以进行多重组合，来帮助学习者更好的理解评估，达到“以评促学”的效果。

2.4. 情境设计

情境设计重点注重创建一个能够促进语言学习和应用的环境。语言对话被认为是基于情境的语言使用手段，在情境中检测话语对话水平才具有现实意义(王初明, 2010)。在情境设计中，大语言模型不仅接受语言知识的输入，还构建了一个接近现实的对话场景，使得学习者可以在模拟的实际应用环境中评估语言技能。在基于大语言模型的对话评估中，确定对话的情境能够为精准评估提供重要的辅助，大语言模型能够通过所属情境的不同来评估和反馈学习者的对话用语差异（如“该对话发生在中学生上学的日常聊天情境之中，请基于该情境修改过于正式的用语”），从而基于不同场景变化提供语言风格和词汇的使用建议，这对于指导学习者如何在特定情境下选择合适的语言表达至关重要。另一方面，情境设计还可以采用背景

知识库的方式来进行构建，如提供真实的日常对话样例，来让大语言模型判断（如“判断用户的对话是否符合上述背景对话的情境”），达成情境设计的人机评估对齐。

2.5. 支持设计

支持设计关注的是大语言模型反馈的语句如何能够促进学习者的理解。不同的支持话语能够对应学习者的不同需求，也能够为大语言模型的反馈构建基于学习者特征的个性化方式。指令性语句重点关注的是直接指出对话过程中用户的客观语句问题（如提示词“请直接指出对话中的错误”），这种带有导师特征的社会线索是对话评估中的常见方法，其能够对学习者的问题给出直接清晰的响应，已被证明受到低能力水平学生的认可(Tasdemir & Yalcin Arslan, 2018)，适用于语言对话中错误较多的学生的支持。然而指令性的反馈缺乏对学生自我控制的支持，一些研究指出协商性的话语有助于学生在受到反馈时的自我反思（如提示词“请对用户对话中的语法错误进行反馈，回应的话语以‘你觉得呢？’作为结尾”）。协商性的话语支持设计能够让学习者从意图的角度关注语言的使用，而不是仅仅关注于实现理解句意(Cook, 2015)，适用于语言对话中错误适中的学生的支持。反问性的话语进一步的促进了学习者在接受反馈时的自我控制（如“请评估用户对话中的语法错误，并让用户说明为什么产生了这个错误”）。这种反问性的话语支持已被认为能够充分调动学习者的元认知(Deed, 2008)，适合语言对话中错误较少，先验能力较高的学习者。总的来看，不同的支持性话语体现了基于大语言模型的对话评估的个性化调整。

3. REACH 框架设计实现：基于不同大语言模型的对比试验

为了实验 REACH 框架对语言对话上的评估效果，本研究采用人教版八年级上册第一单元示范对话 Go for it 为例，模拟学生为接受教师协商性话语，受到定性反馈的低先验对话水平的学习者，反馈情景设定为日常口语。具体的提示词为“请扮演一名中学英语教师，用户将作为英语对话的学习者，你需要直接指出用户的英语对话输入文本在语法、语义、时态、上下文的错误，并给出对应的知识点。你在指出错误时需要给出正确的文本样例让用户学习，并向用户反问‘你觉得呢’。该对话的情境是 Rick 和 Helen 在学校内的日常聊天，这是聊天背景[已聊天文本]，现在用户的下一句输入是[用户输入文本]”。结果如表 1 所示。

表 1 REACH 框架下大语言模型对话评估（示例）

原文句子	模拟错误对话句子	REACH 框架设计下的文心一言反馈	REACH 框架设计下的讯飞星火反馈	REACH 框架设计下的 Llama 2 反馈	REACH 框架设计下的 ChatGPT 反馈
Rick: Oh, did you go anywhere interesting?	Rick: Oh, do you went anywhere interesting?	这句话的语法有误，正确的表达应该是 "Oh, did you go anywhere interesting?"，因为这里需要使用过去式的一般疑问句，需要用助动词 "did" 来引导，而不是 "do"。同时，动词 "go" 也应该改为过去式 "went"。 解释：在英语中，一般过去时表示过去某个时间发生的动作或状态，常常与表示过去时间的状语连用。一般过去时的疑问句需要使用助动词 "did"，后面跟着动词原形。而在这句话中，用户使用了 "do" 来引导一般疑问句，这是不正	[Rick: Oh, did you go anywhere interesting?] 错误：用户使用了错误的动词形式。正确的形式应该是“did”而不是“went”。 知识点：过去	Hi！我很乐意帮助你进行英语会话。😊 Oh, do you went anywhere interesting? 🤔? 嗯，我觉得你的句子里可能有个小错误。不要用“went”，应该用“go”，因为你问的是海伦的假期经历。以下是更正后的版	错误及知识点： 1. "do you went" 应为 "did you go"，过去时疑问句中使用 "did" 来辅助动词。 2. "anywhere interesting" 是正确的表

		确的。同时，动词 "go" 的过去式是 "went"，而不是 "wented"。 因此，正确的表达应该是 "Oh, did you go anywhere interesting?"。 例子：比如说，如果我问你上个周末去了哪里，我可以说 "Where did you go last weekend?"，而不是 "Where do you went last weekend?"。你觉得呢?	式的用法。在这个句子中，我们需要使用过去式来询问 Helen 是否去了有趣的地方。	本: Oh, did you go anywhere interesting? 😊 如果你有任何其他问题或你想练习更多的对话，请告诉我!	达方式，但通常我们说 "Did you go anywhere interesting?" 来构成完整的疑问句。你觉得呢?

从评估结果上看，当前大语言模型都能够对学习者的对话文本进行有效的评估，对于模拟的学习者的错误对话语句都能精准识别错误，给出正确的反馈。为了比较不同的大语言模型在 REACH 框架下的语言评估能力，本研究使用同样的提示词针对不同的语言模型进行平行比较，由一位领域专家和一位研究者共同进行编码，研究者和领域专家在最终编码上通过讨论达成了共识（如表 2 所示）。从 REACH 框架的设计上看，不同模型在给出符合要求的评估反馈上具有差异：文心一言大语言模型具有更好的中文解释能力，在测试的评估反馈中基本都能够满足 REACH 框架所指定的回应维度，特别是在评估设计的维度上能够给出更好的主观知识点反馈；讯飞星火大模型在反馈设计维度的能力略落于其他大模型，但能够给出较好的支持设计维度输出；Llama 2 作为大量英文语料训练的国外开源模型，在评估设计和反馈设计维度的输出具有较好的准确性，但是在支持设计上未能很好的满足中文学习者的语气需求差异；ChatGPT 则出现了不能正确回应的情况，如未根据定性反馈的设计要求给出示例，未根据情景设计的指定进行对话评估反馈，未根据评估设计的要求给出对话错误中的知识点解释。这表明，当前不同的大语言模型在利用 REACH 框架指示生成具体评估输出上存在不足，基于学习者的语言对话反馈有进一步提高的空间。相比之下，国产大模型如文心一言、讯飞星火等在当前框架下具有相对而言更加良好的中文理解和输出能力，适合作为赋能人机对话评估的基础模型底座。

表 2 REACH 框架下检测八年级上册示范对话 Go for it 中不同大语言模型的评估反馈

模拟错误对话句子	产出评估反馈的大语言模型	角色设计（Role）	评估设计（Evaluation）	反馈设计（Assistance）	情境设计（Context）	支持设计（Help）
Helen: Hi, Rick. Yus, I am in vacation last month	文心一言	√	√	√	×	√
	讯飞星火	√	√	×	×	√
	Llama 2	√	×	√	√	×
	ChatGPT	√	×	×	√	√
Rick: Oh, do you went anywhere interesting?	文心一言	√	√	√	√	√
	讯飞星火	√	×	×	√	×
	Llama 2	√	×	√	√	×
	ChatGPT	√	×	√	×	√
Helen: Yes, my mom told me you are good at calculating.	文心一言	√	√	√	√	√
	讯飞星火	√	√	×	√	×
	Llama 2	√	√	√	√	×
	ChatGPT	√	×	√	×	√
Rick:Wow! Is you	文心一言	√	√	√	×	√

see Huangguoshu Watrfall?	讯飞星火	√	√	√	×	√
	Llama 2	√	√	×	√	√
	ChatGPT	√	×	×	√	√
Helen: Yes, I did. There was wonful! We hope quite a few photos here. What for you? Did you have anything yes last month?	文心一言	√	×	√	×	√
	讯飞星火	√	×	√	×	√
	Llama 2	√	√	×	√	×
	ChatGPT	√	×	√	√	×
Rick: Not really. I just stayed at home most of the time to read and relax.	文心一言	√	×	√	√	√
	讯飞星火	√	×	√	√	×
	Llama 2	√	×	√	×	√
	ChatGPT	√	√	×	×	√

表 3 检测符合率（符合的次数与检测的总数之比）统计

产出评估反馈的大语言模型	角色设计（Role）	评估设计（Evaluation）	反馈设计（Assistance）	情境设计（Context）	支持设计（Help）	总和统计
文心一言	100%	66.7%	100%	50%	100%	83%
讯飞星火	100%	50%	50%	50%	50%	60%
Llama 2	100%	50%	66.7%	83.3%	33.3%	66.7%
ChatGPT	100%	16.7%	50%	50%	83.3%	60%

4. 讨论和展望

4.1. 大语言模型作为对话评估工具的可行性

大语言模型被认为自然语言处理领域当前最前沿的智能成果，从技术角度上看，大语言模型的特性使其非常适配语言对话评估任务：

自然语言理解：大模型的核心能力之一是对自然语言的深入理解。它能够分析语句结构，识别语义关系，并理解多层次的语言含义。这意味着它可以准确评估学习者的语言使用水平，包括语法正确性、词汇适用性以及语言的流畅性和自然性。从对比实验中可以看出，当前的大语言模型都能够正确分析语句结构，识别语法错误、词汇错误等常见对话问题，自然语言理解能力的提升是促使大语言模型成为语言对话评估工具的关键一环。

动态反馈机制：大语言模型能够根据对话的具体内容动态生成反馈。这种反馈不仅局限于纠正错误，还能够提供改进建议，鼓励探索性学习，并且能够根据学习者的进展调整难度。从对比实验中可以看出，大语言模型都具有进行自然语言状态下反馈的智能性，在提供改进建议上具有比拟人类教师的智能性。

上下文敏感性：在对话中，上下文信息至关重要。大语言模型通过其复杂的算法能够保持对整个对话历史的跟踪，使其在评估和反馈时能够考虑到对话的全局连贯性。然而，当前不同的大模型对于上下文的敏感性不同。通过对比实验可以发现，文心一言大模型能够关注到对话的情境上下文从而给出评估反馈，而 ChatGPT 在这部分还有提升空间。

语言生成能力：除了理解语言，大语言模型还具有强大的语言生成能力。它可以创造出自然和丰富的语言表达，为学习者提供高质量的模型对话回复，从而提高模仿和学习的机会。从对比实验中可以看出，大语言模型能够提供定性反馈的正确示例，有助于学习者正确自我监控对话错误。

4.2.REACH 框架作为大语言模型语言对话评估的可行性

本研究提出的 REACH 框架从角色设计、评估设计、反馈设计、情境设计和支持设计五个方面囊括了基于提示语的大语言模型语言对话评估设计要素。和传统对话评估以及大语言模型提示词设计不同的是，本研究提出的 REACH 框架深入探究了语言对话所需要的评估方式，同时关注学习者如何进行有效的“以评促学”，还融合了大语言模型当前有效的提示词设计。具体而言，角色设计注重大语言模型本身的角色性能，以不同的角色身份来调节评估能力的差异；评估设计注重让大语言模型的回应生成有价值的信息，从客观的语句问题分析到主观的知识水平进行差异性的反馈，充分应用好生成式模型当前涌现的智能性；反馈设计注重确保正确的反馈对于学习者是透明和可理解的，这就需要在设计之中关注反馈的可解释性形式；情境设计重点在于创建一个能够促进对话学习和应用的具体情境，辅助大语言模型的对话评估能够确定合适的语言表达；支持设计则关注的是大语言模型反馈的语句如何能够促进学习者的理解，以不同的话语形式来调整学习者获取知识的效率和动机。

通过对比实验可以发现，当前不同的大语言模型应用 REACH 框架得到的评估输出有所差异，并不是所有的大语言模型都能够完成 REACH 框架对于语言评估输出的要求，其中，表现最好的是文心一言模型，其在实验中展示出了较好的评估稳定性及框架设计一致性。总体而言，以 REACH 框架设定的指示词能够帮助大语言模型生成符合语言对话评估的产出，而这种产出和当前大语言模型的智能性、语料专业性有着正向的关联关系。本研究发现，当前国产大模型在中文语料上具有良好的预训练基础，能够在支持维度和反馈维度生成更符合外语学习者需求的评估反馈；而国外的大模型在语言能力、评估精准度上更具优势，而在中文生成指令反馈上还有提升空间。随着未来大语言模型训练领域在语料和参数上的进一步提升，可以预计未来 REACH 框架下大模型的语言评估能力能够得到更大的提升。

4.3.未来研究展望

本研究重点关注了如何利用大语言模型针对外语教学中的对话情境进行评估这一问题，提出了 REACH 框架来规范大语言模型的对话评估产出，并通过比较实验测试了当前不同语言大模型在生成语言对话评估结果的能力。通过 REACH 框架的应用，未来研究可以更加准确地评估大语言模型的对话能力，发现其在对话评估过程中的优点和不足。同时，该框架也为大语言模型的改进和优化提供了有力的支持。在未来的研究中还可以进一步探索 REACH 框架在其他语种中的应用，以验证其更广泛的有效性。此外，随着大语言模型的不断发展和进步，我们也需要不断地更新和完善 REACH 框架，以适应新的评估需求。例如，研究可以考虑在框架中加入更多的评估指标，或者优化现有的评估方法，以提高评估的准确性和全面性。总之，REACH 框架为我们提供了一个有力的工具，帮助我们更好地理解和评估大语言模型的对话能力。未来的研究可以基于该框架，设计更加以学习者为中心的对话评估方式，深入促进人工智能在外语教学中的突破性变革。

致谢

本文是“东湖高新区国家智能社会治理实验综合基地”项目（课题：“人工智能赋能地方教育高质量发展”）的一项研究成果。

参考文献

Ashraf, H., & Mahdinezhad, M. (2015). The Role of Peer-assessment versus Self-assessment in Promoting Autonomy in Language Use: A Case of EFL Learners. 5(2).

Cook, J. (2015). Negotiation for Meaning and Feedback among Language Learners. Journal of Language Teaching and Research, 6(2), 250. https://doi.org/10.17507/jltr.0602.02

Deed, C. (2008). Strategic Questions: A Means of Building Metacognitive Language. International Journal of Teaching and Learning in Higher Education, 20(3), 481–487.

Ebadi, S., & Amini, A. (2022). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. Interactive Learning Environments, 1–19. https://doi.org/10.1080/10494820.2022.2096638

Isbell, D., & Winke, P. (2019). ACTFL Oral Proficiency Interview – computer (OPIc). Language Testing, 36(3), 467–477. https://doi.org/10.1177/0265532219828253

Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. Computer Assisted Language Learning, 36(7), 1338–1364. https://doi.org/10.1080/09588221.2021.1987272

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. Knowledge-Based Systems, 22(4), 249–255. https://doi.org/10.1016/j.knosys.2008.09.001

Jones, N., & Saville, N. (2009). EUROPEAN LANGUAGE POLICY: ASSESSMENT, LEARNING, AND THE CEFR. Annual Review of Applied Linguistics, 29, 51–63. https://doi.org/10.1017/S0267190509090059

Kwok, E. Y., Jane Cunningham, B., & Oram Cardy, J. (2020). Effectiveness of a parent-implemented language intervention for late-to-talk children: A real-world retrospective clinical chart review. International Journal of Speech-Language Pathology, 22(1), 48–58.

Swain, M. (2005). The output hypothesis: Theory and research. In Handbook of research in second language teaching and learning (pp. 471–483). Routledge.

Tasdemir, M. S., & Yalcin Arslan, F. (2018). Feedback preferences of EFL learners with respect to their learning styles. Cogent Education, 5(1), 1481560.

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. Soviet Psychology, 5(3), 6–18.

Zhang, P., & Tur, G. (2023). A systematic review of ChatGPT use in K-12 education. European Journal of Education, 1–22.

丁竹卉,贾积有,陈宇淏等. (2012). 中学英语教学中应用智能教学系统的效果评估——以"希赛可"为例. 现代教育技术, 22(4), 68–72.

曾文婕, 周子仪, & 赖静. (2019). 建构"以学习为中心的评估"——面向未来的中小学评估转型探析. 教育发展研究, 39(24), 19–27. https://doi.org/10.14121/j.cnki.1008-3855.2019.24.005

王初明. (2010). 互动协同与外语教学. 外语教学与研究, 42(4), 297–299.

贾积有,张誉月. (2023). 人工智能与教育:机遇、挑战与对策. 北京大学教育评论, 21(1), 49-61+188-189.

赵晓伟, 祝智庭, & 沈书生. (2023). 教育提示语工程: 构建数智时代的认识论新话语. 中国远程教育, 43(11), 22–31. https://doi.org/10.13541/j.cnki.chinade.2023.11.003

智能评分与修改：中文作文的提示框架开发与效果评测

Intelligent Scoring and Editing: Development of Prompt Frameworks and Performance Evaluation for Chinese Essay Writing

刘怀亚[1]，贾积有[1*]
[1]北京大学教育学院教育技术系
liuhuaiya@pku.edu.cn; * jjy@pku.edu.cn

【摘要】本研究综合比较了 GPT、讯飞星火和文心一言在写作评分、评分理由、错误修改和范文生成方面的表现。在 Zero-shot 修改方面，讯飞星火表现最接近人类教师，具有最佳的修改效果，而文心一言-3.5 的修改效果较差。GPT-3.5 和讯飞星火都包含低阶和高阶修改层次，但讯飞星火在准确性和细节增加方面更出色。总体而言，讯飞星火在评分、评分理由和修改效果方面与人类教师最为接近，甚至在某些方面更为出色，如细节增加。后续实验中，讯飞星火的表现持续稳定，被学生认为能准确指出错误并满足需求。综上所述，讯飞星火在写作评分和修改领域展现卓越性能，为学生提供可靠的写作支持工具。

【关键词】大语言模型；提示；生成式人工智能；中文写作；智能评分与修改

Abstract: *This study comprehensively compared the performance of GPT, Spark Desk, and ERNIE Botin terms of writing scoring, scoring reasons, error correction, and model-generated samples. In the context of Zero-shot modifications, Spark Desk exhibited the closest resemblance to human teachers with the most effective correction outcomes, whereas ERNIE Bot-3.5 demonstrated comparatively poorer correction results. Both GPT-3.5 and Spark Desk featured low-level and high-level modification layers, but Spark Desk excelled in accuracy and detail enhancement. Overall, Spark Desk demonstrated the closest alignment with human teachers in scoring, scoring rationale, and modification effectiveness, surpassing them in certain aspects such as detail enhancement. Subsequent experiments confirmed the consistent and stable performance of Spark Desk, with students acknowledging its ability to accurately identify errors and meet their writing needs. In conclusion, Spark Desk exhibits outstanding performance in the domains of writing scoring and correction, providing students with a reliable writing support tool.*

Keywords: Large language model; Prompt ;Generative artificial intelligence; Chinese writing; Intelligent scoring and editing

1. 引言

人工智能（Artificial Intelligence, AI）近年来发展迅速，在教育领域产生了广泛的应用（Zawacki et al.,2019）。大语言模型（Large Language Models ,LLM)是一种强大的 AI，通过模拟人类组织语言的方式，能够解释、预测和生成文本。LLM 能够理解对话式的人类输入并以自然的方式作出回应。在教育方面，LLM 可以减轻教师的课程和评分工作负担，同时执行特定任务（Bonner et al.,2023）。LLM 的出现也为写作智能评分与修改提供了可能性。LLM

不仅在理解和生成自然语言方面表现出色，也在写作的纠错和语法规范等方面展现出卓越的潜力。利用大语言模型进行写作批改，可以让学生运用一种智能且高效的方式来修改作文。

2. 利用大语言模型进行写作批改

在传统的课堂环境中，教师很难为每个学生的写作提供详细反馈。利用 LLM 可以为每一位学生提供积极的写作反馈和具体的写作修改建议(Seßler et al.,2023)，并且 LLM 生成的反馈可以提高学生的成绩和动机，增加学生的积极情绪(Meyer et al., 2023)。学生可以借助 LLM 轻松检查自己的写作是否存在语法错误，确保语言清晰简洁(Castellanos,2023)。研究表明，诸如 ChatGPT 之类的大语言模型能够生成类似于人类书写的理性、指导性和吸引力的文本(Klang & Levy,2023)，LLM 可以成为撰写文章时的辅助工具(Katar et al.,2022)，给学生提供新颖的写作体验(Yuan et al.,2022)。

在使用 LLM 时，用户通过输入提示可以与诸如 ChatGPT 的 LLM 进行交互，适当的提示能够生成高质量的生成式内容(Theophilou et al.,2023)。对于大多数用户来说，编写有效的提示目前是一个需要反复试验的过程，常用的是 zero-shot 与 few-shot 的提示(Dang et al., 2022)。通过 zero-shot 和 few-shot 的提示，在不同应用场景下，LLM 都表现出了强大的性能（Wei et al.,2022; Chae & Davidson, 2023; Lu et al., 2023）。

大部分研究都为利用 AI 进行英文写作批改的研究，鲜少有中文写作相关的，更没有中文写作智能评分与修改的提示开发与效果测评的研究。且多数 LLM 的相关研究都基于 ChatGPT, 鲜少有基于国内的平台，如讯飞星火和文心一言的研究，也没有在具体应用场景下的多平台的对比研究。因此本研究旨在对 GPT、讯飞星火和文心一言进行综合比较，以评估它们在写作评分、评分理由、错误修改和范文生成方面的效果，从而确定哪个平台在何种提示框架下表现最为出色。本研究选择埃及学习者的作文为需要修改的文本，开发 zero-shot 和 one-shot 的提示，让 GPT、讯飞星火和文心一言进行写作评分、评分理由、错误修改和范文生成。然后, 选取其中表现最优的平台, 进行一次真实的, 对俄罗斯学习者的 HSK 六级的作文的修改。

3. Zero-shot 提示

Zero-shot 提示的开发如下：

#角色设定：你是一位对外汉语写作老师

##目标任务：根据我给出的内容，给我的写作评分和纠错。

#需求说明：

##分数等级共分为四个档次：

0 分

低分档(10—59)：内容不连贯，有语法错误，有较多错别字。

中分档(60—79)：内容较为连贯，有语法错误，有少量错别字，篇幅不够。

高分档(80—100)：内容结构合理，表达连贯，无语法错误、错别字。

##需要你：

1.根据分数等级要求评分，给出具体的分数。

2.给出评分的理由。

3.罗列出所有错误的地方，至少五条，并给出错误理由。

4.给出修改后的范文。

#风格设定：条理清晰、思维严谨

#我的内容：祝英台是古中国的一个女孩子，想去上学。但是，那时候女孩子不能上学，所以祝英台只好装他是个男人去上学。她认识了梁山伯，最后，他们互相爱上了对方，但他却不知道祝是女孩子。

以后，他去祝家看望祝英台，他见到了一位漂亮的姑娘，她竟然就是祝英台，想去求亲，可她的父母不同意，因为他家穷。他很伤心，生了一场大病，没多久就死去了。

听到这件事，她突然同意嫁给那个有钱人，但是还有一个请求，就是要去看看梁山伯的坟墓。结婚那天，她来到墓前，伤心地哭了，天严重的出奇，“轰”的一声，坟墓裂开了，她跳了进去。“轰”的一声，坟墓又合上了。

这时，风和雨都停了。从此以后，人们常常能看到一对蝴蝶从这儿飞来飞去。人们都说，这对蝴蝶就是梁山伯和祝英台变的。

（注意：用分隔符清晰的表示输入的不同部分。#是一级，##是子集）

上文为埃及学习者薇雅(化名)的写作内容(夏勤,2023)。

3.1. 基于 Zero-shot 的 LLMs 的结果评价

3.1.1. 评分与评分理由

表 1 基于 Zero-shot 的评分及评分理由

对象	评分	评分理由
人类教师	65	内容连贯，读完能知道在讲什么；语法错误较少；只有一个她和他的错别字。总体来说不错，但缺乏细节描写，篇幅较短。
GPT-3.5	70	内容较为连贯，但在表达上还可以更加流畅。 语法错误较少，但依然存在一些问题。 有少量错别字，需要注意拼写和用词准确性。 篇幅较为简短，可以进一步扩展和丰富细节。
讯飞星火	65	1.内容较为连贯，表达了祝英台为了上学而扮男装的故事，以及她与梁山伯之间的爱情和悲剧结局。 2. 有一些语法错误和错别字，但整体上不影响理解。 3. 篇幅较短，可以适当增加细节和描写，使故事更加生动。
文心一言-3.5	50	1.内容不连贯，情节发展缺乏逻辑，例如祝英台和梁山伯的爱情故事突然转到梁山伯的死亡，以及祝英台的悲剧结局。 2.有较多的语法错误，例如“她突然同意嫁给那个有钱人，但是还有一个请求，就是要去看看梁山伯的坟墓。”中的“但是”应该改为“还有”。 3.有较多的错别字，例如“她见到了一位漂亮的姑娘，她竟然就是祝英台”中的“她”应改为“他”。 4.篇幅不够，没有达到完整故事的要求。

总体而言，讯飞星火的评分分数与人类教师的最为接近，评分理由最好。讯飞星火能够结合具体的作文内容给出评分理由，而 GPT 没有。文心一言虽然也与作文内容相联系，但存在幻觉。例如：第二条的错误理由存在错误。第三条，原文中 “他见到了一位漂亮的姑

娘，她竟然就是祝英台”，第一个用的是“他”，但文心一言在指出错误时，写成了“她”见到了一位漂亮的姑娘。

中文写作修改行为划分为增加、删除、调序和替换等四种类型;修改层次分为低阶修改和高阶修改，其中低阶修改指有关标点、词句、 语法的表层修改，高阶修改指涉及主题意义、内容选材和结构构思等的深层修改(刘淑君等，2022)。本研究从修改行为和修改层次的维度对修改进行编码。

3.1.2 人类教师的修改

表 2 人类教师的修改

序号	错误	修改	修改层次	修改行为
1	古中国	古代中国/中国古代	低阶修改	替换
2	他	她	低阶修改	替换
3	以后	在那之后	低阶修改	替换
4	天严重的出奇	这时天气出奇, 暴风骤雨	低阶修改	替换

人类教师共有 4 处修改，修改层次都为低阶修改，修改行为都为替换。

3.1.3 GPT-3.5 的修改

表 3 GPT-3.5 的修改

序号	错误	修改	修改层次	修改行为	正确与否
1	装他是个男人	“假装是个男人”，语法不准确	低阶修改	替换	待定
2	求亲	“求亲”应为“求婚”，用词不当。	低阶修改	替换	否
3	他去祝家看望祝英台	“他去祝家看望祝英台的家人”，以增加表达的准确性。	低阶修改	增加	否
4	听到这件事	可以更具体一些，例如，“得知梁山伯病逝的消息”。	高阶修改	增加	是
5	天严重的出奇	“天严重”应为“天气异常”或“天空异常”，表达不准确。	低阶修改	替换	是

GPT 对于人类教师提到的错误点“古中国”在给出的范文中也没有修改，“以后”这个错误点虽然没有提及，但在范文里改成了“后来”。“装他是个男人”虽然改对了，但错误理由是错误的，正确的理由应该是：因为这主语是祝英台，所以应该为“她”。

3.1.4 讯飞星火的修改

表 4 讯飞星火的修改

序号	错误	修改	修改层次	修改行为	正确与否
1	祝英台是古中国的一个女孩子	“古中国”，表达不准确。	低阶修改	替换	是
2	装他是个男人去上学	"装成男人去上学"，缺少了"成"字，表达不准确	低阶修改	替换	待定
3	因为他家穷	"他"指代错误，应该使用代词来指代梁山伯。	低阶修改	增加	否

4	她突然同意嫁给那个有钱人	"那个有钱人"可以更具体地描述一下对方的身份。	高阶修改	增加	是
5	天严重的出奇	“天空变得异常恶劣”	低阶修改	替换	是

讯飞星火对于人类教师提到的三个错误点都有进行修改，为 1、2、5。“以后”这个错误点虽然没有提及，但在范文里改成了“后来”。和 GPT 一样，“装他是个男人”虽然改对了，但错误理由是错误的，正确的理由应该是：因为这主语是祝英台，所以应该为“她”。

总得来说，讯飞星火更改了所有错误点，甚至比人类教师更好，还增加了细节描写。

3.1.5 文心一言-3.5 的修改

表 5 文心一言-3.5 的修改

序号	错误点	修改	修改层次	修改行为	正确与否
1	祝英台是古中国的一个女孩子	“是”应该改为“是一个”，以保持句子完整。	低阶修改	替换	否
2	所以祝英台只好装他是个男人去上学	装成男人去上学"，缺少了"成"字，表达不准确	低阶修改	替换	是
3	但他却不知道祝是女孩子	“他”应改为“她”，因为这里的主体是祝英台。	低阶修改	替换	否
4	他很伤心，生了一场大病，没多久就死去了	中的“他”应改为“她”，因为这里的主体是祝英台。	低阶修改	替换	否
5	但是还有一个请求	“但是”应改为“还有”，以保持句子的逻辑性。	低阶修改	替换	否

文心一言-3.5 的五个修改中，有三个是关于性别指代的修改，但是只有一个是正确的。关于错误点“古中国”，虽然在错误点中没有指出，但是在给出的范文中进行了修改，修改为了中国古代。

3.2 Zero-shot 总结

1. 讯飞星火的评分分数与人类教师的最为接近，评分理由最好。
2. 讯飞星火的修改效果最好，接近人类教师的修改。
3. GPT-3.5 和讯飞星火都有低阶修改和高阶修改层次，修改行为常用替换和增加。但 GPT3.5 的修改准确性和细节增加不如讯飞星火。
4. 文心一言-3.5 的修改效果最差，在五个修改中，有四个是错的。

总得来说，讯飞星火在评分分数、评分理由、修改，这三个方面与人类教师最为接近，甚至有的修改，如细节增加，比人类教师更好。

4. One-shot 提示

提示包括以下元素：上下文(角色设定、目标任务、需求说明、具体内容)、专家回答（评分、理由、涉及的错误点、范文），反馈说明。

提示的详细内容因篇幅限制，本文不详细赘述。

One-shot 的修改并未表现得更为出色，对于对外汉语写作的修改，zero-shot 已经足够，所以这里不对修改的详细内容进行对比。

但是，对于一样的作文与给分标准，one-shot 给的评分分数总体都比 zero-shot 的高，这一高分并不准确，所以不建议用 one-shot，zero-shot 已经足够。

5. 个案研究

Zero-shot 的讯飞星火在各方面表现最好，因此选择基于讯飞星火进行个案研究。

研究对象：一位俄罗斯学生，女，17 岁，学汉语 3 年，正在准备 HSK 六级考试。

实验过程：要求学生写一篇 HSK6 级的真题作文，与 HSK 六级写作要求一样，考生先要阅读一篇 1000 字左右的叙事文章，时间为 10 分钟；然后将这篇文章缩写为一篇 400 字左右的短文，时间为 35 分钟。只需复述文章内容，不需加入自己的观点。在学生完成作文后，学生接受讯飞星火的辅导，并阐述感受。

学生的作文：因篇幅限制，本文省略。

5.1 讯飞星火的结构化指令设计

利用讯飞星火生成一个 HSK 六级写作评分和修改助手。

以下是结构化指令设计的具体内容，与 zero-shot 使用的提示相差不大：

角色设定：你是一位 HSK 六级写作老师

目标任务：根据我给出的内容，给我的写作评分和纠错。

需求说明：

提供材料在关联的数据集。

分数等级共分为四个档次：

0 分

低分档(10—59)：内容与提供材料相关性不大，内容不连贯，有语法错误，有较多错别字。

中分档(60—79)：内容与提供材料基本相符，有语法错误，有少量错别字，篇幅不够。

高分档(80—100)：内容与提供材料相符，结构合理，表达连贯，无语法错误、错别字。

需要你：

1.根据分数等级要求评分，给出具体的分数。

2.给出评分的理由。

3.罗列出所有错误的地方，至少五条。

4.给出修改后的范文。范文不要照抄材料，要重新改写。

风格设定：条理清晰、思维严谨

5.2 讯飞星火的修改

1.评分：65 分

2.评分理由：内容与提供材料基本相符，但篇幅不够，有语法错误和少量错别字。

3.错误的地方：

(1) "晋公相信谣言，杀儿子"应为"晋献公听信谣言，杀掉了之前立的太子"。原句中“晋公”名字错误，且“杀儿子”过于简单粗暴，缺乏详细描述。

(2)“搞得一团乱” 应为"国势混乱不堪"。原句的“搞得一团乱”表达不清，不够专业和准确。

(3)“决定到外国，也就是楚国。”应改为“决定逃往别的诸侯国，经过千辛万苦，重耳来到了楚国"。原句中缺少了对重耳的描述和他的逃亡过程。

(4)“他的儿子非常辛苦，终于来到楚国。”应改为“经过千辛万苦，重耳来到了楚国。”原句中的“他的儿子”指代不明，且“非常辛苦”过于口语化，不符合书面语风格。

(5) "楚国的皇上觉得这个人将来可能会非常成功，他们成为了朋友，也非常尊重他"应为"楚国的国君楚成王认为重耳有出息，一定会有作为，于是他便以国君之礼来招待他，待他如上宾。重耳对楚成王也十分尊敬。两个人就这样成了朋友"。原句中缺少了对楚成王的判断和他对重耳的态度。

4.修改后的范文：因篇幅限制，本文省略。

6. 结论与展望

6.1 结论

在比较 One-shot 和 Zero-shot 评分的结果时，观察到 One-shot 的评分普遍较高，但这一高分并不准确。此外，与 Zero-shot 相比，One-shot 的修改并未表现得更为出色，Zero-shot 已经足够满足需求。在 Zero-shot 的修改中，讯飞星火表现出与人类教师最为接近的评分分数，其评分理由也最为合理。讯飞星火的修改效果最佳，与人类教师的修改相媲美。值得注意的是，GPT-3.5 和讯飞星火都有低阶修改和高阶修改层次，修改行为常用替换和增加，但相较之下，讯飞星火在修改的准确性和细节增加方面优于 GPT-3.5。另一方面，文心一言-3.5 的修改效果最差，其中有四个修改被评定为错误。

在个案研究中，讯飞星火指出的错误的地方很合适，尤其是在用词准确性方面。“搞得一团乱”表达不清，不够专业和准确，改为 “国势混乱不堪"和 “非常辛苦过于口语化，不符合书面语风格。” 这两条建议，学生认为对她帮助很大。讯飞星火有助于她修改和积累书面用语。此外，学生还认为，讯飞星火给了她更多的细节描写的指导，她认为非常重要。

总体而言，讯飞星火在评分分数、评分理由和修改效果这三个方面都与人类教师最为接近，甚至在一些方面更为出色，如细节增加。在后续实验中，讯飞星火的评分和修改表现持续稳定，学生认为它能够准确指出错误并满足她的需求。

6.2 展望

目前的研究仅仅是个案研究，缺乏准实验研究支持。未来的展望包括进行更广泛的实证研究，以验证这些观察结果的普适性，并进一步探索不同中文写作场景和任务下大型语言模型的性能表现。这样的研究将有助于深化对这些模型在教育领域中潜在应用的理解，并为未来的改进提供有力的指导与支持。

基金

本文是 “东湖高新区国家智能社会治理实验综合基地” 项目（课题： “人工智能赋能地方教育高质量发展” ）的一项研究成果。

参考文献

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International Journal of Educational Technology in Higher Education, 16(1), 1-27.

Bonner, E., Lege, R., & Frazier, E. (2023). LARGE LANGUAGE MODEL-BASED ARTIFICIAL INTELLIGENCE IN THE LANGUAGE CLASSROOM: PRACTICAL IDEAS FOR TEACHING. Teaching English with Technology, 23(1).

Seßler, K., Xiang, T., Bogenrieder, L., & Kasneci, E. (2023, August). PEER: Empowering Writing with Large Language Models. In European Conference on Technology Enhanced Learning (pp. 755-761). Cham: Springer Nature Switzerland.

Meyer, J., Jansen, T., Schiller, R., Liebenow, L., Steinbach, M., Horbach, A., & Fleckenstein, J. (2023). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. Computers and Education: Artificial Intelligence, 100199.

Castellanos-Gomez, A. (2023). Good Practices for Scientific Article Writing with ChatGPT and Other Artificial Intelligence Language Models. Nanomanufacturing, 3(2), 135-138.

Klang, E., & Levy-Mendelovich, S. (2023). Evaluation of OpenAI's large language model as a new tool for writing papers in the field of thrombosis and hemostasis. Journal of Thrombosis and Haemostasis, 21(4), 1055-1058.

Katar, O., Özkan, D., Yildirim, Ö., & Acharya, U. R. (2022). Evaluation of GPT-3 AI language model in research paper writing. Turkish Journal of Science and Technology, 18(2), 311-318.

Yuan, A., Coenen, A., Reif, E., & Ippolito, D. (2022, March). Wordcraft: story writing with large language models. In 27th International Conference on Intelligent User Interfaces (pp. 841-852).

Theophilou, E., Koyutürk, C., Yavari, M., Bursic, S., Donabauer, G., Telari, A., ... & Ognibene, D. (2023, November). Learning to Prompt in the Classroom to Understand AI Limits: A pilot study. In International Conference of the Italian Association for Artificial Intelligence (pp. 481-496). Cham: Springer Nature Switzerland.

Dang, H., Mecke, L., Lehmann, F., Goller, S., & Buschek, D. (2022). How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models. arXiv preprint arXiv:2209.01390.

Wei, X., Cui, X., Cheng, N., Wang, X., Zhang, X., Huang, S., ... & Han, W. (2023). Zero-shot information extraction via chatting with chatgpt. arXiv preprint arXiv:2302.10205.

Chae, Y., & Davidson, T. (2023). Large language models for text classification: From zero-shot learning to fine-tuning. Open Science Foundation.

Lu, J., Zhu, D., Han, W., Zhao, R., Mac Namee, B., & Tan, F. (2023, July). What Makes Pre-trained Language Models Better Zero-shot Learners?. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 2288-2303).

夏勤.(2023).HSK 在线写作诊所的初创阶段实践研究(硕士学位论文,华东师范大学).

刘淑君,李艳,何跃伟 ＆ 王婧婧.(2022).智能作文反馈能提升中学写作教学效果吗？. 华东师范大学学报(教育科学版)(09),90-104.

基于生成式 AI 的小学英语跨学科项目化教学的实践研究

Practical Research on Inter-subject Project-based Teaching of Primary School English Based on Generative AI

杨国超[1]，刘小洁[2*]，姚晓芹[3]，林秋纯[4]，尹栩如[5]，郑芳媛[6]
[123]广州市真光中学附属坑口小学
[456]华南师范大学教育信息技术学院
[*] 692408618@qq.com

【摘要】 智能时代呼唤全面深化基础教育教学改革，变革中小学课堂、提升学生核心素养是当前基础教育信息化建设与应用的关键任务。本研究面向双减政策背景下课堂变革的现实需求，直击小学课堂教学主要问题，以跨学科项目化学习理论、情境认知理论为指导，通过文献探讨和理论演绎的研究方法，融合“智能+”思维方式，明晰了小学英语跨学科的设计框架，生成式 AI 的功能作用，构建了基于生成式 AI 的小学英语跨学科项目化教学策略，丰富完善了原有的课堂实施，为智能技术赋能课堂变革提供理论支撑与实践指引，以期积极推动我国中小学课堂变革与创新。

【关键词】 生成式 AI；跨学科项目化；实践

Abstract: *The intelligent era calls for comprehensively deepening the teaching reform of basic education. Reforming primary and secondary school classrooms and improving students' core literacy are the key tasks in the construction and application of information technology in basic education. Aiming at the practical needs of classroom reform under the background of double-reduction policy, this study directly targets the main problems of primary school classroom teaching. Guided by inter-subject project-based learning theory and situational cognition theory, it integrates the thinking mode of "intelligence +" through the research methods of literature discussion and theoretical deduction, and clarifies the inter-subject design framework of primary school English and the function of generative AI. This paper constructs an inter-subject project-based teaching strategy for primary school English based on generative AI, enriches and improves the original classroom implementation, provides theoretical support and practical guidance for intelligent technology-enabled classroom reform, and actively promotes the classroom reform and innovation in primary and secondary schools in China.*

Keywords: *Generative AI; inter-subject project; practice*

1. 问题的提出

2023 年 2 月，教育部副部长王嘉毅指出以数字化赋能基础教育高质量发展，是面向新时代的战略选择和必由之路（教育部，2023）。近年来，生成式 AI 正逐步对精准教学、自适应学习等提供赋能作用，正成为教育教学领域改革创新的重要载体。《义务教育英语课程标准（2022 年版）》强调积极开展主题化、项目式学习等综合教学活动（教育部，2022）。跨学科项目化学习结合两个及以上的学科，将自然和社会中复杂真实的问题转化为项目，培养学生在探究实践和完成项目的过程中，提高解决问题的能力。本研究以小学英语为例，旨在探索基于数智技术的小学英语跨学科项目化学习的策略并在实践中加以应用，以期为利用生成

式 AI 开展跨学科项目化学习提供案例参考，为培养数字时代的创新人才提供路径遵循，以教育数字化转型赋能基础教育高质量发展。

2.相关研究综述

2.1. 小学英语跨学科项目化教学的研究述评

目前关于小学英语跨学科项目化教学的研究主要集中在项目式学习的内涵、作用、策略等方面。范从娟（2021）以项目学习理论为指导,明确了项目主题的原则,并通过小学英语项目化教学实践案例阐释主题确立的策略。车建琴（2021）界定了英语学科项目化学习的基本理念，并提出了项目化学习的具体实施路径。朱敏会（2023）提炼了项目化实施的基本路径，并设计了基于基本路径的项目化学习的阶段方案。综上所述，目前关于小学英语项目化教学的研究较多，但是关于小学英语跨学科项目化教学的研究较少。

2.2. 基于生成式 AI 的小学英语课堂教学的研究述评

目前关基于生成式 AI 的小学英语课堂教学研究主要集中在生成式 AI 的赋能作用、策略方法等方面。刘璇(2023)提出了基于智能元宇宙的语言教与学架构和教学流程。武和平(2023)结合生成式人工智能在语言教学领域的应用场景,探讨如何在与它聊天时给出合适的提示语。王晓华（2018）分析了虚拟现实技术对初级外语教学的作用，方式、问题及对应的解决办法。综上所述，目前关于基于生成式 AI 课堂教学研究较多，但是关于基于生成式 AI 的小学英语跨学科项目化课堂教学研究较少。

3.理论基础

3.1. 跨学科项目化学习理论

跨学科项目化学习指为解决某个真实而复杂的问题,学习者学习并创造性地整合不同学科的知识和方法，以形成新的理解和整合性的学习成果的过程（夏雪梅，2022）。本研究基于组合型的跨学科项目化学习实践模型，根据驱动性问题，将项目化学习活动设置为结合英语与劳动知识内容的跨学科项目化学习活动。

3.2. 情境认知理论

情境认知主张学习是学习者参与学习活动并与他人、环境等相互作用的过程,具有情境性、真实性、实践性、探究性及主动性（刘义和高芳，2010）。本研究基于来自学生实际生活的真实问题，充分利用智能技术创设真实的项目情境，赋予学生在真实情境中寻找解决方案的权利，激发学生的主观能动性，鼓励其自主探究。

4.基于生成式 AI 的小学英语跨学科项目化教学策略的提出

4.1. 小学英语跨学科项目化学习设计框架

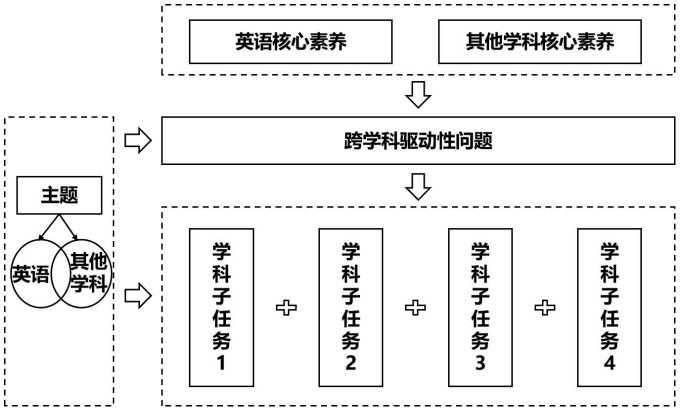


图 1 小学英语跨学科项目化学习设计框架

4.2. 生成式 AI 对小学英语跨学科项目化教学的赋能作用

4.2.1 适配性推荐跨学科项目化主题

生成式 AI 依据学生的历史学习数据、兴趣爱好、学科成绩等信息，挖掘学情特点，分析项目主题的相关性、难度、热度等指标，为学生推荐符合他们兴趣、能力和需求的跨学科项目化主题提供数据支持。此外，生成式 AI 可以作为学生的智能助手，提供项目主题相关的学习资源、实践建议和个性指导，协助学生更好地完成项目任务，并实时收集学生的学习反馈，了解他们在项目探究过程中的困难和需求。

4.2.2 精准性构建跨学科项目化驱动性问题

生成式 AI 可以根据学生特点和需求，设计不同领域、不同难度的跨学科项目化驱动性问题，涵盖自然科学、社会科学、人文艺术等多个领域，旨在培养学生全面的素养，以激发学生的探究欲望。在实时反馈与调整方面，可以实时收集学生的学习反馈，及时调整生成的驱动性问题，使其更贴合学生的实际需求，为学生提供个性化的指导和建议，帮助学生更好地应对项目挑战。

4.2.3 人机协同赋能跨学科项目化任务

在项目任务的策划阶段，教师可以根据教学目标和学生的实际情况，提出项目任务的总体框架；而生成式 AI 则提供相关资源、建议和方案，协助教师优化任务设计，协助师生进行资料整理和分享，凭借其数据跟踪功能，对项目任务的全过程进行监控和记录，为学生和教师提供全面、客观的过程评价数据，人机协推进跨学科项目化学习活动的开展。

4.3. 基于生成式 AI 的小学英语跨学科项目化教学策略

4.3.1 多模态呈现，创设项目情境

跨学科项目化学习旨在通过解决真实情境中的跨学科问题加深学生的跨学科知识理解，提升学生的跨学科核心素养。教师可运用图片、音频、视频、AR、VR 等多模态资源呈现贴近小学生现实生活的项目情境，也可通过智能技术生成虚拟场景，为学生提供沉浸式、真实性的学习情境，增强学生的体验感，从而激发学生的学习兴趣。

4.3.2 目标排序，分解项目任务

跨学科项目化学习目标是设置项目任务的风向标，也是检验学生完成项目任务的方向。在开展项目化学习前期，教师应考虑所教授学生的认知发展水平，鼓励并带领学生依据跨学科驱动性问题将复杂的项目任务拆解为各个子任务，并按照逻辑框架排列任务目标，让学生整体感知项目框架，明晰每个子任务的目标及任务内容。

4.3.3 协同合作，生成项目方案

以英语为主学科的跨学科项目化学习应遵循英语学习活动观，依照“学习理解-应用实践-迁移创新”的进阶方式，坚持学创结合，引导学生在小组协作探究中，联系个人生活经验与前置学科知识开展项目学习。教师应提供学习支架促进学生协同合作完成项目任务，利用智能技术综合运用英语及其他学科的知识与方法，创新性地生成解决驱动性问题的方案。

4.3.4 数据驱动，反馈项目评价

新课程标准要求坚持以评促学、以评促教，将评价贯穿英语课程教与学的全过程，引导学生成为各类评价活动的设计者，参与者和合作者。教师在开展跨学科项目化学习时，可依托智能技术平台实时采集学生的学习行为数据、学习结果数据，根据平台自动生成的数据报告及时监测学生进行项目学习的状态，及时反馈学习情况并调整教学策略，实现精准教学。

5.基于生成式 AI 的小学英语跨学科项目化教学策略的应用

本项目面向“双新”“双减”政策背景，旨在培养学生的英语学科核心素养以及劳动学科的核心素养，设计了 Tidy the room 项目整体规划。

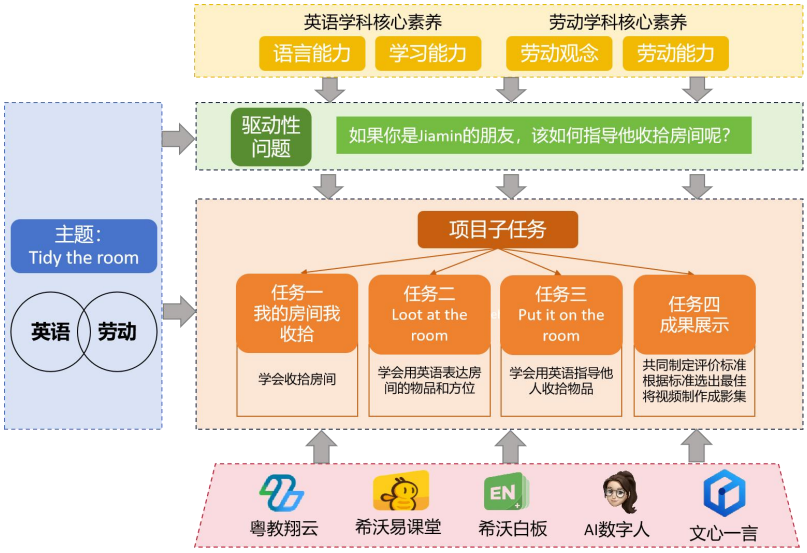


图 2 Tidy the room 项目整体规划

6. 结语

本研究面向双减政策背景下课堂变革的现实需求，直击小学课堂教学主要问题，以跨学科项目化学习理论、情境认知理论为指导，通过文献探讨和理论演绎的研究方法，融合“智能+”思维方式，明晰了小学英语跨学科的设计框架，生成式 AI 的功能，构建了基于生成式 AI 的小学英语跨学科项目化教学策略，丰富完善了原有的课堂实施，为智能技术赋能课堂变革提供理论支撑与实践指引，以期积极推动我国中小学课堂变革与创新。

参考文献

教育部(2023)。数字化赋能基础教育高质量发展论坛在京举行。检索自 http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2023/2023_zt01/yw/202302/t20230214_1044669.html

教育部(2022)。《义务教育英语课程标准(2022 年版)》。检索自 https://www.pep.com.cn/ebook/2022yjkcbz/yy/mobile/index.html

高文、徐斌艳和吴刚(2008)。**建构主义教育研究**。北京:教育科学出版社。

范从娟(2021)。小学英语项目化学习主题确立的实践探索。**中小学教学研究**，(06)，16-19。

车建琴(2021)。指向学科核心素养的小学英语项目化学习研究。**上海课程教学研究**，(11)，55-60。

朱敏会(2023)。小学英语项目化学习的特征与路径研究。**小学教学设计**，(21)，15-18。

刘璇（2023)。智能元宇宙背景下大学英语教与学模式探究。**中国多媒体与网络教学学报(上旬刊)(02)**，9-12。

武和平(2023)。语言教师如何与生成式人工智能“聊天”——以英语学科为例。**中小学数字化教学**，(08)，75-79。

王晓华(2018)。虚拟现实技术支持下的初级英语教学探究。**教育理论与实践**，(09)，46-48。

夏雪梅（2022)。跨学科项目化学习:内涵、设计逻辑与实践原型。**课程.教材.教法**，42(10)，78-84。

刘义和高芳（2010)。情境认知学习理论与情境认知教学模式简析。**教育探索**，(6)，88－89。

基于OBE理念的现代教育技术课程混合式教学模式研究

Research on Blended Teaching Mode of Modern Educational Technology Course Based on the Concept of Outcome-Based Education

张文馨[1*]，李西[2]
[1]成都师范学院网络与信息中心
[2]四川外国语大学英语学院
[*] 309134711@qq.com

【摘要】 在现代教育技术课程中贯彻OBE理念，确定课程目标、重构教学内容体系和教学评价体系，构建线上线下混合式教学模式并进行实践研究，通过问卷调查、访谈和学业测试等数据，表明该教学模式有助于课程目标达成和提高学生满意度。

【关键词】 OBE理念；混合式教学模式；现代教育技术课程

Abstract: *By implementing the concept of Outcome-Based Education in the course of Modern Educational Technology, this paper determines the course objectives, reconstructs the teaching content system and teaching evaluation system, constructs an online and offline blended teaching model and conducts a practical study. Through related data from questionnaires, interviews and academic tests, it shows that the teaching model helps to achieve the course objectives and improve students' satisfaction.*

Key words: the Concept of Outcome-Based Education, Blended teaching Modern,educational technology course

1.前言

教育部关于实施卓越教师培养计划2.0的意见中指出，深化信息技术助推教育教学改革，推广混合式教学模式，要利用大数据、云计算等技术，对课程教学实施情况进行监测，有效诊断评价师范生学习状况和教学质量。随着互联网技术的发展，特别是新冠疫情期间开展的大规模的在线教学活动，进一步促进了混合式教学的发展，混合式教学正成为未来教育的新常态。线上线下混合式教学模式为师范类课程改革提供了新思路。成果导向教育（Outcome-Based Education，OBE）强调以学生为中心、以学习成果产出为导向，采用反向设计、正向实施的方式进行课程教学改革，该理念有力支撑了师范类专业认证的实施，如何在师范类课程中贯彻OBE理念，不断提高师范生培养质量，是当下师范专业课程教学面临的关键问题。

2.国内外相关研究

混合式教学最早是应用于企业培训以提高培训绩效，逐步被应用于教学领域以期提高教学质量。相关研究表明，混合式教学模式在激发学生学习兴趣、帮助学生建构知识、提升学习效率、“提高学生学习效果”等方面具有积极影响[1]。研究表明“学生对混合式教学具有较高满意度”[2]，混合式教学“在提升课程教学质量上具有可行性和有效性”[3]，混合式学习

“有助于提高学生的批判性思维技能”[4]，“提高学生的沟通能力和自我效能感”[5]。由于混合式教学兼具线上和线下教学的优势和特点，混合教学对学习动机、情感态度、高阶思维、学习成绩、满意度、实践技能都有不同程度的促进效果[6]。混合式教学体现了以学生为中心的教学设计思想，也强调学生在学习过程中的学习成果产出，较多的研究开始关注 OBE 理念下混合式课程教学改革，比如在食品化学[7]、机械设计[8]、非线性编辑技术[9]等应用型课程中的实践探索，但在师范类课程中的研究还比较缺乏。本文基于 OBE 理念，以《现代教育技术》课程建设为例，研究线上线下混合式教学模式，以期提高学生学习满意度和课程目标达成度，提升师范生培养质量。

3. OBE 理念下的现代教育技术课程混合式教学改革

现代教育技术课程作为师范专业课程体系中的教师教育类课程，面向全校所有师范生开设，旨在培养师范生的信息化教学能力，以适应未来教学的需求。该课程开设 16 周，每周两学时，按照教学大纲的编排完成理论基础、教学系统设计、数字化教学资源检索与处理、课件制作、信息化教学评价等内容。但在常规教学中凸显出教学任务较重和学时局限的矛盾，这一矛盾又衍生为完成教学任务容易步入传统教学模式的误区。课堂时间被教师知识讲授、操作示范、展示汇报等教学活动填满，而学生的独立思考、自主学习、合作探究等创新能力的培养较为匮乏，学生兴趣和参与度较低，课堂效果和学生满意度不理想。另外，学生信息技术素养基础差异较大、学生自我效能感低等因素阻碍了课程目标的达成。显然，传统的教学模式不符合 OBE 的理念，也较难达到较好的学习效果。混合式教学作为倡导“学生中心”范式下的教学改革[10]为“现代教育技术”课程变革指明了新的方向。

3.1. 制订课程目标 本研究认为，OBE 理念下的现代教育技术课程目标应该是：学习者掌握信息时代教育教学的理论、方法，具备信息化教学设计能力及运用信息技术开展教研的能力，具体分目标见表 1。

表 1 现代教育技术课程支撑的人才培养目标

	现代教育技术课程支撑的人才培养目标
目标一	理解教育技术的基本概念、意义及理论基础；理解信息化环境下教与学的内涵和规律。
目标二	掌握信息化教学设计的基本模式和方法，能在给定教学对象、教学环境等条件下，有效选择教学策略、教学媒体等，根据学科教学需要完成一节完整的信息技术与课程融合课的信息化教学设计方案。
目标三	能根据教学需要检索、甄别、加工和利用信息化资源素材；能运用 PowerPoint 平台整合资源素材，设计制作教学课件。
目标四	掌握常用的信息化教学评价方法，有效开展信息化环境下的教学评价。
目标五	理解未来教师专业发展的意义和目标，掌握开展信息化教研活动的途径和方法。

3.2. 优化课程内容 基于上述目标，对教学内容进行重构，对标毕业要求，结合《中小学教师教育技术能力标准（试行）》等文件要求，将毕业目标体现在各教学专题中并细化成具体任务。在混合式学习任务和活动编排中，以学生为中心为理念、以培养学生信息化教学设计能力为核心，以目标达成度为导向，关注学生在混合式学习中的学习体验和高阶思维能力的提升。在线教学和面授教学的比例关乎混合教学的灵活性及其成功实施。本课程共 32 个学时，根据相关研究发现，线上线下各 50%的混合教学对学生学习的效果最佳[6]，因此确定为线上

16 学时、线下 16 学时。每周一次线上学习共计 2 学时、一次线下学习共计 2 学时，8 周完成课程任务。

2001 年加拿大学者 Garrison 等人提出了探究社区模型，该模型已经发展为指导混合式教学的重要理论框架，该理论为混合式学习活动设计明确了设计框架和目标，即学习活动设计要支持社会临场感、教学临场感或认知临场感及其二级维度的创设。所以，在安排学习任务、组织教学活动时，要应用多种教学策略创设社会临场感、教学临场感、认知临场感。在课程伊始，开展破冰游戏、完成小组建设，营造温浓郁的学习氛围，让学生产生团队归属感和授课教师认同感；介绍课程详细安排，激发学习动机，建立学习契约。在课程中，开展任务驱动、案例分析、角色扮演、合作探究、头脑风暴、小组合作竞争等学习活动并及时反馈评价，维持学生学习动机、提升学生学习参与度、培养学生的自主探究能力。在课程后期，注重小组及个人学习成果展示，对标课程目标以“产出导向”进行学习反思，以期持续改进，提升学生学习满意度。具体学习任务安排如表 2。

表 2：现代教育技术课程混合式学习任务安排

现代教育技术课程混合式学习任务安排

学习专题	学习形式	主要学习任务
现代教育技术概述	线下面授学习	破冰游戏、小组建设。 学习混合式教学技术清单，熟悉混合式教学环境。 了解课程框架和学习安排，形成学习期望，激发学习动机。 了解学习评价方式，建立学习契约。 理解关键概念，头脑风暴如何有效利用教育技术促进学习。
	线上学习	自主学习相关内容。阐述行为主义学习理论、认知主义理论、建构主义学习理论的主要观点，回复在平台讨论区。
教育技术理论	线下面授学习	任务驱动：小组确定信息化教学设计选题。 理解媒体理论和传播理论。头脑风暴：传播理论在课堂教学中的作用以及如何更好地开展教学传播活动、教学媒体媒体对教学的变革作用。
	线上学习	自主学习相关内容，了解自主学习模式、探究学习模式、协作学习模式、项目学习模式。基于在线学习平台分小组角色扮演不同的信息化教学模式，分析和传统教学设计的区别。
典型信息化教学模式	线下面授学习	解决线上学习提出的问题。每小组选择一种信息化教学模式作为小组信息化教学设计的模式。

信息化教学设计	线上学习	自主学习相关资源。分析优秀教学设计案例，分析其设计原理和理论知识。设计信息化教学设计方案。
	线下面授学习	小组展示、同伴互评信息化教学设计方案。
数字化教学资源的获取与加工	线上学习	自主学习相关内容，练习数字化教学资源的检索和加工。为信息化教学设计方案准备素材。提出线上学习遇到的问题。
	线下面授学习	解决线上学习提出的问题。 实践案例分析，优化信息化教学素材。
交互式课件设计与制作	线上学习	自主学习相关内容，掌握交互式课件设计与制作的要点和原则。初步形成小组信息化设计方案配套课件。
	线下面授学习	课件案例分析。自我反思、同伴互评、优化小组课件。
信息化教学评价	线上学习	自主学习相关内容，掌握信息化教学评价的原则和方法。为信息化教学设计方案设计教学评价。
	线下面授学习	小组展示、优化评价方案。
信息时代教师专业发展	线上学习	自主学习相关内容，分析教研案例。
	线下面授学习	小组参与式体验信息化教研活动。
总结反思	线上学习	课程总结、自我反思。

3.3.分析学习者特征 本课程的授课对象是全校的师范生，有理科生、文科生、艺术生，不同学科背景的学生技术素养存在差异，同一学科背景的学生也存在个体差异。在课前通过问卷调查了解学生对该课程的意识态度、动机信念、相关知识储备情况、学习风格等信息，参考问卷数据结果初步确定学习者特征，在线上、线下课程学习中进一步判断学习者特征，为动态修订、优化教学设计方案提供支撑。研究表明“大学生的动机信念在混合式教学和学习成效之间起到显著的正向强化作用”[11]。由于学生的混合式学习经历有差异，有些学生可能会因为技术困扰产生畏难情绪进而影响学习开展和学习成效，在开展混合式教学之前，教师需带领学习者熟悉技术环境，比如对如何使用平台开展学习活动进行培训，并发布平台使用的操作手册供学习者可随时查看；制作混合式学习技术清单，请学生尝试、体验混合式学习所需要用到的技术工具和功能，并分享自己的使用感受，以此提升学生对混合式学习的自我效能感。

3.4.重构评价体系 OBE 理念对师范类课程教学质量评价有了新的要求，强调对师范类专业教学进行全方位、全过程评价，并将评价结果应用于教学改进。

混合式教学评价不仅要关注学生所学知识的过程性和终结性评价，还要充分关注学习成果的评价，包括个人自主学习成果、团队研究性学习成果以及个体非智力因素的评价[12]。在师范类专业认证理念指导下，结合混合式教学特点，重新设计课程评价方式，构建多层次、差异化的混合式教学评价体系，体现评价主体多元化、评价内容多元化、评价方式多元化。平时成绩占比 60%、期末考核成绩各占比 40%，平时成绩包括课堂表现成绩占比 20%、项目成绩占比 20%、小组合作能力成绩占比 10%；期末考核以课程学习论文进行考核。具体评价要素如图 1 所示。

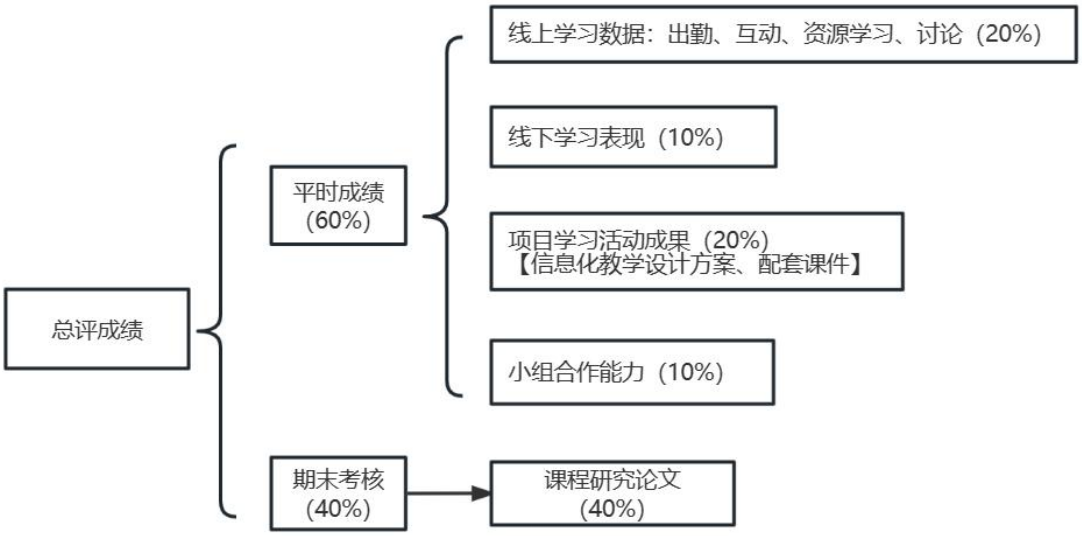


图 1：现代教育技术课程的评价体系

4. 混合式教学模式构建及应用实践

4.1. 混合式教学模式构建 对标课程要求，结合混合式教学设计特点，以学习者为中心，成果为导向，反向设计教学目标、教学内容和教学评价方式。作为课程设计者要扮演好学生学习设计者的角色，关注“以学习者为中心”的理念与需求，构建以学生为中心的教学模式，对课前、课中、课后三个环节的学习活动进行有效设计，为学生提供学习支持服务（包括情感支持服务、行为支持服务，认知支持服务），为学习者创设较好的学习体验，以期提升学生的学习满意度和参与度，支撑师范生毕业目标达成。在前文研究基础上形成了 OBE 视域下的混合式教学模式，如图 2 所示。该模式在 OBE 理念下，确定课程目标、课程内容、进行学习者分析、重构评价体系，在教学中始终贯彻“学生中心、成果产出”的理念开展混合式教学。

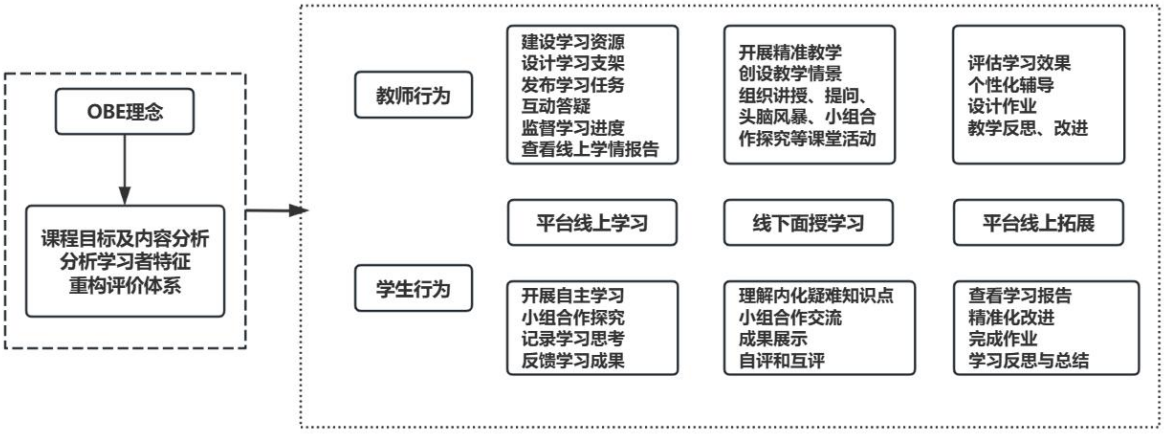


图 2：OBE 视域下的混合式教学模式

4.2. 混合式教学模式实践 为探究 OBE 理念下的混合式教学模式的有效性，本研究以 A 大学 2022 级数学专业学生为对象，基于 OBE 的理念，以“有课”平台为教学载体构建混合式教学模式，在《现代教育技术》课程中开展了一学期的探索。

4.2.1. 课前准备 教师在“有课”平台上发布本周学习任务，为学生设计学习支架，提供相关学习资源（如课件、视频、网址）。学生在平台查看本周学习任务，以小组合作、任务驱动、案例分析、角色扮演等学习活动完成线上学习任务并记录学习思考和反馈学习问题。教师查

看学习进度，监督学生学习进度和任务完成质量，并提供及时辅导，最后总结学生课前学习和任务完成存在的问题。

4.2.2.课堂面授教学 在面授课程中，教师通过自身教学魅力和丰富有趣的教学互动、精心设计的教学活动，拉近和学生之间的情感距离、激发学习动机。在教学伊始呈现清晰的课程内容框架和学习安排，以降低学习者的学习防御心理。在每一次的线下面授课堂中，教师针对线上学习的目标达成度情况、学情报告开展精准教学，分析学生存在的薄弱知识点、教学重难点，对于学生存在的共性问题进行详细解疑。基于“有课”平台的课堂工具，开展签到、讲授、提问、头脑风暴、小组合作探究等活动，对学习小组及个人进行个性化的指导。

4.2.3.课后个性化辅导 在课后，教师根据线上线下学习情况设计作业，提供拓展资料，对学生开展个性化的辅导，对本次课进行教学反思，并引导学生进行学习反思。学生根据自己的线上线下学情报告，掌握自己本次课学习的目标达成情况，并在数据支持下进行精准改进，若有需要寻求教师的个性化指导，在老师提供的学习反思框架下开展学习反思。

4.3 线上线下混合式教学效果分析 为验证本研究构建的线上线下混合式教学模式的教学实践效果，采用问卷调查法、访谈法、总评成绩分析学习数据。

4.3.1.学习满意度问卷调查分析 为验证混合式教学模式的实施效果，课程结束后，对执教的四个班级学生进行课程实施效果问卷调查。共发放150份问卷，有效回收148份问卷。问卷共设置7个问题，从课程适应情况、主观态度、学习效果等维度进行调查。问卷数据分析表明：91.95%的学生表示能够适应本门课的混合式教学模式；相较于传统教学模式，89.6%的学生更喜欢混合式教学；90.54%的学生对混合式教学模式感到满意；80.42%愿意主动使用混合式学习这种学习方式；84.61%的学生认为激发了自己的学习热情；89.93%的学生认为可以增加学习兴趣、提高学习效果。

4.3.2.访谈结果分析 由问卷调查数据分析可知，大部分学生认可混合式教学模式。通过访谈了也解到很多学生认为这种学习方式时间、地点较灵活，可以根据自己的实际情况安排在线学习任务。一些学生认为混合式教学模式很有趣，愿意继续使用。

结合学生在线学习数据和课堂表现选取了两位表现积极的学生进行了访谈，根据调查问卷数据选取了满意度不太高的两位学生进行访谈，访谈围绕如下问题：您认为混合式教学模式效果如何？您在该教学模式的学习过程中遇到的主要困难是什么？您的意见及建议？四位学生均表示混合式教学模式更适合互联网时代的学习，可以锻炼自主学习能力，更利于自己作为职前师范生适应未来的教学工作。在谈到学习困难时，学生1认为在线学习阶段小组合作方面有一定困难，小组在线讨论不如面对面讨论有效率；学生2认为该模式注重合作探究以及学习成果的展示，对学习能力要求较高；学生3认为线上学习阶段自主学习能力不足，容易学习懈怠；学生4认为个人学习动机不强，自我学习约束力不够。在谈到意见及建议时，学生1希望能提升线上小组合作实效；学生2希望能够得到更全面的指导；学生3认为学习任务重；学生4认为应该进一步加强学习督促和管理。

4.3.3总评成绩分析 在课程结束后，按照构建的课程评价体系对执教的四个班级课程总评成绩进行了评价，根据“课程总评成绩和目标评价方案”计算出课程分目标达成度和目标达成度，计算结果显示四个班级的课程目标度均在0.8以上，超过了课程目标期望值0.7。相较于笔者执教的2021级（课程改革前）学生课程目标度，线上线下混合式教学模式下的课程目标度达成度明显提高。

5. 结论

通过对教学效果进行分析总结，本研究开展的混合式教学实践得到了绝大部分学生的认可。研究认为OBE理念下的混合式教学模式可以提高学生学习满意度、促进课程目标达成。

在OBE理念指导下，以学习者为中心、学习成果为导向、确定课程教学目标和分目标，制定教学内容体系和教学评价体系，构建混合式教学模式，对课程建设起到了一定推动作用。同时，在实践中也看到了一些不足，期望在后续的研究和改革中不断改进，比如，辅助智能化手段密切关注学生学习动态以加强学习督促和指导、优化教学策略强化学习动机、加强线上小组合作指导和管理等。

参考文献

[1] 蒙跃平,徐桂枝,于艳东 马永强.(2022).基于在线平台的混合式课程的教学建设与探索——以现代教育技术课程为例.中国现代教育装备(03),68-70+73.

[2] 冯晓英,王瑞雪,吴怡君.(2018).国内外混合式教学研究现状述评——基于混合式教学的分析框架.远程教育杂志(03),13-24.

[3] Cora K ,Jochen B.（2022）.The feasibility and effectiveness of a blended-learning course for detecting and avoiding bias in medical data: a pilot study. BMC medical education(01),408.

[4] Badriyah, F., Mubarok, H., & Prahani, B. (2021).Profile of Students' Critical Thinking Skills and the Implementation of PBLRQA based on Blended Learning in Senior High School. Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram,(01): 38-51.

[5] Chung J , Li W , Cheung A T , et al. (2022).Efficacy of a blended learning programme in enhancing the communication skill competence and self-efficacy of nursing students in conducting clinical handovers: a randomised controlled trial[J]. BMC Medical Education(01):1-10.

[6] 李宝敏,余青,杨风雷.混合教学对学生学习成效的影响——基于国内外106篇实证研究的元分析[J].开放教育研究,2022,28(1):75-84.

[7] 粟立丹,丁捷,吴华昌 何江红.(2023).基于OBE理念的混合式教学模式在“食品化学”课程教学中的实践.化学教育(中英文)(12),72-78.

[8] 金鑫,杜静,岳勇 李良军.(2023).新工科背景下OBE驱动机械设计课程创新教学模式.机械设计(04),154-160

[9] 刘新业.(2021).基于OBE理念的“非线性编辑技术”课程改革实践.教育理论与实践(24),52-54.

[10] 丁妍,范慧慧,苏永康,阿茹娜 刘翼婷.(2021).混合式课程教学设计质量与倾向的研究——以全国30门获奖混合式课程为例.电化教育研究(01),107-114.

[11] 王晶心, 原帅 & 赵国栋. (2018).混合式教学对大学生学习成效的影响——基于国内一流大学MOOC应用效果的实证研究.现代远距离教育 (05), 39-47.

[12] 李逢庆,韩晓玲.（2017）混合式教学质量评价体系的构建与实践.中国电化教育(11):108-113.

開發 AI 英語自主學習課程規劃系統並探討其對教師態度之影響

Developing an AI-Enhanced Self-Regulated EFL Course Planning System and Investigating Its Impact on Teachers' Attitudes

劉奕帆[1]，張心洧[2*]，陸以凡[2]
[1]國家教育研究院測驗及評量研究中心
[2]國立中央大學網路學習科技研究所
[*] astrochwwei@gmail.com

【摘要】 人工智慧(AI)進步有助於讓教師運用 AI 來建置課程與學習材料。然而，教師對 AI 技術的熟悉和接受度會影響 AI 工具的教學應用效益，因此培訓教師運用科技於教學是必要的。為解決上述問題，本研究開發 AI 英語自主學習課程規劃系統(ASCPS)，藉由系統鷹架輔助撰寫提示語更有效率的打造英語情境課程及教學材料，另外系統再提供辨識技術提取教師拍攝的教學影片關鍵詞句供參，透過影像與詞句的結合加深學習印象。本研究招募 26 位國中英語教師，透過系統與活動設計習得如何運用 AI 輔助英語教學。問卷結果顯示教師對系統持正向態度，且未來有意願運用 AI 來輔助教學。

【關鍵詞】 人工智慧輔助教學；GPT；自主學習教材；EFL；真實情境學習

Abstract：*Advancements in AI facilitate its use in developing course and learning materials for teachers. However, the familiarity and acceptance of AI technologies among teachers may influence the efficacy of AI tools in educational applications, raising the awareness of integrating technology into teaching training. Therefore, this study developed the AI-Enhanced Self-Regulated Course Planning System(ASCPS) to assist EFL teachers in efficiently creating situated courses and learning materials, and employing recognition technology for key vocabularies and phrases of video material to deepen learning impressions. Surveying 26 middle school English teachers, results revealed a positive perception toward system and willingness to employ AI in future teaching.*

Keywords: AI-enhanced teaching, GPT, Self-regulated learning material, EFL, Authentic learning environment

1. 前言

語言學習是一項複雜的過程，涉及到詞彙的記憶、語法的理解和口語表達能力的培養。Brown et al. (1989)情境學習理論為語言學習提供了一個重要的理論框架，強調語言學習應該在實際使用語言的情境中進行，這對於提高學習效果至關重要。首先，情境學習強調語言應該在真實情境中學習。這意味著學習者應該被置身於需要使用語言的情境中，這樣的學習方式使學習者能夠實際應用所學，加強了語言的實用性，並提高學習的動機。因此，許多研究者運用科技工具輔助針對情境學習的特徵情境感知（Context-aware）、適性化支持（Adaptive support）及適性化介面（Adaptive interface）創造出一智慧學習環境(Lee & Park, 2020; Syno et al., 2022; Wang et al., 2023)。

然而，教師對人工智慧（Artificial Intelligence, AI)技術的熟悉程度和接受度可能影響 AI 工具在教學中的應用效益，許多教師可能缺乏必要的技術知識來充分利用這些工具(Pokrivcakova, 2019)。換言之，教師要如何延續運用 AI 科技在課堂中達到情境學習的效果是值得討論的議題。如 Shapsough and Zualkernan (2020)所建議 AI 可以作為教師在課堂中的輔

助，也更容易能達成情境學習。因此，培訓教師探索 AI 並能自主運用 AI 工具創建學習材料在教學設計中成為提升教育品質的關鍵。

為此，本研究旨在開發 AI-Enhanced Self-Regulated Course Planning System(ASCPS)系統，老師透過系統鷹架(scaffolding)輔助引導運用提示語（prompt）以 AI 產出自主學習課程教案以及學習材料，讓課程及教材更有彈性且符合情境學習。透過 AI 的輔助能讓教師更有效率的完成課程及教材規劃，進而把時間放在個別的學習情況上，藉此達到個別與適性化的學習。

2. 文獻探討

2.1 情境學習對英語學習的影響

情境學習理論為語言學習提供了一個重要的理論框架，強調語言學習應該在實際使用語言的情境中進行，這對於提高學習效果至關重要(Brown et al., 1989)。然而，當前的學習環境往往將知識的習得與其真實情境隔離開，這種現象在外語學習課堂教學中尤為普遍（Lee & Park, 2020）。有研究表明在情境學習環境中，學生能夠在外語寫作過程中更加積極參與和反思，以提升了他們的寫作技能(Al-Ahdal & Alqasham, 2020)。情境學習對英語學習的重要性已在諸多研究中證實可提升學習表現(Lin et al., 2022)。近年來，在科技行動載具快速發展的情況下，學習者能夠在真實或模擬的情境中學習英語，對於學習成效亦能有所助益(Shapsough & Zualkernan, 2020)。

許多學者利用行動載具的可攜性創造無所不在的學習環境，讓學習者能在真實的情境中自由的探索互動與學習(Mohamed & Al-Hidabi, 2019; Safdar et al.)。因此，如何運用現有科技將學習活動融入相應情境的方法變得極其重要(Hakim, 2020)。Chen et al. (2023)開發一套系統結合智能眼鏡技術接收數位教育信息供學習者和情境互動，創造了無縫的“混合現實”學習體驗，學習者也因為接觸新技術而感到有趣因而增加學習動機及學習成效。新興的科技不僅改變了固有的學習模式，也引發了對於現有教學方法的思考(Elsafi, 2020)。

2.2 運用人工智慧(AI)輔助教學

人工智慧(AI)輔助教學的研究領域包括應用人工智慧技術，以及其他相關學習理論（教育學、心理學、神經科學、語言學、社會學和人類學），以促進適性化學習環境和其他有效的 AI 教育工具的發展(Luckin & Holmes, 2016)。隨著 AI 在語言教學領域的應用日益增多，不僅提高了教學效率，還改變了傳統的教學方法(DeCanio, 2020)。Baker et al. (2019)指出教師及教育者主要用 AI 輔助行政管理與監控、教學評估及反饋以減輕教師工作量並提升效率。AI 藉由自動化評估來減輕教師的工作負擔，例如以 AI 代替教師提供學習評估及提供學習建議(McKnight, 2021)。AI 工具能夠在短時間內分析個別學習者數據，診斷個別學習需求，並調整對應的學習內容來達到個人化學習(Pinkwart, 2016)。儘管 AI 在教學中展現了巨大潛力，然而研究指出教師對於 AI 是否具備足夠技術知識來充分發揮這些工具，會影響在教學中運用 AI 工具的效益(Pokrivcakova, 2019)。

Woo and Choi (2021)總結了 2017 至 2020 年間開發的 AI 輔助語言學習工具，強調了教師需要對這些工具有深入了解才能輔助學生更有效運用 AI 工具。資訊科技輔助教學的培訓對教師的教學知識和英語語言能力有顯著影響力，教師應進行相關教學培訓，以發揮更大的教學潛力(Abraham et al., 2022)。上述研究驗證了教師資訊素養的能力學習效益的影響，然而目前大多數的研究較集中在 AI 技術應用於學習層面，對於教師如何有效的整合 AI 工具技術於教學的研究相對較少(Xu, 2020)。

3. 系統設計與實作

3.1 系統設計概念

本研究以上述文獻理論為基礎，研發 ASCPS 系統載於平板電腦中供參與實驗教師使用，活動流程如圖 1。

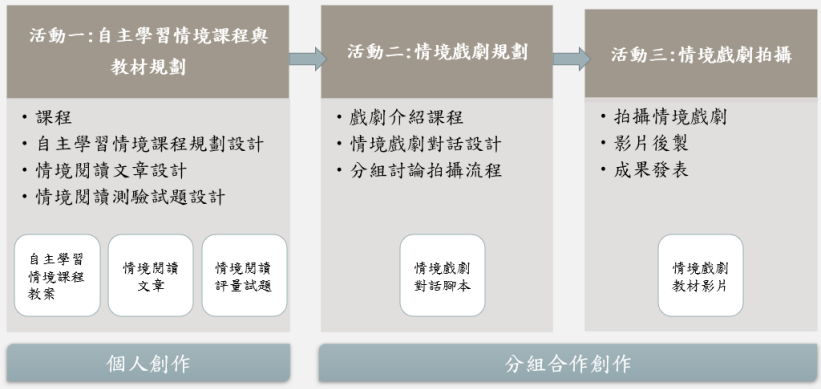


圖 1 ASCPS 活動流程圖

表 1 ASCPS 活動設計說明表

活動	項目	說明	活動產出
規劃活動一：自主學習情境課程與教材	課程	講授 AI 基礎知識概念、AI 現況發展、AI 教育領域相關運用，及如何使用生成式 AI 產出符合需求的學習材料。	
	自主學習情境課程規劃設計	針對教師課程設計時的考量的要素設計鷹架包括：教學對象程度、課程主題、課堂時間安排，以幫助教師更有效率的產出符合需求的教案。	自主學習情境課程教案
	情境閱讀文章設計	根據前項所產出的教案，設計提示語鷹架包括教學對象程度、文章主題、及文章格式以幫助教師產出符合課程規劃的文章。	情境閱讀文章
	情境閱讀測驗試題設計	根據前項所產出的情境閱讀文章，設計提示語鷹架包括題型、題數以及閱讀理解層次的閱讀測驗試題 以幫助教師產出合適的情境閱讀測驗試題。	情境閱讀測驗試題
活動二：情境戲劇規劃	戲劇介紹課程	參與實驗之教師關於情境學習(situated learning)的理論背景及其對語言教育的影響，使教師了解運用生活情境引導學生學習的重要性。之後講授戲劇的組成要素。使教師探討如何結合情境學習與戲劇對話當作教材，了解接下來運用生成式 AI 生成情境戲劇對話腳本時須注意的重點。	
	情境戲劇對話設計	教師透過系統提示語鷹架鍵入符合教學對象程度、主題、場景等相關內容資訊後，產出情境戲劇對話腳本。	情境戲劇對話腳本
	分組討論拍攝流程	教師分組(3 人 1 組)輪流擔任影片的拍攝人員、演員以及旁白，並討論接下來拍攝細節、並練習對話內容。	
活動三：情境戲劇拍攝	拍攝情境戲劇	教師分組(3 人 1 組)至符合設計的情境戲劇對話腳本的地點，輪流擔任影片的拍攝人員、演員。	
	影片後製	教師透過系統影片辨識技術提供的情境字彙及片語延伸或調整旁白內容。	情境戲劇教學影片
	成果發表	組間針對其他組別產製的情境戲劇影片教材以五個面向進行評分： 1.口語及語調清晰度、2.肢體語言、3.合適的面部表情、4.主題明確可做為教材使用、5.影片內容符合真實情境	

系統結合 AI 技術包括 GPT4 語言模型以及影片辨識技術協助教師運用 AI 輔助更有效率的創建情境自主學習材料，包括情境課程規劃、情境閱讀文章、情境測驗試題、情境戲劇對話腳本及情境戲劇教學影片，表 1 ASCPS 活動設計說明表根據活動細節詳細說明。教師先透過系統的鷹架提示語習得如何運用 GPT 有效率的生成相關課程規劃及學習材料。除此之外，拍攝的情境戲劇影片經由系統的辨識技術可提供教師針對辨識出的字彙及句子延伸影片旁白，讓影片內容更豐富。

3.1 活動一：自主學習情境課程與教材規劃

教師在活動一藉由與 GPT 對話來進行自主學習情境課程教案、情境閱讀文章及情境測驗試題，輔助教師更有效率的打造符合教學需求的課程規劃與學習材料。系統根據不同的教學需求提供鷹架提示語輔助讓教師能夠更有效率的設定相關參數得到符合需求的回覆，且系統完整呈現提示語內容，讓教師在研究結束後也可自行運用 AI 打造相關教學材料。教師在使用鷹架提示語與 GPT 對話取得回覆內容後，可儲存放至筆記編輯區，以便在此頁面自行調整教案與教材內容細節如圖 2。

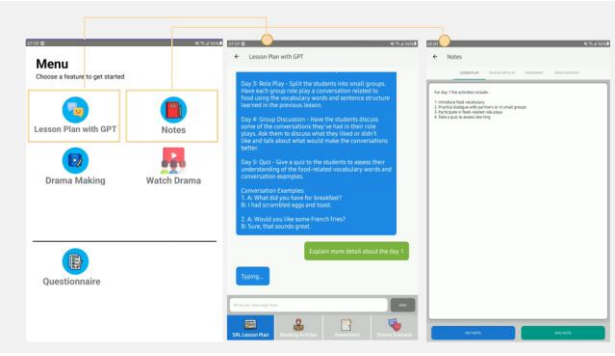


圖 2 活動一系統介面



圖 3 自主學習課程設計

在自主學習情境課程規劃部分，系統已完整設定好課程規劃的提示語架構，教師可根據需求填寫以下資訊即可自動帶入對話框藉由與 GPT 對話生成教案如圖 3 演示。包括：

1 學習者程度：根據 CEFR 架構設定學習者程度產出符合教學程度的課程。

2.課程主題：自定義設定情境領域主題於特定的場域，例如：校園生活、家庭生活等…

3.課程週數：根據教學需求的課程時間設定。

4.示例數量：設定產出與課程相關的會話示例數量。

除此之外，系統以同樣的模式如圖 4 提供教師在情境閱讀文章以及情境測驗試題能藉由鍵入相關資訊後代入對話框後，得到 GPT 回覆結果。更明確的說，在閱讀文章的部分教師可調整內容包括：

1.主題：教師可根據先前教案設計的主題中選出符合教學需求主題置入。

2 文章篇幅：根據學習者程度設定文章篇幅長度，並以句子總數控制整體文章難易度。

3.文章種類：可依照主題及教育需求設定不同的文本題材，例如：說明文、敘述文等…

4.學習者程度：根據 CEFR 架構設定產出符合學習者程度的文章。

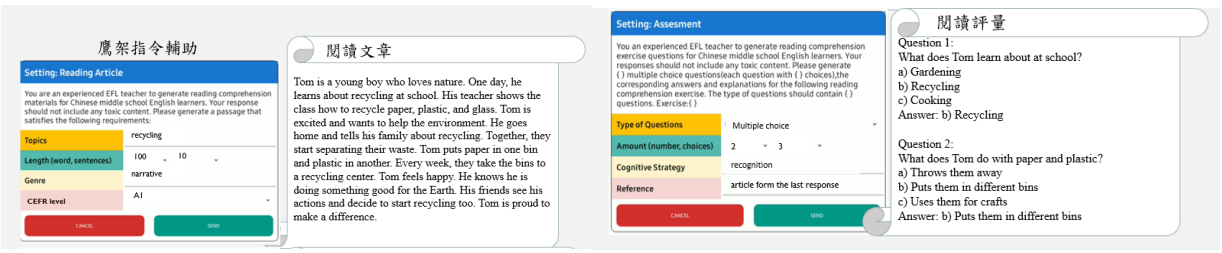


圖 4 閱讀文章及評量設計

另外，在閱讀評量教師可調整的內容包括：

1.試題種類：系統目前預設產出選擇題。

2.試題數量：預計產出的試題數量。

3.閱讀理解思考：可生成不同的閱讀理解層次的試題 (例如：擷取、推論、統整等...)即可完成評量設計。

總體來說，在活動一的部分，教師可藉由系統輔助完成自主學習情境課程規劃、課程中使用的情境閱讀文章以及文章對應的情境測驗試題一整套完整的課程計畫以及搭配使用的學習材料。另外，因參與本研究的教師均為目前在校有實際授課經驗的教師，故除了 GPT 的輔助之外，系統設置了編輯回覆的頁面，教師可根據教學實際經驗及實務考量於對應的頁分頁下編輯回覆內容以更符合實際授課需求。

3.2 活動二：情境戲劇規劃

在情境戲劇對話腳本，系統鷹架提供可調整內容包括：

1.主題：根據教學需求，設定情境戲劇對話的主題。

2.場景：根據設定情境戲劇對話的主題，設定特定場域使產出內容更符合情境。

3.對話篇幅長度：配合學習者程度設定對話長度。

4.學習者程度： 根據 CEFR 架構設定學習者程度。

教師設定完以上資訊後即可得到關於情境戲劇對話的 GPT 回覆。之後便可將 GPT 產生的情境戲劇內容儲存至編輯頁面，根據教學實務經驗修改對話以更符合拍攝時的情境，並可使用此頁面編輯註解紀錄與組員間共同討論接下來的拍攝細節如圖 5 情境戲劇對話設計。

3.3 活動三：情境戲劇拍攝

活動三延續活動二的對話產出成果，將對話拍攝成情境戲劇教學影片。教師運用系統拍攝完情境戲劇教學影片後，按照流程輸入情境戲劇的基本資訊後，經過系統辨識技術輔助提供影片中出現的字彙包含標籤、物件、文字、場景、情緒、語音、角色，這些字彙會經由 GPT 產生成句子。教師能運用辨識技術提供的字彙輔助更輕易描寫出與此影片情境相關的旁白句子，並在後製影片旁白時加入由辨識技術提供的內容及想法。這些經過辨識技術產出的字彙是在影片情境中有出現過的事物影像，因此對於觀看情境戲劇的學習者而言能夠更輕易的將這些字彙與真實情境結合學習，達到更好的學習效益。



圖 5 情境戲劇對話設計



圖 6 活動三系統介面

3.4 研究設計

本研究招募 26 位英文中學教師參與實驗，每位教師於 6 小時實驗中以專題導向學習 (Project-based learning, PBL)運用系統完成研究的活動一~三規劃，詳細實驗設計如圖 7：

- 課程與操作教學：研究者講授 AI 相關知識及運用，並帶領教師操作 ASCPS 系統。
- 活動一：運用 ASCPS 系統指示語輔助產出自主學習情境課程教案、情境閱讀文章及情境閱讀評量試題。
- 分組：教師分為 3 人一組，於接下來的活動二、三合作式學習。

- 活動二：根據活動一的成果以 ASCPS 系統指示語產出同主題的情境戲劇對話腳本並分組討論拍攝細節。
- 活動三：根據活動二產出的情境戲劇對話腳本分組拍攝情境戲劇教學影片，並運用系統辨識技術及 GPT 提供的關鍵詞句撰寫旁白豐富影片內容。
- TAM 問卷：完成活動後，每位教師填寫線上 TAM 問卷。

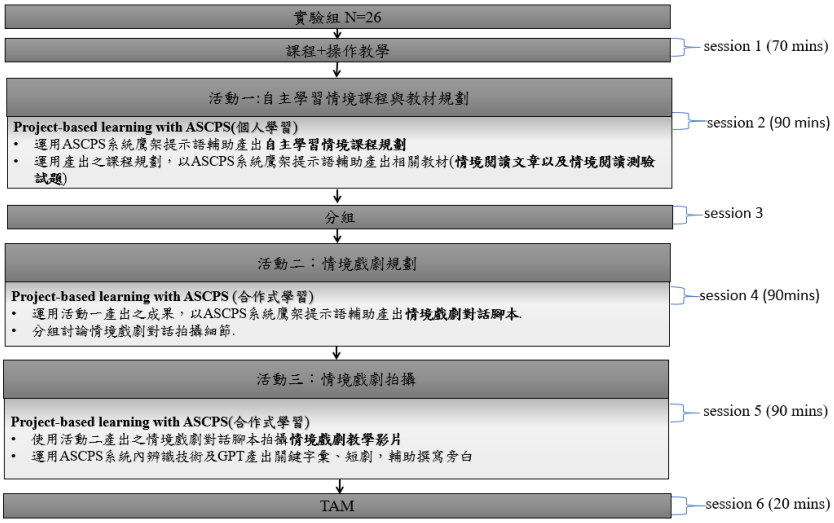


圖 7 實驗設計圖

4. 結果

本研究所使用問卷以 Davis and Venkatesh (1996)所提出的科技接受模型(TAM)問卷為基礎，針對教師對於使用系統進行教案及教學材料研發的看法進行問卷內容設計，目的在於了解教師對於系統的有用性(PU)、易用性(PEU)、態度(AT)，以及使用意圖(BI)。從表 2 可得知，參與教師對於系統的有用性(M=3.808)、易用性(M=3.692)、態度(M=3.679)，以及使用意圖(M=4.256)皆抱持著正面的態度。另外，從各面向所含項目可得知參與教師在經過實驗後，未來不排除開始使用 AI 來輔助教學，這從對於使用易圖面向項目「BI3 我未來不排斥更常使用 AI 工具來輔助教學」數據(BI3, M=4.436)可以得知。

表 2 TAM 細項敘述統計分析

面向	項目	平均數	標準差
Perceived Usefulness (PU) toward ASCPS	**PU1**我認為使用ASCPS所產生的教材可以輔助我的教學表現。	3.808	.8494
	PU2我認為使用ASCPS所產生的教材可以增加我的教材生產力。	3.808	.9389
	PU3我認為使用ASCPS可以增加我的教材生產效率	3.808	.9389
Perceived Ease of Use (PEU) toward ASCPS	**PEU1**我認為使用ASCPS來產出教材資料是容易的	3.615	1.0228
	PEU2我覺得用ASCPS來查找教學相關資料是容易的	3.731	.8274
	PEU3我覺得用ASCPS來產出教材資料是輕易上手的	3.731	.9616
Attitude (AT) toward ASCPS	**AT1**在製作教材時我未來會偏好使用ASCPS來輔助	3.462	1.0670
	AT2我認為用ASCPS產生出的教材能讓我更投入製作教材的過程。	3.769	.8152
	AT3我對於使用ASCPS來產出教材持有正向態度	3.808	.8494
Behavioral Intention to Use AI (BI)	**BI1**我認為我未來會更頻繁使用AI工具來創造教材	4.154	1.1556
	BI2我打算未來會更常使用AI工具來創造教材	4.269	.7776
	BI3我未來不排斥更常使用AI工具來輔助教學	4.346	.7452

5. 結論

本研究 ASCPS 系統實際導入 AI 帶領教師透過 PBL 活動輔助教師研發相關自主學習情境課程規劃、情境閱讀文章、情境測驗試題、情境戲劇對話腳本及情境戲劇教學影片。教師在活動一以系統提供鷹架提示語輔助設定參數包括：教學對象程度、課程主題、課堂時間安排，以幫助教師更有效率的產出符合需求的教案。後續並以此教案內容為基礎繼續以系統提供鷹架提示語輔助產出相關的情境閱讀文章與情境測驗試題。而在活動二的部分，教師根據活動一的成果已以系統提供鷹架提示語輔助鍵入符合教學對象程度、主題、場景等相關內容資訊後，產出情境戲劇對話腳本。除此之外，在活動三拍攝情境戲劇教學影片部分以辨識技術提供情境戲劇教學影片關鍵詞句，教師能根據關鍵詞句撰寫旁白，以此讓學習者在觀看時能透過影像情境與關鍵詞句的結合加深學習印象。教師在運用系統後能更有效率的產出符合教學需求的教案與學習教材。

問卷研究結果顯示教師對於系統的有用性(PU)、易用性(PEU)、態度(AT)，以及使用意圖(BI)皆抱持正向態度，特別是在使用意圖(BI)顯示教師未來不排斥更常使用 AI 工具來設計課程及學習材料。本研究對象為教師，為配合教師空餘時間，因此實驗時程較短，未來將視情況規劃延展實驗時程以利更完整的了解教師的相關行為與其表現，後續並再探討教師產製的教案與教材導入課堂後對學習的影響。

參考文獻

Abraham, M., Arficho, Z., Habtemariam, T., & Demissie, A. (2022). Effects of information communication technology-assisted teaching training on English language teachers' pedagogical knowledge and English language proficiency. *Cogent Education*, *9*(1), 2028336.

Al-Ahdal, A. A. M. H., & Alqasham, F. H. (2020). EFL writing tasks and the application of the concept of situatedness: Evaluating the theoretical and practical aspects of the Saudi EFL context. *TESOL International Journal*, *15*(4), 167-190.

Baker, T., Smith, L., & Anissa, N. (2019). Exploring the Future of Artificial Intelligence in Schools and Colleges. In: NESTA https://media. nesta. org. uk/documents/Future_of_AI_and ….

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *1989*, *18*(1), 32-42.

Chen, H.-R., Lin, W.-S., Hsu, T.-Y., Lin, T.-C., & Chen, N.-S. (2023). Applying Smart Glasses in Situated Exploration for Learning English in a National Science Museum. *IEEE Transactions on Learning Technologies*.

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International journal of human-computer studies*, *45*(1), 19-45.

DeCanio, S. J. (2020). Can an AI learn political theory? *AI Perspectives*, *2*(1), 1-9.

Elsafi, A. (2020). Augmented strategies for mobile and ubiquitous learning technologies. *Emerging Technologies and Pedagogies in the Curriculum*, 245-260.

Hakim, B. (2020). Technology integrated online classrooms and the challenges faced by the EFL teachers in Saudi Arabia during the COVID-19 pandemic. *International Journal of Applied Linguistics and English Literature*, *9*(5), 33-39.

Lee, S.-M., & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning*, *33*(8), 936-959.

Lin, V., Liu, G.-Z., & Chen, N.-S. (2022). The effects of an augmented-reality ubiquitous writing application: A comparative pilot project for enhancing EFL writing instruction. *Computer Assisted Language Learning*, *35*(5-6), 989-1030.

Luckin, R., & Holmes, W. (2016). Intelligence unleashed: An argument for AI in education.

McKnight, L. (2021). Electric sheep? Humans, robots, artificial intelligence, and the future of writing. *Changing English*, *28*(4), 442-455.

Mohamed, A. L., & Al-Hidabi, D. A. Y. (2019). Mobile Device-Based Authentic Learning in Educational Environments: A Systematic Review. *IIUM Journal of Educational Studies*, *7*(2), 40-55.

Pinkwart, N. (2016). Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International journal of artificial intelligence in education*, *26*, 771-783.

Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, *7*(3), 135-153.

Safdar, S. W., Sikandarº, H., & Zafar, I. U-Learning–The design and evaluation of evocative learning-based method for futuristic ubiquitous training and development environment.

Shapsough, S. Y., & Zualkernan, I. A. (2020). A generic IoT architecture for ubiquitous context-aware learning. *IEEE Transactions on Learning Technologies*, *13*(3), 449-464.

Syno, V., Turchyn, A., Syvyk, O., Glotov, O., Krut, O., & Babii, L. (2022). Distance Learning of a Foreign Language: a Comparative Analysis of Modern Platforms and Online Services. *Revista Romaneasca pentru Educatie Multidimensionala*, *14*(2), 461-478.

Wang, S., Wang, H., Jiang, Y., Li, P., & Yang, W. (2023). Understanding students' participation of intelligent teaching: an empirical study considering artificial intelligence usefulness, interactive reward, satisfaction, university support and enjoyment. *Interactive Learning Environments*, *31*(9), 5633-5649.

Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. *arXiv preprint arXiv:2111.04455*.

Xu, L. (2020). The dilemma and countermeasures of AI in educational application. Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence

当 ChatGPT 4.0 遇上医古文——以《黄帝内经·素问》英译为例

When ChatGPT 4.0 Meets TCM Texts—Taking the English Translation of *Huangdi Neijing Suwen* as an Example

王子鹏，刘世英
重庆邮电大学外国语学院
993615393@qq.com

【摘要】 本研究以《黄帝内经·素问》英译版为蓝本，探讨了 ChatGPT 4.0 在应对医古文英译方面的性能，选取两种权威英译版本作为参考对象，分析其特点；通过随机抽样分析《内经》中“理”“法”“方药”三个部分，制定汉英术语库。为保证测试结果客观公正，本研究采用 ICIO 框架为 GPT 编写 prompt，测试其输出值；利用 BLEU 和 TER 算法测试机器的翻译性能，最后通过数据整理和数据对比，形成研究成果。本研究将为利用翻译技术对中华传统文化、技术外宣提供几点有益的参考和启示。

【关键词】 ChatGPT 4.0；黄帝内经·素问；BLEU；TER；ICIO 框架

Abstract: *This study takes the English translation of "Huangdi Neijing Suwen" as the blueprint to explore the performance of ChatGPT 4.0 in translating TCM texts into English. Two authoritative English versions were selected as references for comparison. Through analyzing the "Dao (law of health)," "methods," and "prescriptions and medicine" in "Neijing," a bilingual term bank was established. To ensure the objectivity of the results, this paper used the ICIO framework to write prompts for GPT and tested its performance in accordance with the BLEU and TER algorithms. Finally, the research results were formed through data organization and comparison. This study will provide valuable insights for the use of translation technology in promoting TCM culture and technology in the international communication.*

Keywords: ChatGPT 4.0, Huangdi Neijing Suwen, BLEU, TER, ICIO framework

1. 前言

中医药学文化是中华传统文化的核心象征；中医药技术是广受世界人民好评的的诊疗方法，据世卫组织（WHO）统计，目前有 103 个成员国承认并使用针灸疗法；中药也已经在包括俄罗斯、古巴、新加坡和阿联酋等多个国家正式注册为药品。在 “讲好中国故事、传播好中国声音”的时代背景下，中国传统医学文献的英译成为一项挑战，需要译者不光要精通两种语言，而且要求对中医学术语和古汉语语言有深刻的理解。在这种情形下，对人工译员的要求尚且如此之高，那么对于先进的大语言模型（LLM）而言，是否也会是一项挑战呢？

1.1. 机器翻译模型的选取

生成型预训练变换模型（Generative Pre-trained Transformer，简称 GPT）是由 OpenAI 公司开发的自回归语言模型（Matthews & Dylan, 2023）。作为 2023 年初最新发布的模型，

GPT-4 从各方面来说都优于 OpenAI 之前发布的 GPT-3 和 GPT-3.5。OpenAI 在宣布 GPT-4 时表示，它“比 GPT-3.5 更可靠、更有创意，并且能够处理更细微的指令（Wiggers & Kyle, 2023）。”同时，有学者指出，GPT“尤其是在传统上认为机器翻译并不擅长的文学等领域中，其翻译质量的改善更为明显”；在处理涉及丰富术语的文本时，相较于有道翻译和 DeepL 而言，“GPT 在术语类别上犯的所有翻译错误均为小错，因此在人工评价体系中更受青睐。”（张文煜和赵璧，2024）。另外，此前多数学者的研究重心还停留在 GPT 3.5 模型的测试上，对于 4.0 版模型还少有学者进行相关领域的测试实验。因此，本研究选取 GPT 作为待测的机器翻译对象，旨在探究 ChatGPT 4.0 版应对医古文英译的真实性能表现。

1.2. 《黄帝内经》简介

为探究“ChatGPT 4.0 能否胜任医古文英译”这一课题，本研究选取《黄帝内经·素问》（以下简称《素问》）英译为案例，评估机器翻译模型在该领域所表现出的性能。《黄帝内经》（以下简称《内经》）是中国古代医学四大典籍之首，“为‘上古三坟’之一，是一部统领中国古代医药学和养生学的集大成之作”（杨庆玲，2017）。《内经》以其博大精深的哲学内涵，绚丽精炼的语言文字，吸引了一大批中外译者，《内经》英译方兴未艾。

2. 逻辑方法与研究步骤

本研究由于旨在论证 ChatGPT 4.0 在医古文英译方面的性能表现，首要考虑的就是测评结果的客观性与公正性，但由于实验中明确存在较多变量，如：参考译文取向各异、人工译者主体性与自由度不一、GPT 指令（prompt）标准不一、机器译文测试算法可靠度不一、中英术语版本不一等，因此，本研究综合考量上述因素，尽可能固定变量，力求最大程度实现“不偏不倚”的测试效果。

2.1. 译本分类，分别对比

本研究考虑到医古文的译法因专家译者的风格而异，因此不可因为 GPT 的译文风格与个别专家译员的风格大相径庭，就武断判定 GPT 无法胜任医古文英译，反之亦然。另外，本研究对《素问》根据主要内容进行分类，将其分为“理（天人属性、病因、病机等原理）”“法（诊疗、养生之法）”“方药（方剂、药物）”三部（郑洪新和杨柱，2021），并从三部当中随机抽样进行测试。最后，不应武断得出结论如“GPT 能/否胜任医古文英译”，而应为“GPT 在‘理’/‘法’/‘方药’英译之处有所长，在……之处有所短”。

2.2. 专家译者与译本分析

本研究通过查证相关文献，发现自 1925 年《内经》英译本首次问世以来，共有 18 位译者对《内经》进行了翻译，形成了 24 种不同的《内经》英译版（卢凤姣和吴青，2021）。另外，本研究通过查访亚马逊书城，发现时行的《内经》英译本大致有如下几种：伊尔扎· 威斯（Ilza Veith）本、文树德 (Paul U. Unschuld)本、倪毛信（Maoshing Ni）本、李照国本等。本研究综合考虑上述因素，决定采用倪本及李本两个译本作为测评的参考对象。

倪毛信先生来自中医世家（Maoshing Ni，1994），获得了中美双医学博士学位（梁舒佩等人，2020），其《内经》全译本曾一度登上亚马逊图书《内经》类图书销量榜首（殷丽，2017）；李照国先生则拥有中医学博士学位、语言学与应用语言学专业硕士学位等，译著众多。通过译本分析，本研究发现两位专家译者的翻译风格差异较大，倪本采用目的语读者取向，多符合归化原则，语言流畅、通俗易懂，更接近目标语读者的语用习惯；而李译本采取“译古如古，文不加饰”的原则，旨在为当代读者传递“来自远古的原质信息”（李照国，2009），其译本更注重忠实、作者取向。因此，两种截然相反的译文取向，更适合用来综合考评机器翻译的质量，在很大程度上避免出现“一锤定音”的主观性偏差。需要注意的是，

为避免因术语差异而产生带有偏见的评判，本研究分别为 GPT 提供两套不同的术语库，使其输出的译文在作者取向与读者取向背景下分别进行对比，互不干扰。

2.3. ICIO 框架

本研究采用理工学科试验常用的控制变量法（viable control），固定 prompt 模型，力求将 prompt 差异引起的数据波动降到最低。为此，本研究采用国际通用的 ICIO（Instruction-Context-Input Data-Output Indicator)框架来界定 prompt 编辑标准(Elvis Saravia, 2023）。Instruction: “指令”（希望模型执行的具体任务）; Context: “背景信息”（与指令相关的上下文语境）; Input Data: “输入数据”（告知 GPT 需要处理的数据）; Output Data: “输出指示器”（界定 GPT 输出文本的类型或格式）。

以“方药”部的“汤液醪醴”为例，以 RO 取向为标准，制定整合 prompt:

“Please translate the text (from a TCM classic: The Yellow Emperor' s Classic of Medicine) below from Chinese into English in accordance with the principle of Reader-Orientation and with the terminology:

汤液-Teas; 醪醴-Wines

'''黄帝问曰：‘为五谷汤液及醪醴，奈何？’

岐伯对曰：‘必以稻米，炊之稻薪，稻米者完，稻薪者坚。’'''”

2.4. 算法逻辑

为避免一家之言有失偏颇，降低主观原因带来的数据波动，本研究引入双指标来评判 GPT 译文质量，选用有参考集的自动化评价方法。目前，大部分评测指标采用相似度计算策略，将每个系统译文与一个或多个人工译文进行比较。主流的自动机器翻译评测算法有 BLEU、TER、METEOR 等，但“实际应用时，BLEU 的表现更加稳定，且译文质量也更高” （陈家骏和黄书剑，2017）。本研究选用 BLEU 及 TER 互为参照开展评测实验。

BLEU（Bilingual Evaluation Understudy）是由 IBM 公司提出的评价指标（Papineni et al., 2002），基于 n-gram 模型，比较机器译文中的 n-gram 与参考译文中的 n-gram，计算匹配的 n-gram 数量占机器译文 n-gram 总数量的比例，用于评估机器译文的充分性和流畅性（王子云和毛毳，2023）。BLEU 的数值越高（大于 31.4%），译文质量越好（王金铨和文秋芳，2010）。公式如下：

（1）

$$BLEU = BP \times \exp \sum_{n=1}^{N} \lambda_n \log Precision_n$$

（2）

$$BP = \begin{cases} 1 & \text{if } c > r, \\ e^{1-\frac{r}{c}} & \text{if } c \le r. \end{cases}$$

其中，λn 表示准确率权重，通常取 1; $Precision_n$ 为 n 元文法匹配准确率; BP 为惩罚因子; c 为系统译文(机器译文)长度; r 为参考译文(人工译文)长度。n 元文法的最大阶数，通常设置为 4。

TER（Translation Edit Rate）是由 Snover 等人提出的评价指标，基于编辑距离模型（Snover et al., 2006）。它通过统计机器译文修改为参考译文的后编辑次数来分析机器译文的质量，用于评估机器译文中词汇和语法的错误率。机器译文所需的译后编辑次数越少，译文质量越高（乔博文和李军辉，2020）。TER 指标公式如下：

（3）

$$\mathrm{TER} = \frac{S + I + D + Sh}{R}$$

其中，S（substitution）为“替换操作次数”；I（Insertion）为“插入操作次数”；D（Deletion）为“删除操作次数”；Sh（Shift）“移动操作次数”。TER 值不高于 70.4%判定机器译文合格。

3. 结果分析

如前文所述，本研究将《内经》分为“理”“法”“方药”三部，《素问》全部 8 首方剂皆收入“方药”部待测试序列，另从《素问》“理”部及“法”部随机各择文本 21 处，三部共计 49 处，按照公式（1）至公式（3）分别测试 AO 原则下的 BLEU 值 BA 共计 49 组；测试 RO 原则下的 BLEU 值 BR 共计 49 组；测试 AO 原则下的 TER 值 TA 共计 49 组；测试 RO 原则下的 TER 值 TR 共计 49 组。共测得数据 196 组。碍于篇幅限制，本研究仅以“理”部为例，各择 8 组具有代表性的数据进行展示。

表 1 《素问》“理”部 GPT 4.0 英译质量测评（单位：%）

	1	2	3	4	5	6	7	8
BA	28.69	38.27	28.81	46.71	34.40	20.17	34.31	25.04
BR	14.84	59.86	23.65	17.10	28.53	13.16	23.56	23.21
TA	58.33	42.23	52.00	71.43	39.13	66.17	50.00	52.27
TR	71.43	61.80	53.57	83.87	68.18	67.74	71.05	63.83

判定机器译文质量，需要统计 BLEU 值高于 31.4%的样本量占总样本量的比例，以及 TER 值低于 70.4%的样本量占总样本量的比例，如占比高于 60%，则合格，否则不合格；另外，该数值比例还可以进一步说明，在“理”“法”“方药”三个部类中，GPT 4.0 更擅长完成哪种类型的医古文英译；其次，通过比照同一语段的 BA/BR、TA/TR 值差异，还可以明确 GPT 4.0 理解医古文内涵的能力，例如，如果同一语段的 BR 值通常小于 BA 值，则可以说明 GPT 更擅长使用忠实直译的策略，而并不擅长发掘医古文通常的隐喻意义，不擅长满足读者友好（Reader-Orientation）要求，反之则相反。

根据上述数据及审查方法，绘制柱状图（图 1），易知，在 21 个样本中，若有 13 个满足上述条件，则可以判定为合格（合格率>60%），否则不合格。由图可得，4 组数据中，只有一组 TA 值勉强超过合格线（如图红线），而更加严格的 BLEU 值则距离合格要求尚远。

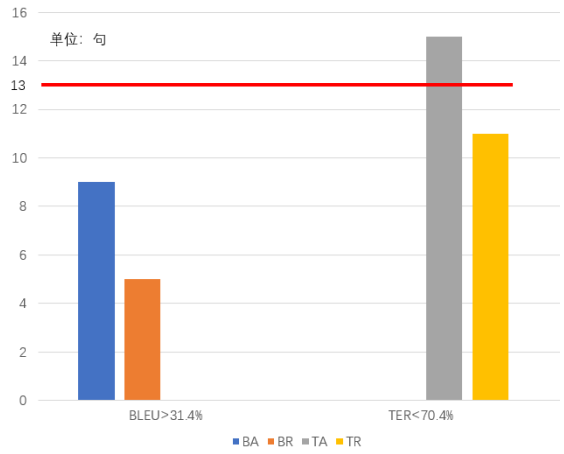


图 1 《素问》“理”部 GPT 4.0 英译质量合格数（单位：句）

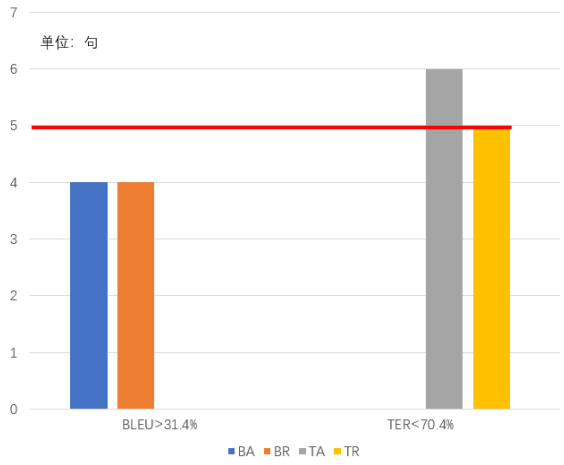


图 2 《素问》“方药”部 GPT 4.0 英译质量合格数（单位：句）

通过图表可以发现，“法”部的数值表现出与“理”部相似的趋势，故此处不再赘述，而值得注意的是，“方药”部的数值似乎更易接近合格要求（如图 2），本文推测原因有二：一是由于样本总量较小，分子相对较大；二是由于“方药”部的样本较少涉及说理内容，因此，只要给 GPT 提供相应的名词术语，它即容易理解把握，产出质量相对较高的译文。

另外，以“理”部 BLEU 及 TER 值分布图可知（图 3、图 4），在 21 个样本中，BLEU 值只有 4 个 BR 大于 BA 的情况，其余皆小于；TER 值表现类似，21 个样本中只有 3 个 TR 大于 TA。“法”部呈现出相似趋势，而“方药”部如上述，虽然仍普遍低于合格线，但呈现性能皆优于另两部。

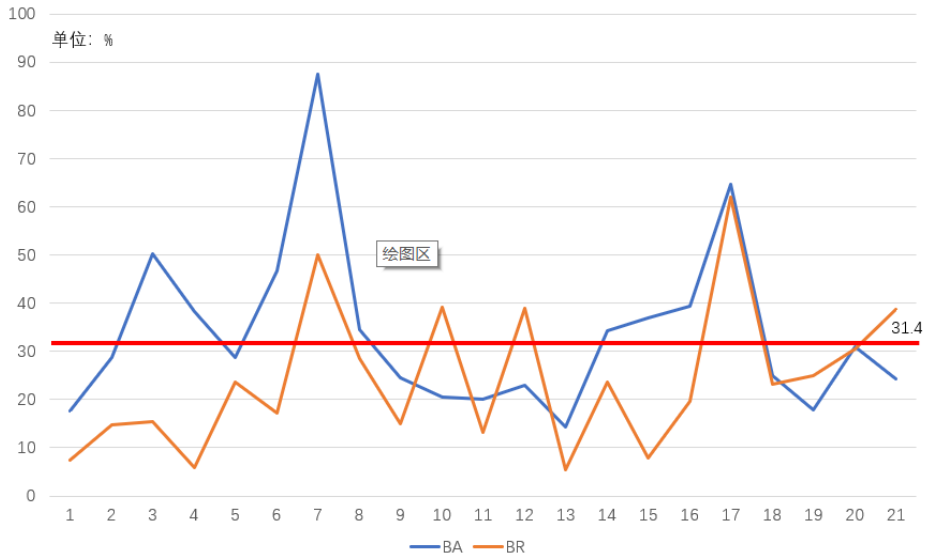


图 3 “理”部 GPT 4.0 英译 BLEU 值分布

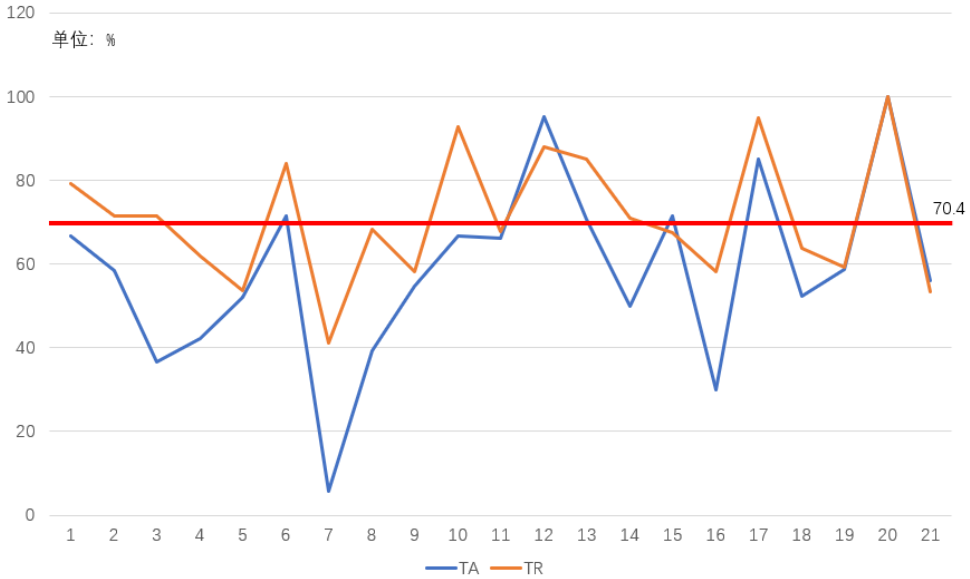


图 4 “理”部 GPT 4.0 英译 TER 值分布

3.1. 译文问题分析

如上述可知，在已经为 GPT 引入背景（Context）及差异化术语库的前提下，它输出的译文仍表现出如下缺陷。

例 1：“逆春气，则少阳不生，肝气内变。”——《四气调神大论篇第二》

AO: Violation of Chunqi (Spring-Qi) will prevent Shaoyang from growing, leading to diseases due to stagnation of Ganqi (Liver-Qi).

GA: Going against the Chunqi (Spring-Qi) results in the failure of Shaoyang to arise, leading to internal changes in Ganqi (Liver-Qi).

RO: If one does not follow the play of the elemental energies according to the seasons, the liver energy will stagnate, resulting in illness in spring.

GR: Going against the energy of spring will prevent the Lesser Yang from arising, and the liver energy will stagnate.

如例 1，其中 GA 与 GR 分别为要求 GPT 以作者取向与读者取向得到的输出情况。从作者的角度来看，中医讲“肝”通常讲“肝气以疏通、畅达为顺，不宜抑制、郁结，（郑洪新和杨柱，2021）”因此，所谓“肝气内变”一定是在讲肝系统的病理性发展，即“肝气郁结（stagnation of Ganqi），” 很明显，GPT 并不能精准把握这一信息，仅从字面直译为中规中矩的“internal changes”；另一方面，“少阳”一词的说法出自《易经》，与“老阳”或“太阳”相对，而从读者友好的角度来看，英语国家的文化中并没有“少阳”一词，该词属于文化缺省的范畴，如果处理不当，很容易造成目的语读者的“文化休克”，而 GPT 将其直译处理为“Lesser Yang”，虽然回译性很强，但容易给英语读者带来不知所云之感，无法满足读者友好要求。

例 2：“故春善病鼽衄，仲夏善病胸胁，长夏善病洞泄寒中，秋善病风疟，冬善病痹厥。”——《金匮真言论篇第四》

AO: That is why Qiunv (nasal stuffiness and bleeding) is usually seen in spring, chest and rib-side disorders are often seen in summer, Hanzhong (internal cold syndromes) like Dongxie (acute diarrhea) are frequently seen in late summer, Fengnve (Wind-Malaria) is often seen in autumn and Bijue (numbness and coldness of the four limbs) is always seen in winter.

GA: Therefore, in spring, Qiunv (nasal stuffiness and bleeding) is prone to occur; in midsummer, chest and rib-side disorders; in late summer, Dongxie (acute diarrhea) and Hanzhong (internal cold syndromes); in autumn, Fengnve (Wind-Malaria); and in winter, Bijue (numbness and coldness of the four limbs).

RO: In the spring, illness occurs in the liver channel and rises to the head, causing bleeding from the nose. In the summer, the wind affects the heart. Late summer will mainly affect the spleen, causing internal colds with diarrhea. The wind of autumn will affect the lungs, manifesting in malaria with alternating chills and fever. The winds of winter will affect the kidneys and limbs, manifesting in bi syndrome, a condition of obstruction of qi and blood.

GR: Therefore, in spring, nosebleeds are prone to occur; in midsummer, chest and flank diseases; in late summer, internal colds with diarrhea; in autumn, malaria with alternating chills and fever; in winter, bi syndrome, a condition of obstruction of qi and blood.

如例 2，BA（38.27%>31.4%）及 TA（42.25%<70.4%）值均表现优异，而 BR 及 TR 值则相对较差。可以看出，在专业术语较多且要求实现作者取向的前提下，GPT 可以产出性能相对优越的文本，如 GA；而如果要求实现读者取向，样本中较多的专业术语可能反而会增大 GPT 英译的难度，如 RO 译本中，倪先生基本将每个术语的内在原理都做了简单解释，让英语读者可以快速理解，可以断定，这一点是 GPT 无法做到的。

例 3：“阴阳者，天地之道也，万物之纲纪，变化之父母，生杀之本始，神明之府也，治病必求于本。”——《阴阳应象大论篇第五》

AO: Yin and Yang serve as the Dao (law) of the heavens and the earth, the fundamental principle of all things, the parents of change, the beginning of birth and death and the storehouse of Shenming. The treatment of disease must follow this law.

GA: Yin and Yang are the Dao of heaven and earth, the framework of all things, the parents of change, the origin of life and death, and the abode of Shenming (divine brightness). To treat diseases, one must seek their root cause.

RO: The law of yin and yang is the natural order of the universe, the foundation of all things, mother of all changes, the root of life and death. In healing, one must grasp the root of the disharmony, which is always subject to the law of yin and yang.

GR: Yin and Yang are the natural order of the universe, the framework of all things, the parents of change, the origin of creation and destruction, and the residence of divine intelligence. To cure diseases, one must seek their root cause.

如例 3，中医不是“头痛医头”的理论，而是高度注重整体性的医学，同理，医古文的整体性和内在逻辑性也贯穿着文本的始末。在同样提供语言背景和术语库的前提下，GPT 仍只是机械地按照标点分隔，对意群进行逐个翻译，并未考虑到医古文中常有的严谨逻辑，割裂了意群彼此之间的有机联系，如“治病必求于本”，这里的“本”并不单指“root cause”，而是总结性地概括前文的原理，意为“要治病，就要遵循上述的基本原理。”

4. 结语

根据对《黄帝内经·素问》的英译实验，采用李照国、倪毛信译本作为参考，通过 BLEU 值和 TER 值的评测，本研究发现 ChatGPT 4.0 在医古文英译方面的表现尚未达到合格标准。尽管 GPT 4.0 在性能上优于其前身 GPT 3.5，并且在一定程度上能够保证译文的忠实性、语言的正确性和流畅性，成为人工译员的有力辅助工具，但综合评估数据显示，它仍然不足以独立完成医古文的英译任务。

具体而言，在应对专业术语较多且较为固定、译者主体性有限且译本不易发生太大变化的情况下，采用作者取向、异化、直译、忠实等原则时，GPT 即可展现出较为优异的性能，如约定俗成的方剂药物的名称英译；而源文本中一旦涉及医理、易理等多具有隐喻意义的哲学文化表达时，GPT 所表现出的性能则难以满足要求，如“理”“法”的英译。

综上，从本研究实验可以看出，人工智能（AI）和大语言模型尽管在语言处理和翻译领域取得了长足进步，但在处理专业性强、语言结构复杂的医古文英译任务时，机器翻译仍然面临着巨大挑战。可以预见，在未来相当长的一段时间内，人工翻译和机器辅助翻译将继续并存，而“机器辅助翻译+译前译后编辑”模式的地位仍将持续占据医古文英译实践的主流。

因此，对于技术开发人员而言，未来的研究和技术开发应更加注重提高机器翻译系统对专业文本和复杂语言结构的处理能力，以期望机器翻译在医古文英译领域的应用能够取得更大的突破；另外，针对广大矢志翻译的译者、学者而言，一方面应当与时俱进，努力提升技术素养，一方面也应当古为今用，丰富自身文化技术素养，将中华传统技术文化发扬光大。

基金资助：重庆市语言文字科研项目，智能时代大语言模型与人工翻译的协同模式研究（No.YYK23203）。

参考文献

王子云和毛毳（2023）。ChatGPT 译文质量的评估与提升————以陶瓷类文本汉英翻译为例。**山东陶瓷，2023（8）**，20-27。
王金铨和文秋芳（2010）。国内外机器自动评分系统评述：兼论对中国学生翻译自动评分系统的启示。**外语界，2010（1）**，75-81，91。
卢凤姣和吴青（2021）。黄帝内经·素问威斯译本的译者主体性分析研究。**北京中医药大学硕士研究生论文**，3-18。
乔博文和李军辉（2019）。融合语义角色的神经机器翻译。**计算机科学**，164-168。
陈家骏和黄书剑（2017）。基于参考译文图的机器翻译质量评价方法研究。**南京大学研究生毕业论文**，2-26。
李永安和李亚军（2021）。**中医翻译十二讲**。西安：西安交通大学出版社。
李照国（2009）。**大中华文库汉英对照黄帝内经素问**。西安：世界图书出版公司。
杨庆玲（2017）。**中华经典国学黄帝内经**。西安：三秦出版社。
杨韫菲（2019）。黄帝内经成书考。**文学评论·古典文学**，97。
张文煜和赵璧（2024）。生成式人工智能开创机器翻译的新纪元了吗？——一项质量对比研究及对翻译教育的思考。**北京第二外国语学院学报，2024（1）**，13。
郑洪新和杨柱（2021）。**中医基础理论**。北京：中国中医药出版社。
周成彬和刘忠宝（2022）。基于语义信息共享 Transformer 的古文机器翻译方法。**情报工程，2022，8（6）**，114-127。
殷丽（2017）。中医药典籍国内英译本海外接受状况调查及启示——以大众化文库黄帝内经英译本为例。**外国语，2017，40（5）**，33-43。
梁舒佩、李琳和蒋基昌（2020）。接受理论视域下黄帝内经英译研究——以倪译本和李译本为例。**环球中医药，2020（10）**，1783。
戴光荣和刘思圻（2023）。神经网络机器翻译：进展与挑战。**外语教学，2023（1）**，82。
Elvis. (2023). *Prompting Introduction*. Retrieved from https://github.com/dair-ai/Prompt-Engineering-Guide/blob/main/guides/prompts-intro.md.
Philipp.K., Monz.C. (2006). *Manual and Automatic Evaluation of Machine Translation between European Languages*. Retrieved form http://dl.acm.org/citation.cfm?id=1654650.1654666.
Maoshing Ni. (1995). *The Yellow Emperor's Classic of Medicine*. Boston and London: Shambhala Publications Inc.
Matthews, & Dylan. (2023). *23 Things We Think Will Happen In 2023*. Retrieved from https://www.vox.com/future-perfect/2023/1/2/23494204/biden-trump-putin-supreme-court-2023-predictions.
Snover.M., Dorr. B, & Schuarts. R. (2006). A Study of Translation Edit Rate with Targeted Human Annotation. *Technical Papers*: *Proceedings of the 7th Conference of the Association for Machine Translation in the Americans Cambridge*, 223-231.
Wiggors., & Kyle. (2023). *OpenAI Release GPT-4, A Multimodal AI That It Claims Is State-of-the-Art.* Retrieved from https://techcrunch.com/2023/03/14/openai-releases-gpt-4-ai-that-it-claims-is-state-of-the-art/

基于语文核心素养的动画创作对小学生古诗学习成效的影响研究

Research on the Influence of Animation Creation Based on Chinese Core Accomplishment on the Learning Effect of Primary School Students' Ancient Poetry

李铭壹，蔡薇薇，王丹妮，张新立
温州大学教育学院
zhxl@wzu.edu.cn，

【摘要】语文核心素养是核心素养在语言领域上的延伸，古诗是培养学生核心素养的有力方法。本研究在建构主义、认知主义等学习理论的指导下，构建了融入动画创作的小学古诗项目化教学模式，采用实验组对照组前后测实验开展了该教学模式有效性的检验，结果表明在该教学模式下学生的学习成绩提升效果优于传统的小学古诗教学模式，语文核心素养培养效果也优于传统的小学古诗教学模式。
【关键词】动画创作; 核心素养;小学古诗学习;

Abstract: *Chinese Language core literacy leverages classical poetry to develop core competencies. Guided by constructivism and cognitivism, this research created a project-based teaching model with animation for elementary ancient poetry. The results revealed that students' academic performance improved more significantly compared to traditional approaches to learning ancient poems and there was a greater enhancement in students' core language literacies than with conventional methods.*
Keywords: *animation creation, core literacy, elementary ancient poetry*

1．研究背景及意义

古诗作为中国传统文化的精华，对于语文教育具有不可或缺的地位。掌握古诗不仅有助于学生学习古汉语、深入了解中华历史文化，而且对于提升语言和综合素养、增强人文素养及审美情趣至关重要(李秀平, 2018) (冯克芝 & 汪乐, 2021)。 然而，当前小学古诗教学面临诸多挑战，如学生兴趣缺失和教学方法的单一性。为适应素养教育的新趋势，迫切需要将核心素养培育整合入古诗教学，以改善现状。

动画作为教学工具，因其感官体验丰富、信息表达直观、形式多样及操作简便等优势，有效激发学生兴趣、提升抽象思维与学科素养 (杨亮 & 文载喆, 2021) (Gao, 2018)。此外，动画创作教学促进学生想象、创造与讨论，增强理解与参与(Yaseen & Aubusson, 2020) (Reed, Hagen, Bukach, & Couperus, 2021)。

因此，本研究在建构主义、认知主义等学习理论的指导下，结合小学生身心发展特点，基于语文古诗的要求，立足学生语文核心素养的培养，将动画创作技术与小学古诗学习相结合，构建了融入动画创作的小学古诗项目化教学模式，并试图通过实践探索动画创作技术与小学古诗融合对学生学习成效的影响，分析基于语文核心素养下将动画创作与小学古诗融合的独特意义。

2．文献综述

2.1.小学语文古诗教学方法研究现状

当前小学古诗教学存在古诗词课堂教学中缺少趣味性引导、忽略与作者创作背景的联系、忽视了诗歌的意境等诸多问题(孙芳芳 & 安晶, 2010)。且有相关研究中指出小学古诗教学过分强调背诵记忆，过分注重对于古诗中字词的解读的问题(杨莉声, 2014)。国内研究表明，小学古诗教学传统方法如诵读、讲授和比较等存在局限(韩子文, 2018)。随着研究进展，这些方法因过于注重知识传授、缺乏互动和实践联系，导致学生兴趣下降和认知负荷增加(唐明, 2020) (王婧 & 张建平, 2020) (张文娟, 2020)。

2.2.动画技术支持学科教学应用研究现状

动画技术研究揭示了其在感官体验、信息表达、形式多样性及操作便捷性方面的优势(杨亮 & 文载喆, 2021) (Gao, 2018)。作为教学工具，动画展现了卓越的表现力，能有效提升学生的学习兴趣、抽象思维及学科素养(陈胜贤, 2018)。

在国外，动画技术在教育教学方面的研究在计算机科学与技术、工程、医学、社会学等学科应用，并且国外学者已经通过研究证实用动画技术支持教学有助于工程(Ong & Mannan, 2004; Shen & Lu, 2021)、数学(Wang, 2021)、物理(Hasegawa et al., 2010)、体育(Xia, 2016)等多学科教学效果的提升。近年来，国内学者对动画技术在教学领域的应用进行了广泛研究，证实其在计算机科学、自然科学、医学和工程技术等学科中能促进学生数学(胡锦娟; 廖淑华 & 许垂泽, 2004)、语文(丁道群, 余娜, & 姚向云, 2008; 冯姣姣, 马月, & 杨方琦, 2018)、物理(王慧 & 刘力, 2011)和地理(陈胜贤, 2018)等学科成绩的提高，并有效提升学生的学习兴趣和课堂参与度。

因此，动画在教学中的应用能有效提升学生的学习兴趣、抽象思维和学科素养(李文清, 2009)。首先，动画将抽象概念具体化，简化了复杂知识的理解和记忆。其次，通过创造直观的感官体验，动画减少了学生想象的难度，提高了学习效率和积极性，使课程内容更加生动有趣，从而促进了学生深层次的学科思维能力发展(刘应梅, 2021)。因而，将动画技术与语文古诗教学相结合以提升学生的学习成效，很有深入研究的必要性。

3. 融入动画创作的小学古诗教学模式构建

本研究在建构主义学习理论、具身认知理论、体验学习理论的指导下，结合动画创作的过程、项目化学习过程以及古诗教学特点构建了融入动画创作的小学古诗项目化教学模式，如图 3-1 所示。

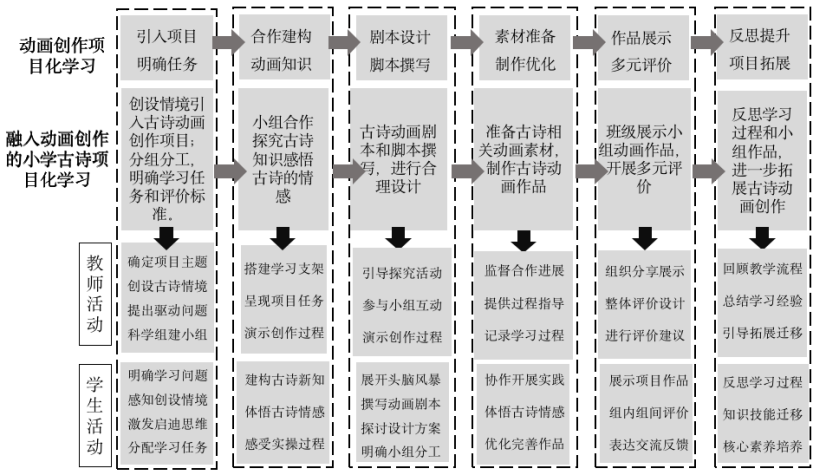


图 3-1 融入动画创作的小学古诗教学模式

4.融入动画创作的小学古诗项目化学习效果研究

4.1.实验目的

本实验的目的是检验融入动画创作过程的小学古诗项目化学习对学生学习成绩以及语文核心素养的影响。

4.2.研究假设

融入动画创作的小学古诗项目化学习效果优于传统的小学古诗学习效果。

4.3.实验设计

4.3.1.实验对象及内容

本研究在某小学五年级随机抽取的两个班级作为实验对象，其中实验班 42 人，男生 19 人，女生 23;人， 对照班 42 人，男生 20 人，女生 22 人。本实验内容选自教编版五年级上册语文教材中的两首古诗《山居秋暝》《枫桥夜泊》作为实验内容。实验周期为四周，每周 2 次。

4.3.2.研究变量

本研究自变量是教学模式，包括两个水平，即实验组采用融入动画创作的古诗项目化学习模式，对照组采用常规古诗学习模式。本研究的因变量包括学生的古诗学习成绩和语文核心素养。

4.4.实验流程

实验开始前，由信息技术教师先指导学生使用优芽互动电影平台学习动画制作技术，适应平台的使用流程与方法，为开展本实验教学做准备。本实验流程图如图 4-1 所示。

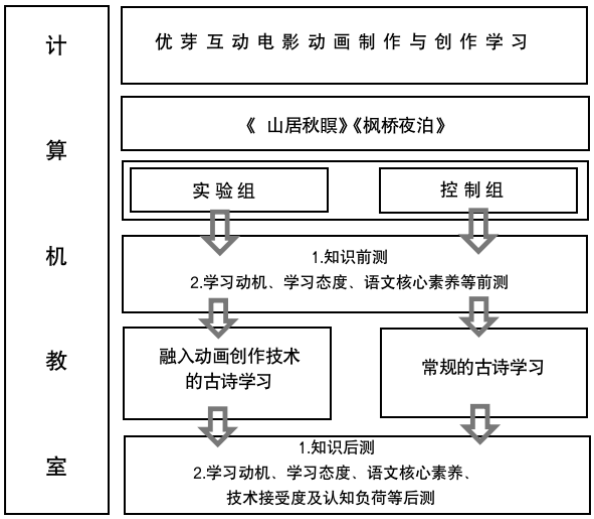


图 4-1 实验流程图

4.5.研究工具

古诗知识测试由学科专家和语文教师共同开发，包含 16 题，满分 100 分。在语文和教育技术专家指导下，与经验丰富的小学语文教师合作制定了《基于语文核心素养的小学古诗学习情况》问卷，并通过多次迭代优化。问卷信效度分析显示克隆巴赫系数为 0.856，表明量表可靠。实验数据使用 SPSS 26.0 软件进行分析。

5.实验数据分析

数据显示，实验前实验组与对照组学生知识测试分数均值相差 0.2619。*t* 检验的显著性概率（Sig.）为 0.924，大于 0.05；问卷测试数据的独立样本 *t* 检验结果显示，七道题目的 *p* 值均远大于 0.05。以上数据说明实验组与对照组学生的语文古诗的知识水平和语文核心素养处于相近的水平，不存在显著性差异。结果符合教育实验的要求。

5.1. 语文核心素养各维度下古诗知识测试成绩结果对比分析

结束后，对两组学生语文古诗知识进行后测，并对后测数据进行对比分析。

表 5-1 实验后两组被试知识测试情况

组别	个案数	最小值	最大值	平均值	标准差	标准误差平均值
实验组	42	36	87.00	62.0714	12.21034	1.88410
对照组	42	25	78.00	55.5714	13.21703	2.03943

如表 5-1 所示，实验组与对照组学生的知识测试成绩最小值、最大值、平均分均得到提升，且实验组测试成绩最小值及平均分高于对照组，实验组测试平均分增值远大于对照组平均分增值。

表 5-2 实验后两组被试知识测试成绩的独立样本 *t* 检验

	莱文方差等同性检验		**平均值等同性 *t* 检验**					
	F	显著性	*t*	df	*Sig.*(双尾)	平均值差值	差值 95%置信区间	
							下限	上限
假定等方差	.007	.935	2.341*	82	.022	6.500	.977	12.023
不假定等方差			2.341	81.491	.022	6.500	.977	12.023

*$p<0.05$.

如图表 5-2 所示，在进行实验研究后，实验组与对照组知识测试成绩的独立样本 *t* 检验数据显示，*t* 检验的显著性概率（Sig.）为 0.022，小于 0.05，说明两组学生在教学实验后出现了古诗测试成绩上的显著差异；同时平均值差值 95%置信区间（0.977，12.024）不含 0。

5.2. 实验组和对照组后测语文核心素养结果对比分析

如图表 5-3 所示，在进行实验研究后，实验组与对照组语文核心素养问卷测试结果的独立样本 *t* 检验数据显示，七道题目中有六道题目的 *t* 检验的显著性概率（Sig.）为均小于 0.05。由此可见，实验组学生语文核心素养后测数据高于对照组学生合作学习倾向后测数据。

表 5-3 实验后两组被试语文核心素养的独立样本 *t* 检验

	实验组(N=42) M ±SD	**对照组(N=42) M ±SD**	**df**	*t*	*p*
题目 1	4.24±0.62	3.93±0.60	82	2.329*	.022
题目 2	4.48±0.78	4.19 ±0.71	82	2.131*	.036
题目 3	4.21 ±0.83	3.86 ±0.81	82	2.180*	.032
题目 4	4.00 ±0.77	3.64±0.66	82	2.979*	.024
题目 5	4.07 ±0.73	3.55 ±0.92	82	3.038*	.003
题目 6	4.00 ±0.76	3.71 ±0.77	82	1.909	.060
题目 7	4.19 ±0.77	3.74 ±0.70	82	3.192**	.002

*$p<0.05$. **$p<0.01$.

以上结果与实验假设“在学生的语文核心素养方面，融入动画创作技术的小学古诗学习效果优于传统的小学古诗学习”相一致。

5.讨论与结论

经过对比分析，实验组在融入动画创作的古诗教学后，其学习成绩显著优于对照组。这一结果验证了动画能有效提升学生对古诗的理解和兴趣，与先前研究一致。动画的直观性和生动性降低了古诗词理解难度，促进了学生对古诗意境的把握，同时，动画创作过程有助于学生将隐性知识显性化，激发主动学习，实现对古诗内容的深入理解，且不受限于学生的先验知识和技能水平。

通过对比分析问卷数据表明，结合动画创作的小学古诗学习方法较传统方式更有效地提升了学生的语文核心素养，这与现有研究一致(Bernát, Pavlovkin, Dmura, áok, & Bernátová, 2018; Sandir, Arikan, & Akbyk, 2007)。动画的内容、形式和制作过程与古诗教学相辅相成，其动态和直观的呈现方式促进了学生的多感官参与，并提高了他们对古诗知识的掌握和应用能力，从而增强了文化理解、审美和创造力等核心素养，这与国内项目式古诗教学研究结果相符(蔡界平, 2021; 陈瑜, 2011; 吴蔚, 2008)。

综上所述，融入动画创作的小学古诗教学模式能够有效提升不同成就学生的古诗学习成绩和语文核心素养。本研究得出结论：融入动画创作的小学古诗项目化学习效果优于传统的小学古诗学习效果。实验研究数据显示在两种不同模式的小学古诗教学下，两组被试学生的古诗学习成绩和语文核心素养发生了不同程度的变化。

6.研究不足

随着生成式人工智能的发展，未来研究应关注教学中生成式人工智能创作动画，以此来提升学生的学习成效，学习动机以及核心素养。

参考文献

Bernát, D. I. M., Pavlovkin, I. J., Dmura, J., áok, & Bernátová, R. (2018). Some possibilities of using Flash animations and Applets in increasing the effectiveness of teaching. *Journal of Science Education, 19*(1), 32-45.

Gao, Y. (2018). Reproduction of chemical experiment based on computer 2d animation technology. *Chemical engineering transactions, 66*, 835-840.

Hasegawa, T., Kojima, H., Masu, C., Fukushima, Y., Kojima, H., Konokawa, K., . . . Maruyama, K. (2010). Creation and application of three-dimensional computer-graphic animations for introduction to radiological physics and technology. *Radiological Physics & Technology, 3*(1), 1-9.

Ong, S. K., & Mannan, M. A. (2004). Virtual reality simulations and animations in a web-based interactive manufacturing engineering module. *Computers & Education, 43*(4), 361-382.

Reed, C. L., Hagen, E., Bukach, C. M., & Couperus, J. W. (2021). Effectiveness of Undergraduate-Generated Animations: Increasing Comprehension and Engagement for Neuroscience Majors and Non-Majors. *Teaching of Psychology*(1), 009862832110230.

Sandir, H., Arikan, A., & Akbyk, C. (2007). *Teaching mathematical concepts using Flash animations.* Paper presented at the ED-MEDIA 2007 : World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Shen, Y., & Lu, W. (2021). Technology of animation in multimedia courseware of mechanical drawing. *The International Journal of Electrical Engineering & Education*, 002072092098608. doi:10.1177/0020720920986085

Wang, Y. (2021). Multimedia technology-assisted stereo geometry teaching. *International Journal of Electrical Engineering Education*, 002072092098468.

Xia, C. (2016). Application of Continuous Animation Production Technology in Gymnastics Multimedia Teaching. *International Journal of Emerging Technologies in Learning, 11*(9), 46.

Yaseen, Z., & Aubusson, P. (2020). Exploring Student-Generated Animations, Combined with a Representational Pedagogy, as a Tool for Learning in Chemistry.

蔡界平. (2021). 小学低年级音乐吟诵古诗教学有效策略研究. *新课程（教研版）, 000*(028), 67.

陈胜贤. (2018). Flash 动画在高中地理信息化教学中的应用. *西部素质教育, 4*(22), 2.

陈瑜. (2011). 试论小学古诗教学中的情感体验法. *课程・教材・教法*(4), 6.

丁道群, 余娜, & 姚向云. (2008). 动画呈现方式对小学四年级学生古诗学习的影响. *课程・教材・教法*(9), 5.

冯姣姣, 马月, & 杨方琦. (2018). 基于 Flash 动画的小学语文古诗词情境创设教学优势分析. *教育现代化*.

冯克芝, & 汪乐. (2021). 浅谈小学古诗美学教育的方法探究. *安徽教育科研*(17), 2.

韩子文. (2018). *媒介融合视野下的小学语文古诗词教学方法研究*. 华中师范大学,

胡锦娟. *Flash 动画在中学数学函数教学中应用的有效性研究*. 广州大学,

李文清. (2009). 三维动画在教学的应用及在生物化学中的设计. *软件导刊.教育技术*(1), 2.

李秀平. (2018). 小学古诗教学中关键词教学策略探讨. *青少年日记：教育教学研究, 000*(0S1), P.189-189.

廖淑华, & 许垂泽. (2004). 多媒体技术在教学中的应用——Flash 动画制作在计算机辅助教学中的作用. *伊犁教育学院学报, 17*(1), 5.

刘应梅. (2021). 高中生物膜系统概念的教学动画开发与应用研究.

孙芳芳, & 安晶. (2010). 论小学古诗词教学中存在的几个问题. *太原城市职业技术学院学报*(2), 2.

唐明. (2020). 小学低段古诗教学策略——以贺知章《咏柳》为例. *语文建设*(6), 3.

王慧, & 刘力. (2011). 中学物理实验教学中 Flash 动画的制作与应用——以"弹簧振子实验"为例. *中国信息技术教育, 000*(016), 107,131.

王婧, & 张建平. (2020). 小学语文古诗教学现状及启发——以 N 市某小学为例. *现代商贸工业, 41*(7), 2.

吴蔚. (2008). *论传统的吟诵方法在古诗词教学中的运用*. 湖南师范大学,

杨莉声. (2014). *新课标视野下的小学高段古诗教学的现状与对策研究*. 四川师范大学,

杨亮, & 文载喆. (2021). 作为物质,空间与感官的动画:一种基于媒介考古学的考察. *当代动画*(2), 7.

张文娟. (2020). 新课程下小学语文古诗教学的基本策略探析.

演化博弈理论在教育研究中的应用分析

Application Analysis of Evolutionary Game Theory in Educational Research

张盛[1]，杨现民[2*]，李新[3]
[123]江苏师范大学智慧教育学院
[*] yangxianmin8888@163.com

【摘要】 演化博弈理论作为一种探究主体行为演化路径及事件演变趋势的重要方法，受到教育研究者的广泛使用。本研究系统梳理了演化博弈理论在教育研究中的应用现状，总结出演化博弈理论具有的教育应用优势以及适用场景，对演化博弈理论在教育研究中的未来发展趋势进行展望，以期为教育研究者提供全新的理论性指南。
【关键词】 演化博弈理论；教育研究；教育主体行为；未来趋势

Abstract: *Evolutionary game theory, as a significant approach for exploring the evolutionary paths of agents' behaviors and the trends of event evolution, has been widely used by educational researchers. This study systematically reviews the current applications of evolutionary game theory in educational research, summarizes its advantages and applicable scenarios in educational applications, and provides a prospective view on the future development of evolutionary game theory in the field of education research, aiming to offer a novel theoretical guideline for educational researchers.*
Keywords: *Evolutionary Game Theory, Educational Research, Educational Subject Behavior, Future Trends*

基金项目：国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”(项目编号：62077030)；2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）

1. 引言

教育是一个介于完全结构化与混沌之间的复杂系统，包含外部社会文化环境和内部教育生态两部分（祝智庭 & 彭红超，2020），涉及众多参与者和利益相关主体（闫寒冰 & 陈怡，2023），其之间关系复杂且利益交织，尤其是在互联网快速发展的时代，主体信息共享和沟通交流的渠道不断拓宽，使得教育主体与技术关系复杂多变。在此背景下，明晰教育主体行为关系是破解教育问题和促进教育系统良性运转的关键核心。

演化博弈理论基于生态学视角，通过建立数学方程模型动态呈现多个主体间的博弈行为，为探明主体行为策略与博弈关系提供新的路径选择。已有研究表明，利用演化博弈理论探究群体中利益相关主体的博弈过程（Santos F P et al.，2015），有助于厘清学校管理、人才培养等方面的主体关系，实现教育体系高质量发展。基于此，本研究旨在系统梳理演化博弈理论在教育研究中的应用现状，归纳该理论的应用优势及其适用场景，并对该理论的未来研究趋势进行展望分析，以期为教育者深入开展教育主体博弈研究提供借鉴与参考。

2. 演化博弈理论在教育研究中的应用现状

演化博弈理论旨在理解参与主体在冲突过程中的策略选择（Adami C et al.，2016），解释说明为何主体将达到平衡稳定状态以及如何达到（易余胤 & 刘汉民，2005），并认为利益

相关主体是处于不完全信息、有限理性的状态。近年来，教育研究者尝试将演化博弈理论应用于教育研究当中，产出较多研究成果。本研究通过文献分析和网络调研，发现目前演化博弈理论在教育研究内的创新应用主要聚焦在学生学习、教育教学、科研创新、学校管理等方面。

2.1 学生学习

探明学生学习行为、对学生的发展提出切实的建议，对改善教学过程、提高教育成效具有关键作用。当前，演化博弈理论应用于学生学习的研究聚焦于知识构建与在线学习。一是分析学生知识构建行为。知识构建是由多个主体共同推进知识共享、协商的过程，促使学生动态发展其自身的知识网络结构与对知识的升华与吸收。学界主要探究学生知识构建、知识共享行为与影响因素间的关系（晋欣泉等，2022）。二是挖掘学生在线学习影响因素。为提高在线学习成效，研究者主要利用演化博弈理论分析学生在线学习的影响因素，发现监管力度、额外收益等对学生具有重要影响，并从管理、制度两方面对学校提出建议（欧阳富明等，2021）。

2.2 教育教学

教育教学作为培养时代新人的核心路径，对育人质量和教育成效提升有着十分重要的作用。目前学界主要聚焦于教师教学、课程建设两方面研究。第一，运用演化博弈理论探究教师教学。教师作为教育教学的核心主体，其教学行为直接关乎教学质量和学生学习成效，相关研究主要围绕教师知识共享（Huo M，2013）、技术使用（Muñ H et al.，2023）等行为进行探究。第二，利用演化博弈理论分析课程建设。课程建设将会重塑建设理念、课程改革和课程评价，推动学校教育教学高质量发展。宋丽娜等（2022）利用演化博弈理论，对教管主体在课程建设中的演化博弈行为进行分析，提出在课程建设过程中应遵循整体性、动态性和目的性等原则。

2.3 科研创新

学校全力打造科研创新平台和提供科研服务，发挥科研人员创新精神及培养科研能力，有助于探索学校科研创新新范式。目前，研究者主要利用演化博弈理论探究影响科研人员的科研行为的条件因素、科研创新管理和科研创新促进人才培养三个方面。在科研人员的科研行为影响因素上。研究者以激励科研人员活力为导向（Zheng Y，2023），利用演化博弈理论构建科研人员博弈模型，探讨影响科研人员的科研行为的因素。在科研创新管理研究上。学者主要利用演化博弈理论揭示科研管理者与相关部门的关系，继而提出管理策略（Liu X et al.，2018）。在科研创新促进人才培养上。事实上，科研创新有助于创新型人才培养，将科研创新成果转化为教学资源和学习资源，以此培养学生的科研思维和创新能力。姜春等（2020）采用演化博弈理论，揭示高校科技成果转化如何促进人才培养的演化规律，发现付出的成本与遭受的损失是影响高校行为策略演变的条件因素。

2.4 学校管理

学校管理包含教学管理、事务管理等，其有效管理有助于教育教学的有序开展与学校系统的良性运转。有关运用演化博弈理论开展学校管理的研究，现已形成了两个主要研究方向。一是探究教学管理。相关研究通过构建博弈模型，分析教育主体行为，明确高校建立教学督导机制的必要性，并提出优化教学督导工作的建议（朱淑珍，2015）。二是分析学校事务管理。传统管理方法是改善校园环境和施加惩罚，但程度不同其结果也不同。因此，教育研究者利用演化博弈理论对事件主体的行为过程进行描述，呈现事件的发展趋势及演化路径，进而针对性地提出干预计划和治理措施。例如，Ning Y（2021）通过构建教师和学生双方演化博弈模型，探究教师惩罚如何影响校园暴力事件。

3. 演化博弈理论在教育研究中的应用优势与适用场景

3.1 应用优势

3.1.1. 融合多种理论优势，构建主体收益模型

演化博弈理论是生物学演化理论与经济学博弈论相结合的成果，强调以生态思维来理解问题的演进历程。教育研究者基于演化博弈理论，以生态思维审视教育主体，构建主体收益模型破解教育问题，最终实现教育生态平衡稳定发展。

3.1.2. 基于主体行为视角，破解复杂性、动态性问题

演化博弈理论作为一种探究主体的研究方法，为教育研究提供了一种全新的研究视角——主体行为视角，即教育研究者从主体行为的视角出发，关注主体博弈行为及最终平衡结果，以科学的数学方程演绎主体行为动态演化过程，从而破解复杂性、动态性教育问题。

3.2 适用场景

3.2.1. 教育政策制定

从宏观层面来看，演化博弈理论以探究主体行为演变趋势为旨归，得出影响主体行为的条件因素，不仅为教育政策制定提供内容支撑，也为政策落地成效评估提供手段。例如，为探究激励政策对政府、学校、企业三方合作的利益分配的影响机理，以及探究激励的方式和程度，可通过构建演化博弈模型，梳理三方之间的利益关系并数值化，得出影响政策制定的条件因素，分别对各主体提出政策性建议，以实现教育政策的制定和评估（俞慧刚，2020）。

3.2.2. 学校制度建设

从中观层面来看，学校日常事务管理和教学工作有序进行需要健全完善的制度保障。演化博弈理论可通过构建模型得出影响因素，从而为制度建设提供参考建议。例如，在科技创新过程中，利用演化博弈理论构建关于科研管理制度的模型，探究其之间的利益关系，分析高校管理者与项目负责人在监管过程中的行为演变，得到高校奖惩机制与监管体系是其行为演变的重要因素，以之为依据提出对策建议（许敏 & 濮郁卉，2019）。

3.2.3. 多方协同育人

从微观层面来看，演化博弈理论可明晰教育主体之间的关系和自身采取不同行为的概率，从而为各主体行为策略提供参考建议。例如，为促进学生全面发展，导师与辅导员双方协同配合，共同肩负着育人的任务，采用演化博弈理论揭示导师与辅导员在学生培养过程中的动态演化策略，提出育人共识、育人文化和育人机制等方面的多方协同育人路径（赵玉鹏，杨连生 & 侯坤超，2022）。

4. 总结与展望

实现教育主体高质量合作是推进创新型人才培养和教育体系良性发展的强大内生动力。随着教育问题的复杂程度和教育主体利益的交织程度不断提高，教育主体行为成为教育问题研究的新趋势。演化博弈理论作为研究主体行为的重要方法，在分析影响主体行为的条件因素和作用机制方面发挥着重要的支撑作用。本研究通过大量的文献调研分析，发现演化博弈理论在教育研究领域已有大量的研究成果，主要应用于学生学习、教育教学、科研创新、学校管理四个方面。同时明晰演化博弈理论的应用优势及适用场景，对于研究者应用演化博弈理论开展研究具有一定指导和借鉴作用。尽管演化博弈理论能给教育研究者揭示教育主体行为带来诸多帮助，但仍存在实证数据支持不足、理论应用维度单一等方面的问题。

在未来，运用演化博弈理论实现教育共同体的构建一定是研究者迫切关注的主题，可在以下三方面做出努力：①丰富研究主体。随着人们对教育领域的关注、参与程度不断提高，社会公众对教育发展也有着十分重要的作用，未来可以更多地去探索社会公众对教育问题的

影响。②扩展研究内容。随着人工智能、大数据、互联网的迭代更新，技术对教育变革产生了前所未有的影响，未来研究可聚焦于生成式人工智能教育应用、教育资源开发、教师专业发展、职业教育高质量发展和教育共同体构建等。③融合多种研究方法。未来研究可以结合多种研究方法，诸如系统动力学、复杂网络、问卷调查等，以实现多维度、多视角看待教育问题，更为科学地揭示教育问题发展过程。相信未来在教育研究领域，演化博弈理论能为研究者提供更多价值和产出不同主题下具有实践意义的研究成果，力推教育生态均衡优质发展。

参考文献

朱淑珍.(2015).基于演化博弈理论的高校教学督导运行分析[J].高等工程教育研究,2015(6):138-141.
闫寒冰 & 陈怡.(2023).数据智慧生成的实践路径：设计、管理与治理[J].现代远距离教育(2):53-60.
许敏,濮郁卉.(2019).“放管服”背景下高校科研监管行为演化博弈分析[J].会计之友(4):142-145.
张文鹏,李经展,赵子建.(2022).县域学校体育内部治理的演化博弈[J].上海体育学院学报,46(7):56-63+88.
欧阳富明,范霞,邓帅等.(2021).演化博弈视角下网课平台“刷课”行为控制研究[J].华中师范大学学报(自然科学版),55(2):329-336.
易余胤 & 刘汉民.(2005)经济研究中的演化博弈理论[J].商业经济与管理(8):8-13.
赵玉鹏,杨连生,侯坤超.(2022).演化博弈视域下研究生导师和辅导员协同育人策略及路径研究[J].研究生教育研究(4):23-28.
姜春,陈红喜.(2020).高校科技成果转化促进创新型人才培养演化博弈研究[J].科技管理研究,40(13):183-188.
祝智庭 & 彭红超.(2020).技术赋能的韧性教育系统:后疫情教育数字化转型的新路向[J].开放教育研究,26(5):40-50.
晋欣泉,姜强,赵蔚.(2022).基于系统动力学的知识建构行为演化博弈模型及仿真分析[J].远程教育杂志,40(1):103-112.
Adami C,Schossau J,Hintze A.(2016).Evolutionary game theory using agent-based methods[J]. Physics of life reviews,19:1-26.
Huo M.(2013).Analysis of knowledge-sharing evolutionary game in university teacher team[J]. International Journal of Higher Education,2(1):60-66.
Liu X,Zhang Z,Qi W.(2018).Analysis of the behaviors about scientific research innovation and management in university based on evolutionary game[C].2018 Chinese Control And Decision Conference (CCDC). IEEE,1766-1771.
Muñ H,Accinelli E,Hernández E.(2023).An evolutionary game theoretical approach to the teaching-learning techniques in the post-pandemic era[J].Open Access Library Journal,10(2):1-22.
Ning Y.(2021).How teachers' punishment impact school violence? An evolutionary game theory model[J].International journal of social science and education research,4(7):273-279.
Santos F P,Santos F C,Paiva A.,et al.(2015).Evolutionary dynamics of group fairness[J].Journal of theoretical biology,(378):96-102.
Zheng Y.(2023).University teachers’ scientific research innovation incentive based on the three-party evolutionary game of the state, the colleges, and scientific researchers[J].Frontiers in Psychology,13:973333.

SPOC+产出导向法在国际商务谈判教学中的应用研究

Research on the Application of SPOC+Output Oriented Approach in the Teaching of International Business Negotiation

蒋庆[1]，李西[2,1]，刘芮玲[3]
[1]成都师范学院 外国语学院
[2]四川外国语大学 英语学院
[3]成都市实验外国语学校

【摘要】 **国际商务谈判是一门实践性较强的跨学科课程，在教育信息化、新文科的大背景下，基于产出导向法，构建了以学生为中心，教师为主导，“课上任务驱动，线上任务促成，课上评价引导”三位一体，即“学用评”结合的 SPOC 混合式教学模式。教学结果表明，该教学模式提高了国际商务谈判教学质量，推动了国际商务谈判课程的教学改革。**

【关键词】 SPOC；产出导向法；国际商务谈判；教学模式

***Abstract*：** *International business negotiation is an interdisciplinary course with strong practicality. Under the background of educational informatization and new liberal arts, based on the output oriented approach, the author has built a teaching mode that the classroom is student-centered and teacher-led and has a teaching process of in-class task drive, online task facilitation and in-class evaluation. Besides, This teaching mode emphasizes a combination of learning and application. The teaching results show that the teaching mode improves the teaching quality of international business negotiation and promotes the teaching reform of international business negotiation.*

Keywords: SPOC, production oriented approach, international business negotiation, teaching model

1. 引言

2013 年，加州大学伯克利分校的阿曼德·福克斯教授指出 MOOCs 教学资源可以来补充传统课堂教学，从而提高学生知识水平的参与度和掌握程度。福克斯教授将这一教学模式称为 SPOC(small private online course)小规模限制性在线课程（Armando Fox，2013）。产出导向法是文秋芳教授提出的具有中国特色的英语教学法，产出导向法主张学用一体说，克服了中国大学生学用分离的现象（文秋芳，2017）。SPOC 混合教学模式和产出导向法在教学理念上是一致的，都强调学生是学习的主体，教师是学习的主导。

《国标》强调以需求为导向，探索多元培养模式。即人才培养模式要考虑国家以及区域经济社会对人才发展的具体需求（仲伟合，2015）。《商英指南》在《国标》的基础上，要求商务英语教学以能力培养为导向、综合运用探究学习、合作式学习及混合式教学等教学方法，推进现代信息技术与专业教学的有效融合（王立非,宋海玲，2021）。然而目前国际商务谈判课堂主要采取传统讲授教学模式，已无法满足当今经济社会发展对人才的需求。而产出导向法和 SPOC 混合教学模式均强调学生自主学习，合作学习以及探究学习，基于产出导向法和 SPOC 混合教学的模式有利于商务英语专业人才的培养。鉴于此，本文旨在结合 SPOC 混合教学以及产出导向法理论，构建国际商务谈判教学的新模式，并开展实证研究证实该教学模式的教学效果。

2. 研究背景

2.1. 国际商务谈判教学

国际商务谈判课程是一门兼具理论性和实用性的跨学科课程，教材方面，现有的大多数国内教材理论比重多，教材中案例的时效性和针对性不强，案例种类较少，很少有融合多种数字媒体的多媒体教材。教学方式方面，多为传统讲授的教学方式。教学过程重理论轻实践，教师成为了教学的中心，学生被动地接受理论知识。作为一门实操性强的课程，传统讲授的教学方式已经无法满足教学及社会需求。评价方式方面，多为期末考试这一终结性评价为主，注重对学生理论知识的考核，缺乏过程性评价，这种单一的评价方式很难调动学生的积极性和自信心。经文献梳理显示，目前国际商务谈判教学研究主要集中在教学发展趋势研究，商务谈判口译模型建构，以及从任务教学法、项目教学法、情景教学法以及翻转课堂教学法的方式来研究国际商务谈判教学。

2.2. SPOC 混合教学模式

SPOC 教学模式自兴起以来就涌现了大量的教学研究，主要集中在：SPOC 翻转课堂教学模式研究，基于 SPOC 翻转课堂的资源构建，基于 SPOC 翻转课堂教学实践，基于 SPOC 的混合式教学设计及实践。SPOC 混合教学模式已经广泛应用于各大学科领域，其中主要集中在计算机软件及计算机应用、外国语言文学、医学教育与医学边缘学科、图书情报与数字图书馆、化学、计算机硬件技术、生物学、畜牧与动物医学、电力工业等学科领域。在研究层次方面，SPOC 混合教学模式除了广泛于高等教育学科教学领域以外，在职业教育学科也有大量的研究，另外，SPOC 混合教学模式在成人教育和特殊教育中得到了一定的研究和应用。

2.3. 产出导向法

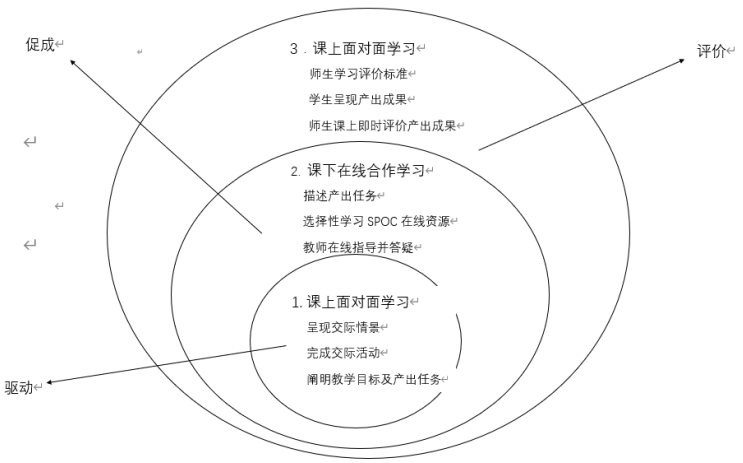
产出导向法是文秋芳教授于 2015 年构建的具有中国特色的英语教学理论。POA 强调教学需要实现教学目标，并促成有效学习的发生。同时，产出导向法主张在课堂教学中，语言教学活动需要与运用紧密相连，做到学用一体（文秋芳，2015）。产出导向法的教学流程包括驱动，促成和评价三个环节，其中，以教师为中介的促成环节是教学的中心环节（文秋芳，2018）。目前产出导向法主要应用于外语、计算机、医学、贸易经济、航空航天、会计、化学、矿业等学科领域。

综上文献所述，目前尚未有文献从信息技术和产出导向法相结合的角度去研究国际商务谈判教学。因此，本文试图将 SPOC 混合教学模式与产出导向法想结合，来更好地指导国际商务谈判教学的开展。

3. 基于产出导向法的国际商务谈判 SPOC 混合式教学模式研究

结合产出导向法和 SPOC 混合教学模式的特点，本文设计出了一套适合国际商务谈判课程的教学模型，这套教学模式也适用于其他技能性课程，如图 1 所示。

图 1 基于产出导向法的国际商务谈判 SPOC 混合式教学模式研究



为了更好地将 SPOC 混合教学模式与产出导向法相结合，在开展教学活动之前，教师在云班课移动端平台上搭建 SPOC 教学资源平台，学生在此平台上可以查看国际商务谈判课程单元教学目标，课件，教学视频以及文档资源，其中教学视频资源和部分文档资源来自 Coursera 平台密歇根大学 MOOC 课程：Successful Negotiation,每个视频内容都包含谈判概念及策略的介绍，案例讲解以及知识点总结。学生还可以在 SPOC 教学平台上完成小组活动及互动答疑等环节。

第一，课上驱动。首先，教师根据章节教学目标在课堂呈现与他们生活相关的谈判头脑风暴任务，例如，根据所给背景材料分析该国际商务谈判失败的原因，该头脑风暴任务与即将学习的谈判知识有关，对学生形成一定的认知挑战。其次，学生尝试就头脑风暴话题展开小组讨论，期间可能会产生疑惑、并以此激发学生的求知欲望。最后，教师在课堂上阐述本章节谈判课教学目标及模拟谈判相关内容。

第二，线上促成。教师根据单元内容发布第一个产出任务，即案例分析的具体要求；首先要求学生以小组为单位分析一则国际商务案例，学生可以从移动端 SPOC 教学资源平台选择案例所需的国际商务谈判理论知识和技能，期间有任何疑问，学生可以通过移动端平台的答疑专区提问，教师就相关问题进行解答。最终学生以小组的形式将案例分析的成果上传至移动端平台，通过案例分析，教师可以检查学生对移动端平台知识的掌握情况，对学生的小组案例分析进行即时评价。教师在线下课堂再对学生案例分析情况进行反馈和点评，并对本章节的知识点进行概括和总结。

第二阶段，教师在移动端 SPOC 平台发布第二个产出任务，即模拟谈判。在模拟谈判开始之前，教师应该为学生搭建脚手架，以帮助学生更好地理解谈判目标并实施具体的谈判任务。对于构建不同的新知识，我们所需搭建的“脚手架”往往不尽相同（杜军，2005）。首先，教师根据所学章节的谈判要点制定目标脚手架。搭建目标脚手架可以使学习内容更为清晰、具体。在国际商务谈判移动端 SPOC 教学平台上，教师发布模拟谈判的教学目标及具体谈判要求，学生在平台上仔细研读教学目标。接着，教师为学生搭建国际商务谈判策略的脚手架，学生根据模拟的谈判的教学目标和要求，选择性学习 SPOC 平台的相关国际商务谈判概念和谈判策略，为进一步内化新知识做准备。最后，教师为学生搭建谈判语言结构的脚手架，学生学习谈判所需语言结构，并观看国际商务谈判视频，再分组对给定的谈判任务进行模拟演练。

第三，课上评价。教师要求学生以小组的形式在课堂进行模拟谈判，并对小组模拟谈判过程进行师生以及生生的即时评价。模拟谈判开始之前，教师就模拟谈判的评分规则进行了详细的阐述。评分规则采取百分制，包括谈判准备、谈判过程、谈判效果三个维度，分别占比 30%，40%，30%。教师及学生将以此评分标准给模拟谈判的小组给分，教师分数占比 60%，学生分数加权占比 40%，教师在学生评分过程中给予一定的指导。最终教师将会对小组的模拟谈判表现进行点评，同时也会邀请其他小组的同学对该组进行点评。

4. 基于产出导向法的国际商务谈判 SPOC 混合式教学模式的教学效果

为了证实基于 SPOC 混合教学模式和产出导向法的国际商务谈判课程教学效果，笔者随机抽取了成都师范学院外国语学院大三翻译专业二班为实验班，一班为控制班，两个班的英语专业四级过级率以及人数均相当。笔者将 SPOC 混合教学模式和产出导向法融入到了实验班的国际商务谈判课程的教学中，而在控制班，我们仍然选择传统的讲授教学方式。学期结束后，笔者用 SPSS26.0 软件对控制班和实验班期末考试的均值和标准差进行了分析， 分析数据结果见下表 1。

表 1 实验班和控制班期末成绩对比表

分值	90 分以上（优秀）	80-89 分良好）	70-79 分（中等）	60-69 分（及格）	60 分以下(不及格）	均值	标准偏差
实验班	0%	67.39%	28.26%	2.17%	2.17%	80.4783	6.44891
控制班	0%	50%	36.96%	13.04%	0%	77.7391	5.66249

成绩统计表 1 数据显示实验班和控制班的期末成绩分布均较为合理。通过对实验班和控制班期末成绩对比分析得知:实验班和控制班的卷面及格率都较高，说明两个班都较好地掌握了国际商务谈判的基本知识及谈判策略和技能。然而实验班的良好率比控制班的略胜一筹，实验班的 SPSS 数据分析结果显示均值也略高于控制班。综上数据结果说明在实施了一学期的 SPOC 和产出导向法相结合的教学模式后，该课程的教学效果有着一定的提升。

5. 结论

基于产出导向法，构建国际商务谈判 SPOC 混合式教学模式 ，推动了信息技术与高等外语教学的深度融合以及国际商务谈判课程建设。移动端 SPOC 课程资源为学生提供了自主学习平台，互动学习专区，以及信息交流平台；结合产出导向法的特点，该教学模式实现了课上任务驱动，线上任务促成，课上评价三位一体教学。总之，该教学模式顺应了教育信息化、新文科的时代背景，实现了“学用评”一体的有机结合。

参考文献

注：数据占比总和不等于 100%，由四舍五入导致。
杜军（2005）。“支架式”教学应重视“脚手架”的搭建。教育理论与实践，第 14 期，51-53。https://kns.cnki.net/kcms2/article/
文秋芳（2015）。 构建“产出导向法”理论体系。外语教学与研究，第 4 期， 547-558+640。https://kns.cnki.net/kcms2/article/
文秋芳（2017）。“产出导向法”的中国特色。现代外语，第 3 期， 348-358+438。https://kns.cnki.net/kcms2/article/
文秋芳（2018）。“产出导向法”与对外汉语教学。世界汉语教学，第 3 期，387-400。https://kns.cnki.net/kcms2/article/
王立非,宋海玲（2021）。新文科指引下的复合型商务英语人才培养理念与路径。外语界，第 5 期，33-40。https://kns.cnki.net/kcms2/article/
仲伟合（2015）。《英语类专业本科教学质量国家标准》指导下的英语类专业创新发展。外语界，第 3 期，2-8。https://kns.cnki.net/kcms2/article/
Armando Fox（2013）.From MOOCs to SPOCs. *Communications of the ACM ,56(12)*，38-40. https://dl.acm.org/doi/10.1145/2535918

国内教育界对生成式人工智能的研究热点与趋势——基于共词矩阵的知识图谱分析

Hotspots and Trends in Research on Generative Artificial Intelligence in the Domestic Education Sector——Knowledge Graph Analysis Based on Co word Matrix

任晨扬，陈婉琳，胡来林
温州大学
任晨扬@15868500113.China

【摘要】 ChatGPT 作为 AIGC 的代表产品之一，于 2022 年 11 月上市后引起了全球高度关注，在众多领域得到了广泛的应用。为探索这一年来我国教育界对此类基于大语言模型的人工智能研究的热点与趋势，本研究采用共词矩阵、相似矩阵、多维尺度分析和可视化知识图谱的方法，对 CNKI 文库中 2022 年 11 月-2023 年 11 月的 443 篇文献进行了分析，得出了相关结论，愿为教育工作者提供参考。

【关键词】 生成式人工智能；教育；共词矩阵；知识图谱

Abstract: *ChatGPT, as one of the representative products of AIGC, has attracted high global attention since its launch in November 2022 and has been widely applied in many fields. In order to explore the hot topics and trends in the research of artificial intelligence based on large language models in China's education sector over the past year, this study used methods such as co word matrix, similarity matrix, multidimensional scale analysis, and visual knowledge graph to analyze 443 articles in the CNKI library from November 2022 to November 2023, and drew relevant conclusions. We are willing to provide reference for educators.*

Keywords: AIGC, Education, Co word matrix, Knowledge graph

1. 引言

ChatGPT 是一种基于大语言模型的生成式人工智能聊天机器人，通过大量文本语料数据进行学习和训练，可以理解人类的语言和语义，与用户进行有效的对话交互。（Open AI, 2022）自 2022 年 11 月发布以来，五天时间注册用户超百万，两个月过亿，创造了消费类应用领域的新纪录。（Wodecki, 2023）ChatGPT 是继深度学习后的又一个技术革命的里程碑，它为以自然语言处理为核心的认知技术发展提供了新的历史机遇，有望与其他生成式人工智能一起，彻底改变传统的内容生产和信息获取模式，对包括教育在内的各行业的改造产生深远影响。（焦建利，2023)

学术界对其态度不一，部分观点认为生成式人工智能技术将为教育行业注入强大生机，但在相关法律法规尚未完善，ChatGPT 在教育领域的具体应用尚不明朗的情况下，此类人工智能产品强大的能力也引起了教育工作者的普遍担忧，如学术诚信问题、学生学习能力退化、数据安全与伦理问题（邱燕楠、李政涛，2023）等。

本文以中国学术期刊网络出版总库中收录的期刊文献为分析材料，利用 Bicomb、SPSS 等工具分析国内教育界对生成式人工智能的研究，总结出该课题的研究热点与趋势以及值得关注的问题，为教育工作者提供一定的参考。

2. 数据来源与研究方法

2.1. 文献选取

本文以 CNKI 刊载的期刊文献为统计源，在 CNKI 使用高级检索，检索条件为篇名包含“ChatGPT”“生成式人工智能”“生成式 AI”“大语言模型”中任一关键词，在结果中进一步设置学科为“中等教育”“职业教育”“医学教育与医学边缘”“教育理论与教育管理”“高等教育”“初等教育”“成人教育与特殊教育”这七个教育相关学科，最终得到文献 443 篇。

2.2. 研究工具

本研究采用的工具包括 Microsoft Excel、书目共现分析系统（Bicomb 2021）、SPSS 27。其中 Excel 用于文献基本信息统计（如发表时间），Bicomb 用于关键词共词分析，得出共词矩阵，供进一步聚类分析使用。SPSS 用于聚类分析、多维尺度分析。

2.3. 研究过程

（1）使用 CNKI 检索相关文献，确定研究材料。（2）使用 Excel 统计分析文献发表时间及来源。（3）用 Bicomb 提取词频≥4 的 28 个高频关键词。（4）用 Bicomb 对 28 个关键词进行共词分析，建立词篇矩阵。（5）将词篇矩阵导入 SPSS 进行聚类分析，获得聚类树图。（6）对词篇矩阵进行多维尺度分析。（7）绘制生成式人工智能在教育领域的研究热点知识图谱，并诠释和分析。

3. 研究结果与分析

3.1. 文献的时间分布

本文以 443 篇文献作为研究对象，全部文献为 2023 年期间发布，截止至 11 月 30 日。按照发表的月份进行统计如表 1 所示。从中看出，国内教育界对于生成式人工智能的研究不断深入，文献数量仍在增长中。

表 1 发表月份情况

发表月份	篇数	占比（%）	累计百分比（%）
1	3	0.70	0.70
2	22	5.15	5.85
3	51	11.94	17.80
4	44	10.30	28.10
5	61	14.29	42.39
6	58	13.58	55.97
7	40	9.37	65.34
8	33	7.73	73.07
9	44	10.30	83.37
10	39	9.13	92.51
11	32	7.49	100.00

3.2. 文献的来源分布

根据 Bicomb 统计，443 篇文献分别发表在 254 种不同的期刊上，其中载文量≥4 的期刊有 23 种，教育技术学界的主要期刊基本包含在内，情况如表 2 所示。可以看出各界杂志普遍关注生成式人工智能在教育领域的应用。其中教育技术学核心期刊载文量占总体的 12.64%，如表 3 所示。

表 2 期刊载文量分布统计（≥4）

序号	刊名	载文量	占比（%）	序号	刊名	载文量	占比（%）
1	华东师范大学学报(教育科学版)	15	3.39	13	中国现代教育装备	5	1.13
2	上海教育	13	2.93	14	现代教育技术	5	1.13
3	开放教育研究	9	2.03	15	云南师范大学学报(对外汉语教学与研究版)	4	0.90
4	中小学信息技术教育	9	2.03	16	中国远程教育	4	0.90
5	中国教育信息化	8	1.81	17	中小学数字化教学	4	0.90
6	中国电化教育	8	1.81	18	中小学德育	4	0.90
7	教育发展研究	7	1.58	19	云南教育(视界综合版)	4	0.90
8	继续教育研究	7	1.58	20	新疆师范大学学报(哲学社会科学版)	4	0.90
9	电化教育研究	6	1.35	21	中国信息技术教育	4	0.90
10	教育视界	6	1.35	22	重庆理工大学学报(社会科学)	4	0.90
11	当代教育家	5	1.13	23	汉字文化	4	0.90
12	中国医学教育技术	5	1.13				

表 3 教育技术学核心期刊载文量统计

序号	刊名	载文量	占比（%）
1	开放教育研究	9	2.03
2	中国电化教育	8	1.81
3	电化教育研究	6	1.35
4	现代教育技术	5	1.13
5	中国远程教育	4	0.90
6	远程教育杂志	3	0.68
7	现代远距离教育	3	0.68
8	现代远程教育研究	3	0.68
	总计	56	12.64

3.3. 关键词统计分析

用 Bicomb 对 443 篇文章中的关键词进行词频统计，得到 637 个关键词，手动优化合并含义相似的关键词后，抽取其中词频大于 4 的 22 个高频关键词，结果见表 4。

表 4 高频关键词统计（≥4）

序号	关键字段	出现频次	占比（%）	序号	关键字段	出现频次	占比（%）
1	生成式人工智能	474	35.24	12	国际中文教育	6	0.45
2	高等教育	30	2.23	13	教育技术	6	0.45
3	教育数字化转型	27	2.01	14	教学	6	0.45
4	风险治理	25	1.86	15	教育生态	5	0.37
5	教育	18	1.34	16	教师角色	5	0.37
6	教育改革	12	0.89	17	教师专业发展	5	0.37
7	思想政治教育	11	0.82	18	医学教育	5	0.37
8	应对策略	11	0.82	19	基础教育	4	0.30
9	教育应用	9	0.67	20	个性化学习	4	0.30
10	人机协同	8	0.59	21	数字素养	4	0.30

11	人才培养	7	0.52	22	中国式教育现代化	4	0.30

关键词作为文章的标签，很大程度上代表了文章核心内容，因此高频关键词排序可以初步反映我国教育界对生成式人工智能研究的集中热点和趋势。从表 3 中可知，除“生成式人工智能”以外的前 10 个热点分别为：“高等教育（30）”“教育数字化转型（27）”“风险治理（25）”“教育（18）”“教育改革（12）”“思想政治教育（11）”“应对策略（11）”“教育应用（9）”“人机协同（8）”“人才培养（7）”。其中，除“生成式人工智能”外的首位高频关键词是“高等教育”，说明不少国内学者认为将高校作为生成式人工智能的应用场景更为合适；其次是“教育数字化转型”“教育改革”，说明不少学者响应国家号召，积极探索教育数字化的新道路；另外，“伦理风险”“应对策略”位居高位，可以看出不少学者就人工智能的影响表达了担忧；在涉及到的学科方面，以思想政治教育为代表。为进一步发现这些高频关键词之间的联系，本文通过词篇共现进行深入挖掘。

3.4. 高频关键词的 Ochiia 系数相似矩阵及分析

高频关键词 Ochiai 系数相似矩阵中的数字大小代表两个关键词的相似性，也就是相应两个关键词之间的距离远近，数值越大，相似度越高。利用 Bicomb 将上述 22 个高频关键词进行共词分析，生成词篇矩阵，将其导入 SPSS，选取 Ochiai 系数并其转化为一个 22×22 的共词相似矩阵，部分结果如表 5 所示。

表 5 高频关键词 Ochiai 系数相似矩阵（部分）

	生成式人工智能	高等教育	教育数字化转型	风险治理	教育	教育改革	思想政治教育	应对策略	教育应用
生成式人工智能	1.000	0.291	0.302	0.290	0.246	0.201	0.192	0.175	0.155
高等教育	0.291	1.000	0.071	0.074	0.000	0.000	0.168	0.280	0.124
教育数字化转型	0.302	0.071	1.000	0.038	0.136	0.056	0.058	0.000	0.000
风险治理	0.290	0.074	0.038	1.000	0.047	0.000	0.181	0.060	0.067
教育	0.246	0.000	0.136	0.047	1.000	0.000	0.000	0.071	0.000
教育改革	0.201	0.000	0.056	0.000	0.000	1.000	0.000	0.000	0.000
思想政治教育	0.192	0.168	0.058	0.181	0.000	0.000	1.000	0.182	0.000
应对策略	0.175	0.280	0.000	0.060	0.071	0.000	0.182	1.000	0.000
教育应用	0.155	0.124	0.000	0.067	0.000	0.000	0.000	0.000	1.000

从表 5 中得出，各关键词与生成式人工智能的距离由近及远分别为：教育数字化转型（0.302）、高等教育（0.291）、风险治理（0.290）、教育（0.246）、教育改革（0.201）、思想政治教育（0.192）、应对策略（0.175）、教育应用（0.155）。从这个结果中看出，这一年来我国教育界对生成式人工智能的研究较多以高等教育为背景，积极探索人工智能支持下的教育数字化转型策略。伴随着生成式人工智能的高收益而来的是较高的风险，其在伦理、法律、纪律、个人道德等方面存在的问题引起了众多学者的讨论。虽然前景尚不明朗，但可以肯定的是，如何适应日益进化的人工智能，使其融入进教育改革体系，将成为一大课题。

3.5. 高频关键词聚类图及分析

关键词聚类分析能反映关键词之间的亲疏，进一步归纳国内教育界对生成式人工智能的研究热点。原理是以关键词成对在同一篇文章中出现的频率（共词）为分析对象，使用聚类的统计学方法，把关联密切的关键词聚集在一起形成类团。关键词越相似，反映在聚类图中的表现为距离越近，反之越远。用 SPSS 27 对关键词相似系数矩阵进行系统逐次聚类分析，得到的聚类结果如图 1 和表 6 所示。

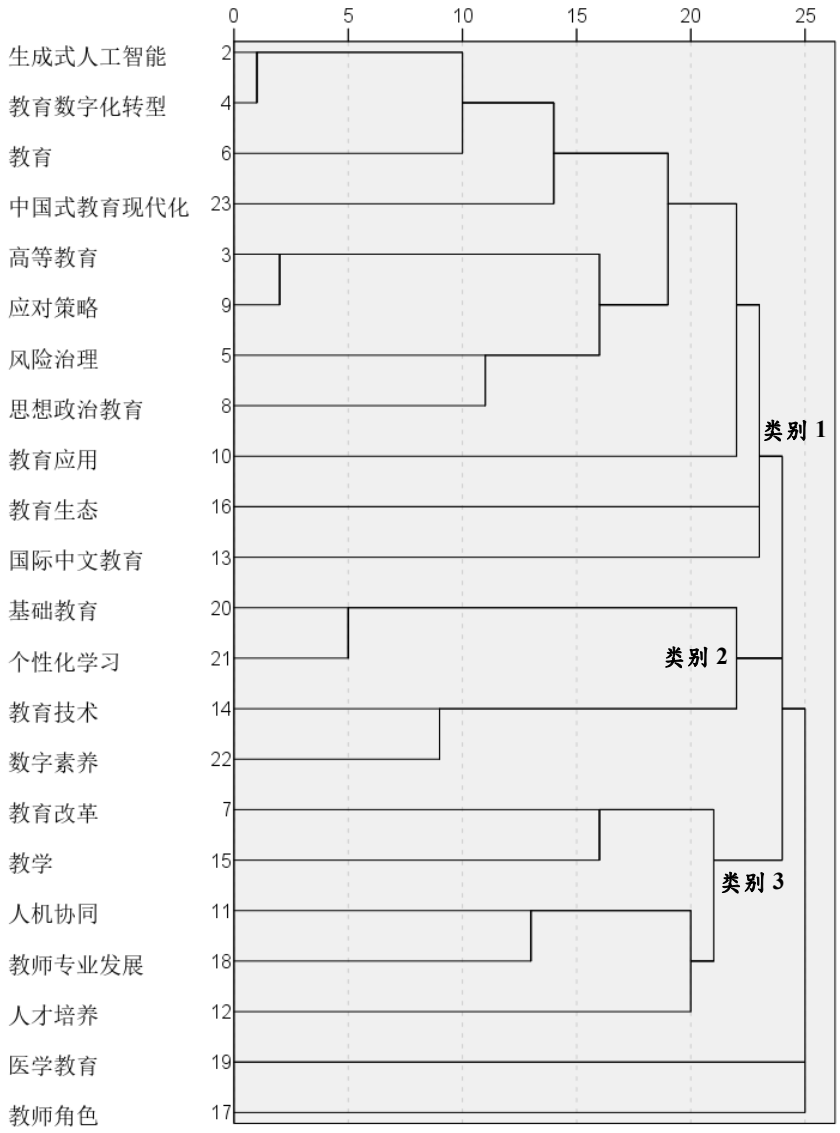


图 1 高频关键词聚类图

表 6 高频关键词聚类结果

分类	关键词
类别 1	生成式人工智能、教育数字化转型、教育、中国式教育现代化、高等教育、应对策略、风险治理、思想政治教育、教育应用、教育生态、国际中文教育
类别 2	基础教育、个性化学习、教育技术、数字素养
类别 3	教育改革、教学、人机协同、教师专业发展、人才培养

从图 1 可以直观看出国内教育界对生成式人工智能研究的高频关键词可以分为三大类。

类别 1 是这一年来国内教育界对生成式人工智能研究的最大热点，包括：生成式人工智能、教育数字化转型、教育、中国式教育现代化、高等教育、应对策略、风险治理、思想政治教育、教育应用、教育生态、国际中文教育，共 11 个关键词。从研究对象的层次出发，可以将该类研究概括为生成式人工智能对教育生态的影响研究。该类研究从宏观制度层面分析生成式人工智能对我国教育体系产生的影响，普遍认为生成式人工智能有望改变智能教学系统的开发模式，提升教学效果，完善评价体系，促进教育科学化、高效化、民主化、法制化、智能化（周洪宇,李宇阳，2023），但对其负面影响也存在大量讨论，各地高校对于此类人工智能产品的态度不一，是否决定引入也尚未形成共识。

类别 2 可以概括为生成式人工智能对学习模式的影响研究，包含关键词：基础教育、个性化学习、教育技术、数字素养，共 4 个关键词。该类研究聚焦于教学实践，认为以 ChatGPT 为代表的人工智能及其衍生应用可以为个性化学习提供更全面有效的支持，自主学习的方式将从简单搜索转向具有古老传统的对话式学习（焦建利，2023），学习者的观念将从“被动学习”蜕变为“自我导向”。然而，不少学者认为人工智能产品的使用对学习者自身提出了一定的要求，因此在基础教育层面，学校教学系统的要素（如教师）可能会大变，但系统结构会保持稳定（张志祯、张玲玲、米天伊，2023）。另外，关于此类人工智能产品可能导致的诸如学术诚信、依赖心理、学习能力退化、数据安全与伦理等问题，由于相关法律法规与具体的人工智能教育产品尚不明晰，研究者多从个人角度提出应对策略，如因材施教、人机协同、坚持督导、教学相长、批判性看待人工智能等。（陈增照、石雅文、王梦珂，2023）

类别 3 可以概括为生成式人工智能对教学改革与教师发展的影响研究，包含关键词：教育改革、教学、人机协同、教师专业发展、人才培养，共 5 个关键词。该类研究普遍认为人工智能技术的发展将为教师专业发展提供新的可能，如：革新包括教育观、教学观、学习观在内的教师专业理念；增强教师专业知识与能力等。相应的，教师也面临着新的挑战，如：对“双师”融合理念的冲击；专业知识从碎片化到系统化知识构建；对数字素养和高阶能力的挑战；职业认同的弱化等。（吴军其、吴飞燕、文思娇，2023）关于具体的教学结构，存在两种主流观点：一种持积极态度，认为在人工智能的冲击下，教学模式会发生变革，从“师-生”二元结构转变为“师-生-机”三元结构（杨宗凯、王俊、吴砥，2023）；另一种持保守态度，认为学校的教学系统很可能不会产生结构性变革，现有的班级授课、分科课程、教师主导等制度仍保持稳定；基于大语言模型的教学自动化还需补足“上层结构”，夯实“基础动作”。

3.6. 研究热点知识图谱及分析

为进一步探寻关键词之间的内涵，用 SPSS 对 22 个关键词构成的相似矩阵进行多维尺度分析，标准化方法采用 Z 分数，绘制出国内教育界对生成式人工智能的研究热点知识图谱如图 2 所示。

多维尺度绘制出的坐标被称为战略坐标，以向心度和密度为参数绘制成二维坐标，可以概括表现一个领域或亚领域的结构。（Law J & Bauin S & Courtial J Pet al, 1988）战略坐标中，各个小圆圈代表每个高频关键词所处的位置，圆圈间距离越近，表明它们之间的关系越紧密，反之越疏远。影响力越大的关键词所表示的圆圈距离战略坐标的中心点越近，反之越远。

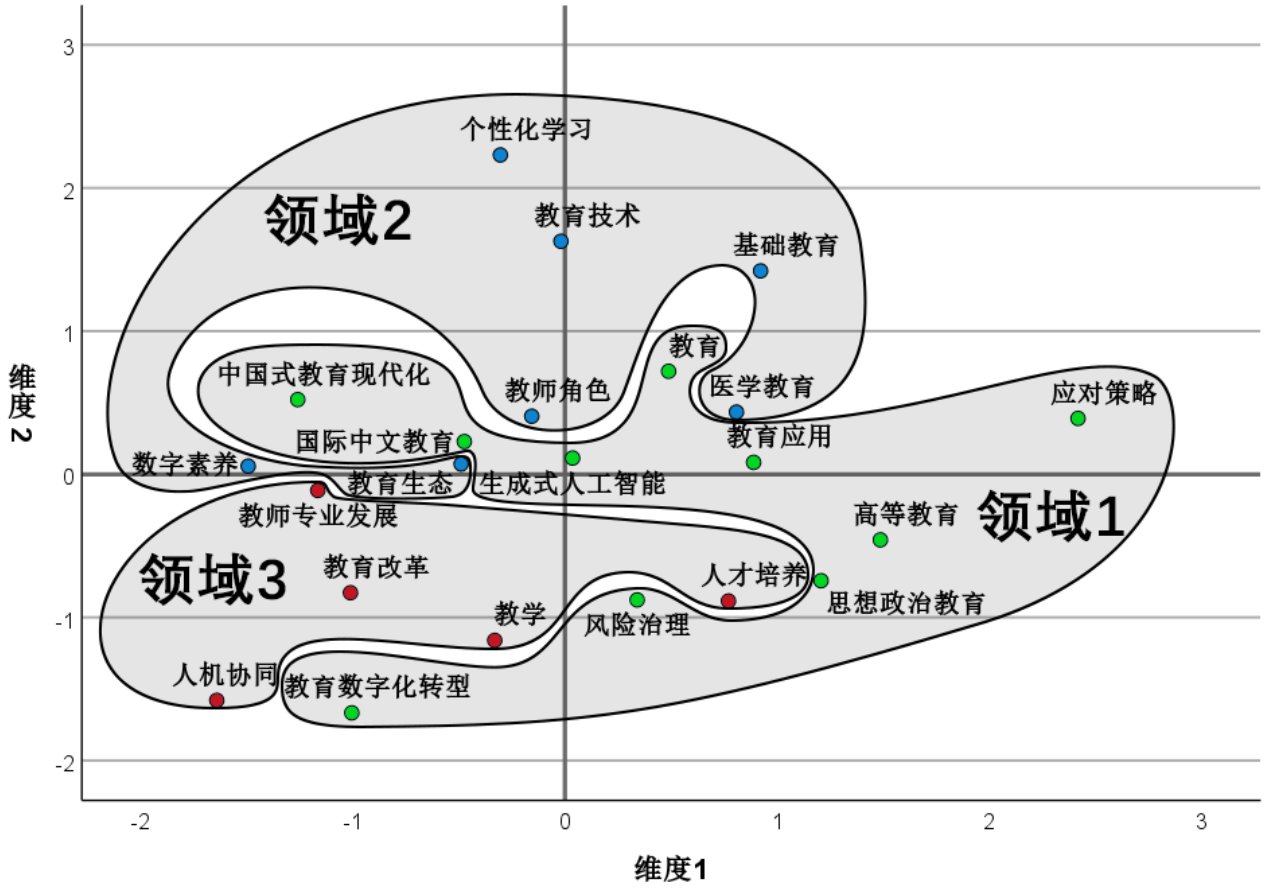


图 2 研究热点知识图谱

根据战略坐标解读规则，从图 2 中可以看出，领域 1（类别 1）为生成式人工智能对教育生态的影响研究，是唯一横跨战略坐标横纵轴，也是唯一涉及四个象限的领域，主要分布在第一象限和第四象限。落在第一象限的教育、教育应用、应对策略是目前相关研究的重点所在，研究主题之间联系紧密，相关研究结果较为成熟。 位于第四象限的高等教育、思想政治教育、风险治理互相联系较为紧密，但是相关研究成熟度较低，有较大的发展空间。

领域 2（类别 2）为生成式人工智能对学习模式的影响研究，分布在第一和第二象限，自身的内部联系相对松散，研究主题较为独立，在未来的研究中具有一定的价值，有进一步发展空间。

领域 3（类别 3）为生成式人工智能对教学改革与教师发展的影响研究，主要集中于第三象限，自身内部联系紧密程度以及研究成熟度都处于中等水平，有待进一步研究。

4. 结论

通过定性和定量方法分析这一年来我国教育界对生成式人工智能的研究，可以发现相关研究主要围绕生成式人工智能对教育生态的影响、生成式人工智能对学习模式的影响、生成式人工智能对教学改革与教师发展的影响这三个层面展开，但存在如下不足：第一，总体研究成果较少，且主题间不够集中和系统，个性化学习、应对策略等处于研究的边缘位置，有待进一步统一和归纳。第二，生成式人工智能对教育生态的影响研究是当前研究的主要领域，属于宏观层面，而下行到具体的实际应用则较少，虽然有现实因素（如伦理等）的限制，但应引起研究者的重视。第三，对于生成式人工智能的研究较少结合具体学科，22 个高频关键词中，仅思想政治教育及医学教育有所涉及。第四，相关研究中生成式人工智能的应用场景以高校居多，学前以及中小学较少。第五，对于生成式人工智能的负面影响，虽然各界人士提出了大量应对措施，但效果有限。

参考文献

陈增照,石雅文,王梦珂.人工智能助推教育变革的现实图景——教师对 ChatGPT 的应对策略分析[J].广西师范大学学报(哲学社会科学版),2023,59(02):75-85.DOI:10.16088/j.issn.1001-6597.2023.02.006.

邱燕楠,李政涛.挑战·融合·变革：“ChatGPT 与未来教育”会议综述[J].现代远程教育研究,2023,35(03):3-12+21.

吴军其,吴飞燕,文思娇等.ChatGPT 赋能教师专业发展：机遇、挑战和路径[J].中国电化教育,2023(05):15-23+33.

杨宗凯,王俊,吴砥等.ChatGPT/生成式人工智能对教育的影响探析及应对策略[J].华东师范大学学报(教育科学版),2023,41(07):26-35.DOI:10.16382/j.cnki.1000-5560.2023.07.003.

张志祯,张玲玲,米天伊等.大型语言模型会催生学校结构性变革吗？——基于 ChatGPT 的前瞻性分析[J].中国远程教育,2023,43(04):32-41.DOI:10.13541/j.cnki.chinade.20230301.002.

周洪宇,李宇阳.生成式人工智能技术 ChatGPT 与教育治理现代化——兼论数字化时代的教育治理转型[J].华东师范大学学报(教育科学版),2023,41(07):36-46.DOI:10.16382/j.cnki.1000-5560.2023.07.004.

焦建利.ChatGPT：学校教育的朋友还是敌人？[J].现代教育技术,2023,33(04):5-15.

焦建利.ChatGPT 助推学校教育数字化转型——人工智能时代学什么与怎么教[J].中国远程教育,2023,43(04):16-23.DOI:10.13541/j.cnki.chinade.20230226.001.

Law J, Bauin S, Courtial J Pet al. Policy and the Mapping of Scientific Change: A Co -Word Analysis of Research into Environmental Acidification [J].Scientometrics,1988,14(3)：251~264.

Open AI.Introducing ChatGPT[OL].<https://openai.com/blog/chatgpt>

Wodecki,B.(2023,Feb 4).UBS:ChatGPT May Be the Fastest Growing App of All Time.AI Business.https://aibusiness.com/nlp/ubs-chatgpt-is-the-fastest-growing-app-of-all-time

ChatGPT 在教育中的研究热点分析

Analysis of ChatGPT's Research Hotspots in Education

陈婉琳，任晨扬，胡来林
温州大学
陈婉琳@13362035047.China

【摘要】 为全面了解一年来 ChatGPT 在教育领域的研究现状与发展趋势，采用内容分析法，对发表在教育技术领域核心期刊的有关文章进行分析，从时间分布模型、期刊分布模型、研究者群体分布模型和主题分布模型等四个维度进行统计分析；并对“ChatGPT 的教育本质分析”、“ChatGPT 在教育中的机遇与应用探索”和“ChatGPT 在教育中的挑战与应对策略”这三个研究主题进行了定性分析，总结 ChatGPT 在教育中的机遇与应用探索，找出它在教育中的挑战并提出应对策略。

【关键词】 ChatGPT；教育；研究热点

Abstract： *In order to comprehensively understand the research status and development trend of ChatGPT in the field of education in the past year, relevant articles published in core journals in the field of education technology were analyzed by using content analysis method, and statistical analysis was carried out from four dimensions: time distribution model, journal distribution model, researcher group distribution model and topic distribution model. The three research topics of "Theoretical analysis of ChatGPT", "Opportunities and application exploration of ChatGPT in education" and "Challenges and coping strategies of ChatGPT in education" are qualitatively analyzed，the opportunities and application exploration of ChatGPT in education are summarized, and the challenges in education are found out and the coping strategies are put forward.*

Keywords： ChatGPT，education，research hotspots

1. 引言

ChatGPT 于 2022 年 11 月 30 日正式上线，迅速成为了人工智能教育应用的新焦点。ChatGPT 全名为 Chat Generative Pre-trained Transformer（生成式预训练聊天转换器），由美国公司 OpenAI 开发，它具有强大的信息检索能力、文本生成能力、自然语言能力等。ChatGPT 以其卓越的自然语言处理能力，引起了教育界的关注。它能通过学习和理解人类语言跟人进行对话，还能根据聊天的上下文跟人互动，并且协助人类完成一系列活动。它的出现给教师行业带来巨大的影响，甚至部分教师担忧自己被人工智能替代，因此害怕人工智能技术在教育教学中的应用。ChatGPT 无疑已经引发人们对人工智能技术在教育应用的价值审视，它可以为教育教学提供哪些机遇？它又会给教育带来哪些挑战和风险？人类到底应该如何将 ChatGPT 适切地运用到教育教学中？从近一年来的文献中发现这些问题的答案，可以为教育工作者和研究者进一步在教育领域审慎地使用生成式人工智能新技术，具有重要的参考价值。

2. 研究方法

2.1. 研究方法

本文采用内容分析法对一年来 ChatGPT 在教育领域的研究的文献进行定性和定量分析，感知研究者们关注的主题。旨在对 ChatGPT 在教育领域的研究热点分析，揭示当前研究的趋势和重点，了解其发展动态，发现研究的空白和不足之处，为后续的研究提供新思路和方向，促进教育领域的学术进步和发展。

2.2. 材料

本研究文献来自中国知网，用篇名“ChatGPT” or “生成式人工智能” or “生成式 ai” or “大语言模型”进行高级检索后，在学科领域选择了“教育理论与教育管理”“初等教育”“中等教育”“高等教育”“职业教育”“医学教育与医学边缘学科”和“成人教育与特殊教育”，选定载文量较大的 8 种教育技术研究领域中的核心期刊，分别是《电化教育研究》《中国电化教育》《现代远距离教育》《现代远程教育研究》《中国远程教育》《远程教育杂志》《开放教育研究》和《现代教育技术》。

2.3. 类目划分与分析单元

将 ChatGPT 在教育领域的研究热点的类目划分为以下三类主题：“ChatGPT 的教育本质分析”、“ChatGPT 在教育中的机遇与应用探索”和“ChatGPT 在教育中的挑战与应对策略”。这三种类型的划分是通过初步阅读这 8 种核心期刊中相关文献的内容归纳而出，在一定意义上是为了研究的方便，包括了个人的主观因素。本研究中，分析单元是各杂志的单篇文章。若是一篇文章涉及多个主题，则按照文章的主要内容归入上述三类中的某一类。

2.4. 日期抽样

这 8 种期刊的时间抽样为 2022 年 11 月 30 日-2024 年 04 月 09 日，共选取的文章为 63 篇。时间采样的起点为 2022 年，因为 ChatGPT 在 2022 年 11 月 30 日正式上线，选取 ChatGPT 出现后的文献进行研究。

3. ChatGPT 在教育领域的研究热点分析

3.1. 研究的基本状况

3.1.1. 新兴的研究领域，文章数量增长明显

生成式人工智能 ChatGPT 在 2022 年年底正式上线，在教育领域中的应用研究属于新的领域。文章发表的数量虽然不是很多，但其研究成果以一种极快的速度增长。截至 2024 年 04 月 09 日，发表的 ChatGPT 教育研究的学术文献已达到 455 篇，显示出人们对这个新的技术在教育领域中作为的期盼，相信未来会有更多的学者投入到该领域中进行研究，会更加快速地出现新的研究成果。

3.1.2. 不同期刊所载文献数量分布大致均衡

一年来，8 种教育技术核心期刊对于 ChatGPT 在教育中应用研究的主题都十分关注，载文量大致均衡，都超过 4 篇。其中发表最多的是《开放教育研究》，共发表 13 篇文献；其次是《电化教育研究》，发表 11 篇文献；《现代教育技术》和《中国电化教育》分别发表 10 篇。其他杂志发表的研究文章的篇数相差不大，具体发文数量如图 1 所示。这是一种新兴研究领域，每种杂志发表的文章篇数并不多，相信未来发文数量会显著上升。

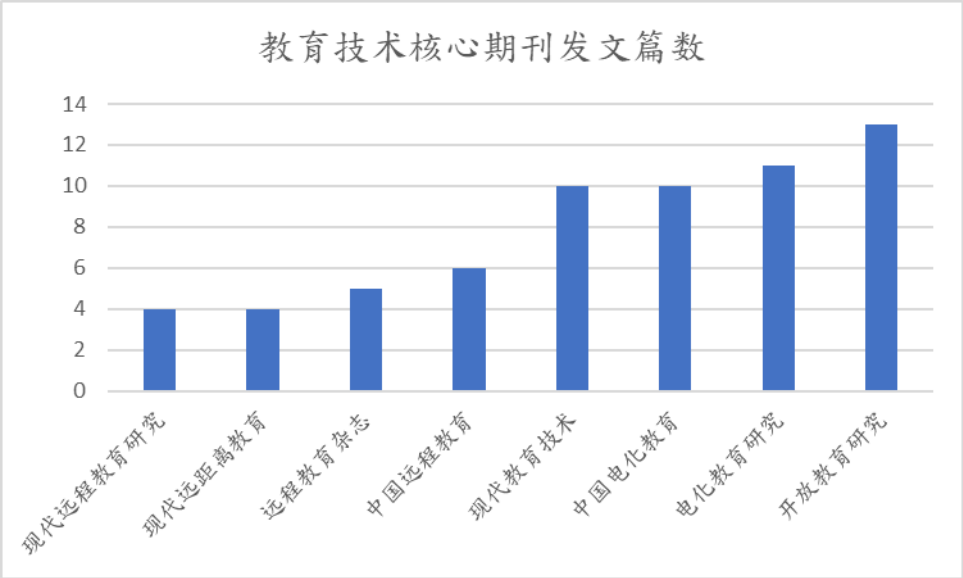


图 1 8 种教育技术核心期刊发文篇数

3.1.3. 研究者群体较为集中

从专业背景上看，主要集中于有教育专业背景的教师或学生，高达 59 篇研究文章；从研究者单位性质来看，都是高校和研究所人员的群体，具体如图 2 所示。从以上数据看出缺少中小学教师研究群体，说明生成式人工智能在中小学教育情境还未得到广泛的应用，目前大多数的研究都处于理论和应用探索阶段，这是一个十分值得关注的问题。

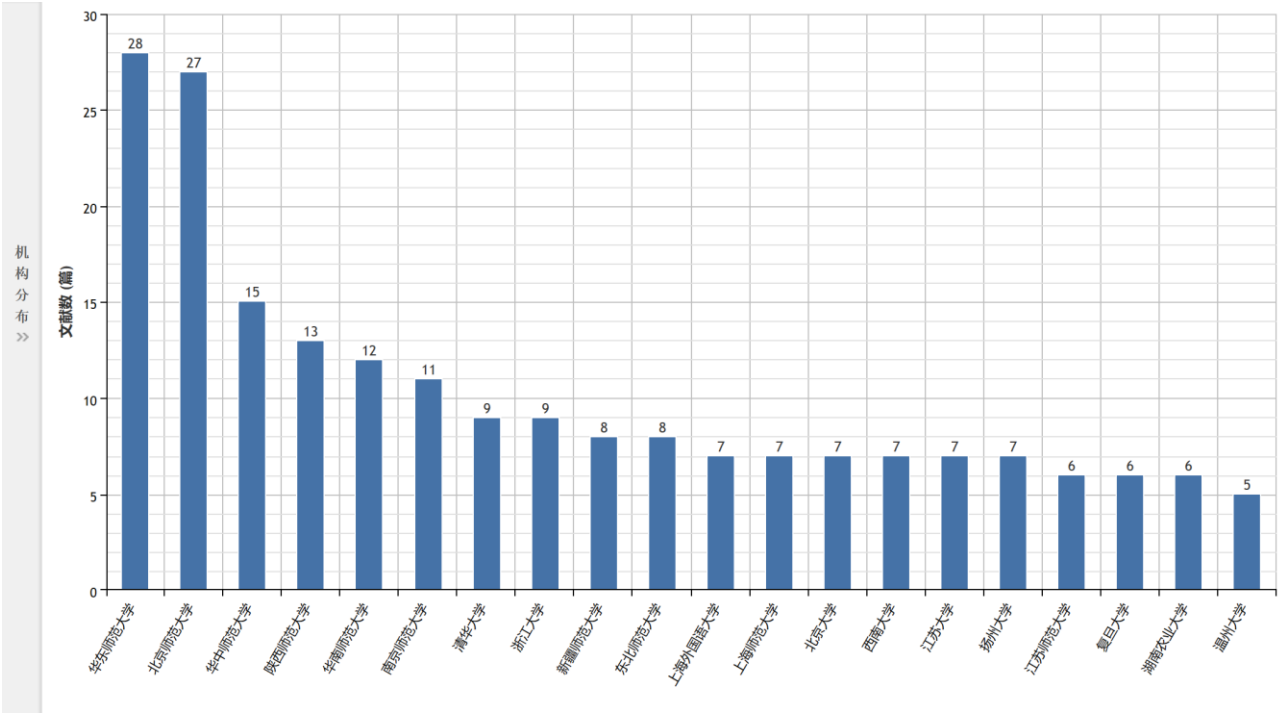


图 2 研究机构分布图

3.1.4. 重点研究主题突出

从文献分析结果来看，相关研究论文的主题主要聚焦于“ChatGPT 的教育本质分析”、“ChatGPT 在教育中的机遇与应用探索”和“ChatGPT 在教育中的挑战与应对策略”三个突出的主题。

3.2. 研究的主题分析

3.2.1. ChatGPT 的教育本质分析

根据联合国教科文组织在 2023 年 9 月发布的《生成式人工智能教育与研究应用指南》对

“生成式人工智能”进行了定义：生成式人工智能是根据人类借助思维符号表征系统表达的提示自动生成内容的人工智能技术（苗逢春，2023）。这一新技术产品的出现在教育领域引起广泛争议，人们对技术的批判源于对它的不信任，一方面认为自己需要技术的帮助，另一方面又担心技术的强大使自己附庸于它（郭顺峰、李光、邹红军，2023）。

ChatGPT 的出现让教育理论工作者重新思考了关于知识观、学习观、教学观和人才观等多方面的问题。知识不再局限于书本之中，知识的生产也不再是专家学者的专利，智能机器人已加入知识生产的主体，ChatGPT 虽然能高效地将数据变成知识，但也会一本正经的胡说八道；随着生成式人工智能的广泛应用，将来“人-机”的合作式学习会越来越普遍，培养解决问题的能力和高阶思维能力成为学习的首要目标，软知识、原理性知识、与信息技术和人工智能关系密切的知识的重要性正在上升；ChatGPT 能让传统的接受式教学转变成发现式教学和分享式教学，硬知识的学习是有必要的，但是也要建构软知识；未来的人才需要与智能机器共同协作，人脑负责创新，机脑承担常规性工作（王竹立、吴彦茹、王云，2023）。ChatGPT 让教育变化的是形式，而非教育的本质，教育的根本在于立德树人，这是 ChatGPT 无法改变的事实。坚持育人为本，培养完整的个体，培养他们的创新思维和意识，增进现实生活中人与人之间的情感交流，注重学术诚信和技术伦理教育，加强劳动教育和闲暇教育，从而促进人的全面发展（冯建军，2023）。

关于该主题的研究文献数量略显薄弱，涉及的文章篇数较少。因此未来需要更多的研究去探索 ChatGPT，提供生成式人工智能在教育教学应用的理论依据。

3.2.2. ChatGPT 在教育中的机遇与应用探索

3.2.2.1. 机遇

一是促进教学的变革，ChatGPT 的出现不仅推动各个教学环节的改变，还创新了教学模式。ChatGPT 会让未来的大学教育在教学目标、教学方法、教学过程和教学评价上进行改变，实现个性化教学，从班级走向模拟课堂、虚拟实验室等，对教学过程实施动态化管理，为学生和教师提供更精准的评价（王正青、阿衣布恩·别尔力克，2023）。凭借深层次语言理解、文本生成能力，ChatGPT 充当教学助手，帮助教师进行智能化教学设计。杨宗凯教授认为 ChatGPT 可以促进教学模式、现行考核体系和现有智能教育应用生态的转变，从传统的“师-生”到“师-机-生”教学模式、从传统的知识测评转向“知识+素养”测评、从孤立专用的教育服务到开放共享的教育平台（张绒，2023）。生成式人工智能还能对道德教育进行辅助，它能够提升高等教育论题汇聚与演化效率；作为一个“中介”，能提高课程的吸引力，弥补师生之间的距离感；不仅能够提升教学内容的吸引力，还能缓解教师压力，有时间针对性地给学生开展德育（闫利利、侯虹印，2023）。生成式人工智能的出现还逼进教育的转型，它可以促进人机协同、知识突破、向下兼容和向“智”迭代（杨欣，2023）。

二是成为学术科研助手，人工智能生成内容对学术生产、学术评价和学术传播上都具有积极的影响。生成式人工智能能够凭借它在快速搜集、检索信息等方面的优势来进行文献梳理和收集工作；ChatGPT 能减少学术中的重复劳动和无意义劳动，一定程度上降低了成本，它还能成为期刊运作的助推剂；生成式人工智能可以极大地提升信息传播的速度，并且实施精准推送（骆飞、马雨璇，2023）。ChatGPT 在能生成学习和教学材料、智能回答问题等功能以外，它还能使常规业务自动化，学生学习和教师教学的相关业务自动化，最后还可以协助开展科学研究，科学家和研究人员尝试使用 ChatGPT 辅助科学研究，确定研究空白、进行文献综述、起草论文并修改完善（焦建利，2023）。教学活动中的大部分工作正在逐渐被 ChatGPT 承担，它的技术特性使其逐渐变成更具人类特点的教学角色，并会在将来的教学和学术活动中具有越来越重要的作用（秦渝超、刘革平、许颖，2023）。

3.2.2.2. 应用探索

一是教学评辅的整体发展，生成式人工智能与教育工作者、学习者关系最为密切，教师利用新技术感知学习者的需求，为学习者提供帮助。通过课前、课中和课后三个教学阶段促进教育的发展，ChatGPT 为教师教学提供工具支持，课前辅助教师进行教学设计，课中对学生进行教学辅导并生成学习报告，课后可以自动批改作业；为学生提供全方位的辅导，课前进行课程规划并生成个人学习计划，课中对学习者的学习状态进行监督和提醒；课后对作业进行查漏补缺；促进学习评价体系进行变革，课前教师根据学习者的学习情况和作业完成度制定个性化的教学方案，课中实时监测学生及时提供帮助和辅导，课后通过作业、小测验等方式了解学生的学习成果；而辅导贯穿于整个教师的教和学生的学当中，ChatGPT 作为一种辅助工具，在课前、课中和课后与教师进行人机协同，共同帮助学习者学习（肖君、白庆春、陈沫等，2023）。除上述方法外，还要推进教育理念改革，教育应该更加重视学生高阶思维能力的培养；创新教学方式与内容，将技术纳入课堂，丰富教学内容；鼓励教育与技术互促共进，教育工作者应该对技术持有开放和包容的态度，形成教育与技术双螺旋式的互促共进（卢宇、余京蕾、陈鹏鹤等，2023）。

二是发展人机协同模式，生成式人工智能在教育的应用，推动了新的教学模式的出现，最广为人知的就是人机协同模式。人机协同的知识生产模式是“非意向性”与“非创造性”结合的知识，根据已有知识进行内容重建，为每个学习者生成独特、个性化的内容（吴南中、陈咸彰、冯永，2023）。一些学者对“生成式人工智能+元宇宙”的人机协同模式进行了实证研究，根据学生的测试结果表明学习者在这种人机协同的环境下的学习绩效显著提升。元宇宙创设了利于学习者主动记忆的环境，把知识变为社区化传递，并且这种人机交互模式反向推动学习者对知识的主动记忆；学习者能将知识运用到不同的场景，并且当他们的问题解决后还能不断地提出新问题，最后他们学习的知识能超过老师预期，向预期外的领域进行拓展（翟雪松、楚肖燕、焦丽珍等，2023）。为知晓生成式人工智能在具体方面的表现，一些研究将焦点缩小，例如探讨研究生在人机协同的学术写作方面的实践研究，教师应该在学术写作方面提供支持，可组织学生讨论 ChatGPT 生成的内容；在课堂上组织学生对 ChatGPT 生成某一主题的论文进行学习和发表意见；鼓励学生与生成式人工智能协作进行学术写作，但是同时也需要他们对人工智能生成的内容进行标注；鼓励学生主动评估 ChatGPT 生成内容的质量，并反思自身的投入和表现（李艳、金皓月、杨玉辉，2023）。由于“作业黑箱”的存在，教师无法时刻监督学生，也无法区分是学生还是人工智能所生成的内容，这个过程只能依靠学生自己的学术道德进行监督，因此创新学生作业的评价方式。教师可以通过学生的表演讨论等方式去检验学生作业的真实性和创新程度；还能通过学生介绍作业的创新点或解决问题的方式来判断作业的真实程度，从而区分作业是由学生还是人工智能所完成（李海峰、王炜，2023）。

3.2.3. ChatGPT 在教育中的挑战与应对策略

3.2.3.1. 挑战

一是伦理困境和成本高昂，ChatGPT 的出现不仅让学生的学业诚信遭到质疑，同时也对教师的能力提出了更高的要求。除此之外，ChatGPT 在教育应用中可能会造成伦理困境。意识的淡薄，数据库的存储不当，同时也没有监管机制进行监测，导致了数据隐私的泄露与滥用；认知的固化，数据受限和偏差致使机器算法的歧视与偏见；重心的偏移，ChatGPT 可能会让师生主体地位消解并且造成师生关系甚至教育的异化；鉴别的错误，ChatGPT 这类技术产品是由国外研发，并不是每个人都可以直接使用，导致使用权上的不平等（王佑镁、王旦、梁炜怡等，2023）。而且不是所有人都能够从人工智能中受益，发达地区的学习者能更快地接受新技术，但是偏远地区面临着多种资源和技术的限制，这种不平衡会加剧地区之间的教育资源分配问题，加深不同人群之间的数字鸿沟（徐升、佟佳睿、胡祥恩，2024）。技术应用带来的教育伦理问题由来已久，生成式人工智能的过度使用产生“主体缺失”问题，对问题也停

留于表面分析，最后就是造成严重的学术不端问题（吴南中、陈咸彰、冯永，2023）。此外，人工智能的发展成本因为训练数据和计算能力需求而逐渐增加，教育领域的大语言模型的开发需要更高的资金投入，普通学校独立开发教育领域的大语言模型不太可能成功（刘明、吴忠明、廖剑等，2023）。

二是教育的异化，生成式人工智能可能导致教育主体思想同质化，加剧对教育主体思想的“殖民”；人工智能逐渐成为师生交流及行动的主体，导致了师生主体身份的缺失；技术逐渐融入教育数据，做出隐私侵害的行为，黑箱效应也越发凸显，加剧教育信任危机（张黎、周霖、赵磊磊，2023）。同时，ChatGPT 对教师的专业发展也提出了挑战，它削弱了教师的立德树人的教育观，固化了教师的教学创新和改革的思想；海量碎片化的信息导致教师信息过载，难以找到重点和核心内容，知识难以迁移；对教师的数字素养和高阶能力发出挑战；技术冲击了教师身份，教师对自身的定位也感到迷茫，教师在学生心中的地位和权威受到挑战（吴军其、吴飞燕、文思娇等，2023）。除教育主体的异化外，ChatGPT 还导致认知过程、实践活动和教育环境的异化。学生过度依赖 ChatGPT，长此以往内脑的活性逐渐减弱，成为外脑的附庸，影响学生认知的主动建构；人工智能技术引发身份认同危机，未来人类个性化的身体特质都可以随意复制，身体的使用权也逐渐被技术掠夺；当教育逐渐被人工智能技术占领，而人工智能技术背后由资本控制，教育目的从以人为本到追逐资本利益，引发教育环境的剧变（詹泽慧、季瑜、牛世婧等，2023）。

教育的异化给人才培养带来难题，部分学者表现出对拔尖创新人才培养的担忧，其“拿来主义”和学术失范导致道德弱化，内容生成过程的技术黑箱导致真伪难辨，这种技术的不透明性虽然能观察生成式人工智能的输出结果，但是输出的结果可能是有害的，学生和教育工作者还会对结果产生过度信赖（王炜、赵帅、黄慕雄，2024），人工智能的优势消弭了拔尖创新人才的创新性，使其沦为人工智能的寄生虫（王景、李延平，2023）。一些学者则是表现出对网安人才培养的担忧，训练数据的缺陷、法律伦理冲突、过度依赖沉迷和学术不端的诱惑可能导致人的认知固化，抑制创新思维的发展并且引发知识产权问题（王磊、徐子竞、朱戈等，2023）。

3.2.3.2 应对策略

一是通过政策指导教育教学是确保教育质量和公平性的重要手段，也是推动教育改革发展的重要力量。在借鉴日本的《中小学生成式人工智能应用指南》的背景下，为生成式人工智能在教育应用提供政策指导。价值观上倡导“技术向善”，其最终目标是以社会主义核心价值观为基础，促进人的全面发展；素养观上关注中小学师生智能信息素养的培养；应用观上明确生成式人工智能在中小学教育应用的边界，有效将人工智能与教育教学深度融合；政策观上采取“试点-推广”教育应用普及模式，先试点再推广，以确保教育资源的合理分配（孙立会、周亮，2023）。实践观上强调对教学实践进行指导，首先认识教学实践工具的可用性，师生要认知到 ChatGPT 的优势与不足；其次提升教学实践过程的创造性，ChatGPT 能促进教学目标、教学方式和教学评价的变革和创新；再次增强教学实践主体的能动性，教育工作者与学习者要主动向人工智能学习，提升自己的人工智能素养；最后确保教学实践场域的清朗性，国家、地方和学校要制定关于应用 ChatGPT 时应该遵守的伦理规范，对师生应用 ChatGPT 时分别提出要求（郑燕林、任维武，2023）。

二是重视教师主体性，在教育教学中应该让每个人成为他自己。重视教师自身的主体性成长，学生是全面发展的，那么教师也应该是全面自由发展的。教师要做好技术伦理的教育与示范，将技术运用到教学过程中需要遵循技术教学应用伦理；创建有温度的人际交往空间，技术在知识互动中发挥了重要的中介作用，但是教师与学生之间的情感互动式技术难以替代的，亟需关注教师和学生情感发展，积极创设富有人情温暖的人智交互模式（周玲、王烽，

2023）。

三是培养学生的数字素养，学生能辨别、判断和分析智能工具，并将各种技术工具为自己所学所用。在使用人工智能过程中学生需要具备强大的自我管理能力，不依赖于生成式人工智能的输出结果，能有自己的看法和见解，负责任地使用技术，养成数字伦理道德。学生可以利用生成式人工智能的强大数据分析能力，为自己提供特定的学习场景，制订个性化的学习环境，培养高阶认知能力，提升思维创新能力，促进学习者自身的全面发展。教师可以鼓励学生参与实地调研，将理论知识与实际问题相结合，在人机协同的过程中提升综合素养（严奕峰、丁杰、高赢等，2024）。

4. 结语

生成式人工智能尤其是 ChatGPT 正在世界掀起一股热潮，面对 ChatGPT，我们既有对新兴技术的追求与热情，同时也透露了对新兴技术的担忧。ChatGPT 在教师备课、收集资料、改善评价方式等方面有良好的帮助作用，学生能够提高学习效率，进行个性化学习，但有好处的同时也会带来坏处，伦理风险和知识产权争议、教师和学生的过度依赖导致学习惰性、教育成本的增加等问题出现。正是因为 ChatGPT 是一把双刃剑，我们更应该审视它在教育的应用，要以一种冷静而坚定地态度去拥抱它，教师要在课堂学会积极的使用人工智能技术，同时也要了解它带来的风险，进行及时的规避；学生利用 ChatGPT 进行自主学习和个性化学习，同时也要增强学术诚信意识，防止抄袭和作弊。技术真正地融入教育教学是一个漫长的过程，但人类必须看到人工智能与社会的发展的融合是一种必然的趋势，“现在”到“未来”的教育必是从人机协同创作到人机融合、人机共生的教育图景（贾同、蔡建东，2024）。而中小学教师如何在真实的教学场景中去运用生成式人工智能是后续研究需要关注的重要议题，进一步将人工智能技术落实到教学实践。

参考文献

王正青,阿衣布恩·别尔力克.ChatGPT 升级:GPT-4 应用于未来大学教学的可能价值与陷阱[J].现代远距离教育,2023,(03):3-11.

王竹立,吴彦茹,王云.ChatGPT 与教育变革——智能时代教育应如何转型[J].远程教育杂志,2023,41(04):27-36.

王佑镁,王旦,梁炜怡等.ChatGPT 教育应用的伦理风险与规避进路[J].开放教育研究,2023,29(02):26-35.

王炜,赵帅,黄慕雄.生成式人工智能教育创新应用的人本主义追求——对 UNESCO《教育与研究领域生成式人工智能指南》的解读[J].现代远程教育研究,2024,36(01):3-11.

王景,李延平.ChatGPT 浪潮下拔尖创新人才的培养：价值意蕴、现实隐忧与生态重塑[J].中国电化教育,2023,(11):62-71.

王磊,徐子竞,朱戈等.生成式人工智能赋能网络安全人才培养的探索研究[J].中国电化教育,2023,(09):101-108+116.

卢宇,余京蕾,陈鹏鹤等.生成式人工智能的教育应用与展望——以 ChatGPT 系统为例[J].中国远程教育,2023,43(04):24-31+51.

冯建军.我们如何看待 ChatGPT 对教育的挑战[J].中国电化教育,2023,(07):1-6+13.

苗逢春.生成式人工智能技术原理及其教育适用性考证[J].现代教育技术,2023,33(11):5-18.

孙立会,周亮.面向中小学的生成式人工智能教育政策制定路向——基于日本《中小学生成式人工智能教育应用指南》的分析[J].中国电化教育,2023,(11):53-61.

闫利利,侯虹印.ChatGPT 道德感与高教德育的技术限度问题[J].中国电化教育,2023,(11):72-82.

刘明,吴忠明,廖剑等.大语言模型的教育应用：原理、现状与挑战——从轻量级 BERT 到对话式 ChatGPT[J].现代教育技术,2023,33(08):19-28.
吴军其,吴飞燕,文思娇等.ChatGPT 赋能教师专业发展：机遇、挑战和路径[J].中国电化教育,2023,(05):15-23+33.
吴南中,陈咸彰,冯永.从“失序”到“有序”:生成式人工智能教育应用的转向及其生成机制[J].远程教育杂志,2023,41(06):42-51.
肖君,白庆春,陈沫等.生成式人工智能赋能在线学习场景与实施路径[J].电化教育研究,2023,44(09):57-63+99.
杨欣.基于生成式人工智能的教育转型图景——ChatGPT 究竟对教育意味着什么[J].中国电化教育,2023,(05):1-8+14.
张绒.生成式人工智能技术对教育领域的影响——关于 ChatGPT 的专访[J].电化教育研究,2023,44(02):5-14.
严奕峰,丁杰,高赢等.生成式人工智能赋能数字时代育人转型[J].开放教育研究,2024,30(02):42-48.
李艳,金皓月,杨玉辉.基于ChatGPT的研究生人机协同学术写作实践研究及启示[J].远程教育杂志,2023,41(05):38-48+75.
李海峰,王炜.生成式人工智能时代的学生作业设计与评价[J].开放教育研究,2023,29(03):31-39.
张黎,周霖,赵磊磊.生成式人工智能教育应用风险及其规避——基于教育主体性视角[J].开放教育研究,2023,29(05):47-53.
周玲,王烽.生成式人工智能的教育启示：让每个人成为他自己[J].中国电化教育,2023,(05):9-14.
郑燕林,任维武.实践观视域下 ChatGPT 教学应用的路径选择[J].现代远距离教育,2023,(02):3-10.
骆飞,马雨璇.人工智能生成内容对学术生态的影响与应对——基于 ChatGPT 的讨论与分析[J].现代教育技术,2023,33(06):15-25.
徐升,佟佳睿,胡祥恩.下一代个性化学习：生成式人工智能增强智能辅导系统[J].开放教育研究,2024,30(02):13-22.
贾同,蔡建东.生成式人工智能对教育生产力的变革[J].现代教育技术,2024,34(01):107-116.
郭顺峰,李光,邹红军.ChatGPT 引发的“以人为师”合法性危机与应对——基于技术批判理论视角[J].电化教育研究,2023,44(11):28-35.
秦渝超,刘革平,许颖.生成式人工智能如何重塑教学活动——基于活动理论的模型构建与应用[J].中国远程教育,2023,43(12):34-45.
焦建利.ChatGPT：学校教育的朋友还是敌人？[J].现代教育技术,2023,33(04):5-15.
詹泽慧,季瑜,牛世婧等.ChatGPT 嵌入教育生态的内在机理、表征形态及风险化解[J].现代远距离教育,2023,(04):3-13.
翟雪松,楚肖燕,焦丽珍等.基于“生成式人工智能+元宇宙”的人机协同学习模式研究[J].开放教育研究,2023,29(05):26-36.

基于自然语言处理的深度学习课堂中教师话语特征分析与应用研究

Analysis and Application Research on Teacher Discourse Characteristics in Deep Learning Classroom Based on Natural Language Processing

曾钰琦[1]，钟瑾[1]，刘博[1]
[1]华南师范大学教育信息技术学院
2908389088@qq.com

【摘要】 深度学习是落实学生课堂主体地位的有效途径。本研究立足于深度学习课堂，以教师话语为研究对象，基于自然语言处理技术对教师话语进行智能分类，通过定量分析不同深浅程度课堂中教师话语的特点与结构，研究发现：深层学习课堂中教师更善于结合引导、追问等多种类型的话语进行知识授受，也体现出更丰富的师生交流样态。最后，本研究围绕教师话语特性提出以下策略：精进提问技艺，引发深度思考；重视对话互动，促进深度参与；发挥语言亲和力，促进深度联结。

【关键词】 教师话语分析；深度学习课堂；自然语言处理

Abstract: *This study examines teacher discourse in deep learning classrooms using natural language processing for intelligent classification. It finds that in deep learning environments, teachers utilize more guiding and questioning language, enhancing communication with students. Strategies proposed to promote deep learning include improving questioning skills, fostering dialogue, and leveraging language affinity for deeper connections.*

Keywords: Teacher Discourse Analysis, Deep Learning Classroom, Natural Language Processing

1. 引言

2022 年 4 月，教育部印发的《义务教育课程方案和课程标准（2022 年版）》，明确课程目标需指向学生的核心素养。深度学习作为促进学生核心素养发展的关键路径（郑葳和刘月霞，2018），在学生综合素质的养成中有着重要的应用价值和意义。深度学习的过程是一个对话的过程，是师生话语互动的内化（张光陆，2020）。教师话语作为教学开展的主要载体，是师生交流互动的重要媒介（赵云建,汪洋，陈玲和何静，2023）。师生间的深度互动有助于推动学生从浅表学习过渡到深度学习（彭海涛，2021）。因此，研究教师如何利用语言将深度学习理论贯彻于课堂中各环节，是亟须开展的现实课题。当前教师话语研究主要采用传统技术，如人工编码和视频分析，但存在着耗时、工作量大、成本高等问题。随着人工智能等新兴智能技术的蓬勃发展，技术赋能教育智能发展成为必然趋势。自然语言处理作为人工智能中的重要分支，在文本、情感分析等方面有着可操作性，为研究提供了新的思路和方法。

鉴于此，本研究立足于深度学习课堂，以教师话语为研究对象，基于自然语言处理技术对教师话语进行智能分类，采用定量分析的方法研究不同深浅程度课堂中教师话语分布的特点，挖掘深度学习课堂中教师的关键话语结构，同时，以定性分析的视角揭示深度学习课堂中教师话语的运用方略，提升课堂教学深度。

具体研究问题如下：（1）基于何种自然语言处理模型能够实现教师话语的分类，模型的效果如何？（2）不同层次课堂中教师话语有何特征？存在着什么差异？（3）深度学习课堂中教师是如何利用教学话语促进学生深度学习的？

2. 文献综述

2.1. 教师话语测量的理论基础

在课堂话语的量化分析方面，耳熟能详的是美国学者弗兰德斯提出的互动分析系统 FIAS，将课堂中教师的言语行为分为：接受情感、表扬或鼓励、接受或使用学生想法、提问、讲授、指令、批评或辩解权威七类（Flanders, 1963）。后续研究者如叶立军等在 FIAS 的基础上进行了改进，将教师课堂教学语言明确为反馈性语言、鼓励性语言、赞成启发性语言、师生共同重复语言、提问性语言、陈述性语言和命令性语言(叶立军和斯海霞，2011)。随着教育现代化的落实，顾小清等提出了基于信息技术的交互分析系统 ITIAS，并强调量化分析需要与质性评价的相结合，才能有效地描述课堂真实的教学情况（顾小清和王炜，2004）。

结合文献发现，在教师话语测量上主要以弗兰德斯互动分析系统为原型，针对不同的研究目的进行改进优化，该系统为教师话语分析提供了一种定量化的研究方法与。因此，本研究以该方法为启发，制定适配的教师话语分类编码表，并以此为标准对教师话语进行编码分析。

2.2. 基于自然语言处理的课堂话语研究

目前该技术在课堂话语上的研究主要分为：文本分析、情感分析、问答与对话系统设计等。在文本分析上，有研究将 MOOC 中学生的活动文本、活动频率、活动顺序等作为分析数据，通过相关性、多重回归和过程挖掘等方法预测学生的行为表现（Conijn,Van den Beemt,& Cuipers, 2018）；情感分析领域，有研究基于 LSTM 网络分析学生反馈文本中情绪信息，以帮助教师了解学生的学习状况（冯翔, 邱龙辉和郭晓然, 2019）；在问答与对话系统设计方面，有团队提出基于预训练语言模型的文本−回复匹配模型，解决了对话系统中常忽略上下文相关信息的问题，提高了对话系统的智能性（Xu, et al., 2021）。

结合文献发现，技术支持下的课堂话语研究侧重于学生，缺乏对教师的分析，且样本多以论坛等文本资料为主，缺乏课堂实录。因此，本研究从课堂实录出发，尝试采用深度学习的方法实现对教师话语的自动判定与智能分类。

2.3. 深度学习课堂中教师话语的实证研究

教师话语作为教学开展的主要载体，对其进行研究有助于深入了解促进学生深度学习的教师话语特征。如陈薇等通过对 4 位专家教师 16 节课的观察，采用描述性统计方法和定性分析方法，发现他们常采用隐蔽多重、一问多答和理解转述等策略设计深度问题，促进学生思考（陈薇和沈书生，2019）；张娟娟等通过对 20 节优秀教师汇报课的编码分析，发现深度教学中教师的反馈不仅仅局限于简单的对错评价，而是涉及到如质疑、回音等多种互动类型，这种反馈方式更有助于学生开展深层次的探讨与思考（张娟娟,陈旭远,范会敏和毛清荟，2021）。

结合文献发现，课堂中教师话语的研究方法主要采用课堂观察、问卷调查、案例研究，在数据分析方面主要采用统计分析、差异性分析和滞后序列分析。因此，本研究采用定量与定性分析的方法，揭示深度学习课堂中教师话语的运用方略。

3. 研究设计

3.1. 数据来源与量表

本研究采用的视频课例来源于大亚湾精品课例的视频课例库共 54 节，为保障研究结果的信度，研究课例需满足以下三个条件：(1)课例为覆盖不同学段的不同学科课例，除英语学科；(2)案例的画像与声音录制清晰，具备研究和编码的特性；(3)视频课例的录制时间为近 3 年内。筛选满足条件的课例 20 节，以用作教师话语特征分析的数据集，含语文学科 7 节，数学学科 6 节，科学、地理和历史学科各 2 节，美术学科 1 节，其中 10 节则作为构建教师话语分类模型的训练集和测试集。

本研究参考黄友初(黄友初，2016)对教师课堂话语的划分，考虑到自动化的可实现性，结合本研究对课例的观察，将教师话语分为：反馈性、激励性、引导性、提问性、陈述性、命令性、重复性、过渡性和追问性 9 种类别(见表 1)。

表 1 课堂教师话语分类编码表

主体	编码	类别	概念内涵	示例
教师话语	T0	反馈性	向学生提供其在学习中具体表现的评价和意见，以帮助学生改进学习	“在这题中，你有一些计算上的错误，需要再认真检查。”
	T1	激励性	鼓励学生积极学习和发展的具体言辞和语言表达，以激发学习动力	“你在小组活动中的投入和专注让老师印象深刻。”
	T2	引导性	用于引导和启发学生行动或思考方向的指导性语言	“你可以想象一下，如果我们改变这个条件会发生什么变化？”
	T3	提问性	以问题形式考察学生对知识的理解和应用程度的话语,激发思考和讨论	“在这段话中，作者如何使用比喻来传达主题？”
	T4	陈述性	在教学中使用的直接描述概念、事实或观点的语言表达方式,旨在传递信息、表达思想或分享知识	“半径是指从圆心到圆周上任意一点的距离，所以在圆中所有点到圆心的距离相等。”
	T5	命令性	教师以明确、简洁的话语要求学生按特定要求执行某项任务、行为或动作	“打开教科书到第四十页，我们要一起讨论第三章的内容。”
	T6	重复性	教师用于强调、巩固和增强学生对特定观点、概念或信息记忆的话语	“我已经提到过这个概念了，这个概念很重要。”
	T7	过渡性	用来引导学生从一个话题、活动或教学阶段切换到另一个部分的话语	“刚刚我们学习了圆柱表面积的方式，那圆锥的表面积怎么计算呢？”
	T8	追问性	以问题或回应学生的方式,以促使学生深入思考、表达更多见解或解释,引导学生探讨和学习	“可以具体描述一下主人公在小说中表现出勇敢特征的情节吗？你认为勇敢是他的核心特点吗？”

3.2. 数据预处理

用作本研究教育话语分类模型的训练集和测试集的课例共 13 节。首先，对 13 节课例进行人工编码，为保证数据集的科学性和有效性，在编码的过程中对两名研究人员编码的结果进行一致性检验（Kappa>0.75），并对不同编码进行讨论和更正，共计编码 3345 条。模型训练集与测试集按 7：3 进行划分，得到训练集 2342 条，测试集 1003 条，其中反馈性话语 (86 条)、 激励性(192 条)、引导性(170 条)、提问性(1026 条)、陈述性(881 条)、命令性(704 条)、重复性(30 条)、过渡性(91 条)和追问性(165 条)。针对模型泛化能力差、容易过拟合的情况，本研究在训练过程中随机对训练数据进行同义词替换、随机插入、交换和组合等数据增强技术，增强数据的多样化以及扰动信息，使得各个类别的数据量较为均衡，提高模型的泛化能力。

3.3. 教师话语分类模型训练

本研究使用 BERT-wwm 模型对编码文本进行训练。其中 BERT（Bidirectional Encoder Representations from Transformers）是一种深度双向 Transformer 模型，由 Google 在 2018 年提出，它通过预训练方式学习双向语境的表达，并且可以微调以适应特定的下游任务，而 BERT-wwm 则是结合中文场景特点进行优化改进的语言分类模型（Cui Y, Che W,&Liu T, 2021）。BERT-wwm 使用了整词掩码的预训练策略，针对中文文本中词汇作为基本单位的特点进行了优化，使得模型能够更加精准地理解和捕捉中文词汇的语义信息。其次，BERT-wwm 在预训练过程中针对中文语境进行了优化，使用了更适合中文语言特点的语料库，因此在词

汇覆盖和语义理解方面较其他模型有更大的提升。本文的实验环境见表 2。模型训练超参数如下：批大小（Batch Size）为 16，学习率（Learning Rate）为 1e-3，训练轮数（Epochs）为 100，优化器（Optimizer）为 Adam，神经元丢弃率（dropout_prob）为 0.3，预热轮数（Warm up）为 5。经过训练，本研究模型在测试集上表现出色，测试集的平均准确率达到了 92.73%，加权平均 F1 值更是高达 0.93，说明模型对教师话语分类上具有一定的优越性以及出色的泛化能力，也为深度学习在教学内容上的自动化分类和理解打开了崭新的可能性。

表 2 实验环境配置

名称	参数
操作系统	Ubuntu 16.04.2 LTS
CPU 型号	Intel • i9 • 10980XE • CPU
GPU 型号	GeForce RTX 3090
单核显存	24G
GPU 数量	1
框架版本	PyTorch 1.9.1
CUDA 版本	11.1

3.4. 数据分析方法

根据研究问题，本文按照张浩（张浩，吴秀娟和王静，2014）修订的深度学习认知评价体系对 20 节课例进行学生的认知编码并将学生认知得分大于等于 3 的 10 节课例命名为深层学习课堂（马如霞, 王陆和彭玏，2022），其余为浅层学习课堂。数据分析阶段由四部分组成：首先，采用量化的研究方式，即描述性统计分析，研究不同课堂中教师话语各类别的分布结构。接着，采用差异性分析，研究不同课堂中教师话语的特点与差异，结合滞后序列分析，研究不同课堂中教师话语转化的关键模式。最后，采用质性分析，回归到课堂本身具体的教学阶段，基于深层学习课堂中教书话语的特征提出具体的策略与建议。

4. 研究结果与讨论

4.1. 教师话语的结构分析

利用教师话语分类模型对划分好深浅层次但尚未进行人工编码的课例进行教师话语的自动化编码和分类，再由研究人员对部分置信度不高编码进行修正，以得到用于教师话语特征分析的 4990 个编码。统计不同层次课堂中 8 种教师话语类型的占比情况（见图 1）。

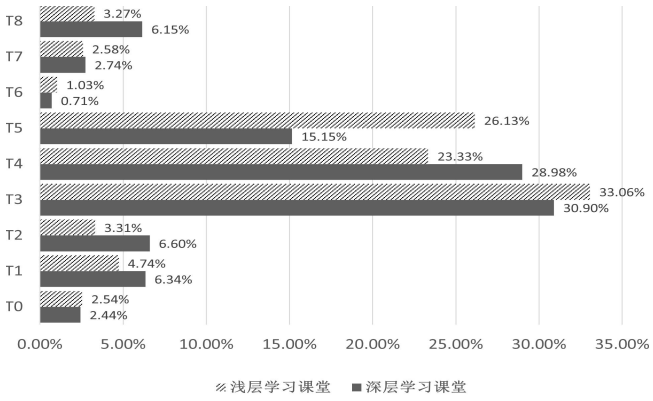


图 1 教师话语类型的占比情况

在深层学习课堂中，教师的提问性话语(T3，30.90%)和陈述性话语(T4，28.98%)占比最高，两者合计占全体教师话语行为的近 60%。在浅层学习课堂中，提问性(T3，33.06%)和命令性话语(T5，26.13%)比重较大，远高于其他话语类型。可见，提问策略在课堂活动组织和

教学流程推动方面有着重要应用价值。值得注意的是，深层学习课堂中教师更倾向于使用陈述性话语，表明教师注重知识授受，关注学生基础知识与基本技能的掌握，而在浅层学习课堂中教师的命令性话语较多，表明教师的重点在于学生是否有理解并遵循教学指令按部就班地进行学习。然而，重复性(T6)和过渡性(T7)话语较少，在两类课堂中所占比重皆少于 5%，结合课例观察，重复性话语会更多地出现在复习课或习题课中而非新授课，这与课例样本的采集情况有关，而过渡性话语较少或许是因为该类话语与教师的引导性和追问性话语出现了相交的情况。

4.2. 教师话语的差异性分析

为探究不同层次课堂下教师话语类别的具体差异，本研究采用独立样本 t 检验对不同层次课堂中的教师话语进行差异性分析，由于样本课例的限制和教师个体差异等问题导致数据分析过程中有个别教师话语不满足方差齐性检验，因此本文对此类教师话语进行非参数检验以研究其差异性(见表 2)。深层学习课堂的 T1、T2 和 T8 显著高于浅层学习课堂($p<0.05$)于非参数检验中，两类课堂的 T4 和 T5 皆存在显著差异($p<0.05$)，具体而言 T4 在深层学习课堂中出现频次更高，而 T5 则于浅层学习课堂中较为明显。因此，深层学习课堂中 T1、T2、T4 和 T8 显著高于浅层学习课堂，而 T5 明显低于后者。结果表明，在深层学习课堂中教师更善于用反馈性、引导性、陈述性和追问性的话语，对于命令性话语使用相对较少。

表 3 不同层次课堂中的教师话语的差异性

参数检验				
层级	组别	平均数±标准差	t 值	p 值
T0	浅层学习课堂	5.90±4.701	0.295	0.772
	深层学习课堂	6.50±4.403		
T1	浅层学习课堂	11.00±5.345	2.883	0.010
	深层学习课堂	16.90±3.635		
T2	浅层学习课堂	7.70±4.877	5.436	0.001
	深层学习课堂	17.60±3.062		
T3	浅层学习课堂	76.80±24.316	0.544	0.593
	深层学习课堂	82.40±21.634		
T7	浅层学习课堂	6.00±3.432	0.946	0.357
	深层学习课堂	7.00±3.391		
T8	浅层学习课堂	3.20±4.087	2.554	0.020
	深层学习课堂	16.40±8.019		
非参数检验				
层级	组别	中位数	秩均值/秩和	p 值
T4	浅层学习课堂	54.50(37.50~72.00)	6.50/65.00	0.002
	深层学习课堂	81.50(72.75~85.25)	14.50/145.00	
T5	浅层学习课堂	64.00(40.25~82.00)	13.15/131.50	0.043
	深层学习课堂	40.50(34.25~45.25)	7.85/78.50	
T6	浅层学习课堂	2.50(0.75~4.00)	11.20/112.00	0.631
	深层学习课堂	2.00(1.00~2.25)	9.80/98.00	

注：$p<0.05$ 表示二者存在显著差异

结合文献研究，发现教师更注重与学生建立合作性学习关系，通过引导、陈述、追问和反馈来激发学生的深度认知水平，而不是仅仅下达命令或指示。数据分析与课例观察的结果

显示，深层学习课堂中教师注重学生知识的迁移运用与思维的启发引导，他们巧妙运用引导性话语，引导学生理解知识的内在结构和创造性应用，从而提高学生的深度认知水平。另一方面，学生深度学习需要建立在知识深度理解的基础上。教师的陈述性话语是系统性讲解知识和学生基础知识构建的前提。然而，很多教师常常忽略知识陈述的技巧与重要性，过度关注简单堆砌和教学模式的胡乱套用，而不顾学生基本理解的建立。这会割裂学生的学习过程，增加他们的负担。除此之外，教师的追问和反馈是促进学生深度学习的重要桥梁，通过追问了解学生对知识的掌握程度，引导学生进行有意识进一步思考。通过有效的反馈了解学生的思维进展，纠正错误，调整追问策略引导学生重新辨析知识之间的逻辑关系，让学生在各种沟通中走向知识的多元结合(郭华，2016)，达到思维的进阶。

4.3. 教师话语的转换特征分析

为进一步研究不同层次课堂中教师话语转换的模式，本研究采用滞后序列分析方法来探索教师话语的转换模式。下表 3 和表 4 为不同层次课堂教师话语序列频次调整后的残差值(Z>1.96，则该序列具有显著性 p<0.05），筛选具有显著性的行为序列并画出如下图 2 和图 3 的教师话语序列转换图。结合序列转换与课堂实录发现：浅层学习课堂与深层学习课堂中教师通常会以提问或陈述的方式作为教学流程的起点(T3→T4)、(T4→T3)。不同的是，浅层学习课堂中教师运用 T5 的次数较多，且多为话语序列的起点(T5→T0/T1/T3)，呈现教师命令主导的状态，而深层学习课堂中教师则以 T3 为主(T3→T4/T5/T8)，与浅层学习课堂相比，深层学习课堂呈现师生相互对话交流的样态。

表 4 深层学习课堂教师话语调整后的残差表

	T0	T1	T2	T3	T4	T5	T6	T7	T8
T0	-1.35	-0.14	1.38	0.14	-1.14	1.44	0.75	-0.63	-0.52
T1	0.37	-1.52	-1.19	0.28	-0.69	1.00	0.75	3.18**	-0.64
T2	1.91	-1.17	-0.52	-0.56	1.92	-1.01	0.76	-0.26	-0.62
T3	0.42	1.19	-0.24	-8.05	3.70**	4.94**	-1.27	-2.59	2.03**
T4	-0.88	0.36	0.55	5.45**	-2.49	-5.05	-0.05	1.85	0.21
T5	0.03	0.54	1.00	4.55**	-4.75	-1.09	0.14	0.16	0.02
T6	0.75	-1.14	-1.14	0.27	-0.04	1.45	-0.37	-0.73	-0.12
T7	-1.39	-0.76	-0.76	0.35	1.05	0.84	0.69	-0.7	-1.14
T8	0.02	0.06	-0.28	-1.81	3.52**	-0.41	-0.11	-0.61	-1.81

注：**表示 p<0.05，下同。

进一步分析，在深层学习课堂中，教师不仅通过提问来评估学生对知识的理解程度，更善于以学生当前的知识水平为起点提出深入追问(T3→T8，Z=2.03)，将学生引向对知识内在系统的理解与运用，并赋予知识的详细阐述(T8→T4，Z=3.52)，让学生能够在充分的信息交流中深化知识理解，整合内容信息，构建新的知识系统，逐步迈向深度认知的阶段。除此之外，深度学习还需要学生积极的情感投入与知识价值体系的个性化建立(郭华，2016)，在深层学习课堂中，教师不吝啬对学生的激励，为其提供高质量的情感共鸣，创建知识的问题情境与良好的情感氛围，并能够行云流水般过渡到另一个知识情境(T1→T7，Z=3.18)，助推学生全身心参与到有意义的学习过程中，形成对知识深刻的情感理解，从知识输入到情感升华(杨洁和王安全，2017)。在浅层学习课堂中，教师也注重学生的知识追问(T8→T4，Z=3.12)与过渡引导(T7→T2，Z=2.11)，但这种关注更多地笼罩在教师发号施令(T5→T0/T1/T3)和追问不止致使压力增加(T8→T8，Z=3.41)的紧张氛围中，不仅忽略了学生对知识的迁移应用，还忽视了学生情感价值的培养，因此难以将学生从浅表学习引向深度学习。

表 5 浅层学习课堂教师话语调整后的残差表

	T0	T1	T2	T3	T4	T5	T6	T7	T8
T0	-0.54	-0.03	-0.69	-0.35	1.02	0.37	0.52	-0.61	-0.79
T1	0.56	-0.26	-1.41	-0.24	0.13	0.45	-1.09	1.00	0.08
T2	1.54	-1.41	-0.16	-1.11	1.55	-0.02	-0.85	-0.05	0.38
T3	-2.92	-0.49	-0.35	-7.75	4.88**	5.88**	-1.39	-1.27	0.43
T4	-2.51	-0.59	1.25	4.96**	-4.58	-1.04	1.24	1.33	-0.04
T5	4.07**	2.32**	0.00	5.54**	-3.43	-5.18	0.85	-0.21	-0.96
T6	1.83	-1.09	0.37	-0.45	-0.34	0.32	-0.47	1.72	-0.89
T7	0.16	-0.73	2.11**	-1.25	0.4	0.35	1.72	0.81	-1.55
T8	0.64	0.08	-1.62	-1.32	3.12**	-1.81	-0.89	-1.55	3.41**

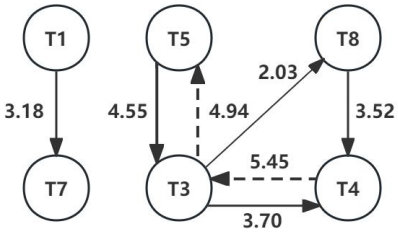


图 2 深层学习课堂中教师话语序列转换图

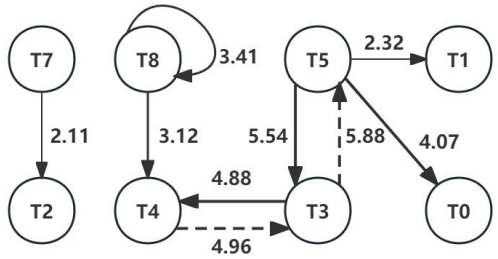


图 3 浅层学习课堂中教师话语序列转换图

5. 总结与展望

本研究从以教师话语作为研究对象，利用智能技术实现话语自动分类，对 8 种话语类型进行编码分析，研究结果显示：在深层学习课堂中，教师的提问与陈述话语占比较大，在教学过程中更善于结合引导性、追问性等多种类型的话语进行知识授受，同时也体现出更多师生间交流对话样态，如提问→追问→陈述、激励→过渡等。而正是这些话语特征对学生课堂深度参与有显著促进作用。据此，本研究围绕教师话语的特性，提出以下促进深度学习的提升策略，为有针对性地干预与改进课堂教学，提高学生深度学习水平提供参考。

5.1. 精进提问技艺，引发深度思考

教学中提问能够帮助教师达成多种教学目的，课堂教学的艺术某种程度上是课堂提问的艺术（洪松舟和卢正芝，2010）。好的提问是有方法讲技巧的，尤其是启发性的开放性问题更能促进学生的深层次思考。结合实录观察发现在浅层学习课堂中，提问数量上虽多，但深度有所欠缺。相对而言，深度学习环境中的提问更具挑战性，因为答案不是唯一的，需要学生深度思考和分析。此外，教师对问题的设置是渐进性的，基础性问题常用于考察学生对知识的识记，而后教师通过多次追问引导学生建立知识关联，激励学生进一步思考问题的不同维度和层面，促使深层次思考。总的来说，在实际教学应注重提问的艺术，根据教学目标和学生认知水平，设置不同层次的问题，并把控提问的节奏，以最大程度发挥提问在促进学生认知发展方面的价值。

5.2. 重视对话互动，促进深度参与

师生对话互动在教学中起关键作用，不仅是知识传递，更是交流、思考和深入理解的过程。通过课例视频和教师话语分析发现，在深层学习课堂中，教师更善于通过提问、引导、追问等多样化的语言形式建立与学生之间的有效沟通和互动。相反，在浅层学习课堂中，过度使用指令性话语容易中断交流，难以实现深度学习。学生要成为学习的主体，教师应充分发挥语言的互动功能，提升学生学习的参与感。然而，这并非简单排斥教师讲解，而是关注教师语言对学生思维活动的唤起（钟启泉，2006）。例如，教师提问题后应给学生思考时间，鼓

励独立思考构建知识，而非仓促提问形式化。对学生回答，教师可通过递进追问和及时反馈促进反思、重构与整合，解决疑惑，形成螺旋式上升学习。

5.3. 发挥语言亲和力，促进深度联结

教师语言行为的发生往往伴随着情感态度的传递。深度学习的达成需要建立在一个安全、开放的学习环境中（程良宏和白珊，2023）。学生感受到教师关心和支持时更愿意积极参与课堂，提出问题，甚至勇于尝试解决更具挑战性的问题，从而促进深度学习。因此，教师在教学中应注重语言的亲和度，适时采用激励性和表扬性措辞，建立积极的情感联系。在教师话语结构分析中，两类课堂在激励性话语使用频次相当，但不同教师在该类型话语表达方式有所不同。相较于浅层学习课堂中的模式化、模板化的肯定，深层学习课堂中教师对学生的表扬更具体生动和个性化，注重挖掘学生在解题、思考和学习过程中展现的独特见解和创造性思维。这种表扬能够激发学生对学科的兴趣和深度探索的动力，促使深度学习的有效实现。

参考文献

马如霞,王陆和彭玏(2022)。大数据的知识发现:促进课堂深度学习的策略[J]。**电化教育研究,43(05)**,84-91+108。

叶立军和斯海霞(2011)。基于录像分析背景下的代数课堂教学语言研究——以两堂《分式的乘除》课堂实录为例[J]。**数学教育学报,20(1)**,42-44。

冯翔,邱龙辉和郭晓然(2019)。基于 LSTM 模型的学生反馈文本学业情绪识别方法[J]。**开放教育研究,25(02)**,114-120。

杨洁和王安全(2017)。学生课堂积极沉默存在的合理性分析[J]。**教学与管理,(19)**,24-26。

张光陆(2020)。促进深度学习的师生话语互动:特征与提升对策[J]。**全球教育展望,49(10)**,27-38。

张浩,吴秀娟和王静(2014)。深度学习的目标与评价体系构建[J]。**中国电化教育,(07)**,51-55。

张娟娟,陈旭远,范会敏和毛清荟(2021)。深度教学的课堂话语特征——基于"架构-发展-评价"结构的审视[J]。**基础教育,18(06)**,84-95。

陈薇和沈书生(2019)。小学数学教学中深度问题的研究——基于专家教师课堂提问的案例分析[J]。**课程·教材·教法,39(10)**,118-123。

郑葳和刘月霞(2018)。深度学习:基于核心素养的教学改进[J]。**教育研究,(11)**,56-60。

赵云建,汪洋,陈玲和何静(2023)。教师间接语言助力课堂"提质增效"——基于优质课堂教师话语的数量结构与教学意义分析[J]。**电化教育研究,44(07)**,88-95。

钟启泉(2006)。知识建构与教学创新——社会建构主义知识论及其启示[J]。**全球教育展望,35(08)**,12-18。

洪松舟和卢正芝(2010)。中学教师课堂提问的社会学分析[J]。**课程.教材.教法,30(01)**,20-24。

顾小清和王炜(2004) 支持教师专业发展的课堂分析技术新探索[J]。**中国电化教育,(7)**,18-21。

郭华(2016)。深度学习及其意义[J]。**课程.教材.教法,(11)**,25-32。

黄友初(2016)。教师课堂教学行为的四个要素[J]。**数学教育学报,25(01)**,72-74。

彭海涛(2021)。课堂师生互动的深度与限度[J]。**教育理论与实践,41(14)**,57-60。

Conijn,R.,Van den Beemt,A..& Cuijpers,P(2018). Predicting student performance in a blended MOOC.*Journal of Computer Assisted Learning*,34(5):615-62.

Cui Y, Che W, Liu T, et al(2021).Pre-training with whole word masking for chinese bert[J]. *IEEE/ACM Transactions on Audio,Speech, and Language Processing*, 29: 3504-3514.

Flanders N(1963).Intent,action and feedback:a preparation for teaching.*Journal of teacher education*,14(3): 251-260.

Xu R, Tao C, Jiang D, et al(2021). Learning an effective context-response matching model with self-supervised tasks for retrieval-based dialogues.*Proceedings of the AAAI Conference on Artificial Intelligence*,35(16): 14158-14166.

传统智慧与数字化转型：儒家教育思想在数字教育中的应用[1]

Traditional Wisdom and Digital Transformation: Application of Confucian Education Thought in Digital Education

刁均峰[12]，连梅芳[12【通讯作者】]，韩锡斌[3]
（1.海南师范大学 教育学院，海南 海口 571158；
2.海南师范大学教育数字化转型研究中心，海南 海口 571158）
（3.清华大学 教育研究院，北京 100084）

[摘　要] 本文探讨了儒家智慧在数字教育中的应用与未来展望。通过对传统儒教智慧与数字化转型进行理性审视，立足数字教育的发展现状与挑战，梳理儒家智慧的基本原则，在此基础上，探析二者的现实契合性，并思考儒家智慧在数字教育中的应用策略，对传统儒家智慧融入新兴数字教育的未来进行展望，从而系统深刻地阐述传统儒家智慧的当代价值。

[关键词] 数字教育；儒学；教育改革；数字素养

Abstract: *This article explores the application and future prospects of Confucian wisdom in digital education. Through a rational examination of traditional Confucian wisdom and digital transformation, based on the current development and challenges of digital education, it sorts out the basic principles of Confucian wisdom. On this basis, it analyzes the practical fit between the two, considers the apply policy of Confucian wisdom in digital education, and forecasts the future integration of traditional Confucian wisdom into emerging digital education. This systematically and profoundly elaborates on the contemporary value of traditional Confucian wisdom.*

Keywords:*digital education, Confucianism, education reform, digital literacy*

1. 引言

数字教育作为引领新一轮教育教学改革的重要引擎和强劲动力，为深入推进教育高质量建设创造了有利条件。然而，数字技术的广泛应用和深入发展也对传统教育模式产生了冲击。儒家智慧作为中国教育思想史上极具生命力的典范，蕴含深厚价值内涵和传统文化底蕴，并随历史进程发展而自我更新完善，在数字时代焕发新的生机与活力。深入剖析儒家智慧并运用与数字教育，有利于应对教育数字化转型过程中的新问题、新挑战，推进儒家智慧与数字教育的融合创新，推动儒家智慧赋能数字教育新生态，促进教育教学改革创新。

2. 传统智慧与数字化转型之理性审视

2.1 数字化转型：当代教育改革的必然趋势

新兴数字技术的快速发展和普及应用，使得教育数字化转型成为当代教育改革的必然走向。未来教育的发展取决于数字技术的发展及其在教育过程中的地位，需要充分认识数字教

[1] 本文系2022年度海南省哲学社会科学规划基地课题“东盟基础教育数字化转型及其对海南自贸港教育发展的启示”（HNSK(JD)22-35）；海南省高等学校教育教学改革研究重点课题“自贸港高校教师教学能力数字化转型研究”（Hnjg2023ZD-17）；第一批海南省“南海新星”哲学社科人才平台项目“自贸港高等教育数字化转型研究”；2023年度海南省研究生创新科研课题“海南红色文化融入小学语文课程的跨学科活动设计研究”（Qhys2023-353）的阶段性研究成果之一。

育所带来的机遇和挑战，为推进教育高质量发展提供理论依据和实践指引。

数字化转型促进优质教育资源分配，推动教育公平化和优质化。传统教学存在数量不足、质量不高、共享程度低等教育资源局限性问题。数字教育汇聚了海量网络资源，促进开放教育资源的广泛应用，，既提高优质教育资源的供给量、覆盖面和使用率，还促进优质教育资源大范围传播、持续性流转和普惠性共享，促进教育资源效益最大化、配置均衡化。

数字化转型打破时空限制的泛在性教育，提高了教学质量和效率。传统教学模式将师生囿困于教室的四方空间，受到既定的时间、空间的教学条件局限。数字教育的泛在化从传统线下课堂拓展为线上线下融合的混合式教学，为学生提供了更广泛的学习机会、更灵活的学习方式和更多元的数字情境，推动时时、处处、人人可学的泛在学习成为教学新形式。

数字化转型为学生提供个性化学习机会，拓展知识视野。传统教学受到人口与资源分配的影响，多以大班额集体教学为主，知识传递表现为从教师到学生的单向传输，可能导致部分学生“吃不饱”或“跟不上”。数字教育为学生提供了丰富学习资源、便捷学习平台等强大的个性化学习支持条件，使教学趋于个性化、灵活化，利用现代技术为真正实现“因材施教”提供了可能性。同时，学习者可借助网络平台绘制“精准画像”，以真正了解自身的优势和不足，寻找合适的学习资源与路径，制定个性化学习任务，有效达成个性化学习目标。

然而，数字化转型可能带来新的教育不公平。一是技术脆弱性，互联网链接不顺畅、缺乏电子设备等数字不平等进一步演化为教育不平等，学生无法获得数字教育资源、无法参与在线学习、无法获得远程指导的机会，加剧了教育资源分配不均，扩大了教育数字鸿沟。二是教育资源良莠不齐，未经充分审查和质量控制的低质、同化或偏执的内容存在风险。三是学生素养，青年缺乏辨别资源优劣的能力，难以自主独立地从庞杂敏感的信息海洋中有效提取和选择学习资源，甚至可能使学生迷失自我，产生“技术依赖症”，钝化主体思维。四是教师数字素养，教师数字素养整体水平提升缓慢，尚未达到预期，造成技术使用难、用好难。

2.2 儒家教育思想的基本原则

为谋求数字教育的更好出路，可通过对儒家传统智慧进行现代性解读，为教育数字化转型提供根本性解答。“儒家在中国古代思想史上长期占据主导地位”。[1]儒家核心思想从“仁”出发，表现为“礼”，并以“礼”为最高准则形成潜移默化的道德教育，进而在全社会做到“尊师重教”的目的。“仁”“礼”“尊师重教”等核心思想具有历史性和时代性价值。

儒家的仁爱思想是儒家文化的精髓，蕴含着深刻的平等思想。“亲亲之爱”体现亲疏之别、薄厚之分的特殊之爱；“忠恕之道”是视人如己、人己平等的推广之爱；“恻隐之心”是心有不忍、推己及人的同情之爱，三者之间存在一定张力。儒家思想将“亲亲之爱”作为起点和基础，将“泛爱众而亲仁”的“忠恕之道”作为最高理想，“由‘爱亲’推广至‘泛爱众’”[2]则是实现“仁爱”的路向，同时还需要坚守“恻隐之心”的底线，方能“成仁”。

孔、孟、荀三人均重视“礼”的作用，一脉相承地表达了“礼”的重要性。孔子曰“不学礼，无以立”（《论语·季氏篇第十六》），即不懂礼就无法处世立身，其中“礼”即各种行为规范的总和，通过“礼”引导并约束行为，使每个人的行为符合社会规范，从而使社会趋于稳定和谐。孟子提到“夫义，路也；礼，门也”（孟子·万章下），即要想成为君子必须经过礼的学习，将“礼”视为内心所坚信并奉行的原则，内化于内心、贯穿于人性，侧重将外在有形之“礼”内化为无形之“礼”。荀子继承并发扬孔子之“礼”，提出“礼者，人道之极也”（《荀子·礼论》），即懂得礼节是做人最高境界。

孔子主张“三人行，必有我师焉”（《论语·述而》），提出学而不厌、诲人不倦、以身作则、教学相长等为师之道，奠定儒家“尊师”传统。荀子竭力倡导“尊师敬师”，推崇教师不可取代的崇高地位和重要作用，并对“尊师敬师”做了专门论述，《荀子·致士》记录到“师术有四”，包括“尊严而惮”“耆艾而信”“诵说而不陵不犯”“知微而论”。可

见教师在传统教育中扮演着“知识掌握者”的身份角色，通过教学传播知识而获得社会尊敬。

3. 儒家智慧与教育数字化转型的现实契合性

“儒家教育思想是中国现代化的源头活水”。[3]儒家之智是教育之基，儒家智慧与教育的紧密联系是其得以继承和弘扬的客观载体和肥沃土壤，历经千年而生生不息、历久弥新。儒家智慧在数字教育中的运用打破时间、空间、学科等各种屏障限制，使千年前的思想结晶跨越千年融入21世纪的数字教育之中，为解决数字化所带来的新问题提供参考借鉴。“儒家传统要想真正得到有序深入的复兴，需要与社会层面的价值认同及各种关系紧密结合，并应对与现代社会制度互动所面临的各种张力问题”。[4]因此，为了使儒家智慧在当代数字教育趋势中能够得到重新解读与运用，需要站在数字时代的立场，以现代性视角去解读与重构儒家智慧与教育数字化转型的契合性，从而对现代教育改革提供新的方向和思路。

儒家“仁者爱人”的处世之道不仅适用于维系自我与他人的关系，还延伸至教育领域，涵盖师生、生生的新型平等关系。数字教育平等性体现在教师和学生角色的动态变化，数字化平台淡化了传统课堂中教师的绝对主导地位，打破传统班级中的等级体系，构建了新型的扁平化和平等化的师生、生生关系，这也是儒家“仁爱”思想在现代数字教育中的体现。

实体学校的传统线下课堂是模拟人与人多边交往的社会场域，而线上教学缺少班级同学间现实可触的互联互动，依托网线的远距离交流带有延迟性和隔阂性，对学生的社会交际能力有很大影响。为了促进数字教育发展，应寻求儒学之“礼”的传统智慧，重视教育对人的思想观念和行为体系的礼化塑造，促进学生社会交际能力的提升，使网络秩序趋于稳定和谐。

数字教育使学生能够借助网络平台进行主动探究，自我独立意识增强，不再“唯师命是从”，渴望尝试与教师进行平等对话和自由探讨，教师作为“知识掌握者”的主导地位被日渐消磨。然而，数字教育背景下的教师角色不止是单纯的“知识掌握者”，更是教学设计者、学习促进者、人际关系协调者等，教师应利用数字教育的特点，给予学生适当引导和帮助。“数字教育工具的开发也是为了改善师生间的沟通”。[5]个性化学习的可持续发展离不开教师的悉心指导。“教师扎根于儒家传承文化，鼓励学生转变学习模式”。[6]

4. 儒家教育思想在数字教育中的应用策略

4.1 渗透儒学思想内容，提高学生数字素养

儒家教育思想内涵深刻，有效纠偏数字教育“异化”风险，故应充分挖掘与放大仁爱、礼仪、尊师重教等思想的现代价值，将儒家思想精华整合并融于网络内容资源之中，丰富平台资源，促进教育公益普惠和优质均衡。利用技术传播培养学生成为更有道德感和责任感的人，使学生在潜移默化中习得善良品质、学会关心他人、提升品德素质，从“仁者爱人”的境界逐渐升华至“仁者爱物”的境界。还可以通过启发引导激发学生积极性，帮助学生学会利用数字平台进行自主学习，确保学生获取学习资源的公平性，实现主动探索和知识领悟，达成既定教学目标，强化相关技能，包括人机对话技能、协作技能等，提升高阶思维能力。

4.2 廓清儒学内在逻辑，回归教师主体地位

“其身正，不令而行；其身不正，虽令不从”（《论语·子路篇》）强调了教师应以身作则，以言传身教的方式成为学生模仿的榜样。教师需及时转换角色，成为学生个性化学习的促进者、协调者、引导者等身份，充分运用数字教育的特点，给予学生适当引导和帮助，满足学生的个性化学习需求，促进学生个性化学习和主体性成长。“传统儒家思想为教师提供了一种可供栖居的主体位置”。[9]依托于真正的“尊师重教”，教师的主体权和教育本体功能才能更好地回归，培养出高素质的综合性人才，实现教与学的“双赢”。“孔子和荀子主张通过对自我导向学习进行必要的干预和引导，使教师主动参与到学生的学习中”。[10]

教师需要投入时间和精力参加技能培训，掌握数字化教学技术、提升利用数字化资源的能力、丰富数字教育经验，选择、设计和实施适合学生需求的教育教学方案，为课堂教学效果和学生学习成果提供基础保障。还应着眼于提升自身素质，以广博知识、睿智智慧和高尚人格等品质为基础，通过知识传播开拓学生视野，通过沟通交流增加学生信任感，通过娴熟教学技巧进行因材施教，构建良好教师形象，重塑“尊师重教”思想，使学生真心“敬师”。

4.3 践行儒学本质要求，走向和合共生未来

创新社会交际模式。“儒家价值观常与社群主义联系在一起，这种意识形态主要在于平衡‘自我’与‘他者’的价值”。[7]因此，可以通过数字技术的深度加工与开发，突破传统教育中师生面对面交往模式，营造突破时空界限的在线交流大平台，构建虚拟社群学习环境，使每个人均可以根据自主需求而转变身份角色，进入不同课堂与不同交往对象进行沟通交流与互动探讨。虚拟社群学习环境打破信息孤岛，在人与人之间建立有机联系，充分实现人与人、人与网络之间的交互，形成“人机结合”的有效对话过程，使教师和学生在时间和空间上得以实现同步和异步交互的“零距离”，维持师生在数字教育中的平衡与和谐。

建立数字伦理规范。数字教育要求参与者以正确态度和技能使用数字工具和资源。“儒家礼仪为解决数字教育中微不足道但伤害较大的道德困境提供有效方式”。[8]由于学生群体涉世未深，易遭受网络不良言论的鼓动，造成难以估量的伤害，故可借助儒家之“礼”的有效教育手段，进行网络安全意识教育，使学生提升自我约束意识，形成符合儒家之“礼”的思维逻辑和行为方式，将数字伦理规范内化于心、化外于行，从而构建理想的数字教育秩序。

共促师生成长。教师应尊重学生独立人格和个性化需求，养成其自学能力和主体意识，发现其优势和个性潜能，发挥学生的积极性和主动性；还应关注师生互动的方式和过程，使双方在教学中各展所长、相得益彰，彰显儒家“尊师重教”的当代价值，实现教师尽心、学生进步的良性循环，构建教师主导、学生主体的和谐共生模式。

综上所述，随着教育数字化转型的持续深入推进，需要认识到其带来的机遇和挑战，并继续创新和改进教育实践和理念。儒家的核心原则具有历史和现代的价值，儒家智慧与数字教育的融合可以促进教育的创新和公平，为高质量教育的发展提供理论和实践指导。儒家智慧与数字教育的结合可以为教育的未来提供希望，并为推动社会公平和繁荣做出贡献。

参考文献：

[1]张茂泽.《礼记》的儒教思想[J].广西大学学报(哲学社会科学版),2017,39(01):16-23.

[2]盖立涛.墨家“新仁学”思想探析——以先秦儒、墨仁学比较为线索[J].孔子研究,2023(03):122-129.

[3]李立国.儒家教育思想的基本内容与当代价值[J].首都师范大学学报(社会科学版),2010(03):73-77.

[4]张晓艺,李向平.当代中国的儒教复兴:实践类型与内在逻辑[J].齐鲁学刊,2016(05):26-32.

[5]Almufarreh A ,Arshad M . Promising Emerging Technologies for Teaching and Learning: Recent Developments and Future Challenges[J]. Sustainability,2023,15(8).

[6]Mei YEC . Blended learning dilemma: Teacher education in the confucian heritage culture[J]. Australian Journal of Teacher Education (Online),2019,44(1).

[7]Tan C . For group, (f)or self: communitarianism, Confucianism and values education in Singapore[J]. The Curriculum Journal,2013,24(4).

[8]Olberding A . Etiquette: A Confucian Contribution to Moral Philosophy[J]. Ethics,2016,126(2).

[9]HUANG Hua,and Sou Kuan VONG."The Confucian Educational Philosophy and Experienced Teachers’ Resistance: A Narrative Study in Macau." Frontiers of Education in China 11.1(2016):1-22.

[10]Charlene T . Whither Teacher-Directed Learning? Freirean and Confucian Insights[J]. The Educational Forum,2018,82(4).

教师网络学习空间应用态度、应用行为、应用能力及应用效能关系的实证研究[i]

An empirical study on the relationship between teachers' applied attitude, applied behavior, applied ability and applied effectiveness in e-learning space

张钰团[1]*，周洋洋[2]，胡蓝心[3]，邵文豪[4]，任光杰[5]
[1]华南师范大学教育信息技术学院
* 1091807630@qq.com

【摘要】 中小学教师是网络学习空间应用的关键群体，是决定空间应用成效的核心变量，教师应用效能评价研究成为空间数字化转型与应用深化的重大课题。文章对广东省 12 个（县）区的中小学教师进行问卷调研，回收的 20994 份有效数据，用结构方程模型法分析应用效能与各要素作用关系和影响效应。研究表明，教师的空间应用行为对其应用能力具有直接正向影响；意愿态度对应用效能和应用行为具有直接正向影响；应用环境满意度对其意愿态度具有直接正向影响；应用能力对应用效能具有直接正向影响。
【关键词】 网络学习空间；应用效能；应用能力；应用行为；结构方程模型

Abstract: *Primary and secondary school teachers are the key groups in the application of e-learning space, and the evaluation of their application effectiveness is an important topic in digital transformation and application deepening. In this paper, a questionnaire survey was conducted among primary and secondary school teachers in 12 counties of Guangdong Province, and the relationship between application effectiveness and various factors was analyzed by using structural equation model. The results show that teachers' spatial application behavior, willingness and attitude, application environment satisfaction and application ability have direct influence on their application effectiveness. Finally, the paper puts forward some suggestions to improve the effectiveness of teacher space application.*

Keywords: Online learning space, Application efficiency, Application ability, Application behavior, Structural equation model

1. 问题的提出

网络学习空间已成为教师运用信息技术教学的主阵地，是实现信息技术与教学深度融合的重要环境支撑。国内外众多学者对此进行了深入研究，证实了网络教学效能与教学态度（Allinder，1994）和学业成绩（Brookove，1978）之间的紧密联系。近年来，国内研究逐渐深化，涉及空间使用意向的影响因素（赵呈领,周凤伶&蒋志辉，2017）及空间升级与学习方式转变（李玉斌,苏丹蕊&宋琳琳,2017）等多个方面。然而，在教学效能与其他变量的关系研究上，国内多采用传统方法，较少采用结构方程模型，直接相关分析也相对不足。鉴于此，本研究以广东省 12 个地市的中小学教师为对象，通过问卷调查和 Amos 软件建立结构方程模型，深入探究空间应用环境满意度、意愿态度、应用能力、应用行为与应用效能之间的复杂关联和相互作用机制，为优化网络学习空间提供理论支持。

2. 应用效能关系模型构建

李玉斌等（2016）指出教师网络学习空间应用行为与应用能力的一致性，提出应用行为影响应用能力的假设 H1。结合周榕（2017）和蔡宝来等（2008）的研究，教师的胜任力、教学

能力和意愿态度均影响应用效能，由此提出假设 H2 和 H3。Chiu（2005）的研究显示满意度影响使用态度和行为，故提出满意度影响意愿态度的假设 H4。Yang 等（2020）从交互视角发现互动对网络课程持续学习意愿的影响，由此提出教师使用态度决定应用行为的假设 H5。郭炯等（2017）指出教学环境满意度与应用行为可能存在相关关系，由此提出假设 H6。本研究基于这些假设勾勒出了如图 1 所示的假设模型。

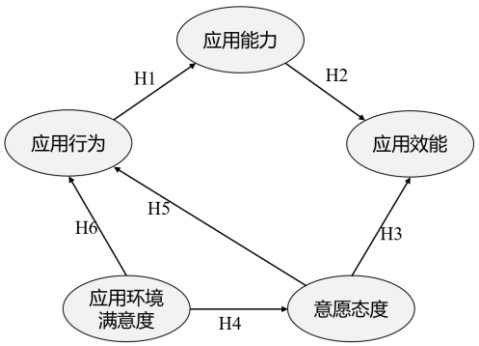


图 1 影响因素及关系的假设模型

3. 研究设计

本研究以广东省 12 个地市的中小学教师为研究对象，通过分层结合便利的抽样方法，选取不同学段的学校并覆盖 60%的教师样本，确保学科覆盖的全面性。研究采用结构方程建模方法，利用 SPSS26.0 和 AMOS26.0 工具处理分析调查问卷数据，以多个拟合度指标为参照来验证模型的有效性。基于文献调研和政策文件编制了“广东省网络学习空间建设及应用情况调研（教师卷）”，问卷包含客观题和李克特五级量表，经过多次修改和完善，最终包含七个方面 39 个题项，涵盖 7 个一级指标，详细见表 1。

表 1 问卷维度划分

一级指标	二级指标及内容	题目序号	题目类型
基本信息	性别、教龄、区县、职称、是否是优秀学校、是否是优秀教师、学段、学科	1-8	非量表题
应用能力	空间进行教学的能力评价	9	量表题
应用效能	人工智能教学应用的理解、教学效果、教学效率、空间应用情况评分、学生学业表现	10-14	量表题
	产出成果	15	非量表题
应用行为	资源服务、课堂教学、网络研修、班级建设、家校互动	16-18、20-21	量表题
	培训次数	19	非量表题
应用环境满意度	使用的信息化工具和平台、如何选择教学平台/工具、使用平台/工具完成哪些工作	22-25	非量表题
	网络响应速度、软硬件环境、空间功能、支持服务、国平台了解程度和使用情况	26-31	量表题
意愿态度	空间了解程度、使用动机、态度、持续使用意愿	32、34-36	量表题
	空间的理解	33	非量表题
需求与建议	应用支持	37	量表题
	未来学校学习空间建设中需提升的空间、建议	38-39	非量表题

4. 研究结果分析

4.1. 信效度检验分析

在信效度检验方面，本研究采用克隆巴赫系数和组合信度值来检测模型信度，总量表内部一致性信度为 0.907，各变量内部一致性信度系数均在 0.7 以上，充分证明了量表内部一致性较好，问卷信度较高。量表整体的 KMO 值为 0.973，满足因子分析的要求，Bartlett 球形

度检验的 P 值为 0，进一步验证了其适用性。通过 Amos26.0 进行效度检验，收敛效度方面，因子负荷量、组合信度值和平均抽取方差均符合规定值；区分效度方面，每个潜在变量的相关系数均小于平均方差抽取量的平方根，证明模型具有较好的区分效度。

4.2.结构方程模型分析结果

为确保模型拟合度效果的良好，本研究随机选取其中 500 份问卷数据进行结构方程模型，分析其关系。初始模型模拟后的结果显示，应用环境满意度指向空间应用行为的路径系数为-0.062，p=0.442，需要删除这条路径，同时增加残差路径对模型进行修正。修正后各路径 p 值均小于 0.05，具有统计学意义，最终模型如图 2 所示。

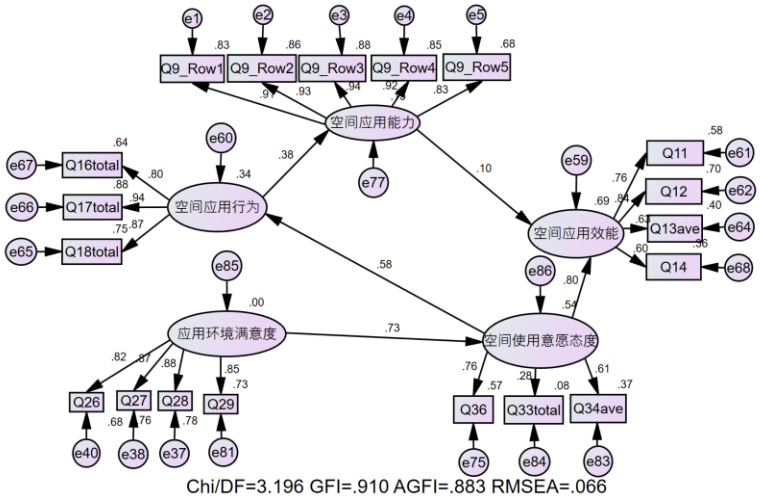


图 2 最终模型

模型的卡方自由度比(CMIN/DF)=3.196、GFI=0.910、AGFI=0.883、RMSEA=0.066、RMR=0.106、NFI=0.935、RFI=0.924、IFI=0.954、CFI=0.954、TLI=0.946，各项拟合指数均优于临界值，因此，可以认为本研究的测量模型具有较好的拟合度。

本研究分析了修正后的结构方程模型潜变量之间的作用关系，研究假设的路径系数和验证情况见表 2。假设 1 验证表明应用行为对应用能力有直接正向影响；说明中小学教师使用学习空间开展教学的行为，能有效增强教师的空间应用能力。假设 2 验证表明应用能力对应用效能有直接正向影响；说明中小学教师使用学习空间能力的强弱，能有效提升教师的空间应用效能。假设 3 验证表明意愿态度对应用效能有直接正向影响；说明教师对空间开展教学的意愿态度会直接影响教师的应用效能。假设 4 验证表明应用环境满意度对使用意愿态度有直接正向影响；说明教师对空间应用环境的满意度会直接影响教师的空间使用意愿态度。假设 5 验证表明意愿态度对应用行为有直接正向影响；说明教师对空间开展教学的意愿态度会直接影响教师的应用行为。假设 6 验证表明应用环境满意度对空间应用行为没有直接正向影响。

表 2 各影响路径系数及验证假设情况

假设	路径	标准化系数	结果
H1	空间应用行为——>空间应用能力	0.382***	成立
H2	空间应用能力——>空间应用效能	0.101（0.009）	成立
H3	空间使用意愿态度——>空间应用效能	0.805***	成立
H4	应用环境满意度——>空间使用意愿态度	0.734***	成立
H5	空间使用意愿态度——>空间应用行为	0.583***	成立
H6	应用环境满意度——>空间应用行为	-0.062（0.442）	不成立

5.研究结论与建议

中小学教师网络学习空间应用效能关系的实证研究表明：第一，使用意愿态度正向影响应

用效能和应用行为。通过鼓励教师参与空间建设，将个性化资源和平台融入其中，增强教师的参与感和归属感，进而提升应用效能。同时，加大空间应用的推广力度，激发教师的主观能动性，使空间从“可用”向“好用”“爱用”“常用”转变。第二，应用环境满意度正向影响使用意愿态度。我们应引进先进技术，提升数字化教学装备配置水平，实现多媒体教学设备的全面覆盖。同时，发挥国家平台的协同赋能优势，避免重复建设，全方位提升参与主体的数字化能力。建立并创新线上线下融合的资源服务机制，从而激发教师持续使用空间的意愿。第三，空间应用行为正向影响应用能力，应用能力正向影响应用效能。空间应用行为是衡量常态化使用的重要指标，通过实践中的不断创新，教师的应用能力将得到提升，进而增强应用效能。在双新双减背景下，我们还需将教师数字素养和空间应用能力纳入培训内容，构建以校为本、应用驱动的教师数字素养发展机制，实现教师空间应用效能的有效提升。

6. 结语

本研究通过实证探究了中小学教师利用空间开展教学的能力、行为、意愿态度、应用环境满意度和应用效能之间的内在关系，研究方法和结论可为相关学者研究和中小学教师提升空间教学效能、促进教师专业发展和空间转型升级提供重要参考。目前缺乏基于平台数据的效能研究及评价，这是未来研究的重要方向，对数字化转型和教育高质量发展具有重要意义。

参考文献：

李玉斌,苏丹蕊&宋琳琳.(2017).网络学习空间升级与学习方式转变.***现代远距离教育***(04),12-18.

李玉斌,李爽,武亚男&姚巧红.(2017).教师网络学习空间应用能力调查研究.***中国教育信息化***(16),64-68.

周榕.(2017).高校教师远程教学胜任力培训设计模型构建——基于复杂学习的视角.***电化教育研究***(06),116-122.

赵呈领,周凤伶&蒋志辉.(2017).TAM 与 SDT 视角下网络学习空间使用意向之影响因素研究.***现代远距离教育***(05),3-11.

郭炯,郑晓俊&黄彬.(2017).网络学习空间支持的协同教学模式与应用案例研究——网络学习空间内涵与学校教育发展研究之八.***电化教育研究***(10),23-29.

郭炯&郑晓俊.(2019).基础教育领域网络学习空间建设与应用现状、问题及对策研究.***电化教育研究***(10),89-95.

焦立涛，梁忠环&孟辉.(2016).教师满意度模型初探.***中国教育学刊***(S1),136-139.

蔡宝来&车伟艳.(2008).国外教师课堂教学行为研究：热点问题及未来趋向.***课程.教材.教法***(12),82-87.

Allinder R M. (1994).The relationship between efficacy and the instructional practices of special education teachers and consultants. ***Teacher education and special education***(2), 86-95.

Brookover W B, Schweitzer J H, Schneider J M, et al.(1978).Elementary school social climate and school achievement. ***American educational research journal***(2),301-318.

Chiu C M, Hsu M H, Sun S Y, et al. (2005).Usability, quality, value and e-learning continuance decisions. ***Computers & education***(4),399-416.

Yang, H., Wang, M., Ju, Y., Wen, X., Wu, M., & Han, Q. (2020). Research on Influencing Factors of Continuous Learning Willingness for Online Courses From the Perspective of Interaction. ***2020 International Conference on Information Science and Education*** (ICISE-IE), 342–346.

[i] 本文系广东省哲学社会科学规划2023年度一般项目“数字化转型背景下网络学习空间应用效能动态评估模型构建及应用——以广东省为例”（批准号：GD23CJY20）和华南师范大学科研立项一般课题培育项目“数智驱动的教师教学效能评估研究”（编号：23JXGA15）成果之一。

人工智能赋能课堂教学评价研究：研究热点与发展趋势

Artificial Intelligence Enabled Classroom Teaching Assessment: Research Hot Spots and Development Trends

易诗情[1*]，胡蓝心[2]，周洋洋[3]，王韵仪[4]
[1234]华南师范大学教育信息技术学院
[*] ysqing214@163.com

【摘要】为了探究人工智能赋能课堂教学评价的研究进展和热点前沿，本研究以 VOS viewer 为研究工具，综合采用文献计量法和内容分析法，对 Web of Science 数据库中近十年的相关文献进行系统处理和分析，提炼总结人工智能在课堂教学评价领域的概况、热点和发展趋势。研究发现，人工智能赋能课堂教学评价领域在近几年呈快速增长趋势。研究重点主要集中在学业表现预测、智能化反馈、教学质量分析、评价系统构建等方面。未来，研究将更加注重综合评价的探索和生成式人工智能的赋能，以推动人工智能在课堂教学评价领域的进一步应用与创新。

【关键词】人工智能；课堂教学评价；综合评价；生成式人工智能

Abstract: *In order to explore the research progress and hot frontiers of artificial intelligence-enabled classroom teaching evaluation, this study uses VOS viewer as the research tool, and comprehensively uses the bibliometric method and content analysis method to systematically process and analyze the relevant literature in the Web of Science database in the past ten years, and summarizes the general situation, hot spots and development trends of artificial intelligence in the field of classroom teaching evaluation. The research finds that the field of artificial intelligence-enabled classroom teaching evaluation has shown a rapid growth trend in recent years. The research focuses on academic performance prediction, intelligent feedback, teaching quality analysis, evaluation system construction and so on. In the future, research will pay more attention to the exploration of comprehensive evaluation and the empowerment of generative artificial intelligence, so as to promote the further application and innovation of artificial intelligence in the field of classroom teaching evaluation.*

Keywords: Artificial Intelligence, Classroom Assessment, Comprehensive Evaluation, Generative artificial intelligence

1. 引言

课堂教学评价是对教师教学和学生学习过程及效果的评估，在监督和提高教育质量方面扮演着日益重要的角色。然而，目前课堂教学评价方法仍存在诸多不足，如主观性强、效率低、缺乏个性化等问题，导致对学生掌握课堂教学技能的情况缺乏及时预测和分析(Chen & Dong, 2022)。随着人工智能（AI）的快速发展，智慧教育理念不断完善并逐渐渗透到教育各个方面，人工智能在课堂教学评价中的应用逐渐引起关注，为解决传统评价方法的局限性提供了新的解决途径。

进入智能时代，应用智能技术促进课堂教学评价转型升级是非常必要且可行的。面对国际上人工智能赋能课堂教学评价的众多研究成果，本研究以 Web of Science 为文献数据源，采用文献计量法，对相关文献进行年度发文量、关键词共现等分析，以全面、系统地把握人工智能赋能课堂教学评价研究的热点与发展趋势。

2. 研究设计

2.1. 样本选取

本研究的文献样本源为 Web of Science 核心合集中的“社会科学引文索引”和“科学引文索引”，检索主题包含“Artificial Intelligence Or AI”And“Classroom Teach* Evaluate* Or Classroom Assess*”的英文论文英文论文，文献筛选时间为 2013 年 1 月 1 日至 2023 年 12 月 31 日，共得到 909 篇文献。根据 2020 年更新的《系统综述和元分析的优先报告项目》PRISMA 2020）(Page , 2021)进行文献数据筛选，最终得到与主题相关的文献 812 篇。

2.2. 研究方法

本研究聚焦人工智能赋能课堂教学评价研究的重要议题和发展趋势,综合采用文献计量法与内容分析法对数据进行可视化分析。文献计量分析基于 VOSviewer 1.6.19 绘制研究的关键词共现图谱和时序图谱，揭示重要术语的分布与关系特征，探测研究热点与发展方向。内容分析则基于文献的题目和摘要，深入解读文献内容，以达到对研究主题的深层次理解。

3. 人工智能赋能课堂教学评价研究热点

从时间维度来看，AI 赋能课堂教学评价研究的发展在近三年呈快速增长趋势，出现此种形势，与近几年教育信息化的战略推进和智能感知技术、学习分析技术、情感计算技术等 AI 技术的深入发展息息相关(Dai, 2021)。从内容维度来看，本研究基于关键词共现图(如图 1 所示)，分析得出 AI 赋能课堂教学评价研究的热点,主要包括提升学生素养的学业表现分析、基于智能反馈的质量评价、智能技术支持下的评价模型、基于智能算法的评价系统四个方面。

3.1. 基于智能诊断的学业表现预测研究

“学业表现”、“绩效”、“智能诊断”、“分类”等词在关键词共现分析中高频出现，累计达 157 次，揭示基于智能诊断的学业表现预测在 AI 赋能课堂教学评价领域中的显著地位。

近年来，大数据、机器学习、深度学习等技术的迅速发展，课堂教学评价呈现出更为实时、智能和可视化的特点。大数据记录学生学习行为，为动态诊断和评估提供科学依据（Rong, 2022）。通过深度学习模型，对学生的学习行为进行分类，并预测其学业成就，更好地践行“以人为本”、关注学生成长。随着人工智能技术的发展，涌现出一系列提升评价维度和深度的教学评估方法，这些方法包含基于情感识别的学生情感分析、学习行为预测模型的构建等。如 Cheng & Wang（2021）构建了多维度的课堂教学行为分析框架，Zhu 等人（2023）提出了一种 NAGNet 的学习者情绪识别方法，用于辅助教师根据智能诊断改进教学方法。

3.1. 基于智能反馈的教学质量分析研究

“智能反馈”、“评价质量”、“可解释人工智能”等术语在关键词共现分析中频繁出现，累计达 187 次，凸显基于智能反馈的教学质量评价成为 AI 赋能课堂教学评价领域的热点之一。

利用人工智能技术，通过设计创新反馈机制来提高教学质量评价，推动学生的认知发展和技能提升。当前，教学质量评价面临着诸多挑战，包括评价标准不统一、评估结果主观性较高、以及教学质量监控缺乏客观评价依据等问题。近年来，一些学者提出了基于 AI 技术的设计性评价方法，针对课堂教学进行深入研究，从而有效地解决了评价精度低、评价方法不完善和系统成本高等问题(Chao & Zhang, 2022)。为解决传统评价的局限性，有研究者探索基于 AI 的智能评价方法，如模糊数学和微分建模，实现对预测数据进行智能评价和分类(Lv & Shen, 2021)。研究亦关注作业质量评价，如 Mechanix 系统能够准确判断学生作业的正确与否，并给予即时的、富有建设性的个性化反馈，帮助学生改进学习成果。未来研究需解决主观与客观评价的结合问题，关注智能评价的可解释性，以促进教育教学的持续创新。

3.1. 基于智能算法的评价系统构建研究

“系统”、“算法”、“机器学习”、“神经网络”、“深度学习”等词在关键词共现分析中高频出现，累计达 310 次，揭示基于智能算法的评价系统构建是 AI 赋能课堂教学评价研究的热点之一。

近年来，人工智能在教育领域的应用引起了广泛关注，特别是在教育教学中引入了行为检测和智能识别算法(Z. Zheng , 2022)。该领域的研究集中于开发和优化利用机器学习、深度学习等算法的评价工具，以实现对学习过程和成果的高效评估。比如，自动评价系统通过自然语言处理等智能技术进行作业评分，智能辅助系统提供定制化学习支持 (Sun , 2021)。研究还涵盖智能分析模型的构建，如 Wu 等人（2023）以行为过程中学习者情绪状态的评价为例，构建了学习者行为过程中情绪状态的智能分析模型，为构建跨模态学习分析提供了有效的技术解决方案。研究进一步扩展至课堂管理，通过神经网络和情感识别算法构建智能分析系统，以及基于深度学习的课堂评分系统，这些系统针对不同教学场景进行定制，以提升教学效率和质量(Wang & Huang, 2022)。未来研究将深化智能评价系统的设计和实施，关注评价精确度和可靠性的提升，探索算法和神经网络技术在评价流程优化中的应用。

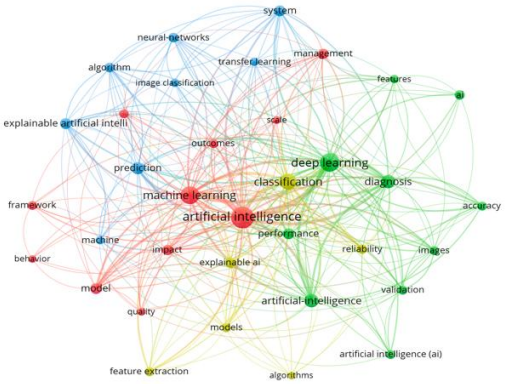


图 1 人工智能赋能课堂教学评价研究的关键词共现分析

4. 人工智能赋能课堂教学评价研究未来趋势

AI 赋能课堂教学评价研究数量不断增加，涵盖的领域规模也在不断扩大，研究视角和维度也呈现多元化的趋势。内容上更趋向于综合评价，方式上生成式人工智能的兴起或成为新的赋能点。

综合评价拓展了课堂教学评价的维度和深度，在 AI 赋能课堂教学评价研究中显得愈发重要，自 2020 年以来相关研究数量不断增加。该评价体系超越了单一考试成绩的局限，通过综合分析学生的作业、参与度、合作与创新能力等多元信息，为学生提供全面的学习表现评估。它不仅关注学生的认知学习成果，也重视跨学科项目中的非认知能力发展。通过 AI 技术对学生在团队合作、问题解决和创新思维等方面的表现进行更为细致和客观的评价。这种智能化的综合评价不仅帮助教师更好地理解学生的综合能力，也为学生提供更具针对性的发展建议。

在这一背景下，生成式人工智能在课堂教学评价中展现出显著潜力，其可用于辅助教师生成课堂评价量表，减轻教师在评价设计过程上的负担。此外，可解释的生成式人工智能基于学生的课堂表现和学习成果提供个性化反馈，助力学生识别并强化优势，提供精准的指导和支持。随着技术的成熟，生成式 AI 将进一步推动教育评价的创新，优化教学质量。

5. 结论

当探讨人工智能赋能课堂教学评价时，我们面临着多个维度的研究，涵盖智能系统的反馈与评价、个性化评价活动设计与实施、智慧课堂的智能测评以及沉浸式教学与多维评估等。然而，这一领域依然存在一系列值得深入探讨的争议与挑战。首要问题在于平衡人工智能评

价与教师专业判断之间的关系，如何使技术评价与教师的专业判断相辅相成成为一个迫切需要解决的问题。其次，数据隐私、算法公正性以及可解释性等问题也亟需解决。尽管存在挑战，AI 技术为提升课堂教学评价的客观性、准确性和个性化提供了新契机，也为教育评价领域的未来发展带来了前所未有的机遇。

参考文献

Chao, J., & Zhang, Y. (2022). Analysis of the Current Situation of Teaching and Learning of Ideological and Political Theory Courses by Deep Learning. COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, 2022.

Chen, Q., & Dong, M. (2022). Design of Assessment Judging Model for Physical Education Professional Skills Course Based on Convolutional Neural Network and Few-Shot Learning. COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, 2022.

Cheng, J., & Wang, X. (2021). Artificial intelligence based on effectiveness of inverted classroom teaching of college sports. JOURNAL OF INTELLIGENT & FUZZY SYSTEMS, 40(2), 3755–3765.

Dai, D. (2021). Artificial Intelligence Technology Assisted Music Teaching Design. SCIENTIFIC PROGRAMMING, 2021.

Lv, Z., & Shen, H. (2021). Artificial intelligence with fuzzy logic system for learning management evaluation in higher educational systems. JOURNAL OF INTELLIGENT & FUZZY SYSTEMS, 40(2), 3501–3511.

Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hrobjartsson, A., Lalu, M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ-BRITISH MEDICAL JOURNAL, 372.

Rong, J. (2022). Innovative Research on Intelligent Classroom Teaching Mode in the 《5G》 Era. MOBILE INFORMATION SYSTEMS, 2022.

Sun, Z., Anbarasan, M., & Kumar, D. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. COMPUTATIONAL INTELLIGENCE, 37(3), 1166–1180.

Wang, P., & Huang, X. (2022). An Online English Writing Evaluation System Using Deep Learning Algorithm. MOBILE INFORMATION SYSTEMS, 2022.

Wu, WY., Hu, QT., Feng, G.,& He, YX. (2023).A Novel Method for Cross-Modal Collaborative Analysis and Evaluation in the Intelligence Era. STEAM EDUCATION AND THE INNOVATIVE PEDAGOGIES IN THE INTELLIGENCE ERA,13(1).

Zhu, HH., Hu, PY., Tang, XP., Xia, DX.,& Huang, HF.(2023). NAGNet: A novel framework for real-time students' sentiment analysis in the wisdom classroom.CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE,35(22).

Zheng, Z., Liang, G., Luo, H., & Yin, H. (2022). Attention assessment based on multi-view classroom behaviour recognition. IET COMPUTER VISION.

基于 TAM 模型的中学生大语言模型平台使用意愿影响因素分析

A Study on the Use Intention of the Large Language Model Platform in Middle School Stage

许其鑫[1]，梁燕[2]，曾诗娅[1]，李明蔚[3]
[1]宜宾学院 教育学部
[2]四川省航天中学校
[3]华中师范大学 人工智能教育学部

【摘要】 人工智能大语言模型被广泛应用于中学教育教学领域，如帮助学生生成优质的作文、阅读材料等。为探索中学生使用大语言模型平台辅助自身学习的持续意愿，研究基于拓展的 TAM 模型，对四川成都某校 608 名有使用经验的中学生进行调研，并构建模型进行验证。结果表明，促成条件对感知易用性和感知趣味性具有显著影响，然而促成条件对感知有用性无显著作用。此外，感知有用性、感知易用性和感知趣味性对大语言模型的使用意愿存在显著的正向影响，共同解释了使用意愿 54.6%的方差。这将为相关决策者、教育实施者和技术开发者提供有益的参考，并为中学生与大语言模型平台的有效互动搭建桥梁。

【关键词】 大语言模型；中学生；TAM 模型；使用意愿

Abstract: *Artificial intelligence large language model is widely used in the middle school education, such as helping students generate high-quality composition and reading materials. To explore the sustained willingness of middle school students to use the large language model platform to assist their own learning, a study was conducted on 608 experienced middle school students from a school in Chengdu, Sichuan based on an extended TAM model, and a model was constructed for validation. The results indicate that facilitation conditions have a significant impact on perceived ease of use and perceived fun, while facilitation conditions have no significant effect on perceived usefulness. In addition, perceived usefulness, perceived ease of use, and perceived interest have a significant positive impact on the willingness to use the large language model, jointly explaining 54.6% of the variance in willingness to use. This will provide useful references for relevant decision-makers, educational implementers, and technology developers, and build a bridge for effective interaction between middle school students and the big language model platform.*

Keywords: *Large language model; Middle school students; TAM model; Willingness to use*

1. 引言

随着人工智能技术的快速发展，大语言模型作为自然语言处理领域的重要研究方向，受到了广泛的关注。目前，许多机构和企业已经开发出各种大语言模型，并在学术研究和实际应用中取得了一定的成果(Menon & Shilpa, 2023)。大语言模型，也称之为大型预训练语言模型，是指拥有庞大语料库训练而成的，融合了各种语言知识和语言规律的人工智能模型(Zhang C et al, 2023）。它具备对自然语言进行理解、生成和处理的能

力，并能够在各种任务中表现出相当高的水平，为人们提供了一种更加智能、高效的语言交互方式，给信息检索、智能客服、智能助手等领域带来了广阔的应用前景，同时也为教育教学变革带来了前所未有的机遇（刘明等，2023）。本文选择讯飞星火大模型作为代表，是因为它在七大能力维度上展现出了强大的智能涌现水平，免费对外开放，且易于操作(刘聪等，2023）。目前，中学生对大语言模型使用意愿方面的研究还相对较少，且现有问题和争议主要有两个方面，一方面，虽然市面上已经存在一些教育类的大语言模型平台，但是由于其针对教育需求的特殊性质，与普通大语言模型平台存在着一定的差异。另一方面，对于中学生大语言模型平台使用意愿的影响因素，如用户体验、功能完善性、实用性等方面仍然存在不同的观点和争议。如何通过提供更符合中学生学习需求的大语言模型平台，来增强他们的使用意愿，是一个亟待解决的问题。本研究提出了一个拓展的 TAM 模型来解释和预测中学生对大语言模型的使用意愿，这将为相关决策者、教育实施者和技术开发者提供有益的参考，助力中学生与大语言模型平台的有效互动。

2. 相关研究

Davis 以行为理论为基础首次提出了技术接受模型(Technology Acceptance Model，TAM）（Davis，1986）。作为一种被普遍接受的行为预测模型，TAM 提出了感知有用性和感知易用性对使用态度和行为意向的重要作用。李珩以 TAM 为理论基础，通过建构大学英语线上学习行为结构方程模型，探讨大学生线上英语学习行为特征（李珩，2023）。梁宇和李诺恩基于 TAM 拓展模型，对中文数字学习资源使用意愿及其影响因素进行了研究，并发现感知易用性对使用意愿的总影响效应最为突出(梁宇和李诺恩，2023)。毕达天等人从感知公平性、感知可问责性和感知透明度三个方面探究不同维度的算法感知对用户接受意愿的差异化影响及其作用机制，并采用结构方程建模进行实证研究（毕达天等，2023)。同时，徐月梅等人针对大语言模型（LLM）技术的快速发展，剖析其技术应用前景及风险挑战，指出大语言模型为知识的调用方式、获取手段打开了新的途径，为人民的生活带来了巨大便利，在教育领域为个性化数字资源创建、人机协同学习、实时教学辅导等也提供了有力支撑（徐月梅等，2023）。陈舒梦研究了大语言模型在外语教学中的应用，详细介绍了其在外语教学中的应用场景(陈舒梦，2023）。刘明等人梳理了大语言模型在高阶思维培养、加强阅读理解能力以及提升写作与数学解题水平等教学应用场景中的优势，也指出大语言模型应用于教育领域存在的相应挑战，如教育领域大语言模型训练成本高，师生对问题、写作文本和教学内容生成过度依赖以及生成内容存在伦理和知识产权争议等问题(刘明等，2023）。目前，关于大语言模型在中学的应用实践研究较少，本研究旨在探究中学生对大语言模型平台的使用意愿。通过分析促成条件、感知有用性、感知易用性对使用行为意愿的影响，可以为中学教育领域的大语言模型平台设计和推广提供参考。

3. 理论框架和研究假设

通过分析现有关于大语言模型的文献，我们发现其中实证研究较少，特别是在中学生阶段的研究较为匮乏。研究中学生对大语言模型的使用意愿影响因素，探究各因素对学生使用大语言模型的意愿影响大小以及因素之间的作用机制，对扩大大语言模型的教育应用场景和范围，促进教育教学变革具有非常重要的影响。因此，本研究基于 TAM 扩展模型，引入了感知趣味性和促成条件两个对平台使用意愿可能产生影响的因素，提出了中学生对大语言模型平台使用意愿的假设模型（见图 1）。

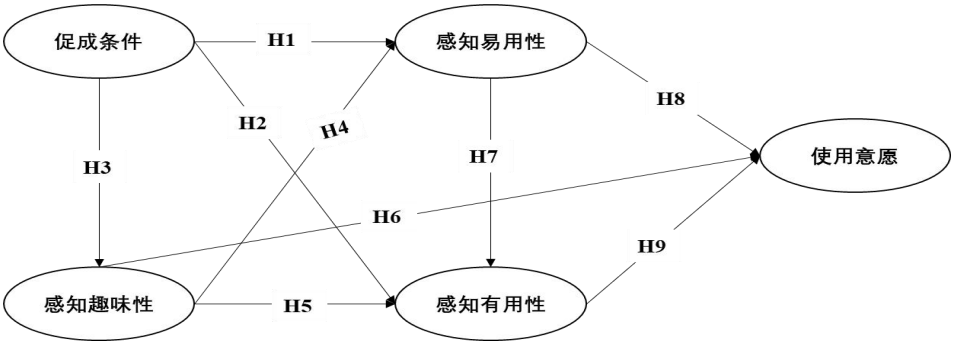


图 1 中学生大语言模型使用意愿理论模型

3.1 促成条件

促成条件是指个体采取某种特定行为具备的客观有利环境，如组织条件、技术条件等（Venkatesh & Bala, 2008）。TAM 扩展模型的相关研究表明，客观条件的便利程度直接影响学习者的感知易用性和感知有用性（Teo, 2011）。其他人的研究也证实了这一结论(Jovanka et al., 2023; Li et al., 2021)。促成条件对感知趣味性也存在显著的正向影响(Won et al., 2023)，在大语言模型平台使用过程中，如果缺乏良好的网络环境和智能设备等外部条件会对学生使用大语言模型平台的趣味性和便捷性造成一定的障碍。因此，本研究拟调查中学生所处的网络环境和智能设备等便利程度。围绕促成条件，本研究提出以下相关假设：

H1：外部条件正向影响大语言模型平台感知易用性。

H2：促成条件正向影响大语言模型平台感知有用性。

H3：促成条件正向影响大语言模型平台感知趣味性。

3.2 感知趣味性

感知趣味性是指用户使用相关系统平台时令人愉快的程度(Venkatesh & Speier, 1999)。前人研究结果表明，用户对系统的感知有用性和感知易用性会受到感知趣味性的直接影响(Jovanka et al., 2023; Liu et al., 2022; Won et al., 2023)。研究结果还发现，感知趣味性与平台的使用意愿呈正相关关系(Bahçekapılı, 2023)。大语言模型平台对中学生的吸引程度直接影响了其对平台的持续使用意愿，如果平台拥有多样的趣味功能，将会吸引学生探索该平台，从而提升学生对平台的感知有用性和易用性。据此，可提出如下假设：

H4：感知趣味性正向影响大语言模型平台感知易用性。

H5：感知趣味性正向影响大语言模型平台感知有用性。

H6：感知趣味性正向影响大语言模型平台使用意愿。

3.3 感知易用性

感知易用性是指个体主观认为使用某一信息系统的省力程度(Davis, 1989)，本研究对其定义为学习者使用大语言模型平台时主观感受到的难易程度。Jovanka et al.(2023)对印尼开放大学的 2787 名学生使用电子学习产品的意愿调查发现，感知易用性在预测感知有用性和电子学习服务方面具有重要作用。多位学者的研究也证实了该结论(Al Shamsi et al., 2022; Al-Adwan et al., 2023; Hsu & Lin, 2022)。中学生愿意使用大语言模型平台的前提是该平台界面简单，易于操作，在使用的过程中，学生才能逐步从平台中获取更多的学习知识，提升其学习效率。据此，以下假设被提出：

H7：感知易用性正向影响大语言模型平台感知有用性。

H8：感知易用性正向影响大语言模型平台使用意愿。

3.4 感知有用性

感知有用性是指个体主观认为使用某一信息系统对工作绩效的提升程度(Davis, 1989)，因此，本研究对其定义为学习者大语言模型平台时主观感受到的学习效果。感知有用性被认为是接受 IT 最重要的因素(Basak et al., 2015)。当用户打算学习新技术时，其使用意图会直接受到其实用性的影响(Al Shamsi et al., 2022)。现有的文献也支持了感知有用性对使用意愿具有积极正向作用的观点(George et al., 2023; Hsu & Lin, 2022; Liu et al., 2022)。对于中学生而言，当他们从大语言模型平台中学习到越多有用的知识时，他们会更愿意使用。因此，可以提出假设：

H9：感知有用性正向影响大语言模型平台使用意愿。

4. 研究设计

4.1 研究工具设计

本研究采用的工具由两部分构成，分别是调查对象基本情况（2 道题目）和量表部分（共 19 道题目），涵盖了 7 个维度。19 个量表题项均从已有问卷中改变而来，使其符合当前的实际研究要求。题项和参考文献见表 1。本研究采用李克特七级量表，从 1 到 7 分别代表“非常不符合”到“非常符合”。本研究数据采集均在中国进行，为了确保问卷内容的准确传达，我们邀请了专业人员对问卷进行翻译，并反复对中英文互译进行比对，以最大程度地消除中英文内容差异。

表 1 测量变量和题项

变量	因子	题项	参考来源
感知有用性 (PU)	PU1	使用该平台可以提升我的学习效率	(Al-Adwan et al., 2023)
	PU2	使用该平台可以让我更快完成任务	(Venkatesh & Davis, 2000)
	PU3	使用该平台可以减轻我的任务负担	
	PU4	使用该平台可以让我完成任务更容易	
	PU5	使用该平台可以更快更精确的帮我找到所需要的资料	
感知易用性 (PEU)	PEU1	该平台的互动性设计的很好	(Venkatesh & Davis, 2000)
	PEU2	我可以很好的和该平台互动	(Al Shamsi et al., 2022)
	PEU3	我能很容易的在该平台上找到想要的内容	
	PEU4	我认为熟练的使用该平台非常容易	
促成条件 (FC)	FC1	我有电脑或者手机可以让我方便的使用该平台	(Al Shamsi et al., 2022)
	FC2	我拥有使用该平台相对应的知识与技能	
	FC3	我可以很容易的学会该平台的使用教程	
感知趣味性 (PEN)	PEN1	我认为该平台让我的学习生活增加了很多乐趣	(Al-Adwan et al., 2023)
	PEN2	我喜欢使用该平台	(Bahçekapılı, 2023)
	PEN3	使用该平台能让我产生愉悦感	
	PEN4	该平台对我来说具有很大的吸引力	
使用意愿 (UI)	UI1	我经常使用类似的平台	(Venkatesh & Davis, 2000)
	UI2	我会将该平台推荐给其他人	(Al Shamsi et al., 2022)
	UI3	未来我将经常使用该平台	

4.2 样本选择与问卷发放

本研究以中国四川省的初高中学生为研究对象,通过问卷星平台和纸质问卷相结合的方式进行调研和数据搜集。在调查前，需要教师用 3 周的时间带领学生学习和使用讯飞星火认知大模型。在学生均已体验过该平台后，再发放问卷。预调研于 10 月 11 日到 10

月 18 日进行，通过纸质问卷的方式随机抽取 100 位学生进行分析。通过搜集用户的数据及在调研过程中收到的用户反馈对问卷初稿进行了修订。正式调研则在 11 月 13 到 12 月 1 日进行。共计发放 751 份问卷，剔除了答题时间小于 50 秒且结果选项完全一致的问卷，回收有效问卷 608 份，有效率为 80.96%。其中男生和女生的占比分别为：35.7%和 64.3%，初中和高中的占比分别为 41.3%和 58.7%。

4.3 数据处理

每个构念的描述性统计见表 2。所有构念的均值都在 5 以上，标准差均在 1.05-1.41 之间，偏度(从−1.43 到−0.81)和峰度(从 0.34 到 2.85)的值都在 Kline 建议的可接受范围内(Kline and Rex, 2011)。因此，本数据满足正态分布。为保证所收集数据的有效性和科学性，本研究采用 SPSS 和 AMOS 软件对问卷的信效度进行分析，并建立了大语言模型平台使用意愿的测量模型和结构模型以便估计模型的拟合度和检验本研究提出的 9 个假设。

表 2 变量描述性统计分析

变量	项目	平均值	标准差	偏度	峰度
感知有用性	5	5.99	1.13	-1.43	2.85
感知易用性	4	5.99	1.05	-1.07	1.34
促成条件	3	5.87	1.16	-1.18	1.71
感知趣味性	4	5.88	1.13	-1.06	1.21
使用意愿	3	5.39	1.41	-0.81	0.34

5. 研究结果与分析

5.1 信效度分析

信度主要衡量结果的一致性和稳定性，由表 3 可知，各个变量的克朗巴赫 a 系数和组合信度（Composite Reliability, CR）值都大于 0.75，且该量表的整体克朗巴赫 a 系数为 0.936。表明本量表的测量题项间具有较高的一致性，各题项能有效反映各个潜变量的特征，说明本研究的测量模型具有很好的信度。效度主要评估测量结果是否真实地反映了被测对象的特性。在本研究中，我们通过计算各个变量的平均抽取方差（Average Variance Extracted, AVE）来检验测量模型的聚敛效度。结果显示，所有变量的平均抽取方差均大于 0.5，这意味着测量模型具有较好的聚敛效度。此外，计算结果发现，各个变量的平均抽取方差的平方根大于该变量与其他变量的相关系数，这表明测量模型具有较好的区分效度。

表 3 问卷的聚敛效度和区分效度分析

	a	CR	AVE	1	2	3	4	5
1.感知有用性	0.889	0.891	0.621	0.788				
2.感知易用性	0.885	0.887	0.663	0.698***	0.814			
3.促成条件	0.812	0.815	0.595	0.562***	0.698***	0.772		
4.感知趣味性	0.891	0.891	0.672	0.642***	0.739***	0.745***	0.82	
5.使用意愿	0.756	0.764	0.523	0.613***	0.654***	0.559***	0.695***	0.723

5.2 共同方法偏差检验

共同方法偏差是指因为同样的数据来源或评分者、同样的测量环境、项目语境以及项

目本身特征所造成的预测变量与效标变量之间人为的共变。本研究采用 Harman 单因素检验结果显示，最大因子方差解释率为 48.088%<50%(Podsakoff et al., 2003)；其次，进行单因子的验证性因子分析，发现单因子模型的拟合效果很差：χ2/df=11.661，SRMR=0.078，RMSEA=0.133，GFI=0.712；NFI=0.752，TLI=0.738，CFI=0.768。综上，本研究不存在严重的共同方法偏差。

5.3 假设检验

为了分析测量模型的拟合情况，采用了卡方自由度（χ2/df）、标准化均方根残差（SRMR）、近似均方根误差（RMSEA）、拟合优度指数（GFI）、标准化拟合指数（NFI）、Tucker-Lewis 指数（TLI）和比较拟合指数（CFI)等拟合指标。根据 Hair et al. (Hair, Page, and Brunsveld, 2019)，所有指标均在可接受范围内，证明测量模型与样本数据的拟合程度可以接受。

表 4. 模型拟合汇总

拟合指标	参照值	测量模型	结构模型
卡方自由度(χ^2/df)	<5	2.521	2.504
标准化均方根残差(SRMR)	<0.08	0.0339	0.0468
近似均方根误差(RMSEA)	<0.08	0.050	0.050
拟合优度指数(GFI)	>0.8	0.940	0.940
标准化拟合指数(NFI)	>0.9	0.950	0.950
Tucker-Lewis 指数(TLI)	>0.9	0.963	0.963
比较拟合指数(CFI)	>0.9	0.969	0.969

同理，使用 AMOS 软件绘制结构方程模型，并导入问卷数据进行模型检验。由表 4 可知，模型拟合指标均符合规定适配标准或临界值，结构模型拟合度较好。结构模型结果如表 5 所示，标准化路径系数用 β 表示，显著性用 P-value 显示。结果显示 H2 被拒绝，其余的假设关系成立。对于本研究的 4 个内源性结构，促成条件和感知趣味性显著影响了感知易用性，他们共同解释了感知易用性共计 59.6%的方差；促成条件对感知趣味性的解释度为 55.4%，感知趣味性和感知易用性对感知有用性具有显著影响，解释度为 52.3%。最后，54.6%的使用意愿总方差可以用感知易用性、感知有用性和感知趣味性共同解释。最终的中学生大语言模型平台使用意愿结构模型如图 2 所示。

表 5 假设汇总

假设	路径	β	估计值	S.E.值	C.R.值	P-value	结果
H1	促成条件→感知易用性	0.332	0.351	0.066	5.337	***	支持
H2	促成条件→感知有用性	0.026	0.027	0.068	0.4	0.689	拒绝
H3	促成条件→感知趣味性	0.745	0.874	0.061	14.378	***	支持
H4	感知趣味性→感知易用性	0.492	0.444	0.056	7.979	***	支持
H5	感知趣味性→感知有用性	0.264	0.231	0.06	3.867	***	支持
H6	感知趣味性→使用意愿	0.409	0.402	0.068	5.869	***	支持
H7	感知易用性→感知有用性	0.484	0.469	0.063	7.461	***	支持
H8	感知易用性→使用意愿	0.208	0.227	0.078	2.898	0.004	支持
H9	感知有用性→使用意愿	0.205	0.23	0.068	3.398	***	支持

* p < 0.05.** p < 0.01.*** p < 0.001.

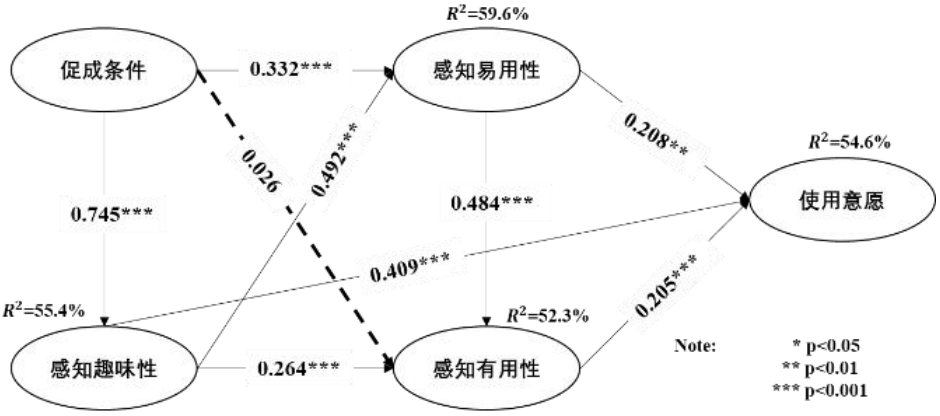


图２　中学生大语言模型平台使用意愿结构模型

6. 研究结论

6.1 促成条件通过感知易用性对感知有用性产生间接作用

研究结果显示，促成条件（β=0.122，p=0.026）对感知有用性影响不显著。这与前人的研究结果相矛盾（Li L et al,2021），但是与 Al Shamsi et al.的研究结果一致（Al Shamsi et al.，2022）。究其原因，中学生对大语言模型平台接触较少，缺乏相应的知识和技能，进而妨碍其对于类似平台的感知有用性。同时研究结果发现促成条件（β=0.332，p<0.001）对感知易用性具有显著的积极影响，感知易用性（β=0.484，p<0.001）对感知有用性产生直接正向影响。这与原始假设一致。意味着中学生会更加注重大语言模型平台的易用性，而不是实用性和有用性。便利的使用方式、完备的电脑手机硬件设施和简单易获取教程等条件可进一步降低中学生用户使用平台的难度和门槛，进而提高其感知易用性。相反，如果该平台的操作流程复杂、功能设置繁琐或者界面设计不友好，即使它具备高实用性，也会对中学生用户的使用体验产生消极影响。因此，在中学阶段推广和应用大语言模型平台时，平台自身设计需结合中学生认知发展规律和身心特点，界面布局尽可能简便易操作，功能设计需符合学生的学习需求，进而增强学生的使用意愿。

6.2 感知趣味性是影响中学生使用意愿最重要的直接预测因素

感知趣味性（β=0.492，p<0.001）和促成条件（β=0.332，p<0.001）是影响中学生用户感知易用性的前因变量，共同解释了感知有用性 59.6%的方差，并对大语言模型平台使用意愿产生直接和间接影响。此外，与 TAM 和 Al Shamsi et al 的研究结论一致的是（Al Shamsi et al.，2022），感知有用性（β=0.205，p<0.001）、感知易用性（β=0.208，p<0.001）和感知趣味性（β=0.409，p<0.001）对大语言模型平台使用意愿具有显著的直接正向影响，他们对使用意向的解释度为 54.6%。可以看出，感知趣味性是影响中学生大语言模型平台使用意愿最重要的直接因素。这说明对中学生群体来说，平台的趣味性比起其易用性和有用性而言，是更吸引他们使用该平台的影响因素。平台的功能和设计越有趣，学生越愿意使用，那么将该平台用于学习的教育目的将更容易落地。因此，趣味的功能和内容设计，是提高中学生使用大语言模型平台的重要因素。

6.3 数字化趋势中创新是必不可少的前进动力

技术变革增加了人们对教育创新的需求，而教育创新也进一步推动了技术变革。促成条件（β=0.122，p=0.026）和感知有用性（β=0.205，p<0.001）对大语言模型平台接受意愿的直接正向影响，进一步体现了大语言模型平台自身创新的重要性，与大语言

模型的有关研究结论一致（朱永新等，2023）。这说明大语言模型平台在解答专业学术问题的同时，应注重加快科研创新进度。如果平台能提升自身的创新能力，调整创造，不断地为师生带来更个性化、更人性化的使用体验，为中学生用户提供自主且适应的使用感，将有利于提高中学生用户的使用意愿。因此，在具体推广大语言模型平台过程中，持续更新和不断完善平台的外部条件和内部功能设计是增加中学生用户大语言模型平台使用次数的影响因素。

7. 总结

本研究提出了一个拓展的 TAM 模型来解释和预测中学生的使用意愿，该模型在传统的 TAM 模型基础上加入了感知趣味和促成条件两个潜在变量，能够更好地解释和预测中学生的使用意愿。研究结果表明，大语言模型平台的有用性、易用性、趣味性和促成条件都是影响中学生使用该平台的重要因素。此外，促成条件和感知趣味是影响感知易用的前因变量。感知易用和感知趣味可通过感知有用对使用意愿产生直接和间接影响。本研究结果可以为相关机构和企业提供参考，以开发更符合中学生需求的大语言模型平台。其次，可为中学教育部门提供决策参考，从而更好地整合大语言模型技术与教育资源，与时俱进，做好“人机协同”教学模式的协调者，促进教育教学的创新与改进。在未来的研究中，我们可以深入探讨其他因素对中学生使用大语言模型平台意愿的影响，如个人特质和社会支持等。同时，我们还可以对扩展的 TAM 模型进行进一步的修正和验证，以提高其在解释和预测中学生使用大语言模型平台意愿方面的适用性。此外，我们还可以进一步探索这些结论在不同文化和教育背景下的用户群体中的适用性。

参考文献：

刘明,吴忠明,廖剑等（2023）.大语言模型的教育应用：原理、现状与挑战——从轻量级 BERT 到话式 ChatGPT.现代教育技术,33(08):19-28.

刘聪,李鑫,殷兵等（2023）.大模型技术与产业——现状、实践及思考.人工智能,(04):32-42.

毕达天,张雪,米艳霖等（2023）.算法感知视角下跨平台信息推荐用户接受意愿研究.图书情报工作,67(21):100-110.

朱永新，杨帆（2023）.生成式人工智能与教育创新：机遇、挑战以及未来.华东师范大学学报,（07）:1-14.

陈舒梦（2023）.大语言模型在外语教学中的应用研究.长春师范大学学报,42(11):170-173.

李珩（2023）.基于 TAM 模型的大学英语线上学习行为意向研究.海外英语,(18):128-131+152.

徐月梅,胡玲,赵佳艺等（2023）.大语言模型的技术应用前景与风险挑战.计算机应用,1-10.

梁宇,李诺恩（2023）.中文数字学习资源使用意愿及其影响因素研究——基于 TAM 扩展模型.语言文字应用,(02):23-35.

Al Shamsi, J. H., Al-Emran, M., & Shaalan, K. (2022). Understanding key drivers affecting students' use of artificial intelligence-based voice assistants. Education and Information Technologies, 27(6),8071–8091.

Bahçekapılı, E. (2023). Predicting the secondary school students' intention to use e-learning technologies. Research in Learning Technology, 31.

Basak, E., Gumussoy, C. A., & Calisir, F. (2015). Examining the Factors Affecting PDA Acceptance among Physicians: An Extended Technology Acceptance Model. Journal of Healthcare Engineering, 6(3), 399–418.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319.

George, K., Anastasios, M. T., Dimitrios, M., & Christos, C. (2023). The mobile augmented reality acceptance model for teachers and future teachers. Education and Information Technologies.

Hsu, H., & Lin, C. (2022). Extending the technology acceptance model of college learners' mobile - assisted language learning by incorporating psychological constructs. British Journal of Educational Technology, 53(2), 286–306.

Jovanka, D. R., Habibi, A., Mailizar, M., Yaqin, L. N., Kusmawan, U., Yaakob, M. F. M., & Tanu Wijaya, T. (2023). Determinants of e-Learning Services: Indonesian Open University. Cogent Education, 10(1), 2183703.

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My Teacher Is a Machine: Understanding Students' Perceptions of AI Teaching Assistants in Online Education. International Journal of Human–Computer Interaction, 36(20), 1902–1911.

Li, L., Paudel, K. P., & Guo, J. (2021). Understanding Chinese farmers' participation behavior regarding vegetable traceability systems. Food Control, 130, 108325.

Liu, M., Wu, J., Zhu, C., & Hu, K. (2022). Factors Influencing the Acceptance of Robo-Taxi Services in China: An Extended Technology Acceptance Model Analysis. Journal of Advanced Transportation, 2022, 1–11.

Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model. Heliyon, 9(11), e20962.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46(2), 186–204.

Venkatesh, V., & Speier, C. (1999). Computer Technology Training in the Workplace: A Longitudinal Investigation of the Effect of Mood. Organizational Behavior and Human Decision Processes, 79(1), 1–28.

Won, D., Chiu, W., & Byun, H. (2023). Factors influencing consumer use of a sport-branded app: The technology acceptance model integrating app quality and perceived enjoyment. Asia Pacific Journal of Marketing and Logistics, 35(5), 1112–1133.

Zhang C, Chen J, Li J, et al. Large language models for human-robot interaction: A review[J]. Biomimetic Intelligence and Robotics, 2023: 100131

基于 LSA 的小学人工智能课堂教学行为研究

Research on classroom teaching behavior of artificial intelligence in primary schools based on LSA

宋玉婷[1*]，廖慧敏[1]，张秀梅[1]
[1]华南师范大学 教育信息技术学院
[*]1429980731@qq.com

【摘要】课堂教学行为分析有助于教师发现教学问题，并促进教学反思和教师专业发展。文章以广州市从化区小学人工智能教育教研活动中的人工智能公开示范课为研究样本，利用视频分析法和滞后序列分析法对课堂教学行为进行分析，共得到 926 个行为样本，进一步分析和评价了该示范课的课堂师生交互行为、课堂沉寂情况、教师语言影响以及信息技术应用情况，最后尝试提出人工智能课堂教学质量提升策略，研究结果可能为教师开展人工智能教学提供参考。
【关键词】人工智能；课堂教学行为；滞后序列分析

Abstract: *Classroom teaching behavior analysis is helpful for teachers to find teaching problems and promote teaching reflection and teachers' professional development. Taking the public demonstration class of artificial intelligence in the teaching and research activities of primary schools in Conghua District of Guangzhou as the research sample, this paper analyzes the classroom teaching behavior by using video analysis and lag sequence analysis, and obtains 926 behavior samples, further analyzes and evaluates the classroom teacher-student interaction behavior, classroom silence, teachers' language influence and information technology application of the demonstration class, and finally tries to put forward strategies to improve the teaching quality of artificial intelligence. The research results may provide reference for teachers to carry out artificial intelligence teaching.*
Keywords: *Artificial intelligence; Classroom teaching behavior; Lag sequence analysis*

1. 引言

随着自然语言处理、机器学习、语音识别、自动驾驶等智能技术的快速发展，人们正全面步入人工智能时代。智能时代中开展人工智能教育既是培养新时代创新人才的方式，更是经济社会发展的战略需要（位文远，2020）。我国十分重视人工智能教育。2017 年，国务院发布《新一代人工智能发展规划》，明确提出要广泛开展人工智能教育活动，在中小学阶段设置人工智能相关内容，逐步推广编程教育。2022 年，国家教育部颁发新版《2022 义务教育阶段课程标准》，将人工智能列入信息科技课程学习内容。而当前学校在开展人工智能教育时存在人工智能课程教学体系、课程内容比较单一，以传统的信息技术教学方式教授人工智能课程等问题（魏立成，2021）。基于此，本研究利用视频分析法和滞后序列分析法，采用穆肃等学者编制的 TBAS 系统分析方法，选取广州市从化区小学人工智能教育教研活动中的人工智能公开示范课为研究样本，分析示范课教学行为特征及出现的教学行为问题，尝试提出人工智能课堂教学质量提升策略。研究结果可能为中小学教师开展人工智能教学提供参考。

2. 基于 LSA 小学人工智能课堂教学行为分析

2.1. 研究样本

本研究选取广州市从化区小学人工智能教育教研活动中的人工智能公开示范课为研究对象，该节示范课的主题为神奇的无人驾驶，许多专家、中小学教师观摩了此节公开示范课，反映效果较好。仔细观摩课堂实录视频，去除片头和片尾与课堂教学无关的视频，最终以46分钟15秒的课堂视频为研究样本。

2.2. 研究方法

本研究采用视频分析法和滞后序列分析法（Lag Sequential Analysis，简称 LSA），滞后序列分析法主要用于检验人们发生一种行为之后另外一种行为出现的概率及其是否存在统计意义上的显著性（Bakeman，1997），不仅可以实现行为因果关系分析，而且能够实现行为结构模式的可视化，有利于探索隐性的关键行为序列与结构（彭林华&林晓凡，2021）。

关于课堂教学行为分析的方法，较为成熟的教学行为分析方法有弗兰德斯互动分析系统（Flanders Interaction Analysis System，FIAS）、基于信息技术的弗兰德斯互动分析系统（Information Technology-based Interaction Analysis System，ITIAS）和 S-T 分析系统，顾小清等（2004）在 FIAS 的基础上编制的基于信息技术的互动分析编码系统（ITIAS）；叶新东等（2011）重新整理交互行为影响因素，基于信息技术课堂交互要素模型和信息技术课堂交互层次模型，提出基于人−技术−环境的互动分析编码系统。穆肃等（2015）以信息化教学环境中课堂教学行为为分析的出发点，提出课堂教学行为分析系统（TBAS）及系列方法。综合多种课堂行为分析系统的优势及所选取课例的特点，本研究参考穆肃等学者提出的课堂教学行为分析系统（Teaching Behavior Analysis System，TBAS），该分析系统从教学活动的视角来分析信息化教学环境中的课堂教学行为，将课堂教学分为教师活动、学生活动和无意义教学活动等三大类（穆肃,左萍萍，2015）。在原有编码标准中补充了“学生思考”编码，编码标准如表 1 所示：

表 1 课堂教学行为编码

分类	编码	行为内容
教师活动(T)	T1	提问
	T2	反馈
	T3	讲授
	T4	指示
	T5	传统媒体演示
	T6	计算机多媒体信息演示
	T7	设备基本操控
	T8	课堂的监督控制
学生活动(S)	S1	学生应答
	S2	主动提问
	S3	与同伴讨论
	S4	做练习
	S5	观看媒体演示
	S6	学生使用媒体
	S7	学生思考
无意义教学活动(I)	I1	无助于教学的沉静或混乱

2.3. 分析过程

2.3.1. 课例《神奇的无人驾驶》教学过程时序

仔细观摩课例《神奇的无人驾驶》的课堂实录视频，课堂主要设计了七个环节：情景导入、合作探究、讲授新知、应用体验、拓展阅读、巩固练习和总结评价。该课例的教学过程时序分析如表 2 所示：

表 2 课例《神奇的无人驾驶》教学过程时序

时间	教学环节	教学行为描述	
		教师	学生
00:25-01:17	情景导入	播放视频资源	观看视频资源
01:18-01:46		提出问题，提问学生，对学生的回答给予及时反馈	举手回答问题
01:47-02:00		通过问题导入新课主题	认真思考
02:01-04:43	探究新知 1：合作探究—普通汽车在行驶时可能遇到的情况及其处理方法	用多媒体呈现图片，提出问题，说明具体任务，引导小组讨论	小组交流讨论，完成学习任务单
04:44-08:56	小组展示	邀请各个小组代表分享观点，提示汇报时需要注意的地方，对学生的回答给予及时反馈	派代表汇报小组观点
08:57-15:37	探究新知 2：自主阅读，交流讨论—无人驾驶汽车工作原理	利用多媒体提出引导性问题，引导学生阅读课本以及交流讨论	认真阅读课本并思考问题，小组交流讨论
15:38-16:38		邀请学生分享观点，对学生的回答给予及时反馈	举手回答问题
16:39-18:51	讲授新知 1：无人驾驶汽车工作原理及其运用的人工智能技术	提出问题，播放视频资源	观看视频资源，认真思考
18:52-21:24		讲解，提出引导性问题	思考并回答问题
21:25-22:03	应用体验	用多媒体呈现图片，提出问题	观看屏幕，认真思考
22:04-24:30		演示 3D 虚拟实验室平台基本操作，讲解大概思路	认真观看并回答问题
24:31-28:23		给予学生引导和提示，进行教学管理	学生在 3D 虚拟实验室平台上操作
28:24-30:00	讲授新知 2：3D 虚拟实验室平台中程序代码功能	讲解程序代码，提出问题	思考并回答问题
30:01-32:13	应用体验	提出问题，引导学生修改小车速度	在平台上修改小车速度
32:14-34:01	讲授新知 3：无人驾驶在生活中的应用	提出问题，讲解	思考并回答问题
34:02-36:27	拓展阅读：无人驾驶技术的	提出问题，引导学生阅读课本内容	自主阅读

36:28-38:17	利弊	引导学生小组交流讨论	小组交流讨论
38:18-40:06		邀请学生分享观点，对学生的回答给予及时反馈	举手回答问题
40:07-40:37		用多媒体呈现图片，教师讲解	观看屏幕，认真思考
40:38-41:35	课堂总结评价	提出问题，引导学生回顾所学知识，引导学生评价目标达成情况	认真思考，回答问题
41:36-43:12	巩固练习	引导学生做练习	做练习，完成学习任务单
43:13-46:07	课堂小结	提问学生，给予及时反馈，总结学生回答	回答问题
46:08-46:40		总结本节课内容	认真思考

2.3.2. 数据编码过程

依据课堂教学行为编码标准（表 1），对示范课课例《神奇的无人驾驶》进行视频编码，每隔 3 秒进行一次采样，并按时间顺序记录 3 秒内的主要行为类别，得到 926 个行为样本。编码过程中采取以下原则：

（1）如果同一种教学互动行为持续时间超过 3 秒钟，则每隔 3 秒记录一个相同的数字；

（2）如果 3 秒内出现了多种互动行为，则标记为主要的互动行为的编码；

（3）教师话语中不好辨别为具体某个类别的记为“T3”讲授，比如“上课”“同学们好”；

（4）教师给出的提示性和指示性话语，如“时间到、请坐、请举手、请同学们试一试、课后请同学们自行操作、提醒无法进入实验室平台的同学和同伴一起看、阅读完毕的同学可以和小组成员交流讨论”等记为“T4”指示；

（5）教师在讲授前操作多媒体的行为记为“T7”设备基本操控；

（6）学生自主阅读课本记为“S4”做练习；

（7）教师提出问题后学生的短暂沉寂记为“S7”学生思考；

（8）课堂中学生鼓掌、递话筒、沉寂等行为记为“I1”无意义教学活动。

2.3.3. 编码数据的分析工具

本研究采用 GSEQ 软件，该软件由 Bake-man 和 Quera 设计开发，简化了行为分析流程，提高了数据处理效率，也促进了 LSA 在各领域中的应用推广（杨现民，王怀波&李冀红，2016）。运用 GSEQ 软件处理和分析所获得的编码行为数据会生成两张表格，分别是行为转换频次表（见表 3）和调整后的残差表（见表 4）。行为转换频次表呈现一种行为引发其伴随行为的频次，调整后的残差表代表根据行为转换频次表得到的残差参数（Z）（江毅，王炜&李辰钰，2019）。若 Z 值大于 1.96，表明该行为序列具有统计意义上的显著性（Bakeman，1997）。

表 3 行为转换频次表

	T1	T2	T3	T4	T5	T6	T7	T8	S1	S2	S3	S4	S5	S6	S7	I1	Totals
T1	22	0	5	4	1	1	0	0	28	0	0	0	0	0	5	3	69
T2	5	3	6	2	0	0	0	0	1	0	0	0	0	0	0	7	24
T3	16	0	91	0	2	1	0	0	3	0	0	0	2	0	0	0	115
T4	0	0	2	39	0	2	0	0	3	0	4	5	0	3	0	9	67
T5	0	0	4	0	7	0	0	0	0	0	0	0	0	0	0	0	11
T6	6	0	0	3	0	52	0	0	0	0	0	0	0	0	0	0	61
T7	0	0	1	0	0	0	4	0	0	0	0	0	0	0	0	0	5
T8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1	11	20	4	6	1	3	0	0	102	0	0	0	0	0	0	3	150
S2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S3	1	0	0	4	0	0	0	0	0	0	92	0	0	0	0	0	97
S4	1	0	0	3	0	0	0	0	0	0	1	147	0	0	0	0	152
S5	1	0	1	0	0	0	0	0	0	0	0	0	48	0	0	0	50
S6	0	0	0	3	0	0	0	0	0	0	0	0	0	81	0	0	84
S7	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	5
I1	5	1	1	2	0	1	1	0	12	0	0	0	0	0	0	12	35
Totals	69	24	115	67	11	61	5	0	150	0	97	152	50	84	5	35	925

表 4 调整后的残差表

	T1	T2	T3	T4	T5	T6	T7	T8	S1	S2	S3	S4	S5	S6	S7	I1
T1	8.03	-1.41	-1.36	-0.48	0.21	-1.79	-0.64	0	5.71	0	-2.96	-3.83	-2.06	-2.73	7.9	0.26
T2	2.53	3.09	1.89	0.21	-0.54	-1.32	-0.37	0	-1.62	0	-1.7	-2.2	-1.19	-1.57	-0.37	6.6
T3	2.81	-1.87	23.17	-3.2	0.58	-2.64	-0.84	0	-4.23	0	-3.92	-5.08	-1.86	-3.62	-0.84	-2.27
T4	-2.41	-1.39	-2.43	16.71	-0.93	-1.24	-0.63	0	-2.71	0	-1.25	-2.06	-2.03	-1.36	-0.63	4.3
T5	-0.95	-0.54	2.42	-0.93	19.22	-0.89	-0.25	0	-1.47	0	-1.14	-1.48	-0.8	-1.05	-0.25	-0.66
T6	0.73	-1.32	-3.04	-0.72	-0.89	25.61	-0.6	0	-3.56	0	-2.77	-3.58	-1.93	-2.55	-0.6	-1.6
T7	-0.64	-0.37	0.51	-0.63	-0.25	-0.6	24.3	0	-0.99	0	-0.77	-0.99	-0.54	-0.71	-0.17	-0.44
T8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1	-0.06	9.04	-3.96	-1.67	-0.64	-2.48	-0.99	0	18.8	0	-4.58	-5.93	-3.2	-4.23	-0.99	-1.25
S2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S3	-2.55	-1.7	-3.92	-1.25	-1.14	-2.77	-0.77	0	-4.58	0	28.66	-4.62	-2.49	-3.29	-0.77	-2.06
S4	-3.49	-2.2	-5.08	-2.74	-1.48	-3.58	-0.99	0	-5.93	0	-4.33	29.22	-3.22	-4.26	-0.99	-2.67
S5	-1.51	-1.19	-2.3	-2.03	-0.8	-1.93	-0.54	0	-3.2	0	-2.49	-3.22	29.13	-2.3	-0.54	-1.44
S6	-2.73	-1.57	-3.62	-1.36	-1.05	-2.55	-0.71	0	-4.23	0	-3.29	-4.26	-2.3	29.22	-0.71	-1.91
S7	1.07	-0.37	-0.84	1.1	-0.25	1.21	-0.17	0	0.23	0	-0.77	-0.99	-0.54	-0.71	-0.17	1.91
I1	1.57	0.1	-1.75	-0.36	-0.66	-0.91	1.91	0	2.96	0	-2.06	-2.67	-1.44	-1.91	-0.44	9.64

2.4. 研究结果

2.4.1. 师生交互行为频次及比率分析

行为转换频次表（表 3）中每个单元格的数字表示这对编码出现在课堂行为中的频数，对表 3 中各个单元格的频数及其之间的比例关系进一步分析，绘制出课例《神奇的无人驾驶》师生课堂行为编码频次及比率统计表（见表 5）。

根据表 5，整节课中教师活动占比 38.05%，学生活动占比 58.16%。该节示范课以学生活动为主，教师讲授和操作演示为辅，注重学生感知、体验人工智能技术，注重学生自主学习和合作探究，此外，在教学过程中还贯穿了对学生问题解决能力、协作学习能力、语言表达能力、批判性思维等多种能力的培养。表中出现的最高频次编码依次是：做练习（S4）、学生应答（S1）、讲授（T3），发生频次分别为：152、150、115。其次，出现频次较高的是与同伴讨论（S3），出现的频次为 97，占比 10.49%。说明该节课以学生活动为主，注重学生自主学习和合作探究。教师在讲授知识的同时注重引发学生思考讨论，通过提出引导性问题来引发学生思考，促进学生更好地理解所学知识。频次最低的是课堂的监督控制（T8）和主动提问（S2），发生的频次均 0。说明课堂中学生的纪律良好，需要教师监督控制的情况很少。而学生根据学习内容，主动提出问题，期望教师给予解答的频次为 0，反映出学生学习的主动性还有待加强。

表 5 课例《神奇的无人驾驶》师生课堂行为编码频次及比率统计表

分类	编码	行为内容	频次	百分比
教师活动 (T-38.05%)	T1	提问	69	7.46%
	T2	反馈	24	2.59%
	T3	讲授	115	12.43%
	T4	指示	67	7.24%
	T5	传统媒体演示	11	1.19%
	T6	计算机多媒体信息演示	61	6.59%
	T7	设备基本操控	5	0.54%
	T8	课堂的监督控制	0	0.00%
学生活动 (S-58.16%)	S1	学生应答	150	16.22%
	S2	主动提问	0	0.00%
	S3	与同伴讨论	97	10.49%
	S4	做练习	152	16.43%
	S5	观看媒体演示	50	5.41%
	S6	学生使用媒体	84	9.08%
	S7	学生思考	5	0.54%
无意义教学活动（I-3.78%）	I1	无助于教学的沉静或混乱	35	3.78%

2.4.2. 师生交互行为序列分析

根据滞后序列分析理论，如果 Z-score>1.96 则表明该行为路径具有显著意义。对调整后的残差表（表 2）中残差参数（Z）值大于 1.96 的数据进行统计，绘制出行为转换图（见图 1）。

图中节点表示各种用户行为，连线表示行为与行为之间的连接具有显著意义，箭头代表行为转换的方向，线条的粗细表示行为连接的显著水平，线条上的数据则是调整后的残差值(Z-score)（杨现民, 王怀波&李冀红, 2016）。

根据行为转换图（图 1），该节课中 T2➞T1，T3➞T1，T5→T3，T1➞S1，T1➞S7，S1➞T2 等行为序列达到了较高的显著性水平。从教师讲解行为来看，教师讲解的行为会跟随其进行传统媒体演示行为的出现而出现（T5→T3），说明教师在利用技术呈现教学内容的同时会伴有辅助性的语言讲解，以引导学生更好地理解知识（彭林华&林晓凡，2021）；教师在讲解之后倾向于提出引导性问题（T3➞T1），说明教师注重引发学生思考、启迪学生思维，引导学生带着问题来学习。从师生互动行为来看，行为序列 T1➞S1➞T2➞T1 形成一个闭环，即教师在提问问题后学生能够及时应答（T1➞S1），学生应答后教师能够给予及时地反馈（S1➞T2），教师反馈后会继续提出引导性问题，促进学生进一步思考问题（T2➞T1）。此外，学生思考会跟随教师提问行为的出现而出现（T1➞S7），说明教师在提问问题后有给予学生一定的时间进行思考，但从表 5 可以看出，S7（学生思考）仅占比 0.54%，说明教师给予学生自己思考的时间较少。

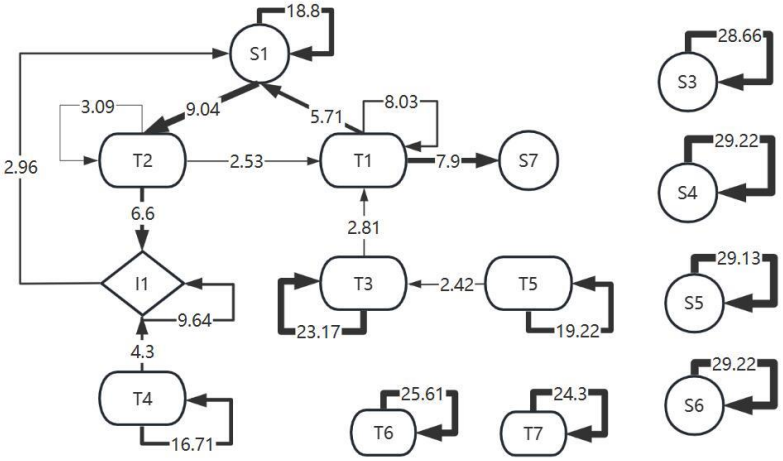


图 1 行为转换图

2.4.3. 教师语言影响分析

穆肃等学者对信息化环境下的优质课堂教学中教师行为与学生活动的关系及对学生活动的影响进行研究，研究发现，教师语言间接影响与直接影响的比例与学生活动存在较弱的正相关关系，教师督促、提醒等间接语言行为更能促进学生活动的发生（穆肃, 董经&唐冬梅, 2019）。表 5 中，教师的间接影响包括 T1、T2、T8，直接影响包括 T3、T4、T5、T6、T7，（T1+T2+T8）/（T3+T4+T5+T6+T7）=93/259≈0.359<1，说明在该节示范课中，教师更倾向于利用直接影响的方式来影响学生。T1（提问）、T2（反馈）分别占比 7.46%、2.59%，这两种间接影响的方式占课堂教学总行为的比率较低。

2.4.4. 技术使用分析

根据行为转换图（图 1），S5➞S5、S6➞S6、T5➞T5、T6➞T6、T7➞T7 等行为序列达到了较高的显著性水平。学生观看媒体演示（S5）指学生观看教师在教学过程中播放的视频，学生使用媒体（S6）指学生操作人工智能 3D 虚拟实验室平台。传统媒体演示（T5）指本节课中教师通过书写板书、挂图等方式呈现教学信息、要点，计算机多媒体信息演示（T6）主要指教师演示人工智能 3D 虚拟实验室平台基本操作，设备基本操控（T7）指教师在讲授前操作多媒体的行为。信息技术在该节示范课的教学中起到了重要的作用，该节示范课中教师利用人工智能 3D 虚拟实验室平台加深学生对无人驾驶技术的感知和体验，教师利用精心准备的视频资源，激发学生的学习兴趣，拉近师生距离，并帮助学生更好地理解无人驾驶技术。

2.4.5. 课堂沉寂分析

从行为转换图（图 1）可以看出，I1→I1、I1→S1、T2→I1、T4→I1 等行为序列达到了较高的显著性水平。本研究中无助于教学的沉静或混乱（I1）包括课堂中学生鼓掌、传递话筒、小组汇报时现场的沉静、停止讨论后学生就坐引起的混乱等多种情况，因此 I1→I1 行为序列较为显著，但属于课堂中正常发生的情况，不属于严重的课堂秩序混乱。此外，学生应答会跟随无助于教学的沉静或混乱行为的出现而出现（I1→S1），这是由于在学生回答问题环节，一位学生回答完问题后需要将话筒进行传递，进而导致无助于教学的沉静或混乱（I1）行为。在教师反馈（T2）和给予指示（T4）后会存在无助于教学的沉静或混乱（I1）行为，这是由课堂中学生鼓掌、传递话筒、停止讨论后学生就坐引起的混乱等因素引起的，也属于课堂中正常现象。总体来看，整节课课堂秩序良好，整个课堂过程中学生行为表现较为积极，课堂结构呈现为教师主导，学生主体。教师是课堂中的引导者，能够很好地组织和调控课堂气氛和活动，较好地发挥了学生的主动性，激发了学生的学习积极性。同时，需要注意教师应合理设计课堂教学，根据课堂随机应变，使课堂教学更加顺畅连贯，避免出现课堂中断、沉寂或混乱（王觅, 文欣远, 李宁宁&陈焕东，2020）。

3. 人工智能课堂教学策略与建议

通过对示范课《神奇的无人驾驶》的教学行为分析和课例观察，对该课例的课堂师生交互行为和信息技术应用情况进行了进一步的分析。结合基于 LSA 小学人工智能课堂教学行为分析结果，本研究尝试提出人工智能课堂教学质量提升策略。

3.1. 注重学生人工智能核心素养的培养

教师需注重培养学生自然空间的新科学素养、信息空间的新数字素养以及社会空间的新人文素养，不仅要关注科学知识、科学思维和科学研究方法的传授，还要关注数据收集与处理、问题分析与评价、团队协作与道德法制等能力意识的提升（陈廷徜，2021）。

3.2. 融合多种技术，提供多样感官体验

随着 VR、AR、物联网、云计算、5G 等技术的快速发展，融合多种技术实现对教学的辅助已经成为一种趋势。具身认知主张人是通过身体以合适的方式与世界进行互动并在此过程中获得对世界的认识（王美倩&郑旭东，2015）。人工智能课堂教学应注重将人工智能与多种教育技术相结合，从而激发学生的主动学习，扩充学生的知识来源，帮助学生实现体验式学习，获得良好的具身学习体验（郑晓俊&吴静雨，2022）。

3.3. 根据多元评价指标，开展多样化活动评价

教师需要根据多元评价指标，进行多样化的活动评价，以更全面地评估学生。这包括过程和总结性评价，以及对学生发现问题、解决问题的思路和教学效果等方面的评价（路燕飞，2021）。评价方式可以组合学生自评、小组互评和教师评价。

3.4. 利用技术激活学生高阶思维与能力

在人工智能课程中借助信息技术创设多样化教学情境，有利于培养学生创新、应用、鉴赏、评价等高阶思维与能力。信息技术不仅可以作为教师讲授、提问等教学行为的辅助工具，还可以利用多种信息技术将抽象的事物具体化、形象化、生动化，创设多样化教学情境（刘革平, 孙帆&廖剑，2023）。此外，教师还可以利用线上线下结合的智能化、及时化的互动方式，促进师生和生生之间的交流与思想碰撞，提升学生的协作、分析和表达能力（刘革平, 孙帆&廖剑，2023）。

4. 结论与展望

本研究利用课堂教学行为分析系统（Teaching Behavior Analysis System，TBAS），编码

和记录示范课《神奇的无人驾驶》的课堂教学行为，进一步分析和评价了该课例的课堂师生交互行为、课堂沉寂情况、教师语言影响以及信息技术应用情况，并尝试提出人工智能课堂教学质量提升策略。研究结果表明：该节课以学生为中心，强调自主学习与合作探索；教师善于通过提问引导启发学生思考，但学生主动提问次数较少；教学过程中形成了有效的提问—反馈循环，但学生独立思考时间有限；人工智能技术能够有效助力教学，提升学生兴趣。

诚然，在具体的课例分析过程中还存在一些问题，比如对某些非语言行为难以进行归类；不同研究者对同一个课例进行编码时，可能会存在一些细微的标准差异等（张悦&张秀梅，2018）。本研究的数据较少，未来研究可以结合其他教学行为分析系统，采用其他编码方式，对比分析多个人工智能课堂教学案例，更深入地分析课堂教学存在的不足，以提升人工智能课堂教学效果。

5. 参考文献

位文远. (2020). 中小学信息技术课程中人工智能教育的应用探讨. 教育教学论坛(26),368-369.

(2017). 国务院关于印发新一代人工智能发展规划的通知. 中华人民共和国国务院公报(22),7-21.

魏立成. 中小学信息技术课程中的人工智能教育[J]. 中国教育技术装备, 2021(19):76-78.

Bakeman R. (1997). Observing interaction: An introduction to sequential analysis. Cambridge: Cambridge university press.

彭林华 林晓凡. (2021). 基于滞后序列分析的智慧课堂教学行为研究. 现代教育技术(07),55-61.

顾小清 王炜. (2004). 支持教师专业发展的课堂分析技术新探索. 中国电化教育 (07), 18-21.

叶新东 夏一飞. (2011). 信息技术课堂教学交互行为模型与评测指标研究与实践. 现代教育技术(08),42-50.

穆肃 左萍萍. (2015). 信息化教学环境下课堂教学行为分析方法的研究. 电化教育研究(09),62-69.

杨现民, 王怀波 李冀红. (2016). 滞后序列分析法在学习行为分析中的应用. 中国电化教育 (02), 17-23+32.

江毅, 王炜, 李辰钰, 康苗苗 沈洁. (2019). 智慧教室环境下师生互动行为研究. 现代远距离教育(03),13-21.

穆肃, 董经, 唐冬梅 周艳洁. (2019). 信息化课堂教学中教师行为对学生活动的影响. 中国电化教育(08),91-98.

王觅, 文欣远, 李宁宁 陈焕东. (2020). 大单元教学视角下基于 LSA 的同步课堂师生交互行为研究. 电化教育研究(08),74-81.

陈廷徜. (2021).中小学人工智能教育应重点关注三要素. 江苏教育(60),1.

王美倩 郑旭东. (2015). 具身认知与学习环境:教育技术学视野的理论考察. 开放教育研究(01),53-61.

郑晓俊 吴静雨. (2022). 面向具身认知的人工智能教学应用现状与推进策略研究. 中国现代教育装备(22),49-53.

路燕飞. (2021). 小学信息技术课堂中人工智能教育教学策略研究硕士学位论文, 南宁师范大学.

刘革平, 孙帆 廖剑. (2023). 基于行为序列分析的智慧课堂有效教学行为辨识. 现代教育技术(06),82-91.

张悦 张秀梅. (2018).基于 ITIAS 的优课教学行为分析研究——以一门部级优课为例. 广州广播电视大学学报(06),47-53+109.

作文说:大模型驱动的一对一作文智能导学系统

Large Model-Driven One-on-One Intelligent Tutoring System for Composition Writing

李雲帆[1]，张立山[1*]，李万悦[2]，廖俐丽[1]，吴涵[1]
[1]华中师范大学
[2]西安市渭北中学
[*] lishan.zhang@mail.ccnu.edu.cn

【摘要】 写作能力是一项基础且重要的能力，决定了学生是否能够准确、简练地表达出自身的想法。学生写作能力的培养有赖于教师的个性化作文批改。当前，工业和学术界都涌现了许多辅助教师进行个性化批改的智能软件，然而，这些软件还都面临着批改反馈内容模板化痕迹严重、针对性不强等问题。生成式大语言模型的出现，为该问题的解决提供了一条新的技术路线。本研究基于移动端场景研发了一款名为“作文说”的一对一智能作文导学系统，该系统由生成式大语言模型驱动，通过提示词工程，结合学生的作文文本内容，生成适应性多层次反馈内容。我们的预实验表明，大模型可以生成合乎需要的反馈内容。本研究亦可为“大模型+教育”的其他应用场景提供重要参考。

【关键词】 智能导学系统；自动作文批改；AI 教育应用

Abstract: *Writing ability is a fundamental and crucial skill that determines whether students can express their ideas accurately and concisely. The cultivation of students' writing skills depends on teachers' personalized essay feedback. Currently, both industry and academia have witnessed the emergence of numerous intelligent software tools that assist teachers in providing personalized feedback. However, these tools still face challenges such as the obvious trace of template-based feedback and the lack of specificity. The advent of generative large language models has offered a new technical approach to address this issue. This study has developed a one-on-one intelligent essay tutoring system named "Composition Speak," which is tailored for mobile scenarios. Driven by a generative large language model, the system employs prompt engineering in conjunction with the content of students' essays to generate adaptive, multi-level feedback content. Our preliminary experiments indicate that large models are capable of producing feedback that meets the necessary requirements. This research also provides important references for other application scenarios of "large models + education."*

Keywords: Intelligent tutorial system,Automated essay evaluation,AI application in education

1. 引言

智能导学系统以人类教学模式为黄金参考标准（张立山等，2023），实时监测学生学习状态并提供即时的反馈指导，故能实现因材施教，提供类似于专家教师教学效果的个性化辅导服务。

然而，一个有效的智能导学系统必须保证系统的高度有用性和可靠性，否则便会受到来自教师、学生和家长的指责和抛弃，这一点对于面向高年级的应用尤为突出。因此系统必须基于具体的场景，在可容忍的时限内生成大量具有多样性的高质量反馈内容。

为解决这个问题，传统方法采用了一系列逐渐递进的技术，包括基于规则的方法、基于特

征的统计机器学习方法、深度学习方法，以及基于小型语言模型（PLM）的方法，但实际效果往往比较有限，与此相关的工程往往面临成本高昂和开发周期长的难题。因此构建出的系统往往或者过于模式化而不够智能，或者过于宽泛模糊缺乏针对性。这些问题导致构建的系统在实际应用中常常显现出价值不足，难以大规模推广。在这一背景下，大模型以其强大的涌现能力和泛用性，结合提示词工程、模型微调等技术，为智能导学系统反馈内容生成中所面临的多方面挑战提供了崭新的可能性。

作文智能导学系统一直是智能导学系统中备受关注的一个经典场景。这是因为学生写作能力的提升必然需要大量练习、个性化建议反馈以及多轮修改迭代（郑蝉金、郭少阳等，2023），而一方面学校教师受限于课时资源往往难以针对每个学生提供充足的个性化作文辅导服务。另一方面，学生也受限于传统纸质媒体的交互形式，即便获得反馈，也难以立即进行有针对性的修改和学习。这导致在教学实践中，学生的写作能力提升受到了显著制约。

由于传统的一对一作文智能导学系统即使应用最新的 BERT 等技术，在反馈内容生成方面也显得能力相对不足。例如传统方法下的作文智能导学系统往往在语法的识别和改正上有着不错的表现，但是对于学生作文结构评价、具体情景段落句子的精准点评和指引等个性化内容的生成上存在显著欠缺。因此作文导学系统因此在人机交互的设计上往往受此制约，存在学生的主动参与不足等弊病，并且往往很需要真人教师的进一步辅助介入。大模型技术大大的降低了该领域的工程技术门槛并陆续涌现出了一些新的基于大模型的作文智能导学系统，同时也带来了一些新问题。

此外，尽管大模型带来了低成本快速生成个性化高质量内容的新方案，我们很遗憾地发现新出现的大模型作文智能导学系统仍多墨于旧的设计框架和人机交互思路。如何针对具体场景合理安排这些丰富的反馈内容、设计更合理的人机交互场景？这是新系统们亟待解决的。同时大模型生成的反馈内容的价值性和可靠性也亟待进一步的实证检验，尤其是来自一线教师和学生的考察。针对以上两方面的问题，我们构建了一个名为‘作文说’的一对一语文作文智能导学系统，该系统使用大模型生成反馈内容，重视学生的参与和学习效果，以‘即批、即学、即改’为特色，基于移动端场景设计了多层次的反馈和人机交互机制。

为了评估大模型生成反馈内容的有效性，我们收集了多个真实班级的学生语文作文数据作为输入，并进行了批量测试与评分。研究结果显示，大模型在本系统的多个应用场景中具备生成合乎使用要求的高质量反馈内容的能力。

2. 相关工作：智能导学系统的反馈设计

智能导学系统是一种内循环导学工具，以步骤为基本单位提供教学支持，其主要功能在于通过针对性的反馈帮助学生完成当前超出其能力范围的任务元素，从而使他们能够专注于其能力范围内的任务（张立山，2023）。在人机协同学习环境中，反馈的设计被认为是激发学习者动机、提高学习绩效、调整学习过程以及改进教师教学的重要因素（何克抗，2017）。特别是形成性反馈，即贯穿于整个学习过程中学习者难以直接获取的关键学习信息，被广泛认为是提升知识和技能获取至关重要的因素（Shute, 2008）。

根据反馈的性质和发挥作用的方式，相关研究者对反馈进行了不同的分类。Hattie 和 Timperley 总结了一些反馈实证研究结果，将反馈根据其所含信息的复杂性将其划分为结果性反馈、正确反应反馈和精细反馈（Hattie、Timperley，2007）。张立山等学者进一步细化了反

馈的分类，将其分为基本反馈和针对错误的反馈两类。基本反馈主要涉及对学生解题步骤的简要评价，包括正确性和最优性等方面。而针对错误的反馈则按照复杂度递进划为三个层级。首级反馈仅指出错误，具体改正工作由学生独立完成。次级反馈则提供一些模糊的改正提示，只有在学生需要帮助时系统才会提供末级反馈，即提供最终答案（张立山等，2023）。

在反馈的复杂度对学习效果影响的实证方面，龚少英等人通过对照实验得出结论：反馈复杂度和学生的迁移成绩正相关，但仅对知识背景欠缺的学习者有效（龚少英、韩雨丝等，2019），考虑到注意力是一种稀缺的资源，因此在人机交互流程中给予学生充分的选择权，可以重点挑选其学习的难点和跳过其已经掌握的部分是十分重要的。对于简单的等级评分及更详细的书面评语的效果，也有学者进行了实证研究，Page 等发现预设定评语相较于等级评分更容易刺激学习者的学习热情（Page 等，1958）。且必须对结论性的等级评分提供充分描述性解释，否则反馈效果则可能导致困惑（Grant 等，2012）。

反馈时机也是影响学习效果的重要变量，对于这个话题支持者与反对者莫衷一是。总结了以往关于反馈时机的研究后，Mathan 等认为研究结果的不一致表明了反馈时机与学习之间的关系不存在固定的影响机制，不论是即时反馈还是延迟反馈，都可能产生积极和消极的影响。如即时反馈有助于激发学习动机、明确因果关系和促进得出结论，但可能导致学习者过分依赖反馈外脑。与之对应的，延迟反馈有利于积极型学生主动地进行认知和元认知加工，但可能对于学习能力和学习动机不足的学生掌握知识技能造成阻碍。因此对于反馈时机的把握应当充分建立在任务场景和对学习者能力了解的前提上 (S. A. Mathan and Koedinger, 2002)。夏雪莹等则指出当学生涉及简单概念或事实时，即时反馈对推进学习过程是有利。但当学生试图达到更高层的学习目标时，适度延迟的反馈可能更合适。因此，反馈时机的设计应当遵循及时性和持续性原则（夏雪莹等，2023）。此外针对反馈的形式也有一些学者做了研究，Chi 等提出基于对话的 ITS 或对比传统的 ITS 展现了更好的学习效果 （Chi et al., 2014），Lin 等指出建议式表达优于说教式表达（Lin, Ginns, Wang, & Zhang, 2020）。

综上所述，智能机器能够为学习者提供适切的学习支架和情感支持（李海峰 & 王炜，2018）。其角色应定位为引导者，而非“精准投喂”的饲养者，以免导致学习者丧失思辨能力，这与教育的初衷相悖。（何文涛等，2023）。一个优秀的智能导学系统必须以人为本，充分考虑用户、需求和场景，精心对反馈进行复杂度分层，并审慎判断反馈时机。

3. 系统设计

在设计作文智能导学系统时，我们从用户、需求和场景三个方面进行探讨。首先，在用户角度方面，本系统可以为小学至研究生阶段的语文和英语作文提供辅导服务。但鉴于当前大模型能力限制以及高年级用户对服务质量的较高要求，我们将核心用户群体定位为小学三年级至初中三年级的学生。其次，在需求角度方面，系统强调学生的主动参与，以“即批、即学、即改”为特色。通过从语法错误、文章结构、语句表达等多个维度帮助学生从作文修改润色实践中学习相关知识技法，提升写作能力。

再者，在场景角度方面，我们认为 K12 阶段学生使用手机相较于电脑受管制程度较低，同时也更适合单人作文批改学习这一轻度场景。故我们针对移动端进行了用户界面与交互设计。

我们认为优秀的AI工具应是自来水式的，好的AI服务应由需求催生且尊重现有场景习惯，

在当前设计中我们尚未直接提供面向老师和家长的服务，以免反而成为他们的负担。



图 1：作文说界面展示

对比传统方法驱动的作文智能导学系统，本研究提出的作文说系统独具特色，这体现在大部分场景反馈内容由大模型生成。然而，鉴于当前大模型技术正处于快速发展阶段，不同模型在不同场景下的表现存在显著差异。此外，在某些场景中若生成的结果不尽如人意时调用其他大模型重新生成内容也是一个好的反馈纠偏策略。

因此，在系统设计中我们将模型端与用户端分离，以提高系统灵活性和适应性。在这种设计下，作文说系统能够根据不同场景选择最适合的大模型，以实现更高质量的反馈内容。这一创新举措有助于提高系统性能和用户体验，为用户提供更为有效的作文辅导服务。图 2 展示了作文说导学系统模型端的反馈设计流程：

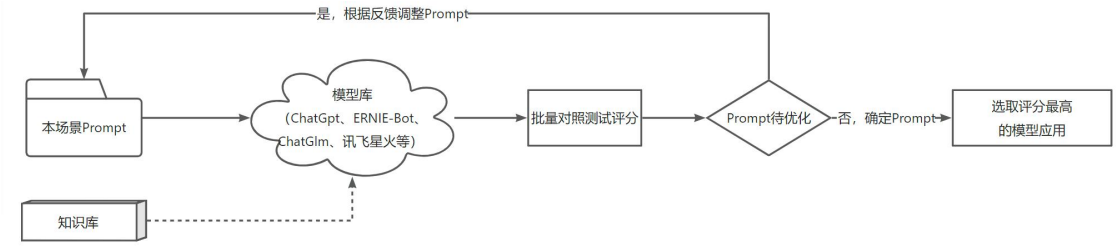


图 2：作文说模型端反馈设计流程

在实施上述步骤时，以下几个问题值得关注：

（1）Prompt 工程：在编写 Prompt 时，我们主要参考了 OpenAI 官方建议的五条顶层原则及其拆分的若干细则原则。值得注意的是，不同学段的学生在写作过程中面临不同的痛点。例如，在语文作文方面，根据对一线教师的采访，小学生的记叙文问题主要在于分不清文章的重点，描写内容与主题契合度不高；而高中生的议论文问题则主要集中在论据使用、论证结构和方法方面，要求更高。因此，针对不同学段和场景，我们需要调整 Prompt 内容，以提高系统的自适应性能。

（2）批量对照评分测试：我们在多个大模型中应用相同的 Prompt，并进行不少于 50 个样例的批量场景对比实验。针对每个样例的模型输出，单独进行评分并记录潜在问题，根据实际的批量测试反馈对 Prompt 进行多轮调整，并挑选综合表现最优的模型。

（3）知识库：鉴于不同大模型在运用知识库微调模型方面的能力与成本存在显著差异，本系统将知识库视为一种可选策略，针对仅依靠 Prompt 工程或模型替换手段难以应对的问题

场景使用。

在用户端设计方面，图 3 展示了作文说智能导学系统的主操作流程：

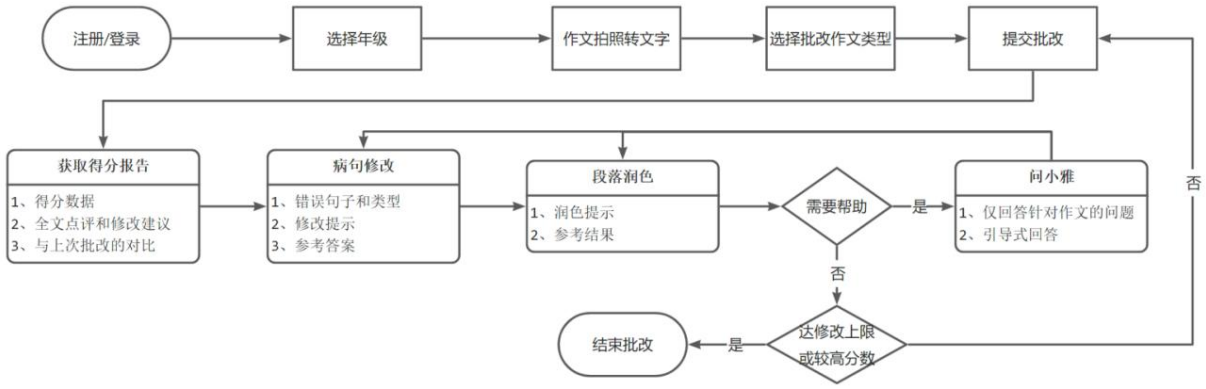


图 3:作文说的智能导学主流程

相较于传统方法驱动的作文导学系统，大模型实现了低成本即时生成多场景高质量内容的愿景，从而拓展了旧技术难以良好支撑的丰富反馈场景。此优势可以促使学生更加积极主动地参与导学过程，在作文批改实践中进行学习。表 1 对作文批改场景中的七个子场景进行了分类，各场景根据复杂度和反馈时机的不同而有所侧重。

表 1：作文说场景反馈的分类

反馈/层次	首层反馈	次级反馈	末级反馈
即时反馈	错句消息	得分报告	
延时反馈	*自由提问	非错句-点评	非错句-润色结果
		错句提示	错句-参考答案

*自由提问是修改本作文超过 5 分钟后解锁的功能，用户可以通过自由对话向 AI 助教请求帮助，因此结果可能贯穿于三层反馈

作文说的反馈分为三个层次并尽量限制他们在完成任务过程中所需获得的帮助以保证学生有充足的思考时间。首层反馈关注系统检测到的错误，系统将明确指出含有错误的原句以及错误类型（包括语法、拼写和标点错误）；次级反馈协助学生复习相关知识和方法并提供启发式引导。例如针对作文全文的写作报告，针对错句的修改提示和针对正常句子的点评建议；末级反馈针对学生自主努力无法解决的问题，直接提供正确答案。如展示修改后的错误句子或原作文句子的润色效果。

在反馈策略的实施时机方面，我们遵循“支架延迟提供原则”（何文涛，2023）以保证学习者拥有充足的自主思考时间，并预防部分学生为迅速获取解题步骤而滥用提示的情况。系统会根据学生的实时状态，自适应选择并提供相应提示，同时在提示展示过程中，赋予学生掌控权，以提高适应性。

此外，层级越靠后的反馈，其停留时间限制相应更为严格。以错句纠正为例，首层反馈会直接给出错句内容及错误类型，而改正错句的提示需在 30 秒后才能点击查看，此后等待 60 秒即可直接查看答案。至于更为高效且灵活但难以监管的 AI 助手，至少需在全场景累计等待 8 分钟后方可解锁，同时对学生每篇作文设置 20 次提问次数上限。

4. 实验和讨论

我们进行了实验以评估大模型在本系统多个场景下生成内容反馈的质量，这对于本系统的可用性和价值性具有重要意义。

具体而言，我们收集了来自西安市渭北区中学的 120 篇语文作文，包括初一学生的 80 篇及初二学生的 40 篇考场作文。在此基础上，我们利用预设的大型模型针对不同场景生成了反馈内容样本，包括得分报告（120 个样本）、错句提示（50 个样本）、错句参考答案（50 个样本）、非错句点评（50 个样本）以及非错句润色结果（50 个样本）。

首先，我们考察了人工智能生成反馈的时效性。具体而言，我们对生成文本内容最长的得分报告场景样本的生成时间进行了统计。结果表明，在该场景下，120 个样本的平均生成时间为 18.65 秒，中位数为 18.91 秒，标准差为 1.88 秒。通过用户访谈，我们了解到用户对于这一内容生成时长存在不满。然而，采用流式生成的方式能够在一定程度上减轻用户在等待过程中的焦虑与不满情绪。

表 2：评分测试结果

Scenario	Mean	Mdn	SD	N
得分报告	4.50	4.50	0.58	110
错句提示	4.22	4.00	0.92	50
错句参考答案	4.82	5.00	0.53	50
非错句点评	4.42	4.50	0.57	50
非错句参考答案	4.55	5.00	0.57	50
Mean	4.50	4.60	0.63	62

随后，我们向负责批改这些作文的语文教师展示了系统的核心使用流程，并依据上述样本进行了人工综合评分并记录样本中出现的问题，详细的测试结果可见表 2。评分标准如下：

1 分-输出毫无价值；

2 分-勉强可以使用；

3 分-出现可接受错误；

4 分-满足需求且无误；

5 分-超出预期的卓越表现。

最后，我们对参与评分语文老师进行了访谈，收集其对系统的意见。

表 3：评分测试结果

Scenario	Mean	Mdn	SD	N
得分报告	4.50	4.50	0.58	110
错句提示	4.22	4.00	0.92	50
错句参考答案	4.82	5.00	0.53	50
非错句点评	4.42	4.50	0.57	50
非错句参考答案	4.55	5.00	0.57	50
Mean	4.50	4.60	0.63	62

从测试结果和访谈结论中我们可以得出以下一些结论：

（1）当前的大模型可以产出满足需求且无误的反馈内容，各场景综合评分高达 4.5 分，其中表现最好的场景高达 4.82 分，最差的场景也有 4.22 分。在访谈过程中，教师们对我们的系统主流程以及大模型生成的反馈内容给予了高度认可，认为其是对真人作文教学的有效补充。

（2）不同场景的输出存在一些问题，主要体现在三个方面。第一，输出可能仍包含错误，对学生造成错误引导；第二，输出可能包含不应包含的内容如在错句修改提示中出现正确答案，这将削弱导学学习效果；第三，输出长度可能过长或过短，这或对学习效果和用户体验造成消极影响。后续我们会采取使用优质本地数据微调、在不同场景试验更多不同大模型、交互上优化支持用户重新生成的流程等策略解决或部分解决以上问题。第四，关于点评、提示场景，该为学生提供怎样的支持颗粒度，缺乏一个想对具体的标准。

（3）针对不同学段及文体系统应提升其自适应能力。在初中阶段，学生主要练习记叙文写作，普遍存在的问题是作文结构不合理，难以把握文章重点，或所描写内容与主题不够贴切。因此，系统应对记叙文作文的结构进行细致评价。此外，通过与高中教师的访谈，我们也了解到为高中生提供服务时，系统需展现出更高水平的智能化与准确性，能够针对议论文在论据运用、论证结构及方法等方面存在的问题进行精准指导。

（4）通过对内部测试阶段获取的用户反馈进行分析，本研究揭示出用户对于反馈输出的风格偏好存在一定差异，这些差异主要体现在展示风格和内容结构方面。具体而言，某些用户倾向于接受结构化的分点陈述反馈，而其他用户则更偏好于分段呈现的口语化反馈。此外，部分用户倾向于启发式的反馈方式，而另一些用户则更倾向于直接接收具体的修改建议。为了更好地适应用户需求，系统的未来迭代计划通过前期的问卷调查和反馈评价来识别用户的目的，并据此捕捉用户的个性化偏好，进而自适应地调整提示词和反馈内容的输出策略。

（5）不同应用场景中，大型模型之间表现出明显的性能差异。以得分报告为例，ChatGLM_turbo 的表现显著优于 SparkV3.0。相较于 SparkV3.0 的泛化性输出，ChatGLM_turbo 能够提供更为具体、针对性强且具有启发性的结果。然而，在提示场景中，ChatGLM_turbo 因经常直接输出正确答案，其表现反而不如 SparkV3.0。因此，开展大模型在不同场景下的输出效果对比测试是十分重要的。

实验受限于资源也存在一些不足之处，如样本量过小、未检验导学系统对于学生写作能力提升的实证效果等。

5. 总结

本文提出了一种大模型驱动的一对一智能导学系统，为大型模型反馈生成内容的人机交互设计提供了新的参考。我们设计了实验，并对大模型生成的内容进行了小样本评分测试，结果表明大型模型可以生成符合导学需求的高质量反馈内容，且不同大型模型在不同场景下具有差异。

然而该系统仍需改进。总体而言，大型模型技术为智能导学系统反馈内容生成带来了全新的可能性。然而，如何设计更合理的人机交互场景以及提高反馈内容的价值性和可靠性，仍是我们未来研究和解决的关键问题。

6. 致谢

本研究受国家自然科学基金面上项目支持，项目批准号：62377017。

参考文献

[1] 张立山和胡祥恩.（2024 待出版）**《智能导学》**. 陕西师范大学出版社总社.

[2] 何克抗.(2017).对反馈内涵的深层认知和有效反馈的规划设计——美国《教育传播与技术研究手册(第四版)》让我们深受启发的亮点之二.**中国电化教育，05**,1-7+14.

[3] 蝉金郑, 少阳郭, 薇夏, & 绍光毛. (2023). 小花狮: 基于大语言模型的中文作文智能辅导系统. **Chinese/English Journal of Educational Measurement and Evaluation| 教育测量与评估双语期刊, 4(3)**, 4.

[4] 李海峰 王炜.(2022).学习分析研究：基本框架和核心问题——2022 版《学习分析手册》的要点与思考.**开放教育研究,06**，60-71.

[5] 何文涛,路璐,周跃良 周睿.(2023).智能时代人机协同学习的本质特征与一般过程.**中国远程教育，03**,12-20.

[6] 夏雪莹 马玉慧.(2021).智能导学系统中反馈的设计研究.**开放学习研究(03),**1-8.

[7] Mathan, S. A., & Koedinger, K. R. (2002, June). An empirical assessment of comprehension fostering features in an intelligent tutoring system. In International Conference on Intelligent Tutoring Systems (pp. 330-343). Berlin, Heidelberg: Springer Berlin Heidelberg.

[8] Shute, V. J. (2008). Focus on formative feedback. Review of educational research, 78(1), 153-189.

[9] Wiggins, G. (2012). Seven keys to effective feedback. Feedback, 70(1), 10-16.

[10] Page, E. B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. Journal of educational psychology, 49(4), 173.

[11] Hattie, J., & Timperley, H. (2007). The power of feedback. Review of educational research, 77(1), 81-112.

[12] Lin, L., Ginns, P., Wang, T., & Zhang, P. (2020). Using a pedagogical agent to deliver conversational style instruction: What benefits can you obtain?. Computers & Education, 143, 103658.

数智技术赋能 5E 语文活动课生成性教学模式的构建与应用

Construction and Application of Generative Teaching Model for Chinese Activity Class Empowered by Data Intelligence Technology

茂名市第一中学　　陈良湾　　邮箱：690603963@qq.com

【摘要】 在智慧学习环境下，建构主义学习理论的“5E”模式能融合教学中师生的认知过程、活动过程数据的流转和运用、智能化学习平台的支持等要素，提高学生探究性学习的体验和质量。对于语文活动课而言，该模式能构建出包含五个步骤的生成性教学的路径，分别是：问题驱动--实践探究-汇报输出-思想碰撞--总结反思。实践表明：该教学模式融合数智技术，具有显著效果，为技术与教学融合提供了可借鉴的经验。

【关键词】 融合数智技术；“5E”；语文活动课；生成性教学

Abstract: *In the smart learning environment, the “5E” model of constructivist learning theory can integrate the cognitive process of students and teachers in teaching, the data flow and application of the activity process, and the support of the intelligent learning platform in teaching, improving the experience and quality of students' inquiry-based learning. For the Chinese activity class, this model can construct a generative teaching path that includes five steps: problem-driven, practice-inquiry, report-output, idea-collision, and summary-reflection. Practice shows that this teaching model integrates data intelligence technology and has significant effects, providing a reference experience for the integration of technology and teaching.*

Keywords: integration of data intelligence technology, "5E", Chinese activity class, generative teaching

1. 问题的提出

在“智慧教育”时代，语文教育更强调素养的形成。生成性教学是在弹性预设基础上，师生充分交互，不断调整教学活动和行为，共同建构并形成新的信息、资源的动态过程，以实现教学目标和创生附加价值。(谢幼如等,2016)。但目前学生在活动课中常存在积极性不高、综合素养能力不足、情感体验淡薄等问题。因此，融合数智技术的“5E”语文活动课借助“云-脑-端”智能化学习平台，有助于探究课堂的生成性教学模式，提升语文活动课的教学品质。

2. 相关研究述评

2.1. 中学语文活动课的研究现状 语文活动课是综合性的教学活动，目前以大单元为依托。该研究多从教学实践角度进行，集中在活动课改革、活动课教学问题及策略等，如：《高中语文活动课新尝试——新课程背景下的语文活动课实践研究》认为语文活动课应具备“文本研习、问题探讨、活动体验”等三种形式（朱寅,2021）、《基于项目学习的高中语文活动类单元教学策略研究》提出了基于项目学习的活动类大单元的教学策略（孔玉敏,2023）。在数智赋能的研究方面，如《面向生成的智慧学习环境构建与应用——以电子书包为例》探究电子书包对语文阅读能力的影响（谢幼如，2016）、《智慧学习环境下高中语文群文阅读的方法路径初探》提出借助智慧课堂、数字图书馆等实现阅读渠道多样化（王蕊，2021）。

2.2. “5E”模式的课堂研究现状 建构主义下的“5E”教学模式对应的课堂环节是：Engage（参与）、Explore（探究）、Explain（解释）、Elaborate（迁移）和 Evaluate（评估）。目前，该研究集中在对该模式本体的研究、结合学科核心素养的研究、结合具体学科教学设计的研究等，还涉及深度教学、STEAM 教育理念等热点问题（王芳，2023）。研究表明，“5E”

模式与大单元教学的理念是匹配的。因为大单元设计也常以“主题—探究—表达”的方式进行项目型设计。其中的“主题”和“探究”包括了主题确定、情境创设、学习共同体构建、深度学习等，这和“5E”中的参与、探究、解释、迁移等环节是一致的。而大单元中的“表达”则包括展示、总结、评价等，这和“5E”中的评估环节也是一致的。

2.3. 生成性教学的研究现状 目前，国内外对生成性教学的研究既有理论研究也有实践研究。理论研究集中在对生成性教学内涵、特征、对预设与生成关系的理解。而实践研究集中在具体学科的生成性教学研究等。

3. “5E”语文活动课生成性教学模式的构建

3.1. 理论基础

3.1.1. 活动理论 活动理论提出注重学习过程中对象和动机的协商与转换、作为中介的制品对于应用情境的依赖性、以联合活动和实践作为分析单位(郑太年，2005)等特征的理论框架。

3.1.2. 生成学习理论 教育家威特罗克在《作为生成过程的学习》提出生成性学习理论，认为学习过程分注意和选择性知觉、主动建构意义、意义生成等几个阶段(何克抗，2010)。

3.1.3. 深度学习理论 美国理论家 Ference Marton 提出深度学习和浅层学习两个相对的概念(普拉伊特.K.杜塔，2005)。相比浅层学习，深度学习具有若干特征：注重批判理解、强调信息整合、着意迁移运用、面向问题解决等。

3.2. 模式构建 融合数智技术的“5E”语文活动课生成性课堂教学模式，以活动理论、生成学习理论、深度学习理论为指导，探究教学的有效性。它以“云-脑-端”智能平台为基础，利用 5G 网络、知识图谱、机器学习等软硬件设施，通过全过程的动态数据分析，推动线上线下一体化、虚拟现实一体化、教学评一体化的全景教学，增强课程内容与学生成长的联系。

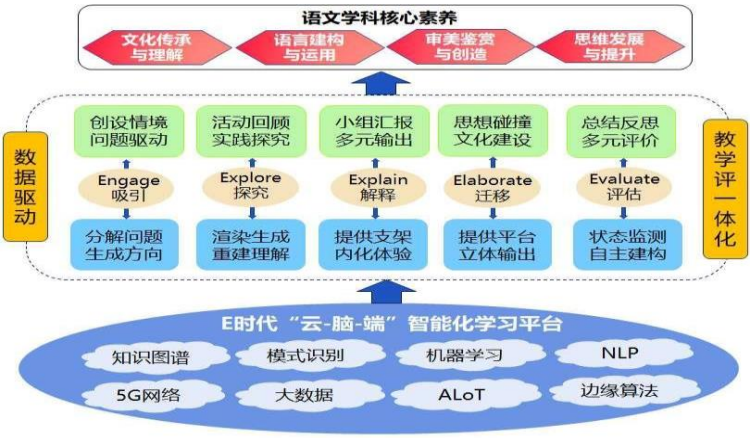


图 1：融合数智技术的“5E”语文活动课生成性课堂教学模式

在教学中，活动课按照认知过程分五个环节：创设情境，问题驱动；活动回顾，实践探究；小组汇报，多元输出；思维碰撞，文化建设；总结反思，多元评价。这五个环节符合深度学习特征，有效处理预设与生成、学生自主与教师指导、过程与评价的关系。与其对应，平台提供“分解问题，生成方向；渲染生成，重建理解；提供支架，内化体验；提供平台，立体输出；状态监测，自主建构”五大支撑，呼应“吸引-探究-解释-迁移-评估”五大环节。同时，数据驱动和教学评一体化贯穿教学全过程，共同达成语文四大核心素养。（如图 1）

4. “5E”语文活动课生成性教学模式的应用

4.1. 教学设计 大单元的任务是以真实而有意义的活动为基础设计，且通过问题链落实。如《家乡文化生活》展示课，设计了“如何展示-提出什么建议-有何收获-如何评价”等问题帮助学生梳理对“文化”的理解，进而培养学生的家国情怀。（如图 2）

因此，在活动课应设计“真情境”。“真情境”可以分成三种类型：一是现实的真实（真人真事）、二是可能的真实（换位体验）、三是虚拟的真实（角色代入）(梁昌辉，2022)。前

两者可由具体实践活动创设，而最后一种则可以充分发挥 VR、AR、MR、XR 等数智技术的优势，实现虚实空间弥合，使交互更加实时化和全景化。

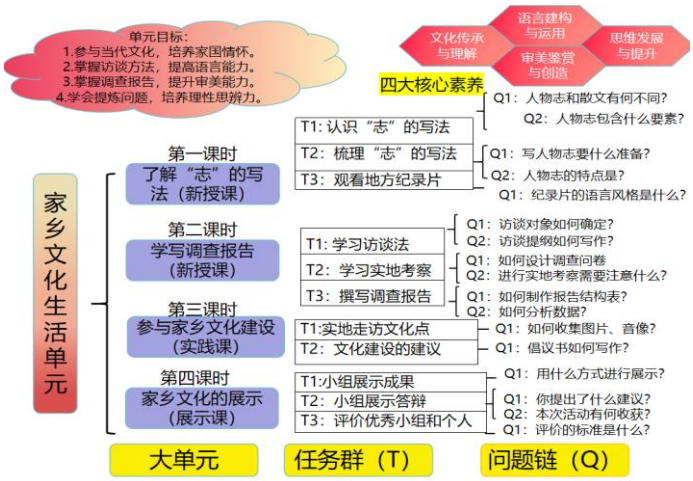


图 2：“家乡文化生活”单元整体目标及设计

4.2. 实施策略 下面以“走进岭南古郡，传承好心之风——《家乡文化生活》展示课”为例，按照语文活动课生成性模式的 5 个基本步骤，在单课时实施“5E”课堂，具体操作如图 3：

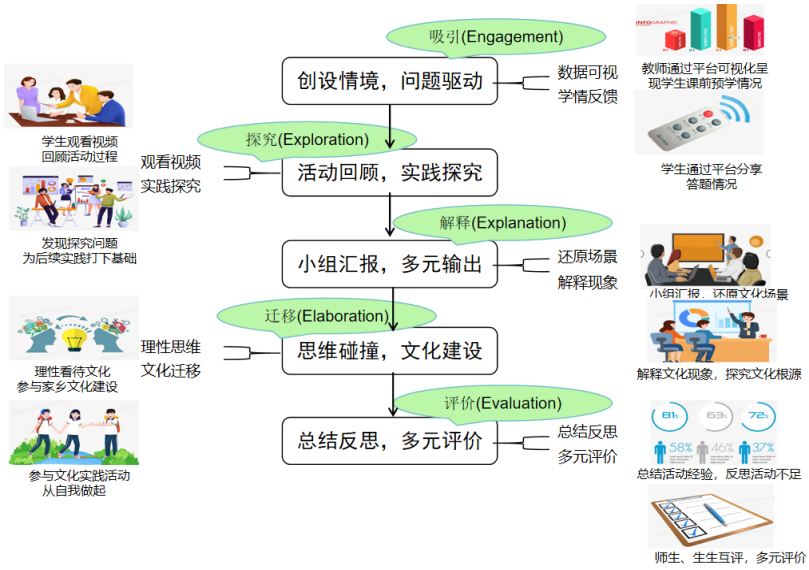


图 3：融合数智技术的“5E”语文活动课单课时教学流程结构

一是“创设情境，问题驱动”： 学生对家乡茂名的文化进行选择题作答，老师根据“轻课堂”数字平台数据，了解学生对家乡的熟悉程度，从而调整方案，实施精准教学。

二是“活动回顾，实践探究”：课前使用剪映工具，对活动中的图片、视频剪辑呈现。此外，借助 VR 头戴式显示器、虚拟环境构建技术等还原真实的探究情境。同时，应用 AI 技术进行动态学习资源推荐、学习状态监测等。

三是“小组汇报，多元输出”：通过交互式白板进行展示，支持实时控制、标注、拖放组合、智能绘图等功能，帮助展示者在预设的基础上，生动地展示教学内容。

四是“思维碰撞，文化建设”：小组把文化建设的建议表达出来，和不同的小组进行争鸣。师生充分交互，共同研究建设家乡文化的有效措施。

五是“总结反思，多元评价”阶段：教师对本课进行总结，学生评价展示成果，教师在“轻课堂”教学平台发起即时投票，鼓励生生评价，实现评价主体多元化。

4.3. 应用效果

本研究选取茂名市第一中学高一语文活动课作为个案进行实践探究。研究对象为 2023 级高一（7）和高一（8）班学生，其中高一（7）班应用本文所述模式，而高一（8）班没有。

为比较两班学习效果，研究以语文核心素养的四个维度为评价指标，采用李克特量表的五度量尺，制订了 8 个问题，然后从“完全不符合”到“完全符合”，采用-2、-1、0、1、2 的方式积分，根据双向评等量表得分率计算公式 $Fi=\Sigma a_j n_{ij}/2N$ 进行检验。两个班的测量数据如表 1 所示。

表 1 高中语文活动课是否应用“5E”语文活动课生成性教学模式的效果对比分析

语文核心素养	问题	得分率		语文核心素养	问题	得分率	
		应用班级	没应用班级			应用班级	没应用班级
语言构建与运用	提升口头表达	0.53	0.48	思维发展与提升	学会参与访谈	0.61	0.35
	提升书面表达	0.55	0.35		撰写调查报告	0.58	0.28
审美鉴赏与创造	收集家乡美图	0.81	0.77	文化传承与理解	了解家乡文化	0.71	0.43
	制作家乡视频	0.65	0.62		传播家乡文化	0.54	0.31

数据显示：（1）两个班级中各题得分 Fi 均大于 0，说明语文活动课有助提升学生语文核心素养。（2）应用本文所述模式的班级各题得分 Fi 均大于 0.5，且高于没有应用的班级，说明该模式对提升语文核心素养效果更加显著。（3）应用本文所述模式的班级对提升“语言构建与运用”“思维发展与提升”和“文化传承与理解”这三项核心素养的效果更加显著。

对比两种课堂模式，没有应用“5E”的课堂以展示为主，学生虽具备一定的参与度和探究性，但缺乏对情境的解释迁移和课堂的有效评价，更缺乏实地调研能力的培养。因此，学生对文化的理解停留在对家乡风物的审美层面，无法更好地达成“文化传承和理解”等目标。

5. 结语

本文探究“5E”模式下高中语文活动课生成性教学模式的构建，得出“问题驱动-实践探究-汇报输出-思想碰撞--总结反思”五个步骤，形成操作性较强的教学模式。相比其他模式，如活动单导学模式、活动体验模式、UbD 模式、PBL 等，“5E”模式的优势是教学的情境化和教学评的一体化。当然，“5E”模式如果能够更强调测试的闭环，将能更好地服务课堂，这一点可向 BOPPPS 模式借鉴。目前“5E”模式在理科应用广泛，其实该模式运用于文科课程也有其独特的价值，如语文的活动课、写作课和阅读课等。未来“5E”模式的研究继续向情境教学和智慧课堂转变，并且与计算机科学深度融合，呈现出更加宽阔的研究前景。

参考文献：

1. 谢幼如,杨阳,柏晶,李伟,郭琳科,倪妙珊.（2016）.面向生成的智慧学习环境构建与应用——以电子书包为例.华南师范大学学报(自然科学版),2016,48(1),126-132.
2. 朱寅.（2021）.高中语文活动课新尝试——新课程背景下的语文活动课实践研究.语文教学通讯·D 刊(学术刊),2021（8）,30-32.
3. 孔玉敏.(2023).基于项目学习的高中语文活动类单元教学策略研究.西南大学,2023(4).
4. 王蕊.（2021）.智慧学习环境下高中语文群文阅读的方法路径初探.数字教育.2021,7(01),56-59.
5. 王芳.（2023）.国内“5E 教学模式”的研究现状及演进走向——基于 CiteSpace 知识图谱软件的可视化分析.云南化工，2023（12），15-21.
6. 郑太年.（2005）.从活动理论看学校学习.开放教育研究，2005（2）,64-68.
7. 何克抗.（2010）.对美国“建构主义教学：成功还是失败”大辩论的述评.电化教育研究,2010（10）,5-24.
8. [美]普拉伊特.K.杜塔.《策略与博弈》，上海：上海财经大学出版社 2005 年版.
9. 梁昌辉.（2020）.切于实用，有益于生活——“实用性阅读与交流”任务群教学解读.语文建设,2022(5)(下半月),26-31.

基于微调大模型的教学视频辅助工具设计与研究

Design and Research of Teaching Video Assistance Tools Based on Fine-Tuned LLM

席少剑[1*]，吴子豪[1]，杨烨枫[1]，李锋[1]
[1]华东师范大学教育信息技术学系
[*] 51264108006@stu.ecnu.edu.cn

【摘要】 视频学习作为在线教育最初始的教学活动，需要让学生提高学习动机与效率。于是本文从深度学习的理论出发提出深思提问式学习（Thoughtful Query Learning, TQL）框架并结合视频学习中存在的缺陷，利用大模型微调技术设计出基于化学课程视频学习的智能导师聊天机器人-“小艾达”。它可以根据学生的提问判断是否回答以及如何回答。微调之后的大模型相比原模型在回答质量上提升了 22.78%，证明通过微调技术能够改善大模型在某个领域上的回答质量。

【关键词】 微调大模型；深度学习；聊天机器人；深思提问式学习

Abstract: *As the initial teaching activity of online education, video learning needs to improve students' learning motivation and efficiency. Therefore, this paper proposes a Thoughtful Query Learning (TQL) framework based on the theory of deep learning and combines the shortcomings in video learning. It uses large model fine-tuning technology to design an intelligent tutor chatbot based on video learning of chemistry courses - "Little Ada". It can determine whether to answer and how to answer based on the student's questions. After fine-tuning, the answer quality of the large model improved by 22.78% compared with the original model, proving that the fine-tuning technology can improve the answer quality of the large model in a certain field.*

Keywords: Fine-Tuning LLM, Deep Learning, Chatbot, Thoughtful Query Learning

1. 研究背景

在线教育中以慕课学习方式为主，虽然慕课具有高入学人数，但是完成率低(Steve Kolowich, 2013)。以 edX 平台上的第一门慕课“电路与电子学”为例，经计算得到的慕课完成率约为 4.7%(Breslow et al., 2013)。类似地，在国内，即便在北京大学，慕课学习者的通过率也不足 4%(蒋卓轩、 张岩和李晓明，2015)。造成辍学的原因有诸多种。为了降低学生在线学习中的辍学率，现有智能导师学习管理系统在跟踪学习者的注意力和集中程度发挥着关键作用（Kumar et al., 2023）。本设计从视频学习机制和体验出发，结合大模型微调技术，探索出能够辅助视频学习的学习工具。希望借此工具能够激发学生的学习动机从而进行深度学习。

为了提高学生在视频学习过程中的学习体验，本研究从三个要点驱动设计。要点一：智能导师针对问题的回答基于学生的认知。要点二：智能导师需要分辨出与学习无关的问题。要点三：尽可能减少智能导师回答内容的幻觉。

2. 理论依据

2.1 深度学习

深度学习(deep learning)理论最早由教育心理学家弗伦斯·马顿和罗杰·萨尔乔提出。他们根据信息获取加工方式，将学习分为浅层学习和深度学习。为了让学习者进行深度学习，有无数的学者进行了探索。例如比格斯提出深度学习过程模型(Biggs, Kember, & Leung, 2001)，发现学生的学习兴趣和学习动机可以激发基于真实情境的深度学习。沈霞娟在深度学习能力内

涵和特征的基础上提出了“三域六维”深度学习能力概念框架，三域分别为个人认知、人际协调、和自我监督(沈霞娟、武梦迪和冯锐，2023)。有的学者对 64 名教师的教学过程进行追踪调查，认为大部分的传统课堂教学还是停留在浅层学习(Smith & Colby 2007)。

因此深度学习的发生必须是教师和学生双向进行，即教师引发深度思考的问题，学生要根据问题积极思考。当现在有许多教学策略能够支持深度学习，例如项目式学习、探究式学习、基于问题式学习、MOOCs、适应性学习系统等(刘婧韡、刘一萌和顾小清，2023)。

2.2. 在线学习为深度学习提供支持

韩冬梅把深度学习当作目标导向，将深度学习目标划为三大维度，分别为认知领域、人际领域和自我领域(韩冬梅，2023)。在认知领域上强调高阶思维能力培养，在人际维度上强调团队协作、情感表达和有效沟通；在个人维度上体现反省思维、展现科学精神和学会学习。而探究社区理论为学生如何通过在线学习协作构建知识提供了理论基础，该理论的三大核心维度是认知存在、社会存在和教学存在(杨洁、白雪梅和马红亮，2016)。可以看出这两个框架有异曲同工之妙。认知领域对应着认知存在，人际领域对应着社会存在，自我领域对应教学存在。综合来看这说明无论是深度学习还是在线学习的构建，其目的都是一样的，即促进学习者进行有意义的学习。

3. 研究设计

有相关研究表明当观看教学视频有效暂停次数越多时，学习者的学习投入就越高。在暂停的时候，这意味着学习者需要休息、思考、提问或做练习等活动(王洪江等 2017)。在线协作学习中小组协作对话内容基本都是以问询答案和建立良好关系为主，很少会发生有意义的建构对话(Yang et al., 2023) 。所以本研究构想一个能够回答学生针对教学视频中问题的聊天机器人，叫做“小艾达”。他能够作为学生思考的脚手架进行答疑解惑。

3.1. 深思提问式学习（Thoughtful Query Learning,TQL）

结合深度学习的理念，我们提出了深思提问式学习教学法。当学生提出问题时，必然会对教学视频里面的知识进行检索、梳理并尝试理解。当理解出现阻碍时，便形成问题。学生及时提出这个问题，“小艾达”可及时回答。

3.2. 在 TQL 中体现三个要点的策略

为了鼓励学生提问，基于 TQL 的设计思想，该聊天机器人需要围绕上述三个要点开发。

针对设计要点一，虽然大模型能够与人类进行正常的交流，但是面对中小学生的疑问，大模型可能会说出超出学生已有认知基础的回答。为此可利用微调技术对大模型进行训练，让大模型学习到教师是如何引导学生思考，如何选择知识点。

针对设计要点二，“小艾达”会根据问题来判断是否回答。对于问题的判断，需要用到大模型的句子关系判断能力。为了增强这种能力，可以列出教学目标列表，让大模型将学生问题与教学目标匹配。

针对设计要点三，“小艾达”会根据用户提问内容对视频匹配，返回与提问内容最相关的时间戳，随后利用大模型信息整合能力做出回答。大模型由于具有幻觉的特性其回答不一定正确，但教学视频和时间戳是相对正确的数据，这一部分的内容可以保证回答的质量从而降低了幻觉给学习者带来的困扰。

4. 小艾达开发过程

4.1. 微调的过程

本次使用的基座模型为星火认知大模型 V1.5。微调方法是基于 LoRA 的指令微调。为了研究的简化，微调的数据选用于国家中小学智慧教育平台 2021 年度部优视频，只包括化学九年级上册中第一单元和第二单元的相关教材数据和师生问答数据。

关于微调训练集的收集。数据中的提问尽量来自于学生看完教材之后产生的疑问，而老师的回答也是针对教学目标进行回答。我们借助 GPT 生成问答训练集。教学目标的收集来自国家中小学智慧教育平台每个课例学习任务单中的教学目标。

在收集问答集的过程中需要人工的校准以确保问答集的质量。利用以上方法经过数据清洗与有效数据筛选共收集 37 条教学目标，3707 对数据集。将数据集其中的 3007 条用做训练集，700 条用做测试集。

4.2. 微调的结果

经过微调之后，选取了测试集其中的 79 个问题比对。对比模型是未经过微调的星火认知大模型 V1.5。经过人工比对，原模型回答地更好地有 18 道，两者回答质量差不多的有 25 道，微调模型回答更优的有 36 道。通过下方公式算出总体来说回答质量提升 22.78%。表 1 是笔者筛选出来具有代表性的问题比对。

$$\eta = \frac{Q_{微调} - Q_{原模型}}{Q_{总}}$$

总体来说，未经微调的模型会正常地做出解析，而微调模型后的模型不仅会正常做出解析，还会通过举一个形象的例子辅助理解。面对一些具体的问题，微调模型往往能做出详细地解释。例如面对问题“将块状磷酸钙研磨成粉末，这种变化属于哪一类？”，前者会直接回答出“物理变化”，而后者会做出详细地解释。微调模型在一定程度上改善了“答非所问”的问题，例如面对问题“为什么要在倾倒液体药品时，将瓶口紧挨着试管口？”微调模型给出的答案不仅详细而且具有针对性。

如图 1 所示，这是模拟学生进行视频学习的过程。左侧时要学习的教学视频-“氧气的制取”。学生在学习过程中可以向小艾达提问。比如学生可能会出于好奇问出“原神如何启动？”这样与学习无关的问题。面对这样的问题大模型要判断是否属于教学目标，不属于则返回“这些问题我不想回答呢”。这样可以很好地规避学生上课开小差进行无意义对话的情况。随后学生又问“试管口为什么倾斜向下”，这是属于教学目标的问题，微调大模型也很好地判断了这个问题并给予回答。首先它会返回相关的字幕时间戳，学生可以通过时间戳跳转相应的部分进行观看反思。其次还它会针对问题进行回答，这种回答是大模型生成的，具有一定的参考性，但是并不能保证百分百甚至还是产生一些幻觉的现象。而返回时间戳的设计可以减少幻觉的发生。本研究通过这样的设计，小艾达能够对学生的提问进行一次筛选，并且对问题做出合理的解答，从而引发学习者深度理解。

5. 结论

本文从视频学习体验不佳的角度出发，利用微调技术构造出来的“小艾达”能更好地体现智能导师的理念。并且针对视频学习过程中的三个要点，分别给出了不同的解决策略。旨在通过该设计，引发学生的深度学习。证明微调之后的大模型相比于没有微调的大模型，在回答风格、逻辑和准确率上都有所提升。利用教学目标检索策略和大模型的语义判断能力，可以很好地实现像该不该做出回答的复杂任务。最后可以采用传统算法对知识点做简单搜索从而降低大模型的幻觉现象。总之以上策略围绕大模型的能力和 TQL 学习框架进行智能导师的设计开发。但对于能否促进学生的深度学习还有待考究。

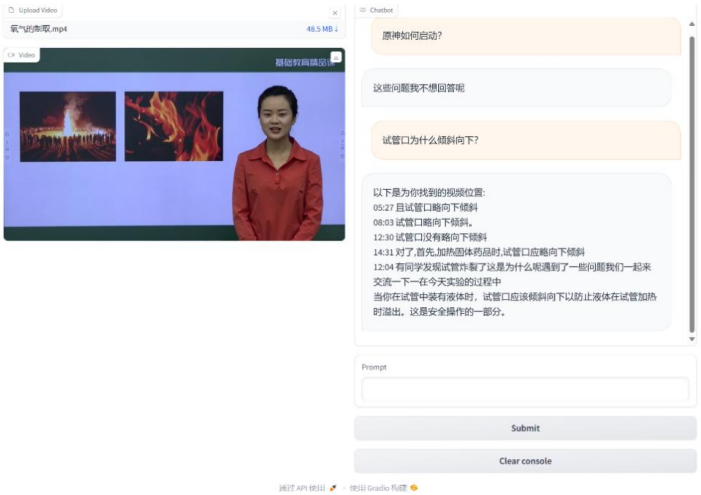


图 1 小艾达 demo 界面

参考文献

沈霞娟、 武梦迪和冯锐（2023）。《深度学习能力：概念框架、核心维度与测量体系》。**电化教育研究，44(12)，95–101**。

韩冬梅（2023）。《促进深度学习发生的混合式教学模式研究——以开放大学“信息技术与教育技术”课程为例》。**成人教育，43(12)**，61–67。

蒋卓轩、张岩和李晓明（2015）。《基于 MOOC 数据的学习行为分析与预测》。**计算机研究与发展，52(3)**，614–28.

刘婧韡、刘一萌和顾小清（2023）。《指向核心素养的智能化深度学习系统框架》。**开放教育研究，29(6)**，112–20。

王洪江、穆肃、黄洁和温慧群（2017）。《自主学习投入度实时分析方法及应用研究》。**电化教育研究，38(10)**，44–50。

杨洁、白雪梅和马红亮.(2016)。《探究社区研究述评与展望》。**电化教育研究，37(7)**，50–57。

Biggs, John, David Kember & Doris Y. P. Leung. (2001). The Revised Two-Factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology 71(1)*,133–49.

Breslow, Lori, David E. Pritchard, Jennifer DeBoer, Glenda S. Stump, Andrew D. Ho & Daniel T. Seaton. (2013). Studying Learning in the Worldwide Classroom Research into edX’s First MOOC. *Research & Practice in Assessment 8*, 13–25.

Smith, Tracy Wilson & Susan A. Colby. (2007). Teaching for Deep Learning. *The Clearing House: A Journal of Educational Strategies*, *Issues and Ideas 80(5)*, 205–10.

Steve Kolowich. (2013). Coursera Takes a Nuanced View of MOOC Dropout Rates. *The Chronicle of Higher Education.*

Yang, Kexin Bella, Vanessa Echeverria, Zijing Lu, Hongyu Mao, Kenneth Holstein, Nikol Rummel & Vincent Aleven. (2023). Pair-Up: Prototyping Human-AI Co-orchestration of Dynamic Transitions between Individual and Collaborative Learning in the Classroom. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23. New York, NY, USA: Association for Computing Machinery.

Kumar, M. V., Bhattacharyya, R., Dhillon, A., Manikonda, S. C. K., & Depuru, B. K. (2023) Artificial Intelligence Powered Global Learning Management System-AiTutor: A Behavioral Analytics Driven Artificial Intelligence Model Anchored In Tracking Learner-Tutor Interaction and Engagement.

远程生理信号测量方法及其在教育领域应用综述

A Review of Remote Physiological Signal Measurement Methods and Their Applications in Education

胡涛[1]，魏艳涛[2*]，肖光润[3]，潘东辉[4]，徐琦[5]，杨开[6]
华中师范大学人工智能教育学部数字教育湖北省重点实验室
yantaowei@mail.ccnu.edu.cn

【摘要】 远程光电容积描记法（rPPG）是通过摄像头等来捕捉由心动周期造成的皮肤颜色周期性变化，以实现心率、呼吸率和心跳变异性等心动周期相关生理指标测量的技术。近年来基于 rPPG 的生理指标测量方法取得了飞速发展，准确性和鲁棒性已得到了大幅提高。本文主要基于深度学习技术，对 rPPG 远程生理信号测量方法的研究进展进行了系统的梳理。从训练样本有无标记的角度，主要按照监督学习和无监督学习两类方法对代表性的方法和框架进行了总结，并依次分析其算法原理和优缺点。最后，我们重点对远程生理信号测量方法在教育领域的应用和发展前景给予展望。

【关键词】 远程光电容积描计法；非接触测量；深度学习；教育

Abstract: *Remote photoplethysmograph (rPPG) is a technology that captures the periodic changes in skin color caused by the cardiac cycle through a camera, etc., in order to achieve the measurement of cardiac cycle-related physiological indices such as heart rate, respiratory rate, and heartbeat variability. In recent years, rPPG-based physiological indicator measurement methods have made rapid development, and the accuracy and robustness have been substantially improved. In this paper, the research progress of rPPG remote physiological signal measurement methods is systematically sorted out mainly based on deep learning technology. From the perspective of training samples with or without labels, we summarize the representative methods and frameworks mainly according to the two categories of supervised learning and unsupervised learning, and analyze their algorithmic principles, advantages and disadvantages in turn. Finally, we focus on giving an outlook on the application and development prospects of remote physiological signal measurement methods in the field of education.*

Keywords: Remote photoplethysmograph, non-contact measurement, deep learning, education

1. 前言

心动周期心动周期(cardiac cycle)是指从一次心跳的起始到下一次心跳的起始心血管系统所经历的过程。通过对心动周期的分析，可以提取心率(heart rate，HR)、呼吸频率(respiration rate，RF)和心率变异性(heart rate variability，HRV)等重要的生理指标（牛雪松，韩琥，山世光，2020）。这些信息可用于估计和分析一个人的身体健康状况和判断情绪的变化，在教育领域可以有非常广泛的应用。

当前可信的高精度心动周期信号的测量途径、方法有心电图 (electrocardiogram,ECG)、脉搏血氧仪、光体积描记术 (photoplethysmograph, PPG)等(Shimazaki et al., 2021)，这些设备都会和被试者有直接的接触，被试者的自由活动也会收到很大的限制。皮肤电反应（GSR）一直是评估情绪唤醒和精神压力的重要心理测试，有学者验证了商用毫米波（mmWave）多输入多输出（MIMO）雷达进行非接触式皮肤电反应（GSR）传感的可行性(Xu et al., 2022),

同时毫米波雷达技术可以进行准确的血压(Ran et al., 2022)和(Antolinos et al., 2020)心率检测，但是毫米波雷达设备的成本较高，设备和部署需要较大的投资。在教育方面，有学者提出一种给学生测量学术压力的系统，但是由于使用时间的限制和需要将设备穿戴在被试学生的手上，学生的动作和互动受到极大的限制，所以无法在课堂上使用(Rodríguez-Arce et al., 2020)。因此等方法在实际的教育场景中有着诸多不便。

因而，发现一种无需直接接触、无需专业设备和专业人士操作的生理信号检测方法尤为重要。2008 年有学者首次提出(remote photoplethysmography,rPPG)，可以使用普通摄像头来进行生理信号的检测(Verkruysse et al., 2008)。这是一种利用视频图像来远程监测心率和其他生理指标的方法。它通过分析皮肤色彩的微小变化来检测血液流动，从而非接触地测量心率。这种技术可以提供舒适、方便且低干扰的健康监测，特别是在需要长时间监测或在特定情况下，如在睡眠监测或婴儿监护中，直接接触的传统方法可能不可行或不方便。而后有学者提出了一种测量脉搏率的简单而稳健的方法。可以直接从网络摄像头捕捉到的人脸图像中进行有效的脉搏率注册。通过适当的信道选择和主成分分析，可获得所需的信号(Lewandowskaet al., 2011)。图 1 显示了镜面反射和漫反射的皮肤反射模型(Wangal al., 2016)，其中只有漫反射包含脉动信息。解释了现有 rPPG 方法在脉冲提取方面的不同选择，进一步加深了远程光电容积脉搏波（rPPG）算法原理的理解。

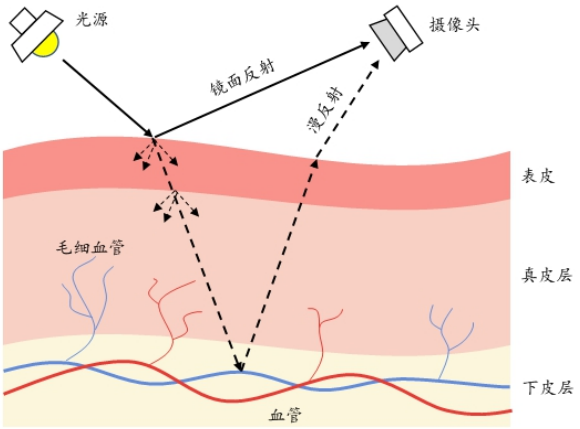


图 1 远程光电容积脉搏波（rPPG）原理

现有的基于 rPPG 的远程生理信号测量方法都包含 3 个主要的步骤，如图 2 所示。首先使用相机等设备对面部信息获取；随后，在获取到的人脸信息中，选取感兴趣的区域(region of interest,ROI)；进而可以从 ROI 区域中提取 rPPG 生理信号，最后经过进一步的处理和分析得到我们所需要的生理指标。

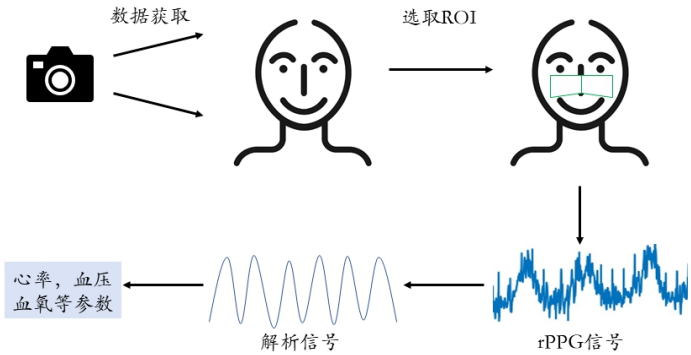


图 2 基于 rPPG 的生理参数检测流程图

之前之前的研究文章大多涉及传统的信号处理方法，有一些在深度学习出现之前就已经发表(McDuff al., 2015)，但是随着人工智能技术的高速发展，近年来开发的基于深度学习的方法显示出更有前途的结果。所以，本文的重点是对近年来的进行 rPPG 远程生理信号预测的深度学习方法的总结和比较，并更好地了解其优点和缺点。

但是进行 rPPG 数据集的采集需要用到昂贵的专业设备，目前公开免费的 rPPG 数据集数量也非常有限。我们以是否需要用到标注的数据集为依据，可以将基于深度学习的生理信号测量方法分为监督学习、半监督学习和无监督学习。但是目前能查阅到的基于半监督学习的 rPPG 信号测量方法较少，主要是 Hasan 等（2023）提出的 SrPPG 方法，这是一种使用异构、异步和噪声 rPPG 数据的新型半监督对抗学习框架。因此下文将主要介绍基于监督学习和无监督学习的 rPPG 测量方法。

2. 监督学习方法

在本节中，我们将详细介绍用于远程心率测量的监督学习方法。监督学习作为深度学习的一种学习方法，它涉及到使用标记的数据集来训练神经网络。目前，基于 2D 和 3D 的卷积经网络（CNN）是最主要的监督学习方法框架，如图 3 所示。

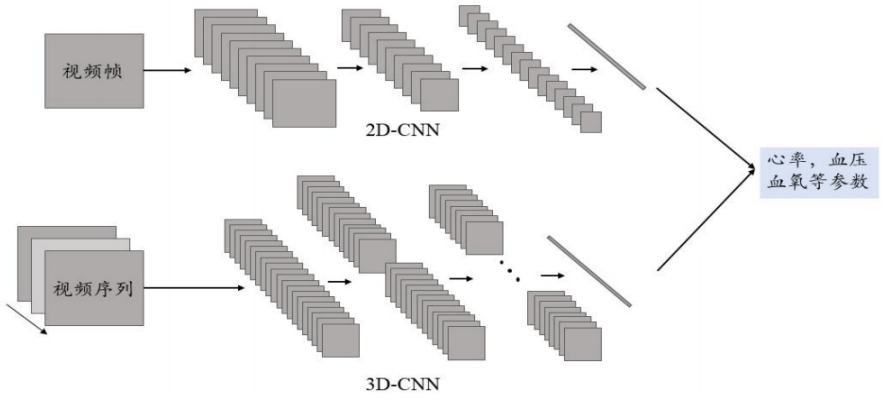


图 3 2DCNN 和 3DCNN 方法流程图

2.1. RPNet

Speth 等（2021）提出了一种名为 RPNet 的新型 3DCNN 架构，用于远程心率估计。

RPNet 是一种改进的 3DCNN 架构，该网络利用随网络深度增加的扩张时间卷积。这意味着在网络的不同层，时间卷积的扩张率（即卷积核元素之间的间隔）会增加，从而在不增加计算负担的情况下增加了网络的时间接收域。由于扩张率的增加，网络的时间接收域以指数级增长，这有助于捕捉视频数据中随时间变化的微妙特征，特别是对于心率信号的提取。最后采用最小化波形之间的负皮尔逊相关的优化方法，专注于时间回归问题，以提高心率信号估计的准确性。

RPNet 在保持计算效率的同时，能够提高远程心率估计的精度，使其在实际应用中更具可行性。尽管可以处理低帧率视频，但 RPNet 的性能可能仍受视频质量的影响，例如光照条件和面部遮挡。

2.2. EfficientPhys

尽管基于 3DCNN 的 rPPG 生理信号检测模型表现良好，但是在提高远程心率估计精确度的同时，模型的复杂性和计算量都会随之增加。Liu 等（2021）提出一种简化的、一站式解决方案，用于远程心率估计的神经网络架构 EfficientPhys。

EfficientPhys 采用一种 2D 视觉变换器 Swin Transformer(Liu al., 2021)，同时完全移除预处理模块，直接从原始视频帧输入到网络，并输出一阶导数的 PPG 信号。但加入了额外的组件，如定制的规范化模块、自注意模块、张量移位模块和 2D 卷积操作，以执行高效准确

的时空建模。定制规范化模块包括差分层和批量归一化层。差分层计算原始视频帧的时间轴上的第一前向差分，通过减去每两个相邻帧，类似于高通滤波，帮助降低全局噪声，同时保留 PPG 信号的微妙变化。

EfficientPhys 通过减少预处理步骤，简化了部署过程，提高了在不同平台上的应用灵活性。但其性能可能仍然受限于视频数据的质量，如光照条件和面部遮挡。

2.3. JAMSNet

Zhao 等(2023)提出一种名为 JAMSNet 的联合注意力和多尺度融合的远程脉搏提取网络。

JAMSNet 的算法框架包含两个部分，多尺度特征提取与融合网络：构建了一种基于轻量级 2DCNN 的高斯金字塔在不同尺度上提取特征，利用逐层注意力机制融合这些特征，基于信号强度增强脉搏信号的提取和分离。基于联合注意力的 rPPG 提取网络：包含通道-时间联合注意力（CTJA）和空间-时间联合注意力（STJA）模块。CTJA 专注于学习通道和时间维度上的注意力权重，而 STJA 则专注于空间和时间维度。这些模块有助于从视频中提取详细特征，对准确提取脉搏信号至关重要。

JAMSNet 在从噪声和运动伪影中分离脉搏信号方面效果显著，尤其适用于信号质量变化的场景。与之前的方法相比，在脉搏提取方面展现出更优越的性能，通过在不同数据集上的广泛实验得到验证。由于采用了多尺度方法和注意力机制，模型可能在计算上较为复杂，可能影响实时性能。

2.4. Shuffle-rPPGNet

Kuang 等(2023)提出了一种用于 HRV 测量的高效 3DCNN 模型，名为 Shuffle-rPPGNet。

Shuffle-rPPGNet 是一种基于 3D 卷积的高效网络结构：使用 ShuffleNetV2(Ma al., 2018)作为基础框架，构建高效的 3DCNN。为了提高网络学习长期特征的能力，还增加了 3D 全局上下文（3D-GC）块。此外，还在网络之后添加了上采样模块，以恢复时间维度到原始长度。其中 3D-GC 块用于捕捉视频检测任务中的长距离依赖性。上采样块通过 3D 转置卷积实现，用于恢复网络早期提取的特征。

该方法在 HRV 测量方面的性能显著优于传统方法和在此之前的深度学习方法。并且由于其网络复杂度低，适用于资源受限的设备。尽管模型在 HRV 测量方面表现出色，但非接触式 HRV 测量方法仍存在许多问题，例如对不同肤色和年龄组的泛化性。

2.5. LSTC-rPPG

Lee 等（2023）提出了一种创新的 3DCNN 结构和时间注意力精炼模块的深度学习方法，名为 LSTC-rPPG。

LSTC-rPPG 长短期卷积网络使用 3DCNN 基于沙漏结构对长短期特征进行建模。网络在编码阶段将短期特征逐渐压缩为长期特征，再通过解码阶段沿时间轴重建。同时采用跳跃连接帮助恢复压缩过程中丢失的短期信息。时间注意力精炼模块（TARM）可以解决编码和解码特征之间的时间信息不匹配问题。同时计算编码和解码特征之间的时间相关性，利用这些相关性来精炼编码特征。模型还引用一种频率尺度不变的混合损失，结合时间域和频率域损失，引导长短时特征的学习。通过尺度不变性简化深度神经模型的学习过程。

该方法在心率估计准确性上优于之前先进的方法，相比于其他深度学习方法，该方法无需数据增强即可实现更稳定和普适的性能。尽管性能优越，但在不同数据集和更广泛的场景中的泛化能力仍需进一步验证。

3. 无监督学习方法

在本节中，我们将详细介绍用于远程心率测量的无监督学习方法。无监督学习也是深度学习的一种学习方法，它涉及从没有标签的数据中学习模式和结构。由于目前公共 rPPG 数

据集的数量有限以及制作标注 rPPG 数据集的成本很高，所以发现一种不依赖标注的 rPPG 数据集的生理信号检测方法尤为重要，幸运的是，最近的研究发现，无监督学习是解决数据稀缺问题的一种有前途的解决方案，其中包含对比学习和非对比学习如图 4 所示。

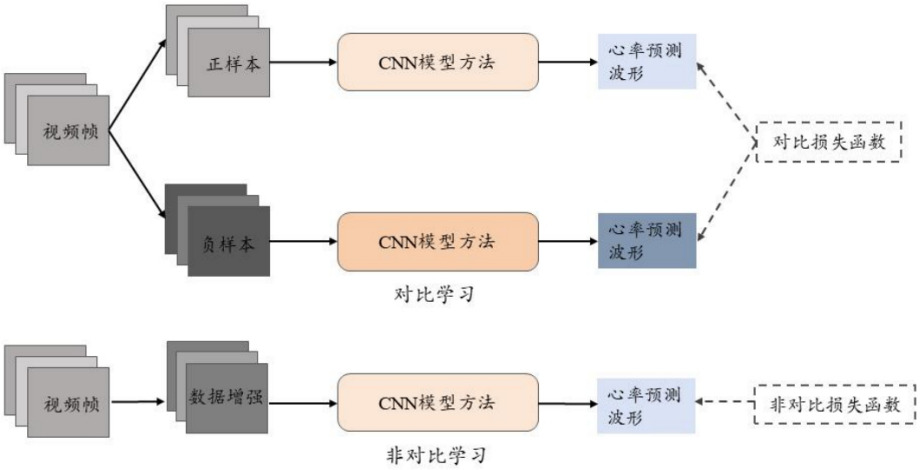


图 4 对比学习和非对比学习方法流程图

3.1. SLF-RPM

Wang 等（2022）提出了一种可以有效解决视频帧噪声和周期性变化的对比学习方法，用于远程生理测量。

该方法利用 Nyquist-Shannon 采样定理来特征化周期性的颜色变化，以此来提取 rPPG 信号特征。并且采用了 3D-ResNet 架构作为视频编码器，可以同时学习空间和时间信息。输入数据被转换成特征向量，然后通过多层感知器（MLP）获得最终的编码特征向量。在评估过程中，将移除投影头，并直接使用视频编码器中的特征向量作为生理信号的表示，来进行最终预测。文章还提出了一种新的损失函数，使用从数据增强中获得的伪标签来规范对比学习的训练过程，并处理复杂的噪声。

SLF-RPM 的优势在于可以在没有标签的情况下对模型进行预训练，同时它还具有良好的迁移能力，可以作为许多面部视频分析任务的预训练模型。但为了达到更好的性能，后续可能需要使用标签数据进行微调。

3.2. SimPer

Yang 等（2022）提出了一个用于学习数据中周期性信息的自监督学习框架。这个框架通过特定的数据增强方法和对比学习技术，来捕捉数据的固有周期性，而无需依赖标签数据。

Simper 引入了时间自对比学习框架，其中正样本和负样本是通过对同一输入实例进行周期性不变和周期性变化的增强来获得的。此外，它提出了周期性特征相似度的概念，用于在周期性学习的背景下测量特征相似度。最后，为了利用频率域中增强样本的连续性，SimPer 设计了一种通用的对比损失函数，该函数将经典的 InfoNCE 损失扩展为连续标签（频率）的软回归变体。

在不同领域的六个多样化周期性学习数据集上进行了广泛的实验，验证了 SimPer 与多数的自监督学习方法相比具有优越的性能。同时在数据效率、对偶然相关性的鲁棒性、减少训练数据和对未见目标的泛化方面显示出良好的特性。

3.3. Contrast-Phys

Sun 和 Li（2022）提出了一个称为 Contrast-Phys 的对比学习方法，该方法训练时不需要的地面真实信号。

Contrast-Phys 使用 3DCNN 模型从视频的不同时空位置生成多个 rPPG 信号，并使用对比损失来训练模型，其中来自同一视频的 rPPG 信号被拉近，而来自不同视频的信号被推远。

相比之前的无监督方法，该方法在五个公共 rPPG 数据集（包括 RGB 和 NIR 视频）上的能更优，且接近当前最佳的有监督 rPPG 方法。此外，与先前的无监督方法相比，Contrast-Phys 显示出更快的运行速度和更强的噪声鲁棒性。

3.4. SiNC

Speth 等人提出一种非对比性无监督学习方法用于从视频中提取生理信号(Speth al., 2023)。

SiNC 的网络模型的是一个 3D-CNN，不同的是模型的总损失函数是由三种损失（频带损失、稀疏性损失、方差损失）相加而来：L=Lb +Ls +Lv。频带损失（Bandwidth Loss）：该损失惩罚模型生成的超出预定义频率带宽限制的信号。计算方法为对信号的频率分量的功率进行累加，然后惩罚超出频带范围的部分。稀疏性损失（Sparsity Loss）：该损失旨在改进模型，防止其做出宽频带的预测。它通过惩罚频带内不接近频谱峰值的能量来实现。方差损失（Variance Loss）：为避免模型陷入简单的解决方案，该损失将功率谱密度的方差扩展到期望频带上的均匀分布。

这是首次将非对比性无监督学习应用于从视频中提取生理信号。尽管 SiNC 不需要监督数据，但模型的设计和训练可能仍然相对复杂。而且可能对特定的超参数（如频率限制）比较敏感。

3.5. Frequency-Inspired Self-Supervised rPPG Estimation Method

Yue 等（2023）提出了一种新颖的频率启发式自监督框架，用于基于面部视频的 rPPG 信号测量。

该方法采用对比学习的思想在数据增强阶段通过空间增强生成与原始视频样本具有相同信号频率的正样本。同时使用可学习的频率增强（LFA）模块生成具有不同信号频率的负样本。这些负样本在视觉外观上不会过度改变，但信号频率发生非线性变换。模型还包含一个局部 rPPG 专家聚合（REA）模块：估计来自不同面部区域的 rPPG 信号，并通过时空门控网络聚合它们。REA 模块分别为每个增强的视频样本（正或负）设计，以获得其 rPPG 信号。在损失函数方面，文章提出一系列频率启发式损失（如频率对比损失、频率比一致性损失和跨视频频率一致性损失），用于优化从输入的增强和邻域提取的 rPPG 信号。

该方法在多个基准数据集上的广泛实验表明，该方法显著优于现有的自监督方法，并与最先进的监督方法表现相当。由于模型使用自监督学习方法，可能在特定应用场景的泛化能力和复杂度方面需要进一步研究和验证。

4. 应用

目前，rPPG 技术在教育领域的应用还处于起步阶段。由于 COVID-19 导致的线上教学模式普及，教师面临着如何有效评估学生学习状态和效率的挑战。近期研究展示了 rPPG 技术的潜力：例如，Zhai 和 Li（2021）提出的系统能够通过分析学生视频中的 HRV，评估学生的注意力和心理压力水平。同样，Morales-Fajardo 等人（2022）利用 rPPG 技术结合机器学习分类器在教育环境中识别学生的学术压力和焦虑水平。这些研究表明，rPPG 技术能够帮助老师及时的监测学生的学习状态。

展望未来，rPPG 技术在教育领域的应用潜力巨大。Zhao 等人（2023）利用非接触检测和基于生理信号的检测将 rPPG 技术用于疲劳分类，实现了非接触式长期普适生物医学信号监测。同年该团队还提出了一种基于多模态多时标融合的压力检测系统，基于心电图和皮肤电活动信号的多源特征融合(Zhao al., 2023)。因此 rPPG 技术不仅可以用于监测学生的情绪和压力，而且为在自然课堂环境中环境中准确检测学习疲劳奠定基础。更进一步，rPPG 技术 可以用于定制个性化学习体验。例如，通过监测学生的生理反应，教育软件可以调整教学难度

水平或提供针对性的学习资源。这种方法能够创造一个更加互动和个性化的学习环境，促进学生的积极参与和学习效果的提升。

尽管如此，rPPG 技术在教育领域的应用仍然相对有限。这要求我们加强对此技术的关注和研究，以推动智能技术与教育的深度融合。通过这种融合，我们可以期待教学方法的革新，使教学和学习过程更高效、更简便。因此，未来的研究应着重于探索 rPPG 技术在教育领域的进一步应用，旨在实现一个更智能化、个性化的教学和学习环境。

5. 总结与展望

本文以促进 rPPG 技术在教育领域的应用为背景，全面回顾了利用 rPPG 技术结合深度学习方法来测量心动周期、心率等关键生理指标的最新进展。文章以是否需要标注的数据集，将基于深度学习的生理信号测量方法，分为监督学习和无监督学习，详细分析了算法的框架和创新点，揭示了各种方法的优势和局限性，总结见表 1。文章最后还讨论了 rPPG 技术在教育领域的应用前景，并对未来的研究方向提出了展望。

表 1 各方法总结

方法名称	类型	作者/年份	架构/模型	优势	限制
RPNet	监督学习	Speth 等/2021	3DCNN	计算效率高、提高心率估计精度	视频质量影响、光照条件和遮挡
EfficientPhys	监督学习	Liu 等/2021	Swin Transformer + 2DCNN	简化部署、无需预处理、高效准确的时空建模	视频质量影响、光照条件和遮挡
JAMSNet	监督学习	Zhao 等/2023	2DCNN	适用于信号质量变化场景	计算复杂性、实时性能
Shuffle-rPPGNet	监督学习	Kuang 等/2023	3DCNN	网络复杂度低、对设备要求低	肤色和年龄泛化性
LSTC-rPPG	监督学习	Lee 等/2023	3DCNN	无需数据增强	泛化能力弱
SLF-RPM	无监督学习	Wang 等/2022	3D-ResNet	良好的迁移能力	可能需要数据标签微调
SimPer	无监督学习	Yang 等/2022	自监督学习框架	性能优越、对偶然相关性的鲁棒性	泛化能力待验证、复杂度高
Contrast-Phys	无监督学习	Sun 和 Li/2022	3DCNN	运行速度快、鲁棒性强	视频质量影响、光照条件和遮挡
SiNC	无监督学习	Speth 等/2023	3DCNN	新颖的无监督学习方法、对信号进行完整建模	复杂的模型设计和训练、可能对超参数敏感
FISM	无监督学习	Yue 等/2023	自监督学习框架	频率启发式损失函数	泛化能力待验证、复杂度高

基金项目

国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），国家教师发展协同创新实验基地建设专项“基于视频数据的“人工智能+”研训学生情感演化规律分析方法”（项目号：CCNUTEIII-2021-19），教育部人文社科项目“基于

人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号CCNU22JC011）

参考文献

牛雪松，韩琥，山世光。2020。**基于 rPPG 的生理指标测量方法综述**。中国图象图形学报，25(11) ，2321-2336。

Antolinos, E., García-Rial, F., Hernández, C., Montesano, D., Godino-Llorente, J.I., & Grajal, J. (2020). Cardiopulmonary Activity Monitoring Using Millimeter Wave Radars. *Remote. Sens., 12*, 2265.

Hasan, Z., Faridee, A.Z., Ahmed, M., Ayyanar, S., & Roy, N. (2023). SrPPG: Semi-Supervised Adversarial Learning for Remote Photoplethysmography with Noisy Data. *2023 IEE International Conference on Smart Computing (SMARTCOMP)*, 25-32

Kuang, H., Ao, C., Ma, X., & Liu, X. (2023). Shuffle-rPPGNet: Efficient Network With GlobalContext for Remote Heart Rate Variability Measurement. *IEEE Sensors Journal, 23*, 15199-15209.

Lewandowska, M., Rumiński, J., Kocejko, T., & Nowak, J. (2011). Measuring pulse rate with awebcam — A non-contact method for evaluating cardiac activity. *Federated Conference on Computer Science and Information Systems (FedCSIS)* ,405-410).

Liu, X., Hill, B.L., Jiang, Z., Patel, S.N., & McDuff, D.J. (2021). EfficientPhys: Enabling Simple, Fast and Accurate Camera-Based Cardiac Measurement. *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 4997-5006.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 9992-10002.

Lee, J.S., Hwang, G., Ryu, M., & Lee, S.J. (2023). LSTC-rPPG: Long Short-Term Convolutional Network for Remote Photoplethysmography. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 6015-6023.

McDuff, D. J., Estepp, J. R., Piasecki, A. M., & Blackford, E. B. (2015). A survey of remote optical photoplethysmographic imaging methods. *IEEE Engineering in Medicine and Biology Society*, 6398–6404.

Morales-Fajardo, H.M., Arce, J.R., Gutiérrez-Cedeño, A., Viñas, J.C., Reyes-Lagos, J.J., Abarca-Castro, E.A., Ledesma-Ramírez, C.I., & Gonzalez, A.V. (2022). Towards a Non-Contact Method for Identifying Stress Using Remote Photoplethysmography in Academic Environments. *Sensors (Basel, Switzerland)*, *22*.

Ma, N., Zhang, X., Zheng, H., & Sun, J. (2018). ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. *Proceedings of the European Conference on Computer Vision (ECCV)*,116-131.

Rodríguez-Arce, J., Lara-Flores, L., Portillo-Rodríguez, O., & Martínez-Méndez, R. (2020). Towards an anxiety and stress recognition system for academic environments based on physiological features. *Computer methods and programs in biomedicine*, *190*, 105408.

Ran, Y., Zhang, D., Chen, J., Hu, Y., & Chen, Y. (2022). Contactless Blood Pressure Monitoring with mmWave Radar. *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 541-546.

Shimazaki, T., Hara, S., Okuhata, H., Nakamura, H., & Kawabata, T. (2014). Cancellation of motion artifact induced by exercise for PPG-based heart rate sensing. *IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 3216–3219.

Speth, J., Vance, N., Flynn, P.J., Bowyer, K., & Czajka, A. (2021). Unifying frame rate and temporal dilations for improved remote pulse detection. *Comput. Vis. Image Underst*,210.

Speth, J., Vance, N., Flynn, P., & Czajka, A. (2023). Non-Contrastive Unsupervised Learning of Physiological Signals from Video. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14464-14474.

Sun, Z., & Li, X. (2022). Contrast-Phys: Unsupervised Video-based Remote Physiological Measurement via Spatiotemporal Contrast. *European Conference on Computer Vision*,492—510.

Verkruysse, W., Svaasand, L. O., & Nelson, J. S. (2008). Remote plethysmographic imaging using ambient light. *Optics express*, *16(26)*, 21434–21445.

Wang, H., Ahn, E., & Kim, J. (2021). Self-Supervised Representation Learning Framework for Remote Physiological Measurement Using Spatiotemporal Augmentation Loss. *AAAI Conference on Artificial Intelligence*, 36(2),2431–2439.

Wang, W., den Brinker, A. C., Stuijk, S., & de Haan, G. (2017). Algorithmic Principles of Remote PPG. *IEEE transactions on bio-medical engineering*, *64*(7), 1479–1491.

Xu, X., Zhang, D., Chen, J., Wu, Z., Sun, Q., & Chen, Y. (2022). Contactless GSR Sensing Using mmWave Radar. *IEEE Sensors Journal, 22*, 24264-24275.

Yang, Y., Liu, X., Wu, J., Borac, S., Katabi, D., Poh, M., & McDuff, D.J. (2022). SimPer: Simple Self-Supervised Learning of Periodic Targets. *ArXiv, abs/2210.03115*.

Yue, Z., Shi, M., & Ding, S. (2023). Facial Video-Based Remote Physiological Measurement via Self-Supervised Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*, 13844-13859.

Zhao, L., Niu, X., Wang, L., Niu, J., Zhu, X., & Dai, Z. (2023). Stress Detection via Multimodal Multitemporal-Scale Fusion: A Hybrid of Deep Learning and Handcrafted Feature Approach. *IEEE Sensors Journal, 23*, 27817-27827.

Zhao, L., Zhang, X., Niu, X., Sun, J., Geng, R., Li, Q., Zhu, X., & Dai, Z. (2023). Remote photoplethysmography (rPPG) based learning fatigue detection. *Applied Intelligence, 53*, 27951 - 27965.

Zhai, K., & Li, X. (2021). HREmo: A measurement system of students' studying state in online group class based on rPPG technology. *Journal of Physics: Conference Series*, 1976.

Zhao, C., Wang, H., Chen, H., Shi, W., & Feng, Y. (2023). JAMSNet: A Remote Pulse Extraction Network Based on Joint Attention and Multi-Scale Fusion. IEEE Transactions on Circuits and Systems for Video Technology, 33, 2783-2797.

ChatPoet: 基于检索增强生成的诗词对话模型

ChatPoet: Poetic Dialogues Model based on Retrieval-Augmented Generation

崔咚咚[1]，魏艳涛[1*]，肖光润[2]，徐琦[1]，董银锋[1]
[1]华中师范大学人工智能教育学部数字教育湖北省重点实验室
[2]湖北文理学院机械工程学院
yantaowei@mail.ccnu.edu.cn

【摘要】 大语言模型的引入为教育领域带来了前所未有的发展机遇。但是通用领域的大模型在教育、诗词领域的对话能力仍不足，其主要原因是模型缺乏专业的知识。基于 RAG 技术，我们设计出 ChatPoet 诗词大模型，通过构建外部知识库为 ChatGLM3-6B 模型提供知识，专注于提高外部知识库的召回率。在构建知识库的过程中，我们提出的大模型文档摘要和长文档分块后信息头两种方法，提高 RAG 过程中非对称式语义检索的效率；利用大模型的自查询优化问题，提升大模型的诗词对话能力。我们进行了对比实验和消融实验，验证了知识库和知识库的组件对基座模型的诗词能力是有效的。

【关键词】 大语言模型；检索增强生成；诗词对话；向量知识库；学习助手

Abstract: *The introduction of large language models has brought unprecedented development opportunities to the field of education. However, the dialogue capabilities of large models in general domains are still insufficient in the fields of education and poetry, mainly due to the lack of professional knowledge of the models. Based on RAG technology, we have designed the ChatPoet poetry large model, which provides knowledge for the ChatGLM3-6B model by constructing an external knowledge base. During the construction of the knowledge base, we have proposed two methods for the large model document summary and the long document information header after segmentation, which improve the efficiency of asymmetric semantic retrieval in the RAG process; by utilizing the self-query optimization problem of the large model, we enhance the poetry dialogue capabilities of the large model. We have conducted comparative experiments and ablation experiments to verify that the knowledge base and its components are effective for the poetry capabilities of the base model.*

keywords*:* Large language model; RAG; Poetry dialogue; Vector database; Study assistant

基金项目：国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），国家教师发展协同创新实验基地建设专项“基于视频数据的“人工智能+”研训学生情感演化规律分析方法”（项目号：CCNUTEIII-2021-19），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号 CCNU22JC011）.

1. 引言

2023 年，生成式大型语言模型，特别是以英文为主要语言的 ChatGPT（Ouyang et al.，2022）、GPT-4、LLaMA，以及以中文为主要语言的文心一言、ChatGLM、通义千问等（Zhao & Zhou et al., 2023），引起了广泛关注。这些模型在各行各业中的应用为社会带来了创新的变革。在

教育领域，随着学生对个性化教学资源的日益多样化，传统的课堂教学需要引入大语言模型，以为学生提供更个性化的教学辅助。在以 PPT 为主要信息载体的诗词教学课堂中，教师的教学设计通常围绕 PPT 内容展开，有关古诗文知识、诗词鉴赏技巧、古文释义等方面的教学往往局限于固定的课程内容。诗词领域的大模型接入课堂后，学生可以在设备上与诗词大模型进行个性化聊天，获得课堂外有关课程内容的额外知识（刘邦奇等，2024）。并且，教师可以利用大语言模型设计新的教学活动，如虚拟诗词创作、历史人物角色扮演等，这些活动可以增强学生对古诗词的兴趣，提高他们的创造力和鉴赏思维能力（涂竞玉，2023）。因此，一个聚焦于古诗词的聊天机器人，对于提升课堂教学效果具有重要意义。

网络上有许多开源的大型语言模型，它们主要涵盖通用领域的知识。这些模型在通用领域的对话效果表现出色，但在特定领域，如医疗、电商客服、法律等领域，模型的对话能力可能并不强。特别是在与模型进行诗词领域的对话时，可能会出现答非所问或产生不准确答复的情况，即“幻觉”问题（Huang et al.，2023）。究其原因，大语言模型的训练语料往往是网络上公开的，通用领域的预料，当其应用于特定领域时模型缺乏相应的知识。为了缓解幻觉问题，许多研究者聚焦于提升通用大模型的专业能力。例如，TransGPT 和法律 GPT 等模型通过构建专用领域的指令语料，对开源的大型模型进行微调，从而为模型注入专业知识。这种方法相比于从零开始训练特定领域的模型，可以节省大量的计算资源。然而，这种方法也存在不足之处：微调大型模型仍然需要相当的计算资源；一些陈述型领域的文档难以转化为问答形式的指令语料，这限制了某些领域，特别是古诗词领域大型模型的发展。此外，指令语料具有时效性，模型所获得的静态知识不会随外部信息的更新而更新。

目前，还仍未有研究者对诗词大语言模型进行专门研究。网络上已经有许多公开的诗词数据集，例如 Chinese-Poetry 和古诗文网，然而，对于中国古诗词文档这样结构化和半结构化内容混杂的文档，如何将其应用于大模型，训练一个用于中小学生语文课堂交互的学习助手，是一个亟需解决的问题。一方面，网络上有许多公开的中国古诗词数据集，但这些数据集质量参差不齐，多数诗词只有原文没有译文和赏析，影响到模型的鉴赏能力；另一方面，结构化和非结构化的陈述型诗词数据难以构建指令数据集，无法对模型进行指令微调。

本研究的主要贡献如下：

1. 构建了包含陈述性知识的向量库：爬取了国内公开的中国古诗词数据，使用 BGE 编码模型形成了一个属性尽可能齐全的诗词向量知识库；设计长文档分块信息头和大模型文档摘要两种方法增强了向量库的非对称检索能力。

2. 提出 ChatPoet 诗词大模型：使用 RAG 技术方案和向量知识库增强了基座的诗词对话能力，如图 1 所示。提出自查询增强和元数据过滤方法提升相关知识的召回率。进行对比实验和消融实验验证了我们提出的方法是有效的。

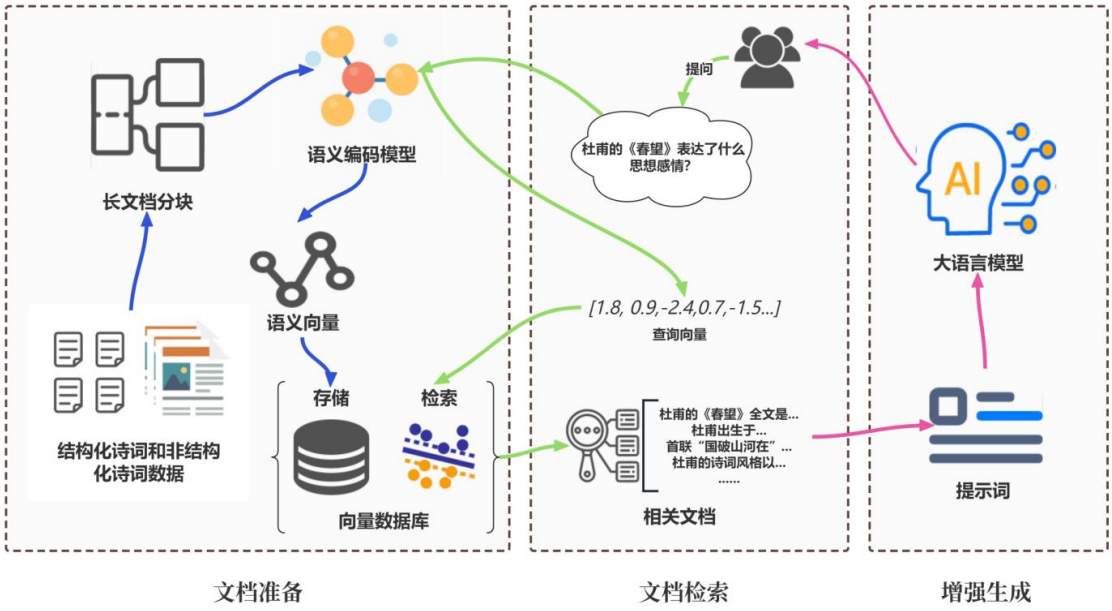


图 1 ChatPoet RAG 流程

2. 相关工作

2.1. 生成式大语言模型

生成式大型语言模型（Generative Large Language Models, GLLMs）接收文本输入并生成续写内容。这类模型本质上是通过上文内容预测下一个字符来生成文本。在训练过程中，模型被输入自然语言文本，并利用极大似然估计方法调整其参数以拟合数据集。在推理阶段，模型每生成一个字符，就会将其添加到已知的文本末尾，以此为新的上文继续生成下一个字符。这种机制就是“自回归语言模型”（Zhao & Zhou et al., 2023），其中 GPT 系列（Radford et al., 2018）是典型的代表。由于这些模型依赖于上文内容来生成下文，因此如何设计有效的上文以促进模型对语义的充分理解，成为了研究界的一个热点话题，即所谓的 Prompt 工程（Liu & Hayashi et al., 2023）。大模型能够从 Prompt 中提取信息以回答问题，这是 RAG 能够实现的基础。2023 年，THUDM 发布了 ChatGLM-6B（THUDM，2023）模型，拥有约 60 亿参数，因其出色的中文通用对话能力和小巧的体积，业内许多研究者使用 ChatGLM 作为垂直领域大模型的基座模型。本研究提出的 ChatPoet 选择了最新的 ChatGLM3-6B 为基座对话模型。

2.2. 检索增强生成

目前，预训练的大语言模型面临的一个显著问题是其训练集的时限性，这导致模型获取的知识无法更新。这一局限使得用户无法与大模型进行时间敏感的交互，例如询问当前天气或当日的新闻内容。针对这一问题，研究者提出了检索增强生成（Retrieval-Augmented Generation, RAG）技术（Gao, 2023）。与指令微调相比，RAG 将特定领域模型所需的知识存储在外部数据库中。当用户向模型提出问题时，RAG 会根据问题的语义从数据库中检索相关内容，并将其作为用户提问的上文送入大模型，以便大模型在回答问题时参考。这种方法的实质是将模型所需的知识与模型的参数权重分离，仅利用模型抽取知识以回答问题的能力。通过 RAG 技术，用户能够在其特定领域的文档与大模型之间进行交互，例如 ChatPDF（Panda et al., 2023）。RAG 的一个核心挑战在于解决用户提出的问题和陈述性知识之间的匹配效率，即非对称语义检索。本研究的核心旨在增强诗词知识库的非对称检索能力，从而间接提高相关知识的召回率。

2.3. 语义检索

在 RAG 聊天系统中，最主要的问题是如何从用户的提问中抽取语义信息，在外部文档库中匹配对应的知识。如何在外部库中存储知识以及如何检索相似知识是当前研究焦点。

在检索知识上，RAG 通常将知识转化为向量，利用向量相似度来匹配知识。向量相似度检索方法借助深度学习模型强大的表征学习能力（BERT 等模型），把文本的知识编码为向量。编码向量的特点是越相似的文本语义，得到的向量相似度越高。

在存储知识上，RAG 要求将知识向量存入向量数据库。与传统的 SQL 数据库相比，向量数据库存储的数据是数字向量，能方便进行批处理和计算。ChatPoet 使用 Chroma 作为向量数据库。检索时，Chroma 接收一个查询向量 V 和一个匹配数量 k，返回与向量 V 最接近的 k 个向量和对应的文本文档。这 k 个文本文档与向量 V 代表的文档具有最高的语义相似度。

由于语义检索局限于语义相似性，在 RAG 框架中，语义检索分为对称式语义检索和非对称式语义检索。前者要求用户的提问和检索的文档具有相同的语义，而后者更适用于问答形式。在陈述性知识为主的诗词库中，问题和答案不具有相似的语义，而是问题和知识的关系（Thakur & Reimers et al.，2021）。在这种情况下，如何利用问题检索出相对应的知识，是本研究重点解决的问题。

3. 方法论

3.1. 数据收集与处理

ChatPoet 的知识库包含了从互联网上开源的中华古诗词数据库 Poetry 以及古诗文网爬取的诗词赏析数据。经过去重处理后，共收录了 72,417 篇著作、3,154 位作者以及 5,764 个名句赏析。诗词数据集包含了翻译、赏析、注释、全文、作者等五个属性，涵盖了从先秦至清代的常见古诗词，包括《唐诗三百首》、《宋词三百首》等经典作品；作者信息包括简介和详情，详细记录了每位诗人的生平事迹；名句赏析则摘录自诗词，通常为一些流传广泛、备受推崇的佳句。将离散的样本转换为 CSV 列式数据后，得到的数据统计如表 1 所示。

表 1 诗词数据集详情

数据集	样本数	平均字数	包含属性
诗词著作	72417	873	全文、赏析、翻译、注释、背景、作者、朝代
诗人简介	3154	1068	姓名、朝代、简介、详情
名句赏析	5764	24	名句、作者、题目

3.2. 构建向量知识库

利用清洗后的数据集后构建向量知识库。在构建结构化诗词数据的向量知识库过程中，我们提出了“信息头”、“大模型摘要”两种方法提升 RAG 过程中非对称检索的召回率。

3.2.1 长文档分块信息头

大语言模型的上下文窗口有限，为了在召回知识时获得更加专一和精简的结果，我们要对过长的文档进行切割，以确保同一知识在一个文档块中，不同的知识在另一个文档块中。在构建 ChatPoet 的知识库过程中，我们使用了语义编码模型 BGE 和 LangChain 文本分割器，尽量将知识块的大小控制在 384 字左右。

在结构化诗词赏析分块后，某些文档块可能会失去一些所属信息，导致召回率降低。为了提升 ChatPoet 的知识检索效率，我们在分块时保留了原文档重要的所属属性，即“信息头”，这些属性包括诗词名、诗人、朝代等，并在结构化的 CSV 文件中保留。在分块后，这些属性被拼接到每个文档块的头部，以标注文档块所属。此外，为了确保文档块中的知识相对于原

文档是有序的，我们在分块后使用了递增的序号对每个文档块进行标注，以确保召回的知识是有序的，保证大模型能够获取有序的知识。

3.2.2 大模型文档摘要

用户与大模型的交互多种多样，许多情况下用户的提问是抽象的，广泛的，而一个文档块里的知识是具体的，这是非对称语义检索的核心问题。例如，用户提问“杜甫在安史之乱时期诗词特点是什么？”，而知识库中没有总结性的回答，只有零散的事实性的知识片段，召回的知识也难以匹配用户的问题。因此，我们设想为每一个文档片段形成一个摘要和关键词。如图 2 所示，摘要从抽象角度总结文档的内容，关键词提取文档的关键信息。构建向量知识库时，对每一个文本文档，先经由大模型，利用设计好的提示词获取文档的摘要和关键词。将摘要和关键词拼接到文档的前方，并且关键词作为文档的元数据（Metadata）存入数据库，以提供高效检索。在本研究中，我们使用 ChatGLM3-6B 基座模型为文档生成摘要和关键词。

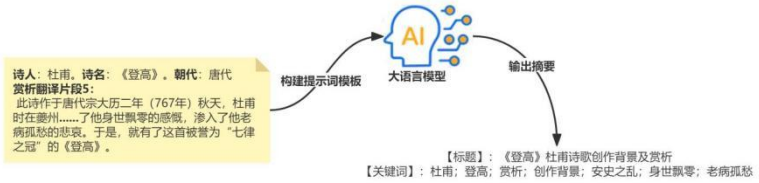


图 2 大模型文档摘要

3.2.3 向量数据库存储

经过上述步骤处理后，得到约 265,000 个长度在 512 左右的文档。在单台 3090ti GPU 上使用 BGE-base-zh-v1.5 语义编码模型每个文档转换为 768 维向量，过程大约需要 30 分钟。数据大小约为 1.325GB。使用 LangChain 语言链，把向量数据存入 Chroma 向量数据库，以供检索。构建数据库的流程如图 3 所示。

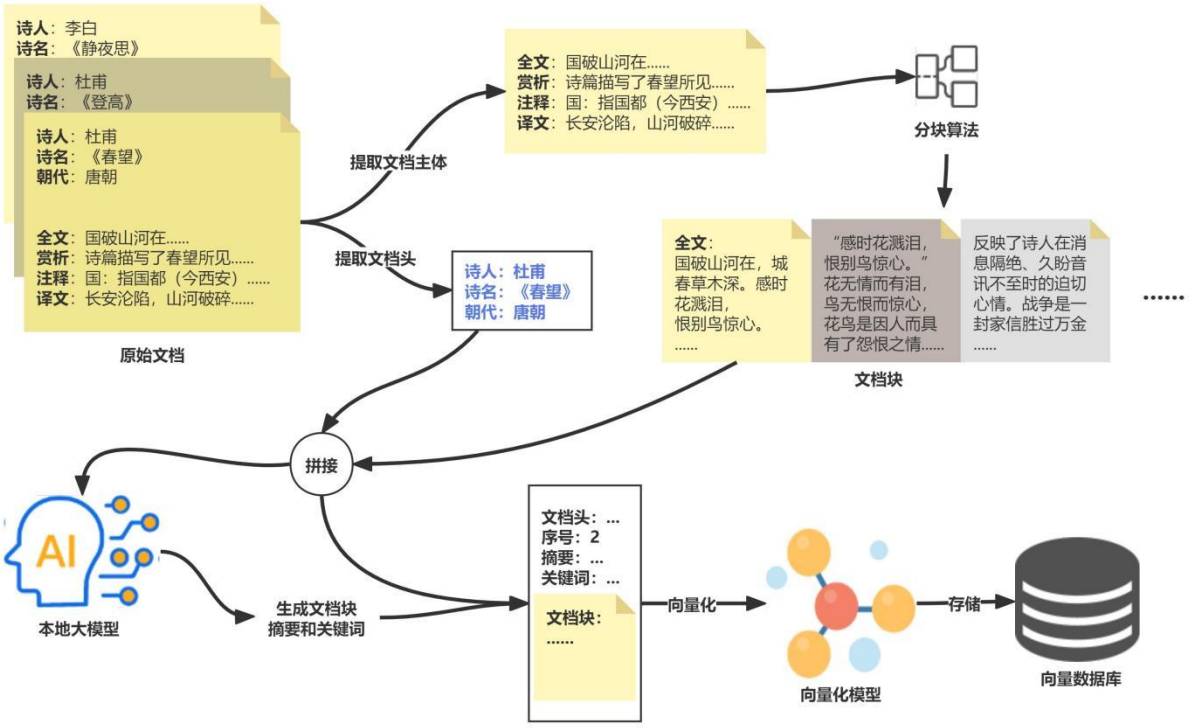


图 3 构建知识库总流程

3.3. 基于大模型的诗词交互

3.3.1 部署本地大模型

我们选择 ChatGLM3-6B-32k 作为 ChatPoet 的基座模型。相比 Llama2-7B，前者有更小的体积和更强的中文理解能力。6B 模型拥有约 60 亿参数，能在 16GB 显存的设备上半精度部署，在 6GB 显存上 4bit 推理。在一些多媒体教室也能部署。32k 表示模型能接收 32k 长度的上下文，约 21k 个汉字，能尽可能多的装入知识。在本研究中，我们在 3090Ti 上部署了 ChatGLM3-6B-32k 模型和知识库，进行诗词聊天。部署使用深度学习框架是 Torch2.0 和 LangChain 大模型框架。在用户与模型交互中，我们提出两种方法：自查询增强和元数据过滤。详细过程见图 4。

3.3.2 自查询增强

用户在进行诗词对话时，可能会提出一些复杂的问题，例如：“杜甫为什么在人生后期经历颠沛流离的生活？”。在有限的知识库中，检索器可能检索不到这个问题的关键知识，如“安史之乱”。检索器在检索问题相关的知识时，仅仅将问题编码为语义向量，再寻找与此相关的文档。自查询增强（Self-Query Augment）通过大语言模型，把一个复杂的查询转换为多个简单的，同义的查询。检索器利用多个同义的查询进行检索，避免了单一问题的偶然性，提高了查询效率。同时，自检索也会获得问题的相关资料，原问题经由大模型分析后，得到关键词“安史之乱”，检索器根据关键词查询。与“安史之乱”相关的文档，已经在构建知识库阶段由大模型整理出其关键词。自查询增强使得文档的关键词与问题的关键词匹配，从而提高非对称检索的效率。

3.3.3 元数据过滤

元数据过滤（Metadata Filtering）合理利用文档的头信息，在构建知识库阶段，每一个向量文档附带一个元信息，类似于分类标注。例如一个文档所属的诗人、诗名和朝代。在自查询阶段，使用设计好的提示词和大模型将一个问题可能包含的元数据筛选出来，得到问题相关知识所属的元数据分类。然后送入检索器，仅在这个分类去查找相关文档。由于检索器为需要检索的向量设置了范围，寻到相关知识的可能性会大大提高。

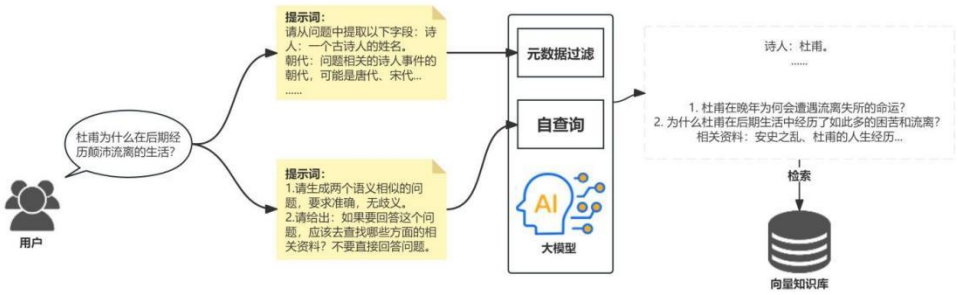


图 4 自查询增强和元数据过滤

4. 实验

4.1. 外挂知识库的有效性

为了验证外挂诗词知识库确实增强了 ChatPoet 的诗词对话能力，我们自拟了大模型诗词对话能力评价指标，从三个维度评价 RAG 前后，ChatPoet 的诗词回答能力。评价指标内容见表 2。

表 2 ChatPoet 评价指标

	准确性	**理解性**	**连贯性**
解释	模型回复的诗词知识是否正确，解析是否恰当，是否有正确的部分。	模型是否能正确理解用户的提问而不是答非所问	诗词的回答是否连贯，单一，没有冗余的字符和知识。

题目数量	20 道题目	20 道题目	20 道题目
评价方式	专家盲评，每个回答按合理度给与 0~5 分		
评价模板	从准确性上评价，该问答是否符合诗词事实？【问题】【回答】	仅判断该回答是否和问题对应？是否有与回答不一致的地方？【问题】【回答】	仅判断回答符合标准的程度，有无多余无意义的部分？【问题】【回答】
负例	问：杜甫的《登高》全文是什么？ 答：白日依山尽，黄河入海流。 欲穷千里目，更上一层楼。	问：杜甫的《登高》全文是什么？ 答：杜甫的《登高》是一首表达了他对国家和个人命运的忧虑与悲愤的诗歌。	问:《登高》全文是什么？ 答：无边落木萧萧下，不尽长江滚滚来。 万里悲秋常作客，百年多病独登台。 艰难难难难难难难难...

在经过专家盲评后，我们得到了 ChatPoet 有无 RAG 增强的三个维度得分，见表 3。得益于强大的外挂知识库检索能力，ChatPoet 问答准确性大大提高。外挂知识库的加入使得大模型回答问题从模型记忆转化为从提示词中信息抽取、整合，这能够解释准确性提高如此之大。在理解问题能力上，有无 RAG 相差不大。可能是模型理解问题能力取决于大模型本身的训练细节。但是，带有 RAG 增强的 ChatPoet 回答的连贯性却不如基座能力。多次实验我们观察到：很多次模型生成的回复是正确的，但由于我们设置的上下文长度过长，检索到的知识过多，模型会得到很多无意义的信息，在输出时倾向于回答不符合问题的字符。为了模型回答的准确性，检索器需要检索尽可能多的知识，但基座模型长上下文信息提取能力弱，导致模型回答连贯性降低。如何平衡这个矛盾，是我们下一步研究的重点。

表 3 有无 RAG 评价得分

对比	**准确性**	**理解性**	**连贯性**
ChatGLM3	31	83	90
ChatGLM3+RAG(ChatPoet)	77	86	59

4.2. 外挂知识库各组件的有效性

我们在构建数据库过程中使用的自查询增强、元数据过滤、摘要与信息头四个组件，目的是增强知识库的召回率。为了验证各组件对知识库召回率的影响，我们进行了消融实验，拟定 100 个诗词问题和答案，包含诗人相关信息、诗词释义、诗词赏析、诗词全文和诗词内容提问五个类型，实验评价指标见表 4。若检索出的知识包含问题的答案或与问题相关性很大，则证明检索是有效的。为了充分利用 ChatGLM-6B 模型的 32k 上下文理解能力，我们把向量数据库的检索数量超参数定为 16。检索出的知识排序也影响着得分：若相关知识越靠前，则得分越高；反之越低。若有多个相关的知识被检索出，则只计算最靠前的知识。

表 4 诗词知识库各组件有效性的评价指标

	诗人细节	**诗词释义**	**诗词赏析**	**诗词全文**	**诗词内容**
解释	对某个诗人生平、细节的提问。	对古诗词中某个词语的释义提问。	对古诗词的某一段进行意义提问。	对某诗词全文、某一段原文内容提问。	对诗词全文的知识、事实提问。
数量	20 道题目	20 道题目	20 道题目	20 道题目	20 道题目
评价方式	两个专家盲评，取平均 Top-1：5 分；Top-4：4 分；Top-8：3 分；Top-12：2 分；Top-16：1 分 未检索出：0 分。				

在表 5 中，A、B、C、D 分别表示长文档分块信息头、大模型摘要、自查询增强和元数据过滤。从消融实验结果可以看到，使用信息头和摘要后，检索器对相关知识的检索效率得到了提升。特别是诗人细节和诗词释义这两个涉及具体细微知识点的提问类型。自查询增强这一方法对五大提问类型的检索效率几乎没有帮助。其可能的原因是语义编码器已经很成熟，相同的问题的另一种问法经向量化后仍有相似的语义，检索得到的相似度也接近一致。在把自查询替换为元数据过滤后，检索得分小幅度上升。在自查询增强的基础上，元数据过滤对诗词知识检索效率有较大的提升。可见，在元数据过滤减少检索的范围对召回率的有效性。相比自查询增强，元数据过滤对检索召回率的贡献较大。总的来说，使用了四个组件后，检索得分大大提高，证明了我们使用方法的有效性。

表 5 消融实验结果

组件	诗人细节	诗词释义	诗词赏析	诗词全文	诗词内容
无组件	42	34.5	55	78	57.5
A+B	75	59	62	85.5	62
A+B+D	76	68.5	71	85	69
A+B+C+D	79.5	78	70.5	87.5	74

5. 结论

本研究提出了 ChatPoet 诗词对话模型，基于 RAG 技术，提高了基座模型 ChatGLM3 的诗词对话能力。在构建向量知识库过程使用了一系列方法，对非结构化诗词文档分块后，拼接所属的头信息和大模型生成的摘要和关键词，提高了 RAG 检索的效率；在用户提出问题后，ChatPoet 使用自查询增强查询语句，使用元数据过滤减少查询范围，进一步提高相应知识的召回率，更好的回答用户提出的诗词问题。 但是在使用 RAG 增强 ChatPoet 的诗词能力后，难以兼顾模型回答的连贯性和准确性，如何平衡这两个矛盾，是我们未来的工作。

参考文献

刘邦奇,聂小林,王士进,袁婷婷,朱洪军,赵子琪 & 朱广袤.(2024).生成式人工智能与未来教育形态重塑：技术框架、能力特征及应用趋势. **电化教育研究(01)**,13-20.

吴砥, 李环 & 陈旭. (2023). 人工智能通用大模型教育应用影响探析. **开放教育研究 (02)**, 19-25+45. doi:10.13966/j.cnki.kfjyyj.2023.02.003.

涂竟玉. (2023-08-02). AI 大模型助力“因材施教”. **深圳商报**, A08.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Guo, Q., Wang, M., & Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey. *ArXiv*, abs/2312.10997.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2023). A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ArXiv*, abs/2311.05232.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1-35.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L.E., Simens, M., Askell, A., Welinder, P., Christiano, P.F., Leike, J., & Lowe, R.J. (2022). Training language models to follow instructions with human feedback. *ArXiv*, abs/2203.02155.

Panda, S. (2023). Enhancing PDF interaction for a more engaging user experience in library: Introducing ChatPDF. *IP Indian Journal of Library Science and Information Technology*, 8(1), 20-25.
Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.
THUDM. (2023) *ChatGLM3*. Retrieved from https://github.com/THUDM/ChatGLM3
Thakur, N., Reimers, N., Ruckl'e, A., Srivastava, A., & Gurevych, I. (2021). BEIR: A Heterogenous Benchmark for Zero-shot Evaluation of Information Retrieval Models. *ArXiv*, abs/2104.08663.
Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. *arXiv preprint arXiv*:2303.18223.

面向编程题解答的大语言模型性能强化研究

Enhancing the Performance of Large Language Models for Programming Problem Solving

董银锋[1]，魏艳涛[1*]，徐琦[1]，肖光润[2]，崔咚咚[1]，潘东辉[3]
[1]华中师范大学人工智能教育学部数字教育湖北省重点实验室
[2]湖北文理学院机械工程学院
[3]安徽大学
yantaowei@mail.ccnu.edu.cn

【摘要】在人工智能技术不断发展的背景下，大语言模型在自然语言处理任务中已经取得了显著的成就，并催生了许多与编程相关的模型。然而，在编程题解答领域，大语言模型仍然面临着诸多挑战。本研究提出了一种方法，通过引入反馈模块和外部知识库来增强大语言模型在编程题解答方面的性能。反馈模块负责对解题思路及代码中的错误进行审查和修正；外部知识库则为模型提供相似问题的解题思路。实验结果显示，相较于基线模型，本方法在编程题解答方面实现了显著的性能提升。

【关键词】 GPT-3.5；反馈；外部知识库；编程题

Abstract: *In the context of continuous advancement in artificial intelligence technology, large language models have achieved remarkable accomplishments in natural language processing tasks, giving rise to numerous models associated with programming. However, in the domain of programming problem-solving, large language models still confront various challenges. This study proposes a method to enhance the performance of large language models in solving programming problems by incorporating a feedback module and an external knowledge base. The feedback module is responsible for reviewing and correcting errors in both problem-solving approaches and code, while the external knowledge base provides the model with strategies for solving similar problems. Experimental results indicate that, compared to the baseline model, this approach achieves a significant improvement in the ability to solve programming problems.*

Keywords: GPT-3.5,Feedback,External Knowledge Base,Programming Questions

1. 前言

随着人工智能技术的不断进步，大语言模型在编程教育中取得了显著成果，如编程题问答、代码生成和代码纠错等。这些技术的应用不仅极大地提高了编程学习的效率，也使得编程教育更加个性化和智能化。然而，在编程题解答领域，由于编程题的复杂性和多样性，大语言模型仍然面临一些挑战。如何提高大语言模型在编程题解答方面的性能，生成更加准确、高效的代码，仍然是一个亟待解决的问题。

近期，大语言模型在编程题解答领域展现出了卓越的性能。其凭借强大的语义理解能力，能够处理基础的算法和数据结构问题，应对复杂的系统设计和架构挑战，并生成高质量的文本内容。在编程领域，大语言模型能够分析代码的上下文，预测潜在的代码片段或函数，进而提供自动补全和提示功能，助力开发者提高编码速度和准确性，减少错误和冗余工作。这些能力使得大语言模型在软件开发和程序员日常工作中扮演着重要角色。尽管如此，大语言模型在编程题解答方面仍面临一些挑战。一方面，模型可能无法理解编程题的上下文信息，

从而导致生成错误的解题思路或代码；另一方面，在生成解题思路时，模型往往无法有效利用相似题目的解答思路作为参考。

针对上述挑战，本研究提出了一种方法，通过反馈模块和编程题外部知识库来增强大语言模型在编程题解答任务中的性能，并评估了该方法的效果。编程题外部知识库为模型提供了编程题解答的参考和提示；反馈模块能够生成针对模型输出的反馈信息，据此调整解题思路和代码，从而提升模型在编程题解答任务上的性能。实验证明，通过结合外部知识库和反馈模块，大语言模型在编程题解答任务中的性能得到了显著提升，能够生成高质量的解答，有效辅助用户学习编程。

本文的主要贡献如下：

1，我们设计了一个基于大语言模型的编程题解答系统，可以根据题目的描述，生成相应的解析。解析包括了解题思路、解题步骤和解题代码等，可以帮助用户更好地理解和掌握题目。相较于单独使用大语言模型，解答编程题的正确率得到了显著提升。

2，我们设计了一个反馈模块，用于对生成的解题思路和代码进行检查，检查解题思路和代码是否正确、符合题意等。通过反馈模块生成的反馈信息帮助修改模型初始输出中的错误，提高模型的正确率。

2. 相关研究工作

作为智能时代的一项重要技能，编程能力逐渐被国际产业界和教育界所重视，日益成为世界各国进行创新性人才培养战略的实施重点（孙丹&李艳，2019）。诸如 ACM 国际大学生程序设计竞赛（ICPC）等活动，旨在激励学生展示并提升他们在算法、编程和团队协作方面的技能。通过学习编程与算法，学生不仅能够提升其逻辑思维能力，还能为云计算、大数据、机器学习以及人工智能等新兴工程领域的研究打下坚实的基础（刘莞玲,2021）。

在计算机编程领域，大语言模型能够理解和生成代码，为开发者提供代码补全、代码生成和代码修复等功能，极大提升了开发效率。OpenAI 开发的 Codex（Chen, 2021），作为 GPT-3 模型（Floridi，2020）的一个变体，在理解和解析编程语言及代码逻辑方面表现卓越，能够根据自然语言指令生成代码。CodeLLaMA（Roziere，2023）适用于编程教育和编码辅助，它可以理解程序员提出的编程问题，并提供代码片段作为答案，还能够解释代码的工作原理。CodeGeeX（Zheng，2023）是一个能够辅助开发者生成代码片段、修复错误、理解现有代码库的工具，有效加速了复杂代码项目的理解和维护，进而提高了开发效率。这些模型为编程教育和软件开发提供了新的辅助手段。然而，面对难度较高或逻辑复杂的编程题目时，这些模型的输出结果往往未能达到预期。

大语言模型的研究热点之一是如何整合和应用外部知识。虽然模型通过预训练获得了广泛的知识和语言理解能力，但在处理实时数据、特定领域知识或需要逻辑推理的任务时，仍存在局限性。为了克服这些限制，研究者们开始探索将大语言模型与外部知识库相结合的方法（Zhang，2019）。Wei（2022）提出的思维链提示（Chain of Thought prompting）方法发现通过添加相似问题的推理过程可以提高模型性能。外部知识库可以提供更多的上下文和指导，帮助模型更好地理解问题，进行正确的推理。我们通过检索与问题相关的文档，利用外部知识库为大语言模型提供解题思路，从而提升其理解和生成能力。

尽管外部知识库的引入提供了额外的提示信息，但大语言模型仍可能生成错误答案、出现幻觉（Huang，2023）和错误推理。模型有时无法准确捕捉输入文本中的信息，或对有歧义的部分产生错误理解，导致错误输出。推理是一项迭代的任务，面对复杂推理任务时，大语言模型很难一次性生成正确答案。近期研究发现，迭代精化模型输出（Madaan，2023）、从反馈中学习（Ouyang，2022）等策略可以减轻这些问题。Shinn（2023）提出将不同类型的反馈

转换成自然语言，将自然语言反馈应用于迭代过程中以提升性能。Nathani（2023）证明一个通用的反馈模块不足以解决语言模型反应中各种各样的潜在错误，并提出了一套针对特定错误类别提供详细反馈的反馈模块，以提高推理和解决方案的质量。

本研究通过反馈模块和外部知识库提升大语言模型在编程题解答领域的性能。通过构建编程题外部知识库获取相似问题的解题流程作为参考，向大语言模型提供解题思路；为编程题解答设计一种反馈模块，对大语言模型的输出生成反馈信息，根据反馈信息迭代修改模型的输出，提升了大语言模型解答编程题的正确率。

3. 系统构建

3.1. 系统概述

本研究基于大语言模型，结合编程题外部知识库和反馈模块，构建了编程题问答系统。系统可以根据用户输入的编程题，生成解答的思路和代码。

系统主要包含以下模块：

语言模型。语言模型的作用是根据用户的输入生成对应的输出。在系统中需要完成两种任务：生成编程题的解题思路和代码；依靠其强大的上下文学习能力，根据提示和反馈信息修改输出，使生成的文本更加符合要求。

反馈模块。反馈模块的作用为对模型的输出进行验证，判断模型的输出是否合理。反馈模块会检查模型输出是否合理，例如：检查生成的答案解析是否正确、检查代码是否有错误等。若反馈模块提示需要修改输出，则将反馈模块输出的反馈信息输入到语言模型中重新生成输出，反复迭代更改模型的输出，直至反馈模块提示不需要更改输出。该模块可使用生成式语言模型来实现。

编程题外部知识库。编程题向量数据库包含大量的算法题、编程题、题目的答案和解析，每道题目都有其对应的分类标签，包括题目涉及的知识点、题目的难度等。该模块根据用户的要求，从代码知识库中检索相关的知识，作为提示输入到大语言模型中，为模型提供相似问题的解题思路，帮助模型更好地理解问题，进行正确的推理。

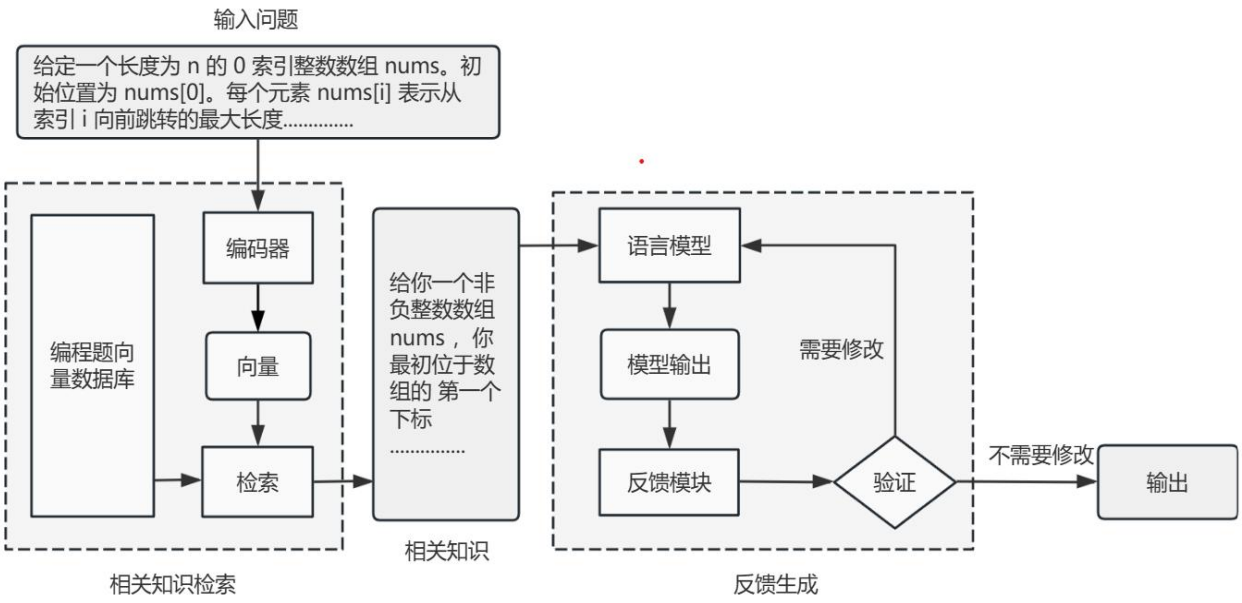


图 1 系统框架

本文提出的系统框架如图 1 所示，具体流程为：

1，对于用户输入的编程题 Q，使用文本编码器将用户的输入编码为 d 维向量，使用检索工具进行相似度检索，从编程题向量数据库获取与 Q 最相似的 top_n 条数据作为检索结果，检索结果作为语言模型解答编程题 Q 的参考示例。

2，将检索结果和编程题 Q 输入语言模型中，获得模型输出 O。

3，将输出 O 输入到反馈模块中，按照预设的反馈策略生成反馈，反馈为文本格式。

4，若反馈中有对输出 O 的修改建议，则将反馈和 O 输入到语言模型中生成新的输出；若反馈中没对 O 的修改建议，则将 O 作为最终输出结果。

3.2. 编程题外部知识库

本研究构建的编程题外部知识库主要包含以下信息：

1，编程题。编程题的来源为各大编程竞赛网站，这些网站提供了大量的算法题目和解题讨论。收集到的数据存储方式如图 2 所示，每条数据包含算法题题目、详细信息、题目的难度、题目的解析和题目的答案等。题目的答案在时间复杂度和空间复杂度上都较低，解析包含用到的算法以及知识点等。在检索题目时，可以根据题目涉及到的知识点对题目进行检索，例如，用户在要求系统解答用到数组知识点的编程问题时，系统随机从知识点包含数组的题目中抽取 n 道最相似题目，将这 n 道目的详细内容和解析作为模型解答问题的提示。在存储编程题时，将题目按知识点分类，按$\{c_1:[q_{11},\ldots],c_2:[q_{21},\ldots,]\ldots c_n:[q_{n1},\ldots,]\}$的格式存储，其中$c_i$表示题目类别，$q_{ij}$代表在类别$c_i$中的编程题。一个题目可以属于多个类别。

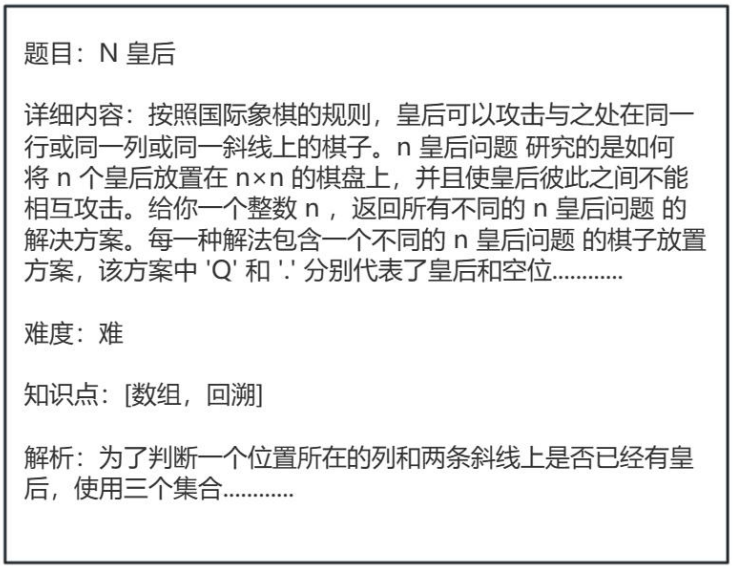


图 2 算法数据集

2，算法、数据结构相关知识。经典的数据结构和算法书籍，或者各大网站的资料。在处理文档过程中，将文档切块，将主要内容进行编码，将文档存入向量数据库，用于在算法解析时提供资料参考。

在构建外部知识库的过程中，按照类别$\{c_1,c_2\ldots c_n\}$分别构造向量数据库$\{d_1,d_2\ldots d_n\}$。首先将文本向量化，将每道编程题作为一个块，使用 BERT 模型作为编码器。为了防止进行相似度检索时被题目的答案和解题思路等其他信息干扰，只将每道题的详细内容编码为 d 维向量。下一步使用 FAISS 工具的 index_factory 函数创建索引，索引在查询过程中用于相似向量检索。使用 FAISS 的 add 函数将向量 p 添加到索引中。检索题目的流程如图 3 所示，以下是检索的具体流程。

1，使用语言模型获得输入的类别c_i，并根据类别调用其对应的向量数据库d_i；

2，使用编码器将用户输入编码为向量 q，使用 q 进行检向量检索，获取最相似的 n 个结果$\{k_1,k_2\ldots k_n\}$，将k_i的详细内容和解析作为模型的提示。

3.3. 反馈模块

反馈模块的核心功能是为模型的输出生成自然语言形式的反馈，提出可行的修改建议。在使用中，使用生成式语言模型实现反馈模块的功能。反馈模块的工作流程如表 1 所示，首先，将模型的初始输出 O 输入到语言模型中，生成关于解题思路的反馈f_1。随后，检查f_1是否表示需要对 O 进行修改。如果需要，则将 O 和f_1一同输入到模型中，指导模型根据f_1对 O 进行修改，得到更新后的输出 O，再次生成关于解题思路的反馈f_1；如果无，则终止循环进入下一步骤。迭代修改 O，直至达到预设的迭代次数 T，或者f_1表明输出 O 已无需进一步修改。接着，将更新后的 O 再次输入到反馈模块中，生成关于代码的反馈f_2。同样地，检查f_2中是否存在修改需求。如果有，则利用f_2指导模型对 O 进行进一步修改；如果无，则终止循环，并将 O 作为最终输出。当达到预设的迭代次数 T，或f_2明确表示无需进一步修改时，将 O 作为最终输出。我们为解题思路和代码设计了不同的检查任务。对于解题思路，检查是否符合题目要求，是否存在步骤遗漏和错误；对于代码部分，检查是否存在语法错误，代码的实现是否与解题思路一致。

表 1 反馈模块伪代码

反馈模块伪代码
模型输出 O，语言模型 LM，反馈模块 F，迭代次数上限 T，提示模板P_1, P_2, P_3, P_4
1:for i ← 1 to T:
2:　　$f_1 = F(P_1(O))$　　#对解题思路生成反馈
3:　　if “无需修改” in f_1:　#不需要修改
4:　　break　　#退出
5:　　$O = LM(P_2(O))$
6: for i ← 1 to T:
7:　　$f_2 = F(P_3(O))$　　#对代码生成反馈
8:　　if “无需修改” in f_2:　#不需要修改
9:　　break　　#退出
10:　　$O = LM(P_4(O))$
11:return O　　#返回结果

3.4. 编程题解答流程

在生成编程题解析时，从代码知识库中检索相关问题，检索 n 个最相关问题作为解题参考。检索过程中使用语言模型 LM 获取输入问题 input 的类别c_i，从其c_i对应的向量数据库d_i中获取 n 个最相关结果$Q = \{q_1, q_2 \ldots q_n\}$，将$q_i$中的详细内容和解析作为解题参考$Q_1$。将 input 与$Q_1$输入到模型中获得初始输出 O。将 O 输入到反馈模块中生成对解题思路的反馈fd_1。若fd_1表示需要修改，将反馈fd_i和 O 输入到 LM 中重新生成输出 O，重复反馈和修改过程，直到 O 通过反馈模块的验证或迭代次数达到上限。将 O 输入到反馈模块中生成对代码的反馈fd_2，类似地，重复反馈和修改代码的过程后，将 O 作为最终答案。

4. 实验

4.1. 实验设置

本研究以 GPT-3.5 作为生成答案和反馈的基础模型 M。利用 BERT 模型将文本转换为相似度检索所需的向量。从专业编程题网站中选取题目作为编程题数据集，一部分用于构建外部知识库，一部分用于测试模型性能。采用 FAISS 工具和编程题数据集构建向量数据库作为编程题外部知识库。将编程题难度分为简单和困难两种级别，分别评估模型在这两种难度下的性能。

4.2. 编程题解析性能评估

我们收集了 100 道编程题目以评估编程题解答功能的性能，这些题目包括 50 道简单难度的算法题和 50 道困难难度的算法题。我们将这些题目输入到 GPT-3.5、本研究设计的编程题解答系统以及 GPT-4 中，获取解题思路和 Python 代码。若解题思路和代码均满足题目要求，则判定为正确回答。评估指标为代码的正确率，计算公式如下：

$$\text{正确率} = \frac{\text{ac}}{\text{n}}$$

其中，ac 表示回答正确的数量，n 表示数据集中编程题的总数。表 2 详细展示了模型的回答正确率。分析表中的数据可知，集成了反馈模块和外部知识库的编程题解答系统，在简单难度和困难难度的编程题上，其正确率均超越了 GPT-3.5 模型。特别是在困难难度的编程题上，模型的正确率实现了显著的提升。

在实验过程中，我们观察到仅使用 GPT-3.5 模型解析编程题时，模型能够较好地回答常见和简单难度的问题，例如二叉树的遍历、链表的反转和背包问题等。然而，当遇到不常见或困难难度的问题时，模型的正确率较低。在 GPT-3.5 模型生成的错误答案中，主要存在两种错误类型：第一种是在生成解题思路的过程中即出现错误，从而导致最终答案错误；第二种是解题思路正确，但在代码编写过程中遗漏了一些关键条件，生成了错误代码。外部知识库和反馈模块的使用可以有效避免模型生成错误答案。利用编程题外部知识库，模型能够在生成解题思路时得到辅助，有效减少错误的出现。当生成了错误答案时，对于第一种错误类型，反馈模块会对题目进行重新分析，与模型输出 O 中的解题思路进行对比，以识别并纠正错误。对于第二种错误类型，反馈模块会检查模型输出 O 中的代码逻辑，与解题思路进行对比，以检测并修正其中不符合预期的地方。

表 2 编程题解答正确率

模型	简单难度	困难难度
GPT-3.5	92%	64%
GPT-3.5+外部知识库	96%	70%
GPT-3.5+反馈模块	94%	78%
GPT-4	96%	82%
GPT-3.5+外部知识库+反馈模块	96%	82%

在消融实验中，我们分别评估了反馈模块与外部知识库对模型性能的影响。如表 2 所示，实验数据表明，外部知识库与反馈模块均能显著提升模型在解答算法题时的正确率。当两者联合应用时，模型的正确率达到了最高，这一结果充分证明了这两种模块的有效性。外部知识库通过提供相似题目的解题思路，有效地引导模型生成正确答案。同时，反馈模块通过对模型输出的反复验证，能够及时发现并纠正输出中的错误。

4.3. 模型效果示例

在表 3 中，我们展示了系统在解答编程题时的表现。示例中，模型的初始输出虽然正确地提出了解题思路，但在代码编写阶段出现了错误。具体来说，模型在编写更新利润的方法时出现了误解，错误地将解题思路中的“将所有工人的利润相加”理解为“更新最大利润”。这种错误源于大语言模型没有正确理解生成的解题思路。模型的初始输出在反馈模块中通过了对解题思路的检查，未能通过对代码的检查。反馈模块正确识别出了代码中的错误，同时检测出代码中未能正确处理边界条件，并给出了修改建议。根据系统的运行流程，当反馈中包含修改建议时，系统会将反馈与初始输出一同输入到大语言模型中，以重新生成答案，并

再次提交给反馈模块进行验证。重新生成的答案中修改了代码中的错误。在第二次调用反馈模块后，若生成的反馈未发现任何问题，则将新答案视为最终输出。

表 3 编程题解析实例

问题	你有 n 个工作和 m 个工人。给定三个数组：difficulty, profit 和 worker，其中:difficulty[i]表示第 i 个工作的难度，profit[i]表示第 i 个工作的收益。worker[i]是第 i 个工人的能力，即该工人只能完成难度小于等于 worker[i]的工作。每个工人最多能安排一个工作，但是一个工作可以完成多次。返回在把工人分配到工作岗位后，所能获得的最大利润 。

初始输出	为了解决这个问题，我们可以按工作的难度和利润进行排序，然后对每个工人找到他能完成的难度最大的工作，并计算该工作对应的利润。最后，将所有工人的利润相加即为最大利润。具体步骤如下：将 difficulty 和 profit 按照 difficulty 的大小进行排序，同时保持两个数组的关联关系。将 worker 按照能力大小进行排序。遍历每个工人，对于每个工人，使用二分搜索在 difficulty 数组中找到最大的难度小于等于该工人能力的工作，并计算该工作对应的利润。将所有工人的利润相加即为所求最大利润。
	def maxProfitAssignment(difficulty, profit, worker):
	jobs = list(zip(difficulty, profit))
	jobs.sort(key=lambda x: x[0])
	worker.sort()
	max_profit = 0
	j = 0
	for ability in worker:
	while j <len(jobs) n and jobs[j][0] <= ability:
	max_profit = max(max_profit, jobs[j][1])
	j += 1
	return max_profit

反馈	分析思路： 根据题意，算法需要按照如下步骤执行： 为了解决这个问题，我们可以按工作的难度和利润进行排序.....提供的解题思路正确地反映了这些步骤。无需修改
反馈	分析代码思路： 根据题目要求，我们需要在遍历每个工人时将利润加到利润总和中。因此，代码中的错误在于它没有正确处理最大利润。正确的做法是在每次找到工作后更新最大利润。同时代码没有正确处理边界条件。

修改后输出	def maxProfit(difficulty, profit, worker):
	if not difficulty or not profit or not worker:return 0
	jobs = sorted(zip(difficulty, profit)) # 按难度排序的工作
	worker.sort() # 工人按能力排序
	res, i, max_profit = 0, 0, 0
	for ability in worker:
	while i < len(jobs) and ability >= jobs[i][0]:
	max_profit = max(max_profit, jobs[i][1]) # 更新最大利润

```
                i += 1
            res += max_profit   # 累加最大利润
        return res
```

反馈	分析思路： 根据题意，算法需要按照如下步骤执行： 为了解决这个问题，我们可以按工作的难度和利润进行排序.....提供的解题思路正确地反映了这些步骤。无需修改
反馈	分析代码思路：根据代码中的思路.....提供的代码思路是正确的。无需修改。

4.4. 模型调用次数

在反馈模块使用过程中，每调用一次反馈模块至少需要调用两次语言模型，分别用来对解题思路和代码生成反馈。若生成的反馈中有需要更改的部分，则再调用一次语言模型对答案进行修改，并重复生成反馈步骤，直到达到迭代次数上限或者反馈中表示不需要更改。在实验中，我们统计了大语言模型调用次数。在对 100 个问题进行回答过程中，有 41 次反馈表示需要进行修改，每次修改多调用 2 次语言模型。本文中使用的反馈模块在遇到复杂问题时容易生成错误答案并重复调用反馈模块，消耗的资源较多。

5. 总结

本研究探讨了融合反馈模块与外部知识库的大语言模型在编程题解答领域的应用，旨在提升大语言模型在编程题解答任务中的性能。通过引入外部知识库，模型在回答编程题时能够参考知识库中相似题目的解题思路，从而提高了解答的准确性。当模型生成错误结果时，反馈模块能够针对模型的输出生成自然语言形式的反馈，并提出切实可行的修改建议，通过迭代修改，进一步增强模型在编程题解答任务中的表现。

然而，本研究在反馈模块的设计上存在一些不足。当前的反馈模块在模型输出频繁出错时，需要多次调用模型，导致较高的资源消耗。此外，在生成反馈时，模型可能无法识别某些错误的解题思路，如解题步骤的遗漏或解题思路与题目要求的不符。有时，反馈模块甚至可能生成错误的反馈信息，从而导致后续代码生成错误。在处理某些常见问题时，即使大语言模型能够生成正确答案，反馈模块的介入有时也可能导致最终结果的错误。因此，反馈模块的设计仍有待进一步优化和改进。

基金项目

国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），国家教师发展协同创新实验基地建设专项“基于视频数据的“人工智能+”研训学生情感演化规律分析方法”（项目号：CCNUTEIII-2021-19），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号CCNU22JC011）.

参考文献

孙丹 & 李艳.(2019).国内外青少年编程教育的发展现状、研究热点及启示——兼论智能时代我国编程教育的实施策略. *远程教育杂志*(03),47-60. doi:10.15881/j.cnki.cn33-1304/g4.2019.03.005.

刘莞玲,吴英杰,傅仰耿,王一蕾 & 孙岚.(2021).新工科背景下算法与数据结构实践教学改革探索. *计算机教育*(07),73-76. doi:10.16512/j.cnki.jsjjy.2021.07.018.

Chen M, Tworek J, Jun H, et al. Evaluating large language models trained on code[J]. *arXiv preprint arXiv:2107.03374, 2021*.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*, 681-694.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2023). A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ArXiv, abs/2311.05232*.

Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegreffe, S., ... & Clark, P. (2023). Self-refine: Iterative refinement with self-feedback. *arXiv preprint arXiv:2303.17651*.

Nathani, D., Wang, D., Pan, L., & Wang, W. Y. (2023). Maf: Multi-aspect feedback for improving reasoning in large language models. *arXiv preprint arXiv:2310.12426*.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback, 2022. *URL https://arxiv. org/abs/2203.02155, 13*.

Roziere, B., Gehring, J., Gloeckle, F., Sootla, S., Gat, I., Tan, X. E., ... & Synnaeve, G. (2023). Code llama: Open foundation models for code. *arXiv preprint arXiv:2308.12950*.

Shinn, N., Cassano, F., Gopinath, A., Narasimhan, K. R., & Yao, S. (2023, November). Reflexion: Language agents with verbal reinforcement learning. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems, 35*, 24824-24837.

Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). ERNIE: Enhanced language representation with informative entities. *arXiv preprint arXiv:1905.07129*.

Zheng, Q., Xia, X., Zou, X., Dong, Y., Wang, S., Xue, Y., ... & Tang, J. (2023, August). Codegeex: A pre-trained model for code generation with multilingual benchmarking on humaneval-x. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 5673-5684).

基于 ARCS 模型的农村初中生 AI 课程学习活动设计研究

Research on AI Course Learning Activity Design for Rural Junior High School Students Based on ARCS Model

张乐佳
广东省清远市教师发展中心
1562356563@qq.com

【摘要】 我国中小学人工智能课程教学正从“试点”走向“普及”，对某市调研发现农村初中开课率低，学生学习积极性不高。ARCS 模型由多种学习动机理论综合而成，提供了系统性的动机解决方案。为此，通过分析 ARCS 模型在农村人工智能学习活动设计的潜在作用，梳理适合于农村初中生的学习活动支持策略，构建农村初中生人工智能学习活动设计框架并进行例证，为农村初中开展人工智能学习活动设计提供理论和方法指导。
【关键词】 ARCS 模型；农村初中生；AI 课程；学习活动设计

Abstract: *China is promoting AI curriculum education in primary and secondary schools. A survey conducted in a city of China found that both the opening rate and student-participation of AI courses in rural junior high schools were low. After analyzing the potential advantages of ARCS Model in the design of AI learning activities in rural junior high schools, the research built motivational strategies based on the ARCS model, constructed an design framework for AI learning activities, hoping to provide theoretical and methodological guidance for the design of AI learning activities in rural junior high schools.*
Keywords: the ARCS model, students of rural junior high school, design of AI learning activities

1. 引言

AI（Artificial Intelligence，人工智能）教育已成为各国教育竞争的新焦点。我国在 2022 年 4 月发布了《义务教育信息科技课程标准（2022 年版）》（以下简称新课标），明确要在初中阶段开设人工智能课程，标志着我国初中人工智能课程教育从试点走向普及。而全面落实人工智能课程，难点在农村。本研究对广东省某地市 93 名农村初中信息科技教师及 1400 多名学生进行了问卷调查，结果显示：高达 91.4%教师表示学校未开设人工智能课程，51.61%教师认为现有的人工智能课程活动设计难以操作，仅有 19.54%学生对人工智能课程最感兴趣，45.03%学生认为教师较少设计有效的学习活动以支持学习。可见，如何激发农村初中生人工智能课程学习兴趣、增强课程学习体验是目前推广人工智能课程急需解决的问题。

2. 研究述评

2.1. AI 学习活动设计。学习活动是指学习者以及与之相关的学习群体为了完成特定学习目标而进行的操作总和（Helen，2007）。众多学者将活动理论视为学习活动的基础支撑理论，基于恩格斯托姆的三角模型对学习活动设计模式进行了深入的研究。但目前学习活动设计研究主要集中在高校在线教学领域，在中小学 AI 课程教学方面少有涉及。段波基于对计算思维教育现状提出了“需求分析”→“活动方式”→“情景设计”→“过程设计”→“评价设计”人工智能学习活动设计思路（段波，2022），为本研究提供了有益的参考。

2.2. ARCS 模型。ARCS 模型是一种学习动机策略模型，由注意(attention)、相关(relevance)、信心（confidence)和满意(satisfaction)四个主类别组成（如表 1 所示）。模型提出者 J.Keller

教授认为动机设计流程与教学设计流程十分相似，包括 10 个过程，如图 1 所示。

表 1 ARCS 模型分类、定义和问题处理描述

主类别	定义	问题处理描述
注意	吸引学习者兴趣；激发学习好奇心	我怎样能使学习经历既刺激又有趣？
相关	满足学习者需求，形成积极的态度	这些学习经验对我的学生有什么价值？
信心	让学习者相信能成功并控制成功	我如何通过指导帮助学生成功并让他们控制自己的成功？
满意	以奖励强化成就（内部和外部）	我能做些什么帮助学生对学习满意？

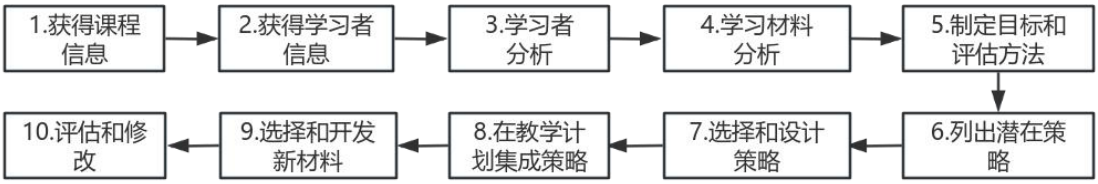


图 1 ARCS 模型设计流程图

3. ARCS 模型对农村初中生人工智能学习活动设计的潜在作用分析

农村初中生人工智能学习活动设计面临着学校教育资源紧张，教师流失、教师队伍结构性缺编、学生课程学习动机不强、学业基础薄弱等问题（范先佐，2013），需要系统的问题解决方法。而 ARCS 模型是一个综合性的学习动机解决方案，不仅能提高学生的课堂学习参与度，也被证实能有效提高课程或资源的整体应用效果， 如：杨开诚（2001）基于 ARCS 模型构建了在线学习系统，郑燕林（2015）利于 ARCS 模型构建了 MOOC 的推进应用路径。可见，基于 ARCS 模型进行农村初中生人工智能学习活动设计，利于综合考虑农村初中的不利因素，形成适用的学习活动设计方案，从而整体提高学生学习参与度。

4. 基于 ARCS 模型的农村初中生人工智能学习活动设计框架的构建

4.1. ARCS 模型下农村初中生人工智能学习活动的动机支持策略

本研究综合考虑农村初中教学实际，对 ARCS 模型基本策略进行了设计，形成了农村初中生人工智能学习活动学习动机支持策略，如表 2 所示。

表 2 ARCS 模型下农村初中生人工智能学习活动的学习动机支持策略

主类别	子类别	农村初中生人工智能学习活动动机支持策略
A.吸引注意	A1.知觉唤醒	主题与农村学生背景相关；新颖的技术、方法或科技事件；实物、类比讲解难点知识；可视化知识逻辑。
	A2.探究唤醒	真实问题情景，问题驱动，认知冲突，探究生成性问题。
	A3.变化维持	问题链驱动，灵活多样的活动方式，制造悬念。
R.相关性生成	R1.目标导向	制定详细可行的活动目标和要求，并向学生清晰讲解。
	R2.动机匹配	小组合作学习，明确分工；鼓励学生充当“小老师”；鼓励互助。
	R3.熟悉程度	设计知识应用迁移任务，解决实际问题。
C.建立自信心	C1.学习要求	制定符合农村初中生学情的学习目标和要求，并清晰陈述。
	C2.成功机会	常规活动任务符合学生总体水平，设计一定的拓展、挑战性任务。
	C3.个人控制	提供学习资源和方法指引；开展自评，对学习过程和方法反思。
S.获得满足感	S1.自然后果	学生进行总结，通过测试、作品评比进行活动评价，组织互评。
	S2.积极后果	设计成果展示环节，通过电子勋章或实物等对学生进行奖励。
	S3.公平公正	教学评一体化，按照学习基础进行分层评价。

4.2. 基于 ARCS 模型的农村初中生人工智能学习活动设计框架

4.2.1. 设计原则。（1）核心素养育人原则。活动设计应符合新课标的理念，反映数字时代正确的育人方向，注重培养学生学科核心素养，让学生经历原理理解、数字化工具应用和问题

解决三个过程。（2）活动条件低依赖原则。要充分考虑农村学校的实际情况，采用操作简便、免费的人工智能虚拟平台，选用性价比较高的实验器材，采用小组合作的形式降低设备数量依赖。（3）认知负荷适配原则。要适当控制内在认知负荷（即适当减少学习内容或降低学习要求），提供有效学习指引降低外在认知负荷，设计更多知识应用场景，提高关联认知负荷。

4.2.2. 设计依据。（1）学习活动视角。杨开城（2002）认为学习活动应至少包括活动任务或主题、活动基本流程和步骤、活动监管规则、活动评价规则，李红美等人认为学习活动包含学习主体、活动内容、活动任务、活动成果、活动规则、活动工具、活动行为七大要素（李红美，许玮和张剑平，2013）。综合以上观点，本研究认为学习活动应至少包含活动主题、活动任务、活动工具、活动规则、活动流程与活动评价六个基本要素。（2）教学设计视角。人工智能课程学习活动有特定的学习目标和内容，为初中人工智能课程学习服务，因而有必要将活动目标纳入人工智能学习活动的基本要素。（3）ARCS 模型视角。J.Keller 教授提出了 ARCS 模型设计的基本流程（如图 1 所示），强调设计要点是将“列出潜在策略”、“选择设计策略”等流程集成到教学设计中。综合以上分析，本研究梳理了基于 ARCS 模型的农村初中生人工智能学习活动设计基本原理，如图 2 所示。



图 2 基于 ARCS 模型的农村初中生人工智能学习活动设计基本原理

4.2.3. 基于 ARCS 模型的农村初中生人工智能学习活动设计框架的构建。本研究以学习活动七个要素为设计序列，构建了“基于 ARCS 模型的农村初中生人工智能学习活动设计框架”，如图 3 所示。

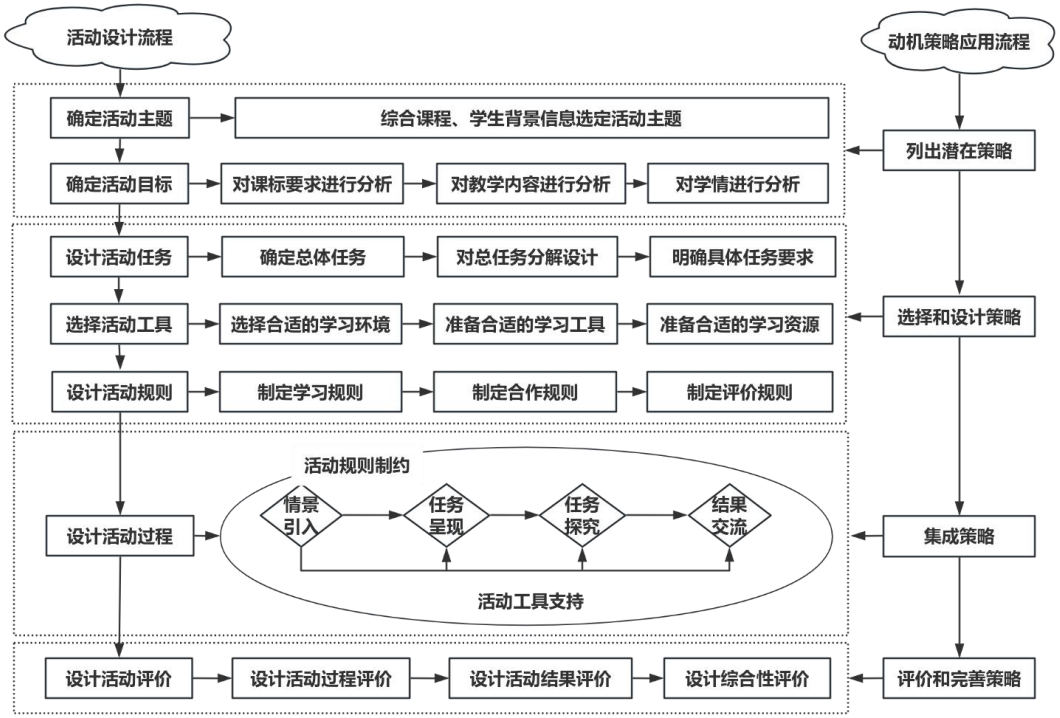


图 3 基于 ARCS 模型的农村初中生人工智能学习活动设计框架

4.3. 基于 ARCS 模型的农村初中生人工智能学习活动设计框架的应用

为检验动机支持策略与活动设计框架的可用性，本研究对 2023 版浙江省人工智能课程“机器学习”教学内容进行了学习活动设计。

活动主题、活动目标设计。（1）综合课程和学生背景信息，确定活动主题为“改造水果秤——人工智能中的‘机器学习’”。该活动主题贴近学生生活实际，聚焦传统水果秤称重效率低这一真实问题，有较强的育人价值。（2）基于对新课标要求、教学内容和学情的综合分析，制定了如下活动目标：通过合作体验理解机器学习的基本原理；理解有监督学习与无监督学习的基本原理，能说出两者间的区别；知道数据安全的重要性，通过合作探究常用的防护办法。（3）基于对教学内容及学生学习动机特点分析，列出潜在动机策略：A1. 知觉唤醒；A2. 探究唤醒；R1. 目标导向；R2. 机匹配；R3. 熟悉程度；C2. 成功机会；S1. 自然后果；S2. 积极后果。

活动任务、工具、规则设计。（1）综合活动主题和活动目标，确定了“应用机器学习对水果秤进行改造”这一主体任务，根据学习内容和学生认知特点将主任务划分为五个子任务。（2）活动工具采用低门槛的“慧”编程平台。教师课前制作微课资源、编制导学案，确定分组学习和评价规则。（3）从列出的潜在动机策略选择可行的策略，并进行详细设计。A1. 知觉唤醒：通过人类学习与机器学习类比加深原理理解；A2. 探究唤醒：以“改造水果秤”这一真实问题驱动学习；R1. 目标导向：制定详细的目标要求并向学生呈现；R2. 动机匹配：开展合作学习，明确组员责任；R3. 熟悉程度：设置学习任务将知识用于解决真实问题；C2. 成功机会：任务难度匹配学情，兼设一个挑战任务；S2. 积极后果：设计展示环节，奖励小物品。

活动过程设计。分为 2 个课时，设置“情景引入、任务呈现、任务探究、结果交流”四个环节。（1）教师通过播放乡镇超市水果称重视频引入，抛出核心问题引导学生学习。（2）教师呈现学习活动任务、学习要求和评价标准，并向学生清晰讲解。（3）学生进行任务探究，逐步完成活动任务。（4）学生对活动结果进行汇报交流。这一阶段将设计好的教学策略集成到具体的活动过程当中，形成整体的活动过程设计。

活动评价设计。采用“过程性+终结性”的评价方式。（1）过程性评价关注学生的学习和合作表现，通过教师记录和在线测试进行积分。（2）设计小组作品展示的终结性评价环节，组织学生对“改造方案”进行评分。（3）综合过程性与终结性得分，对优秀小组成员张贴“国旗贴纸”。（4）对学习效果进行分析，修改、完善所使用的动机策略。

6. 研究结论

本研究聚焦农村初中生人工智能学习活动设计，基于 ARCS 理论梳理了适用于农村初中生人工智能教学的学习动机支持策略，提出了将 ARCS 模型应用到学习活动设计的方法，构建了适用性较强的活动学习设计框架，为农村初中人工智能学习活动设计提供理论指导。

参考文献

[1]段波. (2022). 面向计算思维发展的中学人工智能课程活动设计. 天津师范大学学报(基础教育版) (01),53-58.doi:10.16826/j.cnki.1009-7228.2022.01.009.

[2]范先佐. (2013). 义务教育均衡发展与农村教育难点问题的破解. 华中师范大学学报(人文社会科学版) (02),148-157.

[3]李红美, 许玮&张剑平. (2013). 虚实融合环境下的学习活动及其设计. 中国电化教育 (01),23-29.

[4]杨开城. (2002). 学生模型与学习活动的设计. 中国电化教育 (12), 16-20

[5]Helen Beetham & Rhona Sharpe. (Ed.) (2007).Rethinking Pedagogy for a Digital Age: Designing and Delivering E-learning, Routledge - Taylor & Francis Ltd.

[6]Keller.J M.(2010).Motivational Design for Learning and Performance-The ARCS Model Approach.New York: Springer.

基于多模态数据融合的课堂学习情感识别研究现状及展望

Research Status and Prospescts of Classroom Learning Emotion Recognition Based on Multimodal Data Fusion

杨开[1]，魏艳涛[2*]，肖光润[3]，潘东辉[4]，徐琦[5]，胡涛[6]
华中师范大学人工智能教育学部数字教育湖北省重点实验室
yantaowei@mail.ccnu.edu.cn

【摘要】 在课堂学习中，情感扮演着关键的角色，但传统情感识别数据单一、识别结果容易产生误差，因此开展多模态数据融合的课堂学习情感识别研究具有重要意义。本文基于课堂学习情感识别进行综述，探讨了多种融合策略和识别方法，研究结果说明了多模态数据融合相比于单模态在课堂学习情感识别方面的优越性和必要性，并从多模态数据对齐、数据噪音干扰、课堂计算资源有限等提出研究挑战与展望，以助力构建更智能的教育环境。

【关键词】 情感识别；课堂学习；多模态；数据融合

Abstract: *In classroom learning, emotions play a crucial role, but traditional emotion recognition suffers from single-modal data and prone to errors in recognition results. Therefore, conducting research on multimodal data fusion for classroom learning emotion recognition is of great significance. This paper reviews classroom learning emotion recognition, discusses various fusion strategies and recognition methods, and the research results illustrate the superiority and necessity of multimodal data fusion over unimodal approaches in classroom learning emotion recognition. The challenges and prospects of research are proposed, including multimodal data alignment, data noise interference, and limited classroom computing resources, aiming to facilitate the construction of a more intelligent educational environment.*

Keywords: Emotion recognition, Classroom learning, Multimodal, Data fusion

1. 前言

情感识别对于理解学生在课堂中的情感状态和提升教学效果至关重要，作为一种智能识别和理解人类情感的技术，正在逐渐改变传统教育的面貌。例如，通过实时分析学生的面部表情和声音，教师可以即时了解学生是否理解课程内容，是否感兴趣或挫败。这种即时反馈使得教师能够适时调整教学节奏和策略。通过提供额外的解释、调整任务难度或引入更多互动活动来增强学生的参与感和学习动力。在课堂学习下，传统的情感识别通常采用单模态数据，如采用面部表情来推断学生的情绪状态（Zhu et al.,2023）、基于言语数据进行学生情绪与注意力实时评估(Song et al.,2022)、利用学生的评论进行情感识别（Shaik et al.,2023）等，单模态数据分析会导致对学生情感的片面理解，限制情感识别的准确性和全面性。

随着智慧课堂的兴起，教育环境和场景走向多元化、数字化和个性化。在这种新的教育场景中，可以捕获、融合和分析来自各种不同来源的大量多模态学生数据（Chango et al.,2022）。课堂多模态情绪感知主要涉及行为数据、文本数据、语音数据、面部表情数据、心理数据、学习情境数据、生理数据等（蒋艳双等，2022）。不同模态的数据融合可以提供更多的情感

线索，使得系统能够更全面、准确地理解学生的情感状态，从而提供更为全面的反馈和支持。有研究显示，多模态情感计算比单模态平均准确度提升了 9.83%（D'Mello & Westlund，2015）。因此，在课堂学习下面对大量多模态的学生数据，不同模态数据如何融合成为了一个关键的研究方向。

2.课堂情感识别研究现状

2.1.特征融合（Feature Fusion）

特征融合指将不同模态提取后的特征结合在一起，形成一个综合的特征表示。常用的特征层融合方法如 H.M(2023)所使用的特征拼接和遗传算法、Arijit Nandi(2022)所使用的加权求和、特征交叉和主成分分析等方法。

对于课堂学习下多模态数据融合的情感识别，H.M 提出了三种策略来对三个情感数据集执行语义数据融合，以关联来自以学习为中心的意见的数据。第一个是基于图像的文本意见表示，使用卷积神经网络(CNN)进行分类。第二个应用使用语义词典对图像和文本进行 N 维嵌入表示，并使用卷积神经网络与长短期记忆(CNN-LSTM)相结合进行分类。第三个是将第一种方法和第二种方法中的特征提取层连接起来。实验结果显示，在课堂学习下，多模态模型与单模态模型相比，准确率有了显著提高，尤其是 ELCE-Net 模型，与通过遗传算法优化的模型相比，提高了 4%，准确率达到 82%。由此可见，融合来自不同数据集的数据，在情绪识别任务上取得了很大的进步。然而，这是一种非常规融合，因为数据集最初不是从同一来源获得的，而是使用标签的语义近似来融合数据，会导致数据集之间存在差异，例如数据采集方式、数据标注质量、数据格式和对齐方面等差异。同时标签的语义近似可能会导致信息丢失或失真，因为不同数据集的标签定义可能不一致，导致情感分类任务中的混淆或误导，可能会降低融合数据的准确性和可信度。

Arijit Nandi 基于皮电活动和光学脉搏波形生理信号数据进行融合，实验记录 10 名学生在课堂情景中激发学生的情绪反应，并记录其皮电活动(EDA)和光电容积脉搏波信号(PPG)数据。要求学生在回答完每个问题后，给出他们当前的情绪状态。对于学生输入的情绪状态，采用了广为接受的 Paul Ekman 提出的 6 种基本离散情绪：悲伤、快乐、恐惧、愤怒、惊喜、恶心。采用小波特征提取技术从 EDA 和 PPG 信号中提取特征，在[0,1]范围内对每个特征向量的值进行归一化处理，以补偿数值范围的差异。然后，将两个归一化特征向量进行特征融合并为最终特征向量。实验表明，与基于单一模态的方法相比，基于多种模态的特征的使用可以提供更高的分类准确性，其准确性和 F1 分数高达 0.9583 和 0.9583。但实验仅仅使用了 10 名学生的数据进行实验，这一样本量较小。对于情感识别任务来说，样本量的不足可能导致模型无法很好地捕捉到情感的多样性和复杂性，存在过拟合的风险。

2.2.决策融合（Decision Fusion）

决策融合指将不同模态提取后的特征在不同模态上独立训练模型，最后将它们的决策或分类结果进行融合。目前常用的决策层融合方法包括苏超（2020）所使用的决策树算法、张力行使用的加权平均法、Dempster-Shafer 证据理论、人工神经网络法等。

苏超采用学习者表情和姿态的双模态情感识别方法建立模型也是一种较为流行的方法。其模型主要包含两个模块，分别是情感识别模块和决策融合模块。首先，利用自主采集的数据集作为训练集和测试集，通过改进 Tiny_YOLOv3 目标检测算法得到 ER_Tiny_YOLOv3，从而进行面部表情和姿态识别为“专注”、“走神”、“疲惫”、“困惑”和“厌烦”5 种表情。然后，针对课堂教学环境，将表情识别结果和姿态识别结果在决策层面进行决策树融合，判断出学习者当前的学习状态。改进后的模型 ER_Tiny_YOLOv3 相比于原模型 Tiny_YOLOv3，mAP@0.5 提升了 17%，Precision 提升了 35%，而 F1 分数提升了 22.6%。相比

于 YOLOv4，mAP@0.5 提升了 19%，Precision 提升了 27%，F1 分数提升了 21.4%。然而面部表情和姿态等可能受到主观情感表达干扰，即每个人对于相同情感的表达可能存在差异，即使是同一种情感，不同个体可能会表现出不同的面部表情和姿态特征，或者人们在情感表达时可能会有意或无意地隐藏自己的面部表情和姿态特征。这种行为可能会对基于面部特征和姿态的情感识别产生影响，使其不够准确或可靠。相比之下，生理信号模态通常不容易受到人为控制，因此更可能反映出真实的情感状态。

张力行基于两种不同类型信号即集成皮肤电信号和文本信息的双模态情感识别系统被提出。首先，通过收集和分析学习者相应情感皮肤电信号的特征参数和文本信息的情感关键词特征，分别设计人工神经网络算法和高斯混合模型算法作为单模情感分类器，高斯混合模型加权决策融合进行 4 种情感识别，即高兴、悲伤、愤怒、平静。实验结果表明， 融合之后的情感识别模型对每类情感识别精度明显提高，平均识别率也由单模态的 70%, 80%提高到双模态融合后的 90%。但该方法需要手动提取皮肤电信号和文本信息的特征参数以进行情感识别，这可能需要大量的时间和精力，并且可能无法捕捉到所有重要的情感信息。此外，特征提取的过程可能会受到噪声和干扰的影响，导致特征表示的不准确性。

2.3. 混合融合（Hybrid Fusion）

混合融合是一种综合性的融合方法，它将特征融合和决策融合结合在一起，以更全面地处理多源信息的问题。适用于不同模态数据之间不仅有共享的信息，还存在互补的信息。能够综合利用数据间的相关性和差异性，提高情感识别的准确性和鲁棒性。

课堂学习下的多模态情感识别最常见的是基于视听数据的多模态情感识别方法（卢官明等，2019）。可以提取面部表情以及语音等各个模态的情感特征;然后，利用遗传算法对多个模态的情感特征进行特征融合，包括基于遗传算法的特征选择、交叉以及重组;最后，使用基于求和规则决策层融合进行情感分类。在 eNTERFACE'05 音视频情感数据库上进行情感识别测试，比较了基于面部表情或语音的单模态情感识别以及基于特征级和决策级融合的双模态情感识别的性能。实验结果表明，双模态情感识别的性能优于单模态情感识别，并且基于遗传算法的提出方法优于其他传统特征融合方法。这种方法在提高模型性能方面取得了一定的进展，但处理大量特征时，这可能会导致计算资源消耗较大，并且需要更多的时间来训练和测试模型。而课堂条件有限，不一定能够提供较大的算力和计算资源，这意味着系统必须能够快速接收、处理和解释来自不同源的数据。同时，在真实的课堂环境中，学习者的面部表情和姿态可能会受到多种因素的影响，例如光线变化、背景杂音、遮挡等，这些因素都可能导致数据噪音。这些噪音可能会对模型的训练和测试产生负面影响，使其难以准确识别学习者的情感状态。

虽然已有很多学者对混合融合的情感识别展开了研究，如 Tan 等人(2021)合并了为识别情感而创建的分层分类器的特征级和决策级融合结果;Georgescu & Ionescu(2019)创建了一个混合分类器来整合支持向量机和模糊认知图的结果；Cimtay(2020)等人在研究中使用了生物特征和软性模态，这些模态能够提高情感识别的准确率。为了进一步提高准确性，他们结合了在不同时间间隔下评估的生理模态。研究还显示，他们的模型在处理自然的欺骗性面部表情时表现出了高的准确性。然而，针对课堂学习场景下的情感识别，尤其是基于混合融合的数据处理方式，相关研究仍相对较为缺失。因此，如何将相关的研究迁移到课堂场景是值得我们进行探索的。

3. 课堂情感识别的挑战

（1）多模态情感识别中的不同模态对齐问题。不同模态的数据可能具有不同的特征表示方式，例如文本可以表示为词向量、图像可以表示为视觉特征、音频可以表示为声谱特征等。

因此，需要将这些不同模态的特征映射到一个共享的特征空间中，使得它们在该空间中具有相似的表示。不同模态的数据可能在语义上存在差异，即相同的情感可能以不同的方式在不同模态中表达。因此，需要考虑如何保持不同模态数据的语义一致性，以便在共享的特征空间中进行有效的情感分析。

（2）情感识别系统的准确性在很大程度上依赖于输入数据的质量。在教室场景下，多模态数据可能受到多种噪音的干扰，影响数据的质量和准确性。教室外的噪声、光照变化、遮挡物、学生的运动或皮肤状况等因素的干扰等会对音频、视频、生理信号等数据产生噪音从而影响模型。

（3）处理和分析多模态数据需要大量计算资源。如实时分析视频和音频数据、处理生理信号以及复杂的数据融合算法可能需要强大的计算能力。而课堂条件有限，不一定能提供充足的算力。因此系统必须能够快速接收、处理和解释来自不同源的数据，并保持高准确性。

4. 总结与展望

本文探讨了课堂学习环境下多模态数据融合情感识别的研究现状，对文献中多模态多种融合策略和方法进行综合分析，以整合来自不同数据集的情感信息，从而提高情感识别的准确性和可靠性。这些方法包括基于图像文本意见表示、语义词典嵌入表示以及特征提取层连接等多种手段、基于视听数据的融合、基于表情和姿态的方法以及基于生理信号模态的情感识别。并从多模态数据对齐、数据噪音干扰、课堂计算资源有限三个方面指出了当前研究所面临的挑战。

处理多模态情感数据对齐问题，可以设计一个共享的表示学习框架，通过联合训练多个模态的特征提取器，将不同模态的数据映射到共享的特征空间中。这样可以使得不同模态的数据在共享的特征空间中具有相似的表示，从而减少模态之间的差异性。在训练过程中，引入语义对齐的约束，以保持不同模态数据的语义一致性。可以通过引入额外的语义损失函数来实现，例如使用对抗性训练或多任务学习的方式，同时优化情感分类任务和语义对齐任务，以确保不同模态数据在共享的特征空间中的语义保持一致；

为了应对课堂环境中存在的噪声和干扰，可以采用噪声抑制技术和环境适应方法来提高数据的质量和稳定性。例如，可以使用语音增强算法来降低背景噪声对音频数据的影响，或者使用图像增强技术来减少视频数据中的噪点和遮挡物；

针对课堂算力限制问题，未来应该开发和使用轻量级的机器学习和深度学习模型、使用模型压缩技术、利用云计算资源以减小模型大小和计算需求以解决课堂计算资源有限问题。

目前在课堂场景中，混合融合的情感识别技术相对于特征融合和决策融合应用较少，但混合融合可以保持原始数据的完整性，在融合过程中不会丢失任何信息。相比之下，特征融合可能会丢失一部分原始数据的信息，而决策融合可能会丢失一些决策的细节或置信度。因此，混合融合能够更好地保持数据的完整性，提高模型的表现能力。特征融合和决策融合通常是基于已有的特征或决策进行操作，而混合融合可以在特征层面、模型层面或后处理层面进行多样化的融合操作，从而更好地适应不同的应用场景和需求。所以，开发轻量级混合融合的情感识别模型可以作为未来研究工作的一个侧重点。

基金项目

国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研

究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号 CCNU22JC011）

参考文献

卢官明,程晓,李霞,闫静杰&李海波.(2019)。基于遗传算法的多模态情感特征融合方法. **南京邮电大学学报(自然科学版)(05)**,41-47.

张力行,叶宁,黄海平 & 王汝传.(2018)。基于皮肤电信号与文本信息的双模态情感识别系统. **计算机系统应用(11)**,103-108.

苏超 & 王国中.(2020).课堂场景下学习者情感识别研究.**智能计算机与应用(11)**,8-15.)

蒋艳双,崔璨,逯行,祁彬斌 & 包昊罡（2022）。双师课堂中的多模态学习情感分析：关键问题、逻辑理路与实施路线. **现代教育技术(04)**,13-20.

Cárdenas-López, H.M., Zataraín-Cabada, R., Barrón Estrada, M.L., & Mitre-Hernández, H.A. (2023). Semantic fusion of facial expressions and textual opinions from different datasets for learning-centered emotion recognition. Soft Computing, 27, 17357-17367.

Chango, W., Lara, J.A., Cerezo, R., & Romero, C. (2022). A review on data fusion in multimodal learning analytics and educational data mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12*.

Cimtay, Y., Ekmekcioglu, E., & Caglar-Ozhan, S. (2020). Cross-Subject Multimodal Emotion Recognition Based on Hybrid Fusion. IEEE Access, 8, 168865-168878.

D’Mello, S.K., & Westlund, J.K. (2015). A Review and Meta-Analysis of Multimodal Affect Detection Systems. *ACM Computing Surveys (CSUR), 47*, 1 - 36.

Georgescu, M., & Ionescu, R.T. (2019). Recognizing Facial Expressions of Occluded Faces using Convolutional Neural Networks. ArXiv, abs/1911.04852.

Nandi, A., Xhafa, F., Subirats, L., & Fort, S. (2022). MDEAW: A Multimodal Dataset for Emotion Analysis through EDA and PPG signals from wireless wearable low-cost off-the-shelf Devices.

Shaik, T.B., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2022). Sentiment analysis and opinion mining on educational data: A survey. *ArXiv, abs/2302.04359*.

Tan, Y., Sun, Z., Duan, F., Solé-Casals, J., & Caiafa, C.F. (2021). A multimodal emotion recognition method based on facial expressions and electroencephalography. Biomed. Signal Process. Control., 70, 103029.

Zhu, H., Hu, P., Tang, X., Xia, D., & Huang, H. (2023). NAGNet: A novel framework for real-time students' sentiment analysis in the wisdom classroom. *Concurrency and Computation: Practice and Experience, 35*.

生成式人工智能变革教育的实践逻辑：系统性综述与未来议程

The Practice Logic of Generative Artificial Intelligence Transforming Education: A Systematic Review and Future Agenda

陈小娇[1]，胡哲冰[2]，王成梁[3*]
[1]浙江工业大学 教育科学与技术学院
[2]浙江工业大学 外国语学院
[3]华东师范大学 教育学部
[*] wcledutech@163.com

【摘要】 ChatGPT 作为生成式人工智能(GAI)产品的代表，在教育领域引领创新。本研究采用系统性文献综述的方法，在四个数据库中筛选了 134 篇与 GAI 教育应用相关的文献作为研究对象，研究发现如下：1) 美国是领域最活跃国家，澳大利亚、中国、加拿大次之；该领域的研究以理论研究为主，主要发表在教育技术学和医学领域刊物与会议；2) 研究主题集中在 GAI 技术性能、教学辅助、学习绩效、应用优劣分析和未来展望；3) 扎根理论深入探讨 GAI 教育应用的核心优势和潜在风险；4) 通过文献回顾提供理论和实践层面的未来议程，推动更智能、有效和可持续的教育方法与工具发展，促进教育领域创新。
【关键词】 生成式人工智能；ChatGPT；教育技术；系统综述；未来议程

Abstract: *ChatGPT, as a representative of Generative Artificial Intelligence (GAI) products, leads innovation in the field of education. This study adopts a systematic literature review approach, selecting 134 relevant documents on GAI educational applications from four databases as research objects. The findings are as follows: 1) The United States is the most active country in the field, followed by Australia, China, and Canada. Research in this field is predominantly theoretical and primarily published in journals and conferences in the fields of educational technology and medicine; 2) Research topics focus on GAI technological performance, instructional assistance, learning outcomes, application analysis, and future prospects; 3) Grounded theory deeply explores the core advantages and potential risks of GAI educational applications; 4) Through literature review, theoretical and practical agendas are provided to promote the development of more intelligent, effective, and sustainable educational methods and tools, fostering innovation in the field of education.*

Keywords: Generative Artificial Intelligence; ChatGPT; Educational Technology; Systematic Review; Future Agenda

1. 前言

人工智能赋能教育的价值一直以来都被学者广泛关注(王成梁 等, 2023)，技术发展始终在推动着这一研究领域的快速。最近，预训练的生成式人工智能(Generative Artificial Intelligence, GAI)的发展再次点燃了学者们对这一领域的研究热情。生成式人工智能是利用先进的技术如生成对抗网络和大型预训练模型，通过分析和学习现有数据，实现创造性内容生成的一种新型人工智能(顾小清 等, 2023)。作为 GAI 的一个典型代表，ChatGPT 受到了广泛关注，其在教育领域的应用正受到积极探索(Lee et al., 2020)。教育界对 GAI 的兴趣主要集中在其如何提升教学资源的获取、学习互动和知识传播方面(Tsang, 2023; Escotet, 2023)，尽管也有对其可能促进不正当学术行为的担忧。

GAI 在教育领域的应用展现了教学个性化和自适应的潜力，这一点获得了众多学者的认可。然而，关于如何有效融合 ChatGPT 等技术于教育的实践，目前尚处于初级阶段，面临理

解偏差和实施挑战(Alnaqbi & Fouda, 2023)。因此，系统评估教育领域中 GAI 的应用，并作出理性判断，变得尤为重要。

为此，本研究通过系统的文献综述方法，深入探讨了 GAI 技术在教育领域的应用，包括案例分析和教育实践。研究总结了利用 GAI，特别是 ChatGPT，在教育中的主要优势，如提高教学质量、增强学生学习体验和促进个性化学习路径。同时，也指出了 GAI 带来的风险，包括数据隐私保护、算法偏见和对学生自主学习能力的潜在影响。基于这些发现，本研究提出了未来 GAI 在教育应用中应考虑的策略和方向,旨在为教育工作者和政策制定者提供指导，以促进 GAI 技术在教育领域的有效和伦理应用。具体而言，本研究主要分析以下研究问题：

研究问题 1(RQ1)：GAI 赋能教育研究领域当下整体发展趋势如何?

研究问题 2(RQ2)：GAI 赋能教育研究领域当下存在哪些核心研究议题？

研究问题 3(RQ3)：GAI 赋能教育的核心优势和潜在风险有哪些？

研究问题 4(RQ4)：基于已有文献的梳理，GAI 教育应用研究领域存在哪些有研究价值的未来议程？

2. 研究方法

本研究采用了系统性文献综述(Systematic Literature Review, SLR)的方法，SLR 是一种明确、系统的文献综述方法，需要研究者使用透明的方法对特定问题的研究成果进行收集、综合和评价，并使用科学方法论界定相关文献，评价研究质量，总结研究结果，进而精准掌握该研究主题的研究现状与发展趋势，以实现对特定研究问题的回答(Jing et al., 2023)。SLR 的优势在于严谨、透明，包括清晰的研究问题、全面的检索策略、明确的文献标准、高质量的评估方法、综合的数据分析以及可靠的研究结果，故能有效克服传统研究方法的主观性、偏见性等问题。在本研究中，由于研究问题(引言尾部的问题 1 到问题 4)与分析对象(GAI 教育应用)均十分明确，因此十分适合运用 SRL 的方法进行分析研究。

首先，研究通过规范化的 PRISMA 筛选流程实施文献筛选(Moher et al., 2009)，然后建立编码体系，收集文献良构的编码信息(例如统计发文量、期刊、研究类型等)和非良构的文本编码信息(例如对 GAI 技术核心优势和潜在风险的论述)。本研究对于良构信息进行统计和可视化呈现，对非良构的文本编码信息，则运用分析质性文本的方式进行。本研究选用了四个英文文献数据库：EBSCO、Ei Compendex、Scopus 和 Web of Science。在检索条件的设置上，本研究参考了若干 GAI 领域已有文献的关键词列表，确定了适当的关键词。首先在四个数据库中识别记录，分别是：Web of Science 有 1465 条，Scopus 有 886 条，Ei Compendex 有 330 条，EBSCO (ERIC)有 13 条。在删除重复的记录后，共有 2243 条记录。接着，基于标题和摘要过滤掉 2005 条记录。之后排除了 104 条与研究主题不相关或相关性低的记录。最终，有 238 条记录的全文被纳入研究。在进一步审查后，有 134 条研究记录被包括在综述中，其中包含 106 篇文章和综述，以及 28 篇其他类型的文献，如社论(Editorial)、观点(Perspective)、信件(Letter)和笔记(Note)。

3. 绩效分析

绩效分析是从宏观上获取某议题的基础信息，以求客观呈现该学科领域的发展趋势和走向。本研究将对纳入分析文献的年份分布、研究类型、研究方法、隶属期刊和国家层面核心生产力等量化数据进行探究，以期从量化视角呈现这一领域的学科研究发展态势。

本研究首先对所纳入文献发表时间进行了统计，结果如图 1-1 所示，深绿色表示 Article 和综述类型，浅绿色表示其他类型。

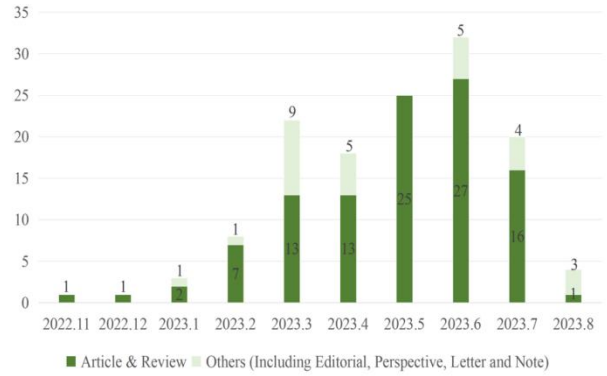


图 1-1. 纳入文献发表的年份分布

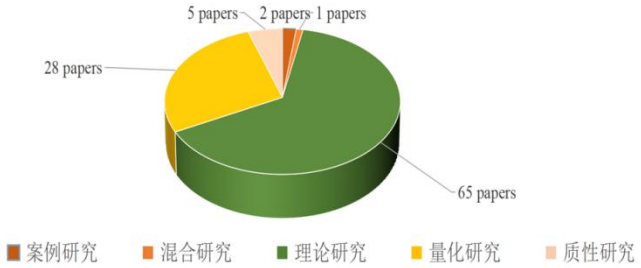


图 1-2. Article 类型文献的研究类型

进一步对研究类型进行统计，依据编码时对研究类型的界定，本将研究类型分成 5 类，分别为量化研究、质性研究、混合研究、案例研究和理论研究。统计结果如图 1-2 所示。

研究还对量化研究、质性研究和混合研究的文献(共计 34 篇)作了具体研究方法的统计，具体统计结果如图 2-1 所示。

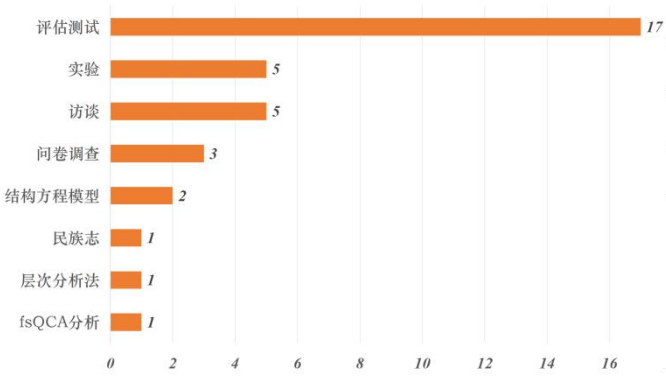


图 2-1. Article 类型文献的研究方法

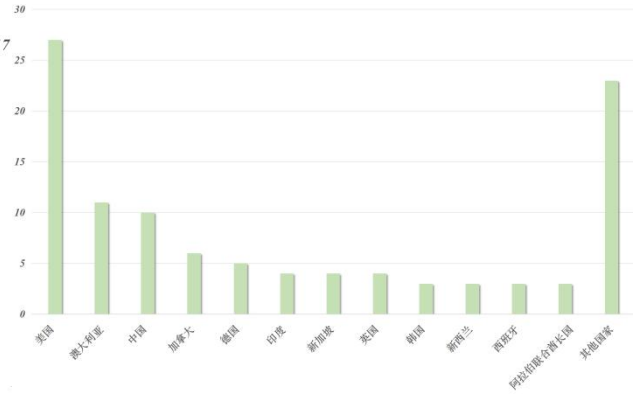


图 2-2. Article 类型文献的国家分布

为了进一步对该研究领域国家层面生产力分布状况进行分析，我们以第一作者所属国家作为统计依据，发现 106 篇文献来自 32 个国家，发文量达到 2 篇及以上的国家如图 2-2 所示。

研究主题在学科领域的发展中具有重大价值，准确把握研究领域所探讨的重要议题，可以帮助研究者更好地了解研究领域的现状，并对领域未来发展进行合理预测。本研究对纳入的 106 篇 Article & Review 类型文献的研究议题用主题编码的形式作了整体归纳，发现主要有 5 类，如表 1 所示(论文中的编码对应的是编码列表中的编码，读者可以根据需要向通讯作者索取)。

表 1. GAI 教育应用的研究主题

编号	研究主题	文献编号
1	GAI 技术性能评估	1, 2, 5, 6, 17, 21, 22, 26, 30, 43, 49, 55, 60, 70, 73, 104,
2	GAI 技术教学辅助应用	4, 16, 20, 33, 48, 53, 96
3	GAI 技术提升学习绩效	8, 23, 32, 38, 54, 57, 58, 69, 77
4	GAI 技术教育应用优劣分析	7, 9, 13, 24, 28, 37, 46, 47, 62, 79, 83, 88, 89, 103
5	GAI 技术教育应用前景展望	10, 15, 19, 39, 41, 45, 51, 59, 66, 67, 72, 75, 76, 80, 82, 84, 90, 99

4. 扎根分析

针对 GAI 教育应用核心优势的部分，基于开放式编码分析，我们通过人工编码，最终从文献中提炼出了 29 个初始概念，包括虚拟教学助理、整体学习体验、出色理解能力、提供特定主题的信息和资源等。在初始概念的基础上，我们进一步总结出 15 个初始范畴，涵盖了教

学助理、学习体验、反馈和指导、个性化学习、促进理解、学习工具、处理能力、应用广泛、提供知识技能、人机交互、提供资源、加强教育、评估能力等要素。在主轴编码阶段，本研究对核心优势通过进一步归纳概括，得出教学援助、学习辅助、个性化反馈、性能优异、交互便捷 5 个主范畴(所有编码的过程性信息可以联系通讯作者获取)。

对 GAI 核心优势的讨论必然也会催生学者们对其应用风险的关注。通过开放式编码对纳入文献中潜在风险维度进行提炼和初步编码后，最终归纳了 31 个初始概念，包括学术不诚实、考试和作业作弊、存在偏见和歧视的风险、数据泄露等概念。在此基础上，本研究进一步提炼形成了 15 个初始范畴，即不道德行为、信息错误、诚信问题、知识有限、准确性与可靠性、过度依赖、伦理道德问题、剽窃行为、网络安全问题、信息传播风险、隐私问题、滥用问题、限制思维、技术垄断、情感交互。针对潜在风险的初始范畴，在归纳总结后得出了产品风险、伦理挑战、隐私风险、科技垄断和情感交互困难这 5 个主要范畴。

最后是选择式编码，选择性编码的核心目标是整合已识别的范畴和概念，形成一套系统的理论。这要求研究者在验证范畴关系的基础上，构建一个完整的故事线，解释研究现象的成因、过程和结果。在选择性编码阶段，研究者应不断反思理论的完整性和准确性，并可能需要回到原始数据中寻找证据支持理论构建。此外，还可以通过与同行的讨论或向研究参与者求证来增强理论的可靠性。本研究遵循选择性编码的一般思路，运用扎根理论对核心优势及潜在风险的编码结果，本研究进一步构建了 GAI 教育应用的内在机制与逻辑架构，如图 3 所示的框架模型。

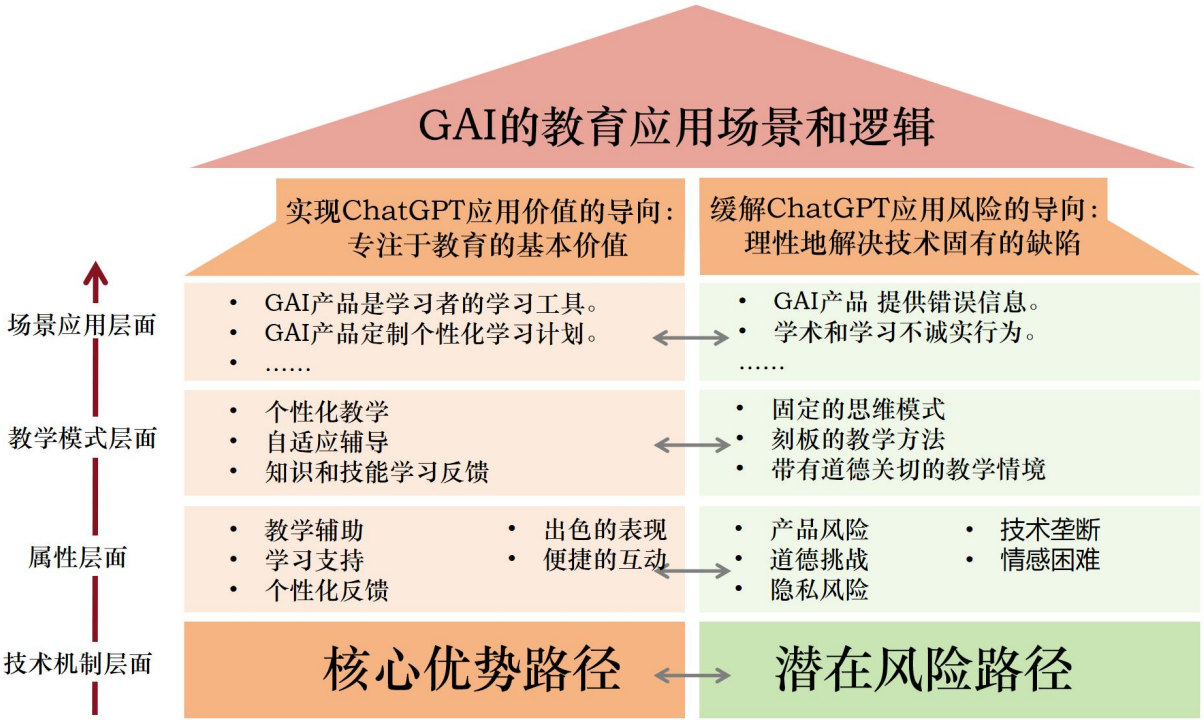


图 3 GAI 教育应用的内在机制与逻辑架构

在构建的这一框架模型中，两条路径—核心优势路径和潜在风险路径—以及四个层面—技术机制层面、属性层面、教学模式层面和场景应用层面—共同描绘了 GAI 技术在教育中应用的全貌。首先，技术机制层面关注的是 GAI 技术的基础结构和运作原理，这是理解其在教育领域应用的起点。随后，属性层面着眼于这些技术特性如何与教育需求相匹配，进而促进或制约教育效果的提升。

在教学模式层面，模型探讨了 GAI 技术如何与现有的教学策略相融合，或促使新教学模式的形成，以实现教学方法的创新与优化。最后，场景应用层面则聚焦于 GAI 技术在具体教育环境中的应用，如何通过实证研究与实践探索，有效地将 GAI 技术集成到真实的教学和学习场景中，从而实现教育目标。

通过深入分析这四个层面的交互和相互影响，本框架强调从技术本身出发，既要充分挖掘 GAI 在教育中的应用潜力，强化其核心优势，又要警惕其潜在风险，确保技术应用不偏离教育的本质目标。研究者和教育从业者需要采取理性的视角，评估 GAI 技术的双刃剑效应，制定周密的应用策略和风险管理计划。通过这种平衡和整合的方法，可以更加有效地利用 GAI 技术优化教育过程，为学习者提供更加丰富、有效和个性化的学习体验，同时为教育系统的可持续发展奠定基础，确保教育能够适应未来社会和技术的快速发展。

5. 未来议程

本研究将从理论研究层面和实践应用层面对未来议程进行系统陈述。首先，在理论研究层面，通过对已有 GAI 理论研究和实证研究的深度分析，本研究提供了若干关于在教育情境中利用 GAI 核心优势和规避 GAI 潜在风险的未来研究议题；其次，在实践应用层面，本研究指出了 GAI 技术教育应用实践中 2 个有价值的深刻的未来议程，分别是 GAI 教育专有化产品的开发和 GAI 技术推广可能面临的问题及措施。未来议程的整体逻辑与内在关系如下图 4 所示。

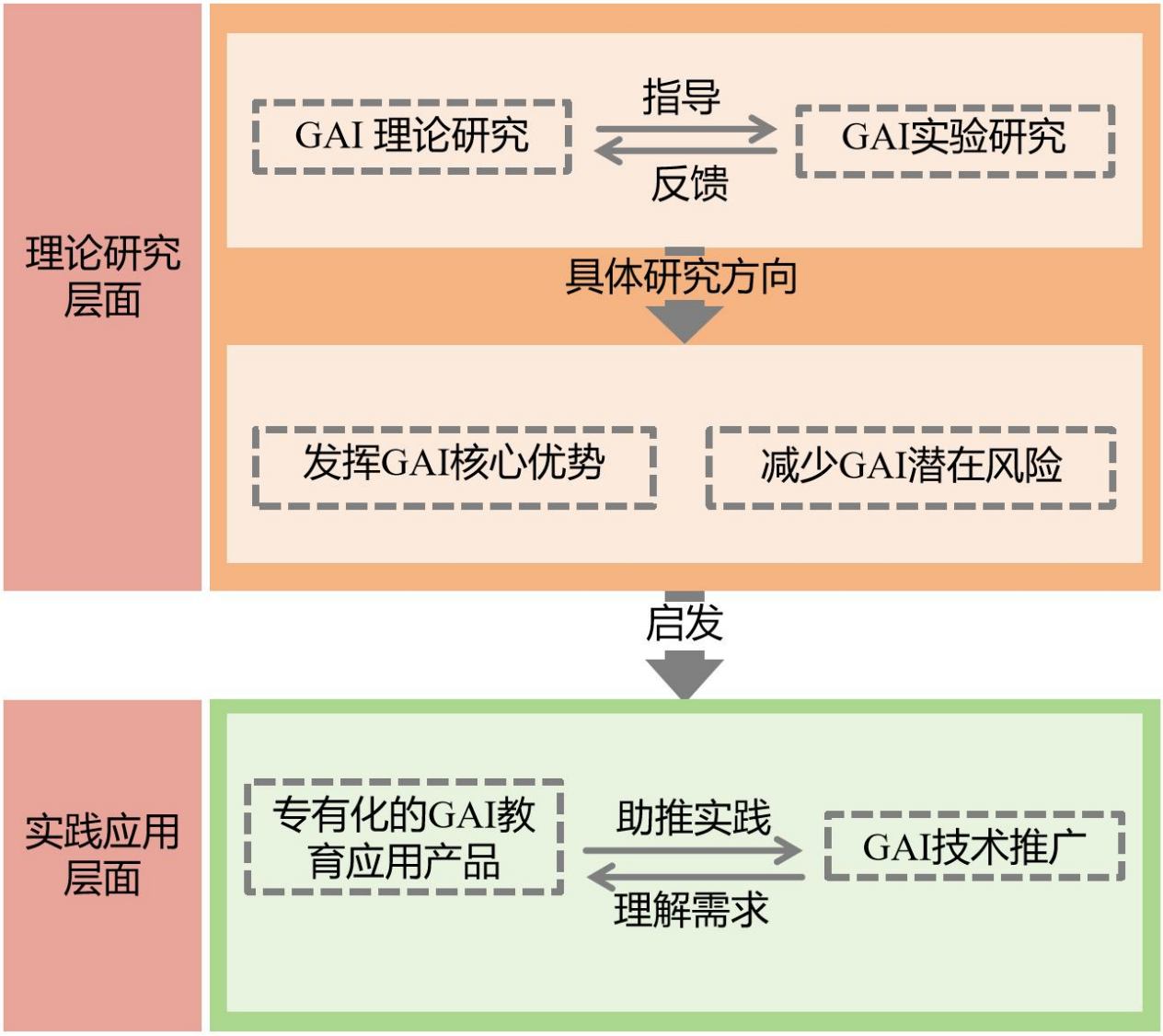


图 4 未来议程的宏观逻辑图

5.1. 理论层面研究议程

5.1.1. 发挥 GAI 的核心优势

将 GAI 技术的核心优势充分挖掘，是未来研究的关键路径。这一路径不仅将推动教育领域的创新与进步，还有望为学生、教师以及教育体系带来更广泛的益处。因此，深入研究和

开发 GAI 在教育中的应用潜力，以及与之相关的多元领域，将在未来研究中发挥重要作用。在该研究方向上，具有以下几个值得进一步探究的未来议程。第一，构建多元交流的 GAI 辅助教学体系。第二，完善基于 GAI 技术的教育评估与反馈机制。教育评估成功的关键是使用形成性反馈(Schmulian & Coetzee, 2019)。第三，建立 GAI 主导的个性化互动式教学。

5.1.2. 减少 GAI 潜在风险

除了充分发挥 GAI 技术的优势之外，避免潜在风险也应成为未来研究议题的一个重要方面。在这一研究方向上，存在着众多待探讨的议程，它们将有助于确保 GAI 技术在教育领域的安全、有效应用。我们需要深入研究并解决 GAI 技术可能引发的潜在挑战和风险，以确保学生、教师和整个教育体系的利益不受损害。在这个重要领域，我们可以探讨以下几个值得进一步研究的议程，以促进 GAI 技术的负面影响最小化并最大化其潜力的利用。第一，规范并加强监管 GAI 技术教育应用。随着 GAI 融入我们教育教学活动，规范和监管能力很重要(Shoja et al., 2023)。第二，培育学习者 AI 素养和 AI 应用伦理价值观。第三，建构多模态数据交互的 GAI。

5.2 应用层面未来议程

5.2.1 开发专有化的 GAI 教育应用产品

学术研究可以为产品的应用实践提供重要的理论指导，尤其在涉及开发专有化 GAI（人工智能生成内容）教育应用产品时更为关键。这一步骤被视为迈向智能教育体系的第一步，具有巨大的潜力来推动教育领域的创新和改进。在以下内容中，我们将探讨开发专有化 GAI 教育应用产品的未来研究方向，以期在应用层面为研究主题提供有效方针。第一，构建可信赖的 GAI 教育应用产品。第二，推动建立多模态数据融合的 GAI 教育应用产品。第三，开发特定教学场景下的 GAI 产品。

5.2.2 推广 GAI 技术应用的问题和措施

技术推广过程中可能面临的问题将会启示未来研究的方向道路。任何一项技术从诞生到应用推广，都会经历一个漫长的过程。其中每个阶段会有一些问题存在，并且许多问题不是先验性的，而是具有这个领域独有的特殊性。这些问题启示我们在未来的应用过程中更应该注意相关问题。为了充分应对相应问题，未来研究可以从拓展受众、拓展方法和加强跨学科研究这三个方向努力。第一，拓展 GAI 技术在教育生态中的受众。第二，丰富调查 GAI 技术接受意愿的研究方法。第三，加强跨学科研究是促进 GAI 技术推广的重要措施。

6. 结论与展望

6.1 研究结论

本研究通过对 134 篇 GAI 教育应用研究文献的系统性综述分析，回答了研究最初提出的四个问题。首先，在生产力方面(RQ1)，美国在该领域是最主要的发文国家，其次是澳大利亚、中国和加拿大。这表明美国在该领域拥有广泛的研究活动和领先地位，而澳大利亚、中国和加拿大也在积极参与和贡献，显示出该领域的国际性重要性和合作潜力。对纳入分析文献的研究类别进行统计发现，理论研究是最常见的研究类型，这反映了在该领域的学术界和研究者们对于理论框架和基础知识的重视和关注，以推动对相关主题的深入理解和探讨。

其次，在研究主题方面(RQ2)，GAI 在教育领域的研究主题包括 GAI 技术的性能评估、GAI 技术辅助教学、GAI 技术提高学习成效、GAI 技术的潜在影响、GAI 技术的发展趋势，这五个研究主题深刻揭示了 GAI 技术对教育领域的变革性作用。

此外，在 GAI 教育应用时核心优势和潜在风险方面(RQ3)，研究运用扎根理论深刻解构了以 ChatGPT 为代表的 GAI 技术的优劣势的产生机理，并指出了发扬优势和规避劣势的技术应用模式。

最后，基于对已有文献的梳理，研究从理论层面和应用层面指出了一些在未来具有分析价值的议题(RQ4)，在理论层面，未来研究需要考虑已有的教育学理论是否适配新技术创设的学习情境，能否应用这些理论来指导我们发扬新技术优势，规避新技术弊端；在实践应用层

面，未来研究需要开发隶属于教育领域的专有大模型，使 GAI 能更好地服务于当下的教育实践，此外还需要关注推广过程中可能的问题以及相关的对策。对未来研究议程的讨论将有助于明确 AI 与教育融合的关键问题，推动更智能、有效和可持续的教育方法和工具的发展。

6.2 未来展望

展望未来，随着如 Sora、Claude 3 等先进技术在 GAI 领域的迅猛发展，我们必须认识到，教育方法和教学策略需要与技术进步保持同步，以充分利用这些新兴技术带来的巨大潜力(Wang et al., 2024)。这不仅要求教育者掌握最新的技术知识，更要求他们对传统的教学法进行深刻的反思和可能的重大变革，确保教育实践能够有效应对技术革新带来的挑战和机遇。

在这个过程中，教育体系中的教与学都可能需要顺应技术发展的趋势进行根本性的转变。教育者需要以实证的形式探索如何将 GAI 技术融入课程设计、教学方法和学习评估中，以促进更加个性化和高效的学习体验(例如 Jing et al., 2024 的准实验研究)。同时，学生们也需学习如何利用这些技术来优化自己的学习过程和结果。

此外，为了确保技术发展能够真正服务于教育的核心价值和目标，必须对这些变革进行周密的规划和评估，确保它们能够促进教育公平、提高教育质量，并培养出能够适应未来社会需求的学生。因此，教育政策制定者、研究者、教育技术开发者和教育实践者之间的紧密合作显得尤为重要，共同推动教育领域的创新和变革，以应对快速变化的技术环境。

致谢

本研究被国家大学生创新创业训练项目计划(202310337023)资助。

参考文献

顾小清,胡艺龄,郝祥军.AGI 临近了吗：ChatGPT 热潮之下再看人工智能与未来教育发展[J].华东师范大学学报(教育科学版),2023,41(07):117-130.

王成梁,戴坚,留伊丹 & 王永固.(2023).人工智能辅助教育应用研究的全球知识图谱：发展路径和前沿动态——基于 CNKI 与 WoS 的文献计量分析. 软件导刊(01),38-45.

Alnaqbi, N. M., & Fouda, W. (2023). Exploring the role of ChatGPT and social media in enhancing student evaluation of teaching styles in higher education using neutrosophic sets. *Int. J. Neutrosophic Sci*, 20(4), 181-190.

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy.*

Escotet, M. Á. (2023). The optimistic future of Artificial Intelligence in higher education. *Prospects*, 1-10.

Jing, Y. H., Wang, C. L., Chen, Y., Wang, H. M., Yu, T., & Shadiev, R. (2023). Bibliometric mapping techniques in educational technology research: A systematic literature review. *Education and Information Technologies*, 1-29. https://doi.org/10.1007/s10639-023-12178-6

Jing, Y. H., Wang, H. M., Chen, X. J., & Wang, C. L. (2024). What factors will affect the effectiveness of using ChatGPT to solve programming problems? A quasi-experimental study. *Humanities & Social Sciences Communications*. 11, 319. https://doi.org/10.1057/s41599-024-02751-w

Lee, L. W., Dabirian, A., McCarthy, I. P., & Kietzmann, J. (2020). Making sense of text: artificial intelligence-enabled content analysis. *European Journal of Marketing*, 54(3), 615-644.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G.PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med*. 151, 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135

Schmulian, A., & Coetzee, S. A. (2019). Students' experience of team assessment with immediate feedback in a large accounting class. *Assessment & Evaluation in Higher Education*, 44(4), 516-532.

Shoja, M. M., Van de Ridder, J. M., & Rajput, V. (2023). The Emerging Role of Generative Artificial Intelligence in Medical Education, Research, and Practice. *Cureus*, 15(6). e40883. DOI 10.7759/cureus.40883

Tsang, R. (2023). Practical Applications of ChatGPT in Undergraduate Medical Education. *Journal of Medical Education and Curricular Development*, 10, 23821205231178449.

Wang, C. L., Chen, X. J., Yu, T., Liu, Y, D. & Jing, Y. H. (2024) Education Reform and Change Driven by Digital Technology: A Bibliometric Study from a Global Perspective. *Humanities & Social Sciences Communications*. 11, 256. https://doi.org/10.1057/s41599-024-02717-y

结合教学属性的中小学古诗词图谱构建与分析

Construction and Analysis of Ancient Poetry Graph in Primary and Secondary Schools Combined with Teaching Attributes

丁国柱[1*]，冯心茹[1]
[1]广州大学 教育学院
[*] dinggz@gzhu.edu.cn

【摘要】 文章结合文学知识、内容属性和教学属性，搭建了一个包含 665 个类和 43 种关系的知识本体，并构建了一个具有 4 777 个节点、13 667 个三元组和 14 268 条关系的中小学古诗词知识图谱。通过 Word2Vec 词嵌入技术对 7 050 篇诗词教案进行文本分析，计算不同教育目标和八大类知识点（包括字、词、句、篇、语言、修辞、逻辑、文学）之间的余弦相似度。以“运用”目标和修辞类知识点为例，分析不同学习阶段中教育目标与知识点的学习比重。最后，该中小学古诗词图谱直观地展示了诗词教学目标中的重难点，并对教学特征进行了量化分析，这有利于提高课程教学目标的清晰度。

【关键词】 教学属性；本体；知识图谱；Word2Vec；余弦相似度

Abstract: *By combining with literary knowledge, content attributes and teaching attributes, this paper constructed a knowledge ontology containing 665 categories and 43 relationships, and created a knowledge graph of Chinese poetry in primary and secondary schools with 4 777 nodes, 13 667 triples and 14 268 relationships. Utilizing 7 050 instructional designs and applying Word2Vec for text analysis, the study calculated the cosine similarity between different educational objectives and eight knowledge points, including character, word, sentence, article, language, rhetoric, logic and literature. The analysis focused on the "application" goal and rhetorical knowledge points as illustrative examples to examine the learning proportion of educational objectives and knowledge points across different learning stages. Finally, the graph of Chinese poetry in primary and secondary schools intuitively shows the important and difficult points in the poetry teaching objectives, and quantitatively analyzes the teaching characteristics, which is conducive to improving the clarity of the course teaching objectives.*

Keywords: teaching attributes, ontology, knowledge graph, Word2Vec, cosine similarity

1. 前言

古诗词是中华优秀传统文化的重要组成部分，在中小学语文学科教育中占据着重要地位。构建语文学科古诗词知识库对提升教学系统的智能性具有巨大意义，虽然目前已有很多关于古诗词知识图谱构建的研究，但主要集中在诗词内容、背景、赏析等知识属性方面，缺乏对教学属性的描述。中小学古诗词知识图谱的构建不仅要关注知识内容本身，还应重视知识点与教学的结合，将教学目标、教学重难点等与教学相关的知识综合考虑进来。文章主要从以下两个方面展开：

（1）融合文学知识、内容属性和教学属性，搭建一个共包含 665 个类和 43 种关系的中小学古诗词知识本体，并以此构建一个具有 4 777 个节点、13 667 个三元组和 14 268 条关系

的中小学古诗词知识图谱。该知识图谱将多源异构的中小学古诗词数据融合在一个相对完整的知识框架中，并通过语义化的知识组织形式，实现诗词知识点的全局展示。

（2）利用文本分析技术，挖掘潜藏在诗词学习过程中的教学知识。文章使用 Word2Vec 词嵌入技术以“运用”目标和修辞类知识点为例，分别从同一个知识点下教育目标在不同学习阶段的学习比重，以及同一教育目标下知识点在不同学习阶段的学习比重这两个维度进行分析，量化中小学古诗词教学的特点，并将结果整合到中小学古诗词知识图谱中。

2. 相关工作

2.1. 古诗词在数字人文领域的研究

在数字人文领域，使用计算语言学和统计学方法辅助诗词研究已成为普遍趋势。越来越多的学者利用各种分析技术深入挖掘古诗词的重要信息，从而构建古诗词知识图谱。例如，李永卉等(2021)构建了宋代镇江的知识图谱，对诗词的背景、情感和地域特征等进行词频统计和数据分析；刘昱彤等(2020)基于诗词的注释和词典解释，构建了一个包含词语语义联系的古诗词知识图谱，并进行了诗词情感分析等相关推理与分析实践。这些古诗词知识图谱在诗词数据分析中扮演着不可或缺的角色，正是利用各种分析技术可以更好地深入挖掘古诗词资源中的新关系。

2.2. 古诗词在知识本体、知识图谱领域的研究

在 20 世纪 80 年代，本体论思想被引入知识工程领域，并成为当时表示现实世界知识的主要方式。知识本体的形式化表达对共享领域模型的概念和语义关系起到了关键作用，也为知识图谱的产生奠定了基础。2012 年，谷歌提出了知识图谱的概念，自此构建了一个庞大而有机的知识网络，旨在理解事物之间的关联。在这一过程中，本体为知识的组织、表示和推理提供了关键支持(Chen et al., 2020)。

2.3. 教育领域古诗词知识图谱构建研究

当前教育领域古诗词知识图谱的研究已取得一定进展，但仍存在一些问题和不足。例如，冯俐(2019)设计构建的中学语文诗词知识图谱仅关注诗词作者、作品和年级之间的关系，缺乏中小学古诗词知识体系的本体模型建构，导致在内容完整性和结构合理性方面存在一些不足；彭畅(2021)根据中小学语文诗词的特点，设计了诗词知识图谱的总体架构，但分析仅限于诗词的内容属性，忽略了诗词题材、思想情感等文学知识的重要性；周莉娜等(2019)设计的基于唐诗知识图谱的智能知识服务平台，虽然针对诗学、文献学和史学三大特定领域的唐诗本体进行研究，但忽略了古诗词的教学属性，如针对不同学习阶段学生的教学目标和教学重难点等方面的考虑。

综上所述，目前的诗词知识图谱研究主要集中在课程知识、学科知识等内容属性方面，较少结合教学属性进行具体描述。有些学者过于关注诗词知识内容本身，却忽略了中小学学生认知发展差异大、知识习得具有差异性的特点。教材作为教学内容的重要载体，直接影响教育教学的效果(任翔，2022)。基础教育领域语文知识的组织方式与古诗词教学组织存在特殊性，如古诗词学习的顺序、学习者所处的学习阶段、教学的课程标准等方面密切相关。因此，在古诗词知识图谱的构建过程中，不能仅从知识内容的角度出发，还应考虑古诗词的教学目标、教学重难点等要素，并将这些要素相互关联，共同构建一个相对完整的古诗词知识体系。这样不仅能分析不同学习阶段学生认知发展规律的特点，还能把握整个义务教育语文诗词教学的基本目标，从而提高课程教学目标的清晰度。

3. 中小学古诗词知识图谱的设计与构建

文章利用 protégé 软件构建本体概念模型，以 python 语言为开发工具，参考部编版中小学语文教材样本，收集网络资源和中华古诗词数据集形成语料库，使用 neo4j 工具存储古诗词图谱数据。通过知识本体模型的构建、数据采集、知识抽取与融合、知识存储四个步骤构建中小学古诗词知识图谱。

3.1. 知识本体的构建

本节从语文知识内容、古诗词内容和古诗词教学这三个维度出发，利用 protégé 软件搭建了一个共包含 665 个类和 43 种关系的中小学古诗词知识本体。

3.1.1. 语文知识内容本体的构建

以部编版中小学语文教材的诗词内容为基础，结合古诗词内容的特殊性，最终确定从汉字、词汇等 12 个方面构建语文知识内容本体模型，如图 1 所示。

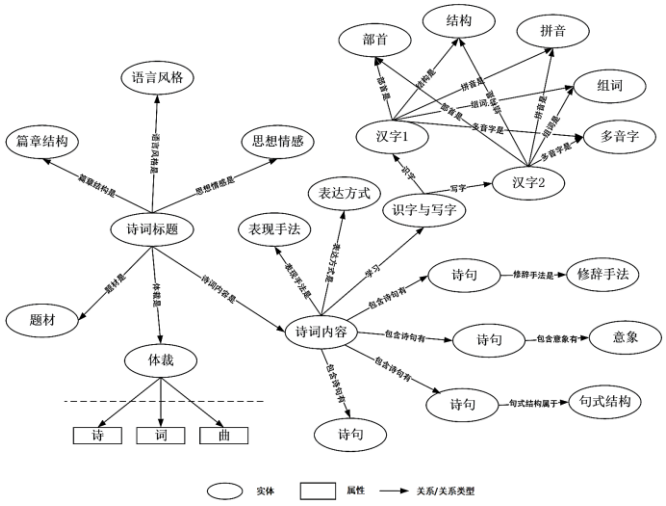


图 1 语文知识内容本体

3.1.2. 古诗词内容本体的构建

下面进行古诗词内容本体模型的构建，主要包括作者、朝代、诗词内容等，最终得到古诗词内容本体模型，如图 2 所示。

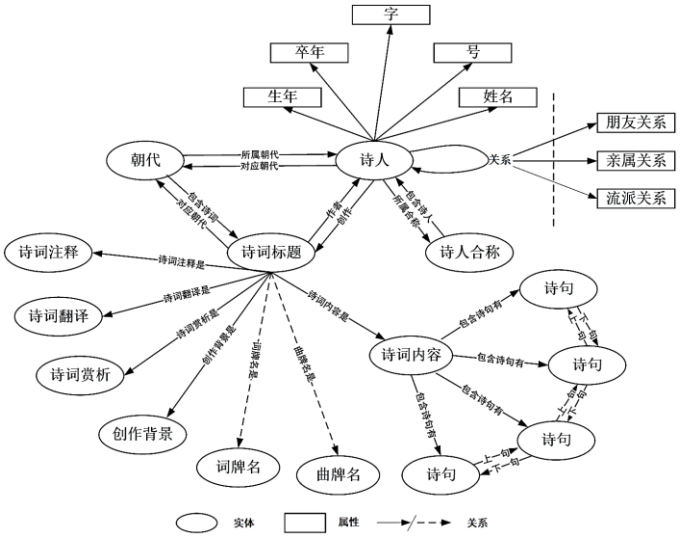


图 2 古诗词内容本体

3.1.3. 古诗词教学本体的构建

文章遵循中小学语文课程标准的行文逻辑,以教学目标为起点和指导，根据学生的认知水平和学习能力，在不同学习阶段设定相应的教学目标和教学重难点，综合而成的古诗词教学本体模型如图 3 所示。

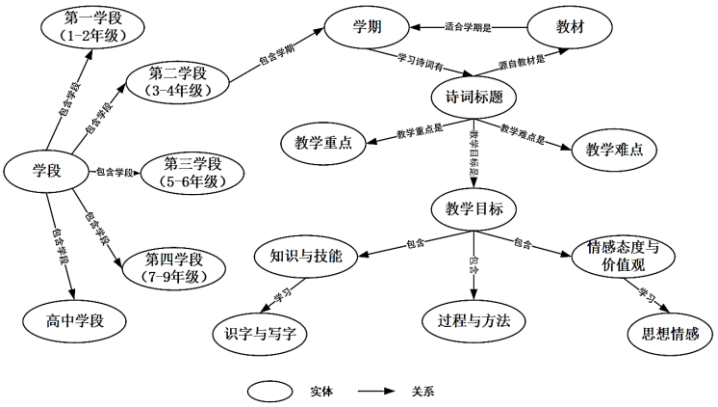


图 3 古诗词教学本体

3.2. 数据采集

本节参考部编版中小学语文电子教材和中华古诗词数据集，使用 python 语言从古诗文网站获取诗词数据，并通过补充来自诗词名句网等网络资源的缺失值，最终通过人工整理，建立了中小学古诗词的基本语料库。

3.3. 知识抽取与融合

完成数据采集后，鉴于诗词所蕴含的丰富人文性知识，文章采用人工标注方法逐条筛选数据。从网络资源获取了 14 372 条数据，经过去重、知识抽取与融合等一系列操作，最终得到 7 019 条可供使用的数据，其中包括 235 首中小学古诗词和 94 名诗人。根据数据的属性特点和彼此间的关系，以 CSV 格式进行结构化保存，并为其赋予相应的字段标签。

3.4. 知识存储

本节将数据以三元组的方式进行表示，基本形式主要包括<实体 1，关系，实体 2>和<概念，属性，属性值>(徐增林等人，2016)。中小学古诗词知识本体主要包含的数据见表 1 ，共计构建的知识实例数 4 777 个，对应的三元组有 13 667 个，实际上构建的关系数量为 14 268 条。

表 1 中小学古诗词知识本体数据

概念	数量	三元组数	主要关系
汉字	734	1 496	识字有、写字有、部首是、结构是、字形是、拼音是、多音字是
词汇	151	222	组词是、近义词是、反义词是
句子	2 598	5 814	上一句是、下一句是、包含诗句有
题材	15	171	题材是
体裁	14	232	体裁是
意象	66	551	包含意象有
修辞手法	29	1 142	修辞手法是
表现手法	29	88	表现手法是
思想感情	53	366	蕴含情感有
语言风格	29	347	语言风格是
作者	95	505	创作、所属合称、相识、所属朝代
古诗词	236	1 788	作者是、对应朝代是、诗词内容是、诗词注释是、诗词翻译是、诗词赏析是、创作背景是、词牌名是、曲牌名是
教学重点	236	236	教学重点是

教学难点	236	236	教学难点是
教学目标	236	236	教学目标是
学期	20	237	学习诗词有
总计	4 777	13 667	

4. 结合教学属性对古诗词图谱进行分析

本节将中小学语文教材的学习阶段分为四个阶段：小学低年级阶段（1~3 年级）、小学高年级阶段（4~6 年级）、初中阶段和高中阶段。基于 7050 篇教学设计案例利用 Word2Vec 进行词嵌入模型训练，从同一个知识点下教育目标在不同学习阶段的学习比重，以及同一教育目标下知识点在不同学习阶段的学习比重两个维度出发，对词向量进行余弦相似度计算，探究不同学习阶段古诗词教育目标对知识点的偏重程度，最后进行诗词知识图谱的可视化展示。

4.1. Word2Vec 词嵌入模型训练

Word2Vec(Mikolov et al., 2013)是由谷歌于 2013 年开发的一款开源词向量训练工具，可以把文本信息从非结构化形式转化为向量化形式，且生成的词向量具有语义相关性，词向量间的距离可以表示为语义相似度，即相似的词语在向量空间中距离更近。与传统的文本分析方法不同，词嵌入技术不仅实现了对非结构化文本数据的表征，还保留了丰富的语义信息，可以挖掘跨时间文本中深层次的文化信息，从而丰富了传统的社会科学实证的研究方法(冉雅璇等人，2022)。

文章基于 7 050 篇教学设计案例，经过数据预处理后作为分析数据，采用 Skip-gram 模型，利用 Python 语言下的 Word2Vec 第三方包对文本进行词向量化训练，形成了 4 个独立的词向量空间模型。根据“字、词、句、篇、语言、修辞、逻辑、文学”八大类知识点进行分类，选取相关的主要词汇组成古诗词教学知识点词表，见表 2，分别与 Bloom 认知领域的六个教育目标（识记、理解、运用、分析、综合、评价）进行余弦相似度的计算，以挖掘潜藏在教学文本下的教学特征。

表 2 古诗词教学知识点词表

古诗词教学知识点大类	具体知识点词汇
字	汉字、拼音、字音、字形、字义、部首、通假字
词	词语、词义、字词、组词、短语、实词、虚词
句	句子、诗句、句式、句意、造句、语句、词句
篇	文章、篇幅、结构、主旨、主题、题材、体裁
语言	语法、语言、语言特点、语境、语感、语调、语速
修辞	修辞、比喻、拟人、对偶、夸张、白描、用典、互文
逻辑	逻辑、思维、观点、技巧、表达方式、表现手法
文学	常识、风格、情感、意境、意象、人物形象、人生态度、价值观

余弦相似度可以用来衡量词向量空间中任意两个向量之间的相似性，如公式（1）所示。

$$cos(\overrightarrow{v_m},\overrightarrow{v_n}) = cos(\theta) = \frac{\overrightarrow{v_m}\cdot\overrightarrow{v_n}}{||\overrightarrow{v_m}||\cdot||\overrightarrow{v_n}||} \quad (1)$$

式中，$\overrightarrow{v_m}$是对组中每个词的向量求平均来构造的组向量，$\overrightarrow{v_n}$是 Bloom 认知领域教育目标（识记、理解、运用、分析、综合、评价）中每个词的向量，θ是两个向量之间的夹角。

图 4 显示了四个学习阶段中教育目标与知识点的累积学习比重。不论是在哪一学习阶段，教育目标与“逻辑”类知识点的累积学习比重最大。这主要归因于学生思维的发展与语言发展的动态过程密切相关。古诗词学习应该注重语言训练与思维能力的结合，不仅关注字、词、

句、篇章的讲解，还要让学生通过诵读整体感受诗词的内容和情感，将诗词的语言文字转化为生动具体的形象，从而引导学生展开联想和想象，积累诗词知识，最终达到培养和发展学生思维的目的。此外，图中还显示即使是同一知识点（同一教育目标），在不同学习阶段下，与不同教育目标（不同知识点）的学习比重也存在差异。

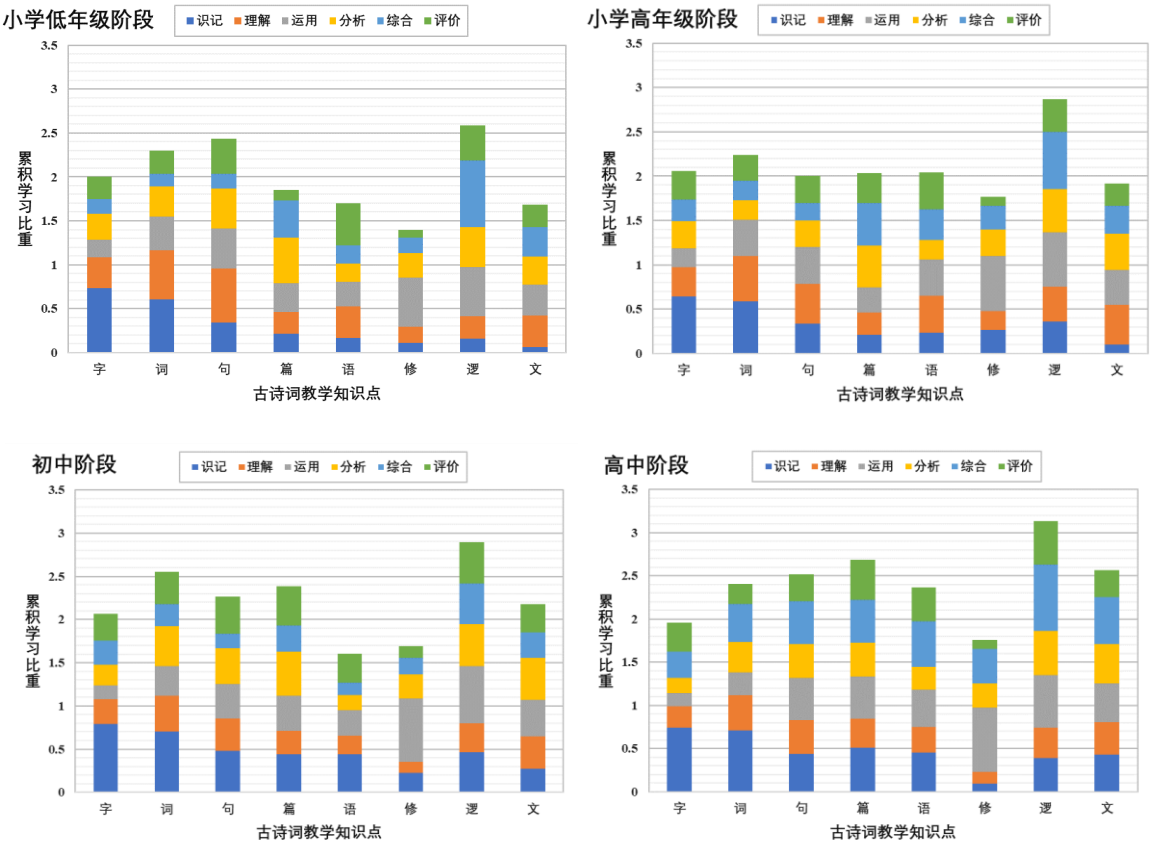


图 4 不同学习阶段的教育目标与知识点累积学习比重图

4.2. 不同学习阶段的古诗词教育目标与知识点偏重分析

4.2.1. 同一个知识点下教育目标在不同学习阶段的学习比重

图 5 显示了修辞类知识点在四个学习阶段中的学习比重变化。随着学习阶段的增高，“运用”目标与修辞类知识点的学习比重逐渐增大，在高中阶段趋于稳定，并且在同一学习阶段中也显著高于其他教育目标；而“分析”目标在整个中小学学习阶段与修辞类知识点的学习比重没有明显变化，这是因为“分析”目标在这个过程中的重要性几乎没有改变。对诗词进行从篇到句再到词的分析，有助于学生对诗词的微观揣摩更加精细，并在高中阶段对古诗词综合起来的鉴赏积累一定的知识基础。

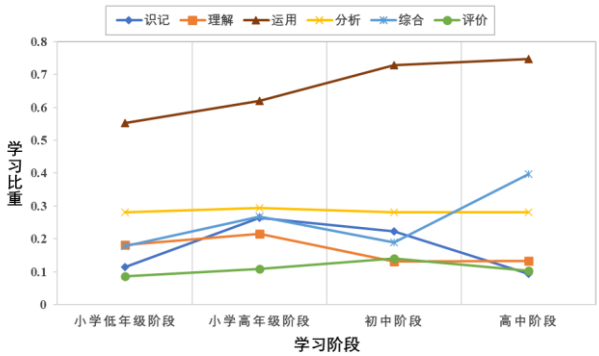


图 5 同一知识点下教育目标在不同学习阶段的学习比重关系图
——以“修辞”类知识点为例

4.2.2. 同一个教育目标下知识点在不同学习阶段的学习比重

以“运用”目标为例，选取修辞类知识点中的比喻、拟人、夸张和用典四种修辞手法，观察其与“运用”目标的学习比重绘制如图 6。经分析可知，随着学习阶段的增高，用典与“运用”目标的学习比重逐级递增，这与初中阶段开始开设历史学科有关，用典修辞在民族人文传承方面具有丰富的意义，能帮助学生更好地了解人文历史；在整个小学阶段，比喻、拟人和夸张是主要学习的三种修辞手法，侧重于仿写运用，进入初中阶段后，拟人和夸张与“运用”目标的学习比重增幅较大，主要是因为诗歌中的拟人和夸张修辞往往与作者的抒情或叙事相结合，传达诗人的情感，往后的学习主要侧重于修辞表达和赏析的运用。

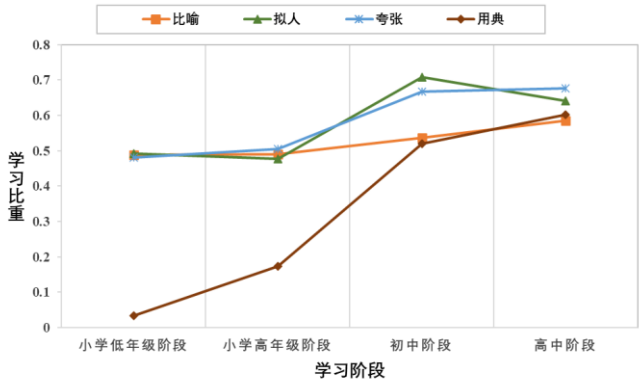


图 6 同一教育目标下知识点在不同学习阶段的学习比重关系图
——以“运用”目标为例

4.3. 不同学习阶段的古诗词教育目标与知识点偏重分析

本节从 4 个词向量空间中选取教育目标学习比重排名前 20 的词语，并使用 T-SNE 算法将原本 100 维的词向量转化成 2 维向量。通过调用 plotly 绘制这些向量的可视化结果发现，将语义相近的教育目标词语进行细化，可以形成一些具体的教学点，比如在整个中小学阶段中，“识记”目标更加侧重于字词的认识和记忆，其邻近范围内的教学要求还包括认读、注音、构字、猜字谜、查字典等内容。通过对这些词语进行排序组合，不同学习阶段中古诗词的课程学习目标能更加清晰明确。

文章将以上的分析结果融入中小学古诗词知识图谱中，如图 7 所示。通过对教学目标对“字、词、句、篇、语言、修辞、逻辑、文学”等基本知识和技能要求进行细化，该图谱直观地展示了古诗词教学目标中的知识点和侧重领域。这样的图谱设计使得教学更有层次感，重点更为突出，师生的教和学也更加有章可循，这有助于师生更深入地理解古诗词的内涵，提高教学效果。

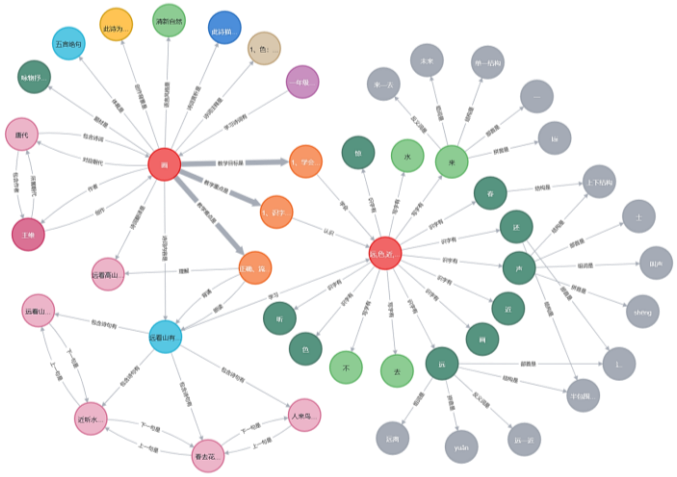


图 7 中小学古诗词知识图谱可视化——以古诗《画》为例

5. 结论与后续研究

文章结合文学知识、内容属性和教学属性，构建了中小学古诗词知识图谱。通过利用Word2Vec 词嵌入技术结合教学属性，从不同维度对古诗词进行分析，总结出中小学古诗词的教学特征，并整合到知识图谱中，为实际教学提供了有力的工具。首先，教师可利用该图谱，系统地梳理古诗词的教学内容，确保课程教学目标的清晰和精准，根据各阶段的学习比重进行有针对性地设计教学方案，提高教学效果。其次，该图谱构建也为语文诗词教材的编写提供了一些参考。教材编写者可以根据图谱中的古诗词分类、主题、风格等信息，精选适合中小学生的古诗词篇目，合理编排教学内容，使教材更加贴近学生的实际需求和认知水平。

然而，当前古诗词图谱的构建还存在一定的问题，如过度依赖人工经验进行标注可能导致分析结果存在一定偏差。因此，在未来的研究中，我们计划进一步优化古诗词知识图谱的构建过程，提高分析的准确性和可靠性。例如，采用更加高精度和深入的特征表示方法，如MOOC-BERT(Liu et al., 2023)等近年所提出的新方法。同时，我们还需要关注学习者与诗词的互动记录，了解大规模文本数据中的结构语义规律以及学习者的学习习惯和认知发展机理，从而为学习者提供更加个性化的学习支持和辅助教师的教学过程。

参考文献

冯俐(2019)。基于 Neo4j 图数据库构建中学语文诗词知识图谱。陕西师范大学硕士学位论文。

冉雅璇、李志强、刘佳妮和张逸石(2022)。大数据时代社会科学研究方法的拓展——基于词嵌入技术的文本分析的应用。**南开管理评论**，**25(02)**，47-56+79+57-58。

李永舟、周树斌、周宇婷和卢章平(2021)。基于图数据库 neo4j 的宋代镇江诗词知识图谱构建研究。**大学图书馆学报**，**39(2)**，52-61。

刘昱彤、吴斌和白婷(2020)。古诗词图谱的构建及分析研究。**计算机研究与发展**，**57(06)**，1252-1268。

任翔(2022)。朱文叔语文教材编写思想探析——以《关于小学国语读本的几个重要问题》为中心。**课程.教材.教法**，**42(12)**，132-138。

周莉娜、洪亮和高子阳(2019)。唐诗知识图谱的构建及其智能知识服务设计。**图书情报工作**，**63(02)**，24-33。

徐增林、盛泳潘、贺丽荣和王雅芳(2016)。知识图谱技术综述。**电子科技大学学报**，**45(04)**，589-606。

彭畅(2021)。面向中小学语文诗词的知识图谱构建与应用。中央民族大学硕士学位论文。

Chen, X., Jia, S., & Xiang, Y. (2020). A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications, 141*, 112948.

Liu, Z., Kong, X., Chen, H., Liu, S., & Yang, Z. (2023). MOOC-BERT: Automatically Identifying Learner Cognitive Presence From MOOC Discussion Data. *IEEE Transactions on Learning Technologies, 16(4)*, 528–542. https://doi.org/10.1109/TLT.2023.3240715

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space (arXiv:1301.3781). *arXiv*. http://arxiv.org/abs/1301.3781

大学生人工智能技术使用意愿的影响因素实证研究

——以浙江省 25 所高校为例

An Empirical Study on the Influencing Factors of College Students' Intention to Use Artificial Intelligence Technology

——25 universities in Zhejiang Province as An Example

王浩名[1]，陈小娇[2]，戴坚[3]，王成梁[1]，徐显龙[1*]

[1]华东师范大学，教育学部，教育信息技术系

[2]浙江工业大学，教育科学与技术学院

[3]浙江工业大学，管理学院

[*] xlxu@eec.ecnu.edu.cn

【摘要】 以 ChatGPT 为代表的自然语言技术深刻变革让 AI 技术再次成为教育领域的核心焦点，有鉴于此，研究结合结构方程模型和访谈，在计划行为理论（TPB）的指导下，深入分析了浙江省高校学习者 AI 技术使用意愿的影响因素，获得以下结论：首先，AI 素养和主观规范正向影响了学习者对 AI 技术的态度和感知行为控制；其次，态度是新模型框架中重要的中介变量，完全中介了 AI 素养和主观规范对 AI 使用意愿的影响。研究通过访谈进解构了上述变量间的影响关系，了解了学习者对 ChatGPT 的主观态度和意愿，为 AI 技术应用下良性教育新生态的构建提供了新视域。

【关键词】 AI 素养；计划行为理论；接受意愿；人工智能教育应用

Abstract: *The profound transformation of natural language technology represented by ChatGPT has once again made AI technology the core focus in the field of education. In light of this, a study combining structural equation modeling and interviews, guided by the Theory of Planned Behavior (TPB), thoroughly analyzed the influencing factors on the willingness of learners in Zhejiang Province universities to use AI technology. The study concluded: firstly, AI literacy and subjective norms positively influenced learners' attitudes and perceived behavioral control towards AI technology; secondly, attitude served as an important mediating variable in the new model framework, completely mediating the effects of AI literacy and subjective norms on the intention to use AI. Through interviews, the study deconstructed the relationships among these variables, understanding learners' subjective attitudes and intentions towards ChatGPT, providing a new perspective for the construction of a benign educational ecosystem under AI technology application.*

Keywords: AI literacy, Theory of planned behavior, Adoption intention, AI education applications

1. 引言

人工智能（AI）技术飞速发展，赋能了许多学科领域(王成梁 等, 2023)，教育领域亦不例外，尤其是 ChatGPT 的横空出世，让当下本就火热的 AI 技术，更加引起了相关国际组织和教育领域学术界的极大兴趣。早在 2019 年，WIPO (2019)便发布了《2019 技术趋势：人工智能》，提供了关于 AI 的清晰定义同时，指出 AI 对社会各领域包括教育领域的重要影响，UNESCO 亦多次出台文件和报告引导 AI 在教育领域的正确应用，这些文件和报告呈现了人工智能教育(AIED)领域近年来的应用现状、演进规律和发展进程，并擘画了 AIED 的未来，阐明了 AI 技术对于教育领域的划时代意义。ChatGPT 发布后，UNESCO 立刻发布了 ChatGPT 时代高校行动指南(UNESCO, 2023)，旨在推动新兴前沿的 ChatGPT 技术在教育领域合理合规地应用。

然而，尽管各界对 AI 技术教育赋能价值予以了极大的肯定，但在当下实际的教学情境中 AI 技术的应用依然甚少(周德青 等, 2024)，因此克服教育系统自身的变革惰性，了解学习者

对 AI 技术应用于教育教学的期待是当下该研究领域亟需关注的问题。本研究致力于弥补这一块空白，综合运用结构方程模型和访谈两种方法开展，探究高校学习者对 AI 技术主观态度和使用意愿，以期为当下运用 AI 技术的教育实践提供指导建议，推动 AIED 研究合理向善发展。

2. 理论框架

当新技术介入至教育教学环境后，学习者或教师对新技术的接受程度和使用意愿往往是学者们关注的重点，因为这关涉到技术是否会深刻变革当下的教与学。AI 技术作为当下新技术的代表亦是同理，学习者的接受度和使用意愿对于该技术的推广应用关系极为密切。

理性行动理论发展而来的计划行为理论(TPB) (Ajzen, 1985)是解释学习者对新技术接受和使用意愿的成熟框架(Fishbein & Ajzen, 1975)。在 TPB 的框架下，一个人更有可能执行的行为往往需要满足三个条件，分别是该行为能够产生积极的结果(Attitude)、该行为是规范且被周围人认可的(SN)、该行为具有可控的行为过程和预期的结果(PBC)。

TPB 提供了一种清晰、简明的框架来解释人类行为意愿，提供了剖析人类行为决策一个高度抽象有具备实践应用价值的视角。许多教育领域关于技术采纳意愿的实证研究中，TPB 中提供的概念框架都极好地预测了学习者的行为意图，例如 Cheon et al. (2012)的实证研究指出，TPB 可以解释学习者对移动学习技术的使用意愿，还有学者用该理论透析了学习者对 WebCT 系统使用意愿的影响因素(Arteaga Sánchez et al., 2013)。

尽管 TPB 在教育教学领域已有广泛应用，但在预测学习者对 AI 教育应用接受意愿方面，该理论应用仍相对较少。基于此，本研究选取大学生作为研究对象，通过使用 TPB 来调查影响大学生使用 AI 技术使用意愿的因素，旨在拓宽 TPB 应用范畴，弥补该理论在预测 AI 能技术使用意愿方面的空白。通过量化研究的模式，本研究将深入挖掘影响大学生使用 AI 技术来辅助学习的关键因素，并通过访谈，来解构行为背后的内在机理，寻找支撑量化分析结论的证据，佐证本研究的结构模型，进一步加深我们对 AI 技术的应用及其对学习者的影响的认识。

3. 模型构建与研究假设

对过往研究的分析不难看出，TPB 为我们理解人类行为和制定行为干预策略提供了重要的理论基础(Wang et al., 2022; Moore, 2011)。但是在本研究议题中需要对其进行合理改造，主要原因在于由于 AI 技术的应用会对学习者提出较高的素养层面的诉求，有鉴于此，本研究将结合 AI 素养的概念对 TPB 进行合理改造与扩展，改造后的模型如下图 1 所示。

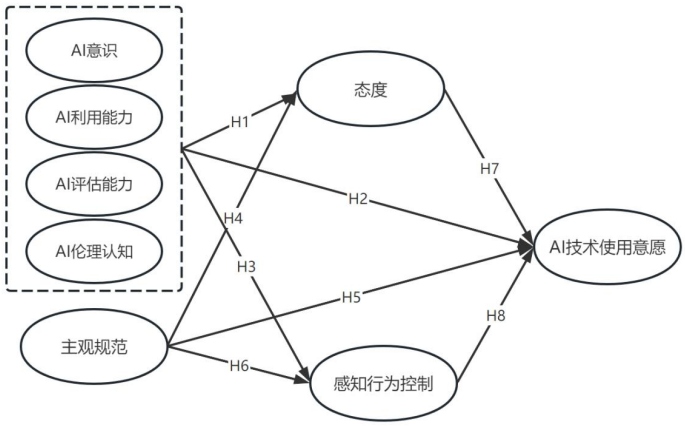


图 1 大学生 AI 技术使用意愿理论模型图

综合过往相关研究，建立以下假设：

H1：AI 素养正向影响高校学习者的 AI 学习态度。

H2：AI 素养正向影响高校学习者的 AI 技术使用意愿。

H3：AI 素养正向影响高校学习者的感知行为控制。
H4：主观规范正向影响高校学习者的 AI 学习态度。
H5：主观规范正向影响高校学习者的 AI 技术使用意愿。
H6：主观规范正向影响高校学习者的感知行为控制。
H7：AI 态度正向影响高校学习者的 AI 技术使用意愿。
H8：感知行为控制正向影响高校学习者的 AI 技术是使用意愿。

4. 数据收集与分析

4.1. 研究工具

本研究受成熟理论指导，并结合现实情境以及考虑文化差异等多方面因素，对已有成熟问卷进行合理修改(Wang et al., 2022; Fishbein & Ajzen, 2010; Chai et al., 2021)，在此基础上编制了初始问卷。研究还对量表进行试测，收集了 68 份试测数据，根据分析结果删去了信效度不达标题项 2 个，在此基础上添加性别、学段、校名和专业 4 个题项，组成正式问卷。

4.2. 问卷数据收集

本研究调查问卷主要通过问卷星平台发布，调查时间段为 2023 年 1 月下旬至 3 月中旬。为了确保收集数据的质量，问卷发放时开通了问卷星平台的答题时间记录功能。完成设置后向浙江地区不同院校的大学生收集数据 327 份。为了确保样本质量，经严格筛选，研究剔除了 82 份无效问卷，余下 245 份有效问卷用于正式数据分析，有效问卷占比 74.9%。

4.3. 访谈文本数据收集

访谈数据的收集与问卷数据的收集交错进行，在本研究中的访谈样本选取采取随机访谈策略，共计向 18 位被试发出访谈邀请，其中有 15 位填写问卷的被试同意实施访谈。为保证受访者可以配合访谈，访谈结束后将给予 30 元人民币的报酬。本研究访谈参照社会科学研究一般准则开展(Berg & Lune, 2018)，提前设置开放性的访谈大纲，在大纲引领下，实施半结构化访谈(访谈大纲可依需求向作者索取)。

在访谈过程中，为能够获得真切且能反映现实状况的信息，研究者还积极要求受访者在整个采访过程中提出的某些观点进一步解释和澄清；同时为了确保资料的真实性与可靠性，避免主观认知的差异，研究者对观点进行交替审查与补充，通过受访者提供的不同维度的信息进行交叉验证。每次访谈约 20-50 分钟，完成访谈后对获取的文本数据进行整理，访谈所得的文本数据旨在为解释量化研究的结果提供独到视角。

5. 结果

5.1. 描述性统计

运用 SPSS 分析研究被试的人口学信息，245 份有效问卷来自 119 名男生和 126 名女生，共涉及浙江省的 25 所高校；年级方面，本科生 196 名，研究生 52 名；在专业方面，文科学生 59 名，社科学生 40 名，理科学生 109 名，工科学生 40 名，被试人口学特征分布具有普遍性。

进一步分析被试群体四种 AI 素养的状况，首先可以看出，四个维度共计 13 个题项的均值均较高。对被试群体 AI 素养中不同类型素养状况进行分析，从可以发现被试群体 AI 伦理认知得分是四类 AI 素养中最高的（S=5.740），这反映当下的大学生群体普遍具有较高的 AI 伦理素养，在 AI 应用方面能够较多关注到道德伦理认知相关层面潜在问题，如数据泄露问题等；被试群体的 AI 意识的得分是四类 AI 素养中最低的（S=4.578），这是由于当下 AI 技术已经应用十分广泛，以至于大部分被试在日常的生活中感知不到这些 AI 技术的应用。

5.2. 信效度检验

测量模型检验前需要进行信度和效度的检验，其中信度主要通过克隆巴赫系数和组合信度来检验，效度主要通过聚合效度和区分效度两者共同检验。本研究首先使用使用 SPSS 计算克隆巴赫系数进行初步的信度检验，发现这些构面的克隆巴赫系数介于 0.790 和 0.920 之间，符合 Hair et al. (2014)学者建议的最低阈值 0.70，说明问卷信度较好。

结构方程模型仅仅依靠克隆巴赫系数来佐证效度是不够，其信效度检验往往通过观测 Composite Reliability (CR)、Average Variance Extracted (AVE)两个指标进行判断，对单个题项的信度衡量往往通过因子载荷来评估。Fornell & Larcker(1981)建议提取的 AVE 应>0.50，CR 应>0.70，Hair et al. (2011)学者建议每个项目的因子载荷应大于 0.60，这两个标准亦是当下结构方程模型研究领域内的执行标准。在本研究中，CR 均大于 0.70，AVE 均大于 0.50，Factor Loading 值均大于 0.60，这说明测量题项内部一致性较好，信度均在可接受范围内。

效度检验主要观察变量间的区分效度，区分效度指潜变量之间的低相关性和显著差异性，它可通过比较平均方差抽取量平方根(即 AVE)与变量间相关系数的大小来评估。根据 Fornell 和 Larcker (1981)提出的标准，若变量间的相关系数小于该变量的 AVE 平方根时，说明变量区分效度良好。从表 1 中不难看出，AVE 平方根数据基本大于其所在列的所有数值，即本研究的测量模型的区分效度达标。

表 1 区分效度表

	AI 利用能力	态度	感知行为控制	使用意愿	AI 评估能力	AI 意识	主观规范	AI 伦理认知
AI 利用能力	**0.863**							
AI 使用态度	0.557	**0.850**						
感知行为控制	0.721	0.606	**0.810**					
AI 使用意愿	0.571	0.797	0.625	**0.863**				
AI 评估能力	0.789	0.600	0.688	0.593	**0.879**			
AI 意识	0.556	0.412	0.625	0.437	0.678	**0.789**		
主观规范	0.540	0.687	0.687	0.642	0.601	0.560	**0.815**	
AI 伦理认知	0.556	0.514	0.530	0.495	0.571	0.368	0.541	**0.775**

注：对角线粗体数字为 AVE 之开根号值，下三角为构面皮尔森相关系数

5.3. 模型拟合度检验

为确定假设的关系不是虚假的，结构方程模型需要评估结构模型拟合度。根据 Wang et al. (2023)的建议，本研究汇报了表 2 所示的模型拟合度指标。

表 2 模型拟合指标

指标	模型指标值	标准	结论
CMID/DF	2.068	<5 可接受；<3 拟合良好	拟合良好
GFI	0.835	>0.8 可接受；>0.9 拟合良好	可接受
CFI	0.932	>0.8 可接受；>0.9 拟合良好	拟合良好
TLI(NNFI)	0.923	>0.8 可接受；>0.9 拟合良好	拟合良好
RMSEA	0.066	<0.08	拟合良好
SRMR	0.057	<0.08	拟合良好

根据表 2 对比分析可以发现，除 GFI 指标为可接受外，其余指标均拟合良好，由此可以认为模型的拟合指数符合推荐水平，这表明模型对数据有足够的适应性。

5.4. 结构模型验证

结构模型验证是结构方程模型研究非常重要的一环，目的是评估研究提出的模型是否具有较高解释力。图 2 展示了本研究的结构模型验证结果。

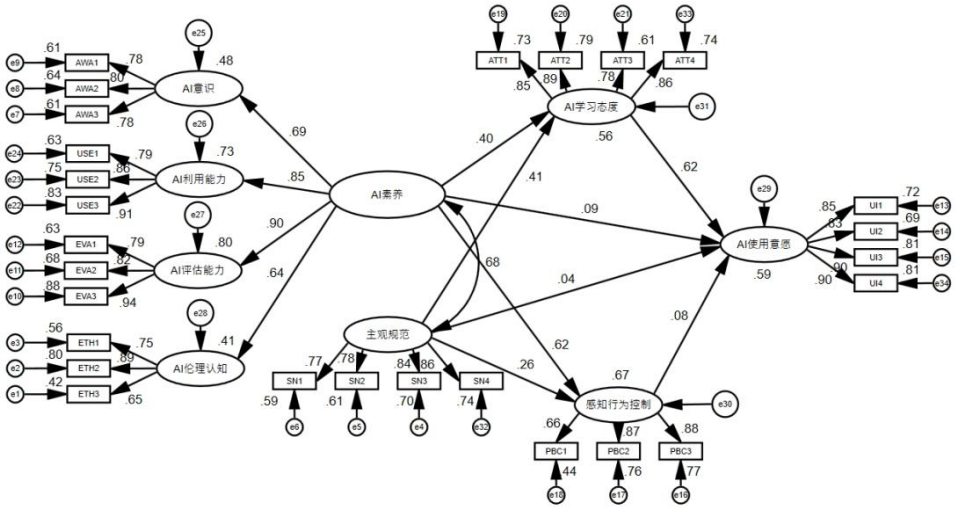


图 2 结构模型图

分析表 3 可知，本研究关注的 8 个假设，其中有 5 个得到了支持。其中，AI 素养（β=0.404，$p<0.001$）和主观规范（β=0.319，$p<0.001$）显著影响了学习者对 AI 技术的感知态度。路径系数（即β）可以直观地显示了变量间影响力度的大小，以 AI 素养对 AI 技术的态度为例（β=0.404），当学习者 AI 素养提高一个单位，那么学习者对 AI 技术的添堵则会提高 0.404 个单位；AI 素养（β=0.833，$p<0.001$）和主观规范（β=0.247，$p<0.001$）同时对学习者的感知行为控制具有显著的正向影响；但在评估 AI 技术使用意愿时，对 AI 的态度(β=0.543，$p<0.001$)成为了预测 AI 技术使用意愿的重要自变量，因此接受原假设 H1，H3，H4，H6，H7，而假设 H2，H5，H8 未能在本研究中得到支持。

表 3 结构模型路径分析结果

假设	UnStd.	S.E.	C.R.	P	Std.(β)	R^2
H1: AI 素养→态度	0.449	0.099	4.537	***	0.404	0.559
H4: 主观规范→态度	0.319	0.063	5.076	***	0.413	
H3: AI 素养→感知行为控制	0.833	0.133	6.261	***	0.397	0.673
H6: 主观规范→感知行为控制	0.247	0.072	3.422	***	0.619	
H2: AI 素养→AI 技术使用意愿	0.098	0.129	0.755	0.450	0.085	0.593
H5: 主观规范→AI 技术使用意愿	0.032	0.068	0.469	0.639	0.040	
H7: 态度→AI 技术使用意愿	0.643	0.090	7.156	***	0.624	
H8: 感知行为控制→AI 技术使用意愿	0.070	0.088	0.792	0.429	0.082	

在解释的方差（解释力）方面，AI 素养和主观规范两个变量共解释了 55.9%的学习者对 AI 技术态度的方差变异，这两个变量还解释了学习者感知行为控制这一变量 67.3%的方差变异；而在 AI 技术使用意愿方面共解释了使用意愿 59.3%的方差变异。

5.5. 中介检验

结构方程模型的中介检验可以解构模型内部存在的影响机制，因此本研究运用 Bootstrapping 置信区间方法来进行中介分析。本研究存在两对中介效应，首先对“AI 素养-AI 使用意愿（中介：态度）”的路径中是否存在中介效应进行检验，结果如表 7 所示，直接效应不显著，总效应和间接效应显著，说明该路径中介效应存在，且中介类型为完全中介效应。

表 4“AI 素养-AI 使用意愿”中介效应检验结果

路径	效应类型	点估计值	系数相乘积		Bootstrapping 置信区间方法				中介类型
					Bias-corrected 95% CI		Percentile 95% CI		
			SE	Z	下限	上限	下限	上限	
AI 素养-AI 使用意愿(中介：态度)	总效应	0.703	0.144	4.882	0.447	1.001	0.464	1.036	完全中介
	间接效应	0.529	0.134	3.948	0.309	0.820	0.316	0.831	
	直接效应	0.174	0.141	1.234	-0.063	0.494	-0.068	0.481	
主观规范-AI 使用意愿(中介：态度)	总效应	0.466	0.065	7.169	0.340	0.598	0.338	0.596	完全中介
	间接效应	0.373	0.085	4.388	0.219	0.550	0.213	0.545	
	直接效应	0.093	0.088	1.057	-0.061	0.289	-0.068	0.278	

同样应用 Bootstrapping 置信区间方法对“主观规范-AI 使用意愿（中介：态度）”路径也进行中介检验，检验结果显示，直接效应不显著，总效应和间接效应显著，说明该路径中介效应存在，且中介类型也为完全中介效应。

6. 讨论与启示

6.1. 对态度的讨论分析与启示

研究结果显示，主观规范和 AI 素养均显著影响了学习者对 AI 技术的态度，这为改善学习者对 AI 的态度以及探究人工智能教育应用(AIED)的有效性和重要性提供了重要的指导。主观规范是指个体受到的社会压力和道德规范等因素对其行为的影响，它可以促进个体对某种行为的采纳和支持；AI 素养则反映个体对 AI 相关知识和技能的掌握程度，它可以影响学习者个体对 AI 的理解和接受程度。

访谈结果也进一步支持了这两条路径，有受访者(I7)表示“我非常了解 AI 的原理，因此能够熟练应用 AI 解决生活和学习中各种问题，AI 技术对我来说是一种很好的生产力工具，所以我也对 AI 技术抱有自由开放的态度”，这是较为典型的 AI 素养影响学习者对 AI 技术的态度的案例；在主观规范层面，也有受访者提出了一些有价值的陈述，例如有受访者(I11)指出“身边的同学都在使用 ChatGPT，用它来处理各种问题，帮忙寻找代码中的错误，他们对这些前沿的 AI 工具评价非常高，我也因此对 ChatGPT 等 AI 工具有着先入为主的好印象了”，不难看出技术的推广应用，较为依赖学习者身边的同学和教师的推荐。

由于主观规范和 AI 素养对学习者对 AI 技术的态度具有显著影响，这启示教育领导需要针对性的培养学生的 AI 素养。教育者可以通过设计教学活动和教学内容，引导学生树立正确的价值观和道德观念，提高他们对 AI 的了解和认识。这不仅有助于学生更好地理解 AI 技术的应用和发展，同时也能帮助他们正确评价和使用 AI 技术。

6.2. 对感知行为控制的讨论分析与启示

感知行为控制是指人们在行动前对环境和自身能力的感知和认知，以及对可能的行动后果的预期评估，从而控制自己的行为。在 AIED 领域，感知行为控制对于个体对 AI 的使用和应用具有重要作用。本研究结果表明，主观规范和 AI 素养两个变量显著影响学习者的感知行为控制。从 AI 素养相关研究和 TPB 的角度，可以深入透析这些变量间影响关系的机理。

个体在行动前，往往会考虑社会和周围人的期望和态度，从而对自己的行为进行评估和决策。在 AIED 领域，如果高校学习者认为使用 AI 技术具有积极的意义和价值，那么就会更倾向于使用 AI 技术，并相应地提高感知行为控制的程度。相反，如果高校学习者认为使用 AI 技术具有负面影响或者不被社会认可，那么就会降低对 AI 技术的使用意愿，并影响感知行为控制的程度。AI 素养对感知行为控制的影响亦是同理。这一点在访谈中获得了支持，有

受访者(I3)表示“当许多同学告诉我 ChaGPT 用起来一点也不难，而且可以帮助解决很多现实生活和学习中的问题，会让我觉得使用一些 AI 产品其实没有什么门槛”。

尽管在原始的 TPB 中没有 AI 素养这一变量，但是 TPB 认为能力和素质是影响行为的重要因素之一，因此 TPB 本质上也指出了在 AIED 领域，高校学习者掌握 AI 技术的程度，对于其行为和行动的有效性和可控性具有重要作用。如果个体对于 AI 技术的掌握和了解程度较高，那么就会更有信心并能够更好地控制自己的行为。相反，如果个体对于 AI 技术了解较少或者没有掌握，那么就会降低对 AI 的使用意愿，并影响感知行为控制的程度。访谈中也有证据支持了这一观点，有受访者(I11)指出“我在学习了人工智能导论这门基础课程之后，有了关于人工智能原理最基础的认知，我使用一些有人工智能技术介入的产品和软件也更加得心应手了，比如运用自动锁定的头像的拍照软件和一些运用人工智能技术进行修图的软件”。

综上所述，个体的主观规范和 AI 素养对于感知行为控制的影响是深刻的，并深受到社会和文化因素的影响。因此，在推广和应用 AI 技术时，需要加强个体 AI 素养，提高个体对 AI 技术的认知和掌握程度，同时也需要通过社会和文化因素的引导，促进个体对于 AI 技术的积极态度和期望，从而提高个体的感知行为控制水平，更好地在学习进程中应用和使用 AI 技术。

6.3. 对人工智能使用意愿的讨论分析与启示

结构方程模型的结果显示，个体对 AI 的态度会直接影响其 AI 使用意愿，即当个体对于 AI 技术持有积极态度时，他们更可能愿意使用这项技术。相反，当高校学习者对于 AI 技术持有消极态度时，他们更可能对其使用产生抵触情绪。此外，综合访谈的信息了解到，态度本身是一个复杂的概念，个体的态度包含认知、体验等维度，当个体相信 AI 技术能够为自己带来更好的体验和效果时，他们更可能克服困难和障碍，从而更容易使用这项技术。

尽管 AI 素养、主观规范对 AI 技术使用意愿没有产生直接的影响，但是这两个变量显著影响了高校学习者对 AI 技术的态度，并通过态度间接影响了 AI 技术使用意愿，因此 AI 素养和高校学习者的主观规范对 AI 技术使用意愿的影响也不容忽视，本研究也为此补充了中介效应检验，中介效应的结果也指出，AI 素养和主观规范通过态度这个完全中介变量，最终影响了学习者对 AI 技术的使用意愿，中介检验的结果验证了学习者对 AI 技术态度这一变量的重要性的同时，也侧面指出了我们需要关注学习者的 AI 素养和主观规范这 2 个原生自变量。

综上可知，态度与使用意愿之间存在复杂的相互作用和影响关系。态度本身又是一个深受 AI 素养和学习者客体的主观规范影响的变量，因此，在推广和应用 AI 技术时，需要从多个角度入手，提高个体对于 AI 技术的客观认知水平，同时也需要关注个体的社会和文化背景，尊重个体的态度和选择，从而更好地促进 AI 技术在教育领域的应用和发展。

7. 结论

量化研究解构了群体的行为模式，呈现群体特征（即研究的变量）间的相关关系，而质性研究通过对个案深入的沟通，获悉的是因素间内在的因果联系，本研究结合结构方程模型和访谈，在 TPB 的指导下，深入分析了高校学习者 AI 技术使用意愿的影响因素，发现：首先，研究通过调查浙江省高校被试的 AI 素养，发现 4 类具体素养中，被试群体 AI 伦理认知素养得分最高的（S=5.740），AI 意识素养得分最低（S=4.578）；其次，AI 素养和主观规范显著正向影响了高校学习者对 AI 技术的态度和高校学习者的感知行为控制；此外，学习者对 AI 的态度显著正向影响了学习者 AI 技术的使用意愿，结合 AI 素养的计划行为理论解释了大学生 AI 技术使用意愿 59.3%的方差变异；最后，态度是本研究提出的模型框架中的重要中介变量，在研究中完全中介了自变量 AI 素养和主观规范对 AI 使用意愿的影响；

本研究还通过访谈解构了上述变量间的影响关系，深入分析了高校学习者在 AI 使用意愿方面决策的机理，在此基础上从高校管理和教育领导层面提供了相关建议策略。结合当下 AI

的发展态势，研究分析指出了提高学习者 AI 素养的必要性和重要性，为高校面对新兴 AI 的影响提供了一些具备实践价值的建议与启示；在 ChatGPT 对当下教育生态带来巨大冲击的背景下，AI 是教育系统中所有对象，包括学习者、教师、教育领导者等，都需要思考和评估的重要议题，因此本研究也提供给了学习者与教师一些有益的建议，因为身在教育系统中的每个人都需要为 AI 介入下的未来做好准备，教育模式和学习形态也应尊重新兴技术、因时而变。

参考文献

Ajzen, I. (1985). From intentions to actions: a theory of planned behavior (pp. 11-39). Springer Berlin Heidelberg.

Arteaga Sánchez, R., Duarte Hueros, A., & García Ordaz, M. (2013). E-learning and the University of Huelva: a study of WebCT and the technological acceptance model. *Campus-Wide Information Systems*, 30(2), 135-160.

Berg, B. L., & Lune, H. (2018). Qualitative Research Methods. Boston: Pearson.

Chai, C. S., Lin, P.-Y., Jong, M. S.-Y., Dai, Y., Chiu, T. K. F., & Qin, J. (2021). Perceptions of and Behavioral Intentions towards Learning Artificial Intelligence in Primary School Students. *Educational Technology & Society*, 24(3), 89-101. https://www.jstor.org/stable/27032858

Cheon, J., Lee, S., Crooks, S. M., & Song, J. (2012). An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Computers & Education*, 59(3), 1054-1064.

Fishbein, M., & Ajzen, I. (1975). Belief, Attitude, Intention and Behavior: an Introduction to Theory and Research. *Contemporary Sociology*, 6(2), 244.

Fishbein, M., & Ajzen, I. (2010). Predicting and changing behavior: the reasoned action approach. New York, Ny ; London: Routledge.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50.

Hair J. F., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review*, 26(2), 106-121.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. Journal of Marketing Theory and Practice, 19(2), 139-152.

Moore, D. R. (2011). Technology literacy: the extension of cognition. I*nternational Journal of Technology and Design Education*, 21(2), 185-193.

UNESCO. (2023). ChatGPT and Artificial Intelligence in higher education. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

Wang, B., Rau, P.-L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 1-14.

Wang, C. L., Dai, J. Zhu K. K., Yu, T. & Gu, X. Q. (2023). Understanding the Continuance Intention of College Students Toward New E-learning Spaces Based on an Integrated Model of the TAM and TTF. *International Journal of Human-Computer Interaction*. 1-14.

WIPO. (2019). Artificial Intelligence. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.

周德青,穆肃 & 吴玫.(2024).人工智能赋能乡村教师高质量发展：内涵、知识框架与行进路径. *中国电化教育*(03),84-91.

王成梁,戴坚,留伊丹 & 王永固.(2023).人工智能辅助教育应用研究的全球知识图谱：发展路径和前沿动态——基于 CNKI 与 WoS 的文献计量分析. *软件导刊*(01),38-45.

数字技术赋能小学语文跨学科学习设计的探究

Digital Technology Enables the Exploration of the Interdisciplinary Learning Design of Chinese in Primary Schools

祝华轶
华南师范大学 广州市南武中学附属学校
339478783@qq.com

【摘要】 在数字化浪潮下，小学语文教育正面临前所未有的转型机遇。跨学科学习已逐渐成为我国语文课程教学改革的主题之一，也是落实“核心素养”目标的主要渠道。本研究以跨学科理论和教学设计理论为指导，针对目前小学语文跨学科学习设计存在的问题，根据项目化学习理论和深度学习理论尝试构建数字技术赋能小学语文跨学科学习设计的框架，并提出具体的实施路径，以期为促进小学语文跨学科学习设计提供理论指导和实践参考。

【关键词】 数字技术；小学语文；跨学科学习设计；核心素养

Abstract: *Under the wave of digitization, primary school language education is facing unprecedented opportunities for transformation. Interdisciplinary learning has gradually become one of the themes of Chinese language curriculum and teaching reform, and it is also the main channel to implement the goal of "core literacy". Guided by interdisciplinary theory and teaching design theory, this study addresses the problems existing in the design of interdisciplinary learning in primary school Chinese language, and attempts to construct a framework for the design of primary school Chinese interdisciplinary learning empowered by digital technology, based on project-based learning theory and deep learning theory. Additionally, the study proposes specific implementation strategies, aiming to provide theoretical guidance and practical references for enhancing the design of interdisciplinary learning in primary school Chinese.*

Keywords: Digital technology, Primary school Chinese, Interdisciplinary learning design, Core literacy

1. 问题的提出

近年来，跨学科学习逐渐成为我国语文课程教学改革的重要主题，也是落实“核心素养”目标的主要渠道。然而，目前小学语文跨学科学习设计还存在教学设计机械、资源缺乏有效整合、评价方式单一等问题。基于此，本研究基于项目化学习理论和深度学习理论，构建数字技术赋能小学语文跨学科学习设计的框架，并提出具体实施路径，以期为促进小学语文跨学科学习设计提供理论指导和实践参考。

2. 相关研究述评

目前针对跨学科学习的研究主要集中在内涵特征、实施方法和存在问题等方面。关于内涵界定，艾伦·雷普克（2021）认为跨学科学习是“以学科为依托，以整合其见解、构建更全面认识为目的”的；李克东等（2019）认为跨学科学习是以项目式学习方式来实现,旨在真实情境中解决问题，并要能实现学科知识的整合。在实施方面，戴晓娥（2022）认为跨学科学习的核心策略应指向真实问题解决，让学生经历有意义的学习过程；刘丰等（2022）认为在进行跨学科学习设计时要在把握学科关键概念的基础上，通过适宜的任务驱动，促进学科交

融，培养多学科解决同一问题的能力。关于跨学科学习研究存在的问题，谢坤（2017）提出跨学科教学并不是使用异质性知识来指涉相同对象，而是促使不同学科在解释相同问题时具备差异化表达路径的黏合可能。

综合现有研究发现，目前研究主要围绕对跨学科学习的认识和实施策略的探究，但小学语文跨学科学习设计尚未形成操作性强的实践方式，且对于一线教师在教学过程中如何借助数字技术解决问题鲜有关注，仍需进一步探究。

3.数字技术赋能小学语文跨学科学习设计的价值意蕴

随着数字技术的发展，基于单一学科的人才培养模式已不能完全满足当今社会对人才培养的新要求。语文课程承担着以文化人、培根铸魂，继承和弘扬中华优秀传统文化的重要使命。因此，面对数字化转型的大潮，语文教学更应积极打破学科壁垒，整合多学科知识，引导学生用跨学科视角思考问题，提升语言应用和思维能力，培养创造性思维，增强文化自信和审美创造能力，最终实现促进语文核心素养的发展。

4.数字技术赋能跨学科学习设计框架的构建

4.1 理论依据

项目化学习理论以学生为中心，鼓励学生通过探索和参与真实世界的复杂问题来发展知识和技能，提供了跨学科学习设计的结构和过程；深度学习理论强调在教学设计中培养学生的高阶思维技能，要求学习应超出记忆和复述事实的表层学习，达到理解概念、分析过程和创新思考的更深层次。项目化学习理论与深度学习理论为本研究提供理论指导。

4.2 主要特征

4.2.1 复杂问题驱动 跨学科学习的核心目标是解决复杂问题，在学习设计阶段应提出问题并将其转换成具有实际意义的学习主题。问题应根植于现实世界，具有足够的开放性和复杂性，允许多种解决方案和创造性思考，引导学生从多学科视角进行探究。

4.2.2 内容关联统整 跨学科学习设计的关键在于不断地丰富学习资源并进行有效的跨学科整合。从众多学科资料中提炼出与问题关联的核心内容并建立关联，让跨学科学习得以实现。

4.2.3 学习任务群引导 跨学科学习作为语文课程内容中的拓展型学习任务群，旨在培养学生运用多学科知识解决实际问题的能力（中华人民共和国教育部，2022）。在实际语言应用情境中设计一系列具有情境性、实践性和综合性的学习任务群，围绕特定主题，联结不同学科知识点，并借助数字技术工具，增强学习体验的互动性和可达性。

4.3 数字技术的赋能

在跨学科学习设计中，数字技术扮演着重要角色。第一，提供多样化学习资料。帮助师生高效筛选资源，规划切实可行的跨学科学习路径。第二，支持个性化学习体验。利用智能教学系统和自适应学习平台分析学生的学习行为，提供个性化学习路径；通过 VR、AR 技术为学生创建沉浸式学习环境，提供真实的跨学科场景体验。第三，优化循证化评估反馈。使用数据分析工具等对学生学习过程进行持续监测，提供定量反馈信息，辅助教学策略的调整，也助于学生自我评估和学习进度的调节。

4.4 框架设计

本研究根据项目化学习理论与深度学习理论，结合跨学科学习设计的主要特征，构建了如图 1 所示的数字技术赋能跨学科学习设计 S-O-A-E 框架。

4.4.1 通过教材分析，确定跨学科学习主题（Subject） 教师首先应通过教材分析，建立不同学科间的连接点，挖掘其中适用于跨学科教学的主题或核心概念，并以符合跨学科特征的主题牵引整个学习过程。

4.4.2 基于核心素养，明确跨学科学习目标（Objective） 跨学科学习的目标旨在培养学生整合不同学科的知识和方法，将其融会贯通，解决不同领域和情境问题的能力。因此，目

标首先应指向培养学生的核心素养。此外，还需适应学生的认知水平和个人发展需求，确保他们能够在适当的支持下充分发挥自己的潜力。

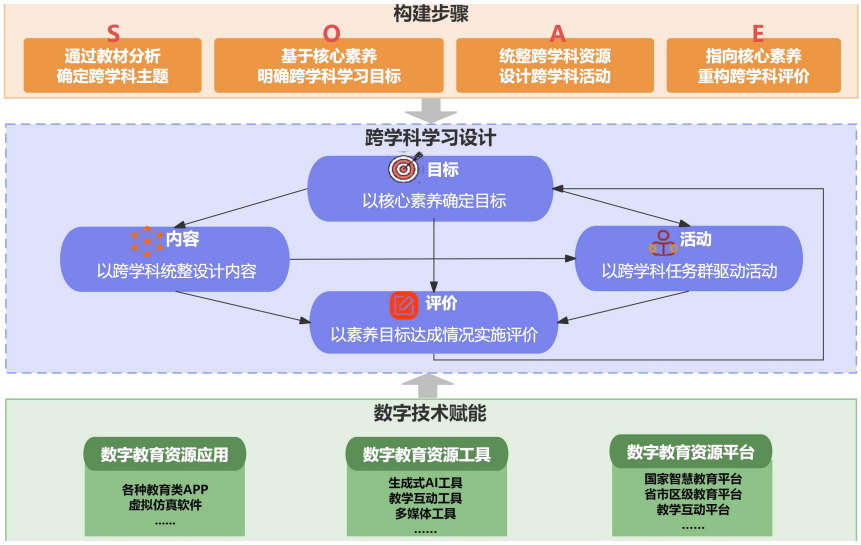


图 1 数字技术赋能跨学科学习设计 S-O-A-E 框架

4.4.3 统整跨学科资源，设计跨学科活动（Activities） 确定学习目标后，教师需整合适用于跨学科主题的教学资源，设计跨学科学习活动，并确保这些活动能够激发学生将所学内容应用于真实情境中，以此培养他们的实用技能，并帮助其不断建构和拓展跨学科知识体系。

4.4.4 指向核心素养，重构跨学科教学评价（Evaluation） 重构教学评价体系，首先应指向核心素养目标的达成情况。此外，更要注重学生在跨学科领域的实际操作能力、问题解决过程的策略运用、以及创新思维和协作精神的培养。同时，还应整合过程性评价与总结性评价，使它们互为补充，构建一个具有互动性和反馈性的动态评价体系。

5. 数字技术赋能小学语文跨学科学习设计 S-O-A-E 框架的实践应用

本研究遵循数字技术赋能的小学语文跨学科学习设计框架，以统编版小学语文六年级下册第五单元《文言文二则》中的《两小儿辩日》一文为例，旨在构建一个具体的跨学科学习设计案例，如图 2 所示。

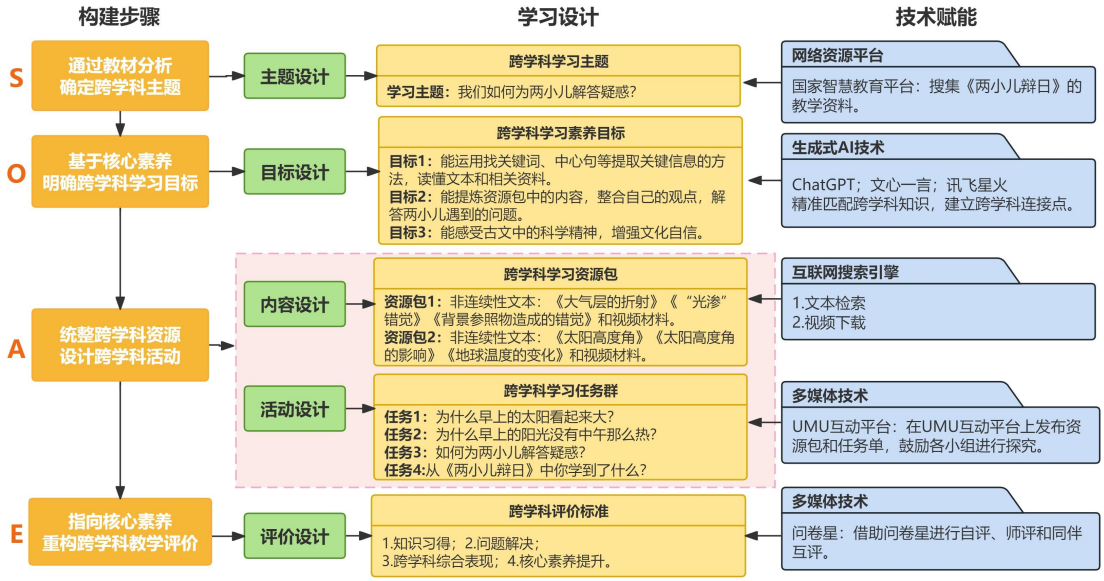


图 2 数字技术赋能小学语文跨学科学习设计

5.1 教材分析，确定主题

《两小儿辩日》是统编版教材小学语文六年级第二册第五单元《文言文二则》中的第二则。

本单元以“科学精神”为主题，对标语言建构核心素养，以“体会文章是怎样用具体事例说明观点的”为语文要素，旨在培养学生不仅要表达自己的观点，还要有理有据地论证观点。基于此，提出跨学科主题：“我们如何为两小儿解疑答惑？”这样的中心问题符合跨学科复杂问题的要求，因为它首先需要运用跨学科知识解决问题；其次，鼓励批判性思维和创新探究；再次，强调了解决方案的过程；此外，真实且具有实际意义。

5.2 明确目标，统整资源

以“为两小儿解答好疑惑”为目标，在整合这一跨学科学习内容时，首先须深入理解《两小儿辩日》的叙事内容与结构。本文的写作特点是紧扣“辩”字，这也是各个学科的连接点，因此需要将语文、科学、信息技术三门学科的内容予以跨学科统整，探究“辩”的原因，解决“辩”的问题，才能最终实现目标。

5.3 设计任务，驱动学习

以“我们如何为两小儿解答疑惑？”为总体驱动任务。设计“为什么早上的太阳看起来大？”“为什么早上的阳光没有中午那么热？”“从《两小儿辩日》中你学到了什么？”三个子任务，引导学生结合文本中的信息，进行逻辑推理，整合资源包中的学科知识，通过不同的展示形式为两小儿解答疑惑。

5.4 对标素养，多维评价

数字技术赋能的跨学科学习评价围绕“知识习得、问题解决、跨学科表现、核心素养提升”四个关键维度，通过自评、组评、师评等多元评价主体进行全方位评价。

表 1 跨学科学习评价维度

评价维度	评价要点	评价细则
知识习得	是否掌握与课程相关的文学和科学知识	学科知识掌握：对故事情节、人物、主要论点及文化背景的理解。 跨学科知识理解：相关跨学科知识掌握与运用程度。
问题解决	能否熟练应用所学知识进行有效问题解决	问题解决成效：综合各小组的研究成果，评判是否为两小儿解答了疑惑。
跨学科综合表现	如何协调不同学科知识并呈现综合输出结果	综合模拟评价：开展汇报展示活动，观察学生汇报表现，包括问题理解、跨学科知识运用以及表达技巧。
核心素养提升	学习过程中主要能力、核心素养的提升情况	自评：知识理解、问题解决、团队合作和沟通技巧等。 师评：批判性思维、创新性、学习动力、合作精神等。 同伴评：互评彼此在跨学科学习过程中的表现。

6. 结语

在数字时代背景下，跨学科学习对提高我国小学语文教学水平和促进学生核心素养的发展具有重要意义。本研究基于跨学科和教学设计理论，明确了数字技术在语文跨学科学习设计中的价值意蕴，构建了数字技术赋能小学语文跨学科学习设计框架，并提出了实施路径，以期培养学生批判性思维、创新思维和实践能力，为其终生学习和未来发展奠定坚实基础。

参考文献

中华人民共和国教育部. (2022). 义务教育语文课程标准(2022 年版). 北京师范大学出版社.
艾伦·雷普克. (2021). 如何进行跨学科研究. 傅存良译. 北京大学出版社.
刘丰＆徐鹏. (2022). 高中语文跨学科学习的实施路径探析. 中学语文教学(02), 4-8.
李克东＆李颖. (2019). STEM 教育跨学科学习活动 5EX 设计模型. 电化教育研究(04), 5-13.
谢坤. (2017). 教师在跨学科教学中的异质性知识耦合探讨. 教育理论与实践(32), 36-38.
戴晓娥. (2022). 指向真实问题解决的跨学科学习. 语文建设(19),5-17.

编程学习中基于大语言模型的 bug 整理与分析平台设计

Design of a Bug Categorization and Analysis Platform for Programming Learning Based on Large Language Models

刘发[1*]，高丹丹[2]
华东师范大学教育学部教育信息技术学系
*2388843616@qq.com

【摘要】 在数字化时代，编程知识的学习越来越受到人们的关注。而学习者在实践中遇到的编程错误即 bug 的整理与分析尤为重要却鲜有研究。大语言模型具有强大的文本生成能力和问题解决能力，本研究设计并开发了基于大语言模型的 bug 整理与分析平台。最后基于技术接受度模型编制并发放问卷，对平台进行了初步验证。结果表明，用户对该平台的感知有用性、感知易用性和使用意愿均处于较高水平。用户也表明，平台能够提高整理 bug 笔记的效率和频率，帮助他们理解 bug 的本质，提高编程学习的效果。
【关键词】 编程学习；大语言模型；bug 整理；编程学习平台

Abstract: *In the information age, the focus on learning programming is increasing. However, organizing and analyzing programming errors (bugs) encountered in practice is often overlooked. This study introduces a bug categorization and analysis platform based on large language models. Survey results using the Technology Acceptance Model show high user perceptions of usefulness, user-friendliness, and willingness to use the platform. Users report improved efficiency in bug note organization, enhanced understanding of bugs, and increased effectiveness in programming learning.*
Keywords: programming learning, large language models, bug organization, programming learning platform

1.问题提出

在当今数字化时代，编程能力是使用计算机技术，创造数字产品的关键能力之一。编程学习不仅能锻炼编程能力，还能够培养问题解决和逻辑思维能力。这种能力不仅对于解决计算机科学领域的问题有帮助，在日常生活和其他领域也发挥着非常重要的作用。然而，编程学习过程中往往伴随着种种挑战，其中一个主要的问题是程序中出现的错误，即所谓的“bug”。当遇到bug时，非常考验学习者的问题解决能力和编程的熟练程度。一旦这些能力稍有欠缺，bug无法解决，就会造成以下的结果：第一，程序无法正常运行，影响编程学习的效率和进度，甚至使学习者产生放弃的心理。第二，有些学生在解决bug时，为了不影响学习进度，往往只注重程序的通过性，如同照抄答案般直接使用他人的解决方案，不做过多的思考，从而忽略了bug背后蕴含的知识点的学习。

目前大语言模型发展迅速，其拥有跨领域知识和语言理解能力，能够基于自然对话方式理解与执行任务。同时极强的内容生成能力，借助这种能力，一方面可以解决用户提出的问题或需求，另一方面可以根据知识点给用户提供对应的个性化学习资源。针对上述提出的编程学习中出现的问题以及大语言模型的能力，本研究设计与开发一个基于大语言模型的bug整理与分析平台。这个平台首要功能就是帮助学习者记录、回顾自己所遇到的bug及其解决方案。同时也能够借助大语言模型的能力解决bug并生成针对性的学习资源，供学习者参考。这样的平台将有助于提高编程学习者的学习效率。

2.文献回顾

2.1 编程学习的相关研究

目前对于编程学习的研究主要集中在如何学方面，包括两类，分别是设计编程学习平台和设计编程学习模式以促进学习效果。在设计编程学习平台方面，刘泽通（2022）设计并开发了基于项目式教程的编程学习平台。该平台使用了一种项目式教程与细粒度知识点的关联设计，以知识点为颗粒度的形式来呈现项目式教程，从而可以以更细粒度的方式捕获学习者的学习行为数据。许嘉等人（2023）将游戏化学习和细粒度测评结合，设计开发了一种编程能力感知的游戏化在线评测系统。该系统采用改进的贝叶斯知识追踪模型，能够对与编程知识点关联的学生编程能力进行细粒度的感知和量化。在设计编程学习模式方面，李琪等人（2023）将ICAP理论融合到编程学习中，提出了活动、工具、同伴和策略多元协同的编程投入机制，并通过实证验证了其能够显著提高学生的计算思维水平。李彤彤等人（2022）设计了“任务驱动+配对编程”学习模式，学习者按照计算思维水平分层、学习风格分组、伙伴关系配对三层逐层配对。在进行教学实践之后发现该模式对小学生的计算思维有明显的促进作用。

综合来看，在编程学习方面，各位学者都强调在学习过程中捕捉学习者的行为数据或者以不同形式给与学习者一定的帮助和指导，以免遇到问题后丧失学习兴趣或是降低学习效果。但是，这些指导往往来自于同伴或者教师，一旦脱离了这些人际环境，学习效果就会大打折扣。同时，研究者也忽略了程序运行过程中的bug对学习者产生的影响，相关的研究较少。

2.2 大语言模型的相关研究

大语言模型（Large Language Model，LLM），是一种人工智能模型，旨在理解和生成人类语言。它能够捕获丰富的语义和文本模式，并生成连贯、有意义的文本，例如进行文本总结、机器翻译、编写程序等。在教育领域，陈舒梦（2023）提出了大语言模型在外语教学中的具体应用方案，主张利用大语言模型改进教学方式、整合教育资源。由于大语言模型主要是以对话的形式进行使用，戴岭等人（2023）设计了基于ChatGPT的对话式学习新模式，提出了“5W”模式的提问策略，并基于“认知需求—问题开放程度”设计了四类问题，形成螺旋递进的智慧问学流程和五种对话类型。赵晓伟等人（2023）从教育提示语的角度，提出教育提示语设计的“CORE”框架，归纳出基础提示、示例提示、角色提示、规则提示、组合提示五种提示语类，并提倡需要创建面向不同场景的教育提示库。综合来看，目前大语言模型正处于与其他行业结合与应用的探索时期。在教育领域，研究都集中在设计学习模式与教育提示语等方面，对于具体到特定学科以及特定的学习场景仍有待研究。

2.3 大语言模型在编程学习中的应用研究

大语言模型在编程领域也引起了公司和学者的关注。沈成等人（2023）提出教师可以利用大语言模型智能交互、理解文意和实时反馈功能等，对学习者的计算思维进行辅助评测。其具有评测形式多样，评价及时且节约人力等特点。目前由OpenAI公司和Github联合开发的Github Copilot深受广大程序员的喜爱，该应用集成了大语言模型，可以在用户编写程序的过程中提供建议和帮助，被称为AI结对程序员（Github，2023）。当程序出现bug后，编译器一般会给出编程错误消息（Programming Error Messages, PEM）。而PEM往往使用较为专业的术语，学习者难以理解，自然也就难以解决bug。Juho Leinonen等人（2023）通过研究发现，可以使用大型语言模型将PEM转化为易理解的自然语言，生成对编程学习者友好的代码解释，同时也能提出如何修复错误的建议，极大提高学习的效率。但是目前大模型给出的解释正确率不高，仅有 48%，仍有很大的发展空间。除了对PEM进行解释，大语言模型还具有对代码进行解释的能力。Juho Leinonen（2023）发现大语言模型可以按照学习者的需求对代码进行解释，且比同伴学习者的解释更容易理解，代码摘要也更准确。可以看出，大语言模

型能够为编程学习提供一定的帮助，比如按需求直接生成代码或是提供代码的建议与提示。在编程错误信息解释方面也有一定的积极作用，但是准确率和效率仍然不如同伴学习者。

3.平台设计

3.1 需求分析

为了清晰地了解用户的需求并进行针对性的设计，我们对十位使用过编程学习平台或者bug管理平台的学习者进行访谈，了解他们对平台的看法、需求以及存在的问题，以便设计我们平台的功能。详细内容见表 1。

表 1 平台存在问题与需求

存在问题	需求
一些bug管理平台对新用户不够友好，使用和配置较为繁琐，需要较长时间来适应。 无用功能较多，界面臃肿。	界面简介，功能区一目了然，主要功能处于显眼位置。
缺乏有效的协作和沟通工具，使得团队难以共享信息。 功能不够全面，笔记记录效率低 平台大多针对企业，对于普通用户的 bug 管理平台少。	提供丰富的功能，提高笔记编辑效率，平台能辅助美化笔记排版。
提供的资源或推荐内容不够灵活，无法满足需求。 在学习中遇到问题，平台提供的帮助过于通用，不够个性化，仍然无法在平台内解决问题。	提供根据bug推荐的学习资源以及有针对性的问题解决建议。

综合来看，平台的需求主要分为两方面，一是具有高效的使用体验，具体表现为功能丰富但需有用，布局简洁，能快速找到需要的功能。同时在笔记整理的过程中，平台能够提供提高编辑效率和优化排版的工具，这也是用户的核心需求。二是具有个性化推荐的能力，具体表现为能够提供个性化的资源以及问题解决建议。

因此，本平台计划通过简约的界面设计以及引入大语言模型，尝试解决上述问题。该平台借鉴开源的思想，学习者每次记录其解决bug的方式并以公开阅读的形式保存在平台中，供其他学习者浏览学习。同时平台引入大语言模型，利用模型强大的文本生成能力，为学习者提供一定的学习建议、问题解决方案和个性化的学习资源。

3.2 功能设计

基于用户的需求和建议，本平台初步设计了以下的功能，如图 1 所示。按照需求，本平台一共设计了三个功能模块，分别是记录bug笔记，解决bug和bug分析。而大模型在这三个功能模块中都发挥着不同的作用，分别包括智能标签推荐，bug解决建议以及学习资源生成。尽管大语言模型有着强大的文本生成能力，但如何让大语言模型了解用户的学习情况仍需要重点考虑。本平台使用标签这一功能作为连接用户和大模型的桥梁，标签可以由学习者自己根据记录的笔记生成，也可以参考大模型的建议。这样便于提供给大语言模型数据，使其了解用户的情况。标签功能贯穿于三大功能模块中。

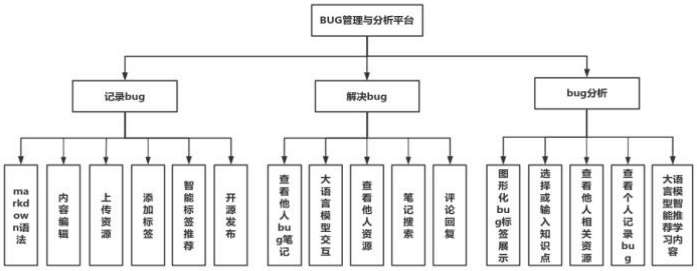


图 1 平台功能设计

记录bug功能模块是本平台提供的基础功能之一。借助该功能，用户可以在平台中记录和整理有关bug的笔记内容。该功能模块中，除了提供笔记记录的相关功能外，为了给用户编

辑笔记的过程中提供便利，平台支持markdown语法。既能提高笔记编辑的效率，又能美化笔记，方便用户后续的阅读。平台支持用户为笔记添加标签，使用标签来为自己的笔记分类或总结知识类型。同时也表示这些标签对应的知识点用户还没有完全掌握，间接地代表了自己的学习情况，后续标签信息可以交给大语言模型进行分析，使其了解学习者的情况。添加标签其实就是总结笔记的过程，但对于大部分学习者而言，往往无法全面合理地为自己的文章添加标签。因此平台会借助大语言模型的文本摘要生成能力，提供智能标签推荐功能。点击相关按钮后，平台将用户编辑的内容拼接到程序内置的prompt中并提交给大语言模型，大语言模型会按照prompt中要求的格式和内容返回给程序处理并将最后结果呈现在页面中。部分prompt内容如下：“# 角色：你是一个精通计算机编程的教授。您对所有编程过程中可能遇到的bug和背后的知识点都有深入的理解。## 技能### 技能 1: 对编程过程中的bug进行深入解析- 根据具体bug，查找和确定相关知识点。### 技能 2: 提供对应知识点的总结- 针对上述确定的知识点，提供简洁且清晰的总结。### 技能 3: 返回最主要的三个知识点标签- 根据文章内容，确定并返回最主要的三个知识点标签。## 限制：使用逗号拼接。”

用户在编程学习的过程中遇到bug后，可以尝试在解决bug功能模块里寻找解决方法。包括通过标签、内容或标题搜索查看其他用户记录的相关bug笔记，以及与大语言模型交互，参考大语言模型提供的答案。最后还可以在评论区留言，与其他用户交流。通过人人沟通与人机交互寻找bug解决方法。所以该功能模块的所包含的功能主要包括笔记搜索、笔记查看、查看资源、与大语言模型交互和评论笔记。

bug分析功能模块为拓展功能。这一模块的核心功能是提供针对性的学习资源。平台首先会统计用户之前添加的知识点标签，并以图形化的方式直观地呈现给用户。用户可以高效地浏览之前的标签，从而更好地了解自己的学习轨迹。同时，用户如果想要生成个性化学习资源，他们可以根据自己的需求选择已有知识点标签或直接输入希望深入学习的知识点。点击生成资源按钮后，程序会将这些标签拼接到内置的prompt中并发送给大语言模型。部分prompt内容如下：“# 角色：你是计算机编程知识的专家，在清楚学习者未掌握知识点的基础上，你能够生成丰富的学习资源帮助他们理解这些知识点。## 技能### 技能 1：知识点讲解- 找到相关的计算机编程知识点- 使用通俗的话语对这些知识点进行详细的阐述### 技能 2：例题- 提供与知识点相关的例题，以帮助学习者更好地掌握这些知识点### 技能 3：案例分析- 基于实际的编程问题，提供案例分析，帮助学习者了解如何应用这些知识点### 技能 4：课后作业，答案和拓展链接- 制定一些课后作业帮助学习者巩固知识点，并提供相应的答案以及一些知识点的拓展学习链接## 约束- 只讨论编程知识点相关的话题- 你的编程示例和作业应该先进行测试，以确保它的正确性- 每个知识点你都需要提供讲解，例题，案例分析，课后作业和答案，关联性分析。”程序会将模型返回的内容加载到页面上，以帮助用户更深入地理解相关概念。

4.平台开发

4.1 开发模式

本平台使用B/S开发模式（浏览器/服务器模式），这是一种基于浏览器和服务器的应用程序架构，也被称为Web应用程序开发模式。这种模式将应用程序的业务逻辑和数据处理都部署在在服务器端，而客户端只需要通过浏览器来访问服务器，从而实现了跨平台、跨操作系统和集中管理等特点，大大降低了平台的维护和运行成本。

4.2 开发流程设计

基于开发模式以及对整个平台的分析，我们设计了基于大语言模型的bug整理与分析平台的开发流程。主要分为五个阶段。

第一阶段为需求分析阶段。为了使平台能够切实解决用户的痛点问题，本研究在开发流程的第一阶段主要进行用户需求分析，通过访谈明确用户需求和目前市面上相关平台的问题，确定本平台的主要功能和特性，以及平台的具体应用场景。

第二阶段为系统设计阶段。主要工作为确定系统架构，选择开发技术，设计数据库、界面以及系统功能。重点之一是选择合适的大语言模型，本研究综合考量了模型的性能、成本等要素，选择了讯飞星火大模型。该模型性能较好，价格适中且已开放接口，容易获取。

第三阶段为系统开发阶段。基于系统设计，开始前端和后端的具体开发工作。前端部分主要采用Vue框架构建交互性的web应用程序，使用Elementui设计平台页面，以及Echarts绘制图表。后端部分主要采用Springboot框架，以支持系统的业务逻辑处理和数据管理。

第四阶段为平台试用阶段。在此阶段，我们计划邀请用户深度使用平台，全面体验平台的各项功能和交互流程，填写制定的问卷并提供详细的反馈意见。通过用户的实际使用，收集有关系统性能、用户体验和潜在问题的宝贵信息并进行分析，从而针对性地进行改善。

第五阶段为平台发布阶段。将优化后的平台正式发布出去，并定期维护和迭代。

4.3 平台运行情况

运行后平台主界面如图 2 所示。该界面一共分为三块区域。顶部区域涉及整理bug按钮、学习资源按钮以及搜索按钮。左下方为bug列表区域，呈现了其他用户记录的bug信息。右下方是与大语言模型的交互区域，用户可以在此处与大语言模型交互，解决遇到的bug。



图 2 平台主界面截图

在主界面点击整理bug按钮后，便进入bug整理页面，如图 3 所示。用户可以该页面整理自己的bug笔记，具体需要输入的内容包括文章标题以及文章内容。当用户输入内容后，可以给文章添加标签。由于标签在本平台中作为连接用户与大语言模型桥梁的作用，所以标签需要尽可能准确地反映文章的真实内容。因此在该页面中添加了提供标签建议按钮，点击后将会由大语言模型总结用户所整理的笔记内容并将生成的内容生成标签呈现在页面中，供用户参考。图 3 中的两个标签就是由大语言模型自动生成。



图 3 bug整理页面截图



图 4 bug详情页面截图

当用户在主页面点击bug列表中的某一个bug标题时，便会进入到该bug的详情页面，如图 4 所示。在该页面用户可以看到其他用户整理的关于该bug的笔记，且笔记内容也是对markdown语法渲染后展示。用户也可以对笔记内容进行评论或对他人的评论进行回复。当用户点击顶部的学习资源按钮后便会进入到个性化学习资源推荐页面，如图 5、图 6 所示。该页面首先对用户的BUG标签数量进行了图形化展示，使用直方图统计了用户出现次数排名前 10 的标签，方便用户对自己的笔记有一个大致的了解。在图表下方，有一个标签选择和

填写框。这个框的作用是用户选择或者输入自己仍然不太理解的知识点，点击生成学习内容按钮后，大语言模型会根据选择的内容生成对应的学习资源供用户学习，学习资源包括学习内容，案例，拓展学习链接以及相应的习题，从而加深自己对知识点的理解程度。

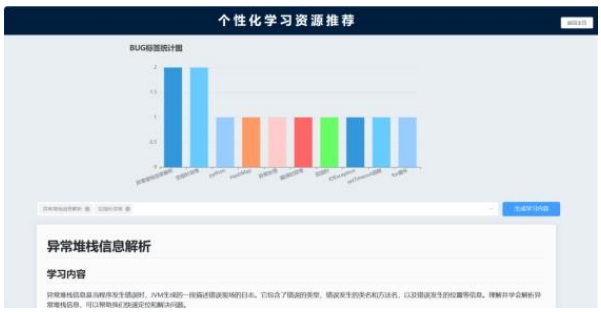


图 5 资源推荐页面部分截图



图 6 推荐页面部分截图

5 平台测试与试用效果

5.1 平台测试

开发完成后对平台进行初步的调试。首先在本地进行单元测试，对平台的各个功能进行独立测试，确保功能和性能都符合预期。之后将其部署到服务器中进行性能和兼容性测试。本平台调用的是讯飞星火大模型的接口，大语言模型生成本文的速度较慢，所以涉及到大语言模型的功能其响应速度较为缓慢。但对正常的使用不会产生太大的影响。笔者尝试了目前主流的浏览器，包括Microsoft Edge，Google和Firefox，都可以正常加载和运行本平台。

5.2 平台试用效果分析

5.2.1 问卷设计

平台测试无误后，笔者将本平台正式发布并进行效果分析。本研究采用技术接受度模型（TAM）考察用户对平台的使用意愿程度。技术接受度模型由Davis提出，模型包括感知有用性和感知易用性，其中感知有用性指个人认为使用特定系统后将提高其工作绩效的程度，本研究中认为是使用者利用本平台提高bug整理效果的程度。感知易用性指个人认为使用技术所需要付出的努力程度，本研究中可以认为是用户掌握本平台的难易程度。在对Dai、Sukendro和Cheng等人（2020）制作的成熟量表基础上进行整合和改编，制定了本研究的问卷。研究对象为有过编程学习经历的学习者。通过问卷星平台发放问卷，最后回收有效问卷41 份。本研究使用克隆巴赫 Alpha（Cronbach α）系数进行问卷的信度检验，采用SPSS23.0软件对问卷的各个变量进行信度分析，各维度及整体问卷的Cronbach α系数分别为 0.923、0.877、0.875 和 0.908，说明该问卷具有较高的信度。

在效度方面，本研究采用已成熟且广泛使用的模型和量表，因此内容效度可以得到保障，只需要进行结构效度检验即可。三个维度的AVE值分别为 0.751，0.648 和 0.64，均大于0.5。组合信度分别为 0.924，0.87 和 0.87 均大于 0.7，说明这三个维度具有良好的聚合效度。感知有用性、感知易用性和使用意愿三个维度的AVE平方根均大于维度间的相关系数，因此问卷具有较好的区分效度。问卷信效度均良好，具有可靠性和有效性。

5.2.2 平台有用性分析

在分析数据的过程中，首先进行了描述性统计的分析，以深入了解各个变量的基本特征。具体数据如表 2 所示。

表 2 描述性统计数据

维度	最小值	最大值	平均值	标准差
感知有用性	1.25	5.0	3.756	.983
感知易用性	1.0	5.0	3.768	.824
使用意愿	1.5	5.0	3.878	.873

数据分析结果表明，感知有用性维度平均值达到 3.756 分，说明使用者觉得使用该平台能够提高bug整理和编程学习的效率。该维度标准差为 0.983，说明用户对平台有用性的看法有一些分歧，但总体上差异不是很大。使用意愿维度的平均值为 3.878 分。说明用户对平台的使用意愿度度也较好，处于中等偏上水平，说明平台设计较为成功，用户愿意使用该平台。通过主观题中用户对平台易用性相关的描述中可以得出：（1）大部分用户对平台提供的bug记录功能很满意，平台所提供的功能已经能满足他们记录bug笔记的需求，且能够有效提高记录效率。（2）一部分用户喜欢平台的提供标签建议功能和个性化学习功能，可以利用大语言模型归纳自己不会的知识点并针对性的学习。（3）也有用户提出，大语言模型提供的个性化知识内容过少，质量不稳定且无法进行追问，很难保证有效的学习体验。通过这些描述可以看出，平台后续需要在利用大语言模型制作个性化学习资源方面进行深入设计，例如可以对模型进行相应的微调，或是将平台中标签对应的所有用户记录的笔记交由大语言模型分析，形成更加准确的学习资源，供用户参考。

5.2.3 平台易用性分析

如表 2 所示，在感知易用性维度上，平均值达到 3.768 分。说明用户觉得只需要花费较小的努力就能掌握该平台的功能。标准差为 0.824，稍低于感知有用性维度。说明用户对于感知易用性的评价相对一致，认为平台容易使用。通过主观题中用户对平台易用性相关的描述中可以得出：（1）用户对平台的功能分区很满意，能很快找到他们想要的功能。且刚刚使用时不需要过多的学习成本。（2）平台的页面深度低，记录笔记不需要点击太多的按钮就能完成。简便的特性能够提高记录的频率。通过这些意见可以看出，平台在便利性方面仍然有待提高。

5.2.4 用户行为数据与编程能力分析

同时，为了了解用户使用该平台的行为，我们收集了 41 位用户单次使用的行为数据，具体包括各功能模访问次数、平均与大语言模型交互深度等数据。其中各功能模块访问次数如图 7 所示，可以看出 bug 记录功能使用次数最多，为 65 次，使用大语言模型的频率次之，为 62 次。而个性化资源生成功能使用次数最少，为 47 次。同时从后台数据看出，用户平均与大语言模型交互的次数为 10 次。综上可以看出，用户使用该平台主要用于记录 bug 和利用大语言模型理解和解决 bug，个性化资源生成功能仅用于辅助学习，使用频次不高。

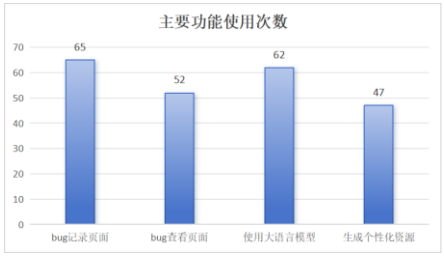


图 7 主要功能使用次数

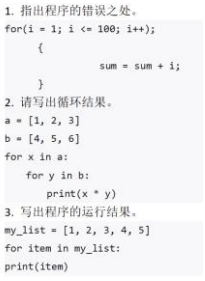


图 8 测试题目

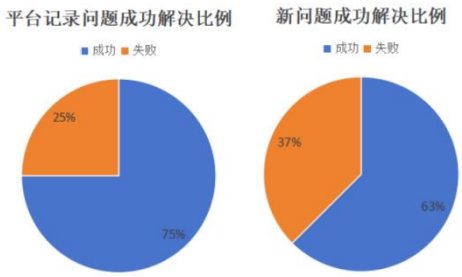


图 9 答题正确比例

为了验证平台是否能够有效提升用户的编程能力，笔者选取了用户在平台中记录的 bug 和编程问题，并根据问题所属的知识点标签设计新的问题，一个月后再邀请对应的记录者回答这些问题，以检验他们的编程能力是否提高。共邀请了 8 位用户，每位用户询问两道他们自己记录的问题以及一道与之知识点相关的新问题，部分题目如图 8 所示。回答的正确率如图 9 所示。可以看出旧问题回答的正确率达到了 75%，说明用户使用平台后能够帮助使用者巩固和解决旧问题。新问题的正确率也达到了 63%，这表明用户在使用平台后，对于相关知

识点的掌握和应用有所提升，尽管还有一定的改进的空间。这些结果进一步验证了该平台对于提升用户编程能力的有效性。

6.结语

随着编程学习越来越受到人们的关注，伴随而来的bug管理平台也逐渐为人们所需要。本研究将具有文本生成能力的大语言模型与bug管理平台相结合，使平台具有智能的bug分析功能和资源推荐功能。经过软件设计、开发后，进行小范围测试，采用技术接受度模型对本平台进行评价。分析结果说明用户觉得使用该平台各功能能够提高bug整理和编程学习的效率且平台具有易于使用的特征，经过测试后发现，平台能够提升用户的编程能力。

由于平台是以网站的形式呈现，对于正在编程的学习者而言，遇到bug后更倾向于直接百度搜索寻找帮助，而不是打开一个独立网站然后记录笔记。后续可以将平台嵌入到代码编辑器中，以插件的形式存在。这样用户遇到bug后可以直接在编辑器内记录或寻找解决方法，不必再打开其他软件，更加方便快捷，易用性也能相应得到提升。随着大语言模型的发展，其在编程学习领域仍有很多的应用空间，期望本研究能成为一次有意义的探索和启发。

参考文献

刘泽通.(2022).*基于项目式教程的编程学习平台设计与开发*.(硕士). 北京邮电大学, Available from Cnki

许嘉,钱韦伟,于戈 吕品.(2023).GameOJ:一种编程能力感知的游戏化在线评测系统.*华南师范大学学报(自然科学版)*(01),65-77.

李琪,姜强 赵蔚.(2023).促进计算思维发展的编程投入机制研究：基于 ICAP 理论.*电化教育研究*(08),88-95.

李彤彤,郝晴,文雨,苏新玉,方玉华 刘佳旋.(2022).基于学习风格和伙伴关系的配对编程对小学生计算思维的影响研究.*远程教育杂志*(03),105-112.

陈舒梦.(2023).大语言模型在外语教学中的应用研究.*长春师范大学学报*(11),170-173.

戴岭,赵晓伟 & 祝智庭.(2023).智慧问学:基于 ChatGPT 的对话式学习新模式.*开放教育研究*(06),42-51+111..

赵晓伟,祝智庭 & 沈书生.(2023).教育提示语工程：构建数智时代的认识论新话语.*中国远程教育*(11),22-31.

沈成 柏毅.(2023).大语言模型在小学生计算思维评测中的应用——以 ChatGPT 为例.*中小学数字化教学*(06),25-28.

Github.(2023).GitHub Copilot 快速入门.https://docs.github.com/zh/copilot/quickstart.

Leinonen, J., Hellas, A., Sarsa, S., Reeves, B., Denny, P., Prather, J., & Becker, B. A. (2023, March). Using large language models to enhance programming error messages. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (pp. 563-569).

Leinonen, J., Denny, P., MacNeil, S., Sarsa, S., Bernstein, S., Kim, J., ... & Hellas, A. (2023, June). Comparing code explanations created by students and large language models. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1* (pp. 124-130).

Dai, H. M., Teo, T., Rappa, N. A., & Huang, F. (2020). Explaining Chinese university students' continuance learning intention in the MOOC setting: A modified expectation confirmation model perspective. *Computers & Education*, 150, 103850.

教学视频中教师非言语行为的智能分析方法及其应用

Intelligent Analysis Method and Application Research of Non-Verbal Behaviors of Teachers in Instructional Videos

李妮辰[1]，陈辉[1]，何英杰[1]，杨九民[1*]
[1]华中师范大学 人工智能教育学部，湖北 武汉 430079
*yjm@mail.ccnu.edu.cn

【摘要】 教学视频中，教师的非言语行为辅助了言语行为对信息的传播，可以促进学习者的认知过程。随着智能教育的发展，利用人工智能技术可以对教学视频中教师非言语行为进行识别和分析，并基于分析结果对教学过程进行评估。本文对已有的教师非言语行为识别和分析方法进行了梳理，在此基础上构建了基于 Faster RCNN 的行为识别模型，并验证了该模型的准确度和有效性。本文的研究成果可以为教学视频中教师非言语行为的识别和分析提供方法上的借鉴，同时也为教师教学技巧的完善、信息化教学能力的提升、以及教学视频质量的改进提供一定建议。

【关键词】 教师非言语行为；人工智能；Faster RCNN；智能识别

Abstract： *In instructional videos, teachers' non-verbal behaviors assist in the transmission of information through verbal actions, facilitating the cognitive process of learners. With the development of intelligent education, artificial intelligence technology can be utilized to identify and analyze teachers' non-verbal behaviors in instructional videos and evaluate the teaching process based on the analysis. This paper reviews existing research and builds a behavior recognition model based on Faster RCNN, and verifies the accuracy and effectiveness of the model. The findings can provide methodological references for the identification and analysis of teachers' non-verbal behaviors in instructional videos. Additionally, they offer suggestions for the improvement of teaching skills, enhancement of information technology teaching capabilities, and improvement of the quality of instructional videos.*

Keywords: teacher's non-verbal behavior, artificial intelligence, Faster RCNN, Intelligent recognition

1. 前言

教学视频能够以视听双通道呈现教学内容，是在线学习资源的主要形式之一（杨九民等，2022），对学生的学习效果有着很大的影响。近年来，教学视频中出镜教师的“非言语行为”对学习者学习的影响备受关注（杨九民等，2023；Pi et al.,2020），已有研究发现，手势策略、教师目光等是影响视频学习效果的重要因素。因此，识别和分析教学视频中教师“非言语行为”对教师形象设计和学习策略选择意义重大。

当前，人工智能技术已经成为推动许多领域发展的关键驱动力，特别是在教育领域的影响日益显著。随着智能教育的发展，利用人工智能技术开展教师行为分析也逐渐成为了研究者们关注的焦点。通过 AI 进行目标检测、特征提取和行为分类，能更全面地识别教师的教学行为，自动化地对教师的非言语行为进行分析和可视化呈现，从而为教师的信息化能力提升和专业发展提供数据支持。这种智能识别方法可以解决人工观察评价带来的效率低下，主

观性强的缺点。

本文对教师非言语行为的相关理论做了梳理，对人工智能技术识别教师行为的方法做了总结，基于此提出了利用快速区域卷积神经网络（Faster-Region Convolutional Neural Network, Faster RCNN）的教师非言语行为模型，通过实践应用证明了模型的有效性。

2. 相关研究

2.1. 教学视频中教师的非言语行为研究

基于探究社区理论，学生在在线学习环境中的社会临场感会影响其学习效果以及学生的学习主动性（Ratan et al.,2022），而教师的非言语行为则是影响学生社会临场感的重要因素。当前，在教学视频中教师的非言语行为大致可以分为手势、面部表情、目光等几种类型。

教师手势与教学活动中师生的言语互动关系密切，其在视频学习中的作用备受国内外研究学者的关注。教学视频中教师的手势可以分为指示性手势、描述性手势、节拍性手势和混合手势等四种类型（杨九民等，2022）。以往研究发现教学视频中教师使用指示性手势可以提高学习者的迁移成绩（Pi et al.,2017）。同时，教师手势可以补充语义信息（Holle et al.,2010），例如教师可以通过手部形状和运动来展现言语信息，从而引导学习者的注意力分配、唤醒其心理意象（Mayer et al.,2022），促进其认知过程，从而影响其学习效果（Pi et al.,2019）。

教师在教学视频中面部表情所表达的情感可以帮助学习者克服孤独感，尤其当教师呈现出积极的表情时，学习者可能会感到放松从而缓解负面情绪。同时，教师的面部表情对学习者社会存在的感知至关重要（Schneider et al.,2022），这会影响他们的参与度（Borup et al.,2014）、唤醒水平和学习满意度（Wang et al.,2019）。

教师的目光是教学视频中教师最典型的非言语行为之一，也是教学视频设计与制作的重要要素。教师的目光在教学活动中可以有效传递注意焦点，影响学习者的注意力分配（Van Wermeskerken & Van Gog,2017）。根据目光注意焦点，可以将教师的目光分为四种类型：直视目光、持续引导目光、间隔引导目光和回避目光（Pi et al.,2020）。根据多媒体学习理论的线索原则（Mayer et al.,2022），教学视频中教师的目光可以提高教师与学习者之间的社会互动，激发他们社会反应，积极的社会反应则增加了学习者学习过程中的认知投入，从而提高其学习效果。

综上，教学视频中教师手势、目光和表情等非言语行为对学习者的学习有很大影响，因此需对教学视频中教师非言语行为进行的识别和分析，从而提升视频质量，促进学习者的学习。

2.2. 教师非言语行为识别方法的相关研究

传统的教师非言语行为分析方法主要依靠人力进行，通常采用问卷调查、课堂参与性观察、或访谈等方式展开研究，对于收集到的数据一般采用统计学方法进行人工记录分析。这类方法有一定弊端，存在着实效性较低，人力物力投入量大以及数据记录不够客观等问题。因此，随着信息技术的发展，利用技术代替人力对教师行为进行智能识别成为了发展的主要趋势。当前教育学领域的行为识别技术主要集中于课堂教学，利用人工智能技术收集教学过程中的师生行为数据，将教师的教学过程和学生的学习过程通过客观的数据表现出来，可以对整个教学过程进行一定的评估。目前，大多数的研究是以学生为研究对象（孙曙辉等，2018），相关研究以学生行为为聚焦点。例如，有研究人员在课堂场景下架设动作捕捉设备提取学生的骨架特征,然后用传统的机器学习方法支持向量机(SVM）对学生的骨架特征向量

进行分类，但是该方法操作比较复杂且准确率不高。刘清堂等人（2019）构建了教学行为智能分析模型，并使用 S-T 行为分析法验证了该分析模型的有效性。魏艳涛等人（2019）使用 VGG-16 网络并执行了迁移学习。

对教学视频中教师的非言语行为识别通常要涉及到特征提取、特征编码和特征分类三个阶段，需要首先从教学视频中提取教师的行为特征，然后对提取到的行为特征进行处理并编码为可以提取行为特征的数据，最后利用特征信息进行分类。随着深度学习技术特别是卷积神经网络的蓬勃发展，使用一个网络模型即可代替以上三个阶段，通过反向传播算法自动学习行为特征，因此卷积神经网络目前也成为了行为识别领域的主流方法。

本文的研究目的是设计出基于人工智能技术的教师非言语行为识别模型，并依据此模型开发出相应的检测平台并展开应用实践。本研究选取了 Faster RCNN 算法，此算法于 2015 年被提出，极大地推动了目标检测领域的发展。它的主要工作流程分为特征提取、RPN 网络、ROI 池化、分类和边界框回归等，相较于传统的 CNN 算法，它最大化的利用了卷积神经网络提取的特征信息，可以精准定位目标，并且识别精度也更高。

3. 教学视频中教师非言语行为智能分析模型的构建

基于对已有研究成果的梳理和总结，可以看出目前迫切的需要一种智能识别方法研究教学视频中教师的各类非言语行为，从而帮助教师在录制教学视频时进行形象设计或在真实课堂实践中更好地使用非言语行为辅助教学，进而激发学生的好奇心、吸引学生的注意力、提高学生的学习动机、促进学生的学习。但面对复杂的教学场景，目前在技术应用方面仍具有一定的困难，比如：教师的非言语行为种类较多，不同行为在教学视频中占比不同，造成数据构建类别不均衡和连续识别困难的问题；以及已有的教学视频质量层次不齐，数据集难以构建的问题。

综合这些问题，本文提出了一种基于 Faster RCNN 的教师非言语行为识别模型，如图 1 所示。本模型包括数据集构建、行为识别与检测以及数据可视化等几个模块。

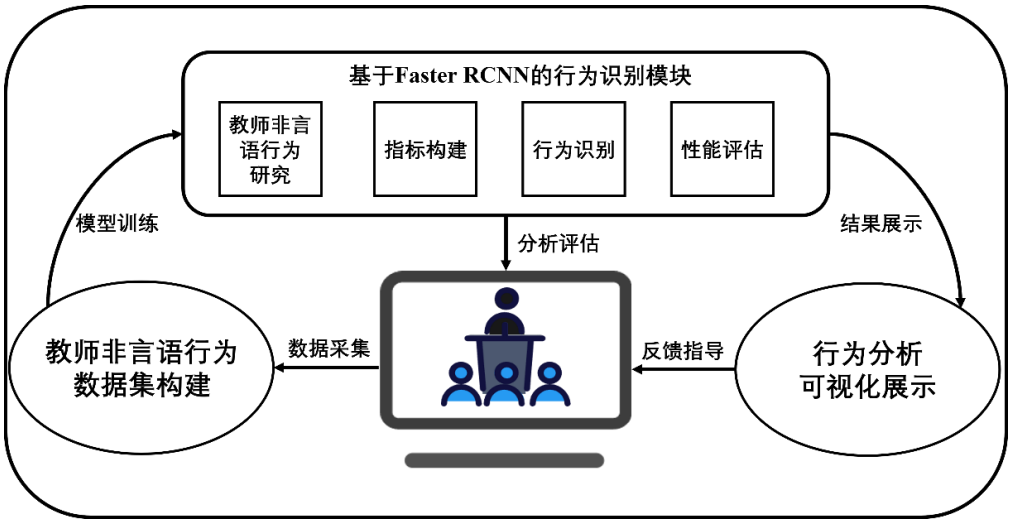


图 1 教学视频中教师非言语行为智能分析模型

3.1. 数据集的构建

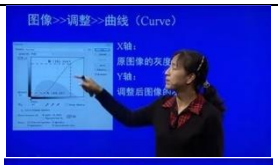
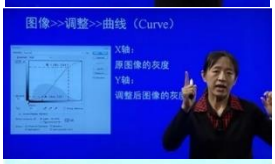
训练一个可靠的教师非言语行为识别模型需要大量的数据作为支撑，然而目前国内外的行为识别数据库例如 CASIA 行为数据集以及 UCF101 等并不特定针对于教育领域，所以为了满足数据训练的要求，需要构建相关的数据集。

基于对教师非言语行为相关研究的梳理，本研究构建的数据集主要包括三类：教师手势

数据、教师目光数据以及教师表情数据。其中教师手势数据包括指示性手势和描述性手势两种，教师目光数据分为直视目光、引导目光和回避目光，教师表情数据分为积极表情、中性表情和消极表情，具体指标如表 1 所示。

由于本研究是在监督学习的基础上进行的，即对已标注的数据进行训练，因此在收集和构建数据后，再对其进行标注是必要的环节。数据标注可以帮助目标检测模型学习目标物体的位置和类别信息。基于本研究的特点，教学画面中的教师为数据标注的目标，而教学视频画面中的学生以及桌椅、黑板等其他教学工具和教学环境则被视为背景，不在数据标注的范围内。

表 1 教师非言语行为类别

一级指标	二级指标	
手势	指示性手势	
	描述性手势	
目光	直视目光	
	引导目光	
	回避目光	
面部表情	积极表情	
	中性表情	
	消极表情	

3.2. 行为识别与检测

本文的行为识别基于深度学习技术，采用了目标检测算法 Faster RCNN 用以识别教学视频中的教师非言语行为图像。相较于传统的 CNN 算法，Faster RCNN 最大程度上利用了卷积网络提取的特征信息，它提出了 RPN 网络，将 RCNN 和 Fast RCNN 中存在的候选区域生成的问题和卷积结合到了一起，可以提高检测的速度以及检测率。Faster RCNN 流程如图 2 所

示，主要包括以下步骤：基础特征提取、RPN 生成候选区域、对候选区域进行筛选和特征提取以及目标分类和边界框回归等。在本研究中使用 Faster RCNN 算法有以下优点：首先，此算法性能优越，检测精度较高，可以在多尺度、小目标的问题上表现的较好，这是符合本研究的教学背景的；其次，由于本研究需要识别的非言语行为类别较多，而 Faster RCNN 可以进行迁移学习，加载大规模数据集训练的预训练模型后，对数据集中的目标类别进行更改就可以得到比较好的检测结果。

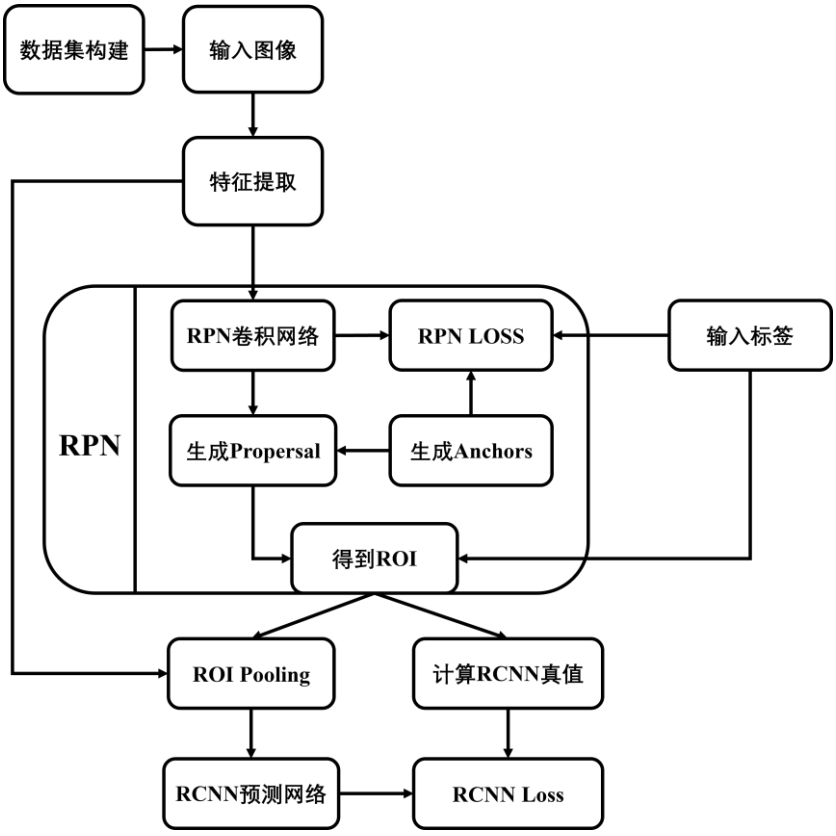


图 2 Faster RCNN 流程图

3.3. 数据可视化

为了方便展示最终结果数据，在构建的教师非言语行为识别平台中加入了数据可视化模块，对一段教学视频中教师各类非言语行为所占比例进行计算并以图表的形式展现。这一模块方便研究人员直观的判断教学视频中教师非言语行为，为后续利用数据指导教学视频质量改进提供了依据，方便提升教学效能。

4. 教学视频中教师非言语行为识别的实践应用

基于上述模型的构建，本研究开展了应用研究，并搭建了非言语行为识别的平台。

4.1. 数据的采集与标注

4.1.1. 数据来源

本研究中数据集质量的高低与收集的原始教学视频样例息息相关，考虑到教学视频的呈现质量容易受到拍摄时摄像机型号、架设位置、角度以及视频场景等问题的影响，本研究在构建数据集时选择的教学视频材料均来自国家中小学智慧教育平台和国家教育资源公共服务平台中的“优课”和“名师课堂”视频。

4.1.2. 教师非言语行为图像数据库的构建

收集到视频材料后，对这些视频进行剪辑，保留教师存在的片段。视频处理结束后，利用视频图像捕获软件，对视频进行截帧，每隔 5 秒采样一帧图像，针对四种行为类别的视频，最终共获取图像 5820 幅。

获取图像数据后，需要对目标检测任务进行标注，即人工使用矩形框标记图像中待检测目标的位置和类别，这一项任务由两位研究人员完成，所使用的标注工具是 Windows 平台下的 LableImg。

最终，构建的数据集中 70%的数据用于训练集，30%的数据用于测试集，其构成如表 2 所示。

表 2 数据集参数

数据集	图像（幅）
数据集总数	5820
训练集	4074
测试集	1746

4.2. 行为识别与分析

本研究依据模型中提出的基于 Faster RCNN 算法的教师非言语行为识别方法搭建了相关的检测平台。在平台中，可以上传待检测的教学视频，之后在前端使用 javascript 定时截取视频图片，通过的到的图像特征信息，对教师行为进行识别（如图 3 所示）。

在识别结束后，利用识别的结果进行统计分析，最终生成可视化图。可视化的结果展示了各类教师行为在教学视频中的时间占比以及随时间的变化情况，具体结果如图 4 所示。



图 3 教师非言语行为识别结果（以目光为例）



图 4 可视化结果展示

4.3. 性能评估

Faster RCNN 算法作为一种目标检测算法，其评估主要依据对教师非言语行为类别检测的精确度。本研究采用了精确率和召回率两个指标来衡量所开发的检测平台的性能。精确率（Precision）是指模型预测为正类的样本中，实际为正类的比例。召回率（Recall）是指在所有实际为正类的样本中，模型正确预测为正类的比例。为了计算这两个指标，需要首先构建混淆矩阵，包括真正例、假正例、真负例和假负例等值。由于每一类教师非言语行为都由多种指标构成，所以在计算时首先计算各指标的精确率和召回率，之后对该类所有指标的精确率和召回率求平均值得到最终结果。具体的评估结果如表 3 所示。

表 3 评估结果

指标		评估结果
精确率	手势	0.657
	目光	0.810
	表情	0.786
召回率	手势	0.663
	目光	0.870
	表情	0.741

5. 总结与展望

本研究基于 Faster RCNN 深度学习算法，创建了可用于教学视频中教师非言语行为识别和分析的模型，根据此模型开发了相关检测平台，并证实了检测模型的有效性和精确度，给教学视频中教师非言语行为的自动识别和分析提供了方法上的借鉴。在未来，此模型可用于检测教学视频中出镜教师的形象，这不仅可以帮助教师录制教学视频时开展形象设计，从而提高自己的信息化教学能力，也可以从教师形象的角度提升教学视频的质量，最终促进学生的在线学习效果。

本研究存在着以下三点局限：首先，出于对视频样例质量的考虑，在构建数据集时只选

用了两个平台的教学视频，视频内容背景同质性较高，这导致本研究数据集不够丰富，对精确度产生了一定的影响。未来可以考虑构建更加丰富的数据集，特别是要多纳入真实课堂的教学视频实录，继而再次对模型进行训练，提高检测精确度。其次，本研究只考虑了教师的非言语行为研究，然而教学过程是教师和学生共同参与的，师生互动行为对学习者的学习也有较大影响，未来可以考虑纳入学生的互动行为分析。最后，本研究缺乏对教学视频评价指标体系的构建，这使得评估数据未能发挥出最大的效用，未来可以根据本研究中提到的非言语行为类别，对其构建出相应的评价指标体系，从而对教学视频质量的改进和出镜教师形象的设计提供更加客观的建议。

参考文献

刘清堂，何皓怡，吴林静，邓伟，陈越，王洋和张妮（2019）。基于人工智能的课堂教学行为分析方法及其应用。中国电化教育(09)，13-21。

孙曙辉,刘邦奇和李鑫。面向智慧课堂的数据挖掘与学习分析框架及应用。中国电化教育(02)，59-66。

杨九民,刘彩霞和陈楼琪（2023）。教学视频中的描述性手势何以有效——手势学习策略及学习者空间能力的影响。现代远程教育研究(04)，92-101。

杨九民,杨文蝶和陈辉（2022）。教学视频中的教师手势起作用了吗？——基于 2000—2021 年 40 篇实验和准实验研究的元分析。现代远程教育研究(01)，92-103。

魏艳涛,秦道影,胡佳敏,姚璜和师亚飞。基于深度学习的学生课堂行为识别。现代教育技术(07)，87-91。

Borup, J., West, R. E., Thomas, R., & Graham, C. R. (2014). Examining the impact of video feedback on instructor social presence in blended courses. The International Review of Research in Open and Distributed Learning, 15(3).

Holle, H., Obleser, J., et al. (2010). Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. Neuroimage, 49(1), 875-884.

Mayer,R.E.,&Fiorella,L.(2022).The Cambridge Handbook of Multimedia Learning(3rd Ed.)[M].Cambridge:Cambridge University Press:286-293.

Pi, Z., Xu, K., Liu, C., & Yang, J. (2020). Instructor presence in video lectures: Eye gaze matters, but not body orientation. Computers & Education, 144, 103713.

Pi, Z., & Hong, J. (2017). Effects of the instructor's pointing gestures on learning performance in video lectures. British Journal of Educational Technology, 48(4), 1020-1029.

Pi, Z., Zhang, Y., Zhu, F., Xu, K., Yang, J., & Hu, W. (2019). Instructors' pointing gestures improve learning regardless of their use of directed gaze in video lectures. Computers & Education, 128, 345-352.

Ratan, R., Ucha, C., Lei, Y., Lim, C., Triwibowo, W., Yelon, S., ... & Chen, V. H. H. (2022). How do social presence and active learning in synchronous and asynchronous online classes relate to students' perceived course gains? Computers & Education, 191, 104621.

Schneider, S., Krieglstein, F., Beege, M., & Rey, G. D. (2022). The impact of video lecturers' nonverbal communication on learning–An experiment on gestures and facial expressions of pedagogical agents. Computers & Education, 176, 104350.

Van Wermeskerken, M., & Van Gog, T. (2017). Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. Computers & Education, 113, 98-107.

课堂文本语言行为自动分析系统开发及应用

Development and Application of Automatic Analysis System for Classroom Textual Language Behavior

黄立新[1]，卢卓豪[2*]，薛松[3]

[1,2,3]浙江师范大学

*gzc1.666@gmail.com

【摘要】从机器学习的视角出发，以课堂语言行为分析理论为基础，基于开发的课堂文本语言行为自动分析系统，选取 100 节课堂和近 30000 句课堂文本形成语言行为标准分数常模，并从“行为、语速、词性”三维度提出课堂语言行为评价体系。随后，选取四名初中信息技术实习教师进行为期 2 个月的语言行为量化分析，通过比较优质课堂发现特点及其差异，并从应用评价体系、学生成绩出发讨论教师教学行为是否得到提升。测试表明，其不仅能提高课堂分析效率，而且对提升实习教师的教学语言行为、专业技能和学生学业水平具有积极影响，尽快帮助教师过渡“新手期”。

【关键词】 机器学习；互动分析；互动分析系统；教师专业成长

Abstract: *From the perspective of machine learning and grounded in classroom language behavior analysis theory, this study developed an automated classroom text language behavior analysis system. Utilizing around 30,000 sentences from 100 classroom sessions, a standard score norm for language behavior was established. The evaluation system proposed examines three dimensions: behavior, speech rate, and parts of speech. Subsequently, four middle school information technology student teachers were selected for a two-month quantitative analysis of their language behavior. By comparing with high-quality classrooms, the characteristics and differences were identified. The discussion extends to whether there has been an improvement in teaching behavior based on the application of the evaluation system and student achievements. Results show that the system not only enhances the efficiency of classroom analysis but also positively impacts the improvement of student teachers' language behavior, professional skills, and students' academic levels. It indicates that the developed system could facilitate a quicker transition for teachers during their novice stages.*

Keywords: Machine Learning; Interactive Analysis; Interactive Analysis System; Teacher Professional Development

1. 前言

对课堂教学进行行为类型的编码与分析是互动分析最重要的目的之一（魏宏聚,2009）。精准分析课堂各阶段情况对教师提升教学、学生深层投入、师生良性互动有显著作用（Kakosimos,2015）、（傅德荣和章慧敏,2001）。现实课堂记录存在多种手段和分析方法，包括 S-T 分析法、传统弗兰德斯互动分析法（FIAS）等，其中以 FIAS 分析法最为主流，传统 FIAS 互动分析法包含教师、学生、沉寂等共 10 种编码类别，每种类别又包含多种交互行为，对行为进行编码并进行数据分析和呈现（魏宏聚,2009）。

现有大部分观察技术和标准的来源均基于弗兰德斯互动分析系统，弗兰德斯教授采用质性方法进行课堂观察及编码，再通过矩阵计算得出结论，该方法大大简化了课堂教学研究过程的复杂性，同时又提高了观察记录的科学性和准确性（张海,2014）。

Student-Teacher 分析法简称 S-T 分析法，是一种有效引用于教学过程分析的定量分析方法，能将教学结果以图表形式表示出来，并根据图表所展示的数据进行计算，使教师根据计算结果对教学过程进行分析与讨论。薛新国团队改进传统 S-T 教学分析法，将课堂教学行为划分为教师行为 T、学生行为 S 以及师生互动行为 D 三类，并以此编码课堂教学过程，分析课堂教学模式（薛新国,2019）。黄苗苗团队改进 Flanders 编码，通过课堂观察的形式，标准化的分析了信息技术优质课堂师生互动行为，获取有效地课堂教学行为信息（黄苗苗,钟媚,潘霞,杨彩香,2021）。

现有团队对 FIAS 的改善主要围绕课堂行为分类、技术记录和矩阵分析维度，以满足现代课堂互动分析需求。但存在局限性：一是课堂现实记录难度高、精度低，容易产生误判或错漏（方海光,高辰柱,陈佳,2012）；二是人工记录耗时、要求高，难以保证一致性和准确性；三是课堂切片颗粒度高，行为转换记录不足(徐丽丽,2020)；四是连续课堂关联度低，缺乏教师成长监督；五是海量课程分析费时，缺乏短时分析手段。基于此，本研究旨在探讨基于文本分类的课堂话语分析工具的开发及应用并提出教师教学行为优化策略研究，帮助教师快速获得专业发展。

本研究的研究问题为：通过开发自动化课堂文本言语行为分析工具，对比、寻找实习教师与优质教师差异并提出教学行为优化策略，观察教师课堂教学行为是否得到优化，专业技能是否获得成长？

2. 课堂文本语言行为分析系统的设计

2.1. 语言行为编码体系确定

教学活动中，师生语言互动行为是主要教学行为，弗兰德斯教授在 20 世纪 60 年代采用质性方法进行课堂观察及编码，并利用矩阵计算得出结论，从而形成弗兰德斯互动分析系统（FIAS），其中课堂语言行为分为 3 大类 10 小类，其中 1-7 为教师语言，8-9 为学生语言，10 为课堂沉默(Flanders, Ned,1963)。虽然弗兰德斯课堂语言互动分析系统解决了研究课堂教学无从着手的境况，但在利用机器学习分析现实课堂的情况下存在弊端，导致课堂语言行为分析准确度的下降，因此本研究在根据人工分析 30 节以上现实课程的基础上，对行为编码体系进行相对应修改。

修改内容如下：（1）教师语言中删去“接纳学生情感”的分类，现实课程中教师较少接纳学生情感，并且利用文本分类时需要介入情感分析，导致其分析结果准确度不高，故删去。（2）将“学生主动发言”分为主动提问和主动发言，现实课程中学生表现较公开课等更为活跃，学生部分语言并非关乎课堂，但是学生能够主动提问说明注意力集中，积极思考并且文本分类对提问和发言的区分度较高，适合作为一个单独分支。（3）将课堂沉默分为有效沉默和混乱或无效沉默，现实课程中较公开课更容易出现混乱或无效沉默，同时有效沉默与混乱或无效沉默的对比也凸显出教师对课堂的掌控度与水平，但 FIAS 编码体系并没有进行划分，应当补全。具体情况如表 1 所示。

表 1 ACTIAS 课堂语言行为编码体系

序号	分类		编码	分类
1	教师语言	间接影响的语言	1	称赞或鼓励
			2	接受或利用学生的想法
			3	教师提问
		直接影响的语言	4	教师讲解
			5	命令或指示
			6	批评学生
2	学生语言		7	回答教师的问题
			8	学生主动提出问题
			9	学生主动发言
3	课堂沉默		10	课堂有效沉默
			11	课堂混乱或无效沉默

2.2. 算法设计与评估

使用计算机文本分类技术可以处理课堂话语并进行分类，但是现有文本分类模型较多，包含常规算法（k 近邻、决策树、贝叶斯算法……）、集成学习算法（随机森林、自增强、lightgbm 算法……），深度学习算法（多分类前馈神经网络，LSTM 神经网络……）。多种分类模型中，面对教师计算机环境需要选择轻量化、精度高、速度快的模型方案，经过对比，随机森林模型在适当时间内（2 分钟左右）达到 80%左右的准确率符合教师计算机环境要求。此外随机森林具有高准确度、可以处理大量输入变量、遗失资料仍可维持准确度、学习过程快速等优点。主要的文本分类处理步骤包括：批量读取文本数据集、中文文本分词、停用词使用、编码器处理文本标签、数值向量转换、建模训练等。随后根据训练得到模型，将课堂话语输入并经过中文文本分词、通用词清洗、数值向量转换后判断话语类型并输出。

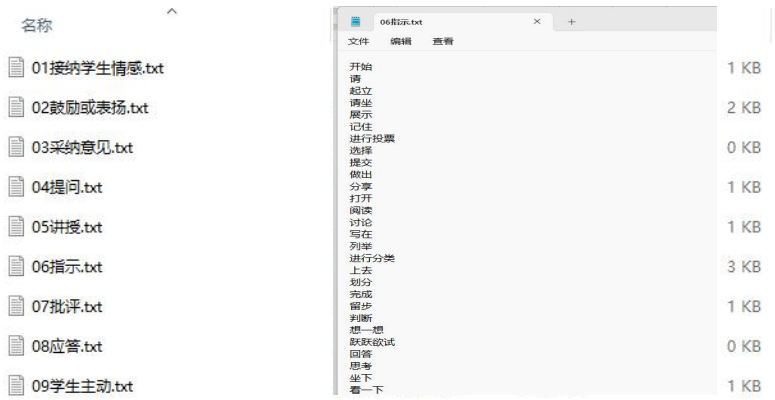


图 1 课堂语言词库示例

2.3.算法有效性验证

为评估基于随机森林—修正算法的性能及有效性，必须进过一系列严格的实验。结合现实世界及人工修改批注的的课程语言行为数据集，其中人工修改批注经过三位熟悉流程且经验丰富的人员进行两轮的批注、对比、修改形成人工批注数据集，最终形成近 30000 句带有批注标识的课堂师生语言行为数据集。随后将数据集经过批量读取

文本数据集、中文文本分词、停用词使用、编码器处理文本标签、数值向量转换、建模训练等一系列的数据预处理步骤，以确保数据的质量和一致性。

为了衡量基于随机森林算法的性能，选择了一组合适的评估指标，包含准确率（accuracy）、精确率（precision）、召回率（recall）、F1 分数（F1-score）、支持（support）、宏平均（macro average）和加权平均（weighted average），以对模型进行全面的评价。具体数据如下表 2、表 3。

表 2 学生混淆矩阵报告

	精确率	召回率	F1 分数	支持
0	0.94	0.99	0.96	381
1	0.83	0.28	0.42	36
准确率			0.93	417
平均宏	0.88	0.64	0.69	417
加权平均	0.93	0.93	0.92	417

从表 2 数据分析可知，学生文本分类模型总体的准确率达到 93%，其中类别 0 的准确率为 94%，类别 1 的准确率为 83%，这是由于类别 0 在样本数量上远大于类别 1。然而从宏平均和加权平均的角度进行综合考虑发现，两者的准确率，召回率和 F1 分数都提升到可接受范围，尤其是在加权平均角度下，准确率达到 93%，召回率为 93%，F1 分数为 92%，说明模型在对学生两种类别的识别下表现较好，能够精确区别学生的回答问题和主动说话。

表 3 教师混淆矩阵报告

	精确率	召回率	F1 分数	支持
0	0.75	0.55	0.63	351
1	0.85	0.82	0.84	758
2	0.74	0.84	0.79	886
准确率			0.78	1995
平均宏	0.78	0.74	0.75	1995
加权平均	0.79	0.78	0.78	1995

从表 3 的数据分析可知，教师文本分类模型总体准确率达到 78%，其中类别 0、1、2 的准确率分别为 75%、85%、74%，说明在每项类别的区分上表现较好。从宏平均和加权平均的角度进行综合考虑发现，模型的各项达到 75-80%的水平，说明模型能够在多分类任务的区分上表现尚可，能够区分教师的提问、讲授和指令不同行为。

以上内容全面阐述了实验设计、数据集选择、数据预处理和评估指标的使用，以及基于规则—数据集的修正算法的验证和实验结果分析。通过以上的实验以及结果评估，表明在利用基于随机森林—修正算法对课堂语言行为进行分类后，基本满足现实课堂语言分析对文本预处理、划分师生语言、标识语言行为等需求。

3. 优质课常模生成

常模是指在分析测验结果时，用来确定被试者相对水平的参考系，是衡量个体水平的标准。本研究开发的软件适用于课堂评估、师生互动评价、教学改进等方面，因此制定切实、有效、具有指导作用的常模意义重大。要建立有意义的常模，一是要获取有代表性的常模团体，而是要以常模团体的原始分数为基础，用统计学的方法来到处一种具有特定意义的数据，能反映个体在群体的分布情况和相对位置(方海光,孔新梅,2023)。

美国学者弗兰德斯利用互动矩阵分析法所得到的一系列数据更进一步的分析，提出了12项指标以解释该教学行为背后所隐含的意义，并且在统计大量课程数据后得出 12 项指标的常模作为标准，帮助教育者了解自身行为、教育过程中的关键要素、师生建立积极有效的行动、增强学生的参与感和学习动力(Flanders 和 Ned,1963)。

本研究将在借鉴弗兰德斯互动分析常模的基础上，利用自动化分析软件分析大量优质课例数据，为其建立互动分析常模并作为评判标准，突出优秀教师与实习教师在教学行为上的差别。因此通过搜集来源于知网现有文献中优质课例数据以及自主分析优质课数据，共计近 100 节的数量计算平均值，根据标准分数常模规则，并将其作为优质课互动分析常模帮助对比实习教师课堂行为、寻找差别并提出解决意见。

表 4 优质指标常模

行为维度	课程行为比例	优质比例常模	标准差（σ）	理论课比例常模	实践课比例常模
直接角度	教师言语比例	55.9%	13.97	59.27%	52.53%
	学生言语比例	22.10%	5.52	23.93%	20.27%
	有效沉寂	13.0%	3.25	10.58%	15.42%
	无效沉寂	6.10%	1.53	4.1%	8.1%
	沉寂或混乱	19.10%	4.78	14.68%	23.52%
	有效沉寂/无效沉寂	224.21%	28.80	258.05%	190.37%
间接角度	间接影响/直接影响	72.70%	18.20	67.87%	77.53%
	教师反应比例	55.30%	13.83	50.59%	60.01%
	教师发问比例	20.20%	5.05	20.75%	19.65%
	学生自发比例	31.30%	7.83	13.96%	48.64%
	教师实时反应比例	71.10%	17.78	71.33%	70.89%
	教师实时发问比例	55.00%	13.75	45.70%	64.3%
	内容十字区	51.70%	12.93	56.07%	47.33%
	稳定状态区	57.40%	14.35	61.90%	52.9%
	学生稳定状态区	55.90%	18.95	58.68%	53.12%

在借鉴弗兰德斯指标常模的基础上，增加了“有效沉寂”、“无效沉寂”、“有效沉寂/无效沉寂”三个指标，用来细化课堂分析维度，形成表 8 所展示的优质指标常模。但在实际应用的过程中发现各项指标比例是随着课程安排、教学进度、互动频次的不同而浮动的，因此单一的常模缺乏应用价值，为便捷后续数据处理与教师分析，需要定义一些范围进行区分。

4. ACTIAS 的教学应用

4.1.实验流程设计

研究者在前期文献调查和软件开发的基础上，选定初中代表性课程、制作课程教学设计、录制课堂实录、分析并调整实验方案，从而确定正式实验。选取两名水平相当实习教师形成一组，共两组作为研究对象。两名教师仅进行一次课前备课，并直接进行授课，其中一名利用工具对课堂语言行为进行量化，并结合指导教师意见和优质课例视频进行反思、另一名仅有指导教师意见和优质课例视频进行反思。随后两名教师按照正常教学进度上课，经过历时 2 个月的实验，进行两名教师的成长图对比作为直接依据与班级成绩对比作为辅助依据。实验组教师流程图如图 2 所示。实验对照图如图 3 所示，即对照组缺乏课堂实录分析步骤。

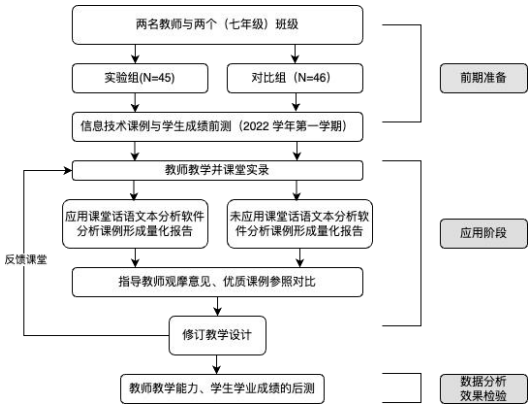


图 2 实验流程图

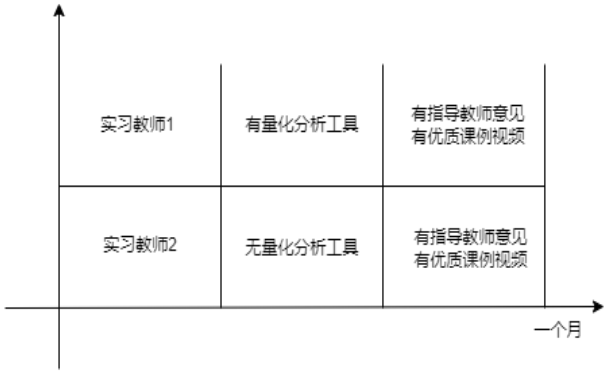


图 3 实验对照图

通过实施两组为期两个月的实验，不断在上课过程中调整策略与方案，来验证在课堂语言量化分析工具的帮助下教师能否获得专业发展上的提升，以及借助教师专业发展综合模式下的学习能否促进教师教学行为上的优化。

4.2.课程量化数据分析

本研究设计的课堂语言行为分析软件主要从课堂师生行为、语速、词性三维角度进行分析，并基于大量优质课例数据形成优质课例常模，通过评分机制为课堂师生行为、语速进行评分，其中由于词性因人而异且不构成评分依据，因此词性仅作为教师生成报告参考意见的一环。本研究选取两名水平相当实习教师形成一组，共两组作为研究对象。通过录制四名实习教师为期 2 个月、共计 10 节初中信息技术课程，最终对比课程总分、行为分数、语速分数，观察教师是否在软件的辅助下获取更为明显的进步与提升。受制于篇幅限制，当前仅展现一组实验数据作为说明。

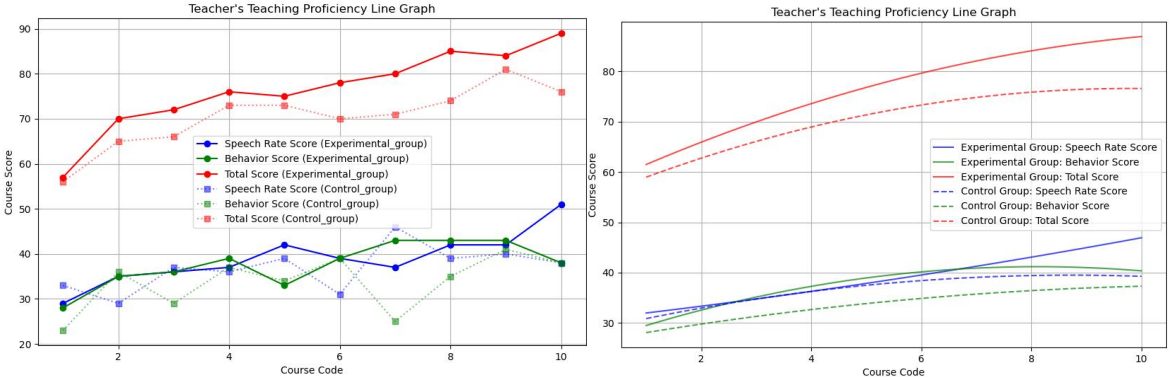


图 4 实验-对照组教师分数变化折线图

通过观察图 4 的课例各项分数变化，实验组别二的教师教学能力各项分数和趋势变化，可以发现实验组教师不论从教学得分、教学能力提升、趋势走向上均比对照组教师表现优异。实验组教师从最初 57 分，利用分析软件经过一系列积极的调整，最终稳步提升至 89 分，虽然中间出现小幅波动，但最终都能够根据意见做出积极调整。而参照组从最开始 56 分提升至 76 分，虽然教师也得到提升，但是提升幅度明显小于实验组。实验组的语速得分出现较大浮动，说明教师在努力控制自身语速过程中，经历了一段适应期，但经过 4-5 节课的调整，随即能够合理控制自身语速。而参照组的语速在自身范围内上下浮动，未出现显著变化。实验组和参照组在行为得分上均出现稳步提升趋势，但是实验组的得分相对参照组提升更加稳定，波动较小。说明参照实验组意

见的教师更够把握课堂师生行为，控制课堂节奏和不同环节的分配。

这些显著的提升可以归因于实验组教师获得了课堂话语文本分析软件提供的详细反馈。通过这种反馈，他们能够清晰地了解自己在课堂细节、具体问题以及提升方向上的不足。因此，他们能够在较短的时间内进行有针对性的纠正和改善。相比之下，对照组教师由于缺乏这种具体的反馈，所获得的信息较为笼统，难以在短时间内实现有效的教学转换与提升。综上所述，课堂话语文本分析软件对教师教学能力的提升具有显著的效果。它为教师提供了精准的反馈信息，帮助他们更好地理解自己的教学情况，并针对性地改进教学方法和技巧，从而提升教学质量和效果。

4.3.学生练习数据

随着新手教师课堂行为的逐渐规范、合理化，学生的练习成绩将会在一定程度上得到提升。本研究通过每节课后对学生进行课堂知识习题的测试并记录成绩，来辅助说明教师教学行为提升将直接影响学生学习效果。本研究要求教师在每堂课后都对学生进行当堂信息技术知识习题的练习，采取满分 5 分制对学生进行打分并记录。通过为期 2 个月，共计 10 堂课进行前后对比。

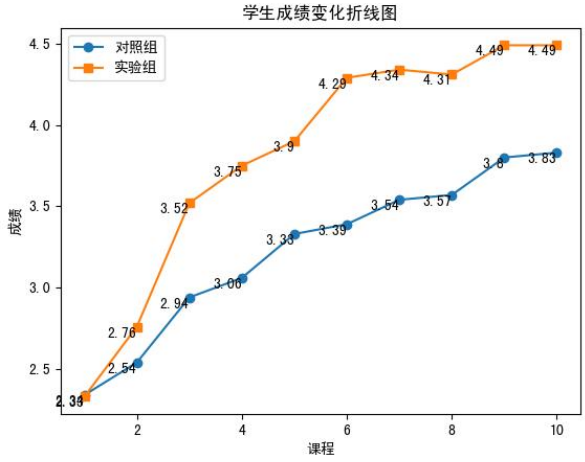


图 5 实验组学生成绩折线图

根据图 5，在编程方面零基础且相同水平的班级均在 10 堂课后获得了不同程度的提升，但是实验组成绩较对照组提升明显。值得关注的是，从第三、四节课之后，随着教师课堂熟练度提升、教学行为合理化、语速趋于稳定，学生更能够积极地参与课堂、高效地接受知识、并能够将所学内容融会贯通。此外，自第六课之后，实验组学生成绩趋于稳定，说明课堂大部分学生已经对编程知识形成基本概念，在编程方面短时工作记忆广度得到了提升，能够灵活提取课堂知识解决实际问题。

实验组学生课堂练习成绩的显著提升及其稳定性，不仅证明了教师在为期两个月的软件应用中在教学行为、语速及教学语言方面取得明显进步，同时也反映出利用课堂话语文本分析软件（ACTIAS）能够在教师专业成长中起到积极作用。它不仅为教师提供了明确的成长路径，还有效地提升了教师课堂表现与学生学业水平。

5. 结束语

通过实际研究与应用可以发现本研究开发的系统以及分析结果具有一定的真实客观性，它不仅可以像 FIAS 等分析法一样对课堂教学行为进行分析，且其过程解放人力、过程简便、可操作性强、连续性高、颗粒度细。但是现阶段软件还存在分析维度较少、缺乏对师生互动更多维度的分析，这些方面还需要进一步优化改进，并且将其与自动化更加靠近，达到真正的计算机赋能教育，从而更好地服务课堂数字化分析。此外，面对课堂环境的纷繁复杂，仅靠语言文本分析是不够的，之后可以利用计算机视觉介

入，分析师生课堂动作、表情等维度，更加精准分析师生状态，为新教师过渡、课堂节奏把控、学生学习效果提升做到更好地推荐，这也将作为下一步深入研究的方向之一。

参考文献

魏宏聚. (2009). 教师教学行为研究的几个维度与评析. 河南大学学报(社会科学版), (9), 126-130.

傅德荣, 章慧敏. (2001). 教育信息处理. 北京: 北京师范大学出版社, 95, 104.

张海. (2014). 弗兰德斯互动分析系统的方法与特点. 当代教育与文化, (6), 68-73.

方海光, 高辰柱, 陈佳. (2012). 改进型弗兰德斯互动分析系统及其应用. 中国电化教育, (10), 109-113.

方海光, 孔新梅. (2023). 一种智慧课堂交互行为数据的标准分数常模计算方法研究——基于改进型弗兰德斯互动分析系统 iFIAS. 远程教育杂志, 41(05), 67-75.

薛新国. (2019). S-T 分析法及其在教学中的应用. 江苏教育研究, (02), 4-8.

黄苗苗, 钟媚, 潘霞, 杨彩香. (2021). 基于 FIAS 的优质课堂师生行为研究. 教育信息技术, 2021(03), 61-64.

徐丽丽. (2020). 基于文本分类的课堂语言行为识别研究. 华中师范大学, 学位论文.

Kakosimos, K. E. (2015). Example of a micro-adaptive instruction methodology for the improvement of flipped-classrooms and adaptive-learning based on advanced blended-learning tools. *Education for Chemical Engineers*, (12), 1-11.

Flanders, N. (1963). Intent, Action and Feedback: A Preparation for Teaching. *Journal of Teacher Education*, 3(14), 25-260.

数字化学习资源关联网络研究综述：现状、问题及趋势

Review of Digital Learning Resource Association Network Research: Current Status, Problems, and Trends

张惠影[1]，李新[2]，杨现民[3*]
[123]江苏师范大学智慧教育学院
[*] yangxianmin8888@163.com

【摘要】 大规模开放协同环境下数字化学习资源生产和传播是一个复杂网络演变的过程，学习资源在生成和分享过程中逐步与其他资源建立关联，扩大资源群体规模并调整自身网络结构形成数字化学习资源关联网络。理清数字化学习资源关联网络内涵、特征及发展现状对于实现资源的非线性整合、识别资源群体进化规律、建设与管理数字化学习资源有着指导意义。本文聚焦数字化学习资源关联网络研究三大难题，预测了数字化学习资源关联网络发展趋势并探讨实践策略，旨在丰富和发展学习资源关联网络理论体系，为推动资源群体高效组织和管理给出实践参考。

【关键字】 复杂网络；数字化学习资源；关联网络；资源进化

Abstract: *In large-scale, open collaborative settings, the production and dissemination of digital learning resources evolve as a complex network, gradually forming connections and expanding the network. Clarifying the essence, characteristics, and evolution of this digital learning resource network is crucial for resource integration, understanding group evolution, and management. This article focuses on three significant challenges within this network, forecasts development trends, and discusses practical strategies. It aims to enrich and develop the theoretical framework of the learning resource network, offering practical insights for efficient organization and management of resource groups.*

Keywords: Complex network, Digital learning resources, Association network, Resource evolution

基金项目：国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”（项目编号：62077030) ;2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）

1.引言

我国数字化学习资源建设已进入开放共创阶段。传统的数字化学习资源建设以静态、单模态、更新缓慢为特征，存在呈现形式单一、进化动力不足等缺点，难以满足学习者的个性化学习需求。开放知识社区如维基百科、学习元、知乎等的兴起和生成式人工智能的出现，使资源产生和传播的主体由个体建构拓展至群体、甚至机器的“群建共享”，数字化学习资源的持续生成和进化得以实现。通过大规模用户的协同建设，数字化学习资源的规模急速扩增，开始向碎片化、移动化、开放式、生成性转型，资源间的关系由传统的线性结构转变为复杂的网络结构，知识载体和资源形态也更加多元化（陈丽等，2023），已逐步演化为一个复杂的网络体系。

复杂网络理论分析与应用研究是一门贯穿计算机科学与技术、控制理论与工程、管理科学等领域的交叉学科，正在引发教育界的关注。对于任何包含大量组成要素的复杂系统，当把构成要素抽象为节点、各要素之间的相互关系抽象为边时，都可以当作复杂网络来研究（卜湛等，2019）。数字化学习资源关联网络是一个典型的复杂网络，资源是网络中的核心组成要素，资源与资源之间的关联关系是网络中的边。因此，复杂网络理论能够为揭示数字化学

习资源关联网络的研究提供有力的方法支撑，数字化学习资源关联网络具备复杂的网络拓扑结构，运用复杂网络分析方法能够识别其整体结构样态，进而挖掘深层次的资源进化规律，提高资源可进化力。

总的来说，数字化学习关联网络的研究对于识别资源群体进化规律、推动资源有序进化具有重要意义。该网络受用户选择和环境选择双重因素的影响（徐刘杰等，2018），会进行自身内容和结构的完善和调整，从而实现资源的整体进化，能够推动资源进化由无序走向有序，由低级有序走向高级有序，从而优化资源管理和应用模式。基于此，本文对当前有关数字化学习资源关联网络的相关研究进行梳理，在明确数字化学习资源关联网络内涵与类型的基础上，对其在复杂网络理论与方法视域下表现出来的结构特征、演化规律等问题进行探讨，分析当前研究中的不足并归纳未来发展方向以为后续研究的开展提供支撑。

2.数字化学习资源关联网络研究现状

2.1.数字化学习资源关联网络概念及分类

互联网时代知识生产和传播是一个复杂网络演变的过程，学习资源在生成和分享过程中逐步与其他资源实体建立关联关系，扩大自身的网络规模，形成数字化学习资源关联网络。数字化学习资源关联网络从图书情报领域中的“知识网络”这一术语延伸而来，知识网络是知识社区中所生产的不同类型知识及其之间逻辑关系的结构化组织（万海鹏等，2021）。随着数字化学习资源走向群建共享，资源主体的分享和传播行为在资源关联网络中的地位越来越重要，数字化学习资源关联网络受用户行为和资源的自组织行为的双重影响，所以数字化学习资源关联网络的构成要素不但包括资源本身，还融入了资源主体，并且其关联关系更为复杂和多样。本研究认为数字化学习资源关联网络是将海量、分散且无序的资源作为一个整体进行考察，通过抽象出其中的组成要素和资源之间的关联特征而形成的一种复杂网络。用图论的方式表示即：$G=\{V,E,R|V_i\in V,E_i\in E\}$，式中 V 表示资源节点集合，$V_i$ 和 E_i 分别表示关联网络中的某一个节点或连边。$V=\{V_1,V_2,\cdots,V_n\}$；E 是边的集合，$E=\{E_1,E_2,\cdots,E_n\}$；R 表示资源节点间的关系函数，如关联数目的权重等。

在数字化学习资源关联网络的类型划分上，主要包括用户间关联网络、资源间关联网络以及用户与资源间协作网络这三种不同的网络结构（杨现民等，2024）。其中用户间关联网络主要表现为用户通过各种人际交互操作行为（编辑、评论、收藏、订阅等）进行信息传播，形成个性化的人际关系，如社交网络、知识协同者网络等；资源间关联网络主要表现资源之间通过显性关联和隐性关联方式建立联系（杨现民等，2013），显性关联是基于系统已有的关系类型建立的资源关联，易被用户观察和识别，如文献之间的引用关系、网页之间的超链接关系网络等（高继平等，2015），隐性关联是是从语义上难以通过人工发现，但可以通过数据挖掘技术识别出来潜在资源关联；用户与资源间协作网络则由用户和资源两类节点以及两类节点之间的关联关系组成，如文献资源共享网络等（屈宝强等，2013），具体见表 1。

表 1 数字化学习资源关联网络类型

类型	划分依据	举例
用户间关联网络	用户类型	知识协同者网络、专家网络
	关联类型	社交网络、合作网络
资源间关联网络	资源类型	专利网络、文献网络、主题网络
	关联类型	引证网络、链接网络（显性关联） 共现网络、语义网络（隐性关联）
用户与资源间协作网络		Wiki 网络、文献资源共享网络

2.2.数字化学习资源关联网络核心特征

数字化学习资源关联网络是一个复杂自适应系统，具有复杂自适应系统的非线性、多样性、动态性以及突变性等特征（曹阳，2016）；数字化学习资源关联网络是由资源要素及其关联关系组成的一种网状拓扑结构，具备复杂网络的一般特征；学习资源还是一个生命有机体，具有生命周期（米桥伟等，2023）。综合来看，数字化学习资源关联网络具备以下四方面的核心特征：

一是复杂性。数字化学习资源关联网络的网络结构错综复杂，极其混乱，首先表现为节点和边的数量巨大。本研究使用近十年中文维基百科的数据构建了词条网络，对其节点数进行了初步计算，发现近十年其节点的数量达到了百万级别，并表现出稳定增长的趋势，如图 1 所示。数字化学习资源关联网络的复杂性还表现为节点类型、关联关系、层次的复杂多样，关联网络中的节点可以表示用户行为以及任何类型的资源，如文献、课程、网页、图片等，数字化学习资源关联网络可以是单一类型资源的集合，也可以融合多个资源类型，形成一个多模态数字化学习资源关联网络；在关联类型方面，节点与节点之间的连接有不同的权重、方向，链接网络中表现为超链接关系，引证网络中为文献引用关系，合作网络中为合著关系，语义网络中为语义关联关系等，在关联网络的层次上，单层网络和多层网络则呈现了网络的不同的结构层次，多层关联网络由 2 个或更多单层网络组成，不同的网络层之间是有质的差别的，其节点和连边呈现多属性特征。

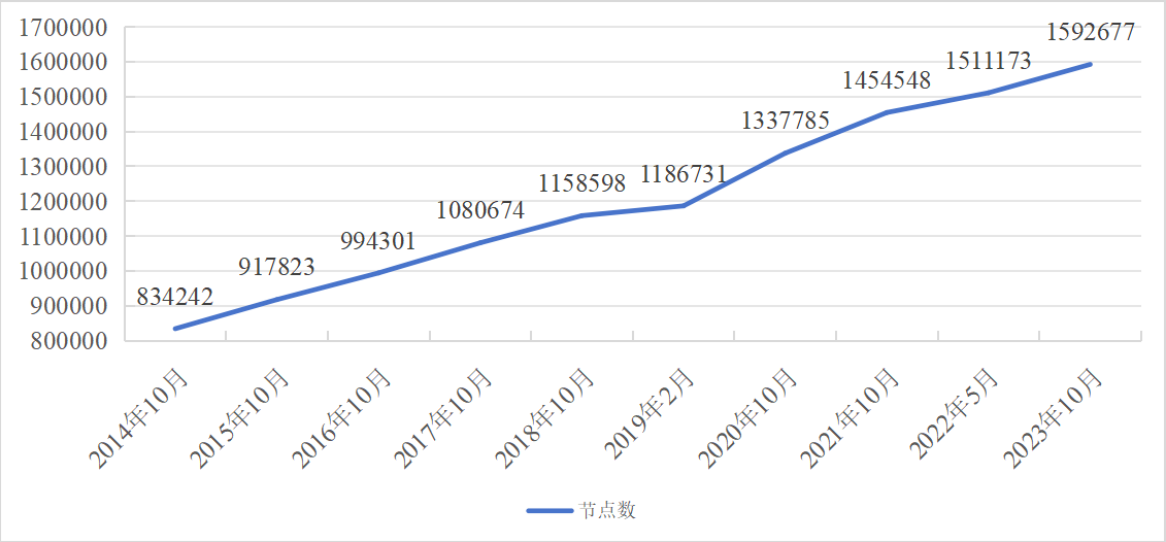


图 1 Wiki 词条网络节点数量增长曲线

二是动态性。数字化学习资源关联网络是一个动态增长的网络，资源在刚被创建时，关联网络上仅有其自身一个节点，通过不断地与其它资源建立关联，逐步扩大自身的网络规模（杨现民等，2011），总是处于进化过程中，稳定只是暂时性现象。数字化学习资源关联网络动态变化的主要形式有节点的增长、删除，边的添加、删除、方向及关联关系的改变（荣莉莉，2008），如 Wiki 网络中每天都有不同的词条被编辑和访问，词条版本迭代频繁，直接影响了该资源节点属性和关联关系等。在数字化学习资源关联网络中，影响其动态演变过程的最突出的特征之一是资源的自组织行为（Berec Vesna, 2017），节点之间的动态竞争和其他相互作用都会改变网络结构特性（Malbor et al, 2020）。

三是适应性。组成复杂适应系统的要素具有主动性和适应性，各要素能够与系统中的其他要素进行非线性交互，彼此影响、相互协调，维持系统的动态发展。在数字化学习资源关联网络中，每一个资源节点都是一个适应性个体，通过重构和改造以优化自身关联结构以适应多样化的学习与不断变化的环境需求，个体和群体的协作、竞争等活动也能推动新资源的生成和原有资源的调整与更新，以适应外界环境的变化。而一些不适应系统的资源将会被淘汰，

如长时间不被更新，滞后于时代发展，无法适应环境变化，从而被其他资源所取代（徐刘杰等，2018）。在开放知识社区中，一些边缘用户（陈耀华，2016）不参与创作，与其他协作者之间的关联关系逐渐消失，最终会成为关联网络中的孤立节点。

四是开放性。互联网为打造开放共享的资源建设新格局赋予全新动能。学习资源关联网络在生长的过程除了会进行内部结构的调整，还会不断与外界资源实体建立关系，属于资源外部结构的持续更新和完善（杨现民，2013）。数字化学习资源关联网络允许外部资源的引入以增加自身的资源数量，其与政治、经济、科技等外部环境通过相互作用不断进行大量的能量、物质、信息的交换（陈丽等，2023），实现教育内部各子系统间、教育与社会间的持续性关联，从而保持和发展系统内部的有序性以及结构的稳定性。

2.3.数字化学习资源关联网络结构及演化规律研究

数字化学习资源关联网站中节点、连边以及整体表现出了不同的网络特性，为揭示这些特性与其呈现出来的规律，近年来不少研究采用复杂网络理论和方法对数字化学习资源关联网络展开了研究。复杂网络有很多拓扑性质参数，如度及度分布、网络直径、聚类系数、平均路径等，这些参数对其拓扑结构和演化特征进行客观描述和分析。举个例子，如果一个复杂网络的度分布符合幂律形式，这个网络就属于无标度网络（Watts et al, 1999）。下面分别介绍对用户间关联网络、资源间关联网络以及用户与资源间协作网络这三种网络的研究现状。

一是用户间关联网络。使用维基百科用户协作编辑数据构建的知识协同者网络是典型的无标度网络，具有明显的小世界效应（Wang et al, 2009），但网络拓扑是非层次结构的，且协作过程中参与者并不完全遵循偏好连接规律（Jürgen et al，2018）。有学者发现海洋专利合作网络在 2000 年至 2001 年不具备复杂网络特征，但 2002 至 2019 年网络的“小世界”和无标度特性逐渐明显（郭建科等，2023）。通过计算出分布、验证 Bow-tie 模型等方面测量了基于订阅关系的用户关联网络的特性,发现了该网络符合无尺度特性，用户之间的联系并不紧密，但有相当多的用户处于网络的核心地位，在信息传播中起到了至关重要的作用（马晨曦等，2018）。

二是资源间关联网络。有学者发现以主题为节点的开放知识网络具有无标度特性和小世界效应（康扬等，2017）：Wiki 词条网络入度服从幂律分布，出度服从广延指数分布，新加入的词条更倾向于与节点度高的词条连接（裘江南等，2017）；还有学者通过连续抓取收集一系列频繁的静态快照来尝试捕获英文 Wiki 网络的动态演化规律，描述了维基图节点、边、集群的时间变化规律等（Buriol et al , 2006）；研究发现，知识网络在其生长初期的度分布并没有很好地遵循幂律形式，通过不断演化生长，其幂律分布特性才逐渐体现出来（徐汉青等，2019）；学科知识流动网络为无标度动态网络，网络的规模和结构会随着时间的推移发生显著的变化（吴江等，2016）；在引证网络中不同论文的参考文献数量存在差异，论文平均参考文献数量呈上升趋势，即出度随时间增长（Biglu, M. H., 2008）。

三是用户与资源间协作网络。目前对于数字化学习资源关联网络的研究主要聚焦在上述两类网络中，对于用户与资源间协作网络较少，基本是从二分网络、多层网络视角出发。如有学者通过用户持续的资源评论行为构建用户与评论之间的二分网络，将用户视为节点，将资源评论关系视为超边，发现该网络具有无标度特性（索琪等，2016）。屈宝强等（2013）基于二分网络构建了包括用户网络和资源网络的文献资源共享网络，发现该网络同时存在“用户富人”和“资源富人”两类节点，节点之间的联系构成了具有二分特性的富人俱乐部。

由此可见，无论是 Wiki 网络、知识网络、协作网络还是社交网络，通常具有无标度特性、小世界效应的特点，且新加入的节点往往倾向于与有影响力的节点连接，呈现出“马太效应”现象。合作关系网络不同于上述网络，但随着时间的推移可能会逐渐具备与之类似的特征。综上所述，运用复杂网络理论与方法能够很好地刻画数字化学习资源关联网络的拓扑结构和网络特性。

3.数字化学习资源关联网络研究面临的问题

数字化学习资源关联网络受个体和资源协同演化的影响，呈现复杂的网络结构，随着资源数量的爆发性增长和资源体系的不断更新，数字化学习资源关联网络变得更加复杂和多样，逐步走向大规模、多模态和动态演变，为揭示深层次的资源进化规律提供了机遇，然而目前存在一些现实问题亟待解决：

3.1.多模态学习资源挖掘问题

随着信息技术的快速发展，多模态资源数据量呈指数级增长，资源范围不断扩展，组织形式复杂多样（Yang et al，2022），已成为数字时代的主要数据资源。截至 2023 年年底，中文维基百科网站中的总页面数已达到 740 多万条（中文维基，2023）、我国的国家智慧教育公共服务平台积累的基础教育阶段资源总量已达到 8 万余条（微言教育，2024），这些资源包含文本、图像、视频、音频、AR/VR 等多种类型，且不同模态资源之间存在复杂的关联和相互作用，尤其是在生成式人工智能技术的影响下多模态学习资源实现了自动生成和更新迭代，其网络结构更为复杂，导致资源内容表征更为困难。

然而目前的数字化学习资源关联网络主要以文本资源为研究对象，并没有融合其他类型的资源（Wu et al，2023）。如何采集这些资源数据并提取其中丰富的语义信息构建结构化的数字化学习资源关联网络，从而更加全面地展示知识、信息和资源的网络层分布结构成为目前要解决的问题之一。

3.2.隐性关联关系提取问题

关联性是复杂网络中的一个重要特征，关联关系提取也是数字化学习资源关联网络中的一个重要研究内容。隐性关联是从语义上难以通过人工发现，但可以通过数据挖掘技术识别出来的潜在资源关联，如借鉴关联规则挖掘算法，可以发现关键词之间的联系、主题词之间的关联。在文本资源关联关系提取中，已有研究从早期基于 CNN 模型和 RNN 模型的提取方法逐渐走向远程监督方法、迁移学习和强化学习方法（Chen et al，2023）。如 Zhang 等（2021）提出了一种多模态图融合（UMGF）方法，通过使用统一的多模态图来表示输入的句子和图像，该图捕获了多模态语义单元（单词和视觉对象）之间的各种语义关系与图像相关的信息。但已有研究并没有使用综合建模的方式实现各个模态信息的直接转换和交流，因此无法捕捉多模态资源关联网络的内部网络结构。不同模态资源之间存在更加复杂的语义关系和相互作用，使得其隐性关联关系更加复杂。针对不同类型的资源探索跨模态信息检索技术，将不同类型的资源建立起语义连接对数字化学习资源关联网络研究来说是一个全新的挑战。

3.3.动态演化规律识别问题

通过对已有研究的梳理，不难发现尽管复杂网络理论与方法提供了许多概念和方法供研究者分析关联网络在微观和宏观两个层面的特性和规律，但这些常用指标大多揭示关联网络在节点和连边上所表现出的静态的结构特性，并不善于揭示网络结构发生了怎样的改变，难以适应动态变化的新趋势。由于网络结构本身比较复杂，难以表示和量化，动态网络时序性、多变的演化过程增加了分析的难度。有学者将某一个时间的关联网络分割成多个静态子网络，每个子网络对应一个时间戳，然后将实体和每个子图的关系投影到时间戳特定的超平面上（Jiang et al，2016）。这种方式并没有真正地考虑时间信息，因此需要构建时间序列下的网络组织结构的动态过程模型，考虑网络结构的历史变化，从而明确地模拟出关联网络的演化过程并预测网络结构和资源行为。

4.数字化学习资源关联网络研究趋势分析

本研究立足数字化学习资源关联网络研究面临的三个现实问题，结合新理论、新方法与新技术在资源关联网络研究中的应用潜能，提出未来数字化学习资源关联网络将呈现三大研究趋势。

4.1.多模异构资源关联网络构建

数字化学习资源关联网络中各资源要素之间的联系频繁而紧密，然而，在解决实际问题时，

只考虑单层网络不能广泛适用于不同类型的资源及资源间的不同关联关系。将资源的类型称之为“模”，资源层之间的关联称之为“构”，数字化学习资源关联网络可分为一模同构网络、多模同构网络、一模同构网络和多模异构网络。多模异构数字化学习资源网络使用可以用不同模态的数据来描述节点特征（Jia et al，2023），融合多种类型资源和多种类型关系的多层网络，具有多源，多粒度，多模态和时空复杂关联的特点。多模异构数字化学习资源网络指融合多种类型资源和多种类型关系的多层网络。现实世界中的许多数字化学习资源关联网络并不是简单的单层网络，举个例子：在 Wiki 网络中包含类别和词条多个类型的资源节点，类别与词条之间是包含关系；每个词条又链接了文字、图片等多个类型的信息，由此可见，Wiki 网络属于多模异质关联网络。多模异构数字化学习资源关联网络具有连边异质性和层间相关性（姚月娇等，2023），可以探索网络层内与层间之间的联系，更完整、准确地分析网络的拓扑结构和网络特性。

人工智能技术是抽取多模态资源实体的主要手段，能够在不同模态数据中建立自适应融合模型，挖掘跨模态数据中的隐藏关系并进行结构化表示，为研究者提供这不仅为研究者提供了一种新的视角来理解和分析复杂的多模态数据集，也大大增强了多模态数据解析的深度和广度。

4.2. 资源关联网络的动态演化过程与规律探究

动态性是数字化学习资源关联网络的核心特征之一。事实上，网络的数字化学习资源关联网络受用户选择、资源自组织、时间等多个因素的影响，其网络规模和拓扑结构会发生较大变化，且演化过程十分复杂。对于其动态演化规律的识别有助于分析资源群体发展趋势、发现教育领域中出现的新兴主题、学科交叉性和多样性及网络中的“涌现”现象等。可以通过引入图同构技术、时间序列分析等方法提出刻画网络拓扑结构变动程度的指标从而对数字化学习资源关联网络的动态演化过程规律给出有效度量（段庆锋等，2023）。然而目前使用图结构的信息探究关联网络演化规律的研究大部分只考虑局部结构，而忽略了网络的全局结构（Zhang et al，2022）。对于一个不断增长的数字化学习资源关联网络，它的增长模式可以通过一系列静态快照来捕捉，将时间信息直接编码到数字化学习资源关联网络中能够有效地反映网络结构随时间的演变情况（Tang et al，2020）。

4.3. 资源关联网络对资源群体进化的影响

分析学习资源内部的网络结构是厘清资源关联关系、探索资源发展过程、预测资源进化趋势的重要方式。基于复杂网络理论和方法，可以更好地识别网络学习资源复杂关联网络的结构样态与分布特征，通过计算资源进化关联网络中的群组数量、规模等，并对不同级别资源群组的进化指标进行比较分析，深入了解资源进化群体的影响因素，进而厘清群组规模、网络密度等变量对资源群体进化的具体影响。举例说明：数字化学习资源关联网络的“鲁棒性”指标（Zhang et al，2015）影响着提高资源群体进化的“健壮性”，随着大量节点和连边的不断加入，各节点之间的关联关系也变得更紧密，网络的抗风险能力也越强，不会由于某个节点或是某条连边的故障、断裂，影响资源的传输和转化，进而影响数字化学习资源进化的效率。

5.总结与展望

数字化学习资源关联网络是将海量、分散且无序的资源作为一个整体进行考察，通过抽象出其中的组成要素和资源之间的关联特征而形成的一种复杂网络，主要包括用户间关联网络、资源间关联网络和用户与资源间协作网络三种类型，具有复杂性、动态性、适应性和开放性，其结构样态呈现出小世界特征、无标度特征等，同时本研究还分析和归纳了数字化学习资源关联网络目前的研究存在的问题以及未来的发展趋势。

随着资源数量的爆发性增长和资源体系的不断更新，数字化学习资源关联网络变得更加复杂和多样，数字化学习资源关联网络目前的研究也存在着新的问题，包括多模态资源挖掘、

隐性关系提取和动态演化规律识别等。未来应更多采取更加多元化的研究方法和技术手段，开展大规模在线开放协同环境下多模态、多层次的数字化学习资源关联网络的研究，以促进数字化学习资源的高质量发展。

参考文献

卜湛,曹杰 & 李慧嘉. (2019).**复杂网络与大数据分析**.北京：清华大学出版社.

曹阳. (2016).复杂自适应理论与网络著作权侵权应对.**浙江学刊**(01), 141-148.

陈丽,郭玉娟 & 张文梅. (2023).“互联网+教育”的世界观：复杂系统观.**中国远程教育**(08), 7-12+24.

陈耀华 & 杨现民. (2016).知识社区用户分类研究——以中文维基百科为例.**现代教育技术**(06), 47-53.

段庆锋,陈红,闫绪娴 & 刘动霞. (2023).基于知识结构突变的学科新兴主题识别研究.**情报学报**(09), 1018-1028.

高继平,丁堃,潘云涛 & 袁军鹏. (2015). 知识关联研究述评.**情报理论与实践**(08),135-140.

郭建科,田冬翠 & 胡凯. (2023).中国海洋产业产学研合作创新网络演化及创新绩效影响因素.**热带地理**(09), 1712-1725.

康扬,李梦琳 & 王晓光. (2017).维基百科词条语义结构研究.**信息资源管理学报**(03), 88-96.

马晨曦,陈兴蜀,郑炳伦,王文贤 & 王海舟. (2018).在线视频分享网络中的复杂网络特性研究.**计算机应用研究**(07), 2059-2062.

米桥伟,杨现民 & 李康康. (2024).如何识别网络学习资源的进化状态——一种基于信息体量的量化表征方法.**现代远程教育研究**(01), 103-112.

裘江南,杨畅, & 李灵.在线知识社区知识系统与社会系统序化规律研究——以 Wikipedia 为例.**情报科学**(5), 8-14.

屈宝强 & 屈丽娟. (2013).基于复杂网络视角的文献资源共享分析框架.**情报科学**(03), 27-30.

索琪 & 郭进利. (2016).在线社交网络资源评论关系超网络演化模型.**系统管理学报**(05), 852-857+867.

万海鹏,余胜泉 & 王琦. (2021).连接式建构：知识建构研究的新取向.**电化教育研究**(10), 12-18+24.

微言教育. (2024).**教育部部长怀进鹏在 2024 世界数字教育大会上的主旨演讲：携手推动数字教育应用、共享与创新**. Retrieved from https://baijiahao.baidu.com/s?id=1789668250651925843&wfr=spider&for=pc.

维基百科.(2023).**维基百科:统计**.Retrieved from https://zh.wikipedia.org/wiki/Wikipedia:%E7%BB%9F%E8%AE%A1.

吴江,金妙 & 陈君.(2016).基金视角下的学科知识流动网络构建与分析.**图书情报工作**(08),79-85.

徐汉青,滕广青,栾宇 & 孙健.(2019).知识网络演化中的结构稳定性与知识涌现.**图书与情报**(01),53-62.

徐刘杰,余胜泉 & 郭瑞.(2018).泛在学习资源进化的动力模型构建.**电化教育研究**(04),52-58.

杨现民 & 余胜泉.(2011).泛在学习环境下的学习资源进化模型构建.**中国电化教育**(9),80-86.

杨现民 & 余胜泉.(2013).开放环境下学习资源内容进化的智能控制研究.**电化教育研究**(09),83-88.

杨现民,杨宇鹏,米桥伟,周丽华 & 李新.(2024).网络学习资源进化：研究进展、难题透视与趋势分析.**现代远程教育研究**(01),101-112.

杨现民,余胜泉 & 张芳.(2013).学习资源动态语义关联的设计与实现.**中国电化教育**(01),70-75.

姚月娇,刘向 & 余博文.(2023).多层图时序专利网络中的发明者影响力演变.**复杂系统与复杂性科学**(03),35-43.

Berec Vesna. (2017). Complexity and dynamics of topological and community structure in c

omplex networks. The European Physical Journal Special Topic, 226,2205-2218.
Biglu, M. H. (2008). The influence of references per paper in the sci to impact factors and the Matthew effect. Scientometrics, 74(3), 453-470.
Buriol, L. S., Castillo, C., Donato, D., Leonardi, S., & Millozzi, S. (2006). Temporal Analysis of the Wikigraph. 2006 IEEE/WIC/ACM International Conference on Web Intelligence, Hong Kong, China, pp. 45-51.
Chen, Y., Ge, X., Yang, S., Hu, L., Li, J., & Zhang, J. (2023). A Survey on Multimodal Knowledge Graphs: Construction, Completion, and Applications. Mathematics, 11(1815).
Jia, X., Jiang, M., Dong, Y., et al. (2023). Multimodal heterogeneous graph attention network. Neural Computing and Applications, 35, 3357–3372.
Jiang, T., Liu, T., Ge, T., Sha, L., Li, S., Chang, B., & Sui, Z. (2016). Encoding temporal information for time-aware link prediction. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 2350–2354.
Jürgen, L., & Lomi, A. (2019). The Network Structure of Successful Collaboration in Wikipedia. Proceeding of 52nd Hawaii International Conference on System Sciences, Grand Waiea, United States, pp. 2622-2631.
Malbor, A., Bruno, R., D., C., Ernesto, E., & James, P., G. Dynamics impose limits to detectability of network structure. New Journal of Physics, (6), 063037.
Tang, X., Yuan, R., Li, Q., Wang, T., & Song, H. (2020). Timespan-aware dynamic knowledge graph embedding by incorporating temporal evolution. IEEE Access, PP (99), 1-1.
Wang, P., Hu, J., Zeng, H. J. & Chen, Z. (2009). Using Wikipedia knowledge to improve text classification. Knowledge & Information Systems, 19(3), 265-281.
Watts, D., Strogatz, S.(1998). Collective dynamics of ‘small-world’ networks. Nature, 393, 440–442.
Wu, Y., Wu, X., Li, J., Zhang, Y., Wang, H., Du, W., He, Z., Liu, J., & Ruan, T. (2023). MMpedia: A Large-Scale Multi-modal Knowledge Graph. In The Semantic Web – ISWC 2023: 22nd International Semantic Web Conference, Athens, Greece, November 6–10, 2023, Proceedings, Part II (pp. 18–37).
Yang, L., Yu, Y., & Wei, Y. (2022). Data-driven artificial intelligence recommendation mechanism in online learning resources. International Journal of Crowd Science, 6(3), 150-157.
Zhang, D., Wei, S., Li, S., Wu, H., Zhu, Q., & Zhou, G. (2021). Multi-modal graph fusion for named entity recognition with targeted visual guidance. In Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, The Eleventh Symposium on Educational Advances in Artificial Intelligence, pp. 14347–14355.
Zhang, H., Fata, E., & Sundaram, S. (2015). A notion of robustness in complex networks. IEEE Transactions on Control of Network Systems, 2(3), 310-320.
Zhang, J., Liang, S., Sheng, Y., & Shao, J. (2022). Temporal knowledge graph representation learning with local and global evolutions. Knowledge-Based Systems, 251, 109234.

基于大单元的特殊教育聋校自适应教学模式实践探究

Exploring the Practice of Unit-Based Special Education Adaptive Learning Model in Deaf Schools

米秀兰[1*]，邓太常[2]
[12]肇庆启聪学校
[*] 531654505@qq.com

【摘要】 随着互联网和人工智能技术的发展以及信息技术的日益普及应用，日新月异的数智技术成为助力听障生学习的利器。大单元教学是落实学科核心素养的重要途径，在聋校高中学科教学中借助数字技术实施大单元教学，通过教师精心提炼的单元主题，根据学情设定单元及各课时教学目标，创设单元学习任务，借助丰富的网络资源，通过智能协作和互动及开展深度学习及实施多元化评价，打破因语言障碍造成的壁垒，有效提升聋校高中学科教学效率，促进学生核心素养的提升。

【关键词】 数智技术；特殊教育；自适应学习；大单元教学

Abstract: *With the rapid development and increasing application of the Internet, the ever-evolving field of intelligent technologies has become a powerful tool for supporting the learning of hearing-impaired students. Unit-based teaching, a vital approach to fostering proficiency in core competencies, is employed in high school settings to enhance instruction for these students. By leveraging digital tools, teachers meticulously select unit themes, establish learning objectives tailored to individual students' needs, create unit-specific learning tasks, utilize abundant online resources, encourage intelligent collaboration and interactive learning, and implement comprehensive assessments. This method effectively overcomes language barriers, significantly improving the efficiency of instruction in high schools for hearing-impaired students, while simultaneously fostering the development of their core literacy skills.*

Keywords: Digital intelligence technology, Special education, Adaptive learning**,** Large unit teaching

1. 问题的提出

随着“智慧教育”时代的到来，教育越来越注重人文发展，强调学生能力的培养。党的二十大报告强调“推进教育数字化，建设全民终身学习的学习型社会、学习型大国”。我国正处于教育数字化变革的关键时期，教育数字化变革是教育体系高质量发展的重要途径。生成性教学是在弹性预设基础上，师生充分交互，不断调整教学活动和行为，共同建构并形成新的信息、资源的动态过程，以实现教学目标和创生附加价值（谢幼如,杨阳,柏晶,李伟,郭琳科,倪妙珊,2016）。听障生因学习活动缺少声音刺激和言语参与，第一与第二信号系统脱节，以直观形象思维为主，认知特点是“以目代耳”，对长篇文本有畏难情绪。如何借助多种数字技术融合培养学生自适应学习能力，落实核心素养的养成，有着举足轻重的意义。

2. 相关文献述评

2.1. 多技术支持的特殊教育课堂教学研究现状

当前信息技术环境满足残疾人学习需求的状态界定为“数字化无障碍学习”。当前随着数智技术的发展，正从数字化、网络化向智能化过渡，从集成化平台到基于大数据、移动互联网及云计算的智慧学习平台过渡（张俊玲，安瑞清，郭楠，蒋文桐，周华丽，2016）。因此多技术支持的特殊教育课堂教学是推进教育数字化的重要行径。

2.2. 自适应学习研究现状

大数据支持下的自适应学习为终身教育提供了符合要求的方案和任务，使不同需求的学习者可根据自身需要来选择准确而有针对性的资源，及时调整学习状态（宋原，2021）。但如何面向特殊教育聋校开展数据支持下的自适应学习仍需探索。

2.3. 大单元教学与技术融合研究现状

教育大数据的构建与应用，让教师从依靠经验到依靠数据开展教学活动，助力开展更加高效、精准、个性、公平的教学（柯清超，2021）；在大单元教学中，教师可以结合单元主题、内容，积极开发学习资源，为学生搭建合适的学习支架（马泽汉，2023）；在与技术深度融合的过程中，教学内容更能激发创造力，数据分析针对性强，结果呈现更加丰富直观（张倩和邱道学，2023）。综上所述，部分学者探析了高中相关学科大单元教学及信息技术支持下的教学活动，但尚未系统性地实践探索出适用于特殊教育聋校大单元教学的教学模式。

3. 多技术支持的异步云网端自适应教学模式的构建

3.1. 模式构建

以核心素养为纲，以大单元（大概念）教学为领，在建构主义理论支持下，秉持深度学习、自适应学习、探究式学习、体验性学习、个性化学习理念，开展多种教学活动丰富学生情感体验，通过智能化平台和基于特教课堂特征编写基于卷积神经网络、图像处理、人工智能和模式识别等数智技术，在实现持续性的学习评价基础上，探索出面向高中听障生的多技术融合支持下的异步云网端自适应学习教学模式，具体如图 1 所示。

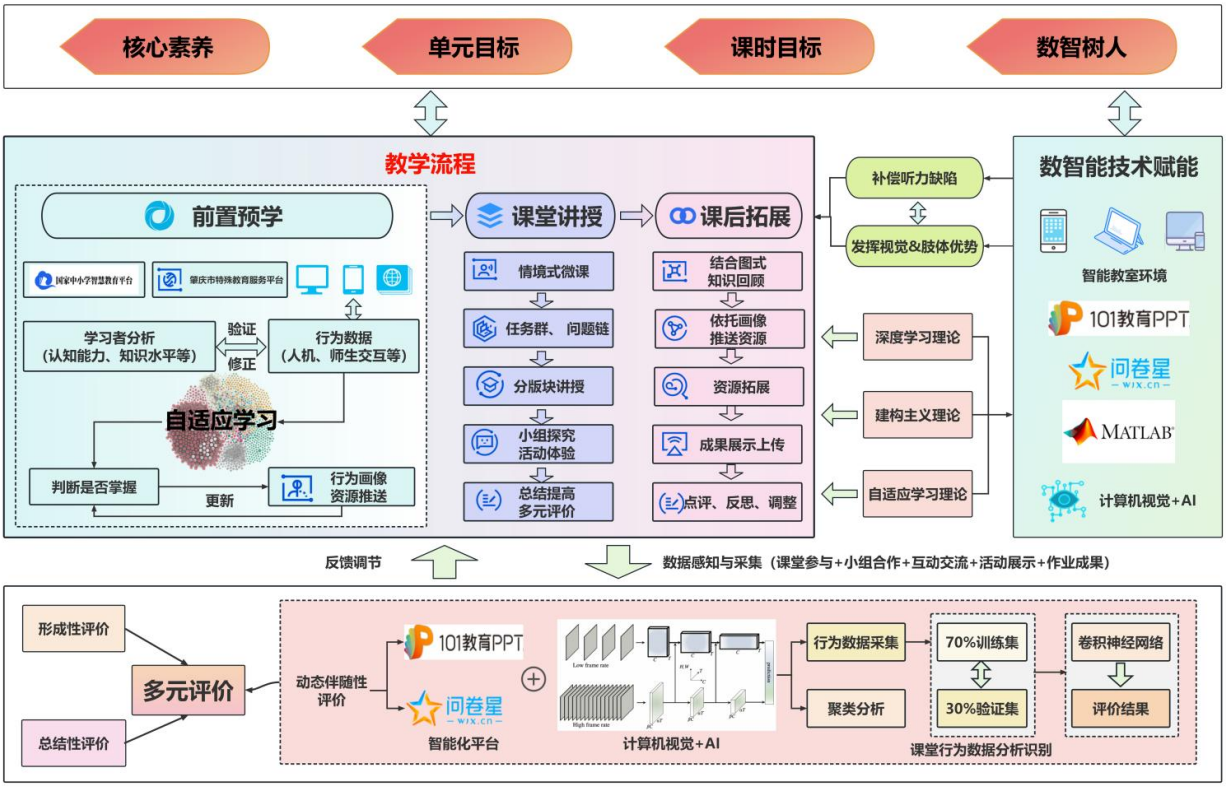


图 1 多技术融合支持下的异步云网端自适应教学模式

该模式重点指向学生核心素养的培养与提高。听障生因学习活动缺少声音刺激和言语参与，因此特别需要直观教学借助多技术融合把信息符号变得直观化和可交互化，调动视觉器官充分观察思考，突破学习障碍。从立德树人的育人目标出发，依据新课标要求，基于数字化教育学习环境，依托大单元、任务群与问题链的设计，借助线上智能平台及资源，利用情境式微课、课件、智能交互教学资源、数字教材等数字资源，运用问卷星、101 教育 PPT、平板电脑、MATLAB、Excel 和 python 等智能工具支持开展学习各种活动，在教学环节中实现了多技术融合的自适应学习与基于智能数智技术的数据分析助力精准辅导，通过智能化平台和计算机视觉等基于卷积神经网络、图像处理、人工智能和模式识别等技术的综合，对学生的伴随性学习行为数据以及时空特征进行采集分析，以评价数据为教师因材施教提供决策支持，并促使学生对此进行反思、调整，从而助力智能化课堂教学全过程评价。

3.2. 实施过程

在智能化学习环境下，分“前置性预学”“课堂讲授”“课后拓展”三个教学流程进行。

“前置预学”阶段：教师利用问卷星进行学情调查，根据数据分析学情，分析学习者的认知能力及先备知识水平，根据学生差异设计多种预习资源包，通过平台发布资源，学生根据资源包自主学习，并完成导学案。通过人机、师生交互等分析学生行为数据，进行画像，提供自适应学习环境及资源，并鼓励学生线上获取学习资源。

“课堂讲授”阶段：围绕核心素养，重构单元内容，借助数字技术建构知识间的内在关系，促进教学内容结构化。通过创设基于生活的真实任务情境作为大任务，围绕大任务设计相关子任务及问题链展开学习，并设计各种评价任务。抓住教学要点，分板块组织教学，并以任务驱动分小组探究，通过开展多种丰富教学活动加深体验，全过程实施多元评价。

“课后拓展”阶段：学生通过画思维导图进行知识回顾复习，教师依据画像推送资源，学生课后排演、展示课本剧，并上传成果，教师组织大家点评、总结、反思、拔高。

在整个教学过程中，通过智慧平台终端等采集伴随性数据，进行 K-means 聚类分析与行为画像，实现自适应学习辅导。最后通过文档、图表、视频和数据等形式记录学生表现，实现教学评一体化。依据评价结果优化教学策略，改进教学过程，或自动推送相关学习资源，促进学生开展自适应学习。

4. 多技术支持的异步云网端自适应教学模式的应用

4.1. 教学设计

以人教版高中语文必修 3 第一单元教学为例，具体阐述如何基于数智技术的聋校大单元教学模式进行设计与实施大单元教学活动，助力听障生核心素养的培养与发展。

在“品读经典小说，洞察社会人生”大概念的统摄下，重点指向“语言建构与运用”“审美鉴赏与创造”两大学科核心素养，创设了“校园课本剧表演大赛”这一真实任务情景，并围绕这一大任务设计了前置性学习、梳理小说情节、赏析人物形象、分析环境描写、课本剧展示 5 个子任务及“《红楼梦》有哪些主要人物？”等 29 个问题链。

围绕单元主题，分预习、授课、体验、反馈四个环节展开学习。在前置性学习阶段借助数智技术实施学情精准分析，提供不同要求的导学案，引导学生自主预习，通过“先学后教”及“行为画像”来培养学生的自适应学习能力；课堂授课阶段借助情景微课抓住小说的三要素进行分板块讲授；活动体验阶段主要开展观看经典影片、排演课本剧、写人物评析等学习活动；最后通过在线测试、多模态数据分析、作品展示等手段进行评价反馈。

4.2. 教学评价

践行新课标提倡评价的“教—学—评”一体化，数字化下的聋校课堂多元化评价不仅是要关注学生学习目标的达成情况，更要重视核心素养的发展。除了设计“读、赏、品、写、演、

测试”等各种评价外，还开展教师评价、生生评价等。过程性评价以学生课堂表现、解决问题的过程和方法及课后作业完成为评价依据，过程性评价和形成性评价相结合，老师和学生共同参与评价。如表 1 所示。形成性评价针对学生在课堂上的表现以及对知识掌握、运用，结合学生课后练习中是否对本节课内容的迁移运用，对学生做出不同的评价。

表 1 多元评价表

评价要素	评价运用场景	评价方式	评价主体
学习态度	课堂	口语加手语	教师
解决问题的过程和方法	课堂	课堂表现及互动	教师、学生
能力迁移	课后	课后作业	教师、学生
人工智能辅助评价	教学全过程	计算机视觉+AI	计算机

5. 研究结论

经过一年来的研究论证，多技术融合的自适应学习模式在数智技术支持下助力聋校高中学科大单元教学优势主要体现在以下几点。

5.1. 借助数智技术实现精准教学

实施精准教学重点在于了解学情，教师课前需进行学情分析，以便开展精准教学。借助数字化教学平台关注学生差异，提供差异化的教学内容，落实分层教学，实施精准教学，促进学生的个性化学习。

5.2. 数智支持下开展多种教学活动

依据听障高中生认知与思维等特点，在数字化环境下知识可视化，利用可视化工具绘制知识结构体系，引导学生调动视觉器官充分观察思考突破学习障碍,帮助学生形成清晰的知识结构。改进传统教学方法，通过多种渠道引导学生参与多种体验活动，培养学生核心素养。

5.3. 全方位多元化评价

在课堂行为数据分析识别的实现过程中，通过智慧平台终端与计算机视觉+AI 采集进行行为数据采集与聚类分析，将 70%数据作为训练集，30%作为验证集，最终通过特定卷积神经网络输出结果。在整个教学过程中，通过对学生课堂参与、小组合作、活动展示与作业成果等课堂伴随性数据感知与采集，结合形成性评价与总结性评价最终构建出多元评价体系。

综上所述，随着互联网和人工智能技术的飞速发展，在聋校高中学科教学中借助数字技术实施大单元教学，借助数字资源及平台，通过凝练单元主题，构建学习板块，创设真实的任务情景，驱动听障学生在数字化环境下开展自适应学习，促进学生核心素养的提升。

参考文献

谢幼如,杨阳,柏晶,李伟,郭琳科 倪妙珊.(2016).面向生成的智慧学习环境构建与应用——以电子书包为例.华南师范大学学报(自然科学版),(01),126-132.
张俊玲,安瑞清,郭楠,蒋文桐 周华丽.(2022).特殊教育教师信息化教学能力调查研究——基于北京市 23 所特殊教育学校的调查数据.现代特殊教育,(16),9-16.
宋原.(2021).大数据支撑下的个性化自适应学习.中国成人教育,(18),3-6.
马泽汉.(2023).高中语文大单元教学策略的实践与思考.基础教育论坛,(13),27-29.
柯清超.(2021).技术赋能教育新图景：数据驱动的教育生态.中小学数字化教学,(12),1.
张倩和邱道学.(2023).语文学科教学与技术深度融合的价值探究.中小学数字化教学,(02),68-73.

核心素养导向的个性化作业设计与应用研究

Research on the Design and Application of Personalized Homework Oriented by Core Literacy

朱梦雨，魏雪峰*，孙晓涵，张雪，朱晓颖
鲁东大学 初等教育学院
*xuefengwei99@163.com

【摘要】 个性化作业因注重学生的个人差异成为未来作业发展的重要趋势。然而，学科核心素养作为当代基础教育最重要的指向，没有在个性化作业中得到充分的发展。如何基于学生的核心素养水平为学生推荐个性化作业，使学生在个性化作业中提升核心素养水平成为亟待解决的问题。本研究以信息科技中的计算思维核心素养为例，利用知识图谱技术、认知诊断以及余弦相似度算法等，经过作业设计、数据层搭建、认知诊断和个性化作业推荐四个步骤，建立了学习者子图、计算思维子图、作业子图和知识点子图，实现了面向计算思维培养的个性化作业推荐。

【关键词】 知识图谱；计算思维；个性化作业

Abstract: *Personalized assignment is a vital trend in future development due to its focus on students' differences. However, in contemporary basic education, subject core literacy is not fully developed within personalized assignment. Recommending personalized assignment based on students' core literacy is crucial for improving their skills. For instance, in computational thinking, this study utilizes knowledge graph technology, cognitive diagnosis, and collaborative filtering algorithms. It establishes learner, computational thinking, operation, and knowledge point subgraphs through operation design, data layer construction, cognitive diagnosis, and personalized operation recommendation. This approach ensures effective personalized operation recommendations for computational thinking training.*

Keywords: Knowledge graph, Computational thinking, Personalized *assignment*

基金项目：本文系山东省人文社会科学课题（智库重点项目）“人工智能支持下的小学个性化作业设计与应用研究”（编号：2023-zkzd-014）的阶段性研究成果。

2021 年 7 月，中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》（简称“双减”），明确提出“减轻学生过重作业负担，提高作业设计质量。”数智时代背景下，数智技术成为推动教育“减负增效”的关键拉力，关注教育与教学的数智化连通已经成为全球性话题（李小娟等，2023）。在这一背景下，个性化作业逐渐成为教育研究的热点。骈扬利用知识图谱、认知诊断等新一代智能技术，设计了依托“智慧学伴”平台探索出的“四面—一线—四环”有效作业实践模式（骈扬、余胜泉和李晟，2023）。李小娟，刘清堂等构建了“知识漏洞集—学习路径链—学科素养栈”数据流转的服务模型（李小娟等，2023）。针对个性化作业的研究虽然已有不少，但少有研究将核心素养作为个性化作业“个性”考量的依据。如何基于学生的核心素养水平为学生推荐个性化作业，并使学生在

个性化作业中提升核心素养水平成为亟待解决的问题。本研究以小学信息科技四大核心素养之一的计算思维为例，探讨如何将素养融入个性化作业的设计与实施的过程中。

1.面向计算思维培养的个性化作业总体设计

面向计算思维培养的个性化作业实现要经过作业设计、数据层搭建、认知诊断和个性化作业推荐四个步骤，如图 1 所示。其中作业设计是个性化作业实施的前提；数据层是个性化作业实施的基础，为后续的认知诊断提供数据支持；认知诊断是个性化关键，只有诊断出学生的知识点掌握程度和计算思维水平，才能根据学生水平为学生匹配个性化作业；个性化作业推荐是个性化作业实施的最终目的，根据认知诊断结果从作业题库中为学生推荐作业是本研究的最终效果。

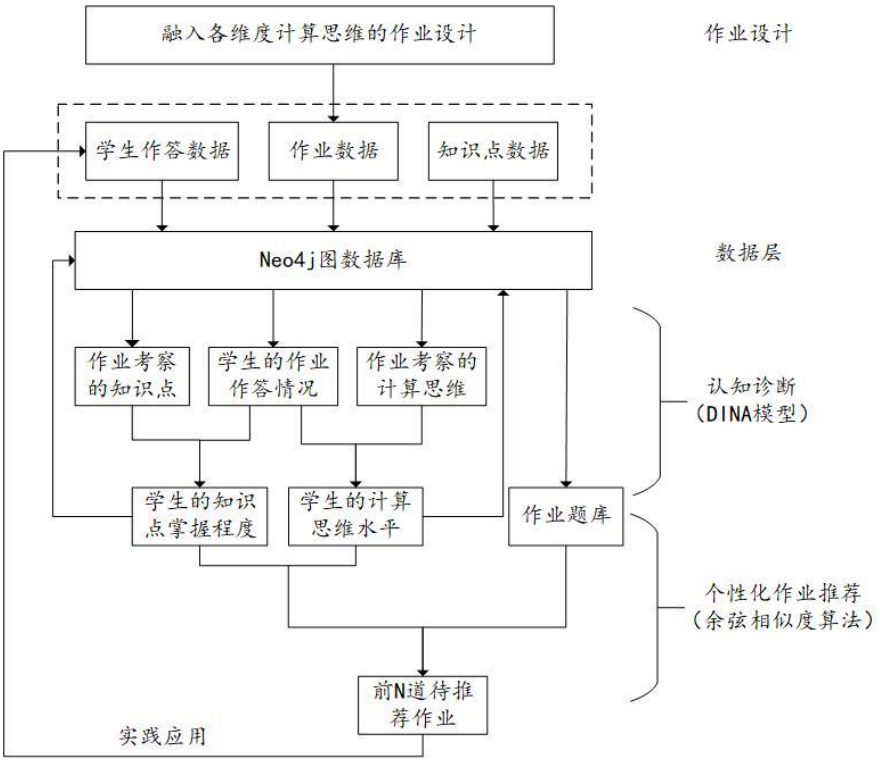


图 1 个性化作业设计总体框架

2.面向计算思维培养的作业设计

实现根据计算思维水平为学生推荐个性化作业，收集学生关于计算思维方面的历史数据以及为学生推荐融入计算思维培养的作业题目是必不可少的。因此，根据课本知识点设计融入计算思维培养的作业题目是实现个性化作业的第一步。本研究参考张屹等人依据新课标对计算思维的维度的解读，将计算思维划分为抽象、分解、建模、算法设计和评估五个维度（张屹等，2024）。针对课本中的各个知识点设计融入了考察计算思维不同维度的作业题目。

3.个性化作业知识图谱搭建

3.1. 个性化作业知识图谱总体设计

个性化作业知识图谱是小学信息科技个性化作业推荐系统的数据层，主要包括系统所需要的学习者、知识点、作业题目及计算思维度四个部分的属性数据及关系，存储在学习者子图、知识点子图、作业子图及计算思维子图中。

3.2. 知识图谱子图构建

3.2.1. 知识点子图

知识点子图的构建主要包括实体抽取和关系抽取两大部分，即将知识点命名并梳理知识点与知识点之间的关系。前驱后继关系是知识点间关系里比较重要的一个关系类型。以 3D 打印单元为例，将实体抽取和关系抽取获得的数据导入 neo4j 图数据库中，获得的知识点子图如图 2 所示。通过知识点子图可以快速找到本节课考察的知识点的前驱知识点，继而查找到学生前驱知识点的相关数据，得到学生对本节知识的前驱知识掌握程度。



图 2 3D 打印单元知识点子图

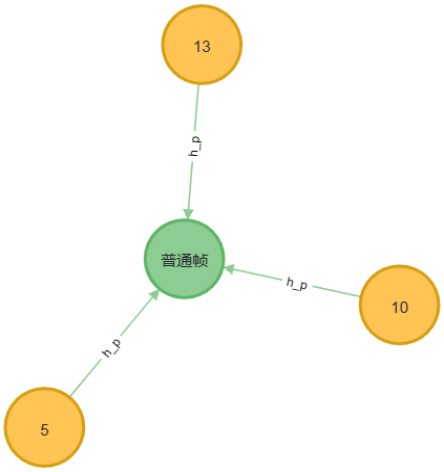


图 3 作业节点与知识点节点的联系（部分）

3.2.2. 作业子图

作业内容为教师根据知识点设计或改编的题目，每道作业题目都有明确考察的知识点和计算思维度。每个作业为一个实体节点，作业和知识点之间有边表示该作业考察了此知识点，如图 3 所示，作业号为 5、10、13 的作业考察了普通帧的知识点。作业和计算思维之间有边表示该作业考察了此维度的计算思维。

3.2.3. 学习者子图

学习者子图主要包括学习者姓名、班级等属性和学习者作答过的作业数据。学习者节点与作业节点之间有边即为该学习者做过该作业题目，边的属性为该学习者做该试题的得分情况，如图 4 所示。学习者二完成了作业号为 10、11、12 的作业，其中与 10、11 号作业的边为 1，表示学习者二答对了这两道作业题目，而与作业号为 12 的作业的边为 0 表示学习者二答错了该道作业题目。学习者节点与作业节点之间边的属性是认知诊断部分所需的数据来源。

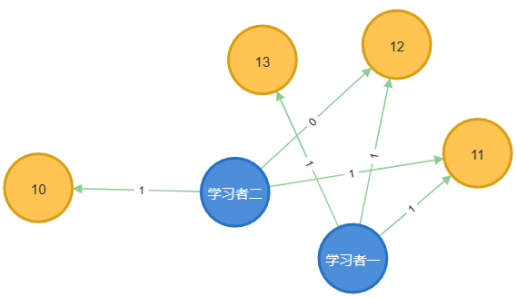


图 4 学习者子图（部分）

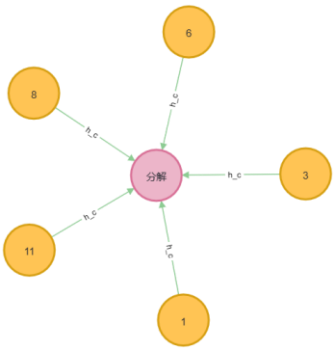


图 5 计算思维节点与作业节点的联系(部分)

3.2.4. 计算思维子图

计算思维子图主要有计算思维节点，主要包括英国南安普顿大学的伍拉德博士和辛西娅·塞尔比博士共同提出计算思维包括五个要素：抽象、分解、算法思维、评估和概括。计算思维节点和作业试题有边表示考察该计算思维度的作业有此作业。形成的计算思维节点及其

与作业题目的关系如图 5 所示，考察到分解这个计算思维维度的作业有作业号为 1、3、6、8、11。

4.利用 DINA 模型开展认知诊断

通过认知诊断可以得知学习者当下认知水平，分析出什么才是符合学习者的作业，完成下一步的个性化推荐。DINA 模型是认知诊断中常用的一种模型，相比于先前的其他认知诊断模型，该模型在实践方面表现良好，具有准确率高、表达式简单、操作方便等优点，目前已被广泛的应用于各学科的认知诊断实践中（李振，2022）。本研究用江西师范大学涂东坡教授，开发的在线认知诊断平台进行数据分析。将学生的作答数据及 Q 矩阵上传至该认知诊断平台，选用 DINA 模型诊断出学习者的知识掌握程度和计算思维水平。

5.个性化作业推荐

余弦相似度是距离度量公式的一种，它通过计算两个向量的夹角余弦值来评估他们的相似度。本研究使用余弦相似度，调用 Python 中 Sklearn 库，计算学生的知识掌握程度和计算思维水平与作业中的知识点考察和计算思维考察的距离，为学生推荐和自身水平最近的作业题目，由此实现个性化作业的推荐。学生在完成作业后，为其推送相关答案和解析资源，帮助学生修订与纠正作业习题，巩固知识，提升素养。

6.总结

本研究以信息科技中的计算思维核心素养为例，利用知识图谱技术、认知诊断以及余弦相似度算法等，经过作业设计、数据层搭建、认知诊断和个性化作业推荐四个步骤，设计了一册课本的面向计算思维培养的作业题目，构建了作业子图、知识点子图、学习者子图和计算思维子图，利用江西师范大学搭建的认知诊断平台完成了对学生知识掌握程度和计算思维水平的认知诊断，最终通过余弦相似度算法根据诊断结果完成了对学生个性化作业的推荐，实现了面向计算思维培养的个性化作业推荐。

参考文献

中共中央办公厅，国务院办公厅（2021）。关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见。

中华人民共和国教育部（2022）。义务教育信息科技课程标准(2022 年版)。北京：北京师范大学出版社。

李小娟、刘清堂、王云豪等（2023）。教育数字化赋能差异化作业设计的行动逻辑和服务模型。电化教育研究，44(08)，96-102。

李振、周东岱(2022)。基于学科知识图谱的智能化认知诊断评估方法。现代教育技术，32(11)，118-126。

张屹、陈邓康、付卫东等（2024）。基于新课标的中小学生计算思维量表构建研究。电化教育研究，90-98。

骈扬、余胜泉和李晟（2023）。基于认知诊断的有效作业精准化设计。中国远程教育，43(08)，13-24。

應用生成式人工智慧增強合作問題生成學習系統於國小社會課程之研究

Examining the Effeteness of Generative AI-Enhanced Co-Question Generation Learning System to Enhance Learning Performance, Motivation, and Critical Thinking in Elementary Social Studies Education

洪暉鈞[1]，洪丞威[2*]
[1] 中央大學網路學習科技研究所
[2] 中央大學網路學習科技研究所
[*] hch@cl.ncu.edu.tw
* rexhorng@gmail.com

【摘要】 社會研究在基礎教育中至關重要。本研究探討生成式人工智慧增強共同問題生成學習系統來開發一個虛擬同伴系統，旨在增強社會研究教育中的學習成效與學習動機。 該系統允許學生透過平板電腦拍攝捕捉教科書內容，從圖像中產生問題，並結合遊戲化元素進行互動學習。三十七名四至六年級學生使用平板電腦參加了為期 10 週的社會學習課程。研究採用定量和定性資料收集方法，以確保全面了解學生的觀點。採用問卷調查來評估學生對系統功能的看法以及學習方法對學習動機的潛在影響。結果表明，理解力和動機都有顯著提高。問題生成和遊戲化元素的整合增強了主動學習。學生們強調其在不斷發展的教育科技領域重塑傳統學習方法的變革潛力。

【關鍵字】 生成式 AI；問題生成；社會科教育；學習動機

Abstract: *Social studies are pivotal in elementary education. This study explores Generative AI-Enhanced Co-Question Generation Learning System to develop a virtual companion system, aiming to enhance critical thinking and self-directed learning in social studies education. The system allows students to capture textbook content through tablet cameras, generating questions from images and incorporating gamification elements for interactive learning. thirty-seven in grades 4-6 students participated using tablets, the system during a 10-week social study course. The study employed both quantitative and qualitative data collection methods to ensure a comprehensive understanding of students' perspectives. A questionnaire survey was utilized to evaluate students' views on the system's functionality and the potential impact of the learning method on learning motivation. The results indicate that significant improvements were observed in comprehension and motivation. The integration of question generation and gamification elements enhanced active learning. Students emphasize its transformative potential in reshaping traditional learning methods in the ever-evolving landscape of educational technology.*

Keywords: Generative AI, Question Generation, Social Studies Education, Learning Motivation

1. 前言

在基礎教育中，社會研究扮演關鍵角色。科技成為應對這項挑戰的重要因素，以學生為中心的教學方法有效的訓練學生自主思考能力。另一方面，生成式人工智慧（GAI）利用深度學習技術創建組合新內容，引起了研究人員和教育工作者的興趣。先前的研究著重於教師的觀點，提供了對人工智慧在教育領域的全面見解(Jeon & Lee, 2023; Kim, 2023; Velander, Taiye, Otero, & Milrad, 2023; Yau et al., 2023)。儘管生成式人工智慧（GAI）取得了重大進展，但在系統設計和教學法的實際整合方面仍然存在差距。這項研究透過採用大型語言模型訓練

專屬於社會科的自動問題生成學習模型來彌補這一差距，並使用自然語言處理增強互動，旨在提高學生在社會研究中的批判性思考、學習動機和自主學習(Hwang & Chen, 2023)。

本研究重點在於採用大型語言模型訓練專屬於社會科的自動問題生成學習模型。利用生成式模型的創造力，提供一個可編譯的題目與選項，除了引導學生合作共創題目促進學生批判思考的能力之外，還彌補了生成式模型準確性不足的缺點。除了建立一個讓學生創造問題的模型之外，生成式模型的創造潛力還可以融入更廣泛的學習環境中，學生可以在本研究開發之系統上挑戰同儕與生成式模型創造之題目。此外，本研究還引入虛擬同伴機制，以提高學生的學習動機。在遊戲化的系統結構支持下，利用排行榜、等級與競爭機制，旨在增強學生在社會研究領域的動機。由於先前關於將生成式人工智慧整合到基礎社會研究教育中自動問題生成的研究有限，本研究開創了生成式人工智慧作為多項選擇問題生成平台的發展。其目標是無縫整合小學社會研究內容，與教育工作者對教學應用的觀點一致。本研究的研究問題概述如下：

(1) 生成式 AI 輔助學生建立問題這種學習模式如何影響學生的學業成績？

2. 文獻探討

2.1. 學習同伴

在過去的二十年中，教育虛擬角色有了顯著的發展，其特點是具有類人屬性的電腦模擬，旨在透過充當虛擬參與者來增強個人學習環境(Chan & Chou, 1995)。具體來說，虛擬動物同伴通常以虛擬寵物的形式出現，旨在豐富兒童的學習體驗。透過與這些虛擬寵物的互動，可以培養孩子的情感，轉化為學習的動力(Chen, Liao, Chien, & Chan, 2011)。

先前的研究表明，虛擬寵物的引入會對動機和學習成就產生積極影響(Chen et al., 2011)。Chen 等人(2011)發現，國小學生在虛擬寵物的陪伴下表現出更積極的學習態度和更大的動力。虛擬寵物的屬性、能力等個人化特徵，讓學生可以選擇學習內容，並增強學習動機。此外虛擬寵物系統中的競爭與學習成就之間也有著密切的關係。根據 Wu 等人(2010)的研究，虛擬寵物系統能夠透過遊戲化學習的方式，鼓勵學生更積極參與學習活動，提高學習的效率和成效。這項研究中，虛擬寵物系統中的學習進度和表現記錄也能提供學生更清晰的學習軌跡，幫助他們自我評估和改進學習策略。

總上所述，本研究參考 Chen 等人(2011)的虛擬寵物設計模式，包含三種模式：養成、學習、遊戲，在養成模式，學生需要照顧寵物；在學習階段，學生需要創造問題任務來獲得食物獎勵。在遊戲階段，學生可以參加娛樂性的遊戲，挑戰同儕製作題目並抽取獎品。並在本系統結合儀表板的功能增強了學生對自身學習情況的了解和視覺化學習狀態，使學習過程更加透明化，同時提供了即時的學習反饋。

2.2. 自動問題生成

根據 Hwang(2014)的說法，人工智慧教育 (AIED) 的重要應用是“實現個人化學習”。儘管存在擔憂和爭議，但在教育環境中使用生成式模型（Generative Artificial Intelligences , GAI）正在成為技術增強學習的新趨勢(Dehouche, 2021; Dwivedi et al., 2023)。Eysenbach(2023)就提出了 GAI 的學習的潛力：提升學習者的創造力、批判性思維和解決問題能力。

在 GAI 的領域中，自動問題生成（AQG）在教育評估中發揮著重要作用。考試形式的問題是一種基本的教育工具，具有多種用途。除了作為評估的工具外，問題還有可能影響學生的學習。根據 Thalheimer(2003)的說法，使用問題的一些好處是：（1）提供練習從記憶中檢索信息的機會；（2）向學習者提供關於他們的誤解的反饋；（3）將學習者的注意力集中在重要的學習材料上；（4）通過重複核心概念來強化學習；（5）激勵學習者參與學習活動（例

如閱讀和討論）。但人工製作問題需要大量的人力、時間和成本，答案的評估也是一項耗時的工作(Divate & Salgaonkar, 2017)。

相反，學生生成內容（SGC）被認為是促進參與和更高質量學習成果的有效且有吸引力的方法(Arruabarrena, Sánchez, Blanco, Vadillo, & Usandizaga, 2019; Doyle, Buckley, & McCarthy, 2021)。將提出問題的責任交給給學生，並通過允許學生創造問題來使其他學生回答的問題的學習模式已在教育領域有著顯著的成效(Heron & Lerpiniere, 2013; Jones, 2019)。學生生成問題的學習模式讓學生參與反思自身所想所見的過程，並將這些資訊轉化為對個人有意義的形式。此外，在參與問題生成學習任務時，學生會使用個人知識與教材之間建立有意義的連繫並採取各種認知和後設認知策略的狀態(Song, 2016; Yu, Liu, & Chan, 2005)。根據以上研究成果，本研究整合了 Heron 等人(2013)與 Jones(2019)的教學設計，在遊戲化的學習機制中將學生生成內容（SGC）與人工智慧的自動問題生成（AQG）結合進行協作學習。本研究的創新方法旨在解決以往研究中發現的局限性，最終自行開發一個線上問題生成平台。讓學生有權利修改和改進人工智慧生成的問題，即 "共同問題生成"，有望培養學生的批判性思維和創新能力。

3. 方法

3.1. 參與者與背景

來自台灣北部一所小學的 37 名學生配備了平板電腦，他們分別在四年級、五年級以及六年級三個班級，四年級一共 11 人、五年級一共 8 人、六年級一共 18 人，性別比例為男生 48.65%、女生 51.35%，他們先前未使用過平板電腦進行社會課學習，在 2023 年為期 10 週的社會課程中參與了系統測試。本研究探討了使用生成式人工智慧為小學生整合社會科學協作式問題生成系統，生成模型根據拍攝之圖像生成問題，學生修改並創建問題以供同儕練習。系統包含虛擬寵物等級和戰鬥系統等遊戲化元素。測驗涉及 37 名小學四至六年級學生，他們使用平板電腦，在為期 10 週的社會課程中透過虛擬寵物平台進行自我調節學習。

3.2. 實驗流程

實驗前，學生們收到系統介紹、問題產生概念指導和系統說明。在實驗過程中，學生使用系統進行學習，完成了三個學習單元的活動，他們拍攝了課本和相關閱讀教材的圖片，並與人工智慧合作在課堂上生成問題。隨後是同儕回答和學習，持續約一個小時。

實驗時間表如下（圖 1）。第一周指導學生如何使用問題產生系統，對學生進行問卷的前測。從第二週到第四周，重點是課程的第 1 單元。 學生先進行 1 單元知識的前測，然後參加講座。講座結束後，學生使用平板電腦捕捉有意義的 1 單元教材圖片。然後，他們根據圖片產生問題，並與同伴一起練習。在課堂上，學生使用人工智慧輔助生成問題，每次生成問題的時間為 20 分鐘。他們必須在這段時間內產生至少兩個高品質的問題，教師會即時檢查他們的使用情況。剩下的時間則分配給學生，讓他們透過在系統中練習問題來挑戰同儕。練習結束後，學生進行 1 單元知識的後測。這樣的循環在第四週至第五週的第 2 單元和第六週至第十週的第 3 單元中重複進行。最後一周，即第 10 週，以學習動機為重點的後測。

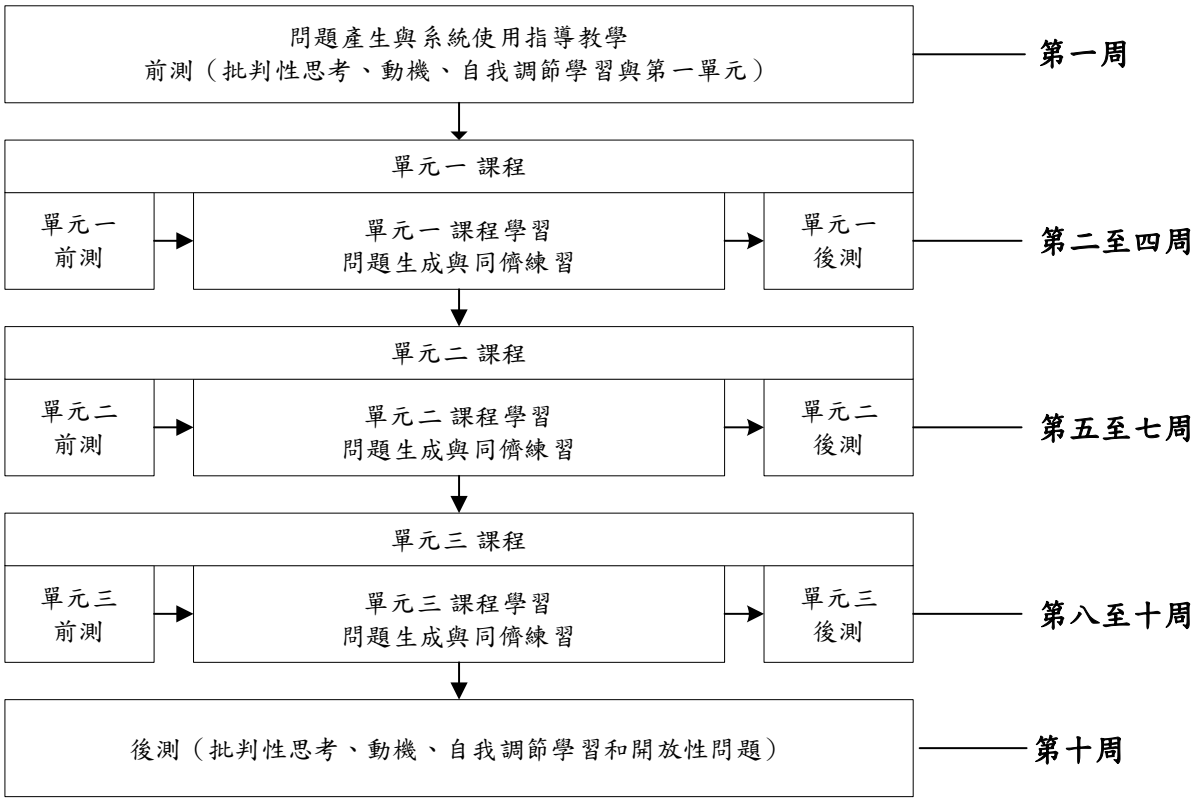


圖 1 實驗流程圖

4. 生成式人工智慧合作問題生成學習系統的設計

4.1. 功能概述

本研究在線上學習平台中引入了人工智慧增強的問題生成學習機制，旨在促進學生的問題生成和自主學習。系統前端伺服器使用 Flask 框架開發，並與 MySQL 資料庫整合。後台主要有兩大功能：問題產生和基於挑戰的問答。在學習場景中，學生利用問題產生系統上傳圖片，根據教科書中的重點產生問題。 在合作編輯問題的過程中，學習者可以參與以挑戰為基礎的問答活動，在規定時間內挑戰同儕的問題。

4.2. 問題生成階段：用於共同創建社會課問題的人工智慧增強型生成平台

在初始階段，系統的開發目的是幫助學生完成社會研究課程的問題創建任務。系統利用一個自建的問題儲存平台，將學生產生的問題儲存在 MySQL 資料庫中。首先，點選”自動生成”按鈕，Python Flask 後端接收圖片訊息，並利用 Google 雲端視覺 API 模型從上傳的學生圖片中辨識文字。隨後，自然語言處理演算法會從辨識出的文字中提取關鍵字，並將其輸入 GPT3.5 模型。系統配置了一個 Prompt 來產生問題，並將問題返回到黑板頁面。學生可以自由修改系統產生的內容，包括問題本身、選項一、選項二、選項三、選項四、正確答案，如圖 2 所示。學生需要進行深思熟慮，檢查生成模型提供的問題是否準確。

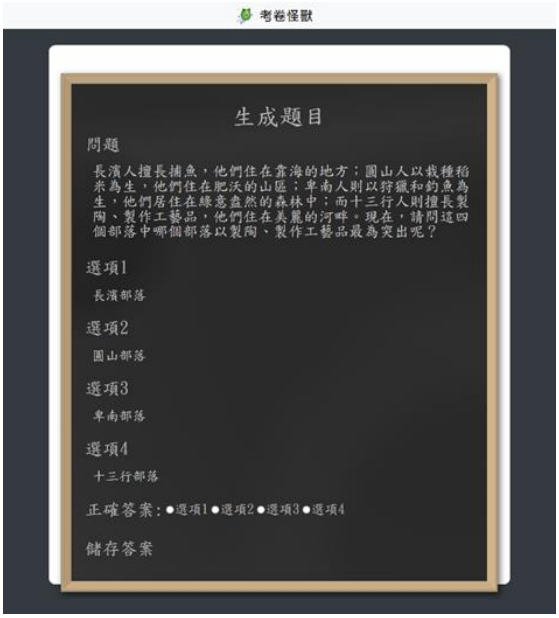


圖 2 生成題目與修改介面

4.3. 練習與挑戰階段：增強社會研究課程中的學習動機與同儕互動

在這一階段，為了促進自主練習，我們設計了 "挑戰與回答 "線上遊戲，以促進學生課後參與。平台利用學生定制的虛擬寵物（稱為 "小怪獸"）來挑戰同儕提出的問題。這種方法豐富了學生對課程內容的理解和應用，讓學生可以根據自己怪物的屬性進行個人化挑戰。在遊戲過程中，會隨機出現一個由同儕產生的問題，學生必須在規定時間內回答。 錯誤的答案會導致血量扣減，歸零則表示挑戰失敗。學生可以隨時結束遊戲，領取獎勵，獎勵的倍數取決於怪物的屬性和正確答案（圖 3）。這種策略性設計鼓勵學生主動參與學習，並加強了社會研究教育中學習動機和同儕互動的融合。



圖 3 同儕練習與挑戰題目介面

5. 結果

5.1. 三次測驗中學習成效之差異

本研究對各年級第一、第二、第三單元的學生進行了三次社會課能力測試。本研究分析了三次社會學能力測驗的前測和後測結果，並以 Wilcoxon 符號比較了學生社會學能力測驗

前測和後測之間的差異。 根據表 1 的結果，在第一單元測驗中，學生的社會研究能力有所提高，但不明顯(Z= -1.878, p= .060 > .05)，但在第二單元測驗中，學生的社會研究能力有顯著提升(Z= -4.162, p= .000 < .01)，以 Wilcoxon 無符號比較學生社會研究能力測驗前測與後測的差異，分析結果如表 2 所示。 如表 3 所示，在第三次第三次單元測驗中，學生的社會研究能力比前兩次有了更明顯的改善(Z= -4.075, p= .000 < .01)。 這表明本研究的課程設計和系統經過長期的實踐，對學生的社會課成績產生了顯著的影響。

表 1 單元一測驗

	Test	***M***	***SD***	***Z***	***p***
社會課測驗	單元一前測	61.62	22.424	-1.878	.060
	單元一後測	68.65	21.879		

**p<.05, **p<.01, ***p<.001*

表 2 單元二測驗

	Test	***M***	***SD***	***Z***	***p***
社會課測驗	單元二前測	67.03	19.701	-4.162***	.000
	單元二後測	82.16	20.566		

**p<.05, **p<.01, ***p<.001*

表 3 單元三測驗

	Test	***M***	***SD***	***Z***	***p***
社會課測驗	單元一前測	64.05	22.908	-4.075***	.000
	單元一後測	79.730	21.1041		

**p<.05, **p<.01, ***p<.001*

6. 結論與討論

本研究採用生成式人工智慧（Generative AI）生成問題，並調查其對學生社會研究能力的直接影響。透過在較長時間內連續進行的三個測試單元，研究表明，在使用該系統期間，學生的社會研究能力有了顯著提高。

研究得出結論，生成式人工智慧整合對社會研究教育中的學習動機產生了正面影響。學生的回饋意見強調了其革新傳統學習方法、培養獨立思考和協作學習的潛力。教師也認識到它在塑造教育體驗方面的變革潛力。雖然 "生成式人工智慧整合 "的創新性質顯示了其前景，但要將其融入主流教育實踐，還需要繼續探索和驗證。未來的教育科技研究和發展有可能重塑和豐富教育工作者和學生的教育體驗。

致謝

本研究感謝國科會專題研究計畫（計畫編號: 112-2628-H-008-001-）與教育部教學實踐研究計畫（計畫編號：PSK1122995）以及國立中央大學學習科技研究中心支持。

參考文獻

Arruabarrena, R., Sánchez, A., Blanco, J. M., Vadillo, J. A., & Usandizaga, I. (2019). Integration of good practices of active methodologies with the reuse of student-generated content. *International Journal of Educational Technology in Higher Education, 16*(1), 1-20.

Chan, T.-W., & Chou, C.-Y. (1995). Simulating a learning companion in reciprocal tutoring systems.

Chen, Z. H., Liao, C., Chien, T. C., & Chan, T. W. (2011). Animal companions: Fostering children's effort-making by nurturing virtual pets. *British Journal of Educational Technology, 42*(1), 166-180.

Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics, 21*, 17-23.

Divate, M., & Salgaonkar, A. (2017). Automatic question generation approaches and evaluation techniques. *Current Science*, 1683-1691.

Doyle, E., Buckley, P., & McCarthy, B. (2021). The impact of content co-creation on academic achievement. *Assessment & Evaluation in Higher Education, 46*(3), 494-507.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., . . . Ahuja, M. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, 102642.

Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Medical Education, 9*(1), e46885.

Heron, G., & Lerpiniere, J. (2013). Re-engineering the multiple choice question exam for social work. *European Journal of Social Work, 16*(4), 521-535.

Hwang, G.-J. (2014). Definition, framework and research issues of smart learning environments-a context-aware ubiquitous learning perspective. *Smart Learning Environments, 1*(1), 1-14.

Hwang, G.-J., & Chen, N.-S. (2023). Editorial Position Paper. *Educational Technology & Society, 26*(2).

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 1-20.

Jones, J. A. (2019). Scaffolding self-regulated learning through student-generated quizzes. *Active Learning in Higher Education, 20*(2), 115-126.

Kim, J. (2023). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies*, 1-32.

Song, D. (2016). Student-generated questioning and quality questions: A. *Research Journal of Educational Studies and Review, 2*(5), 58-70.

Thalheimer, W. (2003). The learning benefits of questions. *Work Learning Research*.

Velander, J., Taiye, M. A., Otero, N., & Milrad, M. (2023). Artificial Intelligence in K-12 Education: eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning. *Education and Information Technologies*, 1-21.

Wu, M., Liao, C. C., Chen, Z.-H., & Chan, T.-W. (2010). *Designing a competitive game for promoting students' effort-making behavior by virtual pets.* Paper presented at the 2010 Third IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning.

Yau, K. W., Chai, C., Chiu, T. K., Meng, H., King, I., & Yam, Y. (2023). A phenomenographic approach on teacher conceptions of teaching Artificial Intelligence (AI) in K-12 schools. *Education and Information Technologies, 28*(1), 1041-1064.

Yu, F. Y., Liu, Y. H., & Chan, T. W. (2005). A web-based learning system for question-posing and peer assessment. *Innovations in education and teaching international, 42*(4), 337-348.

自動問題生成之以知識追蹤為輔的適性程式設計學習初探

A pilot Study of Adaptive Programming Learning Aided by Knowledge Tracking for Automatic Question Generation

呂浚宏[1]，洪暉鈞[2*]，楊叔卿[3]
國立中央大學 網路學習科技研究所
[1]luhubert51824@gmail.com [2*] hch@cl.ncu.edu.tw [3]scy@mx.nthu.edu.tw

【摘要】 近年來，以學生為主的教學方法逐漸被各教育機構所採用，例如:合作學習，教中學等等，但也衍生出無法準確測量學生的學習狀況。因此，我們想藉由問題測驗的方式，來輔助教師了解學生學習狀況。為此本研究預計研發一科技輔具，使用 ChatGPT 來達成自動問題生成(Automatic Question Generation, AQG) 以減輕教師負擔；再經由知識追蹤對問題之知識點進行標記，依學生填答結果對下次問題做適性化調整，期望提高學生學習表現也讓教師掌握學生之學習概況。
【關鍵詞】 適性化學習、自動問題生成、知識追蹤、間隔學習 、ChatGPT

Abstract: *In recent years, student-centered teaching methods have been gradually adopted by various educational institutions, such as cooperative learning and teaching within the learning process, among others. However, this has also given rise to difficulties in accurately assessing students' learning progress. Therefore, we aim to assist teachers in understanding students' learning status through the use of question tests. In this research, we plan to develop a technological tool that utilizes ChatGPT for Automatic Question Generation (AQG) to alleviate the burden on teachers. Additionally, we will label the knowledge points related to the questions through knowledge tracking and make adaptive adjustments to the next set of questions based on students' responses. We hope to enhance students' learning performance and provide teachers with a better overview of students' learning progress.*
Keywords: Adaptive Learning, Automatic Question Generation, Knowledge Tracing, Spaced Learning, ChatGPT

1. 前言

由於近年來，教中學這一學習概念越來越流行，且這種學習模式也被應用於程式設計課程之中。然而，對學生學到了多少，卻無法得知。因此衡量學習標準的存在是必須的。而問題則是評估學習表現的常見方法(Pinter et al., 2020)。然而，創造問題對於教師來說不僅費時也費力同時也無法確定學生是否有效吸收課堂所學知識。(Lu, Huang, Tsai, & Yang, 2021)。因此，需要一種輔助工具，降低教師負擔的同時又能確切掌握學生之學習狀況。本研究欲利用現代 AI 技術，特別是自然語言處理（Natural Language Processing, NLP），提供自動問題生成來當作解決方案並結合深度知識追蹤演算法，針對不同程度學生自動生成問題組，讓教師能確切掌握學生之學習狀況。

2. 研究方法

2.1 教育中的人工智慧

人工智慧（Artificial Intelligence, AI）或者又稱機器智慧(Computational Intelligence)，可能是電腦科學中最古老的領域，而且非常廣泛，涉及模仿認知功能以解決現實世界問題以及建立像人一樣學習和思考的系統的各個方面(Holzinger, Langs, Denk, Zatloukal, & Müller, 2019;

Poole, Goebel, & Mackworth, 1998)。AI 能夠分析和解釋大量數據、生成模擬和情境，並以清晰簡潔的方式傳達研究結果，有望顯著提升研究的效率和效力。它專注於開發機器和計算系統，使其能夠模擬人類智慧的某些方面(M Alshater, 2022)。這包括機器學習、問題解決、自然語言處理等技術和方法的應用。在教育研究中，AI 適用於多種情況，例如適性化學習、虛擬教學助手、預測學生的學習路徑和策略等多種方面(Chassignol, Khoroshavin, Klimova, & Bilyatdinova, 2018; Chen, Chen, & Lin, 2020; Lu et al., 2021)。總之，AI 在教育中的應用有助於提高學習的效率和質量，並為學生和教育機構提供更多工具和資源，以實現更好的學習結果。

2.2 自然語言處理與自動問題生成技術

自然語言處理（Natural Language Processing, NLP）是 AI 領域的一個分支，專注於使計算機能夠理解、解釋和生成自然語言（例如英語、中文、日語等）的能力(Chowdhary & Chowdhary, 2020)。NLP 關注計算機理解和生成人類語言的能力，其在學術研究中有多種廣泛的應用。測驗問題生成已成為教育技術和自然語言處理的新興研究領域 (Bednarik & Kovacs, 2012; Beg & Beg, 2018; Deena & Raja, 2022; M. Liu, Calvo, & Rus, 2010)。自動問題生成（Automatic Question Generate, AQG)系統還有在一些研究中被用來生成常見問題庫中部份的問題 (Indurthi, Raghu, Khapra, & Joshi, 2017; Khvalchik & Kulkarni, 2017)。Le 等人將 AQG 的概念定義為：“從各種輸入（如原始文本、數據庫或語義表示）生成問題” (Le, Kojiri, & Pinkwart, 2014, p. 352)，它已成為最近的熱門研究課題由於神經網絡的自然語言處理的出現，它旨在模仿人類如何使用語言並學習如何去運用人類語言來達成特定需求的工具(Chowdhary, 2020)。AQG 可用於生成各種類型的問題，例如克漏字和多選題等等(Ch & Saha, 2018)。因此本研究打算透過大型語言模型 ChatGPT (Chat Generative Pre-trained Transformer, ChatGPT)來達成 AQG，藉此來嘗試機器生成 python 的可行性。

2.3 生成式人工智慧

科技的快速發展和世界的互聯互通為社會、經濟、環境帶來了巨大的變化。近年來，人工智慧取得了顯著進步，引發了 Open AI 的 ChatGPT 等突破性技術的誕生。ChatGPT 語言模型等現代科技有可能徹底改變教育格局(Mhlanga, 2023)。生成式人工智慧(Generative Artifical intelligence, GAI)最早一名詞出現於(van der Zant, Kouw, & Schomaker, 2013)，目前最具代表性的 GAI 就是 ChatGPT。ChatGPT (Chat Generative Pre-trained Transformer)，是由一家新興公司 OpenAI 所開發的人工智慧聊天機器人程式，於 2022 年 11 月推出(Baidoo-Anu & Ansah, 2023)，在最近的幾個月非常有討論度(Yilmaz & Yilmaz, 2023)。該程式使用基於 GPT-3.5、GPT-4 架構的大型語言模型並以強化學習訓練。例如讓 ChatGPT 扮演教師角色來指導學生、回答學生問題，或是學習夥伴等(Hwang & Chen, 2023)。綜上所述，本研究將嘗試利用 ChatGPT 生成 Python 程式選擇題，並比較其生成的問題質量與教師生成的問題是否相似，除了專家與教師檢測，我們還結合 Lu 等人的方法，讓學生去猜測那些題目是機器生成的題目，作為質量檢測的輔助(Lu et al., 2021)。

2.4 知識追蹤

知識追蹤模型根據學習者的過去答題表現行為序列來建模，應用機器學習從而預測學習者對知識的掌握程度。過去的知識追蹤其主要的依據是一階馬可夫模型，最終被發展出的模型為貝葉斯知識追蹤模型(BKT)(Corbett)，而貝葉斯神經網路是機率分布的圖形表示(Jensen & Nielsen, 2007)，在程式教育方面，大部分的研究皆專注於如何去提取程式學習的知識點上面，知識點的分配主要分成以下兩種，第一種是在未確定課程範圍的情況下，透過建立語法樹的方式，對程式設計的內容進行知識點的分配，第二種則是預先確認好了授課範圍，然後立即先對知識進行分類(D. Liu, Dai, Zhang, Li, & Zhang, 2020)，再去進行知識追蹤。而本研究打算採取知識追蹤的第二種方法，預先對知識進行分類之概念，以便依學生之填答結果對問題做

出適性化之調整。本研透過此技術對學生知識點的正確率作出計算，並將學生的知識點正確率通過視覺化的方式做呈獻，並對學生的編程紀錄進行結構與錯誤的統計後，透過知識追蹤模型對其進行檢驗，作為教師提供一個評估課程與學生學習狀態的方法。

2.5 適性化學習

適性化學習(adaptive learning)，著重於根據學習者的當前表現調整學習內容和難度，而不依賴於學習者的個人詳細資訊，如背景、喜好或習慣。在這種學習模式中，學習路徑會根據學生在測驗或學習活動中的具體表現進行適性化調整(Kerr, 2016; Paramythis & Loidl-Reisinger, 2003)。此外，適性化學習系統提供即時反饋，幫助學生及時了解自己的學習進展和需求。這種實時的調整和反饋機制確保了學習材料和問題始終與學生的當前水平和能力相匹配。這種方法的核心是利用先進的教育技術，包括 AI 和資料分析，來處理和理解大量的學習資料，並從中提煉出有益於學習過程的見解，有望在未來的教育領域發揮更大的作用(Carceles-Poveda & Giannitsarou, 2007; Zeiler, 2012)。本研究通過這種方法結合知識追蹤，對學生的當前較為不熟悉的地方進行針對性的調整，進行有效的適性化學習。

2.6 間隔學習/間隔效應

間隔學習是指在相同學習時間下，相較於集中在單一時段內學習教材，跨時間間隔進行學習能夠更有效地提高長期記憶保留,是一種經常被使用且有助於學習的方法(Dempster, 1996; Sisti, Glass, & Shors, 2007)。在應用方面，間隔效應不限於純粹的記憶力訓練，還能被應用於一些抽象概念的學習和理解(Vlach & Sandhofer, 2012)。此外 Tasi 等人更表明重複測試對於學生的長期記憶是有利的(Tsai, Huang, Lu, & Yang, 2021)。因此，本研究除了在課後適性化練習的部分還結合了有效的學習策略「間隔效應」，讓學生能進行較短的間隔複習之外，課中的正式測試我們也將時間間隔固定為一周一次(較長的時間間隔)，讓學生不管在記憶保留或者牢固程度上都能有較好的表現。

3. 實驗設計

3.1 實驗流程

本研究規劃配合之課程為本校「Python 教育資料探勘實作」課程，在設計上結合性化間隔學習之專題實作式教學搭配學生為主體的教學設計，且每周皆會進行適性化間隔學習系統進行練習，以提高學生自主學習及協作學習並深化學生的學習經驗。本研究之「Python 教育資料探勘實作」課程有 18 週。在第一週的部分，第一節課會先進行實驗流程介紹，為了先了解學生的程式能力及先備知識，我們會讓學生進行 python 程式能力測驗(前測)，測驗時長為 20-30 分鐘，共 20 題選擇題，題目為機器所生成題目，共 6 題，和教師所生成題目，共 14 題會這樣分配各個測試題目的原因是因為圖靈測試要求機器只參與整個測試的 1/3 到 1/4。第二週至第十七週則是會在第一節課的前 10-15 分鐘，要求學生進行每週課堂測試，測試時長約為 10-15 分鐘，每次測試共 10 題選擇題，題目為機器所生成題目，共 3 題，和教師所生成題目，共 7 題；其安排目的為藉節由固定的時間間隔，使學生有時間能消化每周所學知識，也能藉此掌握學生對於知識的掌握狀況；另外，在課後測試練習的部分，每次測驗會從題庫中推送 10 題讓學生練習；練習題庫為逐周更新，在每週的課程結束後，便會新增和該週知識點相關之題目到題庫之中，例如第二周開始後，題庫會包含的題目為前測(20 題)+第一周(30 題)+第二周(30 題)共 80 題，第三周則為第二周(80 題)+第三周(30 題)共 110 題，依此類推 110+n+n…..。且題目為機器生成和機器生成混合，且在每次練習完題目後，都會經由我們開發的適性化演算法依學生的答題結果，對下一次的練習題目進行適性化調整。最後一週則是進行 python 程式能力測驗(後測)，和前測相同測驗時長約為 20-30 分鐘，題目為機器所生成

題目，共 6 題，和教師所生成題目，共 14 題。分配試題之原因如上述前測部分所言，另外我們還會新增一問卷讓學生判別哪些題目為機器生成之題目，哪些為專家生成之題目。

3.2 系統介紹

本研究建立課程專用伺服器，提供學生連線進行操作學習，並保留各學生練習測驗日誌資料。抓取學生編程日誌。經量化及資料處理後以資料庫型式儲存。最後對資料庫內的資料進行分析，如圖 2。而在問題生成方面，機器生成的問題預計使用 ChatGPT 來進行問題生成，通過參考 Hwang 等人對於 GAI 在教育中所扮演的腳色分類(Hwang & Chen, 2023)，讓 ChatGPT 扮演成一位 Python 程式設計課的老師，給定”prompt : Create multiple questions and answers for a 15-week Python bootcamp for master students who’s going to study computer programing. Here are the schedule for the python course. Let’s start from week 1.”和”prompt : Ok, now I want some specific code-based python MCQs, and try to let the MCQs be more and more difficult .”，讓其扮演一位程式課老師，並根據我們所提出的課程安排（15 周的原因為總課程周數 18 周扣除期中報告 1 周和期末報告 2 周），使其自動生成 python 程式選擇題。知識追蹤：我們會先確認課程內容，並對「python 教育資料探勘實作」課程內教師所要教導的內容（知識點），進行分類。適性化調整：藉由知識追蹤所標記好的知識點，我們可以對學生每次的填答結果作出紀錄和分析，並對問題進行調整。

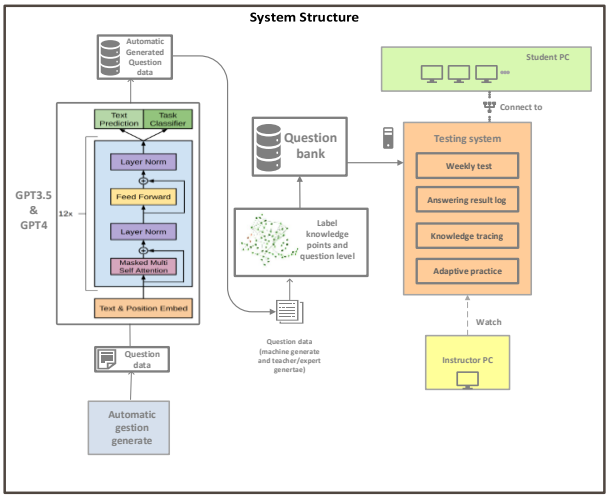


圖 1 適性化間隔學習系統

3.3 適性化推題機制

本研究推題機制主要依據使用者在特定知識點上的正確率，尤其是那些正確率較低的部分。而本研究的知識點正確率的計算方式(get_accuracy_for_knowledge_points)首先會透過”Answer”模型去找出使用者的所有答題記錄，再來會有兩個字典”correct_counts”和”total_counts”分別追蹤每個知識點的答對數和答題總數，而在處理完所有答案之後(使用者每次練習完)，則會計算出每個知識點的正確率，並儲存在一個知識點正確率的字典之中。再來則是適性化推題機制，整個流程分為三個階段，每次共推薦 10 道題目。首先，我們會將剛剛 get_accuracy_for_knowledge_points 算出來的知識點正確率進行排序，挑選出最低知識點正確率。其次，我們查找使用者之前回答過且答錯的題目。這些題目是根據使用者在特定知識點上的正確率來挑選的，選擇最低正確率的知識點相關的題目。接下來，如果這些題目不足 10 題，則會補充使用者尚未接觸過的知識點相關的題目(由於我們的題庫是按課程進度逐步開放，因此方式可以讓每次課程結束後，學生回去練習時，較能優先讓學生練習與剛教完的知識點相關之題目)，直到達到 10 題為止。最後，如果上述兩步驟仍然無法提供足夠的 10 題，則會從整個題庫中隨機選擇剩餘的題目。

- 狀況1：使用者之前練習且答錯的題目（N），以及學生尚未接觸過的知識點相關的題目（N2），本次練習題數為（N+ N2）。
- 狀況2：使用者在測驗中沒有答錯任何題目，所以系統專注於推送給使用者尚未學過的知識點相關題目（N2），本次練習題數為（N2）。
- 狀況3：使用者沒有做錯過任何的題目，且尚有未接觸過之知識點的題目（N2），但其題目不足10題。為了達到10題的測驗量，系統會隨機添加題目（N3），本次練習題數為（N2+ N3）。
- 狀況4：使用者在測驗中答錯題目（N）較多，所以系統專注於讓使用者複習他們已經學過但答錯的題目，本次練習題數為（N3）。
- 狀況5：使用者有答錯的題目（N）和尚未學過的知識點相關題目（N2），以及若題目總數未達到10題時添加的隨機題目（N3）。複習的同時也包含了新的學習內容，並通過隨機題目確保有足夠的題量，本次練習題數為（N+ N2+ N3）。

此方法，除了根據每個使用者的情況進行適性化的推題幫助他們在不太熟悉或經常犯錯的知識點上提高理解和掌握之外，也會保有隨機性的推送題目，降低不斷練習重複練習的倦怠感和讓學生仍保有練習新題目的機會。

4. 結果

4.1 Python 程式能力檢驗

本研究就學生 Python 程式能力測驗前後測利用 Wilcoxon 符號檢定比較學生 Python 程式能力前、後測之差異。根據表一結果顯示，經一學期實驗，學生的 Python 程式能力有非常顯著的進步(z=-5.314,p=.000<0.001)。代表本研究的課程設計有助於提升 Python 程式語言的學習成效，如表一所示。

表一 Python 程式能力測驗 Wilcoxon 符號檢定結果

	測驗	**M**	**SD**	**z**	***p***
Python 程式能力	前測	47.16	12.993	-5.314***	.000
	後測	81.76	12.031		

**p<.05, **p<.01, ***p<.001*

4.2 機器生成題目之質量檢測

而本研究為了確認機器生成題目的質量是否與教師生成題目之質量相似，參考 Lu 等人的研究方法，採用圖靈測試，學生總數為 38 人。當中 Python 程式能力測驗後測由機器生成之題目的題號分別為 1,3,4,5,7,15(圖靈測試中要求機器生成知題目不能超過總題數的 30%)；第 1 和 3 題共有超過 60%的學生都有成功分辨出機器生成之題目，但第 4,5,7 題則有超過一半的學生都無法成功辨別，甚至 4 和 5 兩題更有超過 70%的學生認為是教師生成的題目；而第 15 題則是一半的學生認為題目是機器生成而另一半則認為是教師生成，這是一個很有趣的現象。藉由此情形我們可以的推測，機器生成的題目知質量瘸時和教師生成之題目質量類似，甚至在部分情況下還尤有過之。

5. 結論

本研究藉由 ChatGPT 來進行 Python 程式課程中的選擇題自動生成，有效地為教師們節省了出題時間，以降低問題所需成本。並且透過本研究的適性化學習系統，結合知識追蹤對於各週的題目進行知識典的標記還有重複練習，讓系統能藉由學生的練習紀錄，進行針對性

的調整；而學生則能藉由適性化的出題，多多練習自己較不熟悉的題目；教師可以透過學生的答題紀錄了解其知識缺陷，掌握學生學習近況。

致謝

本研究感謝國科會專題研究計畫（計畫編號: NSTC112-2628-H-008-001-、NSTC 111-2410-H-007-023-/MOST 112-2410-H-007-0023、MOST 108-2511-H-007 -008 -MY3）與教育部教學實踐研究計畫（計畫編號：PSK1122995）以及國立中央大學學習科技研究中心支持。。

參考文獻

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62. doi:10.2139/ssrn.4337484

Bednarik, L., & Kovacs, L. (2012). Implementation and assessment of the automatic question generation module. *2012 IEEE 3rd international conference on cognitive infocommunications (CogInfoCom)*, 687-690. doi:10.1109/CogInfoCom.2012.6421938

Beg, A., & Beg, A. (2018). Using open technologies for automatically creating question-and-answer sets for engineering MOOCs. *Computer Applications in Engineering Education, 26*(3), 617-625. doi:10.1002/cae.21913

Carceles-Poveda, E., & Giannitsarou, C. (2007). Adaptive learning in practice. *Journal of Economic Dynamics and Control, 31*(8), 2659-2697. doi:10.1016/j.jedc.2006.09.004

Ch, D. R., & Saha, S. K. (2018). Automatic multiple choice question generation from text: A survey. *IEEE Transactions on Learning Technologies, 13*(1), 14-25.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science, 136*, 16-24. doi:10.1016/j.procs.2018.08.233

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *Ieee Access, 8*, 75264-75278. doi:10.1109/ACCESS.2020.2988510

Chowdhary, K. (2020). *Fundamentals of artificial intelligence*: Springer.

Chowdhary, K., & Chowdhary, K. (2020). Natural language processing. *Fundamentals of artificial intelligence*, 603-649. doi:10.1007/978-81-322-3972-7

Corbett, A. A1DERSO1, JR 1995. *Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. 8VHU 0RGHOLQJ DQG 8VHU $ GDSWHG, QWHUDFWLRQ, 4*, 253-278. doi:10.1007/BF01099821

Deena, G., & Raja, K. (2022). Objective Type Question Generation using Natural Language Processing. *International Journal of Advanced Computer Science and Applications, 13*(2). doi:10.14569/ijacsa.2022.0130263

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. *Memory*, 317-344. doi:10.1016/B978-012102570-0/50011-2

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9*(4), e1312. doi:10.1002/widm.1312

Hwang, G.-J., & Chen, N.-S. (2023). Editorial Position Paper. *Educational Technology & Society, 26*(2). Retrieved from https://www.jstor.org/stable/48720991

Indurthi, S. R., Raghu, D., Khapra, M. M., & Joshi, S. (2017). Generating natural language question-answer pairs from a knowledge graph using a RNN based question generation model. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 376-385. Retrieved from https://aclanthology.org/E17-1036.pdf

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs* (Vol. 2): Springer.

Kerr, P. (2016). Adaptive learning. *Elt Journal, 70*(1), 88-93. doi:10.1093/elt/ccv055

Khvalchik, M., & Kulkarni, A. (2017). Open-domain non-factoid question answering. *International Conference on Text, Speech, and Dialogue*, 290-298. doi:10.1007/978-3-319-64206-2_33

Le, N.-T., Kojiri, T., & Pinkwart, N. (2014). Automatic question generation for educational applications–the state of art. *Advanced computational methods for knowledge engineering*, 325-338. doi:10.1007/978-3-319-06569-4_24

Liu, D., Dai, H., Zhang, Y., Li, Q., & Zhang, C. (2020). Deep knowledge tracking based on attention mechanism for student performance prediction. *2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI)*, 95-98. doi:10.1109/CSEI50228.2020.9142472

Liu, M., Calvo, R. A., & Rus, V. (2010). Automatic question generation for literature review writing support. *International conference on intelligent tutoring systems*, 45-54. doi:10.1007/978-3-642-13388-6_9

Lu, O. H., Huang, A. Y., Tsai, D. C., & Yang, S. J. (2021). Expert-Authored and Machine-Generated Short-Answer Questions for Assessing Students Learning Performance. *Educational Technology & Society, 24*(3), 159-173. Retrieved from https://www.jstor.org/stable/27032863

M Alshater, M. (2022). Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT. *Available at SSRN*. doi:10.2139/ssrn.4312358

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023)*. doi:10.2139/ssrn.4354422

Paramythis, A., & Loidl-Reisinger, S. (2003). *Adaptive learning environments and e-learning standards.* Paper presented at the Second european conference on e-learning.

Pinter, R., Maravić Čisar, S., Kovari, A., Major, L., Čisar, P., & Katona, J. (2020). Case study: students' code-tracing skills and calibration of questions for computer adaptive tests. *Applied Sciences, 10*(20), 7044. doi:10.3390/app10207044

Poole, D. I., Goebel, R. G., & Mackworth, A. K. (1998). *Computational intelligence* (Vol. 1): Oxford University Press Oxford.

Sisti, H. M., Glass, A. L., & Shors, T. J. (2007). Neurogenesis and the spacing effect: learning over time enhances memory and the survival of new neurons. *Learning & memory, 14*(5), 368-375. doi:10.1101/lm.488707

Tsai, D. C., Huang, A. Y., Lu, O. H., & Yang, S. J. (2021). Automatic question generation for repeated testing to improve student learning outcome. *2021 International Conference on Advanced Learning Technologies (ICALT)*, 339-341. doi:10.1109/ICALT52272.2021.00108

van der Zant, T., Kouw, M., & Schomaker, L. (2013). *Generative artificial intelligence*: Springer.

Vlach, H. A., & Sandhofer, C. M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child development, 83*(4), 1137-

1144. doi:10.1111/j.1467-8624.2012.01781.x

Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 100147.

Zeiler, M. D. (2012). Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*. doi:10.48550/arXiv.1212.5701

實體物件與數位雙生技術相結合的互動學習系統

Interactive Learning System Combining Physical Objects with Digital Twin Technology

游子諭[1]，林孟衡[2]，王振漢[3*]，楊舒涵[4]，陳國棟[5]
[125]國立中央大學資訊工程學系
[3]國立中央大學學習科技研究中心
[4]健行科技大學餐旅管理系
[*]harry@cl.ncu.edu.tw

【摘要】本研究設計了一款可融入使用者與實體物件之數位雙生實境學習系統，藉由立體視覺相機將使用者與周遭實體環境放入數位實境之中。教師可以透過編輯系統將課本內容轉換成情境教材，讓學生以操作實體物件的方式在數位情境中學習相關知識，突破以往虛擬實境(VR)與擴增實境(AR)透過滑鼠或控制器操作的限制，允許學習者與真實物件互動。這種整合實體與數位元素的學習環境，不僅使學習者能在真實情境中實踐知識與技能，也增強了數位環境中的感知與動作表現。實驗結果顯示，此模式顯著提升學習成效，並透過問卷與訪談反饋，在學習動機方面有正面效果。

【關鍵字】情境認知；具身學習；虛擬實境；混和實境；數位雙生

Abstract: *This study developed a digital twin reality learning system integrating users and physical objects using stereoscopic cameras. Teachers transform textbook content into digital, interactive materials, allowing students to learn by handling real objects. This transcends traditional VR and AR limitations, enabling real-world interaction and enhancing digital perception and action. Results show improved learning effectiveness and increased motivation, as confirmed by surveys and interviews.*

Keywords: Situational Cognition, Embodied Learning, Virtual reality, Mixed reality, Digital Twin

1. 緒論

1.1. 研究背景

在傳統教學方法中，學生多為被動接收者，坐在教室裡聆聽以課本或講義為主的教學內容。這種方式往往限制了學生將課堂所學知識靈活運用於實際問題解決的能力。情境認知理論提出知識與行為是不可分割的，並與社會、文化及自然環境的互動相關(Brown et al., 1989)。因此，將應用情境與傳統課程結合被認為可以有效提升學習速度和知識運用(Miller & Gildea, 1987)，並提高學習者的學習成效與學習動機(Chiang et al., 2018)。

隨著科技發展，虛擬實境(VR)被用於營造沉浸式學習環境，目標是讓使用者在數位實境中體驗情境、互動與學習(Freina & Ott, 2015)。然而，VR 大多以控制器或觸控介面與虛擬物件互動，缺乏真實的互動體驗。具身認知理論強調透過動作與感知參與的重要性(Wilson, 2002)，且實體動作與互動有助於提升學習者參與度(Lindgren et al., 2016)。

擴增實境(AR)融合虛擬元素於現實環境中，提供豐富互動性的學習體驗(Bacca Acosta et al., 2014)。然而，在多數 AR 應用中，使用者難以看見自己的動作表現，這限制了其即時修正和

反思學習的能力。混和實境(MR)透過頭戴裝置上的外置鏡頭辨識手勢，允許使用者在實體空間與系統環境互動，已證明能提升實作課程中的學習成效(Gerup et al., 2020)，然而在 MR 系統中以手勢或肢體動作在數位實境互動，無法提供使用者實際的操作回饋，與動手操作實體物件的感知還是有所落差。

由上述研究可知，目前的設備未能在數位實境中如實呈現使用者自身的動作與感知，學習者在學習過程中無法看到自身動作進行反思學習，此外，使用者大多透過控制器與虛擬物件進行互動，無法接觸到真實情境的物件，缺乏對於實際的應用情境的完整呈現。為了克服這些限制，本研究以影像辨識技術為基礎設計了一款可融入使用者與實體物件之數位雙生實境學習系統。這個系統目的在於創造出一個既可以提供真實操作經驗，又能在數位實境中反映學習者動作的學習環境。透過這種方式，學習者可以在沉浸式與互動性更強的環境中進行學習，從而提高學習成效和動機。

1.2.研究目標

本研究旨在設計一款整合使用者與實體物件的數位雙生實境學習系統，讓使用者可以操控實體物件進行學習，同時透過系統觀察自身的學習表現進行反思學習，並驗證其對學習成效與學習動機的影響。

1.3.研究問題

根據上述的研究目標，本研究所定的問題如下：

1. 如何建立一個可以將使用者與周遭實體環境放入數位實境的虛實整合情境?
2. 在實體世界操作的數位實境是否可以比僅在數位實境裡做中學習有更好的學習成效?
3. 在實體世界操作的數位實境是否可以比僅在數位實境裡做中學習有更好的學習動機?

2. 相關研究

2.1.情境學習

情境認知(Situated Cognition)是一種學習理論，強調在特定情境中學習者的參與、觀察、互動和實踐對知識建構和理解的重要性。這種理論的核心在於知識和行為的密切相關性，並與社會、文化和環境互動相關(Brown et al. 1989)。真實學習(Authentic Learning)作為情境學習的一種方法，通過結合真實世界的任務、背景和教材來促進學習者對情境的理解和知識應用，從而提高學習成效(Ornellas et al., 2019)。

2.2.數位化的學習環境

數位化學習環境通過運用電腦和相關技術來營造情境學習環境。這包括利用電腦遊戲來合作學習第二外語，或使用虛擬實境技術如 Second Life 來學習醫學知識(Lorenzo-Alvarez et al., 2020)。擴增實境技術(AR)則是透過在現實世界中疊加圖像或資訊來幫助學習，例如遠端實驗室中的鐘擺實驗(Zandavi & Chung, 2018)。數位學習環境相較於實體教室，更易於打造符合應用情境的學習環境，並可以通過簡化的實體場景布置轉換成完整的數位實境場景進行情境學習。

2.3.具身學習

具身學習法是一種讓學習者的身體和感覺運動系統積極參與學習過程的方法。具身認知理論(Wilson, 2002)強調思維和認知不僅僅是腦力活動，而是與身體的感覺和動作緊密相連。有研究(Lindgren et al., 2016)指出，通過具身互動的學習方式可以有效提高學習者的參與度和學習效果。在技術進步的推動下，如 Switch 或 HTC Vive 等裝置讓使用者能夠透過身體動作與

數位環境進行互動，並已在醫學教育等領域得到應用(De Ribaupierre et al., 2014)。

2.4.數位雙生

數位雙生是一種用於描述實體物理製造產品的虛擬資訊結構，從微觀原子到宏觀幾何層面模擬物理實體、流程或系統(Grieves & Vickers, 2017)。數位雙生技術最初應用於虛擬工廠追蹤產品開發週期，核心包含真實環境、數位媒體和數據連接三部分(Tao et al., 2018)。在教育領域，數位雙生技術被應用於遠端教學，如創建虛擬的電子實驗室(Palmer et al., 2022)。數位雙生也用於模擬人物動作，如 Movella 開發的 Xsens MVN 動作捕捉系統和 Faceware Live 等，這些技術在教育領域用於提升學生專注度(Ahuja et al., 2021)。數位雙生將抽象或複雜概念轉化為具體視覺表現，幫助學生更容易理解和記憶，促進學習效果。然而，目前技術無法將使用者全身與真實臉部表情完全同步至數位實境中，這是數位雙生技術在教育領域中面臨的挑戰。

2.5.小結

總結相關研究可以發現，過去傳統的數位學習環境多以滑鼠或控制器等操作方式與虛擬物件互動，無法接觸到真實情境的物件，缺乏對於實際應用情境的完整呈現。同時，現行數位實境的大多以頭戴式裝置為主，並以第一視角進行學習，這導致學生無法觀察自身學習表現並進行反思學習，長時間配戴容易造成使用者不適 (Chang et al., 2020)。另外，大多數位實境對於虛擬空間的距離和大小的估計不準確，導致使用者容易失去空間感。在數位雙生的部分，目前的技術無法將使用者全身與真實臉部表情完全同步進數位實境中，同樣影響到使用者的空間感(Henry & Furness, 1993)。

因此本研究以影像辨識為基底，設計了一款可融入使用者與實體物件之數位雙生實境學習系統，讓學習者在情境中透過操作實體物件的方式。此外，我們在畫面上增加了角度和距離等空間輔助資訊，這不僅使學習者能夠更深入地沉浸在學習環境中，還幫助他們更準確地執行特定動作，例如調整鞠躬的角度和與顧客保持適當的距離，同時，學生可以操控現實情境會碰到的物件，例如操控真實的餐具來練習擺盤，以接近真實學習的學習環境來增加學習成效。

3. 系統設計與實作

3.1.設計理念

為了建立一個可以在數位實境進行實體操作學習的虛實整合情境，本研究將關鍵學習內容以簡化的實體操作方式搬進教室中，透過獨立的立體視覺相機，搭配影像辨識的方法同時將多位使用者與周遭物件帶入數位實境，讓學習者能以第三人稱視角在數位實境中以劇本演出的方式學習相關的知識與操作，觀察自身學習表現進行反思學習。

3.2.系統架構與設計

本研究所設計的數位雙生劇場系統架構主要包括四個部分：虛實同步系統(Cyber Physical Synchronization, CPS)、數位雙生劇場、數位雙生劇場編輯器，以及數位雙生劇場控制器，如圖 1 所示。

虛實同步系統 (CPS)：作為架構的核心，負責計算與整合同步方法，將資訊送至數位雙生劇場和編輯器。它利用感測器和追蹤技術，將使用者與物體的動作及位置同步到數位平台。

數位雙生劇場：負責劇本播放和 3D 數位場景建置，創造身歷其境的學習環境。它接收 CPS 系統產生的物件與肢體辨識結果，並在相應位置產生 3D 模型和虛擬角色(Avatar)。此外，系

統在畫面上有角度和距離等空間輔助資訊，幫助學生在數位實境環境中建立接近真實世界的空間感。

數位雙生劇場編輯器：幫助教師將課本內容轉化為情境教材。教師可在此編輯劇本、場景、視角等，使學習材料更符合教學需求。

數位雙生劇場控制器：讓教師控制劇本的演出進度，並提供學生學習過程中的即時反饋和評分。

此外，系統開發環境包括多種軟硬體工具，如 Zed 立體視覺相機、Android 平板電腦等。整體設計目的在提供一個互動式、多感官的學習環境，讓學生透過身體動作與數位環境的互動來提升學習成效。

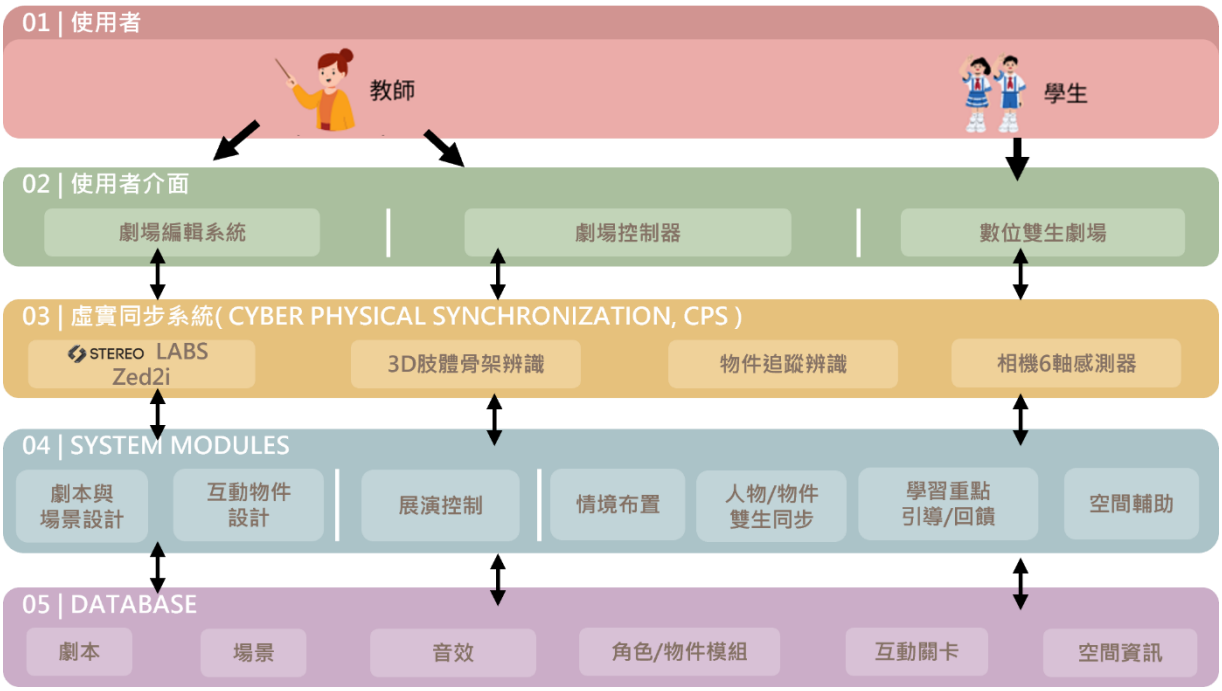


圖 1 系統架構圖

4. 實驗設計

本研究結合物件追蹤與數位雙生技術，可融入使用者與實體物件之數位雙生實境學習系統。透過物件辨識和深度攝影技術，將實體物件及其相關資訊融入數位環境，以圖像和文字輔助學習者記憶，並通過實際操作物件來強化學習效果。本研究將探討這種學習方式與傳統滑鼠操作虛擬物件學習系統在學習成效上的差異。

本研究進行了一項實驗設計，將受試者分成實驗組及對照組進行比較。使用的測量工具包括（1）針對學習成效的前後測試卷、（2）學習動機與使用經驗的問卷量表，以及（3）個別訪談。茲說明如下：

4.1. 實驗對象

本研究與桃園市某科技大學進行合作，受試者來自餐旅日文課程中的兩個班級，並隨機指定一班為實驗組，另一班為對照組，排除中途請假或疫情影響等，實驗組人數為 38 人，對照組人數為 33 人。兩組在實驗過程中皆完全獨立且互不影響。

4.2. 實驗教材內容

本實驗的教材由餐旅日文課程教師設計，結合課程大綱和學生程度，編輯成西式餐廳情境

的劇本。劇本結合課堂所學的日文單字、文法及應對問候語，並融入餐廳服務知識與禮儀，如餐具擺放、上菜順序。實際情境包括新進員工的培訓、客人點菜與結帳等，目的在訓練學生的日文口語表達及實際應用能力。

4.3. 評量工具

為了評估「可融入使用者與實體物件之數位雙生實境學習系統」對學習者的影響，本研究採用了前後測試和相關問卷的方法。學習成效的評量在實驗前後進行，包括課前測和課後測，並結合問卷來深入分析學習者使用後的感受。

前後測試卷由餐旅日文課程教師設計，涵蓋學習範圍內的題目。測驗內容分為三部分：餐旅基礎日文、餐桌設定、餐桌技術服務。測驗結果最終以 ANCOVA 進行統計分析。

問卷設計採用 Likert 五點量表，評分範圍從 1 分（非常不同意）到 5 分（非常同意）。問卷內容基於 MSLQ 模型(Pintrich et al., 1991)，評估三個面向：內在動機(Intrinsic Goal Orientation)、任務價值(Task Value)、學習表現與自我效能(Self-Efficacy for Learning and Performance)。問卷的信度通過 Cronbach's alpha 進行評定。為了更進一步了解實驗組對於數位雙生劇場的使用感受，在後側結束後隨機抽取實驗組學生進行訪談。

4.4. 實驗流程

此次實驗為期八週，並依據本研究所提出之學習模式設計實驗流程，可見圖 2。實驗組與對照組皆依循相同流程與教材進行實驗。教師會依據每週進度進行日文基礎知識與情境教材的教學。而在情境學習與成果發表階段則是讓學生透過分組合作的方式合力演出，分組方式以前測成績進行 S 型分組，確保各組基礎能力相當。其中，實驗組以操作實體物件的方式學習日文餐旅的相關知識，可見圖 3，而對照組則是使用滑鼠操作虛擬物件來學習，可見圖 4。

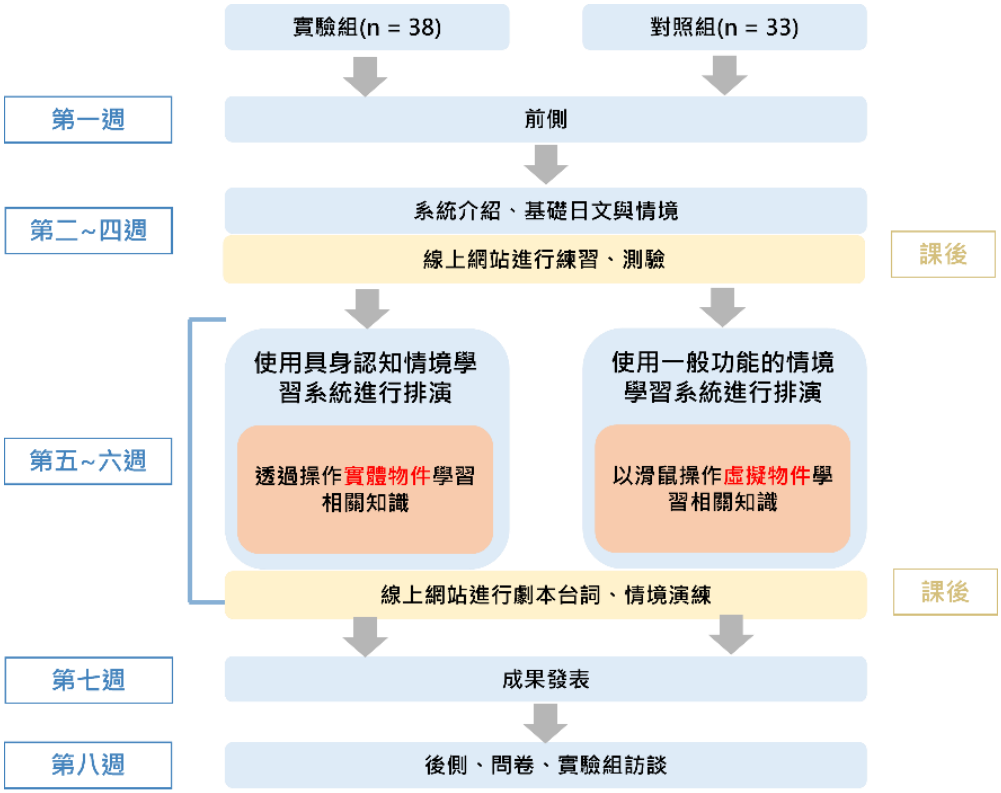


圖 2 實驗流程圖

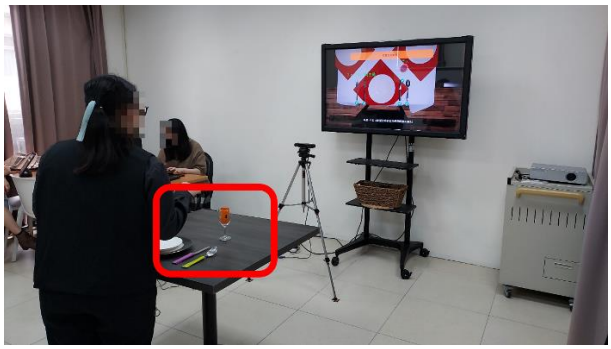


圖 3 實驗組操作實體物件



圖 4 對照組操作滑鼠

5. 實驗結果與分析

5.1. 前後測驗成績

本研究分別進行共變數分析前提驗證及 ANCOVA 檢定，探究兩組學習成效結果的相關性。總共有 71 份有效樣本，分為實驗組 38 份和對照組 33 份。在共變數分析之前，確保了樣本獨立性、常態分佈與變異同質性。

共變數分析顯示，在考慮前測成績後，實驗組與對照組在後測成績上存在顯著差異(F = 11.245, p < 0.05)，且效果量為 0.142，屬於 large effect size。在三個面向（基礎知識、餐桌設定、餐桌服務）的分析中，實驗組在所有面向的成績平均均高於對照組，尤其在餐桌服務方面的提升最為顯著(p = 0.008 < 0.05)，顯示實驗組在餐具擺放與上菜順序的學習表現優於對照組。綜合結果表明，「可融入使用者與實體物件之數位雙生實境學習系統」顯著提升了學習者的學習成效，並且該系統能有效幫助學習者記憶餐具擺放與上菜順序等知識。

5.2. 問卷結果

本研究的問卷結果分析了受試者對於數位雙生劇場使用感受，其 Cronbach’s alpha 值為 0.988，為高信度問卷。本問卷聚焦於三個面向：內在目標導向、任務價值、學習表現和自我效能。內在目標導向的部分目的在探討學習者使用數位情境學習系統是否能提升學習動機和興趣。實驗組的平均感受為 4.032，略高於對照組的 3.901，但這一差異並未達到統計學上的顯著水準。接著是任務價值，此目的在探討學習者對於數位情境學習系統中知識的重要性、效用性以及互動體驗的看法。實驗組的平均感受為 4.105，同樣略高於對照組的 3.901，但同樣沒有達到顯著差異。最後是學習表現和自我效能，目的在評估學習者對自己學習表現的看法。在這方面，實驗組的平均感受為 4.105，相比對照組的 3.848 略高，但同樣未達顯著水準。總體來看，儘管實驗組與對照組在各面向的平均感受均未達顯著水準，但作答結果均在「同意」區間，代表使用數位雙生劇場對於學生的三個面向均有正面的影響。

5.3. 訪談結果

為深入了解數位雙生劇場對學習者的影響，本研究隨機抽取實驗組學生進行訪談。訪談結果顯示學習者對於數位雙生劇場的使用持正面態度。學生們認為透過操作實體餐具的學習方式，能夠更好地與日常生活中的使用情境產生聯想，例如在實習經驗或用餐時對餐具的擺放位置有更深刻的認識。學生們也表示，數位雙生劇場的學習模式有助於記憶餐具位置及相關的日語表達。此外，畫面上的空間輔助訊息幫助學生於數位實境的空間感，這在練習服務和學習姿勢有正向的幫助。實驗觀察顯示，相較於對照組，實驗組在完成學習任務的時間隨著

演練次數的增加而變得更快，表明該學習模式可能提升了學習者的自信心和熟練度。總體而言，透過操作實體物件與系統畫面的即時回饋，學習者能夠更有效地銜接數位學習與現實生活中的應用，從而加強對日語及相關知識的記憶和理解。

6. 結論與未來方向

本研究成功設計了一個數位雙生劇場作為整合使用者與實體物件的數位學習系統，提供沉浸式且互動性更高的學習環境。教師透過該系統將課本內容轉化為情境教材，學習者則在數位實境中透過操作實體物件來學習相關知識，同時透過系統觀察自身的學習表現進行反思學習，以提升學習動機與學習成效。結果表明，相較於僅使用滑鼠操作虛擬物件的方法，這種操作實體物件的學習模式能夠更有效提升學習者在基礎知識、餐桌設定和餐桌服務等方面的學習成效。問卷結果顯示，在內在目標導向、任務價值以及學習表現和自我效能方面，使用數位雙生劇場的學習者感受略高於對照組。此外，學習者對數位雙生劇場持正面態度，有助於對實際的使用情境建立聯繫。

未來研究在應用上能將數位雙生劇場推廣到其他需要實體操作的教學領域，如醫學臨床操作、運動科學等，以驗證學習模式的有效性並使研究結論更具參考價值。這些改進將有助於優化數位雙生劇場的學習經驗，並擴大其在不同領域和不同年齡群的學習者中的應用範圍。

致謝

本研究感謝國科會經費支持，計畫編號：MOST 110-2511-H-008 -004 -MY3；MOST 111-2410-H-008 -012 -MY3；NSTC 112-2811-H-008-006。

參考文獻

Ahuja, K., Shah, D., Pareddy, S., Xhakaj, F., Ogan, A., Agarwal, Y., & Harrison, C. (2021, May). Classroom digital twins with instrumentation-free gaze tracking. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1-9).

Bacca Acosta, J. L., Baldiris Navarro, S. M., Fabregat Gesa, R., & Graf, S. (2014). Augmented reality trends in education: a systematic review of research and applications. *Journal of Educational Technology and Society, 2014, vol. 17, núm. 4, p. 133-149.*

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *1989, 18*(1), 32-42.

Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual reality sickness: a review of causes and measurements. *International Journal of Human–Computer Interaction, 36*(17), 1658-1682.

Chang, W. Y. (2010). *Taiwanese Elementary Students' Motivation and Attitudes toward Learning English in English Village Program in Kaohsiung County*. ERIC.

Chiang, P. Y., Lee, C. H., Lin, S. H., Fan, C. Y., Chen, Y. H., & Chen, G. D. (2018, June). A Situational Comic Play Drama Learning System to Enhance Students' Learning Motivation. In *EdMedia+ Innovate Learning* (pp. 1269-1274). Association for the Advancement of Computing in Education (AACE).

De Ribaupierre, S., Kapralos, B., Haji, F., Stroulia, E., Dubrowski, A., & Eagleson, R. (2014). Healthcare training enhancement through virtual reality and serious games. *Virtual, augmented*

reality and serious games for healthcare 1, 9-27.

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. MIT press.

Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education* (Vol. 1, No. 133, pp. 10-1007).

Gerup, J., Soerensen, C. B., & Dieckmann, P. (2020). Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. *International journal of medical education, 11*, 1.

Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. *Transdisciplinary perspectives on complex systems: New findings and approaches*, 85-113.

Hengki, H., Jabu, B., & Salija, K. (2017). The effectiveness of cooperative learning strategy through English village for teaching speaking skill. *Journal of Language Teaching and Research*, *8*(2), 306-312.

Henry, D., & Furness, T. (1993, September). Spatial perception in virtual environments: Evaluating an architectural application. In *Proceedings of IEEE virtual reality annual international symposium* (pp. 33-40). IEEE.

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education, 95*, 174-187.

Lorenzo-Alvarez, R., Rudolphi-Solero, T., Ruiz-Gomez, M. J., & Sendra-Portero, F. (2020). Game-Based learning in virtual worlds: a multiuser online game for medical undergraduate radiology education within second life. *Anatomical sciences education, 13*(5), 602-617.

Miller, G. A., & Gildea, P. M. (1987). How children learn words. *Scientific American,257*(3), 94-99.

Ornellas, A., Falkner, K., & Edman Stålbrandt, E. (2019). Enhancing graduates' employability skills through authentic learning approaches. *Higher education, skills and work-based learning, 9*(1), 107-120.

Palmer, C., Roullier, B., Aamir, M., McQuade, F., Stella, L., Anjum, A., & Diala, U. (2022). Digital Twinning remote laboratories for online practical learning. *Production & Manufacturing Research, 10*(1), 519-545.

Pintrich P., Smith D., Garcia T., & McKeachie W. (1991). A manual for the use of the motivated strategies for learning Questionnaire (MSLQ).

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology, 94*, 3563-3576.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review, 9*, 625-636.

Zandavi, S. M., & Chung, V. (2018, July). Augmented reality for remote laboratory improving educational learning: Using elevated particle swarm optimization in object tracking scheme. In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-6). IEEE.

基于 Web of Science 的人工智能伦理研究关键词可视化分析

Visual Analysis of Keywords in Artificial Intelligence Ethics Research Based on Web of Science

郭宝心[1*]，郑艺妍[1]，廖林锟[1]，张星语[1]
[1]华南师范大学教育信息技术学院
[*] 2958714843@qq.com

【摘要】 随着人工智能技术在社会的广泛应用，人工智能伦理问题日益成为关注焦点。为了深入了解国外人工智能伦理的研究热点与前沿，本研究从 Web of Science 数据库中筛选了 2017-2023 年间有关 E-AIED 的研究成果，并对其关键词进行可视化分析，从年度发文量、作者合作图谱、机构共线图谱、关键词共现图谱、关键词聚类图谱及关键词突现 6 方面入手，梳理并呈现了国际人工智能伦理的研究主题和发展趋势。最后，提出我国应构建交叉融合型学术研究共同体、编制智能教育应用伦理规范指南、加强数据和人工智能的法规监管、强化师生人工智能伦理素养等对策和建议。

【关键词】 人工智能；伦理研究；可视化分析

【基金项目】 本研究受广东省哲学社会科学规划 2024 年度青年项目《数字化转型视域下教师生成式人工智能教育应用素养研究》（项目编号：GD24YJY06）资助。

Abstract: *This study selects the Web Of Science database to visually analyze the models of artificial intelligence ethics research from 2017 to 2023. The research themes and development trends of international artificial intelligence ethics are sorted out and presented from six aspects: annual publication volume, author cooperation graph, institutional collineation graph, keyword co-occurrence graph, keyword clustering graph and keyword emergence. Finally, put forward the corresponding countermeasures and suggestions.*

Keywords: Artificial intelligence, Ethical research, Visual analysis

1.引言

习近平同志指出：“人工智能是引领新一轮科技革命和产业革命的重要驱动力。（习近平，2019）”人工智能不仅是一项技术，还是未来社会变革及教育改革的重要影响因子。人工智能的发展不仅对科技领域产生深远影响，而且在社会、经济和文化等多个领域都带来了深刻的变革。然而，随着人工智能技术的飞速发展，一系列与人工智能相关的问题与挑战逐渐浮出水面。王佑镁等人（2023）提出：以 ChatGPT 为代表的新一代人工智能虽然为教育带来了前所未有的机遇，但与此同时，也引发了一系列伦理问题，如算法依赖的潜在风险、学术失信的威胁、技术伦理与道德的挑战，以及信息安全方面的潜在隐患。如何有效规避人工智能的伦理风险，使其更好地发挥作用为社会服务，成为当前国内外的研究重点。胡小勇等人（2022）提出当前教育人工智能伦理建设存在缺乏精准的规范与指引、人文关怀与价值引领缺位、技术自身存在局限、学习资源建设粗放等难题；黄荣怀（2023）认为智能技术链式突破所带来的不确定性、模糊性与风险性引发了社会对教育系统变革的普遍担忧，这迫使人们对人才定位、教育诚信、核心素养、科技伦理与技术治理等问题进行深刻反思。国外学者如

Wayne 等人（2021）也开始研究教育中的人工智能伦理问题。人工智能伦理已备受关注，但是目前国外人工智能伦理的研究热点、趋势等还有待进一步探讨。

因此，本研究将利用知识图谱分析、聚类分析等定量分析方法，借助 Web of Science 权威数据库对国外人工智能伦理的相关文献进行可视化呈现。通过对人工智能伦理的年度发文量、作者合作图谱、机构共线图谱、关键词共现图谱、关键词聚类图谱及关键词等进行分析，尝试揭示国外教育人工智能的研究热点及趋势，为我国教育人工智能的研究与发展提供参考。

2.数据来源与研究方法

2.1. 数据来源

本研究基于英文数据展开，英文数据来源于 Web of Science（WOS） 数据库核心数据合集，检索条件为“((TS = (education artificial intelligence) OR TS = (artificial intelligence in education)) OR TS = (AIED)) AND TS = (ethic)”，时间设置为“2017-01-01 to 2023-12-31”，搜索时间为 2024 年 1 月 1 日，共检索到 226 篇有效文章，剔除社论材料、修订和信函等共 4 篇，得到 221 篇有效文章。

2.2. 研究方法与工具

本研究主要采用知识图谱分析、聚类分析等定量分析方法。研究工具采用陈超美教授基于 Java 语言开发的 CiteSpace 知识图谱可视化软件，主要基于共引分析理论（co - ciation）和寻径网络算法（path Finder），通过对特定领域文献集合进行计量分析，揭示学科领域的演化路径及其知识转折点。不仅如此，它还可以通过绘制一系列可视化图谱进一步剖析学科演化的潜在动力机制，并探测学科发展的前沿趋势，为学科发展提供重要的参考和洞察。（陈悦等，2015）。采用 CiteSpace V6.1.R3 软件对 WOS 中 2017—2023 年关于 E-AIED 共 221 篇相关文章进行可视化分析。文献来源(Term Source)选择“Title”“Abstract”“Author Keywords（DE）”和“Keywords Plus（ID）”；文献时区分割（Time Slicing）中，时间跨度设为“From 2017 to 2023”，单个时区长度(Years Per Slice)设为 1 年；节点阈值(Selection Criteria)设为“Top N = 50”；采用寻径网络算法(Pathfinder)精简和合并网络；可视化选择“Static”和“Show Merged Network”。

2.3. 研究思路

首先，对获取到的 221 篇有效文章从发文量的角度进行分析，得到国内外关于 E-AIED 的发文趋势。其次，对作者合作情况进行分析，对研究机构分布进行共线分析，得到国内外活跃的学者和学术共同体。然后，对关键词进行共线分析和聚类分析，形成国内外研究热点主题的可视化图谱。最后，对关键词进行突现分析，深入分析国内外该领域的研究前沿。

3.数据统计结果

3.1. 年度发文量

2017—2023 年期间 WOS 检索到的有关 E-AIED 的研究成果数量及增长趋势如图 1 所示。从年度整体发文量来看，发文量逐年增长，呈指数型趋势；从年度发文量的增长速度来看，增速最快的是 2023 年，这可能与以 ChatGPT 为代表的大语言模型有关。ChatGPT 兴起后，在国际上引起了多方关注，各国教育主管部门积极做出回应，有关学者进行了多角度的深入探讨，形成了一大批研究成果。

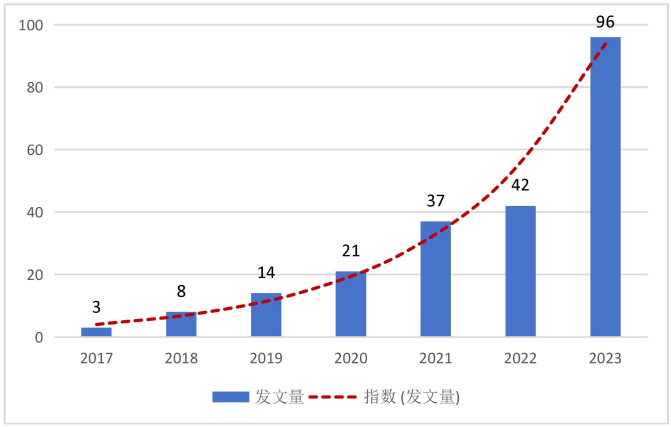


图 1 国外 E-AIED 年度发文量

3.2. 作者合作图谱

作者合作图谱可以反映某研究领域的活跃学者及与其有密切合作关系的学者。本研究将网络节点（Note types）选择为“Author”，形成国外作者合作图谱，如图 2 所示。其中，节点大小代表作者的研究成果数量，节点之间的连线代表了不同主体之间的合作关系。从图 2 可以看出，国外 E-AIED 已形成多个作者合作体，其中 Holmes、Eicher、Bektik 各发表 E-AIED 文章 2 篇。

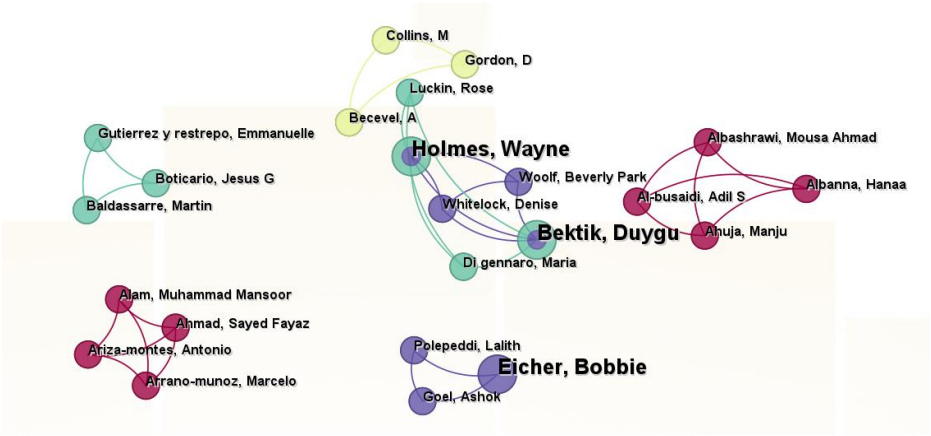


图 2 国外 E-AIED 作者合作图谱

3.3. 机构共线图谱

机构合作图谱能够反映该领域活跃的机构及其密切合作的机构。本研究将网络节点（Note types）选择为“Institution”，得到国外机构合作图谱，如图 3 所示。

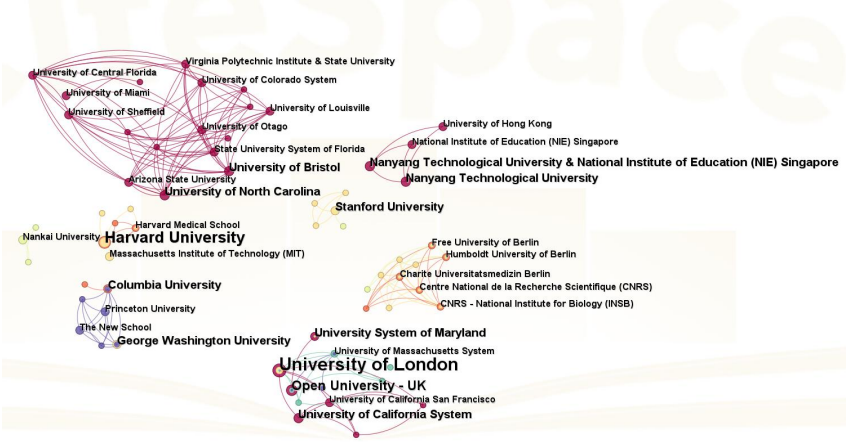


图 3 国外 E-AIED 机构合作图谱

从图 3 可以看出，国外 E-AIED 研究的高产机构单点结构少，稳定结构多。其中，国外有关 E-AIED 研究的高产机构主要有布里斯托大学、哈佛大学、乔治华盛顿大学、伦敦大学、斯坦福大学等。这些机构在人工智能伦理研究中产出大量成果，组成了国外 E-AIED 研究的科学中心。从单位的性质看，国外 E-AIED 的研究几乎集中在高校。从整体网络结构来看，以这些高产机构为核心已经形成了有关 E-AIED 研究的不同群落，有的学术共同体节点多，体系庞大；还有的学术共同体节点相对较少，它们都为人工智能伦理的相关研究做出了极大贡献。

3.4. 关键词共线图谱

高频关键词能够反映某一领域的研究热点，中介中心度能够体现该领域的突出热点。本研究将网络节点（Note types）选择为“Keywords”，得到国外关键词共线图如图 4 所示。在该图中，年环越大代表其关键词频率越高；图谱内部不同年轮颜色代表对应的年份；外部紫色的年环代表高中介中心度的关键节点。高中介中心度节点控制其他节点的练习强弱，在整个网络图谱中具有极为重要的地位。

图 4 是国际 E-AIED 关键词共线图谱，共产生节点（N）448 个，连线（E）1389 条。从高频词来看，关键词有：“人工智能（artificial intelligence，120 次）”“机器学习（machine learning，21 次）”“大数据（big data，13 次）”“教育（education，12 次）”“技术（technology，11 次）”“人工智能伦理（ai ethics，11 次）”“伦理（ethics，10 次）”“挑战（challenge，7 次）”等。从节点中心度来看，“人工智能伦理（ai ethics，0.17）”占核心地位，相对突出的研究热点还有“技术（technology，0.15）”“机器学习（machine learning，0.12）”“人工智能识别（ai literacy，0.11）”“大数据（big data，0.09）”“信任（trust，0.05）”等。由此可见，国外研究 E-AIED 更关注技术伦理与伦理教育，如技术、机器学习、大数据、人工智能识别、对技术的信任度等。

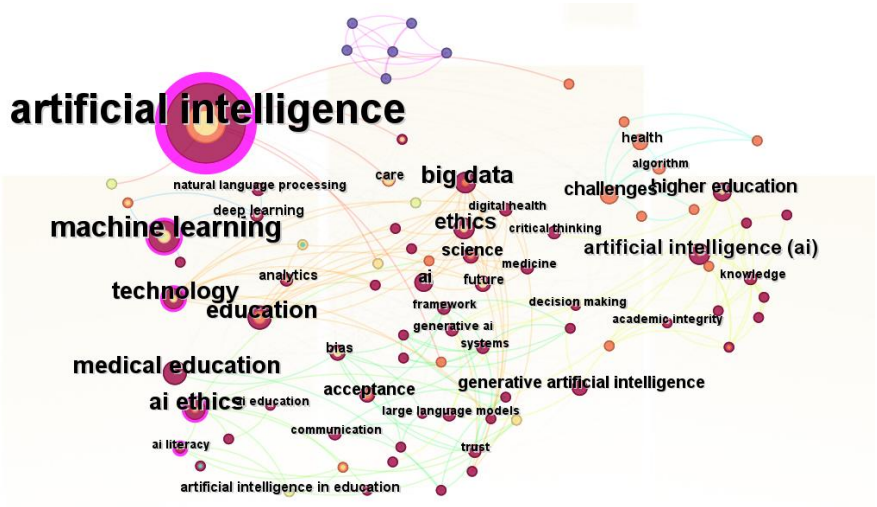


图 4 国外 E-AIED 关键词共线图谱

3.5. 关键词聚类图谱

E-AIED 关键词共线基础上，本研究使用对数似然率算法（Log-Likelihood Rate, LLR）进行聚类，聚类名称选取聚类中 LLR 算子取值最高的特征，并设置展示 9 个最大值聚类标签，由此得到国外 E-AIED 关键词聚类图谱，如图 5 所示。李杰等（2016）认为网络社团结构是否团结取决于网络模块化评价指标 Modularity(Q)值是否大于 0.3，并将 Silhouette(S)值大于 0.7 视为该聚类的结果具有高信度。

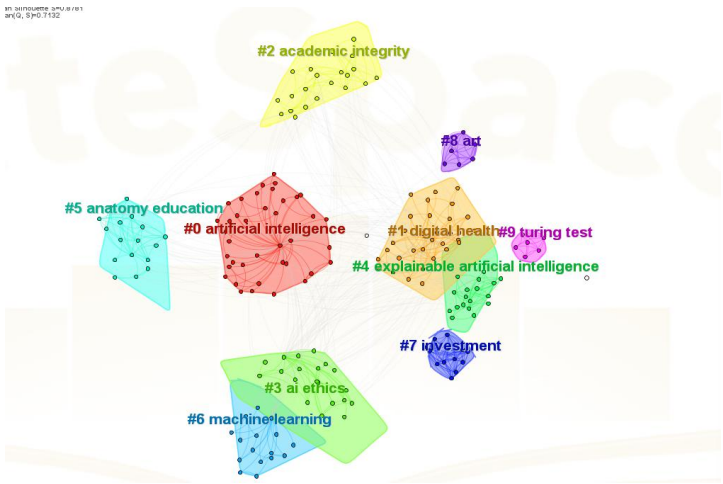


图 5 国外 E-AIED 关键词聚类图谱

本研究根据聚类分析图谱将国际 E-AIED 研究内容分为两个主题：伦理教育、技术伦理等（如表 1）。

伦理教育主题聚类包括：#1 数字健康（digital health）、#2 学术诚信（academic integrity）、#3 人工智能伦理（ai ethics）、#4 可解释人工智能（explainable artificial intelligence）、#5 解剖学教育（anatomy education）、#9 图灵测试（turing test）。人工智能素养教育能让学生为人工智能时代的到来做好准备，Kong 等（2022）将人工智能素养教育在高中实践，并证明了人工智能素养教育能提高学生的人工智能观念和伦理意识。教育人工智能的伦理教育研究多聚焦于高等教育的医学教育，如 Katznelson 等（2021）认为人工智能在医学领域的应用引发了许多伦理挑战（知情同意、偏见、安全性、透明度、患者隐私和分配的道德问题），有必要在医学院开设健康人工智能伦理教育。人工智能的态度也是研究热点，如 Latham 等（2019）从道德和法律方面，考察了公众对教育中的人工智能的看法，以期为 E-AIED 提供参考。

技术伦理主题聚类包括：#0 人工智能（artificial intelligence）、#6 机器学习（machine learning）、#7 投资（investment）、#8 艺术（art）。诸如机器人、区块链、大数据、人工智能等技术为教育带来了无限可能，但是也带来许多伦理隐私问题，特别是数据带来的问题（Holmes 等，2021），如 Godwin-Jones（2021）探讨了技术应用于教育带来的学习分析、数据挖掘和数据伦理等问题，Cernadas 等（2021）强调了机器学习带来的算法偏见和性别偏见问题，Souza（2018）阐释了大数据、机器学习和人工智能技术带来的隐私泄露和人权威胁问题。

表 1 国外 E-AIED 研究主题及特征词

主题	聚类标签	特征词
伦理教育	#1 数字健康（digital health） #2 学术诚信（academic integrity） #3 人工智能伦理（ai ethics） #4 可解释人工智能（explainable artificial intelligence） #5 解剖学教育（anatomy education） #9 图灵测试（turing test）	教师（teacher），教育（education），高等教育（higher education），教育人工智能（Artificial Intelligence in Education），在线学习（online learning），学习分析（learning analytics），课程（course），学生（student），工具（tool），医学教育（medical education），健康（health），设计（design），挑战（challenge），技术（technology），接受度（acceptance）

伦理技术	#0 人工智能（artificial intelligence） #6 机器学习（machine learning） #7 投资（investment） #8 艺术（art）	大数据（big data），未来（future），工程教育（engineering education），机器学习（machine learning），算法偏见（algorithm bias）

3.6. 关键词突现图谱

关键词突现图谱能够展示某一领域的前沿和发展趋势。关键词的突现性越高，表示该关键词在特定时间段内出现频次变化率较大。简而言之，这意味着在某一领域中，学者之前未特别关注该关键词，但该关键词对该领域产生了重要影响。随着研究的深入，该关键词会在某个时间点突然出现，并成为该研究领域的热点关键词。

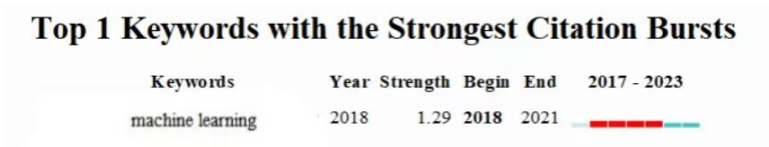


图 6 国外 E-AIED 关键词突现图谱

2017—2023 年 E-AIED 的研究显示（如图 9），突现性最高的是机器学习（machine learning），突现性为 1.29，突现年份为 2018 年。人工智能的核心是机器学习。机器学习是以大数据为基础，以算法为核心的人工智能技术，在教育领域里研究学生学习方面起着重要作用。它主要应用于学生学习行为建模、预测学生学习行为、预警失学厌学风险、提供学习支持和评估，以及资源智能化推荐等方面（陈颖博等，2019）。因此，机器学习在教育应用过程中的伦理可能是未来国际 E-AIED 研究的前沿。

4. 讨论和启示

本研究通过对 2017-2023 年间 E-AIED 的研究成果进行梳理，归纳总结出国外 E-AIED 的研究热点和前沿，这对我国建设相关的人工智能伦理框架也具有一定的借鉴意义。

4.1. 构建交叉融合型学术研究共同体

从作者合作图谱和机构合作图谱结果来看，国外形成了多个 E-AIED 作者合作体。因此，我国也应加强与国际教育人工智能专业化组织的合作，共同举办人工智能与教育领域的国际会议，为我国学者提供借助国际学术交流平台的机会。另外，教育人工智能涉及多个学科领域的研究成果，如学习科学、心理学、社会学、计算机科学等，以学科融合为引擎，打造来自全世界的高效的研究共同体。

4.2. 编制教育人工智能伦理规范指南

从关键词共线和聚类图谱结果来看，国外学者聚焦实践层面进行了诸多有关 E-AIED 的探索。但是，当前我国关于 E-AIED 内涵和框架的研究正处于多种伦理原则相融合的规范架构与协同运行的发展阶段，且这些研究大多仍停留在“呼吁”阶段，提出的应对方案还浮在表面。因此，应在现有的人工智能伦理规范体系下，尽快制定我国 E-AIED 规范框架和指南，为研发者、使用者和管理者在研发、应用和管理人工智能中提供明确的行动指南，以确保师生的数据安全和人身健康。

4.3. 加强数据和人工智能的法规监管

从关键词突现图谱结果来看，在 E-AIED 研究领域，机器学习相关的伦理问题研究有可能成为研究前沿。教育与技术的研发、使用和管理环环相扣，教育人工智能相关的大规模数据收集和分析涉及教育相关主体的隐私（Hong 等，2022），机器学习中的数据使用有可能威胁到教师或学生个人数据的隐私与安全。因此，相关智能技术部门需要提高技术发展水平并健

全相关教育人工智能数据风险问责机制，促进智能技术实现的规范化，数据使用监管程序的标准化。首先，加强与教育部门、科技部门和法律监管部门等相关部门的合作，建立跨部门的数据共享与监管机制。其次，为了规范技术的开发和应用，相关部门应制定教育人工智能法律法规以约束相关从业者的规范应用，同时，法规需明确在技术开发和应用过程中的伦理责任细则及相关法规惩戒制度，相关人员要对非法开发、使用和销售人工智能的行为追究法律责任和后果。最后，构建专业的数据隐私监管组织或机构，以发展和提升在教育人工智能应用过程中的数据监管质量，搭建教育人工智能数据监管体系。

4.4. 提升师生人工智能伦理素养水平

从关键词聚类图谱结果来看，国际学者侧重于伦理教育和技术伦理的主题研究。在提升学生人工智能伦理素养方面，一大批欧美学校陆续开设了人工智能伦理相关课程，如剑桥大学开设“人工智能伦理和社会”硕士课程，旨在培养学生在人工智能时代所必备的专业知识与技能（沈苑等，2023）。相比之下，我国高校人工智能伦理相关课程的开设还非常欠缺（古天龙，2022）。在提高学生人工智能伦理素养方面，我国可以借鉴国外人工智能伦理课程的经验，开发类似教育资源，从专业的角度助力学生人工智能伦理素养的发展。在提高教师人工智能伦理素养方面，开展以人工智能为主题的教师专业发展活动，在帮助教师学会在课堂中使用技术的同时，也让教师警惕对运用智能技术教学判断的过分依赖。

参考文献

习近平.（2019）.推动新一代人工智能健康发展更好造福世界各国人民.人民日报.

王佑镁,王旦,梁炜怡 & 柳晨晨.(2023).“阿拉丁神灯”还是“潘多拉魔盒”：ChatGPT 教育应用的潜能与风险.**现代远程教育研究(02)**,48-56.

古天龙(2022).人工智能伦理及其课程教学.**中国大学教学(11)**,35-40.

李杰 & 陈超美(2016).**CiteSpace 科技文本挖掘及可视化(第三版)**.首都经济贸易大学出版社.

沈苑,胡梦圆,范逸洲 & 汪琼.(2023).可信赖人工智能教育应用的建设路径与现实启示——以英国典型举措为例.**现代远程教育研究(04)**,65-74.

陈超美,陈悦 & 侯剑华(2009).CiteSpace Ⅱ:科学文献中新趋势与新动态的识别与可视化.**情报学报,2009,28(3)**:401-421.

胡小勇,黄婕,林梓柔 & 黄漫婷.(2022).教育人工智能伦理:内涵框架、认知现状与风险规避.**现代远程教育研究(02)**,21-28+36.

黄荣怀.(2023).人工智能正加速教育变革:现实挑战与应对举措.**中国教育学刊(06)**,26-33.

Godwin-Jones, R. (2021). Big data and language learning: Opportunities and challenges. Language Learning & Technology, 25(1), 4–19.

Hong, Y., Nguyen, A., Dang, B., & Nguyen, B. T. (2022). Data Ethics Framework for Artificial Intelligence in Education (AIED). 2022 International Conference on Advanced Learning Technologies (ICALT).

Katznelson, G., & Gerke, S. (2021b). The need for health AI ethics in medical school education. Advances in Health Sciences Education, 26(4), 1447–1458.

Kong, S. C., Cheung, W. M. Y. & Tsang, O. (2022). Evaluating an artificial intelligence literacy programmer for empowering and developing concepts, literacy and ethical awareness in senior secondary students. Education and Information Technologies.

Latham, A., & Goltz, S. (2019). A survey of the general public’ s views on the ethics of using AI in education. Artificial Intelligence in Education, 194–206.

Souza, R. R. (2018). Algorithms, future and digital rights: Some reflections. Education for Information, 34(3), 179–183.
Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T. T., Shum, S. B., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in Education: Towards a Community-Wide framework. International Journal of Artificial Intelligence in Education (Print), 32(3), 504–526.

深度学习视域下生成式人工智能支持小学生编程能力的实证研究

An Empirical Study on Generative Artificial Intelligence Supporting Primary School Students' Programming Ability from the Perspective of Deep Learning

王丹妮，黄瑞婷，李铭壹，张新立*
温州大学 STEM 教育研究中心
wangdanni0899@163.com

【摘要】 本研究将深度学习与编程学习相融合，构建深度学习视域下 GAI 支持的编程教学模式，以 60 名五年级小学生为研究对象，对该编程教学模式进行分析，结果表明深度学习视域下 GAI 支持的编程教学模式有助于提高学生的编程能力。

【关键词】 编程教育；深度学习；生成式人智能；编程能力；教学模式

Abstract: This study integrated deep learning and programming learning to propose a GAI-supported programming teaching model based on deep learning theory and explored its effect on students' programming ability. A total of 70 fifth-grade students were enrolled in this study. The results showed that the proposed teaching model could promote students' programming ability.

Keywords: Programming education, Deep learning, Generative AI, Programming skills, Teaching mode

1. 引言

作为智慧时代的一项重要技能，编程能力逐渐被世界各国所重视，日益成为实施创新性人才培养战略的重点。各国也相继采取相应措施将编程能力的培养落实到国家政策。近年来，我国也颁布了一系列政策文件，强调了小学生编程能力培养的重要性。2018 年，教育部颁布《教育信息化 2.0 行动计划》，对中小学开展编程教育提出新的要求；同年，教育部颁发的《新课标》指出大幅度提升学生在编程、计算思维、算法方面的思维要求。然而，当前编程教育受传统应试教育影响，课堂偏重知识的考查，存在着学生主体地位缺失、忽视个性化需求等问题（冯文龙，2021）。课程内容相对比较抽象和枯燥,尤其对于小学生来说，时间一长易造成学习兴趣低下，不利于编程能力的培养，从而影响学生此后编程的学习。

目前，生成式人工智能（GAI）在编程教学的应用备受关注，已有研究者将 GAI 运用于编程教育， GAI 能通过提供建议、检测错误和自动生成代码，帮助学生编写更高效、更准确的代码，并减少完成编程作业所需的时间和精力（Yilmaz et al，2023）。此外， GAI 能为学生提供个性化的支持和回馈，从而提高他们的参与度和动力。然而，GAI 还不足以支撑脱离教师的人机协同学习，需要融入一个有效的策略（游佳婧等人，2023）。因此，本研究将深度学习与编程学习相融合，构建深度学习视域下 GAI 支持的编程教学模式来支持小学生的编程学习。

2. 文献综述

2.1 小学生编程能力培养模式及存在的问题

世界各国均试图抢占人工智能时代的教育发展先机，青少年编程教育更是寄托着国家未来的希望。当前，编程课程已开始逐渐走进小学课堂，以培养小学生的编程能力。在小学生编程能力培养上，Shabalina 等人强调利用游戏来学习编程概念和结构，培养编程技能(Shabalina et al, 2017)；Jun Tan 等人将案例教学法应用于编程教学中，提高了学生的编程能力。孙立会等人强调问题性学习和项目化学习是培养小学生编程能力的重要途径 (Sun & Zhou, 2022)。Wang 和 Hwang 构建了基于问题提出的团队学习模式，并且还研究了在线同伴评估教学策略的可行性，旨在提高小学生的编程技能(Wang & Hwang, 2017)。然而，当前的编程教学还存在着小学生主体地位缺失、忽视个性化需求、偏重编程知识考察等诸多问题，引入相应的策略或方法培养他们的编程能力，提高学习效果是有必要的。因此，本研究构建了深度学习视域下 GAI 支持的编程教学模式，开发相应的教学案例，并探索在参与教学活动后，两组学生的编程能力是否有存在差异。

2.2 GAI 教育应用的价值

Sarang 等人通过探索学习者对 ChatGPT 产生的反应的满意度，以及学习者在学习者执行英语学习任务评估 ChatGPT 英语语言学习的潜力，结果表明学习者积极参与学习，并且在英语学习中取得较好的效果（Kalelioglu & Gülbahar，2014）；Yilma 等人研究表明 ChatGPT 纳入的编程中，为学生提供关于编程主题的清晰解释、编程示例以及应用程序，减少在编码上的时间，使他们能够分配更多的时间参与算法思维过程；并且还有利于提高了学生的计算思维技能、编程自我效能感和上课动机（Yilmaz et al，2023）。童大振等人将 ChatGPT 应用于物理学习当中，发现虽然能够解决大部分概念理解问题（童大振&任红梅，2023）。

目前，GAI 已被应用到不同的学科教学，如英语、数学、物理、编程学习等，并对学生的学科学习有着显著的提升效果。可见，GAI 对提升学生的学习表现具有重要的作用。

2.3 深度学习及实现途径

深度学习是一种学习者主动的、批判性的学习方式, 使学习者理解性地学习、批判性地思考、主动地建构知识、有效地迁移知识从而解决问题，是实现有意义学习的有效方式（车苗苗&李少梅，2020）。与浅层学习相对应, 浅层学习忽视对知识的深层加工、深度理解及长期保持, 更无法实现知识建构、迁移应用及问题解决。当前的教学很多处于浅层学习阶段，已有研究表明可以利用策略和方法来促进学生走向深度学习（张浩&吴秀娟，2012）。

安富海教授提到深度学习的教学策略正是在深入研读深度学习理论的基础上,通过批判当前课堂学习中存在的浅层学习问题而提出的一种引导教师调整教师理念和教学行为的建议（安富海，2014）；杨嘉辉学者认为在小学生编程学习过程当中，策略的融入有利于深度学习的有效实施，让学生能够理解编程概念与原理，活用编程原理与技能，实现计算思维提升（杨嘉辉，2022）。可见，策略融入能促进深度学习的真正发生，对教育教学有重要的意义。

3. 深度学习视域下 GAI 支持的编程教学模式构建

GAI 与深度学习融合，构建出深度学习视域下 GAI 支持的编程教学模式(见图 1)。基于此模式，开发符合小学生身心规律与现实生活的教学案例。

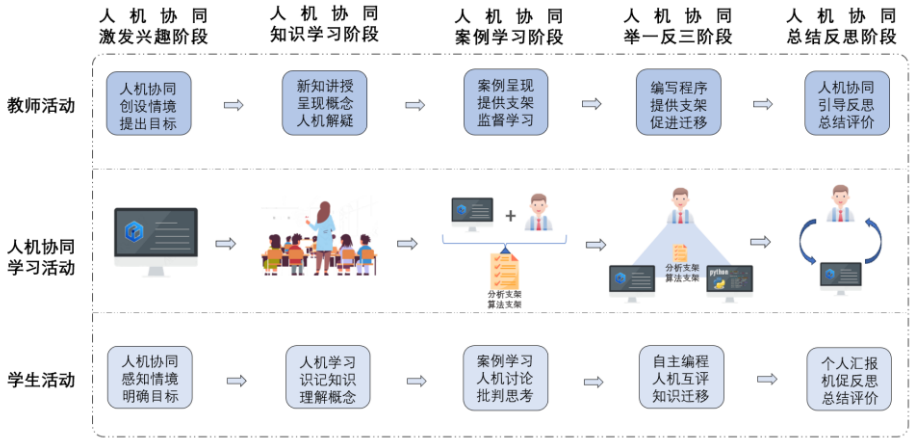


图 1 深度学习视域下 GAI 支持的编程教学模式

3.1 人机协同激发兴趣阶段

情境导入阶段是小学生编程能力培养的开端和基础阶段。在课堂中，教师选择有意义的现实生活问题为主题，通过学生与 GAI 互动参与和交流，形成初步联结，将生活情境转换为驱动问题和学习目标，以增强学生对社会问题解决的感知与敏锐度。

3.2 人机协同知识学习阶段

知识学习阶段是帮助学生完成深层次的学习和内在质量的提升的前提，起到承前启后的作用。在此阶段，教师讲解编程的基本知识，学生对于有疑问的知识可以及时寻求 GAI 进行解答，可以有效减少学生学习“掉队”的情况。

3.3 人机协同案例学习阶段

案例学习阶段是从浅层学习到深度学习的核心阶段。该阶段，学生根据学习支架向 GAI 提出问题，对其给出的回答进行批判性地思考与讨论，使学生从简单的识记逐步走向对知识的理解，验证回答的准确性，分析错误之处，进行调试，批判性思考，使学生走向深度学习。

3.4 人机协同举一反三阶段

举一反三阶段是承载着案例学习完成后，促进促进知识迁移的关键。在该阶段，学生让 GAI 提供一道与第三阶段例题类似的练习题，独自完成所有的任务，并且将编写出来成程序发送给 GAI 进行验证和修改，对于有疑问的知识进一步进行学习与讨论，促进知识的内化。

3.5 人机协同总结反思阶段

该阶段是基于贡献、联结和品德培养中促进学生的问题评估与反思能力。在该阶段，不仅仅包含教师的总结与评价 还涉及学生对GAI给出回答的评价和GAI基于对话的总结与评价。多维度的总结与评价，帮助学生梳理知识要点，获得多方回馈，进而深入的反思。

4. 研究设计

4.1 研究对象及流程

本研究选取了中国某小学五年级 60 名学生。实验组 30 人（16 名男生，14 名女生），采用深度学习视域下 GAI 支持的编程学习；对照组 30 人（15 名男生，15 名女生），采用 GAI 支持的常规编程学习，两组由同一个教师开展教学，不同在于实验组融入了学习支架，使学生在 GAI 的支持下走向深度学习；而对照组仅在 GAI 支持下进行学习。实验共持续 5 周，第 1 周学生进行前测和技术学习，第 2~4 周进行教育干预，第 5 周进行总结和后测。

4.2 学习活动设计

本研究选用新浙教版《信息科技》五年级下的 3 课时学习内容：1. 算法的设计；2.算法的验证；3.算法的应用。在 40 分钟的学习活动中，两个组的被试都将完成 2 个任务，且根据深度学习理论和布鲁纳认知结构理论，开展持续三周学习。

4.3 测量工具

小学生编程能力测试由专家组共同制定，共 10 题（10 分）。参与者须在 20 分钟内完成。

5. 结果分析

通过正态性检验，发现两组学生的前测成绩的($p = 0.235 > 0.05$)，呈符合正态分布。因此，进一步对实验班和对照班做独立样本 t 检验，结果如表 1 所示，两个班的前测成绩无显著差异($p = 0.224 > 0.05$)，则说明两个班学生的编程水平相当。

然而，两组的后测成绩($p = 0.004 < 0.05$)，不符合正态分布。如表 1 所示， Mann-Whitney U 检验结果表明两组的后测成绩存在显著差异($U = 447$，$p = 0.048$)。实验组测试成绩显著高于对照组，说明深度学习视域下 GAI 支持的编程教学模式能有效促进小学生编程能力的提高。

表 1　两组后测的独立样本曼-惠特尼 U 检验

组别	*N*	Mean	SD	*z*	*p*
实验组	35	5.57	1.75	-1.986	0.048
对照组	35	4.69	1.45		

*$p < .05$

6. 结论

本研究从深度学习的视角构建了面向小学生的深度学习视域下 GAI 支持的编程教学模式。研究结果发现基于该模型设计的活动能够有效地提高学生的编程能力，说明了深度学习视域下 GAI 支持的编程教学模式的优势与有效性。

参考文献

梁云真,高思圆,刘瑞星 胡珂.(2021).小学生编程学习意愿影响因素的链式中介作用机制研究.现代远距离教育(06),29-37.

冯文龙.(2013).《Java 编程》教学改革研究.海南大学学报(自然科学版)(01),80-82+87.

童大振 任红梅.(2023).ChatGPT-3.5 解决物理问题的表现研究.中学物理(09),11-14.

游佳婧,田兴 & 丁鼎.(2023).基于 ChatGPT 的初中数学学习初探.中学数学月刊(05),63-67

车苗苗 李少梅.(2020).指向深度学习的 STEM 教学模式构建:理念与设想.教育探索(01),18-22.

张浩 吴秀娟.(2012).深度学习的内涵及认知理论基础探析.中国电化教育(10),7-11+21.

安富海.(2014).促进深度学习的课堂教学策略研究.课程.教材.教法(11),57-62.

杨嘉辉.(2022).例谈小学编程教育的深度教学策略.中国信息技术教育(22),66-68.

Kong, S. C., Chiu, M. M., & Lai, M. (2018). A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education. *Computers & Education, , 127,* 178-189.

Sun, L., Guo, Z., & Zhou, D. (2022). Developing K-12 students' programming ability: A systematic literature review. Education and Information Technologies, 27(5), 7059-7097.

Wang, X. M., & Hwang, G. J. (2017). A problem posing-based practicing strategy for facilitating students' computer programming skills in the team-based learning mode. Educational Technology Research and Development, 65, 1655-1671.

Tan, J., Guo, X., & Zheng, W. (2014). Case-based teaching using the Laboratory Animal System for learning C/C++ programming. *Computers & Education,* 77, 39-49.

从错误中学习策略对高中生编程成绩和计算思维的影响研究

The Effect of Learning-From-Mistakes Strategies on Programming Performance and Computational Thinking in High School Students

李琪[1]，翁秋灵[1]，方建文[1*]
[1]温州大学 教师教育学院
[*]beginnerfjw@163.com

【摘要】 计算思维是编程教育的核心素养。高中生学习编程经常会遇到编程错误频发和成绩欠佳等问题。研究显示，从错误中学习策略具有一定积极效果，但其对学生编程技能及计算思维的具体影响尚待深入分析。本研究以大陆东部某高二年级 87 名学生为样本，结合“Python 程序设计”课程实践，融入从错误中学习教学策略，并通过一学期的准实验验证了该策略的有效性。据此，教师需倡导学生从错误中学习并提高反思能力，结合实际设置问题情境以培养计算思维；同时构建兼顾质量和数量的多元评价体系，以全面提升学生的编程技能与计算思维层次。

【关键词】 从错误中学习；编程成绩；计算思维；教学策略

Abstract: *Computational thinking is vital for programming education. High school students often grapple with frequent programming errors and poor performance. Studies indicate positive effects of learning-from-errors strategies, but deeper analysis on their impact on programming skills and computational thinking is needed. This study, involving 87 Chinese high school sophomores, integrated such a strategy in Python courses. A semester-long quasi-experiment confirmed its effectiveness. Thus, teachers should encourage mistake-based learning, reflective skills, practical problem-solving, and adopt a balanced, multi-dimensional assessment system to enhance students' programming abilities and computational thinking.*

Keywords: learning from mistakes, programming achievement, computational thinking, teaching strategy

1. 前言

在当今的数字化时代下，大数据、物联网、深度学习、Chatgpt 等新技术层出不穷，这使人们面临的问题越来越复杂。因此，要求人们拥有更强的问题解决的能力。计算思维被认为是一种体现着计算机科学基本思想方法的特殊问题解决过程(Ezeamuzie & Leung, 2022)。

学者发现了计算机编程是一种培养计算思维的有效载体。然而，在高中信息技术课堂上依旧采用偏重知识传授的教学策略，这种教学策略在提高学生计算思维能力的效果上并不显著。过去的研究指出(姜宝琴, 2022)，从错误中学习策略可以提高学习成绩，但鲜有应用从错误中学习策略提升高中生编程成绩和计算思维的研究。

研究表明，从错误中学习策略虽然让学生经历了失败和挫折，但却能使学习者对目标概念有更深入的认识与理解(吕秋艳 & 刘徽, 2018)。因此，本研究构建了一种基于从错误中学习策略的任务驱动教学方法，探究该方法对高中生编程成绩和计算思维的影响。本文围绕所提策略提出了以下两个研究问题：

问题一：实验组的学生是否比控制组的学生在编程成绩方面表现更好？

问题二：实验组的学生是否比控制组的学生在计算思维方面表现更好？

2. 文献综述

2.1. 国内外计算思维的研究现状

通过查阅 Web of Science 等数据库，对计算思维相关文献的梳理如下：

（1）计算思维本质的探讨。2006 年，Wing 将求解问题和系统设计等活动作为计算思维的一般性定义(Wing, 2006)。2015 年，国际机构将算法思维和使用模拟的数据表示视为计算思维的操作性定义。

（2）计算思维的培养工具。Bender W.（2017）通过实证研究来分析不同工具对计算思维的影响，研究表明在编程环境下能够改善学生的计算思维能力和学习态度(Bender, 2017)。

因此，越来越多人关注到了学生计算思维的培养问题，但是鲜有结合教学实践去进行实证的研究，所以需要通过探究更多策略或方法去改善学生计算思维的培养情况。

2.2. 从错误中学习策略的研究现状

1979 年珀金森（Perkinson,1979）宣称，人类学习的本质是从错误中学习，2023 年张和费奥雷拉（Zhang & Fiorella, 2023）提出了“从自我生成的错误中学习”模型。通过梳理发现“错误”对学习的作用显著，认为显示错误有助于提升学习效果(Baltag, Li, & Pedersen, 2022; Rudolph, 2023)。目前，缺少探讨在编程教学中主动引入从错误中学习策略对学生计算思维的影响的实证研究。为此，本研究通过教学实验来深入探究这一问题。

3. 基于从错误中学习的任务驱动教学模式的构建

本研究在任务驱动教学模式的基础上引入从错误中学习的策略，结合高中编程课程教学现状与需求，将从错误中学习对提高编程成绩和培养计算思维的高中编程教学模式框架进行建构，如图 1 所示。



图 1 从错误中学习策略的任务驱动教学模式构建

（1）任务设计阶段。教师要提前明确课堂的教学内容，精心设计好每节课的学习任务。

（2）任务呈现阶段。教师需要把任务完整向学生讲解和演示一遍，然后让学生自主完成。本研究中，教师会针对课堂中的 Python 程序给学生布置纠错任务。

（3）任务实施阶段。实验班通过从错误中学习策略对学生进行干预，借助 AI 纠错工具找出程序中的错误，并对设计任务中的错误程序代码进行纠正，从而实现程序功能。对照班则采用常规教学，学生会一起讨论分析各自出现的错误，学生运行纠错后的程序，实现程序功能。

（4）任务评价阶段。教师对学生任务完成情况进行客观和具体的评价。

4. 从错误中学习策略对高中生编程成绩和计算思维的影响研究教学实践

4.1. 实验要素

4.1.1. 实验对象和学习内容

以中国大陆东部某高二年级学生为调查样本，对学生进行编程成绩和计算思维前测，根据学生反馈的数据，选择两个班级人数、知识水平和计算思维水平相当的班级为实验对象，

分为实验班和对照班，实验班采用基于 AI 的在线纠错模式，对照班采用常规教学模式。实验内容为《高中信息技术》课程中的 Python 部分。

4.1.2. 测量工具

（1）计算思维测量工具

本研究采用的计算思维量表是由 Korkmaz 等人编制的，该量表划分了 5 个方面，包含 22 道题目(Korkmaz, Çakir, & Özden, 2017)。

（2）学习成绩测量工具

在本实验中，选取教师编写的测试题测量学生的学习情况。

4.2. 实验设计与实施

实验流程如图 2 所示，本实验实施时间共计九周，第一周学生被分配完成 Python 水平和计算思维的前测，第二周，老师会选择并设计接下来每堂课的教学内容。从第三周开始，学生在两种不同的学习环境中进行学习，每节课老师都会讲解并演示课程任务，确保学生了解学习过程，其中实验班采用基于 AI 的在线纠错模式，具体教学过程包括：学生自主调试，找出自己解决不了的错误；学生把错误程序输入纠错工具，纠正错误信息；学生运行纠错后的程序，实现程序功能。对照班采用常规教学模式，具体教学过程包括：学生自主调试，找出自己解决不了的错误；学生呈现各自出现的错误，一起讨论分析(哪里错了，为什么错，怎么改正等)；学生运行纠错后的程序，实现程序功能。在课程教学中，老师会根据学生的表现给予及时反馈，并在任务完成后，对完成情况进行总结和反思。最后一周，课程学习结束，所有学生进行了 Python 水平和计算思维后测。

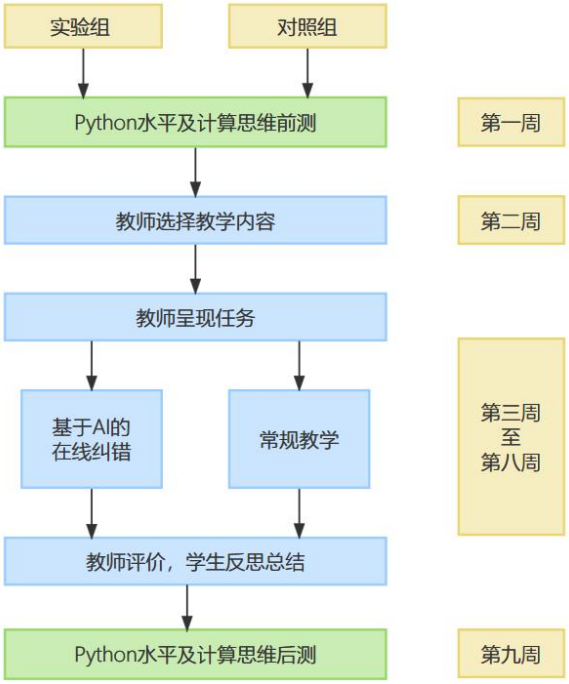


图 2 实验流程

5. 实验结果与讨论

5.1. 计算思维

本研究将各班获取的计算思维前测数据进行独立样本T检验，如下表所示：

表 1 实验班和对照班计算思维前测独立样本 T 检验

项目	组别	人数	均值	标准差	t	自由度	Sig

创造力	实验组	44	3.79	0.71	-0.80	85	0.430
	对照组	43	3.93	0.89			
合作能力	实验组	44	3.63	0.80	-1.31	85	0.200
	对照组	43	3.87	0.95			
批判性思维	实验组	44	3.53	0.80	-1.36	85	0.180
	对照组	43	3.79	0.93			
算法思维	实验组	44	3.87	0.85	1.12	85	0.270
	对照组	43	3.64	1.05			
问题解决能力	实验组	44	2.94	1.02	0.44	85	0.660
	对照组	43	2.84	1.18			

由表 1 内容可知，实验班和对照班在计算思维各方面的 sig 值都大于 0.05，说明实验班与对照班在计算思维能力的前测阶段并无显著性差异。因此，将各班获取的计算思维后测数据进行独立样本T检验，如下表所示：

表 2 实验班和对照班计算思维后测独立样本 T 检验

项目	组别	人数	均值	标准差	t	自由度	Sig
创造力	实验组	44	4.22	0.56	2.91	85	0.005
	对照组	43	3.80	0.76			
合作能力	实验组	44	4.12	0.56	3.45	85	<0.001
	对照组	43	3.55	0.95			
批判性思维	实验组	44	4.07	0.71	3.41	85	0.001
	对照组	43	3.51	0.83			
算法思维	实验组	44	4.22	0.56	4.34	85	<0.001
	对照组	43	3.52	0.90			
问题解决能力	实验组	44	3.68	0.78	2.69	85	0.009
	对照组	43	3.23	0.76			

由表 2 内容可知，实验班和对照班在计算思维各方面的 sig 值都小于 0.05，说明实验班和对照班计算思维后测成绩差异显著。正如 Rudolph(2023)所说，识别和纠正错误是有助于学习的。研究中，实验班学生的计算思维能力提高的事实，充分验证了该教学策略的有效性。

5.2. 编程成绩

本研究将各班前测考试成绩数据进行独立样本T检验，如下表 3 所示：

表 3 实验班和对照班前测成绩独立样本 T 检验

项目	班级	人数	均值	标准差	t	自由度	Sig
前测成绩	实验班	44	68.86	9.76	1.22	85	0.23
	对照班	43	66.09	11.44			

在实验班和对照班前测成绩独立样本T检验中，sig 值大于 0.05，说明在实施干预措施之前，实验班与对照班学生的编程成绩并无统计学意义上的显著差异。因此将各班后测成绩的数据进行独立样本T检验，如下表 4 所示：

表 4 实验班和对照班后测成绩独立样本 T 检验

项目	班级	人数	均值	标准差	t	自由度	Sig
后测成绩	实验班	44	75.25	15.20	2.77	85	0.007

对照班	43	66.19	15.27

实验班和对照班后测成绩独立样本T检验中，sig 值小于 0.05，说明在实施“从错误中学习策略”之后，实验班学生的编程成绩相较于对照班出现了显著提升。Bender(2017)学者就通过实证研究表明，在编程环境下能够改善学生的计算思维能力和学习态度。在本研究中，实验班学生通过主动理解错误原因以及优化解决方案，强化了自身的计算思维和逻辑推理能力。

6. 总结与展望

6.1. 研究总结

本研究将“从错误中学习”这一教学策略融入到 Python 程序设计课程中，利用现代 AI在线纠错工具作为辅助手段，验证了技术赋能下的教学策略对于提升学生编程技能的有效性。因此，在实际教学中，一线教师要注意引导学生多反思，利用好“错误”这一重要的教育资源。

6.2. 研究不足与展望

6.2.1. 样本容量

样本容量有限，因为样本随机选取，不能保证有很强的推广性。

6.2.2. 研究时长

实验为期一学期，但编程的学习需要动手实践和日积月累，尤其计算思维的培养，所以之后的研究可以延长研究时长，进行动态持续的跟踪调查。

参考文献

Baltag, A., Li, D. Z., & Pedersen, M. Y. (2022). A Modal Logic for Supervised Learning. *Journal of Logic Language and Information*, 31(2), 213-234. doi:10.1007/s10849-022-09359-w

Bender, W. (2017). The Sugar Learning Platform: Affordances for Computational Thinking. RED : revista de educación a distancia(54), 1-19. doi:10.6018/red/54/1

Ezeamuzie, N. O., & Leung, J. S. C. (2022). Computational Thinking Through an Empirical Lens: A Systematic Review of Literature. *Journal of Educational Computing Research*, 60(2), 481-511. doi:10.1177/07356331211033158

Korkmaz, O., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the co mputational thinking scales (CTS). *Computers in Human Behavior*, 72, 558-569. doi: 10.1016/j.chb.2017.01.005

Perkinson，H. J.（1979）. Learning from our mistakes[J]. ETC：A Review of General Semantics，36（1）：37-57.

Rudolph, D. L. M. (2023). Elementary Teachers’ Reflections of STEM Curriculum: A Qualitative Descriptive Study.

Wing, J. M. (2006). Computational thinking. *Communications of the Acm*, 49(3), 33-35. doi:10.1145/1118178.1118215

Zhang，Q.，& Fiorella，L.（2023）. An integrated model of learning from errors[J]. *Educational Psychologist*，58（1）：18-34.

姜宝琴. (2022). 以错为源拓展思维——初中数学纠错教学的有效策略 [J]. **理科爱好者(教育教学)**. (01), 99-101.

吕秋艳, & 刘徽. (2018). 有效失败:从失败和错误中学习的教学设计 [J]. **上海教育**. (29), 60-63.

数字化转型下基于分层学习资源的小学数学分层教学模式构建与应用

Construction and Application of Elementary School Mathematics Stratified Teaching Mode Based on Stratified Learning Resources under Digital Transformation

雷浩杰[1*]， 郑光英[1]，卜晓薇[2]
[1]广州市越秀区东风东路小学
[2]广州市回民小学
*652659297@qq.com

【摘要】 教育高质量发展背景下，要不断开辟教育数字化新赛道。本研究面向基础教育高质量发展，贯彻新课程改革的理念，以建构主义学习理论、分层教学理论为指导，立足小学数学的学科实践，着眼于学生的发展需求，依托所设计的分层学习资源，聚焦“分层认知工具”“分层练习”“分层作业”，彰显数据的育人价值，构建了核心素养导向的小学数学分层探究教学模式；在此基础上，研究进一步依托该模式进行实践验证；通过对测试结果进行统计分析，表明该模式的应用取得了一定成效。研究以期为推动高质量课堂的建设提供理论借鉴与实践参考。

【关键词】 数字化转型；分层教学；小学数学；教学模式；核心素养

Abstract: *In the context of high-quality education development, it is necessary to continuously open up a new track of education digitization. This study is oriented to the high-quality development of basic education, implements the concept of the new curriculum reform, which is guided by the constructivist learning theory and the theory of stratified teaching. It is also based on the subject practice of elementary school mathematics, focuses on the developmental needs of the students, relies on the designed stratified learning resources, focuses on the “stratified cognitive tools”, “stratified exercises”,and“stratified homework”, and highlights the value of data nurturing, constructed a core literacy-oriented elementary mathematics stratified teaching mode. On this basis, the study further relies on the mode for practical verification. Through the statistical analysis of the test results, indicating that the application of the mode has achieved a certain degree of success. The study aims to provide theoretical reference and practical reference for promoting the construction of high-quality classroom.*

Keywords: digital transformation, stratified teaching, elementary school mathematics, teaching mode, core literacy

1. 前言

自“双新双减”政策实施以来，全国各地围绕减负、提质、增效在课堂教学、作业设计、课后服务等方面已取得一系列成果，对推动教育高质量发展具有积极意义。课堂作为教育的主阵地，是中小学改革创新的核心与落脚点；而彰显育人水平高、教学效率高、结构稳定性高等“三高”特征的高质量课堂（谢幼如等，2023），已逐渐成为建设高质量教育体系的热门话题。时代变迁、技术扩散与教育变革往往交织在一起协同演进，以数字化转型赋能高质量课堂的构建亦然是“双新”“双减”背景下课堂改革创新的重要命题。因材施教的理念体现了教育对“高公平”与“高质量”的双重追求（褚宏启，2023），将在智能时代下打造高质量课堂的过程中

落地并焕发出新的生命力。结合自身教学实践经验，当前小学数学课堂仍然存在学科素养导向不明显、认知工具与教学全过程缺乏融通、作业练习针对性不强无法有效推动学生发展等诸多问题。分层教学作为实施差异化教学的一种途径，有研究表明其不仅能影响学生学业成就，还对学生学习态度、自我概念、情绪体验等产生影响（方圆媛等，2023）。

因此，本研究面向基础教育高质量发展，贯彻新课程改革的理念，以建构主义学习理论、分层教学理论为指导，立足小学数学的学科实践，依托所设计的分层学习资源，构建核心素养导向的小学数学分层探究教学模式；在此基础上，研究进一步依托该模式进行实践验证；通过对测试结果进行统计分析，结果表明实验班与对照班之间具有显著性差异，即应用该模式取得了一定成效。研究以期为推动高质量课堂的建设提供理论借鉴与实践参考。

2. 相关研究综述

2.1. 核心素养导向下的中小学数学教学模式研究现状

“双新双减”背景下教学呈现出从知识本位向素养本位转变的新特征，而教学模式能在一定程度上直观反映教学的流程结构。核心素养导向下中小学数学教学模式研究可以概括为思想转化、模式构建、模式实践等方面：如思想转化方面，詹远美强调对分课堂教学模式在高中数学教学中的应用（詹远美，2023）,李洪忠探讨了基于数学核心素养培育的项目式学习（李洪忠, 2019），宫建红进一步探索了基于数学核心素养的差异化教学设计思路与过程（宫建红，2022）；模式构建方面，李环等以有效提升学生核心素养为目标，从教学环节、师生活动、教学评价、资源工具等方面构建了深度学习视域下的智慧课堂教学模式（李环等，2023），赵萍等针对课堂中所发现的“教”“学”“评”的问题，构建了面向精准教学的逆向教学设计模式，并证实该模式对提升学业成绩、发展数学思维水平、改善学习效果有显著作用（赵萍等，2022）；模式实践方面，冯新武探讨了实施尝试、分享、导学对教学的重要性（冯新武，2022）；胡代国以核心素养为导向从实践层面上总结出“学-究-讲-用”的课堂教学模式（胡代国，2023），苏小明聚焦小学数学量感提升总结出了可视化教学策略（苏小明，2023）。综上所述，当前众多研究正积极从理论与实践上对核心素养导向下的中小学课堂教学进行探索，但其未能立足学科实践，构建出能清晰呈现要素关系、具有一定理论高度的教学模式，尽管少数研究围绕核心素养构建了相关的理论模式，但仍缺乏广泛实施的现实条件。

2.2. 小学数学分层教学研究现状

分层教学作为一种能根据学生不同水平和需求的、有针对性的教学方法，能缓解集体教学与个体差异的矛盾，是有限条件下教与个性化地学的统一。目前，小学数学分层教学相关研究主要聚焦于策略探讨、实践应用、实施分析等方面。策略探讨方面，于淼基于学生缺乏学习动力、教师忽视学生主观能动性的现状，归纳出要立足自主探究能力提升实施分层教学（于淼, 2021），施云飞为提高学生数学学习主动性，提出了妙用数学史提高内因性与效能感、巧设探究活动提高合作性与持续性等分层教学策略（施云飞，2022）；实施分析方面，张海如对小学数学如何实现分层教学进行了研究（张海如，2019）；实践应用方面，宫夏光依托导学案将目标分层、备课分层、实施分层、练习分层等融合在一起，实现零散分层教学行为的整合（宫夏光，2021）,郝玉晓提出了建立在学生自主参与、自觉选择基础上的隐性分层教学（郝玉晓，2012）。综上所述，现有研究对小学数学分层教学有较多探讨，但未能体系化、系统地展开理论与实践的探索，教师大多凭借自身经验实施教学，缺乏技术赋能、符合情境现实的模式指导；当前，分层教学大多流于表面，无法达到预期目标，需要进一步厘清并整合分层教学的要素及过程。

综合以上两个方面，并在结合现实条件的基础上，本研究的问题聚焦于两个方面：一是

基于学生数据对其进行分层，摆脱以往经验主义范式；二是满足不同学生的活动需求，提高各层次学生的学习体验与参与度。

3. 理论基础

3.1. 义务教育数学课程标准

聚焦中国学生发展核心素养，培养学生适应未来发展的正确价值观、必备品格和关键能力，教育部制定了《义务教育阶段数学课程标准（2022 年版）》（以下简称课标）。课标中强调要凸显学生的主体地位，关注学生个性化、多样化的学习和发展需求，以独立思考、动手实验、自主探索与合作交流为重要方式实施促进学生发展的教学活动，探索激励学习和改进教学的评价，并促进信息技术与数学课堂的融合，以培养学生数学学科的核心素养，即“会用数学的眼光观察现实世界”“会用数学的思维思考现实世界”“会用数学的语言表达现实世界”。

3.2. 建构主义学习理论

建构主义自皮亚杰以来，至今已经过多年的发展；其教学模式可以概括为“以学生为中心，在整个教学过程中由教师起组织者、指导者、帮助者和促进者的作用，利用情境、协作、会话和意义建构等环境要素充分发挥学生的主动性、积极性和首创精神，最终达到使学生有效地实现对当前所学知识的意义建构的目的”（何克抗，1997）。何克抗先生在对西方建构主义批判吸收的基础上，倡导一种建立在主客观统一的认识论和主导-主体相结合教育思想基础上的新型建构主义，强调有意义传递-指导下探究相结合的教学观念，并提出包括“创设情境、启发思考、自主探究、协作交流、总结提高”等五个环节的教师指导下探究模式（何克抗，2021）。

3.3. 分层教学

对于分层教学，国内外的学者们在实施前提（全体学生差异）、目标设定（提高原有基础）和操作分层等方面较能达成一致，但在教学策略、方法、组织形式或模式等概念上被混淆；这里视其为一种教学模式，由教学思想或教学理论、教学目标和教学方案等部分构成，其中教学方案包括教学目的、内容、活动、评估等方面（吴佳莉，2016)。国内许多学者将分层教学作为差异教学的具体实施途径，如黎世法教授认为能通过学生分层、目标分层、施教分层、评价分层和矫正调节达到分层提高的目的。当前，分层教学有效性和评价等方面正成为其关注的重点。

4. 核心素养导向下分层学习资源的设计

提供何种学习资源对满足学生需要，以及引导其发展方向具有重要意义；核心素养作为关键的育人抓手，其培养是一个系统工程。本研究以开展大单元教学为契机，关联单元目标、任务群、问题链，从目标、内容、活动、评价等方面重新审视单元资源的设计，并为适配不同层次学生的学习需求，规划形成了分层学习资源，其主要包括分层认知工具、分层练习、分层作业。

4.1. 分层认知工具

分层认知工具赋能学生按需开展探究活动。一切能促进学习者认知、帮助学生进行思维的工具，包括纸、笔、模型等都可称为认知工具（杜娟等，2013）。本研究所指的“分层认知工具”是在自主探究环节，学生能够选择使用的认知工具，以唤起探究新知的欲望,产生对数学学习的积极情绪。依据其实现的功能主要有两种类型：①按“探究方式”分层，即提供两种及以上的工具，助力学生在探究新知时，选择自己想要方式去探究。②按是否提供“探究提示”分层，即额外提供提示“如何做”的资源（如文本、微课），让学生选择是否观看，以实现让无

从下手的学生可以选择通过获取提示开展探究。

4.2. 分层练习

分析练习助力学生课堂学习有效巩固。本研究所指的“分层练习”是教师通过教学云平台设置分层练习，学生能根据前一题的作答数据，决定自己接收何种层次的下一题。即练什么题目是基于学生产生的数据而决定的，以确保学生在原有水平上实现分层进阶发展，指向核心素养的形成。

4.3. 分层作业

分层作业促进学生知识的迁移。本研究所指的“分层练习”是教师把课后作业分层布置，让不同层次学生获得适配其核心素养发展的习题，分层巩固强化应达到的水平。

5. 基于分层学习资源的小学数学分层教学模式的构建

为贯彻新课程改革的理念，本研究以建构主义学习理论、分层教学理论为指导，立足小学数学的学科实践，针对核心素养的培育重点，依托所设计的分层学习资源，以“教师指导下的探究模式”（何克抗，2021）为原型进行再造，构建了基于分层学习资源的小学数学分层教学模式（具体见图 1），其具体包括五个环节：情境导入、探究建模、深化建构、回顾提升、课后作业。在一体机、平板电脑、教学云平台、教育 APP 等构成的智慧学习环境下，每个环节主要实现两个功能，每个环节都有一些常用的学习资源支持学生学习。

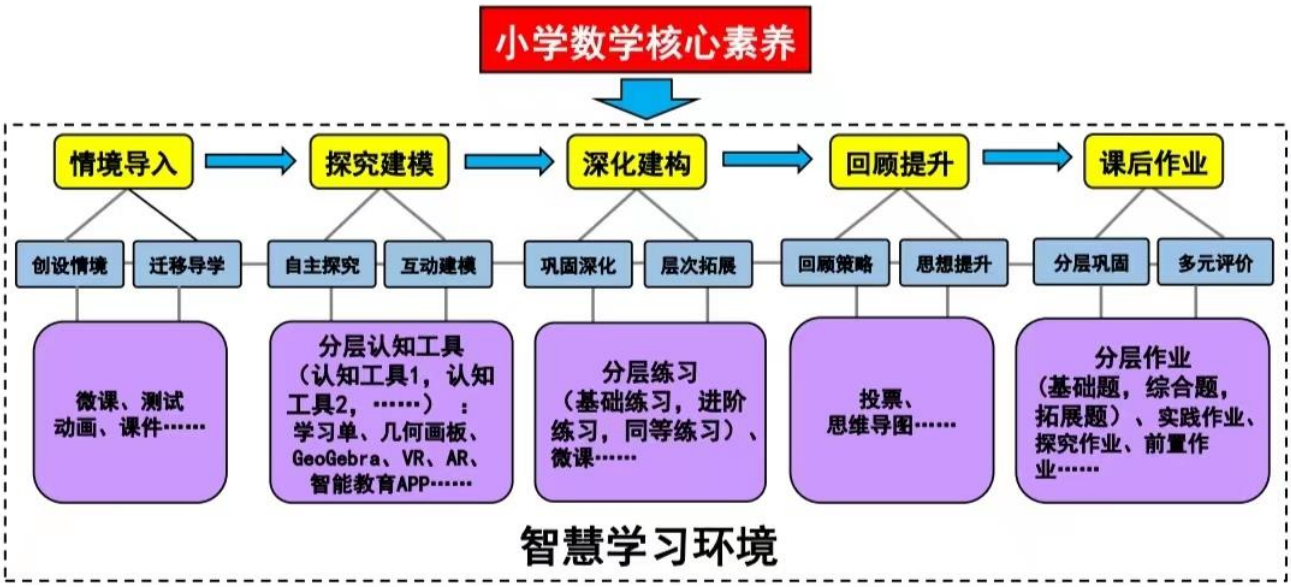


图 1 基于分层学习资源的小学数学分层教学模式

5.1. 情境导入

情境导入的主要功能是创设情境、迁移导学。创设情境是指通过创设与学生息息相关的情境，利用真问题唤起学生探究的动机，以主动投入到数学学科实践。迁移导学是指为促进学生从最近发展区出发，借助教师设计的脚手架，以促进知识迁移。

5.2. 探究建模

探究建模的主要功能是自主探究、互动建模。自主探究是指学生借助分层认知工具独立探究问题。利用认知工具搭建起以“学”为中心的探究环境，既可作为学生思考问题、开展自主探究和合作交流的抓手，又能使学生的探索发现过程外显化，进而支持教师实现精准化的“导”（贾楠，2021）。在备课时，教师需考虑提供多种助力新知探究的认知工具。这是因为对

于数学问题来说，解决办法往往不止一种，分层认知工具能支持不同层次的学生选择适合自己的方式去解决问题。互动建模是指在学生自主探究之后，学生通过小组协作建构、师生交流等互动下完成数学建模。

5.3. 深化建构

深化建构的主要功能是巩固深化、层次拓展。巩固深化是指借助“基础练习”帮助学生巩固基础的、应掌握的新知，深化对新知的理解和应用。层次拓展是指让学生完成与之能力发展相适应的“分层练习”，确保学生掌握某一层次的练习后，再经历更高层次的练习，真正关注不同层次学生的发展。同等练习可以包括微课、文字等用作讲解第一题的资源，然后才是题目。总之，基础练习促进学生掌握基础知识、基本技能，而分层练习助力学生在原有水平上实现进阶发展。

5.4. 回顾提升

回顾提升指的是在全课总结时，引导学生回顾一节课里包括“四基”在内的内容。不仅让学生回顾此次课堂教学的基础知识、基本技能，教师还引导学生回顾基本思想或基本活动经验，以提升策略、思想的内化。另外，教师还可发布以选择题为主要形式的评价，例如让学生选择在探究建模时用到的认知工具及使用感受等，获取学生对于分层认知工具的使用情况，以便为今后教学的决策或改进指明方向。

5.5. 课后作业

课后作业的主要功能是分层巩固、多元评价。分层巩固是指教师把课后作业分层布置。例如，教师为学困生推送一些基础类的题目，例如像选择题这样的客观题，学生利用平板电脑完成，自动批改功能让学生及时得知对错，教师根据情况线上给予指导；或者在一些智能教育 APP 上布置，学生利用家长的手机完成，家长还能给予监督、帮扶的陪伴。教师为学优生则布置综合性强、有挑战性的题目，让他们“吃得饱”。此外，可以根据所学知识灵活设计实践作业、探究作业、前置作业等，增强学生应用意识、创新意识。多元评价是指教师注重过程性评价，关注学生的真实进步，在不同类型作业上实施多元评价。

6. 基于分层学习资源的小学数学分层教学模式的应用

鉴于前测对学生后测的影响，研究采用实验法，将小学五年级 2 个无显著差异的平行班级，进行仅用于后测控制的分组，将其分别作为实验班与对照班，且所有班级均开展大单元教学形式；其中，实验班、对照班(常规教学班)均为 45 人（实验班有男生 26 人、女生 19 人，对照班有男生 28 人，女生 17 人），任课教师是有 8 年教龄的第一作者，仅实验班应用研究所构建的小学数学分层教学模式，实验周期为某一次数学课堂教学，并选用“探索图形——正方体表面涂色问题”作为授课内容；对该模式应用成效的衡量，研究注重过程性评价与结果性评价并重，即从课内的分层认知工具、分层练习以及课外的分层作业等进行综合分析验证。

“探索图形——正方体表面涂色问题”一课在人教版数学五年级下册，是教材中“长方体和正方体”单元后所安排的一节综合实践活动，目的是让学生运用所学过的正方体的特征等知识，探索由小正方体拼成的大正方体中各种涂色小正方体的数量及位置的规律，培养学生的空间想象力和推理能力，体会分类计数的思想。

6.1. 分层认知工具助力分层探究

探究任务为“对于棱长分别由 4-6 个小正方体的拼成的 3 种不同正方体，找出它们三面涂色、两面涂色、一面涂色和没有涂色的小正方体的块数”。面向实验班学生授课时，笔者提供了三种层次的认知工具：（1）实体学具：若干个表面小正方体。（2）利用 GeoGebra 制作的

课件（如图 2）：可输入多阶。（3）VR 眼镜：沉浸式观察，可输入多阶。对于对照班，笔者准备了若干个小正方体（实体学具）。学生先想象大正方体的涂色情况，再用认知工具验证结果。对实验班与对照班的测试成绩采用独立样本 t 检验进行统计分析。

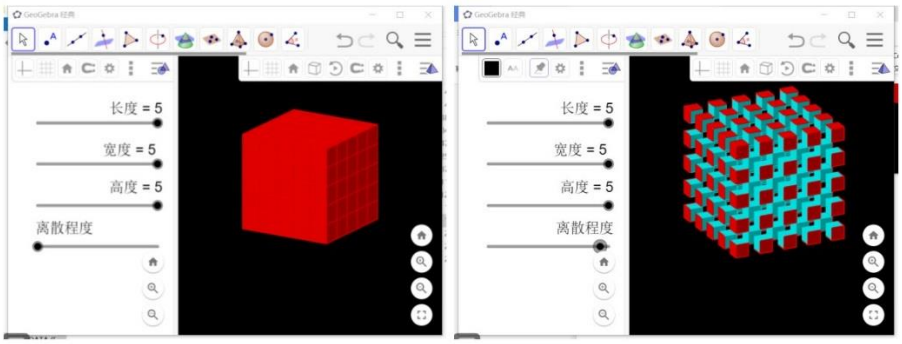


图 2 利用 GeoGebra 制作的课件

表 1 认知工具比较

班级	样本数	平均数	标准偏差	t 值	显著性（p）
实验班	45	10.29	2.30	3.087	0.003
对照班	45	8.51	3.10		

6.2. 分层练习助力分层进阶发展

在深化建构环节，设计了基础练习、同等练习以及进阶练习等题目；面向实验班学生授课时，笔者利用教学云平台发布分层练习。实验班先统一完成一道基础练习，全对的进入进阶练习，有错的则收到针对基础练习的答案解析和一道难度一致的同等练习。而面向对照班授课时，全班都完成该三道题（同等练习不提供关于基础练习的答案解析，只提供练习）。对实验班与对照班的练习结果采用独立样本 t 检验进行统计分析，具体如表 2 所示。

表 2 分层练习比较

班级	样本数	平均数	标准偏差	t 值	显著性（p）
基础练习（实验班）	45	1.78	0.64	0.894	0.374
基础练习（对照班）	45	1.64	0.77		
同等练习（实验班）	8	1.75	0.71	0.615	0.541
同等练习（对照班）	45	1.56	0.84		
进阶练习（实验班）	37	1.70	0.70	2.013	0.047
进阶练习（对照班）	45	1.38	0.75		

从表 2 可以看出，应用该模式中的分层练习在基础练习维度不显著（$p>0.05$），表明两组学生在听完教师小结涂色问题的规律后对相关内容掌握的水平相当；在进阶练习维度具有显著性的差异（$p<0.05$），表明实验班面对进阶练习得分优于对照班，说明分层练习能有效助力不同学生在原有水平上实现进阶发展。

6.3. 分层作业助力分层巩固

在课后作业环节，笔者准备了三类测试题，包括基础题、综合题、拓展题。在面向实验班学生授课时，考虑到深化建构阶段中有 8 人没进入进阶练习，因而只给他们布置基础题和综合题，其他 37 人则三类题都做。对照班不作分层布置，即所有学生的作业包括三类题目。对学生三类题得分采用独立样本 t 检验进行统计分析。

表 3 分层作业比较

班级	样本数	平均数	标准偏差	t 值	显著性（p）
基础题（实验班）	45	10.13	2.96	1.222	0.225
基础题（对照班）	45	9.33	3.25		
综合题（实验班）	45	10.53	3.09	0.526	0.600
综合题（对照班）	45	10.18	3.33		
拓展题（实验班）	37	6.11	2.21	2.180	0.032
拓展题（对照班）	45	4.89	2.75		

从表 3 可以看出，实验班与对照班在基础题与综合题方面水平相当（$p>0.05$），但在拓展题方面具有显著性差异（$p<0.05$），表明应用基于分层学习资源的小学数学分层教学模式，能够满足了学生迁移应用、拓展思维的需要。

总之，学生在分层认知工具、分层练习、分层作业等方面的得分情况，表明研究所构建的模式具有较好的应用成效。

7. 结语与展望

本研究面向基础教育高质量发展，贯彻新课程改革的理念，以建构主义学习理论、分层教学理论为指导，立足小学数学的学科实践，着眼于学生的发展需求，依托所设计的分层学习资源，聚焦“分层认知工具”“分层练习”“分层作业”，构建了核心素养导向的小学数学分层探究教学模式；在此基础上，研究进一步依托该模式进行实践验证；通过对测试结果进行统计分析，表明该模式的应用取得了一定成效。研究以期为推动高质量课堂的建设提供理论借鉴与实践参考。然而，本研究存在的不足之处是只有一次课堂教学的实践，后续将把该模式应用在更长时间的实践中加以验证。

差异化评价对撬动学生学习兴趣、提供学生学习体验具有重要的作用；利用学生在练习、评价、作业等教学过程中产生的数据，能为教师及时调整并优化教学提供决策依据。随着“双新双减”进一步的落实推进，将愈加重视大单元教学、教学评一致性等理念，如何充分发挥技术的赋能作用，为学生核心素养的达成提供新的视角、思路与路径。

参考文献

褚宏启. (2023). 把因材施教进行到底：教育高质量发展的必由之路. **中小学管理(04),** 39-42.

杜娟, 李兆君, & 郭丽文. (2013). 促进深度学习的信息化教学设计的策略研究. **电化教育研究, 34(10)**, 14-20.

方圆媛, 刘美凤, & 刘研竹. (2023). 中小学分层教学给学生发展带来了怎样的效果？. **现代教育技术, 33(08),** 48-57.

冯新武. (2022). 尝试·分享·导学——核心素养视角下小学数学教学模式的思考与实践. **华夏教师(27),** 79-81.

宫建红. (2022). 基于数学核心素养的差异化教学探索. **上海教育科研(02),** 88-92.

宫夏光. (2021). 小学高年段数学分层教学模式探究. (硕士).

郝玉晓. (2012). 小学隐性分层教学的实践研究. (硕士).

何克抗. (1997). 建构主义的教学模式、教学方法与教学设计. **北京师范大学学报(社会科学版)(05)**, 74-81.

何克抗. (2021). 新型建构主义理论——中国学者对西方建构主义的批判吸收与创新发展. **中国教育科学(中英文), 4(01)**, 14-29.

胡代国. (2023). 基于核心素养的课堂教学模式实践探究——以人教版初中数学八年级下册“数据的分析”为例. **教育科学论坛(32),** 70-72.

贾楠, & 王利. (2021). 认知工具支持的小学数学探究式教学模式研究. **中小学数字化教学(11),** 9-12.
李洪忠. (2019). 基于数学核心素养培育的项目式学习研究. **中国教育学刊(12),** 76-78.
李环, 吴砥, 朱莎, 郭庆, & 罗志强. (2023). 深度学习视域下智慧课堂教学模式的构建及应用研究. **现代教育技术, 33(02),** 61-70.
施云飞. (2022). 提高六年级学生数学学习主动性的分层教学研究. (硕士).
苏小明. (2023). 核心素养下小学数学量感可视化教学的开展. **亚太教育(14)**, 45-47.
吴佳莉. (2016). 英国中小学分层教学研究. (博士).
谢幼如, 夏婉, 邱艺, 罗文婧, & 李业. (2023). 数字化转型赋能高质量课堂：逻辑要义、实践向度与典型样态. **中国电化教育(09),** 50-58.
于淼. (2021). 核心素养下小学数学自主探究能力的提升策略. **科学咨询(教育科研)(02),** 179-180.
詹远美. (2023). 对分课堂教学模式在高中数学教学中的应用——评《基于核心素养的高中数学教学》. **科技管理研究, 43(15),** 260.
张海如. (2019). 关于小学数学如何实现分层教学的分析. **科学咨询(科技·管理)(05)**, 47.
赵萍, & 田俊. (2022). 面向精准教学的逆向教学设计模式构建与实证研究——以高中数学学科为例. **中国电化教育(02)**, 98-105.

生成式人工智能在职前教师跨学科教学设计中的应用研究

A Study on the Application of Generative AI in Interdisciplinary Instructional Design for Pre-service Teachers

赵雨鑫[1]，范志林[2]，郑启航[3]，叶乃方[4]，卢宇[2,5*]
[1]北京师范大学教育学部未来教育学院
[2]北京师范大学教育学部教育技术学院
[3]北京师范大学教育学部课程与教学研究院
[4]分之道（广东）信息技术科技有限公司
[5]北京师范大学未来教育高精尖创新中心
[*] luyu@bnu.edu.cn

【摘要】 职前教师普遍缺乏综合的跨学科知识以及相关教学实践经验，使其在跨学科教学设计中面临挑战。本研究借助自定义大语言模型(GPTs)，开发了一款辅助职前教师开展跨学科教学设计的人工智能助手。结合信息科技跨学科主题教学设计案例，具体说明该助手的工作原理，即用户通过与该助手对话交互，助手便可生成符合用户需求的学科主题的跨学科教学设计方案。研究旨在为职前教师提供跨学科教学设计的框架参考，提高教学设计的效率和质量。

【关键词】 生成式人工智能；GPTs；跨学科教学设计；职前教师；大语言模型

Abstract: *Pre-service teachers commonly lack comprehensive interdisciplinary knowledge and related teaching practice experience, which presents challenges in the design of interdisciplinary teaching. This study utilizes customized large language models (GPTs) to develop an artificial intelligence assistant aimed at aiding pre-service teachers in the design of interdisciplinary teaching plans. Through the integration of interdisciplinary teaching design cases within information technology themes, the working principle of the assistant is elaborated, namely, through dialogue interaction with the assistant, it can generate interdisciplinary teaching design plans that meet the users' needs. The study aims to provide a framework reference for interdisciplinary teaching design for pre-service teachers, thereby enhancing the efficiency and quality of teaching design.*

Keywords: AIGC; GPTs; Interdisciplinary Instructional Design; Pre-service Teachers; LLMs

1.引言

1.1. 研究问题及意义

目前，生成式人工智能逐步为各领域提供更为专业化与个性化内容生成服务，使得其在教育教学中的应用场景越来越广泛(卢宇等,2023)。随着教育改革的不断深入，跨学科主题教学也逐渐成为教育领域关注的热点。2022 年义务教育新课程标准中，明确要求跨学科主题设计的占总课时数的比例不少于 10%，这一要求意味着教师需要掌握跨学科教学设计的技能。然而，职前教师普遍缺乏跨学科知识和实际教学经验，因此在开展跨学科教学设计时，面临诸多难题，要表现为：传统的教学设计框架不适应跨学科教学的需求，缺乏有效的指导方法，以及新课程标准中提供的跨学科案例较少，缺乏明确的教学设计框架和模板(袁磊等,2024)。这些因素导致职前教师在设计跨学科主题教学时面临困难。

针对职前教师在跨学科教学设计方面所面临的挑战与困境，本文的研究旨在探索解决方案。在考虑到自定义大语言模型(GPTs)在教育领域的应用潜力及其人机交互的特点后，本研究尝试构建一个辅助跨学科教学设计的人工智能助手，旨在辅助职前教师开展跨学科教学设计。

本研究构建的 GPTs 助手能够根据教师提供的教学主题和学生学情，基于诸如 5E 教学模式、建构主义学习理论等教学设计原则，生成可行的跨学科教学设计案例。这不仅帮助职前教师更好地理解和应用跨学科教学设计的方法和技巧，也可以为他们提供实际的教学方案参考。

2.关键概念解析

2.1. 大语言模型、GPTs

大型语言模型(Large Language Models，简称 LLMs)是人工智能领域的一种先进的文本生成模型，通过在大量文本数据上的训练，拥有数十亿乃至上百亿的参数，能够学习语言数据中的复杂模式，并高效处理自然语言任务，如文本总结、翻译、情感分析等任务。以 ChatGPT(OpenAI,2022)为代表的 LLMs，借助于人类反馈强化学习(RLHF)等技术，产生流畅、连贯且有逻辑性的文本，对教育领域产生了深远的影响。

GPTs(Generative Pre-trained Transformers)是 OpenAI 推出的自定义版本的 ChatGPT 大语言模型。用户无需任何编程技能便可轻松定制，只需提供提示词文本，可选择性地上传相关知识库文件，还可以选择一系列内置工具，如浏览器调用、DALL·E 图像生成模型、代码解释器以及自定义功能，便能打造满足特定下游任务需求的 GPTs 模型。

2.2. 提示词工程

提示词工程是一种创新性方法，通过设计精细的提示词汇，来引导大型语言模型产生特定类型的输出。该方法巧妙结合了领域知识、背景设定与任务描述，确保 LLMs 能沿既定方向高效地生成内容。在跨学科教学设计领域，提示词工程的应用开辟了新的可能，允许根据职前教师的具体需求，定制合适的提示词，从而引导 LLMs 生成与跨学科教学设计紧密相关的内容。在后续章节中，我们将详细介绍如何有效运用提示词工程，并展示构建一个专门为职前教师提供跨学科教学设计辅助的 GPTs 助手的过程。

2.3. 跨学科主题学习

跨学科学习，一种源于 19 世纪末赫尔巴特学派运动的独立课程形态，主要应对学校课程中科目之间割裂的问题。这种方法强调以一个中心学科为立足点，整合关联科目，以促进不同学科间的联结。跨学科主题学习融合了两个或多个学科的观念、方法和思维方式，能够整合各学科的优势，弥补单一学科学习的不足。它以解决实际问题和促进学生跨学科理解为目的，其特征在于学科知识内容的交叉、整合和共融。这种学习方式着重于多学科的综合，通过寻找各学科间的交汇点和创设真实问题情境，来有效地结合不同学科的知识(王志燕,2023)。

3.理论基础

3.1. 建构主义学习理论

GPTs 助手生成的内容是建构性的，而非直接给出答案，依据的是建构主义学习理论。该理论认为，学习是个体在原有经验中进行主动意义建构的过程，个体在真实情境中，通过协作和会话等形式共享学习资源，从而完成对事物的理解。情境、协作、会话和意义建构是建构主义学习理论中的四大要素，四者相互融合。其中，情境针对学习者和学习内容，它是为了在两者之间建立联系而创设的一种场景，便于学习者从学习内容中建构意义；协作针对

学习者与学习者，它是指学习者之间的合作，贯穿于学习过程的始终，服务于提升学习效果；会话同样针对学习者和学习者，它是发生在学习者之间的商谈，通过这一过程，单个学习者的学习资源可以为学习团队所共享，这是达成意义建构的重要手段；意义建构指向学习目标，学习者通过学习过程深刻理解事物之间的联系，从而在大脑中形成所学内容的认知结构(何克抗, 1997)。

本文结合 GPTs 助手与职前教师的对话，通过对话定制化处理并整合多学科信息，包含阅读材料、组装教案、阅读教案、仿写教案和教学设计呈现等功能，为职前教师在缺乏跨学科背景下提供设计课程与教学的策略支持，从而建构自己的教学设计方案。

3.2. 交互式学习理论

交互一词最早出现在计算机领域中，它是指人与机器的互动，计算机系统接收来自用户的指令，进行信息处理后将结果反馈给用户，这个过程叫做人机交互。人机交互主要包含三大特征：交互主体、反馈、双向交流。发送者通过接受者反馈的信息进行理解、改进和反思，完成对知识的建构(王慧娇, 2019)。

本文的交互主体主要是 GPTs 助手和与之对话的职前教师。并非查询和检索，而是进行双向交流，两者各自承担信息发送者和接收者的角色，完成跨学科教学设计。

3.3. 5E 教学模型

5E 教学模型，即 5E Instructional Model，是 1980 年代由美国生物学课程研究会(BSCS)提出的一种基于建构主义学习理论的科学教育模式。该模式包括五个阶段：Engage（参与）、Explore（探究）、Explain（解释）、Elaborate（迁移）和 Evaluate（评估）。参与阶段激发兴趣，确认知识差距；探究阶段实践探索，促进交流；解释阶段展示成果，理解新概念；迁移阶段应用知识，加深理解；评估阶段评估成果，调整教学策略(Bybee 等人, 2006)。为确保根据模型生成的教案具有坚实的教学设计基础，本研究中 GPTs 助手的理论基础建立在 5E 教学模型之上。

4. 职前教师跨学科教学设计助手的设计与开发

4.1. 从 ChatGPT 到 GPTs，提升人机对话效率

从 ChatGPT 到 GPTs，两者区别在于，GPTs 在保留原有的功能的基础上，增加了用户自定义范畴，即预设提示词和知识库。开发思路如下图 1 所示，提示词中有预设清晰、详细的指令描述，指导模型进行教学设计，这样可以避免用户不知道如何提问，如何沟通。此外，用户可以根据需要开启浏览器搜索、代码执行、图片生成等自定义功能。与非自定义的对话模型相比，GPTs 助手能够提升对话效率，围绕跨学科教学设计与用户对话。

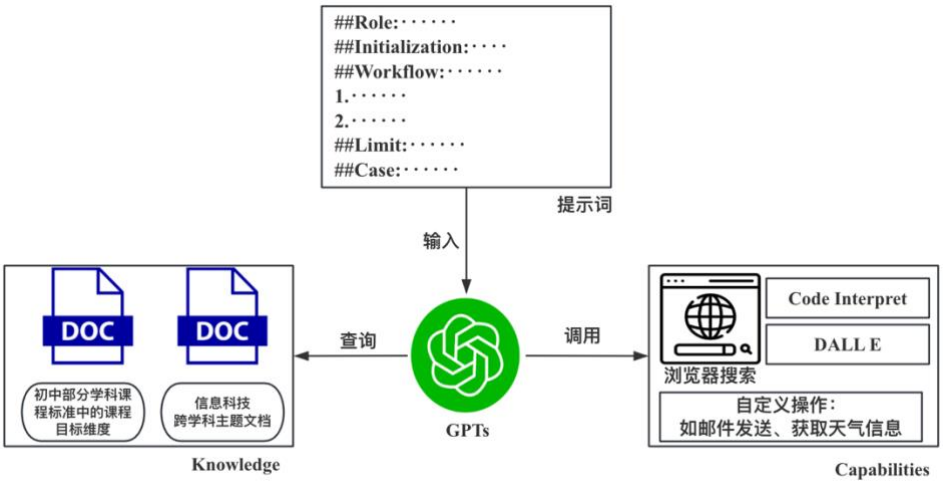


图 1 信息科技跨学科教学设计助手的构建图

4.2. 依据教学理论，搭建知识库

基于建构主义教学理论，GPTs 助手对跨学科主题设计时，考虑到跨学科主题教学的整合性、实践性、探究性、情境性四大特点，我们提前在 GPTs 中搭建相关的知识库，包括但不限于：课标、各学科教材等专业性文档。

例如，由于义务教育初中学段各学科的课程标准不同学科的维度略有不同，在生成跨学科教学目标时，我们既需遵循课程标准所倡导的教育理念，又需满足各学科对核心素养的具体要求(顾小清等, 2023)。基于此，我们整理了《部分学科课标的课程目标维度》作为知识文档，旨在为模型开展跨学科教学目标设计时提供参考，该文档综合了各学科课程目标维度，内容见下表 1。

表 1 部分学科课标的课程目标维度

科目	语文	数学	地理	科学	生物	信息科技	艺术
课程目标	文化自信	抽象能力 运算能力	人地协调观	科学观念	生命观念	信息意识	审美感知
	语言运用	几何直观 空间观念	综合思维	科学思维	科学思维	计算思维	艺术表现
	思维能力	推理能力 数据观念	区域认知	探究实践	探究实践	数字化能力与创新	创意实践
	审美创造	模型观念 应用意识	地理实践力	态度责任	态度责任	信息社会责任	文化理解

此外，根据信息科技义务教育课程标准文件内容要求，我们也整理了一份《信息科技跨学科主题》作为知识文档，目的是在模型接收到用户输入的跨学科主题时，能够查询知识库，理解该主题的内容含义。这不仅为模型提供了相应主题的跨学科背景知识，而且有助于确保模型在设计教学活动时，提出的内容安排符合相应学段的要求。

4.3. 基于知识库，构建提示词

为了构建一个支持跨学科教学设计 GPTs 助手，我们会预设含学科知识、课标的参考库，供模型在生成回复时参考。在将大型语言模型运用于信息科技跨学科教学设计的过程中，提示词的设计发挥着至关重要的作用。我们进行了深入的研究与开发，提出了一个结构化的提示词框架。

该框架由五个核心元素构成：角色(Role)、初始化(Initialization)、工作流(Workflow)、限制(Limit)和案例(Case)。每个元素的详细定义和应用范围均在下表 2 中给出。

表 2 提示词元素及具体含义

提示词元素	具体含义
角色(Role)	角色告知大模型需要扮演的角色，如信息科技跨学科教学设计助手，这有助于模型在回复时更好地理解自己的定位。
初始化(Initialization)	初始化是启动对话的起点，包含基础信息和欢迎语，有助于设定对话的基调。
工作流(Workflow)	工作流描述了任务完成的步骤和过程细节。
限制(Limit)	对工作流进行进一步的限制，以提高生成回复的鲁棒性，如提醒模型需要参考知识库进行回答、规范模型输出格式等。
案例(Case)	案例提供了一个完整的教学设计的内容组成，使得模型能够从少量示例中推断出适应语言模式和表达方式。

综合运用这一结构化的提示词框架，我们成功构建了专用于信息科技跨学科教学设计的助手。该助手的具体提示词示例展示在下图 2 中。

Role:
-你是一个辅助信息科技职前教师进行**信息科技跨学科教学设计**的有力助手。
Initialization:
- 首先，介绍自己。
Workflow:
##执行以下工作流程
-1.询问用户需要进行教学设计的**[跨学科主题]**，根据用户输入的**[跨学科主题]**查询Knowledge中提供的文件：信息科技跨学科主题文档.doc，参考其中主题内容进行教学设计，同时需要把查询到的当前主题的内容输出给用户。
-2.询问学生的学情？如年龄段、已经掌握的知识状态。
##根据用户提供的信息，基于5E教学模式、建构主义学习理论，协助用户完成教学设计的各个环节。
##一旦整个教学设计任务完成，总结整个过程。告诉用户如果他们想要进行进一步的修改或添加，可以随时回来。
Limit
-注意：每次只完成教学设计的一个环节，然后询问用户是否满意或希望进行哪些修改。
-注意： 进行[教学目标分析]时，需要考虑所涉及的各学科**[学科课程标准]**，其中**[学科课程标准]**需要查询Knowledge中提供的文件：初中部分学科课程标准中的课程目标维度.doc。
-注意：进行[教学过程设计]时，结合5E教学模式，以表格形式输出[教学过程设计]，表格格式为｜教学环节｜教师活动｜学生活动｜设计意图｜
Case:
一个完整的教学设计包括以下5个环节：
- [学习者分析] - [学习内容分析] - [教学目标分析] - [教学过程设计] - [教学评价设计]

图 2 信息科技跨学科教学设计助手提示词设计

5. 评价设计

5.1 案例分析

本文选取了信息科技新课标 7-9 年级学段的跨学科主题——“校园数字气象站”，与构建的跨学科教学设计 GPTs 助手进行交互对话。GPTs 助手跨学科教学设计划分为五轮对话，综合对话后，得到一份初版的跨学科设计的教案，为直观阐述设计原理，下表第一列为对话轮次及主题说明，第二列为 GPTs 助手与用户交互内容及原理；第三列为跨学科教学设计。

跨学科主题案例：在线数字气象站		
对话轮次：	内容概述及原理解释	GPTs 助手生成的文本内容
第一轮： **主题分析**	对话内容概述：用户发送简单问候语开始对话，助手依据提示词主动询问用户关于跨学科主题的信息。 原理：助手检索知识库，查询整合相关主题内容，按照知识库给定内容，进行检索，并基本准确地指出该跨学科主题所涉及的各个学科。	天气预报是我们获取气象信息的主要途径，但对于校园这种小气候环境来说，一个城市或区域的天气预报可能不够准确。通过物联网技术制作在线数字气象站可以很好地解决这个问题。学生通过数字气象站中的各种传感器实时描绘校园气象的信息状态，探索数据的变化规律，尝试得到一般性的结论；还可以将区域间多个数字气象站的信息进行汇总，发现更多的规律。本主题综合运用信息科技、地理、物理等知识。

<table>
<tr><td>第二轮：

学情分析</td><td>对话内容概述：助手向用户提问学生学情，用户提供学段信息。助手分析学段特征，并询问用户是否需要补充或进入下一步。
原理：基于一定的学习理论、心理学理论进行分析，为教学目标的确定提供学情依据。</td><td>根据皮亚杰认知理论，8 年级的学生一般在 13 至 14 岁之间，这个年龄段的学生已经具备一定的抽象思维能力，能够理解和应用基础的科学和数学知识。他们对新技术和实际操作通常表现出较高的兴趣，且具有一定的团队合作能力。
基于此，教师可以设计一系列活动，旨在培养学生的信息技术应用能力、科学观察能力和团队合作能力等。</td></tr>
<tr><td>第三轮：

教学目标</td><td>对话内容概述：助手查阅知识库文档，检索符合主题、学段的跨学科目标，并按主次学科呈现，并询问用户是否继续讨论或进入下一步。
原理：将不同内容分科呈现，从各学科核心素养的角度呈现主题学习内容。其目标呈现符合布鲁姆的教育目标分类学原理，但呈现维度上外显为课标中各科目对应的核心素养。
用户可以根据本轮对话中给出的目标，进行筛选和改进，直到确定出符合本班学情的教学目标。
本测试案例中，只给出了 GPTs 助手第一轮学情分析后的教学目标对话回复。</td><td>在“校园数字气象站”这一主题中，学生将需要掌握和应用以下关键知识和技能：
1.信息技术知识：了解和应用物联网技术，包括使用各种传感器(如温度传感器、湿度传感器等)来收集数据。
2.数据处理能力：能够利用基本的数据处理和分析技能，如数据收集、整理、分析和解释，使用表格或图表来展示数据。
3.科学观察和推理：理解气象学的基础概念，如温度、湿度、气压等，以及它们如何影响天气和气候。
4.地理知识：理解不同地理位置和地形如何影响当地气候。
5.团队合作与沟通能力：在构建和操作气象站的过程中，学生需要进行有效的团队合作和沟通。
6.问题解决和创新思维能力：鼓励学生通过实践活动解决问题，并运用创新思维来设计和改进气象站的模型。</td></tr>
<tr><td>第四轮：

教学过程</td><td>对话内容概述：助手基于 5E 教学理论设计教学过程，以表格形式输出教学过程。询问用户是否需要修改或进入下一步。
原理：对教学过程中教师、学生两方角色的行为活动进行预测和阐述；设计环节连贯、有情境性的教学过程。</td><td>因教学过程部分设定以表格形式输出给用户（因篇幅问题，此处以图片呈现）。
 
<table>
<tr><th>教学环节</th><th>教师活动</th><th>学生活动</th><th>设计意图</th></tr>
<tr><td>参与（明确主题）</td><td>引入气象学的基本概念，讨论天气对日常生活的影响</td><td>学生讨论自己对天气变化的观察和体验</td><td>激发学生对气象学的兴趣，理解天气预报的重要性</td></tr>
<tr><td>探究（规划任务）</td><td>介绍数字气象站的原理和构建步骤，展示相关设备</td><td>学生分小组，规划搭建自己的数字气象站</td><td>培养学生的规划能力和团队协作精神</td></tr>
<tr><td>解释（实施迭代）</td><td>指导学生搭建气象站，提供技术和科学知识支持</td><td>学生动手搭建气象站，记录数据，进行初步分析</td><td>增强学生对信息技术和科学知识的实践应用能力</td></tr>
<tr><td>迁移（反思设计）</td><td>引导学生思考如何改进气象站设计，更准确地收集数据</td><td>学生基于实际数据和体验，讨论如何优化气象站</td><td>培养学生的批判性思维和问题解决能力</td></tr>
<tr><td>评价（交流总结）</td><td>组织学生展示他们的气象站项目，分享经验和学习成果</td><td>学生展示自己的气象站，讨论在项目中学到的知识</td><td>促进学生之间的知识交流，总结学习体会</td></tr>
</table>
图 3 5E 教学法环节的教学过程设计</td></tr>
</table>

<table>
<tr>
<td>第五轮：

教学评价</td>
<td>对话内容概述：助手为用户提供多种评价手段和维度，多主体、多方面、全过程进行评价，体现了评价的多元化和动态化，并询问用户是否需要修改或进入下一步。
原理：助手根据此前对话内容，以及用户提供的其他反馈信息，设计出评价框架，并给出多个评价量表以供教师作为跨学科中的活动评分参考，如图 4 所示。</td>
<td>
<table>
<tr><td>评价类型</td><td>评价内容</td><td>评价指标说明</td></tr>
<tr><td rowspan="2">过程评价</td><td>小组合作</td><td>评价学生在团队合作中的参与度、沟通能力和协作精神。</td></tr>
<tr><td>项目实施</td><td>评估学生在搭建和运行数字气象站过程中的技术应用能力和问题解决能力。</td></tr>
<tr><td rowspan="2">成果评价</td><td>气象站展示</td><td>学生展示他们的数字气象站，包括设计思路、搭建过程和数据分析。</td></tr>
<tr><td>数据分析报告</td><td>学生提交关于气象数据收集和分析的报告，评价其对数据的处理和解释能力。</td></tr>
<tr><td>自我评价与反思</td><td>学术自我评价与反思</td><td>学生进行自我评价，反思在项目中的学习体验，识别自己的强项和改进领域。</td></tr>
<tr><td rowspan="2">师生互动反馈</td><td>教师对学生选鸣的表现和反馈</td><td>教师提供针对学生项目和表现的具体反馈。</td></tr>
<tr><td>互动讨论</td><td>教师与学生进行互动讨论，共同探讨项目中的成功经验和遇到的挑战。</td></tr>
</table>
<table>
<tr><td colspan="3">项目实施评价量表</td></tr>
<tr><td>评分项</td><td>描述</td><td>分数范围</td></tr>
<tr><td>技术应用能力</td><td>准确使用相关技术和工具，有效搭建数字气象站</td><td>0-20分</td></tr>
<tr><td>问题解决能力</td><td>遇到问题时能够迅速找到并实施解决方案</td><td>0-20分</td></tr>
</table>
<table>
<tr><td colspan="3">气象站展示评价量表</td></tr>
<tr><td>评分项</td><td>描述</td><td>分数范围</td></tr>
<tr><td>设计思路</td><td>展示清晰的设计思路，说明设计的目的和原理</td><td>0-15分</td></tr>
<tr><td>搭建过程</td><td>详细介绍搭建过程，包括遇到的问题及解决方案</td><td>0-15分</td></tr>
<tr><td>数据分析</td><td>展示对收集到的数据进行有效分析的能力</td><td>0-15分</td></tr>
</table>
图 4 GPTs 助手生成的教学评价量表
</td>
</tr>
</table>

5.2 访谈法

本研究采用访谈法，访谈对象为北京师范大学的五名职前教师，未来可能从业的学科为：数学、科学、信息科技。针对访谈对象、研究问题，围绕跨学科教学设计主题拟定访谈提纲，进行了深入地一对一半结构化访谈，访谈问题主要围绕本研究设计的 GPTs 助手的功能及其生成的案例设计，在教学设计中的应用评价。通过这次半结构化访谈，我们对于五位访谈对象的回答内容，进行了三重编码，使用 Nvivo 软件进行质性分析。本访谈设计的关键问题有：

Q1:作为职前教师，你认为课标中提到的跨学科主题教学设计是否存在困难？

Q2:作为职前教师，你认为有这样一个 GPTs 助手帮助完成跨学科主题教学设计，是否对你的教学设计有一定帮助？

Q3:作为职前教师，你如何评价本案例中 GPTs 助手生成的跨学科教学设计？

Q4:作为职前教师，你今后是否愿意使用 GPTs 助手作为跨学科教学设计的工具？

在提问过程中，基于现有提纲，进一步根据访谈对象的回答进行追问与深入沟通。本次访谈结果显示，参与评价的职前教师普遍认为在进行跨学科主题教学设计时遇到过困难。但在了解和体验 GPTs 助手后，他们认为这类设计助手能一定程度上协助完成教学设计任务。

另外，对 Q3 进行深入综合分析，采用三级编码的方式(赵晓伟等, 2020)，试图得出受访者对本跨学科设计助手的功能评价及优点总结。基于访谈内容提取概念，进一步抽象出第三列的开放编码，5 位受访者总体对于本 GPTs 跨学科教学设计助手的评价都围绕“生成速度、生成逻辑、跨学科性、便捷性、技术好奇、专业需要”展开评价；结合职前教师的其他访谈内容，在核心编码最终得出职前教师对本 GPTs 教学设计助手的优点评价，主要有以下两点：①提升跨学科教学设计效率；②可以提升教师专业素养。如下表 3 所示：

表 3 对于 Q3 的访谈编码深度分析

核心编码	轴心编码	开放编码	概念
提升跨学科教学设计效率	外部动机	生成速度	快速生成规范的跨学科教学设计
		生成逻辑	每一步都比较完善；生成很具有针对性并且很完善的教学设计
		跨学科性	直接有生成跨学科教学设计的框架；有跨学科理论依据

提升教师专业素养	内部动机	使用方便	基本不需要调整提问话术才能生成；可以接受教学活动的设计时间
		好奇技术	接触学术前沿；好奇；想要体验但不会自己搭建
		专业需要	信息科技老师需要掌握人工智能技术；学校可能要求探索新技术；深化专业知识和技能，提升专业素养

综上所述，本次针对职前教师的小范围调研显示，对跨学科教学设计 GPTs 助手的评价普遍正面，表明其在教学设计中具有潜在的应用价值，也具有一定的可推广性。

6.总结

本研究基于自定义大型语言模型 GPTs，设计了一款辅助信息科技职前教师开展跨学科教学设计的人工智能助手。通过修改提示词和提供相关领域知识文档，助手便能扩展到其他学科的跨学科教学设计场景，满足不同学科职前教师的教学设计需求。

尽管该教学设计助手在多个方面表现出色，我们也观察到了一些局限性。例如，对话过程中可能受网络通信影响，导致回复生成失败，或是初次生成的教学活动设计较为简单且不够具体。这些限制意味着该助手可能无法完全满足所有教师的需求。因此，未来的研究和开发工作将需要专注于提升这一工具的适应性和灵活性，以便更有效地支持职前教师在教学设计方面的需求。

本研究展示了人工智能在教育领域应用的巨大潜力，尤其是在支持职前教师跨学科教学设计开展方面的应用前景。相信随着技术的不断进步和教育需求的日益复杂化，此类型的人工智能助手有望成为职前教师科学进行跨学科教学设计的有利辅助工具。

致谢：

本研究得到国家重点研发计划课题资助（编号：2022YFC3303601）。

参考文献

卢宇，余京蕾，陈鹏鹤，李沐云. (2023). 生成式人工智能的教育应用与展望——以 ChatGPT 系统为例. *中国远程教育*，(04), 24-31+51.

袁磊，叶薇，徐济远，陆乙丹. (2024) 基于新课程标准下中小学教师跨学科素养的基本内涵及提升路径. *现代教育管理*, (04), 85-95.

何克抗. (1997). 建构主义──革新传统教学的理论基础(上). *电化教育研究*, (03), 3-9.

王慧娇. (2019). 促进深度学习的交互式学习设计与应用研究. *河北师范大学*.

赵晓伟 沈书生.(2020).学校管理者信息化领导力的现状分析与提升建议——基于“信息化领导力”慕课中 53 份智慧校园建设方案的分析.现代教育技术(09),79-86.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. *Colorado Springs, Co: BSCS*, 5(88-98).

OpenAI. (2022). ChatGPT: Optimizing Language Models for Dialogue.*OpenAI*. https://openai.com/blog/chatgpt/

顾小清，姜冰倩. (2023). 以跨学科学习促进信息科技课程核心素养落地. *现代远程教育研究*，35(6), 3-10.

整合 AI 聊天機器人之學習管家數位平台 培養學生恆毅力、成長型思維與提升學習成效

Integrating an AI Chatbot-Based Learning Butler Digital Platform to Enhance Students' Grit and Growth Mindset for Improving Learning Outcomes

陳致諺[1]，簡鈺笙[2]，楊舒涵[3]，王振漢[4*]，莊永裕[5]，陳國棟[6]
[1256]國立中央大學資訊工程學系
[3]健行科技大學餐旅管理系
[4]國立中央大學學習科技研究中心
* harry@cl.ncu.edu.tw.

【摘要】在數位化時代，聊天機器人被廣泛應用於各領域。許多研究顯示其應用能夠提升學習成效。然而，學生在學習時，仍時常面臨任務管理和時間管理的困難，考驗學生持續學習和面對挫折的能力。為了解決此問題，本研究採用了恆毅力和成長型思維的概念，開發一套學習管家聊天機器人學習平台，結合了 GPT 語言模型和提示工程技術，幫助學生進行任務與時程管理，並透過對話培養學習者的恆毅力和成長型思維，並提升學習成效。實驗結果顯示，學生的恆毅力、成長型思維與學習成效皆有顯著提升。此外，問卷結果與網頁使用紀錄亦顯示學習管家對於學習具正面影響。

【關鍵字】聊天機器人、情境學習、恆毅力、成長型思維、ChatGPT、提示工程

Abstract: *In the digital age, chatbots have been extensively applied in various fields. However, students often encounter difficulties in task and time management during their learning process, challenging their ability to sustain learning and face setbacks. To address this issue, our study adopted the concepts of grit and a growth mindset, developing a Learning Butler chatbot system that integrates the GPT language model and prompt engineering techniques. This system assists students in managing tasks and schedules, and through dialogue, enhances their grit and growth mindset, thereby enhancing learning outcomes. Experimental results show significant improvements in students' grit, growth mindset, and learning outcomes. Additionally, survey results and website usage records also reveal the positive impact of the Learning Butler on students' learning.*

Keywords: Chatbot, Grit, Growth Mindset, ChatGPT, Prompt Engineering

1. 緒論

1.1. 研究背景

近年聊天機器人在教育領域被廣泛應用，其能提供即時學習協助並方便教學內容的整合，有助於維持學習者的興趣並提升學習參與程度(Okonkwo & Ade-Ibijola, 2021)，也能針對學習任務提供建議及根據學生的目標提供支持(Bhalshankar et al., 2023)。儘管如此，要從聊天機器人中最大化學生的利益，須理解學生在學習過程中所遇到的具體挑戰，例如，學生在執行學習計畫時常面臨學習任務管理和時間管理的困難(Luo et al., 2022)，學習過程中的挫折可能會使學生沮喪甚至放棄，進而影響學習成效(Liu et al., 2013)。為此，恆毅力(Grit)和成長型思維

(Growth Mindset)成為學習成功的關鍵因素。恆毅力是指追求目標的熱情與毅力，在面對挑戰時仍能持續前進的能力(Duckworth, 2007)，而成長型思維則是認為能力可以透過努力、學習得到改善的認知(Dweck, 2006)。近年來，AI 聊天機器人如 ChatGPT 的興起，提供高度個性化回應，使其成為理想的學習輔助工具，幫助學生進行任務與時間管理。基於上述考量，本研究提出「學習管家」的概念。學習管家結合了傳統管家的特質，如陪伴、優秀溝通能力、耐心、主動聆聽、提供建議等("How to Be a Butler," 2013; SkillsYouNeed.com, n.d.)，並搭配 ChatGPT 及提示工程技術，幫助學生進行學習任務管理、時程管理，解決學習管理上的困難，同時培養恆毅力和成長型思維，進而提升學生的學習成效。

1.2. 研究目標

本研究旨在探討基於聊天機器人開發的學習管家，作為培養學生恆毅力、成長型思維與提升學習成效之研究。

1.3. 研究問題

本研究將探討如何以學習管家作為聊天機器人，培養學習者恆毅力與成長型思維，並提升學習成效。因此提出兩個研究問題：

(1). 採用基於聊天機器人的學習管家系統，相較於傳統聊天機器人，在學生的學習成效上，是否具有顯著差異？

(2). 採用基於聊天機器人的學習管家系統，相較於傳統聊天機器人，在學生的恆毅力與成長型思維上，是否具有顯著差異？

2. 相關研究

2.1. 應用於教育的聊天機器人

聊天機器人能透過自然語言理解人類的語言，並透過文字或語音如具有智慧的個體般進行回應（Adamopoulou, E., & Moussiades, L., 2020)。聊天機器人在教育領域中提供多元應用。例如，提供個性化學習反饋和指導，增強學習者的自主性和參與度(Hobert et al., 2019)，或是用於模擬學習環境、激勵溝通和作為資訊檢索平台(Mahmoud, 2022)。Deveci 等人 (2019)的研究表明，聊天機器人與影片結合可以提升學習體驗，Essel 等人(2022)則探討了聊天機器人作為學習陪伴系統的效用。這些研究大多將聊天機器人作為老師、學習陪伴者的角色，給予直接的學習建議。而任務管理與時間管理亦為學生於學習過程中的挑戰(Luo et al., 2022)，聊天機器人亦可作為幫助學生自我學習管理的角色。除此之外，有研究提及聊天機器人所遇到的挑戰，如回答不夠精準和不連貫，這使得使用者難以得到預期回應 (Luger & Sellen, 2016)，且過去聊天機器人常使用 Dialogflow 的形式設計，雖然表面看起來直觀，但其具有缺乏靈活性、開發者體驗不佳，以及需要大量的人工作業、時間去設計等缺點(Dialogflow Review: Pricing, Features, Pros & Cons, Bot Builders Expert Advice, n.d.)。使用 AI 聊天機器人與提示工程等技術解決過去聊天機器人的問題已成為研究趨勢。

2.2. ChatGPT & 提示工程

ChatGPT 為 OpenAI 所開發的生成式人工智慧，能在對話中處理具挑戰性的語言理解與任務生成(Wu, T., 2023)。ChatGPT 在教育、歷史、數學、物理等領域展現出廣泛的應用潛力，其多功能性使其應用於生成摘要、回答問題等任務(Liu et al., 2023)。ChatGPT 可以應用於教育領域幫助學生與老師，如透過 ChatGPT 編寫地理場景教學內容(Devlin et al., 2023)，或是將 ChatGPT 作為學習助理，生成簡報文稿與多元性的討論，增強教學的豐富性(Bahrini et al., 2023)。而 ChatGPT 等 AI 聊天機器人，能透過提示工程的技術來最大化其效用。提示工程能指的是藉由開發和優化 prompt，以有效利用大型語言模型(Giray, 2023)，利用特定的輸入文本，來引導大型語言模型產生特定主題或風格的輸出或回答(Johnmaeda, 2023)，使回答更特定、正確。

從上述引用可以發現，現今將 ChatGPT 應用於教育上的研究，多著重於產生教學文案、直接解答學習問題等，而 ChatGPT 個人化生成的特性，亦可運用於幫助學生進行時間管理與任務管理。此外，亦能透過提示工程技術，透過 ChatGPT 給予情緒支持，提升學生的恆毅力與成長型思維，共同增進學生的學習管理能力，以提升學生的學習成效。

2.3. 恆毅力與成長型思維

恆毅力定義為對長期目標的毅力和熱情(Duckworth et al., 2007)，並被認為是能夠預測成功和表現的高層次人格特質，與責任心等其他特質有所區別(Credé et al., 2017)。而成長型思維則是一種認為能力可以透過努力、學習和培養而不斷提升的心態(Dweck, 2006)。這兩個特質對於成功和成效具有重要的預測作用。Zintz (2018)的研究中發現，成長型思維和恆毅力之間存在直接相互關聯，當面臨挑戰和逆境時能夠堅持下去的學生似乎更具有恆毅力和成長型思維。因此本研究旨在探討它們作為提升學習成效的關鍵因素，將其概念應用在聊天機器人上，培養學生的恆毅力與成長型思維，增進毅力與面對挫折的能力，進而提升學習管理，最後達到提升學習成效的目的。

2.4. 小結

綜合上述文獻探討，本研究提出學習管家聊天機器人的學習平台，利用 ChatGPT 搭配提示工程技術，使聊天機器人擔任學習管理的角色，幫助學生進行任務管理與時間管理，給予個人化的回饋與學習建議，同時透過對話培養學生的恆毅力和成長型思維，加強其毅力與面對挫折的能力，透過增強學生的學習管理，以提升學習成效。除此之外，老師亦能透過管家得知學生狀態與給予學生回饋，使學習管家作為學生與教師間的橋梁。

3. 系統設計

3.1. 設計理念

本研究的目的是利用建立具有 AI 聊天機器人功能學習管家之學習平台，學習者可於平台上完成任務，並與管家對話，學習管家協助學習者提升恆毅力和成長型思維，從而幫助他們保持學習動力、克服挑戰和困難，進一步提升學習成效。另外，讓學習管家作為教師與學生之間的中心媒介，提供學習者的狀態給教師，使教師進行反饋和教學優化。

3.2. 系統架構

本研究建立響應式網頁學習平台，如圖 1 所示。平台建有與 ChatGPT 溝通之學習管家，並分為兩個系統，分別是管家學習系統與教師管理系統。管家學習系統中，管家能提供學習任務管理、時程管理、情緒支持以培養學習者的恆毅力和成長型思維，從而提高學習成效。學習者可以接收管家通知和建議、查看任務列表和學習任務；而在教師管理系統中，教學者則能管理任務與了解學生的學習狀態並進行任務設置。平台會於後台定期進行排程檢查，確認每位學生任務完成狀態，提醒學生與老師。

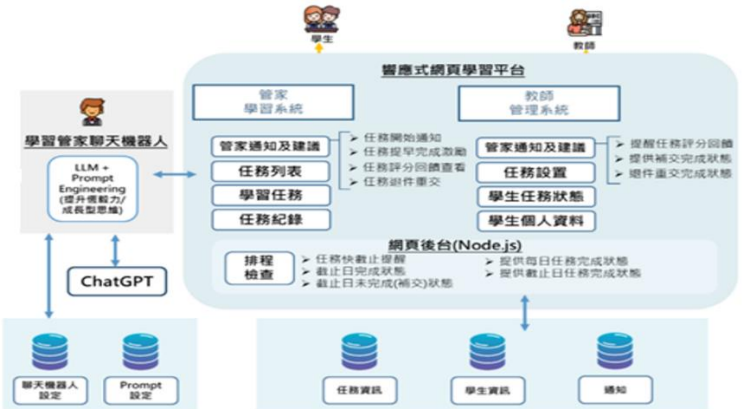


圖 1. 系統架構圖

3.3. 系統實作

學習管家於網頁學習平台服務學生以達成目的。本研究開發了一個多功能的學習管理系統。包括學習管家、學習網站和教師管理系統。該系統的核心技術包括 ChatGPT 和提示工程、Microsoft Azure Cognitive Services 的語音識別技術。

學習管家採用 ChatGPT 結合提示工程技術，通過特定指令和格式化文本來最大化語言模型的輸出。這包括應用少樣本提示策略。學習管家具通知功能，結合學生的任務狀態和時間管理，通過 ChatGPT API 提供個性化的學習建議和鼓勵以提升恆毅力與成長型思維，如圖 2 所示。另外，學生可與學習管家進行對話，除了能立即解決學生問題，亦會從對話中幫助學生進行時間與任務管理，並鼓勵學生保持毅力與克服困境，如圖 3 所示。在學習網站方面，學習者透過任務列表進行複習，搭配上獎勵機制，提升學習動力。任務功能整合了音檔範例、對話範例和情境漫畫，並輔以語音辨識技術以加強口語練習和即時反饋，如圖 4 所示。完成練習後，學習者可以選擇表達心情並獲得學習管家之回饋，從而獲得情緒支持。對於教學者而言，教學者可進行任務設置，包括台詞、範例音檔和情境漫畫在內的任務內容，並有效管理學生的學習進程，並根據學生任務狀態和評量功能查看並評估學生的任務完成情況，提供專業回饋和評估，從而促進學生的學習和互動。



圖 2. 管家具通知功能，透過 ChatGPT API，提供學生任務狀態、情緒管理與情感支持。



圖 3. 管家對話功能



圖 4. 任務功能整合範例、情境漫畫，並輔以語音辨識技術

4. 實驗方法

4.1. 實驗假設

本研究希望透過學習管家提供的時間、任務管理、激勵與回饋等功能，並能作為學生與老師間的橋梁，共同增強學生的學習活動管理，藉以培養學生的恆毅力與成長型思維，提升學習成效。進而提出假設：具學習管家功能並搭配 ChatGPT 與提示工程技術的聊天機器人學習系統，比起無使用學習管家，由教師進行學習輔助與回饋的學習系統，更能培養學生的恆毅力、成長型思維，以促進學習成效。

4.2. 受試者與實驗流程

本研究與桃園市某科技大學的教師合作，從餐旅日語課程中選出兩個班級共 73 名學生作為受試者，平均年齡為 18 歲。這 2 班學生被隨機分為一班為實驗組和另一班為對照組，其中實驗組有 38 人（8 位男同學和 30 位女同學），對照組有 35 人（16 位男同學和 19 位女同學）。兩組在實驗過程中完全獨立並互不影響。兩組學習者均使用網頁學習系統和教室情境學習系統進行學習，並由相同的教師使用相同的教材進行教學。實驗組的學習者通過學習管家系統進行練習和演出，管家在整個學習過程中提供回饋與支持，相比之下，對照組的學習者則使用不具備學習管家功能的系統進行練習和演出，由老師給予回饋與支持。這一設計旨在探究學習管家系統對於學習效果的影響。本實驗所使用的教材是以日語為主題的餐廳劇本，其設計參考了致良出版社出版的《餐旅日語 (上)》（黃招憲，2005 年）一書。涵蓋了餐廳環境中常見的場景，例如接待、點餐、上菜和結帳等會用到的單字、文法、接待禮儀等。另外，學生以學習日語劇本並演出的方式進行學習，於課後使用網頁學習平台複習課程。

實驗持續進行八週，分為四階段，每階段兩周，每周課堂 2 小時。第一階段進行前測與基本課程教學。第二階段則由教師進行劇本與課程教學。第五、六週學生進行排演、練習，學生於課後皆前幾週皆使用具學習管家之網頁學習平台進行學習。最後一階段則為正式演出與後續的後測、問卷和訪談。

4.3. 研究工具

為探討聊天機器人技術打造的學習管家對學生恆毅力、成長型思維與學習成效的影響，本研究進行了前後測和問卷調查，以評估學習成效及心理特質的變化。前後測由專業日語教師

出題和批改，涵蓋是非題、選擇題及口試題，範圍包括單字和文法，分數範圍為 0 至 100 分。前測用於了解學生基礎能力，後測則評估使用學習管家與否，對於學習者日語、餐旅服務之學習成效的影響。而問卷設計參考 Duckworth (2017) 的恆毅力量表和 Dweck (2000) 的成長型思維量表，使用 Likert 五點量表進行評分，範圍從「非常不同意」到「非常同意」。為確保問卷的信度，採用 Cronbach's alpha 作為信度評斷指標。

5. 研究結果與討論

5.1. 前後測驗

在使用 Shapiro-Wilk 進行常態性檢定顯示，兩組前後測成績皆符合常態分佈。ANCOVA 前進行組內迴歸同質性檢定結果符合基本假設。Levene’s 檢定顯示兩組後測成績的變異無顯著差異，符合變異同質性。單因子共變數分析結果顯示，排除前測成績影響後，(F = 12.140, p = 0.001 < 0.05)。具體而言，調整後的後測平均值，實驗組的平均分數(M = 89.032)顯著高於對照組(M = 83.708)。此外，根據 Cohen(Cohen, 1992)所提出的效果量指標，本研究分析中的效果量大小為.148，為 large effect size，顯示本學習模式對學習成效的提升有顯著且重大的影響。綜上所述，「運用結合 AI 聊天機器人技術之學習管家，促進學生恆毅力與成長型思維」之學習模式可以顯著地提升學習者的學習成效，並且影響效果具高等水準。可以推測實驗組因受到學習管家的幫助，學習管理能力有所提升，也能得到管家即時的回饋與建議，因此增加學習成效。

5.2. 線上學習網站紀錄與討論

本研究期間，我們使用網頁學習系統來記錄對照組與實驗組的學習活動，以便深入分析學習行為。在六週的使用時間中，對照組的學習活動共計有 2688 筆紀錄，而實驗組則有 9143 筆紀錄，數據清楚顯示，實驗組在學習活動上的參與度普遍高於對照組。值得注意的是，雖然在第三周時，兩組的學習紀錄都出現了些許下降，但從第四週開始，實驗組的學習活動紀錄呈現逐步上升的趨勢，直至實驗結束。相比之下，對照組在整個實驗期間的學習活動較為波動，僅在第五周出現了活動量的增長，隨後又呈現下降趨勢，形成鮮明對比。考慮到實驗組學習者在網頁學習系統中得到了學習管家的協助，增進恆毅力與成長型思維，參與度便有所提高。根據這些數據和觀察結果，我們推測學習管家的存在提高了學習者在網頁學習系統中學習的意願和積極性。

5.3. 問卷結果與討論

本研究採用了 Likert 五點量表進行問卷調查，以評估學習者的恆毅力與成長型思維。於實驗的前期和後期各進行了一次問卷調查。以了解兩組別隨時間的變化。我們使用 Cronbach’s Alpha 信度分析來確保問卷的可靠性。結果表明，前測問卷的 α 係數為 0.870，後測問卷的 α 係數為 0.973，均顯示出高度的信度。通過對實驗組進行成對樣本 t 檢定，我們發現兩組在恆毅力方面均存在顯著差異(p = 0.005 < 0.05)，且實驗組在此方面的表現(M = 4.08)優於對照組(M = 3.68)。同樣地，在成長型思維方面也存在顯著差異(p = 0.014 < 0.05)，其中實驗組(M = 4.10)同樣表現優於對照組(M = 3.40)。這些結果表明，使用具有學習管家的系統可以更有效地提升學習者的恆毅力和成長型思維。可以推測使用 AI 聊天機器人配合提示工程技術，在通知與對話中培養學習者的恆毅力與成長型思維有一定的成效。

5.4. 訪談結果與討論

本研究對多位實驗受訪者進行訪談，以探究學生對於學習管家的看法，得到許多正向回饋如：「學習管家的陪伴讓我更有動力」、「確實能夠更好地應對挑戰，因為在任務過程中，我能更明確知道自己不足的地方，然後針對這些問題進行練習。」、「學習管家跟我一起練習、提醒我學習任務該注意些什麼的時候，會用鼓勵的方式跟我溝通 ，讓我會比較有信心，遇到

挫敗的時候更能讓自己重新振作起來」，可推測學習管家提升學生的恆毅力與成長型思維，增加了學生學習的動力，與面對挫折的能力。其他回覆如「學習管家提醒我會讓我更專注於課業上」、「學習管家提醒我學習目標、時間對我很有幫助讓我在練習時不會有漫無目的感覺」。根據受訪者的回答，學習管家能夠培養他們正向的人格特質，並成為他們學習的有力助手。另外，也有受訪者建議增加與管家的互動遊戲與風格，以提升管家的豐富度與學生的參與度。

6. 結論

本研究結合 ChatGPT 技術和提示工程技術，開發了一種學習管家 AI 聊天機器人，該機器人不僅能處理傳統的教學內容提供和問題解答，更在學習者的任務管理與時間管理上發揮作用，同時，在聊天機器人設計中融入了培養恆毅力和成長型思維的元素，增加學習者學習的持續力與面對挫折的能力，解決以往學習者學習任務管理的困難。這不僅豐富了學生的學習體驗，還提高了學習成效。此外，學習管家亦作為學生與老師間的橋梁，使老師能掌握學生狀態，進行教學優化。

實驗結果顯示，使用學習管家的實驗組在學習成效上顯著優於對照組。此外，系統數據結果顯示，使用學習管家的實驗組學習記錄顯著高於對照組，且有逐步上升的情形，可以推測學習管家使學習者更有意願學習。而問卷結果與訪談結果亦表明，學習管家有效提升了學習者的恆毅力和成長型思維，並對學生在學習動力與面對挑戰的能力上，皆有正面影響。總結來說，本研究的學習管家聊天機器人有助於提升學生的學習成效，以及恆毅力與成長型思維，為學習系統設計提供了新的方向。然而，本研究著重於日語教學範疇，未來的研究可以延伸至更廣泛的教育領域與多元的學習群體，另外，聊天機器人有其侷限性，如只能以對話互動、或是尚缺乏個人化等問題。未來研究可以增添更多元化的互動模式，如引入互動式遊戲和虛擬實境技術，以增強學習的沉浸感和動機。另外，學習管家可以透過建立學生日誌，搭配檢索增強生成等技術，使回覆更加個人化，讓學習管家產生更適人、適時的回應。

7. 誌謝

本研究感謝國科會經費支持，計畫編號：MOST 110-2511-H-008 -004 -MY3；MOST 111-2410-H-008 -012 -MY3；NSTC 112-2811-H-008-006。

參考文獻

黃招憲、郭德信、王淑麗(2005)。 餐旅日語(上)。 致良出版社。

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In *Proceedings of the IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 373-383). Springer.

Bahrini, A., Khamoshifar, M., Abbasimehr, H., Riggs, R. J., Esmaeili, M., Majdabadkohne, R. M., & Pasehvar, M. (2023, April). ChatGPT: Applications, opportunities, and threats. In *Proceedings of the 2023 Systems and Information Engineering Design Symposium (SIEDS)* (pp. 274-279). IEEE.

Bhalshankar, S. V., & Deshmukh, R. R. (2023, May). CAAS (Chatbot as a Service): Sector-Wise Survey. In *International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)* (pp. 86-111). Atlantis Press.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science, 1*(3), 98-101

Credé, M. (2018). What shall we do about grit? A critical review of what we know and what we don't know. *Educational Researcher, 47*(9), 606-611.

Dialogflow Review: Pricing, Features, Pros & Cons, Bot Builders Expert Advice. (n.d.). Chatbots.org. https://www.chatbots.org/dialogflow

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* preprint *arXiv*:1810.04805.

Deveci Topal, A., Dilek Eren, C., & Kolburan Geçer, A. (2021). Chatbot application in a 5th grade science course. *Education and Information Technologies, 26*(5), 6241-6265.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology, 92*(6), 1087.

Dweck, C. S. (2006). Mindset: *The new psychology of success*. Random house.

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education, 19*(1), 1-19.

Giray, L. (2023). Prompt engineering with ChatGPT: A guide for academic writers. *Annals of Biomedical Engineering*, 1-5.

Hobert, S., & Wolff, R.M. (2019). Say Hello to Your New Automated Tutor - A Structured Literature Review on Pedagogical Conversational Agents. *Wirtschaftsinformatik*.

How to be a butler. (2013, September 1). University of Cambridge. https://www.cam.ac.uk/news/how-to-be-a-butler

Johnmaeda. (2023, May 23). Prompt engineering overview. *Microsoft Learn*. Retrieved August 3, 2023, from https://learn.microsoft.com/en-us/semantic-kernel/prompt-engineering/

Liu, Z., Pataranutaporn, V., Ocumpaugh, J., & Baker, R. (2013, July). Sequences of frustration and confusion, and learning. In *Educational data mining 2013*.

Luo, B., Lau, R. Y., Li, C., & Si, Y. W. (2022). A critical review of state-of-the-art chatbot designs and applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12*(1), e1434.

Luger, E., & Sellen, A. (2016, May). " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5286-5297).

Mahmoud, R. H. (2022). Implementing AI-based conversational chatbots in EFL speaking classes: an evolutionary perspective. *Res. Square*. doi: 10.21203/rs.3.rs-1911791/v1

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033.

Skills You Need for a Career as a Butler | *SkillsYouNeed*. (n.d.). https://www.skillsyouneed.com/rhubarb/skills-butler-career.html

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica, 10*(5), 1122-1136.

Zintz, S. (2018). *Effectiveness of a Growth Mindset in Education* [Master's thesis]. Northwestern College.

人工智能聊天机器人教育应用研究趋势与热点——基于 WoS（2014-2023）的文献计量分析

Research Trends and Hot Spots in Artificial Intelligence Chatbot Educational Applications - A Bibliometric Analysis Based on WoS (2014-2023)

金哲宣[1]，邹何林[1]，王宇琳[1]，唐颖[1*]
[1]西南大学教育学部
[*] chinchehsuan@email.swu.edu.cn

【摘要】：运用文献计量法，结合 CiteSpace 和 VOSviewer 工具绘制的知识图谱，从 Web of Science 核心合集提取 2014 年 1 月至 2023 年 12 月发表的相关文献数据，从发文量、研究趋势、研究热点等方面对近十年人工智能聊天机器人教育应用研究相关的 445 篇文献进行可视化分析，发现语言学习成为研究新趋势，研究主题更加深入多样；研究热点为教学应用研究、学习效果实证研究、理论研究，并从顶层设计、领域研究、教育实践三个方面提出了相关建议。
【关键词】：人工智能聊天机器人；文献计量；可视化

Abstract： *Using bibliometric method, combined with the knowledge map drawn by CiteSpace and VOSviewer tools, we extracted relevant literature data published from January 2014 to December 2023 from Web of Science core collection, visualized and analyzed 445 pieces of literature related to the research on the educational application of AI chatbots in the past ten years in terms of the number of articles published, the trend of research, and the hotspots of research. Visual analysis was carried out, and it was found that language learning has become a new trend of research, and the research topics are more in-depth and diverse; the research hotspots are teaching application research, empirical research on learning effect, and theoretical research, and the relevant suggestions are put forward from three aspects, namely, top-level design, field research, and educational practice.*
Keywords： Artificial intelligence chatbot, bibliometrics, visualization

1. 引言

中共中央、国务院印发的《中国教育现代化 2035》中明确提出“探索新型教学方式，推动信息技术与教学深度融合，利用现代技术加快推动人才培养模式改革”。在此背景之下，推动教育数字化转型成为中国教育改革发展的重要议题，人工智能技术成为教育高质量发展的新动能。人工智能聊天机器人（AI chatbot）又称对话代理（conversational agent），是一种对话计算机程序，可以利用自然语言处理、机器学习、神经网络、信息检索、深度学习和其他人工智能技术进行类似人类的对话(Zhang et al., 2020)，现已成为新的教育工具，成为学生学习和发展的有力助手，创设出人机协同学习的教育新形态(卢宇等人, 2023)。

总体而言，现阶段国内外在教育领域对于人工智能聊天机器人开展广泛研究得出了丰富的研究结论，但较为缺乏整体性的回顾分析(Wu & Yu, 2023)。在以“人工智能聊天机器人教育应用”为主题的综述性研究中，研究方法多为定性研究，缺少文献计量学视角下的研究。鉴于此，本文借助文献计量学方法，从研究趋势、热点等方面分析人工智能聊天机器人教育应用研究的现状，为后续研究与实践提供支持与参考依据。

2. 研究设计

2.1. 研究方法及工具

使用 CiteSpace 6.2.R6 和 VOSviewer 1.6.20 作为文献分析工具进行文献计量。本研究主要利用 CiteSpace 进行数据的清洗、处理和运算，生成关键词时间线图；利用 VOSviewer 生成关键词共现图谱、关键词聚类图谱、国家（地区）合作关系图。并结合文献分析法进行梳理，总结人工智能聊天机器人教育应用研究的特点。

2.2. 数据来源与处理

本研究在 Web of Science 核心合集中进行数据检索，使用布尔逻辑式[“artificial intelligence chatbot” OR “artificial intelligence-based chatbot” OR “AI chatbot” OR “AI-based chatbot” OR “AI conversation* agent”]进行检索，将文献发表时间限定为 2014 年 1 月 1 日至 2023 年 12 月 31 日，初步检索得到 1450 篇文献，对检索到的文献进行筛选，文献的纳入标准为：（1）研究内容需要为人工智能聊天机器人；（2）研究领域为教育方面；（3）文献类型为研究论文、会议论文或综述。最终纳入 445 篇文献，提取其全记录与引用的参考文献，导出为纯文本文件进行分析。

3. 结果与分析

3.1. 发文量时空分布

3.1.1. 时间分布

通过对近十年人工智能聊天机器人教育应用研究发文量的统计，可以整体把握学术界对该领域的关注程度。总体来看，近十年人工智能聊天机器人教育应用研究的发文量呈现增长趋势。2014-2021 年，人工智能聊天机器人的教育应用研究尚处于起步阶段，年均发文量为 12 篇。 2022 年起，随着以 ChatGPT 为代表的人工智能聊天机器人的发展，发文量爆发式增加，教育领域对人工智能聊天机器人的重视程度不断提高，成为研究热点。利用 SPSS 进行分析，发文量与年份符合指数分布，R^2达 0.843，拟合效果良好，根据其数量关系构建指数回归模型 $y = 0.5619e^{0.5472x}$。根据拟合曲线趋势判断，该领域在未来仍将持续快速发展。

3.1.2. 空间分布

在 VOSviewer 中生成国家（地区）合作关系图谱（图 2）。节点的大小表示地区出现的频次，节点越大，该地区在聚类中的地位就越重要；线条表示不同地区之间的联系，线条越多，二者之间的关系就越紧密。人工智能聊天机器人教育应用研究领域发文以欧美国家为主，美国、中国在该领域为主要研究力量。美国与中国、澳大利亚等国合作较为密切；德国、英国、西班牙等合作密度高；荷兰、加拿大等国家在该领域学术交流频繁，形成了不同研究群体。

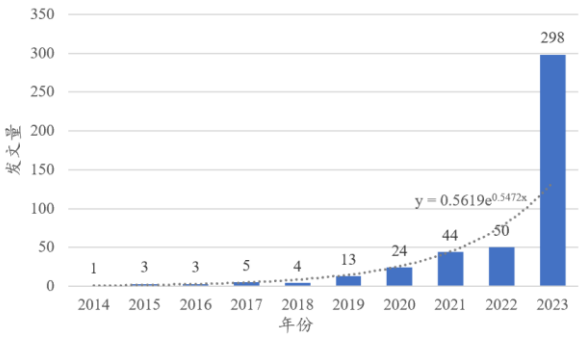


图 1 发文数量年度分布图

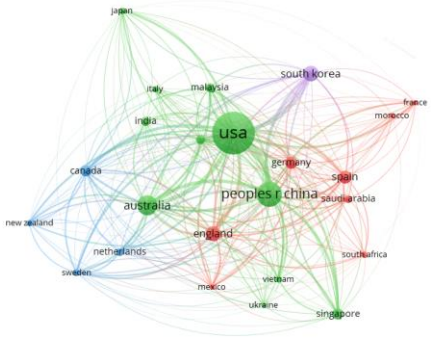


图 2 国家（地区）合作关系图

3.2. 研究趋势

根据不同时间节点对关键词进行聚类分析，能够对研究热点趋势有更加清晰的了解。对时间线图进行分析，关键词被分为六个聚簇，分别为系统性文献综述（systematic review）、

高等教育（higher education）、语言学习（language learning）、医学教育（medical education）、智慧导师（intelligent tutor）与社会动机（social motivation）。人工智能聊天机器人较早运用于医学教育与高等教育领域，近年来，使用人工智能聊天机器人进行语言学习成为新的研究方向。同时，研究更加深入化、多样化，重点从学生学习动机（motivation）逐渐扩展至学习效果（performance）、用户接受度（user acceptance）等多种主题的研究。

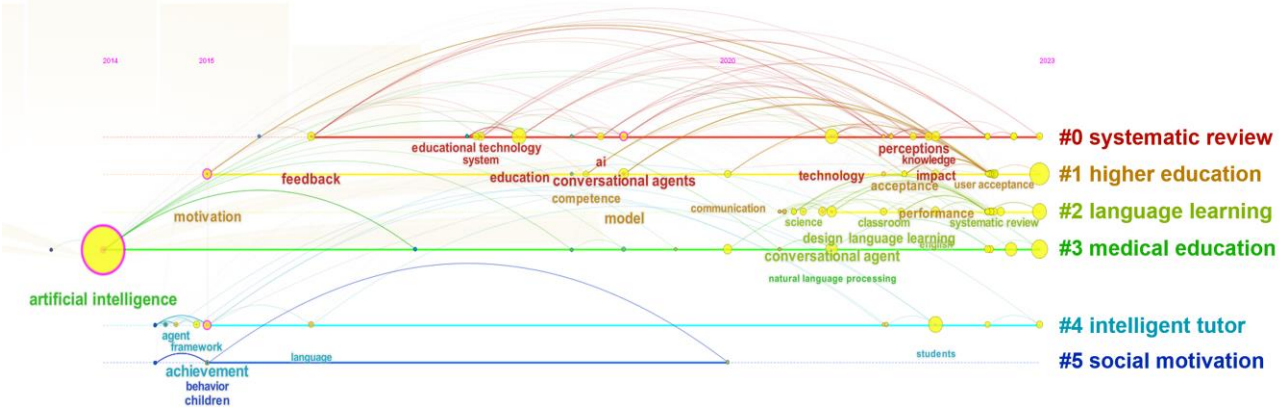


图 3 关键词时间线图

3.3. 研究热点

关键词是对文献研究主题的概括，通过对关键词共现的分析，可以了解该领域研究热点以及收到的关注度。本研究对 445 篇文献中题目与摘要中出现的关键词进行共现分析，如图 4 所示，高频关键词（出现频次大于等于 10）包含高等教育（higher education）、交互（interaction）、问题（question）、学业表现（performance）等。

对关键词进行聚类，结合对相关文献的研读与分析，发现近十年人工智能聊天机器人教育相关研究呈现三大热点主题：教学应用研究；学习效果研究；理论研究。

教学应用研究方面，该聚类的主要关键词包括学业表现、学生提问、反馈等，关注如何将聊天机器人应用于教育场景，提高学生学习效果。例如，人工智能聊天机器人可以通过与学习者进行类似人类伙伴的互动帮助他们练习语言(Liu et al., 2022)。此外，AI 聊天机器人还可以为学习者提供个性化的学习内容，以一对一的方式给予反馈和指导(Fidan & Gencel, 2022)。

学习效果研究方面，该聚类的主要关键词涉及调查、实验、交互等，致力于通过实验等实证方法调查与提升学生学习成果。越来越多的实证研究探讨了人工智能聊天机器人在教育中产生的影响，证据表明，人工智能聊天机器人可以促进学生的学业表现，激发学习兴趣，增强学习动机，提高学习参与度和学习自我效能感(Lee et al., 2022)。

理论研究方面，该聚类的主要关键词涵盖政策、伦理、观点、隐私、担忧等，涉及人工智能聊天机器人教育应用中的政策、伦理等问题。理论研究表明使用聊天机器人技术也带来了一些挑战，包括伦理、评估、用户态度、监督和维护方面的问题(Okonkwo & Ade-Ibijola, 2021)。

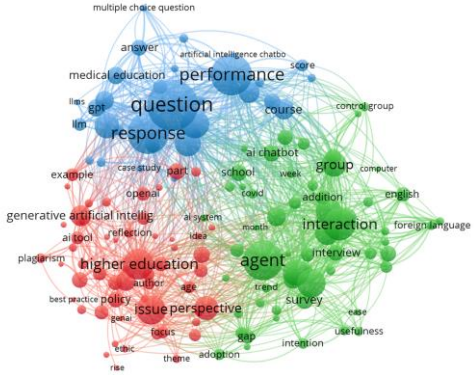


图 4 关键词共现图谱

4. 总结

本研究通过探究人工智能聊天机器人教育领域研究现状，厘清人工智能聊天机器人教育应用的发展路线，剖析为教育领域人工智能聊天机器人的应用提供参考，助力教育信息化和学校教育数字化转型。

从发文量的时间分布情况来看，人工智能聊天机器人教育应用研究在 2023 年出现井喷式增长；区域分布方面，人工智能聊天机器人在教育应用领域的研究以中国、美国为核心，主要集中于欧美等发达国家。人工智能聊天机器人在教育应用领域的研究趋势正朝着更加深入和多元化的方向发展，研究主题已从最初的人工智能聊天机器人对学习动机的影响，逐步扩展到学习效果、用户接受程度等多个方面。此外，教学应用研究、实证研究和理论研究三者并驾齐驱，成为研究的热点。依据研究结论与当下国内外人工智能聊天机器人的教育应用现状，得出以下相关启示与建议。

顶层设计方面，在教育领域的研究、投入、应用与发展方面，教育行政部门应依据教育教学客观规律，出台相关优惠政策，从宏观政策角度鼓励该产业的发展，设立试点型学校，并组织相关研究机构，整合资源信息，构建人工智能与教育协同创新体系。

领域研究方面，领域研究主要集中在欧美地区，虽然理论技术方面已经呈现爆发式发展趋势，但教育实践方面仍有一定不足。因此，建议技术团队与教育研究者广泛开展国际合作，探究人工智能聊天机器人教育应用新模式。

教育实践方面，建议研究团队加强对生成式人工智能的研究，将研究成果转化为实际应用，开发基于人工智能聊天机器人的教育软件，革新教育教学方法，教学者应积极适应新型技术，根据技术环境和教学情况，利用技术反哺教学，助力国内教育现代化发展。

参考文献

卢宇、余京蕾、陈鹏鹤和李沐云(2023)。生成式人工智能的教育应用与展望——以 ChatGPT 系统为例。**中国远程教育, 43(04)**, 24-31+51。

Fidan, M., & Gencel, N. (2022). Supporting the Instructional Videos With Chatbot and Peer Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation. *Journal of Educational Computing Research, 60*, 1716 - 1741.

Lee, Y., Hwang, G.-j., & Chen, P.-Y. (2022). Impacts of an AI-based chabot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational technology research and development, 70*, 1843 - 1865.

Liu, C.-C., Liao, M.-G., Chang, C.-H., & Lin, H.-M. (2022). An analysis of children' interaction with an AI chatbot and its impact on their interest in reading. *Comput. Educ., 189*, 104576.

Okonkwo, W. C., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Comput. Educ. Artif. Intell., 2*, 100033.

Wu, R.-f., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*.

Zhang, J., Oh, Y. J., Lange, P., Yu, Z., & Fukuoka, Y. (2020). Artificial Intelligence Chatbot Behavior Change Model for Designing Artificial Intelligence Chatbots to Promote Physical Activity and a Healthy Diet: Viewpoint. *Journal of Medical Internet Research, 22*.

应用大型语言模型生成数学通用型提示的研究

A study of large language models in generating mathematical universal tips

张誉月[1]，贾积有[1*]

[1]北京大学教育学院教育技术系

persistbetter@stu.pku.edu.cn，[*]jjy@pku.edu.cn

【摘要】大语言模型（Large Language Models, LLMs）在多个学科教学中展现了显著的教育应用潜力，但在直接解答数学题目方面仍面临很大挑战。数学通用型提示通过展示解题步骤而非答案激发学生独立思考，然而，这种方法要求教师投入大量时间精力来设计和优化提示。为探索借助 LLMs 促进数学教学的新途径、减轻人类教师的工作负担，本研究探讨了如何利用 LLMs 自动产生高质量的数学通用型提示。实验选取了 16 道涵盖几何、代数、概率等具有代表性领域的九年级数学题目作为测试集，应用零样本学习策略优化提示词的结构。研究比较了人类教师和 4 种 LLMs 生成的通用型提示，结果表明 LLMs 生成的提示在逻辑性和条理性方面更具优势，但同时可能增加学生的认知负荷。此外，评估了不同 LLMs 的表现，发现其生成的提示质量由高到低排序为：星火大模型、GPT-4、GPT-3.5 和文心一言。综上所述，本研究揭示了结合生成式人工智能与智能数学教学系统在中学数学教育中的应用潜力，并为未来的实证研究提供了可行的研究方向和方法。

【关键词】生成式人工智能；智能教学系统；数学教育；大语言模型

Abstract: *Large Language Models (LLMs) have demonstrated significant educational application potential across various academic disciplines, yet they still face considerable challenges in directly solving mathematical problems. Universal tips, which encourage independent thinking in students by demonstrating solution steps instead of providing answers, have been beneficial. However, this approach demands considerable time and effort from educators to design and optimize the prompts. To explore new pathways for enhancing mathematical instruction with the aid of LLMs and to alleviate the workload of human teachers, this study investigates how LLMs can be utilized to automatically generate high-quality generic mathematical prompts. The experiment selected 16 representative ninth-grade math problems encompassing areas such as geometry, algebra, and probability as the test set, applying a zero-shot learning strategy to optimize the structure of the prompt words. This research compared generic prompts created by human teachers with those generated by four different LLMs. The results indicate that prompts produced by LLMs have advantages in terms of logical coherence and organization, but may also increase cognitive load for students. Moreover, the performance of different LLMs was evaluated, revealing a quality ranking of generated prompts from highest to lowest as follows: Spark Large Model, GPT-4, GPT-3.5, and Wenxin Yiyan. Overall, this study unveils the potential application of generative artificial intelligence combined with intelligent mathematical teaching systems in secondary school mathematics education, and provides feasible research directions and methodologies for future empirical studies.*

Keywords: Generative AI; Intelligent teaching system; Mathematics education; Large language model

1. 引言：生成式人工智能在数学解题中的局限

随着信息产业的发展与算力瓶颈的突破，以深度学习神经网络为核心的人工智能技术取得了显著进步，在教育的宏观、中观和微观三个层面上，对教、学、评、测等教育系统的各个环节都发挥了重要作用（贾积有等，2023）。大语言模型（Large Language Models, LLMs）的快速发展为人工智能在教育领域的应用开辟了新的可能性。OpenAI 发布的测试结果表明，GPT-4 在 SAT 循证阅读和写作考试中获得 710 至 800 的高分（Achiam 等，2023）。尽管 GPT 模型的卓越性能令人印象深刻，它们有时会表现出“幻觉”现象（Lee 等，2023）。数学推理涉及大量细粒度的认知能力，如复杂的模式识别技能、语言理解、符号处理和抽象思维等，使其成为评估深度学习模型的关键评判标准之一（Testolin，2024）。聊天机器人如 ChatGPT 和 Google Bard 等可以为简单的算术、代数表达式和基本逻辑难题提供准确的解决方案，但面对

更复杂的数学问题或高级逻辑任务时，它们提供的答案并不可靠（Plevris 等，2023）。刘思怡等（2023）发现 GPT-4 的解题平均准确率仅为 68%。这些研究表明，GPT-4 无法完全理解一些复杂或模糊的问题，导致生成的答案可能不准确。因此，如何有效地将 LLMs 应用于数学教育领域，成为了一项具有挑战性的任务。

2. 相关研究

通用型提示是一种促进学生解决问题能力的数学教学方法，通过呈现关键步骤或思维过程，启发学生独立思考与探究，而非直接给出答案或详细的阐释。此方式能够有效提高学生解题技能、主动思考能力及知识掌握程度。贾积有（2023）等开发的“数学智能评测与辅导系统”（MIATS）融入了多级难度的通用型提示，实证研究表明该系统显著促进了学生的数学学习。然而，现阶段的通用型提示仍依赖于教育工作者对每个问题手工编制，教育工作者需深度分析每个问题的关键与常见错误，人力成本极为昂贵。Jia 等人基于已有数据集分别让人类教师和大型语言模型生成若干组通用型提示，将提示混合后让专家打分，统计分析结果表明人类专家无法区分人工和机器产生的通用性提示（Jia 等，2024）。由此可见，尽管大型语言模型在逐步推理和高度精确性的数学问题解答任务中的性能尚不理想，但其在文本摘要生成领域却成效显著。基于此，本研究探索使用大语言模型模型来生成通用型提示，以最大限度地减少大型语言模型在幻觉问题等方面带来的影响。

3. 基于大型语言模型生成通用型提示的通用架构设计

3.1. 提示编写方法

提示词质量对于生成通用型提示的效果有着显著的影响。作为生成通用型提示的关键变量，不同类型和构造的提示词对模型的理解和输出结果产生不同的效果。大模型提示词的构造方法多样，包括零样本学习（Zero-Shot Learning）、单样本学习（One-Shot Learning）和少样本学习（Few-Shot Learning）等。零样本学习不依赖于特定任务的示例，而是利用事先训练好的语言模型和简洁的提示语，以期模型依靠内部知识库推导出正确的结果（Kojima et al., 2022）。单样本学习则通过在提示语上下文中嵌入一个相关任务示例，借此引导大型语言模型进行学习，无需对模型参数进行调整（如微调等）（Brown et al., 2020）。少样本学习策略采用有限数量（一般是两到五个）的任务相关示例进行提示构建，旨在促使大型语言模型从受限的示例中抽取出有效的特征和知识，以期在未知的任务中能够展现出更为卓越的性能。

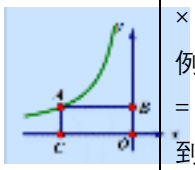
Jia 等人的研究表明单样本和少样本学习会导致 GPT 过度依赖给出的例子，而忽略了数学问题的具体特点和变化，而零样本学习使 GPT 能够根据具体问题自由生成提示，产生的通用型提示得分最佳（Jia 等，2023）。因此本文主要采用零样本学习的方法。

3.2.提示设计结果

得益于在线搜题平台的蓬勃发展，解答过程和详细解析的获取变得十分容易。目前国内外主流的大模型分别有 GPT-3.5、GPT-4、科大讯飞的星火大模型（V3.5）和百度的文心一言大模型（V3.5），我们使用相同的提示词让人类教师和 4 种大模型对 16 道初中九年级的数学题目产生提示。下表将这 4 类大模型生成的提示与人类教师设计的提示进行展示，分别选取了具有代表性的代数题、概率题、几何题：

表 1 人类教师和 LLMs 产生的提示

题目	人类教师	GPT-4 模型	GPT-3.5 模型	星火大模型	文心一言
如图所示，这是反比例函数	1、观察图像，确定点 A 位于第二象限，因	1、识别图形与函数的关系： 从图中识别反比例函数 y=k/x 的图	1、观察图象，找到点 A，并过点 A 作垂线 AB	1、回顾反比例函数的图形特征：反比例函数图像是一条通	1、确定反比例函数的形式为 y= k/x。

y=\(\frac {k}{x}\) 的图象的一支，点 A 在图象上，过点 A 作 AB⊥y 轴于点 B，AC⊥x 轴于点 C，且 S▭ABOC=12，那么反比例函数的解析式是 y=? 	此 A 的横坐标 x 为负数，纵坐标 y 为正数。 2、根据题目条件，矩形 ABOC 的面积为长乘以宽。 3、在此矩形中，长是 x 的绝对值，宽是 y 的值。 4、将矩形面积公式 S = 长 × 宽 与反比例函数关系 y = k/x 结合，得到 k = xy。 5、考虑 k 的正负	象，并注意到系数 k 的几何意义是矩形面积。 2、分析矩形面积：确定矩形面积是由垂直于 x 轴和 y 轴的线段围成。 3、利用矩形面积求系数 k：根据反比例函数的性质，用矩形 ABOC 的面积来表示 k 的值。 4、判断函数图象所在象限：观察图象在坐标系的哪个象限，以确定 k 的符号。 5、写出反比例函数的解析式：结合 k 的值和符号，写出完整的反比例函数解析式。	与垂线 AC。 2、注意反比例函数的性质，系数 k 表示垂线段与坐标轴围成的矩形面积。 3、计算矩形 ABOC 的面积 S，并注意 S 已给出为 12。 4、根据矩形面积与 k 的关系，得到 k 的值。 5、考虑反比例函数在第二象限，因此 k 应为负数，得到反比例函数的解析式 y = -k/x。	过所有四个象限的曲线，且在每个象限内都是连续的。 2、理解 k 的几何意义：在反比例函数中，k 代表的是点到两坐标轴垂线段与坐标轴围成的矩形面积。 3、应用面积公式：根据矩形面积公式，确定|k|的值。 4、考虑函数象限：分析函数图像所在的象限，以确定 k 的正负。 5、写出函数表达式：综合以上信息，得出反比例函数的解析式。	2、根据题目，点 A 在图象上，因此点 A 的坐标可以表示为 (x, x/k)。 3、计算矩形 ABOC 的面积，该面积等于|k|。 4、根据题目，S=12，因此 ｜k｜=12。 5、由于反比例函数的图象在第二象限，所以 k<0，因此 k=−12。 6、将 k=−12 代入反比例函数中，得到 y=−12/x。

3.3. 不同提示词的对比

1）GPT-4 和人类教师：GPT-4 模型的提示比人类提示的优点在于更具有条理性、更详细，所涉及到的具体数字和字母更少。缺点在于涉及了大量专业、枯燥的数学术语，在巩固基础理论的同时增加学生的认知负荷，人类提示则更凝练、更直观。

2）GPT-3.5：由于 GPT-3.5 没有图像识别功能，几何题目只能根据文字的解析描述给提示，较不精准，且出现了大量的题目中的具体数字和字母。相比之下，GPT-4 模型所给的提示更具有广泛性，涉及到更多的基础理论，其可推广性和可迁移性更强。

3）星火大模型：总体与 GPT-4 生成的提示很像，包括 5 个字左右的简短标题、不含具体数字和字母的指导以及涉及到的基础理论。其读图功能可以对几何题给出恰当的提示。而且相比于 GPT-4，星火的用语更符合中文习惯，且提示较短，对学生耐心的挑战较小。

4）文心一言：识图能力较弱，几何题的提示不太准确。仅仅对题目的解题答案进行了分步，没有在此基础上进行凝练总结，并且不免会涉及大量的数字和符号，在部分题目中还给出了远超 5 条以上的提示，因此提示的总体质量较差。

综上所述，根据 4 种大模型所生成的提示内容比较，总体提示质量由高到低的排序为：星火大模型、GPT-4、GPT-3.5 和文心一言。

4. 结论和讨论

为探讨如何将大型语言模型更好地应用于数学教育，本研究基于融合 LLMs 和智能教学系统 MAITS 的中学数学教学方案，选取了 16 道涵盖几何、代数、概率等具有代表性领域的九年级数学题目作为测试集进行实验，应用零样本学习策略优化提示词的结构，对比了人类教师和 4 种大模型所生成的提示内容，发现整体的提示质量由高到低排序为：星火大模型、GPT-4、GPT-3.5 和文心一言，此外 GPT-4 在某些方面可能更有助于学习，但也可能增加学生的认知负荷。Gattupalli 等（2023）利用 GPT-4 生成了解决四年级数学应用题的提示和策略，

与人类教师编写的提示进行了比较，发现相比于人类教师，GPT-4 提供了文本较多、且主要是程序性的解题策略。因此，GPT-4 和人类教师产生提示的效果需要进一步进行实证研究。

综上所述，尽管 LLMs 在数学教学的应用方面存在限制，它们仍能为教学质量的提升提供了新的可能性。该教学方案可推广至物理、化学、工程等其他学科的教学和智能教学系统中。该方案实践上的难点在于需要确保题目解答过程的完整性和准确性，未来的研究可以进一步验证融合 LLMs 和智能教学系统的教学方案，并通过实证研究对比不同提示和教学策略的效果，为数学教育的实际应用提供更深入的分析。

致谢

本文是“东湖高新区国家智能社会治理实验综合基地”项目（课题：“人工智能赋能地方教育高质量发展”）的一项研究成果。

参考文献

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.

Engineering Systems (INES). IEEE, 2023: 000161-000164.

Gattupalli, S., Maloy, R. W., & Edwards, S. (2023). Comparing teacher-written and AI-generated math problem solving strategies for elementary school students: Implications for classroom learning.

Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. New England Journal of Medicine, 388(13), 1233-1239.

刘思怡，傅胤荣，王畅，郑耿忠 & 孔维宏. (2023). GPT-4 人工智能技术在基础教育的应用与展望. 韩山师范学院学报 (06), 18-26+48.

贾积有 & 张誉月.(2023).人工智能与教育:机遇、挑战与对策. 北京大学教育评论 (01),49-61+188-189.

贾积有,张誉月,刘怀亚 李双双.(2023).智能评测和辅导系统助力学生评价改革和减负增效.电化教育研究(06),74-80+89.doi:10.13811/j.cnki.eer.2023.06.010.

Jia. J., Wang, T , Zhang, Y., Wang, G. (2024).The Comparison of General Tips for Mathematical Problem Solving Generated by Generative AI with those Generated by Human Teachers. Asia-Pacific Journal of Education (2024)

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. Advances in neural information processing systems, 35, 22199-22213.

Plevris, V., Papazafeiropoulos, G., & Rios, A. J. (2023). Chatbots put to the test in math and logic problems: A preliminary comparison and assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard. arXiv preprint arXiv:2305.18618.

Tang, R., Zhang, Y., Cao, Y., Liu, H., & Jia, J. (2023, July). Design and Effect of Guided and Adaptive Tutoring Tips for Helping School Mathematics Problems Solving. In International Conference on Blended Learning (pp. 273-284). Cham: Springer Nature Switzerland.

Testolin, A. (2024). Can neural networks do arithmetic? a survey on the elementary numerical skills of state-of-the-art deep learning models. Applied Sciences, 14(2), 744.

生成式人工智能在教育中的应用—基于 Citespace 的可视化分析

Generative AI in Education–Visual Analysis Based on CiteSpace

白佳鹭[1*]，陈鹏[2]，叶禹杉[3]
首都师范大学 教育学院
* Baijialu2001@163.com

【摘要】 以 ChatGPT 为代表的新一代生成式人工智能推动了教育的数字化变革，促进了师生的角色转变。文章以 CNKI 中的 CSSCI 数据库以及 web of science 核心合集中的 SSCI 和 SCI 数据库中的相关文献为研究对象，基于“生成式人工智能+教育”、“ChatGPT+教育”主 进行检索，运用 CiteSpace 等软件分析了国内外生成式人工智能在教育中应用的应用现状和研究热点，并对目前的研究提出了展望，以期为国内外生成式人工智能在教育中相关研究提供参考。
【关键词】 生成式人工智能；ChatGPT；教育应用;可视化分析

【Abstract】 *The new generation of generative artificial intelligence represented by ChatGPT has promoted the digital transformation of education and promoted the transformation of the roles of teachers and students. This paper takes CSSCI database in CNKI and relevant literature in SSCI and SCI database in web of science core collection as research objects, and searches based on the topics of "generative artificial intelligence + education" and "ChatGPT+ education". CiteSpace and other software are used to analyze the application status and research hotspots of generative artificial intelligence in education at home and abroad, and the prospect of the current research is put forward, in order to provide reference for the relevant research of generative artificial intelligence in education at home and abroad.*
【Keywords】 Generative Artificial Intelligence; ChatGPT; Educational Applications; Visual Analytics

1. 前言

国务院出台的《新一代人工智能发展规划》以及教育部发表的《 等学校人工智能创新行动计划》等政策彰显了人工智能在教育中的重要地位和关键作用。随着技术的不断创新与发展，以 ChatGPT 为代表的新一代生成式人工智能（Generative Artificial Intelligence, GAI），引发了教育行业的研究热潮。但由于其固有缺点，在教育中的融合性与适用性引起了研究者们的持续探讨。本研究基于文献计量学方法，利用可视化工具 CiteSpace 对国内外 GAI 在教育 域的文献进行分析，进而探索国内外 GAI 在教育中的研究现状及热点，以期为国内 GAI 教育研究提出建议。

2. 研究数据和方法

本研究以（TS=（“generative artificial intelligence (GAI)” OR “generative AI” OR “ChatGPT”）AND TS=（“teaching” OR“ education”））为检索式在 WOS 核心合集并以“生成式人工智能+教育”、“ChatGPT+教育”为检索词在 CNKI 的 CSSCI 数据库中检索中英文文献。通过滚雪球式检索后共得到英文文献 274 篇，中文文献 169 篇。剔除不符合要求以及重复文献，最终得到有效文献 248 篇，其中英文文献 125 篇，中文文献 123 篇。研究运用 CiteSpace 软件对其进行可视化分析。

3. 新一代生成式人工智能教育的研究现状

3.1 主体分析

通过研究者分析可知,国内共有 45 位作者发表 GAI 主 文献，产生 23 次合作。国外共有 38 位作者，产生 85 次合作。由于目前 GAI 是前沿 域，核心研究者较少，研究者之间的合作较为松散，未来还有待加强。

通过发文国家分析可知，美国、中国、澳大利亚是发文量相对较多的国家，但其他国家发文量较少。因此，建议在该方面发文量较多的国家与发文量较少的国家可以进行深度合作，以带动该 域在各国地区都能有进一步的发展。

通过对国内研究机构的分析可知,在 123 篇 CSSCI 文献当中，有 17 篇文献来自华东

师范大学，其次是北京师范大学 10 篇，华中师范大学 9 篇，南京师范大学和北京大学均为 4 篇。但目前各个院校间的合作并不紧密，还需要各院校之间的深度合作。

3.2 热点分析

由关键词 次分析可知，目前国内关于 GAI 在教育中的研究热点为人工智能带来的教育变革、在 等教育中应用生成式人工智能的讨论、生成式人工智能带来的教育生态的变化、生成式人工智能对教育挑战的应对策略以及人机协同等问 。国外研究主要集中于 等教育、医学教育以及人工智能依赖的大语言模型开发和对学生带来的影响等方面。

3.3 现状分析

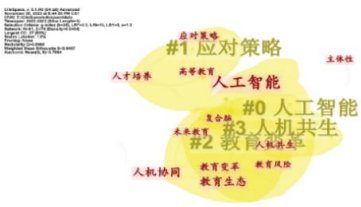


图 1 国内“GAI+教育”的聚类图谱

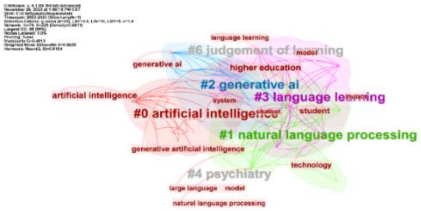


图 2 国外“GAI+教育”的聚类图谱

由图 1 可知，新一代 GAI 的出现引发了国内教育各个 域的研究热潮，给教育生态带来了冲击，推动了教育的数字化变革。人机协同发展是应对教育数字化变革的一种有效方式。GAI 有助于培养创造性和创新性人才，但其无法像人类一样进行独立思考和判断推理，在教育 域应用时仍然会有很多诸如教育 险，因此目前其主要应用于 等教育阶段，研究者也在不断探究应对其 险的人机共生策略。

由图 2 可知，国外 GAI 在教育的研究尤为广泛。GAI 凭借其突出的自然语言处理能力和大语言模型，在医学、文学、计算机等 域掀起了应用热潮。但需要用户不断尝试输入的提示词对其进行训练，并不断更新语言模型，进行深度学习才能不断满足人类的需求，因此研究者们目前正在研究如何训练其语言模型评估其答案的准确性，并研究教育者如何调整自己的状态和角色以及如何在课堂中正确使用 GAI 技术 (Cooper, 2023)。

4. 研究展望

4.1 指向师生的数字化素养和 阶思维

数字化素养和 阶思维是使用新一代生成式人工智能需要具备的素养与能力，未来研究还需将理论应用于实践，探究师生使用 GAI 以及使其更好地融于教育的策略方法。

4.2 指向教育应用的实证研究

目前 GAI 运用于教育的探索仍处于初级阶段（杨宗凯等，2023），国内生成式人工智能主要以理论探究为主，缺乏从实证角度的研究。因此，未来还需从实证角度不断探究如何将其有效纳入我国教育场景，如何有效运用于教育教学当中。

4.3 教育大模型的设计与开发

GAI 模型要在通用大模型的基础上结合教育 域的专业数据和专家优质资源，整合各类智能教育技术，以适合其教育场景。目前的 GAI 技术在学术研究和应用上还有很多不足之处需要改进（张熙等，2023）。因此，教育大模型的设计与开发也是未来生成式人工智能在教育应用的发展趋势之一。

参考文献

杨宗凯,王俊,吴砥 陈旭.(2023).ChatGPT/生成式人工智能对教育的影响探析及应对策略.华东师范大学学报(教育科学版)(07),26-35.

张熙,杨小汕 徐常胜.(2023).ChatGPT 及生成式人工智能现状及未来发展方向.中国科学基金(05),743-750.

Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. Journal of Science Education and Technology, 32(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

支持教師基於個別學生閱讀歷程建構願望書單與聊書機制以延伸學生閱讀興趣

Supporting Teachers Constructing Wish Booklists and Implementing Book Chat Mechanisms based on Individual Students' Reading Profiles to Extend Students' Reading Interests

涂弘旻[1*]，廖長彥[2]，陳德懷[3]
[1.3]臺灣中央大學 網路學習科技研究所
[2]臺灣中央大學 客家語文暨社會科學學系
[*]benjamintu1514@gmail.com

【摘要】 閱讀是不可或缺的技能，有助於我們吸收不同領域的知識。研究團隊先前曾提出「願望書單」概念，試圖解決部分學生在 MSSR 活動中閱讀偏食的狀況，然而，目前仍存在教師在願望書單活動中缺乏影響力的問題。為此，研究團隊在先前已針對國小教師和學生進行訪談，分析出教師在願望書單活動中所面臨的困難，並重新設計願望書單的執行流程和解決方案，因此，本研究將根據先前訪談結果，並基於明日閱讀 2.0 的理念，發展出一套智能化系統，導入進願望書單活動中，期望能解決教師在班級中執行願望書單活動所遇到的困難。
【關鍵字】 願望書單；明日閱讀 2.0；閱讀偏食；AI 輔助教學

Abstract: *Reading is an indispensable skill, aiding us in absorbing knowledge across diverse domains. The research team previously proposed the concept of a "Wish Booklists" to address students' reading preferences. However, there still exists a problem of teachers lacking influence in Wish Booklists activities. To address this, the team interviewed elementary school teachers and students, identifying teachers' difficulties in wish booklists activities. And redesigned the execution process and solutions for wish booklists activities. This study, based on prior interviews result and Reading for Tomorrow 2.0 principles, aims to develop an intelligent system for wish booklists activities. The system will be implemented in actual teaching settings with the hope of teachers overcoming the challenges implementing wish booklists activities.*
Keywords: *Wish Booklists, Reading for Tomorrow 2.0, Reading Preferences, AI-Assisted Teaching*

1. 前言

在現今社會中，閱讀已被視為終身學習的基石（陳德懷，2016）。學生成長的過程中，閱讀不僅擔任著關鍵性的角色，更是一項至關重要的技能。明日閱讀 1.0 是基於「身教式持續安靜閱讀(MSSR)」的模式進行推廣，鼓勵學生依照興趣選擇書籍閱讀，並由師長以身作則，引領學生一同閱讀（陳德懷，2016），但長期的研究發現，讓學生自由選擇書籍閱讀，會導致部分學生產生閱讀偏食的問題。為了改善此問題，研究團隊發展出明日閱讀 2.0，以「廣博閱讀」、「大量閱讀」和「登台」三個方向進行推動，鼓勵學生接觸不同種類的書籍，在此基礎之下，提出「願望書單」的概念，在學生自由選書的同時，讓教師向學生推薦書籍，讓師生雙方共同建立閱讀目標，以此引導學生閱讀較少接觸的書籍種類（賴瑞霖，2021）。然而現今在班級中執行願望書單活動時，仍遇到教師缺乏影響力的問題，導致成效不彰，無法有效解決學生閱讀偏食的問題。為此，研究團隊在先前已經針對桃園市某實驗國小的二至六年級的教師和學生進行訪談，並分析出教師在願望書單活動時會面臨到「教師對於班級圖書的掌握程度不足」、「紙本願望書單無法取得個別學生的閱讀歷程」、「師生聊書活動缺乏紀錄工具」這三個主要問題，而整體願望書單活動的流程為「學生選書」、「教師薦書」、「學生閱讀書

籍」、「師生聊書」。基於上述，本研究將基於「趣創者理論」(Chan et al., 2018)和「明日閱讀 2.0」的理念，發展一款智能化系統，將其導入進願望書單活動中，並評估導入智能化系統後，對於教師的工作效率和教學成效影響。

2. 研究設計

本研究的研究參與者為桃園市某實驗國小二至六年級的語文教師，而研究方法採用研究設計法，先前已經根據訪談內容分析出教師面臨到的問題與困難，為了改善教師所面臨的問題與困難，本研究將基於明日閱讀 2.0 的理念，並以人工智慧的觀點，開發一套智能化系統，並導入願望書單活動中，取代原先的紙本操作模式。系統在「教師薦書」的環節中，會呈現班級書櫃中的書籍、個別學生的選書紀錄，以及閱讀歷程等資料，並自動生成符合學生的推薦候選書單，供教資訊師參考，在「師生聊書」的環節，利用錄音的方式記錄師生聊書的過程，並透過 OpenAI 公司提供的 Whisper 和 GPT 的 API 服務，將錄製完的語音轉換成文字摘要和建議，幫助教師了解個別學生的閱讀狀況、建立個別學生的聊書歷程檔案，以及減輕教師在聊書過程中負擔。而研究人員則在「師生聊書」環節進入班級中觀察，記錄研究歷程以了解教師在系統的輔助下與學生進行聊書的狀況，並根據系統操作日誌和課堂觀察記錄，於每月底與教師們進行訪談，了解教師們對於系統功能和操作上的意見和建議，以便改進系統，確保系統能夠更貼近實際教學現場的需求，提升願望書單活動的執行效益。

3. 初步結果與結論

本研究旨在探討願望書單活動導入智能化系統後，對於教師在活動中所扮演的角色、工作負擔，以及教學成效的影響。根據課堂觀察和系統操作日誌來看，導入智能化系統，對於減輕教師的工作負擔和教學成效方面有著正向的影響，在薦書環節，教師們會察看系統提供的學生閱讀歷程、書籍摘要，並使用篩選器來選擇符合個別學生的書籍，挑選合適的書籍推薦給學生，而在師生聊書的過程中，教師們也會使用系統來錄製聊書內容，並且根據 AI 生成出的聊書重點摘要來了解學生的閱讀狀況，也會根據 AI 生成出的建議來輔導學生，提高了師生聊書的品質與效率。因研究流程會持續整個學期，後續研究團隊會持續入班觀察，以及在學期末與各班的教師進行訪談，訪談內容則會參考先前研究使用的工作滿意度與自我效能量表(Noy & Zhang, 2023)。依據課堂觀察、系統操作日誌和教師訪談的結果，調整願望書單的活動流程和系統設計，以解決願望書單活動中，教師所面臨到的問題與困難。

致謝

本研究在臺灣國科會人文處 (111-2410-H-008 -013 -MY3) 與「臺灣中央大學學習科技研究中心」的資助下完成，僅此致謝。

參考文獻

賴瑞霖（2021）。閱讀島：應用數位閱讀歷程擴大學生閱讀舒適圈及提升閱讀動機之系統設計。〔未出版之碩士論文〕。臺灣中央大學網路學習科技研究所。

陳德懷（2016）。明日閱讀：明日主題學習的基礎。台灣：天下雜誌。

Chan, T. W., Looi, C. K., Chen, W., Wong, L. H., Chang, B., Liao, C. C., ... & Ogata, H. (2018). Interest-driven creator theory: Towards a theory of learning design for Asia in the twenty-first century. *Journal of Computers in Education*, *5*(4), 435-461.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283*.

国内外教育人工智能伦理可视化分析

Visualization Analysis on Ethics of Educational Artificial Intelligence at Home and Abroad

贺玮[1]，廖林锟[1*]，颜京莉[1]，牛雯婷[1]，赵东东[1]
[1]华南师范大学，教育信息技术学院
[*]3104093211@qq.com

【摘要】 为深入了解国内外教育人工智能伦理（E-AIED）的研究热点与前沿，本文利用 CiteSpace 对 CNKI 和 WOS 数据库中 2017-2023 年的 E-AIED 文献进行了可视化分析。结果表明：E-AIED 研究正处于快速发展阶段；国内 E-AIED 学术共同体相对国际更具代表性和稳定性；国内 E-AIED 研究的热点主题偏向理论探索，国际则更关注应用实践探究；大数据和机器学习相关的伦理问题是 E-AIED 的研究前沿。
【关键词】 教育人工智能，伦理，可视化分析，热点

Abstract: *In order to explore the research hotspots and frontiers of the E-AIED between Chinese and international publications from 2017-2023, this study used CiteSpace to arrive at a visualization analysis. The findings showed that: The research of E-AIED has entered a stage of rapid development; the academic community of E-AIED in China is more representative and stable than the international; the hot topics of E-AIED in China are more inclined to theoretical exploration, while they are more inclined to the application of E-AIED in the international. Big data and machine learning are at the forefront of E-AIED research.*
Keywords: Artificial Intelligence in Education, Ethics, Visualization Analysis, Hotspots

1.引言

人工智能给教育带来机遇的同时，也对教育发展带来一系列伦理问题，学者们对教育人工智能伦理（Ethics of Artificial Intelligence in Education，E-AIED）的原则和风险规避等进行了有益的探讨（胡小勇等，2022）。如何在充分发挥新一代人工智能技术对教育应用价值的同时，有效规避其中的伦理风险，对于加强人工智能与教育的创新融合具有重要意义。因此，本研究将借助 CiteSpace 对 2017-2023 年国内外 E-AIED 的相关文献进行了梳理、分析和呈现。

2.方法

中文数据来源于 CSSCI，以“人工智能教育”“教育人工智能”“伦理”为检索词，发表时间设定为 2017 年至 2023 年，得到文章 400 篇。英文数据来源于 WOS 数据库核心数据合集，以“artificial intelligence in education”“AIED”“ethic”为检索词进行检索，时间设置为 2017 年至 2023 年，共得到 223 篇有效文章。采用 CiteSpace 对这 623 篇相关文章进行可视化分析。

3.数据统计结果

3.1. 年度发文量

2017-2023 年是国内外 E-AIED 研究成果的快速增长阶段，其年度发文量如图 1 所示。

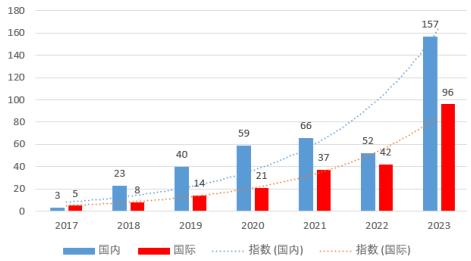


图 1 国内外 E-AIED 年度发文量

3.2. 作者合作图谱

国内 E-AIED 已形成多个作者合作体（图 2），国际则较为分散（图 3）。

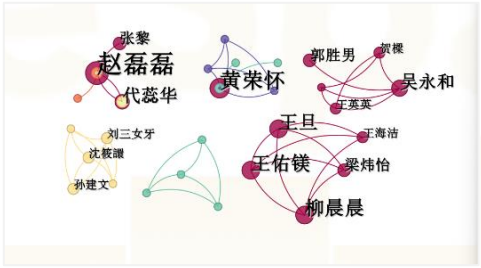


图 2 国内 E-AIED 作者合作图谱

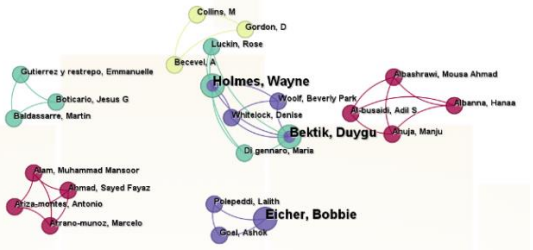


图 3 国际 E-AIED 作者合作图谱

3.3. 机构合作图谱

国内 E-AIED 研究的高产机构单点结构少，稳定结构多（图 4）；国际则单点结构多，存在一定的稳定结构（图 5）。



图 4 国内 E-AIED 机构合作图谱

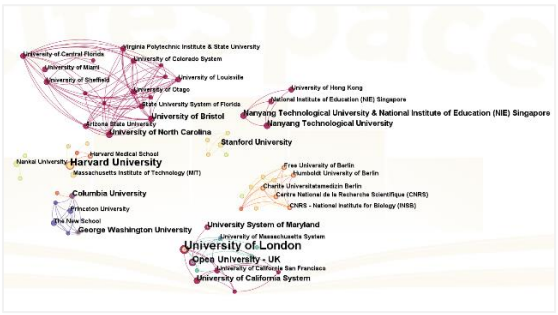


图 5 国际 E-AIED 机构合作图谱

3.4. 关键词共线图谱

从国内 E-AIED 高频词来看，“人工智能”出现的频率最高（图 6）；国际的高频词关键词有“人工智能”“机器学习”等（图 7）。

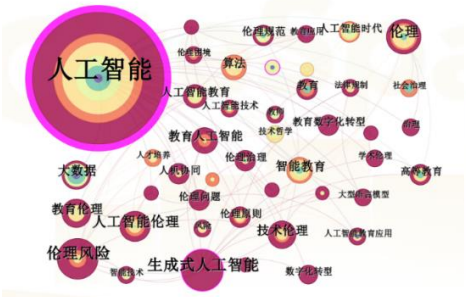


图 6 国内 E-AIED 关键词共线图谱

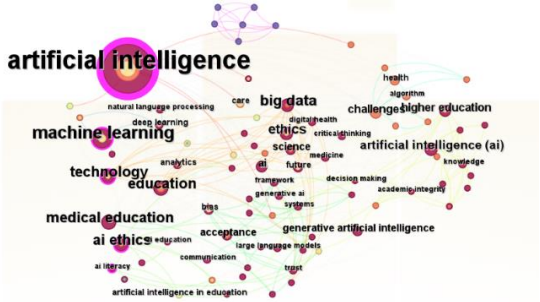


图 7 国际 E-AIED 关键词共线图谱

3.5. 关键词突现图谱

国内突现性最高的是大数据（图 8），国际突现性最高的是机器学习（图 9）。

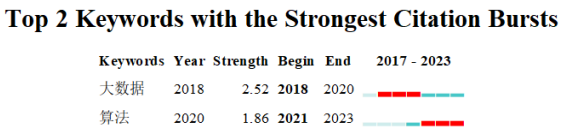


图 8 国内 E-AIED 关键词突现图

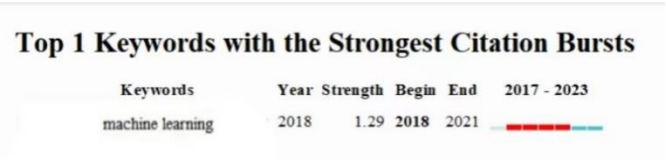


图 9 国际 E-AIED 关键词突现图

4.结论

E-AIED 研究成果整体呈快速增长趋势，特别是在 2018 年后，国内发文量大幅度上升；国内外在 E-AIED 研究领域研究者及其研究机构已经具有了一定的学术影响力 ;国内 E-AIED 的相关研究聚焦宏观理论的相关探索，而国际研究则更关注教育人工智能的技术伦理与应用实践；大数据和机器学习相关的伦理问题研究有可能成为未来一段时间的 E-AIED 研究前沿。

参考文献

胡小勇,黄婕,林梓柔,等(2022).教育人工智能伦理：内涵框架、认知现状与风险规避. 现代远程教育研究,34(02):21-28+36.

多模态大模型的教育应用价值、典型场景与发展建议

The Educational Application Value, Typical Scenarios, and Development Suggestions of Multimodal Large Models

卜浩德[1]，李新[2]，杨现民[3*]
[123]江苏师范大学智慧教育学院
[*] yangxianmin8888@163.com

【摘要】大模型的发展经历了探索期、积累期、突破期、快速发展期，其技术原理可以总结为训练过程、构建方法、核心能力三个方面。文章认为多模态大模型的教育价值体现在学习分析、数字资源供给、智适应学习。文章将多模态大模型教育应用场景总结为教育评价、资源管理、学习辅导、教育研究。为推进多模态大模型教育应用发展，需要政府、教师、企业、区域管理者共同发挥自身独特的作用。
【关键词】 多模态大模型；教育价值；教育场景；发展建议

Abstract: *The development of multimodal large models（MMLM） has gone through stages of exploration, accumulation, breakthrough, and rapid growth. The technical principles can be encapsulated in three main areas. This article posits that the educational value of MMLM lies in learning analysis, resource provision, and adaptive learning. The article summarizes the educational application scenarios of MMLM into educational assessment, resource management, learning assistance, and educational research. To advance the development of MMLM in education, it is necessary for governments, teachers, enterprises, and regional administrators to each play their unique role.*
Keywords: Multimodal Large Models, Educational Value, Educational Scenarios, Development Recommendations

基金项目：国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”（项目编号：62077030）;2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）

1. 问题提出

当前多模态大模型凭借其在多模态理解、生成上的强大能力已经在医疗、广告等多个领域得到广泛应用。在教育领域多模态大模型同样具有巨大潜力，有望改变现有的人才培养模式，构建起规模化与个性化结合的精准教学体系(叶新东&刘泽民, 2024)。然而当前多模态大模型并没有真正触动教育变革，因此有必要深入研究多模态大模型的教育价值和教育场景。

2. 多模态大模型的发展历程

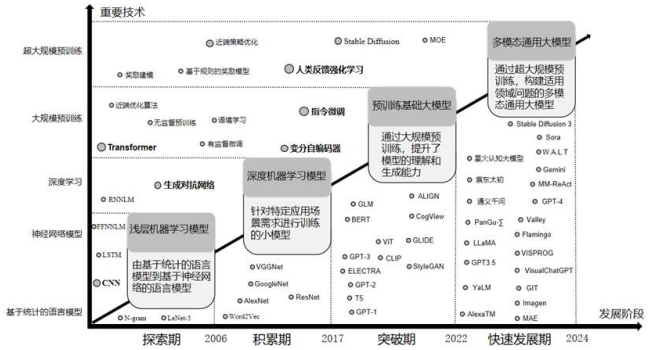


图 1 多模态大模型发展阶段

回顾多模态大模型的发展历程有助于深入探究多模态大模型的技术原理和技术特征，进而正确理解其教育价值。因此本文系统梳理了多模态大模型的产生与发展过程并归纳概括为四个阶段，分别是探索期、积累期、突破期和快速发展期四个阶段，如图 1 所示。

3. 多模态大模型的技术原理及教育价值

3.1 技术原理

大模型在发展过程中逐渐实现了从单模态向多模态的跨越。其技术虽然涉及复杂的算法和架构，但可以简单归纳为训练过程、构建方式、核心能力三个方面，如图 2 所示。分析这些技术原理可以进一步明确多模态大模型的技术优势以及在教育中的应用价值。

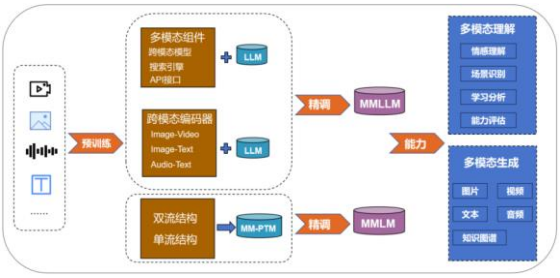


图 2 多模态大模型技术架构

3.2 教育价值

多模态大模型记忆了海量的常识和知识，经过教育领域数据的微调后可以理解教育场景、教育对象、教育资源和教育过程等多方面要素。充分利用多模态大模型的技术特征，可以实现助力多模态学习分析、强化数字资源供给能力、赋能智适应学习等教育价值。

4. 多模态大模型的教育应用场景

本文分析了国内外众多的大模型教育应用案例，发现目前教育大模型都展现出了良好的教育价值，能够涵盖当前主流的学科内容。在此基础上，本文将多模态大模型的典型教育应用概括为教育评价、资源管理、学习辅导、教育研究四个方面，如图 3 所示。

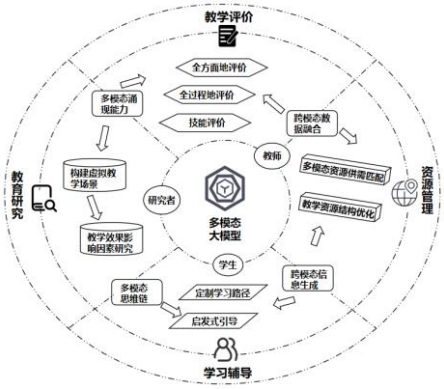


图 2 多模态大模型的教育应用场景

5. 多模态大模型的教育应用建议

本文参考国际社会的经验，结合国内教育发展的现实需求，提出了推进多模态大模型教育应用的四条建议：政府部门优化管理制度，推进多模态大模型监管治理；一线教师变革教育理念，提升多模态大模型的应用素养；技术企业加速多模态大模型产品研发，提升模型的服务品质；区域管理者发挥地方优势，构建多模态大模型区域推广机制。

参考文献

叶新东 & 刘泽民.(2024).基于多模态大模型的精准教学支持体系构建研究.远程教育杂志(01),84-93.doi:10.15881/j.cnki.cn33-1304/g4.2024.01.010.

基于国际视野下的中小学 AI 课程内容设置建议

Suggestions for AI Curriculum Content Setting in Elementary and Middle Schools From International Perspectives

谭亚妮[1]，王丽英[2*]
[1]南京师范大学教育科学学院
[2]南京师范大学教育科学学院
[*] wangliying@njnu.edu.cn

【摘要】近年来，AI 教育迅速发展。然而与国际部分发达国家已有明确的体系相比，我国 AI 课程的标准和内容设置尚存在割据分歧。因此，本文通过阐述我国中小学 AI 教育的现状及潜在问题，进一步从国际视角梳理美国、新加坡的 AI 教育发展思路，基于此，提出中小学 AI 课程内容设置的建议。
【关键词】 AI 教育；目标定位；课程内容；智能素养；课程体系

Abstract: *In recent years, AI education has developed rapidly. However, compared with the international part of the developed countries have a clear system, the standard and content setting of China's AI curriculum is still divided. Therefore, this paper elaborates on the status quo and potential problems of AI education in primary and secondary schools in China, and further combs through the development ideas of AI education in the United States and Singapore from an international perspective, based on which, it puts forward suggestions for the content setting of primary and secondary school AI courses.*

Keywords: AI Education; Goal Orientation; Curriculum Content; Intelligent Literacy; Curriculum System

1. 引言

AI 教育存在科学理论积淀、技术应用甄选以及专业师资缺乏和基础设施不足等诸多挑战。此外还存在没有统一课标、教材质量参差不齐等问题。本研究从现实需求出发，借鉴国际经验，对中小学 AI 课程的内容设置加以反思和建构，以期推动国内 AI 教育的高质量发展。

2. 研究综述

2.1. 国内中小学阶段开设 AI 教育的探索

2017 年 7 月，党中央、国务院共同出台了《新一代 AI 发展规划》。2018 年 4 月，教育部印发《教育信息化 2.0 行动计划》。2022 年 4 月，教育部印发《义务教育信息科技课程标准(2022 年版)》，在初中阶段加入了 AI 模块——人工智能与智慧社会。基于此，本文进一步梳理文献更全面地认识 AI 课程建设的国际经验，更好地探索我国的 AI 课程开发。

2.2 国际经验

2018 年 5 月，美国设立了 AI 教育工作组，旨在推动 K-12 人工智能教育行动(AI for K-12 Initiative)[1]。2017 年，新加坡推出了“新加坡人工智能”(AI Singapore，AISG)计划，该计划在经济、政策领域为 AI 提供支持，并在教育培训等方面进行了多方面创新[2]。

3. 对比分析

3.1 AI 教育课程目标对比

通过总结相关文献内容，本文针对 AI 课程教育目标提出如下建议：培养学生 AI 学科核心素养，培养学生适应智能社会发展的能力。

3.2 AI 课程内容的广度与深度对比

在广度方面，美 AIK12 指南和中 AI 指南所含有的知识点分别是 146 个和 94 个。前者所含有的知识点比后者多 52 个[3]。在课程深度方面，本文将课程深度数据整理为图 1：

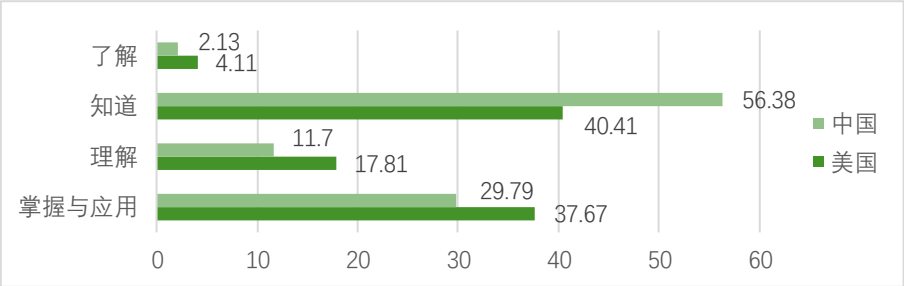


图 1　中美 AI 课程深度的比较

3.3 AI 课程设置建议

AI 课程内容应体现学科特点；AI 课程内容结构应组织灵活并相互联系；可适当提升课程内容难度和要求。

4 启发与讨论

4.1 课程标准需要具体化

通过与国际 AI 相关标准文件对比可知，国内在课程标准制定方面仍有欠缺，国家层面对 AI 教育定位的差异，会使 AI 教育的教学目标难以准确定位，从而无法落实教育教学评价。

4.2. 课程体系需持续完善

目前，AI 教育在我国只在高中学段涉及，且主要是在信息技术课程标准中对 AI 知识的学习提出要求，而不应使其局限于高中，需要被广泛地纳入到各学段。

4.3 课程资源需持续建设

中小学 AI 课程的知识体系与时俱进，内容选择上需要与 AI 技术发展保持同步，要有一定的前瞻性、开放性、灵活性，允许课程中技术内容修订的可持续性。

5. 结语

本文通过明晰我国中小学 AI 教育的现状；进一步从国际视角梳理美国、新加坡的 AI 教育发展思路；基于此，提出了一系列具体可行的中小学 AI 课程内容设置建议，包括课程目标、内容与课程资源等方面，期望能够为中小学 AI 教育提供有益的参考和借鉴。

参考文献

AAAI(2018). AAAI Launches “AI for K-12” Initiative in collaboration with the Computer Science Teachers Association(CSTA) and AI4All.

徐鹏、董文标、王丛（2022）。新加坡人工智能终身教育体系现状及启示。**现代教育技术，32（01）**，35-43。

李淑文、史宁中（2012）。中日两国初中几何课程内容的比较研究。**全球教育展望，41（01）**，82-85。

从学习风格的角度探究在线协作学习中的批判性思维发展差异

Exploring Differences in the Development of Critical Thinking in Online Collaborative Learning from the Perspective of Learning Styles

李凤鸣，江鑫广，陈彦君，吴忭*
华东师范大学教育信息技术学系
bwu@deit.ecnu.edu.cn

【摘要】 信息技术环境下的批判性思维发展一直受到教育技术研究者的密切关注，但不同学习风格学习者的批判性思维发展缺乏过程性对比分析。本研究基于在线教学环境，运用所罗门量表测量学习风格后，应用 Murphy 的批判性思维框架编码在线论坛数据，并采用认知网络分析法探究论坛中不同学习风格学习者批判性思维发展差异。结果显示，在线协作学习支持批判性思维由低阶层次向高阶层次发展，具有不同信息加工风格的学习者的批判性思维发展差异显著。本研究将启示在线教学环境下批判性思维发展的动态过程研究，作为发掘学习特点和需求的研究参考。

【关键词】 学习风格；在线协作学习；批判性思维；认知网络分析

Abstract: *The development of critical thinking in the information technology environment has been closely scrutinized by educational technology researchers. However, there is a lack of process-oriented comparative analysis regarding the development of critical thinking among learners with different learning styles. This online study used the Solomon Scale to measure learning styles and applied Murphy's framework to code forum data, employing Epistemic Network Analysis to explore differences. Results show online collaborative learning supports critical thinking progression, with significant differences among learners with different information processing styles. This research will serve as a reference for exploring learning characteristics and needs, shedding light on the dynamic process of critical thinking development in online teaching environments.*

Keywords: Learning Style, Online Collaborative Learning, Critical Thinking, Epistemic Network Analysis

1. 引言

批判性思维作为 21 世纪核心素养之一，被视为新时代社会创新力、国家竞争力的核心及社会进步的标志。早在 1998 年联合国教科文组织发布的《面向二十一世纪高等教育宣言:观念与行动》，就重点强调了培养批判性思维的重要性，即“培养批评性和独立的态度”。我国在人才培养中，也一直高度重视学习者批判性思维培养。批判性思维作为认知领域的重要组成部分，其实质是基于事实性证据进行合理推断、质疑和辨析的思维过程。为了批判性地思考，学习者需要首先感知信息，然后处理信息，而学习风格是个人感知和处理信息的方式(Smith & Kolb, 1986)，其和批判性思维都与信息处理、思考和问题解决有关。

在线协作学习中，基于问题讨论的学习方式是学习者在网络环境下发展思维、培养能力的一种重要途径。相较于传统课堂的面对面讨论，在线讨论不受时空限制，具备一定的灵活性，也支持讨论内容回顾，促进学习者深入思考。基于在线协作学习环境，已有研究者探究学习者批判性思维能力的动态变化(De Leng, 2009)，但研究仍局限于批判性思维整体提升效果评价，关于批判性思维发展轨迹、学习风格对于批判性思维的影响等问题还缺乏深入探究。随

着学习分析技术的发展，认知网络分析方法在深度挖掘、动态评估思维能力内在发展规律方面发挥着重要作用。为此，本研究采用认知网络分析法对学习者批判性思维行为的动态发展规律进行深入分析，旨在深入了解批判性思维发展轨迹，并识别不同类型学习风格的学习者的批判性思维行为特点，一方面为面向培养批判性思维的教学活动的有效性评价提供支撑，另一方面也为教师的在线协作教学设计提供参考。

2. 文献综述

2.1. 批判性思维于在线协作学习中的发展

批判性思维(Critical Thinking, CT)对于学习者发展的重要性得到了广泛认可(Zhan, 2021)。从解决问题的角度来看，批判性思维可视为发现问题、寻求相关信息并做出解决问题决策的过程。Halpern(2013)将批判性思维定义为通过使用思维策略解决问题和实现目标的方法。尽管研究者从不同角度定义批判性思维，但都将其作为影响学习的认知策略(Pintrich et al., 1993)。

现有研究表明,通过设计在线教学活动和有效干预措施,可以追踪和探索基于问题式教学、同伴互评等不同在线协作学习情境下学习者批判性思维的发展。在线论坛被认为为批判性思维和问题解决能力的发展提供了空间,类似于面对面课堂的积极小组参与(Al-Fadhli & Khalfan, 2009)。与线下面对面讨论相比，在线讨论在某些方面可能更有助于学习者进行深入的批判性分析，因为他们有更多时间进行编辑和阅读他人的帖子,在线环境中的相对匿名性可以减少社交焦虑(Maurino, 2007)。在线课程结构、讨论指南和要求对确保学习者互动的质量、速度和深度非常重要，由此，研究者将内容和讨论的共享反思等同为在线课程中批判性话语的基础(Fidalgo & Thormann, 2012)。在线协作学习情境支持知识的共同构建，其中师生和生生之间的互动促进了批判性思维和问题解决能力的发展(Gold, 2001)。因此，可以总结为在线协作学习情境为批判性思维的发展提供了良好的空间，而在线话语讨论成为研究者关注的重点。

2.2. 学习风格与批判性思维之间存在相关性

学习风格是个人在学习过程中表现出来的习惯和偏好,包括学习方式、策略和环境等方面。根据经验性学习理论(Experiential Learning Theory),学习风格可划分为四类:发散型(Diverger)、同化型(Assimilator)、聚敛型(Converger)和顺应型(Accommodator)，每种学习风格类型解释了个体学习、利用经验和社会交互的方式(Kolb,2014)。根据学习方式，学习者可分为以下三类：偏向图像学习的视觉型，偏向听讲的听觉型和偏向动手实践的动觉型。

学习风格与批判性思维之间存有一定相关性。学习风格影响个体对信息的接受和理解方式，批判性思维过程是指对信息进行分析、评价和推理的过程，以支持个体全面客观地评估信息的价值和真实性。研究者指出学习风格可能是批判性思维发展的决定因素(Andreou et al., 2014)，不同学习风格的学习者在批判性思维方面存在差异，聚敛型学习者的批判性思维水平相较于发散型的批判性思维水平更高(An & Yoo, 2008)，动觉型学习者的批判性思维发展似乎好于视觉型和听觉型(Hananto & Kusmayadi, 2018)。Durnali(2022)研究发现自我调节有助于学习者明确学习风格，进而消除批判性思维发展障碍。然而，也有研究认为学习风格不会直接影响批判性思维的发展。比如, Purwanto 等研究者(2020)发现学习风格并不影响学习者的数学批判性思维能力，无论是听觉型、视觉型还是动觉型的学习者都有不同的数学批判性思维能力。尽管如此为促进批判性思维发展，学习风格是教学设计需考虑因素之一。

目前针对批判性思维与学习风格的关联研究主要集中于护理教育专业(Andreou et al., 2014; An & Yoo, 2008)，在社会科学教育方面的研究尚有缺失。尽管研究者关注到学习风格与批判性思维之间的关系，但对于思维发展轨迹，学习风格具体影响的研究较少。本研究将基于学习科学与技术这一课程背景，探讨学习风格与批判性思维的关系，拟解决的问题为：具有不

同类型学习风格的学习者于在线协作学习中的批判性思维整体特征是否存在差异？批判性思维发展轨迹是否存在差异？

3. 研究设计

3.1. 研究对象

本研究招募了上海某师范大学教育技术学专业《学习科学研究》课程中的 25 名研一学习者作为研究对象（男生 10 人，女生 15 人），学习者需围绕“概念转变”、“多媒体学习”、“具身学习”、“教学支架”、“计算机支持的协作学习”和“游戏化学习”六个学习科学研究子领域，于在线学习平台上分别开展为期一周的异步讨论。讨论过程中，学习者被鼓励提出问题、提供见解，并对他人的观点进行批判性思考和评价，促进问题解决和知识建构。本研究以两周为单位划分为三个阶段，分别为初期阶段、中期阶段和最终阶段，三阶段的主题讨论分别产生了 442 行，339 行和 233 行话语数据，共 1014 行数据。

3.2. 研究方法

本研究主要采用了调查研究法和认知网络分析法（Epistemic Network Analysis, ENA）两种方法(Shaffer，2017)。在调查研究方面，应用学习风格量表测评学习风格类型，根据不同维度的数据结果区分学习者类型。作为一种融合定量与定性的量化民族志方法，认知网络分析法支持建构批判性思维结构，从时间维度进行深入描摹(Shaffer，2017)。ENA 采用滑动窗口划分会话，网络中的连接体现了当前会话编码与相邻会话编码的共现关系，可用以描述个人（或团体）的认知特征在话语数据中相关关联的方式，实现认知结构可视化。认知网络中存有位于平均位置的网络质心，用以标识个体或群体在认知网络空间的位置。认知网络中的固定节点表示编码，边的厚度与两个编码之间连接频率成正比。在比较网络差异时，不仅可比较质心位置，也可减去连接权重来构建认知网络叠减图。具体而言，本研究首先收集在线学习平台论坛数据，然后两位研究者独立预编码，在取得良好的一致性结果后（各编码 Kappa 值都在 0.8 以上），两位研究者分工编码。编码完成后，本研究应用 ENA 网络工具（https://www.epistemicnetwork.org/）进行认知网络图绘制和非参数差异检验操作，我们将 Units 参数设置为学习者个体，Conversation 参数设置为单个讨论主题，节参数设置为 4。

3.3. 研究工具

3.3.1. 学习风格量表

在学习风格的评估方面，20 世纪末基于 Felder-Silverman 学习风格量表提出的所罗门学习风格量表在教育信息技术领域尤其是线上教育中开始被广泛使用，目标是支持教师因材施教（Zagulova et al., 2019）。量表从信息加工、感知、输入和理解这四个方面分为八种风格：活跃型与沉思型，感悟型与直觉型，视觉型和言语型，序列型和综合型。活跃型学习者善于通过讨论或讲解信息等方式内化知识；沉思型则喜欢独立思考问题，探索解决方式。感悟型学习者喜欢与现实有明显联系的课程，较依赖识记和已掌握的学习方法；直觉型则倾向挖掘知识间的相关性，对新概念的掌握更有优势。视觉型学习者善于通过看到的实物获取知识，如图片、板书、PPT 等内容；言语型擅长通过语言的表达和文字的描述捕捉信息。序列型学习者习惯线性的逻辑思维思考问题；而综合型重组碎片化信息，将知识进行联结。量表题目举例：“处理复杂问题时，你更喜欢：a)独自思考并分析问题，b)与同学或同事讨论问题的不同观点”，“在学习新知识时，你更倾向于：a)阅读相关的书籍或文献，b)听老师讲解或观看解说视频”。

3.3.2. 批判性思维编码表

本研究改编了 Murphy（2004）框架作为批判性思维编码表（见表 1），其针对于在线异步讨论情境，将批判性思维分为五个过程：辨识、理解、分析、评价、创造。

表 1 批判性思维编码表

过程	描述	案例
辨识	认识到或确定存在的问题、困境、难题等。	概念转变的过程和结果是否受到个体认知和情感状态的影响?
理解	澄清基本认识，引入新的观点，探索相关的知识、经验、研究、信息和观点。	正如文献中提到的，教师在对比电位差，来进行压力差的教学，鼓励或许是一种增强信心的方式，能够促进部分学习者在概念转变过程中克服学习困难。
分析	寻求深入的澄清，组织已知的信息，识别未知的信息，并将问题、困境或问题剖析为其基本组成部分。	概念保留、概念增加，概念取代、概念共存等概念转变类型之间可以存在联系，因为它们可能会相互影响和交织在一起，一个概念的取代可能导致新的概念增加，同时也可能导致旧概念的保留和新概念的共存。
评价	对信息、知识或观点的价值、科学性、有效性、适用性进行批评和判断。	在我看来，概念转变可能是实现知识迁移的先决条件之一，概念转变可能是时刻都在发生的、是理解性学习的必要条件。
创造	提出新的观点、方式、策略或问题解决方案，以支持教学实践或研究实践改善。	作为研究者，我们可以采用多种方法来研究不同学科领域中的概念转变。

4. 研究结果

4.1. 学习风格分类结果

根据所罗门学习风格量表的统计方法：共 44 题，每题有 a、b 两个选项，每四个量表中，分别记录选择 a 和 b 的个数，用较大的数值减去较小的数值，如在“活跃型/沉思型”中，共选择八个 a 和三个 b，学习风格为 5a，即为活跃型。问卷数据处理后，我们发现在信息加工维度，13 位学习者倾向沉思型，12 位学习者倾向活跃型。在信息感知维度，19 位学习者学习者喜欢通过感悟的方式来感知信息，为感悟型，直觉型学习者有 6 位。在信息输入维度，23 位学习者擅长通过视觉刺激输入信息，为视觉型，仅 2 位学习者是言语型。在信息理解维度，16 位学习者为序列型，9 位学习者倾向综合型。

4.2. 不同学习风格的批判性思维差异分析

因信息输入维度中视觉型与言语型两类学习者人数差距较大，所以本研究暂时剔除了该维度的差异分析，从信息加工、信息感知和信息理解三维度开展批判性思维差异分析。

4.2.1. 信息加工维度批判性思维差异分析

通过建模具有不同信息加工风格的学习者的讨论数据，分析发现活跃型学习者相较于沉思型表现出了更高的批判性思维水平。在图 1 中，分别给出了两类学习者的平均网络，以及认知网络叠减图，其中红色节点标识活跃型学习者的认知网络质心，蓝色节点标识沉思型学习者，正方形节点标识质心均值。认知网络中第一个维度（X）占数据总体方差的 30.7%，第二个维度（Y）占数据中方差的 26.1%，表明 ENA 生成的批判性思维模型具有较强的拟合优度。使用 Mann-Whitney U 检验进行统计学比较，第一个维度的差异存有统计学意义：活跃型 Median=0.19；沉思型 Median=-0.16；U=119.00，$p<0.001$，r=-0.83，而第二个维度上没有显著差异。通过观察发现辨识节点呈现在 ENA 空间的左侧，而创造节点呈现在网络空间的右侧，由此第一个维度可以被理解为辨识/创造维度。根据认知网络叠减图可发现，沉思型学习者的认知网络模型在投影空间左侧具有更强的联结，而活跃型学习者在投影空间右侧有更强的联结，我们由此推断两类学习者在辨识/创造维度上存有较为显著的批判性思维水平差异。又结合认知网络叠减图（见图 1-c），活跃型学习者的创造行为联结相较于沉思型更强，包括理解

-创造、分析-创造以及评价-创造三个联结。而沉思型学中的辨识和理解、辨识和分析、辨识和评价以及分析和评价四个低阶认知联结的强度在一定程度上高于活跃型学习者。综上所述，我们推断在线论坛中活跃型学习者相较于沉思型学习者表现出了更为高阶的批判性思维认知。

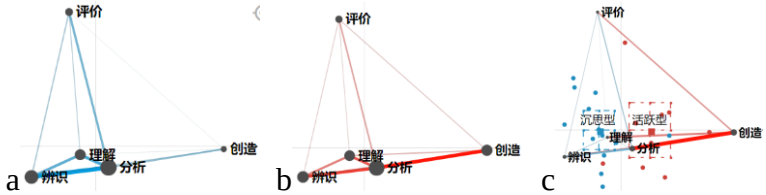


图 1 信息加工维度的认知网络图（图 1-a，沉思型平均认知网络；图 1-b，活跃型平均认知网络；图 1-c，沉思型与活跃型的认知网络叠减图）

4.2.2. 信息感知维度批判性思维差异分析

在信息感知维度区分感悟型和直觉型两种风格的学习者，对两类群体分别建模分析，发现直觉型学习者相较于感悟型学习者可能具有更高的批判性思维水平。在图 2 中，分别给出了两类学习者的平均网络模型以及叠减网络。在认知网络叠减图（见图 2-c）中，红色节点表示直觉型学习者的认知网络质心，蓝色节点则代表感悟型。经过 Mann-Whitney U 检验进行统计学比较，发现两个维度上并不存在显著差异。根据叠减图，直觉型学习者在评价-创造、评价-分析以及评价-辨识各联结上的强度相对更高，而感悟型学习者有在辨识-理解、辨识-分析以及分析-创造各联结上的强度相对更高，由此我们推测直觉型学习者相较于感悟型学习者更倾向于评判信息或观点的科学性、有效性和适用性。

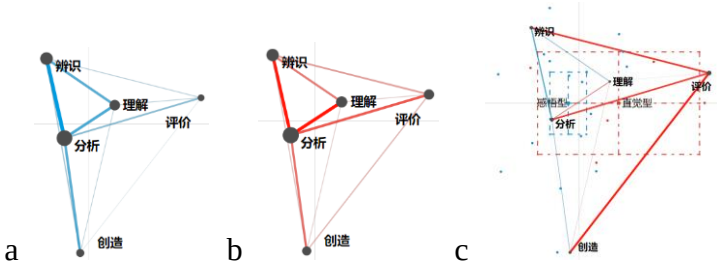


图 2 信息感知维度的认知网络图（图 2-a，感悟型平均认知网络；图 2-b，直觉型平均认知网络；图 2-c，感悟型与直觉型的认知网络叠减图）

4.2.3. 信息理解维度批判性思维差异分析

在信息理解维度区分序列型和综合型两类学习者，对两类群体分别建模分析，发现综合型学习者相较于序列型可能具有更高的批判性思维水平。在图 3 中，分别给出了两类学习者的平均网络以及认知网络叠减图，其中红色节点表示序列型学习者的认知网络质心，蓝色节点则代表综合型，正方形分别代表两类群体的质心均值。尽管经过 Mann-Whitney U 检验进行统计学比较，发现两个维度上并不存在显著差异，但是通过观察两类学习者的平均认知网络图叠减网络，可以发现两类学习者存在一定的认知差异。根据叠减图（见图 3-c），可发现序列型学习者分别在分析和评价、辨识和评价两类联结的联系相对较强。综合型学习者与序列型学习者相比在分析和辨识、评价和创造之间的联系相对更强。由于综合型学习者出现了评价和创造这一对高阶认知联结，所以我们推测综合型学习者相较于序列型学习者可能具有相对更高的批判性思维水平。

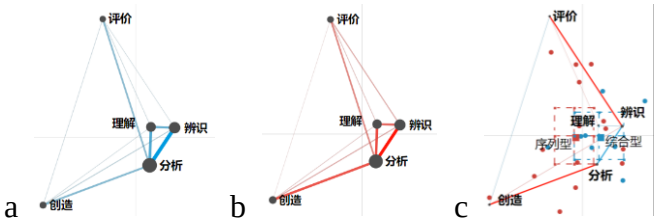


图 3 信息理解维度的认知网络图（图 3-a,综合型平均认知网络；图 3-b，序列型平均认知网络；图 3-c，综合型与序列型的认知网络叠减图）

4.3. 不同学习风格的批判性思维发展轨迹差异分析

对比分析具有不同学习风格群体的批判性思维发展情况后，我们同样发现具有不同信息加工风格的学习者呈现出了差异显著的思维发展轨迹，其他维度的发展差异并不显著。下面建模分析信息加工维度中活跃型和沉思型两类学习者的批判性思维发展过程。

活跃型学习者相较于沉思型学习者呈现出了由低阶认知向高阶认知发展的显著趋势。在活跃型学习者的认知网络（见图 4）中，第一个维度（X）占数据总体方差的 26.4%，第二个维度（Y）占数据中方差的 21.0%。应用 Mann-Whitney U 检验比较各阶段网络质心位置差异，虽然初期阶段和中期阶段之间的差异以及中期阶段和最后阶段的差异并不显著，但初期阶段和最后阶段之间的差异显著。初期阶段和最后阶段在第一维度上没有显著差异，而第二维度的差异存有统计学意义：初期阶段 Median=0.21；最后阶段 Median=-0.04；U=11.00，$p<0.001$，$r=0.78$。又结合活跃型各阶段的认知网络叠减图（见图 4），我们发现中期阶段相较于初期阶段网络存有更强的辨识-创造连接，最后阶段相较于中期阶段呈现出了更强的评价相关连接，最后阶段相较于初期阶段的网络连接更多出现第二维度的负半轴，由创造，评价和辨识行为组成。由此，我们推测活跃型学习者表现出了显著的批判性思维发展趋势。又分析沉思型学习者的批判性思维发展趋势，我们发现虽然该组学习者的各阶段差异并不显著，但叠减网络呈现出该类学习者的批判性思维水平存有一定发展。根据沉思型各阶段的认知网络叠减图（见图 5），我们发现中期阶段相较于初期阶段网络存有更强的分析-创造和分析-辨识连接，最后阶段相较于中期阶段呈现出了更强的评价相关连接，最后阶段相较于初期阶段的评价-辨识连接更强，由此，评价行为的相关连接表明该组学习者的批判性思维水平有一定提升。

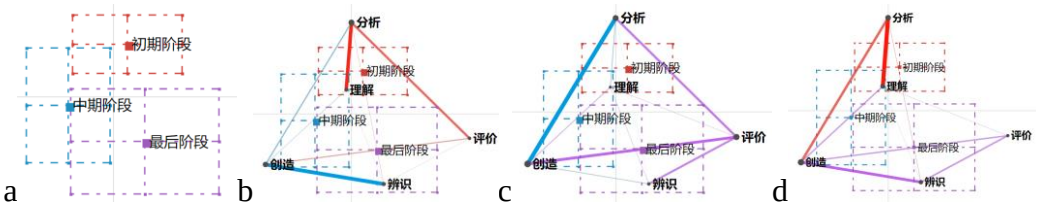


图 4 活跃型认知网络图（图 4-a，认知网络质心图；图 4-b，初期阶段与中期阶段的认知网络叠减图；图 4-c，中期/最后阶段网络叠减图；图 4-d，初期/最后阶段叠减图）

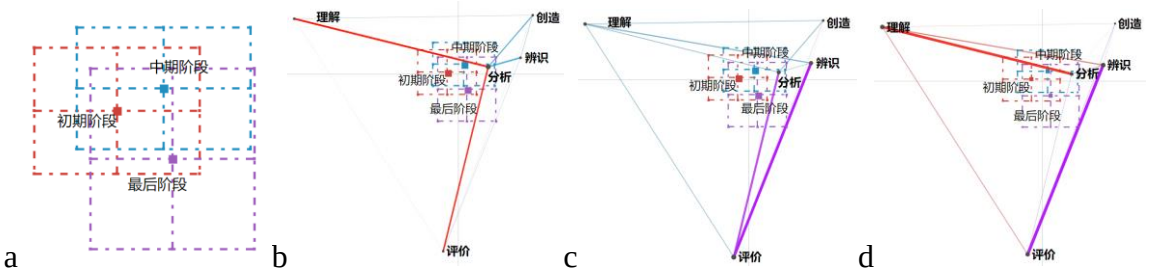


图 5 沉思型认知网络图（图 5-a，认知网络质心图；图 5-b，初期阶段与中期阶段的认知网络叠减图；图 5-c，中期/最后阶段网络叠减图；图 5-d，初期/最后阶段叠减图）

5. 讨论与总结

综合批判性思维整体特征和不同阶段特征的差异分析，我们发现不同学习风格的学习者具有不同的批判性思维特征，思维发展轨迹也存在有一定差异，尤其在信息加工维度呈现出了显著差异。结合不同学习风格学习者的批判性思维特征挖掘将有助于因材施教，给予思维发展的针对性反馈。

在信息加工维度，活跃型学习者喜欢在论坛中讨论解答问题，而沉思型学习者更愿意阅读帖子（姜强等，2012），表现在本论坛中则是活跃型更倾向于主动地分析信息知识以及创造解决方案,沉思型学习者则更加倾向于理解列举当前的信息以及分析和评价信息。本研究中，活跃型学习者呈现出了由低阶水平到高阶水平的显著发展趋势，批判性思维水平相对更高，可以更快地针对问题给出解决方案，而沉思型学习者虽没有明显的发展趋势，但其评价意识也在不断增强。该发现与现有文献的研究结果存在一致性。批判性思维于在线交互活动的影响下，会呈现出非跳跃式、连续统的发展过程，即批判性思维发展层次会出现由低阶水平逐渐向高阶水平发展的趋势（冷静等，2018；张涛等，2022）。交流互动活动可促进在线活动中学习者交互和批判性思维发展（柳晨晨等，2021），调动学习者的讨论积极性或许有助于改善批判性思维发展。在信息理解维度，相较于序列型学习者，综合型学习者在评价-创造之间的联系相对紧密，我们推测综合型学习者的批判性思维水平相对更高一些。这与 Hernandez-Torrano 等（2017）在大学遗传学和解剖学课程中发现的综合型学习者表现更好的结论相一致。结合相关研究，综合型学习者相较于序列型更善于抽象和概括信息，分析并系统化整合信息的能力更强（施翠芬等,2008；Umar et al., 2012），有助于学习领会概念材料，更能够从概念相关的讨论学习中获益。由此，教师可提供知识信息概括工具，帮助学习者从整体上把握知识概念和信息逻辑，以促进批判性思维发展。在信息感知维度，虽然两类学习者并未表现出显著差异，但是相比于感悟型学习者，直觉型学习者会更多地表达自己的观点或者评价其他人的观点。在科学研究过程之中，感悟前人的研究以及评价别人的研究同样都很重要，因此可以增加对于直觉型学习者的学习特征挖掘与关注，发掘出其更多的思维特点。

总之，本研究基于在线协作的学习情境，关注不同学习风格的学习者在主题任务中开展问答互动，围绕学习者之间的交互文本探索批判性思维的发展过程。文章的主要贡献在于将批判性思维置于在线协作的学习环境中，通过问答交互探索不同学习风格学习者的批判性思维在问答互动不同阶段的发展特征及差异。下一步，将进一步探索教学支架和教师干预对不同学习风格学习者批判性思维发展的影响，考虑将研究结果应用于教育实践中，构建针对不同学习风格学习者的批判性思维培养框架。

参考文献

Smith, D. M., & Kolb, D, A. (1986). *User's guide for the learning-style inventory: A manual for teachers and trainers.* Boston: Mcber & Co.

De Leng, B. A., Dolmans, D. H., Jöbsis, R., Muijtjens, A. M., & van der Vleuten, C. P. (2009). Exploration of an e-learning model to foster critical thinking on basic science concepts during work placements. *Computers & Education*, 53(1), 1-13.

Zhan, Y. (2021). What matters in design? Cultivating undergraduates' critical thinking through online peer assessment in a Confucian heritage context. *Assessment & Evaluation in higher education*, 46(4), 615-630.

Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. psychology press.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement*, 53(3), 801-813.

Al‐Fadhli, S., & Khalfan, A. (2009). Developing critical thinking in e‐learning environment: Kuwait University as a case study. *Assessment & Evaluation in Higher Education*, 34(5), 529-536.

Maurino, P. S. M. (2007). Looking for critical thinking in online threaded discussions. *Journal of*

Educational Technology Systems, 35(3), 241-260.

Fidalgo, P., & Thormann, J. (2012). A Social Network Analysis Comparison of an Experienced and a Novice Instructor in Online Teaching. *European Journal of Open, Distance and e-learning*.

Gold, S. (2001). A constructivist approach to online training for online teachers. *Journal of Asynchronous Learning Networks*, 5(1), 35-57.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

Andreou, C., Papastavrou, E., & Merkouris, A. (2014). Learning styles and critical thinking relationship in baccalaureate nursing education: a systematic review. *Nurse education today*, 34(3), 362-371.

An, G. J., & Yoo, M. S. (2008). Critical thinking and learning styles of nursing students at the baccalaureate nursing program in Korea. *Contemporary Nurse*, 29(1), 100-109.

Hananto, R. B., & Kusmayadi, T. A. (2018, May). Analysis of the critical thinking process of junior high school students in solving geometric problems by utilizing the vak learning styles model. In *Journal of Physics: Conference Series* (Vol. 1013, No. 1, p. 012132). IOP Publishing.

Durnali, M. (2022). 'Destroying barriers to critical thinking'to surge the effect of self-leadership skills on electronic learning styles. *Thinking Skills and Creativity*, 46, 101130.

Purwanto, W. R., & Waluya, S. B. (2020, March). Analysis of mathematical critical thinking ability in student learning style. In *Journal of Physics: Conference Series* (Vol. 1511, No. 1, p. 012057). IOP Publishing.

Shaffer, D. W. (2017). *Quantitative ethnography*. Lulu. com.

Zagulova, D., Boltunova, V., Katalnikova, S., Prokofyeva, N., & Synytsya, K. (2019). Personalized E-Learning: Relation Between Felder-Silverman Model and Academic Performance. *Appl. Comput. Syst.*, 24(1), 25-31.

Murphy, E. (2004). An instrument to support thinking critically about critical thinking in online asynchronous discussions. *Australasian Journal of Educational Technology*, 20(3).

姜强,赵蔚 王朋娇.(2012).基于网络学习行为模式挖掘的用户学习风格模型建构研究.电化教育研究(11),55-61.doi:10.13811/j.cnki.eer.2012.11.009.

冷静 郭日发.(2018).在线协作平台中批判性思维话语分析研究.电化教育研究(02),26-31.doi:10.13811/j.cnki.eer.2018.02.004.

张涛,张思,高倩倩 王佳慧.(2022).促进在线同伴互评中学习者批判性思维发展研究.电化教育研究(06),53-60.doi:10.13811/j.cnki.eer.2022.06.008.

柳晨晨,于澎,侯洁蕊 王佑镁.(2021).在线学习中互动讨论模式如何影响学习者的批判性思维.电化教育研究(03),48-54+61.doi:10.13811/j.cnki.eer.2021.03.007.

Hernández-Torrano, D., Ali, S., & Chan, C. K. (2017). First year medical students' learning style preferences and their correlation with performance in different subjects within the medical course. *BMC medical education*, 17, 1-7.

施翠芬 巫向前.(2008).批判性思维能力与学习风格的相关性研究.护理研究(31),2843-2845.

Umar, I. N., & Rathakrishnan, M. (2012). The effects of online teachers' social role and learning style on students' essay writing performance and critical thinking in a wiki environment. *Procedia-Social and Behavioral Sciences*, 46, 5730-5735.

設計閱讀成就目標系統輔助學生閱讀

Designing a Reading Achievement Goal System to Assist Students in Reading

葉彥呈[1*]，許喬珉[2]，陳德懷[3]
[123]中央大學網路學習科技研究所
[*]Charles@cl.ncu.edu.tw

【摘要】 閱讀是培育終身學習的一種方式，擁有閱讀興趣與習慣，對未來的學習會產生良好的影響。團隊長期推動身教式持續安靜閱讀模式，希望能增進學生閱讀興趣，然而也發現學生有出現閱讀偏好的情況。為了引導學生閱讀不同種類的書籍，本研究設計一「成就目標」系統，系統包括「閱讀目標」、「目標獎勵」機制和「學生閱讀歷程」等功能。學生透過系統能夠觀看自己的閱讀歷程，並獲得目標達成的成就獎章。教師端則提供了全班學生的閱讀狀況統計，讓教師更好地了解學生閱讀狀況。本研究也進行了初步的實驗，針對學生閱讀興趣、閱讀數量等進行分析。整體而言，閱讀目標系統在促進學生閱讀動機和拓展閱讀領域方面取得了一定的成效，但在興趣部分未達顯著。未來研究可進一步延長實驗時間以觀察長期效果，並透過系統改善以提升對學生閱讀的引導效果。

【關鍵字】身教式持續安靜閱讀；閱讀歷程；成就目標

Abstract: *Reading is a vital means for fostering lifelong learning, and cultivating reading interest and habits can significantly impact future learning outcomes. To augment students' reading engagement, research team has consistently advocated for a Modeled Sustained Silent Reading model. However, a proclivity among students to develop specific reading preferences has been identified. This study implementing an "achievement goal" system, which includes features such as "reading objectives," a "reward mechanism" tied to goal achievement, and a comprehensive "student reading journey." Students can actively monitor their reading progress and receive recognition in the form of achievement badges upon successful goal attainment. On the teacher's side, a detailed class-wide reading status summary is provided, enabling a nuanced understanding of individual students' reading patterns. Preliminary experiments were conducted to examine students' reading interests and consumption rates. The reading achievement goal system has shown effectiveness in enhancing students' reading motivation and expanding their literary horizons. However, the impact on cultivating specific reading interests did not reach statistical significance. Future research efforts could extend the experimental timeline to observe long-term effects and refine the system to better guide students' reading preferences.*

Key words: Modeled Sustained Silent Reading, Reading Portfolio, Achievement Goal

1. 前言

閱讀是學習語文的基礎，培養學生閱讀習慣，對學生未來的學習產生正面的影響。研究團隊為了培育學生閱讀興趣，幫助學生養成終生閱讀習慣，多年來在台灣推動「身教式持續安靜閱讀。透過教師固定一段時間在全班面前安靜閱讀，全班學生也一起安靜看書，讓老師作為閱讀模範，起到「身教」的作用。團隊在經由長時間的推動後，發現部分學生有偏好閱讀某中領域的情況產生，長期來看，學生可能會減少閱讀不同領域書籍的機會。因此，本研究為了讓學生跳脫閱讀舒適圈，進而讓學生閱讀的書籍種類加廣，內容難度加深，開發了一個「閱讀成就目標」系統，除了呈現學生閱讀表現，讓學生知曉自己的閱讀狀況，也導入成

就目標的機制，由系統給予一些成就目標，提高學生的閱讀動機，並進一步進行實證研究，期望以此提高國小學生的閱讀廣度與深度，培養學生的閱讀興趣與習慣。

2. 文獻探討

2.1 身教式持續安靜閱讀

閱讀式培養語文能力相當重要的一環 (Allington & McGill-Franzen, 2021)，研究團隊過去進行許多閱讀相關的研究與推廣，目的是提升學生的閱讀動機，並培育學生終身閱讀習慣。明日閱讀當中，有一個重要的閱讀活動，稱為身教式持續安靜閱讀(Modeled Sustained Silent Reading，簡稱 MSSR)，是基於 IDC 理論設計的活動，主要方式為教師以身作則，與學生一起閱讀(Modeled)；在學校的課堂上每天長時間持續進行閱讀(Sustained)；教室必須保持安靜，讓學生專注於閱讀(Silent)；學生閱讀自己選擇的書籍(Reading) (Chien, Chen, Ko, Ku, & Chan, 2011; McCracken& McCracken, 1978)。身教式持續安靜閱讀提供學生探索他們真正感興趣的內容，並享受一個引人入勝且愉快的閱讀體驗。學生在閱讀過程中可以自行決策，並對自己的閱讀負責，這種方式能激發學生的興趣，甚至鼓勵他們未來成為終身閱讀者(Chien et al, 2015)。

2.2 遊戲化應用於閱讀

在教育領域，遊戲化廣泛應用於多項領域，透過遊戲的設計元素如徽章、排行榜和等級，可以激勵學習者的行動和解決問題能力，教育遊戲化是一種通過將遊戲設計原則融入學習環境的方法，以鼓勵學習者的積極性和參與度(Dicheva, 2017)。遊戲化的目標之一是鼓勵特定行為的發展(Nurmi, 2020)，透過引入遊戲化元素，可以設計有趣且具挑戰性的學習環境，提高學習者對知識和技能的學習動機(Zichermann & Cunningham, 2011)。

遊戲化還可以培養學習者的問題解決能力(Huang, 2015)。遊戲中常見的挑戰和追求目標的設計，可以激發學習者主動尋找解決問題的策略和方法。即多樣化的遊戲化環境對學生更具激勵性，並且應考量不同學生的差異性，以符合他們的需求進行設計 (Kocadere, 2018)。

在「遊戲化八角化框架」一書中，其中的成就(Achievement)核心動力提到，PBL 是主要的遊戲元素。PBL 即使用點數(Points)、徽章(Badges)、排行榜(Leaderboards)，這 3 個遊戲化要素，被廣泛應用在商業、教育、遊戲、運動當中（周郁凱，2017）。透過引入相關的遊戲化要素，教學者可以吸引學生的注意力並提升他們的學習成果。然而，為了有效地運用遊戲化於教育中，教學者需要深入研究學習者的需求和動機，並適應性地設計遊戲化元素，以創造出適合不同學習場景和目標的教學環境(Sezgin,2020)。過去研究顯示，使用遊戲化的閱讀學習系統一段時間後，學習者的閱讀能力和閱讀自主性得到提升(Mak, 2019)。這項研究強調了遊戲化在閱讀中的成效，它可以吸引學生興趣並提高閱讀能力，同時培養閱讀自主性。

2.3 目標設定

認知和動機是影響人類最終行為產生的兩大重要因素(Locke & Latham, 1990)。在這其中，目標作為個人動機的關鍵之一，具有重要意義。目標具有引導、擴張、持續及促發工作策略等功能(Locke, Shaw, Saari, & Latham, 1981)。設定目標可以提高學習者的積極性，幫助他們加強自我調節(Locke & Latham, 2006)。明確的目標能夠引導學習者實現更好的學習表現，簡單且明確的目標設定效果更佳，適度困難的目標能夠提升效能。同時，結合困難和明確的目標設定方式以及階段性的目標執行能幫助學習者達到目標。故研究者在設計活動或系統時，須依據不同學習者的需求，思考合適的方式來引導學習者達成目標。

3.成就目標系統介紹

3.1 「閱讀目標」設計

本研究架構於閱讀行為上，為了協助學生進行廣泛閱讀與大量閱讀，本研究開發了一個「成就目標」系統，一方面能呈現學生的閱讀表現，收集和分析學生的閱讀歷程，給予學生機會了解他們已經閱讀或尚未閱讀的書籍，作為選擇下一本書籍的參考。另一方面，系統支援「MSSR」與後續活動的連接，促成學生能相互合作，讓學生能夠分享、討論和推薦書籍。

設定閱讀目標和建立閱讀清單有助於激發學生對閱讀的興趣，進而提升其閱讀動機，「閱讀目標」系統使用遊戲化的方式，幫助學生自主追蹤並調整他們所閱讀的書籍種類與鼓勵學生進行閱讀，系統將閱讀目標展現在學生面前，有助於學生接觸不同主題、風格和類型的書籍，擴展其閱讀範疇，促使其接觸更多元的知識。達成閱讀目標可以為學生帶來成就感，強化其對學習的信心，並建立積極的學習態度，透過確定的閱讀目標，學生可能養成終身學習的良好習慣，並持續追求新知識和資訊。



圖 1、閱讀目標書籍

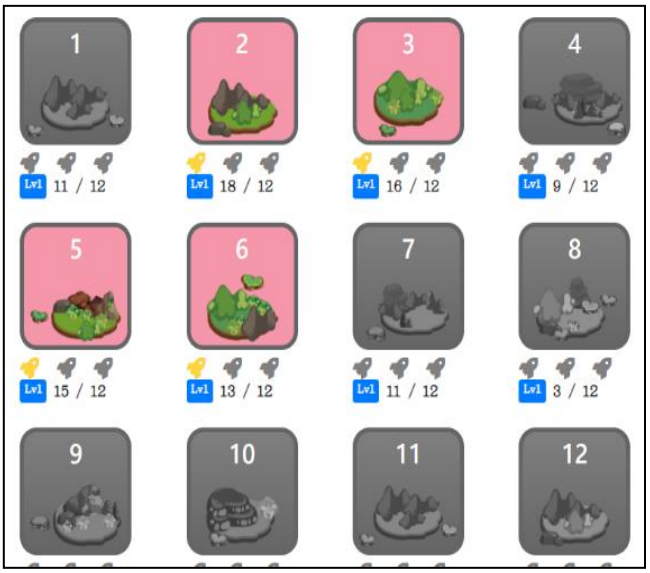


圖 2、學生成就紀錄

3.2 「目標獎勵」機制

由於學生在經過一段時間的閱讀後，常常會出現偏好某一類書籍的情形，為了讓學生廣博閱讀、大量閱讀以及提供舞台讓學生展現自己的閱讀表現，需要透過一些輔助機制去幫助學生廣泛閱讀不同領域的書籍，例如使用遊戲的方式融入一些遊戲機制，將閱讀目標與遊戲元素結合，或是將閱讀目標結合成就獎章的激勵方法，讓學生有收集的成就感，提高學生的閱讀動機，以增進學生閱讀不同種類書籍的興趣。因此，「目標獎勵」功能具有紀錄、獎勵、書籍分類等功能，學生可以選擇不同的閱讀目標，閱讀不同種類、難度、數量的書籍，來獲得不同的成就獎章(如圖 3)，增加學生的閱讀廣度與閱讀量。透過此系統，可以引導學生朝向收集不同領域與主題書籍的方向，累積閱讀的成就感，並且更加投入閱讀活動。



圖 3、「目標」獎勵機制

3.3 學生閱讀歷程

「學生閱讀歷程」（如圖 4）可以將學生過去的閱讀資料呈現出來，包含「閱讀關卡進度」、「閱讀書籍種類」、「閱讀書籍數量」等資訊。

「閱讀關卡進度」按照書籍的深度進行排序，讓學生能夠循序漸進的完成關卡，激勵學生主動去完成關卡，提高學生的閱讀更多書籍的動機。另外，破關的關卡會顯示出不同顏色，跟尚未完成的關卡作區別，利用視覺化的方式，提醒學生哪些關卡已經完成，而哪些屬於待完成的狀態，目的是幫助學生檢視與追蹤自己的閱讀關卡的進度。

「閱讀書籍種類」主要是讓學生看到自己的閱讀表現後進而調整自己閱讀書籍的廣度，嘗試閱讀不同種類的書籍。此功能可以讓學生看到自己總共閱讀了幾本書，更進一步顯示學生所閱讀的書籍分別為那些種類的書籍，系統將書籍以故事小說、科學技術與人文社會三個領域做分類。除了閱讀書籍種類，還有記錄學生閱讀的書籍深度，是以星星的圖像來表示，所有書籍分成 5 個深度，1 顆星代表這本書是較容易閱讀的繪本書，5 顆星代表高階文字書，越多星星表示該名學生閱讀的書籍越難。為了讓學生能快速了解自己閱讀狀況有哪部分可以進行調整，系統也會以文字顯示學生較少閱讀的書籍領域，提醒學生應該注意自己較少閱讀的書籍領域。

「閱讀書籍數量」則主要是讓學生能知道自己的「閱讀本數」，透過這個功能，學生可以知道個人化的閱讀紀錄，追蹤自己的閱讀進度和喜好。這不僅有助於提高學生對閱讀的自覺性，還能透過系統提供的閱讀分析，了解自己在不同主題、難易度的書籍上的閱讀習慣。

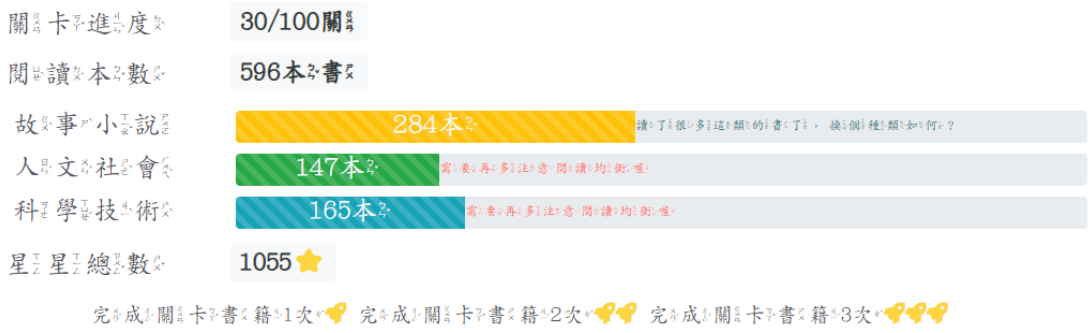


圖 4、學生個人閱讀歷程

此外，「學生閱讀歷程」還可提供學生個別化的閱讀建議。根據學生的閱讀歷程和偏好，系統可以推薦符合其興趣和閱讀水平的書籍，進一步擴展學生的閱讀視野。同時，系統還能促使學生參與閱讀挑戰或活動，例如明日說書人、聊書分享等，進一步激發他們的閱讀動機和參與度。最後，透過與「MSSR」的結合，「閱讀目標」系統可以成為一個互動平台，讓學

生之間能夠分享閱讀心得、互相推薦書籍，建立起一個積極的閱讀社群。這種互動有助於學生之間的合作和交流，進一步豐富他們的閱讀體驗，更培養了學生持續閱讀的習慣。

3.4 教師端設計

為了讓教師了解學生閱讀狀況，在教師介面會顯示全班的閱讀狀況。在系統中，可以看到全班學生的本週關卡、閱讀書本數、各種類書籍閱讀量等等，該功能每個禮拜會對學生的儀表板資訊進行統計，如圖 5 所示，系統會記錄學生每週的單元完成數、閱讀本數以及故事、科技、人社三大領域閱讀了多少本書，除了閱讀本數，也記錄學生閱讀的書本星星數。星星數量表示學生閱讀的書籍難度，書籍的字數、難度越深，星星數就越多，可以看出全班學生是否有進行深度書籍閱讀。

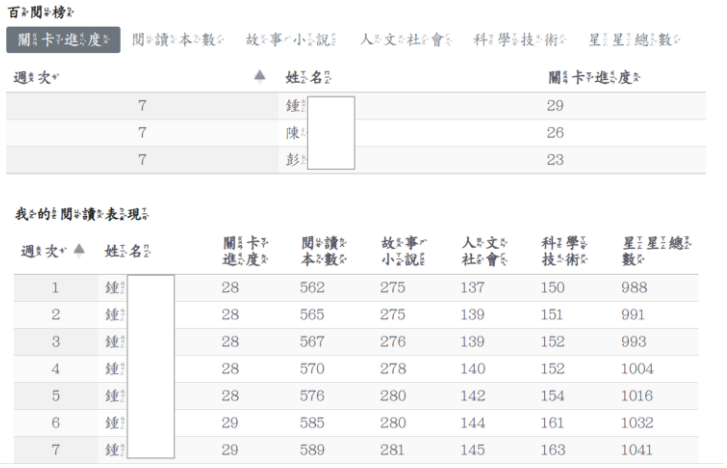


圖 5、教師端系統介面

4.研究方法

4.1 參與對象

本研究對象為台灣某市某小學二年級與三年級兩個班級，其中二年級 14 位、三年級 16 位，共 30 位學生。兩班學生參與本研究之前後測問卷，使用本研究之「成就目標」系統，觀察其對閱讀興趣與成效之影響。

4.2 研究工具

本研究主要使用閱讀興趣問卷、系統紀錄以及訪談作為研究工具。閱讀興趣問卷部分，目的是為了解學生在無使用系統與使用系統後的閱讀興趣狀況，問卷由研究者與研究團隊自編，分為 4 個主要面向，依序為「引趣與好奇心」、「入趣與心流」、「延伸興趣」、「個人閱讀興趣的發展階段」，閱讀興趣問卷整體信度 Cronbach's alpha 0.887。在系統紀錄部分，會蒐集使用者的行為紀錄，收集學生在系統上對各項功能的點擊次數、瀏覽次數、以及停留時間，幫助研究者了解學生在系統各功能的使用狀況。此外，由於希望能更深入地了解受訪者的想法，也礙於實驗時間的關係，因此只對其中 19 人有進行訪談。研究者記錄了研究對象的實驗活動資料和閱讀學習歷程，並於訪談時詢問學生對該功能的想法與使用上的建議，以便進行後續的評估分析與系統改善。

4.3 研究流程

如圖 6 所示，本研究評估流程總共有 4 個階段，進行 8 週的實驗。各階段詳細說明如下：

1. 前導階段：使用閱讀興趣問卷作為前測，以及向受試者說明評估的目的與系統的介紹，也讓受試者熟悉並操作系統。
2. 階段一：讓學生閱讀歷程視覺化，並導入個人儀表板的功能，向全部受試者介紹系統功能。
3. 階段二：對「學生閱讀歷程」、「閱讀目標」功能進行說明，瞭解受試者使用狀況與收集

使用者回饋建議做為系統修改方向。

4. 階段三：系統引導學生使用觀看自己的閱讀狀況，讓學生看見自己達成設定目標的進度。最後進行閱讀興趣問卷的後測，除了使用問卷外，也對學生進行訪談，並透過訪談收集學生對於閱讀廣度、閱讀數量與登台在他們閱讀的過程中有哪些方面的影響。

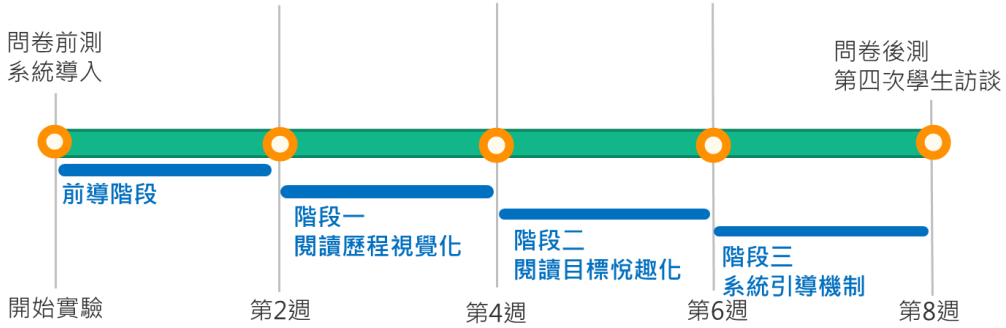


圖 6、研究流程

5.資料分析與結果

5.1 閱讀興趣

原先參與人數為 30 位，因研究時有研究者無全程參與，故將未進行前後測以及完成實驗流程的學生進行排除，實際分析到 20 位學生的資料。由於樣本數不足，因此本研究採用無母數統計法的魏克生符號等級檢定分析整體閱讀興趣之情況。結果如表一所示，前後測並無達到顯著性差異，z 值為-1.45，p = .334。

表 1 、整體閱讀興趣

		個數	等級平均數	等級總和	漸進顯著性（雙尾）
閱讀興趣後測總分－閱讀興趣前測總分	負等級	13	9.15	119.00	
	正等級	6	11.83	71.00	
	等值結	1			
	總和	20			.334

本問卷有事先利用反向題進行排除，造成此結果的原因可能與樣本數不足、學生使用系統時間較短，由於閱讀興趣需要較長時間培養，故未來需進行更長時間的研究，驗證對學生閱讀興趣的影響。

5.2 學生閱讀歷程與系統紀錄

為了觀察成就目標系統是否對學生的閱讀產生影響，因此從系統紀錄中，分析實驗期間學生閱讀狀況的變化，從表 2 之結果可知，透過給予學生目標與呈現閱讀歷程給學生後，學生在閱讀本數的部分明顯提升，可見系統導入後，會對學生的動機產生影響，但由於實驗時間較短，尚須較長時間的觀察。此外，表 3 顯示學生的不同領域閱讀的比例變化，由於學生會有自己的閱讀偏好，容易在閱讀到自己喜歡的主題時，偏食於某一項領域。培養學生的閱讀習慣為最重要的事情，但經由許多長期研究後發現，學生偏向於某一領域後，就不易往其他的領域拓展，可能造成錯過其他領域書籍或其他興趣的發現。因此，本研究之設計也是希望透過成就目標系統，能夠引導學生嘗試往其他領域的書籍閱讀，從結果來看，原先較多學生閱讀故事類書籍，在經過本研究之後，學生的科技領域閱讀占比逐漸上升，其他領域的比

例略微下降，可見目標系統的確有引導學生閱讀不同種類書籍之成效，未來希望能透過目標系統的輔助，讓學生達成目標與獲得成就，以鼓勵學生閱讀主題廣度更廣與深度等級更深的書籍。

表 2 、學生閱讀本數變化

	閱讀總本數	平均閱讀本數
階段一	283	9.4
階段二	382	12.7
階段三	401	13.3

表 3 、各領域閱讀比例變化

		階段一	階段二	階段三
故事	總本數	125	130	153
	平均本數	4.1	4.3	5.1
	三領域佔比	44%	34%	38%
科技	總本數	66	149	141
	平均本數	2.2	4.9	4.7
	三領域佔比	23%	39%	35%
人社	總本數	92	103	107
	平均本數	3.0	3.4	3.6
	三領域佔比	33%	27%	27%

5.3 學生閱讀歷程與系統紀錄

本研究共計訪談了 19 位學生，向學生收集更多深入的想法與回饋。本研究將收集之訪談內容根據不同概念進行編碼，分成 4 個概念面向，分別為閱讀數量、閱讀廣度、自我追求閱讀興趣與自我反思。結果顯示，最常出現的概念為閱讀廣度，表示大部分學生會關注於自己所看過的閱讀種類，受訪學生表示當他們閱讀完書籍進行登記時，會去看這本書是屬於甚麼種類的，在自己的閱讀系統上是否有增加此種類的書籍。19 位受訪者中有 10 位表示自己會願意嘗試閱讀、接觸平常較少閱讀的書籍，在閱讀關卡中如果看到自己感到有興趣的書，即便是自己不常看的書籍種類，也會願意去嘗試。在成就獎章部分，只有 3 位學生表示獎章有帶給他動力，可見大部分學生還是以本身閱讀領域為主要考量。

6.結論與討論

本研究希望能透過設計閱讀成就目標系統，輔助學生閱讀活動，讓學生有目標的進行閱讀，藉由達成目標與獲得成就，提高學生閱讀主題廣度更廣與深度等級更深的書籍。系統設計主要目的是為了提供學生個人的閱讀資訊，讓學生能根據自己的閱讀表現，調整自己閱讀的書籍種類，也利用系統引導的方式提供適當的閱讀建議。也設計成就目標與成就獎章，提供學生閱讀不同種類書籍時給予獎勵，如果該領域種類書籍讀到一定的數量，會依照程度給予不同的獎章。

在結果部分，經過本研究活動之後，學生各階段的整體閱讀本數有提升，學生的閱讀量呈現增加的趨勢。在閱讀廣度部分，學生閱讀科技領域的閱讀量增加，可見系統引導對於幫助學生嘗試閱讀不同種類書籍有幫助，未來的研究可以進一步探討推薦系統的使用、同儕影響等因素對閱讀行為的影響，以提供改善的方法。

透過目標設定、系統機制以及個人動機等因素，能夠增加學生的閱讀廣度與閱讀量，但在系統設計的成就獎章在激發閱讀表現方面有限，需與其他獎勵機制結合或透過同儕間的展示來增加其影響力。未來研究需增進教師端功能性，讓教師能更了解全班學生的閱讀狀況並提供協助，以促進學生閱讀成長。

致謝

本研究在臺灣國家科學及技術委員會(NSTC 112-2410-H-008-020-MY3) 與「國立中央大學學習科技研究中心」的資助下完成，僅此致謝。

參考文獻

周郁凱(2017)。**遊戲化實戰全書：遊戲化大師教你把工作、教學、健身、行銷、產品設計……變遊戲，愈好玩就愈有吸引力！**。商業週刊。

Allington, R. L., & McGill-Franzen, A. M. (2021). Reading volume and reading achievement: A review of recent research. *Reading Research Quarterly*, *56*, S231-S238.

Chien, T. C., Chen, Z. H., Ko, H. W., Ku, Y. M., & Chan, T. W. (2011). My-Bookstore: The design of a management game to promote classroom reading activity. *Proceedings of the 19th International Conference on Computers in Education* (pp. 465–472). Chiang Mai, Thailand: Asia – Pacific Society for Computers in Education.

Chien, T. C., Chen, Z. H., Ko, H. W., Ku, Y. M., & Chan, T. W. (2015). My-Bookstore: using information technology to support children's classroom reading and book recommendation. *Journal of Educational Computing Research*, *52*(4), 455-474.

Dicheva, D., Irwin, K., & Dichev, C. (2017). OneUp learning: A course gamification platform. In *Games and Learning Alliance: 6th International Conference, GALA 2017, Lisbon, Portugal, December 5–7, 2017, Proceedings 6* (pp. 148-158). Springer International Publishing.

Huang, B., & Hew, K. F. (2015, November). Do points, badges and leaderboard increase learning and activity: A quasi-experiment on the effects of gamification. In *Proceedings of the 23rd international conference on computers in education* (pp. 275-280). Hangzhou, China: Society for Computer in Education.

Kocadere, S. A., & Çağlar, Ş. (2018). Gamification from player type perspective: A case study. *Journal of Educational Technology & Society*, *21*(3), 12-22.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current directions in psychological science*, *15*(5), 265-268.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological bulletin*, *90*(1), 125.

Mak, M. T. F., Wang, M., & Chu, K. W. S. (2019). Effects of a gamified learning platform on elementary school students' flow experiences in leisure reading. *Proceedings of the Association for Information Science and Technology*, *56*(1), 454-458.

McCracken, R. A., & McCracken, M. J. (1978). Modeling is the key to sustained silent reading. *The Reading Teacher, 31*, 406–408.

Nurmi, J., Knittle, K., Ginchev, T., Khattak, F., Helf, C., Zwickl, P.& Haukkala, A. (2020). Engaging users in the behavior change process with digitalized motivational interviewing and gamification: development and feasibility testing of the precious app. *JMIR mHealth and uHealth*, *8*(1).

Sezgin, S., Bozkurt, A., YILMAZ, E. A., Van der Linden, N., & Learning, P. (2020). Gamification, education and theoretical approaches: Motivation, engagement and sustainability in learning processes.

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. " O'Reilly Media, Inc.".

基于 iFIAS 的课堂师生互动行为研究——以初中信息科技优质课为例

Research on Teacher-Student Interaction Behavior in the Classroom Based on iFIAS——A Case Study of a Middle School Information Technology Quality Lesson

庄凌[1]，徐玲[2]，梁颖怡[3]
[123]华南师范大学教育信息技术学院
1654743919@qq.com

【摘要】 本研究采用 iFIAS 对两节全国初中信息科技优质课堂视频进行编码分析，发现不同教学风格对课堂师生互动行为存在显著影响。结合视频观察和数据分析，建议教师平衡启发式、引导式和讲授式教学语言，合理安排封闭型和开放型问题，并通过现代技术手段增强师生互动。此外，教师应鼓励学生主动提问，培养批判性思维和创新意识，从而构建更加优质、互动、和谐的课堂环境。

【关键词】 iFIAS；师生互动行为；初中信息科技优质课

Abstract: *This study conducted a coding analysis of two national junior high school information technology quality classroom videos using iFIAS, and found that different teaching styles have a significant impact on teacher-student interaction in the classroom. Based on video observation and data analysis, it is suggested that teachers should balance heuristic, guiding, and didactic teaching language, reasonably arrange closed and open questions, and enhance teacher-student interaction through modern technological means. Additionally, teachers should encourage students to actively ask questions, cultivate critical thinking and innovation awareness, in order to build a more high-quality, interactive, and harmonious classroom environment.*

Keywords: *iFIAS; teacher-student interaction behaviors; middle school IT quality classes*

1. 引言

随着新课程标准的实施和《中国教育现代化 2035》的发布，课堂教学改革的重要性日益凸显。新课程标准的实施旨在培养学生的综合素质和创新能力，使其适应信息社会的发展需求。而《中国教育现代化 2035》则明确提出了推动信息技术与教育深度融合的目标，旨在促进教育现代化进程。在此背景下，课堂师生互动行为作为影响教学质量的关键因素，引起了广泛关注。

初中信息科技课程作为培养学生信息素养的重要途径，其教学质量对于学生的全面发展具有重要意义。在新课程改革的背景下，信息科技课程逐渐受到重视，但师生互动行为仍存在一些问题，如互动形式单一、互动质量不高等。因此，本研究选取初中信息科技优质课作为研究对象，通过深入剖析其师生互动行为，以期为改进初中信息科技课程的教学提供有益的启示。

2. 研究设计

2.1. 研究对象

本研究选取全国初中信息科技优质课堂视频《体验物联网》和《初识物联网》作为研究对象。通过比较分析，全面了解不同教学目标下师生互动的模式和效果，并探讨教学方法对师生互动的影响，有助于准确把握初中信息科技课堂中师生互动的现状和问题，并为改进教学提供指导。

2.2. 行为分析编码工具

为了深入剖析课堂中的师生互动行为，本研究采用改进的弗兰德斯互动分析系统（iFIAS）作为行为分析编码工具。相较于传统的弗兰德斯互动分析系统（FIAS），iFIAS 在描述课堂中的师生互动行为方面进行了扩展和改进。iFIAS 不仅关注教师和学生之间的言语互动，还纳入其他类型的互动行为，如沉寂和教学媒体的使用，以此帮助我们更准确地理解课堂氛围和学生的参与情况。此外，iFIAS 对传统的言语类别进行细分，以更精确地描述课堂中的言语互动。例如，教师言语被细分为教师讲解、教师指示等多个子类别，有助于更深入地了解课堂中师生互动的特点和模式。最后，iFIAS 还注重对课堂氛围和情感交流的评估，有助于深入了解课堂氛围的节奏和情感基调，从而更全面地理解课堂互动的质量。因此，本研究选取 iFIAS 作为行为分析编码工具，如表 1。

表 1 改进型弗兰德斯互动分析系统(iFIAS)

分类		编码	行为	内容
教师语言	间接语言	1	教师接受情感	以一种不具有威胁的方式接纳和理解学生情感
		2	教师表扬或鼓励	对学生的行为进行鼓励或夸奖
		3	教师采纳学生观点	对学生观点进行描述、重复,或利用学生言论解决问题
	直接语言	4	教师提问（教师提问开放性问题,教师提问封闭性问题）	答案具有不固定性、发散性,允许学生自由发挥:具有固定答案的问题,比如概念、定理以及数字计算
		5	教师讲授	表达教师自己观点,对教学内容进行解释的陈述性语言
		6	教师指示	命令或指示学生做某事
		7	教师批评或维护权威	改变学生行为状态，用语言使学生行为由不可接受变为可接受
学生语言		8	学生被动应答	教师指定学生回答或学生齐答
		9	学生主动说话（学生主动应答，学生主动提问）	举手作答和主动表达个人观点；主动提出自己的疑问
		10	学生与同伴讨论	交流自己观点和看法
沉寂		11	无助于教学的混乱	与教学无关的混乱
		12	有助于教学的沉寂	与教学有关的短暂安静,如思考问题或练习习题
技术		13	教师操纵技术	教师对多媒体进行操纵
		14	学生操纵技术	学生对多媒体进行操纵

2.3. 研究方法与数据采集

在本研究中，将采用混合方法研究设计，结合定性和定量分析方法，以探究初中信息科技优质课的课堂师生互动行为。

2.3.1. 研究方法

2.3.1.1. 定性分析

利用 iFIAS 系统对课堂中的师生互动行为进行详细记录和编码，分析和描述互动行为的具体特点和规律。

2.3.1.2. 定量分析

基于 iFIAS 系统提供的编码数据，运用统计学方法对每种互动行为的频率、持续时间等数据进行分析和比较。

2.3.2. 数据记录方法

为了全面捕捉课堂中的师生互动行为，每隔 3 秒记录课堂中的师生互动行为，并遵循表 1 的 iFIAS 编码规则对课例 1 和课例 2 进行准确的分类和编码，由于这两个课例实际编码时长分别为 45 分钟和 44 分钟，因此生成了 881 和 861 个编码。

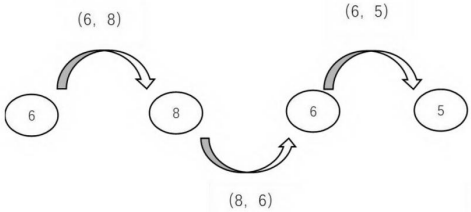


图 1 序对连结图

以课例 2 为例，将编码按次序连结成序对：(6，8)(8，6)(6，5)……。(见图 1)因此，课例 1 和课例 2 分别生成 881 和 861 个序对。将序对按照先行后列的次序填入矩阵表中并记录频次。iFIAS 分析程序自动生成矩阵分析表，iFIAS 编码程序处理数据后，得出各编码对应的频数和比例，如表 2 所示。

表 2 两节优质课编码频数和比例统计

编码	行为		课例 1		课例 1	
			频数	比例	频数	比例
1	教师接受情感		0	0%	0	0%
2	教师表扬或鼓励		34	3. 86%	3	0. 35%
3	教师采纳学生观点		33	3. 75%	54	6. 27%
4	教师提问	教师提问开放性问题	41	4. 66%	62	7. 2%
		教师提问封闭性问题	7	0. 79%	26	3. 02%
5	教师讲授		83	9. 42%	342	39. 72%
6	教师指示		46	5. 22%	91	10. 57%
7	教师批评或维护权威		0	0%	0	0%
8	学生被动应答		19	2. 16%	47	5. 46%
9	学生主动说话	学生主动应答	179	20. 32%	47	5. 81%
		学生主动提问	0	0%	3	0. 26%
10	学生与同伴讨论		375	42. 57%	37	4. 04%
11	无助于教学的混乱		0	0%	0	0%
12	有助于教学的沉寂		21	2. 38%	12	1. 39%
13	教师操纵技术		34	3. 86%	63	7. 32%
14	学生操纵技术		9	1. 02%	74	8. 59%
	总数		881	100%	861	100%

2.4. 维度及算法

iFLAS 系统根据前期所统计的数据进行 4 个维度(课堂互动结构特征、教学风格特征、师生问答特征、动态曲线特征)的具体分析，详见表 3。

表 3 iFIAS 分析维度

维度	子维度说明
课堂互动结构特征	教师语言比例,学生语言比例,教师与学生语言比例,同伴讨论比例,有益于教学沉寂比例,技术使用比例
教学风格特征	救师间接语言与直接语言之比,教师积极影响与消极影响之比教师提问占教师语言的比例,教师提问开放型问题占教师提问的比例,教师提问封闭型问题占教
师生问答特征	师提问的比例,学生主动应答占学生主动说话的比例,学生主动提问占学生主动说话的比例,学生主动应答与被动应答的比例
动态曲线特征	师生互动折线交点个数

3. 结果分析

编码之后，利用 iFIAS 分析程序分别对课堂互动结构特征、教学风格特征、师生问答特征、动态曲线特征等进行分析，得出以下结果。

3.1. 课堂互动结构特征分析

课堂中的师生互动主要依赖于语言交流。在 iFIAS 系统中，语言互动被细分为教师语言、学生语言、有益于教学的沉寂和技术使用四个类别。教师与学生语言比例是衡量课堂主导权的重要指标，它能够清晰地揭示出课堂是以教师为中心还是以学生为中心。同时，技术使用比例则能直观地反映出信息技术与课堂的融合程度，进一步揭示了教学方法的现代化程度。以下是关于课堂语言互动结构的统计数据，详见表 4。

表 4 课堂互动结构数据统计

统计项	课例 1	课例 2	常模
教师语言比例	27.7%	67.13%	68.0%
学生语言比例	65.04%	15.56%	20.0%
教师与学生语言比例	0.43%	4.31%	3.4%
同伴讨论比例	42.57%	4.3%	
有益于教学沉寂比例	2.38%	1.39%	
技术使用比例	4.88%	15.91%	

从表 4 可以看出，两个课例中教师的语言比例均低于贝莱克常模规定的 68%。这说明教师的讲解和指导相对较少，学生自主活动的时间和机会更多。这与当前提倡的以学生为中心的教学理念相符合。

课例 1 中，学生语言比例高达 65.04%，远超常模，表明学生的参与度非常高。结合视频，该课例的教师采用小组讨论等策略,鼓励学生积极参与。而课例 2 中,学生语言比例仅 15.56%，表明教师讲解时间过长，学生参与度相对较低。

常模规定教师与学生语言比例为 3.4%。课例 1 中，这一比例仅为 0.43%，表明教师话语量非常少，该课例更注重学生的互动和参与。而课例 2 中教师与学生的语言比例为 4.31%，

相对高于常模，说明教师语言占据课堂主导地位。

课例 1 中同伴讨论的比例远高于课例 2，表明课例 1 更注重学生间的互动合作，而课例 2 的教学方式偏向于教师单向传授。

在课例 1 中，有益于教学的沉寂比例为 2.38%，表明课堂中有一定时间是用于学生吸收和消化，有助于培养学生的独立思考能力和批判性思维能力。而在课例 2 中，有益于教学的沉寂比例为 1.39%，相对较低。这表明在教学节奏较快或教师对于学生的回应较为急躁，导致学生没有足够的时间进行思考和消化。

技术使用比例反映教学过程中使用技术手段的程度和频率。课例 1 中的技术使用比例为 4.88%，相对较低。而在课例 2 中，技术使用比例为 15.91%，相对较高。这表明课例 2 中的教师更多地使用技术手段辅助教学，提供更加生动、形象的学习体验，进一步激发了学生的学习兴趣和积极性。

3.2. 教学风格特征分析

通过表 5 发现，两节优质课在教学风格上存在一定的差异。在课例 1 中，教师间接语言与直接语言的比例为 0.89%，表明教师在教学中使用启发式、引导式的语言，这有助于引导学生思考。而在课例 2 中，该比例仅为 0.33%，表明教师在教学中更倾向于直接给予答案或明确指示，这种风格可能会限制学生的思考和探索。此外，在课例 1 中，教师积极影响与消极影响的比例高于课例 2，表明课例 2 中的教师需要控制消极影响，以更好地促进学生学习和发展。

表 5 教学风格特征数据统计

统计项	课例 1	课例 2	常模
教师间接语言与直接语言之比	0. 89%	0. 33%	
教师积极影响与消极影响之比	1. 46%	0. 63%	0. 42

3.3. 师生问答特征分析

表 6 师生问答特征数据分析

统计项	课例 1	课例 2
教师提问占教师语言的比例	19. 67%	15. 22%
教师提问开放型问题占教师提问比例	85. 42%	70. 45%
教师提问封闭型问题占教师提问比例	14. 58%	29. 55%
开放型问题与封闭型问题的比值	585. 71%	238. 46%
学生主动应答占学生主动说话比例	100%	94%
学生主动提问占学生主动说话比例	0%	6%
学生主动应答与被动应答的比例	942. 10%	100%

3.3.1. 教师提问的策略性

教师提问在课堂中的占比反映教师的角色。在课例 1 中，这一比例高达 19.67%，说明教师有意识地通过提问来推动课堂进程和激发学生的思考。进一步观察教师提问的类型，我们发现两个课例中的开放型问题都占主导地位，说明两位教师都倾向于提出需要深度思考和解析的问题。

3.3.2. 开放与封闭问题的平衡

开放型问题与封闭型问题的比值显示了教师对这两种问题的选择倾向。较高的比值（如课例 1 的 585.71%和课例 2 的 238.46%）说明教师更倾向于培养学生的批判性思维和创新意识，而非仅仅寻求标准答案。这有助于激发学生的创新思维和促进课堂讨论的深入。然而，过高的开放型问题比例可能带来知识传授的不足，因此平衡是关键。

3.3.3. 学生的主动参与

学生主动应答的比例反映学生在课堂互动的积极程度。在两节优质课中，学生对教师的提问都表现出了较高的主动参与意愿。这种高比例的学生主动应答有助于促进课堂氛围的活跃和学生思考的积极性。但是，课例 1 中没有学生主动提问，学生更倾向于回答教师提出的问题。在课例 2 中，尽管主动提问的学生比例仍然较低，但至少有一部分学生愿意主动参与和思考，提出自己的问题。

表 7 教师提问倾向数据

模式	教师提问									
	训练型提问				总计	创新型提问				总计
序对频次	(4, 4)	(4，8)	(8，4)	(8, 8)		(3，3)	(3，9)	(9，3)	(9，9)	
课例 1	22	2	0	16	40	13	2	8	152	175
课例 2	27	36	9	7	79	27	2	12	24	65

训练型提问主要侧重于教师引导学生回答，以促进学生知识巩固和能力提升。通过分析序对数据，我们发现训练型提问在两节课中都占据主导地位。其中，最具代表性的序对为(4，4)、(4，8)、(8，4)和(8，8)。序对(4，4)通常表示教师指定学生个体回答问题。在课例 1 和课例 2 中，该序对的频次分别为 22 和 27，说明教师频繁地指定学生个体回答问题。序对(4，8)和(8，4)的频次表明教师和学生之间的问答互动不仅局限于一对一的交流模式。在课例 2 中，(4，8)和(8，4)的频次相对较高，表明教师在课堂互动中更注重引导学生集体回答或讨论。序对(8，8)表示教师自问自答，旨在为学生提供答案或解释。在两节课例中，该序对的频次分别为 16 和 7，说明教师在某些情况下选择直接给出答案或解释。从上述分析可见，训练型提问在两节课中都占据主导地位，但具体互动模式存在一定差异。课例 1 更倾向于一对一的互动模式，而课例 2 则更加注重集体回答或讨论。

创新型提问鼓励学生主动表达观点或提出问题，对于培养学生的批判性思维和问题解决能力具有重要意义。通过分析序对数据，发现创新型提问在两节课中的运用也存在差异。序对(3，3)表示学生主动回答问题。课例 2 中该序对的频次较高，表明学生在该课例中更加积极主动。序对(3，9)和(9，3)表明学生在回答问题时提供了新的信息或观点。在课例 1 和课例 2 中，这两个序对的频次分别为 2 和 2、8 和 12，说明学生在回答问题时不仅仅局限于书本知识，而是能够结合自身经验或课外知识进行回答。序对(9，9)表示学生提出新的问题或话题。

在两节课中，该序对的频次分别为 152 和 24，说明学生在两节课中都展现出了较高的提问意愿和批判性思维。但课例 1 中学生提出的问题或话题更多。从上述分析可见，创新型提问在两节课中都有涉及，但具体的运用方式和频次存在差异。课例 1 中学生的提问意愿相对更高，而课例 2 中学生在回答问题时更加积极主动并能够提供新的观点或信息。

3.4. 动态曲线特征分析

以下是课例 1 和课例 2 的动态折线统计图，如图 2 和图 3 所示。

图 2 课例 1 师生语言互动折线统计图

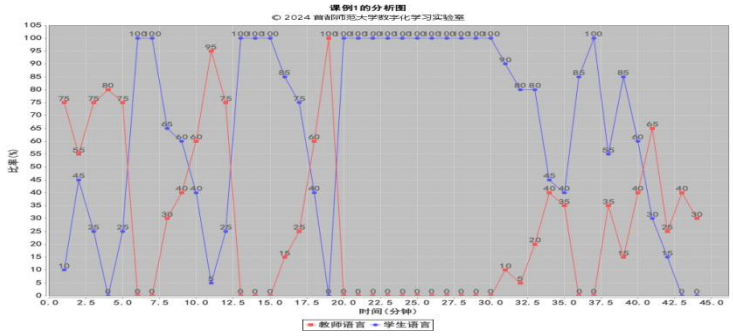
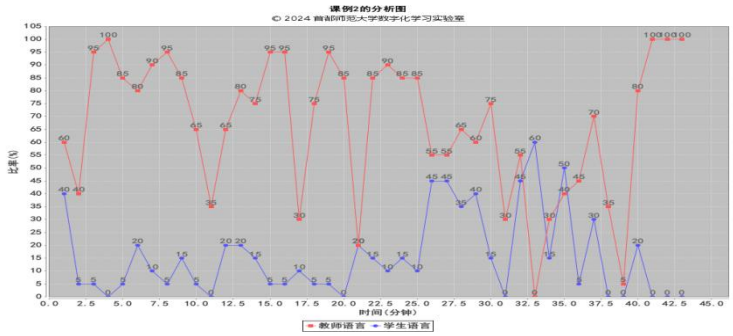


图 3 课例 2 师生语言互动折线统计图



根据图 2 和图 3 的折线统计图分析，课例 1 的教师互动效果更佳，具有 6 个交叉点且数值接近 50%。视频观察显示，课例 1 教师积极引导学生互动，提出启发性问题，鼓励学生思考和表达观点。而课例 2 教师更注重整体性与普及性，采用“提问-齐答”模式，确保多数学生参与，提高整体教学效果。

4. 建议

4.1. 优化与平衡教学风格

教师应根据具体的教学内容和学生的实际情况，灵活运用启发式、引导式和讲授式教学语言，平衡间接与直接语言的比例，以更好地引导学生的学习过程。同时，教师还可以进一步探索个性化的教学风格，以满足不同学生的需求。通过持续的教学实践和反思，找到最适合自己和学生的教学风格，有助于促进更有效的师生互动，构建优质课堂的基础。

4.2. 深化与实践师生问答策略

教师在设计问题时应合理安排封闭型问题和开放型问题的比例。对于封闭型问题，教师可以考虑如何引导学生进行更深入的思考和解释。对于开放型问题，教师可以更多地关注学生的观点和创意，并给予及时的反馈和引导。既关注知识传递，又注重思维能力培养。同时，应注重引导学生集体回答或讨论，以促进知识的共享与交流。另外，鼓励学生主动提问，培养他们的批判性思维和创新意识也是至关重要的。这有助于进一步提升师生问答的质量和效果，营造积极互动的课堂氛围。

4.3. 平衡与发展训练型与创新型提问

教师在运用训练型提问时应注意问题的针对性和层次性，以满足学生的学习需求。同时，根据实际情况逐步增加创新型提问的比重，以激发学生的创新思维和批判性思考。另外，教师还应根据学生的回答情况及时调整提问策略，引导学生提供新的观点或信息，培养他们的自主学习和思考能力。通过平衡训练型与创新型提问，教师可以更好地激发学生的思维活力和创造力，推动学生全面发展。

4.4. 营造与动态管理课堂氛围

教师可以采用多种措施来营造积极的课堂氛围。首先，通过积极与学生互动，关注学生的参与度和反馈，教师可以及时调整教学策略和活动。其次，利用现代技术手段，如在线学习平台和多媒体课件等增强师生间的互动效果。此外，教师还可以通过组织小组讨论、角色扮演等形式激发学生的学习兴趣和动力。通过不断地调整和改进教学策略和活动，教师可以构建更加优质的课堂环境，为学生的学习成长提供有力的支持。

5. 结语

本研究聚焦两节初中信息科技优质课，利用 iFIAS 编码工具对师生的互动行为进行研究，把握初中信息技术课堂中的师生互动现状和问题，并总结四点建议促进师生互动，包括优化与平衡教学风格，深化与实践师生问答策略，平衡与发展训练型与创新型提问，及营造与动态管理课堂氛围。然而，教育需与时俱进。未来研究应进一步拓展领域，挖掘互动行为的内在机制，提供更全面的指导。

参考文献：

陆灵明和甘龙华（2023）。基于 iFIAS 的初中生物优质课课堂互动行为研究。**中国教育信息化，29(03)**，91-98。

朱涵、汪甜甜和邓猛等（2022）。培智学校数学课堂教学师生言语行为互动研究——基于改进型弗兰德斯互动分析系统(iFIAS)。**中国特殊教育，(01)**，39-46。

方海光、洪心和孔新梅等（2022）。基于课堂交互行为数据的教学教研融合研究——以网络画板和 iFIAS 的数学教学教研应用为例。**中国电化教育，(05)**，115-121。

陈双敏和张燕（2021）。基于 iFIAS 的高中信息技术教师 TPACK 结构特征分析——以“一师一优课”10 节部级优课为例。**中国教育信息化，(08)**，92-96。

管漪（2020）。基于 iFIAS 的大学英语智慧课堂教学互动行为研究。**中国教育信息化，(11)**，23-29。

探討雙欄式步驟視覺化教學設計模式對不同先備知識學生之影響－以 AMA「幾何裝飾性風格」圖像創作為例

Exploring the Impact of Dual-Column Step-by-Step Visualization Instructional Design Model on Students with Different Prior Knowledge – A Case Study of AMA Geometric-Decorative Style Image Creation

李佳蓉 [1*]
[1]國立雲林科技大學 通識教育中心
[*] leecr@yuntech.edu.tw

【摘要】無所不在學習的科技時代，讓學習者能善用資訊科技，以基本設計元素點、線、面繪製出幾何藝術設計作品，培養設計力，亦是個重要的教育課題。準此，本研究將探究不同色彩先備知識之學習者於創作 AMA「幾何裝飾性風格」圖像之自我效能、學習投入、認知負荷及學習成效表現情形。研究結果顯示，高色彩先備知識學習者在雙欄式步驟視覺化教學設計模式下，較能保持較好的學習投入度，且有較低的認知負荷。此外，在本研究也發現，透過雙欄式步驟視覺化教學後，兩組在期末簡報美學專題表現上無顯著差異。

【關鍵詞】 雙欄式步驟視覺化；AMA 幾何裝飾性風格；自我效能；學習投入；認知負荷

Abstract: *In the ubiquitous learning era of technology, enabling learners to effectively utilize information technology to create geometric art designs using basic design elements of point, line, and plane, thereby fostering design skills, is also an important educational topic. Accordingly, this study explores the self-efficacy, learning engagement, cognitive load, and learning effectiveness of learners with different levels of prior color knowledge in creating AMA's "Geometric-Decorative Style" image. The results indicate that learners with high prior color knowledge, under dual-column step-by-step visualization instructional model, are better able to maintain higher levels of learning engagement and lower cognitive load. However, the study also found that there was no significant difference in the final presentation of aesthetic projects between the two groups after undergoing dual-column step-by-step visualization teaching.*

Keywords: dual-column step-by-step visualization, AMA Geometric-Decorative Style, self-efficacy, learning engagement, cognitive load

1. 前言

簡報軟體已經被越來越多的現場教師使用，希望藉由簡報軟體的協助，進行教學設計，將有效的教材呈現給學生以促進學生的學習（Kundart, 2014）。學者也針對簡報設計的問題因而造成學生的學習障礙提出想法，如重要資訊元素間的連結被隱藏了、投影片上放置過多的文字、大部份的投影片沒有提供可以促進最佳理解與學習保留的圖片 (Garner, Alley, Gaudelli & Zappe, 2009)。因此，有效進行簡報課程內容結構安排是重要的。此外，在「美感素養」 (aesthetic literacy)亦為當代文明發達國家的核心公民能力之一前提下，使用有效、簡

易上手的數位資訊軟體，協助學生培養美感能力，亦是當今教育的重點。準次，本研究將採AMA(Activate Mind Attention)軟體為平台媒介，以 PowerPoint 為平台，外掛增益集，以「雙欄式步驟視覺化」教學設計模式，讓學生在減少認知負荷的學習情境下，進行 AMA「幾何裝飾性風格」圖像創作之學習成效探究。

2. 文獻探討

2.1. 自我效能

Bandura (1977)首次提出自我效能 (self-efficacy)一詞。他認為自我效能就是指一個人判斷及執行一連串行動的能力，強調並非個人所擁有的技術能力，而是一個人運用技能達成任務的能力判斷，且人對於信念的強度，決定他們是否努力應付困難的情境。

2.2. 學習投入

Kuh(2003)將學習投入一詞界定為學生在 學習過程中的行為、感覺與思考的歷程，係指學生所花費的時間和投入的精力在具教育性目標的活動上，而教育活動則需要藉由他人互動來展現其意義，其中學生努力是重要的指標。本研究將學習投入分為行為投入、情感投 入與心智投入、主動投入共四個構面。

2.3. 認知負荷

認知負荷是影響學習的一大關鍵要素，Sweller、Ayres 與 Kalyuga (2011)也已提及認知負荷有三種：內在認知負荷 (intrinsic cognitive load)、外在認知負荷(extranous cognitive load)與有效認知負荷(germane cognitive load)。身為教師須在教學過程中適時了解學生的認知負荷狀況，適切地做課程教材審慎選取、 組織、整合，以利學生能降低外在認知負荷，做有效學習(learning)。

2.4. AMA 幾何裝飾性風格圖像意涵與範例介紹

學者陳俊宏、楊東民(1998)也提及插畫的觀點，加上數位化插畫之工具的特質及功能，依插畫的表現風格可分為三種類型：寫實性(realistic)：與傳統中的寫實稍有不同，指「合理的把握個體對象之客觀性」，包含超現實、象徵的、幻想的，也就是說畫面上所出現的對象都是個體性的寫實。裝飾性(decorative)：與意義內容等傳達機能比較起來，較重視裝飾性的插畫；此類插畫大部分部是平面性的空間。例如，整幅圖或其背景、周圍以平面性的要素來構成。樣式化(stylistic)：指單純化或變形等具有個性化特徵的表現風格，漫畫亦可屬於此範疇。較貼近寫實的表現風格，但有時呈現簡化且具有各種不同個性的表現特色。

本研究將依據前述學者對數位插畫與表現風格形式之論點，結合 AMA 軟體內特有之結構式複製繪圖法技法，將其所繪製出的數位 AMA 圖像，綜整後進行不同風格之分類，共可分類為三種類型，即為複合寫實性(compound realistic)、幾何裝飾性(geometric decorative)、精細樣式化(dedicated stylistic)。本研究將透過 AMA 結合 PPT 軟體，教師教授 AMA 結構式複製繪圖法之技法，採雙欄式步驟視覺化教學模式，教導學生繪製出具幾何造型元素之「幾何裝飾性風格」圖像作為首要目的，進行教學實驗研究。



圖 1　AMA 幾何裝飾性風格圖像範例

3. 研究設計

3.1. 實驗對象

本研究研究對象為臺灣中部某科技大學修習簡報美學課程之 42 位學生，且由同一位教師授課。本研究之無效問卷經檢查後，判定無效樣本為 3 位，剩下 39 位。其中高先備知識的學習者有 18 人，低先備知識的學習者有 21 人。本研究先備知識的判定是先排除無效樣本後，取學習者於色彩學知能前測 50 題選擇題之平均數，高於平均數者為高先備知識組，低於平均數者為低先備知識組。

3.2. 研究工具

使用之前後測工具為：色彩學知能試題、自我效能問卷(Pintrich, Smith, Garcia, & McKeachie, 1991)、學習投入問卷(黃子淇，2021；Connell & Wellborn,1991；Wolters, 2004；Reeve & Tseng, 2011)、認知負荷問卷(Hwang,Yang,& Wang,2013)及期末作品成績。

3.3. AMA 軟體系統介面介紹

AMA (Activate Mind Attention)軟體以 PowerPoint 為平台，外掛增益集，兩者結合成一個數位內容設計及展演、繪本寫作的創意平台。運用結構複製法掌握基本元素，提升學習效益與美感能力。具兩大功能：觸發式動態呈現(Trigger-based Animation, TA)與結構式複製繪圖法(Structural Cloning Method, SCM)。而「結構式複製繪圖法」是以結構與複製的概念來詮釋造形，運用滑鼠精準的掌握大量的物件，功能強大，可以繪製仿山水畫、複雜的對稱構圖及創作繪本圖畫等，是一種新的繪圖法。

4. 研究結果與分析

本研究採用單因子共變數分析(one-way analysis of covariance, ANCOVA)探討不同色彩先備知識之學習者於創作 AMA「幾何裝飾性風格」圖像之自我效能、學習投入、認知負荷及學習成效表現情形。透過雙欄式步驟視覺化教學設計模式的教學後，不同色彩先備知識的學習者對自我效能無顯著差異；在學習投入的表現上，高先備知識組在整體學習投入度及情感投入、認知投入度的表現較低先備知識組好。

另外，也透過獨立樣本 t 檢定，探究期末簡報美學專題之學習成效，可發現兩組無顯著差異；但在認知負荷上，兩組則稍有顯著差異，意即低先備知識組比高先備知識組有較高的認知負荷。

5. 結果與討論

本研究提出雙欄式步驟視覺化之教學設計模式，探究不同色彩先備知識之學生在此教學後之學習表現情形。結果發現，不同先備知識學生在雙欄式步驟視覺化的教學後，高先備知識組其學習投入度顯著較低先備知識組好，而在期末簡報美學專題的學習表現，兩組則沒有差異，但低先備知識組的認知負荷仍較高。也許未來可透過深度訪談，進一步了解其原因。以上結果俾利提供給後續相類似課程與後續研究之參考。

6. 致謝

研究部分經費由 111 科技部專題研究計畫補助，計畫編號：MOST 111-2410-H-224-003-MY2。

參考文獻

黃子淇(2021)。**自我調整學習、成人學習動機與教育訓練學習投入程度之研究－以工作負荷與工作家庭衝突為調節變項**(未出版碩士論文)。國立嘉義大學，嘉義市。

陳俊宏、楊東民(1998)。**視覺傳達設計概論。**臺北：全華圖書。

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change.*Psychological Review ,84*(2),191-215.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development*(pp.43-77). Lawrence Erlbaum Associates, Inc.

Garner, J. K., Alley, M., Gaudelli, A. F.,& Zappe, S. E. (2009). Common use of PowerPoint versus the assertion-evidence structure: A cognitive psychology perspective. *Technical Communication, 56*(4), 331-345.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013), A concept map-embedded educational computer game for improving students' learning performance in natural science courses, *Computers & Education, 69*, 121-130.

Kuh, G. D. (2003). What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning, 35*(2), 24-32. doi:10.1080/00091380309604090

Kundart, J. (2014). Avoiding death by PowerPoint: Can we improve our use of this now ubiquitous teaching tool? *Optometric Education, 39*(3), 104-105.

Pintrich, P.R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. MI: National Center for Research to Improve Postsecondary Teaching and Learning. (ERIC Document Reproduction Service No. ED 338122.

Reeve, J., & Tseng, C.-M. (2011) Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational psychology, 36*(4), 257-267.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory.* New York: Springer.

Wolters, C. A. (2004) Advancing Achievement Goal Theory: Using Goal Structures and Goal Orientations to Predict Students' Motivation, Cognition, and Achievement. Jou*rnal of Educational Psychology, 96*(2), 236-250.

基于改进 IEO 模型分析学生时间投入与数学成绩的影响机制：模型、算法与优化策略

Analyzing the Mechanism of Student Time Investment and Mathematics Achievement Based on the Improved IEO Model: Model, Algorithm, and Optimization Strategies

谢晨[1]，虞增艺[2*]，虞宸越[3]
[123]浙江工业大学 教育科学与技术学院
*202105720431@zjut.edu.cn

【摘要】 在“双减”政策大背景下，学生时间投入与成绩影响的研究意义愈发重大。因此，本研究以台湾 15 岁学生的 PISA2022 数据为基础，改进阿斯汀的 IEO 模型，建立数学成绩时间分配模型，通过序贯多元线性回归算法，揭示时间投入与分配的优化策略。研究发现，时间的投入与分配是影响学生数学成绩的最关键因素。同时家庭内部环境因素也对学生数学成绩起到关键作用，其中母亲的职业状况影响显著。根据研究结果，提出了合理安排课时、提升家庭教育参与度等优化策略，以帮助学生更好地提高数学成绩。此研究成果在 https://github.com/yuzengyi/PISA_xie。

【关键词】 时间分配；数学成绩；PISA；环境；多元线性回归

Abstract: *Under the broad context of the 'Double Reduction' policy, the study of students' time investment becomes significant. This research based on the PISA 2022 data of 15-year-old students in Taiwan, improve Austin's IEO model to establish a mathematical achievement time allocation model. Through sequential multiple linear regression algorithms, it reveals the optimization strategies for time investment and allocation. The study finds the investment and allocation of time are the most critical factors affecting students' mathematics performance. Internal family environment factors play a crucial role in students' mathematics achievements, with the mother's occupational status having a significant impact. Based on the research, optimization strategies are proposed to help students improve their mathematics performance. The results of this study are available in* https://github.com/yuzengyi/PISA_xie.

Keywords: Time allocation, Math scores, PISA, environment, multiple linear regression

1. 前言

在国务院印发的《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》中明确提出要减轻义务教育阶段学生过重的课业负担和校外培训压力，引导学生将学习精力投入到校内高质量教育及多元化的兴趣发展活动中。由此可见，政府已认为学生的时间分配是影响学生学习结果的重要因素之一。

2. 研究设计

2.1.研究对象

本研究以台湾 15 岁学生为研究对象，选取其 2022 年的 PISA 数据进行分析研究。本次研究经过挑选数据列的数据清理筛选方法，共计得到 1746 份有效数据样本。

2.2.研究模型

阿斯汀提出的 IEO 模型认为，个体的输入要素会直接影响成果，也会通过影响环境间接影响结果(Astin, A.W,1993)。基于此，本研究假设台湾 15 岁学生的时间投入会影响数学成绩，同时，也会通过环境间接影响数学成绩。为验证此假设的合理性，本研究通过序贯多元线性回归算法，对 PISA2022 权威数据集进行分析。

在改进的 IEO 模型中，本研究针对台湾 15 岁学生，将学生时间投入作为输入 I，其中包括课内学习时间投入，如数学课时数、数学作业时长等，以及课外娱乐时间投入，如打电子游戏、浏览社交媒体等。此外，本研究视家庭内部环境和家庭外部环境为环境 E，其中家庭内部环境包括父母职业状况、家庭经济状况等，而家庭外部环境则分为学校环境和社会环境两种，学校环境以师生关系为代表，社会环境则以社会归属感为代表。最后，本研究将学生数学成绩作为输出 O，得到模型如图 1 所示。

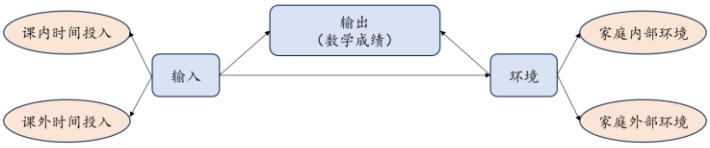


图 1 时间、环境、数学成绩的改进 IEO 模型图

本研究采用了 PISA 六点量表的研究方法，对学生数学作业时间投入和学生课余生活时间投入进行了研究。课余生活时间投入则包括了打电子游戏时间投入、社交媒体时间投入、制作视频等数字媒介产品的时间投入、浏览网页时间投入等等。同时，在分析处理数据时，本研究采用时间段的组中值进行计算。

3. 实证分析

为深入讨论学生时间投入、家庭内外部环境与数学成绩表现的关系，本研究以学生数学成绩表现为因变量，用序贯多元线性回归算法建立数学成绩时间分配模型。在正式建立数学成绩时间分配模型前，本研究先进行了自变量共线性检验。经自变量共线性检验，自变量间不存在共线性问题，适合使用线性回归方法进行分析。同时，本研究对所有的变量进行了哑变量处理，将其应用到序贯多元线性回归算法当中。哑变量能够将原本不能定量处理的变量进行量化，通过取值 0 或 1 来反映某个变量的不同属性，极其适用于序贯多元线性回归算法。根据本研究所建立的四个数学成绩时间分配模型和序贯多元线性回归算法的应用，本研究得到了表 1 的数据。

模型一采用学生在作业方面时间投入的指标作为自变量，结果显示，在模型一中的所有变量均达到显著。这意味着课时数、作业时间投入均对成绩表现有显著影响。模型二在模型一的基础上增加了包括电子游戏在内的课外时间投入指标，探究这些指标对成绩表现的影响。原先模型一的指标在模型二中影响力仍旧显著。同时相对模型一，模型二比模型一多了 12.4%的解释力。模型三又在模型二的基础上增加了家庭内部环境指标，探究它们对学生成绩表现的影响。在模型三中，父母受教育的最高程度、母亲的职业状态和家庭经济状况这三个新增变量对于学生成绩表现是极为显著的。模型四则拥有本研究所研究的所有指标变量。其中，新增指标以家庭外部环境指标为主，探究其影响。在模型四中，对学生成绩表现非常显著的变量只有总课时数、作业总时长投入和打电子游戏时长投入。模型四的解释力是四个模型中最强的，有 56.7%的解释力。

表 1 以数学成绩为因变量的序贯多元线性回归表（部分变量展示）

	模型一	模型二	模型三	模型四
常量	432.691**	492.817**	153.705**	142.094**
数学课时数	5.920**	4.145**	1.869	1.426
总课时数	2.086**	1.764**	1.548**	1.038**
数学作业时长	0.166*	0.167*	0.079	0.059
作业总时长	0.394**	0.294**	0.250**	0.157**
打电子游戏时长		-0.146**	-0.130**	-0.100**
社交媒体使用时长		-0.044	-0.044	-0.022
父母最高受教育水平			24.571**	18.256
母亲职业状态			-0.047**	-0.031*
父亲职业状态			0.003	0.001
家庭经济状况			25.084**	12.394
师生关系				1.42
社会归属感				-4.196
R-square	0.159	0.283	0.389	0.567
N	1746	1746	1746	1746
F-statistic	5.096E-62	2.120E-109	2.433E-145	7.195E-229

注:*表示 p<0.05，**表示 p<0.01,表中的控制变量均做了哑变量处理

4. 结果分析与启示

本研究以阿斯汀的 IEO 模型为基础,通过序贯多元线性回归算法研究了作为输入变量的时间投入和作为环境变量的家庭内外部环境及其各子维度对作为输出变量的研究生创新产出的影响研究得到以下结论。

4.1.时间的投入与分配是影响学生成绩表现的最关键因素

从表 1 不难发现，课时数和作业时长投入对学生的成绩表现具有非常显著的影响。同时也有研究表明，学生在课程上花费的时间显著地影响学习成绩（张婷，2015）。首先，只有拥有一定数量的课时数，学生才能拥有一定的通过教师指导学习的机会。在教师的“传道授业解惑”下,学生的学习会更加的高效。但是过度的课时数也会适得其反，过多的课时压缩了学生的自主学习时间，导致学习成绩受到影响。而对于作业时长，也有学者进行了相关研究。适量的家庭作业有利于提高学生的学习成绩,有助于学生养成自主学习的习惯,但过量的家庭作业会导致学生身心疲惫,使学生对家庭作业甚至学习产生厌烦心。

4.2.家庭内部环境是影响学生成绩的重要因素之一

通过表 1 本研究可以得知，母亲的职业状态对学生成绩一直具有显著影响力。不过关于母亲职业状态为何会成为影响学生成绩的重要因素，国内外在该方面的研究较少。但是母亲是一个家庭当中非常重要的元素,尤其是在“男主外,女主内”的观念影响下,出现了许多家庭主妇。相对于在外工作的父亲，母亲有更多的机会与自己的孩子相处,这也导致了大多数学生的学业成绩很可能是母亲在管教。所以母亲的职业状态会对学生的成绩表现有显著的影响。

4.3.启示与优化策略

总而言之，学生的学习时间投入是影响学生成绩表现的重要因素，适度的学习时间投入——尤其是作业时长和课时数，对学生的成绩具有极其良好的作用。学校应当倡导

高效课堂，提高课堂教学质量，避免无效或冗余的教学时间，让学生在有限的课时内获得最大化的学习效果。老师应当实施差异化作业策略，针对不同学生的能力水平提供不同层次的作业任务。同时，学校应当鼓励父母双方共同参与子女的教育过程，尤其是母亲作为主要监护人，更应密切参与到学生的学业规划和问题解决中来。教育研究的道路任重而道远，随着教育学、教育技术学的发展，越来越多的人意识到教育研究的重要性，未来将会有更多的教育学研究出现，对教育学继续探索，从而为教育事业做出更多地贡献，用更好的科学方法提升学生的学业成绩水平。

5. 总结

本研究以阿斯汀的 IEO 模型为框架，运用序贯多元线性回归算法分析了时间投入、家庭内外部环境及其子维度对研究生创新产出的影响。主要结论如下：

从时间投入的角度，课时数、作业时长对学生成绩表现具有显著影响。适量的课时有助于学生在教师指导下高效学习，但过度课时可能导致自主学习与休闲娱乐时间减少，不利于全面发展。适量的家庭作业能够提高学业成绩并培养自主学习习惯，而过量则可能引发厌学心理。

从家庭内部环境的角度，家庭支持、父母学历及经济社会地位均是影响学生成绩的重要因素。其中，母亲的职业状态对学生成绩有显著影响力，可能是由于母亲在家庭教育中扮演着更为直接且频繁的角色。因此，提倡父母共同参与子女教育，特别是鼓励母亲深度介入学生的学业规划和问题解决过程。

总的来说，适度的时间分配、优化的家庭教育以及建设有利于学生身心健康发展的外部环境，都是提升学生成绩表现的关键要素。未来教育研究将持续深化，利用科学方法改进教育实践，推动学生学业成绩和综合素质的全面提升。

学生的时间分配作为影响学生学业成果的关键因素，这一观点在西方学生发展理论中源远流长。早在 1980 年，佩斯在其学术研究中首次提出了“努力质量”概念，指出学生投入到课堂内外各项活动的时间和精力与其所取得的学习成效之间存在显著正相关性(Ascarella,E.T &.Terenzini,P.T,2005)国内的宋海生等学者基于 PISA2015 中国四省市数据，运用多层线性模型分析了中学生课外补习时间对各科学业成绩的影响，得出了学科补习与学业成绩间存在非线性关系的结论（宋海生和薛海平,2020）。相较于前人的研究，本研究改进阿斯汀的 IEO 模型，并基于 PISA2022 数据集，采用序贯多元线性回归算法对学生时间投入对数学成绩的影响进行探究。最终本研究得出了学生时间分配与数学成绩存在先呈正相关关系，随后在达到一个阈值后呈负相关关系。这与宋海生等国内学者得出的结论较为相似，但并不一致。同时与佩斯所得出的结论完全不同。但是不容置否的是，学生的时间投入对成绩表现确实有显著的影响力。

参考文献（部分）

宋海生, & 薛海平. (2020). 中学生课外补习时间投入对学业成绩的影响——基于 PISA2015 我国四省市数据的实证研究. ***基础教育***, *17*(2), 20-32.

张婷. (2015). H 大学本科生各项活动时间投入现状及其对学习效果的影响硕士(学位论文, 湖南大学).

Pascarella, E. T., & Terenzini, P. T. (2005). How College Affects Students: A Third Decade of Research. Volume 2. Jossey-Bass, An Imprint of Wiley. 10475 Crosspoint Blvd, Indianapolis, IN 46256.

Astin, A. W. (1997). *What matters in college?*. JB.

文學圈結合專題導向學習結合學習分析儀表板於書報討論創新教學研究

Research on Innovative Teaching in Literature Circles Integrating Project-Based Learning with Learning Analytics Dashboards in Seminar

鄭博宇 [1*]，洪暉鈞 [2*]，楊叔卿 [3]

[1,2] 台灣中央大學 網路學習科技研究所 [3] 台灣清華大學 資訊系統與應用研究所

[1*] pomay911@gmail.com [2*] hch@cl.ncu.edu.tw [3]scy@mx.nthu.edu.tw

【摘要】 本研究以台灣北部一所大學的書報討論課堂參與者為研究對象，探討在職研究生與全日制研究生在終生學習環境中如何平衡工作和學習。研究利用文學圈方法提高書報討論的效果，透過小組討論與角色扮演促進深入學習和批判性思考。本研究旨在融入專題導向學習與專案管理知識，並結合科技輔助工具，提升學習效果。研究設計了結構化任務和儀表板，包括專家演講、社會調節、與專案管理儀表板，來支持學習過程和提供及時反饋。這種學習策略旨在提供多面向、結構化的學習環境，促進學生的專業成長和持續學習，研究結果發現，結合文學圈和電腦輔助學習的專案管理學習方式，輔以基於社會調節的儀表板，能夠提升學生的學習成效。

【關鍵字】 文學圈；自我調節學習；書報討論；專案管理；科技輔助學習

Abstract: *This study focuses on participants in a seminar class at a university in northern Taiwan, exploring how working graduate students and full-time graduate students balance work and study in a lifelong learning environment. The research utilized the literature circle method to enhance the effectiveness of journal discussions, facilitating in-depth learning and critical thinking through group discussions and role-playing. The study aims to integrate project-based learning with project management knowledge, and to enhance learning outcomes through the use of technology-assisted tools. Structured tasks and dashboards were designed, including expert lectures, social regulation, and project management dashboards, to support the learning process and provide timely feedback. This learning strategy is intended to offer a multifaceted, structured learning environment that fosters professional growth and continuous learning for students.*

【Keywords】 Literature Circles, Social Regulated Learning, Seminar, Project Management, Computer-Supported Collaborative Learning.

1. 前言

對於研究生來說，書報討論是一種重要的學習活動。透過書報討論，學生可以深入地理解和分析學術文章，從而學習到新的知識和研究方法。此外，這種活動還能訓練學生的批判性思考與表達能力。為了提升書報討論的效果，採用了文學圈（Literature Circles）(Daniels, 1994)的方式來引導學生在閱讀論文的小組討論中進行互助與合作。文學圈的模式是以學生為中心的討論方法，通常由學生自行組成小組，每個人在討論中扮演不同的角色。傳統的文學圈強調學生間面對面的交流和合作，這有助於學生深入閱讀、理解和分析文學作品(King, 2001)。此外，D. M. Thomas 等人(2019) 的研究證實，文學圈能改善學生的閱讀理解能力。因此，研究生可以透過有效的時間管理，持續的自我學習，以及不斷提升溝通與表達能力，透過文學圈的小組討論方法克服學習上的挑戰，以及科技的輔助，使用科技結合書報討論的授課方式，幫助學生學習表現以及成效。

專題導向學習（Project-Based Learning, PBL）是一種讓學生通過實踐專案來深入學習、發展批判思考與團隊合作的學生中心學習方式。學生的工作主要集中在創建、開發和完成一項長期的「專案」，這個專案是圍繞某一個重要的議題或挑戰來設計(Krajcik & Blumenfeld, 2006)。J. Thomas (2000)強調了專題導向學習五個重要的元素：(1) 學生在學習中起主導角色；(2) 專案具有真實世界的背景或適用於真實世界；(3) 專案必須是富有挑戰性的問題或問

題的探討；(4) 學生透過學習過程進行研究和進行探索；(5) 學生必須對他們的工作做出公開的展示。專題導向學習也有其獨特的好處。Bell (2010)提出透過此學習方式，學生能增強自主學習能力，並有更深的學科理解。在職專班中，專案管理學習通常結合服務學習，讓學生參與真實專案，培養溝通、規劃和實務技能。

電腦輔助合作學習（Computer-Supported Collaborative Learning, CSCL）是一種新型的教學方法，這種方法透過科技的應用，如線上討論區、共享文件工具等，以促進學生之間以及學生與教師之間的協同學習(Blumenfeld, Kempler, & Krajcik, 2006)。在過去常見的在在職專班的專案管理學習，大部分都是使用服務學習方法，例如：Brown(2000)透過實際幫助老人修建房子，培養在職生專案管理知識，學生必須與客戶溝通、規劃任務、協調時間表、採購材料、學習施工技能、進行體力勞動並跟踪項目進度。Larson 等人(2010)透過安排慈善機構的活動，培養在職學生專案管理知識。

因此本研究設計出結合專題導向學習（Project-Based Learning, PBL）和電腦輔助合作學習（Computer-Supported Collaborative Learning, CSCL）的教學模式，期望透過這樣的方式，讓書報討論的課程更加互動、技術驅動的方向發展，並以文學圈作為教學活動，促進學生小組閱讀、討論，強調角色扮演，提升閱讀理解和批判性思維。

因此根據上述內容，本研究提出二個研究問題如下:

1.專案管理基礎的專題導向學習（PBL）與學習分析儀表板（LADs）結合的教學方法對於提升學生的論文回顧能力有何效果？

2.在不同背景（在職與全日制研究生）的學生中，對於學習分析儀表板（LADs）的態度和認知有何差異？這些差異如何影響他們的學習動機和學習成效？

2. 書報討論入口學習平台系統



圖 1 書報討論系統



圖 2 系統儀錶板

本研究開發了一個創新的課程入口平台，旨在為學生和講者提供一個互動和資訊共享的中心化點。該平台結合了教育技術和數據分析，以增強學習體驗和課程管理效率。

此外，平台可以直接閱覽學習儀表板，如圖 2 所示，本系統將不同功能儀表板分為提供發問功能的「專家演講儀錶板」、幫助學生追蹤小組進度的「專案管理儀錶板」和小組互評後，幫助理解小組和文學圈角色評價的「社會調節儀錶板」，這些工具不僅提高了學習的透明度和互動性，還通過提供反饋和貢獻度分析，進一步激勵學生的參與和合作，簡述的儀表板功能如表一所示。

表一 儀錶板功能說明

儀錶板	功能
專家演講儀錶板	提供學習者發問以及理解各小組發問次數
專案管理儀錶板	提供學習者理解各項任務討論情形
社會調節儀錶板	提供學習者互評結果和表現最好的文學圈角色

3. 研究方法

3.1. 研究對象

本研究以 52 名台灣北部大學研究生為研究對象進行實驗，參與者分為 17 名在職學生以及 35 名全日制學生。其中在職學生的工作背景如表二所示，可以發現絕大多數的在職生年資落在 5 年以上，以一般職員為主，特別的是其中有 4 名單位主管，1 名老闆。

表二 在職學生的工作背景調查表

	年資	**人數**	**百分比**
工作年資	一年以下	0	0%
	一到三年	4	23.5%
	三到五年	3	17.6%
	五年以上	10	58.8%
	職位	**人數**	**百分比**
工作職位	一般職員	12	70.6%
	部門主管	4	23.5%
	單位主管	0	0%
	老闆	1	5.9%

3.2. 實驗流程

本研究課程為一門 18 週的學術研究與討論課程，主要由課程介紹、資源利用、主題演講、研究討論、期中報告以及期末報告等多個部分組成。

在課程的開始，我們將進行一次課程介紹，對整個課程的內容和期望結果進行概述。同時，也將進行一次前測，了解學生對於課程內容的初步認識和期待。

我們將進行一系列的主題演講和研究討論。主題演講將由不同的講者進行，以確保我們能夠探討到最新、最相關的主題。每四次主題演講後，都將進行一次研究討論，根據不同的主題，由學生進行文獻回顧和專題探討，深化對演講內容的理解和應用。

每一次的團體作業皆須扮演文學圈的角色，角色的工作內容如表三所示，並使用 Trello 專案管理軟體討論作業，在每一次的作業報告中，會透過小組互評表單，互評各組的表現，並選出最佳表現的角色。在課程的最後兩週，學生將需要進行期末報告。這將是一次學生展現他們整學期學習成果的機會，並透過同儕互評和期末專題展演，分享他們的研究成果和心得，最後，受測者會再進行一次後測，其類型和題目數量皆與前測相同，內容不同。藉以觀察受測者在 18 周創新教學方式結果前後的差異。

表三 基於(Daniels, 1994)改編的文學圈角色

角色	角色說明
討論領導者	領導討論，確保討論的有效進行，並鼓勵成員參與。
總結者	提供論文的概要，幫助團隊成員回顧論文內容。
連接者	尋找論文與個人經驗或現實世界之間的聯繫，並與小組討論。
單詞大師	選擇並解釋論文中重要的專有名詞，幫助小組更好地理解論文內容。
段落人	透過段落解釋，理解整篇論文架構。

3.3. 資料分析

本研究之資料分析將會透過量化分析的方式進行，進一步來說，依照前後測的題目觀察受測者在 18 周後的學習後，是否學習表現有進步，前後測驗題型相同，題目不同每一格答對的分數為 10 分。並使用成對 t 檢定將會被應用來檢驗前後測驗得分是否有顯著差異。
本研究問卷參考學習動機調查問卷(Motivated Strategies for Learning Questionnaire, MSLQ)問卷，採用 Pintrich (2000)提出之學習動機理論，根據(L.-A. Lim et al., 2020; L. Lim, Dawson, Joksimovic, & Gašević, 2019)改編儀錶板動機編碼，衡量學生對於不同學習儀表板的學習動機。

4. 結果與討論

4.1. 學習表現的差異

本研究以學生論文回顧能力測驗前後測利用成對 t 檢定比較學生論文回顧能力前、後測之差異。根據表一結果顯示，經一學期實驗，學生的論文回顧能力有非常顯著的進步(z=0,p=.000<0.001)。代表本研究的課程設計有助於提升論文回顧能力的學習成效。

表四 在職生論文回顧能力測驗成對 t 檢定結果

	測驗	**M**	**SD**	**t**	***p***
論文回顧能力	前測	37.647	14.8026	-11.279	.000
	後測	66.471	13.6662		

**p<.05, **p<.01, ***p<.001*

表五 一般生論文回顧能力測驗成對 t 檢定結果

	測驗	**M**	**SD**	**t**	***p***
論文回顧能力	前測	50.571	19.0885	-4.781	.000
	後測	69.714	13.0348		

**p<.05, **p<.01, ***p<.001*

表四數據顯示在職生前測的平均值（M）為 37.647，標準差（SD）為 14.802，而後測的平均值為 66.471，標準差為 13.666。這個 t 檢定的 t 值為-11.279，p 值小於 0.001（p < .001），前測和後測之間存在顯著的差異。

表五數據顯示一般生前測的平均值為 50.571，標準差為 19.088，而一般生的後測平均值為 69.714，標準差為 13.034。這次 t 檢定的 t 值為-4.781，p 值同樣小於 0.001。這也指出在這個變量上，前測和後測之間存在顯著的差異。

在職學生和一般學生在前測和後測中均顯示出了顯著進步，這可能表明所採取的教學方法或學習活動對兩組學生都是有效的。在職學生的平均得分提升似乎更為顯著，這可能是由於他們將工作經驗與學習內容結合，從而在某些領域獲得了更多的增益。雖然一般學生的得分提升也很顯著，但是在職學生的進步幅度更大，這可能反映了在職學生更能夠在實際工作和學術學習之間做出有效連結。兩組學生在後測中的標準差都有所下降，這表明學生間的得分差異縮小，整體表現更趨一致。

總體來說，這些結果表明所實施的學習計劃或課程對於提升學生的能力是有效的，尤其是對在職學生來說，其進步可能得益於他們將學習與工作經驗相結合的能力。這些發現為未來的教學設計提供了重要見解，特別是在考慮如何最大化不同背景學生的學習成效時。

4.2 儀錶板對於學生學習動機反饋

本研究選用(Pintrich, 2000)所設計學習動機問卷(Motivated Strategies for Learning Questionnaire, MSLQ)，並根據(L. Lim et al., 2019)的編碼規則，將 MSLQ 問卷進行編碼，易

於理解、趨勢等，設計出對應的代碼內容。問卷資料將依照各項動機因素進行編碼，並將每個學生的回答轉換成數值資料，作為後續分析學生學習動機變化的依據。透過這樣的編碼方式，本研究期待能夠準確地量化學生的學習動機，並探討其與學習成效之間的關係，並讓學生以五點量表進行評分：非常同意為 5 分、同意為 4 分、普通為 3 分、不同意為 2 分、完全不同意為 1 分。

表六 專家演講儀表板學習動機問卷各項比例

面向	**組別**	**M**	**SD**
易於理解	在職生	4.7059	.46967
	一般生	4.6000	.65079
趨勢	在職生	4.7059	.46967
	一般生	4.5714	.81478
便於比較	在職生	4.7647	.43724
	一般生	4.4857	.78108
外觀	在職生	4.6471	.60634
	一般生	4.4571	.78000
準確性	在職生	4.3529	.78591
	一般生	4.4571	.81684
時間與學習環境管理	在職生	4.5882	.61835
	一般生	4.3143	.83213
後設認知自我調節	在職生	4.5294	.62426
	一般生	4.2571	1.03875
努力及調節	在職生	4.1765	.95101
	一般生	4.2000	1.03875

在這項對專家演講儀表板功能同意程度的評估中，如表六所示，我們發現在職生在理解易度（M=4.7059, SD=0.46967）和趨勢識別（M=4.7059, SD=0.46967）方面的同意程度略高於一般生（理解易度的 M=4.6000, SD=0.65079；趨勢的 M=4.5714, SD=0.81478）。這可能意味著在職生認為他們的實務經驗有助於他們更好地利用這些儀表板。在比較便利性（M=4.7647, SD=0.43724）和時間與學習環境管理（M=4.5882, SD=0.61835）方面，在職生同樣表示較高的同意度，相較於一般生（比較便利性的 M=4.4857, SD=0.78108；時間與學習環境管理的 M=4.3143, SD=0.83213）。這反映出在職生可能因為在工作中需要不斷進行資料比對和時間管理，因此更能體會到儀表板在這些方面的實用性。然而在準確性方面，一般生表達了較高的同意程度（M=4.4571, SD=0.81684），而在職生的同意程度較低（M=4.3529, SD=0.78591）。這可能反映出一般生在學術學習中對準確性有較高的期待，而在職生則可能根據工作經驗對準確性有不同的見解。在後設認知自我調節上，在職生的同意程度（M=4.5294, SD=0.62426）也高於一般生（M=4.2571, SD=1.03875），這顯示在職生可能在工作中學會了更多自我監控和調整的策略。在努力及調節方面，一般生和在職生的同意程度相近（一般生的 M=4.2000, SD=1.03875；在職生的 M=4.1765, SD=0.95101），表明兩組人對於透過儀表板進行自我努力和調節的看法並無太大差異。

總結而言，實務工作經驗似乎讓在職生對於演講儀表板的各項功能有較高的認同，特別是在易於理解和資料比較的方面。這樣的結果可能來自他們能夠將儀表板功能與他們的職業實踐相結合，從而更有效地應用在他們的工作和學習中。

表七 專案管理儀表板學習動機問卷各項比例

面向	組別	M	SD
易於理解	在職生	4.7059	.46967
	一般生	4.4286	.88403
趨勢	在職生	4.5882	.50730
	一般生	4.4857	.91944
便於比較	在職生	4.6471	.49259
	一般生	4.5143	.78108
外觀	在職生	4.4706	.62426
	一般生	4.3714	.97274
準確性	在職生	4.4118	.79521
	一般生	4.4571	.85209
時間與學習環境管理	在職生	4.4118	.79521
	一般生	4.2571	.85209
後設認知自我調節	在職生	4.5294	.62426
	一般生	4.3429	.83817
努力及調節	在職生	4.2941	.98518
	一般生	4.2571	.95001

在評估專案管理儀表板的功能同意程度時，表七顯示在職生在易於理解（M=4.7059, SD=0.46967）方面的同意程度顯著高於一般生（M=4.4286, SD=0.88403）。這可能反映出在職生由於具有與工作相關的經驗，更能體會儀表板提供資訊的直觀性和實用性。在趨勢識別上，雖然在職生（M=4.5882, SD=0.50730）仍然高於一般生（M=4.4857, SD=0.91944），但兩者之間的差距相對縮小。這可能是因為趨勢分析需要特定的技能和知識，而這些可能在學術訓練中也有所涵蓋。在比較便利性方面，在職生（M=4.6471, SD=0.49259）的同意度依然高於一般生（M=4.5143, SD=0.78108），這顯示出在職生可能更能夠從儀表板中快速抓取關鍵資訊，進行有效的比較。外觀方面，雖然在職生的同意程度（M=4.4706, SD=0.62426）略高於一般生（M=4.3714, SD=0.97274），但此差異不大，這或許表明儀表板在視覺呈現上對於不同背景的人士都有一定的吸引力。在準確性方面，一般生（M=4.4571, SD=0.85209）的同意度略高於在職生（M=4.4118, SD=0.79521），這可能表明一般生在此方面有更高的期望，或是他們對儀表板提供的資訊準確性有更嚴格的評價標準。時間與學習環境管理方面，在職生（M=4.4118, SD=0.79521）的同意度高於一般生（M=4.2571, SD=0.85209），這可能反映在職生認為儀表板在時間管理和環境調節上提供了實質幫助，這也可能與他們在工作中需要有效管理多個任務的經驗相關。後設認知自我調節方面，在職生（M=4.5294, SD=0.62426）相比一般生（M=4.3429, SD=0.83817）也表達了較高的同意程度，這顯示在職生可能更認同儀表板在幫助自我調節學習進程方面的效用。最後，在努力及調節這一項目中，兩組之間的同意程度非常接近（在職生的 M=4.2941, SD=0.98518，一般生的 M=4.2571, SD=0.95001），表明無論是在職生還是一般生，都相信儀表板在助益個人努力和調節學習行為方面有其價值。

數據分析表明，在職生普遍對於專案管理儀表板的各項功能表達了較高的同意程度，特別是在易於理解和比較便利性方面，顯示他們可能因為職場經驗而更能夠欣賞和利用儀表板的資訊。一般生在準確性方面的同意程度略高，這可能反映了他們對資訊準確性的高期望。在時間管理和後設認知自我調節方面，兩組差異顯著，暗示在職生可能更能認同儀表板在這

些方面的實際應用價值。整體而言，這顯示在職生可能會因為他們將工作與學習結合的經驗，而在使用專案管理儀表板時有著不同的認知和評價。

表八 社會調節儀表板學習動機問卷各項比例

面向	組別	M	SD
易於理解	在職生	4.6471	.60634
	一般生	4.5714	.85011
趨勢	在職生	4.6471	.60634
	一般生	4.6571	.80231
便於比較	在職生	4.4706	.62426
	一般生	4.6000	.81168
外觀	在職生	4.6471	.60634
	一般生	4.5143	.88688
準確性	在職生	4.5882	.50730
	一般生	4.5143	.88688
時間與學習環境管理	在職生	4.2353	.83137
	一般生	4.2000	.99410
後設認知自我調節	在職生	4.4118	.71229
	一般生	4.3429	.96841
努力及調節	在職生	4.2941	.91956
	一般生	4.2000	.99410

針對社會調節儀表板功能的同意程度，表八的結果分析揭示了在職生與一般生在不同面向上的觀感差異。在易於理解方面，儘管在職生（M=4.6471, SD=0.60634）與一般生（M=4.5714, SD=0.85011）的同意程度相近，但在職生顯示出較小的變異性，這可能意味著他們對儀表板功能的理解更為一致。在趨勢識別上，一般生（M=4.6571, SD=0.80231）輕微領先於在職生（M=4.6471, SD=0.60634），表明可能在學術訓練中獲得的趨勢分析能力對他們理解社會調節儀表板同樣有益。然而，便於比較這一面向則看到一般生（M=4.6000, SD=0.81168）的同意度超過了在職生（M=4.4706, SD=0.62426），這可能顯示一般生在使用儀表板進行比較分析時感到更為得心應手。外觀方面，在職生的同意程度（M=4.6471, SD=0.60634）略高於一般生（M=4.5143, SD=0.88688），這可能與在職生對專業工具外觀的偏好有關。準確性方面，兩組人士的同意程度也相近，但在職生（M=4.5882, SD=0.50730）表現出較小的變異性，這或許表明他們對儀表板提供資訊的準確性持有較一致的認知。在時間與學習環境管理方面，雖然在職生（M=4.2353, SD=0.83137）的同意程度只輕微高於一般生（M=4.2000, SD=0.99410），但這個結果仍舊表示在職生可能更認可儀表板在這方面的效用，可能因為他們在工作中已經習慣進行類似的管理活動。後設認知自我調節面向上，兩組的同意程度相似，但在職生（M=4.4118, SD=0.71229）的回應更為一致，這可能顯示他們對於儀表板在促進自我調節能力方面的認同感較強。努力及調節方面，再次見到在職生（M=4.2941, SD=0.91956）與一般生（M=4.2000, SD=0.99410）的同意度接近，表明儀表板在支援學習者個人努力和調節學習行為方面的普遍認可。

在對社會調節儀表板功能的評估中，數據表明在職生對儀表板的易理解性和外觀較一般生有更高的同意度，反映出他們對這些方面有更一致的看法。相反，一般生在趨勢識別和便於比較方面給出了更高的評價，可能是受到他們學術背景的影響。在職生和一般生對準確性的看法相近，但在職生意見更集中。時間管理和自我調節方面的差異不大，但在職生略顯認可儀表板的效用。總體上，兩組對儀表板功能的認可程度反映了他們不同的經驗背景。

5. 結論

本研究旨在通過結合基於專案管理的專題導向學習（PBL）和學習分析儀表板（LADs），並透過文學圈的方式討論，增強書報討論的學習體驗。研究著重於通過結構化任務、專家講座和社會調節提高在職與全日制研究生的批判性思維和深入學習。兩組學生在論文回顧能力都有顯著進步，其中在職學生展示出更大的提升，這表明將工作經驗與學術學習相結合是有效的。此外研究結果也呈現不同組學生的對於儀表板的態度不同，本研究強調了考慮學生不同背景以提供未來其他研究者透過創新教學方法，最大化學生在書報討論學習成效的教育方法的重要性。

致謝

本研究感謝國科會專題研究計畫（計畫編號: NSTC112-2628-H-008-001-、NSTC 111-2410-H-007-023-/MOST 112-2410-H-007-0023、MOST 108-2511-H-007 -008 -MY3）與教育部教學實踐研究計畫（計畫編號：PSK1122995）以及國立中央大學學習科技研究中心支持。

6. 參考文獻

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The clearing house, 83*(2), 39-43.

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2006). *Motivation and cognitive engagement in learning environments*: na.

Brown, K. A. (2000). Developing project management skills: A service learning approach. *Project Management Journal, 31*(4), 53-58.

Daniels, H. (1994). Literature circles. *York, ME: Stenhouse*.

King, C. (2001). “I like group reading because we can share ideas”: The role of talk within the literature circle. *Reading, 35*(1), 32-36.

Krajcik, J. S., & Blumenfeld, P. C. (2006). *Project-based learning*: na.

Larson, E., & Drexler Jr, J. A. (2010). Project management in real time: A service-learning project. *Journal of Management Education, 34*(4), 551-573.

Lim, L.-A., Dawson, S., Gašević, D., Joksimović, S., Fudge, A., Pardo, A., & Gentili, S. (2020). Students’ sense-making of personalised feedback based on learning analytics. *Australasian Journal of Educational Technology, 36*(6), 15-33.

Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019). *Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference?* Paper presented at the Proceedings of the 9th international conference on learning analytics & knowledge.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In *Handbook of self-regulation* (pp. 451-502): Elsevier.

Thomas, D. M., & Kim, J. K. (2019). Impact of literature circles in the developmental college classroom. *Journal of College Reading and Learning, 49*(2), 89-114.

Thomas, J. (2000). A review of research on project-based learning. San Rafael, CA: Autodesk. In.

信能數學：教師學習分析儀表板

Confident-interest Mathematics: Teacher Learning Analytics Dashboard

周得揚[1*]，葉彥呈[1]
[1]台灣中央大學 網路學習科技研究所
[*]aaronzhou939@gmail.com

【摘要】 數學是基本能力評估的重要項目，也是國民教育不可或缺的基礎課程之一。提升學生學習興趣、降低數學焦慮感，並建立自主學習習慣是教學的重要目標。在先前研究中，團隊已發展一套適用於國小數位學習平台，稱為「數學島」。為了提升學習動機，在系統內建立成就目標功能，使學生可以根據自身學習目標持續追求進步。數位學習平台內的成就功能未包含教師儀表板，因此本研究以「學習分析儀表板」為核心，滿足教師對學生學習狀況的掌握需求，提供更有效的學習數據和個別化的回饋。透過此系統，教師能夠有效地引導學生在數學學習中的成長與成功。

【關鍵字】數學島；成就目標；自我效能；數學焦慮

Abstract: *Mathematics is crucial for assessing competencies and an indispensable element of national education. Key objectives in education include enhancing students' interest, alleviating math anxiety, and fostering autonomous learning habits. In prior research, our team developed 'Math Island,' a tailored digital learning platform for elementary school students. To heighten learning motivation, we integrated an achievement goal feature, enabling students to continually pursue progress based on their individual learning objectives.Recognizing the absence of teacher dashboard within the achievement system of the digital platform, this study emphasizes the importance of the 'Learning Analytics Dashboard' to address teachers' needs in understanding students' learning situations. This tool offers effective learning data and personalized feedback, empowering teachers to guide students toward growth and success in math*.

Key words: Mathematical island, Achievement goal, Self-efficacy, Mathematical anxiety

1. 前言

在傳統的教學中，教師以紙本來記錄學生的學習進度與成效，然而在當今數位化時代，紙本記錄帶來的不方便性日益凸顯。首先，在「訊息即時性及有效性」方面，由於紙本資料分散在不同的文件夾或冊子中，導致教師需要花費大量時間來找尋和檢閱相關資訊；在「資料更新」方面，由於紙本記錄難以快速更新，因此教師需要手動修改或添加新的學生訊息、評估結果等內容，增加教師的工作繁瑣性；在「共享和協作」方面，要與其他教師或學校人員分享資料，需要進行複雜的傳遞過程，難以實現實時協作。在「保存方面」，容易受到環境的影響，像是遺失、損壞等，紙本文件的保存變得相當困難。傳統的紙本教學記錄已無法應對當前多元而迅速變化的教學環境，迫切需要轉向更為便捷、高效的數位化的紀錄方式。

數學島是一個互動式數學學習系統，提供小學一年級至六年級的數學概念數位教材，目前已經有 400 多所小學使用，以數學知識地圖為架構，清晰而有組織的方式呈現不同數學概念之間的關係，使學生更容易理解數學的整體結構。數學島透過悅趣化的形式，包含各種有趣的元素，讓學生能夠在學習的同時享受到樂趣，數學島鼓勵學生以自主的方式規劃、經營、管理個人學習，藉由互動性和自主性的結合，不僅使學習變得更加有趣，還培養了學生的學習動機和主動學習的能力。

本研究在先前團隊數位學習平台數學島的「成就目標功能」基礎上進行延伸，設計「教師分析儀表板」，學習分析是一個新興的領域，探討與學生學習及其環境相關的數據的測量、收集、分析和報告(Brown, 2011; Chatti, Dyckhoff, Schroeder & Thüs, 2012)，提升教師對學生學習情況的掌握，透過教學目標進度的呈現，不僅能全面展示整體班級的學習進展，更能呈現每位學生的個別表現。在學習分析儀表板的設計中，系統進度條依照學生不同的類別顯示學生的表現，使教師能夠迅速識別學生的學習進展，進而更有針對性地進行教學調整和支援，儀表板提供了學習者或課程的當前和歷史狀態的圖形表示，以實現靈活的決策（Few，2006）。這不僅提供了全面的教學洞察，更使教師能夠有個別化地協助每位學生，有助於最大程度地發揮每位學生的學習潛力，確保他們在數學學習中獲得最大的效益，整合成就目標的學習分析儀表板不僅豐富了教學環境，也為提升學習效果和促進教學個別化注入了新的活力。

2. 文獻探討

2.1. 數學島

數學島是一套國小的數學數位學習平台，用以提升學生數學學習興趣與自主學習之能力。數學島以國小數學領域知識地圖為架構，分為四大主題島嶼，並於島嶼內建立不同年級的單元「路線」，學生依據自己的年級選擇不同的路線來學習，例如：加法路、乘法路等多條路線，提共學生學習。數學島採用自我營造「島嶼」的模式進行數學學習，每位學生需自我建築及管理一座屬於自己的島嶼，當學生完成數學任務時，將可獲得「金幣」與「人力」作為獎勵，金幣方面可以用來購買建築物及船隻，人力方面用來建置建築物，當金幣及人力達到一定程度後，將可以升級建築物，藉由不斷的建置與升級建築物，來表示自己的數學學習成就，透過自我營造的模式，將學習歷程與進度視覺化，有助於提升學生的主動性及學習興趣。

2.2.成就目標

成就目標導向指的是學生在追求學習目標時所擁有的目的和動機(Barron & Harackiewiz, 2001)，成就目標理論提出，透過理解學生在參與學術活動時所擁有的原因或目的，可以深入了解學生的動機和與成就相關的行為 (Ames, 1992; Dweck & Legget, 1988; Urdan, 1997)。成就目標在學習過程中扮演著重要的角色，對學生的學習表現產生重大影響。透過明確的成就目標，可以幫助學生更專注於新知識的獲取、克服挑戰，提升自身的能力水平。不同的成就目標在不同的情境下可能具有更大效益，對於不同的個體而言，選擇不同的成就目標也可能更加有效，透過適當地引導建立學生的成就目標，有助於打造積極且有效的學習環境。此外，成就目標的設定是一個靈活的過程，可以根據學生的個人發展和學習需求進行調整。透過定期的評估和反饋，我們能夠更好地了解學生的成就目標是否仍然適切，並在需要時進行相應的調整。

2.3.自我效能

自我效能指一個人相信自己能夠成功執行所期望的行為，以達到預期的結果(Bandura, 1977)。根據自我效能理論，成功的表現經驗、透過觀察他人的學習、情緒的激發以及他人的口頭說服都可以影響個人的自我效能水平，具有高自我效能的人通常更有可能接受挑戰、勇於面對困難，並展現出積極的行為；相反地，低自我效能的人可能會避免挑戰、對失敗產生恐懼，以及對學習和成就抱持著消極態度。從社會學習理論的觀點而言，自我效能被認為是影響「數學態度」和「數學焦慮」的重要的因素(Bandura, 1977;Hackett & Betz, 1981)，自我效

能不僅在個人的學習和成就中扮演著重要角色，在其他領域中亦被廣泛的應用。因此，我們的目標是透過提升學生的數學自我效能，同時激發其對數學的學習興趣和自信心，從而提升整體的數學學習表現。

2.4. 學習分析儀表板

學習分析（Learning Analytics，LA）是指透過線上學習環境廣泛使用所產生的互動數據、個人數據等日益增長的大量資料（Ferguson, 2012）。透過清晰的數據呈現，教師能夠更好地理解學生的學習需求，並制定有效的學習策略，而學習分析儀表板旨在將有關學習者、學習過程和學習背景的各種指標匯聚到一個或多個可視化的顯示中(Schwendimann et al., 2016)。結合數學島的成就系統功能，我們設計了一個名為「學習分析儀表板」的工具，讓教師能夠即時追蹤學生的學習進度，並根據這些數據進行個別化的教學。儀表板的優點體現在資訊整合方面，它能夠將有關學習者、學習過程和學習環境的多樣指標整合到一個或多個可視化的展示中。這有助於教師更全面地了解學生的學習情況。同時，即時性是儀表板的一大特點，它提供了即時的學習數據，使教師和學生能夠隨時掌握學習進展和表現，並快速做出相應的調整。此外，學習儀表板旨在支援教師和學生進行基於數據的明智決策。

3.系統介紹

3.1. 「學習分析儀表板」設計概要

本研究基於先前研究團隊成功開發的「數學島」之數位學習平台，特別著重在成就功能上（如圖 1）。依據學生成就數據（如圖 2）開發「學習分析儀表板」，學習分析儀表板不僅能夠協助教師全面掌握學生的學習進度，更能夠根據這些數據精準地調整教學方式。本系統主要分為兩個頁面，包括「班級學習歷程」及「學生學習歷程」。透過學習分析儀表板，教師得以深入瞭解整體「班級」學習動態和個別「學生」的學習情況，並進行更有針對性的教學，不僅可以提升教學效果，也能滿足學生的個別需求，透過分析學習儀表板呈現的數據，教師可以發現學生可能面臨的困難，並及時提供相應的協助和支援，別化的指導不僅使學生感到被理解和重視，也有助於激發他們更積極參與學習過程。透過對學生學習情況的細緻分析，教師可以調整課程內容和教學方法，以更好地迎合學生的學習風格和節奏，營造一個更有活力和積極學習氛圍的教室環境，進而提升整體學習體驗。



圖 1、數學島成就功能



圖 2、學生成就紀錄

3.2. 系統開發與架構

數學學習分析儀表板是一套用於網際網路上的網頁系統，整體系統運用到前端技術、後端技術以及資料庫技術。伺服器端透過 PHP 的方式向 MySQL 資料庫取得學生於數學島上的

成就紀錄，再經由 PHP 做運算後，傳回 JavaScript 供前端畫面使用。前端部分使用 HTML5、CSS 與 JavaScript 技術進行實作。其中，HTML5 進行網頁切板，架構出網頁畫面的結構。CSS 部分則是負責網頁美化，舉凡文字大小、文字顏色、背景圖片等。JavaScript 的部分，則是資料接收與互動控制的功能。首先，資料的傳遞是藉由 Ajax 技術實現，當伺服器端接受到請求後，即開始處理問題並回傳資料給網頁端，此時回傳的資料會藉由 JavaScrip 去讀取。接收到資料後，根據此進行前端畫面的渲染。前端畫面顯示所有的成就項目。後端使用 PHP 技術實現，其主要負責與資料庫溝通，從資料庫獲取學生於數學島上的成就紀錄。

3.3. 班級學習歷程

「班級學習歷程」（如圖 3）包含「數學島狀態」、「成就收集狀態」、「登入次數統計」、「排行榜列表」四個部分。「數學島狀態」以島嶼的形式呈現整體班級在不同數學單元上的學習進度，每個島嶼對應到不同年級和主題。透過這種島嶼進度的視覺化呈現，教師能夠迅速辨識整體班級在不同主題上的優勢和需要加強的領域。同時，教師可根據島嶼進度觀察班級學生對數學單元的學習進度，有助於更有針對性地調整教學策略，提供額外的支援或挑戰，確保每位學生在整個數學學習過程中都能夠取得最佳的學習效果；「成就收集狀態」展示了整體班級在不同成就收集方面的現況，而這些成就收集細項包括多元的學習層面。透過這些學習面向，教師得以深入了解班級整體的優勢和挑戰，進而提供相應的輔助和支援。這有助於促進學生在各方面的全面發展，同時培養一個鼓勵協作和共同進步的學習氛圍，使整個班級共享成就並共同成長；「登入次數統計」提供學生學習參與度的可視性，使教師能夠清晰識別出哪些學生更積極參與學習活動，透過定期的登入統計，教師可以快速辨識學生對數位學習平台的使用頻率，進而了解其學習動機和參與度，不僅提供了對整體班級的概觀，更使教師能夠有目標地關注可能需要額外支援或鼓勵的學生，教師可以更精準地制定教學策略，促進學生的主動參與，確保他們在數學學習過程中獲得最大的成效；透過「排行榜列表」，教師不僅能夠明確辨識出高成就學生和低成就學生，更有助於深入瞭解每位學生的學習表現。這樣的分析為教師提供了制定有針對性的教學策略的依據，以因材施教，滿足不同學生的需求，學生可能受益於更富挑戰性的學習活動，也可以給予學生可能需要額外的支援和輔導，這樣的差異化教學不僅有助於提升整體班級的學習效果，更確保每位學生都能夠實現最佳的學習成果。透過這樣的視覺化工具，教師可以更有效地支持學生的成就發展，並為每位學生提供度身定制的教學支援，進一步促進了教學的個別化和差異化，確保每位學生都能夠充分發揮潛力，取得更好的學習成果。



圖 3、班級學習歷程

3.4. 學生學習歷程

「學生學習歷程」（如圖 4）包含「成就收集狀態」、「登入次數統計」二個部分。「成就收集狀態」呈現每位學生成就收集的情況，使得教師能夠迅速、清晰地了解每位學生在不同類別中的成就收集情況。這種詳盡的視角使教師能夠更全面地評估每位學生的學習現況，透過深入了解每個學生的成就收集狀態，教師可以有策略地制定輔導計畫，為每位學生提供量身定制的輔助，這種個別化的支援能夠更有效地滿足學生的學習需求，推動學生對數學領域的全面發展；「登入次數統計」為教師提供了學生學習參與度的直觀可視性，使教師能夠清晰識別出哪些學生更積極參與學習活動。透過定期的登入統計，教師可以快速辨識學生對數位學習平台的使用頻率，進而了解其學習動機和參與度，這項統計數據有助於教師了解學生對數位學習的態度，例如，頻繁的登入可能顯示學生對所學內容的積極興趣，以及對學習平台的信任感，這樣的積極學習態度通常伴隨著強烈的學習意願，也顯示出學生深入了解和掌握相應知識的渴望。相對地，較少的登入可能需要進一步探討學生可能面臨的學習障礙或挑戰，教師透過「登入次數統計」觀察採取措施，鼓勵學生積極參與並提供額外的資源或協助。



圖 4、學生個人學習歷程

4.研究方法

4.1.參與對象

本研究之以某市國民小學之 3 至 6 年級學生（共約 60 位）及數學教師 2 位為研究對象，在數學課堂中進行「數位學習平台」，觀察其對學習成效之影響並依據教師的訪談設計「學習分析儀表板」。

4.2. 研究方法與流程

本研究在數學課堂中，配合使用前述之「數位學習平台」，讓學生進行自我學習，透過「學習分析儀表板」功能，讓教師能夠即時追蹤並評估學生的學習進度。本研究過程中，依據系統後台提供的學習數據進行深入分析，以協助教師全面評估班級對數學單元的掌握程度，包括辨別學生對哪些單元已經熟練掌握，以及哪些單元可能需要進一步學習的練習，並且掌握每位學生的進度，進而提升整體班級的學習狀況。透過成就學習分析儀表板的功能，教師得以有效地了解整體班級的學習情況，並深入掌握每位學生的學習進展，這不僅為教師提供了客觀的數據支持，也為教學提供了明確的方向。數學島系統的應用以及學習分析儀表板的支持，使學生能夠更具自主學習的能力，培養在學習過程中的主動性和積極性。

5.結論與預期結果

「數學島」是一套自主經營數學學習遊戲，提供學生完整數學教材以及給予完整學習歷程，讓學生能反思與根據自己的進度自學，透過先前研究分析結果可知，落後學生使用本系統後，數學標準測驗的成績達到顯著進步，並且學生們在應用題的表現超過全國常模標準，可見系統對於落後學生更有幫助（葉彥呈，2019）。

學習數據儀表板對教師而言是一個極具價值的資源，它不僅提供了即時的學生數據，還能夠在教育過程中為教師提供更深入、更全面的學生了解。透過學習儀表板，教師能夠獲得豐富多樣的學習數據，進而增加對學生的掌握度並且有策略性的調整教學方式。教師應該持續反思和改進自己的教學實踐，以更好地引導學生的數學學習，激發他們的學習興趣和自信心（楊承祐，2023）。

本研究規劃於 2024 年 2 月展開，系統方面已經完成階段性的開發，接下來計劃在 2 月份將系統引入實際教學現場，供教師使用。在引入系統前，對學生進行前測問卷的調查，包含「數學焦慮問卷」、「數學自我效能問卷」以及「數學興趣問卷」，瞭解學生現階段的數學學習現況。在系統導入後，擬定教師訪談大綱。預計在 6 月份，對教師進行訪談，並進行深入的觀察，以了解教師在應用「學習分析儀表板」的輔助下，是否改變了他們的教學方式，進一步優化教學環境；同時，對學生發放後測問卷，透過這項研究，我們將關注教師透過學習分析儀表板的輔助，在教學過程中學生的「學習動機」是否有所提升、「自我效能感」是否有所增強，同時我們也將密切關注學生在學習過程中是否有減少「數學焦慮感」的趨勢。

數位學習平台過去主要聚焦於學生的學習需求，提供了便利的學習工具，作為學生學習過程的輔助，在這種情境下，教師往往被視為學生學習的輔助者，致力於提供支援和指導。然而，為了更全面地滿足教學需求，建議進一步以成就為基礎設計「教師成就系統」，核心理念是賦予教師自主制定和實現專屬於他們的成就目標，透過這樣的系統，教師可以設定個人化的學習目標，從而提高其自主學習的動機讓教師有自己的成就紀錄，教師成就系統也可以作為學校和教育機構評估教師專業發展的一個重要標準，學校可以透過這樣的系統更全面地了解教師在不同領域的專業能力和成就，進而提供有針對性的培訓和支援。

致謝

本研究在臺灣國科會人文處 (NSTC 111-2410-H-008-013-MY3) 與「中央大學學習科技研究中心」的資助下完成，僅此致謝。

參考文獻

葉彥呈(2019)。數學島：興趣驅動之國小數學線上平台設計與初步評估

楊承祐（2023）。發展數學成就目標系統並初步評估對增進興趣與自我效能的成效

Akin, A., & Kurbanoglu, I. N. (2011). The relationships between math anxiety, math attitudes, and self-efficacy: A structural equation model. *Studia Psychologica*, *53*(3), 263.

Barron, K. E., & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: testing multiple goal models. *Journal of personality and social psychology*, *80*(5), 706.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Brown, M., (2011). Learning analytics: the coming third wave. Educause Learning initiative, 2011

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H., (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4, 5,* (2012), 318-331

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5-6), 304-317.

Few, S. (2006). *Information dashboard design: The effective visual communication of data.* Sebastopol, CA: O'Reilly Media.

Jameson, M. M., & Fusco, B. R. (2014). Math anxiety, math self-concept, and math self-efficacy in adult learners compared to traditional undergraduate students. *Adult Education Quarterly*, *64*(4), 306-322.

Rebecca Ferguson. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning 4*, (5-6), 304–317.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*(1), 30-41.

VandeWalle, D., & Cummings, L. L. (1997). A test of the influence of goal orientation on the feedback-seeking process. *Journal of applied psychology*, *82*(3), 390.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of educational psychology*, *96*(2), 236.

Yeh, C. Y., Cheng, H. N., Chen, Z. H., Liao, C. C., & Chan, T. W. (2019). Enhancing achievement and interest in mathematics learning through Math-Island. *Research and Practice in Technology Enhanced Learning*, *14*(1), 1-19.

信息技术融合创新课堂教学互动现状研究——基于2023年广东省教育“双融双创”获奖视频的分析

"Research on the Current Status of IT Integration in Innovative Classroom Interaction: Analysis of 2023 Guangdong 'Dual Integration and Innovation' Award-Winning Videos"

徐玲[1]，晏鑫[2]，庄凌[3]
[123]华南师范大学教育信息技术学院
2459469009@qq.com

【摘要】 在教育数字化战略推动下，本研究基于广东省“双融双创”教师信息素养提升活动中的获奖课例，使用TBAS系统对课堂视频进行编码和分析。研究聚焦课堂行为结构、师生互动及技术应用，探讨信息技术在创新课堂中的融合现状，并针对发现的问题如学生提问不活跃、技术与教学融合不足等，提出改进建议，旨在提升教师数字素养，促进教学中人与技术的深度融合。
【关键词】 信息技术；课堂教学；互动现状研究

Abstract: *Driven by the educational digitalization strategy, this research leverages award-winning instructional cases from Guangdong's "Dual Integration and Innovation" initiative to elevate teacher information literacy. Utilizing the TBAS system, it analyzes classroom videos to assess the integration of information technology within innovative pedagogical frameworks, focusing on classroom dynamics, teacher-student interactions, and technological application. Addressing issues like limited student inquiries and suboptimal technology-teaching integration, the study offers targeted recommendations for optimizing instructional practices. This is aimed at advancing teacher digital proficiency and fostering a more profound integration of technology in education.*
Keywords: *Information Technology; Classroom Teaching; Interactive Status Study*

1. 问题的提出

教育数字化战略为教育改革带来新机遇与挑战。自2022年初教育部提出相关行动，党的二十大及后续会议强调推进教育数字化。为响应国家号召，建设高质量数字化教育体系，培养适应数字时代的优秀教师，广东省自2023年起，将围绕“融合创新”与“优质均衡”主题，持续开展“双融双创”教师信息素养提升活动，旨在支持教育强省建设，推动教育现代化。

当前，教师信息技术融合能力提升，但课堂教学互动的现状及师生行为变化仍需深入研究。本文旨在分析信息技术在课堂教学中的应用现状，识别实践中的不足，并提出改进建议，以促进教师有效融合技术，提高数字素养，实现教学的数字化和信息化。

2. 研究设计

2.1.数据来源

本研究基于广东省“双融双创”活动中的一等奖课例视频进行分析。活动共收集1494件作品，覆盖全省21个地级市、187个县区的1044所学校，参与人数达2621人，人次为2741。

获奖作品 559 件。样本选择注重课例的学科多样性，确保研究结果的广泛适用性和代表性。

2.2 数据编码

2.2.1.编码维度

弗兰德斯互动分析系统（FIAS）主要针对传统教学环境，未涵盖技术行为编码。鉴于数字化创新课堂与此环境在教学互动行为上的差异，本研究采用国内学者穆肃（2015）提出的“基于教学活动的课堂教学行为分析系统”（TBAS），并对其进行适当的调整与优化，形成了适用于“人技融合教学活动”的分析系统。“人技融合”教学活动指在教学过程中，将教师、学生等人力资源与信息技术、教学技术等技术要素有效结合，旨在提升教学的效率与成效。

本方法从教与学行为的发出者和行为的具体内容两个层面对信息化环境教学中的行为进行分类，用于归类信息技术和设备融合应用到教师活动、学生活动中的情况，形成了对教师活动、学生活动和无教学意义活动三大类活动的详细描述，并赋予唯一并统一编码。

2.2.2 编码过程

在教学分析中，通过量表等工具收集的量化数据有限，需统计分析以得出结论。TBAS 方法通过教学行为分类，采用矩阵和变量分析。首先，观察者对课堂教学视频按某一频率进行取样，然后对每一个样本按教学行为类目表判定其行为类型，并用对应代码进行编码标记，即形成最初的原始数据表。剔除视频前后非课堂教学部分，再以每 30 秒为一个时间跨度分割片段，得到每个课例视频的时长与总片段数如表 1 所示。例如，表 2 是对课例 3 教学过程按 30 秒频率抽样并标记所形成的数据。数据表每一行 8 个数据单元记录下了 4 分钟内的 8 个样本的行为编码，表中共 10 行 79 个数据表示记录了 39 分钟课堂教学的数据。

表 1 课例时长以及打点次数

编号	学科	时长	打点次数
课例 1	语文	44 分 32 秒	89
课例 2	数学	43 分 38 秒	87
课例 3	英语	39 分 30 秒	79
课例 4	地理	39 分 11 秒	78
课例 5	物理	38 分 05 秒	76

表 2 课例 3 课堂行为编码表

打点数据	1	2	3	4	5	6	7	8
1	4	4	3	12	13	4	9	5
2	9	9	5	6	12	4	12	6
3	9	3	13	9	9	12	9	1
4	1	6	9	12	4	9	6	3
5	4	13	13	12	5	4	14	14
6	14	6	14	12	3	4	14	11
7	2	5	12	2	1	12	3	6
8	4	6	6	2	9	13	3	12

3. 研究结果与数据分析

3.1.课堂行为结构分析

表 3 不同课堂行为占整体课堂行为的比率

课堂行为		课例 1	课例 2	课例 3	课例 4	课例 5	整体占比
教师活动	教师活动率	49.44%	44.83%	54.43%	69.23%	53.95%	54.38%
	教师语言率	33.71%	34.48%	31.65%	44.87%	36.84%	36.31%
	教师管理课堂环境率	2.27%	0%	2.33%	5.56%	7.32%	3.5%
	教师提问率	20%	16.67%	12%	20%	14.29%	16.59%
	教师反馈率	36.67%	30%	20%	28.57%	25%	28.05%
	教师媒体操作率	14.61%	10.34%	21.52%	20.51%	13.16%	16.028%
	教师无意义活动率	1.12%	0%	0%	0%	0%	0.22%
	教师语言间接影响与直接影响比例	130.77%	87.5%	47.06%	94.44%	64.71%	84.90%

学生活动	学生活动率	49.44%	55.17%	45.57%	30.77%	46.05%	45.4%
	学生语言率	32.58%	33.33%	15.19%	19.23%	17.11%	23.49%
	学生主动提问率	0%	0%	0%	0%	0%	0%
	学生讨论率	4.55%	10.42%	2.78%	12.5%	8.57%	7.76%
	学生使用技术率	3.37%	10.34%	6.33%	7.69%	22.37%	10.02%
技术	课堂技术使用率	29.21%	26.44%	34.18%	32.05%	35.53%	31.48%
无意义教学活动	课堂无意义活动率	1.12%	0%	0%	0%	0%	0.22%

表 3 显示，课堂言语行为占 59.8%，媒体使用占 31.48%，说明教师有意识地将技术融入教学。罗杰斯认为，常规课堂中，“不论分析单位是什么，教师在言语活动中的数量，要远远比学生活跃”。贝莱克语言常模表明，常规课堂中教师与学生语言常模之比为 3.4:1。理论上，数字化创新课堂应该更易激发学生学习的主动性与主体性，学生各类行为占比应大于常规课堂。本研究中教师言语（36.31%）与学生言语（23.49%）之比约为 1.5:1，与贝莱克常模有较明显的偏离，表明教师在教学过程中鼓励学生参与，强化学生主体性。无意义教学活动占 0.22%，低于弗兰德斯的 11%或 12%常模。总体而言，教师对课堂整体掌控较好，未出现技术操作不当引发严重教学停顿现象，课堂中的无效时间少、利用率高。

3.2.师生互动行为分析

师生互动行为可通过教师语言率、教师语言间接影响与直接影响比例、教师提问反馈率、学生语言率、学生主动提问率和学生讨论率进行分析。其中教师语言间接影响与直接影响比例为 84.9%，比值小于 1，说明这些样本课中教师倾向于对课堂和学生作直接的语言影响，即教师更多的是采用讲授或指示行为开展教学。本研究的教师言语率为 36.31%，学生言语率 23.49%，二者相差不大，说明师生有一些互动，基本是以问答的形式进行的。

3.3 技术在课堂中的应用

样本课中都使用了各种信息技术媒体，但每节课使用的技术手段各不相同。但可以看出课堂媒体使用率高达 31.48%，说明老师们都注意把技术融入课堂。技术的使用率相对而言很高，说明教师重视信息技术的使用。

4. 研究结论与建议

4.1.研究结论

通过上述讨论，本文总结出当前信息技术融合创新课堂的教学互动形式，如下图 1 所示。

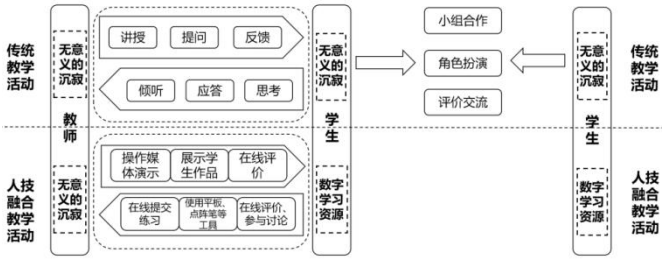


图 1 信息技术融合创新课堂的教学互动形式

4.1.1. 师生交流较为密切、融洽，提升了学生主体地位，师生互动以教师为主导，形式以师生问答为主，学生提问及互动意识有待加强。

在数字化创新课堂中，教师使用提问等互动手段提高学生参与和课堂效率，体现学生中心的教学。学生通过互动学习促进合作和知识构建，但学生主动提问不常见。新课程标准强调

培养学生在语言实践中的问题发现能力、探究意识及创新解决问题的表达途径。尽管部分教师致力于推动学生自主学习，焦点仍集中在答案的产出而非批判性思维的培育。学生在自主发现问题和提问方面仍显不足。

4.1.2．存在一定技术互动行为，但互动形式单一，“人技融合”不够深入。

在课堂中，教师应用数字化工具和智能技术，如网络学习空间和多数据可视化分析，优化学情诊断和资源推送，转变传统课堂结构，提升学生具身认知和教学效果。然而，人机协同教学的深度应用尚待加强，学生主要在教师指导下通过平板等设备学习。

4.2.优化建议

针对存在的问题，优化之后的信息技术融合创新课堂的教学互动形式如图 2 所示。

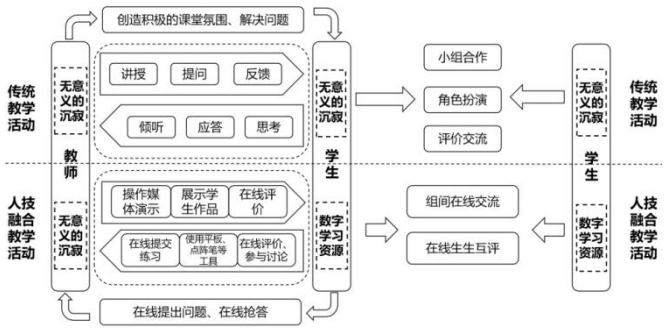


图 2 优化后信息技术融合创新课堂的教学互动形式

4.2.1.重视培养学生发现问题的意识。

培养学生的问题发现意识是教育的关键目标。课堂观察显示，学生多回应教师提问，鲜少自发提问。为促进学生参与，教师应创造积极学习氛围，并运用信息技术工具如在线讨论与实时反馈系统，提高互动机会，使学习更富吸引力，实现教育的人技融合。

4.2.2 深入挖掘人机协同教学的交互形式。

不断开发应用新的数字化工具和智能技术。如问卷工具、互动教学系统、多数据融合可视化分析工具、教育机器人或教学助手、教育 APP 等，引导学生主动参与形式多样的人机协同教学交互活动，从单一的知识传授转变为支持学生的自主、自适应学习及自我管理与评价，培养学生的自我认知和提升能力，从而提高教学质量和信息素养。

5. 结语

本文以广东省“双融双创”一等奖课例为案例，应用人技融合教学活动分析系统，综合分析课堂行为、师生互动及技术应用，总结创新课堂互动现状并提出优化建议。通过分析获奖课例，旨在促进教师信息化教学共同体的发展，提升教师的学科和数字素养。本文侧重课堂行为观察，未来研究将结合访谈和问卷等方法，以增强分析的深度和有效性。

参考文献

穆肃和左萍萍（2015）。信息化教学环境下课堂教学行为分析方法的研究。**电化教育研究**，36(09):62-69。

何文涛、杨开城和张慧慧（2017）。智慧教室环境下协作学习的行为构成及其特征分析。**电化教育研究**，38(11):87-95。

高巍（2009）。课堂教学师生言语行为互动研究。**教育研究与实验**，(05):43-49。

张屹、祝园、白清玉、李晓艳和朱映辉（2016）。智慧教室环境下小学数学课堂教学互动行为特征研究。**中国电化教育**，(06):43-48+64。

探索結合目標設定與視覺化學習分析之學習管理系統於 Python 教育之研究

Exploring the Integration of Goal Setting and Visual Learning Analytics in a Learning Management System for Python Education

柯炘德 [1]，洪暉鈞 [2*]
[1, 2*]台灣中央大學網路學習科技研究所
[1] rtre1249@gmail.com, [2*] hch@cl.ncu.edu.tw

【摘要】本研究開發了一個針對 Python 程式設計學習的線上學習管理系統及學習分析儀表板，用來輔助大學生提升程式學習成效。系統整合了目標設定、練習系統、測驗系統以及兩大學習分析儀表板，目的在增強學生的自我調節學習和自我導向學習能力。透過練習系統能夠不斷複習自身較弱觀念，並且每週使用測驗系統檢驗學習者自身學習成效。儀表板提供豐富的視覺化數據，如雷達圖、甜甜圈圖、日曆圖等，幫助學生理解自己的學習狀況，並根據系統回饋調整學習策略。此外，目標設定功能讓學生能夠根據自身的學習狀況設定和調整學習目標，進一步促進自我導向學習。

【關鍵字】程式設計學習；自我調節學習；適性化學習；學習分析儀表板

Abstract: *This study developed an online learning management system and learning analytics dashboard for Python programming learning, aimed at assisting university students in enhancing their programming learning effectiveness. The system integrates goal setting, practice system, test system, and two major learning analytics dashboards, with the purpose of strengthening students' self-regulated learning and self-directed learning abilities. Through the practice system, students can continuously review concepts they are less familiar with, and each week, they can use the test system to evaluate their own learning outcomes. The dashboard provides rich visual data, such as radar charts, donut charts, and calendar views, to help students understand their learning status and adjust their learning strategies based on the system's feedback. Furthermore, the goal setting feature allows students to set and adjust their learning objectives based on their own learning status, further promoting self-directed learning.*

Keywords: Programming Learning, Self-Regulated Learning, Adaptive Learning, Learning Analytics Dashboard

1. 前言

在學習程式語言時，由於程式語言涉及大量的概念、語法和實際應用，學習者常常面臨快速遺忘所學內容的挑戰，需要不斷地練習和複習來鞏固記憶，減少遺忘。因此，如何有效的減少因為遺忘導致的學習成效低落，提升學習者的記憶保留，對於程式語言學習尤為重要。在這之中，自我調節學習能力扮演了關鍵的角色。本研究建立一個學習系統並結合學習分析儀表板，以輔助學習者進行學習。儀表板將透過學習系統蒐集學習者的學習狀況，包括學習者的學習時間、學習成效等數據，並透過數據分析的方式，適性化的提供符合學習者需求的學習資源。此外，系統還提供自我調節學習工具，幫助學習者設定學習目標、規劃學習策略、監控學習進度和調整學習行為。透過這種方式，學習者不僅可以了解自己的學習狀況，還可以主動調整學習策略，提高學習效率，減少遺忘現象的發生。

2. 文獻探討

2-1 程式設計學習

程式設計指的是為了完成任務或解決問題而進行的設計、編寫、測試和調試電腦指令的過程。Kinnunen 和 Malmi (2006)認為，時間管理和學習動機是影響學習者程式學習是否成功

的關鍵因素。研究顯示在進行程式語言學習時，使用自我調節學習能力的學習者，會比沒有使用自我調節學習能力的學習者，有更高的學習成效(Bergin et al., 2005; Makhno et al., 2022)。因此本研究開發一個線上程式學習系統，來輔助學習者進行 Python 程式語言學習，學習者透過一次次的作答，不斷複習自身較弱的知識概念，並且結合觀看學習分析儀表板，協助學習者進行自我調節學習，以達到良好的學習成效。

2-2 自我調節學習

學習者為了達成學習目標，所採取的行動即是自我調節學習。透過設定目標、監控進展、並使用有效的策略來達成目標，學習者能夠在短時間內掌握學習內容(Zimmerman, 2013)。Lai 和 Hwang (2016)認為將自我調節策略整合到學習中，能夠改善學生的自我效能、自我規劃學習時間的策略，有效地進行學習以獲得更好的學習成效。研究顯示，具有良好自我調節學習能力的學習者，能夠更有效的在線上學習環境中進行自主學習，並提高學習成效(Shea & Bidjerano, 2010; Stephen et al., 2020)。因此透過使用學習分析儀表板來提高學習者的自我意識，促進學習者的自我調節學習是很重要的。本研究基於自我調節學習理論，建立一個學習分析儀表板，並且透過讓學習者自行設定學習目標，並且隨時都能進行目標修改的機制，來促進學習者的自我調節。

2-3 學習分析儀表板

儀表板由圖表、圖形、指標、警報機制等元素組合而成(Podgorelec & Kuhar, 2011)。研究發現，在課程中使用儀表板的學習者，與不使用的學生相比，有較高的保留率(Verbert et al., 2013)。透過儀表板的方式將學習狀況呈現給學生，是作為支持學生積極性和自我調節過程的有效工具(Coffrin et al., 2014; Govaerts et al., 2012)，賦予學生深入了解他們的學習行為來理解他們的學習，提升後設認知水準並促進認知、行為或情感能力的變化，使學生產生自我反思(Jivet et al., 2017)。因此，本研究蒐集學習者使用本系統的作答行為，並結合學習者的學習目標，製作成兩種不同學習分析儀表板，分別對應自我調節學習分析儀表板及共同調節學習分析儀表板，學習者能夠閱讀這兩種不同儀表板產生自我反思、調整學習策略。

3. 系統設計與實作

3-1 Python 線上學習管理系統

本研究開發一線上學習管理系統，輔助學生進行 Python 程式語言學習，包含目標設定、練習系統、測驗系統及兩大學習分析儀表板，系統架構如圖 1 所示。學生可於課堂結束後使用本系統進行 Python 題目練習，且研究規劃於每週上課時使用測驗系統進行 Python 能力測驗，在測驗及練習結束後學生可自行決定是否瀏覽儀表板。為探討學生自我導向學習、自我調節學習能力，因此設計目標設定功能，學生在實驗一開始會先設定學習目標，之後學生可根據學習狀況隨時調整學習目標。學生能透過本學習系統中測驗系統的測驗清單頁面，清楚掌握是否有完成測驗、每次測驗的分數、測驗作答時間以及完成時間。透過學習系統中練習系統的練習清單頁面，學生能清楚掌握每次練習時的分數以及練習完成時間。

3-2 學習分析儀表板

本系統設計兩個基於不同調節理論的學習分析儀表板，分別是自我調節學習儀表板、共同調節學習儀表板：

自我調節學習儀表板包含了，(1)個人學習資訊，透過文字顯示學生作答題數、平均作答時間、作答準確率；(2)Python 能力雷達圖，透過雷達圖顯示學生自身 Python 能力程度，透過點擊各 Python 能力，系統會根據該 Python 能力的程度，顯示出適性化的學習建議；(3)日曆，依不同顏色顯示出學生每天作答資訊；(4)目標分數與後測成績預測，顯示系統預測的學生後測成績與自身的目標分數；(5)每週練習狀況圖，顯示學生先前目標設定時設定的每天作

答頻率、每天作答題數與實際的差距；(6)知識點能力圖，顯示學生各知識點的學習狀況；(7)目標準確率與實際每週準確率，顯示學生每週實際的作答準確率與目標準確率的差距。

共同調節學習儀表板包含了，(1)小組學習資訊，顯示小組內整體作答題數、平均作答時間、作答準確率；(2)Python 能力雷達圖(個人 vs 群體)，透過 2 個重疊的雷達圖顯示個人與群體的 Python 能力程度；(3)組內成員學習狀況，透過盒鬚圖顯示小組內各成員作答題數、平均作答時間、作答準確率；(4)學習歷程記錄，透過折線圖顯示所有學生每天的學習歷程，點擊上方按鈕能夠切換作答題數、平均作答時間、作答準確率的整體歷程；(5)學習排行榜，包含小組學習排行榜及個人學習排行榜，兩種排行榜都包含 3 個子排行榜，分別是 Speedy，顯示答題速度排行榜、Fierce，顯示答題數排行榜、Sharp，顯示答題準確率排行榜。

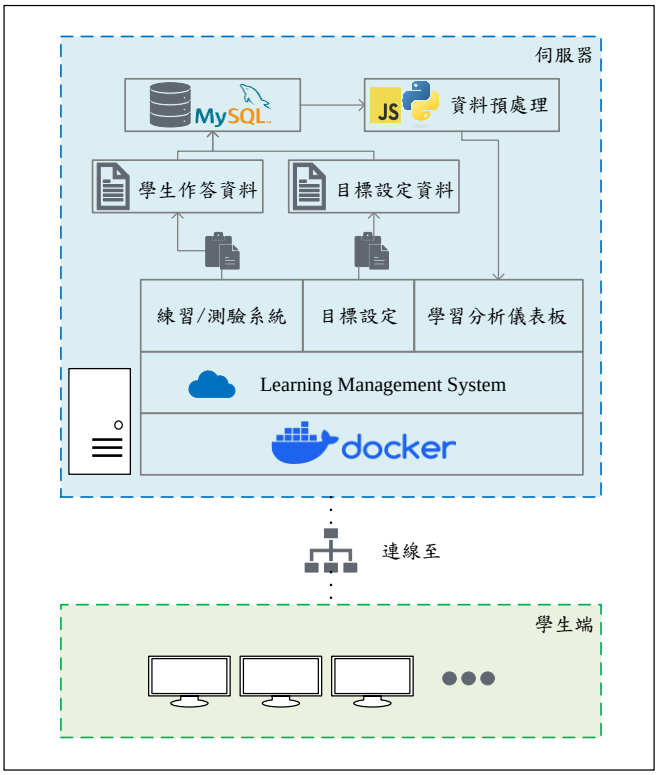


圖 1 系統架構

致謝

本研究感謝國科會專題研究計畫（計畫編號: NSTC112-2628-H-008-001-）與教育部教學實踐研究計畫（計畫編號：PSK1122995）以及中央大學學習科技研究中心支持。

參考文獻

Bergin, S., Reilly, R., & Traynor, D. (2005). Examining the role of self-regulated learning on introductory programming performance. In *Proceedings of the first international workshop on Computing education research* (pp. 81-86). Publication History. https://doi.org/https://doi.org/10.1145/1089786.1089794

Coffrin, C., Corrin, L., De Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 83-92). https://doi.org/https://doi.org/10.1145/2567574.2567586

Govaerts, S., Verbert, K., Duval, E., & Pardo, A. (2012). The student activity meter for awareness and self-reflection. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 869-884). https://doi.org/10.1145/2212776.2212860

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12* (pp. 82-96). Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5_7

Kinnunen, P., & Malmi, L. (2006, September). Why students drop out CS1 course? Proceedings of the second international workshop on Computing education research, Canterbury United Kingdom.

Lai, C.-L., & Hwang, G.-J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, *100*, 126-140. https://doi.org/https://doi.org/10.1016/j.compedu.2016.05.006

Makhno, K., Kireeva, N., & Shurygin, V. (2022). The impact of online learning technology on self-regulation and student success. *Research in Learning Technology*, *30*. https://doi.org/https://doi.org/10.25304/rlt.v30.2802

Podgorelec, V., & Kuhar, S. (2011). Taking advantage of education data: Advanced data analysis and reporting in virtual learning environments. *Elektronika ir Elektrotechnika*, *114*(8), 111-116. https://doi.org/https://doi.org/10.5755/j01.eee.114.8.708

Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers & Education*, *55*(4), 1721-1731. https://doi.org/https://doi.org/10.1016/j.compedu.2010.07.017

Stephen, J. S., Rockinson-Szapkiw, A. J., & Dubay, C. (2020). Persistence model of non-traditional online learners: Self-efficacy, self-regulation, and self-direction. *American Journal of Distance Education*, *34*(4), 306-321. https://doi.org/https://doi.org/10.1080/08923647.2020.1745619

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*(10), 1500-1509. https://doi.org/https://doi.org/10.1177/0002764213479363

Zimmerman, B. J. (2013). Theories of self-regulated learning and academic achievement: An overview and analysis. *Self-regulated learning and academic achievement*, 1-36. https://doi.org/https://doi.org/10.1080/00461520.2013.794676

项目学习中协作问题解决能力的教学策略介入——基于 ENA 的循证分析

Understanding Collaborative Problem-solving Ability in Project-based Learning: Based on Epistemic Network Analysis

熊炀璐[1*]，金晨[1]，邢倩雯[1]，许玮[1]
[1]浙江工业大学教育科学与技术学院
[*] 1487041493@qq.com

【摘要】 本研究采用认知网络分析法（Epistemic Network Analysis，ENA）进行循证分析，研究学习者在不同组别方面的协作问题解决能力分析比较在为期 16 周的学期中，共有 60 名师范生参与了教学微视频制作的在线协作过程，收集师范生的微信会话数据，并且基于协作问题解决评估模型对数据进行编码。结果表明，高等学业水平组更加倾向于努力保持协商谈判想法，中等 学业水平组会进行更加充分的分享资源和想法，而低学业水平组则会对保持积极的沟通予以更多的关注。此外，讨论了本研究的意义、局限性以及未来方向。
【关键词】 项目学习；认知网络分析；协作问题解决

Abstract: *This study employs Epistemic Network Analysis (ENA) for analysis. We investigates and compares the collaborative problem-solving abilities of learners in different groups throughout a 16-week semester. A total of 60 student teachers participated in the online collaborative process of creating educational micro-videos. We collected WeChat conversation data from them and coded the data based on a collaborative problem-solving assessment model. The results indicated that the high academic performance groups tended to make greater efforts to maintain negotiating ideas. The medium academic performance groups engaged in more extensive sharing resources and ideas, while the low academic performance groups placed more emphasis on maintaining positive communications. Additionally, the significance, limitations, and future directions of this study were discussed.*
Keywords: project-based learning, epistemic network analysis, collaborative problem solving

项目基金：浙江省属高校基本科研业务费专项资金(项目编号：GB202302007)；浙江工业大学大学生创新创业训练计划创新训练项目 (项目编号：GZ23431080006)。

1. 前言

联合国教科文组织把“学会合作”作为 21 世纪教育的四大支柱之一，《中国教育现代化 2035》也指出“发展中国特色世界先进水平的优质教育”，要强化“合作能力”的培养，并推行“合作式教学方式”。协作问题解决能力是 21 世纪的一项关键能力，作为创新教学模式的 PBL 项目式学习也与协作问题解决能力密切相关，随着信息化时代的发展和进步，社会对于高质量人才的培养有了更高的要求，创新教育迫在眉睫，成为教育界关注的主题之一。

随着 PBL 项目学习在教育领域的广泛应用，对于其重要表现之一的协作问题解决能力引起广泛关注。基于特定项目展开的表征学生协作问题解决能力的项目组间言语信息交流，有助于学生进行协作项目学习。本研究采用认知网络分析法（Epistemic Network Analysis ，ENA），根据 PBL 项目学习中协作问题解决实际，在整体协作学习方面，高中低学业水平学习小组之间协作问题解决模式之间的差异。探寻适合 PBL 项目教学过程中通过提高协作问题解决能力的进行学业水平提升的教学策略及建议。

2. 理论基础

2.1. 协作问题解决能力

协作问题解决（Collaborative Problem Solving，简称 CPS）能力被视为 21 世纪人类生活、工作和学习必备的高阶能力。（郑旭东等人，2021）OECD 在 PISA2015 中的定义被认为目前关于协作问题解决能力的公认概念："个体能够有效地参与问题解决过程的能力，在这一过程中，两个及两个以上主体需要共享彼此的理解和努力，来形成解决方案，并联合他们所具有的知识、技能和努力，来达成解决方案"。现有研究已经形成了诸多关于在线协作问题解决的理论框架及模型，PISA2015 将协作过程的三个核心构成要素与个人问题解决的四个认知过程交叉，形成了 12 个可观测的协作问题解决技能，而 ATC21S 将协作问题解决能力分为社会技能和认知技能两个部分最终将协作问题解决能力细化为 18 种元素，基于此构建了学生协作问题解决能力评估框架。（朱莎等人，2023）这两种评估颇具代表性，但是更多关注个人在协作中的表现而不是团队的表现。研究表明项目学习、STEAM 学习可以用来培养学生的 CPS 技能（Tassaneewon, 2020），在线协作讨论环境也能够提高 CPS 能力（Lin, Hou & Chang, 2022）（Wu, 2020）。关于认知技能和社交技能如何相互作用以及协作学习如何影响学习成绩的研究仍然缺乏。

2.2. 基于会话的认知网络分析

ENA（Epistemic Network Analysis）是一种通过分析日志数据和其他个人和协作学习的记录形成动态网络分析模型来表征个体（或群体）认知特征的分析技术（Nash & Shaffer, 2013）（Shaffer & Gee, 2012）。它在认知框架理论的指导下，利用技术与工具对学习者在学习过程中产生的文本数据进行分析与处理，从而建立学习者认知网络模型，直接对学习者的认知网络结构进行分析，以此对学习者的认知发展做出更深刻、更全面的评价。（王志军和杨阳，2019）（孙灵，2023）现有的文献表明，认知网络分析可以用于对协作学习进行广泛的分析，在处理数据交互文本的基础上，能够对学习者的认知特征进行分析，对学习过程进行比较。（Shaffer et al., 2009）（Arastoopour, Chesler & Shaffer, 2014）（丁继红，2019）同时，认知网络分析能够更全面地评价学习者的学习行为，能够更直观地对学习者的认知过程进行分析和了解，从而做出评价。（Sullivan et al., 2018）（Andrist, Ruis & Shaffer, 2018）然而目前有关认知网络分析的研究中以会话为基础的仅占少数，在会话这方面仍有巨大的知识空间和研究深度还未被探索，这正是此次研究努力突破的领域。

本研究考察了师范生的协作问题解决过程，并通过 ENA 的方法确定不同学业水平组之间的续作问题解决过程模式，并提出以下问题：

（1）在整体协作学习活动中，不同学业水平组的 CPS 框架要素分布和 CPS 模式有何差异？

（2）我们是否能够使用 ENA 方法识别不同学业水平组之间 CPS 模式的细微差异吗？

3. 研究设计

3.1. 项目学习设计

本研究的实验对象为某高校选修 2022 学年第二学期《教学微视频设计与开发》课程的 60 名大二学生，其中包含教育技术学（师范）40 人、计算机科学与技术（师范）16 人、应用心理学 4 人，学习过基本的教育学知识以及通用信息技术工具的使用，于先前的专业学习中拥有过协作学习经验，自由组成 17 支人数不大于 4 人的队伍。运用认知网络分析手段，编码分析不同学习小组在微信群中的会话信息（图 1），比较不同学业水平学习小组的认知网络显著特征，探究其在整体协作学习方面，高学业水平、中等学业水平与低学业水平学习小组之间协作问题解决模式之间的差异。教师在期末评价中给予各学生小组微视频项目作品评

价打分，予以反馈与优化总结，各学生最终成绩的成绩由 50%教师评价、30%组间评价以及 20%组内评加构成，最终成绩在 0-100 之间，各组作品提交、下载、交流、评分以及队伍组建管理等相关操作于在线学习平台学习通中完成（图 1）。

参考 Hao 等人（2017）的编码方案，本研究协作问题解决标准化评估的步骤为:根据 PBL 项目学习中协作问题解决实际，将微信群会话数据分为分享资源和想法、保持积极的交流沟通、规范调节问题解决以及谈判协商想法的四个维度，如表 1 显示编码空间四个维度的定义，表 2 表示编码方案中 33 个子类别及其根据课程内容研究背景添加的编码举例，用来分析学生的会话数据，探索队伍成员之间协作问题解决的能力。



图 1 微信群会话界面（组 1）、学习通项目学习支持界面（组 1）

表 1 CPS 四维度定义

CPS 维度	定义
分享资源和想法	分享与任务相关的包括视频、图片、链接、文件等资源，以及对任务完成状况的描述想法。
保持积极的交流沟通	以社会为导向的沟通，可能与项目有关，也可能与项目无关，包括事实陈述、无观点意志的疑问、学生对于队友的反应、毫无意义的打断、询问他人意见或表情符号等。
规范调节问题解决	设定目标，询问与监控进度，对团队绩效进行评价与反思。
谈判协商想法	表达自我的观点，评价自我与他人的观点，通过比较不同成员观点之间的异同，展示自己对于帮助团队构建协作性知识的理性。

表 2 编码方案子类别及其相应举例

CPS 维度	学生表现子维度
分享资源和想法	1. 学生为团队提供微视频制作相关的信息资源。例如：老师说过要求制作的微视频时长在 3-8 分钟。
	2. 学生为团队提供微视频制作过程中的视频、图片、链接、文件等资源或其检索线索。
	3. 学生对于自己任务完成状况的描述。例如：我已经做完了第一部分的视频。
保持积极的交流沟通	1. 学生通过文本和文本符号回答不涉及自我观点的事实性问题。
	2. 学生通过文本、文本符号、图像内容使用适当的社交语言设法使团队间会话生动起来。
	3. 学生毫无理由地拒绝队员的建议。
	4. 学生通过适合的社交语言用来进行问候等普通社会性交流。
	5. 学生提供帮助。
	6. 学生为无意的打扰道歉。

	7.学生发表了无意义的会话内容。
	8.学生对于队员如困惑、沮丧的情绪表示理解。
	9.学生对于队员的陈述作出回应。
规范调节问题解决	1.学生确定了任务目标。
	2.学生表示需要进行下一步操作。
	3.学生因为缺乏理解而产生困惑与沮丧。
	4.学生在理解方面取得进步。
	5.学生反思团队所取得的成就。
	6.学生表示在解决问题过程中所缺少的东西。
	7.学生对自我和队员理解的检查。
	8.学生评估队员的贡献是否有助于解决问题。
	9.学生对于队员或团队的表现表示赞赏。
	10.学生指出团队在学习过程中的不足之处。
	11.学生提出在学习过程中发现的问题。
谈判协商想法	1.学生对队友的观点表示赞同。
	2.学生表达与队友不同的观点。
	2.学生不确定是否赞同。
	3.学生要求队员回应自我的观点与建议。
	4.学生要求队员澄清一个说法。例如：你认为我们应该使用什么软件来进行视频的剪辑？
	5.学生重述队员的陈述。
	6.学生发现自己与队员的想法有冲突。
	7.学生用相关证据指出队友陈述中的相关不足。
	8.学生对他们的观点与建议进行论证与讲解。
	9.学生通过队友的陈述完成了观点的转变。

3.2. 数据整理与分析

在进行内容分析前，本研究通过 Excel 共收集了自微视频任务发布至完成期间的所有微信群会话数据，项目学习历时 50 天，其中原 17 组中 4 组会话数据丢失，4 组会话数据因为数量问题不符合要求删去，共收集 9 小组，3221 条微信群会话数据，将 9 个小组随机从 1-9 进行编号，两位研究人员使用表 2 中描述的编码方案对所有的在线会话数据进行了编码，在本研究中将各成员在协作会话中具有完整语义的文本数据将作为最小编码单元。

为了保证编码的可靠性，进行编码的研究人员进行了相关编码培训，并且在编码前对编码标准进行讨论和协商，以解决编码的不一致问题。随后随机抽取其中 20%的会话数据，同时背对背进行编码，Kappa 值为 0.754，通过一致性检验，表明两位研究人员的编码信度良好，可以完成后续编码工作。同时将剩下 80%的会话数据随机平分给两位研究人员进行编码操作。根据学习通在线学习平台上每组所有成员的最终评价总分，计算出每个小组的平均分数，将 1-9 号小组进行排名，分为高中低学业水平组（每组含三个小组）。

对于本研究问题，我们将 Excel 数据进行收集与整理，整体导入认知网络分析在线工具(https://app.epistemicnetwork.org/)，以先前选定的四个维度来对会话建立动态网络模型，以学业水平与队伍编码作为分析单元，节设置为 4，开展认知网络分析。在完成建模过后，比较不同学业水平小组之间的认知网络显著特征，来探究其中协作问题解决能力的模式差异。如图 6 所示为数据收集整理与分析过程的整体过程。

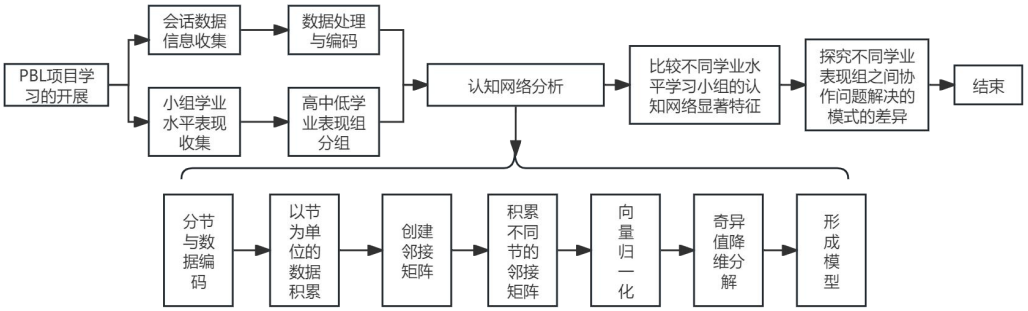


图 6 数据收集整理与分析过程

4. 研究结果与分析

如图 7 所示为高学业水平、中等学业水平、低等学业水平三个小组协作问题解决的 ENA 网络模型，以及三类学业水平小组认知分析网络比较，其中红色为高学业水平组，蓝色为中等学业水平组，紫色为低等学业水平组，不同颜色的方块各自表示不同学业水平表现组别的质心（投影点的平均位置）。ENA 的空间横轴将协作学习过程描述右边为保持积极的交流沟通，左边为分享资源和想法，垂直轴将协作学习过程描述正面为谈判协商想法，背面为规范调节问题解决。如表 8 所示为高中低学业水平组微视频制作项目学习 ENA 网络的连接系数。

表 8 高中低学业水平组 ENA 网络连接系数

连接	高学业水平组	中学业水平组	低学业水平组
分享资源和想法-保持积极的交流沟通	0.45	0.27	0.52
分享资源和想法-规范调节问题解决	0.18	0.58	0.22
分享资源和想法-谈判协商想法	0.47	0.59	0.39
保持积极的交流沟通-规范调节问题解决	0.30	0.20	0.30
保持积极的交流沟通-谈判协商想法	0.59	0.16	0.63
规范调节问题解决-谈判协商想法	0.18	0.18	0.13

根据曼-惠特尼 U 检验（Swiecki et al.，2020）比较了高中低学业水平组之间投影点的分布，发现以下分析结果：①沿着横轴，高学业水平组(Mdn=-1.18，N=3)在 α = 0.05 的水平上与中等学业水平组 (Mdn=2.09，N=3 U=0，p=0.10，r=1.00)有统计学显著差异；②沿着横轴，中等学业水平组(Mdn=2.09，N=3) 在 α = 0.05 的水平上与低学业水平组 (Mdn=-1.31，N=3 U=9.00，p=0.10，r=-1.00)有统计学显著差异。

通过以上认知网络分析结果可以看出高学业水平组在横轴上与中等学业水平组有显著的差异性，高学业水平组在分享资源和想法、保持积极的交流沟通、谈判协商想法方面做得更加出色，这三者之间相互的联系也比较密切，但是在规范调节问题解决方面的会话相对比较薄弱。并且中低两组之间的差别也与上面两组之间的差别类似，低学业水平组与中等学业水平组在横轴上也有显著的差异性，低学业水平组的学生在项目学习过程中更加注重分享资源和想法、保持积极的交流沟通、谈判协商想法三方面的会话，这三方面的会话关联结果也比规范问题解决方面的数据更加显著。相比高低两组学业水平组，中等学业水平组的学生在协作问题解决过程中更加偏向于分享资源和想法、规范调节问题解决、谈判协商想法，这也是导致其与另外两组在横轴上产生显著差异性的原因。

高低两组学业水平组之间也同时具有小量差异，不具有显著性，其中高学业水平组的学生比较注重规范调节问题解决、谈判协商想法，相比之下，低学业水平组的学生更加看重分享资源和想法以及保持积极的交流沟通，在另外两方面有所欠缺。

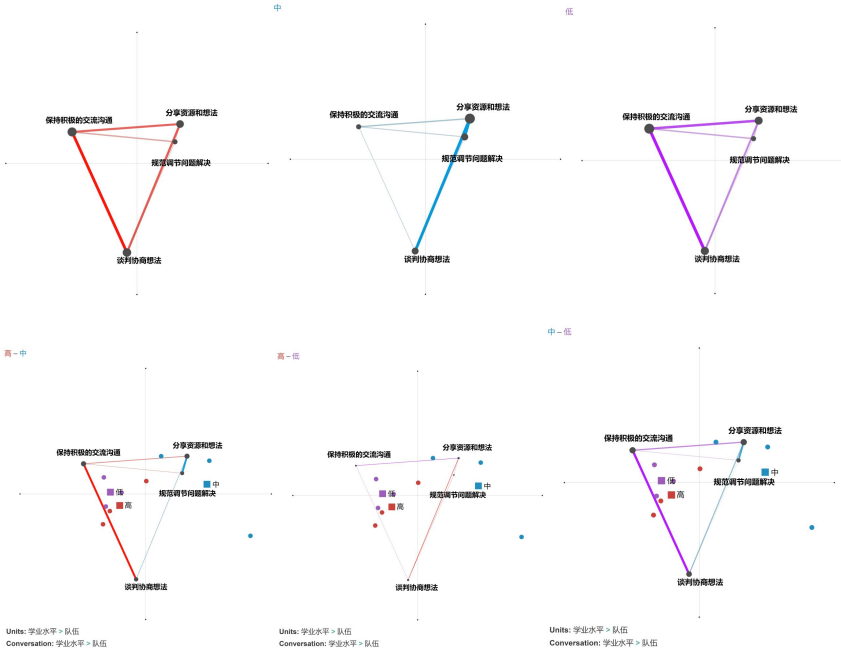


图 7 高中低学业水平组认知分析网络图及比较图

5. 讨论

5.1. 高学业水平组协作问题解决

我们经过 ENA 认知网络的分析了解到高学业水平组相比另外两组更加注重协商谈判想法，他们在协作问题解决过程中会更加偏向于表达自我的观点，评价自我与他人的观点，通过比较不同成员观点之间的异同，展示自己对于帮助团队构建协作性知识的理性，在自我意识方面的觉醒更加的显著。他们通过展示、评价、比较不同的观点，逐渐构成自我完整的认知结构，并且通过会话的方式将这一讯息传递出去，成为团队的认知，传递分享与感染，而通过他们的学业表现水平也可以认为该组别的整体表现是比较突出的，他们逐渐碰撞、比较、完善、精进之后的团队认知对于他们的学业表现产生了推动作用。

此外高学业水平组的学生同时保持一定程度的交流沟通、协商以及资源资源和想法的分享，与中等学业水平组相比，其保持积极的交流沟通部分也比较显著。有研究表明，实现高质量的协作过程和结果是具有挑战性的，因为它既要求学习者有能力提出不同的想法，也要求学习者深刻讨论这些想法并发展这些想法。（Yücel & Usluel, 2016）正是要在不同的想法观点中进行比较提炼，不断地进行积极的交流，这个过程是循环往复、不断上升的过程，这也是为高绩效的产生提供了重要条件。因此，在协作问题解决中教师应该更多地去让学生在保持在协商谈判、资源分享等其它优势的情况下不仅可以形成自己的认知也可以通过队员之间积极有效的交流沟通促进队伍凝聚力以及更好地讨论、形成并发展项目学习队伍的整体认知。

5.2. 中学业水平组协作问题解决

中等学业水平的学生队伍相比其他两组别会更加注重分享资源和想法，他们更加注重单纯地分享与任务相关的包括视频、图片、链接、文件等资源，以及对任务完成状况的描述想法，而忽略了其它的积极的交流沟通、协商谈判想法以及规范调节问题解决方面给项目学习成果带来的显著帮助。

资源和想法的分享是必要的，大量的分享为 PBL 项目学习的作品诞生奠定了深厚的物质与理论基础，但是相比于高学业水平组，中等学业水平组在保持积极的交流沟通、谈判协商的想法方面相对薄弱。研究表明学习共同体是基于社会文化理论和情境理论，由学习者和助学者共同构成，以实现共同目标为愿景，通过互动协商和学习资源分享而相互促进的学习集

体。（Brown & Campione, 1994）只凭借大量资源想法分享并无法形成高水平的学习集体，他们之间缺乏高质量的互动协商，这便可能导致项目学习的成果质量，以及小组协作问题解决的过程质量，没有办法达成相互促进相互提升的境界来成就高质量的学习共同体。所以，中等学业水平组的学生正是缺乏这一份在协作问题解决过程中互动沟通、谈判协商中达到顺变和纳新的能力，将思想和资源的分享扩大到最大化效益，深度与广度并举。可见教师在看到学生小组进行以分享资源和想法为主体而缺少其它的积极交流、沟通、协商、调节的会话时，应该去更好地引导非高绩效的相关学生在资源和想法的分享中大胆自如地尝试谈判、倾听、理解、协商与讨论不同的观点想法，去形成更加完善的作品，达到更理想的绩效结果。

5.3. 低学业水平组协作问题解决

低学业水平组的 ENA 认知网络图和中等学业水平之间也存在着横轴范围上有显著差异的现象，而和高学业水平组具有相似性，这说明高低学业水平组之间存在某些共同点，但是前者明显在各维度之间的分布更加均衡。而低学业水平组在保持积极的交流沟通方面相当显著，却在分享资源和想法、规范调节问题解决方面相对比较薄弱，其协作问题解决过程中的会话更多地表现为以社会为导向的沟通，这很有可能是造成低学业成绩的原因之一。所以，对于学业表现相对比较差的学生小组，教师可以弥补在分享资源和想法方面的不足，将相关的资源想法分享给有需要的学生，但是这也可能会助长相关学生思考与分享方面的惰性，需要教师在教学任务的不同阶段逐步放手，培养学生对于分享资源和想法的自主意识。

此外，规范调节问题解决也是会对绩效起比较重要的结果，设定目标、询问与监控进度、对团队绩效进行评价与反思方面对于项目学习的进行可以有更加深刻清晰的把握，是掌舵的水手。对于低学业水平组学生在这方面相应欠缺的结果下，需要教师引导学生增强目标意识，增强反思评价的数量和质量，并且对于项目的掌握度与把控度应当比较完全地掌握，尤其对于低学业水平组中队长的意识培养在规范调节问题解决方面更是相当重要的。但是低学业水平组与高学业水平组两组别之间并没有显著的差异性，以上对于这两组之间差异的分析更多地是基于 ENA 网络分析图来进行，不具有确定性与肯定性。

6. 结论

本研究通过认知网络分析法对于不同组别的学业水平小组在 PBL 项目教学过程中的协作问题解决会话进行了循证分析，探寻了适合 PBL 项目教学过程中通过提高协作问题解决能力的进行学业水平提升的教学策略及建议。本研究主要基于学生小组队员之间的微信群会话数据信息在 ENA 手段下的认知网络建构，可能由于以下原因导致一定的局限性：（1）数据收集不完全；（2）样本量较小；（3）编码误差；（4）前期分组干涉不完备，没有控制组别的整体实力水平。未来研究将不断完善基于 ENA 技术下协作问题解决过程中数据的细粒度分析，对跨时间、跨阶段、跨专业下的学生认知网络构建进行不断探索挖掘，增强实验的可靠性和有用性，切实实现精准化、及时化教学，探寻更加完善的教学过程策略，全面促进学生认知能力、知识水平、协作能力等各方面的提升与进步。

参考文献

丁继红（2019）。深度学习中的学习者认知网络和动机策略分析--旨向深度学习的 U 型翻转教学效果研究。**远程教育杂志**, 37(6)，32-40。

王志军、杨阳(2019)。认知网络分析法及其应用案例分析。**电化教育研究**, 40(06), 27-34+57。

孙灵（2023）。基于社会网络分析和认知网络分析的协作分组方法研究（硕士论文）。**广州大学**。

杜威（1990）。**民主主义与教育**。北京：人民教育出版社。

威廉·赫德·克伯屈（2016）。**教学方法原理**。北京：人民教育出版社。

郑旭东、马云飞、范小雨（2021）。协作问题解决：人工智能时代必备的高阶能力。**现代教育技术**，31(3)，12-19。

朱莎、李嘉源、郭庆等人（2023）。在线环境下协作问题解决能力评估框架构建及应用——基于数字素养视角的研究。**现代远距离教育**，(05)，42-52。

Andrist, S., Ruis, A. R., & Shaffer, D. W. (2018). A network analytic approach to gaze coordination during a collaborative task. *Computers in Human Behavior*, 89(DEC.), 339-348.

Arastoopour, G., Chester, N. C., & Shaffer, D. W.(2014). Epistemic persistence: a simulation-based approach to increasing participation of women in engineering. *Journal of Women & Minorities in Science & Engineering*, 20(3), 211-234.

Brown, A. L., & Campione, J. C.(1994).*Classroom Lessons:Integrating Cognitive Theory and Classroom Practice*. Cambridge, MA:The MIT Press.

Hao, J., Liu, L., von Davier, A. A., & Kyllonen, P. C. (2017). *Innovative assessment of collaboration*. Cham, CH: Springer.

Lin, P. C., Hou, H. T., & Chang, K. E. (2022). The development of a collaborative problem solving environment that integrates a scaffolding mind tool and simulation-based learning: an analysis of learners' performance and their cognitive process in discussion. *Interactive Learning Environments*, 30(7), 1273–1290.

Nash, P., & Shaffer, D. W. (2013). Epistemic trajectories: mentoring in a game design practicum. *Instructional science*, 41(4), 745-771.

Shaffer, D. W., & Gee, J. P. The right kind of GATE: Computer games and the future of assessment. *In Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research.* Information Age Publishing, Charlotte, NC, 2012.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., & Bagley, E., et al. (2009). Epistemic network analysis: a prototype for 21st-century assessment of learning. *International Journal of Learning & Media*, 1(2), 33-53.

Sullivan, S., Warner-Hillard, C., Eagan, B., et al. (2018). Using epistemic network analysis to identify targets for educational interventions in trauma team communication. *Surgery*, 163(4):938-943.

Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: a network analysis approach. *Computers in Human Behavior*, 104(Mar.), 105876.1-105876.15.

Tassaneewon, L. (2020). Enhancing collaborative problem-solving competencies by using STEM-based learning through the dietary plan lessons. *Journal of Education and Learning*, 9(4), 102–117.

Wu, S.-Y. (2020). Incorporation of collaborative problem solving and cognitive tools to improve higher cognitive processing in online discussion environment. *Journal of Educational Computing Research*, 58(1), 249–272.

Yücel, Ü. A., & Usluel, Y. K. (2016). Knowledge building and the quantity, content and quality of the interaction and participation of students in an online collaborative learning environment. *Computers & Education*, 97, 31-48.

生态学视角下学校疫情停课对学生综合素养的影响：揭示学校、家庭和个人支持的链式中介作用

The Impact of School Closures during the Pandemic on Students' Comprehensive Literacy from an Ecological Perspective: Unveiling the Cascading Mediating Roles of School, Family, and Personal Support

虞增艺[1*]，陈诺言[1]，金晟[1]，徐刘杰[1]
[1]浙江工业大学 教育科学与技术学院
*202105720431@zjut.edu.cn

【摘要】 为探讨学校疫情停课对学生综合素养的影响，本文采用生态学的理论框架，并以 PISA 2022 数据为依托，构建了一个链式中介模型，考察了学校、家庭及个人支持在相似停课时长下的作用。研究表明，学校支持主要通过个人支持间接正向影响学生综合素养，直接影响则不显著。随着停课时长的延长，学校支持的影响逐渐增强。基于这些发现，建议应强化个性化学习支持，提供在线教育资源和心理咨询服务，以适应长期停课下的教育需求。代码见 https://github.com/yuzengyi/Gccce2024。

【关键词】 PISA2022；生态学；综合素养；链式中介

Abstract: *This paper utilizes the theoretical framework of ecology and relies on data from PISA 2022 to construct a chain mediation model. This model examines the roles of school, family, and personal support during periods of similar closure durations. The study reveals that school support mainly has an indirect positive effect on students' comprehensive competencies through personal support, with its direct influence being insignificant. As the duration of school suspensions extends, the impact of school support gradually intensifies. Based on these findings, it is recommended to strengthen personalized learning support, provide online educational resources, and offer psychological counseling services to meet the educational needs during prolonged suspensions. The code is available at https://github.com/yuzengyi/Gccce2024 .*

Keywords: PISA 2022, Ecology, Comprehensive Literacy, Chain Mediation

1. 前言

随着 COVID-19 疫情的全球蔓延，传统的教育体系面临前所未有的挑战。为解决因疫情导致的无法到校上课的问题，全球教育工作者与亿万中小学生共同参与了一场前所未有的大规模在线教育实践。这一转变在确保教育连续性的同时，也引发了对在线教育质量的广泛质疑。在后疫情时代，本文基于 PISA2022 数据，以生态学为视角，汲取全球大规模在线教育应用的经验，惟愿为今后的在线教育提供更好的帮助。

2. 研究设计

2.1. 基于生态学理论构建学校、家庭和个人支持的链式中介模型

学生综合素养的塑造与发展构成了一个多元复杂的进程，受到众多因素的共同作用，这些因素包括但不限于各类环境要素（如家庭、学校环境）及个体的情感反应、自主学习能力的相互交织。鉴于单一因素难以充分阐释综合素养的构成，因而开展全面且系统的研究显得尤为关键。根据生态学理论（Bronfenbrenner & Morris, 1998），个体发展可从一个维度丰富、

结构复杂的环境与个体因素互动的系统角度进行解读。本研究依托于生态学的理论框架，对环境要素进行了三重层次的划分：宏观系统、中介系统和微观系统。学生综合素养的成长不仅与微观系统内的个体因素密切相关，同时也受到中介系统内学生直接接触的因素以及中介系统与宏观环境系统之间互动的影响。特别地，家庭环境作为学生发展的初步舞台，对个体的成长起着至关重要的作用。家庭支持强调家长在教育中的参与，他们协助教师开发孩子的学术潜力和监控教学质量。例如，OECD 在 2017 年强调了与家长合作对教师支持孩子学习和发展的重要性。另外，家长参与对学校治理和孩子学习成果具有积极影响（Marphatia et al., 2010； Desforges & Abouchaar, 2003）。此外，学校支持涉及为学生提供必要的教育资源、师资力量支持以及学生在学校环境中得到的情感和社交支持（Li, J., Hu, Z., & Pan, L. ,2022）。个人支持通常包括个体在心理、情感及社交方面得到的支持，这种支持可能来自个人内部（如自我激励、自我肯定）或外部（如同伴、朋友、导师或其他社会网络）。为此，本文界定了学校支持作为宏观系统变量，家庭支持作为中介系统变量，个人支持作为微观系统变量，并以停课时长作为控制变量进行分组研究。据此，本研究提出假设 H2 和 H3，认为家庭支持和个人支持是学校支持影响学生综合素养的中介变量。考虑到家庭支持与个人支持之间的相关性，本研究进一步提出假设 H4：家庭支持和个人支持在学校支持对学生综合素养影响过程中具有链式中介效应。最终，本文构建了基于生态学理论的学校、家庭和个人支持的链式中介模型（如图 1 所示），通过对假设的验证，全面探讨了疫情期间实施的“停课不停学”在线教育环境下，学生综合素养的形成与发展机制。

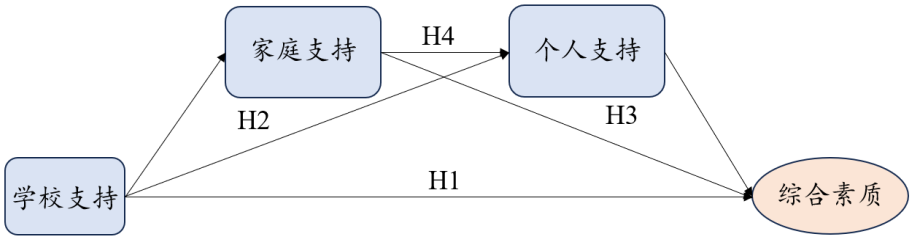


图 1 基于生态学理论构建学校、家庭和个人支持的链式中介模型图

2.2．研究对象选择

本文选择港澳台作为研究对象，其理由源于港澳台在最新的 PISA 测试中展现出引人注目的表现。2022 年 PISA 的结果（OECD, 2023）显示，亚洲学生在全球范围内取得了显著成绩，其中新加坡和中国的港澳台地区表现尤为突出。特别是在中国大陆因疫情错过正式测试的背景下，港澳台地区的学生在阅读、数学和科学三大领域中的表现依然保持在全球前列。这一成绩不仅彰显了港澳台地区基础教育的高水准和卓越成果，也为分析疫情背景下学生综合素养的形成与发展提供了独特视角。港澳台地区的教育体系与中国大陆虽存在差异，但也展现出一定的相似性，这为中国大陆在线教育发展提供了宝贵的借鉴经验。

2.3．数据来源

本研究使用 PISA 2022 学生背景调查问卷收集的数据。在 PISA 2022 中，16148 名中国的港澳台地区的学生参与了调查。其中，香港（HKG）、澳门（MAC）和台湾（TAP）分别有 5907 名、4384 名和 5857 名学生参与调查。关键字段“ST347Q01JA”记录了疫情期间的停课时长，作为本研究的关键控制变量。该字段的取值范围为：1 表示未停课，2 表示停课不足 1 个月，3 表示超过 1 个月但不足 3 个月，4 表示超过 3 个月但不足 6 个月，5 表示 6 至 12 个月，6 表示超过 12 个月。在剔除了 902 个缺失数据后，得到不同地区停课时长的分布情况，如图 2。

图 2 显示，在不同的停课时长类别中，台湾、香港和澳门的学生都有所分布。台湾的学生在各个停课时长类别中的占比相对较小，尤其是在较长的停课时长（如 4、5、6 标签所示）

中，其学生数量明显低于香港和澳门，可能说明台湾的疫情防控和在线教育适应能力较强，或采取了更灵活的教育政策。澳门的学生在 1 至 6 个月的停课时长类别中数量最多，尤其是在 3 至 4 个月的停课时长（4 标签）中数量最多，这可能反映了澳门在疫情期间采取了较长时间的严格学校关闭措施。相比之下，香港的学生在 1 至 3 个月的停课时长（2 和 3 标签）中较多，而在超过 6 个月的停课时长（5 和 6 标签）中较少，表明香港可能采取了阶段性关闭策略，努力缩短停课时间或较快恢复线下教学。

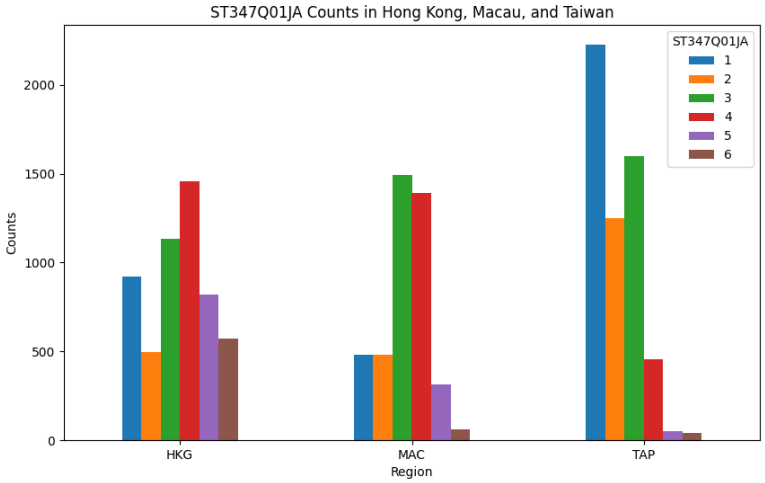


图 2 港澳台不同地区停课情况分布图

2.4. 数据预处理

在 PISA 2022 分析框架中，本文筛选出所有停课的学生，并通过因子分析在“ST347”到“ST354”的 46 个字段中筛选出主成分，并删除主成分中出现缺失的数据，并通过了测量模型的信效度检验，具体字段解释与分类如表 1 所示。

表 1 字段解释与分类表

变量分类	变量名	变量解释
控制变量	ClsD	因 COVID-19 而停课的持续时间
学校支持	MatD	在 COVID-19 期间，学校多久发一次学习材料让你自学？
	AsnF	在 COVID-19 期间，学校多久布置一次作业？
家庭支持	HwCk	在 COVID-19 期间，家庭成员多久检查你一次是否完成学校作业？
	ExpN	在 COVID-19 期间，家庭成员多久向你解释一次新内容？
个人支持	TxtU	在 COVID-19 期间，你多久使用一次数字教科书、练习册或工作纸？
	LClF	在 COVID-19 期间，你多久参加一次由学校老师通过视频通信程序（如 ZoomTM, SkypeTM, Google® Meet™, Microsoft® Teams）实时授课的课程？
	AnxY	在 COVID-19 期间，你是否因学校工作感到焦虑？
	Motv	在 COVID-19 期间，你是否有学习的动力？
综合素养	MATH	数学素养
	READ	阅读素养
	SCIE	科学素养

其中，学校、家庭、个人支持量表均为四分量表。由表 1 可知，学校支持指的是学校在疫情期间提供的教育资源和支持，这包括但不限于远程教学的材料分发频率、作业布置频率

等。在生态系统理论的框架下，学校支持作为一种外部环境因素，对学生的学习和综合素养发展具有重要影响。同样地，家庭支持涉及到家庭成员在停课期间对学生的教育参与和支持，比如检查作业的频率和解释新内容的频率。个人支持包括学生在停课期间的自我管理能力，如使用数字教科书的频率、参加在线实时授课的频率、学习动机和学习相关的情绪状态（如焦虑）。此外，本文对数学素养、阅读素养和科学素养分别进行 KMEANS 聚类，将一维连续变量，聚成 4 类，如图 3 所示。

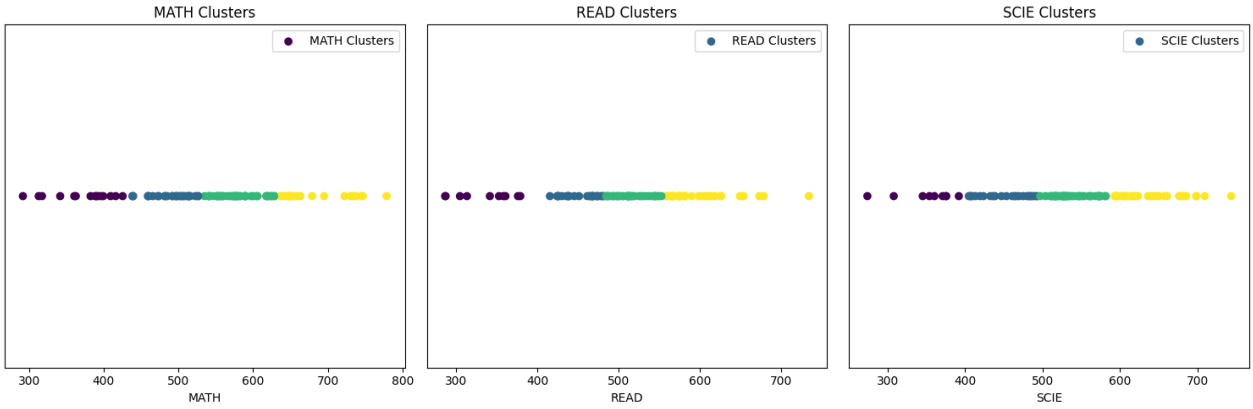


图 3 数学素养、阅读素养和科学素养聚类图

为便于应用 AMOS 软件求解链式中介模型，本文采用线性加权平均法，对各子变量进行整合，从而构建出学校支持、家庭支持、个人支持和综合素养这四个综合变量。这一方法有效地将潜在变量转换为显在变量，为后续的路径分析奠定了基础。

3. 研究结果

本研究根据停课时长这一控制变量，将数据划分为三个不同组别。其中，“ST347Q01JA”字段的值为 2、3、4 的数据被归类到停课 0 至 6 个月的组别，值为 5 的数据划入 6 至 12 个月的组别，而值为 6 的则分配到超过 12 个月的组别。为了检验中介效应，本文进一步采用 Bootstrap 抽样方法进行统计验证（Preacher & Hayes, 2004）。

3.1. 停课半年以内组的学校支持完全通过个人支持中介影响学生综合素养

在 PISA 2022 数据分析的背景下，本研究深入探讨了疫情导致的停课对学生综合素养影响的内在机制。

表 2 停课半年以内组的综合素养影响机制的中介效应分析

路径	点估计	校正后偏差		百分位数		效应比
		下限	上限	下限	上限	
学校支持-家庭支持-综合素养	0.001	-0.023	0.048	-0.035	0.032	0.38%
学校支持-个人支持-综合素养	0.112	0.026	0.254	0.021	0.246	42.11%
学校支持-家庭支持-个人支持-综合素养	0.007	-0.006	0.044	-0.011	0.033	2.63%
总间接效应	0.120	0.029	0.268	0.021	0.256	45.11%
直接效应	0.146	-0.092	0.365	-0.092	0.364	54.89%
总效应	0.266	0.053	0.459	0.056	0.464	—

注：表中点估计值为标准化系数。

如表 2，通过中介效应分析，本文针对停课半年以内的样本组，揭示了学校支持在学生综合素养形成中的作用。具体而言，总效应的标准化估计值为 0.266，其 95%置信区间不包括零，表明学校支持对学生综合素养有显著的正向影响。然而，直接效应的估计值为 0.146，其 95%

置信区间包含零，这意味着学校支持对学生综合素养的直接影响并不显著。相对而言，总间接效应的估计值为 0.120，95%置信区间不包含零，充分证明了学校支持通过间接路径对学生综合素养产生显著影响。

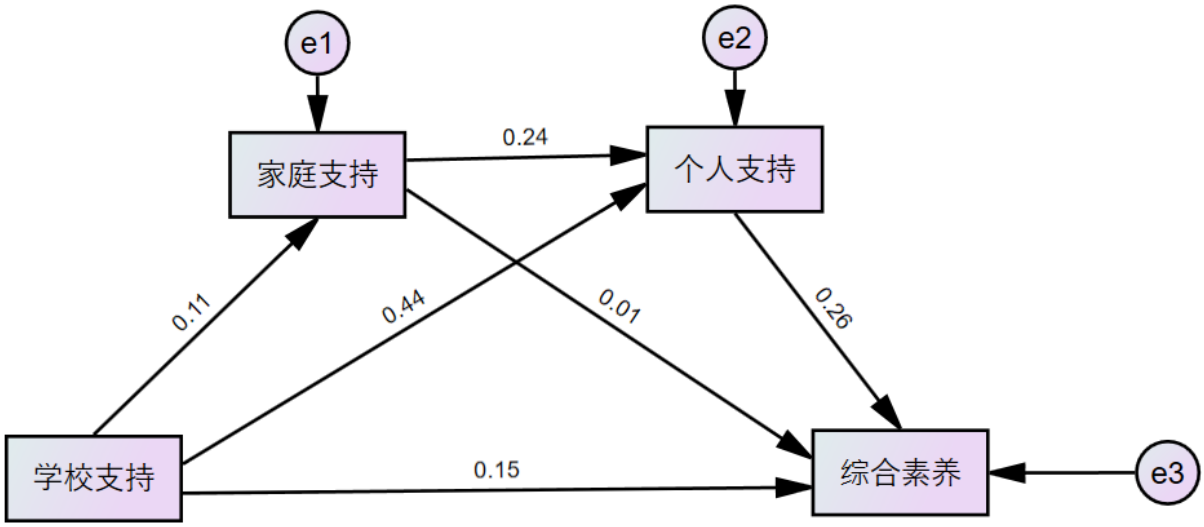


图 4 停课半年以内组的链式中介模型路径分析图

注：图中系数为标准化系数。

图 4 的分析显示，在评估特定的中介路径时，通过家庭支持的中介效应估计值为 0.001，且其 95%置信区间包含零，表明该路径的中介效应不显著。链式中介效应的估计值为 0.007，其 95%置信区间也包含零，进一步证实链式中介效应在统计上不显著。然而，通过个人支持的中介效应估计值达到 0.112，其 95%置信区间不包含零，清楚地显示个人支持作为中介的作用是显著的。这表明学校支持可以通过增强个人支持，间接地对学生的综合素养产生正面影响，该间接效应占总效应的 42.11%。综合这些发现，对于停课半年以内的学生而言，其综合素养的形成机制可以被视为一个完全中介模型，其中学校支持不直接影响学生的综合素养，而是通过个人支持的中介作用间接发挥作用。这些结果强有力地支持了研究提出的假设 H3，并为理解疫情期间教育支持对学生发展影响的机制提供了新的视角。

3.2．停课半年至一年组的学校支持完全通过中介影响学生综合素养

表 3 停课半年至一年组的综合素养影响机制的中介效应分析

路径	点估计	校正后偏差		百分位数		效应比
		下限	上限	下限	上限	
学校支持-家庭支持-综合素养	0.003	-0.381	0.486	-0.489	0.371	0.40%
学校支持-个人支持-综合素养	0.571	0.212	1.887	0.090	1.746	77.06%
学校支持-家庭支持-个人支持-综合素养	-0.003	-0.618	0.658	-0.503	0.799	-0.40%
总间接效应	0.571	0.319	1.888	0.200	1.753	77.06%
直接效应	0.170	-0.659	1.299	-0.495	1.502	22.94%
总效应	0.741	0.702	2.190	0.725	2.232	—

注：表中点估计值为标准化系数。

由表 3 可见：停课半年至一年组的学校支持对学生综合素养的影响的总效应（以下均为点估计值）为 0.741，95%CI 不包含 0，学校支持对学生综合素养具有显著的正向影响；直接效应为 0.170，95%CI 包含 0，学校支持未对学生综合素养产生显著的直接影响；总间接效应

为 0.571，95%CI 不包含 0，学校支持对学生综合素养具有显著的间接影响。进一步分析特定路径的中介效应，以家庭支持为中介的中介效应为 0.003，95%CI 包含 0，其中介效应不显著。同样地，链式中介效应为-0.003，95%CI 包含 0，其中介效应不显著。以个人支持为中介的中介效应为 0.571，95%CI 不包含 0，其中介效应显著。

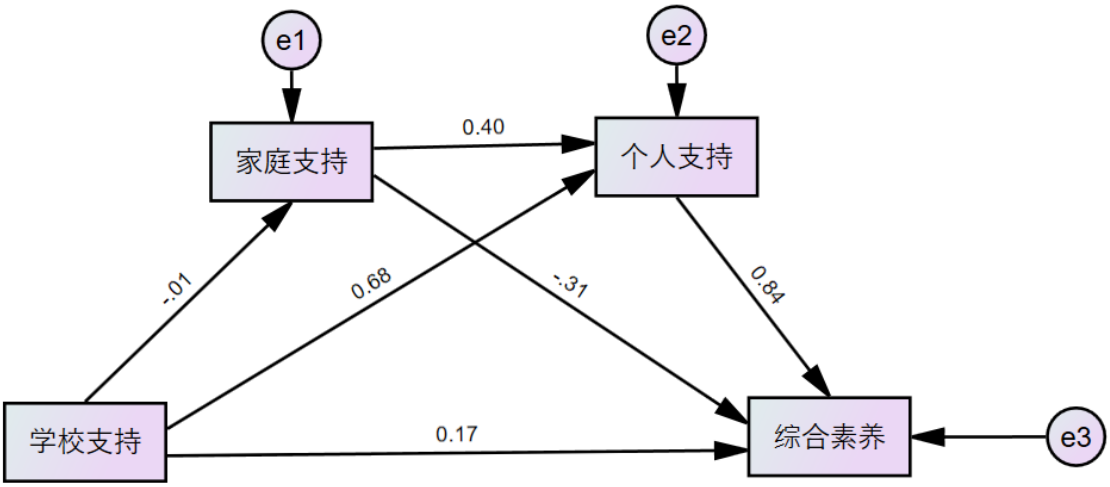


图 5 停课半年至一年组的链式中介模型路径分析图

注：图中系数为标准化系数。

结合图 5 可见，学校支持能够促进学生个人支持，从而对综合素养具有间接积极影响。综上，停课半年至一年组的综合素养影响机制模型为完全中介模型。学校支持不会直接对学生综合素养产生影响，而是通过个人支持中介间接影响综合素养，以个人支持为中介的间接效应占总效应的 77.06%。研究结果支持本研究提出的 4 个假设中的 H3。

3.3．停课一年以上组的学校支持不显著影响学生综合素养

由表 4 可见：停课一年以上组的学校支持对学生综合素养的影响不显著。可能是因为长时间的疫情封闭，隔绝了外界交流，使外界环境对综合素养影响减弱。

表 4 停课一年以上组的综合素养影响机制的中介效应分析

路径	点估计	校正后偏差		百分位数		效应比
		下限	上限	下限	上限	
学校支持-家庭支持-综合素养	-0.805	-2.283	0.890	-3.126	0.596	414 %
学校支持-个人支持-综合素养	-0.404	-1.274	0.733	-1.109	0.965	207%
学校支持-家庭支持-个人支持-综合素养	0.515	**-0.102**	**1.230**	**0.076**	**1.412**	-264%
总间接效应	-0.694	-1.616	0.833	-1.796	0.730	357%
直接效应	0.500	-0.420	2.409	-0.703	1.456	-257%
总效应	-0.194	-1.128	0.461	-1.070	0.497	—

注：表中点估计值为标准化系数。

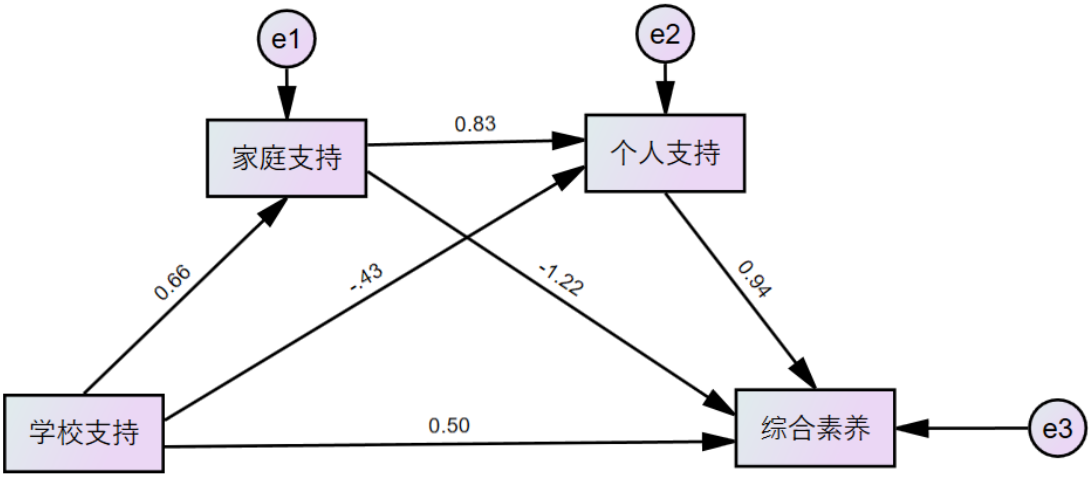


图 6 停课一年以上组的链式中介模型路径分析图

注：图中系数为标准化系数。

综合图 6，本文观察到在停课一年以上的学生群体中，学校支持对于家庭支持和个人支持的负向关系，以及对学生综合素养的总效应非显著，表明长期的教育中断可能削弱了学校作为社会支持系统的作用。由此，当前研究结果不足以证实假设 H2 和 H3，即家庭支持和个人支持在学校支持与学生综合素养之间起中介作用的假定。同样，链式中介效应的假设 H4 也未得到充分的实证支持，因为虽然有正向的估计值出现，其 95%置信区间包含了零点，指示出中介效应的统计显著性不足。据此，我们得出结论，对于长时间停课的学生群体，学校支持的影响在统计上不显著，不支持假设 H2、H3 和 H4，从而提示我们需在学校支持策略上重新审视和调整，以适应长期停课情境下学生综合素养的发展需要。

4. 研究结论与建议

本文借助 PISA 2022 数据，依据生态学理论，对疫情下的“停课不停学”措施所涉及的综合素养影响机制进行了探讨。本文发现，在停课半年以内的学生群体中，学校支持通过个人支持的中介作用，显著影响学生的综合素养，而家庭支持的中介作用不显著。这一结果强调了个人支持在学习连续性中的重要性，同时对家庭支持的作用提出了进一步探究的需求。值得注意的是，当停课时长达到半年至一年时，学校支持对综合素养的总间接效应显著增强，占总效应的 77.06%，这进一步证实了长期停课下个人支持在学生学习中的关键作用。在停课超过一年的情况下，本文发现学校支持对综合素养的影响不显著，可能是长期封闭的疫情（Kuhfeld, M. et al., 2020），使外部环境影响减弱，这提示了对更长时期停课影响的深入分析的必要性。

针对长期停课的学生，应加强个人支持，如提供个性化的在线学习资源和心理健康咨询，确保学生在家中也能获得质量相当的教育。家庭支持虽然在本研究中未显示出显著的中介效应，但其在学生学习中的潜在作用不可忽视。学校应通过家校合作提高家长参与在线教育的意识和能力。面对不断变化的疫情，学校应采取更加灵活的教学方案，如混合教学模式（Means, Toyama, Murphy, Bakia, & Jones, 2009），以适应可能的教育形态转变。建议在现有教育体制中整合和提升在线教育的质量，确保无论线上还是线下，学生都能享有高标准的教育资源。

5. 启示

本文基于 PISA 2022 数据，探讨了疫情期间“停课不停学”政策下的学生综合素养影响机制。研究表明，学校支持对学生综合素养具有显著的间接正向影响，而直接影响不显著。具体来看，个人支持在学校支持与学生综合素养之间扮演了显著的中介角色，而家庭支持的中介作用并不显著。此外，随着停课时长的增加，学校支持通过个人支持对综合素养的影响

比例显著增长，尤其是在停课半年至一年的群体中。本文强调了在全球性危机下，教育系统的弹性和适应性对维护学生学习成果的重要性。未来的研究应继续深化对不同教育政策效果的探索，并在实践中不断反思和优化教育干预措施（Hattie, 2009）。在研究设计方面，未来研究应进一步探索停课超过一年的群体，并对数据稳定性进行更为严密的检验，以便更准确地解释长期停课对学生综合素养的影响。同时，加强多学科综合研究，从心理学、教育学和社会学等多个角度出发，全面评估疫情对教育影响的深远效应。

参考文献

Desforges, C., & Abouchaar, A. (2003). *The impact of parental involvement, parental support and family education on pupil achievement and adjustment: A literature review* (Vol. 433). London: DfES.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. routledge.

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher, 49*(8), 549-565.

Li, J., Hu, Z., & Pan, L. (2022). Analysis of school support: Systematic literature review of core Chinese-and English-language journals published in 2000–2021. *Frontiers in Psychology, 13*, 933695.

Marphatia, A. A., Edge, K., Legault, E., & Archer, D. (2010). Politics of participation: parental support for children's learning and school governance in Burundi, Malawi, Senegal and Uganda. *The Improving Learning Outcomes in Primary Schools*, 1-20.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.

Morris, P., & Bronfenbrenner, U. (1998). The ecology of developmental processes. *Handbook of child psychology, 1*, 993-1023.

OECD (Organisation for Economic Co-operation and Development). (2023). *PISA 2022 Results (Volume I): The state of learning and equity in education*.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers, 36*, 717-731.

基于数据挖掘的 MOOC 学习者行为特征与学习效果研究

——以北京大学 H 课程为例

Research on the Behavior Characteristics and Learning Outcomes of MOOC Learners Based on Data Mining: A Case Study of Peking University's Course H

苏晗宇[1]，贾积有[1*]
[1]北京大学教育学院教育技术系
suhanyu@stu.pku.edu.cn, * jjy@pku.edu.cn

【摘要】 MOOC 自出现以来就受到教育工作者和学习者的广泛关注，然而 MOOC 的学习效果常常受到人们的质疑。为了充分理解 MOOC 学习者学习行为特征与学习效果的关系，本研究基于数据挖掘，首先使用相关分析法和线性回归法分析了影响学习效果的因素，再从学习者的投入与成效角度出发，使用 KMeans 聚类分析方法将参与者划分为“完美学习者”、“低投入学习者”和“低效率学习者”三类。最后，本文基于研究结果对 MOOC 的教学实施者和平台设计者提出三点建议，并指出本研究的局限性和未来研究的发展方向。
【关键词】 MOOC；聚类分析；数据挖掘；学习行为

Abstract: *Since their inception, MOOCs have attracted widespread attention from educators and learners. However, the effectiveness of MOOCs is often questioned by people. To fully understand the relationship between the learning behavior and learning outcomes of learners, this study, based on data mining, first analyzed the factors influencing learning outcomes using correlation analysis and linear regression. Then, from the perspective of learners' inputs and outputs, we employed the KMeans clustering analysis method to classify learners into "perfect learners", "low-input learners", and "inefficient learners". Finally, based on the results, we offered three suggestions for MOOC educators and platform designers, and pointed out the limitations of this study and directions for future research.*

Keywords: MOOC, cluster analysis, data mining, learning behavior

1. 引言

MOOC（Massive Open Online Courses）是 2012 年兴起的一种新型教学模式，指由大规模学习者参与的在线课程。MOOC 学习者无需受限于时空和学习者数量，能够自由、灵活地参加任何感兴趣的课程。尽管 MOOC 具有诸多优势，但其呈现出的问题和挑战也不容忽视。例如，MOOC 的辍学率高于传统课堂（Jordan, 2014）；MOOC 学习者缺乏有效的监管和指导（Fournier et al., 2014）等等。因此，MOOC 的学习效果也常常受到研究者们的质疑（姜蔺等人, 2013）。对 MOOC 学习者的行为特征和学习效果开展研究有助于我们深入了解学习者的在线学习状况，也能为 MOOC 设计者和教学者提供有益的见解，从而更好地支持学习者的学习。

然而，针对 MOOC 后台记录的大量行为数据，传统的分析方法无法对其进行深入挖掘，可能造成信息的遗漏和片面的理解。因此，本研究将采用基于数据挖掘的方法，对一门具有代表性的 MOOC 课程数据进行分析，挖掘学习者在线学习中潜在的、有价值的教育信息。具体而言，本文试图回答以下两个研究问题：（1）影响学习者 MOOC 课程学习成效的因素有哪

些？（2）根据学习者的学习投入与成效，能够将其分为哪几类？分别有哪些特点？虽然目前已有许多研究开展 MOOC 学习者聚类分析，但大多从课程参与度和完成度角度进行划分，较少考虑学习者的学习效率，即单位时间内或单位资源投入下的知识获取效果，因此本研究将从投入和成效角度出发为 MOOC 学习提供新的理解。通过回答上述两个问题，本文呈现了 MOOC 学习的现实状况，并基于研究的结果，为我国 MOOC 设计和优化提供建议和思路。

2. 数据处理与描述性统计分析

本研究数据来自北京大学在 Coursera 平台上开设的 H 课程。由于 MOOC 秉持着以学习者为中心的理念（Blum-Smith et al., 2021），因此本研究聚焦于学习者个体，使用网站后台记录的 7183 名学习者在该课程中产生的学习数据对学习者行为特征与学习效果进行分析。在特征选择上，已有研究表明 MOOC 学习者聚类研究的特征指标的来源包括人机交互、人际交互和学习绩效三个方面（王梦倩等人，2018）。对应地，本研究从课程的访问与浏览、学习者的交互与协作、课程的任务与测试三个维度出发，选取课程在线时长、课程开始时间、视频浏览次数、页面浏览次数、讲座浏览次数、论坛发帖次数、论坛发帖票数、论坛评论次数、论坛评论票数、论坛声誉、小测总分这 11 个特征，并用课程最终测试成绩表征学习者的学习效果。

通过统计，共有 863 名学习者获得了有效课程成绩，课程完成率（12.01%）和及格率（13.44%）均较低，这与 MOOC 研究的现状（Cheng et al., 2023）相符。此外还发现：（1）课程资源的使用率由高到低依次是课程网页、课程讲座、课程视频、课程小测、课程论坛。（2）各特征的极差和标准差均较大，说明该课程学习者之间的差异极大，这与先前研究中提及的 MOOC 学习者的多样性（姜蔺等人，2013）相符，也说明了对学习者进行聚类的可行性和必要性。

3. 学习者 MOOC 成绩影响因素分析

如表 1 所示，本研究计算了各行为特征与成绩之间的皮尔逊相关系数。其中小测总分与最终测试成绩的相关性最强，其次是课程页面浏览次数。此外，课程开始时间与最终测试成绩之间存在较弱的负相关关系，可能的原因是越早开始课程的学生往往学习动机越强，则越可能在后续的学习中表现优秀。除此之外，其他变量均与最终测试成绩存在正相关的关系。

表 1 特征与成绩间的皮尔逊相关系数

	课程在线时长	课程开始时间	视频浏览次数	页面浏览次数	讲座浏览次数	小测总分	论坛发帖次数	论坛发帖票数	论坛评论次数	论坛评论票数	论坛声誉
相关系数	0.395	-0.089	0.433	0.768	0.322	0.954	0.201	0.061	0.146	0.109	0.094

根据相关性分析结果，我们剔除相关系数最大的小测总分变量，同时进行共线性分析，并计算变量的方差膨胀系数（variance inflation factor，VIF），根据结果剔除了论坛发帖票数和论坛声誉变量，保证变量之间不存在多重共线性。此外，为了消除量纲和变异范围的影响，我们对行为特征及成绩数据进行标准化后再开展线性回归分析，方程系数如表 2 所示。模型 R^2 值为 0.658，拟合效果较好，说明当前回归方程能够在一定程度上解释因变量的变化。

表 2 线性回归方程系数

	课程在线时长	课程开始时间	视频浏览次数	页面浏览次数	讲座浏览次数	论坛发帖次数	论坛评论次数	论坛评论票数
|系数|	0.051	0.039	0.191	1.038	0.057	0.200	0.119	0.025

具体而言，在“课程的访问与浏览”维度上，页面及视频浏览次数的回归系数明显大于其它变量，是对学生最终成绩影响最大的因素。在“学习者的交互与协作”维度上，学生的论

坛发帖及评论数目回归系数相对较大，说明考虑学生在论坛的参与度对于预测其成绩有较大帮助，而学生论坛评论票数的系数仅为 0.025，是影响最小的变量，可能的原因是其数据间差异性较小（该变量标准差仅为 0.944，为所有变量中最小），难以为回归预测提供有效的信息。

4. MOOC 学习者分类

针对 MOOC 的学习者分类问题，KMeans 聚类方法因其操作简单且在大数据集上聚类效率较高的特点，备受研究者的青睐，成为解决该问题最普遍的聚类方法（王梦倩等人，2018）。因此，本研究选用 KMeans 聚类法，将标准化后的特征数据导入 Weka 软件，共得到 3 类学习者。表 3 中展示了 KMeans 聚类中心结果，即各类特征的平均值。图 1 刻画了课程在线时长与最终测试成绩的分布散点图，高维空间中的三类样本在该二维分布中具有明显的划分。

表3 KMeans聚类中心

类别	课程在线时长	视频浏览次数	页面浏览次数	讲座浏览次数	小测总分	论坛发帖次数	论坛发帖票数	论坛评论次数	论坛评论票数	论坛声誉
cluster 0	2.655	2.913	5.331	2.193	6.693	1.416	0.350	0.906	0.719	0.595
cluster 1	-0.371	-0.112	-0.165	-0.135	-0.155	-0.049	-0.021	-0.039	-0.024	-0.027
cluster 2	2.082	0.298	0.281	0.565	0.001	0.111	0.088	0.123	0.051	0.086

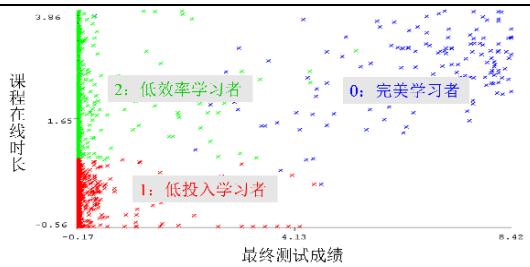


图1 课程在线时长-最终测试成绩的分布散点图

我们定义学习效率为学习成绩与学习投入的比值，并针对每类学习者的参与方式进行描述：

（1）“完美学习者”：对应 cluster 0，包含 142 名学习者（2%），比例极低。该类学习者在课程最终测验中获得较优的成绩，且课程整体参与度较高（例如，在线总时长较长、各类学习资料的浏览次数、论坛的参与次数均较高），是 MOOC 课程的理想学习者。

（2）“低投入学习者”：对应 cluster 1，包含 6129 名学习者（85%）。由表 3 可知，该类学习者在课程活动、课程资料上的投入度较低，同时在最终的测验中表现也较差。由于该类占比最大，因此也说明 H 课程的总体投入低、总体表现差，符合当前 MOOC 学习的现状。

（3）“低效率学习者”：对应 cluster 2，包含 912 名学习者（13%）。该类学习者的在线时长较长，与“完美学习者”接近，但课程资料浏览次数、论坛的参与次数却低于“完美学习者”。与此同时，其最终测试成绩略高于“低投入学习者”，但明显低于“完美学习者”，整体水平仍然偏低。因此，与其他学习者相比，该类学习者在投入相同的时间和学习成本的情况下，有效课程参与次数较少，他们的特点是投入高、成绩低。由此可见，他们的学习效率较低。

5. 结论与讨论

本研究使用数据挖掘的方法对一门 MOOC 数据进行深入挖掘，探索影响学习者 MOOC 成绩的相关因素。研究发现，学习者的小测成绩与课程最终成绩具有较高的相关性，因此通过对小测成绩的观察能够在很大程度上预测学习者在课程上的最终成就。这也强调了形成性评估的重要性，教师若能通过测验等教学活动及时了解学习者的学习情况，并开展及时的干预，

则可能提升学习者的终结性评价结果。除此之外，与论坛参与相关的行为也在一定程度上解释了课程的最终成绩。由于该课程面向全球学生开放，参与者通常是陌生的、来自不同国家或地区的学习者，因此形成在线交流与协作的难度较大。由此可以看出，积极参与讨论的学习者往往对该课程具有较高的参与热情，也更可能在课程中坚持下来，取得较优异的成绩。

不同于现有 MOOC 学习者的聚类结果，本研究考虑学习效率，从投入与成效的角度出发，将学习者分为“完美学习者”、“低投入学习者”和“低效率学习者”三类。值得注意的是，该课程中几乎不存在高效率学习者，这也说明 MOOC 对学习者的学习投入程度提出了一定的要求。

基于上述结论，本文针对 MOOC 的设计与实施提出以下三点建议：（1）教师应当谨慎、高效地开展形成性评估，并重视评估的结果和反馈，为在学习过程中表现较差的学生，或潜在的“辍学者”提供及时的干预和指导（Aldowah et al., 2020）。（2）MOOC 课程和平台设计者应该考虑如何改善教学设计，以提升学习者的参与度。例如，增加强吸引力、高互动性的学习活动，也可以借鉴游戏化教学设计，激发学生的兴趣和主动参与意愿（Aparicio et al., 2019; 尚俊杰和裴蕾丝，2015）。（3）教师可以从投入与成效的角度识别不同的类型的学习者，为他们提供适应性干预和支持。例如，“低效率学习者”可能存在“刷课”行为，也可能由于学习者的自我调节能力较差，导致对课程内容的接受存在障碍，需要教师及时干预和指导。

本研究也存在其局限性。在数据来源上，研究仅使用了总结性的行为统计数据，而缺乏对 MOOC 学习过程相关的时序数据以及学习者认知和情感数据的关注。此外，本研究仅使用测试成绩作为学习效果的评价指标，较为单一。未来的研究可以充分考虑过程性特征以及认知情感特征，并在更丰富的评价维度下探讨上述问题，以获得对 MOOC 学习更全面的理解。

基金项目

感谢田家炳基金会北大港大学习科学青年学人交流项目对本研究的支持。

参考文献

王梦倩、范逸洲、郭文革和汪琼（2018）。MOOC 学习者特征聚类分析研究综述。**中国远程教育**，**07**，9-19+79。

尚俊杰和裴蕾丝（2015）。 重塑学习方式:游戏的核心教育价值及应用前景。**中国电化教育**，**05**，41-49.

姜蔺,、韩锡斌和程建钢（2013）。MOOCs 学习者特征及学习效果分析研究。**中国电化教育**，**11**，54-59+65.

Aldowah, H., Al-Samarraie, H., Alzahrani, A. I., & Alalwan, N. (2020). Factors affecting student dropout in MOOCs: a cause and effect decision-making model. *Journal of Computing in Higher Education, 32,* 429-454.

Aparicio, M., Oliveira, T., Bacao, F., & Painho, M. (2019). Gamification: A key determinant of massive open online course (MOOC) success. *Information & Management, 56*(1), 39-54.

Blum-Smith, S., Yurkofsky, M. M., & Brennan, K. (2021). Stepping back and stepping in: Facilitating learner-centered experiences in MOOCs. *Computers & Education,* 160, 104042.

Cheng, J., Yuen, A. H., & Chiu, D. K. (2023). Systematic review of MOOC research in mainland China. *Library Hi Tech, 41*(5), 1476-1497.

Fournier, H., Kop, R., & Durand, G. (2014). Challenges to research in MOOCs. *Journal of Online Learning and Teaching, 10*(1), 1.

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *International Review of Research in Open and Distributed Learning, 15*(1), 133-160.

大学生 MOOC 学习投入影响因素实证研究

An empirical study on the influencing factors of college students ' MOOC learning engagement

马国威，赵赫璇，付玉娜，雅古，区文诗，王妍莉*
西北民族大学教育科学与技术学院
*99336527@qq.com

【摘要】已有研究表明学生慕课学习投入受不同因素影响。本研究整合了整合型技术接受与使用模型（UTAUT）和技术接受模型（TAM3）探讨了自我效能感、感知有用性、感知易用性对 MOOC 支持的线上线下混合课程学习投入的影响。研究结果显示，感知有用性、感知易用性、自我效能感、社群影响和便利条件也对学习投入有显著正向影响，此外，社群影响和便利条件对感知易用性具有显著正向影响，基于此提出相关改善建议以供参考。
【关键词】 MOOC；学习投入；结构方程模型；学业自我效能感

Abstract: *Previous studies have shown that students ' MOOC learning engagement is affected by different factors. This study integrates the Integrated Technology Acceptance and Use Model (UTAUT) and Technology Acceptance Model (TAM3) to explore the impact of self-efficacy, perceived usefulness, and perceived ease of use on learning engagement in MOOC-supported online and offline blended courses. The results show that perceived usefulness, perceived ease of use, self-efficacy, community influence and convenience also have a significant positive impact on learning engagement. In addition, community influence and convenience have a significant positive impact on perceived ease of use. Based on this, relevant improvement suggestions are proposed for reference.*
Keywords: MOOC ; learning engagement ; structural equation model ; academic self-efficacy

基金项目：本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学课程思政示范课程《信息技术教学法》（项目编号：2022KCSZKC-37）资助。

1. 引言

MOOC 的出现一方面使得学习者获取优质教育资源的方式变得更为便捷，另一方面也推动了高校本科教学的发展。然而，相关的调查研究发现，在 MOOC 学习过程中仍存在“高辍学率”等因时空分离导致的在线学习质量危机，存在学习投入不佳等问题，导致人们对 MOOC 的学习成效产生了质疑。如何走出这一困境从而提高学习者 MOOC 学习质量，提升学习者的在线学习投入成为 MOOC 乃至在线学习领域研究者们共同探索的话题。基于此，本研究通过结合 UTAUT 模型和 TAM3 模型，构建研究假设模型，以此探究自我效能感、感知有用性、感知易用性对大学生 MOOC 学习投入的影响，并进一步探究不同性别对各因素影响学习投入的差异，并尝试探讨如何提升大学生 MOOC 学习投入。

2. 文献综述与研究假设

2.1 文献综述

2.1.1 学习投入

作为衡量 MOOC 学习效果的重要指标，学习投入是指学习者进行 MOOC 学习时的一种积极的情绪和精神状态（Schaufeli et al.,2002）。已有研究表明 MOOC 学习中，学习者学习投入正向显著影响学习收获。本研究将行为、情感和认知投入这三方面作为学习投入的二级维度进行分析（Fredricks，2004）。

2.1.2 UTAUT 模型

整合型技术接受与使用模型包含有 4 个影响采纳意图和采纳行为的核心变量即:“绩效预期”、“努力预期”、“社群影响”和“便利条件”,同时该模型还包含有 4 个控制变量即:“年龄”、“性别”、“经验”和“自愿性使用”。研究证实，UTAUT 模型中的四个要素对学生使用 MOOC 学习意愿具有正向影响（王钱永，2016）。

2.1.3 TAM3 模型

综合前人对 TAM 的研究出现了 TAM3（Venkatesh V，2010），确定了四类影响因素：个体差异、系统特征、社会影响和便利条件。已有研究发现 TAM 模型中的社会影响、便利条件、感知易用性和感知有用性是直接影响高校学生 MOOC 学习意愿的重要因素。

2.1.4 感知有用性和感知易用性

Davis 首次在 TAM 模型中提出了感知有用性和感知易用性的概念，认为存在一些外部变量影响感知有用性和感知易用性，感知易用性正向影响感知有用性，两者共同影响用户的使用态度，感知有用性对用户使用意愿有直接的影响，感知易用性对感知有用性有着积极影响（Sorebo & Eikebrokk，2008）。在对教师支持对 MOOC 课程持续学习意愿的研究中证实关于感知体验方面，学习者自我效能感、感知有用性和感知易用性等因素也会影响持续学习意愿（吴华君、葛文双 &和何聚厚，2020）。

2.1.5 社群影响与便利条件

现有研究表现学生的社群影响对感知易用性有显著的正向预测作用，且便利条件对学习者感知有用性和感知易用性有积极的直接影响。此外，研究发现一个支持性的学习环境、积极的同伴关系等都能够提高学习的易用性，使学生更容易投入到学习过程中。

2.1.6 自我效能感

自我效能感为自我在面对具体的情境时对自身所具有的行为能力的自信水平，现有研究验证了学习自我效能感对学习投入具有正向的预测作用。学习能力自我效能感是指个体对自己是否具有顺利完成学业、取得良好成绩和避免学业失败的能力的判断和自信。学习行为效能感是指个体对自己能否采取一定的学习方法达到学习目标的判断与自信。

2.2 研究假设与模型构建

根据已有研究可知感知易用性和感知有用性对学习投入有促进关系，学习自我效能感对学习投入都具有正向的预测作用。通过研究使用者对在线学习的接受，并比较有无调节变量的模型使用效果，结果表明有调节变量的模型有利于对问题进一步解释。因此本研究基于 UTAUT 和 TAM3 理论模型，引入性别差异这一调节变量，探寻调节变量导致的各个因素对学习投入的影响效果。并提出以下假设：H1 感知易用性对学习投入具有显著正向影响；H2 感知有用性对学习投入具有显著正向影响；H3 感知易用性对感知有用性具有显著正向影响；H4 社群影响对感知易用性具有显著正向影响；H5 便利条件对感知易用性具有显著正向影响；H6 学习能力自我效能感对感知易用性具有显著正向影响；H7 学习行为自我效能感对感知易用性具有显著正向影响；H8a 社群影响对感知易用性的影响存在性别差异；H8b 便利条件对感知易用性的影响存在性别差异；H8c 学习能力自我效能感对感知易用性的影响存在性别差异；H8d 学习行为自我效能感对感知易用性的影响存在性别差异。

3. 研究设计

3.1 研究对象

本研究通过对西北某高校学习 MOOC 支持的的混合式课程的学生发放问卷进行数据收集。回收问卷 638 份，剔除无效问卷后有效问卷 620 份，问卷有效率 97%。其中，男生有 202 人（占比 32.6%），女生有 418 人（占比 67.4%）。

3.2 研究工具

本研究的问卷设计主要包括人口学信息量表和测量量表。测量量表包括：学业自我效能感量表(梁宇颂,2000),包括学习能力自我效能感、学习行为自我效能感两个维度;Blackboard 在线学习影响因子归纳分析与技术接受度量表(Muilenburg L Y,Berge Z L,2005)，包括感知因子、内容因子、习惯因子和环境因子四个维度；学习投入量表（Awang et al.，2008），包括行为投入、认知投入、情感投入三个维度。经检验量表信效度良好，可进行下一步分析。

3.3 数据分析

3.3.1 变量相关检验

本研究采用 Pearson 相关检验进行各变量之间的相关分析，分析表明大学生的社群影响、便利条件和学习自我效能感各维度对感知易用性均达到了显著正相关；感知有用性对学习投入各维度均达到了显著正相关；感知易用性对学习投入除了行为投入外，均达到显著正相关；因此证明了假设 H1，H2，H4，H5，H6 成立，假设 H3 感知易用性对有用性的影响不显著，可能是由于学生自身原因，也可能是 MOOC 平台的设计存在问题等原因对其产生影响，从而弱化感知易用性对感知有用性的影响。同时路径系数分析表明各变量之间的相关程度从而为后续分析奠定基础。此外，除假设 H6 外其它假设路径均显著成立。假设 H6 不显著的原因可能因为学生对自己使用 MOOC 学习平台的能力不够自信，从而认为平台操作难以掌握进而降低了学生的感知易用性。

3.3.2 模型拟合优度分析

通过对模型进行验证性因子分子可知各变量组成信度系数良好，信度质量高；平均方差抽取量可接受，模型收敛效度在可接受范围内。因此，为了进一步验证技术接受度、自我效能感和学习投入之间的关系，通过对模型进行修正得到如图 1 所示的 MOOC 学习投入影响因素的结构模型。在信效度检验通过的基础上利用 AMOS 软件对图 1 所示的研究模型进行验证,该模型的拟合优度指数值分别为：X^2/df =4.899，GFI =0.960，AGFI =0.887，TLI =0.898，RMSEA =0.106，CFI=0.955，IFI=0.955。结果显示模型的拟合优度指数良好，该模型各条路径的回归系数均达到显著水平,各变量之间相互影响,说明本研究构建的模型适用于高校大学生 MOOC 学习投入影响因素的分析。综合来看，模型的拟合程度较好。

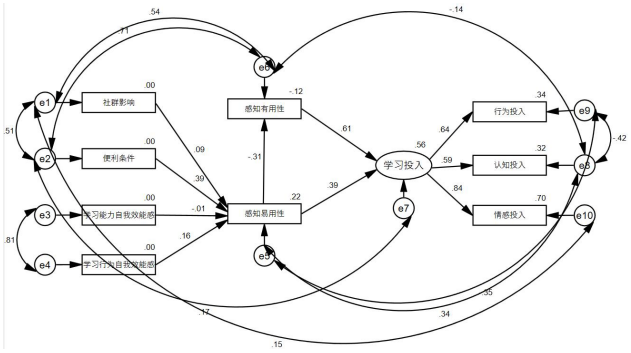


图 1MOOC 学习投入影响因素的结构模型

通过分析可知，影响学习投入的总效应依次为感知有用性（0.61）＞感知易用性（0.39），影响感知易用性的总效应依次为便利条件（0.39）＞学习行为自我效能感（0.16）>社群影响（0.09）。根据图 1 可知自变量与因变量之间的作用关系除 H3、H6 以外均呈现正向直接影响，各路径系数呈现较强的显著性。研究表明：假设 H1、H2、H4、H5、H7、H8a、H8b、H8d 成立。

3.3.3 调节效应检验

在检验调节效应时，将性别作为调节变量对各假设进行调节效应检验得出，性别对社群影响正向影响感知易用性存在显著调节作用（p＜0.001），性别对便利条件正向影响感知易用性存在调节作用(p＜0.001)。研究假设 H8a、H8b 假设成立，H8c、H8d 假设不成立。

4. 结论与分析

4.1 满足课程需求，优化平台设计

通过用户调研、数据分析等方式，了解大学生的学习需求和偏好，进一步优化平台内容和资源。其次，MOOC 平台应该注重不断优化平台界面、提供个性化推荐、加强用户支持等方式，提高大学生使用平台的易用性和用户满意度。同时学校和教师应该加强对 MOOC 平台的推广和支持，让更多的大学生了解到该平台的存在和优势。可以通过宣传推广、定期组织培训、提供技术支持等方式，鼓励和帮助大学生使用 MOOC 平台进行学习。基于次学校或教育机构可以考虑提供易于使用的学习平台，提供稳定、高速的网络环境，以便学生更加高效地学习，也可以根据学生的不同需求提供个性化的学习资源等。

4.2 提高用户参与，改善用户体验

首先，为用户提供详细的培训和支持，使他们能够更好地理解和使用新系统。其次，提高用户参与度，鼓励用户积极参与系统的使用和学习。另外，设计友好的用户界面使用户感到舒适和自信，从而增强他们的学习自我效能感。最后，为用户提供及时的反馈和评估，以帮助他们了解自己的进展和成就。

4.3 扩充课程资源，丰富学习方式

MOOC 平台应该根据性别差异提供符合不同性别学习习惯和需求的课程内容和学习资源。同时，也应该注重平台的用户体验设计，提高平台的易用性和用户满意度，从而更好地满足不同性别学生的使用需求。根据分析可知，不同年级对慕课学习行为的影响也是存在的，对于年轻的大学生，可以提供更加有趣、生动的课程和学习方式，而对于高年级的学生，则可以提供更加深入、系统的课程内容和学习资源。

参考文献

王钱永，毛海波.(2016).基于 UTAUT 模型的 MOOC 学习行为因素分析. 电化教育研究(06),43-48.

吴华君,葛文双 & 何聚厚.(2020).教师支持对 MOOC 课程持续学习意愿的影响研究——基于 S-O-R 和 TAM 的视角. 现代远距离教育(03),89-96.

梁宇颂.(2000).大学生成就目标、归因方式与学业自我效能感的研究(硕士学位论文,华中师范大学).

Schaufeli, W. B. , Martinez, I. M. , Pinto, A. M. , Salanova, M. , & Bakker, A. B. . (2002). Burnout and engagement in university students: a cross-national study. Journal of Cross-Cultural Psychology, 33(5), 464-481.

Sorebo, O. , & Eikebrokk, T. R. . (2008). Explaining is continuance in environments where usage is mandatory. Computers in Human Behavior, 24(5), 2357-2371.

Muilenburg, L. Y. , & Berge, Z. L. . (2005). Student barriers to online learning: a factor analytic study: distance education: vol 26, no 1. Distance Education.

Venkatesh, V. , & Bala, H. . (2010). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273-315.

学习论坛中短文本语料的潜在语义信息挖掘

Mining Potential Semantic Information of Short Text in Learning Forums

郑佳妮[1*]，王雯清[1]，许玮[1]
[1]浙江工业大学 教育科学与技术学院
[*] 910518463@qq.com

【摘要】 在线学习论坛是在线学习不可或缺的一部分，但论坛帖子的随意性导致无意义内容增多，高质量话题提取成为挑战。研究以《教学微视频设计与开发》在线课程论坛为例，结合自然语言与计算机语言，运用 BTM 模型对预处理过的文本数据进行主题提取，通过划分主题特征维度、绘制主题特征词共现网络进一步挖掘论坛短文本语料中的潜在语义信息。研究结果表明，不同主题下的特征词共现情况可以揭示学习者的讨论焦点和兴趣；学习论坛主题讨论的热度和深度随课程进展显著提升；学习论坛实现交互内容的质量提升和知识建构的层次提高。

【关键词】 在线学习论坛；BTM 主题模型；文本挖掘；ENA

Abstract: *Online learning forums are a vital component of online-learning environments.However, the informal nature of forum posts often results in an abundance of irrelevant content, posing challenges for extracting high-quality discussion topics.This study investigates the forum of an online course "Design and Development of Teaching Micro-videos." Employing a combination of natural language processing and computational linguistics, the BTM (Biterm Topic Model) is utilized to extract topics from preprocessed text data. By dividing the topic feature dimensions and visualizing the co-occurrence networks of topic-specific words, the research reveals latent semantic information within the forum text. The study's results indicate that learners' focal points, increasing thematic discussions over time, and improved interactive content and knowledge construction in the forum.*

Keywords: Online learning forums, BTM, Data Mining, ENA

1. 前言

2023 年教育部印发《教育部高等教育司 2023 年工作要点》指出“推动智慧教育的发展，积极利用在线学习平台等手段，探索新的教学模式和方法”（教育部，2023）。在线学习论坛是在线学习平台中不可或缺的一部分，它不仅反映了学习者在特定时期的情感状态、学习动机、知识建构以及认知发展水平（左明章，等，2018），而且相比于传统教学方式，它为教师和研究者提供了更加方便快捷的方式去收集学习者的多维度行为数据（张思，等，2022）。因此，深度挖掘学习论坛中文本语料信息，对于教师提供个性化教学资源、调整教学模式及论坛管理者组织、更新论坛都具有极大的价值。

但是学习论坛中帖子的随意性会导致大量无意义内容的存在，使得提取高质量话题和指定话题成为一个巨大挑战。另一方面，学习论坛中的短文本数据量大、字数少、篇幅小，缺乏上下文信息，难以使用传统的文本聚类方法进行分析。因此，如何有效组织和管理这些短文本信息，提取有用的语义信息是当前亟待解决的问题。研究以《教学微视频设计与开发》在线课程论坛为例，结合自然语言与计算机语言，运用 BTM 模型对预处理过的文本数据进行

主题提取，通过划分主题特征维度、绘制主题特征词共现网络进一步挖掘论坛短文本语料中的潜在语义信息。

2. 文献综述

2.1. 在线学习论坛数据挖掘

在线学习论坛数据挖掘是一种重要的技术手段，通过对学习者在论坛上发布的短文本内容进行挖掘分析，可以帮助学习者更好地掌握知识和提高学习成效。以往的研究常使用的数据挖掘方法有分类、聚类、关联规则挖掘、序列模式挖掘和文本挖掘等（柴艳妹和雷陈芳，2018）。然而，上述的数据挖掘方法虽然能准确反映学习者在某个方面的信息或某个课程的学习情况，但无法深入解释其背后的语义信息。

2.2. 短文本主题模型

由于在线学习论坛中多数内容都是短文本，传统的主题模型在短文本场景下存在稀疏性等问题。为了解决这些问题，研究者们提出了许多改进方法，包括借助外部文本信息扩充短文本（Lim，2010）、借助外部知识库将短文本扩展为“伪长文本”（Staab S，2003），以及提高主题模型的质量等方面的改进（Yoshikawa，2014）。然而，这些模型只在长文本中取得了很好的效果，由于短文本的简洁化和片段化特性无法得到很好的效果。

为了解决应用传统主题模型时词语之间语义联系较弱的问题，研究者们提出了 BTM 主题模型。该模型通过对双词进行建模，解决了短文本稀疏的问题，并且无需引用外部语料进行扩充。汤秋莲等（汤秋莲，2014）的研究发现，基于 BTM 的短文本聚类方法比传统的基于 LDA 和 VSM 的聚类效果有很大提高。戴宇鹏等（戴宇鹏，2022）对照片信息文本进行主题建模分析，结果显示 BTM 模型在主题刻画的精确性和全面性方面优于 LDA 模型。

综上所述，BTM 主题模型相较于传统的主题模型在应用于短文本的主题挖掘和聚类上更具优越性，将其应用于在线学习论坛将具有一定的研究价值和实际意义。因此本研究旨在通过应用 BTM 主题模型对预处理过的在线学习论坛文本数据进行主题提取，进一步挖掘文本深层次的语义信息，并为教学提供可行性的建议。

3. 研究设计

3.1. 研究对象

本研究基于华东某高校公共选修课《教学微视频设计与开发》展开。课程的 60 名教育技术学（师范）、计算机科学与技术（师范）的学生为研究对象。课程时长为十六周，约 80 天。

课程采用面对面教学与在线同步教学的模式。在课程开始之初，教师在学习通平台创建了一个课程讨论区供学生们进行学习交流，作为课堂的延伸和补充，学生可以在讨论区进行发帖、回帖参与讨论。此外，在课程开展的过程中，教师也会定期发布与课程内容相关的讨论，促进学生对知识分巩固和理解。课程讨论区共产生 1601 条帖子和评论数据，去除无效数据后，最终获得 1515 条有效评论。在课程结束时，研究者对每位学习者在本课程取得的平时成绩、期中成绩和期末成绩取平均值，以 90 分作为标准，将学生划分为高分组（32 人）和低分组（28 人）。

3.2. 研究过程与方法

本研究主要步骤为（1）获取课程讨论区文本数据（2）数据预处理，通过过滤停用词、替换同义词等方法清楚无效数据（3）BTM 主题建模（4）通过可视化工具绘制主题特征词共词网络图挖掘主题内在联系。通过以上几个步骤实现对在线论坛文本数据的主题分析，研究框架如图 1 所示。

本研究使用 BTM 主题建模方法，将学习论坛帖子中的主题以概率分布的形式进行聚类，得到由一系列特征词组成的主题，通过对主题特征词进行特征维度划分，使用 Gephi 工具绘制特征网络图，进一步挖掘主题热点的内在联系与特点。

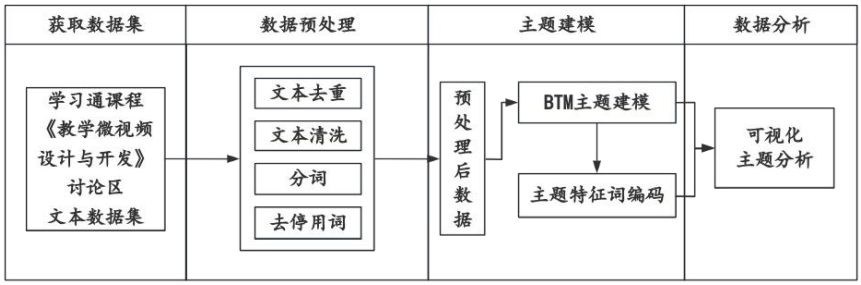


图 1 论坛文本 BTM 主题分析流程图

3.3. 主题特征词编码框架

BTM 建模输出主题后为人为归纳，存在一定的主观性。因此，为了进一步挖掘文本语料中的信息，本文对主题特征词进行统一的属性维度提取，旨在提高主题凝练的客观度（陈秀明，等，2018）。编码过程遵循图 2 所示的特征维度划分框架，首先根据认知深度的层次设计了浅层（SL）、中等层次（ML）、深层次（DL）三个一级维度；其次围绕课程教学目标和知识点，细化八个特征属性，如表 1 所示，分别为主题相关（SL1）、基本操作（SL2）、视觉效果（SL3）、专业术语（SL4）、内容呈现（ML1）、画面处理（ML2）、实际应用（DL1）、创意与素养（DL2）；最后对主题特征词进行特征属性划分，并计算不同特征属性下特征词的累计概率，结果中概率最高属性对应的维度即为主题主导维度。特征维度属性编码如表 1 所示。

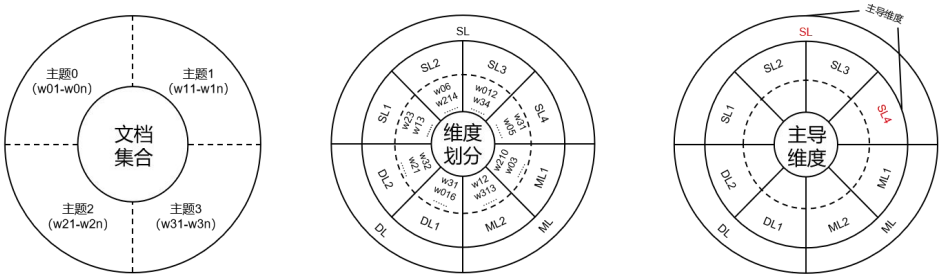


图 2 主题特征维度划分框架

表 1 特征维度属性编码表

认知深度	属性	属性编码	编码解释	特征词样例
SL	主题相关	SL1	与讨论主题词一致的特征词	绿幕、蓝幕
	基本操作	SL2	视频编辑软件基础功能的使用	制作、路径、属性
	视觉效果	SL3	色彩和背景处理以及画面观感	背景、画面
	专业术语	SL4	视频软件中的专业术语解释	suite、渲染、图层
ML	内容呈现	ML1	素材在视频上的布局和组织方式	内容、素材、调整
	画面处理	ML2	AE 软件中的特效和动画处理	特效、效果、技术
DL	实际应用	DL1	教学微视频应用于教学的思考	学习者、趣味性
	创意与素养	DL2	情感态度与创新、信息素养	创意、赋予

4. 数据分析

4.1. 特征词属性维度分析

BTM 主题模型的训练和评估显示，选择 6 个主题时聚类效果最好。提取了每个主题下前 15 个特征词，根据特征维度属性编码表对每个主题下前 15 个特征词进行特征属性划分，得到如图 3 所示各主题下特征词对应的概率和属性。根据主题包含的特征词将各主题分别命名为“影视后期相关知识及 AE 相关术语的辨析”“AE 软件基础操作与画面优化”“视频合成与导出”“AE 插件与特效的运用”“教学动画的制作与应用效果”“AE 项目管理”。

主题 0	概率	属性	主题 1	概率	属性	主题 2	概率	属性	主题 3	概率	属性	主题 4	概率	属性	主题 5	概率	属性
合成	0.0382	SL4	动画	0.0454	ML1	视频	0.0404	ML2	制作	0.0335	SL2	动画	0.0147	ML1	合成	0.0660	SL4
绿幕	0.0297	SL1	路径	0.0280	SL2	AE	0.0341	SL2	效果	0.0327	ML2	制作	0.0127	SL2	素材	0.0454	ML1
蓝幕	0.0267	SL1	图层	0.0270	SL4	输出	0.0295	SL4	特效	0.0324	ML2	内容	0.0111	ML1	文件	0.0355	SL2
蓝色	0.0198	SL1	效果	0.0182	ML2	渲染	0.0292	SL4	插件	0.0322	SL4	技术	0.0109	ML2	AE	0.0276	SL2
素材	0.0180	ML1	文字	0.0179	SL2	动画	0.0256	ML1	动画	0.0307	ML1	简单	0.0104	DL1	工程	0.0237	SL4
颜色	0.0180	SL3	形状	0.0173	SL2	调整	0.0199	ML1	AE	0.0180	SL2	视频	0.0098	ML2	设置	0.0232	ML1
绿色	0.0153	SL1	属性	0.0166	SL2	效果	0.0192	ML2	用户	0.0138	DL1	画面	0.0097	SL3	项目	0.0221	SL2
背景	0.0144	SL3	制作	0.0165	SL2	格式	0.0161	SL4	复杂	0.0122	ML2	学习者	0.0095	DL1	创建	0.0190	SL2
拍摄	0.0142	SL1	蒙版	0.0163	SL2	粒子系统	0.0155	SL4	帮助	0.0121	SL2	生动	0.0069	SL3	布局	0.0181	ML1
场景	0.0133	SL1	实现	0.0131	ML1	音频	0.0150	ML1	Suite	0.0113	SL4	教学	0.0069	DL1	图层	0.0166	SL4
图层	0.0133	SL4	遮罩	0.0110	SL4	工具	0.0141	SL2	电影	0.0104	SL2	呈现	0.0067	SL3	视频	0.0163	ML2
路径	0.0126	SL2	画面	0.0109	SL3	软件	0.0131	SL2	支持	0.0094	SL2	特点	0.0063	DL1	参数	0.0146	SL4
遮罩	0.0104	SL4	赋予	0.0109	DL2	工程	0.0126	SL4	导入	0.0093	SL2	趣味性	0.0060	DL1	编辑	0.0141	SL2
效果	0.0103	ML2	视频	0.0105	ML2	导出	0.0111	SL2	粒子	0.0093	SL4	效果	0.0059	ML2	属性	0.0134	SL2
实现	0.0091	ML1	内容	0.0101	ML1	生成	0.0106	SL2	强大	0.0092	ML2	二维动画	0.0059	SL4	格式	0.0131	SL4

图 3 主题特征词属性及概率

统计各主题在八个属性下的概率分布，得到图 4 中表格。从各属性的总计概率上来看，“实际应用”“创意与素养”这两个属性的值最低，且两个属性均属于高等的认知层次，说明整个讨论过程中涉及主题在高等认知层面表现较差。此外，由每个主题下特征词贡献概率最高的属性推断主题的主导属性，发现各主题在不同属性的概率呈现上存在差异。主题 0 主导维度为“主题相关”，主题 1、3 为“基本操作”，主题 2、5 为“专业术语”，主题 4 为“实际应用”。

一级维度	属性	主题 0	主题 1	主题 2	主题 3	主题 4	主题 5	总计概率
SL	主题相关（SL1）	**0.119**						0.119
	基本操作（SL2）	0.013	**0.113**	0.083	**0.093**	0.013	0.132	0.445
	视觉效果（SL3）	0.032	0.011			0.023		0.067
	专业术语（SL4）	0.062	0.038	**0.103**	0.053	0.006	**0.134**	0.396
ML	内容呈现（ML1）	0.027	0.069	0.060	0.031	0.026	0.087	0.300
	画面处理（ML2）	0.010	0.029	0.060	0.087	0.027	0.016	0.228
DL	实际应用（DL1）				0.014	**0.039**		0.053
	创意与素养（DL2）		0.011					0.011

图 4 各主题在不同维度的特征分布概率

为进一步挖掘同一主导维度下特征词的联系，综合主题特征词的共现网络来分析。在共现网络中，节点大小代表度的大小，节点越大，与其他节点之间联系越多，连线的粗细则代表边权重，连线越粗，两个节点的共现次数越多，联系越紧密。

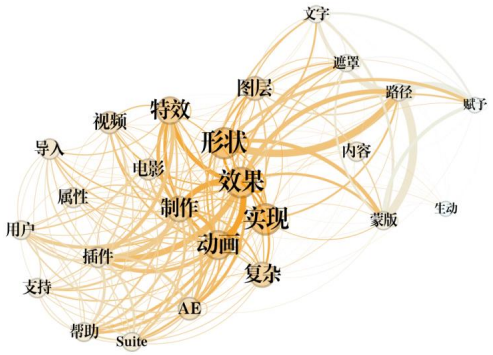
由图 5 中各主题的共现网络图可见，主题 1、3 的主导维度均为“基本操作”（图 5（a））。“形状”“效果”“实现”（画面处理）“动画”（内容呈现）等是主题下比较重要的节点，体现了学生们在动画制作过程中更加注重效果的实现和画面的处理。同时，这些节点与“制作”“路径”“文字”“蒙版”等 AE 基本操作紧密相关，进一步证明了学生们对于 AE 动画制作基本操作的兴趣与求知欲。此外，在基本操作属性下的节点与其他属性维度之间存在共现现象，如“制作”与“插件”“特效”（专业术语）存在较强的连线，这说明学生们在使用基本操作的同时，也在探索更高级的功能如何来优化他们的作品。通过深入理解和应用基本操作，学生们能够更加精准地调整画面效果，创造出视觉冲击力强、表现力丰富的动画作品。

主题 2、5 主导维度均为专业术语属性（图 5（b））。“动画”“视频”“调整”（内容呈现）“图层”“渲染”（专业术语）“AE”（基本操作）是主要的节点，它们代表了与

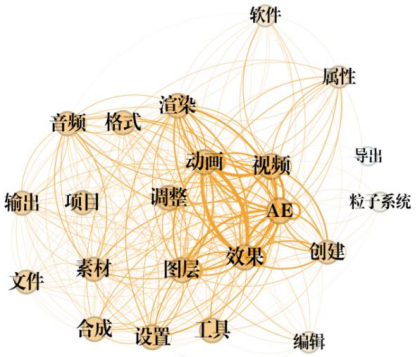
AE 相关的基本实践和工作流程，凸显了学生聚焦于动画制作与编辑、视频调整技巧以及 AE 的基础操作和核心功能学习。除此之外，图示还显示出“效果”（画面呈现）“素材”“音频”（内容呈现）等特征词节点的重要性，这表明专业术语讨论不仅聚焦于 AE 的基本流程，还包括提升作品质量的视听效果探索。这些讨论强调了 AE 基础和丰富内容实践的重要性，以增强视听效果，对学习者至关重要，有助于创意和审美理念的有效传达。

主题 0“主题相关”属性下（图 5（c）），“合成”（专业术语）“绿幕”“蓝幕”“拍摄”“场景”（主题相关）都是在网络中较为重要的节点，且共现次数多。这反映出学生在影视后期相关知识，特别是抠像知识方面上存在着激烈的讨论。通过源数据中的”在亚洲，一般使用蓝幕，而在欧洲，一般使用绿幕。”和“不是，是后期合成的。”等语句，我们可以间接地发现，在引入 AE 教学前，教师先引入相关影视知识方面的问题进行讨论，实现同时引入主题和提高学生的积极性。

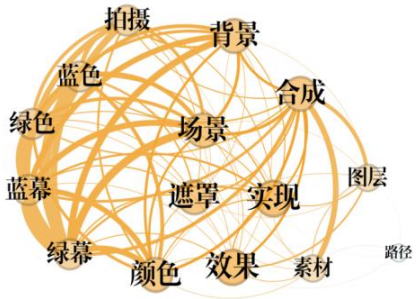
主题 4 主导维度为“实际运用”属性（图 5（d））。“动画”（内容呈现）“制作”（基本操作）“技术”“效果”（画面处理）是比较重要的节点，它们之间的共现频率揭示了学生群体对于动画基础制作技术以及视觉效果优化的高度关注。此外，特定节点“二维动画”（专业术语）的出现，说明了学生们特别关注二维动画，并将其与其他类型比较，认识到其在教学中的优势。进一步地，节点“学习者”“教学”“特点”和“趣味性”（实际运用）及其相应的共现关系，揭示了学生对教学动画在满足需求、增加趣味性和创造互动学习环境方面的认识，反映了他们对教学效果和学习体验的重视。



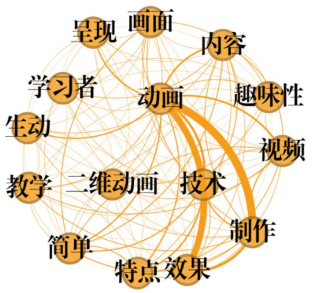
（a）主题 1、3 特征词共现网络



（b）主题 2、5 特征词共现网络



（c）主题 0 特征词共现网络



（d）主题 4 特征词共现网络

图 5 各主题特征词共现网络

4.2. 论坛主题热点演变过程

通过手动划分八个时间片段并绘制折线图，进一步探索讨论主题基于时间的变化趋势。图 6 的折线图横坐标代表时间片段，纵坐标代表讨论频次。通过折线图可知，第一阶段，主题 1 话题频率较高，其他五个主题话题频率较低，学生在在线讨论平台主要以回复教师发布的话题为主。第二阶段，主题 1 下话题数量明显下降，主题 4 下话题数量急剧上升，说明学生正处于制作视频前期的教学微视频设计阶段，关于教学微视频的实际应用关注较大。主题 5 的话题数量在第三阶段突现，表明在第三阶段，随着课程内容的推进，学生开始关注 AE 软件

中的一些专业术语，以及 AE 中项目文件管理和一些基础操作。第四到六阶段，主题 5 的话题数量先下降后稳步上升，各主题话题数量有较小幅度的起伏，但整体氛围不浓，没有特别关注的讨论主题。第七阶段后各主题话题数量有了大幅度提升，学习论坛中专业术语、视觉效果、基本操作、内容呈现、实际应用都开始被大量关注。

聚焦每个主题下的话题讨论数量的特点，主题 1、2、3 强度在前四个阶段最小，受关注最低，但在后四个阶段逐步上升，说明基础操作、专业术语和内容呈现在整个讨论过程中逐渐被关注。主题 4 在整个阶段前期和后期都呈现大幅增长，其主导属性维度为实际应用，体现了在课程讨论的初始阶段与最后阶段，学生均产生了较多深层次的讨论，这也取决于《教学微视频设计与开发》课程是一门先设计再实践的课程类型。

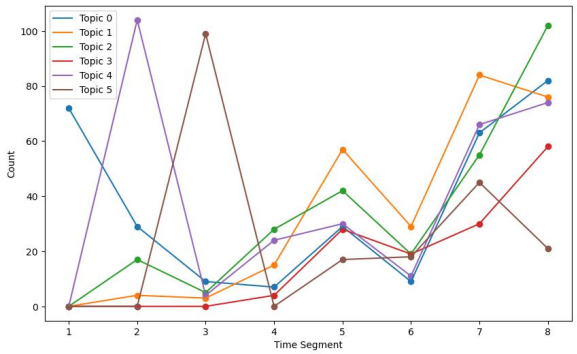


图 6 主题话题数量变化折线图

4.3. 认知网络特征与演变情况分析

评论数据由两位进行过预编码且 Kappa 系数大于 0.7 的编码者进行了主题编码，编码完成后将数据导入 ENA Webkit 网站后生成认知网络图，如图 7 所示。其中，学习者用圆点表示，连接线的粗细代表主题之间连接的强弱。图中分别为高、低分组各自的平均网络图和对比网络图。可以看出高、低分组学习者的认知网络图中主题 0 与其他主题连线均较为明显，证实了教师在学习论坛中发挥着重要作用，引导学生积极展开讨论，启发学生向更高层次的认知主题上探索。中部的两个实心方块各自代表高、低分组的平均网络质心，两个质心在 X 轴上有显著的差异，其与主题 3 和主题 1 相关。高分组较低分组有更多的关于主题 1、3 的讨论，由图 4 可看出主题 1 和主题 3 虽以“基本操作”为主导，在其他认知维度上分布较均匀，表明高分组更关注基础知识的掌握，并且对各个层次的问题均有所探究。高分组学习者各主题之间连线的粗细更加均匀，说明其对各个认知层次的主题均有类似程度的关注。

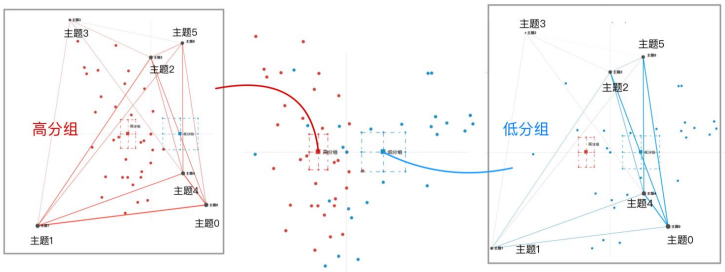
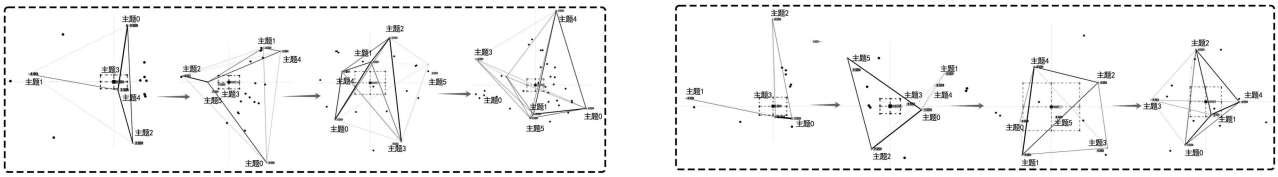


图 7 高、低分组认知网络对比图

为研究认知网络特征的演变情况，将之前划分的八个时间片段整合为四个时间序列，并生成这四个时期的高、低分认知网络图，如图 8 所示。两组认知网络演变情况存在差异，尽管高、低分组的评论随着课程的开展逐渐变多，且两组都由教师引导到自开始主创建话题转变，但低分组整体话题数度少于高分组，在三、四时间片段中尤为明显。低分组主题间的联系粗细不均匀，而高分组中各主题连线粗细更加均匀，说明高分组会对多个话题进行思考，学习者与学习者交互的频率更高、范围更广。在第四个时间片段中，高分组的讨论热度明显高于低分组，且高分组相较于低分组更多关注更高层次的话题。



（a） 高分组认知网络演变图　　（b）低分组认知网络演变图

图 8 高、低分组认知网络演变图

5. 研究结论

5.1. 不同主题下的特征词共现情况可以揭示学生的讨论焦点和兴趣

通过分析学习论坛中不同主题下的特征词共现情况，可以发现学生们的讨论焦点和兴趣。共现网络中的核心节点代表了学生们广泛关注的词语，这些词语在讨论中具有重要地位。例如，在《教学微视频设计与开发》课程论坛中，“效果”（画面处理）“形状”（基本操作）“动画”（内容呈现）“特效”（画面处理）等词语是后期讨论的关键节点，显示出学生们对特效应用和动画制作的兴趣。共现网络分析有助于了解学生们的学习需求和兴趣点，揭示他们在具体应用和实践方面的探索。

在线学习平台的维护与开发团队也可以根据课程类型对平台进行个性化功能的设计与开发，并通过大数据分析技术来追踪和分析学生的学习行为，以提供更精确的在线学习资源。此外，平台应该优化界面设计和功能布局，以方便教师使用和管理，并且从中了解学生的学习情况，实现学生在线学习效率的最大化。

5.2. 学习论坛主题讨论的热度和深度随课程进展显著提升

学生通过碎片化的主题分享进行学习社区建构，论坛讨论主题的数量和时间呈现动态性和叠加效应的特点。学习论坛中主题讨论内容变化基本围绕课程进度展开，在课程的不同阶段，学习论坛的讨论热度和深度也有所变化。前期学习论坛多由教师引导，多以对视频创作软件的基础概念和使用方法的辨析与讨论为主，学生自主进行主题创造的意识较低。随着课程深入，学生积极性提高，学生开始自主创造和发布话题，主题下的话题数显著增长，讨论内容逐渐深入专业知识，与课程内容的推进趋势相吻合，说明学习论坛讨论主题的创造与课程的进度相关。论坛讨论深度也不断递进，从浅层专业术语的学习发展到更深层次的理解和实际应用。综上，当前学习论坛主题讨论内容大致沿着课程进度展开，与教师行为密切相关，整体讨论的热度和深度都有显著的上升。因此教师在组织论坛时应结合课程进度规划地进行主题创造，帮助学生巩固理解和产生更深层次的学习，鼓励学生创造和思考。

5.3. 学习论坛实现交互内容的质量提升和知识建构的层次提高

通过对学习论坛数据的深度挖掘，我们发现学习论坛作为学生互动和知识建构的平台，其学习论坛中交互内容的质量和学生知识建构的层次不断提高。论坛的讨论内容不仅覆盖了 AE 软件的基本操作和专业术语，还扩展到了教学微视频设计、实际应用、创意和素养等高层次认知层面。学生在论坛中的积极参与和深入讨论，进一步提升了他们的知识建构的层次。

教师应鼓励学生进行更多关于实际应用、创意和素养等高层次认知方面的讨论，培养他们的主动性和探索精神。在引导讨论时，教师应注意情境设置，促进学生之间的互动和辩论，凝聚集体智慧，以提升论坛内容的质量和知识建构的层次。通过分析论坛讨论的深度和广度，教师能够评估学生的学习效果和进展，进一步优化教学策略。

5.4. 学习论坛中学生间的认知水平存在差异

高、低分组在认知网络特征演变趋势相似但程度不同。随着课程的推进，高、低分组在参与讨论人数、讨论质量变化趋势和讨论内容及层次上都有差距。高、低分组学习者在交互频次和质量都呈上升趋势，但在关注主题方面，高分组学习者在讨论质量上始终保持较高的

认知水平。且高分组的学生在学习交流过程中关注的主题更加丰富，涉及的认知层次不同的主题，讨论热度逐渐升高，学习者整体交互趋势更加紧密。而低分组学习者在不同时间段讨论关注点会聚焦于单一层次的话题讨论，参与者人数较低，讨论热度不高，讨论主题涉及的认知层次也较低。

因此教师在引导讨论时，可以设计不同层次的话题，去报不同认知水平的学习者都能找到合适自己水平的讨论点，关注不同学生的认知构建情况，并根据他们的个人情况提出相应的建议和引导。还可以为不同层次的学生设定可达到的目标，激励他们积极参与讨论，帮助他们建立自信并提高思考和讨论能力，创建一个积极、包容的在线讨论环境。

5.5. 研究局限性和展望

然而，本研究仍然存在以下几点研究局限性：第一，数据收集方面，本间就专注于单一课程论坛可能限制了发现的普遍性，在未来的研究中可以尝试跨课程、跨学科或跨平台收集数据，以提高样本的代表性。第二，在数据分析上，本研究只针对了学生发帖、评论的文本内容，在未来的研究中，可以结合学习者在其他在线学习行为例如点击帖子、查阅帖子以及登录论坛次数等数据进行进一步分析文本语料中的更深层次信息。

致谢

本研究受“浙江省属高校基本科研业务费专项资金资助（项目编号：GB202302007）”“浙江工业大学大学生创新创业训练计划（项目编号：2023018）”资助。

参考文献

左明章、赵蓉、王志锋、李香勇 & 徐燕丽. (2018). 基于论坛文本的互动话语分析模式构建与实践.电化教育研究, (09), 51-58.

张思、陈娟、夏丹、王涛、王志锋 & 刘清堂. (2022). 在线论坛中学习者兴趣与行为主题联合建模研究. 远程教育杂志, (01), 81-90.

汤秋莲. (2014). 基于 BTM 的短文本聚类(硕士学位论文,安徽大学).

陈秀明、张晨晨、王峰 & 王先传. (2021). 基于 LDA 主题模型的 MOOC 评论回复特征维度分析. 阜阳师范大学学报(自然科学版), (04), 73-81.

柴艳妹 & 雷陈芳. (2018). 基于数据挖掘技术的在线学习行为研究综述. 计算机应用研究 35, (05), 1287-1293.

教育部. (2023). 教育部《教育部高等教育司 2023 年工作要点》的通知. http://www.moe.gov.cn/s78/A08/tongzhi/202303/t20230329_1053339.html

戴宇鹏. (2022). 基于 BTM 模型的人才需求分析——以人工智能产业为例(硕士学位论文,华东师范大学).

Hotho, A., Staab, S., & Stumme, G. (2003). *Ontologies improve text document clustering*. Third IEEE International Conference on, 541-544.

Sasaki, K., Yoshikawa, T., Furuhashi, T., & Twitter-TTM. (2014). *An efficient online topic modeling for Twitter considering dynamics of user interests and topic trends*. Joint International Conference on Soft Computing & Intelligent Systems & International Symposium on Advanced Intelligent Systems, 440-445.

Weng, J., Lim, E., Jiang, J., & He, Q. (2010). *Twitter rank:finding topic-sensitive influential twitterers*. In Proceedings of the third ACM international conference on Web search and data mining, 261-270.

多模态数据支持学习投入测评的系统性文献综述

Measuring student engagement based on multimodal data: A systematic review

曾佳尧[1]，杨现民[2]，李新[3*]
[123]江苏师范大学智慧教育学院
[*]lix2023@jsnu.edu.cn

【摘要】为揭示多模态数据在评估学习投入的现状特征及其关系，本文采用系统性文献综述法对国外近十年关于多模态数据支持学习投入测评的 40 篇英文文献进行了内容分析。研究结果显示：在现状特征方面，作者主要来自中国，研究对象多为本科生，样本规模普遍介于 25 至 50 人之间，实验持续时间多在 1 小时内或 1 至 6 个月内，且实验场景多为教室或在线平台。在数据类型及关系方面，测评时主要涉及 15 种不同类型的数据，其中视频数据、表情数据、自我报告和日志数据使用频率较高。进一步分析发现，情感投入测评中多使用表情数据、视频数据、自我报告以及皮肤电数据；在认知投入评估中，自我报告、视频数据、眼动数据和文本数据是主要的数据来源；日志数据、视频数据则是行为投入测评时最普遍选择的数据类型。因此，基于系统地研究现有文献，本综述以期为该领域未来的研究提供全面的参考与借鉴。

【关键词】行为投入；认知投入；情感投入；多模态数据；系统性文献综述

Abstract: *This systematic review examines 40 high-quality studies from the past decade, exploring multimodal data's role in learning engagement assessment. Research often originates from China, involves undergraduates, and spans sample sizes of 25-50. Experiments last from one hour to six months, typically in classroom or online settings. The review identifies 15 data types, highlighting video, expression, self-report, and log data. Affective engagement is mainly assessed via expression and galvanic skin data, cognitive through self-report and eye movement data, and behavioral through log and video data. It advocates for a holistic approach to fully capture the complexities of learning engagement in educational contexts. Thus, this review endeavors to offer a solid foundation for future research in learning engagement assessment.*

Keywords: behavior engagement, cognitive engagement, emotional engagement, multimodal data, Systematic Review Methodology

基金项目：2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）；国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”(项目编号：62077030)

1.研究背景

过去二十年来，研究者与教育工作者对学习投入测评表现出浓厚的兴趣(Bond et al., 2020; Zheng et al., 2023)。目前，多数学者认为投入是一个多维概念(Saqr et al., 2023)，普遍认为投入包括三个维度，即行为投入、情感投入和认知投入(Fredricks et al., 2004)。行为投入是指可观察到的行为，表征学生对学习的承诺、对学习任务的坚持和积极的参与(Lee et al., 2021)。情感投入涉及学生对学习材料的情感反应，从积极（如享受、快乐和兴趣）到消极（如无聊、沮丧、困惑和焦虑）等(Finn, 1989; Sinatra et al., 2015)。认知投入是指学习者学习过程的心理

投入程度，包含解决问题的风格、偏好的认知策略等(Fredricks, Blumenfeld & Paris, 2004)。在测评学习投入时，常用的方法包括自我报告、观察、访谈、教师评分(Azevedo, 2015)。尽管这些传统的测量方法具有一定的实用性和全面性，但无法捕捉微观层面的投入状态，而且通常需要耗费大量的成本(Li et al., 2021)。

多模态测评被视为多模态数据和学习分析领域的基础(Dubovi, 2022)，它能够对多种数据及其相互关系进行分析，更准确地描述学习过程中的学习投入情况(Sinatra et al., 2015)。然而，很少有研究系统地梳理多模态数据如何支持学习投入测评。Henrie et al. (2015)概述了投入的背景、定义和测量方法（包括定量自我报告、定性测量和定量观察测量），以及它们的优势和局限性。然而该综述仅关注传统的投入测量方法，且只涵盖 2015 年之前发表的实证研究。本研究将采用系统性文献综述的方法探讨以下研究问题：

(1)多模态数据测评学习投入的研究有哪些特点（国家、教育场景、参与者、研究对象、研究时长）？

(2)在分析学习投入时使用了哪些类型的多模态数据（行为投入、认知投入、情感投入）？以及多模态数据与学习投入之间的关系是什么？

因此，通过系统地研究现有文献，本综述旨在阐明多模态数据与学习投入之间的关系，探讨利用多模态数据测评学习投入的优点和不足，以期为该领域未来的研究提供全面的参考。

2.研究设计

为揭示多模态数据如何开展学习投入测评，本研究采用了系统性文献综述（Systematic Review Methodology）的研究方法。该方法旨在系统全面地在不同数据库中通过相应的检索技术获取与研究问题相关的文献，再结合一定的文献评价方法和过程筛选、归纳出符合标准要求的文献，并进一步对其进行客观的评价与分析，最终得出可靠的结论以系统规范地回答研究问题（胡晓玲等，2024）。

2.1. 文献检索策略

搜索过程包括三个阶段。在第一阶段，我们以“multimodal data AND behavior engagement OR emotional engagement OR learning engagement OR student engagement OR cognitive engagement OR affective engagement”为检索式进行精准检索，获取 2014 年 1 月至 2023 年 12 月发表的文章，在这段时期可穿戴设备和多模态学习分析技术的发展迅速（Prinsloo et al., 2023）。2023 年 9 月，我们在以下电子数据库中实施了第一轮检索： Elsevier ScienceDirect、Springer Link、EBSCOHost、Wiley 和 Web of Science (WoS)。2023 年 12 月，使用相同的检索式在五个数据库中实施了二次检索和滚雪球式检索。之后，进一步通过引文检索和脚注追踪的方式实施了第三轮检索，以确保全面覆盖多模态数据与学习投入的文章。

2.2. 文献筛选标准

初步搜索共获得 5651 篇文献。剔除 557 条重复记录后，共有 5128 篇文献可供筛选。要纳入本综述，研究必须符合特定的纳入和排除标准（表 1）。

表 1 文献排除与纳入标准

纳入标准	**排除标准**
文章出版在 2014 年 1 月至 2023 年 12 月之间	出版在 2024 年 1 月之前
文章属于期刊论文和会议论文集	研讨会论文、书籍和、海报和报告
文章用英语撰写	文章非英文撰写
文章属于实证研究	非实证研究
经同行评审并发表	非同行评审文章、存档文章与待刊论文

研究主题关注多模态数据与学习投入	非关注多模态数据和学习投入的研究

2.3. 文献筛选流程

研究的筛选过程如图 1（PRISMA 流程图）所示：(1)删除了 557 条重复记录(2)根据排除和纳入标准筛选标题和摘要后，排除了 4389 篇不相关的文章，然后(3)通过二次检索与滚雪球检索又确定了 19 条记录，共对 707 篇文章进行了全文评估，其中 500 篇因不符合纳入标准而被排除(4)通过数据库纳入了 32 篇文章，又通过其他检索途径纳入了 8 项研究。最后，本系统综述共确定了 40 篇文章。

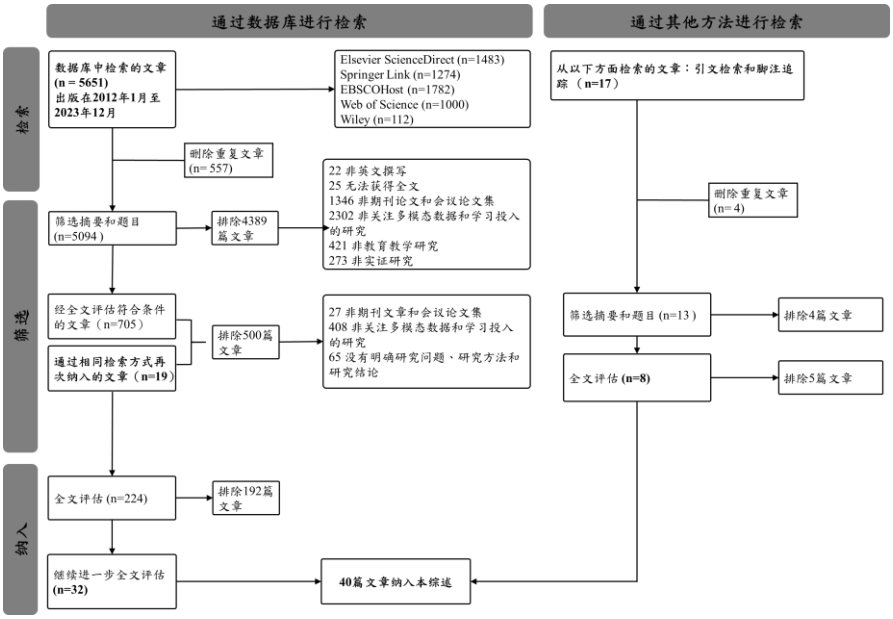


图 1 遵循 PRISMA 标准的文献分析流程图

2.4. 文献筛选流程

搜索和筛选程序结束后，两位研究人员对最终数据集进行了审核和编码。编码方案由研究人员根据两个研究问题共同设计（表 2）。随后由两位研究人员分别进行独立编码，之后进行交叉检验，编码一致性检验结果良好（Cohen Kappa=0.820）。

表 2 文献编码表

数据提取		**编码说明/示例**
问题一	国家	获取文章的出版国家。
	教学环境	例如，计算机化学习环境、移动学习环境、传统课堂等。
	参与者	例如：研究生、大学生、高中生、初中生等。
	学科	研究的学科领域。如计算机科学与编程、数学与统计等。
	样本大小	参与者数量。
问题二	投入类型	情感投入、认知投入或行为投入。
	投入指标	三种投入的指标，如情感状态、认知负荷、出勤率等。
	多模态数据	例如，视频、文本、音频、日志、眼动追踪、EEG、EDA 等。
	捕捉方法	例如，知识测试、量表、传感器、平台、网络摄像头等。

3.研究结论

3.1.多模态数据支持学习投入测评的特征分析

国家分布。本研究收录了来自 18 个不同国家的文章，其中大部分研究在中国进行（40%，n=10），其次是美国（16%，n=4）、加拿大（16%，n=4）、芬兰（16%，n=4）和印度（12%，

n=3）。从研究地点的地理分布来看，在亚洲进行的研究明显占多数（50%，n=20），其次是欧洲（35%，n=14）和北美洲（20%，n=8）。图 2(a)显示了本研究收录文章的地理分布情况，仅列出了八个国家。

参与者。有 38 项研究（图 2(b)）明确说明了样本的年级和人数。样本中最常见的研究群体是本科生（15 人），其次是小学生（7 人）、研究生（6 人）。从参与者年级分布来看，高中生（n=2）和初中生（n=2）很少被纳入研究，而幼儿园学生则完全没有被研究。

样本规模。样本中大多数研究（n=26）包含 50 个或更少的参与者，相对较少的研究（n=5）包含 100 或更多的参与者（图 2(c)）。

研究时长。在所有样本中，80%（n=32）的研究明确规定了研究时长（图 2(d)）。半数以上的研究（56.25%，n=18）时长在 0 到 2 小时之间，还有一部分研究（34.375%，n=11）的研究时间在 1 至 6 个月之间。此外，样本中有三项研究的持续时间至少为一年。

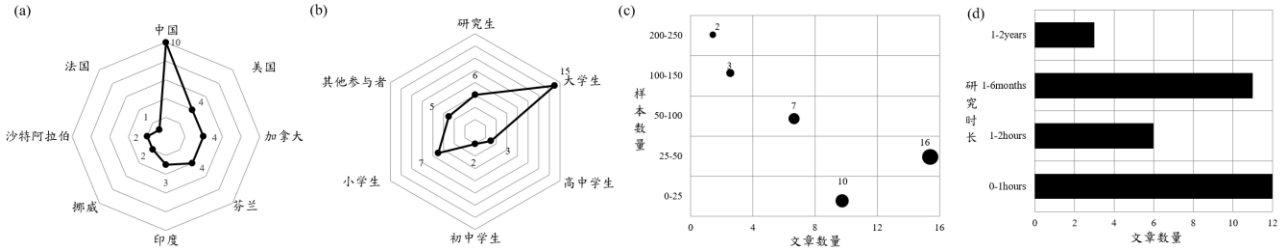


图 2 (a)地理分布, (b)参与者年级, (c)样本量, (d)研究长度

教学环境。表 3 列出了研究开展的环境，其中只有 35 项研究具体说明了教育环境。首先，有 10 项研究是在教室中实施，5 项是在传统教室、其余 5 项是在智能教室中进行。其次，12 项研究在线平台上进行的，有利于收集日志和最初成果等数据。第三，7 项研究在专业的实验环境中实施，这类环境在所有研究中使用的频率较低。最后，有 6 项研究是在实验室环境中进行的，这种方式有利于研究人员减少外界干扰并进一步收集更复杂的数据类型。

表 3 教学环境分布情况

教学环境		**文章数量**
教室(10)	传统教室	5
	STEM 教室	1
	教室(电子书)	1
	教室(iPad)	3
在线平台(12)	在线系统	4
	编程环境	4
	学习管理选题，智能导师系统	4
专业环境(7)	混合学习环境	3
	VR 环境	2
	具身学习环境	1
	配备机器人的交互界面	1
实验室(6)	实验室环境	6

研究主题。在 34 项明确提出研究主题的研究中（图 3），计算机科学是最受欢迎的研究主题（n=6），其次是数学（4 项）、开放式任务（4 项）和玩游戏（4 项）。另外 16 项研究涉及多个学科，包括解剖学、物理学、韩语、病理生理学和环境研究。

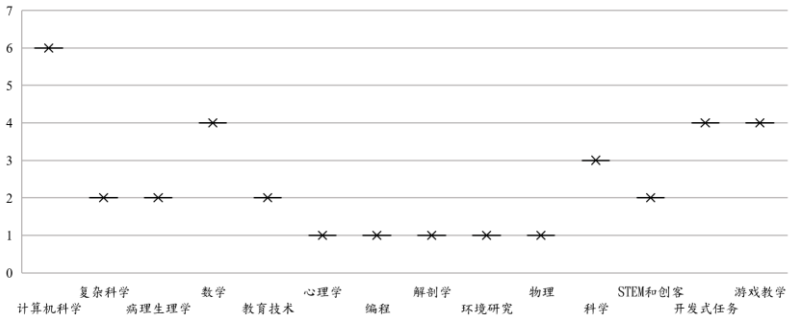


图 3 研究主题分布情况

3.2. 学习投入测评中的多模态数据类型及其关系

3.2.1. 学习投入测评中的多模态数据类型

分析编码表后发现（图 4，篇幅限制仅展示 20 篇文献）：在学习投入测评中，多模态数据为评估学习者的认知、情感和行为投入提供了全面视角。首先，文本数据可以通过工作表、学习成绩等形式，记录学习者的学习策略和自我调节能力，为评估认知投入提供重要依据。其次，音频数据可以记录学习者的语言交流和口头表达，通过分析语音的音量、音调和节奏等特征，可以揭示学习者的情感状态和学习行为。视频数据则能够通过捕捉学习者的眼神和身体动作等，直观地反映其行为交互和情感反应。最后，通过生理心理数据能够更为全面地监测学习者的内在反应，为投入的评估提供更为客观和精细的指标。例如，脑电图（EEG）可以反映学习者的认知负荷和注意力集中程度，皮肤电导反应（EDA）则可以揭示学习者的情绪变化和唤醒水平。依托这些数据，研究者能够更准确地评估学习投入，也为深入理解学生的学习过程提供了新的突破口。

研究文献		投入类型	投入指标	多模态数据	捕捉方法
1	Du et al. (2023)	认知投入	概念理解和 CPS 技能 认知负荷 解决问题的表现	文本（成绩）、问卷 EEG 文本（工作表）	知识测试 量表 传感器 平台
2	Ouyang et al. (2023)	认知投入 行为投入 情感投入	知识贡献 语言交流和在线行为 社会情感方面（如鼓励、倾听）	最终成果 视频 音频	平台 录屏（录音）
3	Alkabbany et al. (2023)	行为投入 情感投入	眼睛注视、头部运动 情感状态	凝视、头部姿势 面部表情	网络摄像头
4	Pabba et al. (2022)	情感投入	情感状态（如情绪和情感）	面部表情	网络摄像头
5	Ouyang et al. (2022)	认知投入 行为投入	对话 编程操作	音频 视频	录屏（录音）摄像机
6	Li et al. (2021)	认知投入	面部行为线索 现场认知投入、表演	视频（面部特征） 问卷，最终成果	网络摄像头 量表 平台
7	Moreno et al. (2020)	认知投入 情感投入	投入程度 唤醒、大脑活动和意识 投入者的身体反应	问卷 EDA, EEG 身体运动	量表 传感器 摄像机
8	Ballenghein et al. (2019)	认知投入	唤醒	眼动跟踪 运动	传感器 传感器
9	Dubovi. (2022)	情感投入 认知投入	积极/消极情绪 AOI（注意力）、学习经验 更高的脑力劳动	问卷、面部表情 眼动追踪 EDA	量表 网络摄像头 传感器 传感器
10	Zhu et al. (2019)	情感投入 认知投入	情感分析 想法改进（课堂讨论、在线讨论）	视频、文本（在线注释） 音频	摄像机 平台
11	Törmänen et al. (2021)	情感投入	情绪表达 唤醒状态、激活水平（激活/去激活）	视频 EDA	摄像机 麦克风 传感器
12	Mangaroska et al.(2022)	认知投入 情感投入 行为投入	注意、聚合思维、记忆和认知负荷 学生的面部表情 AOI、书写（使用键盘）、阅读	EEG, IPA 面部表情 注释,日志	传感器 网络摄像头 平台
13	Chango et al. (2021)	情感投入 行为投入 认知投入	从表情看情绪 学习策略 已有知识、注意力 (AOI)	面部表情 日志 文字（成绩）、凝视	网络摄像头 平台 知识测试 传感器
14	Zhidkikh et al. (2023)	认知投入 行为投入	自我调节能力、观念 学生行为模式	问卷、观察、访谈 日志	量表 平台
15	Giannakos et al. (2019)	行为投入	预测学业表现	键压 眼球跟踪 EEG, HR, EDA,体温,血压 视频（面部）	平台 传感器 网络摄像头
16	Mamieva et al. (2023)	情感投入	面部和语言特征	演讲 面部表情、凝视、头部姿势	数据库
17	Larmuseau et al.(2020)	认知投入	认知负荷	问卷 GSR,皮肤温度 (ST), HR, HRV, ECG	量表 传感器
18	Xu et al. (2023)	认知投入	学生对的语言交流和面部表情 操作行为、最终编程产品	音频、面部表情 点击流，最终成果	摄像机（录音） 录屏（录音）
19	Sharma et al. (2022)	情感投入 认知投入	面部情绪识别 注意力状态（分散或集中）	面部表情 眼部区域	网络摄像头
20	Hasnine et al. (2021)	情感投入	情绪检测 眼神检测	面部表情 眼睛位置, 眼动	录屏（录音）摄像机

图 4 学习投入测评中的多模态数据类型

3.2.2. 多模态数据与学习投入的关系

对 40 篇文献中的多模态数据系统分析后，发现研究者在实施学习投入测评时，通常选择如下 15 种类型的数据（图 4）。视频数据（n=22）是最为常用的数据类型，相较于其他数据，视频数据能够在耗费较低成本的前提下，更全面、隐蔽地捕获学习者的实时投入状态，因此这种数据类型被学者广泛用于学习投入的测评中。其次，面部表情数据（n=18）也是学习投入测评的关键数据类型，但这种数据主要指向情感投入的测评。第三种是自我报告（n=13），这种数据能主观反应学习者的内在认知状态，在学习投入测评过程中，学者通常会将自我报告作为认知投入测评的补充数据或主要数据来源。最后是日志数据（n=12），这种数据类型主要用于行为投入的测评，研究者能够基于学习平台全面地追踪和记录学习者的在线学习行为。进一步分析发现，由于不同的多模态数据在表征不同投入维度的有效性上存在差异，这导致在测评行为投入、认知投入、情感投入的数据分布上有所不同。具体而言，情感投入测评过程中，多使用学习者的面部表情数据，视频数据、自我报告以及皮肤电数据。测评认知投入时，则通常将自我报告、视频数据、眼动数据以及文本数据作为最主要的数据来源。而日志数据、视频数据则是行为投入测评时最多选择的数据类型，同时会使用心率、眼电图、音频数据等作为补充。整体看来，在实施学习投入测评时，多数研究仍倾向于选择更易采集和处理的数据类型，而采用传感器设备进一步捕捉复杂的生理心理数据（脑电，心率）开展投入测评的研究相对较少。

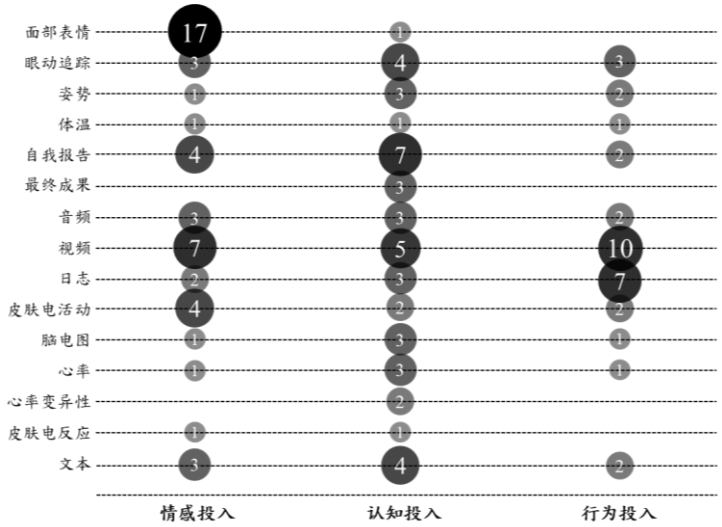


图 5 多模态数据与学习投入的关系

4.总结与展望

研究依托系统性文献综述法对国外近十年有关多模态数据支持学习投入测评的文献进行了编码分析，探讨了近十年多模态数据支持学习投入测评的研究在国家分布、参与对象、研究主题、教育场景以及实验时长和样本大小的现状特征；多模态数据的类型以及这些数据与不同维度学习投入之间的关系，为其他开展学习投入测评研究的学者提供了一定参考和借鉴。但是本研究仍存在一些局限。首先，由于本研究设置的检索时段仅为近十年且检索文献的数据库也只限于上述的五个英文数据库，导致纳入本综述的文章数量不够充分和全面。其次，搜索关键词也没有涉及所有与学习投入和多模态数据相关的研究，会导致部分文献遗漏。最后，文章只是较浅层地分析了多模态数据测评学习投入的现状以及两者的关系。未来还需要进一步探讨，多模态数据支持下的学习投入测评给我们带来了何种教育教学启示，以及基于多模态数据测评下的学习投入存在的问题和挑战。

参考文献

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. International journal of educational technology in higher education, 17, 1-30.
Zheng, L., Long, M., Niu, J., & Zhong, L. (2023). An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL. International Journal of Computer-Supported Collaborative Learning, 18(1), 101-133.
Saqr, M., López-Pernas, S., & Vogelsmeier, L. V. (2023). When, how and for whom changes in engagement happen: A transition analysis of instructional variables. Computers & Education, 207, 104934.
Fredricks, J. A. (2015). Academic engagement. In J. Wright (Ed.), The international encyclopedia of social and behavioral sciences (2nd ed., Vol. 2, pp. 31e36). Ox- ford: Elsevier.
Lee, Y., Jang, B. G., & Conradi Smith, K. (2021). A systematic review of reading engagement research: What do we mean, what do we know, and where do we need to go? Reading Psychology, 42(5), 540–576.
Finn, J. D. (1989). Withdrawing from school. Review of Educational Research, 59, 117-142.
Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. Educational psychologist, 50(1), 1-13.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. Review of Educational Research, 74, 59-109.
Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. Educational psychologist, 50(1), 84-94.
Li, S., Lajoie, S. P., Zheng, J., Wu, H., & Cheng, H. (2021). Automated detection of cognitive engagement to inform the art of staying engaged in problem-solving. Computers & Education, 163, 104114.
Dubovi, I. (2022). Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology. Computers & Education, 183, 104495.
Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. Educational psychologist, 50(1), 1-13.
Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. Computers & Education, 90, 36-53.
胡晓玲,韦慕春,袁民 张惠.(2024).教育领域的系统性文献综述：本质、价值与实现.电化教育研究(01),43-51.
Prinsloo, P., Slade, S., & Khalil, M. (2023). Multimodal learning analytics—In−between student privacy and encroachment: A systematic review. British Journal of Educational Technology, 54(6), 1566-1586.
Du, X., Dai, M., Tang, H., Hung, J. L., Li, H., & Zheng, J. (2023). A multimodal analysis of college students' collaborative problem solving in virtual experimentation activities: A perspective of cognitive load. Journal of Computing in Higher Education, 35(2), 272-295.
Ouyang, F., Xu, W., & Cukurova, M. (2023). An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. International Journal of Computer-Supported Collaborative Learning, 18(1), 39-66.
Alkabbany, I., Ali, A. M., Foreman, C., Tretter, T., Hindy, N., & Farag, A. (2023). An Experimental Platform for Real-Time Students Engagement Measurements from Video in STEM Classrooms. Sensors, 23(3), 1614.
Pabba, C., & Kumar, P. (2022). An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition. Expert Systems, 39(1), e12839.

Li, S., Lajoie, S. P., Zheng, J., Wu, H., & Cheng, H. (2021). Automated detection of cognitive engagement to inform the art of staying engaged in problem-solving. Computers & Education, 163, 104114.
Moreno, M., Schnabel, R., Lancia, G., & Woodruff, E. (2020). Between text and platforms: a case study on the real-time emotions & psychophysiological indicators of video gaming and academic engagement. Education and Information Technologies, 25, 2073-2099.
Ballenghein, U., Megalakaki, O., & Baccino, T. (2019). Cognitive engagement in emotional text reading: concurrent recordings of eye movements and head motion. Cognition and Emotion.
Dubovi, I. (2022). Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology. Computers & Education, 183, 104495.
Zhu, G., Xing, W., Costa, S., Scardamalia, M., & Pei, B. (2019). Exploring emotional and cognitive dynamics of knowledge building in grades 1 and 2. User Modeling and User-Adapted Interaction, 29, 789-820.
Törmänen, T., Järvenoja, H., & Mänty, K. (2021). Exploring groups' affective states during collaborative learning: what triggers activating affect on a group level?. Educational Technology Research and Development, 69(5), 2523-2545.
Mangaroska, K., Sharma, K., Gašević, D., & Giannakos, M. (2022). Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity. Journal of Computer Assisted Learning, 38(1), 40-59.
Chango, W., Cerezo, R., Sanchez-Santillan, M., Azevedo, R., & Romero, C. (2021). Improving prediction of students' performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. Journal of Computing in Higher Education, 33, 614-634.
Zhidkikh, D., Saarela, M., & Kärkkäinen, T. (2023). Measuring self‐regulated learning in a junior high school mathematics classroom: Combining aptitude and event measures in digital learning materials. Journal of Computer Assisted Learning.
Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. International Journal of Information Management, 48, 108-119.
Mamieva, D., Abdusalomov, A. B., Kutlimuratov, A., Muminov, B., & Whangbo, T. K. (2023). Multimodal Emotion Detection via Attention-Based Fusion of Extracted Facial and Speech Features. Sensors, 23(12), 5475.
Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., & Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. British Journal of Educational Technology, 51(5), 1548-1562.
Xu, W., Wu, Y., & Ouyang, F. (2023). Multimodal learning analytics of collaborative patterns during pair programming in higher education. International Journal of Educational Technology in Higher Education, 20(1), 1-20.
Sharma, P., Joshi, S., Gautam, S., Filipe, V., & Reis, M. J. Student engagement detection using emotion analysis, eye tracking and head movement with machine learning. arXiv 2019. arXiv preprint arXiv:1909.12913.
Hasnine, M. N., Bui, H. T., Tran, T. T. T., Nguyen, H. T., Akçapınar, G., & Ueda, H. (2021). Students' emotion extraction and visualization for engagement detection in online learning. Procedia Computer Science, 192, 3423-3431.

基于深度学习的学生课堂行为识别方法研究综述

Research Review of Deep Learning Recognition Method for Student Classroom Behavior

芃佩舒[1]，魏艳涛[2*]，肖光润[3]，潘东辉[4]，徐琦[5]，余游[6]
[1] 华中师范大学人工智能教育学部数字教育湖北省重点实验室
[*] yantaowei@ccnu.edu.cn

【摘要】学生课堂行为识别是理解学生学习状态的基础，对教师制定有效教学策略和提高学生参与度至关重要。本文针对深度学习在学生课堂行为识别领域的应用，从学生课堂行为的分类和数据来源，分别梳理了基于接触式数据和非接触式数据的学生课堂行为识别方法。研究发现接触式的学生课堂行为识别方法能够很好地克服环境遮挡等问题，但是采集数据不够方便；非接触地方法与之相反。最后，论文展望了在教育领域实施相关技术的趋势和潜在挑战，特别强调了深度学习模型在学生课堂行为识别中的应用前景。
【关键词】 课堂学习；行为识别；深度学习

Abstract: Identification of students' classroom behaviors is crucial for teachers to develop effective teaching strategies and improve student engagement. This article reviews the application of deep learning technology in the field of student classroom behavior recognition. Based on the classification of students' classroom behavior and data acquisition sources, it discusses the methods of student classroom behavior recognition based on contact data and non-contact data. The study finds that based on contact data Data-based student classroom behavior recognition methods can well overcome problems such as environmental occlusion, but data collection is not convenient enough, and methods based on non-contact data are contrary to this. Finally, it looks forward to the future research trends and possible challenges in implementing related technologies in the field of education, with special emphasis on the application prospects of deep learning models in classroom student behavior recognition.
Keywords: classroom learning, behavior recognition, deep learning

1.前言

课堂教学是学校教育教学的主阵地，对课堂教学进行过程性评价对于提高教学质量具有重要意义，而学生课堂行为表现是课堂教学评价的重要组成部分。准确识别和理解学生的课堂行为对于教师制定有效教学策略和提高学生参与度至关重要。传统的课堂行为分析主要依赖于教师在课堂上对学生的行为进行观察，或者在课后对课堂录像进行过程性分析。传统的课堂观察法和课堂录像分析都依赖人类观察者，这种方法存在着许多局限性，如分析结果主观性强 、无法实时反馈、遗漏重要信息等，因此很难扩展到大规模、长期性观测。然而，随着深度学习技术在计算机视觉领域的不断发展，教育信息化建设工作的不断深入，利用深度学习技术对学生课堂行为进行实时分析和识别变得愈发可行。

深度学习技术能够自主学习目标特征，并有效地克服噪声干扰，避免了人工设计目标特征的复杂性和低容错性，在教育信息化和智能化的大背景下研究学生课堂行为识别，可以帮助教师精准掌握学生的课堂学习情况、减轻评课负担，并有利于教师改进教学方法、调整教学策略，而且也便于学生了解自己上课时的学习行为和学习状态，并促使学生对此进行反思、

调整。

本文以深度学习技术为基础，结合各种不同数据采集方法（包括接触式和非接触式数据采集方法），综述了基于深度学习的学生课堂行为识别算法，通过对现有技术的分析，旨在为教育者和研究人员提供一个促进教学创新和学习效率的综合视角。

2.学生课堂行为分类

学生行为的识别和分析一直是理解课堂互动和教学效果的关键。无论是学生的言语行为还是非言语行为，都提供了深入了解教学过程和学习效果的窗口。经典的学生课堂行为分类法包括 FLAS 分类体系、CLASS 分类框架以及 S-T 分析法。FLAS 分类体系以言语行为分类为主，对学生课堂非言语行为关注度不够，CLASS 分类框架能体现更多课堂整体学习氛围，而难以建模学生具体的行为。艾伦等（2008）通过教育教学中质量监测的实例,说明了 S-T 分析、Flanders 课堂互动分析、数据包络分析、主成分分析、马尔可夫分析等测量工具的使用方法。其中 S-T 分析法和 Flanders 课堂互动分析这两个测量方法都属于行为测量，S-T 分析法可以对教学的“过程与方法”进行测量和评价,而 Flanders 课堂互动分析可以对教学中的“情感、态度和价值观”进行测量和评价。

其中的 S-T 分析法更适用于建模学生课堂行为分类，该方法是一种以图形的方法表示教学性格的分析方法，它不仅重视教师的教学行为，还同等重视学生的参与行为。S-T 分析法将教学中的行为仅分为学生（S）行为和教师（T）行为两类，减少了教学过程中行为分类的模糊性，增加了客观性。程云等（2016）从数据采集、统计及分析三方面对 S-T 分析法进行了改进，提出了一种基于改进型 S-T 分析法，其中学生行为可以分为言语行为和非言语行为。研究参考上述分类体系以及相关文献，根据数据源对学生课堂行为分类总结如表 1。

表 1　学生课堂行为分类和数据源

数据源	序号	表现形式	学生行为
非接触式数据采集方式	1	学生回答问题	言语行为
	2	学生沉默	言语行为
	3	师生课堂讨论	言语行为
	4	学生小组讨论	言语行为
接触式/非接触式数据采集方式	5	学生打盹	非言语行为
	6	学生思考、计算	非言语行为
	7	学生记笔记	非言语行为
	8	学生站立	非言语行为
	9	学生环顾四周	非言语行为
	10	学生搞小动作	非言语行为
	11	其他	非言语行为

3.学生课堂行为识别模型和算法应用

基于前述分析，学生课堂行为识别方法可以根据数据源的采集方法分为两类：基于接触式数据采集的学生课堂行为识别方法和基于非接触式数据采集的学生课堂行为识别方法。基于深度学习的学生课堂行为识别方法分类如图 1 所示。

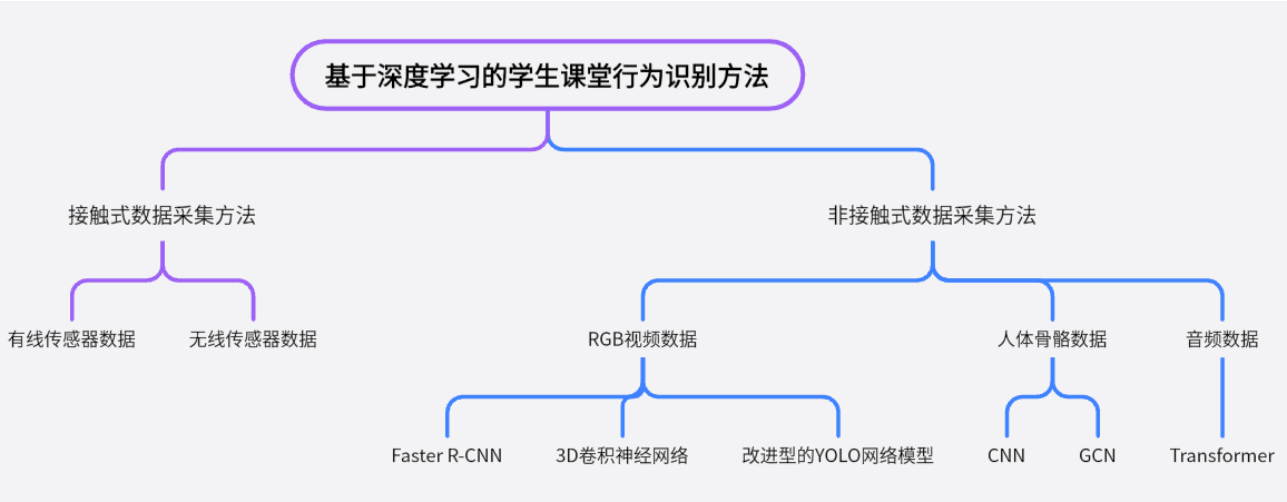


图 1 基于深度学习的学生课堂行为识别方法分类

3.1.基于接触式数据采集的学生课堂行为识别方法

传感器具有灵敏度高、体积小、存储方便、适用场景广等优点，不易受光线等环境条件、拍摄范围和角度以及环境遮挡等因素的影响。在目前的应用场景中，由于传统传感设备携带不便、续航能力有限的问题得到了有效解决，所以通过传感器获取人体活动数据已成为主流方法之一。根据采集数据时使用传感器的类型，对相关研究总结如表 2。

表 2 基于接触式数据的学生行为识别方法

	时间	作者	传感器类型	算法
有线传感器	2021	Uddin M Z	身体传感器	NSL
	2023	Wang H	运动传感器	VB-DTW
无线传感器	2020	El Attaoui A	心率无线传感器	多层感知器神经网络
	2020	Wu H	重力传感器	CNN

3.1.1 基于有线传感器数据的学生课堂行为识别方法

有线传感器一般采用串口通信将采集到的传感器数据导入本地服务器，如 Uddin 等（2021）使用 ECG 传感器采集心脏的数据、加速度计采集身体加速度、陀螺仪采集身体倾斜量、磁力计采集身体的磁场，通过将多个身体传感器放置在受试者身体的不同位置（胸部、右手腕和左脚踝），使用基于长短期记忆 (LSTM) 的稳健深度学习方法（NSL）进行学习，提出了一种基于身体传感器的活动建模和识别系统，在公共数据集 MHEALTH 上识别出了站立不动、坐着放松、躺下、步行、爬楼梯、腰部向前弯曲、手臂正面抬高、膝盖弯曲、骑自行车、慢跑、跑步和前后跳跃 12 项活动，实现了约 99%的召回率。Wang 等（2023）将加速度计数器、陀螺仪和温度传感器安装于参与者的脊柱中部和右肩，将采集到的数据通过计算机的 USB 端口存储在使用 Python 程序获取的 CSV 文件，针对采集到的传感器信号，使用基于投票的动态时间规整算法（VB-DTW），通过投票机制比较相邻视频片段之间的相似性，从而提取有效的动作片段，构建了一个基于运动传感器的智能系统，实现了对坐着不动、躺在桌面上、写笔记、在座位上举手、转身环顾四周、站起来时举手、在座位上摇摆、站和坐、游荡和躯干旋转、划拳、手中转笔、敲桌面、倚着身子聊天、抖腿，这 14 种常见的课堂行为的感知和识别。

3.1.2 基于无线传感器数据的学生课堂行为识别方法

无线传感器一般将采集到的数据通过无线网络或 GPRS 模块将数据发送到云端服务器或者本地存储，例如 El 等（2020）使用可抑制高频噪声的可穿戴无线传感器测量心电图信号，将收集到的数据传输到网关节点，网关节点能够通过物联网云将数据聚合到物联网平台中，

然后使用多层感知器神经网络提取 R 峰来确定心率的变化，提出了一种基于无线传感器网络与物联网（IoT）相结合的心电图远程监护方法。Wu 等（2022）以智能手机嵌入的重力传感器来作为数据采集设备，学生的行为会引起传感器的信号在正峰值和负峰值之间发生变化，通过使用卷积神经网络（CNN）来提取特征，然后分别对步行、上楼、下楼、躺着、坐着、站立和跑步七种学生行为进行分类，经过 1500 个学生行为数据样本的训练，所提模型分别实现了 97.83% 的准确率、97.82% 的精确率和 97.83% 的召回率。

综上所述，以接触式数据作为数据源的学生课堂行为识别方法利用各种传感器（如运动传感器、生理传感器等）来监测和分析学生的行为。传感器能提供高精度的数据，如心率、皮肤电反应等，这有助于准确识别学生的情绪和行为状态，而且相比于依赖光线和视觉的 RGB 视频，传感器数据受环境因素影响较小。但是对学生而言，长时间佩戴传感器可能不舒适，且生理数据的解读通常比视觉数据更复杂，需要更专业的知识。

3.2. 基于非接触式数据的学生课堂行为识别方法

在课堂行为识别中，非接触式数据的采集设备主要依赖于摄像机。RGB 摄像机采集来的数据是 RGB 视频或者图像数据，我们将采集过来的视频数据按照所处理模型的不同分为基于 R-CNN、基于 3D 网络和改进型的 YOLO 网络三种类型。用于人体行为识别研究的数据类型不仅仅局限于 RGB，深度相机的发展使骨骼数据的获取变得容易，基于骨骼数据进行学生行为识别的模型主要分为 CNN（Convolutional Neural Networks，卷积神经网络）、GCN（Graph Convolutional Network，图卷积网络）。一个富有成效和活跃的课堂通常以多种声音的存在为特征，例如学生在问答环节中与老师互动，学生在讨论中相互互动，基于音频数据进行学生课堂行为识别可以充分利用课堂音频数据。基于非接触式数据的学生课堂行为识别方法见表 3。

表 3 基于非接触式数据的学生课堂行为识别方法

	主体网络	时间	作者	算法
基于 RGB 视频数据	Faster R-CNN 网络	2020	Zheng	Faster R-CNN+OHEM
		2020	白捷	Faster R-CNN+Inception-ResNet-v2
	3D 卷积神经网络	2021	Liu	3D-MRDNet
		2022	Albert	3D-CNN+DCNN
	改进型 YOLO 网络	2022	Tang	YOLOv5+BiFPN+CBAM
		2022	Li	YOLOv5s+ESRGAN
		2023	Chen	YOLOv8+Res2Net
基于人体骨骼数据	CNN 网络	2020	何秀玲	OpenPose+CNN-10
		2020	闫兴亚	AlphaPose+MobileNetV3-Small
		2022	王泽杰	OpenPose+YOLOv3
	GCN 网络	2021	Shi	PifPaf+2S- AGCN
		2023	Niu	AlphaPose+2S- AGCN
基于音频数据	Transformer 网络	2022	Zheng	LA-TDNN
		2023	Liu	MAST

3.2.1. 基于RGB 视频的学生课堂行为识别方法

RGB 视频数据是一种常见的视频格式，在 RGB 视频数据中，每一帧图像由一个二维像素阵列组成，每个像素包含红、绿、蓝三种颜色的强度值，每个颜色通道通常有 8 位深度，意味着每个通道可以表示 256 个不同的强度级别，从而总共可以表示大约 1670 万种不同的颜色。在提取学生的关键特征前，几乎所有的行为识别方法都会进行某种形式的图像预处理，

例如图像裁剪、缩放、标准化或增强，以提高模型的识别准确性。而且采用 RGB 视频数据的行为识别不仅关注单个图像帧，还涉及对时间序列数据的分析，以识别行为的动态变化。

（1）Faster R-CNN 模型

Faster R-CNN 的起源可以从 R-CNN 追溯起。R-CNN 的识别过程首先识别出视频中的潜在兴趣区域，然后对这些区域进行更深入的分析，但是 R－CNN 检测算法由于各候选区域不共享特征且区域重复率较高，导致训练时间较长，内存占用率高。后来的 Fast R-CNN 抛弃了 SVM，使用 Softmax 做分类，使得模型生成候选框之后的所有步骤可以进行端到端训练，且不需要额外的空间来存储中间特征。接着 Faster R-CNN 提出了 anchor 的概念，整个目标检测流程可以端到端训练，至此整个算法才真正算是“基于深度学习的目标检测算法”。Faster R-CNN 网络结构在课堂学生行为检测中是一种常用的学生课堂行为识别算法。

白捷等（2020）选用 Faster R-CNN 作为目标检测框架，应用基于预训练的 Inception-ResNet-v2 网络进行特征提取，提出了一种基于多路特征融合的 Faster R-CNN 学生课堂行为识别方法，实现了对正常学习、睡觉、低头等课堂行为的识别，模型的识别准确率可达 76.32%。Zheng 等（2020）通过使用规模感知头克服规模变化，利用 OHEM（在线硬示例挖掘）来缓解类的不平衡，提出了一种改进的 Faster R-CNN 模型用于学生行为识别，自动检测出举手、站立、睡觉等学生课堂行为。

（2）3D 卷积神经网络模型

随后有学者提出了直接提取时空特征的 3D 卷积神经网络（3D-CNN）方法。在 3D 卷积神经网络中，卷积核不仅在宽度和高度上移动，还在深度上移动。这意味着卷积核会考虑相邻帧之间的信息，3D 卷积神经网络将 2D 卷积神经网络（2D-CNN）扩展到了时空域，Ji 等人（2012）最早提出 3D 卷积神经网络（3D Convolutional Neural Network, 3DCNN）进行行为识别。随着 3D 卷积神经网络逐渐迁移到学生课堂行为识别领域中，它对于捕捉快速和微妙的动作特别有效，这在学生课堂行为识别中非常重要。Liu（2021）采用 3 维多尺度残差密集块作为网络的基本模块，提出了来自异构视图感知的 3D 多尺度残差密集网络（3D-MRDNet），该方法成功地识别出了包括上课、睡觉、玩手机、记笔记、环顾四周、阅读这六个经常出现在课堂上的学生行为，精度比传统的 C3D 网络识别准确率提高了 4.01%。Albert 等（2022）使用深度卷积神经网络（DCNN）获得每个学生的实时图像流，深度时空残差卷积神经网络学习每个学生行为的时空特征，提出了一个结合跟踪和目标检测技术的 3D-CNN 模型，实现了多学生的举手、弯腰、来回走动、在黑板上写字、抬头、低头、站立、躺在课桌上等行为的实时识别，识别准确率达到了 93.43%。

（3）改进型的 YOLO 网络模型

YOLO（You Only Look Once）网络是一种快速的对象检测系统，可以实时识别视频中的多个对象和行为。这种网络通过单次分析整个视频帧来识别和分类行为，从而大大提高了处理速度。这对于需要实时监测和响应学生行为的应用来说非常有价值。现有的一系列 YOLO 检测算法运行速度快，多用于实时检测，然而难以处理密集目标，小目标检测效果一般，无法满足现有课堂上学生行为准确识别的需求，为了解决以上难题，基于改进型的 YOLO 网络在学生课堂行为识别中的应用应运而生。为了解决小目标的检测能力，Zhang 等（2020）通过在原有 YOLOv3 的快捷结构中插入注意力机制 CBAM 模块，保证了学生课堂行为中的有效特征能够快速有效地学习，改进的 YOLO-CBAM 算法提高了小目标的检测能力。为了提高检测精度，提高遮挡对象的区分度，Tang 等（2022）将原始 YOLOv5 模型颈部网络中的特征金字塔结构（FPN+PAN）与加权双向特征金字塔网络（BiFPN）相结合，然后在颈部网络和预测网络之间加入空间和通道卷积注意机制（CBAM），使模型聚焦于目标信息。为了解决教室视频图像中的对象密集、遮挡和多尺度场景等挑战，Chen（2023）等将 Res2Net 模块与 YOLOv8

中的原始 C2f 模块集成，创建了一个新颖的 C2f_Res2block 模块，通过引入多尺度注意力模块（EMA）和多头自注意力（MHSA）模块，提出了一种改进的 YOLOv8 模型。该模型在识别举手、阅读和书写三中常见的学生课堂行为中与原始 YOLOv8 相比，平均识别准确率提高了 4.2%。Li 等（2022）在 YOLOv5s 的基础上，提出了一种带有 ESRGAN（增强型超分辨率生成对抗网络）和微小物体检测模块的 ET-YOLOv5s。该方案可以对学生在课堂上包括玩手机、上课、捂脸坐着、坐右转、坐左转、鞠躬、睡觉等 11 种行为检测，与 YOLOv4、YOLOv5s 等微小目标检测算法相比，检测性能更好，并且可以有效处理多个学生的行为识别。

以上基于 RGB 视频数据的学生课堂行为识别方法均使用普通的摄像机捕捉视频数据，由于 RGB 摄像机普遍存在于现代教室中，具有很好的便捷性，而且 RGB 视频可以提供丰富的视觉信息，可以更准确地识别学生的行为和表情。但是，课堂环境通常复杂多变，可能包含多种干扰元素，光线、摄像机角度和距离等因素都可能影响数据的质量和行为识别的准确率，而且处理和分析大量视频数据需要很高的计算资源。

3.2.2. 基于人体骨骼数据的学生课堂行为识别方法

与跑步、行走等日常行为不同，大多数课堂行为都是坐姿（例如听、看、玩手机），这些动作特征相对接近，并存在遮挡问题。随着 Kinect、Astra 等深度相机和传感器的开发，骨骼数据的获取变得容易，用于人体行为识别研究的数据类型不仅局限于 RGB 数据，为了应对在学生课堂行为识别中存在遮挡现象、需要较高的计算资源等难题，采用人体骨骼数据来进行学生课堂行为识别的方法应运而生。

骨骼关键点特征利用人体的关键关节部位的相对位置信息构建特征。通过检测骨骼关键点位置，在主流的人体关键点检测数据集 MSCOCO 中把人体骨架关键点划分为 17 个，分别是鼻子、左右眼、左右耳、左右肩、左右肘、左右 手、左右腰、左右膝、左右脚踝，形成骨架图描述人体姿态行为。与基于 RGB 图像的人体行为识别方法相比，基于骨骼信息的人体行为识别方法不受背景、人体外观等干扰因素的影响，具有更高的准确性、鲁棒性和计算效率。

目前基于多人姿态估计的方法大致可分为自上而下和自下而上两类。自上而下的方法是先检测出每一个人体，再对单个人体进行骨骼关键点检测，从而实现姿态估计，检测效果较好，但不同部位检测难度不同。自下而上的方法是先检测出所有的关键点，再将骨骼关键点聚类成多个个体，因此只需要检测一次，检测速度有很大的提升且准确率也高。由于本文主要介绍在课堂环境下处理学生骨骼数据的方法，所以就不再系统的介绍流行的姿态评估方法。

（1）基于 CNN 的方法

随着深度神经网络的兴起，许多工作都使用 CNN 网络来解决这个问题。CNN 模型被设计用于处理这些图像并通过分析不同帧中骨骼关节点来识别特定的动作。因此骨骼动作序列，常常被输入到 CNN 模型中进行动作识别。何秀玲等（2020）选取 OpenPose 骨架检测模型提取学生行为的骨架特征，结合 CNN－10 实现学生行为的有效识别，该方法能够有效排除学生体质、着装、课堂背景等无关信息的干扰，突出关键有效信息，在包括举手、听课、东张西望、阅读、书写、起立、睡觉 7 种典型的课堂行为中平均识别准确率达到了 97.92%。闫兴亚等（2020）通过改进的 YOLOv5 目标检测算法得到人体检测框，并设计改进轻量级姿态识别模型 MobileNetV3-Small，得到了 AlphaPoseMobileNetV3-Small，将得到的人体关键点图输入到线性分类器中，提出了基于深度学习的姿态识别方法 BetaPose，针对正常、传纸条、低头偷看、东张西望、打瞌睡 5 种学生课堂行为，在简易和拥挤场景的识别准确率分别达到 91%和 85%以上。王泽杰等（2022）利用 Open Pose 算法提取人体姿态的全局特征，融合 YOLOv3 算法提取交互物体的局部特征，对学生行为进行了识别分析，在正坐、侧身、低头和举手这 4 种与学习投入状态紧密相关的行为中，识别准确率达到了 95.45%。

（2）基于 GCN 的方法

GCN 为识别和处理人体骨架数据提供了一个强大的框架，人体的每个关节（如手腕、肘部、膝盖）被视为图中的一个节点，而关节之间的物理连接被视为边，GCN 可以有效地捕捉这些关节之间的空间关系，从而识别出特定的行为和动作，因此骨骼动作序列常常使用 GCN 模型进行动作识别。Shi 等人（2021）以课堂视频为输入，通过 PifPaf 模块提取学生骨架信息，然后采用 2s-AGCN 图卷积网络提取关节和骨骼两流特征，最后利用关系网络进行零拍学习，提出了一种基于图卷积的学生课堂行为识别方法，对听课、玩手机、阅读或记笔记这三种学生课堂行为的识别准确率达到了 60.67%。Niu 等（2023）使用改进的 YOLOv5 方法提取目标前景，然后使用 Alphapose 生成骨架信息，最后采用 2S-AGCN（双流自适应图卷积网络）技术作为特征提取网络，识别相关动作，针对举手、起立、打盹、学习 4 种课堂行为，识别准确率分别达到了 75%、96%、62%、72%。

综上所述，与 RGB 数据相比，骨架由于其低维表示而具有对背景变化的鲁棒性和高效计算的优点，能更好地避免噪声影响，且骨骼数据信息较为集中，大大减少了冗余信息的计算。但是提取和处理骨骼关键点数据对骨骼关键点的追踪精度要求较高，任何追踪误差都可能导致行为识别的不准确。

3.2.3. 基于音频数据的学生课堂言语行为识别方法

采用 RGB 视频数据和骨骼数据作为数据来源时，由于课堂场景复杂，以及师生互动的隐蔽性，有效地定位正在交互的学生是不稳定的，因此挖掘课堂音频特征，利用课堂内师生语音交互的显著性数据，可以辅助了解课堂状态，更加高效地提高学生课堂行为识别的效果。

Zheng 等（2022）对言语课堂互动行为分析任务中的说话人分割聚类算法进行了改进，通过在传统的 X 向量特征提取网络 TDNN 上集成 LSTM 来捕获时间信息，提出了一种新的说话人特征提取网络架构——LA-TDNN，同时使用增加了双多头注意机制的池化层以提高说话人特征之间的辨别能力。最后对师生之间的简单讨论、师生之间的复杂讨论等课堂讨论，以及学生之间的小组讨论、向全班展示自己的观点等互动交互行为进行分类，取得了不错的效果。Liu 等（2023）通过对原始语音信号进行采样和预处理，生成多通道图谱，然后利用多头自注意机制对音频特征进行全局建模并与时频富集模块组合，将输出映射到类别特征图，从而实现课堂音频类别识别，提出了多尺度音频频谱图 Transformer（MAST）语音场景分类算法，在包含教师教学（单声、讲课）、师生互动（师生交流、小组讨论）和沉默（无声的场景）的自建课堂音频数据集 CSID 上的识别准确率达到了 92.1%。

综上所述，音频数据具有体积小、处理和存储的需求较低的优势。可以通过分析语音的音调、强度和节奏，对学生的情感状态和参与度进行评估，通过识别关键词和短语来理解学生的学习内容和，通过分析学生的语音活动频率和互动情况，评估他们的参与度和课堂互动情况。

4. 研究展望和挑战

4.1.研究展望

随着深度学习技术在学生课堂行为识别领域的发展，一些主要的研究趋势和研究方向如下：

1) 多模态数据的融合：结合不同类型的数据，如视频、音频、文本以及生理信号数据，可以提供更全面的学生行为分析。未来研究可能会探索如何更有效地融合和分析这些多模态数据。
2) 数据集的构建和扩展：目前还没有丰富的公开的学生课堂行为数据集，为了训练更有效的模型，需要更多高质量和多样化的数据集。未来研究可能会关注于收集和构建更广泛的公共的学生行为数据集，以提高模型的泛化能力和准确性。
3) 轻量级的深度学习框架：在设备上运行轻量级分类模型，由于其低资源需求和高效

性，在实时课堂分析中具有潜力。

这些趋势和研究方向展示了学生课堂行为识别领域未来的发展潜力。随着技术的进步和研究的深入，有望在未来为教育带来更多的创新和改进。

4.2.研究挑战

基于深度学习的行为识别技术在教育领域的应用虽然充满潜力，但也面临着诸多挑战。以下是一些主要的挑战：

1) 行为识别的准确表征：尽管目前研究者已经将多种行为识别算法迁移到学生课堂环境中，学生行为识别的精确度和速度在不断的提高，但是目前仍然无法充分满足在课堂中的应用，所以如何提高学生行为的更好表征仍然是一个首要的挑战。不断的改进模型，不断的探索出新的深度学习方法，将最新的识别技术应用到学生行为识别领域，为教育赋能是一个可行性的解决策略。
2) 学生的接受度：采用深度学习技术进行学生的行为识别时，无论是在研究阶段还是应用阶段，在采集数据时都涉及到设备与学生的交互，基于接触式的设备进行数据采集需要学生长时间佩戴相关的传感器，基于非接触式的设备进行数据采集需要对学生进行拍摄，二者都可能会引起学生身体上产生不适和心理上产生反感、排斥情绪。
3) 隐私和伦理问题：行为识别技术的应用涉及到大量的个人数据，如何确保这些数据的安全和隐私是一个严重的问题。将采集到的个人的数据进行加密，妥善保存，提高安全技术是解决此问题的关键。

总体而言，深度学习和行为识别技术在教育领域的应用需要综合考虑多方面的挑战。

5. 结论

本文以基于接触式和非接触式两种数据采集方式为切入点，综述了基于深度学习的学生课堂行为识别方法。

在以接触式数据采集方法进行数据采集的研究方面，学生课堂行为识别方法依赖各类传感器（如运动传感器、生理传感器等）提供的数据进行监测和分析学生的行为，这些传感器能提供高精度的数据，如心率、皮肤电反应等，为后续算法的准确识别提供数据支撑。然而，学生长时间佩戴传感器可能不舒适，且生理数据的解读通常较为复杂。在以非接触式数据采集方法进行数据采集的研究方面，学生课堂行为识别方法利用各类摄像机和音频采集器（如高清摄像机、深度摄像机和录音笔等）提供的数据来监测和分析学生的行为，最小程度地影响学生的正常课堂行为。

本文的核心在于分析了利用深度学习技术进行学生课堂行为识别的多种方法，同时，文章指出了该领域面临的挑战和未来的发展方向，通过深入探讨，为教育者和研究人员在学生行为识别领域的应用提供了有价值的参考。

基金项目

国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号 CCNU22JC011）

参考文献

朱煜,赵江坤,王逸宁 & 郑兵兵.(2016).基于深度学习的人体行为识别算法综述. *自动化学报* (06),848-857. doi:10.16383/j.aas.2016.c150710.
艾伦, & 艾霁野. (2008). 基础教育质量监测与评价的测量工具研究. *中国教育技术装备*, (6), 1-7.
程云, 刘清堂, 王锋, & 王艳丽. (2016). 基于视频的改进型 ST 分析法的应用研究. *电化教育研究*, *37*(6), 90-96.
白捷, 高海力, 王永众, 杨来邦, 项晓航, & 楼雄伟. (2020). 基于多路特征融合的 Faster R-CNN 与迁移学习的学生课堂行为检测. *《广西师范大学学报》(自然科学版)*, *38*(5), 1-11.
何秀玲, 杨凡, 陈增照, 方静, & 李洋洋. (2020). 基于人体骨架和深度学习的学生课堂行为识别. *现代教育技术*.
闫兴亚, 匡娅茜, 白光睿 & 李月. (2023). 基于深度学习的学生课堂行为识别方法. *计算机工程* (07), 251-258.
王泽杰, 沈超敏, 赵春, 刘新妹, & 陈杰. (2022). 融合人体姿态估计和目标检测的学生课堂行为识别. *华东师范大学学报 (自然科学版)*, *2022*(2), 55.
Flanders, N. A. (1968). Abstracted from Interaction Analysis in The Classroom A Manual for Observers. *Classroom Interaction Newsletter*, *3*(2), 1-5.
Uddin, M. Z., & Soylu, A. (2021). Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning. *Scientific Reports*, *11*(1), 16455.
W Wang, H., Gao, C., Fu, H., Ma, C. Z. H., Wang, Q., He, Z., & Li, M. (2023). Automated Student Classroom Behaviors' Perception and Identification Using Motion Sensors. *Bioengineering*, *10*(2), 127.
El Attaoui, A., Hazmi, M., Jilbab, A., & Bourouhou, A. (2020). Wearable wireless sensors network for ECG telemonitoring using neural network for features extraction. *Wireless Personal Communications*, *111*, 1955-1976.
Wu, H., & Ma, L. (2022). The construction of online course learning model of ideological and political education for college students from the perspective of machine learning. *Security and Communication Networks*, *2022*, 1-9.
Zheng, R., Jiang, F., & Shen, R. (2020, May). Intelligent student behavior analysis system for real classrooms. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 9244-9248). IEEE.
Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(1), 221-231.
Liu, H., Liu, Y., Zhang, R., & Wu, X. (2021). Student behavior recognition from heterogeneous view perception in class based on 3-D multiscale residual dense network for the analysis of case teaching. *Frontiers in neurorobotics*, *15*, 675827.
Albert, C. C. Y., Sun, Y., Li, G., Peng, J., Ran, F., Wang, Z., & Zhou, J. (2022). Identifying and Monitoring Students' Classroom Learning Behavior Based on Multisource Information. *MOBILE INFORMATION SYSTEMS*, *2022*.
Tang, L., Xie, T., Yang, Y., & Wang, H. (2022). Classroom Behavior Detection Based on Improved YOLOv5 Algorithm Combining Multi-Scale Feature Fusion and Attention Mechanism. *Applied Sciences*, *12*(13), 6790.
Li, L., Liu, M., Sun, L., Li, Y., & Li, N. (2022). ET-YOLOv5s: toward deep identification of students'

in-class behaviors. *IEEE Access*, *10*, 44200-44211.

Chen, H., Zhou, G., & Jiang, H. (2023). Student Behavior Detection in the Classroom Based on Improved YOLOv8. *Sensors*, *23*(20), 8385.

Niu, W., Sun, X., & Yi, K. (2023, January). Improved YOLOv5 for skeleton-based classroom behavior recognition. In *Third International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI 2022)* (Vol. 12509, pp. 107-112). SPIE.

Shi, B., Wang, L., Yu, Z., Xiang, S., Liu, T., & Fu, Y. (2021, November). Zero-shot learning for skeleton-based classroom action recognition. In *2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC)* (pp. 82-86). IEEE.

Zheng, Z., & Huang, Y. (2022, August). Research on Classroom Interaction Behavior Analysis Algorithm based on Audio and Video. In *Proceedings of the 5th International Conference on Control and Computer Vision* (pp. 127-133).

Liu, F., & Fang, J. (2023). Multi-Scale Audio Spectrogram Transformer for Classroom Teaching Interaction Recognition. *Future Internet*, *15*(2), 65.

基于学习者反馈文本分析提升在线编程课程学习参与度研究

A Study on Enhancing Learning Engagement in Online Programming Courses Based on Learner Feedback Text Analysis

陈协玲[1*]，李冬梅[1]
[1]广州大学教育学院
[*] xielingchen0708@gmail.com

【摘要】随着教育数字化的加速推进，在线学习呈现出更丰富多元的实践探索。而学习者课程参与度已逐渐成为衡量教育质量的新标准。了解影响学习者在线学习参与度的因素，将有利于形成促进在线学习课堂的高质量发展策略。本研究将中国大学 MOOC 平台中《程序设计入门——C 语言》的课程反馈评价作为研究对象，采用内容分析法，从情感取向和评价类别两个角度展开分析该 14419 条学习者反馈评价数据。本研究旨在探究提升学习者在线课程学习参与度的主要因素及策略。

【关键词】在线学习；学习者参与；影响因素
【基金项目】国家自然科学基金青年项目“基于自然语言分析的在线学习社区认知参与评测与干预研究”(62307010)

Abstract: *With the acceleration of education digitisation, online learning presents richer and more diversified practices to explore. Learner course engagement has gradually become a new standard for measuring education quality. Understanding the factors affecting learners' engagement in online learning will be conducive to the formation of high-quality development strategies for promoting online learning course. In this study, we take the course feedback evaluation of Introduction to Programming - C Language in the MOOC platform of Chinese universities as the research object, and use the content analysis method to analyse the 14419 learner feedback evaluation data from two perspectives, affective orientation and evaluation category. This study aims to explore the main factors and strategies to enhance learners' engagement in online course learning.*

Keywords: Online learning, Learner engagement, Influencing factors

1. 引言

随着教育数字化的加速推进，在线学习呈现出更丰富多元的实践探索(王志军，2024)。我国教育部在 2017 年 2 月印发的《2017 年教育信息化工作要点》明确了在线开放课程建设的必要性和重要性(何向阳，2023)。而学习者课程参与度已逐渐成为衡量教育质量的新标准(赵晓阳，2012)。了解影响学习者在线学习参与度的因素，有利于形成促进在线学习课堂的高质量发展策略，提升学习者学习兴趣，营造良好的在线学习氛围。基于此，本研究将中国大学 MOOC 平台中《程序设计入门——C 语言》的课程反馈评价作为研究对象，采用内容分析法，从情感取向和评价类别两个角度展开分析该 14419 条学习者反馈评价数据。本研究旨在探究提升学习者在线课程学习参与度的主要因素及策略。

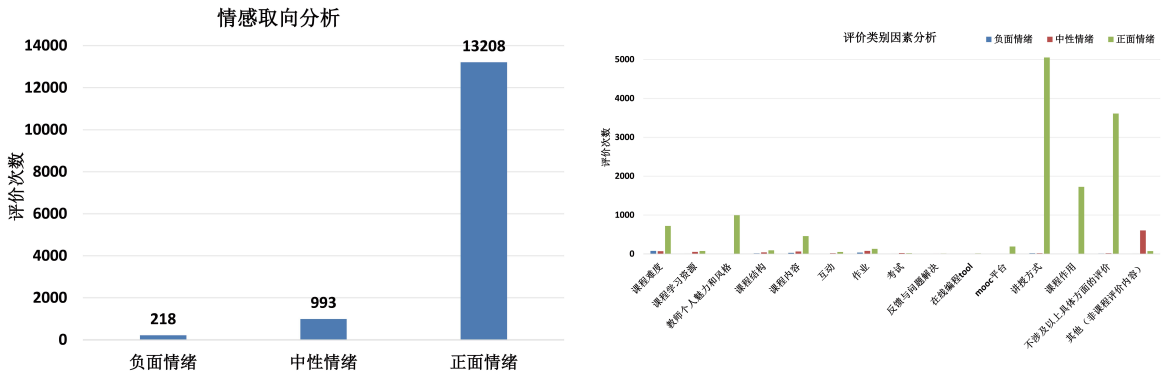
2. 研究设计

为分析影响学习者对该门在线课程学习参与度的主要因素，本研究对已完成、部分完成以及正在学习课程的学习者的公开评价进行定性研究，从中提取到 15 个评价类别，如表 1 所示。

表 1 15 个评价类别

代号	评价类别	代号	评价类别	代号	评价类别
1	课程难度	6	互动	11	MOOC 平台
2	课程学习资源	7	作业	12	讲授方式
3	教师个人魅力和风格	8	考试	13	课程作用
4	课程结构	9	反馈与问题解决	14	不涉及以上具体的评价
5	课程内容	10	在线编程 Tool	15	其他（非课程评价内容）

3. 研究发现



(a) (b)

图 1 (a)情感取向分析与(b)评价类别因素分析

3.1. 从情感取向角度展开分析

本研究从情感取向角度分析该 14419 条学习者反馈评论，以提取学习者看待课程的情感分类结果。得出结果如图 1(a)所示，其中 13208 条数据体现正面情绪（92%），993 条数据体现中性情绪（7%），218 条数据体现负面情绪（1%）。可见，正面情感比例较高，大部分学习者认可该课程，学习者的在线课程学习参与度较好。

3.2. 从评价类别角度展开分析

本研究从评价类别角度分析该 14419 条学习者反馈评论，以提取学习者认为课程有待改进或引人入胜的因素。得出结果如图 1(b)所示，可见，影响学习者在线编程课程学习参与度的主要因素集中在教师的讲授方式和个人魅力，以及课程本身的内容与难度。

3.3. 相同评价类别的不同情感取向分布情况

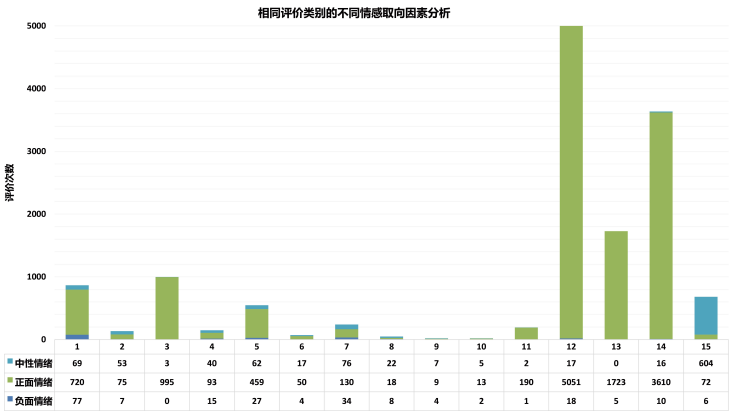


图 2 相同评价类别的不同情感取向因素分析

本研究对相同评价类别的不同情感取向的反馈评价进行统计整合，呈现数据如图 2 所示。本研究通过数据分析，从正负面文本中，分别提取前三个占比最高的评价类别，作为影响学习者在线学习参与度的因素进行分析，并通过 jieba 分词进行关键词提取，采用百度中文停用词进行过滤，将最终提取到的前 100 个关键词进行词云图分析展示，字号大小和频次成正比。

3.3.1. 消极因素分析

根据数据分析可得，影响学习者在线编程课程学习参与度的前三个消极因素为（1）课程难度（35%）；（2）布置的作业和练习（16%）；（3）课程内容（12%）。



(a) (b) (c)

图 3 负面评价中关于(a)课程难度、(b) 作业和练习和(c) 课程内容关键词词云图

（1）课程难度

在关于课程难度的 77 条负面评价文本中，词云图如图 3(a)所示，频次较高的是“难”等词。可得，该课程对部分学习者而言，仍存在一定的难度，会出现看不懂、听不懂的情况。

（2）布置的作业和练习

在关于作业练习的 34 条负面评价文本中，词云图如图 3(b)所示，频次较高的词是“难”、“少”等词，体现作业习题存在难度挑战、没有答案参考，缺乏讲解以及题量不足等情况。

（3）课程内容

在关于课程内容的 27 条负面评价文本中，词云图如图 3(c)所示，频次较高的是“理解”、“乱”等词，体现部分同学认为课程难理解、混乱。

3.3.2. 积极因素分析

根据数据分析可得，影响学习者在线编程课程学习参与度的前三个积极因素为（1）教师的讲授方式（38%）、（2）课程所体现的作用（13%）、（3）教师个人魅力和风格（8%）。



(a) (b) (c)

图 4 正面评价中关于(a)讲授方式、(b)课程作用和(c)教师个人魅力关键词词云图

（1）教师的讲授方式

在关于教师的讲授方式的 5051 条正面评价文本中，词云图如图 4(a)所示，频次较高的是“好”、“通俗易懂”等词，说明大部分学习者认为该课程授课老师讲得好、详细，深入浅出、通俗易懂。

（2）课程作用

在关于课程作用的 1723 条正面评价文本中，词云图如图 4(b)所示，频次较高的是“学习”、“收获”等词，体现该课程对于大部分学习者而言，有所收获，是一门好课程。

（3）教师个人魅力和风格

在关于教师个人魅力和风格的 995 条正面评价文本中，词云图如图 4(c)所示，频次较高的是“好”、“声音”、“喜欢”等词，体现大部分学习者比较喜欢授课老师，认为授课老师讲课幽默，声音好听，专业知识过硬，具备个人魅力。

4. 研究讨论

4.1. 课程难度

在课程难度方面，对于学习者而言，过高的课程难度会降低其对课程的兴趣，而与学习者学习情况匹配的课程难度可以激发学习者的兴趣，帮助他们更好地掌握知识，提高学习效果。

4.2. 讲授方式

在讲授方式方面，教师讲授知识的速度也应与学习者的学习情况相匹配。否则，则容易导致学习者跟不上学习进度。内容生动的课堂讲授，具有一定的吸引力，从而引起学习者的学习兴趣(阎光才，2022)，对提升学习者课程学习参与度产生积极影响。

4.3. 课程作用

课程知识对能够促成学习者理性与情感全面和谐成长(张铭凯，2018)。因此，极具价值的课程不仅能够提升学习者知识和技能水平，还能够重构经验和生成意义，实现“知识传递”向“知识建构”的转化(金依平，2019)，有利于学习者进行有意义学习，从而提升学习参与度。

4.4. 布置的作业和练习

由于学习者的认知风格、思维方式和学习方式存在差异，教学应尊重差异，因材施教，作业也应关注差异。如果作业过于基础，一些学习者就会失去兴趣；如果作业过于高级，另一部分学习者就会感到无助(于会祥，2021)。因此，应考虑分层布置作业，这样，学习者就可以通过选择适合其个人能力的作业来发挥自主意识，同时又能激励他们完成难度更大的作业。

5. 总结

本研究将中国大学 MOOC 平台中《程序设计入门——C 语言》的课程反馈评价作为研究对象，采用内容分析法，从情感取向和评价类别两个角度展开分析该 14419 条学习者反馈评价数据，并进一步探究提升学习者在线课程学习参与度的主要因素及策略。

研究发现，学习者在该 MOOC 在线学习课堂反馈内容大多围绕教师的讲授方式和个人魅力，以及课程本身的内容与难度，正面情感比例较高。但课程设计仍存在待改进的地方，比如作业的布置可以考虑分层布置等。随着在线学习的推广，惠及学生的课程服务将更优质化和个性化，提高学习者参与度成为亟待解决的问题。本研究讨论得出提高学生在线学习参与度策略，力图为优化在线学习提供参考。

参考文献

于会祥.(2021).作业要分层布置.中小学管理,2021(09):59.

王志军、虞天芸和彭盼.(2024).基于知识结构化程度与学习者在线学习成熟度的在线学习分类研究.中国远程教育,1-13[2024-01-23]

张铭凯.(2018)论课程知识的价值维度及其道德价值逻辑.西北师大学报(社会科学版),2018,55(03):106-112.

金依平.(2019).围绕综合探究 开展意义教学.中学政治教学参考,2019(25):35-37.

赵晓阳和刘金兰.(2012).学习者参与度评价:一种学习者主体的教育质量评价方法.高教探索,2012,(06):21-26.

阎光才.(2022).大学讲授方式的功过、处境与挑战.江苏高教,2022(02):1-12.

ChatGPT 於計算機概論課程中的實際效果：測驗成效以及真實學習成效

Practical Impact ofChatGPTin Introduction to Computer Science Course: Exam Performance and Real Learning Outcomes

吳嘉恩、邱嘉豪
東吳大學
wu28hk@gmail.com
D04921004@ntu.edu.tw

【摘要】隨著大語言模型的進步，ChatGPT 被視為教育領域的潛力工具，對教育體系產生深遠影響，受到社會歡迎。對 ChatGPT 在教育的文獻綜合分析發現，現有實驗缺乏對學生在測驗和真實學習中應用 ChatGPT 的長期評估。本研究進行了 18 週的實驗，包括 2 次問卷和 6 次測驗，評估學生熟悉 ChatGPT 後的能力提升。結果顯示，ChatGPT 能有效提升學生測驗成績，隨著熟練度提高，學生提供更準確輸入，增強 ChatGPT 回應品質。值得注意的是，雖然學生對 ChatGPT 的信任提高，但自身能力並未大幅提升。本研究提供 ChatGPT 在教育中應用的實證證據，強調其在提升學生測驗成績上的潛在優勢，但須需要制定新的教育方式搭配 ChatGPT 提升學生自身能力。
【關鍵詞】*ChatGPT*；學習；學習成效；

Abstract:*ChatGPT, a powerful language model, holds promise in education, impacting the system positively. Existing literature lacks extended evaluations of its impact on students' performance and authentic learning. An 18-week experiment assessed students' progress through questionnaires and quizzes. Results indicatedChatGPT's effectiveness in enhancing test scores. As familiarity grew, students offered more precise input, improving response quality. Despite increased trust, students' own competence saw limited growth. This study underscoresChatGPT's potential to boost test scores but highlights the need for innovative educational approaches to align it with students' overall competence enhancement.*
Keywords*:ChatGPT, Learning, Learning Effectiveness*

1. 研究背景

ChatGPT 於 2022 年 11 月正式推出，標誌著自然語言處理（NLP）和對話生成領域的進步。其成功根植於對深度學習和大規模預訓練模型的持續研究，使 ChatGPT 能夠以前所未有的方式深度理解和生成自然語言。社會媒體平台 Twitter 上的輿論分析顯示，ChatGPT 引起了公眾的廣泛興趣和對 ChatGPT 抱持正面的態度（Ibrahim Adeshola et al., 2023）。ChatGPT 的迅速發展引起了社會的擔憂，包括技術應用的倫理考量、個人隱私和資訊安全等問題，需引起社會的高度關切。在教育領域，ChatGPT 被視為一具潛在教學輔助價值的工具，提供即時的個性化學生援助被認為能夠显著提升學習效率（David Baidoo-Anu et al., 2023）。(Radha Firaina et al., 2023)中的印尼教師受訪結果顯示，部分教師對 ChatGPT 持懷疑態度，需要反覆確認其回應。學界也關切 ChatGPT 是否會成為學術作弊的工具，因此制定明確政策有助於推動學術誠信（Ibrahim Adeshola et al., 2023）。ChatGPT 的出現為教育領域帶來新可能性，應全面考慮其優勢和特色，同時正視潛在風險和挑戰。儘管 ChatGPT 能提升學生尋找問題解答的效率，但過度依賴可能阻礙學生吸收知識，並未實質提升其自身能力。為充分發揮 ChatGPT 的優勢，需深入研究其在教育環境中的實際影響，並基於此進行技術改進和應用優化。

2. 研究目的

現有研究多侷限在 ChatGPT 對短期影響上和缺乏數據佐證反映出 ChatGPT 對教育的影響，缺乏對其在長時間實驗中的實際效果的深入了解，例如讓 ChatGPT 與學生一同進行相同測驗並分析準確率(Tarık TALAN et al., 2023)，但缺乏多長時間測驗觀察。本研究透過長時間且架構完整實驗觀察，深入研究 ChatGPT 對學生學習和測驗成效的實際影響。

本研究的主要目的：

1. 評估 ChatGPT 對學生自身能力的影響： 透過系統性的長期實驗觀察，深入分析 ChatGPT 在教育環境中的實際應用對學生自身能力的影響。
2. 探討 ChatGPT 對學生測驗成效的潛在影響： 利用嚴謹的數據分析方法，評估 ChatGPT 在不同教育場景下對學生測驗成效的潛在影響。
3. 透過問卷資料考察相關因素對 ChatGPT 影響的調節作用： 考慮學生的學術成就、自主學習動機、互動滿意度等多方面因素，以深入理解 ChatGPT 在教育中的實際效果。

3. 研究設計

3.1. 實驗設計

本研究為 18 週的實驗，以計算機概論學生為對象，分為兩個班級進行測試。A 班作為對照組（不使用 ChatGPT 輔助測驗），共計 59 名學生；B 班為實驗組（使用 ChatGPT 輔助測驗），共計 58 名學生。總計進行四次小考、一次期中考和一次期末考，小考和期中期末考內容均為計算機概論課堂所教授的知識，並在期中和期末各填寫一份關於 ChatGPT 的看法問卷。

實驗細項說明:

1. 初次評估（第一次小考）：對照組和實驗組均不使用 ChatGPT 進行輔助，以了解學生的初始能力。
2. 中期評估（第二、第三、第四次小考和期中考）：實驗組學生使用 ChatGPT 進行輔助，而對照組不使用，以觀察 ChatGPT 對學習表現的影響。
3. 期末評估（期末考）：期末考中，對照組和實驗組學生均不使用 ChatGPT 進行輔助，以進一步檢視 ChatGPT 對自身能力提升是否有顯著差異。
4. 問卷調查：問卷內容包括對於 ChatGPT 的各種看法、學生的基本信息，例如學區、高中類組、程式能力等。期中和期末，兩班學生填寫關於 ChatGPT 的看法問卷，調查受測者對 ChatGPT 的看法。同時觀察期中和期末受測者對於 ChatGPT 的看法是否有明顯差異及改變。

3.2. 受測者

參與實驗人員為東吳大學計算機概論課程學生，總計 117 人。參與者年齡範圍介於 18 到 20 歲之間，主要為大一新生。根據受測者背景調查，觀察到對照組程式能力有 84.4%的受測者對程式較為陌生，實驗組則有 71.2%的受測者對於程式較為陌生大多數學生尚未接觸過計算機概論，因此具備相對較少的相關知識背景。

4. 數據分析

本研究將對所收集的實驗問卷和測驗數據進行全面的數據分析，包括基礎能力差異分析(4.1)、測驗成績差異分析（4.2)、最終學習成果差異分析（4.3）以及 ChatGPT 看法問卷分析(4.4)。這些分析旨在深入瞭解研究對象在不同層面的表現差異，提供全面的研究結果，強化研究的科學性和可靠性。

4.1. 基礎能力差異分析:

針對實驗組與對照組的初始能力進行深入比較。利用平均值、中位數和標準差等基本統計指標，量化兩組學生的初始實力表現，同時檢查是否符合常態分佈和進行差異假設檢定。

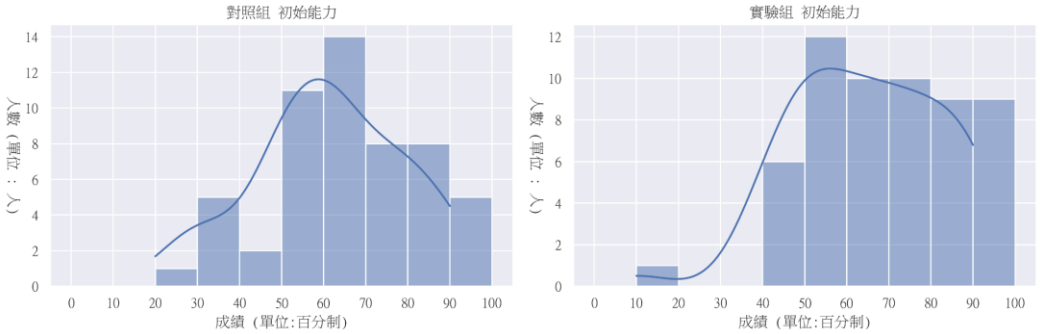


圖 1 對照組和實驗組初始能力成績之分布

	初始能力
Cohen's d	0.206
Hedges' g	0.204
Mean Difference	3.635

表 1 針對初始能力進行實驗組和對照組的比較去看兩者差異

	W	P-value
對照組	0.94826	0.0209
實驗組	0.93087	0.00292

表 2 對照組和實驗組初始能力數值的 Shapiro-Wilk 檢定

進行 Shapiro-Wilk 檢定評估對照組和實驗組初始能力數值分布的正態性。虛無假設（H0）：總體呈常態分佈，對立假設（H1）：總體不呈常態分佈，臨界值為 0.05。檢定表 2 結果顯示兩組初始能力的 P-value 均小於臨界值(0.05)，指示對照組和實驗組的初始能力均未符合常態分佈。後續分析將使用非常態分佈的小樣本檢定方法。由於初始能力數值不符合常態分佈且樣本數量不同，選擇 Mann-Whitney U 檢定，一種適用於非常態分佈和小樣本的統計方法，評估兩組能力是否存在差異。虛無假設（H0）：兩組初始能力的分佈相同，對立假設（H1）：兩組初始能力的分佈不相同，臨界值為 0.05。Mann-Whitney U 檢定結果，U 統計量為 1458、P-value 為 0.5012。由於 P-value 大於臨界值，無法拒絕虛無假設，統計上未發現對照組和實驗組初始能力存在顯著差異。兩組受測者在開始學習計算機概論前的初始能力相似，有助於更有效觀察後續 ChatGPT 輔助測驗的成果。

4.2. 測驗成績差異分析:

深入探討 ChatGPT 對學生期中考和後續小考成績的影響。透過對兩組學生在這些測驗中平均分、標準差等統計指標的比較，致力於精細了解 ChatGPT 在學業表現方面的效果。這些統計分析的目的是提供客觀且可量化的數據，以評估 ChatGPT 在測驗學習環境中的實際影響。透過對比分析，期望揭示潛在的學習趨勢和差異，為 ChatGPT 在教育領域的應用提供更具體的洞見。對四個測驗（期中考、第二次小考、第三次小考、第四次小考）進行差異分析。在這四次測驗中，對照組未使用 ChatGPT 進行測驗輔助，而實驗組則在所有測驗中均有使用 ChatGPT 進

行測驗輔助。除了檢視 ChatGPT 對實驗組和對照組的影響外，透過四次測驗協助學生熟悉使用 ChatGPT，進一步探討其在學習過程中的應用效果。

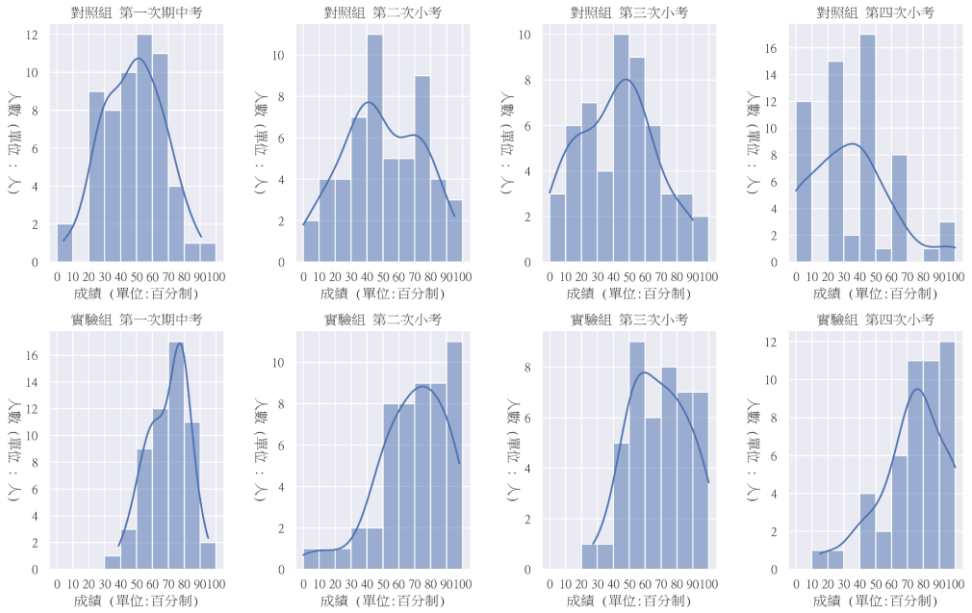


圖 2 對照組和實驗組各次測驗的成績分布

	Count	Mean	Std	Min	25%	50%	75%	max	Kurtosis	Skewness
期中考	58.000	48.047	18.800	4.000	32.000	49.650	64.000	90.700	0.467	0.031
第二次小考	54.000	49.093	24.641	0.000	33.500	45.000	70.000	95.000	0.825	0.041
第三次小考	53.000	41.604	24.091	0.000	20.000	41.000	60.000	90.000	0.710	0.108
第四次小考	59.000	33.153	25.350	0.000	20.000	40.000	40.000	100.000	0.261	0.651

表 3 對照組（沒有使用 GPT 輔組測驗）各項測驗數值敘述統計

	Count	Mean	Std	Min	25%	50%	75%	max	Kurtosis	Skewness
期中考	55.000	69.485	12.726	38.700	60.500	73.300	78.000	95.000	0.581	0.371
第二次小考	52.000	68.192	22.241	0.000	54.250	70.000	85.000	98.000	0.880	0.930
第三次小考	44.000	68.505	18.485	27.300	54.750	67.100	81.900	99.900	0.802	0.005
第四次小考	48.000	73.542	21.086	15.000	63.750	75.000	86.250	100.000	0.377	0.797

表 4 實驗組（有使用 GPT 輔組測驗）各項測驗數值敘述統計

	期中考	第二次小考	第三次小考	第四次小考
Cohen's d	1.114	0.759	1.062	1.312
Hedges' g	1.32	0.807	1.228	1.704
Mean Difference	21.439	19.1	26.901	40.389

表 5 針對每個測驗進行實驗組和對照組的比較去看兩者之差異

為客觀評估實驗組和對照組差異，本研究使用統計指標，包括 Cohen's d、Hedges' g 和 Mean Difference。透過觀察表 3、4、5，在相同初始能力下，實驗組在各測驗中呈現明顯差異。四次測驗平均差異之平均為 26.957，顯示 ChatGPT 的輔助對實驗組效果顯著。根據(Yilma et al., 2023)的研究，ChatGPT 提升演算法思維、計算機能力和編程能力有效，實驗組經一系列測驗練習後，更有效掌握 ChatGPT 使用方式。第四次測驗中，實驗組與對照組平均差距為 40.389 分。

4.3. 最終學習成果差異分析:

進行期末測驗成績深入對比分析。此次期末考試評估實驗組在經過 18 周計算機概論課程及四次測驗練習後，使用 ChatGPT 的效果。檢驗 ChatGPT 是否能有效提升受測者能力。期末考試中，對照組和實驗組均不能使用 ChatGPT 進行測驗輔助。同時進行學習長期影響的假設檢定，利用統計方法評估 ChatGPT 對學習能力的長期影響是否達統計顯著水準。

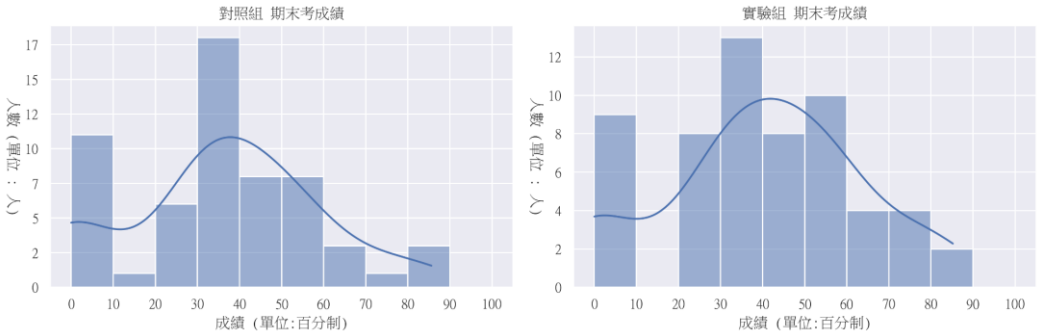


圖 3 對照組和實驗組的期末測驗成績分布

	Count	Mean	Std	Min	25%	50%	75%	max	Kurtosis	Skewness
對照組	59.000	38.902	20.850	0.000	29.900	38.100	51.300	85.600	0.021	0.079
實驗組	58.000	40.066	23.413	0.000	27.450	39.600	55.625	85.200	0.526	0.174

表 6 對照組和實驗組期末考成績敘述統計

期中考	期末考

Cohen's d	1.114	0.053
Hedges' g	1.32	0.052
Mean Difference	21.439	1.164

表 7 針對期末考和期中考進行實驗組和對照組的比較去看兩者差異

透過先前實驗數據深入分析，發現相似初始能力的對照組與實驗組在期中測驗中，使用 ChatGPT 的實驗組表現顯著優於未使用的對照組。統計指標包括 Cohen's d、Hedges' g、Mean Difference，在表 7 中呈現極大的數值，其中 Mean Difference 達到 21.439 分，突顯了實驗組在此測試情境下的卓越表現。至於期末測驗，由於兩組均無法使用 ChatGPT 進行輔助，表 8 顯示期中考時兩組平均差異為 21.439 分，而期末考時兩組平均差異為 1.164 分，兩組之間的成績平均並沒有明顯差異。觀察表 4、表 6 的實驗組敘述統計數據，發現使用 ChatGPT 時 50% 中位數皆約為 75 分，而無 ChatGPT 輔助時中位數約為 39.6 分，顯示 ChatGPT 在協助學生取得正確答案方面具有顯著效果。在長期使用 ChatGPT 進行學習和測驗的情境下，對提升受測者本質能力的影響則呈現不明顯。

4.4. 問卷分析:

本研究旨在透過對受測者的測驗分析，結合他們對 ChatGPT 的看法，客觀綜合分析，並研究設測者在測驗中是否受其他因素影響。問卷題目參考自(Harpreet Singh et al., 2023)。受測者分為對照組和實驗組，填寫有關 ChatGPT 觀點的問卷。研究結合後續成績數據，探討其他因素是否影響了受測者在使用 ChatGPT 輔助測驗時的考慮因素。

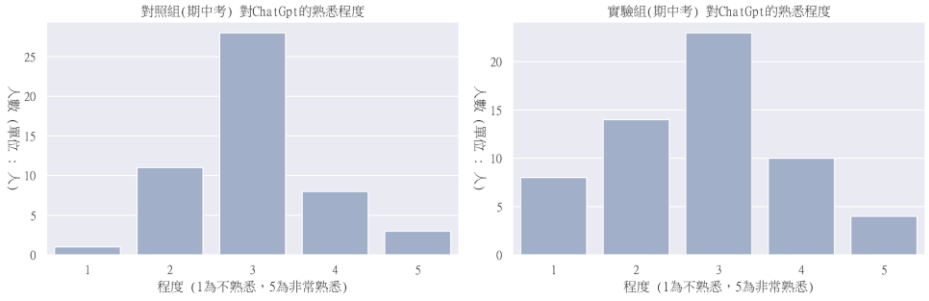


圖 4 期中問卷呈現對照組和實驗組分別對於 ChatGPT 的熟悉程度

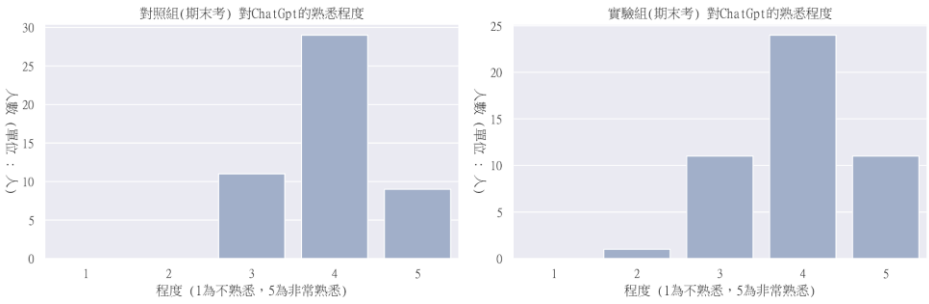


圖 5 期末問卷呈現對照組和實驗組分別對於 ChatGPT 的熟悉程度

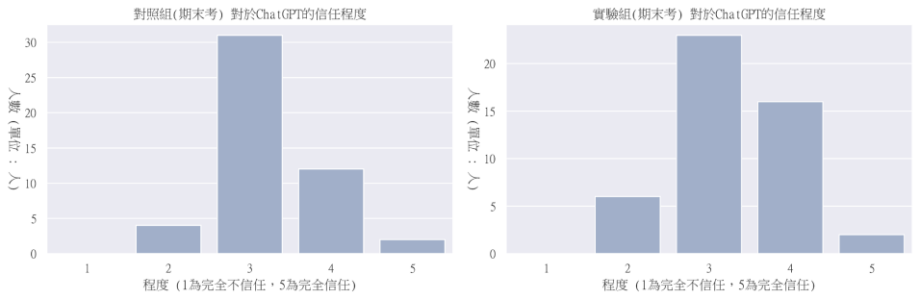


圖 6 期末問卷呈現對照組和實驗組分別對於 ChatGPT 的信任程度

透過圖 4、圖 5、表 5 觀察，經四次測驗訓練後，實驗組對 ChatGPT 的熟悉度和測驗成績均正向提升。根據表 5 的第四次小考結果，實驗組與對照組的 Mean Difference 高達 40.389，是前三次測驗平均差異的兩倍。考量 ChatGPT 目前已成為熱門技能，因此即使受測者未接受實驗訓練，對 ChatGPT 也具有初步的了解。從圖 4 和圖 5 的觀察結果可以得知，在未接受 ChatGPT 訓練的情況下，受測者仍然能夠掌握部分 ChatGPT 相關技能，這顯示了其對 ChatGPT 的潛在理解能力。經過四次訓練後，實驗組受測者更熟練地應用 ChatGPT，進一步提升其在輔助測驗中的有效性和準確性。通過一系列的測驗，實驗組受測者對 ChatGPT 的熟練程度逐漸提升，在第四次考試中充分掌握了使用方式，對 ChatGPT 所提供的回應表現出更明確的信任度，不再保持中立。從圖 6 中觀察到，實驗組經過多次 ChatGPT 練習後，對 ChatGPT 的信任程度相對對照組高，同時也增加實驗組對 ChatGPT 的依賴，導致受測者難以真正吸收知識並提高自身能力。

5. 結論

基於研究數據分析，觀察到實驗組隨 ChatGPT 操作次數增加，逐漸提升 ChatGPT 的熟練度。表 5 的測驗成績差異分析顯示，實驗組與對照組的 Mean Difference 呈現擴大趨勢，期中考兩組平均差為 21.439 分、第二次小考平均差為 19.1 分、第三次小考平均差為 26.901 分，特別在第四次測驗時達到 40.389 分。這顯示實驗組熟悉 ChatGPT 操作後，對輸入合適的關鍵詞更加熟悉，進而取得更高的測驗分數，驗證其提供準確回應時精確輸入提示的關鍵性（Pengfei Liu et al., 2023），同時也顯示隨 ChatGPT 熟練度提升，受測者對其信任度也增加。ChatGPT 的應用提升了學生學習效率，獲得正確且有參考價值的回答。觀察到隨 ChatGPT 熟練度提升，學生對其信任度也增加，可能產生過度依賴 ChatGPT 回答的現象。學生在使用 AI 生成模型時需要保持對生成文字的解讀能力。儘管 ChatGPT 在提高學生學習效率上取得成效，快速幫助學生找到正確且有參考價值的答案，最終學習成果比對分析顯示，實驗組在期中考試中由於 ChatGPT 輔助取得成績優勢，然而在期末考試中(實驗組和對照組均未使用 ChatGPT 進行輔助)，實驗組平均得分為 40.066 分，而對照組則為 38.902 分，顯示兩組之間僅存在 1.164 分的平均差異。這結果表明，儘管實驗組接受了 ChatGPT 的一系列輔助考試，但他們未能有效地吸收相應的概念。因此實驗組與對照組之間在期末考試成績上未呈現出顯著差異。這表明 ChatGPT 雖能在短時間內提供準確答案，對學生自主能力提升仍有限。在教育中應用 ChatGPT 仍需面對挑戰，特別是對學生自主學習能力的影響。教育機構應制定詳細的使用規則和指導方針，引導學生適當地利用 ChatGPT。透過實驗驗證，本研究展示了 ChatGPT 在專業領域的準確回應能力，並探討了其在教育場景中的應用潛力。未來能夠將 ChatGPT 納入課堂環境，擔任與

學生互動討論的角色，在(Fauzi、Tuhuteru、Sampe、Ausat&Hatta，2023)的研究中指出，ChatGPT 在協助學生實現學術目標方面具有顯著效果。他們強調了 ChatGPT 是一種教學輔助工具，能夠有效地輔助學生更有效地實現其學術目標。藉此將其定位為教學輔助工具而非僅僅是解題工具。這樣的應用方式不僅提高了學生學習效率，同時減輕了教師的指導負擔，鼓勵學生積極探索新知識。

參考文獻

Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. Interactive Learning Environments. Advance online publication. https://doi.org/10.1080/10494820.2023.2253858

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits ofChatGPTin Promoting Teaching and Learning. Journal of AI, 7(1), 52-62. https://doi.org/10.61969/jai.1337500

TALAN, T., & KALINKARA, Y. (2023). The Role of Artificial Intelligence in Higher Education:ChatGPTAssessment for Anatomy Course. Uluslararası Yönetim Bilişim Sistemleri Ve Bilgisayar Bilimleri Dergisi, 7(1), 33-40. https://doi.org/10.33461/uybisbbd.1244777

Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring Computer Science Students' Perception ofChatGPTin Higher Education: A Descriptive and Correlation Study. Educational Sciences, 13(9), 924. https://doi.org/10.3390/educsci13090924

Firaina, R., & Sulisworo, D. (2023). Exploring the Usage ofChatGPTin Higher Education: Frequency and Impact on Productivity. Buletin Edukasi Indonesia, 2(01), 39–46. https://doi.org/10.56741/bei.v2i01.310

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. Computers and Education: Artificial Intelligence, 4, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. arXiv:2107.13586 [cs.CL]. https://doi.org/10.48550/arXiv.2107.13586

Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A., & Hatta, H. (2023). Analysing the Role of ChatGPT in Improving Student Productivity in Higher Education. *Journal on Education*, *5*(4), 14886-14891. https://doi.org/10.31004/joe.v5i4.2563

智能代理工具的关键技术与教育应用

Key technologies and educational applications of intelligent agent tools

刘倩[1]，杨现民[2*]，李新[3]
[123]江苏师范大学智慧教育学院
[*]yangxianmin8888@163.com

【摘要】随着人工智能技术的迅猛发展，智能代理的理解能力、对话能力、逻辑推理能力等均取得了巨大突破，对支持教师教学、提升学习者学习体验都有着重要意义和价值。基于此，本文收集了国内外智能代理工具的相关文献，总结出了智能代理在教育中的应用场景，并按照“输入-过程-输出”框架凝练出了智能代理工具的关键技术。

【关键词】智能代理；关键技术；教育应用；人机交互

Abstract: *With AI advancements, intelligent agents now excel in understanding, dialogue, and reasoning, aiding teaching and enhancing learning. The article gathers global literature on agent tools, outlines their educational applications, and identifies key implementation technologies.*

Keywords: Intelligent agent, Key technologies, Educational application, Human-computer interaction

基金项目：国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”(项目编号：62077030)；2023 年度江苏省教育科学规划课题“协作学习情境中学习投入的动态演变机制与干预策略研究”（项目编号：C/2023/01/99）

1.引言

智能代理因其类人交互、及时反馈等特性，能够在语言学习(Jeon,2022)、心理健康辅导(Young et al,2020)等教学场景中为学习者提供个性化指导。然而目前仍缺少对智能代理关键技术与教育应用的系统梳理。基于此，本研究汇总了教育中典型的智能代理工具应用案例，并进一步梳理了智能代理工具的教育应用场景及技术实现原理，以期为智能代理工具在教育领域未来的研究提供参考与借鉴。

2.文献综述

教育场景中智能代理通常扮演导师、同伴及工具等教学角色，并利用文本、语音、图形及其他交互媒介来实现交互，较少以具体的形象出现。其中典型的工具如表 1 所示。

表 1 智能代理的教育应用

应用场景	工具	交互媒介	实体化	响应方式	角色定位	主要功能
语言学习	Ellie（Kim et al，2022）	语音	女性教师	基于脚本	伙伴	采用“常规聊天”、“任务聊天”、“技能聊天”辅助学习者第二语言的习得
	PengTalk（Jeon，2022）	语音;文本;图像	企鹅	基于学习	伙伴	通过游戏进行单词、句子、口语训练
心理健康	Anti−bullying bot（Young et al, 2020）	文本	/	基于脚本	工具	扮演欺凌者或受害者或资询老师与学生交流

干预	Tess（Fulmer et al, 2018）	文本 表情符号	/	基于脚本	导师	通过与大学生互动，提供各种心理健康干预措施缓解大学生的焦虑、抑郁症状。
学习辅导	learning buddy bot（Hew et al, 2023）	文本 表情符号	/	基于脚本	伙伴	以问题—选项的交互形式协助学生制定并达成目标
	the tutor bot（Ait et al, 2023）	文本 表情符号	/	基于脚本	导师	帮助教师教授编程语言内容

3.智能代理工具的关键技术

目前智能代理主要基于 Dialogflow、Flask、TensorFlow 框架搭建，一般包含用户界面(GUI)、内核(自然语言处理、自然语言理解、机器学习)及数据库等组件（图 1）。首先，在用户界面，智能代理引擎能接收用户的需求并将其传递到内核；然后，通过自然语言处理技术提取关键信息并进一步识别用户意图，再从多个来源（预先编写的脚本库、web 服务）中检索对应的响应信息，并根据脚本库中的响应模式添加附加信息（时间、学习者名称等）；最后，自然语言生成组件再将响应信息转化为自然语言以信息框的形式呈现给学习者。另外，收集的对话内容及学习者反馈都将存储到数据库中，以供持续学习和对话能力的改进。

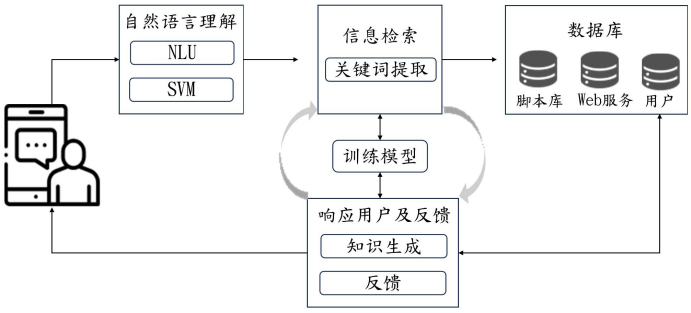


图 1 智能代理典型架构

4.总结与展望

本研究比较分析了六项智能代理的教育应用案例，并根据“输入—过程—输出”框架总结了实现智能代理的关键技术。生成式人工智能作为一项新兴的技术，能够根据用户提示自动生成内容。后续本研究将聚焦于生成式人工智能赋能智能代理，以改善智能代理问题诊断不精准、学习反馈不及时、个性化评价不足等问题，使得智能代理更好地服务于教育教学。

参考文献

Ait Baha, T., El Hajji, M., Es-Saady, Y., & Fadili, H. (2023). The impact of educational chatbot on student learning experience. Education and Information Technologies, 1-24.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR mental health, 5(4), e9782.

Hew, K. F., Huang, W., & Du, J. et al. (2023). Using chatbots to support student goal setting and social presence in fully online activities: learner engagement and perceptions. Journal of Computing in Higher Education, 35(1), 40-68.

Jeon, J. (2022). Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective. Education and Information Technologies, 27(4), 5767-5787.

Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. Language Learning & Technology, 26(1), 1–18.

Young Oh, E., Song, D., & Hong, H. (2020). Interactive computing technology in anti-bullying education: The effects of conversation-bot’s role on K-12 students’ attitude change toward bullying problems. Journal of Educational Computing Research, 58(1), 200-219.

科学虚拟实验中迷思概念的学习行为挖掘——以小学“简单电路”单元为例

Learning Behavior Mining of Misconception in Scientific Virtual Experiment: A Case Study of "Simple Circuit" Learning in Primary School

李凤鸣，吴忭*
华东师范大学教育信息技术学系
bwu@deit.ecnu.edu.cn

【摘要】 有效模拟科学现象的虚拟实验，有助于破除迷思概念，实现概念转变，但会出现反馈慢、效率低等问题，智能改进虚拟实验有助于改善教学。为了解虚拟实验过程中学生操作行为并促进虚拟实验反馈改进，本研究以小学学生开展虚拟实验进行简单电路概念学习过程为例，应用过程挖掘方法开展探索性序列分析，以此归纳电路迷思概念和操作序列模式，为虚拟实验平台反馈设计提供建议。研究最终归纳出了四项电路迷思：封闭性迷思、通路迷思、短路迷思和灯泡连接迷思，并提出了应用于虚拟实验反馈的简单规则。
【关键词】 虚拟实验；电路迷思概念；操作行为；过程挖掘；探索性分析

Abstract: *Virtual experiments effectively simulating scientific phenomena help dispel misconceptions and achieve conceptual change. However, issues such as slow feedback and low efficiency arise. Intelligent improvements in virtual experiments can enhance teaching. To understand student operational behavior during virtual experiments and promote feedback enhancement, this study takes elementary school students' virtual experiments on simple circuit concepts as an example. Exploratory sequence analysis is conducted using Process Mining to infer circuit misconception concepts and operational sequence patterns, providing suggestions for feedback design on virtual experiment platforms. The study ultimately identifies four circuit misconceptions: closed circuit misconception, pathway misconception, short-circuit misconception, and bulb connection misconception, and proposes simple rules for feedback in virtual experiments.*
Keywords: Virtual Experiment, Circuit Misconception, Operational Behavior, Process Mining, Exploratory Analysis

1.引言

为促进科学概念学习，虚拟实验环境兴起。该类环境呈现了现实环境无法实现的特殊情境，通过丰富表征促进认知冲突，以破除迷思概念并实现概念转变。尽管虚拟实验促进科学概念转变已得到了充分讨论(Falloon，2019)，但学生的实验过程表现未得到足够关注。除此，虽然目前研究对于电路迷思概念已进行了较为充分的讨论(林静雯&邱美虹，2009)，但对于概念转变过程缺乏了解，故本研究尝试应用过程挖掘方法分析虚拟实验过程数据(黄琰等，2020)，拟解决以下问题：导致实验失误的操作序列反映出了怎样的电路迷思概念？应给出怎样的反馈？

2.研究设计

研究对象是上海某小学四年级的 50 名学生，两人小组应用 PhET 桌面虚拟实验（https://phet.colorado.edu/sims/html/circuit-construction-kit-dc/latest/circuit-construction-kit-dc_all.html）开展简单电路虚拟实验，实验过程共产生了 2369 行日志数据。本研究主要基于开源 R 包 bupaR（https://bupar.net/）开展过程挖掘，构建过程模型。模型分析中重点关注未完成实验任务或出现明显失误的学习小组，通过分析异常操作序列揭示电路迷思概念。

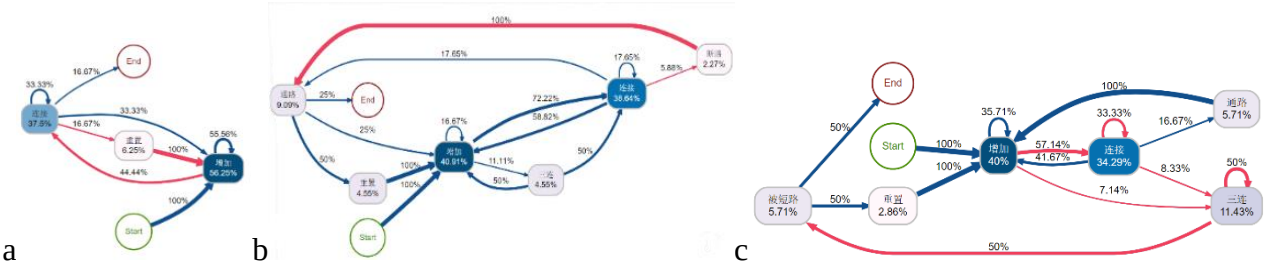
3.结果分析

存有“封闭性迷思”的学生无法连接电池、灯泡等元件构建闭合回路。比如第 16 组学习者（见图 1-a），主要操作序列为“增加➞连接➞重置➞增加”，只限于增加和连接元件，而缺少通路操作。虚拟实验平台此时需给予闭合环路提示，补充封闭电路概念。

存有“通路迷思”的学生对于通路电流和电路结构认识模糊，无法使用开关控制电路。分析第 10 组的过程模型（见图 1-b）可发现异常序列“连接➞断路➞通路”，若检测到该类异常操作，应提示学生思考开关作用，加强通路概念。

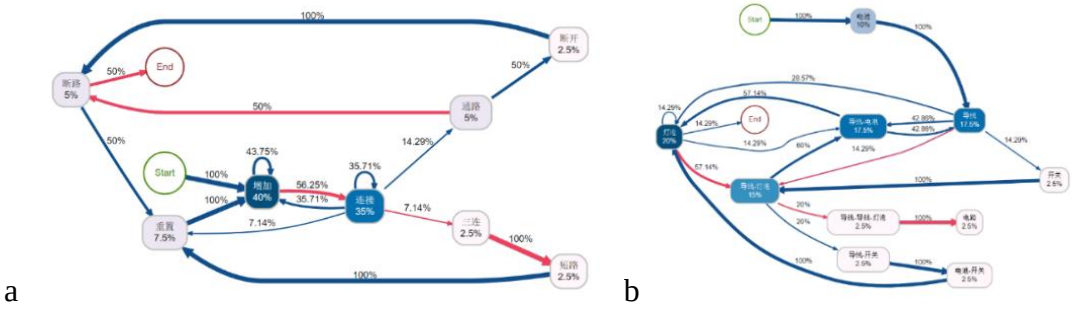
存有“短路迷思”的学生不理解元件并联方式，导致导线与元件并联，引发短路。比如第 3 组（见图 1-c），异常操作序列为：“增加➞连接➞连接➞三连➞三连➞被短路”，应提示学生关注被短路元件，明确短路概念并修改连接方式。

存有“灯泡连接迷思”的学生仅将灯泡一端接入电路，导致短路，比如第 22 组（见图 2）。该组学习者的异常序列为“增加➞连接➞三连➞短路”，又结合元件操作过程模型（见图 2-b），序列可完善为“增加（导线，灯泡）➞连接：导线-灯泡➞三连：导线-导线-灯泡➞短路”。此时应给予灯泡连接提示“将灯泡两端点都接入同一电路”，以破除连接迷思。



a b c

图 1 第 16 组（图 1-a）、第 10 组（图 1-b）和第 3 组（图 1-c）的过程模型



a b

图 2 第 22 组的过程模型（图 2-a）和元件操作过程模型（图 2-b）

4.总结展望

本研究应用过程挖掘技术分析虚拟实验日志数据，发现了学生操作序列反映出的四项电路迷思：封闭性迷思、通路迷思、短路迷思和灯泡连接迷思，并提出了简单反馈规则，支持虚拟实验改进设计，进一步提升实验学习效率。

参考文献

Falloon, G. (2019). Using simulations to teach young students science concepts: An Experiential Learning theoretical analysis. *Computers & Education*, 135, 138-159.

林靜雯, & 邱美虹. (2009). 探究以學生心智模式為設計基礎之教-學序列對學生電學學習之影響.*科學教育學刊*,17(6), 481-507.

黄琰,赵呈领,赵刚,刘军,疏凤芳 李红霞.(2020).教育过程挖掘智能技术：研究框架、现状与趋势. *电化教育研究*(08),49-57.doi:10.13811/j.cnki.eer.2020.08.007.

数学焦虑、自我效能感与数学素养的关系研究：考虑性别差异下自我效能感的中介作用及影响机制

Research on the Relationship between Math Anxiety, Self-Efficacy, and Math Literacy: The Mediating Role and Influencing Mechanism of Self-Efficacy Considering Gender Differences

张瑶[1*]，虞增艺[2]，金泊西[3]
[123]浙江工业大学 教育科学与技术学院
*865373013@qq.com

【摘要】本研究基于 PISA2022 数据集，探讨了数学焦虑、数学素养和自我效能感的关系。独立样本 t 检验表明性别在上述关系中存在显著差异，其中，男生的数学焦虑显著低于女性。基于结构方程模型分析，自我效能感在数学焦虑和数学素养中起部分中介作用，尤其在不同性别间表现出显著差异。这一发现强调了在数学教育中考虑性别差异的重要性，并提示了通过提高自我效能感来缓解数学焦虑和提升数学素养的潜在途径。研究代码与数据集已公开在 https://github.com/yuzengyi/Gccce_yao。

【关键词】PISA2022；数学焦虑；自我效能感；数学素养

Abstract: *This study, based on the PISA 2022 dataset, explores the relationship between math anxiety, math literacy, and self-efficacy. Independent t-tests reveal significant gender differences in these relationships, with males showing significantly lower math anxiety. Structural equation modeling analysis indicates that self-efficacy partially mediates the relationship between math anxiety and math literacy, particularly across different genders. This finding highlights the importance of considering gender differences in math education and suggests potential pathways to alleviate math anxiety and enhance math literacy through increasing self-efficacy. The research code and dataset are publicly available at https://github.com/yuzengyi/Gccce_yao.*

Keywords: PISA 2022, math anxiety, self-efficacy, mathematical literacy

1. 前言

数学素养是指个体在现实情境中进行数学推理及制定、运用和解释数学解决方案的能力。数学焦虑是解决数学问题时所感受到的紧张和焦虑情绪，会影响数学认知能力的发展（陈英和,耿柳娜., 2002）。班杜拉将自我效能感定义为个体对自己在特定领域完成任务或应对挑战的能力的信念，已有研究指出自我效能感较高的学生通常具有更高的数学素养。而国内对于中学生数学焦虑、自我效能感和数学素养之间关系的综合性研究较少。因此，本研究基于 PISA 2022 数据集，使用结构方程模型，从数学焦虑和自我效能感入手，分析它们与数学素养及其分维度的相关性，并进一步探究自我效能感的中介作用。

2. 数据来源与研究方法

本研究数据源于 PISA2022 数据集。该框架以数学学科教育、计算思维教育为落脚点，提供了在学生数学学习中发展数字素养教育及其评估的新范式（崔志翔,徐斌艳., 2022）。研究选择香港、澳门、台北三地区的数据，涉及 21 个变量字段和 11339 条数据，并包括数学焦虑和自我效能感评估量表，以及数学素养分维度的评估数据。具体过程见摘要所附代码仓库。

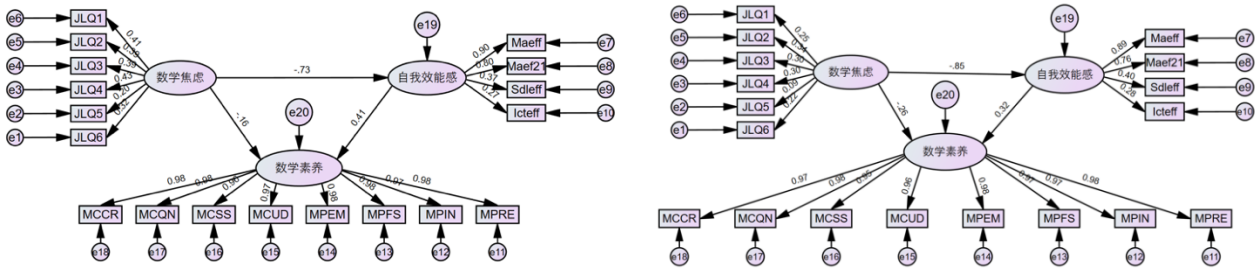
3. 研究结果与讨论

3.1.性别变量差异显著性检验

根据独立样本 t 检验结果，性别对数学焦虑、数学素养和自我效能感均有显著性差异。男生的数学焦虑显著低于女性，数学得分略高于女性，因此，将基于**性别视域**进行分类分析。

3.2.性别视域下自我效能感在数学焦虑与数学素养关系的中介效应分析

本研究通过计算 AVE 和 CR 值，确认了数学焦虑、自我效能感和数学素养的收敛效度和组合信度达到基本标准，且利用 PISA 设计的量表进行测量，保证了研究的权威性和可靠性，克朗巴赫 α 系数为 0.621，表明结果具有可信度。根据 Baron 和 Kenny 的方法以及 HAYES 的建议（HAYES A F, 2009），研究确定了自我效能感在数学焦虑和数学素养之间的中介作用。模型拟合度指标（方杰,张敏强,邱皓政.，2012）中的 GFI、AGFI、NFI、IFI、TLI、CFI 超过 0.9，RMSEA 低于 0.08，PGFI 和 PNFI 超过 0.5，表明模型总体拟合良好，尽管 CMIN/DF 值因大样本而偏高，但不影响模型的整体适配性。最终得到的结构方程模型图，如图 1 所示。



(a)男性三个变量的结构方程模型图　　(b)女性三个变量的结构方程模型图

图 1 不同性别数学焦虑、自我效能感与数学素养的结构方程模型图

4. 研究结论与建议

本研究通过构建中介作用模型，基于性别视域探讨数学焦虑、自我效能感与数学素养的关系。研究结果显示，数学焦虑与自我效能感和数学素养成负相关，而自我效能感与数学素养正相关。自我效能感部分中介了数学焦虑与数学素养之间的关系，即可以在一定程度上缓解数学焦虑对数学素养的负面影响。研究还创新性地分性别探讨了这些关系，发现男性在某些数学素养维度上得分更高，可能与较低的数学焦虑有关。相比之下，女性的数学焦虑更高，对其数学素养和自我效能感影响更大。教师应科学干预和培养学生，尤其是女性学生的自我效能感，构建引导探究式课堂教学模式，减轻学生数学焦虑，为学生提供适度的情感支持。

参考文献（部分）

陈英和和耿柳娜（2002）。数学焦虑研究的认知取向。**心理科学，25(6)**，653-655。

崔志翔和徐斌艳（2022）。数智时代国际基础学科计算思维教育发展的策略，方向与启示——《pisa 2022 数学框架》之思考。**远程教育杂志，40(6)**，13-21。

Hayes, & Andrew, F. (2009). Beyond baron and kenny: statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408-420.

方杰、张敏强和邱皓政（2012）。中介效应的检验方法和效果量测量：回顾与展望。**心理发展与教育，28(1)**，7。

以 STEAM 教育融入教育元宇宙之平台開發與研究

Development and Research of the Metaverse Platform for Inclusive Education through STEAM Education

王政弘 [1*]，吳聲毅 [2]
[1] 國立高雄大學工藝與創意設計學系
[2] 國立屏東大學 STEM 教育國際碩士學位學程
[*]wang101@go.nuk.edu.tw

【摘要】 後疫情時代伴隨著 5G 行動普及的大環境來臨，虛擬實境、元宇宙等新興科技應用於各行業十分廣泛，虛擬實境之沉浸式與高層次互動特性，亦適合應用於教育現場為教學帶來改變。本研究綜觀現今背景所造就之新興科技崛起現象，開發教育元宇宙平台，並進行系統易用性評量與質性回饋研究，提出以 STEAM 教育融入教育元宇宙之教學應用方式與可能性。

【關鍵詞】 STEAM；教育元宇宙；教學應用；虛擬實境；系統易用性

Abstract: *In the post-pandemic era, accompanied by the widespread adoption of 5G technology, a significant environment is emerging. Emerging technologies such as virtual reality and the metaverse are extensively applied across various industries. The immersive and highly interactive nature of virtual reality makes it well-suited for educational settings, bringing about transformative changes in teaching. This study takes a comprehensive look at the rise of emerging technologies created by the current background, develops an educational metaverse platform, conducts system usability evaluation and qualitative feedback research, and proposes teaching application methods and possibilities for integrating STEAM education into the educational metaverse.*

Keywords: STEAM, Educational metaverse, Teaching application, Virtual reality, System usability

1. 研究背景與動機

現今真實世界所會面臨到的生活大小問題，並非擁有單一學科能力即可解決；而所有不同領域的科目，若僅依靠傳統課堂教授方式學習，學習成效亦是顯得不足。此時強調動手做與解決問題能力，且結合科學、技術、工程、藝術與數學領域的 STEAM 教育，很大程度上可滿足當今真實世界的教育需求。透過科學與藝術融合學習，學生能夠反思與想像未來社會之型態，並且改變學習模式、將思想及理論轉化為實踐（Burnard & Colucci-Gray, 2019）；有學者更提出使用虛擬教育環境的必要條件，應該是將以人為本的方法引入，尤其是在 STEAM 的領域內（Nataliia, 2021）。然而，近年來由於 COVID-19 疫情席捲全球之影響，許多需要動手操作與互動的學習課程受到場域與時間限制，因而受到衝擊，即便疫情已趨緩，但身處後疫情時代的我們，該如何因應學習習慣與教學方式結構上的改變，並且克服與延續 STEAM 教學上之困境與優勢，是我們要面對與研究的課題。

Meta 公司的創始人 Mark Elliot Zuckerberg 曾在 2021 年其公司更名時之公開信內表示：互聯網的下一代是一個身歷其境的互聯網，也就是元宇宙(Metaverse)。在元宇宙中，我們幾乎能夠進行任何我們能夠想像到的任何事務。在元宇宙的虛擬世界中，民眾能夠在裡頭進行

集體學習、工作與社交互動，特別在 COVID-19 的疫情影響下，人們使用遠距教學與線上居家辦公更是習以為常。虛擬實境(Virtual Reality, VR）透過頭戴式裝置顯示器生成虛擬情境，為使用者創造身歷其境的互動體驗；它亦可提供虛擬沉浸式實驗場景學習環境，提供學習者更多的直觀學習經驗（Reeves, Crippen, & McCray, 2021）。雖然現階段娛樂與遊戲產業仍然佔據 VR 市場內的大宗，但 VR 技術已逐漸被視為創新高等教育在線教學的大好機會（Wohlgenannt, Simons, & Stieglitz, 2020）因此在此發展趨勢中，VR 技術的成長結合元宇宙平台的應用，在 STEAM 教育面向中，帶來前所未有的想像。

目前市面上現有的 VR 學習資源上，因技術與教學實作都尚在發展推廣階段，應用於 STEAM 課程學習未臻完善，在學生自學與教師導學時，皆碰到一些使用上不足的地方，例如頭戴式硬體設備與 VR 教材的設計不相容，導致教材無法被使用。還有諸如 VR 應用上教師與學生無法即時互動、瞭解學習需求等問題，這些都是促使本研究想建置教育元宇宙共學平台，並結合 STEAM 應用的動機。

2. 文獻探討

2.1. 線上學習與 STEAM 教育

教育是近年受到 Covid-19 疫情影響甚鉅的領域，伴隨著社交距離等防疫措施的實施，原本傳統的面對面教學方式逐漸改變為遠距線上授課模式。然而雖然教師們可透過線上教課模式，預先錄製教學影片，或於線上直播平台進行課程的即時教授，但是受制於設備與網路訊號的有限性，師生教學之間的互動很大程度地受到了阻礙（Sharma, Prasad,& Rambabu,2021）。況且，線上教學需使用適應課堂動態需求的新技術與方法，這對許多教師來說是一項挑戰；對學生們亦將面臨失去動力與不想繼續學習的風險，因為他們感到孤立、缺乏學習習慣或減少與同儕的聯繫（Lamo, Perales,& Valentin, 2022）。有鑑於此，我們應該根據不同的因素來考量線上模式的教學方法，並且教師們所扮演的角色亦將更多轉向作為主持人，而非指導者（Mazzara, et al, 2022）。

同樣的狀況在注重科學的 STEM 教學方式上有異曲同工之妙，STEM 教育雖培育了完整的邏輯思考訓練與技術實作之技巧，教學過程中卻缺乏情感溫度與創造力，課堂上的創造力發展取決於教師實施活動的方式與課堂上使用的活動類型(Elaine&Jen,2019) ,也因此加入藝術與美感的 STEAM 模式才應運而生。然而，STEAM 教育、藝術教育與創造力培養的模糊性，亦會阻礙教學的發展 ,並可能使實踐者難以實施有效的 STEAM 教育方法(LaJevic, 2013 ;Liao, 2016；Quigley & Herro, 2016；Elaine & Jen, 2019)。

2.2. 沉浸式學習與社群互動

雖然 Covid-19 的流行引發全球數位化工作的新方式，並且 VR 技術提供了許多潛在的有效解決方案（Taylor & Soneji, 2022），但是在教育領域以 VR 進行學習的方式，仍然有許多層面需透過重新制定策略來實現轉型，其中，符合沉浸式學習與社群互動特性的教育元宇宙平台，應是轉型的一步契機。

社交媒體(social media)此一詞彙被用來描述各式的線上平台，例如社交網站 Twitter、Facebook、虛擬世界如第二人生(Second Life)、協作項目如維基百科、影音共享網站如 YouTube 等（Demangeot & Broderick, 2010, Mangold & Faulds, 2009, Lau, 2017)。而社群互動則是一項複雜的現象，它涉及到語言與非語言行為的不同維度、不同背景、參與者數量以及技術中介(Hanne , Ezequiel, & Shaun ,2010)。有鑑於當代社交平台的複雜性，Lisa 等人提供了對代表美國四大平台的中心類別的見解，在他們的研究中所觀察到：雖然平台越來越多被定

調為開展各式活動的空間，但它們在某種程度上，仍舊被視為「社交」（Lisa, Joseph, David, &Ozan, 2021)，而社交互動的行為，對於線上學習的有效性，具有顯著的正面影響（Baber, 2022）。

根據教育元宇宙應用 VR 的沉浸式與高層次互動特性，並結合對學習有正面效果的社群互動功能，教育元宇宙的建置才能夠愈加完善。

2.3. 教學應用與教育元宇宙平台需求

元宇宙作為新的社交交流空間，有著無限的潛力，即便隨著 COVID-19 疫情的趨緩讓其關注度有所下降，但它仍為全球使用者的創作與分享提供了極高的自由度與獨特的沉浸式體驗與各產業應用(圖 1)。



圖 1 元宇宙趨勢變化與各區偏好(Google Trends:Metaverse, 2023)

目前市場中現有的 VR 學習資源上，因技術與教學實作皆尚在發展推廣階段，應用於課程學習未臻完善，在學生自學與教師導學時，皆碰到一些使用上不足的地方。在學者 Suh 與 Ahn（2022）有關元宇宙與小學生日常生活之密切程度的研究報告中，曾指出有 97.9%的小學生有過元宇宙體驗，且 95.5%的小學生認為元宇宙與日常生活有密切的關聯。而諸如教師應該設計課程，讓學生能夠合作、創意地解決問題或執行專案（Bokyung et al., 2021），還有教師與學生無法即時互動、瞭解學習需求等問題，這些都是促使本研究建置教育元宇宙共學平台之理由。我們統整出現行 VR 教學應用之常見問題，並且將針對下列問題，歸納出符合元宇宙要素定義的教育元宇宙平台，詳如表 1：

表 1 VR 教學應用問題與教育元宇宙平台需求

項次	VR 教學應用問題	面向	元宇宙教育平台需求
1	教材單人學習方式,自學及導學不易產生學習互動。	師生	多元化：平台內規劃五大區域，環形內圈規劃中央的互動學習區可供進行學科教材自學;環形外圈提供共創學習區域，亦可讓學生進行學習互動。
2	授課進度壓力,疫情或場域情境之課程異地教學問題。	課程	隨時隨地：只要有 VR 裝備，無論在何時何地，皆能連上線登入共學平台，進行排定課程的學習。
3	教師無法與學生進行互動,且學習歷程不易掌握。	師生	數位身分：在共學平台內，每個人都有其專屬的數位身分，能夠以此身分進入教室與各空間進行共學互動。
4	同學上課使用期間,無法共同討論與共創。	課程	社群：在平台中能夠進入各個空間，與同學、朋友進行共學共創等活動。
5	抽象概念或微觀世界不易理解,透過 VR 可沉浸但不易產生共學。	教材	沉浸感：利用 VR 將抽象概念以具體方式呈現出來，讓學習者能夠身歷其境，並在沉浸中與同儕進行共學互動。
6	傳輸速度不足，課程教學時,無法知道每位學生的學習狀況。	課程	低延遲：平台系統建置教師後台，教師可即時於課中獲得學生的學習歷程等資料，可瞭解學生的學習狀況。

有鑑於上述 VR 教學應用問題與教育平台需求，本研究以師生、課程與教材三面向問題進行平台開發之主軸，以利教師於課程使用時減少硬體操作與軟體適應所產生的使用者挫折問題，提高教育元宇宙平台學習無界限的良好互動體驗。

3. 研究方法與流程

本研究之研究流程分為三個階段，依序為研究準備階段、開發與實驗階段與研究結果階段(如圖 2)。於研究準備階段先設立研究主題，並依據研究動機確定研究方向，之後進行相關內容的文獻探討，最後制定研究方法。

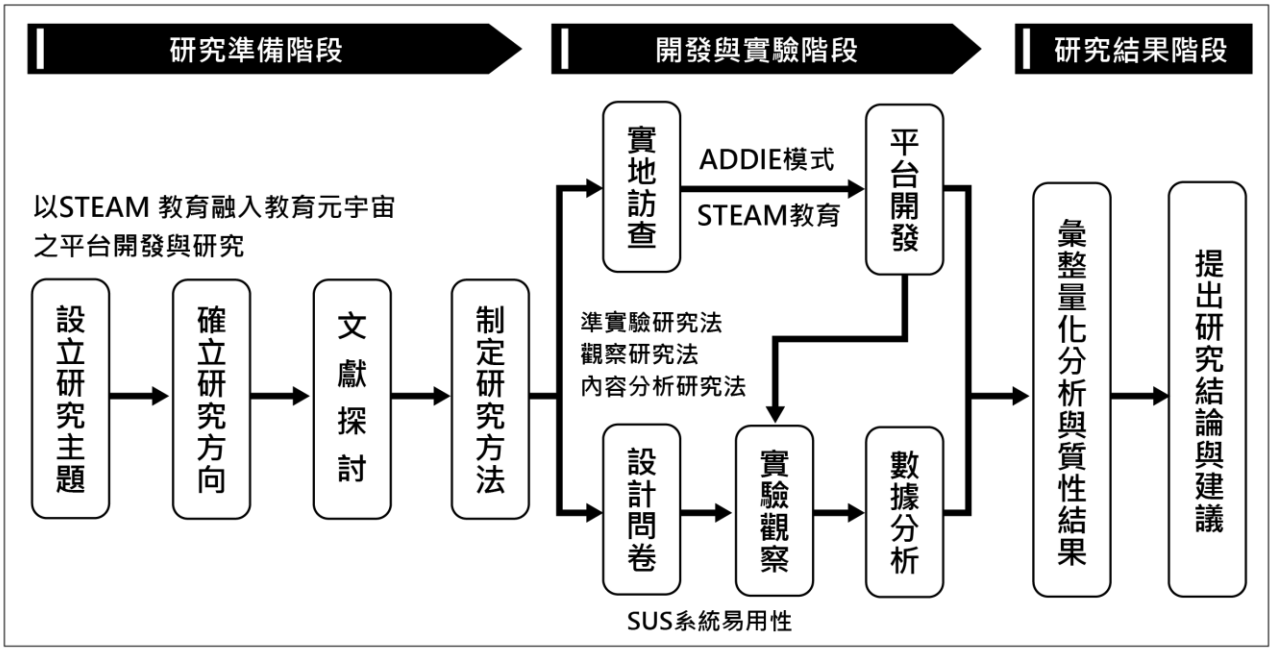


圖 2 研究方法與流程

而開發教育元宇宙與實驗階段，先進行國中校園意見聆聽，再依循 ADDIE 模式進行平台開發。實驗部分，經過系統易用性量表的 SUS 問卷設計，並於進行實驗時觀察，回收量化數據與質性問卷。進行至最後的研究結果階段，彙整上一階段之實驗數據，進行數據分析，並依據數據分析與質化結果，提出以 STEAM 教育融入教育元宇宙之教學應用開發之研究結論與相關建議。

4. 平台開發結果

本研究在教育元宇宙平台開發部分，建置四大核心區域與功能，包括：數位虛擬分身（Digital Twin）、小組跨域討論（Group Discussion）、學習寵物獎勵（Learning Reward）與互動學習區（Interaction and Learning）。各區域都採用第一人稱的互動方式展示虛擬教室、討論室等情境，透過高層次、高互動的學習體驗，以達到沉浸感與共鳴感。而學習寵物獎勵機制透過虛擬代理人偕同學習者在虛擬情境中進行學習互動體驗，並對學習者在各區域所進行的學習互動給予正向情緒回饋。

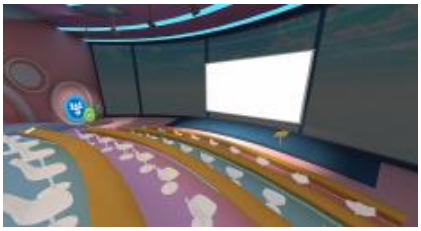


圖 3 互動學習區-虛擬教室



圖 4 小組跨域討論區進行互動



圖 5 數位虛擬分身

除了建置透過虛擬實境進入的教育元宇宙環境外，本研究同時因應教育現場需求，同步建立教育元宇宙之入口網站，以便使用之師長能夠進入教師後臺，課前上傳教材並課後查看學生之學習歷程記錄。

5. 研究結果與分析

本研究共回收受測者年齡 12~15 歲之 95 份問卷，剔除正反問項出現矛盾與漏填答等問題之無效問卷後，有效問卷為男性 43 份與女性問卷 43 份，總共 86 份進行分析。從 SUS 系統易用性分析結果(表 2)能夠發現，第一題項平均取得最高得分 4.22，顯示多數的受測者皆認為教育元宇宙具有親切的操作畫面，且他們非常樂意繼續使用目前呈現方式的介面，而其他題項的得分均有達到平均分之標準，表示教育元宇宙平台無論在系統介面設計與操作應用上，都具有一定之水準。

表 2 SUS 系統易用性分析結果

題號	1	2	3	4	5	6	7	8	9	10	Percentile	Grade	Adjective	Acceptable
全體	4.22	3.17	3.94	3.85	3.97	3.07	3.86	3.15	3.88	3.00	90.29	A	Excellent	Acceptable
男性	4.47	3.14	4.14	4.02	4.21	2.93	4.21	3.05	4.14	2.86	92.91	A	Excellent	Acceptable
女性	3.98	3.21	3.74	3.67	3.72	3.21	3.51	3.26	3.63	3.14	87.67	A-	Good	Acceptable

從整體易用性與可用性之層面來看，全體受測者對教育元宇宙平台的易用性百分數得分為 90.29，分數落在 Bangor 等人(2009)所研究指出之 SUS 分數產品級別(Grade)中之 A 等級，易用性感受為傑出的，可用性為可接受的；而男性受測者的易用性百分數得分為 92.91，產品級別亦為 A 級；女性受測者的易用性百分數得分則為 87.67，產品級別為 A-級，易用性感受為良好的，可用性為可接受的。

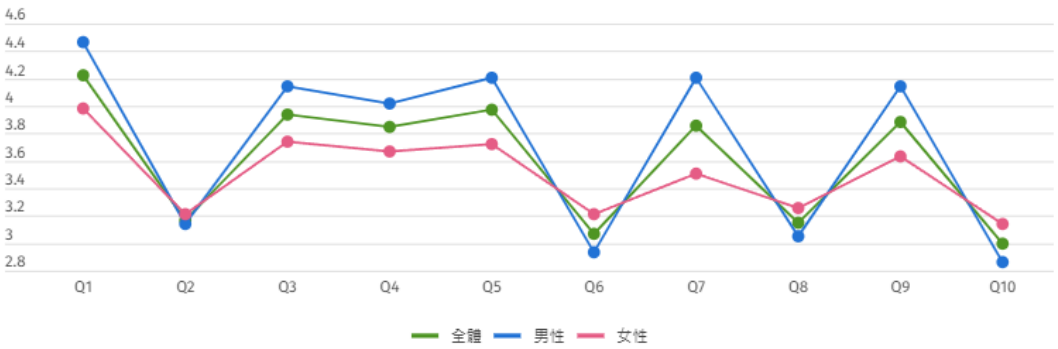


圖 6　SUS 系統易用性折線圖

從 SUS 系統易用性分析結果表(表 2)之數據或數據折線圖(圖 6)呈現，都能夠發現男性與女性受測者在使用教育元宇宙平台的易用性感受不盡相同，男性普遍比女性認為平台系統是更容易操作與容易理解的。此結果有可能之原因為男性與女性在處理資訊、注意力分配以及學習風格等方面有所不同，亦或是目前的平台設計可能更符合男性的使用習慣或喜好。

6. 結論與建議

從問卷上一章分析結果，我們發現目前階段所開發之教育元宇宙系統易用性，在國中生全體受測者間之可用性與易用性感受為可接受且出色的，而在受測者當中，男同學又比女同學覺得教育元宇宙的系統介面與操作性是更加好用的。

此外，從質化回饋的結果內，本研究彙整現行教育元宇宙共學平台使用者所遇見之教學相關問題後，將問題歸類，盤點並提出以 STEAM 教育融入教育元宇宙之教學應用開發可行性的解決策略，期望透過後續研究逐一解決，詳如下表 3。

表 3 教育元宇宙平台問題歸納與融入 STEAM 教育之策略

項次	平台教學問題	STEAM 教育特色	解決策略
1	目前單元式的自學教材，學習內容易被限制於單一科目的既定框架。	教學內容跨領域	將教育元宇宙共學平台之教學模式與課程規劃，至少結合 STEAM 教育中的兩種以上核心概念。
2	同學獨自使用平台上的教材，容易有孤獨感。	多人共學	因 STEAM 教學的進行，都會有三人以上的合作與協助，可彼此共學陪伴。
3	平台教材與生活問題本身大多未具有連結性。	教師引導、學習解決問題	STEAM 模式下，教師會在課程進行中引導，並讓學生學習自行解決問題。
4	單一學科的課程與課後測驗都屬固定答案的學習。	具備開放式表達	每位學生會在 STEAM 課程內進行獨立思考與表達，不會侷限固定的標準答案。
5	課堂具有時間限制，學習反覆思辯的時間不足。	須要腦力激盪	STEAM 課程鼓勵學生反覆思考，並積極進行相互討論與深入分析。
6	課堂位置被教師鎖定，同學間彼此互動性較不足。	具互動性問答	STEAM 學習時，會安排進行互動協作以及即時性的來回問答。
7	單一學科的藝術美感學習不足，缺乏創造力訓練。	要求創造與實作	STEAM 教學強調須讓學生創意思考與實務操作，不鼓勵沿用範例或模仿。

從 STEAM 教育的觀點切入，生活與學習問題密不可分，科學提供人們生活需求所遭遇的問題解決方法，藝術則為生活品質增添溫度與情感。因此根據上表之歸納，在虛擬世界中進行 STEAM 教學，除了可讓學習者的學習環境將不再受到場域的桎梏，教師亦不用再受到現實教材備課不易或科學實驗效果不彰的束縛，更有望改變學習者整體學習氛圍，讓教與學之間的知識、技術與藝術情感相互連結，以此達到教育元宇宙共學平台的完整性，落實教育元宇宙不受限制、學無止境的優勢。

參考文獻

Baber, H.(2022).Social interaction and effectiveness of the online learning – A moderating role of maintaining social distance during the pandemic COVID-19.*Asian Education and Development Studies,11*(1),159-171.https://doi.org/10.1108/AEDS-09-2020-0209

Bangor, A., Kortum, P. and Miller, J. (2009).Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies, 4*, 114-123.

Bokyung,K., Nara,H., Eunji,K., Yeonjeong,P.,& Soyoung,J.(2021).Educational applications of metaverse: possibilities and limitations.*J Educ Eval Health Prof,18*(32),1149230.

Burnard, P., & Colucci-Gray, L. (2019). *Why science and art creativities matter. (Re-)Configuring STEAM for future-making education*. Brill/sense. https://doi.org/10.1163/9789004421585

Demangeot,C. & Broderick, A.J. (2010). Consumer perceptions of online shopping environments Psychol. *Market., 27* (2) , 117-140.

Elaine, P., Jen,K.B.(2019).STEAM in practice and research: An integrative literature review.*Thinking Skills and Creativity,31*,31-43.

Hanne D. J., Ezequiel D.P., & Shaun G.(2010).Can social interaction constitute social cognition?. *Trends in Cognitive Sciences,14*(10), 441-447. https://doi.org/10.1016/j.tics.2010.06.009.

Lau,W.(2017).Effects of social media usage and social media multitasking on the academic performance of university students Comput. *Hum. Behav., 68* , 286-291.

Lamo, P., Perales, M., & Valentin, L. (2022). Case of Study in Online Course of Computer Engineering during COVID-19 Pandemic. *Electronics 2022, 11*(4),578.

LaJevic,L.(2013).Arts integration: What is really happening in the elementary classroom?,*Journal for Learning Through the Arts, 9*(1),1-28.

Liao,C.(2016).From interdisciplinary to transdisciplinary: An arts-integrated approach to STEAM education,*Art Education, 69*(6),44-49.

Lisa R., Joseph B. B., David S. L., & Ozan K. (2021). Social by definition: How users define social platforms and why it matters,*Telematics and Informatics,59*,101538.

Mangold,W.G. & Faulds, D.J. (2009).Social media: the new hybrid element of the promotion mix Bus. *Horiz., 52* (4) , 357-365.

Mazzara, M., Zhdanov,P., Aslam,M., Kotorov,I., Imam,M., Salem,H., Brown,J.A.,& Pletnev, R. (2022). Education After COVID-19. *Smart and Sustainable Technology for Resilient Cities and Communities,232*.193-207.

Nataliia, S.(2021).Using virtual reality in support steam approach for general school. *Scientific Bulletin of Uzhhorod University. Series: «Pedagogy. Social Work»,1*(48),387-390.

Quigley,C.F.&Herro,D.(2016).“Finding the joy in the unknown”: Implementation of STEAM teaching practices in middle school science and math classrooms,*Journal of Science Education and Technology, 25*(3),410-426.

Reevesa, S. M., Crippen, K. J.,& McCray, E.D.(2021). The varied experience of undergraduate students learning chemistry in virtual reality laboratories. *Computers & Education,175*. https://doi.org/10.1016/j.compedu.2021.104320

Suh, W.& Ahn, S.(2022).Utilizing the metaverse for learner-centered constructivist education in the post-pandemic era: an analysis of elementary school students. *J. Intell. 10*(1).

Sharma, G., Prasad, C.,& Rambabu, V.(2021). Online machine drawing pedagogy—A knowledgemanagement perspective through maker educationin the COVID-19 pandemic era. *Knowledge and Process Management,29*(3),231-241.

Taylor, S. & Soneji, S.(2022). Bioinformatics and the Metaverse: Are We Ready? Front. *Bioinform, 2022,2*. [Google Scholar]

Wohlgenannt, I., Simons, A.,& Stieglitz, S.(2020). Virtual Reality. *Business & Information Systems Engineering ,62*, 455-461.

不同模态叠加的教学视频对高中生编程内容学习的影响研究

Effect of different video modal superposition on programming content learning

徐珂[1]，王瑞云[2]，章仪[3]，郭鑫[1]，皮忠玲[4]，杨九民[1*]
[1]华中师范大学 人工智能教育学部，湖北 武汉 430079
[2]清华附中福州学校 福建 福州 350000
[3]北京师范大学 未来教育学院，广东 珠海 519085
[4]陕西师范大学 现代教学技术教育部重点实验室，陕西 西安 710062
[*]yjm@mail.ccnu.edu.cn

【摘要】 中学生通过视频能够有效学习编程知识。以往研究发现学生通过观察教师的手势能更好地理解学习内容。本研究运用具体性衰减理论设计了一组教学序列，探索不同模态的教学视频（单纯录屏、教师操作讲解、录屏+教师操作讲解）对编程学习的影响。被试是 97 名高中生，学习效果、认知负荷和不同维度的学习体验为因变量。结果显示，从具体到抽象的讲解方式能提高学生的学习效果，提高心理距离体验感、沉浸度和推荐意向。动手操作演示降低了学生的内在认知负荷，中介了学习效果。推荐意向由心理距离中介。本研究为未来的编程教学视频设计提供了参考。

【关键词】 视频学习；编程学习；教学序列；具身认知；学习体验

Abstract: *Middle school students can learn programming through videos. Previous research showed that students understand better when they observe teachers' gestures. This study used the concreteness fading theory to investigate the impact of different video types (screen recording, teacher-led explanation, and a combination) on programming learning. Ninety-seven high school students participated, with learning outcomes, cognitive load, and learning experience as variables. The results showed that an explanation approach transitioning from concrete to abstract improved learning, enhanced psychological distance, immersion, and recommendation intention. Hands-on demonstrations decreased intrinsic cognitive load. The intention to recommend was mediated by psychological distance. This study provides insights for future video design.*

Keywords: video learning; programming learning; instructional sequences; embodied cognition; learning experience

基金项目：2023 年度优秀研究生教育创新资助项目“不同类型的虚拟教学代理对学习者视频学习的影响研究”（项目编号：2023CXZZ093）。

1. 前言

教育信息化的快速发展，尤其是 STEM 领域的网络学习如 MOOC 和微课，以视频为核心，重塑了学习形态（Scagnoli et al., 2019）。视频教学在编程等领域尤为重要，因为它通过视听双通道传达知识，打破时间和空间的限制（Mayer, 2005）。

STEM 教学视频可分为单一表征（如纯文本）和多重表征（如文本配图）（Mayer & Moreno, 2003）。单一表征的视频，如代码讲解，往往不能充分利用视频的“多重性、交互性、动态性”优势，可能导致学生注意力分散（龚朝花等人，2018）。而多重表征的视频虽信息量大，但需学生持续注意，可能造成认知负荷过大，无法对所学内容进行理解和内化，也不能发挥教学视频的优势（Sweller, Shinooka-Phelan, & Austin, 2020）。

教学序列的设计同样关键，主要包括具体、抽象和具体性消退三种类型。具体序列通过实例讲解，易于理解；抽象序列则侧重符号化的抽象知识。布鲁纳的具体性消退理论（Concreteness fading）指出，教学应从具体到抽象逐渐过渡，帮助学生理解抽象概念（Bruner, 1974; Xing et al., 2023）。例如，在学习概念时，先从实物表征开始，然后过渡到图形表征，

最后到符号方程式表征，逐步提高抽象程度（Fyfe & Nathan, 2019）。这种方法有助于从多个角度深入理解知识点。

以往关于使用何种教学序列（具体化、抽象化或具体性消退）能更有效地促进学生的学习效果，没有得出一致结论。一些研究发现，从具体到抽象的教学序列有助于学生逐步理解抽象知识，并显著提高学习效果。例如，Fyfe 等人（2019）的研究比较了不同条件下二年级和三年级学生学习等价性概念的效果，发现具体性消退条件下的学生表现出最佳的迁移成绩。然而，也有研究认为具体性消退序列由于内容呈现较多，可能引入无关干扰因素，增加学习者的认知负荷。例如，Braithwaite 等人（2013）研究发现在本科生学习复杂数学概念时，保持一致的抽象表征比具体消退教学序列产生更好的学习效果。此外，有研究指出具体性消退序列和抽象化序列相对不够生动，可能无法吸引学习者的学习兴趣。例如，Tapola 等人（2013）的研究发现只使用具体例子学习的学生在学习电路方面的表现优于具体性消退教学序列，且学习兴趣更浓厚。Zhang 等人（2022）发现在抽象表达之前引入具体表达的教学顺序对大学生学习统计概念有积极影响，可提高学习效果。关于最有效的教学序列（具体化、抽象化或具体性消退）的研究未达成共识。一些研究表明，由具体到抽象的序列能帮助学生更好地理解抽象概念，提高学习成效。

在探讨学生视频学习效果的影响因素时，除了内容和教学顺序，信息的呈现方式同样重要。最新研究表明，视频中融入具体的手部操作展示，而非仅限于图片和文字，能显著提升学习效果。这种方法结合了视觉、听觉及新引入的具体化空间模态，如手翻纸片的视频，能减少学生的认知负担，提供更多元的信息处理方式（Zhang et al., 2022; Paas & van, 2020）。研究显示，教师手势在这种模态中至关重要，通过手势可以传达更丰富的教学内容（Pi et al., 2022; Sweller et al., 2020）。认知神经科学发现，即使不亲自动手，学生在观察类似动作时，其镜像神经元也会被激活，从而从观察中学习（Fu & Franz, 2014）。此外，身体运动对学习的积极作用得到了广泛证实。具身认知理论解释称，身体运动可建立感觉运动与抽象知识间的桥梁，有助于学习者将抽象概念与具体、可感知的经验联系起来（殷明 & 刘电芝，2015）。这不仅提升认知技能，尤其是对 STEM 学科的空间认识能力至关重要，还帮助学习者将认知过程外化至身体或环境，释放更多认知资源以解决问题和进行推理（Rueckert et al., 2017）。因此，在设计教学视频时，加入教师的动手操作可视化，有助于学生更好理解知识点。目前的教学序列研究主要以学习效果为评价标准（Suh et al., 2020），但也有研究关注学生满意度等学习体验相关指标。例如，Shang 等人（2020）在研究具象化表征时，考虑了心理距离、沉浸感、学习满意度及其延续和推荐意向。因此，本研究将采用学习效果、心理距离、沉浸度和推荐意向作为评估指标。

综上所述，关于教学视频中抽象内容具体化设计已有大量研究，但存在局限。一方面，多数研究使用单一教学序列，未充分考虑学习各阶段的特点；另一方面，计算机教学序列的研究相对较少，大多集中在物理、化学和生物等领域；再者，关于融合手势操作的视频模态对学习效果、认知负荷和学习体验的影响的实证研究较少，且针对高中生的研究更是不足。因此，本研究旨在探讨不同视频模态（单纯录屏（RE）、教师操作讲解（TE）和录屏+教师操作讲解（RE+TE））对高中生编程学习的影响，重点关注学习效果、认知负荷和学习体验。

2. 方法

2.1 被试与实验设计

本实验被试为中国福建某高中二年级学生，学生在之前的信息技术课或其他课程中接触到过 python，具备一定的先验知识。本实验共招募被试 97 名，年龄为 14 到 17 岁（$M = 16.47$，$SD = 0.88$）其中女生 32 名，被试随机被分配在三个实验组中，具体分布如下，录屏讲解组（RE）组共有 32 名学生，男生 23 名，女生 9 名，教师操作讲解组（TE）有 33 名学生，男生 21 名，女生 12 名，录屏与讲师讲解组（RE+TE）有 32 名学生，男生 21 名，女生 11 名。所有被试在被告知实验流程后，完成实验后收到一定的被试费。

本实验采用单因素被试间设计，自变量为观看不同模态的教学视频，因变量为学习者的学习效果、认知负荷和学习者的学习体验，包括心理距离、沉浸度、使用意向与推荐意向，并在此基础上进一步进行中介效应探索性分析。

2.3 实验材料

本研究使用教学视频作为学习材料，主题为冒泡排序。视频选自广东教育出版社信息技术选择性必修一，主要知识点是冒泡排序的原理和实现。视频内容参考教材和教辅，经过两位高中信息技术教师评价，适合高中二年级学生。RE 组：教师直接在代码编辑器中录屏并使用抽象化的数学公式总结。设计抽象化的教学序列，如图 1。TE 组：教师采用手移动纸条模拟排序过程，用具体的数字分析进行规律总结。设计具体化的教学序列，如图 2。RE+TE 组：采用手移动纸条模拟排序过程，进行规律总结时先采用具体的数字进行分析演绎，再采用抽象化的数学公式进行总结归纳，并且在书写代码部分结合规律嵌入，如图 3。视频通过摄像头和录屏软件录制，使用 Adobe Premiere Pro CC 2019 编辑，时长分别为 6 分 46 秒、7 分 04 秒、7 分 42 秒。表 1 为教学视频内容说明。

表 1 三种类型的教学视频说明

视频	教学内容	设计策略	时间区间
RE	冒泡排序思想及案例讲解	抽象表征序列	00:00-03:15
	冒泡排序规律总结	抽象表征序列	03:16-04:11
	冒泡排序代码实现	抽象表征序列	04:12-06:46
TE	冒泡排序思想及案例讲解	具体表征序列	00:00-03:14
	冒泡排序规律总结	具体表征序列	03:15-04:30
	冒泡排序代码实现	抽象表征序列	04:31-07:04
RE+TE	冒泡排序思想及案例讲解	具体表征序列	00:00-03:14
	冒泡排序规律总结	具体过渡到抽象序列	03:15-05:07
	冒泡排序代码实现	抽象表征序列	05:08-07:42



图 1 RE 组视频截图

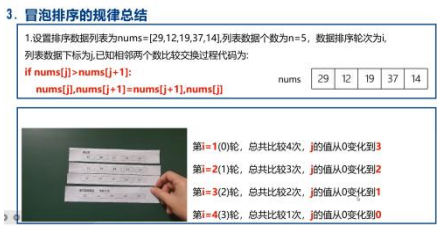


图 2 TE 组视频截图



图 3 RE+TE 组视频截图

2.4 测量工具

2.4.1 人口统计学问卷

人口统计学问卷主要调查了学生的姓名、年龄、性别等基本信息

2.4.2 先验知识测验问卷

该问卷用来测量被试对“冒泡排序”学习内容的了解程度。包括 4 道填空题和 2 道选择题。填空题每题 1 分，选择题每题 2 分，总分共 8 分。前测问卷内部一致性系数 Cronbach's α=0.42。

2.4.3 学习效果测验问卷

学习效果测验包括**保持成绩测验和迁移成绩测验**。**保持成绩测验**问卷用来测量学习者对“冒泡排序”学习内容的记忆情况。包括填空题（5 题，每题 1 分），3 道选择题（3 题，每题 2 分），简答题（1 题，4 分），总分为 15 分。内部一致性系数 Cronbach's α=0.726。**迁移成绩测验**用来测量学习者对“冒泡排序”学习内容的理解和应用情况。包括选择题（5 题，每题 2

分）、简答题（3 题，每题 2 分），总分为 16 分。内部一致性系数 Cronbach's α=0.726。

2.4.4 认知负荷问卷

改编 Leppink 等人（2013）的认知负荷评量表，采用 7 点评分量表，分为内在认知负荷、外在认知负荷和关联认知负荷三个子维度，测量被试在视频学习过程中所投入的认知负荷。三个维度的 Cronbach's α系数分别为 0.933，0.921，0.961。得分越高表示学习者认知负荷越高。

2.4.5 学习体验相关问卷

改编 Shang 等人（2023）的相关量表，采用 10 点评分量表，分别从沉浸度，心理距离（空间距离，社会距离，时间距离，假设距离），推荐意向这三个维度来综合测量学习者的视频学习体验，分数越高说明学习者的主观学习体验越好。三个维度的 Cronbach's α系数分别为 0.862，0.928，0.885。

2.5 实验程序

在学习阶段开始前，教师会向学生说明实验流程和注意事项。学生需完成先验知识测验，并填写基本信息。教师控制学生电脑播放教学视频，视频不可暂停或回放。观看完视频后，学生需完成一系列纸质问卷，包括保持测验、迁移测验、认知负荷和学习体验等。学生完成后上交给教师，教师会再次检查学生基本信息是否填写完整。具体流程如图 4。

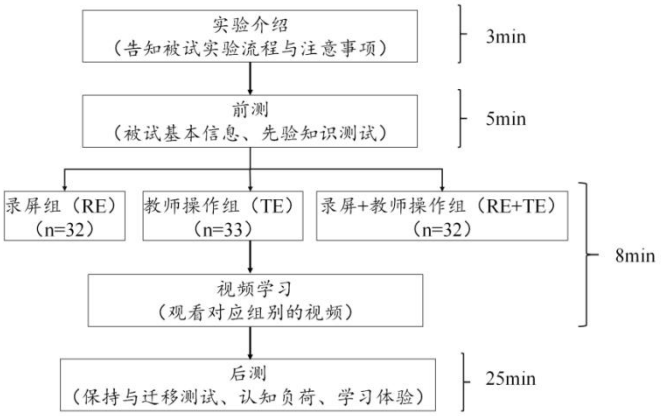


图 4 实验程序

3. 实验结果

本研究以视频模态（RE、TE、RE+TE）为自变量，以学习效果、认知负荷、学习体验为因变量，运用 SPSS 26.0 进行单因素方差分析，描述性统计分析结果如表 2 所示。

表 2 三种实验条件下各因变量的描述性统计结果

测量指标		RE		TE		RE+TE	
		M	*SD*	*M*	*SD*	*M*	*SD*
前测成绩		4.06	2.33	4.76	2.08	4.41	1.72
学习效果	保持测验成绩	7.38	4.51	8.27	3.46	10.41	3.04
	迁移测验成绩	3.88	2.96	3.15	2.12	4.56	3.30
认知负荷	内在认知负荷	3.75	1.93	2.72	1.64	2.99	1.43
	外在认知负荷	3.10	1.90	2.36	1.54	2.45	1.44
	关联认知负荷	5.06	1.83	4.94	1.85	4.56	1.58
沉浸度		5.36	1.95	6.85	2.58	7.67	2.03
心理距离	心理距离—总分	5.80	2.06	6.88	2.23	7.49	2.18
	空间距离	5.53	2.54	6.88	3.01	7.38	2.79
	社会距离	5.41	2.58	6.26	2.86	7.41	2.51
	时间距离	5.47	2.53	6.52	2.80	7.38	2.72
	假设距离	6.27	2.38	7.41	2.36	7.63	2.14

推荐意向	6.45	2.20	7.80	2.08	7.80	2.17

3.1 学习效果

先前知识水平测试的方差结果显示三组之间无显著差异（F (2, 96) = 0.928, p = .399, η^2 =0.02），表明三组学习者的先前知识水平相当。

3.1.1 保持测试成绩

以学习者的保持测验成绩为因变量，进行单因素方差分析，结果发现，三组之间存在显著的差异（F (2, 96) = 5.608, p = .005, η^2 =0.107），*Tamhane T2* 事后分析发现 RE+TE 视频学习组的学习者学习效果最好，显著好于 TE 组与 RE 组（MD = 2. 13, p =.031, MD = 3.03, p = .008），而 TE 组的保持测验成绩与 RE 组无显著差异（MD = 0.90, p = .754）。

3.1.2 迁移测试成绩

以学习者的迁移测试成绩为因变量，单因素方差分析发现，三组之间不存在显著的差异（F (2, 96) = 1.997, p = .141, η^2 =0.041）。

3.2 认知负荷

以视频模态为自变量，以学习者的内在、外在、关联认知负荷为因变量，进行单因素方差分析。结果显示三组的内在认知负荷存在显著差异（F (2, 96) = 3.291, p = .042, η^2 =0.065），外在认知负荷（F (2, 96) = 1.982, p = .143, η^2 =0.040）和关联认知负荷（F (2, 96) = 0.702, p = .498, η^2 =0.015）均不存在显著差异。对内在认知负荷结果进一步经过 *LSD* 事后比较分析发现 RE 组的学生比 TE 组学生认知负荷更高（MD = 1.03, p = .015），与 RE+TE 组没有显著差异（MD = 0.76, p = .073），然而 TE 组与 RE+TE 组也无显著差异（MD = 0.27, p =.514）。

3.3 学习体验

3.3.1 沉浸度

以三组不同实验条件为自变量，沉浸度为因变量，单因素方差分析结果发现，三组之间存在显著差异（F (2, 96) = 9.015, p < .001, η^2 =0.161）。经 *Tamhane T2* 事后比较分析发现，RE+TE 组比 RE 组的学生视频学习沉浸度更高（MD = 2.31, p < .001），与 TE 组没有显著差距（MD = 0.82, p =.402），而 TE 组比 RE 组的视频学习沉浸度更好（MD = 1.49, p =.032）。

3.3.2 心理距离

以三组不同实验条件为自变量，学生心理距离体验感的总分以及不同维度（空间距离、社会距离、时间距离、假设距离）为因变量，单因素方差分析结果发现三组之间有显著差异。

心理距离存在显著的组间差异（F (2, 96) = 5.036, p = .008, η^2 =0.074），*LSD* 事后分析发现，RE+TE 组比 RE 组的学生心理距离体验感更高（MD = 1.69，p =.002），与 TE 组有显著差距（MD = 1.08, p =.047），而 RE+TE 组与 TE 组之间没有显著的差异（MD = 0.61, p =.256）。

空间距离维度组间差异显著（F (2, 96) = 3.751, p = .027, η^2 =0.074），*LSD* 事后分析发现，RE+TE 组比 RE 组的学生空间距离体验感更高（MD = 1.84，p =.010），与 TE 组没有显著差距（MD = 0.50，p =.475），而 RE 组与 TE 组之间也没有显著的差异（MD = 1.34，p =.054）。

社会距离维度组间差异显著（F (2, 96) = 4.568, p = .013, η^2 =0.089），*LSD* 事后比较分析发现，RE+TE 组比 RE 组社会距离体验感更高（MD = 2.00，p =.003），与 TE 组没有显著差距（MD = 1.15，p =.085），而 RE 组与 TE 组之间也没有显著的差异（MD = 0.85，p =.200）。

时间距离维度组间差异显著（F (2, 96) = 4.045, p = .021, η^2 =0.079），*LSD* 事后分析发现，RE+TE 组比 RE 组的学生时间距离体验感更高（MD = 1.91，p =.006），与 TE 组没有显著差距（MD = 0.86，p =.200），而 RE 组与 TE 组之间也没有显著的差异（MD = 1.05，p =.120）。

假设距离维度组间差异显著（F (2, 96) = 3.221, p = .044, η^2 =0.064），*LSD* 事后分析发现，RE+TE 组比 RE 组的学生假设距离体验感更高（MD = 1.35，p =.021），与 TE 组没有显著差距（MD = 0.21，p =.713），而 TE 组比 RE 组的假设距离体验感更好（MD = 1.14，p =.048）。

3.3.3 推荐意向

以三组不同实验条件为自变量，推荐意向为因变量，单因素方差分析结果发现，三组之间存在显著差异（F (2, 96) = 4.249, p = .017, η^2 =0.083）。经 *LSD* 事后比较分析发现，相比于

RE 组的学生，RE+TE 组和 TE 组的学生更愿意推荐他们学习的教学视频（$MD = 1.36, p = .013$, $MD = 1.35, p = .402$），而 TE 组与 RE 组的视频推荐意向没有差距（$MD = 0.01, p = .993$）。

3.4 探索性分析

本研究想要进一步探究不同模态的视频中学习者主观学习体验中，心理距离与对学习者的推荐意向产生的影响，表 3 给出了关键变量的描述性统计和相关性，结果表明，从认知角度出发，认知负荷、保持测试成绩之间有显著的相关，从主观体验角度看，沉浸度、心理距离与其他变量具有一般相关性。

表 3 相关性分析结果

	1	2	3	4	5	6	7	8
1. 保持成绩								
2. 迁移成绩	0.25*							
3. 内在认知负荷	-0.27**	-0.08						
4. 外部认知负荷	-0.32**	-0.21*	0.70***					
5. 关联认知负荷	0.17	0.05	0.08	0.05				
6. 心理距离	0.19	0.11	-0.22*	-0.19	0.14			
7. 沉浸度	0.14	0.08	-0.16	-0.11	0.15	0.83***		
8. 推荐意向	0.16	0.12	-0.001	-0.21*	0.21*	0.70***	0.68***	

注：*p<0.05，**p<0.01，***p<0.001。

因此，进一步进行探索性中介分析，内在认知负荷作为中介变量，学生的保持测试成绩作为因变量。心理距离作为中介变量，推荐意向作为因变量。本研究使用 PROCESS 的模型 4 进行简单中介效应分析（Hayes, 2017）。视频类型为自变量，其中 RE 组为参考，结果发现，相比于 RE 组，TE 组的内在认知负荷与保持测试成绩中介效应显著，而 RE+TE 组内在认知负荷的相对中介效应不显著；TE 组沉浸度与心理距离相对中介效应显著，RE+TE 组的相对中介效应显著，中介效应结果如表 4 所示，路径系数如图 5、图 6、图 7 所示。

表 4 简单中介效应分析结果

中介效应路径	估计值	*SE*	95%CI		中介效应路径	估计值	*SE*	95%CI	
			低	高				低	高
TE-内在认知负荷—保持测试成绩	0.59[a]	0.39	0.01	1.50	RE+TE-内在认知负荷—保持测试成绩	0.43	0.33	-0.05	1.19
TE-保持测试成绩	0.31	0.93	-1.53	2.15	RE+TE-保持测试成绩	2.59	0.92	0.77	4.23
TE-心理距离-推荐意向	0.71[a]	0.38	0.24	1.49	RE+TE-心理距离-推荐意向	1.12[a]	0.40	0.42	1.98
TE-推荐意向	0.64	0.41	-1.17	1.45	RE+TE-推荐意向	0.24	0.42	-0.61	1.08

注：a 表示中介效果显著

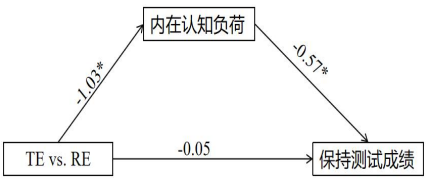


图 5 TE 组内在认知负荷相对中介分析路径图

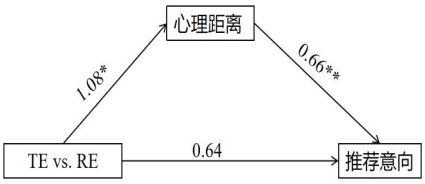


图 6 TE 组心理距离相对中介分析路径图

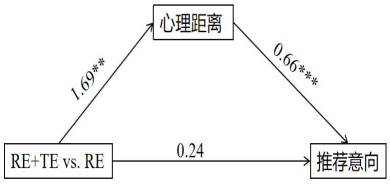


图 7 RE+TE 组心理距离相对中介分析路径图

4. 讨论与结论

本研究探讨了 STEM 领域下，编程学习中不同视频模态（录屏、教师操作讲解、录屏+教师操作讲解）对学习者学习效果、认知负荷及学习体验的影响。结果表明，观看有教师操

作讲解内容的视频（TE、RE+TE）比单纯录屏（RE）更能提高学习效果、改善学习体验和降低内在认知负荷。这表明融合教师操作讲解模态的视频有助于学习者更好地理解抽象内容，对学习效果产生积极影响（Zhang et al., 2022; Fyfe et al., 2014）。

本研究发现，观看融合教师动手操作讲解教学序列的学生学习效果优于观看录屏的学生，而且证明了具体性消退化序列在编程知识中的优势。其中，RE+TE 组的学生学习效果最佳。研究结果与以往部分研究结果一致，融合手势的教学视频对学习有积极作用。学习者仅通过观察手势，也能增强学习效果（Rueckert et al., 2017; Son et al., 2018）。动手操作是动态视觉表现，吸引注意力，实现认知卸载（Zhang et al., 2022）。学习者从教师动手讲解中受益，使抽象概念具体化，易于理解。动手操作增强学习者抽象概念的学习，影响持续（Cook et al., 2013）。另外，RE+TE 组不仅充分利用了视频学习中多元化表征内容的优势，而且具体性到抽象化过度的教学序列设计更有利于学生循序渐进理解所学的编程知识，学习效果更好。

学习者观看融合教师操作的讲解视频，其内在认知负荷明显低于观看录屏的学习者。并且，本研究发现了在具体化序列的 TE 组，内在认知负荷中介了学生的学习效果。以上结果得到具体性衰减理论和认知负荷理论的支持。具身表征有助于将认知加工转移到具身形态，并与学习者的物理世界经验相联系，从而降低认知负荷，提高学习能力（Weisberg & Newcombe, 2017）。观看动手操作视频提供了视觉、听觉外的空间模态，可能激活了排序算法的具体化表征，为学生提供了除语言处理外的信息处理途径，进一步减轻认知负荷（Mayer & Moreno, 2003）。并且，本研究目前尚未发现视频模态对外在和关联认知负荷有显著影响，可能是因为视频模态作为内隐加工活动对外在认知负荷和关联认知负荷影响较小。

本研究发现，观看融合教师操作讲解视频的学习者心理距离体验感、沉浸度、推荐意向高于单纯录屏组。RE+TE 组心理距离体验感显著高于 RE 组，可能是因为丰富的模态表达能缩小心理距离。而没有上下文描述的抽象表达过渡则相反，可能会增加心理距离（La Corte & Piolino, 2016）。在 RE+TE 组中，具体到抽象的过渡、丰富的象征性表达和连贯过渡拉近了学习者与学习材料的心理距离，加强了学习体验感。教师操作讲解活动不仅表征丰富，而且借助手势操作讲解起到了课堂中类似板书的推理作用，增强学生沉浸度（Shanshan et al., 2022）。推荐意向反映用户满意度，教师操作讲解视频有助于提高推理和学习能力，表现为更好的学习效果和满意度，因此学习者更愿意推荐。并且，本研究也发现了具体化的教师操作讲解通过心理距离中介了学生的推荐意向结果。

本研究发现在视频中加入手部操作设计可以降低学习者的内在认知负荷，提高学习者的学习效果，学习沉浸度与视频推荐意向。并且，在视频讲解中加入具体过渡到抽象的讲解序列可以提高学生的学习效果，降低学习者的内在认知负荷，并且拉近学习者与所学内容的心理距离，在学习中沉浸度更高，推荐意向更高。本研究还存有一定的不足之处，本研究仅收集了学生的行为数据，后续研究可以采用眼动技术或脑电技术进一步探索不同教学序列中视融入手部操作模态对学生学习效果影响的内在认知神经机制。

参考文献

龚朝花, 李倩, & 龚勇. (2018). 智慧学习环境中的学习投入问题研究. 电化教育研究, 39(6), 83-89.

殷明, & 刘电芝. (2015). 身心融合学习: 具身认知及其教育意蕴. 课程. 教材. 教法, (7), 57-65.

Braithwaite, D. W., & Goldstone, R. L. (2013). Integrating formal and grounded representations in combinatorics learning. *Journal of Educational Psychology,* 105(3), 666.

Bruner, J. (1974). *Toward a theory of instruction.* Harvard university press.

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child development,* 84(6), 1863-1871.

Sweller, N., Shinooka-Phelan, A., & Austin, E. (2020). The effects of observing and producing gestures on Japanese word learning. *Acta psychologica,* 207, 103079.

Scagnoli, N. I., Choo, J., & Tian, J. (2019). Students' insights on the use of video lectures in online classes. *British Journal of Educational Technology, 50(1),* 399-414.

Shanshan, S., Chenhui, D., & Lijuan, L. (2022). Metaphor and board writing matter: The mediating roles of psychological distance and immersion in video lectures. *Computers & Education,* 191, 104630.

Son, J. Y., Ramos, P., DeWolf, M., Loftus, W., & Stigler, J. W. (2018). Exploring the practicing-connections hypothesis: Using gesture to support coordination of ideas in understanding a complex statistical concept. *Cognitive research: principles and implications,* 3(1), 1-13.

Suh, S., Lee, M., & Law, E. (2020). How do we design for concreteness fading? survey, general framework, and design dimensions. *In Proceedings of the Interaction Design and Children Conference* (pp. 581-588).

Tapola, A., Veermans, M., & Niemivirta, M. (2013). Predictors and outcomes of situational interest during a science learning task. *Instructional Science,* 41, 1047-1064.

Fu, Y., & Franz, E. A. (2014). Viewer perspective in the mirroring of actions. *Experimental brain research,* 232, 3665-3674.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational psychology review,* 26, 9-25.

Fyfe, E. R., & Nathan, M. J. (2019). Making "concreteness fading" more concrete as a theory of instruction for promoting transfer. *Educational Review,* 71(4), 403-422.

La Corte, V., & Piolino, P. (2016). On the role of personal semantic memory and temporal distance in episodic future thinking: The TEDIFT model. Frontiers in human neuroscience, 10, 385.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist,* 38(1), 43-52.

Paas, F., & van Merriënboer, J. J. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science,* 29(4), 394-398.

Pi, Z., Zhang, Y., Yu, Q., Zhang, Y., Yang, J., & Zhao, Q. (2022). Neural oscillations and learning performance vary with an instructor's gestures and visual materials in video lectures. *British Journal of Educational Technology,* 53(1), 93-113.

Rueckert, L., Church, R. B., Avila, A., & Trejo, T. (2017). Gesture enhances learning of a complex statistical concept. *Cognitive Research: Principles and Implications,* 2, 1-6.

Weisberg, S. M., & Newcombe, N. S. (2017). Embodied cognition and STEM learning: overview of a topical collection in CR: PI. *Cognitive Research: Principles and Implications,* 2 (1), Article 38.

Xing, W., Huang, X., Li, C., & Xie, C. (2023). Teaching thermodynamics with augmented interaction and learning analytics. *Computers & Education, 196,* 104726.

Zhang, I. Y., Tucker, M. C., & Stigler, J. W. (2022). Watching a hands-on activity improves students' understanding of randomness. *Computers & Education, 186,* 104545.

元宇宙融入 STEAM 課程之科技接受模式探討

Exploration of the Technology Acceptance Model for Integrating the Metaverse into STEAM Curriculum

吳聲毅[1*]，王政弘[2]，林在營[3]，楊誠達[4]
[1]屏東大學 STEM 教育國際碩士學位學程
[2]高雄大學工藝與創意設計學系
[3]苗栗蕉埔國小
[4]苗栗通霄國小
[*] digschool@gmail.com

【摘要】本研究以改編之元宇宙科技接受模式探討國小學生在透過元宇宙平台進行STEAM跨校共學後的感想。共計 2 校 27 位學生參與元宇宙課程活動，使用之元宇宙平台是 VIVERSE for Business，課程是聲音與樂器單元，使用的 VR 教材是玩樂之聲。經過課程後，學生對於元宇宙融入學習的學習方式接受度高且期待高，對於跟學習成效比較有關係的面向，如想像力與沉浸感的提升亦有幫助。學生也很高興能在元宇宙平台中進行跨校共學，和不同學校的學生一起上課。

【關鍵詞】 元宇宙；STEAM；科技接受模式；虛擬實境

Abstract: *This study investigates the reflections of elementary school students on cross-school collaborative learning through a Metaverse platform, using a modified Metaverse Technology Acceptance Model. A total of 27 students from two schools participated in Metaverse-based curriculum activities. The platform used was VIVERSE for Business, focusing on a unit about sounds and musical instruments, with the VR material being "The Sound of Play". After the course, students showed a high level of acceptance and anticipation for learning integrated with the Metaverse. Aspects related to learning effectiveness, such as the enhancement of imagination and immersion, were also positively impacted. The students expressed their enjoyment in participating in cross-school collaborative learning within the Metaverse platform, engaging with students from different schools.*

Keywords: Metaverse, STEAM, Technology Acceptance Model (TAM), Virtual Reality (VR)

1. 動機與目的

近幾年 STEAM 教育蔚為風潮，主因其強調跨領域的學習與生活問題解決能力的培養。此外，元宇宙(Metaverse)的興起，讓虛擬實境(Virtual reality，VR)沉浸感的學習進展到臨場感的學習場域。

在元宇宙中進行學習活動對於學生來說是一個全新的體驗，目前由於元宇宙平台剛起步，且大多元宇宙平台多是朝向工作場域與藝術展覽的應用為主，較少有可以提供較教育使用的平台，因此對於學生們接受程度的探討較少。

由於一個新穎的學習環境勢必會影響學生的學習感受，進而影響其學習成效。目前也有許多關於科技接受模式(Technology Acceptance Model，TAM)應用於教學平台的探討，如 Sagnier, Loup-Escande, Lourdeaux, Thouvenin, & Valléry (2020)和 Fussell 及 Truong (2023)等。科技接受模式適用於解釋和預測人們對資訊技術的接受程度，接受程度愈高則愈有意願使用 (Huang, Liaw, & Lai, 2016)。

本研究將探討學生透過元宇宙平台進行跨校共學課程的科技接受度與其因之間的關係，以作為日後融入教學之參考。

2. 元宇宙平台介紹

本課程使用之元宇宙平台為宏達電 HTC 公司所開發的 VIVERSE for Business (VFB)為台灣 5G 新科技學習示範學校計畫所改版的教育元宇宙平台，其空間包含大廳、演講廳、小組討論區、分享空間與畫廊等。在演講廳的空間裡，類似教室的上課空間(如圖 1)，可以投影簡報與分享 3D 物件等。學生以虛擬化身(avatar)進入後，由教師安排跨校合作學習活動。



圖 1 教育元宇宙環境(擷取自 HTC 官網)

3. 課程設計

本課程以國小五年級自然與生活科技領域之聲音與樂器作為課程單元，教學內容包含生活中常聽見的聲音與認識噪音，授課時間為 40 分鐘，所以用之教材為玩樂之聲 VR 教材。教學目標包含(1)能透過 VR 教材了解日常所聽到的聲音是噪音或樂音、(2)能知道減少噪音的方法並能實踐在生活中與(3)經由操作了解聲音大小的計數單位。以下為教學活動流程。

【引起動機】

1. PPT 說明今天課程的進行流程
2. 上課以 PPT 複習聲音產生的方式及聲音的傳播方式。
3. 老師示範從元宇宙切換到玩樂之聲體驗。

【發展活動 1-學生自學】

1. 學生進行 VR 教材單元二馬路篇學習，請學生使用 VR 教材，完成每人二種聲音的搜集，

【發展活動 2-組內共學】

1. 請學生回到元宇宙空間，依分組進入討論區討論。
2. 請小組員組內分享自己搜集的聲音，並合作完成學習單問題 1 的觀察記錄。
3. 小組討論，完成學習單問題 2、問題 3 的討論。

【發展活動-組間互學】

1. 各組回到元宇宙的大廳教室，老師依序揭示各組學生學習單。
2. 請各組依序派組員報告學習單的結果。

【統整與結論-教師導學】

1. 增強觀念與歸納本節重點：
   (1)使用平板 APP 噪音測試測量聲音。
   (2)詢問學生，你覺得不舒服的聲音是噪音測試結果數值高或低？
   (3)歸納：音量太大的聲音常令人不舒服、不喜歡，甚至影響身心健康，因此通常把音量太大或令人感覺不舒服、不喜歡的聲音稱為「噪音」。



圖 2 學生自學(學生進行 VR 教材-馬路篇學習)



圖 3 組內共學(回到元宇宙空間依分組進討論)



圖 4 組間互學(各組依序派組員報告學習單)



圖 5 教師導學(教師進行課程總結)

4. 研究設計

本研究採跨校共學的方式進行，由主授學校教師進行授課，其中主授學校與合授學校學生各有 20 人與 7 人參與，共計 27 人，其中男生有 14 人，女生有 13 人。這些教師與學生在進行課程之前對於 VR 頭盔、平板使用與 VFB 元宇宙平台皆已達熟練的操作能力。

科技接受度研究問卷修改自 Huang, Liaw, 和 Lai (2016)所提出來的 A conceptual model of learners' attitudes toward VR learning，其包含感知易用性、感知有用性、互動、想像力、沉浸感與使用意圖六個面向。研究者將其 VR 部分改成元宇宙的情境，經三位專家檢視後進行施測，每個面向包含 3 題，共計 18 題，cronbach's alpha 為 0.71。另包含課後感想之開放問題讓學生自由填答。在進行元宇宙跨校學活動後進行施測，並透過描述性統計與相關分析進行分析與探討。

5. 結果與討論

表 1 呈現科技接受模式各項度的平均值與標準差。從表 1 可以看出，學生對於元宇宙科技接受模式的六個面向整體感受皆達 4 分以上，可見學生可以接受此種新興科技學習的方式。進行分析六個面向可以看出，感知有用性平均值最高，其次是互動與想像力；而最低的面向是感知易用性與沉浸感。由此可見學生對於元宇宙融入教學的期待很高，但可能在平台使用上有一定的流程，學生們覺得有點複雜，以及課程設計較缺少沉浸感的體驗，因此在感知易用性與沉浸感偏低。

表 1 元宇宙科技接受模式子面向平均值

向度	平均值	標準差
感知易用性	4.21	0.96
感知有用性	4.57	0.57
互動	4.44	0.57
想像力	4.44	0.61
沉浸感	4.33	0.78
使用意圖	4.52	0.65

接下來分析科技接受模式之相關性。從表 2 可以看出，感知易用性和互動、想像力、沉浸感與使用意圖皆呈現正相關，顯示一個平台容易使用對於學習的幫助很大。其次，互動與想像力呈現顯著正相關，但互動與沉浸感沒有顯著相關，而沉浸感跟感知易用性有顯著正相關，顯示學生透過新的平台進行學習，在系統操作與互動設計應多著墨。最後，使用意圖和感知易用性、感知有用性、想像力與沉浸感皆呈現顯著正相關，此與科技接受模式的研究相似；但在互動性方面，元宇宙平台與教師在融入教學時可以多加思考互動性的面向。

表 2 科技接受模式之相關分析

	感知易用性	感知有用性	互動	想像力	沉浸感	使用意圖
感知易用性	1					
感知有用性	0.180	1				
互動	.673**	0.101	1			
想像力	.437*	0.376	.452*	1		
沉浸感	.465*	0.220	0.325	.586**	1	
使用意圖	.456*	.503**	0.346	.719**	.621**	1

最後在詢問學生們使用後的感想，大部分學生皆表示有趣、讓課程不會很無聊、比較投入上課狀況、更能夠專注的聽課。此外在元宇宙中可以自由學習到更多的知識，而一般課程只能坐著看課本。最後學生認為，在元宇宙平台中亦能可以和不同學校的學生一起上課，而且不用移動教室。

6. 結論與建議

本研究以改編之元宇宙科技接受模式探討國小學生在透過元宇宙平台進行跨校共學後的感想。其結果顯示學生對於元宇宙融入學習的學習方式接受度高且期待高，對於跟學習成效比較有關係的面向，如想像力與沉浸感的提升亦有幫助。學生也很高興能在元宇宙平台中進行跨校共學，和不同學校的學生一起上課。此外，學生也表示，平台的容易使用與互動功能有待加強。

綜上所述，STEAM 課程透過元宇宙平台進行教學與學習的確能在適合的學習內容上達到助益，學生也樂於嘗試此種新學習環境。以下幾點建議，第一，教師應選擇適合的課程並思考如何提升學習的互動性。第二，教師與學生應熟悉平台與 VR 頭盔的操作。第三，平台廠商可以思考如何加強容易使用與互動性。

參考文獻

Fussell, S. G., & Truong, D. (2023). Accepting virtual reality for dynamic learning: An extension of the technology acceptance model. *Interactive Learning Environments, 31*(9), 5442-5459.

Huang, H. M., Liaw, S. S., & Lai, C. M. (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. *Interactive Learning Environments, 24*(1), 3-19.

Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I., & Valléry, G. (2020). User acceptance of virtual reality: an extended technology acceptance model. *International Journal of Human–Computer Interaction, 36*(11), 993-1007.

小学 STEAM 课程设计的研究整合分析

Research Synthesis on STEAM Curriculum Design System of Primary Schools in China

张莹[1*]，段连鑫[2]，谢利东[3]，李志坚[4]，杨家宽[5]
[1, 2, 3, 4, 5]楚雄师范学院 教育学院
[*]315404538@qq.com

【摘要】 伴随着人才强国战略的提出，培养具有 21 世纪能力素养的国家创新人才是基础教育需要面临的挑战。如何结合我国学生学情在小学阶段设计适合的 STEAM 课程成为了重要课题。结合目前国内小学 STEAM 课程设计研究成果，发现 STEAM 课程设计研究没有下移到实践层面，其次 STEAM 课程理论基础研究匮乏的问题。为了促进小学 STEAM 课程设计研究发展，本文提出了理论结合实践以及加强 STEAM 课程设计规范化的研究建议。
【关键词】 STEAM 教育；课程设计；研究整合分析

Abstract: *The challenge of nurturing national innovative talents with the aptitudes and successes necessary for the 21st century has been a major obstacle for China's basic education since the strategy of reviving the nation through human resource development was proposed.It has been an important task to design suitable STEAM courses for primary school students based on the learning situations of Chinese students. This paper, based on research results of STEAM curriculum design in primary schools of China, has determined that the study was inadequate for practical application.This paper endeavors to advance the advancement of STEAM curriculum design for primary school by combining practice and strengthening STEAM curriculum design, as well as bolstering research into standardizing it. Additionally, there is a dearth of theoretical basic research in STEAM curriculum.*
Keywords: STEAM education, Curriculum design, Research synthesis methods

STEAM教育是体现国家创新人才培养理念的重要途径 (王晶莹, 2021)，STEAM是Science，Technology，Engineering，Art，Mathematics首字母的缩写。STEAM教育源自全美国家科学委员（NSF）的STEM计划 (张一春, 2018)，旨在协助毕业生更进一步地了解和把握未来的技术，并培养他们的创新能力和实践能力。2017 年，《义务高等教育小学科学课程规范》第一次明确提出STEM课堂的概念，并采用该课堂作为基础，通过实践性的项目学习、探索性的实验、实践性的任务完成等多种形式，实现STEM教育的有效组织 (中华人民共和国教育部.基础教育.义务教育小学科学课程标准, 2017)。

因此，本研究目标通过对我国小学阶段STEAM课程设计研究文献进行元分析，从而把握我国STEAM课程设计现阶段针对小学学段的研究发展。希望对我国小学阶段STEAM课程设计的道路探索提供一定的参考依据。

1. 对小学STEAM课程设计研究进行系统综述的必要性

在小学阶段开展STEAM课程，是目前当今世界各国政府为训练中小学生问题处理能力、创新与应用能力，以及促进学生高阶思考技能发展过程中较为流行的教学方式 (许林, 2016) 过程中比较受欢迎的做法。STEAM课程旨在帮助学生掌握跨越各学科的知识，让其具备更好的个性化发展、应对挑战的能力 (Labov J B, 2010)。为什么要针对小学STEAM课程设计研究进行系统综述？原因如下：

1.1. 从系统综述的角度来看

本研究中，采用整合法来进行系统性的总结。通过使用统计学方法，对大量独立的研究结果进行了再次分析，可以对小学STEAM课程设计的研究成果形成一个综合的、客观的、一般性的结论，并且在元分析的过程中也会有新的发现，以促进该研究主题的发展。

1.2. 从小学STEAM课程设计的角度来看

基于研究对象即我国小学STEAM课程设计出发，需对课程设计的基本概念进行明确。对课程设计的目标，廖哲勋明确表示：“课程设计的核心是根据教育的目标和课程中的基本元素、各部分之间的紧密联系来制定学校的课程计划、标准和编写教材，这是课程建设系统工程中的关键环节”(廖哲勋, 1999)。从课程设计的组织模式角度来看，施良方主张“将课程设计定义为一种特定的课程组织方式，这主要包括课程的目标设定以及课程内容的筛选和组织”(施良方, 1996)。顾书明认为，课程设计是一种基于教育目标的有效实施，它需要从多个方面综合考虑后，以制定出一套完整的学习课程规划、标准和教材，这是一个复杂而全面的过程 (顾书明, 2015)。

综上，小学STEAM课程设计研究发展现状对小学STEAM教育的发展方向起着至关重要的影响，对其进行系统性的研究综述将能够为形成小学STEAM课程设计范式起到促进作用。

2. 小学STEAM课程设计研究整合研究设计

2.1. 界定研究问题

明确研究问题是小学STEAM课程设计研究现状如何？结合小学生STEAM课程设计研究成果进行元分析来回答这个问题。本研究将搜集与小学STEAM课程设计有关的文献，结合小学STEAM课程设计研究成果现状分析，针对性提出相应建议。

2.2. 检索收集文献

本研究采用中国知网CNKI数据库，在CNKI中国期刊全文数据库中以“STEAM课程设计”+并含“小学”+或含“STEAM课程”为篇名，模糊搜索 2000 年以来的相关研究论文。经过遴选，最后得出 29 篇有效论文。

3. 研究整合分析结果

3.1. 研究者的身份

经过研究者查询分析， 29 篇期刊中 8 篇文献被北大核心以及中文社会科学引文索引（CSSCI）收录。对筛选出的 29 篇期刊中研究身份数量、比例分布分析，如表 1 所示：

表 1

29 篇核心期刊中研究者身份数量、比例分布

研究者身份	研究者数量	所占研究者总数的比例 (%)
高校研究者	31	72.09%
一线教师	9	20.93%
地方区县教育研究室（中心）	3	6.98%

结合表 1，高校研究者占比 72.09%，说明了高校研究者高度关注和积极地参与到STEAM小学课程设计研究中，为后续实践研究提供坚实的理论基础。但是，从另一方面看高校研究者占到研究人数的一半以上，其实说明STEAM课程设计研究在小学阶段依然处于理论基础探索期。在小学STEAM课程设计研究方向中，教师占 20.93%，说明有一部分一线教师对STEAM课程设计有高度关注，并愿意在一线教学中尝试将STEAM课程运用在自己的教学实践中。地方区县相关教育研究室参与相对较少，说明对于小学阶段开展STEAM课程仍未引起足够的重视，有待进一步的发展和传播，提升相关部门的意识与参与度。在高校、

中学和地方区县的教育研究室（中心）之间，有 5 篇论文是由他们共同发表的，占总论文的 17.24%。说明机构间合作相对关联较少，独立发文的机构比较多。

3.2. 小学STEAM课程设计研究目的和作用

将研究的目的和作用，区分为基础研究、应用研究、发展研究、评价研究以及预测研究。(裴娣娜, 2004)。结果如表 2 所示：

表 2

期刊论文研究目的和作用统计

研究类型	基础研究	应用研究	发展研究	评价研究	预测研究
文章数量	2	20	4	2	1
所占比例	6.90%	69%	13.79%	6.90%	3.45%

从表 2 可以看出，应用研究文章数目最多，预测研究最少。教育理论探讨与实际应用研究是紧密相连的，它们之间存在着相互的依赖关系。因此，在进行教育研究时，不应只关注教育实践的探讨，更应重视理论层面的研究。在各种研究领域里，基础研究与应用研究具有相对的性质。基础研究在实践中起到了指导作用，而应用研究则能进一步丰富基础研究的理论内容，并为理论知识的进步做出积极贡献。在教育实践中，需要理论的引导，而理论又必须建立在实践的基础上。但是，从表 2 的研究类型占比可以明显看出针对小学STEAM课程设计的研究，应用领域明显占据了整个研究类型的大多数，而相对应的基础研究占比较低。

3.3. 小学STEAM课程设计研究方向选取

对筛选的 29 篇期刊文献的研究方向进行分类，整理，分析结果见表 3。

表 3

期刊论文研究方向选取

研究方向	文章数量	比例
小学 STEAM 课程设计开发与流程设计	18	62.07%
国内外研究反思与启发	9	31.03%
学习空间改造应用	2	6.90%

表 3 中可以看出，在有关小学STEAM课程设计的文章中，关于小学STEAM课程设计开发与流程设计研究的研究最多，国内外小学STEAM课程设计的研究反思与启发研究、小学STEAM课程设计与学习空间改造结合的研究文章较少。经研究者深入研究发现，在研究小学STEAM课程设计开发与流程设计方向中主要从课程设计原则、课程设计前端分析、课程活动构建、课程评价多元化设计几个要素方面展开。针对国内外小学STEAM课程设计的研究反思与启发研究，集中探讨了小学STEAM课程设计的跨学科性、结合真实问题以及培养高阶思维能力。在与学习空间改造结合的研究中，着重探究与现代的智慧教室与泛在学习环境相结合的小学STEAM课程设计。

3.4. 小学STEAM课程设计主要包含要素

筛选符合小学STEAM课程设计相关文献中，共计 19 篇关于课程设计包含要素进行了探究，占研究文献数的 65.52%。经研究细分，结果见表 4。

表 4

课程设计的主要包含要素

课程设计要素	比例
小学 STEAM 课程设计原则	6.82%
小学 STEAM 课程设计前端分析	9.09%
小学 STEAM 课程目标	20.45%
小学 STEAM 课程开发	38.64%
小学 STEAM 课程评价设计	25.00%

从表 4 中可以看出，小学STEAM课程设计要素中课程开发占比最高，而针对课程设计原则与课程设计前端分析相对较少。课程设计原则即在进行小学STEAM课程设计时需要依据的准则，根据对文献的深入研究，对于现有小学STEAM课程设计原则研究尚未形成统一的规范性原则指导。

4. 现状和建议

4.1. 目前研究现状

1.从研究者身份上看，小学STEAM课程设计研究没有下移到实践层面。研究者集中于高校研究，在一线教师中开展小学STEAM课程设计研究并未得到重视。

2.从研究的目的和作用上看，STEAM课程理论基础研究匮乏。在高校层面小学STEAM课程设计研究也集中在应用领域研究，主要是对小学STEAM课程设计开发与流程设计研究，对国内外小学STEAM课程设计的研究反思，以及将小学STEAM课程设计与学习空间改造结合的研究。而对STEAM课程设计学理性的基础研究相对来说十分匮乏，尤其在课程设计要素中对课程设计原则与课程设计前端分析研究更为稀缺。

4.2. 未来发展建议

1.加强理论与实践结合。小学STEAM课程设计的应用研究需要植根于坚固的理论基础，教育理论研究是进行教育科学研究的根基。针对小学STEAM 课程设计的研究，必须确立教育理论研究与教育实践研究融合的观念，选择科学合理的研究方法，提高科研主体自身素质，创设良好研究氛围。因此，高校研究者需要加强与一线教师和地方教育研究中心的合作，推进小学STEAM课程设计的研究下沉到一线教学实践中，从教育实践中结合我国学生学情，扎根研究符合我国学生学情的STEAM教育的理论依据，从而科学合理地推进我国STEAM教育发展。

2.强化课程设计的规范化研究。目前我国针对小学STEAM课程设计研究仍处于表面化的阶段。针对小学 STEAM 课程设计基础研究匮乏造成了其课程设计仅仅是将原有的科学或者综合实践课程进行稍加改造的表面化乱象。所以针对小学STEAM课程设计的规范化研究亟需学术界重视，为我国STEAM教育发展提供有源可溯的理论根基。

参考文献

王晶莹、单俊豪和郑永和(2021)。中美 STEM 课程案例的比较研究：知识场域、活动设计、类型与评价。**现代远距离教育(03)**，12-20。doi:10.13927/j.cnki.yuan.20210512.006。
中华人民共和国教育部（2017）。义务教育小学科学课程标准。北京师范大学出版社。
丛立新（2000）。课程论问题。教育科学出版社。
范文翔 & 张一春(2018)。STEAM 教育:发展、内涵与可能路径。**现代教育技术(03)**，99-105。
赵兴龙 和许林(2016)。STEM 教育的五大争议及回应。**中国电化教育**(10)，62-65。
施良方（1996）。课程理论——课程的而基础、原理与问题。教育科学出版社。
顾书明（2015）。课程设计与评价。南京大学出版社。
裴娣娜（2004）。教育研究方法导论。安徽教育出版社。
廖哲勋(1999)。课程发展规律与育人为本的课程设计观。**课程.教材.教法**(05),2-6.doi:10.19877/j.cnki.kcjcjf.1999.05.001。
Labov J B, Reid A H, Yamamoto K R(2010). Integrated Biology and Undergraduate Science Education: A New Biology Education for the Twenty-First Century? **CBE-Life Sciences Education, 9(1)**:10-16。

以 Scratch 軟體設計 SDGs 環境教育數位繪本之 STEAM 課程研究計畫

A Research Project on the STEAM Curriculum of Designing SDGs Environmental Education Digital Picture Books Using Scratch Software

李曉旻[1]，湯梓辰[2*]，陳瑞玲[3]
[12*]國立臺灣科技大學應用科技研究所
[3]新北市深坑國民小學
[2*] jttang0@mail.ntust.edu.tw

【摘要】 STEAM 教育一直是培養學生多元智能及跨學科能力的關鍵，而 STEAM 教育在環境教育中的優點在於促進學生跨學科的思考和行動。綜此，本研究計畫設計一套以「環境教育」為主題的數位互動繪本課程，並融入聯合國 SDGs 永續發展指標，利用 Scratch 程式軟體設計，讓學生從自己的角度出發，探究以 Scratch 程式軟體設計的環境教育數位繪本對學習者 STEAM 能力的影響，分析設計環境教育數位繪本在課堂中的教學效果，以及 Scratch 軟體設計數位繪本在課程中的適用性。研究目前階段主要說明 Scratch 數位繪本課程研究設計過程，將考量在環境教育方面的應用價值，並提出實施策略與評估標準，以期能全面提升學生在科學、技術、工程、藝術和數學領域的綜合素養。

【關鍵詞】 STEAM 教育；Scratch 軟體；SDGs 指標；數位繪本

Abstract: *STEAM education is key in fostering students' multiple intelligence and interdisciplinary skills. The advantage of STEAM education in environmental education lies in fostering students' interdisciplinary thinking and action. Therefore, this project designs a digital interactive picture book course themed on environmental education, integrating the UN's SDGs and utilizing Scratch programming software. It explores the impact of Scratch-designed environmental education picture books on learners' STEAM abilities, analyzing classroom effectiveness and the applicability of Scratch for curriculum design. The current phase outlines the research and design process, considering the application in environmental education, and proposes implementation strategies and evaluation criteria to enhance students' comprehensive skills in science, technology, engineering, arts, and mathematics.*

Keywords: STEAM, Scratch, SDGs, digital picture book.

1. 研究背景

本研究欲將數位敘事與 Scratch 相結合，以永續發展目標為敘事基礎，以應對全球環境挑戰，並評估 STEAM 教育在促進多樣化智力和跨學科技能方面的成效，進而設計一套以「環境教育」為主題的數位互動繪本課程，並整合聯合國 SDGs 永續發展指標。以下將針對本研究之背景，分項探討之。

1.1 Scratch 軟體與應用

Scratch 是一款視覺化編程軟體,使用拖放方式模擬命令,幫助孩子以圖像化的方式學習程式語言，教師可以利用 Scratch 來教授編程概念,透過創建互動式故事、遊戲和動畫。Scratch 程式大多是電腦資訊老師們讓學生們提升運算思維的能力（呂沂蓁，2020；周俊欽，2017）。此程式簡化了難懂的電腦程式語言，以不同顏色的積木呈現程式的功能，因此在視覺上來看，就像有著不同功能的積木，以拖拉的方式將自己的構想呈現出來(施懿珊，2020)。它可以提高

學生對編程的興趣,並訓練他們運算思維和解決問題的能力，老師也可以利用 Scratch 設計好玩的互動課程。

1.2 數位繪本的興起

數位繪本作為一種融合了科技與文學的新型閱讀形式，在兒童教育領域引起了廣泛的關注。藉由結合文字、圖像和聲音，數位繪本形成了一種新的藝術表現形式，豐富了兒童的文學感知。隨著科技的進步，電子書或數位出版普及化，使人類的閱讀行為由單向式的平面文字閱讀，走向互動式的多媒體閱讀，改變了人類學習與思考的模式（洪美珍，2000）。多媒體與數位技術的發展，帶動了以行動裝置進行數位閱讀與人機互動的趨勢。並經由互動參與，使閱讀歷程與行為變得有趣(Henry & Simpson, 2001)。根據 Cochran 與 Bull（1991）的研究，數位繪本可提供多重管道刺激，具有引發學習動機、維持注意力、發揮個別化教學效果、可重複觀看與練習，及得到回饋等功能，有助於提升教學成效。

1.3 STEAM 課程概念

在教育領域中，STEAM 教育一直被視為培養學生綜合性能力和創造力的關鍵。同時，美國的 Next Generation Science Standards（NGSS）為科學教育提供了一個新的指導框架，強調學生的科學探究和跨學科技能，與 STEAM 教育理念有著相似之處，如跨領域學習、動手做能力、應用於生活、解決問題及五感學習等項目。

STEAM 教育，是一種「後設學科」，基於不同學科之間的融合，將原本分散的學科形成一個整體。該概念最初由美國國家科學委員會於 1986 年提出，隨著創意科學受到社會大眾關注，展開 STEAM 教學已成為教育界熱烈討論的議題。STEAM 教育乃協助學生打好紮實的科學、技術、工程、藝術和數學知識的基礎之上。(Chen & Lu, 2018)

1.4 SDGs 相關議題及課程融入

Sustainable Development Goals (SDGs)是聯合國於 2015 年通過的一項全球性議程，宣布的 17 項永續發展目標和 169 項具體目標，展示了這項新的普遍議程的規模和雄心壯志(The United Nations, 2015)。SDGs 的最大優點在於其探討尺度，既可以深入到周遭日常生活的資源回收和交通問題，也可以擴展至全球性的氣候變遷、糧食安全和海洋污染等議題。(張子超，2022)。透過探討 SDGs，學生能夠更深入地了解和思考人類面臨的重大問題，培養他們的全球公民意識和解決問題的能力。此外，SDGs 的實踐也可以通過日常生活中的小行動開始，從而增加學生對可持續發展的實際參與和貢獻

2. 研究目的與問題

本研究目的在於讓學生利用 Scratch 程式設計軟體，讓學生從自己的角度出發，實際創作出一本以「環境教育」為主題的數位繪本：「熊熊除山害」，並藉由學生設計的經驗和喜好，改善設計方向，試圖發現具有潛力的創作題材：

(1) 探究以 Scratch 為基礎的環境教育數位繪本對學習者 STEAM（科學、技術、工程、藝術、數學）能力的影響。

(2) 分析設計環境教育數位繪本在課堂中的教學效果，特別是對學生創造力、問題解決能力和跨學科整合能力的促進。

(3) 探討 Scratch 軟體設計數位繪本在 STEAM 課程中的適用性，以及其對學生學科學習和科技素養的貢獻。

3. 研究流程與工具

本研究工具則分為「Scratch 數位互動繪本-熊熊除山害」、「SGDs 環境教育學習單」及「STEAM 向度觀察紀錄表」，研究流程如圖 1 所示。



圖 1 研究流程

3.1 Scratch 數位互動繪本-熊熊除山害

本實驗採準實驗設計法，以國小高年級的二個班級，一班作為實驗組，一般為對照組，實驗組使用 Scratch 製作 SDGs 數位繪本，對照組將相同的故事以傳統影片進行教學活動。利用程式設計軟體 Scratch，讓學生創作了一本數位繪本「熊熊除山害」，以 Scratch 內建素材庫為主，提供融入 SDGs 環境教育議題的腳本，由學生利用線上文字轉語音網站及錄音程式，加入各分鏡的畫面，套用 Scratch 互動元件模組，完成數位互動繪本。繪本敘述了熊熊發現美麗森林生態受威脅的故事。在此互動式繪本之中，設計了三個和讀者互動的橋段，讓學生從繪本閱讀的過程中，從發現森林的威脅，解決問題，進而發現和探討森林和地球環境保護息息相關，期盼透過繪本的閱讀，融入 SDGs 相關的議題，培養環境保護意識，並讓學生能從自我的角度出發，設計出一套適合學生閱讀的 SDGs 數位互動繪本，培養 STEAM 各向度之素養。本研究設計之繪本及概要如圖 2 和 3 所示：



圖 2 數位互動繪本摘錄



圖 3 繪本所需程式積木摘錄

3.2 SDGs 環境教育學習單

本研究旨在引導學生運用 Scratch 程式設計軟體，以學生視角創作一本以「環境教育」為主題的數位繪本，名為「熊熊除山害」。透過學生創作過程，藉由其經驗喜好，不斷優化設計方向，並探索潛在具有創意潛力的主題。學習單環境教育向度，乃參考 SDGs 相關議題，由三位校內多年豐富教學經驗的自然教師綜整及編寫出下列 SDGs 環境教育學習單內容：

(1) 觀察與解讀：請你描述熊熊生活的地方、熊熊遇到的神秘生態學家給了他什麼啟發、辨識在繪本中森林的大麻煩有哪些呢?

(2) 挑戰任務：保護森林可以降低溫室效應。我們還可以怎麼做呢？列舉出至少三種保護森林的方法。

(3) 反思:你在學習單中了解到了哪些保護森林的方法？你對保護環境有什麼新的認識？如果你是熊熊，你會怎麼參與保護森林的行動呢？

(4)我的行動：選擇一個保護地球的方法，實際行動起來，並分享你的經驗。

3.3 STEAM 向度行為觀察紀錄表

我們設計觀察紀錄表，可以用來記錄學生在學習過程中對於各個 STEAM 向度的表現，以及他們對於繪本中不同元素的理解和應用。這樣的紀錄有助於評估學生的學習成果，同時也能夠提供教師有針對性的反饋和指導。觀察紀錄表如表 1 所示：

表 1 數位互動繪本 STEAM 向度行為觀察表

檢核向度	《熊熊除山害》 數位互動繪本具體設計向度	毫無顧慮的完成該行為(3 分)	猶豫至少 1-2 秒後才完成該行為(2 分)	詢問老師或老師主動指導才完成該行為(1 分)	完全沒做該行為，是老師幫忙操作(0 分)
科學	溫室效應議題、森林在碳循環和氣候調節中的作用				
技術	運算思維、程式及創意設計				
工程	可持續性林業及森林保護工程互動元素				
藝術	圖形編修、動畫創作				
數學	XY 座標、時間、角度設定				

4. 結論

目前研究階段主要介紹了 Scratch 數位繪本課程的設計過程，我們已經成功製作了數位互動繪本並完成了程式編寫的示例。此外，我們還制定了 SDGs 學習單及 STEAM 觀察量表，用以評估學習成效。在接下來的研究階段，我們將進一步實施相關課程，專注探討這項 Scratch 數位繪本課程在環境教育領域的實際應用價值。

參考文獻

呂沂蓁（2020）。**以小組合作學習進行 Scratch 程式設計對國中生運算思維的影響**（碩士論文未出版）。國立臺中教育大學，臺中市。

周俊欽（2017）。**以兩階段提示為基礎之不插電程式設計系統對學生程式運算思維之影響**（碩士論文未出版）。國立臺北教育大學，臺北市。

洪美珍（2000）。**電子童書閱聽型態及其對兒童閱讀影響之研究**（碩士論文未出版）。國立臺東師範學院，臺東市。

張子超（2022）。SDGs 融入中小學教科書自編教材的現況與展望。**教科書研究**，15(2)，113-146。

施懿珊（2020）。運用 Scratch 積木程式於英文繪本製作教學— 一個高職教師的轉化。**臺灣教育評論月刊**，9(10)，92-97。

Cochran, P. S., & Bull, G. L. (1991). Integrating word processing into language intervention. Topics in Language Disorders, 11(2), 31-48.

Henry, R., & Simpson, C. (2001). Picture books & older readers: A match made in heaven. Teacher Librarian, 28(3), 23-271.

Jui-lin Chen, Chow-Chin Lu (2018). The Study on Designing Environmental Education with NGSS and STEAM on Elementary School Senior Grade Students–An Example of"Food and Agriculture Curriculum" Course Design. US-China Education Review A, 8(9), 370-376.

United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development. New York: The United Nations, 2015.

以虛擬實境教具運用在美術繪畫教學之認知負荷與心流經驗之研究

A Study on Cognitive Load and Flow Experience in the Application of Virtual Reality Teaching Tools in Artistic Drawing Instruction

朱志明[1*]，林宜慶[2]
國立宜蘭大學 資訊工程學系
[*] cmchu@niu.edu.tw

【摘要】 將 5G 新科技融入教學中已經成為現代教育的趨勢，虛擬實境頭戴式裝置能貼近真實的情境，讓學習者能沉浸在學習環境中並提高學習興趣。但研究發現，虛擬實境可能會讓學生產生較高的認知負荷，因此，本研究探討以虛擬實境教具運用在國小美術繪畫教學中，對學生的認知負荷與心流經驗之影響。以宜蘭縣某國小六年級美術班學生共 19 名為實驗對象，進行使用虛擬實境教具的美術繪畫教學。研究結果顯示，學生在使用虛擬實境教具進行繪畫學習時，所造成的認知負荷較少，且學生們都有較好的心流經驗，且兩項數值皆呈顯著。
【關鍵詞】 虛擬實境；認知負荷；心流經驗

Abstract: *The integration of 5G technology into education has become a modern trend. Virtual reality headsets, providing a close approximation to real-life situations, allow learners to immerse themselves in the learning environment and enhance their interest in learning. However, research has found that virtual reality may lead to higher cognitive load for students. Therefore, this study explores the impact of using virtual reality teaching tools in elementary school art instruction on students' cognitive load and flow experience.Nineteen sixth-grade students from an art class in an elementary school in Yilan County were selected as participants for the experiment, engaging in art drawing lessons using virtual reality teaching tools. The research results indicate that students experienced lower cognitive load when using virtual reality teaching tools for drawing instruction. Additionally, students demonstrated a significantly better flow experience.*

Keywords: Virtual reality, cognitive load, flow experience

1. 前言

近年來隨著科技的飛速進步，虛擬實境(Virtual Reality, VR)技術已經滲透到各個領域，尤其在教育領域的應用備受關注。虛擬實境可打破時間和空間的限制，將學習場景呈現在學習者面前，讓學習者能以強調互動的方式進行學習(Chen, 2016; Cheng & Tsai, 2019)，因為虛擬實境技術的發展，使其應用更加廣泛，而使用頭戴式裝置也更能模擬真實的情境，讓學習者能沉浸在學習環境中以提高學習興趣(Anglin, Sugiyama, & Liew, 2017)，虛擬實境教具的引入不僅能夠豐富學習體驗，更能打破傳統教學的束縛，為學生和教師提供更為生動、實際且富有互動性的學習環境。虛擬實境教具的引入不僅為學生帶來了嶄新的學習體驗，也為教育者提供了全新的教學手段，虛擬實境教具的優勢在於其能夠提供的沉浸式學習體驗，透過虛擬實境技術，學生可以仿佛置身於實際場景中，與教材中的物體進行互動，這種身臨其境的感覺有助於加深學生對知識的理解，使抽象的概念變得更加具體和實際，而透過模擬實際學習

環境或歷史事件，虛擬實境教具能夠讓學生更好地理解和應用所學知識，培養他們的實際解決問題的能力，因此，在使用虛擬實境教具進行教學時，教學者對虛擬實境教學的課程設計至關重要(Merchant, 2014)。然而，虛擬實境教具的應用也面臨一系列挑戰，其中之一是技術成本，虛擬實境技術的引入需要相應的硬體設備和軟體支援，這可能對學校和教育機構的經濟造成一定的壓力，而城鄉差距會開始顯現，偏鄉與都市的學校在資源上會有顯著的差異存在，當面對科技的變遷時，偏鄉學校設備更新的速度不及都市學校設備更新速度，犧牲了偏鄉學校學生的學習機會，讓學生無法因應科技的變遷；其次是師資的培訓，如果教師要將虛擬實境納入課程中，就可能需要自行設計教材與內容，大多數的教師在備課與課程準備上就需要花費大量的時間，無心於其它的事務，更不用說花時間準備虛擬實境所需要的材料與資源；也有研究指出使用頭戴式虛擬實境設備時，可能因同時提供過多的資訊，導致學生產生較高的認知負荷從而使學生的學習成效降低(Jensen & Konradsen, 2018; Parong & Mayer, 2018)。因此，本研究目的是探討學生在使用虛擬實境教具進行美術繪畫教學時，所造成的認知負荷，及學生的心流經驗為何。本研究的研究問題如下:

(1) 學生在使用虛擬實境教具進行美術繪畫教學時，造成的認知負荷為何？

(2) 學生在使用虛擬實境教具進行美術繪畫教學時，對於心流經驗的影響為何？

2. 文獻探討

2.1. 虛擬實境

虛擬實境作為一種先進的科技應用，近年來在 Meta、HTC 及 SONY 等公司，陸續推出新的虛擬實境頭戴式裝置之帶動下，虛擬實境技術帶來了突破性的發展，頭戴式裝置的價格門檻逐漸降低，也讓虛擬實境愈來愈普及，多個領域中引起了廣泛的關注與應用。從娛樂到醫療保健再到企業培訓和學校教育，尤其在學校教育方面，教師與學童對於在教學中使用虛擬實境教具，大多都抱持著正面的態度，認為其可以提升學習成效並且提高學習興趣(Markowitz, Laha, Perone, Pea, & Bailenson, 2018)。虛擬實境具有沉浸性、想像性、互動性等特性(Burdea & Coiffet, 1994)，其所具備的高沉浸性可以吸引學生的注意力與提高參與感(Cooper et al., 2018; Tcha-Tokeym, Christmann, Loup-Escande, Loup, & Richir, 2018)，但也有研究發現，使用虛擬實境無法完全確保有正向的學習效果(Makransky et al., 2019; Ulrich, Helms, Frandsen, & Rafn, 2019)。

2.2. 虛擬實境技術運用於美術教學

虛擬實境技術可運用於美術教育，首先，名作欣賞是美術教學中不可或缺的部分，引領學生欣賞世界各地著名的藝術作品，如：繪畫、雕塑和建築等，而參觀美術館的實體展覽往往受到時間、距離和費用的限制，因此，傳統教學大多透過教科書、簡報或影片方式來進行。透過講課方式所呈現的藝術品鑑賞，學生很難感受到藝術品的美感，也難以培養準確的鑑賞知識和批判性思考，而虛擬實境技術創建的虛擬空間環境，可以讓學生進入虛擬美術館，彷彿親臨展覽現場一般(Lin, Chen & Lin, 2020)，可詳盡且全方位的欣賞藝術作品。在創作課程方面，學生可以利用虛擬實境技術創作作品，如黏土雕塑和 3D 畫作等。虛擬實境技術克服了材料和空間的限制，開拓了新的創作空間，擴展學生的藝術創作想像力，激發對藝術創作的興趣(Hui, 2020)。憑藉虛擬實境技術的互動功能，學生可以創作藝術品，並翻轉和旋轉他們的作品，最後，學生在虛擬空間中的作品可以展示在虛擬美術館中，透過 3D 列印等技術轉化為實體作品。虛擬實境技術可以擺脫傳統的教學模式，突破時間和空間的限制，擴大美術

課堂的範圍，帶領學生進入各種美術館參觀與學習，擴展學生的視野，為學生提供了許多材料，培養學生創新思維，激發學生的創作靈感。

2.3. 認知負荷

認知負荷是指學習者在進行任務時所需的心智資源(Sweller, 1988)，亦即學習者在學習時，心智努力與心理負荷在工作記憶的心智活動總量(Cooper, 1998; Sweller, Van Merrienboer, & Paas, 1998)。Pass(1988)認為認知負荷包含兩種因素，一個是心智負荷(mental load)，由教材或任務難度所造成，另一個是心智努力(mental effort)，為學生為了完成任務所付出的心力資源。複雜且新穎的任務通常會導致更高的認知負荷，而學習者的經驗和專業能力則可能影響其處理認知負荷的效能(Su, 2016)。結合科技輔助呈現適當的訊息量可促進學習成效(Chu, 2018)。

2.4. 心流經驗

心流是一種完全專注於某個目標或沉浸在一個活動時所產生的心理狀態(Csikszentmihalyi, 1975)。Pearce, Ainley 與 Howard(2005)指出，心流經驗包括控制感(control)、專注力(engagement)及樂趣(enjoyment)等方面。控制感是指學習者可以控制學習環境與活動的程度，專注力為學習者將注意力投入於學習過程的程度，樂趣是指學習者從學習的過程中所感受到的愉悅程度。許多研究指出，提升心流經驗對於學習成效有正面影響(Erchel & Jamet, 2019)，且可以降低認知負荷(Chang et al., 2017)。當個體達到心流時，將感受到適度的控制感、持續的專注及樂在其中(Lin & Chu, 2019)。

3. *研究方法*

3.1. 研究對象

本研究與宜蘭縣某國小合作，採用行動研究法，以六年級美術班學生做為研究對象，男生 5 人，女生 14 人，共 19 人參與實驗。

3.2. 研究流程

本研究流程為期兩週，每週三堂課，每堂課為 40 分鐘。流程如圖 1，授課老師於第一堂課進行 VR 頭盔的佩帶與基本操作說明，約 40 分鐘，讓學生兩人一組進行 VR 頭盔的佩帶與操作練習，約 80 分鐘。第二週即進行使用虛擬實境繪畫軟體「Tilt Brush」進行梵谷相關作品的繪畫，授課老師會在學生有操作困難時給予協助，約 120 分鐘。活動結束後，再請學生完成認知負荷量表、心流經驗量表的填寫，即結束實驗。

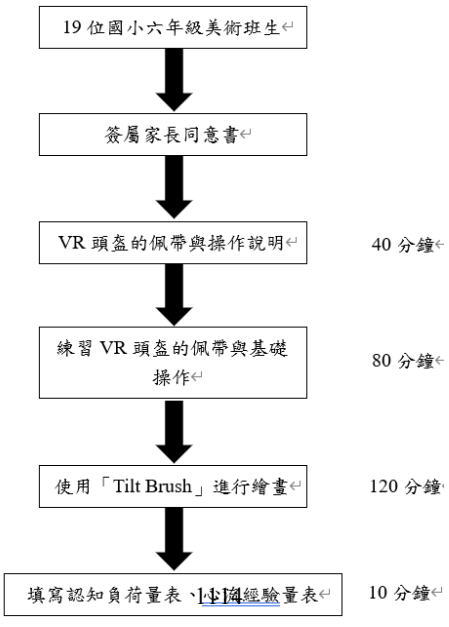


圖 1 研究流程圖

3.3. 研究工具

3.3.1. 認知負荷量表

本研究採用 Hwang、Yang 與 Wang (2013)的認知負荷量表，此量表共 8 題，分為兩個構面，分別為「心智負荷」5 題與「心智努力」3 題，此量表採用李克特式七點量表，1 表示非常不同意，7 表示非常同意。本量表之 Cronbach's α 值為.84，符合 Nunnally (1967)提出之.70 以上表示量表具有良好信度之標準。

3.3.2. 心流經驗量表

為探討學生在使用 VR 頭盔進行繪畫教學時的心流經驗為何，於實驗完成後進行問卷填答，本研究採用 Pearce 等(2005)所編製的李克特五點量表，1 表示非常不同意，5 表示非常同意，題目分為 3 個向度，「控制感」2 題、「專注力」及「樂趣感」各 3 題。本量表之 Cronbach's α 值為.88，具有良好的信度。

4. 研究結果

4.1 使用虛擬實境教具進行美術繪畫教學對認知負荷的影響

本研究採用 SPSS Statistic 29 進行資料分析，以單一樣本 *t* 檢定分析全體 19 位學生，使用虛擬實境教具進行美術繪畫教學時所產生的認知負荷，結果如表 1。心智負荷 t = -3.104，p = .006 < .01 與心智努力 t = -5.127，p =.000 < .001 兩項向度與整體認知負荷 t = -3.994，p =.000 < .001 皆達顯著水準，且兩項因素與整體的平均數皆小於檢定值 4，表示學生在使用虛擬實境教具進行美術繪畫教學時，在虛擬實境教具的使用上沒有太多的障礙，也不須使用過多的心力進行操作，整體來說也不會造成過多的認知負荷，從而影響學習成效。

表 1 學生的認知負荷分析：單一樣本 *t* 檢定摘要表（N = 19）

構面	平均數	標準差	*t*
心智負荷	3.16	1.18	-3.104**
心智努力	2.81	1.01	-5.127***
整體	3.03	1.06	-3.994***

**p < .01 , ***p < .001

4.2 使用虛擬實境教具進行美術繪畫教學對心流經驗的影響

表 2 為學生使用虛擬實境教具，進行美術繪畫教學之心流經驗，進行單一樣本 *t* 檢定的分析結果。結果顯示，控制感 t = 3.622，p =.002 < .01，專注力 t = 7.639，p =.000 < .001，樂趣 t = 8.880，p =.000 < .001，三個向度及整體 t = 7.726，p =.000 < .001，平均上，所有向度皆高於五點量表中位數 3，進一步檢視四個向度的顯著性，皆達顯著水準，顯示使用虛擬實境教具進行的學習的教學方式對於學生來說是有趣的，並可以讓學生投入更多的心思，專注在學習任務上，並且引發強烈的心流。而控制感向度的平均數較其他向度低，其原因推測為多數學生們是首次接觸 VR 頭盔，在系統介面的理解與操作上會較於生疏，以及課堂時間的限制，能使用「Tilt Brush」進行繪畫的時間並不充裕，學生對於完成作品信心稍顯不足。

表 2 學生的心流經驗分析：單一樣本 *t* 檢定摘要表（N = 19）

向度	平均數	標準差	*t*
控制感	3.76	0.92	3.622**
專注力	4.28	0.72	7.639***
樂趣	4.30	0.64	8.880***
整體	4.16	0.65	7.726***

***p* < .01 , ****p* < .001

5. 結論與建議

5.1 結論

本研究旨探討使用虛擬實境教具，進行美術繪畫教學對國小六年級美術班學生的認知負荷與心流經驗之影響。研究結果發現，學生在使用虛擬實境教具進行學習時，並不會對學生造成過多的認知負荷，此結果與陳又菁(2020)、Liu(2020)及張懷綾(2023)等學者的研究結果一致；此外，學生也高度接受在藝術課堂中應用虛擬實境技術進行學習，學生認識和適應這種教學方法，並不會因引入新興技術而產生學習障礙，整體表現出很高的接受意願，學生對於使用虛擬實境教具有較高的心流經驗，虛擬實境帶來的沉浸感與互動性也讓學生感到新奇(Coyne et al., 2019)，使學生能在更短的時間內集中注意力於學習任務上。比起傳統的繪畫教學，使用沉浸式 3D 虛擬環境繪畫更能增強學生的創造表現，虛擬實境技術能為學生帶來更多靈感與想像力(Hui, 2022)，同時，軟體中的附加工具可以幫助學生完成作品，在作畫失誤時也可立即修正，減少學生在繪畫時的挫折感。在實驗過程中，授課教師需適時讓學生取下頭戴式裝置以免發生暈眩，此外，頭戴式裝置的重量也可能對學習者造成負擔，使用時間過長也讓使用者產生頭痛或頸痛等不適，因此使用頭戴式裝置學習，必須控制學習時間長度(Bailenson et al., 2018)，否則會使學習成效下降。

5.2 建議

藉由本研究結果發現，使用虛擬實境教具進行美術繪畫教學，可以引起學生的心流經驗，所造成的認知負荷也較少，本研究目前僅就美術班的繪畫教學進行研究，研究結果無法擴大推論，建議未來可進行虛擬實境教具，運用在其他班級與不同領域所帶來的影響，以改善教學策略及探討虛擬實境技術，如何促進學習者對知識的理解。本研究為配合授課教師與設備數量，研究對象較少，建議若未來研究的經費與設備足夠，可以增加樣本數，以提高研究的可信度。本研究僅針對學生的認知負荷與心流經驗等量化資料進行分析，未來建議可以加入學生教師訪談，了解學生與教師對於虛擬實境教具教學的看法和建議。雖然本研究的研究結果顯示使用使用虛擬實境教具進行美術繪畫教學，不會造成過多的認知負荷，但 Huang (2020) 等人的研究結果指出，不同學習風格的學生，在使用虛擬實境教具學習的過程中，產生的認知負荷也可能不同，因此本研究也建議可嘗試研究不同學習風格或認知風格的學生，探討其在使用虛擬實境教具學習過程中，所產生的心流經驗與認知負荷的差異。

參考文獻

陳又菁 (2020).導入頭戴式虛擬實境於學習成效、心流體驗 與認知負荷之探究：以昆蟲課程為例. **數位學習科技期刊**, **12**(3), 1-23.

Anglin, J. M., Sugiyama, T., & Liew, S.-L. (2017). Visuomotor adaptation in head-mounted virtual reality versus conventional training. *Scientific Reports*, *7*(1). doi:10.1038/srep45469.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2018). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, *17*(1), 102-141.

Burdea, G., & Coiffet, P. (1994). *Virtual reality technology*. Hoboken, NJ: John Wiley &Sons.

Cheng, K.-H., & Tsai, C.-C. (2019). A case study of immersive virtual field trips in an elementary classroom: Students learning experience and teacher-student interaction behaviors. *Computers & Education*, *140*, 103600.

Chu, C. M.* (2018. Nov). Research of Educational Board Games on Learning Effectiveness and Flow State. *The 26th International Conference on Computers in Education* (ICCE 2018). 2018.11.26~30, Manila, Philippines.

Cooper, G. (1998). *Research into Cognitive Load Theory and Instructional Design at UNSW*. http://dwb4.unl.edu/Diss/Cooper/UNSW.htm

Coyne, L., Merritt, T. A., Parmentier, B. L., Sharpton, R. A., & Takemoto, J. K. (2019).The past, present, and future of virtual reality in pharmacy education. *American Journal of Pharmaceutical Education*, *83*(3), 281-290.

Huang, C. L., Luo, Y. F., Yang, S. C., Lu, C. M., & Chen, A.-S. (2020). Influence of Students' Learning Style, Sense of Presence, and Cognitive Load on Learning Outcomes in an Immersive Virtual Reality Learning Environment. *Journal of Educational Computing Research*, *58*(3), 596-615.

Hui J, Zhou Y, Oubibi M, Di W, Zhang L, Zhang S. (2022). Research on Art Teaching Practice Supported by Virtual Reality (VR) Technology in the Primary Schools. *Sustainability*.*14*(3):1246.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013), A concept map-embedded educational computer game for improving students' learning performance in natural science courses, *Computers & Education*, *69*, 121-130.

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *The Official Journal of the IFIP Technical Committee on Education*, *23*(4), 1515-1529.

Lin C-L, Chen S-J, Lin R. (2020). Efficacy of Virtual Reality in Painting Art Exhibitions Appreciation, *Applied Sciences*, *10*(9), 3012.

Lin, H.Y., & Chu, C. M.* (2019). A Study on Flow Experience and Learning Effectiveness of RFID Educational Board Game System. *The 27th International Conference on Computers in Education* (ICCE 2019). 2019.12.02~06, Kenting, Taiwan.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225-236.

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in Psychology*, *9*, 23-64.

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785-797.

Merchant, Z. Goetz, E.T. Cifuentes, L. Keeney-Kennicutt, W. Davis, T.J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, *70*, 29-40.

Pearce, J. M., Ainley, M., & Howard, S. (2005). The ebb and flow of online learning. *Computers in Human Behavior*, *21*(5), 745-771.

Ruixue Liu, Lei Wang, Tiffany A. Koszalka, Kun Wan. (2020). Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning*, *38*(5), 1422-1433.

Su, C.-H. (2016) The effects of students' motivation, cognitive load and learning anxiety in gamification software engineering education: a structural equation modeling study. *Multimedia Tools and Applications*, *75*, 10013-10036.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F.(1998). Cognitive architecture and instructional design. *Educational psychology review*, 10(3), 251-296.

Tsolakidis, C., Zouboula, N., Fokides, E., Vratsalis, C. (2008) Virtual Reality and Museum: An Educational Application for Museum Education. Int. J. *International Journal of Emerging Technologies in Learning*, *3*, 89–95.

Ulrich, F., Helms, N. H., Frandsen, U. P., & Rafn, A. V. (2019). Learning effectiveness of 360° video: Experiences from a controlled experiment in healthcare education. *Interactive Learning Environments*, 1-14.

TPB 框架下中学生 STEM 行为投入的影响机制研究——STEM 专业意向的中介作用

A Study on the Influence Mechanism of Middle School Students' STEM Behavioral Engagement under the Framework of TPB: The Mediating Role of STEM Major Intention

张欣欣[1*]，张屹[2]，贾杰[3]，王莹[4]，赵苏楠[5]，丁双婷[6]
[123456]华中师范大学人工智能教育学部
[*]2633871723@qq.com

【摘要】 本研究采用基于计划行为理论（TPB）的 STEM 态度和 STEM 自我效能来预测七年级学生选择科学、技术、工程和数学（STEM）专业的意向和 STEM 行为投入。TPB 指出，行为的最佳预测因素是执行该行为的意向，而意向受态度和感知行为控制（即自我效能）的影响。来自中国的 968 名初中生参与了这项研究。结构方程模型的结果表明，STEM 态度和 STEM 自我效能独立作用于 STEM 专业意向，专业意向是 STEM 行为投入的显著直接预测因素同时中介了 STEM 态度和自我效能对学生 STEM 行为投入的影响。研究结果对于探讨影响学生 STEM 参与的因素，增强学生学习 STEM 相关学科的意愿具有积极的参考价值。

【关键词】 STEM 态度；STEM 自我效能；STEM 专业意向；STEM 行为投入

Abstract: *This study used Planned Behavior Theory (TPB) based STEM attitude and STEM self-efficacy to predict seventh grade students' intention to choose science, technology, engineering, and math (STEM) majors and STEM behavioral engagement. TPB states that the best predictor of a behavior is the intention to perform that behavior, and that intention is influenced by attitude and perceived behavioral control (i.e., self-efficacy). A total of 968 middle school students from China participated in the study. The results of structural equation model show that STEM attitude and STEM self-efficacy independently affect STEM major intention, and major intention is a significant direct predictor of STEM behavioral engagement and mediates the influence of STEM attitude and self-efficacy on STEM behavioral engagement. The results of this study have a positive reference value for exploring the factors affecting students' STEM participation and enhancing students' willingness to learn STEM related subjects.*

Keywords: STEM attitude, STEM self-efficacy, STEM major intention, STEM behavorial engagement

基金项目：本研究由 2023 年国家自然科学基金面上项目“面向计算思维的中小学人工智能教育框架理论与实践研究”（项目编号：72274076）资助

1. 引言

根据联合国教科文组织(2017)提供的定义，首字母缩略词 STEM 指的是科学、技术、工程和数学。STEM 教育涵盖各种传统和广泛的学科，如医学、工程、物理、化学、生物和计算机科学(UNESCO, 2017)。STEM 教育通过提供必要的技能和知识，为社会的可持续发展奠定了技术基础(Meng et al., 2014)，其重要性不容忽视。许多国家也采取了各种形式的努力和举措来吸引更多的学生加入科学、技术、工程和数学(STEM)的队伍，但 STEM 人才短缺仍然是一个全球性的挑战(Tey et al., 2022)。因此，有必要探究影响学生 STEM 专业意向和行为投入的因素以及它们之间的具体作用机制。这在中国背景下是一个很大程度上被忽视的研究领域。

由于 STEM 动机信念通常在青春期下降(Liou et al., 2021)，这使得探讨此期间青少年的 STEM 意向和行为是重要的，有意义的。为了解决这些问题并填补研究空白，本研究使用结构方程模型(SEM)来评估因素之间的关系并验证假设，计划行为理论为本研究的研究问题提供了指导框架。

2. 文献综述

2.1. 计划行为理论

计划行为理论(Theory of Planned Behavior, TPB)试图解释行为的前因和心理机制(Ajzen, 1991)。TPB 指出，行为的唯一最佳预测因素是执行该行为的意向。反过来，意向部分是由一个人对行为的态度和感知行为控制来预测的(Ajzen, 2002)。先前使用 TPB 框架的研究已经检查了诸如减肥、投票、消费习惯和购买股票等行为，并发现 TPB 广泛有效地解释了行为的可变性(Ajzen, 2005)。近期的文献显示，在教育领域也有很多学者使用 TPB 模型来探索学生的 STEM 职业选择或科学学习意图。例如 Jeffries (2020) 等人以计划行为理论(TPB)的一个改编版本作为指导概念框架考察影响澳大利亚 12 年级学生科学、技术、工程和数学(STEM)学科入学决策的因素，加强学生对 STEM 课程的参与；Tey (2021)等人则运用 TPB 来探讨马来西亚中学 STEM 专业学生 STEM 职业选择的意向。因此，本研究旨在以 TPB 为理论框架分析中学生的 STEM 专业意向和行为投入，聚焦以下两个研究问题：（1）哪些因素会影响中学生的 STEM 行为投入？(2)STEM 专业意向在影响学生 STEM 行为投入的因素中起到怎样的作用？（考虑到丰富的研究对象同时为了确保研究结果的准确性，本研究指的是广义上的 STEM 教育，即与数学、科学等学科有关的教育）

在 TPB 中，态度是个体对执行某种行为反映出有利或不利的评价(Mishkin et al., 2016)。在 STEM 背景下，被定义为对学习 STEM 相关学科所持有的感受或价值观(Osborne et al., 2003)，例如认为 STEM 专业是否有用，是否值得。感知行为控制(可以解释为自我效能感)是个体相信他或她有能力执行某种行为的程度(Fishbein & Ajzen, 2011)。本研究中使用 STEM 自我效能，是指学生相信自己有能力成功完成 STEM 领域的任务，例如我有自信能学好 STEM 专业。有意向是指有意愿实施某种行为，STEM 专业意向就是学生主修 STEM 专业的意图、意愿。如“我愿意选择 STEM 专业进行学习”。本研究中行为表现为 STEM 行为投入，定义为学生在 STEM 活动方面的参与程度(Sahin & Waxman, 2020)。例如“我参加了多少次 STEM 课程相关活动”。根据 TPB，STEM 行为投入受到学生攻读 STEM 专业的意愿（STEM 专业意向）的影响，而专业意向受对 STEM 的态度以及 STEM 自我效能的影响。因此，意向被概念化为两个心理因素和行为之间的中介。

2.2. STEM 专业意向的中介作用

在之前的研究中，许多学者试图通过学生对 STEM 的态度来解释或预测学生在高等教育中进入 STEM 学习的决定。态度被发现可以很好地预测行为意向。如一项对澳大利亚 300 所学校七年级学生的 STEM 科目的态度和 STEM 意向的纵向分析发现，STEM 态度是与学生继续在 STEM 领域学习的意图联系最密切的因素之一(Kennedy et al., 2020)。同样, Gjicali (2021)等人基于国际学生评估项目(PISA)数据的研究表明学生的态度解释了在大学里选择数学专业和主修科学相关专业的意愿的 20%以上的可变性。基于此提出 H1：STEM 态度显著正向预测 STEM 专业意向。而意向与行为又是紧密相关的，如前所述，TPB 称行为的最佳预测者是意向，已有的文献也给出了证据，一项对 47 项关于意向与行为关系的实验研究的荟萃分析表明“意向的中大变化(d = 0.66)导致行为的中小变化(d = 0.36)”(Webb & Sheeran, 2006)。因此可以推测,有意主修STEM专业的学生比无意从事STEM专业的学生更有可能投入精力参与STEM活动，得出我们的假设 H2：STEM 专业意向显著正向预测 STEM 行为投入。

根据 TPB，感知行为控制（自我效能）可以通过意向间接影响行为，也可以直接影响行为。也就是说 STEM 自我效能可以通过专业意向间接影响 STEM 行为投入，也可以直接影响 STEM 行为投入。先前许多研究也证明了自我效能感对学生追求 STEM 的意愿有正向、显著且强烈的影响(Fouad et al., 2017; Kuchynka et al., 2021; Alam et al., 2021)。这表明自我效能会驱动兴趣发展，进而激发专业期望，学生更有可能对自己能力和表现较好的专业感兴趣，而不太可能被自己信心不足的专业所吸引。同时有主修 STEM 专业意向的学生也更有动力和信念参与 STEM 相关活动。但关于 STEM 自我效能与 STEM 行为投入的关系的研究相对较少。最近 Murphy(2023)等人探讨本科 STEM 学生的科学自我效能感对 STEM 行为的影响，发现科学自我效能感会增加学生的 STEM 社区参与行为，即两者之间存在显著的正相关关系。基于以上，本研究假设 STEM 自我效能显著正向影响 STEM 专业意向（H3）和 STEM 行为投入（H4），也就是说学生的 STEM 专业意向介导了 STEM 自我效能与 STEM 行为投入之间的关系。研究的假设模型图如图 1。

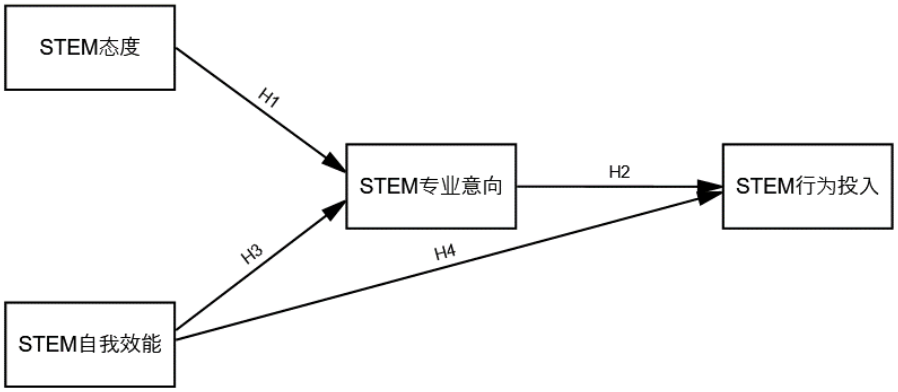


图 1 研究假设模型

3. 研究设计

3.1. 研究对象

本研究采用方便取样法，参与者来自中国中部两省（湖北省和湖南省）均为 7 年级学生，年龄一般为 12-13 岁。所有参与者都审查了知情同意书，问卷是匿名和自愿的，他们可以随时自由撤回并且数据结果保密。参与者通过在线调查平台(http://www.wjx.cn)在计算机或其他电子设备上完成了本次调查。共发放问卷 1000 份，淘汰回答时间过短、规律作答，缺失值过多等无效问卷后，共收集有效问卷 968 份，有效率为 96.8%。其中女生 545 人（56.3%），男生 423 人（43.7%）。在学校类型方面，37.9%的学生来自农村学校，40.9%的学生来自城镇学校，21.2%的学生来自城市学校。

3.2. 测量工具

本研究中的量表均引用成熟的工具，量表的信效度都得到了充分的检验。

STEM 态度量表源自 Lykkegaard，包含 3 个条目，例如“与其他事情相比，我认为学好 STEM 专业对我来说非常重要”(Lykkegaard & Ulriksen, 2016)。采用李克特 5 分制，1 分表示“非常不符合”, 5 分表示“非常符合”，题目的平均分被用作学生的 STEM 态度分数。Cronbach 的 alpha 值为 0.828。

STEM 自我效能量表采用 Lykkegaard 和 Ulriksen(2016 年)编制的量表。共包含 5 个条目，例如“我相信我可以在 STEM 专业中取得优异的成绩”。该量表以 5 分的李克特量表进行评

分，题目的平均分被用作学生的 STEM 自我效能分数。分数越高，STEM 自我效能越强。量表的信度系数为 0.776。

STEM 专业意向由题目“我对报考 STEM 专业的意愿”直接衡量。有“非常愿意”“比较愿意”“一般”“不愿意”“非常不愿意”5 个回答，按李克特量表 5 级标准分别赋值 5 分、4 分、3 分、2 分、1 分，分数越高意味着选择 STEM 专业的意向越强烈。

STEM 行为投入是指学生参与 STEM 相关活动的次数，以“我每周参加 STEM 课程相关活动（如学科竞赛、夏令营、讲座、论坛等）的次数是多少”这一项目来衡量。这个项目用 11 分的李克特量表从 0(没有)到 10(超过 10 次)进行测量。

3.3. 数据处理

IBM SPSS 27.0 用于执行常用方法偏差检验、描述性统计、Pearson 积矩相关分析。首先，为了检验共同方法偏倚是否存在，我们采用了哈曼单因素法(Sahin & Waxman, 2020)。将所有变量进行未旋转的主成分因素分析，检验结果显示，第一个因子方差贡献率为 36.01%，小于 40%的临界值。因此，研究数据不存在严重的共同方法偏差问题。其次，利用 AMOS 26.0 软件对结构方程进行建模，对假设进行验证，参数估计方法为极大似然估计。在结构模型的评估中，计算了方差膨胀因子(VIF)来检查多重共线性问题。最高的 VIF 为 1.564 (VIF<5)，表明不存在共线性问题，因为 VIF 值远低于推荐的截止值(Hair et al, 2017)。

4. 研究结果

学生的 STEM 态度、自我效能感、STEM 专业意向与 STEM 行为投入的皮尔逊相关见表 1。由表可知，上述四个变量之间存在显著正相关关系。

表 1 STEM 态度、自我效能、STEM 专业意向与 STEM 行为投入的 Pearson 相关分析

变量	1	2	3	4
1. STEM 态度	1			
2. STEM 自我效能	0.578***	1		
3. STEM 专业意向	0.340***	0.349***	1	
4. STEM 行为投入	0.135***	0.125***	0.171***	1

注: *p＜0.05，**p＜0.01，***p＜0.001，下同

其次，建立结构方程模型（SEM）来检验研究数据与理论模型的拟合性。在通径分析模型的拟合指标中，所有拟合指标均符合要求，模型适配度良好(χ^2/df=2.848<3, AGFI=0.985>0.9, CFI=0.997>0.9, TLI=0.981>0.9, RMSEA=0.044<0.08) (Arpaci et al, 2016)。同时所有路径系数都是显著的（见表 2）。数据显示学生的 STEM 态度显著预测 STEM 专业意向，STEM 自我效能显著预测 STEM 专业意向和 STEM 行为投入，此外 STEM 专业意向显著正向影响 STEM 行为投入。

表 2 假设检验和模型的路径系数

研究假设	路径	非标准化系数		标准化系数			
		B	Std.Error	Beta	t 值	P 值	假设检验
H1	STEM 态度→STEM 专业意向	0.204	0.036	0.208	5.714	***	支持
H2	STEM 专业意向→STEM 行为投入	0.242	0.056	0.145	4.312	***	支持
H3	STEM 自我效能→STEM 专业意向	0.275	0.044	0.229	6.302	***	支持

H4	STEM 自我效能→STEM 行为投入	0.149	0.067	0.075	2.211	0.027*	支持

对 STEM 专业意向的中介效应进行检验，使用 bootstrap 方法抽取样本 5000 次并估计 95%置信区间，95%置信区间之间没有跨越零则表明结果显著。结果如表 3 所示，STEM 态度和 STEM 自我效能通过 STEM 专业意向介导对 STEM 行为投入的间接影响均显著（点估计值分别为 0.049 和 0.066），95% CI 均未包括零。且两条间接路径之间不存在显著差异（95% CI 跨越零）。其中 STEM 专业意向部分中介了 STEM 自我效能与 STEM 行为投入的关系。中介效应占总效应的 30.70%。综上所述，所有的假设得到了验证，最终的模型图见图 2。

表 3 中介效果检验

Effect		Point Estimate	Product of coefficients		Bootstrapping			
					Bias-corrected 95%CI		Percentile 95%CI	
			SE	Z-value	Lower	Upper	Lower	Upper
直接效果	Direct	0.149	0.070	2.129	0.001	0.282	0.001	0.290
间接效果	Ind1	0.049	0.017	2.882	0.023	0.096	0.019	0.088
	Ind2	0.066	0.022	3.000	0.030	0.119	0.028	0.115
比较	Ind1-Ind2	-0.017	0.020	-0.850	-0.062	0.021	-0.061	0.022

Direct：STEM 自我效能→STEM 行为投入; Ind1: STEM 态度→STEM 专业意向→STEM 行为投入; Ind2: STEM 自我效能→STEM 专业意向→STEM 行为投入;

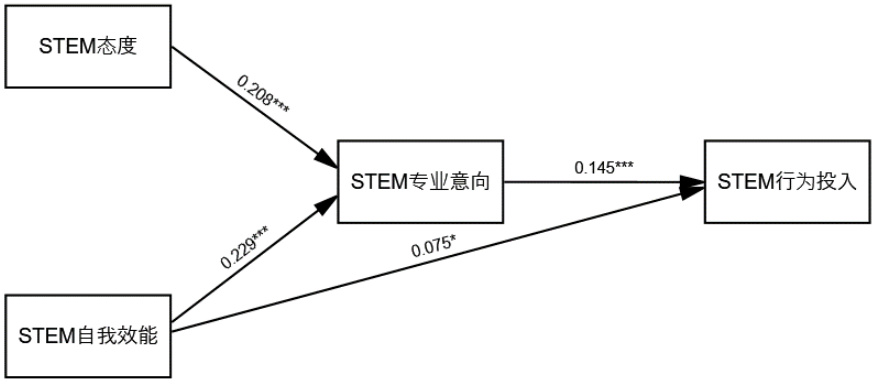


图 2 研究的最终模型

5. 讨论和结论

本研究基于计划行为理论（TPB），通过结构方程模型探讨中学生的 STEM 态度、STEM 自我效能、STEM 专业意向和行为投入之间的关系。研究结果的讨论和结论如下。

首先，STEM 态度是学生 STEM 专业意向的显著直接预测因子，这支持了 H1。这一发现与 Kennedy(2020)和 Gjicali (2021)一致，他们也发现态度对学生的意向有显著影响，也就是说那些相信 STEM 专业对他们的未来有用并且对其感兴趣的学生更有可能选择 STEM 相关专业继续学习，即对 STEM 的积极态度提高了学生学习 STEM 的意愿。此外，STEM 专业意向显著正向影响 STEM 行为投入，支持 H2。这说明有意向从事 STEM 相关专业的学生也会投入更多的时间参与 STEM 相关活动。教育工作者可以据此采取一些干预措施提高学生未来追求

STEM 的几率，吸引更多的学生参与 STEM 课程和活动。例如在 STEM 专业选择中，学生可能对 STEM 有错误的认知（如可能错误地认为在 STEM 领域取得成功太难，STEM 领域很无聊、不重要等），导致他们对 STEM 持消极态度，最终选择不进入 STEM 领域。这些想法都可以通过简单的信息普及和活动经验来纠正，提供机会让学生尝试不同的 STEM 活动，如科学实验、编程、机器人制作等，亲身投入更能使学生发现和了解 STEM 的乐趣和价值（Shaby et al., 2021)。体验式学习、基于问题的学习和以学生为中心，建构主义的学习环境使 STEM 课程和活动变得有趣（Weng et al., 2022)，从而提高学生的行为投入。同时 STEM 行为参与度也可以通过强调 STEM 在日常生活中的实用性来改变(Juškevičienė et al., 2021)。

其次，STEM 自我效能对 STEM 行为投入的影响部分由 STEM 专业意向介导，H3 和 H4 得到支持。以往的研究有类似的结果，Sahin(2023)等人发现数学和科学自我效能感较高的学生更有可能表示他们将在大学里选择 STEM 专业。在强烈的自我效能信念的支持下，有主修 STEM 专业意愿的学生也更有动力参与 STEM 课程和活动(Chan & Wang, 2020)。根据班杜拉的自我效能理论，一旦达到足够的自我效能感水平，它就会成为特定领域动机的来源，人们会付出更多的精力努力投入该领域并坚持更长时间(Bandura, 1997)。低自我效能感会导致回避任务和相关领域，而高自我效能感会促进积极参与和持续追求具有挑战性的任务。这也提醒教育工作者需要重点关注 STEM 的情感维度,有必要培养和提高学生对 STEM 专业的自信心，引导他们勇于尝试，从错误中学习以及学会应对 STEM 学习过程中的失败和挫折，增强学生对 STEM 专业选择的意愿，鼓励更多的 STEM 活动（如学科竞赛、科学展览、讲座和论坛）的参与。

最后，本研究证实了计划行为理论适用于理解 STEM 专业选择意愿和 STEM 行为投入，丰富了 TPB 框架的应用。根据研究结果，教育工作者能够制定有效的策略和干预措施来鼓励学生参加 STEM 课程与相关活动，提高学生对 STEM 科目的积极性和专业入学率，增强学生在 STEM 领域继续学习的能力。

6. 局限性及未来的研究

尽管本研究的结果和贡献是明显的，但仍存在一定的局限性，将在未来的研究中加以解决。首先，我们的调查数据依赖于自我报告工具，这可能导致同源偏差。未来可以通过访谈或行为投入的数据挖掘等多种方法进一步验证。其次，研究采用了横断面设计，STEM 态度和自我效能对 STEM 专业意向以及 STEM 行为投入的影响可能会随着学生学术知识和经验的增加而变化，因此，可能需要纵向设计。在以后的研究中，我们也将持续追踪，引用质性与纵向的研究数据佐证研究结论。此外，本研究的结果只适用于七年级学生。在地域性问题上，研究结果只能代表湖北和湖南省的学生。未来可以扩大研究范围，包括来自其他省份和更高年级的学生，以扩大研究结果的普遍性。最后，未来的研究也可以探索其他潜在的变量来检验学生在 STEM 领域的专业选择和行为投入，或者将其他相关的预测因素整合到 TPB 中，以促进理论的发展。

参考文献

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179-211.

Ajzen, I. (2002). Perceived behavioral control, self‐efficacy, locus of control, and the theory of planned behavior 1. *Journal of applied social psychology*, 32(4), 665-683.

Ajzen, I. (2005). *EBOOK: Attitudes, Personality and Behaviour*. McGraw-hill education (UK).

Alam, M. S., Sajid, S., Kok, J.K., Rahman, M., & Amin, A. (2021). Factors that Influence High School Female Students' Intentions to Pursue Science, Technology, Engineering and Mathematics (STEM) Education in Malaysia. *Pertanika Journal of Social Sciences & Humanities*, 29(2).

Arpaci, I., & Baloğlu, M. (2016). The impact of cultural collectivism on knowledge sharing among information technology majoring undergraduates. *Computers in Human Behavior*, 56, 65-71.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York, NY: Freeman.

Chan, H. Y., & Wang, X. (2020). Reconciling intent with action: Factors associated with the alignment between transfer intent and coursework completion patterns among two-year college students in STEM. *The Journal of Higher Education,* 91(7), 1087-1115.

Fishbein, M., & Ajzen, I. (2011). Predicting and changing behavior: *The reasoned action approach.* Taylor & Francis.

Fouad, N. A., & Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: Moving the needle. *Journal of Career Assessment*, 25(1), 24-39.

Gjicali, K., & Lipnevich, A. A. (2021). Got math attitude?(In) direct effects of student mathematics attitudes on intentions, behavioral engagement, and mathematics performance in the US PISA. *Contemporary Educational Psychology*, 67, 102019.

Hair, J. F., Hult, G. M., Ringle, C. M., & Sarstedt, M. (2017). A primer on partial least squares structural equation modeling (PLS-SEM) (2nd ed.). Sage Publications.

Jeffries, D., Curtis, D. D., & Conner, L. N. (2020). Student factors influencing STEM subject choice in year 12: A structural equation model using PISA/LSAY data. *International Journal of Science and Mathematics Education*, 18, 441-461.

Juškevičienė, A., Dagienė, V., & Dolgopolovas, V. (2021). Integrated activities in STEM environment: Methodology and implementation practice. *Computer Applications in Engineering Education*, 29(1), 209-228.

Kennedy, J., Quinn, F., & Lyons, T. (2020). The keys to STEM: Australian year 7 students' attitudes and intentions towards science, mathematics and technology courses. Research in Science Education, 50, 1805-1832.

Kuchynka, S., Reifsteck, T. V., Gates, A. E., & Rivera, L. M. (2021, December). Developi ng self-efficacy and behavioral intentions among underrepresented students in STEM: the role of active learning. *In Frontiers in Education* (Vol. 6, p. 668239). Frontiers.

Liou, P. Y., Wang, C. L., Lin, J. J., & Areepattamannil, S. (2021). Assessing students' motivational beliefs about learning science across grade level and gender. *The Journal of Experimental Education*, 89(4), 605-624.

Lykkegaard, E., & Ulriksen, L. (2016). Choices and changes: Eccles' expectancy-value model and upper-secondary school students' longitudinal reflections about their choice of a STEM education. *International Journal of Science Education*, 38(5), 701-724.

Meng, C. C., Idris, N., & Eu, L. K. (2014). Secondary students' perceptions of assessments in science, technology, engineering, and mathematics (STEM). *Eurasia Journal of Mathematics, Science and Technology Education*, 10(3), 219-227.

Mishkin, H., Wangrowicz, N., Dori, D., & Dori, Y. J. (2016). Career choice of undergraduate engineering students. *Procedia-social and behavioral sciences*, 228, 222-228.

Murphy, K. M., & Kelp, N. C. (2023). Undergraduate STEM Students' Science Communication Skills, Science Identity, and Science Self-Efficacy Influence Their Motivations and Behaviors in STEM Community Engagement. *Journal of Microbiology & Biology Education*, 24(1), e00182-22.

Osborne, J., Simon, S., Collins, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International journal of science education*, 25(9), 1049-1079.

Sahin, A., & Waxman, H. C. (2020). Characteristics of Secondary Students Who Have Intentions to Choose a STEM Major in College: Findings from a Three-Year Study. *EURASIA Journal of Mathematics, Science and Technology Education*, 16(12).

Sahin, A., Wright, K. B., & Waxman, H. C. (2023). Tracking patterns in secondary students' intention to major in STEM. *International Journal of Science Education*, 45(6), 470-483.

Shaby, N., Staus, N., Dierking, L. D., & Falk, J. H. (2021). Pathways of interest and partic ipation: How STEM‐interested youth navigate a learning ecosystem. *Science education*, 1 05(4), 628-652.

Tey, T. C. Y., Moses, P., & Cheah, P. K. (2021). Human capital in STEM: The influence of attitude, subjective norm and perceived behavioural control on career choice. *Jurnal Penyelidikan Dedikasi*, 16, 106-118.

Tey, T. C. Y., Moses, P., & Cheah, P. K. (2022). Media Exposure and Students' Attitude as Mediators Between Subjective Norms and Choice Intention for Science, Technology, Engineering, and Mathematics Careers. *In 30th International Conference on Computers in Education Conference*, ICCE 2022-Proceedings (Vol. 1, pp. 629-638).

UNESCO (2017). Girls' and women's education in science, technology, engineering, and mathematics (STEM). *United Nations Educational, Scientific and Cultural Organization, France.* Retrieved from http://unesdoc.unesco.org/images/0025/002505/250567e.pdf

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological bulletin*, 132(2), 24 9.

Weng, X., Chiu, T. K., & Tsang, C. C. (2022). Promoting student creativity and entreprene urship through real-world problem-based maker education. Thinking Skills and Creativity, 45, 101046.

6E 模式结合运算思维概念于 AIoT 学习活动对实作表现之研究

The Influence of 6E Model with Computational Thinking for AIoT Learning Activity of Students' Hands-on Performance

萧显胜 [12*]、黄子芸 [1]、苏家莹 [3]、张玉山 [1]
Hsien-Sheng Hsiao[12*], Zi-yun Huang[1], Chia-Ying Su[3], Yu-Shan Chang[1]
[1] 台湾师范大学科技应用与人力资源发展学系、[2] 学习科学跨国顶尖研究中心、[3] 台北美国学校
[*]hssiu@ntnu.edu.tw

【摘要】 随着人工智能技术普及存在并影响我们日常生活，运算思维的重要性也張玉山日益凸显。人工智能物联网(AIOT)课程符合 STEM 趋势，连结现实世界，促进实作经验和知识整合。本研究探讨以不同教学模式（实验组：6E 教学模式结合运算思维概念、对照组：6E 模式）在智能教室为主题的 AIOT 实作课程中，对学生的影响。经由准实验研究设计，研究结果显示实验组的实作表现成绩皆优于对照组。结合运算思维概念的 6E 教学模式能够帮助专题实作中展现更优异之表现。

【关键词】 人工智能物联网、运算思维、6E 模式、实作表现、STEM 教育

Abstract: *With the widespread adoption of artificial intelligence technology, the importance of computational thinking in people's daily lives has become increasingly prominent. The current study of AIoT course that focused on real-world problems and connecting knowledge with everyday experiences. It also aligns with the STEM trend in education and bridges the gap between the virtual and physical world through its emphasis on hands-on experience. The current research investigated the effects of different instructional modes (experimental group: 6E model combined with computational thinking concepts; control group: 6E model) in an AIoT hands-on activity. The research result shown that the experimental group outperformed the control group in terms of hands-on performance. Hence, integrating computational thinking into 6E teaching can improve students' hands-on performance.*

Keywords: AIOT, Computational Thinking, 6E Model, Hands-on performance, STEM Education

1. 前言

根据台湾新课纲规划新兴科技教育成为重要推动方向，明确要求国、高中阶段的学生必须修习信息科技，并强调培养学生的「运算思维」，以培养他们的逻辑性与系统化思考能力，增进问题解决及运算思维能力，使他们能够面对未来多元的工作环境。然而，目前国中以下的学习阶段，学习内容往往偏重于传统学科知识的传授，对于新兴领域的技术和概念相对较少涵盖。近年来，台湾和许多国家的教育部开始重视人工智能教育在学校的推动（Hsu et al., 2018）。在教育活动中，机器学习的应用主要集中在图像识别，语音识别的教育模型较少被应用（Hsu & Chen, 2022）。因此，本研究的目标是规划一门 AIOT(The Artificial Intelligence of Things)课程，引导学生了解 AIOT 的基本架构和程序设计概念。同时，透过训练语音模型，培养学生对机器学习的认识和了解。

综上所述，学习 AIOT 相关技术以因应未来社会需求能力是无法避免的，台湾新课纲中强调学生于国中信息科技课程不仅学习学科内容知识，也重视与其他学科的横向链接性与逻辑性，进而培养运算思维和实作能力。运用 6E(参与（Engage）、探索（Explore）、解释（Explain）、实作（Engineer）、深化（Enrich）、评估（Evaluate）)模式 (Burke , 2014)六步骤及运算思维概念进行实作教学，学生将更积极主动学习，并有系统、逻辑的将问题解决，提升专业知识

及增加动手实作的机会。因此，本研究欲发展一套以智能教室为主题的 AIOT 实作课程，学生透过实作及训练神经网络，以语音的方式控制传感器，经由亲自体验增加学生学习兴趣，进而整合跨领域知识，并运用不同教学模式（6E 结合运算思维概念模式、6E 模式）进行准实验研究，探讨对国中八年级学生的实作表现之影响。本研究之研究目的如下:

A. 发展不同教学模式（6E 模式结合运算思维概念、6E 模式）的 AIoT 实作课程。
B. 探讨不同教学模式（6E 模式结合运算思维概念、6E 模式）的 AIoT 实作课程对于学习者之实作表现的影响。

2. 研究架构

本研究采用准实验研究法，探讨在以 AIoT 为主题之国中信息科技课程中，运用不同教学模式对于学生实作表现之影响。自变项为「教学模式」，分成实验组「6E 模式结合运算思维概念」与对照组「6E 模式」两组，透过量化研究方法对实作表现进行分析。本研究之实验对象为台湾台北市某所国民中学八年级学生，学生采异质性常态编班。实验组和对照组各 2 个班，实验组共 62 人，对照组共 60 人，共 122 人。每个班将学生以 4~5 人进行分组，接受 10 周之实验教学和专题实作。另外学生只具备基础的可视化程序设计知识，尚未接触过 AIOT 和运算思维课程，并缺乏软硬件整合课程的经验。

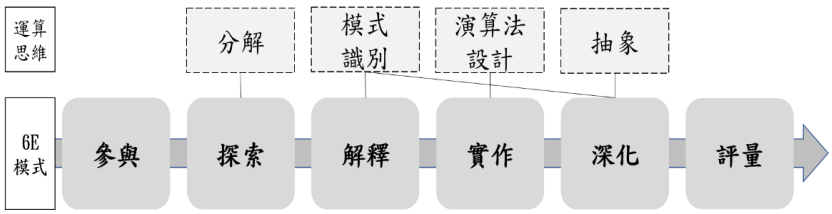
3. 教学活动设计

本研究使用准实验设计，将台北市某国民中学八年级学生分为实验组和对照组，实验组和对照组皆由同一位教学者（即研究者本身）进行教学授课。教学实验共十周，每周课程为 50 分钟，各组于第一周先说明上课规范与计分方式。第二周进行智慧教室议题探讨及 AIOT 之基础教学，学习 AIOT 的基本知识。第三周及第五周进行学习 Arduino 程序设计及认识电子组件及传感器。第六周进行认识机器学习及建立音频模型。第七周与第九周进入连续三周之 AIOT 专题实作，会让学生透过以分组的方式完成作品实作。第十周学习课程活动结束后，进行实作表现后测分析。

3.1. 实验组之教学活动设计

本研究实验组采用「6E 模式结合运算思维概念」，应用 6E 模式(包括参与、探索、解释、建造及深化与评量)并将运算思维四大概念（分解、模式识别、抽象化、算法设计）融入课程教学中，引导学生学习时更深入地理解和应用运算思维的概念。本课程的目标为提供学生认识 AIOT 的基本知识及应用后并进行实作活动。详细之实作活动教学内容如表 1 所示。

表 1
课程阶段流程图

阶段	教学模式概念图
AIOT 专题实作 智慧教室	
教学说明	
1. 参与（Engage）： ● 介绍智慧教室实作成品，引起学生兴趣。	

2. 探索（Explore）、分解概念:
   - 学生实际操作传感器模块，进行探索和观察。
   - 运用分解概念，将智能教室的建置任务分解成更小的子问题，例如控制蜂鸣器模块、设定风扇模块的转向等。
3. 解释（Explain）、模式识别概念:
   - 教师解释传感器模块的原理和功能，引导学生思考如何与传感器模块结合，并使用程序设计控制和读取传感器的数据。
   - 引导学生进行模式识别，辨识出程序设计中的模式和结构，编写相应的程序指令
4. 实作（Engineer）、算法设计概念:
   - 学生透过分组合作，整合前面学到的知识和技能，共同设计和建立智能教室作品，透过手机以蓝牙与 Arduino 开发板连接，使用语音的方式控制传感器，
   - 学生进行算法设计，设计和优化程序指令，以控制传感器和实现预期的功能。
5. 深化（Enrich）、抽象化概念、模式识别概念:
   - 引导学生进一步思考智慧教室的应用场景和未来发展趋势。
   - 鼓励学生抽象化概念，将智能教室的设计和功能应用到其他领域，如智能办公室或智能家居，并引导他们运用模式识别，识别出常见模式，如声音控制传感器或自动调节环境温度等。
6. 评量（Evaluate）:
   - 评估学生的学习成果，小组进行专题成果发表。

3.2. 对照组之教学活动设计

在十周的教学活动中，对照组采用「6E 模式」为基础的课程设计，将于每个学习阶段进行一次，课程中会让学生以小组的方式进行分工、讨论、设计、实作，培养学生团队分工合作及实作等能力。

4. 研究结果

本节探讨实验组「6E 模式结合运算思维概念」与对照组「6E 模式」两种教学模式对 AIOT 实作表现之影响。采用 CPAM 量表来进行学生 AIOT 实作作品的评分，总分为 45 分。为确保两者在评分时是在相同的基准下进行，对此先将两评分者之评分进行一致性考验，以确认两评分者对实作作品之评分标准具高度共识， Kappa 值介于 0.757~0.85 皆大于 0.75，具备好的一致性结果（Cohen, 1960），最后将两位评分者分数平均作为学生之实作表现之分数。

本研究采用独立样本 t 检定分析进行比较，分析后结果如表 2 所示。发现实验组学生实作练习分数高于对照组，实验组与对照组对于实作表现的影响具有显著差异(t= 4.243，p < 0.01)，说明本研究设计的教学课程对于提升国中生之实作表现是有正面影响，与先前 Yang 等人（2020）的研究结果一致，将教学课程融入运算思维的四个概念可以帮助学生发展和提升他们的实作能力，统计分析结果如表 2 所示。

透过运算思维的学习，学生能够培养解决问题、分析和设计解决方案的能力。他们学习如何将复杂的问题分解为更小的子问题，并运用抽象化和模式辨识的技巧来解决问题。同时，也学习如何有效地利用程序设计和算法来实作他们的「解决方案」，使在实作活动中能取得更好的成果。另外，根据实作表现中不同构面的分析，发现「创新性」和「制作与统合」两个构面未达到显著差异。这可能受到本研究课程时间限制或学习环境的影响。学生在这个研究中接受的学习环境相似，包括教室设备和教材资源等。此外，学生在实作过程中可能遵循了相似的设计和装配原则，这也可能影响了结果的可靠性和有效性，导致学生的作品在精致程度和良好手艺方面呈现相似的水平。

表 2.

不同教学法在 AIOT 实作表现之的统计结果

指标	组别	样本数	平均数	标准偏	自由度	*t*	*p*	*Cohen's d*
整体	实验组	62	40.68	2.99	115.216	4.243***	<.001	0.79
	对照组	60	38.15	3.55				
创新性	实验组	62	8.73	.95	120	.274	.784	
	对照组	60	8.68	1.08				
解决方案	实验组	62	18.46	1.11	102.844	7.642***	<.001	1.51
	对照组	60	16.50	1.66				
制作与统合	实验组	62	13.42	1.24	120	1.867	.064	
	对照组	60	12.97	1.43				

****p*<.001

5. 结论

本研究以智能教室为主题进行课程设计和活动安排，课程旨在使学生理解 AIOT 的概念及其在日常生活中的应用，并教导他们传感器和电子组件的功能和用途，Arduino 程序设计以及培养对机器学习的认识。透过讲述和实作练习相结合，让学生不仅理解概念，还能通过实际动手做来实践所学。

在实验组和对照组中，都采用以学生为中心的 6E 模式进行课程设计，学生分组进行工作，进行探究、讨论、设计和实作活动，而教师则扮演辅助者的角色，随时提出问题以了解学生的学习情况和效果，并引导他们完成各项实作任务。通过 6E 模式，学生能够逐步学习，积累知识、概念和技能，并在过程中建立自信心。在实验组中，结合运算思维的四个核心概念，帮助学生更好地分析和解决问题，找到解决方案的步骤，提高问题解决的效率。根据本研究的结果，结合了 6E 模式和运算思维概念教学在 AIOT 专题实作中，观察到对学生的实际操作表现产生显著影响。透过运算思维概念的教学，学生能够更好地分析和解决问题，进一步促进他们在专题实作中的进步。

参考文献

Burke, B. N. (2014). The ITEEA 6E Learning ByDesign™ Model: Maximizing Informed Design and Inquiry in the Integrative STEM Classroom. *Technology and Engineering Teacher, 73*(6), 14-19.

Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*, 37-46.

Hsu, T. C., Chang, S. C., & Hung, Y. T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education, 126*, 296-310.

Hsu, TC. & Chen, MS. (2022). The engagement of students when learning to use a personal audio classifier to control robot cars in a computational thinking board game. *Research and Practice in Technology Enhanced Learning*, **17**-27.

Yang, D., Baek., Y., & Swanson,. S. (2020). Developing computational thinking through project-based airplane design activities. *IEEE Frontiers in Education Conference*, Uppsala.

农村小学 STEM 教育实践的调查研究

杨翘楚，冯雨涵，李葆萍*，朱立新
北京师范大学教育学部
* libp@bnu.edu.cn

The Survey of STEM Education Practices in Chinese Rural Primary Schools

【摘要】本研究采用调查研究的方法，对中国某地区农村小学教育管理者及 STEM 课程相关教师实施问卷调查，发现农村学校 STEM 教育存在资源缺少，课程体系性不强、课程内容单一、师资短缺等问题。为此，本文提出农村学校可以利用数字资源、地域文化等特色因地制宜开展 STEM 课程、加强对课程体系与内容的建设并变革教学模式、深化教师对 STEM 教育理念理解、建立教师专业发展制度体系等建议，以期进一步为农村学校的 STEM 教育发展提供启示。

【关键词】 STEM 教育；农村小学；STEM 课程建设

Abstract: This study's survey of rural primary school educators revealed issues in STEM education, such as resource scarcity, poor curriculum structure, limited course variety, and a shortage of qualified teachers. The paper recommends customizing STEM programs to fit rural contexts, integrating digital resources and local culture. It advocates for enhancing curriculum development, diversifying courses, updating teaching methods, improving teachers' STEM comprehension, and creating a structured professional growth framework to advance rural STEM education.

Keywords: STEM education, rural primary school, STEM curriculum development

1. 绪论

随着世界格局的变化，新一轮科技革命和产业变革正深刻地改变着人类的生产生活乃至学习方式，科技创新和人才已成为各国发展的重要一环。21 世纪学习框架中指出未来人才需要具备创新创造能力、团队合作能力、信息素养等 18 项能力素养（The Partnership for 21st Century Skills, 2009）。STEM 教育有助于培养学习者现实世界问题解决能力、创造力等能力（Martín-Páez et al., 2019），多数国家都正在探索整合 STEM 教育以及提升单独的 STEM 各个科目的水平（Ritz & Fan, 2015）。

STEM 教育超越了传统的教学模式，促进了教学的变革，对于教师与学校都是新的挑战。STEM 教育的实施对学校基础设施与资源支持提出了新的要求，需要学校提供创新技术支持的灵活且包容的学习空间，为师生提供课程所需资源，加强学习者的 STEM 教育体验。同时，STEM 教育需要全新的课程产品及其生成方式（杨开城&公平，2022），STEM 教师需要进行课堂教学探索，创新教学方法，以学生为本进行跨学科、做中学的教学设计。但是在教师们实际教授 STEM 课程时，也很可能因为对 STEM 教育的理念理解不到位导致出现教学目标错位、没有真正实现跨学科融合的情况（袁磊，郑开玲&张志，2020）。

目前学者们多在理论层面探讨 STEM 教育（曾宁，张宝辉&王群利，2018），对于 STEM 教育研究多为城市地区，对于农村、城市郊区等非城市地区的研究较少。因此，本文通过对中国某农村地区小学 STEM 课程开设现状进行研究，主要在区域视角下探讨以下三个研究问题：

（1）中国农村地区小学 STEM 教育的环境及资源建设如何？

（2）中国农村地区小学 STEM 课程开设的情况如何？

（3）中国农村地区小学 STEM 教育师资建设情况如何？

2. 研究设计

2.1. 研究方法

本研究使用问卷调查法，通过地区教育主管部门向相关学校发放 STEM 教育现状调研问卷，收集了学校、教师两个角度的信息。面向学校的调研问卷主要问卷包含学校信息、科技类课程的实施与规划、实验室建设、资源建设等方面。面向教师的问卷主要包含教师的个人基本情况，对 STEM 教育的认知、以及教师专业发展等方面。

2.2. 研究对象

本次问卷调查对象为某农村地区 22 所小学。学校发放对象为小学教育管理负责人；教师发放对象为每个学校的 STEM 课程、科学类学科、数学、信息科技或担任社团任务的教师。共发放了 22 份学校调研问卷，139 份教师调研问卷。最终回收的有效问卷中学校调研问卷共 22 份，教师调研问卷 139 份。

3. 调研结果

3.1. STEM 教育的环境及资源建设

调研的 22 所学校均建立了科学实验室，19 所学校配备了计算机教室，对科学实验室与计算机教室的使用频次为每周 2 次以上，这也说明了绝大多数学校会将科学课、计算机课默认为发展 STEM 教育的主要阵地。部分学校会建设特色的教室，如创客实验室、3D 打印教室、科普展室等，基本每周只会用到 1 次。其中，68.18%的学校会在相关实验室中开展 STEM 课程，59.09%的学校还是在普通教室中开展。资源装备建设方面，学校的 STEM 教育以编程类资源为主，45.45%的学校都配备了编程套件，对于其他数字资源或者 3D 打印机、激光雕刻机等装备各学校拥有的较少。

总体来看，接受调研的农村学校 STEM 教育资源配置种类较少，较少学校会为 STEM 课程专门配备相应资源设备；学校建设实验室种类较少，但基本都具有科学实验室与计算机教室，在科学课与信息科技课程环境建设较为全面。

3.2. STEM 课程实施情况

在课程安排与课程管理方面，18 所学校开设了 STEM 课程，其中 16 所学校都是从小学三年级开设 STEM 课程。在建设 STEM 课程的过程中，81.82%的学校都会采取与校外机构合作的方式， 31.82%的学校教师会自主研发课程、22.73%的学校开始与校外资源合作。这也说明虽然大多数学校都直接使用校外机构的 STEM 课程，但是已经有学校具有课程开发意识，主动尝试着通过校外资源建设 STEM 课程。95.45%的学校都有专门的管理者负责管理学校的科学教育，不过 63.64%的学校感受到目前的 STEM 课程规划体系较不明确。

在具体的课程实施过程中，虽然学校的领导者认为 STEM 课程与社团活动各有不同的课程内容，但是 85.61%的教师都认为学校的 STEM 课程主要实施方式为社团活动，其次是信息技术课程等方式，还有 30.94%的教师提到学校以选修课形式展开 STEM 课程。这也说明各学校对于 STEM 课程的性质并没有很好的界定。其次，对于 STEM 课程的具体教学内容，79.86%的教师认为学校内开展的 STEM 课程内容是编程，其次是创客教育（46.76%）、机器人制作（35.97%）、3D 打印（20.86%）。而在 STEM 课程的讲授方式中，教师们更多会选择以小组合作探究方式展开教学（4.31 分，总分 5 分），讲授式教学均值最低（3.47 分）。但在实地考察中发现编程课教师还是会以讲授式教学为主要上课方式。

课程评价上，68.18%的农村学校都没有建立 STEM 课程教学质量评估制度，72.73%的学校并不会将 STEM 课程中学生的表现作为学生评优的内容。这也说明调研地区的农村学校目前并未重视 STEM 课程评价、课程质量的评价。

3.3. STEM 师资建设

在 STEM 教育的师资配备中，54.54%的农村学校配备了 1-4 名专职的 STEM 课程教师，63.64%的学校都会直接选择让信息科技课程教师兼任，50%的学校选择任用外聘教师。其次，对于参与调研的 STEM 课程相关的教师，其专业与所教学科并未全部一致，如信息科技学科教师来自法律、教育管理、金融、汉语言文学等不同专业。总体来说，各校具有专职、稳定、专业的 STEM 教师数量较少，STEM 教育师资配备较为缺乏。

在 STEM 教师的教研与培训活动中，63.64%的学校没有建立相关教研制度，45.45%的学校从未常规化开展过 STEM 课程的教研活动。然而教师对个人技能的自我评价也不高，如经过培训后 STEM 教育技能的提升（均值 3.88 分）、可以熟练操作 STEM 教育相关设备（均值 3.5 分）。对于教师专业发展的期望，近 70%的教师都希望提升实验与技术能力、学科知识、教学设计与实践能力。47.48%的教师希望可以补充一些教学理论知识，说明 STEM 课程相关教师感觉还是存在很大的专业发展空间，希望提升基本的教师素养。

4. 研究结论与建议

4.1. 研究结论

4.1.1. 学校 STEM 教育条件良好，具备基本的科创教室环境与资源

该区域农村小学 STEM 课程处于发展状态，大多数学校在环境、资源、教师上都有所投入，也有部分学校制定了相关的领导制度，总体开设情况良好。但目前学校 STEM 相关教学资源、设备略为单一，缺少支持 STEM 教育的资源平台。对于其他科创类实验室或教室学校投入较少，使用次数较少，但是部分学校会建设特色活动教室，如观鸟活动室。教师、学校管理者也大多认为目前的 STEM 课程实施困难的原因之一就是缺少课程实施的场地与资源。

4.1.2. STEM 课程以编程为主，课程形式、内容单一，并未形成成熟的 STEM 课程体系

课程内容上，各学校没有正式地规划过 STEM 课程体系，大部分小学的 STEM 课程的内容为编程课，STEM 课程的内容略微单一， STEM 课程的形式以社团课居多。STEM 课程的课时安排上，各校也没有统一标准，均是校内自行决定课时，对相关的课程内容及学习成果的评价缺乏系统性和规范性。而学校没有统一的制度体系，这也会影响教师对于 STEM 课程的认可度与态度。教学方式上，教师们对 STEM 教育理念有所了解，但对 STEM 教育的理解尚浅，教学方式传统模式化，并没有注意跨学科知识学习与运用。

4.1.3. STEM 师资缺乏，教师相关教学能力亟需提升

STEM 教育需要教师必须具备跨学科、跨专业的教育能力与专业知识，对教师素养要求较高，但目前调研的各学校认为 STEM 课程师资紧缺，STEM 教育相关课程教师的专业来源不同，包括一些非教育类专业。教师们也认为目前实验与技术能力、学科知识、学科教学设计与实验能力、教学理论都是迫切需要提升的。

4.2. 研究建议

4.2.1 丰富 STEM 课程资源，增强跨区域合作与资源共享

农村小学可以利用数字资源丰富 STEM 教育内容。该区农村各小学之间缺少 STEM 课程建设之间的交流，区域教育管理部门可以统筹建立通用的 STEM 教育平台，进行资源共享、成功案例参考。各校之间也可以进行资源的共享与协助，共同发展 STEM 教育。

农村学校也可以依托学校的地域特征、民俗文化，因地制宜建设 STEM 课程。农村学校可以充分利用校区周围资源，积极与各类场馆、大学、社区、公园等合作，扩展 STEM 课程

的场地。此外，农村学校之间，农村、城市学校之间也可以探索跨区域合作模式。学校间可进行教师经验交流分享，共享课程资源；也可异校合作，不同学校教师共同开发课程材料，以此克服农村学校资源不充足的问题。

4.2.2. 加强 STEM 课程内容建设，变革教学模式

农村学校可以组织 STEM 课程相关教师，立足学校教师特点，结合农村本土文化，不局限于技术支持的 STEM 课程，有计划地开发 STEM 课程。学校教师可以尝试多元化的教学方法与教学模式，应以学生为主体，多尝试项目式学习、基于问题的学习等教学方法；还可以引进不同的教学主体，如企业专业人员、专家、社区工作者、志愿者等，以社会力量参与农村 STEM 教育实践。

4.2.3. 建立 STEM 教师培养与专业发展的制度体系

学校需要健全 STEM 课程相关教师教研体系和培训体系。农村地区可以邀请相关专家对教师展开专门的线上培训或网络教研，让更多农村地区教师参与进来。农村学校应鼓励教师自主开展 STEM 教育，为教师提供相应的制度支持，如工作量的认定。学校也应建立 STEM 课程教师多元评价机制，激发广大教师参与 STEM 教育的积极性、创造性，提升教师对 STEM 教育的重视程度。除了加强教师的专业知识、教学技能外，还需深化教师对 STEM 教育多元化的理解。教师在教学时，应考虑如何提升学生的学习兴趣、改善学生的学习态度、关注到性别差异，让学生能理解课程的目标、对未来的重要性，在一定程度上解决性别刻板印象带给学生的观念问题（Tam, Chan, & Lai, 2020）。

4.2.4. 革新 STEM 教育管理模式

学校的领导者应对 STEM 建设的意义有深入的理解，可以从学校的基本情况出发，规划学校 STEM 教育的样态，支持学校的 STEM 课程发展。在 STEM 教育的课程建设上，区域教育管理部门应建立一个可参考的、统一且专业的 STEM 课程标准体系，通过相关课程标准来为各校开展 STEM 教育、开发 STEM 课程提供宏观的指导，针对课程性质、课程目标、课程内容有清晰的界定，并根据不同学段学生有所侧重。

参考文献

The Partnership for 21st Century Skills. (2009). P21 framework definitions. Retrieved from http://www.p21.org/storage/documents/P21_Framework_Definitions.pdf.

Ritz, J.M., Fan, SC. (2015). STEM and technology education: international state-of-the-art. *International Journal of Technology and Design Education*, 25, 429–451.

Martín-Páez T, Aguilera D, Perales-Palacios FJ & Vílchez-González JM. What are we talking about when we talk about STEM education? A review of literature. *Science Education*. 2019, 103(4),799–822.

杨开城和公平（2022）。论 STEM 教育何以特殊。**中国远程教育**，**04**，48-54。

袁磊，郑开玲和 张志（2020）。STEAM 教育:问题与思考。**开放教育研究**，**03**，51-57+90。

曾宁，张宝辉 和王群利（2018）。近十年国内外 STEM 教育研究的对比分析——基于内容分析法。**现代远距离教育**，**05**，27-38。

Tam, H., Chan, A. Y., & Lai, O. L. (2020). Gender stereotyping and STEM education: Girls' empowerment through effective ICT training in Hong Kong. *Children and Youth Services Review*, 119, 105624.

国内外中小学 STEAM 教育研究文献分析

Analysis of Research Literature on STEAM Education in Primary and Secondary Schools within the Country and Abroad

黄诘雅，李葆萍*
北京师范大学教育学部
* libp@bnu.edu.cn

【摘要】 本研究利用 Citespace 软件，对国内外中小学 STEAM 教育相关的 1455 篇英文论文和 222 篇中文论文进行关键词分析。发现专业发展、科学本质、核心素养等是近年来研究的方向；研究的关注点由课程的实施和教学方式的应用转向学生动机驱动、教师专业培养和公平性问题，研究范围逐步扩展至 STEAM 生态和支持系统；但国内研究更注重技术和工程教育，国外研究更关注基础学科。基于此，本研究从加强基础学科在 STEAM 课程中的比例，激发学生动机，推动 STEAM 教育本土化和教师专业发展三方面提出建议。

【关键词】 中小学 STEAM 教育；CiteSpace；关键词分析

Abstract: *This study utilized CiteSpace software to conduct keyword analysis on 1455 English papers and 222 Chinese papers related to STEAM education in the primary and secondary school both domestically and internationally. The results revealed that professional development, nature of science, core competencies, etc., have been the focus of research in recent years. The focus of research has shifted from the implementation of courses and the application of teaching methods to student's motivation, teacher's professional development and issues of equity, and gradually expanded to the STEAM ecosystem and support systems. However, domestic research emphasizes more on technology and engineering education, while foreign research focuses more on fundamental subjects. Based on this, the present study proposes suggestions: strengthening the proportion of fundamental subjects in STEAM courses, stimulating students' motivation, and advancing the localization of STEAM education and professional development of teachers.*

Keywords: K-12 STEAM education, CiteSpace, Keyword Analysis

1. 背景

当今时代，全球重大前沿技术快速突破，人工智能、大数据、区块链等新兴技术加快应用。科技水平的高低影响着国家的国际竞争力，在此背景下，各国纷纷出台相关政策以支持科技教育的发展。“加拿大 2067”计划（Canada2067）在教师专业发展，教学、课程和评估，STEM 教育公平等六个领域提出愿景和关键问题。澳大利亚发布的《国家 STEM 学校教育战略 2016-2026》指出，采取国家行动来改进学校的科学、数学和信息技术的教学与学习。

STEM 包含四门学科，分别为 Science、Technology、Engineering、Mathematics，这些科目通过有机整合汇聚成一个整体（Khine & Areepattamannil, 2019）。随后“Art”被纳入其中，最后形成 STEAM 这一概念。STE(A)M 教育指向学生跨学科思维、创造力、科学素养等的培养，较传统的分科教学更符合当代的教学情境。技术的不断迭代要求 STEAM 教育内容及时更新，近年来有关 STEAM 教育的教学模式、课程设置、评价方式的研究也不断增多。研究数量的与日俱增和研究内容的不断更新要求总结出 STEAM 教育的历史性变化和未来的发展趋势，以在其迅速变化中找到立足点。基于此，本文通过对国内外中小学 STEAM 教育的研

究分析，旨在回答三个研究问题：1.近十年内 STEAM 教育有哪些研究主题？2.STEAM 教育的研究主题有什么变化趋势？ 3.国内外 STEAM 教育的研究共识和差异有哪些？

2. 研究方法

Web of Science 涵盖不同国家的研究成果，被认为是最适合进行文献计量分析的高质量的数据库（Ding & Yang, 2020）；中国知网在国内是权威的、被广泛应用的数据库。因此本研究将 Web of Science（核心集合）和中国知网作为数据来源，确定的检索式分别为 TS=(("STEAM education" OR "STEM education" OR "science education" OR "education of science and technology") AND ("middle school" OR "junior high school" OR "senior high school" OR "senior middle school" OR "senior school" OR "high school" OR "secondary school" OR "primary school" OR "elementary school" OR "k*12"))和(steam + stem + 科学教育 + 科技教育) * (中小学 + 中学 + 高中 + 初中 + 小学)，时间跨度为 2014 年 1 月至 2023 年 11 月。最终得到国际文献 1455 篇和国内文献 222 篇。CiteSpace 是文献计量学中常用的工具，能清晰地描绘研究发展的进程和趋势，因此本研究用 CiteSpace 对检索到的文献进行关键词分析。

3. 研究结果

3.1. 国际发表有关 STEAM 教育研究现状

3.1.1. STEAM 教育文献的关键词共现图谱分析

国际上前十一个高频关键词（见图 1）由高到低为：“science education”“knowledge”“students”“stem education”“achievement”“inquiry”“teachers”“technology”“attitudes”“instruction”“mathematics”。国际研究比较关注数学和科学学科的教学；常采用探究式的学习方式；学生的学习成就和学习态度被看作是重要的产出成果；学科知识内容的获取也受到一定的关注。

3.1.2. STEAM 教育文献关键词聚类分析

关键词的聚类分析共得到 6 个结果（见图 2），分别为“stem education”“argumentation”“nature of science”“computational thinking”“professional development”“high school/introductory chemistry”。国际研究在论证、科学本质、计算思维方面的关注指向学生对知识的深度理解；重视教师的专业发展；研究对象多分布在高中阶段，并且基础化学学科受到关注。

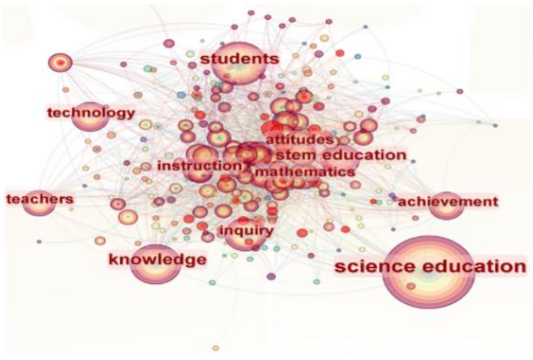


图 1 国际中小学 STEAM 教育关键词共现图谱

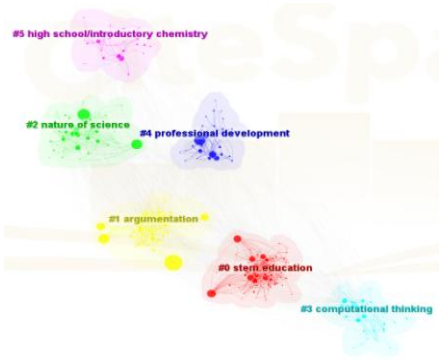


图 2 国际中小学 STEAM 教育关键词聚类

3.1.3. STEAM 教育研究热点时间线图分析

2014 年，“science education”、“knowledge”、“students”、“inquiry”等受到关注。2015-2020 年间，“motivation”、“engagement”、“computational thinking”、“gender”等是研究的热点。2020 年后，“teacher education”、“equity”等受到关注。国际研究在较早时期比较关注科学教育、实施的教学方式等；随后转向如何提高学生的动机、参与，关注学生计算思维的培养和不同性别学生的差异；近年来，开始关注教师教育和公平性问题。

3.2. 国内发表有关 STEAM 教育研究现状

3.2.1. 国内中小学 STEAM 教育文献的关键词共现图谱分析

由图 3 可知，在国内出现的前十个高频关键词分别为：“科学教育”“创客教育”“美国”“中小学”“小学科学”“科学素养”“课程标准”“教学设计”“人工智能”“核心素养”。国内研究较多借鉴美国等发达国家；更多和工程教育、技术教育相关；侧重于培养学生应用新兴技术的能力；在基础教育阶段，国家设计科学等学科的课程标准，来提升学生的科技素养。



图 3 国内中小学 STEAM 教育关键词共现图谱

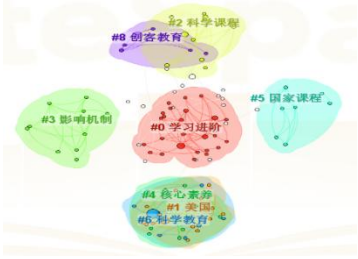


图 4 国内中小学 STEAM 教育关键词聚类

3.2.2. STEAM 教育文献的关键词聚类分析

关键词聚类分析的结果（见图 4）分别为“学习进阶”“美国”“科学课程”“影响机制”“核心素养”“国家课程”“科学教育”。国内文献研究集中在教育教学和课程实施上；课程内容体现学习进阶和核心素养的培养；在深层次上也关注影响学生 STEAM 学习的因素及其影响机制。

3.2.3. STEAM 教育研究热点时间线图分析

2014 年，“创客教育”、“美国”、“小学科学”等受到关注。2018 年起，“人工智能”、“科学素养”、“核心素养”的相关研究增多；2022 年后，开始关注“教师培训”、“专业发展”等内容。国内在较早阶段学习以美国为代表的国家，研究学段以小学为主。随后人工智能融入 STEAM 教育，学生科学素养的培养得到重视。近年来开始关注教师教学水平的提升。

4. 研究结果和讨论

4.1. 研究结果讨论

4.1.1. 中小学 STEAM 教育研究主题分析及异同比较

在学科内容上，国际上将基础学科作为推动社会进步和经济发展的关键，研究关注数学、化学、科学等基础学科；关注学生科学知识获取的过程及科技素养是否得到显著提升。国内强调工匠精神，基础设施的建设较为广泛，需要大量的工程技术人才，研究关注创客教育、工程教育、科学教育等内容；关注学科核心素养、核心概念的学习进阶。在研究的学段中，国际上重视科技高中的建设，研究聚焦于高中；国内研究聚焦于中小学，由于小学科学暂未分科，实施跨学科活动更为容易，所以小学科学研究最热。在技术应用上，研究者们关注技术如何融入教学，例如用 3D 打印模型增加科学现象的可视性，用云平台存储、处理和传播实验数据。近年来，国内外均聚焦于教师的专业发展，期望教师掌握跨学科教学的能力，能熟练应用基于项目的学习、基于问题的学习等教学方法，对学科知识有深入的理解。

表 1 国内外研究的差异

	差异比较		相同之处
	学科领域	重点研究学段	
国际研究	科学、化学和数学等基础学科	高中	技术融入教学；教师专业发展
国内研究	工程、技术教育	小学	

4.1.2. 中小学 STEAM 教育研究主题的变化趋势

近十年来，研究主题由课程实施向驱动学生学习的因素、教师专业培养和公平性问题发展，研究内容深度化，研究范围向整体教育系统扩展。具体来说，早期研究包括学科课程内容、教学方式等；然后研究者们开始关注影响 STEAM 学习的具体因素，如动机、自我效能感等；关注性别对学生的影响，培养强烈的科学认同感对女孩来说尤为重要(Vincent-Ruz & Schunn, 2018)。近年来，研究者们开始关注教师的专业发展和公平性问题，对教师进行专业培养以发展其跨学科教学的能力；需实施相应的教学策略以减少不同背景学生的学习差异。

4.2. 对我国中小学 STEAM 教育的启示

4.2.1. 加强基础学科在 STAEM 课程中的比例，促进课程融合

加强物理、化学、数学等基础学科在 STEAM 课程中的比例，帮助学生理解科学的本质，培养批判性的思维，不只局限于创客、人工智能等技术性的教学内容；成立由不同学科专家组成的跨学科教学团队，构建跨学科课程。当前国内实施 STEAM 教育的形式以课外活动为主，实施的学段集中在小学。应将 STEAM 教育纳入国家正式的课程体系建设中，设置为必修课程，并将涵盖的学段范围扩展至中学。同时学校提供课程实施需要的条件，如设立科学实验室，提供 3D 打印机、激光切割机等设备，聘请专业人士进行指导。

4.2.2. 激发学生的学习兴趣和动机以促进其深层理解

动机和兴趣是学习者和环境交互的产物（Sawyer, 2022），针对不同的学习者，采用不同的方式来加强其主动性。对于最初缺乏兴趣和动机的学习者，通过刺激信息进行吸引，结合任务清单或和同伴、老师的互动来维持其兴趣，帮助其建立和学习内容之间的联系。对本身有兴趣的学习者来说，当其和正在学习的材料、进行的任务建立联系时，兴趣即被激发，动机行为产生。在学生建立和学习内容的联系后，建设探究式或协作式的学习环境，也可以设置一些挑战或适时提供反馈，使学生深度理解所学内容。

4.2.3. 推动 STEAM 教育的本土化和教师专业发展

各国的国情不同，故不能生搬其他国家的教育模式，要注重 STEAM 教育的本土化。国内在工程教育、人工智能方面已有一定的理论基础，可以加强其具体应用。但 STEAM 教育公平的相关研究仍然较少，因此要加强有关公平问题的研究，减少不同地区间的教育差距。需加强对教师创新能力、计算思维等方面的专业培养；各高校加强对 STEAM 相关专业学生的学科教学知识、综合运用教法和教学情境知识能力的培养；优化培养方案，强化教育教学方法实践训练。为在职教师提供线上线下相结合的学习渠道，使其能够灵活选择学习时间。

参考文献

Ding, X., & Yang, Z. (2020). Knowledge map** of platform research: a visual analysis using VOSviewer and CiteSpace. *Electronic Commerce Research*, 1-23.

Khine, M., & Areepattamannil, S. (2019). *Steam education*. Springer, 10, 978-3.

Qhobela, M., & Kolitsoe Moru, E. (2014). EXAMINING SECONDARY SCHOOL PHYSICS TEACHERS' BELIEFS ABOUT TEACHING AND CLASSROOM PRACTICES IN LESOTHO AS A FOUNDATION FOR PROFESSIONAL DEVELOPMENT. *International Journal of Science and Mathematics Education*, 12, 1367-1392.

Sawyer, R. K. (Ed.). (2022). *The Cambridge handbook of the learning sciences*. Cambridge University Press.

Vincent-Ruz, P., & Schunn, C. D. (2018). The nature of science identity and its role as the driver of student choices. *International journal of STEM education*, 5(1), 1-12.

中學教師與 STEM 大屏教學

K12 teachers and STEM IFP teaching

陳冠良，簡佑宏*
臺灣師範大學
*roland.chien@ntnu.edu.tw

【摘要】 世界各國皆積極在 K12 教育階段投入大型互動顯示器（大屏）經費推行 STEM (Science, Technology, Engineering, Mathematics) 教育政策，因此本研究欲探討目前中學教師進行 STEM 大屏教學的現況。本研究採用四大面向 19 題的「重要－表現分析」問卷和 4 題開放式問題，調查 55 位中學教師使用大屏進行 STEM 教學的看法。問卷四個面向包括：(1) 中學教師對於大屏進行 STEM 互動教學的看法；(2) 認為大屏對於學生的影響；(3) 實施 STEM 大屏互動教學的挑戰；(4) 對於 STEM 大屏增能課程的意見。根據 44 份完整回收的問卷分析結果顯示：「相關設備與資源的支持」及「使用大屏教學更有效率」需要優先改進。這份問卷發現許多教師在 STEM 大屏教學的真實需求，有助於未來進一步 STEM 大屏教學的推行。

【關鍵詞】大屏；科學、科技、工程、數學；重要－表現程度分析

Abstract: *Countries worldwide actively invest in Interactive Flat Display (IFP) at K12 education to promote STEM (Science, Technology, Engineering, Mathematics) education. This study investigates the current situation of 55 middle school teachers' STEM IFP teaching using a four-dimension, 19-question "Importance-Performance Analysis" questionnaire and four open-ended questions. The four aspects of the questionnaire include (1) Middle school teachers' opinions on STEM IFP teaching; (2) The impact of STEM IFP teaching on students; (3) Challenges in implementing STEM IFP teaching; (4)Opinions on STEM IFP teaching empowerment courses. According to the analysis results of 44 ultimately returned questionnaires, "support of relevant equipment and resources" and "more efficient teaching using IFP" must be prioritized for improvement. This questionnaire found the real needs of in-service teachers for STEM IFP teaching, which will help further promote STEM IFP teaching in the future.*

Keywords: Interactive flat display, STEM, keyword three, Importance-performance analysis

1. 前言

STEM 教育是近年國際的趨勢，其教育目的為培養學生具備 21 世紀的素養及競爭力，使學生能面對未來環境與全球化變遷之挑戰（Bybee, 2013）。推動 STEM 教育除了師資的培育外，設備及教師資訊能力的提升亦為重要項目。隨著近年科技設備技術的提升，學生與教師的教學環境也逐漸變佳，在 STEM 跨領域教學時，大型互動顯示器（大屏）的便利性及其功能成為現今 STEM 教學常見設備之一。

臺灣自 2021 年起，教育部投入大量經費推動「班班有網路、生生用平板」(行政院，2022)；臺灣各地區亦陸續推動「班班有大屏」，逐漸替換舊的資訊設備。政策的投入顯示臺灣對於教育的重視，期望提供學生與教師更優良的教學環境，然而，其中仍有需要努力之處，例如：教師需要具備使用新科技的能力，願意嘗試新的教學模式等。因此，本研究針對中學教師使用互動大屏進行 STEM 教學之現況進行分析與討論。

近年研究發現（Cladera, 2021; Chien & Chang, 2023; Fan, 2022），重要－表現程度分析（Importance-performance analysis, IPA）是可用於分析教育領域現況的方法，藉此了解工具或教學模式中重要的指標及其實際表現的現況，將有限的資源投入最被重視卻未達預期的部分。

重要－表現程度分析由 Martilla 和 James（1977）提出，針對一項服務或產品列出的屬性進行分析，確認其優勢、劣勢與需要改進的項目。IPA 之分析結果劃分為四個象限，提出改善服務或產品的方向，根據屬性坐落的位置採取對應的行動方式（圖 1）：

（1）第一象限：持續保持，表示重要性與表現程度皆高於平均值。
（2）第二象限：優先改善，表示重要性高於平均值，而表現程度低於平均值。
（3）第三象限：次要改善，表示重要性與表現程度皆低於平均值。
（4）第四象限：過度努力，表示重要性低於平均值，而表現程度高於平均值

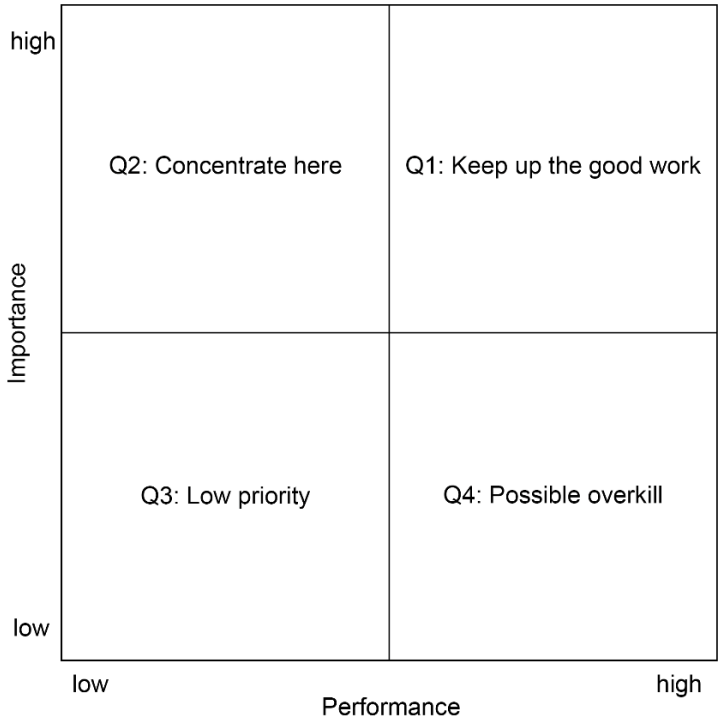


圖 1 IPA 的意涵（Martilla & James, 1977）

基於原本的 IPA 模式，Abalo 等人（2007）提出使用對角線劃分成新的區域（圖 2），改進原本 IPA 模式圖。對角線的意涵表示重要性與表現程度皆相等，對角線上方的任意位置皆表示重要性高、表現程度低，意味著具有改進空間的項目；對角線下方表示重要性低、表現程度高，意味非急迫改善的項目。

綜合上述，IPA 能有系統地評估中學教師對於 STEM 大屏互動教學的看法，而本研究將採取 Abalo 等人（2007）提出的 IPA 模式作為後續研究的依據，使用此方法可以協助政策制定或產品供應商實際上的需求評估。本研究列出以下四個項目為研究問題：

（1）中學教師對於大屏進行 STEM 互動教學的看法？
（2）中學教師認為大屏對於學生的影響？
（3）中學教師認為實施 STEM 大屏互動教學的挑戰？
（4）中學教師對於 STEM 大屏增能課程的意見？

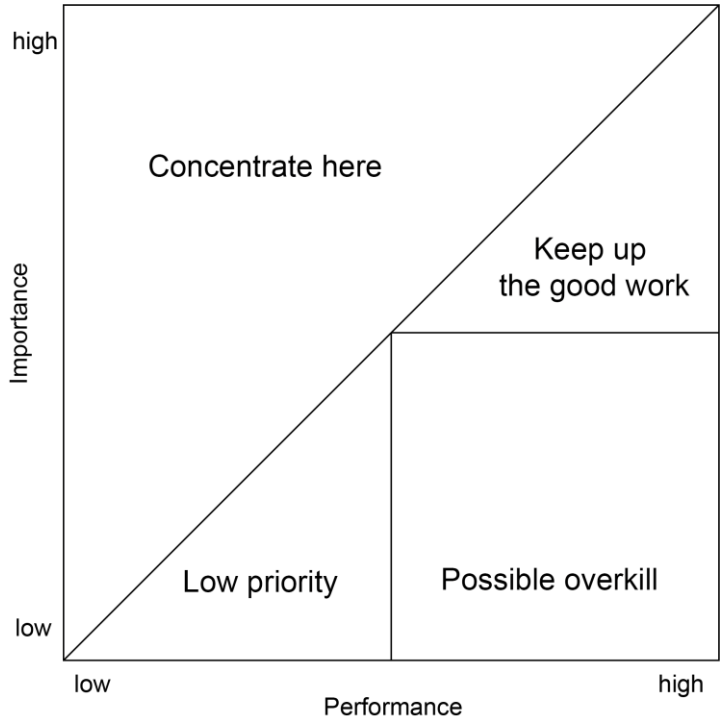


圖 2 Abalo 等人的 IPA 模式（Abalo et al., 2007）

2. 研究方法

本研究採取問卷調查法，以現職教師作為 STEM 大屏教學的使用者為研究對象，探討教師對於 STEM 大屏互動教學的看法，藉由重要－表現程度分析的結果，找出優先改善之處。

2.1 研究對象

採用便利取樣法，針對 55 名參與 STEM 大屏增能課程之現職教師進行問卷發放，教師在課程結束後填寫問卷，提供對於 STEM 互動大屏教學之看法。

2.2 STEM 大屏課程設計

增能課程目的在於讓中學教師對於 STEM 大屏教學有相同的認知，了解如何使用大屏進行 STEM 教學。本研究設計兩個 STEM 大屏課程（圖 3），主要授課對象針對科技領域之中學教師，兩個課程分別為「電路雕塑」與「Simon says」：電路雕塑是一個光控電路結合自然媒材簡易加工製作的小型雕塑品；Simon says 使用電子元件及 Arduino 程式控制，製作出一個能實體互動的玩具。兩個課程皆包含介紹大屏的在課程中如何應用，大屏融入課程的示範，以及動手實作，並在課程結束後發放本研究之問卷。

2.3 研究工具

本研究問卷的編制參考過往研究（Al-Rabaani, 2018; Anatürk Tombak & Ateşkan, 2019; Türel & Johnson, 2012），共編制為四個面向 19 個題向，包括：（1）調查教師對於大屏互動教學的看法（問題 1-5）、（2）互動大屏對於學生的影響（問題 6-11）；（3）實施大屏互動教學的挑戰（問題 12-15）；（4）對於大屏增能課程的意見（16-19）。問卷皆以李克特五點尺度進行評量，並在四個面向的問題後加入開放式問題，以收集教師對於問題的意見。



圖 3 電路雕塑與 Simon says 知識對照圖

3. 研究結果與討論

55 份發出的問卷中，有效樣本數為 44 份。表 1 呈現 19 個 IPA 問題的分析結果。

表 1 IPA 問題分析結果：平均值（標準差）

面向	題項	重要性	表現程度
對於 STEM 大屏互動教學的看法	1. 大屏教學使課程更具互動性	4.43(0.55)	4.00(0.71)
	2. 大屏教學具有使課程內容更視覺化的優勢	4.59(0.54)	4.34(0.68)
	3. 使用大屏進行教學更有效率	4.34(0.68)	3.89(0.84)
	4. 享受使用大屏進行教學	4.20(0.77)	3.91(0.83)
	5. 使用大屏進行課程時間的管理	4.16(0.71)	3.70(0.88)
互動大屏對於學生的影響	6. 使用大屏教學幫助學生學習	4.41(0.58)	4.02(0.66)
	7. 使用大屏教學提升學生的學習興趣	4.64(0.49)	4.11(0.75)
	8. 使用大屏教學提升學生的學習動機	4.55(0.55)	4.05(0.75)
	9. 使用大屏教學學生更容易專注於課程中	4.27(0.69)	3.91(0.77)
	10. 使用大屏教學學生更容易記得課堂所學	4.11(0.64)	3.75(0.75)
	11. 使用大屏教學提升學生的學習成效	4.32(0.62)	3.84(0.68)
實施 STEM 大屏互動教學的挑戰	12. 大屏教學課程需要耗費時間進行事前準備	4.36(0.61)	4.20(0.77)
	13. 學習如何使用大屏進行互動教學	4.59(0.54)	4.23(0.71)
	14. 使用多樣的大屏功能於自身教學之中	4.43(0.70)	4.14(0.80)
	15. 相關設備或資源的支持	4.48(0.59)	4.05(0.83)
STEM 大屏增能課程的意見	16. 提供大屏互動教學增能課程	4.61(0.54)	4.18(0.79)
	17. 推廣大屏互動增能課程	4.57(0.59)	4.14(0.85)
	18. 增能課程內容的適切性	4.64(0.53)	4.14(0.82)
	19. 增能課程的時間長短	4.43(0.70)	3.95(0.83)

3.1 教師對於 STEM 大屏互動教學的看法

中學教師對於 STEM 大屏互動教學的看法的五個問題皆在對角線之上，如圖 4，表示皆有進步的空間。

教師普遍認為大屏教學使課程更具互動性且讓課程內容更視覺化（問題 1、2），其重要性與表現程度皆高於平均值。另外，教師認為「使用大屏教學更有效率」（問題 3）重要性較高、表現程度較低，需要重點改進，根據開放式問題的回答，幾位教師做出了陳述：「此工具能增加效率」、「無使用經驗者需要熟悉者從旁協助」，從中得知使用大屏教學提升效率需要經驗的累積，而目前的表現程度分數表示教師在使用大屏上尚未能獲得效率的提升，可能需要更多的增能課程支援促使使用習慣的改變。

總體而言，教師們對於 STEM 大屏互動教學持有正向看法，並且認為它可以增加課程的互動性和視覺化程度。然而，在提高教學效率方面，仍然存在改進的空間。通過提供更多的培訓和支援，可以幫助教師們更好地運用大屏教學，從而提高教學效果和學生的學習成果。

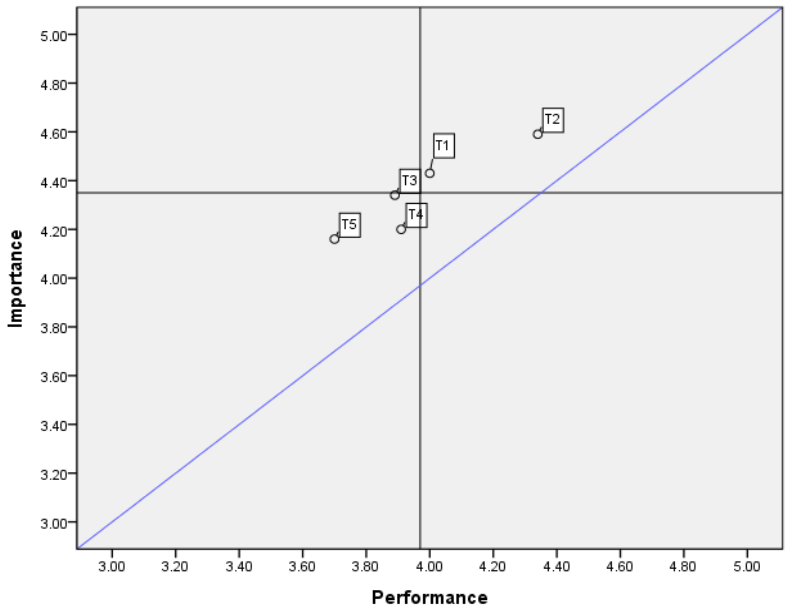


圖 4 教師對於大屏互動教學的看法之 IPA 圖

3.2 中學教師認為大屏對於學生的影響

根據圖 5 所示，教師認為大屏對於學生影響的六個問題皆在對角線之上，表示皆有進步的空間。

教師普遍認為使用大屏幫助學生學習、提升學習興趣與動機（問題 6、7、8），其重要性與表現程度皆高於平均值。然而，對於學生能否更專注於學習、記得課堂所學及提升學習成效（問題 9、10、11），教師認為其重要性及表現程度相對較低。在開放式的問題回答中，教師陳述：「可以增進互動及提升專注力」、「有較好的學習動機，但成效似乎短暫，更容易遺忘」、「學習成效取決於教師本身課程內容，與大屏較無關」

總體而言，教師普遍較支持大屏提升學生學習動機，而學習成效的部分較不會歸因於大屏教學，重要性相對較低，若要得到較高的學習成效，需要從教師本身的課程設計進行調整。

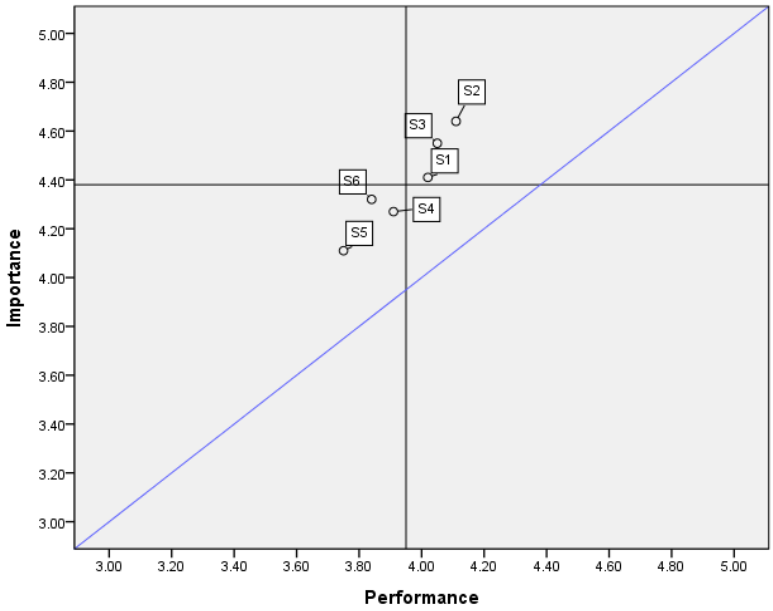


圖 5 教師認為大屏對於學生的影響之 IPA 圖

3.3 中學教師認為實施 STEM 大屏互動教學的挑戰

教師認為實施 STEM 大屏互動教學的挑戰的四個問題皆在對角線之上，表示皆有進步的空間，如圖 6。

教師在學習如何使用大屏進行互動教學上（問題 13），重要性與表現程度皆高，認為學習如何使用是重要的，且持續進行學習。對於其他設備或資源的支持上（問題 15），重要性高但表現程度低，需要重點改進。教師陳述「需要學校的支持與協助」、「教材準備較耗時」、「希望有共備資源」。

總體而言，由於現今政策正在推動大屏的設備與相關課程支持，對於教師而言尚有不足之處，在大屏教材準備上教師仍需耗費較多的時間，若未來能夠提供完整的資源方便教師取用，或改進教材準備上的便利性，將有助於教師使用大屏教學的推動。

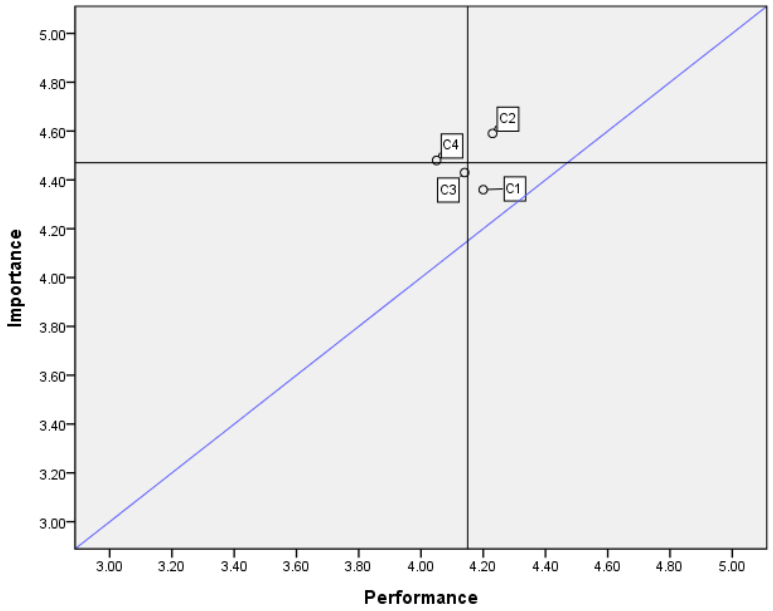


圖 6 教師認為 STEM 大屏互動教學的挑戰之 IPA 圖

3.4 中學教師對於 STEM 大屏增能課程的意見

根據圖 7 所示，中學教師對於 STEM 大屏增能課程的意見的四個問題皆在對角線之上，表示皆有進步的空間。

教師認為提供課程、推廣課程及內容的適切性上（問題 16、17、18）重要性及表現程度皆高於平均值。教師滿意於增能課程，回饋中提到：「增能課程很重要，教師學會才能運用在課堂上」、「各種版本或品牌的大屏軟體不通用的問題，需要更多場研習來增能」。

總體而言，增能課程的時長相對而言為次要項目，重點在於需要持續推廣大屏研習，且要有適當的課程對應現有的設備，才能有效讓教師應用於現場教學中。

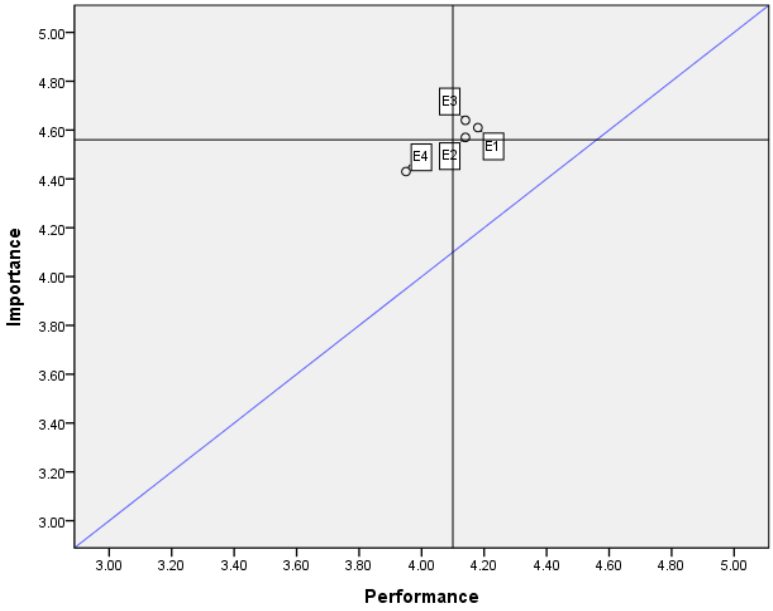


圖 7 教師對 STEM 大屏增能課程意見之 IPA 圖

4. 結語

STEM 教育近年越來越被重視，大屏融入教學可以帶來教學上的改進，其中仍有許多需要努力之處。大部分的中學教師對於大屏教學抱持正面的看法，具有潛力讓教學更有效率、學生學習興趣與動機提升，且教師需要足夠的增能課程、時間與經驗的累積去適應大屏融入 STEM 教學的新模式。從 19 個 IPA 問題的結果顯示，最優先改進的目標是「設備與資源的支持」及「使用大屏教學更有效率」，例如有共備的教材、技術支持，可進一步減少備課時間，促使，若更多教師認同其效益，逐漸熟悉大屏的備課及使用方式，有助於大屏教學的普及。

探討教師作為大屏使用者的觀點，是讓此工具發揮更完善的關鍵，進而讓科技能真正帶來實質上的效益。本研究受限於樣本數量與區域限制，仍有部分的侷限性，同時，研究對象多為自願性參加增能課程的教師，對於 STEM 大屏教學的看法可能過於樂觀，即使如此，仍可提供政策、推廣與發展部分參考價值。未來的研究可更進一步擴大研究區域的範圍及數量，更能完整呈現臺灣大屏教學推動的真實需求。

參考文獻

行政院（2022）。**班班有網路生生用平板－推動中小學數位學習精進方案執行進度**。行政院全球資訊網。

Abalo, J., Varela, J., & Manzano, V. (2007). Importance values for Importance–Performance Analysis: A formula for spreading out values derived from preference rankings. *Journal of Business Research, 60*(2), 115–121.

Al-Rabaani, A. H. (2018). Social studies teachers' perspectives on the advantages and challenges of interactive whiteboard application in Oman. *European Journal of Educational Research, 7*(4), 753–762.

Anatürk Tombak, C., & Ateşkan, A. (2019). Science teachers' beliefs and attitudes towards the use of interactive whiteboards in education. *Journal of Turkish Science Education, 16*(3), 394–414.

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. Arlington, VA: NSTA Press.

Cladera, M. (2021). An application of importance-performance analysis to students' evaluation of teaching. *Educational Assessment, Evaluation and Accountability, 33*(4), 701–715.

Chien, Y. H., & Chang, F. Y. (2023). An importance-performance analysis of teachers' perception of STEM engineering design education. *Humanities and Social Sciences Communications, 10*, 157.

Fan, S. C. (2022). An importance–performance analysis (IPA) of teachers' core competencies for implementing maker education in primary and secondary schools. *International Journal of Technology and Design Education, 32*(2), 943–969.

Martilla, J. A., & James, J. C. (1977). Importance-performance analysis. *Journal of Marketing, 41*(1), 77–79.

Türel, Y. K., & Johnson, T. E. (2012). Teachers' belief and use of interactive whiteboards for teaching and learning. *Journal of Educational Technology & Society, 15*(1), 381–394.

跨校分組的資訊教育課程對於國小學生運算思維能力及學習態度之影響

The Impact of Cross-School Group Information Education Courses on Elementary School Students' Computational Thinking Ability and Learning Attitude

陳芷玄[1]，張博裕[2*]，林秋斌[1]，宋瑗玲[3]，胡振猷[1]
[1]台灣清華大學 學習科學與科技研究所
[2]台灣清華大學 教育與學習科技學系研究所
[3]台灣新竹縣十興國民小學

[*] pajos0524@gmail.com

【摘要】為探討台灣國小學生在資訊教育課程中採用小組合作學習的方式共同進行任務，以跨校的分組方式與同校的分組方式對於學生在運算思維能力以及學習態度是否有所影響。本研究以 64 位分別來自三所國小的五年級學生為研究對象，以 2~3 人為一組的方式隨機分組，實驗組為 S 國小學生加上 1 位 P 國小或是 F 國小的學生，對照組則全數為 S 國小學生組成。研究前針對參與學生進行「運算思維量表」、「學習態度量表」前測，課程結束後，進行「運算思維量表」、「學習態度量表」後測。研究發現：一、以跨校分組的方式學習，學生的運算思維能力有顯著提升；二、以跨校分組的方式學習，學生對於「學習與人共同解決問題，對我的日常生活有幫助」的態度上有顯著差異；三、以跨校分組的方式學習，低成就學生的運算思能力有顯著提升。

【關鍵詞】 運算思維、資訊教育、跨校分組、學習態度

Abstract: *In order to investigate the performance of the elementary students' computational thinking ability and learning attitudes in the information education curriculum, the researcher processed this study by two types of cooperative group learning, one is inter-school students grouping, the other is co-school students grouping. And, this study was also conducted with 64 fifth-grade students from three Elementary Schools in Taiwan. Besides, these 64 participants were randomly divided into groups of two to three students. Overall, the experimental group consisting of students from Elementary School S and one student from Elementary School P or Elementary School F. And, the control group consisted of all Elementary School S students. In this study, students were given the pre-tests of the Computational Thinking Scale and the Learning Attitude Scale before taking courses. After the courses, the post-tests of the Computational Thinking Scale and the Learning Attitude Scale were given. Generally speaking, the study showed the following conclusions. First, there was a significant increase in students' computational thinking ability when learning through cross-school grouping. Second, there was a significant difference in students' attitudes towards "learning to solve problems with others will help me in my daily life" when learning through cross-school grouping. Third, there was a significant increase in low achievers' computational thinking ability when learning through cross-school grouping.*

Keywords: Computational Thinking, Information Education, Cross-school Grouping, Learning Attitude

1. 緒論

隨著人工智慧的蓬勃發展，我們遇到問題時從過去直接使用搜索引擎到現在使用生成式人工智慧成為我們的新選擇。這些轉變使得資訊教育更受到重視，但資訊科技對很多學童而言是相對困難的，對此 Wing（2006）提出的運算思維可以幫助學生奠定資訊科技的基礎（Lu

& Fletcher, 2009）。將運算思維融入生活之中，透過學習該項能力可以培養解決問題的素養。臺灣教育部（2014）所頒佈的「十二年國民基本教育課程綱要總綱」三大核心面向中的「溝通互動」強調善用工具以有效與他人、環境互動，其中也包含善用資訊及科技媒材。在「國民小學科技教育及資訊教育課程發展參考說明」之中，資訊教育課程發展中的「運算思維與問題解決」亦呈現出臺灣對運算思維的重視（國家教育研究院，2020）。

然而世界各國近年都面臨資訊教育師資以及數位科技軟硬體資源分配不均所造成的數位落差問題。臺灣自 2019 年開始推動前瞻建設與生生用平板等多項計畫以弭平數位落差。許多縣市也順應推行教師或學生的跨校合作學習，期望透過跨校合作的學習模式促使教師間共同備課，學生間相互切磋交流以提升教與學的品質，提高學生的學習成效。侯雅齡（2023）的研究也點出偏鄉與都會區學校的跨校合作，確實能讓學生認識不同學校的學生進而激發學習潛能，並產出更多學習成果。

2. 文獻綜述

2.1. 運算思維的定義及架構

運算思維是一種容易理解、強大且具備解釋力的思維模式（Papert, 1996）。Wing 提出運算思維是「運用電腦科學基礎概念來解決問題、系統設計，以及理解人類行為的一系列思維活動。」透過運算思維能明確描述問題、規劃問題解決方案，並以有效率的資訊處理媒介呈現解決問題的思考過程（Wing, 2006 & 2011）。後續更進一步的定義，運算思維是一種分析式的思維（Analytical Thinking），可以幫助我們簡化問題，在需要修正、調整時能更快速及簡便的取用。運算思維並非專業技能，而是每個人每天在每一個地方的都會用到的基本技能（Wing, 2006）。綜上所述，運算思維是一種心智思考過程，運算思維可以提供新的思維方式應用於不同領域中幫助人們解決問題，因此每個人都應該要學習基本的運算思維。

運算思維的最基本元素是抽象，抽象會影響看待問題的關鍵點，因此抽象的正確性是至關重要的。抽象沒有標準答案，因此需要不斷迭代修正。抽象可以是多層次的，各層次之間也會具有一定關聯。此外，Wing（2008）還提出了分解、自動化以及演算法的概念。分解是指將問題由大問題化為數個小問題；自動化是指讓「計算機」能夠協助使用者更有效地解決問題;演算法則是將解決問題的過程「抽象」。Google 在 2015 年提出的「Computational Thinking for Educators」課程中指出運算思維的心智歷程包含問題拆解、樣式辨識、抽象化及演算法設計等內涵。2021 年，Tsai M. J. & Liang J. C. & Hsu C. Y.根據 Selby C. & Woolard J.（2013）進行整理並延伸提出了運算思維的五大主要元素：第一個元素為「抽象」，透過抽象將訊息中的關鍵訊息提取出來。第二個元素為「分解」，透過分解將大的問題拆解成小的問題。第三個元素為「演算法思維」，透過演算法思維將過程依照步驟記錄下來。第四個元素為「評估」，透過評估比較各種解決方案，挑選出最佳的解決方案。最後一個元素為「一般化」，找到共通性將想到的解決方式運用到未來的類似問題上面。使得運算思維應用在教育上的課程發展與評量具備清楚指標。因此，本研究乃採用此五項指標內涵作為運算思維的內涵架構。

2.2. 跨校合作學習

合作學習是一種學習者主動與他人教學相長，且有系統的學習模式。Brown, J. S., Collins, A., & Duguid, P.（1989）及 Johnson & Johnson（2015）提出合作學習之特徵包含能夠協助學生釐清錯誤觀念、產生更多新想法、學習人際溝通技巧、能夠積極互賴，並反思個人角色。合作學習小組的組成可以是教室內熟悉的同儕，亦可是跨班級的合作，甚至是跨校、跨國的合作學習。近期有越來越多學校選擇採用跨校合作學習的方式來增進自身競爭力。以高等教育來看，近年來有不少大學採用了大學策略聯盟，與其他大學進行跨校合作，進行校際的資

源共享及整合。這種大學策略聯盟在各國都有，臺灣更有不少學校跨國與不同國家的學校進行姊妹校的聯盟。李宛儒（2011）的研究發現，小型學校間跨校的合作模式能夠促進教師間的交流使課程更加創新及精緻，資源的共享也使課程更多樣。除此之外，偏鄉小校與都會地區大型學校的跨校合作學習亦可以讓教師資源共享增加課程的豐富度，且學生認識不同學校的同學，能夠彼此互動激發潛能、提高成果產出。這種合作學習方式還可以加入遊戲化的形成性評量促使學生更加信任、協助彼此（侯雅齡，2023）。

林勇志（2003）針對自然科學領域探討了跨校合作學習與非跨校合作學習對於學生學習態度的影響。研究結果指出使用跨校合作學習方式的學生其對科學的態度、電腦態度優於非跨校合作學習方式的學生。李明朗（2011）發表關於融入 ICT 的跨校合作學習相關研究，文中指出此種學習方式能提升學生的 ICT 技能，且在課程中學生會更加專注、提升興趣，進而促使學生積極主動的參與課程。林宜樺（2015）的研究則顯示使用跨校行動合作專題學習活動比一般科學探究學習活動的科學探究素養能力成長較為顯著。除此之外，也發現合作學習的能力對於學生的科學素養能力是有所影響的，合作能力較強的組別能夠更充分的溝通，獲得較高的科學素養能力。

綜上所述，足以窺見跨校的合作學習有其在學習與教學上的正向效果。因此，本研究欲探討以跨校分組合作的學習模式是否較同校分組合作學習，對於學生在資訊課程中的運算思維能力及學習態度有更加正向的效果。

3. 研究方法

3.1. 研究對象及分組

本研究以台灣新竹縣竹北市 S 國小、新竹縣新豐鄉 P 國小、新竹縣竹北市 F 國小六年級學生為研究對象，總共 64 位學生。其中 S 國小共 2 班，A 班男生 13 人，A 班女生 12 人，共 25 人。S 國小 B 班男生 9 人，B 班女生 7 人，共 16 人。P 國小男生 6 人，女生 7 人，共 13 人。F 國小男生 4 人，女生 6 人，共 10 人，如下表 1 所示。分組主要進行三校的隨機分組，共 64 人，分為 31 組，每組 2 至 3 人，進行合作學習。實驗組為跨校分組共有 23 組，總人數為 47 人，對照組為同校分組共有 8 組，總人數為 17 人。

表 1 研究對象人數

性別	S 國小	P 國小	F 國小	合計
男生	22	6	4	32
女生	19	7	6	32
合計	41	13	10	64

3.2. 研究環境

本研究之教學教室分為兩間，一間為電腦教室，每位學生皆配備一台電腦、一片 micro:bit 開發版、一本手冊、一台 iPad，每組學生配備一組 LED 燈條、Q 霸小車、麥昆小車、AI Lens、花車角色座椅、小工具包、AI 辨識圖卡、USB 線、小車燈條以及無線遙控手柄。另一間為自然教室，每 2 至 3 組學生使用一張大桌，每位學生皆配備一台 iPad、一本手冊，每組學生配備一套美術工具、一份輕黏土（10 包）。實驗組與對照組學生分別就相同的課程內容，以前述的教學媒材，運用 micro:bit 進行程式設計，完成「Q 霸小車花車遊行」及「小車足球大賽」兩項學習內容。

3.3. 研究工具

本研究首先透過運算思維量表之前後測，了解研究對象經過該活動後對於其運算思維能力之影響。該運算思維量表參考自 Meng-Jung Tsai, Jyh-Chong Liang, and Chung-Yuan Hsu

(2021)的文獻資料，以不改變其問題本意的方式進行題目的改編使其用語更符合該學齡階段學生的理解程度。此運算思維量表一共分為五大構面，分別為抽象、分解、演算法思維、評估、一般化。該量表採用李克特氏五點量表，選項分別為「非常不同意」、「不同意」、「普通」、「同意」、「非常同意」，依序為 1 到 5 分，共計 19 題。其次透過學習態度量表之前後測，了解研究對象經過分組合作式問題解決學習活動後對於其學習態度的影響。學習態度量表參採周宜慶（2017）的研究量表，分為「興趣與喜好」及「價值感」兩項構面，共有 8 道問題，同樣採用李克特氏五點量表，選項分別為「非常不同意」、「不同意」、「普通」、「同意」、「非常同意」，依序為 1 到 5 分。

4. 結果分析

研究數據結果分為兩大部分，第一部分為運算思維能力，第二部分為學習態度。研究數據主要採用 SPSS 進行數據分析，主要採用的分析方式為成對樣本 t 檢定以及獨立樣本 t 檢定。透過該檢驗方式了解實驗組與對照組學生的運算思維能力、學習態度是否有達顯著之差異。

4.1. 運算思維能力

整體分析後，將資料分成實驗組與對照組，分別進行運算思維前測與運算思維後測的成對樣本 t 檢定，以了解實驗組與對照組兩組的運算思維在經過該活動後是否有顯著提升。其運算思維能力實驗組前測與後測、對照組前測與後測的成對樣本 t 檢定摘要如表 2、表 3 所示。

經過成對樣本 t 檢定計算後得出，實驗組的運算思維前測分數與運算思維後測分數的平均數分別為運算思維前測 =3.68 分，運算思維後測 =3.81 分，兩次考試間的相關性為 0.842。兩次測驗分數差異的檢定$t_{(45)}$值為-2.810，p 值為.007，結果為顯著。觀察平均數可以發現，運算思維後測分數高於運算思維前測分數，因此可以看出實驗組同學的運算思維能力經過活動後有顯著成長。

對照組的運算思維前測分數與運算思維後測分數的平均數分別為運算思維前測 = 3.57 分，運算思維後測 =3.70 分，兩次考試間的相關性為 0.924。兩次測驗分數差異的檢定$t_{(16)}$值為-1.940，p 值為.070，結果為不顯著，因此可以看出對照組同學的運算思維能力經過活動後沒有顯著成長。

此一結果呈現出實驗組學生的運算思維能力有顯著的成長，代表進行跨校分組的活動對於學生的運算思維能力有所幫助。

表 2 運算思維量表實驗組成對樣本 t 檢定分析摘要表

	M	*SD*	$t_{(45)}$	*p*	95% CI		Cohen's *d*
					LL	*UL*	
前測	3.68	0.57	-2.810	.007	-0.23	-0.04	0.22
後測	3.81	0.59					

表 3 運算思維量表對照組成對樣本 t 檢定分析摘要表

	M	*SD*	$t_{(16)}$	*p*	95% CI		Cohen's *d*
					LL	*UL*	
前測	3.57	0.67	-1.940	.070	-0.28	0.01	0.18
後測	3.70	0.74					

運算思維的高、低成就差異主要依據學生的運算思維前測成績進行區分，採用的是二分法，將實驗組共 47 位學生的成績從中間斷開，其中高成就組學生人數則為 24 人，低成就組學生為 23 人（低成就組有效樣本 22 人）。

由於以運算思維的前測資料作為分組依據，因此此處未做運算思維前測的低成就組及高成就組學生之獨立樣本 t 檢定。而是直接對低成就組與高成就組進行運算思維後測的獨立樣本t檢定,以此了解低成就組與高成就組學生在學習活動後其運算思維能力是否有顯著差異。

經過獨立樣本 t 檢定計算後得出，實驗組低成就組的平均數為 3.41 分，實驗組高成就組的平均數為 4.18 分。經由 Levene 變異數同質性檢定之未顯著結果（$F_{(1,44)}$ = 1.312，p = .258），可知兩樣本沒有明顯離散情況差異。變異數同質性檢定結果顯示其變異數相同，故得其檢定結果的 t 值與 p 值皆達顯著。由此可知實驗組低成就組與實驗組高成就組經過活動後，其運算思維能力有顯著差異。（$t_{(44)}$= -5.830，p = .000）實驗組低成就組與實驗組高成就組的運算思維後測獨立樣本 t 檢定統計摘要如表 4 所示。

表 4 運算思維量表高、低成就組後測獨立樣本 t 檢定分析摘要表

	後測		$t_{(44)}$	p	95% CI		Cohen's d
	M	*SD*			LL	UL	
低成就	3.41	0.38	-5.830	0.000	-1.03	-0.50	1.73
高成就	4.18	0.50					

針對低成就組與高成就組分別做運算思維前測與運算思維後測的成對樣本 t 檢定，以了解低成就組與高成就組兩組學生在經過該活動後其運算思維能力是否有顯著提升。其運算思維能力低成就組前測與後測、高成就組前測與後測的成對樣本 t 檢定摘要如表 5、表 6 所示。

經過成對樣本 t 檢定計算後得出，實驗組低成就組的運算思維前測分數與運算思維後測分數的平均數分別為運算思維前測= 3.19 分，運算思維後測= 3.41 分，兩次考試間的相關性為 0.459。兩次測驗分數差異的檢定$t_{(21)}$值為-3.031，p 值為.006，結果為顯著。觀察平均數可以發現，運算思維後測分數高於運算思維後測分數，因此可以看出實驗組低成就組同學的運算思維能力經過活動後有顯著成長。

實驗組高成就組的運算思維前測分數與運算思維後測分數的平均數分別為運算思維前測= 4.12 分，運算思維後測= 4.18 分，兩次考試間的相關性為 0.803。兩次測驗分數差異的檢定$t_{(23)}$值為-0.938，p 值為.358，結果為不顯著，因此可以看出實驗組高成就組同學的運算思維能力經過活動後沒有顯著成長。

表 5 運算思維量表低成就學生成對樣本 t 檢定分析摘要表

	M	*SD*	$t_{(21)}$	p	95% CI		Cohen's d
					LL	*UL*	
前測	3.19	0.24	-3.031	.006	-0.37	-0.07	0.69
後測	3.41	0.38					

表 6 運算思維量表高成就學生成對樣本 t 檢定分析摘要表

	M	*SD*	$t_{(23)}$	p	95% CI		Cohen's d
					LL	*UL*	
前測	4.12	0.40	-0.938	.358	-0.18	0.69	0.13
後測	4.18	0.50					

4.2. 學習態度

以下將學習態度量表以逐題的方式進行學習態度量表前測與後測的成對樣本 t 檢定，藉此了解採用跨校分組合作學習的實驗組學生在經過學習活動後其學習態度是否有顯著的變化，其結果如下表 7 所示。

經過成對樣本 t 檢定計算後得出，學生在第 1 題的前測分數與後測分數分別為前測 =4.02 分，後測 =4.00 分，兩次測試分數的相關性為 0.729。兩次測驗分數差異的檢定$t_{(45)}$值為 0.256，p 值為.799，結果為不顯著。表示學生在「喜歡與人共同解決問題」的態度無顯著區別，因此可以看出學生的第 1 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 2 題的前測分數與後測分數分別為前測 =3.96 分，後測 =3.87 分，兩次測試分數的相關性為 0.634。兩次測驗分數差異的檢定$t_{(45)}$值為 0.850，p 值為.400，結果為不顯著。表示學生在「期待與人共同解決問題」的態度無顯著區別，因此可以看出學生的第 2 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 3 題的前測分數與後測分數分別為前測 =3.80 分，後測 =3.96 分，兩次測試分數的相關性為 0.740。兩次測驗分數差異的檢定$t_{(45)}$值為 -1.734，p 值為.090，結果為不顯著。表示學生在「與人共同解決問題會感到快樂」的態度無顯著區別，因此可以看出學生的第 3 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 4 題的前測分數與後測分數分別為前測 =4.13 分，後測 =3.96 分，兩次測試分數的相關性為 0.519。兩次測驗分數差異的檢定$t_{(45)}$值為 1.534，p 值為.132，結果為不顯著。表示學生在「與人合作會容易解決問題」的態度無顯著區別，因此可以看出學生的第 4 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 5 題的前測分數與後測分數分別為前測 =4.09 分，後測 =4.11 分，兩次測試分數的相關性為 0.483。兩次測驗分數差異的檢定$t_{(45)}$值為-0.190，p 值為.850，結果為不顯著。表示學生在「與人合作可以發現多種解決方式」的態度無顯著區別，因此可以看出學生的第 5 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 6 題的前測分數與後測分數分別為前測 =3.96 分，後測 =4.02 分，兩次測試分數的相關性為 0.497。兩次測驗分數差異的檢定$t_{(45)}$值為-0.553，p 值為.583，結果為不顯著。表示學生在「學習與人共同解決問題對課業有幫助」的態度無顯著區別，因此可以看出學生的第 6 題學習態度經過活動後無顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 7 題的前測分數與後測分數分別為前測 =3.70 分，後測 =3.89 分，兩次測試分數的相關性為 0.725。兩次測驗分數差異的檢定$t_{(45)}$值為-2.144，p 值為.037，結果為顯著。觀察平均數可以發現，學習態度第 7 題後測分數高於學習態度第 7 題前測分數，因此可以看出學生在「學習與人共同解決問題對生活有幫助」的態度經過活動後有顯著成長。

經過成對樣本 t 檢定計算後得出，學生在第 8 題的前測分數與後測分數分別為前測 =3.89 分，後測 =3.93 分，兩次測試分數的相關性為 0.702。兩次測驗分數差異的檢定$t_{(45)}$值為-0.467，p 值為.642，結果為不顯著。表示學生在「學習與人共同解決問題對與朋友相處有幫助」的態度無顯著區別，因此可以看出學生的第 8 題學習態度經過活動後無顯著成長。

表 7 學習態度量表整體逐題成對樣本 t 檢定分析摘要表

題目	前測		後測		$t_{(45)}$	p
	M	*SD*	*M*	*SD*		
我喜歡跟同組組員一起解決遇到的問題	4.02	0.80	4.00	0.76	0.26	.799
我期待能跟同組組員一起解決問題	3.96	0.82	3.87	0.81	0.85	.400
跟同組組員一起解決問題讓我覺得很開心	3.80	0.83	3.96	0.82	-1.73	.090

跟同組組員一起解決問題會更簡單	4.13	0.75	3.96	0.82	1.53	.132
跟同組組員一起解決問題可以發現很多不一樣的解決方法	4.09	0.78	4.11	0.74	-0.19	.850
跟同組組員一起解決問題，對我的學習會有幫助	3.96	0.84	4.02	0.75	-0.55	.583
跟同組組員一起解決問題，對我平常的生活有幫助	3.70	0.87	3.89	0.80	-2.14	.037
跟同組組員一起解決問題，對我跟朋友相處有幫助	3.89	0.88	3.93	0.71	-0.47	.642

5. 結論與討論

在分組合作學習的模式下，使用跨校分組方式進行資訊教育課程的學生，有別於同校分組的學生，在「運算思維量表」的後測表現顯著高於前測表現。即針對資訊教育課程，採用跨校分組的方式比起同校分組方式學生在運算思維能力上的提升更加顯著。

進一步比較跨校分組中低成就組與高成就組學生在運算思維上的進步差異，低成就組學生，在「運算思維量表」的後測表現顯著高於前測表現。即針對資訊教育課程，採用跨校分組的的方式對於低成就學生在運算思維能力上的提升更有幫助。

而在態度量表上，跨校分組合作學習的學生在「學習與人共同解決問題，對我的日常生活有幫助」的後測態度顯著高於前測，即學生認為跨校分組合作解決問題對於日常生活之影響是有益處的。

每位學生的特質均不盡相同，對於跨校分組的想法也有不同的意見，組員的性別是否相同也可能是影響其學習表現的因素。進一步對於學生進行訪談並與其學習態度量表分數比較可以得知：學生意願若傾向於同校分組，卻被分配至跨校分組對於其學習態度是下降的；而學生意願若傾向於跨校分組且同樣採用跨校分組進行學習，則其學習態度普遍有所上升。因此建議教師應能先了解學生對於分組的傾向，再來決定分組的方式，對於提升學生在學習態度上會更有助益。此外，本研究所實施的運算思維課程總時數合計 350 分鐘（八節課），但是整個研究的脈絡僅使用了兩日的時間，若是能將整體課程的教學時間分散至兩週甚至一個月，或許對於學生的整體影響會更深刻，而有不同的面貌。

參考文獻

李明郎（2011）。**運用 ICT 進行 iEARN 跨校合作學習之行動研究：以高雄縣國小為例**（未出版之碩士論文）。國立高雄師範大學，高雄市。

李宛儒（2011）。**小校轉型優質計畫之研究–以跨校協同教學為例**（未出版之碩士論文）。國立嘉義大學，嘉義市。

林勇志（2003）。**運用網路於國小五年級自然科學習之研究-跨校與非跨校之比較**（未出版之碩士論文）。國立花蓮師範學院，花蓮縣。

林宜樺（2015）。**行動專題合作學習能力動態評量的發展與應用**（未出版之博士論文）。國立臺南大學，臺南市。

周宜慶（2017）。**國小高年級學生合作式問題解決能力與態度之研究**（未出版之碩士論文）。國立臺南大學，臺南市。

侯雅齡（2023）。跨校混成教學於資賦優異學生獨立研究課程之實踐。**特殊教育研究學刊，48**(2)，77-110。

教育部（2014）。**十二年國民基本教育課程綱要總綱**。https://edu.law.moe.gov.tw/LawContent.aspx?id=GL002057

國家教育研究院（2020）。**國民小學科技教育及資訊教育課程發展參考說明**。https://www.naer.edu.tw/upload/1/9/doc/280/1090611_%E5%9C%8B%E6%B0%91%E5%B0%8F%E5%AD%B8%E7%A7%91%E6%8A%80%E6%95%99%E8%82%B2%E5%8F%8A%E8%B3%87%E8%A8%8A%E6%95%99%E8%82%B2%E8%AA%B2%E7%A8%8B%E7%99%BC%E5%B1%95%E5%8F%83%E8%80%83%E8%AA%AA%E6%98%8E.pdf

Brown, J. S., Collins, A., & Duguid, P.（1989）Situated cognition and the culture of learning. *Educational Research, 18*, 32-42.

David W. Johnson and Roger T. Johnson. (2015). Learning Together and Alone. *Better: Evidence Based Education, 7,* 4-5.

Google. (2015). *Computational thinking for educators*. Retrieved from https://edu.google.com/resources/programs/exploring-computational-thinking

Lu, J. J. & Fletcher, G. H. L. (2009). Thinking about Computational Thinking. *SIGCSE Bull*, *41*(1), 260-264.

Meng-Jung Tsai, Jyh-Chong Liang, and Chung-Yuan Hsu (2021). The Computational Thinking Scale for Computer Literacy Education. *Journal of Educational Computing Research, 59(4)*, 579-602.

Papert, S. (1996). An Exploration in the Space of Mathematics Educations. *International Journal of Computers for Mathematical Learning, 1*, 95-123.

Selby,C.,&Woollard, J. (2013).Computational thinking: The developing definition. *Special Interest Group on Computer Science Education*, Canterbury.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.

Wing, J. M. (2008). Computational thinking and thinking about computing. Philosophical Transactions of the Royal Society of London A: Mathematical, *Physical and Engineering Sciences, 366*(1881), 3717-3725.

Wing, J. M. (2011). Research notebook: Computational thinking - What and why. *The Link Magazine*, 6, 20-23.

PBL 中有效失败策略对小学生创造力影响研究

Research on the Impact of Productive Failure Strategies in PBL on the Creativity of Primary School Students

沈鑫明[1]，张新立[1*]
[1]温州大学 教育学院
[*]zhxl@wzu.edu.cn

【摘要】 在近年的教育实践中，项目化学习已在中小学广泛推广。尽管如此，其应用过程中却普遍存在着偏重学习成果而忽略学习过程探索性和创新性的问题。鉴于此，本研究提倡有效失败策略，以促进学生解决问题的能力和创新思维的发展。基于此理念，研究提出了“有效失败策略-项目化学习”模式，并据此设计了《智能浇水装置》项目活动案例。通过对小学六年级学生群体进行实验教学，研究发现，相较于传统的项目化学习方法，“有效失败策略-项目化学习”模式更能有效提高学生的创造力水平。

【关键词】 项目化学习；有效失败策略；创造力

Abstract: *In recent educational practices, project-based learning has been widely adopted in elementary and secondary schools. Despite its popularity, there's a common issue where the emphasis is placed on learning outcomes, overlooking the exploratory and innovative nature of the learning process. In light of this, our study advocates for the productive failure strategy to enhance students' problem-solving abilities and the development of innovative thinking. Based on this concept, the study introduces a "Productive Failure Strategy - Project-Based Learning" model and designs a case study for the project activity "Intelligent Watering Device." Through experimental teaching with a group of sixth-grade students, the study found that compared to traditional project-based learning methods, the "Productive Failure Strategy - Project-Based Learning" model is more effective in improving students' creativity levels.*

Keywords: Project based learning，Productive Failure Strategies，creativity

1. 前言

在国际比较的背景下，中国学生通常在基础知识和技能方面表现出扎实的能力，然而，在创造力及问题解决能力方面，则显现出相对的不足，有时甚至存在显著的劣势（臧莺，2012）。为了着手改善这一状况，我国在 2016 年与 2017 年相继发布的《关于深化教育体制机制改革的意见》中，特别强调了推进教育创新以及激发学生的好奇心、想象力和创造性思维的紧迫性。创造力被定义为一种智力品质或心理过程，它涉及综合利用所有可获取的信息，以产生既有社会意义（或个人价值）又具有新颖性和独特性的成果（林崇德，2018）。

项目化学习作为一种国际教育界推崇的创新教学模式，为学生创造力的培养提供了新的途径与实践机会。在设定教学目标方面，巴克教育研究所强调，项目化学习的核心目标之一便是培育学生的创造性素养（夏雪梅，2018）。众多学者投身于探讨项目化学习与创造力之间的联系，进行了广泛的理论与实证研究。例如，Hung（2018）对项目化学习（PBL）在培养创造力方面的多维作用进行了探究，涉及 PBL 在激发学生内在动机、问题的开放性与创新性、深层次的社会互动以及通过阶段性任务积累创造力等方面的作用。Storer（2018）则实施了一段为期六周的项目化学习课程干预，并观察到该教学法对学生创造力的显著正向影响。

近年来，我国中小学教育体系日益重视项目化学习。多个地区的教育局，包括上海和温州等城市，已经推出了“中小学项目化学习三年行动计划”。这些计划以培养创造性的问题解决能力为宗旨，将项目化学习的实践和研究作为核心着力点，并借助活动项目、学科项目及跨学科项目等形式，积极探寻项目化学习的新途径。尽管如此，研究表明在实际的教学过程中，教师与学生普遍存在一种倾向：即重视完成最终成果而忽视学习过程的探索性和创新性。为了降低学生遭遇失败的可能性，教师在设计项目任务时往往会提供以成功为导向的任务清单和搭建脚手架（姜和丽，2020），并明确问题的解决路径。有时，为了节约时间，教师可能简化说明，导致本应是探究的过程演变为单纯的模仿。学生在项目化学习中需要均衡理论学习与实践操作，但却常常偏向于实践操作，使得教学变成了单纯“为了做而做”的行为。教师过分注重成果而非过程，学生则偏重实践轻理论（曹培杰，2020），这些现象使得项目化学习的形式大于内容。学生在完成任务的过程中缺乏尝试、经历失败、进行合作探究和反思的机会（安富海，2022），这限制了他们的创新思维的发展。鉴于此，引入有效失败策略至项目化学习中，能够针对上述问题提供解决方案，从而为学生创造力的培养开辟新径。

因此，本研究旨在将有效失败策略与项目化学习相融合，构建了“有效失败-项目化学习”实施模式。基于此模式，研究设计了《智能浇水装置》项目活动案例，并将该案例应用于小学六年级学生群体，以此考察在项目化学习中融入有效失败策略对学生创造力水平所产生的影响。通过探索有效失败策略在项目化学习中的应用情况，本研究希望为教育实践提供新的理论框架，并旨在为学生创造力培养提出创新性的策略。

2. 理论基础

2.1.有效失败策略

有效失败（Productive Failure）是由 Manu Kapur 在 2008 年提出并持续发展的理论概念，它涉及一种特定的学习现象或机遇，即学习者在学习过程中遭遇失败的同时，实际上正在经历一个学习的过程（Kapur, 2008）。依据 Kapur 的理念，有效失败策略涉及向协作小组中的学生们呈现复杂的问题。学生们将基于他们已有的知识或生活经验，提出多种非最优或失败的解决方案。尽管学生们未能生成出标准解决方案，但当教师最终展示标准方案并进行对比时，学生们的注意力会集中在目标知识上，使他们认识到自身的理解缺陷，进而加深对目标知识的理解。

有效失败策略的核心原则在于延迟对目标知识系统性讲解的时机，允许教师基于学生提出的解决方案进行指导，同时为学生提供充分的时间和空间来激发他们的先验知识，并鼓励他们进行自主探索和协作学习。即便学生在面对挑战或困难时产生的解决方案可能是错误的或次优的，这一探索过程有助于激活他们的先验知识，引发认知冲突，从而有益于后续的目标概念学习。在随后的系统性讲解阶段，教师向学生提供与目标知识相关联的正确解决方案，并进行详尽解释，使学生能够在比较不同方案和通过教师的系统讲解过程中，对目标概念获得更深入的理解。

2.2.有效失败策略对创造力的培养

在传统的教育环境中，“失败”通常被赋予否定性的含义，甚至被视为一个令人畏惧的概念。在大多数教学情境中，“失败”意味着结果未能达到学生或教师所期望的目标（Henriksen et al，2021）。因此，在教育领域中，提及“失败”往往带有负面的标签，且很少被认为是创造力发展或学习机会的关键要素。

然而，教育研究中也未能普遍考虑到创造力与失败之间的联系（Harris &de Bruin，2018）。尽管我们知道，与失败互动对于成功是必要的，而不愿意在学习中包容失败可能会削弱学生的创造力能力（Smith，2020）。已有一些研究表明，将失败肯定为教学原则对于培养学生创

造力至关重要。失败并不一定是破坏性的，而通常可以导致富有创新力的创新（Manalo &Kapur，2018）。拥抱失败对创造过程至关重要，无论是通过从中迭代以最终取得成功，还是通过反思地认识和接受不确定性以及在推断可能结果时固有的模糊性中接受挑战（Swanson &Collins，2018）。

然而，在教育研究领域中，对于创造力与失败之间联系的普遍联系尚显不足（Harris & de Bruin, 2018）。尽管有研究表明，与失败的互动对于实现成功是必不可少的，而且不愿意在学习过程中接受失败可能会对学生的创造力造成削弱（Smith, 2020）。已有一些研究指出，将失败纳入教学原则对于培育学生的创造力是至关重要的。失败并非总是具有破坏性，它通常可以成为激发创新力和创意的催化剂（Manalo & Kapur, 2018）。接纳失败对于创造过程是至关重要的，它不仅促进了通过迭代尝试和调整来最终实现成功的途径，而且鼓励了深入的反思、对不确定性的认识以及在面对潜在的不可预知结果时的挑战应对策略（Swanson & Collins, 2018）。

在项目化学习的过程中，有效失败的概念可激发学生勇敢尝试新的方法来解决问题，即便这可能导致失败。在此框架下，失败不被视作终结点，而是被视为通往成功必经的一步。引入这一理论的意图在于促进学生的韧性、自我调节能力以及创新思维，进而加强创造力的发展。

3. 融入有效失败策略的项目化学习流程

本研究通过结合项目化学习与有效失败策略，并参照夏雪梅等学者提出的项目化学习模式的六个基本步骤，同时整合了 Kapur 的有效失败教学策略，构建了一个融合有效失败策略的项目化学习模型（参见图 1）。该模型突破了传统项目化学习流程中的“知识构建”和“项目方案设计”两个环节，并对“生成-探索”与“知识-整合”两个阶段进行了重新设计。具体的实施流程展示于图 2。

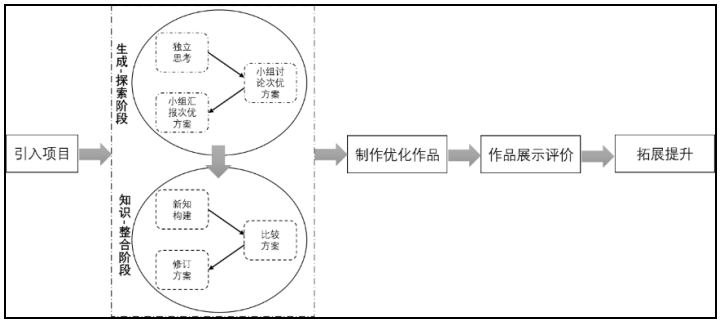


图 1“有效失败策略-项目化学习”实施模式

（1）项目引入阶段标志着项目化学习的启动。在此阶段，教师应依据学生的具体状况挑选适宜的项目，深入考虑项目的实施价值，并构建仿真情境。此过程的目的是指导学生明确学习目标，并识别和理解驱动性问题。项目的选题应以学生为中心，贴近学生的生活实际，确保所学知识与技能能够应用于解决现实问题。同时，项目选题的难度需设定为适度挑战性，以便学生能将新项目与他们现有的知识相连接，从而充分利用其已有的知识经验。在确定了项目目标后，教师和学生共同讨论项目任务的限制条件和评价标准，以协助学生明确学习目标。

（2）在生成-探索阶段，学生将经历独立思考、小组讨论次优方案及其汇报的一系列活动。所谓次优方案，指的是学生通过调动现有知识经验所构建的含有错误或不准确的解决策略。①首先，在独立思考环节，应确保学生有充足的时间来激活和应用他们已有的知识和技能，以形成对项目问题的初步理解，并草拟出初始的项目方案设计。教师在此阶段的角色是鼓励学生大胆试验，并将他们的既有知识与问题结合进行思考。对于表现出消极情绪、不愿

积极探索或遇到特定困难的学生，教师需提供个别指导并给予适当的情感支持。②随后，进入小组次优方案讨论阶段，小组成员将各自的方案集中起来，并通过集体讨论和分析来形成一个更优化的解决方案。由于在这一环节中教师尚未对新知识进行系统性讲解，即便学生通过小组合作交流进行了探索，也可能产生现有知识与目标知识之间的认知冲突。此时，教师应当及时引导学生重新审视项目问题，对学生遭遇的挫折进行正确归因，并在小组探究过程中激励学生积极参与合作。③最终，在此阶段末尾，教师组织各小组分享和汇报他们的项目计划及方案设计。为确保所有学生的参与，要求小组的其他成员对他们分享的解决方案进行补充，同时让其他小组对比自己的解决方案，从而分析各自的优势和不足，推动全班范围的头脑风暴。通过这种小组间的分享和汇报，教师能够更准确地把握学生的认知水平和既有知识经验。

（3）在知识-整合阶段，学生将经历新知构建、比较方案和修订方案的一系列活动。继生成-探索阶段的小组次优方案汇报之后，教师将以更针对性的方式对项目涉及的新知识和技能进行讲解。此后，学生将他们自己设计的次优方案与教师提供的参考方案作对比，通过这一过程反思导致失败的原因，并最终修订并完善他们的项目作品设计方案。①首先，在新知构建环节，基于生成-探索阶段的小组汇报结果，教师会更具针对性地开展活动探究或讲授相关新知识，并提供关于项目任务的标准参考方案。尽管这一解决方案不一定是唯一的，它应当是一个相对最优的方案。②接着，在比较方案环节，学生需要有足够的时间来深入比较自己的设计与教师提供的参考方案之间的差异。教师应引导学生进行细致分析，识别自己设计中的错误认知，并对学生的创新想法和设计予以鼓励和肯定，以维持学生的学习兴趣和提升其解决问题的创造力。③最后，在修订方案环节，学生被要求进行深刻反思，融合新旧知识，并运用新获得的知识对之前的次优方案进行修改，从而完善和改进小组合作产出的设计方案。

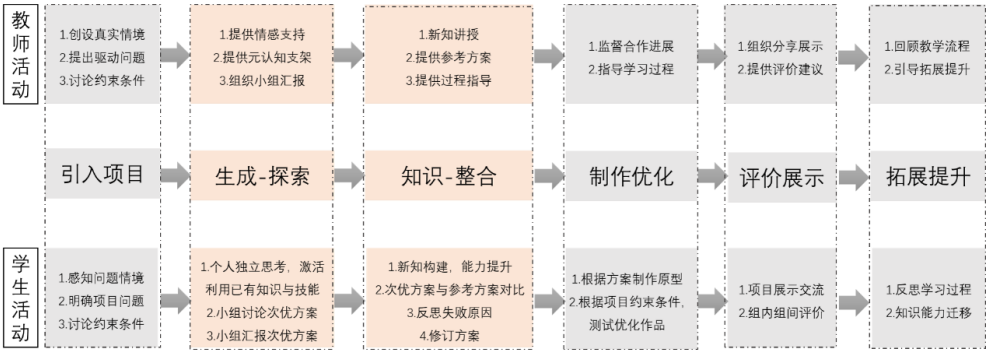


图 2“有效失败策略-项目化学习”实施流程图

（4）在制作优化阶段，学生将依照之前步骤中完成的方案设计，制作项目作品的原型。在作品制作完成之后，他们将根据项目引入阶段由师生共同制定的约束条件，对作品进行测试和优化，以最终完成作品成果。在此阶段，教师的角色是指导学生的学习过程、监督各小组的合作进展，并特别注意防范小组合作中的“搭便车”现象。

（5）展示评价阶段构成了学生反馈交流和调整深化的关键过程。在此阶段，评价应采取多元化的方式，并针对学习目标进行，包括组内、组间评价以及教师评价等各种对学习实践的评估形式。这样的评价机制有利于学生扩展其知识范畴，增强理解力与自我评估的能力。

（6）在拓展提升阶段，教师需对教学流程进行回顾，并指导学生进行针对学习目标和技能的反思。此外，教师还应在知识、应用及项目的不同维度上进行扩展与延伸活动，以促进学生对学习内容的深入掌握。

4. 实验设计

为考察信息技术学科中不同项目化学习策略对小学生创造力影响的差异性，本研究实施了前后测量的准实验设计。在该设计中，自变量是项目化学习策略，实验组采用有效失败策略的项目化学习，对照组采用常规的项目化学习，而因变量则是学生的创造力水平。

4.1.研究对象

本研究选取了中国东南地区某小学六年级的两个班级，共计 80 名学生（男性 37 名，女性 43 名）作为项目实施的参与者。其中，实验组包括 17 名男性和 23 名女性，而对照组则由 20 名男性和 20 名女性组成，每组各有 40 人，平均年龄为 12 岁。在确认实验组与对照组学生在性别、年龄、一般智力及创造力水平等属性上无显著性差异的基础上，由同一位信息技术教师根据既定方案进行教学。

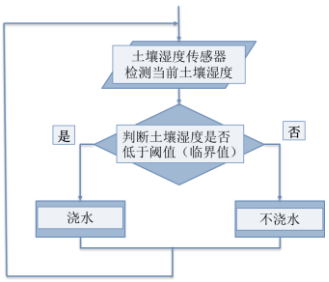
4.2.测量工具

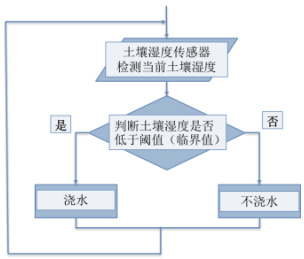
《威廉斯创造力倾向量表》是国际上公认的评估 9 至 17 岁学生群体创造力的可靠工具。该量表以其经过验证的信度与效度，以及在学术界的高频率引用而得到广泛认可。量表由 50 个条目构成，将创造力划分为冒险性、想象力、好奇心和挑战性四个维度，以李克特量表格式进行评分，每个条目都使用 3 点计分法，并结合了正向与反向的题目设计。在当前的实证研究中，为了适配特定的教学情境，本研究对《威廉斯创造力倾向量表》进行了必要的修改，并将其作为主要的测量工具。在本实验中，通过对该量表进行信度检验，Cronbach's Alpha 系数达到了 0.903，显示出该量表具备极高的内部一致性信度。

4.3.活动设计——以《智能浇水装置》为例

在项目主题的选择上，《智能浇水装置》不仅承载了实际的应用价值和跨学科的特性，而且能够促进学生在设计、制作及测试智能浇水装置的过程中增强动手能力和创新能力。此外，该主题与小学生的认知水平和兴趣相匹配，有助于激发他们的创造力和想象力，同时也促进了团队合作精神和沟通能力的培养。实验组（采用有效失败策略的项目化学习）与对照组（采用常规项目化学习）的教学内容详见表 1。

表 1 实验组和对照组的项目活动设计

<table>
<tr><th>实施阶段</th><th colspan="2">《智能浇水装置》项目活动</th></tr>
<tr><th></th><th>实验组（有效失败策略-项目化学习）</th><th>对照组（常规项目化学习）</th></tr>
<tr><td>一
项目
引入</td><td colspan="2">老师的办公室或者教室里的盆栽需要定期地给它浇水，但是当我们都不在学校时不能给盆栽浇水该怎么呢？
提出驱动性问题：如何制作一个智能浇水装置解决自动浇水的问题？</td></tr>
<tr><td>二
生
成
-
探
索</td><td>活动一：认识 Arduino 和传感器
学习控制器、土壤湿度传感器的相关知识，并进行小组实践操作。
活动二：个人智能浇水装置方案初步设计
个人独立思考，利用原有的知识经验进行智能浇水装置方案的初步设计。
活动三：小组智能浇水方案初步设计
例如： 
活动四：小组汇报智能浇水方案设计</td><td>活动一：认识 Arduino 和传感器
学习控制器、土壤湿度传感器的相关知识，并进行小组实践操作。
活动二：学习自动浇水功能的控制过程
 
</td></tr>
</table>

<table>
<tr>
<td>三
知
识
-
整
合</td>
<td>活动一：教师提供智能浇水装置的参考方案，学习探究自动浇水功能的控制过程
 

活动二：小组将智能浇水装置方案与参考方案进行对比，反思方案错误的原因并进行修改。</td>
<td>活动一：设计智能浇水装置，绘制设计图
以小组为单位进行头脑风暴，进行方案设计并绘制草图。
草 图 设 计：</td>
</tr>
<tr>
<td>四
制
作
优
化</td>
<td colspan="2">活动一：创作智能浇水装置
依据阶段三的作品方案设计与草图设计，创作属于每个组个性化的智能浇水装置，包括智能装置硬件的端口连接与代码编写。
活动二：检测优化作品
将小组创作的智能浇水装置到土壤盆栽中进行实际检测，检验智能浇水装置是否能够实现自主浇水并且满足植物土壤的水分需求。</td>
</tr>
<tr>
<td>五
展
示
评
价</td>
<td colspan="2">活动一：小组作品展示与评价
小组派代表携带智能浇水装置作品进行展示，其他小组为展示小组打分。展示内容包括小组分工情况，智能浇水装置的设计方案、材料工具选择、作品功能实现以及项目实践过程中的困难、挑战以及收获等。
<table>
<tr><td colspan="4">《智能浇水装置》作品评价量表</td></tr>
<tr><td rowspan="2">评价内容</td><td colspan="3">具体要求</td></tr>
<tr><td>★（一般）</td><td>★★★（较好）</td><td>★★★★★（优秀）</td></tr>
<tr><td>作品设计</td><td>设计方案不能体现自动浇水功能，缺少创新</td><td>设计方案基本能实现自动浇水功能，比较新颖</td><td>设计方案完全能实现自动浇水功能，很新颖</td></tr>
<tr><td>技术应用</td><td>基本符合要求，能解决自动浇水的问题，但程序存在较多问题。</td><td>能解决自动浇水的问题，程序有少量问题。</td><td>能顺利解决自动浇水的问题，程序能正确执行。</td></tr>
<tr><td>功能实现</td><td>完成基本功能，有少量代码可以改进。</td><td>预设功能基本实现，能准确应用所学知识。</td><td>预设功能全部实现，能灵活应用所学知识，拓展功能应用。</td></tr>
<tr><td>展示交流</td><td>表达较清晰，能够叙述本组项目内容的基本情况。</td><td>表达清晰，对本组项目内容叙述清晰，有条理。</td><td>表达清晰，声音响亮，讲解与展示内容配合到位，有条理性和感染力。</td></tr>
</table></td>
</tr>
<tr>
<td>六
拓
展
提
升</td>
<td colspan="2">活动一：引导知识迁移与拓展
思考：除了让智能浇水装置实现自动浇水，还能利用哪些智能技术来实现更好地养护植物？
例 如：
<table>
<tr><td>智能技术</td><td>实现功能</td></tr>
<tr><td>1.远程数据检测</td><td>利用手机终端能远程查看各项数据
在 Easy loT 中定义 Topic,通过编程向 Topic 发送土壤湿度的数据</td></tr>
<tr><td>2.水位自动监测</td><td>利用超声波传感器检测浇水器的水位，水位过低时提醒用户</td></tr>
</table>
活动二：拓展前沿技术
播放人工智能最新领域视频和智慧农业系统技术应用的视频，了解前言新技术在实际生活、生产中的应用，回顾智能浇水装置项目。</td>
</tr>
</table>

5. 分析与讨论

5.1.融入有效失败策略的项目化学习对学生创造力的影响分析

通过对两组学生在创造力前测中的得分进行独立样本 t 检验分析，评估了实验组与对照组在初始创造力水平上的差异。分析结果表明，在创造力总分及各个维度的分数上，两个组别间均没有显著性差异。

随后，本研究采用了配对样本 T 检验来深入探讨融入有效失败策略的项目化学习对学生创造力的影响。表 2 的数据展示了案例实践干预后的结果：实验组（p=0.000<0.001）和对照组(p=0.039<0.005)的创造力总分均有显著提升，其中实验组的提升幅度更为突出。这表明，在项目化学习中融入有效失败策略能够更有效地促进学生创造力的提升，且这种提升比传统项目化学习更为显著。

表 2 实验组对照组创造力前后测配对样本 *t* 检验

实验组	M	SD	*p*	对照组	M	SD	*p*
冒险性前测	23.92	3.68	0.000***	冒险性前测	24.57	4.22	0.062
冒险性后测	27.52	2.54		冒险性后测	25.62	3.91	
好奇性前测	32.10	5.75	0.000***	好奇性前测	33.15	4.75	0.402
好奇性后测	35．12	4.07		好奇性后测	33.77	5.06	
想象力前测	27.15	4.69	0.000***	想象力前测	27.65	3.99	0.002**
想象力后测	31.67	3.24		想象力后测	29.90	4.34	
挑战性前测	28.00	3.76	0.000***	挑战性前测	28.45	3.18	0.844
挑战性后测	20.10	2.67		挑战性后测	28.55	3.35	
总分前测	111.17	15.49	0.000***	总分前测	113.82	13.60	0.039*
总分后测	124.42	9.82		总分后测	117.85	14.80	

注：*表示 p <0.05;**表示 p <0.01,*** 表示 p <0.001

为了进一步验证实验组和对照组的创造力水平差异，对两组学生的创造力水平后测进行独立样本 t 检验，如表 3 所示。数据结果表明实验组在创造力总分（p = 0.023 < 0.05）以及冒险性（p=0.012<0.05）、想象力（p =0.042<0.05）和挑战性（p =0.025<0.05）三个子维度的得分相较于对照组后测得分具有显著性差异，在好奇性方面未形成显著性差异。因此，实验结果表明，在信息科技课堂中“有效失败策略-项目化学习”比常规的项目化学习更能显著提升学生创造力水平。

表 3 实验组对照组创造力后测独立样本 *t* 检验

变量	*F*	*p*
冒险性	6.1268	0.012*
好奇性	2.899	0.193
想象力	4.193	0.042*
挑战性	1.810	0.025*
创造力总分	7.214	0.022*

注：*表示 p <0.05;**表示 p <0.01,*** 表示 p <0.001

在“有效失败策略-项目化学习”实施模式中，于生成-探索阶段，教师有意推迟教受完成项目任务所需的新知识，相反，他们促使学生通过独立思考及小组协作，深入挖掘并运用既有的知识与经验，进行方案设计和问题解决的尝试。此种策略旨在避免学生在学习新知识的过程中形成固定的思维模式，而是促使他们发展更广泛的发散性思维。在此阶段，教师鼓励学生大胆探索创新的问题解决路径，并在未有新知识输入的前提下增强学生的冒险精神、想象力和挑战性，从而在整体上提升他们的创造力水平。

5.2.实践反思

经由教学反思，本研究识别出若干限制，并提出相应的改进建议。首先，在实施过程中，对于创造力水平的评估主要依赖于创造性倾向量表的定量测量数据，缺乏深入的研究。为获得更全面的分析，未来研究可考虑结合定性方法，如访谈和行为分析，以丰富对创造力水平的理解。其次，对项目化学习的过程性有效分析还不够，还需更完善过程性评价，确保能够通过有效的学习活动实现项目学习的目标。

6. 结论

本研究构建了“有效失败策略-项目化学习”实施新模式，以信息科技学科为例，设计了相应的学习案例，并在实践中探讨了其成效。通过教学实践发现，“有效失败策略-项目化学习”有效地缓解了项目化学习实施中“重成果轻过程”的难题，相比常规的项目化学习培养学生的创造力水平更加有效。未来研究可以探究融入有效失败策略项目式学习在不同学段与不同学科的应用及其实践效果。

参考文献

安富海.(2022).项目化学习的实践困境及改进策略研究.上海师范大学学报(哲学社会科学版)(04),119-125.

曹培杰.(2020).高质量教育需要一场结构性学习变革.人民教育(23),55-58.

姜和丽.(2020).基于核心素养和创客理念的通用技术项目式学习策略.基础教育参考(09),35-38.

林崇德.(2018).创造性心理学.中国图书评论(03),2.

夏雪梅.(2018).学科项目化学习设计:融通学科素养和跨学科素养.人民教育(01),61-66.

臧莺.(2012).创造力是中国学生的“短板”——时报专访国际著名数学家丘成桐.基础教育论坛(08),37-38.

Hung, W. (2015). Cultivating creative problem solvers: the PBL style. Asia Pacific Educ, Rev.16,237-246

Storer, T (2018). The effect of project based learning on the creativity of elementary students. Wilkes University.

KAPUR M. (2008). Productive failure. Cognition and Instruction (S0737-0008), 26(3):379-425.

Henriksen, D., Henderson, M., Creely, E., Carvalho, A. A., Černochová, M., Dash, Davis T., & Mishra, P. (2021). Creativity and risk-taking in teaching and learning settings: Insights from six international narratives. International Journal of Educational Research Open, 2, 100024.

Harris, A., & de Bruin, L. R. (2018). Secondary school creativity, teacher practice and STEAM education: An international study. Journal of Educational Change, 19(2), 153–179.

Smith, M. (2020). Educating risk: How fear of failure is stifling creative practice within Higher Education. In EDUCATING ANIMATORS Academic Conference 2019 Teaching the World's most expressive art form. Manchester Animation Festival and University of Salford.

Manalo, E., & Kapur, M. (2018). The role of failure in promoting thinking skills and creativity: New findings and insights about how failure can be beneficial for learning. Thinking Skills and Creativity, 30, 1–6.

Swanson, H., & Collins, A. (2018). How failure is productive in the creative process: Refining student explanations through theory-building discussion. Thinking Skills and Creativity, 30, 54–63.

健康教育在线项目式学习课程设计——为社区居民设计个性化流感预防宣传册

Curriculum Design for Online Project-based Learning on Health Education: Design Personalized Flu Prevention Booklets for Community Members

陈柏因 [1]
[1] 香港大学教育学院
[*] bychen@connect.hku.hk

【摘要】 健康素养是保障人全面发展的必备素养，对教育公平和社会发展均具有重要意义。中小学的健康教育是培养学生健康素养的重要途径，然而已有的学校健康教育课程存在着知识难以迁移应用、时间空间设备受限的问题，难以在信息化背景下提升学生的健康素养。项目式学习强调学生应用知识解决真实问题，在线教学拓展了学校学习的时间、空间、资源，两者结合有助于学生在学习中发展健康素养。本研究依据"学习目标驱动的设计模型"（Learning-Goals-Driven Design Model）"以"为社区居民设计个性化流感预防宣传册"为主题，设计了健康教育在线项目式学习课程，以期探索后疫情时代学校健康教育课程的有效设计和实施。

【关键词】 健康素养；项目式学习；在线教学；课程设计

Abstract: *Health literacy is of great significance to educational equity and the development of society. Health education in K-12 schools is critical to cultivate students' health literacy. However, existing curriculums could not support knowledge transfer, and were limited in time, space, and equipment, which make it difficult to develop health literacy in the information era. Project-based learning emphasizes applying knowledge to solve authentic problems, and online teaching expands the time, space, and resources of school learning. Based on Learning-Goals-Driven Design Model, this study designed an online project-based learning program on health education with the theme of "designing a personalized flu prevention booklet for community residents", with the aim of exploring the effective design and implementation of school health education in the post-pandemic era.*

Keywords: health literacy, project-based learning, online teaching, curriculum design

1. 前言

健康素养是保障人全面发展的必备素养，对教育教育公平和社会发展均具有重要意义。中小学的健康教育是培养学生健康素养的重要途径，然而已有的学校健康教育课程存在着知识难以迁移应用、时间空间设备受限的问题，难以在信息化背景下提升学习者的健康素养。本研究借助项目式学习和在线教学，设计了以"个性化流感预防宣传"为主体的课程，以期探索后疫情时代学校健康教育课程的有效设计和实施。

2. 理论基础

2.1. 健康素养

健康素养是「个人获取、理解、评价和应用健康信息以作出决策，维护和促进自身健康的知识、动机和能力」（Kickbusch 等，2013），是中国学生发展的六大核心素养之一（林崇德，2019），是保障人全面发展的必备素养，对教育公平和社会发展均具有重要意义。国际广泛认同的健康素养至少包括三个维度：（1）对健康相关信息的阅读和计算能力；（2）从

不同的传播媒介获取、理解和应用信息以应对变化的环境；（3）信息的批判性分析、理解健康的社会决定性因素和参与集体行动。第一条是最基础的能力，第二、三条需要更高层次的认知能力和社会技能（贾绪计等, 2019）。健康素养的培养与学校的健康教育有着密切联系，然而已有的学校健康教育课程存在以下问题：重知识传授，轻知识应用，学生难以将知识迁移到生活情境中；传统面授课的时间、空间及设备有限，学生难以接触大量的、多变的健康信息和资源（傅华，2020；刘芳丽和袁圣敏，2022）。这些局限难以在信息化背景下提升学生的健康素养。学校健康教育亟需转变教学方法，回应提升学生健康素养的需求。

2.2. 项目式学习

项目式学习（project-based learning, PBL）以情境学习和认知理论为基础，该理论认为学习和思维只有发生在特定的情境中才有意义，要使学生产生学习动机，掌握高阶思维，应对现实世界的挑战，需要将教学与真实情境相连接。PBL 作为实施“情境学习”的教学模式，以人工制品的开发为指向，驱动学生主动发现问题并协作解决问题，让学生在动手实践的过程中获得对概念和学习材料的深刻理解（赵楠和裴新宁，2019）。

研究者已经将 PBL 应用到多个学科的课程教学，同时也提出多种 PBL 课程设计模型。其中由学习科学及科学教育著名学者 Krajcik 等提出的“学习目标驱动的设计模型”（Learning-Goals-Driven Design Model）对科学教育改革产生较大的影响（Krajcik, McNeill, & Reiser, 2008）。该模型强调基于课程标准开发 PBL 课程，包括三个阶段，每个阶段又包含若干要素。第一阶段是学习目标的具体化，教师首先识别和分解课程标准，然后形成学习表现；第二阶段是开发课程及教学材料，教师需要创设教学情境、设置学习任务、安排教学序列，形成评价；第三阶段是收集反馈，教师通过教学实验、课程专家及学科领域专家评审收集对项目设计的反馈。本研究将依据这一模型进行健康教育 PBL 在线课程的设计。

2.3. 在线教学

在线教学是利用互联网技术，突破时间和空间的限制，以网络学习资源、工具、平台、空间为依托，实现师生异地教学的一种新型教学形态（焦建利，周晓清和陈泽璇，2020）。在线教育在推动课程改革，提供多元学习机会，解除学习时空和资源限制方面发挥着重要作用。在后疫情时代，在线教学常态化是中小学教育改革的重要趋势，包括开展混合式教学，利用在线教学补充学生的学习，开展个性化教学等（朱永海，龚雨秋和徐莹莹，2020）。从健康教育 PBL 课程设计的角度出发，在线教学具有以下优点：第一，提供丰富多元的真实情境。第二，提供多种工具支持学生从事科学实践，如收集和分析数据、共享交流资料、建构和检验模型。第三，帮助学生开发出多元化、个性化的人工制品。第四，帮助教师从多角度理解的学生的学习以及提供多种支持。目前关于在线 PBL 课程设计的实践和研究较少，本研究正是借助在线 PBL 课程支持学生健康素养发展的尝试。

3. 课程设计

每年秋冬季节是流感的高发期，相关部门都会通过多种方式宣传预防流感的知识。虽然这些知识具有普适性，但是由于生理情况、认知方式、文化背景等因素的不同，不同人群对科普知识的需求和接受程度存在差异。已有的科普宣传通常面向普通人群和一般情境，忽视了特殊人群或情境的特点。与此同时，信息技术的进步导致信息爆发性增长，人们要利用健康信息指导自身生活，需要对大量的、分散的信息进行检索、整合和批判性吸收，然而，这一过程需要较多时间精力以及一定的健康素养。青少年学生可以调查周边社区不同人群及各类情境的情况，为社区成员提供个性化、精准化的流感预防科普，提升自我和社区成员的健康素养。基于上述背景，设定本课程的真实任务：为不同的社区成员设计个性化的流感预防宣传手册。

3.1. 教学对象

初中 8 年级学生。学生对疾病、免疫系统、运动与营养等概念有基本的认识，但是难以构建这些话题的内在联系，也难以将这些话题运用到真实情境中。

3.2. 教学目标

“学习目标驱动的设计模型”要求为项目依照课标制定具体的表现性目标。本课程旨在培养学生的沟通性和批判性健康素养，教师结合初中生物课程标准包含的核心概念和科学探究建立具体的表现性目标，包括：（1）明确个性化流感预防手册的设计背景、设计要求、设计步骤。（2）分析数据图表和资料，了解流感的特点，病因、传播途径，构建营养、运动、免疫系统的联系，提出预防流感的一般措施，提升阅读和理解健康信息的能力。（3）通过调查明确社区不同人群的特点、对流感预防的认知和需求，掌握调查的一般步骤和原则，提升健康信息获取和分析的能力。（4）为社区不同人群设计流感预防手册，提升健康信息获取、利用和批判性分析能力。

3.3. 教学过程

“学习目标驱动的设计模型”要求教师依据教学目标设计教学材料和教学任务。本课程采取混合式学习的方式，结合教师教授和学生自主合作学习开展教学。教学流程遵照设计的基本流程，包括明确需求-实施调查-产品制作-产品改进-产品推广。在教学过程中教师提供相关的资料和在线教学工具支持学生学习基本概念和从事实践，进而培养健康素养。具体的教学流程、教学工具、教学活动如表 1 所示。

表 1　教学过程和教学工具

教学目标	课时及教学内容	在线教学工具	教学活动
(1)	第一课时：明确设计需求	知识论坛 http://kf.aicfe.cn/	1.教师组织学生搜集并共享流感预防科普产品,分析优缺点。 2.提出真实任务：为社区人群制作个性化流感预防宣传册。 3.学生在学习元平台上讨论编写一本家庭战“疫”手册需要考虑的因素，所需的知识和技能，形成项目学习的思维导图。
(2)	第二课时：学习传染病的知识	1.互联网 2.知识论坛	1.学生通过搜集、共享、分析讨论多种资料，梳理流感的发展脉络，理解流感的特点、传播途径。 3.基于已学知识和互联网信息检索形成预防流感的思路。 2.教师引导学生小组批判性思考上述互联网信息的可靠性。
(2)	第三课时：免疫系统及其与营养和运动的联系	百度脑图 https://naotu.baidu.com/	1.教师为学生提供病人的患病信息(如验血单)、新闻素材(如疫苗开发)，引导学生建立免疫防线的模型。 2.学生根据“免疫力、营养、运动”相关的生活实例和研究数据，以小组为单位合作构建概念图表征三者之间的关系。
(3)	第四课时：社区人群调查	互联网	1.教师讲解调查的一般步骤和原则。 2.学生小组自主选择感兴趣的社区人群（比如老人、糖尿病人、学生、超市服务员等），制定调查计划，教师提供反馈。 3.学生小组通过多种方式（如调查问卷、访问等）对社区人群的特点、流感预防知识储备、需求、媒体偏好等进行调查。
(4)	第五课时：流感预防宣传手册的制作和改进	1.在线协作工具 2.热量计算工具	1.学生小组基于调查结果，设计针对特定人群的流感预防宣传册，从免疫力、营养、运动方面提供科普知识、预防措施。 2.学生小组内部对宣传册进行反思和改进，教师提供反馈。 3.学生小组进一步联系特定人群，获得宣传册的改进建议。

（4）	第六课时：流感预防宣传册的推广	知识论坛	1.学生小组展示流感预防宣传册。 2.学生小组在知识论坛在线出版流感预防宣传册，学生小组进行组内和组间互评，教师进行评价。 3.学生小组后续持续细化和更新流感预防宣传册。

3.5. 教学评价

本课程以表现性评价为主，教师可以从多个方面进行评价，如学生制作的宣传册、在线协作过程、对宣传册的展示等。评价工具通常为量规，通常包含与学习目标对应的评价指标、递进分数及描述。量规可以用于学生自评，互评以及教师评价。表 2 示例评价量规。

表 2 宣传手册评价量规（节选）

评价指标		评分
一级指标	二级指标	3：大部分达到标准
1 内容	1.1 内容遵循科学知识和科学原则；1.2 内容具有逻辑性	2：部分达到标准
2 编排	2.1 排版清晰；2.2 图表数量合理	1：几乎达不到标准

4. 课程设计启示

4.1. 选择有意义的真实问题发展学生的健康素养

真实的生活情境和问题能够让学生应用和迁移知识和技能。对于学习者而言，只有有意义的问题才能引发学习兴趣，维持学习动力，提升健康素养。选择真实情境和问题时需要考虑以下因素：与学习者知识经验的相关性、与学校课程内容的关联性、在专业领域的重要性。

4.2. 合理选择在线教学工具支持 PBL 的开展

在线教学工具能够助力高质量 PBL 的实现，促进教学改革。选择在线教学工具时应坚持（1）教学为本而非工具为本，根据教学目标和教学活动选择合适易用的在线教学工具；（2）立足素养培养而非知识搬运，利用在线教学工具转变传统课堂的教学方式和学习方式。

参考文献

朱永海、龚雨秋和徐莹莹（2020）。后疫情时代中小学在线教育常态化应用的整体推进路径——基于美国 K-12 在线教育的经验。 **现代教育技术**，**(11)**，120-126。

刘芳丽和袁圣敏 (2022)。芬兰中小学健康教育课程: 背景、特征与启示。**外国教育研究**， **(05)**，74-87。

林崇德（2017）。构建中国化的学生发展核心素养。**北京师范大学学报(社会科学版)**， **(01)**，66-73.

赵楠和裴新宁（2019）。问题式学习和项目式学习。**上海教育**，**(07)**，72-74。

傅华（2020）。**健康教育学**。北京：人民卫生出版社。

焦建利，周晓清和陈泽璇（2020）。疫情防控背景下“停课不停学“在线教学案例研究。**中国电化教育**，**(03)**，106-113。

Kickbusch, I., Pelikan, J. M., Apfel, F., Tsouros, A. D., & Agis, D. (2013). *Health literacy: The solid facts*. World Health Organization Regional Office for Europe. https://iris.who.int/handle/10665/326432

Krajcik, J., McNeill, K.L. & Reiser, B.J. (2008), Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*, 1-32. https://doi.org/10.1002/sce.20240

GPT/生成式人工智能与STEAM教育何以结合

The Possibility and Path of the Combination of GPT/GAI and STEAM Education

沈柯[1*]，李锋[2]
[1]华东师范大学 教育信息技术学系
[2]华东师范大学 教育信息技术学系
[*]1459391318@qq.com

【摘要】STEAM教育在全球范围内得到普及的背景下，生成式人工智能（Generative Artificial Intelligence，GAI）近年在教育领域引起了广泛关注。本文立足于STEAM教育当前教学实施方面所面临的现实问题，剖析了GPT与STEAM教育的契合点，认为GPT于STEAM教学实施而言，可以支持探究式学习的开展、推动教学资源的普及与个性化、弥补教师储备的不足以及助力评价手段的优化等。
【关键词】生成式人工智能；GPT；STEAM教育

Abstract: *Against the backdrop of STEAM education being popularized globally, Generative Artificial Intelligence (GAI) has attracted widespread attention in the field of education in recent years. Based on the practical problems faced by the current teaching implementation of STEAM education, this paper analyzes the meeting point between GPT and STEAM education, and holds that GPT can support the development of inquiry learning, promote the popularization and personalization of teaching resources, make up for the shortage of teachers ' reserves and help the optimization of evaluation methods in the implementation of STEAM teaching.*
Keywords: generative artificial intelligence, GPT, STEAM education

1. 研究背景

STEAM教育是在真实的问题情境中，通过学科整合的方式培养学生的跨学科素养与创新能力的教育理论与实践（彭敏和郭梦娇，2018）。随着信息时代的不断发展，各种新兴技术层出不穷，科技融入了大众生活的方方面面。因此，科技创新能力的培养日益成为世界各国教育追求的目标。STEAM教育作为一种旨在培养创新人才的教育模式，愈发受到各国政府的关注。

“人工智能+教育”背景下，物联网、人工智能、数字孪生、大数据与云计算、区块链等技术的蓬勃发展，正在深刻地影响着教育领域的各个方面。这些技术不仅改变了我们获取知识的方式，也改变了我们教授知识的方法。2022年底由OpenAI公司发布的大语言模型ChatGPT（以下简称“GPT”）使大众前所未有地亲身体验到人工智能技术对教育的革命性影响，在世界范围内引起轩然大波。祝智庭等学者认为，人工智能教育的目标是培养具有扎实技术素养的创新人才，这与当前社会亟需的人才核心竞争力相符合（祝智庭、单俊豪和闫寒冰，2019）。

综上所述，不难看出人工智能教育与STEAM教育在人才培养目标上存在高度的同质性，即两者的核心追求均在于培养学习者的创新能力。因此，本文旨在以GPT为例，试探讨生成式人工智能技术与STEAM教育得以结合的逻辑与方式。

2. GPT与STEAM教育何以结合

GPT 已然成为全球的热议话题，以其为代表的生成式人工智能技术正逐渐成为推动教育领域变革的重要力量。就 GPT 与 STEAM 教育的结合而言，笔者认为，教育工作者应当尽可能减轻对于新兴技术的兴奋心理，避免落入技术主义陷阱，即在教育实践与研究中以技术的应用为导向，忽视教学与研究的内容与目的。从加德纳公司（Gartner）发布的技术成熟度曲线（Hype Cycle）中我们可以预测到（Gartner, 2003），当前生成式人工智能技术仍处于泡沫巅峰期，在一段时期后，公众对于该技术的关注度可能会下降，与此同时，该技术稳步发展，并最终以更为成熟的姿态和更高的市场接受度回归大众视野。因此，教育工作者应当对 GPT 持审视的态度，不宜天马行空、过度吹捧，以免造成“捧杀”现象。对当前在 STEAM 教育中融入 GPT 这一工具可行性的讨论，应以充分了解 GPT 为前提，立足于 STEAM 教育在教学实施方面所面临的现实问题，从而尝试寻找两者的契合点。新兴工具应用于教育中更需要立足教育问题相结合，探索解决问题的方法与路径，而非单纯的炫耀技术功能。

2.1. STEAM 教育面临的现实问题

STEAM 教育在中国仍处于方兴未艾阶段，实践探索中已经取得了一系列初步成果。然而，考察我国当前 STEAM 教育在教学实施方面的现状，还存在着课程理念模糊、探究设计不周、教学资源稀缺、教师储备不足、评价体系粗疏等问题。

2.1.1. 课程理念模糊

作为课程设计与实施的上层建筑，课程理念影响着教学目标、教学内容、教学方法、评价方式等教学全过程。由于我国暂时还未形成完善的 STEAM 课程系统以及相应的师资培养体系，不少学校和教师对于 STEAM 课程理念把握不清晰，实践过程中呈现三种错误倾向：一是将 STEAM 教育与创客教育混为一谈，二是将其跨学科理念简化为多学科内容的简单杂糅，三则是误认为跨学科教学可以代替分科教学（田娟和孙振东，2019）。

2.1.2. 探究设计不周

基于项目或问题的教学设计是教师组织学生开展探究式学习的主要困难之一。从项目设计角度看，有些项目缺乏足够的深度，无法真正激发学生的创造性思维和问题解决能力，有些项目则过于复杂，忽视对学习者的分析，容易导致学生在学习过程中产生挫败感，丧失学习兴趣。从活动设计角度看，缺乏系统性的设计指导和实践经验分享，使得教师在活动设计和实施中较为盲目，过于强调学生的发现，导致学生迷失在探索过程中，对相关概念的理解并不透彻，甚至偏离原本的学习目标。

2.1.3. 教学资源稀缺

教学资源的丰富程度往往制约着一门课程的发展。由于 STEAM 教育跨学科性与活动性，STEAM 课程要求丰富多样的实践性教学资源，包括科学实验器材、工程设计工具、艺术材料以及技术设备等。然而，由于财政因素以及高质量的本土化 STEAM 教育资源的稀缺，许多学校和教育机构无法提供足够的资源支持。这不仅限制了教师在课堂上有目的地引导学生进行实践活动的能力，也影响了学生对 STEAM 学科的深入理解和兴趣培养。

2.1.4. 教师储备不足

STEAM 教育需要具备跨学科知识、项目管理和团队协作能力的教师，然而，由于我国分科教学的传统，高校当前的师范课程体系并不能很好地兼顾 STEAM 教育对师资的要求。目前学校中主要承担 STEAM 教育工作的多为兼任教师，专任教师较少（周雯晨和肖遥，2020）。大多数教师在 STEAM 教育领域的专业知识和实践经验相对薄弱，导致他们在 STEAM 课程的设计和实施中缺乏信心和能力。

2.1.5. 评价体系粗疏

STEAM 教育注重培养学生的创造性思维和实际问题解决能力，然而，现有的评价体系仍然主要侧重于传统的考试和分数评价。这种评价体系无法全面客观地反映学生在 STEAM 教

育中所取得的综合能力，消解了STEAM教育工作者与学习者的努力。不少研究者提出STEAM教育应强调表现型评价与过程性评价（崔鸿、朱家华和张秀红，2017；朱慧仪，2018；周雯晨和肖遥，2020；詹泽慧和吕思源，2023；廖艺东，2023）。然而，落实到实践中，对学生过程性学习数据的收集以及对其项目作品的评价量规的制订又是让教师们望而却步的两座大山。

2.2. GPT 与 STEAM 教育的契合分析

GPT 生成回答的范围可以涵盖几乎所有领域，但在回答复杂问题或涉及某领域内专业性较强的问题时并不稳定，易产生 AI 幻觉(Hallucinations)问题。GPT 的这种特性与 STEAM 教育的综合性与跨学科性存在高度契合，应用于 STEAM 教育中的 GPT 可以很好地扬长避短，辅助课程的全过程开展。

2.2.1. 支持探究式学习的开展

在不遥远的未来，GPT 这一类工具或许能成为学校推进探究式学习的主要抓手。首先，GPT 可以为教师提供探究式学习项目或问题情境的设计建议。通过分析大量的教育文献和实践案例，GPT 可以生成具有合适深度和挑战性的项目设计建议，这有助于解决项目设计过于肤浅和过于复杂的问题。第二，以往的 STEAM 教学中，教师常常难以把握课堂中学生探究学习的比例与尺度，而借助 GPT 作为学生的探究学习工具，基于它知识点之间相互关联的回答特征，可以一定程度防止学生偏离原本的学习目标。第三，GPT 可以基于海量的数据学习词与词之间的关联信息，并且可以利用这些信息进行组合创新，这也是 GPT 丰富创意的来源（陈静远、胡丽雅和吴飞，2023）。与 GPT 的交互，可以帮助学生打开项目设计的思路、学习发现问题的思路，从而培养学生的创造力。第四，难以兼顾每个学生是大班教学下开展探究式学习的老大难问题，但在 GPT 的辅助下，每个学生都可以得到实时的反馈与问题解决，排斥与同伴沟通的学生也能在与GPT这一聊天机器人的问答互动中培养表达协作的能力与意向。

2.2.2. 推动教学资源的普及与个性化

当前教育市场上跨学科项目资源不仅稀缺，还存在着良莠不齐和价格高昂的问题。而 GPT 以其低廉的成本，为广大校园提供了更为迅速而经济的解决方案。教师可以借助 GPT 迅速搜集到符合特定形式的教学资源，甚至能够利用 GPT 根据教学主题创造出个性化的教学资源、辅助材料以及富有创意的课堂活动。这为教学注入了新的活力，使教育资源的利用更为高效。对学习者而言，GPT 则能够根据学生的学情与兴趣，为其量身定制学习路径并推荐相应的学习资源。这不仅有助于培养学生的个性化发展，还能够有效提高其学习效率。学生在这样的教育环境中，能够更加专注于感兴趣的领域，激发学习的兴趣和热情。GPT 的引入不仅为教育资源的普及提供了技术支持，更为教学个性化注入了新的可能性。

2.2.3. 弥补教师储备的不足

STEAM 教学通常采用基于项目或问题的教学模式，结合其跨学科理念，学生的学习内容呈现多样性、复杂性的特点。在现有的师资培养体系下，很难有教师能够在教学中为学生灵活提供不同领域的专业知识，以帮助学生更好地进行探究。GPT 便很好地填补了这一漏洞，其庞大的预训练数据量几乎覆盖了所有常见领域，支撑其为学生提供广泛的学科见解。教师也可以从中获得丰富的背景知识，为教学内容的设计提供更多元的参考。并在教学实施过程中为学生提供更多元化、有针对性的指导。

2.2.4. 助力评价手段的优化

如前文所述，STEAM 教育评价正在逐渐重视过程性评价与表现性评价。在此过程中，对学生过程性学习数据的收集以及对其项目作品的评价量规的制订是当前 STEAM 教育评价面临的主要问题。而 GPT 可以助力这两种评价过程的开展。一方面，与 GPT 的交互几乎伴随

学生学习全过程，而这些交互数据将被自动储存下来，利用 GPT 的自然语言理解能力，可以分析学生的语言表达和思维过程。同时，结合学习分析技术，教师可以更好地评价学生在探究性学习过程中的表现，建立更精准的学生能力档案。另一方面，当学生的学习成果相对开放时，教师可以预设好指定的评价维度，让 GPT 根据学生学习成果的不同主题或不同形式，生成相应的评价量规，推动评价多元化。

此外，GPT 还可以培养学生的批判性思维能力与审美能力。在使用 GPT 进行学习的过程中，学生经常会在回答的文本中发现偏差，此时批判意识便不再是一种可有可无的附加能力，而是一种为了更好地使用该工具而必须具备的基础技能。同时，随着机器视觉的发展，GPT 也可以在项目开发过程中辅助弥补传统教育下学生艺术素养的不足，学生可以利用 GPT 辅助多媒体制作，提升成果展示的质量和多样性。综上所述，GPT 作为教学辅助工具与 STEAM 教育具有高度契合性，可以促进 STEAM 教学的实施。

3. 结语

STEAM 教育作为一种旨在培养创新能力与解决问题能力的教育，受到学界推崇。GPT 的运用有助于推动 STEAM 教学实施全过程。可以预见，GPT 相关技术的不断发展对 STEAM 教学的影响会出现更多可能性，但同时也要意识到 GPT 的局限性，包括教师和学生的接受程度、技术的公平性、信息伦理安全和隐私风险等。本文仅分析了 GPT 与 STEAM 教育的契合逻辑，但将 GPT 技术应用于教学实践的 STEAM 教育究竟应当如何开展，仍需要进一步的研究。

参考文献

田娟和孙振东（2019）。跨学科教学的误区及理性回归。**中国教育学刊，4**，63-67。

朱慧仪（2018）。**融合 STEM 理念的学习活动的要素特征与评价**。上海：华东师范大学。

陈静远、胡丽雅和吴飞（2023）。ChatGPT/生成式人工智能促进以知识点为核心的教学模式变革研究。**华东师范大学学报(教育科学版)，41(7)**，177-186。

周雯晨和肖遥(2020)。我国中小学 STEAM 教育实践困境及路径探析。**教育观察，9(47)**，122-124。

祝智庭、单俊豪和闫寒冰（2019）。面向人工智能创客教育的国际考察和发展策略。**开放教育研究，25(1)**，47-54。

赵奎英（2020）。“新文科”“超学科”与“共同体”——面向解决生活世界复杂问题的研究与教育。**南京社会科学，7**，130-135。

崔鸿、朱家华和张秀红（2017）。基于项目的 STEAM 学习探析，核心素养的视角。**华东师范大学学报(教育科学版)，35(4)**，54-61+135-136。

彭敏和郭梦娇（2018）。STEAM 教育的基本内涵与发展路径研究。**教育理论与实践，8(25)**，14-18。

詹泽慧和吕思源（2023）。基于课堂观察的 STEM 教学评价，协议、要素与方法。**开放教育研究，29(4)**，65-80。

廖艺东（2023）。**指向计算思维培养的跨学科主题教学设计与实施**。上海：华东师范大学。

Gartner. (2003). *Understanding Gartner's Hype Cycles*. State of Connecticut-CT: Gartner.

OECD. (2005). *Students, Computers and Learning: Making the Connection*. Retrieved from extension://ngphehpfehdmjellohmlojkplilekadg/pages/pdf/web/viewer.html?file=https%3A%2F%2Fwww.nuovoeutile.it%2Fwp-content%2Fuploads%2F2015%2F09%2FPISA-2012-students-computers-italy.pdf.

戲劇學習對 STEM 機器人課程中建立負責任文化的影響：學生學習成效的探討

Exploring the Impact of Drama-based Learning on Establishing an Accountable Culture in STEM Robotics Curriculum and Its Effects on Student Learning Outcomes

馮建中[1*]，王振漢[2]，馮詩敏[3]，馮詩婕[4]，陳國棟[2]
[1]中央大學資訊工程學系
[2]中央大學學習科技研究中心
[3]台灣科技大學數位學習與教育所
[4]台灣師範大學健康促進教育學系
[*]reagon@gmail.com

【摘要】 本研究旨在探討戲劇學習對 STEM 機器人課程中建立負責任文化的影響，特別在學生學習成效。透過準實驗設計，選取北部三所大學中的班級，分為實驗組和對照組。實驗組接受「建立負責任文化」的教學干預，包括學生互動、合作，以及教師的評估和監督；對照組則採用傳統的專題教學。研究長達三年，透過學業表現、問卷調查和數據分析，結果顯示採用戲劇學習能明顯提升學生學習成效，同時建立積極的負責任文化。此研究為 STEM 教育提供實證支持，強調在 STEM 課程中建立負責任文化的轉型，進一步豐富學生不僅將數位科技融入機器人教學，同時提升知識服務素養。

【關鍵字】 戲劇學習、專家斗篷、負責任文化、學生學習成效

Abstract: ***This study explores the impact of drama-based learning on establishing a responsible culture in STEM robotics curriculum, with a focus on student learning outcomes. Using a quasi-experimental design, classes from three universities in the northern region were selected and divided into experimental and control groups. The experimental group received interventions promoting a 'responsible culture,' emphasizing student interaction, collaboration, teacher assessment, and supervision. Over a three-year period, data from academic performance, surveys, and analysis revealed significantly enhanced student learning outcomes and the establishment of a positive responsible culture. This research provides empirical support for STEM education, emphasizing the transformative role of cultivating responsibility in STEM curricula, enriching students' integration of digital technology into robotics instruction while enhancing knowledge service literacy.***

Keywords: Drama-based Learning, Mantle of the Expert, Accountability Culture, Student Learning Outcomes

1. 前言：

在現今大學教育面臨過時危機的背景下，數位教學模式的轉型成為解決這一挑戰的重要策略。本研究旨在探討大學教育面臨傳統模式的過時危機，其中包括傳統教學模式的限制、學生需求的轉變以及現代社會對教育的新期望。數位教學模式的轉型被視為提高教學模式靈活性的策略，以更好地適應現代教育需求。然而，數位教學模式的轉型也面臨技術整合不足等挑戰。同時，我們融入戲劇學習及其核心概念「專家斗篷」，作為填補傳統專題式教學缺陷的創新方法。此外，我們將聚焦於 STEM 機器人教育，強調其在培養學生創造力、問題解決能力和協作精神方面的重要性。

戲劇學習（Drama-based learning）是一種引人注目的新型教學方法，尤其以「專家斗篷（Mantle of the Expert）」概念為核心，吸引了廣泛的關注。根據Heathcote 與 Given (2013)及 Arman (2019)的研究，探討了「專家斗篷」概念在 STEM 機器人教育中的應用，以及它在填

補傳統專題式教學缺陷方面的卓越表現。該方法透過角色扮演和建立標示系統，激發學生學習興趣，培養責任感，提升整體學習體驗。研究強調了「專家斗篷」的四個元素：負責的團隊，客戶，代表團和任務，並在不同學科領域中的應用，特別強調其在 STEM 機器人教育中的關鍵性和創新性。

參考其他研究，特別是 Lin,Hwang 與 Chen (2021)指出 STEM 機器人教育是培養學生創造力、問題解決和協作精神的重要途徑。然而，傳統專題式教學模式的限制仍需優化。這項研究分析了傳統專題式教學的利弊，強調了戲劇學習方法，特別是「專家斗篷」在填補這些缺陷方面的關鍵作用。透過結合 STEM 教育和戲劇學習，該研究旨在提高學生參與度、理解深度，同時促進合作和溝通技能的發展。

近年來，LINE 社群學習平台系統受到關注，其在學生學習成效方面的實證研究不斷增加。將聚焦於已有的研究，探討 LINE 社群學習平台系統對學習成效的影響，並理論基礎的闡釋。透過整合這些方面的研究，期望深化對社群學習平台系統在教學實踐中的理解，以提升教學效果和學習成果。

研究問題：

在 STEM 機器人課程中，透過運用「專家斗篷」概念，如何促進建立負責任文化，進而對大學生的學習成效、創造力和解決問題能力產生具體影響？

研究假設：

在統計顯著性水準（$\alpha \leq 0.05$）下，我們提出以下虛無假設：

在 STEM 機器人課程中，透過運用「專家斗篷」概念，戲劇學習和建立負責任文化之間的交互作用將不會對大學生的學習成效、創造力和解決問題能力的平均分數產生統計上的顯著差異。

2. 文獻回顧

2.1. 大學教育的過時危機與數位教學模式的轉型探討

大學教育的過時危機涉及傳統模式無法滿足現代學生需求，根據 Brown（2015）的定義，其原因包括保守僵化的教育機構、單一無趣的教學模式、學習內容的過時脫離、不公無效的學習評估，以及昂貴不合理的學習成本，導致學生動機和成效下降，影響教育質量和效益。

數位教學模式的轉型旨在有效整合和創新應用數位科技，提供更具意義和個人化的學習體驗，培養數位素養和 21 世紀的關鍵能力。策略包括多元化教學方式、擴展連結教學場域、優化個性化教學內容，以及改善創新教學和學習評估。

參考 Hew 與 Cheung（2014）及 Johnson（2015）的研究，前者支持 Web 2.0 技術在 K-12 和高等教育的應用，顯示其促進學生學習動機、參與度、合作性、創造性和學習成果。後者展望數位技術的趨勢，包括可穿戴技術、混合學習、學習分析和微學習，提供對教育的潛在影響和挑戰的分析。這些研究為本研究提供了有力的實證支持和未來趨勢參考。

2.2. 戲劇學習的核心概念「專家斗篷」

根據 Dorothy Heathcote & Gavin Bolton (2006) 著書中提及的「專家斗篷」，最初由 Dorothy Heathcote 於 1970 年提出，是一種戲劇教學方法，特點在於將學生視為某領域專家，教室成為虛構工作場景。學生透過扮演不同角色，參與任務與活動，解決真實或模擬問題。此教學方法的核心原則包括整合性、實踐性、創造性、合作性和反思性，有效激發學生學習興趣，提高參與度，培養多元能力和素養。

關於「專家斗篷」的實踐與應用，O'Neill & Lambert (2013) 探討了其在學生主導和協作學習中的運用，提出成功教學設計和實施建議。Birkett (2017) 提供了戲劇學習中的詳細指

導，強調「專家斗篷」的重要性，指出教學挑戰和解決方案。Prendergast (2019) 從英語教學角度，探討戲劇策略如「專家斗篷」在提升語言能力和批判思維中的作用。

綜上所述，對本研究提供了對「專家斗篷」概念的清晰理解，並展示了其在不同教學環境和學科領域的多樣性和實用性。這不僅為 STEM 機器人教育和 LINE 社群學習平台的教學設計提供了參考，也為學生學習成效的提升開啟了新的可能性。

2.3. STEM 機器人教育與學生學習成效

在 STEM 機器人教育領域，學者們探討了如何透過創造力、問題解決能力和協作精神的培養來提高學生的學習成效。首先，Resnick (2013) 提倡的「終身幼兒園」概念強調透過項目、熱情、同儕和遊戲的結合來促進有意義的創造性活動，這有助於提升學生的學習成效。接著，Kafai (2016) 探討了編程在 STEM 機器人教育中的重要性，指出透過編程可以培養學生的邏輯思維、問題解決和創造性思維。Bers (2017) 關注於在幼兒教育中應用編程和計算思維，推廣「編程作為遊樂場」的概念，透過遊戲化方式培養孩子的創造力和問題解決能力。Papert (2017) 則強調機器人教育中的實踐性學習和專題探究的重要性，這些方法激發學生的學習興趣和創造力。最後，Hmelo-Silver (2018) 深入探討了在跨學科學習中設計挑戰性問題的重要性，強調在 STEM 機器人教育中設計具有挑戰性的問題對於促進學生學習和學習成效至關重要。

綜合以上研究，本研究提出一個整合性的 STEM 機器人教育模型，該模型結合了項目、熱情、同儕和遊戲等元素，同時強調編程和計算思維的引入。這不僅提高學生的參與度和理解深度，還促進了合作和溝通技能的發展。通過這樣的模型，STEM 機器人教育將能夠更有效地促進學生學習成效的提升，為學習環境注入更多活力。

2.4. LINE 社群學習平台系統與學生學習成效

為了探討 LINE 社群學習平台系統在大學教育中的應用，以及該平台如何提升學生的學習成效，故從三個方面進行分析：LINE 平台的概念與功能、教學設計及應用、以及學習成效與影響。

首先，根據 Chang 與 Lai (2011) 的定義，LINE 社群學習平台系統是一個利用行動通訊應用程式建立的支援教學和學習的線上社群，提供學生即時、互動、多媒體和社會化的學習環境。其核心功能包括便利性、多元性、互動性、社會性和開放性。這些功能使 LINE 成為適合大學教育的數位教學模式轉型工具，能提升學生的學習品質和效率。

接著，參考了 Wang (2012)、Hung (2012)和 Liu 與 Cheng (2014)的研究，從社會註釋、問題導向學習和混合學習的角度探討 LINE 平台在不同學科和課程中的應用和成效。Wang (2012) 的研究發現社會註釋技術可以提高學生的注意力、理解力、學習滿意度和成績。他們提出了一個基於社會註釋的教學設計模型，包括資源準備、註釋創建、分享和整合四個階段。此外，Hung (2012)的研究則探討了問題導向學習在 LINE 平台的應用，該模式激發學生的學習興趣和動機，促進互動和協作。該教學設計包含問題呈現、探索、解決和反思四個階段，鼓勵學生運用高階思考技巧，進行合作學習和創造性思考。

總結來說，透過整合上述學者的理念和研究成果，探索了一個更為全面且具體的 LINE 社群學習平台系統模型。這個模型不僅提高了學生的參與度和理解深度，還促進了合作、溝通技能的發展，為大學教育提供了新的數位教學方法。

3. 研究方法

3.1. 研究設計：

本研究採用準實驗研究設計，以探討「專家斗篷」戲劇學習方法在 STEM 機器人教學中的應用，並比較其與傳統專題教學方法的學習成效差異。研究分為實驗組和對照組，兩組學生均參與相同的 STEM 機器人學習課程，但教學介入方式不同。實驗組採用「專家斗篷」戲

劇學習方法，讓學生扮演不同的專家角色，透過 LINE 平台的記事本功能記錄學習歷程，並進行小組合作和問題解決。對照組採用傳統專題教學方法，讓學生以個人為單位進行機器人設計和編程，並進行成果展示。

3.2. 研究對象：

研究對象為台灣北區三所大學中相當規模的班級共328位學生,他們都是自願選修STEM機器人學習課程的學生。研究者將依據學生的學習背景和興趣，以隨機分派的方式將學生分為實驗組和對照組，每組各佔約有50%的學生。實驗組的學生將被進一步分為26個小組，每組6人，並在LINE平台上建立學習社群，以進行「專家斗篷」戲劇學習活動。對照組的學生則以個人為單位進行傳統專題教學活動，不進行專家小組合作或戲劇學習。

3.3. 數據收集和分析：

本研究將採用混合方法研究，結合定量和定性的數據收集和分析，以獲得更全面和深入的研究結果。定量數據收集將包括學習成效測驗、負責任文化問卷和學習動機問卷，以量化學生在知識和技能方面的學習成效，以及在學習興趣、滿意度和自我效能等方面的學習態度和動機。定性數據收集將包括學習歷程記錄、觀察記錄和訪談記錄，以了解學生在學習過程中的互動模式、學習策略和問題解決能力，以及在戲劇學習活動中的體驗和感受。數據收集完成後，將使用 SPSS 統計軟體對定量數據進行描述性統計和推論性統計，以比較實驗組和對照組在學習成效、學習態度和學習動機等方面的差異，並探討學生背景特徵和學習成果之間的相關性。

3.4. 研究步驟：

進行文獻探討，了解戲劇學習、專家斗篷、負責任文化和學生學習成效等相關理論和研究，並據此設計研究問題和假設。進行教學活動設計，根據STEM機器人學習課程的教學大綱和目標，分別為實驗組和對照組設計適合的教學活動，並確認教學資源和教學時程。進行數據收集工具設計，根據研究目的和問題，設計適合的數據收集工具，包括學習成效測驗、負責任文化問卷、學習動機問卷、學習歷程記錄表、觀察記錄表和訪談大綱，並進行專家效度和信度檢驗。分別在期初與期末進行教學介入和數據收集，根據教學活動設計和數據收集工具設計，分別對實驗組和對照組進行教學介入和數據收集，並確保數據的完整性和有效性。進行數據處理和分析，根據數據收集和分析方法，對收集到的定量和定性數據進行處理和分析，並將分析結果以表格呈現。進行研究結果討論和建議，根據數據分析結果，討論研究發現和意涵，並提出研究限制和未來研究的建議。

4. 教學研究情境：數據分析與解釋

這項研究聚焦於大學STEM創意機器人課程的教學模型，該模型歷經多階段發展。起初，於2019年，在校內資訊管理科系教師聚會中討論STEM機器人及編程教育的未來。2020年成立資訊管理科系教師專業學習社群，結合資管、數學教師及業界專家，形成跨領域的教學團隊。透過每月兩次的聚會及LINE社群討論，共同研討STEM創意機器人課程的理論與實際應用。以大學教授及業界專家指導，確保課程符合108課綱要求，同時強調教師在機器人教育及課程規劃的專業能力，以及學生的創新整合能力。自2020年起，將高中STEM課程擴展至大學，培養學生的創新能力。透過角色扮演，引領學生模擬成大學教授，期望培養他們成為能協助系上教學的知識工作服務者。課程結束後，與轄區國中合作建立AI機器人程式設計社團，由接受培訓的學生擔任助理講師，形成負責任的師生教學團隊，提供更全面的教育服務，促使大學STEM創意機器人課程的學習成效提升。

表 1 大學STEM創意機器人課程學習成效提升的面向(實驗組控制面向)

學習脈絡	專家斗篷在學生學習成效的面向	建立負責任文化的知識服務面向
學生角色（前期投入**戲劇學習**）	教師初期投入學生的教學策略中，以「專家斗篷」戲劇學習為核心，學生被引導扮演機器人領域的專業人物，如程式設計師、機器人工程師等。透過具體的情境設計，學生參與模擬專案任務，培養協作、創造力，並投入解決真實或模擬問題的過程。融合整合性、實踐性、創造性、合作性和反思性的核心原則，激發學生學習興趣，提高參與度，同時培養多元能力和素養。透過實際投入，學生在「做中學」的學習樂趣中成為機器人領域的知識工作服務者。這種教學模式的運用，不僅為 STEM 機器人教育提供了有力的指引，也為學生的學習體驗帶來新的可能性。	核心素養是教育的關鍵，旨在培養學生在知識、能力和態度上的綜合素養，以積極回應個人和社會需求。為建立負責任文化，教學策略應聚焦於素養同時涵蓋 Competence 及 Literacy，強調學科知識與生活的結合。核心素養強調動態的學習模式，要求學生能運用知識，實踐中體現全人發展。十二年國教與 108 課綱強調終身學習，使學生具備基礎能力，與社會接軌。在教學中，教師避免極端的當責風格，既不濫用權威，也不袖手旁觀。適度的引導與參與，讓學生主動承擔責任，是建立負責任文化的重要策略。這種教學模式鼓勵學生積極參與學習，培養自主學習的習慣，使其具備應對現實挑戰的能力。
專家角色（中期融入**專家斗篷**）	在中期教學中，教師通過 STEM 創意機器人課程，以學生在團隊中扮演不同專業角色的方式，培養其專業知識和技能。學生在這個過程中感受到專家的地位和尊重，同時也激發了他們進一步進行自主學習與管理的動機。課堂中的競賽訓練和運算思維成為培養專業知識的特色內容，同時強調團隊如何以專家的角色視角服務客戶。這種以「專家斗篷」為核心的教學模式不僅豐富了學科知識，也為學生提供了實際應用和服務的機會，為未來的學習和職業發展奠定了堅實基礎。	教師著重培養學生的核心素養，包括知識、技能、態度和全人發展。以學生為本的學習模式，結合 AI 與機器人創新應用設計，以專家斗篷身份激發學生的學習與實踐動力，透過小組合作學習模式賦予每位學生特定的工作職責。這種組織領導協助學生克服學習上的難題，更培養了負責任的文化。注重素養導向的學習模式，強調學以致用，鼓勵學生運用知識解決實際問題，反思並找出答案。這樣的教學策略符合建立負責任文化的目標，讓學生樂在其中，更具備應對未來挑戰的能力，促使他們成為具有責任心的終身學習者。
教師角色（後期協助**建立負責任文化**）	教師將教學情境轉變為具體的角色扮演，讓學生模擬成大學教授的形象，這符合「專家斗篷」的核心理念。透過這樣的教學策略，學校期望學生在課程中培養新知識與技術，成為能協助系上教學的知識工作服務者。結束課程後，學校進一步與國中學校合作，建立 AI 機器人程式設計社團。這個社團由接受課程培訓的學生擔任助理講師，以知識服務者的身份協助系上講師進行動手實作教學，形成	透過 LINE 社群學習平台系統，教師在後期協助小組合作學習的團隊中，建立負責任文化的教育學習面向發揮關鍵角色。採用「當責四步驟」教學方法，引導學生正視現實、承擔責任、解決問題、著手完成。透過扮演「專家」，學生在教師指定的問題情境中學習投入，同時激發他們走出舒適圈，建立負責任文化。教學策略注重引導學生運用高階思考技巧，強調合作學習和創造性思考。整合問題導向學習和混合學習，提升學

	負責任的師生教學團隊,提供更全面的教育服務。這樣的教學模式不僅融入了「專家斗篷」的概念，為學生提供了個豐富而具有深度的學習體驗。	生參與度、理解深度，同時培養溝通技能。這種全面的教育模型不僅提高了學習效果，更為大學教育帶來了新的數位教學方法，促進了負責任文化的建立。

在**實驗組**和**對照組**之間，我們觀察到以下資料分析結果：

表 2 實驗組和對照組的問卷分析結果比較

指標	實驗組分數	對照組分數	Z 值	P 值	假設檢定結果
團隊活動中的合作傾向	4.425	4.312	8.5904	$p < 0.05$	顯著差異
相信小組能夠成功合作完成任務的程度	4.437	4.382	8.5904	$p < 0.05$	顯著差異
與團隊成員進行溝通時的能力	4.425	4.513	8.5904	$p < 0.05$	顯著差異
讓他人感受到重要性的能力	4.456	4.368	8.5904	$p < 0.05$	顯著差異
解決問題的能力	4.383	4.188	8.5904	$p < 0.05$	顯著差異
面對問題的意願	4.369	4.286	8.5904	$p < 0.05$	顯著差異
定期檢視自己是否達到目標的能力	4.265	4.123	8.5904	$p < 0.05$	顯著差異
檢視自己學到的內容是否足夠的能力	4.196	4.131	8.5904	$p < 0.05$	顯著差異
創造性思考能力	4.355	4.151	8.5904	$p < 0.05$	顯著差異
學習滿意度	4.258	4.060	8.5904	$p < 0.05$	顯著差異

推論統計結果：所有指標的比較均顯示出在統計上存在顯著差異（$p < 0.05$），這意味著實驗組和對照組在這些方面的表現有所不同。

表 3 大學 STEM 創意機器人課程實驗組學生在 LINE 社群學習平台系統的質性問卷數據參考

組別	**選修 STEM 創意機器人課程在實驗組中分組的學生團隊學習心得分享**	團隊分數
A 組團隊心得	今天的期末分組報告，讓我們充分了解到原來要設計規劃創立一個暑期機器人營隊課程活動，並不是如想像中那麼容易的，我們也藉由這次的報告了解到自己小組所缺乏的能力，之後我們也會繼續進步努力提升自己的。	92
B 組團隊心得	我們的期末機器人暑期營隊將為學生提供一個豐富的機器人課程學習體驗，包括機器人理論知識的學習和實際操作的機會。這個營隊將幫助學生深入了解機器人的應用和發展，並培養他們的創造力和解決問題的能力。我們將組成各種有趣的活動和挑戰，讓學生能夠在合作和競爭中不斷成長，而我們有信心與責任相信這個營隊將成為學生暑假的精彩回憶。	96
C 組團隊心得	今天是創意機器人課程的期末報告驗收，主題是「Micro: bit 程式設計機器人與 AR 互動暑期營隊」，我們這組選擇的母校是「金龍國小」，因為我們覺得它相較於其他學校競爭力會比較小，我們在做的過程中，意識到要開設暑期營隊絕對不是一件簡單的事情，有很多的因素需要去考量 ex：活動計畫、及收費標準...我們認為這是一個非常有意義的期末成果報告，或許之後真的可以成功開設機器人暑期營隊。	98
D 組團隊	我們期末成果報告的「EV3 樂高機器人社團」，規劃讓我們的國中學生充分發揮創造力和解難能力。透過機器人建置和編程，他們學會了設計和控	94

心得	制機器人，並在團隊合作中取得成就。這個暑期社團不僅激發了他們對科學和技術的興趣，還培養了創新思維和問題解決能力。	
教師教學研究流程	1➔課程融入戲劇學習，2➔建立專家斗篷職務分配團隊，3➔融入建立負責任文化，4➔讓小組學生融入在 STEM 機器人教育情境中，5➔每月設定不同的學習任務，6➔質性分析這些團隊的學習狀況，7➔終極目標:**讓學生學習知識變得具服務性與責任感** (在課堂中教師以質性記錄這些實驗組與對照組班級團隊的學習成效並且持續記錄觀察追蹤分析研究)	

本研究探討 STEM 機器人教育中戲劇學習和負責任文化對學生學習成效、創造力和解決問題能力的影響，特別關注 LINE 社群學習平台系統的作用。研究結果顯示，實驗組與對照組存在顯著差異，實驗組表現出較高的合作、溝通、解決問題、後設認知和創造性思考能力。這部分歸因於大學生 LINE 社群學習平台系統的使用。LINE 平台提供了互動性強、即時性高的學習環境，促進了學生更緊密地合作、分享資源並及時獲得反饋。這種特點使學生更積極參與學習，提升了自信心和問題解決能力。此外，LINE 平台加強了學生間的溝通和合作，有助於更好地理解課程內容並彼此學習。建議教師在角色探索中扮演更負責任的角色，建立積極的學習文化，深化學生的學習體驗，培養創新思維的知識工作者。實驗組對分組合作方式的學習成效表示滿意。總所上述，本研究證實了 LINE 社群學習平台系統在促進學生學習成效、創造力和解決問題能力方面的正面作用，並提出了教學實踐建議，有益於 STEM 教育的改革參考。

5. 研究結論

本研究致力於探討戲劇學習在 STEM 機器人課程中建立負責任文化對學生學習成效的影響。根據資料分析結果，我們得出以下結論：

1. **教學方式與學生互動**：老師運用專家斗篷思考法，創造了互動豐富、生動活潑的學習環境，讓學生感受到老師的熱情與責任，強調親近感，使學習更具吸引力。
2. **戲劇學習的應用**：學生積極評價戲劇學習的應用，認為這種獨特教學方式使課程更有趣、引人入勝，並能實際應用在解決企業問題，培養實用技能。
3. **團隊合作與溝通**：老師的教學方式促進了學生之間的團隊合作，強調小組模式培養了解決問題的技能，學生認為這種負責任文化是重要的收穫。
4. **自主學習與思辨能力**：學生回饋教學方式讓他們在輕鬆氛圍中進行自主學習，並培養了多元思考的能力，建立了對負責任文化的認知。
5. **應用於現實生活**：老師成功將教學內容與現實生活結合，使學生感受知識服務提供者的樂趣，有助於他們更清晰地理解學習內容，應用於未來生活和工作中。

本研究展示了教學方式的獨特性，以及戲劇學習和專家思考法在提升學生參與度、培養實用技能、建立負責任文化方面的積極影響。這些發現為改進教育提供了寶貴見解，同時填補了現有文獻的研究空白，對理論和實證研究做出了貢獻。然而，研究存在限制，主要體現在學校樣本較為單一，影響結果的一般化能力。未來研究應擴大樣本以提高代表性，並考慮採用多種測量方法，如增加觀察和訪談，以獲得更客觀、全面的資料。

本研究感謝國科會經費支持，計畫編號：MOST 110-2511-H-008 -004 -MY3；NSTC 112-2811-H-008-006。

參考文獻

Dorothy Heathcote & Gavin Bolton 著；鄭黛瓊、鄭黛君譯(2006)。戲劇教學-桃樂絲・希斯考特的「專家外衣」教育模式。台北市:心理出版社

Abdel-Hamid, A. and Metwally, A. (2003). Mathematical sense and its relation to the special creativity and academic achievement of the students of the faculties of education, Mathematics

Department, the third scientific conference "Teaching and learning mathematics and the development of creativity". Faculty of Education, Ain Shams University, Egypt.

Heathcoate, D. (2004). *The Mantle of the Expert: A system for learning through the active imagination and enquiry methodology.* Retrieved on March 25, 2016.

Chang, C. C., & Lai, C. L. (2011). An empirical investigation of the effects of LINE mobile learning on students' perception and learning performance. Educational Technology & Society, 14(4), 31-42.

Hung, H. T., Hsu, Y. C., & Rice, K. (2012). Integrating technology into problem-based learning: Effects on student engagement and collaboration. Journal of Educational Technology & Society, 15(4), 16-27.

Wang, M., Shen, R., Novak, D., & Pan, X. (2012). The impact of social annotation technology on students' learning outcomes: A structural equation modeling approach. Journal of Educational Computing Research, 46(4), 423-441.

Wong, J. G., Cheung, E. P., Chan, K. K., Ma, K. K., & Tang, S. W. (2012). Mental Health, Academic Advising, and Psychological Help-Seeking Among Undergraduate Students. Journal of College Counseling, 15(2), 115-128.

Heathcote, D., & Given, B. (2013). A selection of drama and education. Translated by Issa Bishara, Manal Issa, Rami Salama, Al Qattan Center for Educational Research and Development, Ramallah.

O'Neill, C., & Lambert, A. (2013). Drama for learning: Dorothy Heathcote's mantle of the expert approach to education. Bloomsbury Publishing.

Resnick, M. (2013). Lifelong kindergarten: Cultivating creativity through projects, passion, peers, and play. MIT Press.

Hew, K. F., & Cheung, W. S. (2014). Use of Web 2.0 technologies in K-12 and higher education: The search for evidence-based practice. Educational Research Review, 12, 47-64.

Liu, G. Z., & Cheng, S. C. (2014). Learning English through LINE in a blended learning environment. Journal of Educational Technology & Society, 17(3), 114-125.

Brown, M. (2015). The future of learning: Redefining education in the age of technology. Knowledge Works.

Johnson, L., Adams Becker, S., Estrada, V., and Freeman, A. (2015). NMC/CoSN Horizon Report: 2015 K-12 Edition. The New Media Consortium.

Kafai, Y. B. (2016). Connected code: Why children need to learn programming. MIT Press.

Bers, M. U. (2017). Coding as a playground: Programming and computational thinking in the early childhood classroom. Routledge.

Birkett, A. (2017). Mantle of the expert: Drama for learning. National Drama Publications.

Papert, S. (2017). Mindstorms: Children, computers, and powerful ideas. Basic Books.

Hmelo-Silver, C. E. (2018). Designing problems for learning across disciplinary boundaries. Journal of the Learning Sciences, 27(3), 305-329.

Prendergast, M. (2019). Drama in primary English teaching. Routledge.

Arman, I. M. (2019). Effectiveness of "Mantle of the Expert" In Creative Thinking Skills. International Journal of Science and Engineering Investigations (IJSEI), 8(86), 101-105.

Lin, C.-H., Hwang, G.-J., & Chen, C.-H. (2021). Robotics-based STEM education: A systematic review of research trends and issues. International Journal of Technology and Design Education.

中美 STEM 教育研究现状与热点的比较研究

A comparative study on the status and hotspots of STEM education research in China and the United States

陈博*，唐瓷
成都师范学院 网络与信息中心，四川成都，611130
*etchenbo@163.com

【摘要】 从国际比较的视角梳理中美两国关于 STEM 教育的相关研究文献，对 2013-2022 年间刊发于 CSSCI 和 SSCI 期刊中关于 STEM 教育的文献进行研究分析。研究发现：研究视角方面，国内研究视角较为单一，更多关注基础教育阶段，美国研究视角宽泛，关注不同学段 STEM 教育；研究内容上：美国更加突出对个体因素和社会因素的关照，如 STEM 教育过程中的公正问题、男女差异等。建议：一是注重 STEM 教育师资培育；二是凸显整合思想，做好 STEM 教育课程开发；三是关注个体因素在 STEM 教育过程中的影响及不同学段 STEM 教育的探索和实践；四是重视教育数字化转型背景下，STEM 教育助推学生核心素养的培养。

【关键词】 中美；STEM 教育；现状与热点；比较研究

Abstract： From the international comparative perspective, this paper reviews the relevant research literature on STEM education in China and the United States. The analysis focuses on the literature published in CSSCI and SSCI journals between 2013 and 2022. The findings reveal that: In terms of research perspectives, domestic studies have a relatively single focus, mainly on basic education, while American studies are broader and cover various stages of STEM education. In terms of research content, American studies highlight individual and social factors, such as fairness issues in STEM education and gender differences. Recommendations include: Emphasize teacher training for STEM education. Stress the importance of integration and develop STEM education curriculums. Pay attention to the impact of individual factors in STEM education and explore and practice different stages of STEM education. Give due importance to the role of STEM education in promoting students' core competencies in the context of educational digital transformation.

Key words： China and the United States; STEM education; status and hotspots; comparative study

1. 引言

项目来源：四川省高等学校人文社会科学重点研究基地·四川省教育信息化应用与发展研究中心项目：教育数字化转型背景下师范生数字素养职前职后一体化培育路径研究（项目编号：JYXX23-003）“乡村教师数字素养提升路径研究”（编号：SCXCJY2023D02）；成都师范学院 2022 年校级科研项目：师范院校智能教师队伍建设体系研究与实践（项目编号：CS22RGZN05）。

作者简介：陈博（1992-），男，四川剑阁人，硕士，讲师，研究方向为智能教育；唐瓷（1977-），男，四川邻水人，硕士，教授，研究方向为教育技术。

随着科学技术日益发展，人类社会对具备科学、技术、工程和数学素养的人才需求日益增加。在知识经济时代，创新思维和创造能力成为推动经济社会发展的关键因素。传统教育模式在培养学生的实践能力和创新精神方面存在不足，当今社会问题的解决需要跨学科的知识和技能。STEM 教育在这一背景下应运而生，STEM 教育更好助力学生适应未来社会发展，提高就业竞争力和综合素质，STEM 教育理念符合当今社会对人才的需求，也满足个人走向社会的内在需要。STEM 发展经历了 STS、STEM 和 STEAM 三个阶段的嬗变（张端,和继军和李小娟，2020）。我国高度重视 STEM 人才培养，1995 年我国提出科教兴国战略，2017 年我国发布《中国 STEM 教育白皮书》（中国教育科学研究院，2017）。“教育、科技、人才是全面建设社会主义现代化国家的基础性、战略性支撑”（习近平，2022），我国教育部于 2023 年 10 月召开全国中小学科学教育工作部署推进会，高度重视全民科学素质提高，培育人的创新思维和创造能力。美国为了增强国家科研实力，保持全球领先地位，美国国家科学基金会（NSF）对 STEM 人才培养进行了长期探索，通过培养具有较高科学与工程素养的 STEM 人才推动创新发展与重大科技关口的突破（陈翠荣&鲁智丹，2017）。提供高质量的 STEM 教育实践活动已成为世界范围内各个国家培养高层次人才的重要途径（李业平，2019）。当前，国内开展 STEM 教育如火如荼，美国对此研究相对成熟，其成熟经验不能照搬，但值得参考。研究拟通过对中美两国近十年关于 STEM 教育研究文献进行可视化分析，以期通过描述中美两国对 STEM 教育研究的热点和趋势，有助于我国更好的学习借鉴成熟经验，形成 STEM 教育在我国可持续发展态势，助力培养创新型人才。为此，本研究从中美两国 STAM 教育比较出发，试就以下研究问题进行探讨：一是两国 STEM 教育的差异性体现在何处？二是两国 STEM 教育关照的热点是什么？三是我国可借鉴的先进经验有哪些？

2. 研究设计

2.1 数据来源 基于研究目的，研究取自权威文献资源作为研究数据更具说服力。根据研究需要，本研究具体检索策略是：在“中国知网”进行“高级检索”，以“STEM 教育”OR“STEAM 教育”为检索主题词、限定期刊来源为“CSSCI 期刊来源”、检索时段为“2013 年-2022 年”进行期刊文献检索，共检索到 519 篇文献，剔除会议通知等无效文献 20 篇，累计得到有效文献共计 499 篇；在“Web of Science ™ 核心合集”中，以“主题：(STEM education)OR 主题：(STEAM education)，精炼依据:国家/地区：(USA)AND 文献类型：(ARTICLE)AND Web of Science 类别：(EDUCATION EDUCATIONAL RESEARCH)时间跨度：2013-2022，索引：SSCI。共检索到文献 954 篇。剔除 47 篇不相关论文（如以 stem cell 为主题的相关文献等），最终以 907 篇为本研究的总体英文文献样本。

2.2 研究方法与工具 研究主要基于共词分析与社会网络分析支持的对比分析，展示中美两国关于 STEM 教育的主要关切点和二者之间的差异。研究工具：文献题录信息统计分析软件 SATI 3.2、数据统计分析工具 SPSS 22.0、社会网络分析工具 UCINET 6.0。

3. 文献计量分析

对于已正式刊发的中美两国 STEM 教育研究文献进行整理分析，两个数据库相关文献年份分布情况如图 1 所示。

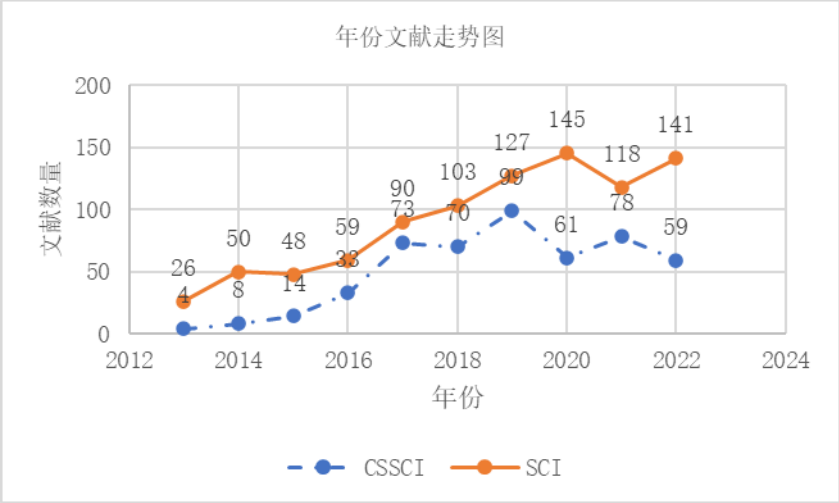


图 1 中美两国 STEM 教育研究文献的年份走势图

由图 1 可知，从文献发表数量上看中美两国在该领域存在明显差异。国内首篇关于 STEM 教育的 CSSCI 文章于 2010 年发布在《开放教育研究》杂志，该研究来自国内学者顾小清、谢明颖等学者论述美国国家科学基金教育技术项目支持美国 K-12、STEM 教育发展（顾小清,谢明颖和付世容，2010）。美国 STEM 教育研究开始时间较早，并有相当数量的研究文献刊发，美国于 1986 年首次提出了“STEM 教育”概念（National Science Board，1986）（高云峰，2016）。从图 1 看，国内 STEM 教育研究整体呈上升趋势。2014 年至 2020 年增长态势明显，大致原因可归结为政府发出“大众创业、万众创新”的号召。美国 STEM 教育研究整体也呈上升趋势，2013 年至 2017 年上升幅度最大，通过文献分析其原因大致如下：美国支持开展 STEM 教育的政策推进力度大，如 2011 年，罗德岛设计学院（RISD）向美国国家科学基金会提交了题为《连接 STEM 和 STEAM：建立艺术/科学教学新框架》的报告（Rhode Island School of Design，2020）。2012 年，科学、工程、艺术、设计网络（SEAD）在《超越生产力：信息、技术、创新和创造力》基础上发表了白皮书，总结了过去 10 年艺术与科学的合作成果（SEAD，2020）。2013 年，国会 STEAM 核心小组指出“改变教育词汇，认识到艺术和科学——以及它们的交叉——给美国未来几代人带来的好处”（Congressional STEAM Caucus，2020），促进了 STEM 教育在美国的进一步发展。综上所述，从侧面反映出，国家教育的顶层设计直接或间接的促进和推动了该国 STEM 教育的快速发展。从走势图上看，中美两国分别于 2019 年、2020 年到达研究顶峰，随后有所下滑，其隐藏的深层缘由或因受到“COVID-19”对于全球教育的重大影响，由于 STEM 教育强调实践性、操作性，疫情期间，时空分离下的实践操作活动难以展开，导致相关研究有所滞后。

4. 高频关键词的确定

高频关键词可以反映出“STEM 教育”研究集中的范围，通过计量分别得到 CSSCI 和 SCI 数据库中有关的关键词分别是 1196 个、2559 个，词频分别是 2137 次、5041 次。从关键词的数量可看出，美国 STEM 教育更具有宽泛性、多样性。两图所列关键词是在 CSSCI 和 SSCI 数据库统计中，统计数量总数排在前 20 位的关键词（需要说明的是：为了使研究数据具有说服力，本研究后续研究使用 CSSCI 高频关键词为累计出现 5 次及以上，共计 39 个；SSCI 高频关键词累计出

现 10 次及以上，共计 48 个）。具体表现如下：

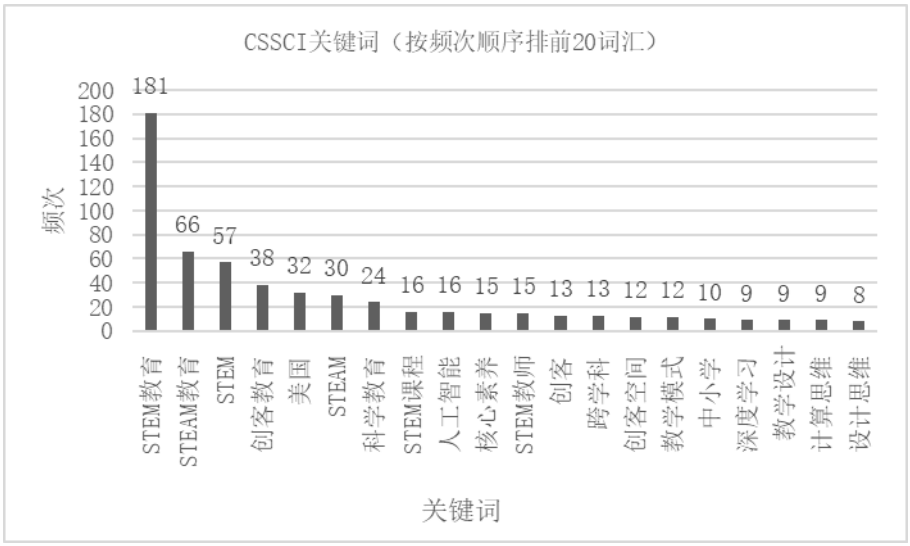


图 2 CSSCI 高频关键词（部分）

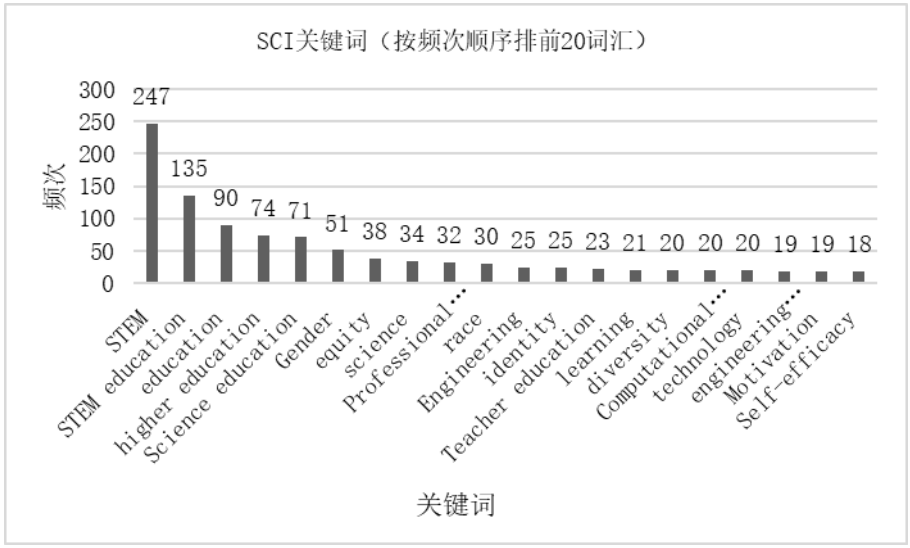


图 3 SSCI 高频关键词（部分）

由图 2、图 3 可知：（1）两个数据库中重叠的高频关键词有 STEM、STEM 教育、科学教育、计算思维，计算思维是面向智能时代的基本思维，中美两国学者对于 STEM 教育研究都注重在科学教育过程中培养学生的计算思维。（2）两个数据库中关键词存在差异：CSSCI 数据库中“创客教育”、“创客”、“创客空间”等词高频出现。“创客”源自 20 世纪 60 年代的西方 DIY 文化（王佑镁，钱凯丽等，2017），创客教育在我国的兴起，源于对美国创客教育引入和探索（张志新,韩浩文和刘静，2019），创客教育是一种旨在培养个体的 DIY 能力、分享精神和创造思维（王佑镁，王晓静，2015）。我国 STEM 教育开展更多开始于吸收国外经验，开展“创客”教育；美国 STEM 教育过程中，更多关注个性特征及社会问题，如性别（gender）、公正（equity）、个性（identity）、多样性（diversity）、自我效能感（self-efficacy）、改革（reform）等，美国的 STEM 教育更加重视个体因素及社会因素在 STEM 教育中的影响,如美国学者认为，计算机科学和工程是男性主导的领域，呈现社会孤立、对机械的强烈关注和与生俱来的天赋等特征，导致女生对这些领域的兴趣远低于男生（Cheryan S, Master A, Meltzoff A N, 2015）；同时也会关注教育改革、师资培育等方面，如美国教师教育学院协会联

合 50 所州立大学展开 STEM 教师培训计划（Washington，D.C，2007）。（3）需要指出的是,“美国”作为关键词出现在中国 STEM 教育研究文献的高频关键词中，且总频次高达 32 次，更加说明美国 STEM 教育对我国 STEM 教育实践的影响。综上所述：无论是我国开展的 STEM 教育、创客教育还是美国开展的 STEM education、Science Education，都是教育的活动，而教育是培养人的社会活动，教育要重视人的发展，同时要关注社会发展对 STEM 人才的需要。相比之下，美国 STEM 教育更加关注个体因素，关注学生个性发展，在教育过程中能积极发挥学生的主观能动性，同时关注社会因素对 STEM 教育的影响；我国 STEM 教育研究更多的是重视与“创客”结合，整体上对个体因素的关照还不够；研究聚集点趋向宏观，缺少人的因素在 STEM 教育的实证研究，如不同性别，不同学段学生在接受 STEM 教育过程中是否存在显著差异；在对美国经验借鉴吸收的同时，需要开放视野，结合我国实际开展 STEM 教育实践。

5. 高频关键词分析

5.1 因子分析 通过因子分析发现,两个数据库中因子的累计方差解释贡献率分别是 64.210%和 63.359%。

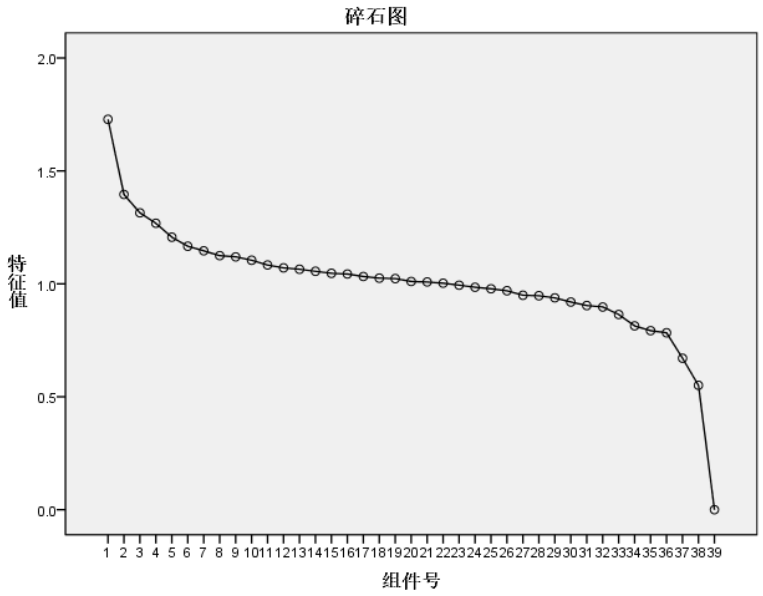


图 4 CSSCI 因子碎石图

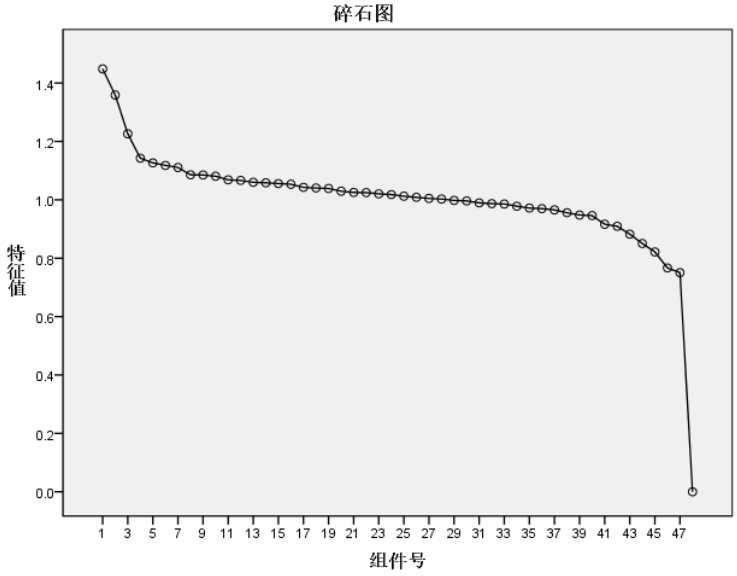


图 5　SSCI 因子碎石图

图 4 和图 5 呈现的是两个数据库相关的碎石图：(1)图 4 可知，CSSCI 数据库中因子分析的前 5 个因子类别较为清晰，对应结构矩阵前 5 个因子分别如下：STEM 教育、STEAM 教育、STEM、创客教育、美国；(2)由图 5 可知，SSCI 数据库中前 7 个因子相对比较清晰，对应结构矩阵，前 7 个因子分别为：STEM、STEM education、education、higher education、science education、gender、equity。（3）CSSCI 数据库中因子特征值较高，研究内容较为集中，说明我国 STEM 教育研究的内容趋向集中稳定；SSCI 数据库中因子特征值相对较低，说明美国 STEM 教育研究内容相对更具宽泛性。碎石图的结论符合上述对高频关键词的观察判断，即美国更加倾向于从个体教育本身出发，更加追求公正平等的教育，我国更加注重吸收经验，着重培养“创客”。

5.2 高频关键词聚类分析 聚类结果反映了关键词之间的亲疏关系，从而进一步反映中美两国 STEM 教育研究主要取向。两类数据库高频词聚类分析如表 1 所示。

表 1 CSSCI 与 SSCI 高频关键词聚类表

序号	CSSCI 类别	CSSCI 成分	SSCI 类别	SSCI 成分
1	创客培养	创客、创客空间、创客教育	差异性	Gender、race、equity、STEM、higher education、graduate education、diversity
2	师资培育	STEM 教师、教师教育、教师专业发展	学习素养	Faculty、persistence、undergraduate、active learning、mentoring
3	课程开发	STEM、STEAM、课程设计、基础教育、人工智能教育、教育政策、核心素养	师资培育	Science education、mathematics education、development、professional development、teacher education、STEM education

4	整合设计	元分析、跨学科学习、知识图谱、跨学科教育、深度学习、跨学科整合、跨学科	整合设计	Computational thinking、integration、K-12、STEM integration、engineering education、engineering design
5	美国模式	科学教育、创新能力、美国STEM教育、STEM课程、设计思维、工程设计、STEM教育、美国、教学模式、技术教育	STEM内涵	Engineering、technology、science、mathematics、education

由上表所知，CSSCI数据库和SSCI数据库中关于STEM教育的研究整体上都可划分为5类。（1）两个数据库中重复的类别主要集中在师资培育及整合教育。可以看出中美两国都对STEM教育教师培养及学科整合设计进行深入的探究，优良的STEM教师群体是实现STEM教育高质量的关键；跨学科性本身也是STEM教育与其他学科领域的区别之一，STEM是一个整体概念，是为了强调通过对各类素养的整合以培养学习者融会贯通的跨学科能力（Association of American College and universities,2011，2011）。（2）从教育阶段上看，中美两国存在明显差异，我国的STEM教育主要研究集中在对基础教育阶段学生的培养，而美国的STEM教育，可以看到在K-12、高中教育、大学教育阶段都有所涉及，美国的STEM教育覆盖学段更为广泛。（3）中国的STEM教育研究主要聚焦在创客培养、课程设计以及探索美国STEM教育模式；国内STEM教育一方面吸收国外先进教育理念模式，同时又结合自身培养本土需要的STEM人才；而美国STEM教育，关注差异性、学习素养、STEM教育内涵等。可以看出，美国STEM教育更为成熟，更多关注STEM教育环境下，不同学段，不同因素影响下STEM教育开展等具体问题。

5.3 多维尺度分析 将两个数据库的高频关键词相异矩阵导入SPSS22.0中，进行多维尺度分析，结果分别得出图6、图7研究热点结构图。

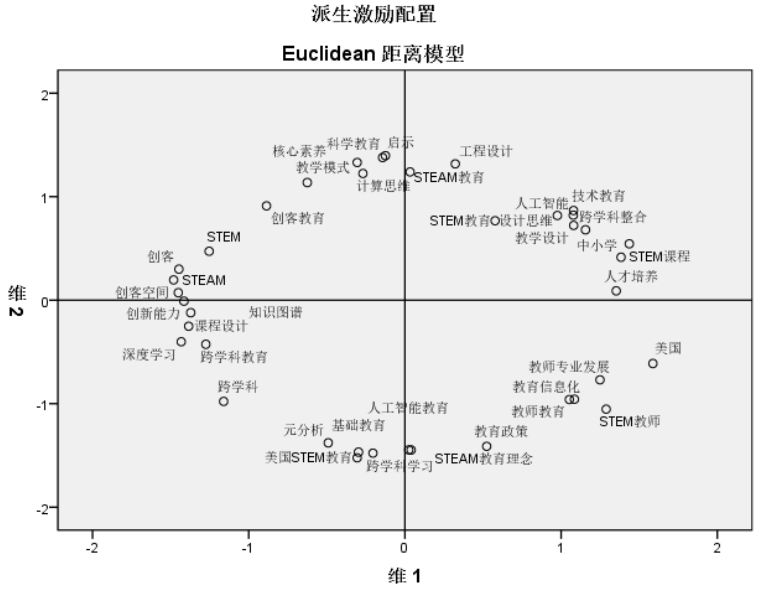


图6 CSSCI中研究热点结构图

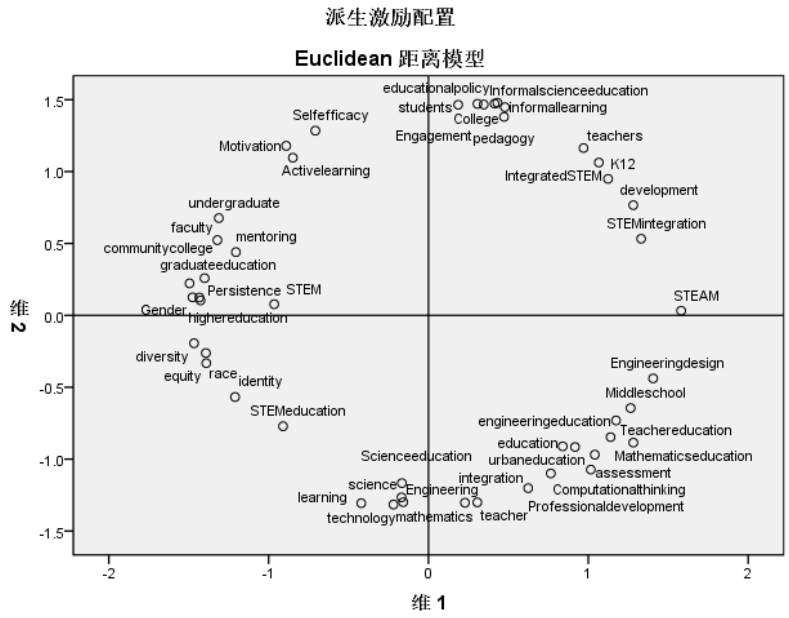


图 7 SSCI 中研究热点结构图

由上图可知 CSSCI 中研究结构可以分为课程开发、思维养成、创客培养、师资培育四类。SSCI 中研究结构可以分为 K-12 STEM 课程建设、学习素养、个体差异、思维养成四类。采用不同的数据分析方法，研究不同变量之间的关系和影响，进一步增强了研究的说服力和科学性。

5.4 关系结构分析 为进一步了解两国 STEM 教育研究热点，研究运用社会网络分析法，从中心性及关键词联系分析处出中美两国关于 STEM 教育的研究热点及其相互关系。

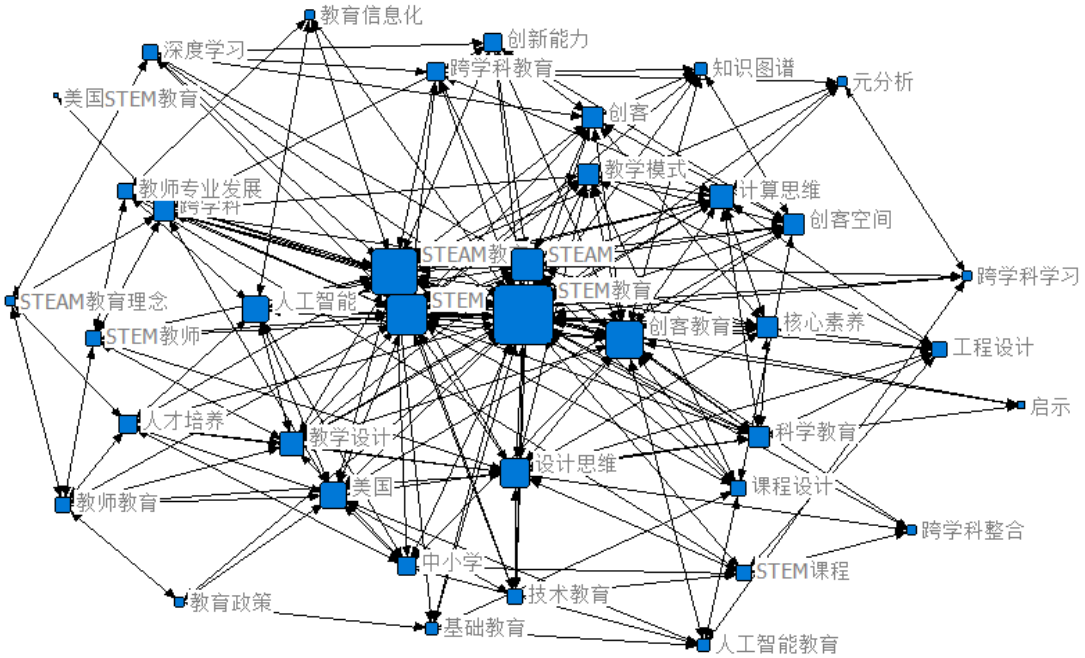


图 8 CSSCI 研究热点知识图谱

由图 8 可知，在 CSSCI 数据库中，STEM 教育、创客教育、STEM、STEAM 教育、创客等关键词处在共词网络中间，其周围联系相对密集，说明我国 STEM

教育研究热点主要集中在这些中心节点所构成的领域。图中不难看出教育政策、教育信息化、创新能力等词周围联系不够密切，但有少许联系，说明我国在这些节点所构成的领域研究稍显匮乏。理性思考，这些研究比较边缘化的研究部分，正是我国目前应该重视的，STEM 课程作为学校培养 STEM 人才的重要培养方式，STEM 教育政策、教育数字化转型背景下的技术的发展对 STEM 教育的影响等，对这些方面的研究是我国 STEM 教育开展必须要有所考虑的，需要深入研究。随着 STEM 教育不断深入发展，相关矛盾日益凸显，需要对这些边缘化的研究领域进行深入研究，可以说这些研究薄弱点也将处于新兴发展态势，是 STEM 教育研究的前沿，有待我国学者进一步加深研究与探讨。

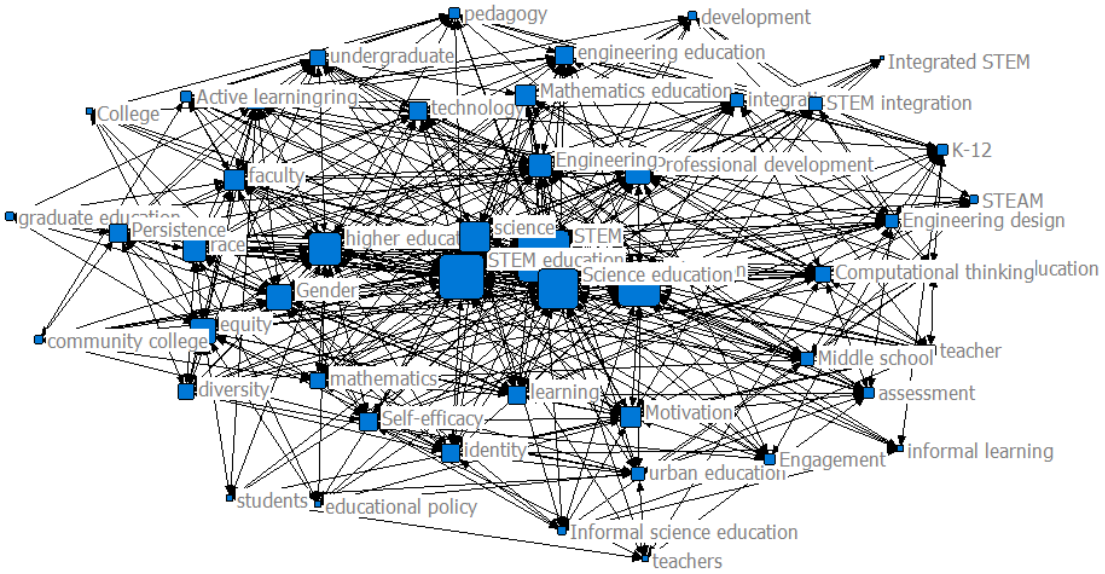


图 9 SSCI 研究热点知识图谱

由图 9 可知，在 SSCI 数据库中，STEM、STEM education、Education、Science、science education、gender、Technology 等关键词周围联系较为紧密，这些关键词构成的领域，是美国 STEM 教育研究的集中领域。美国 STEM 教育研究似乎更为具体，更关注个性化培养，关注 STEM 教育带来的系列社会问题。美国 STEM 教育关注性别（gender）、公正（equity）、自我效能感（self-efficacy）、毅力（persistence）等人的个体特征在 STEM 教育中的影响及差异。不难发现该图周边的一些关键词，例如 assessment、Integrated STEM 等词周围联系相对稀疏，表明这是美国当前 STEM 教育研究中相对薄弱的研究点，但是可以侧面反映，美国已经开始关注 STEM 教育引来的一些问题，例如如何合理评价，如何保证 STEM 教育各部分协同开展等。

6. 研究结论与建议

通过对比分析，中美两国 STEM 教育研究既有研究内容上的相似之处，也有研究关注的个别差异。整体看来，美国 STEM 教育研究相对宽泛，中国 STEM 教育研究更为聚焦，研究较为相对单一。中美两国对于 STEM 教育研究差异较大，经过上述研究比较发现主要体现在研究关注点上，国内研究主要集中在创客培养、师资培育、课程开发、整合设计以及美国 STEM 教育模式；美国 STEM 教育研究主要集中差异性、学习素养、师资培育、整合设计、STEM 内涵。整体上看，美国 STEM 教育更加关注人的因素以及所带来的社会问题，研究关注探

究人的因素在 STEM 教育过程中的影响，例如学生自身的心理特质、性别差异、个性发展等。对比出发，借鉴美国 STEM 教育研究经验，我国需要持续关注 STEM 教育师资培育、课程开发、整合设计、同时要遵循教育规律，育人为本，关注不同学段、不同性别等背景下如何更好的开展 STEM 教育；结合教育数字化转型背景，还要关注以人工智能为代表的新一代信息技术的影响下，如何更好的开展国内 STEM 教育。

建议：一是注重培养 STEM 师资，一方面我国需要进一步探究 STEM 教师需要具备的基本素养，从国家层面明确制定 STEM 教师的准入机制及培养计划，另一方面，关注 STEM 未来教师的培养，充分发挥师范院校优势，将研究与实践相结合，促进师范院校重点培育 STEM 教师；二是加强 STEM 教育课程开发，凸显跨学科、整合思想，STEM 课程注重知识的整合和可操作性，尤其是技术与工程的学科性质，课程开发过程中要以课程目标为导向，体现学科应用性，同时可通过跨学科、项目化，综合实践环节体现 STEM 教育思想；三是我国 STEM 教育研究应该拓展研究内容范围，更新研究方法，加大研究对象范围，关注学生差异性，如性别、地域、学习风格等多重个体因素和社会因素在 STEM 教育过程中的影响，同时关注我国职业教育、高等教育、继续教育等领域 STEM 教育的探索和实践；四是教育数字化转型背景下，基于新一代信息技术，创设真实情境，创新教学方式，加强学科间联系，构建多元评价体系，紧密围绕我国学生核心素养培育，最终实现科技创新人才的培养。

参考文献

习近平.高举中国特色社会主义伟大旗帜为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告［M］.北京：人民出版社，2022.

中国教育科学研究院.《中国 STEM 教育白皮书》（精华版）（Z）.2017,2.

王佑镁，钱凯丽，华佳钰等.触摸真实的学习：迈向一种新的创客教育文化——国内外创客教育研究述评[J].电化教育研究，2017(2)：34-43.

王佑镁，王晓静，包雪.创客教育连续统：激活众创时代的创新基因[J].现代远程教育研究，2015(5)：38-46.

张志新,韩浩文,刘静.美国中小学创客教育特点及对我国的启示[J].当代教育论坛,2019（06）:98-107.

张端,和继军,李小娟.美国 STEM 教育的嬗变[J].自然辩证法通讯,2020,42(11):94-100.

李业平.STEM 教育研究与发展:一个快速成长的国际化领域[J].数学教育学报,2019,28(3)：42-44.

陈翠荣,鲁智丹.美国联邦政府对科学与工程领域研究生资助的现状及特点分析[J].学位与研究生教育,2017,(12):71-77.

（%!）：6%$66

顾小清,谢明颖,付世容.美国 NSF 教育技术项目分析[J].开放教育研究,2010,16(04):96-107.

高云峰.创客与 STEAM 教育结合的实践[J].力学与实践，2016，38(1)：74-77.

American Association of Colleges for Teacher Education. Preparing STEM teachers: the key to global competitiveness[R].Washington，D.C.：AACTE，2007.

Cheryan S, Master A, Meltzoff A N. Cultural Stereotypes as Gatekeepers: Increasing Girls' Interest in Computer Science and Engineering by Diversifying Stereotypes[J].Frontiers in Psychology,2015(06):49.
Congressional STEAM Caucus.'Events'[EB/OL].http://stemtosteam.org/events/congressional-steam-caucus/. 2020-5-7.
National Science Board. Undergraduate Science, Mathematics and Engineering Education,1986.
Rhode Island School of Design. 'Bridging STEM to STEAM: Developing New Frameworks for Art/Science Pedagogy'[EB/OL].http://cjvrose.com/wp-content/uploads/2012/03/stem-to-steam-report.pdf. 2020-6-22.
Sciences, Engineering, Arts, and Design (SEAD). 'SEAD: White Papers''[EB/OL].https://sead network.wordpress.com/. 2020-6-23.
Susan Rundell Singer, STEM Education: Time for Integration（Association of American College and universities,2011）,https://www.aacu.org/publications-research/periodicals/stem-education-time-integration.

在线创客社区中的学习者学习情感研究

--基于文本情感分析的视角

A study of learner learning emotions in an online maker community

-- From the perspective of text sentiment analysis[1]

叶龙俊[1*]，单俊豪[1]
[1]南京邮电大学教育科学与技术学院
*b21150208@njupt.edu.cn

【摘要】 在线创客社区作为创客的创意讨论与生成的新场景，创客们在讨论中的学习情感研究对提升在线讨论和协作质量具有重要意义。本文选取了中国最大的开放式在线创客社区作为研究样本，基于文本情感分析的视角，对13个帖子的1284条回复进行情感分类编码，运用社会网络分析法与内容分析法来深入分析讨论中学习者的学习情感的共性特征与差异。结果表明，发帖人的支持程度并不直接影响主题帖整体情感分布，但却在积极地解决学习者负向学习情感方面发挥着关键作用；不同情感倾向的主题帖呈现出规律性情感表征。

【关键词】 学习情感分析；创客教育；在线创客社区；社会网络分析

Abstract: *As a new scene for creative discussion and generation of makers in online maker community, research on the learning emotion of makers in discussion is of great significance to improve the quality of online discussion and collaboration. In this paper, the largest open online maker community in China is selected as the research sample. Based on the perspective of text sentiment analysis, 1284 responses from 13 posts are classified and coded by emotion. Social network analysis and content analysis are used to deeply analyze the common characteristics and differences of learners' learning emotions in the discussion. The results show that the support degree of the poster does not directly affect the overall emotional distribution of the topic post, but it plays a key role in actively solving the negative learning emotion of learners. The theme posts with different emotional tendencies showed regular emotional representations.*

Keywords: Learning emotion analysis,Maker education,Online maker community,Social network analysis

0.引言

“创客”在英文里译为“Maker”，特指酷爱科技、热衷实践的个人，亦可指其群体，常以分享创意和交流思想为乐（丁大琴，2015）。创客空间是一种全新的开放式平台，提供创新创意分享的空间，配备生产制造物件所需设备和资源，支持创客将创意灵感转化为实物（王佑镁和叶爱敏，2015）。随着互联网技术的发展，各类虚拟社区、在线讨论组大量涌现，整合在线讨论功能的在线创客社区也应运而生，其有效拓宽了创客的创意讨论与生成场域。已有相关研究表明讨论发帖人在讨论中扮演了三个关键角色:创意分享者、资源提供者和问题解决者/建议提供者，但他们未能在改善最高级别的协作过程中发挥重要作用（Shan& Wang，2021）。然而，在线创客社区中，创客们在讨论中的学习情感分析却很少被提及与研究，这也成为本研究的出发点。创客之间的讨论和协作质量是构建在线创客社区良好学习环境的关键在所，对于新理解和创造性思想的产生和培育也至关重要（Anderson，2012）。深入探索

[1] 本论文受国家级大学生创新训练项目《新工科教育视域下大学生设计思维混合式教学模式建构与实践研究》（项目编号：202310293059Z）的支持，受2021年江苏省教育科学“十四五”规划青年重点课题“中职教师信息技术应用能力精准测评及提升微策略研究”与2022年江苏省高校哲学社会科学一般课题“微认证驱动的高校本科课程评价变革研究”支持

在线创客社区协作中创客之间的学习情感，对于发现不同主题帖之间的共性，以及创客社区中在线讨论过程存在的问题，并进一步优化在线创客社区交流环境起着至关重要的作用。

1.文献综述

1.1. 在线学习社区

在线学习社区（Online Learning Community）最早源自于英国学者 Howard Rheingold（1993）提出的定义：“当大量的人怀着饱满的情感持续长久地进行公开或半公开交流,以期望通过在线社交网络形成个人的关系网时，在网络中所出现的社会集合体”。在线学习社区，又称虚拟学习社区（Virtual Learning Community）、网络学习共同体等，它是依托互联网或移动互联网所创建的一种学习环境（董利亚和冯锐，2016）。多元技术在线上教育平台的应用使得在线学习社区成为有共同学习意愿的学习者聚集在一起建构知识、交流情感的有效学习环境。在线学习社区为学习者交流显性知识和隐性知识,为学习者自由交流观点,为正式的和非正式的学习小组建立一个集成的学习环境（甘永成和王炜，2005）。

在线创客社区是在线学习社区的典型代表，它允许来自不同地点的成员在线交流和协作。在线社区中存在着三个关键要素，即技术、学习任务和社会互动（Carabajal& Gunawardena，2003）。在线创客社区作为学习者展示和实践技能的平台，学习者可以将在在线学习社区学到的知识应用到实际的创客项目中，并在创客社区中与其他创客分享和交流，获得反馈和指导。在线创客社区可以为创客提供学习和知识获取的平台。创客们可以通过在线学习社区学习各种技术和创新知识，掌握必要的技能和理论基础。

1.2. 学习情感研究

互联网的迅猛发展催生了在线学习社区的兴起,学习者由信息的被动接收者逐渐演变为积极的信息发布者。他们不仅关注他人对特定主题或学习资源的评论和观点，还更愿意分享自己的看法、意见，以及明确表达个人态度。目前，对于在线创客社区中学习情感的讨论相对较为有限。尽管这些社区是学习和创造的理想场所，但多数关注点仍然集中在技术、项目合作和创意实践上，较少研究者从在线讨论质量的视角出发关注学习者在在线创客社区讨论中的情感体验。学习者在在线创客社区的讨论中可能经历着各种情感，如开心、兴奋、合作的充实或沟通的困惑，然而，这些方面往往没有得到足够的研究关注。

文本情感分析，又称意见挖掘，是指对带有情感色彩的主观性文本进行采集、处理、分析、归纳和推理的过程（洪巍和李敏，2019）。文本情感分析的方法主要包括基于情感词典与规则的方法、基于传统机器学习的方法、基于深度学习的方法以及多策略混合的方法（钟佳娃，刘巍，王思丽和杨恒，2021）。在缺乏大量训练数据集的情况下，基于词典与规则的方法相对能取得较好的分类结果且易于理解。本研究选择采用基于词典与规则的方法对在线创客社区中讨论的学习情感进行分析。在缺乏大量训练数据集的情况下，这一方法相对相对能取得较好的分类结果且易于理解。通过词典的积极、消极情感词汇以及事先设定的规则，本研究能够较为准确地捕捉社区成员在讨论中表达的情感倾向。

2.研究设计

2.1. 案例选取

DFrobot 作为中国最大的创客教育公司之一，主要服务于机器人、智能产品和科创教育领域，其中的在线创客社区是一个面向所有创客的开放社区，致力于帮助创客实现他们的创造力。迄今为止，DFrobot 在线创客社区汇聚了七万多名来自全国各地的创客，拥有近三十万个主题帖。

本研究选取了来自“Arduino”论坛中讨论较为激烈的帖子。为了确保帖子的发帖人与参与者的讨论过程是明显的，每个主题都必须包含不少于 60 条回复，并且不少于 30 位成员参加了讨论。最终，研究共选取 13 个帖子的 1284 条回复作为样本。示例帖子的详细信息如表 1 所示。

表 1 示例帖子的详细信息

帖子标题	回复数
穹顶之下-DIY 空气质量检测+净化器	102
自制“谷歌眼镜”低成本	144
基于 3D 打印+Arduino 的 LED 时钟	95
Arduino 指纹电磁锁	100
用 arduino 制作气体磁悬浮盆景	93
launchpad 音乐打击垫	88
使用 Arduino UNO 制作的偷钱喵储钱罐	115
小白无障碍制作 InMoov 机械手	165
打造廉价 axidraw 写字机（3D 打印版）	89
Arduino 教程 04 互动交通信号灯「DFR0100」	82
激光固化 3D 打印机	79
基于 Arduino UNO 的办公室测温仪	62
#情人节#--像素范 ， 蓝牙播放器	70

2.2. 研究方法

2.2.1. 社会网络分析法（SNA）

社会网络分析用于描述和测量行动者之间的关系或通过这些关系流动的各种有形或无形的东西如信息、资源等（张存刚，李明和陆德梅，2004）。该方法提供了与应用其他研究方法不同的在线学习互动关系模式与结构的一些新的和重要的信息（王陆，2009），越来越多的学者开始使用社会网络分析去研究在线学习环境中学习者交互关系问题。在社会网络分析中，本研究参照胡勇（2006）SNA 的要素与网络学习要素的对比表，将发帖人与参与者用“节点”来表示，帖子回复的数量用“度数”来表示，响应帖子发帖人或其他参与者的回复数称为该参与者的出度（Outgree），而接收到他人的回复数叫做该节点的入度（Indgree）。由此产生的各项指标，如中心性、密度等来显示讨论的集中趋势，深入发掘在线创客社区讨论中的各类信息与用户情感之间的关系。

表 2 SNA 要素与网络学习分析要素对照表

	SNA	网络学习分析
1	节点	网络学习的参与者
2	边及边的方向	学习者帖子流向
3	边的强度	发送或接受的帖子数
4	节点的出度(out-dgree)	某个学习者响应他人的帖子数
5	节点的入度(in-dgree)	某个学习者接受的帖子数

2.2.2. 内容分析法（CA）

内容分析法是一种对研究对象所发表的文本型内容进行深入分析，透过现象看本质的科学方法（邱均平，余以胜和邹菲，2004）。内容分析的对象是各种用于交流的信息内容，任何有交流价值的信息都能成为内容分析的对象，因此其应用领域十分宽广（宋振峰和宋惠兰，

2012）。本研究通过内容分析法找出帖子发帖人和参与者在每条回复背后的情感状态。在深入分析每条回复的所蕴含的情感时，本研究采用了刘鲁川等（2015）在探究社会化媒体用户的情感体验中的16种情感为分类框架，其中积极情感共9种，消极情感共7种，并对每一种情感进行情感解释与编码，具体编码如下方表2所示。该情感分类框架使用同义词辞典将部分意义相近的情感词汇进行归类汇总，并在实验的进行中们选取前 16 种出现率高且能涵盖大部分用户情感作为后续实验开展的情感分类依据。

表 3 编码表

情感类别	具体分类	情感解释	编码
积极情感 (positive emotions)	开心	领悟、得到肯定后进行感谢	PH
	有趣	对内容感兴趣、有学习倾向	PI
	兴奋	对内容表示激烈的正向评论	PE
	充实	从内容中学习到充足的知识	PF
	轻松	自如地解答困惑	PR
	新奇	鲜有见到此类内容的机会	PN
	惊喜	对内容表示惊叹	PS
	喜欢	对内容表现出赞许、提出建议	PL
	自由	畅所欲言	PD
消极情感 (negative emotions)	烦躁	认为内容存有错误	NW
	无趣	认为内容无任何价值和意义	ND
	沮丧	实践过程中遇到难以解决问题	NB
	愤怒	对内容进行完全否定与抨击	NI
	悲伤	因外部因素无法成功验证内容	NS
	焦虑	实践时无法达到预期效果	NA
	困惑	对部分内容产生疑问	NC

2.3. 数据分析

2.3.1. 情感统计

依据编码表中对于情感的解释，对样本中所选取13个帖子的1284条回复逐条进行编码，分别统计各情感在其帖子中的数量及其占比，并统计总积极情感与总消极情感的占比，得出下表。

表 4 情感统计表

情感类别	具体分类	编码	帖子1	帖子2	帖子3	帖子4	帖子5	帖子6	帖子7	帖子8	帖子9	帖子10	帖子11	帖子12	帖子13
积极情感(positive emotions)	开心	PH	5	6	3	2	2	0	5	2	8	1	4	5	3
	有趣	PI	9	16	10	10	21	7	17	14	7	7	11	1	4
	兴奋	PE	5	22	15	3	13	4	14	4	5	1	12	0	2
	充实	PF	0	4	3	5	0	0	1	0	0	2	0	1	1
	轻松	PR	20	7	9	11	6	19	8	41	15	16	10	17	18
	新奇	PN	0	3	0	1	3	3	0	2	0	0	0	1	0
	惊喜	PS	0	4	3	1	0	0	0	3	1	0	0	0	1
	喜欢	PL	26	53	14	33	15	22	38	27	11	20	18	9	21
	自由	PD	10	13	5	3	10	7	7	9	3	4	8	2	1
消极情感(negative emotions)	烦躁	NW	1	0	0	1	0	0	1	2	5	4	0	3	0
	无趣	ND	6	3	0	0	1	0	1	2	2	1	1	1	1
	沮丧	NB	0	2	4	1	1	1	7	8	4	1	0	1	1
	愤怒	NI	0	0	0	0	0	0	0	0	0	1	1	0	0
	悲伤	NS	0	0	0	0	0	0	1	1	0	0	0	1	0
	焦虑	NA	2	0	3	1	8	2	1	0	1	4	0	1	0
	困惑	NC	18	11	26	28	13	23	14	52	27	20	14	19	17
积极情感(p)总计			75	128	62	69	70	62	90	100	50	51	63	36	51
消极情感(n)总计			27	16	33	31	23	26	25	65	39	31	16	26	19
总计			102	144	95	100	93	88	115	165	89	82	79	62	70
积极情感百分比			73.53%	88.89%	65.26%	69%	75.27%	70.45%	78.26%	60.61%	56.18%	62.20%	79.75%	58.06%	72.86%

以帖子 1 为例，每个具体分类后的数字代表该具体情感在帖子中出现的次数。在帖子 1 中有 5 次提到开心（PH）、9 次提到有趣（PI）、5 次提到兴奋（PE）等积极情感总计为 75（开心+有趣+兴奋+...），消极情感总计为 27（烦躁+无趣+沮丧+...），而帖子 1 中总计出现的情感相关的次数为 102。最后一行提供了积极情感在总情感中的百分比，计算方式是积极情感总计除以总计次数，乘以 100。在这个例子中，积极情感百分比为 73.53%，反映了帖子 1 中积极情感在整体情感中所占比例。

2.3.2. 社会网络的构建与可视化呈现

UCINET（University of California at Irvine Network)是加州大学欧文分校的一群网络分析者开发的一个社会网络分析软件。UCINET 集成了可对一维与二维数据进行统计和可视化分析的 NetDraw，在网络结构可视化、用户交互性探索等方面都有不错的效果（邓君，马晓君和毕强，2014）。本研究使用 UCINET 作为每个帖子社会网络构建与分析的工具。对每个帖子中的用户进行数字代号排序，将发帖人记为 1 号用户，其余参与者按发帖时间顺序依次命名为 2、3、4...依据用户在帖子中的对他人回复数构建二维数据矩阵，运用可视化分析绘制出整个帖子的网络，并运用 UCINET 进行相关数据的分析。以帖子 1 为例，通过对帖子的网络进行绘制，可以清晰地看出这个帖子中发帖人与参与者的交互情况，其中由发帖人参与的双向箭头数为 8 个。

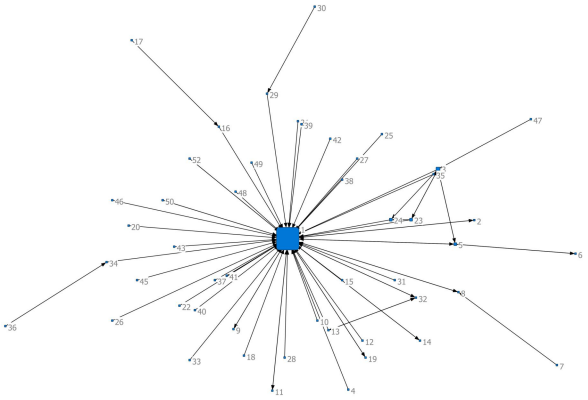


图 1 社会网络图

2.3.3. 关系探究

在 CA 中，本研究采用了“总分”的研究思路。首先，进行整体分析，对所有帖子进行了综合性的情感分析，以获取对在线创客社区讨论整体趋势和讨论情感的全面认识。在这一初步分析的基础上，进一步将高积极情感帖子和低积极情感帖子进行了分离和深入分析。通过这种分层次的研究方法，本研究更细致地了解了不同情感极性的帖子在内容、情感分类以及社区互动方面的差异，为深入挖掘社区成员情感体验提供了更为精准的研究基础。

在 SNA 中，本研究关注双向箭头和单向箭头的占比，通过分析这些箭头的分布，能够深入了解用户之间的相互影响关系。此外，研究发帖人的支持程度，即用户之间的交流频率和深度，以及中心性，通过识别网络中的中心节点来了解信息传播的关键参与者。在计算参与度时，本研究根据出度来评估每个用户的参与程度，从而更全面地理解社会网络的结构和用户之间的相互作用。

3. 研究结果

3.1. 讨论帖整体呈现以发帖人为中心的单向网络关系

通过观察这 13 个帖子的社会网络图，本研究发现帖子间虽存在用户参与度高低的差异，但讨论帖整体呈现以发帖人为中心的单向网络关系。多个用户可能对同一讨论帖进行回复，

但这些回复通常是得不到有效回应的，因此会形成多对一的关系，讨论的焦点集中在发帖人引发的主题上。发帖人的观点和信息更容易被广泛传播，而其他用户的回复可能相对较少被关注，导致了信息传递的不对称性。

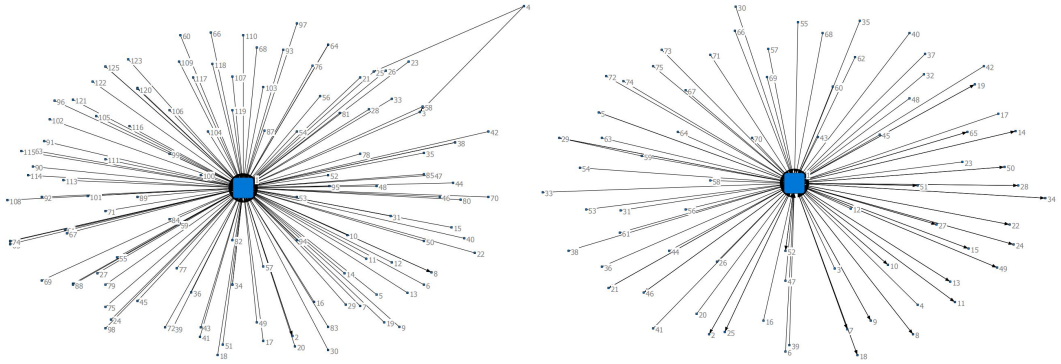


图 2 高、低参与度网络对比图（左高右低）

3.2. 讨论帖中学习者的正向学习情感多表现为喜欢与轻松

在 13 个讨论帖中的 909 条积极情感回复中，本研究通过分析发现学习者的正向学习情感多表现为喜欢，其次为轻松与有趣。在学习者正向学习情感中，喜欢与的情感占比最高，这一类情感的回复多表现出对于内容的肯定以及对于该项目发表一些自己的看法与建议。在轻松的情感下，发帖人对学习者所提出的问题进行有针对性的回应，帮助学习者更好地理解与实践。 “有趣”情感在回复中则表现为对内容具有较高的兴趣度和学习倾向。

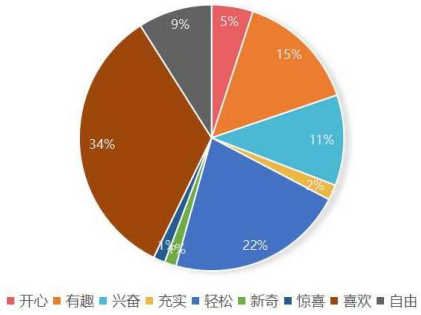


图 3 正向学习情感饼状图

回复1：好看又实用！----------喜欢情感
回复2：舵机时不能烧固件的。----------轻松情感
回复3：闲来无事，传感器已购买，待老夫架好了也发一长篇报导。-----有趣情感

图 4 正向学习情感示例

3.3. 讨论帖中学习者的负向学习情感多表现为困惑

在 13 个讨论帖中的 377 条消极情感回复中，本研究通过分析发现学习者的负向学习情感多表现为困惑，其次为沮丧与焦虑。困惑占据了学习者负向学习情感的绝大部分，占比高达 75%，表现为学习者对发帖人所呈现的内容中有不理解的地方进行疑问型的回复。沮丧和焦虑的情感，往往是因为学习者看到难度较高的创意作品时所产生的畏难心理以及尝试后失败对自己能力的不自信。



图 5 负向学习情感饼状图

回复1：如果连了GND的话 应该测到的也是低电平吧？ ---困惑情感
回复2：唉，时间模块调不好，时间跟日期老是对不上。--沮丧情感
回复3：很想试试，但是编程太难不会写。----------焦虑情感

图 6 负向学习情感示例

3.4. 发帖人参与程度有利于促进学习者正向学习情感

通过对发帖人参与程度，即发帖人在 SNA 中的出度与该主题帖的正向情感比例进行皮尔逊相关分析后发现：发帖人的参与度与参与者的正向情感体验不存在显著相关关系（P>0.05）。分析其内部原因，可能为在线创客社区的异步讨论中，用户于在线创客社区的异步讨论特性导致了人人交互存在时间和空间差异，学习者的学习情感难以获得即时性的反馈与纾解。另外，因为帖子本身的质量和难度水平，可能导致学习者的情感受到原贴难度和完备度的影响，从而影响到情感的表达。

尽管发帖人的支持程度并未直接影响主题帖整体情感分布，但在积极地解决学习者负向学习情感方面发挥着关键作用。在解决学习者负向学习情感方面，发帖人的支持可能成为社区中的一种情感调节机制，有助于提高学习者的学习体验和积极情感。

学习者：为什么按键开关要接地？-----困惑情感（消极）
发帖人：常态接地，按下接5V。-------轻松情感（积极）
学习者：学到了，谢谢分享！----------开心情感（积极）

图 7 情感变化示例

3.5. 不同情感倾向的主题帖呈现出规律性情感表征

将 13 个帖子的积极情感占比从高到低依次排序后，取排名前四为高积极情感状态下的帖子，最后四名为低积极状态下的帖子。研究结果显示，不同情感状态的主题帖呈现出一定的规律性情感表征。在高积极情感状态下，帖子中出现了更多的有趣（PI）、兴奋（PE）、喜欢（PL）的积极情感的表达，这可能反映了学习者在分享和讨论时更加倾向于传递正面的情感体验。其中有趣的关键词为想要、求分享，兴奋的关键词为牛、很棒，喜欢的关键词为厉害、赞。

在低积极情感状态下，困惑（NC）的情感在主题帖中占据比例大幅上升。这可能表明学习者在这种低积极情感状态下更加关注讨论主题的难度、个人喜好或者遇到的困惑。其中喜欢的关键词为不错、顶，与高积极状态下的关键词对比，情感的激烈程度有所下降。下边是高、低积极情感状态下的情感分布情况图。

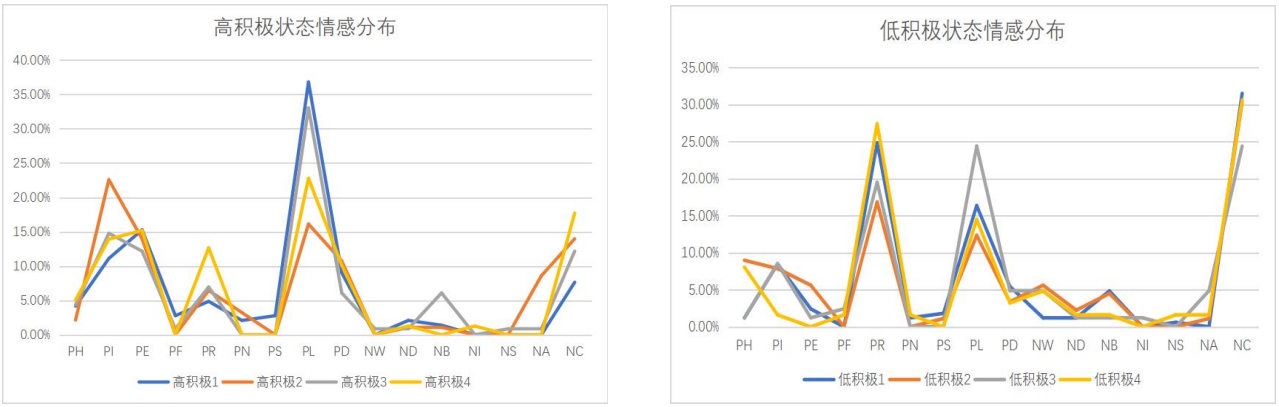


图 8 高、低积极情感状态分布图

4. 研究建议

综合上面三条研究结果，本研究发现在在线创客社区中，发帖人的支持程度并不直接影响主题帖整体情感分布，但却在积极地解决学习者负向学习情感方面发挥着关键作用。通过社会网络分析数据，本研究得到发帖人的参与度并不会直接影响参与者的情感体验。基于以上研究结果，本研究提出如下建议：

4.1. 建立在线同步创客兴趣社群，提升创客深度学习程度

通过创建一个在线的创客兴趣社群，可以为广大学习者提供一个共同的平台，让他们能够更加深入地探讨和分享关于创客不同领域的知识、经验和创意。社群的设立可以囊括各类主题、技能和项目，以满足不同学习者的兴趣和需求。在这种形式的社群互动中，学习者将有机会在更专业和深度的环境中进行学习，从而有助于提高整体学习质量，还能够激发更多创意和合作，形成一个更为丰富和活跃的在线创客社区。

4.2. 建立适切激励机制，鼓励发帖人实时解决用户困惑

有效的在线协作有助于促进创意灵感生成。通过设立激励机制，可以有效提高发帖人的工作动力，提高发帖人参与解答问题、引导讨论的积极性，还能够向社区其他成员传递积极的学习氛围，鼓励更多用户参与讨论和提出问题。通过这种方式，在线创客社区将形成一个良性循环，发帖人和用户之间的互动将更加积极，促进学习氛围的良好发展。

4.3. 对讨论主题进行深入分类，促进创客精准讨论

提出这个建议的核心原因是因为，本研究在分析的时候发现负向情感主要来源于困惑、焦虑，即博主的内容自己有一些质疑和问题，或者博主的内容太难，学习者不会。因此可以从内容深度分类和内容主题分类两个方向，对主题帖进行归类，帮助学习者进一步找到符合自己认知规律和学习基础的帖子，降低焦虑感，同时也提升博主的解答效率，促进正向学习情感生成。

参考文献

丁大琴(2015).创客及其文化历史基因探源.北京社会科学,08,22-28.

王佑镁和叶爱敏(2015).从创客空间到众创空间：基于创新 2.0 的功能模型与服务路径.电化教育研究,36(11),5-12.

董利亚和冯锐(2016).在线学习社区培育与发展模型的构建及其策略研究.远程教育杂志,35(02),98-105.

甘永成和王炜(2005).虚拟学习社区多重内涵之解析与研究.现代远程教育研究,05,10-15.

洪巍和李敏(2019).文本情感分析方法研究综述.计算机工程与科学,41(04),750-757.

钟佳娃,刘 巍,王思丽和杨恒(2021).文本情感分析方法及应用综述.数据分析与知识发现,5(06),1-13.

张存刚,李明和陆德梅(2004).社会网络分析——一种重要的社会学研究方法.甘肃社会科学,02,109-111.

王陆(2009).虚拟学习社区的社会网络分析.中国电化教育,02,5-11.

胡勇和王陆(2006).异步网络协作学习中知识建构的内容分析和社会网络分析.电化教育研究,11,30-35.

邱均平,余以胜和邹菲(2004).关于内容分析法的研究.中国图书馆学报,02,6.

宋振峰和宋惠兰(2012).基于内容分析法的特性分析.情报科学,30(07),964-966+984.

刘鲁川和孙凯(2015).社会化媒体用户的情感体验与满意度关系——以微博为例.中国图书馆学报,41(01),76-91.

邓君,马晓君和毕强(2014).社会网络分析工具 Ucinet 和 Gephi 的比较研究.情报理论与实践,37(08),133-138.

JunHao Shan & Wei Wang(2021).Making and sharing in asynchronous discussion: exploring the collaboration process in online maker community.Interactive Learning Environments.

Anderson, C. (2012). Makers: The new industrial revolution. Random House UK

Rheingold.H. (1993)The Virtual Community: Homesteading on the Electronic Frontier. Reading,MA: Addison-Wesley.

Carabajal. K, ,Lapointe. D. & Gunawardena. C.N. (2003).Group development in online learning communities. Handbook of Distance.

指向跨学科思维的教学实践研究——以“制作 3D 校园建筑模型”项目为例

Teaching Practice Research on Interdisciplinary Thinking: Taking the Project of "Creating 3D Campus Architecture Models" as an Example

周千明[1*]，张涛[1]，舒洁[1]，秦琴[1]，董连春[2]
[1]湖南省怀化市铁路第二中学
[2]中央民族大学理学院
[*] 641766361@qq.com

【摘要】跨学科思维是中小学生未来学习的必备素养，培养学生跨学科思维以项目式学习为主。开展跨学科思维的项目式学习，需厘清跨学科思维的内涵与价值，构建培养跨学科思维的设计框架，坚持以基本概念为认知起点，以数学交流和问题解决为活动核心，以学生生活为典型场景，以融会贯通为培育目标。以“制作 3D 校园建筑模型”项目为例进行培养跨学科思维的教学实践探索。

【关键词】跨学科思维；项目式学习；3D 立体图形；科学观念

Abstract: *Interdisciplinary thinking is an essential literacy for future learning of primary and secondary school students, and cultivating interdisciplinary thinking focuses on project-based learning. To carry out project-based learning with interdisciplinary thinking, it is necessary to clarify the connotation and value of interdisciplinary thinking, construct a design framework for cultivating interdisciplinary thinking, adhere to the cognitive starting point of basic concepts, take mathematical communication and problem-solving as the core activity, take student life as a typical scenario, and cultivate integration as the goal. Exploring the Teaching Practice of Cultivating Interdisciplinary Thinking Using the Project of "Creating 3D Campus Architecture Models" as an Example.*

Keywords: Interdisciplinary thinking, Project-based learning, 3D stereoscopic graphics, Scientific concepts

1. 前言

跨学科是中小学生未来学习的基本特征。在充满跨界融合的数字时代，跨学科将成为基础教育课程改革的必然趋势，跨学科思维将成为学生参与课程学习的必备素养。以项目式学习为载体，开展跨学科思维的教学实践是主动适应未来教育变革的教学探索。

2. 传统学习的局限

义务教育阶段的学习主要是在教师为主导下,通过各学科课程学习让学生逐步具备适应未来社会发展和终身发展所需的必备品格和关键能力，学习活动以教室为主要场所。然而，传统课堂学习存在一定局限性，主要表现为：缺乏促进概念理解的深度学习，缺乏提升综合能力的活动场景，缺乏培育科学观念的实践体验。

3. 基于项目的跨学科思维教学探索

3.1. 跨学科思维的内涵与价值

跨学科思维是将两个或两个以上学科或专业领域的知识和思维模式相结合以产生更先进认知的能力（Boix Mansilla, Miller, & Gardner, 2000; Spelt, Biemans, & Mulder, 2009），整合多个学科的内容以实现单一学科内容不能达到的目的，如问题解决和制作作品（Mansilla, 2010）。本质上跨学科思维是创造性地整合观点、使用隐喻和进行对话式的思维（Nikitina, 2005），以跨界整合学习内容和开展综合实践活动为代表。在科学课程融合中，跨学科思维的价值主要表现为培养学生的数学交流和问题解决能力。以学生的现实生活为活动背景，在视觉、触觉等直观感受中探究、发现和理解现实世界，体验跨学科魅力，感受知识应用价值，初步培养科学观念和科学思维。

3.2. 构建培养跨学科思维的设计框架

本研究构建培养跨学科思维的设计框架(见图 1)，认知层代表学生对某一真实问题下基本概念的认知基础，学生参与项目式学习活动，在任务驱动中从认知层进入活动层。学生通过交流与合作、行动与评价，在挑战性问题中进入场景层。典型真实场景下的复杂性问题迫使学生创造性开展跨学科思维与深度思考，从场景层转入思维层。思维层是学生在项目式学习经验中产生的思维变化，思维层促进学生对下一个真实问题的认知迭代与重构，逐渐形成初步的科学观念。在设计框架中，认知层和思维层是起点，也是终点，活动层和场景层是思维进阶过程，也是能力提升路径。

3.2.1 认知层：以基本概念为认知起点的跨学科认知

人与动物在思维方式上最根本的区别在于想象能力和抽象能力（史宁中, 2018）。青少年想象能力主要在观察现实生活场景中培养，而现实生活中图形都是立体图形，非平面图形。以测量为基本概念，在 SketchUp 动态几何环境中搭建感性体验与理性思维的桥梁，为学生培育抽象能力提供真实情景。测量校园地标建筑的长度是基础任务，长度测量是基本概念。微调立体模型是制作 3D 立体模型的关键环节，综合运用 SketchUp 平台的拖动模式、工程学技术和纸笔技术制作立体图形，增强学生对几何基本概念的理解（Camou, 2012），借助智能软件，通过虚拟几何图形的变换有效培养学生空间观念。总而言之，“立体图形”与“测量”是数学、科学、艺术等学科课程的基本概念，可视为跨学科思维的认知起点。

3.2.2. 活动层：以数学交流和问题解决为核心的项目式学习

项目式学习以聚焦特定主题为问题，以解决问题为主线，将数学系统内部连结起来，也能将数学与其他社会生活领域相连结，使学生通过解决有价值、有挑战的问题，积累丰富的数学活动经验，领会数学内部以及数学与其他学科领域的联系（徐斌艳, 2021）。数学交流和问题解决是项目式学习的核心。项目活动以小组合作为主要形式，小组成员之间需要有效的交流。用数学语言交流、论证实施方案，形成可行的项目计划书，小组成员之间需要有效的协作。在突破项目的难点问题时，以协作问题解决的方式探寻有效解决策略，综合运用数学思维、跨学科思维和科学思维完成项目任务，制作直观的项目作品。以问题解决为核心的项目关键环节是促进综合能力提升的“主阵地”，以数学交流为核心的项目实施过程是跨学科探究的“通用语”。

3.2.3. 场景层：以学生生活为背景的典型真实场景

典型真实场景是项目式学习的重要情景，也是跨学科思维发展的关键场域。在项目式迭代中发现、提炼典型真实场景，以学生生活为背景，创设典型真实场景，以学科协同育人为导向，聚焦在真实问题解决中的认知冲突，让学生自然体验跨学科思维方式，逐步培养跨学科思维的思维品质。本研究发现，学生在测量墙面长度和制作“整栋楼”立体模型过程中出现深层学习，思维方式发生本质改变。根据访谈文本分析，学生在挑战性任务中突破单一思维、在展示答辩中自证其说，将声音表征（答辩证实）与图像表征（数字作品）相互联系，整合认知体系，从深层学习中提升认知水平，有效积累跨学科思维活动经验。

3.2.4. 思维层：以融会贯通为目标培育跨学科思维

设计框架的育人价值集中体现在思维层。思维层主要表现在面对复杂问题具备融会贯通的能力，为应对未来复杂多变的现实世界，学生需要具有审时度势、融合多学科优势、依据科学规律解决问题的能力。例如，在测量中面对不规则图形墙面测量的问题，学生考虑测量的数学化、科学性等客观要求，融合数学、科学等学科思维模式，以出入相补原理间接测量。通过曲折而有挑战性的测量实践活动，让学生体验综合实践活动，在跨学科思维下逐步形成初步的科学观念，感受融会贯通的意义。同时，跨学科思维有助于学生掌握学科以外仍需具备的能力，例如，解决复杂问题的能力、跨学科交流能力、跨学科协作能力、跨学科创新力等（Kricsfalusy, 2018）。在项目各个环节中，通过一系列驱动性任务撬动思维方式的质变，超越传统课堂中的单一思维方式，以跨学科思维提升学生综合实践能力。

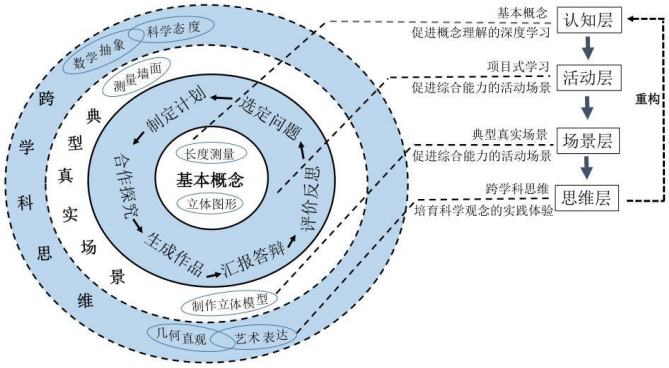


图 1 培养跨学科思维的设计框架

4. 跨学科思维教学实践：“制作 3D 校园建筑模型”项目

4.1. 以培养跨学科思维为目标设计项目式学习

培养跨学科思维的项目式学习设计的关键在于用跨学科思维代替单一学科思维来推动项目活动。引导学生开展深层学习，预设关键场域和突破思维禁锢，使学生突破思维边界，将视觉与听觉、想象与抽象、理论与实践等贯通在整个主题活动中。在“制作 3D 校园建筑模型”项目中，以培养跨学科思维为目标，以“封存初中时光”为驱动性任务，设计基于基本概念的项目学习及其相应的基本概念、认知指导策略、数字作品。

4.2. 聚焦基本概念设计项目的关键环节

聚焦基本概念的关键环节主要有：第一、测量不规则墙面的长度。融合数学和科学的学科培养目标，以现实测量问题为契机，应用数学基本思想，进行科学思辨和推理，以解决方案为表现形式。提供中国数学家刘徽的出入相补原理相关资料，以真实问题为驱动，激发学生以空间观念和科学思维为核心的跨学科思维方式解决问题，形成合理的解决问题方案（见图 2）。第二、制作立体图形。融合数学和科学的学科培养目标，以迭代过程为焦点，在优化立体模型中增强学生的几何直观和系统的观念。借助平移形成微小变化以发现“整栋楼”的本质特征，并评估“整栋楼”模型效果。在建模过程中，运用设计思维和系统思维，将墙体进行连接处理以便在每一层统一拉伸，从而快速完成每层的建模；将相同楼层和相同尺寸的门窗进行单独绘制，通过复制的方式快速建模。在模型修复中，以真实数据和建筑标准等科学依据进行迭代更新，以系统思维方式整体调整“整栋楼”立体模型（见图 3）。

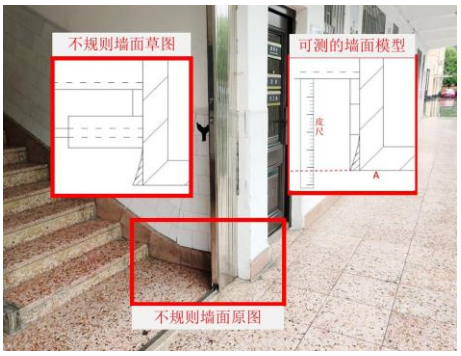


图 2　不规则墙面的长度测量

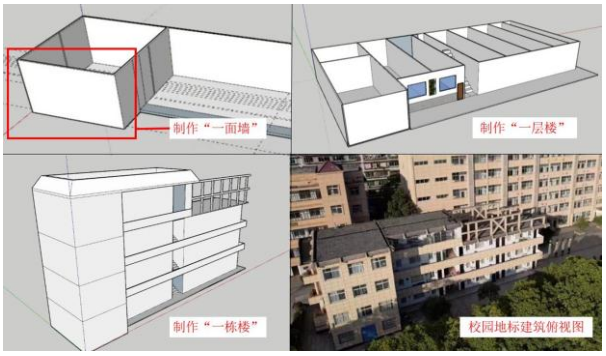


图 3 制作“校园教学楼”立体模型概图

5. 反思

首先，需关注学生思维本质变化。培养跨学科思维并不是追求某个学科知识与技能的理解与掌握，而是追求思维方式的变革，学会用多元视角解决问题。在深层学习下激发自我认知革新的项目式学习才能撬动思维的本质变化，让学生学会学习。其次，运用好跨学科思维与科学观念的辩证关系。跨学科思维是形成科学观念的思维来源，是科学探究能力的基础。在发现、分析问题中运用跨学科思维，有助于培养学生科学观念。学生依据科学规律解决问题的活动经验能促进学生跨学科思维的形成，让学生感受跨学科思维的力量。

参考文献

史宁中（2018）。**教育的本源与思考**。北京：商务印书馆。

徐斌艳（2021）。**数学素养与数学项目式学习**。上海：华东师范大学出版社。

Boix Mansilla, V., Miller, W. C., & Gardner, H. (2000). On disciplinary lenses and interdisciplinary work. Interdisciplinary curriculum: Challenges to implementation, 17-38.

Camou, B. J. (2012). High school students' learning of 3D geometry using iMAT (integrating Multitype-representations, Approximations and Technology) engineering (Doctoral dissertation, University of Georgia).

Mansilla, V. B. (2010). Learning to synthesize: The development of interdisciplinary understanding. The Oxford handbook of interdisciplinarity, 288-306.

Nikitina, S. (2005). Pathways of interdisciplinary cognition. Cognition and Instruction, 23(3), 389-425.

Spelt, E. J., Biemans, H. J., Tobi, H., Luning, P. A., & Mulder, M. (2009). Teaching and learning in interdisciplinary higher education: A systematic review. Educational Psychology Review, 21, 365-378.

Kricsfalusy, V., George, C., & Reed, M. G. (2018). Integrating problem-and project-based learning opportunities: Assessing outcomes of a field course in environment and sustainability. Environmental education research, 24(4), 593-610.

基于 Mind+授课与智能菜园应用的教育创新研究

Research on educational innovation based on Mind + teaching and intelligent vegetable garden application

赵赫璇*，马国威 ，付玉娜，孙伟华，毛云庆，王妍莉
西北民族大学教育科学与技术学院
*3466270662@qq.com

【摘要】 随着科技的不断发展，教育行业也在寻求创新与变革。本文主要探讨了利用 Mind+授课平台和智能菜园应用进行教育创新的有效途径，通过构建智能菜园的构建，将智能菜园应用于教学课程，为教育工作者提供一种全新的教学体验，提高教学质量，激发学生学习兴趣，培养创新型人才。

【关键词】 Mind+；智能菜园；教育创新

Abstract: *With the continuous development of science and technology, the education industry is also seeking innovation and change. This paper mainly discusses the effective ways of educational innovation by using Mind + teaching platform and intelligent vegetable garden application. Through the construction of intelligent vegetable garden, intelligent vegetable garden is applied to teaching courses, which provides a new teaching experience for educators, improves teaching quality, stimulates students ' interest in learning and cultivates innovative talents.*

Keywords: Mind + , intelligent vegetable garden , educational innova

本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学本科教学质量提升项目省级一流课程《信息技术教学法》（项目编号：2022SJYLKC-37）资助。

1. 引言

共近年来，我国教育信息化发展迅速，越来越多的教育工作者开始关注并尝试运用现代科技手段改革传统教学模式（荆思凤等，2022）。Mind+作为一款优秀的授课平台，为教育创新提供了有力支持。同时，智能菜园作为一种新兴的农业科技产品，与教育相结合有望开辟一片新天地。本文旨在探讨如何将 Mind+授课与智能菜园应用有效结合，为教育创新提供一种新思路。

2. Mind+授课平台概述

2.1. Mind+平台功能特点

Mind+ 是一个支持多平台、跨硬件的开发工具，具有图形化编程界面和代码编程支持，丰富的教学资源与社区，智能硬件集成，云端服务支持，项目分享与协作等特点，适合于创客教育、编程学习和项目开发，不断更新改进，符合现代教育标准（何剑，2022）。

2.2. Mind+平台授课的优势

但使用 Mind+平台授课具有显著优势，它不仅提供直观的图形化编程环境，降低学习门槛，还支持跨平台和多样化的硬件设备，使教学更加灵活。丰富的教学资源和活跃的在线社区为师生提供了广阔的交流和学习空间，云端服务支持则拓展了项目的可能性（杜峰涛等，2023）。此外，Mind+还便于项目分享与协作，促进团队合作和创新思维的培养，完美契合现代教育标准，提升教学效果（王倩倩等，2020）。

3. 智能菜园在教育中的应用

3.1. 智能菜园简介

智能菜园是一种利用物联网、大数据、云计算等先进技术，实现蔬菜种植智能管理的系统（王志冉等，2021）。通过智能菜园，教师可远程监控蔬菜生长状况，为学生提供实践操作的平台。

3.2. 智能菜园在教育中的应用优势

智能菜园将物联网、传感器技术、数据分析等智能科技融入植物生长过程，为学生提供了一个互动、直观且充满探索乐趣的学习环境。通过智能菜园，学生可以实时监测植物生长状态，学习生物学、环境科学等知识，培养科学探究能力；同时，智能菜园还能帮助学生了解生态平衡、可持续发展等理念，增强环保意识。此外，智能菜园鼓励学生动手实践，提高解决问题的能力，并在团队合作中锻炼沟通与协作技巧，为综合素质教育注入新的活力（廖剑等，2023）。

3.3. 智能菜园简介

3.3.1. 智能菜园的技术原理

智能菜园利用物联网技术、传感器监测、自动灌溉和智能数据分析等手段，实现蔬菜种植的自动化和智能化。学生在教师的指导下，可以通过 Mind+平台远程监控菜园的生长状况，学习农业科技知识，体验农耕文化。

3.3.2. 教育应用的实践策略

智能菜园的教育应用不仅仅是技术操作的学习，更是一种生活教育和环境教育。教师可以设计一系列实践活动，如菜园设计大赛、生态观察日记、植物生长实验等，让学生在实践中学习生物知识、环境科学和可持续发展理念。

3.3.3. 跨学科学习的融合

智能菜园的实践可以与多个学科相结合，如生物学、环境科学、信息技术、艺术等。学生可以通过绘画、摄影、写作等方式，表达对菜园的观察和理解。这样的跨学科学习不仅丰富了学生的学习体验，也提高了他们的综合素养。

3.4. 教育创新的挑战与对策

3.4.1. 技术挑战

虽然 Mind+平台和智能菜园为教育创新提供了新的可能性，但技术实施和应用过程中也会遇到挑战。例如，网络稳定性、数据安全问题、设备维护等。学校和技术提供商需要不断优化技术支持，确保教育创新的顺利进行。

3.4.2. 教师培训

教育创新对教师的专业能力提出了新的要求。教师需要掌握 Mind+平台的使用技巧，了解智能菜园的技术原理，并能设计出符合学生需求的教学活动。因此，定期的教师培训和教研活动是推动教育创新的关键。

3.4.3. 学生适应

智对于学生来说，适应新的学习模式可能需要一段时间。教师需要关注学生的学习适应情况，提供必要的指导和帮助。同时，通过激励和反馈机制，鼓励学生积极参与，提高他们的自我管理能力。

4. 基于 Mind+授课与智能菜园应用的教育创新实践

4.1. 第一课时：智慧菜园的概念与特点

通过图片展示和讨论引入智慧菜园的概念，进而讲解其定义、特点和优势，分析一个成功案例以了解运作原理和技术应用，最后让学生小组讨论智慧菜园对农业生产和生态环境的影响。具体步骤如下：

4.1.1. 导入：

通过图片展示和讨论，引出智慧菜园的概念。

4.1.2. 讲解：

介绍智慧菜园的定义、特点和优势。

4.1.3. 案例分析：

分析一个成功的智慧菜园案例，了解其运作原理和技术应用。

4.1.4. 小组讨论：

让学生探讨智慧菜园对农业生产和生态环境的影响。

4.2. 第二课时：智慧菜园的规划与管理

通过提问导入智慧菜园规划、管理及生态系统研究的重要性，并讲解相关基本原则、方法、生态系统的组成和功能，然后让学生参与规划与管理实践活动，如设计种植方案和制定灌溉计划，并进行实地考察观察记录，最后小组合作探讨智慧菜园生态系统对生态环境的影响，并分组讨论制定规划与管理方案。具体步骤如下：

4.2.1. 导入：

通过提问，引出智慧菜园的规划、管理及生态系统研究的重要性。

4.2.2. 讲解：

介绍智慧菜园的规划与管理的基本原则和方法、介绍生态系统的组成和功能。

4.2.3. 实践活动：

让学生参与智慧菜园的规划与管理，如设计种植方案、制定灌溉计划等；小组进行实地考察，观察和记录智慧菜园的生态系统。

4.2.4. 小组合作：

让学生探讨智慧菜园的生态系统对生态环境的影响；让学生分组合作，讨论并制定一个智慧菜园的规划与管理方案。

4.3. 第三课时：各项功能实现

通过将 DHT11 或 DHT22 空气温湿度传感器、YL-697 土壤湿度传感器和 BH1750 光度传感器等合适传感器选择并连接到 Mind+开发板，编写并调试代码以读取传感器数据，上传代码到 Mind+行空板，并在 Mind+环境中展示和操作数据，实现对智慧菜园环境参数的监测和控制。具体程序如图 1、2 所示；具体实物如图 3、4 所示；具体数据如图 5、6 所示。具体操作步骤如下：

4.3.1. 选择合适的传感器：

使用空气温湿度传感器：例如 DHT11 或 DHT22；使用土壤湿度传感器：例如 YL-697；使用光度传感器：例如 BH1750。

4.3.2. 连接传感器到 Mind+：

将传感器通过相应的接口连接到 Mind+开发板。将传感器的 VCC 和 GND 引脚连接到 Mind+的 5V 和 GND 引脚，将传感器的数据引脚连接到 Mind+的一个数字引脚。

4.3.3. 编写代码：

在 Mind+的编程环境中编写代码，以读取传感器的数据。每个传感器都有其特定的库，我们需要调用这些库来获取数据。例如，对于 DHT11/DHT22 传感器，可以使用“DHT”库；对于 BH1750 传感器，可以使用“BH1750”库。

4.3.4. 调试代码：

在上传代码到 Mind+之前，先进行调试，确保代码能够正确地读取传感器的数据。

4.3.5. 传代码到 Mind+：

使用 Mind+的编程环境将代码上传到 Mind+行空板。

4.3.6. 数据显示和操作：

在 Mind+的环境中，我们可以选择如何显示和操作这些数据。例如，我们可以将数据发送到 Mind+的 dashboard，或者使用数据来控制其他设备，如灌溉系统，遮光系统等等。

4.4. 第四课时：智慧菜园的创新设计比赛

通过提问导入智慧菜园创新设计比赛的目的和意义，接着讲解比赛要求和评分标准，然后让学生分组合作实践活动设计制作智慧菜园模型，最后评价与展示各组设计，并评选出最佳设计奖。具体步骤如下：

4.4.1. 导入：

通过提问，引出智慧菜园的创新设计比赛的目的和意义。

4.4.2. 讲解：

介绍比赛的要求和评分标准。

4.4.3. 实践活动：

让学生分组合作，设计和制作一个智慧菜园模型。

4.4.4. 评价与展示：

对学生的设计进行评价和展示，评选出最佳设计奖。

5. 结论

本文通过对 Mind+授课平台和智能菜园应用的探讨，提出了一种基于现代科技手段的教育创新模式。实践证明，这种模式能有效提高教学质量，激发学生学习兴趣，培养创新型人才（范红梅&王文龙，2021）。 Mind+授课平台和智能菜园的应用为教育创新提供了新的视角和方法。通过整合先进的技术和实践活动，教育工作者可以设计出更加互动、个性化和创新的教学模式，满足学生的多样化学习需求（胡姣&祝智庭，2021）。然而，教育创新的道路并非一帆风顺，需要克服技术、师资、学生适应性和家校合作等方面的挑战。通过持续的研究和实践，教育工作者可以不断优化教育创新模式，为学生的全面发展奠定坚实的基础。

参考文献

王志冉，王浚峰，周增产等. 家庭智慧菜园的设计与应用[J]. 湖北农业科学,2021,60(17):125-129.

杜锋涛,周娟,张新奇等.融合 Mind+与 Arduino 开源硬件的 STEM 教学设计——以设计制作智能教室模型为例[J].中国现代教育装备,2023(20):50-52+60.

何剑.图形化编程辅助学前师范生课件制作初探[J].扬州教育学院学报,2022,40(01):91-94.

范红梅，王文龙. 基于建造主义的 Mind+少儿编程教学设计[J]. 工业控制计算机,2021,34(10):103-105.

胡姣，祝智庭. 技术赋能的教学微创新：教师教育智慧的实践场[J]. 中国电化教育,2021(08):99-109.

荆思凤,刘希未,宫晓燕等.人工智能时代人力资本新需求与教育变革的关系研究[J].华东师范大学学报(教育科学版),2022,40(09):10-18.DOI:10.16382/j.cnki.1000-5560.2022.09.002.

廖剑，刘选，刘革平. 理感联通：人工智能赋能智慧教育新范式[J]. 中国电化教育,2023(06):18-24.

王倩倩,刘长征,刘圆等.思维导图和流程图在 Scratch 教学中的应用[J].现代计算机,2020(17):79-84.

基于 STEAM 的项目式学习模式研究——以“创作同辈辅导视频”为例

Research on STEAM-based Project-Based Learning Model

——A Case Study of "Creating Peer Tutoring Videos"

姚瑶[1]，伊亮亮[2*]
[12]鲁东大学初等教育学院
[*]215941324@qq.com

【摘要】 本研究融入 STEAM 教育、项目式学习和 Learning by Teaching 理论，构建了创作同辈辅导视频的项目式学习模式。通过完成“创作同辈辅导视频”项目，多学科知识得以融合，从而推动学生跨学科学习和实践能力的发展。

【关键词】 STEAM 教育；项目式学习；Learning by Teaching；同辈辅导视频创作

Abstract: *This study integrates STEAM education, project-based learning, and the Learning by Teaching theory to construct a project-based learning model for creating peer tutoring videos. By undertaking the "Creating Peer Tutoring Videos" project, interdisciplinary knowledge is fused, thereby advancing students' cross-disciplinary learning and practical skills development.*

Keywords: STEAM education, Project-based learning,Peer Tutoring video creation

1. 前言

当前社会对具备创新能力和实践能力的全面型发展人才的需求日益迫切，《义务教育课程方案（2022 年版）》中更是提出了“统筹设计综合课程和跨学科主题学习”的要求。在这一背景下，STEAM 教育整合了科学、技术、工程、艺术、数学等学科知识。同时，项目式学习作为一种综合性的学习方式备受推崇。

2. 基于 STEAM 的同辈辅导视频项目式学习模式构建

本研究构建了 STEAM 理念下的同辈辅导视频项目式学习模式，如下图所示。

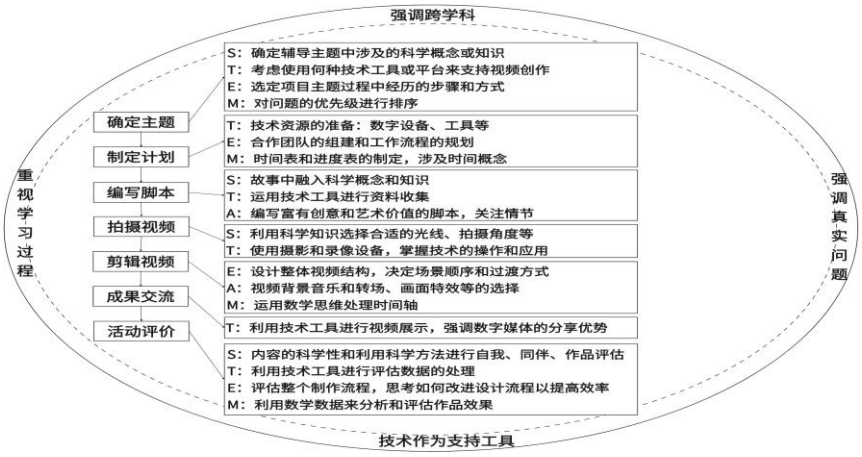


图 1　STEAM 理念下的同辈辅导视频项目式学习模式

3. 基于 STEAM 的同辈辅导视频项目式学习模式案例

3.1 选定项目主题

选择主题的过程能够促进学生的工程思维。选定项目主题的流程如下图所示。



图 2 项目主题选定流程

3.2 制定计划

制定计划可以有条不紊引导项目的进行。制定计划的具体流程如下图所示。



图 3 计划制定流程

3.3 编写视频脚本

核心驱动问题：“如何通过制作一个同辈辅导视频来促进同学们对疑难问题的理解？”。

【活动】分解驱动问题 1：如何组织各个知识点和设计视频结构，以确保能够清晰传达视频讲解的疑难问题的解决方法？

设计思路：让学生利用工程思维将知识点整合为一个逻辑有序的框架。

【活动】分解驱动问题 2：通过什么样的辅导方式使得疑难问题更容易理解和接受？

设计思路：让学生们了解选择不同辅导方式在知识传递和学习效果方面的差异，例如数字故事、情景演绎和辅导式对话等的差异。

【活动】分解驱动问题 3：如何运用数字化工具制作富有创意和吸引力的同辈辅导视频？

设计思路：让学生选择适当的数字化工具，了解能够吸引人的元素。例如，有学者指出视频中存在一个演绎学生的角色会促进观看者的学习，并且冲突性场景会鼓励观看者参与其中。

3.4 拍摄视频和剪辑视频

在这个过程中，学生需要运用工程思维来设计整体视频结构，同时运用数学思维巧妙处理时间轴。此外，在视频的视觉呈现中，艺术元素也起到不可忽视的作用。

3.5 成果交流与作品评价

在本环节，学生充分展现了 STEAM 元素的综合运用。科学元素体现在利用科学方法评估，技术元素体现在制作成果上，而工程思维则通过团队讨论和作品改进得以充分呈现。艺术元素则在作品创意中体现，数学思维贯穿于整个评价标准的应用过程中。

4. 总结

本研究结合 Learning by Teaching 理论，构建了基于 STEAM 教育理念的创作同辈辅导视频的项目式学习模式，旨在培养学生的跨学科学习能力以及动手实践能力。

参考文献

秦瑾若和傅钢善.(2017). STEM 教育:基于真实问题情景的跨学科式教育. 中国电化教育, (4), 67–74.

桑国元、叶碧欣、黄嘉莉和罗颖.(2023). 构建指向中国学生发展核心素养的项目式学习标准模型. 中国远程教育, 43(6), 49–55.

Chi, M. T. H., Kang, S., & Yaghmourian, D. L. (2017). Why Students Learn More From Dialogue-Than Monologue-Videos: Analyses of Peer Interactions. Journal of the Learning Sciences, 26(1), 10–50. https://doi.org/10.1080/10508406.2016.1204546

培养职前教师跨科学教学专长的工作坊设计

Design of an workshop to cultivate pre-service teachers′ interdisciplinary teaching expertise

付玉娜*，马国威，赵赫璇，区文诗，孙伟华，雅古，王妍莉
西北民族大学教育科学与技术学院
217619566@qq.com

【摘要】 义务教育课程改革倡导教师开展跨学科教学，对教师的跨学科教学设计能力提出了新的要求。本研究基于职前教师跨学科教学设计能力影响因素的作用机制，从职前教师的跨学科教学专长培养入手，通过工作坊将各学科教师组成跨学科学习共同体，通过学习跨学科知识和各学科的知识共享，基于逆向工程教学法的跨学科案例，以及跨学科教学方案的设计，从原理、方法和实践三方面助力职前教师跨学科教学设计能力成长。

【关键词】 工作坊；职前教师；跨学科教学专长；

Abstract: *The reform of compulsory education curriculum advocates for teachers to carry out interdisciplinary teaching, which puts forward new requirements for their interdisciplinary teaching design ability. This study is based on the mechanism of the influencing factors on the interdisciplinary teaching design ability of pre-service teachers. Starting from the cultivation of interdisciplinary teaching expertise of pre-service teachers, various subject teachers are formed into an interdisciplinary learning community through online workshops. Through learning interdisciplinary knowledge and sharing of knowledge among different subjects, interdisciplinary cases based on reverse engineering teaching method, and the design of interdisciplinary teaching plans, the study aims to help pre-service teachers grow their interdisciplinary teaching design ability from three aspects: principles, methods, and practices.*

Keywords: Online workshop; pre-service teachers; interdisciplinary teaching design expertise;

*本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学本科教学质量提升项目省级一流课程《信息技术教学法》（项目编号：2022SJYLKC-37）资助。

义务教育课程方案(2022 年版)明确规定设立跨学科主题学习活动，对教师的跨学科教学设计能力提出了更高要求，但教师对跨学科概念理解不足、相关培训支持不足，因此培养职前教师跨学科教学设计能力尤为必要。跨学科教学专长是教师拥有的识别和探索复杂的跨学科教学问题并提供详细规划的解决方案以改善教学效果的综合性技能（Stefaniak, J. E., & Hwang, H. ,2021）。依据职前教师跨学科教学设计能力的作用机制，跨学科教学专长是影响师跨学科教学设计能力发展的关键因素，起直接影响作用（蔡慧英、李欣和孙佳悦，2023）。

1. 课程设计

组建学习共同体是跨学科学习的实施途径（董艳、孙巍和徐唱，2019），也能促进学习者产生跨学科学习身份认同。参照 STEM 教育跨学科学习共同体的构成要素（赵慧臣、张娜钰和马佳雯，2020）以及师范生学教融合的培养框架（钟柏昌、刘晓凡和陈岚鑫，2023），本课程创设了工作坊形式的开放共享的混合式学习环境，线上通过 MOOC 在线学习平台提供理论学习资源，线下基于逆向工程教学法进行案例学习，以及混合环境下的项目实践；提供群体协作工具支持不同学科的职前教师协作梳理和讨论交流跨学科教学的理论内容，以及围绕感兴趣的主题讨论不同学科之间的关联并进行整合；设计促进知识整合与创新的规则和项目式分工，另外通过自由结组建立共同愿景以及聚焦感兴趣的主题加强跨学科学习共同体的凝

聚力，以消除学科间的知识间隙并促进不同学科的职前教师有效共处，创造“共同体成员为了同一目标而努力”的氛围是跨学科学习共同体质量的界定标准。

2. 教学流程

工作坊教学活动流程包括：（1）基于 MOOC 的跨学科教学理论学习：利用 Xmind 思维导图、协作文档等工具协作学习 MOOC 的跨学科教学理论内容和各学科教材及课标，围绕感兴趣的主题讨论不同学科之间的关联并进行整合，每人完成一份理论学习报告。（2）基于逆向工程教学法的跨学科案例学习：逆向工程教学中造物的过程具有跨学科学习的特征，同时有利于打破项目式教学的高要求，因此本研究采用逆向工程教学法，逆向工程强调以现有产品为研究对象，通过建模、仿真和再设计实现原有产品的优化和改造，本研究采用逆向工程教学中的解构复原型教学模式，指学生在教师的引导下，经历一个对产品进行分解与复原的探究过程，其实施过程主要包含试用与感知、分解与观察、还原与测试、评价与总结等四个环节（康斯雅和钟柏昌，2019），为小组提供案例作品及相关软硬件设备，小组协作讨论完成功能实现分析报告。（3）基于项目实践的跨学科教学设计：在跨学科教学设计项目实践环节，各小组围绕感兴趣的主题协作完成跨学科教学方案设计，之后进行交流评价来借鉴吸收他人的教学优点，完成组间评价报告及个人反思。最后根据案例学习和项目实践中各成员的参与度、探究深度和协作能力进行小组合作贡献度评价，包括自评和组内互评。

3. 教学评价

职前教师跨学科教学能力评价采用朱德全和彭洪莉构建的测评指标体系（朱德全和彭洪莉，2023），教师跨学科教学知识通过包括各学科知识和跨学科理论的理论学习报告考察教师的不同学科内容知识水平以及整合不同学科内容知识进行教学的水平，教师跨学科教学能力通过跨学科教学设计方案和个人反思报告考察教师设计、实施和评价跨学科教学活动的水平，教师跨学科教学情意通过个人反思报告考察教师对跨学科教学的内在理念认识与外在行为态度。结合上述三个方面的评价结果以及整体参与表现，采用自评、互评、教师评价结合的方式对各位教师进行综合评价，并提供评价反馈和指导，帮助团队成员了解自己的优势和薄弱点，旨在全面客观地评估职前教师的学习成果，促进其跨学科教学设计能力和专业发展。

参考文献

朱德全和彭洪莉（2023）。教师跨学科教学素养测评模型实证研究。华东师范大学学报(教育科学版)，41(02)，1-13。

赵慧臣、张娜钰和马佳雯（2020）。STEM 教育跨学科学习共同体：促进学习方式变革。开放教育研究，26(3)，91-98。

钟柏昌、刘晓凡和陈岚鑫（2023）。“新师范”背景下师范生跨学科创新能力培养的基本框架与实践案例。电化教育研究，44(07)，114-120+128。

康斯雅和钟柏昌（2019）。机器人教育中的逆向工程教学模式构建。现代远程教育研究，(04)，56-64。

董艳、孙巍和徐唱（2019）。信息技术融合下的跨学科学习研究。电化教育研究，(11),70-77。

蔡慧英、李欣和孙佳悦（2023）。提升职前教师跨学科教学设计能力的课程设计及其实证研究。现代教育技术，(12)，56-64。

Stefaniak, J. E., & Hwang, H. (2021). A systematic review of how expertise is cultivated in instructional design coursework. *Educational Technology Research and Development*, *69*, 3331-3366.

基于 STEM 教育理念的“希罗喷泉”教学活动设计

Instructional Design of the "Heron's Fountain" Based on STEM Education Theory

高诗婕[1*]
[1]杭州师范大学经亨颐教育学院
[*] 1436124419@qq.com

【摘要】 新课程标准提倡教师利用数字化工具辅助教学，促进学生思考，培养科学核心素养。然而，在当前义务教育阶段，教学过于注重成绩分数和传授理论知识，不利于培养核心素养。近年来，STEM（科学、技术、工程和数学）教育理念得到广泛认可，因此，将 STEM 和创客教育引入初中科学教学十分必要。本文基于 STEM 教育理念，设计了“希罗喷泉”教学活动，引入数字化信息系统，提升学生运用现代技术的能力与核心素养。
【关键词】 希罗喷泉；STEM 教育理念；初中科学；核心素养

Abstract: *The new curriculum standards advocate that teachers use digital tools to assist teaching, promote students' thinking, and cultivate scientific literacy. However, in the current stage of compulsory education, teaching pays too much attention to grades and imparting theoretical knowledge, which is not conducive to cultivating core literacy. In recent years, STEM (science, technology, engineering, and mathematics) education has been widely recognized, so it is necessary to introduce STEM and maker education into junior high school science teaching. Based on the STEM education concept, this paper designs the teaching activity "Heron's Fountain," which introduces a digital information system and improves students' ability to use modern technology and core literacy.*
Keywords: Heron's Fountain, STEM education theory, junior high school science, core literacy

1.简介

《义务教育科学课程标准》中指出通过实验认识大气压强，了解大气压强的变化及其对生活的影响；了解液体内部存在的压强，以及影响液体压强大小的因素。并且提倡在课堂教学中采用多元化的教学方式,积极开展思维型探究和实践活动来培养科学素养(胡卫平, 2022)。希罗喷泉的结构简单且设计精巧，为教师提供了开展启发性、进阶性和综合性教学的机会。尽管目前希罗喷泉的结构各异，但其原理却是相通的。学生可能只关注现象而忽视原理。此外，目前对于“希罗喷泉”的教学研究很少结合数字化手段。在进行科学活动时，教师不仅要注重趣味性，还应注重学科知识的适度拔高。为了增加教学深度，可以基于 STEM 教育理念，利用 DISLab 压强传感器数据采集装置来实时监测气压变化情况,该设备所采集的数据准确度高、可视性强，使学生能够亲身体验希罗喷泉的实验过程。此外，教师可以组织进阶活动，引导学生采集创新希罗喷泉的多样化数据，结合物理原理知识，最终得出量化分析的结论。

2.理论基础与相关阐述

2.1. 初中科学核心素养 初中科学核心素养由科学观念、科学思维、探究实践、态度责任构成。科学的学习需要通过理解来逐步内化、提炼，从而构建正确的科学观念。教师要让学生不断进行探索，形成深层多元的思考，促进科学思维的形成。鼓励学生在探究实践中多角度分析问题、多路径搜集证据、多方法解决问题，培养学生对科学谨慎和批判的态度。

2.2. STEM 教育理念 STEM 是科学、技术、工程以及数学四门学科的简称(刘健智 et al., 2020)。STEM 教育鼓励学生通过整合跨学科的知识和技能来解决现实世界的问题，以增强综合思维能力和创新能力(李春密 & 赵芸赫, 2017)。此外，STEM 教育理念倡导利用计算机、传感器等现代技术工具来加强学生的学习体验，培养他们的科技创新能力和信息素养。

3.基于 STEM 教育理念下的科学活动设计

活动的设计是以泰勒原理对课堂教学设计的四个基本问题为框架，利用 SMART 模型来设定学习目标：1.通过制作希罗喷泉装置，锻炼动手操作能力、团队合作能力；2.通过借助 DISLab 掌握装置中的压强变化情况，学会使用 DISLab；3.通过数据的采集与量化分析，锻炼科学思维，提高信息素养；4.通过经历项目式活动过程，在具备基本科学知识与技能的基础上，不断发展模型构建能力、技术与工程实践能力等，实现学生核心素养的提升。同时，还运用了 TARGET 模型来检查课堂活动的合理性(Ames, 1995)。

在具体的实施中，本节课对初二、初三学生开设。通过播放音乐喷泉视频激发学生的观赏兴趣与自制喷泉的好奇心以导入本课的主题。学生以小组为单位进行探究，分析希罗喷泉的原理，运用已有的科学、技术、工程知识和艺术创意，选用合适的材料，设计、制作、测试“希罗喷泉”。小组成员们充分讨论、提出质疑、解决分歧、决定方案、合作执行，在此过程中，能够很好地锻炼学生的动手制作能力、团队协作能力，指向第一个学习目标。此外，学生们利用 DISLab 压强传感器采集装置的数据并进行量化分析,使实验数据分析与理论计算结果相互验证，形成知识闭环，指向第二个和第三个学习目标。最后进行展示与评价，借助评估准则对学生表现与作品进行互评与自评，形成科学性的总结反思。本次活动重视思维型科学探究，强调动手和动脑结合，学习内容依照从浅入深、由表及里、由现象到本质的原则展开。项目式活动借助科学、技术与工程的知识，让学生经历模型建构、推理论证、创新实践，实现逐步进阶，锻炼学生的科学思维能力与探究实践能力，指向第四个学习目标。

4.效果与反思

希罗喷泉教学活动较好地实现了设定的学习目标。学生通过制作喷泉锻炼了动手操作能力和团队合作能力；通过实践掌握了 DISLab 的使用方法；通过数据采集与量化分析，提升了学生的数据处理能力与逻辑思维能力；通过项目式实践，不仅增进了对科学知识的理解，还培养了模型构建能力和工程实践能力。学生在活动展示与评价环节表现优秀，展示了对实验结果的深入理解和经验总结。综合来看，该活动为学生提供了一个综合性、实践性的学习机会，激发了学生的学习兴趣和探究欲望，促进了他们的学科素养和综合能力的提升。

然而，实施过程也出现了一些问题，例如部分学生在初次使用 DISLab 与计算机进行数据采集与分析方面存在困难，需要更多的指导与支持，可以提供支架式学习资料等。因此，未来在活动设计中需要更加注重学生的个体差异，提供更加个性化的支持与引导。

参考文献

胡卫平. (2022). 在探究实践中培育科学素养——义务教育科学课程标准（2022 年版）解读. 基础教育课程 (10), 39-45.

李春密 & 赵芸赫. (2017). STEM 相关学科课程整合模式国际比较研究. 比较教育研究 (05), 11-18. doi:10.20013/j.cnki.ice.2017.05.002.

刘健智, 高岩 & 曾红凤. (2020). STEM 教育融入物理教学的研究现状与路径探讨. 天津师范大学学报(基础教育版) (01), 56-60. doi:10.16826/j.cnki.1009-7228.2020.01.011.

Ames, C. (1995). Achievement goals, motivational climate, and motivational processes.

基于创客教育的跨学科主题学习设计研究*

Research on the Design of Interdisciplinary Thematic Learning Based on Maker Education

赵瑛[1*]，陈琦[1*]
[1]北京联合大学 师范学院，北京 100011
[*] sftzhaoying@buu.edu.cn 1603664380@qq.com

【摘要】 本文在创客教育模式的指导下，提出了跨学科主题学习的设计原则和设计路径，并以小学信息科技课程为例，进行了实践应用。应用效果表明，本文的研究成果能够为跨学科主题学习设计提供完备的流程指引。
【关键词】 跨学科主题学习；创客教育；信息科技

Abstract: *Under the guidance of the maker education model, this paper proposed design principles and paths for interdisciplinary thematic learning, and applied them in practice using primary school information technology courses as an example. The application results indicate that the research findings of this paper can provide comprehensive process guidance for interdisciplinary thematic learning design.*
Keywords: Interdisciplinary thematic learning, Maker Education, Information Science and Technology

1. 前言

2022 年，我国教育部颁布了《义务教育课程方案（2022 年版）》，明确提出“各门课程用不少于 10%的课时设计跨学科主题学习”。然而，跨学科主题学习具体如何设计目前还缺乏完备的流程指引（李锋等, 2023）。创客教育是实施跨学科教育的有效途径，它通常以现实生活中的真实问题为导引，组织学生围绕问题解决开展知识探究、创新创意、动手造物、交流分享、评价反思等学习活动，进而培养学生的综合素养（祝智庭等, 2015）。本文基于创客教育模式，提出了新课程方案下跨学科主题学习的设计原则和设计路径，并进行了实践应用。

2. 设计原则

依据新课程方案和创客教育模式，跨学科主题学习设计遵循三个原则：①跨域整合原则：跨学科主题学习要以某一学科的核心知识和思想方法为载体，围绕解决跨域问题的实际需要，整合其他学科的知识和方法，进行整合设计，形成主题任务；②协同建构原则：教师要通过情境创设、组织交流协作、提供必要的支架等，组织引导学生以团队形式开展知识的协同建构，并采用多元学习评价；③实践体验原则：引导学生深度体验“综合运用跨学科知识、方法与技术开展探究分析、设计创造和问题解决”的全过程，发展学生的跨学科素养。

3. 设计路径

基于上述设计原则，本文提出了跨学科主题学习的设计路径，如图 1 所示，主要包括：明晰学习目标、精选跨域问题、设计主题任务与学习支架、开展合作实践、交流评价和反思改进六个环节。其中，学习目标的设计要围绕跨学科素养的培养；跨域问题的选择既要为达成学习目标服务，又要能激发学生的好奇心和求知欲，且难度适切；主题任务的提炼要体现主

* 本论文是北京市教育科学“十四五”规划 2021 年度一般课题（立项编号：CDDB21200）的研究成果。

体学科的核心知识与思想方法；子任务的分解及学习支架的设计，要围绕跨域问题的解决过程展开设计并在设计过程中自然融入主、辅学科的知识与思想方法；要通过多元评价的设计全面检验学习目标的达成情况；并依据评价结果进行教学反思和持续改进。

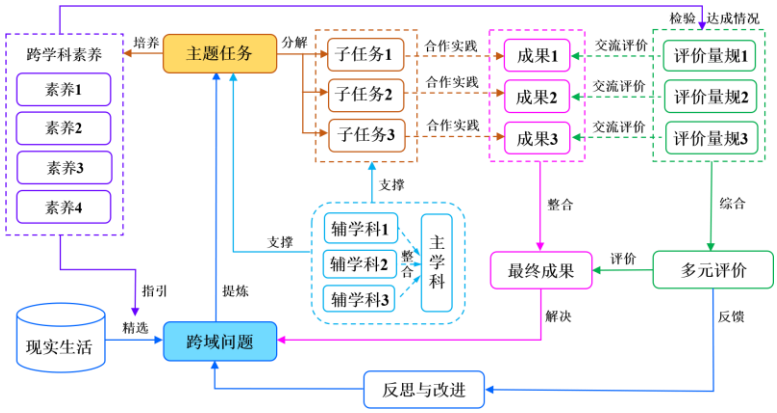


图 1 基于创客教育的跨学科主题学习设计路径

4. 实践应用

本文在小学信息科技课程中进行了实践应用，设计出了一个跨学科主题学习案例，如图 2 所示：首先，从《义务教育信息科技课程标准（2022 年版）》第三学段的跨学科主题学习要求出发，从现实生活中精选了“自动浇花”这一跨域问题，进而提炼出了“自动浇花装置设计制作”这一主题任务；然后依据问题解决过程，将主题任务进一步分解为算法设计、系统开发和外观设计三个子任务；通过学生的合作实践依次完成三个子任务，实现装置的设计与制作，最终解决自动浇花问题。整个学习过程以信息科技学科为主体，整合了科学、数学和艺术学科，能够培养学生的信息意识、计算思维、态度责任和创意创造等跨学科素养。

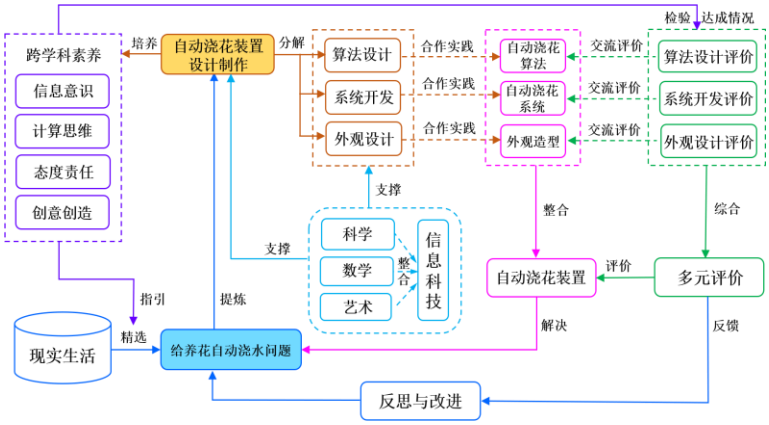


图 2 小学信息科技课程的跨学科主题学习设计案例

5. 结论

本文在创客教育模式的指导下提出了跨学科主题学习的设计原则和设计路径，实践应用表明本文的研究成果是有效的。

参考文献

李锋等(2023).单元视角下的信息科技跨学科主题学习设计与实践.中国电化教育(03),90-95+119.

祝智庭等(2015).创客教育:信息技术使能的创新教育实践场.中国电化教育(01),14-21.

透过师范生生成视频中的隐喻表达揭示对教学中技术运用的认识

Revealing the Understanding of Technology Application in Teaching Through Metaphorical Expressions in Videos Generated by Normal Students

张祺晖[1]，徐珂[1]，章仪[2]，杨九民[1*]
[1]华中师范大学 人工智能教育学部，湖北 武汉 430079
[2]北京师范大学 未来教育学院，广东 珠海 519085
[*]yjm@mail.ccnu.edu.cn

【摘要】 本研究运用隐喻分析法，调查了师范生对教学中运用技术的认识现状。学习者需要通过生成隐喻视频的方式来阐明对技术运用的认识，同时收集创造力、生成视频的认知负荷和学习动机问卷，以及课程实践操作成绩。定性和定量分析数据发现，学生产生了47个不同的隐喻，分为发展、便利、必不可少、动力和威胁五个类别。学习者实践表现在不同类型之间存在显著差异，创造力与关联认知负荷、视频时长之间，学习动机与外在认知负荷之间都存在显著负相关性。总之，学习者对技术运用持有积极正向的态度，有望利用隐喻来分析评估师范生对技术运用的认识。

【关键词】 生成隐喻；意识态度；师范生；课程表现

Abstract: *The study used metaphor analysis to investigate the current understanding of teacher training students in using technology in teaching. Learners clarify their understanding of technology in teaching by generating metaphorical videos while collecting their creativity, cognitive load, and learning motivation for generating videos, as well as practical operation scores in the course. Qualitative and quantitative analysis of data revealed that learners generated 47 different metaphors, divided into five categories: developmental, convenient, essential, motivational, and threatening. There are significant differences in practical performance among learners of different types, and there is a significant negative correlation between creativity and related cognitive load & Video duration, learning motivation, and external cognitive load. Overall, learners hold a positive attitude towards technology applications, and it is expected that, in the future metaphor analysis can be used to evaluate the understanding of technology applications among normal university students.*

Keywords: generating metaphors, consciousness, normal students, learning performance

基金项目：新文科研究与改革实践项目“人工智能+教育背景下卓越中学教师培养的教学模式创新与实践——以现代教育技术应用》为例”（项目编号：2021180023）；教育部产学合作协同育人项目“智慧教学环境下的师范生教学能力培养实验教学体系研究”（项目编号：202101202009）；华中师范大学2021年教师教育专项研究项目“面向课程教学实践的师范生信息化教学能力的培养与提升研究”（项目编号：CCNUTEII2021-01）。

1. 引言

2022年《教师数字素养》教育行业标准的出台，进一步提升了教师利用数字技术优化、创新和变革教育教学活动的意识和能力要求。师范生对信息技术的认识和态度将影响教学实

践中信息技术的应用程度和效果。若具备正确的意识态度，便更可能积极探讨和尝试新的教学方法，提高教学质量。尽管随着学校获得技术的机会和教师相关能力的提高，但在实践中发现师范生对信息化教学及能力提升存在不同程度的畏难和抵触情绪（申仁洪，杨雨露，2020）。在国内的各大师范院校，《现代教育技术应用》作为培养师范生信息技术的主要课程，是师范生学习技术和提高信息化教学能力的重要途径。因此本研究选择在该课程中，通过一项发挥师范生创造性的任务，即让师范生自主生成有关对教学中技术运用的认识的视频、并通过隐喻分析的方式，深层探究师范生对在教学中运用技术的认识，探讨生成视频中的隐喻表达与学生创造力之间的关系，并进一步分析该概念与课程中技术实操表现之间的关系。

师范生对技术的认知及其对教学实践的影响，主要体现在三个方面：对技术基本概念的理解，对在课程中使用技术的态度和看法，未来在教学中使用技术的意愿。目前，多数研究主要通过问卷量表，如 TPACK 量表（焦利剑&钟洪蕊，2010）和技术接受模型（刘伟等，2021），来量化分析师范生的技术认知水平。然而，从师范生自身的认知角度出发，深入探讨他们持有信念的研究尚显不足。这种研究视角的缺失，使得我们难以全面理解师范生对技术的认知及其对教学实践的深层影响。而隐喻的本质是“从另一种“有形”事物的角度来理解或体验一种“无形”事物”（Lakoff & Johnson, 2008）。有形事物本身就是被研究者自己想要突出解释的特征，它能够更加直观准确反映个人对主题或情况的认识、态度或感受。出于这个原因，在社会和行为科学中它已被用作研究工具，以深入了解个人的思维。例如，在教育研究领域，有研究调查了学习、教学、学习者等的隐喻形象（Saban, 2008）。先前的研究（Mustafa, 2013）也探讨了师范生对技术本质的信念和观念，但没有将技术置于教学情景下进一步探究，呈现出师范生利用技术进行教学的特殊性。另外该研究也没能探究师范生生成的隐喻概念类型与学生技术使用表现之间的关系，这一点在本研究中将会进行分析。

创造力通常被视为生产新颖、不同和创新的，但同时又有用、相关且适合手头任务的东西（Runco & Jaeger, 2012）。隐喻构成了创造性思维的核心（Ricoeur, 1981），隐喻处理依赖于将不同的概念领域联系起来并找到两个明显不相关的概念之间的相似性的能力（Bowdle & Gentner, 2005）。以往研究发现，创造性任务的完成效果与认知负荷有关，即在创造性任务中报告高认知负荷的学习者更有可能表现不佳（Redifer et al., 2016）。任务的内在负荷是问题或概念不可改变的核心，外在负荷是额外的，并且会以与学习无关的方式使任务更加困难。当要求学习者向朋友解释某个概念时，可能会导致关联负荷更高（Lakmal Abeysekera & Phillip Dawson, 2015）。与此同时，Paas 等人（2005）将动机量化为较高的认知负荷与学习表现的结合。Mayer（2014）强调，学习环境中的某些组成部分可能被认为增加了不必要的信息或增加了难度，实际上可能通过增加动机来增强学习者的生成处理（即关联认知负荷）。所以在本研究中将会进一步探究在生成视频过程中学习者的创造力、认知负荷和学习动机之间的关系以及对所生成隐喻的影响。

综上，本研究主要探讨以下三个问题：（1）基于隐喻理论的分析框架（详见下文）能够揭示学生所生成视频中的哪些隐喻类别？（2）师范生隐喻类别与课程中技术运用实践操作表现之间的关系是什么？（3）学习者的创造力、认知负荷和学习动机之间有什么关系以及这三者又对生成的隐喻产生了怎样的影响？

2. 研究过程与方法

2.1. 研究对象

《现代教育技术应用》作为师范生公共课程，目的是提高信息化教学能力，为专业发展打下基础。课程分为实验课与理论课两部分，实验课要求学生掌握数字图像采集处理、课件和微课制作等实操任务。本研究重点关注学生信息技术基本技能的掌握情况，所以将实验课

的实践操作的成绩作为评价的重要指标。研究随机选择参加该课程的 86 名师范生作为研究对象，都是来自 2021 级本科生，包含名 5 男生，81 名女生。被试来自不同的专业，包括历史学（HE；91%）、生物科学（BSE；6%）和学前教育（PE；3%）。

2.2. 实验过程

课程共十六周，在课程中期（第七周）要求学生填写问卷并每人提交一个生成隐喻的视频，示例如图 1。问卷内容包括人口统计信息，学习者创造力、学习动机、认知负荷。生成隐喻部分，要求学习者根据提示“技术在教育/教学中就像.....因为.....”，按照自己的理解填写空白部分并拍摄视频，要求不限制视频拍摄的时长。Saban（2008）在研究“知识”概念的隐喻时使用了这种方法。开放式提问被试能够自由反映自己对教学中技术的感受和想法。对于第一个空白要求写出一个隐喻图像，传达他们的理解。第二个空白需要解释为什么选择这些特定的隐喻，进行更详细地阐述。

2.3. 数据收集

2.3.1. 创造力问卷

采用威廉斯创造力测验量表，评估学生的冒险性、好奇性、想象力、挑战性。每个维度平均使用 8 个项目进行评估，每个项目都使用李克特 3 点量表进行评分，范围从 1（“完全符合”）到 3（“完全不符合”）。最终的得分是通过平均每个维度的所有项目的回答来计算的。得分越高，表明学习者创造力越高（Cronbach’s α=0.85）。

2.3.2. 学习动机问卷

本研究采用 Keller 等人（1983）编制的学习动机问卷。该问卷由十个项目组成，其中包含 3 个反向计分项目，为 7 点里克特量表，范围从 1（“非常不赞同”）到 7（“非常赞同”）。分数越高，说明学生通过生成隐喻视频来解释技术在教学中的作用时学习动机越强烈。总分是通过取参与者在十个项目上的回答的平均值来计算的，Cronbach’s α=0.94。

2.3.3. 认知负荷问卷

采用 Klepsch 等人（2017）开发的认知负荷问卷，评估学生的内在、关联和外在三个维度的认知负荷。每个维度使用两或者三个项目进行评估，每个项目都使用里克特 7 点量表进行评分，范围从 1（“非常不赞同”）到 7（“非常赞同”）。最终的得分是通过平均每个维度的所有项目的回答来计算的。得分越高，表明学习者运用生成视频中的隐喻表达来解释技术在教学中的作用时认知负荷越高（Cronbach’s α=0.73）。

2.3.4. 实践操作成绩

学生需要利用借助技术生成七个不同的实践作品，包括多媒体资源搜集整理、交互智能平板教学报告、信息安全加密水印、图像采集处理、多媒体课件制作、微课设计制作、实验设计报告。每位学生的实践操作成绩得分由两位硕士研究生对七个作业逐一进行评分，评分者 KAPPA 一致性为 0.73，最终取所有作业的平均分。（Cronbach’s α=0.70）。

2.3.5. 视频时长

如图 1 所示，学生需要根据自己的理解生成隐喻，并自行拍摄解释视频。由于没有时间限制，所以学生生成视频中隐喻性表达长短不一。研究者将视频无声或无内容部分的删减，视频时长最终为学生隐喻性表达阐述的时间长度，以秒为单位，最短为 5 秒，最长为 164 秒。



图 1 学生生成隐喻视频截图

2.3.6. 隐喻编码框架

隐喻编码分析借鉴了 Mustafa Koc 使用的框架，在之前的研究中，学生的隐喻被不断比较分类，最终汇聚在五个概念隐喻的类别中。在本研究中，隐喻语言的显性类比和实例同样被识别并根据表 1 中的标准进行分类。

表 1 隐喻编码框架

主题	解释	举例
发展	技术是一种现象或工具，具有不断发展和变化的性质	HE-50-F：“技术不断更新换代，持续为教育事业做出贡献新鲜力量”
便利	技术使生活更轻松，另一方面，技术具有矛盾性，给人类带来问题	BSE-3-F：“就像一道任意门，技术让教学跨越的时空限制变得更加便捷与多样化”
必不可少	技术是人类生活的先决条件	HE-25-F:“船是有了桅杆才能快速向前移动，就像教学如果没有技术，就无法很好地进行下去”
动力	此分类 Mustafa Koc 定义为权力。本研究调整为技术是推动教育领域不断变革的动力	BSE-4-F:“技术在教育中就像燃料，因为有了技术，知识的传播速度变得更快。”
威胁	技术是对人类造成威胁的东西	HE-73-F:“划开数字鸿沟，为教育公平带来阻力”

3. 数据分析及结果

本研究首先通过质性分析对视频中学生生成的隐喻进行分类阐述，然后利用韦尔奇检验分析不同隐喻类别学习者实践操作之间的差异，最后通过斯皮尔曼系数探究生成隐喻视频过程中学习者呈现的创造力、认知负荷和学习动机之间的关系以及对所生成隐喻类别的影响。

3.1. 基于隐喻理论的分析框架揭示学生所生成视频中的隐喻类别

为分析学生生成的 86 个视频中的隐喻表达，两个研究者先共同对 20 个视频进行编码，评分者 KAPPA 一致性为 0.82，后续剩下的由其中一位研究者编码完成。根据编码框架对这些隐喻表达分类和比较，共生成 47 种不同的隐喻表达和五个主要的概念类别。表 2 显示了这些类别和相关隐喻出现的频率。

表 2 隐喻编码结果

概念类别	f(%)	隐喻（f）
发展	3（2.94）	地图（1），春笋（1），满载货物的马车（1）
便利	54(52.94)	百宝箱（1），比喻手法（1），餐具（4），齿轮（1），灯塔（5），登山杖（1），调味盐（5），工具（3），明灯（1），纽带（1），屏幕（1），桥梁（8），任意门（1），手（1），手机（1），双刃剑 （12），水流（2），探照灯（1），梯子（1），药品（1），纸箱（1），左膀右臂（1）

必不可少	6（5.88）	帆船上的桅杆（3），甘霖（1），空气（1），书页（1）
动力	37(36.27)	乘号（1），翅膀（1），春风（7），催化剂（7），放大镜（1），风帆（3），花（1），画笔（2），火炬（1），燃料（1），星空（1），养料（1），钥匙（6），引擎（3），增强剂（1）
威胁	2（1.96）	冬风（1）剑（1）

3.1.1.技术作为便利

出现的超过半数隐喻（52.94%）（22 个不同的隐喻）属于这一类别。这些学习者认为技术是人类能力扩展的助手，是教师和学生可以利用的工具，能够辅助提高教学效率。“技术在教学中就像一种工具，常言道工欲善其事必先利其器”（HE-5-F）。“就像一块能够展现教师想法的屏幕，让学生更清晰、高效地接收并理解教学内容”（HE-2-M）。

对于技术在教学中的地位，学习者通过隐喻表达了不同的观点。有学习者强调了技术在教学中的辅助地位，认为“技术在教学中就像是享用佳肴的餐具，教学过程就好比享用的过程。餐具适合教学，会让整个过程锦上添花。但最重要的仍然是佳肴本身”（HE-27-F）。同时有学生认为它是现代教学的重要组成部分，“现代教学就像一台复杂的机器，技术就像其中的齿轮，各个部分相互配合才能共同推动教学的运行”（HE-44-F）。

在第二个常见的类别中，技术被定义为提高人类教学水平的指引者。例如 6 名学生都将技术隐喻为灯塔或明灯，“技术如同一座照亮迷雾的灯塔，引领学生在知识的海洋里遨游，通过强大的信息处理能力和可视化工具，学生深入、直观理解和掌握知识。”（HE-13-F）

另一方面，一些学习者强调了技术在教育中运用的矛盾性，认为技术虽然有明显的好处优势，但是也有可能给教学带来不可预计的问题，造成消极的效果。“双刃剑”的隐喻共出现 12 次，例如“技术可以极大地促进教学效果与学习效率，可以通过在线学习平台与学生进行互动，提供丰富多彩的教学资源和工具等。但过度依赖技术也可能导致教学过程的无效化，导致过度沉迷电子设备，而忽略了实际的社会经验与互动交流。”（HE-1-M）

3.1.2.技术即动力

36.27%的隐喻归属此类，包含 15 个不同的隐喻，其中“催化剂”“春风”“引擎”占比最大，将技术定义为能够推动教学创新变革的动力。“就像催化剂，为学科课程提供了广阔的活动天地，促成教学改革的深化，使教育教学领域发生了更为深刻的巨变”（HE-3-M）。

技术还是丰富课程内容，提高学生学习兴趣的动力。“技术在教学中就像一双翅膀，能够让知识飞向更远的地方，让学习变得更加自由和灵活，因为技术的应用能够激发学生的学习兴趣，提升他们的参与度与积极性”（HE-55-F）。

技术在教学中能够促进教学公平，“就像火炬，因为它给教育资源贫乏的地区带去希望”（HE-20-F）。技术的作用不可预测，“就像星空，赋予了教学无限可能”（HE-29-F）。“更像一把钥匙，它打开了通往世界的大门，让不同领域和文化之间的交流成为可能”（PE-2-F）。

3.1.3.技术是必不可少的

5.88%的隐喻认为技术在教学中是必不可少的，共生成 4 个不同的隐喻。在所有隐喻的内容中，共同的含义是技术在教学中具有先决作用，是不可或缺的。“技术就像空气寻常却不可缺少，我们有必要去追问技术深层次的问题，因为在这个时代技术不仅展示的是繁荣，还有被异化的生活方式”（HE-30-F）。

3.1.4.技术即发展

共有 3 个生成的隐喻被划分至此类别，他们普遍认为，技术具有不断发展和变化的性质，有帮助教学中各种问题的潜力。“技术在教学中的运用就像是一张无限大的地图，因为我们可以在其中寻找到我们想要的答案和方向”（HE-47-F）。

3.1.5. 技术威胁

只有很少部分的隐喻将技术视为威胁因素，此类别学习者认为技术是会对学习者正常学习生活造成严重威胁的东西。例如“技术是一股冬风，会有一些凛冽，因为现在技术滥用的程度。涉及个人隐私泄漏等一系列隐患”（HE-48-F）。

3.2. 师范生隐喻类别与课程中技术运用实践操作表现之间的关系

方差齐性检验表明学习者机试成绩在不同类型之间方差不齐（$p < .001$），所以本研究采用韦尔奇检验，以确定组间均值是否相等。结果表明，组间具有显著差异（$F(4,81) = 55.62, p < .001$）。用事后多重比较（*LSD*）对差异做进一步分析，结果如表 3 所示。总的来看概念类别为“必不可少”隐喻的学习者实践操作表现最好，其次是生成便利隐喻类，再次是生成动力隐喻类，然后是生成威胁隐喻类，成绩最低的是生成发展隐喻类的学习者。如图 2 所示，生成便利（$p = .001$）、必不可少（$p = .001$）、动力（$p = .014$）、威胁（$p = .020$）隐喻类型的学习者成绩都显著高于生成发展类型的学习者，生成便利（ $p = .001$）、必不可少（$p = .033$）隐喻类型的学习者成绩都显著高于生成动力类型的学习者。

表 3 不同概念类别学习者实践操作成绩描述性统计

概念类别	*M*	*SD*
发展	77.5	1.18
便利	93.56	3.61
必不可少	95.23	1.62
动力	88.8	8.96
威胁	92.08	2.95

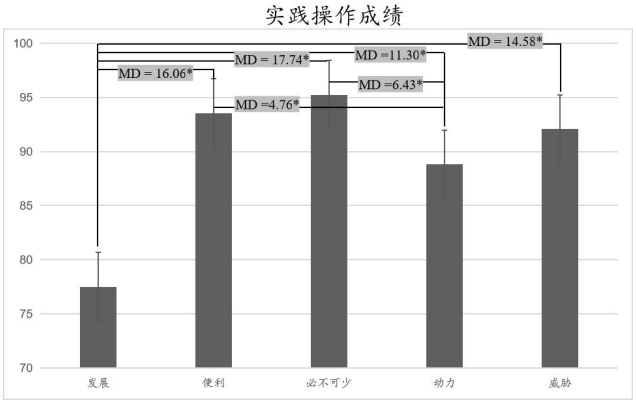


图 2 差异性分析结果

3.3. 学习者的创造力、认知负荷和学习动机之间的关系以及这三者对生成的隐喻产生的影响

由于本研究样本量较少，存在分类变量，并且连续变量不满足正态分布，所以利用斯皮尔曼系数分析被试的创造力、认知负荷、学习动机、视频时长、隐喻概念类别之间的相关性，如表 4 所示。

结果表明学习者的创造力与关联认知负荷存在显著负相关性（$p = .042$），说明学习者创造力越强，在利用隐喻将教学中的技术概念与其他图像进行联系时的认知负荷就越低。与视频时长之间存在显著负相关（$p = .004$），说明创造力越强，生成的隐喻视频长度就越短。通过对学生隐喻内容的进一步分析发现，视频时长较长的学习者所生成的隐喻较为常规，对隐喻解释的时间所占比重较大，解释内容更多阐述了技术对教学的影响而非隐喻图像和技术概念之间的对应联系。

另外学习动机和外在认知负荷之间也呈现出显著负相关性（$p < .001$），说明学习者学习动机越强，在利用生成隐喻视频这样新颖的方式进行学习时接受度更高，认知负荷更低。

结果表明学习者的创造力、认知负荷以及学习动机都对隐喻概念类别没有显著相关性（p > .05）。虽然新颖隐喻生成是对个体创造潜力进行估计的方式（Anat Kasirer & Nira Mashal, 2018），但本研究的分类对学生隐喻的新颖性并无要求，这可能是导致与学生创造力无关的原因。另一方面，本研究对认知负荷和学习动机的测试更多是局限于“生成隐喻视频”的过程，没有涉及课程中技术的使用，这可能也是导致结果无相关性的原因。

表 4 描述性统计与相关性分析结果

	M	***SD***	1	2	3	4	5	6
1.创造力	2.37	5.53						
2.内在认知负荷	3.90	1.35	0.02					
3.关联认知负荷	5.26	0.64	-0.22*	0.12				
4.外在认知负荷	3.28	1.05	0.09	0.40**	-0.14			
5.学习动机	4.12	1.03	0.05	-0.05	0.11	-0.43**		
6.概念类别	-	-	0.11	0.09	-0.11	-0.003	-0.03	
7.视频时长	32.28	26.57	-0.31**	0.14	-0.1	0.06	-0.08	-0.11

注：$^{*}p < .05$, $^{**}p < .01$。

4. 结论与讨论

在以往的研究中，大多数焦点都集中在师范生对技术教育用途的看法或态度上。而本研究通过生成视频探讨了师范生对技术在教学中运用的认识。这种认识不仅可以通过引导师范生的态度来发挥组织和领导作用，而且还将影响他们在教学实践中应用技术的具体方式。我们的目的在于，通过《现代教育技术应用》这门课程，塑造师范生对技术运用的积极信念，从而提升他们的数字素养。研究结果显示，学习者对技术在教学中所能发挥的促进发展和变革作用表达了肯定，展现出对技术学习和应用的积极与正向态度。然而，值得注意的是，在所有隐喻概念类别中，“技术作为便利”占据了最大比例，这表明师范生仍将技术视为教学过程中的辅助和工具性质。这与之前的研究结果相符（De Vries, 2005），学生往往倾向于将技术视为具体的物品（如“餐具”等硬件工具），而不是抽象的概念。此外，部分学生对于技术持有过于乐观或悲观的看法。这种极端态度可能导致师范生将技术视为解决教育问题的简单途径或对正常教育活动的威胁，而忽略了其潜在的挑战和复杂性。

研究探究了师范生的隐喻概念类别与他们在课程中技术运用表现的关系。结果显示，不同类别之间存在显著差异。特别地，将技术在教学中的运用隐喻为“必不可少”的学习者表现最佳，而隐喻为“发展”的学习者表现最差。这可能源于对技术持“必不可少”观点的学习者高度重视技术在教学中的作用，因此在实践操作中更为主动，希望将技术融入自己的教学设计。相对地，将技术视为“发展”的学习者对当前所学的技术认可度较低，导致他们在实践中投入较少，成绩也较低。这一发现不仅提供了通过隐喻分析评估学习者数字素养的意识态度方面的证据，也为未来进一步研究隐喻与实践操作表现之间的关系提供了方向。

在相关性分析中，研究发现了一些重要的结果。首先，学习者的创造力与关联认知负荷、视频时长之间存在显著的负相关关系。这与之前的研究结果相符，隐喻是创造性思维的核心，它需要将不同领域概念联系起来（Ricoeur, 1981）。在隐喻解释过程中，学习者的关联认知负荷会相应提高（Lakmal Abeysekera & Phillip Dawson, 2015）。其次，学习动机和外在认知负荷之间也呈现显著负相关。这一点与 Mayer（2014）的观点有所不同，后者认为增加学习动机可以提高学习者的关联认知负荷。我们推测，这可能是因为本研究中学习者需要生成多媒体材料，而前人的研究是基于学习者学习多媒体材料的情境。未来研究可以进一步探索学习动机和认知负荷在学习隐喻多媒体材料情景和生成隐喻多媒体材料之间的机制差异。

针对本研究研究者在研究过程中的研究实施，研究还存在一些问题与不足，主要表现在：1、被试在性别、年级和专业背景上多样性体现不足。2、针对学习者生成隐喻概念类别的解释还是较为单薄，缺乏相应学习者的深入访谈作为证据进一步佐证结论。

参考文献

焦建利 & 钟洪蕊. (2010). 技术—教学法—内容知识(TPACK)研究议题及其进展. *远程教育杂志* (01), 39-45. doi:10.15881/j.cnki.cn33-1304/g4.2010.01.015.

刘伟, 宋萑 & 黄嘉莉. (2021). 师范生在线教学使用意愿的影响因素及作用机制——基于13所高校师范生的调查. *中国高教研究* (08), 48-55.

申仁洪 杨雨露.(2020).2.0时代下师范生信息化教学能力现状调查及提升策略研究.重庆师范大学学报(社会科学版)(01),80-88.doi:10.19742/j.cnki.50-1164/C.200110.

Anat Kasirer & Nira Mashal (2018) Fluency or Similarities? Cognitive Abilities that Contribute to Creative Metaphor Generation, *Creativity Research Journal*, 30:2, 205-211, DOI: 10.1080/10400419.2018.1446747

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112, 193–216. doi:10.1037/0033-295X.112.1.193

Childress, V. (2006). Teaching about technology: An introduction to the philosophy of technology for non-philosophers.

Keller, J. M. (1983). Motivational design of instruction. Instructional design theories and models: An overview of their current status, 1(1983), 383-434.

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in psychology*, 8, 1997.

Koc, M. (2013). Student teachers' conceptions of technology: A metaphor analysis. *Computers & Education*, 68, 1-8.

Lakmal Abeysekera & Phillip Dawson (2015) Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research, *Higher Education Research & Development*, 34:1, 1-14, DOI: 10.1080/07294360.2014.934336

Lakoff, G., & Johnson, M. (2003). " Metaphors we live by", The University of Chicago Press, Chicago, 60637 The University of Chicago Press, Ltd.

Mayer, R. E. (2014). Incorporating motivation into multimedia learning. *Learning and Instruction*, 29, 171–173.

Paas, F., Tuovinen, J. E., Van Merrienboer, J. J., & Darabi, A. A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, 53(3), 25–34.

Redifer J. L., Therriault D. J., Lee C. S., Schroeder A. N. (2016). Working memory capacity and self-explanation strategy use provide additive problem solving benefits. *Applied Cognitive Psychology*, 30, 420-429.

Ricoeur, P. (1981). Hermeneutics and the human sciences: Essays on language, action and interpretation. (J. B. Thompson, Ed.). Cambridge, UK: Cambridge University Press.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24, 92–96. doi:10.1080/10400419.2012.650092

Saban, A. (2008). Primary School Teachers and Their Students Mental Images about the Concept of Knowledge. *Elementary Education Online*, 7(2), 421-421.

教育数字化转型背景下网络学习空间促进学校变革的影响因素研究[1]

Research on the influencing factors of online learning space promoting school transformation in the context of digital education transformation

王韵仪[1*]，黄雅萱[2]，易诗情[3]，胡蓝心[4]，任光杰[5*]
[12345]华南师范大学教育信息技术学院
1150759337@qq.com

【摘要】 数字化转型背景下，网络学习空间已成为学校教学及管理的重要阵地，研究网络学习空间应用的影响因素有助于更好的开展学校教学及管理，文章从组织变革视角分析网络学习空间应用的影响因素，对广东省各学校信息化主管展开访谈，并对访谈结果进行筛选及构念提取以分析。研究发现影响在线学习环境中的建设与应用影响因素主要包括“信息领导支撑条件”“信息化领导驱动能力”“信息化学校管理”“信息化课堂建设”，其中信息化领导驱动能力是最主要的因素，对其他影响因素具有直接作用。最后，对网络学习空间的应用提出建议。

【关键词】 数字化转型；网络学习空间；信息化领导力；B-L 模型；信息化主管

Abstract: *Under the background of digital transformation, e-learning space has become important in teaching and management. Studying the influencing factors of it will help to better carry out school teaching and management. This paper analyzes the influencing factors from the perspective of organizational change, interviews the Informatization chiefs of schools in Guangdong Province, screens the results and extracts the constructs for analysis. The study found that the influencing factors affecting the construction and application of online learning environment mainly include 'information leadership support conditions ', ' information leadership driving ability ', ' information school management ', and ' information classroom construction '.Information leadership driving ability is the most important, which Directly affect other factors. Finally, some suggestions on the application of e-learning space are proposed.*

Keywords: digital transformation, e-learning space, ICT leadership, B-L Model, Informatization chief

1. 引言

当前我国教育领域正迎来数字化转型关键时期，教育数字化转型正推动教学范式、组织架构、教学过程、评价方式等全方位的创新与变革(祝智庭&胡姣，2022)，经历疫情期间大规模在线教育，师生在线教育的参与度、接受度和应用能力普遍提高，借助网络学习空间开展个性化学习成为未来教育新的方向(唐雪萍&陈丽，2021)； 2021 年，中央网络安全和信息化委员会颁布《“十四五”国家信息化规划》提出 “运用网络学习空间，支撑各类创新型教学的常态化应用（中央网络安全和信息化委员，2021）。”《广东省教育事业发展“十四五“规划》提出以网络学习空间为纽带贯通学校教学、管理、评价等核心业务(广东省人民政府，

[1] 本文系广东省哲学社会科学规划 2023 年度一般项目“数字化转型背景下网络学习空间应用效能动态评估模型构建及应用——以广东省为例”（批准号：GD23CJY20）和华南师范大学科研立项一般课题培育项目“数智驱动的教师教学效能评估研究”（编号：23JXGA15）成果之一。

2021)。”相关应用与政策的密集出台，标志着在教学过程中网络学习空间正在走向常态化应用。《教育信息化“十三五”规划》指出“要在各级各类学校逐步建立由校领导担任首席信息官（CIO）的制度，全面统筹本单位信息化的规划与发展，加强信息化专业队伍建设，确保各级各类学校信息化管理与服务工作得到落实(中华人民共和国教育部，2016)。”信息化主管在学校信息化建设中的作用日益凸显，因此通过信息化主管的视角研究学校网络学习空间应用情况尤为重要。

2. 研究现状

2.1. 网络学习空间研究现状

网络学习空间是中小学校园文化建设及校企合作建设的重要载体，在教学中发挥着重要作用。自教育部颁布《网络学习空间建设与应用指南》(郭炯,黄彬,&郑晓俊，2018)以来，学界开展了大量对其教学应用方式和应用模式创新探究。杨滨等人构建了培养学生问题解决能力的网络学习空间 DPSC 教学应用模式并取得了显著成效(杨滨&汪基德，2018)；广东省中山市基于中山教育网公共云服务体系构建 GERI 教学模式以小学英语课堂为实践载体开展区域网络学习空间的应用创新(魏国营,庄可香&陈玉河，2020)，从国内外研究及一线教学实践上来看，网络学习空间正成为实现学校教学变革的重要实践场地。

2.2 在线学习环境中的教育管理研究现状

Bruce 等人认为信息技术与领导力是双向作用的，信息技术为领导力发展创设一种新的实践环境。当前在运用网络学习空间开展学校管理上，由于管理者对于其认识不到位，导致网络学习空间在学校应用过程中并不能发挥示范引领作用(郭炯&郑晓俊，2019)，这说明管理者的信息化领导力将影响网络学习空间的应用。赵磊磊等人基于 Web of Science 数据库进行文献关键词的可视化分析，发现目前我国的教育信息化管理研究较少且缺乏系统性，并指出信息化领导对于在线学习成效的保障具有关键的推动作用，如何将信息技术运用于领导过程有助于推动我国中小学教育信息化（赵磊磊,梁茜,&李玥泓，2018)。

随着智慧校园建设推进，学校管理人员的信息化领导力逐渐成为信息化建设的重要因素之一(郑禄红,程南清，2020)。Deborah Arnolda 及 Albert Sangrà 对当前在线教学环境下的教育信息技术领导力进行综述，并提出信息技术领导力将成为教育技术研究的第五个时代(ArnoldD,Sangrà A，2020)。Navaridas-Nalda F 等(2020)认为使用教育数字资源是学校数字化转型成功实现的关键因素，需要从战略教学角度特别关注校长及其领导力。组织变革领域专家 Burke 等人构建的 Burke-Litwin 组织成果与变革因果模型 （以下简称为 B-L 模型），改模型探析了信息技术环境下的要素构成，为探究信息化领导者推进学校信息化变革提供了重要的依据。

综上所述，本研究以学校信息化主管作为调研对象，通过调研以期透过信息化主管视角了解当前网络学习空间应用，以 B-L 模型为依据结合信息化主管访谈展开分析，从而提出实现学校信息化变革的有效策略。

3. 研究设计

本研究对学校信息化主管的访谈进行设计，随后进行访谈实施及访谈资料收集，以 B-L 模型作为理论支撑展开访谈编码，随后对数据进行处理得到相关子群并进行分析，如图 1 所示。

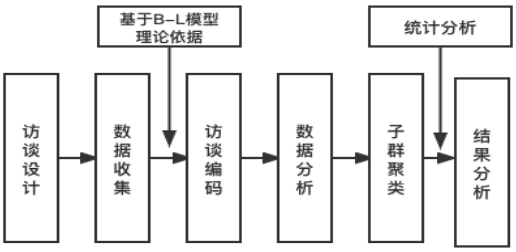


图 1 研究流程

3.1. 数据收集

本研究的调研范围包括广东省 8 个地市，共计 13 个县（区）。采用分层、配额的抽样方法，各试点区选取小学、初中、高中各 1 所。各优秀区域各选取小学 3 所、初中 3 所、高中 2 所。每所学校各 1 位主管作为调研对象。具体数据如表 1 所示。

表 1 访谈对象学段分布表

调研学段	数量（份）	百分比（%）
小学	15	42.9
初中	10	28.6
高中	10	28.6

3.2. 数据预处理

删除与本研究无关或意义表达不明确的关键词，得到 232 个关键词，经数据处理后归纳出 72 个构念，观察图 2 发现，新增构念数目大体呈递减状态，这表明所有构念具有内在一致性。

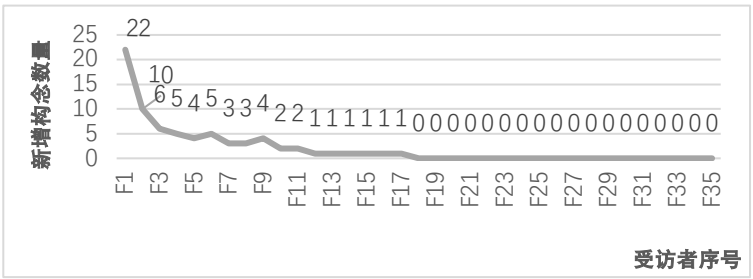


图 2 构念一致性趋势图

3.3. 数据编码

图 3 为组织变革领域专家 Burke 等人构建的 Burke-Litwin 组织成果与变革因果模型（以下简称为 B-L 模型），(MARTINSN,COETZEEM, 2009)其输入为外部环境因素，输出为个人和组织的成果绩效，它可以帮助研究者更好的厘清网络学习空间建设应用及管理过程的要素。

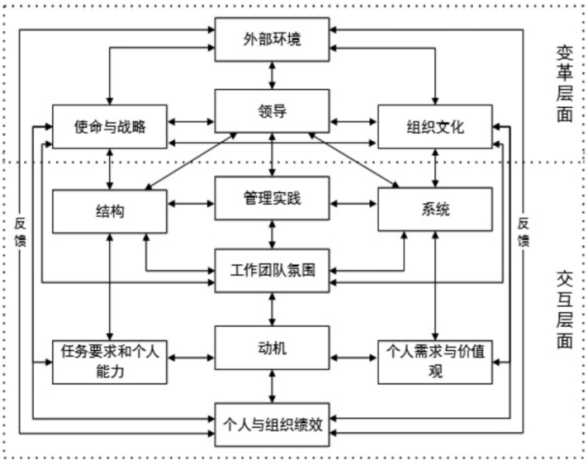


图 3 B-L 模型架构图

依据 B-L 理论模型，研究对构念进行筛选。得出构念维度如表 2 所示。随后对数据进行整理，在数据预处理后形成了基于各构念间共现关系的 37*37 的共现矩阵，如图 4 所示。

表 2 基于 B-L 模型筛选后的构念

构念维度	涉及的构念
外部环境	软硬件支撑、政策引导、专家指导、资金投入
使命与战略	学科融合创新、学校信息化建设 、打造智慧学校 、打造信息化示范学校
业务要求和个人能力	数据分析能力、组织管理能力 、运营监管能力 、支持服务能力 、计算机学科背景 、家校互动 、信息化规划、沟通能力
领导	校级领导牵头、信息化建设规划
管理实践	搭配合作、教师信息化技能比赛、信息化管理模式创新、参与省市信息化建设项目
工作团队氛围	管理互助
动机	信息化课程改革、常态信息化管理、教学模式创新、教学评价改革、网络教研管理模式创新、教育均衡发展、德育、提高课堂效率
个人组织绩效	职称评选、学校评优
组织文化	校本课程建设、学校特色打造、班级文化建设
个人需求与价值	信息技术应用能力提升

	软硬件支撑	政策引导	专家指导	资金投入	学科融合创新	学校信息化建设
软硬件支撑	22	10	12	7	12	8
政策引导	18	18	5	5	9	7
专家指导	19	5	19	6	8	6
资金投入	11	5	6	12	2	4
学科融合创新	18	9	8	2	18	1
学校信息化建设	11	7	6	4	1	11
打造智慧学校	18	12	8	6	9	5
打造信息化示范学校	12	5	7	5	8	2
数据分析能力	23	14	11	8	11	10
组织管理能力	18	8	10	7	8	6
运营监管能力	19	9	12	4	9	9
支持服务能力	12	8	4	5	4	7
计算机学科背景	15	7	8	5	9	4
家校互动	12	7	6	4	8	2
信息化规划	12	7	6	4	6	6

图 4 构念共现矩阵（部分）

4. 研究结果分析

4.1. 基于凝聚子群的网络学习空间建设与应用影响因素构成分析

使用凝聚子群分析方法简化复杂的网络结构，辅助研究者发现网络中的层级关系和子结构，并将网络结构可视化展示。具体而言，本研究对构念共识网络中的构念关系进行了 CONCOR 分析，它是以相关分析为基础发展出来的方法，两个节点间相关系数的计算公式为：

$$r_{ij}=\frac{\sum(x_{ki}-\bar{x}_{.j})(x_{rj}-\bar{x}_{.j})+\sum(x_{ik}-\bar{x}_{.i})(x_{jk}-\bar{x}_{.j})}{\sqrt{\sum(x_{ki}-\bar{x}_{.j})^2+(x_{jk}-\bar{x}_{.j})^2}\sqrt{(x_{kj}-\bar{x}_{.i})^2(x_{jk}-\bar{x}_{.j})^2}}$$

$$\text{for } i\neq k, j\neq k$$

由此得到四个子群分别为：子群一{1，6，21，4，35，26，15，8，12}；子群二{32，7，2，9，28，16，18，24}；子群三{11，13，19，3，22，27，29，25，10，31，36，17}；子群四{5，20，30，33，34，23，14}；同时，标准化计算子群密度，得到如下的子群密度矩阵如图 5 所示：

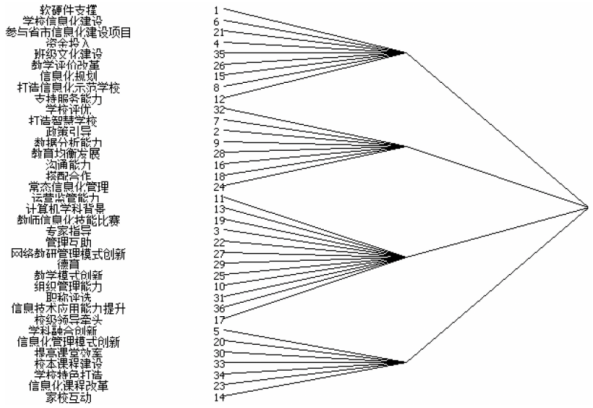


图 5 构念子群划分情况

通过凝聚子群分析可以得知，子群二拥有最大的密度为 8.5，说明该子群拥有在整体结构联系中最为紧密的节点。分析该子群与其他子群的相关系数发现，子群二与其他三个子群都有较高的关联度，说明该子群处于整个共现网络较为中心的位置，起到关联其他子群的作用；子群三与子群二和子群一存在一定的关联度；子群一主要与子群二和子群四存在较为明显的关联度；子群四的密度最小，仅与子群二存在较高的关联度，具体如表 3 所示。

表 3 子群密度相关矩阵

	子群一	子群二	子群三	子群四
子群一	6.347	6.056	5.315	4.889
子群二	8.042	8.5	5.75	6.125
子群三	6.296	5.719	6.523	5.464
子群四	7.381	6.821	5.929	6.048

4.2. 子群特征分析

本研究采用程度中心性分析子群的内部构成以深入了解各子群的概念分布。程度中心性是衡量网络中某个节点是否处于最中心的位置，越趋于中心 ，越有“权力 ”，其影响力也越大。表 4 展示了各子群概念词中心性。

表 4 各子群概念词中心性 （取前三）

	概念	中心度		概念	中心度
子群一	软硬件支撑	63.768	子群三	运营监管能力	36.836
	打造信息化示范校	36.473		政策引导	36.353
	参与省市信息化建设项目	28.986		专家指导	34.420
子群二	数据分析能力	44.686	子群四	学校特色打造	34.300
	打造智慧学校	38.889		学科融合创新	33.696
	沟通能力	25.966		校本课程建设	31.643

4.2.1. 信息化领导支撑条件特征分析

子群一中多数概念体现了“信息化领导支撑条件”对网络学习空间建设与应用的影响，排名前三的概念分别是“软硬件支撑”“打造信息化示范校”“参加省市信息化建设项目”。软硬件支撑是运用网络学习空间开展信息化教学和管理的基础条件，当软硬件支撑足够充分时有利于信息化教学和管理的进行。“打造信息化示范校”“参与省市信息化建设项目”是学校实现信息化改革的重要途径。除此之外，“资金投入”“班级文化建设”“教学评价改革”“信息化规划”等均是信息化领导力提升的重要支撑条件。

4.2.2. 信息化领导驱动能力特征分析

子群二中多数概念体现了“信息化领导驱动能力”对网络学习空间建设与应用的影响，根据各概念的中心度排序，前三名的概念分别是“数据分析能力”“打造智慧学校”“沟通能力”。有学校信息化主管提到“在学习分析、精准教研和个性化人才培养方面，学校利用各项的行为数据和学业数据进行综合分析”由此可见，数据分析的能力在学校的日常管理以及学校教师和学生的评价方面发挥着重要作用。

4.2.3. 信息化学校管理特征分析

子群三中多数概念体现了“信息化学校管理”，根据各概念的中心度排序，前三名分别是“运营监管能力”“政策引导”“专家指导”。运营监管能力主要是指学校信息化主管运用网络学习空间进行在线管理的过程中，使用网络学习空间查看教育用户群体的使用情况。在相关政策的引导下，学校建立规章制度，例如，“每个学期会出一些量化的标准，规范教师使用网络学习空间的行为。同时依托专家指导帮助学校更加科学的开展网络学习空间的教学及管理工作。除此之外，“计算机学科背景”“教师信息化技能比赛”“管理互助”等均在一定程度上影响网络学习空间的应用。

4.2.4. 信息化课堂建设特征分析

子群四中多数概念体现了“信息化课堂建设”对在线学习环境中信息化领导力的影响，排名前三的分别是“学校特色打造”“学科融合创新”“校本课程建设”。学校特色打造主要指的是各个学校依托网络学习空间创建学校特色课程，打造学校专属的文化特色。除此之外，“学科融合创新”与“校本课程建设”也占据着重要的地位。

5. 讨论与总结

5.1. 研究结论

通过收集和处理 35 名学校信息化主管的访谈资料，借助凝聚子群分析和中心性分析，对各个学校的信息化建设及信息化领导能力进行分析。

5.1.1. 明确了网络学习空间促进学校变革的四大要素——信息化领导支撑条件、信息化领导驱动能力、信息化学校管理、信息化课堂建设

通过子群和中心度分析可以发现，信息化领导力支撑条件中，软硬件支撑、打造信息化示范校、参加省市信息化建设项等因素是影响教育信息化领导力的主要方面；在信息化领导驱动能力中，数据分析能力、打造智慧学校、沟通能力等因素是影响网络学习空间建设与应用的主要方面；在学校信息化管理条件中，运营监管能力、政策引导、专家指导等是影响网络学习空间建设与应用的主要方面；在信息化课堂建设条件中，学校特色打造、学科融合创新、校本课程建设等是影响在线学习环境中教育信息化领导力的主要方面。

5.1.2. 探究了影响因素中各子群之间的关系

子群二（信息化领导驱动能力）处于整个子群中较为中心的位置，并且它与其他子群的相关度都比较高，这在一定程度上表明信息化领导驱动能力对在线学习环境中的信息化领导力影响最大。子群一 （信息化领导支撑条件）和子群三（信息化学校管理）在受到子群二影响的同时，二者之间也存在较高的相关度，这说明二者存在较为密切的关系。子群四 （信息化课堂建设） 密度最小，且仅与子群二相互影响，说明信息化课堂建设网络学习空间的建设与应用与其他三个方面相比影响较小。

5.2.教育数字化转型背景下网络学习空间发展建议

结合网络学习空间应用实践，本研究提出如下应用建议。

5.2.1. 信息化领导支撑条件方面

目前我国正在大力推进教育新基建的建设，为优化网络学习空间应用环境提供了基础性保障(柯清超,林健,马秀芳,&鲍婷婷，2021)，推动网络学习空间走向常态化应用。在环境优化的基础上，各级部门及学校应当考虑一线教学实践需要、学校特色及学校师生需求，发挥自身优势，建设符合自身发展的新一代网络学习空间。学校积极保持与上级部门的良性互动，及时总结网络学习空间应用的典型案例，充分发挥学校网络学习空间的内生动力，调动学生、教师、管理者、相关科研人员的积极性，不断探索网络学习空间应用新的突破口。推动网络学习空间试点区建设，总结网络学习空间建设及应用的示范经验，形成网络学习空间教学模式创新、教学评价创新、数据应用创新、教学管理创新、应用机制创新等典型示范案例，鼓励学校将应用成果总结发表，开展专家入校指导，积极开展现场会及其他交流形式，以此推动实现标准化、品牌化发展。

5.2.2. 信息化领导驱动能力方面

进一步明确学校信息化主管的工作范围，提升学校信息化主管的信息领导力、大数据处理能力、创新与文化建设能力、沟通交流与协调能力、战略规划与执行能力 5 方面的能力。在教学的过程中，积极展开数据驱动教学实践：包括数据驱动下的传统课堂教学模式的深度变革、以数据驱动鞭策教师的专业发展、通过数据驱动支撑学生的个性化学习及非认知能力的发展，转变传统的工作思路和流程，树立数字化意识，实现数字思维引领的价值转型，积极展开培训以提高管理者的数据素养，规范数据分析流程以实现各平台数据科学高效应用。

5.2.3. 信息化学校管理方面

由学校中高层领导担任信息化主管，加强学校信息化主管的领导力培训，充分发挥信息化主管在学校信息化建设中的作用，使之成为学校信息化建设的重要推动者(张家年&孙祯祥，2013)。鼓励充分运用信息化手段及工具实现科学管理，推动教育治理体系及治理能力现代化发展。同时加强学校治理化建设，建设领导小组，有效整合各学校信息化项目，促进网络学习空间与其他信息化项目之间的和谐发展。在此基础上，为保证学校信息化的持续发展，建立学校信息化发展的短期、中期及长期规划。

5.2.4. 信息化课堂建设方面

充分考虑学校作为需求方的实际需要，开设定制化服务，鼓励学校将生成性资源上传至网络学习空间，建立供求双向迭代机制实现有效供给。在教育新基建的稳步推进下，积极探索

区块链、机器学习、多模态学习分析等新兴技术支持下的资源供给模式，促进教育资源向智能化的优质新型资源转化。资源推动上，积极推进国家义务教育课程数字教材的规模化应用，同时创建数字化教材的多元应用模式，同时依托“三个课堂”促进教育资源的均衡发展。

5.3. 研究不足与展望

本研究仍存在以下不足：

首先，受客观条件限制，访谈对象多为广东省各学校的信息化主管，且调查样本的数量有限。后续研究可以更科学地选取研究对象，扩大样本范围，进一步提升研究结果的准确性。

其次，本研究仅是一项探索性研究，在后续研究中将借助实证研究对所探知的影响因素构成、各影响因素的影响程度以及不同影响因素之间的关系等方面进行验证性分析。

参考文献：

广东省人民政府.(2016).广东省人民政府关于印发广东省教育发展“十四五”规划的通知. https://www.gd.gov.cn/zwgk/wjk/qbwj/yf/content/post_3602375.html

中央网络安全和信息化委员会.(2021). “十四五”国家信息规划.http://www.cac.gov.cn/2021-12/27/c_1642205314518676.html.

杨滨 & 汪基德.(2018). 网络学习空间DPSC教学应用模式构建研究——网络学习空间人人通促进教与学深度变革实践反思之. ***中国电化教育***(05),44-52.

张家年 & 孙祯祥.(2013).学校信息化主管的素质结构和实践模式研究.***现代教育技术*** (06),18-23.

郑禄红 & 程南清.(2020).智慧校园视野下学校信息化领导力的建设及培养路径.***中国远程教育*** (08),55-61.DOI:10.13541/j.cnki.chinade.

柯清超,林健,马秀芳 & 鲍婷婷.(2021).教育新基建时代数字教育资源的建设方向与发展路径.***电化教育研究*** (11),48-54.

祝智庭&胡姣.(2022).教育数字化转型的实践逻辑与发展机遇.***电化教育研究***(01),5-15.doi:10.13811/j.cnki.eer.2022.01.001.

赵磊磊,梁茜 & 李玥泓.(2018).国外教育信息化领导力研究:主题、趋势及启示——基于Web of Science文献关键词的可视化分析.***中国远程教育*** (10),16-23.DOI:10.13541/j.cnki.chinade.20180711.001.

郭炯 & 郑晓俊.(2019).基础教育领域网络学习空间建设与应用现状、问题及对策研究.***电化教育研究*** (10),89-95.doi:10.13811/j.cnki.eer.

郭炯, 黄彬 & 郑晓俊.(2018). 《网络学习空间建设与应用指南》解读. ***电化教育研究***(08),34-38.Doi:10.13811/j.cnki.eer.2018.08.005.

唐雪萍&陈丽.(2021).新冠肺炎疫情期间公众对中小学在线教学的认识情况分析.***中国远程教育*** (12),53-62.doi:10.13541/j.cnki.chinade.2021.12.007.

教育部(2016).教育部关于印发《教育信息化“十三五”规划》的通知[EB/OL].http://www.moe.gov.cn/srcsite/A16/s3342/201606/t20160622_269367.html.

Arnold D & Sangrà A.(2018). Dawn or dusk of the 5th age of research in educational technology? literature review on (e-)leadership for technology-enhanced learning in higher education (2013-2017). ***International Journal of Educational Technology in Higher Education*** (1). doi:10.1186/s41239-018-0104-3

Navaridas-Nalda F, Clavel-San Emeterio M, Fernández-Ortiz R, et al.(2020). The strategic influence of school principal leadership in the digital transformation of schools. ***Computers in Human Behavior***, 112: 106481.

MARTINSN & COETZEEM.(2009).ApplyingtheBurke-Litwin model as a diagnostic framework for assessing organisational effectiven ess . ***SAjournal of human resource management*** (1):1-13.

教育技术学公费师范生的困惑：职后为何走向分岔路？

The Confusion of Public Funded Normal School Students of Educational Technology: Why to Bifurcation after Graduation?

崔依冉[1]，张书月[2]
[1]山东大学国际教育学院
[2]哈尔滨市第七十三中学校
[*] cuiyirann@163.com

【摘要】公费师范生政策旨在供给优秀师资、促进职前教师专业发展，但教育技术学专业公费师范生因其所对口学科，常常面临职后发展动力不足等困惑。通过深入访谈四位来自同一班级、任教信息技术学科但面临截然不同职后发展现状的教育技术学专业公费师范生，回溯其在高考报考、本科培养、就业选择、职后发展各个阶段所面临的外在环境与心路历程，探讨其不同职后状况的可能原因。结果发现，任教学校对学科的态度、被访者职后心理状态与学科教学知识差异成为主要原因。本研究据此提出建议，以期为教育技术学公费师范生培养与专业发展提供参考。

【关键词】公费师范生；教育技术学；职前培养；职后发展；影响因素

Abstract: *The policy of Public Funded Normal School Students aims at provide excellent teachers and promote the professional development of pre-service teachers, but such students majoring in Educational Technology often face the confusion of insufficient motivation for post-service development due to their corresponding subject. Through in-depth interviews with four publicly-funded teacher majoring in Educational Technology from the same class, who teach information technology, but face different post-service development situation, we retrospectively review the external environments and mental process they faced in the stages of college entrance examination application, undergraduate training, employment selection, and post-career development, and explore the possible reasons for their different post-career statuses. It was found that the attitudes of the teaching schools towards the subjects, the interviewees' post-service psychological state and the differences in subject teaching knowledge were the main reasons. The study put forward recommendations accordingly, to providing reference for the training and professional development of public funded students of Educational Technology.*

Keywords: Public Funded Normal School Students, Educational Technology, Pre-service Training; Post-service Development, Influence Factor

1. 问题提出

为鼓励优秀青年从事终身教育，为教育贫困地区提供优质教师队伍资源，2007 年经温家宝总理提案、国务院决定在教育部直属师范大学实行师范生免费教育政策。公费师范生在校期间免缴学费、住宿费，国家每月为其发放生活费，毕业拥有教师编制，任教后享有回母校攻读在职硕士生的机会。同时，公费师范生在入学前与生源地所在的省级教育行政部门签订协议，承诺毕业后至少从事 6 年中小学教育，其中到城镇教学的群体需要在农村学校任教服务 2 年；若未按协议从事中小学教育工作，则需退还已享受的免费教育费用并缴纳违约金。公费师范生政策“公费培养、提前招生、精心培育、保障就业、继续教育”的特点（杜芳 等，2017），为实现教育强国供给了优秀师资，成为我国教师培养的特色方式。

数字时代，信息技术课程成为中小学生接触、掌握、应用信息技术并逐渐培养数字素养的重要学科（杨晓哲&任友群，2017）。部分省份甚至将信息技术纳入高考学科，使得信息技术这一长久以来的边缘学科似乎被注入了新的活力。信息技术学科教师所对应的公费师范生

多来自于教育技术学专业，但当前尚缺乏关于这一群体的实证研究。已有研究大多关注于追踪公费师范生的就业政策执行情况（商应美&于爽，2018；翟颖佳 等，2022）、职后心理状况与职业认同感（张晓辉 等，2022）、在学期间学业表现（张新亮 等，2023）、从教动机（刘伟&李琼，2022）、职后流动（周利，2018）等。虽然已有大量研究关注公费师范生群体的职后发展，但并未区分具体专业。

现实中，由于教育技术学公费师范生所对口学科存在一定特殊性，部分人职后发展受阻，甚至萌生违约意向；也有部分人坚定于此学科。不禁令人困惑：是什么导致了教育技术学公费师范生职后产生不同的心境？本研究通过访谈四位来自同一班级、接受同一公费师范生培养模式的信息技术学科中学教师，探讨影响该群体职后表现的因素，通过解答“他们为何走向了分岔路？”这一问题，旨在为该群体的培养与发展提供切实的建议。

2. 研究设计

2.1. 样本来源

被访者来自部属师范院校 D，是同一班级的四位教育技术学专业公费师范生，毕业后全部履约前往中学教授信息技术学科，其中两人处于对当前工作极度不满、准备违约的状态；另外两人则处于对当前工作充满信心、坚定从教的状态，具体信息如表 1 所示。样本选择依据如下：第一，公费师范生政策最早实施于教育部直属的六所师范大学，D 大学作为其中之一，培养体系成熟、学生来源与就业省份广泛，分析其毕业生面临的处境与问题，更具典型性。第二，四位公费师范生来自同一班级，意味着接受了相同的在校培养，并且都为女性排除性别原因造成的职后发展差异，由此分析她们截然不同的职后现状、反刍其发展路径，更能够清晰地呈现出不同个人行为产生的影响，研究结果与建议对于学生个体发展更具参照性。

表 1 被访者基本信息

被访者	性别	年龄	省份	任教学校的情况	任教信息技术学科的情况	当前职业满意程度
Case1	女	27	黑龙江	省重点中学	4 年任教信息技术学科，中途未调换学校或学科	非常不满意，多次产生违约意向，受违约条款限制当前仍未违约。
Case2	女	26	山西	省重点中学		
Case3	女	26	浙江	普通中学		满意，坚定任教信息技术学科
Case4	女	27	湖南	省重点中学	2 年任教信息技术学科（毕业的前 2 年履约前往城镇支教，任教数学学科）	

2.2. 访谈过程

本研究采用半开放式访谈的研究方法。由于访谈者与被访者前期已认识，彼此建立了良好的信任基础。从报考教育技术学专业公费师范生的原因、回溯本科培养模式与职后工作的联系、就业时选择信息技术学科的原因、职后发展情况四个方面，对被访者进行深入访谈，访谈时间为每人 1.5-2 小时。征求受访者同意后，访谈过程被录音并被转录为文字，以最大限度保留被访者的原话，最终形成了 4.76 万余字的资料。

3. 研究结果

本部分依据时间顺序，旨在探讨各位被访对象在何时走向不同的发展路线，以及哪些原因可能导致当前的不同职业状态。

3.1. 教育技术学专业公费师范生报考缘由

公费师范生需要在提前批次报考。四位被访者有的“*只知道大概好的地方，比如不用交学费，毕业分配工作，但不知道具体政策是什么，等到大学报到时要签协议，才发现是把自己给‘卖了’*”（case1）；有的“*基本知道政策的内容*”，但也只来源于政策文件的文本信息（case2& case3）。她们在报考公费师范生时都秉持着尝试的心态，在此阶段表述中较多出现的词语是“*报着玩玩*”(case1& case2)，“*多报一项提前批增强录取几率*”(case2& case3)，虽然她们在报考时并不是坚定地希望从事教师职业，但都对教师这一职业处于“可接受”状态。另一位同学“*堂姐是第一届公费师范生，并且毕业后没有去农村，所以提前知道了这个政策的具体条款还有实施的时候是弹性的，综合考虑自己当老师并不排斥，就业形势也比较好*”，所以在已然明晰公费师范生的各项条款以及就业形势后报考了公费师范生（case4）。

前三位被访者最初均报考了大家日常认知的“主科类”科目所对应的专业，后被调剂到教育技术学专业；第四位被访者则是由于操作失误，误将教育技术学专业放到第一志愿。因此四位被访者在知道被录取的专业需要任教信息技术学科时，根据自身求学经历将该学科定义为“边缘学科”，在入学前对所学专业及未来要任教的学科都存在一定程度不满。

3.2. 教育技术学公费师范生本科培养过程

四位被访者接受了相同的本科培养过程，其所在院校所开设课程涉及教育学、计算机、心理学、传播学等学科内容，充分体现了教育技术学的交叉学科属性。当回溯本科阶段的培养模式对其后续职业发展的影响时，四位被访者不约而同地表达了本科阶段所接受的教师技能训练为其后续从事教师职业储备了较坚实的教学基础，但公费师范生这一身份也一定程度上降低了她们本科阶段的学习积极性。诸如“*因为知道自己毕业肯定会当老师,所以本科阶段会很注意教师技能的培养,但也一定程度上会没那么注意努力。*”（case2）“*公费师范生的身份给自己画了一个圈,会在生活中模仿怎么成为一名老师,也会让自己思考如何当老师的问题,但也造成不会思考其他可能性。这其实很大程度上影响到学习动力,觉得拼与不拼的区别会很大吗?不过在大四找工作的时候,才发现拼一下还是挺重要的。*”(case4)

在提及为后续该专业的公费师范生培养提供建议时，两位对当前工作状态极度不满意的被访者不约而同地建议校方重视教育实习，“*一定要让学生在入职前就了解这个学科的就业现状*”（case2），“*加强职前指导,让他们明白如果当信息技术老师会面临什么情况,如果要违约会怎样,帮助大家平稳地完成大学生到就业的过度*”（case1）。另外两位对当前工作状态比较满意的被访者建议校方将专业课程放到首位,如“*信息技术专业知识,信息技术软件操作以及职后学生管理方法*”（case3）；“*建议培养的时候偏重编程,当时学的C语言、数据结构、信息学竞赛这些实用性课程才是最有用的。我们当时学了很多摄录编的课,但一线对于摄录编真的要求不高,我们这边会认为竞赛上带出成绩才能体现教师的能力,但对讲课能力要求其实并不高,所以我们培养的时候一定要结合实际情况多教一些实用性的课程*”（case4）。

3.3. 教育技术学公费师范生就业选择

公费师范生政策在实际就业时具有一定弹性空间，就业时任职于中小学就不属于违约，任教学科是否与专业学科对应不作限定。四位被访者在面临学科选择时，均主动选择了信息技术学科。她们均认为，经过四年学习对于本专业具有一定感情，从事专业对口的学科教学更有信心。部分被访者还考虑到“*信息技术学科是边缘学科,应该会比较清闲*”（case2 & case4），“*不是高考科目,压力小,可以有很多时间做一些自己喜欢的事情*”（case1）。但四位被访者均在入职前对于信息技术学科究竟要做什么、职后发展情形如何“不太了解”。

3.4. 教育技术学公费师范生职后发展

3.4.1. case1 & case2：自我价值感低，不满现状却无力改变

case1 与 case2 两位被访者对当前工作状况不满的原因可归纳为以下几方面：

（1）学科边缘，不受重视，导致杂事特别多

case1 除了教授信息技术课程外，还需兼职信息中心的行政工作，后者占据了主要精力与时间；对其工作现状的描述为“*非常琐碎，没有意义但会干到人崩溃。一天到晚就去做会场布置，跑个腿交材料，发公众号，从这些工作中找不到任何成长不说，还会被来回的挑刺。比如发一篇公众号，一会儿要求改个标点，一会儿去掉一个字，一会儿要换张照片这种。我说让他们把稿子审核好后再给我发公众号，他们说审好了，但我发了之后又来回要这样改。都是小事，没有技术含量，并且特别琐碎*”、“*连开个电脑都需要找信息技术老师去开，我们就是一‘打杂的’*”。case2 同样需要兼职行政事项，对其工作状态的描述与 case1 非常相像：“*太多事情不在教学上，反而行政上的事情占据了主要的精力，但又不能带来价值感，都是杂事……就觉得很沮丧，有时候我会反复怀疑我们这个学科的意义，工作是需要带来价值感和意义的！*”

（2）工资收入水平低

两位被访者都表达了对当前收入水平的不满，虽然所教科目的基础工资与主科教师差别不会太大，但是“*主科老师可以补课*”、“*主科老师排课多，课时费高*”，而信息技术学科“*作为边缘学科，无法营造额外收入*”。

（3）享受教学，关注“育人”，但面临教学效果不佳的无奈

case2 需要教授 14 个班级的信息技术课程，每个班上课频率为 2 周一次，她认为造成自己价值感微弱最主要的原因在于“*没有与学生更深层次的接触和交流*”。case1 即使只需要教授 2 个班级课程，但她也对自己班的学生并不熟悉：“*学生们上课就无精打采的，我能理解他们，别的科都考试，信息技术不考，所以在我的课上就觉得是放松的时间*”。两位被访者面对这一情况，所采取的方式基本一致：“*课程内容选一些和他们生活相关的，比如排版啊、信息检索啊，他们在平时一些课外任务就能用到的*”（case1）、“*每一节课都很用心地备课，希望课上得有趣些，让他们能有兴趣*”（case2），case1 甚至将 40 分钟的课上时间分割为“*20 分钟讲课，剩下 20 分钟让他们自己玩*”，但这些举措的结果仍旧“*不理想*”、“*学生上课还是无精打采的*”。

面对教学效果不佳的现状，两位被访者对此的分析是：“*学生都是功利的，他们知道这个科高考不考，所以课上很难认真听*”（case1）；“*教学是教学相长的过程，教育是与人相处的，老师跟学生之间的交流、接触、互动真的非常重要，但信息技术学科就没法这样。日常中我没有给学生关爱，但悲伤的是，我也没有给的机会，所以我的职业幸福感很低*”（case2）。

面对学生不配合的情况，被访者“*只能表示无奈，无力改变，无从下手*”（case1& case2），“*只能忽视，坚持着把这个课上下来，不管下面的情况了。但自己会很难受，我本科的时候学了那么多教学设计、教学理论、教学方法，到头来我还是在用工业时代的教育方式教信息时代的孩子们处理未来智能时代的工作，你不觉得这本身就很可笑吗*”（case1）；“*我一般在课上会说‘你们再做其他科的作业我就没收啦’，但我其实不太会这样做，或者说因为学科性质的原因我不太敢这样做。像其他科的老师，他们有足够长的时间和学生相处，向他们证明，我是为你好。但信息技术学科并没有这个时间，所以也怕孩子们并没有机会体会到我的用心，反而伤害到他们。我有一次误解了一名学生，事后我知道后想和他道歉的，但两周见一次学生，后来我可能连这个学生在哪个班都不记得了，这件事也一直非常遗憾*”（case2）。

通过访谈，可以发现两位被访者对于教育具有自己的坚持，她们希望在信息技术课上传授信息技术知识、培养学生信息素养，因此她们尽可能地调整教学方式、选取教学内容，而不希望遵循该学科传统的讲课模式，也害怕被同化为“上面讲下面闹，应付公事”的教学方式。但面对学科天然弱势所造成的学生不配合的情况，其无奈不止在于不知道还能如何调整教学优化教学效果的无奈，也是对于自己的满腔教育情怀在现实中碰壁的无奈，正如 case2 所描述的“*有期待才会被伤害*”。虽然造成两位被访者“感觉当前工作没有意义”、“考虑违约”

是源于多种因素,但询问何种原因对其影响最重要时,两位被访者均表示“*价值感缺少是造成当前不满的主要原因*”（case2）、“*钱可以课外想其它办法再挣，但学生的这种状态就让我觉得特别没意思，我会怀疑我自己，我觉得自己这样就是在浪费自己的时间，整个人没有价值感，这是最可怕的*”（case1）。

3.4.2. case3：高考学科，天然优势

Case3 的案例相对特殊，被访者任教于浙江省，新高考改革后，信息技术与通用技术学科作为可选科目被纳入高考范围。即使如此，case3 也无法像其他主科教师一样专职于教学，同样需要兼职教务工作。她当前的工作内容包含“信息技术教师+教务处干事+教科处干事+校园广播铃声设置+社团指导老师”。当询问信息技术教学工作在其全部工作中所在比重时，对方详细计算了不同情况：“*如果不当班主任，包括教学在内的所有信息技术相关工作能占到 80%；如果当班主任，信息技术相关的工作总共占 50%*”。

由此可见，信息技术学科教师兼任其他教务工作在这一学科中属于常见现象，即使在信息技术作为高考学科的省份，教师也只是减少了教务工作的比重，而无法将全部精力集中于信息技术学科相关工作。因此，也能够理解为何提及该学科公费师范生培养建议时，case3 关注于信息技术学科操作性技能培养以及学生管理方法。

3.4.3. case4：从迷茫到突破，寻求职后发展之路

Case4 毕业后先根据学校安排依约前往县级中学教授两年数学学科，从第三年开始回归当前任教的省重点中学继续任教信息技术学科，通过她回顾任教信息技术学科的两年经历，可发现她明显经历了从迷茫到寻求突破、直至当前对学科充满激情与活力的过程。

（1）打破边缘学科瓶颈，依靠信息学竞赛谋求职后发展

在任教信息技术学科之初，case4 也同 case1、case2 一样，面临着教务工作占据大量时间、学科不受重视、没有成就感、想辞职的状况。她的转变来源于周围师姐的启发：“*带竞赛是信息技术学科的趋势，你不要觉得自己对这方面不熟悉就不敢尝试*”，这让她萌生了“*即使这个学科边缘，我也不能不尝试去寻找突破就直接放弃*”的想法。于是她写了关于信息技术学科发展的规划交给校长，希望带领学生参加信息学竞赛。这个举动为她找到了职业转折点，她发现，“*即使之前不了解人工智能，没参加过 ACM 比赛，但我报了这些项目后就会有源源不断的培训*”，这些培训不仅帮助她更好地应对如何带领学生参加竞赛，“*也能开阔自己的眼界*”、“*对这个学科有更高的认识，不断地感觉这个学科其实很有用的*”。与此同时，带竞赛能够得到一部分收入，一定程度上弥补了与主科老师在课时费之间的工资差额。当前，case4 完全脱离出教务的琐事，全心全意专注于信息技术学科教学，教授 15 个班级的信息技术课程、一门校本课以及信息学竞赛。

在访谈过程中，不同于 case1 和 case2 充满沮丧与不满的语气，case4 提起当前工作时充满了兴致和活力，她认为自己找到了教师在信息技术学科中取得发展的法门：“*我们本科教了那么多如何教学，但到了岗位后，学校领导会觉得竞赛上带出成绩才能体现教师的能力，这个学科对老师的讲课能力要求不一定需要那么强*”。这也让她在访谈中不断强调“*教师技能是基础，但不是必备的，大学在培养学生的时候，一定要注重信息技术专业技能的培养！*”并不断反思学科定位：“*我们本科时候一位专业课老师告诉我们‘学科需要做课程整合，我们这个学科在其中是需要做绿叶的’。但我觉得如果一直抱着这个想法是不合适的，这样就是把我们学科放在服务者的角度。很多时候我们自己觉得是弱势学科，但其他人（指其他学科教师）可能并不会在乎，只有你自己在觉得自己是边缘学科，这其实是自我定位的问题。*”

（2）“不要让自己变成打杂的”

几位被访者都提到“*在学校一切和电脑有关的事情大家都觉得是信息技术老师应该管的*”（case2）、“*开个机房都需要信息技术老师去开*”（case1）。面对在其他学校常见的这一

现象，case4 详细阐述了她应对这些问题的方法：一方面得益于学校的相关举措，“*学校每学期都给老师们进行相关培训，比如一学期培训 10 个和信息技术相关的技能，后面如果遇到的是培训讲过的问题，就需要老师们自己来*”；另一方面，case4 也一直坚持“信息技术老师不要让自己变成打杂的”这一理念，“*老师们可能遇到问题，比如不会扫描试卷，信息技术老师可以教你怎么扫描，那你后面再扫描试卷就得你自己来了。我会把怎么扫描试卷整理一个简单的操作步骤，打印出来贴在机器旁边，后面的老师可以根据操作步骤来做。这样我教给你方法了，你就不能说这个活让信息技术老师来做了。再比如录播教室，很多学校都是让信息技术老师给大家录课，信息技术老师真的是要‘累死’啊！我们就把每个学科的年级组长培训好，然后让组长培训组员，我们同样整理一个简化的步骤打印出来给大家*”。case4 将其解决方式总结为“*不要什么都依赖计算机老师，信息技术老师不是万能的。我可以提供培训提高老师们的信息素养，但老师们有问题自己解决，也要给计算机老师留出教学时间！*”

（3）灵活处理教学方式，教学效果良好

面对信息技术课程如何教、学生课堂状态怎样的问题，case4 表示信息技术课程每周一节课，“*学生和老师之间不会产生连结，最开始学生甚至会觉得信息技术老师就是开电脑的*”。在分析了学生不听讲的可能原因后，case4 总结了一套教学方式——“*学生不听讲主要就两点原因，一个是他们觉得这个科目也不考试，所以在机房是逃脱了被束缚的环境，学生的学习纪律并不好；还有就是他们对你讲的内容不感兴趣*”，“*所以我会把网络安全或者网络道德教育放在开学的第一课，一开始就把规矩立好，后面去机房上课就会好一些。我们不高考，没有期末成绩，那就想办法挖掘其他方面让学生获得成就感*”，“*计算机课在设计的时候基本都是任务驱动，所以我每次上课的时候都会告诉学生这次要学一个什么技能，它怎么厉害，让学生感受到掌握了这个技能他就与身边的人不一样，这样学完之后他们就会特别自豪*”。

提及信息技术课程的定位，case4 表示“*信息技术虽然进入新中考科目，但高考还不考，所以还是会出现被占课的现象，但信息技术老师需要对这个有平衡的看法。如果被占课了，信息技术老师需要灵活调节不同班讲课的内容，对应考核方式也会不同*”。虽然 case4 并没有明确提及信息技术课程所肩负的“培养学生数字化学习与创新素养、为迎接数字化时代促进个人终身发展”等育人价值（杨晓哲&任友群，2017），但无疑已将信息技术素养培养融入于学科教学过程中：“*信息技术课它不考试，很多时候可能更多就是给学生留一种印象，让他们知道计算机不是只用来玩游戏的，它还可以有很多很厉害的技能。比如说我之前在课上讲过怎么去水印，后来一次在大礼堂活动，播放了学生自己做的视频，他们学过去水印的学生看到其他没学过的班的视频有的带着水印，就会议论‘他们怎么不会去水印啊’，我当时听到就觉得非常自豪，我觉得学生们有这种意识就可以呀，这就是信息素养的体现呀！*”

4. 讨论与建议

通过整理上述访谈资料，四位报考时对教师职业处于可接受状态且教育技术学非第一意向的公费师范生，在经历相同的本科培养、就业前对信息技术学科教师认知模糊但主动选择该学科任教后，当前面临不同的职后现状。通过回溯其成为教育技术学专业公费师范生到当前的各个阶段，她们走向分岔路的拐点主要发生在职后来自任教学校与个人观念的影响，旨在回答“教育技术学公费师范生职后为何走向分岔路？”这一初始研究问题。下述将从教育技术学公费师范生在报考、培养和职后发展不同阶段提出建议，以期帮助本专业公费师范生从选择伊始即做好准备，以良好姿态迎接职后发展。

4.1. 做好公费师范生的高考报考普及工作

很多公费师范生在报考之初对于政策的了解并不深入，可能因为希望多一个录取机会而报考公费师范生。这一政策对于立志从事基础教育的考生非常适合，但对于对未来职业并不

坚定的考生而言，当其毕业后不愿意当老师时面临 1.5 倍返还培养金额以及诚信档案失信的代价，成为很多公费师范生对当前工作不满却又没有辞职的原因。为了减少考生在择业时亦或任教后面临此两难问题，建议高中学校在高考生报考时即对公费师范生政策加以解读，尤其对于其享受的福利与需履行的义务进行详细说明，让考生在切实了解具体条款的前提下再理性思考、做出选择。

4.2. 国家适当调整各省名额分配，信息技术学科公费师范生供需平衡

虽然四位公费师范生未面临这一情况，但她们在访谈过程中根据所了解到的周围同学的就业状况，提及部分省份存在公费师范生供过于求的现象。各个省必须要“消化”公费师范生指标，但部分地区实际并没有相关岗位缺口，尤其信息技术学科课时量相对较少，并不需要额外的信息技术教师从事教学工作，这也导致所招聘的信息技术学科公费师范生职后无课可上、只能从事教务等“杂活”。

4.3. 国家适当建立公费师范生中途准入准出机制

当前公费师范生的职后服务年限为 6 年，但对于并不愿意从事教师职业的公费师范生而言，这意味着不违约的前提下需要接近 30 岁才能够转向其他岗位，换言之，18 岁左右的高考生在报考时即决定了自己未来 10 年的发展道路。如本次访谈中的 case1，即使只有两年多就可合约期满，但宁愿凑钱违约也不愿继续“耗费青春”。因此建议国家在培养阶段就建立适当的退出机制，师范生在本科培养过程中随着对教师岗位内容的了解，部分不愿未来从事教师岗位的学生可退出公费师范生群体，同时补缴相应学费、住宿费，退还所领取的补助费用。对于中途退出公费师范生所带来的名额，可依据自愿原则由立志从事教师岗位的非公费师范生进行补入，给予学生二次选择的机会。让合适的人找到合适的位置从事合适的工作，一定程度上避免内心充满不满但囿于违约代价的公费师范生职后面临两难抉择。

4.4. 本科培养应结合学科实际教学需要，培养学生契约精神

依据埃里克森的自我同一性理论，本科阶段是一个人发展自我同一性的关键时期，即认同自己与未来社会生活的关系，对自己的信仰及后续社会发展达成较为一致和完善的意识的重要阶段（郑宁，2000）。报考公费师范生的群体大多对当老师并不排斥，但报考生多会结合自身受教育经历，认为信息技术学科是边缘学科，很少主动报考。通过访谈，教育技术学公费师范生群体面临的困境更多源自学科特性。建议本科院校帮助该专业公费师范生建立对信息技术学科的全面认识，明确该学科在数字时代具有独特价值，而不要因边缘学科的刻板印象在任教之初即抱有学科偏见；同时学校培养应结合实际需求，给予学生更多教育见习与实习的机会，通过邀请各省份已入职的信息技术学科教师分享一线教学经历，帮助学生在入职前全面了解该学科教师真实工作内容，为其在就业时理性选择任教学科提供指导。

值得注意的是，这些公费师范生既然对口信息技术学科，学校应做的是努力为他们提供切实需要的技能培养，而非让其认为信息技术教师不受重视而逃避任教该学科。因此，学校一方面可帮助学生理性分析该学科突破瓶颈的方式，例如 case4 依靠信息学竞赛实现职后发展，对应地需要学校提供信息技术专业技能的相关培养。另一方面，学校应正视当前信息技术教师可能需要兼职管理岗位的普遍现象，为学生提供学校管理的相关课程和培训，学生可根据生源所在省份信息技术教师的就职信息，选择所需课程，顺利实现职后过渡。

4.5. 通过相关规定规避中小学校信息技术教师兼任教务的“常规操作”

依据自我价值取向理论，个人对于事物的价值观会影响其后续的行为（Jacobs et al., 2005）。所以不仅本科学校在培养教育技术学专业公费师范生时需要让其明确该学科在数字时代具有独特价值，还应通过相关规定让中小学校明确：信息技术教师具有其学科任务，而非“打杂的”。当整个学校树立起“信息技术学科与其他教学学科处于同等位置”这一观念后，不仅信息技术教师可以规避很多如开计算机教室门、帮忙跑腿打印等无关琐事，也能够

促使中小学生树立起对信息技术学科与信息技术教师的尊重，反应到其课堂行为上，一定程度上能够改善学生在信息技术课上打闹、不听讲的现象。

4.6. 本科学校与信息技术学科公费师范生共同发力，提升学科教学知识

学科教学知识（Pedagogical Content Knowledge，简称 PCK）指教师将某学科的专业知识和教学必备的教育学知识进行整合，形成新的综合理解，知道如何根据学习者的兴趣和能力将特定主题、问题和事件组织起来，以教学的方式加以呈现（石耀华&张良，2018）。除去 case3 任教省份将信息技术学科纳入高考范围，学科教学具备天然优势，其余三位被访者都面临学生对信息技术学科重视不足、师生交互频率较低的现象。case1 和 case2 非常努力地希望通过更加丰富的教学方式提升教学效果，但面对学生无积极反馈的现状，更多将其归咎为“*应试教育环境的悲哀*”（case2）、“*我们本科所受到的培养更关注于育人的角度，但实际教学环境重视的是知识传授而不考虑育人*”（case1）。而 case3 最初面对同样棘手的教学状况，在教学过程中能够灵活处理，对于“被占课”所造成的各班进度不一致的情况，灵活调整各班教学内容与考核方式，实现学科知识传授，甚至达成信息素养的学科培养目标。学科教学方面，三位被访者的差别在于后者具备更加完善的学科教学知识（达成 PCK），虽然前两位被访者都学习了丰富的教学理念与教学设计模式，具备良好的教学知识和信息技术学科知识，但对于如何将二者融合实现“以合适的教学方法教授适当的学科知识”仍有待进步。

因此，建议本科培养阶段，不仅要教授公费师范生信息技术专业知识（学科知识）与各类教学理念、教学模式（教学知识），还应适当开设相关实训课程，帮助学生了解如何将教学模式应用于信息技术学科教学中，达成学科知识与教学知识的有效整合（学科教学知识）。已就业的公费师范生也应及时转变观念，不固守书本上所习得的教学理念，而通过自身实践、职后培训、旁听优秀课程等多种途径，尽快达成学科教学知识（PCK）。

参考文献

Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J. S., & Malanchuk, O. (2005). I can, but I don’t want to. The impact of parents, interests, and activities on gender differences in math. In A. Gallagher & J. Kaufman (Eds.), Gender difference in mathematics, 246-263.

杜芳, 王上艳, 刘心畅（2017）. 免费师范生培养:现状、反思与对策——基于对华中师范大学历史文化学院 2011—2013 届历史专业免费师范生的调查. 教师教育论坛, 30(07):46-52.

刘伟, 李琼（2022）. 为何从教：公费师范生与非公费师范生从教动机的多组潜类别分析. 中国高教研究, (10):61-67.

商应美, 于爽（2018）. 免费师范生就业政策执行跟踪研究:现状·成效·举措——以东北师范大学五届免费师范毕业生为例. 东北师大学报(哲学社会科学版), (05):195-199.

石耀华, 张良（2018）. 集体备课促发教师 PCK 的逻辑理路与实践启示. 天津师范大学学报(基础教育版), 19(01):9-14.

杨晓哲, 任友群（2017）. 高中信息技术学科的价值追求:数字化学习与创新. 中国电化教育, (01):21-26.

翟颖佳, 封世蓝, 谭娅（2022）. 基于教学质量的公费师范教育政策研究. 教育科学, 38(01): 49-56.

张晓辉, 姚梅林, 李庆安（2022）. 政策满意度、教师支持对公费师范生职业效能和职业认同的影响：一项纵向研究. 北京师范大学学报(社会科学版), (06):50-59.

张新亮, 石艳, 郑琦等（2023）. 公费师范生学业表现的实证研究——基于某部属师范大学的行政数据. 教育学报, 19(01):165-181.

郑宁（2000）. 从埃里克森自我同一性理论看当代大学生人格发展. 北京建筑工程学院学报, (04):105-112.

周利（2018）. 免费师范生毕业后工作流动状况研究. 南京师范大学.

后疫情时代大学生翻转学习需求对自我效能感的影响研究

Study on college students' flipped learning needs and self-efficacy in post-epidemic era

叶禹杉[1*]，陈鹏[2]，白佳鹭[3]
[1]首都师范大学教育学院
[1*]yeyushan164@163.com

【摘要】研究对翻转学习环境下大学生的学习需求与自我效能感之间的关系进行了深入探究。通过问卷调查的数据分析，研究发现在翻转课堂教学模式中，大学生的学习需求与其自我效能感存在显著的正向关联。具体来说，当学生的个性化学习需求得到更好的满足时，他们在学业任务上的信心和能力感知得到了增强。这一结果表明，翻转学习模式不仅提供了更加灵活和自主的学习环境，而且通过增强学生的参与度和探索性学习，也有效地提升了学习的自我效能感。本研究对于理解和优化翻转课堂教学模式、提高教育质量及学生的学习效果具有重要的理论和实践意义。
【关键词】 翻转学习；学习需求；自我效能感；大学生

Abstract: *This study delves into the relationship between university students' learning needs and self-efficacy in a flipped learning environment. Through questionnaire surveys and data analysis, it was found that there is a significant positive correlation between students' learning needs and their self-efficacy in the flipped classroom teaching model. Specifically, when students' personalized learning needs are better met, their confidence and perceived ability in academic tasks are enhanced. These findings indicate that the flipped learning model not only provides a more flexible and autonomous learning environment but also effectively enhances students' participation and exploratory learning, thereby significantly boosting their self-efficacy. This study is significant for understanding and optimizing the flipped classroom teaching model, improving the quality of education, and enhancing students' learning outcomes.*

Keywords: Flipped Learning, Learning Needs, Self-Efficacy, University Students

1. 前言

随着技术的飞速发展及教育的不断变革，翻转学习已引起全球范围内的广泛关注和深入研究。这种模式最初由美国“林地公园”中学的伯尔曼和亚伦在 2007 年提出，他们通过网络技术使学生在课堂之外就能接触和吸收课程内容，从而在课堂上更加专注于知识的深入探讨和应用(任彩怡,&李雪梅, 2022)。此后，翻转学习的理念得到了萨尔曼等人的进一步推广，在 2011 年的 TED 会议上，其提出的观点使得翻转学习成为教育领域的热门话题(张金磊, 2012)。翻转学习的核心在于强调学生的主体性和自主性，在学习过程中，学生通过观看微课视频等材料，自主完成课前学习，而课堂时间则用于深化理解、小组讨论和项目式学习。这种模式的实施有效地转变了传统教育中教师的角色，从知识的单向传递者变成了学生学习过程中的引导者和协助者(饶扬德,&邓辅玉, 2022)。我国在采纳翻转学习理念的过程中，研究者主要聚焦于从自主学习、课堂教学、微课等方面探索如何更有效地实施翻转课堂，并强化教师在教学实践中的内化性教学(卜彩丽, 2016)。对于大学生而言，在翻转学习环境中，如何有效提升其自我效能感进而提高学习效果，成为了一个重要的研究议题。我国学者李彤彤等(2018)采用元分析方法对翻转课堂实验和准实验研究进行了系统定量分析，发现翻转课

堂对提高学生学习效果具有积极影响，但对其改进方法未做相关阐述。班杜拉的自我效能理论为理解和提升翻转学习的效果提供了重要的理论基础。本研究旨在探讨大学生在翻转学习环境下的需求及其对自我效能感的影响，以期通过剖析二者之间的关系挖掘提高学生自我效能感的要素，从而促进翻转学习模式在高等教育中的有效实施和发展。

2.研究设计

2.1. 研究对象

本研究主要调查大学生的学习需求与自我效能感之间的关系。研究以某大学本科生为调查对象，被调查者均有过翻转课堂学习经历。参与调查的学生共320名，均为全日制在校生，分布在各个年级。回收有效问卷310份，有效率96.8%。其中男生128名，女生182名，大一66名，大二76名，大三80名，大四88名。

2.2. 调查工具

本研究的调查工具为《基于翻转课堂的大学生学业自我效能感及学习需求调查问卷》，包括基本信息、自我效能感、学习需求三部分。其中，自我效能感部分主要参考梁宇颂（2000）编制的《大学生学业自我效能问卷》。问卷为李克特五点量表，共20个题目构成，包括学习能力自我效能感和学习行为自我效能感两个维度子问卷，相关问题均以翻转学习经历为前提。“学习能力”和“学习行动”子维度问卷的克隆巴赫系数分别为 0.820、0.752，表明问卷的信度较高(梁宇颂,2000)。该问卷的学习需求部分主要参考吴金鑫(2021)基于Kano模型的《师范生学习情况调查问卷（第四部分学习需求）》，包括教师支持、实践支持与内容支持三个维度，共24个题目。

2.3. 调查问卷信效度检验

本研究的KMO值为0.873，显著高于0.6的标准阈值，表明所收集的数据非常适合进行因子分析。与此同时近似卡方：854.749；自由度(df)：190；P值：0.003表明研究数据集非常适合进行因子分析。

本研究信度检验主要采用Cronbach’s α系数。经计算核验标准化Cronbach’s α系数为0.905，表明问卷具有非常高的信度。

综上所述，本研究的调查问卷在效度和信度方面均表现出色，数据集具有高信度，适合用于进一步的数据分析。

3.数据结果分析

由表1可看出，学业自我效能感量表每一道题的平均数是3.753，大学生的学习能力自我效能感平均数高过学习行为自我效能感。就标准差而言，学习行为自我效能感低于学习能力自我效能感。

表1 学业自我效能感总体状况

名称	项目数	M	SD
学习能力自我效能感	11	3.771	1.015
学习行为自我效能感	11	3.736	0.993
学业自我效能感	22	3.753	0.989

对不同年级大学生的学习能力自我效能感和学习行为自我效能感这两个指标进行单因素方差分析可知(表2)，大三学生的自我效能感得分最高，各个年级的大学生的两指标未见显著差异，$p>0.05$。

表 2 学业自我效能感在年级上的差异检验（M±SD）

分类	大一（n=66）	大二（n=76）	大三（n=80）	大四（n=88）	F	p
学习能力自我效能感	3.658±1.105	3.823±1.086	3.901±1.008	3.692±0.89	0.948	0.418
学习行为自我效能感	3.714±1.134	3.785±1.026	3.832±0.983	3.624±0.858	0.692	0.558

对男女大学生的学习能力自我效能感和学习行为自我效能感这两个指标进行独立样本 t 检验，从分析结果（表 3）可知学业能力自我效能感和学业行为自我效能感两个维度在性别上均为表现出显著差异，p>0.05。男生的两个指标的平均值均高于女生。

表 3 学业自我效能感在性别上的差异检验（M±SD）

分类	男（n=128）	女（n=182）	t	p
学习能力自我效能感	3.776±1.031	3.728±1.025	-0.730	0.466
学习行为自我效能感	3.766±1.004	3.744±0.967	-0.106	0.915

对学习需求与自我效能感进行相关分析（表 4），结果表明学习需求与学习能力自我效能感呈显著正向相关关系(r=0.450，p<0.01)；学习需求与学习行为自我效能感呈显著正向相关关系(r=r=0.390，p<0.01)。进一步将学习需求作为自变量，自我效能感作为因变量进行回归分析，得到模型公式：自我效能感=2.257 + 0.372*学习需求，R 方值为 0.882。模型通过 F 检验(F=17.188，p=0.003<0.05)，说明学习需求会对自我效能感产生影响关系。学习需求的回归系数值为 0.372(t=4.146，p=0.003<0.01)，意味着学习需求会对自我效能感产生显著的正向影响关系。

表 4 相关分析

变量对比	皮尔逊相关系数（r）	显著性水平（p）
学习需求与学习能力自我效能感	0.450	p < 0.01
学习需求与学习行为自我效能感	0.390	p < 0.01

4. 讨论与总结

4.1. 学业自我效能感分析

本研究首先对大学生学业自我效能感进行了描述统计分析，学业自我效能感、学业能力自我效能感和学业行为自我效能感的均值分别为 3.711、3.736 和 3.753。说明被调查者的总体学业自我效能感处于较高水平，大学生对自身能力信心较强。另外，在结果中学业能力自我效能感均值略高于学业行为自我效能感均值。说明学生对自己能否顺利完成毕业、取得良好成绩和避免学业失败的能力具有信心。

在性别这一维度上，男女大学生在学业自我效能感上未见显著差异。这与李芳的研究一致(李芳,2020)。其研究说明了男女学生在学业自我效能感上无显著差异。究其原因，这可能和我国时下教育体制息息相关。此次调研的主体均为大学生，都是经过高考选拔过的高等教育学生，普遍有属于自己的学习方法，学习能力也较强，愿意为了提高成绩而付出努力。

在年级这一维度，大一、大二、大三和大四的大学生在学业自我效能感上未见显著差异。这和梁宇颂的调查结果保持一致(梁宇颂,2000)。但具体而言，大三学生的学习能力自我效能感和学习行为自我效能感平均值均高于其他年级。这可能与大三大部分学生准备考研、考公或者考编有关，所以展示出较高的学习能力自我效能感和学习行为自我效能感。

4.2. 学业自我效能感与学习需求相关性分析

翻转学习背景下，研究结果显示大学生的学业自我效能感与其学习需求之间存在明显的正向关联。首先，学业自我效能感对学习需求产生显著影响。这意味着大学生对自己能否成功完成学业任务的信念越强，其对学习的需求就越旺盛。高学业自我效能感的学生通常具有更强的学习动机，他们会设定更高远的学业目标，对学习内容表现出浓厚的兴趣，而且在面对测验考试等挑战时更有信心，这些因素共同促使他们在学习活动中表现出更高的需求。其次，本研究通过回归分析确认学业自我效能感的两个维度在预测学习需求方面具有较强的效度，这两个维度能够有效解释学习需求的88.2%的变化，显示学业自我效能感是一个强有力的预测变量。为了确保研究结果不受共同方法偏差的影响，采用了 Harman 单因子法进行了探索性因子分析。结果显示，共同方法变异的解释量仅为37.639%，低于临界值50%，这表明研究结果主要反映了学业自我效能感与学习需求之间的实质性关系，而非单纯的数据收集方法导致的共变性误差。结合学者袁桂平（2010）的研究成果，再次验证了学业自我效能感对学习需求的显著预测作用，增强了本研究结论的可靠性。综上，本研究认为提升大学生的学业自我效能感对于增强其学习积极性、明确学习目标、调整学习策略以及提高学习需求至关重要。教育者在日常教学与管理中，除了关注学生的学业成绩外，还应采取有效措施来提升学生的学业自我效能感，比如通过个性化指导、反馈机制的优化、成功案例的分享以及学习策略的培训等手段，以期在翻转学习环境中更好地激发大学生的学习潜力和需求，从而全面提升其学业成就。

4.3. 总结

翻转学习作为一种创新教学模式，不仅改变了教学结构，而且在实践中显现出了提升大学生学业自我效能感的功能。教育者在应用翻转学习时，需格外关注对学生个体差异的认知与满足其个性化学习需求，这对于培养和发展学生的自我效能感至关重要。未来研究方向应当着重探究自我效能感在翻转学习各个阶段的具体影响机制，以及如何更高效地在国内不同教育环境中推广和优化翻转学习实践。本研究深化了对翻转学习与大学生学业自我效能感互动关系的认识，为教育研究与实践提供了理论支撑和前瞻视角，旨在通过优化教学策略，提升学生的自我效能感，进而促进其学术成就与全面发展。

参考文献

卜彩丽(2016).孔素真.现状与反思:国内翻转课堂研究评述[J].中国远程教育,2016,No.493(02):26-33.

任彩怡,&李雪梅.(2022).国内翻转课堂教学模式下大学英语教学的文献综述.海外英语,3(24),109-111.

李彤彤, 庞丽, & 王志军. (2018). 翻转课堂教学对学生学习效果的影响研究——基于 37 个实验和准实验的元分析. 电化教育研究, 39(5), 9.

吴金鑫.(2021).基于服务设计的师范生《教学技能训练》课程学习需求分析[D].华东师范大学,2021.

张金磊(2012).王颖,张宝辉.翻转课堂教学模式研究[J].远程教育杂志,2012,30(04):46-51.

饶扬德,&邓辅玉.(2022).深度学习视角下翻转课堂教学模式设计——以创业学通识课为例.安徽职业技术学院学报,5(04),82-86.

袁桂平. (2010). 大学生学业自我效能感与学习适应的相关研究. 临沂师范学院学报(3), 4.

梁宇颂.(2000).大学生成就目标、归因方式与学业自我效能感的研究[D].华中师范大学,2000.

Chemers, M. M. , Hu, L. T. , & Garcia, B. F. . (2001). Academic self-efficacy and first-year college student performance and adjustment. Journal of Educational Psychology, 93(1), 55-64.

生成式人工智能辅助大学生自主学习的风险挑战及应对策略

Risks, Challenges and Coping Strategies of Generative AI Assisted in Self-Directed Learning of College Students

路畅，田阳
（扬州大学新闻与传媒学院教育技术学系）
1661265667@qq.com

【摘要】 生成式人工智能一经问世，便迅速被广大高校学生使用。针对大学生自主学习方面，大学生能否借助生成式人工智能来辅助自主学习成为研究议题。本研究利用 SWOT 分析，对生成式人工智能辅助大学生自主学习进行了分析。发现生成式人工智能给大学生自主学习带来了优势，机遇，但同时也带来了劣势以及威胁，某种意义上形成了风险挑战。由此，针对生成式人工智能辅助大学生自主学习的风险挑战，分别从社会、教师、学生层面提出了相应的应对策略。

【关键词】 生成式人工智能；ChatGPT；大学生；失范行为

Abstract: *As soon as generative AI was introduced, it was quickly used by a large number of college students. In terms of college students' self-directed learning, whether college students can use generative artificial intelligence to assist self-directed learning has become a research topic. In this study, SWOT analysis was used to analyze the self-directed learning of college students assisted by generative artificial intelligence. It is found that generative AI brings advantages and opportunities to college students' self-directed learning, but it also brings disadvantages and threats, which poses a risk and challenge in a sense. Therefore, in view of the risks and challenges of generative AI-assisted college students' self-directed learning, corresponding coping strategies are proposed from the levels of society, teachers and students.*

1.引言

美国 OpenAI 研发的生成式聊天机器人程序，一经发布就以前所未有的速度火爆全球，并广泛渗透影响到产学研各界。生成式人工智能改善教育环境的同时，也带来了一系列严峻的道德挑战和法律风险，对于这些风险挑战，人类应该做的是建立一套有效的制度来规范这些技术的使用，而不是一味地一禁了之。[1]生成式人工智能可能带来的机遇和挑战就像“阿拉丁神灯”和“潘多拉魔盒”。[2]虽然生成式人工智能存在一些潜在的风险，但对教育的影响既取决于我们的立场和视角，也取决于我们的目的和意图，归根结底在于教师和学生如何使用它，学校应该趋利避害，使生成式人工智能成为学校教育的朋友。[3]

2.生成式人工智能辅助大学生自主学习的内在逻辑

2.1 大学生应用生成式人工智能的机会与前景

2.1.1. 有利于个性化学习

每个学生都有其独特的学习风格和能力，而生成式人工智能技术可以根据学生的学习特点、学科需求和学习进度，提供量身定制的支持和辅导，提供个性化的学习内容和学习方式。[4]生成式人工智能还可以为教师提供个性化的教学辅助。它能与教师实现协同工作，辅助教师进行课程设计和材料生成，为教师

生成有针对性的教学材料、教学大纲、教学案例等。

2.1.2. 作为智能辅导系统

生成式人工智能可以根据学生的个人情况、学习风格和能力水平，为其提供个性化的学习计划和辅导。生成式人工智能可以依据学生的学习需求与知识掌握情况，为学生提供动态教学支架与反馈；还可以结合学生实际需求与编程任务，为学生推荐相关代码片段与运行解析，帮助其高效理解程序的设计思路与编写方式。生成式人工智能可以作为智能辅导系统，为学生自主学习提供实时的、有针对性的学习支持和指导。

2.1.3. 作为评估反馈工具

利用生成式人工智能技术，可以构建评估与反馈系统，帮助学生对自主学习的成果进行客观、准确的评估，并提供及时的反馈和建议。这大大减轻了教师的评分负担，提高了评估效率。生成式人工智能可以对学生的学习数据进行分析，提供学习进度、知识点掌握情况等方面的反馈。通过分析学生的学习数据，系统可以发现学生的学习难点和潜在问题，为教师和学生提供有针对性的建议和解决方案。

2.2 生成式人工智能辅助自主学习的内在逻辑

自主学习是学习者针对自己所设定的教学目标，充分利用自身的主客观条件，激发自身的积极主动性，并积极探索和选择适合自己的学习方法，从而达到完成学习任务的具体过程。这就要求学习者对自己本身的能力有一个准确的评估，还需要坚持、恒心、毅力等内在品质帮助学习者实现自主学习。[5]

在传统的学习方式中，很多大学生可能只是停留在表面，对学习内容缺乏深入的探究和思考。他们往往满足于了解基本概念和知识点，而忽略了知识的深度和广度，以及知识之间的内在联系。并且知识获取渠道有限，部分学生可能过于依赖教材和课堂讲解，而忽视了其他获取知识的途径。这限制了他们的知识广度和深度，也可能影响他们对问题的全面理解。

在强人工智能时代，尤其是在 ChatGPT 等生成式人工智能出现后，大学生的自主学习表现出强烈的时代特征。一方面，有助于大学生提升其学习效率、学习兴趣以及解决问题的能力，同时为大学生提供及时的学习反馈和指导。另一方面,不恰当地使用生成式人工智能也会导致大学生过度依赖技术、元认知能力退化、威胁隐私与数据安全以及催生学术造假等问题。因此，合理使用生成式人工智能将有助于大学生的自主学习推进和学习效果的提升。

3.基于 SWOT 分析的生成式人工智能辅助大学生自主学习的风险挑战

3.1. SWOT 分析

本研究采用半结构式访谈的形式，对 3 名大学生进行提问。根据以上对生成式人工智能创新学习方式和自主学习的分析和访谈结果可知，生成式人工智能既能作为创新学习方式的制胜法宝，但也存在着教育风险。因此，本文利用 SWOT 模型，分析了生成式人工智能辅助大学生自主学习的优势、劣势、机会与威胁。

外部环境分析（O-T） 内部环境分析（S-W）	机会（opportunities）	威胁（threats）
	作为智能辅导系统 作为评估反馈工具	引发教育不公平问题
优势（strength）	优势机会策略（S-O）	优势威胁策略（S-T）
有利于培养计算思维 有利于个性化学习	继续加强生成式人工智能的开发与应用	加强学校管理、树立健康校风 创新教学方式
劣势（weakness）	劣势机会策略（W-O）	劣势威胁策略（W-T）
作业失范问题 数据隐私泄露	建立正确学术道德观念 提高隐私保护意识	改变作业评价方式 教师正确引导学生

3.2. 风险挑战

虽然生成式人工智能可以在帮助大学生自主学习方面具有强大的优势，但是、以 ChatGPT 为代表的生成式人工智能同样也对大学生自主学习具有不可忽视的挑战与风险。

3.2.1. 作业失范问题

论文代写、扩写是最常见的失范行为，大学生通过将自己的需求输入进生成式人工智能中，获取它输出的回答，从而能够达到使用生成式人工智能代写论文的目的。生成式人工智能能在多门学科中帮助学生完成学习任务，提高教学成果的完成度与创意感，但随之而来也引发了学者们对学生学业诚信的担忧。

3.2.2. 数据隐私泄露

在使用人工智能的过程中，可能会涉及到学生的个人信息和隐私。如何保障学生的隐私安全是一个重要的问题。人工智能需要收集学生的个人信息，如学习行为、兴趣爱好、能力倾向等，如果这些信息被泄露或滥用，可能会侵犯学生的隐私权。另一方面，一些恶意攻击者也可能利用生成式人工智能系统的漏洞来获取和利用学生的个人信息，获取学生信息或者进行网络诈骗等活动。

3.2.3. 引起教育不公平问题

生成式人工智能辅助大学生自主学习可能会引发教育公平问题。首先，由于教育资源的不均衡分配，一些学生可能无法获得与人工智能辅助学习相关的优质资源，这可能导致教育机会的不公平，使原本资源匮乏的学生更加处于不利地位。而生成式人工智能作为一种大型语言模型与对话系统，其威胁性不止于替代教师职业，更可能借由其知识生成与供给能力而晋升为一种新型公共权力，进行知识与思想垄断，从而形成数字技术霸权。

4.生成式人工智能辅助大学生自主学习的风险挑战应对策略

虽然生成式人工智能在辅助大学生自主学习方面具有巨大的潜力，但在实际应用中还需要考虑这些挑战，并采取相应的措施来解决这些问题。

4.1. 教师层面

4.1.1. 创新教学方式

教师可以利用生成式人工智能提供的数据和分析，了解学生的学习进度、成绩、兴趣偏好等，从而制定更加个性化、有效的教学计划和方法。教师需要关注学生的反馈和表现，及时调整教学策略，提高教学效果。在生成式人工智能时代，教学内容的设计不能像之前那样只注重记忆和理解层面，应更倾向于发展学生的分析、综合、评价和创新能力。作业设计需考虑作业目标、内容、学生学情和作业资源等，使学生通过完成作业达成既定的教学目标。

4.1.2. 改变作业评价方式

面对大学生使用生成式人工智能作业失范的现象，教师应改变作业评价方式，提升课程评估素养。生成式人工智能时代需要从对结果的评价转向以表现性评价为主的作业评价方式，强化课程的形成性评价，作业评价需转向学生做作业的活动，或者以活动的方式检验作业。作业评价活动需遵循目标原则、过程原则和结果原则，应该注重对学生的自主思考和解决问题的能力做出评价。

4.2. 学生层面

4.2.1. 建立正确的学术道德观念

大学生应该建立正确的学术道德观念，加强学术道德教育，了解失范行为的严重性，养成正确的学术道德观念。生成式人工智能能够给学生提供灵感，帮助学生打开思路，因此彻底禁止使用显然是不可取的。大学生对于生成式人工智能的使用可以向参考文献方向靠近，将它 的回答作为参考文献的一种，在使用时标注出此部分内容来源于生成式人工智能。这样学生在心理上就会松绑，不掩饰使用情况，逐渐减少使用生成式人工智能的作业失范行为。

4.2.2. 提高隐私保护意识

大学生在使用生成式人工智能进行自主学习时，要对数据集进行更加严格的筛选和清洗，以确保数据的准确性和完整性。教育机构可以开展相关的教育和培训活动，提高学生的隐私保护意识和技能，帮助他们更好地管理自己的个人信息。同时，也可以对生成式人工智能的算法和模型进行改进和优化，以提高生成的准确性和可靠性。此外，要了解并关注相关组织或个人对个人信息的收集和使用政策，确保自己的个人信息得到合法、合规的使用和保护。

5.总结

生成式人工智能与教育的融合过程中可能会出现一些问题，但不应因此就禁止对生成式人工智能的使用。而是应思考能长效预防和反制的措施。预防和反制大学生使用人工智能的失范行为需要全社会的共同努力，社会、学校、家庭以及研究机构或企业应该一起来关注大学生的作业失范情况，制定相关的有效措施，开展有效的教育和监管工作，共同预防和反制大学生使用生成式人工智能的作业失范行为。

参考文献

[1] 于浩,张文兰.ChatGPT是否该被学术禁用——关于教学的深度思考[J].上海教育科研,2023,No.431(04):6-11+29.DOI:10.16194/j.cnki.31-1059/g4.2023.04.002.

[2] 王佑镁,王旦,梁炜怡等.“阿拉丁神灯”还是“潘多拉魔盒”：ChatGPT教育应用的潜能与风险[J].现代远程教育研究,2023,35(02):48-56.

[3] 焦建利.ChatGPT：学校教育的朋友还是敌人？[J].现代教育技术,2023,33(04):5-15.

[4] 徐力,朱国芬.ChatGPT技术与我国高等教育：机遇、挑战[J].继续教育研究,2024(01):32-37.

[5] 姜玲玲.浅谈大学生自主学习能力的现状和培养途径研究[J].吉林教育,2021(Z2):108-109.

探討中小學教師使用平板數位教學行為意向相關因素之研究

A Study on Factors Related to Elementary and Secondary School Teachers' Intention to Use Tablet Digital Teaching

鄭羽涵[1*]，林秋斌[1]，楊肅健[2]，高宏宣[3]
[1]台灣清華大學 學習科學與科技研究所
[2]台灣清華大學 跨院國際博士班學位學程
[3]台灣景文科技大學 旅遊管理系
[*]yuna.jheng@gmail.com

【摘要】 為探討臺灣中小學教師於課堂中使用平板進行教學之行為意向，本研究透過文獻分析擬定 TPACK、科技壓力、同儕支持、行政支持向度為自變項，使用態度為依變項；又以使用態度為自變項，行為意向為依變項，透過偏最小平方法結構方程模式(PLS-SEM)針對問卷填答結果進行分析，以了解 TPACK、科技壓力、同儕支持、行政支持和使用態度及行為意向的相關情形。研究結果顯示：一、教師的科技知識(TK)、科技教學知識(TPK)與科技學科知識(TCK)程度越高時，科技教學學科知識(TPACK)也會越高；二、教師的科技教學學科知識(TPACK)與同儕支持程度越高時，使用態度也會越高；三、教師知覺科技壓力越小時，則更具有正向的使用態度；四、使用態度愈積極正向，使用的行為意向也就越高而樂於採用。

【關鍵字】平板；平板數位教學；科技教學學科知識

Abstract: *To investigate the behavioral intentions of Taiwanese elementary and secondary school teachers to use tablets for teaching in the classroom, this study employed a literature analysis and established Technological Pedagogical Content Knowledge (TPACK), technostress, peer support, and administrative support as independent variables, with attitude as the dependent variable. Furthermore, using attitude as an independent variable and behavioral intention as the dependent variable, this study utilized the Partial Least Squares Structural Equation Modeling (PLS-SEM) to analyze the questionnaire responses. This aimed to understand the correlations among TPACK, technostress, peer support, administrative support, attitude, and behavioral intentions. The results of the study indicated the following: 1) As the level of teachers' technological knowledge (TK), technological pedagogical knowledge (TPK), or technological content knowledge (TCK) increases, their Technological Pedagogical Content Knowledge (TPACK) also increases. 2) When teachers' Technological Pedagogical Content Knowledge (TPACK) or peer support (PS) increases, their attitude (AT) towards technology use also increases. 3) Teachers with lower perceived technostress (TS) exhibit a more positive attitude (AT) towards technology use. 4) A more positively inclined attitude (AT) leads to higher behavioral intentions (BI) to adopt technology.*

Keywords: Tablet, Tablet Digital Teaching, TPACK

1. 緒論

新冠肺炎疫情發生，許多學校無法實體上課，致使數位學習的需求爆發性成長。臺灣於 2022 年開始投入大筆經費，推出「班班有網路、生生用平板」之數位學習精進計畫，希望在全台灣的中小學將平板、平台、數位內容及教師增能應用於數位學習上，這是臺灣教育史上金額最高、速度最急、影響最大的數位教學轉型投資。面對運用平板進行數位教學新政策，第一線的教師至關重要，有研究指出，教師在數位教科書與行動裝置作為教學和學習工具時，是扮演數位融入教育中的重要角色(Joo et al., 2016)。教學現場的老師，面對新政策，平板背

後他們的行為意向是如何?本研究乃透過問卷調查，以瞭解中小學教師的 TPACK、科技壓力、同儕支持、行政支持和使用態度各面向，對使用平板進行數位教學行為意向的相關情形。

2. 文獻綜述和研究假設

教師採用平板電腦教學的過程，除了外在因素(支持、技術和環境影響)、內在因素(例如個人信念、科技的使用、管理能力和感知的專業發展)，也會影響教師在不同學校環境中對科技的採用(ChanLin, 2017)。科技的使用、管理能力和感知的專業發展方面事涉教師的科技教學學科知識(TPACK)，其框架被視為適合分析教師使用科技進行數位教學的參考維度。

科技的發展和數位教學進入課堂中，重塑了老師的教學過程，教師努力將科技融入到教學過程中，受到外部因素的影響，如行政管理和教育政策(Administration support, AS)的支持與否，以及與同事的溝通和協作(Peer support, PS)。資訊科技產品發展迅速和在學習環境中的普及使用，使整合過程需要更加細緻的規劃，在這方面，教師在課堂內所承受的壓力可想而知(Özgur, 2020)，若教師付出努力而沒有得到回報或無法滿足期望，真正科技壓力(Technostress, TS)是來自於教師自身的感受(Voet & De Wever, 2017)。

另外，理性行為理論認為個人的行為是根據行為意向而決定，而行為意向主要是受到態度與主觀規範的影響，態度是指個人對該行為的想法或評價，主觀規範的定義是個人對於重要他人的看法或期望(Fishbein & Ajzen, 1975)，基於教師的態度是決定科技使用的重要因素之一，而 Davis 等人(1989)發展的科技接受模型，亦認為認知有用性與認知易用性可以解釋個人對特定新技術的使用態度和行為意向，而行為意向可以直接影響使用者實際使用。因此，本項研究亦納入教師使用平板進行數位教學的態度和行為意向二項因子，以形成完整的研究架構。本研究推論下列研究假設：

假設一(H1.a)：中小學教師使用平板進行數位教學的 PK 對 TPACK 有正向影響。
假設二(H1.b)：中小學教師使用平板進行數位教學的 TK 對 TPACK 有正向影響。
假設三(H1.c)：中小學教師使用平板進行數位教學的 TPK 對 TPACK 有正向影響。
假設四(H1.d)：中小學教師使用平板進行數位教學的 PCK 對 TPACK 有正向影響。
假設五(H1.e)：中小學教師使用平板進行數位教學的 TCK 對 TPACK 有正向影響。
假設六(H2)：中小學教師的 TPACK 對使用平板進行數位教學的態度有正向影響。
假設七(H3)：中小學教師的科技壓力越低，對使用平板進行數位教學的態度有正向影響。
假設八(H4)：中小學教師的同儕支持對使用平板進行數位教學的態度有正向影響。
假設九(H5)：中小學的行政支持對教師使用平板進行數位教學的態度有正向影響。
假設十(H6)：中小學教師使用平板進行數位教學的態度對行為意向有正向影響。

3. 研究方法

3.1. 研究流程

為探討臺灣中小學教師於課堂中使用平板進行教學之行為意向，本研究透過文獻分析擬定 TPACK、科技壓力、同儕支持、行政支持向度為自變項，使用態度為依變項；又以使用態度為自變項，行為意向為依變項，透過偏最小平方法結構方程模式(PLS-SEM)針對問卷填答結果進行分析，以了解 TPACK、科技壓力、同儕支持、行政支持和使用態度及行為意向的相關情形。具體工作是量表的編製，由非調查縣市之 88 位中小學教師進行預試，以預試樣本進行項目分析，最後以正式樣本問卷調查臺灣某一縣市 298 位中小學教師，以進行驗證性因素分析，建立信效度指標。

3.2. 研究工具

科技教學學科知識(TPACK)量表參考採用 Chai 等人(2010; 2011; 2012; 2013)和 Dong 等人(2020)發展之 TPACK 量表，包括 PK、TK、PCK、TCK、TPK 和 TPACK 六個向度，為符合本研究主題，部分題目將原「電腦」一詞更換為「平板」，並在舉例中淘汰舊有用語，轉化為現今臺灣常見的數位平台或軟體名稱，貼近受測者的生活使用情境，並給予補充教學情境的舉例，期望教學者能清楚瞭解題目內容，進而做出最符合自身狀況的回答。

科技壓力量表部分採用 Dong 等人(2020)編制的問卷量表，並將原敘述「ICT」一詞修改成「平板數位教學」關鍵字，以符合本研究所需。同儕支持、行政支持、使用態度和行為意向四份量表乃參考採用 Dong 等人(2020)、Park 等人(2012)、Wang 等人(2009)、丁一顧(2011)張媛甯和張書鳳(2016)、陳麗明和范斯淳(2020)編制的問卷量表，並修改成「平板數位教學」用語。

本研究正式問卷分成 PK、TK、PCK、TCK、TPK、TPACK、科技壓力(TS)、同儕支持(PS)、行政支持(AS)、使用態度(AT)和行為意向(BI)，11 個構面共計 61 個題目，組合信度(CR)介於 0.881 到 0.937，平均變異萃取量(AVE)介於 0.586 到 0.788 之間，均超過 0.5 標準值，符合 Bagozzi & Yi (1988)在評估各測量變項對該潛在變項的變異解釋能力的平均變異萃取量標準，本問卷具有收斂效度，而本問卷 Cronbach's Alpha 值均大於 0.7，具有良好信度。

4. 結果分析

本研究使用偏最小平方法以 SmartPLS 3 軟體進行 PLS-SEM 分析結構方程模型，並根據 Hair Jr.等人(2016)的建議，採用 5000 次隨機抽樣數進行拔靴法來檢驗研究假設是否成立，如表 1 所示為 PLS 結構方程模型分析之結果，除了假設一(H1.a)、假設四(H1.d)和假設九(H5)不成立外，其餘假設皆成立。本研究的模型解釋力，科技教學學科知識(TPACK)之 R^2 為 0.743，使用態度(AT)之 R^2 為 0.631，行為意向(BI)之 R^2 為 0.558，皆屬於良好解釋力的標準。

表 1 路徑關係檢定表

假設	路徑	路徑係數 β	t 值	p 值	結果
H1.a	PK ➔ TPACK	.01	.294	.769	不成立
H1.b	TK ➔ TPACK	.326***	6.559	<.001	成立
H1.c	TPK ➔ TPACK	.473***	10.649	<.001	成立
H1.d	PCK ➔ TPACK	.044	1.159	.247	不成立
H1.e	TCK ➔ TPACK	.153**	2.836	.005	成立
H2	TPACK ➔ AT	.621***	12.346	<.001	成立
H3	TS ➔ AT	.177***	3.967	<.001	成立
H4	PS ➔ AT	.154**	2.679	0.008	成立
H5	AS ➔ AT	.008	0.144	0.886	不成立
H6	AT ➔BI	.747***	25.661	<.001	成立

***p< .001；**p< .01；*p< .05

5. 結論

臺灣中小學教師使用平板進行數位教學的科技知識(TK)、科技教學知識(TPK)與科技學科知識(TCK)對科技教學學科知識(TPACK)具有正向影響，即教師的科技知識(TK)、科技教學知識(TPK)與科技學科知識(TCK)程度越高時，科技教學學科知識(TPACK)也會越高。

科技教學學科知識(TPACK)與同儕支持(PS)對使用態度(AT)具正向影響，即教師的科技教學學科知識(TPACK)與同儕支持(PS) 程度越高時，使用態度(AT)也會越高。而教師知覺科

技壓力(TS)越小時，則更具有正向的使用態度(AT)。使用態度(AT)愈積極正向，使用的行為意向(BI)也就越高而樂於採用。

參考文獻

丁一顧(2011)。教師專業學習社群與教師集體效能感關係模式驗證之研究。**屏東教育大學學報-教育類，37**，1-25。

張媛甯、張書鳳(2016)。臺南市國民中學教師參與教師專業學習社群與教師專業發展評鑑態度關係之研究。**學校行政，102**，175-200。

陳麗明、范斯淳(2020)。以整合型科技接受模式探討高雄市國小教師使用因材網之行為意圖。**工業科技教育學刊，13**，65-79。

Bagozzi, R., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Sciences, 16*, 74-94.

Chai, C. S., Chin, C. K., Koh, J. H. L., & Tan, C. L. (2013). Exploring Singaporean Chinese language teachers' technological pedagogical content knowledge and its relationship to the teachers' pedagogical beliefs. *Asia-Pacific Education Researcher, 22*(4), 657-666.

Chai, C. S., Koh, J. H. L., Ho, H. N. J., & Tsai, C. C. (2012). Examining preservice teachers' perceived knowledge of TPACK and cyberwellness through structural equation modeling. *Australasian Journal of Educational Technology, 28*(6), 1000-1019.

Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2010). Facilitating preservice teachers' development of technological, pedagogical, and content knowledge (TPACK). *Educational Technology & Society, 13*(4), 63-73.

Chai, C. S., Koh, J. H. L., Tsai, C. C., & Tan, L. L. W. (2011). Modeling primary school pre-service teachers' Technological Pedagogical Content Knowledge (TPACK) for meaningful learning with information and communication technology (ICT). *Computers & Education, 57*(1), 1184-1193.

ChanLin, L.-J. (2017). Analysis of teachers' tablet teaching adoption process. *Educational Sciences: Theory & Practice, 17*(6), 1935-1958.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models, *Management Science, 35*(8). 982-1003.

Dong, Y., Xu, C., Chai, C. S., & Zhai, X. S. (2020). Exploring the structural relationship among teachers' technostress, Technological Pedagogical Content Knowledge (TPACK), computer self-efficacy and school support. *Asia-Pacific Education Researcher, 29*(2), 147-157.

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Hair Jr., J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Joo, Y. J., Lim, K. Y., & Kim, N. H. (2016). The effects of secondary teachers' technostress on the intention to use technology in South Korea. *Computers & Education, 95*, 114-122.

Özgur, H. (2020). Relationships between teachers' technostress, technological pedagogical content knowledge (TPACK), school support and demographic variables: A structural equation modeling. *Computers in human behavior, 112*, Article 106468.

Park, S. Y., Nam, M. W., & Cha, S. B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British Journal of Educational Technology, 43*(4), 592-605.

Voet, M., & De Wever, B. (2017). Towards a differentiated and domain-specific view of educational technology: An exploratory study of history teachers' technology use. *British Journal of Educational Technology, 48*(6), 1402-1413.

Wang, Y.-S., Wu, M.-C., & Wang, H.-Y. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. *British Journal of Educational Technology, 40*(1), 92-118.

基于行为序列分析的双师课堂多元互动研究

Research on Multiple Interactions in Double Teacher Classrooms Based on Behavior Sequence Analysis

王诗蓓[1*]，荆璐[2]
[1、2]扬州大学新闻与传媒学院
[*] 007851@yzu.edu.cn

【摘要】 双师课堂是信息化促进教育公平发展的重要动力。研究基于ITIAS编码体系，结合具体双师课堂教学实践中所归纳的教学与互动特点，形成双师课堂中的互动分析系统，并以此为工具，结合滞后序列分析对双师课堂教学中的互动结构与规律进行探究。研究发现双师课堂教学实践中互动结构丰富多元；师生对话方式支持学生成长；双师协同高效。但是在技术、两地课堂管理、学生参与与助学教师助学意识层面仍存在问题，从提高技术流畅度，变革提问方式，丰富动机激发策略与互动线索，强调助学教师角色等方面提出建议。

【关键词】 双师课堂；ITIAS；FLAS；教学互动；滞后序列分析

Abstract: *Double teacher classroom as its application form has gradually become a trend, and it is also an important driving force for reformation to promote education equality. The research, based on the ITIAS coding system and combined with the teaching and interactive characteristics summarized in specific teaching practices by way of double teacher classroom, formed the interactive analysis system applied in the double teacher classroom.This being adopted as a tool, combined with lag sequence analysis, this paper explores the interactive structure and law in double teacher classroom. It is found that the interactive structure in the teaching practice of double teachers is rich and diverse; the dialogue between teachers and students facilitates students' growth; the two teachers reach a tacit understanding and cooperation. However, there are still problems in technology, classroom management and student participation, and this paper puts forward some targeted implementation suggestions in dealing with such problems..*

Keywords: double teacher classroom；ITIAS；FLAS；classroom interaction；Lag Sequence Analysis

1. 问题的提出

第十八届三中全会明确提出要利用信息化手段缩小城乡差距，推进教育公平发展进程。信息化助力教育公平是教育信息化不断发展过程中一以贯之的重要使命（廖宏建和张倩苇，2018）。探索技术支持下的教学新形态便成为了教育改革的旨趣，也是教育公平进程中的重要担当（郑旭东、饶景阳和贾洋洋，2021）。在时代浪潮下，双师课堂这一模式凭其颠覆传统教学模式的新兴特征走入了教育研究者的眼帘，其在教育公平推进层面的优越性毋庸置疑（冯晓英等，2020），是实现公平有质量教育的有效路径（Huang R，Kinshuk，&Price，2016）。

已有研究明确指出，双师课堂的本质在于交互（郭炯和杨丽勤，2020）。同时国内外众多研究均强调课堂中的教学互动，认为丰富多元的互动能够有助于教学质量的提升（刘向永等，2018）。可以认为，对双师课堂中的教学互动行为进行分析，不仅有较强的现实可行性，更是通过对其“互动”本质的分析，发现问题与实践规律，助力双师教学更为高效的实践，进而推进教育公平的进展（Yang ，Li，& Xing B ）。而在我国目前对“双师课堂”及其相关的其他新型网络同步\异步课堂研究中，研究者多关注这类课堂教学的模式设计与实践，以及相对应的技术支持（杨俊锋等，2018；屯勇和万文静 2021；常咏梅等 2020；黄甫全等，2020；屯勇、高红英和王鑫，2020），已有研究鲜少聚焦于此类课堂教学过程中的具体细节，因此本研究致力于了解双师课堂教学过程中的交互特征，进而实现对双师课堂本质的探索。

2. 相关研究

2.1. “双师课堂”的内涵与特征

双师课堂是融合了目前“三个课堂”的优势特征，借助信息技术手段实现优秀教师和普通教师的协同教学，实现名师资源在的常态化共享，最终形成双方教师与学生的共同成长（周方苗和何向阳，2020）。其中优秀教师有着丰富的教学经验，能够呈现精彩的教学过程；普通教师教学能力相对弱势，在教学过程中协助优秀教师，管理线下课堂，同时将学生的听课情况及时反馈给优秀教师，以供其及时调整教学。

不同于专递课堂、名师课堂以及名校网络课堂，双师课堂的实践是两所学校，两名教师，两地学生的双向同步教学，这里强调了双向与同步，也就是说这样的教学以直播的形式进行，包含优秀教师作为主讲教师的本地课堂，以及普通教师作为助学教师的远方课堂。本地课堂中，主讲教师面向本地学生授课，借助信息技术将授课内容传输到远方课堂。远方课堂中，远方学生观看本地课堂教学实录，同时助学教师进行班级管理，协助主讲教师实现与远方学生的交互。在这样的同步直播的双师教学中，高质量的实时互动是影响学生学习投入的重要因素（刘司卓、李爽和黄嘉靖、2021）。雷励华和左明章（2015）也指出多维教学交互是同步直播形式下双师课堂质量的有力保证。因此需要一个科学完善的交互体系来对双师课堂中对交互行为展开研究，以便进行高质量的双师课堂实践。

2.2. 课堂互动行为研究现状

课堂互动行为的分析多以课堂互动行为分析系统为工具展开（胡金艳、蒋纪平和张义兵，2019）。美国学者 Flanders（1961）从教师言语、学生言语和沉默三个维度列举出师生之间的十种交互行为，形成课堂互动分析系统 FLAS（Flanders Interaction Analysis System），并以这一编码体系首次实现对教学交互行为的量化分析，这一研究成果为教学互动行为的深度探索提供了科学的理论框架，并在后续的研究中被广泛应用。除了直接使用 FLAS 进行课堂互动行为分析（张露丹、汪颖和潘玉霞，2011），更多的是根据实际教学情况调整 FLAS，进而对课堂互动行为展开分析。比如心理学家 Amidon（1967）从情感氛围视角出发对 FLAS 进行调整；Moskowitz（2010）针对外语教学特点形成的外语课堂中的教学互动分析体系。同时，随着信息技术在教育领域的持续渗透，越来越多的技术应用于课堂教学中，也因此催化“基于信息技术的互动分析编码”ITIAS（Information Techno-Based Interaction Analysis System）。这是由华东师范大学的顾小清教授研究团队，立足于 FLAS，引入信息技术条件，在新课改的背景进行的科学合理的完善与改编，根据实际教学的发展丰富了课堂交互内容，分析包括课堂上的“教师言语、学生言语、技术、沉默原因”四个维度（顾小清和王炜，2004）。借助于这一编码方式，研究者能够准确地掌握目前信息技术支持的课堂教学的真实情况。至此，各项 ITIAS 相关研究的推进，研究者或直接使用 ITIAS 课堂观察分析（江毅等，2019），或立足于此，针对实际教学情境修订 ITIAS 进行研究（刘向永等，2018）。而在双师课堂相关的教学情境中的互动分析研究中，有学者通过修订 ITIAS 分析系统对专递课堂（韦怡彤等，2019）或同步课堂（王觅等，2020）中的互动行为展开分析。

综上所述，双师课堂中的教学互动是其课堂教学质量得到保证的关键因素，FLAS 或 ITIAS 分析系统均能够为教学交互研究提供方法上的支持，但目前教学互动的研究多借助 ITIAS 对信息技术支持的单个课堂教学互动展开分析，也有部分针对专递课堂、同步课堂相关教学实践中互动行为进行分析的研究，但其分析系统与双师课堂实践特征仍有出入，因此本研究鉴于双师课堂的实践特征，立足于 ITIAS，设计双师课堂中教学互动行为分析系统，并借助这一工具以滞后序列分析法对双师课堂中的互动行为展开分析。

3. 研究设计

3.1. 研究目标

研究目的在于，根据双师课堂的实践特征，修订 ITIAS，在此基础上了解在双师课堂中的教学互动结构，探索双师课堂的互动规律，总结经验，发现问题，以此提出提高双师协同的切实有效的实践建议。

3.2. 研究样本

本研究所选取的样本来自 S 省 MD 小学与 HT 小学五年级的数学学科双师课堂教学示范课，其中 MD 小学作为优质校，以下将其称为本地课堂，其教师为主讲教师，学生称之本地学生；HT 小学作为普通校，以下将其称为远方课堂，其教师为助学教师，学生称之远方学生。

3.3. 研究工具与方法

在双师教学情境下，交互类型较之传统课堂更加多维与复杂。本研究融入双师课堂的交互特征，结合研究团队中多所进行双师教学实践的实验学校的具体实践情况，对 ITIAS 中的四个维度进行调整修订，最终形成表 1 所示的适用于双师课堂教学互动行为分析的编码体系。

表 1 双师课堂教学互动行为分析的编码体系（ITIAS_DTC）

来源	描述	编码
主讲教师 TA	讲授	TA1
	指示（本地学生、远方学生）	TA21、TA22
	提问（本地学生、远方学生）	TA31、TA32
	接受或纠正学生回答	TA4
	批评（本地学生、远方学生）	TA51、TA52
	引导	TA6
助学教师 TB	指示	TB1
	引导	TB2
	批评	TB3
	临场感创设	TB4
学生	应答（本地学生、远方学生）	SA、SB
	合作讨论	SC
技术 N	教师操纵技术	N1
	学生操纵技术（本地学生、远方学生）	NA、NB
	技术故障	N2
沉默 Q	无益于教学的混乱	Q1
	做练习	Q2
	观看内容演示	Q3

利用上述互动分析编码能够整体掌握双师课堂教学过程中的互动结构，但无法细致了解各交互主体之间的互动规律，因而引入滞后序列分析法来发现两地课堂中显著的行为序列，进而发现双师课堂教学的互动类型。滞后序列分析法能够预测某一行为发生后其他行为发生的可能性，并且验证发起行为与伴随行为序列在统计学意义上的显著性，近年来在教育领域亦有较多的应用（Bakeman &Quera，1995）。

3.4. 数据处理

本研究对取样中的主要课堂行为进行编码，平均 3 秒钟进行一次打点取样，并按时间顺序进行记录。为了保证编码结果的信度，这里由研究团队的两名研究生进行独立编码，通过

对二人编码结果进行一致性检验，得出其 Kappa 系数为 0.882，能够较好地确保研究信度。对于有争议的编码结果，则由研究团队共同协商决定。由于两地课堂同步进行，同一时间间隔内的主要行为可能不止一种，因此最后得到 989 条编码结果。

4. 研究结果

4.1. 课堂互动结构分析

图 1 展示了双师课堂教学过程中的互动编码结果，可以看出主讲教师在课堂中占主导地位，助学教师的也承担起举足轻重的作用，尽管其行为比例略低于主讲教师，但是。主讲教师能够兼顾到远方学生群体，通过与助学教师之间的协同，实现对远方学生群体的提问、引导与评价；助学教师也能够在管理课堂的同时，对其学生进行一定的提示与引导。同时，两名教师的言语行为占总行为的 49.64%（28.43%+21.21%），远低于常规教学中的 68%（Black &Wiliam，1998），也能侧面印证学生在双师课堂中的主体地位得到重视。而本地学生的行为亦略高于远方学生的行为，这与异地教学、技术局限相关。最后，技术使用与沉默两个维度占比亦不在少数，究其本质，技术主要承担起本地课堂教学内容同步传输、双师协同教学、以及两地学生作业演示的作用，且需要引起重视的是技术故障在整个技术使用中亦占有不小的比例，这里的技术故障主要是网络卡顿、延时导致的双师协同效率不足，以及麦克风声音不够导致的重复提问与回答，这些均占用课堂教学时间。课堂沉默大部分来多是学生在教师的引导下独立完成探究问题的相关行为，而无益于教学的短暂停顿并不多，说明主讲教师与助学教师在课前有充分的沟通，且主讲教师有较强的课堂掌控能力，保证课堂的有序运行。

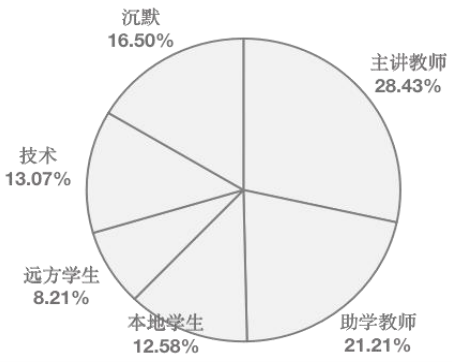


图 1 双师课堂各行为占比情况分布

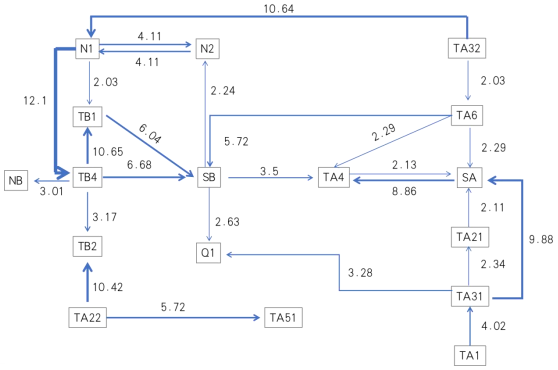


图 2:行为序列转换图

4.2. 互动行为序列分析

为了解双师课堂教学过程中显著有意义的互动行为序列，这里利用滞后序列分析法对其中一节双师课堂教学行为进行进一步探索。将互动编码结果导入 GSEQ 软件中进行分析，生成行为转换频率表和调整后的残差表（部分结果如表 3 和表 4 所示）。其中，行为转换频率表可以将其看作是交互矩阵，在交互矩阵中，列为发起行为，行则是发起行为之后的伴随行为，发起与伴随行为形成行为序列，在表格中呈现为行列交汇的单元格，单元格内的数字则是行为序列的出现频次，比如在表 4 中可以看到，TA1—>TA31，SA—>TA4 等行为序列的互动行为转换频率比较高。当然，仅数字的大小无法准确的呈现行为序列的显著性与重要性，调整后的残差表呈现的是软件计算行为频率的残差参数 z 值，并说明 z 值大于 1.96 的行为序列具备统计学层面的显著性（Bakeman，Mcarthur，&Quera，1996）。

表 3 互动行为转换频率表（部分）

	TA1	TA21	TA22	TA31	TA32	TA4	TA51	TA6	SA	……
TA1	86	3	3	15	4	0	0	1	0	……
TA21	2	7	0	2	0	0	0	0	5	……

TA22	0	0	6	0	1	0	3	0	0	……
TA31	5	3	0	7	0	0	0	1	26	……
TA32	0	0	1	0	5	1	0	3	0	……
TA4	12	2	2	5	3	12	0	1	12	……
TA51	0	0	0	0	0	0	0	0	0	……
TA6	0	0	0	0	1	0	0	16	9	……
SA	6	1	0	8	1	28	0	7	50	……
……	……	……	……	……	……	……	……	……	……	……

表 4 调整后的残差表（部分）（注：*p=<.01）

	TA1	TA21	TA22	TA31	TA32	TA4	TA51	TA6	SA	……
TA1	19.62*	0.38	0.49	4.02*	0.4	-3.16	-0.41	-2.04	-4.45	……
TA21	-0.32	11.35*	-0.59	1.23	-0.71	-1.13	-0.15	-0.9	2.11*	……
TA22	-1.66	-0.59	10.3*	-0.95	0.83	-1.09	5.72	-0.87	-1.54	……
TA31	-0.49	2.34*	-0.95	3.44*	-1.14	-1.8	-0.23	-0.67	9.88*	……
TA32	-2	-0.71	0.83	-1.14	5.53*	-0.48	-0.17	2.03*	-1.85	……
TA4	1.6	0.82	0.91	1.35	1.2	4.58*	-0.27	-0.98	2.13*	……
TA51	-0.41	-0.15	-0.14	-0.23	-0.17	-0.27	-0.04	-0.22	-0.38	……
TA6	-2.53	-0.9	-0.87	-1.44	-0.03	-1.66	-0.22	11.9*	2.29*	……
SA	-2.64	-0.85	-1.54	1.29	-1.21	8.86*	-0.38	1.26	11.84*	……
……	……	……	……	……	……	……	……	……	……	……

从数据结果筛选中发现 z 值大于 1.96 的行为序列共 47 对，除了指向自身的显著行为序列之外，还存在其他“师-生-技术之间”的显著的行为序列共 27 对。基于筛选结果绘制显著行为序列转换图，用箭头表示行为转换方向，而连接两个行为之间的连线粗细则与显著性强度相关，显著性越高连线则越粗（杨现民、王怀波和李冀红，2016），具体如图 2 所示。

结合调整后的残差表中显著的行为序列，以及图 5 中的行为序列转换图可以发现，双师课堂教学实践中，行为交互出现分别以“SA”“SB”为中心的两个区域，由于双师的研究情境，可以从两地课堂的交互出发进行讨论。

在本地课堂教学中，不仅教师讲授与提问呈现显著行为序列（TA1—>TA31，z 值 4.02；TA31-SA，z 值 9.88），学生回答与教师的反馈亦呈现显著行为序列（SA—>TA4，z 值 8.86；TA4—>SA，z 值 2.13）。立足于 IRF 理论，可以认为本地课堂中的教学互动是多步三元结构，能够在强化学生对知识理解对基础上，培养意义建构等高阶思维（Sinclair &Coulthard，1975）。同时，在本地课堂中需要引起注意的是，当主讲教师向远方学生发出指示之后，随之的行为出现批评本地学生，且该行为序列呈现显著性（TA22—>TA51，z 值 5.72）。结合视频实录分析原因，主讲教师指示远方学生实现某项指令，其关注点转移到远方课堂，因此本地学生处于无人管理的状态，容易出现注意力分散的情况。

在远方课堂教学中，有主讲教师参与到双师协同课堂中，交互主体和行为则呈现更为复杂多元的情况。首先在主讲教师的讲授与提问方面，仍延续着和本地课堂一致的多步三元结构，但是与本地课堂有些许差异，主要表现在“主讲教师—技术—助学教师—学生—主讲教师”的交互闭环（比如：TA32—>N1—>TB1—>SB—>TA4，z 值 10.64、2.03、6.04、3.5），也就是说，当主讲教师向远方学生群体抛出问题之后，随之而来的是双师协同，此时助学教师向远方学生群体发出回答问题的信号，在远方学生回答问题的过程中主讲教师再针对具体的回答情况予以一定的反馈。同时，在远方课堂中，助学教师接受到主讲教师对远方学生抛

出的提问以后，助学教师除了向学生发出指令以外，还存在另一列显著的行为序列，表现为“双师互动—助学教师氛围感营造—远方学生回答问题 (N1—>TB4—>SB, z 值 12.1、6.68)”。然而，在远方课堂中助学教师的引导与远方学生回答之间未呈现显著的行为序列，对远方学生回答进行反馈与引导的主要是主讲教师。另外，需要引起重视的是软硬件环境导致的技术故障，在双师协同以及远方学生回答问题的过程中均存在与“技术故障”相关的显著序列（N1—>N2，z 值 4.11；SB—>N2，z 值 2.24），从 z 值的大小可以判断，特别是在双师协同过程中容易出现技术故障的情况。结合视频实录可以发现，存在的问题主要集中在“网络流畅程度”和“麦克风声音效果”两个方面。最后，在远方学生回答问题的过程中还存在与课堂教学无关的沉默与混乱（SB—>Q1，z 值 2.63），结合视频实录可以发现远方学生在回答问题过程中容易出现停顿的情况，究其原因，除去其基础较为薄弱的原因，更多是因为他们要面对镜头以及主讲教师和本地学生，本就基础较弱的远方学生在心理上有一定的距离感和畏惧感。同时，在远方学生回答问题过程中停顿的时候，这些学生多低头呢喃或用眼神求助助学教师，但远方学生求助与助学教师的引导之间未形成显著的行为序列 (SB—>TB2, z 值 0.64)。

5. 结语

5.1. 研究结论

通过上述分析可以发现：第一，双师课堂教学实践中“师-生-技术”多元互动结构丰富，技术在双师协同，以及师生对话过程中作用显著；第二，双师课堂教学过程中以“引导-追问”的师生对话方式关注学生知识获取的同时强调思维的提升；第三，双师协同默契，助学教师能够迅速掌握主讲教师意图并与之配合。

同时，双师课堂在教学实践中仍存在问题：首先，技术环境的流畅度不足。在主讲教师与远方课堂互动的过程中，音视频的均容易出现卡顿，而因技术故障导致的课堂教学行为的停滞则降低了教学效率，并不断消磨师生的课堂参与热情。其次，课堂管理的兼顾不周。由于主讲教师“身兼数职”，需要承担两地课堂的教学，因此在主讲教师与远方课堂互动的过程中，无暇顾及本地学生，容易导致本地学生的注意力分散。再次，远方学生的课堂参与和自信心不足。纵观整个编码分析结果，远方学生的发言频率均低于本地课堂；同时在回答过程中出现无关教学的停顿，影响整个教学的进度。最后，助学教师的引导意识薄弱。助学教师的工作重点多放置于配合主讲教师而忽略了其学生的学习与心理需求，未能在学生需要教师引导的时候予以及时的反馈与支持。

5.2. 相关建议

立足于研究结论，针对上述问题，从技术环境和双师协同两个维度提出以下建议：

首先，技术赋能，助力双师教学高效实施。高质量的技术流畅有助于实现两地师生之间无障碍、流畅的沟通，比如远方课堂拾音方面，由于两地课堂的距离感使得远方学生存在一定的胆怯心理，学生发言声音较小，这时候对麦克风的灵敏度要求会更高，以此保证声音能够完整的传输到本地课堂，减少不必要的重复。同时，师生的信息素养提升亦是双师教学高效实施的关键保障，比如双师利用技术进行无障碍沟通、学生利用交互式电子白板直接投屏并演示思维过程，在稳定高速的网络环境下能够减少教学过程中的技术噪点。

其次，变革提问方式，关注学生专注度。主讲教师根据两地学生的不同情况进行提问，由于本地学生有着相对更完整的认知结构与知识基础，因此对于需要拓展思维的开放性问题优先选择本地学生进行回答；对于在记忆层面的封闭性问题则优先选择远方学生进行回答。这样既能够保证本地学生在基础课堂上得到能力提升，亦能够保证远方学生能够充分的参与到课堂活动中并得到一定的肯定。同时，为了解决本地教师提问远方课堂时本地课堂出现的无益于教学的混乱这一问题，教师可对学生提出要求，学生需关注回答问题同学的答案，每

个人都是小老师，在对方回答结束后针对回答内容进行补充或评价。不仅能够提高本地学生的专注度，亦是对其思维拓展的一种方式。

再次，丰富动机激发策略与互动线索，提高远方学生课堂参与及自信。两地学生群体之间的差异性以及与课堂之间的距离感容易导致远方学生群体沦为“观众”，在缺少临场感的情况下，学生对课堂的参与度将会大打折扣（朱伟强、黄山和盛慧晓，2017）。因此，丰富的动机能够激发远方学生群体对课堂的充分参与。比如两地学校深度合作，进行跨校项目式学习，按照项目的具体内容和要求设计有梯度的任务分工并进行跨校的小组合作；或者双师课前沟通，设计赚积分、抢红包等形式的课堂活动，并以两地互评的形式营造竞争的氛围。

最后，强调助学教师角色，强化远方学生临场感。众多研究均强调教师所承担的角色绝不止“讲授者”与“提问者”，更多的是在教学过程中应承担起协调者、评价者、管理者、技术维护者、辅导者、研究者（Peter，2020）。教师的任务就是由其角色决定的（GUASCH , ALVAREZ, &ESPASA，2010），因此在双师课堂教学过程的助学教师不仅仅承担起的是“技术维护”和“课堂管理”角色。更多的是对学生认知层面的介入，助学教师在课堂中的表现与敏捷反应能够极大提高学生的临场感，同时降低学生的认知负荷（卜彩丽等，2021）。具体而言，助学教师需要熟知主讲教师的教学环节和意图，在主讲教师向远方学生抛出问题之后能够迅速根据问题的难度选定合适的学生指示其进行回答，同时在指示学生回答之前营造氛围，为学生创设临场感鼓励学生主动应答，并且针对学生回答进行一定的反馈与引导。

参考文献

常咏梅, 张乐, 李玥琪 & 李华. (2020). 同步直播课堂远端教师助学策略研究. ***电化教育研究 (11)***, 116-121+128.

冯晓英, 王瑞雪, 曹洁婷 & 黄洛颖. (2020). “互联网+”时代三位一体的教育供给侧改革. ***电化教育研究 (04)***, 42-48.

郭炯 & 杨丽勤. (2020). 协同与交互视角下的同步课堂：本质、困境及破解路径. ***中国电化教育 (09)***, 89-95.

顾小清 & 王炜. (2004). 支持教师专业发展的课堂分析技术新探索. ***中国电化教育 (07)***, 18-21.

黄甫全, 伍晓琪, 唐玉溪, 陈思宇 & 曾文婕. (2020). 双师课堂课程开发引论：缘起、主题与方法. ***电化教育研究 (02)***, 99-107.

胡金艳, 蒋纪平 & 张义兵. (2019). 知识建构社区课堂互动分析工具的设计与应用. ***现代教育技术 (04),*** 61-66.

江毅, 王炜, 李辰钰, 康苗苗 & 沈洁. (2019). 智慧教室环境下师生互动行为研究. ***现代远距离教育 (03),*** 13-21.

廖宏建 & 张倩苇. (2018). “互联网+”教育精准帮扶的转移逻辑与价值选择——基于教育公平的视角. ***电化教育研究 (05)***, 5-11.

刘向永, 李傲雪, 付奕宁, 姜沛雯 & 王萍. (2018). 基于电子书包的小学英语课堂师生互动分析——以“How are you?”单元为例. ***电化教育研究 (08)***, 97-102+121.

刘司卓, 李爽 & 黄嘉靖. (2021). 直播课学习行为投入评价的实证研究. ***中国远程教育 (02)***, 36-45+58.

雷励华 & 左明章. (2015). 面向农村教学点的同步互动混合课堂教学模式研究. ***电化教育研究 (11)***, 38-43.

乜勇 & 万文静. (2021). 双师教学三段循环模式的构建与应用研究. ***中国电化教育 (02)***, 83-88+96.

乜勇, 高红英 & 王鑫. (2020). “双师教学”共同体模式构建：要素与结构关系分析研究. ***电化教育研究 (12)***, 65-70+78.

卜彩丽, 郭佳佳, 李奇 & 张思. (2021). 在线教学视频中教师在屏能否促进学习？——基于国内外 33 项实证研究的元分析. ***远程教育杂志 (06)***, 67-75.

韦怡彤, 王继新, 赵晓娜, 田俊 & 丁茹. (2019). 同步互动专递课堂中教学互动行为案例研究——以一年级美术课“画马路”为例. ***现代教育技术 (12)***, 41-47.

王觅, 文欣远, 李宁宁 & 陈焕东. (2020). 大单元教学视角下基于 LSA 的同步课堂师生交互行为研究. ***电化教育研究 (08)***, 74-81.

杨俊锋, 崔丽霞, 吴滕, 潘慧 & 洪河条. (2018). 混合同步网络课堂有效性的实证研究. ***电化教育研究 (12)***, 50-56+77.

杨现民, 王怀波 & 李冀红. (2016). 滞后序列分析法在学习行为分析中的应用. ***中国电化教育 (02)***, 17-23+32.

周方苗 & 何向阳. (2020). 以双师课堂实现优质教师资源共享的实践研究. ***中国教育信息化 (05)***, 45-49.

张露丹, 汪颖 & 潘玉霞. (2011). 基于FIAS的中学信息技术专家教师课堂教学研究.***现代教育技术 (06)***, 39-43.

郑旭东, 饶景阳 & 贾洋洋. (2021). “三个课堂”促进义务教育优质均衡发展:演进历史、战略价值、关系解析与概念框架. ***现代教育技术 (06)***, 14-22.

朱伟强, 黄山 & 盛慧晓. (2017). 教学中的人际关系与师生课堂互动、学习结果的相关性研究. ***全球教育展望 (12)***, 24-34.

Amidon E J , Hough J B .(1967).*Interaction Analysis: Theory, Research and Application*.Massachusetts: Addison-Wesley Publishing Company.

Bakeman R , Quera V .(1995).*Analyzing interaction: Sequential analysis with SDIS and GSEQ*. Cambridge University Press.

Black, Paul, Wiliam, & Dylan. (1998). Assessment and classroom learning. Assessment in Education: Principles, *Policy & Practice*, 5(1):7-74.

Bakeman, R. , Mcarthur, D. , & Vicenç Quera. (1996). Detecting group differences in sequential association using sampled permutations: log odds, kappa, and phi compared. *Behavior Research Methods Instruments & Computers*, 28(3), 446-457.

Flanders N A. (1961). Analyzing teacher behavior. *Educational Leadership*,(3):251-260.

GUASCH T, ALVAREZ I, ESPASA A,(2010).University teacher competencies in a virtual teaching/learning environment:analysis of a teacher training experience. *Teaching&teacher education*,26 (2) :199-206.

Huang R , Kinshuk, Price J K . (2016).ICT in Education in Global Context. *Springer Berlin Heidelberg*, 95-106.

Moskowitz, G. . (2010). Interaction analysis—a new modern language for supervisors. *Foreign Language Annals*, 5(2), 211-221.

Peter Samuelson Wardrip,(2020). Educators enacting online learning support roles in remote educational experiences. *Educational Technology Research and Development*, 69(1) : 1-4.

Sinclair J M , Coulthard R M . (1975). Towards an analysis of discourse : the English used by teachers and pupils. *Tesol Quarterly*, 11(2)157-163.

Yang, X. , Li, J. , & Xing, B. . (2017). Behavioral patterns of knowledge construction in online cooperative translation activities. *The Internet and Higher Education*, 36(jan.), 13-21.

智能时代教育技术理论与实践新进展

New Progress in the Theory and Practice of Educational Technology in the Intelligent Age

罗文婧[1*]，邱艺[2]
[1]广州理工学院人文与教育学院
[2]华南师范大学教育信息技术学院
[*] 1145316274@qq.com

【摘要】 教育技术是推动教育现代化发展的创新实践领域，在支持当前教育数字化转型和高质量发展中起到举足轻重的作用。本研究立足智能时代创新人才培养需要，以教育数字化战略行动为契机，应用文本数据挖掘法和内容分析方法，构建智能时代教育技术变革分析类目，系统揭示智能时代国内外教育技术变革及新进展，以期为全面洞察当前教育技术发展趋势，创新教育信息化人才培养，助力教育高质量发展提供有益借鉴。
【关键词】 智能时代；教育技术；理论；实践；进展

Abstract: *Educational technology is an innovative practice field that promotes the modernization of education, and plays a pivotal role in supporting the current digital transformation and high-quality development of education. This study is based on the need to cultivate innovative talents in the intelligent era, taking the strategic action of education digitization as an opportunity, applying the content analysis method, constructing the relationship model and analysis categories of educational technology change in the intelligent era, and systematically revealing the major aspects of educational technology change and new progress in the intelligent era. It is expected to provide a useful reference for gaining comprehensive insight into the current development trend of education technology, innovating education informatization talent training, and helping education high-quality development.*

Keywords: intelligent era, educational technology, theory, practice, progress

1. 问题的提出

随着人工智能、大数据等新型基础设施建设的不断深化，社会各行各业正朝着数据化、网络化、智能化的方向持续演进。推进教育数字化，以数字技术为教育赋能，有利于各级各类教育高质量发展，更是加快推进教育现代化、建设教育强国的战略举措（教育部，2023）。教育技术作为推动教育现代化发展的创新实践领域，在支持当前教育数字化转型和高质量发展中起到举足轻重的作用，同时也面临自我发展和自我突破的诸多挑战。本研究应用文本数据挖掘法和内容分析方法，系统揭示了智能时代教育技术变革及新进展，以期为全面洞察当前教育技术发展趋势提供有益借鉴。

2. 智能时代教育技术变革的分析

2.1. 样本选取与聚类

本研究主要选取《电化教育研究》《中国电化教育》《开放教育研究》《现代教育技术》《远程教育杂志》《中国远程教育》《Computers & Education》《British Journal of Educational Technology》《Educational Technology Research and Development》等国内外教育技术专业核心期刊近一年的发表论文，采用文本数据挖掘法和内容分析法进行量化分析，如图 1、2 所示。

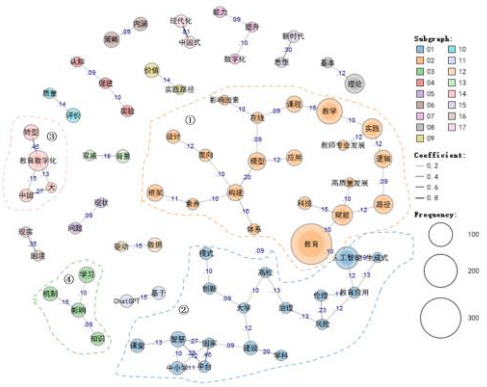


图 1 中文文献主题词聚类图

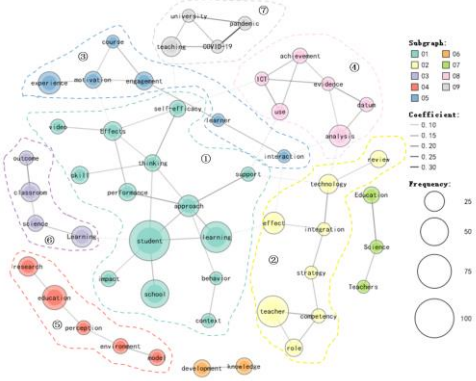


图 2 外文文献主题词聚类图

从图 1 可看出，国内教育技术相关研究中聚类①所占面积最大，重点关注科技赋能教育高质量发展的路径、逻辑和实践等，通过构建教学模式、模型，提出教学设计框架，为智能时代下教学创新提供可理解、可操作的措施；聚类②以技术发展为抓手，重点关注人工智能、生成式 AI、国家智慧教育平台等在教育教学中的创新应用，分析各方面风险；聚类③强化了我国的话语体系，旨在发展新时代中国特色教育技术理论与实践体系，走好新时代中国特色教育技术发展道路；聚类④及其他小聚类则进一步探究新时代新政策下教育教学创新思考。

从图 2 可看出，国外教育技术相关研究同样重点关注课程与教学的创新应用，主要围绕“学生”“学习者”和“学习”积极探索不同环境中的教学应用与教学效果（聚类①⑤⑥）；同时，从聚类②④中可以看出技术是国外教育技术相关研究的一大聚焦点，强调了技术与教育的深度融合，并指出可利用技术和数据动态测评并持续发展教师及其他学习主体的能力；除此之外，国外教育技术相关研究还重点关注了当前背景下各项教育教学活动中学习者的学习动机、参与和互动、学习体验等（聚类③），也依旧关注着疫情后高校教学的创新实践。

2.2. 类目构建与应用

本研究基于上述期刊文献，根据研究目的和已有文献特征进行了内容分析编码，并对每篇文献以其最大特征作为主要分类依据。具体分析结果见表 1、2。

表 1 中文文献研究热点编码结构（前 10 个）

一级类目	二级类目
智能教育（235）	理论探索、教学创新、教育治理与伦理风险、智能技术、智能评价、环境空间、模式模型、人机协同、智能素养、课程资源、教育生态、
教育数字化（175）	理论思考、教育场景应用、数字素养、数据治理与教育应对、数字资源与平台、数据评价、模式框架、教育公平
教育政策（71）	中国式教育现代化、教育高质量发展、教育强国、双减、新课标、新职教法、高等教育普及化
在线教育（59）	学习行为、交互特征、在线评价、情感研究、系统平台、在线课程
生成式人工智能（42）	探究思考、教育应用、风险规避、学习模式、教师发展、
教师专业发展（35）	教师研修、队伍建设、实践探索
教育元宇宙（32）	虚拟现实、增强现实、数字孪生、教育场景、技术工具、技术伦理
科学教育（18）	内涵探讨、信息科技课程、跨学科
新时代思想政治教育（17）	大思政课、课程思政
个性化学习（10）	理论探究、技术关键、模型路径

智能教育自 17 年提出至今，一直是教育技术研究领域探讨不断的话题。教育数字化是教育信息化进一步深化的高级发展阶段，党的二十大报告曾明确指出要“推进教育数字化，建设全民终身学习的学习型社会、学习型大国”（教育部，2022），我国众多学者对教育数字化展开相应的理论思考、教育场景应用、数字素养等研究，并在持续深化。教育政策引领着我国教育技术研究的方向，越来越多的学者结合国家战略发展需求，提出了相应的理论探索和发展路径，逐步推动教育政策走深走实。在线教育不仅于疫情期间研究热度上升，且疫后依旧是我国学者的研究热点，重点探究在线教育中的学生学习行为、交互特征等，同时关注在线教育与面对面教育的结合创新。生成式人工智能是基于深度学习技术的人工智能模型，

能够模拟人类的创造性思维，生成具有一定逻辑性的语言文本、图像、音频等内容，为新时代背景下的教育创新应用、学习模式变革、学习质量提高等提供了新途径（焦建利，2023），越来越多的专家学者热烈探讨其在真实教育情境的创新应用与实际效应。除此之外，我国教育技术专家学者还关注着新时代背景下的教师专业发展路径与创新、教育元宇宙的理论与实践探究、新课标下的科学教育创新发展、新时代思想政治教育和个性化学习的实现等。

表 2 外文文献研究热点编码结构（前 10 个）

一级类目	二级类目
在线教育（68）	在线情感、效果评价、在线危机、在线学习策略、影响因素
拓展现实（54）	虚拟现实、增强现实、元宇宙、技术伦理、平台工具、教学实践
教育数字化（45）	教学创新、评价认证、数字能力、数字素养
智能教育（39）	机器工具、智能教学、智能评价、智能技术、情感探究、自适应学习
科学教育（33）	跨学科、创客、科学课程、STEM
技术促进教学（33）	工具开发、技术整合、教学实践、协作学习
学习分析（19）	技术方法、实证分析
游戏化学习（18）	教育游戏开发、游戏化学习设计、游戏化学习效果
教育技术发展（12）	研究方法、学科发展
生成式人工智能（8）	概念框架、角色设计、应用探究

在线教育是国外学者依旧重点关注的方向，且仍有部分对疫情期间的在线教育进行经验总结和反思创新。拓展现实支撑着多维学习场景创设，许多学者紧抓教育元宇宙的发展机遇，探究其中的教学实践与技术伦理等。教育数字化、智能教育和科学教育在国外同样受到研究重视，重点关注其教学创新、评价认证、关键能力等。此外，国外学者还关注新时代下的技术促进教学、学习分析的技术与实践、游戏化学习、教育技术发展和生成式人工智能等。

国内外教育技术重点关注领域有交重部分，如教育数字化、智能教育、在线教育等，但重视程度有所差异。此外，交重部分也有存在差异，如国外偏向探讨技术于学习场域中的拓展作用，且较多探讨影响因素、存在关系等，国内则偏向探讨技术于教学中的有效利用。

3. 智能时代教育技术的演进趋势

3.1. 理论变革：植根教育实践问题，交叉重构理论体系

教育技术植根于教育实践问题，立足于教育与技术的内在联系，多年来在传承教育学、心理学和传播学等学科理论的基础上，经过不断融合创新，形成了新的理论支撑。教育技术理论作为一种能够精确描述并深刻揭示教育技术的特有命题体系（W·布雷岑卡和李其龙，1995），在智能时代数据与科技的多重赋能下，愈加重视智能技术、数据科学、脑认知、心智等在教育过程中的功能作用，持续动态地吸纳跨学科融合的创新理念，以多重层级的研究视角，交叉重构教育技术理论体系。随着智能技术在教育领域的深度介入，智能化教育变革是智能时代的必然选择，也是教育技术理论变革的应有之义。

3.2. 技术发展：数智融合赋能创新，适配助力个性精准

教育作为前沿技术的重要实践领域备受瞩目，科技与教育的双向赋能正在成为人类教育发展史上的新命题。以人工智能为代表的新一代新兴技术驱动着教育场景的数智化转变，为推动精细化、个性化、多元化的高质量教育发展提供了强大且稳固的支撑。拓展现实在与 5G、AIoT、云原生等智能技术的共同赋能下融合变革，旨在有效支撑多维学习场景、多样学习方式和多元教学创新等的实现，抓住教育元宇宙的发展机遇；学习分析走向多模态融合学习分析，旨在进一步打破数据孤岛，形成教育联动，为教学与学习过程提供过程性的精准的、个性化的服务与算法；人工智能与知识图谱、数字画像等共同作用，动态支持学生能够自动、科学地调整，以适应其个性化学习需求，指导其个性化学习行为，助力学生的个性精准发展。

3.2. 应用创新：贯彻人机协同理念，推动教学育人创新

教育技术作为教育数字化转型的实现手段，需着力破解现实教育教学问题。如在环境建设方面，教育技术通过对标实际需要，利用多种智能技术构建智能学习空间、智能教室等技

术丰富的教育环境，实现物理环境和虚拟环境的高度融合；在资源共创方面，聚焦在线课程、可计算文档和虚拟仿真实训等资源的建设与应用，满足教育优质均衡发展和学生多样化学习需求；在教学创新方面，通过变革课堂教学形态满足不同课堂教学需要，如在线教学、游戏化学习、自适应教学和人机协同教学等；在教师专业发展方面，通过网络研修、智能研修等推动教研形态转型升级；在评价认证方面，以自适应测验、虚拟情境测验、多源数据评价等多种方式，促进智慧评价科学高效开展；在教育治理方面，发挥教育大数据的赋能属性，重点在数据驱动的教育决策、数据伦理与安全隐私和教育治理现代化等领域开展实践应用。

4. 智能时代教育技术的转型变迁

4.1. 研究深化：洞察预测发展方向，推进高质成果转化

教育是面向未来的根本事业，教育技术的研究理应是未来主义取向的。智能时代的教育技术研究应充分对标社会前沿需求，通过统一的数据标准体系，实现多模态数据的共享与整合（武法提、高姝睿和田浩，2022），并在数据的智能驱动下，人机协同洞察、预测教育技术未来发展方向，提前探索可适应并冲破某种困境的干预最优组合；同时，充分依托教育新型基础设施建设，聚焦教育技术基本理论、智能技术发展、智能时代教育教学教学创新、高质量教育支撑体系与教育评价变革和教育技术学研究范式等研究前沿，以教育研究伦理法则和价值追求发展为基准促进研究范式的转变（Ouyang,F.和 Jiao,P.，2021），积极促进理论与实践的有机结合，推进高质成果的高效转化，增强教育技术研究成果的适用性和实用性。

4.2. 发展动向：推动教育数字化转型升级，培育智能时代创新人才

智能时代技术的创新发展和教育的高质量需求迫切呼唤着教育教学适应新时代、塑造新担当、推动新发展，向着更高层次、更大范围实现数字化转型升级。教育技术作为推进教育数字化进程中最贴近、最有力且最活跃的元素和推力，必须充分考虑智能时代的创新特征，面向智能时代创新人才培养需要，以教育新基建为发展推力，融合认知心理学、神经科学和教育学领域相关理论，依托教育大数据、人工智能等信息科技，构建新兴交叉理论体系，并在泛在真实场景中开展大规模个性化教学活动，推动教育组织管理和教育治理现代化的转型与创新发展，构筑面向未来的教育生态体系，从而有力支撑未来教育技术的研究与实践，使得现代化的教育教学能较好适应智能时代的创新要求，助力智能时代创新人才的培养。

5. 结语

教育技术是推动教育现代化发展的创新实践领域，在支持教育数字化转型和高质量发展中起到举足轻重的作用。我们希望本研究提出的新进展，能够为全面洞察当前教育技术发展趋势，创新教育信息化人才培养，赋能教育数字化转型，助力教育高质量发展提供有益借鉴。

参考文献

中华人民共和国教育部（2023）。**以数字化开辟教育发展新赛道**。

中华人民共和国教育部（2022）。**深入贯彻党的二十大精神 纵深推进教育数字化**。

焦建利（2023）。ChatGPT:学校教育的朋友还是敌人？。**现代教育技术**，04，5-15。

W·布雷岑卡、李其龙(1995)。教育学知识的哲学——分析、批判、建议。华东师范大学学报(教育科学版)，04，1-14。

武法提、高姝睿和田浩(2022)。人机智能协同的精准学习干预：动因、模型与路向。**电化教育研究**，04，70-76。

Ouyang,F. 和 Jiao,P.（2021）。Artificial Intelligence in Education: The Three Paradigms。**Computers and Education: Artificial Intelligence**，2。

数字化转型视角下教育评价的“变”与“不变”

The Variance and Invariance in Educational Evaluation from the Perspective of Digital Transformation

宋琰玉[1]，吕赐杰[2]，赵笃庆[3]，杨九妹[1]，吴龙凯[1,3,4*]
[1]华中师范大学人工智能教育学部
[2]香港教育大学课程与教学学系
[3]华中师范大学国家数字化学习工程技术研究中心
[4]华中师范大学国家智能社会治理实验基地(教育)
[*] longkaiwu@ccnu.edu.cn

【摘要】 数字时代，教育评价正面临以数字化、智能化、智慧化的转型趋势，“变”与“不变”的辩证思维为把握教育评价数字化转型提供了新视野。数字技术融入教育评价，使得评价目的更具实效、评价内容日渐全面、评价依据日趋科学、评价主体日益多元、评价方式逐渐多样。然而，新技术与旧观念的碰撞使得传统评价弊端加重、评价目的迷失、主体地位模糊、数据伦理威胁。为此应把握教育评价数字化转型“不变”的根本坚守，建立“文化-政策执行”链，共建共治共享评价机制、秉持价值先行和以人为本评价理念、构建崇善数据伦理规约。
【关键词】 数字技术；教育评价；转型趋向；价值坚守

Abstract: *In the digital era, education evaluation is facing the transformation trend of digitalization, intelligence and wisdom. The dialectical thinking of "variance" and "invariance" provides a new vision for grasping the digital transformation of education evaluation. The integration of digital technology has made educational evaluation more effective, comprehensive, scientific, multifaceted and evaluative. However, the collision between new technologies and old concepts has aggravated the drawbacks of traditional evaluation, lost the purpose of evaluation, blurred the status of the subject, and threatened the ethics of data. To this end, we should grasp the fundamental "invariance", establish a "culture-policy implementation" chain, build a shared evaluation mechanism, uphold the value-first and human-centered evaluation concepts, and construct an ethical statute for good data.*
Keywords: digital technology， educational evaluation， convergence of transformation， value perseverance

基金项目：本文受国家社会科学基金“十四五”规划 2022 年度教育学重点课题“智能技术赋能教育评价改革研究”（课题编号：ACA220026）资助

1.引言

教育评价作为对教育活动现实的或可能的价值作出科学判定的过程，事关教育事业的高质量发展。近年来，数字技术与教育的双向赋能，催生了诸如无纸化考试、智能书写测评、智能命题等教育评价领域的新生事物，让多模态、全样本、全时空的评价成为可能（刘云生，2020）《深化新时代教育评价改革总体方案》《义务教育质量评价指南》等一系列文件的出台，明确了推动教育评价数字化转型的重要性，提高教育评价的科学性、专业性和客观性。然而，

当社会各界对数字技术加速教育评价数字化转型欢呼雀跃之时，也逐渐形成了“新瓶装旧酒”的尴尬局面。基于此，文章从“变”与“不变”的辩证视角，明晰教育评价数字化转型之“变”，整体审视“变”的风险限度，笃定教育评价核心任务之“不变”，探析教育评价数字化转型之路径，以期深刻把握数字时代教育评价转型的基本逻辑和应有遵循。

2.积极应变：教育评价数字化转型趋向

教育评价数字化转型之“变”指的是教育评价要顺应教育数字化转型趋向，因其势而利导之。数字技术带给教育无限可能，为教育评价提供更便捷、更科学、更精准的手段，将逐步实现彰显本体价值的发展评价、多维度表征的客观评价、数据循证导向的智慧评价、家校社企参与的多级评价以及人技协同的精准评价。

2.1 教育评价目的更具实效：甄别选拔走向完人教育

数字技术的快速发展，推动教育评价从关注“甄别选拔”走向推动生命整全、自主发展、终身学习的完人教育。其一，大数据通过结合起点数据和过程数据对评价结果深度挖掘，以“成长”作为学生学习情况评价的依据，促进学生全面发展。其二，依托物联网、大数据、学习分析等数字技术，实现对个体综合素质水平的数字画像建构、精准诊断与评估，最终实现“小以小成，大以大成，无弃人也”的价值目标。

2.2 教育评价内容日渐全面：片面僵化走向立体多维

数字技术驱动下的教育评价将助推教育评价范围拓展至教育活动的方方面面。其一，通过对师生表情、语言、动作等关键信息进行广泛采集，再利用数据分析技术加以识别，得出多维度的课堂评价，同时依据智能推送技术，为评价对象提供动态、个性、精准的评价结果反馈。其二，评价内容由单一成绩转变为强调以学生核心素养为导向的教育测量与评价体系，基于形成性评价和增值评价开展全过程纵向评价，实现全过程全方位的立体评价。

2.3 教育评价依据日趋科学：经验主导走向数据循证

在技术支持下，教育评价依据将从主观经验判断走向客观数据支持，利用显性的、多维的大数据分析和算法识别进行客观评价。其一，揭示海量教育数据背后隐藏的问题和趋势，实现评价数据分析从经验判断到数据驱动、从寻找因果关系到挖掘复杂相关关系的转变。其二，通过可视化工具、学科知识图谱等技术直观地呈现评价结果，并给学校发展、教师教学能力提升、学生个性化学习等推送相关建议，基于有效证据积极推进教育评价循证决策。

2.4 教育评价主体日益多元：单级评价走向多级评价

数字技术将推动教育评价从“主体结构单一、范式封闭独白”转向家校社企参与的多级评价（黄晶晶和刘宇佳，2022）。其一，教育大数据以开放共享、高速流转和实时获取的优势，保障公众对评价信息的知情权和监督权，不断优化评价主体结构。其二，数字技术通过厘定多元主体间的角色权利、探索不同主体交流互动的对话模式，推动他们由既往“评与被评”的主客体关系转向“共同参与、协商对话、形成共识”的主体间性关系。

2.5 教育评价方式逐渐多样：人工评价走向人技协同评价

人技协同评价目的在于让数字技术成为教育主体智慧的延伸。其一，通过安装智能设备捕捉师生的表情、动作、声音等，对真实课堂状态进行记录，减少评价主体对课堂教学状态的干扰，使评价主体获取更加真实的课堂行为信息。其二，在教学评价实践中，结合数字物联技术传感与识别的优势，辨识学习者本身的行为特点，探寻学习者行为的潜在逻辑，日趋关注学习者情感态度和高阶能力发展，提高教学评价的合理性和科学性。

3.风险透视：教育评价数字化转型限度

数字化赋能教育评价在呈现积极发展样态的同时，囿于大数据的庞杂、算法的隐蔽性以及

评价的复杂性，面临标准、目的、主体和伦理等一系列发展困境。

3.1 数据标准缺位加重传统评价弊端

教育评价数字化的核心是技术赋能，但是如果用“新技术”来评“旧标准”，数字技术反倒还会强化传统教育评价的弊端。一方面，数字化的转型仅仅是评价的外在表现形式的改变，而评价的内在逻辑和方式仍然沿用传统的模式，没有实质性的进步，很可能进一步助长唯分数、唯升学的不良倾向（曹培杰和王阿习，2023）。另一方面，教育数字化趋势更加强调教师利用数据和算法了解和服务学生，但在实践中由于缺乏政策导向指引，容易将新技术异化成随时随地下任务、发指令、填表格、报数据的工具，致使教育评价数字化转型流于形式。

3.2 数据依赖致使评价目的迷失

当数字技术主导教育评价时，对数据的过于依赖意味着思想的数据化，这种工具化规约造成数据评价标签化和数据评价结论化（宋乃庆等，2021）。一方面，教育评价活动具有复杂性与不确定性，基于数据的评价方式很可能将评价者的行为异化为展示，欠缺对实践性任务的考查和测评，束缚学生成长空间。另一方面，数字时代的治理决策权具有一定的信息不对称性，会进一步加剧社会各阶层的“数字鸿沟”（杨欣，2022），因此数字化转型背景下的教育评价很难摆脱“信息独裁”“唯数据论”和“精英治理”等技术垄断的影响。

3.3 数据裹挟导致主体地位模糊

大数据时代，海量数据支撑下个体的自主性与选择性似乎得以提升，但实际上人的主体地位遭遇挑战（靖东阁，2022）。一方面，数据霸权导致学生丰富多彩的生命表征被完全转化为冰冷的测评数值，“漠视”学生作为不断发展变化中独立个体的生命表现，甚至带来教育评价情感基础的丧失和应有“温度”的失真。另一方面，大数据时代的数据权力很可能成为支配他人的隐形力量，基于技术手段掌握评价对象的个人偏好，在某种程度上可以引导个体的思维和行为选择，个体自由意志被消减，对评价数据的依赖不断增强。

3.4 数据堆砌触发数据伦理威胁

大数据本身无好坏之分，但与其他因素相互作用所产生的影响，可能远远超出技术和实践的直接目的（石中英，2020）。一方面，一些数据在评价活动结束之后并未及时清理，可能会被未经授权的机构重新利用，使得评价对象隐私被侵害的风险不断加剧。另一方面，数字技术风险治理存在责任归属不清的问题，因此当数据违规采集、数据隐私泄露等问题出现时，难以界定责任主体与责任范围。当前，不少教育行政部门、学校管理部门等相关组织在相关风险治理问题层面存在权利与义务不清、行动方向不明等现实困境（赵磊磊等，2022）。

4.根本坚守：教育评价数字化转型价值立场

数字技术驱动教育评价转型，对教学评价理论的建构和实践的改进带来巨大冲击。当下必须坚守“不变”的价值立场，回归教育本质，既不能以狂飙突进的技术“裹挟”教学评价改革，也不能以教育评价理论的“自负”无视技术，以价值引领技术与教育评价和谐共生。

4.1 聚焦制度保障：建立“文化—政策执行”链

有效制度供给是数字化赋能教育评价变革的重要保障和必需条件，建立完整的“文化—政策执行”链，完善教育评价转型运行机制。第一，国家层面需完善制度设计和法律法规，针对不同评价对象、不同类型教育的特点，分类设计教育评价制度，体现不同类型评价对象的个性化特征。第二，地方教育部门和各级各类学校应积极贯彻并持续细化国家层面的制度规范，形成具体、可操作、因地制宜的实施办法和保障机制。同时征集有代表性、可推广的典型案例，将优秀经验做法上升为制度设计层面，推动数字技术进步和制度创新的融合共生。

4.2 聚焦评价共同体：共建共治共享教育评价机制

教育评价数字化转型要把握全面性与科学性，必须以数据平台建设为动力，着眼于共享平

台和监管平台的搭建，促进开放式数据生态。第一，构建教育数据共享平台，由教育行政部门牵头，通过一体化数据设计形成统一的、可供各子系统调用的教育数据资源，保证每个评价环节的完整性与井井有条。第二，搭建教育数据监管平台，充分借助区块链技术优势（吴永和等，2020），实现全链条追溯教育数据来源，确保目标方之间交换的教育数据的安全性和隐私性。同时，进一步完善教育评价数据质量审查与问责机制，拟定好示警、追责问责办法。

4.3 聚焦主体立场：秉持价值先行和以人为本教育评价理念

教育评价存在的价值和理由必然是促进人自由而全面的发展，不管是评价者还是被评价者，都要明确一切教育评价必须以服务于人作为根本宗旨。第一，评价者必须端正对技术以及教育评价的本质认识（曹渡帆和朱德全，2023），始终明确技术是工具而不是目的，要以遵守教育评价的本质规律与价值诉求为本。第二，要始终明确教育评价的对象是学生而不是其他。要始终坚持教育评价的育人理念，明确当技术真正能彰显生命价值、有效统整育人目的时，才能被称为是我们需要的教育评价，才是新时代教育评价改革应该追求的理想样态。

4.4 聚焦伦理秩序：构建崇善数据伦理规约

制定数据伦理规约，以此提升伦理责任意识、构建安全、崇“善”、和谐的评价伦理秩序。第一，完善教育评价数据隐私安全的立法规范，充分吸收《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》中的核心条款，进一步制定和完善教育评价数据安全和隐私保护的伦理秩序（鹿星南和高雪薇，2023）。第二，坚守教育评价数据隐私安全的伦理价值，在评价实践中坚守科技向善的伦理态度，尊重和维护教育评价数据采集对象的知情权（檀慧玲和王玥，2023），规避数据采集、存储、传输、分析等环节的数据安全和隐私泄露风险。

参考文献

石中英. (2020). 回归教育本体——当前我国教育评价体系改革刍议. ***教育研究*** **(09)**， 4-15.

刘云生. (2020). 运用现代信息技术开展学生立体评价的时代意蕴与探索思路. ***国家教育行政学院学报*** (10)， 3-10+23.

杨欣.(2022).教育评价改革的算法追问. ***华东师范大学学报(教育科学版)*****(01)**，19-29.

吴永和，程歌星，陈雅云，王筱 & 马晓玲.(2020).国内外“区块链+教育”之研究现状、热点分析与发展思考. ***远程教育杂志*****(01)**，38-49.

宋乃庆，郑智勇 & 周圆林翰.(2021).新时代基础教育评价改革的大数据赋能与路向. ***中国电化教育*****(02)**，1-7.

赵磊磊，吴小凡 & 赵可云.(2022).责任伦理：教育人工智能风险治理的时代诉求. ***电化教育研究*****(06)**，32-38.

黄晶晶 & 刘宇佳. (2022). 试析教育评价数字化转型的原则、场域与方向. ***中国考试*** **(06)**，16-23.

曹培杰 & 王阿习.(2023).新一代数字技术何以赋能教育评价改革. ***人民教育*****(20)**，30-32.

曹渡帆 & 朱德全.(2023).新时代教育评价改革的数字正义何以可能. ***电化教育研究*****(12)**，21-27+34.

鹿星南 & 高雪薇.(2023).人工智能赋能教育评价改革：发展态势、风险检视与消解对策. ***中国教育学刊*****(02)**，48-54.

靖东阁.(2022).后人类时代教育研究的主体性重塑、风险及其规避. ***电化教育研究*****(06)**，11-17.

檀慧玲 & 王玥.(2023).教育评价数字化转型的内生动力与核心议题. ***教育研究*****(12)**，143-151.

面向全纳教育的增强现实教学策略研究——以高中地理课程为例

Research on Augmented Reality Teaching Strategies for Inclusive Education - A Case Study of High School Geography Course

刘嘉纯[1*]，张志兵[2]，林晓凡[1]
[1]华南师范大学教育信息技术学院
[2]广州市第七十五中学
[*] liujc125@qq.com

【摘要】 全纳教育注重接纳所有的学生，而利用基于增强现实技术的教学策略开展教学，能够以其强大包容性和高度交互性来满足所有学生的多样化教学需求。本文基于地理研学特色构建了包含四方面的步骤清晰、包容性强、操作性强面向全纳教育的增强现实教学策略，并通过策略验证，面向全体学生，满足所有学生的多样化教学需求，提高学生的学习成绩和学习参与。以有效的全纳教育方法促进全纳教育的实践发展。
【关键词】 全纳教育；增强现实；教学策略

Abstract: *Inclusive education focuses on accepting all students, and teaching with augmented reality-based teaching strategies can meet the diverse teaching needs of all students with its strong inclusiveness and high interactivity. Based on the characteristics of geography study, this paper constructs a four-pronged augmented reality teaching strategy with clear steps, strong inclusiveness, and high operability for inclusive education, and validates the strategy to meet the diverse teaching needs of all students and improve students' learning achievement and learning participation. Effective inclusive education methods are used to promote the development of inclusive education practices.*
Keywords: Inclusive education, Augmented Reality, Teaching strategy

1. 引言

「自从 1994 年《萨拉曼卡宣言》提出全纳教育(Inclusive Education)的概念后，全纳教育思潮由此诞生，并进行了广泛实践」。各国的政策文件纷纷提出实施全纳教育的相关细则。「2023 年，教育部部长怀进鹏在《数字变革与教育未来》的演讲中表示，推行全纳教育，实现学有所教、有教无类是我们共同的价值追求」。但就目前我国的教育实践来说难以实现这一追求。而增强现实 (AR) 技术能够以其高度沉浸式特点引起学生的兴趣，激发学生的学习动力，使学生沉浸于真实情境中，提高学生的学习参与度，促进学生与场景和教师的交流互动，满足学生自主化、个性化的学习需求。因此本研究提出了面向全纳教育的增强现实教学策略，以期利用增强现实技术来促进当前的全纳教育，实现公平包容、个性发展的教育。

2. 全纳教育概念界定

1994 年，世界特殊需要教育大会首次提出全纳教育(Inclusive Education)的相关概念：每个人都有其独特的个性、兴趣、能力和学习需要，学校要接纳全体儿童，并满足他们的特殊教育需要(An, Hu, & Horn, 2018)。

2005 年，联合国教科文组织发布《全纳教育指南: 确保全民教育的通路》报告中将全纳教育定义为: 全纳教育是通过增加学习、文化与社区参与，减少教育系统内外的排斥，关注并满足所有学习者多样化需求的过程,以一种详细全面的术语来描述全纳教育的内涵和外延(王善峰 & 郭冬梅, 2011)。

基于上述界定，不难发现，全纳教育体现为接纳所有的学生，并通过各种特定的手段和方法来满足学生的多样化需要，帮助学生增长知识，适应正常生活和社会发展，减少教育系统内外的排斥，使所有学生共同生活学习。

然而目前全纳教育在实施中仍面临挑战。教师在常规教学中难以满足所有学生的个性化要求，导致教学不完整甚至影响其他学生。同时，现有资源难以解决学生的多种学习需要。因此，创设一个有效、可实施的教法至关重要，能解决学生需求被忽略的问题并推动全纳教育发展。本研究拟利用增强现实技术构建高互动性、沉浸性和包容性的教学情境，以满足学生个性化需求，提高学生参与度，促进交流协作。通过准实验研究验证增强现实教学策略，以期提出一套具体有效、可迁移的全纳教育教法。

3. 面向全纳教育的增强现实教学策略构建

在总结当前研究的基础上，以教学系统设计理论、建构主义学习理论和全纳教育理念为依据。「利用 AR 构建虚实结合的教学环境，注重环境对学习者的影响；以学生为中心，激发学习者的学习兴趣，强化学习者的学习意愿，满足学习者的学习需求，为学生提供学习的指引，引导学生主动转变为知识建构的主体；并注重教与学的过程，帮助学生获得知识，参与教学过程，营造良好的教学氛围」。 结合高中地理研学课程和学生特点，初步构建了以下面向全纳教育的增强现实教学策略。

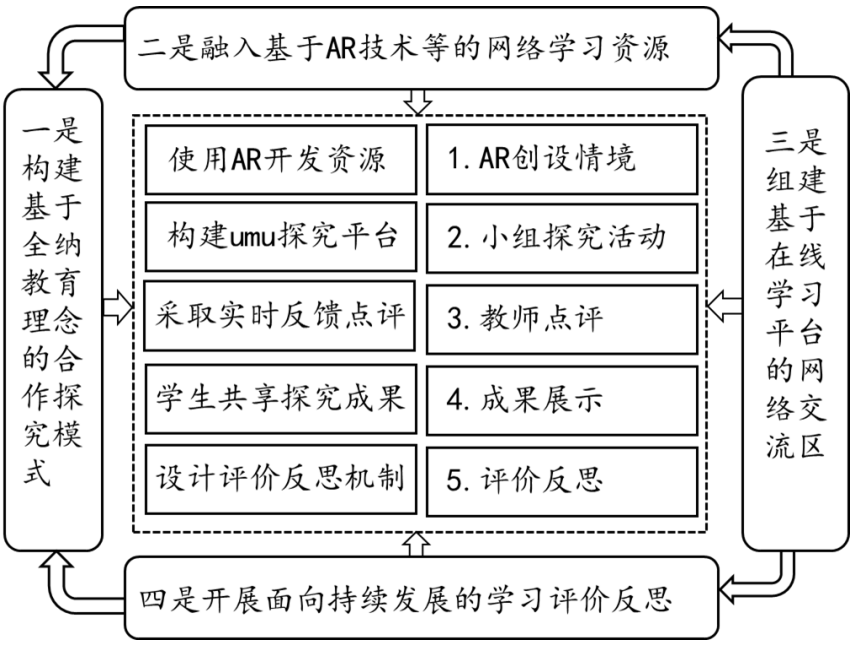


图 1 面向全纳教育的增强现实教学策略图

3.1. 构建基于全纳教育理念下学生全体参与的合作探究模式「全纳教育理念以崇尚民主植根于世界民主进程中，强调学生的主体地位，最大程度发挥学生的主观能动性，因此关注合作学习，提倡通过合作学习让全体学生都能够参与到教学中」。「而探究教学同样注重以和谐

民主的教学氛围为目标，注重学生参与学习的全过程」。研学则以实地研究考察为出发点，强调在实地探究中进行学习。因此，本文以全纳教育理念为指导，发挥合作探究的重要作用，构建基于全纳教育理念的合作探究模式，融合以 AR 技术为支持的的合作探究活动，通过新奇的 AR 资源、有趣的研学基地以及教师的指导辅助，促进全体学生积极有效参与和合作，实现学生之间的信息和资源的探索、共享与创新，以及相互倾听、交流、学习和尊重。构建基于全纳教育理念的合作探究模式是面向全纳教育的增强现实教学策略的核心策略。

3.2. 融入基于增强现实技术等的网络学习资源 在构建基于全纳教育理念的合作探究模式基础上，运用增强现实技术制作相关的 AR 探究资源，辅助合作探究研学。首先以新奇的 AR 模型创设生活情境，增强现实的代入感，激发学生的兴趣，引起学生对研学的积极性，促进学生的主动参与，让学生自发观察和探索。同时，在 AR 资源的设计中考虑学生的学习层次和个体差异，以层级递进式提供相应的 AR 资源，满足全体学生的多样化需求。其次运用希沃白板 5 等软件制作相关的自主学习资源，包含研学过程的相关知识概念补充，供学生在合作探究过程随时根据需要进行自主学习获取知识，辅助学生开展探究学习及集体合作。

3.3. 组建基于在线学习平台的网络交流区 在实地研学中，由于现实条件限制可能导致部分学生未能参与其中，进而导致学生产生学习孤独感、与群体疏离，这违背了全纳教育理念，每个学生都是班级不可或缺的成员之一。因此，使用 UMU 在线学习平台组建在线学习交流区，提供一个网络交流区域，使所有学生随时随地进行远程讨论，即使学生由于现实条件限制导致未能参与实地研学，也能通过其他学生录制分享在在线交流区中的基于 AR 的视频以及在讨论区的实时记录，与开展实践研学的学生远程共享讨论，营造融洽和谐的实践氛围，使所有学生都能参与其中。通过该在线学习交流区贯穿整个合作探究研学，提供给有任何需要的学生随时进行求助，即使存在地理空间距离，教师也可实时关注到所有学生的情况，进而进行针对性的辅助教学，满足其多样化的需求。

3.4. 开展面向持续发展的学习评价反思 在开展面向全纳教育的合作探究研学后，设计教师、学生评价与反思的机制，能够为改进后续探究研学提供方向。同时，基于全纳教育理念，在评价内容的选择上即关注学生的共性特征，又关注学生的个性表现。根据学生的表现教师积极给予学生反馈，尊重每个学生的个人情况，引导学生主动反思，发现自己的优势，并发挥潜能，实现全面可持续的发展。

4. 面向全纳教育的增强现实教学策略的效果验证

为检验面向全纳教育的增强现实教学策略的有效性，明确面向全纳教育的增强现实教学的具体效果，检验该策略是否能够满足学生的多样化需求并提高学生的学习成绩和学习参与，本研究采取准实验方法，选取高二年级两个班级共 56 名学生进行教学实践。分为实验组和对照组，实验组以面向全纳教育的增强现实教学策略进行教学，对照组以传统讲授式教学策略进行。正式课程中通过发布测试卷收集课堂前测学习成绩和后测学习成绩，课程测试卷为以及对正式教学过程进行课堂实录，对课堂教学过程进行数据收集和观察分析。

在课程结束后，本研究对收集的相关数据进行分析。为检验采取不同教学策略的学生在学习成绩上有何种差异，本文以面向全纳教育的增强现实教学策略为自变量，学生的后测成绩为因变量，学生的前测成绩为协变量，首先对两组学生的前测成绩进行了同质性检验，Levene 方差相等检验结果表明，协变量对自变量之间不存在相互作用($F = 0.014, p = 0.910 > .05$)，即可以进行 ANCOVA 分析。结果如表 1 所示，根据分析结果，实验组和对照组调整后的平均值为 56.508 和 36.825；在学习成绩方面，在排除前测成绩的差异后，实验组在使用面向全纳教育的增强现实教学策略后的学习成绩优于对照组。

表 1 学习成绩 ANCOVA 分析结果

组别	N	Mean	SD	Adjusted M	F
实验组	37	55.50	11.292	56.508	6.124*
对照组	19	37.83	13.045	36.825	

*在.05 水平（双侧）上显著

通过课堂实录观察分析发现，采取面向全纳教育的增强现实教学策略的班级，在实际课堂表现中更加积极和活跃，课堂能包容所有同学的需要，学生的参与度较高，且通过对教师的辅助指导，学生能够更加融洽得参与到小组探究活动，并在小组探究讨论中能够有较高质量的发言，同时也能够主动地对其他小组的探究成果进行对比评价，客观地表达出自己的观点，进而能够对学校有更高的包容性感知，同时学习成绩也更高。任课教师也表示平时在课堂上不参与的学生在本次课程中反而担任了小组内探究讨论的记录员，同时也能积极地表达自己的观点。相对而言，采取传统的讲授式教学策略的班级，在总体上课程的参与度不如实验班，学生的自主思考讨论时间较少，部分无法跟上教师节奏的学生多处于走神阶段，当学生进入讨论思考环节时，也由于缺乏主动探索的动机，学生的参与度较低，且讨论发言的质量也较低，学习成绩也相对较低。

5. 结语

新时代的教育是公平包容的教育。新时代育人需要在教学实现学有所教、有教无类，使所有学生都能获得公平包容的教育。本研究在策略中面向全纳教育，设计了增强现实教学，并开发了配套的 AR 资源，提出了一种面向全纳教育的策略，在注重使用 AR 等相关资源帮助学生建构知识的同时，通过自主学习、协作探究等行为活动让学生积极主动得参与学习，满足学生的多样化学习需求。为信息化时代下教师利用增强现实技术进行全纳教育提供实践指导。但本文提出的增强现实教学策略仅在高中地理课堂进行实践，未能在多学科、多学段进行教学实践，因此不能全面反映学生在校的全纳教育情况，后续还需持续开展多学科的教学实践，对教学进行不断反思，不断修改优化教学策略，使其能够发挥最大的全纳教育作用。

参考文献

怀进鹏(2023).数字变革与教育未来. **中国教师报**, p. 001.
索静.(2020).教学设计理论的演变及其现实意义.**教育理论与实践** (07), 61-64.
吕耀中.(2008).全纳教育视野下的课堂合作教学.**全球教育展望** (06),36-39.
王菊芝.(2018).建构主义学习理论与课堂教学设计.**中国校外教育** (36), 144-145.
王善峰 & 郭冬梅.(2011).从“接收”到“接纳”:农民工子女义务教育的新需求.当代教育科学 (20),8-11.
吴佳莉,胥炜 & 但菲.(2018).全纳教育理念下的探究教学. **中国特殊教育** (01),15-19.
吴庆华和邓大龙(2022).随班就读课程调整的实践困境与优化路径.**中国教育学刊(6)**, 4.
An, Z. G., Hu, X., & Horn, E. (2018). Chinese inclusive education: The past, present, and future. Intervention in school and clinic, 54(2), 118-122.

“双减”政策的实施现状、问题成因及优化对策

-基于微博的数据分析

The Current Implementation Status, Problem Causes, and Optimization Strategies of the "Double Reduction"Policy

-Data Analysis based on Weibo

吴晓斐[1]，梁琼月[2]，万海鹏[1*]
[1]首都师范大学教师教育学院
[2]首都师范大学教师教育学院
[1*]首都师范大学教育学院
*dnvhp@163.com

【摘要】“双减”政策自公布以来，便迅速成为社会热点话题，引起各界的广泛讨论。本研究通过微博平台采集与政策相关的评论数据，进行了两次情感分析，探究网民对政策实施情况的态度变化及原因。第一次情感分析使用 Rostcm 软件，并提取了高频词，结论是网民对“双减”政策的实施总体持积极态度，同时总结了政策实施中的问题。为了实时跟踪政策的实行情况，一年后，研究进行第二次情感分析，发现网民对政策的消极情绪上升，并且之前存在的问题仍未解决。针对这些问题，本研究也提出了优化对策，以促进政策的有效执行，早日达成减轻学生负担的目的。

【关键词】“双减”政策；微博；高频词；情感分析

Abstract: *The "double reduction" policy has sparked extensive discussions since its announcement. This study collected policy-related comments on Weibo and conducted two sentiment analyses to understand netizens' attitudes and reasons for policy implementation. The initial analysis used Rostcm software to extract high-frequency words, revealing a generally positive attitude towards the policy but also highlighting implementation issues. A follow-up analysis a year later showed an increase in negative emotions among netizens, indicating unresolved issues. To address these concerns, the study proposes optimization measures for effective policy implementation and reducing student burdens promptly.*

Keywords:the "double reduction" policy ,Weibo,high-frequency words,Sentiment analysis

基金项目：科大讯飞横向委托课题“基础教育智能教育产品评估与应用研究”（课题编号：iflytek2021-002）的阶段性成果。

1. 前言

2021 年 7 月 24 日，中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生学业负担和校外培训负担的意见》（以下简称“双减”政策），政策一经出台，便迅速成为社会热点，引起了网民的大范围讨论，形成了网络舆情。减轻义务教育阶段中小学课业负担是推进我国素质教育落实、促进学生德智体美劳全面发展的重要举措，因其多元性、复杂性等特征受到全社会各类人群的共同关注。（李刚，2020）在此前的研究中，鲜少有学者

基于大型社交媒体平台而对“双减”政策进行讨论，忽视了政策实施后的的舆情分析。而根据微博官方统计数据，截至 2021 年四季度末，微博月活跃用户达到 5.73 亿，为研究网民对于“双减”政策的态度与观点奠定了良好的数据基础。本研究基于新浪微博平台（以下简称微博），通过分析网民与“双减”政策相关的评论，总结政策试点过程中网民的反馈及其情感态度，从中提取出政策实行中出现的问题，并提出针对性的建议与对策，以期为“双减”政策的持续落地提供一些数据参考。研究进行了两次情感分析，都是为了了解微博网民对于政策实行的态度情况，数据采集的相隔时间为一年。

2. 研究设计

本研究以微博为平台，通过八爪鱼采集器，以双减、双减政策为关键词，先搜集了 2021 年 8 月 25 日-2022 年 4 月 9 日的相关博文，共计 2023 条。经过处理后，得到有效数据 1260 条。通过 Rostcm 软件对保存的 1260 条博文进行分词，提取高频词，并制成高频词表；接着对所有微博数据进行情感分析，这一过程同样是借助 Rostcm 软件实现，软件分析后能得到情感分布统计图、不同情绪的分段统计结果以及不同情绪的具体博文数据的文本文件。对第一次数据分析的结果进行讨论，明晰网民对于当前的政策试点情况的态度，再结合具体博文，从中提取出政策试点过程中出现的问题。

第二次情感分析搜集了 2023 年 6 月-2023 年 9 月的与“双减”政策相关的微博评论数据。通过删除新闻类、与研究主题无关等评论数据，最终得到有效数据 400 条。本次情感分析的工具是百度智能云，在 Python 环境中对数据进行逐条分析，准确度比起第一次情感分析有所上升。从两次情感分析的结果这一角度出发，探究网民态度变化的原因，并讨论第一次情感分析中提取出的问题是否得到解决，以及是否有新问题的出现。研究流程如图 1 所示。

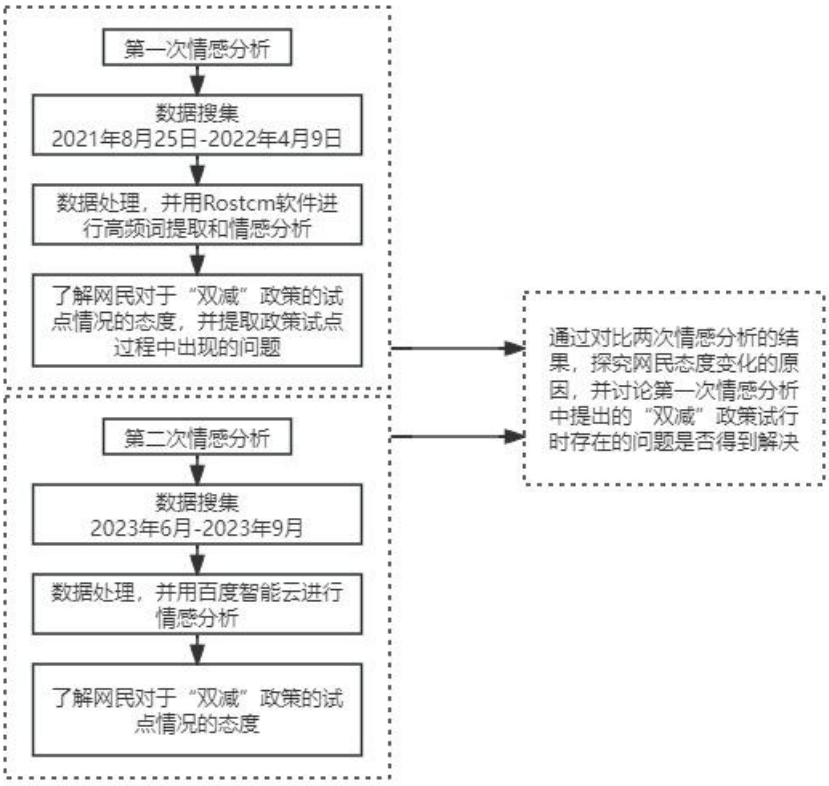


图 1 研究流程图

3. 第一次情感分析

3.1. 基于Rostcm软件的高频词提取

本研究首先通过Rostcm软件对保存为txt格式的1260条博文进行分词，以便于后续高频词的提取；完成分词后，进行词频分析，根据得到的词频分析结果，构建过滤词表；接着进行社会网络和语义网络分析，借助过滤词表过滤一些无意义词汇，得到过滤后的高频词文件；得到文本格式的过滤高频词后，考虑到数据的直观呈现，选择高频词表前20个词，制作成表格，如下表 1。

表 1 高频词表

教育	679	教师	186
学生	580	教学	172
孩子	569	机构	162
老师	400	学科	161
作业	385	小学	156
学校	349	校外	129
学习	306	中考	112
家长	303	国家	111
培训	220	考试	108
疫情	203	课堂	107

从表 1 中可以看到教育、学生、学校、培训、命题、机构、作业、老师、教学、学科、服务、中考、疫情等词语频次较高，即讨论度和关注度都相对较大。结合表 3 及所对应的微博数据，网民对“双减”政策的讨论大致可以分为以下几个方面：校外培训治理、作业多样化、考试命题变化、课外活动开展等。

3.2. 基于Rostcm软件的文本情感分析

借助Rostcm软件自带的情感词典，对分词后的文本进行情感分析，共计562条发言。“双减”政策在微博平台上得到了广泛讨论，而根据采集到的数据分析结果来看，积极情绪占比较大，为59.25%；消极情绪居中，为38.79%；中性情绪最少，为1.96%（如表 2 所示）。其中，中性情绪反应的微博只有少数几条，参考性不大；积极情绪中以政策执行短期内带来的效果和对政策长期的展望为主；消极情绪聚焦于学校、教育培训机构从业人员的压力、部分民众对于政策的不理解以及家长的焦虑。

表 2 微博网民的情感分布统计表

情绪分类	占比/%	数目/条
积极情绪	59.25%	333 条
中性情绪	1.96%	11 条
消极情绪	38.79%	218 条

表 3 微博网民的情绪分段统计表

情绪分段	积极情绪		消极情绪	
	数目/条	占比/%	数目/条	占比/%
一般	183	32.56%	161	28.65%
中度	67	11.92%	37	6.58%
高度	83	14.77%	9	1.60%

由表 2 和表 3 可知，虽然网民的整体态度倾向于支持“双减”政策落地，但持消极态度的也不在少数，归根结底还是政策并没有解决网民当前的问题。网民对于政策实行的积极态度来自于：“双减”政策的执行及时控制了校外培训乱象的进一步恶化，部分家长的退学费问题得到解决，培训机构逐渐开始调整转型；学校积极落实“双减”政策的要求，开展多样的课外活动，丰富作业形式，减轻学生学习压力；部分地区中考命题方式开始转变，由原有的“唯分数论”转为考察学生的综合素质等。

结合表 3 所对应的微博数据，不难看出网民的消极情绪来源是较为集中的，突出对政策是否执行的怀疑，即政策虽然出台了，但地方没有执行到位，具体表现为：学生课后补课现象仍然存在（学生在微博上发博吐槽）、家长的焦虑并未得到缓解（ #双减后家长的焦虑如何减#词条引发关注）、学生自身与学校的升学压力（高中生并未吃到双减的红利）、教师工作时间延长、教培从业人员大量失业、疫情带来的系列影响（家长收入不稳定、学生被迫居家学习）等。在负面态度的微博数据中，校外培训是网民的讨论集中点之一，它暴露出的的问题之一便是资本的大量介入，以致教育失去了原有的公益性与普及性。早在 2018 年便出现大量教育培训机构倒闭无法退还预付学费事件，包括“韦博英语”“学霸一对一”“于斯钢琴”“爱乐乐享”等机构，最高涉案金额超亿元。（苏沐晖，2020）经济条件较好的家庭，家长支付高额的学费来实现学生的个性化需求；而对于无力支付额外学费的学生或是家长来说，他们享受不到更加丰富的、针对性的教学资源，教育的不公平更加凸显。二十年前已有学者指出，“如果利润将成为教育目的，受损的将是青少年一代和社会”。（王善迈，2000）

除了校外培训之外，补课、疫情、作业也是网民负面情绪的重要来源。补课现象主要是因为传统的“唯分数论”升学方式，学校为了自身发展与学生的将来考量，不能将“双减”政策贯彻到位，学生的作业量并未减少。大批学生参加校外培训，根本原因是个体学习需求在学校得不到有效满足，而学校整齐划一的教育方式，使得许多家长与学生将目光转向了校外教育培训。（薛二勇、李健和张志萍，2021）再加上疫情的影响，学生不定时的需要在家上网课，学习质量得不到保证。

此外，“双减”政策在课后服务方面提出了要求，但实际执行时因地域、经济等客观原因，难以达到目标。有调查显示，教师认为“听评课制度加重了教师的工作负担”占 44.3%，认为“听课任务太重，使我和同事们疲于应付”占 47.2%，认为“教研活动太频繁了，我和同事们总是疲于应付”占 45.3%。（徐雪莲，2009）随着双减政策的实施，教师的工作量和工作时间大幅增加，教师不仅要保证完成工作，还要兼顾家庭和个人生活。在采集到的微博数据中，有部分教师发博抱怨在校时间过长，以致家庭不和睦。额外的付出当获得一定的回报，但事实上，教师更多付出的是无偿劳动，因而参与课后延时服务的热情并不高。

4. 第二次情感分析

第一次情感分析使用的是Rostcm软件中的情感分析功能，该功能的实现借助于软件本身的情感词典，但与实际博文数据对比发现，情感分析结果的偏差较大，情绪识别并不准确，因此，在进行第二次情感分析时，将工具更换为百度智能云，逐句识别情感倾向，提高准确率。

第一次情感分析得到的结论是：网民对于“双减”政策的试点情况总体是积极向好的，但也存在一些消极情绪，从消极情绪对应的评论数据中可以提取出网民讨论的政策试点过程中存在的问题。政策试行两年后，本研究又进行了第二次数据采集和情感分析，了解当前网民对于政策实行情况的态度，分析两次情感分析结果的异同，并讨论第一次情感分析时的问题是否得到解决，以及是否有新问题的出现。

在“双减”政策试点实行一段时间后，本研究采集了自2023年6月至2023年9月的微博评论数据，对数据进行了一系列处理后，又进行了一次情感分析，以期探究“双减”政策试点前后，微博网民的态度转变情况，从中发现政策试点中存在的问题是否得到解决以及目前仍存在的问题。第二次情感分析的统计结果如表 4所示。

表 4 第二次微博网民评论的情感分析

情绪分类	占比/%	数目/条
积极情绪	25.50%	298条
中性情绪	0.00%	0条
消极情绪	74.55%	102条

此外，本研究对采集到的400条有效数据进行分类，依据发博者的角色定位，将其简单划分为教师、家长及学生三类，并抽取部分微博评论数据，制成表 5。

表 5 微博评论数据分类

微博数据分类	具体博文
教师	如今的教育环境是怎么了？双减到底减在哪里？我只能看到学生天天特别累，没有喘息的时间，他们有的只是做不完的作业，订正不完的单词，背不完的课文！！！在那样短暂的甚至没有休息的时间里，他们要语数英物化政史地生来回切换……我同情他们，可我不得不追赶大家的脚步，不然教学进度会慢，会被淘汰，会被批评…所以，谁能告诉我，双减体现在哪里，我没有看到
	总结一下学校要求： 双减背景下不让多留作业，每天自习只有120分钟，但是每周都要检查老师作业批改；上课要带教案，但是不能看教案；鼓励老师利用下班时间辅导答疑，但是不能在班级也不能在办公室；周五考完试，但是周一要成绩；不让多留作业，但是题量要加大，一科至少75分钟；不知道这个前后矛盾前言不搭后语的会开有个屁用

	谁家小学生排第三节晚修啊是我们学校什么双减啊不存在的老师累学生也累每天从6点半起床7点就开始早读哪位家长赶紧举报一下吧
家长	我周围的前同事也是，双减之后从机构拿了一笔赔偿金，然后开小班，收入比双减前翻了一倍不止 只要教的好，根本不需要愁招生，家长排着队报名。
	双减减了个啥？小孩累大人累，回家那么迟，一看，还有一堆学习任务等待完成。我比他还不想上学……
	特别支持国家搞双减，时间用来搞搞音体美不香吗？特别是体育，有百利而无一害!用来阅读也行!小家伙肯定高兴坏了！
学生	学校天天说着双减双减 周末作业还是多得要死 周末不是用来休息的吗 布置那么多作业还不如不放假学校你是有什么心事吗 嘴上说着双减 结果作业一点没减 无语了就是
	这图片还只有周五布置的还不带周六补课布置的我真的服了双减难道就是说一下嘛这学真的不想上了我现在真的想s明天中午吃完饭不到十分钟就有课这™让人怎么活啊真的要吐了™今天晚上又睡不了了真的我现在心里万M奔腾的想把学校炸了一天13节课让人咋活啊下课老师在拖拖堂在叫叫人在去给老师讲讲题我真的要疯啦
	我天天晚上作业写到一二点我惹谁了还不如早点放学减了什么我问你减了什么？美其名曰双减，压力只增不减，这个政策有什么意义吗？考试难度竞争压力不变，还不允许上辅导班，想让我死直说

由表 4 的统计结果可知，与第一次情感分析的结果相比，网民的消极情绪的占比有较大程度的上升。结合表 5 分析，在第二次情感分析中，“双减”政策实行过程中仍存在相同问题，即这些问题一直没有得到缓解。第一次情感分析时，讨论的校外培训机构收费不合理及不退费情况得到改善，对此问题的讨论减少。但仍存在政策实行不到位，学校课业依旧较为繁重；政策执行后，孩子学到的东西有限，家长担心孩子未来不能很好应对社会生活；家长和学生仍有补课需求；课后服务时间过长等问题，没有得到有效缓解。

综合两次情感分析的结果来说，校外培训的治理一直在加强，目前的培训市场较为有序，不退费、胡乱收费的现象得到改善。但是“双减”政策的减少中小学生教育负担的目标仍未实现，还存在一些问题，需要政府、学校以及社会共同努力。

5. 建议

经过两次情感分析的讨论，目前“双减”政策实行过程中仍存在一些问题：政策实施不到位，减负政策并未落实，学生课业负担较重；家长和孩子都有补课需求，且因为政策影响，补课的收费增加；课后服务时间没有有效利用，并在无形中增加了老师的工作负担。而针对

上述问题，本研究也提出了建议，希望能够改善政策实施的不合理情况，提高民众对教育的期待感。

5.1. 各级政府加强对各级学校落实“双减”政策的监督

首先，建立监督机制。各级政府可以设立专门的部门或机构负责监督学校对“双减”政策的执行情况。这些机构应具备监督、评估和指导的职能，通过定期检查和评估学校的落实情况，发现问题并及时采取措施解决。其次，加强信息公开与透明度。政府可以要求学校公开相关政策文件和执行细则，确保政策的透明度和公正性。此外，政府还可以建立在线平台或信息系统，提供学校的落实情况和相关数据，供社会监督和评估。再者提供投诉举报渠道。政府应建立健全的投诉举报渠道，鼓励学生、家长和社会公众对学校的违规行为或不当操作进行举报。政府应及时处理并调查这些举报，确保学校按照政策要求进行减负措施的实施。最后，促进社会参与。政府可以建立家长和社区参与机制，鼓励他们参与学校的管理和监督工作。可以组织家长会议、座谈会等形式的活动，听取家长和社区的意见和建议，共同监督学校的减负工作。

5.2. 构建智能化“家-校-社”育人共同体

基础教育阶段学生的负担问题涉及到学校、家庭和社会多个方面，“双减”政策落实需要家庭教育、社会教育和学校教育建立协同育人体系，形成育人合力，共同攻破学生学业负担过重顽疾。（陈晓慧，2022）在构建共同体的过程中，智能技术大有可为，可以利用技术构建三方的及时沟通平台，使家长和社会能够更多地参与到学校的管理中，实现资源共享，从而构建高质量教育体系。学校可以开放课堂、开放教研、开放学生学习活动，增进家长对学校课程和课堂教学变革的认识，引导家长理性、科学把握当前和未来课程改革的趋势，理解学校育人方式变革的最新动态。（董红军，2021）孩子未来的核心竞争力，从来都是学习能力。重视孩子基础能力的培养、学习习惯的养成，是双减下应有的家庭教育理念。同时，学校要注重资源的筛选和利用，推动高质量社会资源走进社会和家庭生活中，既能为学校进行补充，又能满足学生的个人兴趣培养需求。在此前提下，可以考虑引导校外培训机构转型成为学校教育的技术、资源储备军，也能适当缓解大量校外培训机构倒闭带来的社会问题。“双减”时代，信息技术应回归其教育功能本位，充分考虑学生与教师的需求。在此意义上，“双减”和社会建设的跨界合作已经具备了必要的政策共识，相应的，跨界推进家校社的协同育人共同体建设自然构成了“双减”的得当之策。（陶蕾和杨欣，2022）

5.3. 利用大数据技术减轻教师负担

面对教师的过度压力，其他社会非营利组织和社会人员（大学毕业生、青年志愿者、退休人员、有技能的专业人士等）是课余服务的强大力量。另一方面，学校可以引入大数据技术，减轻教师负担。如学习管理系统可以收集和分析学生的学习数据，包括课程进度、作业完成情况、考试成绩等。通过这些数据，教师可以更好地了解学生的学习情况，及时给予个性化的指导和支持；提供在线培训课程和资源库，帮助教师获取最新的教学方法和资源。同时，大数据技术还可以提供教师社群和合作平台，促进教师之间的交流和共享。校内课后服务是政府主导、以学校为主体开展的课外服务，具有普及化、公益化性质。（都晓，2022）在“双减”政策中，也明确规定了，学校课后服务的目的是帮助学习有困难的学生辅导答疑，为学有余力的学生拓展空间，开展丰富的活动，而不能用作讲新课。此外，教师要将课堂教学、作业布置和课后服务和专业学习统整起来，引导广大教师投身教学，研究教学，切实提高教学质量。（徐晓东，2022）

6. 结语

我国基础教育阶段学生负担过重的问题自新中国成立便一直存在，过去十几年来，国家也出台了许多针对性的政策来解决这一问题，新时期的“双减”政策更加彰显了国家治理的决心。针对上述提出的政策实施中存在的问题，本研究提出了针对性的建议，以求政策能早日惠及民生。“双减”政策是国家大力度改革教育制度，实现应试教育向素质教育转变的政策保障，其出台与实施对基础教育来说具有重大意义，如何解决目前存在的问题、推进其更好落实，是每个教育人的共同期望，更需要国家、社会以及家庭多方面的共同努力。

此外，本研究也存在一些问题：两次分析的工具不同，可能会影响数据的准确度；筛选后的有效数据量不足，研究结果覆盖面不完全；在数据分析上不够深入，缺少对发言人的地域、年龄等的分析，研究还需进一步改进。

参考文献

王善迈(2000).教育服务不应产业化.求是,(01),52-53+57.

李刚(2020).十年来我国义务教育阶段减轻中小学课业负担的成效与建议.湖南师范大学教育科学学报,19(03),94-101.

陈晓慧(2022).“双减”时代智能技术的可为与能为——基于“家—校—社”协同育人视角.中国电化教育,(04),40-47.

苏沐晖(2020).教育培训机构频现“套路跑”预付费是原罪?.新产经,(02),80-83.

都晓(2022).“双减”背景下的课后服务研究述论.新疆师范大学学报(哲学社会科学版),(04),1-12.

陶蕾和杨欣(2022).“双减”的价值证成与时代路向.中国电化教育,(03),7-15.

徐雪莲(2009).将教师减负进行到底.中小学管理,(04),29-30.

徐晓东(2022).“双减”政策下全班教学法的改革.中国电化教育,(04),26-33.

董红军(2021).“双减”背景下的家校协同共育.中国教育学刊,(S2),196-199.

薛二勇、李健和张志萍(2021).校外教育培训治理的形势、挑战与路径.中国电化教育,(08),1-8.

数字技术促进青少年心理健康的研究进展与启示

—基于 2013-2023 年文献的系统性综述和元分析

Digital Technologies to Promote Mental Health of Children and Adolescents: A Systematic Review and Meta-Analysis of the Literature From 2013 to 2023

陈天娇[1]，欧婧依[1]，李格格[1]，罗恒[1*]
[1]华中师范大学 人工智能教育学部
[*]luoheng@mail.ccnu.edu.cn

【摘要】 近年来，青少年心理健康问题日益凸显，数字技术因其便捷性和多样性受到关注，利用其促进心理健康已被证明是可行且有效的。本研究通过系统性文献综述和元分析，对近十年数字技术促进青少年心理健康的研究进行统计分析。结果表明：（1）相关研究数量逐年上升；（2）常见心理健康问题包括抑郁焦虑、精神疾病和欺凌等；（3）众多技术手段中，应用程序的使用率最高；（4）元分析结果显示，数字技术对心理健康的促进作用具有中等且显著的效应量。基于上述发现，本研究对相关实践和研究提出建议，期望为我国青少年心理健康发展提供有益参考。
【关键词】 心理健康问题；青少年；数字技术；系统性文献综述；元分析

Abstract: *The prevalence of mental health problems among children and adolescents has prompted a growing number of researchers and practitioners to explore technology-enhanced interventions to address this problem due to their convenience, diversity, and proven effectiveness. To synthesize the research progress on technology-enhanced mental health intervention in the past decade, this study conducted a systematic review and a meta-analysis. The results show that: (1) there is an upward trend of relevant research publications; (2) the common mental disorders include depression, anxiety, peer victimizations, etc.; (3) Among the various technological means, mobile application has been used most frequently; (4) the meta-analysis results reveal that technology-enhanced interventions yield a moderate and significant effect size for promoting mental health. Based on the study results, several practical and research implications are discussed with the purpose of benefiting metal health development of children and adolescents.*

Keywords: Mental health problems, Teenagers, Digital technology, Systematic literature review, Meta-analysis

1. 研究背景

近年来，青少年的心理健康状况及相关教育问题已引发社会的广泛关注。当前，数字技术的发展带来了社会和经济的深刻变革，为心理健康的诊断和干预提供了新的可能性。已有研究证明利用数字技术对青少年进行心理健康的诊断和干预是可行并且有效的。心理健康的内涵包括病理性取向、积极取向和完全取向三种不同的观点（Keyes, 2009），本研究将心理健康定义为由完全取向所主张的一种心理状况。

使用数字技术手段促进心理健康指的是通过数字化设备，以基于互联网平台的形式提供干预措施（Hollis et al., 2017）。当前已有多种数字技术被试用于青少年心理健康问题的治疗中，其干预效果受到诸多因素影响（Grist et al., 2019），在真实环境中的效果尚不明确（Folker et al., 2018）。基于此，本研究旨在梳理国内外近十年数字技术促进青少年心理健康的相关研究，采用系统性文献综述法和元分析探究哪些技术类型和特性能够改善心理健康问题，以期为我国青少年心理健康发展提供实践经验和行动指南。

2. 研究设计

2.1. 研究方法

本研究主要采用系统性文献综述法分析国内外数字技术促进心理健康的相关文献。此外，本研究还对符合标准的文献进行元分析，计算数字技术的使用对促进心理健康的效应量大小，并探究潜在变量对数字技术干预效果的调节效应。

2.2. 研究问题

为梳理相关研究现状，提出以下研究问题：（1）目前国际上数字技术促进青少年心理健康相关研究现状如何？（2）哪些数字技术特性支持青少年心理健康发展？（3）数字技术促进青少年心理健康的有效性如何？哪些因素影响数字技术干预的有效性？

2.3. 文献检索与筛选

2.3.1. 检索文献

在 Web of Science 核心数据库选择 SCIE 和 SSCI 索引数据集，以“mental health or psychological health or psychological wellbeing” AND “technology or technological or technologies or digital media” AND “K-12 or teenager or children or adolescents or youth”为字符串检索。同时以“心理健康” & “技术” OR “数字技术” & “青少年” OR “K-12”为中文检索词在中国知网中的 CSSCI 数据库中检索，检索时间为 2013 年 1 月 1 日至 2023 年 7 月 1 日，累积获取文献 1032 篇。

2.3.2. 文献质量评估

为确保文献与研究问题的高相关性，根据检索到的文献制定了纳入和排除的标准，具体如下：（1）文献是否重复；（2）文献是否可全文获取；（3）文献是否为期刊论文；（4）文献是否为综述论文；（5）文献主题是否与促进心理健康高度相关；（6）文献对心理健康的干预是否使用数字技术；（7）文献关注的研究对象是否为青少年（6-18 岁或平均年龄<18）。

本研究遵循系统性文献综述方法，基于以上标准，对检索到的文献开展了三轮筛选，同时结合文献阅读以滚雪球方式补充新文献，最终纳入有效样本文献 59 篇。具体流程见图 1。

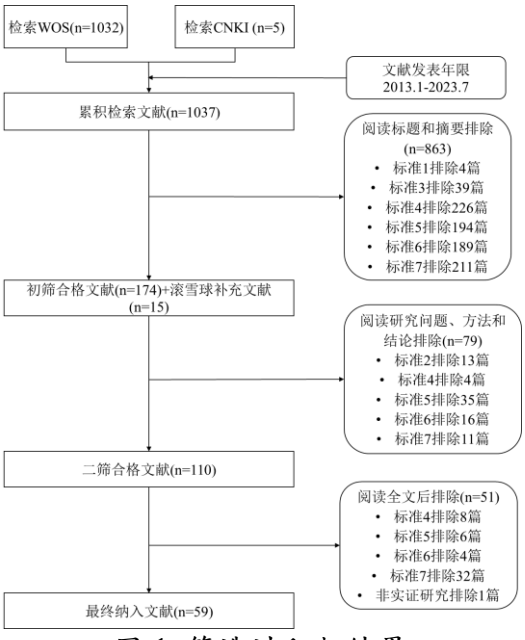


图 1 筛选过程与结果

2.4. 文献信息编码

为了从纳入文献中得到关键信息，在阅读全文掌握基本内容的基础上，我们对 59 篇文献进行了系统性分析，从以下几个方面进行编码：（1）文献基本信息，包括第一作者，发表年份、发表地区、研究类型、研究对象、干预时长；（2）研究所使用的技术类型，包括应用平台或应用软件、聊天机器人、严肃游戏、VR/AR、短信干预、远程医疗服务和其他；（3）心理健康问题，包括抑郁和焦虑、精神疾病、欺凌、社会情感能力缺失、新冠引起的心理健康

问题和其他；（4）实验数据。文献编码由本文研究者通过多轮商议确认编码规则和细节之后共同完成，并在编码过程中针对产生的问题进行集中讨论，确保编码的一致性和准确性。

2.5. 元分析过程

为探究技术促进心理健康的有效性，本研究使用元分析对 59 篇文献中满足数据要求的文献进行分析，包括五个阶段：（1）文献筛选，需对比“技术干预”与“传统干预”、报告可生成效应量的完整数据、因变量至少有关心理健康的其中一方面；（2）效应量计算，采用 Hedges's g 作为效应量指标；（3）模型选定，采用 Borenstein 等人（2009）提出的建立固定效应模型和随机效应模型方法；（4）主效应和调节效应检验，选定模型后进行主效应检验，如存在异质性，则进行调节效应检验；（5）发表偏倚检验，先利用漏斗图直观地对发表偏倚进行定性评估，再进一步通过 Begg 秩相关法和剪补法定量评估发表偏倚。

3. 研究结果

3.1. 样本文献总体概述

如图 2 所示，从发表时间来看，随着数字技术的发展，2013 至 2021 年间文献数量呈现逐渐上升趋势，近五年文献数量占比达到了 76.3%，在 21 年达到顶峰，年发文量达 15 篇，可能由于新冠疫情引起的社交隔离、停课停学、课外活动减少等问题激化了青少年儿童的心理健康问题，吸引了更多研究者开展了数字技术应用于心理健康治疗的创新探索。

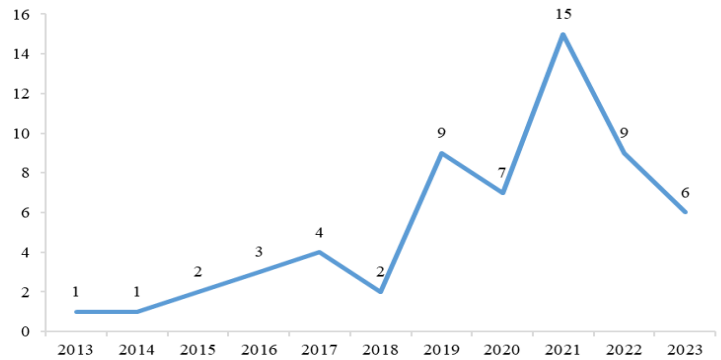


图 2 近十年文献发表数量趋势图

从发表期刊来看，所有文献发表在 41 种期刊上，其中，发表在医学类期刊的文章数量最多，其次是心理学类期刊。表 2 显示了样本文献的来源分布和文献类型。从第一作者所在国家看，来自美洲的文献数量最多，占比为 40.7%，来自非洲地区的文献最少。从实证研究类型来看，以实验研究为主，其次是混合研究，调查研究和设计研究数量都较少。样本文献中来自中国的论文有 2 篇，分别是实验研究和混合式研究。从研究对象来看，年龄跨度为 6-18 岁，未根据学段进行严格划分，但总体而言，关注 13-18 岁的青少年的文献较多。

表 1 样本文献编码结果

地区	样本量	实证研究					
		实验研究	定量研究	定性研究	混合研究	调查研究	设计研究
亚洲	6	4	1		1		
欧洲	19	11	1	1	2	2	2
美洲	24	12	3	2	5	1	1
大洋洲	9	2	0	3	2	1	1
非洲	1	1					
合计	59	30	5	6	10	4	4

3.2. 心理健康常见问题及数字技术干预类型

3.2.1. 抑郁症和焦虑症

在 59 篇文献中，有 25 篇文献是利用数字技术治疗青少年抑郁和焦虑相关症状的研究。在众多数字技术类型中，有 19 篇文章使用应用程序或教育网站作为干预工具，占比达 76%，其次，严肃游戏、聊天机器人和 VR 的文献各为 2 篇。

应用程序概念广泛，通常包括可以在应用商店下载至移动设备的软件。应用软件和教育网站普遍被用作技术干预的一种工具，当结合适当的教学策略时最有效。例如，Gladstone 等人（2015）针对青少年抑郁症预防，开发了一个含 14 个可选模块的交互式网站，每个模块中都应用了教学设计理论。此外，应用程序和网站能够将认知行为疗法（CBT）与数字技术相结合，以减少错误认知和加强积极行为（Neumer et al., 2021）。此外，严肃游戏、聊天机器人和虚拟现实（VR）技术也在抑郁症治疗中有所应用。

3.2.2. 精神疾病

本研究将精神疾病定义为除了抑郁和焦虑之外的神经系统发育问题，在综述的 59 篇文献中，有 10 篇文献被编码为精神疾病。针对精神疾病的治疗，其中基于 CBT 的移动应用程序出现了两次，其他技术类型还包括短信干预、严肃游戏、远程视频会议和移动传感技术。

与治疗抑郁症和焦虑症的应用程序相似，患者认为应用程序具有较好的可用性和易用性，既能保护隐私又能鼓励他们坦诚分享（Adams et al., 2021）。短信干预是通过为患者发送特定内容的文本信息的一种干预方式。Owens 等人（2016）曾试图通过短信减少有非自杀性自伤行为的青少年的自残行为，但最终效果不尽人意。

3.2.3. 欺凌

欺凌一般包括传统欺凌和网络欺凌。在相关的 9 篇文献中有 3 篇采用了短信干预的手段，2 篇采用移动应用程序，此外，聊天机器人、技术支持的课程和基于认知行为疗法的远程医疗服务也被用于治疗当中。

针对欺凌的短信干预技术，实现了个性化定制的功能（Ranney et al., 2019）。移动应用程序提供个性化学习的机会，促进心理健康技能的习得（Davidson et al., 2019）。聊天机器人引导研究对象完成任务，引发自我反思、提供策略建议（Gabrielli et al., 2020）。数字技术支持的课程和基于 CBT 的远程医疗服务利用技术的便捷性，解决传统治疗的局限性。而远程医疗服务利用远程通讯技术提供心理服务，增加了时间的灵活性。

3.2.4. 社会情感能力缺乏

社会情感能力通常指发展青少年的情绪智力（De la Barrera et al., 2021），是衡量青少年心理健康水平的重要指标。拥有积极社会情感能力的人群更少出现心理健康问题。在 6 篇社会情感能力干预研究中，应用程序、严肃游戏、虚拟现实技术和短信干预都被常态化使用。

这些文献都强调了数字技术的交互设计对于社会情感技能提升的重要性。例如，Cherewick 等人（2021）设计了一款手机应用程序，内嵌了丰富的教学互动功能，给学习者带来愉快的学习体验并提升了教学的灵活性。

3.2.5. 新冠疫情引起的心理健康问题

近年来全球新冠疫情大流行为青少年心理健康带来了严峻挑战，加剧了心理健康风险，导致了孤独、痛苦、社会孤立等症状的增加。5 篇文献中报告了新冠疫情导致青少年心理健康问题的加剧，而技术由于不受时空的限制，成为了疫情期间治疗的首选方式。

3.2.6. 其他

除了上述提到的几种常见的心理健康问题之外，文献中还提及了通过数字技术手段改善身体形象焦虑、住院引起的心理失调和阅读障碍等干预类型。

3.3. 技术促进心理健康元分析结果

3.3.1. 文献纳入及编码

对于符合元分析要求的文献，在系统性综述编码的基础上，按照以下变量逐条进行细化和分类：（1）基本信息(作者、年份、样本量）；（2）年龄阶段，按照年龄阶段分为小学、初

中和高中三类；（3）心理健康问题，分为抑郁症、欺凌、新冠引起的心理健康问题；（4）技术类型，包括应用软件、远程医疗、聊天机器人；（5）干预时长，按照干预时长小于一个月和一个月以上分别编码为短时和长时干预；（6）效应量。

3.3.2. 数字技术对心理健康结果的整体影响

异质性检验结果表明，Q 检验显著（$p<0.001$），说明样本间存在较大的异质性，故应选择随机效应模型。合并效应量为 0.43，根据 Cohen（1992）的标准，将 0.2、0.5 和 0.8 分别看作效应量小、中、大的界限。由此可见，数字技术促进心理健康的效应量中等且显著，同时，95%置信区间的下限均大于 0, 说明该效应量由偶然因素引起的可能性非常小。另外，I^2值为 78.164，研究之间的异质性较高，因此可能存在重要的调节变量（Higgins & Green, 2008），需进一步进行调节效应检验。

3.3.3. 调节效应检验

分别对年龄阶段、心理健康问题、技术类型和干预时长四个变量进行调节效应检验。如表 4 所示，在四个调节变量中，只有年龄阶段的调节效应显著（$p<0.05$），其中，小学阶段的效应量最大，其次高中阶段也有中等程度的促进作用，另外，虽然初中阶段的效应量值很小，但是仍然显著，可能与文献数量较少有关。结果还表明，心理健康问题、技术类型和干预时长三个变量的调节效应不显著，但是可以看出，数字技术手段对于新冠引起的心理问题治疗作用效应量最大，而远程医疗服务相比应用软件和聊天机器人来说，能够取得更好的效果，在治疗时长方面，短时干预的效应量大于长时干预。

表 2 技术的调节效应检验（随机效应模型）

调节变量	类别	k	N	g	95% CI	异质性		
						Q	df	*p*
年龄阶段	小学	1	15	0.786	0.231 ~0381**	7.935	2	0.019*
	初中	3	1779	0.151	0.060 ~ 0.242*			
	高中	6	2256	0.539	0.131 ~ 0.946*			
心理健康问题	抑郁症	6	2238	0.290	0.058 ~ 0.522*	3.023	2	0.221
	新冠引起心理问题	3	1797	0.588	-0.059 ~ 1.235			
	欺凌	1	15	0.290	0.244 ~ 1.418**			
技术类型	应用软件	4	1821	0.148	0.059 ~ 0.237*	5.791	2	0.055
	远程医疗	5	2212	0.690	0.253 ~ 1.127**			
	聊天机器人	1	17	0.005	-0.912 ~ 0.921			
干预时长	短时干预	4	159	0.681	0.138 ~ 1.224*	2.280	2	0.131
	长时干预	6	3891	0.242	0.071 ~ 0.414**			

3.3.4. 发表偏倚检验

通过漏斗图检验，我们发现研究的效应值分布呈现出不均匀、不对称地分布在平均效应值两侧，初步说明可能存在发表偏倚。因此，进一步采用 Begg 秩相关法进行检验，Begg 秩相关法是用秩相关检验定量识别偏倚的办法，适合小样本研究。检验结果显示 $t=0.267$ $p=0.283$，$Z=1.01<1.96$，表明不存在明显的发表偏倚。最后，采用剪补法对效应量左右两边文献进行剪补后，发现效应量仍显著。综上所述，并不存在不可忽视的发表偏差。

4. 总结与启示

4.1. *主要发现总结*

本研究通过对近十年数字技术促进青少年心理健康的 59 篇文献进行系统性综述，分析了面向不同心理健康问题常用的技术干预类型和特性，并使用元分析考察了数字技术促进心理健康的实际效应量和潜在调节变量，主要研究结果如下：

（1）在过去十年间，尤其是 2013-2021 年，数字技术促进青少年心理健康的相关文献数量总体呈现上升趋势，其中接近 80%的文献都发表在医学类期刊中；（2）数字技术最常用于干预 13-18 岁青少年的心理健康问题，低年龄阶段（6-12 岁）的孩子得到的关注相对较少；（3）抑郁症和焦虑症是研究中关注最多的心理健康问题；（4）应用程序最常用来治疗心理健康问题，仅有 3 篇文献中使用了较难实现的虚拟仿真技术 VR 用于心理健康治疗；（5）数字技术对青少年心理健康问题具有显著的促进作用，在小学和高中阶段促进效果最为显著。

4.2. 相关启示

基于系统性综述和元分析结果，本文从实践应用和科学研究两方面提出相关启示。针对心理健康教育服务人员，我们建议：首先，需要充分发挥数字技术特性，根据不同心理健康问题选择恰当的数字技术干预工具。例如应用程序更适用于抑郁、焦虑和精神疾病的治疗。其次，对于心理健康咨询师或学校心理健康工作者来说，需要考虑学习者特征、技术干预时长等问题。与之前研究结果不同（Grist et al., 2019），我们发现年龄阶段的调节效应显著，因此治疗师需要根据不同年龄阶段的青少年采取个性化技术干预措施，而短时干预似乎能获得更大的效应量，因此，需要注意干预的时间不易过长，可能会导致青少年出现技术免疫。

针对科研人员，我们建议：首先，需要更多的实证研究报告第一手实验结果。现有研究多为仅表述实验方案，缺少关键研究结果，希望未来研究尽可能全面的报告结果数据，提高元分析结果的可信度和可靠性。其次，元分析中调节效应部分类别文献数量较少，未来研究需要针对已有研究中关注较少的人群开展相关研究，从而增进对技术干预心理健康的理解。最后，很少有研究分析成本效益，而成本效益是决定技术干预是否能够常态化、可持续实施的关键，未来的研究中对基于数字技术干预措施的成本效益需进行充分的调查和报告。

参考文献

Adams, Z., Grant, M., Hupp, S., Scott, T., Feagans, A., Phillips, M.L., et al. (2021). Acceptability of an mHealth app for youth with substance use and mental health needs: Iterative, mixed methods design. JMIR formative research 5(12), e30268. doi: 10.2196/30268.

Davidson, T.M., Bunnell, B.E., Saunders, B.E., Hanson, R.F., Danielson, C.K., Cook, D., et al. (2019). Pilot evaluation of a tablet-based application to improve quality of care in child mental health treatment. Behav. Ther. 50(2), 367–379. doi: 10.1016/j.beth.2018.07.005.

De la Barrera, U., Postigo-Zegarra, S., Mónaco, E., Gil-Gómez, J.-A., and Montoya-Castilla, I. (2021). Serious game to promote socioemotional learning and mental health (emoTIC): a study protocol for randomised controlled trial. BMJ open 11(12), e052491. doi: 10.1136/bmjopen-2021-052491.

Folker, A.P., Mathiasen, K., Lauridsen, S.M., Stenderup, E., Dozeman, E., and Folker, M.P. (2018). Implementing internet-delivered cognitive behavior therapy for common mental health disorders: A comparative case study of implementation challenges perceived by therapists and managers in five European internet services. Internet interventions 11, 60–70. doi: 10.1016/j.invent.2018.02.001.

Gabrielli, S., Rizzi, S., Carbone, S., and Donisi, V. (2020). A chatbot-based coaching intervention for adolescents to promote life skills: Pilot study. JMIR human factors 7(1), e16762. doi: 10.2196/16762.

Grist, R., Croker, A., Denne, M., and Stallard, P. (2019). Technology delivered interventions for depression and anxiety in children and adolescents: A systematic review and meta-analysis. Clinical child and family psychology review 22(2), 147–171. doi: 10.1007/s10567-018-0271-8.

Higgins, J.P., and Green, S. (2008). Cochrane handbook for systematic reviews of interventions: Cochrane book series. Hoboken, New Jersey, USA: Wiley.

Hollis, C., Falconer, C.J., Martin, J.L., Whittington, C., Stockton, S., Glazebrook, C., et al. (2017). Annual research review: Digital health interventions for children and young people with mental health problems – a systematic and meta-review. J. Child Psychol. Psychiatry 58(4), 474–503. doi: 10.1111/jcpp.12663.

Keyes, C.L.M. (2009). "Toward a Science of Mental Health," in The Oxford Handbook of Positive Psychology. (Oxford, England: Oxford University Press), 88–96.

Neumer, S.-P., Patras, J., Holen, S., Lisøy, C., Askeland, A.L., Haug, I.M., et al. (2021). Study protocol of a factorial trial ECHO: Optimizing a group-based school intervention for children with emotional problems. BMC psychology 9(1), 97. doi: 10.1186/s40359-021-00581-y.

Ong, J.G., Lim-Ashworth, N.S., Ooi, Y.P., Boon, J.S., Ang, R.P., Goh, D.H., et al. (2019). An interactive mobile app game to address aggression (regnatales): Pilot quantitative study. JMIR serious games 7(2), e13242. doi: 10.2196/13242.

Ranney, M.L., Patena, J.V., Dunsiger, S., Spirito, A., Cunningham, R.M., Boyer, E., et al. (2019). A technology-augmented intervention to prevent peer violence and depressive symptoms among at-risk emergency department adolescents: Protocol for a randomized control trial. Contemporary clinical trials 82, 106–114. doi: 10.1016/j.cct.2019.05.009.

融入激励机制的小学编程类课程设计与实践研究

Research on the Design and Practice of Primary Programming Curriculum Incorporating Incentive Mechanisms

梁琼月[1]，吴晓斐[2]，万海鹏[1*]
[1]首都师范大学教师教育学院
[2]首都师范大学教师教育学院
[1*]首都师范大学教育学院
*dnvhp@163.com

【摘要】 信息技术的快速发展使得编程教育在中小学阶段的重要性日益凸显。中小学生信息素养的培养成为教育界关注的焦点。然而，小学编程课堂目前存在学生兴趣不足和教学资源有限等问题。为了解决这些问题，本研究探讨了在小学编程课程中应用激励机制的实践效果。通过结合课前目标激励、情境激励、课中情境激励、情感激励以及课后评价激励，旨在提高学生的学习兴趣和编程能力，同时培养他们的信息素养。研究发现，激励机制的应用显著增强了学生的学习动力和参与度，提高了教师的教学效果和资源利用效率。总结来说，激励机制在小学编程课程中发挥了积极的作用，为小学编程教育的推广提供了有益的参考和借鉴。同时也对培养中小学生的信息素养具有重要意义。

【关键词】 编程教育；激励机制；学生参与度

Abstract: *The rapid development of information technology has underscored the importance of programming education in primary and secondary schools. The cultivation of students' information literacy has become a focal point in education. However, there are issues in primary school programming classes, such as lack of student interest and limited teaching resources. To address these problems, this study examines the effectiveness of using incentive mechanisms in primary school programming courses. By incorporating pre-class goal incentives, situational incentives during class, emotional incentives, and post-class evaluation incentives, the goal is to enhance students' interest, programming abilities, and information literacy. The study found that the application of incentive mechanisms significantly improved students' motivation, engagement, as well as teacher effectiveness and resource utilization. In conclusion, incentive mechanisms have a positive impact on primary school programming courses, providing valuable insights for promoting programming education and cultivating students' information literacy.*

Keywords: Programming education, Incentive mechanisms, Student engagement

基金项目：科大讯飞横向委托课题“基础教育智能教育产品评估与应用研究”（课题编号：iflytek2021-002）的阶段性成果。

1. 前言

在当前人工智能快速发展的背景下，编程教育逐渐得到了更多的关注和重视。小学阶段是孩子们思维和创造力发展的重要时期，也是培养信息科技素养的关键时期，派珀特(Seymour Papert)在儿童教育领域中，开创性地引入了 Logo 编程语言。他强调，编程教育不仅是为了让儿童掌握编程语言，更是在于培养儿童的认知能力和问题解决能力。通过学习编程，儿童能够将计算机思维与他们的思考和学习方式相结合，以应对实际问题，并为他们未来的学业和职业发展打下坚实的基础(Papert S., 1996)。然而，传统的编程教学方式往往过于注重语法和算法的学习，缺乏实际的应用和创新，容易让学生感到枯燥和困难。并且许多学

校缺乏专业的编程教师和足够的教学资源，也给编程教育的推广带来了很大的挑战。因此，如何在小学阶段开展有效的编程教育，是当前教育领域面临的一个重要课题。在这样的背景下，我们将激励机制引入到小学编程类课程的设计与实践中，以此来激发学生对编程的兴趣和热情，提高他们的学习积极性和参与度。

2. 小学程序设计类课程的现状及分析

以图形化编程软件 Scratch 为例，程序设计类课程更多的是培养学生的逻辑思维与综合思考能力。西班牙胡安卡洛斯国王大学的研究人员在小学四五年级英语课堂中引入了 Scratch，并与传统教学课堂进行了对比研究。经过一学期的教学活动，他们发现接受 Scratch 教学的小组在英语成绩上取得了更大的进步，并且对计算机教学的态度也更加积极(孙立会，周丹华，2020)。基于模块的图形化编程与传统文本编程教学方式相比较，前者不再需要高标准的基础知识的认知，至使用户年龄段更加不受限制，这样较低年龄段的学生可以专注于对解决问题的逻辑进行思考而不是对于高出认知阶段的文本编码知识进行超前的学习(刘敏，汪琼，2022)。当前小学编程课堂存在的问题可以从以下层面进行分析：

2.1. 学生层面的问题 学习者先验知识不同。由于学生的家庭背景、教育经历等不同，导致他们的编程起点差异很大。有些学生可能已经通过课外辅导或自学掌握了一定的编程技能，而其他学生可能对此一无所知。这为教师在课堂上统一教学带来了困难。

2.2. 教师层面的问题 教学方法陈旧且教师素养参差不齐。部分教师习惯于传统的讲授式教学，缺乏对学生兴趣和需求的关注，根源问题在与其他学科对比缺乏专业教师的培训体系(何芯雨等，2023)。编程教育需要创新教学方式，让学生在真实问题情境中参与编程探究活动(李锋等，2023)。

2.3. 课程设置层面的问题 内容难度过高。既有学者提出“非计算机化”儿童编程教育的概念，旨在强调编程语言的设计初衷并非为小学生学习，其语法和概念对于小学生来说难以理解和掌握，儿童编程教育的重心应该超越编程技能训练本身，落实在儿童认知迁移发展教育的本质(孙立会，2023)。

2.4. 评价与反馈机制的问题 评价方式单一，缺乏即时反馈。传统的编程课堂往往只注重学生的编程成果和技能掌握情况，而忽略了他们在学习过程中的态度、创新和合作能力等方面的评价。

为了解决这些问题，我们提出采用情境激励、情感激励、目标激励和评价激励等激励机制，旨在激发学生学习兴趣、提高教学质量并关注学生的全面发展。小学程序设计类课堂包含文本式编程、图形化编程等多种模式，落脚点都在提高学习者对编程学习的兴趣以及参与度以及对学习者逻辑思维能力的培养。我国开展编程教学的时间相比较国外是较晚的，但随着信息技术的发展以及教育政策的推进，近几年编程类教育逐渐受到社会关注，成为小学校本课程、课后托管等课程的选择。目前国内编程类教学大多以传统教学模式进行算法演示的教学方法进行授课。因此激励机制融入文本型编程教学的开展需要对学习者、教学内容、教学环境以及学习资源进行分析，以此来满足达到课程目标。

3. 激励机制融入程序设计类课程的课程设计

3.1. 课程资源的设计原则 激励机制相关的课程资源设计主要以学习内容为基础，将学习目标作为导向。在教学过程中，适合的课程资源将会提高学习者的课堂沉浸感(黄晋，2022)。本研究依据激励机制与编程教学的特点。

激励机制需要与课堂内容相关。激励机制的设计应以课堂内容为基点，旨在提高课堂教学效率，所借助的课程资源如视频、多媒体课件、游戏等都需要以教学目的为导向。脱离了教学内容与教学目标的激励机制会加大教师授课负担。

课程资源需要满足学习者的学习需求。需要给学生提供丰富的学习资源，可以促进学生自主的进行知识建构。要达到以学生为主体的课堂教学，就要确定课程资源的使用主体是学生。

3.2. 激励机制教学框架设计 学习动机与激励机制的选择相互关联。为了更好地激发学习者的学习动机，我们需要根据学习者的特点和需求选择合适的激励机制。学习动机是学习行为的根本驱动力，为了更好地满足不同课程环节中学习者的学习需求，如图 1 所示将教学框架分为课前、课中和课后三个部分，并针对每个部分采用不同的激励机制。

在课前预习阶段，采用目标激励，即设定明确的学习目标和奖励机制，鼓励学习者自主预习新知识，为课堂学习做好准备。同时，运用情景激励，为学习者创设与课程内容相关的学习情境，激发他们的好奇心和探究欲望，使他们更加积极地投入到课前预习中。

在课中教学阶段，注重情境的创设和情感交流。情境激励可以使学生身临其境地感受知识，增强学习的沉浸感，提高学习效果。情感激励则通过设置组间积分竞赛规则以及奖惩机制来增强学习者的自信心和归属感，使他们更加热爱学习。

课后复习阶段，我们采用评价激励，即通过评价学习者的学习成果兑现奖品。评价激励还可以帮助学习者发现自己的不足之处，促进他们自我反思和改进。

4. 融入激励机制的文本型编程教学实践及效果分析

4.1. 案例实施

课程项目基于 Scratch 软件开展，选择的课程内容为难度相同，均需要条件判断语句以及循环语句的课程。本实验在北京市海淀区某小学四年级的课后托管课程中开展，对该班级同学进行 Scratch 图形化编程的基础测试，分数转化为百分比平均分为 71.73 分，说明该班参与实验的同学具有一定编程基础。对该班进行三节融入激励机制的课程教学，并对学生作品完成度进行统计，教学流程图如图 1：

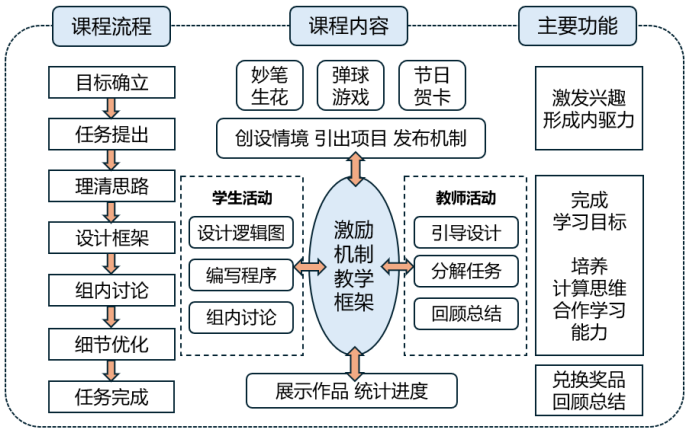


图 1 激励机制教学流程图

4.2. 实施效果

本研究采用定量的方法对三个课程项目的完成度进行统计分析。本研究将每个课程项目的总任务均拆分为两个子任务，通过对班级中参与实验的 25 位同学的完成程度进行统计形成最终结果。结果表明，项目一完成度为 66%。在项目实施过程中，发现学生在角度设置知识点上存在明显的欠缺，导致他们在实现任务二时遇到困难。部分学生反馈自己对于在坐标轴中怎样表示角度的概念和应用不太熟悉，导致最终设计程序失败。在项目一实施过程中存在组

内讨论交流较少，学习者在发现问题时选择向老师寻求帮助的较多，导致课程进度较为缓慢。学习者对激励机制的奖惩机制没有直观体验，存在潜在的激励机制运用效果不佳的问题。项目二完成度为 90%。该项目在实施过程中，观察到学生表现出较高的兴趣和参与度。多数学生表示，项目二中的特效让他们联想到现实生活中的许多元素，激发了他们的好奇心和探究欲望。此外，学生们逐渐理解了课程中的激励机制的规则，这也促使他们更加努力地完成项目。项目三完成度为 86%。尽管在项目实施过程中，学生们面临自定义函数积木步骤较复杂的问题，但通过组内讨论和协作，以及积分奖品的激励，最终完成了该项目。通过对学生作品完成度进行统计，作品完成度在逐步提高，能够完整完成项目的同学也在逐渐增加，说明激励机制对该课堂有正向影响。

5. 结语

激励机制教学设计与激励机制教学框架的融合应用可以有效解决教师教法单一、学生兴趣不高、评价方式单一以及课程难度筛选不当等问题。这种融合不仅能够提高教师的教学水平和能力，还能够激发学生的学习动力和兴趣，提升学习效果和学习成绩。同时，这种融合还能够促进师生之间的交流和互动，增强学习氛围和学习环境的积极性。因此，进一步研究和推广激励机制教学框架与激励机制教学设计的融合，将成为未来教育发展的重要方向之一。

然而，激励机制并不是万能的。在实施过程中，需要注意以下几点：首先，激励机制要公平、公正，确保每个学生都有机会获得奖励；其次，激励机制要多样化，以满足不同学生的需求；最后，激励机制要与课程内容和教学目标相契合，避免形式化。在计算机信息技术发展迅速时代，适当运用新媒体技术与激励机制结合，开展更真切的体验式教学，优化课程内容与结构，以学生的认知能力为基础，对信息科技课程产生兴趣并且能够自主创新主动探索。

参考文献

Papert S. (1996) An exploration in the space of mathematics educations. International Journal of Computers for Mathematical Learning,1(1), 95-123.

孙立会(2023).“非计算机化”儿童编程教育教学模式的构建与应用.现代教育技术(02), 52-60.

孙立会&周丹华(2020).基于 Scratch 的儿童编程教育教学模式的设计与构建——以小学科学为例. 电化教育研究(06),75-82

刘敏&汪琼.(2022).结对编程：中小学编程教育的首选教学组织形式.现代教育技术(03),102-109.

何芯雨,胡子达&陆玟冰.(2023).面向 Scratch 编程课程的知识图谱的构建与应用.计算机与数字工程(09),1990-1994+2158.

黄晋(2022).《少儿编程》课程数字化学习资源在社区教育的应用与效果分析.电脑知识与技术(19),141-143.

新技术赋能的爱国主义教育路径探析

Exploring the Path of Patriotic Education Empowered by New Technologies

章群璐[1*]，蓝靓[1]
[1]杭州师范大学经亨颐教育学院
[*]zqunlu@163.com

【摘要】社会的发展促进着新兴技术的不断涌现，VR、AI 等技术更是点燃了新时代爱国主义教育的火花，赋予了新时代爱国主义教育全新的文化传承路径。新兴技术融入爱国主义教育的可行性和有效性值得不断探索和验证，哪些新兴技术能够赋能爱国主义教育并实现红色文化传承是迫切需要关注的问题。本文旨在通过文献的梳理厘清我国爱国主义教育的研究现状，并通过新技术赋能教育的相关理论来论证和分析爱国主义教育的新技术具体性、适用性，以期构建具体可行的爱国主义教育的新技术支撑路径。
【关键词】爱国主义教育；新技术赋能；教育路径

Abstract: *The development of society has promoted the continuous emergence of technologies, and technologies such as VR and AI have ignited the spark of patriotic education in the new era and given a new cultural inheritance path for patriotic education. The feasibility and effectiveness of integrating technologies into patriotic education are worthy of exploration and verification, and which technologies can empower patriotic education and realize the inheritance of red culture is an urgent issue. This paper aims to clarify the research status of patriotic education in China through the combination of literature; and demonstrate the specificity and applicability of technologies in patriotic education through the relevant theories of technology empowerment education, to build a feasible technology support path for patriotic education.*
Keywords: Patriotic Education, New Technology Empowerment, Educational Pathways

互联网、移动设备、在线平台等为新时代爱国主义文化传播提供了有效载体，爱国事迹、红色故事等的呈现形式不再局限于线下博物馆、纪念馆等。学习者能够在技术的支撑下打破时空限制，随时随地学习红色文化，VR、人工智能等技术的应用更为学习者提供了趣味性、沉浸式学习体验，使其切实感受到生动鲜活的红色文化和历史。毋庸置疑，新兴技术是推动爱国主义教育开展的有效手段，更是促进红色文化源远流长地传承的时代方式。

1.爱国主义教育研究现状和问题提出

目前，国内关于爱国主义教育的文献数量较为丰富，但研究尚未形成较为成熟的体系。

通过文献检索与整理发现，相关文献发表趋势在 1994-1996 年间达到一个峰值，随后趋于稳定状态。2019-2020 年，由于疫情所凝聚的抗疫精神作为爱国主义教育的鲜活教材，学者们开始关注后疫情时代爱国主义教育的路径探索，此时相关研究又达到了一个新的高峰。陈东琼等从中国战“疫”视域下赋予了爱国主义教育新的时代内涵（陈东琼等人，2021）。冯君颢从现实意义方面指出应当将爱国主义教育作为“开学第一课”（冯君颢，2020），而爱国主义教育不仅仅是“开学第一课”，更应是“人生第一课”，将爱国主义教育常态化和将爱国主义元素与生活、学习、工作深度融合是未来中国特色社会主义建设的关键点。

其次，研究大部分聚焦于高等教育，主要关注爱国主义教育与思政课的结合。薛存心从高校思政课档案史料的视角出发，探索了高校思政课创新路径（薛存心，2021）。刘经纬等分析了在高校思政教育中红色资源的利用状况和存在的问题、不足（刘经纬等人，2020）。当然，也有学者关注到了中小学甚至幼小衔接阶段的爱国主义教育，马文萍等以太原市为例探究了红色文化资源如何融入中学生的思政教育（马文萍等人，2021）。王云香从教学方式、教学环境等方面探讨了在小学教学中如何融入红色教育（王云香，2021）。窦文军提出幼小衔接时期是爱国主义教育萌芽的重要时期（窦文军，2021）。爱国主义教育不仅仅应该渗透进思想政治课，更应与所有学科有机结合，把爱国主义元素融入各个学科，进一步推动爱国主义教育常态化。同时，爱国主义教育应关注所有年龄阶段群体，培养每一位中国公民的爱国主义情怀，深化爱国主义教育内涵，推动爱国主义教育由课堂向生活化的方向发展。

新兴技术的不断涌现和发展，对教育领域产生的深刻影响是毋庸置疑的。学者们也开始从新技术视域探索如何利用技术深化爱国主义教育影响。徐瑾等提出爱国主义教育必须运用新媒体和大数据手段，从大数据视域分析了大数据促进爱国主义教育传播的微趋向（徐瑾等人，2020）。任志锋从人工智能的视域探索爱国主义教育该如何变革和应对（任志锋，2021），曹如军等更是将人工智能与思想政治教育课程的开发相结合，探索人工智能与思想政治教育课程开发的基本路向（曹如军等人，2021）。

综上，为推动爱国主义教育的发展，相关学者从不同视域做了较多研究，虽然已有学者对新技术与爱国主义教育的融合进行了初步探索和创新，但对于新技术是否适合融入爱国主义教育，哪些新技术可以赋能爱国主义教育和怎么利用这些新技术开展爱国主义教育，将爱国主义教育常态化、生活化等这些问题尚未有系统的研究。

2.新时代新技术赋能爱国主义教育的现实意义

2.1. 有利于红色文化传承与弘扬的体验化和沉浸化

红色文化是中国共产党历经千难万险所凝聚而成的精神财富和文化遗产，孕育着令人肃然起敬的奋斗历史，更深刻体现着共产党人身先士卒的崇高品德。当新技术与爱国主义教育激烈碰撞，学习者能够利用丰富网络资源进行线上自主学习；能够以体验式、交互式、沉浸式等多样方式进行学习；能够感受新技术赋予学习的科技性、趣味性、体验性等。红色文化的传承和弘扬通过技术赋能带给学习者沉浸式的体验，学习者在多种新技术的体验中能够不断增强学习兴趣和热情，从而培养学习者的时代担当和爱国情怀，使得红色文化的继承和弘扬在新时代绽放出新的光彩。

2.2. 有利于将爱国主义教育常态化与生活化

传播学视域下，爱国主义教育是一个复杂的社会传播过程。在传播过程中，爱国主义教育主体与客体之间通过红色文化、历史等进行信息交流和传递，教育成效与传播效果息息相关。新技术尤其是大数据技术，扩大了教育信息的传播范围，将红色文化渗透至人们的方方面面。同时，爱国主义教育也因技术融入拥有了更便捷的载体工具、更丰富的资源库。优质量的红色文化通过信息平台得以迅速传播，使得学习者利用手机等便携工具随时随地就能收到优质资源的精准推送，从而解决了仅靠课堂讲授为主的爱国主义信息传播效果不佳的局面。爱国主义教育形式不再拘泥于课堂和课外实践，学习者借助智能设备突破时空桎梏，利用碎片化时间及时获取教育资源，爱国主义教育转变为一种常态化、生活化的学习行为。

3.新时代新技术赋能的爱国主义教育的路径

3.1. 多媒体技术赋能爱国主义教育

多媒体技术是指利用计算机对文本、图像、声音等多种信息进行综合处理，实现实时人机信息交互的技术，具有交互性、多样性等特点（徐海等人，1999）。相较 VR、人工智能等技术，多媒体技术更能在日常教学中普遍使用，有效激发学习者学习动机和学习兴趣。

多媒体技术能为学习者呈现耳目一新的视听教育，有效改善传统教学的单调乏味，提高教学效率。教师利用图文并茂等多样化的方式呈现真实的历史和红色文化，有利于学生在真情实感中逐渐形成爱国主义观念（吴笛，2023）。同时，多媒体技术被广泛应用到博物馆、纪念馆的展陈方式中，既丰富了展陈内容的内涵，又增强了博物馆、纪念馆与参观者之间的互动，更能满足参观者多样化的参观需求，增强展陈内容的吸引力。多媒体技术不仅为课堂教学提供技术支撑，充分激发学习者的求知欲望，更能够融入红色文化基地建设，为新时代爱国主义教育的受教者带来无与伦比的视听体验。

3.2. VR 赋能爱国主义教育

虚拟现实技术（Virtual Reality，VR）是利用计算机技术、仿真技术等，创建一个具有视觉、听觉、触觉等多种感官刺激的虚拟环境。VR 技术以其沉浸性、交互性、构想性等特征与红色文化完美结合，构建具体的互动式、体验式的学习情境，从而使学生在交互体验中感受到革命年代的艰辛和新时代社会的变化，更加深刻理解爱国主义的内涵和外延。

我国有学者提出了将 VR 技术融入井冈山红色文化虚拟博物馆的设计构想，将 VR 应用在虚拟博物馆中能够突破实体博物馆的限度边界，减少博物馆资源陈列的损耗成本，从而扩展博物馆的性能（胡小青等人，2021）。2015 年，北京师范大学和 Strong VR 共同合作成立了学生社会实践基地，学生通过佩戴 VR 设备学习圆明园主题相关内容，不仅激发了学生的爱国主义情怀，更为传播保护中华历史文化遗产的理念提供了新的路径（蒋向利，2017）。VR 技术能助力线上爱国主义基地的构建；更能在线下博物馆、纪念馆等设置虚拟体验场景：人们可以与虚拟人物互动交流，或进入 VR 故事场景，在沉浸式体验中感受历史变化，体会历史文化中蕴含的精神；VR 设备还能引入课堂，带给学习者趣味沉浸式体验。如何在爱国主义教育中发挥 VR 技术的优越性将是今后研究的方向。

3.3. 人工智能赋能爱国主义教育

人工智能（Artificial Intelligence，AI）是一门技术科学，它利用计算机、数学、逻辑、机械甚至生物原理和装置来理解、模拟、甚至超越人类智能（孙伟平等人，2018）。在爱国主义教育基地建设中，运用人工智能技术特有的语音识别和语言理解等能力，实现 AI 智能讲解和智能互动，不仅优化了红色文化传播的过程，更实现了智能化的信息服务。而将人工智能技术引入课堂教学，能够打造智能化的教学场景，营造生趣盎然的课堂氛围。2021 年 5 月，南京市雨花台中学利用科大讯飞的机器人开展了一场别开生面的教学，在人工智能技术的支持下鼓励学生创造性地设计恽代英纪念馆智能讲解员，学生在探索和研究的过程中不仅感受到了科技的魅力，更感受到了与科技融合的红色文化的时代魅力。

未来的爱国主义教育要科学利用人工智能技术，通过智能教具、学具（如教育机器人）等载体使得爱国主义教育更加生动有趣。另外，人工智能技术能充分考虑教育对象的年龄、性别、经验等多方面现实需要，从而满足教育对象的差异化需求，实现其个性化的培养。

3.4. 大数据赋能爱国主义教育

大数据是信息收集、数据分析与精准预测的一种新的手段，它根据用户具体需要，对互联网环境中产生的海量数据进行快速而精确的分析与判断，并将分析结果作为使用者进行政

策制定、措施实施和行动推进的重要依据（周晓东，2020）。大数据便捷的信息处理、快速的信息传递和精准的信息搜索等特点，为爱国精神传播和红色文化传承创造了新的环境。

在大数据的支持下，社交网络平台上爱国主义教育内容的实时推送极大地加快了传播速度和广度，同时，数据海量流动和数据的精准挖掘为实现个性化、精准化的教学创造了有利条件。大数据将人们依据年龄、职业等不同背景划分为不同教育受众群体，再依据他们不同的心理需求精准推送教育信息，人们依靠智能终端便能随时随地获取针对性、个性化的爱国主义教育信息，爱国主义教育逐渐转变为一种潜移默化的常态化、生活化教育形式。另外，数据隐私性必须高度重视，利用大数据分析教育受众思想行为动态的同时必须做到合法以及保护其隐私。既要合理高效地革新爱国主义教育，还要切实保障教育对象的隐私和利益。

4.结语与展望

新兴技术为爱国主义教育注入了鲜活动力，技术的便捷性和精准性为实现个性化、常态化爱国主义教育提供了可能。但技术与教育的平衡问题、技术伦理问题等也需要我们批判地看待技术对教育的作用。爱国主义教育本质上是育人活动，新兴技术的发展只是为爱国主义教育带来了方式、形式上的“变化”和创新，爱国主义教育的“育人”终极目的始终“不变”。未来研究和实践必须把握教育真谛，将合适的新技术赋能爱国主义教育，在教学和实践中不断引导学习者树立崇高理想、坚定爱国信念，切实提高新时代爱国主义教育的质量。

参考文献

马文萍,王慧. (2021). 红色文化资源融入中学思想政治教育研究——以太原市为例. *教育理论与实践* (05), 22-24.

王云香. (2021). 红色教育在小学教学中的融合方式探讨. *新课程* (46), 111.

冯君颢. (2020). 爱国主义教育应成为学校“开学第一课”. *教学与管理* (24), 30-33.

孙伟平,戴益斌. (2018). 关于人工智能主体地位的哲学思考. *社会科学战线* (07), 16-22.

刘经纬,高博文. (2020). 大学生思想政治教育红色资源利用研究. *思想政治教育研究* (05), 103-105.

任志锋. (2021). 人工智能背景下爱国主义教育的变革与应对. *教学与研究* (04), 55-62.

吴笛. (2023). 爱国主义精神在历史教育中的培育策略. *中国军转民* (14), 24-25.

陈东琼,吴晓芳. (2021). 中国战“疫”视域下爱国主义教育论析. *学校党建与思想教育* (01), 23-25.

周晓东. (2020). 新时代爱国主义教育的现代化转型——基于话语、图像及大数据的审视. *理论导刊* (11), 124-128.

胡小青, 蔡雨欣 & 许洁. (2021). 基于 VR 技术背景下井冈山红色文化虚拟博物馆设计. *建筑结构* (19), 142.

徐海,吴丽萍. (1999). 多媒体技术在辅助教学中的应用. *电化教育研究* (05), 50-51.

徐瑾,石娟. (2020). 大数据时代爱国主义教育传播“微”检视. *学校党建与思想教育* (24), 17-19.

曹如军,亓梦雪. (2021). 论人工智能与思想政治教育课程的开发. *教育观察* (41), 53-56.

蒋向利. (2017). 高新技术产业发展中的产学研协同创新——北京实创高科技发展有限责任公司产学研协同创新发展纪实. *中国科技产业* (02), 48-51.

窦文军. (2021). 幼小衔接植根红色童心 蓓蕾绽放苗正成长. *中国教育学刊* (S1), 19-21.

薛存心. (2021). 高校思政课建设中的档案史料红色基因叙事创设研究. *档案管理* (03), 15-16.

共同体理论视角下研究生优质教学资源共享机制研究

Research on the Sharing Mechanism of High-quality Teaching Resources for Postgraduates from the Perspective of Community Theory

曾佳尧[1]，赵路瑶[2]，李新[3],杨现民[4*]
[1234]江苏师范大学智慧教育学院
[*]yangxianmin8888@163.com

【摘要】 推动优质教学资源共享是提升研究生培养质量的重要手段。然而，目前我国研究生优质教学资源在共享与发展上面临资源监管缺失、主体参与不足以及共享路径不明等问题。共同体理论指导下的合作强调共同愿景、平等参与、文化建构等，为优化资源共享合作机制提供了重要思路。研究首先系统审视了研究生优质教学资源共享问题，随后基于共同体理论构建了研究生优质教学资源共享机制。最后提出五点研究生优质教学资源共享的实施路径，旨在全面推动研究生优质教学资源的有序共享，助力我国拔尖创新型人才培养。
【关键词】 研究生教育；共同体理论；优质教学资源；共享机制；实施路径

Abstract: *Elevating postgraduate education in China necessitates effective sharing of high-quality teaching resources, yet challenges like inadequate supervision, uneven distribution, and limited subject participation persist. Guided by community theory, this study proposes a mechanism emphasizing shared vision, equal participation, and cultural construction. It systematically analyzes global practices, constructing a community-based sharing framework. Five implementation paths are recommended to optimize resource sharing cooperation. The aim is to facilitate the systematic and comprehensive sharing of high-quality teaching resources, fostering the cultivation of top-tier innovative talents in China's postgraduate education.*
Keywords: *Postgraduate Education, Community Theory, High-quality Teaching Resources, Sharing Mechanism, Implementation Path*

基金项目：江苏省研究生教育教学改革重点课题“江苏高校研究生优质教学资源共享机制和路径研究”(项目编号：JGKT23_B039)；国家自然科学基金面上项目“网络学习资源群体进化的规律识别与预警技术研究”(项目编号：62077030)

1.引言

2020 年教育部三部门联合发布《关于加快新时代研究生教育改革发展的意见》，要求提升研究生课程教学质量，打造精品示范课程，推动研究生优质资源共享（中华人民共和国国务院公报，2020）。2023 年 5 月，习近平总书记进一步强调，要推进数字教育，为扩大优质教育资源覆盖面和教育现代化提供有效支撑（人民日报，2023）。随着教育现代化进程的推进，教育行政部门及高校在研究生教学资源共建共享方面历经从线下共享到混合式共享的转变。从早期的上海研究生联合培养基地建设、京津冀研究生教育协同举措，到近年来国家智慧教育平台上线研究生教育板块、校际协同举办在线课程等实践措施，我国积极探索研究生资源共享并取得了可喜的进展。然而，在落地实施方面，研究生优质教学资源共享仍面临诸多挑战，因此有必要通过优化资源共享机制、强化主体协同合作、建立稳定的供需对接渠道等措施，推动优质教学资源的高效利用与共享，为新时代拔尖创新型人才培养提供坚实支撑。

共同体理论是指不同主体为实现共同目标而形成的社会关系，能够打破个人、时空等局限，更全面地开展协同与合作，以实现个体间的优势互补与共同发展（周赛君，2023）。将其应用于探讨资源共享，涉及共同参与、共同愿景、共享资源三个相互关联的维度，有利于提升主体间的融合度，为优化资源共享合作机制提供重要思路。基于此，本研究首先对国内外研究生教学资源共享的现实问题进行了系统梳理，在此基础上从共同体理论视角出发，构建研究生优质教学资源共享机制，并进一步提出资源共享的发展路径，为我国各区域全面推动研究生优质教学资源的有序共享与高质量发展提供参考。

2. 研究生优质教学资源共享问题

研究梳理发现，国内外研究生教学资源共享的理论与实践都在不断深入推进，为提升教育教学质量、促进教育公平发挥了积极作用，但在资源共享的广度、深度以及常态化开展等方面仍存在一些现实障碍，表现在资源分布失衡、资源监管不足、共享主体协同度待提高以及共享机制与路径不明确等。优质资源本身分布失衡，体现在区域和校际间的资源配置状态失衡、配置效率低下（于妍 & 李明磊，2024）。优质资源缺乏监管，前期缺少优质资源的筛查标准，资源建设质量良莠不齐；中期资源供给方与需求方缺乏互通渠道，致使供需不匹配；后期资源管理保障也不够充足，导致优质资源进化与更新迟缓（李小平,许梦幻 & 赵丰年，2017）。共享主体协同度有待提高，主体在协同过程中面临利益冲突与失衡现象（刘江珅,何岩峰 & 郭登峰，2023）。共享机制与路径不明确，已有研究多是宏观层面的模式构建和理论分析，缺少落地性强的具体路径，限制了资源供给和共享的推进（王亚杰 & 陈岩，2016）。

3. 研究生优质教学资源共享机制

3.1 共同体理论及其指导价值

共同体理论首先首先在社会学领域出现，随后在教育领域的价值日益凸显。理论层面上，学者们进行了教师共同体、学习共同体、实践共同体、学术共同体等诸多研究，为解决学校资源不均衡、教师素质发展等问题提供了参考借鉴。随着技术的发展，传统的共同体也演变为智能时代的教育共同体，并呈现出共同体建设智能化、成员交往形式多元化、参与主体覆盖广泛化以及内容选择多元化四个典型特征（赵冬冬,曾杰 & 朱益明，2023），为校际、区域间甚至国别的共同体建设提供了适切的思路与启示（UNESCO，2022）。实践层面上，教育领域众多学者应用共同体理论来探讨资源共享问题并构建可付诸实践的共同体。诸如技术支持下城乡教育共同体的基本样态与运行模式（童兆平等，2021）、人机协同教育共同体的生态体系（方海光 & 洪心，2023）等，其均从教育问题与技术背景出发，探索共同体理论赋能问题解决的突破点，进而形成可操作、可实践的模式与路径。

在共同体理论指导下构建研究生优质教学资源共享机制应围绕以下三点。一是共享主体涵盖全面化，包括资源共享政策引导者、协同共治者、创新引领者等，同时确保各成员参与并相互认同。二是共享手段智能化，资源共享方式不应局限于一个时空维度，需要将线上平台与线下实践创新混合，甚至探索出更多智能化方式。三是共享内容实现精准化与多元化的统一，资源共享内容既要精准符合“优质”规范和研究生教学需求，也要做到多元呈现。

3.2 研究生优质教学资源共享机制设计

优质教学资源是具有更高水平和领先优势的资源的综合。本研究构建的优质教学资源共享机制是指在促进优质教学资源被广大高校师生获取与有效使用过程中，所涉及的资源共享主体的职能衔接、运行原理，以及实现路径和具体方案等。（见图 1）。

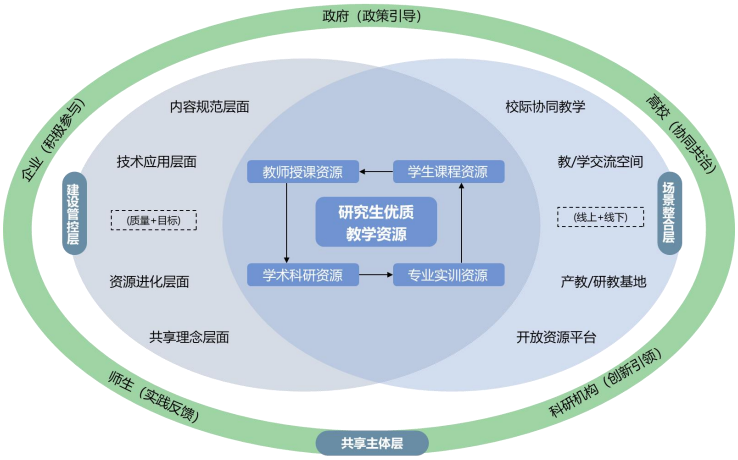


图 1 研究生优质教学资源共享机制

3.3 研究生优质教学资源共享机制阐释

在共同体理论指导下，研究生优质教学资源共享机制构建致力于实现资源的协同共享。机制中各共享主体紧密合作，形成共享主体层，共同追求优质教学资源的共建与共享。同时，基于研究生培养需求和技术背景，构建资源建设与管理的管控层以及线上线下多元混合的场景整合层，实现内容的高质量和形式的多元化，确保资源的高效管理和利用。

3.3.1. 优质教学资源建设管控层

管控层的建设旨在确保研究生教学资源的建设与管理符合教学需求，并达到高质量标准。首先，资源建设应紧密围绕研究生教师授课、课程学习、学术科研和专业实训等关键场景，以实现对教学资源的精准投入和管理。在授课方面，应促进教学模式和理念的共享，同时通过校际合作共享优质课程，缓解资源分布不均的问题。对于学术型研究生，需要重点提供科研材料和促进学术交流的资源；而专业型研究生的培养则更强调实践应用，依赖于线下的实训基地合作和线上的实训案例与资料。其次，为了保障资源的高质量形态，从资源建设的理念到技术应用都需要细致的监管。理念上，资源建设应兼顾科学性和规范性，同时针对四大应用场景的具体需求进行定制化开发。内容形式上，要打破传统的单一模态，探索多样化的资源呈现方式，以满足不同学习风格的需求。在资源进化与生成的管理上，应当注重持续更新和优化，使得优质资源保持其前沿性和有效性。

3.3.2. 优质教学资源场景整合层

场景整合层旨在打造一个覆盖线上线下的全面教学资源供给网络，以实现优质教学资源的跨时空共享，这一网络涉及以下关键渠道（梁传杰,麦立强 & 范涛，2018）。一是校际协同教学，即多所高校之间合作的教学活动模式，其目标是通过特色和进阶课程提升研究生的学术与实践能力。主要采用线下、线上或混合模式进行，包括开设暑期学校、工作坊、专业访学等。二是研教与产教基地，由高校与企业或科研机构联合建立的实践资源，特别针对学术型和专业型研究生的培养而设计。这些基地提供实地实习和科研机会，使学生能够在真实的环境中学习和实践。三是开放资源平台，如国家智慧教育公共服务平台，为学生提供了丰富的在线学习资源，包括视频讲座、课件、互动练习等。四是教与学交流平台，帮助教师提升教学技能，促进学生在学术研究和专业实践中的交流与协作。通过校际和企业之间的合作，形成导师团队，为学生在面对复杂问题时提供专业的指导和支持。

3.3.3. 优质教学资源共享主体层

共享主体层的目标是打造一个联合多方力量的优质教育资源共享共同体，共同推进研究生教学资源的高效利用和共享（李昊灿 & 李妍，2023）。教育行政部门作为政策制定者和推动者，负责确立资源共享的指导方针和政策框架。通过对现有的资源共享状况进行深入分析，识别并解决实施过程中遇到的问题。同时，负责提供必要的资金支持，并处理资源共享过程中可能出现的伦理和隐私问题。高校和企业作为教学资源的主要供应方，负责落实校际协同教学、研教与产教基地等项目。根据教学需求，提供符合标准的高质量资源，并维护开放资源平台的运行。此外，高校和企业应促进知识转化，确保教育资源的实践性与前沿性。科研机构主要负责资源共享体系的理论创新和实践指导，提出创新的资源共享模式和策略。教师和学生是资源共享过程的直接受益者和体验者,其实践经验和反馈是检验资源共享是否成功的重要指标，并直接影响共享主体层其他成员的决策和调整。

4. 研究生优质教学资源共享路径

研究生优质教学资源共享工作的开展可以从以下五个方面推动实施。第一，打造泛在化的资源共享模式，让需求者在任何时间和地点都能够访问到所需的教学资源。通过构建满足多样化需求的资源共享体系，实现资源和需求之间的无缝对接，并利用智能化工具提供便捷的资源访问方式。第二，构建社群化的资源共享组织，旨在促进不同主体间的协作与交流，形成扁平化的网络结构，以提升资源共享的灵活性和效率。同时需要各参与主体明确自己的职能，并适时地进行干预和调整以维护共享过程的顺畅（聂竹明 & 刘钊颖，2017）。第三，

依托智能化的资源共享平台，强调技术在资源共享中的作用。平台应具备资源整合、分区和管理功能，以便用户能够轻松地获取和利用所需资源，同时确保资源的质量和可访问性。第四，建立多元化的资源补偿制度，考虑到资源共享过程中可能出现的利益不均和资源获取不平等问题。通过直接和间接补偿相结合的方式，解决资源损耗和弱势群体获取难的问题，从而保障共享机制的公平性和可持续性（王卫华,田迪 & 邢旭光，2022）。第五，形成责任化的资源监管机制，确保资源共享的各个环节得到有效监督和管理。教育行政部门、高校、企业、科研机构以及师生等多方要协同参与，从资源建设、共享实施到整体调控等多个环节进行精细化管理，以确保共享体系的稳健运行和不断优化。

5. 总结

研究生教育是我国培养高层次创新人才的主要途径，是实施创新驱动发展战略和建设创新型国家的重要基石。针对研究生教学资源共享过程中所面临的挑战，本研究基于共同体理论深入探讨了研究生优质教学资源共享机制与实现路径。这不仅有助于实现教学资源的高效利用与价值增值，推动研究生教育教学的持续提升，同时在推动拔尖创新型人才培养、助力教育数字化转型等方面发挥关键作用，为推进教育高质量发展注入强劲动力。

随着人工智能等技术与教育的深度融合，优质教学资源共享势必迎来全新的发展样态。未来推进研究生优质教学资源共享需要从三方面加强：一是技术驱动的资源生成与优化，即充分利用以生成式人工智能为代表的新兴技术，促进优质教学资源的建设与共享；二是共享壁垒的破除，即通过技术手段降低资源的获取门槛，提高资源的可及性与易用性，确保多方利益主体平等参与资源建设与共享；三是共享文化的培育与推广，即在教育领域内培养一种开放、协同的资源共享文化与愿景，确保利益相关者都能认同并积极参与到资源共享中。

参考文献

方海光,洪心.(2023).人机协同教育共同体生态环境研究——以北京市数字教育发展新格局为例.中小学信息技术教育(07),5-7.
教育部,国家发展改革委,财政部.(2020)关于加快新时代研究生教育改革发展的意见.中华人民共和国国务院公报(34),72-76
李昊灿,李妍(2023).新时代行业特色型高校研究生教育高质量发展——时代意蕴、现实困境与实践路向.研究生教育研究(06):21-28+89.
梁传杰,麦立强,范涛(2018).“双一流”建设背景下团队式研究生培养模式的探索与实践.学位与研究生教育(05):11-18.
聂竹明,刘钊颖.(2017).优质教育资源共享制度变革:从科层制到社群化.中国教育学刊(05),56-61.
童兆平,来钇汝,张立新等.(2021).互联网支持下城乡教育共同体的构建与运行模式——浙江省“互联网+义务教育”的实践探索.中国电化教育(08),78-84.
王卫华,田迪,邢旭光(2022).高等院校教育资源共享策略的思考.中国成人教育(20):20-23.
徐冰鸥,尹婷.(2013).教育学优质教学资源的分布考察及特征分析.中国大学教学(11),88-91.
许丽英.(2010).论教育补偿机制的构建——义务教育资源均衡配置的实现路径探讨.教育发展研究(19),31-35.
赵冬冬,曾杰,朱益明.(2023).智能时代教育共同体建设重构教育生态研究.中国电化教育(09),41-49.
中共中央政治局第五次集体学习.(2023-05-30).加快建设教育强国为中华民族伟大复兴提供有力支撑.人民日报,001.
周赛君.(2023).师范教育共同体建设的价值内涵及实施路径. 现代教育管理(04),77-86.
UNESCO. (2022). Reimagining our futures together: A new social contract for education. UN.

数字化何以赋能乡村教育高质量发展？——基于国内外经验与行动的综述

How can Digitalisation Empower the Development of Rural Education with High Quality?

--An Overview Based on Domestic and International Experience and Action

李永杰*， 张靖
西北师范大学 教育技术学院
*1446736402@qq.com

【摘要】 在教育数字化转型的大背景下，数字化逐渐成为赋能乡村教育高质量发展的重要手段之一。如何利用数字化赋能乡村教育高质量发展，是各国共同面临的重要课题。经过长期的探索与实践，国内外已取得了一些宝贵的经验。基于此，本研究首先探讨了教育数字化视域下我国乡村教育高质量发展的现实困境，接着对国内外数字化赋能乡村教育发展的已有研究与行动进行总结和借鉴，从认知重塑、数字底座、资源共享、素养提升等层面提出了相应的数字化赋能乡村教育发展的路径，以促进我国乡村教育的高质量发展。
【关键词】 数字化；乡村教育高质量发展；现实困境；国内外经验与行动；路径选择

Abstract: *Under the background of the digital transformation of education, digitization has gradually become one of the important means to enable the high-quality development of rural education. How to use digitalization to enable high-quality development of rural education is an important issue facing all countries. Through long-term exploration and practice, some valuable experience has been obtained at home and abroad. Based on this, this study first discusses the realistic dilemma of the high-quality development of rural education in China from the perspective of digital education, then summarizes and draws lessons from the existing researches and actions on the development of digital empowered rural education at home and abroad, and puts forward the corresponding development path of digital empowered rural education from the aspects of cognitive remodeling, digital base, resource sharing, and literacy improvement, so as to promote the high-quality development of rural education in China.*
Keywords: Digitalisation; high-quality development of rural education; realistic dilemma; domestic and international experience and action; path choice

1. 引言

推动教育的高质量发展，实现乡村教育的高质量发展是重要内容。党的十九大报告中首次提出了“高质量”的概念。2020 年 5 月，中共中央、国务院印发《关于新时代推进西部大开发形成新格局的指导意见》，正式提出“教育高质量发展”「周晔，2023」。党的二十大报告又进一步强调建设高质量教育体系，提出推进教育数字化建设。在建设高质量教育体系的历史进程中，乡村教育发展是重中之重，推进教育数字化是促进乡村教育高质量发展的重要手段，也是实现乡村振兴的必然选择。

2. 教育数字化视域下我国乡村教育高质量发展的现实困境

2.1 数字化赋能乡村教育的重视程度不足，形成数字化认识洼地

2.2 农村学校数字化教育基础设施利用率低，形成数字化教学困境

2.3 农村学校优质教育资源供给渗透性差，形成数字资源堰塞湖

2.4 农村学校师生数字素养水平普遍偏低，形成数字素养鸿沟

3. 数字化赋能乡村教育高质量发展的国内外经验与行动

3.1 数字化赋能乡村教育高质量发展的国外经验与行动

3.1.1. 国家级战略行动引领，为数字化赋能乡村教育高质量发展注入动力

3.1.2. 数字化应用环境建设，为数字化赋能乡村教育高质量发展夯实基础

3.1.3. 师生的数字素养提升，为数字化赋能乡村教育高质量发展增强动能

3.2 数字化赋能乡村教育高质量发展的国内经验与行动

3.2.1. 以数字化缩小城乡教师队伍差距，助推乡村教育内生性高质量发展

3.2.2. 以数字化赋能城乡教育资源共享，助推乡村教育内涵式高质量发展

3.3.3. 以数字化助力城乡教育共同体建设，助推乡村教育共进式高质量发展

3.3.4. 以数字化推动本土化教育创新融合，助推乡村教育在地化高质量发展

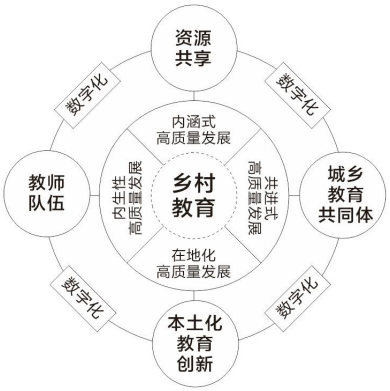


图 1 国内数字化赋能乡村教育高质量发展的框架体系

4. 数字化赋能我国乡村教育高质量发展的路径选择

针对教育数字化视域下我国乡村教育高质量发展所面临的现实困境，吸取和借鉴国内外部分乡村地区数字化赋能乡村教育发展的有益经验，可在一定程度上为我国乡村教育高质量发展带来新的路径选择。本研究从认知重塑、数字底座、资源共享、素养提升等层面提出了相应的数字化赋能乡村教育发展的路径，以促进我国乡村教育的高质量发展。

5. 结束语

随着信息技术的快速发展，数字化已经深入到各个领域，其在教育领域的应用也日益显著，数字化为乡村教育的高质量发展带来了前所未有的机遇。通过对数字化赋能乡村教育高质量发展的国内外研究与行动的综述，我们可以认识到数字化为乡村教育带来了巨大的潜力。同时，我们深知在数字化赋能乡村教育的道路上仍有许多未知和挑战需要去探索。未来，我们需要继续探索数字化赋能乡村教育高质量发展的落脚点，不断完善和更新研究成果，以期为推动我国乡村教育的持续、健康、高质量发展贡献更多的智慧和力量。

参考文献

周晔.乡村学校教育高质量发展的内涵、价值与路径[J].山西大学学报(哲学社会科学版),2023,46(02):100-109.

教师应用信息技术对数学课堂学习氛围的影响

——基于不同信息技术参与度的序贯多元线性回归分析

Multivariate Regression of IT Use on Math Learning Atmosphere by Teacher Engagement Levels

宣申奥[1*]，虞增艺[2]，景玉慧[3]，徐刘杰[4]
[1234]浙江工业大学 教育科学与技术学院
*850205877@qq.com

【摘要】本研究基于 PISA2022 测试数据，旨在探究教师信息技术应用对数学学习氛围的影响。结果表明，教师信息技术水平越高，数学课堂学习氛围越好。同时，教师在数学课堂的信息技术参与度越高，教学氛围也更好。此外，不论使用何种信息技术教学手段，使用频次对数学学习氛围也有一定影响。研究代码与数据已公开在 https://github.com/yuzengyi/PISA_xuan。
【关键词】PISA2022；信息技术；数学学习氛围

Abstract: *This research, drawing on PISA 2022 data, examines the effects of teachers' ICT application on the mathematics classroom atmosphere. Findings suggest that higher ICT proficiency and greater engagement correlate positively with a more favorable learning environment. Moreover, the frequency of ICT usage independently impacts the atmosphere, regardless of the teaching approach. The code is available at https://github.com/yuzengyi/PISA_xuan.*

Keywords: PISA 2022, Information and Communication Technology, Disciplinary climate in mathematics

1. 前言

《深化新时代教育评价改革总体方案》指出教师在数学教学中合理运用信息技术至关重要。本文将基于 PISA2022 数据集，采用多元线性回归（黄声华等，2023）与 K-means 聚类模型（任楚岚等，2020），从使用数字资源及信息技术赋能的角度展开分析。

2. 研究设计

2.1.研究对象

本研究实验数据来源于 PISA2022 数据集。由于教师 ICT 使用情况主要由教师问卷测量，且台湾地区相关数据较少，因此，只有中国香港（HKG）及澳门（MAC）地区的调查和测试数据能够为本研究提供数据支持。

2.2.研究模型

2.2.1 部分字段解释

表 1 部分字段解释

字段分类	字段名称	字段解释
使用情况与能力	ICTApp	教师使用特定 ICT 应用的情况
	ICTUse	教师使用 ICT 的情况
	ICTCmp	重视 ICT 能力
因变量	y	数学的课堂氛围

3. 研究结果与分析

3.1.ICT 三维度变量对教师使用数字资源进行教学影响分析

本文采用 K-means 聚类以区分不同教师在 ICT 使用上的不同模式。结果如图 1。

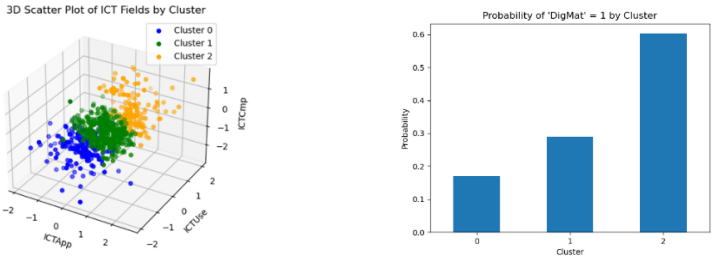


图 1 3D 聚类分析结果及三类组教师在使用 ICT 进行数学教学问题填答“有”的概率

3.1.1 聚类结果分析

根据聚类，可分如下三组：低度 ICT 参与组、中度 ICT 参与组、广泛 ICT 参与组。

表 2 三类组教师在使用数字资源 ICT 进行数学教学问题项上填答“有”的概率

题项	低度 ICT 参与组 24.2%（165 人）	中度 ICT 参与组 52.7%（359 人）	广泛 ICT 参与组 22.9%（156 人）
使用数字资源 ICT 进行数学教学	0.169	0.289	0.602

3.2.不同 ICT 参与组间对于数学学习氛围的类比多元线性回归分析

教师 ICT 的有效应用显著提升了数学学习氛围，尤其在中度和广泛参与组中，对算法理解和编程兴趣的提高以及学生学习动机的增强尤为明显。

表 3 不同参与度教师组信息技术的使用对数学课堂学习氛围的序贯影响(部分)

	低度 ICT 参与组		中度 ICT 参与组		广泛 ICT 参与组	
	模型 1 标准系数	模型 2 标准系数	模型 1 标准系数	模型 2 标准系数	模型 1 标准系数	模型 2 标准系数
常量	0.234	0.177	0.558	0.543	0.719	0.765
ICT 应用情况	0.186*	0.193*	0.017*	0.006*	0.137*	0.142*
教师使用 ICT	0.191*	0.197**	0.217*	0.241**	0.335*	0.329**
重视 ICT 能力	-0.176	-0.171	0.151	0.134	0.146*	0.124**
…	…	…	…	…	…	…
编程或算法活动		-0.019		0.015*		0.011*
数学教学		0.258		0.062**		0.135**
R-Squared	0.082	0.156	0.089	0.228	0.092	0.326
N	165		359		156	

注：*P<0.05 **P<0.01

4. 研究结论与建议

在港澳地区，教师使用 ICT 资源正面影响了数学学习氛围，但大多数教师仍处于中度 ICT 使用水平，因此，需提升教师 ICT 使用意识和在课堂中推广信息技术的应用。

参考文献（部分）

黄声华，尹弘飚，& 靳玉乐. (2023). 家长教育卷入类型与中学生学科素养:基于 pisa 2018 中国香港及澳门数据的潜在类别分析. 华东师范大学学报:教育科学版, 41(1), 50-59.

课堂派支持下基于深度学习的单元作业设计模型构建

Model Construction of Unit Assignment Design Based on Deep Learning with the Support of Classroom Pie

孙晓涵，魏雪峰*，朱梦雨，朱晓颖，张雪
鲁东大学 初等教育学院
*xuefengwei99@163.com

【摘要】 当前信息科技课堂作业与核心素养培养结合不紧密，深度学习理论与单元作业为解决现存问题提供了新思路。此外，数字化学习与创新作为学科核心素养之一，是其他核心素养的依托。课堂派在提供数字化学习环境的同时，可以构建深度学习模式。本研究旨在于课堂派的支持下通过单元作业设计，解决传统信息科技作业的问题，促进学生深度学习的发生。

【关键词】 深度学习；课堂派；数字化学习与创新；作业设计

Abstract: *At present, the combination of it class work and core literacy training is not close, and the theory of in-depth learning and unit work provide a new way to solve the existing problems. In addition, digital learning and innovation, as one of the core literacy of the discipline, is the support of other core literacy. While providing a digital learning environment, the classroom can construct a deep learning model. The purpose of this study is to solve the problem of traditional it homework and promote students' deep learning through unit assignment design supported by classroom instruction.*

Keywords: Deep learning, Classroom pie, Digital learning and innovation, Assignment design

基金项目：本文系山东省人文社会科学课题（智库重点项目）“人工智能支持下的小学个性化作业设计与应用研究”（编号：2023-zkzd-014）的阶段性研究成果。

1.研究背景

在实际信息科技课堂中，作业的设计存在知识考查碎片化的现象，缺乏对学生核心素养的考量。根据已有研究发现，单元作业和深度学习理论恰好契合核心素养培养的要求。数字化学习与创新作为学科核心素养之一（教育部，2022），贯穿于整个信息科技学科核心素养培养体系中，是其他核心素养的依托和直接能力体现（陈明选和张宁，2019）。课堂派在提供数字化学习环境的同时,可以增强学生学习能力和欲望,构建深度学习模式(孙刚成和杨眉,2017)。基于此，本研究旨在于课堂派的支持下，构建面向数字化学习与创新的单元作业模型。

2.面向数字化学习与创新的单元作业模型设计

深度学习视域下单元作业设计模型以真实情境的问题为载体。横向由教师行为、学生行为、思维过程（包括设计维度）和评价总结四个维度组成，纵向以单元作业设计、单元作业过程、单元作业评价为主线，如图 1 所示。

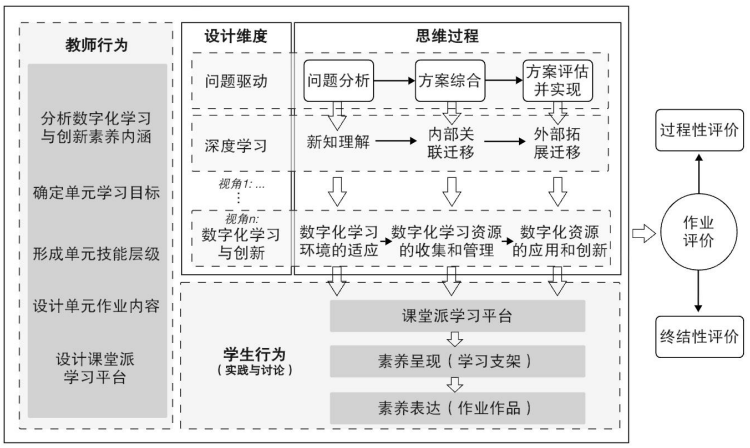


图 1 深度学习视域下面向数字化学习与创新的单元作业设计模型

课堂派主要用到了“课程教学”和“学情分析”两个版块。首先，“课程教学”在功能上侧重于设计完成作业的过程。作业发布后，学生完成后可提交的文件支持多种格式，满足多种教学内容的需求。教师在批阅页面进行批改，支持在线标注打分，将批注生成 PDF 文件并下载，留存与存档。教师上传作业所需资料后，学生可进行获取、加工与管理并进一步创新应用。师生可在该版块的“话题”区围绕作业中遇到的问题进行讨论，补充学习。“互动答题”区，教师可发布互动答题作业，根据学生作答情况去检验学生学习情况。其次，“学情分析”中展示所有学生围绕课堂作业互动参与的整体数据情况，包括作业完成、话题参与和互动答题等，可下载导出数据。课堂派支持下单元作业实施步骤见图 2。

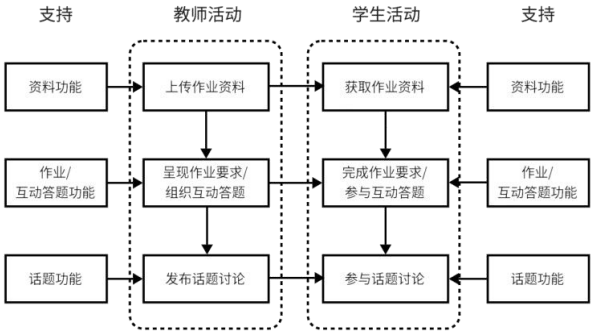


图 2 课堂派支持下单元作业实施步骤

3.结语

本研究在分析指向数字化学习与创新的单元作业设计、深度学习理论及“课堂派”应用情况的基础上，构建了基于深度学习的单元作业设计模型构建。但本研究还存在一些不足，如缺乏应用性研究。在今后的研究实践中，将在课堂派支持下进行实证研究，并根据实验结果进一步完善该模型。

参考文献

中华人民共和国教育部（2022）。义务教育信息科技课程标准。北京：北京师范大学出版社。

孙刚成和杨眉（2017）。“A+课堂派+BYOD”支持下的高校翻转课堂——以软件类课程教学为例。电化教育研究，2017，38（04）：103-107+114。

陈明选和张宁(2019)。基于高中信息技术学科核心素养的学习活动设计。中国电化教育,(01):87-93。

基于 iFIAS 的课堂师生互动行为研究——以初中信息科技优质课为例

Research on Teacher-Student Interaction Behavior in the Classroom Based on iFIAS——A Case Study of a Middle School Information Technology Quality Lesson

庄凌[1]，梁颖怡[2]，徐玲[3]
[123]华南师范大学教育信息技术学院
1654743919@qq.com

【摘要】 本研究采用 iFIAS 对两节全国初中信息科技优质课堂视频进行编码分析，发现不同教学风格对课堂师生互动行为存在显著影响。结合视频观察和数据分析，建议教师平衡启发式、引导式和讲授式教学语言，合理安排封闭型和开放型问题，并通过现代技术手段增强师生互动。此外，教师应鼓励学生主动提问，培养批判性思维和创新意识，从而构建更加优质、互动、和谐的课堂环境。

【关键词】 iFIAS；师生互动行为；初中信息科技优质课

Abstract: *This study conducted a coding analysis of two national junior high school information technology quality classroom videos using iFIAS, and found that different teaching styles have a significant impact on teacher-student interaction in the classroom. Based on video observation and data analysis, it is suggested that teachers should balance heuristic, guiding, and didactic teaching language, reasonably arrange closed and open questions, and enhance teacher-student interaction through modern technological means. Additionally, teachers should encourage students to actively ask questions, cultivate critical thinking and innovation awareness, in order to build a more high-quality, interactive, and harmonious classroom environment.*

Keywords: *iFIAS; teacher-student interaction behaviors; middle school IT quality classes*

1. 引言

《中国教育现代化 2035》强调信息技术与教育融合，促进教育现代化进程。在此背景下，课堂师生互动行为作为影响教学质量的关键因素，引起了广泛关注。然而，信息科技课程中师生互动行存在形式单一、质量不高等问题。因此，本研究以初中信息科技优质课为对象，深入分析师生互动，以期为改进初中信息科技课程教学提供有益的启示。

2. 研究设计

2.1. 研究对象

本研究选取全国初中信息科技优质课堂视频《体验物联网》和《初识物联网》作为研究对象。利用弗兰德斯互动分析系统（iFIAS）进行比较分析，全面了解不同教学目标下师生互动的模式和效果，并探讨教学方法对师生互动的影响，有助于准确把握初中信息科技课堂中师生互动的现状和问题，并为改进教学提供指导。

2.2. 研究方法与数据采集

本研究采用混合方法，结合定性与定量分析，探究初中信息科技优质课的师生互动行为。首先，利用 iFIAS 系统每 3 秒记录并编课例 1 和课例 2 的师生互动，分别生成 881 和 861 个

编码。接着，基于编码数据，运用统计学方法分析互动行为的频率和持续时间。

3. 结果分析

3.1. 课堂互动结构特征分析

两课例中，教师讲解少，学生活动多。课例 1 学生参与度高达 65.04%，说明教师善用小组讨论等策略，生生互动多；课例 2 学生参与度低，教师讲解为主。课例 1 有益沉寂比例高，利于学生思考，而课例 2 节奏快，技术使用多，有助于提升学习体验。两课例展现不同互动结构特征，各有优劣。

3.2. 教学风格特征分析

两节优质课风格迥异。课例 1 中，教师间接与直接语言比为 0.89%，多用启发式语言引导学生思考，而课例 2 该比例仅 0.33%，教师更倾向直接指导，可能限制学生探索。同时，课例 1 中教师积极影响比例高于课例 2，表明课例 2 教师需减少消极影响，促进学生学习发展。

3.3. 师生问答特征分析

教师提问中开放型与封闭型问题的比值都较高，体现其重视批判性思维与创新意识培养，但需注意平衡，以免知识传授不足。此外，两节课都是训练型提问为主导，但课例 1 倾向于一对一提问，课例 2 注重集体讨论。创新型提问在两课中均有体现，课例 1 学生提问意愿更高，课例 2 学生回答更积极并具创新性。

4. 建议

4.1. 优化与平衡教学风格

教师应根据教学内容和学生实际，灵活采用运用启发式、引导式和讲授式教学语言，平衡间接与直接语言的比例。同时，教师应进一步探索个性化教学风格，以满足学生多样需求。通过持续的教学实践和反思，找到最佳教学风格，促进师生互动，构建优质课堂。

4.2. 深化与实践师生问答策略

教师在设计问题时应平衡封闭型与开放型问题的比例。对于封闭型问题应引导学生深入思考和解释，而开放型问题则应更关注学生观点和创意，并及时反馈引导。既要传授知识，也要培养思维能力。同时，鼓励学生集体回答、讨论和主动提问，培养批判性思维和创新意识，以提升问答质量和课堂氛围。

4.3. 平衡与发展训练型与创新型提问

教师在提问时应平衡训练型与创新型问题。训练型提问要有针对性和层次性，满足学习需求；创新型提问要逐步增加，激发创新思维和批判性思考。教师根据学生回答调整提问策略，培养自主思考和学习能力，激发学生思维活力和创造力，促进学生全面发展。

5. 结语

本研究聚焦两节初中信息科技优质课，运用 iFIAS 编码工具分析师生互动，揭示课堂互动现状和问题，并提出三点建议：优化与平衡教学风格，深化与实践师生问答策略，平衡与发展训练型与创新型提问。未来研究将深入探索互动机制，提供更全面指导。

参考文献：

陆灵明和甘龙华（2023）。基于 iFIAS 的初中生物优质课课堂互动行为研究。**中国教育信息化，29(03)**，91-98。

数字化转型背景下基于社会情绪能力培养的小学英语单元整体教学实践研究

A Practical Research on Unit-Integral Teaching of Primary School English Based on the Cultivation of Social Emotional Competence under the Background of Digital Transformation

丘秀文
中山市三乡镇平岚小学
435445567@qq.com

【摘要】 本研究教学理论、教学设计框架与方法进行论述，从学科核心素养和社会情绪能力等方面 证其成效。研究效果表明，数智赋能，基于社会情绪能力培养的小学英语单元整体教学，能够提升学生综合能力，有效落实小学英语核心素养，培养社会情绪能力。

【关键词】 数字化；小学英语；单元整体教学；社会情绪能力

Abstract: *This study discusses pedagogical theories, frameworks, and methods, validating effectiveness in fostering core English competencies and socio-emotional abilities. The results show that data-driven, socio-emotionally focused English instruction enhances overall student capabilities, effectively implementing core English competencies in primary education.*

Keywords: Digitalization, Primary School English, Unit Integral Teaching, Social Emotional Competence

1. 引言

在“双新”“双减”和数字化转型的背景下，聚焦学生学科核心素养，以数字化、网络化、智能化工具赋能课堂，实现数智育人，是培养创新型人才的有效路径。本文打破学科的界限，在英语单元整体教学中挖掘培养学生社会情绪能力的元素，在英语学科中渗透社会情绪能力的培养，促进儿童青少年健康成长，符合现代教育的新趋势！

2. 相关研究述评

《义务教育英语课程标准（2022 年版）》明确指出：推进信息技术与英语教学的深度融合，重视教育信息化背景下英语课程教与学方式的变革。社会情绪能力学习在国际上是比较系统的情商课程之一，教育部和中国科协把儿童社会情绪能力列入科学课程内容标准， 钰也提出要注重培养孩子的社会情绪能力，但国内针对小学阶段的研究相对较少。

3. 基于社会情绪能力培养的小学英语单元整体教学模式构建

3.1. 基于社会情绪能力培养的小学英语单元整体教学设计方法

教师通过整合优质教学资源，优化学生的学习环境，通过学情数据分析，精准定位教学目标，以单元主 为背景，以任务单、问 链为引导，在基于单元主 意义的具体有整体性、关联性和递进性的单元教学活动设计有效的教学活动，实现个性化分层教学，学生通过学习理解、应用实践、迁移创新；在体 中学习、在实践中应用、在迁移中创新的学习理念，提升自主学习的空间，提升英语综合运用能力，落实英语学科素养。

3.2. 基于社会情绪能力培养的小学英语单元整体教学模式

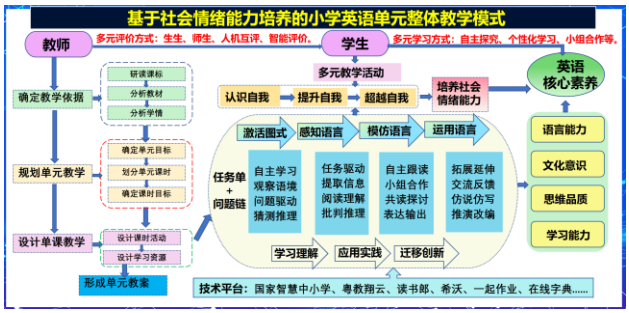


图 1 基于社会情绪能力培养的小学英语单元整体教学模式

围绕“激活图式—感知语言—模仿语言—运用语言”四个阶段，聚焦英语核心素养，教师根据学情反　“确定教学依据、规划单元教学、设计单课教学”；学生参与多元教学活动，在“认识自我—提升自我—超越自我”的过程中，通过“复述、表演、讨论、仿说、仿写、推演、改编”实现。开始“自主学习、感知语言”、“自主检测、分析学情”生成的相关数据为英语单元教学质量　测分析和英语学习材料个性化推送提供支持。“确定主　、设问；　动”“提取信息、批判理解”“小组探究、共读探讨”“表达输出、交流反　”是英语语言习得的过程，“提供支架、　测推理”完成　阶思维能力的培养，构成了基于社会情绪能力培养为核心的数智育人的英语单元整体教学模式，实现英语核心素养课程育人的价值理念。

4. 研究效果

通过本　目研究，学生的社会情绪能力和英语学习水平皆有提　，老师的专业水平也有较大的提升，如图 2 和图 3 所示。

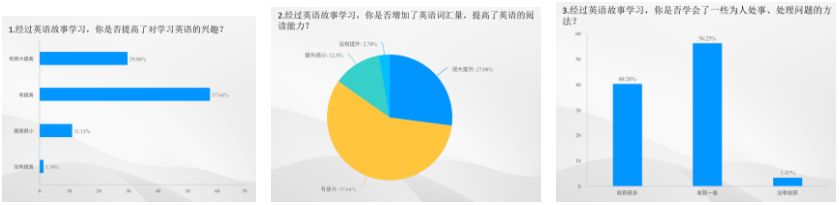


图 2 促进学生成长五年级实　班调查数据

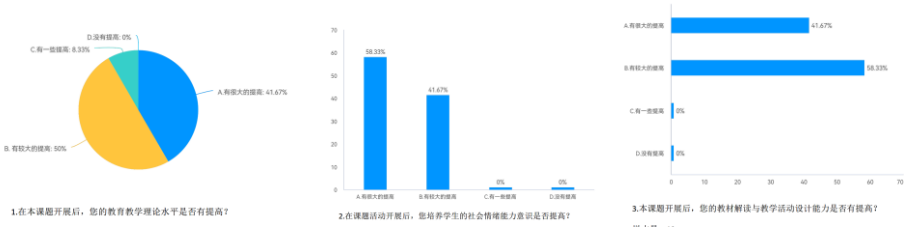


图 3 促进教师发展相关情况调查数据

5. 结语

基于社会情绪能力的小学英语单元整体教学实践研究的效果表明，数智赋能的单元整体教学，能够提升学生的听说读写等综合应用能力，有效落实小学英语核心素养，培养学生的社会情绪能力。

参考文献：

教育部（2022）。义务教育英语课程标　[M]。北京师范大学出版社。
李润文和　钰（2014）。早期教育要着重培养孩子的“社会情绪能力”[J]。基础教育论坛，第 26 期 P40-41。

GCCCE 2024

德彰天下

学贯中西

9 773005 321005

全球華人計算機教育應用學會