西南大學
SOUTHWEST UNIVERSITY

GCCCE2024

第28屆全球華人計算機教育應用大會

THE 28TH GLOBAL CHINESE CONFERENCE ON COMPUTERS IN EDUCATION

中國 重慶 | 西南大學(北碚校區)

CHONGQING,CHINA
SOUTHWEST UNIVERSITY(BEIBEI CAMPUS)

2024年6月1日 — 6月5日

JUNE 01,2024 — JUNE 05,2024



工作坊論文集

WORKSHOP PROCEEDINGS



出版者：全球華人計算機教育應用學會
書名：第 28 屆全球華人計算機教育應用大會工作坊論文集
作者：游師柔、洪榮昭、江波、張琪、侯惠澤、江豐光、王舒民、夏綠荷、孫之元、馬志強、葉新東、趙建華主编
出版年月：2024年 6月
版次：初版
ISBN：9786269747856



第 28 届全球华人计算机教育应用大会

The 28th Global Chinese Conference on Computers in Education

GCCCE 2024 工作坊论文集

GCCCE 2024 Workshop Proceedings

主编 Editors

游师柔 阳明交通大学

Shih-Jou YU, National Yang Ming Chiao Tung University, Taiwan

朱蕙君 东吴大学

Hui-Chun CHU, Soochow University, Taiwan

孙丹儿 香港教育大学

Daner SUN, The Education University of Hong Kong, Hong Kong

卢 宇 北京师范大学

Yu LU, Beijing Normal University

孙之元 杨明交通大学

Jerry Chih-Yuan SUN, Yang Ming Chiao Tung University, Taiwan

江绍祥 香港教育大学

Siu Cheung KONG, The Education University of Hong Kong, Hong Kong

刘革平 西南大学

Ge Ping LIU, Southwest University

责任编辑 Executive Editor

曾詩倩，東吳大學

Shih-Chien TSENG, Soochow University, Taiwan

目录 CONTENTS

一、 序言 Message from Organizer xiii

二、 大会组织 Conference Organization xiv

工作坊 Workshop

W01 迎接「元宇宙」的世代，如何融入新科技于教学工作坊

1 親子關係和小學生學習毅力影響學生在家長支持在疫情後手機成癮之研究

何美慧

8 探究教育科技應用於國小學童執行功能與創造力遊戲表現之差異

蔡仲倫

16 Quizizz 融入國語領域教學對國小學生的自我效能、遊戲興趣與形成焦慮感之研究

李逸銘

24 以 Picasso 互動數位遊戲學習探討性別對意境想像力、繪圖認知負荷、情境命名認知負荷與趣味性價值之差異

宋紀萱

32 Research on Immersive Virtual Experimental Environment based on Visual Emotional Design

Liqiao Nong, Lanlan Zhang and Geping Liu

40 LearningverseVR: Intergrating Generative AI and VR Technology in an Immersive Game-based Learning Platform

Yanjie Song, Kaiyi Wu and Jiaoyang Ding

46 Measuring EFL Learners' Perceptions of Using ChatGPT in Translation: An Exploratory Study

Huiqi Fan, Bowen Jing, Yanqing Yi, Lin Luan

50 我国职业教育教师信息素养发展可视化分析

刘妍萱、叶建宏

56 英国数字化推动教师专业发展的变革契机、变革框架与借鉴启思

梁冰冰、叶建宏

W02：教育大语言模型工作坊

62 Fine-Tuning Large Language Models for Educational Support: Leveraging Gagné's Nine Events of Instruction for Lesson Planning

Linzhao Jia, Xiaozhe Yang, Han Sun, Changyong Qi and Yuang Wei

70 Exploring Learner's Knowledge Inquiry Level and Experience on Utilizing ChatGPT during Programming Learning

Dan Sun

77 AI Agent for Education: von Neumann Multi-Agent System Framework

Yuan-Hao Jiang, Ruijia Li, Yizhou Zhou, Changyong Qi, Hanglei Hu, Yuang Wei， Bo Jiang, and Yonghe Wu

85 比較自然語言模型在分析數位學習文獻的效能

陳一嘉，陳攸華

92 基于大语言模型微调的数学习题认知属性自动标注

胡航磊，姜元昊，李瑞佳

100 人机协同视域下 ChatGPT 辅助师范生开展教学设计的创新探索

许洁，孙丹，李艳

108 A Study on Educational Data Analysis and Personalized Feedback Report Generation Based on Tags and ChatGPT

Yizhou Zhou, Mengqiao Zhang, Yuan-Hao Jiang, Naijie Liu, and Bo Jiang

116 IntelliChain: An Integrated Framework for Enhanced Socratic Method Dialogue with LLMs and Knowledge Graphs

Changyong Qi, Linzhao Jia, Yuang Wei, Yuan-Hao Jiang and Xiaoqing Gu

122 中国古典诗词人工智能生成图像创作与评价：提示词生成与优化

陈弘正，许怀之，李威，萧巨昇

130 A Large Language Model-Enabled Solution for the Automatic Generation of Situated Multiple-Choice Math Questions

Ruijia Li,Yuan-Hao Jiang,Yiting Wang,Hanglei Hu and Bo Jiang

137 教育大语言模型应用的 SWOT 分析：机遇、挑战与对策

徐琦, 魏艳涛, 刘清堂, 肖光润, 师亚飞

145 基于大语言模型的素养知识图谱构建——以信息科技课程为例

叶宇雷

153 智能代理驱动的课程智能助教构建研究——以材料力学为例

胡震，补熠暄，李艳洁，段庆龙

W03：学习投入与学习行为建模工作坊

161 C-STEAM 教育模式下文化统整课程设计研究

詹梓灵，王炜，崔萌

169 视频直播教学中互动策略对学习投入的影响研究

杨琳，胡航

177 隐私保护的学习者学习投入度识别

陈晓雨，魏艳涛，马玉建，王涛

181 多模态教学视频分析技术驱动的职前教师集体反思中的认知与情感分析

叶如梦，王婷，龙陶陶

189 工程翻转学习情境下基于过程挖掘的学生自我调节学习行为分析

陈香均，龙陶陶，朱晓萌，伍冬

197 基于课堂生态学视角的城中村小学科学教育信息化的案例研究

徐莹，赵雅，龙陶陶

203 跨学科课例研究下教师合作对话模式的社会认知网络特征分析

赵雅，徐莹，龙陶陶

211 六帽思维技术对职前教师集体反思的作用：协同调节模式的滞后序列分析

梅傲雪，龙陶陶

217 基于在线教学视频的弹幕现象分析

徐佳慧，刘勇，刘妙韵

225 损益权衡后的自主调节：行为经济学视角下教师网络研修学习投入决策逻辑研究

王玥、梁云真、沈子琦、郭琳

233 基于系统性文献综述的 XR 赋能空间能力培养实证研究分析

刘西庆、张双秀、张慕华

241 沉浸式虚拟现实赋能社会和情感技能学习的系统性文献综述研究

王博雯、顾双玲、张慕华

249 基于信息化平台的全过程评价——以《明朝的对外关系》一课为例

王誉颖、吴欣熙、郭辉群、李静茹、崔萌

253 课堂场域下学生社交投入评测的研究现状及展望

余游、魏艳涛、肖光润、徐琦、芪佩舒

263 数智赋能师范生教学设计能力提升的教学评一体化研究——以《小学信息技术教育》课程为例

吴欣熙、陈思婷、钟卓华、崔萌

270 混合式教学模式下《学习分析技术与方法》课程设计与评价研究——以《数据驱动教学》为例

蒲圻彬、于慧敏、张琪

278 多模态学习分析：理论基础、融合策略与多维数据案例

马宇慧、李璇、蒲圻彬、张琪

W04：数字游戏式学习(Game-based Learning)与游戏化(Gamification)教学策略运用于关键力教育工作坊

282 結合動態情節互動鷹架之行政倫理數位敘事教育遊戲之設計與評估

劉書瑋，倪紳煬，侯惠澤

290 高中化學教育桌游開發設計與教學實踐研究：以綠色工廠為例

杨景淇，胡玲燕，江丰光

298 結合三維擬真互動情境式腳本鷹架之職場壓力調適能力培訓教育遊戲

簡志忠，吳培菁，侯惠澤

304 EMI 數位課程遊戲融入議題探究與 SDGs 教學策略運用於關鍵力教育

許雯晴，陳聖智，林永騰，蔣旭政

312 如何利用實境解謎遊戲激發元認知以提升價值教育

李浩然

320 遊戲式學習應用於培育 21 世紀關鍵能力趨勢 – 應用文獻計量分析之初探

王舒民

324 開發《食物忍者》互動式遊戲繪本以提升國小學生之創造力思維

藍文君，莫德駿，湯梓辰

326 探討後疫情幼兒情緒學習在「數位繪本」之應用

湯梓辰，趙彥俞

328 《ANesT 》最強蟻巢 - AR 教育桌遊設計與開發

徐美恩，陳怡瑄，周文己，鄭夢慈

330 核心素養視域下皮影戲融入初中語文課堂實施初探

郭泽宇，林昕悦，唐伟，江丰光

332 開發以學習人工智慧路徑演算法之 TAMGPA 學習教育模式初探

朱志明，李品諺

W05：电脑支援个人化与合作学习

334 針對視障生的實體程式積木感測板之開發

孔崇旭，陳俊憲，郭承諺，湯志炫

338 第二外語對話練習平台系統開發

洪霈羽，施彥安，張立杰

342 A Research Proposal for a Dynamic Assessment-based Robot-Assisted Language Teaching Approach on Elementary School Students' English Learning Performance and Learning Experience

Shu-Yun Chien

346 ADDIE 模型指导下翻转课堂与 VR 技术融合的教学设计探究

梁楚渝

350 以體驗式學習的虛擬實境美術館增強學生的藝術鑑賞能力

邱敏棋，夏綠荷，黃國禎

358 混成式教學於大學體育課對動機信念和合作能力影響之探究

陳偉瑀，陶淑瑗，張家榮

365 提問引導同儕回饋策略對新詩創作線上學習社群互動參與的影響

梁心怡

373 Using Chatbots to Assist English Learners in Reading Strategies

Chiou-Hui Chou and Si-Yu Liu

377 Students' Capacities to Provide Feedback in L2 Writing After Receiving Teacher and Peer Feedback

Peiling Ren, Shuya Wang, Jiaojiao Ji and Yong Wu

381 個人化音樂指揮機器人學習系統之開發

黃國豪，劉晉均，陳宥瑋，黃暄涵

389 先備知識對遊戲化程式設計翻轉教室之影響

黃國豪

397 網路反思提示下小組友誼對高中生科學探究歷程的影響

邱瓊慧

401 Revitalizing Mathematical Proficiency through ScaffoldiaMyMaths: Empowered by Adaptive Cognitive Scaffoldings

Daner Sun, Zhizi Zheng, Lan Yang, Kee-lee Chou, Chee-Kit Looi, and Zhixuan Song

W06：创新互动回馈科技提升学习动机

403 A Game design and development for Enhancing Self-Efficacy

Yu Bai, Wenhao Wang, Fuzheng Zhao, Chengjiu Yin

405 Embedding Student Teachers' Feedback Literacy Development in Online Peer Assessment Activities

Ying Zhan, Daner Sun, Jan Christian C Gube, Ngok Cheng Chan, Kam Wing Chan, Zhi Hong Wan and Junjun Chen

409 基於情境感知與機器學習引導機制之遊戲式學習模式對學習者學習成效之影響

陳君銘，葉鎮源

413 How to Design a Competitive Serious Game to Improve University Students' Concept of RGB Mixing

Hsi-Hsun Yang

421 基於概念圖的元宇宙探究式學習模式在地理課程學習成效之影響

曾詩倩，朱蕙君

428 A Proposal on Designing a Scaffolded Flipped Classroom Environment to Improve Students' Engagement: a Design-based Research

Hantong Zhao and Qiyun Wang

432 使用 Line 社群與 ChatGPT 的翻轉學習 來翻轉高中生對英文寫作的學習成效與影響

Chin Chun Lin, Ken Zen Chen and Yu-Zhen Dai

440 結合雙層次測驗之適性輔助學習模式對於國小數學學習之問題解決與後設認知能力之影響

蘇芳玉，王俊傑

448 探討不同VR學習模式對國小學生地震防災學習成效與情境興趣之影響

Jhen-Yuan Chen and Jerry Chih-Yuan Sun

456 生成式 AI 聊天機器人結合自我調節學習理論應用於大專院校學習者之實作專題課程

徐唯芝，孫之元

461 AI 輔助之問題導向自主學習

王俊傑

465 運用互動式視覺化工具促進人工智慧概念的理解：從自我調整學習的觀點

鄭年亨

W07：人机协同视域下的协作学习设计、分析与评价

469 小学生配对编程脚本设计与应用研究：基于有效失败视角

王靖，李会，陈小汾

477 Personalized Learning Supported by Learning Analytics in Smart Education: A Systematic Literature Review

Jing Ning, Binger Zhang and Zhiqiang Ma

485 CSCL 视域下元认知支架对群体感知的促进作用研究

李红慧，曹兰英，张思

493 基于反馈理论的课堂教学资源智能推荐框架设计

刘欣鑫，钟林容，段金菊

501 基于人机协同的学生提问交互活动设计

付玉娜，赵赫璇，马国威，雅古，区文诗，孙伟华，王妍莉

508 素养本位下生成式人工智能在小学科学探究实践中的应用研究

杨宁，杨锦，包正委

516 智慧学习环境中指向创新素养提升的协作探究活动设计与实践研究——以《莓农匠心大比拼》课程为例

李梦兴，李琳，夏伟婕，刘妍

524 在线协作工具支持的六顶思考帽对学生自我效能感与群体效能感的影响

黄益品，郑晓丽，潘盈盈，涂芸芳

532 新课标背景下生成式人工智能在小学科学教学评价中的功能与角色

杨宁，李晓璠，包正委

540 面向项目化学习的表现性评价设计：基本模型与具体路径

沈诗淼，马志强

553 混合式教学中同伴互评对共享调节能力的影响作用研究

赵赫璇，马国威，付玉娜，孙伟华，区文诗，王妍莉

561 生成式 AI 支持的人机协同论证：基于两种提示设计的实证比较

崔鑫，尤欣雅，马志强

W08：中小学新一代人工智能教育:培育未来的创新者

569 四维共筑：普及创客、人工智能教育之平阳行动

谢贤晓

575 专业标准视域下的中小学人工智能教育教师培训课程设计

夏正仁

580 校地合作背景下中小学人工智能教育实施路径探索与实践

郭晓戈，涂芸芳，方建文，胡开胜，梁晨

586 信息科技核心素养框架下人工智能课程设计与策略研究—以《语音识别技术》为例

王千姿，陆盈莹

595 生成式人工智能赋能精准教学的实践路径与框架设计

廖龙燕

603 基于人工智能教育机器人的中学实验安全课程设计与开发

林晓婧，连丽雯，吴仙勇

611 基于 Citespace 的作文自动评价系统研究的可视化分析与概述

Zheng Kunhan, Huang Biyun and Xie Shiyi

618 Supporting Students' Artificial Intelligence Literacy Development by Promoting Authentic Learning Experience Using Programmable Robots

Yunsi Tina Ma and Henry Chi-Fuk So

626 OpenHydra：一个本地化部署的开源人工智能教学平台

谢作如

632 基于 OpenInnolab 平台的初中人工智能校本课程设计及其教学实践探索

陈婧，孟晓培，应利波

640 高中人工智能教育实验室建设--来自温州科技高级中学的经验

陈恩慈，林淼焱

651 借势·借域·借力—基于浦育平台在农村小学开展新一代人工智能教育的实践

陈慧娴

657 人工智能领航山区教育创新——以校本课程开发为例

张艳超，徐昊宇，赵杰

667 AIGC 在幼儿家庭教育中的应用与思考——基于微信公众号“扑满的杂记团”的案例分析

扑满，周纲，黄鲁凡

W09：人工智能赋能知识建构创新研究与实践

675 Epistemic Agency for STEM Learning Through Design-based Integrative Approach

刘家亮，Carol Chan，赵建华，孔晶

683 小学知识建构课堂中协作认识代理的模式研究

张晓洁，陈桂涓，冯诗惠，赵建华

687 How Can Children Advance Collective Knowledge? An Exploratory Study Using the CICR Approach in Primary One Science

Xueqi Feng, Jianhua Zhao, Yuqin Yang, and Carol K.K. Chan

691 核心素养背景下知识建构促进小学生科学模型建构的设计研究

姜羽，吴少文

695 基于会话代理促进学习者反思性思维的框架构建研究

缪静敏、林婉芸

703 Empowering Teachers to be Learning Designers: A Transformative Model to Teacher Professional Learning

Kun Liu, Jianhua Zhao and Nancy Law

707 促进知识建构的信息科技课堂学习空间设计研究——基于发生认识论视域

丁添，唐璇

715 数智时代基于知识建构的教师数字素养提升逻辑与实践

王贤灿

721 基于知识建构的素养课堂教学策略研究——以一年级小学语文《影子》为例

肖利猛，黄芳，彭家益，肖霞洁，冯雪琦

725 人工智能技术支持下的协同知识建构教学模式研究

杨文静，赵建华

733 College of Design Students' Intention to Use Generative AI Tools: The Role of Intrinsic and Extrinsic Motivation

Hsi-Hsun Yang

737 A Systematic Review of AI-supported Public Speaking Instruction in the Past Decade

Ping Lu and Chunping Zheng

745 基于人工智能技术的教学过程优化思考

康健

753 人工智能赋能书院制高校学生管理：内涵、应用及创新路径

于越、王令军、陈劲松

761 机器评估与人工评估视角下不同学业投入学生创造性思维特征探究

赵亚南、毛云庆，普琳洁，唐天奇，沙景荣

769 基于知识图谱可视化表征的形成性反馈设计

张灵琳，马志强

777 AI 绘画以图解文的中职诗词教学策略研究

陈志雄

785 基于知识建构原则的“人工智能+教育”实践探索——以一年级跨学科课程《可持续发展目标 11：可持续城市和社区》为例

林洪全、冯雪琦、李燕莹、胡秀、李娜、吴珍科、周辉、李慧哲

一、 序言 Message from Organizer

第 28 届全球华人计算机教育应用大会（GCCCE 2024）于 2024 年 6 月 1 日至 6 月 5 日由重庆西南大学举办。本届工作坊于 2024 年 6 月 1 日至 2 日举行，采用线下会议形式，汇聚世界各地的研究者们交流分享最新研究成果，拓展合作创新之契机。

GCCCE 2024 工作坊以计算机教育应用领域的各种新兴主题为核心，旨在加强社群内特定主题研究群的凝聚力，为与会者提供一个学术交流与讨论的平台。本届鼓励计算机教育应用相关领域的研究者以各种新兴主题、实作导向或具有深度互动的方式策画与安排各自的工作坊。为达到深度交流互动与专业讨论的目的，本届工作坊除论文发表，亦欢迎其他形式的活动，包含但不限于小型论坛、教学分享展示、互动系统展示与实作活动等。

本论文集共收录九个独具特色的工作坊主题，经由严谨的同侪审查后，最终收录 124 篇论文。 其主题分别为：
工作坊 1：迎接「元宇宙」的世代，如何融入新科技于教学工作坊
工作坊 2：教育大语言模型工作坊
工作坊 3：学习投入与学习行为建模工作坊
工作坊 4：数位游戏式学习(Game-based Learning)与游戏化(Gamification)教学策略运用于关键力教育工作坊
工作坊 5：电脑支援个人化与合作学习
工作坊 6：创新互动回馈科技提升学习动机
工作坊 7：人机协同视域下的协作学习设计、分析与评价
工作坊 8：中小学新一代人工智能教育:培育未来的创新者
工作坊 9：人工智能赋能知识建构创新研究与实践

本届工作坊筹备历程能有效进展，由衷感谢本届工作坊副主席刘子韬教授（暨南大学）的协助，诚挚感谢各工作坊的主席、副主席与团队付出的心力和给予高度支持，使本届的论文征集、收件、评审与活动进行之相关工作得以顺利推进，以及感谢所有作者的参与，为创新计算机教育应用开创崭新的一页。最后，祝福各工作坊的组织者与参会者获得丰富的学术交流与美好的互动经验，满载而归！

游师柔 阳明交通大学
工作坊主席

二、 大会组织 Conference Organization

主办单位 Organizer:

全球华人计算机教育应用学会
Global Chinese Society for Computers in Education (GCSCE)

承办单位 Host:

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

大会主席 Conference Chair:

江紹祥 香港教育大学
Siu Cheung Kong, The Education University of Hong Kong

大会顾问 Consultants:

陳德懷 中央大学
Tak-Wai Chan, National Central University, Taiwan
施如齡 中央大学
Ju-Ling Shih, National Central University, Taiwan

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

朱蕙君 東吳大学
Hui-Chun Chu, Soochow University, Taiwan

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

孫丹兒 香港教育大学
Daner Sun, The Education University of Hong Kong
盧　宇 北京师范大学
Yu Lu, Beijing Normal University
孫之元 陽明交通大学
Jerry Chih-Yuan Sun, National Yang Ming Chiao Tung University, Taiwan

组织委员会主席 Local Organizing Chair:

劉革平 西南大学
Geping Liu, Southwest University

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

謝濤 西南大学

Tao Xie, Southwest University
韩继宁 西南大学
Jining Han, Southwest University
梁樱 西南大学
Ying Liang, Southwest University

工作坊主席 Workshop Coordination Chair:

游师柔 阳明交通大学
Shih-Jou Yu, National Yang Ming Chiao Tung University, Taiwan

工作坊副主席 Workshop Coordination Co-Chair:

劉子韜 暨南大學
Zitao Liu, Jinan University

工作坊委员会 Workshop Program Committees

W01 迎接「元宇宙」的世代，如何融入新科技於教學工作坊

洪榮昭　臺灣師範大學(主席)
Jon-Chao Hong, National Taiwan Normal University, Taiwan (Chair)
葉建宏　北京師範大學(副主席)
Jian-Hong Ye, Beijing Normal University (Co-Chair)

W02 教育大语言模型工作坊

江　波　华东师范大学 (主席)
Bo Jiang, East China Normal University (Chair)
卢　宇　北京师范大学(副主席)
Yu Lu, Beijing Normal University (Co-Chair)
魏艳涛　华中师范大学(副主席)
Yantao Wei, Central China Normal University (Co-Chair)
周　杰　华东师范大学(副主席)
Jie Zhou, East China Normal University (Co-Chair)

W03 学习投入与学习行为建模工作坊

张　琪　淮北师范大学 (主席)
Qi Zhang, Huaibei Normal University (Chair)
童名文　华中师范大学(副主席)
Mingwen Tong, Central China Normal University (Co-Chair)
胡　航　西南大学(副主席)
Hang Hu, Southwest University (Co-Chair)

W04 數位遊戲式學習(Game-based Learning)與遊戲化(Gamification) 教學 策略運用於關鍵力教育工作坊

侯惠泽　台湾科技大学(主席)
Huei-Tse Hou, National Taiwan University of Science and Technology, Taiwan (Chair)
江丰光　上海交通大学(副主席)
Feng-Kuang Chiang, Shanghai Jiao Tong University (Co-Chair)
王舒民　中国文化大学(副主席)
Shu-Ming Wang, Chinese Culture University, Taiwan (Co-Chair)

W05 电脑支援个人化与合作学习

夏绿荷　勤益科技大学(主席)
Lu-Ho Hsia, National Chin-Yi University of Technology, Taiwan (Chair)
张立杰　中央大学(副主席)
Ben Chang, National Central University, Taiwan (Co-Chair)
陈攸华　中央大学(副主席)
Sherry Y. Chen, National Central University, Taiwan (Co-Chair)
黄国豪　台湾云林科技大学(副主席)
Gwo-Haur Hwang, National Yunlin University of Science and Technology, Taiwan (Co-Chair)
林秋斌　清华大学(副主席)
Chiu-Pin Lin, National Tsing Hua University, Taiwan (Co-Chair)
林珊如　陽明交通大學(副主席)
Sunny S. J. Lin, National Yang Ming Chiao Tung University, Taiwan (Co-Chair)
于富云　成功大学(副主席)
Fu-Yun Yu, National Cheng Kung University, Taiwan (Co-Chair)

W06 创新互动回馈科技提升学习动机

孙之元　阳明交通大学(主席)
Jerry Chih-Yuan Sun, National Yang Ming Chiao Tung University, Taiwan (Chair)
陈攸华　中央大学(副主席)
Sherry Y. Chen, National Central University, Taiwan (Co-Chair)
朱蕙君　東吳大学(副主席)
Hui-Chun Chu, Soochow University, Taiwan (Co-Chair)
陈铿任　阳明交通大学(副主席)
Ken-Zen Chen, National Yang Ming Chiao Tung University, Taiwan (Co-Chair)
游师柔　阳明交通大学(副主席)
Shih-Jou Yu, National Yang Ming Chiao Tung University, Taiwan

(Co-Chair)
张琼方　阳明交通大学(副主席)
Chiung-Fang Chang, National Yang Ming Chiao Tung University, Taiwan (Co-Chair)

W07 人机协同视域下的协作学习设计、分析与评价

马志强　江南大学(主席)
Zhiqiang Ma, Jiangnan University (Chair)
张　思　华中师范大学(副主席)
Si Zhang, Central China Normal University (Co-Chair)
段金菊　西南大学(副主席)
Jinju Duan, Southwest University (Co-Chair)
王　靖　江南大学(副主席)
Jing Wang, Jiangnan University (Co-Chair)

W08 中小学新一代人工智能教育:培育未来的创新者

叶新东 温州大学(主席)
Xindong Ye, WenZhou University (Chair)
谢作如 上海人工智能实验室/温州科技高级中学 (副主席)
Zuoru Xie, Shanghai Artificial Intelligence Laboratory/Wenzhou High School of Science and Technology (Co-Chair)
杨启凡 福建师范大学(副主席)
Qi-Fan Yang, FuJian Normal University (Co-Chair)

W09 人工智能赋能知识建构创新研究与实践

赵建华 南方科技大学(主席)
Jianhua Zhao, Southern University of Science and Technology (Chair)
陈桂涓 香港大学(副主席)
Carol K.K. Chan, The University of Hong Kong (Co-Chair)

親子關係和小學生學習毅力影響學生在家長支持在疫情後手機成癮之研究

Children's Smartphone Addiction after COVID-19 Mitigated by Parent-Child Relationships and Supporting, and Child's Persistence in Study

何美慧
國立臺灣師範大學創造力發展碩士在職專班
pureho@gmail.com

【摘要】 受新冠病毒流行影響,迫使學生轉而在家進行線上學習。這種學習方式使得孩子們更加依賴使用智能手機，許多研究指出這種使用方式對孩子繼續使用智能手機產生了影響。因此，本研究探討了在疫情爆發後在家進行線上學習對親子關係以及對孩子智能手機成癮的影響。自 2023/10/3~2023/11/10 止使用網路問卷 Google 表單，進行線上調查收集數據，並進行結構方程模型分析，本研究的結果顯示，在疫情期間，在線上學習的情境中，親子關係和父母的支持之間存在正相關。父母的支持和孩子在學習上的堅持表現出與智能手機成癮之間的負相關。

【關鍵詞】 親子關係；支持度；堅持度；手機成癮

Abstract: *Affected by COVID-19, students were forced to take more online learning at home. During this period, children have become more much relied on using smartphone on their studies, and because this familiarity which affected children continuous intention to use smartphone. Therefore, this study explored the impact of parent-child relationships and affecting children's smartphone addiction after the epidemic. From 10/3 2023, to 11/10 2023, this study collecting data through online surveys using Google Forms and conducting structural equation modeling, the results of this study indicate that in the context of online learning, parent-child relationships can positively predict parental supporting in using smartphone, and persistence in studying. Parental supporting and children persistence in studying show a negative correlation with smartphone addiction.*

Keywords: parent-child relationship, support, persistence, mobile phone addiction

1.前言

先前的實務與理論研究顯示，負面的親子關係可能是智能手機成癮的重要風險因素（Buctot et al., 2020；Nur et al., 2021）。負面的親子關係指的是涉及冷漠、拒絕、不關心的或忽視的育兒方式，這導致孩子缺乏反應能力、缺乏溫暖和過度控制（Greene et al., 2020；Tehrani et al., 2023）。研究已經確認，負面的親子關係對兒童的社會情感發展產生有害影響，例如情緒調節、自我控制和社會適應（Xu＆Yan, 2023）。儘管已有多項研究探討了親子關係對智能手機成癮的影響（Wang et al., 2023；Yao et al., 2022），但這些影響背後的機制仍然不明確。因此，本研究的目的是檢驗在親子關係和智能手機成癮之間的支持和/或監管態度之間的中介作用。研究結果將不僅有助於促進對智能手機成癮機制的理解，還將為干預大學生智能手機成癮提供潛在選擇。

疫情嚴重期間，個人在家隔離期間主要透過網路科技與世界聯繫、取得疫情資訊（Hu et al.，2022）。更重要的是，為了逃避和緩解因孤立而產生的壓抑情緒，最可能、最便捷的方式就是透過智慧型手機上網（Blasi et al., 2019）。因此，由於居家隔離的空間封閉性和智慧型手機的便利性，人們最有可能出現智慧型手機成癮，即個體強烈或無法控制地專注於智慧型手

機的使用，以至於忽視了生活的其他方面，從而導致心理上的依賴和行為問題（Horvath et al., 2020）。然而，認知控制機制已被證明參與成癮行為的發展和持續（Wang et al., 2020）。在神經層面上，成癮行為是大腦前額葉、皮質和頂葉神經基質的運作（Schmitgen et al., 2020），而影響個人的成癮行為的持續性，逐漸形成認知控制機制是否可以抑制成癮行為。根據這一觀點，Covid-19 時期後與智慧型手機成癮行為認知控制的持續發展有密切相關，因而，疫情後的手機沈癮的現象是值得探索的。

根據補償性網路使用理論（Kardefelt-Winther, 2014），個人可能對智能手機上癮，是因為他們在現實中感到不快樂，因此使用智能手機逃避現實並緩解焦慮。糟糕的親子關係可能使青少年感到在現實中不受歡迎，這也是智能手機成癮的風險因素（Gong et al., 2022）。相反，良好的親子關係是防範智能手機成癮的保護因素（Gong et al., 2022）。父母接納-拒絕理論（PART）強調，父母的接受/拒絕可能影響青少年的行為、認知和情感發展（Fard et al., 2015）。PART 強調父母的支持促進青少年的心理健康，而父母的拒絕則產生相反的效果（Yang et al., 2023；Yue et al., 2022）。從以上觀點可以看出，親子關係、支持、監管和智能手機成癮之間存在著緊密的關聯。然而，我們尚未找到任何實證研究來驗證它們之間的具體聯繫。因此，具體研究目的在探討在親子關係中，父母的支持和監管對智能手機成癮的影響，以及其是否發揮中介作用。

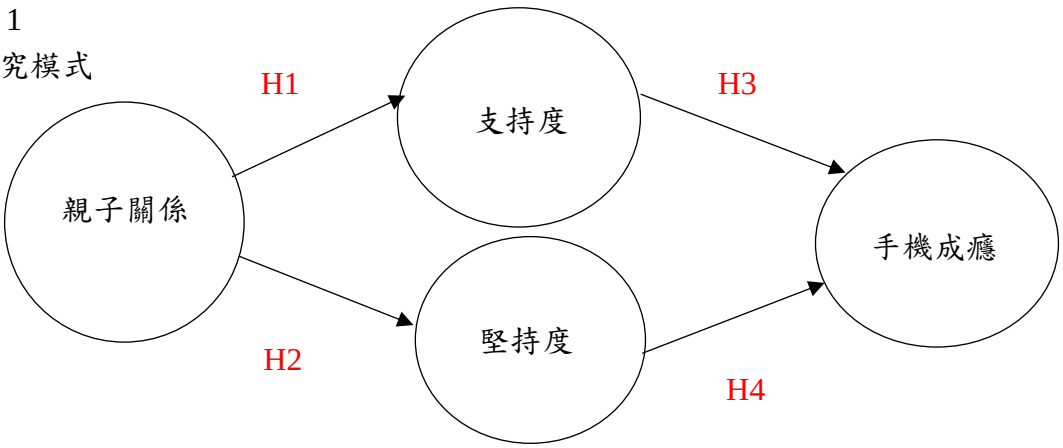
2.研究方法

2.1.研究架構與假設

值得注意的是，父母心理控制在很大程度上不同於過度干預/保護(Nelson et al., 2013)。具體而言，親子關係對子女的自主性、學業表現認同和能力感有相關，不良的親子關係導致子女通過過度使用互聯網來滿足其心理需求，並且在學業上失去自信心 (Shek et al., 2018)。有證據顯示，親子關係與智能手機成癮存在顯著正相關關係(Jiang et al., 2022)。問題行為理論強調，不良的親子關係會影響子女問題行為（如智能手機成癮 (Yao et al., 2021)。親子關係與智能手機成癮行為之間是否存在顯著關聯可能取決於中介因素。有學者指出，親子關係與子女學習的堅持性和父母對子女學習的支持之間存在中介關聯（Wu et al., 2022）。因此，本研究提出架構如圖 1 所示。本研究擬檢驗之假設如下，以檢驗研究變量之間的關係。

H1：親子關係與支持度呈現正相關。
H2：親子關係與堅持度呈現正相關。
H3：支持度與手機成癮呈現負相關。
H4：堅持度與手機成癮呈現負相關。

圖 1
研究模式



2.2.研究程序與對象

本研究以 Google 表單建立網路問卷，並利用臉書、Instagram、PTT 等社交平台廣發問卷，採取隨機抽樣方式來。以小學生家長為施測對象，主要為在 COVID-19 疫情期間，因學校停課措施而進行線上學習的學生家長擔任支持者的角色。透過網路問卷收集樣本進行分析與了解。問卷經回收後，一共獲取了 205 份樣本，當中共有 201 份有效問卷，問卷之有效回收率為 98.05%。其中，男性參與者共有 68 人 (33.8%) ；女性參與者共有 133 人 (66.2%)。

2.3.研究工具

本研究之工具採用問卷調查法，使用「疫情期間國小學生居家上課對親子關係的影響」問卷請家長以匿名方式進行填答，問卷包含四個部分(如附件一)，分別為第一部分個人基本資料、第二部分親子關係(生活相關) 、第三部分支持度(課業相關) 、第四部份堅持度、第五部分手機成癮、研究者使用網路問卷工具（Google 表單）建立了一份結構化的問卷，收集家長的經歷感受。

3.研究結果

3.1.題項分析

(一)本研究在項目分析中採用一階驗證性因素分析作為內部效度檢定方法。其檢定如表 1，如若未達此數值則須將此題刪除 (Hair et al., 2019)。(二)本研究在項目分析外部效度。Green 與 Salkind (2004) 指出 t 值超過 3 (p*** < .001)，即達到顯著之標準。表 1 之數值介於 27.22 至 95.32 之間，表示是具有外度效度。經分析後得知，親子關係之題項從 6 題刪至 4 題；支持度之題項從 6 題刪至 4 題；Persistence (堅持度)之題項從 6 題刪至 4 題；手機成癮之題項從 6 題刪至 4 題。

表 1

各構面之一階驗證性因素分析整合表

適配度	*χ2*	*df.*	*χ2/df.*	RMSEA	GFI	AGFI	FL	*t*-value
臨界值	---	---	<5	<0.10	>0.80	>0.80	>0.50	>3
親子關係	4.72	2	2.36	0.08	0.99	0.94	0.74-0.77	51.16-67.88
支持度	4.99	2	2.50	0.09	0.99	0.94	0.57-0.89	50.68-95.32
堅持度	0.72	2	0.36	0.00	0.98	0.88	0.81-0.85	55.24-75.75
手機成癮	2.24	2	1.12	0.02	0.99	0.97	0.81-0.88	27.22-35.84

3.2.構面信度與效度分析

Emerson (2019)指出 Cronbach's α 值若約為 0.7 之標準，方可視為具有較良好的內部一致性；CR 值高 0.7 之標準，方可視為具有良好的信度。在測試構面的收斂效度時，Hair 等人 (2019)也建議因子負荷 (FL)和平均提取變異數(AVE)應高於 0.5。表 2 顯示，本研究 Cronbach 的 α 值虛擬 0.71 至 0.89 間；CR 值虛擬 0.83 至 0.92 間；FL 值閒置 0.05 至 0.88 間；AVE 值閒置 0.56 至 0.74 間。相應地，構面的內部和外部信度以及收斂效度的值是可以接受的。在檢驗構面的區別效度，Zainudin (2015) 強調 AVE 的平方根值應高於構面間 Pearson 相關係數的絕對值。表 3 顯示構面的區別效度是可以接受的。

表 2

各構面信度與收斂效度一覽表

構面	M	SD	Cronbach's α	CR	*FL*	AVE
親子關係	4.13	.65	0.71	0.84	0.87	0.56
支持度	4.22	.67	0.74	0.83	0.86	0.56
堅持度	4.13	.72	0.85	0.90	0.88	0.69
手機成癮	3.23	1.12	0.89	0.92	0.05	0.74

表 3

構面的區別效度分析

構面	1	2	3	4
1. 親子關係	(0.76)			
2. 支持度	0.59	(0.75)		
3. 堅持度	0.63	0.65	(0.83)	
4. 手機成癮	0.51	0.52	0.61	(0.86)

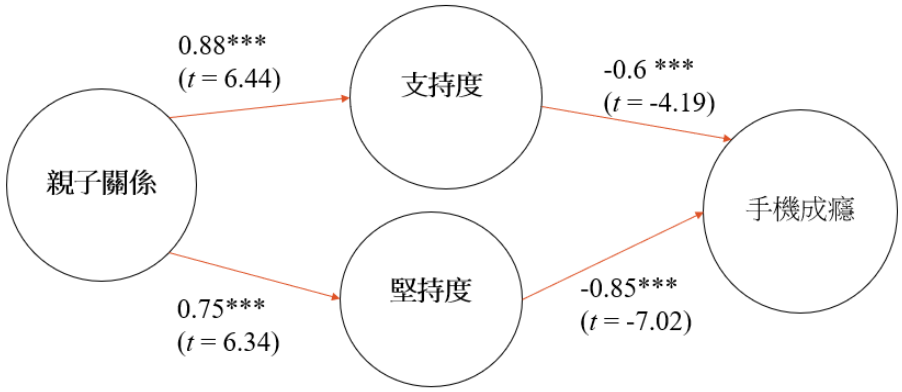
3.3. 整體適配度分析

本研究以模型適配度來檢驗研究模型的優劣，其指標為：χ2/df 數值須小於 5 ；RMSEA 數值小於 0.1；GFI、AGFI、NFI、NNFI、CFI、IFI 與 RFI 數值皆須大於 0.8；PNFI 與 PGFI 數值皆須大於 0.5 (Hair et al., 2019) ，符合上述即為有良好的模型適配度。經分析得知，χ^2/df. 為 1.89；RMSEA = 0.07；GFI = 0.9；AGFI = 0.86；NFI = 0.88；NNFI = 0. 93；CFI = 0.94；IFI = 0.94；RFI = 0.86；PNFI = 0.74；PGFI = 0.66。

3.4.路徑分析

本研究四條假設經分析得知，親子關係可以正向顯著支持次(β =0.88, t = 6.44***, p < .001) ；親子關係可以正向顯著堅持度 (β =0.75, t = 6.34***, p < .001)；支持度可以負向手機成癮(β =- 0.6, t = -4.19, p < .001) ；堅持度可以負向預測手機成癮(β = -0.85, t =-7.02, p < .001)。另外, 圖 2 顯示，親子關係對支持度為 56%（R^2=0.56），具高效果量為 f^2= 0.41；親子關係對堅持度解釋力為 68%（R^2=0.68），具高效果量為 f^2= 0.36；支持度對手機成癮解釋力為 37%（R^2 =0.37），具中效果量為 f^2= 0.26；堅持度對手機成癮解釋力為 10% (R^2 = 0.10)；具低效果量 f^2 = 0.03.。

圖 2



路徑驗証分析（*p＜.05, **p＜.01, ***p＜.001)

4.研究討論

從路徑分析測結果資料分析看來，親子關係和支持度及堅持度是正相關，支持度和堅持度對手機成癮是負相關。進一步討論如下：

4.1.假設驗證

一、親子關係與支持度的正相關：在疫情期間進行線上學習的情境中，研究揭示家長和孩子之間的親子關係和支持度之間存在正相關。發現當親子之間的關係更加親密且充滿支持時，孩子更有可能在線上學習中保持積極的態度。這樣的現象表示，擁有支持性的家長能夠建立一個鼓勵孩子參與學習的環境，進而提升其學習效果。

二、親子關係與堅持度的正相關：研究發現了親子關係與堅持度之間存在正相關。這發現表示當親子之間的關係更加親密時，孩子在面對學習上的困難時更容易保持堅持度。這可能是由於良好的親子關係提供了一個溫馨的環境，鼓勵孩子克服挑戰，培養積極的學習態度。三、支持度對手機成癮的負相關：研究結果顯示，支持度與手機成癮之間存在負相關。這表示當國小學生感受到更多的支持時，他們傾向於展現較低程度的手機成癮。這樣的結果即獲得正面支持和關愛的國小學生較少依賴手機作為應對壓力的逃避途徑。

四、堅持度對手機成癮的負相關：研究發現堅持度與手機成癮之間呈現負相關。這表示當國小學生展現較高的堅持度時，他們傾向於呈現較低程度的手機成癮。結果可能顯示，擁有較強的堅持力的個體更能夠控制手機使用，不易陷入過度使用或成癮的狀態，在疫情期間學習方式的長期影響下，手機成癮仍然可能是一個具有挑戰性的問題。

4.2.建議

研究結果顯示，針對國小學生手機成癮現象進行的樣本調查及統計分析發現，手機成癮的平均數為 3.23，呈現中等程度。這表示國小學生對手機的依賴已相當普遍，這僅是平均水平的呈現，並不足以反映手機成癮現象的全貌，因為這平均值暗示著手機成癮可能呈現持續的趨勢，針對國小學生的手機成癮問題，我們應該持續關懷，以確保他們的心理健康和全面發展。透過提供相應的教育、支持和資源，我們可以協助他們建立健康的手機使用習慣，並應對可能帶來的負面影響。

本研究則專注於國小學生在疫情期間停課居家線上學習的情境，探討這一族群的親子關係。同時，提供心理健康資源，在社會層面，廣宣對手機成癮的認識，同時強調培養良好的

數位使用習慣，有助於實現平衡，減少不良影響。總之，確保提供的支援措施全面而有效，需要深入研究和專業的參與。

5. 結論

本研究旨在深入研究新冠病毒大流行期間，實施居家線上學習措施，導致學生進行居家線上學習的情境，並探討這種學習情境對親子關係的相關影響。我們探討親子關係的多個層面，包括品質、親密度、支持、學習堅持度和手機成癮等方面。綜合上述，本研究強調在疫情期間及疫情後，家長支持和親子關係對孩子居家線上學習體驗的關鍵性。我們提供了實際可行的建議，以改善學生的學習環境。在研究限制方面：因本研究針對國小學生進行，未全面對國高中生進行研究。因此，研究結果僅代表國小學生的情況，不適用於其他年齡層次的學生。未來研究建議:可以擴大對象，包括國高中生，以比較不同年齡層學生的親子關係。此外，也可以考慮質性研究深入探討，以獲得更全面的了解，可以進行長期追蹤研究，以觀察其在成長過程中的變化和發展。

參考文獻

Blasi, M. Di, Giardina, A., Giordano, C., Coco, G. L., Tosto, C., Billieux, J., et al. (2019). Problematic video game use as an emotional coping strategy: Evidence from a sample of MMORPG gamers. *Journal of Behavioral Addictions, 8*(1), 25–34. https:// doi.org/10.1556/2006.8.2019.02

Buctot, D. B., Kim, N., & Kim, J. J. (2020). Factors associated with smartphone addiction prevalence and its predictive capacity for health-related quality of life among Filipino adolescents. *Children and Youth Services Review, 110,* Article 104758.

Fard, Z. S., Mousavi, P. S., & Pooravari, M. (2015). Predictive role of parental acceptance, rejection and control in the Internet addiction of the female students. *International Journal of Applied Behavioral Sciences, 2*(2), 42–51.

Greene, C. A., Haisley, L., Wallace, C., & Ford, J. D. (2020). Intergenerational effects of childhood maltreatment: A systematic review of the parenting practices of adult survivors of childhood abuse, neglect, and violence. *Clinical Psychology Review, 80,* Article 101891. https://doi.org/10.1016/j.cpr.2020.101891

Grolnick, W. S., & Pomerantz, E. M. (2022). Should parents be involved in their children's schooling? *Theory into Practice, 61*(3), 325–335. https://doi.org/10.1080/ 00405841.2022.2096382

Horvath, J., Mundinger, C., Schmitgen, M. M., Wolf, N. D., Sambataro, F., Hirjak, D., et al. (2020). Structural and functional correlates of smartphone addiction. *Addictive Behaviors, 105.* https://doi.org/10.1016/j.addbeh.2020.106334. Article 106334.

Hu, Q., Liu, Q., & Wang, Z. (2022). Meaning in life as a mediator between interpersonal alienation and smartphone addiction in the context of Covid-19: A three-wave longitudinal study. *Computers in Human Behavior, 127,* 107058. https://doi.org/10.1016/j.chb.2021.107058

Jiang, X. Y., Lin, Y., & Liu, Q. X. (2022). The association between parental psychological control and adolescent smartphone addiction: The role of online psychological needs satisfaction and environmental sensitivity. *Psychological development and education, 38*(2), 254–262.

Kardefelt-Winther, D. (2014). A conceptual and methodological critique of internet addiction research: Towards a model of compensatory internet use. *Computers in Human Behavior, 31*(1), 351–354. https://doi.org/10.1016/j.chb.2013.10.059

Nelson, D. A., Yang, C., Coyne, S. M., Olsen, J. A., & Hart, C. H. (2013). Parental psychological control dimensions: Connections with Russian preschoolers' physical and relational aggression. *Journal of Applied Developmental Psychology, 34*(1), 1–8.

Nur, H., Setyaningrum, P., & Novandita, A. (2021). Permissive, authoritarian, and authoritative parenting style and smartphone addiction on university students. *Journal of Educational, Health and Community Psychology, 10*(3), 419.

Schmitgen, M.M., Horvath, J., Mundinger, C., Wolf, N.D., Sambataro, F., Hirjak, D., & Wolf, R.C. (2020). Neural correlates of cue reactivity in individuals with smartphone addiction. *Addictive Behavior, 108,* 106422
https://doi.org/10.1016/j. addbeh.2020.106422

Shek, D. T., Zhu, X., & Ma, C. M. (2018). The influence of parental control and parentchild relational qualities on adolescent internet addiction: A 3-year longitudinal study in Hong Kong. *Frontiers in Psychology, 9*, 642.

Tehrani, H. D., Yamini, S., & Vazsonyi, A. T. (2023). Parenting styles and big five personality traits among adolescents: A meta-analysis. *Personality and Individual Differences, 216*, 112421. https://doi.org/10.1016/j.paid.2023.112421

Wang, H., Sun, Y., Lan, F., & Liu, Y. (2020). Altered brain network topology related to working memory in internet addiction. 9(2), 325. doi:10.1556/2006.2020.00020.

Wang, J., Li, M., Geng, J., Wang, H., Nie, J., & Lei, L. (2023). Meaning in life and selfcontrol mediate the potential contribution of harsh parenting to adolescents' problematic smartphone use: Longitudinal multi-group analyses. *Journal of Interpersonal Violence, 38*(1–2), 2159–2181.

Wu, J., Barger, M. M., Oh, D., & Pomerantz, E. M. (2022). Parents' daily involvement in children's math homework and activities during early elementary school. *Child Development, 93*(5), 1347–1364. https://doi.org/10.1111/cdev.13774

Xu, C., & Yan, W. (2023). Negative parenting styles and social adjustment of university students: A moderated chain mediation model. *Current Psychology, 42*(31), 27719–27732. https://doi.org/10.1007/s12144-022-03809-1

Yang, L.L., Guo, C., Ligny., Gan, K. P., & Luo, J.H. (2023). Mobile phone addiction and mental health the roles of sleep quality and perceived social support. *Frontiers in Psychology, 14*, 1265400.

Yao, X., Wu, J., Guo, Z., Yang, Y., Zhang, M., Zhao, Y., & Kou, Y. (2022). Parental psychological control and adolescents' problematic mobile phone use: The serial mediation of Basic Psychological need experiences and negative affect. *Journal of Child and Family Studies, 31*(7), 2039–2049.
https://doi.org/10.1007/s10826-021- 02169-x

Yue, Y., Aibao, Z., & Tang, T. H. (2022). The interconnections among the intensity of social network use, anxiety, smartphone addiction and the parent-child relationship of adolescents: A moderated mediation effect. *Acta Psychologica, 231*, 103796.

探究教育科技應用於國小學童執行功能與創造力遊戲表現之差異

A Study of Independent Sample t-Test on the Performance of Executive Function and Creativity Game among Primary School Children with Educational Technology

蔡仲倫
國立臺灣師範大學創造力發展碩士在職專班
kevin103504023@gmail.com

【摘要】本研究旨在探討國小四年級學童在執行功能與創造力遊戲表現的關係，並探究性別及高低分組之差異。以立意取樣方式進行問卷調查及平板操作實測，研究工具包含「青少年執行功能行為評量表」及「創造力遊戲應用程式」。60 份有效樣本經由描述統計及獨立樣本 t 檢定進行資料分析後，本研究結果發現：（一）執行功能測驗中呈現顯著性別差異；（二）創造力遊戲測驗未呈現顯著性別差異；（三）執行功能測驗高低分組學生在創造力遊戲測驗未呈現顯著差異；（四）創造力遊戲測驗高低分組學生在執行功能測驗呈現顯著差異。

關鍵詞：國小學童、執行功能、創造力、創造力遊戲

Abstract: *This study aimed to explore the relationship between the performance in executive function and creative game. Questionnaires and tablet-based practical tests were conducted on fourth-grade student through purposive sampling. The research tools included the " Behavioral Rating Scale for Executive Function (BRSEF)" and the app "Assessing Creativity ". 60 valid samples were analyzed using "descriptive statistics" and "Independent Sample t-test". In the BRSEF, girls significantly outperformed boys; students in the high- and low- scoring group on the performance of creative game showed significant difference. However, no significant gender differences can be found in the creative game, and students in the high- and low- scoring group on the performance of BRSEF did not show significant differences in the creative game.*

Keywords: Primary school children, Executive function, Creativity, Creative game

1. 前言

近年來，各國教育當局強調培養未來人才對社會的重要性，並著手從基礎教育開始培育具有關鍵學習能力的兒童。在兒童的認知能力發展中，執行功能（Executive Function, EF）被確認為一個極為重要的因素，其被定義為一系列能夠調節思維和行為的認知控制過程（Miyake & Friedman, 2012），包含更新、計劃、組織、自我控制等高層次的認知功能。各國研究強調執行功能在兒童到青春期發展階段中扮演關鍵角色，對學習和學業成就具有基礎性的影響（Cragg et al., 2017）。此外，執行功能的發展程度還會影響一個人的行為規範、獨立思考、自我主動學習，以及問題解決等多方面的能力（Jacob & Parkinson, 2015）。這凸顯了學術界對執行功能研究的重視，反映了對兒童全面能力發展的迫切需求。

創造力同執行功能一樣被視為關鍵的認知發展能力，在過去研究中顯示，前額葉是負責執行功能的大腦區域，也是進行各類創意思考時共同激發的區域（Dietrich & Kanso, 2010）。

目前心理學家廣泛使用兩種評估創造力的測試問卷，分別為 Torrance Tests of Creative Thinking（TTCT）以及 Williams Creative Thinking Test（WCTT），兩者皆將創造力視為多種能力的綜合表現，而近年來在數位科技迅速發展和數位遊戲蓬勃興起的趨勢下，結合數位科技工具將使教育更貼近生活。

越來越多的研究顯示執行功能與創造力之間存在正向關聯性（Zabelina et al., 2019）。然而，過去的研究多使用紙筆心理計量方法，對於數位遊戲對執行功能和創造力的影響較少探討。因此，本研究希望透過使用創造性遊戲應用程式，探索學生在與過去不同的情境中展現創造性思維能力的表現，採用臺灣師範大學數位學習實驗室所開發的 APP 來評量受試者的創造力表現，並探究學生在執行功能和創造力之間的差異及表現，以及可能存在的影響因素。本研究之目的為下列二項：(1)探究國小四年級男女學童在執行功能與創造力遊戲表現之差異為何？(2)探究國小四年級學童在執行功能與創造力遊戲不同方面間之差異及表現為何？

2.文獻探討

執行功能（Executive Functions, EF）統稱由大腦前額葉所掌管的一系列基礎與高層次的認知功能，能夠調節我們日常生活中的思想和行為，其中工作記憶（working memory/updating）、抑制控制（inhibitory control）和認知彈性（cognitive flexibility/shifting）被認為是組成執行功能的三個核心要素（Diamond, 2013）。近年來，研究將其分為冷（cold）和熱（hot）兩種類別，冷執行功能是在沒有情感影響的環境中觸發的認知能力，包括規劃、語言分析、問題解決、工作記憶、認知彈性、抑制和持續注意等。相反，熱執行功能是在情感環境下引發的認知能力，包括社會認知、情感調節、情感決策以及延遲滿足等（劉惠美，2021）。過去研究發現性別在執行功能表現上存在差異，例如女生在在情緒反應控制、認知重估和情緒應對方面表現較佳（Nolen-Hoeksema, 2012），而男生在持續注意力表現上更快速且更一致（Riley et al., 2016），故本研究欲深入探討性別在執行功能表現上之差異。

在認知發展研究中，執行功能和創造力的發展一直是研究的焦點。創造力（Creativity）被定義為一種能創造出原創、新穎、適當或有意義的東西的能力（Amabile, 1983）。根據美國學者 Guilford（1956）理論，創造力的思考歷程分為聚斂性思考和擴散性思考兩種，其中「擴散性思考」與創造力的表現密切相關，且包含四種特質，分別為流暢力、變通力、獨創力及精進力。過去的研究對性別與創造力之間的關係得出了不同的結果。有研究指出發現女生在流暢力、獨創力和精進力方面表現較優越（Krumm et al., 2014），然而，也有研究顯示男生在某些創造力方面的表現可能優於女性（張秀雯，2010）。因此，本研究欲進一步探討性別在創造力表現上之差異。

學界普遍認為執行功能和創造力皆是認知發展中的關鍵歷程，同時也與個體的認知活動密切相關（游孟諺，2020）。在 Collin 與 Koechlin（2012）研究中發現，當執行功能無法順利運作時，創造性思考能力就會介入，開始探索並選擇新的行為策略。Sharma 與 Babu（2017）的研究結果也指出，具有高於平均水平創造力的個體在執行功能方面呈現明顯的優越表現。在兒童求學階段中，要有好的表現需要創造力、注意力、靈活性及自我控制能力等與執行功能相關的核心能力（Diamond & Lee, 2011）。由此可知，執行功能與創造力等多重能力的共同發展是學生獲得成功的關鍵。

另外，有些研究針對執行功能的核心要素與創造力指標之關係進行關注及檢驗，且得到不同的結果。Benedek 等人（2014）使用創造力的擴散性思考測驗，同時也針對執行功能三大核心能力（更新、轉換和抑制）進行評估。研究結果指出，更新和抑制能力與擴散性創造力的表現呈現預測性關係，然而轉換能力則未能預測擴散性創造力的表現。

綜上所述，執行功能與創造力在個體發展中扮演著複雜而多面的角色。針對不同的研究結果，執行功能與創造力在不同面向之間存在特定的相互關係。創造力的表現可能受到執行功能不同能力的調節和影響，同時，創造力的提升則可能反過來影響執行功能的運作，兩者之間可能存在雙向的連結。故本研究欲深入探討執行功能和創造力之間的關係，以期深化對這兩個重要認知能力如何相互影響的理解。

3.研究方法

3.1. 研究架構假設.

執行功能的發展可能有助於培養學生的思維能力，例如創造力中的擴散性思考，這與性別和年齡相關聯（Pasarín-Lavín et al., 2024）。因此，本研究旨在了解國小四年級學童在執行功能與創造力遊戲表現的關係，並深入探究不同性別在這兩者能力之間的差異及表現，同時亦探討執行功能各分向能力與創造力遊戲各指標在高低分組下之差異及表現。本研究採用問卷調查法及平板操作實測，問卷資料以自陳式量表進行蒐集。依據研究目的以及相關文獻，本研究架構如圖 1 所示。

H1：國小四年級學童在執行功能測驗中呈現顯著性別差異，且女生的表現優於男生。

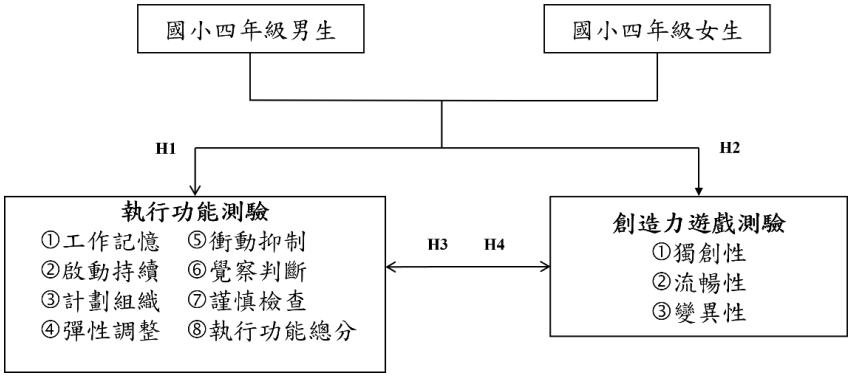
H2：國小四年級學童在創造力遊戲測驗中呈現顯著性別差異，且女生的表現優於男生。

H3：執行功能測驗高分組的學生在創造力遊戲測驗表現上顯著優於低分組的學生

H4：創造力遊戲測驗高分組的學生在執行功能測驗表現上顯著優於低分組的學生。

圖 1

研究架構圖



3.2.程序與參與者

本研究針對臺北市某公立國小四年級學童進行立意取樣，共計 93 名學生參與。其中，33 名學生因未完全參與問卷和遊戲測驗而被排除，最終得到 60 份有效樣本。參與者中，男性占 56.7%，女性占 43.3%，平均年齡為 10.4±0.5 歲。研究於 2023 年 10 月準備各項評估工具，11 月 3 日徵詢校方同意，11 月 14 日正式進入校園施行測量。在施測預先完成所有前置作業，並向班級老師領取家長同意書和名條。在課堂上進行創造力遊戲和執行功能量表測試，並提醒學生遵守測試規則。測試過程中保持安靜，並提供必要協助。整個測試過程共計 50 分鐘。完成測試的學生獲得小禮物以示感謝。

3.3.研究工具

本研究之工具共計二項，分別為「青少年執行功能行為評量表」及「Assessing Creativity 遊戲應用程式」。為不影響量表之內容與信效度，作答方式及計分方式皆依原始版本執行。此外，為檢驗每份量表之有效性，於問卷中置入二題「檢驗題」，分別於「青少年執行功能行為評量表」完整版第 71 和 72 題。以下將說明各項測量工具之內容。

3.3.1. 青少年執行功能行為評量表（Behavioral Rating Scale for Executive Function, BRSEF）

本研究採用「青少年執行功能行為評量表」(陳心怡等人，2020)，其蒐集了 7,925 位年齡範圍在 9 至 17 歲學生（包括國小、國中、和高中/職階段）自陳的執行功能行為表現，建置了具穩健性及代表性的臺灣常模資料。量表共 72 題（包含 70 題正式題和 2 題檢驗題)，提供一個執行功能總分與七個分向度分數，分別為「工作記憶」、「啟動持續」、「計畫組織」、「彈性調整」、「衝動抑制」、「覺察判斷」、「謹慎檢查」，並採用 Likert 五點量表，分數越高表示執行功能越佳。經多元檢驗顯示，該量表具有良好的信度及效度，於執行功能總分信度達.96 -.98，各分量表信度則均在.82-91 之間；效度方面，該量表與語意流暢度（Verbal Fluency）及兒童彩色路徑描繪測驗（CCTT）得分之間呈現顯著相關，顯示良好的效標關聯效度，也與一般教師在課堂中觀察學生的執行功能評分之間存在顯著相關性，顯示該評分工具擁有良好的生態效度。

3.3.2. Assessing Creativity 遊戲應用程式 本研究採用由臺灣師範大學工業教育學系洪榮昭教授研究團隊所開發之創造力遊戲應用程式，其操作型定義為參考 Torrance（1966）所發展出來之「陶倫斯創造思考測驗（TTCT)」。本研究採用其中三項創造力評量電腦化作業（Tsai et al., 2022)，以觸控手機、平板的遊戲方式，測量個別學生之精緻性創意或創新性思考力表現，以下說明其設計原理與作業施測方式：

3.3.2.1. 一筆劃遊戲（One-line Drawing）

該測驗主要評估使用者在創造力思考中「獨創性」、「流暢性」、「變異性」之表現。參與者於測驗期間需以手指按住螢幕的方式畫出各種不同的圖，並為所繪製的圖畫輸入名稱。此外每一張圖只能以一筆完成，畫完前手指不可以離開螢幕，測驗時間共 8 分鐘。

3.3.2.2.屬性聯想遊戲（Attributes Association）

該測驗主要評估使用者在創造力思考中「獨創性」、「流暢性」、「變異性」之表現。參與者需於畫面上所出現的物體進行點選分類，且每個類別至少要有兩個，並將分類的原因以文字敘述的方式呈現。測驗時間共 8 分鐘，參與者需於時間內盡量想各種不同的分類方式組合，且越多越好。

3.3.2.3. 簡圖聯想遊戲（Sketch Association）

該測驗主要評估使用者在創造力思考中「獨創性」、「流暢性」、「變異性」之表現。參與者需於螢幕上所提供之圖片進行圖形聯想，並將聯想到的圖案以文字敘述的方式呈現。測驗時間共 8 分鐘，參與者需於時間內盡可能想出具創意且符合圖形邏輯之答案。

3.4. 資料處理與分析

本研究運用 SPSS 23.0 統計分析軟體進行資料處理與分析，以獨立樣本 t 檢定檢驗性別在執行功能總分、各向度能力與創造力指標是否具有顯著差異。本研究亦針對執行功能七個分向能力與創造力遊戲三個指標進行高低分組的差異檢驗，探討兩項能力在高低分組下，彼此間是否呈現顯著性差異。

4.研究結果

4.1. 主要結果分析

本研究以獨立樣本 t 檢定，檢驗性別在執行功能總分、各向度能力與創造力的表現是否存在顯著性差異，亦針對執行功能七個分向能力與創造力遊戲三個指標進行高低分組，檢驗彼此的表現是否呈現顯著性差異。

4.1.1.性別與執行功能、創造力遊戲指標之分析

表 1 顯示，執行功能總分及各向度能力的整體平均得分，及男女生在各向度能力的得分差異，可發現女生在執行功能表現上皆優於男生，且在衝動抑制（t = -2.04, p < .05）、覺察判斷（t = -2.33, p < .05）和謹慎檢查（t = -2.33, p < .05）呈現顯著差異，其餘分向能力則未呈現顯著差異（p > .05）；而在創造力遊戲指標方面，研究結果顯示男生雖然在「獨創性」、「流暢性」、「變異性」的表現上皆優於女生，但各項指標皆未呈現顯著差異（p > .05）。

表 1

性別與執行功能總分、各向度能力之 t 檢定差異分析（N=60）

	總人數 (*N*=60)		男 (*N*=34)		女 (*N*=26)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*t* 值
執行功能總分	3.26	.75	3.11	.69	3.46	.78	-1.83
工作記憶	3.02	.96	2.96	.86	3.10	1.10	-.56
啟動持續	3.39	.89	3.28	.77	3.52	1.02	-1.05
計畫組織	3.42	.83	3.29	.72	3.59	.95	-1.35
彈性調整	3.24	.78	3.08	.79	3.45	.73	-1.87
衝動抑制	3.28	1.00	3.05	.96	3.57	1.00	-2.04*
覺察判斷	3.54	.83	3.33	.82	3.82	.77	-2.33*
謹慎檢查	3.59	.90	3.36	.89	3.89	.83	-2.33*

*p < .05.

4.1.2.執行功能分向能力與創造力遊戲指標之高低分組分析

本研究將執行功能七個分向能力分數之前 27%（高分組）與後 27%（低分組），以獨立樣本 t 檢定檢驗兩極端組在創造力遊戲指標之表現是否有顯著差異，結果顯示，執行功能七個分向能力之高低分組在創造力遊戲各指標之表現皆未呈現顯著差異（p > .05）。在創造力遊戲指標方面，本研究將三個指標分數之前 27%（高分組）與後 27%（低分組），以獨立樣本 t 檢定檢驗兩極端組在執行功能七個分向能力之表現是否有顯著差異，結果顯示，獨創性和流暢性的高低分組在執行功能各分向能力之表現未呈現顯著差異（p > .05）。而變異性的高低分組（如表 2）在工作記憶（t = -2.09, p < .05）與彈性調整（t = -2.44, p < .05）的表現上呈現顯著差異，其餘分向能力則未呈現顯著差異（p > .05）。

表 2

變異性與執行功能分向能力之 t 檢定差異分析（N=32）

	高分組(27%) (*N*=16)		低分組(27%) (*N*=16)		
	Mean	*SD*	*Mean*	*SD*	*t* 值
工作記憶	3.33	1.13	2.53	1.03	-2.09*
啟動持續	3.59	0.96	3.08	.96	-1.51
計畫組織	3.49	1.05	3.23	.76	-.81
彈性調整	3.55	.85	2.83	.82	-2.44*
衝動抑制	3.36	1.03	3.17	.96	-.53
覺察判斷	3.73	.91	3.40	.92	-.103
謹慎檢查	3.64	1.15	3.58	.85	-.18

*p < .05.

5.討論

5.1.結果意涵

本研究假設一：「國小四年級學童在執行功能測驗中呈現顯著性別差異，且女生的表現優於男生」獲得支持。研究結果顯示，女生在衝動抑制、覺察判斷和謹慎檢查方面的表現顯著優於男生，符合 Barnett 等人（2007）的主張，即在兒童發展階段，女生比男生表現更為不衝動，亦與 Bezdjian 等人（2009）的研究發現一致，在兒童至青少年發展時期，男生相較於女生更加衝動。推測男生可能具有比女生更快的反應時間，而這可能是導致衝動行為的增加，或在判斷和檢查等能力方面存在差異的原因。

本研究假設二：「國小四年級學童在創造力遊戲測驗中呈現顯著性別差異，且女生的表現優於男生」未獲得支持。此結果與過去的研究不盡相符（Garin et al., 2016）。雖然男生在創造力遊戲各指標的表現上優於女生，然經分析後發現各項指標未呈現顯著差異。推測幾個可能的原因：一，本研究所蒐集之樣本數性別比例不夠平均，故未能呈現真實的性別差異；二，國小四年級學童在數位化的創造力遊戲的使用上可能受到其他因素的影響，如：學習環境的改變、遊戲設計的差異等，影響了性別差異的顯著性。然本研究結果與 Abraham（2016）的研究一致，顯示男女生之間在創造力測量上沒有顯著差異，故仍有待未來進一步探究。

本研究假設三：「在高低分組下，執行功能測驗高分組的學生在創造力遊戲測驗表現上顯著優於低分組的學生」未獲得支持。研究結果顯示，執行功能七個分向能力之高低分組在創造力遊戲各指標之表現皆未呈現顯著差異，與過去的研究並不一致（Huntsinger & Ray, 2016）。認為造成的原因為研究工具使用的差異，本研究在執行功能測驗上採用問卷調查法，然在創造力遊戲測驗上則採用數位應用程式進行，可能影響了研究結果的一致性。未來研究可進一步探討不同工具之間的比較，以更全面地理解執行功能和創造力之間的關係。

本研究假設四：「在高低分組下，創造力遊戲測驗高分組的學生在執行功能測驗表現上顯著優於低分組的學生」獲得支持。研究結果顯示，變異性的高低分組在工作記憶與彈性調整的表現上呈現顯著差異。本研究結果與 Zabelina 等人（2016）的研究發現相似，當個體擴散性思考能力越好，其轉移能力和認知彈性較佳。研究者認為能夠產生多樣性且非傳統性想法的個體可能更能靈活應對新情境，並展現更好的認知適應性。Orzechowski 等人（2023）的研究指出，擴散性思考中的靈活性與工作記憶機制的效能呈現正向關聯性。這顯示在進行擴散性思考時，個體展現出更佳的靈活性，同時也表現出更好的工作記憶，與本研究結果相似。

5.2.研究限制及未來研究建議

本研究因受限於時間及人力等因素，在研究推論上有其限制。由於本研究的參與者僅限於臺北市某公立國小四年級學童，在樣本的蒐集上未能廣泛涵蓋多個學校或年級，對研究結果的全面性可能產生一定的影響。建議未來研究可擴展研究對象與研究範圍，使其研究結果更具代表性。此外，本研究因時間與資源的限制，在執行功能測驗以及創造力遊戲測驗上僅以各一種測量工具來進行研究，未來建議可以增加不同的測量工具，有助於更全面地評估研究參與者之執行功能與創造力的表現，透過結合多元測量工具，研究結果將更具可信度且更具有廣泛的應用價值。

6.結論

本研究旨為探究國小四年級學童在執行功能與創造力遊戲表現之關係，並深入探究不同性別在這兩項能力之間的差異及表現，同時亦探討執行功能各分向能力與創造力遊戲各指標在高低分組下之差異及表現。研究結果顯示國小四年級學童在執行功能測驗中呈現顯著性別差異，且女生在衝動抑制、覺察判斷和謹慎檢查方面的表現顯著優於男生。而在創造力遊戲測驗中則未呈現顯著性別差異，此結果可能顯示在創造性思考方面，性別差異相對較小或不

明顯。針對執行功能各分向能力與創造力遊戲各指標在高低分組下之差異及表現，結果顯示僅在創造力遊戲測驗中，高低分組之間的變異性指標與工作記憶和彈性調整存在顯著差異。此結果可能意味著參與創造性活動之個體，其越能表現出變異性，在工作記憶和彈性調整兩個方面表現更為優越。本研究期待未來能更深入探索性別在執行功能和創造力方面的差異，並進一步研究這兩者之間的相互作用，以促進學童在這兩個能力上的發展。

參考文獻

張秀雯(2010)。**國小學童知覺體育教師教學風格與創造力之相關研究**〔未出版之碩士論文〕。臺北市立大學。

陳心怡、陳美芳、胡心慈（2020）。**青少年執行功能行為評量表**。中國行為科學社股份有限公司。

游孟諺（2020）。**執行功能訓練融入國語課程對國小學童執行功能、創造性思考及國語學習表現之效果研究**〔未出版之碩士論文〕。國立臺灣師範大學。

劉惠美（2021）。**推動以「提升執行功能」和「培育素養」為核心之教育評估與介入方案**（計劃編號：MOST 109-2420-H-003-002）。臺北市：科技部。

Abraham, A. (2016). Gender and creativity: an overview of psychological and neuroscientific literature. *Brain imaging and behavior*, *10*(2), 609-618. https://doi.org/10.1007/s11682-015-9410-8

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357-376. https://doi.org/10.1037/0022-3514.45.2.357

Barnett, M. L. (2007). Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. *Academy of management review*, *32*(3), 794-816. https://doi.org/10.5465/AMR.2007.25275520

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73-83. https://doi.org/10.1016/j.intell.2014.05.007

Bezdjian, S., Baker, L. A., Lozano, D. I., & Raine, A. (2009). Assessing inattention and impulsivity in children during the Go/NoGo task. *British Journal of Developmental Psychology*, *27*(2), 365-383. https://doi.org/10.1348/026151008X314919

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition, 162*, 12-26. https://doi.org/10.1016/j.cognition.2017.01.014

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168. http://dx.doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959-964. https://doi.org/10.1126/science.1204529

Dietrich, A., & Kanso, R. (2010). A Review of EEG, ERP, and Neuroimaging Studies of Creativity and Insight. *Psychological Bulletin and Review*, *136*, 822-848. https://doi.org/10.1037/a0019749

Garin Vallverdu, M. P., Lopez Fernandez, V., & Llamas Salguero, F. (2016). Creativity and Multiple Intelligences Students by Gender in Primary Education. *REIDOCREA-REVISTA*

EECTRONICA DE INVESTIGACION Y DOCENCIA CREATIVA, *5*, 33-39 https://doi.org/10.30827/Digibug.40068

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*(4), 267. http://dx.doi.org/10.1037/h0040755

Huntsinger, J. R., & Ray, C. (2016). A flexible influence of affective feelings on creative and analytic performance. *Emotion*, *16*(6), 826-837. https://doi.org/10.1037/emo0000188

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement. *Review of Educational Research*, *85*(4), 512-552. https://doi.org/10.3102/0034654314561338

Krumm, G., Lemos, V., & Filippetti, V. A. (2014). Factor structure of the Torrance Tests of Creative Thinking Figural Form B in Spanish-speaking children: Measurement invariance across gender. *Creativity Research Journal*, *26*(1), 72-81. http://doi.org/10.1080/10400419.2013.843908

Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: The role of gender. *Annual review of clinical psychology*, 8, 161-187. https://doi.org/10.1146/annurev-clinpsy-032511-143109

Orzechowski, J., Gruszka, A., & Michalik, K. (2023). The impact of working memory on divergent thinking flexibility. *Thinking & Reasoning*, *29*(4), 643-662. https://doi.org/10.1080/13546783.2022.2109730

Pasarín-Lavín, T., García, T., Rodríguez, C., Núnez, J. C., & Areces, D. (2024). Divergent thinking and Executive functions in children: A developmental perspective based on intellectual capacity. *Thinking Skills and Creativity*, *51*, 101466. https://doi.org/10.1016/j.tsc.2024.101466

Riley, E., Okabe, H., Germine, L., Wilmer, J., Esterman, M., & DeGutis, J. (2016). Gender differences in sustained attentional control relate to gender inequality across countries. *PloS one*, *11*(11), e0165100. https://doi.org/10.1371/journal.pone.0165100

Sharma, S., & Babu, N. (2017). Interplay between Creativity, Executive Function and Working Memory in Middle-Aged and Older Adults. *Creativity Research Journal*, *29*(1), 71-77. https://doi.org/10.1080/10400419.2017.1263512

Torrance, E. P. (1966). *Torrance tests of creative thinking: Norms-technical manual*. Personnel Press.

Tsai, C. R., Hong, J. C., & Tai, K. H. (2023). Correlates between imagination types and abilities in designing works. *International Journal of Technology and Design Education*, *33*(3), 841-861. https://doi.org/10.1007/s10798-022-09747-0

Zabelina, D. L., Friedman, N. P., & Andrews-Hanna, J. (2019). Unity and diversity of executive functions in creativity. *Consciousness and Cognition*, *68*, 47-56. https://doi.org/10.1016/j.concog.2018.12.005

Zabelina, D., Saporta, A., & Beeman, M. (2016). Flexible or leaky attention in creative people? Distinct patterns of attention for different types of creative thinking. *Memory & Cognition*, *44*, 488-498. https://doi.org/10.3758/s13421-015-0569-4

Quizizz 融入國語領域教學對國小學生的自我效能、遊戲興趣與形成焦慮感之研究

Enhancing Students' Gameplay Interests and Anxiety in Using Quizizz to Gamifying Chinese Contents for Promoting Learning Value of Elementary School Students

李逸銘
國立臺灣師範大學創造力發展碩士在職專班
liyiming0301@gmail.com

【摘要】 學習內容遊戲化能激勵學生學習不同科目。Quizizz 是否能成為國小生學習國語的輔助工具？本研究利用 Quizizz 讓研究參與者輔助國語課堂學習，並調查研究參與者的自我效能、遊戲興趣、遊戲焦慮感和遊戲價值之間關係。本研究對 193 名五年級學生進行單組準實驗研究。然後採用驗證性因素分析和結構方程模型對資料進行分析。結果顯示，學生的自我效能與遊戲興趣呈現正相關；但與遊戲焦慮感呈現負相關；學生的遊戲興趣與遊戲價值呈現正相關，但與遊戲焦慮感呈現負相關。本研究表明，將國語學習內容遊戲化，可以激發學生的積極情緒，消除消極情緒，提高學習價值。

【關鍵字】 自我效能；遊戲興趣；遊戲焦慮感；Quizizz；國語領域

Abstract: *Gamification can motivate students to learn different subjects. Could Quizizz be supportive tool for elementary school students to learn Chinese? This study used Quizizz for participants to learn Chinese, and investigates the relationship between self-efficacy, gameplay interest and, anxiety, and perceived value. Conducted a single-group quasi-experimental study with a total of 193 fifth grade students. Confirmatory factor analysis and structural equation modeling were then used to analyze the data. The result revealed that students' self-efficacy can positively predict their gameplay interest; but negatively related to gameplay anxiety; their perceived value can be positively predicted by students' gameplay interest; but negatively by their gameplay anxiety. This study suggested gamifying Chinese learning contents with Quizizz can activate students positively emotion and deactivate negative emotion and enhance their learning value.*

Keywords: Self-efficacy, Gameplay Interest, Gameplay Anxiety, Quizizz, Chinese Language learning

1. 前言

教育部於 **2022 年**發表「推動中小學數位學習精進方案」，提出「班班有網路，生生用平板」的想法，以期在教育現場帶給學生更生動的教材、更多元的教學、更有效的學習（教育部，2022）。在此背景下，許多教師開始在課堂中融入數位教學，讓學生可以使用平板來輔助學習，其中即時反饋系統（Interactive Response System，簡稱 IRS）受到許多教師的喜愛。IRS 運用在課堂中，具有三大特性：「遊戲性」、「即時性」與「互動性」，不僅可以活絡課堂氣氛，

引起學生在課堂上的積極性，也能協助教師及時釐清學生盲點並調整教學流程（李育勳、歐陽誾，2015）。

本研究中的 Quizizz 遊戲即為 IRS 中的一種。近年來，時常被教師運用於各領域的課堂中。許多研究都顯示，在教學中融入 Quizizz 遊戲，對於學生學習是有很大的助益。Capuno (2023)研究指出，將 Quizizz 遊戲整合到課堂中能增進學生學習興趣並激發其自主學習。Rozi & Rustipa(2023) 研究指出，將 Quizizz 遊戲運用在寫作教學中，可以提升學生的學習興趣。

故本研究目的為:探討將 Quizizz 遊戲融入國語領域的課堂教學中,國小學童的自我效能、遊戲興趣、遊戲焦慮感與遊戲價值的相關。

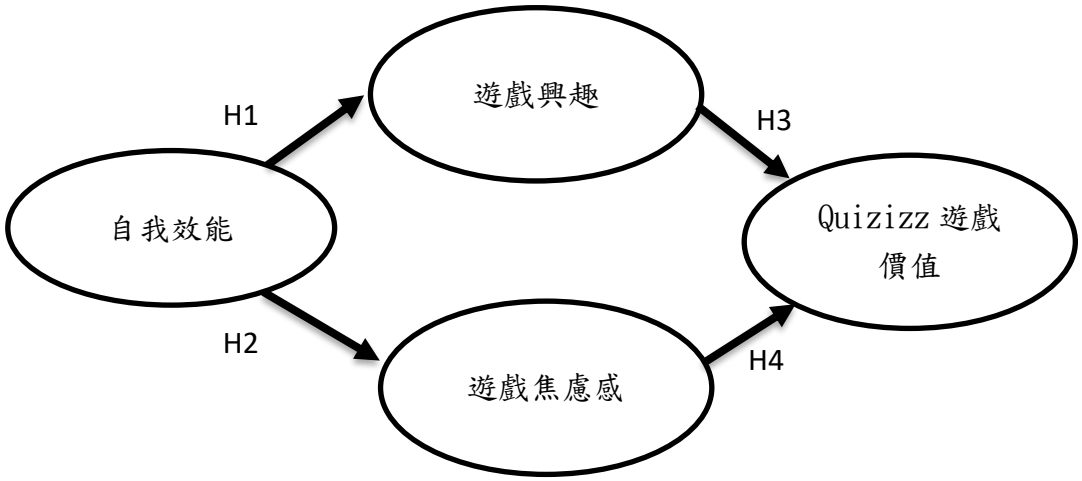
2. 研究方法

2.1. 研究模式與假設

本研究採取單一群組準實驗研究設計（single group quasi-experimental design）與問卷調查法，以此探討在使用 Quizizz 遊戲融入國小國語領域教學，學生的自我效能、遊戲興趣、遊戲焦慮感與 Quizizz 遊戲價值之關聯。並藉由 SPSS 23.0 和 AMOS 20.0 等統計軟體來分析資料。最後，透過資料分析的結果提出結論，以利後續研究者或教師作為參考。本研究亦輔以課程回饋，瞭解學生對出題與遊戲之感受。本研究發展出之研究模式如下圖所示：

圖 1

研究架構圖



H1：國語領域融入 Quizizz 遊戲，學生自我效能對遊戲興趣具有顯著正相關。

H2：國語領域融入 Quizizz 遊戲，學生自我效能對遊戲焦慮感具顯著負相關。

H3：國語領域融入 Quizizz 遊戲，學生的遊戲興趣對 Quizizz 的遊戲價值具有顯著正相關。

H4：國語領域融入 Quizizz 遊戲教學，學生的遊戲焦慮感對 Quizizz 的遊戲價值具有顯著負相關。

2.2 研究參與者

本研究以新北市某公立國民小學為實驗參與學校，該所學校為一中型學校，全校班級數共 59 班，其中五年級共 11 班。本研究採立意抽樣，樣本為全部五年級學生，年齡介於 10-11 歲，學生人數共 260 人，其中同意參與研究的學生人數共 193 人，不同意參與研究的學

生人數共 67 人，共獲得 193 份有效樣本，男性學生 96 位(49.7%)，女性學生 97 位(50.3%)。

2.3 研究工具

本研究所採用的研究教材為翰林國語課本第十冊，研究工具共 2 種，包括「Quizizz 對國小學生提升國語領域的學習興趣與形成焦慮感之調查問卷」、「Quizizz 即時反饋系統」。

2.3.1 Quizizz 對國小學生提升國語領域的學習興趣與形成焦慮感之研究問卷

本研究經由文獻探討以及以研究架構為依據，參考國內外相關之問卷量表，編製 Quizizz 對國小學生提升國語領域的學習興趣與形成焦慮感之調查問卷。編製問卷後，由指導教授和其他專家審查，進行內容效度與表面效度之確認。本問卷內容包含兩個部分，第一部分為「背景資料」；第二部分配合研究主題，含有「自我效能」、「遊戲興趣」、「遊戲焦慮感」、「遊戲價值」等量表。

表 1

本研究問卷各構面之資料衡量尺度與衡量方式表

構面	資料衡量尺度	衡量方式
性別，喜好科目，專長科目	名義變項	單選題
班級	名義變項	填充題
自我效能	比例變項	Likert5 點量表
學習興趣		
焦慮感		
遊戲價值		

表 2

Quizizz 對國小學生提升國語領域的學習興趣與形成焦慮感之調查問卷

基本資料
1. 性別:男/女
2. 班級
3. 喜好課目:國語、英文、數學、自然、社會
4. 專長科目:國語、英文、數學、自然、社會
自我效能
2.使用 Quizizz 時，我覺得只要我努力，就有辦法成功回答問題。
3.使用 Quizizz 時，我有信心能成功解決問題。
4.使用 Quizizz 時，碰到難題，我認為我能想到辦法解決。
6.使用 Quizizz 時，遇到難題，我認為只要我努力就能夠克服。
遊戲興趣
2.我覺得很喜歡 Quizizz 上課的互動方式。
3.我覺得使用 Quizizz 上國語課很有趣。
4.使用 Quizizz 上國語課，我覺得很興奮。
7.使用 Quizizz 上國語課，會讓我很專注。
遊戲焦慮感
4.在 Quizizz 中，我會擔心別人比我先解完題目。

5.在 Quizizz 中，我會因為分心別人而影響表現。
6.在 Quizizz 中，我會擔心剩下的時間太少。
8.在 Quizizz 中，我碰到不會的題目時，我會驚慌得不知所措。
Quizizz 遊戲價值
4.我在使用 Quizizz 上國語課時，會使用生活經歷過的事情推理解題。
5.我在使用 Quizizz 上國語課時，可以學習如何傾聽別人意見，增加大家答對的機率。
6.我在使用 Quizizz 上國語課時，可以學習如何運用小組討論，減少答錯。
7.我在使用 Quizizz 上國語課時，可以學習如何運用道具幫助我解題。

3. 研究結果

3.1 信、效度分析

下表顯示符合以上收斂效度三項標準，各構面之題項 factor loading 值介於.610 至.775 間，皆符合建構效度的檢驗標準(Byrne, 2001)，因此本研究數據具備建構效度。內部一致性信度可以利用組合信度(CR)以及 Cronbach's α 檢驗，各構面之組合信度(CR)須大於.7，Cronbach's α 值以大於.5 為佳(Nunnally, 1978)；本研究數據中構面之 CR 介於.778 至.836 間，Cronbach's α 值介於.697 至.753，皆符合內部一致性信度的標準。

表 3

各構面信度與收斂效度一覽表

構面	M	SD	Cronbach's α	CR	*FL*	AVE
自我效能	3.94	0.91	.733	.834	.746	.557
遊戲興趣	4.07	0.92	.711	.823	.733	.538
遊戲焦慮	2.71	1.29	.753	.844	.758	.576
遊戲價值	3.70	1.05	.697	.816	.725	.527

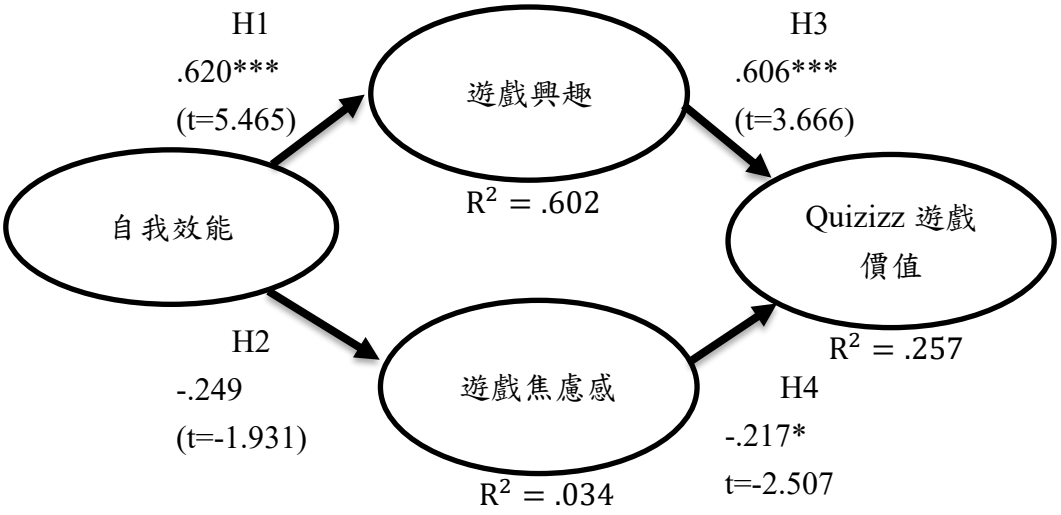
3.2 適配性分析結果

本研究運用 AMOS 20 進行 SEM 分析藉以檢驗模式的適用程度及各別構面的關聯性。透過統計分析得指標如下表，χ^2/df =1.156、RMSEA= .029、GFI= .932、AGFI= .907，NNFI=.972、CFI= .977、IFI= .978，PNFI= .713、PGFI= .685 以上指標皆符合 Hair et al. (2009)建議的規範，雖然相對適配度中的 NFI 和 RFI 應超過 0.9 才能達到良好水準，但 Bentler, P. M. (1990)指出：當研究模型中 NFI=.826 時，研究模型也能達到滿意水準。Baumgartner 與 Homburg (1996)也指出：RFI 須大於. 90 是一個參考點，雖然 RFI 值不到.90 以上，因此在解釋是要特別警慎，但當 RFI 大於. 80 時，許多研究模型仍然具有很高的顯著性。故在本例中，NFI= .855、RFI= .826，雖然這兩個指數略低於標準，但根據這些指數，樣本與研究模型的適合度是可以接受的。

由表 7 與圖 9 可以得知，本研究所設立之 H1、H2、H3、H4 皆成立，使用 Quizizz 遊戲後，學生的自我效能對國語領域的學習興趣具有非常顯著正相關(β=.620，t = 5.465***，p <.001)，使用 Quizizz 遊戲後，學生的自我效能對國語領域的焦慮感無相關，但接近負相關(β=-.249，t = -1.931)，使用 Quizizz 遊戲後，學生對國語領域的學習興趣對 Quizizz 的遊戲價值具有非常

顯著正相關(β=.606，*t*=3.666***，*p*<.001)，使用 Quizizz 遊戲後，學生對國語領域的焦慮感對 Quizizz 的遊戲價值具有顯著負相關(β=-.217，*t*=-2.507* *p* <.01)。此外，國語領域的學習興趣的解釋力為 60.2%，國語領域的焦慮感的解釋力為 3.4%，Quizizz 軟體的遊戲價值的解釋力為 25.7%。

圖 2 模式驗証



4. 討論與建議

本研究旨在探討使用 Quizizz 遊戲融入國語領域教學後，國小學生的自我效能、遊戲興趣、遊戲焦慮感與 Quizizz 軟體的遊戲價值之間的關係，並瞭解 Quizizz 軟體融入國語領域教學後，對學生學習成效之影響。本章節共分為三節，第一節先針對研究結果進行討論，第二節再根據研究分析得到的結果提出研究之貢獻，最後，第三節針對國小教師未來教學以及研究方面提供建議。

4.1 假設驗證

國小學童的自我效能是指，國小學童在國語領域中使用 Quizizz 遊戲學習的自我效能，表示學童有能力去主動去完成測驗，克服困難的信念（Bandura，1986）。國小學童的遊戲興趣是指，國小學童對於在國語領域課堂中融入 Quizizz 遊戲所產生的遊戲興趣，表示學童對 Quizizz 遊戲表現關注、願意積極投入的心理傾向（張春興，2007)。國小學童的遊戲焦慮感是指，學童對於 Quizizz 遊戲缺乏自信、預期自己會無法完成測驗，無法得到好成績(Mayer & Morento, 2003)。Quizizz 遊戲價值是指 Quizizz 遊戲運用在教學有許多的助益。

根據研究結果顯示，國語領域融入 Quizizz 遊戲，學生自我效能對遊戲興趣具有顯著正相關，學生自我效能對遊戲焦慮感無相關，接近負相關，學生的遊戲興趣對 Quizizz 的遊戲價值具有顯著正相關，學生的遊戲焦慮感對 Quizizz 的遊戲價值具有顯著負相關。

根據作者、王志美（2020）遊戲自我效能、遊戲興趣、認知負荷與地理桌遊的遊玩自信心提升之相關研究，研究結果提到國中學生對於地理科遊戲自我效能對於遊戲興趣呈現正相關。楊書旻（2016）基於概念影響關係之遊戲式評量學習模式對國小學童數學學習成效、認知負荷與自我效能之影響，研究結果提到國小學生數學自我效能與數學學習興趣呈現正相關。賴眉潔（2021）即時反饋策略在國小成語教學對高低學習能力學習成效和學習動機之影響，

研究結果提到國小學生成語教學的自我效能與學習興趣呈現正相關。

根據吳慧慈（2023）數位遊戲式學習融入英語教學對偏鄉國小學生學習動機與焦慮感影響之研究：以 Baamboozle 為例，研究結果提到國小學生對於英語領域自我效能與焦慮感呈現負相關。作者、詹瓊華（2018）共變推理遊戲:遊戲自我效能與後設認知影響遊戲中的焦慮，興趣及表現之研究，研究結果提到高一學生的遊戲自我效能與焦慮感呈現負相關。

根據蘇怡璇（2021）即時反饋系統融入國語文教學對國小學童閱讀理解學習成效與閱讀動機之研究—以 Quizizz 為例，研究結果提到國小學生對於國語領域學習興趣與 Quizizz 遊戲價值呈現正相關。黃詩淳（2018）即時反饋系統輔助技術型高中英語教學對學生學習態度與英語字彙學習成就之影響，研究結果提到高中學生對於英語字彙學習興趣與 Quizizz 遊戲價值呈現正相關。

根據蘇怡璇（2021）即時反饋系統融入國語文教學對國小學童閱讀理解學習成效與閱讀動機之研究—以 Quizizz 為例,研究結果提到國小學生對於國語領域的學習焦慮與 Quizizz 遊戲價值呈現負相關。黃詩淳（2018）即時反饋系統輔助技術型高中英語教學對學生學習態度與英語字彙學習成就之影響，研究結果提到高中學生對於英語字彙的學習焦慮與 Quizizz 遊戲價值呈現負相關。

綜上所述，許多研究都與本研究的研究結果相等，顯示本研究獲得支持，證明國語領域融入 Quizizz 遊戲，學生自我效能對遊戲興趣具有顯著正相關，學生的遊戲興趣對 Quizizz 的遊戲價值具有顯著正相關，學生的遊戲焦慮感對 Quizizz 的遊戲價值具有顯著負相關。雖然本研究的研究結果顯示 Quizizz 遊戲融入國語領域教學後，學生的自我效能對遊戲焦慮感無相關，跟許多研究結果有些差異，但其數據已經很接近負相關，推測可能亞洲小孩在負向選項填答時，較為保守，使得遊戲焦慮感的數值較為不明顯。

如同陳慧娟(2015)的研究提到的，即時反饋系統(IRS)能引發學生的學習興趣、改善學習態度，十分適合教師進行多元化的教學。本研究也證實了這一點，Quizizz 遊戲能增進不同學生對於學習的自我效能與學習興趣，因此學生的學習成就也能有所提升。

4.2 實務性建議

在現行對即時反饋系統(IRS)相關的許多研究中，Quizizz 遊戲軟體的研究較為稀少，並且教學領域多為英語領域與數學領域，本研究將研究領域拓展到國語領域，可以進一步的了解 Quizizz 遊戲在國語領域教學中使用，對學生的自我效能、學習興趣與焦慮感的影響。在許多的研究中都顯示在教學中融入即時反饋系統，對於學生學習是有很大的助益(王全世, 2000)，本研究對現場教學的教師提供了不同教學活動的參考。傳統教學模式固然有其方便性，但在現今強調學生自主學習的 108 課綱中，要刺激學生主動思考、找尋答案、自主學習，選擇以遊戲化的方式進行教學活動，或許是一個可以嘗試的方法。

4.3 研究限制與未來研究建議

本研究的研究樣本只從研究者任職的學校蒐集，導致樣本數量不足。未來研究者可以多選擇幾間國小進行研究樣本的蒐集，獲取更多的樣本，研究也能更有代表性。以「有」、「無」使用數位遊戲平台為分組，進行相關研究

本研究結果顯示在使用 Quizizz 遊戲融入國語領域教學後，學生的自我效能與遊戲興趣有正相關的影響。未來研究者可以以此參考，進行有使用經驗與無使用經驗之相關研究，了解學生的自我效能、焦慮感與學習興趣在這兩種情況之下所發生之影響。

5. 結論

本研究結果顯示，國語領域融入 Quizizz 遊戲，學生自我效能對遊戲興趣具有顯著正相關，但與遊戲焦慮感無相關；學生的遊戲興趣對 Quizizz 的遊戲價值具有顯著正相關，但遊戲焦慮感對 Quizizz 的遊戲價值具有負相關。透過文獻探討可以發現許多研究都與本研究的研究結果相等，顯示本研究獲得支持。綜上所述，本研究表明，將國語學習內容遊戲化，可以激發學生的積極情緒，消除消極情緒，提高學習價值。

參考文獻

中文部分：

王全世（2000）。資訊科技融入教學之意義與內涵。**資訊與教育，80**，23-31。

吳慧慈（2023）。**數位遊戲式學習融入英語教學對偏鄉國小學生學習動機與焦慮感影響之研究：以 Baamboozle 為例**（未出版碩士論文）。國立臺中教育大學。

李育勳（2014）。**即時反饋系統應用於國小三年級學生修辭教學之行動研究**（未出版碩士論文）。國立臺南大學。

李育勳、歐陽誾（2015）。即時反饋系統應用於國小三年級學生修辭教學之行動研究。**師資培育與教師專業發展期刊，8**（3），77-101。
https://doi.org/10.3966/207136492015 1208030

洪榮昭、王志美（2020）。遊戲自我效能、遊戲興趣、認知負荷與地理桌遊的遊玩自信心提升之相關研究。**教育科學研究期刊，65**（3），225-250。
https://doi.org/10.6209/JORIES.202009_65(3).0008

洪榮昭、詹瓊華（2018）。共變推理遊戲:遊戲自我效能與後設認知影響遊戲中的焦慮，興趣及表現之研究。**教育科學研究期刊，63**（3），131-162。
https://doi.org/10.6209/JORIES.201809_63(3).0005

張春興（2006）。**張氏心理學辭典**。臺北市：臺灣東華。

張春興（2007）。**教育心理學－三化取向的理論與實踐**（重修二版）。臺北市：東華。
https://doi.org/10.6251/BEP.19930701.1

陳慧娟（2015）。**IRS 即時反饋系統輔助教學對社會科學習成效影響之研究-以國小六年級學生為例**（未出版碩士論文）。萬能科技大學。

黃詩淳（2018）。**即時反饋系統輔助技術型高中英語教學對學生學習態度與英語字彙學習成就之影響**（未出版碩士論文）。國立臺北科技大學。

楊書旻（2016）。**基於概念影響關係之遊戲式評量學習模式對國小學童數學學習成效、認知負荷與自我效能之影響**（未出版碩士論文）。東吳大學。

賴眉潔（2021）。**即時反饋策略在國小成語教學對高低學習能力學習成效和學習動機之影響**（未出版碩士論文）。國立臺中教育大學。

蘇怡璇（2021）。**即時反饋系統融入國語文教學對國小學童閱讀理解學習成效與閱讀動機之研究—以 Quizizz 為例**（未出版碩士論文）。中原大學。

英文部分：

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ*, *1986*(23-28), 2. https://doi.org/10.4135/9781446221129.n6

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing*, *13*(2), 139-161. https://doi.org/10.1016/0167-8116(95)00038-0

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588. https://doi.org/10.1037/0033-2909.88.3.588

Capuno, J. G. C. (2023). Quizziz: A Game-based Formative Assessment Tool for Enhancing Students Self-Regulated Learning. *International Journal of Social Learning (IJSL)*, *3*(3), 329-340.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43-52. https://doi.org/10.1207/S15326985EP3801_6

Maming, K., Sani, A. A., Putra, D. P., & Radiana, S. P. (2023). Get Ahead with Quizizz: Advancing Junior High School Students' Vocabulary Mastery in Online Learning. *Elsya: Journal of English Language Studies*, *5*(2), 224-235.

Rozi, F., & Rustipa, K. (2023). The Use of Quizizz to Develop Instructional PJBL Practices in Teaching Writing. *English Education Journal*, *13*(1), 128-139. https://doi.org/ 10.5539/ijel.v7n1p126

以 Picasso 互動數位遊戲學習探討性別對意境想像力、繪圖認知負荷、情境命名認知負荷與趣味性價值之差異

Exploring Gender Differences in scenario Imagination, Game Cognitive Load, Naming Cognitive Load, and Enjoyment Value through Picasso Interactive Digital Game-Based Learning

宋紀萱
國立臺灣師範大學教育學院創造力發展碩士在職專班
hsuan0956@gmail.com

【摘要】 本研究目的在探討性別在互動數位遊戲學習的成效差異，研究者使用是以繪圖呈現的數位遊戲 APP，採用接龍的方式完成各別圖畫，並在限定時間內嘗試最大命名數量，最後評估意境想像力、繪圖認知負荷、情境命名認知負荷、趣味性價值之性別比較向度。研究方法採問卷方式，以獨立樣本 t 檢定男性與女性之學習成效差異。結果顯示：意境想像力、繪圖認知負荷與趣味性價值方面，男性與女性無顯著差異。情境命名認知負荷方面：女性圖畫命名認知負荷小於男性，有顯著差異。根據研究結果，本研究設計之繪圖聯想命名 App 是結合想像力與語言聯想的方式學習。

【关键词】 數位遊戲學習；意境想像力；認知負荷；趣味性價值

Abstract: *Imagination and creativity are essential for students to be developed in the era of AI. Accordingly, this study utilized an app, named Picasso developed by Digital game-based learning of Taiwan Normal University. The App centered around drawing for players to collaboratively image those picture they generated and image scenarios within a limited time. The study evaluated gender differences in aesthetic imagination, game cognitive load, imaging cognitive load, and hedonic value. The results indicated: No significant gender differences in psychological scenario imagination ,imaging interest and hedonic value, Significant gender differences in imaging cognitive load, the level of cognitive load for females was lower than male. To advance the statistical analysis, it is need to increase sample size in the future study.*

Keywords: *Digital game-based learning, Aesthetic imagination, Cognitive load, hedonic value, imagination, creativity*

1. 前言

想像力是一種從具體到新的具體的轉化和創造活動。從既定的具體事物走向創造的具體事物，只有藉助抽象才能實現創造性建構的可行性(Vygotsky, 1998, p. 163)。根據心理研究的觀點，兒童很容易並經常在他們的想像中表現出不可能甚至不可能的結果；想想兒童讀物、電視節目和電影中大量出現的幻想現象吧(Lane et al., 2016)。Vygotsky (2004)將遊戲概念化為兒童創造和經歷的想像情境的體驗。我們從這些研究中了解到，想像遊戲可以作為正式環境中教學和學習的起點，帶來動態的學習體驗，並提供獲得抽象和象徵性概念的途徑。兒童從出生到八歲的心理發展過程中，從嬰兒期到學齡前期末，他們的想像力都在不斷發展（Vygotsky, 1998）。與普遍的看法相反，想像力的發展並沒有消失，而是在整個童年時期都在發展(Perone & Goncu, 2014)，儘管在學校裡並不總是能看到(Fleer, 2023)。

重點特別放在以教學實踐為導向的教學模式上，這些模式主要是培養學童的探究或好奇心並激勵想像力(Vygotsky, 2004)。然而，研究顯示，遊戲是學齡前兒童發展想像力的最佳形式，這種發展是透過學齡前兒童在情境中想像，並利用想像力延伸視覺領域以賦予其新的意義而顯現出來的(Fleer, 2023)。兒童的想像遊戲以多種方式延伸、轉換和擴展了兒童的現實世界。此外，在教育環境中，集體想像被認為是兒童學習和社會活動的關鍵方式。集體想像在這裡被定義為幼兒意識到、加入、參與同情境中進行並促進想像情境的活動。集體想像可以讓兒童共享一個抽象的智力空間(Fleer et al., 2021)。在這個空間裡，兒童創造和分享新的意義和理解，並向他人介紹想像及想法和討論，如此可再重塑想像情境的意義;然後發展及匯聚與共同想像情境。集體想像被認為是一種高級、更發達、更成熟的想像形式，並維持想像活動的動機(Fragkiadaki et al., 2021)。因此，本研究改編了一個名為 「畢卡索」的應用程序，讓兒童旋轉繪畫，並通過用形容詞和名詞為場景命名來想像他們完成的畫面，例如，黑暗的街道、移動的小船等。洪家萱(2002)同一學習內容會有個體本身產生的內在認知負荷，同時也會有因為不量的教學設計而產生的外在認知負荷，然而此兩者是累加的。若教學者能夠提供適當的教材、學習內容的呈現方式，便能使學習者更加投入其中，更專注於學習內容上，雖然過程中會增加個體的認知負荷，但亦能促進認知基礎模型的建立，有助於學習者進行學習。根據這項活動，本研究觀察了學生對性別差異的認知和情感過程。

2. 研究方法

2.1. 研究假設

Kurt Squire 和 Henry Jenkins 曾在 Video Games and the Future of Learning 提到數位遊戲的學習原則、身分建構、社群合作，以及樂趣和動機如何影響物來學習的發展。認知控制過程是指負責處理資訊、調節想法和行動以適應特定目標或任務的所有功能。一些研究強調了分析認知控制過程與情緒過程的相關性（Ramírez et al.,2023） 認知能力的性別差異代表了男孩、女孩和不同性別青年之間的生物過程和社會化過程的協同作用（Kheloui et al.,2023) 參與者的性別可能與情緒處理有關，因為男性和女性的表現取決於所呈現的刺激的價值（Bian et al.,2017）。因此, 本研究為比較研究，研究目的在探究性別與數位遊戲學習之關係，以數位遊戲融入美術學習，最後透過問卷了解學童於意境想像力、繪圖認知負荷、情境命名認知負荷、趣味性價值的表現差異情形。重視反饋歷程的誘惑會引發學生個人興趣，提升自我成就感，增加學生學習興趣及動機（尤怡婷，2021）。

本研究擬檢驗之假設如下：

假設一：不同性別在意境想像力具有差異。

假設二：不同性別在繪圖認知負荷具有差異。

假設三：不同性別在情境命名認知負荷具有差異。

假設四：不同性別在趣味性價值具有差異。

2.2. 研究工具: Picasso App 遊戲

本研究採用由國立臺灣師範大學數位學習實驗研究室所開發之「Picasso」 遊戲 App 來進行。Picasso App 是一款以互動接龍繪圖的方式提升想像力的遊戲 App，透過參與者之間的互動刺激其想法與學習，並同時也具有跨領域融合語言應用之特色，其中能設定不同的人數、繪圖回合、限制時間、命名次數，使教學者能依照學習者的程度，循序漸進的創造不同的遊戲房間。除了遊玩模式，學習者也能在遊戲結束後欣賞來自不同遊戲者的作品。Picasso App 遊戲最大的特色是完成的圖畫能使用「放大」、「縮小」、「旋轉」等功能引導學習者從不同的角度思考。

2.3. 問卷工具

本研究屬量化研究，透過 Picasso App 互動數位遊戲學習後實施問卷調查，進行資料蒐集，針對本次研究內容進行延伸、修改來符合研究要件，內容包含四部分，並經由專家學者審查編修，以確認問卷內容的適當性及語詞使用的流暢度。意境想像力: 我喜歡以觸感感受週遭的事物；繪圖認知負荷: 我覺得想加一條線或塗色，要花很大的心力才能完成；情境命名認知負荷: 我覺要花很大的心力才能想出情境；趣味性價值:玩 Picasso App 我覺得很開心。

2.3.1. 意境想像力特質問卷

意境想像力指的是想像力的一般性特質，想像力為可能性的思考，而非事實性的思考，其思考事物可能是什麼，而非現存事實應如何（White, 1990），是形成意象、知覺和概念的能力，也是創造力中重要的特質之一。本研究的意境想像力是指進行「Picasso App」的過程中，學習者想像力之性別差異。本研究共設計 6 題。題項內容如表 1：

表 1

使用「Picasso App」學習者之「意境想像力」問卷題項

題號	選項內容
I1	我喜歡隨手塗鴉和繪畫表達對環境的感受。
I2	我喜歡感受大自然巧妙的變化。
I3	我喜歡想像物品不同尺寸、顏色等變化的樣子。
I4	我喜歡以觸感感受週遭的物品。
I5	我喜歡一面聽音樂一面想事情。
I6	我喜歡以文字述說的方式描繪內心世界。

2.3.2. 遊戲認知負荷問卷

「認知負荷」係指個體在執行某種工作、作業或任務的過程中，個體感受到心智負荷與心智努力的負載狀態（陳密桃，2003）。本研究的遊戲認知負荷是指進行「Picasso App」的過程中，學習者在遊戲中所產生的認知負荷之性別差異。本研究共設計 6 題，如表 2：：

表 2

使用「Picasso App」學習者之「遊戲認知負荷」問卷題項

題號	選項內容
G1	我覺得我花了極大的心力，才能跟上遊戲的進行。
G2	我覺得我花了極大的心力，才能掌握遊戲的內容。
G3	我必須付出極大的注意力，才能適應這個遊戲的方式 。
G4	我覺得進行這個遊戲，讓我感受到不知所措。
G5	我覺得我需要花很多腦筋來畫出遊戲中的圖案。
G6	我覺得進行這項活動時，我心理壓力很大 。

2.3.3. 命名認知負荷問卷

認知負荷是「多方位」的結構，是將一個特別的任務，加在學習者的認知系統上。而構成學習者認知負荷的本質，乃是學者在心智努力上所耗費的「努力」強度，越需要努力的，認知負荷就越大（Pass,1992）。本研究的命名認知負荷是指進行「Picasso App」的過程中，學習者在命名階段所產生的認知負荷之性別差異。本研究共設計 6 題。題項內容如表 3：

表 3

使用「Picasso App」學習者之「命名認知負荷」問卷題項

題號	選項內容
N1	我覺得我花了極大的注意力，才能完成此命名活動。
N2	我覺得進行命名活動，讓我感受到極大的挫折感。

題號	選項內容
N3	我覺得命名活動對我來說非常困難。
N4	我覺得我需要花很多腦筋來進行命名活動 。
N5	我覺得進行這項命名活動時，我心理壓力很大。
N6	遊戲限定的時間讓我心中感受到極大的壓力。

2.3.4. 趣味性價值問卷

趣味性被認定為主動性內在價值，趣味性的交易行為是內在反應的快樂，來自引人入勝的活動所產生的一種內在感受（Unger and Kernan, 1983；Huizinga, 1995）。本研究的趣味性價值在探討進行「Picasso App」的過程中，學習者所產生的趣味性價值之性別差異。本研究共設計 6 題。題項內容如表 4：

表 4

使用「Picasso App」學習者之「趣味性價值」問卷題項

題號	選項內容
T1	我認為使用數位遊戲學習非常有趣。
T2	我認為使用數位遊戲學習會讓我很投入。
T3	我認為使用數位遊戲學習會讓我感覺很興奮。
T4	我認為使用數位遊戲學習會讓我感覺時間過得特別快。
T5	我認為使用數位遊戲學習會讓我忘掉別的事。
T6	我很喜歡這次數位遊戲的互動方式。

2.4. 研究程序與參與者

Picasso App 是一款以互動接龍繪圖的方式提升想像力的遊戲 App，透過參與者之間的互動刺激其想法與學習，並同時也具有跨領域融合語言應用之特色，其中能設定不同的人數、繪圖回合、限制時間、命名次數，使教學者能依照學習者的程度，循序漸進的創造不同的遊戲房間。此遊戲為多人互動模式，以學生的班級、座號設定帳號密碼，登入後能記錄每次遊戲的繪圖步驟及命名資料。Picasso App 遊戲過程總共分為三階段，第一階段：繪圖、第二階段：圖畫命名、第三階段：故事聯想，本研究著重繪圖階段及圖畫命名階段，本次遊戲房間設定：人數 3-5 人、繪製線條 3 回合、上色 2 回合、每張圖命名 3 次、每回合繪圖時間限制 60 秒、命名總時長限制 7 分鐘。繪圖階段如圖 1，為先繪製線條再進行上色，玩家首次繪製線條必須通過中心點且只能畫一筆，完成後，圖畫交換規則為 1 號玩家傳給 2 號，2 號傳給 3 號以此類推最後一位傳給 1 號玩家，第二回合繪製線條需至少與上一位玩家的線條產生 1 個交點，完成後圖畫再做交換，第三回合繪製線條需至少與前兩位玩家的線條各產生 1 個交點，線條回合結束後進入上色回合，上色 2 回合能透過改變顏色、粗細、筆刷進行變化，每一回合圖畫皆會進行交換，回合結束後則進入第二階段圖畫命名如圖 2。此階段玩家可以使用「放大」、「縮小」、「旋轉」等功能將觀察圖畫並依照「形容詞＋名詞」的規則命名，在限定時間內進行最大數量的命名，每張圖畫最多命名 3 次。本研究以 Google 表單製作網路問卷，於數位遊戲學習後發問卷，來施測 8-9 歲之國小三年級學生。本研究為臺北市某國小（市區中型學校）三年級學生為施測對象，學童年齡為 7、8 歲，扣除無效問卷後得 54 份有效問卷。有效問卷中，男性 21 位，佔總參與者 38.9%，女性 33 位，佔總參與者 61.1%。

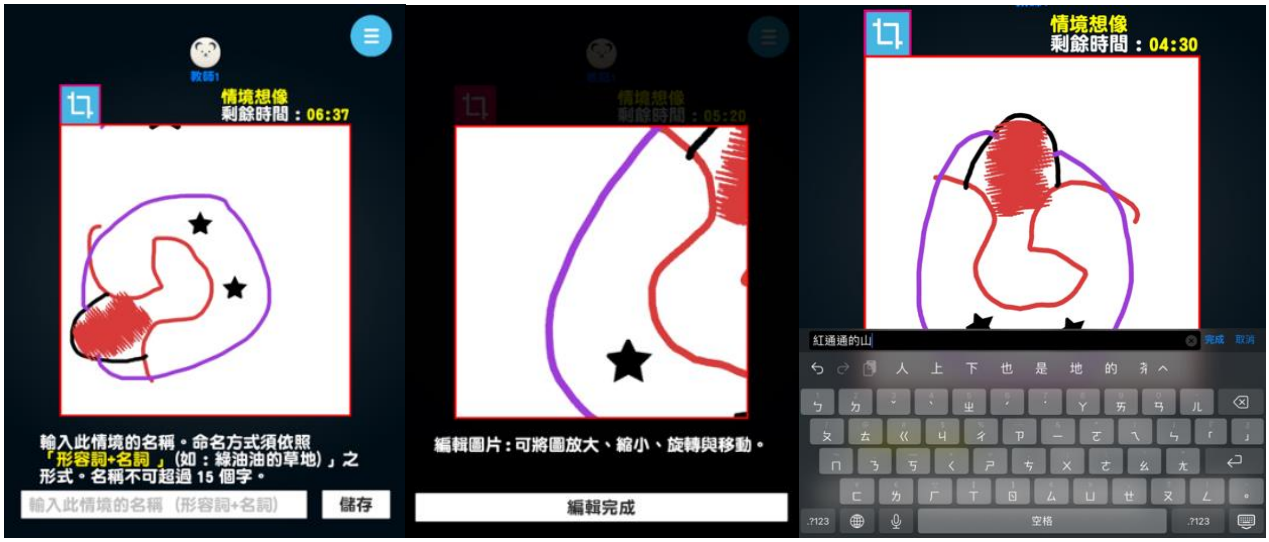
圖 1

第一階段繪圖



圖 2

第二階段命名



3. 研究結果

3.1. 構面信效度分析

本研究有 4 個構面有量表題項，經 Cronbach's alpha 檢驗後刪除不良試題，從表 1 可以得知本研究 4 個構面的 CR 值 介於 .80~ .92，皆大於 .7 ，由顯示可知本研究模式內部有良好的一致性，皆符合 Hair 等人 (2019) 建議 CR 值應該超過 .7 的標準。

本研究透過 Cronbach’s α 及組合信度(composite reliability, CR)檢驗量表的信度，Tabbakh 與 FreelandGraves （2016）指出 Cronbach’s α 及組合信度值若高於 .60 則視為可接受標準。從表 5 中可以得知本研究經檢驗，Cronbach’s α 值介於 .61~.89，皆符合建議標準。本研究進行收斂效度檢測，檢驗收斂效度需透過因素負荷量(factor loading, FL)和平均萃取變異量(average variance extracted, AVE)以檢驗兩數值是否皆通過標準。DeVellis 與 Thorpe (2021)亦認為因素負荷量及 AVE 若大於 0.6 表示該構面具備適合的收斂效度,從表 5 中可以得知 FL 及 AVE，皆符合建議標準。

表 5

各構面信度分析彙整表

構面	Cronbach's α	CR	FL	AVE
意境想像力	.61	.80	.75	.57
繪圖認知負荷	.84	.89	.78	.61
情境命名認知負荷	.89	.92	.85	.70
趣味性價值	.77	.86	.80	.60

3.2. 差異性分析

如表 6 所示學習者於意境想像力之構面問卷，男女性別的平均數差異的考驗結果未達顯著水準（ t = 1.46 , p > . 0 1 ），表示接受虛無假設，即受試者之性別在意境想像力無顯著差異。學習者於繪圖認知負荷之構面問卷，男女性別的平均數差異的考驗結果未達顯著水準（ t = 0.21 , p > . 0 1 ），表示接受虛無假設，即受試者之性別在繪圖認知負荷無顯著差異。學習者於情境命名認知負荷之構面問卷，男女性別的平均數差異的考驗結果達顯著水準（ t = 2.12 *, p < . 0 1 ），表示拒絕虛無假設，即受試者之性別在繪圖認知負荷達顯著差異，表示男性學生的情境命名認知負荷明顯大於女性學生。學習者於趣味性價值之構面問卷，男女性別的平均數差異的考驗結果未達顯著水準（ t = 1.18 , p > . 0 1 ），表示接受虛無假設，即受試者之性別在趣味性價值無顯著差異。

表 6

性別獨立樣本 t 檢定分析

構面	性別	樣本數 (n)	平均數 (M)	標準差 (SD)	自由度 (df)	t 值 (t)	p 值 (p)
意境想像力	男	21	3.73	1.00	52.00	1.46	.15
	女	33	4.07	.72			
繪圖認知負荷	男	21	2.67	1.20	52.00	.21	.84
	女	33	2.61	.93			
情境命名認知負荷	男	21	2.77	1.28	29.98	2.12	.04*
	女	33	2.11	.80			
趣味性價值	男	21	4.30	.72	52.00	1.18	.25
	女	33	4.00	.97			

4. 結論

本研究目的主要探究透過「Picasso」App 互動數位遊戲學習，探討國小三年級學童在意境想像力、繪圖認知負荷、情境命名認知負荷、趣味性價值上，不同性別的表現是否會受到影響而呈現差異。研究結果顯示：（一）性別與意境想像力未達顯著差異：假設一與結果不一致；（二）性別與繪圖認知負荷未達顯著差異：假設二與結果不一致；（三）性別與情境命名認知負荷達顯著差異：假設三獲得支持；（四）性別與趣味性價值未達顯著差異：假設四與結果不一致。命名與語言學習能力相關，最近的研究數據表明，在不同的語言成績衡量標準上，女性學習者的得分往往高於男性學習者(Iwaniec, J, 2019)，女孩比男孩更擅長語言學習 (Chavez, 2000)，與相關研究相符。

數位遊戲學習應用於教育的相關研究日趨增加，然於美術領域的比較研究較少。在數位遊戲融入教學與學習相關研究中，也少有以性別來進行比較的研究。是以，本研究可作為數

位遊戲學習認知與情意理論應用於美術領域的參照，數位遊戲學習之性別差異亦可應用於教學與學習參考。

美術教學知識廣博，還須兼顧學習者的情意與技能，於實務現場，學習者的想像力扮演著重要角色，但想像力特質該如何去評量一直是教學者的一大難題，且缺乏具體的教學引導作法，教學者之想法與作為，深深影響學習者的理解與感受。數位遊戲已經成為新興學習利器，亦成為教學改革之方向。本研究的教學活動執行完成後，給許多老師帶來嶄新的教學方向與體驗，且產生熱烈迴響與正面回饋。

本研究之數位遊戲學習工具 Picasso App，除了包含美術領域的繪圖部分外，還能延續圖畫進行命名及故事創作，但考量受試者年紀因素，本次研究只進行至第二階段的命名活動較為可惜，第三階段的故事創作可以作為評量 Bloom 的認知領域六層次中的「創造」，結合語文進行跨領域的學習，於語文能力影響想像力的學習表現，此議題也可做為未來研究方向的建議。

參考文獻

洪家萱(2022)。以認知負荷的端點探討情境佈題對異分母數加法學習之影響(未出版碩士論文)。國立臺中教育大學

尤怡婷. (2021). 以 [WordWall] 數位遊戲探討印尼初學者在華語文學習焦慮, 遊戲學習興趣, 遊戲焦慮與知覺遊戲學習的價值. 臺灣師範大學華語文教學系海外華語師資數位碩士在職專班學位論文, 2021, 1-120.

Berk, R. A. (1996). Student ratings of 10 strategies for using humor in college teaching. Journal on Excellence in College Teaching, 7(3), 71-92.

Bian, L., Leslie, S.J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. Science, 355 (6323), 389–391. https://doi.org/10.1126/science.aah6524.

Chavez, M. (2000). Gender in the language classroom. London: McGraw Hill.

DeVellis, R. F., & Thorpe, C. T. (2021). Scale development: Theory and applications. Sage publications.

Draper, S. W. (1999). Analysing fun as a candidate software requirement. Personal Technologies, 3, 117-122.

Fleer, M. (2023). The role of imagination in science education in the early years under the conditions of a Conceptual PlayWorld. Learning, Culture and Social Interaction, 42, 100753. https://doi.org/10.1016/j.lcsi.2023.100753

Fleer, M., Fragkiadaki, G., & Rai, P. (2021). Collective imagination as a source of professional practice change: A cultural-historical study of early childhood teacher professional development in the motivated conditions of a Conceptual PlayWorld. Teaching and Teacher Education, 106, 103455

Fragkiadaki, G., Fleer, M., & Rai, P. (2021). The social and cultural genesis of collective imagination during infancy. Learning, Culture and Social Interaction, 29, 100518. https://doi.org/10.1016/j.lcsi.2021.100518

Garner, R. L. (2006). Humor in pedagogy: How ha-ha can lead to aha!. College teaching 53-54, 177-180.

Iwaniec, J. (2019). Language learning motivation and gender: The case of Poland. International Journal of Applied Linguistics, 29(1), 130-143.

Kheloui, S., Jacmin-Park, S., Larocque, O., Kerr, P., Rossi, M., Cartier, L., & Juster, R. P. (2023). Sex/gender differences in cognitive abilities. Neuroscience and Biobehavioral Reviews, 152, 105333. https://doi.org/10.1016/j.neubiorev.2023.105333

Lane, J. D., Ronfard, S., Francioli, S. P., & Harris, P. L. (2016). Children's imagination and belief: Prone to flights of fancy or grounded in reality? Cognition, 152, 127–140. http://dx.doi.org/10.1016/j.cognition.2016.03.022

McManus, I. C., & Furnham, A. (2010). "Fun, fun, fun": Types of fun, attitudes to fun, and their relation to personality and biographical factors. Psychology, 1(03), 159.

Ramírez, V. A., Lipina, S. J., & Ruetti, E. (2023). Cognitive and emotional processing in tasks with emotional valence: Analysis of age and gender role on child development variations. Trends in Neuroscience and Education, 33, 100212. https://doi.org/10.1016/j.tine.2023.100212

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. Phi Delta Kappan, 87(2), 105-111. https://doi.org/10.1177/003172170508700205

Squire, K. (2006). From content to context: Videogames as designed experience. Educational Researcher, 35(8), 19-29. https://doi.org/10.3102/0013189X035008019

Vygotsky, L. S. (1998). The collected works of L.S. Vygotsky. In T. M. J. Hall, & R. W. Rieber (Eds.), Vol. 5. Child psychology. New York: Kluwer Academic and Plenum Publishers (Ed. English translation).

Vygotsky, L. S. (2004). Imagination and creativity in childhood. Journal of Russian and East European Psychology, 42(1), 7–97.

Research on Immersive Virtual Experimental Environment based on Visual Emotional Design

Liqiao Nong[1], Lanlan Zhang[2], Geping Liu[3]*
[1]Guangxi Polytechnic of Construction
[2]Kunming No. 5 Middle School
[3]Southwest University
* Corresponding author: qiaoshining2021@foxmail.com

Abstract: *Virtual reality technology applied in the field of education is mainly used in virtual classroom, virtual experiment and role playing teaching nowadays. However, there are still some problems in the practical application of virtual reality technology in the teaching of digital courses, such as paying too much attention to knowledge and skills learning, ignoring students' emotional experience, and causing high cognitive load of learners. With the theoretical support of visual cognitive theory and cognitive emotion theory of multimedia learning, this study conducted correlation analysis of visual elements and corresponding emotional states in virtual environments, analyzed the visual cognitive mechanism, and explored the intersection point with the characteristics of learners in virtual experiments.*

Keywords: virtual reality, immersive virtual experiment, visual cognition

1. Introduction

The application of virtual reality in the field of education is gradually widespread, and learners are faced with a wide variety of learning software, and their expectations for products are constantly changing and improving. The current design gradually tends to stimulate users, improve life and trigger emotions (Jensen, 2021). Mayer's (2005) cognitive theory of multimedia learning focuses on the cognitive characteristics of learners in the process of multimedia learning. With the deepening of research and the needs of learners, researchers have realized that they should not only pay attention to the cognitive characteristics of multimedia learning, but also pay attention to factors such as emotion and cognitive load. Moreno proposed the cognitive emotion theory of multimedia learning, which takes into account the cognitive effects of emotional factors on learners (Lu Jiamei, 2016). It is believed that positive emotion can promote learners' cognition of knowledge. In the traditional classroom, students' emotional communication always exists with the progress of the class, which is closely related to the teacher's teaching style and language expression. When learners learn in a virtual environment, the features of the virtual environment, such as color and spatial layout, will affect learners' emotional state. However, current virtual experiments simply present the real world, ignoring students' emotional experience (Erez, & Isen, 2022). Therefore, more attention should be paid to the emotional module in the design of virtual scenes, such as stimulating learners' senses, stimulating emotions and strengthening their learning motivation through visual effects. Therefore, we should pay attention to learners' emotional experience in the process of learning knowledge and skills. This study intends to explore an optimal design scheme to stimulate learners' emotions in a virtual learning environment. The core issue of this study is how to explore the optimal virtual environment design scheme to stimulate learners' emotions and improve learners' emotional experience under the guidance of visual cognitive theory.

2. Literature Review

2.1Review of research on virtual experiment

The developed countries have a long history of virtual experiment research. The concept of virtual experiment was first proposed by experts and scholars in the United States. Through the analysis and sorting of previous literatures, it was found that it could be classified into four aspects: design, development and implementation, teaching application practice, investigation and evaluation, according to the development life cycle of virtual experiment. Foreign countries have gradually combined evaluation and design, forming an iterative design, that is, the process of design, then use and evaluation, and then design again. For example, Potkonjak et al. believe that virtual experiments provide students with a better learning environment than physical experiments, and can provide students with effective teaching to acquire knowledge skills and practical experience (Li Hongxia, 2020). Corter and Nickerson both pointed out in their research that virtual experiment plays an important role in subject curriculum teaching, which can deepen students' understanding of teaching curriculum concepts and experimental operation skills (Diana Ibet Roman-Sanchez, et al., 2018). The above researchers highly recognized the application of virtual experiments in teaching, and analyzed the relevant data through several empirical studies, and reached a conclusion to further support their views, which is relatively convincing. A large number of studies have been done on the application of virtual experiments in teaching abroad, including research objects in higher education and primary education, which confirm the value of virtual experiments in the field of education application.

2.2 Review of research on emotional design

The research on emotional design in multimedia learning environment has achieved a lot of results. For example, learning effect can be enhanced by enhancing learners' learning interest and motivation, that is, emotion can act as a learning facilitator. It can be seen that color can affect people's emotions, cognitive processes and learning effects. In view of the different views on the influence of emotion on learning effect. Um et al. (2022) conducted a study on the influence between positive emotion and multimedia learning effect on how to design virtual environment to arouse learners' emotions based on virtual world as a platform, and proved that positive emotion can promote learners' emotions. Their research also shows that emotional design is the first step in the field of multimedia learning, and the positive effect of emotional design of multimedia learning materials on learning outcomes. The positive emotions generated by learners in the learning process can be used as clues to extract information from long-term memory, and learners will show more freedom in cognitive processing when they are in positive emotional states. In the process of problem solving, positive emotions can promote creative thinking, divergent thinking and achievement. Emotional states not only affect information processing, but also induce and regulate learners' learning interest and motivation (Erez, & Isen, 2022). Therefore, visual and emotional design of learning materials may not induce the emotions of learners, but it will negatively affect the learning effect, which indicates that learners do not have to produce positive emotions in the learning process to have an effect on the learning effect. A well-designed learning material has a

significant effect on stimulating the emotions of learners. Further promote learners' understanding of knowledge and improve learning motivation.

3. Visual cognitive model of virtual experiment based on emotional design

3.1Visual characteristics of learners in virtual experiments

In people's visual perception of space, some similar visual elements will be classified into a group, and the similarity judgment is mainly based on the color, shape and scene layout of the current virtual space. Therefore, virtual experiment learners mainly follow the following three principles in terms of visual features: the principle of proximity is that an element in virtual space is regarded as a whole by its adjacent elements; the principle of similarity is that elements with similar properties in virtual space are regarded as a whole. Continuity principle: It means that human beings will perceive similar objects continuously when they recognize external things. Therefore, the visual characteristics of virtual experiment learners are holistic.

Learners' previous learning experience will have certain psychological expectations for the things they have never come into contact with, such as virtual experiment environment. When learners enter the virtual laboratory, they will automatically study the entire environment layout of the virtual laboratory and how to control their own motion information in the virtual environment.

3.2Cognitive characteristics of learners in virtual experiment

Cognition is the internal process and result of human brain's cognition of things, including feeling, perception, memory, thinking and other psychological processes (Liu Yaru, 2018). According to the rules of psychologist Piaget's division of children's development period, the virtual experiment learners in this paper belong to the formal operation stage, and the learners in this stage have strong perception and observation, rich emotional experience, and different ways of thinking and aesthetic perception. Sensory perception mainly includes vision, hearing, space and time perception, etc. In the aspect of the emotional, with the growth of age, college students' sense of self-experience is gradually enriched. In the aspect of the way of thinking, at this stage, learners in a very rigorous logical thinking stage, and the independent logic of thinking is mainly reflected in the strong thirst for knowledge, and they will try their best to obtain knowledge in the field of their interest. Creative thinking is mainly the novelty of the viewpoint, the divergence of thinking, flexibility, uniqueness, abruptness, the value of the result and other characteristics (Nurdyani F., 2018). The cultivation of creative thinking in virtual experiment learning is inseparable from the relaxed and comfortable environment. In the learning process, learners need to be inspired to gradually cultivate their creative thinking. In the aspect of the aesthetic, in daily life, everyone's aesthetic view is not the same, different personalities, different cultivation of people, their aesthetic appreciation ability is different.

3.3Visual cognitive process of virtual experiment learners

In the cognitive process, people first summarize the stimulus information according to visual cognitive habits, then make targeted groups for the modeling, and compare with the information in the memory bank, and finally make information processing (Zhu Shiyu, 2018). Visual cognitive habits are related to learners' original visual selective attention and designers' visual guidance in the

virtual environment. Learners' cognition of the virtual laboratory starts from surface features (color, spatial layout, target object location), and the received information stimuli combine with learners' cognitive habits to form emotional stimuli. Colin Ware proposed a three-stage model of visual perception (Ware C., 2012), which shows three stages of visual perception by users after receiving visual information. The first stage is that the environmental information is repaired and its basic features are extracted. In the second stage, the active pattern perception process extracts the visual structure and divides the visual scene into regions with different colors, textures and motion patterns. The third stage is to reduce information to a small number of objects in visual working memory through active attention mechanisms, forming the basis of visual thinking (Shen Zhuqi, 2019).

3.4Visual cognitive model of virtual experiment learners

In the field of cognitive psychology, human beings have many sensory perceptions to perceive the characteristics of external things, but visual perception is the most valuable and important one (Wang Haiyan, et al., 2016). However, human attention is limited, only specific resources can be allocated to the visual system for perception, and visual selective attention will play an important role in this process. According to the three-stage model of visual selective attention and perceptual perception, a core viewpoint of visual cognition theory is that humans do not absorb all the received information, but selectively digest some of the received information according to their own visual habits and original experience. Under the guidance of visual cognitive theory and cognitive emotion theory of multimedia learning, this study extracts visual elements that trigger students' emotions in the design of virtual experiment environment based on three levels of emotional design, and analyzes the visual and cognitive characteristics of virtual experiment learners according to the previous two sections, so as to build a visual cognitive model of virtual experiment learners. This model helps us to understand the visual cognitive process of virtual experiment learners from different aspects.

4. Experiment and results of immersive virtual experiment environment

The specific objectives of the experiment are explore the influence of warm and cold colors on the affective assessment of virtual experiment learners. And the research hypothesis are as follows. Hypothesis (1): Learners' emotions are different in the virtual experimental environment of cool color and warm color. Hypothesis (2): Learners' emotions are different in the virtual experimental environment with simple scenes and complex scenes. The number of subjects recruited in this experiment was 60, the experiment samples are all college students, 25 boys and 35 girls, aged at 25 years old. All subjects had normal vision and had not participated in similar experiments. All subjects were right-handed and were tested separately. All experiments were completed. The content of this experiment is the tensile experiment of low carbon steel, a compulsory course of architectural mechanics major. We uses Unity3D software to build a virtual environment and design interaction program. The content is mainly about "the relationship between force and displacement of low carbon steel in the drawing process of low carbon steel", which lasts 10 minutes and includes the following key steps: (1) The low carbon steel is correctly installed on the electronic universal machine; (2) Installation of extensometer; (3) Open the operating software and set parameters; (4) Complete the whole process of low carbon steel drawing; (5) Generate

experimental report data. We apply PAD three-dimensional emotion model to measure user emotions. The simplified Emotion Scale is an emotional scale developed by Mehrabian based on PAD's three-dimensional emotion model. PAD three-dimensional emotion model holds that emotion has three independent dimensions, P, A and D, where P stands for pleasurable degree, which represents the positive and negative characteristics of individual emotional states. A is for activation - represents the individual's physiological activation level; D is for dominance - indicates how much control an individual has over a situation or others. The PAD Affective Scale is based on the three-dimensional affective model of PAD developed by Mehrabian and Russell in 1947 (Mehrabian, A. & J. Russell, 1977). PAD emotional dimension scale believes that emotion has three dimensions: pleasurable degree, arousal degree and dominance degree. The original scale has been slightly modified to be more suitable for testing learners' emotional state in a virtual environment.

4.1The influence of visual elements on learners' affective assessment in virtual experiments

In the cognitive process of visual selective attention, spatial location and object features are the two factors that preferentially select objects (Han Xinli, 2018). The emotional state of the stimuli in the virtual experiment scene was taken as the dependent variable, and the color (warm color, cool color) and scene complexity (concise and complex) were taken as the independent variables. A 2×2 analysis of variance was carried out in the mixed design of multiple factors. Multivariate analysis of variance is used to analyze the influence of color and scene complexity on learners' emotions, which is completed in two steps. The first step is to analyze the independent influence of two variables on learners' emotions; The second step is to analyze the influence of the interaction of the two variables on emotion.

In this multivariate analysis of variance, the isogeneity of the covariance matrix is shown in the Box's test results in Table 1, F (18,11081) = 1.336, p = 0.154, greater than 0.05, which has two meanings. The first meaning is homogeneity of variance. In other words, it is considered that the overall variance of the GPA of learners' emotional scores in the four scene designs (simple scene + warm tone/simple scene + cool tone/complex scene + warm tone/simple scene/cool tone) is homogeneous, and AN variance analysis can be performed to meet the prerequisite requirements of AN variance analysis. The second implication is that in "multivariate test", wilks' Lambda test statistic is used to explain the meaning represented by the data in the results of multivariate analysis of variance.

Table 1. Box test for equality of covariance matrix

Box's M	F	df1	df2	Sig.
26.674	1.336	18	11081.816	.154

Test the null hypothesis that the observed covariance matrix of the dependent variable is equal in all groups

a. Design: intercept + color + complexity + color * complexity

The results of "multivariate test" in multivariate analysis of variance are shown in Table 2. There was no significant difference in the main effect of color in the laboratory in the virtual environment, wilks' Lambda=0.973, F(3,54)=0.492, p>0.001,η2=0.027, that is, it was believed that the color in the laboratory in the virtual environment had no significant effect on learners' emotions. There are significant differences in the main effects of the laboratory scene complexity in the virtual environment, wilks' Lambda=0.564, F(3,54)=13.905, p<0.001, η2=0.436, that is, it is believed that the complexity of the laboratory scene in the virtual environment has a significant impact on learners' emotions. In the virtual environment, lab color and scene complexity have no significant interaction with learners' emotion test results. Wilks' Lambda=0.946, F(3,54)=1.035, P=0.385, η2=0.054, In other words, the interaction between lab color and scene complexity in virtual environment has no significant influence on learners' emotions. Since the multivariate effect is too small, the α level was adjusted to α=0.016 before testing the results of one-way ANOVA, because learners' emotions were analyzed from three dimensions in this study.

Table 2. Multivariate test

Effect	**Wilks' Lambda**	**F**	**Hypothesis df**	**Error df**	**Significance**	**Effect size**
Colour	.973	.492[b]	3.000	54.000	.689	.027
Scene complexity	.564	13.905[b]	3.000	54.000	.000	.436
Colour * Scene complexity	.946	1.035[b]	3.000	54.000	.385	.054

a. The result of the calculation using alpha=.05

b. Precise statistics

c. Design：Intercept+Colour+Scene complexity

4.2 The effect of learner's gender on emotion in virtual experiment

In Wang Ying's DSMS model (Wang Ying, etc., 2016), the physiological characteristics of the demographic characteristics refer to the differences in gender and age of learners, the social characteristics refer to the educational level and family economic situation of learners. Among the 60 subjects, there were 25 male subjects and 35 female subjects. In the multivariate analysis of variance, covariance matrix equality Box's test, F (6,18256) = 2.356, p = 0.028, p value is less than 0.05, in the "multivariate test", pillai Trace test statistics are used to explain the meaning of the data represented in the results of multivariate analysis of variance. The results of "multivariate test" in multivariate analysis of variance show that gender of virtual experiment learners has significant difference on emotion, pillai's Trace=0.883, F(3,56)=3.095, p=0.034,η2=0.142. The p value is less than 0.05, indicating that gender of learners in virtual experiment has a significant difference in

emotion. Based on the analysis results of multivariate test data, the data of "test of intersubjective effect" in multivariate analysis of variance were further analyzed, and the effect of gender of learners on emotional pleasure was F(3,56)=1.435, p=0.956, $\eta 2$=0.00. p value is greater than 0.05; Activation F(3,56)=0.032, p=0.858, $\eta 2$ = 0.001, p is greater than 0.05; Dominance F(3,56)=7.608, p=0.008, $\eta 2$=0.116, and p value is less than 0.05, indicating that learners' gender has no significant difference on pleasure and activation in virtual environment, but has significant difference on dominance. In the dimension of emotional dominance, male students (M=2.16, SD=0.8) were more likely than female students (M=1.39, SD=1.23) to autonomously control emotions induced by virtual environments.

Table 3. Data table of gender's influence on emotion

	Variable source						
	Gender	Mean value	Standard deviation	Degree of freedom	F	Significance	Effect size
Happiness	Male	2.3700	.91036	1	.003	.956	.000
	Female	2.3571	.85165				
Activation degree	Male	2.1100	1.2626	1	.032	.858	.001
	Female	2.0571	1.00748				
Degree of dominance	Male	2.1600	.80325	1	7.608	.008	.116
	Female	1.3857	1.22676				

5. Conclusion

Virtual experiment can solve many problems existing in traditional experimental teaching, improve the quality of practical teaching and cultivate students' ability of independent learning. Therefore, virtual experiment plays a very important role in traditional teaching, and its environment design will directly affect learners' sense of experience. Based on the previous emotional design of multimedia learning materials and the "visual cognition theory", this study uses questionnaire survey to measure the emotions of learners in virtual experiments, and deeply discusses the mapping relationship between visual elements and emotions in virtual experiments. On this basis, the visual emotion design scheme of virtual experiment is obtained, in order to optimize the design of virtual experiment environment and provide reference for improving the emotion of virtual experiment learners. The participants in the experiment show a good distribution in terms of personnel structure, and most learners hold a positive attitude towards virtual experiment. The application of virtual experiment in teaching can greatly improve learners' learning motivation. The differences between the four emotionally-designed scenes in terms of emotional activation and dominance are not great, and there are obvious differences in pleasure. The influence of scene complexity is more significant. The mapping relationship between virtual experiment environment and learners' emotions reflects the combination of warm colors and complex scenes in optimal design scheme. To sum up, this study examined the influence of visual element design in

virtual experimental environment on learners' emotions. Although the design of the experiment only found the positive effects of color and scene complexity on stimulating learners' emotions, the positive effects of sound and learning partners were also mentioned in the interview records. Virtual experiment designers can consider the important role of design-related elements of virtual environment in the learning process from multiple channels such as hearing and touching.

References

Diana Ibet Roman-Sanchez, Teresa Lopez-Arenas,& Mauricio Sales-Cruz. (2018). An interactive virtual laboratory for process engineering education. Computer Aided Chemical Engineering, (43): 1707-1708.

Erez, & Isen. (2022)The influence of positive affect on the components of expectancy motivation. Journal of Applied Psychology, 87 (6) :1055-1067.

Han Xinli. (2018). Research on the application of visual guidance in Web Design. Qingdao University of Technology.

Jensen R. (2021). The Dream Society: How the Coming Shift from Information to Imagination Will Transform Your Business. Library Journal, (7): 112.

Li Hongxia. (2020). Research on the problems and countermeasures of physics experiment teaching in junior middle school. Central China Normal University.

Liu Yaru. (2018). Research on Emotional design of electronic products for school-age children based on user experience. Xi 'an Polytechnic University.

Lu Jiamei. (2016). On Affective Teaching Model. Educational Research, (12):55-60.

Mayer, R. E. Cognitive theory of multimedia learning (2nd ed.). (2005). New York: Cambridge University Press.

Mehrabian, A. and J. Russell, Evidence for a three-factor theory of Emotions [J]. Journal of Research in Personality, 1977. 11: p. 273-294.

Nurdyani F,Slamet I,& Sujadi I. (2018). Creative thinking level of students with high capability in relations and functions by problem-based learning. Journal of Physics: Conference Series, 983:102-112.

Wang Haiyan, HUANG Yamei, & Chen Mo. (2016). Analysis of ACT-R cognitive Model for Icon visual Search Behavior. Journal of Computer-Aided Design and Graphics, 28(10):1740-1749. (in Chinese)

Wang Ying, An Yujie, Huang Ronghuai. (2016). Research on the construction of distance learner Feature Model. Open Education Research, (01):47-53.

Ware C. (2012). Information visualization: perception for design. Elsevier.

Shen Zhuqi. (2019). Research on Visual Cognitive Behavior Model of Information Visualization Interface. Southeast University.

Zhu Shiyu. (2018). Research on Doll Expression Design Based on Children's Visual Cognition [D]. Xiangtan University.

LearningverseVR: Intergrating Generative AI and VR Technology in an Immersive Game-based Learning Platform

Yanjie Song [1*], Kaiyi Wu [2], Jiaoyang Ding [3]
[1]The Education University of Hong Kong, Hong Kong, PRC
[*] ysong@eduhk.hk

Abstract: *Rapid advances in generative Artificial Intelligence (AI) and Virtual Reality (VR) technologies are revolutionizing a variety of fields, including the education and gaming industries. However, there is still little research on how to integrate AI and virtual reality technologies to enhance immersive game-based learning. Therefore, the article describes the design and development process of "LearningverseVR", an immersive game-based learning platform developed using generative AI and VR technologies, which is based on the metaverse platform "Learningverse" developed by the lead author and her research team. The platform's design framework consists of two main components: generative AI for game-based learning and virtual reality for immersive learning. Generative AI utilizes an ecosystem of large language models to create NPCs with different personalities and life backgrounds, enabling learners to interact with the NPCs in a VR environment without scripted dialogues, resulting in an interactive and immersive game-based learning experience. The creation of the 'LearningverseVR' platform provides a new perspective on utilizing advanced technologies in education.*

Keywords: Generative AI, Metaverse, Virtual Reality (VR), Game-based learning, Immersive

1. Introduction

Recent developments in generative AI have shown the world its great potential and have dramatically changed the landscape in many fields. Generative AI can generate text, images, audio, and even video content based on large-scale training data to generate new, creative, and unique content for users. In the field of education, generative AI is increasingly used in teaching and learning to provide personalized learning and enhance immersion. For example, digital games are enhanced by using generative AI to make them more personalised and authentic (Gozalo-Brizuela & Garrido-Merchán, 2023), while avoiding cumbersome scripting through automated test execution (Prasetya et al., 2022). In digital games, generative AI has the capability to create non-player characters (NPCs) with different personalities and life backgrounds (Karaca et al., 2023); generate real-time dialogue and behaviour (Westera et al., 2020); and adjust the personalities and behaviours based on player preferences and habits. Thus, it allows for a flexible and unique gaming experience and enhances player engagement and immersion. In specific, Large Language Models (LLMs) can be used to simulate human behaviour (Park et al., 2023), giving NPCs more realistic and authentic characteristics.

Virtual reality has the potential to improve interactivity, immersion, and sensation of presence (Burdea & Coiffet, 2003). According to Girvan (2018), the virtual reality environment offers "shared, simulated spaces which are inhabited and shaped by their inhabitants who are represented as avatars" and permits multi-user interactions. Realistic virtual worlds and characters (such as avatars and NPCs)

can be created using generative AI which can enhance students' motivation in the game-based learning environment (Papastergiou, 2009). Increased presence, interaction, and immersion during the learning process can benefit students' motivation, engagement, embodiment, and capacity for knowledge generation (Helsel, S. 1992; Virvou et al., 2005; Makransky & Petersen, 2021). However, not much research has been done to change digital game-based learning through the use of generative AI and virtual reality.

In view of this, the aim of this project is to create "LearningverseVR," an immersive game-based learning platform that uses generative AI and virtual reality to improve learners' immersive and interactive experiences and offer novel views on digital game-based learning.

2. Development of the Immersive Game-based Learning Platform – LearningverseVR

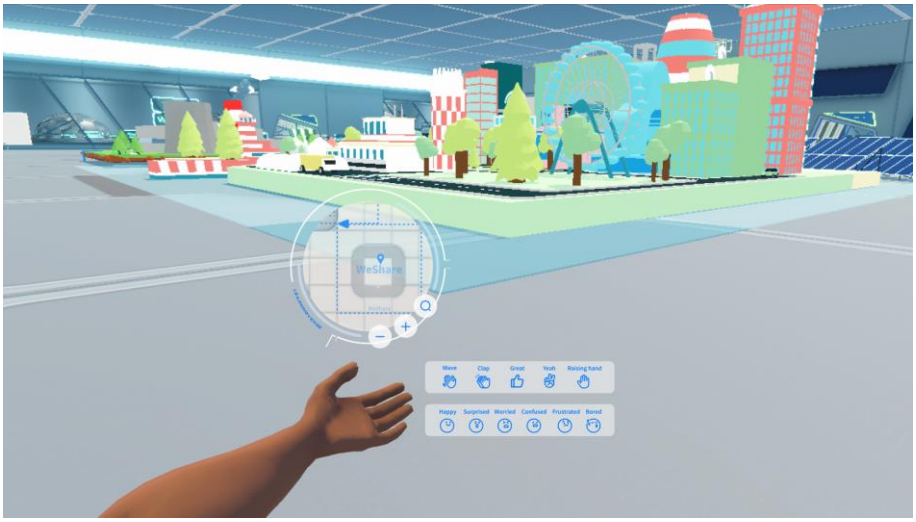
"LearningverseVR" is an immersive game-based learning platform based on "Learningverse" - a 3D educational metaverse platform developed by the first author and her research team (Song et al., 2023). The platform was created using generative artificial intelligence (AI) and virtual reality (VR) technologies that provide an immersive learning environment in which learners can act as avatars to interact with virtual environments and other avatars. Fig.1 displays an avatar interacting with objects in a first-person view in the LearningverseVR environment. Also, the platform provides custom collaborative tools for learning activity design on a common computer, lowering the threshold for learners to create a learning environment and access the metaverse platform.

The successful integration of VR and generative AI technology into digital game-based learning required a robust technological infrastructure. With this, the development of "LearningverseVR" builds upon our current metaverse platform, "Learningverse," to accomplish the smooth integration of VR and generative AI for game-based learning through the integration of several technologies.

The "LearningverseVR" platform uses Unity as the client and VR engine. The concurrent service offered by Unity enables online multiplayer capability, enabling students to participate in engaging and dynamic learning experiences. GPT-3.5 (OpenAI, 2023) is selected as the general-purpose Large Language Model (LLM) that functions as the agents' brain. It is responsive, affordable, and has strong all-around capabilities. For logic agents like trigger agents and data masking agents, a local LLM – 'LLaMa' serves as the brain. The combination of local deployment and the use of OpenAI offers several benefits includes: (1) Cost control: Cloud services can save money by being especially well-suited for frequent and large-scale usage. (2) Data security: By processing user data locally, the dangers related to data transmission and storage are decreased. (3) OpenAI's state-of-the-art LLM, which offers superior natural language processing, sophisticated conversational skills, and extensive training data and computational resources. In summary, a more complete and effective solution in terms of cost management, data security, and feature growth is offered by the hybrid usage of local deployment and cloud services. To provide multimodal capabilities, Stable Diffusion is used as the text-to-image generation model from text to visuals, and a powerful text-to-speech converter is employed in addition to generative AI to enable the NPC can speak and listen.

Fig. 1

First-person view of an avatar in LearningverseVR environment



3. Design Framework of “LearningverseVR” with Generative AI

3.1. Generative AI-driven NPCs

To produce NPCs with a variety of personalities and life experiences, generative AI is employed to provide the background and context. In the LearnigverseVR environment, these NPCs may converse with the user and have their own avatars and opinions in the advancement of technology. Using generative AI to generate an NPC requires three steps: (1) Retrieval and prompt engineering based on the LLMs, which guarantees the NPC's foundational knowledge; (2) Generative visual model and text-to-speech converter creation of an agent, which energises the LLMs; and (3) Agent-client communication, which permits the generated agent to interact with the client and have its embodied avatar as an NPC in the learning environment.

First, a client allows the LLMs to be context-aware and reasonable, and provide communication between the environment and NPC. Second, to improve the accuracy of responses and prevent hallucinatory responses from LLM-based agents, a knowledge base is introduced with a vector database and retrieval enhancement, as well as Retrieval Augmented Generation (RAG) (Lewis et al., 2020), in which external data is retrieved and then passed to the LLM during the generation step. The knowledge base is a vector database designed to store embedded data and execute vector searches. In addition, a nested prompt engineering is used, in which users enter information into prepared fields. Third, the agents communicate with "LearningverseVR" via API. Users can speak with 3D NPCs, which are physical avatars of LLM-based, and have the option of manually entering or using speech recognition. The content entered by users is routed to the agent for processing. After obtaining responses from the agent, the client sends text-to-speech (TTS) requests sentence by sentence to generate speech and reduce latency. Furthermore, emotions and body gestures are added to NPCs by predefining system prompts. When NPCs receive information comprising tone and action, the client sends tone markers to the TTS system, which then plays the appropriate body movement animations in "LearningverseVR".

3.2. Generative AI to Support the Game-based Learning Process

In game-based learning, generative AI improves students' perceptions of freedom and control in virtual worlds and enriches interactions with non-player characters. For starters, generative AI may significantly enrich the grid narrative and increase the flexibility of the game-based learning process. In the classic game-based learning method, the NPC's conversation flow and options are linear or tree-like, with branches occurring only at a few crucial nodes. With generative AI, we only need to set the NPCs' personal information and limit the knowledge base they have, but not the precise content of their discussions. Outside of this linear procedure, the learner can divert his or her thoughts to create knowledge networks. These increase the level of involvement and investigation of his or her learning process, perhaps improving learning outcomes (Rowe et al., 2011).

Generative AI can be used to add gamification components such as points, badges, leaderboards, and achievements through LLM communication with the image generation module. It can also be dynamically adjusted according to the students' learning progress, mastery of knowledge points, completion of learning activities, and assessment of learning outcomes. And it provides a reference and basis for teachers' evaluation.

4. Design Framework of “LearningverseVR” with VR Technology

4.1. Enhancing the Interactive and Immersive Experience with LLM-driven NPCs in Virtual Reality

In the VR environment, interactive experiences are essential for establishing a sense of immersion. Common virtual reality interactions include clicking, gripping, and throwing objects, which are frequently replicated using physics engines to replicate physical impacts. By transferring the player's interaction status with things to an NPC's, we may extend object interactions to include interactions with NPCs, who will respond via their embodied avatars based on the learner's interactions with specific objects. Interactions with NPCs provide learners with a more authentic and rewarding experience, allowing them to form deeper connections with NPCs in the virtual world. This type of interaction not only increases student engagement and interest but also allows for the expression and sharing of emotions.

To improve immersion in the VR environment, we built the LLM ecosystem that involves three categories of agents: world agents, organic agents, and inorganic agents. These three categories of agents perceive and influence one another, resulting in a single environment. The framework of the LLM ecosystem improves immersion and presence in virtual reality situations. Inorganic items (e.g., weather, temperature, seasons) are important aspects that traditional design approaches typically ignore. However, the LLM ecosystem incorporates these elements, making the scene appear more genuine and colourful. Learners can have a wide range of interactions with virtual characters and environments, improving immersion, interactivity, and engagement.

4.2. Implementing Memory, Affinity, and Familiarity Mechanisms for NPCs in VR Environment

The client captures dialogues and interactions between learners and NPCs and stores them in a vector database to develop long-term memory. In the following exchange, the client retrieves the

prior chat records. The NPC can learn about the learners' interests and preferences by recording their conversations and interactions with it. When used with the background environment, the NPC may deliver individualized discussion experiences for each learner.

To improve the immersion of the interaction, two variables are added to the design: likability and familiarity between the NPC and learner. When the learner interacts with the NPC, the client receives the likability and familiarity values from the database and sends them to the NPC's agent, and the LLM determines whether these values should be changed and updates the database appropriately. If likability and familiarity are low, the interaction may appear lacking in excitement and warmth, and NPC's comments may be brief and incomplete. On the other hand, when likability and familiarity are high, the interaction becomes more amicable and focused on the learner. The NPC provides further information and comprehensive responses, actively participates in conversation, and expresses encouragement and appreciation.

5. Conclusion

To improve students' interactive and immersive learning experiences, this research presents the creation of "LearningverseVR," an immersive game-based learning platform powered by generative AI and VR technology which provides engaging and dynamic learning activities. This research uses generative AI to create NPCs with diverse personalities and life backgrounds to provide an interactive game-based learning environment, and integrate with VR technology, which further enhances learner immersive and interactive experiences by leveraging the Large Language Model (LLM) ecosystem. This research contributes to educational practice which involves integrating cutting-edge technologies in an innovative pedagogical design; constructing an immersive learning environment to activate student motivation and engagement; and providing an innovative framework for designing and developing NPC. The development of the LearningverseVR platform offers innovative perspectives on game-based learning.

Acknowledgements

The study was funded by the Fund for Innovative Technology-in-Education (FITE) (Ref. 05404 and 05408) (2023-26), University Grant Committee, Hong Kong.

References

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology*. John Wiley & Sons.

Girvan, C. (2018). What is a virtual world? Definition and classification. *Educational Technology Research and Development*, *66*(5), 1087-1100. https://doi.org/10.1007/s11423-018-9577-y.

Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023). A survey of Generative AI Applications. https://doi.org/10.48550/arXiv.2306.02781

Helsel, S. (1992). Virtual reality and education. *Educational Technology*, *32*(5), 38-42.

Karaca, Y., Derias, D., & Sarsar, G. (2023). AI-Powered Procedural Content Generation: Enhancing NPC Behaviour for an Immersive Gaming Experience. https://www.researchgate.net/publication/376480816

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 1-22. https://doi.org/10.1007/s10648-020-09586-2.

OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. https://chat.openai.com/chat

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & education*, *52*(1), 1-12. https://doi.org/10.1016/j.compedu.2008.06.004

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023, October). Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (pp. 1-22). https://doi.org/10.1145/3586183.3606763.

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education*, *21*(1-2), 115-133. https://doi.org/10.3233/JAI-2011-019

Song, Y., Cao, J., Wu, K., Yu, P. L. H., & Lee, J. C. K. (2023). Developing 'Learningverse'-a 3D Metaverse Platform to Support Teaching, Social and Cognitive Presences. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/TLT.2023.3276574

Virvou, M., Katsionis, G., & Manos, K. (2005). Combining software games with education: Evaluation of its educational effectiveness. *Journal of Educational Technology & Society*, *8*(2), 54-65.

Westera, W., Prada, R., Mascarenhas, S., Santos, P. A., Dias, J., Guimarães, M., ... & Ruseti, S. (2020). Artificial intelligence moving serious gaming: Presenting reusable game AI components. *Education and Information Technologies*, *25*, 351-380. https://doi.org/10.1007/s10639-019-09968-2.

Measuring EFL Learners' Perceptions of Using ChatGPT in Translation: An Exploratory Study

Huiqi Fan, Bowen Jing, Yanqing Yi, Lin Luan*
Beijing University of Posts and Telecommunications, China
*luanlin@bupt.edu.cn

Abstract: *ChatGPT, one of the most intelligent AI-powered chatbots developed by OpenAI, has introduced numerous opportunities and challenges in the field of English as Foreign Language (EFL) learning. For EFL learners, translation can be recognized as a helpful tool to improve their EFL proficiency. Due to its remarkable ability to provide translations within minutes, ChatGPT has attracted growing attention from scholars to delve into its role and use in translation. However, the acceptance and use of ChatGPT in translation from EFL learners' perspective remains largely unexplored. To fill this gap, the study aimed to investigate EFL learners' intentions to use ChatGPT in translation and their perceptions of applying the AI-based chatbot in their translation process. 26 EFL learners participated in this study and both questionnaire and semi-structured interview were used to collect quantitative and qualitative data respectively. The results demonstrated that the majority of students held positive attitudes towards the use of ChatGPT and exhibited high intention to use ChatGPT in their translation process. Analysis of interviews also revealed students' perceptions of the benefits (e.g., high speed and low cost) and limitations (e.g., output variability) as well as the challenges introduced by ChatGPT. Such findings open the door for empirical research to confirm the perceived benefits and limitations of ChatGPT in this study and generate implications for future research in ChatGPT for translation purposes.*
Keywords: ChatGPT, EFL learners, learners' perception, translation

1. Introduction

With the advancement of artificial intelligence (AI) technologies, chatbot technology, with ChatGPT being the representative, has received growing recognition in language learning due to its ability to communicate with language learners using natural language (Huang et al., 2022). The rapid development of ChatGPT in recent years has led to increasing attention from researchers and educators in the domain of EFL learning and translation. As an important component in second language acquisition (SLA), translation can be recognized as a helpful tool to improve their EFL proficiency (Putrawan, 2019). Previous studies have shown that ChatGPT possesses significant benefits in the translation process and that it raises public concerns about its potential risks (Siu, 2023). Nevertheless, EFL learners' perceptions of using ChatGPT in translation is still insufficiently investigated. Therefore, this study sets out to explore the extent to which EFL learners intend to use ChatGPT in translation and their perceptions of ChatGPT use in their translation process.

2. Literature Review

2.1. ChatGPT in EFL Teaching and Learning

Since its first public debut, ChatGPT has sparked heated debate on its opportunities for and threats to all walks of life. Particularly, there are a series of studies giving heed to the potential opportunities and challenges for the introduction of ChatGPT in higher education. For instance, Farrokhnia et al. (2023) employed the SWOT analysis framework to outline ChatGPT's strengths and weaknesses that can identify opportunities for and threats to education respectively. Despite many challenges related to the application of ChatGPT, the studies mentioned above also combine to stress the potential of ChatGPT as a valuable and powerful digital learning tool in higher education. In the context of EFL learning, ChatGPT has the potential to assist students in developing various skills, including reading, writing, listening, speaking and the

learning of grammar and vocabulary (Ali et al., 2023).

2.2. The Application of ChatGPT in Translation

In terms of English language learning, the translation process between the source language and the target language are inextricably linked (Roza & Zulhirawati, 2023). With impressive performance in natural language processing tasks, such as question answering, text generation and language translation (Siu, 2023), ChatGPT has demonstrated its potential in translation tasks especially (Jiao, 2023), which has garnered growing attention in the field of translation. For example, Siu (2023) delved into the application of ChatGPT in translation, examining its strengths and limitations. However, most existing studies have explored the impact of ChatGPT in translation, yet little research has been focused on the acceptance and use of ChatGPT in translation within the EFL contexts. Therefore, two research questions were proposed as follows:

(1) To what extent do EFL learners intend to use ChatGPT in translation?

(2) How do EFL learners perceive the application of ChatGPT in translation?

3. Method

3.1. Research Context and Participants

This study was conducted at a comprehensive university in China. A total of 26 postgraduate students across years 1-3 from EFL backgrounds (22 females, and 4 males) participated in the study. The participants aged from 22 to 28 years were from Translation (n=23) and Foreign Language and Literature (n=3). The participants were chosen purposefully based on the first author's conversations with students and classroom observations to ensure all participants had at least heard of ChatGPT and had better use ChatGPT in translation before this study.

3.2. Instruments

The questionnaire consisted of two sections. The first section contained six questions to collect information about participants' gender, age, year of study, major and experiences of using ChatGPT in translation. Based on Davis (1989), the second section comprised four Technology Acceptance Model (TAM) constructs (Perceived Ease of Use, Perceived Usefulness, Attitude and Behavioral Intention) with 18 close-ended items that were measured with 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5). The items were adapted from TAM questionnaires concerning ChatGPT use within EFL contexts (Liu & Ma, 2023; Yilmaz et al., 2023; Zou et al., 2023). In addition, semi-structured interviews were conducted to explore how the participants perceive the application of ChatGPT in translation. Three interview questions are as follows: (1) What are the advantages of using ChatGPT in translation? (2) What are the disadvantages of using ChatGPT in translation? (3) What are the challenges that ChatGPT has brought to translation?

3.3. Data Collection and Analysis

In this study, the questionnaire and semi-structured interviews were conducted to obtain quantitative and qualitative data respectively and all the questions were presented in Chinese, participants' native language. To address the two research questions, descriptive statistical analysis was first employed to analyze quantitative data using SPSS 26. Then participants' responses in semi-structured interviews were analysed through the general inductive approach.

4. Results

4.1. Analysis of Questionnaire Results

To investigate the extent to which EFL learners intend to use ChatGPT in translation, descriptive analysis was performed on data collected from the questionnaire. The mean score for each dimension lay in the range of 3.65-3.96, which demonstrated students' positive perceptions and attitudes towards the use of ChatGPT in translation. Specifically, the participants had the highest scores in the attitude dimension (Mean=3.96, SD=0.54), as they were interested in and

positively assessed the use of ChatGPT in their translation process. For instance, most participants particularly stated that ChatGPT is an attractive tool for translation purposes (Q12, Mean=4.23, SD=0.51). However, the dimension with the lowest score was "Perceived Ease of Use" (Mean=3.65，SD=0.52). In addition, the mean scores for "Perceived Usefulness"（Mean=3.92, SD=0.45）and "Behavioral Intention" (Mean=3.91, SD=0.63) were almost equal, both of which were close to 4, indicating that most students agreed that ChatGPT was useful in their translation process, and they had high intentions to use this AI-powered tool.

4.2. Analysis of Interviews

Table 1 demonstrated EFL learners' perceptions of ChatGPT use in translation from three perspectives, benefits, limitations and challenges. In terms of benefits, all the participants claimed that ChatGPT performs translation tasks in a short time. Over half of the students agreed with ChatGPT's strengths in different aspects, including low cost (73.08%), interactivity (57.69%) and flexibility (50.00%). Moreover, 38.46% of the students favored ChatGPT's versatility as ChatGPT could assist them handle a variety of tasks. In contrast, agreement with limitations ranged from 50.0% for "Misleading information and Bias" to 84.6% for "Output variability." Furthermore, the majority of students worried that ChatGPT "Generates inaccurate or inappropriate content" (80.77%) and "Undermines their initiative and autonomy" (84.62%); and more than half of participants expressed their concerns about "the unemployment situation in the translation industry" (65.38%), "Ethical and moral issues" (61.54%) and "Data security and privacy issues" (53.85%).

Table 1. Students' Interview Results.

Theme	Code	N	Percentage
Advantages	High speed	26	100%
	Low cost	19	73.08%
	Versatility	10	38.46%
	Flexibility	13	50.00%
	Interactivity	15	57.69%
Disadvantages	Poor performance in some specialized texts	20	76.92%
	Misleading information and bias	13	50.00%
	Limited support for Chinese language	15	57.69%
	Output variability	22	84.62%
Challenges	Generates inaccurate or inappropriate content	21	80.77%
	Exacerbates the unemployment situation in the translation industry	17	65.38%
	Data security and privacy issues	14	53.85%
	Undermines the initiative and autonomy of translation learners or translators	22	84.62%
	Ethical and moral issues	16	61.54%

5. Discussion

The study investigated Chinese EFL learners' perceptions of using ChatGPT in translation. The findings revealed that most participants held positive attitudes toward using ChatGPT in translation, which corroborated that of Roza and Zulhirawati (2023), indicating that ChatGPT was easy to use, useful and enjoyable for translating English texts. Further analyses of interviews revealed that ChatGPT has considerable advantages especially in terms of high speed and low cost, making it attractive for generating automatic translations rapidly and accessible to individuals with limited budgets, which aligned with the findings of Siu (2023). Furthermore, most participants thought that ChatGPT has limited capabilities for some specialized texts, and they mostly reported that ChatGPT tends to generate multiple translations for the same source text at different times (Chen et al., 2023; Siu, 2023). In addition, students were also concerned about the challenges brought by ChatGPT, and they were most afraid that their initiative and autonomy may be undermined by the powerful AI tool, which also aligned with the findings of Sahari et al. (2023) and Siu (2023), suggesting that students must not rely too much on ChatGPT and it is essential for them to amplify professional skills.

6. Conclusion

This study was conducted to explore Chinese EFL learners' acceptance of ChatGPT and how they perceived this AI-powered tool in their translation process. Through the descriptive statistics of items in the modified TAM questionnaire, the study first revealed students' positive attitudes towards and high intentions to ChatGPT for their translation needs. It was also found that ChatGPT has considerable advantages such as high speed, low cost and high interactivity. However, students also expressed their concerns about the disadvantages of and challenges brought by ChatGPT.

This study also has some limitations. First, the sample size was relatively small with only 26 students from one single university. Future research is encouraged to recruit larger group members, in order to guarantee the generalization of the research findings. Secondly, all data were mainly collected from a self-report survey. To this end, advanced technologies such as artificial intelligence and multimodal learning analytics may be employed to obtain more in-depth insight into students' perceptions of ChatGPT in translation.

Acknowledgements

This research was supported by the Teaching Reform Projects of Beijing University of Posts and Telecommunications (Grant number 2022Y013 and 2023ZD08).

References

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on Learning Motivation: *Journal of English Studies in Arabia Felix*, *2*(1), 41–49.

Cheng, Y., Wang, R., Chen, J., Chao, Y., Maimaitili, A., & Zhang, H. (2023). Context-Based AI Translation From a Globalization Perspective: A Case Study of ChatGPT. *Sino-US English Teaching*, *20*(9), 370-380.

Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *00*(00), 1–15.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237–257.

Jiao, W., Wang, W., Huang, J. T., Wang, X., Shi, S., & Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv preprint arXiv:2301.08745*.

Liu, G., & Ma, C. (2023). Measuring EFL learners ' use of ChatGPT in informal digital learning of English based on the technology acceptance model learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14.

Putrawan, G. E. (2019). The role of first language and translation in EFL learning: A brief literature review. *International Journal of Linguistics, Literature and Translation (IJLLT)*, *2*(2), 170-174.

Roza, V., & Zulhirawati, Z. (2023). Higher Students' Perception of Using Chat GPT in Translating English Texts. *BiCED Proceeding*, *1*, 64-73.

Sahari, Y., Al-Kadi, A. M. T., & Ali, J. K. M. (2023). A Cross Sectional Study of ChatGPT in Translation: Magnitude of Use, Attitudes, and Uncertainties. *Journal of Psycholinguistic Research*, *52*(6), 2937-2954.

Siu, S. C. (2023,). *ChatGPT and GPT-4 for Professional Translators: Exploring the Potential of Large Language Models in Translation*. Social Science Research Network. https://doi.org/10.2139/ssrn.4448091

Yilmaz, H., Maxutov, S., Baitekov, A., & Balta, N. (2023). Student attitudes towards chat GPT: A technology acceptance Model survey. *International Educational Review*, *1*(1), 57-83.

Zou, B., Liu, Q., Han, Y., Li, Z., & Zhang, W. (2023). Exploring students' acceptance of an artificial intelligence speech evaluation program for EFL speaking practice: an application of the Integrated Model of Technology Acceptance. *Computer Assisted Language Learning*, 1–26.

我国职业院校教师数字素养发展可视化分析

Visual Analysis of Vocational Teachers' Digital Literacy in China

刘妍萱，叶建宏*
北京师范大学教育学部
* kimpo30107@hotmail.com

【摘要】 教师数字素养是技术变革引领下职业教育教学实践的必然要求。本文基于 2010 年至 2024 年间 CNKI 数据库中与职业教育教师数字素养相关的 CSSCI 及核心期刊文献，运用 CiteSpace 可视化软件对文献数量、关键词等进行提取和解析，梳理并呈现近十五年来我国职教教师信息素养的研究热点、发展现状及趋势，以期为后续研究提供参考。研究发现，职教教师数字素养领域的热点包括数字素养的丰富内涵与构成要素、数字素养的评价指标体系构建等，研究尚处于探索阶段，未来应更加注重理论与实践的结合，推动职教教师数字素养研究的全面发展。

【关键词】 职业教育；教师数字素养；CiteSpace；可视化分析

Abstract: *Teachers' digital literacy is an inevitable requirement for the teaching practice of vocational education led by technological change. Based on the literature of CSSCI and core journals related to information literacy of vocational education teachers in the CNKI database between 2010 and 2024, this paper extracts and analyzes the number of articles, issuing numbers, keywords, etc. using the CiteSpace visualization software, to organize and present the research hotspots, the current development status and the trends of the information literacy of vocational education teachers in China in the past 15 years, in order to provide a reference for subsequent research. The study found that the hotspots in the field of digital literacy of VET teachers include the rich connotation and constituent elements of digital literacy, the construction of evaluation index systems of digital literacy, etc. The study in this area is still in the exploratory stage; in the future, more attention should be paid to the combination of theory and practice, and to promoting the comprehensive development of the research on the digital literacy of VET teachers.*

Keywords: Vocational Education, Teachers' Digital Literacy, CiteSpace, Visual Analysis

1. 前言

习近平总书记指出："数字技术正以新理念、新业态、新模式全面融入人类经济、政治、文化、社会、生态文明建设各领域和全过程，给人类生产生活带来广泛而深刻的影响。"在数字技术的全面加持下，职业教育大步迈向了数字化高速发展的时代，"数字技术＋职业教育"的在场形态得到全面更新（李晓娟等人，2021）。教育部职业教育与成人教育司（2022）发布的《职业教育与继续教育 2022 年工作要点》明确指出，"以新一代信息技术为重点，加强师生信息素养教育"。目前，教师作为现代职业教育数字化建设的核心要素，其数字素养的缺失是制约发展的关键。2019 年发布的《中国教育信息化蓝皮书》指出，虽然我国的教育信息化建设取得了一定成效，但教师的数字化能力整体水平仍然偏低，数字教育教学应用还没有充分发挥其应有的作用。其中，高职教育是数字化能力较为薄弱的一环，数字化教学资源利用率相对较低（许倩倩等人，2023）。在此背景下，本研究使用知识图谱可视化软件 CiteSpace，

梳理并呈现近十五年以来我国职业教育教师数字素养的研究现状、热点和前沿，为促进我国职教教师数字素养的研究与发展提供参考。

2. 研究设计

2.1. 研究的数据来源

本研究的数据来源于中国知网（CNKI）数据库，以“主题”为检索项，以“教师数字能力”or“教师数字素养”or“教师数字胜任力”or“教师信息化”or“教师信息素养”or“教师信息能力”为关键词（关键词来自马丁内斯布拉沃等人（Martínez-Bravo et al., 2020）对 1968 年至 2017 年数字素养元研究的综合分析内容，并基于本研究的参与者群体在关键词前加上“教师”），选定核心期刊和 CSSCI 期刊，设置检索时段为 2010 年 1 月 1 日—2024 年 3 月 25 日，学科为“职业教育”，共检索文献 129 篇。经精简并处理后，得到有效文献 51 篇。

2.2. 研究方法与工具

CiteSpace 知识图谱具有多元、分时、动态的特征，可以将一个知识的演进历程集中展现在一个网络图谱上，并自动将图谱上引文节点文献和共引聚类的研究前沿表示出来（陈悦等人，2015）。本研究以文献计量分析为主要研究方法，运用 CiteSpace .6.3.R1 软件对所获取的有效文献进行知识图谱分析，并采用 Excel 进行辅助统计。通过分析职教教师数字素养研究的时间分布、研究机构、核心作者、关键词等方面，揭示教师数字素养的研究现状、研究热点和研究趋势，使得未来研究更加有目的性和方向性。

3. 研究结果及分析

3.1. 研究概况分析

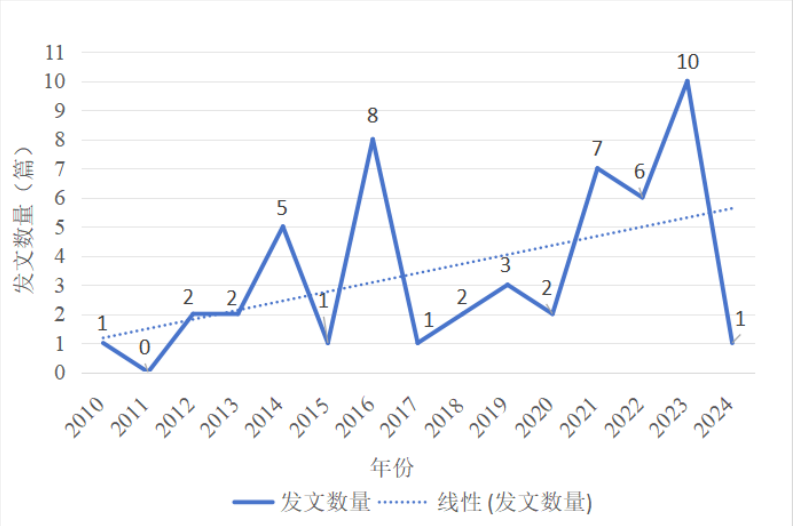
为全面了解职教教师数字素养领域的研究现状，本研究统计了 2010—2024 年相关文献的发表情况，并绘制了如图 1 所示的文献数量变化趋势图。

图 1 显示，2010 至 2024 年间，职教教师数字素养相关论文数量持续上升，尤其在 2016 年和 2021 年显著增长，并 2023 年达到顶峰，共发表 10 篇论文。中华人民共和国教育部（2010）颁布《国家中长期教育改革和发展规划纲要（2010—2020 年）》，确立了到 2020 年基本实现教育现代化的目标。教育信息化的发展程度决定着教育现代化的实现，而教师的信息化水平则是教育信息化的核心要素。在此背景下，我国开始深入探究职教教师的数字素养，开启了职业教育信息化发展的新篇章。

国务院（2014）发布的《国务院关于加快发展现代职业教育的决定》中明确指出：“加强现代信息技术应用能力培训，将现代信息技术应用能力作为教师评聘考核的重要依据。”教育部（2016）印发的《教育信息化“十三五”规划》强调将现代信息技术融入职业教育，促进教育体系重构与教学方式革新。这促使各地各校积极开展职教教师数字素养的研究与实践，导致 2016 年相关研究迅猛增长。教育部等九部门（2020）联合发布《职业教育提质培优行动计划（2020-2023 年）》，明确提出实施职业教育信息化 2.0 建设行动要大力提升职业教育信息化建设水平，推动信息技术与职业教育教学的深层次融合，全面提升教师的信息素养和能力。同时，全球数字技术的快速发展和新冠疫情的影响使在线教学成为新常态，促使职教教师数字素养研究在 2021 年迅猛增长。

图 1

2010-2024 年职教教师数字素养文献发表数量



3.2. 研究热点分析

在文献计量学领域，一般通过对高频关键词的分析，可以发现学科的研究热点（姜春林等人，2008）。此外，中心性体现的是节点对图谱中关键词网络的控制能力，一定程度上能够代表研究的热点领域。在 CiteSpace 中设定节点类型为“keyword”，并设置 TopN=50，TopN%=10，运行后生成关键词共现图谱（如图 2），并进一步统计排名前 20 的高频关键词的频次和中心性，如表 1 所示，以此确定我国职教教师信数字素养领域的研究热点。

表 1

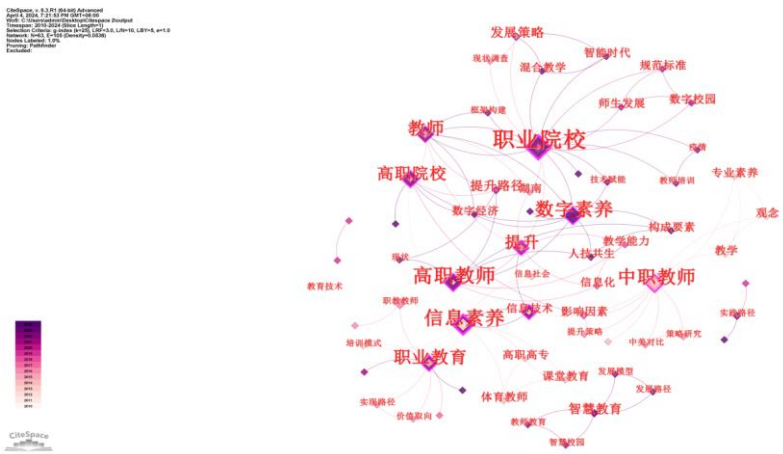
2010-2024 年职教教师数字素养文献关键词共现频次及中心性

序号	关键词	频次	中心性	序号	关键词	频次	中心性
1	职业院校	10	0.5	11	发展策略	2	0
2	信息素养	8	0.58	12	提升路径	2	0.02
3	高职教师	7	0.28	13	职教教师	2	0.05
4	高职院校	7	0.31	14	智慧教育	2	0
5	中职教师	6	0.32	15	影响因素	2	0.09
6	职业教育	6	0.35	16	实践路径	2	0
7	数字素养	5	0.25	17	教育技术	1	0
8	教师	5	0.13	18	教学	1	0
9	信息技术	4	0.32	19	湖南	1	0
10	提升	3	0.29	20	实现路径	1	0

根据图 2 所示的聚类结果和表 1 所示的统计结果，职业院校、信息素养、高职教师、高职院校、中职教师、职业教育、数字素养、教师、信息技术、提升等关键词都是研究热点，但是，考虑到职业院校、信息素养、职业教育等词专指度较低，本研究通过归纳分析，对具有高度中心性和高频次的节点进行整合，将我国职教教师数字素养研究热点概括为四个方面：一是数字素养的丰富内涵与构成要素；二是数字化教学的现状、困境、影响因素及提升路径；三是数字素养的评价指标体系构建；四是数字素养的国际经验及比较研究。

图 2

2010-2024 年职教教师数字素养关键词共现图



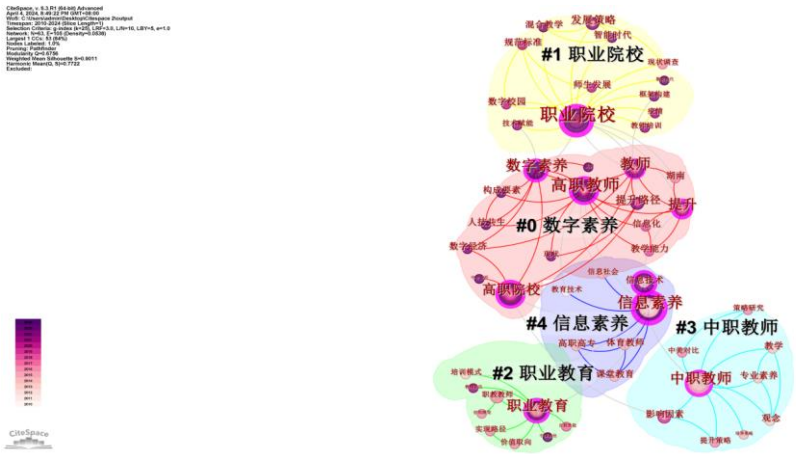
3.3. 研究现状分析

高频关键词共现网络图谱只表征研究热点外在的显性特征，而关键词聚类视图能够体现聚类间的结构特征,突出关键节点及重要链接。运用 CiteSpace 绘制关键词聚类图谱（见图 3）。

图 3 显示，数字素养、职业院校、职业教育、中职教师、信息素养为职教教师数字素养研究的主体。第一聚类包括关键词：数字素养、高职教师、数字素养、提升路径、提升、高职院校、教学能力、人技共生、数字经济、构成要素等；第二聚类包括关键词：职业院校、发展策略、规范标准、框架构建、现状调查、数字校园、教师培训等；第三聚类包括关键词：职业教育、培训模式、价值取向、实现路径、自我效能等；第四聚类包括关键词：中职教师、专业素养、策略研究、影响因素、观念、提升策略、中美对比等；第五聚类包括关键词：信息素养、信息技术、高职高专、信息社会、教育技术、课堂教育等。此外，聚类阴影的位置反映了该关键词相关研究与其他研究的交汇点。具体而言，第一类数字素养与第五类信息素养存在显著的交叉，这表明两者之间存在相互影响的关系。同时，第三类职业教育与第五类信息素养的交叉也较为明显，显示出职业教育与信息素养研究的紧密联系。这些交叉现象为本研究揭示了不同研究领域之间的内在联系，为后续研究提供了有益的启示。

图 3

2010-2024 年职教教师数字素养文献关键词聚类图



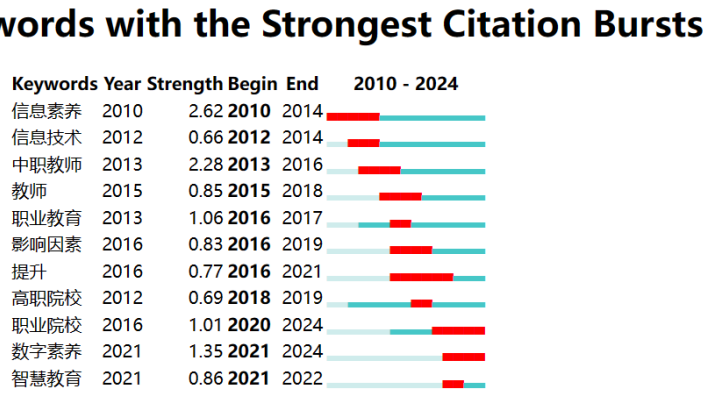
3.4. 研究趋势分析

图 4 显示，我国职教教师数字素养研究前沿关键词随时间不断演化，约每四至五年呈现一次突现。以突现的强度序列来看，排除类似于“教师”、“职业教育”“职业院校”等指向性过于宽泛的词语，可以得到高突现率的词汇有“中职教师”（突现强度 2.28）、“智慧教育”（突现强度 0.86）、“影响因素”（突现强度 0.83）、“提升”（突现强度 0.77）、“高职院校”（突现强度 0.69）、“信息技术”（突现强度 0.66）等，说明这些词汇在 2010-2024 年间均是研究者集中关注的前沿主题。

本研究综合 CiteSpace 生成的突发性关键词列表，划分出我国职教教师数字素养研究前沿的两个阶段。一是研究的早期阶段（2010-2020 年），筛选的突现词有“信息技术”、“影响因素”、“提升”，可见这一阶段的研究围绕职教教师的信息技术、信息素养、信息化教学，具体包括教师信息技术的应用现状、信息素养的提升及信息化教学能力的培养。二是研究的近期阶段（2021 年-2024 年），筛选的突现词有“数字素养”、“智慧教育”，说明这一阶段的研究集中在职教教师数字素养、数字胜任力，具体包括教师数字教学能力提升策略、数字胜任力框架构建、数字素养提升路径。从内涵发展来看，教师数字素养的内涵经历了从信息素养（Digital Literacy）到教师数字技能（Digital Skills）到教师数字素养（DigitalLiteracy）再到教师数字胜任力（Digital Competence）的演变过程。

图 4

2010-2024 年职教教师数字素养文献关键词突现率图



4. 总结与启示

随着科技创新的不断推进和产业的深度转型，新质生产力已经成为推动经济社会发展的关键因素。职业教育作为与产业发展紧密相关的教育类型，其数字化转型和创新直接关系到新质生产力的发展水平。教师作为现代职业教育数字化建设的核心要素，其数字素养的缺失是制约发展的关键。因此，深入研究职教教师数字素养对职业教育赋能新质生产力、提升职业教育的适应性和创新性具有重要意义。

本文综合高频关键词统计及前沿分析，我国职教教师数字素养研究主要聚焦于教师数字素养的内涵、要素、实现路径等，尚处于探索阶段，理论丰富但实践不足。未来，我国职教教师数字素养研究可进一步深化和拓展。其一，加强实践应用研究，将理论成果转化为具体的教学实践，探索提升教师教学素养的实践路径。其二，深化对数字化教学影响因素及实施路径的研究，为优化教学环境和提升教学质量提供科学依据。其三，关注教师数字素养发展的动态变化，及时应对新挑战和新问题。其四，构建完善的教师数字素养标准框架和评价体系，

为教师的专业发展提供明确指导。综上，未来研究应更加注重理论与实践的结合，推动职教教师数字素养研究的全面发展。

参考文献

中华人民共和国教育部（2010）。《国家中长期教育改革和发展规划纲要（2010—2020 年）》。http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/201008/t20100802_93704.html

中华人民共和国教育部（2016）。关于印发《教育信息化“十三五”规划》的通知。http://www.moe.gov.cn/srcsite/A16/s3342/201606/t20160622_269367.html

陈悦、陈超美、刘则渊、胡志刚、王贤文(2015)。CiteSpace 知识图谱的方法论功能。**科学学研究，(02)，**242-253。

姜春林、杜维滨、李江波（2008）。经济学研究热点领域知识图谱:共词分析视角。**情报杂志，(09)**，78-80+157。

教育部、国家发展改革委、工业和信息化部、财政部、人力资源社会保障部、农业农村部、国务院国资委、国家税务总局、国务院扶贫办（2020）。关于印发《职业教育提质培优行动计划（2020—2023）》。http://www.moe.gov.cn/srcsite/A07/zcs_zhgg/202009/t20200929_492299.html

李晓娟、王屹（2021）。技术赋能：职业院校教师数字素养的要义、挑战及提升。**中国职业技术教育，23**，31-37+45。

许倩倩、吴雪萍（2023）。数字化转型背景下职业院校教师数字胜任力：发展逻辑、内涵要素与提升策略。**职业技术教育，44(23)**，13-20。

教育部职业教育与成人教育司（2022）。关于印发《职业教育与继续教育 2022 年工作要点》的通知。**教职成司函，1 号**，03-14。

国务院（2014）。关于加快发展现代职业教育的决定。http://www.moe.gov.cn/srcsite/A16/s3342/201203/t20120313_133322.html

Martínez-Bravo, M. C., Sádaba-Chalezquer, C., & Serrano-Puche, J. (2020). Fifty years of digital literacy studies: A meta-research for interdisciplinary and conceptual convergence. *Profesional De La Información, 29(4)*, 25.

英国数字化推动教师专业发展的变革契机、变革框架与借鉴启思

Window of Opportunity for Transformation，A Framework for Change and the Enlightenment for Digitally Driven Teacher Professional Development in the United Kingdom

梁冰冰，叶建宏*
北京师范大学教育学部
* kimpo30107@hotmail.com

【摘要】 在教育数字化时代，教师专业发展离不开数字技术的支持。英国因教育政策“断层”下的发展诉求、设计公平可行教育政策的目标指引及教育公平的需求与制度公正的重塑的需要，积极通过数字化浪潮推动教师发展，出台了一系列教育政策积极作为，打造“数字化+教师”的发展理念，建立较为先进可靠的数字基础设施建设，形成了高效简捷稳定的教育数字系统三维变革框架，并明晰了教育数字化推动高质量发展的未来路向：明确教师职责，提升教师数字胜任力，以数字安全为旨归，保障教师发展环境。

【关键词】 *教育数字化 教师专业发展 变革框架 借鉴启思*

Abstract: *In the era of digitalization of education, the professional development of teachers cannot be separated from the support of digital technology. The United Kingdom actively promotes the development of teachers through digitalization because of the educational policy development, the design of fair and feasible education policy target guidelines, the need for fairness in education, and the need to reshape the system of justice. A series of education policies has been introduced to actively act on the development concept of “digitalization + teachers” to establish a more advanced and reliable digital infrastructure, to form a three-dimensional framework for the transformation of an efficient, simple and stable education digital system, and to make the future direction of the digitalization of education to promote high-quality development: clarifying the responsibilities of teachers, improving their digital competence, and safeguarding the environment for teacher development with digital security.*

Keywords: Digitization of education, Teacher Professional Development, Framework for Change, Enlightenment

1. 引言

数字化转型作为一种系统性变革，正在形塑社会发展的各维度和层次。无远弗届的数字化发展也正在显著影响着教育的发展，数字技术作为教育生态系统架构的底层技术逻辑存在，对教育发展具有建构性价值，正在升级且重塑教育方式、知识感知等方方面面。根据摩尔定律，数字技术变革的指数速度已经超过了社会、政治和许多经济机构的适应能力（Hanna, 2018），如何及时适应数字技术变革，对这一问题的回答不仅直接关系到我国人才培养的质量和层次，也深刻影响我国结构性变革的持续发展。

数字化转型方兴未艾，其研究仍如火如荼地进行着。理清数字化专业概念的相关研究，是我们持续开展研究的基础。关于数字化转型的认识主要有两个维度，一是数字化转型是一种策略、过程或者模式（Cichosz et al., 2020），二是数字化是人与社会之间的交流和互动范式的改变（Kaur, 2021）。这两种观点以技术功能角度予以审视，数字化无论是一种过程亦或

是一种变化，其本质都是由技术推动。教育数字化的发展，就是作为技术物的教育数字媒介的集合作用于教育发展的过程，不仅仅是具有表象上的功能可供性，亦有承载着主体与技术关系的嵌合性。

教育数字化的发展与教师专业发展是矛盾统一的关系。教育数字化重构教师专业发展实践形态。首先，教育数字化促进教师发展素质，教育数字化为教师服务提供更为广阔及丰富的知识存储库，以技术形式推动教师发展之间的交流与合作，有助于教师专业发展。其次，教育数字化改变教师知识传播方式。教育数字化以技术行为改变了教育传播的媒介，改变了以往以图文书本为主的知识传播方式，增加以音频图像等多种传播媒介形式。最后，教育数字化在一定程度上会挑战教师权威。教育数字化改变了教师对于知识的垄断性地位，由原来的教师主导地位变为“师-生-机”为主体的三角形教育关系，这在一定程度上需要教师让渡一部分课堂权威，主动融入教育数字化浪潮之中，与时俱进更新知识观点，依托数字技术创新优化教育教学活动（董云川、韦玲，2021）。同时，教育数字化也在不断适应教师专业发展。教育数字化的本质是技术，教师专业发展在使用数字化过程中对数字技术并非全盘接受，会因自身的技术偏好、发展需求等对技术物予以改造，最终技术物也会因其人性化趋势不断印证进而加强其物质技术性发展水平。

我国教育数字化在推动教师专业发展中仍存在着一些明显劣势，我国教育发展整体不均衡，在数字化设施建设方面存在着一些资源型不均衡的问题；数字化助力教育专业发展需要时间适应，现阶段教育数字化发展仍未得到完全熟稔适应，会产生由于获取和使用数字资源的程度不同而导致的“数字鸿沟”等问题。因此本研究借鉴英国教师专业发展的数字化经验，以期为我国教师专业发展提供镜鉴启思。

2. 英国教育数字化推动专业发展的变革契机

英国教育数字化的发展并非是一蹴而就，而是在现实教育情境下产生一系列问题，通过专业人员对问题的识别，经由政府管理者及民众共同推动，进而使得教育数字化推动教师专业发展成为可能。

2.1 问题契机：教育政策“断层”下的发展诉求

英国教师专业发展中的问题是阻碍其教育变革的阻力，但对问题的识别与寻求解决的决心是运用教育数字化推动教师专业发展的重要契机。

首先是教育与技术之争。教育数字化转型归根结底其本质是数字技术在教育领域中的内嵌与耦合。一方面，数字技术尚不稳定，在教育中的运用存在诸多不稳定因素。正如“科林格里奇困境”对当代技术和社会发展之间张力关系的描述：我们往往限于时代和知识等的局限性，难以预测和控制数字技术的“非预期后果”（贾向桐、胡杨，2021），故教育数字化在推动社会变革时亦存在诸多风险，难以融入于教师专业发展中；且数字技术的学习运用需要一定的周期，使用主体之间能力有差，无法兼顾教育公平。另一方面，数字数字化的运用易流于表面形式，忽视教育性，在提高教学效率、创新教育方式的同时，其中的不确定性风险也悄然而至（董倍宏、杨剑，2023）。

其次是英国政府对“数字经济”高地的抢占。新一轮信息技术的发展成为催生教育结构性变革的驱动动因，持续改变着知识生产方式（祝智庭等人，2023），影响着经济社会与人类文明的发展进程。英国政府认识到跻身于全球数字化创新浪潮前列的重要性，这不仅能为后续发展创造基本条件，也是提升国家竞争力的关键（丁声一等人，2016），数字经济的发展需要高技术技能型人才，而学校教育是培养优秀技术技能型人才的重要环境，教师质量在其中发展着至关重要的作用。

最后是教育质量的隐忧。近年来，英国在国际学生评估计划中，数学和阅读成绩始终未能超越国际平均水平，国际测试的欠佳表现引发了社会各界对教师质量的担忧（谭娟、饶从满，2019）。这一学业成绩表现差强人意，也在一定程度上反映出英国教师团队质量问题，英国近年来难以雇佣合适的任课教师，教师的专业水准不断下降，教学质量难以达到标准，这是英国教师队伍亟需解决的问题。

2.2 政策契机：设计公平可行教育政策的目标指引

面对英国教师专业发展现状，英国政府及相关群体并非无所作为，而是发挥各自角色，积极推动公平可行的教育政策的出台。

英国高校认识到数字技术对组织文化的影响，积极建立以数字为核心的组织文化，推动组织变革和有效治理。同时，教育数字化的发展不仅仅技术问题，还涉及组织文化和治理结构的变化，英国高校鼓励跨部门合作，打破传统边界，促进知识创造与创新、知识管理和运用、知识交流与合作关系的发展。

学术界对文化目标的寻求。首先，学者们以理论视角关注教育数字化发展过程，利用自身知识深度与经验，通过发表文章、文化宣讲等方式引起社会对于教育数字化重视，而且，学者们可以通过为政府提供政策建议，成为教育数字化发展中的“智库”，利用自身知识积极开展对教育发展的课程研讨或培训，最后，学术界持续评估教育数字化的效果，通过研究和反馈机制，不断优化和改进教育技术的应用。

2.3 政治契机：教育公平的需求与制度公正的重塑

在分析英国如何成为教育数字化转型的领导者时，需要关注公众对于这一变革的态度变化，还要考虑执政党如何通过其治理能力和理念来引导和推动教师的专业发展，确保教育系统能够有效地适应和利用数字化技术。这意味着，英国在推动教育数字化转型的过程中，既要确保政策得到民众的理解和支持，也要展现执政党的领导力和清晰的政策方向。

公众的意见和情绪在很大程度上能够影响政策的制定和优先级，这可以被看作是国家氛围或民意的体现。数字经济的发展为提高教学效率和学习体验提供了新的基础，但也带来诸多挑战和担忧。英国大部分公民认为自身不具备工作岗位所要求的新的数字素养，这种技能短缺不仅影响了个人的就业前景，也对国家的经济发展和社会稳定构成了潜在威胁。教育工作者同样面对这一问题，教师们呼吁政府采取更加积极的措施，以确保教育系统能够有效地支持数字化转型。此外，教育专家和学者也在探索如何通过教育改革来解决技能差距问题，教育体系需要迅速响应技术变革和市场需求的变化，这包括更新课程内容，引入更多与数字技术相关的学科，以及提供持续的职业发展和终身学习机会。

执政党的执政理念与决策变化是政治源流中的核心推力，其理念的转变推动了英国教育数字化转型（曲铁华、王洪晶，2022），在英国政治体系中，执政党的指导思想和政策决策对于国家政治发展具有决定性的作用，随着治理理念的更新和变化，英国教育体系也随之发生了变革，特别是在采纳和利用数字技术方面，以实现教育的现代化和数字化，执政党的思想和决策直接影响了教育领域的发展方向，推动了教育向数字化转型的进程。尽管英国执政党不断迭代变化，但是每个时期的英国政府均十分关注教育数字化的发展，教育政策的演进与发展具有连续性。英国的数字化转型工作最先也是开始于 2000 年左右的政务电子化、信息化活动，随着数字经济的不断发展，英国数字发展政策也在不断完善和改进，2022 年英国发布《英国数字战略》，重点强调英国数字基础、创意和知识产权、数字技能和人才，为数字增长融资等内容（GOV.UK，2022）。

3. 英国教育数字化推动教师专业发展的变革框架

英国教育数字化推动教师专业发展是一个系统工程，其着力点不仅仅在于改善基础设施建设，还在于理念层面的转变及稳定的数字教学系统。英国教育数字化推动教师专业发展变革框架主要有三个层级，理念层面是数字化应用于教师专业发展的理论基础和指导思想，主要指教育观念的更新和思维方式的塑造；环境设备层主要是教育数字化中对教师专业发展所提供的必要的硬件设施和技术支持，完善的基础设施是教师能够开展数字化教学和学习活动的基础；系统发展层是对于教师专业发展各个要素的系统化整合，确保教师可以通过参与在线课程、远程培训等方式获得专业发展，这是数字化转型应用于教师专业发展的关键点，也是教育改革的最终目标所在。

3.1 变革理念层：打造"数字化+教师"发展理念

只有在理念层面进行变革，教育数字化才能更好地推动教师专业发展（冯晓英等人，2021）。一是通过智能技术的集成和应用，教师能够更加精准地识别和满足学生的学习需求。例如，个性化学习路径的设计与实施，使得教师可以根据每个学生的学习数据和进度，提供定制化的教学支持和资源，从而提高教学的针对性和有效性。二是智能技术还为教师提供了强大的数据分析和反馈工具，使他们能够实时监测学生的学习进度和理解程度。这种即时反馈机制不仅帮助学生及时调整学习策略，也为教师提供了改进教学方法的依据，促进了教学与评估的紧密结合。此外，智能技术促进了教育资源的共享和优化配置，教师可以通过在线平台和协作工具，与其他教育工作者共享最佳实践、教学材料和研究成果。这种跨界合作和知识共享，有助于提升教师的专业能力和创新教学方法。再次，智能技术还为教师提供了新的职业发展机会，如在线教育、远程培训和虚拟课堂等新兴领域。这些新机会不仅拓宽了教师的职业道路，也要求教师群体不断更新技能和知识，以适应教育领域的快速变化。最后，智能技术的应用还有助于构建更加开放和包容的教育环境，通过提供无障碍的学习工具和资源，确保所有学生都能获得高质量的教育，无论他们的背景或能力如何。这种普及化的教育方式，有助于实现教育公平和终身学习的目标。

3.2 支撑性力量：较先进可靠的数字基础设施建设

互联网、大数据、人工智能等技术提供了丰富的在线资源和平台，使教师能够随时随地进行学习和研究。这些环境不仅包括在线课程和研讨会，还有专业的教育论坛和社区，教师可以在这些平台上交流思想、分享经验，并获取最新的教育资讯和研究成果。现代技术如智能教室、虚拟现实（virtual reality, VR）和增强现实（augmented reality, AR）等，为教师提供了模拟和实验教学的新工具。这些工具可以帮助教师更有效地展示复杂的概念，提供互动和沉浸式的学习体验，从而提高学生的学习兴趣和参与度。学习分析和人工智能技术可以对教师的教学活动进行跟踪和评估，提供关于教学效果的反馈。这些数据驱动的评价工具可以帮助教师了解自己的教学强项和需要改进的地方，从而有针对性地调整教学策略和方法。

3.3 核心竞争力：高效简捷稳定的教育数字系统

在数字化时代，教师专业发展需要一个全面的数字学习生态系统以帮助教师适应并利用数字技术，实现专业自主发展。首先，创建教师数字学习生态，建立系统性的学习环境，其中包含数字学习平台、资源、场景和组织结构，旨在满足教师的学习需求，并支持教师通过数字技术进行创新学习，教学生态系统鼓励开放性、适应性和持续性，以支持教师的专业发展需求；其次，定向开发教师数字服务，根据教师的专业发展需求，定制和提供数字服务，英国正在加强公共数字服务体系建设，为教师提供智能推送的、针对性的数字学习服务，以满足教师多元化和个性化的学习需求；最后，提升教师数字素养，这强调了教师掌握数字技

术的重要性，包括使用数字设备、处理数据、开发教育资源、遵守数字伦理和增强数据安全意识，英国将数字资源和服务融入教学设计、课堂管理和评价中，以促进课堂教学和课程育人的创新；此外，教师需要具备数字领导力，即利用数字技术进行沟通、团队建设和技术变革的能力。

4. 英国教育数字化推动教师高质量发展对我国的借鉴启思

教育数字化不仅仅是教育领域的一个附加元素，而且是推动整个教育系统变革的关键力量。教师的数字胜任力，即教师有效利用数字技术进行教学的能力，正处于教育变革和技术创新的交汇处。这种能力不仅是数字化时代教育改革的自然结果，也是英国对能够适应并推动教育发展的高素质教师的要求，英国通过强化教师的数字技能，推动教育系统的现代化，培养能够在全球舞台上竞争的人才，从而支持国家的长远发展。

4.1. 明确教师职责，加强教师专业素养

教师不仅是知识的传递者，还是学生学习过程的引导者、激励者和支持者。在教育的各个方面，教师都需要有一个清晰的自我认识和定位，这有助于他们更好地履行教育职责和提升教学质量。英国强调教师的责任和使命是整合数字知识并生成数字智慧，目的是培养出能够在数字化世界中成功竞争的学生。这要求教师不仅要有深厚的专业知识，还要有前瞻性的教育理念和创新的教学方法；教师不仅要掌握数字技术，还要能够在教学中创造和应用这些技术；教师需要不断学习和提升自己的专业知识、教学技能和教育理论，同时也要培养良好的职业道德，如对学生的关爱、对教育事业的责任心和对个人行为的自律，这些素质共同构成了教师的“立身之本”，是教师在教育工作中取得成功和赢得尊重的关键。通过强化这些基本素质，教师能够更有效地促进学生的全面发展和社会的进步。

4.2 提升教师数字胜任力，助力教师专业发展

提升教师的数字胜任力不仅是英国推动数字化时代教育变革、培养合格数字人才的迫切需要，也是多个国家及国际组织就教师数字能力要求达成的共识。教师数字胜任力的提升是一个多维度的发展过程，主要有三个领域，一是认知领域，教师需要加深对数字技术的理解，建构完善的数字知识体系，清楚地认识到数字胜任力对于提升教学质量和学生学习体验的重要性。二是技能领域，在学校和教育机构内部建立积极的数字文化和技能支持系统，推动教育数字化技能经验交流和资源共享；三是社会领域，教师需要提升对数字技术的社会责任感，使用数字技术时遵守法律和道德规范，以自身发展能力为基点影响社会数字化能力，主动承担起培养下一代数字公民的责任。

4.3 以数字安全为旨归，保障教师发展环境

数字技术具有双刃剑特征，虽然数字技术带来了便利和效率，但同时也带来了诸多数据安全问题。首先，政府应当从政策角度出台相关法律法规以保障数据安全，为保障数据安全提供参照依准，定期对学校的网络和信息系统进行安全审计，检查潜在的安全威胁，并采取相应的预防和修复措施；其次，组织专业的网络安全培训课程，涵盖网络安全基础、网络体系结构、计算机病毒防护等内容，确保教师具备必要的网络安全知识和技能，提高教师的安全防范能力；最后，要建立网络安全事件应急响应计划，增强教师的信息安全处理能力和水平。

参考文献

丁声一、谢思淼、刘晓光（2016）。英国《数字经济战略(2015-2018)》述评及启示。**电子政务，04**，91-97。

冯晓英、郭婉瑢、黄洛颖（2021）。智能时代的教师专业发展：挑战与路径。**中国远程教育，11**，1-8+76。

曲铁华、王洪晶（2022）。新世纪英国教师教育改革探析。外国教育研究，49（10），31-46.

李振东、陈劲、王伟楠（2023）。国家数字化发展战略路径、理论框架与逻辑探析。**科研管理，44**（07），1-10.

祝智庭、赵晓伟、沈书生（2023）。技能本位的学习范式：教育数字化转型的认识论新见解。**电化教育研究，44**（02），36-46。

贾向桐、胡杨（2021）。从技术控制的工具论到存在论视域的转变——析科林格里奇困境及其解答路径问题。**科学与社会，03**，26-39。

董云川、韦玲（2021）。人工智能促进高等教育发展的伦理纠偏。**重庆高教研究，9**（02），51-58。

董倍宏、杨剑（2023）。教育数字化转型：不确定性风险及其治理。**电化教育研究，44**（11），52-59。

谭娟，饶从满（2019）。英国基础教育教师队伍建设的现实困境与改革对策。**外国中小学教育，10**，64-72。

Cichosz, M., Wallenburg, C. M., & Knemeyer, A, M. (2020). Digital transformation at logistics service providers: Barriers, success factors and leading practices. The International *Journal of Logistics Management, 31*(2),209-238.

GOV.UK (2022). *UK digital strategy*. https://www.gov.uk/government/publications/uks-digital-strategy/uk-digital-strategy

Hanna, N. (2018). A role for the state in the digital age. *Journal of Innovation and Entrepreneurship, 7*(1), 5-15.

Kaur, N. (2021). Digitalization and work behaviour: A paradigm shift. In A. Singh (ed.), *Big data analytics for improved accuracy, efficiency, and decision making in digital marketing* (pp. 102-123). IGI Global.

Fine-Tuning Large Language Models for Educational Support: Leveraging Gagné's Nine Events of Instruction for Lesson Planning

Linzhao Jia[1, 2, 3, 4, *], Changyong Qi[1, 2, 3], Yuang Wei[1, 2, 3], Han Sun[1, 2, 3, 4], Xiaozhe Yang[1, 2, 4]

[1] Lab of Artificial Intelligence for Education, East China Normal University

[2] Shanghai Institute of Artificial Intelligence for Education, East China Normal University

[3] School of Computer Science and Technology, East China Normal University

[4] Institute of Curriculum and Instruction & Classroom Analysis Lab, East China Normal University

[*] lzjia@stu.ecnu.edu.cn

Abstract: *Effective lesson planning is crucial in education process, serving as the cornerstone for high-quality teaching and the cultivation of a conducive learning atmosphere. This study investigates how Large Language Models (LLMs) can enhance teacher preparation by incorporating them with Gagné's Nine Events of Instruction, especially in the field of mathematics education in compulsory education. It investigates two distinct methodologies: the development of Chain of Thought (CoT) prompts to direct LLMs in generating content that aligns with instructional events, and the application of fine-tuning approaches like Low-Rank Adaptation (LoRA) to enhance model performance. This research starts with creating a comprehensive dataset based on math curriculum standards and Gagné's instructional events. The first method involves crafting CoT-optimized prompts to generate detailed, logically coherent responses from LLMs, improving their ability to create educationally relevant content. The second method uses specialized datasets to fine-tune open-source models, enhancing their educational content generation and analysis capabilities. This study contributes to the evolving dialogue on the integration of AI in education, illustrating innovative strategies for leveraging LLMs to bolster teaching and learning processes.*

Keywords: lesson planning, large language models, Chain of Thought, Gagné's Nine Events of Instruction, fine-tuning

1. Introduction

The advent of large language models (LLMs) has revolutionized various domains, notably in education. This development is not only reshaping educational methodologies and analytics but also highlights a growing trend: an increasing number of individuals are leveraging LLMs, including tools like ChatGPT, to analyze teacher dialogue in classroom settings or assist teachers in lesson preparation (Forero & Herrera-Suárez, 2023; Sun & Hoelscher, 2023; Uchiyama, Umemura, & Morita, 2023). However, the current landscape lacks a streamlined pipeline for educators to efficiently leverage LLMs, posing challenges in directly applying these models for instructional support. Given that LLMs are not specifically trained on educational datasets, educators are required to invest time in designing prompts, which may not always yield satisfactory results. Our research aims to bridge this gap by proposing a comprehensive pipeline that not only standardizes the use of LLMs in educational settings but also optimizes them for educational content, enhancing their utility in teaching design and evaluation.

In this study, we aim to fine-tune LLMs through *prompt tuning* and *model tuning* within the framework of *Gagné's Nine Events of Instruction*, thereby assisting teachers in classroom dialogue assessment or in generating classroom dialogue teaching plans. Robert Gagné is considered to be the foremost researcher and contributor to the systematic approach to instructional design and training (Kruse, n.d.). By integrating Gagné's Nine Events of instruction into the design of classroom dialogue, educators can leverage a proven instructional design framework to enhance the effectiveness of their teaching, aligning closely with the cognitive, affective, and psychomotor objectives outlined in Bloom's Taxonomy.

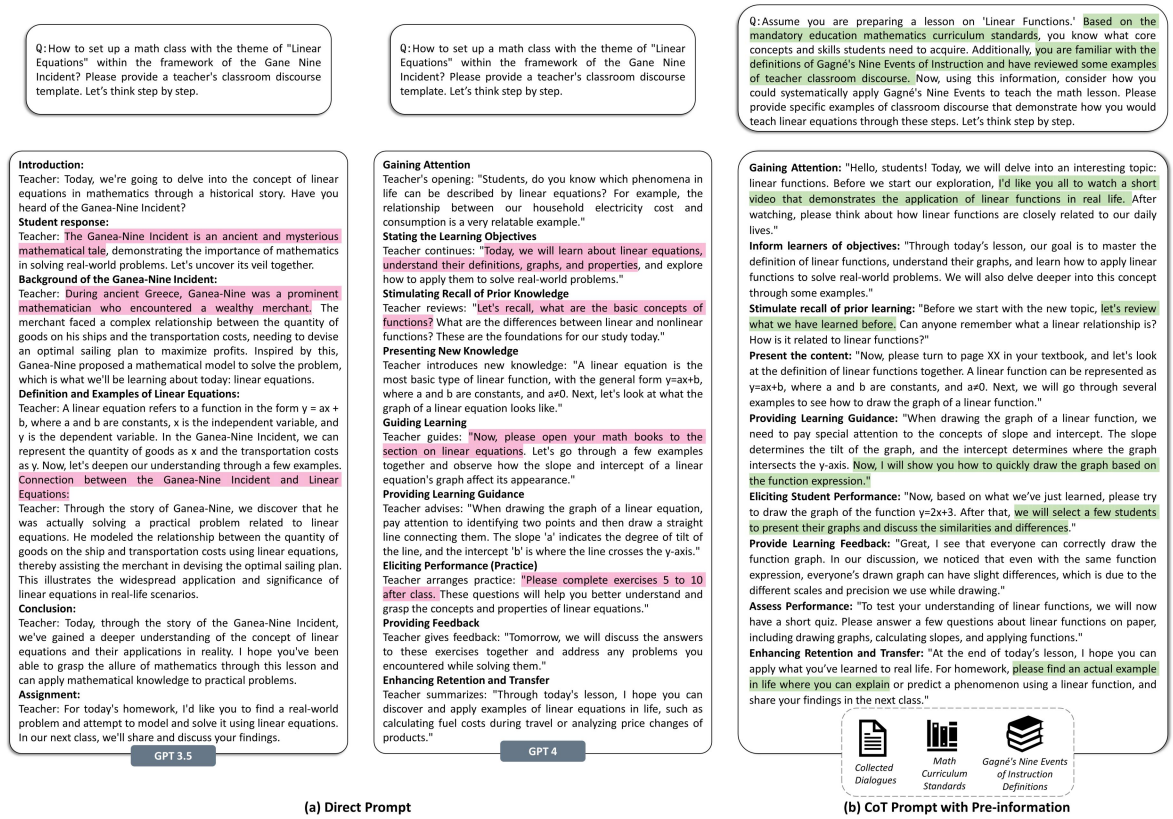


Figure 1. Comparative analysis of CoT and Direct Prompts

For prompt tuning, we design scenarios that prompt non-open-source models, like ChatGPT (OpenAI et al., 2024), to guide the LLMs to generate dialogue that enriches teaching and learning experiences. As shown in Figure 1, through a comparative analysis of CoT and direct prompts, this research aims to illuminate the benefits of employing structured, thought-provoking prompts to optimize the educational utility of language models in creating and evaluating teaching dialogue. As for model tuning, we fine-tuned on open-source models, including ChatGLM (Du et al., 2022) and Llama-Chinese (Touvron et al., 2023). Training these LLMs on our specially designed dataset imbues them with a deeper understanding of instructional design, significantly enhancing their capacity to produce content that is not only relevant but also robust. Our contributions to the field of instructional technology and educational methodology enhancement can be summarized as follows:

- **Open-access dataset:** We have developed an open-access dataset specifically tailored for training and evaluating LLMs within the context of educational dialogue generation. This dataset is designed to encapsulate a wide range of instructional scenarios, facilitating the development of LLMs capable of understanding and generating pedagogically valuable content.
- **Novelty Pipeline:** Our research introduces a novel pipeline for employing LLMs as central tools in advancing educational methodologies. By integrating structured prompt tuning and fine-tuning processes, we aim to unlock the full potential of LLMs in creating dynamic, interactive, and enriching teaching and learning experiences.
- **Fine-tuned LLM:** The fine-tuned LLMs resulting from our research can be seamlessly integrated into classroom analysis systems. These models have been specifically enhanced to comprehend and generate educational content, making them invaluable assets for teachers and educational content creators. By embedding these models within classroom analysis systems, educators can leverage real-time, AI-driven insights to tailor teaching strategies and content, thereby optimizing the teaching impact.

2. Related work

2.1 Large Language Models (LLMs) in Education:

The integration of LLMs into educational practices represents a significant leap forward in the use of artificial intelligence to enhance teaching and learning processes. The study (Golchin, Garuda, Impey, & Wenger, 2024) assesses GPT-4 and GPT-3.5's grading of MOOC courses using enhanced Zero-shot-CoT prompting, showing GPT-4's superior alignment with instructor grades, particularly in subjects requiring imaginative thinking. Long et al., (Long, Luo, & Zhang, 2024) explored the application of GPT-4 in labeling classroom dialogues. The findings demonstrate that employing GPT-4 for this task not only achieves a high degree of similarity with human labeling but also significantly reduces the time required for the process. Qian et al., (Qian, Shea, Li, Fryer, & Yu, 2023) developed a user-adaptive generative chatbot that facilitates language learning through constrained decoding, dynamically aligning its curriculum to user requests. Macina introduces MATHDIAL, a semi-synthetic dialogue dataset for math tutoring that enhances pedagogical effectiveness through equitable practices and scaffolding, offering over 3,000 conversations for benchmarking and advancing tutor response generation models, and setting a new standard in dialogue-based tutoring research in NLP (Macina et al., 2023).

2.2 Prompt Tuning and Model Tuning in Educational Settings:

Recent studies have illustrated that augmenting the number of tasks within constructed instructions can enhance the generalization capabilities of LLMs. Notably, prompting with Chain-of-Thought (CoT) (Wei et al., 2023) has been proven to effectively tackle intricate problems by systematically producing steps of reasoning. This methodology's success has sparked significant interest in the educational sector. Specifically, the Socratic teaching method (Dan et al., 2023), which guides students towards self-discovery and understanding by asking progressively complex questions, mirroring the stepwise elucidation process found in CoT prompting. Building on these insights, a study (Uchiyama et al., 2023) presents a system leveraging LLMs, notably ChatGPT, to facilitate preparation learning in flipped classrooms, featuring automatic student query responses and enhanced prompt clarity through video subtitles, aiming to mitigate common flipped classroom challenges and contribute to learning analytics through data collection.

In the domain of educational technology, the fine-tuning of large models has gained significant attention. Among these advancements, Dan et al. (2023) have introduced EduChat, notable for its innovative approach to interactive learning using fine-tuned language models. This system is distinguished by its ability to tailor educational content and provide dynamic, context-aware interactions with students. Collectively, these studies embody diverse strategies for leveraging language models, each offering unique perspectives on the depth, relevance, and educational utility of the generated content, thereby enriching the landscape of LLM application in educational technology.

3. Framework

In this section, we discuss the Dataset Preparation and Fine-tuning processes, which are essential for customizing Large Language Models (LLMs) to adeptly analyze and generate teacher dialogue aligned with Gagné's Nine Events of Instruction.

3.1. Dataset Preparation

The initial input data we used in this study mainly consists of three parts, (1) Definitions and examples of Gagné's Nine Events of Instruction, which serving as a foundational guide for model training and evaluation. (2) Math Curriculum Standards: By obtaining the official published curriculum standards and using manual extraction methods, the curriculum requirements for each knowledge point are organized and input to LLMs as a limiting condition for

generating teacher dialogue templates. (3) Collected Dialogues: Extract teacher dialogues and teacher-student interaction from real classroom environments, and manually annotate them as examples of real-life Gagné's teaching events. To obtain more diversity and comprehensiveness data, we manually write prompt templates and have GPT4 generate some examples of teacher dialogue templates. These prompts are designed to guide GPT-4 in generating teacher dialogue that is not only subject-specific but also structured around Gagné's instructional framework.

Table 1. Generated Dataset.

Gagné's Nine Events of Instruction	Example	Number
Gain attention	Let's begin with a quick challenge: Can anyone guess which shape we'll be exploring today with this image?	715
Inform learners of objectives	Today, we'll learn how to calculate the area of different triangles, a skill essential for many real-life applications.	422
Stimulate recall of prior learning	Think back to last week when we studied polygons. How does understanding their properties help in calculating the area of triangles?	769
Present the content	We'll look at formulas for the area of right, isosceles, and equilateral triangles using visuals and demonstrations.	768
Provide "learning guidance"	I'll guide you through each formula step-by-step. Remember, the base and height are key in these calculations.	862
Elicit performance	Now, let's try calculating the area of these triangles together. Start with this right triangle.	680
Provide feedback	That's a good attempt. Make sure to measure the height correctly—it's perpendicular to the base.	823
Assess performance	I'll hand out a short quiz now. You'll calculate the area of a few triangles to show your understanding.	476
Enhance retention and transfer to the job	As homework, find objects at home shaped like triangles and calculate their area. It'll help solidify what we learned today.	665

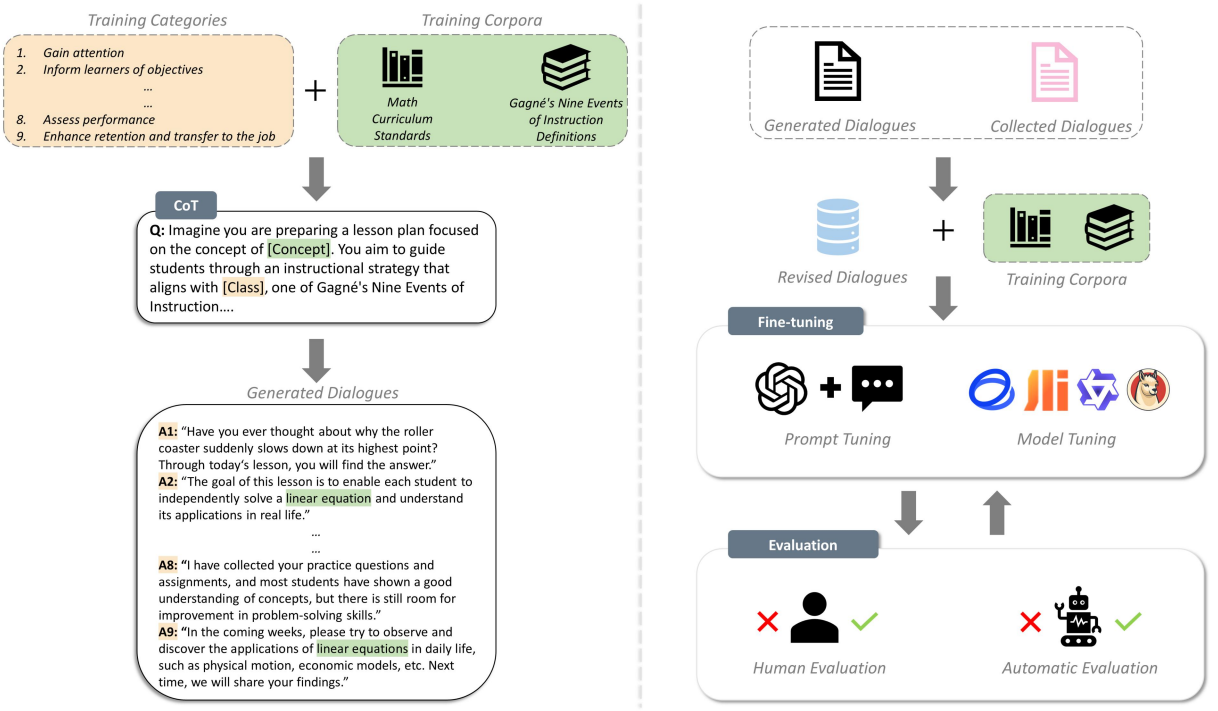
As shown in Figure 2.(a), the CoT prompt template is designed to lead the language model through a comprehensive thought process that covers all of Gagné's Nine Events, tailored to teaching the mathematical concept of linear equations. By replacing [Concept] with " linear equation " and adapting the template to each specific event [Class], the model can generate detailed and instructionally sound teacher dialogue for each stage of the lesson.

After the data generation phase, a team of educators and subject matter experts conducts a thorough review of GPT-4's outputs. This involves checking for accuracy, alignment with the curriculum standards, and adherence to Gagné's Nine Events. Any inaccuracies or misalignments are manually corrected or removed to ensure the final dataset comprises high-quality teacher dialogue templates. In Table 1, we present examples of the speech templates generated based on Gagné's nine events, along with a statistical count of each event. Due to the dataset being manually curated and revised, the quantities across categories are not uniform.

3.2. Fine-tuning

3.2.1. Prompt tuning

To enable LLMs to better handle teacher dialogue generation tasks that comply with the Gagné's Nine Events, we first build the agent by utilizing the ChatGPT API and connecting downstream tasks to complete this task. As shown in Figure 2.(b), we take the revised dialogues and training corpus as input, and employ the CoT method to guide the



(a) Data Preparation **(b) Fine-Tuning and Evaluation Pipeline for Dialogue Generation**

Figure 2. Illustration of the Process for Generating Educational Dialogue via LLMs

ChatGPT model towards generating outputs that adhere to Gagné's Nine Events. The CoT approach involves structuring prompts that not only direct the model's attention to the task but also encourage it to follow a logical sequence of thoughts. These prompts are designed to elicit detailed and structured responses from ChatGPT, ensuring that the generated dialogue is both pedagogically sound and closely aligned with the instructional objectives outlined in the curriculum standards.

For instance, to generate content that aligns with the 'Gain Attention' event, the CoT prompt might begin with a real-world problem that can be solved using the concept being taught, followed by questions or scenarios that lead ChatGPT to explore and elucidate the concept in a manner that would captivate students' interest. The output from ChatGPT, post-CoT prompting, is then evaluated for alignment with the instructional principles of Gagné's Nine Events. Through iterative refinement and tuning of the CoT prompts, we optimize the model's performance, ensuring that the generated teacher dialogues are aligned with the desired instructional outcomes. This methodological approach demonstrates a novel application of LLMs in educational content creation, enhancing the quality and efficacy of teacher-student interactions in the learning process.

3.2.2. Model tuning

In the realm of model fine-tuning for enhancing the generation of teacher classroom dialogue aligned with Gagné's Nine Events of Instruction, we adopt a nuanced approach that diverges from traditional extensive corpus retraining methods. Our strategy involves selectively adjusting critical parameters within the ChatGLM model, leveraging a retrieved dataset meticulously constructed to embody the essence of educational dialogue as delineated by Gagné's framework. Initially, we identify the pivotal aspects of Gagné's Nine Events that are most crucial for effective classroom dialogue. The constructed dataset serves as a comprehensive resource, embedding these instructional principles within authentic and synthesized educational contexts. This dataset not only provides the model with a rich variety of pedagogically aligned content but also exposes it to the nuanced dynamics of classroom interaction.

Subsequently, employing the ChatGLM model as our foundational framework, we delve into fine-tuning, specifically employing Low-Rank Adaptation (LoRA) techniques. This approach enables us to refine the model's parameters subtly, enhancing its ability to process and produce content that resonates with the instructional strategies inherent in Gagné's Nine Events. The ultimate goal is to validate the model's enhanced capability through a structured

experimental workflow, assessing its performance across various educational dialogues and scenarios. This fine-tuning endeavor marks a significant stride towards realizing the potential of LLMs in education, promising a future where AI-driven tools synergize with pedagogical methodologies to foster enriched learning experiences.

4. Experiments

4.1. Dataset

The dataset, central to our study, is meticulously crafted around Gagné's Nine Events of Instruction and includes a rich blend of content generated by ChatGPT and real teacher utterances gathered from live classroom settings. This comprehensive collection forms the backbone of our research, offering a nuanced perspective on educational dialogue. These utterances have been meticulously annotated and refined manually, providing a rich and diverse basis for our analysis.

4.2. Baseline Models

In the process of developing our prompt tuning model, we utilized GPT3.5 and GPT-4 (Brown et al., 2020; OpenAI et al., 2024). GPT-3.5, released by OpenAI, is acclaimed for its exceptional generative power and wide-ranging applications, symbolizing a crucial advancement in language model development. Following this, GPT-4 was introduced with substantial refinements in model architecture, training dataset breadth, and diversity, thereby significantly boosting the performance capabilities of the model.

Continuing, we turn our attention to open-access models that can be deployed locally, including Llama2-7B (Touvron et al., 2023), ChatGLM3-7B (Du et al., 2022), Qwen-7B (Bai et al., 2023), and Baichuan2-7B (Yang et al., 2023). These models are not only innovative in their technical design but also offer the flexibility of local deployment due to their open-access nature, facilitating research and practical applications. We conducted fine-tuning on these models using the **Llama-Factory** (Zheng et al., 2024) platform.

4.3. Evaluation Metric

Evaluate the models' efficacy through quantitative metrics and undertake human assessments by involving frontline educators to appraise the models' outputs for their consistency with Gagné's events and the overall quality of instruction.

- **BLEU-4 (Bilingual Evaluation Understudy):** BLEU-4 is used to evaluate the quality of machine-translated text against a set of reference translations. It focuses on the accuracy of four consecutive words (4-grams) in the translated text. A higher BLEU score indicates better translation quality.
- **ROUGE (Recall-Oriented Understudy for Gisting Evaluation):** ROUGE is a set of metrics for evaluating automatic text summarization and machine translation by comparing the overlap between the generated text and reference texts. This set encompasses several variants: ROUGE-1, which quantifies the overlap of unigrams, thus assessing the extent of basic content similarity; ROUGE-2, which measures the overlap of bigrams, offering insights into the prevalence of contiguous word pairs; and ROUGE-L, which evaluates the longest common subsequence, serving as an indicator of the text's coherence over extended sequences.

4.4. Human Evaluation

To validate the practical application of our method, we are designing a human evaluation study where frontline teachers will assess the quality and applicability of the LLM-generated content in real-world classroom settings. We intend to create a questionnaire focusing on the following key dimensions:

- **Alignment with Instructional Principles:** We will examine the extent to which the generated dialogue adheres to Gagné's Nine Events of Instruction, evaluating its consistency with established educational methodologies.

- **Content Quality:** The pedagogical soundness, accuracy, and alignment with the math curriculum standards will be scrutinized to ensure the content's educational validity.
- **Engagement and Clarity:** This criterion focuses on the dialogue's ability to capture students' interest and clearly communicate mathematical concepts, pivotal for fostering an effective learning environment.

By meticulously examining these key dimensions, we aim to bridge the gap between advanced language technology and real-world educational needs, ensuring that the innovative potential of LLMs is fully realized and effectively harnessed to enhance teaching and learning outcomes.

5. Discussion

This study aims to verify whether large language models can effectively analyze and generate teacher classroom dialogue that aligns with Gagné's Nine Events through precise prompt design or dataset fine-tuning. The model can be instrumental in curriculum development and assessment design, ensuring that educational objectives are met effectively across different cognitive domains. And this blend not only provides a more nuanced understanding of pedagogical effectiveness but also offers a scalable method to enrich teaching materials.

The findings from our study illuminate the substantial potential of LLMs like ChatGPT in enriching educational content creation. Moving forward, it is imperative to continue refining these models, ensuring they are both pedagogically sound and practically applicable. Collaborations between AI researchers and educational practitioners will be key in achieving a symbiotic relationship between technology and teaching, ultimately leading to enhanced educational outcomes.

6. Conclusion

In summary, our research endeavors to bridge the gap between the capabilities of current LLMs and the evolving needs of educational practitioners. By providing open-access resources, introducing a novel application pipeline, and delivering fine-tuned models ready for integration into educational systems, we contribute substantially to the enhancement of instructional technologies and methodologies. Our exploration into the use of LLMs for generating and analyzing teacher dialogue within the framework of Gagné's Nine Events of Instruction has underscored the viability and effectiveness of these models in educational settings. The successful application of both prompt tuning and model fine-tuning strategies has set the groundwork for future research, particularly in the systematic integration of AI in curriculum design and instructional evaluation.

References

Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., … Zhu, T. (2023, September 28). *Qwen Technical Report*. arXiv. https://doi.org/10.48550/arXiv.2309.16609

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901. Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Dan, Y., Lei, Z., Gu, Y., Li, Y., Yin, J., Lin, J., … Qiu, X. (2023, August 4). *EduChat: A Large-Scale Language Model-based Chatbot System for Intelligent Education*. arXiv. Retrieved from http://arxiv.org/abs/2308.02773

Du, Z., Qian, Y., Liu, X., Ding, M., Qiu, J., Yang, Z., & Tang, J. (2022). GLM: General Language Model Pretraining with Autoregressive Blank Infilling. In S. Muresan, P. Nakov, & A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 320–335). Dublin, Ireland: Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.26

Forero, M. G., & Herrera-Suárez, H. J. (2023, December 4). *ChatGPT in the Classroom: Boon or Bane for Physics Students' Academic Performance?* arXiv. Retrieved from http://arxiv.org/abs/2312.02422

Golchin, S., Garuda, N., Impey, C., & Wenger, M. (2024, February 29). *Large Language Models As MOOCs Graders*. arXiv. Retrieved from http://arxiv.org/abs/2402.03776

Kruse, K. (n.d.). *Gagne's Nine Events of Instruction: An Introduction*.

Long, Y., Luo, H., & Zhang, Y. (2024, February 22). *Evaluating Large Language Models in Analysing Classroom Dialogue*. arXiv. https://doi.org/10.48550/arXiv.2402.02380

Macina, J., Daheim, N., Chowdhury, S. P., Sinha, T., Kapur, M., Gurevych, I., & Sachan, M. (2023, October 23). *MathDial: A Dialogue Tutoring Dataset with Rich Pedagogical Properties Grounded in Math Reasoning Problems*. arXiv. Retrieved from http://arxiv.org/abs/2305.14536

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., … Zoph, B. (2024, March 4). *GPT-4 Technical Report*. arXiv. Retrieved from http://arxiv.org/abs/2303.08774

Parker, M. J., Anderson, C., Stone, C., & Oh, Y. (2023, September 29). *A Large Language Model Approach to Educational Survey Feedback Analysis*. arXiv. Retrieved from http://arxiv.org/abs/2309.17447

Qian, K., Shea, R., Li, Y., Fryer, L. K., & Yu, Z. (2023, April 11). *User Adaptive Language Learning Chatbots with a Curriculum*. arXiv. Retrieved from http://arxiv.org/abs/2304.05489

Sun, G. H., & Hoelscher, S. H. (2023). The ChatGPT Storm and What Faculty Can Do. *Nurse Educator*, *48*(3), 119. https://doi.org/10.1097/NNE.0000000000001390

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., … Scialom, T. (2023, July 19). *Llama 2: Open Foundation and Fine-Tuned Chat Models*. arXiv. https://doi.org/10.48550/arXiv.2307.09288

Uchiyama, S., Umemura, K., & Morita, Y. (2023, July 21). *Large Language Model-based System to Provide Immediate Feedback to Students in Flipped Classroom Preparation Learning*. arXiv. Retrieved from http://arxiv.org/abs/2307.11388

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., … Zhou, D. (2023, January 10). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. arXiv. https://doi.org/10.48550/arXiv.2201.11903

Yang, A., Xiao, B., Wang, B., Zhang, B., Bian, C., Yin, C., … Wu, Z. (2023, September 20). *Baichuan 2: Open Large-scale Language Models*. arXiv. https://doi.org/10.48550/arXiv.2309.10305

Zheng, Y., Zhang, R., Zhang, J., Ye, Y., Luo, Z., & Ma, Y. (2024, March 21). *LlamaFactory: Unified Efficient Fine-Tuning of 100+ Language Models*. arXiv. Retrieved from http://arxiv.org/abs/2403.13372

Exploring learner’s knowledge inquiry level and experience on utilizing ChatGPT during programming learning

Dan Sun[1, *]
[1] Hangzhou Normal University
* dansun@hznu.edu.cn

Abstract: *Generative AI can provide efficient and personalized intelligent services and technical support for education. As a typical generative AI language model, ChatGPT's application in programming has been widely concerned by the industry. However, few scholars have explored how learners use ChatGPT for programming learning from the perspective of empirical research. This study analyzes the learners' knowledge inquiry quality and learning experience through fine-grained analysis. The results show that: (1) learners mainly pay attention to the exploration of superficial-level and middle-level knowledge when using ChatGPT; (2) student interviews revealed four highly concerned themes from students’ user experience about ChatGPT: the services offered by ChatGPT, the stages of ChatGPT usage, and the pros and cons of ChatGPT. In summary, this study puts forward relevant suggestions on how to use ChatGPT to assist college students in programming learning, in order to provide reference for improving the efficiency of programming learning.*

Keywords: ChatGPT, programming learning, knowledge inquiry, experience, college student

1. Introduction

In recent years, with the rapid development of technology, generative artificial intelligence technology has been widely used in various fields. In the field of education, the launch of ChatGPT (Chat Generative Pre-Trained Transformer), a generative AI tool, has aroused heated debate (Rahman & Watanobe, 2023; Yilmaz & Karaoglan Yilmaz, 2023). As an intelligent dialog system based on a large language model, ChatGPT can assist learning and teaching through the input, interaction and generation of natural language text, providing students and teachers with more convenient and efficient learning tools (Sun et al., 2024). Among them, ChatGPT’s efficient code generation capability makes it possible to assist programmers in accomplishing various tasks in the programming process, such as code refinement, code generation from text, etc., and ChatGPT can be integrated into programming environments in order to improve the efficiency of developers’ work (Yilmaz & Karaoglan Yilmaz, 2023). Despite the growing interest in programming education, learning to program is still a challenging and complex process for college-level learners. During the process of program writing, learners have difficulty debugging errors in the programs they write and often need some support and guidance. Generative AI tools (e.g., ChatGPT) can provide solutions to the problems learners encounter in programming by quickly giving essentially correct answers and providing them with support and guidance. In addition, in-depth analysis of learners’ programming process can help teachers and learners to adopt more effective teaching strategies and learning styles (Sun et al., 2021a; Sun et al., 2023). Therefore, this study takes ChatGPT as an example and analyzes learners' knowledge inquiry level and user experiences in the programming process using ChatGPT through the lens of behavior analysis. A variety of learning analytic methods are used to explore the quality of learners' knowledge inquiry level and experience on utilizing ChatGPT during programming learning, and to study the effectiveness of ChatGPT in university programming education, with a view to providing new ideas for improving the teaching effectiveness of programming courses.

2. Literature review

2.1 ChatGPT in programming education

Unlike traditional programming tools, as an artificial intelligence language model, ChatGPT can use natural language to communicate with users and can be used even by learners with no programming experience. OpenAI, Inc. states that ChatGPT is programmed to understand user needs and respond to them (OpenAI, 2023). Traditional programming tools, which usually consist of software development environments, programming languages, libraries, and other related components, require learners to have specific programming knowledge. For example, Python IDEs are designed only for the Python language and require learners to master the baSFI concepts and operations of the language. However, while cross-platform IDEs such as Flutter or Microsoft Visual Studio support multiple languages, they still require learners to master specific syntax and concepts in order to operate.

There are still different views on the effectiveness of ChatGPT in programming education. On the one hand, researchers have highlighted various advantages of ChatGPT over other programming tools, such as providing easy access, offering quick responses, facilitating personalized learning, supporting multiple languages, providing clear explanations and programming examples, allowing for queries and searches, and providing resources on advanced topics that enhance students' computational thinking, programming self-efficacy, and motivation to learn. The researchers found that with the support of tools such as ChatGPT (mainly by providing code explanations and debugging), programming training can be effective in improving students' programming skills. On the other hand, there are limitations to the use of ChatGPT in programming education. The study highlights relevant limitations such as unstructured learning, dependence on assistive tools and environments due to lack of applications, limited support for data structures and algorithms (Yilmaz & Karaoglan Yilmaz, 2023), as well as issues such as lack of common sense, potential bias, difficulty in complex reasoning, and inability to process visual information (Rahman & Watanobe, 2023). The researchers also emphasized that the ethical issues associated with ChatGPT (e.g., bias, discrimination, privacy, security, misuse of technology, transparency, and societal impact) are intricate and need to be taken seriously.

2.2 Process-oriented analysis of programming learning

Although programming achievement is often a concern in programming education, the use of multiple learning analysis methods to evaluate the programming process is gaining increasing attention. The evaluation of the learning process can highlight the essence of promoting the programming quality of learners through practice (Pereira et al., 2020) . The research emphasizes that procedural and terminal evaluation complement each other, which can provide a comprehensive insight into learners' programming process and performance, and better understand programming phenomena and potential factors affecting the programming process (Sun et al., 2021b). Therefore, in the evaluation of programming education, in addition to focusing on the final results, but also pay attention to the performance of students in the programming process, and use a variety of analytical methods to evaluate students' programming ability.

Previous empirical studies have employed a variety of analytical methods to demonstrate various aspects of the programming learning process. Turkle and Papert (1992) pointed out that in the process of programming learning, patching learners usually build a solution gradually by repeatedly debugging and modifying the code in small increments, while planning learners tend to code after planning a more complete solution, and edit a large section of code each time. Sun et al. (2021a) used clickstream analysis, delayed sequence analysis and quantitative content analysis to analyze the pair programming behavior, discourse and cognitive level of the three groups. Studies have shown that high, middle and superficial-level groups have different characteristics in terms of social interaction, cognitive participation and final programming performance. The above research shows that a variety of learning

analysis methods can be used to analyze students' programming learning process, which is conducive to showing different dimensions of programming learning effects.

Since the characteristics of current learners in the process of programming learning with ChatGPT have not been fully explored, the purpose of this study is to use learning analysis technology to collect, analyze and mine the quality of knowledge inquiry and experience of learners in the learning process. Try to put forward the guidance of using ChatGPT to assist programming learning. The research questions are as follows:

1. What is the quality of learners' inquiry questions when using ChatGPT to facilitate programming learning?
2. What is the student's experience on utilizing ChatGPT during programming?

3. Research methodology

3.1 Participants

This study takes 36 students majoring in educational technology in the course of Object-oriented Programming in a normal university in the spring of 2023 as the research objects. The study collected conversation data and questionnaire data in the process of programming learning with ChatGPT as samples. In the first four sessions, students learn the basic concepts of Python programming and experience the application of ChatGPT. In the last class, students are asked to make programming projects using ChatGPT.

3.2 Data collection and analysis

Table 1. Coding framework of students' behavior

Category	Code	Level	Description
Knowledge inquiry (KI)	Superficial-level knowledge inquiry (SKI)	1	Students present questions related to the topic, but without personal processing and perspective
	Medium-level knowledge inquiry (MKI)	2	Students ask questions related to the topic, which are processed by themselves
	Deep-level knowledge inquiry (DKI)	3	Students ask questions related to the topic, clearly state their own ideas, arguments or opinions, and provide detailed steps to solve the problem
Feedback Inquiry (FI)	Superficial-level feedback inquiry (SFI)	1	Student feedback on ChatGPT is simply an expression of (no) agreement, or a simple assessment.
	Medium-level feedback inquiry (MFI)	2	Students evaluate ChatGPT's feedback and ask it to give simple feedback
	Deep-level feedback inquiry (DFI)	3	Students rate the feedback from ChatGPT and provide specific examples for further feedback

Firstly, the study recorded learners' knowledge inquiry level in ChatGPT through the log of ChatGPT platform. In view of such data, this study adapts the framework of open coding of knowledge inquiry developed by Ouyang and Dai (see Table 1), and analyzes the quality of learners' initial questioning (knowledge inquiry) and follow-up questioning of ChatGPT feedback from three aspects: superficial-level, medium-level and deep-level (Ouyang & Dai, 2022). First of all, all the data are sorted into Excel according to the order of learners and question content; Secondly, it identifies the questions raised by the learners and divides the units according to the order of questions, forming 516 coding units in

total. Then the two researchers manually coded according to the coding table, determined the final coding category, and conducted consistency comparison (score reliability is 0.85). Finally, all the data was converted into the format required by cognitive network analysis (Shaffer et al., 2016), and the online cognitive network analysis tool (webENA) was used for data analysis.

Secondly, at the end of the class, we conducted semi-structured interviews with learners to obtain their experience with the use of ChatGPT in programming. The interview focused on learners' feedback after using ChatGPT. Thematic analysis was used to analyze the interview data (Cohen et al., 2013). A six-step sequence was used to identify themes: (1) formatting the text data, (2) coding the data separately by two coders, (3) recording specific coded segments of data, (4) comparing segments with the same codes, (5) integrating the codes, and (6) double checking the final coded themes.

4. Results

4.1 Knowledge inquiry level

Detailed observation of Figure 1 reveals interesting patterns in how learners establish connections within their knowledge inquiry. A significant portion of their focus is dedicated to creating bridges between superficial and middle-level knowledge inquiries (SKI/MKI). This denotes that learner often engage with content that extends their foundational understanding and propels them towards more complex concepts and themes. In parallel, learners also give considerable attention to forming links between superficial knowledge inquiries and the superficial feedback content inquiries (SKI/SFI). Essentially, this implies that learners are keen to validate their basic understanding or initial interpretation of the learning material by seeking feedback that corresponds to their level of knowledge. This is indicative of learners' active participation in closing their learning feedback loops.

However, the figure demonstrates that the connection pathways involving deep-level knowledge inquiry (DKI) or deep feedback content inquiry (DFI) are rather faint. These weaker connections suggest that there may be less engagement with more complex, nuanced or higher-order elements of the subject matter. Learners may find it challenging to establish these connections due to various factors, such as the lack of pertinent background knowledge, inadequate foundational skills, or perhaps the absence of adequate guidance or prompts that can stimulate deeper inquiries.

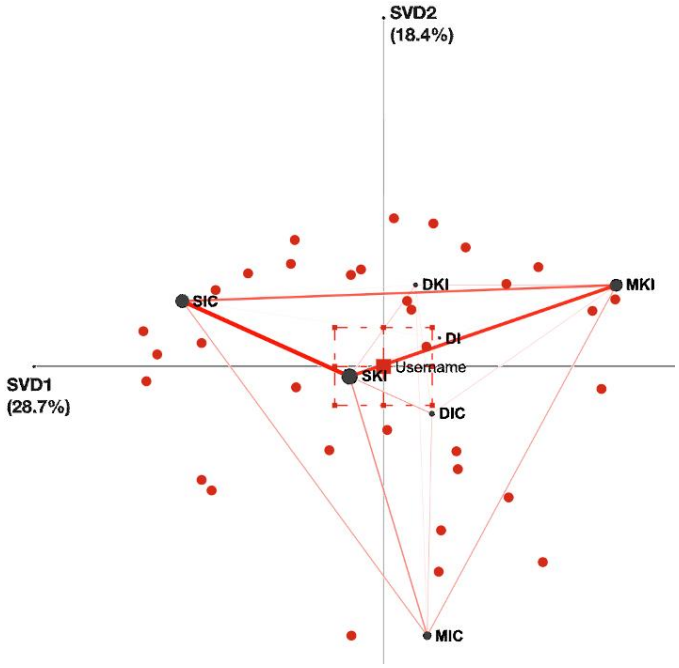


Figure 1. Cognitive network distribution map of learners' knowledge inquiry activities

4.2 User experience on utilizing ChatGPT

We gathered data from students through semi-structured interviews and utilized thematic analysis to explore different themes (Table 2). The services offered by ChatGPT, stages of ChatGPT usage, benefits of ChatGPT, and drawbacks of ChatGPT were the four main themes that emerged from the thematic analysis of learners' interview data. The first theme explored the services offered by ChatGPT that are most commonly utilized by the students. Among the 36 students in the CPF group, programming-related services such as checking errors, learning new codes, writing codes, modifying codes, and interpreting codes were mentioned 34 times. Wang remarked, "ChatGPT is a versatile tool for addressing coding inquiries and a valuable resource for learning and gaining insights into programming concepts." Additionally, domain-general services, including acquiring unfamiliar knowledge, searching for materials, answering general questions, analyzing problems, and summarizing information, were mentioned 13 times. Xia stated, "ChatGPT can identify and correct code errors, provide coding ideas, suggest relevant libraries, and offer line-by-line explanations for me to understand different codes."

The second theme highlights the utilization of ChatGPT by students across various stages of the programming process. Among these stages, coding (mentioned 34 times), debugging (mentioned 37 times), and acquiring unfamiliar knowledge (mentioned 19 times) emerged as the top three. Additionally, the students mentioned the importance of decomposing problems (mentioned 8 times) and designing projects (mentioned 4 times) in their interactions with ChatGPT. Yu stated, "Our group used this tool to generate ideas. Afterward, we chose a favorite theme, and ChatGPT helped us expand on it." Meanwhile, Chen highlighted that "ChatGPT's most advantageous feature in programming is its targeted debugging feedback. By copying and pasting error messages into ChatGPT, it helps identify problematic areas in our code."

Among the various advantages highlighted by CPF students, several key points emerged. First of all, ChatGPT mentioned "extensive programming knowledge", which includes offering coding solutions and facilitating coding tasks like writing, comprehending, and analyzing code, 29 times. Secondly, the "contextualized and varied responses" provided by ChatGPT, including contextualized feedback and diverse, targeted answers, were mentioned 20 times. According to students, ChatGPT surpasses traditional search engines like Bing or Baidu by providing not only quick programming solutions but also considering previous queries and offering multiple answers to a given problem. Thirdly, "accurate and efficient feedback" generated from integrating vast resources and saving search time was mentioned 9 times. Fang found ChatGPT more helpful than popular programming websites like CSDN or GitHub. In contrast to fruitless searches on forums, Fang could input source code or issues into ChatGPT and receive specific and tailored responses, saving significant time. Fourthly, "human-like interaction" was mentioned 2 times. Yang perceived ChatGPT as a chatbot with a high degree of humanization, stating, "ChatGPT seems to have a human-like understanding of my natural language." On the other hand, Xue pointed out an exceptional aspect of ChatGPT that sets it apart from other tools: its willingness to acknowledge incorrect answers, apologize, and accept suggestions for correction.

The fourth aspect highlighted the disadvantages of ChatGPT, as pointed out by CPF students. Firstly, issues related to inaccurate codes arising from un-updated Python libraries, insufficient code compatibility, and inadequate decomposition of large projects were mentioned 11 times. For example, Wang expressed, "Expecting ChatGPT to generate complete and flawless code for a large programming project is unrealistic." Secondly, "limited input and output" aspects were raised 4 times. This includes constraints on input and output data types (e.g., mainly text responses), insufficient understanding of "human language", inaccurate comprehension of lengthy conversational information, and demanding requirements for questioning techniques (e.g., prompts). Yang noted, "formulating questions in ChatGPT can be challenging as it requires logical and clear queries for the desired feedback." Thirdly, technical problems such as instability, slow response times, incomplete feedback (with portions of the question unanswered), and limited comprehension of the Chinese language were mentioned 3 times.

Table 2. Results of students' user experience on utilizing ChatGPT

Themes	sub-themes	*N*	Themes	sub-themes	*N*
Services offered by ChatGPT	·Programming-specific service	34	**Pros of ChatGPT**	·Expansive programming knowledge	29
	·Domain-general service	13		·Contextualized and varied response	20
Stages of ChatGPT-using	·Designing projects	4		·Accurate and efficient feedback	9
	·Decomposing problems	8		·Human-like interaction	2
	·Acquiring unfamiliar knowledges	19	**Cons of ChatGPT**	·Inaccurate codes	11
	·Coding	34		·Limited input and output	4
	·Debugging	31		·Technical problems	3

5. Discussion and implications

It can be seen from knowledge inquiry questions that learners mainly focus on superficial and middle-level knowledge inquiry when using ChatGPT, while the proportion of deep knowledge inquiry and feedback content inquiry is relatively small. Therefore, in the process of using ChatGPT to assist programming learning, learners should pay attention to the quality and depth of questions, and stimulate ChatGPT to provide more in-depth knowledge by raising challenging and open questions (Rahman & Watanobe, 2023). You can also try to ask questions related to deeper knowledge and ask ChatGPT for more in-depth explanations and guidance. In addition, learners can improve the quality and accuracy of ChatGPT responses by using Prompt words to clarify their questions, needs, or desired responses. for example, Andrew Ng pointed out in his ChatGPT Prompt Engineering for Developers course that designing prompt words (Deeplearning, 2023) by the two principles of "prompt clarity and give AI more time to think" can improve the quality of inquiry questions. In addition, learners can design or leverage open prompt word resources (such as masks or identities preset by platforms such as ChatGPT) to prompt ChatGPT to provide more in-depth and creative answers.

In addition, the thematic analysis of interview data from learners revealed four main themes: the services offered by ChatGPT, the stages of ChatGPT usage, and the pros and cons of ChatGPT. The services offered by ChatGPT include programming-specific services and domain-general services. The stages of ChatGPT usage include designing projects, decomposing problems, acquiring unfamiliar knowledge, coding, and debugging. The advantages of using ChatGPT include having extensive programming knowledge, contextualized and varied responses, and accurate and efficient feedback. Other studies have also advocated for the advantages of ChatGPT, such as providing rapid and accurate responses, availability at any time and location, ease of access, utilizing natural language, facilitating personalized learning, supporting multiple languages, offering clear explanations and programming examples, as well as providing resources for advanced topics. Besides, the study identified several disadvantages of using ChatGPT, including inaccurate codes, limited input and output, and technical problems. Other studies have also identified additional disadvantages, such as causing occupational anxiety, providing inadequate information, lacking a real programming environment, and potentially increasing laziness among programmers (Yilmaz & Karaoglan Yilmaz, 2023). Overall, the findings suggest that college students view ChatGPT as more beneficial and user-friendly and are willing to incorporate it into their future professional and educational endeavors. This study emphasizes the need to understand the potential benefits and limitations of incorporating AI language models into programming learning and how these technologies might affect individuals’ perspectives and attitudes toward technology, including learners and educators.

6. Conclusion

This study takes 36 students in a normal university as the research objects, collects the learning process data of their programming tasks by means of platform log and interview, and uses different learning analysis methods to analyze and compare the learners' knowledge inquiry quality and user experience. By analyzing the programming process of learners, the inquiry question coding framework is developed to analyze the difference in the quality of learners' inquiry questions. In view of these differences, some suggestions on how to make better use of ChatGPT for programming learning are put forward in order to improve the programming learning effect and promote the development of higher order thinking ability. There are also some limitations in this study, such as small sample size, single age and major distribution of learners, etc. These deficiencies will be explored and improved in subsequent studies. Future research will further explore and refine these limitations to better utilize generative AI as a powerful tool to support students' programming learning.

Acknowledgements

We acknowledge the students participated in the research.

References

Cohen, L., Manion, L., & Morrison, K. (2013). Research methods in education. Routledge Deeplearning.ai 2024-02-10, Retrieved from https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/

OpenAI. (2023). Introducing ChatGPT. 2024-02-10, Retrieved from https://openai.com/blog/chatgpt.

Ouyang, F., & Dai, X. (2022). Using a three-layered social-cognitive network analysis framework for understanding online collaborative discussions. *Australasian Journal of Educational Technology, 38*(1), 164-181.

Pereira, F. D., Oliveira, E. H., Oliveira, D. B., Cristea, A. I., Carvalho, L. S., Fonseca, S. C. et al. (2020). Using learning analytics in the Amazonas: Understanding students' behaviour in introductory programming. *British Journal of Educational Technology, 51*(4), 955–972. https://doi.org/10. 1111/bjet.12953

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

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on episteMFI network analysis: analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9-45.

Sun, D., Ouyang, F., Li, Y., & Chen, H. (2021a). Three Contrasting Pairs' Collaborative Programming Processes in China's Secondary Education. *Journal of Educational Computing Research, 1*(8), 54.

Sun, D., Ouyang, F., Li, Y., & Zhu, C. (2021b). Comparing learners' knowledge, behaviors, and attitudes between two instructional modes of computer programming in secondary education. *International Journal of STEM Education, 8*(1), 54–54.

Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education, 21*(1), 14.

Turkle S, Papert S. (1992). Epistemological Pluralism and the Ｒevaluation of the Concrete. *Journal of Mathematical Behavior, 11*(1), 3-33.

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

AI Agent for Education: von Neumann Multi-Agent System Framework

Yuan-Hao Jiang[1, 2, 3], Ruijia Li[4], Yizhou Zhou[5], Changyong Qi[1], Hanglei Hu[4],
Yuang Wei[1, 2, 3, *], Bo Jiang[1], Yonghe Wu[4]

[1] Lab of Artificial Intelligence for Education, East China Normal University
[2] Shanghai Institute of Artificial Intelligence for Education, East China Normal University
[3] School of Computer Science and Technology, East China Normal University
[4] Department of Educational Information Technology, East China Normal University
[5] School of Design and Engineering, National University of Singapore
[*] philrain@foxmail.com

Abstract: *The development of large language models has ushered in new paradigms for education. This paper centers on the multi-Agent system in education and proposes the von Neumann multi-Agent system framework. It breaks down each AI Agent into four modules: control unit, logic unit, storage unit, and input-output devices, defining four types of operations: task deconstruction, self-reflection, memory processing, and tool invocation. Furthermore, it introduces related technologies such as Chain of Thought, Reson+Act, and Multi-Agent Debate associated with these four types of operations. The paper also discusses the ability enhancement cycle of a multi-Agent system for education, including the outer circulation for human learners to promote knowledge construction and the inner circulation for LLM-based-Agents to enhance swarm intelligence. Through collaboration and reflection, the multi-Agent system can better facilitate human learners' learning and enhance their teaching abilities in this process.*

Keywords: large language model, AI Agent, multi-Agent system, AI for education, knowledge construction

1. Introduction

In recent years, large language models (LLMs) have achieved remarkable success in the field of artificial intelligence (AI) (Brown et al., 2020; P.-L. Chen & Chang, 2023), with their rapid increase in intelligence level being attributed to the effective utilization of comprehensive training datasets and a large number of model parameters. Benefiting from LLM's significant advancements in various downstream tasks such as machine translation (Radford et al., 2019), question-answering systems (Touvron, Lavril, et al., 2023), and dialogue systems (Touvron, Martin, et al., 2023), numerous real-world applications based on LLM have emerged rapidly. By introducing high-quality educational data and educational corpora, LLM's capabilities in educational tasks such as essay assessment, Socratic teaching, and emotional support have been rapidly enhanced (Dan et al., 2023). With the core ability of self-reflection in LLM, AI Agents can utilize LLM as the brain to empower their environment perception (Russell & Norvig, 2010) and task-solving abilities (Radford et al., 2019), effectively facilitating collaborative learning (Y. Chen et al., 2023), online learning (Živojinović et al., n.d.), and learning scenario simulations (Spoelstra & Sklar, 2007) in educational settings.

Thanks to the automated decision-making ability of AI Agents, the System composed of Multiple Agents can each perform its duties and solve complex problems through collaboration among AI Agents (X. Chen et al., 2019; Radford et al., 2019; Touvron, Lavril, et al., 2023; Touvron, Martin, et al., 2023). Similar to natural phenomena reflecting collective intelligence such as biological foraging (X.-F. Wei et al., 2024) and genetic evolution (Jiang et al., 2023), multiple intelligent agents can also produce emergent phenomena of "1+1>2" through collaboration among individuals. Through collaboration and interaction, the Multi-Agent System (MAS) can possess extraordinary intelligence and automatically complete complex tasks such as software development, and even learn to deceive other Agents in games to achieve victory (Hong et al., 2023). MAS can also replace traditional Intelligent Tutoring Systems (ITS) to better

promote learners' knowledge construction process (Deng et al., 2023). Thanks to LLM's rich responsiveness and real-time response capabilities, multiple AI Agents in MAS can concurrently play different roles such as teachers (Chi, 1996) and learning companions (Simon et al., 2023) to facilitate learners' learning process.

To comprehensively analyze the capabilities of MAS, inspired by the von Neumann Machine structure (Zhang et al., 2020), this paper proposes the von Neumann MAS Framework (vNMF). vNMF covers four operational stages: task decomposition, self-reflection, memory processing, and tool invocation, and divides MAS into elements such as logic unit, storage unit, control unit, and input-output devices. vNMF also includes a bidirectional cycle model, covering the inner circulation of Agent swarm intelligence emergence and the outer circulation of learner knowledge construction. Based on vNMF, the structural composition and operational logic of MAS can be more easily understood. With vNMF, researchers can better understand how LLM-related technologies, represented by Chain of Thought (J. Wei et al., 2023), influence the final output of MAS, and teachers or students can better utilize LLM-based-MAS to promote teaching and learning.

In this paper, Chapter 1 introduces the background of the research, Chapter 2 presents the proposed von Neumann MAS framework. Then, Chapter 3 introduces the ability enhancement cycle of MAS for education, including the outer circulation for human learners to promote knowledge construction and the inner circulation for LLM-based Agents to enhance swarm intelligence. Finally, Chapter 4 summarizes the work of this paper.

2. von Neumann MAS Framework

Upon deliberation of information flow, the architecture of LLM-based AI Agents is delineated into four core components: planning, action, tools, and memory (OpenAI, 2023). AI Agents engage in problem analysis through the planning module, amalgamating historical insights with the memory module to foster cogitation, harnessing the tools module to assimilate external resources, and ultimately operationalizing thought outcomes via the action module. For a comprehensive comprehension of AI Agent constituents, further classification into perception, brain, and action is proposed (Xi et al., 2023). Through the perception module, AI Agents procure environmental data, leveraging LLM to enhance cognitive and planning proficiencies, culminating in environmental influence via the action module. In contrast to prevailing frameworks, we introduce the vNMF model, drawing inspiration from the von Neumann Machine's architecture, to interlink AI Agent components, as depicted in Figure 1. AI Agent operations are delineated into four categories: task decomposition, self-reflection, memory processing, and tool invocation. Furthermore, AI Agents are deconstructed into four modules: logic unit, memory unit, control unit, and input-output (IO) devices. Notably, the logic unit and control unit jointly comprise the AI Agent's "Central Processing Unit," serving as the bedrock for enhancing the Agent's self-capabilities.

In the proposed vNMF model, the logic unit operates akin to the limbs of the AI Agent, endowed with the capability to activate external tools and execute specific tasks. The memory unit retains the memories of AI Agents, facilitating the recollection of past experiences to refine their actions. Serving as the brain of the LLM-based-Agent, the control unit orchestrates the coordination among multiple AI Agent modules. Functioning as the core component of the AI Agent, the control unit collaborates with the logic unit for task decomposition and with the memory unit for self-reflection.

Furthermore, the input-output devices are tasked with procuring external environmental data for the AI Agent and disseminating thinking or execution outcomes. This module encompasses components such as a Graphical User Interface (GUI), multimodal sensors, or external robotic entities. Through the synergy of these four modules in task decomposition, self-reflection, memory processing, and tool invocation, the individual AI Agents' capabilities can be enhanced, leading to improved performance of the MAS particularly in educational settings. Subsequently, leveraging the proposed vNMF model, we will expound on the perspectives of task decomposition, self-reflection, memory processing, and tool invocation.

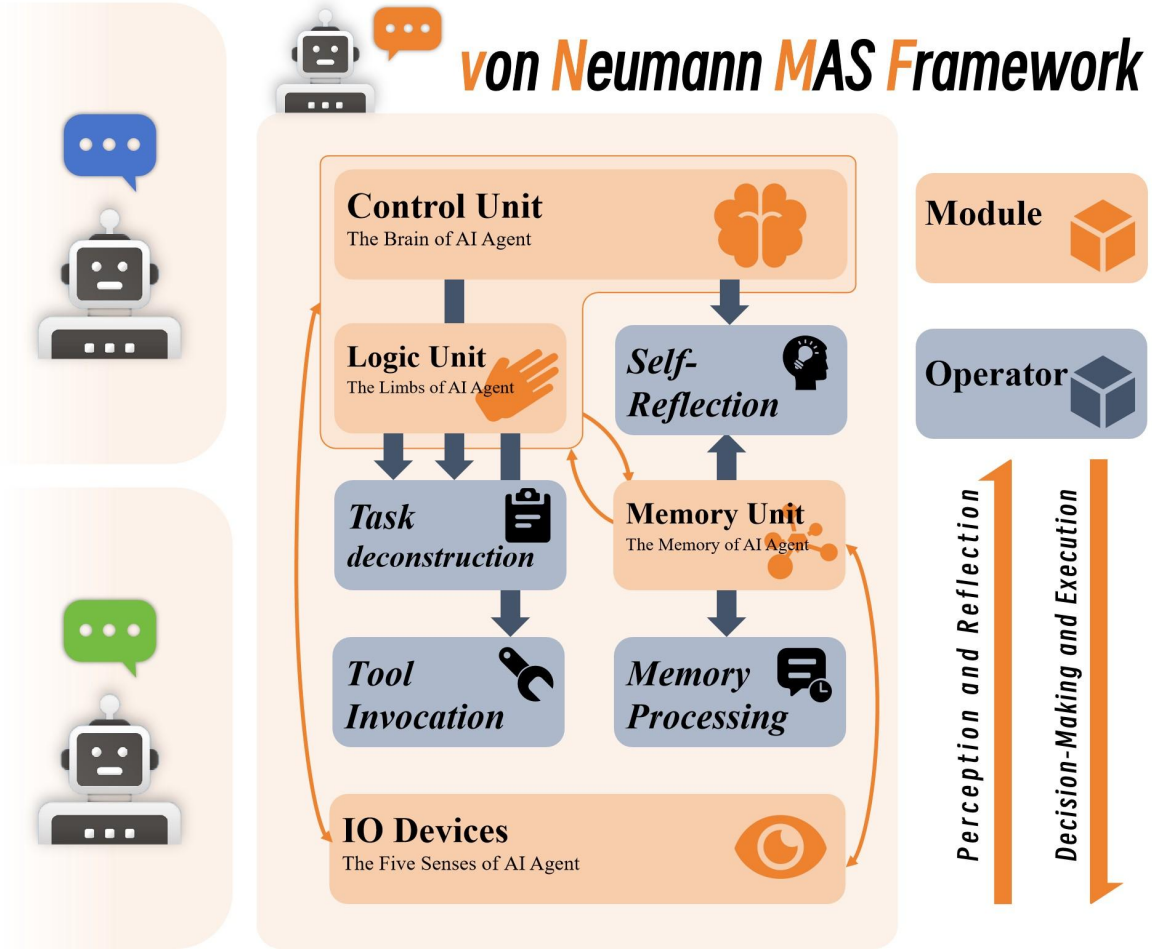


Figure 1. von Neumann multi-Agent system framework

2.1. Task Deconstruction

Through task decomposition, AI Agents can partition large tasks into smaller, more manageable sub-goals, thereby enhancing their efficiency in handling complex tasks. For instance, when tackling intricate mathematical problems, AI Agents employ task decomposition to segment them into multiple steps, resulting in heightened accuracy and clearer presentation of the thought process to learners. A prominent method of task decomposition, Chain of Thought (CoT) (J. Wei et al., 2023), has evolved into a widely embraced Prompting technology in the industry, significantly enhancing model performance in complex tasks. By instilling a "think step by step" approach in the model, CoT dissects intricate tasks into simpler subtasks, elucidating the model's thought process with greater clarity and precision.

To further refine CoT's performance, Tree of Thoughts (ToT) (Yao, Yu, et al., 2023) is proposed to circumvent potential erroneous paths during problem decomposition. ToT initially dissects the problem into logically rigorous steps, with each step corresponding to a distinct thought, thereby forming a tree-like structure of thoughts where each node represents a possible direction of thought. During the exploration process, ToT can adopt breadth-first search (BFS) or depth-first search (DFS) strategies. The BFS strategy systematically explores the first-level child nodes of the current node before proceeding to the subsequent level, while the DFS strategy delves deeply into each branch, exhaustively exploring all possible paths. This extension of CoT by ToT enables comprehensive and in-depth problem-solving, yielding more accurate and comprehensive solutions. Moreover, by incorporating aggregation, backtracking, loops, and other intricate operations, Graph of Thoughts (GoT) (Besta et al., 2023) is introduced, concurrently supporting multi-chain, tree-like, and arbitrary graph-shaped Prompt schemes, thereby equipping AI Agents with the capability to tackle complex problems.

In addition to adopting CoT, ToT, GoT, and other LLM-based task decomposition methods, AI Agents can harness external classical Planners for broader sequence overall planning, a technique termed LLM+Planner (LLM+P) (Liu et al., 2023). LLM+P utilizes planning domain definition language (PDDL) as an intermediary interface for describing planning problems. Initially, LLM translates the problem into "problem PDDL," subsequently soliciting the classical Planner to generate a PDDL plan based on the existing "domain PDDL," and ultimately translating the PDDL plan back into natural language. Through LLM+P, AI Agents can effectively leverage existing problem planners to enhance their task decomposition proficiency in specific domains, thereby averting incorrect paths or insoluble loops.

2.2. Self-Reflection

In the instructional setting, errors or excessively intricate content might yield counterproductive outcomes, exacerbating confusion among learners. Through self-reflection, AI Agents can engage in self-critique and retrospection of past actions, assimilate lessons from errors, and refine future actions to elevate the quality of the ultimate output. To augment the self-reflective capacity of AI Agents, Reson+Act (ReAct) (Yao, Zhao, et al., 2023) integrates reasoning and action within LLM, facilitating more effective interaction of AI Agents with their environment. ReAct delineates clear cognitive steps, guiding AI Agents to structure their thinking across three phases: thought, action, and observation. Grounded in ReAct, AI Agents can achieve heightened performance in tasks demanding a substantial degree of knowledge or decision-making prowess (Yao, Zhao, et al., 2023).

In contrast to the ReAct, Reflexion (Shinn et al., 2023) presents a framework rooted in the principles of reinforcement learning, to enhance AI Agents' reasoning capabilities through the endowment of dynamic memory and self-reflection capacities. This framework adheres to standard reinforcement learning (RL) protocols, where the reward model furnishes simplistic binary rewards, and the action repertoire is modeled after ReAct, introducing linguistic elements into the action domain of specific tasks to execute intricate reasoning steps. Upon detecting suboptimal planning by AI Agents, Reflexion resets the state of the AI Agent and amalgamates unsuccessful trajectories with ideal reflections to guide subsequent planning iterations, leveraging historical experiential lessons to foster reflection and glean insights from failure.

To emulate the multi-threaded reasoning processes inherent in multi-agent collaborations, multi-Agent debate (MAD) (Du et al., 2023) is proposed. MAD endeavors to mitigate illusions or misconceptions through iterative debates among multiple Agents. In each debating round, each AI Agent must contemplate not only its response but also responses from other AI Agents, crafting new responses based on this collective input. Within this framework, each agent assumes the dual roles of validator and improver, tasked with generating novel responses predicated on the received inputs. This iterative process can be iterated across multiple rounds to incentivize AI Agents to reflect on their conduct and enhance their ultimate performance.

2.3. Memory Processing

Through the synthesis and retrospection of memory, AI Agents can manage information of longer duration and greater complexity. Within the educational realm, interrogating memory allows for the analysis of a learner's current grasp of knowledge, thereby facilitating the provision of higher-quality feedback to enhance teaching efficacy. Memory can be classified into short-term memory and long-term memory based on its temporal generation. Short-term memory retains information that is presently within our awareness, whereas long-term memory has the capacity for prolonged information storage. Short-term memory functions in the context of learning, constrained by the context window length of the transformer structure, while long-term memory serves as an external vector storage that agents can attend to during retrieval. Long-term memory is further subdivided into declarative memory, housing facts and events, and procedural memory, which encompasses unconscious skills that can be executed automatically. MetaGPT presents an effective memory processing framework aimed at facilitating the rapid response of MAS to high-concurrency requests

(Hong et al., 2023). By adeptly processing and organizing the information embedded within memory, multiple AI Agents within MAS can better preserve their distinct individual traits and demonstrate enhanced anthropomorphism within classroom simulation environments (Jinxin et al., 2023).

2.4. Tool Invocation

Utilizing tools represents a significant trait of high-level human intelligence. AI Agents acquire the ability to access external application APIs, with the external tools at their disposal significantly augmenting the model's functionality. Proficiency in invoking external tools enables AI Agents to enhance the teaching process. For instance, they can initiate the analysis of input question images using computer vision models, employ optical character recognition techniques to derive text representations from educational resources and employ external computing tools for solving calculus or other intricate operations. HuggingFace stands out as an open-source platform boasting a plethora of models accessible for AI Agents. Leveraging the HuggingFace platform, HuggingGPT (Shen et al., 2023) serves as a framework employing ChatGPT as a task planner to select models based on descriptions and succinctly summarize responses derived from execution outcomes. Furthermore, through Tool Augmented Language Models (TALM), AI Agents can adeptly identify suitable tools (Parisi et al., 2022). By enabling AI Agents to acquire proficiency in utilizing existing tools, their operational scope is broadened, allowing them to execute more intricate operations.

3. The Ability Enhancement Cycle of MAS for Education

In the preceding sections, we have delved into the structure of MAS. To delve into how MAS fosters the learning journey of students in educational settings and continually evolves within this context, we introduce an ability enhancement cycle model tailored for MAS in education. This cycle encompasses both inner circulation for LLM-based-Agents and outer circulation for human learners, as depicted in Figure 2.

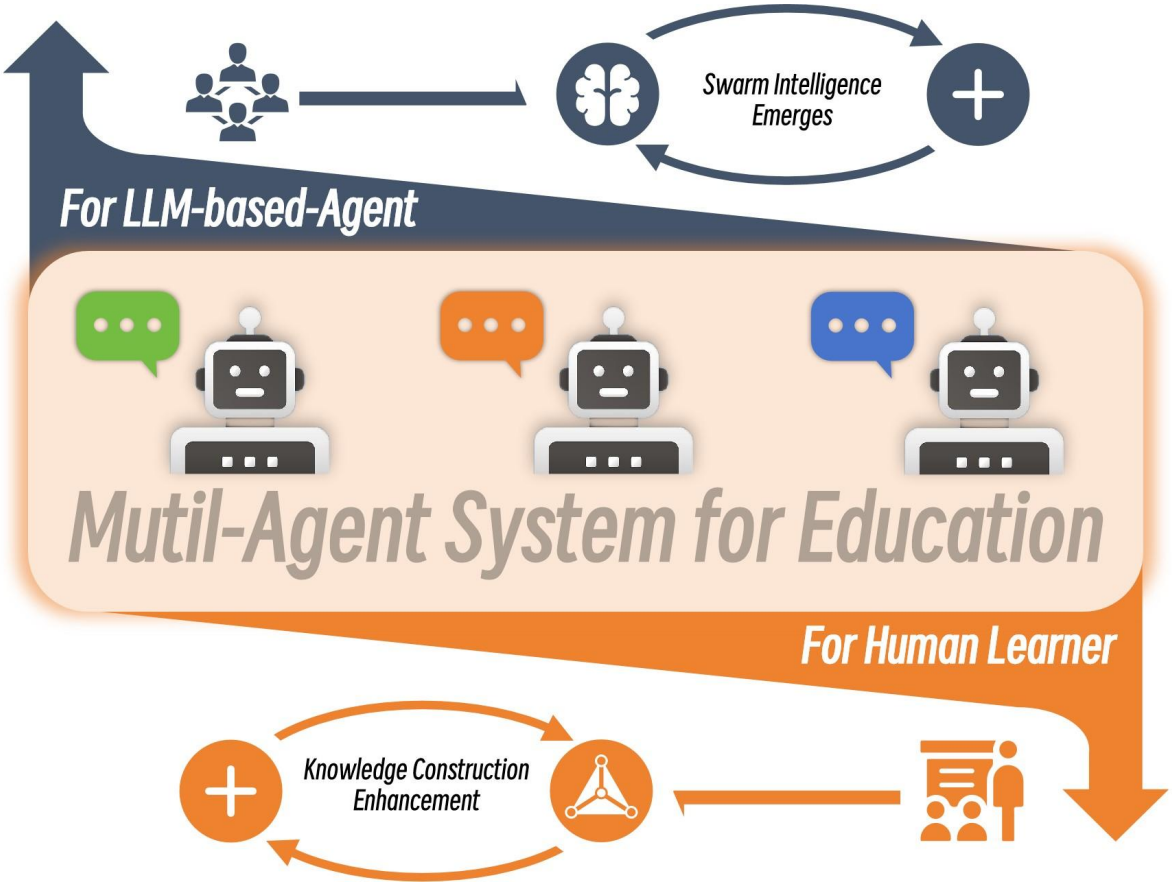


Figure 2. The ability enhancement cycle of MAS for education

3.1. Outer Circulation for Human Learners: Knowledge Construction Enhancement

Within the outer loop of MAS, human learners enrich their learning journey through human-computer collaboration, aiming for effective knowledge acquisition. MAS operates as a unified entity, refining output outcomes and bolstering pedagogical capacities through collaborative efforts among multiple intelligent agents (Du et al., 2023). Moreover, it facilitates the simultaneous adoption of various roles by multiple AI Agents, including of educators, learning companions, and more, to streamline the learning process (Jinxin et al., 2023). As learning companions, AI Agents contribute to enhancing the learning experience. Leveraging artificial intelligence technology, they furnish learners with tailored learning experiences and educational resources by delivering timely support, explanations, inquiries, and content. They also undertake metacognitive scaffolding responsibilities, intervening promptly in learners' educational journeys (Neira-Maldonado et al., 2024). Through the integration of teaching resources grounded in extensive data analysis and real-time access to network resources facilitated by AI Agents, the system amalgamates diverse educational materials and courseware. This provision of abundant teaching resources and case studies aims to guide learners through online learning effectively, akin to instructors (J. Wei, 2021). This approach fosters the knowledge construction process among learners and enhances learning outcomes.

3.2. Inner Circulation for LLM-based-Agent: Swarm Intelligence Emerges

As the outer loop of MAS operates, it enriches the knowledge construction process among human learners. Concurrently, within the inner loop of MAS, it fosters the emergence of swarm intelligence among LLM-based Agents. In the natural realm, the macroscopic intelligent behavior displayed by social organisms through collaboration is termed swarm intelligence (Tang et al., 2023; Zhou et al., 2023). Through effective collaboration among multiple AI Agents, MAS's capabilities are enhanced, facilitating the adept handling of intricate tasks such as software development (Hong et al., 2023). Educators can simulate and forecast collaborative outcomes among learners at varying proficiency levels through collaboration and simulation among multiple Agents, thus providing a foundation for differentiated instruction (Spoelstra & Sklar, 2007). Additionally, multiple AI Agents can refine output structures through debate. Despite the potential for erroneous information output by LLM, the stochastic nature of these errors is likely to be mitigated in this cross-validation process (Du et al., 2023). Through the effective promotion and utilization of swarm intelligence in MAS, the teaching capabilities of MAS are significantly augmented. Furthermore, MAS can leverage interaction data for model fine-tuning, continually enhancing output results throughout the teaching process (Casper et al., 2023).

4. Conclusion

This paper introduces the von Neumann MAS framework and delineates the ability enhancement cycle of MAS for education based on this framework. We posit that MAS can realize both the outer circulation for human learners, fostering knowledge construction, and the inner circulation for augmenting Swarm Intelligence among LLM-based Agents. By bolstering teaching capabilities, MAS can better assume the roles and responsibilities of learning companions, educators, or educational tools for learners, aiding them in effectively constructing knowledge.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 61977058, and the Natural Science Foundation of Shanghai under Grant 23ZR1418500.

References

Besta, M., Blach, N., Kubicek, A., Gerstenberger, R., Gianinazzi, L., Gajda, J., Lehmann, T., Podstawski, M., Niewiadomski, H., Nyczyk, P., & Hoefler, T. (2023). Graph of Thoughts: Solving Elaborate Problems with Large Language Models (arXiv:2308.09687). arXiv. https://doi.org/10.48550/arXiv.2308.09687

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Casper, S., Davies, X., Shi, C., Gilbert, T. K., Scheurer, J., Rando, J., Freedman, R., Korbak, T., Lindner, D., Freire, P., Wang, T., Marks, S., Segerie, C.-R., Carroll, M., Peng, A., Christoffersen, P., Damani, M., Slocum, S., Anwar, U., … Hadfield-Menell, D. (2023). Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback (arXiv:2307.15217). arXiv. https://doi.org/10.48550/arXiv.2307.15217

Chen, P.-L., & Chang, C.-S. (2023). Interact: Exploring the potentials of chatgpt as a cooperative agent. arXiv Preprint arXiv:2308.01552.

Chen, X., Li, S., Li, H., Jiang, S., Qi, Y., & Song, L. (2019). Generative adversarial user model for reinforcement learning based recommendation system. 1052–1061.

Chen, Y., Arkin, J., Zhang, Y., Roy, N., & Fan, C. (2023). Scalable Multi-Robot Collaboration with Large Language Models: Centralized or Decentralized Systems? (arXiv:2309.15943). arXiv. https://doi.org/10.48550/arXiv.2309.15943

Chi, M. T. H. (1996). Constructing Self-Explanations and Scaffolded Explanations in Tutoring. Applied Cognitive Psychology, 10(7), 33–49. https://doi.org/10.1002/(SICI)1099-0720(199611)10:7<33::AID-ACP436>3.0.CO;2-E

Dan, Y., Lei, Z., Gu, Y., Li, Y., Yin, J., Lin, J., Ye, L., Tie, Z., Zhou, Y., Wang, Y., Zhou, A., Zhou, Z., Chen, Q., Zhou, J., He, L., & Qiu, X. (2023). EduChat: A Large-Scale Language Model-based Chatbot System for Intelligent Education (arXiv:2308.02773). arXiv. https://doi.org/10.48550/arXiv.2308.02773

Deng, Y., Ren, Z., Zhang, A., Lei, W., & Chua, T.-S. (2023). Towards Goal-oriented Intelligent Tutoring Systems in Online Education (arXiv:2312.10053). arXiv. http://arxiv.org/abs/2312.10053

Du, Y., Li, S., Torralba, A., Tenenbaum, J. B., & Mordatch, I. (2023). Improving Factuality and Reasoning in Language Models through Multiagent Debate (arXiv:2305.14325). arXiv. http://arxiv.org/abs/2305.14325

Hong, S., Zhuge, M., Chen, J., Zheng, X., Cheng, Y., Zhang, C., Wang, J., Wang, Z., Yau, S. K. S., Lin, Z., Zhou, L., Ran, C., Xiao, L., Wu, C., & Schmidhuber, J. (2023). MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework (arXiv:2308.00352). arXiv. http://arxiv.org/abs/2308.00352

Jiang, Y.-H., Gao, S., Yin, Y.-H., Xu, Z.-F., & Wang, S.-Y. (2023). A control system of rail-guided vehicle assisted by transdifferentiation strategy of lower organisms. Engineering Applications of Artificial Intelligence, 123, 106353.

Jinxin, S., Jiabao, Z., Yilei, W., Xingjiao, W., Jiawen, L., & Liang, H. (2023). Cgmi: Configurable general multi-agent interaction framework. arXiv Preprint arXiv:2308.12503.

Liu, B., Jiang, Y., Zhang, X., Liu, Q., Zhang, S., Biswas, J., & Stone, P. (2023). LLM+P: Empowering Large Language Models with Optimal Planning Proficiency (arXiv:2304.11477). arXiv. https://doi.org/10.48550/arXiv.2304.11477

Neira-Maldonado, P., Quisi-Peralta, D., Salgado-Guerrero, J., Murillo-Valarezo, J., Cárdenas-Arichábala, T., Galan-Mena, J., & Pulla-Sanchez, D. (2024). Intelligent Educational Agent for Education Support Using Long Language Models Through Langchain. In Á. Rocha, C. Ferrás, J. Hochstetter Diez, & M. Diéguez Rebolledo (Eds.), Information Technology and Systems (Vol. 932, pp. 258–268). Springer Nature Switzerland. https://link.springer.com/10.1007/978-3-031-54235-0_24

OpenAI, R. (2023). Gpt-4 technical report. Arxiv 2303.08774. View in Article, 2, 3.

Parisi, A., Zhao, Y., & Fiedel, N. (2022). TALM: Tool Augmented Language Models (arXiv:2205.12255). arXiv. https://doi.org/10.48550/arXiv.2205.12255

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI Blog, 1(8), 9.

Russell, S. J., & Norvig, P. (2010). Artificial intelligence a modern approach. London.

Shen, Y., Song, K., Tan, X., Li, D., Lu, W., & Zhuang, Y. (2023). HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face (arXiv:2303.17580). arXiv. https://doi.org/10.48550/arXiv.2303.17580

Shinn, N., Cassano, F., Berman, E., Gopinath, A., Narasimhan, K., & Yao, S. (2023). Reflexion: Language Agents with Verbal Reinforcement Learning (arXiv:2303.11366). arXiv. https://doi.org/10.48550/arXiv.2303.11366

Simon, J., Fürstner, I., & Gogolák, L. (2023). Usage of Automatic Guided Vehicle Systems and Multi-agent Technology in Higher Education. Gradus, 10(1). https://doi.org/10.47833/2023.1.ENG.004

Spoelstra, M., & Sklar, E. (2007). Agent-based simulation of group learning. 69–83.

Tang, K., Wei, X.-F., Jiang, Y.-H., Chen, Z.-W., & Yang, L. (2023). An adaptive ant colony optimization for solving large-scale traveling salesman problem. Mathematics, 11, 4439.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., & Azhar, F. (2023). Llama: Open and efficient foundation language models. arXiv Preprint arXiv:2302.13971.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., & Bhosale, S. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv Preprint arXiv:2307.09288.

Wei, J. (2021). Study on the Effective Mechanism of Network English Independent Learning Platform Based on Multi Agent of Big Data. 2021 2nd International Conference on Information Science and Education (ICISE-IE), 1073–1076. https://doi.org/10.1109/ICISE-IE53922.2021.00244

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models (arXiv:2201.11903). arXiv. https://doi.org/10.48550/arXiv.2201.11903

Wei, X.-F., Tang, K., Chen, Z.-W., Chen, H.-J., Shi, Y.-H., & Jiang, Y.-H. (2024). ACDO: An Ant Colony Dynamic Optimization Framework for Tourism Route Planning. Proceedings of the 2023 4th International Conference on Computer Science and Management Technology, 851–856. https://doi.org/10.1145/3644523.3644675

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., & Zhou, E. (2023). The rise and potential of large language model based agents: A survey. arXiv Preprint arXiv:2309.07864.

Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T. L., Cao, Y., & Narasimhan, K. (2023). Tree of Thoughts: Deliberate Problem Solving with Large Language Models (arXiv:2305.10601). arXiv. https://doi.org/10.48550/arXiv.2305.10601

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing Reasoning and Acting in Language Models (arXiv:2210.03629). arXiv. https://doi.org/10.48550/arXiv.2210.03629

Zhang, H.-T., Panda, P., Lin, J., Kalcheim, Y., Wang, K., Freeland, J. W., Fong, D. D., Priya, S., Schuller, I. K., Sankaranarayanan, S. K. R. S., Roy, K., & Ramanathan, S. (2020). Organismic materials for beyond von Neumann machines. Applied Physics Reviews, 7(1), 011309. https://doi.org/10.1063/1.5113574

Zhou, K.-L., Wu, Y.-C., Jiang, Y.-H., & Zhou, F. (2023). Face Recognition Based on SVM Optimized by Improved Sparrow Search Algorithm. Software Guide, 22(5), 35–41.

Živojinović, L., Naumović, T., Barać, D., & Despotović-Zrakić, M. (n.d.). An Application of Agent Based Simulation in E-education. Retrieved 22 March 2024, from http://ipsitransactions.org/journals/papers/tar/2019jan/p4.pdf

比較自然語言模型在分析數位學習文獻的效能

The Comparison of the Effectiveness of Natural Language Models in the Analysis of Digital Learning Literature

陳一嘉[1]，陳攸華[2,*]
[1] 中央大學網路學習科技研究所
[2] 中央大學網路學習科技研究所
[*] sherry@cl.ncu.edu.tw

【摘要】 自然語言的語言模型可強化資訊檢索的搜尋功能，而 TF-IDF 與 Sentence-Transformer/all-mpnet-base-v2 可完成此項強化，此外，數位學習領域文獻正大量增長，因此本研究旨在基於數位學習領域文獻，來全面性的比較 TF-IDF 及 Sentence-Transformer/all-mpnet-base-v2 的效能差異。結果顯示兩者表現出相互彌補的優劣勢，前者具有好的精確度和錯誤率，但召回率和遺漏率方面卻表現不佳，後者則擁有好的召回率和遺漏率，但精確度和錯誤率卻有缺失。此項發現不僅可以幫助自然語言的研究者理解 TF-IDF 與 Sentence-Transformer/all-mpnet-base-v2 之優勢與劣勢，進而在未來能選擇合適的語言模型，也可以為數位學習領域提供一個有效參考，以能建立具有強化搜尋功能的數位學習文獻資料庫。

【關鍵字】 自然語言；語言模型；文獻搜尋；數位學習

Abstract: *Among various Natural language models, TF-IDF and Sentence-Transformer/all-mpnet-base-v2 can enhance capabilities of information retrieval. This study aims to comprehensively compare effectiveness differences between them in the analysis of digital learning literature, where there is a rapid growth recently. Results revealed complementary strengths and weaknesses existed between these two models. Specifically, TF-IDF showed good precision and error rates while their recall and omission rates were poor. On the other hand, Sentence-Transformer exhibited superior recall and omission rates, but failed in precision and error rates. These findings can not only aid researchers in selecting appropriate language models by understanding their advantages and disadvantages, but also offer effective references for the establishment of a versatile literature database in the domain of digital learning.*

Keywords: natural language, language models, literature searching, digital learning

1. 前言

自然語言處理（Natural Language Processing）屬於一種人工智慧領域的科技，此科技可以讓電腦能能夠像人類一樣閱讀、聽懂以及生成人類語言(王穎潔 et al., 2021)，在此過程，需將各種輸入的語言轉換成有意義的符號，讓電腦能夠判斷、記憶以及學習人類語言，進而分析它的深層涵義、重點結構以及潛在因素，以提供多種不同的相關應用，包括：資訊檢索(Dien et al., 2019)、翻譯(Khan et al., 2020)、情緒分析(Guo, 2022)、問題答覆(Lende & Raghuwanshi, 2016)等。而其中資訊檢索(Ajmal et al., 2023)則是一項常見的應用，理想的資訊檢索系統，不僅需要搜尋出詞彙完全符合條件的文件,而且還需要找出詞彙雖然不同,但是在意義、概念 上接近的文件，以避免遺漏相關的文件(陳光華 & 莊雅棻, 2001)。因此需要仰賴自然語言，才能夠將檢索詞彙以及文本轉為有意義的數值，之後電腦能夠依據這些數值，來分析文本結構與重點，從而選出與檢索詞彙最具相關性的文本作為檢索結果。

這個過程需要藉助語言模型來完成，而 Term Frequency-Inverse Document Frequency(TF-IDF)和 SentenceTransformer/all-mpnet-base-v2(ST/AMB2)則是常被用於資訊檢索的語言模型。TF-IDF 以數值方式呈現檢索詞彙在各文本中的重要程度，此數值包含兩個部分，其一是詞彙在單一文本中的出現頻率，也就是 TF 分數，其二是文本之稀有度，也就是 IDF 分數，而 TF 分數乃是依據詞彙出現次數佔文本總詞數的比例來計算，也就是詞彙出現次數除以文本總詞數，此計算方式可以避免篇幅大小所造成出現次數不平衡的問題。而 IDF 分數則是根據包含詞彙之文本數佔文本總數的比例來計算(Qaiser & Ali, 2018)，此種方式能夠過濾掉那些頻繁出現的虛詞，例如：“is”, “the”..等，從而專注更具區分性的重要性詞彙，也就是 Log e (the total number of documents / the number of documents containing the term + 1) 。上述 TF 分數與 IDF 分數兩者相乘的結果則可以判斷出某個詞彙在文本中是否具有重要性，分數越高則越具有重要性，分數越低則越不具有重要性。(劉文平 et al., 2023)

而 ST/AMB2 的詞庫中儲存各式各樣的詞彙，這些詞彙乃是由單字組成，並且每個單字會被賦予不同的數值，模型會根據上下文以及單字的位置，判斷每個單字在句中的涵意，進而了解單字與單字間是否具有關連性，若是具有關連性則組成一個詞彙，然後給予詞彙一個數值，以符合句中的涵意(Reimers & Gurevych, 2019)。此文本中各詞彙數值的平均值則構成了文本向量，而某個詞彙和文本向量的距離就是文本分數，當文本分數越高時，表示詞彙和文本向量的距離越近，也就是相似度越高，因此，這個詞彙的重要程度則可能越高(Song et al., 2020)。

前述兩種語言模型各自擁有不同的優缺點，關於優點部分，TF-IDF 的考慮層面較為完整，因為它不僅考慮到詞彙在單一文本中的出現頻率，同時也重視該詞彙在所有文本當中的稀有度，兩者合併的考慮，可以具有較強的區別能力(Zhou, 2022)，進而能判別出重要的關鍵詞彙。在另一方面，ST/AMB2 則以向量距離做為相似性的判斷，注重詞彙之間的關聯程度，而非依賴完全的匹配(Beermann et al., 2023)。換句話說，它的匹配更多的乃是依靠相似程度，只要兩者之間存在著一定的關聯性，儘管是非常微小的部分匹配也能被識別出來，所以它具有較強的包容能力。

上述的優點也產生了它們的缺點，TF-IDF 的缺點來自於採用完全匹配的方式來選取重要詞彙。換言之，在無法完全匹配的狀況下，可能會遺漏一些重要的詞彙。另外，它的稀有度雖然可以有較強的區別能力，但卻可能忽略了那些高頻且重要的詞彙。而 ST/AMB2 的缺點來自於以相似程度作為分數的判斷基準，使其不受限於完全的匹配，擁有較大的包容能力，故所產生的詞彙也許並沒有太高的重要性，也就是說它選取的方式較為鬆散，可能會降低精準的程度。

綜言之，基於 TF-IDF 和 ST/AMB2 的原理不同，它們有不同的優缺點，可是卻缺少實證研究來比較兩者的效能，特別是在檢索數位學習領域的學術文獻，而此項比較卻非常重要，因為數位學習領域的研究正在蓬勃發展(Aulakh et al., 2023)，因此本研究的目的是以數位學習領域文獻為基礎，來比較 TF-IDF 與 ST/AMB2 之效能差異，為了實現上述目的，本研究分析了三個研究問題，這些研究問題如下：

- 基於召回率、精確率、錯誤率以及遺漏率之效能評估標準，TF-IDF 對於識別數位學習領域文獻之關鍵詞表現為何？
- 基於召回率、精確率、錯誤率以及遺漏率之效能評估標準，ST/AMB2 對於識別數位學習領域文獻之關鍵詞表現為何？
- 基於召回率、精確率、錯誤率以及遺漏率之效能評估標準，TF-IDF 和 ST/AMB2 對於識別數位學習領域文獻之關鍵詞是否有所差異？

2. 研究方法

2.1. 蒐集與分析工具

本研究使用的工具包括關鍵詞清單與實驗數據集，詳述如下：

- *關鍵詞清單：*為建立有效的關鍵詞清單，蒐集 2016 至 2023 年之六本數位學習領域 Q1 期刊常用的關鍵詞，包括“British Journal of Educational Technology”,“Computer & Education”, “Educational Technology Research and Development”, “Interactive Learning Environment”, “Journal of Computer Assisted Learning”, 及“Journal of Educational Computing Research”。這六本數位學習領域期刊共計 5180 筆文獻，而所收錄的關鍵詞共計有 102 個，這些關鍵詞將被列入在本研究的關鍵詞清單。
- *實驗數據集：*我們利用 Python 的 Beautiful Soup 爬蟲技術，蒐集了 100 篇數位學習相關領域之發表文獻，作為此次研究的實驗數據集。然後以關鍵詞清單作為基礎，用 TF-IDF 與 ST/AMB2 來分析文獻的標題和摘要，以擷取最重要的相關詞彙。而這些最重要的相關詞彙會再與文獻的關鍵詞比對，以進行效能之評估。

2.2. 分析流程

為了兼顧 TF-IDF 與 ST/AMB2 不同的原理與優缺點，我們所使用之關鍵詞清單不僅包含了詳細精準的詞彙，例如 “Game-based learning” 或 “Gamification”，而且也涵蓋了較為廣泛的詞彙，例如 “Game”。上述的關鍵詞清單則被用來進行 TF-IDF 與 ST/AMB2 之分析，關於 TF-IDF，首先設定最後 25%之標題詞彙給予加權提升，以避免忽略出現次數頻繁但仍具重要性的詞彙。然後計算清單上之詞彙於所有文本之稀有度，以得到 IDF 分數，以及藉著分析詞彙在標題與摘要出現的頻率，來取得 TF 分數，之後則再透過 TF 與 IDF 分數相乘，計算出每一個詞彙所得到的 TF-IDF 分數，並給予排序。

利用 ST/AMB2 進行分析時，為避免不足的模型訓練所造成的偏差，則先將十篇數位學習領域的文獻轉至詞庫中，以使此模型熟悉該領域的專門詞彙。另外在計算文本分數和關鍵詞分數之前，也將沒有意義的虛詞從文本中刪除，以避免干擾。而文本分數則是以 60%的標題和 40%之的摘要來計算，以凸顯標題之重要性。在另一方面，先將文本切割成不同的單字，而再將這些單字與關鍵詞清單做比對，若是某些單字之組合屬於清單中的其中一個關鍵詞，則是擷取這些單字和所屬關鍵詞在詞庫中的數值，然後以數值上的差異來評估兩者之間的距離，每一個距離之間所得數值之平均則為這個關鍵詞在此文本的分數，也就是分數越高時，這個關鍵詞在此文本越為相關。接下來則是用 40%的文本分數及 60%的關鍵詞分數來計算每一個詞彙最終的分數，並依此分數加以排序。

最後依據每一個模型所生成的詞彙數量來選擇較具相關的詞彙，TF-IDF 所生成的詞彙並沒有太多，故選取前 50%者做為效能評估。相反地，ST/AMB2 所生成的詞彙較多，因此只選入分數屬於前三分之一的詞彙，以便能選擇和文本非常相關的詞彙做為效能評估。上述所選入 TF-IDF 與 ST/AMB2 兩者的關鍵詞彙，則被用來和文本發布之關鍵詞做人工方式的比對，然後進行效能評估，包括：召回率、精確率、錯誤率、遺漏率、F1-Score。最後，透過視覺化的圖表將效能評估的結果完整呈現。圖 1 描述上述兩種模型的分析流程，總共有六個階段，此六個階段皆是透過 Window 11 x64 筆記型電腦，搭配 Intel-i7-12700H 處理器，於 Anaconda 使用 Python 版本 3.7.16，同時以 NVIDIA GeForce RTX3060 6GB GPU 進行加速處理與執行。

2.3. 效能評估定義

此研究包含五種效能評估：

- 召回率(Recall)：模型與文本發布關鍵詞相同之總數 / 所有文本發布關鍵詞總數
- 精確率(Precision)：模型與文本發布關鍵詞相同之總數 / 模型生成詞彙總數

- 錯誤率(False)：模型與文本發布關鍵詞不相同之總數 / 模型生成詞彙總數
- 遺漏率(Missing): 模型遺漏文本發布關鍵詞之總數 / 所有文本之發布關鍵字總數
- F1-Score: 2 x (精確率 x 召回率) / (精確率 + 召回率)

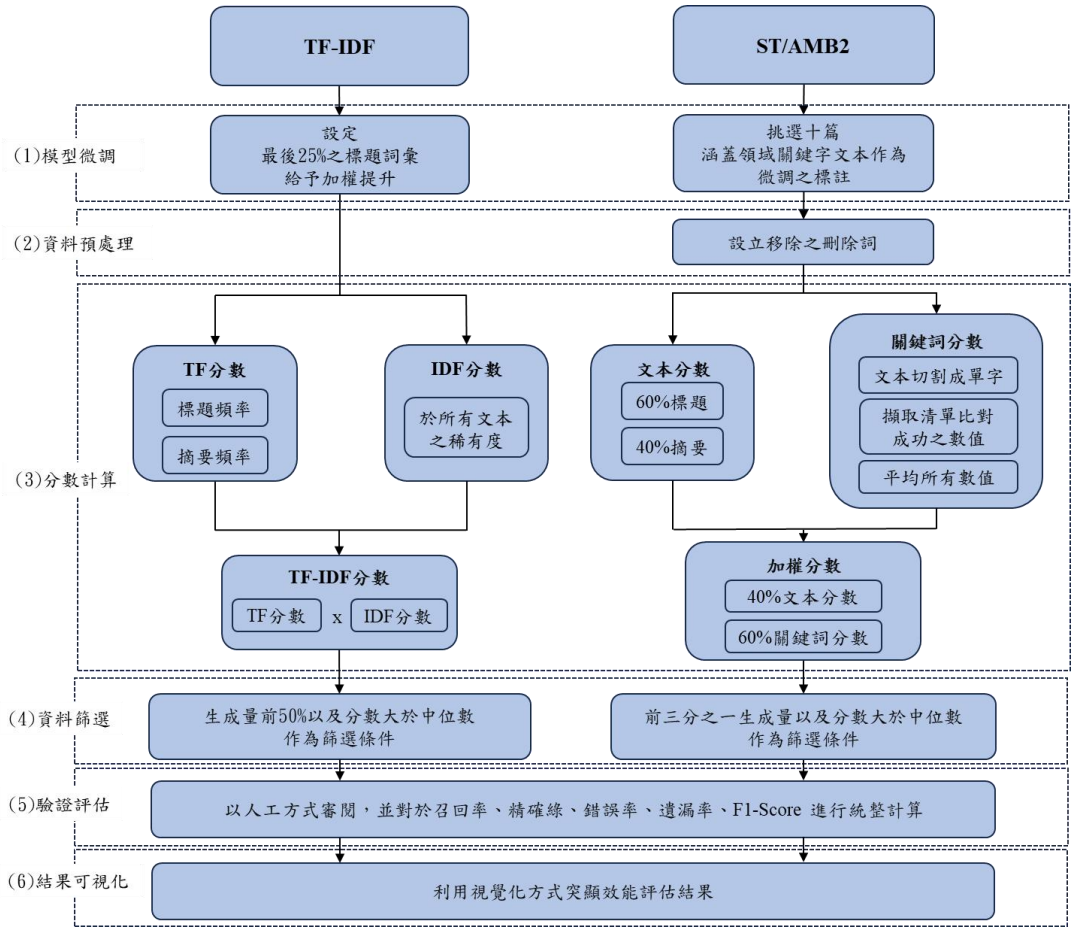


圖 1 資料分析流程圖

3. 結果與討論

3.1. 研究結果

依據圖 1 所顯示之流程,此章節將描述 TF-IDF 和 ST/AMB2 在每一個階段所獲得的結果,以下先列出有關 TF-IDF 的結果:

- 第一階段--標題中所出現的詞彙，可以顯示出該文章的重點內容，因此有給予加權處理，更明確的說，屬於最後 25%的詞彙若有出現在標題當中，則予以加權提升。
- 第二階段--TF-IDF 不需要資料預處理。
- 第三階段--所得的 TF 分數之範圍介於 0.067 至 1 之間，而 IDF 分數則介於 2.2 至 5.63 之間，而每個詞彙所得的 TF-IDF 分數則是 TF 分數與 IDF 分數相乘的結果，所以範圍則介於 1.91 至 4.92 之間。
- 第四階段--只有前 50%之詞彙才會被收錄，也就是只收錄相乘分數高於 1.4 的詞彙，共計 94 個詞彙。而與所有文本之發布的 419 個關鍵字相比對，則有 84 個比對成功。
- 第五階段—進行效能評估後的結果顯示：召回率為 20.29%、精確率為 90.43%、錯誤率為 9.57%、遺漏率為 79.71%以及 F1-Score 為 32.81%。
- 第六階段--圖 2 用視覺化的方式顯示效能評估的結果。

上述結果顯示 TF-IDF 具有高度的精確率和較低的錯誤率,此結果與 Kadhim (2019)之研究相呼應，也就是 TF-IDF 具有精準的檢索能力，而之所以造成精準檢索能力的原因是 TF-IDF

進行查詢時，需要出現完整的關鍵詞。然而，在召回率和遺漏率方面則表現不佳，因為 TF-IDF 運用完全匹配的方法，而無法識別同義詞之變化，就有可能遺漏具有相同意義的重要詞彙(趙漢清 & 王志國, 2019)。

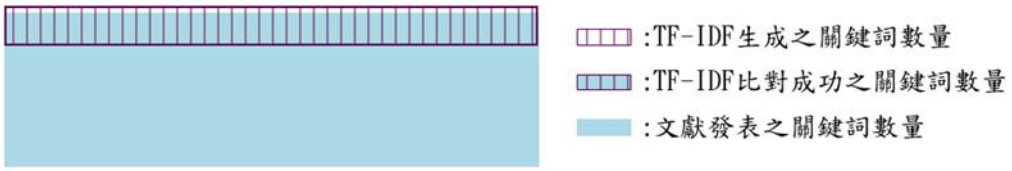


圖 2 TF-IDF 的效能評估

以下則是列出 ST/AMB2 每一個階段的結果:

- 第一階段--為了使模型能夠熟悉數位學習領域的專門詞彙，則以十篇該領域的研究文獻來微調模型。
- 第二階段--從文本中刪除介係詞或冠詞，以使模型更專注於重要之詞彙。
- 第三階段--所得的文本分數範圍介於 0.008 至 0.3 之間，而關鍵詞分數介於-0.03 至 0.66 之間，而以 40%的文本分數及 60%的關鍵詞分數來加權，最後所得的結果則是每個詞彙的加權分數，其介於 0.24 至 0.43 之間。
- 第四階段--排序每個詞彙的加權分數，然後只收錄前三分之一的詞彙，也就是僅收錄加權分數高於 0.2395 的詞彙，共計有 444 個詞彙。而與所有文本之發布關鍵字相比對，則有 184 個比對成功。
- 第五階段--進行效能評估後的結果顯示：召回率為 45.10%、精確率為 42.57%、錯誤率為 58.56%、遺漏率為 62.05%以及 F1 Score 為 43.8%。
- 第六階段--圖 3 用視覺化的方式顯示效能評估的結果。

上述結果顯示 ST/AMB2 擁有較高的召回率和較低的遺漏率，此乃是因為模型中的詞庫記錄了各類字詞的變化(Galli et al., 2024)，故擁有較高的包容性(Kosenko & Zharikova, 2023)。然而，精確率以及錯誤率方面則不佳，這種情形尤其會發生在特定領域，因為詞庫沒有學習過該領域的專有名詞,可能會造成相關性的誤判,進而減低了精確率(Kosenko & Zharikova, 2023)。

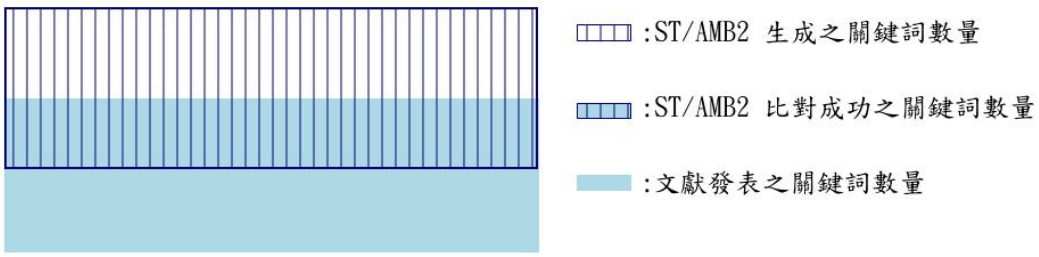


圖 3 ST/AMB2 的效能評估

4. 結論

一項理想的資訊檢索系統不僅應該幫助使用者獲得完全符合條件的檢索結果，也需要使他們能夠搜尋到文字表達不同但意義相近的檢索結果。此部分需要仰賴自然語言處理的語言模型，而 TF-IDF 與 ST/AMB2 正是常被用於資訊檢索的語言模型。在另一方面，數位學習領域在近年來產生了大量的文獻，因此，本研究目的乃是探討 TF-IDF 與 ST/AMB2 對於判別數位學習文獻關鍵詞之效能差異。而研究結果可以為前述的三個研究問題提供完整的答案。關於第一個研究問題，TF-IDF 具有高度的精確率和較低的錯誤率，但在召回率和遺漏率方面表現

不佳，關於第二個問題， ST/AMB2 擁有較高的召回率和較低的錯誤率，然而，在精確率以及錯誤率則表現有所缺失。關於第三個問題，TF-IDF 與 ST/AMB2 擁有不同的優缺點，可以彼此相互彌補，前者具有精準的判斷能力，卻缺乏好的召回率和遺漏率，而後者缺乏精準判斷能力，但卻擁有較好的召回率和遺漏率。

上述的發現對於自然語言處理以及數位學習領域皆可有具體的貢獻，在自然語言處理方面，本研究進行了完整的評估，包括：精確率、錯誤率、召回率以及遺漏率與 F1-Score。這些完整評估後的結果可以幫助該領域的研究者更明瞭 TF-IDF 與 ST/AMB2 之優勢與劣勢，以能夠在未來做為選擇適當語言模型之依據。在數位學習領域，本研究使用 TF-IDF 與 ST/AMB2 來分析該領域的文獻，這些分析後的結果可以做為未來建立數位學習文獻資料庫的參考，以使該資料庫能夠善用不同語言模型的強項，來強化搜尋功能，以滿足不同使用者的檢索需求。而儘管具有以上的貢獻，本研究仍存在著侷限，其中一個限制是本次研究的資料總數量較小，而未來則需要以更大的數據量來驗證這項資料的結果，另一個限制是我們只比較了兩個語言模型，未來應加入多個語言模型，以提升研究結果的適用性。

參考文獻

王穎潔, 朱久祺, 汪祖民, 白鳳波, & 弓箭. (2021). 自然語言處理在情感分析領域應用綜述. *電腦應用*, 0. https://doi.org/10.11772/j.issn.1001-9081.2021071262

陳光華, & 莊雅蓁. (2001). 應用於資訊檢索的中文同義詞之建構. *資訊傳播與圖書館學, 8 (1), 2001 年 9 月*, 59-75.

趙漢清, & 王志國. (2019). 基於機器學習的中醫學派文本分類研究. *中華醫學圖書情報雜誌, 27*(12), 7-11. https://doi.org/10.3969/j.issn.1671-3982.2018.12.002

劉文平, 李艷春, 張賀, 張宇馳, 丁鼎, 於泉, & 王傳煬. (2023). 基於 TF-IDF 和 jieba 分詞的交通運輸綜合執法語音檔和文字檔關聯匹配技術. *Open Journal of Transportation Technologies*, *12*, 377. https://doi.org/10.12677/OJTT.2023.125041

Ajmal, S., Ahmed, A. A. I., & Jalota, C. (2023). Natural Language Processing in Improving Information Retrieval and Knowledge Discovery in Healthcare Conversational Agents. *Journal of Artificial Intelligence and Machine Learning in Management*, *7*(1), 34-47.

Aulakh, K., Roul, R. K., & Kaushal, M. (2023). E-learning enhancement through educational data mining with Covid-19 outbreak period in backdrop: A review. *International journal of educational development*, 102814.

Beermann, J., Benfer, R., Both, M., Müller, J., & Diedrich, C. (2023). Comparison of Different Natural Language Processing Models to Achieve Semantic Interoperability of Heterogeneous Asset Administration Shells. 2023 IEEE 21st International Conference on Industrial Informatics (INDIN),

Dien, T. T., Loc, B. H., & Thai-Nghe, N. (2019). Article classification using natural language processing and machine learning. 2019 International Conference on Advanced Computing and Applications (ACOMP),

Galli, C., Donos, N., & Calciolari, E. (2024). Performance of 4 Pre-Trained Sentence Transformer Models in the Semantic Query of a Systematic Review Dataset on Peri-Implantitis. *Information*, *15*(2), 68. https://doi.org/10.3390/info15020068

Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. *Journal of Intelligent Systems*, *31*(1), 113-126. https://doi.org/10.1515/jisys-2022-0001

Kadhim, A. I. (2019). Term weighting for feature extraction on Twitter: A comparison between BM25 and TF-IDF. In 2019 international conference on advanced science and engineering (ICOASE) (pp. 124-128). IEEE

Khan, N. S., Abid, A., & Abid, K. (2020). A novel natural language processing (NLP)–based machine translation model for English to Pakistan sign language translation. *Cognitive Computation*, *12*, 748-765. https://doi.org/10.1007/s12559-020-09731-7

Kosenko, D. V., & Zharikova, D. R. (2023). KRGP: knowledge-based response generation with persona. *Записки научных семинаров ПОМИ*, *529*(0), 72-85.

Lende, S. P., & Raghuwanshi, M. (2016). Question answering system on education acts using NLP techniques. 2016 world conference on futuristic trends in research and innovation for social welfare (Startup Conclave),

Qaiser, S., & Ali, R. (2018). Text mining: use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, *181*(1), 25-29. https://doi.org/10.5120/ijca2018917395

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (2020). Mpnet: Masked and permuted pre-training for language understanding. *Advances in Neural Information Processing Systems*, *33*, 16857-16867.

Zhou, H. (2022). Research of text classification based on TF-IDF and CNN-LSTM. journal of physics: conference series,

基于大语言模型的素养知识图谱构建——以信息科技课程为例

Constructing an Ontology-Based Knowledge Graph for K-12 Computer Science Competency via Human-AI Collaboration

叶宇雷*
华东师范大学教育信息技术学系
*yyl_ray@126.com

【摘要】 素养在教育领域中是一个被广泛应用的概念。素养的定义通常以非结构化的自然语言方式呈现，这种呈现方式为素养的测量和评价带来了阻碍，而形式化表达的素养本体能够明确素养的范围及其内部结构。但本体构建需要投入时间成本并且依赖于领域专家。借助 ChatGPT 等大语言模型强大的自然语言理解能力，本研究提出一种人机协同的本体和知识图谱构建方式。与 ChatGPT 协作，从义务教育阶段的信息科技课程框架中抽取、调整和确定知识概念、关系和实体，构建素养本体及基于本体的知识图谱。我们验证了 ChatGPT 协同抽取方式的结果，评价指标 Kappa 值超过 0.6，精确率均大于 0.75。指标表明 ChatGPT 的抽取结果与人工抽取的结果相对吻合，且其抽取任务至少能够覆盖人工抽取结果的 75%。

【关键词】 知识图谱；本体；人机协同；ChatGPT；素养

Abstract : *Competency is a broad concept in education, often described in unstructured natural language within curriculum frameworks, which makes assessing it difficult. The construction of a competency ontology serves to mitigate this drawback by illustrating its boundary and internal structure. However, construction of ontology is a very time-consuming work that need huge effort from several domain experts. Inspired by the powerful language understanding capability of large language models (LLMs) such as ChatGPT, this work proposed a human-AI collaboration approach to accelerate the construction of competency ontology. This paper demonstrate how to extract knowledge concepts, relations and entities from a Chinese national computer science curriculum frameworks to construct a competency ontology and ontology-based knowledge graph through collaboration with ChatGPT. We evaluate the extraction result of ChatGPT by comparing with manual result. The evaluation reveals that Kappa scores surpass 0.6, indicates that the result aligns relatively well with manual result. ChatGPT's outcomes demonstrate that it is capable to cover at least three-quarters of human efforts in extraction.*

Keywords: knowledge graph, ontology, human-computer collaboration, ChatGPT, competency

1. 引言

素养本位的教育和学习已经成为了世界各国 K12 教育的热点。近年来，部分国家以学生素养为依据更新各国的 K12 课程标准和教学大纲。中华人民共和国教育部发布了新版《义务教育阶段信息科技课程标准》（以下简称《新课标》），指明信息科技课程包含的四个素养：信息意识、计算思维、数字化学习与创新及信息社会责任。然而，《新课标》中四个素养的定义着重于定性描述，难以指导素养相关的实践工作顺利展开。明确素养定义是开展素养本位的教育的前提条件，但目前素养概念仍缺乏统一而明确的定义(El Asame & Wakrim, 2018)。

在现有研究和讨论中，不同学者和组织对于素养的定义有着明显区别。素养定义的模糊性和宽泛性为教育者和教育领域研究者围绕素养开展工作带来了挑战。使用自然语言的表达方式存在多义性、模糊性等局限，以这种形式表示的素养使定义结构化和任务处理更复杂。自然语言易于被人类阅读，却难以被计算机处理。从这一角度来看，自然语言形式的定义阻碍了使用数字设备开展素养相关活动的过程。因此，形成清晰的、结构化的素养定义是推进素养本位教育的重要手段(Kiesler, 2020)。

现有研究主要通过构建模型的方式，提供精确的描述和高效的表达方式，解决定义模糊的问题。素养模型的作用不仅限于描述和定义，模型还能用于绩效评估、目标设置和信息管理等。本体模型以特定形式进行表示，能够清晰地表示概念的边界和内部结构。本体模型能够提供一种既可计算，又具有一定可读性的表示形式。因此，构建素养本体模型能够促进素养相关实践工作的数字化进程。除此之外，本体模型定义了清晰的边界和内部结构，能够消除自然语言定义的模糊性。传统的本体构建方法需花费大量时间，也高度依赖于领域专家。

大语言模型（Large Language Models, LLMs）的出现缩减了自然语言与计算机语言间的鸿沟，LLMs 使得用户能用自然语言进行交互。同时，用户通过某些方法调整模型使 LLMs 能够在特定任务上有更好的表现，这些方法包括预训练、微调和提示词。其中，提示词是一种简单而广为应用的方式。以 ChatGPT 为代表的 LLMs 在自然语言处理任务上展现出强大的能力，尤其在处理自然语言生成任务方面。ChatGPT 在问答和写作上表现出色，生成的文本不仅内容上难以与人类文本区分，且语言流畅，风格多样。除文本生成任务外，ChatGPT 在基础自然语言处理任务中也有可观表现，如文本分类任务、命名实体识别等(Qin 等, 2023)。用户能通过聊天框，使用自然语言指示 GPT 模型完成相应任务，而无需掌握专业知识或是花费大量时间训练一个巨大参数量的模型。

通过与 ChatGPT 协作完成实体抽取任务，有望提高本体构建和维护效率，加速构建过程，并减对于领域专家的依赖，减少资源投入。因此，本文提出以下问题：ChatGPT 能否用于素养模型构建过程？ChatGPT 具体能为构建提供哪些帮助，以及提供帮助的有效性如何？

2. 文献综述

2.1 素养定义

不同研究者和组织在他们的研究或文档中都给出了对于素养的不同定义。然而，素养定义的多种多样反映了素养定义存在着争议性，其定义在各个领域内都尚未达成共识。

在教育领域，Christiaens 使用了 HR-XML 提供的素养定义框架。在该框架中，“素养”是指“与发挥效用相关的个人的特征，这些特征指的是特定的、可识别的、可定义的和可测量的知识、技能、能力和/或其他相关个人特质（如态度、行为和身体素质），这些特征是人力资源所拥有的，或是在特定的业务环境中，开展某项活动所必要或是至关重要的”(Christiaens 等, 2006)。Sampson 和 Fytros 致力于对素养进行统一的定义，分析不同的素养定义，并最终构建除了一个素养的模型。该研究确定了素养的三个维度，即个人特性、熟练程度和情境。个人特性包括知识、技能、态度、行为和价值观等。熟练程度表示个人通过完成活动表现出来的素养等级。情境则指素养发挥作用的场景(Sampson & Fytros, 2008)。

IEEE LTSC 组织发布了相关标准文件，指明素养由特定情境中的知识、技能、态度、能力和学习目标构成(IEEE LTSC, 2006)。OECD 教育司长认为素养包含了知识、技能、态度和价值观(OECD, 2005)。

尽管学界对素养定义未达成统一共识，但现有研究种素养的定义仍有相同之处。许多研究对于素养的定义都涵盖了知识、技能、态度和价值观的概念范围，将素养视为基本属性或

是学习内容。除此之外，大部分研究普遍认为，素养与个人表现密切相关，或是将个人表现视作需达成的目标，或是将其作为预设的选拔条件。

2.2 基于本体的素养模型构建

本体包括概念和关系两个要素，构建本体的过程即是明确本体包含概念及相互关系的过程。目前，人工构建是主流的本体构建方式。领域专家通过分析现有的文件、标准和应用，构建领域本体。或是研究者采用德尔菲法，通过访谈或问卷征集专家意见，并分析采集数据，以确认领域本体包含的元素。本体元素确认后，构建者根据现有标准形式化表示本体。Miranda 分析了不同研究的素养概念界定，基于分析结果提出了素养模型，并将评估也融入其中(Miranda 等, 2017)。El Asame 等人基于 IMS RDCEO 和 HR-XML 标准提出了素养模型，该模型中素养由子素养、技能、知识、表现和情境组成(El Asame & Wakrim, 2018)。Paquette 使用 RDF 语言构建了一个素养本体模型。其核心部分包括知识、通用技能、表现、元领域和参考文档，其他构成部分包括活动者、活动、资源和证明记录(Paquette 等, 2021)。Katja 采用了德尔菲法构建出可持续发展核心素养框架，将素养定义为任务成功完成所需具备的知识、技能、动力和态度的集合(Katja 等, 2021)。

除人工构建方法外，有部分研究从自然语言处理任务的角度出发，通过算法实现本体构建过程的自动化。Wang 等人使用 RF-IDF 等算法，构建计算机科学领域的本体，展示了自动化构建本体的过程，并通过检索应用作为结果的验证，证明了算法的有效性(B. Wang 等, 2019)。

人工构建本体的方法得益于专家的参与，保证构建结果的准确性和权威性。但该种方法依赖于专家，导致无论是初次构建还是后续的修订和迭代都需要花费大量的时间和其他资源。自动化构建本体的方法能够减少对专家依赖并提高效率，但需要具备一定的算法、自然语言处理及其他相关领域的知识，或是大量相关数据及计算资源的支撑。这对于计算机科学专业领域以外的研究者来说无疑是一道门槛，为非专业研究者带来了额外的研究负担。

2.3 信息抽取和提示词工程

信息抽取是指从不同形式和结构的信息源中通过系统性的处理过程抽取出有效信息的过程。这些数据以文本、文档、表格形式出现，具有结构化、半结构化或是非结构化的属性。一般情况下，数据源中包含最多的是非结构化的数据，尤其是以文本形式出现居多，因此从非结构化的数据中抽取有用的信息是自动构建本体模型和知识图谱基本步骤。信息抽取任务最基本的方法是基于规则或者词典分别抽取实体和关系，统计学习和机器学习的方法也是广泛应用的方法。随着 GPT 模型展现出色的自然语言处理能力，部分研究者以命名实体识别任务和关系抽取为指向，针对性地训练和调整 GPT 模型。

提示词工程是指开发和优化提示词以帮助大语言模型能够更好地理解用户意图，从而提升大语言模型的表现。目前，研究者已提出了多种提示词设计策略。Few-shot 是一种被广泛使用的提示词设计方法(Brown 等, 2020)。用户为 GPT 模型提供特定任务的样例，ChatGPT 能够在整体上提升输出结果的质量。思维链（Chain of Thought, CoT）也是一种常用的提示词设计方法(Wei 等, 2023)。用户在设计提示词时，清楚地说明特定任务的解决过程，这能够显著提升模型在复杂推理等任务的表现。CoT 方法提出后，不少研究者扩展了此方法，形成了各种变体，如自洽的 CoT(X. Wang 等, 2023) 等设计策略。CoT 的设计策略和其他设计策略，如思维树（Tree of Thought, ToT）、思维图（Graph of Thought, GoT）等，被统称为思维万物 XoT(Ding 等, 2023)。总之，上述研究表明根据特定原则设计提示词对于提高模型性能的有效性。与预训练、微调或其他针对性调整大语言模型的方法相比，提示词工程是一种简单的、高效的、经济的调整方式。本文希望提出一种半自动化的本体构建方式，使用户无需深入学习计算机科学相关的专业知识，通过与 ChatGPT 协作和简洁的代码片段，即可构建出素养模型。这种方法为本体构建提供了实用的解决方案，并提高了构建过程的可用性和简易性。

3. 半自动本体和知识图谱构建方法

3.1 本体构建

半自动化的本体构建过程主要分为四个步骤：首先，进行数据预处理。第二，完成实体和关系抽取。将预处理完的数据放入提示词中，并输入给 GPT 模型。本研究使用的是 ChatGPT3.5 模型，在这个过程中需数次修订和迭代提示词以优化模型输出。第三，定义概念和关系。筛除模型输出错误的结果并进行加以调整和修正，并根据结果概括和归纳概念和关系，不断更新素养模型。最后，进行本体的表示、管理与可视化工作。

本研究以中国义务教育阶段信息科技课程中的素养为例，将《新课标》为素养模型构建的依据。初始的数据源是非结构化的，信息包含在以图片形式存储的 PDF 文件中。应用格式转化工具将初始文件转化为 TXT 格式的文件。转化后的文件中，既包括了非结构化的文本数据，也包括半结构化的表格信息。然后对文本进行切分，并使用正则表达式将划分的段落和句子整合入提示词中，确保输入长度在 ChatGPT 的字数限制以内。

ChatGPT 通过对话的交互方式完成信息抽取任务，提供关于素养本体的必要信息。为了 GPT 模型能够给予更好的反馈，我们参考提示词设计原则对提示词进行设计和优化。首先，提示词中需确保无拼写或语法错误。其次，用户应在提示词中对任务进行清晰的描述，所写句子应当尽量简短、用词应当准确、描述应当简洁。此外，提示词的设计是一个由简到繁、逐步完善的过程。初始提示词应简单明了，在后续的迭代中逐渐补全内容，以达到理想的输出结果。将任务、情境、输入数据、输出格式写入提示词中能增强 GPT 模型对于任务的理解，有助于 GPT 模型输出所需的结果。除严格遵守这些原则外，在设计提示词时还有其他的技巧，提示词设计技巧及对应内容如表 1 所示，本研究中所使用的提示词具体如图 1 所示。

表 1 提示词设计技巧及对应内容

提示词设计方法	对应内容
在情境阐述部分包含身份信息	你是一名资深的信息科技课程教师.....
将输出统一为JSON格式，同时限制键值对的形式。	.....将最终的识别结果以JSON格式输出，每个JSON元素中都包含了学习内容和对应的掌握要求，每个元素都有三个键值对，键分别是“句子”、“表现”和“内容”.....
提供在遇到可能出现的情况时的不同应对方法	.....如果句中无表现和内容，跳过该句子。..... .....如果同一句子中有多个表现和内容，请将所有表现和学习内容都抽取出并放入不同的JSON元素。.....
根据ChatGPT的错误结果，添加要求和限制	.....请输出所有的JSON元素。..... .....请确保每个JSON元素中的表现和内容在句子中有对应关系。.....
拆解抽取任务并描述任务的每一个步骤	.....为了能够完成抽取，需要依次进行以下的操作： 步骤一：以句号(。)和分号(；)为分隔符，以句子为单位分隔文本，并对每一个句子进行如下操作； 步骤二：识别句子中包含在信息科技课程中的学习内容； 步骤三：识别步骤二中学习内容对应的表现； 步骤四：生成当前句子JSON元素，其中“句子”键的值为当前句子，“表现”键的值为步骤三的识别结果，“内容”键的值为步骤二中识别出与表现对应的学习内容。 .....

3.2 知识图谱构建

完成本体构建后，本研究进而以本体为框架，基于 GPT 模型抽取出的概念和实体构建知识图谱，从而表示素养的结构和内涵。知识图谱的构建过程如下：第一，完成实体分类和链接。将实体根据概念进行分类并链接到对应概念，同时根据对应概念之间的关系确定实体之间的关系。其次，进行实体消歧，消除指向同一客体的冗余实体。最后，可视化知识图谱。

好的知识图谱应当考虑经济性和针对性，减少人力和机器的时间成本，使得用户能够高效地管理、检索和查询。为此，消除冗余的信息是必要之举。在实际过程中，自然语言的表达可能使用同义词或多种表示方式指代同一实体。因此，本研究使用现有的中文同义词数据库，如 OpenKG 中的中文知识同义词库等，进行实体消歧。消除重复实体后，本研究使用 Neo4j 数据库存储实体及关系数据，以为用户提供检索和可视化的服务。

提示词示例

你是一个资深的信息科技课程教师，需要阅读以下的文本并从中识别出信息科技课程相关的内容以及与内容对应的掌握要求。将最终的识别结果以JSON格式输出，每个JSON元素中都包含了学习内容和对应的掌握要求，每个元素都有三个键值对，键分别是“句子”、“表现”和“内容”。

注意：
待处理文本以三个单引号(‘’’)为分隔符。
请准确地识别出所有的表现和内容，将识别结果以JSON格式输出，JSON元素包含“句子”、“表现、”内容“键。
请确保每个JSON元素中的表现和内容在句子中有对应关系。
请输出所有的JSON元素。
如果句中无表现和内容，跳过该句子。
请确认每个JSON元素中仅包含有对应关系的一个表现和一个学习内容。如果有多个表现和内容，请将所有表现和学习内容都抽取出并放入不同的JSON元素。
如果表现和内容中有代词，请使用具体内容将代词替换。

为了能够完成抽取，需要依次进行以下的操作：
步骤一：以句号(。)和分号(；)为分隔符，以句子为单位分隔文本，并对每一个句子进行如下操作；
步骤二：识别句子中包含在信息科技课程中的学习内容；
步骤三：识别步骤二中学习内容对应的表现；
步骤四：生成当前句子的JSON元素，其中“句子”键的值为当前句子，“表现”键的值为步骤三的识别结果，“内容”键的值为步骤二中识别出与表现对应的学习内容。

```
示例一：
  文本: '''.....具有信息意识.....'''
  输出: '''
  [
    ...
    {
      "句子": ".....具有信息意识....."
      "表现": "具有",
      "内容": "信息意识",
    },...
  ]'''
示例二：
  文本: '''.....探索大规模数据.....'''
  输出: '''
  [
    ...
    {
      "句子": ".....探索大规模数据.....",
      "表现": "探索",
      "内容": "大规模数据",
    },...
  ]'''

文本: '''[素养描述文本]'''.
```

图 1 输入的提示词内容

4. 结果

4.1 素养本体

经过本体构建过程，我们构建出了中国义务教育阶段信息科技课程的素养本体模型。该模型采用了 OWL 的表示标准，具体构成如下：

知识、技能、态度和价值观类。本研究参考 DeSeCo（Definition and Selection of Competencies）提出的定义(OECD, 2005)，并创建相应类：知识类、技能类、态度和价值观类。

表现类。表现类是此本体中基础的类之一。尽管对素养概念的定义尚未达成共识，但与表现相似的词语却经常出现在素养的定义中，表现是素养的重要观测指标。因此，记录表现对素养测量是必要的。教师设定教学目标时，也需要根据具体课程和具体情景明确具体表现。

领域类。根据布鲁姆分类方法，划分出认知、情感和动作技能三个领域。每个领域中都有相应的目标，这些目标与本体中的表现等级相对应。知识、技能、态度和价值类中的元素都能够依次归入各个领域中。表现类也根据领域和表现等级进行更细粒度的分类。

为了本体的准确性和全面性，应调整初始本体以囊括被排除在外的概念。在本研究中，通过抽象具有相似特征的概念和关系，最终形成以下类别：课程、素养、元领域、领域、学段、EKSAPO、知识、技能、态度和价值观、学习产出、学习工具。其中，元领域是对领域的抽象，EKSAPO 则表示知识、技能、态度和价值、学习产出、学习工具实体组成的类别。

考虑到不同概念类别之间的联系，本研究确定了“包含素养”、“包含 EKSAPO”、“属于学段”、“属于等级”、“属于领域”等关系。其中“包含 EKSAPO”关系又被划分为“包含知识”等，表示 EKSAPO 及其子类间的关系。信息科技课程素养的本体框架如图 2 所示。

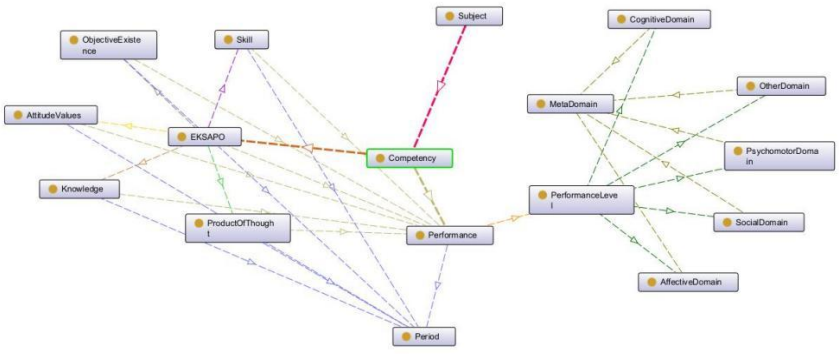


图 2 素养本体框架

4.2 素养知识图谱

完成构建过程后，构建出两个主要的词库：EKSAPO 词库，包含 327 个词元素；以及表现词库，包含 164 个词元素。EKSAPO 词库包含了与知识、技能、态度和价值、学习产出等相关的实体。此外，人工添加了课程、素养、学段、领域等词库。完成实体消歧后，表现词库的元素个数由 164 个降低至 70 个。最终，知识图谱中共有 417 个表示实体的节点。

文本材料被分割为不同的部分，每部分都描述了一种子素养。因此，可以根据各实体所在句子、段落或章节的位置这一隐含信息，推断实体对应素养和学段，确认素养与 EKSAPO 实体、学段与表现及学段与 EKSAPO 实体间的关系，分别是“包含 EKSAPO”（包含知识、包含技能等）、“包含表现”、“属于学段”。“具备表现”关系表示 EKSAPO 实体与表现实体间的联系，输出 JSON 元素中“表现”与“内容”默认具备这种关系。同时，“具备素养”、“分为领域”、“属于领域”、“属于等级”关系分别是学科与素养、元领域与领域、表现等级与领域及表现与表现等级间的关系。在最终的知识图谱中，实体之间总共存在 1390 条关系。图谱的概况如图 3 所示。以“算法”为关键词进行检索所得结果如图 4 所示。

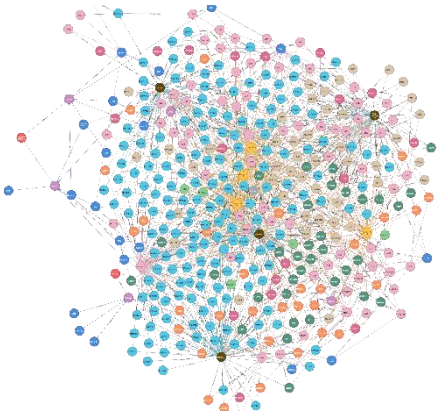


图 3 素养图谱概况

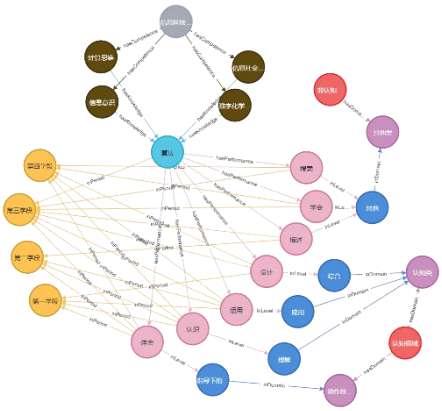


图 4 以“算法”为关键词检索结果

5. 验证

本研究使用文本数据抽取实体和关系，构建出领域本体和知识图谱。完成半自动构建任务后，我们对原始抽取结果进行了修正和调整。因此，我们将人工抽取结果与 GPT 模型抽取的结果进行对比，以评估 ChatGPT 在抽取任务上的表现。本文使用准确率（Accuracy）、精确率（Precision）、召回率（Recall）、F1 值和 Kappa 值来评估本体和知识图谱半自动化构建过程的有效性。我们不仅考察表现与学习内容的抽取结果，同时也将元组中表现与内容的对应关系进行检测。各项指标的计算结果如表 2 所示。

表 2 两个词库和抽取元组的验证结果

指标	表现词库	学习内容词库	抽取元组
精确率	0.90	0.79	0.80
召回率	0.67	0.59	0.60
准确率	0.62	0.51	0.52
F1	0.77	0.67	0.69
Kappa	0.79	0.65	0.67

在 GPT 模型的计算结果中，Kappa 值均大于 0.6，表明 GPT 模型抽取的实体和元组与人工抽取的结果有较好的对应关系。精确率也均高于 0.75，说明 GPT 模型抽取的有效结果能够

涵盖四分之三以上的人工抽取结果。然而，GPT 模型输出结果的召回率和准确率值在 0.5-0.65 区间，GPT 模型忽略、误抽，甚至捏造了相当一部分实体和元组。表现实体抽取结果的所有指标均为最高，元组结果其次，学习内容指标结果最低。GPT 模型识别表现类实体效果较好，但识别学习内容实体的表现欠佳。元组抽取效果下降的主要原因可能是模型对学习内容的识别效果不佳，模型匹配表现与学习内容时产生的失误影响较小。

6. 讨论

针对抽取结果，我们分析了谓词、客体和元组抽取产生错误的情况，并总结了常见的问题类型。多数情况下，抽取结果的问题体现在无法准确设定谓词和客体的边界从而囊括多余信息。特殊句式结构是抽取频繁出错的重要原因之一。以下是易出错的两类特殊句式结构。

ChatGPT 在处理具有特殊结构的动词和短语时，容易出现失误。特殊动词，如处理 “能够（做）”结构时， ChatGPT 会将“能够”作为谓词，并将真正的谓词混入学习内容。“对…（做/进行）…”结构中的客体和谓词顺序与普通句式相反，且谓词以名词形式作为另一个宾语出现。同时，“对”本身有动词含义，如“面向”、“应答”等，也可能被误判为谓词。“使用…（做）…”是双宾语短语，包含两对“谓词-客体”，前面的“使用”与其客体作为后面动作的方式。但 GPT 模型错将“使用”之后的整个部分都视为该词的客体输出。

ChatGPT 在处理并列关系时难以控制抽取粒度。并列关系可能出现在谓词中、客体中，或是二者均有，甚至是描述谓词的副词和形容客体的定语也会出现并列关系，如 “信息存储和传输过程中的编码及解码步骤”。并列的谓词和客体应拆分到不同元组中，确保抽取结果的最小颗粒度，而 ChatGPT 往往将整部分的谓词和整部分的客体全部输出到同一元组中。

除了上述两种最常见的情况以外，跳过应抽取的元组也是频繁发生的问题。同时，尽管在提示词中要求 GPT 模型将代词替换为所指内容，但 GPT 模型并没有很好得完成指代消解任务。针对以上的输出情况，我们调整了提示词，尤其是针对特殊动词和短语的情况，在提示词中加入了这些特殊结构的例子。同时，人工修订提取的元组结果，并反馈给 GPT 模型。

7. 总结

本文提出了基于文本数据的半自动化本体构建方法，与 ChatGPT 协作进行概念和关系抽取。结果表明，ChatGPT 的抽取结果能够覆盖大部分人工抽取的结果，促进构建领域本体的过程。后续工作可以聚焦于知识图谱的优化和应用方面。虽然通过文本的位置关系获得了一些隐藏在材料中的关系，但仍有忽略的部分。例如，知识、技能等实体之间存在着父子关系和先后关系。其次，三元组的表示形式也存在局限，如在面对“利用数字化平台处理数据”这种两个谓词-客体对之间存在关系的情况时，元组间的关系无法体现。另外，如何将课程素养的本体和知识图谱应用于素养的培养、管理和评估等工作是图谱发挥其价值的重要问题。

参考文献

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Christiaens, S., De Bo, J., & Verlinden, R. (2006). Competency Model in a Semantic Context: Meaningful Competencies (Position Paper). R. Meersman, Z. Tari, & P. Herrero , *On the*

Move to Meaningful Internet Systems 2006: OTM 2006 Workshops (Page 1100–1106). Springer. https://doi.org/10.1007/11915072_12

Ding, R., Zhang, C., Wang, L., Xu, Y., Ma, M., Zhang, W., Qin, S., Rajmohan, S., Lin, Q., & Zhang, D. (2023). *Everything of Thoughts: Defying the Law of Penrose Triangle for Thought Generation* (arXiv:2311.04254). arXiv. http://arxiv.org/abs/2311.04254

El Asame, M., & Wakrim, M. (2018). Towards a competency model: A review of the literature and the competency standards. *Education and Information Technologies*, *23*(1), 225–236. https://doi.org/10.1007/s10639-017-9596-z

IEEE LTSC. (2006). *Draft Standard for Learning Technology Standard for Reusable Competency Definitions*. https://www.ieeeltsc.org/wg20Comp/wg20rcdfolder/IEEE_1484.20.1.D3.pdf

Katja, B., Barth, M., Gisela, C., Cohen, M., Diaz, L., Doucette-Remington, S., Weston, D., Geoffrey, H., Niki, H., Meghann, J., Kealalokahi, L., Yoko, M., Rieckmann, M., Roderic, P., Walker, P., & Michaela, Z. (2021). Key competencies in sustainability in higher education—Toward an agreed-upon reference framework. *Sustainability Science*, *16*(1), 13–29. https://doi.org/10.1007/s11625-020-00838-2

Kiesler, N. (2020). Towards a Competence Model for the Novice Programmer Using Bloom's Revised Taxonomy—An Empirical Approach. *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, 459–465. https://doi.org/10.1145/3341525.3387419

Miranda, S., Orciuoli, F., Loia, V., & Sampson, D. (2017). An ontology-based model for competence management. *Data & Knowledge Engineering*, *107*, 51–66. https://doi.org/10.1016/j.datak.2016.12.001

OECD. (2005). *The definition and selection of key competencies:Executive summary*. https://www.oecd.org/pisa/35070367.pdf

Paquette, G., Marino, O., & Bejaoui, R. (2021). A new competency ontology for learning environments personalization. *Smart Learning Environments*, *8*(1), 16. https://doi.org/10.1186/s40561-021-00160-z

Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., & Yang, D. (2023). *Is ChatGPT a General-Purpose Natural Language Processing Task Solver?* (arXiv:2302.06476). arXiv. http://arxiv.org/abs/2302.06476

Sampson, D., & Fytros, D. (2008). Competence Models in Technology-Enhanced Competence-Based Learning. 收入 H. H. Adelsberger, Kinshuk, J. M. Pawlowski, & D. G. Sampson (编), *Handbook on Information Technologies for Education and Training* (页 155–177). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74155-8_9

Wang, B., Luo, J., & Zhu, S. (2019). Research on Domain Ontology Automation Construction Based on Chinese Texts. *Proceedings of the 2019 8th International Conference on Software and Computer Applications*, 425–430. https://doi.org/10.1145/3316615.3316685

Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E., Narang, S., Chowdhery, A., & Zhou, D. (2023). *Self-Consistency Improves Chain of Thought Reasoning in Language Models* (arXiv:2203.11171). arXiv. https://doi.org/10.48550/arXiv.2203.11171

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models* (arXiv:2201.11903). arXiv. https://doi.org/10.48550/arXiv.2201.11903

人机协同视域下 ChatGPT 辅助师范生开展教学设计的创新探索

Innovative Exploration of ChatGPT-Based Instructing Design for Normal Students from the Perspective of Human-Computer Collaboration

许洁[1]，孙丹[2]，李艳[3,*]
[1]浙江大学教育学院
[2]杭州师范大学经亨颐教育学院
[3]浙江大学教育学院
[*] yanli@zju.edu.cn

【摘要】 教学设计能力是师范生应具备的核心教学能力，当前在提升师范生教学设计能力过程中仍存在教师反馈滞后等问题。以 ChatGPT 为代表的生成式人工智能为教学工作提供了全新机遇，本研究尝试通过准实验研究来探究 ChatGPT 辅助师范生教学设计，师范生需要在为期 8 周的课程中完成三个版本教学设计。研究发现，不同版本教学设计存在差异，ChatGPT 和师生在一定程度上能够辅助师范生，以提高教学设计作品质量；ChatGPT 辅助教学设计过程中，大多数师范生保留了思考力，但存在伦理问题。鉴于此，研究提出 ChatGPT 辅助师范生教学设计的相关建议，期望为师范生培养工作提出新思路。

【关键词】 人机协同；ChatGPT；师范教育；教学设计

Abstract: *Instructional design ability is the core teaching ability that normal students should possess. There are still problems such as the lag of teacher feedback in the process of improving normal students' instructional design ability. Generative artificial intelligence, represented by ChatGPT, offers opportunities for instructing. This study explores the instructional design of ChatGPT-assisted through quasi-experimental. Students were required to complete three versions of instructional design works during the 8-week course. The findings show that there are differences between different versions of works, ChatGPT and teachers can help normal students, so as to improve their instructional design's quality. In the process of ChatGPT-assisted, most normal students retain thinking ability, there are ethical problems. The study puts forward some suggestions, which are expected to provide new ideas for the training of normal students.*

Keywords: human-computer collaborative, ChatGPT, teacher education, instructional design

1. 前言

师范生是未来教师队伍的主力军，培养师范生的教学设计能力是师范教育中的重要环节。2021 年，教育部颁布的《中学教育专业师范生教师专业标准（试行）》中强调教学设计是教师专业能力的重要组成部分。在大学课堂上，师范生撰写教学设计是教师教育模块课程中必不可少的环节，以此提高教学设计能力。教师反馈能够有效促进学习者的学习投入，提高其学业表现（董艳等，2021），然而，针对师范生的教学设计撰写需求，教师难以做到分身乏术，及时反馈（余明华等，2024）。此外，在教师教育实践中，少有课堂采用人机协同教学模式，师范生在未来实际教学中，将难以应对人机协同教育中的教学（周跃良等，2022）。

ChatGPT 作为对话式学习机器人，其强大的文本生成功能给教育工作者带来了极大的便利。多位研究者提出，ChatGPT 的有问必答特性可以辅助教师开展教学设计，减轻教师备课的负担（何文涛等，2023）；也可以为学生提供及时反馈，减轻教师课上的教学压力（余胜

泉等，2024）。然而，ChatGPT 能在多大程度上帮助师范生开展教学设计，是否会对师范生的教学设计过程产生影响，以及如何影响师范生教学过程中的行为模式等一系列问题仍缺乏实证研究，有待进一步探索。鉴于此，本研究基于人机协同视角，尝试将 ChatGPT 融入师范生课堂，同时结合师生反馈，探索 ChatGPT 及师生反馈辅助师范生开展教学设计的相关实践。由此探索 ChatGPT 对教学设计的应用潜能，为人工智能技术融入未来教育教学提供思路和参考。

2. 文献综述

2.1. 师范生教学设计能力培养过程中存在的问题

教学设计能力是师范生应具备的核心教学能力，美国教育与传播技术协会（AECT, 2012）将其划分为设计、开发、利用、管理和评价五种能力；陈花等（2024）提出师范生教学设计能力包括 6 个基本要素和 24 个能力要点，其中 6 个要素是教学内容分析能力、学生情况分析能力、教学目标设计能力、教学方法设计能力、教学媒体设计能力和教学过程设计能力；教学设计能力正是从多种教学能力提取出的上位能力。在大学课堂中，撰写教学设计是师范生在教学实践中重要的环节。然而，当前在培养师范生教学能力过程中仍存在一些问题。一方面，理论基础薄弱是主要问题之一（陈薇，2018）。部分师范生对教学设计理论知识掌握不深，难以进行科学有效的教学设计。另一方面，实践能力不足也是问题突出所在（李飞等，2022）。师范生缺乏上课实战经验，难以将理论转化为操作能力，个人教学设计水平参差不齐，对于所教学生的最近发展区把握较弱。与此同时，缺乏实践反馈机制也影响师范生的成长。在一般的课程教学中，课程人数较多，教师无法照顾每位师范生并为其提供个性化的教学设计反馈。然而，师范生需要得到教师或同伴的及时反馈，才能继续开展接下来活动。鉴于以上困境，师范生亟需得到能够及时反馈的个性化助手，为其提升教学设计能力提供帮助。

2.2. ChatGPT 辅助教学设计的机遇与挑战

ChatGPT 可以为教师提供帮助和建议，规划课程思路，生成课程材料（Wang et al., 2023），为辅助教学设计提供了可能性。已有研究明确指出由 ChatGPT 担任的智能教师可以根据特定主题设置任务并提供相关资料，如问题设计，测验构思等（王天健, 2023）。宋葭和林敏（2023）两位学者提到，在备课时，ChatGPT 可以为教师提供教学资源，如教学的背景知识，生成课堂讲义、练习题等。具体而言，在辅助教师生成个性化教学方案的过程中，ChatGPT 等依据教师的教学需求，制定学习目标，创建合适特定需求的课程大纲与课程计划，分步骤生成多种适切的教学设计，为教师在备课过程中提供思路启发与多种备选方案，整合和推荐网络上的教学资源，提升教师的备课效率与授课质量。

然而，虽然 ChatGPT 在帮助教师完成诸如生成课程材料、提供建议、通过回答问题和促进合作为学生充当虚拟导师等任务方面具有巨大潜力，但它也提出了一些挑战，如生成不准确或捏造的信息（Wang et al., 2023）。王天健学者（2023）提出通过 ChatGPT 开展教学设计目前尚存在两大问题，首先，ChatGPT 并不能为自己提出的具体方案负责，其回答仍然存在可靠性问题；此外，ChatGPT 所提供的方案普适性强、针对性弱，缺乏围绕特定任务、情境以及对象开展教学设计的能力。因此，ChatGPT 所提供的教学设计仅仅是对人类教学设计在语言逻辑上的模仿，并不真正具备设计教学的能力。面对真实的教学情境，人类教师仍具有不可替代的育人价值（冯志伟等，2023），教师的反馈对师范生完善教学设计至关重要。

2.3. ChatGPT 辅助教学设计相关实证研究

Megahed 等（2023）要求 ChatGPT 为本科统计学课程量身定制课程大纲，并指出这些教学建议只需要很少的修改。Zhai（2023）表示，ChatGPT 在特殊教育背景下提供了有价值的建议，他认为这对有特殊学习要求的学生尤其有利。在德国教育实践中，ChatGPT 对教师教

学设计的工作效率提升十分显著（王天健，2023），德国教育界对 ChatGPT 介入教学设计的态度较为积极，对其未来发展前景保持乐观（Deutscher Bundestag，2023）。然而，当前我国通过 ChatGPT 辅助教学设计的相关实证研究较少，未能看出 ChatGPT 在辅助教师教学设计方面提供哪些具体的帮助。因此，相关实证研究亟需加强。

综上所述，应当迅速调整当前教学中的教学方法以应对这些问题和挑战。本研究拟聚焦人机协同的教学设计实践，探究融入 ChatGPT 和师生反馈对师范生教学设计的影响。师范生在实践过程中需要完成三个版本的教学设计，分别为小组首次完成的教学设计 V1 版本，与 ChatGPT 协作之后完善的教学设计 V2 版本，师生提供反馈后完善的教学设计 V3 版本。研究包含以下四个问题：

（1）V1、V2 和 V3 三个版本的教学设计是否存在差异，如果存在，具体有哪些差异？

（2）在使用 ChatGPT 辅助教学设计时，师范生存在怎样的行为模式？

3. 研究设计

3.1. 研究对象

研究选取 Z 大学选修“科学与技术教育教学设计与实施”课程的 39 名学生（男生 16 人，女生 23 人）作为研究对象。研究对象分别是来自 Z 大学国优计划和科学与技术教育专业的学生，其专业背景多样，之前均没有相关教学经验。该课程目标即系统地为学生教授教学设计模块，为中小学培养专业素养卓越、教学基本功扎实的优秀教师。

3.2. 研究过程

研究对象全程参与到课程学习中，实验课程持续 8 周，每周 4 个课时。前四周，教师讲解教学设计基本理论，学生解读与自身专业背景相符的中小学学科课程标准，寻找优秀课程案例；第五周，研究对象随机分组，1-3 人为一组，共形成 20 组，每组成员根据自己的专业背景选择合适的教学设计选题，查阅相关教学资源，完成教学设计初始版本 V1；第六周，教师向学生发放与 ChatGPT 交互的提问工程语资料，小组成员熟悉并使用 ChatGPT-3.5，然后与 ChatGPT 协同完善教学设计，完成教学设计二次版本 V2；第七周至第八周，各小组汇报展示，师生对汇报小组进行反馈，各小组根据师生反馈完善教学设计，提交教学设计最终版本 V3。

3.3. 数据收集与分析

本研究收集的数据包括小组完成的三个版本教学设计和小组与 ChatGPT 的交互对话数据，拟从多个角度探究 ChatGPT 辅助师范生开展教学设计的效果。针对研究问题一，将每组教学设计作品的 V1，V2，V3 版本进行对比，通过 SPSS 进行配对样本 t 检验，得出不同版本教学设计作品的教材分析、学情分析、教学目标、教学重难点、教学过程等维度是否存在差异，以及有哪些差异；针对研究问题二，根据小组与 ChatGPT 交互的对话数据，以及小组教学设计的 V1 和 V2 版本，通过滞后序列分析，得出 ChatGPT 辅助师范生教学设计时的行为模式。

研究参考 AECT（2012）颁布的教学设计能力标准和胡久华等（2022）对师范生教学设计技能的维度划分及水平层级，分别从分析教学内容、分析教学对象、设计教学目标、设计教学重难点、选择教学策略、选择教学媒体、设计教学内容、设计教学活动、设计教学评价、设计板书十个维度对师范生三个版本的教学设计作品进行打分，每个维度占比 10 分，满分为 100 分。两位研究人员结合教学设计作品的学科背景，通过协商后分别对分析教学内容、分析教学对象、设计教学目标等维度进行打分，并进行一致性比较（Kappa 系数大于 0.8）。

为更好地了解师范生在 ChatGPT 辅助教学设计中的行为模式，研究者将行为分为师范生的提问与 ChatGPT 的反馈两类，提问行为参考丹东尼奥等（Dantonio, 2006）的提问结构，该结构由多种提问方式如开放性提问、引导性提问、评价性提问及各种追问组成，结合师范生提问的实际情况，将分析维度划分为指示（ANQ）、评价（EGC）、示范（DRE）、追问（FGC）

和继续（CGC）五类（见表 1）。对 ChatGPT 提供的反馈内容及学习者的教学设计作品分析，为更好地了解学习者对 ChatGPT 反馈内容的回应情况，参考 Yan 等（2023）提到的学生对 ChatGPT 反馈内容的行为描述，即不修改、少量修改、大量修改人工智能生成的文本等，将分析维度划分为原样拷贝（CW）、部分原样拷贝（CP）、修改后拷贝（CM）、无需拷贝（CN）四类（见表 1）。根据表 1，两位研究人员对师范生的提问行为及对 ChatGPT 反馈内容的拷贝行为进行独立编码，编码一致性检验结果较好。

表 1 ChatGPT 辅助师范生教学设计行为编码表

分类	编码	活动名称	行为描述
提问	ANQ	Asking new questions	学生在 ChatGPT 平台自主提出新的问题
	EGC	Evaluating generated content	学生对生成内容正误进行评价
	DRE	Demonstrating the right example	学生提供教学设计范例或正确的内容生成
	FGC	Following-up generated content	学生针对生成内容进一步追问
	CGC	Continuing to generate content	学生要求 ChatGPT 继续生成未完成的内容
反馈内容	CW	Copying the whole	学生将 ChatGPT 平台中全部反馈内容拷贝
	CP	Copying in part	学生将 ChatGPT 平台中部分反馈内容拷贝
	CM	Copying after modification	学生将 ChatGPT 平台中反馈内容修改后拷贝
	CN	Copying of nothing	学生无需将 ChatGPT 平台中的反馈内容拷贝

4. 研究发现

4.1. 不同版本教学设计作品比较

研究使用配对样本 t 检验对三个版本的教学设计作品得分进行分析，结果如表 2 所示。V2 和 V3 在设计教学内容、设计教学活动、设计教学评价三个维度上均比上一版有显著提升（差值均值>0，且 $p<0.05$）；此外，V2 相较 V1 在分析教学内容、设计教学目标、选择教学策略、选择教学媒体维度有显著提升，V3 相较 V2 在分析教学对象维度有显著提升，但三个版本在设计教学重难点、设计板书维度上不存在显著差异。结合总体得分，V3 得分更高（Mean (V1)=61.08, Mean (V2)=71.55, Mean (V3)=75.05）。实验结果表明，ChatGPT 辅助师范生开展教学设计，能够在分析教学内容、设计教学目标、选择教学策略、选择教学媒体、设计教学内容、设计教学活动、设计教学评价等维度提供帮助；师生反馈提供的帮助体现在分析教学对象、设计教学内容、设计教学活动、设计教学评价方面。

表 2 三个版本教学设计作品得分的配对样本 t 检验（n=20）

三个版本	V2 — V1		V3 — V2		V3 — V1	
教学设计作品	M±SD	t	M±SD	t	M±SD	t
分析教学内容	0.35±0.67	2.333*	0.45±1.39	1.443	0.80±1.44	2.491*
分析教学对象	0.35±0.81	1.926	0.50±0.89	2.517*	0.85±1.13	3.344**
设计教学目标	0.78±1.11	3.134**	0.25±0.81	1.926	1.13±1.23	4.077**
设计教学重难点	0.35±1.46	1.071	0.35±0.99	1.584	0.70±1.69	1.853
选择教学策略	1.40±1.76	3.559**	0.03±0.11	1.000	1.43±1.74	3.659**
选择教学媒体	1.15±1.27	4.056**	0.10±0.45	1.000	1.25±1.41	3.966**
设计教学内容	1.65±1.50	4.931***	0.25±0.44	2.517*	1.90±1.59	5.357***

三个版本	V2 — V1		V3 — V2		V3 — V1	
教学设计作品	M±SD	t	M±SD	t	M±SD	t
设计教学活动	1.65±1.53	4.819***	0.20±0.41	2.179*	1.85±1.60	5.176***
设计教学评价	1.95±1.85	4.717***	1.25±1.12	5.000***	3.20±1.64	8.718***
设计板书	0.85±1.84	2.062	0.03±0.11	1.000	0.88±1.83	2.133*
总分	10.48±8.26	5.670***	3.50±2.70	5.787***	13.96±8.84	7.066***

p<0.05*, p<0.01**, p<0.001***

4.2. ChatGPT 辅助教学设计时师范生的行为分析

研究共收到 20 个小组 511 个行为转换序列，为了分析学习者在使用 ChatGPT 辅助教学设计过程中的行为，本研究采用滞后序列分析法（LsA）对编码结果进行分析，使用 Yule's Q 来表示过渡关联的强度，该参数具有一定的描述性（范围从-1 到+1，其中 0 表示无关联）（Chen et al., 2017）。节点代表学习者行为及其出现的频次，边的数值代表序列转换的强度（见图 1）。

根据统计结果，学习者的行为分布具有以下特征：（1）提问类行为和反馈内容类行为各占整个行为分布的一半，这与数据来源的属性较为符合。（2）在提问类行为中，频次最高的行为是“指示”（ANQ），频次为 152，占所有行为 30%；其次是“追问”（FGC，13%）和“继续”（CGC，5%），频次分别为 67 和 28；最后是“评价”（EGC，1%）和示范（DRE，1%），频次分别为 5 和 6。这说明学习者在教学设计过程中能够不断提出新问题，并且对 ChatGPT 反馈的内容产生兴趣，会加以追问生成内容。（3）在反馈内容类行为中，除了“无需拷贝”（CN，25%），出现频次较多的行为是“修改后拷贝”（CM），频次为 64，占比 13%；最后是“部分原样拷贝”（CP，7%）和“原样拷贝”（CW，5%），频次分别为 34 和 28。这说明学习者在接收到 ChatGPT 的反馈后，会对其思考并修改，适当地将相关内容摘录到小组的教学设计作品中。

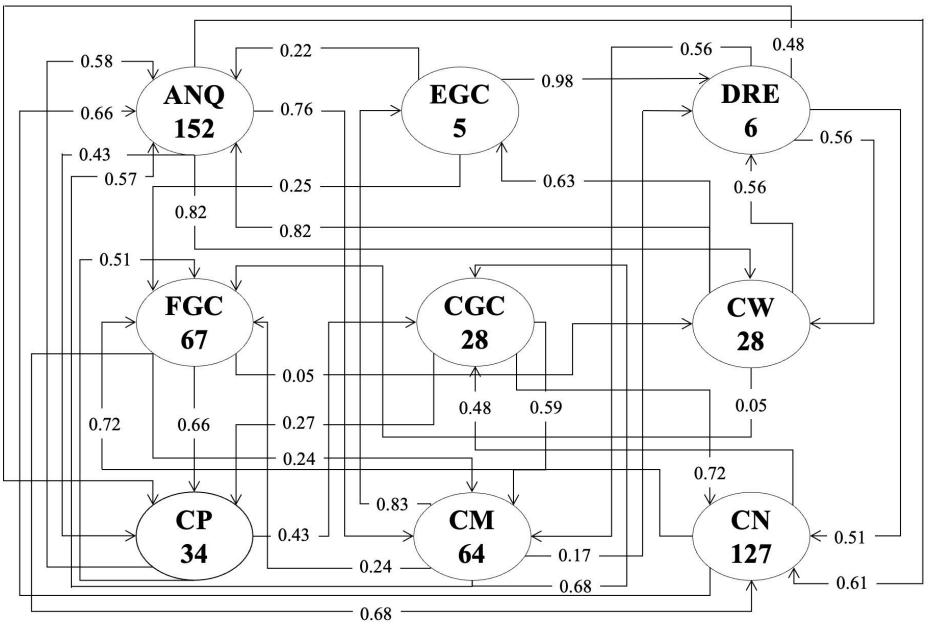


图 1 学习者行为序列转换图

学习者的行为转换序列（除去节点内部的转换序列）存在以下典型特征：（1）评价->示范（EGC -> DRE: Yule’s Q = 0.976），即学习者在收到 ChatGPT 的反馈信息后，会对其给出

的内容正误进行评价，并提供教学设计范例或正确的内容生成。可见，学习者在使用 ChatGPT 辅助完善教学设计的过程中会认真思考，并将其作为对话的工具。(2) 指示->原样拷贝（ANQ -> CW: Yule's Q = 0.817），指示->修改后拷贝（ANQ -> CM: Yule's Q = 0.765），即学习者会向 ChatGPT 提出新问题，在接收到 ChatGPT 的反馈后，学习者会原样拷贝或修改后拷贝。可能对于较为简单的问题，ChatGPT 的反馈容易使人接受，学习者会将其原样拷贝在自己的教学设计作品中；但相对于较为复杂的问题，ChatGPT 给出的反馈需要学习者进行修改后拷贝，表明学习者在使用 ChatGPT 辅助教学设计过程中也会有所思考。（3）继续->无需拷贝->追问（CGC -> CN: Yule's Q = 0.724；CN -> FGC: Yule's Q = 0.720），这意味着学习者在输入了"继续"对话后，得到的反馈内容并不满意，因此又会对 ChatGPT 进行追问，直到得到自己满意的反馈内容为止。由此可见，ChatGPT 的反馈内容在一定程度上能够指导学习者完善教学设计，学习者可能认为不断地与 ChatGPT 交互才能获得满意的结果。

5. 结论与讨论

通过对人机协同视域下 ChatGPT 辅助师范生开展教学设计进行创新探索，根据以上研究结果，得出以下结论并进行讨论：

第一，每个小组 V1、V2、V3 版本教学设计作品存在差异，ChatGPT 和师生反馈在一定程度上能够辅助师范生完善教学设计，提升其教学设计作品质量。具体来说，V2 版教学设计相较于 V1 版，在分析教学内容、设计教学目标、选择教学策略、选择教学媒体、设计教学内容、设计教学活动、设计教学评价等维度都有显著提高，这说明 ChatGPT 对师范生教学设计的提升是多维度的（余胜泉等，2024）。在这些维度上，ChatGPT 能够根据小组不同的教学设计需求提供具体的反馈信息。一位学习者提到，ChatGPT 在确定教学目标、构思活动内容以及寻找合适的教学资源方面起了重要作用，例如，在讨论如何通过实际例子解释复杂的三角恒等式时，ChatGPT 提供了一系列创新的教学方法和实际应用场景。除此之外，第三版教学设计相较于第二版，师生反馈除了在设计教学内容、设计教学活动、设计教学评价方面有显著提高外，在分析教学对象维度也起到了一定的促进作用。可能是师生反馈更容易捕捉到教学设计作品中教学对象的阐述和分析，而 ChatGPT 缺乏人性感知（李艳等，2024），在分析教学对象维度不够具体和全面。对于 V3 和 V1 教学设计，除了设计教学重难点没有显著提高，其他维度均显著提高。可能是设计教学重难点对于师范生来说相对容易，教学设计作品中对其改动较小。此外，九个维度的显著提高也意味着在 ChatGPT 和师生反馈的共同辅助下，师范生的教学设计作品质量得到了提升。

第二，ChatGPT 辅助开展教学设计过程中，大多数师范生保留了思考力，但存在伦理问题。从 ChatGPT 辅助师范生教学设计的行为频次和行为转换序列可以看出，师范生会对 ChatGPT 提供的反馈内容进行思考和评价，再决定是否会将其拷贝到自己的教学设计作品中。具体来说，在行为频次方面，师范生的提问（ANQ）行为远远高于其他行为，在提问的基础上，会对 ChatGPT 的反馈进行追问（FGC），或让其继续生成内容（CGC），表明师范生在此过程中会思考 ChatGPT 给出的反馈，进行自主探究和反思，进一步获得满意的回答（孙丹等，2024；余明华等，2024）。行为序列转换中的评价（EGC）->示范（DRE），继续（CGC）->无需拷贝（CN）->追问（FGC）也可以证明这一点。对于指示（ANQ）->原样拷贝（CW），指示（ANQ）->修改后拷贝（CM）这两种行为序列，表明 ChatGPT 给出反馈若合理，师范生会原样拷贝，若反馈不太合理，师范生会进行修改后拷贝，也意味着 ChatGPT 的存在帮助师范生减轻了教学设计压力，提高了效率（何文涛等，2023）。然而，师范生也应当注意其中存在的伦理问题，减少原样拷贝的频率，不可过度依赖，以 ChatGPT 给出的反馈为启发，通过自身的思考和判断，提高教学设计作品的原创性。

6. 启示与建议

基于以上研究发现，针对 ChatGPT 及师生反馈在辅助师范生完善教学设计的实践过程，本研究得出以下启示：

第一，在师范教育尤其是师范生教学设计能力培养的过程中，除了课堂上师生反馈外，鼓励 ChatGPT 在相关课程中的应用，但要求师范生保持独立思考能力，注意伦理问题。当前学习者使用 ChatGPT 的基本情况较为乐观，在课程学习等多个场景都会使用到该工具（李艳等，2024），师生反馈和 ChatGPT 的反馈并行能够提高师范生教学设计作品的质量，同时提高课堂学习效率。然而，考虑到 ChatGPT 的双刃剑作用，学习者应规范自身的使用行为。一方面，学习者应时刻关注我国教育部门及高等教育机构出台的相关指导规范，确保使用 ChatGPT 的行为符合规范，时刻保持负责任的态度。学习者也可以多多交流互动，分享经验和教训，形成学习共同体。另一方面，ChatGPT 存在的虚假信息可能会扭曲科学事实并散播错误信息，并且削弱学生的思考力。因此，需要学习者辩证采取 ChatGPT 的建议，学习者应该明白 ChatGPT 只是一个智能工具，学习者应加强自身“人类优势”的培养，在 ChatGPT 的使用过程中保持人的主动性和独立思考能力，不能完全依赖 ChatGPT 给出的结果，降低“原样拷贝”的频率，防止滥用。学习者在学习的过程中通过与 ChatGPT 合作，共同思考和解决问题，将 ChatGPT 视为人机协同的伙伴，从中获得新的见解和观点，构建人机协同新范式。

第二，指导师范生如何与 ChatGPT 开展高质量对话，以此提高师范生的教学设计作品质量，培养其教学设计能力。由研究结果可知，V2 相较于 V1 版本的教学设计作品得分在多个维度有显著提升，然而，在实践过程中一些学习者仍反映与 ChatGPT 交流起来有困难。首先，培养师范生的教学设计能力需要多种教学策略的支持，教师应当在课堂中融入人机协同的教学模式，引导学生参与实际教学设计项目，在这些项目中，学生将与对话系统进行互动，并从中学到如何整合 ChatGPT 的观点和建议，将其应用于实际的教学设计和实践中。其次，教育者在课堂上需要培养师范生对话的技巧和技能，包括提出明确、有针对性的问题，理解 ChatGPT 的反馈内容，并能够就回应展开深入的思考和探讨。学生需要了解如何引导对话，以便从 ChatGPT 那里获得有价值的反馈和信息。最后，教育者需要帮助学生理解 ChatGPT 的反馈，通过具体的教学设计范例，对反馈内容进行批判性思维，并教导学习者如何将 ChatGPT 的反馈内容融入到自己的教学设计中，从而提高其对教学设计的理解和运用。

参考文献

陈花,王爱玲,段巍等(2024).职前教师教学设计能力检核模型的建构研究[J].教育理论与实践,44(01):43-48.

陈薇(2018).师范生教学专长培养的困境及发展系统策略[J].江苏高教,(11):68-72.

丹东尼奥等(2006).课堂提问的艺术:发展教师的有效提问技能[M].宋玲,译.北京:中国轻工业出版社.

董艳,罗泽兰,杨韵莹等(2021).教育信息化 2.0 时代视角下的教师反馈素养研究[J].电化教育研究,42(08):35-42+58.

冯志伟,张灯柯,饶高琦(2023).从图灵测试到 ChatGPT--人机对话的里程碑及启示[J].语言战略研究,(02):20-24.

何文涛,张梦丽,逯行等(2023).人工智能视域下人机协同教学模式构建[J].现代远距离教育,(02):78-87.

胡久华,褚童,王璇等(2022).师范生化学教学设计能力构成、水平层级及发展进阶研究[J].化学教育(中英文),43(24):70-78.

李飞,黄青青(2022).师范生实践创新素养培育的困囿及其纾解[J].教育学术月刊,(11):45-51.
李艳,许洁,贾程媛等(2024).大学生生成式人工智能应用现状与思考——基于浙江大学的调查[J].开放教育研究,30(01):89-98.
宋萑,林敏(2023). ChatGPT/生成式人工智能时代下教师的工作变革:机遇、挑战与应对[J]. 华东师范大学学报(教育科学版), 41(7): 78-90.
孙丹,朱城聪,许作栋等(2024).基于生成式人工智能的大学生编程学习行为分析研究[J].电化教育研究,45(03):113-120.
王天健(2023).德国教育应对 ChatGPT 的理念与经验——对德国联邦议院调研报告的分析[J].比较教育学报,(05):66-78.
余明华,王龚,卜洪晓等(2024).未来教师如何培养？——人机协同师范教育创新的理论模型与实践进路[J].现代教育技术,34(01):117-126.
余胜泉,熊莎莎(2024).基于大模型增强的通用人工智能教师架构[J].开放教育研究,30(01):33-43.
中华人民共和国教育部. (2021). 教育部办公厅关于印发《中学教育专业师范生教师职业能力标准（试行）》等五个文件的通知. [EB/OL]. 资料引自 http://www.moe.gov.cn/srcsite/A10/s6991/202104/t20210412_525943.html
中华人民共和国教育部. (2023). 教育部关于实施国家优秀中小学教师培养计划的意见. [EB/OL].资料引自 http://m.moe.gov.cn/srcsite/A10/s7011/202307/t20230726_1070952.html
周跃良,吴茵荷,蔡连玉(2022).面向人机协同教育的教师教育变革研究[J].电化教育研究,43(10):5-11.
AECT (2012). *AECT Standards, 2012 version* [EB/OL]. Retrieved from https://www.aect.org/docs/AECTstandards2012.pdf.
CHEN B, RESENDES M, CHAI C S, HONG H Y (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse[J]. *Interactive learning environments, 25*(2): 162–175.
Deutscher Bundestag (2023). *Forschungsausschuss gibt Studie zu Auswirl kungen von ChatGPT auf Bildung und Forschung in Auftrag* [EB/OL]. Retrieved from https://www.bundestag.de/ausschuesse/a18_bildung_forschung/pressemitteilungen/ta-chatgpt-10-02-2023-934002/.
F. M. Megahed, Y.-J. Chen, J. A. Ferris, S. Knoth, and L. A. Jones- Farmer (2023). How generative ai models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study [J]. *Quality Engineering,* 1–29.
Wang M., Wang M., Xu X., et al., (2023). Unleashing ChatGPT's Power: A Case Study on Optimizing Information Retrieval in Flipped Classrooms via Prompt Engineering [J]. *IEEE Transactions on Learning Technologies, 14*(08):1-13.
X. Zhai (2023), Chatgpt for next generation science learning [J]. *The ACM Magazine for Students, 29*(3): 42–46.
Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation [J]. *Education and Information Technologies,* 1-25.

A Study on Educational Data Analysis and Personalized Feedback Report Generation Based on Tags and ChatGPT

Yizhou Zhou [1, *], Mengqiao Zhang[2], Yuan-Hao Jiang[3, 4, 5], Naijie Liu[6], Bo Jiang[3]

[1] School of Design and Engineering, National University of Singapore
[2] National Institute of Education, Nanyang Technological University
[3] Lab for Artificial Intelligence for Education, East China Normal University
[4] Shanghai Institute of Artificial Intelligence for Education, East China Normal University
[5] School of Computer Science and Technology, East China Normal University
[6] School of Management Science and Engineering, Shandong University of Finance and Economics
* e1010686@u.nus.edu

Abstract: *This study introduces a novel method that employs tag annotation coupled with the ChatGPT language model to analyze student learning behaviors and generate personalized feedback. Central to this approach is the conversion of complex student data into an extensive set of tags, which are then decoded through tailored prompts to deliver constructive feedback that encourages rather than discourages students. This methodology focuses on accurately feeding student data into large language models and crafting prompts that enhance the constructive nature of the feedback. The effectiveness of this approach was validated through surveys conducted with over 20 mathematics teachers, who confirmed the reliability of the generated reports. This method can be seamlessly integrated into intelligent adaptive learning systems or provided as a tool to significantly reduce the workload of teachers, offering timely and precise feedback to students. By transforming raw educational data into interpretable tags, this method supports the provision of efficient and timely personalized learning feedback that offers constructive suggestions tailored to individual learner needs.*

Keywords: large language model, ChatGPT, educational data analysis, adaptive learning systems, study feedback

1. Introduction

The development of Large Language Models (LLMs) such as GPT-4 has significantly impacted various domains, especially education (Doughtly et al., 2024). These models are adept at understanding and generating natural language, thus offering significant capabilities in text generation, question-answering, and more (Liu et al., 2024). Particularly in education, LLMs are poised to enhance learning experiences and provide immediate, constructive feedback to students. However, effectively integrating the extensive and complex data typical of educational settings into these models presents notable challenges (Keuning et al., 2018).

Immediate feedback is a strong reinforcement mechanism, underscored by B.F. Skinner's operant conditioning theory, which asserts that immediate rewards or corrections can significantly influence future behavior by reinforcing desired learning activities (Skinner, 1963; Arifin & Humaedah, 2021). This is particularly vital in educational contexts where timely feedback can help students quickly recognize and correct mistakes, fostering a better understanding and retention of material (Hao et al., 2022). Moreover, feedback customization according to individual learning trajectories allows for personalized educational guidance, making learning more effective.

Nevertheless, while immediate feedback is invaluable, deploying LLMs like GPT-4 for generating comprehensive educational feedback presents significant challenges, particularly in data handling. Traditional LLM applications tend to focus on singular tasks such as grading individual questions or essays, providing feedback limited to specific prompts

without a broader contextual understanding. This narrow approach often misses the need for cumulative feedback that offers constructive insights over a learning period. To address this, we propose a tag annotation method that refines raw educational data into a structured format of predefined tags. These tags represent essential attributes of the student data comprehensively, enabling the model to process vast information more efficiently and generate feedback that is both integrative and strategically supportive of the student's educational journey. This method allows for the production of rich, context-aware feedback that supports holistic student development, overcoming the limitations of traditional single-point assessments.

Our research utilizes data from an adaptive learning system tested in a Shanghai primary school, captures detailed records of students ' performances across various parameters including correctness, difficulty levels, knowledge categories, ability levels, and task completion time. By transforming these data points into tags, we enable the LLM to generate nuanced feedback. This feedback is not only precise but also supportive, encouraging students to engage positively with their learning material. This approach of tag annotation simplifies the model's data processing tasks and enhances the relevance of its outputs, making LLMs more practical for real-world educational applications. It marks a significant advance in integrating sophisticated AI technologies into education, paving the way for scalable and impactful personalized learning feedback.

2. Related Work

2.1 Traditional Feedback Generation with LLMs in Education

The integration of Large Language Models like ChatGPT into education has highlighted the potential for automated feedback. Traditionally, such models provide immediate responses to individual tasks, such as answers to questions, essays, or coding projects (Hao et al., 2022). However, they have yet to fully address the challenge of analyzing large datasets for more comprehensive feedback. There is a growing need for methods that can efficiently encode broader educational data into these models. This would enable the generation of detailed, personalized feedback reports reflecting the student's learning journey over time.

2.2 Effective Reporting in Educational Feedback

Research indicates that the most effective feedback reports are those that provide holistic insights into a student's learning progress, addressing both strengths and areas for improvement. Kobus (2007) found that reports combining detailed analysis with positive reinforcement encourage students more effectively (Kobus et al., 2007). These studies underscore the importance of balanced feedback that motivates students while guiding them towards academic improvement. Periodic, comprehensive feedback reports are more beneficial as they allow for adjustments in teaching strategies and student learning approaches, aligning more closely with ongoing educational needs (Lyon et al, 2016). Additionally, the tone of the feedback is crucial; emphasizing constructive and supportive language helps maintain a positive learning environment (Kobus et al., 2007). Avoiding overly critical or negative remarks is essential to prevent discouraging students. Nurturing a student's confidence and interest in learning through supportive feedback significantly enhances learning outcomes (Henderson et al, 2005), which not only corrects misconceptions but also bolsters the student's abilities and self-esteem, fostering a conducive atmosphere for continuous learning and growth.

3. Experiment and Data Analysis

3.1 Experimental Procedure and Data Acquisition

The study collected experimental data from two primary school classes in Shanghai, utilizing a three-dimensional adaptive learning system developed by the Lab for Artificial Intelligence in Education at East China Normal University.

This system addresses knowledge, ability, and affective attitude, providing a variety of multiple-choice questions, each with two possible outcomes: correct (1) and incorrect (0). Students engaged in the system's online learning sessions during designated class periods on Monday and Wednesday afternoons each week.



Figure 1. Adaptive learning system interface display.

The experiment was carried out within a real online educational platform, specifically the aforementioned adaptive learning system, which included advanced data tracking modules. This setup not only ensured the accurate collection of comprehensive learning data but also strictly adhered to data protection measures to safeguard student privacy. Specifically, the final dataset provided by the system allowed for the careful extraction of data on students' performance in multiple-choice questions, including their accuracy rates. Additionally, we gathered data on the knowledge categories, ability levels associated with each question, and the time students took to complete these questions. Conducting the experiment during afternoon class periods ensured students were engaged and not distracted by other academic responsibilities, thus reflecting their true learning capabilities more accurately. This careful arrangement provided a solid foundation for subsequent analyses, aiming to evaluate the effectiveness of tag-based feedback and adaptive learning strategies employed.

3.2 Data Processing and Tag Annotation

Effective use of educational data in generating meaningful feedback through Large Language Models (LLMs) like GPT-4 requires thorough preprocessing and organization of raw data into a structured and interpretable format. The initial dataset included a vast array of knowledge categories and ability levels—more than fifty and thirty distinct types, respectively. To enhance model practicality, these were consolidated into manageable classes.

Knowledge categories were refined into six primary groups: Calculations of Speed and Time, Geometric Shapes and Properties, Data Statistics and Probability, Algebra and Functions, Arithmetic Operations and Properties. This categorization aimed to cover a broad spectrum of subjects while simplifying complexity. Ability levels were similarly reduced and organized into six groups: Practical Mathematical Application Skills, Data Organization and Statistical Skills, Computational Skills, Geometric Thinking Skills, Reasoning and Logical Thinking, and Innovative and Abstract Thinking. These groups reflected the core skills essential for academic success in structured learning environments.

A major challenge was managing multiple entries per student per question, often including brief, non-essential attempts likely from navigational actions rather than genuine problem-solving efforts. To resolve this, only the longest duration of attempt per question was retained, considered to represent the most substantial effort. Additionally, entries with zero duration were discarded, presumed to be guesses or inactive interactions, thus providing little educational value.

The streamlined data led to the development of a tag annotation system, organized into three primary categories: Performance Tags, Knowledge Domain Tags, and Ability Level Tags. Performance Tags (12 tags) were based on a combination of difficulty level (1-3), accuracy, and speed. Accuracy was split into 'adequate' (above 65%) and

'struggling' (below 55%). Speed was assessed by comparing each student's performance against their peers; top or bottom performers were tagged if the sample size exceeded 40, otherwise, the fastest or slowest 50% received the tag. Knowledge Domain Tags (10 tags) reflected proficiency or challenges in each of the five knowledge areas, with each capable of producing a positive or negative tag based on performance relative to benchmarks. Ability Level Tags (12 tags) assessed whether students excelled or struggled in each of the six domains.Specific tag settings are detailed in Table 1.

Table 1. Labels correspond to data features.

Tag Category	Tag	Description
Basic Analysis	Tag_1_1	Correctly and quickly on easy questions.
	Tag_1_2	Correctly and quickly on medium difficulty questions.
	Tag_1_3	Correctly and quickly on difficult questions.
	Tag_1_4	Correctly but completed slowly on easy questions.
	Tag_1_5	Ccorrectly but completed slowly on medium difficulty questions.
	Tag_1_6	Correctly but completed slowly on difficult questions.
	Tag_1_7	Incorrectly but completed quickly on easy questions.
	Tag_1_8	Answered incorrectly but quickly on medium difficulty questions.
	Tag_1_9	Incorrectly but completed quickly on difficult questions.
	Tag_1_10	Incorrectly but completed slowly on easy questions.
	Tag_1_11	Incorrectly but completed slowly on medium difficulty questions.
	Tag_1_12	Incorrectly but completed slowly on difficult questions.
Knowledge Category Analysis	Tag_2_1	Outstanding performance in calculations involving speed and time.
	Tag_2_2	Outstanding in the identification and properties of geometric shapes.
	Tag_2_3	Outstanding in data statistics and probability problems.
	Tag_2_4	Outstanding in algebraic equations and functions.
	Tag_2_5	Outstanding in arithmetic operations and properties.
	Tag_2_6	Struggling with calculations involving speed and time.
	Tag_2_7	Struggling to recognize and work with properties of geometric shapes.
	Tag_2_8	Finding data statistics and probability problems challenging.
	Tag_2_9	Struggling with algebraic equations and functions.
	Tag_2_10	Struggling with arithmetic operations and properties.
Ability Analysis	Tag_3_1	Strong capabilities in practical application of mathematics.
	Tag_3_2	Strong capabilities in statistical analysis.
	Tag_3_3	Strong computational skills.
	Tag_3_4	Strong geometric thinking skills.
	Tag_3_5	Strong logical reasoning skills.
	Tag_3_6	Strong innovative and abstract thinking skills.
	Tag_3_7	Challenged by the practical application of mathematics.
	Tag_3_8	Challenged by statistical analysis.
	Tag_3_9	Challenged by computational skills.
	Tag_3_10	Challenged by geometric thinking skills.
	Tag_3_11	Challenged by logical reasoning.
	Tag_3_12	Challenged by innovative and abstract thinking.

In total, 34 tags were designed to cover the potential spectrum of student performance scenarios comprehensively. This system ensured that each student could be associated with multiple tags, providing a detailed picture of their learning behaviors and outcomes, critical for generating personalized educational feedback.

3.3 Tag Parsing and Report Generation

The conclusion of our methodology involves transforming processed data into structured tags and combining these with the GPT-4 model to generate personalized educational feedback reports. After data refinement and tag annotation as outlined in the previous sections, we create a complete dataset named 'student_tag'. This dataset stores each student's 34 identified tags, where each tag is denoted as either 0 or 1, indicating the absence or presence of a specific learning characteristic or challenge.

The process of Prompt Design begins with the retrieval of student tags using the get_student_tags function from the 'student_tag' dataset. These tags provide customized inputs for generating reports, ensuring that each student's feedback is uniquely tailored to their individual performance data. The prompts designed for interaction with GPT-4 are carefully structured to include several distinct sections: an Overview, Basic Analysis, Knowledge Category Analysis, Ability Analysis, Learning Strategies and Recommendations, and a Summary. This structured approach aids the model in understanding the purpose of each section and in generating corresponding content. Each segment of the prompt includes specific instructions and guidelines that outline the type of content needed and its intended purpose, helping the model produce more accurate and relevant text outputs. The use of "you" in the prompts to directly address the student enhances the relatability and specificity of the reports, fostering greater engagement from the student.

The use of GPT-4 employs openai.ChatCompletion.create to utilize the model. Control over the randomness and length of the generated text is achieved by setting the Temperature to 0.4, which ensures that the generated text is consistent and minimally random, suitable for the needs of a professional report (Zong & Krishnamachari, 2023). The Max_tokens is set to 1000 to allow ample space for a thorough report. The Top_p is configured to 1, utilizing the full potential vocabulary to enrich the text's comprehensiveness and coverage. Controls on repetition and innovation, such as frequency_penalty and presence_penalty, are set to 0. This decision indicates a preference to maintain the natural flow and consistency of the content, which is deemed more crucial in this context (Savelka, 2023).

For practical illustration, a sample report for randomly selected student number 2965 is provided in the appendix. This example demonstrates the tailored feedback generated by the system.

Through these mechanisms, the system effectively synthesizes complex educational data into actionable insights, providing students with constructive feedback that is both encouraging and tailored to promote educational advancement and personal growth.

4. Evaluation

The potential of large language models (LLMs) like GPT-4 for educational feedback has been demonstrated through an evaluation with primary level mathematics teachers. The assessment, conducted via a questionnaire, examined five key aspects of the generated reports — comprehensibility, practicality, motivation, clarity, and organizational structure. Each aspect was rated on a scale from 0 to 10, with a maximum score of 50.

We received 63 questionnaires, discarding three for unjustifiably low scores and another 28 for non-genuine perfect scores, leaving 32 valid responses for analysis. Overall, the average scores exceeded 7 points in most dimensions, indicating that the reports were well-received. Teachers found them helpful for understanding and supporting the learning process, particularly in identifying student strengths and challenges.

The evaluation highlighted the need for simpler language and better data presentation, as clarity received the lowest average score of 6.59. A boxplot analysis revealed varying teacher perceptions across different aspects. Understanding Level had a median score of 7.5, indicating effective capture of students' strengths and challenges. Practicality scored

slightly lower, reflecting varied teacher opinions on the suggestions' applicability. The Motivation effect, with a median near 8, showed a positive impact, despite a few low outliers. Clarity and Organizational Structure had median scores of around 7 and above 7.5 respectively, suggesting the need for clearer reports and confirming the reports' logical structure facilitated information assimilation.

Table 2. Questionnaire items for the assessment of personalized feedback reports.

Category	Question Description
Understanding Level	Do you believe this report accurately analyzes students' main strengths and challenges in their learning?
Practicality	Are the recommendations proposed in the report practical and feasible? Do you think these suggestions are easily applicable in actual teaching practices?
Motivation effect	Do you feel that this report effectively motivates students to improve their learning methods and enhance their learning attitudes?
Clarity	Is the language used in the report clear? Did you encounter any parts that were difficult to understand while reading it?
Organizational structure	Do you find the structure of the report to be logical and conducive to quickly grasping key information? Are the various sections of the report well-organized and easy to follow?
Advise	Do you have any additional suggestions for this report?

.

In summary, teachers affirmed the reports' effectiveness in highlighting understanding, practicality, and organizational structure, with a recommendation to improve clarity to avoid misinterpretations. Such insights are crucial as they guide the refinement of algorithms that underpin report generation. This reaffirms the viability of using LLMs like GPT-4 to provide personalized educational feedback, enhancing the learning experience.



Figure 2. Boxplot of questionnaire results across five dimensions.

5. Discussion and Implications

This study has demonstrated the potential of the GPT-4 large language model in generating personalized educational feedback reports in the field of education. Most teachers gave positive reviews of the learning reports produced, particularly in terms of analyzing students' learning strengths, challenges, and providing specific suggestions. These feedback reports can serve as a powerful tool to help teachers understand and support students' learning processes more effectively, echoing the importance of immediate feedback in education.

The lower evaluations for report clarity highlight areas for improvement, suggesting that future report generation should focus more on the simplicity of language and the clarity of data presentation to ensure that both teachers and students can understand them better. Additionally, the presence of low outliers also indicates potential shortcomings in how the reports motivate students to improve their learning methods.

The significance of this study lies in providing a new perspective and method for using AI technology to support personalized educational feedback. Through the process of tag annotation and tag parsing, it is possible to transform vast amounts of complex learning behavior data into comprehensive and accurate input.

Future adjustments to the prompts can also be made to ensure the report outputs better meet the expectations and needs of math teachers in the classroom, effectively helping to reduce teacher workload and providing students with timely, personalized analysis of their learning status.

6. Conclusion

This study explored a method of using tag annotation and tag parsing to implement the large language model ChatGPT in generating personalized, real-time educational feedback reports, confirming its potential to play a key role in personalized education. By converting students' multidimensional learning data into structured tags, the model is capable of generating precise, personalized learning reports. These reports not only summarize students' performance but also provide specific strategies for improvement.

Feedback from teachers indicates that such reports are substantially helpful in understanding students' learning situations and providing targeted guidance. However, variations in scores for clarity and motivational impact also reveal areas for improvement, pointing out directions for future research and development.

In summary, this research provides a successful case study demonstrating the practicality and effectiveness of LLMs in the field of education, offering valuable insights and guidance for future AI research in similar areas. It not only reinforces the concept of data-driven educational feedback but also opens new avenues for educators and researchers to further explore the role of artificial intelligence in enhancing teaching quality and the learning experience.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 61977058, and the Natural Science Foundation of Shanghai under Grant 23ZR1418500.

References

Arifin, Z., & Humaedah, H. (2021). Application of Theory Operant Conditioning BF Skinner ' s in PAI Learning. Journal of Contemporary Islamic Education, 1(2), 101-110.

Doughty, J., Wan, Z., Bompelli, A., Qayum, J., Wang, T., Zhang, J., ... & Sakr, M. (2024, January). A comparative study of AI-generated (GPT-4) and human-crafted MCQs in programming education. In Proceedings of the 26th Australasian Computing Education Conference (pp. 114-123).

Hao, Q., Smith IV, D. H., Ding, L., Ko, A., Ottaway, C., Wilson, J., ... & Greer, T. (2022). Towards understanding the effective design of automated formative feedback for programming assignments. Computer Science Education, 32(1), 105-127.

Henderson, P., Ferguson-Smith, A. C., & Johnson, M. H. (2005). Developing essential professional skills: a framework for teaching and learning about feedback. BMC Medical Education, 5, 1-6.

Keuning, H., Jeuring, J., & Heeren, B. (2018). A systematic literature review of automated feedback generation for programming exercises. ACM Transactions on Computing Education (TOCE), 19(1), 1-43.

Kobus, T., Maxwell, L., & Provo, J. (2007). Increasing Motivation of Elementary and Middle School Students through Positive Reinforcement, Student Self-Assessment, and Creative Engagement. Online Submission.

Liu, Z., Ping, W., Roy, R., Xu, P., Shoeybi, M., & Catanzaro, B. (2024). ChatQA: Building GPT-4 Level Conversational QA Models. arXiv preprint arXiv:2401.10225.

Lyon, A. R., Lewis, C. C., Boyd, M. R., Hendrix, E., & Liu, F. (2016). Capabilities and characteristics of digital measurement feedback systems: Results from a comprehensive review. Administration and Policy in Mental Health and Mental Health Services Research, 43, 441-466.

Savelka, J., & Ashley, K. D. (2023). The unreasonable effectiveness of large language models in zero-shot semantic annotation of legal texts. Frontiers in Artificial Intelligence, 6, 1279794.

Skinner, B. F. (1963). Operant behavior. American psychologist, 18(8), 503.

Zong, M., & Krishnamachari, B. (2023, June). Solving math word problems concerning systems of equations with gpt-3. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 13, pp. 15972-15979).

Appendix

Personal Analysis Report of Student 2965

Overview:

Your overall learning performance exhibits a blend of strengths and areas that need improvement. You demonstrate a strong ability to answer both easy and difficult questions with speed and accuracy, and you excel in identifying and working with geometric shapes. However, there are areas such as calculations involving speed and time, solving algebraic equations, and understanding functions where you need to improve. You also find the practical application of mathematics and logical reasoning challenging.

Basic Analysis:

You have shown a consistent ability to answer both easy and difficult questions correctly and quickly. This indicates a strong grasp of the fundamental concepts and an ability to apply them effectively under time constraints. Your speed and accuracy are commendable and will serve you well in timed examinations and competitive scenarios.

Knowledge Category Analysis:

Your performance in geometry is noteworthy, with a strong ability to identify and work with geometric shapes. However, there are areas in your mathematical knowledge that need improvement. You need to work on your calculations involving speed and time, which are fundamental to physics and practical applications of mathematics. Additionally, your understanding of algebraic equations and functions needs to be strengthened. These are key areas in advanced mathematics and are crucial for problem-solving in many scientific disciplines.

Ability Analysis:

Your ability to apply mathematical concepts in practical scenarios and logical reasoning needs improvement. These are essential skills not only for mathematics but also for many other areas of learning and everyday life. Developing these skills will help you understand and solve complex problems more effectively.

Learning Strategies and Recommendations:

1. Practice: Regular practice of calculations involving speed and time, and solving algebraic equations will help you improve in these areas. Use online resources and textbooks to find practice problems.

2. Real-world Applications: Try to find real-world examples of mathematical concepts to understand their practical applications. This will also help improve your logical reasoning skills.

3. Seek Help: Don't hesitate to ask for help from your teachers or peers when you're struggling with a concept. Group study can also be beneficial in understanding different approaches to problem-solving.

Summary:

Remember, every learner has strengths and areas for growth. Your speed and accuracy in answering questions and your proficiency in geometry are your strengths to build upon. The areas that need improvement are not weaknesses, but opportunities for you to grow and learn. With consistent effort and the right strategies, you can overcome any challenge. Your potential is limitless, and every step you take towards improvement is a step towards success. Keep up the good work, and always strive for progress, not perfection.

IntelliChain: An Integrated Framework for Enhanced Socratic Method Dialogue with LLMs and Knowledge Graphs

Changyong Qi[1,*], Linzhao Jia[1], Yuang Wei[1], Yuan-Hao Jiang[1], Xiaoqing Gu[2]
[1] Lab for Artificial Intelligent for Education, East China Normal University
[2] Department of Educational Information Technology, East China Normal University
[*]changyongqi@stu.ecnu.edu.cn

Abstract: *With the continuous advancement of educational technology, the demand for Large Language Models (LLMs) as intelligent educational agents in providing personalized learning experiences is rapidly increasing. This study aims to explore how to optimize the design and collaboration of a multi-agent system tailored for Socratic teaching through the integration of LLMs and knowledge graphs in a chain-of-thought dialogue approach, thereby enhancing the accuracy and reliability of educational applications. By incorporating knowledge graphs, this research has bolstered the capability of LLMs to handle specific educational content, ensuring the accuracy and relevance of the information provided. Concurrently, we have focused on developing an effective multi-agent collaboration mechanism to facilitate efficient information exchange and chain dialogues among intelligent agents, significantly improving the quality of educational interaction and learning outcomes. In empirical research within the domain of mathematics education, this framework has demonstrated notable advantages in enhancing the accuracy and credibility of educational interactions. This study not only showcases the potential application of LLMs and knowledge graphs in mathematics teaching but also provides valuable insights and methodologies for the development of future AI-driven educational solutions.*

Keywords: LLMs, Knowledge Graphs, Multi-Agent Systems, Chain-of-Thought, Personalized Learning

1. Introduction

The advancement of artificial intelligence technology, particularly the development of Large Language Models (LLMs), is propelling educational innovation forward. These models demonstrate significant potential in complex interactions, natural language understanding, and generation, paving the way for new possibilities in creating intelligent educational agents(McShane & Leon, 2022). In the realm of education, LLMs support more personalized and interactive learning experiences, which are crucial for meeting diverse learning needs and enhancing the quality of education (Kumar et al., 2024). However, despite LLMs' remarkable capabilities in enhancing the intelligence of educational applications, they still face challenges regarding accuracy, reliability, and relevance to specific educational content (Kaddour et al., 2023).

This study is dedicated to optimizing the information exchange and collaboration mechanisms among intelligent agents through the chain-of-thought dialogue approach, combining LLMs and knowledge graphs, to enhance teaching effectiveness and the learning experience. The chain-of-thought dialogue, an interaction mode that simulates the human thought process, can promote a deeper understanding and response among intelligent agents, offering learners richer and more targeted learning content (Chae et al., 2023). However, achieving this goal requires addressing the inherent shortcomings of relying on LLMs, including improving the accuracy of interactions and ensuring the credibility of educational content (Chu et al., 2023).

To overcome these challenges, this research proposes an innovative solution: integrating knowledge graphs to enhance the credibility of LLMs' application in educational settings, while optimizing the design and collaboration of the multi-agent system. As an effective tool for organizing and representing information, knowledge graphs provide rich, structured background knowledge for LLMs, aiding intelligent agents in more accurately understanding and responding to educational content (Kommineni et al., 2024). Furthermore, by optimizing the design of the multi-agent system, this

study aims to achieve more effective collaboration and information sharing among intelligent agents, thereby enhancing the educational outcomes based on the chain-of-thought dialogue approach. The goal of this research is to develop a reliable and effective educational multi-agent system that can leverage LLMs and knowledge graphs to provide personalized learning experiences and promote efficient collaboration among intelligent agents through the chain-of-thought dialogue method. In this way, we aim to offer a new and effective AI application mode in the field of educational technology, especially in supporting complex teaching tasks and facilitating students' learning processes.

2. Related work

Amidst the rapid advancements in artificial intelligence, LLMs have unveiled unprecedented potential in the field of education. Generative models have been demonstrated to effectively support adaptive learning systems, automated content generation, and intelligent assessment mechanisms (Bahroun et al., 2023). Their capacity to understand and generate complex natural language offers robust support for personalized educational pathways. However, the academic community has extensively discussed the limitations and accuracy challenges of LLMs when dealing with domain-specific knowledge (Pezeshkpour, 2023). Despite their revolutionary capabilities in language processing, how to effectively leverage these capabilities to enhance learning efficiency and outcomes remains an open question in educational applications.

2.1 Chain-of-Thought Dialogues in Educational Technology

Chain-of-thought dialogue, as a method simulating human thought processes, has gained considerable attention in educational technology research in recent years. By facilitating deeper and more coherent exchanges, this approach aims to promote critical thinking and deep understanding among learners. Research has highlighted that the implementation of chain-of-thought dialogue significantly enhances learning motivation and engagement in online learning environments (Wang et al., 2023). The integration of LLMs in chain-of-thought dialogue represents a significant leap forward in educational technology research. These models, when applied to simulate and generate chain-of-thought dialogues, offer the potential to facilitate more nuanced and deep learning interactions (Yin et al., 2023). Recent studies have explored how LLMs like GPT-3.5 can be fine-tuned to produce sequential reasoning steps in problem-solving tasks, effectively mimicking the human thought process(Koralus & Wang-Maścianica, 2023). This application not only enhances the dialogue's relevance and depth but also promotes a more engaging and interactive learning experience. However, despite these advancements, challenges persist in ensuring the dialogues' consistency with factual accuracy and pedagogical objectives, indicating a fertile area for future research.

2.2 Leveraging Knowledge Graphs and Multi-Agent Systems to Transform Educational Interactions

Knowledge graphs, as a structured representation of knowledge, play a critical role in enhancing the knowledge base and reasoning capabilities of LLMs. By integrating rich domain knowledge in the form of graphs, LLMs can achieve more accurate contextual understanding and reasoning within educational settings, significantly improving the quality and accuracy of responses (Meyer et al., 2023). Additionally, knowledge graphs provide detailed entity relationships and attribute information, effectively aiding LLMs in avoiding the generation of hallucinations—content that appears plausible but is factually incorrect (Agrawal et al., 2024). This capability is especially important for educational applications, as it directly relates to the reliability of learning content and the quality of education. Simultaneously, the development of multi-agent systems based on LLMs represents a new zenith in intelligent teaching and interactive learning within the field of educational technology. Agents within these systems leverage the computational power and advanced language processing abilities of LLMs to facilitate efficient communication and collaboration amongst themselves. More importantly, they can continuously learn from interactions with learners, thereby dynamically adapting to learners' needs and providing personalized learning support. Furthermore, by

simulating complex interactions of the real world, multi-agent systems offer learners a more rich and immersive learning experience. The flexibility and dynamism of these systems are unparalleled by traditional educational methods.

Integrating knowledge graphs and multi-agent systems based on LLMs, especially in supporting chain-of-thought dialogues based on LLMs, not only enhances the accuracy and relevance of educational content but also greatly increases the personalization and interactivity of educational interactions. This transforms the educational process from a one-way transmission of knowledge into a dynamic, bidirectional, and interactive learning experience.

3. Framework

In this study, we introduce IntelliChain, a comprehensive educational support architecture that synergistically integrates LLMs, knowledge graphs, and a multi-agent system to facilitate an efficient chain-of-thought dialogue mechanism, as shown in *Figure 1*. Specifically, IntelliChain is built upon modular principles with a focus on optimizing the collaborative efficiency of intelligent agents in three main areas: the strategy design of chain-of-thought dialogue, the construction and integration of knowledge graphs, and the optimization of the multi-agent system. This framework not only aims to enhance the accuracy and relevance of educational content but also leverages the unique strengths of each component to support nuanced educational dialogues, thereby setting a new standard for interactive learning environments.

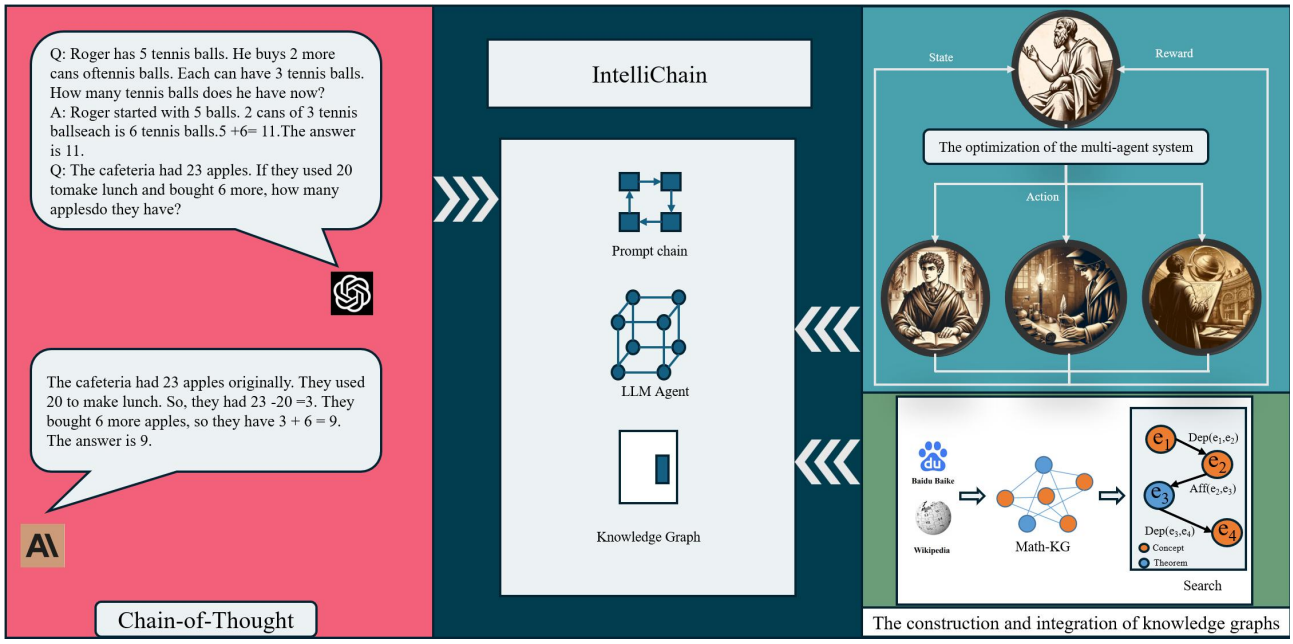


Figure 1. The Framework of IntelliChain.

Within the IntelliChain framework, an innovative dialogue strategy is articulated, leveraging a sophisticated chain-of-thought approach to enhance the pedagogical efficacy of educational dialogues. This methodology is underpinned by an integration of advanced pedagogical principles, encompassing strategies such as guided questioning, sequential analytical reasoning, iterative feedback mechanisms, and the facilitation of exploratory inquiry. The design premise of IntelliChain advocates for a structured dialogue progression, where learners are methodically navigated through complex cognitive tasks, thereby augmenting their analytical acumen and fostering critical thinking capacities. The framework introduces a novel role-based interaction schema, wherein agents designated as "instructors" and "learners" partake in a reciprocal exchange of investigative queries and insights. This orchestrated educational dialogue not only promotes active learner engagement through sustained cognitive involvement but also ensures pedagogical alignment via continuous feedback loops.

The construction and integration of the knowledge graph represent a cornerstone in facilitating enriched educational dialogues, specifically tailored to the domain of mathematics education. This process is undergirded by a meticulous amalgamation of domain expertise and high-caliber educational resources, culminating in the development of a knowledge graph that meticulously catalogues key mathematical concepts, principles, and illustrative examples. The resultant knowledge base serves as a robust and comprehensive repository, furnishing LLMs with precise and extensive domain-specific information. To achieve an efficacious integration of the knowledge graph with LLMs, this study introduces an advanced querying mechanism designed to harness the knowledge graph's potential fully. Prior to each dialogue iteration, the system autonomously conducts a query within the knowledge graph based on the specific knowledge points implicated in the dialogue content. This procedure extracts pertinent information to serve as prompts for dialogue generation, thereby ensuring the educational content's relevance and accuracy.

The optimization of the multi-agent system represents a strategic deployment of intelligent agents, precisely tailored to the specific demands of educational tasks. This strategy is designed to fully leverage the unique expertise of each agent, thereby enhancing the quality of instructional content and the efficiency of the learning process. At the heart of system optimization lies the adoption of advanced reinforcement learning algorithms, establishing a learning and adaptation mechanism. This mechanism enables intelligent agents to iteratively adjust their behaviors and strategies based on feedback received from interactions with learners. This adaptive process allows the system to self-modulate based on the correlation between the agents' actions and learners' responses, ensuring that educational interventions are optimally aligned with the learners' evolving needs and preferences.

4. Results

In the present study, the IntelliChain framework was utilized to examine the differential outputs of teacher agents employing the Socratic method in solving the classical chicken-rabbit problem under three distinct system configurations: without agents, with agents but without knowledge graph integration, and with agents integrated with a knowledge graph, as shown in *Figure 2*. This comparative analysis aimed to elucidate the impact of knowledge graph integration and multi-agent system optimization on the quality of pedagogical dialogue and teaching efficacy.

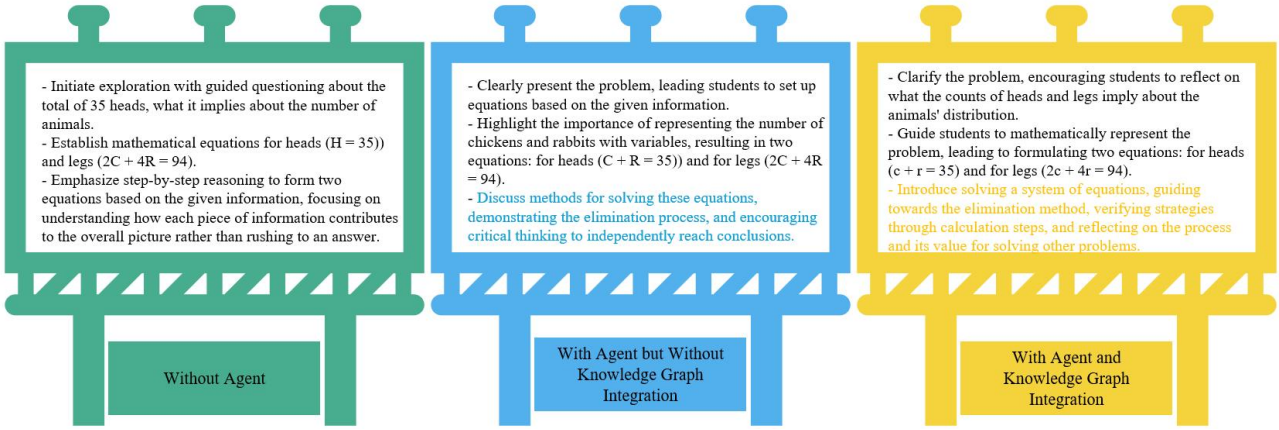


Figure 2. Differential Teaching Outcomes Using the IntelliChain Framework Across Three System Configurations.

In the present study, the IntelliChain framework was utilized to examine the differential outputs of teacher agents employing the Socratic method in solving the classical chicken-rabbit problem under three distinct system configurations: without agents, with agents but without knowledge graph integration, and with agents integrated with a knowledge graph, as shown in Figure 2. This comparative analysis aimed to elucidate the impact of knowledge graph integration and multi-agent system optimization on the quality of pedagogical dialogue and teaching efficacy.

Without Agent Configuration: In the absence of agent intervention, the teaching dialogue primarily relied on open-ended questioning and step-by-step reasoning to stimulate learner exploration. While this approach facilitated

learner engagement, it lacked the specificity and efficiency that could be achieved through the utilization of specific algorithms or educational resources, rendering the teaching process somewhat generic and less targeted. With Agents but Without Knowledge Graph: The configuration of employing teaching agents without knowledge graph integration showed a marked improvement in teaching dialogue by adhering to the Socratic method, characterized by structured problem presentation and algebraic problem-solving guidance. Despite the potential of agents in enhancing the learning process, the absence of knowledge graph support limited the depth and breadth of instructional content, underutilizing the potential benefits of agent involvement in teaching. With Agents Integrated with Knowledge Graph: The integration of teaching agents with a knowledge graph significantly enhanced the quality of teaching dialogue. This configuration not only enabled agents to guide the problem-solving process through algebraic methods effectively but also deepened the discussion on problem context and related concepts utilizing the rich information from the knowledge graph. Such an integrated approach not only deepened learners' understanding of the problem but also promoted personalized learning and active learner participation by dynamically adjusting teaching strategies and methods.

5. Discussion

The IntelliChain framework constitutes a pivotal technological advancement in the realm of educational technology, particularly in augmenting Socratic method through the integration of chain-of-thought dialogues, knowledge graphs, and an optimized multi-agent system. This comprehensive strategy not only facilitates a deeper and structured inquiry, akin to the Socratic method, but also pioneers a personalized educational pathway via adaptive learning strategies. Despite the framework's promising capabilities, challenges such as its applicability across diverse learning scenarios and the imperative for maintaining up-to-date, unbiased knowledge graphs warrant further investigation. Future research directions might include refining the framework's dialogue strategies to accommodate various learning styles and broadening the scope of knowledge graphs to cover a wider spectrum of disciplines. Additionally, the potential integration of emergent technologies presents an exciting frontier for creating more immersive and interactive learning environments, underscoring IntelliChain's transformative potential in educational methodologies.

6. Conclusion

This study introduced the IntelliChain framework, an innovative approach that enhances educational dialogues through the integration of LLMs, knowledge graphs, and a multi-agent system, optimized for the Socratic method. It demonstrated the framework's capability to improve the precision and relevance of educational content while facilitating personalized learning experiences. Comparative analysis across different configurations underscored the significance of knowledge graph integration and multi-agent system optimization in augmenting teaching efficacy. Despite promising outcomes, challenges such as adaptability across diverse learning scenarios and maintaining unbiased knowledge graphs remain. Future efforts will focus on refining dialogue strategies and exploring emerging technologies to further enhance the learning environment. IntelliChain represents a significant stride towards advancing AI-driven personalized education, promising to elevate the quality and effectiveness of teaching and learning methodologies.

References

McShane, M., & Leon, I. (2022). Language Generation for Broad-Coverage, Explainable Cognitive Systems (arXiv:2201.10422). arXiv. https://doi.org/10.48550/arXiv.2201.10422

Kumar, H., Musabirov, I., Reza, M., Shi, J., Wang, X., Williams, J. J., Kuzminykh, A., & Liut, M. (2024). Impact of Guidance and Interaction Strategies for LLM Use on Learner Performance and Perception (arXiv:2310.13712). arXiv. https://doi.org/10.48550/arXiv.2310.13712

Kaddour, J., Harris, J., Mozes, M., Bradley, H., Raileanu, R., & McHardy, R. (2023). Challenges and Applications of Large Language Models (arXiv:2307.10169). arXiv. https://doi.org/10.48550/arXiv.2307.10169

Chae, H., Song, Y., Ong, K. T., Kwon, T., Kim, M., Yu, Y., Lee, D., Kang, D., & Yeo, J. (2023). Dialogue Chain-of-Thought Distillation for Commonsense-aware Conversational Agents (arXiv:2310.09343). arXiv. https://doi.org/10.48550/arXiv.2310.09343

Chu, Z., Chen, J., Chen, Q., Yu, W., He, T., Wang, H., Peng, W., Liu, M., Qin, B., & Liu, T. (2023). A Survey of Chain of Thought Reasoning: Advances, Frontiers and Future (arXiv:2309.15402). arXiv. https://doi.org/10.48550/arXiv.2309.15402

Kommineni, V. K., König-Ries, B., & Samuel, S. (2024). From human experts to machines: An LLM supported approach to ontology and knowledge graph construction (arXiv:2403.08345). arXiv. https://doi.org/10.48550/arXiv.2403.08345

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. Sustainability, 15(17), 12983. https://doi.org/10.3390/su151712983

Pezeshkpour, P. (2023). Measuring and Modifying Factual Knowledge in Large Language Models (arXiv:2306.06264). arXiv. https://doi.org/10.48550/arXiv.2306.06264

Wang, H., Wang, R., Mi, F., Deng, Y., Wang, Z., Liang, B., Xu, R., & Wong, K.-F. (2023, December 1). Cue-CoT: Chain-of-thought Prompting for Responding to In-depth Dialogue Questions with LLMs. The 2023 Conference on Empirical Methods in Natural Language Processing. https://openreview.net/forum?id=FRRlmKxuf2

Yin, Z., Sun, Q., Chang, C., Guo, Q., Dai, J., Huang, X., & Qiu, X. (2023). Exchange-of-Thought: Enhancing Large Language Model Capabilities through Cross-Model Communication. In H. Bouamor, J. Pino, & K. Bali (Eds.), Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 15135–15153). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.936

Koralus, P., & Wang-Maścianica, V. (2023). Humans in Humans Out: On GPT Converging Toward Common Sense in both Success and Failure (arXiv:2303.17276). arXiv. https://doi.org/10.48550/arXiv.2303.17276

Meyer, L.-P., Stadler, C., Frey, J., Radtke, N., Junghanns, K., Meissner, R., Dziwis, G., Bulert, K., & Martin, M. (2023). LLM-assisted Knowledge Graph Engineering: Experiments with ChatGPT (arXiv:2307.06917). arXiv. https://doi.org/10.48550/arXiv.2307.06917

Agrawal, G., Kumarage, T., Alghamdi, Z., & Liu, H. (2024). Can Knowledge Graphs Reduce Hallucinations in LLMs? : A Survey (arXiv:2311.07914). arXiv. https://doi.org/10.48550/arXiv.2311.07914

中国古典诗词人工智能生成图像创作与评价：提示词生成与优化

On the Creation and Evaluation of AI-generated Artworks of Classical Chinese Poetry: Prompt Generation and Optimization

陈弘正[1,*]，许怀之[2]，李威[1]，萧巨昇[3]
[1] 黄冈师范学院 教育学院
[2] 黄冈师范学院 文学院（苏东坡书院）
[3] 黄冈师范学院 美术学院
[*] hungcheng@hgnu.edu.cn

【摘要】 本研究运用人工智能技术生成可体现中国古典诗词精髓的艺术创作。以苏轼《念奴娇·赤壁怀古》为例，研究广泛汇总相关文本资料，借助大语言模型 ChatGPT 分析并生成诗意描述作为 Midjourney 的提示词。通过经典译文及 ChatGPT 对话方式，分别生成多幅具有中国水墨画韵味的 AI 绘画作品。研究导入 ChatGPT 的图像识别能力，从江水表现、浪的渲染、历史氛围、情感氛围及艺术风格等方面对作品进行评价，并与文学领域及绘画领域专家的评价结果进行对比分析。最后，研究归纳了运用 ChatGPT 来评价古典诗词生成画作的特点，并总结了运用提示工程技术引导大语言模型分析和生成知识的策略，为在艺术创作等复杂领域应用之生成式人工智能提供了新思路。

【关键词】 人工智能艺术创作；中国古典诗词；大语言模型；提示词工程

Abstract: *This study explores the potential of generative AI technology in creating artistic representations that capture the essence of classical Chinese poetry. We utilized Su Shi's renowned poem, "Charm of a Maiden Singer: Thinking of Ancient in Red Cliff," as a case study and conducted extensive textual analysis, drawing from diverse scholarly interpretations and historical contexts. Subsequently, we employed the large language model ChatGPT to generate evocative descriptions, serving as prompts for the AI image generator Midjourney. Through established translations and interactive dialogues with ChatGPT, we produced a series of AI-generated artworks reminiscent of traditional Chinese ink-wash paintings. To evaluate these creations, we harnessed ChatGPT's image recognition capabilities, focusing on critical aspects, including the portrayal of the river's dynamism, the rendering of waves, the evocation of historical and emotional atmospheres, and the adherence to artistic styles. Comparative analyses were then undertaken, juxtaposing these AI-driven evaluations with those provided by experts in literature and painting. Our findings elucidate the unique characteristics and potential of utilizing ChatGPT to assess artworks generated from classical poetry. Additionally, we identify effective strategies for employing prompt engineering techniques to guide large language models in the nuanced tasks of knowledge analysis and generation. This research contributes novel insights into the application of generative AI within the intricate realm of artistic creation of classical Chinese poetry.*

Keywords: AI artistic creation; classical chinese poetry; large language model; prompt engineering

1. AI 赋能古典诗词，探索艺术创作新路径

近年来，AI 技术在艺术领域的应用引发广泛关注，尤其是在中国古典诗词与中国古画领域的交叉应用方面，研究人员探索运用 AI 图像生成技术创作出既具诗意又能展现古画韵味的作品 (Chen & Chen, 2023)。然而，AI 生成图像是否能超越人类创作仍存在争议 (Rasyad et al., 2023)。现有研究尝试从不同角度切入，如基于古画自动创作诗词 (Chen et al., 2021)，模拟传统水墨画风格 (Li & Li, 2006)，以及根据诗词生成带有书法的诗意水墨画 (Yuan et al., 2022)

等，这些研究为本研究提供了宝贵启示。此外，运用 ChatGPT 生成提示词已成为 AI 艺术创作的新趋势 (Chen & Chen, 2023)。与人工撰写提示词的传统方式相比，ChatGPT 能通过对话和提示工程技术，生成更精准、更具诗意的描述，进而引导 AI 图像生成平台 (如 Midjourney) 创作高质量作品 (ChatGPT 4 & Sabit, 2023)。

本研究以创新的思维将生成式 AI 技术应用于中国古典诗词的艺术创作，并构建了多维度的评价体系。首先，利用 ChatGPT 强大的文本分析能力，以大量诗词研究及赏析的文本为基底，将苏轼的《念奴娇・赤壁怀古》转化为更具深度和诗意的描述，以此作为 AI 图像生成的精准提示词。其次，从江水表现、浪的渲染、历史氛围、情感氛围及艺术风格等方面，对 AI 生成的图像作品进行评价，并与文学和绘画领域专家的评价进行对比分析，展现了艺术评价的多元性和主观性。此外，本研究巧妙地运用知识提示、链式思考提示和方向性刺激提示等多种提示工程技术，引导 ChatGPT 进行深度思考和分析，从而生成更全面、更贴合需求的知识和评价结果。这些创新不仅拓展了生成式 AI 在艺术创作领域的应用潜力，也为提示工程技术在引导大语言模型进行知识分析和生成方面提供了新的思路和策略，为 AI 技术在更复杂领域的应用开辟了新的道路。

2. 文本基础构建：深度解读苏轼《念奴娇・赤壁怀古》

本研究选择苏轼的《念奴娇・赤壁怀古》作为案例，探究 AI 如何将古典诗词转化为图像。这首词在中国文学史上的地位举足轻重，其文学艺术价值和深刻的历史、哲思内涵，使其成为历代学者研究和赏析的焦点。

《念奴娇・赤壁怀古》不仅频繁入选历代诗词选本，更引发了当代学界的广泛讨论和研究 (王兆鹏、郁玉英和郭红欣，2012)。据统计，截至 2012 年，已有 186 篇学术文章专门探讨这首词，且中国知网数据库中的相关论文数量迄今已达 578 篇，涵盖了文学、历史、文化、教育、艺术等多个领域，反映了学界对其持续关注和深入研究的热情。

为了使 AI 模型能够更全面地理解这首词的意境和内涵，本研究广泛收集整理了相关的研究文献成果。这些资料包括：

（1）历代文人重要的注释和评笺：如刘乃昌等 (1988) 编著的《唐宋词鉴赏辞典》，以及徐中玉和金启华 (1999) 主编的《中国古代文学作品选》；

（2）各种视角的赏析文本：如蘅塘退士等 (2009) 编选的《唐诗三百首・宋词三百首・元曲三百首》，以及刘滋培 (1981)、彭秀模和吴广平 (1984)、陈国华 (2015)、盛大林 (2023) 等学者的研究论文；以及

（3）教育领域的相关论文：如汪婉青和周庆华 (2000)、朱文成 (2006)、叶丽清 (2020)、孙笛和李振峰 (2023) 等探讨《念奴娇・赤壁怀古》在语文教学中的应用价值的论文。

通过收集整理以上资料，本研究初步汇集了近一万五千字的文本，构建了丰富的文本基础，为后续利用 ChatGPT 进行提示词的分析与生成提供了可靠的依据。

上述文本基础构建步骤的重要性在于：（1）确保提示词的深度和准确性：丰富的文本基础确保了 ChatGPT 能够更全面地理解词义、意境和历史文化背景，从而生成更精准、更具诗意的提示词；（2）体现研究的严谨性：广泛参考权威的研究成果，保证了研究结果的可靠性和学术价值；以及（3）为后续评价提供参考：这些文本资料也为后续专家对 AI 生成图像的评价提供了重要的参考依据。

3. AI 图像创作与提示词设计策略

本研究分别以（1）经典英译及（2）大语言模型文本探究生成来决定 Midjourney 提示词的主体，并辅以不同的绘画风格来生成图像。其中，经典英译的文本采用许渊冲教授的译本

（许渊冲，2021），而大语言模型文本探究则导入本研究汇整的赏析文档，透过对话来生成适合的提示词。此外，限于篇幅，本研究将专注探索「大江东去，浪淘尽，千古风流人物」此句的图像面貌。

3.1. 经典英译提示词

图 1 (a)–(d) 为根据许渊冲教授英译本生成的图像，译文为：

The endless river eastward flow, with its huge waves are gone all those Gallant heroes of bygone years.

其中，图 1 (a) 融合了中国水墨画与林风眠的风格，图 1 (b) 则融合了中国水墨画与日本浮世绘的风格，图 1 (c)与图 1(d) 则融合了中国水墨画与林风眠的风格之外，加入 Midjourney 的风格化参数（--stylize）并重覆生成的结果。其中，--stylize 的默认值为 100，并可接受介于 0 至 1000 的整数值。根据 Midjourney 的官方文件的说明，低风格化值生成的图像与提示词非常接近，但艺术性较低；高风格化值生成的图像则偏艺术化，但与提示词的联系较少。本研究依主观偏好选取风格化值为 150。

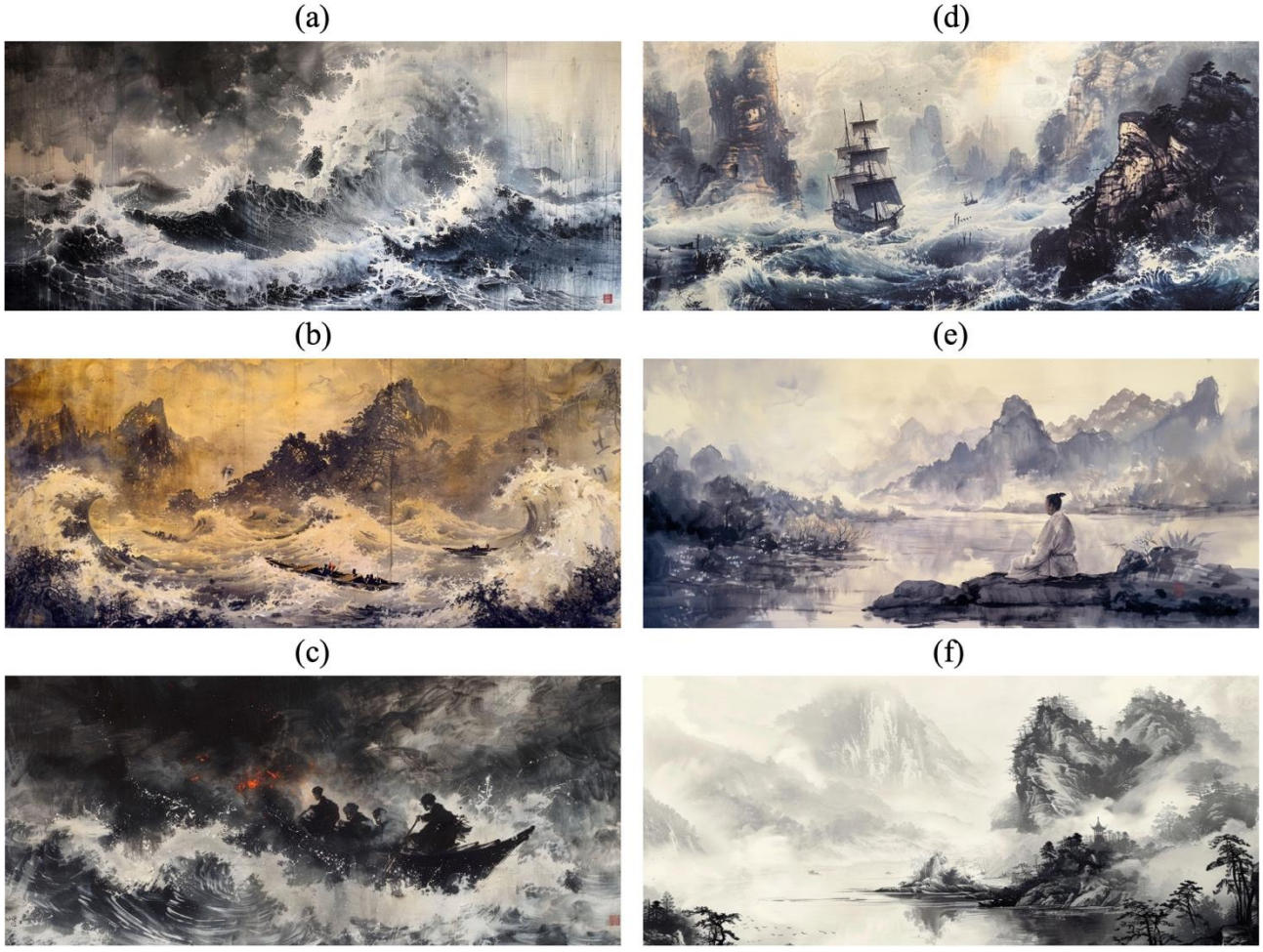


图 1 Midjourney 生成图像。(a)画作一，(b)画作二，(b)画作三，(b)画作四，(b)画作五，及(b)画作六。

3.2. 大语言模型文本探究生成提示词

首先在 ChatGPT 中导入本研究汇整的赏析文档，并提出下列几个问题引导 ChatGPT 回答，藉此梳理文本中重要的思想，包括（1）读这首《念奴娇．赤壁怀古》，你觉得苏轼怀念了哪些往事呢？（2）“大江东去，浪淘尽、千古风流人物。”风流人物是指哪些人？（3）读“浪

淘”的注释，从以往白居易的“浪淘”到眼前“浪淘尽、千古风流人物”有什么变化？（4）你觉得东坡的“神游”是到了什么地方？

ChatGPT 对上述问题的回答中有不少内容颇为深入，提供了本研究撰写提示词的重要指引，举例：

白居易的“浪淘”源自其《浪淘沙》词，比喻世事如沙粒，终被时间浪潮淘尽，体现世间无常与生命短暂。苏轼在“大江东去，浪淘尽、千古风流人物”中，继承并扩展了这一思想，将“浪淘”寓意历史英雄人物的不朽记忆与无情消逝，深化了对物质世界不稳定和生命逝去的认识。他的这种反思不仅是对历史英雄的哀悼，更透露出对生命无常和历史循环的超越性哲学思考，显现其作为文学家和哲学家的深邃境界。

在上述的背景下，进一步导引 ChatGPT 想像苏东坡在写下“大江东去，浪淘尽、千古风流人物”词句时心中的景像，並以中国水墨画及文人画的风格来呈现。图 1(e) 为 Midjourney 生成的图像，提示词为：

In a serene setting, Su Dongpo sits contemplatively by a vast river, embodying the eternal flow of time. This traditional Chinese ink and wash scene captures him composing his verse about the fleeting nature of human achievements. Misty mountains in the background and delicate brushstrokes emphasize the spiritual essence of literati art, focusing on the poet's reflective pose and the timeless landscape around him, all rendered in nuanced shades of black ink.

这段提示词描述了一个宁静的场景：文人苏东坡静静地坐在宽阔的河边，体现了时间的永恒流转。背景中朦胧的山脉和细腻的笔触强调了文人画的精神本质，并以黑色水墨的细微色调巧妙呈现；既展现了苏东坡的个人风采，也体现了中国水墨画的艺术魅力。

除了上述的探究方式，本研究进一步引导 ChatGPT 参考文档中各种赏析的观点，以较宏观的视野来勾勒这一词句的图像。图 1(f)为依照此思路下生成的图像，提示词为：

Visualize a scene inspired by Su Dongpo's verse, in traditional Chinese ink and wash painting style. A vast river flows, symbolizing time's unending passage, with misty mountains adding depth and eternity. Rendered in nuanced black ink, this tranquil scene contemplates the transient nature of human achievements against eternal nature's backdrop.

比较上述两段提示词发现，第一段更多着眼在苏东坡个人的侧写，第二段则更侧重藉外在景物的描绘抒发人生哲理的感悟。

4. AI 图像评价与提示词设计策略

4.1. 大语言模型探究 AI 生成图像与评价

为了进一步探究生成图像的细节以及与原词句之间的对映，首先以经典英译为提示词的生成图像（图 1(a)–(d)）为标的，請 ChatGPT 依序探究四幅图像：（1）分析画作的特点，（2）評价这四幅画作与“大江东去，浪淘尽、千古风流人物”词句的贴切程度，以及（3）由画作特点建议可用来评价图像与词句贴近程度的面向。

依据上述的探究方式，ChatGPT 由生成图像与词句的比较汇整出五个面向及其评价依据，依序如下：（1）江水的表现：词句中的“大江东去”描绘了长江的宏大流动，画作中江水是否广阔、有力，流向是否明确，是评价的重点；（2）浪的渲染：“浪淘尽”体现了浪的力量和洗刷一切的意象，画中浪的动态是否强烈，是否能体现出时间的流逝和历史的洗礼；（3）历史氛围的营造：词句中回顾了“千古风流人物”，画作是否能唤起对过去英雄豪杰的联想，或是通过一些象征性元素来表现历史的深远影响；（4）情感氛围与哲思：蘇東坡的词具有深沉的情感和哲理思考，画作是否能传达出类似的情感深度和引人深思的氛围；以及（5）艺术

风格与技巧：考虑到水墨画和文人画的风格特点，画作是否展现了传统水墨的韵味以及文人画的意境。

表 1 为依据上述五个面向及其评价原则，由 ChatGPT 来评价六幅画作的结果。其中，各个面向的分数介于 1 至 10 分之间，代表该画作与每个面向的贴切程度。

表 1 ChatGPT 评价 AI 图像

	江水的表现	浪的渲染	历史氛围的营造	情感氛围与哲思	艺术风格与技巧	总分
画作一	9	10	7	9	10	45
画作二	8	7	9	8	8	40
画作三	9	9	7	7	9	41
画作四	7	8	8	6	8	37
画作五	6	5	6	10	7	34
画作六	8	6	10	9	7	40

由评价结果来看，画作一获得最高的评价，画作三次之，画作二及画作六并列第三，画作四、画作五依序为第五及第六名，针對各幅画的综合评价依名次简述如下。

画作一：在五个面向上都有很高的得分，尤其是在浪的渲染和艺术风格与技巧上，这幅画的强烈动感和大气磅礴的风格，与苏东坡词中的江水和历史英雄的壮阔气魄非常契合，是所有作品中最能体现词意的画作。

画作三：在江水的表现和浪的渲染上得分较高，艺术风格也表现出色，展示了大江东去的气势和浪淘尽的视觉冲击力，但在历史氛围的营造上稍显不足。

画作二与画作六：这两幅画作的总分相同，都非常好地营造了历史氛围，第二幅画在情感表达上稍胜一筹，而第六幅画则在历史氛围的营造上更为出色，两幅作品都有各自的特点和魅力。

画作四：在所有的面向中得分均衡，表现了与自然力量相抗衡的主题，有动静结合的特点，但在情感氛围和哲思上略显不足。

画作五：在情感氛围与哲思上得到最高分，展现了词中的静谧与内省，虽然与诗句中的江水气势不符，但静谧的情感氛围与苏东坡的哲思相呼应，呈现了文人画的静谧感。

最后，进一步由 ChatGPT 比较上述结果与本研究汇整的赏析文档的一致性可知，在文档分析中提及的苏东坡词中的江水壮阔、浪花汹涌，以及其所蕴含的哲理，可能最能被第一幅画所体现，因此从文档分析来看，第一幅画应该是最能代表词句精神的画作。相比之下，其他画作虽然在不同程度上反映了词句中的某些元素，但可能在表达赤壁之战的历史氛围和哲学深度上不如第一幅画作深刻。

4.2. 文学领域专家探究 AI 生成图像与评价

表 2 为依据上述五个面向及其评价原则，由某师范学院文学院教师以文学赏析的个人视角来评价六幅画作的结果。

由评价结果来看，画作五获得最高的评价，依序为画作六、画作四、画作二、画作三与画作一，该专家針對五个评价面向的综合评价如下。

江水的表现：第五和第六幅画透过山崖倒影凸显江水的静谧辽阔，尤其眼前是江湾后的景色，更有长江渊远流长的含意。

浪的渲染：第四幅画以高崖，大船进行衬托，最能显现江面宽阔，江浪奔流的宏大气象。

历史氛围的营造：第五幅画将人景结合，突出史以人记的意象。第六幅画是江水自云雾中的远山而来，更见江流千古，时空变换之妙。

表 2 文学领域专家评价 AI 图像

	江水的表现	浪的渲染	历史氛围的营造	情感氛围与哲思	艺术风格与技巧	总分
画作一	6	8	6	5	7	32
画作二	6	9	6	7	7	35
画作三	7	8	7	6	6	34
画作四	7	9	6	7	7	36
画作五	8	7	8	8	8	39
画作六	8	6	8	7	8	37

情感氛围与哲思：第五幅画所绘之江畔端坐的诗人，仿佛读者以第三视角观赏苏轼当年的写作，对词作的情感更能身历其境。

艺术风格与技巧：第五与第六幅画皆是用静景烘托动感，刺激读者思考，以更深入词意。

4.3. 绘画领域专家探究 AI 生成图像与评价

表 3 为依据上述五个面向及其评价原则，由某师范学院美术学院国画专业教师以绘画赏析的个人视角来评价六幅画作的结果。

表 3 绘画领域专家评价 AI 图像

	江水的表现	浪的渲染	历史氛围的营造	情感氛围与哲思	艺术风格与技巧	总分
画作一	5	8	5	5	6	29
画作二	5	8	5	6	7	31
画作三	5	5	5	4	5	24
画作四	5	8	7	6	7	24
画作五	5	5	7	7	6	30
画作六	5	5	8	8	8	34

由评价结果来看，画作六获得最高的评价，依序为画作二、画作五、画作一，画作三与画作四殿后，该专家針對五个评价面向的综合评价如下。

江水的表现：整体来说，浪表现很好，但對於江浪的细节表现必須再考证研究。特别是江浪與海浪不分，誤将海浪及湖水视為江浪。

浪的渲染：整体来说的渲染效果極佳，然而以古典水墨视角来看则缺乏筆性和線性，几乎全为西洋画法。

历史氛围的营造：整体来说歷史氛圍的营造過於概念化，缺乏歷史感知的厚度。

情感氛围与哲思：整体上可以再加強。

艺术风格与技巧：技巧很好，還談不上風格層面。总体風格還是偏向於商品繪畫，有很好的技法，但與水墨畫關係不大。

4.4. 綜合比較

在比较 ChatGPT、文学领域专家和绘画领域专家对六幅画作的评价结果时，可以发现他们在评分和综合评价上的重点和侧重点有明显的差异。

首先，从评分差异来看，ChatGPT 的评分相对较高，特别强调浪的渲染和艺术风格与技巧，同时在情感氛围与哲思方面也给出了高分，显示出对画作的动态感和艺术表达的重视。而文学领域专家的总体评分较为均衡，但更注重情感氛围与哲思和历史氛围的营造，强调画作在表达诗歌内容和历史文化背景方面的能力。至于绘画领域专家所给出的总分普遍较低，对所

有画作的艺术风格与技巧持较为严格的评价，重视画作在传达历史氛围和情感表达的深度上的表现。

其次，在评价的侧重点与解读上，ChatGPT 侧重于画作如何体现苏东坡词意的壮阔气魄和哲理，对技术性较高和动感强烈的画作给予更高评价（如画作一、画作三）。而文学领域专家则更关注画作是否能深刻表达文学作品的情感和哲学思考，对能够引发深度思考和情感共鸣的画作评价较高（如画作五、画作六）。至于绘画领域专家则从绘画技术和风格的角度进行评价，对传统水墨风格的忠实度和技术执行的精细度给予较大关注，批评现代化或西化的表现手法，此点可从评价较高的画作六、画作二及画作五窥其一二。

这三种评价反映出不同背景下的专业视角对艺术作品的多样解读和评价标准。ChatGPT 作为 AI，可能更倾向于从数据中提取的特征和总体印象进行评价；文学领域专家从文学和文化的角度深入探讨画作与原文的关联和情感表达；而绘画领域专家则从传统艺术技法和风格的保持程度进行严格审视。这些差异展示了艺术评价的主观性和多元性，不同的评价者根据自己的专业背景和审美标准，会有不同的感知和评价结果。

5. 结论与建议

综上所述，ChatGPT 与文学领域专家及绘画领域专家在评价六幅画作时存在显着的差异。这种差异既体现在排名和评分上，也反映在综合评价上。举例来说，画作五及画作六虽未获得 ChatGPT 较高的评价，却同时在文学及绘画领域专家眼中得到较高的评价；而获得 ChatGPT 较高评价的画作一及画作三却在文学及绘画领域专家眼中近乎敬陪末座。这种不小差异不啻反映了不同评价者在审美和评价标准上的不同，也提醒吾人在运用 ChatGPT 等大语言模型来评价艺术或图像作品时需要综合考虑多种因素，特别是更全面的导入专业的领域知识（如文学及绘画），以避免过于片面和主观的判断结果。

本研究巧妙地运用了多种提示工程技术，借以引导和深化问题的分析，简要统整如下。

首先，吾人统整丰富的研究文本及赏析资料以提供 ChatGPT 清晰的目标和背景信息，这属于生成知识提示 (Knowledge Prompting)，有助于大型语言模型建立背景知识，理解文本的上下文和意图。此外，本研究采用生成知识提示的技巧，通过分析特定的画作与词句内容的关联来生成新的知识或理解。这种方式巧妙的从生成的知识中再萃取出基于生成知识的评价面向及对应的评价标准，据此得以对画作给出客观的评价与诠释。

其次，吾人将任务分解成多个子问题，逐步引导模型进行思考，例如询问“苏轼怀念了哪些往事”以及“风流人物”的指代，这体现了链式思考 (Chain-of-Thought Prompting) 的技巧，让模型逐步推理和分析，得出更全面和深入的答案，包括画作的客观与系统性的分析与评价。此外，本研究也融入方向性刺激提示（Directional Stimulus Prompting）的手法，通过提供具体的评价维度和要求，如“江水的表现与浪的渲染”、“历史氛围的营造”等，明确的给出方向性的刺激，促使生成的回答更加贴合问题的具体需求。

综上所述，本研究结合多种提示工程技术来引导对复杂问题的全面性理解和结构化分析。具体的成果不仅提供了基于丰富的词意赏析文本所生成的 AI 绘画提示词，也结合了 ChatGPT 的识图能力，逐步给出对生成画作的客观分析，有效提升了吾人对词意理解与赏析的深度与广度，也促进了更精细和全面的知识生成。

参考文献

王兆鹏、郁玉英和郭紅欣（2012）。宋词排行榜。上海：中华书局。

叶丽清（2022）。《念奴娇・赤壁怀古》教学内容的调查与分析。佛山科学技术学院。DOI:10.27960/d.cnki.gfskj.2020.000125。

刘乃昌等（1998）。唐宋词鉴赏辞典（唐．五代．北宋）。上海：上海辞书出版社。
刘滋培（1981）。苏轼《念奴娇・赤壁怀古》赏析。固原师专学报(社会科学版)，1，49-52。
朱文成（2006）。《念奴娇・赤壁怀古》教学案例。语文建设，9，36-38。
许渊冲（2021）。许渊冲译宋词三百首。北京：中译出版社。
孙笛和李振峰（2023）。心物同形理论下的高中古诗词教学策略——以《念奴娇・赤壁怀古》教学设计为例。汉字文化，14，105-107。
汪婉青和周庆华（2000）。从《念奴娇・赤壁怀古》谈语文教学的美育作用。安徽教育学院学报，1，98-99。
陈国华（2015）。一首壮词中的幻灭迷梦——浅析苏轼《念奴娇・赤壁怀古》。语文教学通讯・D 刊(学术刊)，7，68-69。
徐中玉、金启华（1999）。中国古代文学作品选（二）。上海：华东师范大学出版社。
盛大林（2023）。苏轼《念奴娇・赤壁怀古》新考新论。商丘师范学院学报。39，40-53。
彭秀模和吴廣平，1984；彭秀模和吴广平（1984）。苏轼《念奴娇・赤壁怀古》辨析。吉首大学学报(社会科学版)，1，42-48。
蘅塘退士等（2009）。唐诗三百首．宋词三百首．元曲三百首。北京：华文出版社。
ChatGPT 4 & Sabit Ekin (2023). Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices. *TechRxiv*. May 04.
Chen, H. C., & Chen, Z. (2023). Using ChatGPT and Midjourney to Generate Chinese Landscape Painting of Tang Poem 'The Difficult Road to Shu'. *International Journal of Social Sciences and Artistic Innovations, 3* (2), 1-10.
Chen, J., et al., (2021). Automatic Poetry Generation Based on Ancient Chinese Paintings. *Journal of Computer-Aided Design & Computer Graphics*, *33* (7), 1038-1944.
Li, X., & Li, Y., (2006). Simulation of Chinese Ink-Wash Painting Based on Landscapes and Trees. *2006 Second International Symposium on Plant Growth Modeling and Applications*, 328-334.
Rasyad, F., et al., (2023). A Systematic Literature Review of Generative Adversarial Network Potential In AI Artwork. *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, Jakarta, Indonesia, 853-857.
Yuan, S., Dai, A., Yan, Z., Liu, R., Chen, M., Chen, B., Qiu, Z., & He, X. (2022). Learning to Generate Poetic Chinese Landscape Painting with Calligraphy. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)*, 5019-5022.

A Large Language Model-Enabled Solution for the Automatic Generation of Situated Multiple-Choice Math Questions

Ruijia Li[1,*], Yuan-Hao Jiang[2, 3, 4], Yiting Wang[5], Hanglei Hu[1], Bo Jiang[2]

[1] Department of Education Information Technology, East China Normal University
[2] Lab of Artificial Intelligence for Education, East China Normal University
[3] Shanghai Institute of Artificial Intelligence for Education, East China Normal University
[4] School of Computer Science and Technology, East China Normal University
[5] Software Engineering Institute, East China Normal University
[*] 10214507406@stu.ecnu.edu.cn

Abstract: *In this research, we introduce a methodology employing the LangChain framework to automatically generate contextualized multiple-choice mathematics questions aimed at core literacy. This methodology utilize multiple iterations of large language model dialogues and auxiliary external tools to ensure high accuracy and relevance to real-world scenarios. Furthermore, we present and implement the integration of this methodology with an online learning platform, while also considering the pertinent ethical and security concerns. The code is available online on the project's GitHub page: https://github.com/youzizzz1028/MCQ-generation-Chain*

Keywords: automatic question generation, LangChain, ChatGPT, core literacy, prompt engineering

1. Introduction

1.1. Research Background

With the advancement of artificial intelligence technology, various types of adaptive learning systems are being implemented in educational settings. Multiple-choice questions (MCQs) are utilized as a testing tool in online learning, mobile learning, and blended learning (Nagasaka, 2020) due to expedite testing, automate grading, and rapid assessing. Nonetheless, the manual creation of questions is laborious and time-intensive. Despite available to question bank resources, the processes of formatting, tagging, and integrating questions demand substantial human effort. Investigating methodologies for automatically generating MCQs with a predetermined level has emerged as a pivotal aspect of learning system development.

The previous methods such as random number generation (Nagasaka, 2016), named entity recognition (Ali et al., 2010), and the Transformer architecture (Vachev et al., 2022) are applied to generate questions. These studies mainly focus on generating questions in the humanities and linguistic disciplines. However, in the field of mathematics, predominantly template-based random number generation methods are commonly used due to high demand for logical rigor. Consequently, this leads to lower diversity and higher costs in question generation, making widespread applicability difficult. This study explores how to utilize large language models(LLMs) to generate mathematical questions. Indeed, large language models exhibit deficiencies in mathematical reasoning abilities (Frieder et al., 2024), thus necessitating the incorporation of external tools. Utilizing retrieval augmented generation

method with external knowledge databases, code interpreters and PoT (Program of Thought) (W. Chen et al., 2022), the quality and diversity of questions could be improved.

Furthermore, the "Mathematics Curriculum Standards for Compulsory Education (2022 Edition)" points out that the three-dimensional core competencies of the mathematics discipline are: observing the real world with mathematical perspective, thinking about the real world with mathematical reasoning, and expressing the real world with mathematical language. The context plays a role of connection and support in acquisition of knowledge and the cultivation of core competencies, serving as an essential bridge to establish connections between the domain of knowledge and the real world (Siddiq et al., 2016). Therefore, it is necessary to integrate quiz with real-life scenarios and craft questions reflecting students' lives. (Xu et al., 2023).

1.2. Research Contents

This study aims to investigate an efficient approach for automatically generating contextualized MCQs in mathematics education and applying them to online learning platforms. Therefore, the specific research questions addressed in this study are:

(1) What is the technical solution of generating contextualized mathematical MCQS utilizing LLMs?

(2) How to integrate the questions generated by LLMs into online learning platforms?

2. Related Work

Due to the rise of large-scale online course platforms and adaptive learning platforms, the construction of question banks is receiving recognition. For learners, questions are plagued by non-reusability. Phenomena such as item exposure and item fatigue (Revuelta & Ponsoda, 1998) require constant updates to question banks to maintain the validity and discrimination of assessments. However, manually creating questions is a complex task that incurs costs in terms of time, labor and resources. As a result, an increasing number of researchers turn their attention to the field of Automatic Question Generation (AQG).

(Kurdi et al., 2020) classifies and describes AQG systems across seven dimensions: (a) the purposes and uses of question generation, (b) the domains of knowledge under assessment, (c) the sources of knowledge, (d) the methods of generation, (e) the types of questions, (f) the formatting of question structures, and (g) the evaluation of question quality. Presently, question generation technologies primarily aim to produce test questions, bolster online learning platforms, or aid students in self-directed learning, with a focus on disciplines including language, mathematics, humanities, and medicine. Drawing upon a text, knowledge graphs, or ontologies (Sewunetie & Kovacs, 2022), classical methods for generating questions can be categorized into three main types: (a) Syntax-based approaches, which involve identifying key concept words using techniques such as syntax trees and POS (Part-of-speech tagging), and generating questions by replacing these key concepts. (b) Semantic-based approaches, which primarily utilize natural language processing techniques to identify word properties and dependencies among them and generate questions based on extracted dependencies. (c) Template-based approaches, which predefine question templates and fill them using random number generation or keyword extraction. Subsequently, as deep learning technologies have advanced, neural networks (Zhou et al., 2018), reinforcement learning (Y. Chen et al., 2019; Wang et al., 2020), and pre-trained language models (Chan & Fan, 2019) have been integrated into AQG

systems. Furthermore, some researchers pay attention to the utilization of LLMs for generating questions (Lee et al., 2023).

However, owing to the requirements of computational accuracy and logical rigor in the field of mathematics, it is not suitable to solve problems using connectionist artificial intelligence algorithms, such as neural networks. Instead, symbolic methods are often employed for solutions (Zong & Krishnamachari, 2023). Numerous technical solutions in AQG systems are inapplicable to question generation in the field of mathematics. The predominant method still relies on generating random numbers using predefined templates, which can result in problems such as rigid question formats, detachment from real-life scenarios, and low scalability. (Grévisse, 2024) investigated the quality of multiple-choice questions generated by GPT models. Its result shows that while questions are linguistically fluent, they still have many issues including overly long stems, absence of correct answers in the options, ambiguous wording, and low relevance to the topic. In our experimentation, it is also confirmed that directly using GPT models and prompts to generate questions, especially multiple-choice questions, is challenging to achieve optimal outcomes. Thus, it is necessary to introduce external knowledge libraries to serve as inputs for knowledge and complementary teaching strategies. It also incorporates code interpreters and PoT (W. Chen et al., 2022), regulates the process of question generation and adjusts super-parameters of LLMs.

3. Contextualized Math Multiple Choice Question Generation

3.1. Technical Solutions

The process of creating and generating multiple-choice questions involves six stages: (Das et al. (2021): (a) Background input, (b) Preprocessing, (c) Question generation, (d) Answer generation, (e) Distractor generation, and (f) Formatting. This study adopts this workflow and proposes a solution for automatically generating mathematical MCQs based on the LangChain framework. The solution integrates multi-turn large language model dialogues and pluggable interfaces, such as SerpAPI and llm-math. The solution consists of four core modules: the Domain Agent, the QG (Question Generation) Agent, the Solution Agent, and the Option Generation Agent. These modules are interconnected, ultimately generating mathematical MCQs based on real-world scenarios while ensuring quality assurance.

3.1.1. Domain Agent

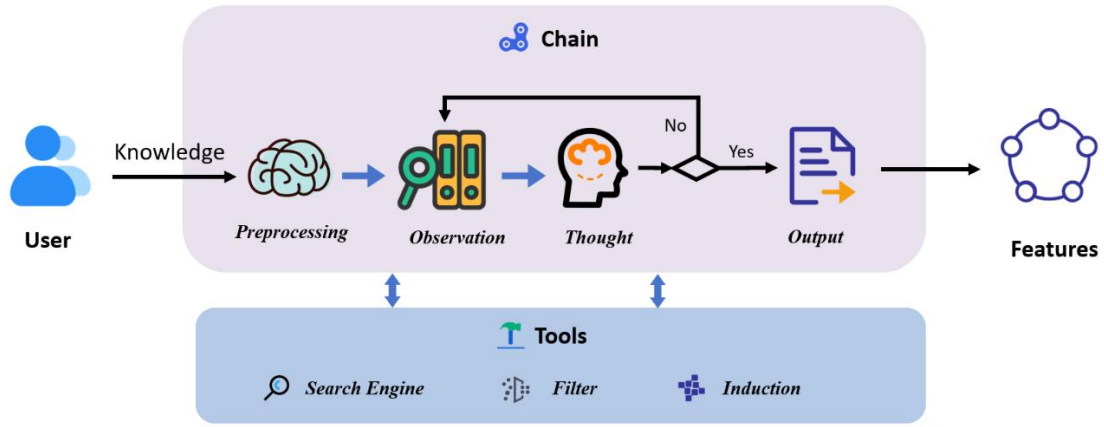


Figure 1. Domain Agent

This agent utilizes SerpAPI to crawl web resources, analyze the main assessment methods and provide quiz examples, which assist the LLMs in establishing a domain foundation related to specific knowledge points, thereby reducing hallucinations and knowledge errors.

3.1.2. Question Generation Agent

The Question Generation Agent receives background information passed from preceding agent and user-defined parameters, including knowledge points, number of questions and difficulty. Additionally, this agent, which utilizes a large amount of corpus on the internet, is assisted by search engine to validate whether the context of questions aligns with real-life scenarios.

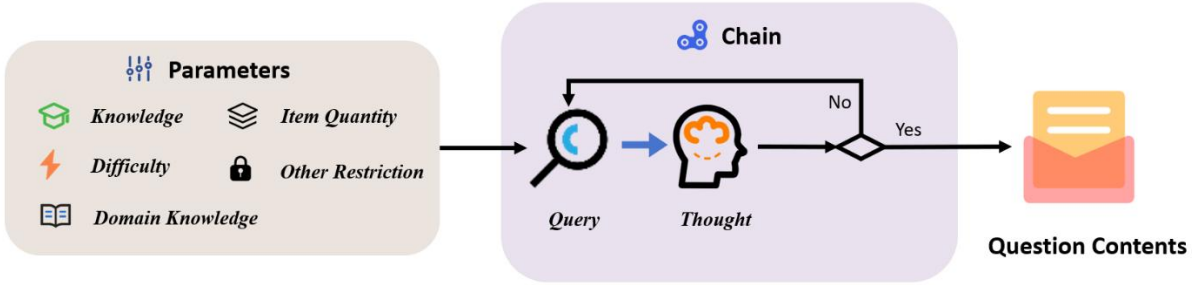


Figure 2. Question generation Agent

3.1.3. Solution Agent

The Solution Agent conducts numerical calculations by utilizing llm-math, POT and Python interpreter. During question analysis, the LLMs supplies Python expressions suitable for problem-solving and executes the code step by step. This process results in determining the correct answer, which is ultimately presented in natural language form.

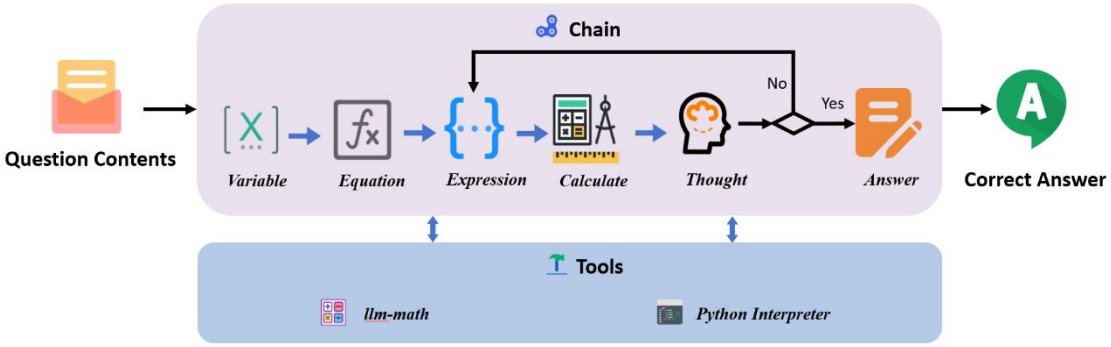


Figure 3. Solution Agent

3.1.4. Option Generation Agent

This agent provides templates for generating MCQs. After embedding the questions and correct answers generated from the preceding steps into the prompt, it generates distractors and shuffles the order of them to format the questions. Additionally, it incorporates formatting hints and checks for data type coercion, ultimately producing contextually multiple-choice questions in strict JSON format.

3.2. Platform Design and Ethical Reflection

Some studies suggest that ChatGPT should serve to guide and collaborate in the generation of quizzes, aiming to alleviate teachers' workload rather than supplanting their tasks (E Sabzalieva & A Valentini, 2023; Lo, 2023). This study aims to enhance the quality of generated questions while considering the sensitivity of educational applications, proposing that machine-generated quizzes and

practice materials serve as a baseline, rather than being directly distributed to students. Thus, in terms of platform design, we have provided teachers with convenient tools for question generation, modification, and question bank management. Additionally, supporting exportation into various file formats (such as docx, excel, JSON, etc.) facilitates compatibility with traditional conventional assessments and other learning platforms.

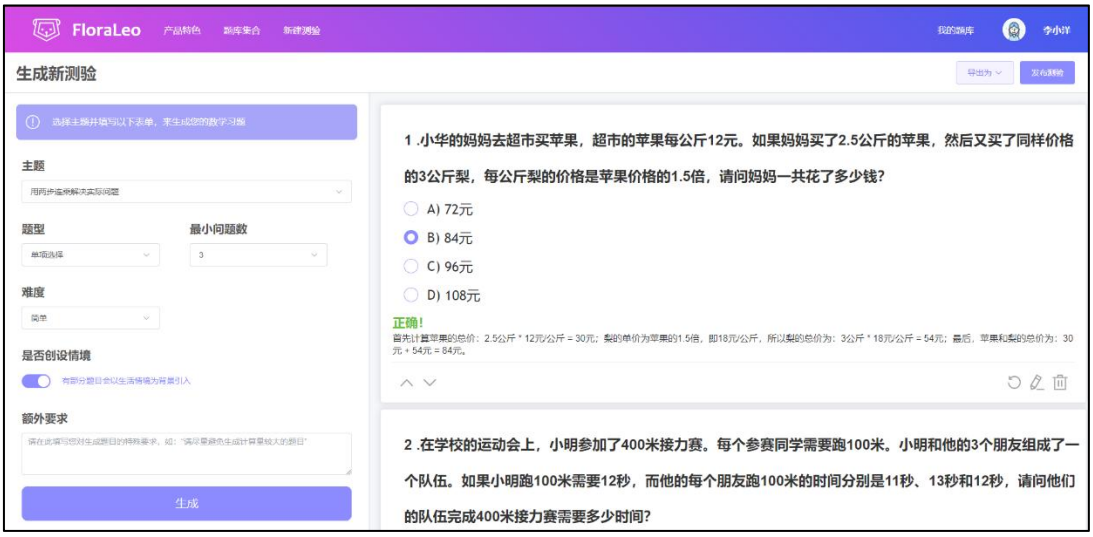


Figure 4. Exercise generation page

4. Discussion and Conclusion

In conclusion, this study proposes a multi-agent solution utilizing large language models within the LangChain framework and pluggable external tools for the automatic generation of contextualized mathematical multiple-choice questions. The accuracy of the questions generated by this approach is assured to a certain extent. Their quality is greater than those directly generated by GPT4 and highly resembles manually crafted questions in various aspects. By leveraging the semantic expression capabilities of large language models and data from external search engines, the generated questions are more closely aligned with real-life scenarios. As a result, the generated questions address the cultivation and assessment of mathematical core literacy.

However, this approach exhibits certain limitations. Firstly, the support for question types is not yet comprehensive, as the current version is solely designed for multiple-choices questions, lacking mechanisms for generating fill-in-the-blank and true/false questions. Secondly, the method for controlling the difficulty of question generation is crude, as it currently relies solely on adjusting prompt words based on the understanding of the large language model, lacking a more precise quantification method. Lastly, the generated questions are limited to topics related to numbers and algebra, with no provision for multi-modal support for geometry questions and statistical chart questions.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 61977058, and the Natural Science Foundation of Shanghai under Grant 23ZR1418500.

Reference

Ali, H., Chali, Y., & Hasan, S. A. (2010, July). Automatic question generation from sentences. In Actes de la 17e conférence sur le Traitement Automatique des Langues Naturelles. Articles courts (pp. 213-218).

Chan, Y. H., & Fan, Y. C. (2019, November). A recurrent BERT-based model for question generation. In Proceedings of the 2nd workshop on machine reading for question answering (pp. 154-162).

Chen, W., Ma, X., Wang, X., & Cohen, W. W. (2022). Program of thoughts prompting: Disentangling computation from reasoning for numerical reasoning tasks. arXiv preprint arXiv:2211.12588.

Chen, Y., Wu, L., & Zaki, M. J. (2019). Reinforcement learning based graph-to-sequence model for natural question generation. arXiv preprint arXiv:1908.04942.

Das, B., Majumder, M., Phadikar, S., & Sekh, A. A. (2021). Automatic question generation and answer assessment: a survey. Research and Practice in Technology Enhanced Learning, 16(1), 5.

Sabzalieva, E., & Valentini, A. (2023). ChatGPT and artificial intelligence in higher education: quick start guide.

Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P., & Berner, J. (2024). Mathematical capabilities of chatgpt. Advances in Neural Information Processing Systems, 36.

Grévisse, C. (2023, October). Comparative Quality Analysis of GPT-Based Multiple Choice Question Generation. In International Conference on Applied Informatics (pp. 435-447). Cham: Springer Nature Switzerland.

Kurdi, G., Leo, J., Parsia, B., Sattler, U., & Al-Emari, S. (2020). A systematic review of automatic question generation for educational purposes. International Journal of Artificial Intelligence in Education, 30, 121-204.

Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2023). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education. Education and Information Technologies, 1-33.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410.

Nagasaka, K. (2020). Multiple-choice questions in mathematics: Automatic generation, revisited. In The 25th Asian technology conference in mathematics, virtual format, Radford University, Virginia, USA and Suan Sunandha Rajabhat University, Thailand.

Revuelta, J., & Ponsoda, V. (1998). A comparison of item exposure control methods in computerized adaptive testing. Journal of Educational Measurement, 35(4), 311-327.

Sewunetie, W. T., & Kovács, L. (2022, November). Comparison of automatic question generation techniques. In 2022 IEEE 22nd international symposium on computational intelligence and informatics and 8th IEEE international conference on recent achievements in mechatronics, automation, computer science and robotics (CINTI-MACRo) (pp. 000025-000030). IEEE.

Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I., & Scherer, R. (2016). Taking a future perspective by learning from the past–A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. Educational research review, 19, 58-84.

Vachev, K., Hardalov, M., Karadzhov, G., Georgiev, G., Koychev, I., & Nakov, P. (2022, April). Leaf: Multiple-choice question generation. In European Conference on Information Retrieval (pp. 321-328). Cham: Springer International Publishing.

Wang, L., Xu, Z., Lin, Z., Zheng, H., & Shen, Y. (2020, December). Answer-driven deep question generation based on reinforcement learning. In Proceedings of the 28th International Conference on Computational Linguistics (pp. 5159-5170).

Xu, X., Zhang, Q., Zhang, Y. & Shen, H.How to Design Contextualized Assessment of Information Literacy for Students in Primary and Secondary Schools. (2023). Modern Distance Education Research: (pp. 1-10)

Zhou, Q., Yang, N., Wei, F., Tan, C., Bao, H., & Zhou, M. (2018). Neural question generation from text: A preliminary study. In Natural Language Processing and Chinese Computing: 6th CCF International Conference, NLPCC 2017, Dalian, China, November 8–12, 2017, Proceedings 6 (pp. 662-671). Springer International Publishing.

Zong, M., & Krishnamachari, B. (2023, June). Solving math word problems concerning systems of equations with gpt-3. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 13, pp. 15972-15979).

教育大语言模型应用的 SWOT 分析：机遇、挑战与对策

SWOT Analysis of the Application of Educational Large Language Models: Opportunities, Challenges, and Strategies

徐琦[1]，魏艳涛[1,*]，刘清堂[1]，肖光润[1,2]，师亚飞[3]
[1]华中师范大学人工智能教育学部　数字教育湖北省重点实验室
[2]湖北文理学院
[3]河南师范大学
[*] yantaowei@mail.ccnu.edu.cn

【摘要】 大语言模型（LLMs）引起了教育领域的极大关注，是教育数字化转型面临的新机遇和新挑战。基于此，研究围绕现有教育 LLMs 应用现状，提出教育 LLMs 应用的 SWOT 分析框架，从教育环节上分析其优势、劣势、机遇与挑战。研究发现，教育 LLMs 具有文本生成等七项优势，不具备类人创造性等五项劣势，其应用在存在助力智能教学、加快数字治理等六项机遇，同时也存在教学模式不清晰等、管理权责易纠纷等六项挑战。基于分析结果提出了相应对策，包括提升数字素养、发展高阶思维，培养学习意识、开展情感关怀，权衡使用度、健全制度规范等，以规避潜在教育风险，推进 LLMs 应用赋能教育高质量发展。

【关键词】 大语言模型；SWOT；教育应用

Abstract: *Large Language Models (LLMs) have attracted significant attention in the education, presenting new opportunities and challenges for the educational digital transformation. In this context, research focuses on the current applications of educational LLMs, proposing a SWOT analysis framework for the application of educational LLMs. This framework analyzes the strengths, weaknesses, opportunities, and challenges of educational LLMs across educational aspects. The study reveals that educational LLMs possess seven advantages, such as text generation, while lacking in five areas, including human-like creativity. The application of LLMs presents opportunities in areas such as enhancing intelligent teaching and accelerating digital governance. However, challenges exist, such as unclear teaching models and potential disputes over management responsibilities. Based on the analysis, corresponding strategies are proposed, including enhancing digital literacy, fostering higher-order thinking, cultivating a culture of learning, providing emotional support, balancing usage, and strengthening institutional regulations. These strategies aim to mitigate potential educational risks and promote the empowerment of high-quality education development through the application of LLMs.*

Keywords: large language models, SWOT; education application

1. 前言

新一代人工智能已经迈向大模型时代，大模型也随之成为数字时代的新基座。大语言模型（Lage Language Models，LLMs）作为大模型的核心底层技术，具备生成多模态、跨领域、且与人类表达高度相似的内容，这深刻影响着教育的组织形式、治理方式、教学结构和学习生态等。随着教育数字化转型的推进，学界已经开始研究关注和应用教育 LLMs，如 MathGPT 等，但较少反思教育应用中存在的诸多挑战（申丽萍等，2024）。此外，教育相对于商业、医疗等领域，对 LLMs 的精准性、伦理道德等要求更高（苗逢春，2023），人类目前对于 LLMs 的风险仍充满担忧，LLMs 能否满足教育领域的应用需求仍需要多重考究。

基于此，本文采用态势分析法（Strengths Weaknesses Opportunities Threats，SWOT），探索并回应教育 LLMs 的外部优势（S）和劣势（W），以及它应用于教学管评测研等教育环节的外部机遇（O），和对教学者、学习者、管理者、评价者、测试者和教研者可能带来的外部挑战（T）。根据分析结果，聚焦人机主体制定对应措施，以期提高广大教育工作者与研究者对 LLMs 的认知，帮助其做出更科学的教育决策，推进数字教育高质量发展。

2. 教育大语言模型应用的 SWOT 框架

根据现有教育 LLMs 的应用案例，如表 1 所示，采用 SWOT 分析法对 LLMs 的内部优势（S）和内部劣势（W）进行归纳，分析教育 LLMs 应用的外部机遇（O）和外部挑战（T），从而进行优劣势组合，提出教育 LLMs 应用的发展策略。

表 1 教育 LLMs 应用案例

序号	名称	研发单位	采用的 LLMs	主要功能	应用
1	星火语伴	科大讯飞	星火认知大模型	语言学习	https://sj.qq.com/appdetail/com.iflytek.aistudy.aitalk
2	EmoGPT	华东师大	自研大模型	心理疏导	https://shmh.ecnu.edu.cn/0a/b0/c32784a527024/page.htm
3	MathGPT	好未来	自研大模型	数学学习	https://www.mathgpt.com/
4	智海-三乐	阿里云	通义千问	知识学习	https://github.com/zhihaiLLM/wisdomBot
5	Khanmigo	可汗学院	GPT-4	多学科学习	https://www.khanmigo.ai/
6	子曰	网易有道	自研大模型	AI 口语学习	https://ziyue.youdao.com/#/home
7	阿尔法蛋儿童认知大模型	淘云科技	星火认知大模型	儿童练习表达机器人	https://www.toycloud.com/AlphaEggApp/download.html
8	银河大模型	作业帮	自研大模型	课后辅导	https://www.zybang.com/official/pages/news/

图 1 为教育 LLMs 应用的 SWOT 分析框架，从其主要功能分析它的优势（S）和劣势（W）。通过梳理 LLMs+教学（教、学、研）和评估（管、评、测）典型应用，分析其赋能教育的潜在机遇（O），根据聚焦教学者、学习者、管理者、评价者、测验者、教研者六大主体，分析其对教育的挑战（T）。最后，根据上述分析结果，围绕教育 LLMs 融入教育过程中所涉及的人机核心主体，提出对应策略，化风险为发展的动力，不断推进教育数字化高质量发展。

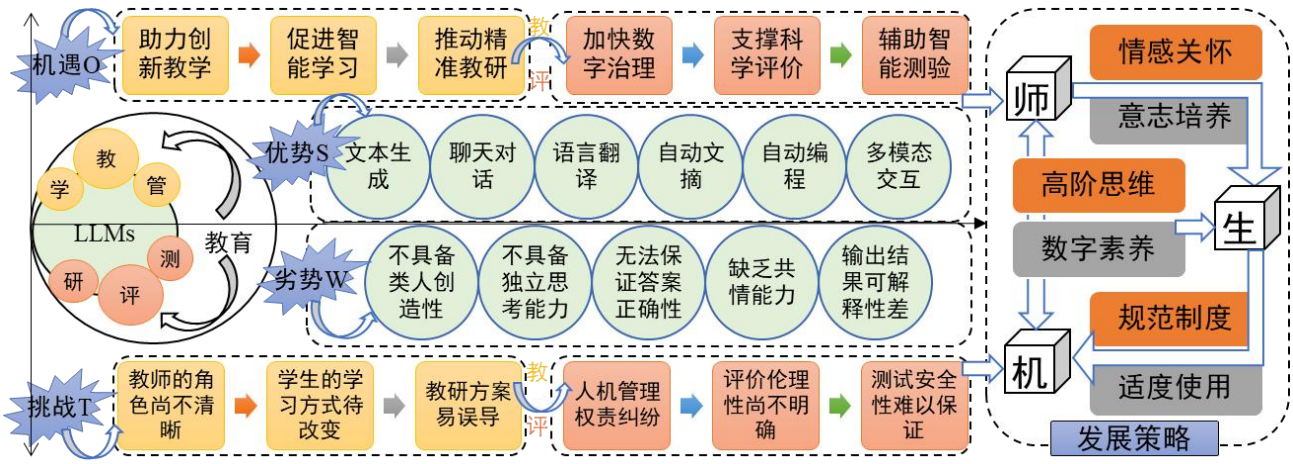


图 1 教育大语言模型应用的 SWOT 分析框架

3. 研究结果

3.1. 教育大语言模型应用的内部优势（S）

语言模型实际上是一个基于上下文的分类器，通过对上下文进行描述，模型就可以预测文本中下一个单词或句子。随着语言模型的参数增多、训练数据量增大，语言模型进入了大语言模型时代。目前，主流的 LLMs（例如，谷歌 LaMDA、百度文心一言）均基于 Transformer 进行预训练、自监督训练、自回归或自编码模型等，只是在参数、训练周期、优化算法、指令微调等细节上存在差异（Gordijn et al., 2023）。从教育 LLMs 的应用来看，其优势（S）体现在文本生成、聊天对话、语言翻译、自动文摘和多模态交互这七个方面。

文本生成是为用户生成基础文本内容（骆飞等，2023），通过单轮会话帮助用户快速筛选信息，完成文字推介、介绍等任务，提升信息传播的速度与效率。聊天对话是根据用户的需求和反馈，调整并生成响应式文本。重点根据上下文信息作答，并根据用户的要求更新回复内容（Wu et al.，2023）。语言翻译是对需要翻译的多国语言进行分析和处理，将其转化为目标语言的对应文本。这个过程包括文本分词、词性标注、句法分析、语义理解等。自动文摘是从大量文本中提取关键信息，以抽取式+生成式的方式组合形成有效文摘，节省用户的时间和精力（周玲等，2023）。自动编程是根据目标程序语言需求，生成符合逻辑的新代码，生成式代码可作为初步蓝图，也可用于代码 BUG 检查，为用户提供建议。多模态交互是通过生成和接收语言、图像、视频等模态与用户进行交互，这其能够处理更加复杂的任务（Adesso，2022）。

3.2. 教育大语言模型应用的内部劣势（W）

虽然 LLMs 结构复杂，但是它仍然无法完全模拟人类思维和推理过程，它的表现依赖于训练的数据，由于现有数据的局限性和偏差，导致模型在具体教育环节上可能表现不佳，这些劣势（W）主要体现在不具备类人创新性、不具备独立思考能力、无法保证答案准确性、缺乏共情能力和结果可解释性差这五个方面。

第一，不具备类人创造性。真正的创造性一定有智力难度，主要是创造概念和理论、发现规律或提出定理等。LLMs 不能凭借对现实世界的感官体验在“实践”中形成原创性的知识，不能经历从实践中来、到实践中去的认识升华过程，它的学习进化也只是工作技能和质量的进化（卢宇等，2023），它没有创造性。第二，不具备独立思考能力。LLMs 以输入-输出方式工作，可以根据上下文关系将聊天持续下去，而且已经可以识别分析长文本输入指令，但一直是被动应答，不能主动提出问题并通过平等对话启发提问者的思路，只能“言传”而不能“意会”。它不能借助没有答案的问题启发人的思维，也没有独立思考能力。第三，无法保证答案正确性。LLMs 并不能保证它提供的答案始终是正确的。因为它的回答是基于其训练数据和算法生成的最有可能正确的回复，并不一定是最新的或完全准确的。第四，缺乏共情能力。LLMs 没有人的情感、人的意志品质，不能与人进行情感交流与共鸣，不懂得爱和被爱，没有作为人的社会关系和个体经验的积累，缺乏独特性和个性（Wesley，2023）。第五，结果可解释性差。它基于 Transformer 结构的黑盒模型的可解释性差（Chaoning et al.，2023），也一直是深度学习领域在不断探索的问题。可见，探讨 LLMs 在教育领域中的应用时，需要时刻关注其内部劣势（W），避免带来新的问题，如知识错误、价值观偏离等。

3.3. 教育大语言模型应用的机遇（O）

基于上述 LLMs 的内部优势（S）和劣势（W），教育作为一个复杂的系统，具有独特的环节和流程，这些典型环节包括教学和评估两个场景，具体包括教学、学习、教研、管理、评价和测试六个方面。因此，分别以教育环节和 LLMs 的功能优势作为横纵轴进行交叉分析，梳理出教育 LLMs 应用的六项机遇（O），下面将具体阐述。

在教学方面，第一，助力智能教学。这就包括在备课阶段辅助课程设计（Official Microsoft Blog，2024），在课堂活动阶段丰富资源呈现方式，如语文意境创设、函数图像生成、几何学图像演示、地理区域绘制、音乐作品赏析、医学影像解释、艺术创作、体育动作设计等，在课后阶段归纳课程内容，辅助教学反思等。第二，促进智能学习。教育 LLMs 可以通过问答、追问和反问等方式加深对学生对知识的理解和应用，拓展学习思维和提高学习效率。还可用于学习智能推荐、答疑解惑、知识文摘、识图解题等方面，在语言学习和编程学习方面也具有优势，如帮助学生查词译句和译文优化，生成代码指导或 BUG 提示等，满足学习者多维学习需求（沈书生等，2023）。第三，推动精准教研。教育 LLMs 能在教研方案、研究方法、教学分析报告等方面提供见解（段俊吉，2023），尤其能避免新手教研者地盲目探索。目前，GPT-4 已装上了 Wolfram 插件（Stephen wolfram，2024），可以提供数千种可用算法、权威数据和准确计算，有望协同教研者完成教育数据分析和讨论，提升教师的综合素质。

在评估方面，第一，加快数字治理。LLMs 可以在资源调配、课程建设和师生信息管理方面提供便利。例如，筛选教师简历、统计学籍信息，生成通知和规章制度等（王蕾，2023）。目前，GPT-4 已整合到 Microsoft 365 Copilot 应用中，既可以总结图文信息，还能轻松切换 Word、PPT、Excel 等数据格式，为便携式管理提供了参考。第二，支撑科学评价。这就包括筛选评价指标、优化评价方案等。随着 LLMs 的长文下长度的增加，更多细粒度高精度多场景的评价信息将有望生成，尤其为学生写作、报告、手册、论文等复杂的内容提供了评价参考（Patel et al.，2023），以缓解人工评价的费时耗力，提升教学评价的科学性。第三，辅助智能测验。LLMs 能辅助命题设计，如为生成精准化的题型、题项和参考答案等。还可以辅助试卷生成，例如，命题拆分、试题重组、难易搭配、随机组卷等（焦建利，2023），真正实现“百人白卷”，进一步确保考试公正公平。它还具备自动阅卷的功能，节约考评者的时间精力，提高阅卷效率。

3.4. 教育大语言模型应用的挑战（T）

教育 LLMs 应用的挑战隐藏在教育系统的六大环节，研究从人机协同的视角深挖生成式人工智能在教学和评估两方面对教学者、学习者、管理者、评价者、测验者和教研者带来的潜在挑战（T），避免其在教育领域盲目推广和普及，以发挥其实有价值和贡献。

在教学方面，主要有教学模式不清晰、学习方式待改变、教研过程易误导三重挑战。从教育角色上，教育 LLMs 正逐渐形成一类新型教育者，人师与机师的界限仍不清楚，教师角色面临着威胁。从教师能力上，新的数字鸿沟正在形成，这既存在于会与不会者之间，也存在于过度依赖与常规使用者之间（Marshall，2024）。而这些处于数字边缘的教师可能面临着淘汰危机。在思维模式上，教育 LLMs 所采用的描述迭代、综合搜索等机制恰好模拟了建立新旧知识联动的学习过程（Taha et al.，2023），如何培养更高阶的思维层次将成为现代教育的难题。在思考方式上，过度依赖生成的“答案”会导致学生缺少原创性思考，加剧“学习偷懒”和“投机取巧”现象，LLMs 的生成机制仍需要人工干预。在学习能力上，生成式答案往往真假难辨（Tyna et al.，2023），学生容易固化在狭隘的认知世界，降低思维灵活性，难以形成正确的价值观和良好的学习习惯。此外，由于教育 LLMs 仍未达到通用人工智能水平，生搬硬套的痕迹仍很明显，对教研的实用价值仍需要教师重视与考量。目前，已有高校出台了“生成式人工智能教师使用指南”，但是教师与技术应用的共建共享机制仍有待完善。教研是科学严谨的过程，需要多方参与配合，LLMs 的误导行为仍待谨慎规避。

在评估方面，存在管理权责易纠纷、评价伦理性不明确和测试安全性无保障三重挑战。从用户需求上，一些学生希望通过“一键式答案”来“减负”，教育 LLMs 和 AI 抄袭检测技术同时受到了广泛关注（Nature, 2024），LLMs 生成的人工制品版权问题随之而来（Unesco, 2024），更多复杂的权责关系正不断涌现。其次，当管理者采纳了模型输出的失误决策时，机器的责

任将难以追究。随着模型的拟人化程度不断提升，更多跨国治理的难题有待解决。教育 LLMs 在诸多专业测试中都表现出优异水平（张云等，2024），新形势下的人才评价机制仍有待探索。对于教育 LLMs 生成的评价内容，它是否会生成攻击性或有偏见的内容尚不明确，是否能遵守不伤害人、尊重人、有益于人、公正待人等伦理原则仍待考究。在基本人权维护方面（Openai，2024），师生隐私信息可能难以得到保障，无疑给评价工作增加了额外负担。在组卷可信度方面，教育 LLMs 是否会捏造虚假、低质、恶意试题，试卷的检查工作仍需完善（Saib et al.，2022）。在考题安全性方面。教育考题泄露问题屡见不鲜，如何加强生成式考题的安全性仍待思考。在考试时效性方面，对于政治类考题而言，时事热点问题是考试材料的重要来源，虽然 LLMs 已具备联网搜索的功能，但其可靠性仍难以保证（张志祯等，2023）。

4. 教育大语言模型应用的对策建议

虽然教育 LLMs 在应用过程中仍存在诸多挑战（T），但其所具备的核心能力已开始对教育各环节产生直接影响和启示作用。因此，根据 SWOT 分析结果，围绕师生机协同的角度厘清应对措施，以更加开放和包容的态度去应对 LLMs 的快速发展，让技有所用、人有所长，推进教育数字化的高质量发展。

4.1. 提升数字素养，发展高阶思维

在人工智能时代，未来的工作场景需要具备高级技能的人才来面对日益复杂的人机交互任务。《提升全民数字素养与技能行动纲要 2022~2035》中提出了数字素养的基本框架，但是并没有明确具体要素，LLMs 以技术为载体助力提高学生的素养，这并不意味着技术和学习的混合叠加，还需要探讨技术赋能前后的素养提升具体路径。从人机协同的角度，学习者不仅需要出色的技术准备，如正确使用生成式人工智能技术，还需要师生、生生协同来适应不断变化的任务情境，加强技术操作、信息管理、创造力、批判性思维和解决问题能力等核心技能，以及道德意识、文化意识、灵活性、自我指导和终身学习技能等（钟秉林等，2023），这些素养与当前的经济和社会发展形态密切相关，也成为了全球现代学校培养人才的共识和目标。

同时，LLMs 技术搭建了学习者从记忆力与理解力等低阶思维能力跨越至高阶思维的桥梁，这不仅是对技术赋能学习异化的基本回应，也是技术赋能学习的现实需要。《中国教育现代化 2035》曾明确指出学生主动思考以及启发式、探究式教学的重要性。教育 LLMs 不直接产生思维，但其使用影响着思维的方式，包括其扮演的助学角色、思维工具、人机环境或应用空间。在新型生机协同过程中，高阶思维不再指向单一的技术素养和思维过程，而将成为技术延伸式的多种认知成分交织系统，包括问题分析、知识关联、信息综合、知识生成以及对思维过程的监督、管理和调节等成分。在人机交互环境下，人类“回答问题”的方式可能不再是首要的学习方式，而有效发问、维持对话、追加新问题、寻找反例、探索未知才是更重要的思维进阶学习方式，问题质量越高，个体的成长可能也会越快。

4.2. 培养学习意志，开展情感关怀

学习是一项需要意志投入和控制的活动，教育 LLMs 应用在知识获取与表达上的优异并非意味着人类学习过程不再重要，没有知识的头脑也难以评价信息质量和做出有效决策。从平庸走向成功需要更多树立理想、持之以恒、全神贯注、适时调整、乘风破浪、直面苦难和精益求精的过程，从师生协同的角度，这需要自己参与，也离不开教师的监督和引导。人可以暴力地控制机器，使其无休止地训练直至达到目标。但是，教师无法在学生身上复制这一过程，教师需要指引学生树立明确的奋斗目标与方向，引导学生持之以恒地追求自己的理想，坚定不移地信仰自己的理想。最后，人要学而有成，还要学会慎独自责，努力克制不利于自身志向实现的私心杂念，如贪婪、骄傲、吝啬、急躁等等。

教育 LLMs 的推理能力较强，对于自然科学，它会通过计算规律给出几乎正确的答案，但涉及主观、感性和人情世故等社会科学的内容，它可能难以解释这类复杂规律，也无法满足学习者的情感需求。从人机协同的角度，教育并非只关注知识层面的问题，而是建构于机器之上的开放性系统，教育需要知识，但也需要情感、想象、尊重和个性等。很多情况下，教育并不需要科学意义上的精确，它追求人文意义上的模糊。频繁的人机协作加剧了师生疏离与情感淡漠问题，我们更需要重拾师生交互中“抽象、模糊”的表达，并重新理解和打开人类情感的世界。根据马斯洛的需求层次论，学生需要知识更需要情感，例如学生做出夸张的动作吸引教师注意时，教师要发现其情感表达，增强情感交流与人文关怀。同时，课堂气氛也是学习的重要条件，这些条件也需要教师来营造和加强。

4.3. 权衡使用的度，健全制度规范

麦克卢汉认为“媒介是人的延伸”，教育 LLMs 正是这种模拟、延伸和扩展人的智能产品。当师生缺乏技术信任和数字技能时，容易降低与他人的竞争优势；而过分关注和过度使用新技术也容易在技术怪圈之中迷失自我，从而带来基础能力的退化与不良影响。同时，人工智能等技术的更新和演进速度非常快，第一代 GPT-1 到 GPT-4 经历了四代更新，每一代的性能都有明显提升，人们难以在每一轮技术更新中都找到其教育应用价值。因此，我们需要在师生机之间找到平衡点，把握使用的合适度，一是提升技术在教育中的应用质量，基于生成式内容，集中精力去关注教学再加工和知识再处理；二是将教书的重心提升至育人层次，发挥技术的知识传授等“教书”优势，重新审视“育人”工作，把教学难点、价值培养和教学研究等工作向前推进。

面向 LLMs 的潜在劣势，各国都提高了警惕，采取了相应措施来积极对应。但是现行相关法规制度却难以解决与快速进步的技术手段存在的问题，而且国内教育领域也缺乏相关使用标准、规范和指南。因此，教育领域需要推进制定相关法律法规，形成技术与教育双螺旋式的互促共进。第一，提高隐私保护意识。贯彻落实《新一代人工智能伦理规范》，切实将伦理与隐私意识融入教育应用之中。第二，落实相关责任制度。对违反相关制度的师生或机构施以相应惩罚（García-peñalvo，2023），为生成式人工智能技术使用人的行为设置“护栏”。第三，规范相关使用制度。加强监测和评估，要求师生标注应用生成式人工智能技术生成的文本内容，适当报告其使用情况等，同时研发反抄袭检测工具，保障教学公平而有序的发展。第四，构建人工智能算法伦理全过程监管体系以及教育 LLMs 应用的风险治理体系，确保未来人工智能教育应用的规范化。

5. 结语

LLMs 的教育应用势不可挡，但是师生需要保持尊重事实、谨慎判断和小心求证的心态，拥抱 LLMs 技术，直面其不良影响和潜在风险，制定契合未来教育发展的应对措施，推动技术赋能教育创新。研究提出教育 LLMs 应用的 SWOT 分析框架，从教育 LLMs 的应用现状中分析教育 LLMs 应用的内部优势（S）、内部劣势（W）、外部机遇（O）和外部挑战（T）。结果表明，教育 LLMs 应用具有文本生成、聊天对话、语言翻译、自动文摘和多模态交互 7 项功能优势（S），但是也存在不具备类人创新性、不具备独立思考能力、无法保证答案准确性、缺乏共情能力和结果可解释性差 5 项内部劣势（W），LLMs 赋能教学管评测研过程具有助力智能教学、促进智能学习、推动精准教研、加快数字治理、支撑科学评价和辅助智能测验 6 项机遇（O），也对教育主体可能带来教学模式不清晰、学习方式待改变、教研过程易误导、管理权责易纠纷、评价伦理性不明确和测试安全性无保障 6 项潜在挑战（T）。最后，综合 SWOT 分析结果，聚焦人机核心主体提出相应对策，包括提升数字素养、发展高阶思维，

培养学习意识、开展情感关怀，权衡使用的度、健全制度规范等内容，规避其可能带来的教育风险，推进 LLMs 赋能教育高质量发展。

6. 致谢

本研究受到了以下项目的资助：国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目“融合多模态数据的教师教学能力智能评价技术及应用研究”（项目号：CCNU22JC011）。

参考文献

Adesso G (2022). GPT4: The ultimate brain.*Authorea Preprints*.

Alshater M M (2022). Exploring the role of artificial intelligence in enhancing academic pe rformance: A case study of ChatGPT. *Available at SSRN*.

Chaoning Z, Chenshuang Z, Zhengsheng Z, et al. (2023). A Complete Survey on Generativ e AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 All You Need?[2024-03-10].*arXiv*.10.1 3140/RG.2.2.29980.16001.

García-peñalvo F J (2023). La percepción de la Inteligencia Artificial en contextos educativ os tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge S ociety (EKS)*, 24: e31279-e31279.

Gordijn B, Have H (2023). ChatGPT: evolution or revolution?.Medicine, *Health Care and P hilosophy*, 26(1):1-2.

Marshall C (2024). Noam Chomsky on ChatGPT: It’s "Basically High-Tech Plagiarism" and "a Way of Avoiding Learning". *Open-culture*.[2024-03-10].https://www.openculture.com/2 023/02/noam-chomsky-onchatgpt.html.

NATURE (2024).Tools such as ChatGPT threaten transparent science; here are our ground r ules for their use.[2024-03-10].https://www.nature.com/articles/d41586-023-00191-1.

Official Microsoft Blog (2023). Introducing Microsoft 365 Copilot–your copilot for work. [2 024-3-10].*https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your -copilot-for-work/*

Openai (2024). ChatGPT plugins.[2024-03-10].*https://openai.com/blog/chatgpt-plugins*.

Patel S B, Lam K, Liebrenz M(2023). ChatGPT: friend or foe. Lancet Digit. *Health*,5: e10 2.

Saib M O, Rajkoomar M, Naicker N, et al. (2022). Digital pedagogies for librarians in hig her education: a systematic review of the literature. *Information Discovery and Delivery*, 51(1): 13-25.

Stephen Wolfram (2024). ChatGPT Gets Its “Wolfram Superpowers”![2024-03-10].*https://writ ings.stephenwolfram.com/2023/03/chatgpt-gets-its-wolfram-superpowers/*

Taha B A, Uzair M, Ibtehaj U (2023). The future of medical education and research: Is C hatGPT a blessing or blight in disguise? *Medical Education Online*,28:1.

Tyna E, Sam M, Pamela M, Daniel R (2024). GPTs are GPTs: An Early Look at the Labor or Market Impact Potential of Large Language Models.[2024-03-10].*https://arxiv.org/abs/2303.10130.*

Unesco (2024). Launch of Futures of Education report.[2024-03-10].*https://en.unesco.org/futuresofeducation/news/launch*

Wu C, Yin S, Qi W, et al. (2023). Visual chatgpt: Talking, drawing and editing with visual foundation models. *arXiv preprint* arXiv:2303.04671.

Wesley J.W.（2023）.与机器共情：将人类意图与智能工具相结合.中国社会科学报,[2023-0306]. http://sscp.cssn.cn/xkpd/tbch/tebiecehuaneirong/202303/t20230306_5601387.html.

段俊吉(2023).教育数字化转型中的数字代沟：数字反哺的教育镜像.**中国远程教育**,43(03):46-54.

冯志伟,张灯柯,饶高琦(2023).从图灵测试到 ChatGPT——人机对话的里程碑及启示.**语言战略研究**, 8(02):20-24.

焦建利(2023).ChatGPT 助推学校教育数字化转型——人工智能时代学什么与怎么教.**中国远程教育**, 43(04).

卢宇,余京蕾,陈鹏鹤,李沐云(2023).生成式人工智能的教育应用与展望——以 ChatGPT 系统为例.**中国远程教育**,43(04):24-31+51.

骆飞,马雨璇(2023).人工智能生成内容对学术生态的影响与应对——基于 ChatGPT 的讨论与分析.**现代教育技术**,2023,33(06):15-25.

苗逢春(2023).生成式人工智能技术原理及其教育适用性考证.**现代教育技术**,33(11):5-18.

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

沈书生,祝智庭(2023).ChatGPT 类产品:内在机制及其对学习评价的影响.**中国远程教育**,43(04):8-15.

王蕾(2023).人工智能生成内容技术在教育考试中应用探析.**中国考试**, (08):19-27.

张云,何声清(2023).学科教育领域学习进阶的测评研究与应用研究进展.**中国考试**, 371(03):68-76.

张志祯,张玲玲,米天伊,丘诗萍(2023).大型语言模型会催生学校结构性变革吗？——基于 ChatGPT 的前瞻性分析.**中国远程教育**, 43(04):32-41.

钟秉林,尚俊杰,王建华,等(2023).ChatGPT 对教育的挑战(笔谈).**重庆高教研究**,11(03):3-25.

周玲,王烽(2023).生成式人工智能的教育启示：让每个人成为他自己.**中国电化教育**, (05):9-14.

基于大语言模型微调的数学习题认知属性自动标注

Finetuning Large Language Models to Automatically Classify Cognitive Skills in Mathematical Problems

胡航磊[1,*]，姜元昊[2,3,4]，李瑞佳[1]
[1] 教育信息技术学系，华东师范大学
[2] 智能教育实验室，华东师范大学
[3] 上海智能教育研究院，华东师范大学
[4] 计算机科学与技术学院，华东师范大学
[*]51254108023@stu.ecnu.edu.cn

【摘要】 数学在日常生活和许多职业领域中都是必不可少的，因此学生学好数学课程的重要性不言而喻；同时，随着课程改革的陆续推进，亟须一种更为智能地衡量和提升学生数学学习知识与能力的方法与工具。在提高学生数学能力一系列方式，做题毫无疑问是核心之一，但是，只有明确题目锻炼学生何种数学认知能力，才能进行查漏补缺。现有的题目通常直接关联知识点种类，而非认知能力，且大部分都源自于特定规则和专家标注，尽管这种方法可以达到预期目标，但人为标注的高度专业性和主观性使得分类存在一定的局限性。在本研究中，我们通过使用 TIMSS 对数学认知的分类，构建了一个小学数学题与其认知能力的数据集。同时还介绍了一种使用大语言模型对题目进行自动分类的方法，并对三种不同的大语言模型和不同的认知分类标签进行实验评估。我们的发现，对于基础文本分类，二者的分类效果相差并不大；但在多标签场景下，更大参数量的语言模型能够发挥更好的效果。

【关键词】 数学认知；大语言模型；多标签分类

Abstract : *Mathematics is essential in daily life and many professional fields, so the importance of students mastering mathematics courses is self-evident; Meanwhile, with the continuous advancement of curriculum reform, there is an urgent need for a more intelligent method and tool to measure and enhance students' mathematical learning knowledge and abilities. Among the various ways to improve students' mathematical abilities, problem-solving is undoubtedly one of the cores. However, only by clarifying what kind of mathematical cognitive ability the problem exercises can we identify and fill in gaps. The existing problems are usually directly related to the types of knowledge points, rather than cognitive abilities, and most of them originate from specific rules and expert annotations. Although this method can achieve the expected goals, the high professionalism and subjectivity of human annotations make classification somewhat limited. In this study, we constructed a dataset of elementary school math problems and their cognitive abilities by using TIMSS to classify mathematical cognition. At the same time, a method of automatically classifying questions using a large language model was also introduced, and experimental evaluations were conducted on three different large language models and different cognitive classification labels. Our findings indicate that there is not much difference in classification performance between the two for basic text classification. But in multi label scenarios, language models with larger parameter quantities can perform better.*

Keywords: math cognitive skills, large language model, muti-label classification

1. 引言

高质量的数学教育，在这个技术不断突破的时代，对于学生的成长至关重要，尤其是低年龄段的学生。智能辅导系统（Intelligent Tutoring System, ITS）已被证明其在提高数学成绩

方面的有效性(Fancsali and Ritter 2014; Roschelle et al. 2016)。同时，许多 ITS 都会有一个在学生完成练习后向他们提供自动反馈的相关组件。借助这一组件与一些人为制定的规则，教师能够大规模的提供自动化、个性化的反馈，确保不同学业水平的学生能够基于不同的反馈完成相应练习。

在 ITS 中，这种直接、及时的反馈可以高效辅助学习，帮助学生依据自己的需求与节奏逐步提高能力，自动查漏补缺。至于如何生成该类反馈，其首要条件是成体系的题目与相应类别的标签，即定位于每一项练习，明确其隶属的题目类型，考验的知识内容以及提升学生的技能水平类型。以此为依托，系统才能根据学生的答题情况，智能判别学生的知识点掌握程度与技能变化，在提供反馈的同时，根据学生能力的欠缺及时推送相关题目，形成更好的干预。

本研究主要专注于建立题目与对应标签之间的关联，一项与此相关的工作是对数学应用题(Math Word Problem, MWP)进行分类 (Carpenter, Moser, and Hiebert 1981; Heller, Riley, and Greeno 1984; Mitra and Baral 2016; Nesher, Greeno, and Riley 1982; Vergnaud 2020)，其中最主要的一种分类囊括了三个大类型，以及八个子类型——更改(Change)｛结果未知、更改未知、开始未知｝、部分整体(Part-Whole)｛总量未知、部分未知｝和比较(Compare)｛差异未知、大量未知、少量未知｝，并使用基于关键词检索与位置检索的方法来训练机器学习模型，最终目的是评估学生在完成某些数学应用题后的学习能力水平。然而，相比于单纯的数学应用题分类，对于数学认知技能分类的分类难度更高，因为前者主要是依据题目本身属性进行分类，专注于题目内容，例如题目使用何种运算方式、题目如何进行设问等；后者是更为抽象的分类，主要定位的是经过该题训练后，学生何种认知技能能够提升，因此，认知技能的分类不单单需要理解题目内容的相关结构，同时还需要有对如何解决题目有更为深刻的认识。

在方法层面，传统上的题目分类主要依托于专家制定的规则，这些规则高度概括了题目可能出现的各种不同结构与特征，例如，显式的加法关键词：sum, add, join, combine；显式的减法关键词：left, change, rest, remain, empty, out, off, leftover；以及显式的比较类别关键词 some 和 less 等(Mandal and Naskar 2021)。因此，这些系统的创建往往都是资源密集型，且对于设计者无法预见的问题较难产生响应。

相对的，如果需要推广到更加泛化的领域，可能需要更大量的数据或是更详细的规则进行驱动。定位于这一问题，我们尝试结合大语言模型在海量数据中进行无监督预训练后所能发挥出来的涌现能力，再使用带标注的少量认知分类数据进行有监督的微调，实现即使不依靠规则，也能使模型达到较为优秀的分类效果。

2. 方法

在本节中，我们将概述传统的数学题自动分类任务；并说明我们的方法，该方法用大语言模型的微调对数学题进行分类。

2.1 传统文本分类任务

最早期的文本分类往往采用机器学习的方法，例如使用朴素贝叶斯，基于题目中可能出现的数学术语、关键词和语法结构等特征进行分类；使用支持向量机，选择合适的核函数，有效地处理题目的复杂特征，例如数学公式、图形信息和文字描述；使用决策树和随机森林，对特征进行逐步分割，并构建多个决策树确定最终的类别。但是传统的分类方法在数学题分类任务中往往具有一定局限性，主要体现在对复杂特征和长距离依赖关系的处理能力上。

而随着深度学习技术的发展，模型能够更好地捕获题目的语义信息和全局特征，提高分类的准确性和泛化能力。例如使用循环神经网络处理题目中的序列信息；使用卷积神经网络可以从题目中提取数学符号、图形等其他视觉特征；使用 Transformer 模型捕捉题目中的全局依赖关系。但是，现有的方法往往只适用于文本进行单一标签的分类，但在真实场景的应用

中，数学题的复杂性导致其不会只局限于单一标签，因此，对于数学题的分类，使用多标签分类的方法更为合适。

2.2 基于大语言模型的分类方法

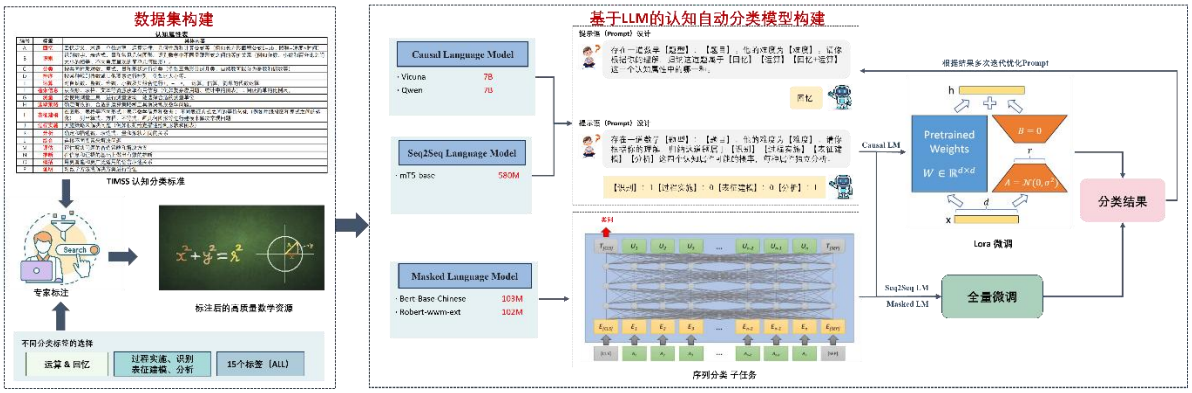


图 1 本研究的方法路线

经典的多标签分类方法是使用 Bert 进行微调。同理，我们的方法也是对大语言模型进行微调，但除了基础的 Bert 模型外，我们还额外选择了序列到序列大语言模型以及因果大语言模型，这些模型相比以 Bert 为代表的遮蔽语言模型，都在更大数据量上进行了预训练，具备更大的参数量，拥有更高的知识储备和能力与对话能力，我们希望模型能够依靠其参数中蕴含的知识，再使用统一格式的数据进行微调明确任务，最终在达到与专家分类相似的效果。

不同于遮蔽语言模型在微调后能够直接输出类型，序列到序列语言模型和因果语言模型的输入和输出都是一段文本序列，因此，在我们预设的场景之下，对于输入，我们设计不同的提示信息，为模型提供除题目内容外更多的参考依据；对于输出，考虑到因果语言模型的本质是预测后一个词出现的概率，因此我们认为，只输出类型的名词并不能拥有较好的效果，相反，以特定自然语言模板形式的输出，可能对于模型的分类而言更为有效。

3. 实验与评估

在本节中，我们详细介绍了实验与评估相关的内容。首先，我们展示了有关数据集构建的细节，对于不同数据量的认知类型，我们采取了多种不同的方式进行分类。然后，我们展示了我们的定量实验，将我们的方法在多种不同的模型上训练并对性能进行了整体的比较。

3.1 数据集细节

在数据集的构造过程中，我们邀请了多位专家使用 TIMSS 中对于学生认知技能的分类标准(Mullis and Martin 2013)进行数据的标注。该数据集具有 27193 道小学数学题，其题型主要包括解答、填空与选择题，每道题均含题号、考察知识点、难度、解析、参考答案以及认知技能类型，其中难度为 1-5 之间的整数，认知技能类型与具体内容见图 2：

编号	标签	具体内容
A	回忆	回忆定义，术语，单位进率，运算定律，几何性质和计算公式等（例如长方形面积公式S=ab；路程=速度×时间）
B	识别	识别数字、表达式、量和常见几何图形。识别数学中不同呈现形式之间的等价关系（例如分数、小数和百分比之间大小的相等；不同角度呈现的简单几何图形）。
C	分类	按共同性质对数、算式、量和形状进行分类（例如三角形分成几类、自然数可以分为奇数和偶数等）
D	排序	按某种规则将数或其他要素进行排列，例如比大小等。
E	运算	对自然数、整数、分数、小数及其组合进行+，−，×，÷运算，估算，简单的代数运算
F	检索信息	从图形、表格、文本等资源获取有用信息（例如复杂应用题、统计中的图表）；阅读简单的比例尺。
G	测量	会使用测量工具，进行测量活动，能选择合适的测量单位
H	选取策略	确定有效的、合适的运算策略和工具解决常规数学问题。
I	表征建模	在图形、表格等不同形式中表示数学信息和数据；不同表征方式之间的等价转化（例如用线段图和算式之间的转化）；列出算式、方程、不等式、画几何图形等进行建模来解决常规问题
J	过程实施	实施策略来解决问题（例如根据给定描述绘制形状和图表）
K	分析	确定和描述数、表达式、量和形状之间的关系
L	综合	连接不同要素来解决问题
M	评估	评估解决问题的备选策略和解决方案
N	推断	在信息和证据的基础上做出有效的推断
O	概括	用更普遍和更广泛适用的语言表述关系
P	证明	对数学方法或解决方案进行论证

图 2 认知技能相关标签与具体描述

经过专家标注后，为了方便模型的训练，又对数据进行二次清洗，主要排除涉及图片（部分题目忽略图片只传输文本会影响解读）与表格的题目，例如“解下列方程”等，最终获得有效数据 22081 条（鉴于证明相关题目均与几何高度相关且涉及图片，因此清洗完成后中【证明】标签的题目数量为 0），其中各类别数据分布如下：

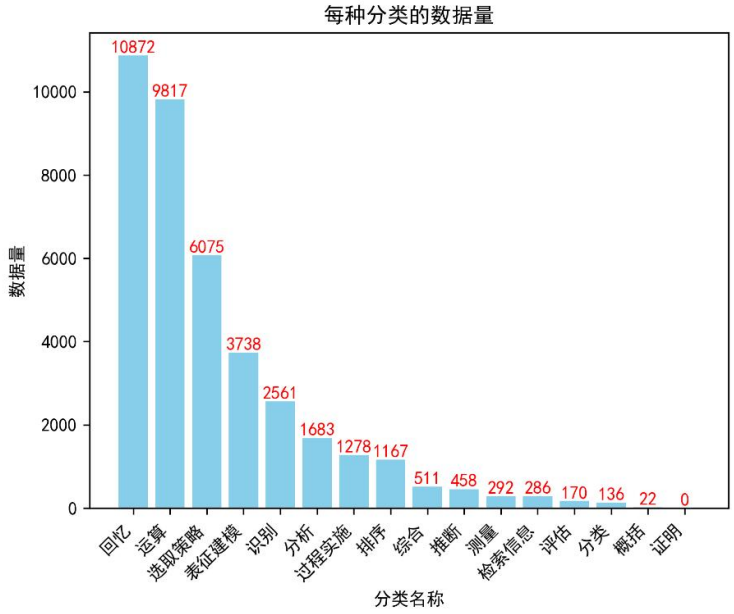


图 3 认知技能相关标签的不同数据分布

此外，由于标签数量分布不均匀，存在严重的长尾现象，我们尝试了三种不同的数据标签选择：

（1）回忆 运算 回忆+运算

共计 19281 条数据，其中仅含有回忆的 7738 条，仅含有运算的 5346 条，同时含有回忆与运算标签的占 6197 条。

（2）识别、过程实施、表征建模、分析

共计 10391 条数据，其中 2561 条含有识别标签，1278 条含有过程实施标签，3738 条含有表征建模标签，1683 条含有分析标签。

（3）除【证明】外的其余 15 类标签

每种标签所含有数据条目量，见图 3。

3.2 指标与基线

对于认知技能的分类，基于不同类型的标签选择，我们的主要指标也不全相同，对于简单的回忆、运算与回忆运算分类，我们的主要指标是分类准确度（Accuracy）；而对于多标签分类任务，我们除了基础的准确度、精确度（Precision）、召回率（Recall）和 F1 外，我们还额外选择了以下评估指标：

汉明损失（Hamming Loss）：衡量模型的错误分类率，即模型在每个标签上错误分类的平均比例。

Jaccard 相似度指数（Jaccard Similarity Score）：衡量模型预测的标签集合与真实标签集合的相似度。

并且，我们主要基于以下三大类模型进行对比分析：

Masked Language Model（遮蔽语言模型）：

遮蔽语言模型在训练时随机遮蔽输入序列的一些标记，然后模型需要预测这些被遮蔽标记的真实值，它的预训练过程使得模型能够捕捉上下文中的双向关系，适用于多种下游任务，如文本分类、命名实体识别等，代表模型：BERT(Kenton and Toutanova 2019), RoBERTa(Liu et al. 2019)。

Seq2Seq Language Model（序列到序列语言模型）：

序列到序列模型包括一个编码器（Encoder）和一个解码器（Decoder），编码器接收输入序列，将其编码成上下文向量；解码器接收上下文向量并生成输出序列。其主要用于序列生成任务，如机器翻译、文本摘要、对话生成等。Seq2Seq 模型在处理不定长的输入和输出序列时表现出色。代表模型：T5(Raffel et al. 2020)。

Causal Language Model（因果语言模型）：

因果语言模型是一种解码器的模型，通常基于 Transformer 架构。在生成过程中，模型逐步生成输出序列的每个元素，每个元素的生成都依赖于已生成的元素。其主要用于生成任务，如文本生成和语言建模。GPT(Radford et al. n.d.)就是因果语言模型的代表。

3.3 实现细节

我们对我们的数据集进行了 9：1 的训练、测试分割，并进行大语言模型的微调。对于不同类型的模型，我们采用了不同的微调方法以及超参数设置，具体细节如下。

表 1 模型选择与超参数具体信息

	Causal Language Model		Seq2Seq Language Model	Masked Language Model	
	Vicuna1.5	Qwen1.5	mt5-base	bert-base -chinese	Robeta-wwm -ext
提出机构	斯坦福	阿里	谷歌	谷歌	哈工大
参数量	7B	7B	580M	103M	102M
Epoch	5	5	8	10	10
Batch_size	64	64	4	16	16
学习率	3e-4	3e-4	3e-4	2e-5	2e-5
微调方法	Lora	Lora	全量	全量	全量
Lora_rank	8	8	-	-	-
（min/epoch）	30	20	8	1	1

首先，所有大模型的微调均使用 HuggingFace(Wolf et al. 2020)的 transformer 库。其次，在微调上，对于 Vicuna1.5 和 Qwen1.5 这两个因果语言模型，由于参数量限制，我们采用了 Lora 的高效参数微调方法，batch_size 为 64，lora_rank 为 8，学习率为 3e-4，均进行了 5 轮的微调；对于 bert 和 roberta，我们选择全量微调，batch_size 为 16，学习率为 2e-5，均进行了 10 轮微调；对于 mt5，我们选择全量微调，batch_size 为 4，学习率为 3e-4，进行了 8 轮微调。以上所有模型的微调均是在一个 NVIDIA RTX 4090 GPU 上进行的。对于每个 epoch，微调所需时间不等，具体取决于所选的大语言模型参数量与训练数据。

此外，在输入输出数据的处理上，鉴于模型对于自然语言处理方面所擅长的子任务不同，我们也进行了相应格式上的修改与优化。

对于 Bert 和 Roberta 这两类遮蔽语言模型，其本身就存在能实现序列分类的子任务，因此，对于标签选择三分类任务，直接转译为 0、1、2 三类，对于多标签分类任务，则转译为每个标签的独立分类，例如[0, 1, 0, 1]。

而对于 T5 等序列到序列语言模型以及 Qwen 等因果语言模型，由于其输入输出均为文本序列，因此我们统一采用设计 prompt 的形式实现分类任务，我们在输入题目内容的基础上又增加了难度与题型，最终整合成一段完整的话语注入模型，而在输出方面，对于三分类任务，我们选择直接输出分类的类型；对于多标签文本分类，我们则选择以每个标签加 0 或 1 的形

式控制模型输出，来让模型预测其是否属于该分类。

表 2 不同模型不同任务标签数据输出格式

模型	分类类型	标签数据输出格式
Masked LM	三分类	0, 1, 2
	多标签文本分类	[1, 0, 0, 0]
Seq2seq LM	三分类	回忆、运算、回忆+运算
	多标签文本分类	【识别】:1【过程实施】:0【表征建模】:0【分析】:0
Causal LM	三分类	运算、回忆、回忆+运算
	多标签文本分类	【识别】:1【过程实施】:0【表征建模】:0【分析】:0

3.4 结果与分析

【回忆】【运算】与【回忆+运算】三个标签的识别结果见表 3，由此可以看出，定位于基础的分类任务，模型之间分类准确率没有显著差距，但是考虑到微调时间与算力成本，可能在单标签分类任务上，依旧是经典的方法更具性价比。

表 3 对于回忆与运算标签进行三分类结果

模型	**类别**	**Precision**	**Recall**	**F1-score**	**Accuracy**
BERT	回忆	72%	67%	69%	80%
	运算	81%	84%	83%	
	回忆+运算	85%	86%	86%	
RoBERTa	回忆	90%	77%	83%	79%
	运算	77%	88%	83%	
	回忆+运算	67%	71%	69%	
T5	回忆	85%	84%	85%	79%
	运算	80%	80%	80%	
	回忆+运算	68%	69%	68%	
Vicuna	回忆	86%	87%	87%	81%
	运算	82%	83%	83%	
	回忆+运算	72%	69%	71%	
Qwen	回忆	88%	87%	88%	**82%**
	运算	84%	82%	83%	
	回忆+运算	71%	74%	72%	

两种多标签文本分类的结果见表 4。从整体结果表现来看，参数量更大的模型在各个指标上都更占优势，特别是在标签数量达到了十五种，各标签数量分布严重不均的情况下，毫无疑问 Bert 等模型的综合指标都是下降的，但是 7B 的模型在各方面指标都占据优势的前提下，甚至还能有所提高。

表 4 四个标签与十五个标签的多标签分类结果

识别、过程实施、表征建模、分析

	Hamming Loss	ACC	Micro				Macro			
			P	R	F1	Jaccard Similarity	P	R	F1	Jaccard Similarity
BERT	**0.12**	67%	80%	79%	80%	66%	80%	76%	78%	65%
RoBERTa	0.15	61%	81%	68%	74%	58%	79%	65%	71%	56%
T5	0.13	67%	79%	77%	78%	64%	77%	75%	76%	62%
Vicuna	**0.12**	**88%**	**88%**	**88%**	**88%**	78%	**86%**	85%	**86%**	75%
Qwen	**0.12**	**88%**	**88%**	**88%**	**88%**	**79%**	**86%**	**86%**	**86%**	**76%**

除【证明】外的 15 类标签

	Hamming Loss	ACC	Micro				Macro			
			P	R	F1	Jaccard Similarity	P	R	F1	Jaccard Similarity
BERT	**0.06**	49%	81%	76%	78%	64%	70%	57%	62%	48%
RoBERTa	0.07	42%	78%	71%	74%	59%	70%	43%	49%	37%
T5	**0.06**	50%	79%	78%	78%	64%	73%	63%	67%	52%
Vicuna	0.07	**93%**	**93%**	**93%**	**93%**	86%	87%	84%	85%	76%
Qwen	0.07	**93%**	**93%**	**93%**	**93%**	**88%**	**88%**	**86%**	**87%**	**78%**

4. 讨论、结论与未来工作

通过以上实验的结果，我们发现，在数据量较为均衡的单标签分类数据集中，7B 参数量的大语言模型使用 Lora 微调的方法，并未存在明显优势；但是在多标签分类的场景下，参数量更大的模型效果普遍更优。然而，在现实场景中，不同认知技能的题目量往往不是我们能够控制的，同时，一个数学题所考察的学生认知技能，也往往是多样的，因此，多标签分类任务才是更为贴近真实环境的，因此，如何结合少量的高质量数学认知技能标注数据，仅依靠模型自身能力，实现较好的分类效果，本身就是一个值得探索的问题。

在本研究的一系列实验中，不管是基础的文本分类还是多标签的文本分类，其最终的分类效果都是因果语言模型占据优势，我们人为，一方面是因为其参数量的提升，不管是知识储备还是计算，都更有优势，另一方面，不同于 Bert 模型只给了题目内容，其他模型我们还额外提供了题型与难度（与认知相关性最高的两个属性），这也为模型分类提供了依据。同时，对于不同数据总量的标签，虽然我们没有逐一分析，但是就整体结果而言，即使存在长尾现象，模型也能够有较好分类效果。

基于以上思考与结论，我们对于提升分类的效果，重点应该在以下两方面：一是提高模型自身参数量，使用全量微调的方法，并在预训练与 prompt 设计的过程中注入更多数学领域知识，即让这个“专家”变得更加专业与聪明；另一方面，对不同认知属性结果逐一定量定性分析，其实在 TIMSS 中将以上十六种认知属性首先划分为了三个维度，我们猜测，不同维度的认知属性或许对于模型分类也会存在一定影响，或许能够基于维度单独设计 prompt，降低模型的分类难度。

参考文献

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2016). Instruction based on adaptive learning technologies. Handbook of research on learning and instruction, 2, 522-560.

Carpenter Thomas P., Moser James M., and Hiebert James. 1981. “Problem structure and first-grade children’s initial solution processes for simple addition and subtraction problems.” Journal for Research in Mathematics Education 12(1):27–39. doi: 10.5951/jresematheduc.12.1.0027.

Fancsali Stephen E., and Ritter Steven. 2014. “Context personalization, preferences, and performance in an intelligent tutoring system for middle school mathematics.” Pp. 73–77 in Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, LAK ’14. New York, NY, USA: Association for Computing Machinery.

Heller Joan, Riley Mary S., and Greeno James G. 1984. Development of children’s problem-solving ability in arithmetic. Academic Press.

Mandal, Sourav, and Sudip Kumar Naskar. 2021. “Classifying and Solving Arithmetic Math Word Problems—an Intelligent Math Solver.” IEEE Transactions on Learning Technologies 14(1):28–41. doi: 10.1109/TLT.2021.3057805.

Mitra, A., & Baral, C. (2016, August). Learning to use formulas to solve simple arithmetic problems. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 2144-2153).

Mullis, I.V.S. & Martin, M.O. (Eds.) (2013). TIMSS 2015 Assessment Frameworks. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2015/frameworks.html

Nesher P., Greeno J. G., and Riley M. S. 1982. “The development of semantic categories for addition and subtraction.” Educational Studies in Mathematics 13(4):373–94. doi: 10.1007/BF00366618.

Roschelle Jeremy, Feng Mingyu, Murphy Robert F., and Mason Craig A. 2016. “Online mathematics homework increases student achievement.” AERA Open 2(4):2332858416673968. doi: 10.1177/2332858416673968.

Vergnaud Gérard. 2020. “A classification of cognitive tasks and operations of thought involved in addition and subtraction problems” edited by Carpenter T. P., Moser J. M., and Romberg T. A. 39–59. doi: 10.1201/9781003046585-4.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. n.d. “Improving Language Understanding by Generative Pre-Training.”

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2019). Huggingface's transformers: State-of-the-art natural language processing. arXiv preprint arXiv:1910.03771.

智能代理驱动的课程智能助教构建研究——以材料力学为例

Research on the Construction of Course Intelligent Teaching Assistants Driven by Intelligent Agents —— Taking Materials Mechanics as an Example

胡震[1]，补熠暄[1]，李艳洁[1*]，段庆龙[2]
[1]北京林业大学 工学院
[2]易鑫集团
[*] liyanjie@bjfu.edu.cn

【摘要】助教需要协助教师解决学生在学习过程中遇到的问题，是教育体系中不可或缺的重要角色。本文将大语言模型、智能代理、知识图谱等智能技术与教育领域的实践经验融合，开发能提供个性化、高效学习支持的智能助教系统。本文对智能助教的设计原理与实现机制进行了探索，并以材料力学课程为例，通过多项实验验证了其在教育实践中的可行性。结果表明，该系统能良好地胜任助教的部分工作，在知识问答场景具有显著优势，解放了教师的时间和精力，帮助学生自主高效地学习，为高等教育的智能化发展提供参考和借鉴。
【关键词】 人工智能教育；个性化学习；大语言模型；智能代理

Abstract: *Teaching assistants need to help teachers solve the problems encountered by students in the learning process, which is an indispensable and important role in the education system. In this paper, the large language model, intelligent agent, knowledge map and other intelligent technologies are integrated with the practical experience in the field of education to develop an intelligent teaching assistant system that can provide personalized and efficient learning support. This paper explores the design principle and implementation mechanism of intelligent teaching assistant, and takes the course of mechanics of materials as an example to verify its feasibility in educational practice through a number of experiments. The results show that the system can be well qualified for part of the work of teaching assistants, and has significant advantages in the knowledge question and answer scene. It frees up teachers' time and energy, helps students learn independently and efficiently, and provides reference for the intelligent development of higher education.*

Keywords: artificial intelligence education, personalized learning, large language model, intelligent agent

1. 引言

全球教育变革的步伐不断加快，提供高质量的数字学习已成为世界多国共识。在此背景下，中国已开启了一场自上而下，落实国家教育数字化战略行动的系统工程。人工智能作为推动新一轮科技革命的关键引擎和人类社会发展的战略技术，在深化教育改革、催化教育新形态方面发挥着重要作用（方海光 等，2022）。人工智能与教育应用深度结合，逐渐形成教育人工智能，推动未来教育的创新发展（祝智庭、韩中美和黄昌勤，2021）。

在课程学习中，助教具有重要的作用。他们能协助教师完成教学任务，如辅导答疑、习题解析等，提高了教学效率，确保了教学质量。相比教师，助教更贴近学生，能够更好了解学生的学习需求，提供个性化的指导和帮助。然而，人类助教精力和时间有限，不能及时回应每位学生的问题。同时，人类助教的知识储备有限，需要不断学习和提升，增加了教育成本。随着数字化教学不断普及，使用智能化技术构建智能助教是当前亟待解决的重要课题。

2022 年，以 ChatGPT 为代表的大语言模型在语言理解和生成方面的强大能力吸引了各方的广泛关注。大语言模型赋能教育领域，助力教学数字化变革成为当下的焦点（卢宇 等，2023b）。为了有效提高教师的教学效果、学生的学习效率，同时降低教育成本，拓展教育资源，本文将大语言模型与智能助教相结合，同时借助知识图谱，智能代理等自然语言处理技术，研究能提供个性化、高效学习支持的智能助教系统。

2. 相关研究

智能助教（张旭 等，2023）指基于人工智能技术实现的智能辅助教学系统或机器人，无需人类教师的直接参与，它能够胜任重复、繁琐的教学任务，同时还能与人类教师协作开展教学活动，实现优势互补。智能助教可以提供 24/7 的即时答疑服务，帮助学生节省查找资料的时间，让教师有更多的精力关注课程设计。学生的学习需求和进度各不相同，智能助教可以根据学生的学习历史和行为，提供个性化的问题解答和学习建议，帮助学生更有效地学习。

传统的智能助教以功能划分为主，大部分用于响应学生学习过程中的常见问题，本质是一种问答系统，通常基于检索式编码模型实现，通过向量空间内的语义相似度计算，召回相近问题的答案。这种方式构建的智能助教适用场景有限，不仅需要人工维护问答对知识库，在面对复杂的学生问题时，由于语义表征的不充分性导致其匹配效果不够理想。借助人工智能技术进行学生意图识别，根据不同意图进行不同的处理与响应是构建智能助教的常见方法，但传统深度学习模型需要大量的训练数据，且训练完成后不易扩展，交互逻辑较为死板。大语言模型拥有很强的语义理解和文字生成能力，基于大语言模型构建智能助教具有更高的智能化。不同课程的学习内容和学习方法有所不同，以课程为单位对智能助教进行设计和管理，更符合教学实际。本文主要研究大模型驱动的，可应用于不同教学场景的课程智能助教。

2.1. 大语言模型

大语言模型属于生成式人工智能（Generative Artificial Intelligence），主要以 Transformer 为系统架构，在海量无标签语料数据集上进行自监督或无监督训练，模型参数较大，能很好地适配各种下游任务（卢宇 等，2023a）。

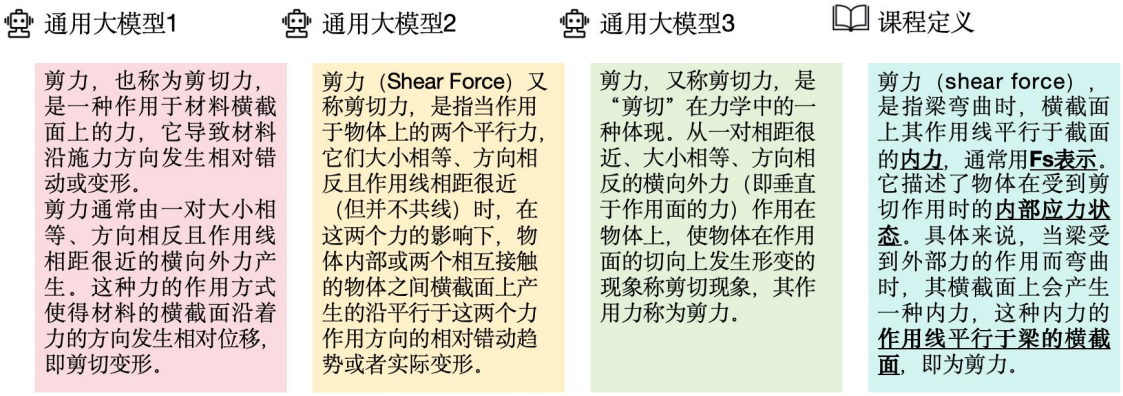


图 1 通用大模型回答“什么是剪力”

由于训练数据和训练过程的影响，通用大模型在教育应用中存在显著的幻觉和逻辑缺失等问题（余胜泉和熊莎莎，2024）。简单地基于大模型构建的智能助教可用性有限，存在诸多风险。其可能会向学生传递错误的知识，降低了智能助教的可信度。通用大模型在知识问答过程中，往往会生成宽泛的回答，不能严格从定义、原理等角度深入解析问题，见图 1。智能助教不仅需要对知识问答进行响应，还需对相关的习题进行推荐，帮助学生巩固知识点。

然而通用大模型对课程习题的记忆量有限，不能针对具体知识点生成多样且合理的习题。以材料力学为例的部分课程，包含大量的多模态信息，如机械结构图、实验原理图等。大语言模型还不具备多模态回复的能力，需要借助其他智能技术进行增强。

2.2. 智能代理

智能代理（Xi et al., 2023）是基于大语言模型的新型人工智能技术，是一种能够感知环境、进行决策和执行动作的智能实体。智能代理能够自主制定和调整工作计划，与外界实体进行交互，根据环境感知做出相应反应，并始终以实现目标为导向。

智能代理为教育带来了前所未有的智能化变革。大语言模型作为智能代理的“大脑”，可以充分理解学生在学习中的不同诉求，并针对诉求自主制定完善的解决方案，最终通过使用丰富的教学工具（乐惠骁和汪琼，2022）完成回复。在大模型思考诉求和制定方案的过程中，可以借助 COT 等（Wei et al., 2022）方法，使大模型模仿人类思考问题的方式，减少幻觉和错误率，保证方案的可行性。通过不断丰富工具集合，可以使智能助教胜任更多的教学任务，在不同的教学场景中发挥作用，从而实现教育教学的智能化迭代和可持续发展。

2.3. 知识图谱

知识图谱是由节点和边构成的图形数据库，将复杂的知识领域借助数据挖掘、知识计量、信息处理和图形绘制进行显示。知识图谱中，节点代表实体，边表示实体间的关系，这种结构化的表示形式极大地提高了计算机对信息的理解、处理与检索能力。

将大语言模型和知识图谱协同使用（Pan et al., 2024），可以增强事实知识，有效缓解大模型的幻觉问题，同时让答案生成具有更好的可解释性（江波，丁莹雯和魏雨昂，2023a）。大模型生成答案时，先在知识谱图中检索出相关信息，将这些信息作为先验知识和学生问题一起输入大模型中，限制大模型的生成范围，有效防止了大模型错误回答、发散回答。通过搭建学科知识图谱（周东岱 等，2024），不同学科的智能助教可以依赖对应学科的专有知识图谱进行知识检索和答案生成，提高了课程智能助教的专业性和可用性。随着知识图谱的不断完善，智能助教能自如应对知识的更新，不需要反复训练，具有良好的扩展性。

3. 智能助教系统的设计与构建

本文研发的智能助教系统由三部分构成：决策大脑、记忆中心和工具集合，详见图 2a。

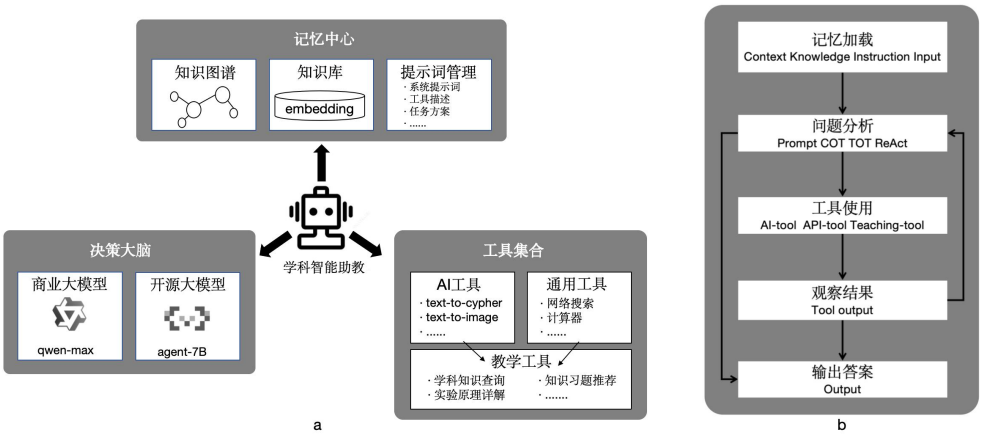


图 2 智能助教架构与工作流程

3.1. 系统工作流程

收到学生的提问后，智能助教先进行记忆加载，并分析学生问题制定执行方案。对于简单的问题，智能助教直接生成答案并输出，当需要使用工具时，智能助教选择合适的教学工

具，根据工具所需参数从学生的输入问题中抽取对应的信息，进行调用。工具成功处理完成后，智能助教观察工具返回的结果并进行下一步规划，思考是否需要使用其他工具，或直接生成答案。在工作中，智能助教保持思考（Thinking）-执行（Action）-观察（Observation）三种状态不断循环，直至解决学生问题，如图 2b 所示。

3.2. 记忆中心

记忆中心用于存储和管理智能助教工作所需的记忆信息，分为长期记忆和短期记忆两种，长期记忆主要包括知识图谱和知识库等外部知识，短期记忆指学生和智能助教对话的上下文历史（Context），以及指令提示（Prompt）等。

3.2.1. 长期记忆

智能助教通过知识库和知识图谱引入外部知识，作为长期记忆对通用大模型进行学科知识增强。知识库主要包含材料力学课程教材和相关书籍等。针对课程教材，本文按照书中的知识结构生成关键词索引，在同一知识关键词下，以自然段落为单位对所属的文本进行切分，尽量保证知识信息的完整性，对其他书籍或文档，直接按照段落进行切分。最后使用编码模型对获得的所有知识片段向量化，以知识向量的形式存储在向量数据库中。本文借用 ChatGPT 大语言模型，基于教材、教学课件、教学大纲等多种教学资源，自动化抽取知识实体和知识关系（唐晓晟 等，2024），并以《材料力学》教材的层级结构为主要脉络，自顶向下不断细化知识元（袁满, 仇婷婷和胡超，2019）。同时，本文将材料力学课程习题按照题目类型分类，设置习题节点，与图谱中最细粒度的知识节点进行连接，最终通过教师人工校验和完善，形成材料力学学科知识图谱。

3.2.2. 短期记忆

上下文历史用于帮助大模型获取多轮对话信息，与提示词一起作为短期记忆输入到大模型中。提示词通过向大模型描述助教的角色信息、可供使用的教学工具、以及一些其他的生成条件，从而限制大模型的生成内容和格式。本文通过提示词告诉大模型正在扮演材料力学课程助教，加强大模型的自我认知，使其按照人类助教的思维方式进行推理，提高了学生与智能助教系统的交互体验。工具描述影响着智能助教对教学工具的调用，本文对智能助教所有工具的功能和调用参数设计了详细的描述提示词，保证了教学工具使用的成功率。

3.3. 决策大脑

智能代理驱动的智能助教系统以大语言模型作为决策大脑。本文选择了在中文场景更有优势的阿里巴巴 Qwen-max 大模型。Qwen-max 有强大的基础语义理解能力，能够良好地完成对学生意图的识别和关键信息的抽取，同时具有优秀的推理能力，能够针对学生不同的问题制定对应的解决方案，在反复思考和观察中精准提炼出问题的答案。千亿级的参数使得 Qwen-max 有更强的指令跟随和工具调用能力，在对话过程中能够始终铭记自己助教身份，不仅能根据需要自主选择工具，还能按照工具要求的格式正确生成调用。本文使用 ReAct 方法（Wei et al., 2022），要求大模型在选择工具时给出明确的选择理由，使得智能助教能够动态创建、维护和调整行动计划，并将工具返回结果整合到推理中。不仅提升了大模型解决问题的能力，同时增加了思考过程的可解释性和可信度（江波, 丁莹雯和魏雨昂，2023b）。

同时，本文选择了参数量较小的开源模型用于本地化部署方案。ModelScope-Agent-7B 是在阿里巴巴 Qwen-7B-Chat 模型的基础上，使用 MSAgent-Bench 数据集（Li et al., 2023）训练增强的 Agent 模型，相较于同参数量的其他开源模型，有更好的工具调用能力。

3.4. 工具集合

智能助教完成学生问题的响应需要用到各种不同的工具，由工具集合定义和管理。工具集合中包含 AI 工具、通用工具和教学工具。AI 工具是由一系列人工智能模型实现的，本文使用的 AI 工具包括文本翻译、查询语句生成等，以及文本转语音、文生图等多模态模型，使

智能助教支持语音、图片等多模态信息交互能力。通用工具主要是一些程序组件或其他第三方 API 接口，除常用的天气查询组件，计算器组件外，本文主要使用了网络搜索 API，用于对知识库外的知识进行检索，给大模型提供更多的知识信息。教学工具是在 AI 工具和通用工具的基础上实现的，通常由一系列基础工具组成，共同完成某一教学任务。

本文针对材料力学课程学习中学生最常见的问题，设计了学科知识查询工具、知识习题推荐工具、实验原理详解工具三种教学工具。其中，学科知识查询工具和知识习题推荐工具主要借助知识图谱进行知识和习题检索，实验原理详解工具用于对课程涉及的核心实验进行解答，基于实验知识库进行匹配召回。当知识、习题和实验涉及到图片，工具调用会返回图片存储的地址，并通过提示词让大模型在生成答案时保留图片的地址。教学工具可以根据教学目标不断扩充，以适应各类教学场景。在定义教学工具的过程中，需要对工具有清晰的功能描述和明确的参数要求，以提高智能助教对教学工具的正确使用。



图 3 智能助教效果展示

4. 智能助教可行性验证

4.1. 数据来源

为反映智能助教在使用过程中的真实情况，本文选取校内材料力学课程班级中的学生进行实验，共计 62 名同学（男 37 人，女 25 人）。实验分为 A，B 两组，每组随机分配 31 名同学，A 组同学在课程学习过程中使用智能助教，B 组同学使用通用大模型，实验持续一个月。本实验采用问卷调查法，通过“问卷星”电子问卷的形式进行各组的使用反馈数据采集。

本文还针对智能助教的答案回复内容和教学工具使用情况进行了实验，使用了 MMQA（Material Mechanics Question and Answer）和 MMTAGENT（Material Mechanics Teaching AGENT）两个数据集，数据样例见表 1。MMQA 数据集是由 257 条材料力学领域的知识问题构成，部分来自于教学过程中学生高频提问的问题，共计 57 条，其余 200 条数据借助大语言模型针对材料力学的重要知识点生成，基本能覆盖本科教学阶段材料力学课程主干知识。MMQA 数据集用于测评智能助教的知识回答效果，验证在外部学科知识的增强下，大模型生成内容的正确性和专业性。MMTAGENT 包含了 500 条教学工具调用数据，标签为 4 类，分别是知识查询 300 条，习题推荐 100 条、实验详解和其他类各 50 条，均由大语言模型生成。MMTAGENT 数据集用于评估智能助教使用教学工具解决问题的能力，验证智能代理驱动智能助教的可行性。本文在生成数据的过程中选择了和决策大脑不同分布的大模型 ChatGPT，同时对数据进行人工校验和筛选，以保证实验的有效性和泛化性。

表 1 实验数据集样例

数据集	数据样本	标签
MMQA	如何计算斜截面上应力？	/
MMTAGENT	扭矩的定义是什么？	知识查询
MMTAGENT	推荐两道压杆稳定相关的计算题。	习题推荐
MMTAGENT	低碳钢和铸铁拉伸实验中“力-变形的关系曲线”有何不同？	实验详解
MMTAGENT	树上有三只鸟，打死一只还剩几只？	其他

4.2. 实验设置

在调查问卷中，本文从生成答案的正确性、流畅性、专业性，以及情感陪护、是否愿意替代人类助教等多维度对智能助教进行评估。正确性指生成答案是否真实，不存在谬误。流畅性指生成答案是否通顺，前后有无矛盾。专业性指答案是否属于材料力学学科范畴，要精准不能过于宽泛。情感陪护关注学生在使用过程中的情感体验，智能助教不仅是一种辅助学习系统，更希望它可以在与学生交互的过程中扮演好助教的角色，与学生互动更自然并提供个性化的回复。以上四个维度由学生在 0-10 打分。是否愿意替代人类助教能够反映学生使用辅助学习系统的整体感受，有是与否两个选项。

在 MMQA 数据集上，本文对比了通用大模型 Qwen-max 和基于该大模型实现的材料力学智能助教的效果，每条知识问题同时由通用大模型和智能助教进行回答，由两位材料力学教师独立选择较好的答案，分为智能助教优、通用大模型优、两种答案均可和两种均不可四种情况。在 MMTAGENT 数据集上，本文选择了 Qwen-max、讯飞星火商业大模型以及 ModelScope-Agent-7B 开源大模型，通过工具选择情况对比不同大模型对学生意图的理解及思考规划能力，采用的评价指标为准确率（Accuracy），即选择工具正确的样本数量占总量的百分比。智能助教正确选择教学工具后能顺利得到工具返回的结果，依赖于大模型对工具调用所需参数的抽取能力，因此，本文同时统计了工具成功调用的样本数量占总量的百分比。

4.3. 实验结果与分析

4.3.1. 使用问卷调查

研究统计了 A、B 两组同学的使用情况，其中正确性、流畅性、专业性和情感陪护四个维度分别取组内均值，是否代替人类助教计算是或否占组内百分比，具体结果如表 2 所示。生成答案无论是正确性、流畅性还是专业性，A 组均高于 B 组。情感陪护方面，A 组也优于 B 组。A 组有 48%（15/31）的同学选择替代人类助教，对智能助教的使用意愿相对较高。B 组同学显著倾向于人类助教，仅有 29%（9/31）的同学选择替代。

实验证明，在知识问答场景中，智能助教已经能很好地满足学生的需要，相较于通用大模型，智能助教生成的答案更加优质。在情感陪护方面，智能助教未能达到人类助教的水平，大语言模型有丰富的通用知识和语言能力，但缺乏人类的情感理解和共情能力。目前，智能助教已初步胜任部分助教工作，但距离替代人类助教还有一定差距。

表 2 问卷数据统计

组别	正确性	流畅性	专业性	情感陪护	是否替代人类助教
A 组	8	8	7	5	48%
B 组	5	6	3	2	29%

4.3.2. 答案生成评估

在 MMQA 数据集上，两位材料力学教师对智能助教和通用大模型生成结果的选择如图 4 所示，在 257 条知识问题中，教师 1 有 64%倾向于智能助教的生成结果，13%持平，9%选择通用大模型，14%两者均不选，教师 2 有 67%倾向于智能助教，12%持平，10%选择通用大模型，11%两者均不选。实验证明，相较于通用大模型，智能助教生成的答案更符合教师的期望。总体上，两位教师认为智能助教生成答案的准确性基本一致，分别为 77%和 79%，说明智能助教已经能够良好解决学生在学习过程中的知识问题。

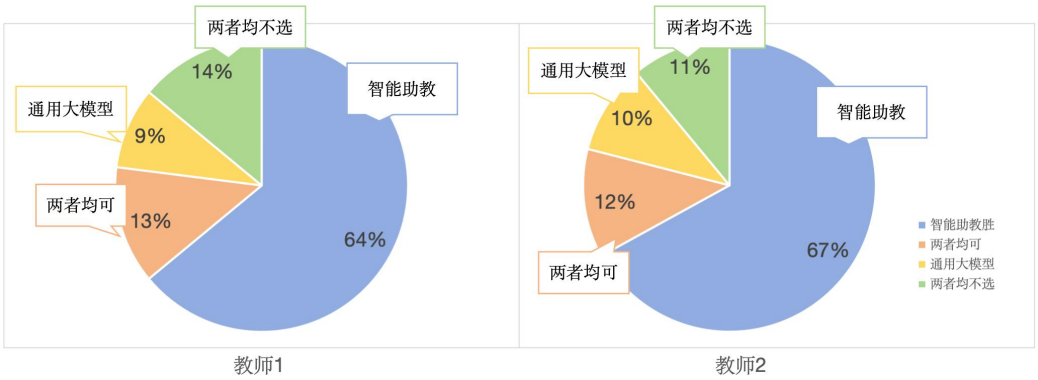


图 4 答案生成评估

4.3.3. 工具使用实验

MMTAGENT 数据集上的实验结果如图 5 所示，Qwen-max、讯飞星火选择教学工具的准确率分别为 98.3%、97.9%，ModelScope-Agent-7B 的准确率为 95.4%。在使用教学工具成功率方面 Qwen-max、讯飞星火分别为 91.5%、89.7%，ModelScope-Agent-7B 的成功率为 81.3%。三个大模型在教学工具选择中均有较高的准确率，最大差值为 2.9%，但使用成功率不足，其中 7B 开源模型较商业模型有较大差距，比 Qwen-max 降低了 10.2%。实验证明，参数量较小的模型已基本能够准确识别学生的意图，判断是否使用工具及使用工具的类型，但在参数抽取和工具调用格式方面还有较大提升空间，需要通过微调或其他手段加强模型本身的抽取能力和指令跟随能力。对比可知，相较于 7B 量级的开源大模型，参数量更大的商业模型具有更好的性能。

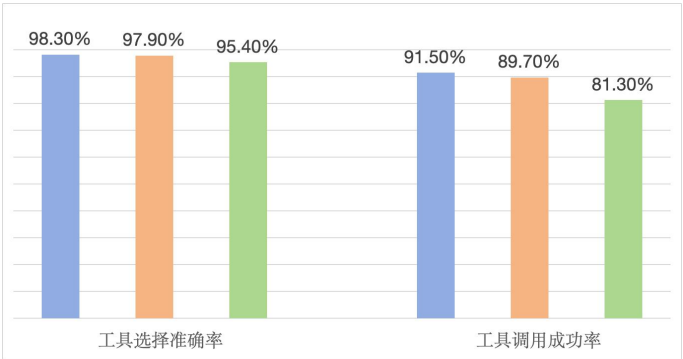


图 5 教学工具使用评估

5. 结束语

融合大语言模型、智能代理、知识图谱等多种智能技术的学科智能助教，为学生提供不限时空的辅助学习服务，为高等教育带来了更加智能化、个性化和高效的学习支持。本文以

材料力学为例，通过多项实验证明了其在教学实践中的可行性。通过对教学工具的不断完善，智能助教系统将支持更多功能，适应更多教学场景。在未来，智能助教不仅是学生学习的好帮手，同时也能协助老师更好地完成教学任务，如帮助教师智能备课、跟踪学习进度等。

虽然大语言模型具有强大的自然语言理解和生成能力，但在某些特定的、复杂的语言环境中，其表现不如人类助教准确和灵活。其次，大语言模型缺乏人类的情感理解和共情能力。目前，智能代理驱动的智能助教仍然无法完全替代人类助教。两者各有优劣，可以相互补充，共同提高教学效果和学生的学习体验。

参考文献

方海光、 孔新梅、 李海芸和郑志宏（2022）。 人工智能时代的人机协同教育理论研究。 **现代教育技术 (07)**， 5-13。

江波、 丁莹雯和魏雨昂（2023a）。 可解释学习者模型:可信个性化学习的技术关键。 **远程教育杂志 (02)**， 48-57。

江波、 丁莹雯和魏雨昂（2023b）。 教育数字化转型的核心技术引擎：可信教育人工智能。 **华东师范大学学报(教育科学版) (03)**， 52-61。

乐惠骁和汪琼（2022）。 人机协作教学：冲突、动机与改进。 **开放教育研究 (06)**， 20-26。

卢宇、 余京蕾、 陈鹏鹤和余胜泉（2023a）。 多模态大模型的教育应用研究与展望。 **电化教育研究 (06)**， 38-44。

卢宇、 余京蕾、 陈鹏鹤和李沐云（2023b）。 生成式人工智能的教育应用与展望——以 ChatGPT 系统为例。 **中国远程教育 (04)**， 24-31+51。

唐晓晟、 程琳雅、 张春红和朱新宁（2024）。 大语言模型在学科知识图谱自动化构建上的应用。**北京邮电大学学报(社会科学版) (01)**， 125-136。

余胜泉和熊莎莎（2024）。 基于大模型增强的通用人工智能教师架构。 **开放教育研究 (01)**，33-43。

袁满、 仇婷婷和胡超（2019）。 细粒度课程知识元组织模型及知识图谱实现。 **吉林大学学报(信息科学版) (05)**， 526-532。

张旭、 满文琪、 孔新梅和方海光（2023）。 人工智能助教机器人中小学课堂应用模式及效果评价模型构建研究。 **中国现代教育装备 (24)**， 54-57。

周东岱、 董晓晓和顾恒年(2024)。 教育领域知识图谱研究新趋向：学科教学图谱。 **电化教育研究 (02)**， 91-97+120。

祝智庭、 韩中美和黄昌勤(2021)。 教育人工智能(eAI)：人本人工智能的新范式。 **电化教育研究 (01)**， 5-15。

Li, C., Chen, H., Yan, M., et al. (2023). Modelscope-agent: Building your customizable agent system with open-source large language models. arXiv preprint arXiv:2309.00986.

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering.*

Wei, J., Wang, X., Schuurmans, D., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.

Xi, Z., Chen, W., Guo, X., et al. (2023). The rise and potential of large language model based agents: A survey. arXiv preprint arXiv:2309.07864.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:2210.03629.

C-STEAM 教育模式下文化统整课程设计研究

Research on the Design of Cultural Integration Curriculum Under the C-STEAM Education Mode

詹梓灵[1]，王炜[1*]，崔萌[1]
[1]广东第二师范学院 教师教育学院
*345979411@qq.com

【摘要】 在“双减”背景下，跨学科主题学习是落实“立德树人”的一个重要切入点。结合广州市天河区非物质文化遗产“押花”传统手工技艺，文章以“学科统整+文化”为研究导向，设计出 C-STEAM 主题课程——《百变押花秀》，将多学科和文化教育落实到教与学过程中，充分发挥区域文化赋能教育高质量发展，培养“知国风、明国事、扬国情”的新时代青年人才。
【关键词】 学科统整；课程设计；本土文化；校本课程

Abstract: *In the context of " double reduction, " interdisciplinary thematic learning is an important entry point for the implementation of " moral education. " Based on the traditional craftsmanship of the intangible cultural heritage ' Yahua ' in Tianhe District of Guangzhou City, this paper takes ' discipline integration + culture ' as the research orientation, and designs the C-STEAM theme course ' Changeful Yahua Show ', which integrates multidisciplinary and cultural education into the teaching and learning process, gives full play to the high-quality development of regional culture-enabled education, and cultivates young talents in the new era of ' knowing the national style and national conditions '.*

Keywords: Integration of Disciplines, Curriculum Design, Local Culture, School-based Curriculum

1. 问题提出

近年来，随着科学技术的快速发展和人才培养模式的转变，STEAM 教育逐渐成为了教育领域的研究热点。在《义务教育信息科技课程标准（2022年版）》中，每一学段的课程内容都增设了跨学科主题的相关教学指导（中华人民共和国教育部，2022）。可见，跨学科主题教学已成为新一轮课程改革的风向标，基础教育不仅是信息科技课程，在其他学科课程中，跨学科主题教学也是课程改革研究的主要方向。“双减”背景下（中共中央办公厅、国务院，2021），为落实“立德树人”教育根本任务，结合《中华优秀传统文化进中小学课程教材指南》中提出的传统文化课程育人体系建设是落实“立德树人”的重要切入点，也是培养核心素养的有效途径，更是增强民族文化自信的关键载体（中华人民共和国教育部，2021）。基于此，本文提出了 C-STEAM 教育的顶层设计，以广州市天河区非物质文化遗产“押花”传统技艺为项目主题，有效落实项目课程的教与学并迭代发展为特色校本课程，推动创新人才的培养和传统文化的传承。

2. 文献研究现状述评

文献的聚类分析可以将相关的文献聚集在一起，形成一个或多个聚类，每个聚类代表一个主题或领域，有助于深入了解某个领域的研究现状和趋势。本研究通过在 CNKI 期刊库进行高级检索，输入关键词：“C-STEAM”、“课程设计”、“文化本位”，检索文献发表日期为 2020 年 1 月至 2024 年 1 月，通过分析主要主题，最后筛选出 179 篇相关主题文献，借助 Citespace 软件对文献的关键词进行聚类分析并生成可视化图表，如相关关键词出现的频率,时间上的热点词变动等等，从而了解研究领域的前沿动态，挖掘其中的知识结构和内在联系，预测未来的发展趋势和研究方向。

在 Citespace 中，Q 值（模块度）和 S 值（平均轮廓系数）是用来评估图谱的聚类效果和节点间关系的指标，Q 值大于 0.5，S 值大于 0.7，通常被认为是良好的聚类效果。而本次文献研究的相关指标系数为 Q=0.5092,S=0.8862，说明生成的图谱能够很好地反映数据之间的聚类关系，聚类之间有明显的区分度，且聚类内部的关系较为紧密。结合关键词共现图谱结果和相关文献，发现近五年的文献研究大多聚焦于课程设计、校本课程、项目式学习、STEAM 教育、核心素养等主题，强调文化素养、劳动教育、6C 模式、计算思维、创客教育等理念。关键词共现图谱如下图 1 所示。

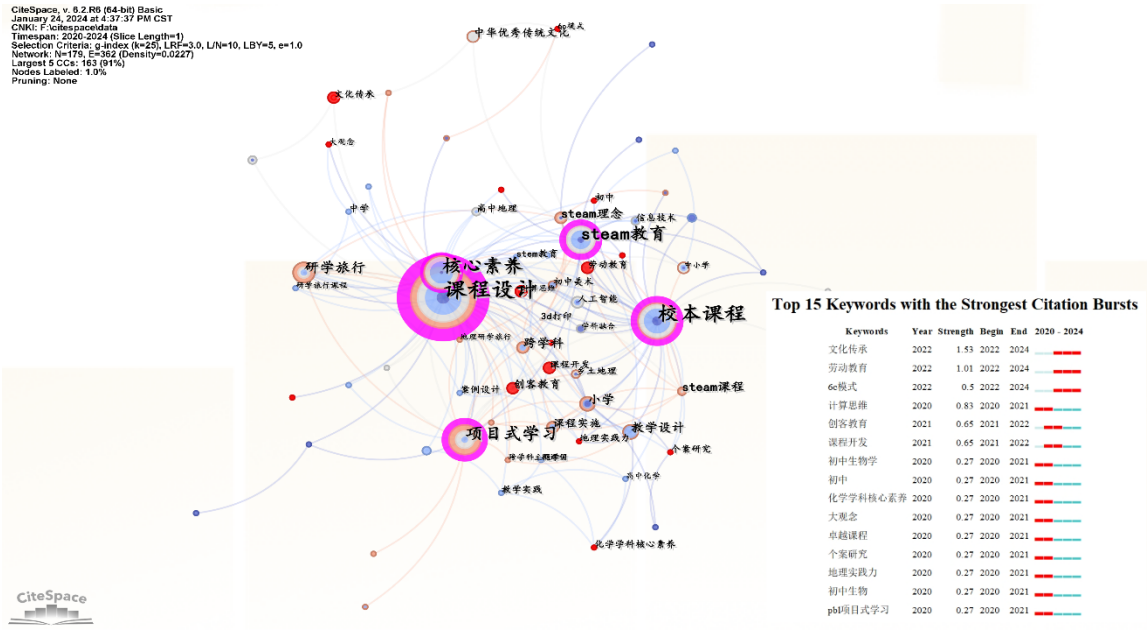


图 1 研究热点共现关系图

此外，通过分析文献关键词的时间轴分布情况，如下图 2 所示，发现近五年 C-STEAM 研究领域前期侧重于理论方面的研究，后期主要进行课程开发和实践研究，而如何在中小学有效开展基于文化本位的跨学科教学是 C-STEAM 教育研究出发点及落脚点。通过分析时间轴上的热点分布，可见中华优秀传统文化、核心素养、校本课程和项目式学习是 C-STEAM 教育研究的重中之重。

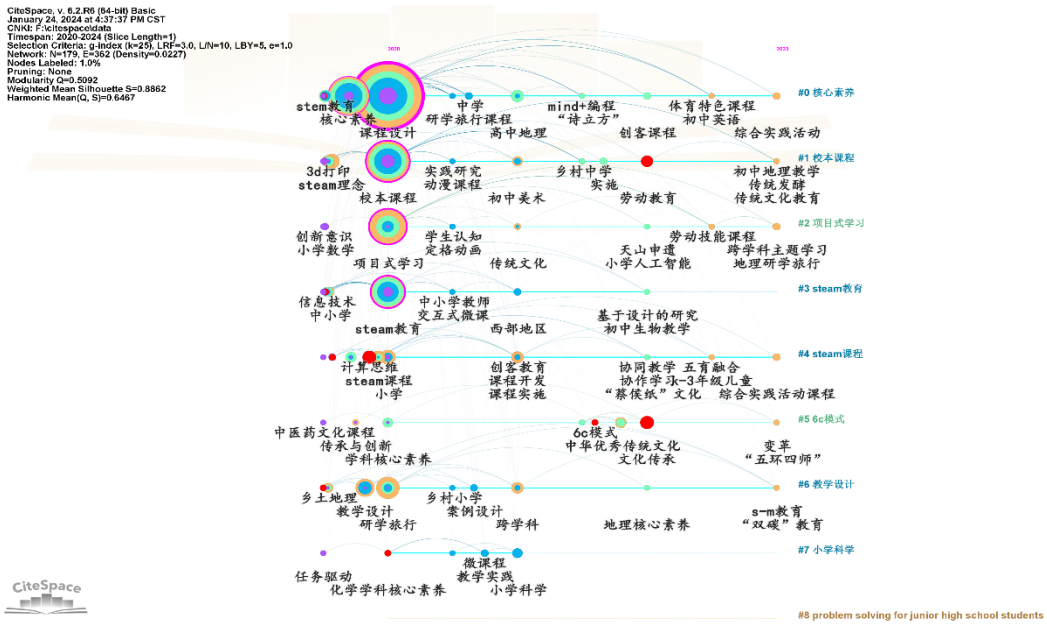


图 2 研究热点时间变动图

近几年，国内外迎来了跨学科融合的 STEAM 教育热潮，美国、加拿大、澳大利亚等发达国家把 STEAM 教育纳入国家课程体系中（Bybee R. W.，2013）（English L. D.，2016）。2022 年，我国教育部在新版义务教育课程方案和课程标准中，明确提出各学科要安排 10%的学时用于开展跨学科的主题学习活动（中华人民共和国教育部，2022）。可见，STEAM 教育深受我国教育界的重视。但是，国内一直缺乏对 STEAM 教育进行本土化重构。2019 年，一些学者为凸显 STEAM 中的“A”，提出人文引领下的学科统整教育（A-STEM）(詹泽慧、李克东和林芷华等, 2020)。然而 A-STEM 只强调人文的基本元素，并未确切落实到文化传承的终极目标上(詹泽慧、钟柏昌和霍丽名等, 2020)。为解决这一研究问题，2020 年，詹泽慧等学者((詹泽慧、李克东和林芷华等, 2020)结合中国优秀传统文化发展的国情，研究出与中国具体国情相适配的 STEAM 教育模式 (占位符 1)——面向文化传承的学科统整教育，即 C-STEAM 教育(詹泽慧、李彦刚和高瑞翔等, 2022)，强调在培养学生的计算思维、创新意识和人文素养等跨学科核心素养的同时(董黎明和焦宝聪, 2020)，注重创设以优秀传统文化为主题背景的教学情境，提升学生对文化的感知力和创造力。然而，C-STEAM 教育研究领域也面临着迫切需要解决的难题——如何更好地将文化元素融入到 C-STEAM 教育中，如何更好地培养学生的文化意识和跨文化交流能力，如何有效开展 C-STEAM 主题课程教学等(袁磊、张瑾和滕洁梅等, 2022)。

综上所述，不同时期、不同国家的 STEAM 教育的培养目标可能会有很大不同，而国内文化本位的 C-STEAM 教育仍存在一些问题需要进一步探讨和研究(詹泽慧和曾筱洵, 2022)。基于此，依托詹泽慧教授提出的面向文化传承的 C-STEAM 教育模式(詹泽慧、李克东和林芷华等, 2020)，本研究基于“双减”背景和新课标要求，以“立德树人”为核心目标，结合相关政策引领及区域文化赋能，开发具有校本特色的 C-STEAM 主题课程，以实现教育助推文化传承，为培养全面发展的新时代人才增添文化教育底色。

3. 统整项目课程设计

3.1 课程背景概述

本研究依托所选取的学校具备“广州高水平美育团队”办学特色以及深入挖掘区域本土文化，最终选取“天河区押花非遗传统技艺”作为项目课程主题。押花的制作工序并不复杂，首先，收集天然植物花卉并进行干燥处理；然后，进行巧妙构思将处理过的花卉通过拼、摆、剪等方式组建成一种图形或一幅画；最后，将制作出的押花画进行封装保存。作为广州市非物质文化遗产之一，天河区押花非遗传统技艺不仅展现出艺术价值，还承载着教育价值，对学生思维发展和美育的培养都具有重要的启示作用(许等, 2020)。基于此，文章借鉴詹泽慧教授等提出的 C-STEAM 理念下的 6C 模型设计及相关案例(詹等, 2023)，结合天河区押花非遗技艺，以《百变押花秀》课程为例，搭建课程框架并彰显课程育人价值(詹, 钟, 等, 2020)，具体教学框架内容如下图 3。

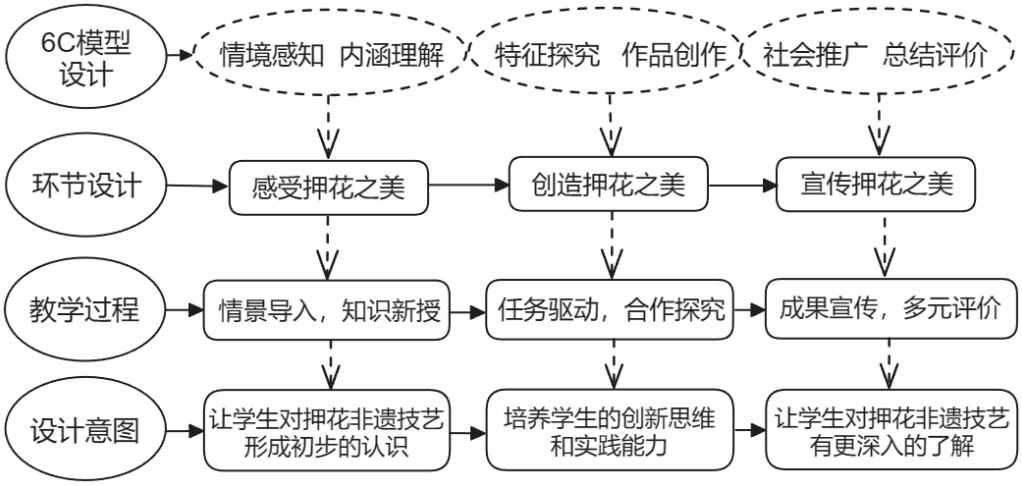


图 3 基于 6C 模型设计的统整项目课程教学框架

3.2 学科统整设计

以《百变押花秀》C-STEAM 主题课程为例，本次主题学习项目融合四个学科进行设计，分别是信息科技、美术、科学、语文。课程以认识押花非遗技艺、制作电子押花名片为主线，其中，信息科技为本课程提供资源保障，如在信息化教学环境下，教师通过数字化教学资源进行授课并提供资源让学生利用演示文档制作电子押花名片；美育贯穿在整个教学环节之中，让学生通过观看视频、欣赏作品去发现、感受押花之美；科学的学习则是蕴涵在教师提供的植物卡片拓展学习资料包中，学生可以边实践边认识对应植物的名称；语文学科的内容要求在学生了解押花非遗技艺之后，构思并撰写一份宣传讲稿，在作品展示环节对自己的押花名片进行宣传。学科统整情况如表 2 所示。

表 1 《百变押花秀》C-STEAM 课程学科统整

统整学科	教学内容	教学目标
信息科技	➢ 运用演示文档工具制作电子	➢ 让学生学会图层的排列和组合的具体操作步骤 ➢ 通过制作电子押花，培养学生运用信息化工具解

	押花	决问题的创新意识和实践能力
	➢ 对文字和图片进行排版，形成宣传名片	➢ 通过任务驱动学习，提升学生的独立思考和动手能力
美术	艺术欣赏与画面设计	通过播放押花宣传视频和展示押花工艺品，让学生感受押花之美，为创造美和宣传美做铺垫
科学	认识植物名称	提供植物学习电子卡片，学生不仅能利于植物卡片制作电子押花，还能在制作作品过程中认识植物的名称
语文	撰写宣传讲稿	结合对押花的认识，在作品展示环节对自己的押花名片进行宣传

4. 统整项目课程实施

4.1 课程实施情况

2023 年 12 月下旬，本次项目课程于广州市某小学进行线下授课，面向的教学对象为小学五年级学生， 图 4 为教师进行现场授课。本节课结合天河区押花传统非遗技艺，并在信息科技的基础上，融合多个学科，如美术、语文、科学进行课程设计，设计了“发现押花之美”、“感受押花之美”、“创造押花之美”三个探究环节进行展开，以培养学生的创新精神和实践能力为核心，注重让学生在轻松愉快的氛围中探索和学习。通过引导学生运用 PPT 工具，让他们亲手制作电子押花，在感受非遗押花技艺魅力的同时，体验创作的乐趣，同时提升他们的信息技术素养和审美能力。



图 4 教师进行现场授课



图 5 学生作品展示

4.2 课程设计的优势与不足

本节课以综合培养学生的创新精神及实践能力为核心目标，通过结合非物质文化遗产——押花技艺与信息技术的应用，设计了一节寓教于乐的信息技术课。教学内容和方法都十分贴合当代教学理念，强调了学生在学习过程中的主体地位。课程明确知识与技能的传授目标，并采用了情境导入、任务驱动、示范教学等多样化的教学策略，有助于激发学生的学习兴趣，提高他们的学习效率。通过分层教学和小组合作，不仅能够照顾到不同水平的学生，也促进了学生间的交流与合作。教学过程中设计的三个探究环节“发现押花之美”、“感受押花之美”、“创造押花之美”极具创新创意，能有效地引导学生从感知美、欣赏美到创造美的过程中学习和体验。此外，不仅强调创新鼓励，这对于培养学生敢于尝试和创新的精神至关重要，还强调跨学科的课程设计，使学习活动更富教育意义，激发学生的发展潜能。

然而，《百变押花秀》课程设计也存在一些不足。由于材料限制和技术难度，学生可能在制作过程中遇到挑战。同时，评价体系的不完善可能影响对学生创造力和设计水平的全面评估。为了克服这些不足，需要进一步完善课程设计，优化教学资源，并建立更为全面的评价体系，以更好地推课程发展。

5. C-STEAM 校本课程开发的研究局限及未来展望

C-STEAM 教育模式在当前教育体系中具有显著的意义和价值，但在其校本课程开发的过程中，也面临着一些研究局限。首先，当前 C-STEAM 校本课程开发在组织形式上尚显单一，课程的组织者多数是学校教师，课程主要是在校内课堂中进行，没有真正做到“走出去”，让学生与社会建立联结，让知识在实际生活中得以运用。课程内容的融合和校内外社会多方合作也面临一定的挑战，难以实现真正的超学科统整。其次，尽管 C-STEAM 教育模式强调文化传承与创新教育的结合，但在实际操作中，往往过于注重知识的传授，而忽视了对学生创新思维和实践能力的培养。对于如何具体落实培养学生的人文精神、文化认同等目标，也缺乏明确的方法和手段。最后，C-STEAM 作为一种新型教育模

式，其品牌建设和推广力度尚显不足。很多学校对 C-STEAM 的理解不够深入，缺乏对其价值和意义的充分认识，导致在校本课程开发中缺乏积极性和主动性。基于上述问题，本研究从组织形式、培养目标、品牌建设三个方面对未来研究进行展望。

5.1 依托“文化+家校社”多方力量协同育人作用

未来 C-STEAM 校本课程开发应更加注重组织形式的创新，可以探索更多元化的教学方法，如项目式学习、探究式学习等，以激发学生的学习兴趣和主动性。同时，加强家庭、学校和社区的交流与合作，通过多方力量，实现协同育人作用。在学校方面，作为课程的主要实施者，学校应负责制定课程计划、组织实施教学活动等；在家庭方面，家长应积极支持学生的学习，提供必要的帮助和指导；在社区方面，社区可以为学生提供实践场所、文化资源等方面的支持。而作为项目课程的核心建设者，学校要加强沟通与合作，共同构建协同育人的 C-STEAM 教育优质生态圈。同时，C-STEAM 主题课程是一个不断发展和完善的课程体系，需要学校、家庭和社区不断探索与创新，实现“家校社”联合培养具有国际竞争力的创新型人才。

5.2 强调以“多元学科知识+社会生活技能”为培养目标

在培养目标方面，应更加注重学生的创新思维和实践能力的培养。在 C-STEAM 主题学习课程的设计与实施过程中，笔者发现不同于传统学科是通过教授某一学科的基本知识和技能、以提升学科核心素养为教学目标，C-STEAM 多学科融合课程在教授多个学科知识的同时，更注重教师创设情境教学和学生探究合作学习，可以通过设计具有挑战性和实践性的学习任务，引导学生主动思考和探索，让学生在实际生活场景中能够联结和运用多元学科知识发现问题、分析问题、解决问题，培养他们的解决问题能力和创新精神。同时，加强对学生人文精神、文化认同等方面的培养，帮助他们形成正确的价值观和世界观。

5.3 打造“本土文化+学校特色+多学科融合”的地方学校品牌课程

加强 C-STEAM 的品牌建设和推广力度，提高其在教育体系中的知名度和影响力。在 C-STEAM 主题学习课程的设计与实施过程中，笔者发现教学实施不局限在课堂上，更多的是在课外，一方面，让学生真实地感受传统文化的魅力，在文化熏陶中有助于深度学习；另一方面，给予学生更多的实践空间，加强社会联结，促进学生全面发展。学校要有意识、有魄力、有行动，结合学校和当地的实际情况，设计相关的跨学科主题，打造“本土文化+学校办学特色+多学科融合教学”的地方学校品牌课程，通过学校开展文化统整学科教育带动整个区域的文化宣传工作，发挥文化育人作用的同时，助推优秀传统文化的社会性推广，让更多的人了解 C-STEAM 的价值和意义。同时，加强与其他学校和机构的合作与交流，共同推动 C-STEAM 教育模式的发展和完善。

参考文献

中华人民共和国教育部(2022).义务教育信息科技课程标准（2022 年版）[EB/OL].http://www.moe.gov.cn/srcsite/A26/s8001/202204/W020220420582361024968.pdf.

中共中央办公厅、国务院（2021）.《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见[EB/OL].http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202107/t20210724_546576.html

中华人民共和国教育部（2021）.《中华优秀传统文化进中小学课程教材指南》[EB/OL].http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202102/t20210205_512630.html.

中华人民共和国教育部(2022).教育部关于印发义务教育课程方案和课程标准(2022 年版)的通知2[EB/OL].http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html.

董黎明, & 焦宝聪. (2020). 创客教育中创新思维发展要素模型构建. 开放学习研究, 25(5), 14–22. https://doi.org/10.19605/j.cnki.kfxxyj.2020.05.003

许秋璇, 杨文正, 卢雅, & 周琴英. (2020). 融入“大概念”的 STEM 整合课程设计模型构建与应用研究. 电化教育研究, 41(7), 86–93. https://doi.org/10.13811/j.cnki.eer.2020.07.012

袁磊, 张瑾, 滕洁梅, 覃颖, & 刘丽霞. (2022). 我国 STEAM 教育生态的反思与重构. 电化教育研究, 43(11), 31–37. https://doi.org/10.13811/j.cnki.eer.2022.11.004

詹泽慧, & 曾筱洵. (2022). 培养有文化涵养的跨学科创新人才——C-STEAM 的理念、方法与愿景. 中国信息技术教育, 18, 4–9.

詹泽慧, 季瑜, & 赖雨彤. (2023). 新课标导向下跨学科主题学习如何开展：基本思路与操作模型. 现代远程教育研究, 35(1), 49–58.

詹泽慧, 李克东, 林芷华, 钟柏昌, 麦梓莹, & 李炜贤. (2020). 面向文化传承的学科融合教育(C-STEAM):6C 模式与实践案例. 现代远程教育研究, 32(2), 29-38+47.

詹泽慧, 李彦刚, & 高瑞翔. (2022). 文化本位的跨学科学习：STEAM 教育本土化的一种重要途径. 中国电化教育, 1, 95–104.

詹泽慧, 钟柏昌, 霍丽名, & 黄美仪. (2020). 面向文化传承的学科融合教育(C-STEAM)：价值定位与分类框架. 中国电化教育, 3, 69–76.

Bybee R. W.（2013）. The case for stem education：Challenges and opportunities[M]. NSTA Press：82-87.

English L. D.（2016）. STEM education K-12：Perspectives on integration[J]. International Journal of STEM Education，3（1）：3,s40594-016-0036-1.

视频直播教学中互动策略对学习投入的影响研究

The Effect of Interactive Strategies on Learning Engagement in Live Video Teaching

杨琳[1*]，胡航[1]
[1] 西南大学教师教育学院
[*] 2301950402@qq.com

【摘要】 视频直播教学作为最为广泛的一种教学手段在学习和生活中具有举足轻重的作用，教学互动的形式也逐渐趋于多样化、个性化、动态化。本研究通过对 118 名大学生的问卷调查，采用描述性分析、皮尔逊相关分析和逐步回归分析，结合文献研究探究了师生互动(T-S)、生生互动(S-S)、学生和内容互动(S-C)三个维度十二小项的教学互动策略，探讨了视频直播课程教学中的互动策略对学习投入，包括行为投入(BE)、情感投入(EE)和认知投入(CE)三个维度影响，数据分析表明：教师的互动策略对提升视频直播课程教学过程中学生学习投入具有显著的积极影响。基于此，本研究从设计任务驱动模型、优化互动策略实施、技术赋能教学实践三方面提出建议，以期加强学生认知投入、丰富学生情感投入、深化学生行为投入，同时为教师在线直播教学时的策略选择提供依据，有助于学生在线学习效果的提升。

【关键词】 学习投入；互动策略；视频直播教学

Abstract: *Live video teaching has a pivotal role in learning and life as one of the most widely used teaching means, the form of teaching interaction has gradually tended to be diversified, personalized and dynamic. In this study, through a questionnaire survey of 118 college students, descriptive analysis, Pearson's correlation analysis and stepwise regression analysis, combined with literature research to explore the teaching interaction strategies of twelve small items in the three dimensions of teacher-student interaction (T-S), student-student interaction (S-S), student-content interaction (S-C), and to explore the interactive strategies in the teaching of the live video course on the learning inputs, including behavioral inputs (BE), emotional engagement (EE), and cognitive engagement (CE) dimensions. The data analysis showed that the teacher's interaction strategy has a significant positive impact on enhancing students' learning engagement in the teaching process of live video courses. Based on this, this study proposes suggestions from designing task-driven models, optimizing the implementation of interactive strategies, and technology-enabled teaching practices in order to strengthen students' cognitive input, enrich students' emotional input, and deepen students' behavioral input, as well as to provide a basis for teachers' strategy choices during live online teaching, which will help improve students' online learning effects.*

Keywords: learning engagement, interactive strategies,live video instruction

1. 前言

视频直播教学以其参与性强、交互性高的独特优势，已经成为了在线教育开展的重要形式。5G 时代的到来，为这种实时性强的直播教学打造了更加稳定成熟的网络教学环境，网络教学平台和视频会议等软件的互动功能也趋于多样化，但是技术的革新与进步并没有使互动功能得到很好地应用，普遍存在的现象是学生互动积极性不高、教师互动意识不强等问题，这也就会直接影响到学生的在线学习投入度和学习质量。随着直播教学的不断发展，如何设计视

频直播教学中的互动策略和方法，分析其对学生的学习投入的影响因素和作用机制，以提升在线学习体验和学习效果已经成为在线教学过程中亟待解决的问题。

2. 文献综述

2.1 视频直播教学

对于视频直播教学的概念还没有一个被广泛认可的界定，而且也经常存在多种说法比如在线直播课（杨姗姗和龙秋菊，2019）、直播教学（林晓凡和刘思琪，2019）等，聂赵育认为视频直播教学是指以视频为媒介进行的实时性教学活动,形式上以直播为主，复播为辅（聂赵育，2019）。视频直播教学就是依托于网络直播平台来开展在线学与教活动的一种方式，这种能够进行实时互动的形式能够使学习者开展异地同步、实时有效的在线学习。相关研究表明，交互性是评价在线课程教学质量的重要指标（肖广德和黄荣怀，2014）。所以，提升在线学习质量的首要任务就是要让学习者积极参与到教学互动中，有学者也指出了在线交互行为活动是在线学习投入的发展趋势之一（尹睿和徐欢云，2016）。

2.2 在线学习投入

从心理学的角度出发，在线学习投入主要由行为投入、认知投入和情感投入三个维度构成（Christenson et al., 2012）。迄今为止，关于学习投入度的研究大多从教师支持策略（栾琳等人，2022；刘斌等人，2017）、网络学习环境（齐媛，2019）、学习动机（黄庆双等人，2019）等视角入手，依据学业自我效能感和学业拖延（郭文斌和苏蒙，2021）等中介预测大学生的在线学习投入度，但很少有研究探究在线直播教学中的互动策略与学习投入之间的关系。已有研究表明, 学生、教师和内容三者之间的互动与学生的学习投入度有关（Wang et al., 2022)），学习投入是衡量学习满意度和学业成就的重要指标，能较为准确地预测学习者的网络学习质量和学习过程中的心理状态（Martin & Bolliger, 2018）。

2.3 视频直播教学互动策略

依据目前在线学习现状，相关学者提出了“在线学习”和“课堂学习”的学习文化、学习共同体、学习环境、协商交往、建构以及绩效六大互动策略（周世平，2006）。也有学者基于混合式学习理念设计了师生共建、问题情景、协作与竞争、激励与调控策略（胡政权等人，2020）等诸多互动策略。总览教学互动策略研究，绝大部分学者依据师生关系互动提出相对应的策略，对于学生和学生以及知识和平台等主体研究较少。由于平台只是作为教学的媒介，所以本研究只考虑学习者，提出师生（T-S）、生生(S-S)以及学生和学习内容间的三大类十二小项的互动策略（如表 1 所示），进一步探究视频直播教学中教师、学生和教学内容三者之间的互动关系，分析在线直播教学互动与学习投入的关系，揭示互动策略影响学习投入的内部作用机制，以期促进大学生在线学习的有效开展，使其在互动过程中表现出更强的学习动机和参与感。基于此，本研究提出以下假设：视频直播教学中的互动策略可以显著影响学习投入度。

表 1

视频直播教学互动策略

分类	互动策略	具体描述
师生互动(T-S)	氛围营造	教师可以通过课前点名、播放音乐和视频等营造轻松的学习氛围
	鼓励引导	教师通过任务驱动式学习方法以及言语、手势等启发学生交流互动
	提问测验	教师针对重点内容进行实时提问，每一节课内容采取当堂测验
	评价激励	教师对于学生的回答给予及时的肯定，对不同学生进行个性化评价

生生互动(S-S)	交流讨论	学生通过与其他学习者互相交流学习内容和课后作业来提升兴趣
	组内合作	小组合作学习我利用各种学习工具，使用各种学习策略来寻求答案
	公开分享	评选标兵、优秀案例分享激发学生竞争意识，起到榜样示范作用
	组间辩论	学生通过小组协作开展项目式学习，在互评过程中发现解决问题
学生和内容互动(S-C)	知识内化	学生通过深入理解并内化课本的阅读材料来掌握学习内容
	资源整合	学生通过运用各种学习策略来整合学习资源的内容
	主动搜集	学生主动去查阅相关文献、书籍资料，积极构建自己的知识体系
	动手操作	制作微课视频和电子课件素材等，将制作的资源与老师同学共享

因此，本研究包括两个研究步骤，首先，从影响因素出发，梳理在线直播课程中学习者与教师、学生和教学内容三者之间互动的策略，将各项在线直播教学策略进行分类整合；其次，通过问卷调查和访谈，探究在线直播教学互动策略对大学生行为投入、情感投入和认知投入三方面的影响。

3. 研究过程及数据分析

3.1 研究对象

本研究选取参与在线直播课程的高校本科生作为研究对象，通过问卷星发放并回收问卷共118 分，剔除作答时间小于 30 秒、量表总得分为满分或零分的问卷后，共获得有效问卷 99 份，有效率为 83.90%。其中，男生 28 名，女生 71 名。按照被试的年级从大一到大四分别占比 47.47%、8.08%、10.10%、34.34%。验证视频直播课程的三种互动方式与学生学习投入度之间的关系。

3.2 研究工具

3.2.1. 教学互动策略对大学生在线学习投入度的影响问卷

本研究对《大学生在线学习投入中教学交互存在问题》调查问卷（王明慧，2022）进行改编，调查问卷共有题项 14 项，分为师生互动（T-S）、生生互动(S-S)、学生和内容互动(S-C)三个维度，每个维度下设四个小的教学互动策略问题。积分采用里克特五点计分，从 1 分“完全不符合”到 5 分“完全符合”，得分越高表明师生互动水平越高。本研究中，该量表的标准化 Cronbach α系数为 0.964，内部一致性良好，KMO 值为 0.921（>0.6），表明该量表的结构效度较好。

3.2.2. 在线学习投入量表

Sun 和 Rueda 等人修订的“远程学习投入量表” (SEDE)（Sun & Rueda, 2012）。该量表由行为投入(BE)、情感投入(EE)和认知投入(CE)三个维度组成，计分方式采用李克特五级量表，随后利用问卷星发放问卷，从 1 到 5 分别表示“完全不符合”到“完全符合”。分数越高表明在线学习者的体验感和投入度越强。在本研究中，该问卷行为投入、认知投入、情感投入三个维度以及问卷整体的内部一致性信度（Cronbach's 系数）分别为：0.850、0.852、0.904、0.936，问卷信度较好，KMO 值为 0.914（>0.6），表明该量表的结构效度较好。

3.2.3. 访谈法

本研究从在线直播课程互动情况以及互动对在线学习投入影响情况对参与在线直播教学的一些大学生进行了访谈，访谈结果作为调查问卷的补充研究。虽然这些同学来自不同的年级、不同的大学但是对于在线直播课程互动以及学习投入都有着各自的见解，通过访谈结论如下：大多数同学自新冠肺炎疫情以来已参与在线学习三年以上的时间，对于直播课程已十分熟悉并且能够很好地适应这种学习方式。从在线直播课程互动情况来看，大多数同学表示

愿意与老师进行课堂上的交流讨论疑难问题，对于老师布置的任务也能够及时完成，但是也有少部分同学表示并不愿意与老师在线互动而且表明这种实时点名和提问会给自己的学习带来一定的压力和负面情绪；从学习方式来说，大多数同学认为小组合作的学习效果要优于个人学习，因为这可以很好地调节在线学习孤独感和学习焦虑情绪；面对如此海量的学习资源，很多同学表示会存在一定的资料查询障碍，并且对于如何分辨良莠不齐的学习内容有一定的困难。从互动对在线学习投入影响情况来看，普遍情况是学生能够投入大量的时间和精力去关注直播课程，而且对于何种互动方式能够更好地提高自己的学习效果这个问题，同学们各抒己见，并未有明显的策略喜好偏向，而且部分同学对于学习投入的自我认知定位不够清晰，而且大部分同学表示现有的互动总是停留在浅层次的问题，如点名、签到等活动，互动策略的实施深度明显不够。

3.3 数据处理及分析

本研究使用 Excel 进行数据录入、导出与管理，使用 SPSS AU 进行分析。数据处理的过程包括三个环节：一是描述性统计，便于比较投入度的三个维度影响大小；二是 Pearson 相关分析，探究在线直播教学互动策略对大学生行为投入、情感投入和认知投入三方面的影响；三是逐步回归分析，将教学互动策略作为自变量、投入度作为因变量，探索三大类十二小类的教学策略能否有效预测在线学习投入度。

3.3.1. 大学生在线学习投入的描述性统计

数据的描述性统计分析结果显示，大学生在线学习投入度的平均值为 3.76，三个维度的平均值分别为：行为投入（CE）3.75、情感投入（EE）3.63、认知投入（BE）4.10。由此可见，大学生在线学习投入度的整体水平处于中等偏上的水平，其中认知投入的水平要明显优于行为投入和情感投入两个维度。

3.3.2. Pearson 相关分析

利用相关分析去进一步探究行为投入 BE、情感投入 EE、认知投入 CE 分别和师生互动、生生互动、学生和内容互动共 3 项之间的相关关系，使用 Pearson 相关系数去表示相关关系的强弱情况如表 2 所示。由此可以看出 X 轴各维度之间与 Y 轴各维度之间全部均呈现出显著性，且有着正相关关系，说明视频直播教学中互动策略对学习投入的程度有很大的影响。

表 2

互动策略与在线学习投入度的 Pearson 相关分析

	行为投入 BE	情感投入 EE	认知投入 CE
师生互动	0.615**	0.532**	0.559**
生生互动	0.644**	0.587**	0.559**
学生和内容互动	0.726**	0.591**	0.635**

* $p<0.05$ ** $p<0.01$

3.3.3. 逐步回归分析

首先，将行为投入 BE 作为因变量。经过模型自动识别，如表 3 所示，最终余下 3 项在模型中，R 方值为 0.560，意味着(S-S)交流讨论、(S-C)主动搜集、(S-C)动手操作可以解释行为投入 BE 的 56.0%变化原因。而且模型通过 F 检验(F=40.256，p=0.000<0.05)，说明模型有效。即交流讨论、主动搜集、动手操作这三种互动策略会对行为投入产生显著的正向影响关系。

表 3

行为投入逐步回归分析结果 (n=99)

	非标准化系数	标准化系数	*t*	*p*	VIF

	B	标准误	*Beta*			
常数	1.396	0.256	-	5.450	0.000**	-
(S-S)交流讨论	0.246	0.087	0.287	2.819	0.006**	2.237
(S-C)主动搜集	0.246	0.081	0.307	3.056	0.003**	2.176
(S-C)动手操作	0.189	0.083	0.247	2.263	0.026*	2.562
R^2			0.560			
调整 R^2			0.546			
F			*F* (3,95)=40.256, *p*=0.000			
D-W 值			2.113			

因变量：行为投入 BE

* $p<0.05$ ** $p<0.01$

其次，将情感投入 EE 作为因变量。经过模型自动识别，如表 4 所示，最终余下 2 项在模型中，*R* 方值为 0.410，意味着(S-S)组间辩论，(S-C)知识内化可以解释情感投入 EE 的 41.0%变化原因。而且模型通过 *F* 检验(*F*=33.348，*p*=0.000<0.05)，说明模型有效。即组间辩论和知识内化两种互动策略会对情感投入产生显著的正向影响关系。

表 4

情感投入逐步回归分析结果 (n=99)

	非标准化系数		标准化系数	*t*	*p*	VIF
	B	标准误	*Beta*			
常数	1.415	0.281	-	5.040	0.000**	-
(S-S)组间辩论	0.260	0.083	0.324	3.148	0.002**	1.721
(S-C)知识内化	0.294	0.079	0.381	3.706	0.000**	1.721
R^2			0.410			
调整 R^2			0.398			
F			*F* (2,96)=33.348, *p*=0.000			
D-W 值			1.804			

因变量：情感投入 EE

* $p<0.05$ ** $p<0.01$

最后，将认知投入 CE 作为因变量。经过模型自动识别，如表 5 所示，最终余下 2 项在模型中，*R* 方值为 0.412，意味着(T-S)氛围营造，(S-C)知识内化可以解释认知投入 CE 的 41.2%变化原因。而且模型通过 *F* 检验(*F*=33.576，*p*=0.000<0.05)，说明模型有效。即氛围营造和知识内化两种互动策略会对认知投入产生显著的正向影响关系。

表 5

认知投入逐步回归分析结果 (n=99)

	非标准化系数		标准化系数	*t*	*p*	VIF
	B	标准误	*Beta*			
常数	1.441	0.290	-	4.967	0.000**	-
(T-S)氛围营造	0.236	0.104	0.277	2.274	0.025*	2.414
(S-C)知识内化	0.366	0.110	0.405	3.327	0.001**	2.414

R^2	0.412
调整 R^2	0.399
F	F (2,96)=33.576, p=0.000
D-W 值	2.016

因变量：认知投入 CE

* p<0.05 ** p<0.01

总的来说，学生和内容互动的策略实施对大学生的行为投入、情感投入和认知投入三方面均具有显著的影响（如图 1 所示）。由此可见，在线直播课程教学中，教师如果能注意主动搜集、动手操作和知识内化等策略的实施，促进学生与教学内容之间的互动，会极大地提高学生再行为、情感和认知等方面的投入度，也会提高学生的在线学习积极性，激发其自主学习的动机。

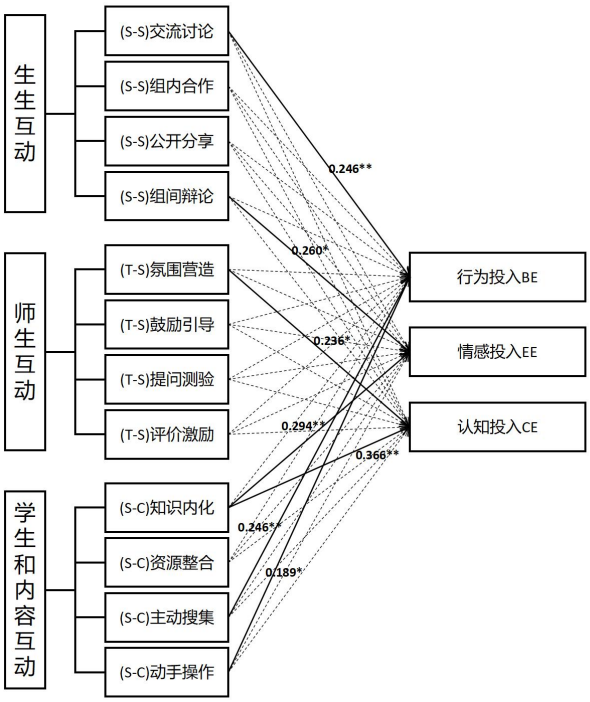


图 1

在线直播教学互动策略与投入关系模型

3.4 研究结论与讨论

本研究的创新之处在于：一是国内外关于在线课堂互动策略的研究大多是理论化的策略，很少应用于实践，缺乏一定的可操作性；二是大多研究策略只针对于师生之间的互动，少有提及学习者与学习者之间以及和学习内容之间的互动；三是深入分析对学习投入的影响最为显著的视频直播教学互动策略，并将其应用于预测学习投入度。而实验结果证明了视频直播教学中的互动策略可以显著影响学习投入度。主要得出以下结论：大学生的在线学习投入度整体水平较高，情感投入和行为投入水平还有待提高；提升大学生学习投入度可以从三方面互动策略入手，其中要重点关注学生和学习内容之间的互动，对提升视频直播课程教学过程中学生学习投入具有显著的积极影响。

4. 对策建议

4.1 设计任务驱动模型，加强学生认知投入

在教学过程中，教师要精心设计任务驱动式的问题，在课堂导入时创设好学习氛围，在教学法方面，教师可以运用跨空间小组协作学习活动策略增强远程学生的课堂融入感（师亚飞等人，2021），以便加强学习者的认知投入，引导学习者实现知识的内化和顺应。

4.2 优化互动策略实施，丰富学生情感投入

推进情感教育为魂的数字教育，促进信息技术与教育教学的深度融合，可以深入分析学生特点，借助 AR、VR 等技术手段来丰富学生的情感体验，更好地实施大规模因材施教，关注人的全面发展，丰富情感投入，使学习者对在线直播教学有了更深层次的归属感（刘繁华和易锡添，2021）。

4.3 技术赋能教学实践，深化学生行为投入

我国教育评价改革进入前所未有的重视阶段，教育教学过程中可以利用多模态数据监测学习投入的变化规律，较全面地揭示了学生在协作过程中行为、认知和情感的投入状态（吴军其等人，2022），制定更具准确性和解释性的技术测评方案，来推动在线直播教学的顺利开展（骆方等人，2021）。通过技术赋能教学实践，深化行为投入，促使学习者参与更有价值的学习活动。

5. 结语

本研究探讨了在线直播课程中三维度互动策略对大学生学习投入的影响，有助于通过互动策略的选择加强学生认知投入、丰富学生情感投入、深化学生行为投入。同时，本研究进行实验时间较短，实验对象较少，得出的实验结果具有一定的局限性。此外，本研究所提及的三大类策略还可以进一步的发展完善，未来可以在本研究的基础上进一步探究基于其他中介作用影响下的互动策略对学习投入的影响以及不同中介之间的关系，以提升在线学习的学习效果，达到学习者的高参与度和高投入度的在线学习新常态。

参考文献

郭文斌和苏蒙（2021）。网络学习空间教师支持对大学生学习投入的影响研究——基于学业自我效能感的中介作用。**教育理论与实践**，**30**，50-54。

胡政权、赖于树和邱刚（2020）。基于 SPOC 混合式学习的教学设计及互动策略。**无线互联科技**，**07**，104-106。

黄庆双、李玉斌、任永功、郑玉敏和黄悦悦（2019）。大学生成就目标定向对在线学习投入影响研究:学业拖延的中介作用。**现代远距离教育**，**05**，77-85。

黄荣怀、张慕华、沈阳、田阳和曾海军（2020）。超大规模互联网教育组织的核心要素研究——在线教育有效支撑“停课不停学”案例分析。**电化教育研究**，**03**，10-19。

林晓凡和刘思琪 (2019)。面向高阶思维能力培养的直播教学策略。**现代教育技术**, **03**, 99-105。

刘斌、张文兰和刘君玲（2017）。教师支持对在线学习者学习投入的影响研究。**电化教育研究**，**11**，63-68+80。

刘繁华和易锡添（2021）。在线学习投入的分析模型构建及应用研究。**电化教育研究**，**09**，69-75。

栾琳、董艳和刘金金（2022）。教师支持策略对大学生在线学习投入的影响研究。**现代教育技术**，**03**，119-126。

骆方、田雪涛、屠焯然和姜力铭（2021）。教育评价新趋向：智能化测评研究综述。**现代远程教育研究**，**05**，42-52。
聂赵育（2019）。视频直播教学:“互联网+教育”新形势。**软件导刊(教育技术)**，**09**，80-81+86。
齐媛（2019）。网络学习环境对小学生学习投入的影响研究——以教客平台为例。**中国特殊教育**，**12**，83-88。
师亚飞、童名文、孙佳、戴红斌、龙陶陶和王建虎（2021）。混合同步学习环境对学生认知投入的影响机制研究。**中国远程教育**，**09**，29-38+68+77。
王明慧（2022）。教学交互对大学生在线学习投入的效用及优化策略研究，延安大学。
吴军其、吴飞燕、张萌萌、戴新菊和张影（2022）。多模态视域下智慧课堂协作学习投入度分析模型构建及应用。**电化教育研究**，**07**，73-80+88。
肖广德和黄荣怀（2014）。学习过程导向的网络课程教学交互设计研究——面向网络课程实施的视角。**电化教育研究**，**09**，85-90。
杨姗姗和龙秋菊（2019）。基于在线直播课的大学英语“线上线下”混合式教学模式研究。**百色学院学报**，**06**，135-140。
尹睿和徐欢云（2016）。国外在线学习投入的研究进展与前瞻。**开放教育研究**，**03**，89-97。
周世平（2006）。“在线学习”与“课堂学习”的互动策略研究，上海师范大学。
Christenson, S.L., Reschly, A.L., & Wylie, C. (2012). *Handbook of Research on Student Engagement*. Springer US.
Martin, F. & Bolliger, D.U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning 22*(1), 205-222.
Sun, C. Y. , & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: their impact on student engagement in distance education. *British Journal of Educational Technology, 43*(2), 191-204.
Wang, Y., Cao, Y., Gong, S., Wang, Z., Li, N., & Ai, L. (2022). Interaction and learning engagement in online learning: The mediating roles of online learning self-efficacy and academic emotions. *Learning and Individual Differences, 94*, 1041-6080.

隐私保护的学习者学习投入度识别

Privacy-Preserving Learning Engagement Recognition

陈晓雨[12]，魏艳涛[12]，马玉建[12]，王涛[12*]
[1]华中师范大学 人工智能教育学部 湖北 武汉 430079
[2]数字教育湖北省重点实验室（华中师范大学） 湖北 武汉 430079
[*] tmac@ccnu.edu.cn

【摘要】基于人脸数据的投入度识别已成为智慧教育背景下学习分析的重要途径，但人脸数据涉及隐私风险。为平衡数据开发和隐私保护，本研究提出差分隐私保护的学习投入度识别模型 DPLER(differential privacy-preserving learning engagement recognition)，解耦身份与识别特征，通过差分隐私添加拉普拉斯噪声保护身份隐私，并构建基于 LRCN(Long-term Recurrent Convolutional Networks)的模型进行投入度识别。实验结果显示，DPLER 在保护人脸隐私的同时，与无隐私保护下的原始人脸识别准确率相当，为安全可信的学习分析环境提供技术路径。

【关键词】 智慧教育；学习投入度；隐私保护；差分隐私；人脸数据

Abstract: *Recognizing the learning engagement based on facial images has become an important approach in analyzing learning within the context of smart education. However, facial images contains privacy risks. To Balance the relationship between data development and privacy preservation, this study proposes a differential privacy-preserving learning engagement recognition model called DPLER, which decouples identity and recognition features. It preserves identity privacy by adding Laplace noise through differential privacy, and constructs a model based on LRCN for engagement recognition. Experimental results show that DPLER achieves comparable accuracy to the original facial recognition without privacy preservation while preserving facial privacy. This provides a technical pathway for constructing a secure and trustworthy learning analysis environment.*

Keywords: learning engagement, privacy preservation, differential privacy, facial image, learning analysis

1. 前言

人工智能时代下学习分析致力于以多模态、即时性和个性化的数据探索教育本质，成为推动教育数字化转型的重要力量。个性化、精准性和实时性的学习数据在带来巨大教育价值的同时也蕴藏着隐私泄露和侵犯等风险，引发潜在的信任危机（田贤鹏，2020）。2021 年教育部等七部门发布关于加强教育系统数据安全工作的通知，重点保障个人信息安全。2023 年 3 月第十三届学习分析与知识国际会议主题为“迈向可信学习分析”，聚焦于学习者的隐私、公平和自主发展等（吴永和等，2023）。随着智慧教育环境的全面升级，学习者行为数据被广泛而实时的进行检测和采集，例如学习者的高敏感人脸数据等。基于此，本研究针对采集课堂学生人脸数据，提出了一种差分隐私保护的学习投入度识别模型 DPLER，以期为促进教育大数据安全应用与挖掘提供理论框架与技术路径，促进数智时代教育生态和谐优质发展。

2. 研究方法

本研究的主要目的是探究构建保护学生隐私的隐私化数据集的方法，确保蕴含学情的学生人脸数据能够安全可信地被分析利用，平衡数据可用性和学生隐私之间的关系冲突。

首先要构建基于人脸数据的学习投入度识别模型。本研究的学习投入度结合相关基于视觉感知的研究（Ömer 等，2021；赵春等，2021；Pabba& Kumar，2022），主要采用面部表情、面部动作单元 AUs (action units)和头部姿态实现表征，分别记为为高投入，中投入和低投入。高投入主要涉及情绪表现为常态、惊讶、开心等表情，面部 AUs 为 AU5 或 AU14 或 AU5+AU24，头部处于正常状态。中等投入主要涉及情绪表现为常态、自然、放松等表情，头部处于正常状态或轻微偏转。低投入主要涉及情绪表现为乏味、疲惫、无聊等表情，面部 AUs 为 AU27 或 AU7+27 或 AU43，头部产生轻微或严重偏转。

其次是实现学生人脸数据的隐私保护机制。本研究使用差分隐私技术解决学生人脸采集分析中的隐私保护需求。差分隐私通过在数据或查询结果中添加随机噪声，使得攻击者无法通过结果推测特定个体的信息，从而保护了个体的隐私（Muneeb 等，2020）。

最后是验证基于隐私化数据集的学习投入度识别模型训练结果的准确性。本研究将进一步构建学习投入度识别模型，并运用模型训练隐私化后的数据集和原始数据集，比较隐私化数据集和原始数据集在分类识别上的准确性，从而验证所提模型的有效性。

3. 隐私保护的学习投入度识别模型

本研究提出一种基于差分隐私增强的学习投入度识别模型 DPLER (differential privacy-preserving learning engagement recognition)，DPLER 模型构建包括人脸隐私化和学习投入度识别两个阶段。在人脸隐私化阶段，DPLER 模型对学习者人脸中与身份相关的敏感特征和与识别相关的普通特征解耦，身份特征添加拉普拉斯噪声并保证视觉相似性，从而保护学生隐私。在学习投入度识别阶段，基于 LRCN 构建学习投入度识别模型，有效评测学习者的学习投入度。

3.1 第一阶段：人脸隐私化

人脸隐私化的处理过程包括三个步骤，如图 1 所示。

（1）对人脸数据中的身份表征和属性表征进行分离，模型由三个部分组成：身份编码器，属性编码器和融合生成器。我们使用最先进的人脸识别模型作为身份编码器在潜在空间中提取精确的面部身份特征，表示为：$z_{id}(X)=f(X)$。接下来，我们将人脸图像输入到全卷积分割网络 U-Net 中，从 U-Net 解码器生成的特征映射作为属性表征 $z_{attr}(X)=g(X)=\{z^1_{attr}(X),..., z^n_{attr}(X)\}$，如姿态、表情、光照和背景等，其中 $z^k_{attr}(X)$ 表示为 U-Net 解码器的第 k 级特征图，n 表示特征层数。然后，我们需要将这些身份表征和属性表征直接拼接在一起，在构建融合生成器时采用自适应注意力反标准化生成器(Adaptive attention denormalization, AAD)（Zhang 等, 2023）在多个特征层次上实现特征集成，自适应调整身份嵌入和属性嵌入的有效区域，通过融合生成器恢复原始人脸 $X=h(z_{id}, z_{attr})$。

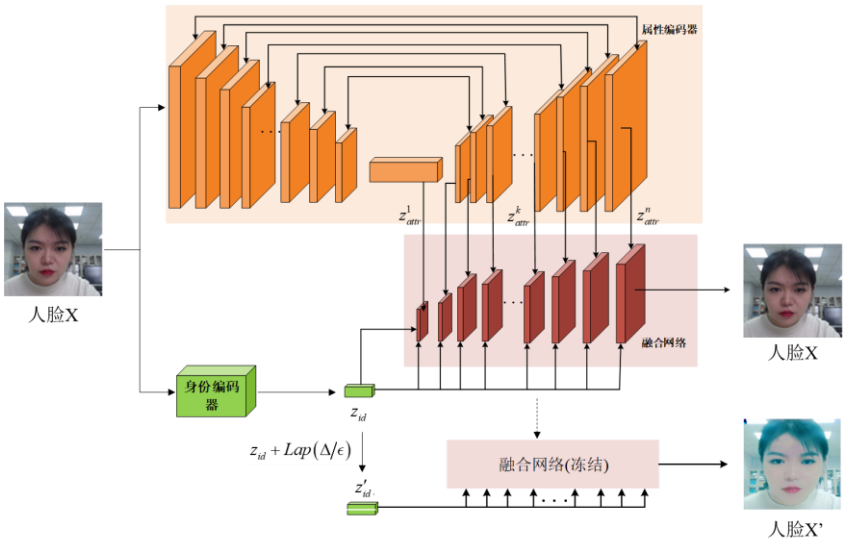
（2）向身份表征上加入噪声。对于人脸 X，如果一随机算法 M 对身份特征添加拉普拉斯噪声后 $M(X)= z_{id}(X)+Lap(\Delta/\epsilon)= z'_{id}(X)$，满足 $Pr[z_{id}(X)\in O]\leq e^{\epsilon}Pr[z'_{id}(X)\in O]$，其中 O 是所有输出的集合，则 M 满足 ϵ-差分隐私保护，其中$\Delta=\max\|z_{id}(X)- z'_{id}(X)\|_1$ 为敏感度函数。更具体地说，在该步骤使用噪声发生器在我们选定的隐私预算 ϵ 上生成拉普拉斯噪声,然后直接 z_{id} 上添加拉普拉斯噪声，得到一个扰动身份表征 z'_{id}。

（3）合成身份表征和属性表征，生成隐私化人脸。该步骤冻结了第一阶段训练好的融合生成器的所有参数，结合融合生成器的扰动的身份表征 z'_{id} 和原始的属性表征 z_{attr} 生成隐私化人脸 $X'=h(z'_{id}, z_{attr})$。 在这个过程中，我们需要比较学习投入度识别模型在不同隐私预算参数下中的准确率，因此需要对模型进行多轮训练，以得出准确率最高的隐私保护方案。

3.2 第二阶段：差分隐私保护的学习投入度识别

第二阶段是在隐私化处理的学生人脸数据集上构建学习投入度识别模型。本研究基于

图 1 人脸隐私化的处理过程



LRCN（Jeff 等，2016）构建学习投入度识别模型，LRCN 同时具备处理时序视频或单帧图片的能力，以及单值预测或序列预测的能力，这使得 LRCN 成为处理序列输出输入信息的集大成网络。本研究通过 LRCN 模型预测学生的学习投入状态属于低投入、中投入或高投入。

4. 隐私保护的学习投入度识别模型的实践验证

4.1 隐私化数据集构建

由于目前开放的课堂学生隐私化人脸数据库并不多，且大都是国外学习者群体（李新&李艳燕，2021），鉴于不同区域人脸的差异，为了保证数据真实性和实用性，本研究通过在线学习场景，收集了 30 位学习者的学习视频，然后将每个学习者的学习视频剪切成若干个 10s 的视频，形成用于学习投入度识别的人脸数据集 LeEngFace；然后对 LeEngFace 隐私化处理，构建隐私化人脸数据集 LeEngDPFace，并对数据集中的数据进行学习投入标注。

4.2 实验环境及参数设置

实验环境为 64 位 Ubantu18.04 操作系统,NVIDIA GeForce RTX 3090 开启 GPU 加速,网络模型搭建使用 python3.7 编程语言、pytorch1.7.1 深度学习框架。

4.3 学习投入度识别结果分析

基于 LeEng-DPFace 数据集和 LeEng 数据集所训练的 LRCN 学习投入度识别模型的识别准确率如表 1 所示。基于 LeEng-DPFace 数据集的训练数据集的准确率为 96.70%，测试集训练准确率为 66.87%, 相比基于 LeEng 数据集的准确率略有下降, 以极小的识别准确性为代价, 提高了模型和数据的安全性。

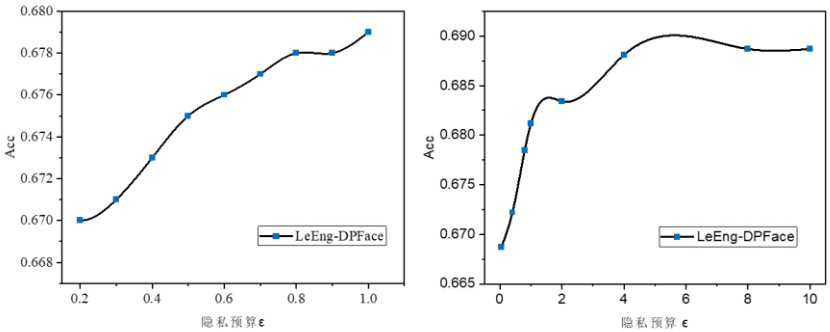
表 1 基于 LRCN 的学习投入度识别模型性能

数据集	训练集准确率	测试集准确率
LeEng	98.7%	69%
LeEng-DPFace	96.7%	68.87%

人脸的隐私保护强度取决于隐私参数 ϵ，ϵ 越小，意味着添加的噪声扰动强度越大，可能会影响识别结果的准确性；否则，则反之。图 2 为在不同隐私参数 ϵ 下学习投入度识别模型的表现情况。结果表明，随着隐私参数的增加，拉普拉斯扰动过程向身份表征注入的噪声减少，隐私保护的级别逐渐降低，准确率随之提高。我们可以根据实际情况选择合适的隐私预算参数来平衡隐私保护和模型性能之间的关系。由此可以证明，本文的 DPLER 模型在保护人

脸数据隐私的前提下能够保护模型的性能，解决了隐私化数据应用于课堂投入度分析存在的问题，设计具有更高服务可用性的隐私保护方案，克服了隐私化数据应用于课堂投入评测的技术瓶颈。

图 2 不同差分隐私方案对学习投入度识别准确性的影响



5. 总结与展望

学生人脸数据能够准确反映学习者的学习投入度状态，对于评估学习效果和提供即时反馈具有重要价值，但同时也引发了学生对于隐私泄露的恐慌。本研究提出一种隐私保护的学习投入度识别模型 DPLER，达到学习投入分析和隐私保护的双重目的。隐私保护机制有助于提高教育质量，实现个性化教育的目标，让学生在安全可信的数字教育生态中全面发展。

参考文献

田贤鹏（2020）。隐私保护与开放共享：人工智能时代的教育数据治理变革。**电化教育研究**，**41**，05，33-38。

吴永和、钱雨、郭胜男和许秋璇（2023）。迈向可信学习分析：报告解析、结构模型与未来生态——2023 年学习分析与知识国际会议评述。**现代远距离教育，05**，3-15。

Ömer, S., Patricia, G., Sidney, D., Peter G., Ulrich, T., & Enkelejda, K. (2021). Multimodal Engagement Analysis from Facial Videos in the Classroom. *IEEE Transactions on Affective Computing*,*14*(2), 1012-1027.

赵春、舒杭和顾小清（2021）。基于计算机视觉技术的学生课堂学习行为投入度测量与分析。**现代教育技术**，**31**，06，96-103。

Pabba, C. & Kumar, P. (2022). An Intelligent System for Monitoring Students' Engagement in Large Classroom Teaching Through Facial Expression Recognition. *Expert Systems*, *39*(1): 1-28.

Muneeb, U., Mubashir H., Jinjun, C. (2020). Differential Privacy Techniques for Cyber Physical Systems: A Survey. *IEEE Communications Surveys & Tutorials*, *22*(1), 746-789.

Zhang, Y., Liu, y., Hu, r., Wu, Q., & Zhang, J. (2023). Mutual Dual-Task Generator with Adaptive Attention Fusion for Image Inpainting. *IEEE Transactions on Multimedia*, *26*, 1539-1550.

Jeff, D., Lisa, A., Marcus R., Subhashini, V., Sergio, G., Kate S., & Trevor, D. (2016). Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(4), 677-691.

李新和李艳燕（2021）。基于系统性文献综述的国外学习投入实证研究分析。**现代远程教育研究**，**33**，2，73-83+95。

多模态教学视频分析技术驱动的职前教师集体反思中的认知与情感分析

Cognitive and Emotional Analysis of Collective Reflection of Pre-service Teachers Driven by Multi-modal Teaching Video Analysis Technology

叶如梦[1]，王婷[2]，龙陶陶[3*]
[123]华中师范大学人工智能教育学部
[*]taotaolong@ccnu.edu.cn

【摘要】 集体反思已逐渐成为提升教师专业发展的关键一环。但相关研究表明，教师的集体反思缺乏深入性和批判性。为此，本研究将基于视频的自动多模态分析技术整合到集体反思中，比较了有无基于视频的多模态分析报告的集体反思，通过认知网络分析可视化职前科学教师的集体反思的认知和情感模式，从认知和情感变化角度探讨该技术在集体反思中的作用。研究结果表明，职前科学教师在结合了多模态分析报告的集体反思中表现出更高的认知水平。定性数据也证实，通过对报告的讨论，他们对教学实践和如何改进未来教学有了更深入和批判性的反思。

【关键词】 集体反思；基于视频的自动多模态分析技术；认知网络分析；职前科学教师

***Abstract*:** *Collective reflection has gradually become the key to improving teachers' professional development. However, relevant research shows that teachers' collective reflection lacks depth and criticality. Therefore, this study integrated video-based automatic multimodal analysis technology into collective reflection, compared collective reflection with and without video-based multimodal analysis reports, visualized the cognitive and emotional patterns of collective reflection of pre-service science teachers through Epistemic Network Analysis, and explored the role of this technology in collective reflection from the perspective of cognitive and emotional changes. Findings suggest that pre-service science teachers exhibit higher cognitive levels in collective reflection with the reports. Qualitative data also confirmed that through discussion of the reports, they developed deeper and critical reflections on teaching practice and improving future teaching.*

Keywords: *collective reflection, Video-based automatic multi-modal analysis technology, Epistemic Network Analysis, pre-service science teacher*

1.引言

近年来，党和国家对科学教育高度重视，相继出台了《关于新时代进一步加强科学技术普及工作的意见》《全民科学素质行动规划纲要(2021—2035 年)》《关于加强新时代中小学科学教育工作的意见》等一系列政策文件，强调要大力发展科学教育，培养具备科学家潜质、愿意献身科学研究事业的青少年群体。然而，科学教师队伍质量一直是影响科学教育发展的关键因素。为此，如何促进科学教师的专业发展成为了科学教育界的重要议题（Wilson, 2013; 季娇、伍新和燕婷，2014; 王晓生，2023）。教学反思被认为是促进科学教师专业发展的重要手段。长期以来，教学反思多由科学教师独自开展，被视为科学教师的个人实践活动。由于职前和在职科学教师的教学实践和专业知识技能水平有限，他们的个体反思活动往往效果不佳，因此在教师教育领域，集体反思受到了极大的重视。集体反思是指社区成员通过社会互动共同反思的过程及其结果（Epler et al., 2013; Kager et al., 2022）。许多研究表明，集体反思能更好地帮助教师开展反思活动（Harford & MacRuairc, 2008; Arefian, 2022; 张晓光，

2020）。他人的介入有助于教师从不同角度重新审视自己的教学行为，产生新的教学思路（Sorensen，2014），进而提升反思成效。然而，针对职前科学教师的集体反思研究仍相对有限。也有证据表明，教师的反思性思维水平较低，他们在参与集体反思的过程中能够积极表达自己的观点，但缺乏更深入的问题理解和批判性评价（Tang et al., 2023）。

为弥补这些不足，本研究在集体反思中采用了基于视频的自动多模态分析技术。这种技术能够从整体、微观和具体的视角对教师的课堂行为进行分析，为教师提供有关其工作的有用信息（Davidsen & Vanderlinde, 2014）。但少有量化研究探讨将基于视频的自动多模态分析技术融入集体反思后，教师的认知和情绪变化。因此，本研究的目的是使用认知网络分析（Epistemic Network Analysis，简称 ENA）来可视化职前科学教师在将基于视频的自动多模态分析报告融入集体反思前后的认知和情绪状态，从而探讨基于视频的自动多模态分析技术在促进职前科学教师集体反思中的作用。

2.集体反思和基于视频的自动多模态分析技术的研究综述

集体反思是指社区成员通过社会互动共同反思的过程及其结果（Martin & Double, 1998）。在此过程中，成员们通过观察、协作和知识交流进行集体体验式学习（Dixon, Lee, & Corrigan, 2021）。与个体反思不同，集体反思发生在社会背景下，注重群体之间开展反思性对话，互相咨询和比较（Subramaniam, 2013）。鉴于其独特优势，集体反思正逐渐应用于教师教育中（Hong et al., 2019; 初胜华和张卫东，2014）。通过集体反思，教师有机会重新审视自己的教学实践（Akyel, 2000），通过分享和交流彼此的教学实践，教师对教学实践产生新的认识，从而提高教学效果（Dyke, 2006）。然而，随着教师教育领域集体反思研究的增多，许多研究者发现，虽然职前教师在实际进行集体反思时能够得到反思性思维广度的支持，但由于职前教师教学经验有限，他们的反思往往较为肤浅，难以进行更深入和批判性的反思（Mcgarr, Mccormack, & Comerford, 2019; Shin, 2021）。如何更有效地发挥集体反思的作用，提高职前教师反思的深度仍需要思考。虽然关于职前教师教育中集体反思的研究日益增多，但针对职前科学教师的集体反思研究仍相对有限。

基于视频的自动多模态分析技术是视频分析发展的第三阶段，前两个阶段分别是基于视频的分析和基于视频的手动多模态分析。视频分析的发展从最初的观察和手动分析，到利用多模态分析工具进行手动编码，再到借助多模态支持下的智能编码分析。视频分析的每一次发展都帮助教师进一步理解和改进自己的教学实践。在第一阶段，基于视频的分析主要是在教师教学完成后作为回顾反思的工具（Harford & MacRuairc, 2008）。教师只能执行最基础的操作，如播放和暂停视频，但在访问视频的特定时刻和进行详细编码方面存在挑战，因此教师对视频的分析相对有限。在第二阶段，基于视频的手动多模态分析主要使用手动编码的多模态分析工具来分析视频。教师在此过程中可以对视频进行更多操作，如逐帧分析、使用标注工具以及标记视频的特定片段（He & Pao, 2020; Fang et al., 2023）。此外，多模态分析利用机器学习方法来分析和挖掘视频中言语和非言语行为的各个方面，如停顿、目光或肢体语言（Ding, Glazewsk, & Pawan, 2022），从而使教师能够对教学视频进行更详细和全面的分析。然而，手动编码视频分析非常耗时，不利于在日常教学实践中普及使用。第三阶段的基于视频的自动多模态分析技术是人工智能技术快速发展的产物。它利用视频录像作为输入，机器能够自动识别和分析教师与学生的话语、语音情绪、面部表情、身体动作及动作时机等多个方面（Mingzhe et al., 2020）。由于视频是自动编码的，视频编码的时间大大缩短，这使得教师可以更高效地分析大规模的教学视频，理解自己的教学风格、学生的学习状态以及课堂气氛，发现教学实践中的各种问题，并进行更有针对性的改进。作为视频分析技术的最新发展，基于视频的自动多模态分析技术有助于教师全面、具体、客观地分析和理解教学，在教师教育领域具有强大的应用潜力。

因此，本研究旨在探讨基于视频的自动多模态分析技术在促进职前科学教师集体反思中的实际应用效果。我们希望通过这一技术的引入，能够协助职前科学教师更深入地剖析自己的教学实践，发现潜在的问题，并找到改进的方向。同时，我们也期望通过这一研究，为职前教师教育的创新提供新的思路和方法，推动教师教育的现代化和专业化发展。

3.研究方法

3.1. 研究情境

本研究关注的是华中地区一所研究型大学科学教育协会的一群职前科学教师所进行的集体反思。应《科学教育》的要求，同时出于职前教师的信念和责任感，科学教育协会的学生们自发地在一所薄弱小学开展了支教活动。每次在 J 小学教授完他们设计的校本科学课程后，这些职前教师们会聚集在学校会议室进行集体反思。在参与者知情且同意下，本研究使用了他们的集体反思视频共 8 次。前三次反思是由职前教师独立进行的，后五次由职前教师参考多模态分析报告进行。转录后的文本随后进行统一编码和分析。本研究将职前科学教师集体反思分为两个阶段，即(1)设计阶段，侧重于教学设计和模拟教学；(2)实践阶段，侧重于对课堂教学实践的反思，该阶段融入了反思视频的多模态分析报告，以提高职前科学教师的批判性思维及对教学的反思更加深入。以视频录像为输入，机器可以自动识别并分析师生的互动、教师的面部表情、注意力状态、身体姿势、言语情绪等多个方面（D'Mello, Dieterle, & Duckworth, 2017）。集体反思的第一阶段持续了近两个月，第二阶段持续了三个月。参与者每周或每两周进行一次集体反思，反思内容集中在对上周支教实践的自我评价、支教实践的优势和不足以及改进建议上。

3.2. 参与人员

每次参与集体反思的人数并不固定，一般在 7-12 人，但有 5 位同学每次都会参与并积极发言。他们都来自中国中部一所著名的研究型大学。并且他们参加此项目完全是自愿的。在加入本次自愿教学之前，他们都没有任何 K-12 教学经验，也没有从事过与 K-12 教学相关的反思活动。

3.3. 数据的收集与分析

表 1 教学反思认知情绪编码系统

认知状态					情绪状态			
其他	触发	拓展	整合	解决	其他	积极	消极	困惑
0	1	2	3	4	0	1	2	3

本研究是在参与者知情同意的情况下进行的，首先我们录制了他们集体反思的视频，以记录每次职前科学教师教授完课程的情况。随后反思中的口头对话被转录成书面形式。本研究采用的研究方法受到在线互动讨论环境中广泛使用的探究社区(CoI)模型的启发（Shea et al., 2012）。CoI 模型将认知存在分为四个不同的阶段：触发、探索、整合和解决（马志强、刘亚琴和孔丽丽，2018）。由于反思会议的频率高，持续时间长，最终转录的文本将近 12 万字，因此我们使用了在业界得到广泛认可的预训练语言模型 BERT 进行数据的处理和编码（Liu et al., 2023）。BERT 可以利用大量未标记的文本数据来捕获丰富的上下文信息，避免对注释较少的数据进行过拟合（余同瑞等人，2020）。BERT 已被许多研究人员引入到各研究领域中，且都取得了良好效果。例如，Lee 等人开发了用于生物医学文本挖掘的 BioBERT（Lee et al., 2020）。Beltagy 等人提出 SCIBERT 来提高科学评价的绩效（Beltagy, Lo, & Cohan, 2019）。本研究使用了由刘智等人（Liu et al., 2023）开发的 BERT，该 BERT 已被用于各种研究，可通过不同语境下的讨论文本来识别学习者的认知存在水平。我们用数字“0~4”对认知的五种状态进行顺序编码。同样采用数字“0~3”来系统编码四种情绪状态，如表 1 所示。研究使用认知网络分析(ENA)可视化分析了编码数据，有助于直观地探索职前教师在集体反思过程中

认知和情感方面的变化。在本研究中，第一阶段的反思记录了职前教师在学校实际教学前所进行的教学设计。第二阶段的反思则是职前教师在 J 小学结束教学后的反思。

4.研究结果及分析

4.1. 第一阶段中认知和情绪变化情况

图 1 展示了最初三次反思期间的认知转变，图 2 展示了同一时期的情绪变化。在图 1 中，触发、扩展和其他之间有更紧密的联系，这表明在此阶段职前教师主要在家讨论课程设计，从而促进了他们想法的产生并进行扩展。从图 2 可以看出，消极与主动、积极与困惑之间的联系更加紧密，这说明职前教师在课程设计中可能遇到了一些困难，他们原本的积极心态变得很困惑。并且职前教师的情绪在消极和积极之间波动，很不稳定，这可能会影响反思的参与度，这可以通过参与者分享的“我认为上节课几乎没有什么有价值的东西”，“现在看来完全没有教室的气氛。”来支持。很明显，在认知方面，职前教师还没有进入更深层次，而是更倾向于关注拓展与反思无关的领域。第一阶段职前教师的情绪主要处于消极和困惑，这可能会削弱职前教师的信心，降低他们参与反思的意愿。有些人用“我也觉得很尴尬，不知道该说什么”，“我没有任何想法或建议。”这样的方式表达他们的消极情绪。

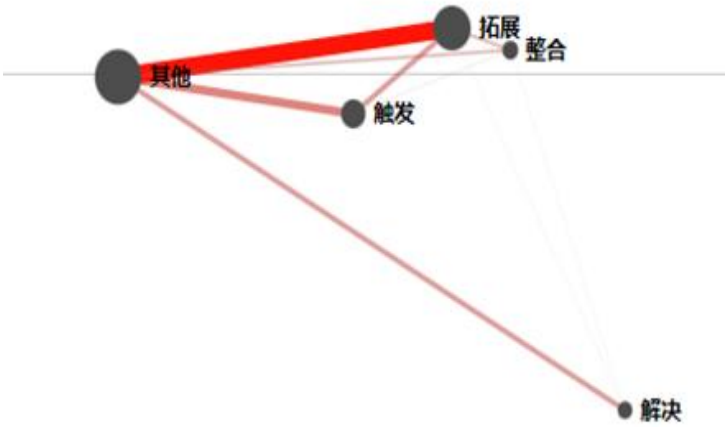


图 1 第一阶段的反思认知图

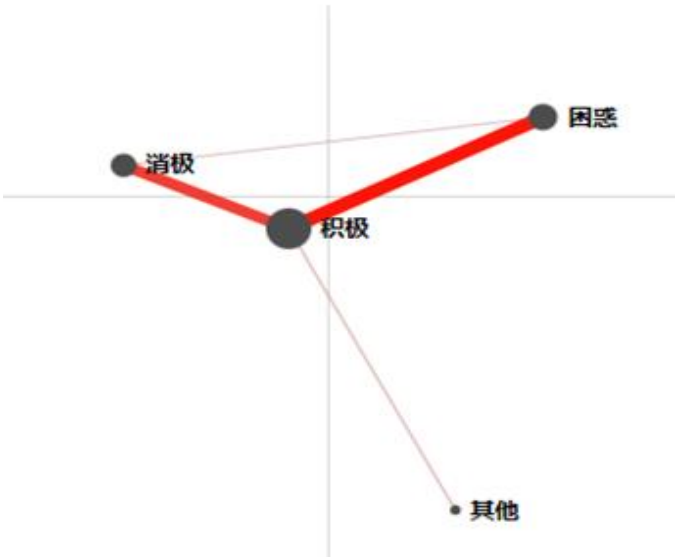


图 2 第一阶段的反思情绪图

4.2. 第二阶段中认知和情绪变化

图 3 显示了在加入多模态反思报告后，后续五场反思的认知变化。整合了反思报告的后五次反思期间的情绪变化如图 4 所示。图 3 中触发、扩展和整合之间的连接非常紧密。相比于

第一阶段的扩展和触发，第二阶段的整合非常重要，这表明他们在反思后会进行统一分类和整合。这可以从参与者的分享中得到支持：“我们的教学设计没有很好地考虑到学生，我们对学生的预期反应也没有那么准确，我们没有考虑到他们的真实情况”。由此可见，职前教师在第二阶段进行了深刻的反思，认知也达到了更高的水平。在图 4 中，消极、积极和困惑之间有更密集的连线。由于各位职前教师是第一次进行实际教学，课堂中可能会有一些突发情况，从而导致职前教师产生一定程度的挫败感，因此可能会出现轻微的负面情绪。但随后进行的集体反思和头脑风暴中，大家提出了各种解决方案和自己的想法，整体氛围还是较为轻松的。这可以从职前教师的话中体现，比如“我认为你可以放慢节奏，必要时在黑板上写下来”，“我觉得我们遇到的很多困难都是新手老师必经的，其实这很正常”。同样值得注意的是，在第二阶段，职前教师的活跃程度显著高于其他两种情绪。以多模态分析报告为参考，职前教师反思更有目的性和规范性，参与度更高，整体氛围也更好。这一点也能很好地被参与者的话所支持，比如“这样的话，下节课我就会更清楚了”，“我明白了，我认为这应该....”。

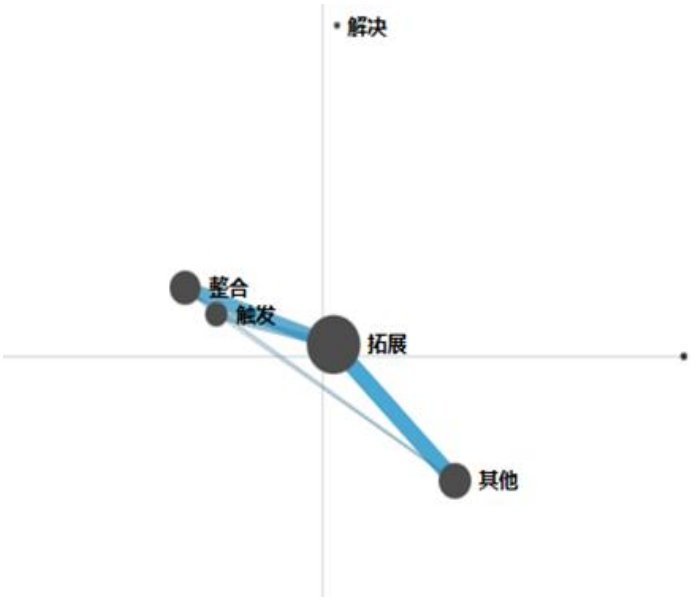


图 3 第二阶段的反思认知图

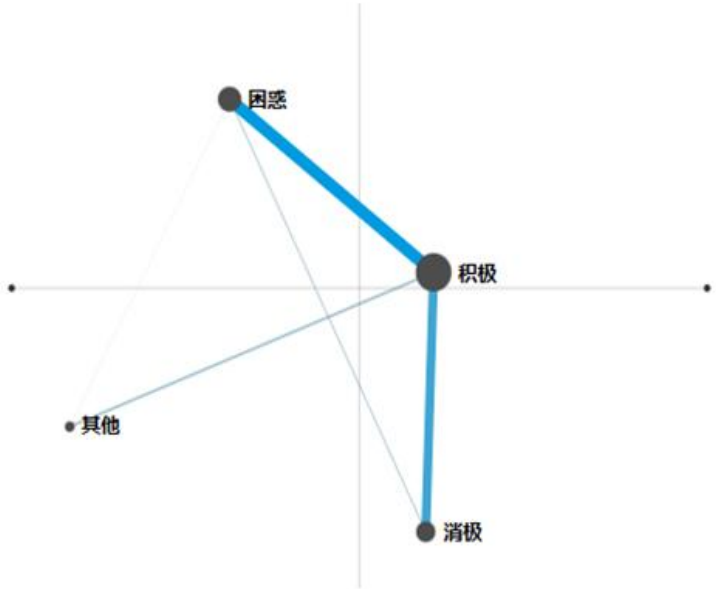


图 4 第二阶段的反思情绪图

5.讨论

5.1. 集体反思有利于实现深度反思

在本研究中，义务教学实践的协作任务是一种通过集体反思来培养职前科学教师批判性思维和高阶认知过程的手段。这种形式的集体反思可以被视为一种协作过程，用于集体识别并解决个别职前科学教师可能发现但难以独立解决的挑战。集体反思有助于参与者沟通想法，提高高阶思维能力，并丰富他们提出问题、分享不同观点、寻找遗漏信息的机会，促进小组成员之间进行讨论。

5.2. 作为证据的多模态分析报告“驱动”了集体反思

与肤浅的反思或内省相比，批判性反思的目的不仅是了解教学行为和情境的意义，而且是了解支配教师行为和信念的潜在假设和前提，这就对教师提出了更高的要求。以往的研究表明，基于证据的反思往往更深入，因此业内专家都认可基于证据的反思。本研究对这一点进行了延伸和深化，多模式报告是职前教师深入思考的重点，它作为证据推动了反思的深化。使用基于视频的多模态分析技术可以帮助职前教师更好地理解正式教师的行为和信念，促进职前科学教师进行彻底且批判性的反思。

5.3. 多模态分析报告可以活跃反思的气氛

当没有专业报告指出当前存在的问题时，职前教师在反思时可能会更加迷茫，缺乏目的性。在整合了多模态分析报告后，每个人都可以了解自己的问题和上次反思中存在的问题，使反思更加容易进行且高效。大家一起分享经验，共同应对挑战，不仅能给彼此提供情感支持，还能增强职前教师团队的凝聚力和责任感。

6.启示

基于集体反思过程中出现的问题以及对结果的研究，对于改进集体反思本研究得出了以下启示。首先，文献普遍同意反思必须基于具体证据，如实践经验或数字报告。本研究中使用的多模态报告从数据科学的角度提供了一个独特的视角，值得借鉴。其次，基于视频的多模态分析报告提供了一个客观且全面的教学过程概览，这有利于促进职前教师进行批判性反思。为了确保能获得对该分析报告的全面理解，这些报告应该与引导者的解释一起纳入反思过程，这种整合可以让所有职前教师理解这些报告，并使得他们将该反思与个人反思进行有意义的比较。

7.局限性

本研究也存在一些局限。首先，参与者的数量较少，而且他们都来自同一所大学，因此结果的泛化性较为有限。研究结果可能仅适用于大型研究型大学的职前教师群体。其次，考虑到课堂效果的真实性，没有设置对照组作为比较来检验职前科学教师集体反思前后的学习效果。第三，本研究仅分析了职前科学教师的集体反思，并没有将其扩展到其他学科或其他情境。

8.结论

本研究采用认知网络分析(ENA)可视化职前科学教师的认知和情感状态，研究将整合了基于视频的自动多模态分析报告的集体反思与没有该报告的集体反思进行对比，旨在探索基于视频的自动多模态分析技术在促进职前科学教师集体反思中所起的作用。研究结果表明，整合了多模态分析报告的集体反思不仅可以使职前科学教师进行更深入、批判性的思考，还能提高职前科学教师的反思参与度，使反思的整体氛围和谐轻松。但也存在一定的挑战，如如何使职前教师在过程中注意力更聚焦于反思，减少其他行为的发生。在未来，我们将考虑进行手动编码，与 BERT 编码的结果进行比较，以验证 BERT 的成熟度。此外，我们也在准备使用其他编码框架进行手动编码，关注其他课题的认知发展。

参考文献

马志强、刘亚琴和孔丽丽(2018)。网络探究学习社区理论与实证研究发展脉络。现代远程教育研究（03）,39-48。

王晓生(2023)。小学科学教师队伍建设：价值使命、现实羁绊与实践路径。中国教育学刊(06),91-95。

余同瑞,金冉,韩晓臻,李家辉和郁婷(2020)。自然语言处理预训练模型的研究综述。计算机工程与应用，(23),12-22。

初胜华和张卫东(2014)。基于教师合作反思教学模式的高校外语教师专业发展研究。外语与外语教学(01),67-72。

张晓光(2020)。芬兰教师入职教育中的同伴小组指导:理论、模式、条件与效果。比较教育研究(08)，89-96。

季娇、伍新春和燕婷(2014)。探索职前科学教师专业发展的新途径——非正式学习情境的促进作用。课程.教材.教法(03),106-112。

Akyel, A. (2000). Collaboration to Explore Teaching: A Case Study Report. TESL Canada Journal, 18, 58-74.

Arefian, M.H. (2022). Collaborative action research as a reflective tool for pre-service EFL teachers' inclusion. Reflective Practice, 23, 651-662.

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A pretrained language model for scientific text. arXiv preprint arXiv:1903.10676.

Davidsen, J., & Vanderlinde, R. (2014). Researchers and teachers learning together and from each other using video-based multimodal analysis. Br. J. Educ. Technol., 45, 451-460.

Ding, A.E., Glazewski, K.D., & Pawan, F. (2022). Language teachers and multimodal instructional reflections during video-based online learning tasks. Technology, Pedagogy and Education, 31, 293-312.

Dixon, M., Lee, C.J., & Corrigan, C. (2021). 'We were all looking at them quite critically': collaborative reflection on a university-based coach education program. Reflective Practice, 22, 203-218.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. Educational psychologist, 52(2), 104-123.

Dyke, M. (2006). The role of the 'Other' in reflection, knowledge formation and action in a late modernity. International Journal of Lifelong Education, 25, 105-123.

Epler, C.M., Drape, T., Broyles, T.W., & Rudd, R.D. (2013). The Influence of Collaborative Reflection and Think-Aloud Protocols on Pre-Service Teachers' Reflection: A Mixed Methods Approach. Journal of Agricultural Education, 54, 47-59.

Fang, Y., Wu, L., Jumelle, F., & Shi, B.E. (2023). Integrating Holistic and Local Information to Estimate Emotional Reaction Intensity. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 5934-5939.

Harford, J., & MacRuairc, G. (2008). Engaging student teachers in meaningful reflective practice. Teaching and Teacher Education, 24, 1884-892.

He, J., & Pao, H.K. (2020). Multi-modal, Multi-labeled Sports Highlight Extraction. 2020 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), 181-186.

Hong, H., Lin, P., Chai, C.S., Hung, G., & Zhang, Y. (2019). Fostering design-oriented collective reflection among preservice teachers through principle-based knowledge building activities. Comput. Educ., 130, 105-120.

Kager, K., Jurczok, A., Bolli, S., & Vock, M. (2022). “We were thinking too much like adults”: Examining the development of teachers’ critical and collaborative reflection in lesson study discussions. Teaching and Teacher Education.

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 36(4), 1234-1240.

Liu, Z., Kong, X., Chen, H., Liu, S., & Yang, Z. (2023). MOOC-BERT: Automatically identifying learner cognitive presence from MOOC discussion data. IEEE Transactions on Learning Technologies.

Martin, G.A., & Double, J.M. (1998). Developing Higher Education Teaching Skills Through Peer Observation and Collaborative Reflection. Innovations in Education and Training International, 35, 161-170.

Mcgarr, O., Mccormack, O., & Comerford, J. (2019). Peer-supported collaborative inquiry in teacher education: exploring the influence of peer discussions on pre-service teachers’ levels of critical reflection. Irish Educational Studies, 38, 245-261.

Mingzhe, L., Chen, X., Gao, S., Chan, Z., Zhao, D., & Yan, R. (2020). VMSMO: Learning to Generate Multimodal Summary for Video-based News Articles. ArXiv, abs/2010.05406.

Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Pickett, A., ... & Jian, S. (2012). Learning presence: Additional research on a new conceptual element within the Community of Inquiry (CoI) framework. The internet and higher education, 15(2), 89-95.

Shin, J. (2021). Preservice music teachers in Korea and their collaborative reflection with peers. International Journal of Music Education, 39, 371-382.

Sorensen, P. (2014). Collaboration, dialogue and expansive learning: The use of paired and multiple placements in the school practicum. Teaching and Teacher Education, 44, 128-137.

Subramaniam, K. (2013). Examining the Content of Preservice Teachers’ Reflections of Early Field Experiences. Research in Science Education, 43, 1851-1872.

Tang, Y., Zhang, S., Sun, M., Wen, Y., An, S., & Liu, Q. (2023). Understanding student teachers’ reflective thinking using epistemic network analysis and fine-grained trace data. Thinking Skills and Creativity.

Wilson, S.M. (2013). Professional Development for Science Teachers. Science, 340, 310-313.

工程翻转学习情境下基于过程挖掘的学生自我调节学习行为分析

Analysis of Students' Self-Regulated Learning Behavior Based on Process Mining in Engineering Flipped Learning Environment

陈香均[1]，龙陶陶[2*]，朱晓萌[3]，伍冬[4]
[12]华中师范大学人工智能教育学部
[34]武汉理工大学材料科学与工程学院
[*] taotaolong@ccnu.edu.cn

【摘要】 自我调节学习（SRL）在线上学习、翻转学习等学习情境中扮演着重要角色，因此，探索学习者的自我调节学习行为模式备受教育研究工作者的关注。尽管自我调节学习是当前研究的热门主题之一，但较少研究关注工程翻转课程中学生的自我调节学习行为模式。本研究旨在探讨在以基于问题的学习（PBL）为导向的工程翻转课堂中，学生的自我调节学习行为模式。通过对 177 名学生学习行为数据进行过程挖掘，发现学生的自我调节学习行为之间的依赖关系与 Pintrich 提出的自我调节学习行为模型有所差异，且学生在课程里表现出不同的学习路径。

【关键词】 过程挖掘；自我调节学习；学习行为；工程翻转课堂

Abstract: *Self-regulated learning (SRL) plays an important role in various learning situations such as online learning and flipped learning. Exploring SRL behavior patterns has attracted the attention of many educational researchers. Although SRL is one of the hot topics in current research, few studies focus on the SRL behavior patterns of students in engineering flipped courses. This study aimed to explore students' SRL behavior patterns in a flipped engineering course of which the in-class learning was conducted in the way of Problem-Based Learning. Based on the process mining of 177 participants' learning behavior data, the results showed that the dependence relationship between students' SRL behavior was not completely consistent with the Pintrich's SRL model, and students had different learning paths.*

Keywords: Learning behavior, SRL, Process mining, Flipped engineering classroom

1.引言

自我调节学习能够促进学习者高效学习，是 21 世纪的关键技能之一(Manukaram, K., & Abdullah, M. N. L. Y., 2021; Du, J., Huang, W., & Hew, K. F., 2021)。自我调节学习是指学习者可以控制自己的学习过程，并且可以调节自身认知、动机、行为等方面来实现特定的学习目标（Loyens, S. M., Magda, J., & Rikers, R. M., 2008）。已有研究表明，自我调节学习能够影响学生的学习投入度(杨燕清、吴明、林艳芹和洪霞, 2022)与学习成绩(李文昊、任晓曈、朱希雅和陈冬敏, 2022)，特别是在在线教育中，自我调节学习发挥的作用显著。后疫情时代下，翻转课堂作为一种线上线下结合的混合式教学模式备受关注。翻转课堂强调学生需要在课前进行在线资源的学习(张金磊、王颖和张宝辉, 2012)，这就要求学生具备一定的自我调节学习能力。近年来，教育领域中，与自我调节学习主题相关的研究主要集中在研究促进学生高效学习和应用自我调节学习技能的策略(李运福, 2021; 徐晓青、赵蔚、刘红霞和姜强, 2021)，以及研究影响学生自我调节学习行为的因素上(An, F., Xi, L., & Yu, J., 2023)。然而，在大多数研究中，学生的自我调节学习行为过程性的数据收集未能被重视，数据收集方式较为单一，主要来源

于学生的自我报告，这使得最终生成的学生自我调节行为报告不可避免的具有较强的主观性。

过程挖掘，又称工作流挖掘，是一种从系统内部的事件日志中提取信息，用于发现、监控和改善实际运作流程的方法(Van Der Aalst, W., 2011)。当前，过程挖掘技术已被广泛应用于教育研究中(邓国民、徐新斐和朱永海, 2021)。然而，采用过程挖掘技术分析工程类本科生自我调节学习行为的研究尚属罕见。并且，较多研究关注的学习场景通常为线上学习情境，如：MOOC，对翻转课堂中的自我调节学习关注较少。翻转课堂在工程类教学中应用较广，但也有研究者指出其课前阶段存在诸多问题。

本研究在学生自我调节学习行为视域下，使用过程挖掘技术分析了工程类翻转课堂中学生的学习行为，以期为探索和优化工程类翻转课堂教学提供基于客观数据的参考。

2.文献综述

2.1. 自我调节学习(Self-Regulated Learning, SRL)

自我调节学习源自于班杜拉提出的社会认知理论，随后，Zimmerman 等学者在此基础上对自我调节学习进行了具体阐释。Zimmerman(1989)认为，自我调节学习是指学习者能够从元认知、动机和行为方面积极参与学习过程，主动获取知识与技能，并强调自我调节策略、自我效能感和学业目标是自我调节学习的三个要素(Zimmerman, B. J., 1989)。随着研究的不断深入，关于自我调节学习的模型也逐渐丰富。Zimmerman(2009)认为自我调节学习应具有三个过程，即计划阶段、执行阶段、自我反思阶段。Pintrich (2004)则将自我调节学习划分为四个阶段：预见、计划和激活；监测；控制；反应和反思。

目前的研究已运用多种方法来调查学生的自我调节学习行为，如：收集反思性日记、半结构性访谈(Hromalik, C. D., & Koszalka, T. A., 2018; Liu, R., Zhu, J., LI, W., Ju, T., & Chen, L., 2022)。学生的自我调节行为虽能以此被反映出，但却难以直观呈现。除此，研究者们也较为关注自我调节学习在翻转课堂中的应用，如： van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2020)探索了在翻转课程里为学生提供自我调节学习支持后，学生学习成果的变化。然而，与线上学习情境相比，仅有较少研究关注翻转课堂情境中学生的自我调节学习的过程，且其研究数据主要来源于学生的自我报告，因而，难以做到较为客观。

此外，许多研究已表明自我调节学习对于学生在编程等学科课程的学习方面有积极影响(Öztürk, M., 2022)。然而，关注自我调节学习应用于工程类课程的研究却较少，且已有的研究主要探索中小学生 STEM 过程中的自我调节学习情况(Li, S., Chen, G., Xing, W., Zheng, J., & Xie, C., 2020)。在高等工程教育背景下，对工程类本科生的自我调节学习行为研究尚属罕见。

2.2. 工程翻转课堂

作为近年来的热门议题之一，翻转课堂是一种学生在课前独自学习视频、文本等资源，课堂中再与教师、同学互动交流的教学模式(Bergmann, J., & Sams, A., 2012; Kim, M. K., Kim, S. M., Khera, O., & Getman, J., 2014)。研究表明，翻转课堂能够促进学生学业进步以及创新思维等高阶技能发展(王翠如和胡永斌, 2018; Rodríguez, G., Díez, J., Pérez, N., Baños, J. E., & Carrió, M., 2019)。将翻转课堂融入高等工程教育中已被证明是一种有效促进学生学习的方法(孟军、刘冰璇、翟洪江和兰淇, 2018)，然而，Nielsen, K. L. (2023) 指出在使用翻转课堂的过程中，学生可能会感到挫败与沮丧。

为提高工程翻转课堂的实施效果，研究者们尝试将项目式学习(Chua, K. J., & Islam, M. R., 2021)、基于问题的学习(Gómez-Coma, L., Díaz-Sainz, G., Fallanza, M., Ortiz, A., & Ortiz, I., 2023)等策略融入到教学设计中，且相关研究结果表明这些策略对于促进学生学习是有效的。与其他学习策略相比，基于问题的学习导向的工程翻转课堂强调构建工程实践中的真实问题(Chen, X., Long, T., Cheng, L., Gan, X., & Zhu, X., 2023)，在引导学生解决工程实践问题的过程中，提

高对专业知识的理解并促进工程能力的发展。

2.3. 过程挖掘

过程挖掘自 20 世纪末被提出以来，发展迅速。过程挖掘通过利用系统中的事件日志数据来发现操作者与系统的具体交互过程(徐晓青、赵蔚、姜强、刘红霞和乔丽方, 2022)。过程挖掘通常被认为包含三种类型：过程发现，一致性检查和过程增强(黄琰、赵呈领、赵刚、刘军、疏凤芳和李红霞, 2020)。近年来，过程挖掘被视作一种新技术来探索学生的自我调节学习行为，例如，乔丽方、赵蔚、段红和徐晓青(2022)使用过程挖掘技术探索不同类型的学习者的自我调节学习过程差异。过程挖掘技术能够直接分析学生真实的在线的学习日志，从而将学生的自我调节学习较为全面的可视化的进行呈现。然而，在现有的研究里，研究者们更常运用滞后序列分析、分层聚类分析等方式分析学生的学习行为(Chen, W., & Cheng, H. N., 2020; Han, F., & Ellis, R. A., 2023)，而不是应用过程挖掘技术发现学生的学习行为。

2.4. 本研究

尽管一些研究已经使用了过程挖掘技术探索在线学习情境下学生的自我调节学习行为，但进一步使用过程挖掘技术探索不同类型课堂里学生的自我调节学习行为仍有价值。本研究旨在运用过程挖掘技术调查 pbl 导向的工程翻转课堂中学生的自我调节学习行为，以期为解决当前工程翻转课堂中学生自我调节学习行为不明晰的问题提供一些参考。

3.方法

3.1. 背景

本研究在华中地区的某工程高校进行。研究团队选取《计算机在材料科学与工程中的应用》这门工程专业课程进行翻转课堂教学实践，并根据基于问题的学习策略设计了课程活动。课程共有 4 个 pbl 活动，且都含有在线学习与课堂学习环节。所有学生需在课前登录“小雅”在线学习平台完成相关学习任务，接着，在课堂中与教师和同伴进行交流与互动。

3.2. 参与者

本研究共选取 177 名工程专业的大三学生。这些学生自入学以来，没有参与过工程翻转课堂、系统的基于问题的学习活动以及从未接受过任何与自我调节学习相关的训练。

3.3. 工具

表 1 SRL 阶段与学习行为事件对应表

SRL 行为阶段	学习行为事件	编码	学习行为的具体解释
预见、计划和激活	总览大纲	V-S	浏览课程大纲或课程章节
	生成计划	G-P	编辑节点、更新资源生成学习计划
	分析任务	A-T	打开任务导航栏或任务导图
监测	观看资源	V-R	查看具体学习资源，如：视频、文档等
	查看任务	V-T	查看任务说明或任务提交界面
	移动资源	M-R	整理与移动小组资源
	下载资源	D-R	下载任务数据
控制	执行任务	P-T	打开任务资源，执行任务
	提交作业	S-W	提交 4 个 pbl 学习任务、小测和思考题
反应和反思	查看论坛	V-F	查看论坛或在论坛发表观点
	提交反思	S-R	提交个人反思报告
	查看作品	V-W	查看其他同学的优秀作品

采用 Pintrich (2004)的四阶段自我调节学习行为模型来对学生的学习行为进行分类，即：预见、计划和激活；监测；控制；反应和反思。同时，在自我调节学习理论视域下，对学生的

在线学习行为进行了微观层面的划分，得到 12 种学习行为事件，如表 1 所示。

3.4. 数据收集和分析

研究团队收集了所有课程参与者在学习平台上产生的事件日志数据，共计 67,105 条。每个事件日志都包括了参与者的交互行为以及一些其他必要的信息，如：id、时间戳。表 2 呈现了学生观看视频的事件日志数据的摘录。在分析数据之前，对所得数据进行预处理，排除了一些与研究无关的行为事件数据，如：学生添加课程、签到，最终获得数据 62,276 条数据。使用 ProM 6.12 对事件日志数据进行过程挖掘分析。

表 2 事件日志的示例

Student Id	Detailed activity	Timestamp
12345	Play video	2022-09-29T11:03:12.820Z
12345	Play video	2022-09-29T11:03:20.859Z
12345	Play video	2022-09-29T11:03:25.672Z
12345	Play video	2022-09-29T11:03:27.698Z
12345	Play video	2022-09-29T11:03:12.820Z

4.结果

对所收集的数据主要运用 ProM 软件中基于启发式矿工算法概念并融合了分类技术的 iDHM 工具中的 Dependency Graph 过程模型进行分析。

4.1. 学生在 pbl 导向的工程翻转课堂中的自我调节学习行为

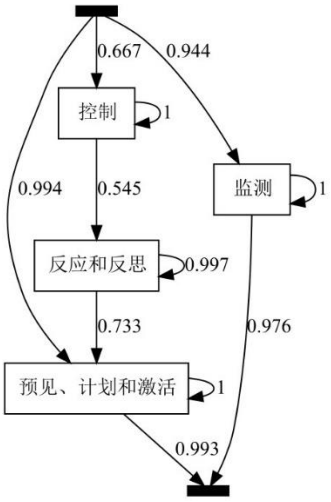


图 1 SRL 阶段 Dependency Graph 过程模型

如图 1 所示，Dependency Graph 过程模型被用于挖掘学生的自我调节学习行为过程，其结果显示了四个 SRL 阶段之间的依赖关系。图 1 中的矩形表示行为事件，其中，两个黑色矩形分别代表开始与结束。矩形之间的有向箭头表示事件之间发生的顺序关系，箭头上的值表示事件之间的依赖强度。正如图 1 所示，在开始学习之后，多数人会开展预见、计划和激活活动（依赖度=0.994）或监测活动（依赖度=0.944），随后直接结束学习活动。还有一部分学生在开始学习后，会开展控制活动（依赖度=0.667），随后进行反应和反思活动，其中，控制活动与反应和反思活动之间的依赖度为 0.545，接着，学生再进预见、计划和激活活动（依赖度=0.733），最后结束学习。

4.2. 自我调节学习视域下，学生在 pbl 导向的工程翻转课堂中的具体学习过程

如图 2 所示，Dependency Graph 过程模型被用于分析 SRL 视域下学生在 PBL 导向的工程翻转课堂中的学习行为过程。从依赖关系来看，以下载资源（D-R）为出发点，总览大纲（V-

S）为结束点，清晰的展示了五条路径。(a)第 1 条路径是“下载资源（D-R）→总览大纲（V-S）”,这两个行为事件之间存在很强的依赖关系，依赖强度为 0.935。(b)第 2 条路径是“下载资源（D-R）→提交作业（S-W，依赖度为 0.700）→查看作品（V-W, 依赖度为 0.875）→总览大纲（V-S, 依赖度为 0.792）”。(c)第 3 条路径是“下载资源（D-R）→提交作业（S-W，依赖度为 0.700）→执行任务（P-T, 依赖度为 0.800）→提交反思（S-R, 依赖度为 0.948）→总览大纲（V-S, 依赖度为 0.817）”。(d)第 4 条路径是“下载资源（D-R）→生成计划（G-P，依赖度为 0.667）→移动资源（M-R, 依赖度为 0.288）→提交反思（S-R, 依赖度为 0.750）→总览大纲（V-S, 依赖度为 0.817）”，其中，生成计划（G-P）和移动资源（M-R）之间的依赖关系较弱，两者直接的依赖度仅为 0.288。(e)第 5 条路径是“下载资源（D-R）→生成计划（G-P，依赖度为 0.667）→查看论坛（V-F, 依赖度为 0.462）→提交反思（S-R, 依赖度为 0.667）→总览大纲（V-S, 依赖度为 0.817）”。

由图 2 可知，查看任务（V-T）与其他学习行为活动事件没有绝对的依赖关系。总览大纲（V-S）与分析任务（A-T）之间的依赖强度为 0.667，而分析任务（A-T）与观看资源（V-R）之间的依赖强度为 0.750。相比之下，观看资源（V-R）和下载资源（D-R）之间的依赖关系较弱，依赖度为 0.339，这表明下载资源事件的发生对观看资源事件的依赖性不强。

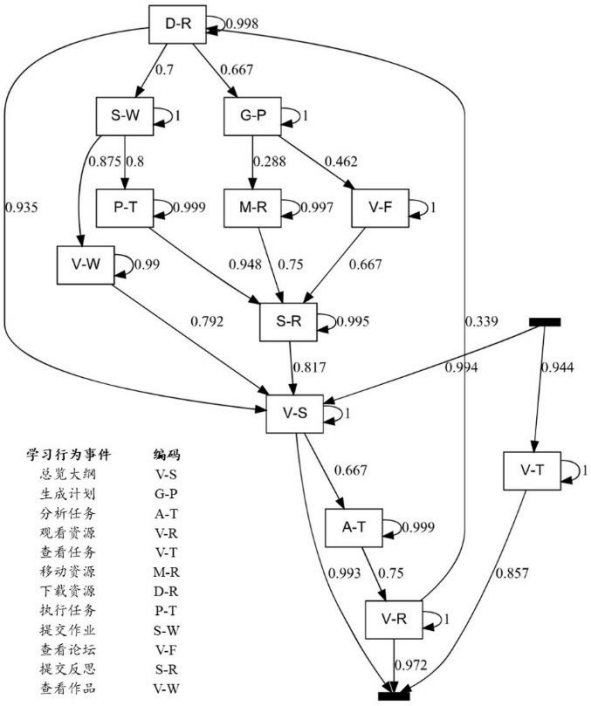


图 2 自我调节视域下学生工程翻转课堂学习行为 Dependency Graph 过程模型

5.讨论和启示

本研究运用过程挖掘方法探索了自我调节学习视域下，学生在 PBL 导向的工程翻转课堂中的学习行为过程。研究发现，学生的自我调节学习行为在工程翻转课堂中与纯在线课堂有差异，即 SRL 的四个阶段之间未能显示出“预见、计划和激活→监测→控制→反应和反思”的依赖关系。然而，这并不代表在本研究中，学生的四种自我调节学习行为之间没有相关关系。除此，研究还发现，学生在课前学习阶段的学习路径呈现出多样化的特点。

首先，在学生的自我调节学习行为中预见、计划和激活阶段和监测阶段之间以及监测阶段与控制阶段之间没有呈现出顺序依赖关系。这与 Pintrich (2004)制定的自我调节学习周期有所差异。Pintrich 提出的自我调节学习的四个阶段是按照“预见、计划和激活→监测→控制→反应和反思→预见、计划和激活”顺序进行循环的。其原因可能是，本研究是在翻转课堂情境

下开展的，学生在教师线下的指导下就解决了学习过程中遇到的一些问题，从而导致学生不会在线上学习中表现出预见、计划和激活阶段与监测阶段、控制阶段强烈的依赖性。在翻转课堂中，教师的干预会影响学生课前的自主学习，在本研究中，学生会定期参与课堂学习，与教师交流，从而可能出现这种现象。此外，学生对翻转课堂的学习模式不熟悉，特别是对智能学习平台操作不熟练以及对融合 PBL 策略的工程活动不熟悉，使得学生的自我调节学习表现不佳，从而难以顺利完成自我调节学习周期循环。已有研究证实，学生在未适应翻转课堂学习时，其课前自主学习的动机较低且容易出现学习习惯不良的情况(Nielsen, K. L., 2023)。

其次，在课程中，学生的学习行为路径呈现出了明显的差异。工程翻转课堂能够增强学生对工程内容的主动学习，并支持学生的自主学习(Cho, H. J., Zhao, K., Lee, C. R., Runshe, D., & Krousgrill, C., 2021)。本研究为学生创建了智能学习空间，如：提供了丰富的在线学习资源和基于 PBL 策略的活动任务，为学生的自主学习提供了条件，这可能是学生出现不同学习路径的原因之一。从局部的学习行为来看，学生观看资源与下载资源的行为之间没有较强的依赖关系，可能是因为大部分学生在观看资源后并未计划立即完成学习任务，因此没有下载资源。研究结果还显示，多数学生在查看任务或总览大纲后会直接结束学习。这与以往的研究相似，Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018)在挖掘 MOOC 中学习者的学习行为时发现，有一部分学习者只停留在形式上的查看内容，没有完成任何任务的目标，这可能与学习者需要花费时间与精力来理解课程有关。

基于研究结果，提出以下建议，以便教育工作者或者相关研究人员进一步改善工程翻转课堂。首先，为学生提供种类、数量足够丰富的学习资源，从而支持学生的自主学习。值得注意的是，在开展翻转课堂之前，可以对学生进行一段时间的培训，确保学生能够熟练使用智能平台或能够掌握一定的 SRL 策略。此外，过程挖掘技术被应用于发现学生的学习行为过程，这使得学生的学习过程能够被可视化呈现。

6.结论和局限性

本研究探讨了 SRL 视域下，PBL 导向的工程翻转课堂中学生的学习行为路径。采用过程挖掘技术分析了学生的行为日志事件，结果表明，学生的自我调节学习行为之间的依赖关系与 Pintrich 的 SRL 模型中的循环顺序不完全一致，且学生在课程里出现了不同的学习路径。

这项研究存在几个局限性。首先，研究仅选择了一门课程的数据进行收集，其代表性可能不足。此外，过程挖掘算法较多，研究只选择了一种算法，可能不太全面。未来的研究还可以选用更多的工程课程来设计翻转课堂，从而拓展数据收集的来源。除此，还可以运用更多的算法或者过程发现模型来挖掘学生的学习行为。

参考文献

王翠如和胡永斌(2018)。翻转课堂真的能提升学习成绩吗?——基于 38 项实验和准实验研究的元分析。开放教育研究(04)，72-80。

邓国民、徐新斐和朱永海(2021)。混合学习环境下学习者的在线自我调节学习潜在剖面分析及行为过程挖掘。电化教育研究(01)，80-86。

乔丽方、赵蔚、段红和徐晓青(2022) 。基于隐马尔可夫模型的学习者在线自我调节学习过程挖掘——时间动力学视域下的分析。电化教育研究(10)，57-64。

李文昊、任晓曈、朱希雅和陈冬敏(2022)。认知弹性设计下自我调节学习行为影响学习成绩的机制研究。电化教育研究(02)，64-71。

李运福(2021)。大学生在线学习焦虑、自我调节与学习绩效的关系。教育评论(09)，99-108。

杨燕清、吴明、林艳芹和洪霞(2022)。自我调节学习在高职护理专业学生在线学习准备度与学

习投入间的中介作用。中华护理教育(06)，535-539。
孟军、刘冰璇、翟洪江和兰淇(2018)。大数据背景下高校翻转课堂学习评价的研究——以 A 校“工程热力学”课程为例。高等工程教育研究(05)，166-171。
徐晓青、赵蔚、刘红霞和姜强(2021)。自我调节学习中学习分析的应用框架和融合路径研究。电化教育研究(06)，96-104。
徐晓青、赵蔚、姜强、刘红霞和乔丽方(2022)。过程挖掘赋能教育数据分析：三种挖掘算法的应用探析。远程教育杂志(03)，45-55。
黄琰、赵呈领、赵刚、刘军、疏凤芳和李红霞(2020) 。教育过程挖掘智能技术：研究框架、现状与趋势。电化教育研究(08)，49-57。
An, F., Xi, L., & Yu, J. (2023). The relationship between technology acceptance and self-regulated learning: the mediation roles of intrinsic motivation and learning engagement. Education and Information Technologies, 1-19.
Bergmann, J., & Sams, A. (2012). Before you flip, consider this. Phi Delta Kappan, 94(2), 25-25.
Chen, W., & Cheng, H. N. (2020). Analyzing Graduate Students' Behaviors of Self-regulated Learning in a Blended Learning Environment. In Proceedings of the 2020 8th International Conference on Information and Education Technology (pp. 53-57), New York: Association for Computing Machinery.
Chen, X., Long, T., Cheng, L., Gan, X., & Zhu, X. (2023). Using WPBL to Improve Engineering Undergraduates' Computational Thinking Performance in Flipped Classroom. In International Conference on Blended Learning (pp. 151-162). Cham: Springer Nature Switzerland.
Cho, H. J., Zhao, K., Lee, C. R., Runshe, D., & Krousgrill, C. (2021). Active learning through flipped classroom in mechanical engineering: improving students' perception of learning and performance. International Journal of STEM Education, 8, 1-13.
Chua, K. J., & Islam, M. R. (2021). The hybrid Project-Based Learning–Flipped Classroom: A design project module redesigned to foster learning and engagement. International Journal of Mechanical Engineering Education, 49(4), 289-315.
Du, J., Huang, W., & Hew, K. F. (2021). Supporting students goal setting process using chatbot: implementation in a fully online course. In 2021 IEEE International Conference on Engineering, Technology & Education (TALE) (pp. 35-41). Wuhan: IEEE.
Gómez-Coma, L., Díaz-Sainz, G., Fallanza, M., Ortiz, A., & Ortiz, I. (2023). Integration of chemical engineering skills in the curriculum of a master course in industrial engineering. Education for Chemical Engineers, 45, 68-79.
Han, F., & Ellis, R. A. (2023). Self-reported and digital-trace measures of computer science students' self-regulated learning in blended course designs. Education and Information Technologies, 1-16.
Hromalik, C. D., & Koszalka, T. A. (2018). Self-regulation of the use of digital resources in an online language learning course improves learning outcomes. Distance Education, 39(4), 528-547.
Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. The Internet and Higher Education, 22, 37-50.
Li, S., Chen, G., Xing, W., Zheng, J., & Xie, C. (2020). Longitudinal clustering of students' self-regulated learning behaviors in engineering design. Computers & Education, 153, 103899.
Liu, R., Zhu, J., LI, W., Ju, T., & Chen, L. (2022). Investigating Engineering Students' Experiences of Self-Regulated Learning in Project-Based Learning Activities. International Journal of

Engineering Education, 38(5), 1422-1433.

Loyens, S. M., Magda, J., & Rikers, R. M. (2008). Self-directed learning in problem-based learning and its relationships with self-regulated learning. Educational psychology review, 20, 411-427.

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. Computers in Human Behavior, 80, 179-196.

Manukaram, K., & Abdullah, M. N. L. Y. (2021). The Influence of Family and Psychological Factors on Self-regulated Learning Behavior of Elementary School Students. The Journal of Behavioral Science, 16(3), 1-15.

Nielsen, K. L. (2023). Why can the flipped classroom frustrate students? experiences from an engineering mathematics course. Education Sciences, 13(4), 396.

Öztürk, M. (2022). The effect of self-regulated programming learning on undergraduate students' academic performance and motivation. Interactive Technology and Smart Education, 19(3), 319-337.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. Educational psychology review, 16, 385-407.

Rodríguez, G., Díez, J., Pérez, N., Baños, J. E., & Carrió, M. (2019). Flipped classroom: Fostering creative skills in undergraduate students of health sciences. Thinking Skills and Creativity, 33, 100575.

van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2020). Self-regulated learning support in flipped learning videos enhances learning outcomes. Computers & Education, 158, 104000.

Van Der Aalst, W. (2011). Process Mining: Discovery, Conformance and Enhancement of Business Processes. Heidelberg: Springer.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. Journal of educational psychology, 81(3), 329.

Zimmerman, B. J. (1989). Models of self-regulated learning and academic achievement. In Self-regulated learning and academic achievement: Theory, research, and practice (pp. 1-25). New York, NY: Springer New York.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In Handbook of metacognition in education (pp. 299-315). Routledge.

基于课堂生态学视角的城中村小学科学教育信息化的案例研究

A Case Study on the Informationization of Science Education in Urban Village Primary Schools from the Perspective of Classroom Ecology

徐莹[1]，赵雅[2]，龙陶陶[3*]
[123]湖北省武汉市华中师范大学人工智能教育学部
[*] taotaolong@mails.ccnu.edu.cn

【摘要】 城中村的落后问题是推进城市化进程中一个急需解决的问题，而城中村的义务教育均衡作为影响城中村发展的重要因素应该被赋予更多的关注。同时，教育信息化作为推进教育均衡的一个重要手段在城中村义务教育的发展中发挥着重要作用。而科学课堂作为教育信息化的实践基地也更应被纳入研究的范畴内。由于历史发展、地理位置等原因，城中村小学科学课堂的信息化发展存在着诸多问题。对武汉市化工区 J 小学的个案研究发现，城中村小学教育信息化发展生态失衡，具体表现在信息化设备与学校发展的失衡、教师群体内部结构失衡、学生主体与课堂发展的失衡。所以，应该从学校生态和课堂生态的改善着手，提出通过建立地方高校与城中村小学帮扶政策来改善其科学课堂的教育信息化发展的策略。
【关键词】 教育信息化；城中村小学；科学课堂；课堂生态学

Abstract: *Education informatization plays an important role in promoting educational balance. With the increasing attention to science education, education informatization is gradually integrating into science classrooms to promote the development of science education. However, there still exists the problem of uneven distribution of science education resources, which is even more scarce in urban village primary schools. Due to historical development, geographical location, and other reasons, there are many problems with the informatization development of science classrooms in urban village primary schools. This study adopts a case study of J Primary School in Wuhan Chemical Industry Zone to investigate the development of education informatization in this school from the perspective classroom ecology. The study results show the low level information literacy of the teaching staff, the backwardness of information technology equipment in J Primary School, the shortcomings of informatization in science classrooms, and the weak acceptance of information devices of students. Therefore, educators should pay great attention to improve the education informatization development by establishing assistance policies between local universities and urban village primary schools.*

Keywords: Education informatization, Urban Village Primary School, Classroom Ecology

1. 研究背景及研究意义

社会形态的变迁和媒介技术的发展，使智慧教育成为未来教育的主要形态，人类也将面临着新一轮的教育变革（胡钦太，刘丽清，郑凯，2019）。而教育信息化作为教育变革的核心推动力，在提高学校行政运作、高质量教师队伍建设和改善学生综合评价方式等方面具有重要的作用，也是推动教育现代化和办好人民

满意的教育的重要途径（任友群，冯仰存，郑旭东，2018）。因此，为了建设智慧校园，实现智慧教育，推进教育信息化必不可少（陈婷，甘梦蝶，2022）。但是，在教育信息化发展的过程中，由于城乡义务教育发展的不均衡，不同地区的教育信息化也就呈现出不同的质量，这对实现全国智慧教育产生了巨大的阻碍。

关于城中村的定义，不同学科研究者有不同见解，但对城中村本质上属于农村、物质形态构成和生活结构方面明显非农化、城中村的居住主体是城市外来流动人口等几个属性上已达成共识（仝德，冯长春，2009）。这些属性也使它面临着需要同时满足当地外来务工人员和原驻民子女义务教育需求的挑战。目前关于教育信息化发展现状的调查集中在经济发展相对落后的省份和市区，而忽略了对经济状况较好的城市中的城中村教育信息化发展状况的研究。比如，黎兴祥通过信息化基础设备、数字资源、信息技能、信息化应用和信息素养这五个方面对偏远山区中小学信息化建设现状进行调查（黎兴祥，2020）；黄斌和王妮丹关注贫困地四川革命老区，在了解基础环境建设、教育资源开发、教师信息素养、以及保障机制的基础上调查其基础教育教育信息化现状（黄斌，王丹妮，2020）；还有研究者对西藏、湘东、广西和西部少数民族等地进行了中小学教育信息化现状调查。虽然有研究针对城中村小学进行调查，但关注焦点仍然在落后地区。中国原教育工会主席方明曾说“教育是服务于人的，哪里有人哪里就需要教育”（方明，2021），而由于城中村处于城乡的边缘地带，在推进城乡义务教育均衡发展的任务下就更容易被忽视。

随着科技创新政策的不断推进，科学教育在中小学教育体系中占据着越来越重要的位置。但由于科学课具有教学内容丰富、可操作性强、灵活度高等特点，对科学教师的综合素质和教学资源的要求也相较于更高，对教育信息化的依赖性也相对更大。因此，科学课既是中小学素质教育改革的试验基地，又是城中村薄弱学校开成、开齐和开好最困难的课程。城中村薄弱学校教育的发展引起了社会各方的关注，已有研究涉及学校的校园环境的改善、校园文化的形成、数学课和英语课的提高、校本课程的开展等领域，但鲜有实证研究来探究如何提高薄弱学校的科学教育。

城中村小学的发展是促进我国城乡义务教育均衡发展的关键所在，其教育信息化的发展应该被赋予更多的关注，作为教育信息化实践重地的科学课堂也更应被纳入研究范畴。由于我国城市化的发展方向大体一致，城中村的发展也就存在着许多相同之处，这也使不同地区的城中村小学科学课堂的教育信息化存在着一定的共性。个案研究作为一种在特定的情境下城中村小学教育信息化发展提供阐释的研究方法，为探究城中村小学教育信息化的健康发展的研究路径提供了一个崭新的视角。因此，本研究以武汉市化工区 J 小学科学课堂教育信息化的发展情况为个案，分析其发展过程中存在的问题以及改善策略。

2. 研究对象及研究方法的选择

本研究采取了个案研究的方式，通过个别访谈、集体访谈和田野观察的方法收集研究资料， 采用扎根理论方法对研究资料进行分析 ，从课堂生态的视角对研究对象的信息化基础设备、数字资源、信息技能、信息化应用和信息素养等方面作出解释性理解。

武汉作为中国中部地区的中心城市，从 2003 年就开始了对主城区的城中村的改造，但至今改造出路仍未卜（张锦翌，彭建东，李志刚，2021）。J 小学所在的化工区作为新兴工业城区，是城市建设发展的重要支柱，需要大量的劳动力，也吸引了大量的外来务工人员。J 小学作为该村唯一一所小学，承担着该村外来务工人员和原住民子女的教育责任。

位于武汉市化工区的胡澈村属于青山区正在开展的东部地区第 1 期五个村“城中村”改造项目的范围，其项目完工时间预计在 2023 年 10 月底。作为改造的重点项目，J 小学直接接受武汉市青山区教育局的财政拨款，依照党和国家教育方针，已初步具备符合国家要求的义务教育均衡发展办学条件。在 2022 年，J 小学的学区内适龄儿童少年入学率已达 100%，基本实现开齐开足开好国家课程、地方课程、学校课表上墙、上网公布等基本要求，也形成了“经典诵读”和“足球”的学校特色。截止到 2022 年，J 小学师资力量为 25 人，在校学生为 150 人。拥有较好的绿色生态校园环境。但同时，笔者也发现 J 小学在优化方面侧重于学生德育和体育的发展，而对学生信息素质发展的关注度不高。

此次访谈对象为该小学的 3 名教师和五年级学生。教师分别为该校校长、科学教师、信息技术教师（见表 1）。由于人数均只有 1 名，因此以职位来匿名。五年级全体学生现在一共 32 名，其中在 2021 年下半年转入一名外来务工子弟。

表 1 教师相关信息

职称	性别	年龄	在校任职时间（年）	专业背景	学位	主要工作
校长	男	48	3	行政管理专业	本科	全面负责学校的行政工作
科学教师	女	33	2	语文专业	本科	负责 3-6 年级学生科学教学工作
体育教师	男	38	13	体育专业	本科	体育教学和学校总务工作

3. 城中村小学教育信息化发展生态失衡的具体问题

课堂生态学是教育生态学的重要内容。在课堂生态学中，课堂被类比为生态系统，因此就可以运用生态学的视角、方法和理论来研究课堂现象，分析课堂问题，以提高课堂教学效果（樊桂芳，2009）。同时，由于课堂也具有了生态系统所具有的某些特点、功能、规律，如能量流动、物质交换等（张惠薇，2021）。而课堂生态系统正是一种以教师和学生为主体，软硬件环境等客观因素制约的特殊生态（胡胜，黄镘，2021），由教师、学生和环境三个要素构成（黄远振，陈维振，2008），具有整体性、协变性和共生性三大生态特征。

3.1 教师群体内部结构失衡

新时代下，信息技术逐渐走进了校园，并成为学校教学工作的重要工具。为了解决校际教育信息化发展不平衡的问题，需要引进先进的信息技术来改善学校的教学条件。这同时也需要具有良好信息素养的老师来统筹学校的教学工作。而教师群体作为课堂生态系统的子系统，是教育发展和教学质量的重要保障，其内部结构失衡也会直接影响学校生态系统的平衡。

从年龄结构上来看，J 小学现有教师 25 人，年龄都集中在 35 岁以上，整体存在老龄化现象。从信息素养上来看，中年教师大多对信息技术的敏感度低，接受能力差，信息化教学能力不足。从教师结构上来看，J 小学缺乏专职的信息技术教师，必须依靠兼职或者由比较懂电脑操作技术的教师兼任，只能够在规定时间内完成最基本的教学需要，无法开展高质量的课堂。例如，在教育信息化发挥重要作用的科学课堂中，J 小学只有一名科学教师，负责 3-6 年级学生的科学教学工作。在对科学教师的采访中发现这名科学教师的本科是语文专业，由于她在

本科学习中从未接触过科学教育，因此在入职前期存在着很大的困难，担心无法保证科学课教学质量。

3.2 信息化设备与学校发展的失衡

虽然J小学属于小规模的城中村学校，但是在武汉市政府实施相应的帮扶政策后，校园内也具备了功能齐全的电脑室，同时基本实现了多媒体的全覆盖。但由于具有良好信息素养教师的缺乏，这些信息化设备不能完全融入教学中，得不到充分的利用。在对J小学校长的访谈中发现，J小学存在着由于帮扶政策的偏向而导致的信息化设备被闲置的问题。J小学的帮扶政策重点在改善学校的教学环境，如操场的重建和功能室的完备上，如此可以看出此政策着重于改善学校的物理环境，并没有从教师结构不合理，教师岗位定位不清晰等方向来培养教师的信息技能，从而使信息化设备能够被妥善地使用到教学中去。同时，J小学“足球特色学校”的定位也使其更关注学生体育方面的发展，而相对地忽略了对学生信息素养等方面的培养，对信息课和科学课的重视程度方面也存在着不足。而J小学目前仍使用的是传统的科学教室，与“智慧教室”“创客教室”之间的差距较大，教育信息化在科学课堂中仍没有足够的渗透度。教室中的大部分教学用具也使用年限较长，并多为直观教具，缺少可操作的实验器材，不利于科学实验的开展等问题。同时，现有的教学用具的利用率也很低，缺少信息化用具，需要新的教学资源的加持。

3.3 学生主体与课堂发展的失衡

学生是课堂生态的主体，其在课堂的各种行为和感受会直接影响课堂的发展。在对J小学五年级的全体学生发布调查问卷，对他们的科学知识和信息素养进行调查后发现，大部分学生整体的知识面较窄，他们对于科学知识的理解多根据日常的生活经验，多是一种朴素的、感性的认识，缺少对科学知识概念定义的理性认识。由于信息素养的培养需要以大量科学知识为基础，因此科学知识的缺失会直接影响信息素养的形成和信息技能的提高。在调查中发现，学生对科学课堂的活动仅局限在做实验、做笔记和看视频三个活动上，并没有参与需要信息素养的活动中，而这直接影响了科学课堂信息化的发展。

4. 城中村小学教育信息化发展困境的生态改善策略

从城中村薄弱学校的自身发展来看，促进薄弱学校的教育均衡需要结合学校的实际情况，因地制宜。而服务地方是地方高校的重要使命。在呼吁义务教育均衡发展的大背景下，地方高校应在追求自身进步与卓越发展的同时，理应主动承担起推动校际教育均衡发展的重任。因此，地方高校肩负着发展自我与推动区域薄弱学校发展的双重使命（谭劲，2014）。可见，在对薄弱学校进行针对性的调研后，地方高校可与其建立起相应的帮扶机制，利用高校的研究和资源优势，推动薄弱学校个性化的优质与均衡发展。

位于武汉市的华中师范大学是教育部直属的师范高校，具有充足的教学资源和教学研究的经验。校内的科创协会由华中师范大学人工智能教育学部科学教育专业的学生创办，专门从事开展科学课教学的志愿活动。通过对J小学科学课堂教育信息化的现状分析发现，J小学科学课堂生态的改善需要新鲜力量的融入。教师层面，需要为教师减轻负担的同时，提升教师的专业能力。学生层面，需要尝试多样化的教学活动，接触像人工智能这样的科技前沿知识，拓宽知识面。课堂环境层面，需要融入更多的新型教学资源，提高教学效率，并在此基础上提升资源的利用程度。因此，以改善小学科学课堂生态为共同意识，促进华中师范大学科创协会生态系统和J小学科学课堂生态系统之间的物质与能量流动，使J小学课堂生态变得更加完善发展。

4.1 满足J小学教师、学生、课堂环境三个方面的需求

科创协会生态系统由协会成员、协会指导教师和活动资源构成，其中协会成员主要为华中师范大学科学教育与教育信息技术专业学生，具备一定的专业能力，同时有实践能力提升需求，因此参与积极性较强。协会指导老师是华中师范大学的在职教师，有指导教师能力提升的经验，可以在能量流动过程中给予建议，甚至参与整个过程。活动资源主要为创客实验室，其中包含现阶段比较前沿的科学教学工具，如：VR眼镜、机器人、开源硬件等。这三者正好满足J小学教师、学生、课堂环境三个方面的需求，同时也可以提升协会成员的实践经验，达到互助共赢。

4.2 设计两个生态之间活动的具体措施

具体措施的总体规划是通过课程入手，从课程开发、课程实施和课程改善三个环节来设计两个生态之间的活动，其中每个环节都包含生态系统的物质、能量与信息的流动，并在流动过程中促进生态系统的改善。

课程开发阶段，主要由科创协会成员结合J小学的实际情况与课程标准进行开发，在此过程中，J小学科学教师提供信息支持，如：学生的喜好、学生的知识水平等。这个过程由协会成员先进行学习者分析后，教师再进行的修改与完善。在课程开发完成后，需要与J小学教师进行交流讨论，结合实情对课程设计进行改善，使课程设计更具有可操作性。课程实施阶段，科创协会成员进行主讲，J小学根据其意愿担当助教或评课教师。课程实施完成后，J小学教师给予关于主讲教师教学行为和课程流程等方面的改进意见。课程反思阶段，将课程录像给协会指导教师观看，收集意见。在此之后，课程协会的参与成员进行讨论交流，根据收集的意见与建议对自身进行反思并分享收获。

在整个过程当中，J小学科学教师可以减轻课程创新的负担，同时提升自己的信息技术教学能力，而科创协会的成员将理论与实践结合，巩固学校所学知识的同时提升教学实践能力。

4.3 关注这两个生态系统之间循环流动的可持续性

由于J小学的一个根本性问题是科学教师资源的缺乏，因此需要考虑两个生态系统之间循环流动的可持续性。鉴于此，本研究从大学社团入手，既保证了人员的稳定性，同时又提升了参与者的多样性和丰富性。在华中师范大学科创协成员的不断迭代更新下，通过协会老成员带新成员，将与J小学的交流合作作为协会的主要活动，在保证互动顺利进行的同时，满足协会成员对于实践能力提升的需求，进而成员参与互动交流的积极性也可以得到保证。综上，可持续性的问题可以在逻辑上成立，不可否认这个过程可能会出现塌陷，但是生态系统从因改善需求发生失衡到恢复平衡也需要时间来巩固，因此这个变化时期需要有详细的规划与操作，并对生态系统进行有效评估。

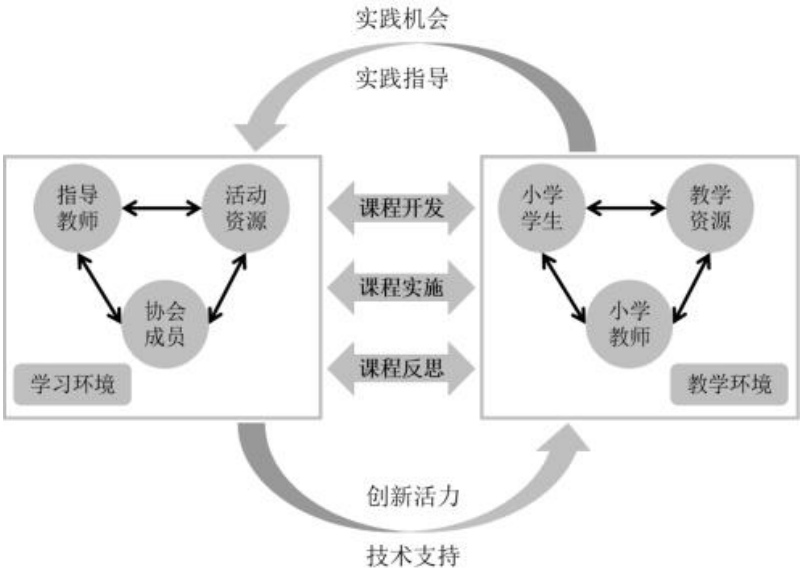


图 1 生态流动图

这些措施的设计需要在实践中证明，因此在后期的研究中，我们将以此为基础来进行实操，通过多方面对设计进行改进与完善，如：J 小学科学教师与协会成员的成长、学生对课程的反馈等。希望给予其他研究者一些思路上的借鉴。

参考文献

胡钦太,刘丽清,郑凯.工业革命 4.0 背景下的智慧教育新格局[J].中国电化教育,2019(03):1-8.

任友群,冯仰存,郑旭东.融合创新,智能引领,迎接教育信息化新时代[J].中国电化教育,2018(01):7-14+34.

陈婷,甘梦蝶.教育信息化促进基础教育高质量发展：逻辑、困境与路径[J].教育与教学研究,2022,36(09):120-128.DOI:10.13627/j.cnki.cdjy.2022.09.003.

仝德,冯长春.国内外城中村研究进展及展望[J].人文地理,2009,24(06):29-35.

黎兴祥.山区中小学教育信息化建设与应用现状调查研究——以肇庆市怀集县为例[J].中国教育技术装备,2020(09):11-13+16.

黄斌,王丹妮.四川革命老区基础教育信息化现状调查及对策研究[J].中国教育信息化,2020(01):78-81.

方明.乡村教育要放到城乡融合背景下统筹布局[J].中国乡村发现,2021(03):85-87.

张锦翌,彭建东,李志刚.中部大城市城中村的“在地边缘化”问题——对武汉先建村的实证[J].现代城市研究,2021(07):112-117.

樊桂芳.外语教育课堂生态观探究[j].教育与管理（理论版）,2009（1）.

张惠薇.课堂生态学视域下负反馈调节在中学地理教学中的影响分析[J].科学咨询(教育科研),2021(01):227.

胡胜,黄镘.立足课堂生态本质构建课堂教学新生态[J].中国教育学刊,2021(08):104.

黄远振,陈维振.课堂生态的内涵及研究取向[J].教育科学研究,2008(10):48-50.

谭劲.名校推进区域义务教育校际均衡发展的策略——基于重庆市珊瑚实验小学的实践[J].教育探索,2014(01):76-78.

跨学科课例研究下教师合作对话模式的社会认知网络特征分析

What did they say in an Interdisciplinary Lesson Study?Capturing Teachers' Collaborative Talk Pattern for Teacher collaborative learning

赵雅[1]，徐莹[2]，龙陶陶[3*]
[123]华中师范大学
[*] taotaolong@mails.ccnu.edu.cn

【摘要】 跨学科课例研究对跨学科教育视域下的教师专业发展具有重要意义。教育生产性对话（PPT）为更深入地理解跨学科课例研究中教师合作对话提供了过程性数据。本研究以 12 名职前教师和 3 名大学教师为研究对象，采用社会认知网络特征（SENS）分析法研究跨学科课例研究过程中不同群体的互动过程和认知特征。研究结果显示活跃和多样化的跨学科互动已经形成，不同群体的认知特征都主要集中在对问题的发现和解决。研究为教师开展跨学科教育环境下的跨学科课例研究提供了借鉴，为分析参与者的互动规律和认知特征提供了新思路。

【关键词】 跨学科课例研究；教育生产性对话；社会认知网络特征分析

Abstract: *With the development of interdisciplinary education, interdisciplinary lesson study has garnered significant recognition. The framework of Pedagogically Productive Talk (PPT) has been established to facilitate the analysis of teachers' collaborative talk pattern in an interdisciplinary lesson study, with dialogue serving as procedural data. Utilizing the Social Epistemic Network Analysis (SENS) approach, this study explored 12 student teachers and 3 faculty members' collaborative talk patterns with regards of the interactive processes and epistemic patterns in an offline interdisciplinary lesson study. Quantitative analysis revealed active and diverse cross-disciplinary teacher interaction, and cognitive characteristics demonstrate the practical problem identification and resolution in different groups. Besides, there were distinctive interaction characteristics in different stages.*

Keywords: Interdisciplinary lesson study, Teachers' collaborative talk, Social epistemic network signature

1. 前言

2022 年 4 月教育部发布的《义务教育课程方案和课程标准（2022 年版）》明确将跨学科教育作为优化课程内容和结构的重要推动力（中华人民共和国教育部，2022）。跨学科教育的高质量发展离不开跨学科课例研究的有力支持（钟柏昌和赵国雄，2023）。研究发现跨学科课例研究对促进课程改革、优化教师教学和提高学生成绩都具有重要作用（许华和陈晓均，2023）。同时，跨学科课例研究中教师合作学习的类型和策略也对其质量产生了影响（王建良和冯霞，2018）。然而，由于跨学科课例研究中教师合作学习的复杂性和动态性，很少有研究涉及对其内在过程机制的探究（杨文正和许秋璇，2021）。近年来，教师合作对话已经成为教师合作学习中的重要研究领域。教育生产性谈话(Pedagogically Productive Talk, PPT)作为新提出的理论框架已经被研究证实可以揭示教师合作对话的内在机制（Lefstein, A., Louie, N., Segal, A., & Becher, A, 2020; Moe, A., Consiglio, P., & Katz, I, 2022）。社会认知网络特征

(Social Epistemic Network Signature，SENS)分析法是一种合作学习分析方法，综合运用了社会网络分析（Social Network Analysis，SNA）与认知网络分析（Epistemic Network Analysis，ENA）的优势，可以从社会和认知层面直观全面地揭示合作者在认知过程中产生的交互数据，已经被广泛应用于教师教育领域。因此，本研究采用社会认知网络特征(SENS)分析法分析跨学科课例研究中的教师合作对话模式，并对不同群体在不同阶段的互动行为和认知特征进行可视化分析，以提升教师的跨学科合作能力，为开展跨学科教育背景下的跨学科课例研究提供借鉴与参考，为其他研究者动态表征和评价教师合作对话过程提供新的研究思路。

2. 理论基础

2.1. 跨学科课例研究

跨学科课例研究对跨学科教育的高质量发展至关重要 (Zhong B. Liu X, Zhan Z,et al. , 2022)。跨学科课例研究强调跨越学科间的边界，将不同学科进行整合以应对教学挑战（陈刚，王淑萍和丛明宇，2016）。具有创新精神和高素质的 STEAM 人才培养离不开跨学科课例研究的兴起（管光海，2017）。然而，跨学科课例研究的实施和发展需要在因地制宜的基础上消耗大量的时间和精力(于国文和曹一鸣，2017）。目前对跨学科课例研究主要集中在模式构建与案例介绍方面（邵俊峰，2016；刘洁和陈畅频，2019；陈金海，2020），缺少对跨学科课例研究中教师合作对话有效性和发展路径的探究。

2.2. 教师合作对话

教师合作对话立足于教师合作学习，指教师通过话语进行专业成长与发展。研究表明，教师合作对话可以改善教师教学和学生学习。目前，有关教师合作对话理论框架层出不穷，如课例研究（Lesson study），有关知识的学习（Opportunities to learn）等（Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E, 2015; Warwick, P., Vrikki, M., Vermunt, J. D., Mercer, N., & van Halem, N, 2016）。Lefstein 等人提出的教育生产性谈话（PPT）作为一种新兴的理论框架直接关注话语本身，可以识别教师合作对话中的精细模式，以评估教师合作对话的质量（Saunders,et al., 2023）。

综上所述，跨学科课例研究的研究和实施仍尚待开展。同时，教学生产性谈话（PPT）可以评估和促进跨学科课例研究的质量。在跨学科课例研究的不同阶段，教师合作对话的生成和发展需要更多的关注。因此，本研究旨在通过回答以下问题来研究跨学科课例研究中教师合作对话的内在机制：

（1）在跨学科课例研究过程中，职前教师和大学教师是如何互动以促进教师合作对话的产生和发展呢?

（2）跨学科课例研究中不同阶段的教育生产性谈话是什么样的?

（3）跨学科课例研究中不同群体的教育生产性谈话是什么?

3. 研究设计

3.1. 研究对象

本研究选取华中地区某师范大学 12 名来自不同年级和背景的科学教育、软件工程和教育技术学的职前教师, 以及 3 名大学教师组成的跨学科课例研究团队为研究对象。以设计一门“人工智能+科学”跨学科课程为样本开展为期 5 周的实证研究。研究将研究对象分为三个组，分别是技术组、教学组和指导组。技术组共 5 名成员，专业为教育技术学和软件工程，负责用于跨学科课程的信息技术工具的开发和调整。教学组共 7 名成员，专业为科学教育，科学课教学经验丰富，负责跨学科课程的设计与实施。指导组共 3 名成员，分别是 1 名科学教育专

业的大学教师和 2 名教学技术学专业的大学教师，负责监督跨学科课程的实施并提供资金和技术上的支持。在整个研究过程中，小组分配保持稳定，此前，没有人有过跨学科课例研究的经验。

3.2. 研究背景

本研究选取的为期五周的“人工智能+科学”跨学科课程的开发和实施主要针对一所教学条件薄弱的城中村小学，旨在通过人工智能辅助科学课教学的方法促进教育公平。此跨学科课程旨在将平板电脑、聊天机器人和虚拟实验应用程序这三种信息工具融入到小学科学课堂中。本研究的跨学科课例研究共分为四个阶段，分别是学习（study）、计划（plan）、教学（teach）和反思（reflect）。其中，除计划阶段召开了两次线下会议外，其余三个阶段都只开了一次线下会议，会议时长平均为 50 分钟，共 5 次会议。

3.3. 数据收集与分析

本研究采用社会认知网络特征（SENS）分析方法，结合社会网络分析和认知网络分析，探究跨学科课例研究中的不同群体在不同阶段的互动行为和认知特征（Nash, P., & Shaffer, D. W, 2012）。

（1）选用 UCINET6.0 作为社会网络分析工具，分四个阶段对跨学科课例研究中职前教师和大学教师的社会互动网络结构进行剖析，探究职前教师和大学教师在线下会议讨论过程中的交互特征。通过对跨学科课例研究中的话语数据进行分析，建立职前教师和大学教师间的交互关系矩阵，输入到 UCINET6.0 中，结合点度中心度和中间中心度对网络结构量化分析。

（2）基于 Saunders 等人于 2023 年修订的教育生产性框架（PPT），对跨学科课例研究中职前教师和大学教师的话语内容进行双重编码。原始的编码包括发现问题、结合理论知识分析问题产生的原因、结合表征性数据分析问题产生的原因、包容和思考其他人的想法、寻求解决方法以及支持和批判他人的观点这六个维度。由于该编码框架最初是为针对单一学科课程进行教研的成熟教师团队设计的，因此为了评估跨学科课例研究中教师合作对话的质量和有效性，增加了维护成员关系的维度，见表 1。由两位经验丰富的编码者对跨学科课例研究中产生的话语数据进行编码，其编码一致性 Cohen's Kappa 系数为 0.914>0.8，信度良好。将编码好的数据导入认知网络分析在线工具 Web Tool(version 1.7.0)(http://app.epistemicnetwork.org)构建基于跨学科课例研究的认知网络可视化分析，从认知角度评估教师合作对话的有效性（Dawson, S., Gašević, D., Siemens, G., & Joksimović, S, 2014）。

表 1 跨学科课例研究中职前教师和大学教师的认知网络分析编码表

维度	释义	编码
发现问题	引发或启发大家思考可能存在问题	POP
结合理论知识分析问题产生的原因	结合理论知识，从逻辑证明的角度梳理问题产生的原因	PR
结合表征性数据分析问题产生的原因	结合真实经验，图像，录像，对话等分析问题产生的原因	ROP
包容和思考其他人的想法	提供多元的视角、探索、解决冲突	MV
寻求解决方法	提供不同的方法去解决教学问题	GO
支持和批判他人的观点	表达赞同或提出质疑	SC
维护成员关系	维持秩序、提供鼓励	SUE

4. 数据分析

4.1. 社会网络分析结果

中心性是社会网络分析的一个研究重点，它可以测量节点接近网络中心的程度，从而表征个体在社会网络结构中所处的位置及重要性。本研究主要从点度中心度和中间中心度这两个方面对不同群体在跨学科课例研究的社会交互行为进行分析。

节点的点度中心度（Degree Centrality）指网络结构中与该节点有联系的节点个数，包括点入度和点出度。节点的中间中心度（Betweenness Centrality）测量一个节点担任其他两个节点间最短桥梁的次数。表 2 揭示了跨学科课例研究中部分成员的点度中心度和中间中心度，节点值越大表明成员的中心性越高。在互动次数方面，F2 是最活跃的成员(264,216)，其次是 F1 (171,209)， S3 (161,163)， F3(81,87)和 T1(76,84)，这表明在跨学科课例研究过程中，三个群体保持相对平等的地位，没有一个占主导地位的小组。在互动中介方面，S1（40.343）作为其他两个节点间最短桥梁的次数最多，其次是 F1(30.790)和 T1(19.288)。这表明每个群体中都有一个成员作为促进者，起着凝聚社会互动的作用。

点度中心度社群图的基本轮廓和图 1 基本一致。节点表示参与的成员，节点的大小反映该节点与其他节点的连接数量。节点间的连线表示节点间的互动关系，线的粗细表示两个节点之间交互的程度，总共有 1374 个话语被编码分析。通过对点度中心度社群图的分析发现 F2 和 S3、F1 和 F2、F1 和 S2、S2 和 T5、F2 和 F3 以及 F2 和 T1 是互动相对频繁的组合。这表明，在跨学科课例研究的过程中，大学教师和职前教师之间，以及不同学科的教师之间均发生了积极多样的互动。

表 2 部分成员社会网络结构中心度值

成员编号	点入度	点出度	中间中心度
F2	264	216	5.798
F1	171	209	30.790
S3	161	163	6.579
F3	81	87	2.514
T1	76	84	19.288
S2	43	50	0
S1	38	28	40.343

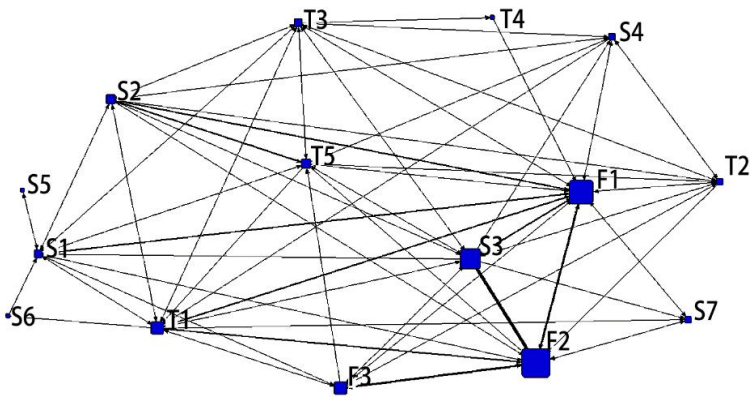


图 1 跨学科课例研究过程中不同群体间的点度中心度社群图

4.2. 认知网络分析结果

表 2 展示了跨学科课例研究四个阶段的教育生产性对话（PPT）的认知网络特征差异。节点代表了教育生产性对话（PPT）的七个维度，节点的大小表示每个维度的占比程度，节点间的连线表示不同维度间的互动关系，总共有 2228 个话语用于编码分析。

在学习（study）阶段，发现问题（PoP）是出现频率最高的维度，其次发现问题分别与寻求解决方法（PoP-GO,0.9）和维护成员关系（PoP-SUE,0.84）间保持着较为紧密的关系。参与者在维系团队关系的基础上主要专注于提出和解决实际问题。在计划（plan）与教学（teach）阶段，突出的维度是发现问题（PoP）、寻求解决办法（GO)和支持和批判他人的观点（SC），其中发现问题与寻求解决方法（PoP-GO,0.94）、发现问题与支持和批判他人的观点（PoP-SC，0.53）以及寻求解决方法与支持和批判他人的观点(GO-SC,0.41)之间存在显著联系。这表明，虽然参与者继续优先考虑发现问题和寻求解决方法，但逐渐将注意力转向支持和批评他人的观点，同时，维护成员关系的话语相对减少。

在反思（reflect）阶段，教育生产性对话（PPT）的七个维度分布得更为均匀。具体而言，发现问题(PoP)、包容和思考其他人的想法(MV)、寻求解决方法(GO)和支持和批判他人的观点(SC)是最突出的维度。值得注意的是，发现问题、寻求解决方法和支持和批判他人的观点之间的联系有所减弱(PoP-GO，1.00; PoP-SC, 0.54; GO-SC, 0.63)。然而，寻求解决方法(GO)、支持和批判他人的观点(SC)和包容和思考其他人的想法(MV)之间的联系尤为突出(GO-SC, 0.43; MV-GO, 0.60; SC-MV, 0.82)。这表明在反思阶段，跨学科课例研究团队更强调团队内部的包容性和对不同观点的思考，特别是在发现问题和寻求解决方案方面，同时仍会对他人观点表示支持或批判。

表 3 揭示了在跨学科课例研究过程中，三个小组教育生产性对话（PPT）的独特特征。三个小组的特征也有很多许多相似之处。例如，三个小组的发现问题(PoP)都是最突出的特征，然后是寻求解决方法(GO)和支持和批判他人的观点(SC)。同时，发现问题(PoP)和寻求解决方法(GO)之间、发现问题(PoP)和支持和批判他人的观点(SC)之间以及寻求解决方法(GO)和支持和批判他人的观点(SC)之间的联系在技术组、教学组和指导组这三个群体中都发挥了重要作用。此外，与技术组和教学组相比，指导组的发现问题(PoP)与维护团队关系(SUE)之间的联系更为强烈(PoP-SUE 分别为 0.49,0.33 和 0.35)。这表明三个小组都优先考虑发现问题、寻求解决方案和批判性地支持观点，教师小组更强调在整个过程中团队关系保持亲密的重要性。

表 2 跨学科课例研究不同阶段的认知网络图

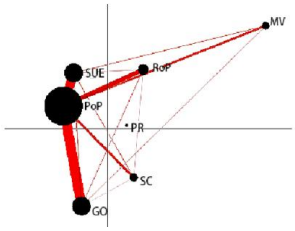
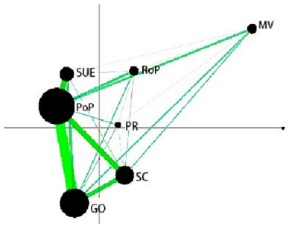
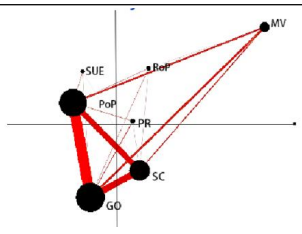
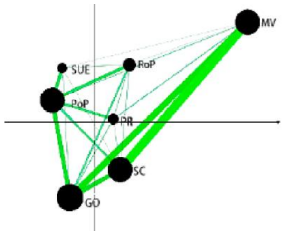
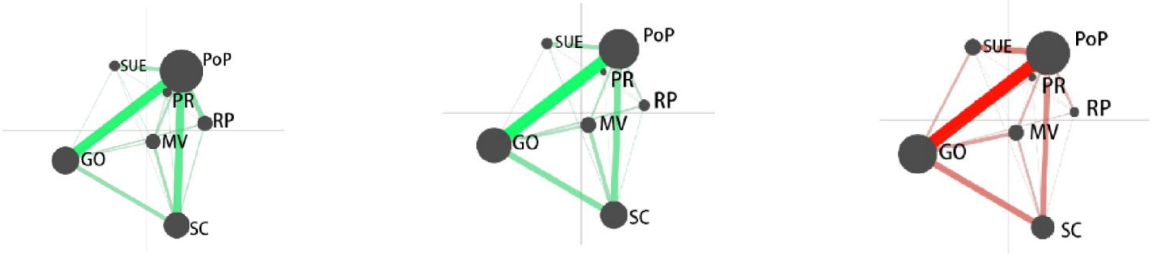
学习（study）	计划（plan）	教学（teach）	反思（reflect）
			

表 3 跨学科课例研究不同群体的认知网络图

技术组	教学组	指导组



5. 结论与展望

本研究应用教育生产性对话（PPT）框架，探究职前教师和大学教师在跨学科课例研究中的合作模式。研究结果表明在跨学科课例研究的过程中，不同背景的教师之间存在着不同的互动方式，构建了一个实践共同体(CoP)。同时，在跨学科课例研究的每个阶段，三个小组都侧重于对实际问题的发现和解决，同时还表现出了其他的合作特征。

社会网络分析(SNA)结果表明跨学科课例研究的群体不再是孤立的个体，而发生了积极地互动。Wenger(1998)指出实践共同体(CoP)是指一群在某一领域拥有共同兴趣的个体通过积极地参与集体学习，来共享技能和完成任务（Wenger E, 1998; Wenger E, 2000）。实践共同体(CoP)在促进不同背景的成员参与活动方面发挥了关键作用（Zhao G, Yang X, Long T, et al., 2019; Wang Q & Lu Z, 2012）。本研究可视化地揭示了跨学科课例研究内部多元的交流和互动模式，跨学科教师跨越了学科壁垒，积极地参与知识整合和集体讨论，促进了实践共同体(CoP)的形成，营造了一个充满活力和包容的跨学科学习环境。

认知网络分析的结果揭示了跨学科课例研究中不同阶段和小组的教育生产性对话（PPT）的认知特征。该模式下的跨学科教师合作对话对跨学科教育的高质量发展至关重要。研究结果表明跨学科课例研究四个阶段的认知焦点和动态在逐渐转变。在学习阶段（study），确定关键问题(PoP)占据首要位置，同时保持协作氛围(GO)和维护团队关系(SUE)也非常重要。先前的研究也强调了清晰的问题识别和积极的群体动力对于有效合作学习的重要性（Jonassen D, 2011; Johnson D W & Johnson R T, 2009）。随着跨学科课例研究进展到计划（plan）和教学（teach）阶段，支持和批评他人观点(SC)开始出现，这表明参与者对主题的更深入探究以及更具批判性和建设性对话的出现。然而，在教学阶段（teach），发现问题(PoP)和维护团队关系(SUE)之间的联系明显减少，这表明此阶段参与者更倾向于完成任务，而不再维持团队凝聚力。这表明参与者遇到了平衡任务与维持关系的挑战，这个挑战对合作团队保持生产力和发展创新能力至关重要（Carmeli A, Tishler A & Edmondson A C, 2012）。在反思阶段（reflect），该教师合作对话模式展示了一个成熟的群体将不同的观点整合到建设性的反思对话中的方式，这有利于在合作学习环境中提高群体的创造力和解决问题的能力（Engeström Y, 2001）。同时研究结果还强调了发现问题和合作在不同学科背景的跨学科课例研究中有效促进教学的重要意义（Fortus D, Dershimer R C, Krajcik J, et al., 2004）。在反思阶段（reflect），发现问题(PoP)、支持与批判他人观点(SC)和包容和思考其他人的想法(GO)之间的紧密联系反映了一种识别、评估和解决教学挑战的结构化方法，这有助于从反思角度探究如何促进跨学科教育的进步（Häkkinen P, Järvelä S, Mäkitalo-Siegl K, et al., 2017; Korthagen F & Vasalos A, 2005）。

此外，研究结果有助于在跨学科教育中理解和应用教育生产性对话(PPT)框架。通过对教育生产性对话(PPT)框架的认知网络分析，揭示了跨学科课例研究中不同群体之间合作过程的细微动态，这证实了 Lefstein 等人(2020)强调的教育生产性对话(PPT)在分析教师合作对话中的作用，展示了教育生产性对话(PPT)框架在不同教育群体中的适用性，这为未来的研究提供了新的视角。本研究还存在着局限性，例如仅对一项案例进行了分析，且仅收集了话语数据。

未来的研究应更关注职前教师和信息技术在跨学科教育中的整合作用，以优化策略。总而言之，本研究的发现有潜力改变跨学科教育实践的方向、培养更具包容性、更有活力和更有效的跨学科学习环境。

参考文献

于国文，曹一鸣（2017）。跨学科教学研究:以芬兰现象教学为例。**外国中小学教育，2017(07)**，57-63。

中华人民共和国教育部（2022）。**义务教育课程方案和课程标准（2022 年版）**。北京：北京师范大学出版社。

王建良，冯霞（2018）。基于教学项目实施的跨学科教研。**人民教育，2018（2）**，21-23.

许华，陈晓均（2023）。小学跨学科教研实施路径与策略。 **基础教育课程, 2023, (19)**，52-57.

刘洁，陈畅频（2019）。立足同伴资源的跨学科教研探索。 **教育理论与实践，2019（11）**，39-40。

杨文正，许秋璇（2021）。融入“大概念”的 STEAM 跨学科教研：模式构建与实践案例。**远程教育杂志，2021（2）**，103-112。

陈刚，王淑萍，丛明宇（2016）。高校跨学科教学研究型团队组织建设探讨。**中国成人教育，2016（14）**，53-55。

陈金海（2020）。高中跨学科校本教研体系的构建与实践。 **教学与管理，2020（31），**33-35。

邵俊峰（2016）。共生理念下中学跨学科教研基地的创设。 **中小学教师培训，2016（6）**，23-25。

钟柏昌，赵国雄（2023）。 跨学科教研的困境分析与破解之道——基于共同内群体认同模型的分析。**现代远距离教育, 2023, (05)**，16-23。

管光海（2017）。基于标准聚焦核心概念的 STEM 整合教育模式———美国 I-STEM 模式的特点及启示。**开放教育研究，2017（6）**，87-93。

Carmeli A, Tishler A, Edmondson A C. (2012). CEO relational leadership and strategic decision quality in top management teams: The role of team trust and learning from failure. *Strategic Organization, 2012, 10(1)*, 31-54.

Dawson, S., Gašević, D., Siemens, G., & Joksimović, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. *Proceedings of the fourth international conference on learning analytics and knowledge (pp. 231–240).* New York, NY, USA: ACM.

Engeström Y. (2001).Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of education and work, 2001, 14(1)*, 133-156.

Fortus D, Dershimer R C, Krajcik J, et al. (2004).Design-based science and student learning. *Journal of Research in Science Teaching, 2004, 41(10)*, 1081-1110.

GASEVIC D, JOKSIMOVIC S, EAGAN B R, & SHAFFER D W. (2018). SENS：network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in human behavior, 2018, 7（3）*,1-16.

Häkkinen P, Järvelä S, Mäkitalo-Siegl K, et al. (2017). Preparing teacher-students for twenty-first-century learning practices (PREP 21): a framework for enhancing collaborative problem-solving and strategic learning skills. *Teachers and Teaching, 2017, 23(1)*, 25-41.

Jonassen D.(2011). Supporting problem solving in PBL. *Interdisciplinary Journal of Problem-Based Learning, 2011, 5(2)*, 95-119.

Johnson D W, Johnson R T.(2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational researcher, 2009, 38(5)*, 365-379.

Korthagen F, Vasalos A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and teaching, 2005, 11(1),* 47-71.

Lefstein A, Vedder-Weiss D, Segal A（2020）. Relocating research on teacher learning: Toward pedagogically productive talk. *Educational researcher, 2020, 49(5)*, 360-368.

Moe, A., Consiglio, P., & Katz, I. (2022). Exploring the circumplex model of motivating and demotivating teaching styles: The role of teacher need satisfaction and need frustration. *Teaching and Teacher Education*, 118, 1e12.

Nash, P., & Shaffer, D. W. (2012). Epistemic trajectories: Mentoring in a game design practicum. *Instructional Science, 41(4)*, 745–771.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17e40.

Wang Q, Lu Z. (2012). A case study of using an online community of practice for teachers' professional development at a secondary school in China. *Learning, media and technology, 2012, 37(4)*, 429-446.

Warwick, P., Vrikki, M., Vermunt, J. D., Mercer, N., & van Halem, N. (2016). Connecting observations of student and teacher learning: An examination of dialogic processes in lesson study discussions in mathematics. *Zdm*, 48, 555e569.

Wenger E. (1998). Communities of practice: Learning as a social system. *Systems thinker, 1998, 9(5),* 2-3.

Wenger E. (2007). Communities of practice and social learning systems. *Organization, 2000, 7(2),* 225-246.

William Saunders,Taylor Topham, Bryant Jensen,David Marcelletti, Kevin McCarthy, Lindy Lee（2023）.What's in a teacher team meeting? Testing pedagogically productive talk as a framework for teacher collaboration. *Teaching and Teacher Education,131(2023)104176*.

Zhong B, Liu X, Zhan Z, et al.(2022). What should a Chinese top-level design in STEM Education look like? *Humanities and Social Sciences Communication, 2022(1),*1-8.

Zhao G, Yang X, Long T, et al.(2019). Teachers' perceived professional development in a multi-regional community of practice: Effects of beliefs and engagement. *Learning, Culture and Social Interaction, 2019, 23,* 100347.

六帽思维技术对职前教师集体反思的作用：协同调节模式的滞后序列分析

The Role of the Six-hat Thinking Technique on Pre-service Teachers' Collective Reflection: Lag Sequential Analysis on the Co-regulation Pattern

梅傲雪[1]，龙陶陶[2*]
[1 2]华中师范大学 人工智能教育学部
[*] taotaolong@mail.ccnu.edu.cn

【摘要】 集体反思已成为教师培训和专业发展的一项重要活动。然而，集体反思还存讨论混乱，效率低，难以达到通过集体反思促进个体反思效果等不足。本研究制定了基于六帽思维技术的集体反思框架：按照红-白-黄-黑-绿-黑-绿…绿的顺序(直到没有新的黑帽出现)进行反思，并采用协同调节编码方案和滞后序列分析(Lag Sequential Analysis)对进行三轮教学设计项目学习的职前教师集体反思行为进行分析，揭示了六帽思维技术增强的集体反思协同调节模式。本研究还发现了高水平理解组和低水平理解组在理解六帽思维本质时所表现出的调节过程和调节焦点的差异。

【关键词】 集体反思；协同调节；六帽思维技术；职前教师

Abstract: *Collective reflection has become a rigorous activity in teacher training and professional development. However, collective reflection still has the disadvantages of chaotic discussion, low efficiency and difficult to achieve the effect of promoting individual reflection through collective reflection. This study developed a collective reflection framework based on six thinking caps: follow the rules of red-white-yellow-black-green-black-green... Green sequential (until there is no new black hat idea) reflection, and the collaborative adjustment coding scheme and delayed sequence analysis (LSA) were used to analyze the collective reflection behavior of pre-service teachers who had three rounds of instructional design project learning, revealing the collaborative adjustment mode of collective reflection enhanced by six thinking hats technology. This study also found differences in co-regulation processes and co-regulation focus between the high-performing and low-performing groups in understanding the nature of six-hat thinking.*

Keywords: collective reflection, co-regulation, six-thinking hats, pre-service teachers

1. 前言

集体反思可以增强教师的批判性思维和反思能力，促进教师的专业发展（Jiang & Zheng, 2021）。此外，教师参与深入的集体反思，可以促进教师之间的相互学习，提高教师的教学实践质量（Hong et al, 2019）。然而，目前的研究表明，当教师面对有争议或困难的问题时，会出现害羞、不喜欢与他人交流、从众心理等问题，这使得反思和发展批判性思维变得更加困难（Darmawansah, Lin & Hwang, 2022）。因此，有必要建立一个能够引导教师集体反思的框架，使反思过程更具组织性、深度和批判性。

在群体活动中，协同调节被认为是多个成员以社会互动和认知过程的形式对彼此的行为进行社会调节的共同行动，也是合作的创造性结果（Sibo et al., 2023）。集体反思的过程也需要教师集体调节自己的思想，激发和利用所有成员的创造性思想进行改进（Sibo et al, 2023）。尽管已有研究试图对协作学习中的群体协同调节机制提供一个整体的理解，但现有的研究框架并未充分考虑教师集体反思过程的协同调节。

六帽思维技术是一种广泛使用的跨领域有效协作的思维策略工具（Jiang& Zheng, 2021），它包含六种不同颜色的帽子，每种颜色代表一种特定的思维方式。当思考者戴上其中一顶帽子时，他们需要采用某种特定的思维方式。通过在教师集体反思中采用六帽思维技术，可能有助于使反思过程更有结构、批判性和深度。六帽思维技术在讨论不同的观点时，也可以防止思维定式，使各种观点得到充分讨论而不偏离主题，从而对群体内部的协同调节产生积极影响（Segal, 2023）。

2. 研究现状

2.1. 集体反思

集体反思是学习者反思他们的小组练习、设计、项目或学习成果的过程（Darmawansah, Lin & Hwang, 2022）。与个人反思不同，集体反思发生在社会背景下，可以通过社会关系反映个体的认知视角，在帮助学习者理解批判性的概念并主动感知批判性学习方面表现出更大的潜力（Lord et al., 2017）。然而，有效的集体反思只有在所有参与者都理解对话的背景并以有组织和积极的方式参与对话时才能发生（Marcos, Sanchez & Tillema, 2011）。缺乏批判性思维、脚手架或指导的过程阻碍了集体反思的有序进行（Marcos, Sanchez & Tillema, 2011）。

集体反思在教师教育中也至关重要。教师反思工作的目的是整合个人和集体的观点、思想、课程知识和课堂经验（Blau & Shamir-Inbal, 2017）。教师通过在教学实践中反思自身的优势和不足，可以更好地了解自身的教学潜力和局限性。集体反思可以提供良好的合作氛围，使教师对合作伙伴更加尊重，形成比个人反思更深的理解（Blau & Shamir-Inbal, 2017）。由于课堂教学实践经验有限，职前教师难以进行深入反思（Allal, 2020），但充分的支持和指导能促进他们在学习过程中进行批判性和有组织的集体反思，从而促进他们的专业发展。因此，要解决职前教师集体反思中的问题，促进专业发展，还需要给予适当的支持。

2.2. 协同调节

学习调节是指对自己的学习进行有意识的、主动的管理，具体表现为学生在元认知、行为、情感、动机等方面的积极参与。当个体学习者进行合作学习时，有三种类型的调节共同起作用（Chan, 2012）。第一，每个小组成员都有学习自律的责任；第二，每个学生支持其他小组成员共同调节学习。协同调节学习是指“自我与他人之间的自我调节暂时协调”，需要与他人合作规划、监控和评估学习（Segal, 2023）；第三，小组共享学习调节。共享调节是指群体成员在协同规划、任务执行和反思等阶段的深思熟虑和战略性适应。

协同调节是多个自我调节者对彼此的学习进行社会调节的共同行为（Segal, 2023）。现有的研究主要集中在构建协同调节的理论模型、调节能力的发展轨迹和培养模式，探索如何促进学习者在学习合作中更好地实现协同调节，提高学习绩效（Su et al., 2018）。然而，现有的研究主要集中在中小学群体，而不是教师教育领域。教师教育研究表明，自我调节和协同调节是教师主动学习的关键。协同调节学习的能力对于职前教师来说可能是非常重要的。然而，简单地把学习者放在一起并不能自动带来合作和富有成效的学习，学习者需要知道如何规范自己的学习和协作[16]。因此，为学习者提供合适的协同学习框架成为加强学习调节的关键。

2.2. 六帽思维技术

六帽思维技术是 Edward de Bono 博士为思维训练而创造的一种思维策略工具，其核心思想是“平行思维”（De Bono, 2017）。平行思维允许思考者在单独或集体思考时一次专注于一件事。当每个人都从同一角度思考时，思考者可以有效地接收和输出信息及想法（Jiang, & Zheng, 2021）。“六帽思维技术”代表红、黄、蓝、绿、白、黑六种不同颜色的帽子，每种颜

色代表一种特定的思考方式。当思想者戴上其中一顶帽子时，他们需要采用一种特定的思维方式。

大多数现有研究探讨了帽子组合的使用及其对协作学习的影响。例如，通过分析专家设计师的话语，Chien（2021）发现了“白蓝”和“黑绿”帽子的组合，“白蓝”组合对设计问题的关注起着重要作用。“黑绿”组合中，黑帽将之前模糊的想法和方案中的遗漏和错误罗列出来，然后再用绿帽思维快速产生新的想法，并进一步细化为设计方案的主要功能点。Hu（2021）探究了不同帽子的转换模式，发现被试是否改变帽子主要取决于该思维模式中是否存在连续的思维生成。然而，现有的研究只探讨了帽子组合对受试者创造性思维、批判性思维和解决问题能力的影响，而没有形成一个更结构化的反思过程框架。

3. 研究方法

3.1. 研究对象

共有 71 名的职前教师自愿参与本研究，他们分别来自地理（12.68%）、语文（26.76%）、数学（22.54%）、英语（30.99%）、体育（7.04%）专业。他们在学科领域有两年以上的学习经历，在各自的学科领域积累了比较丰富的知识。部分职前教师还通过支教、家教等活动积累了学科教学经验。然而，他们对教育技术和教学设计仍然不熟悉。在参加课程之前，他们中没有人了解或使用过六帽思维技术。

3.2. 研究设计

前两周，所有职前教师都通过微课、讨论和问答的方式接受了六帽思维技术的培训。因为掌握程度对六帽思维技术的使用有很大的影响。第二周结束时，所有职前教师都需要完成一项由 De Bono（2017）设计的测试，以检查他们对六帽思维技术的理解程度，本研究计算各组的平均得分，将前 27%的得分归为高水平理解组，后 27%的得分归为低水平理解组。随后，高水平理解组和低水平理解组的参与者被分成 5-6 人一组，每组建立一个 QQ 群进行讨论。

参与者被要求采用六帽思维反思框架来集体反思他们的合作工作。他们按照固定的顺序使用思考帽。蓝帽虽然没有直接出现在固定的顺序中，但在切换帽子颜色时，可以体现出蓝帽在过程控制中的整体作用。固定的帽子序列意在表达设计的直观感受（红色），回顾教学设计的内容（白色），总结教学设计的优点（黄色）和缺点（黑色），改进缺点（绿色）。然后讨论是否有其他需要改进的地方（黑色），并进行改进（绿色），直到所有团队成员都没有提出新的缺点。这种方法将增强讨论的顺序，并从隐含的角度发挥蓝帽子的作用。

在线合作中，他们的在线讨论被记录为他们的讨论话语。

3.3. 数据编码

本研究采用两种编码方案来分析职前教师合作学习的协同调节过程和重点。所有编码方案都改编自先前的研究，这些研究类似地考察了学生在协作学习环境中的协同调节行为（Zhang et al., 2021）。在本研究中，调节重点的编码方案包括任务理解、内容监控和过程监控三个编码（见表 1），其码间信度系数为 0.90。协同调节过程的编码方案包括计划、监控、调节和评估四个编码（见表 2）。评分者信度系数为 0.68。

表 1 调节重点编码

编码内容	描述
任务理解 (TU)	谈论对任务目标和对需求的理解。
内容监控 (CM)	监管举措侧重于检查、细化、修改和改进小组成员的任务响应。

过程监控 (PM)	讨论时间管理。

表 2 调节过程编码

编码内容	描述
计划 (P)	对学习任务的决策或初步安排。
监控 (M)	对行为、认知、动机和情绪状态的意识和监控。
调节 (R)	选择和使用各种认知策略以确保完成学习任务。
评估 (E)	职前教师对其学习表现、教学设计的判断或评价归因。

3.4. 数据分析

本研究采用滞后序列分析对高、低水平理解组的调节过程和调节重点进行了研究。为了分析结果，研究人员将高水平理解组和低水平理解组的编码数据收集成两组数据。为了确定各组的调节行为模式，使用 GSEQ version 5.1 应用进行滞后序列分析。计算每个序列之间的残差值（z 值），以确定该连接是否具有统计显著性。z 值大于 1.96 表示特定序列达到显著性水平（$p<0.05$）。

4. 研究结果

图 1 显示了低水平理解组和高水平理解组调节过程的行为序列转换模式。使用六帽思维技术，所有的小组成员在一定的时间内从同一个角度进行讨论。因此，在高、低水平理解组中都存在着相同行为的过渡模式。从调整后的残差值来看，低水平理解组更有可能具有相同的行为转变模式。在规划学习任务后，高水平理解组会继续讨论计划(3.03)或根据计划监控自己的行为和认知(2.1)。当监控行为发生时，他们更倾向于继续监控行为(5.37)或回到修改计划(2.5)。从行为序列转换模式可以看出，高水平理解组具有规划行为和监控行为的循环模式。

图 2 展示了低水平理解组和高水平理解组调节重点的行为序列转换模式。同样，在低水平理解组和高水平理解组中也出现了相同行为之间的行为转换模式。与高水平理解组相比，低水平理解组在 TU->TU(4.53)行为序列中具有更高的残差值(7.82)。然而，高水平理解组出现 PM->TU(3.57)的行为序列，他们在检查任务目标和需求之前进行了时间管理。收集小组聊天记录并对小组成员话语进行标注后，发现绝大多数学习者对于六帽思维技术的理解比较到位，并且能够正确、充分使用各种颜色的帽子来表达自己的情绪、观点。绝大多数小组长在讨论中扮演着“蓝帽”的角色，组织、控制讨论流程，比如“现在大家戴绿色思考帽，思考课堂导入问题”、“那我们现在对微课设计进行分析，考虑一下可行性、技术的可操作性、使用的环境、软硬件条件吧”。以蓝色帽子为主导和引导，保证了在同一时间内小组成员均佩戴同样的帽子，并站在同一立场思考问题，这有效地扩大了每个人在不同方向上思维的广度，防止受到某一思维的限制。“六帽思维技术”的使用让小组在线讨论过程变得更加有序，几乎避免了讨论混乱、意见不一致的情况，提高了小组的在线讨论的效率，有利于学习活动的进行和学习任务的高效完成。

但也有小组对“六帽思维技术”工具的使用方式不恰当，主要是对各顶帽子的使用意义掌握不当，在讨论过程中出现了错误的搭配情况，为了避免错误情况持续发生，课程中安排了助教加入各个小组的讨论群，一旦讨论出现问题，助教能够及时与小组成员进行沟通，尽可能保证学习者对六帽思维技术工具的使用不产生偏差。

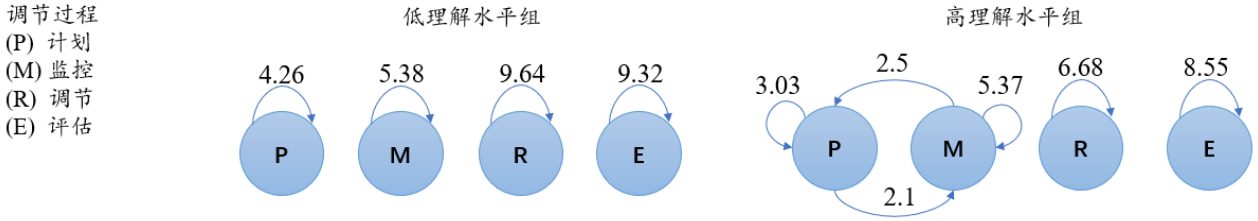


图 1 调节过程模式

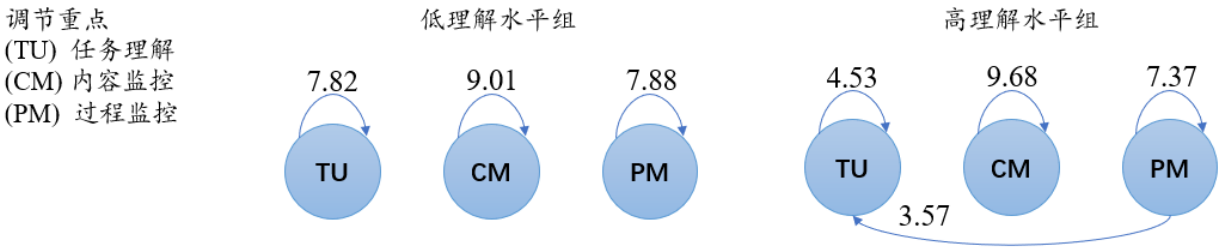


图 2 调节重点模式

5. 讨论与总结

本研究将六帽思维技术运用于职前教师的集体反思中，基于高、低理解水平群体在六个思维帽上的可观察模式，探讨职前教师在教学设计协同学习项目中的调节行为转换模式。

研究结果揭示了职前教师协同调节的行为顺序和重点，有助于运用六帽思维技术进行集体反思。高水平理解组以计划与监控之间的过渡为特征共同调节集体反思，而低水平理解组则不存在这种协同调节模式。“计划”可以确定反思过程的方向，“监控”可以确保反思过程在确定的方向上进行。在高水平理解组中出现了 P->M 和 M->P 的行为序列，反映了蓝帽可以帮助学习者在反思过程中反复确认计划，以及讨论是否按计划进行。在高、低水平理解组中，成员在同一角度反思时都非常重视计划和任务理解，并且出现了相同的调节行为过渡模式（P->P, TU->TU）。然而，由于对六帽思维技术的使用不熟悉，低水平理解组将更多的时间花在自己与他人之间的自我调节的临时协调上，有些人甚至需要更多的调节行为来让反思回到正轨。所以出现了更多相同的行为序列，缺乏计划和监控之间的行为序列。这导致小组成员反复暂停思考，以判断是否正确使用了思考帽，从而揭示了相同调节行为的过渡模式。在高水平理解组中，同一行为的过渡模式是以集体反思过程本身为基础的，其中包含协调部分的行为。高水平理解组创造了一种更持续和顺畅的沟通，试图避免集体中可能会抑制分享信息和发展创新思想的情况。

蓝帽的作用在高水平理解组中更为明显。调节重点结果显示，高水平理解组对学习任务、任务要求和任务目标进行了回顾（PM->TU），并在安排和规划教学设计反思时发挥蓝帽主导作用。而低水平理解组更关注于讨论任务本身。对六帽思维技巧的理解程度较高，更有利于小组明确教学任务，保证讨论按计划进行。他们对蓝帽的使用并不直接体现在以六帽思维技术为基础的集体反思顺序上，而是更加注重蓝帽在过程中的主导作用。从这个意义上说，蓝帽固定了思考帽的顺序，以提供一个框架来规范群体内部的思考和讨论。框架通常构成了一种有价值的综合和创造机制，因为它们允许小组成员通过以其他人可以理解的方式重新构建他们的讨论（Sibo et al., 2023）。六帽思维技术提供了一个趋同的框架，表明团队成员努力提供一个良性的反思的环境，以实现共同的解释。低水平理解组没有认识到六帽顺序与反思过程的对应关系，容易出现反思阶段对任务不清楚，需要花费大量时间来确定任务目标和需求。

本研究运用滞后序列分析方法，以强调六帽思维技术为框架，可视化职前教师在集体反思过程中的协同调节行为模式。本研究揭示了职前教师对六帽思维技术的理解水平在集体反思中的协同调节模式差异。同时，本研究丰富了六帽思维技术在教学教育领域的应用，可以更高效、更有针对性地引入到未来教师教育的课堂中，促进职前教师学术能力和专业能力的发展。

参考文献

Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. Assessment in Education: Principles, Policy & Practice, 27(4), 332-349.

Blau, I., & Shamir-Inbal, T. (2017). Re-designed flipped learning model in an academic course: The role of co-creation and co-regulation. Computers & Education, 115, 69-81.

Chan, C. K. (2012). Co-regulation of learning in computer-supported collaborative learning environments: A discussion. Metacognition and learning, 7, 63-73.

Chien, C. W. (2021). A case study of the use of the Six Thinking Hats to enhance the reflective practice of student teachers in Taiwan. Education 3-13, 49(5), 606-617.

Darmawansah, D., Lin, C. J., & Hwang, G. J. (2022). Empowering the collective reflection-based argumentation mapping strategy to enhance students' argumentative speaking. Computers & Education, 184, 104516.

De Bono, E. (2017). Six Thinking Hats: The multi-million bestselling guide to running better meetings and making faster decisions. Penguin uk.

Hong, H. Y., Lin, P. Y., Chai, C. S., Hung, G. T., & Zhang, Y. (2019). Fostering design-oriented collective reflection among preservice teachers through principle-based knowledge building activities. Computers & Education, 130, 105-120.

Hu, Y., Ren, Z., Du, X., Lan, L., Yu, W., & Yang, S. (2021). The shifting patterns based on six thinking hats and its relationship with design creativity. Thinking Skills and Creativity, 42, 100946.

Jiang, Y., & Zheng, C. (2021). New methods to support effective collaborative reflection among kindergarten teachers: An action research approach. Early Childhood Education Journal, 49, 247-258.

Marcos, J. M., Sanchez, E., & Tillema, H. H. (2011). Promoting teacher reflection: What is said to be done. Journal of Education for Teaching, 37(1), 21-36.

Segal, A. (2023). Rethinking Collective Reflection in Teacher Professional Development. Journal of Teacher Education, 00224871231188702.

Sibo, I. P. H., Celis, D. A. G., Liou, S., Koh, B., & Leung, A. K. (2023). The role of argumentation on high-and low-creative performing groups: A structuration analysis of undergraduate students' group discussion. Thinking Skills and Creativity, 47, 101217.

Su, Y., Li, Y., Hu, H., & Rosé, C. P. (2018). Exploring college English language learners' self and social regulation of learning during wiki-supported collaborative reading activities. International journal of computer-supported collaborative learning, 13, 35-60.

Zhang, S., Chen, J., Wen, Y., Chen, H., Gao, Q., & Wang, Q. (2021). Capturing regulatory patterns in online collaborative learning: A network analytic approach. International journal of computer-supported collaborative learning, 16, 37-66.

基于在线教学视频的弹幕现象分析

Analysis of Barrage Phenomenon Based on Online Teaching Video

徐佳慧[1*]，刘勇[1,2]，刘妙韵[1]
[1]北京邮电大学人文学院 [2]北邮-润尼尔虚拟现实创新技术与应用联合实验室
[*]xujh@bupt.edu.cn

【摘要】 在线视频学习过程中缺乏面对面课堂环境下的及时交流和相应技术支持，师生、生生之间缺乏互动交流，不利于学习者学习。弹幕的出现不仅增强师生之间的关系，而且提升生生之间的交流水平，一定程度上增进教师和学习者、学习者和学习者之间的情感沟通，从而缩短学习者、教师之间的距离。本研究以哔哩哔哩网站初中化学“化学式的书写”视频弹幕文本为研究对象，利用话语分析法从微观-弹幕语言分析、中观-弹幕文本分析、宏观-弹幕文化分析三个角度对弹幕内容进行分析，以助于教师更好地理解在线教学视频中的弹幕现象，进而促进在线教学发展。

【关键词】 弹幕；话语分析；弹幕现象；在线视频学习；哔哩哔哩

Abstract: *Online video learning lacks real-time interactive and corresponding support not as in the face-to-face teaching environment. The lack of interaction between teachers and students and students is not conducive to learners' learning. The emergence of bullet screen not only enhances the relationship between teachers and students, but also improves the communication level between students and students, and enhances emotional communication to a certain extent, thus shortening the distance between them. The paper takes "Writing Chemical Formula" video barrage text as the research object, and conducts discourse analysis from three perspectives: micro-barrage language, middle-view barrage text and macro-barrage culture, so as to help teachers better understand the barrage phenomenon in online teaching videos and promote the development of online teaching.*

Keywords: Barrage; Discourse Analysis; Barrage Phenomenon; Online Video Learning; Bilibili

1. 问题的提出

随着互联网技术的发展，在线学习越来越普遍。但观看课程视频时，视频内容和学习者之间是单向传递信息的过程，师生、生生之间缺少交流互动，导致学习单调且教学质量无法保证[12]。弹幕为解决这一问题提供了有效途径，学习者以弹幕形式发表言论，教师利用弹幕强调课程重点深化学习者对教学内容的认识[14]，促进师生沟通、缩短师生距离[26]，提高用户参与交互的积极性[15]。然而，弹幕在线教学视频中仍存在局限性：一是当视频中的内容引起共鸣时，学习者纷纷发表弹幕讨论，使屏幕上充满大量的弹幕信息，满屏弹幕会遮挡视频内容，影响学习者观看体验[9]。二是有研究发现弹幕信息多为日常口语、语气词和网络流行语等，信息价值不高[17]。三是弹幕中的有些内容根本与教学无关，甚至在教学视频中宣传广告或发布一些恶劣言论，以上这些都对学习者的学习环境造成不利影响，甚至会让学习者对在线学习失去兴趣[16]。

本研究通过对一段在线学习视频中发送的弹幕情况进行分析，分析弹幕的特征以及其反映的文化现象，以便教师更好地了解在线学习视频中的弹幕现象，进而开展良好的基于弹幕交互的在线教学，促进在线学习的发展。

2. 研究现状

“弹幕（Danmaku）”原指电脑上一种射击类游戏，现指在网络上观看视频时弹出的评论性字幕[9]。弹幕会在视频播放过程中从右至左在屏幕上出现，一条条飘过[25]，而能够支持发送弹幕的视频就称为弹幕视频。弹幕视频指用户在观看时可在视频上留言，且留言会以字幕的形式显示在屏幕上，如河南博物院的 AR 弹幕服务[11]。随着弹幕功能的不断完善和交互技术不断提高，弹幕在在线视频教学中开始大展身手。

已有研究在弹幕对学习效果、认知负荷、学习满意度的影响等教育应用方面取得一定成果，在弹幕的学习效果研究方面，林钦杰[18]发现弹幕视频比无弹幕视频学习效果更佳，且当弹幕与学习内容相关时，会促进学习效果；而钱爱平[21]研究则进一步指出，教师发送与视频内容相关的引导性弹幕能显著提高学生的学习成绩，但此类弹幕也会增加学生的认知负荷。在弹幕对满意度影响方面，郭磊[8]研究发现用户认为发弹幕可以消除孤寂感，增加乐趣。同时，与学习内容相关的弹幕能增强学生的学习满意度和社会存在感，从而提高学习成绩。也有研究认为弹幕技术自身的特点可以为解决 MOOC 中“互动性不强”问题指出方向[23]。

上述研究主要从教育应用角度出发，探讨弹幕对学习者在某些方面的影响。但将弹幕作为在线视频播放过程中发送的一种语言的文本表达的研究较少，丁国栋[3]等人从文本、话语实践、社会文化实践三个向度对教学视频弹幕文本进行分析。本研究将采用话语分析法，从弹幕本身的文本特征出发，分别从微观、中观、宏观三个维度对弹幕内容包含的信息开展递进式分析。

3. 研究方法

3.1. 理论基础

弹幕信息多为日常口语、语气词和网络流行语等，基本是日常话语的文本体现。本研究采用话语分析法对弹幕内容进行分析，其目标是解释话语与现实之间的关系，解释隐藏意义[20]，在研究中，语境分为四类：微观话语(对语言的具体研究）、中观话语(仍然是语言研究，但视角更加广泛）、宏观话语(研究话语系统，例如文化）、宏大话语(例如全球化现象)[29]，在话语分析中，微观和宏观研究方法综合起来使用，可以在话语分析中起到解释性的作用。

鉴于本文对弹幕的研究集中在对中文弹幕的分析，并不包括国际化的其他语种弹幕，因此，研究仅从微观、中观、宏观三个维度对在线视频中的弹幕文本进行分析，以此解释我国在线学习视频中的弹幕现象，进而促进我国在线学习的发展。

3.2. 研究对象和研究方法

本研究选取的视频为哔哩哔哩网站（简称为 B 站）于 2019 年 12 月 6 日发布的一节微课“初中化学—化合式的书写（化合价）”[24]，总时长 34 分 24 秒。选取 B 站的原因是它的弹幕文化氛围非常浓厚，其上视频广受青少年的欢迎。B 站开设知识课堂教学栏目，满足青少年既能发弹幕又能进行知识学习的需求，B 站正成为年轻人学习的首要阵地。研究选择的视频面向对象主要为初中生，中学生思维活跃，善于接受新的事物，是 B 站的主要用户。

本研究于 2024 年 1 月 24 日对该视频弹幕信息进行统计时发现，该视频已有 12.5 万的播放量，共计有 1625 条弹幕，可见此视频的受欢迎程度。由于弹幕样本数量足够大，本研究将其作为弹幕分析的研究对象。本研究获取该视频的弹幕文本信息，并存入 Excel 文件中以进行后续数据分析，存储内容（部分）如图 1 所示。图中按照弹幕字数由少到多分别统计对应的弹幕条数，并给出具体的弹幕内容，不同弹幕内容以分号隔开。

基于提取出的弹幕文本，本研究通过话语分析法，分别从微观、中观、宏观三个维度对弹幕信息进行分析。

弹幕字数	弹幕条数	弹幕内容（部分）
1	78	啊；懂；额；负；氯；妙；难；嗯；是；氧；有；这；
2	123	FeO；爱了；不懂；不难；不是；懂了；反着；废了；感谢；好难；
3	193	666；CO2；Na+；Cl-；抵消了；懂了!；二三铁；二四铜；感谢。；
4	216	NaCl；铵根那里；抱佛脚了；CI是氯；抵消了啊；懂啦懂啦；还有酸根；好好听课；哪配平呢；
5	135	不用带符号；给老师鞠躬；核外电子数；恍然大明白；讲的很棒!；老师辛苦啦；那是碘分子；
6	133	1应该是省略；2和2抵消了；纯净物才有!；抵消了没括号；电荷怎么去?；二八稳定结构；二三铁二四碳；
7	131	246硫都齐全；单质化合价为零；单质没有化合价；二价氧钙钡镁锌；氟氯溴碘负一价；钙钡镁锌正二价；

图 1　弹幕文件信息（部分）

4. 研究结果

本研究通过话语分析法对弹幕内容进行分析后的结果如图 2 所示。由图 2 可知，对弹幕内容主要从三个维度开展分析：微观-弹幕语言分析、中观-弹幕文本分析、宏观-弹幕文化分析。其中微观维度是对弹幕语言本身的具体研究，包括对弹幕的字数、内容和句型的特征分析；中观维度是对弹幕文本特点的研究，包括表达特点、文本类型特点以及文本修辞特点；宏观维度即是对弹幕文化的研究，研究弹幕在在线视频中进行互动本身的特征。

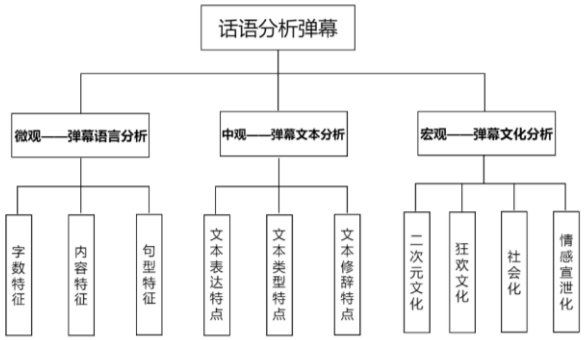


图 2 弹幕内容分析图

4.1. 微观——弹幕语言分析

微观维度主要是对弹幕语言特征的描述，本文主要从弹幕语言的字数特征、内容特征、句型特征三个方面进行弹幕语言分析。

4.1.1. 弹幕语言的字数特征

通过对采集到的 1625 条弹幕文本进行字数分布统计，结果如图 3 所示，其中蓝色标识弹幕字数，红色标识弹幕条数，从左至右，弹幕字数从 1 到 98。

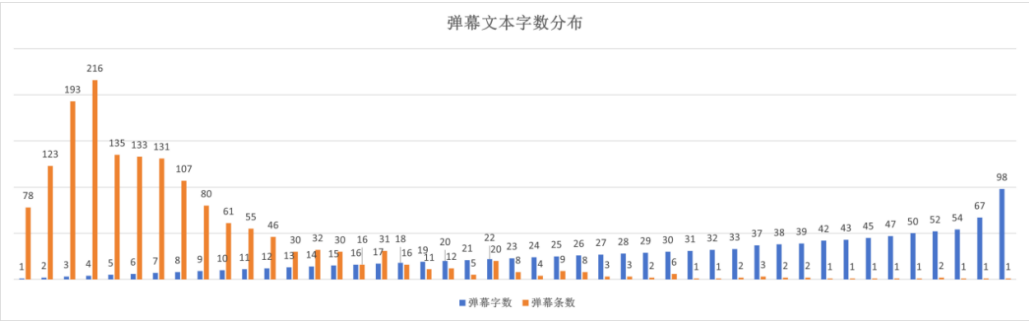


图 3 弹幕文本字数分布图

由图可知，弹幕字数与弹幕条数的变化趋势正好相反，随着弹幕字数从左至右依次增加，弹幕条数呈现减少的趋势。其中 4 个字的弹幕文本数量最多，为 216 条，字数最多的弹幕共 98 个字，只有 1 条。经过统计，每条字数<10 的弹幕为 1115 条，占总条数的 68.6%，可见学习者在学习视频过程中发布的弹幕，更多的是几个词语或者简短的几句话。因此通过对弹幕语言字数进行分析可知弹幕具有字数少、表达意思不完整、内容与所学密切相关的特征。这一特征也符合认知负荷理论[2]，短小精炼的弹幕语言能够减少学习者的外在认知负荷，使学习

者利用有限的认知资源达成更好的弹幕信息获取效果，最终使得弹幕更容易被捕捉和阅读。如果字数过多，再加上弹幕的快速移动，无疑会增加学习者的认知负荷，最终放弃阅读弹幕。

4.1.2. 弹幕语言的内容特征

通过北京佳禾田园科技发展有限公司的在线工具微词云，对所选视频的弹幕文本进行词频分析，得到频数>20 的高频词，统计结果见表 1。结合图 4 的词云图可知，“谢谢、老师”等词语出现的频率最高，这反映在课程学习过程中，学生会对教师的讲解和答疑表示感谢的现象。高频词中的“不是、为什么、可以”等词语反映的是学生在学习过程中思考的过程，“就是”则偏重于口头语，“原子团、离子”则与学习内容相关。由此可知，弹幕语言的内容与所学视频内容“化合式的书写”、师生间互动以及学生学习态度有密切关系。弹幕语言这一内容特征充分反应弹幕能够带来良好健康的学习氛围，促进师生间良好的互动。

表 1 哔哩哔哩视频初中化学——化合式的书写的 1625 条弹幕文本高频词统计

序号	1	2	3	4	5	6	7	8	9	10
高频词	老师	谢谢	不是	化合价	为什么	可以	括号	原子团	就是	离子
词频	56	36	33	29	27	25	24	23	22	21



图 4 弹幕文本词云图

4.1.3. 弹幕语言的句型特征

通过对存入 Excel 的弹幕语言进行分析，发现弹幕中陈述语句 1300 条，感叹语句 122 条，疑问反问句 203 条。陈述语句大多反映的是学习者对视频教学内容的二次重组、学习者之间交互过程以及学习者对视频的吐槽，如“正负化合价和为零、一价氯氢钾钠银、双减时代前来补课”等；感叹语句更多反映的是学习者自己对视频内容掌握情况的表露以及自己的态度，如“涨芝士了！！！！！、就这个不会！！！、懂了懂了！”等；疑问反问语句更多的是学习者针对教学内容来提出自己的问题，希望能得到肯定的答复，在一定程度上反应的是师生之间、生生之间的互动情况，如“核外电子数？、为什么是负二价、三个的怎么交叉”等。选择性理解理论认为受众总是依据自己的价值观念和思维方式对接触到的信息做出独特的个人解读[7]。所以弹幕语言的句型特征更多的是学习者对教学内容的二次重组以加深自己对学习的理解，以及学习者自己对视频内容的态度，使得学习者在同样时间内产生更多的内容价值。

4.2. 中观——弹幕文本分析

中观维度是从更宽泛的文本角度描述弹幕语言特征，本文主要从弹幕文本表达特点、文本类型、修辞特点三个方面进行弹幕文本分析。

4.2.1. 弹幕文本表达特点

弹幕语言多具有网络语言的特点，不单单采用纯文字的表现形式，还有字母、数字、标点符号等非汉语文字的表现形式[10]。本研究通过分析发现，弹幕主要包括以下 3 种表现形式。

纯文字弹幕。纯文字弹幕是弹幕语言的主要形式，如“正负相抵消、我也初二的、五氧化二磷”等，尽管语句并非主谓宾完整，但是纯文字弹幕语义表达清晰，有利于他人理解。

文字加非文字弹幕。B 站作为 ACG（动画、漫画、游戏）内容创作与分享的视频网站，其视频弹幕语言也具有独特的二次元特点，语言中有非汉字网络符号夹杂，包含数字的弹幕如“老师太 6 了、269 报道”等；包含英文字母的弹幕如“我 day 到了、woc 谢谢 up 主”等；包含日语的弹幕如“我狠狠の悟了”等；包含表情的弹幕如“好(✪▽✪)”等等，这类文字夹杂非文字弹幕语句，多为有特定含义的潮词、热词，形成弹幕的一种潮流文化。

只有字母、数字或者其他符号弹幕。字母弹幕如“ohhhhhh！、yyds”等；其他符号弹幕如“ヾ(@^▽^@)ノ、🔔🔔🔔🔔🔔🔔🔔🔔🔔🔔”等；数字弹幕，如“666”等。这类弹幕语言形式需要观众对一些二次元文化有一定了解，否则观众很难理解网络符号，如“666”是用来表示某人或者某物很厉害或很牛的词语，有时也有反讽的含义。此类弹幕需要根据视频内容来进行理解，多比较随意，与上面非文字弹幕的特点相同，也是一种潮流文化的体现。

4.2.2. 弹幕文本类型分析

周舟将弹幕分为预警、版聊、评论、空耳、解说以及狂欢六种[27]。评论、解说弹幕是对当前视频内容的阐述或解释说明。预警弹幕是提醒观众即将出现震惊的画面。版聊弹幕是指与观看内容无关的，纯粹闲聊的弹幕。空耳弹幕是根据原句子的发音，以恶搞或者双关为目的采用谐音法表达出其他意思的句子或短语等[8]。狂欢弹幕类似于刷屏弹幕[27]。钱爱平将弹幕分为与视频内容无关和与视频内容相关的弹幕两种[21]，其中与视频内容无关的弹幕又分为与教师相关、与弹幕界面相关、与生活环境相关、娱乐以及情感宣泄五类弹幕。通过对“化学式的书写”视频的弹幕文本进行分析，将弹幕分为五类，分别为对学习内容的解释重组、空耳、情感宣泄、师生生生互动、与教学无关的弹幕。其示例如表 2 所示。

表 2 弹幕文本类型及示例

弹幕类型	示例
对学习内容的解释重组弹幕	“正三负二”“正负化合价和为零”“ 二价氧钙钡镁锌”
空耳弹幕	“我不李姐”“萌混过关”“芜湖”“哈哈”
情感宣泄弹幕	“好厉害”“太难了”“完蛋了”
师生、生生互动弹幕	“谢谢老师”“好好听课”“别发弹幕了”
与教学无关弹幕	“华夏的同学好”“解元会元状元”“宁波的在吗”

4.2.3. 弹幕文本修辞特点

弹幕语言同时也是网络语言，不受传统语言、语义规范标准约束。弹幕语言中常见修辞手段包括谐音、双关、缩略、反复、反语、用典等[10]。本研究选取的视频弹幕主要采用其中三种修辞手段：谐音、缩略、反复。谐音修辞，如“芜湖”意为“呜呼”；“我不李姐”意为“我不理解”；“饿氧化碳”意为“二氧化碳”等，谐音类弹幕充分体现了弹幕的二次元文化特征。缩略修辞，如“二三铁”其完整语句为“铁元素有化合价有正二价和正三价”；“铁三亚铁是二”其完整意思是“铁离子是正三价，亚铁离子是正二价”等，缩略类弹幕其完整含义与教学视频内容密切相关。反复修辞，有强调的意思，如“懂了 懂了”表明自己对知识点已经学会的态度；“妙啊 妙啊”表达学习者对老师讲课质量的满意以及自己赞叹的心情等，反复类弹幕则是以学习者为中心，表达学习者的强烈态度。上述弹幕语言修辞特点不同于我们在汉语中理解的修辞，而是具有二次元特点，并且要与视频语境相结合才能理解。

通过对弹幕文本表达特点、类型以及修辞三方面进行分析，发现弹幕文本的生成主要源于学习者对学习内容的需求以及情感的需求。马斯洛需求层次理论认为人的需求主要分为生理需求、安全需求、爱与归属需求、尊重需求及自我实现需求五个层次[28]。学习者通过发送弹幕进行师生、生生互动，实现宣泄情感，同时可以根据自己对视频的理解，对视频内容进行二次创作，一切都是基于对爱与归属的需求以及自我实现的需求。

4.3.宏观——弹幕文化分析

宏观维度主要从弹幕交互的二次元文化、狂欢文化、社会化、情感宣泄化四个方面对弹幕文化进行分析。

4.3.1.弹幕交互的二次元文化

“二次元”是一个ACGN亚文化圈的专门用语[1]，其中ACGN为英文Animation（动画）、Comic（漫画）、Game（游戏）、Novel（小说）的首字母合并缩写。“二次元”来自于日语，字面意思是“二维”。日本早期的动画、漫画、游戏等作品都是以二维图像构成，因而被称为“二次元世界”，简称“二次元”。“二次元”正逐渐发展派生出与主流文化相对独立的次文化体系，二次元文化。在线视频中能够体现二次元文化的一个典型例子是空耳弹幕，例如“我不李姐”其原意是“我不理解”，学习者表达对知识的疑惑和不解；“hhhh”和“哈哈哈哈”是一样的意思，表示学习者大笑的心情。二次元文化是与现实若即若离、需要在虚拟世界寻求自我价值的亚文化，B站则是二次元的聚集地。在B站发送弹幕，需通过B站社区考试成为会员才具有发弹幕等权限，在测试题目中专门有一组题目涉及弹幕礼仪。弹幕礼仪是在弹幕视频互动过程中每人都应该遵守的行为准则，如不能恶意狂刷弹幕遮挡视频内容、不用弹幕争吵或说脏话等。正是因为这种严苛的准入机制过滤大量对于二次元文化接受度低的受众。

4.3.2.弹幕交互的社会化

弹幕发展是网络交流媒介发展的必然产物，弹幕的即时互动性满足人们对于互动的要求，用户利用弹幕发表自己的意见，可以与相关群体进行有效互动[19]。如弹幕中的“华夏的同学好、九班的在吗、宁波的在吗”等，是学习者表明自己身份，渴望得到身份认同的表现。弹幕中的“为什么不是三氧化二铝、我懂了”等则是学习者对学习内容的反馈，对于疑惑希望得到他人解答或者认可。

B站的准入机制导致弹幕视频受众具有极高的同质性和相似的价值取向和认同意识[5]。社会认同理论认为“个体认识到他（或她）属于特定的社会群体，同时也认识到作为群体成员带给他的情感和价值意义”[30]。这些弹幕的实时互动为学习者与学习者之间、学习者与教师之间创设一个“动态”的多向交流空间。不同学习者基于兴趣或价值观进行互动，产生一种大家同步学习、彼此给予的现场式陪伴和交流的感觉。

4.3.3.弹幕交互的狂欢文化

B站具有滚动弹幕、彩色弹幕和流沙弹幕等多种弹幕形式，滚动弹幕从右至左进行滚动移动，彩色弹幕中字体颜色可以设置为彩色，流沙弹幕可以呈现流沙飘散效果。学习者发送的弹幕从屏幕上一条条划过，五颜六色、哗哗流淌，营造出一种热闹狂欢的氛围，带给学习者在课堂上集体学习的身临其境的感觉，如视频播放中满屏的“谢谢老师、好好听课”弹幕。这种强烈的临场感和狂欢感有效缓解学习者学习时的孤独感，使学习者学习的积极性提高，同时学习者也通过狂欢式的弹幕来表达学习者自身的情感和教学互动。

在当前弹幕发展过程中可以发现，人们除利用弹幕进行与所观看视频内容有关的实时交流交互外，还利用弹幕对其生活等进行吐槽，这实际是新一轮的文化狂欢，它不仅体现人们的参与性要求，而且为人们战胜社交恐惧提供必要的力量，是一种群体狂欢的新形式[19]。视频上线初期只有小部分学习者发送弹幕以表达自己所感，在此感染和影响下，越来越多的学习者跟随发送弹幕，逐渐形成沸腾的群体效应。心理学认为，个人在面对群体时会受到群体行为的影响从而进一步融入群体，同时增加个体对群体的影响。群体共振之下的个体同步化产生的振荡，会使个体产生群体快感，每位发送弹幕的学习者共同作用就会是一次群体狂欢[6]。

4.3.4.弹幕交互的情感宣泄化

弹幕是学习者表达自身态度的载体，学习者可以通过发送弹幕宣泄自己的情绪。如弹幕中“太难了”是学习者吐槽视频内容的难度，“豁然开朗”则是学习者明白视频内容的态度等。

哈佛大学神经科学家简森·米歇尔和戴安娜·塔米尔认为自由表达对自己本身就是一种内在的奖励，分享个人观点能够取得与获得钱财与食物一样的脑电波[22]。弹幕内容体现的是学习者的知识基础以及学习过程中的困境或情绪，呈现的是学习者的学习状态以及情感状态。

随着视频中弹幕的积累和沉淀，人们在观看弹幕视频的过程中可以看到其他群体在观看视频的不同阶段所产生的不同体验，带给人们情感的共鸣[19]。所谓情感共鸣，是指在他人情感表现或造成他人情感变化的情境（或处境）的刺激作用下，所引起的情感或情绪上相同或相似的反应倾向[13]。弹幕的互动性和即时性使得人们能够更加方便地将自己的观点与态度分享出去，使得更多的人在同一时刻可以得到最佳的情感共鸣体验。

5. 建议

本研究运用话语分析法从微观、中观、宏观三个维度对在线学习视频弹幕进行分析。在微观维度上，从字数、内容、句型三个方面进行特征分析，发现弹幕语言字数较少，表达意思碎片化；弹幕内容与视频本身、师生及生生间互动有密切联系；弹幕语言句型特征更多反映的是学习者对视频内容的二次加工重组以及自我理解。在中观维度上，从文本的表达、类型、修辞三方面对弹幕文本进行分析，弹幕文本包含纯文字弹幕、文字与非文字组合弹幕以及数字符号等非文字弹幕；文本类型分为对学习内容的解释重组、空耳、情感宣泄、师生与生生互动、与教学无关的弹幕五种；弹幕文本修辞包括谐音、缩略、反复三种。在宏观维度上，从弹幕交互的二次元文化、狂欢文化、社会化、情感宣泄化四个方面对弹幕文化进行分析。

为更有效地开展在线教学，促进基于弹幕交互的在线学习，本文提出以下两点建议。

5.1. 规范弹幕语言，建立健康文明学习氛围

弹幕语言是一种典型的青年亚文化，特征是具有青春性、不确定性和反叛性。文化上偏向非主流、边缘亚文化和次文化[4]。弹幕语言深受二次元文化影响，许多网络文字、网络符号在弹幕语言中经常出现。由于国家层面对弹幕的监管只能依据已出台的网络监管法律法规进行，而无法细致贯彻到具体的内容管理，导致在对弹幕的内容审查上出现许多目前无法克服的问题，一些不雅、低俗词语在弹幕上频频出现，严重影响学习者的健康成长，并且不利于在线视频的良性发展。

为更好的规范弹幕语言，建立健康文明氛围，建设绿色和谐健康的网络环境，需要利用技术及时过滤不健康弹幕内容，加强对弹幕内容的监督。通过弹幕屏蔽列表，屏蔽暴力、不健康词汇，为学习者提供健康良好的学习氛围。除利用技术对弹幕进行监控外，在每门在线视频课程上配备一位助教，由助教监控整个在线视频课程的开展，如有不健康弹幕出现，助教及时过滤，并对学习者做出提醒并警告，同时可以建立奖罚机制，每位学习者相互监督，以此建立良好的在线视频学习氛围。

5.2. 提供在线支持服务，助力在线视频教育发展

弹幕虽然可以解决在线视频学习过程中互动不及时的问题，有助于学习者之间加强沟通交流，助于老师与学习者之间的情感互动，但若弹幕信息不加以管理，只是杂乱无章的数据而已。学习者在视频播放过程中，可以根据教学内容中自己不懂的知识点或者感兴趣的内容发表自己的看法或者意见，也可以对其他学习者提出的疑问进行解答，在这过程中，可以由助教老师为学习者提供在线支持服务，为学生提供个性化服务，如对发布的弹幕信息进行及时收集，对学习者共同问题及时进行答疑解惑，个别学习者的问题课下解决，及时解答所有学生问题；同时根据学习者发布的弹幕内容对学习者的学习状况、学习风格、知识点掌握情况、学习满意度等信息进行分析，为教师实施教学提供依据。对于学习者而言，在视频学习过程中，弹幕互动形式可以减轻学习的孤独感，缓解负面情绪，增强在线学习的信心。

综上，通过对哔哩哔哩网站初中化学式的书写这一节视频为例对弹幕进行分析，运用话语分析法从微观、中观、宏观三个维度，对弹幕语言的字数、内容、句型、表达特点以及弹幕文本类型、修辞特色进行分析，分析基于弹幕交互特征的弹幕文化现象及其所反映的在线学习文化，从而更好地理解基于弹幕交互的在线学习文化，以助于高效地开展在线教学。

参考文献

[1]曹超群(2021)。B站二次元文化传播对青年用户群体的影响研究硕士。湘潭大学。

[2]崔宁和方媛(2015)。论网络学习环境下学习者认知负荷影响因素及优化策略。软件导刊(教育技术),09,19-22。https://doi:10.16735/j.cnki.jet.2015.09.008.

[3]丁国栋和杜华(2021)。教学类视频弹幕文本的批判话语分析——以哔哩哔哩网站的某套高中物理教学视频为例。现代教育技术，07，72-79。

[4]付晶晶(2016)。新媒体时代的弹幕文化现象分析。南京邮电大学学报(社会科学版)，02，9-16+89。

[5]付宁和崔政韬(2019)。弹幕视频广告的话语分析。东南传播，06，128-131。

[6]古斯塔夫·勒庞(2007)。乌合之众：大众心理研究。广西师范大学出版社。

[7]广陵(1996)。传播学的选择性理论。新闻爱好者，07，16-17。

[8]郭磊(2015)。我国弹幕视频网站的受众研究硕士。云南大学。

[9]何敏(2019)。“弹幕”视频技术在在线教学中的应用研究。广西广播电视大学学报，01，9-12。

[10]何永波(2021)。新媒介视域下的语言表征研究——网络视频弹幕语言特点及其受众心理分析。大众文艺，19，133-135。

[11]河南省人民政府(2022)。河南博物院推出全国首个AR弹幕服务。河南省人民政府。

[12]黄文有(2017)。弹幕视频技术支持下的在线视频课程交互模型设计。中国教育信息化，14，23-27。

[13]黄希庭(2004)。简明心理学辞典。安徽人民出版社。

[14]李海峰和王炜(2015)。弹幕视频:在线视频互动学习新取向。现代教育技术，06，12-17。

[15]李欢和张筱兰(2016)。弹幕视频应用于MOOC学习中的研究。中国教育信息化，18，84-87。

[16]李健和张新明(2015)。弹幕在在线教育中的应用初探。福建广播电视大学学报，01，5-7+31。

[17]李振华(2015)。弹幕视频在视频教学中的应用研究。软件导刊，02，167-168。

[18]林钦杰(2017)。弹幕视频学习效果及其影响因素研究硕士。福建师范大学。

[19]马黛(2021)。新媒体时代的弹幕文化现象分析。西部广播电视，05，79-81。

[20]毛浩然、徐赳赳和娄开阳(2018)。话语研究的方法论和研究方法。当代语言学，02，284-299。

[21]钱爱平(2017)。在线教学视频中弹幕对学习者的影响硕士。华中师范大学。

[22]乔纳伯杰(2014)。疯传—让你的产品、思想、行为像病毒一样入侵。北京电子工业出版社。

[23]汪基德、冯莹莹和汪滢(2014)。MOOC热背后的冷思考。教育研究，09，104-111。

[24]笑忘书(2019)。初中化学—化学式的书写(化合价)。哔哩哔哩。https://www.bilibili.com/video/BV11E411a7Ur?share_source=copy_web.

[25]张钗(2015)。弹幕视频的互动现状及发展策略。青年记者，05，65-66。

[26]张婧婧、杨业宏和安欣(2017)。弹幕视频中的学习交互分析。中国远程教育，11，22-30+79-80。

[27]周舟(2015)。传播学视野下的网络青年亚文化硕士。西南大学。

[28]Maslow, A H. (1981). Motivation and personality. Prabhat prakashan.

[29]Mats, A. (2000). Varieties of discourse: On the study of organizations through discourse analysis. Hum relat,9,53.

[30]Tajfel, H. (1978). Differentiation between social groups: Studies in the social psychology of intergroup relations. London: Academic press.

损益权衡后的自主调节：行为经济学视角下教师网络研修学习投入决策逻辑研究

The Decision Logic of Teachers' Online Training Participation under the Lens of Behavioral Economics: Balancing Costs and Benefits for Self-Regulation

王　玥[1, 2*]，梁云真[1]，沈子琦[1]，郭琳[1]
[1]河南师范大学 教育学部，河南 新乡 453007;
[2]河南开放大学 教师网络学院，河南 郑州 450000
[*] 949889264@qq.com

【摘要】 作为时间贫困的教师群体，他们会从“性价比”的角度出发，将网络研修成本收益比保持在合理区间，不断调整工学矛盾的调节意愿，进而影响学习投入。研究基于行为经济学理论，分析教师网络研修学习投入的选择逻辑。面向参与网络研修的中小学教师发放 3026 份问卷，根据 SEM 和 fsQCA 互证结果，构建出教师自主调节工学矛盾意愿诱发的四种类型（控“损”逐“益”、拒“延”增“效”、注“力”求“益”、“损”“益”平衡）的选择逻辑，并结合教师的性别、教龄等因素给予各教师群体工学矛盾自主调节意愿的激发策略。以期为减少教师网络研修情景下的工学矛盾冲突，改善研修体验、提升研修效果提供参考借鉴。

【关键词】 工学矛盾；自主调节；决策逻辑；学习投入；模糊集定性比较分析（fsQCA）

Abstract: *As a group experiencing time poverty, teachers approach online training from a cost-benefit perspective, maintaining a reasonable ratio between the costs and benefits of network research to continuously adjust their willingness to resolve work-study conflicts, thereby affecting learning engagement. Grounded in behavioral economics theory, this study analyzes the selection logic behind teachers' decisions to invest in online training. A total of 3026 questionnaires were distributed to primary and secondary school teachers participating in online training. Based on complementary SEM and fsQCA results, four types of decision-making logics are identified for teachers' self-regulation of work-study conflicts: Cost-Controlled Benefit-Seeking, Delay-Avoidance Efficiency-Increasing, Effort-Focused Benefit-Seeking, and Balanced Cost-Benefit. Furthermore, considering factors such as gender and teaching experience, strategies are proposed to stimulate the willingness for self-regulation of work-study conflicts among different teacher groups. This research aims to provide valuable insights for reducing work-study conflicts during online training scenarios, enhancing training experiences, and improving training outcomes.*

Keywords: Work-Study Conflict,Self-Regulation, Decision Logic,Learning Engagement,fuzzy set qualitative comparative analysis (fsQCA)

1. 引言

教育部印发《新时代基础教育强师计划》强调网络研修通过优秀教师资源共享和教学效益最大化促进教育公平与均衡发展，成为教师专业发展主要途径之一(刘永泉、孟凡丽和魏炜，2021)。网络研修初衷是缓解教师工学矛盾，但时间贫困教师因规定课时限制，常选择“倍速刷课”或“挂课”，导致“伪学习”和资源浪费。因此，减少工学矛盾冲突、提升教师网络研修学习投入是教师研修首要任务和薄弱环节, 亟待从理论上寻找解决策略(田爱丽和于天贞, 2017)。

以往研究多聚焦教师网络研修学习投入中项目要素、资源供给和研修模式等，认为教师会基于自我利益最优原则作出理性决策，主动提升学习投入。然而，外部归因和“理性人”假设无法解释教师在研修过程中“非理性”选择，例如：面对更智能化的研修资源推送，部分教师

却不愿接触新内容；教师有能力却仅满足最低任务的完成标准等。问题根源在研究者忽略教师作为决策者解决工学矛盾的主观能动性，过于依赖教师研修外部干预。因此，寻求对“理性人”假设超越、从内生视角激发教师研修意愿，是优化教师网络研修学习投入的重要突破口。

2. 文献综述

2.1. 教师网络研修学习投入影响因素

网络研修通常以教师是否完成视频学习、参与讨论和作业情况等为评判标准(刘璇和崔永鹏，2021)。繁重研修任务、不合理时间安排及重“完成”轻“投入”的考核机制，与教师繁重工作产生工学矛盾，导致部分教师选择“挂机”完成学时(张妮、刘清堂、曾祥翊、李阳和胡小勇，2019)。针对这一问题，研究者多以向外归因为主探索影响因素。张婧鑫等从资源供给角度提出课程资源适配程度、及时性等正向影响学习者的深入程度(张婧鑫、姜强和赵蔚，2019)；李宝敏等从流程模式与培训项目管理提出，工作坊模式以及情境化研修活动可提高教师的研修讨论深度 (李宝敏、黄庆玲，2018)。向外归因的研究可验证技术支持下研修支持服务要素对教师学习投入的影响，为提升研修服务针对性提供理论依据。但研修外部支持仍需通过增强教师研修意愿来促进其学习投入 (张晓娟和吕立杰，2020)。因此，张思等探讨了动机信念、任务感知和研修胜任力对学习投入的直接影响或间接影响 (张思、何晶铭、邓露、吴林静和杨玉芹，2020)；刘清堂等分析了教师的学习意愿、个人体验等因素对网络研修学习投入影响(刘清堂、卢国庆、吴林静、邓伟和王娇娇，2021)。然而，研究若能关切教师网络研修过程中“学习者”与“工作人”角色权衡的复杂性，教师学习投入的决策逻辑可能会有更完整的理解。

2.2. 行为经济学为分析教师学习投入选择逻辑提供新视角

目前，研究者已经将行为经济学的前景理论等分析框架运用在成人学习者网络学习积极性分析等研究领域 (贾云鹏、狄昱函、曾红权和刘青秀，2020)。行为经济学在分析教师网络研修学习投入领域聚焦个体决策的有限理性、深入研究心理因素、关注选择环境和激励机制，成为理解教师学习投入决策有力工具。它认为学习投入决策是平衡预期成本和收益，选择最优方案的过程，即“损益权衡”(刘林澍和刘儒德，2014)。也有研究者认为，成人学习者更看重学习短期收益与成本对比，而忽视长期效用 (李烨、刘云波、刘奎，2022)。由此，合理推论研修教师会从“性价比”角度出发，希望网络研修学习成本和学习收益之间的比例保持合理。

3. 研究模型与假设

3.1. 预期学习成本对教师自主调节工学矛盾意愿的影响

教育领域已有研究者参考现有成本的内涵，将学习情境中决策心理成本细分为预期努力、情感和社会成本等维度(刘林澍和刘儒德, 2014)。这也适用于分析教师参与网络研修影响因素。

预期努力成本是决策者预期完成研修任务时付出的精力、时间等维度的努力。张思等认为，研修教师对学习任务的感知直接或间接影响学习投入(张思、何晶铭、邓露、吴林静和杨玉芹，2020)。预期情感成本指决策者预期任务将引发的负面情绪。刘哲雨等认为，学习者感到能力不足而产生焦虑情绪会导致研修倦怠和学习投入不足 (刘哲雨、周继慧和周加仙，2022)。因此，教师网络研修预期的努力成本、情绪成本越低，越容易产生自主调节工学矛盾意愿。预期社会成本来源于反馈寻求行为，指个体对自身特定行为的合理性和正确性所采取的积极确认(Ashford, S. J. &Cummings, L. L. , 1983)。丁桂凤等认为，学习投入不足导致的社会成本损失越大，可能越促使其参与学习 (丁桂凤、李伟丽、孙瑾、张露和古美乔，2019)。研修教师未完成或低质量完成任务可能会受到同期研修者、单位领导同事的负面评价。因此提出假设：

H1：教师网络研修预期的努力成本对自主调节工学矛盾产生负向影响作用。

H2：教师网络研修预期的情绪成本对自主调节工学矛盾产生负向影响作用。

H3：教师网络研修预期的社会成本对自主调节工学矛盾产生正向影响作用。

3.2. 预期学习收益对教师自主调节工学矛盾的影响

从收益角度看，学习投入是长期和短期收益选择困境：前者指向更好未来职业、收入和社会地位等，后者是解决当前问题和缓解压力等(杨永赞和曾雪云，2022)。李烨等(李烨、刘云波、刘奎，2022)、Lavecchia 等(Lavecchia, A. M., Liu, H., Oreopoulos & P，2016)都认为成年学习者更看重短期收益与成本对比，忽视考虑未来。由此推论，教师网络研修预期长期和短期学习收益越高，越易自主调节工学矛盾，且短期收益作用高于长期收益。因此，提出假设：

H4：教师网络研修预期长期收益对积极自主调节工学矛盾产生正向影响作用。

H5：教师网络研修预期短期收益对积极自主调节工学矛盾产生正向影响作用。

H5.1：短期收益对积极自主调节工学矛盾的正向作用高于长期收益的正向影响作用。

3.3. 教师自主调节工学矛盾意愿积极参与研修的作用

当前，面向教师的各种培训都存在工学矛盾，脱产培训是连续教学任务与教师全身心投入之间无法分身的“场景矛盾”，网络研修是教育活动与研修活动的“时间矛盾”。因此，在关注培训周期和制度保障的同时，需重视教师自发调节工学矛盾、参与网络研修的意愿。教师权衡学习成本与收益后，会做出积极或消极决策，这取决于其调节工学矛盾的意愿程度。这种意愿是引发教师学习的“诱因”并贯穿培训全过程 (王钱永和毛海波，2016)。因此，提出假设：

H6：教师自主调节工学矛盾意愿对研修投入产生正向影响作用。

3.4. 默认偏好对教师自主调节工学矛盾意愿、学习投入之间的影响

已有研究发现，学习决策者常因认知惰性和拖延惯性易选默认选项，可能导致背离理性决策的行为障碍，因其认知和风险成本低 (刘骥和王易昕，2020)。因此，学习者常放弃最优选项，选择熟悉的默认选项，这可能增加辍学风险。(贾云鹏、狄昱函、曾红权和刘青秀，2020)。因此，提出假设，默认偏好对教师工学矛盾的调节意愿、网络研修投入会产生负向影响。

H7：默认偏好对教师网络研修学习投入产生负向影响。

H8：默认偏好对教师自主调节工学矛盾意愿产生负向影响。

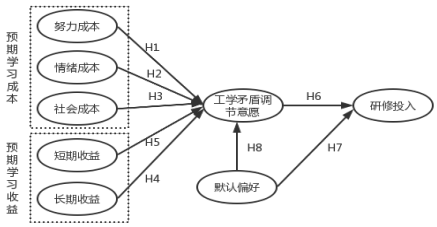


图 1“成本—收益”视角下教师网络研修学习投入的影响因素模型

现有研究主要采用问卷、实验等方法来确定影响因素，并使用结构方程模型（Structural Equation Modeling，简称 SEM）来分析因果变量的净效应，但无法探讨多因素间的协同效应。模糊集定性比较分析法（Fuzzy-Set Qualitative Comparative Analysis，简称 fsQCA）能研究多个不同条件变量如何以“组合”形式共同导致最终结果 (杜运周和贾良定，2017)，可识别出不同条件变量的不同组合对结果贡献程度，有广泛应用前景。总之，本研究试图回答以下问题：

①教师主动调节工学矛盾意愿是否直接影响到他们的学习投入程度；同时，考察预期的学习成本、预想的学习收益、默认偏好等因素对他们调节工学矛盾的意愿的单一贡献。②探究预期学习成本、预期学习收益及默认偏好三个要素对教师调整工学矛盾形成组合效应及其类型。③不同教龄和性别的教师群体是否存在不同的诱发因素组合。

4. 数据收集与分析

4.1. 量表设计

本研究综合了刘林澍等人的决策心理成本框架(刘林澍和刘儒德，2014)、李烨等人提出的收益预期框架(李烨、刘云波、刘奎, 2022)以及 Yi Jiang 等人提出的默认偏好量表(Jiang, Y., Kim, S. I., & Bong, & M.2020)。采用了李克特五级量表。正式发放问卷前进行网络问卷前测，分析 60 份样本，并使用 SPSS 17.0 检验信效度，结果显示，各变量测量量表的信度较高。

4.2. 样本选择

本研究采用网络方式发放问卷，共收集到河南、湖北等省份的 3360 份问卷，经过数据清洗后，有效的问卷数量为 3026 份，有效率为 90.59%。

4.3. 数据的分析结果

4.3.1. 基于 SEM 的净效应分析

研究借助 Amos26.0 检验模型适配度，各个模型配适度指标均显示达到合理水平，表明模型拟合效果良好，结果见表 1。结果表明教师自主调节工学矛盾意愿正向影响教师网络研修学习投入，又受到多重因素影响。其中，除假设 H1 不成立，H2-H8 假设均成立。另外，长期收益比短期收益对教师自主调节工学矛盾意愿影响的标准化回归系数高, H5.1 假设不成立。

表 1 学习投入的影响因素模型回归系数汇总表格

假设	X		Y	非标准化回归系数	SE	Z（CR 值）	p	标准化回归系数	假设是否成立
H1	努力成本	→	调节意愿	0.018	0.056	0.314	0.754	0.014	不成立
H2	情绪成本	→	调节意愿	-0.108	0.036	-2.988	0.003	-0.132	成立
H3	社会成本	→	调节意愿	0.091	0.017	5.36	0	0.128	成立
H4	长期收益	→	调节意愿	0.522	0.018	29.135	0	0.553	成立
H5	短期收益	→	调节意愿	0.038	0.019	2.009	0.045	0.05	成立
H6	调节意愿	→	学习投入	0.751	0.017	42.969	0	0.728	成立
H7	默认偏好	→	学习投入	-0.104	0.012	-9.036	0	-0.126	成立
H8	默认偏好	→	调节意愿	-0.067	0.024	-2.828	0.005	-0.084	成立

4.3.2. 基于 fsQCA 的组态效应分析

研究在已有结论的基础之上，着重分析教师工学矛盾自主调节意愿的影响因素间的组合效应，将成本—收益的 5 个影响维度结合默认偏好因素共 6 个条件进行组态分析。

4.3.2.1. 变量的选取与校准

本研究使用 fsQCA 3.0 软件校准变量。对各变量的测量题项分值求平均值运算发现数据整体非正态分布，因此未将锚点固定为 5、3、1，而是根据各变量数据分布确定，详见表 2。

表 2 变量校准标准

类别		校准标准		
类别	变量	完全隶属	交叉点	非完全隶属
前因构件	社会成本（SC）	5	3	1
	情绪成本（EMC）	5	3	1
	努力成本（EFC）	5	3	1.5
	长期收益（LT）	5	4	2.6
	短期收益（ST）	5	3	1
	默认偏好（DP）	5	3	1
结果	调节意愿（AW）	5	3.8	2.8

分析以 AW 和~AW 为结果的各单项前因条件必要性，一致性水平均未超 0.9，表明任一独立前因条件均不足以构成教师网络研修工学矛盾自主调节意愿必要条件（拉金 Ragin, Charles C，2019)。

4.3.2.2. 整体组态分析

本研究设定一致性阈值 PRI 为 0.8，案例频数阈值为 6。分析简约解和中间解的组态路径，核心条件为两者共有，边缘条件为仅中间解具备(拉金 Ragin, Charles C，2019)。如表 3 所示。

共得到 4 条诱发高调节意愿的前因路径(H1-H4)和 4 条诱发低调节意愿的前因路径(L1-L4)。单条路径与整体一致性值均高于 0.7，总体覆盖率分别为 0.76（H1-H4）和 0.84（L1-L4）。这表明已获组态路径可以有效研究教师工学矛盾调节意愿的诱发机制。

表 3 教师工学矛盾调节意愿诱发路径

前因构件	高调节意愿				低调节意愿			
	H1	H2	H3	H4	L1	L2	L3	L4
努力成本			•	•	•	•	●	
情绪成本	⊗	⊗		•	•		•	⊗
社会成本	⊗		⊗	●		⊗		⊗
短期收益		⊗		•	•		●	●
长期收益	●	●	●	●	⊗	•	•	•
默认偏好	⊗	⊗	⊗		•	⊗	⊗	⊗
一致性	0.78	0.8	0.75	0.77	0.92	0.79	0.86	0.82
原始覆盖率	0.55	0.5	0.45	0.42	0.43	0.36	0.34	0.33
净覆盖率	0.03	0.02	0.02	0.02	0.19	0.05	0.03	0.02
总体一致性	0.77				0.73			
总体覆盖率	0.76				0.84			

注：●表示核心条件存在；•表示边缘条件存在；⊗表示核心条件缺失；⊗表示边缘条件缺失。

4.3.2.3. 分层组态分析

为揭示教师性别和教龄差异，研究进行了组态分析，详见表 4、5。结果表明，教龄分层与总路径相似度达 86%和 75%，性别分类与总路径相似度达 100%，验证了模型预测性稳定。

表 4 高工学矛盾调节意愿诱发路径的教龄比较与性别比较

前因构件	高工学矛盾调解意愿										
	教龄比较							性别比较			
	SH-1	SH-2	MH-1	MH-2	MH-3	LH-1	LH-2	MAH-1	MAH-2	WH-1	WH-2
努力成本		•			⊗		•		•		•
情绪成本	⊗	•	⊗	⊗		⊗		⊗	•	⊗	•
社会成本	⊗	●		⊗	⊗	⊗	⊗	⊗	●	⊗	●
短期收益		•			⊗				•		•
长期收益	●	●	●	●	●	●	●	●	●	●	●
默认偏好	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗		⊗	
一致性	0.76	0.76	0.81	0.79	0.81	0.79	0.76	0.79	0.78	0.78	0.76
覆盖率	0.62	0.36	0.53	0.57	0.46	0.51	0.42	0.48	0.5	0.57	0.39
净覆盖率	0.38	0.11	0.03	0.06	0.01	0.12	0.02	0.18	0.2	0.17	0.13
对应总路径	H1	H2+H4	H1	H1	无	*H3	H3	*H3	H4	H1	H4
总体一致性	0.76		0.78			0.78		0.78		0.76	
总体覆盖率	0.73		0.73			0.74		0.75		0.772	
与总路径相似度	86%							75%			

注：LH-1，MAH-1 与 H3 路径只在是辅助条件缺失情况不一致，也在一定程度上支持了 H3。

表 5 低工学矛盾调节意愿诱发路径的教龄比较与性别比较

前因构件	低工学矛盾调解意愿												
	教龄比较								性别比较				
	SL-1	SL-2	SL-3	ML-1	ML-2	LL-1	LL-2	LL-3	MAL-1	MAL-2	MAL-3	WL-1	WL-2

努力成本	.	.		.	.	.	●		.	●	●	.	●
情绪成本	.	⊗	⊗	.		.		⊗	.	.		.	
社会成本		⊗	⊗					⊗			⊗		⊗
短期收益	.	⊗	●	.	●	⊗		●	.	●		.	
长期收益	⊗	⊗	.	⊗	.	⊗	.	.	⊗	.	.	⊗	.
默认偏好	.	⊗	⊗	.	⊗	.	⊗	⊗	.	⊗	⊗	.	⊗
一致性	0.83	0.95	0.83	0.92	0.89	0.91	0.76	0.81	0.92	0.85	0.82	0.93	0.76
覆盖率	0.4	0.37	0.35	0.43	0.36	0.45	0.38	0.31	0.46	0.34	0.34	0.42	0.4
净覆盖率	0.09	0.04	0.04	0.1	0.02	0.2	0.05	0.01	0.2	0.03	0.04	0.12	0.04
对应总路径	L1	无	L4	L1	L3	*L1	L2	L4	L1	L2+L3	L2	L1	L2
总体一致性	0.82			0.77		0.76			0.77			0.77	
总体覆盖率	0.62			0.66		0.62			0.62			0.64	
与总路径相似度	75%								100%				

注: LL-1 与 L1 路径只在短期收益是辅助条件的存在还是缺失情况不一致，也部分支持了 L1。

5. 结果讨论

5.1. 教师自主调节工学矛盾意愿决策类型的组态视角讨论

5.1.1. 控“损”逐“益”型：情绪等成本管理与长期收益拉动

网络研修中高工学矛盾调节意愿的组态中，路径 H1 原始覆盖率最高为 55%。其中预期长期收益与情绪成本缺失是高调节意愿的核心条件，社会成本、默认偏好作为边缘条件缺失。由此可见在预期学习情绪和社会成本低、长期收益高时，教师更愿调节工学矛盾。持此逻辑的教师预期网络研修带来长期收益，因此重视积极情绪对学习效果的影响，采取策略降低情绪成本(王雪、乔玉飞、王鉴羽、成雨薇和李慧洋，2022)。低工学矛盾调节意愿组态 L1 呈现的该群体教师不相信长期收益，倾向于拖延和推迟决策，并可能在面对较大的努力或情绪成本时，更容易选择回避而不付诸行动。仅凭短期收益拉动只有较低的工学矛盾调节意愿。

5.1.2. 拒“延”增“效”型：研修偏好改善与职业长远发展牵引

高工学矛盾调节意愿组态路径 H2 中，预期长期收益的存在和默认偏好的缺失是该路径的核心条件，预期短期收益和情绪成本作为边缘条件缺失。同时，在结构方程模型中，这种默认偏好对调节意愿产生了显著的负向影响。教师有高长期收益预期，意识到默认偏好可能导致拖延和不利选择。因此，教师会采取策略如设定目标、制定计划、建立自律和寻求支持来优化学习效果(刘骥和王易昕，2020)。低工学矛盾调节意愿组态路径 L3、L4 中，即使考虑长期收益，教师仍更倾向追求即时回报，即使能克服惰性，也难产生较高工学矛盾调节意愿。

5.1.3. 注“力”求“益”型：长期收益驱动下主动付出努力成本

高工学矛盾调节意愿的组态路径 H3 中，预期长期收益作为唯一的核心条件、努力成本作为边缘条件存在，与社会成本、默认偏好作为边缘条件缺失的组合作用下，产生较高的工学矛盾调整意愿。持此种逻辑的教师对网络研修的长期回报有高预期，认为学习需长期积累，因此愿付出额外努力提升技能。相反，在组态路径 L2 体现的预期努力成本较高、预期长期收益有限情况，教师调节工学矛盾的意愿较弱，易退缩和回避压力，不愿追求更高长期收益。因此，权衡长期收益和努力成本时，对两者看重程度不同会导致不同决策结果。

5.1.4. “损”“益”平衡型：多重收益期待与压力下的组织服从

由高调节意愿解释路径 H4 可知，预期学习成本和预期学习收益的各要素均出现在该路径，其中社会成本、长期收益作为核心条件。结构方程分析显示，长、短收益预期均对调节意愿有显著的正向影响，且长期收益预期（β=0.553，p<0.001）影响系数高于短期收益（β=0.05，p<0.05）。此种逻辑主要体现为，高度关注长期收益并综合权衡努力成本、情绪成本、社会成本、默认偏好等因素。面对学习中的风险和挑战，如情绪和组织压力，教师会灵活应对，采取策略以提高研修回报并应对压力，包括情绪管理、寻求支持、合理安排学习及调整方法。

5.2. 教师自主调节工学矛盾意愿群体差异的分层组态视角讨论

5.2.1. 成熟型教师和男性教师往往以追求长期收益为动力主动调节工学矛盾

在组态分析的教龄比较中，长期收益的单一驱动作用在成熟型教师调节工学矛盾意愿过程中最突出，两条高调节意愿诱发路径都属于该类型。新入职教师和成长型教师激发高调节意愿，长期收益预期则需结合其他核心条件。这与已有研究结论“教师培训 MOOC 中个体参与度最高的是处于平静保守期（46~55 岁）和职业消退期（55 岁以后）的资深年长教师”保持一致 (詹泽慧、邵芳芳、范逸洲、何国庆、姚佳静和汪琼. 2020)，此时期教师自主学习和发展需求强烈。同样地，男性教师调节意愿与长期收益追求相关，女性教师则无此路径（详见表 5）。因此，应充分重视成熟型及男性教师的长期收益预期，为他们提供自主学习和发展机会。

5.2.2. 成长型教师、女教师和部分新入职教师通过管理负面情绪调节工学矛盾

分层组态分析显示，在女性教师、新入职教师和成长型教师部分的路径（WH-1、SH-1、ML-1、ML-2）显示，教师在较低情绪成本和较高长期收益的共同作用下，工学矛盾调节意愿较高。由于女性教师在整体中占比较大，情绪成本的影响不容忽视。因此为提高教师网络研修积极性，应注重及时疏导和科学管理负面情绪(刘哲雨、周继慧和周加仙，2022)，特别是针对女性教师、新入职教师和成长型教师，以减轻他们在工学矛盾中的情绪成本负担。

5.2.3. 新入职教师存在组织压力下克服惰性被迫调节工学矛盾的特殊群体

研究发现，社会成本对调节意愿有显著的正向影响（β=0.128，p<0.001）。为避免因低质量完成研修任务而面临的负面评价和不良后果，教师会主动调节工学矛盾并积极参与研修。在分层组态分析中，男女教师均因组织压力而参与网络研修。在教龄的比较中，新入职教师呈现出一条组织压力型影响路径，占据该分层高调节意愿路径的 50%，而成长型和成熟型教师则不受此影响。因此，新入职教师在组织压力下，为避免负面评价，会克服惰性，调节工学矛盾并积极参与研修。

6. 总结与建议

本研究总结出，控“损”逐“益”、拒“延”增“效”、注“力”求“益”、“损”“益”平衡四类教师工学矛盾调节意愿较高的决策类型。同时，分层组态分析结果表明高调节意愿的诱发路径呈现多元类型。性别方面，女性教师的情绪成本有效控制对工学矛盾调节意愿有重要影响；教龄方面，成熟型教师更多依靠长期收益驱动，成长型教师依靠长期收益与情绪成本降低的组合作用，新入职教师的诱发路径最为多元，除长期收益、社会成本的核心条件外，短期收益以及努力成本、情绪成本都会对其产生协同影响。因此，提高教师网络研修学习投入度，可以针对不同选择逻辑、不同性别、不同职业发展阶段的教师提供更为精准的对策支持。

参考文献

丁桂凤，李伟丽，孙瑾，张露，& 古美乔. (2019). 小微企业创业失败成本对创业失败学习的影响——内疚的中介作用. **研究与发展管理, 31(4)**, 11.

王钱永, & 毛海波. (2016). 基于 utaut 模型的 mooc 学习行为因素分析. **电化教育研究, 37(6)**, 6.

王雪, 乔玉飞, 王崟羽, 成雨薇, & 李慧洋. (2022). 教育智能体如何影响学习者情绪与学习效果? ——基于国内外 39 篇实验或准实验研究文献的元分析. **现代教育技术, 32(8)**, 8.

田爱丽, & 于天贞. (2017). 任务导向的慕课研修模式分析. **教师教育研究, 29(5)**, 7.

刘永泉孟凡丽魏炜. (2021). 在线教育助推义务教育优质均衡发展的价值意蕴及作用机理. **中国电化教育, 000(010)**, 86-90,119.

刘林澍, & 刘儒德. (2014). 远程学习者辍学影响因素分析与对策——学习情境中决策心理成本的视角. **中国电化教育(8)**, 6.

刘哲雨, 周继慧, & 周加仙. (2022). 教育神经科学视角下促进心流体验的智慧教学活动设计. **现代教育技术, 32(7)**, 14-21.

刘清堂,卢国庆,吴林静,邓伟 & 王娇娇.(2021).中小学教师参与混合式研修满意度的影响因素研究.**现代远距离教育(01)**,3-12.

刘璇, & 崔永鹏. (2021). 非线性学习视域下教师智能研修系统设计与应用研究. **中国电化教育(9)**, 7.

刘骥, & 王易昕. (2020). 行为经济学在教育领域的前沿实践与启示. **教育与经济, 36(4)**, 7.

张妮, 刘清堂, 曾祥翊, 李阳, & 胡小勇. (2019). 信息技术支持下的区域研修现状及发展研究——基于全国 25 省(市)35 县(区)的调研. **中国电化教育(10)**, 8.

张思, 何晶铭, 邓露, 吴林静, & 杨玉芹. (2020). 网络研修社区中教师的动机信念对学习投入的影响. **中国电化教育(4)**, 9.

张晓娟, & 吕立杰. (2020). 精准扶贫背景下教学点教师远程培训路径探索--以 spoc 引领式培训模式为支持. **中国电化教育(2)**, 58-66.

张婧鑫, 姜强, & 赵蔚. (2019). 在线学习社会临场感影响因素及学业预警研究——基于 coi 理论视角. **现代远距离教育(4)**, 10.

李宝敏 & 黄庆玲.(2018).助学者支持与教师工作坊在线讨论深度关系研究——以信息技术应用能力提升工程教师工作坊为例.**电化教育研究(07)**,97-102+113.

李烨, 刘云波, & 刘奎. (2022). 行为经济学视角下终身学习困境的破解路径. **现代远程教育研究, 34(1)**, 8.

杜运周 & 贾良定.(2017).组态视角与定性比较分析(QCA):管理学研究的一条新道路.**管理世界(06)**,155-167.

杨永赞, & 曾雪云. (2022). 未来自我连续性对大学生学习投入的影响:未来结果预期的中介作用. **中国健康心理学杂志(004)**, 030.

美 拉金 Ragin, Charles C. (2019). **重新设计社会科学研究**. 机械工业出版社.

贾云鹏,狄昱函,曾红权 & 刘青秀.(2020).教育经济学的新发展: 教育行为经济学介评.**教育与经济(02)**,79-87.

詹泽慧, 邵芳芳, 范逸洲, 何国庆, 姚佳静, & 汪琼. (2020). Mooc 与教师专业化发展:积极老龄化的大数据发现. **中国远程教育(6)**, 12.

Ashford, S. J. , & Cummings, L. L. . (1983). Feedback as an individual resource: personal strategies of creating information. Organizational Behavior & Human Performance, 32(3), 370-398.

Lavecchia, A. M., Liu, H., & Oreopoulos, P. (2016). Behavioral economics of education: Progress and possibilities. In Handbook of the Economics of Education (Vol. 5, pp. 1-74). Elsevier.

Jiang, Y., Kim, S. I., & Bong, M. (2020). The role of cost in adolescent students' maladaptive academic outcomes. Journal of School Psychology, 83, 1-24.

基于系统性文献综述的 XR 赋能空间能力培养实证研究分析

Empirical research and analysis on XR-enabled spatial ability training based on systematic literature review

刘西庆 张双秀 张慕华*
首都师范大学初等教育学院
6960@cnu. edu. cn

【摘要】 空间能力对科学、技术、工程和数学（STEM）领域的学习至关重要。扩展现实（XR）已被证明是培养空间能力的有效工具。鉴于空间能力的内隐性和复杂性，鲜有研究厘清 XR 如何赋能空间能力培养。研究采用系统性文献综述的方法，旨在全面描绘和梳理 XR 赋能空间能力培养的研究现状、空间能力测评方法，以及 XR 在空间能力培养方面的已有证据。在对国内外 2000 年以来发表在 SSCI/CSSCI 期刊的 29 篇实证研究系统分析后发现：XR 赋能空间能力培养的研究集中在教育学、医学和计算机科学三个领域，心理旋转测试 Mental Rotations Test (MRT)是空间能力测量的主要方法，XR 能够显著提升心理旋转技能和空间可视化方面的能力，但同时 XR 学习可能也对学生的空间能力提出了更高的要求。

【关键词】 空间能力；扩展现实；实证研究；研究特征；测量方法

Abstract: Spatial ability is crucial for learning in the fields of science, technology, engineering, and mathematics (STEM). Extended reality (XR) has been proven to be an effective tool for cultivating spatial ability. Given the implicit and complex nature of spatial ability, few studies have clarified how XR can empower the cultivation of spatial ability. This study adopts a systematic literature review method to comprehensively depict and sort out the research status of XR-enabled spatial ability cultivation, spatial ability assessment methods, and existing evidence of XR in spatial ability cultivation. Based on a systematic analysis of 29 empirical studies published in SSCI/CSSCI journals since 2000, it was found that XR-enabled spatial ability cultivation research focuses on three fields: education, medicine, and computer science. The Mental Rotations Test (MRT) is the main method for measuring spatial ability. XR can significantly improve mental rotation skills and spatial visualization abilities, but at the same time XR learning may also put higher demands on students' spatial ability.

Keywords: Spatial ability; Extended Reality; Empirical Research; Research features; Measuring instruments

1.研究背景

空间能力是一种重要的智能，其在科学、技术、工程和数学领域的学习中有着极其重要的作用（Buckley et al., 2018）。空间能力可以较好地预测学生在 STEM 领域的学业表现，同时空间能力可以帮助学生识别视野中的物体，分析其形状大小并建立对应的心理表征（Carroll, 1993）。空间能力因素分析也被认为有助于设计干预措施以提高 STEM 教育绩效（郭守超, 2022）。空间能力的个体水平不同，它会受性别（Gonzales et al., 2020）、与运动相关的活动、一般儿童活动或电子游戏（Tumkor & Serdar, 2018）的影响。已有研究证明在虚拟技术辅助学习中拥有高空间能力水平的学生能够表现地更好（Chen & Chen, 2022）。空间能力具有可塑

性，这一特性在低空间能力的学生中尤为显著（Lee & Wong, 2014）。然而，目前仍缺乏关于如何最有效地提升这些学生的空间能力的研究与探索（Uttal et al., 2013）。

空间能力的概念最初由美国心理学家高尔顿（Galton）提出，他将其定义为：对个体最近考察或研究过的任何对象，能够随意地唤起一个清晰、稳定、完整的心理形象的能力，包括从任何方面自由地观想这个对象、并可以随意的改变对象、把对象的任何图像投射到纸上等能力。加德纳多元智能理论认为，每个人具有言语语言、数理逻辑、视觉空间、音乐韵律、身体运动、人际沟通、自我认知、自然观察八项智能(钟志贤，2004)。加德纳通过实践证明，每一种智能都具有同等的重要性，在人类认识和改造世界的过程中发挥着巨大的作用。其中视觉空间智能是指人对色彩、形状空间位置的正确感受和表达能力，突出特征为对视觉世界有准确的感知，产生思维图像，有三维空间的思维能力，能辨别感知空间物体之间的联系。虽然加德纳将空间能力描述为一种智能，但生产、生活、研究中仍采用空间能力（郭守超, 2022）。越来越多的文献表明，空间能力具有适度的可塑性，并且能够在特定的环境下得到有效的提升（Di & Zheng, 2022）。一些研究人员就开始探究如何在虚拟技术支持的环境中（VR/AR/MR）培养学生的空间能力（Conesa et al., 2023；Darwish et al., 2023）。许多的研究报告了虚拟技术对学生空间能力的积极影响，尤其是在训练学习者的空间能力方面（Zeynep & Ömer 2020）。

虚拟现实(VR)、增强现实(AR)和混合现实（MR）是指 HCI 研究人员几十年来研究的空间界面技术和概念主张的总和（Çöltekin et al., 2020）。在社会生活中，“扩展现实”（或 XR）已被采用作为 VR/AR/MR 技术的总括术语，本文所提及的 XR 也均指 VR/AR/MR 三种技术的总称。近十年，随着技术发展，更便宜、更轻便的设备以及更丰富软件生态出现，XR 开始广泛应用在社会各个领域。XR 不仅仅是简单的 3D 展示或环境虚拟，沉浸式界面提供基于可视化和交互设计的强大体验，改变了我们的空间感（Çöltekin et al., 2020）。正因为这两个特性，XR 可以作为构建与培养空间能力的重要媒介。本文对 XR 应用于培养空间能力的研究作综述，旨在通过回顾现有空间能力和虚拟现实、增强现实、混合现实、扩展现实相关的研究，尝试阐明 XR 技术应用对培养学生空间能力的影响。我们特别关注研究的文献特征，筛选实证研究为主的文献；探究空间能力的测量和特征；以及 XR 对培养空间能力的潜在贡献。

2.研究设计

2.1. 研究方法

本研究应用系统性文献综述法开展研究。系统性文献综述法符合研究的科学性，研究过程清晰，是一种系统的教育研究方法。系统性文献综述法首先通过全面检索和严格筛选保证研究对象的可靠性与客观性，进而对作为研究对象的文献进行编码分析，最终完成对知识的整合以及获得更客观的结论。

2.2. 研究问题

已有的大部分综述研究集中在基于 AR 与 VR 培养空间能力的发展上，但很少有研究关注 XR 与空间能力之间的关系。因此，本文探讨了进一步收集和分析这方面数据的必要性。为了填补文献上的空白，本研究确定了如下研究问题：

（1）国内外基于 XR 培养空间能力的实证研究主要呈现哪些特征（研究地区、研究对象、研究领域、研究时间、研究方法等）？

（2）国内外空间能力的维度划分以及使用什么评估工具测量不同维度空间能力的提升？

（3）XR 培养空间能力的实证研究聚焦哪些研究主题？

2.3. 样本获取

2.3.1. 文献检索策略

为获取国内外基于 XR 培养空间能力的高质量实证研究文献，本研究于 2024 年 1 月在

Web of Science 文献数据库进行检索。因为考虑以 XR 为检索词会出现检索内容不全面的情况，因此将 XR 分为 AR、VR、MR 和 XR 四个部分，限定研究领域为教育教学，进行分别检索。

基于 AR 培养空间能力的文献检索词为“AR”OR“augmented reality”OR“augmented reality system”OR“AR tools”OR“AR application”（Topic）and“spatial ability"OR“spatial skills”（Topic），得到 41 篇论文。VR、MR、XR 检索式与 AR 类似，分别得到 58 篇、4 篇和 2 篇论文。

2.3.2. 文献筛选标准

本研究基于研究问题，制定了针对初步检索到的 105 篇文献的文献筛选标准。其中 1-5 条是系统性文献综述法为了保证研究样本的准确性与权威性而普遍采用的筛选标准。6-8 条的目的是保证所选文献是教育领域聚焦于空间能力培养。根据第一轮筛选的纳入标准，检查了 105 篇文献的标题和摘要，69 篇文献被剔除，剩下 36 篇文献有待进一步研究。在筛选的第二阶段，根据排除标准审查了 36 篇论文的全文，剔除 7 篇文献。最终，筛选出 29 篇满足 meta 分析标准的文献。

表 1 文献纳入\排除标准

序号	纳入标准	排除标准
1	全文下载	全文无法下载
2	文章只出现一次	重复文章
3	英文论文	非英文论文
4	期刊论文	会议论文、评论文章等
5	实证研究	非实证研究
6	研究主题聚焦基于 XR 培养空间能力	研究主题非基于 XR 培养空间能力
7	研究领域为教育领域	研究领域非教育领域
8	研究的变量包括空间能力	研究的变量不包括空间能力

2.3.3. 文献筛选流程

本研究遵循系统性文献综述及元分析方法（Preferred Reporting Items for Systematic Reviews and Meta-Analyses，PRISMA）的研究思路，通过检索、筛选、合格和纳入 4 个阶段对基于 XR 培养空间能力的相关文献进行筛选，具体 PRISMA 筛选过程如图 1 所示。针对筛选后的 29 篇文献，研究从研究地区、研究对象、研究领域、研究时间、研究方法等维度对其进行编码分析。

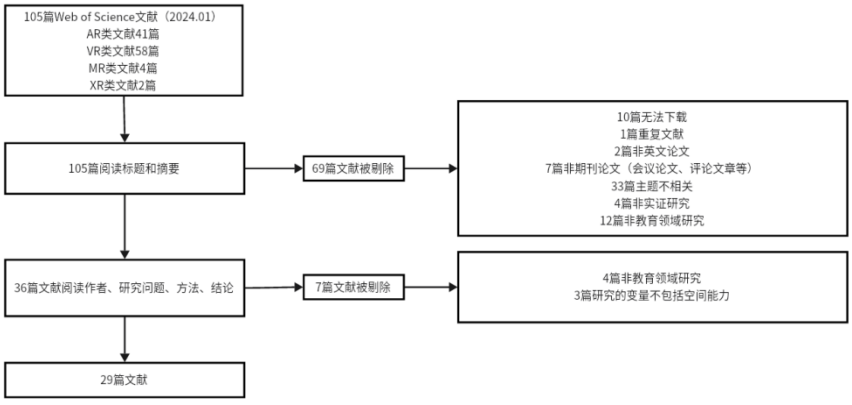


图 1 PRISMA 流程图

3.基于文献计量分析的 XR 培养空间能力实证研究现状

本研究聚焦国内外 XR 应用于培养空间能力的实证研究，重点关注 XR 培养空间能力的具体方式与所培养空间能力的具体维度等。以下将按照文献发表、作者分布、研究设计、研究工具以及研究方法等来具体分析国内外 XR 培养空间能力实证研究的基本特点。

3.1. 文献发表与引用情况

从发表时间来看，国内外有关 XR 应用于培养空间能力的实证研究起步较早，可以追溯至 2007 年，并在随后的 17 年时间内受到持续关注。其中值得注意的是 2019 年至今，相关研究有一个明显的增长趋势（见图 2），表明国内外研究者长时间关注该领域，特别是近 4 年有更多研究者关注到了该研究领域，并有了更多成果发表。从发表的期刊分布来看，29 篇文献发表在 22 种不同的期刊上，其中发表在教育类期刊上的文章有 19 篇（65.5%），发表在医学类期刊上的文章有 4 篇（13.7%），发表在工程技术类期刊上的文章有 3 篇（10.3%），其余为综合性期刊（13.7%）。

从文献引用率来看，有 12 篇文章的引用率在 20 次及以上，超过 50 次的文章有 4 篇。其中引用率最高的文章是 Lee 等（2014）研究使用桌面虚拟现实学习：空间能力低的学习者受到更积极的影响（229 次），其次是 Jang 等（2017）在三维虚拟现实环境中学习解剖学时，直接操作比被动观看更好（186 次）。其中值得注意的是引用超过 20 次以上的 12 篇文献中有 7 篇（58.3%）是和医学教育相关，引用超过 50 次以上的 4 篇文献中有 2 篇（50%）是和医学教育相关，这可能是由于 VR 应用于医学培训相关研究起步较早，相关研究众多的原因。

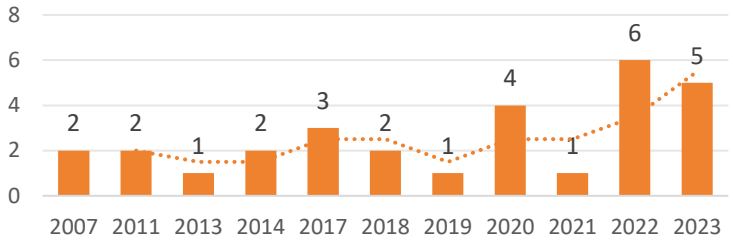


图 2 发表时间分布图

3.2. 作者分布特征

为展现国内外有关 XR 应用于培养空间能力实证研究的作者分布情况，本研究以第一作者为分析单位，对 29 篇样本文献进行分析，发现作者总共来自 13 个国家，其区域分布（图 3）分别是亚洲（n=10，34.5%）、北美洲（n=9，31.0%）、欧洲（n=8，27.6%）、大洋洲（n=1，3.4%）以及非洲（n=1，3.4%）。其中发文作者数量最多的三个国家和地区分别是：美国 7 位（24.1%）、中国台湾 5 位（17.2%）、西班牙 3 位（10.3%），可见美国学者是 XR 应用于培养空间能力实证研究的主要群体。

3.3. 研究设计特征

为进一步分析 XR 应用于培养空间能力的研究情境与实验设计，本研究对样本文献的学科背景、研究对象、样本数量以及研究周期等指标进行了系统分析。

学科背景分析发现，在教育学领域中，国内外 XR 应用于培养空间能力的研究主要集中在医学、计算机科学这两类学科的高等教育阶段，化学、地理学、工程学、材料科学等学科的研究者也在关注 XR 与空间能力。

从研究对象的类别来看（图 4），国内外 XR 应用于培养空间能力的研究中，以大学生为研究对象的占 65.5%（n=19），以中学生为研究对象的占 17.2%（n=5），以小学生为研究对象的占 10.3%（n=3），其他（学龄前儿童和新手志愿者）占 6.9%（n=2）。可见，目前国内外 XR 应用于培养空间能力的研究对象主要是大学生，对中小学生的关注较少。

从样本数量来看（图 5），国内外 XR 应用于培养空间能力的研究中，样本数量可以大致分为 3 类，31.0%（n=9）的研究样本数量为 40 人以下，38.0%（n=11）的研究样本数量为 40-80 人，31.0%（n=9）的研究样本数量为 80 人以上。这一特点可能和 XR 应用于培养空间能力相关研究主要采用准实验法有关。分析表明，当使用受试者内设计（组内设计）时，至少需要 22 名受试者才能达到 80%的把握度（Ho et al., 2022）。在控制成本和确保实验结果权衡下，实验样本控制在 30 人左右或者 70 人左右为最优解，符合已有研究样本量的分布情况。

从研究周期来看（图 6），大部分研究的研究周期都在一年之内（90%，n=26），周期为一天以内的研究最多，占比为 38.0%（n=11），其次为数天到数周但小于 1 个月的短周期，占比为 34.5%（n=10）。可以发现，当前国内外围绕 XR 应用于培养空间能力的研究大多都是数个小时到数天的短周期实验，前测-实验-后测在连续的时间段内完成，避免其他各类因素干扰实验结果。

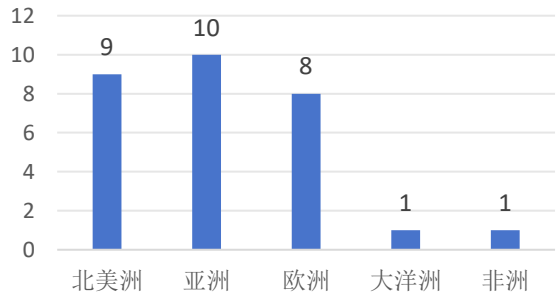


图 3 发表地区分布图

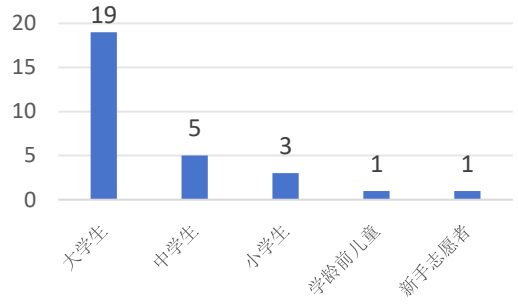


图 4 研究对象分布图

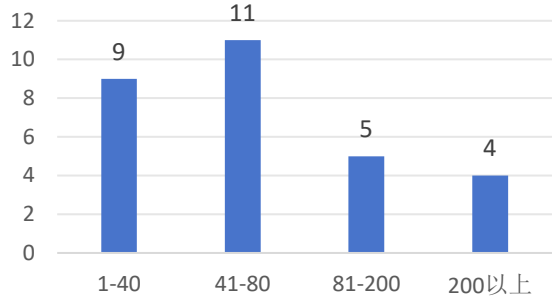


图 5 样本数量分布图

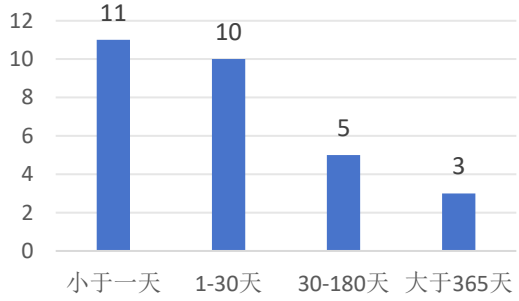


图 6 研究周期分布图

4.XR 培养空间能力实证研究特征分析

4.1. 对空间能力的测量方法不尽相同

空间能力是一种重要的认知能力，在各个学科中都有体现。因为空间能力没有统一的定义，通常被认为是几个影响因素共同组成的，所以关于空间能力的研究依据影响因素划分目前主要有两条研究路线（Roca-González et al., 2017）。一是分析空间感知、空间取向或心理旋转、空间可视化三个因素。第二种是空间关系或心理旋转、空间可视化。在这些研究中，存在许多关于空间能力的标准化测试(如表 2)，可以测量空间能力不同方面的水平。

表 2 空间能力的标准化测试

量表	测量空间能力维度	频次
Mental Rotations Test (MRT)	心理旋转/空间关系	13
Purdue Spatial Visualization Tests: Visualization of Rotations (PSVT: R)	心理旋转	4
Purdue Visualisation of Rotations Test (ROT)	空间能力	3
Spatial Reasoning Instrument (SRI)	心理旋转和空间可视化	2
Perspective Taking/Spatial Orientation Test(PT/SOT)	空间取向	2
Differential Aptitude Test (DAT-5: SR)	空间可视化	2
Picture Rotation Test (PRT)	心理旋转	1
Spatial Perception Scale (SPS)	空间感知能力	1
Spatial ability test from Barrett and Williams (2003)	空间可视化	1

Standard intelligence test (I-S-T 2000 R)	空间可视化	1
General spatial ability test	空间能力	1

心理旋转测试 Mental Rotations Test (MRT)是范登堡和库斯 1978 年设计的测试，是测量空间关系最常用的测试之一（Vandenberg & Kuse, 1978）。MRT 含 20 题，每题左侧为目标图形，右侧为四个（或三个）测试图形，其中两图形为目标图形的不同旋转版本。参与者需从四选项中选出正确旋转的目标图形。

普渡大学空间可视化测试 PSVT:R（Purdue Spatial Visualization Tests: Visualization of Rotations）由 Guay(1976)开发，被广泛应用于测试年龄为 13 岁及以上受教育者的三维心理旋转能力。普渡大学空间可视化测试涉及 30 个问题，包括 13 个对称问题和 17 个非对称问题，从最简单到最难。图 7 展示了 PSVT：R 例题。

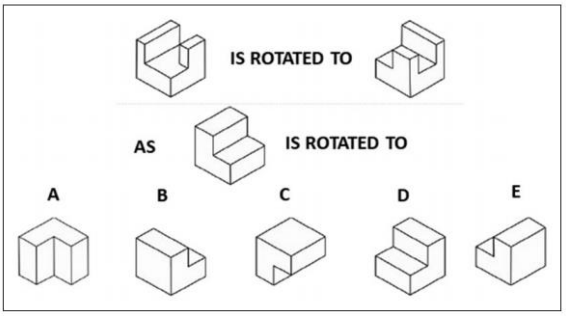


图 7 PSVT：R 例题

普渡旋转可视化测试 ROT(Purdue Visualisation of Rotations Test)是一种在科学教育中广泛使用的空间能力测量方法（Bodner & Guay, 2017）。ROT 要求测试者在 10 分钟内完成 20 个题目。每个问题都会出示一个 3D 物体的旋转范例，要求测试者在不同的物体上进行相同的旋转后，从五个选项中选择正确的结果。

SRI 是一个标准化空间推理测试，包括心理旋转、空间方向和可视化三个子测试，共 30 题，需在 45 分钟内完成。PT/SOT 测试空间取向，要求测试者通过纸笔回答 12 个问题，涉及不同物体间的视角和方向判断。DAT-5: SR 评估空间可视化，需从四个选项中选出与三维物体展开图对应的答案。PRT 测量儿童心理旋转能力，涉及人物和动物的图片识别。SPS 测试六岁儿童的空间感知，包括物体匹配、位置确定和视角解释。空间能力测试则评估空间可视化，包含 75 个可折叠或形成的图案。I-S-T 2000 R 测量智力因素，包括立方体旋转判断。一般空间能力测试则通过纸笔形式，30 分钟内完成 45 个问题，全面评估空间能力。

4.2. XR 用于培养学生的空间能力可行且效果明显

为了更好的展示研究 XR 培养空间能力实证研究的结果，我们根据空间能力在研究中不同的变量类型（自变量、因变量），将所有研究分为 XR 的应用对学生空间能力的影响以及不同空间能力水平对学生在 XR 环境下学习的影响两种类型。13 项研究报告了 XR 的应用对学生空间能力的影响，12 项研究考虑了不同空间能力水平对学生在 XR 环境下的影响。

这些发现表明，在 XR 环境中培养空间能力是可行的，并且可以通过训练提高空间能力（Zhou et al., 2022）。协作沉浸式虚拟环境（VR）中的空间训练可能是一种有效的方法来提高空间技能（Onesa et al., 2023）。运用增强现实技术（AR）能明显提高学习者的空间能力（Elford et al., 2022）。具体到空间能力的不同类型，大部分研究都证实了心理旋转能力在接受训练后会显著提高（Yu & Denham, 2024；Zeynep & Ömer 2020；Roca-González et al., 2017）。此外，也有研究证明可以通过 XR 环境中的学习训练提高空间定位、空间可视化、空间推理能力（Roca-González et al., 2017；Carlos & Luis 2017；）。电子游戏、XR 影像的消费和生产

对空间能力的培养有积极的影响（Baumgartner et al., 2017 ； Tumkor & Serdar 2018）。部分研究认为应用 XR 可以减少认知负荷，从而提高空间能力（Baumgartner et al., 2017）。

4.3. 学生的空间能力水平影响其在 XR 环境中的学习效果

空间能力作为协变量，探究不同空间能力水平的学生在 XR 辅助学习中的表现也是一大研究热点。学生参与 XR 环境中的学习活动或者培训时的表现与空间能力水平直接相关，空间能力高的学生取得的效果更好（Awuor et al., 2022），具有高空间能力的学生在虚拟现实辅助学习中的表现明显优于具有低空间能力的学生（Chen & Chen, 2022）。同时，XR 对空间能力低的学生有积极的补偿影响（Safadel & White, 2020），与空间能力高的被试相比，空间能力低的被试往往从操作和交互的优势中获益最多（Jang et al., 2017），低空间能力学习者的学习成绩受到基于虚拟现实的桌面学习环境的更积极的影响（Lee&Wong, 2014）。也有部分研究证明了不同的空间能力水平对学生的学习效果和满意度没有影响（Cathy et al., 2023）。但普遍认为无论学生的空间能力如何，XR 课程都能提升其学习效果（Laricheva & Ilikchyan, 2023）。

5.研究结论与启示

国内外有关 XR 应用于培养空间能力的实证研究起步较早，并于近五年相关研究的数量快速增加。关于 XR 应用于培养空间能力的研究主要集中在医学、计算机科学的高等教育阶段；主要研究对象为大学生，对中小学生的关注较少；研究方法主要采用准实验法；大部分实验研究周期都在一年之内，其中周期为一天以内的研究最多；从研究量表来看，心理旋转测试 MRT 使用频率最高。XR 对培养空间能力有积极影响，空间能力也影响着 XR 辅助课程的教学效果。

未来研究应进一步拓展 XR 技术在教育中的应用领域，尤其关注其在中小学生空间能力发展中的作用。借鉴医学教育领域的成功实践，我们可以探索将 XR 技术应用于建筑、工程、艺术设计等空间能力需求较高的学科。同时，研究应采用多样化方法，开展长期纵向研究，以全面评估 XR 对空间能力的长期影响。此外，我们还应辩证看待 XR 学习对学生空间能力提出的更高要求，深入研究个体差异、认知风格等因素，以制定更加精准有效的 XR 辅助学习方案。

综上，XR 技术在培养空间能力方面具有巨大潜力，仍有许多值得探索和研究的内容。我们需要持续关注这一领域的进展，不断优化研究方法和 XR 技术的应用，以更好地发挥 XR 技术在教育领域的作用。

参考文献

郭守超. (2022). 数据驱动的空间能力评价研究 (Doctoral dissertation, 华东师范大学).

钟志贤. (2004). 多元智能理论与教育技术. 电化教育研究, (3), 7-11.

Awuor, N. O., Weng, C., Matere, I. M., Chen, J. H., Puspitasari, D., & Nguyen Phuong Tran, K. (2022). The effect of 3D-stereogram mobile AR on engineering drawing course outcomes among first-year vocational high schoolers with different spatial abilities: a Bloom's taxonomy perspective. Interactive Learning Environments, 1-19.

Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. Educational Psychology Review, 30(3), 947-972.

Bodner, G. M., & Guay, R. B. (1997). The Purdue visualization of rotations test. The chemical educator, 2(4), 1-17.

Baumgartner, E., Ferdig, R. E., & Gandolfi, E. (2022). Exploring the impact of extended reality (xr) on spatial reasoning of elementary students. TechTrends, 66(5), 825-836.

Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies (No. 1). Cambridge university press.

Chen, C. C., & Chen, L. Y. (2022). Exploring the effect of spatial ability and learning achievement on learning effect in VR assisted learning environment. Educational Technology & Society, 25(3), 74-90.

Carlos Carbonell Carrera & Luis A. Bermejo Asensio (2017) Landscape interpretation with augmented reality and maps to improve spatial orientation skill, Journal of Geography in Higher Education, 41:1, 119-133

Cathy Weng, Dani Puspitasari, Khanh Nguyen Phuong Tran, Pei Jie Feng, Nicholas O. Awuor & Isaac Manyonge Matere (2023) The effect of using theodolite 3D AR in teaching measurement error on learning outcomes and satisfaction of civil engineering students with different spatial ability, Interactive Learning Environments, 31:5, 2722-2736

Conesa, J., Mula, F. J., Bartlett, K. A., Naya, F., & Contero, M. (2023). The Influence of Immersive and Collaborative Virtual Environments in Improving Spatial Skills. Applied Sciences, 13(14), 8426.

Çöltekin, A., Lochhead, I., Madden, M., Christophe, S., Devaux, A., Pettit, C., ... & Hedley, N. (2020). Extended reality in spatial sciences: A review of research challenges and future directions. ISPRS International Journal of Geo-Information, 9(7), 439.

Di, X., & Zheng, X. (2022). A meta-analysis of the impact of virtual technologies on students' spatial ability. Educational technology research and development, 70(1), 73-98.

Elford, D., Lancaster, S. J., & Jones, G. A. (2022). Exploring the effect of augmented reality on cognitive load, attitude, spatial ability, and stereochemical perception. Journal of Science Education and Technology, 31(3), 322-339.

Gonzales, R.A., Ferns, G., Vorstenbosch, M.A.T.M. and Smith, C.F. (2020), Does spatial awareness training affect anatomy learning in medical students?. Anat. Sci. Educ., 13: 707-720.

Ho S, Liu P, Palombo DJ, Handy TC, Krebs C. 2022. The role of spatial ability in mixed reality learning with the HoloLens. Anat Sci Educ 15: 1074–1085. https://doi.org/10.1002/ase.2146

Jang, Susan; Vitale, Jonathan M.; Jyung, Robert W.; Black, John B. . (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. Computers & Education, 106(), 150–165. doi:10.1016/j.compedu.2016.12.009

Lee, Elinda Ai-Lim; Wong, Kok Wai (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. Computers & Education, 79(), 49–58. doi:10.1016/j.compedu.2014.07.010

Laricheva, E. N., & Ilikchyan, A. (2023). Exploring the Effect of Virtual Reality on Learning in General Chemistry Students with Low Visual-Spatial Skills. Journal of Chemical Education, 100(2), 589-596.

Roca-González, C., MARTIN GUTIERREZ, J., GarcÍa-Dominguez, M., & Mato Carrodeguas, M. D. C. (2017). Virtual Technologies to Develop Visual-Spatial Ability in Engineering Students. Eurasia Journal of Mathematics, Science and Technology Education, 13(2), 441-468.

Safadel, P., & White, D. (2020). Effectiveness of computer-generated virtual reality (VR) in learning and teaching environments with spatial frameworks. Applied Sciences, 10(16), 5438.

Tumkor, Serdar (2018). Personalization of engineering education with the mixed reality mobile applications. Computer Applications in Engineering Education, (), –. doi:10.1002/cae.21942

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: a meta-analysis of training studies. Psychological bulletin, 139(2), 352.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and motor skills, 47(2), 599-604.

Yu, J., & Denham, A. R. (2024). Utilizing augmented reality for embodied mental rotation training: A learning analytics study. British Journal of Educational Technology, 55, 499–518.

Zeynep Gecu-Parmaksiz & Ömer Delialioğlu (2020) The effect of augmented reality activities on improving preschool children's spatial skills, Interactive Learning Environments, 28:7, 876-889,

Zhou, Y., Xu, T., Yang, H., & Li, S. (2022). Improving spatial visualization and mental rotation using forspatial through shapes and letters in virtual environment. IEEE transactions on Learning Technologies, 15(3), 326-337.

沉浸式虚拟现实赋能社会和情感技能学习的系统性文献综述研究

A systematic literature review on immersive virtual reality empowering social and emotional skill learning

王博雯 顾双玲 张慕华*
首都师范大学
6960@cnu.edu.cn

【摘要】 社会和情感能力是 21 世纪学习者的关键技能。沉浸式虚拟现实被证实对于培养学习者的社会与情感技能有积极作用。但是当下对于沉浸式虚拟现实赋能社会与情感技能学习尚未形成体系化的认识。研究采用系统性文献综述的方法，以 2018 年到 2024 年间的 53 篇实证研究为样本，旨在获得沉浸式虚拟现实技术在培养社会与情感技能方面的全貌。研究发现，沉浸式虚拟现实技术在培养学习者社会与情感技能的研究集中在教育学和医学领域，其中自我效能感、成就动机、自信、移情的应用最为丰富且都具有促进作用，在测量方式上，采用自我报告法应用最为广泛。

【关键词】 沉浸式虚拟现实；社会和情感技能；测量方法；实证研究

Abstract: Social and emotional abilities are key skills for 21st century learners. Immersive virtual reality has been proven to have a positive effect on cultivating learners' social and emotional skills. However, there is currently no systematic understanding of how immersive virtual reality empowers social and emotional skill learning. The study adopts a systematic literature review method, using 53 empirical studies from 2018 to 2024 as samples, aiming to obtain a comprehensive view of the immersive virtual reality technology in cultivating social and emotional skills. Research has found that research on immersive virtual reality technology in cultivating learners' social and emotional skills is concentrated in the fields of education and medicine. Among them, the applications of self-efficacy, achievement motivation, confidence, and empathy are the most abundant and have promoting effects. In terms of measurement methods, self-report is the most widely used.

Keywords: immersive Virtual Reality; Social and emotional skills; Measurement method; Empirical research

一、 研究背景

1. 社会与情感技能是新时代学习者必备的关键能力

社会情感技能作为与外部世界交往的综合能力，已成为国际教育领域关注的热点问题，世界各地的教育系统越来越认识到，社交和情感技能对学生和社会都至关重要。社会情感技能作为心理学和教育学的重要研究领域，不仅对于个人的学业成就、幸福感等方面有重大影响，而且在提高受教育程度、就业能力、工作表现以及社会参与等方面也发挥着重要作用（郭忠辉等，2021）。已有研究表明，社会与情感技能与认知能力之间存在内在关联性（Cunha et al., 2010），社会与情感技能有助于提高认知能力，二者的相互作用有助于学生在多方面取得成功，因此社会与情感技能是新时代学习者必备的关键能力。

经合组织将社会情感技能定义为个体通过正式或非正式学习习得，表现为认知、情感、行为模式相一致并对个体产生重要影响的能力（OCED, 2020）。经合组织以“大五人格模型”作为社会与情感技能的调查框架，将其划分为六大领域：任务执行表现、情绪调节、与他人

互动、协作、开放性以及社会与情感调查项目的其他技能。按照框架与标准，在六大领域下确定了 19 项二级技能，包括成就动机、责任、坚持等（Stephanavičius et al., 2023）。

2. 沉浸式虚拟现实（Immersive Virtual Reality，以下简称 IVR）的应用现状

在数字和智能技术时代，随着虚拟现实技术的发展和软硬件设备的突破，具备交互性的 IVR 技术被应用在多个领域。美国科技创新智库信息技术与创新基金会于 2021 年发布报告《沉浸式学习的前景：增强现实和虚拟现实在教育中的潜力》，介绍了 IVR 在 K-12 教育、高等教育和教师培训的应用场景，认为沉浸式虚拟现实技术不受时空限制、成本低的特点有助于其在教育领域的发展（武建鑫等，2023）。同时有研究表明，在国外 IVR 已经被应用于工程、化学等 18 个学科领域的教学中。在国内 IVR 同样被应用于空间认知与空间推理技能学习、写作学习、科学概念学习与理工科实验教学等多个学习情境之中（张慕华等，2022）。

3. 沉浸式虚拟现实技术赋能培养学习者的社会和情感技能方面的研究

IVR 能够使用户发挥自身在虚拟世界中的主动性，通过综合调动视觉、触觉、听觉多种感觉器官产生思维共鸣。因此，IVR 可以被用于社会与情感技能领域之中，对于培养学习者的社会与情感技能方面存在积极作用，具有较深的实践意义。但是当下学者对于 IVR 赋能社会与情感技能的关注与实证研究较少，尚未形成体系化的认识。本文以 IVR 技术在社会与情感技能方面的实证研究为鉴，厘清 IVR 技术在社会与情感技能领域的应用思路，探究其在实际应用中发挥的重要作用，从而为我国学者开展 IVR 与社会与情感技能研究提供实践参考。

二、 研究设计

1. 研究方法

本研究采用系统性文献综述法开展研究。系统性文献综述法的优势在于严谨、透明，包括清晰的研究问题、全面的检索策略、明确的文献标准、高质量的评估方法、综合的数据分析以及可靠的研究结果，能够有效克服传统研究方法的主观性、偏见性等问题（Sutherland, 2004）。

2. 研究问题

为了解国内外当前以沉浸式虚拟现实技术为依托来培养学习者的社会和情感技能方面的研究情况，本研究确定了如下研究问题：（1）国内外沉浸式虚拟现实技术用于发展学习者的社会和情感技能方面的研究主要呈现哪些特征（研究地区、研究时间、研究对象、研究方法等）？（2）沉浸式虚拟现实技术用于培养学习者的社会和情感技能的测量方法有哪些？（3）沉浸式虚拟现实技术对于提升学习者的社会和情感技能方面的策略有哪些？

3. 样本获取

（1）文献检索策略

为有效获取国内外近几年沉浸式虚拟现实技术赋能社会与情感技能的实证研究文献，本研究在 Web of Science 文献数据库中进行检索，在检索前，笔者对经合组织确定的 19 项社会与情感技能的二级技能进行翻译，并寻找英文同义词替换以保证文献检索的完整性，而后采用布尔运算方式在 Web of Science 进行逐一检索，例如“immersive Virtual Reality/IVR and (emotion control or emotional regulation)”，数据库范围限定为 Web of Science 核心合集，研究领域限定为教育研究，共获得 523 篇文献。

（2）文献筛选标准

为保证文献分析结果的准确性和可靠性，精确呈现国内外沉浸式虚拟现实技术用于发展学习者的社会和情感技能方面的研究情况，基于研究问题，本研究针对初步检索到的 523 篇文献制定了如表 1 所示的文献纳入/排除标准（Indriasari et al., 2020）。其中 1-6 条为系统性文献综述法为了保证样本的准确性与权威性而普遍采用的筛选标准；第 7 条是为了筛选实验类实证研究，剔除部分采用大规模问卷调查，但是缺乏研究问题、严谨的实验过程以及清晰的

研究方法的文献；第 8 条旨在将文献的研究主题聚焦于沉浸式虚拟现实与社会和情感技能的研究，研究问题和测量方法要与社会和情感技能相关，剔除未测量相关社会和情感技能或非研究重点的研究文献。

表 1 文献纳入/排除标准

纳入标准	排除标准
1. 英文论文	非英文论文
2. 实证研究	非实证研究
3. 全文可获取	全文不可获取
4. 文章是期刊论文	书稿、会议论文、报告等
5. 文章至少包含三页	少于三页的海报、段论文或简介等
6. 文章标题只出现一次	重复性标题
7. 研究包括明确的研究问题、方法及结论	研究没有呈现明确的研究问题、方法、结论
8. 研究主题聚焦 IVR 与社会情感技能	研究主题非 IVR 与社会情感技能

（3）文献筛选过程

本研究遵循系统性文献综述及元分析方法（Preferred Reporting Items for Systematic Reviews and Meta-Analyses，PRISMA）的研究思路进行文献综述。该方法是国际上常用的系统性综述方法，包含 27 项指标（如标题、摘要、方法、结果、讨论等）和 4 个阶段(Liberati et al.,2009)。需清楚地呈现文献识别、筛选、纳入或排除的过程及原因，以提升系统性文献综述和元分析报告的准确性。基于该研究思路，本研究最终获得符合条件的论文 53 篇。PRISMA 流程图如图 1 所示。针对筛选后的 53 篇文献，研究从作者、年份、国家、期刊、研究主题、研究对象、研究方法、测量方法、测量工具、研究结论等维度对其进行编码分析。

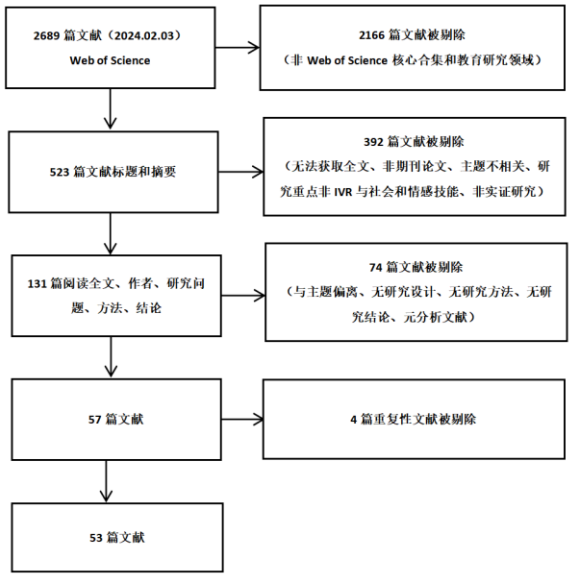


图 1 PRISMA 流程图

研究主题非 IVR 与社会和情感技能的文章将被剔除。只在结论简短地提到社会情感技能但未在文中写清研究设计和研究方法的文章将被剔除。由于采用元分析方法的论文无法计算样本数量，元分析研究也在多次筛选后被剔除。

三、 基于文献计量分析的社会和情感技能实证研究现状

实证研究是由“真实问题—方法运用—数据分析—结论诠释”构成的统一体(刘选， 2017)，是教育走向科学的必要途径(袁振国， 2017)。本研究聚焦国内外以沉浸式虚拟现实技术为依托对学习者社会和情感技能的实证研究，重点关注社会和情感技能实证研究的设计、方法、结论等。以下将按照文献发表、作者分布、研究设计以及研究方法等来具体分析国内外沉浸式虚拟现实技术培养学习者的社会和情感技能实证研究的基本特点。

1. 文献发表与引用情况

从发表时间来看，目前国内外社会与情感技能的实证研究整体上呈现逐年递增的趋势，表明社会与情感技能持续且正在以上升的趋势受到国内外研究者的关注。从发表的期刊分布来看，53 篇文章发表在 27 种不同的期刊上。发表数量排前三的分别是《英国教育技术杂志》

（BRITISH JOURNAL OF EDUCATIONAL TECHNOLOGY）、《计算机辅助学习杂志》（JOURNAL OF COMPUTER ASSISTED LEARNING）、《交互式学习环境》（INTERACTIVE LEARNING ENVIRONMENTS）、《计算机与教育》（COMPUTERS & EDUCATION）。

从文献引用率来看，有 17 篇文献的引用率在 20 次以上，超过 50 次的文章有 9 篇。其中，引用率最高的文章是 Makransky 等（2018）研究的使用沉浸式或桌面 VR 管理虚拟科学学习模拟方面是否存在差异，其中测量了动机的变量（272 次）；其次是 Makransky 等（2019）研究的有关 IVR 作为提供实验室安全培训媒介的有效性，在对学习结果进行考察时，测量了内在动机和自我效能感的前后变化（190 次）可以发现引用率较高的主题主要集中在 IVR 技术对学习者成就动机和自我效能感的影响，这也是国内外社会和情感技能的研究热点。

2.作者分布特征

为展现国内外社会和情感技能实证研究的作者分布情况，研究以第一作者为分析单位，对 53 篇样本文献进行分析，发现作者总共来自 18 个国家，其区域分布分别是亚洲（n=21，39.6%）、欧洲（n=20，37.7%）、北美洲（n=8，15.1%）、大洋洲（n=3，5.7%）以及非洲（n=1，1.9%）。从国内外作者发文数量来看，中国作者共 15 篇，占总体篇数的 28.3%；国外作者共 38 篇，占总体篇数的 71.7%，说明国外学者是社会与情感技能实证研究的主要群体。

3.研究设计特征

为进一步分析 IVR 技术用于培养学习者的社会和情感技能的研究情景与实验设计，本研究对样本文献的研究对象和样本数量指标进行了系统分析。

从研究对象的类别来看，国内外近几年 IVR 技术赋能社会和情感技能实证研究中，以大学生为研究对象的占 64.2%（n=34），以初高中为研究对象的占 18.9%（n=10），以小学生为研究对象的占 11.3%（n=6），其他（学前儿童）占 5.7%（n=3）。可见，国内外 IVR 技术与社会和情感技能实证研究的对象主要是大学生和初高中学生，对小学生及学前儿童的关注相对较少。这可能是由于 IVR 技术对小学生和学前儿童来说操作和实施起来比较困难，并且对面向其开展实验研究时，需要征得教师和父母的同意，这为实证研究的开展增大了难度。

从样本数量来看，国内外近几年 IVR 技术与社会和情感技能实证研究中，35.4%（n=19）样本研究数量为 100 人以上，24.5%（n=13）的研究样本数量为 40 人以下，其他研究的样本数量分布相对平均。这说明国内外 IVR 技术与社会和情感技能实证研究的研究样本量都是相对较少，可能是因为技术应用实施相对复杂。

4.研究方法特征

为更好地借鉴国内外社会和情感技能的实证研究范式，本研究将实证研究方法划分为实验设计、准实验设计及非实验设计三类。实验设计是指控制实验条件和安排实验程序的研究，其目的在于分析实验条件和实验结果之间的关系，一般采用随机分组；准实验设计是指既不能直接操纵自变量又不能对研究中的额外变量进行严格控制的研究，不采用随机分组，而是进行多组别的分类和测量，如设置对照组和控制组等。其他属于非实验设计。对 53 篇实证研究文献的研究范式和数据分析方法进行分析后发现，有 18 篇文章分别采用了实验设计和准实验设计（34%），17 篇文章属于非实验设计（32.1%），采用研究方法分布较均匀。

四、沉浸式虚拟现实技术培养学习者社会和情感技能研究特征分析

1.社会和情感技能的测量方法

（1）数据收集方式

对 53 篇文献数据收集方式进行系统分析（见表 2）后发现，问卷（45.80%）是 IVR 和社会与情感技能研究中数据收集的主要方式。该方法操作简单，能够在短时间内获取研究对象的背景资料。其次是测试（24.30%）。研究者结合教育教学主题设计测试题目，依据研究对

象的测试成绩判断学习者的知识习得情况。第三种方式是访谈（11.22%）。研究人员通过对访谈内容进行文本分析来衡量研究对象的社会与情感技能，如 Lee（2022）对研究对象进行焦点小组访谈，将录音音频转换成文本逐字分析，针对访谈数据进行编码处理(Lee et al.,2022)。第四种方式是录像/录音（6.54%）。此方式主要通过录像或者录音的方式记录学习者在学习过程中的学习表现，进而对视频或音频文字进行编码分析。此外，可穿戴设备（4.67%）、日志（4.67%）和观察（2.80%）等方法也常被用来收集数据，从而保证数据收集方式的多样性。值得强调的是，统计分析发现有 69.81%的研究采用了两种或两种以上的数据收集方式，说明已经有越来越多的研究试图通过融合多种数据源的方式来消除单一数据来源的局限性，保证数据分析结果的可靠、准确。

表 2 社会与情感技能的数据收集方式

数据收集方式	频率	百分比
问卷	49	45.80%
测试	26	24.30%
访谈	12	11.22%
录像/录音	7	6.54%
可穿戴设备	5	4.67%
日志	5	4.67%
观察	3	2.80%

（2）测量方式

社会与情感技能可以使用自我报告量表、教师评分、文本分析、观察评价、生理反应、视频编码等方式进行测量。通过系统分析 53 篇文献发现（见表 3）：首先，自我报告法（50.00%）的应用最为广泛，其中比较经典的测量量表包括：①人际反应指数量表（Interpersonal Reactivity Index，IRI），包括透视、幻想、移情关注、个人痛苦四个分量表，由戴维斯提出用来衡量研究对象的倾向移情水平；②考夫曼创造力领域量表（Kaufman Domains of Creativity Scale ，K-DOCS），包括日常，学术，性能，机械和科学，艺术 5 个维度，被证明是评估创造力的有效量表；③沉浸感量表（Igroup Presence Questionnaire，IPQ），包括身临其境感、空间感知、参与度、真实感四个维度，被认为衡量参与者在 IVR 中自我报告的存在感的适当工具。其次是教师评分（21.88%）。研究者对研究对象的表现及作品进行评分，依据最终评分结果分析 IVR 对学习者社会与情感技能的影响。再次是文本分析（11.45%），研究者以编码的形式分析访谈数据和学生的写作内容，探究学习者使用 IVR 后社会与情感技能产生的变化。第四种是观察评价（7.29%），研究者观察学习者在实验过程中的学习表现来评价其社会情感技能。随着可穿戴设备在教育领域的逐渐普及，生理反应（6.25%）作为非传统测量方式也逐渐得到应用，包括心率、脑电波等。Xiaozhe Yang（2018）在实验中通过可穿戴设备记录参与者在设计作品过程中的脑电波数据，对脑电波数据进行编码，从而分析参与者在活动过程中的注意力变化（Yang et al. , 2018）。此外，也有部分研究选用视频编码（3.13%）的方式，研究者依据一定标准对录像视频进行编码分析。整体来看，当下对于社会与情感技能的测量方式应用较为单一，相关测量方法的进步仅局限于对传统方法的组合或迭代，亟需一种本质上完全不同的测评方法。

表 3 社会与情感技能的测量方法

测量方法	频率	百分比
自我报告法	48	50.00%
教师评分	21	21.88%

测量方法	频率	百分比
文本分析	11	11.45%
观察评价	7	7.29%
生理反应	6	6.25%
视频编码	3	3.13%

2.沉浸式虚拟现实技术对于培养学习者社会与情感技能具有显著影响

对53篇文献数据收集方式进行系统分析（见图2）后发现，IVR已应用在社会与情感技能领域的多个二级技能中。其中，IVR在培养学习者自我效能感、成就动机、自信、移情四个方面的应用最为丰富，在好奇心、创造力、元认知等方面也有应用，但研究案例较少。

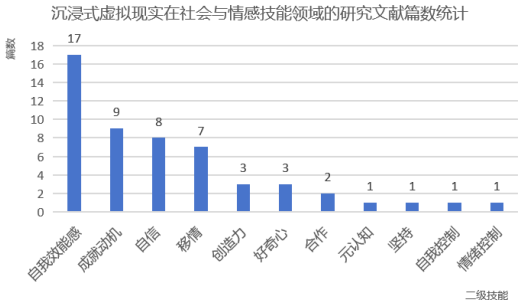


图2 沉浸式虚拟现实在社会与情感技能领域的研究文献篇数统计

（1）沉浸式虚拟现实对于提升社会与情感技能中自我效能感具有显著效果

通过分析IVR与自我效能感相关的18篇文献，发现其中17篇文献通过研究数据表明IVR的应用对于提升学习者自我效能感具有显著影响，占比高达94.44%。此外，表明未发现具有显著影响的研究中，研究者认为学习者在研究过程中始终保持较高的自我效能感和内在动机，学习者认为自身在模拟过程中受益，证实了IVR在激发学生兴趣方面具有明显效果。

（2）沉浸式虚拟现实在社会与情感技能中关注度有待提高的二级技能

研究者对IVR与成就动机、自信、移情这三个二级技能的关注度较高。在IVR与成就动机相关的9篇文献中，5篇文献通过数据表明IVR对于提升学习者的成就动机具有统计学上的显著差异，4篇文献表明在研究中实验组和对照组学生的学习成绩和学习动机差异无统计学意义以及IVR对于提升学习者成就动机发挥的影响很小。IVR与自信相关的8篇文献均表明IVR能够有效提升参与者自信。值得注意的是，其中7篇文献均与医学生教育相关，研究中使用IVR技术对医学生进行培训，参与者在知识、信心和同理心的大多数衡量标准上都有了实质性的提高，由此可见IVR技术在医疗和教育领域的应用价值之高。

此外，通过阅读文献发现IVR在移情、好奇心、创造力、合作、元认知、坚持、自我控制和情绪控制方面也已开展相关研究，研究数据均认可了IVR在应用中发挥的正向作用，表明学习者得到了积极正向的学习与情感体验。

（3）沉浸式虚拟现实尚未关注的社会与情感技能

通过查阅资料发现，目前沉浸式虚拟现实在宽容、信任、批判性思维、乐观、抗压能力、社交、活力等二级技能领域中的应用较少，尚未有研究证实沉浸式虚拟现实对于提升以上二级技能具有积极作用。宽容、信任、社交、抗压能力等二级技能是学习者与在社会生存与表达情感的必备技能，因此，在以上领域开展应用IVR的研究具有重要的实践意义，有助于完善IVR在培养学习者社会与情感技能领域的研究体系。

综上所述，纳入分析的53篇文献中大部分研究者认为IVR在提升学习者社会与情感技能方面具有显著影响，IVR的应用领域较为广泛，具有较高的实用价值。

3.应用沉浸式虚拟现实技术提升学习者的社会和情感技能策略

基于以上分析，本研究认为有必要在教育领域应用IVR技术，从而发挥对培养学习者社会与情感技能的促进作用，结合现有研究总结如下相关策略。

充分挖掘IVR技术特有的沉浸感、临场感和交互性特征在培养学习者社会与情感技能方面的效果。多个实证研究已经表明了IVR对于培养学习者社会与情感技能具有显著效果，IVR能够提供更强的沉浸感、存在感和真实感，被认为是一种可行的和有价值的方法，在教育活动中使用IVR技术可以不同程度地培养学生的社会与情感技能和其它知识与技能。因此，教育者应当提升自身应用新兴技术的能力，结合教育主题有针对性地选取IVR技术作为辅助教学工具，从而提升教学效果和学习者的学习体验。

重视IVR环境中的反馈设计，注重IVR学习与传统学习活动的结合。教师应该在教学内容的高沉浸要求和特定的教育目标之间保持平衡，从促进教学效果提升的角度使用IVR，围绕教学内容因地制宜技术的使用。同时，教育者可以使用适当的反馈设计，及时了解学习者在学习过程中的感受与体验，结合反馈内容调整教学安排，以提高学习者与学习材料互动的程度。此外，教育者不应将IVR课程视为所有学习活动的替代品，而是选择正确的方法，使其在特定的教育环境中提高学习效果。

重视学习者个体差异，采用分组教学等教学策略提升IVR教学效果。有研究表明学习者学习能力水平的差异会影响使用IVR教学的效果和学习者的学习体验感，因此，教育工作者可以考虑在实施IVR学习活动时，将具有不同水平的学生进行异质分组，从而最大化IVR的教学效果。

五、 研究结论与启示

1. 研究结论

本研究利用系统性文献综述法，筛选国内外近几年社会和情感技能领域的53篇英文实证文献，从研究主题、数据收集、测量方式以及研究关注点对社会和情感技能实证研究现状进行系统分析，研究发现：国外近几年社会和情感技能的实证研究数量整体上呈逐年上升趋势，且主要发表在教育学和医学类期刊上。从作者地理分布来看，社会情感技能的研究仍以国外学者为主；从研究设计来看，国内外学习投入的实证研究主要聚焦大学生，其次是初高中学生和小学生；研究方法以定量研究和混合研究为主。从数据收集来看，以问卷、测试和访谈为主。从测量方式来看，自我报告法在社会与情感技能的应用最为广泛。从研究关注点来看，沉浸式虚拟现实技术在培养学习者自我效能感、成就动机、自信、移情四个方面的应用最为丰富，在好奇心、创造力、元认知等方面也有应用，但研究案例较少。

2. 研究启示

本研究在分析国内外社会和情感技能实证研究的基础上，总结出其对我国学者开展社会和情感技能实证研究的两大启示：

（1）开发更加有效和全面的评估方式

IVR 赋能社会与情感技能作为近年来新兴的技术与概念，具有明显的时代特征，因此，评估方式的创新与发展显得尤为重要。研究者需引入更多的评估指标和方法，综合考虑多方面影响因素，利用现代科技手段对评估数据进行深度挖掘和分析，从而得出更加科学的评估结果。同时，评估方式的开发应注重过程性和动态性。本研究建议相关研究可以通过对学习者进步和变化的跟踪观察验证基于IVR的社会与情感技能培养是否会对学习者发挥长远作用以及发挥作用的时长，以此提升IVR在社会与情感技能领域研究的完整性。此外，评估方式应该在实践中不断探索和创新，从而为被评估对象提供更加全面、准确、科学的评估服务。

（2）拓宽IVR在社会与情感技能领域的应用范围

IVR 技术在社会与情感技能领域的使用仍在蓬勃发展，在很大程度上尚未探索，因此需

要更多的实证研究来更好地揭示 IVR 技术如何更好地在社会与情感技能领域发挥价值。技术开发者应深入了解用户的需求，通过市场调研、用户访谈等方式，收集用户在使用 IVR 过程中的反馈和建议，以便不断优化系统功能和用户体验。同时，研发者应不断投入研发力量来提升技术水平，开发更多的 IVR 场景，加强与多领域的跨界合作，推动 IVR 技术的创新发展。

六、 参考文献

郭忠辉, 王佳 & 刘淑杰. (2021). 经合组织社会情感技能测评框架及其对我国的启示. 教育测量与评价 (02), 21-26+35. doi:10.16518/j.cnki.emae.2021.02.004.

武建鑫 & 王嘉琳.(2023).虚拟现实/增强现实技术将把沉浸式学习带向何方——《沉浸式学习的前景：增强现实和虚拟现实在教育中的潜力》解析. 内蒙古师范大学学报(教育科学版)(03),95-104.

张慕华,祁彬斌,黄志南,朱永海 & 李云文.(2022).沉浸式虚拟现实赋能学习的内在机理——沉浸感和情感体验对学习效果的多重影响. 现代远程教育研究(06),92-101.

刘选(2017).实证研究怎么做:让研究者困惑的地方——来自华东师大第二届全国教育实证研究论坛的启示[J].现代远程教育研究,(3):18-25.

袁振国(2017）.实证研究是教育学走向科学的必要途径[J].华东师范大学学报(教育科学版),35(3):4-17,168.

Cunha, F. , Heckman, J. J. , & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. Social Science Electronic Publishing[J].Econometrica,78(3):883-931.

OECD. (2020). Personality Matters: Relevance and Assessment of Personality Characteristics [EB/OL]. http://dx.doi.org/10.1787/8a294376-en

Sutherland, S. (2004). An Introduction to Systematic Reviews[J]. Journal of Evidence Based Dental Practice, 4 (1): 47-51.

Indriasari, T., Luxton-Reilly, A., & Denny, P. (2020). Gamification of Student Peer Review in Education: A Systematic Literature Review[J]. Education and Information Technologies, (3):1-30.

Steponavičius, M., C. Gress-Wright and A. Linzarini (2023), "Social and emotional skills: Latest evidence on teachability and impact on life outcomes", OECD Education Working Papers, No. 304, OECD Publishing, Paris, https://doi.org/10.1787/ba34f086-en.

Liberati, A., Altman, D., & Tetzlaff, J. et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Healthcare Interventions: Explanation and Elaboration[J]. Journal of Clinical Epidemiology, (10):1-34.

Makransky, Guido; Lilleholt, Lau (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educational Technology Research and Development, (), –. doi:10.1007/s11423-018-9581-2

Makransky, Guido; Borre-Gude, Stefan; Mayer, Richard E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. Journal of Computer Assisted Learning, (), jcal.12375–. doi:10.1111/jcal.12375

Lee,T.,Wen, Y.,Chan, M. Y.,Azam, A. B.,Looi, C.K.,Taib, S.,Ooi, C.H.,Huang, L. H.Xie, Y.,&Cai, Y.(2022).Investigation of virtual & augmented reality classroom learning environments in university STEM education [J].Interactive Learning Environments,https://doi.org/10.1080/10494820.2022.2155838-en.

Yang, XZ.,Lin, L.,Cheng, PY.,Yang, X.,Ren, YQ.,&Huang, YM.(2018).Examining creativity through a virtual reality support system[J].ETR&D-Educational Technology Research and Development,66(5):1231-1254.

Lin, YJ.,&Wang, HC.(2018).Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom[J].Education and Information Technologies,26(4):4487-4505.

Cheng, K.,&Tsai, CC.(2020).Students' motivational beliefs and strategies, perceived immersion and attitudes towards science learning with immersive virtual reality: A partial least squares analysis[J].British Journal of Educational Technology,51(6):2139-2158.

基于信息化平台的全过程评价——以《明朝的对外关系》一课为例

End-to-end Evaluation Based on the Information Platform - Taking the Lesson "Ming Dynasty's Foreign Relations" as an Example

王誉颖[1]，吴欣熙[1]，郭辉群[1]，李静茹[1]，崔萌[1*]
[1]广东第二师范学院教师教育学院
[*] mengcui@m.scnu.edu.cn

【摘要】 以义务教育教科书（人教版）七年级下册《历史》“明朝的对外关系”一课的教学设计为例，在教学设计中充分运用信息化教学平台辅助教学分析与教学评价，落实教学重难点，掌握学生学情。
【关键词】 信息化平台；教学设计；教学分析工具；教学评价

Abstract: *Taking the teaching design of the seventh grade textbook "History" lesson on "Ming Dynasty's Foreign Relations" as an example, fully utilizing the information technology platform in the teaching design to assist in teaching analysis and evaluation, effectively implementing teaching priorities, and understanding student learning situations.*

Keywords: Information technology platform, Teaching design, Teaching analysis, Teaching evaluation

随着信息技术的迅猛发展和普及，信息化平台在教育教学中的应用越来越广泛。这种新型的教学模式不仅丰富了教学手段，提高了教学效率，更为学生的全面发展和深度学习提供了有力支持。全过程评价作为一种注重学生学习过程和学习成效的评价方式，与信息化平台的结合，可以更加科学精准地反映学生的学习状态，对教学成效意义重大，促进教学质量的提升，具有双向促进作用（夏云青，2021）。本文以《明朝的对外关系》一课为例，探讨如何利用信息化平台进行全过程评价的教学设计。深入分析信息化平台如何应用于课前准备、课中互动、课后反馈等各个环节中，对学生学习过程和学习成效实现全过程评价，为教学设计提供有益的参考和启示。

1. 教材解析

本课例选自义务教育教科书（人教版）七年级《历史》第三单元第 15 课——明朝的对外关系。《明朝的对外关系》这一课在教材体系中扮演着承上启下的重要角色。它既是对之前学习内容的延伸，又为后续的历史学习奠定了基础。

在本课之前，学习内容涉及明朝的建立、政治制度、科举制度等多个方面，学生已经对明朝的基本概况有了一定的了解。在此基础上，《明朝的对外关系》这一课进一步拓宽了学生的学习视野，将关注点从国内转向国外，探讨明朝与周边国家及地区的关系与互动。这种转变不仅有助于学生更全面地了解明朝的历史面貌，还能够加深他们对古代中国与世界联系的认识。其中本节课的重点分别是“郑和下西洋”以及“戚继光抗倭”。

本课的学习内容为后续的历史学习提供了重要的线索和背景。明朝的对外关系不仅影响了当时的国际格局和地区稳定，还对后世的对外政策和文化交流产生了深远的影响。此外，还暗示了明朝由盛转衰的基调，为后面明朝的灭亡埋下了伏笔。

因此，可以说《明朝的对外关系》这一课在教材内容上是承上启下的一节内容，它有助于学生更好地理解历史发展的连续性和变迁性，构建一个更加完整、连贯的历史知识体系，提高他们的历史素养和综合能力。

2. 教学过程设计

2.1 课前调查

在课堂开始前一天，教师提前下发有关本节课中关于“明朝对外关系”的三件大事的预习检测，学生完成后提交至小雅平台，教师在课前进行检查并总结学生作答数据，分析学生理解难点。

2.2 课堂导入

通过展示“马来西亚三宝佛洞”、“马来西亚三宝井”、“泰国三宝公寺”、“柬埔寨三保公庙”等海外纪念场所，引出本节课第一个重点事件的主人公——郑和，亦称“马三宝”。

2.3 课中学习

2.3.1 第一部分内容：海之昌——航海壮举

在第一部分，教师通过导入部分的海外纪念场所引入郑和的详细生平介绍及其“下西洋”这一壮举。接着教师放映“世界地图局部图”，要求学生结合课本内容以及地理知识，在图上找出西洋的地理范围以及郑和下西洋的路线，学生自主思考后在点阵笔授课系统上传自己所绘路线图，教师通过点阵笔授课系统分析学生作答情况并且抽取学生案例进行讲解，讲述郑和下西洋的正确地理位置及路线。

然后学生通读教案材料，分析郑和下西洋的目的，学生进行抢答，教师结合学生答案进行补充讲解。而后教师展示 15-16 世纪中西方主要航海活动（郑和下西洋与哥伦布航海、达·伽马远航印度等）比较表，请同学们通过表格分析郑和下西洋的地位，学生思考后教师进行解答。

最后进入小组合作探究，各小组结合课本及材料，分析以下两个问题：①郑和下西洋成功的原因和条件有哪些?②郑和下西洋的影响有哪些；以小组为单位进行回答，教师进行归纳总结并补充。

2.3.2 第二部分内容：海之战——倭寇来犯

在第二部分，教师以《凯歌》这首军歌为导入，引入其作者“戚继光”及“倭寇”这一概念。接着教师分析倭寇的形成原因以及倭患猖獗带来的危害，从而引入戚继光组建“戚家军”抗倭并最终基本消除东南沿海“倭患”这一壮举。然后小组进行讨论，探究戚继光抗倭能够取得胜利的原因，以小组为单位进行回答，教师最后进行归纳总结并补充。

2.3.3 第三部分内容：海之殇——殖民侵略

在第三部分，以前面讲到的郑和下西洋后，西方开始了地理大发现的时代，从 16 世纪开始向外进行海外殖民为过渡，引出“葡萄牙人攫取澳门的居住权”这一部分内容。教师讲解16世纪开始欧洲殖民者对我国沿海地区进行侵略活动以及葡萄牙殖民者攫取了在我国广东澳门的居住权的相关史事。

其中在课堂各互动环节，教师运用班级优化大师记录学生课堂表现，分析学生课堂情况，进行过程性评价。

2.4 归纳总结

教师对本节课三大部分内容进行板书总结，在明朝前期，国力强盛下，中国有了和平的对外交往——郑和下西洋；明朝中期，国力衰弱，有了外来的倭寇和葡萄牙的入侵，说明了国力的强弱决定着对外关系。最后教师发放学生自我评价量表，学生在点阵笔授课系统对本节课堂的学习过程情况进行自我评价。

2.5 课后巩固

教师布置课后巩固习题，要求学生课后自主完成习题并上传至小雅平台上，教师结合平台智能分析学生作答情况，了解学生课堂知识掌握情况，完成结果性评价。

3. 教学分析与全过程评价的实施

3.1 诊断性评价

诊断性评价是对学生的知识、技能以及情感等多个方面的情况所进行预测的活动。通过应用诊断性评价，教师可以了解学生的知识掌握情况以及运用情况，同时教师可以根据评价结果优化的教学计划和教学设计，为学生个性化制订学习计划和学习方法。

在本课《明朝的对外关系》中，教师在课前发布本节课教案，学生预习后完成教案并上传至小雅平台，教师通过平台数据呈现，可以分析学生理解难点，从而调整课堂教学方案，对学生理解难点着重讲解。

3.2 过程性评价

过程性评价是一种在课程实施的过程中对学生的学习进行评价的方式，对学生的学习动机效果、过程以及与学习密切相关的非智力因素进行全面的评价，主张评价过程与教学过程的交叉和融合，促进学生对学习的过程进行积极的反思，从而更好地把握学习方式方法。理解和掌握评价的方法，作为与终身学习相呼应的一个方面，能够促进终身可持续发展的实现（高凌飚，2004）。

在本课案例中，教师在课堂组织上，使用了班级优化大师对学生的课堂积极程度、表现情况小组合作参与度等方面进行评价；在课堂学习上，学生使用点阵笔填写课堂思考习题的答案以及自我评价量表，即时同步上传到点阵笔授课系统，系统以柱形图、雷达图可视化、实时、直观反馈学生的课堂思考、学习情况、学习反思，实现了学习过程的评价。



图 1 学生课堂表现评价

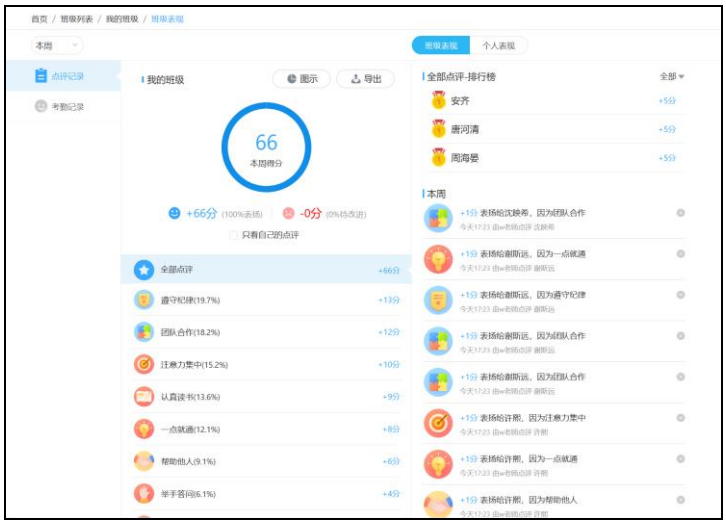


图 2 学生课堂过程性评价统计

3.3 结果性评价

结果性评价是指在一个学习阶段或一门课结束后对学生学习效果进行的评价，是判断学生达到教学目标的评价，是检验教学成果最重要的教育评价方式（周建芳等人，2023）。

本节课最后，教师布置课后巩固习题，检测学生课堂内容掌握情况，学生课后完成习题后上传至小雅平台上，教师结合平台智能工具分析学生作答情况，判断学生知识掌握情况，实现结果性评价。

4. 结语

本论文以《明朝的对外关系》一课为例，深入呈现了基于信息化平台的全过程评价与教学设计的融入。通过充分利用信息化平台的优势，在课前、课中、课后各个环节中实现对学生学习全过程的评价。这一评价方式不仅关注学生的学习成果，更重视学生在学习过程中的表现与成长。

在教学设计过程中，我们贯彻了全过程评价的理念，通过设定明确的教学重点、设计多样化的教学活动、运用信息化工具进行实时反馈与评价，确保学生能够在学习过程中获得充分的指导与支持。同时，我们也注重使用信息化平台激发学生的主动性和创造性，鼓励他们积极参与课堂讨论、小组合作等活动，以实现知识与能力的共同提升。

展望未来，随着信息技术的不断发展和教育教学改革的深入推进，基于信息化平台的全过程评价将在教学设计中发挥更加重要的作用。我们将继续探索和完善这一评价方式，为培养具有创新精神和实践能力的新时代人才贡献力量。

参考文献

周建芳、周振和张楚俊（2023）。数据赋能课程结果性评价的应用研究[J]。高教学刊，2023,9(36):95-98。

夏云青（2021）。基于雨课堂的过程性评价与终结性评价相结合的学习评价探索与实践——以大学数学混合式教学为例[J]。河南农业，2021(30):28-29。

高凌飚（2004）。关于过程性评价的思考[J]。课程.教材.教法，2004(10):15-19。

课堂场域下学生社交投入评测的研究现状及展望

Research Status and Prospect of Student Social Engagement Evaluation in Classroom

余游[1]，魏艳涛[1*]，肖光润[1,2]，徐琦[1]，芣佩舒[1]
[1]华中师范大学人工智能教育学部 数字教育湖北省重点实验室
[2]湖北文理学院
[*]2335404347@qq.com

【摘要】 社交投入是指学生有意识地与他人进行意义协商与协同创造所付出的努力。传统课堂存在数据类型复杂、数据采集困难等问题，导致课堂场域下的社交投入评测研究面临着挑战。所以，本文通过对国内外相关研究进行深入分析，发现大多研究采用自我报告、视频采集等方法进行学生社交投入数据采集，统计分析、内容分析及社会网络分析等方法常用来对学生社交投入进行评测，此外，基于机器学习的方法也逐渐被关注。本文对以上几个方面作出归纳总结，能够为课堂场域下学生社交投入相关研究提供借鉴。

【关键词】 课堂场域；学习投入；社交投入；评测方法

Abstract: *Social engagement refers to students' efforts to consciously engage in meaning negotiation and collaborative creation with others. There are some problems in traditional classroom such as complex data type and difficult data collection, which lead to the challenge of social engagement evaluation in classroom field. Therefore, through in-depth analysis of relevant research at home and abroad, this paper finds most studies use self-report, video capture to collect data on student social engagement. Statistical analysis, content analysis and social network analysis are commonly used to evaluate student social engagement. In addition, machine learning-based methods are gradually being concerned. This paper summarizes the above aspects, which can provide reference for student social engagement into related research in classroom field.*

Key words: classroom field; learning engagement; social engagement; evaluation method

1. 引言

学习投入是影响学生学习成效的关键影响因素，对学生的学习投入进行评测有助于教师改进教学设计、学生进行自我调整，从而提升学习效果。因此，如何精确全面地对学生的学习投入进行评测是值得探究的问题。社会学理论指出，学习本质上就是一个交互的过程，学生通过个性交互和社会性交互获取新知识，而且高效的社会性交互可以促进个性交互，进而提高交互水平。在此背景下，学生社交投入的研究逐渐受到重视。

有研究指出，学习社交投入与学习成效有密切关系，学习社交投入是影响学业成就的重要因素。社交是支撑学习社区融入学习内容、促进非正式讨论和学习的重要主体。对于在线学习，社交投入有助于缓解在线课程学习者的孤独感，维持学习动机，提高课程的黏性和完成率，尹睿等人（2017）基于结构方程模型揭示出社交投入是影响在线学习投入的最重要、最突出的因素；唐璐玢（2020）从交互促进策略、交互实现模式、交互过程及交互质量等方面对在线学习环境下的社会交互进行了研究。社交互动促进了信息和资源的共享，通过同伴间及师生间的互动和讨论可帮助学生更深入地理解学习内容，提高学生的参与感和学习动机并

促进了学习者的社会性学习，以及提供情感支持和归属感，这些都对学业成就产生积极影响，在真实课堂环境中，社交互动也可营造浓厚学习氛围，推动学生融入学习情境中，促进学生分享经验和内化群体知识。

目前研究结合在线协作学习、翻转课堂等不同教学场景对社交投入进行了定义阐述及评测模型建构，但主要关注学生在虚拟空间中的社交互动、合作和网络形成等。教育部基础教育司司长吕玉刚（2021）指出“课堂是学校教育工作的主阵地，学校要把提高课堂教学质量作为核心任务”。从小学到大学的教育中，课堂学习仍占主导地位，在真实的课堂环境下，学生与学生、教师与学生之间更容易建立面对面的互动和关系，在同步交流互动过程中，学生可以保持积极、高效的参与度并进行深层次的思考，这样教师和学生都能得到及时的反馈，有利于促进学生吸收知识及开展讨论、实验、小组活动等互动学习。因此课堂中学生的社会交互同样值得关注，但社交投入涉及学生参与课堂互动、同伴合作、师生关系等多个维度，不同研究对其内涵的界定各有侧重，缺乏普遍认同的概念界定和研究框架，相关文献也存在定义模糊，维度划分不清等情况，这给评测带来了一定的难度。目前学术界少有对社交投入的发生机制及评测方法做系统的归纳，基于以上问题，本文主要围绕其定义、数据采集和评测方法进行一系列探索。

2. 文献检索

为了更好的理解社交投入理论、社交投入评测方法及社交投入如何影响学生学习投入，本文回顾国内外学生社交投入理论和实证研究的相关文献，以此为后续论述提供理论支持。本文使用国内具有代表性的文献来源于中国知网，主要使用“学习投入测量”、“学习投入分析”“混合式学习投入”、“社会交互”、“社交互动”、“协作学习交互研究”、“混合学习”、“多模态学习分析”作为关键词，时间选择 2019 年及以后，一共检索到 825 条结果，为了最大程度保证研究的可靠性和可信度，将期刊来源限定为 cssci 后，剩余 91 篇文献。国外具有代表性的文献分别来源于 Google 学术、Sci-Hub、Springer、IEEEXplore、Web of Science，主要使用“student engagement”、“classroom learning”、“learning engagement recognition”、“measures of blended learning engagement”、“collaborative learning”、“social engagement analysis ”、“classroom social engagement”、“social interaction”、“students’ interactions”作为检索关键词，去除与主题不符的文献，共剩余 57 篇。对选取的文献进行定性分析，以期为新的研究提供理论支持。

3. 社交投入与学习投入概念

学习投入是指个体在学习过程中表现出的一种积极的、充实的精神状态，涉及个体对学习活动的兴趣、动机、积极参与程度以及在学习过程中的努力和持久性。从上世纪 80 年代 Mosher 等人（1985）最早提出学习投入这一术语，随后 Finn（1989）探讨了学生在学校环境的参与程度并提出“参与—认同模式”，将学习投入划分为行为投入与情感投入两个维度；Martin（2008）从行为投入与认知投入两个层面构建了学习投入二维框架；2004 年 Fredricks 等人（2004）基于学校教育研究提出了学术界普遍认可的学习投入三维框架，分别为行为投入、情感投入和认知投入，其中行为投入指积极地参与学术或非学术活动并表现出积极的行为，情感投入包括对教师、同伴和课堂活动的积极情绪反应以及对学校活动的重视和对学习内容的兴趣，认知投入被定义为自我调节学习，使用深度学习策略并运用必要的认知策略来理解复杂的想法和掌握困难的技能。

随着技术和教育理论的不断发展，Kahu（2013）认为学习投入不仅受制度和个人因素的影响也要考虑广泛的社会背景，将社会文化观点与所讨论的情感和行为观点相结合，突出学习

投入的社会文化背景。基于教师和学生对学习投入现象学的定性描述,Fredricks 等人(2016)在之前提出的三因素框架中增加了社交投入（社会投入）维度，并指出该维度由社会情感和社会认知组成。社交投入这一概念考虑了社会互动、合作和寻求帮助在学生学习和学校教育过程中的重要性并代表了超越个人学习的一种进步。Linnenbrink 等人（2011）根据课堂任务参与的社会形式，认为社交投入包括学习投入和社会互动的质量。Finn 等人（2012）将社会投入定义为学生在课堂上的亲社会行为以及围绕教学内容与同龄人互动的质量。 Wang 等人（2016）提出社会投入包括学生与同伴和成年人的社会互动质量以及在学习过程中用于建立和维持关系的意愿程度；Muramatsu 等人（2016）指出社交投入涉及积极的社会感交互，它既可以反映小组成员之间的相互尊重程度，也可以反映小组的内部凝聚力，同时也把任务概念化为一个团队的努力成果，而不是个人的活动。尹睿等人（2017）将社交投入界定为学生与教师、同伴、学习共同体之间的交互，如师生之间的反馈交流、同伴作业互评、学生的资源与经验分享、小组的协作学习活动等。李新等人（2023）最新研究揭示社会投入考虑了学习过程中同伴之间的社会性互动、协作、寻求帮助等行为，丰富了以往针对个体学习投入的研究范畴，为更准确全面测量学生投入提供了新的视角。

此外，互动作为学生社交投入的表现方式在不同程度上影响着学生的学习投入。Lu 等人（2014）采用了定性和定量的方法进行了社交互动对 13 名学生的学习投入作用的研究，得出社交互动可使他们的学习投入达到较高的程度。Nennig 等人（2023）基于学习者群体理论，通过分析大学普通化学课堂上完成课堂任务时的小组对话来描述学生的社会和认知投入并指出学习投入是一个跨越认知、行为和情感维度的多维结构，受学生先验知识、动机、情绪及社交互动等因素的影响。李军靠（2021）表明有效的教学互动可以整合情感和智力因素，提高认知投入；促进并发展师生互动与合作，可提高行为投入；减少学生学习孤独感，促进情感投入。本研究社交投入概念图如下图 1：

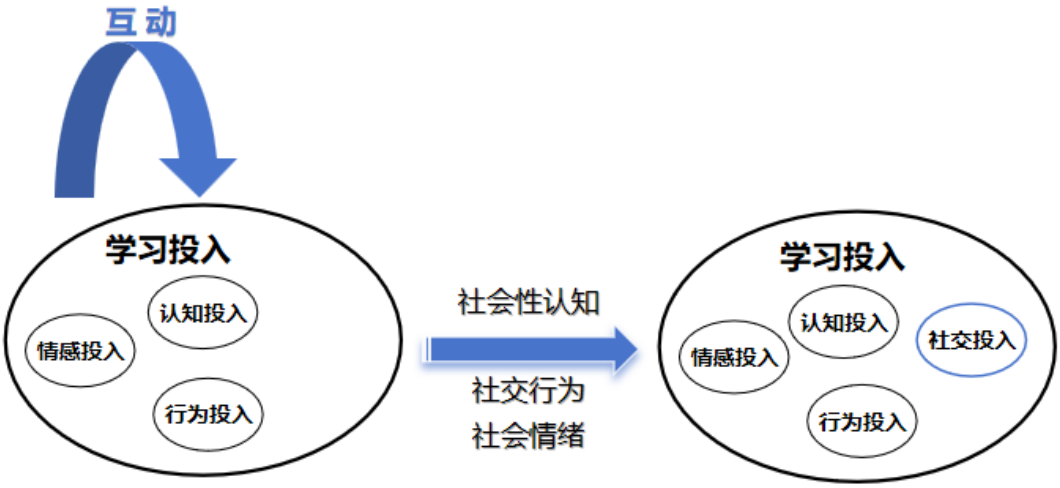


图 1 社交投入概念图

综上所述，课堂场域下学生社交投入是指学生在课堂这一特定社会情境中与同伴和教师互动交流所付出的努力及表现出的心理状态，其不仅反映了学生的社交技能，还对学生学习动机、情感发展和学业成就有重要影响，通过课堂讨论、小组合作学习以及与教师和同伴有效

的沟通等可促进学生的社交投入进而提高课堂学习效率，是课堂教育过程中不可忽视的重要因素。

4.学生社交投入的数据采集方法

4.1.自我报告法

自我报告法是通过发放问卷了解学生的学习态度、动机、投入度等学生社交投入关键因素。Kaufman 等人（2015）探究师生互动质量和学生性别对五年级数学学生学习投入的影响程度中对学生进行问卷调查，其中社交投入维度包含的项目有关于解释学术背景和在数学课课堂上讨论想法等。Wang 等人（2016）以数学和科学领域为重点，采用顺序混合方法设计了社会投入量表（五分制的李克特量表），例如：我借鉴了别人的想法；在数学/科学课上我试图理解别人的想法等,反映了围绕教学内容的社会互动和对同伴的群体互动情感。Hunsu 等人(2018)基于社交网络框架提出社交投入三因素模型（信任、互惠和归属感），以此开发一个有 11 项题目的用于课堂的量表来评估学生在社交网络中发展社会资本时的社会投入度，如：当我与这个同学交流的时候，我得到了我想要的好处。Zelhendri Zen 等人（2019）在探讨翻转课堂模式可加强学生社交投入研究中，所采用的题项主要围绕由社会化能力、人际交往能力和与他人合作的自我控制三个要素展开，这进一步丰富了学生社交投入自我报告的内容。

4.2.访谈法

通过结构化、半结构化、非结构化访谈等,可深入了解学生的社交互动、主观感受和态度，访谈内容可包括学生在课堂内外的社交投入情况，与同伴或教师的互动经历及遇到的社交挑战或障碍等问题，有助于获得学生个人视角的深刻洞见，可为了解学生的社交投入提供丰富的深层次信息。Fredricks 等人（2016）在评测数学和科学投入时采用半结构化访谈，访谈问题主要有：投入和脱离对你意味着什么；是什么影响了你对数学和科学的投入或脱离。Martin 和 Collie (2019)采用半结构化访谈,结合经验取样法,多次访谈了高中生,捕捉他们在日常学习生活中的社交投入动态变化,揭示了个体内部社交投入的波动规律。Wang 等人(2019)通过采访学生与同伴一起工作的感受以及制定社交策略是否支持他们有成效地参与学术活动等问题，以反映青少年课堂学习及在更大的学校环境中社会互动的重要性。

4.3.视频采集

视频采集可以提供客观的记录，捕捉学生在学习过程中的言语表达、面部表情、姿势等多个方面的数据，提供了更多维度的信息，允许研究者重复观看特定场景，以更全面地评估学生的社交投入。Sanna 等人（2016）在数学教学课程中设置分组学习，收集了为期 7 周 84 小时的视频记录，并对小组的认知和社会情感互动及互动中自我调节的三个阶段进行编码。Balaman 等人（2017）针对线下协作学习环境，利用高清摄像头采集了学生的表情、面部、姿态、手势、互动行为数据。李新等人（2022）通过摄像机、录音笔、录屏软件、学习平台等方式收集学生的录屏数据、对话文本数据、学习视频数据和学业表现数据来分析协作学习情境中学习投入，其中对学习录像进行编码来分析学生的社交投入。Hannula 等人（2022）在推进视频研究方法以捕捉社会互动和多模态信息的过程中强调了视频记录为深入分析同伴互动和学习过程提供了丰富的数据。Vuorenmaa 等人（2023）收集物理任务协作学习场景下的 92 名中学生录像，并从投入、社会互动、共同和社会共享调节的群体调节类型等方面进行分析。

5.学生社交投入的评测方法

统计分析法、内容分析法、社会网络分析法和基于机器学习的方法是评测学生社交投入度较为关键的研究工具，每种方法各具特色且相互补充。统计分析法适合处理大规模数据集，揭示学生社交投入模式或趋势，但高度依赖数据质量；内容分析法特别适合于分析学习平台上的文本和互动内容等，可提供关于社交行为的深入见解，但可能受主观解读的影响；社会网络分析法揭示学生社交结构和互动模式，适用于客观分析社交动态，但容易忽略个体的内在动机；基于机器学习的方法是新兴趋势且具有巨大潜力，尤其在处理复杂数据集和揭示难以察觉的模式方面，但需高水平技术支持并要注意算法应用问题，未来发展可能会聚焦于机器学习与其他方法相融合，以实现更全面、更精确的评估。

5.1.统计分析法

统计分析法是一种用于处理、分析和解释数据的数学方法，它包括描述性统计分析、推断性统计分析和多变量统计分析，这种方法能够帮助研究人员从数据中提取有意义的信息、识别模式和趋势以及做出基于数据的预测和决策，增加了研究的普适性和推广性。Zelhendri Zen 等人（2019）将翻转课堂模式和社交投入设为自变量，学生高阶思维能力设为因变量来探究翻转课堂的实施与社会投入度的强化和学生学业成绩提高之间的关系。Wang 等人(2019)采用双因素建模技术和混合方法开发并验证投入度多维概念化的学校投入度学生调查测量表，将验证性因素分析（CFA）作为检验因素结构的主要方法。杨新荣等人（2021）基于结构方程模型对中学生数学学习投入问卷进行分析，揭示出社会投入是影响中学生数学学习投入最重要的因素，它对学生在数学学习过程中行为、情感和认知投入都有着显著的正向影响作用。

5.2.内容分析法

内容分析法主要体现在对学生作业、课堂互动等的深入分析，定量或定性地分析文本或口语数据，以识别学生学习过程中的模式、理解程度和投入程度。相比传统的调查方法，内容分析法可以快速地处理和分析大量数据，不仅可以评估某一时间点的社交投入，还能够跟踪学生社交投入的变化趋势。Lu 等人（2014）通过文本分析和描述性统计，研究了社会交互对本科阶段的学生在社交网络环境下学习投入的影响。Soffer 等人（2019）通过分析学习管理系统日志文件来评价学生的学习社交投入，包括在线作业、同伴交互、学习活动等指标。Mustafa 等人(2022)在探究翻转中学数学课堂中的学生投入时，通过对学生访谈、日记及问卷等记录进行内容分析，以确定学生在非翻转和翻转教学方法中的数学学习体验，得出要成功地进行翻转学习，重要的是有一个完善的交互设计并在数学教学中鼓励社会互动，加强学生在社会环境中的参与和数学学习。

5.3.社会网络分析

社会网络分析（SNA）提供了一种量化社交互动和连接强度的方式，可以用来分析学生之间的交流模式、合作关系以及知识共享情况，通过了解这些社交网络的结构和动态，教育者可以更好地理解学生的投入度、协作学习的效果及社交互动如何影响学习成果，有助于客观评估学生的社交投入程度。Lee 等人（2016）通过社会网络分析法探究学生个体在群体中的位置和亲密度，进而评估学生社交投入。马志强等人（2020）将社会认知网络分析引入协作交互研究中，并指出该方法能够揭示协作学习小组的社会认知网络结构特征与关键指标，及描述小组网络结构演化的基本过程。尹睿等人（2023）在混合学习投入评价中利用该方法将学生社交投入水平表征为交互深度、交互广度，以此探究了班级整体的社会网络分布，以及不同类型学生的社会关系情况。

5.4.基于机器学习的方法

随着教育数字化转型的发展，学生在学习过程中产生大量的数据，这些数据包括结构化数据和非结构化数据等（例如图像、音频）。鉴于学生在社交投入过程中产生数据类型的复杂性，学术界目前还没有明确提出专门用于测量学生社交投入的机器学习或深度学习方法。因此，有研究者将社交投入概念化为社交情绪，社交互动等不同维度，从情感和行为方面采用机器学习方法对社交投入进行评估。

Pekrun 等人（2012）指出社交情绪可以直接影响学生对学业任务的投入，尤其是当处于师生或生生互动中时，Isohätälä 等人（2020）将积极的社会情绪互动分为两类：社会情绪支持（即鼓励、表扬、恭维、道歉等）和幽默（即善意的玩笑、大笑等）。对于学生情绪检测，Zhu 等人（2023）提出了 NAGNet 网络，通过分析面部表情来推断学生的情绪状态。Sümer 等人（2023）采用 Attention-Net 和 Affect-Net 等深度网络来提取其注意力和情感特征，分析课堂上学生的面部视频。也有基于语音的分析，记录学生在课堂中的发言、讨论数据,利用语音识别和自然语言处理技术提取语义和情感特征,评估学生的参与度和投入质量。例如使用循环神经网络(RNN)对语音序列建模,判断学生发言的活跃程度。因此，可利用深度学习网络提取学习者情绪特征，然后根据社会情绪类别进行编码进而分析学生的社交投入度。

Dewaele 等人（2021）在探究教师积极性与学生的社会行为学习投入研究中主要关注社会行为投入，如学生在小组中与同学的互动或与老师的互动。美国学者弗兰德斯曾提出一种课程观察工具--互动分析系统（FIAS），对课堂中师生的互动语言行为进行统计分析。因此，在行为互动过程中我们也可聚焦在学生所表现出的社交行为动作并对其进行分析，利用机器学习方法提取出学生在互动过程中表现出的肢体行为，经编码进而分析学生的社交投入度。比如何秀玲等人（2020）选取 OpenPose 骨架检测模型提取学生行为骨架特征，结合 CNN-10 实现学生行为的有效识别。学生与学生之间互动主要围绕小组活动展开，在小组中识别多人行为可更有利于提取出学生互动行为。如 Albert 等人（2022）提出结合实时目标检测来获得每个学生的实时图像流，利用深度时空残差 CNN 学习每个学生行为的时空特征，可实时识别课堂教学场景中多个学生的课堂行为。考虑到学生社交投入是一个连续的过程,也有一些研究采用序列模型如隐马尔可夫模型(HMM)、条件随机场(CRF)等,对学生一段时间内的行为序列进行建模,捕捉社交投入的动态变化模式,进而实现连续的投入状态评估。

6.挑战与展望

6.1.挑战

社交投入数据的多样性和复杂性。学生社交投入的评估涉及多种类型的数据,如面部表情、语音、文字交流、身体姿势语言等，具体而言,学生社交投入的言语数据包括与教师、同伴的对话内容、提问回答、小组讨论等,这些蕴含着丰富的语义、情感等信息;非言语数据涉及面部表情、眼神交流、手势姿态、身体朝向等,可反映学生的情绪状态和参与程度;生理数据如心率、脑电等,提供了学生社交投入的生理唤醒和认知加工的客观指标;学生的课堂笔记、作业、测验等学习产出数据,也蕴含着学生社交互动对学习的影响信息。此外，在真实课堂环境中采集学生社交行为数据,会受到教室布局、师生位置、设备安放等物理环境因素的制约,数据质量的一致性难以保证。课堂活动的多样性和教学流程的动态性以及学生在不同教学情境下的社交行为的差异性,这些都增加了数据采集的难度。面对处理在不同教学平台下收集的大量的、复杂的多模态数据时，选择合适的数据融合策略、机器算法及模型对数据进行分析也是一个挑战。

评测社交投入的工具与算法模型具有局限性。传统的评估方法如自我报告问卷、观察法等

存在一定局限，问卷调查容易受到学生主观性的影响，观察法虽然更加客观，但耗时耗力，难以大规模实施。已有评估工具大多聚焦于社交投入的行为层面，对学生的情感体验、认知加工等内部过程关注不够，未能深入挖掘社交投入的心理机制。评估内容的选取也缺乏统一的标准，不同研究者从不同维度进行测量，评估结果的精准性有待提高。以及学生的社交互动往往涉及大量非结构化数据，如自然语言、面部表情等，处理这些数据要选择正确的算法和模型，例如，一些算法可能对特定类型的社交互动（如小组讨论）更敏感。另外不同文化背景、不同年龄段及教育阶段的学生所展现的社交投入特点不同，需开发具有适应性的模型。根据具体的应用场景和数据特点要适当地进行算法的调优，以提高模型的准确性，更需要采用严格的验证方法及评估指标来全面衡量模型的表现，包括准确率、召回率、F1 分数等。

6.2. 展望

多模态数据融合精确评估学生社交投入。通过收集学生在课堂中的视觉、语音、文本等多渠道行为数据，利用机器学习和深度学习技术进行联合建模和分析，可全面刻画学生社交投入的多维度表现，综合考虑学生的面部表情、语言互动、文本表达等多种社交线索，克服了单一数据源的局限性，实现对学生社交投入的客观、动态、连续的评估，精细刻画学生在不同时间点、不同情境下的投入状态变化。多模态社交投入分析的实现，离不开机器学习和深度学习技术的支持，未来研究可进一步探索多模态数据的高效采集、标注和处理方法，提高评估模型的可解释性和泛化性，还需探索多模态数据的融合架构(如早期融合、晚期融合、混合融合等)，以发挥不同模态数据的优势，通过多模态数据的融合分析，可以从行为、情感、认知等多个层面对学生的社交投入进行立体化、多角度的刻画，促进多模态数据融合在教育实践中的应用，为评测学生社交投入水平提供科学支撑。

跨学科研究助力评测学生社交投入。评测学生社交投入是一个复杂的问题，涉及学生的行为表现、心理体验、认知加工等多个层面，单一学科的理论和方法难以完全解释其内在机制。通过整合教育学、心理学、社会学、脑科学、计算机科学等多个学科的研究成果，可以构建更加综合、立体的社交投入评测框架。同时，跨学科研究也有助于厘清社交投入与学习投入、学业成就、心理健康等变量之间的关系，揭示社交投入在学生发展中的重要作用。未来研究可将教育理论与数据分析技术相结合，利用数据科学的方法分析学生的学习和社交行为。比如，心理学和行为科学的整合有助于解读学生的社交行为和情感状态，为数据模型提供深层次的见解；计算机科学家和 AI 专家可开发高级算法和工具，用于收集和分析学生的社交互动数据；社会学家和文化研究专家可提供关于不同文化背景对学生社交行为影响的意见，帮助开发更具包容性和适应性的评测工具等。

7. 研究结论

本文综合分析了课堂环境中学生社交投入的评测方法及其重要性，为理解学生在课堂中的社交投入提供了新的视角，对教育实践和未来研究具有重要意义。研究发现，准确评估社交投入对于学生的社交技能、情感交流、合作学习及社交需求等方面有促进作用，可指导教育者判断教学设计的效果，调整教学内容、方法和策略，创设有利于学生积极互动、主动参与的课堂氛围，进而提供个性化支持和有效干预措施，促进学生全面发展。面对当前在课堂环境中社交投入评估工具的局限性，多维数据的采集和整合困难等问题，本文提出未来研究应考虑多模态数据融合技术和跨学科合作研究，深入地了解学生的非语言交流及不同社交场景下学生的互动模式，从而提供更准确的学生社交投入评估。

基金项目

国家自然科学基金“面向同步直播课堂的可解释学习投入自动评测方法研究”（项目号：62277029），教育部人文社科项目“基于人工智能的在线学习参与度识别研究”（项目号：20YJC880100），武汉市知识创新专项“同步直播课堂中远端学生学习投入自动评测方法研究”（项目号：2022010801010274），华中师范大学交叉团队项目"融合多模态数据的教师教学能力智能评价技术及应用研究"（项目号 CCNU22JC011）

参考文献

马志强，管秀.(2020).面向多维关联的社会认知网络分析——协作学习交互研究的新进展.远程教育杂志(06),96-103.

尹睿,徐欢云.(2017).在线学习投入结构模型构建——基于结构方程模型的实证分析.开放教育研究(04),101-111.

尹睿，何淑茵.(2023).循证教学评价视角下混合学习投入评价研究.电化教育研究(05),120-128.

李军靠,王明慧.（2021）.教学交互对大学生在线学习投入的效用及提升策略[J].延安大学学报(社会科学版),43(06):106-111.

李新,李艳燕.(2022).协作学习情境中学习投入数字化表征模型的构建与应用.中国远程教育(08),70-77.

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

何秀玲,杨凡,陈增照,方静 李洋洋.(2020).基于人体骨架和深度学习的学生课堂行为识别.现代教育技术(11),105-112.

杨新荣,林明 马林.(2021).中学生数学学习投入结构探索——基于结构方程模型的分析.西南大学学报(自然科学版)(04),1-9.

姚晓丹.(2021).“双减”后第一个学期,中小学校将有哪些变化——访教育部基础教育司司长吕玉刚.中学语文(26),86.

唐璐玢.(2020).在线 MOOC 学习中社会交互对学习投入的影响机制研究硕士学位论文,电子科技大学).

Albert, C. C. Y., Sun, Y., Li, G., Peng, J., Ran, F., Wang, Z., & Zhou, J. (2022). Identifying and Monitoring Students' Classroom Learning Behavior Based on Multisource Information. *MOBILE INFORMATION SYSTEMS*, 2022.

Balaman Olcay Sert.(2017).The coordination of online L2 interaction and orientations to task interface for epistemic progression.*Journal of Pragmatics*115-129.

Cevikbas, Mustafa Kaiser, Gabriele.(2022).Student Engagement in a Flipped Secondary Math ematics Classroom.*International Journal of Science and Mathematics Education*(7),1-26.

Dewaele, J. M., & Li, C. (2021). Teacher enthusiasm and students’ social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes. *Language Teaching Research*, 25(6), 922-945.

Finn, J. D. (1989). Withdrawing from school. Review of educational research, 59(2), 117-142.

Finn,J.D., Zimmer, K.(2012).Student engagement: what is it and why does it matter? In S.Christenson,A. L.Reschy, & C.Wylie(Eds.), Handbook of research on student engagement (pp. 97e131).*New York: Springer*.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of

the concept, state of the evidence. *Review of educational research*, 74(1), 59-109.

Fredricks J. A, Ming-Te Wang, Jacqueline Schall Linn, Tara L. Hofkens, Hannah Sung, Aly ssa Parr & Julia Allerton. (2016). Using qualitative methods to develop a survey meas ure of math and science engagement. *Learning and Instruction* 5-15.

Hunsu, N., Simmons, D. R., Brown, S. A., & Adesope, O. (2018). Developing an instrume nt of classroom social engagement. Association for Engineering Education-Engineering Library Division Papers.

Hannula, Markku S., Haataja, Eeva, Löfström, Erika, Garcia Moreno Esteva, Enrique, Salmi nen Saari, Jessica F. A. Laine, Anu. (2022).Advancing video research methodology to c apture the processes of social interaction and multimodality. ZDM – *Mathematics Educ ation*(2),1-11.

Isohätälä, J., Näykki, P., & Järvelä, S. (2020). Convergences of joint, positive interactions and regulation in collaborative learning. *Small Group Research*, 51(2), 229-264.

Jie Lu Daniel Churchill.(2014).The effect of social interaction on learning engagement in a social networking environment. Interactive Learning Environments (4),401-417.

Kahu,E.R.(2013).Framing student engagement in higher education. *Studies in Higher Education*,38(5),758–773.

Linnenbrink-Garcia, L., Rogat, T., Koskey, K. (2011).Affect and engagement during small group instruction. *Contemporary Educational Psychology*,36,13e24.

Lu J,Churchill D.(2014).The effect of social interaction on learning engagement in a social networking environment[J].Interactive Learning Environments,22(4):401-417.

Lee, J., Bonk, C. J. (2016). Social network analysis of peer relationships and online interactions in a blended class using blogs. *The Internet and Higher Education*, 28, 35-44.

Mosher, R., & MacGowan, B. (1985). Assessing Student Engagement in Secondary Schools: Alternative Conceptions, Strategies of Assessing, and Instruments.

Martin A J.Enhancing student motivation and engagement:The effects of a multidimensional intervention[J].*Contemporary Educational Psychology*,2008,33(2):239-269.

Muramatsu,K.,Tanaka,E.,Watanuki,K.,& Matsui,T.(2016).Frame-work to describe constructs of academic emotions using ontological descriptions of statistical models. *Research and practice in technology enhanced learning*,11(1),5.

Martin, A. J., & Collie, R. J. (2019). Teacher–student relationships and students' engagement in high school: Does the number of negative and positive relationships with teachers matter?. Journal of Educational Psychology, 111(5), 861.

Nennig, H. T., Macrie-Shuck, M., Fateh, S., Gunes, Z. D. K., Cole, R., Rushton, G. T., ... & Talanquer, V. (2023). Exploring social and cognitive engagement in small groups through a community of learners (CoL) lens. Chemistry Education Research and Practice, 24(3), 1077-1099.

Pekrun, R., Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In Handbook of research on student engagement. Boston, *MA: Springer US*, 259-282

Rimm-Kaufman,Sara E., Baroody,Alison E., Larsen,Ross A. A., Curby,Timothy W. Abry,Tash ia.(2015).To what extent do teacher–student interaction quality and student gender contri bute to fifth graders' engagement in mathematics learning?.*Journal of Educational Psyc*

hology(1),170-185.

Sanna Järvelä, Hanna Järvenoja, Jonna Malmberg, Jaana Isohätälä Márta Sobocinski.(2016).How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement?. *Learning and Instruction*39-51.

Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., & Kasneci, E. (2023). Multimodal Engagement Analysis from Facial Videos in the Classroom. *IEEE Transactions on Affective Computing*, 1012–1027.

Tal Soffer Anat Cohen.(2019).Students' engagement characteristics predict success and completion of online courses.*Journal of Computer Assisted Learning*(3),378-389.

Vuorenmaa Eija,Järvelä Sanna,Dindar Muhterem Järvenoja Hanna.(2023).Sequential Patterns in Social Interaction States for Regulation in Collaborative Learning. *Small Group Research*(4),512-550.

Wang，Ming-Te,Jennifer A. Fredricks,Feifei Ye,Tara L. Hofkens Jacqueline Schall Linn.(2016).The Math and Science Engagement Scales: Scale development, validation, and psychometric properties.*Learning and Instruction*, 16-26.

Wang, Ming-Te,Fredricks, Jennifer,Ye, Feifei,Hofkens, Tara Linn, Jacqueline Schall.(2019).Conceptualization and assessment of adolescents' engagement and disengagement in school: A Multidimensional School Engagement Scale.*European Journal of Psychological Assessment*(4),592-606.

Zelhendri Zen Syamsuar.(2019).Implementation of Flipped Classroom Model to Strengthening Students' Social Engagement and Its Relation with Students' High Order Thinking Performance.(eds.)

Zhu, H., Hu, P., Tang, X., Xia, D., & Huang, H. (2023). NAGNet: A novel framework for real-time students' sentiment analysis in the wisdom classroom. *Concurrency and Computation: Practice and Experience*, e7727.

数智赋能师范生教学设计能力提升的教学评一体化研究

——以《小学信息技术教育》课程为例

Research on the integration of teaching and evaluation for the improvement of teaching design ability of mathematics intelligence-enabled normal students

-Taking the ' Information Technology ' course as an example

吴欣熙[1]，陈思婷[1]，钟卓华[1]，崔萌[1*]
[1]广东第二师范学院 教师教育学院
*mengcui@m.scnu.edu.cn

【摘要】 提高师范生教学技能是新时代发出的“强师计划”吁求的应有之义。以教学评一体化为理论基础，数智赋能驱动模式引领，在“课堂教学前➡课堂教学中➡课堂教学后”等情境中将“教学评”有机融合，促进师范生教学设计能力提升。具体来说，是师生在完成评价任务的过程中，通过可见的学习证据将评价融入教学的全过程，依据评价推论教与学的成效，优化学习计划，再以可理解的方式呈现。并将这种教学方法，凝练成一个旨在提升信息科技课程师范生教学设计能力提升的教学评一体化教学设计方案，以期为中国式现代化全力打造高质量教师队伍提供借鉴。

【关键词】 教学评一体化；职前教师；数字素养；教学设计

Abstract: *Improving the teaching skills of normal students is the proper meaning of the " strong teacher plan " issued in the new era. Based on the integration of teaching evaluation, guided by the model of digital intelligence empowerment, the " teaching evaluation " is organically integrated in the situations of " before classroom teaching → in classroom teaching → after classroom teaching, " so as to promote the improvement of normal students ' teaching design ability. Specifically, in the process of completing the evaluation task, teachers and students integrate evaluation into the whole process of teaching through visible learning evidence, infer the effectiveness of teaching and learning based on evaluation, optimize the learning plan, and then present it in an understandable way. This teaching method is condensed into an integrated teaching design scheme of teaching evaluation aimed at improving the teaching design ability of normal students in information technology courses, in order to provide reference for the Chinese-style modernization to build a high-quality teacher team.*

Keywords: Integration of teaching evaluation, pre-service teachers, digital literacy, teaching design

1. 引言

百年大计，教育为本；教育大计，教师为本。党的十八大以来，以习近平同志为核心的党中央将教师队伍水平提升建设摆在突出位置，并结合国内教师现状作出一系列重大决策部署，教师队伍建设取得显著成就。其中各地政府和师范类院校也积极出台细化政策助力教师队伍水平提升建设的落实，在 2022 年广东省教育厅印发的《广东“新师范”建设实施方案（2022—2025 年）》中明确指出，目前师范生培养存在学科知识不扎实，师范基本技能不强，信息化素养不足，人才培养模式单一等问题。由于师范生的教学能力和教学理念对其未来成为正式

在职教师之后的实际教学行为有着直接的影响，因此提升师范生的培养模式的深化改革势在必行。

本研究将基于师范类专业认证的视域构建高师院校师范生核心素养的评价指标体系，选取广东省内高校师范生开展实证研究，准确把握当下高师院校师范生核心素养培育能力中的教学设计能力提升的实践困囿，并在此基础上为我国师范教育质量的提升与优化提供精准建议，切实推动以师范生为中心的课堂教学方法和评价方式变革，培养师范生反思性学习能力、实践能力和创新能力。并且利用信息技术变革教师教学方式和师范生学习方式，提升师范生信息素养和利用信息技术促进教学的能力。坚持以生为本，推进评价方式变革，树立“全面考核、注重过程，强化能力、促进发展”等的理念，逐步建立“内容综合化、形式多样化、过程全程化”的评价模式，使评价的重点真正落实到学生能力培养和自身素质提高上。

2. 师范生教学能力培养提升的驱动要素

2.1 师范生实践素养提升是新师范建设的重要任务

教育强国，师范先行。高质量教师是高质量教育发展的中坚力量，师范生的培养是决定未来教师教育高质量发展的关键。随着“互联网”的发展带动着教育信息化逐步迈向“创新引领”转变，而师范生是构建新时代高质量教师队伍中的“新鲜血液”和主力军，他们的教学技能水平影响着其未来专业化发展的进程，也直接影响着基础教育的质量和水平。为顺应时代发展的需求，新时代的师范生应该具备应用信息技术的能力，将学科和技术有机融合，实现教育创新发展（马莲姑、黄寿孟和贾晓婷，2022）。

强化教育实践环节，重点提升师范生实践素养是“新师范”建设的重要任务。《教育部关于实施卓越教师培养计划的意见》中明确指出信息技术对提升教学质量的重要作用同时也对师范生的信息素养提出了明确的要求，并表明了“大力提升师范生的信息素养”；教育部等八部门在 2022 年联合印发的《新时代基础教育强师计划》中提出，鼓励开展师范生培养改革，创新师范生教育实践机制模式，提升教师教育质量，是构建优质均衡的基本公共教育服务体系，推动教育高质量发展的切实可行之策（朱旭东，2022）。因此，本篇文章结合广东省政府颁布的《广东“新师范”建设实 施方案（2022—2025 年）》文件，提出深化师范生培养模式改革，强化师范生实践能力培养，建设广东特色“新师范”。

2.2 数据密集型科学研究范式助力数智赋能教育实践

数据密集型的科学研究范式推动着大数据与教学的深度融合，助力数智赋能下师生的教与学的进阶重塑。信息技术与互联网技术的飞速发展让教育行业进入了大数据时代。由于学习者在整个学习过程中的学习行为数据能够被记录下来，因此这种多源过程性数据为基于大数据的教育评价提供了一种全新的可能。在传统教育评价体系中，学习者的评价多为形成性评价与终结性评价相结合的方式，然而这两种评价方式都存在着评价内容单一、评价方法单一等的问题，不能全面地反映出学生的真实水平。

随着可收集的学习者过程性数据日益丰富，传统的教育评价正在向大数据驱动的科学评价转变。十九大以来，我国的教育信息化发展也开始步入 2.0 阶段，教育大数据和 AI 支持下的教学数据分析也作为技术触点推动着教育生态智能化发展。信息技术与教育的融合也从简单的“应用融合”转变为更强调信息技术支持下教育教学创新发展的“融合创新”（李毅、何莎薇、邱兰欢，2020）。“数据驱动”型教学的运行机制是教师、学生、教学媒介和教学内容等围绕教学大数据所进行的多边互动。这种集数据论、工具论和方法论于一体的大数据与教学的深度融合重塑了师生的教与学，同时也促使教学范式从“基于经验”向“数据驱动”转型；教学内容从“学科知识”进阶“个体智慧”；教学方法从“规范讲授”迈入“具身体验”；教

学评价从“观察判决”转向“证据评估”。数据密集型科学研究范式、计算教育学、教育人工智能等创新理念助力数据驱动的教育实践。

2.3 教学评一体化设计是促进有效学习的重要路径

“教学评一体化”起源于 20 世纪 80 年代美国掀起的一场“基于标准的教育改革运动（standard-based reform，简称为 SBR）”，其要求将教师的教学、学生的学习、教学的评价与课程的标准保持一致（McLaughlin, Milbrey W. & Lorrie A，1995）。随后，各国也先后开展了以教学评为核心的“为了学习而评价”的研究，倡导将评价与教学相结合，实现“教学评一体化”。2022 年《义务教育课程标准》明确指出强化考试评价与课程标准、教学的一致性，促进“教-学-评”有机衔接，注重评价与教学过程“伴随式”开展。“教、学、评”一体化是指依据课程目标、教材和学生的学情这“三大依据” 确定教学目标，将教学目标达成度的测评嵌入到教学过程中，依据从学生的学习表现和测评的结果及时调整教学，精准引领学生走向学习重点，有效落实“教学评一体化”（褚艳秋，2023）。“评价促进学习”是教育评价改革重要的理念转变。新课标理念下倡导的 课堂教学评价是过程性与终结性有机结合的、多元化的、灵活的、伴随性的。即评价不仅仅是教学前后的一个诊断和对于教学成果检验的单一标准；这标志着教育评价从“对学习的评价”转向“为学习的评价”和“作为学习的评价”。

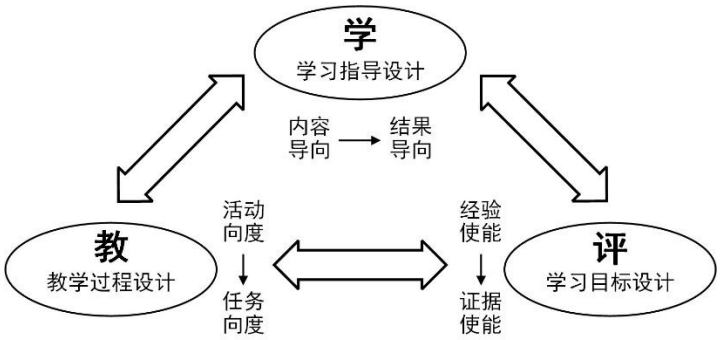


图 1 教学评一体化思路侧重点

2.4 信息科技课程优质师资培养极为关键且迫在眉睫

随着信息时代 2.0 的到来，信息科技课程优质师资的培养也刻不容缓。人工智能融入教育是新时代社会发展与时代进步的必然之选。而作为与新技术关联最为密切的信息技术教师更应该以一种开放的姿态去拥抱技术，并将其应用于日常的教育教学中。2022 年《义务教育课程标准》首次独立设置“信息科技”国家课程，体现出国家对于信息科技教育的重视。教师是新课程的执行者，相对于其他专业，信息科技教师的知识迭代、技能更新速度快，这就对其专业素养、能力发展提出了更高要求（李海英，2022）。“信息科技”课程标准的单独设置是良好的开端，优质的信息科技师资是落实新课标的“最后一公里”。所以在信息时代 2.0 的背景下，推进人工智能下的教师教育改革，培养和发展富有智能教育素养的信息科技课程师范生培养极为关键且迫在眉睫。

3. 数智赋能下师范生实践素养中教学设计能力提升的教学评一体化培养模式及应用路径研究设计

3.1 明确研究目的和研究设计理念：利用数据赋能下的教学评一体化推动以评定教，学评融合的师范生培养

本研究的主要目的是：在数据驱动的理念下，利用信息技术将学生学习过程中的教学-学习-评价三个阶段有机的融合在一起，利用数据赋能实现“教学评一体化”设计，从而提升师范生实践素养中的教学设计能力。如何基于 AI 教学支撑平台的数据分析技术，将“教学-学

习-评价”三者有机融合，真正实现以评定教、学评融合的一体化设计，助推师范生信息素养中教学设计能力的提升，是主要的研究内容之一。

其中在研究的设计上，研究拟在广东省的师范类院校中选取部分信息技术方向的师范生作为研究对象，选择《小学信息技术教育》这一门课程作为实验的课程，在课程的教学设计和实施中，融入教学评一体化的思想，在师范生实践能力的教学设计编写能力的培养过程中始终倡导“学生本位”，实现以评定教，学评融合的面向信息技术学科核心素养的、具有教学实用性的教学模式。

3.2 锚定研究方向：融合文献资料与实际痛点问题锚定研究的具体方向

教学设计是师范生实践素养中的一个重要的环节，同时也是一种复杂的教学技术。美国教育加杜威（J.Dewey）在 1900 年的时候提出了应建立一门链接学习理论和教学实践的“桥梁学科”（linking science），其目的主要是建立一套系统的与设计教 学活动也有关的理论知识体系，以达到优化教学的目的（J.Dewey，1990）。并且教学设计是师范生所需要掌握的基本技能之一，也是在教师教学过程中，所必不可少的一个部分。通过教学设计，教师可以将已有的教学理论和研究成果运用于实际的教学中，以指导教学工作的实施；同时在另一个方面，教学设计可以让教师的教学经验升华成教学理论，指导教学实践。

其中如何基于“教学评”一体化，利用 AI 智能辅助教学平台——小雅平台实现线上线下融合教学的方式助推师范生教学设计能力的提升是本项目研究的主要解决的实际问题之一。本研究主要设计在《小学信息技术教育》课程中的教学设计教学环节部分。

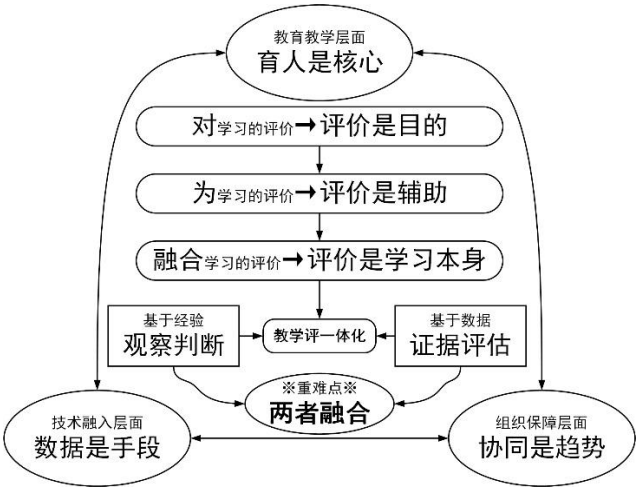


图 2 拟解决的关键问题

3.3 研究设计：联合 AI 智能辅助学习平台与伴随式测评协同赋能师范生教学设计能力提升

3.3.1 课程开始教学前：

在针对于师范生的教学设计能力提升的教学内容开始前，教师将会在课前给本节课上课的学生发布一个作业，作业的主要的要求是——让学生自主完成一个与信息技术相关的教学设计，并且将自主完成的教学设计提交到我们的 AI 智能辅助教学平台上面（如图 8）。学生提交的教学设计是基于学生当前能力范围内的教学设计，这一份教学设计是一个基础，在后续的教学过程中，学生可以结合课程学习的知识去对自己制作的教学设计进行不断的优化。同时教师在课程开始前，需要先分析学生提交上来的教学设计，分析学生在在教学设计能力具体环节方面的优点与不足，最后将分析出来的数据形成“学生学情画像”并将其结合课程内容与课程目标确定教学的起点与教学的内容。



图 3 AI 智能辅助教学平台内课程建设

3.3.2 课程项目教学时：

教师在课程教学中可以根据教学设计中的各个不同的部分对课程进行正常拆分如划分为：学习者分析部分、教学目标部分和课堂环节部分等，在课程内容的教学中，可以结合学生在课前提交的教学设计。选取部分学生所提交的教学设计中于本节课相关联的部分，作为典型案例与课堂教学内容相结合进行针对性分析点评，并提出相对应的改进建议。做到在课堂上能够帮助学生打通知识与实际应用之间的桥梁。

对应不同课程的每个知识点都可以采取相似的方式，选取不同学生的教学设计分别进行针对性的点评，将教师的教学、学生的学习和评价进行一个有机的结合，帮助学生能够更好地认识到自己在教学设计能力方面存在的不足和能够继续提升优化的方向，将课堂上学习到的知识更好地运用到实际中。

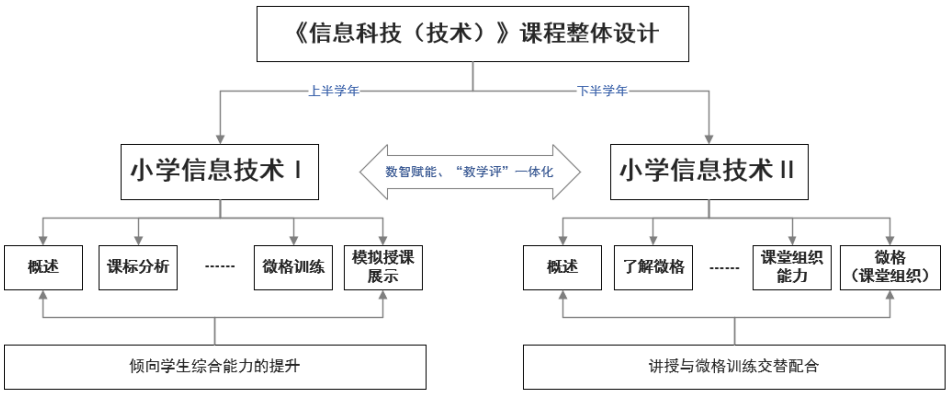


图 4 项目课程建设证明材料

3.3.3 课程教学后：

在课程学习结束之后，学生可以依据本节课学习到的内容对自己教学设计中对应的部分进行修改并提交到 AI 智能辅助教学平台上面，同时学生还可以在平台上对其他同学所提交的教学设计结合自己本节课学习到的知识提出修改性的建议，教师也可以在后台实时观察学生的互动情况，并对在课堂上没有来得及点评的同学的教学设计及时给予修改优化建议。在收到教师或者其他同学提出的修改建议之后学生可以结合自己学习到的知识对其他同学的教学设计进行多次优化修改，重新上传到平台中。

在下一节课的备课期间，教师可以结合学生提交的教学设计修改及时调整课程的设计和内容，同时可以依靠 AI 智能辅助教学平台中分析出的学生在不断优化中产生的过程性数据，发现学生的薄弱点，上传线上学习资料到学习平台给学生自行学习。

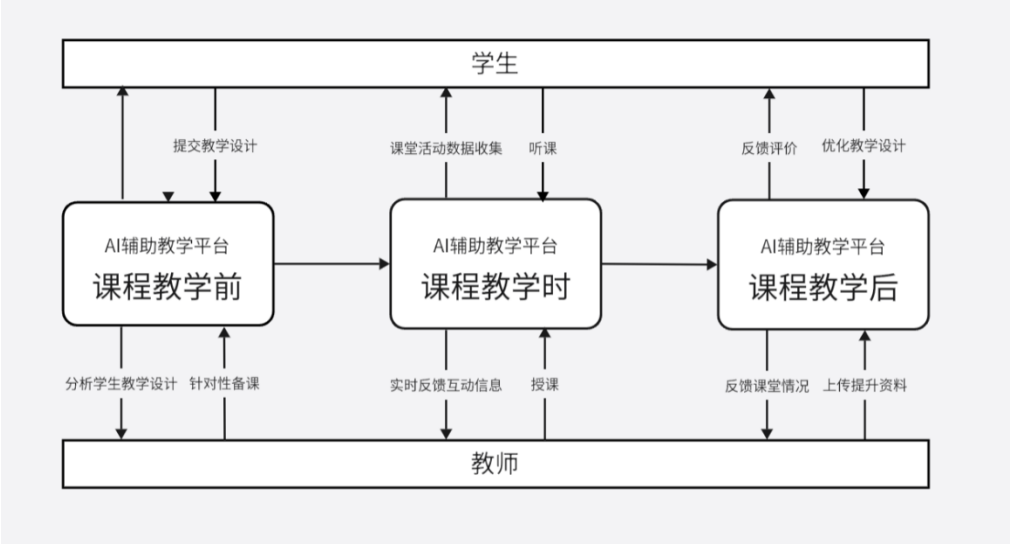


图 5 数智赋能师范生教学设计能力提升的教学评一体化课堂教学流程图

4. 小结

教学评一体化作为一种教学理念对师范生教学设计能力的培养进行优化提升，其主要目的是使师范类院校培养的师范生能够符合大数据时代下社会发展的要求，满足基础教育的需求。师范类学生培养的核心理念是“学生中心，产出导向，持续改进”，这一核心理念也是师范类专业人才培养工作的行动指南（李海英，2022）。

其中《小学信息技术教育》是教育技术专业师范生的教学法核心课程，该课程直接对标小学信息科技课程师资培养，其教学目标定位在现代教育观念与学习理论指导下，运用信息化教学设计方法，对中小学信息科技课程进行教学设计、资源开发、教学评价等综合实践，为师范生能够独立承担中小学信息科技课程的设计与教学打好基础，并且在其中怎样完成一个优质的教学设计是本节课的核心教学内容。现有的教学中，教师教学与学生学习、学生评价三者相割裂现象居多，教师缺乏“教-学-评”是一个整体、应该一体化去思考和谋划的意识。因此本研究基于“教-学-评”一体化的思想优化相应的课程体系，为信息科技教师人才质量的培养保驾护航，切实提高师范生实践能力。

参考文献

马莲姑、黄寿孟和贾晓婷（2022）. “互联网”时代师范生信息技术应用能力培养研究. 海南开放大学学报，2022, 23(04):127-133.

李毅、何莎薇和邱兰欢（2020）. 教育信息化 2.0 时代下师范生信息素养评价指标体系研究. 中国电化教育, 2020, (06):104-111.

任友群、黄荣怀和熊璋（2022）. 从信息技术到信息科技——关于《义务教育信息科技课程标准(2022 年版)》的对话. 课程. 教材. 教法, 2022, 42(12):21-31.

李海英（2022）. 师范类专业认证背景下师范生实践能力培养策略研究. 贵州师范学院学习报，2022，38(04):71-77.

朱旭东（2022）. 《强师计划》是基础教育高质量发展的强大引擎. 中国教育学刊，2022(09):1.

褚艳秋（2023）. 优化关键环节：让“教学评一体化”真实落地. 中小学管理, 2023(11):53-55.

Dewey J（1900）. Psychology and social practice[J]. Science, 1900, 11(270): 321-333.

McLaughlin, Milbrey W., & Lorrie A. Shepard（1995）. Improving Education through Standards-Based Reform. A Report by the National Academy of Education Panel on Standards-Based Education Reform. National Academy of Education, Stanford University, CERAS Building, Room 108, Stanford, CA 94305-3084., 1995

混合式教学模式下《学习分析技术与方法》课程设计与评价研究

——以《数据驱动教学》为例

Research on the Course Design and Evaluation of Learning Analysis Technology and Methods Under the Hybrid Teaching Model - Taking Data-Driven Teaching as an Example

蒲圻彬[1]，于慧敏[1]，张琪[1]
淮北师范大学 教育学院
* 1651614788@qq.com

【摘要】 随着教育领域数字化转型的持续推进，学习分析技术日益受到关注，成为提升学习体验和教育效果的关键。《学习分析技术与方法》课程应运而生，旨在全面系统地传授学习分析的理论知识与实践技能。本文对该课程的内容及其采用的混合式教学方法了深入剖析，以深化对其内涵与价值的理解。本文以该课程中的《数据驱动教学》为例，对教学活动进行设计，展示了混合式教学的具体实施，有效促进学生的深入理解和学习。经过对课程的全面评估，学生的成绩提升和积极反馈进一步验证了混合式教学模式的有效性。
【关键词】 学习分析；《学习分析技术与方法》；混合式教学；数据驱动教学

Abstract: *As digital transformation in the education field continues to advance, learning analytics technology has received increasing attention and has become the key to improving learning experience and educational effects. The course Learning Analysis Technology and Methods came into being, aiming to comprehensively and systematically teach the theoretical knowledge and practical skills. This article provides an in-depth analysis of the content and the blended teaching method used to deepen our understanding of connotation and value. This article takes Data-Driven Teaching as an example to design teaching activities and demonstrate the specific implementation of blended teaching, which effectively promotes students' in-depth understanding and learning. Comprehensive course evaluations, including improved student performances and positive feedback, affirm the blended teaching model's effectiveness.*

Keywords: Learning Analytics, *Learning Analytics Technology and Method*s, Blended Learning,Data-Driven Teaching

1. 前言

在数字化转型推动教育变革的当下，学习分析技术脱颖而出，成为一种关键的新兴技术，它利用大数据分析来优化学习体验和教育成果。2011 年，美国新媒体体联盟的《地平线报告》首次提出学习分析的概念（2022，陈明选）。本研究以“学习分析”为关键词在中国知网数据库进行检索，检索时间跨度为 2012 年至 2023 年，采用 Citespace 对其时间线进行分析（如图 1），关键词的出现频率和相互之间的关联性清晰地描绘了学习分析领域的发展轮廓。此外，还对学习分析进行聚类分析（如图 2），进一步揭示了该领域内的主题分布和演进路径。图 1 和图 2 提供了学习分析在过去几年中的演变和拓展的宏观视图，它们不仅展示了研究活动的量化增长，还揭示了学术讨论中的新主题和趋势（祝歆，2024）。而《学习分析技术与方法》课程自在中国大学 MOOC 上开设以来，备受好评。该课程通过对学习分析这一领域的深入研究，为混合式教学提供了理论与实践指导。2003 年北京师范大学何克抗教授将混合式教学 这一概念引入国内以来，国内学者对混合式教学展开了大量研究，认

为混合式教学是“把传统学习方式的优化和数字化或网络化学习的优势结合起来”（许缦，2024）。《学习分析技术与方法》这门课程的目标是构建一个课程数字化转型体系，以促进学生互动和深入学习的能力。该课程内容涵盖了数据收集、测量与分析，并对学习分析的理论基础、技术手段、实际应用、研究进展以及未来趋势进行了全面的阐释。作为一种数据密集型的研究方法，学习分析在教育领域的应用旨在提升个体学习者的绩效和成就，已经成为教育研究的热点和一个充满潜力的实践领域。该领域的探索并不局限于传统教育学，还包括信息科学、认知科学等多个学科的交叉融合，这不仅反映了教育数字化的新范式，也代表了新思维与新方法的兴起。学习分析的发展呈现出对教育质量学习分析的发展呈现出对教育质量和个体学习体验提升的持续承诺，预示着个性化和精准化教育的未来趋势（孔晶，2022）。

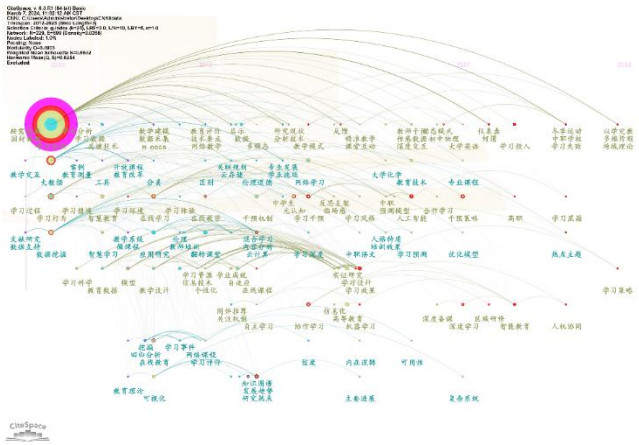


图 1 学习分析的时间线谱

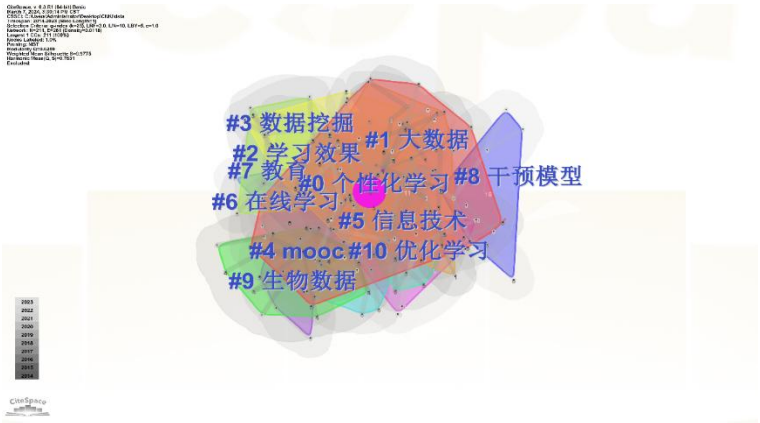


图 2 学习分析的关键词聚类分析

2. 课程设计

本课程从教育学、信息科学、认知科学等领域展开剖析。将学习分析模型、学习者画像、多模态学习分析、学习推荐系统、自适应学习等纳入课程内容，同时关注学习的认知机理、师生角色变化、技术与人文的融合、智能时代的数据伦理等要素。同时，实施“人在回路”混合教学本课程采用展和关键能力的培养为落脚点，进行“40%+60%”指向深度学习的混合教学，前置低阶学习目标，分配给以学生为主体、老师监督的线上学习，整合集中与分散、同步与异步学习优势，充分利用智能技术促进思维技能和元认知导向的深度探究（陈裕先，2022）。 课程的整体流程框架图（如图 3），清晰地展示了课程设计的宏观结构，从教学设计到教学实施的每一个环节都被详细规划。

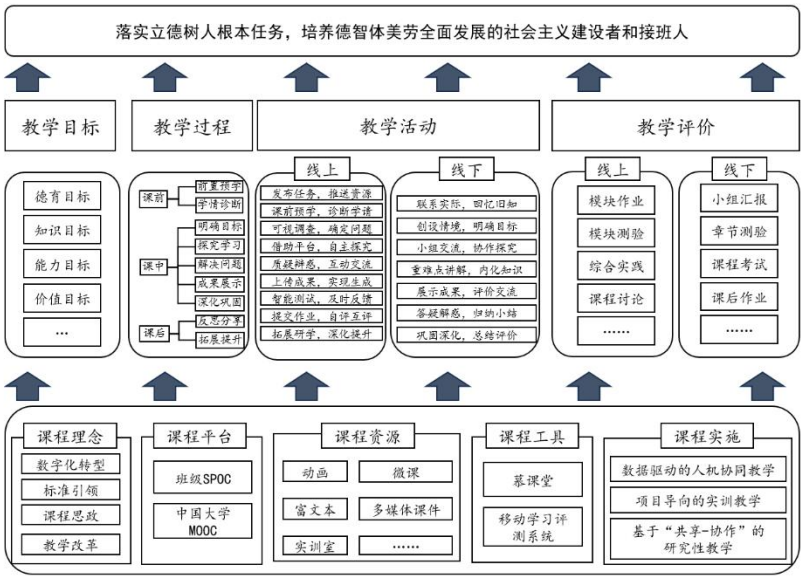


图 3 《学习分析技术与方法》整体框架图

2.1.混合式教学过程

课程采用“诊断改进的深度反思教学”即协作探究学习，运用项目导向的实训教学、基于“共享-协作”的研究性教学等完成进阶式教学任务，引导学生通过多重案例和多维视角思考；数据驱动精准教学，分步教学与综合练习结合，通过课堂提问和随堂测试，对学情数据综合分析，遴选典型问题；元认知反思教学，采用案例研究、虚拟社群等策略，在教学过程中加入反思、对话与综合评估，实现对学习分析理念、范式与应用的深度理解（冯玲玉，2021）。本课程采用混合式教学，将按照以下教学过程（如图 4）进行实施：

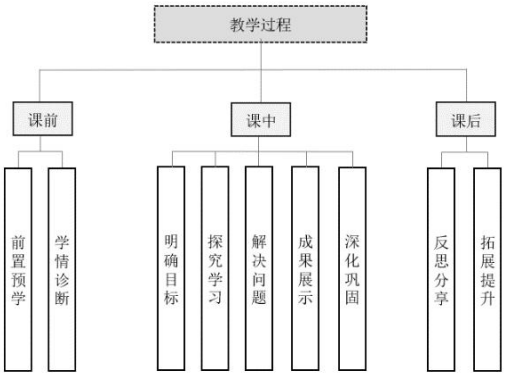


图 4 混合式教学过程

在课前阶段，前置预学，学情诊断。学生根据自主学习任务单的相关内容，利用网络学习平台上的相关资源开展自主学习，完成教师设定的任务，并将自主学习过程中遇到的相关困惑及建议提交至学习平台，形成课前自主学习反馈，教师以此了解学生学习情况。在课中阶段，明确目标、探究学习、解决问题、成功展示、深化巩固。教师明确教学目标，针对学生在任务完成过程中存在的共性问题，采用集中讲授或组织讨论的方式进行答疑解惑。在课中的典型任务探究阶段，学生可以根据不同的探究问题，采用自主探究或合作学习的方式来开展研究性的学习活动。学生对其成果进行展示和交流，教师不仅要对学生的学习成果予以点评和指导以达到深化巩固，引导学生反思在知识、技能上的收获，还要引导学生进行学习过程、学习态度、学习经验、学习方法等方面的反思和总结，并开展自我评价，建构自我意识。在课后阶段，反思分享、拓展提升。学生根据教师和其他同学的建议，修改、完善、提炼自己的学习成果和反思总结并提交至学习平台，以进行更大范围的交流和传播。

2.2.课程内容

课程内容包括学习准备、学习分析的概念与特征、学习分析支撑教与学变革、数据挖掘与建模、数据分析技术以及学习分析实践与趋势 6 个章节。教学内容采用“渐进分化、螺旋上升”方式进行整体设计。本课程将线上教学、面授讨论、同伴反馈环节等有机结合。课前组织学生完成自主学习任务，收集问题反馈（如图 5）；课中梳理主题主线，采用任务驱动、支架渐隐等方式，解析重点难点，精准概括提升。

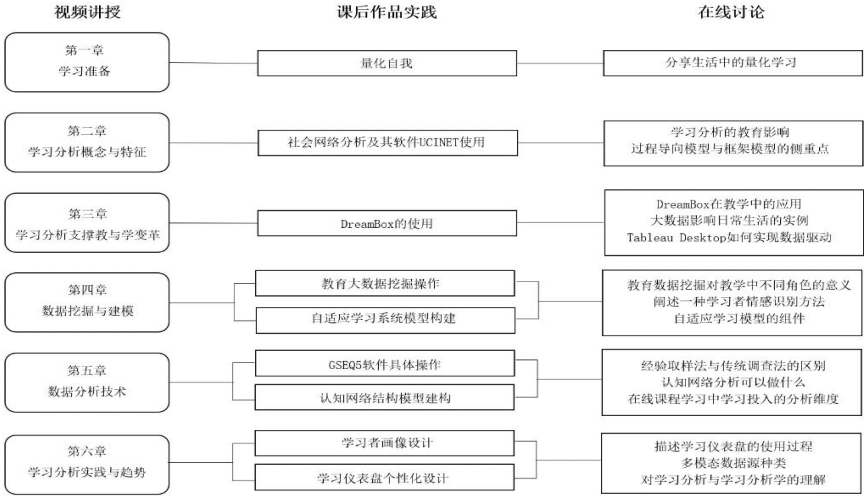


图 5　《学习分析技术与方法》课程内容体系

3. 活动案例设计

本研究以《数据驱动教学》为例，将其作为混合式教学活动设计的典型案例进行探讨。《数据驱动教学》为《学习分析技术与方法》第二章“学习分析的本质模型”第 1 节“学习分析的本质”的内容，该节内容在学习分析本质与模型中占据重要地位，学生通过理解与掌握数据驱动范式在教育教学中的应用，尝试运用数据来理解学习过程。在开放、协同的环境中，实现互动共生的深度建构，为学生提供更多理解教育技术并利用技术赋能教学的机会，对学习者的自主学习能力和合作探究能力都提出了较高的要求。

3.1.总体策略

采用线下课堂教学和线上课程教学相结合的混合式教学，该方法整合了传统的线下授课方式与先进的线上学习平台，旨在最大化教学效果与学习效率。在线上教学方面，以视频教学为主，学生利用碎片化时间，依据个人学习节奏自主学习相关内容并完成指定的课前练习题，此过程不仅促进了学生的自我学习能力，还为线下课堂互动提供了充分的预备。教师角色在本模式中转变为学习引导者和知识的桥梁。教师通过慕课堂线上备课，包括但不限于设计针对性的课堂练习题、构建高阶思维讨论主题，以及定期更新课程公告，确保学生在学习过程中的问题能够得到及时的回应和解决。在线下课堂的设置中，教师侧重于解决学生在线上学习过程中遇到的疑难问题，通过慕课平台发布的练习和互动环节，进一步加深学生对知识点的理解。此外，线下课堂也是引导学生探索新知识、培养批判性思维和解决问题能力的重要环节。本教学策略的设计，旨在实现教与学的最佳结合，通过线上与线下的互补，为学生构建一个全面、互动、反思的学习环境，同时也为教育工作者提供了一个有效的教学框架，以适应当前教育领域中持续变化的需求（张嫚晴，2024）。通过此种模式的实施，预期将能够促进学生综合素质的提升，以及对所学知识的深入理解与应用。

3.2 教学设计

在总体策略的基础上，教学过程的设计致力于将混合式教学的理论转化为实践，确保学生能够在线上自主学习和线下课堂互动中获得无缝的教学体验。通过细化每个教学环节，本设计旨在促进学生的主动学习，加深对课程内容的理解，同时提高教学策略的执行效率和教学质量，以下是具体的教学设计（如表 1）。

表 1 教学设计

一、教学目标： 1.识记数据驱动方法的概念，掌握大数据学习分析范式的特征，理解数据驱动范式在人工智能中的应用，辨别数据驱动教学与传统教学范式的异同。 2.通过自主探究，获取、梳理、概括大数据的特征和本质。小组合作探究，辨别数据驱动教学与传统教学范式的异同，理解数据驱动范式在教学中的应用，明确数据驱动范式在教育教学中应用的价值。 3.形成良好的教育数据意识与教育数据思维，提高运用数据解决教学实际问题的能力，遵守教育数据伦理规范。
二、教学重难点： （一）教学重点 通过自主探究，获取、梳理、概括大数据的特征和本质。小组合作探究，辨别数据驱动教学与传统教学范式的异同，理解数据驱动范式在教学中的应用，明确数据驱动范式在教育教学中应用的模式和价值。 （二）教学难点 理解数据驱动范式在教学中的应用，明确数据驱动范式在教育教学中应用的价值。
三、教学方法：讲授法 人机协同教学法 项目导向教学法
四、教学过程：
（一）导入环节（7 分钟） **活动 1：**在线上学习阶段，学生自主学习，教师在线答疑并讲解课前布置的习题。 **活动 2：**复习旧知 回顾和讨论大数据的本质及关键特征，如 3V 特征；互信息：用多种信息去衡量一个事物；交叉熵：比较信息间的一致性；黑天鹅效应：一些事一般不会发生，发生就会颠覆以前的认知。
二、探索环节（27 分钟） **活动 3：**引入新知 （1）教师发布学习视频让学生观看并小组讨论，自主探究大数据驱动的特征和价值。 （2）通过线上线下交互，随堂发布练习题 1，学生用手机答题，教师实时了解学生答题情况，了解学生掌握情况，教师总结学习分析范式的两个特征。 **活动 4：**情境创设，引导探索 （1）在模拟教学场景中，利用学习管理系统（LMS）数据进行学习者行为的分类分析，探讨数据驱动在定制化教学中的角色。 （2）学生在教师的讲解下，掌握学习者学习行为的基本类型。 （3）学生讨论数据驱动赋能因材施教，发布练习题 2，学生作答。 **活动 5：**合作探究，深入领悟 （1）团队协作探究人工智能中数据驱动的应用，包括机器翻译、图灵测试等案例。 （2）线上线下交互，观看视频资源从语音识别到“鸟飞派”案例，理解数据驱动范式在人工智能中的应用。明确机器智能最重要的是能够解决人脑所能解决的问题，而不在于是否采用和人脑解决问题一样的方法。
三、应用环节（16 分钟）（ **活动 6：**翻转课堂，融会贯通 （1）实施翻转课堂，围绕数据驱动教学的案例如 Course Signals、D2L 和 Knewton 平台

进行讨论，发布练习题 4。 （2）教师讲解自适应平台，该平台会依据学生学习痕迹做数据分析，为学生做推荐，哪种方式对学生学习更有效，符合个性化学习。 **活动 7：**教学决策，总结升华 （1）通过线上与线下交互，学生利用手机观看 DDI 演进案例，发布练习题 4 让学生思考对三种教学范式的理解以及讨论观看视频后的感受。 （2）通过讨论翻转课堂与数据驱动的关系，促进学生对教学模式的深入理解，增强批判性思维和创新能力，注意“以学定教 因学定导”。
四、总结环节（6 分钟） 1.教师带领学生回顾数据驱动决策的关键点，推送练习题 5，并总结“数据驱动教学”的核心理念。 2.通过平台发布课后作业，强化学生对数据驱动教学理念的理解和应用。

3.4. 教学反思

教学达成方面，在丰富技术支持的智慧教室环境下，教师利用人机协同教学法、项目式教学法等多种教学方式，充分利用资源、工具，构建师生互动环节，开展多元化教学，将学生作为学习的主体，充分调动学生学习的积极性与参与度，达成学习目标，提升了学生素养。但同时也应注意，线下教学时知识点剖析应更加精炼，调控学生讨论时间分配问题，学生线上学习质量还可以有所提升。

4. 课程评价

《学习分析技术与方法》这门课程采用混合式教学模式，在现代教育教学实践中展现出其显著优势。该课程对学生学习成绩采用取了综合评分方法：平时成绩占 20%，线上自主学习成绩占 20%，期末考试占 60%。具体来说，平时成绩包括慕课堂互动（30%）、小组测评（20%）、课后作业（20%）、以及期末综合作品（30%）；线上自主学习成绩则由 MOOC 课程讨论（20%）、模块测验（30%）、模块作业（20%）、以及在线期末考试（30%）构成。在教育技术专业的教学实践中，本研究探讨了混合式教学模式对大二学生学习成绩的影响。针对两个不同年级的学生群体——2020 级学生采用传统教学模式，而 2021 级学生实施了混合式教学策略。

表 3 2020 级和 2021 级学生成绩描述统计

	班级	个案数	平均值	标准差	标准误差平均值
成绩	2020	61	80.66	4.672	.598
	2021	65	83.98	5.481	.680

表 4 2020 级和 2021 级学生成绩独立样本检验

		莱文方差等同性检验		平均值等同性 t 检验						
									差值 95% 置信区间	
		F	显著性	t	自由度	显著性（双尾）	平均值差值	标准误差差值	下限	上限
成绩	假定等方差	7.165	.008	-3.657	124	<.001	-3.329	.910	-5.130	-1.527
	不假定等方差			-3.676	122.889	<.001	-3.329	.906	-5.121	-1.536

初步数据分析揭示（如表 3 和表 4），2020 级学生的平均成绩为 80.66（标准差=4.672），相较之下，2021 级学生的平均成绩提升至 83.98（标准差=5.481）。深入分析采用独立样本 t 检验方法，结果显示显著性差异（F=7.165，p<0.05），这表明混合式教学对学生成绩的提升有显著影响。这一统计学证据支持了混合式教学相较于传统教学模式，在提升学生学习成绩方面的有效性。这表明混合式教学对学生成绩的提升有显著影响。独立样本检验结果进一步验证了这一发现，说明采用混合式教学的 2021 级学生相比于 2020 级的学生，在学习成绩上有了明显的提升。

通过有机结合线上和线下的教学方法，课程大大增强了教学的灵活性与学生的学习体验。根据学生评价的词云分析（如图 6）显示，关键词如“有趣”、“实用”和“互动”频繁出现，反映出学生们对课程内容的丰富性、教学风格的互动性、以及对学习感受的积极性给予了高度评价，这些都在加强他们的知识吸收和专业发展方面起到了重要作用。



图 6 学生评价词云

在对学生评价进行社会网络分析后发现课程评论文本围绕的五大主题（如图 7）：课程内容、教学风格、学习感受、知识量和专业发展。分析结果明显显示，这门课程有效地促进了学生的深度学习能力，激发了他们对课程的浓厚兴趣，并显著拓宽了他们的学科视野，同时提升了他们将知识应用于实践的能力。网络图中密集的连接和词云图中的关键词展现了学生评价的多个方面，揭示了他们对课程的全面认可和积极反馈。这些数据的可视化不仅显示了课程内容与学生专业成长的紧密联系，也证明了混合式教学模式在促进学生深度学习能力方面的卓越表现。

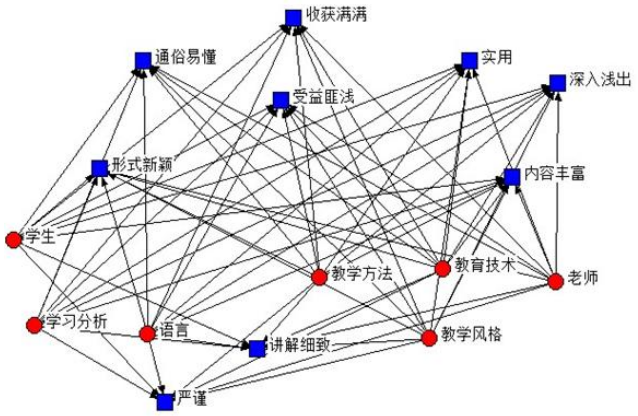


图 7 学生评价社会网络分析

混合式教学的设计兼顾了自主性线上学习与面对面的教师引导互动，提供了个性化的学习途径，不仅激发了学生的学习热情，也加强了他们对知识的应用和理解。从学生的反馈来看，他们对这种互动性强且实用性高的教学方法表示赞赏，并认为这种教学模式增进了他们的知识和技能，为未来的职业道路打下了坚实的基础。因此，混合式教学确实能够提高学生的学习成效和满意度，这一教学方法在当前和未来的教育领域中具有广泛的应用前景。

5. 结语

学习分析技术通过深入了解学习者的需求，优化了教学内容和方法，从而推动了学生的深度学习和个性化发展。通过《学习分析技术与方法》课程的混合式教学实践，既见证了数字化转型在教育领域中的深刻影响，也体验了学习分析技术如何有效融入教育实践，优化教学过程，提高学生的学习效能。课程设计巧妙结合了线上自主学习和线下教师指导，实现了知识传授与能力培养的双重目标，加强了学生对教育科技的理解，并有效促进了学生自主与合作探究能力的提升。

学生的积极反馈，以及通过词云和社会网络分析揭示的评价主题，证实了混合式教学在提升教育效果方面的显著成就。学生们的学习兴趣、学科视野的拓宽以及知识应用能力的提升，均昭示了这种教学模式深厚的教育价值和广阔的未来应用潜力。同时，这也指向了未来教育的发展趋势——更加个性化、数据驱动且以学生为中心的教学模式，这将引领学习者走向更高效、更具创新性的学习境界。

综上所述，混合式教学不仅是一种教学模式的变革，更是教育理念和教育方法创新的体现。随着技术的不断进步和教育需求的不断演变，我们有理由相信，未来的教育将更加多元化、个性化、智能化，为每一位学习者提供定制化的学习路径，实现其潜力的最大化释放。

参考文献

陈明选, 周亮 & 赵继勇. (2022). 学习设计与学习分析的联结：现状、挑战与实现路径. 开放教育研究 (06), 27-36.

孔晶 & 刘家亮. (2022). 学习分析支持元认知培养的学习设计理论与实证研究. 开放教育研究 (03), 102-110.

许缦. (2024). 新文科建设背景下基于 BOPPPS 教学模式的管理类课程混合式教学设计——以“企业战略管理”课程为例. 黑龙江教育(理论与实践) (03), 36-41.

冯玲玉, 甄宗武 & 虎二梅. (2021). “以学习活动为中心教学设计”视角下的混合式教学机理分析. 电化教育研究 (11), 100-106.

张嫚晴. (2024). 混合式教学模式下学生学习评价分析——以“基础会计”课程为例. 新课程研究 (03), 19-22.

祝歆, 楚心茹 & 丁庆洋. (2024). 混合式教学研究现状及趋势——基于 CiteSpace 可视化分析. 北京联合大学学报 (01), 13-23+41.

陈裕先. (2022). 基于 SPOC 理念的混合式教学模式分析与设计——以数据库技术课程为例. 新余学院学报 (03), 105-111.

多模态学习分析：理论基础、融合策略与多维数据案例

Theoretical Basis, Fusion Strategies, and Multidimensional Data Cases of Multimodal Learning Analysis

马宇慧[1]，李璇[2]，蒲圻彬[3]，张琪[4*]
[1234]淮北师范大学教育学院
*zqzqhata@sina.com

【摘要】智能感知技术的发展为学习分析领域的变革提供了跨数据支持，实现智能技术与教育研究的深度融合。该文立足于智能时代“多模态学习分析”的本质，从认知理论、情感理论等剖析多模态学习分析的实质，比较了三种融合方法，结合来自视频记录、皮肤电活动的多维数据分析自我调节学习的各阶段的交互演变过程，阐释了对应的三个分析案例，以实现多模态学习分析与真实教育场景的深度融合。
【关键词】多模态学习分析、理论基础、融合策略、数据表征

***Abstract*:** *The development of intelligent perception technology provides cross-data support for the change in the field of learning analysis, and realizes the deep integration of intelligent technology and educational research. Based on the essence of "multimodal learning analysis" in the age of intelligence, this paper analyzes the essence of multimodal learning analysis from cognitive theory and emotion theory, compares three fusion methods, combines the interactive evolution process of each stage of self-regulation learning with multidimensional data analysis from video recording and skin electrical activity, and explains the corresponding three analysis cases. To realize the deep integration of multi-modal learning analysis and real education scene.*
Keywords: *Multimodal learning analysis, theoretical basis, fusion strategy, data representation*

1.前言

认知是理解人类如何学习的基础。认知不仅是单纯的大脑生物活动，更贯穿学习者“感知—行动”始终。利用多模态学习分析技术构建面向个体学习者的全空间、多场景、时序性的建模分析，使用多模态学习分析的跨模态感知与推理有助于建立语义空间的联结，纳入结构化知识来帮助计算机深度认知（张 琪，李福华，孙基男，2020）。认知理论是学习投入多模态数据建模的重要理论依据。除了认知维度外，认知、情感和社会活动在学习投入中被外化，形成相互依存和动态适应性系统。多模态学习投入的直接目标是以瞬时方式来考量学习者的认知、情感和社会状态，并分析这些过程的相互作用，以此进一步了解学习的本质。结合教育学、心理学、脑科学、认知神经科学的相关研究内容对学习者内部潜在特征进行系统解释，探究学习者“认知发展状况—行为表达模式—情感发生机制—学习交互模式”的协同进化机理，在更深层次上探究学习者的学习发生机制。

2.理论基础

2.1. 认知理论

多模态学习研究中使用最广泛的理论是来自认知科学的理论框架，尤其是具身认知理论。具身认知的核心观点是认知基于感知和行动，即使对于离线认知（offline cognition），在没有相关环境支持下发生的认知活动也是如此。由此产生的具身学习涉及通过感官、运动和内

省进行感知，通过协作进行语言和知识学习，能够激活大脑参与，有助于吸收更好的知识（Clark J, Crandall P G, Pellegrino R, et al. 2019）。

具身认知理论揭示了通过可观察的生理响应推断不可观察心理状态的机制，是多模态数据表征学习投入的重要理论依据。其中一些生理反应是无意识的，如避让、情绪传染等；另一些生理反应则具有明确的指向性，如元认知反射与主观感受。现有研究阐明了可穿戴设备和多模态数据源的价值，包括眼动追踪、脑电和皮肤电导数据等。这些传统方式难以获取的信息对于学习情感、行为倾向、压力水平、认知负荷等具有直接的解释力，如高学习投入期间交感神经系统中的刺激增加。面部表情和情感状态之间也存在明确的关系，如思考过程中学习者的皱眉，达成任务时眉毛扬起。已有研究确定了一些生理特征与认知状态的相关性，例如，眼动特征与注意力的联系，以及脑电与认知负荷的关系等。随着便携脑电、脑机接口、自然用户界面与非接触式伴随采集技术的成熟，更有助于自然、深入、全面地分析学习投入的过程与机制。

另一个重要的认知理论是认知负荷理论。认知负荷理论主要依赖于以容量有限的工作记忆为核心的认知结构的支持。在教学设计和学习过程中，应尽可能减少外部认知负荷，增加相关认知负荷，以促进学习学习活动的达成，并体现高学习绩效。

2.2. 学业情绪的控制 ——价值理论

除认知理论外，多模态学习分析普遍关注情感理论，尤其是学业情绪的控制——价值理论（Giannakos, M., & Cukurova, M. 2023）。其中包括学习环境、控制-价值评价、成就情绪和学习成就四个要素。并将成就情绪按照效价（积极或消极）、生理唤醒（激活或未激活）和关注点（与活动相关或与结果相关）3 个维度进行分类。该理论的中心观点是学习过程中成就情绪的产生或唤起有两个核心前提，即行为与结果间的因果感知控制以及其价值感知。

3.三类融合方式比较

沃斯利描述了三项多模态学习分析的典型案例（Pekrun R. 2006），每项研究对应相应的融合模式。研究群体为人文科学专业的学生，探索其工程设计背景下的专业投入水平。学生先前的经验用于界定新手、中级或专家。多模态数据包括Kinect传感器采集音频、手势和视频，高分辨率网络摄像头记录学习过程，皮肤电传感器用于测量压力与唤醒水平，传感器数据均通过采集软件同步，提取如下特征并计算每种模态特征的最小值、最大值和平均值。

音频：网络摄像头和Kinect音频通道组合，基于Sphinx语音识别工具包开发，任一通道在给定时间内检测到语音，则认为音频存在。手部运动数据：由Kinect生成，存储十二个上身关节的三维数据。使用左手腕、右手腕、手和肘部数据点，计算左手腕和左手、左手腕和左肘、右手腕和右手、右手腕和右肘连接向量的角位移。皮肤电信号：收集皮肤电基线数据。

3.1. 基本融合

基本融合直接整合各模态的特征，不考虑特征之间的相互作用。分析时通过特征约简算法完成特征选择。基于特征训练无监督模型，比较不同融合与建模算法的差异性，可以对建模过程进行更好优化。

3.2. 低阶融合

利用低阶融合方法，在较小的时间尺度上进行多模态数据融合，前提是研究者事先知道各模态在时间上的特定关联。低阶融合建立在认识论框架的基础上，根据多模态的行为分析学生的认知过程。在工程设计场景中存在两种不同方法：一是基于实例的推理，即使用现实世界中的实例作为解决任务的切入点；二是基于原理的推理，即将工程原理作为设计的基础。

3.3. 高阶融合

高阶融合与低级融合类似，不同之处在于多模态数据融合发生在不同的尺度上。高阶融合首先进行语义分析，试图理解这些数据。如不同类型的手势以不同的方式表达具体的知识，手势不仅在解释中产生，当学习者用语言表达想法时，手势无处不在（Worsley M. 2014）。

麦克尼尔进一步界定了四类手势（Sebeok T A, Umiker-Sebeok D J. 1995），包括：指向手势，象征手势，隐喻手势，节拍手势。上述手势可以进行组合构成更广泛的表征性手势类别。基于上述研究，阿利巴利等阐释了不同类型手势与认知过程的关系（McNeill D. 2008）。其中，指向手势反映了认知在物理环境中的基础；表征手势体现了对行为和感知的心理模拟；隐喻性手势反映了基于身体的概念隐喻；指向手势通常与语音一起使用，体现了说话者将语音内容定向到物理环境中的相应对象位置。

根据上述理论，可以假设言语和皮肤电反应更多与特定类型的手腕运动有关，而非位移量。分析特定类型的手势学习者语言出现时机，将学习行为界定为6种类型，如表1所示。

表1 学习行为分类

计划	原型设计想法或检查材料
评估	测试机制或测试系统
修改	对现有设计进行修改
无关	不主动参与活动
恢复	撤销先前设计的一个或多个部分
实现	将部件组合在一起，形成结构

利用除手腕位移之外的多模态数据，根据表1的分类进行数据融合，具体分析过程与低阶融合相同。研究发现，虽然高级融合中使用的算法与低级融合相同，但改变数据融合的方式会极大地影响数据类型和分析维度。这种方法可以分析不同场景学习者操纵过程的语义运动，提供了预测学生学习困难的最有效方法。综上所述，基本融合作为基本的探索性方法，能够在统计水平上进行分析，为解开多模态数据以及构建分类器提供了切入点。低级融合提供了一种更密切考虑每个参与者过程时间元素的手段，可以在一定粒度上分析该数据与其他模式数据之间的对应关系，采用低级融合方法可以提供相当丰富的信息，提供跨越多个时间尺度的见解。高级融合试图从原始的低级数据中提取额外的语义，可以正确地将低级融合的模态正确地进行归类，具有很大的潜力，能够深入了解语义行为或行为是如何产生的。

4.基于多维数据的心理网络分析

结合来自视频记录、皮肤电活动的多维数据分析自我调节学习各阶段的交互演变过程。基于上述视角，研究阐释了对应的三个分析案例。教育场景为物理课程中协作学习过程。学生以三人一组的形式进行协作学习，任务包括设计实验来测量光速和头发厚度，并使用激光、镜子、透镜、棱镜和双分光镜进行动手实验探索反射、折射、色散和干涉。课程全程视频录制，学生佩戴智能腕带测量皮肤电导反应。使用定性编码方式对视频进行分析，分析侧重于个体不同调节阶段相关的言语与行为，包括监控计划和目标设定，监控任务理解，监控任务实施，监控评估，以及监控动机和情绪5种维度。此外，协议还包括与学习监控相关的动作和手势。传感器收集的学生皮肤电数据，使用基于MATLAB进行处理，通过连续分解分析检测皮肤电导反应（SCR），返回检测到的峰值的开始时间和幅度。

4.1. 群组水平分析——监测事件和 EDA 峰值之间的相互作用

利用跨主体网络探索监测事件和EDA之间的关系，分析监测事件之间的相互作用以及如何相互影响。简言之，监测事件集中在任务定义（TD）、目标设定和规划（GP）、任务实施（TE）、评估（EV）、动机和情感（ME）和表示同意（YES）阶段。

使用横截面网络研究监测事件和EDA峰值（即SCR）之间的关系。汇总每个学生的每个

动作的频率，生成无向网络。图1呈现了四组受试者之间的平均静态网络，显示了彼此在时间上相互依赖的事件。连接节点的蓝线表示与监测目标有显著关联，从一个节点到另一个节点的红线表示有显著的负关联，线的粗细对应于关联的强度。

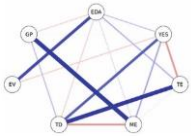


图 1 群组监测事件与皮肤电活动（EDA）之间的关联

上图证明了不同监测事件之间的关联，以及EDA和观察到的监测事件之间的关联。其中，TD与学生 EDA 之间存在微弱的负相关，而TE和YES与 EDA 之间存在微弱的正相关。此外，EDA 与EV之间存在强烈的正相关。根据图中的监控关联， GP与ME之间尤为强烈。

4.2. 个体水平分析——监测事件和 EDA 峰之间的顺序相互作用

该案例表明3名学生的个人同期网络，个体水平分析揭示了个人网络的异同。当关注三个学生（学生1、2、 3）的同期网络时，其网络出现了相似性（图2）。对于每个学生，EDA与同意呈负相关， EDA与TD、ME和TE却呈正相关，只有学生1的 EDA和GP呈正相关，表明与其他两个学生相比，他更频繁地制定计划。此外，与学生2相比，学生1、3在TE和YES以及TD和YES之间有更强的关联，表明前两个学生在监控方面更活跃。

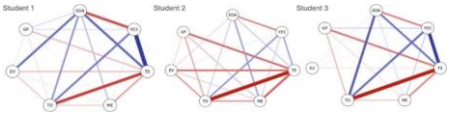


图 2 个体监测事件的同期网络分析

参考文献

张 琪，李福华，孙基男（2020）。多模态学习分析：走向计算教育时代的学习分析学。中国电化教育（09）, 7-14。

Clark J, Crandall P G, Pellegrino R, et al. 2019. Assessing smart glasses-based foodservice training: An embodied learning theory approach[J]. Canadian Journal of Learning and Technology, 45(2): 1-19.

Giannakos, M., & Cukurova, M. 2023.The role of learning theory in multimodal learning analytics[J]. British Journal of Educational Technology, 54: 1246–1267.

Pekrun R. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice[J]. Educational Psychology Review, 18(4): 315-341.

Worsley M. 2014. Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors[C]. Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge: 1-4.

Sebeok T A, Umiker-Sebeok D J. 1995. Advances in Visual Semiotics[M]. New York: Mouton de Gruyter.

McNeill D. 2008. Gesture and Thought[M]. Chicago: University of Chicago Press.

Alibali M W, Nathan M J. 2012. Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures[J]. Journal of the Learning Sciences, 21(2): 247-286.

結合動態情節互動鷹架之行政倫理數位敘事教育遊戲之設計與評估

Evaluation of a Digital Narrative Educational Game on Administrative Ethics Combined with Dynamic Scenario Interactive Scaffolding

劉書瑋[1]，倪紳煬[2]，侯惠澤[3,4*]
[1]臺灣科技大學應用科技研究所
[2]臺灣科技大學應用科技研究所
[3]臺灣科技大學應用科技研究所
[4]臺灣科技大學技職賦能研究中心

[*] hthou@mail.ntust.edu.tw

【摘要】 行政倫理在公共服務體系中一直佔有相當重要的地位，但是在行政倫理的教學上，卻未受到足夠的重視。本研究結合動態情節鷹架之機制，發展出一套學習行政倫理的數位敘事教育遊戲，讓學習者在擬真的公務機關與事件中，認識行政倫理的知識及概念，試圖改善現有行政倫理教學的互動性與學習支援問題，並促進他們的批判思考傾向。初步結果指出此遊戲能成功增進學習者的知識，並能維持高度的心流狀態，同時擁有較低的焦慮情況，也提升了他們在活動前至活動中的動機，亦能產生明顯的批判思考傾向，證明此遊戲的學習效益。
【關鍵词】 行政倫理；遊戲式學習；動態情節互動鷹架；數位敘事

Abstract: *The administrative ethics have been playing an important role in governance of the country, but it is not given enough attention in its instruction. This study developed an educational digital narrative game for learning administrative ethics by integrating the mechanism of dynamic scenario scaffolding, allowing learners to recognize the knowledge and concepts of administrative ethics and participate in the progression of ethical events in a simulated public administration, trying to improve the interactivity and learning support of the existing instruction of administrative ethics, and to facilitate their tendency to think critically. Preliminary results indicated that the game was successful in increasing learners' knowledge of administrative ethics, and maintaining a high flow state with low anxiety, increasing their motivation from pre-activity to in-activity, and generating a significant tendency to critical thinking, which demonstrated the learning benefits of the game.*

Keywords: administrative ethics, game-based learning, dynamic storytelling scaffolding, digital narrative

1. 研究動機與目的

行政倫理一直是公共服務及政治最注重的議題之一，其關乎到正義、公平與平等的概念（Chapman, 2019）。但是行政倫理的教育，卻長期未受到重視。全球多數國家在公共關係中，並未將重點放在倫理的教育上，亦缺乏對教學方法的討論（Austin & Toth, 2011）。其中行政倫理的教學環境，又尤為重要。在不同的學習環境下，行政倫理的學習可能產生正向或負向的影響（Menzel, 1997）；行政倫理的教學，需考量到學習者的互動與回饋（Ni, 2013）；能否使學習者產生對倫理的判斷及思考能力，是行政倫理的學習關鍵（Hejka-Ekins, 1988; Zavattaro, 2022）。再者，就教學的內容來說，我們也須注意到倫理事件的發展，經常隨著時

間而有所變化（Cooper, 2012）。因此環境能否讓學習者產生批判性思維，注意到時間與事件情節的推移，並提供足夠的互動與支援，便是行政倫理教學可能需要注意的部分。

而搭配情境的遊戲，被認為是可能的解方之一。情境式學習遊戲可以讓學習者在模擬的現實環境中解決問題，除能提升學習者的參與程度及學習成效之外，也能讓他們擁有更多的批判思考傾向，並做出正確的決策（Chang et al., 2020; Hwang & Chang, 2023）。若適時搭配鷹架的輔助，則更能協助學習者進行思考，維持心流並使他們解決問題（Chen & Wu, 2023; Hou & Keng, 2021）。因此本研究試圖發展出一套建立於虛擬線上環境中的行政倫理教育遊戲，強調遊戲中的情境，同時將時間與事件情節的發展納入鷹架當中。而為了解此遊戲在學習上之效用，研究亦量測學習者的心流與焦慮狀態，以及不同時間點的動機狀態。其能觀察此遊戲是否能讓學習者產生積極的行為，並確保他們能持續進行遊戲（Lin & Hou, 2023）。本研究同時也觀察學習者的批判思考傾向，了解此遊戲能否讓學習者對倫理產生更多的思考，並找出此種教學方式的優缺點。故本研究之研究目的，主要為了解學習者在此遊戲的學習成效、心流狀態、動機狀態、焦慮狀態及批判思考傾向情況。

2. 文獻探討

2.1. 行政倫理教育之內涵與挑戰

倫理教育首要的目標，為提高學習者對於倫理的判斷與思考能力（Hejka-Ekins, 1988）；或許可利用虛構或真實的案例，來鼓勵學習者推理並促進他們的批判性思維（West, 2021）。近年有關行政倫理的教學研究，亦有不少採用案例的方式將知識傳授給學習者。Meyer（2021）曾運用電視節目於研究所的行政倫理課程，認為貼近生活與日常經驗的教學方式，能引發更多討論並讓課堂更有趣；Zavattaro（2022）則透過電影案例傳遞行政倫理的概念，認為產生教學效益的關鍵在於讓學習者產生批判性思考。然而在教學時，由於各地的倫理習慣不同，仍須尋找適合當地的文化方法來教學（Austin & Toth, 2011）；倫理為一種持續的過程，教學時應該注意到事件的進展，情況時常隨著時間推移而有所改變（Cooper, 2012）。再者，能否提供足夠的互動與回饋，也是值得關注的焦點之一（Ni, 2013）。

2.2. 教育遊戲在行政倫理教育的潛力

為了解決互動性的問題，遊戲式學習被認為是具有潛力的教學方式。遊戲式學習指設計遊戲來讓學生在遊戲中學習，將具有娛樂性的互動與學習活動相互結合（Prensky, 2003）。透過有趣與動態的內容，將使學習者提升動機並減少焦慮，同時增進學習興趣與積極程度（Hung et al., 2014; Pho & Dinscore, 2015）。若加入情境式學習的概念，讓學習者在擬真的遊戲空間探索，更是使他們在解決任務的過程中成功吸收知識，甚至提升批判思考的能力（Chang et al., 2020; Hwang & Chang, 2023）。近年有關行政倫理的研究，亦曾透過遠距或遊戲的方式，提升學習者的知識。Ni（2013）認為相較於傳統面對面課程，在線上學習公務行政的課程，可以有效消除地域障礙並提升互動性與回饋程度，並增加內向學習者的參與程度。Exmeyer 與 Boden（2020）則指出數位遊戲可以增加公共服務領域學習倫理的娛樂性，並能作為有效的學習工具，而擬真的學習環境則能提升沉浸感，並促進對自我、關係與環境的反思。因此我們認為應用情境式學習的遠距學習遊戲，在行政倫理的教育上具有潛力。

2.3. 鷹架設計

不過設計教育遊戲，仍須注意支援與回饋的機制。結合適合的學習策略或工具，將學習內容融合至遊戲當中，才能提升學習者的學習效益，並促使他們持續互動與反思（Cheng & Su, 2012; Chang & Hwang, 2019）。其中經常使用的策略或工具，即為「鷹架」。鷹架的概念來自於 Wood 等（1976），指的是由教學者從旁給予協助，讓學習者的能力逐漸增加，再隨程度逐漸拆除。其能引導學習者的探索，使其有目標可依循，以維持學習者的學習樂趣（Chen &

Law, 2016; Ch'ng et al, 2020）。而若與行政倫理教育相結合，則須更注意回饋與事件情節的推移，數位敘事（digital narrative）的概念可能是可行的方法之一。其強調系統、使用者、流程與輸出，能作為鷹架讓學習者參與其中，同時獲得知識並產生共鳴（Gil & Sylla, 2022）；其屬於任務導向類型的活動，結合敘事、遊戲和角色扮演體驗，並能提供人物發展、時間順序與多種結果的敘事方式（Roth & Koenitz, 2016）。更進一步的，其能提供事件當中的因果連結，能透過時間與空間解釋框架（Murray, 2018）。因此，若將其概念納入鷹架的設計當中，作為「動態情節互動鷹架」讓學習者與劇情互動，同時隨著事件進展或學習者的選擇，給予不同的回饋與支援，則更符合行政倫理的學習需求，使其成為完整並具效益的學習遊戲。

3. 研究方法

3.1. 研究對象與程序

本研究共招募 24 位學習者，全數皆為網路招募，年齡分布於 18 歲至 60 歲之間，性別則為 7 位男性與 17 位女性。所有實驗活動均為線上舉辦，在實驗開始之前，所有學習者均被要求簽署詳細說明參與者權利與保護措施的知情同意書。

實驗開始時，學習者將在 Google Meet 中集合，並填寫知情同意書與學習成效前測問卷，時間約 15 分鐘。填寫完畢之後，學習者將前往 Gather Town 開始行政倫理學習遊戲「正直與清廉都回來了」，依據任務指示前往公務機關任職，在閱讀文件以及與同事交談的過程中，學習有關行政倫理的知識，同時偵查該機關是否有違反行政倫理的事件產生，並做出結論及判斷，總體時間為 45 分鐘。遊戲結束之後，學習者將填寫後測量表及問卷，包含學習成效後測問卷、心流量表、活動焦慮量表、學習動機量表與批判思考傾向量表，時間約 20 分鐘。

3.2. 教學遊戲設計

為了解結合動態情節互動鷹架之教育遊戲對學習行政倫理知識的影響，本研究設計出一套教育遊戲「正直與清廉都回來了」，學習行政倫理中有關工作態度、行政處理、生活態度及人際交往等四項原則。其運用 Gather Town 作為主要遊戲平台，亦結合 Google 表單作為遊戲 NPC，依據現有公務機關之空間與人力配置，於平台中建構出擬真的遊戲場景與對話情境，同時亦會隨著故事情節的進展發生變化。而學習者則化身為機關職員參與運作並與同事互動，使遊戲在故事脈絡、場景與操作等方面，皆有高度的擬真，並期待學習遷移的產生。

遊戲舞台設定於台灣某虛構的公務機關，學習者在遊戲中為飾演一位新到任的公務人員，其為了保護自身與家人的廉潔，決定花上一段時間，仔細審視機關中的各個細節，包含同事之間的互動與公務運作的各個文件，找出任何違反行政倫理的線索，並依據所蒐集到的事證，做出相對應的行動，而遊戲亦將依照其所採取的行動，給予相應的好壞遊戲結局。

在遊戲開始之前，學習者將在教學關卡中了解遊戲的基本操作，接著再正式進入遊戲場景，在家人 NPC 的引導下，獲得行政倫理的基礎知識並前往公務機關任職，與多位同事及主管 NPC 對話，同時使用遊戲中的各個設施，尋找各個有關行政倫理的線索，以完成遊戲任務。其任務依據事件進展並結合「動態情節互動鷹架」推進，一共分為三個階段（時間點），各個階段的場景相同，但人物的對話或設施，會隨著時間推展或學習者的互動選擇而有所改變。第一個階段為剛到職，可學習到基礎的倫理知識，而若運用技巧與同事互動，選擇到關鍵的對話選項後，則可隱約察覺該機關似乎具有問題，學習者須回報正確的推理，才能前往下個階段；第二個階段為學習者到職約一個月，一開始學習者將接受主管臨時指派的任務，與主管對話後劇情才會繼續推進，隨著情節推展並與同事交談過後，學習者將逐漸深入理解行政倫理的概念，並發現機關部分的行政程序有問題，但由於缺乏最直接的違法證據，僅能擬定因應措施為下個階段做準備，回報正確的措施後即可前往下個階段；第三個階段為到職後約兩個月，學習者須搜尋到最關鍵的線索，讓整體事件水落石出，指出違法者的錯誤倫理行為，

並向政風處回報，此階段被設計成僅能回報一次，故學習者須審慎思考，以獲得最好的遊戲結局。遊戲除與動態情節互動鷹架結合之外，亦設置有輔助機制，協助學習者進行遊戲，並成功學習行政倫理的知識，詳細可見「鷹架設計」一節。此遊戲之遊戲畫面如圖 1 所示。



圖 1 遊戲畫面

3.3. 鷹架設計

為了讓學習者順利完成遊戲活動，並增進遊戲內的情境效果，本遊戲設計了多維度的鷹架機制，以協助學習者獲得關鍵線索並成功推進任務，提升對於行政倫理的知識，同時加深遊戲情境因素，提升遊戲的擬真程度。遊戲內包含認知鷹架、程序鷹架、情意鷹架，以及動態情節互動鷹架等，多維度鷹架機制如表 1 所示。

表 1 多維度鷹架機制

鷹架類型	遊戲設置	內容
認知鷹架	遊戲空間	學習者能透過觀看家中、辦公室的物件或標語，了解到行政倫理的基礎知識，例如行政中立等等。
	個人電腦（Google 表單）	學習者能獲得行政倫理的基礎知識，以及事件的相關線索，亦能回報任務，讓學習者評估自身知識及對事件的掌握程度。
	遊戲 NPC（Google 表單）	學習者能透過與家人、同事與主管的對話，了解到行政倫理的知識，並能運用溝通技巧掌握事件細節，以完成遊戲任務。
程序鷹架	遊戲空間	學習者在第一個任務階段時，遊戲空間中將有額外的文字提示，讓學習者熟悉遊戲與任務的執行方式，使其有步驟可依循。
	個人電腦（Google 表單）	學習者能透過其逐步回憶任務內容，並了解此階段的任務應該要做什麼。

情意鷹架	遊戲 NPC（Google 表單）	學習者透過與家人的對話，能感受到彼此的濃厚情感，將更有動力完成遊戲；與同事及主管的對話，則能感受到真實的公務職場情境，並能感受到不同的人格特質與面對主管的緊張氣氛。
動態情節互動鷹架	遊戲空間	學習者在不同的任務階段時，雖然身處同個遊戲空間（家中、辦公室），但景物卻會隨著時間的進展而有所變動，能讓學習者感受到事件的發展，具有高度擬真性。
	遊戲 NPC（Google 表單）	學習者若選擇不同的對話選項，NPC 將給予不同的互動反應；而在不同的任務階段時，以及與特定的人物對話過後，亦會產生不一樣的對話內容，呈現出職場的業務指派及人際關係，以及人際相處的潛規則，具有高度擬真性。

4. 研究工具

4.1. 學習成效問卷

此問卷由研究者發展，目的為測量學習者經由遊戲過後，對於行政倫理概念的提升情況，包含知識與情境等兩部分。知識部分包含工作態度、行政處理、生活態度及人際交往等四個面向，共有 14 題；情境部分則為模擬真實公務職場環境，以 Google 表單呈現，學習須分析職場狀況，依據行政倫理原則做出最適當的反應，並可區分為 4 個等級。題目加總分數為 100 分，均通過公共行政專家檢閱，具備專家效度。

4.2. 心流量表

此量表改編自 Kiili（2006）心流量表，並由 Hou 與 Li（2014）翻譯為中文版本，目的為測量學習者在活動中的心流狀態，包含心流前提與心流經驗等兩個主要項目，採用 Likert 五點量表，共有 22 題。在本研究中，其量表之 Cronbach's α 值為.917，具備高度信度。

4.3. 活動焦慮量表

此量表發展自 Krashen（1987）的情意濾網假說量表（Affective Filter Hypothesis），並由 Hung（2001）翻譯為中文版本，目的為測量學習者在活動中的焦慮狀態，採用 Likert 五點量表，共有 8 題。在本研究中，其量表的 Cronbach's α 值為.862，具備高度信度。

4.4. 學習動機量表

此量表由研究者發展並修改自 Pekrun（2005）的情緒階段性量表，目的為測量學習者遊戲活動前、中、後的動機狀態，採用 Likert 五點量表，共有 3 題。在本研究中，其量表之 Cronbach's α 值為.849，具備高度信度。

4.5. 批判思考量表

此量表引用由 Lin 等人（2019），目的為測量學習者在活動中的批判思考傾向，採用 Likert 五點量表，共有 6 題。在本研究中，其量表之 Cronbach's α 值為.904，具備高度信度。

5. 研究結果

5.1. 學習成效

為了解學習者在活動進行前後的學習成效差異，本研究使用 Wilcoxon 符號等級檢定進行檢驗。結果顯示學習者在經由遊戲活動之後，學習成效之後測（M = 75.63）與前測（M = 64.32）相比，為顯著優於前測（Z = 2.92, p = .003），表示此教育遊戲具有正向的學習效果。

5.2. 心流

為了解學習者遊戲活動的心流狀態，本研究使用單一樣本 Wilcoxon 符號等級檢定，與 Likert 五點量表之中位數（亦即 3）進行檢驗，結果如表 2 所示。結果顯示學習者在心流前提（Z = 4.09, p = .000）、心流經驗（Z = 4.27, p = .000）與總體心流（Z = 4.27, p = .000）狀態均顯著大於 3，表示學習者均擁有良好的心流狀態，擁有動力能繼續進行遊戲。

表 23 心流之單一樣本 Wilcoxon 符號等級檢定摘要表

	M	SD	*Z*	*p*
心流前提	4.03	0.66	4.09***	.000
挑戰與技能的平衡	4.08	0.60	4.04***	.000
清楚的目標	3.94	0.87	3.46**	.001
明確的回饋	3.85	0.74	3.53***	.000
控制感	4.31	0.57	4.24***	.000
行動與知覺的融合	3.96	0.95	3.50***	.000
心流經驗	4.12	0.58	4.27***	.000
專注度	4.29	0.72	4.18***	.000
時間扭曲	3.73	0.89	3.05**	.002
自成經驗	4.27	0.65	4.28***	.000
喪失自我意識	3.85	0.96	3.17**	.002
總體心流	4.08	0.51	4.27***	.000

**p < .01, ***p < .001

5.3. 活動焦慮

為了解學習者遊戲活動的焦慮狀態，本研究使用單一樣本 Wilcoxon 符號等級檢定，與 Likert 五點量表之中位數(亦即 3)進行檢驗。結果顯示學習者的活動焦慮顯著小於 3(M = 2.28, Z = -3.47, p = .000），表示此遊戲活動並不會對學習者造成太多的焦慮。

5.4. 學習動機

為了解學習者在遊戲活動前、中、後對學習主題的學習動機，本研究使用 Wilcoxon 符號等級檢定，評估學習者在活動前至中、活動中至後的動機狀態。結果顯示學習者在活動前（M = 3.46）至活動中（M = 3.83）的動機變化具有顯著差異（Z = 2.12, p = .034），活動中（M = 3.83）至活動後（M = 3.92）則無顯著提升（Z = 0.71, p = .480）。表示此遊戲能明顯提升學習者在學習活動中的動機狀態，讓學習者對學習行政倫理更有動機。

5.5. 批判思考傾向

為了解學習者在遊戲活動中的批判思考狀態，本研究使用單一樣本 Wilcoxon 符號等級檢定，與 Likert 五點量表之中位數（亦即 3）進行檢驗。結果顯示學習者的批判思考傾向顯著高於 3（M = 4.47, Z = 4.29, p = .000），表示此遊戲活動能讓學習者產生更多的思考，並判斷倫理事件的合理性。

6. 結論

本研究發展出一套學習行政倫理的教育遊戲，其結合了動態情節互動鷹架，以及數位敘事的概念，讓學習者在擬真的遊戲空間中，學習行政倫理的知識與概念。我們藉由讓學習者扮演公務人員，參與公務機關的運作，並體驗真實且生動的對話過程，改善了行政倫理教學的互動性。而以動態情節互動鷹架為首的多維度鷹架，除注重時間與事件的推移之外，也加

強了遊戲對學習者的支援與回饋。在經由初步的實證研究後，此遊戲能成功提升學習者學習成效，並維持在高度的心流狀態，同時保有較低的焦慮情形，也能提升他們的學習動機。這和 Lin 與 Hou（2023）的發現相似，表示學習者除了能持續進行遊戲之外，亦對學習主題逐漸湧現興趣。而在批判思考傾向的部分，亦能產生相當明顯的正向效果。比照 Chang 等（2020）、Hwang 與 Chang（2023）的研究結果，代表此遊戲的擬真環境，可能促進了學習者對於行政倫理的思考，並讓他們試圖去判斷事件中倫理或價值的合理性。後續研究將加入準實驗研究法的設計，與目前常見的線上講述教學做比較，以深入確認此遊戲對行政倫理教學所產生的效果，此外亦新增對於倫理立場的討論，了解不同立場是否會有不同的情況發生。

參考文獻

Austin, L. L., & Toth, E. L. (2011). Exploring ethics education in global public relations curricula: Analysis of international curricula descriptions and interviews with public relations educators. *Public Relations Review, 37*(5), 506-512.

Ch'Ng, E., Li, Y., Cai, S., & Leow, F. T. (2020). The effects of VR environments on the acceptance, experience, and expectations of cultural heritage learning. *Journal on computing and cultural heritage (JOCCH), 13*(1), 1-21.

Chang, C. Y., & Hwang, G. J. (2019). Trends in digital game-based learning in the mobile era: a systematic review of journal publications from 2007 to 2016. *International Journal of Mobile Learning and Organisation, 13*(1), 68-90.

Chang, C. Y., Kao, C. H., Hwang, G. J., & Lin, F. H. (2020). From experiencing to critical thinking: a contextual game-based learning approach to improving nursing students' performance in electrocardiogram training. *Educational Technology Research and Development, 68*, 1225-1245.

Chapman, R. A. (2019). *Ethics in public service*. Edinburgh University Press.

Chen, C. H., & Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. *Computers in Human Behavior, 55*, 1201-1212.

Chen, H. L., & Wu, C. T. (2023). A digital role-playing game for learning: Effects on critical thinking and motivation. *Interactive Learning Environments, 31*(5), 3018-3030.

Cooper, T. L. (2012). *The responsible administrator: An approach to ethics for the administrative role*. John Wiley & Sons.

Exmeyer, P. C., & Boden, D. (2020). The 8-Bit Bureaucrat: Can Video Games Teach Us About Administrative Ethics?. *Public Integrity, 22*(5), 409-424.

Gil, M., & Sylla, C. (2022). A close look into the storytelling process: The procedural nature of interactive digital narratives as learning opportunity. *Entertainment Computing, 41*, 100466.

Hejka-Ekins, A. (1988). Teaching ethics in public administration. *Public administration review*, 885-891.

Hou, H. T., & Keng, S. H. (2021). A dual-scaffolding framework integrating peer-scaffolding and cognitive-scaffolding for an augmented reality-based educational board game: An analysis of learners' collective flow state and collaborative learning behavioral patterns. *Journal of Educational Computing Research, 59*(3), 547-573.

Hou, H. T., & Li, M. C. (2014). Evaluating Multiple Aspects of a Digital Educational Problem-solving-based Adventure Game. *Computers in Human Behavior, 30*, 29-38.

Hung, C. M., Huang, I., & Hwang, G. J. (2014). Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. *Journal of Computers in Education, 1*(2), 151-166.

Hung, M.H. (2001). *The Effects of Captioned, Subtitled and Non-captioned Television Videotapes on Foreign Language Learning*. Master's thesis, National Cheng Kung University Institute of Education, Tainan, Taiwan.

Hwang, G. J., & Chang, C. Y. (2023). Facilitating decision-making performances in nursing treatments: A contextual digital game-based flipped learning approach. *Interactive Learning Environments, 31*(1), 156-171.

Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, 2*(2), 187-201.

Krashen, S. D. (1987). Principles and Practice in Second Language Acquisition. *Prentice-Hall International.*

Lin, H. C., Hwang, G. J., & Hsu, Y. D. (2019). Effects of ASQ-based flipped learning on nurse practitioner learners' nursing skills, learning achievement and learning perceptions. *Computers & Education, 139*, 207-221

Lin, Y. C., & Hou, H. T. (2023). Multi-dimensional evaluation of an educational board game using real-time diagnostic procedure scaffolding: analysis of learners' learning effectiveness, flow, anxiety, and emotion. *Educational Psychology, 43*(8), 874-894.

Menzel, D. C. (1997). Teaching ethics and values in public administration: Are we making a difference?. *Public Administration Review*, 224-230.

Meyer, S. J. (2021). Everything is fine! Using "The Good Place" to teach administrative ethics. *Journal of Public Affairs Education, 27*(2), 126-140.

Murray, J. H. (2018). Research into interactive digital narrative: a kaleidoscopic view. In Interactive Storytelling: 11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11 (pp. 3-17). Springer International Publishing.

Ni, A. Y. (2013). Comparing the effectiveness of classroom and online learning: Teaching research methods. *Journal of public affairs education, 19*(2), 199-215.

Pekrun, R. (2005). Progress and Open Problems in Educational Emotion Research. *Learning and Instruction, 15*, 497-506.

Pho, A., & Dinscore, A. (2015). Game-based learning. *Tips and trends*.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE), 1*(1), 21-21.

Roth, C., & Koenitz, H. (2016, October). Evaluating the user experience of interactive digital narrative. In Proceedings of the 1st international workshop on multimedia alternate realities (pp. 31-36).

West, J. P. (2021). Ethics education in public administration: An introduction to the symposium. *Journal of Public Affairs Education, 27*(4), 392-397.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines.*

Zavattaro, S. M. (2022). "We'll See Who's Powerless Now!" Using WALL-E to Teach Administrative Ethics. *Public Integrity, 24*(7), 702-716.

高中化学教育桌游开发设计与教学实践研究：以“绿色工厂”为例

The Design and Teaching Practice of a High School Chemistry Education Board Game: Taking “Green Factory”as an Example

杨景淇[1]，胡玲燕[2]，江丰光[3*]
[13]上海交通大学教育学院
[2]上海市实验学校
[*] fkchiang@sjtu.edu.cn

【摘要】 本研究以高中化学知识为基础，结合心流理论开发设计出一款以化工生产为主题的教育桌游——“绿色工厂”；并在中国海南省某高中展开教学实践，以高二年级 42 名学生为研究对象，分析“绿色工厂”桌游教学对学生化工思维的影响，学生的心流体验以及对桌游的接受度如何，综合评估桌游的教学效果和使用价值。结果显示，学生在使用“绿色工厂”桌游时产生了较高的心流体验，桌游教学后学生的化工思维显著提升。
【关键词】 化学教育；教育桌游；心流；化工思维

Abstract: *This study designed an educational board game named ‘Green Factory’, focused on chemical production and incorporating the Flow theory. The research was conducted at a high school in HaiNan with 42 second-grade students as participants. The impact analyzed the effects of the ‘Green Factory’ board game on students’chemical engineering thinking, their experience of flow, and their acceptance of the game to provide a comprehensive assessment of its pedagogical effectiveness and utility. The results showed that students experienced an enhanced state of flow while engaging with the ‘Green Factory’ board game, and the game is beneficial to significantly improve their chemical engineering thinking.*

Keywords: Chemical Education, Educational Board Game, Flow, Chemical Engineering Thinking

1. 前言

基于游戏的学习（Game -Based Learning, GBL）是指将教育游戏（Educational Game）融入到课程教学和自主学习活动中（Krath et al., 2021）。作为一种新兴的教育方式，GBL 强调寓教于乐的教育理念，凸显游戏对教育的支持作用。已有研究表明， GBL 对学生的学习成效有正面影响(李玉斌等人, 2019)，并且有利于调动学生的学习积极性(Jabbar & Felicia, 2015; 曾萍, 2022)，以及培养学生计算思维、创造力、问题解决能力等高阶能力（张屹等人, 2022）。教育桌游作为 GBL 环境下的一种重要应用形式，相比于其他电子类教育游戏，桌游由于易用性高、交互性强、以学生为中心等特点，更容易被学校和家长所接受，适合在基础教育阶段引入教学课堂。目前，将桌游用于教学已成为国外教育领域的热点方向，就化学教育而言，国外开发了多种知识主题的化学桌游，例如元素周期表（Franco et al., 2012; Moreno et al., 2014)，化学用语(Klara et al., 2013; Yenikalaycı et al., 2019)和有机化学(Knudtson, 2015; Gogal et al., 2017）等；而国内化学教育桌游的发展起步相对较晚，相关研究较少；且目前国内外的化学桌游相关文献大多只停留在开发设计阶段，实证研究较少或者缺乏更多维度的分析来进一步验证桌游对学生学习的影响。

整合化学知识难点，融入桌游教学是一种面向未来教育的新思路和新方法，符合新课程

改革发展素质教育的基本要求，有助于真正落地学生化学核心素养的培养。本研究立足于高中化学知识，结合心流理论开发设计出一款以化工生产为主题的教育桌游——“绿色工厂”，并通过实证研究探讨桌游的使用效果和应用价值。本研究选择化工生产为主题的原因如下：（1）目前高中化学教学越来越重视联系化工生产实际情境，《普通高中化学课程标准》（2017年版 2020 年修订）中明确提出该部分知识的内容要求和学业要求等；（2）该模块知识点在新教材中分布零散，本身有一定难度，不利于学生系统学习和掌握；（3）化工生产是将化学知识转化成社会生产的关键桥梁，该模块有助于培养学生的化工思维；（4）传统讲授式的课堂难以让学生体验工艺流程的各个环节，理解较为抽象困难。因此，本研究的目的在于：

（1）以高中化工生产知识为基础，结合心流理论设计一款以化工生产为主题的教育桌游——“绿色工厂”。

（2）通过准实验研究，分析“绿色工厂”桌游教学对学生化工思维的影响，学生在过程中的心流体验以及接受度如何，进而评估“绿色工厂”桌游的教学效果和应用价值。

2. 文献探讨

Csikszentmihalyi（1975）最早提出心流（Flow）概念，指的是个体因内在动机而完全沉浸于某项活动或情境中的一种状态。在学习过程中，学习者在心流状态下会对学习内容保持较高的兴趣，并积极投入学习过程，伴有较好的学习表现。经历心流体验后，学习者会为了再次获得心流体验而主动投入到下一次学习活动中，持续学习的动机由此产生。相关研究证实，心流理论可运用于教育游戏的设计指导和效果评估。Hsieh（2013）等研究发现，在教育游戏过程中，具有较高心流体验的学生往往具有更高的学习表现；Li（2022）发现中学生在进行化学教育桌游活动时具有持续的心流体验，促使他们积极参与活动。

Kiili（2014）在 Csikszentmihalyi 心流理论的基础上提出了适用于教育游戏心流框架，将心流体验组成元素分为心流前因（Flow Antecedents）和心流状态（Flow State）两部分：明确的目标、及时的反馈、挑战与技能的平衡、可玩性和控制感是心流体验产生的前提条件，归属于心流前因；注意力、时间失真、自主参与和自我意识消失是心流体验过程的认知特征，归属于心流状态。

在桌游过程中，如何让学习者进入到心流状态是促进学习效果和持续学习的关键。因此，教育桌游的设计要满足心流前因要素，其中明确的目标、及时的反馈、挑战与技能的平衡被认为是引发心流体验最重要的三要素（Chen, 1999; Kiili, 2006）。明确的目标具有导向功能和激励功能，可以激发学习者的兴趣，帮助其更快进入状态。桌游过程中，反馈的主要目的是告知学习者的个人表现和实现目标的进展，学习者的每次行动都会有结果，但结果呈现的快慢会很大程度上影响学习者的体验：及时的反馈可以激发积极情绪，让学习者保持专注，思考下一步决策；而延时的反馈会导致学习者不清楚游戏进程，产生消极情绪，抗拒反馈结果。此外，当学习者的技能不足以完成桌游挑战时，就会陷入焦虑的状态；当学习者面临的任务或者目标太简单时，技能水平高于完成挑战的需要就会陷入厌倦的状态，只有当技能与挑战平衡，才能产生心流。因此，本研究选取心流理论作为指导桌游设计的核心理论，并对学生在桌游实际学习过程的心流体验进行测量和评估。

3. 游戏设计

“绿色工厂”桌游主题为化工生产，知识内容围绕高中化学阶段中的三个重要化学工业：硫酸工业、合成氨工业和纯碱工业展开，以《普通高中化学课程标准》学业要求和高中化学教材（人教版、沪科版）为基础，联系实际生产情境适度拓展；涉及到化工生产中的反应原理、物料选择、能源消耗、工艺流程、环境保护等等问题。在游戏过程中，学生扮演化工工

程师角色，依据化学原理，收集相应的物质原料和机器设备，选择正确的生产条件，统筹多方资源，建立起硫酸、纯碱、合成氨等化工厂；并监控生产过程，及时解决遇到的技术问题；同时生产过程注重发展绿色化学，保护生态环境，实现可持续发展。

游戏目标分数包括：成功建厂（30%）、经济效益（25%）、绿色化学（45%）三部分；分数最高者获得胜利。学生通过达成游戏目标的过程，进而实现掌握化工生产相关知识，运用“绿色化学”思想分析解决生产的相关问题，发展化工思维的学习目标。图 3-1 为“绿色工厂”的桌游设计展示。



图 3-1“绿色工厂”桌游设计展示

“绿色工厂”桌游可供 2-8 名学生使用，每套桌游组件类型与数量如表 3-1 所示，表 3-2 介绍了不同卡牌所涉及的化学知识和对应游戏功能。

表 3-1 “绿色工厂”桌游组件

组件类型	种类*数量	组件类型	种类*数量
游戏版图	1*1	工厂卡	3*4
集团工厂版图	4*1	价格卡	6*5
化工集团招牌卡	4*1	游戏代币	8*20
物质卡	12*4	工厂 token	3*4
设备卡	24*4	玩家代表物	4*1
百宝箱	7*4	四面骰	1*1
事件卡	9*3	工艺指南	1*1
专家站	4*4	游戏说明书	1*1
生产条件卡	11*4		

表 3-2 部分卡牌对应的化学知识和游戏功能

卡牌类型	知识联系和游戏功能
	物质卡牌面包含物质名称、化学式、外观图、主要性质用途等内容，提供信息，作为学习支架；同时标明价格和绿色化学分数，学生可自主权衡进行选择和计算。物质卡涉及三大工业生产的原料物质（例如硫磺、硫铁矿、空气、原盐等）、催化剂（铁触媒、五氧化二矾）、三废处理物质（石灰乳、氨水等）。

卡牌类型	知识联系和游戏功能
	设备卡牌面包含设备名、图例、用途以及发生反应的化学方程式，为提供信息，作为学习支架。设备卡涉及三大工业生产的主要反应设备、除杂净化设备以及三废处理设备等。学生根据生产流程自主选择设备，完成建厂分数和绿色化学分数。
	生产条件卡包含温度和压强两种类型，学生可根据化学反应速率、化学平衡移动原理和综合成本考虑选择适宜的反应条件。
	工厂卡包含硫酸、合成氨、纯碱工厂三种类型，学生根据工艺指南收集完每个工厂对应的物质、设备和生产条件卡时，可获得对应工厂卡上的成功建厂分数和经济收益。
	事件卡表示化工生产中可能出现问题，例如催化剂中毒、停电、生产安全等；也会出现相关部门的抽检事件，例如环保抽检和市场质检等；还会带来相关的机遇，例如原料补贴、设备补贴、打击垄断等；学生需做好准备迎接相应的事件挑战。
	专家站卡包含安全员、技术员、质检员和环保专家四种类型，不同专家聘请薪资不同，职能不同，能够为化工厂带来一定的经济收益和事件支持。学生可根据需求聘请专家，助力化工生产。
	百宝箱卡能够提供给学生不同的技能，以增添游戏进程的随机性和趣味性。例如：市场回收、免税、加速、交易、定价等等。

4. 研究设计

4.1 研究对象与方法

本研究在中国海南省某高中开展桌游教学，研究对象为高二年级 42 名高考选考化学的学生。在一周内共开展两次桌游教学活动，每次时间控制在 80~100min；首先由教师背景讲解，知识引入（10-15min）；其次教师解释桌游规则流程、游戏布置（5-15min）；接着学生自主进行桌游活动，教师指导（45-60min）；最后交流分享，课程总结（5-10min）。研究采用准实验研究，单组前后测设计。




图 4-1“绿色工厂”桌游教学实践

4.2 研究工具

本研究用于评估桌游应用效果的测量工具包括：化工思维量表、心流量表和桌游接受度量表。针对化工思维量表用于评估学生的化工思维水平，量表参考段南（2023）编制的《高中生化工思维量表》，Cronbach's alpha 值为 0.83，信度较好。根据化工思维的构成要素，量表分为实践创造、社会价值、绿色安全和复杂权衡四个维度，采用李克特 5 点量表的形式，共 16 道题目。该量表在桌游教学活动开始前和结束后分别填写。

心流量表用于测量学生在化学桌游使用过程中的心流体验。采用的是由 Kiili（2006）所编制的教育游戏心流量表，Cronbach's alpha 值为 0.91，信度较高；量表分为心流前因和心流状态两部分；采用李克特 5 点量表的形式，共 22 道题目。每次桌游教学结束后及时填写。

桌游接受度量表用于评估学生对桌游的有用性和易用性的态度看法。量表参考了 Lin（2022）所修订的游戏接受度量表，Cronbach's alpha 系数值为 0.93；量表分为桌游有用性和桌游易用性两部分，采用李克特 5 点量表的形式，共 8 道题目。

4.3 数据处理与分析

使用 SPSS26.0 软件对收集数据进行处理分析，采用配对样本 t 检验对学生的化工思维水平变化进行分析，对学生的心流体验和桌游接受度数据作描述性统计分析。

5. 研究结果

5.1 化工思维

学生在桌游活动前后化工思维平均得分分别为 54.90±7.302，63.10±4.689，经检验，前后得分具有显著性差异（t=-9.826，p<0.001）；具体到四个维度的前后测得分结果如图 5-1 所示，通过“绿色工厂”桌游实践，四维度得分都显著上升；其中实践创造思维上升最显著，说明学生通过桌游活动，在理解化工理论联系实际生产、掌握工艺流程知识方面自评提升最高；复杂权衡思维在四维度中相对提升最小，说明学生在以宏观整体视角看问题，综合权衡物料、工艺、能源、经济、环境等因素上还需进一步提高巩固。

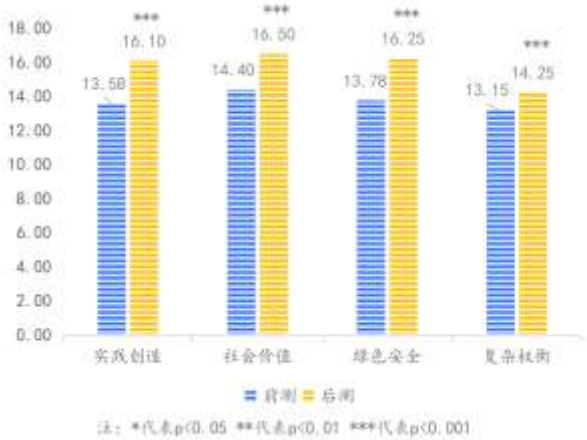


图 5-1 化工思维各维度得分前后测变化

5.2 心流体验

学生在两次桌游活动中总体心流体验的平均得分分别为 4.17±0.602 和 4.47±0.552，心流前因和心流状态各要素得分结果如图表 5-1 所示，所有子维度得分均高于中位数 3 以上，说明桌游设计达到了引发心流产生的前提条件，学生在活动中获得了较高的心流体验，进入到心流状态中，具体表现为两次活动中学生都体现了高度的注意力集中、自主参与以及忽略了时间的流逝。同时，在第二次活动中，学生在心流前因各个维度得分较第一次有所上升，且都达到 4 以上，说明学生在熟悉了解桌游规则和玩法后，更容易进入心流状态。

表 5-1 心流各维度结果分析

心流维度	第一次桌游活动		第二次桌游活动	
	平均值	标准差	平均值	标准差
心流前因	**3.88**	**0.621**	**4.37**	**0.612**
挑战与技能的平衡	4.41	0.553	4.56	0.572
明确的目标	3.89	0.937	4.54	0.674
及时的反馈	3.99	0.720	4.40	0.700
控制感	3.62	0.835	4.24	0.866
可玩性	3.51	0.614	4.12	0.720
心流状态	**4.44**	**0.430**	**4.57**	**0.479**
注意力	4.38	0.551	4.59	0.517
时间失真	4.73	0.466	4.71	0.535
自主参与	4.74	0.447	4.79	0.402
自我意识消失	3.95	0.741	4.21	0.908

5.3 桌游接受度

学生对“绿色工厂”桌游的反馈结果如表 5-2 所示。桌游的有用性和易用性得分分别为 4.75±0.461、4.62±0.610。说明学生认可“绿色工厂”用于教学的价值，对桌游的接受度较高。例如在桌游的有用性方面上，超过 75%的学生非常认同桌游对学习化工生产知识游有帮助以及通过桌游强化了相关的知识点；在易用性上，超过 70%的学生非常认同桌游学习是一种灵活的学习方式。也有少部分学生对桌游的规则和玩法表示不是很清楚，这说明桌游需要进行一些机制和规则的难度调整，进行梯度划分，考虑制作简易版和进阶版桌游。

表 5-2 桌游接受度结果分析

	题目	五点量表					平均值	标准差
		5	4	3	2	1		
桌游有用性	Q1.《绿色工厂》可以帮我学习化工生产相关知识	78%	20%	2%	-	-	4.75	0.488
	Q2.《绿色工厂》桌游课比传统课堂更能帮助我理解化学工艺流程	76%	24%	-	-	-	4.75	0.434
	Q3.学习过程中增加化学桌游可以缓解听课疲劳	78%	20%	2%	-	-	4.75	0.488

	题目	五点量表					平均值	标准差
		5	4	3	2	1		
	Q4.桌游帮助我学习/复习了硫酸、合成氨纯碱工业生产以及绿色化学等相关知识点	73%	27%	-	-	-	4.73	0.448
					桌游有用性		**4.75**	**0.461**
桌游易用性	Q5.《绿色工厂》的规则玩法和知识呈现都容易理解	68%	27%	5%	-	-	4.63	0.581
	Q6.《绿色工厂》游戏过程顺畅	61%	27%	12%	-	-	4.48	0.711
	Q7.《绿色工厂》桌游学习是一种灵活的学习方式	78%	20%	2%	-	-	4.75	0.488
	Q8.总的来说，我认为桌游作为教学工具容易使用	66%	27%	7%	-	-	4.58	0.631
					桌游易用性		**4.62**	**0.610**

注： 5-非常同意，1-非常不同意

6. 结论

桌游进课堂是一种创新的教学方式，学生利用桌游可以实现“玩中学”，在轻松愉快的氛围中学习知识和技能。本研究以高中化学知识为基础，结合心流理论开发设计出一款以化工生产为主题的教育桌游——“绿色工厂”，并在高中展开教学实践，评估桌游的教学效果和使用价值。研究结果显示，学生在使用“绿色工厂”桌游时产生了较高的心流体验，化工思维显著提升，学生对桌游的接受度较高。本次桌游教学实践也存在一定的局限，例如进行桌游教学干预的周期较短，只进行了两次课程。在未来的教学实践中，会进行更长周期的干预研究，进一步评估桌游教学效果。同时会在桌游设计中融入更多的化学学科知识，结合学生的心流体验结果和接受度反馈，继续优化桌游的机制和玩法；根据不同的受众群体，对桌游难度进行梯度分级，制作简易版和进阶版桌游，提升“绿色工厂”作为教学工具使用的有用性和易用性。希冀本研究能为未来高中化学课堂教学提供新的思路。

参考文献

中华人民共和国教育部（2020）。普通高中化学课程标准（2017 版 2020 修订），人民教育出版社。

李玉斌，宋金玉，姚巧红（2019）。游戏化学习方式对学生学习效果的影响研究——基于 35 项实验和准实验研究的元分析。电化教育研究，40(11)，56-52。

张屹，马静思，周平红等（2022）。人工智能课程中游戏化学习培养高中生计算思维实践的研究——以“挑战 Alpha 井字棋”为例。电化教育研究，43(9)，63-72。

段南，（2023）。基于 STEAM 教育理念发展高中生化工思维能力的教学研究。硕士学位论文，延安大学。

曾萍（2022）。 游戏化学习提升学生自主学习能力。中小学信息技术教育(7)，36-38。

Chen, H., Wigand, R. T., & Nilan, M. S. (1999). Optimal experience of Web activities. *Computers in Human Behavior*, 15(5), 585-608.

Csikszentmihalyi, M. (1975). Flow: *The psychology of optimal experience*. Harper and Row.

Franco Mariscal, A. J., Oliva Martínez, J. M., & Bernal Márquez, S. (2012). An Educational Card Game for Learning Families of Chemical Elements. *Journal of Chemical Education*, 89(8), 1044-1046.

Gogal, K., Heuett, W., & Jaber, D. (2017). CHEMCompete: An Organic Chemistry Card Game To Differentiate between Substitution and Elimination Reactions of Alkyl Halides. *Journal of Chemical Education*, 94(9), 1276-1279.

Jabbar, A. I. A., & Felicia, P. (2015). Gameplay Engagement and Learning in Game-Based Learning: A Systematic Review. *Review of Educational Research*, 85(4), 740-779.

Krath, J., Schürmann, L., & von Korflesch, H. F. O. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, 125, 106963.

Klara, K., Hou, N., Lawman, A., & Wang, L.-Q. (2013). Developing and Implementing a Collaborative Teaching Innovation in Introductory Chemistry from the Perspective of an Undergraduate Student. *Journal of Chemical Education*, 90(4), 401-404.

Knudtson, C. A. (2015). ChemKarta: A Card Game for Teaching Functional Groups in Undergraduate Organic Chemistry. *Journal of Chemical Education*, 92(9), 1514-1517.

Kiili, K., Lainema, T., de Freitas, S., & Arnab, S. (2014). Flow framework for analyzing the quality of educational games. *Entertainment Computing*, 5(4), 367-377.

Kiili, K. (2006). Evaluations of an experiential gaming model. *An Interdisciplinary Journal on Humans in ICT Environments*, 2(2), 187-201.

Moreno, L. F., Hincapié, G., & Alzate, M. V. (2014). Cheminoes: A Didactic Game To Learn Chemical Relationships between Valence, Atomic Number, and Symbol. *Journal of Chemical Education*, 91(6), 872-875.

Li, C.-T., Hou, H.-T., & Lin, W.-S. (2022). Chemistry education board game based on Cognitive Mechanism: Multi-dimensional evaluation of learners' knowledge acquisition, flow and playing experience of board game materials. *Research in Science & Technological Education*, 1–21.

Lin, Y.-C., & Hou, H.-T. (2022). The evaluation of a scaffolding-based augmented reality educational board game with competition-oriented and collaboration-oriented mechanisms: Differences analysis of learning effectiveness, motivation, flow, and anxiety. *Interactive Learning Environments*, 1–20.

Yenikalaycı, N., Çelikler, D., & Aksan, Z. (2019). Ion Hunters: Playing a Game To Practice Identifying Anions and Cations and Writing Their Names and Formulas. *Journal of Chemical Education*, 96(11), 2532-2534.

Hsieh, Y.-H., Lin, Y.-C., & Hou, H.-T. (2013). Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning. Interactive Learning Environments, 24(1), 178–193.

結合三維擬真互動情境式腳本鷹架之職場壓力調適能力培訓教育遊戲

Educational Game for Workplace Stress Adjustment Ability Training with 3D Realistic Interactive Situational Script Scaffolding

簡志忠[1]，吳培菁[2]，侯惠澤[3, 4*]
[1]臺灣科技大學應用科技研究所
[2]臺灣科技大學應用科技研究所
[3]臺灣科技大學應用科技研究所
[4]臺灣科技大學技職賦能研究中心
[*] hthou@mail.ntust.edu.tw

【**摘要**】職場新人適應不良而造成離職意向提高，將產生公司在新人招募和培訓方面的投資浪費。本研究結合內心思維與外在表達互動劇本之機制，發展出一款有助於新進同仁面對工作不適應，提升正向心理思維及良好溝通方式之 3D 情境擬真教育遊戲。本研究希望促進學習者在擬真的公司環境中，認識新人工作壓力大的三個層面的分析及概念，並透過與 NPC 角色互動事件的發展，讓學習者瞭解自己工作抗壓及溝通能力，並透過鷹架反思，提升學習者職場抗壓能力與正向思維。初步結果指出本研究在學習過程可以提升學習成效、維持高度學習心流、並降低學習焦慮。本遊戲具備輔助學習者發覺自身壓力調適問題並引導正向思維的潛力。
【**關鍵詞**】職場抗壓力；遊戲式學習；互動鷹架；情境擬真；離職意向

Abstract: *The increase in the turnover intention due to the maladjustment of newcomers in the workplace will result in the waste of company's investment in recruiting and training newcomers. This study designed interactive scripts that combined with internal thinking and external expression to develop a 3D situational simulation educational game that can help newcomers face workplace maladjustment and enhance positive mental thinking and good communication style. In this study, we hope to promote learners' understanding of the three phases of newcomers' work stress in a simulated company environment, and through the development of interactive events with NPC characters, learners can understand their own ability to adapt to stress and communication skills, and through the scaffolding reflection, enhance learners' ability to adapt to stress and positive thinking. Preliminary results indicated that this study can enhance learning effectiveness, maintain a high level of flow, and reduce learning anxiety during the learning process. The game has the potential to assist learners in identifying their own stress adjustment problems and to guide positive thinking.*
Keywords: Workplace stress tolerance, Game-based learning, Interactive scaffolding, Situational simulation, Turnover Intention

1. 前言

人員離職是全球許多公司所面臨到成本最高且棘手的人力資源挑戰之一(Shamsuzzoha & Shumon, 2007)。對公司來說，員工的離職意味著在人員招募和培訓方面的投資浪費，且員工離職率高也會影響現有人員的積極性且增加工作量並使工作計畫變得困難(Belete, 2018)。Deloitte 調查顯示，新世代員工的離職意向（Turnover Intention）率從 2017 年的 38%上升至 2019 年的 49% (Lu et al., 2023)。在所有離職族群中，新世代員工表現出最典型的特質，較容易形成鮮明的個性和高度的自我意識，他們的離職行為大多沒有明顯跡象，對現代企業人事管理和社會穩定帶來不確定的風險和挑戰(Zhao et al., 2019)。同時，新世代員工容易感到孤獨並有社會支持感低的傾向(Wang & Zhang, 2019)。組織心理學對離職意向進行了大量研究，發現離職意向是代表個體心理因素的傾向(Meyer & Allen, 1984)，有必要將動機和需求滿足等心

理變化作為問題的根源。Lu et al. (2023)研究提到，公司關注員工心理健康，可以提高員工的心理安全感和依賴感，進而降低離職意向，促進公司內部制度環境的穩定。另外，Al-Kahtani & Allam（2015）研究指出，在員工的溝通氛圍中，防禦性溝通氛圍對組織認同有顯著的負面影響，並增加了員工離職率。King & Emmons (1990) 提出「情緒表達矛盾」(ambivalence over emotional expression)的概念，意指個體內心思維與外在表達不同。由於刻意地不要表露真實感受，因此內心有較大的矛盾與緊繃，容易產生身心壓力。因此，提升新進同仁面對工作壓力的正向心理思維及良好的溝通方式的重要性，為本研究的研究動機。

此外， Doolittle & Camp (1999)研究指出，遊戲式學習活動可加入與學習主題相關的情境脈絡之中，將學習主題知識轉化成遊戲環境中的實際任務，可以較容易地吸引學習者的注意力並激發學習動機。因此在擬真場景上，同時導入互動式情境體驗學習，藉此模擬人際互動模式，可以進一步作為反思鷹架(Chien et al., 2023)，引導激發學習者的反思及批判性思考，或能積極主動進而維持高度動機及控制感，並在擬真環境中直觀的分析解決分配給他們的任務(Krajčovič et al., 2021)，透過人際互動脈絡機制並導入擬真情境與動機促進是有意義的線上學習重要關鍵 (Bovermann et al., 2018)。Chien et al. (2024) 研究發現，在擬真遊戲多路徑互動中，透過不同的溝通言語的選擇將產生不同的資訊內容跟劇情發展，將有助於引導學習者建構全面性歸納與推理能力。

有鑑於此，本研究設計一款有助於新進同仁面對工作不適應，提升正向心理思維及良好溝通方式之 3D 情境擬真教育遊戲。在遊戲中，期望藉由角色扮演來達成與職場同事或主管互動的擬真情境來促進學習動機(Gillispie et al., 2010)，並搭配各種鷹架來引導學習者反思學習與問題解決(Hou, 2022)，以提升新進同仁職場抗壓能力正向思維。情境遊戲式學習中的鷹架主要集中在遊戲過程中的三個方面：互動、敘事和知識(Zuo et al., 2022)。故本研究目的以數位遊戲設計，透過 3D 擬真情境、並結合「內心思維與外部對話」之互動情境式腳本鷹架、互動簡報的認知鷹架、知識探索與引導的即時反饋鷹架，初探學員們的學習成效、心流、焦慮感以及對於遊戲擬真度、遊戲接受度、遊戲元素之心理狀態。

2. 研究方法

2.1. 研究對象與程序

本研究經案例分析初步測試，參與者為台灣 20 名 20~60 歲之成年人，採自由於網路報名。參與者在施測前均未曾以 *Google Form* 互動劇本機制脈絡感知鷹架學習方式進行過相關課程。在實驗開始之前，所有學習者均被要求簽署詳細說明參與者權利與保護措施的知情同意書及保密合約書。實驗流程首先由研究者進行活動講解及施測流程說明(10 分鐘)，接下來進行教學活動(40 分鐘)，後測問卷施測評量 (15 分鐘)。

2.2. 教學遊戲設計

本研究設計一套 3D 擬真情境互動劇本教育遊戲《菜鳥順順飛》。遊戲任務是讓學習者化身一位新進同仁，在遊戲中循序與三位代表不同事件的關鍵同仁的非玩家角色(NPC: Non-player character)對話，對話內容包含內心思維與外部對話兩部份。每個 NPC 的問話，學習者的回答過程會有三種不同選項，學習者可選擇接近自己的「內心思維」與「外在表達」選項，同時不同的選擇也將產生不同的劇情路徑，作為互動情境式腳本鷹架，可以進一步達到反思學習，促進學習者藉由反覆思考各種選擇而產生不同情境結果。透過不同路徑的結果將會產生一組「指數碼」。所以當依序完成與三位 NPC 對話，分別獲得三組「指數碼」後，再輸入於「指數分析系統」中，即可呈現學習者在遊戲三個事件中分別的抗壓指數及溝通指

數。若指數低於 70%為不合格，可再一次探索遊戲中在虛擬場景電視牆上的互動簡報所提供的認知鷹架的提示。反思後進行第二輪的關鍵同仁(NPC)對話，此目的讓學習者透過隨時諮詢引導與探索知識的即時回饋鷹架，達成提升其學習成效。這些互動腳本與選擇項目的設計均由具有 25 年以上的企業高階經理人專家設計，具備專家效度。

本研究採用 Spot Virtual 平台建構 3D 場景，如圖 1。除擬真辦公室擺設及格局外，也設計探索解謎關卡，增加學習者在學習過程中的不確定性、控制感、成就感及趣味性。Hou (2015)指出，具有情境式學習的情境模擬遊戲可能有機會提高學習者的後設認知過程，達到高學習心流。另外，對話內容採用 Google Form 設計 NPC 互動劇本，如圖 2。讓學習者在高度擬真的公司環境及對話氛圍，同時沈浸故事及角色擬真感受，引導學習者體驗自己內心思維與外在表達不同現象，並瞭解自己在職場工作上可能的抗壓程度，達成動態即時反思學習。因此，本研究學習目的是透過《菜鳥順順飛》遊戲式學習後，提升新人工作抗壓能力與溝通能力。

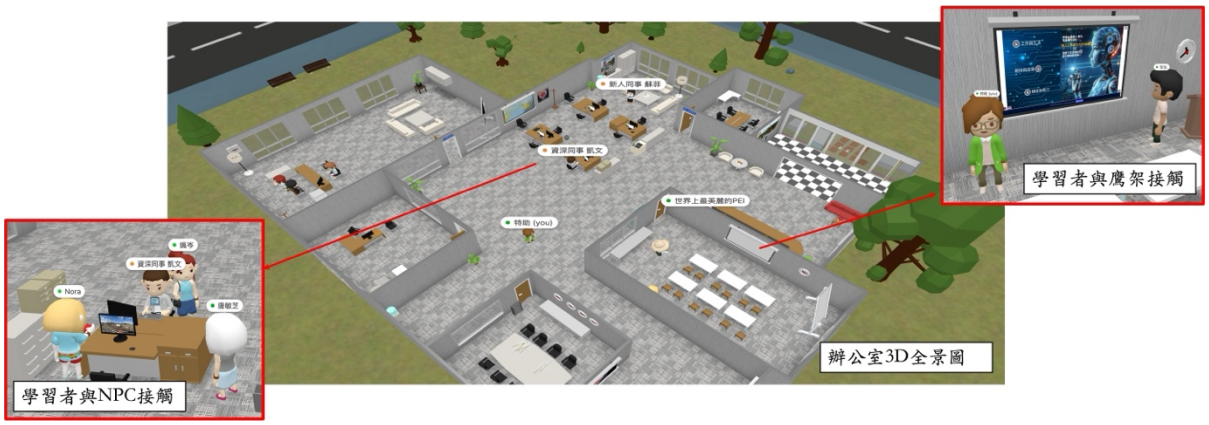


圖 1 遊戲畫面



圖 2 結合內心思維與外在表達的多路徑遊戲互動劇本鷹架

3. 研究工具

3-1 心流量表

在探討學習者學習活動中的心流狀態，使用 Kiili (2006)的心流問卷作為心流評量，由學者 Hou & Li (2014)翻譯為中文版本。共有兩大維度，分別為心流前提 (Flow Antecedents)以及心流經驗(Flow Experience)，共計 22 題。此問卷採用李克特氏五點量表。

3-2 活動焦慮量表

在探討學習者在參與活動時，焦慮程度如何，使用 Krashen (1987) 開發的情意濾網假說量表(Affective Filter Hypothesis)，由洪美雪(2001)改編之「學習經驗量表」，共計 8 題。此問卷採用李克特氏五點量表。

3-3 遊戲回饋

在探討學習者在遊戲中的使用心得部分，本研究修正 Davis (1989)提出的科技接受度量表，共有兩大維度，分別為認知有用性、認知易用性；此外，遊戲元素評估則參考侯惠澤 (2016) 所提出的促進遊戲動機的元素進行評估，共計 12 題。此問卷採用李克特氏五點量表。

3-4 學習成效

學習者經過與三種 NPC 互動事件後，本研究分析學習者於遊戲中獲得的「抗壓指數及溝通力指數」(路徑腳本與評估準則皆經由具備科技產業 25 年以上的高階經理人設計)。學習者於每一階段兩項指數皆需達 70%以上為合格。我們透過比較學習者兩次探索階段後，以學習總人數來計算面對三種同仁事件全合格比例是否提高。其中抗壓指數與溝通力指數分開比較。

4. 研究結果

表 1 為學習者完成任務後，各項維度的描述統計分析。在心流方面：整體心流($M=4.43$，$SD=0.55$)、心流前提($M=4.38$，$SD=0.60$)、心流經驗($M=4.47$，$SD=0.53$)皆顯著高於量表中位數(即 3)，其他各心流子維度平均值也皆高於 3。其中，子維度專注度平均值更高達 4.64。Chien et al. (2023) 的類似情境模擬教育遊戲的研究也指出，在探索互動機制下可提升玩家們的控制感，產生高度投入遊戲任務的心流體驗，進而提升線上學習專注力。在焦慮感方面：整體焦慮（ $M=2.02$，$SD=0.65$ ），低於 3 並達顯著。在遊戲活動中適度的低焦慮是持續心流的重要指標。在遊戲心得方面：整體遊戲回饋($M=4.37$，$SD=0.75$)、遊戲有用性($M=4.35$，$SD=0.75$)、遊戲易用性（ $M=4.40$，$SD=0.83$ ）、遊戲元素（ $M=4.33$，$SD=0.63$ ）也顯著高於 3。Hassan et al. (2021) 指出，良好的遊戲元素會引起學習者的成就感、參與感和提升動機。在擬真度方面：整體擬真($M=4.29$，$SD=0.45$ ）顯著高於 3，其他各擬真子維度平均值也皆高於 3。Hallinger et al. (2020) 研究提到，情境擬真可以支持學習者達到更佳組織概念化和應用知識的能力。在學習成效方面：工作抗壓指數在三項事件皆合格者，總比例由 10%提升為 60%。溝通指數三項事件皆合格者總比例由 10%提升為 50%。由此推論 3D 高擬真情境結合鷹架導向的 NPC 互動劇本鷹架機制的遊戲設計，應有助於提升學習成效。

表 1 單一樣本 Wilcoxon 符號等級檢定摘要表

	(N=20)			
	M	*SD*	*Z*	*Sig.*
整體心流	4.43	0.55	3.92***	0.000
心流前提	4.38	0.60	3.89***	0.000
心流經驗	4.47	0.53	3.93***	0.000
焦慮感	2.02	0.65	-3.71***	0.000
整體遊戲回饋	4.37	0.75	3.80***	0.000
有用性	4.35	0.75	3.83***	0.000
易用性	4.40	0.83	3.73***	0.000
遊戲元素	4.33	0.63	3.84***	0.000
整體擬真度	4.29	0.45	3.94***	0.000
操作擬真	4.10	0.79	3.51***	0.000
角色擬真	4.40	0.50	4.05***	0.000
劇情擬真	4.25	0.55	3.99 ***	0.000
場景擬真	4.40	0.68	3.84***	0.000

$^{***}p < .001$

5. 結論

《菜鳥順順飛》是本研究開發一款以公司培訓『職場新鮮人壓力適應』為主題的 3D 擬真情境線上遊戲，透過結合內心思維與外在表達的互動劇本，讓學習者體驗內心情緒與表達的矛盾概念，同時透過鷹架引導學習者瞭解自己工作上的心理壓力及給予如何正向面對的參考建議。綜合以上數據分析，顯示本研究以公司情境擬真體驗式學習，在導入鷹架導向的 NPC 劇本互動機制，包含互動情境式腳本鷹架、互動簡報的認知鷹架、知識探索與引導的即時反饋鷹架。在學習過程可以維持高度學習心流、並降低學習焦慮，同時也提升學習成效。另外，在學習者開放性反饋中也提到「遊戲會引導我說出心裡的想法」、「透過抗壓力回饋指數，可以知道自己內心想法不能太軟爛，而是要正向面對」。研究結果初步發現，本研究設計之 3D 情境擬真機制具有了解自身情況並引導正向思維來解決職場適應不良的潛力。未來研究可以再增加更多的樣本量，並探討鷹架導向的 NPC 劇情對話互動機制，與一般鷹架機制在線上學習的成效差異、擬真感受及後設認知之差異，可做更深入的全面比較分析。

參考文獻

Al-Kahtani, N. S., & Allam, Z. (2015). Research Article Mediating Effect of Role Ambiguity on Communication Climate: A Meta Analysis. *Research Journal of Applied Sciences, Engineering and Technology*, *10*(3), 350-356.

Belete, A. (2018). Turnover Intention Influencing Factors of Employees: An Empirical Work Review. *Journal of Entrepreneurship & Organization Management*, *07*(03). https://doi.org/10.4172/2169-026x.1000253

Bovermann, K., Weidlich, J., & Bastiaens, T. (2018). Online learning readiness and attitudes towards gaming in gamified online learning–a mixed methods case study. *International Journal of Educational Technology in Higher Education*, *15*(1), 1-17. https://doi.org/10.1186/s41239-018-0107-0

Chien, C. C., Ho, Y. T., & Hou, H. T. (2023). Integrating Immersive Scenes and Interactive Contextual Clue Scaffolding Into Decision-Making Analysis Ability Training Game. Journal of Educational Computing Research, 62(1), 376-405. https://doi.org/10.1177/07356331231205058

Chien, C. C., Ho, Y. T., Ngu, P. C., & Hou. H. T. (2024) Design of a Complex Problem-Solving Ability Training Game with Simulation Spaces and Realistic Plots, paper presented at the The Southeast Asian Conference on Education (SEACE2024), February 15-19, 2024, Chiang Mai, Thailand.

Doolittle, P. E., & Camp, W. G. (1999). Constructivism: The Career and Technical Education Perspective. *Journal of Career and Technical Education*, *16*(1) , 23-46. https://doi.org/10.21061/jcte.v16i1.706

Gillispie, L., Martin, F., & Parker, M. A. (2010). Effects of a 3-D video game on middle school student achievement and attitude in mathematics. *Electronic Journal of Mathematics and Technology, 4*(1), 68–80.

Hassan, M. A., Habiba, U., Majeed, F., & Shoaib, M. (2021). Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments*, 29(4), 545-565. https://doi.org/10.1080/10494820.2019.1588745

Hou, H. T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: A video-based process exploration. *Computers in Human Behavior*, 48, 424-435. https://doi.org/10.1016/j.chb.2015.02.010

Hou, H.-T. (2022). *Augmented reality board game with multidimensional scaffolding mechanism: A potential new trend for effective organizational strategic planning training*. *13*. https://doi.org/10.3389/fpsyg.2022.932328

King, L. A., & Emmons, R. A. (1990). Conflict over emotional expression: Psychological and physical correlates. *Journal of Personality and Social Psychology*, *58*(5), 864–877. https://doi.org/10.1037/0022-3514.58.5.864

Krajčovič, M., Gabajová, G., Matys, M., Grznár, P., Dulina, Ľ., & Kohár, R. (2021). 3D Interactive Learning Environment as a Tool for Knowledge Transfer and Retention. *Sustainability*, *13*(14), 7916. https://doi.org/10.3390/su13147916

Lu, J., Guo, S., Qu, J., Lin, W., & Lev, B. (2023). "Stay" or "Leave": Influence of employee-oriented social responsibility on the turnover intention of new-generation employees. *Journal of Business Research*, *161*, 113814. https://doi.org/10.1016/j.jbusres.2023.113814

Meyer, J. P., & Allen, N. J. (1984). Testing the "side-bet theory" of organizational commitment: Some methodological considerations.. *Journal of Applied Psychology*, *69*(3), 372–378. https://doi.org/10.1037//0021-9010.69.3.372

Shamsuzzoha, A. H. M., & Shumon, M. R. H. (2007). Employee Turnover- a Study of its Causes and Effects to Different Industries in Bangladesh. *Manufacturing Engineering/ Vyrobne Inzinierstvo*, *6*(3), 64-68.

Wang, M. M., & Zhang, J. (2019). The impact of authentic leadership on new generation employee creativity: the mediating role of perceived insider status. *Sci Manage ST*, *40*(3), 127-141.

Zhao, Y. X., Zhao, S. M., & Xu, Y. F. (2019). A study of intergenerational differences of Chinese employees based on two years of achieving styles data. *Chin. J. Manag*, *16*, 1751-1760.

Zuo, T., Jiang, J., Spek, E. V. der, Birk, M., & Hu, J. (2022). Situating Learning in AR Fantasy, Design Considerations for AR Game-Based Learning for Children. *Electronics*, *11*(15), 2331. https://doi.org/10.3390/electronics11152331

EMI 數位課程遊戲融入議題探究與 SDGs 教學策略運用於關鍵力教育

Integrating Issue-Based Inquiry Learning into English as a Medium of Digital Instruction Gaming Curriculum and Applying SDGs Teaching Strategies to Key Competencies Education

許雯晴[1]，陳聖智[2*]，林永騰[3]，蔣旭政[4]
[1]政治大學傳播學院傳播碩士學位學程
[2]政治大學傳播學院數位內容碩士學位學程
[3]臺北教育大學師資培育處
[4]台灣師範大學大眾傳播研究所
[*] scchen222@gmail.com

【**摘要**】本研究目的主要在於探討關鍵能力教育下資訊科技融入國小五年級英語議題探究課程對學習成效、學習動機與學習行為的影響。透過三大面向關鍵能力的探究，包括學習與創新、數位素養、生活與解決問題能力。本研究分析了英語數位課程、SDGs 教學策略與關鍵力教育的融合，如何將英語數位課程與 SDGs 結合，應用於關鍵力教育，學習內容包括蒐集資訊、表達想法、規劃活動等。研究方法採用前實驗設計法，對學生進行為期三週的實驗課程。結果顯示，資訊科技融入教學對英語學習成效和動機有正面影響，並改善了學生的創造力、認知負荷、易用性和環境保育意識。

【**关键词**】關鍵力教育；議題探究學習；英語數位課程；SDGs 教學策略；學習行為分析

Abstract: *The main purpose of this study is to explore the impact of integrating information technology into the fifth-grade English issue-based inquiry curriculum under key competencies education on learning outcomes, motivation, and behavior. Through the exploration of three key competencies, including learning and innovation, digital literacy, and life and problem-solving abilities, this study analyzes the integration of the English digital curriculum, SDGs teaching strategies, and key competencies education. It examines how to combine the English digital curriculum with SDGs and apply them to key competencies education, with learning content including information collection, idea expression, and activity planning. The research method used is a pre-experimental design, conducting a three-week experimental course for students. The results show that the integration of information technology into teaching has a positive impact on English learning outcomes and motivation and has improved students' creativity, cognitive load, usability, and environmental conservation awareness.*

Keywords: Key Competencies Education, Issue-Based Inquiry Learning, English Digital Curriculum, SDGs Teaching Strategies, Learning behavior analysis

1. 前言

關鍵能力（key competencie）的重要性包括知識、技能和態度的綜合，並且可以在不同情境中運用，是可移動的（portable）、多功能的，能夠解決不同問題，對未來發展至關重要。許多研究認為問題解決、合作學習和溝通是重要的能力；適應力、文化契合性和成長潛力亦不可或卻 從教育的角度來看檢視，台灣的 108 新課綱與美國 21 世紀關鍵能力聯盟(Partnership for 21st Century Skills)提出的能力有高度相似性(曾子旂, 2019; Beckett & Miller, 2006; Larmer, Mergendoller, & Boss, 2015; Chang, Kuo, Hou, & Koe, 2022）。從未來人才所需的三大面向關

鍵能力觀點來檢視，包括學習與創新、數位素養、生活與解決問題能力。而創造力是指產生新穎有價值想法的能力，這種能力可以透過橫向思考，類比新的方案，進行推理，解決開放式的問題，並運用判斷來產生新的想法。創造力不僅限於英語領域，它也是科學探索和日常問題解決的重要工具，兒童創造力的培養，可透過遊戲和實際操作進行培養策略來支持和鼓勵兒童的創造力發展，包括動手做、分享靈感、迎接挑戰、建立自信和欣賞創意。

EMI 係指在英語非母語的教育機構（non-English speaking institutions）提供的學習課程，其內容的傳遞、師生互動、學習及學術支持教材、學習材料以及學習成果的展示和評量（such as oral presentation, assignments, or tests）都應 100%使用英語。本研究於教學現場實際觀察，學生對於英文課的參與度與積極度普遍不高。多數學習者將學習英語的負面失敗經驗與自身能力產生錯誤歸因，不當的認知投射到學習英語的心態。因此如何改善學生對英語的既定印象，增進學生的正面學習經驗、專注力、學習動機與學習成效，是需要深入探討的議題(Ansyari, 2015; Huang, Yang, Chiang, & Su, 2016; 許雯晴，2022; Raygan & Moradkhani, 2022; Taghizadeh & Basirat, 2022）。本研究主要探討學生能具有關鍵能力，可以運用到不同的情境、達成不同目標、解決不同問題的能力，包含知識、技能、態度所組合成的能力，透過在日常生活中的經驗積累知識，進而將知識與生活情境做連結，藉由研究者自行設計的跨領域議題探究英語課程，以海洋污染議題為主軸，讓雙語教育進入學生的課堂中又能更加對海洋國家的議題進一步認識。課程中使用數位學習的形式，並透過教師從旁引導輔助的策略，探究資訊融入議題探究課程對於學習者的學習動機與學習成效之影響，以及學習者在學習過程中的是否具有學習行為的改變，如：課程的認知負荷、參與度及創造力是否有所變化、對環境保護的意識是否提高等。因此，本研究目的在於創立一套跨領域的英語學習課程，並探討數位科技導入英語教學的應用，對於學生學習行為、學習動機及學習成效的影響，盼望能激發學生對知識探究的渴望，培養學以致用的態度與問題解決的能力。基於此，本研究的問題為：

1. 應用數位遊戲於議題探究 EMI SDGs 課程，對學習者之學習成效？
2. 應用數位遊戲於議題探究 EMI SDGs 課程，對學習者之學習動機？
3. 應用數位遊戲於議題探究 EMI SDGs 課程，對學習者之學習行為有何影響？能否提高學生參與度、創造力、環境保育意識、認知易用性與認知有用性，降低英語課堂對於學生的認知負荷？

2. 文獻探討

本研究評析彙整傳播理論科技接受模型及教育理論，並將其應用於本研究的實驗課程當中，以探究平板、電子繪本、數位遊戲以及數位學習平台作為輔佐學習者的數位學習工具，檢視學習者的學習動機與學習成效是否有正面影響，以及其學習行為的改善。科技接受模型由（Davis, 1989）提出，用來解釋使用者是否願意使用特定科技資訊系統，並透過使用者的認知有用性與認知易用性分析使用者後續的使用態度、意圖與實際使用行為的關係，找出影響使用者接受新科技資訊系統的因素，做為判斷使用者行為意圖的預測工具。

建構主義盛行於 1970 至 1980 年代，由認知心理學家 J. Piaget 與美國哲學家暨教育家 John Dewey 為其創始者。主張人類的學習行為，除了受到內在認知結構影響外，外在的環境也是一大因素。Dewey（1933）認為知識是從解決問題的過程形塑而成，並將解決問題的歷程分為五步驟：發現問題、定義問題、提出解決問題的假設、推論假設與驗證假設。建構理論教學本著以人為主、以學習為重的核心理念，以學生為中心，重視整體知識構建過程的主動性與評量的多元化，使學生有更高的動機去貼近英文與海洋議題等等生活化的題材(Zhang, 2022; Njiku, 2023）。

本研究之課程希望將構建主義的精神結合研究者自身英文課程的班級，在國小學童的學習場域中分組討論與情境融入參與，進行海洋議題教育的英文數位課程，期許能讓學生結合自身對於海洋的了解與知識，更進一步與英文課程做連結，達到雙語教育的同時又能夠對於海洋知識加深加廣；同時結合平板結合電子白板的功能，以及數位遊戲導入，提升互動性與專注度；又能兼具數位遊戲所帶來的動機與學習成效顯著提升。

3. 研究方法

本研究針對國小五年級學生，將平板結合電子繪本、數位遊戲與數位學習平台，藉以探討科技資訊媒介導入議題探究與 SDGs 英語課程，對於提升國小五年級學生在英語科之學習動機、學習態度與課堂學習行為之影響。本研究採前實驗研究設計法（pre-experimental design）進行單組前後測設計（one-group pretest-posttest design）之研究，針對國小五年級四個班共計 84 名學習者，以資訊科技輔助議題探究英語課程的方式，以平板、電子繪本、數位遊戲、數位學習平台做為學習者的數位學習媒介，施以三週共三節課的實驗課程，以此比對學習者在參與課程前、後的差異。每位受試者在參與課程前後皆有作答英語認知測驗，且在參與本實驗前皆無參與過議題探究英語課程，亦未接受過結合多項資訊科技做為學習媒介的課堂，研究有效問卷數量為 80 份。四班級的學習者皆會有 PowerPoint 投影、影片資源、心智圖、學習單、聽力練習等學習輔助工具，以確保學習者的學習權益。在後續的研究觀察，將會依照學生程度區分為「高」、「中」、「低」三群而加以分析，以檢視不同程度的學生，在學習行為與成效上的進度幅度。

本研究為使授課班級於其他課堂能正常上課，故受試班級由研究者根據自身對研究班級的了解，主觀判斷選取最適合研究目的的樣本，為立意抽樣，故本研究為前實驗研究法。資料分析的方式則包含質化與量化：質化工具為「教學紀錄手札」；量化工具為涵蓋了英語學習動機、認知負荷、創造力、環境保育意識、認知有用性、認知易用性等面項的「問卷量表」與學習成效檢測的「英語認知測驗」。資料蒐集方法與工具上，本研究之英語學習動機量表改編 Keller 的 IMMS 量表（Instructional Materials Motivation Survey questionnaire）共 36 題，由注意力（Attention）、相關性（Relevance）、自信程度（Confidence）與滿意程度（Satisfaction）四個構面來評估學習者之學習動機。認知負荷評量的主要目的是測量受測者在學習過程中，自身所知覺到的認知負荷感，藉以評估學生的心智努力程度。依據 Paas（1992）和 Sweller, van Merriënboer, & Paas（1998）所提出的觀點，可分為心智負荷（Mental load）與心智努力程度（Mental effort）兩個構面，以評估學習者的認知負荷量表。創造力量表參考 Furtwengler（2021）使用的創造力導向量表（Creativity Orientation Scale, COS）（Scott-Barrett, Johnston, Denton-Calabrese, Mcgrane, & Hopfenbeck, 2023）共 30 題，分為創造力識別（Creative Identity）、創造力偏差（Creativity Bias）、不確定性（Uncertainty）、社會支持（Social Support）四項構面。本研究為測量學習者在實施實驗課堂前後之環境意識是否提高，參考 Dunlap, Van Liere, Mertig, & Jones（2000）提出的新生態典範量表（New Ecological Paradigm Scale, NEP Scale）做為環境保育意識量表，共 15 題。新生態典範量表（Dunlap & Van Liere, 1978）被廣泛應用於某活動或是干預行為的影響之前後測，如教育課程對環境世界觀的影響，也是目前應用最廣泛的環境價值觀與態度衡量標準（Anderson, 2012）。認知有用性與認知易用性量表採用 Davis（1989）科技接受模型的認知有用性與認知易用性量表（Initial Scale Items for Perceived Usefulness and Perceived Ease of Use）共 28 題。

在學習者參與實驗課程前後，分別發放「英文認知測驗」與「問卷量表」做為本研究前、後測資料蒐集，並於後續資料分析探討其中差異。本研究所實施的課程，參考 108 課綱 19 項議題融入課程，透過每週一堂的「彈性學習課程」，主以海洋教育、輔以資訊科技，將議題

融入英語教學，教案教材與課程規劃為研究者自行設計，共設計三堂資訊融入的議題探究英語課程，每堂課實施時間為 40 分鐘。Dewey（1933）的建構主義強調學習者從實作中學習，主動建構知識，套用於本研究的實驗課程中有兩部份，一為第二堂課學習者使用平板查找海洋污染之圖片、影片、新聞做為支持自身論述的佐證，二為第三堂課小組活動，學習者需與組員一同合作，針對不同的解決方案提出認同或不認同的觀點，過程中可透過平板搜尋資訊、與組員們討論等方式，建構自身的學習脈絡，強化整體學習效益，資訊科技融入議題探究英語課程之簡易課程活動。三堂課具有連貫性，課堂分別對應的教學內容與教學目標如表 1：

表 1

議題探究英語課程分別對應的教學內容與教學目標，資料來源：許雯晴整理

堂數	教學內容	教學目標
1	十個海洋生物英語詞彙	能正確認出並聽懂十個海洋目標單字。
2	八個一次性塑膠產品單字 七大感官動詞與句型	將前堂課的十個目標單字，加上此堂課所學，運用感官動詞描寫出正確且符合圖片畫面的句子。
3	七項減緩海洋污染的片語 表達看法的句型	結合前兩堂課的單字與此堂課的片語，嘗試用英語表達自身該如何減緩海洋污染之看法。

本研究之課後學習成就檢測將分成「形成性評量」與「總結性評量」。前者將於每堂課結束前十至十五分鐘的課程活動中實施，包括「聽力練習單」、「seesaw 平台個人創作」、「seesaw 平台小組創作」；後者則於本研究實驗課程開始前、後十五分鐘施測，作為本研究學習成效分析依據。檢測內容為研究者根據此實驗課程教學目標與授課內容自編「英語認知測驗」。除了從量化工具來檢視學生課堂參與度是否提升，研究者將於課程實施期間，將針對學生課堂參與、同儕互動、課堂中的肢體語言進行觀察，並於授課完畢當日，由研究者回想學生於上課過程中的學習行為表現，撰寫成「教學紀錄手札」，做為學習者參與度的檢測工具。質化工具為教學記錄手札，將資料所描述的特定事件或情境，透過編碼與設計代碼等方式分割成有意義的組織元素，並賦予每個事件的主題與代碼，以便研究了解研究對象於實施課堂過程中的學習行為。

4. 研究結果分析與討論

本研究探討議題探究融入 EMI 數位遊戲課程對於學習者的學習成效、學習動機、學習行為、創造力、認知易用性、認知有用性、認知負荷及環境保育意識是否產生正面影響。對於學習高成就以及英語程度中等的學習者，有明顯差異；對於部分學習低成就的學習者，選擇題的答對率增加，填答意願也有所提高。本研究觀察學習成效的檢測依據為「英語認知測驗」、「創作作品」、「聽力檢測」以及「口說音檔」，英語認知測驗包含實驗課程的學習目標與學習內容；創作作品為教案中的角色扮演作業，學生設身處地思考海洋生物可能會遇到的情境，並發揮想像力運用課堂所學的單字與文法，完成此項個人作業；聽力檢測為第一堂的課程，每題英語聽力檔案包含其中兩個海洋生物的目標單字，共有三題，學生將填寫聽到的單字於學習單中，並於課程進行時點名回答，以此檢視學生對於目標單字的聽力能力；口說音檔的部分則為第三堂課的活動三，套用句型來錄製影片對於海洋污染的看法。

「英語認知測驗」的部分，整體平均從 55.17 進步到 80.46。學習成就高的學生，前測的成績約落在 70 至 80 分左右，後測的成績約落在 100 至 110 分之間，共有三位同學獲得滿分，前後測總平均的數值四捨五入後分別得到 78.54 與 102.92 分，兩者相差 24.37 分。學習成就中等的學生，前測的成績約介於 50 至 60 之間，考到 65 分者有 3 人，前測平均值為 56.5；後

測的成績約落在 75 至 85 之間，考到 90 分的人數有 3 人，後測總平均為 80.63，整體分數進步為 24.13。而學習成就低的學生，前測總平均為 29.6，前測的成績約落在 20 至 40 間，部分學生可以拿到 45 分，以獲取 35 分的人數為最多，但仍有學生僅能得到 10 分；後測約為 50 至 70 之間，後測總平均為 57.7，整體進步分數為 28.1 分。觀察前後測的作答狀況來解釋此現象，發現因為程度高的學習者原先就已具備部分單字基礎，但後測相較於前測，依舊可以看出學生提升英語認知能力。以「使用環保餐具」的題目為例，前測時學生大多能撰寫「use」來表達「使用」的意思，而關於「環保」的翻譯可能是直接空著或是寫「eco」與「environment」來表示，至於「餐具」則寫出「knife」、「fork」或「chopstick」等詞彙，沒有人寫出總稱「tableware」；而在後測的作答狀況發現，學生會使用上課所講的「eco-friendly」以及「reusable」來表答「環保」之意，也能正確的撰寫出「tableware」來表達「餐具」。至於 2 題的句子翻譯與 2 題段落閱讀，前測時因為缺乏對目標單字與文法的認知能力，故句子翻譯無法全數得分，但亦能從作答中看得出學習高成就的學生已具有基本單字與文法概念。例如題目「我查覺海鷗正在咬寶特瓶蓋（plastics caps）」會翻成「I see a seagull bite the plastics caps.」。

「創作作品」的部分，從三個標準評估學生的學習成效，各標準之詳細說明如表 3 所示，若學生三者皆有達成，則獲得「A」等級，即為 100 分；三者達成兩者，則屬於「B」等級，即為 90 分；僅達成一者，則為「C」等級，即為 80 分。將學生依照英語程度高中低檢視，發現不論程度高低，大部分的學生能獲得「A」等級，加上因為是圖像創作，學習成就低的學生投入程度相較於平時英文課來得高，也更願意花時間去完成，這部分能從其作品製作與呈現方式觀察到，因此創作作業並沒有「C」等級的學生；而獲得「B」等級的學生，英語程度高中低的皆有，可見其等級與英語程度並沒有關聯。觀測變項與檢測方式對應如表 2。

表 2

本研究觀測變項與檢測方式對應，資料來源：許雯晴整理

觀測變項	觀察與檢測方式
學習成效	英語認知測驗、創作作品、聽力檢測、口說音檔
學習動機	問卷量表、課程回饋單
學習行為	課堂筆記、課堂參與度、回答問題
創造力	問卷量表
認知易用性	問卷量表
認知有用性	問卷量表
認知負荷	問卷量表
環境保育意識	問卷量表、課程回饋單

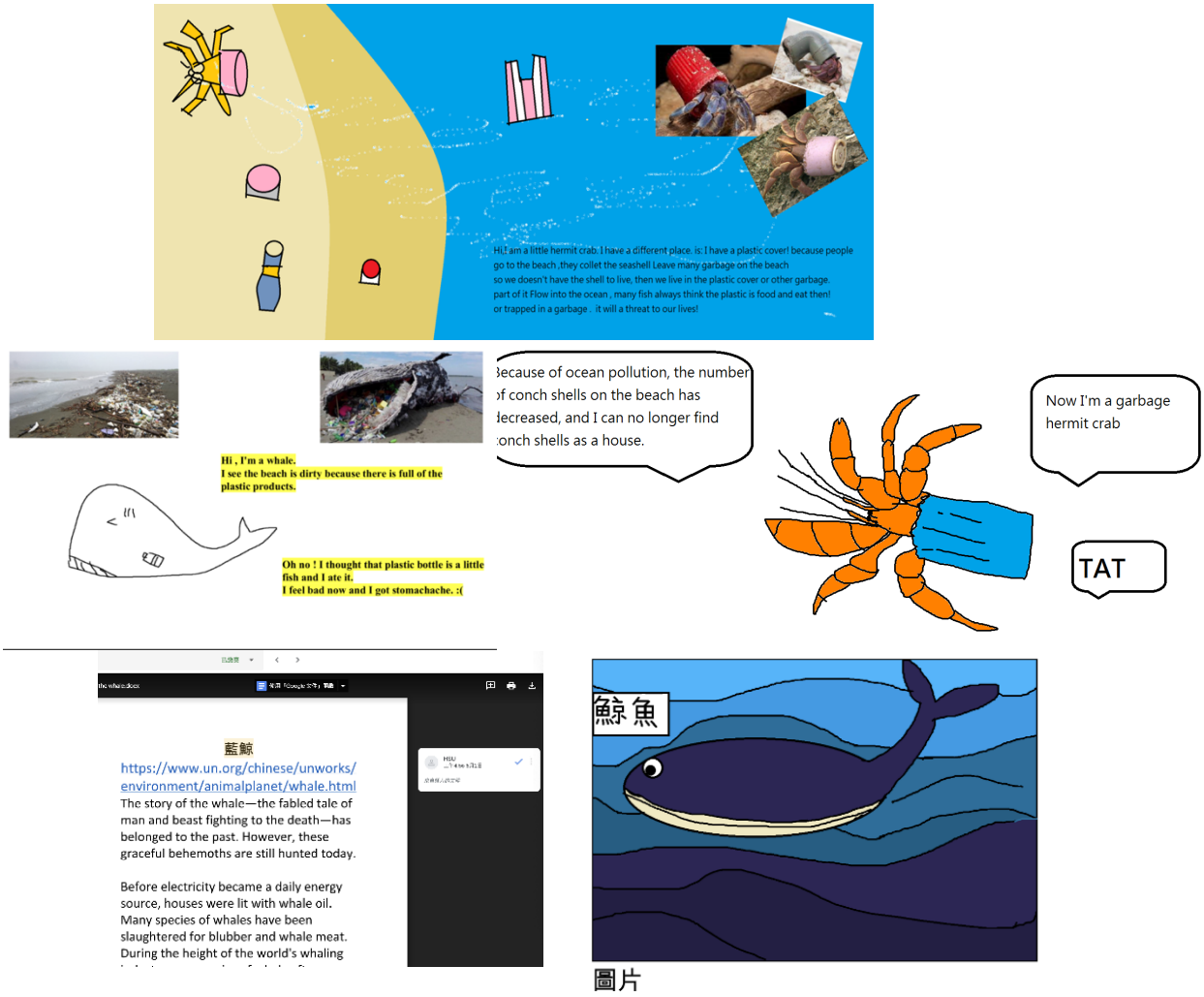
表 3

創作作品評估標準之詳細說明，資料來源：許雯晴整理

評估標準	詳細說明
句子陳述	撰寫的句子使用到目標單字或是至少一個海洋生物單字，且沒有拼音錯誤，即符合。
佐證資料	佐證資料與繪製的圖像有所關聯，以此檢視學生查找資訊的能力。若是畫海龜結果找螃蟹的圖片，或是找全英文的新聞但是在講天氣預報，則不符合。
圖畫用心程度	採用電腦、紙筆或是其他創作方式皆可，但作品需有一定完整度。若是句子只寫一半、畫圖筆跡草率、看不出畫作傳達的訊息或是主題與本次作業不相關等情況，則被歸為「作品不完整」。

圖 1

英語程度高階學生的創作作品



圖片

再仔細觀察不同程度的學生於作品中的句子陳述，發現程度高的學生會傾向於使用更多課外或是他所學過的單字，撰寫更多篇幅的敘述，以此來將自己想像的故事或想法陳述出來，或者，會嘗試使用課外文章作為佐證資料，如上圖 1 所示。從作品中亦可觀察出，即便是高程度的學生，依舊有可能有大小寫或是標點符號使用錯誤的情況發生，但整體而言並不影響觀者對於作品訊息的理解程度。

「口說音檔」的部分，參考初級英語檢定考試口說標準的評分項目，包含其「發音」、「語調」、「句子的流暢度」以及「文法、詞彙使用的正確性」來給予學生回饋。口說內容為問句「How to reduce marine pollution?」以及自身對於此問題的看法。此部分的課程設計旨在讓學生能將第三堂課所學的「七大片語」以及「表達觀點的句型」活用於日常生活中。「問卷量表」的部分，學生在課程實施前後的學習動機相差的 t 統計值為-5.801，雙尾顯著性 p 值 =0.000<0.05，如表 5 所示，故拒絕虛無假說，「在本實驗課程實施前後，學生的學習動機有顯著差異」成立。整體而言，不論是學習程度高或低的學生，在參與本實驗課程後，其「學習成效」與「學習行為」皆有不同幅度的改善，英語能力在聽說讀寫四項不同層面亦有所提升。從問卷量表來分析其餘的六項變項，除「認知有用性」前後差異並不顯著外，「學習動機」、「創造力」、「認知易用性」、「認知負荷」、「環境保育意識」皆有顯著改善，如表 6 所示；其中「學習動機」與「環境保育意識」兩個面項，亦可從學生的課程回饋中呼應。

表 5

學習動機之成對樣本檢定

成對 1	成對變數差異					t	自由度	顯著性(雙尾)
	平均數	標準差	平均數的標準誤	差異的 95% 信賴區間				
				下界	上界			
學習動機 a 學習動機 b	-.46196	.71231	.07964	-.62048	-.30345	-5.801	79	.000

表 6

本研究觀察面項之顯著性與不顯著性統整

觀測變項	雙尾顯著性（p 值）	顯著	不顯著
學習動機	.000(***)	V	
創造力	.042(*)	V	
認知易用性	.000(***)	V	
認知有用性	.672		V
認知負荷	.011(*)	V	
環境保育意識	.000(***)	V	

5. 結論

本研究創立一套跨領域英語教學教案，分析了 EMI 英語數位課程、SDGs 教學策略與關鍵力教育的融合，搜集了學生的學習成效、學習動機和學習行為的數據，進行統計分析，以評估資訊科技融入教學對學生的影響。研究發現，不論學習者程度屬於學習成就高或是學習成就低，在成績評量的分數上約進步 25 分，學習後段的學習者進步的原因為目標單字的選擇題答對率變高，填答試卷的意願也有所提升，而學習成就高的學習者則是中翻英的句子文法與單字拼法正確而得分。創作作品的評分機制按照句子陳述、佐證資料、圖片用心程度三個標準給予學生級別，這部分學生的英語程度與作品等級並沒有關聯性，大部分的學生能獲得最高等級 A，而作品中可以觀察到程度高的學生偏好使用課外詞彙與較多的文字來完整呈現其想傳達的訊息，學習程度中後段的學生則傾向使用課堂中的目標單字與文法來完成其作品，從學生繳交的作品中也可觀察到後段學生的態度轉變，從消極轉為積極。而課程回饋單有同學對於使用平板當作課本上課感到有趣、對海洋生物跟塑膠污染相關的單字有所印象，可見學生在參與完本研究的實驗課程後，察覺到英文對於日常生活的重要性，並且有想要提升自己英語能力的意願。因此，透過科技與數位遊戲導入主要能提升互動性與專注度，對於 EMI 與 SDGs 課程三者具有交互作用並影響學習成效。最後，學習行為面項根據問卷量表的檢測結果顯示，學生的「創造力」、「認知負荷」、「認知易用性」以及「環境保育意識」，在實施課程前後皆有顯著差異。

參考文獻

曾子旂（2019）。從二十一世紀關鍵能力的養成論以專案式學習促進核心素養。**課程研究**，**14**(2)，85-106。

許雯晴（2022）。**國小五年級議題探究英語數位課程對學生學習成效、學習動機與學習行為之影響：以科技接受模型的觀點**，國立政治大學傳播碩士學位學程未出版碩士論文。

Anderson, M. W. (2012). *New Ecological Paradigm (NEP) Scale*. Berkshire Publishing Group.

Ansyari, M. F. (2015). Designing and evaluating a professional development program for basic technology integration in English as a foreign language (EFL) classrooms. *Australasian Journal of Educational Technology, 31*, 699–712.

Beckett, G. H., & Miller, P. C. (2006). *Project-based second and foreign language education: Past, present, and future*. Greenwich, CT: Information Age.

Chang, C. H. S., Kuo, C. C., Hou, H. T, & Koe. J. J. Y. (2022). Design and Evaluation of a Multi-Sensory Scaffolding Gamification Science Course with Mobile Technology for Learners with Total Blindness. *Computers in Human Behavior, 128*, 107085.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340.

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Boston, MA: D.C. Heath & Co Publishers.

Dunlap, R. E. & Van Liere, K. D. (1978). The new environmental paradigm. *Journal of Environmental Education, 9*(4), 10-19.

Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, E. R. (2000). Measuring endorsement of new ecological paradigm: A revised NEP scale. *The Journal of Social Issues, 56*(3), 425-442.

Furtwengler, S. R.(2021). Development of a creativity orientation scale using EFA. *Journal of Creativity, 31*, 100004.

Huang, C. S. J., Yang, S. J. H., Chiang, T. H. C., & Su, A. Y. S. (2016). Effects of Situated Mobile Learning Approach on Learning Motivation and Performance of EFL Students. *Journal of Educational Technology & Society, 19*(1), 263–276.

Larmer, J., Mergendoller, J., & Boss, S. (2015). *Setting the standard for project based learning: A proven approach to rigorous classroom instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.

Njiku, J. (2023). Attitude and technological pedagogical and content knowledge: The reciprocal predictors? *Journal of Research on Technology in Education, 55*(6), 1020–1035.

Paas, F. G. W. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*, 429434.

Raygan, A., & Moradkhani, S. (2022). Factors influencing technology integration in an EFL context: investigating EFL teachers' attitudes, TPACK level, and educational climate. *Computer Assisted Language Learning, 35*(8), 1789–1810.

Scott-Barrett, J., Johnston, S., Denton-Calabrese, T., Mcgrane, J. A., & Hopfenbeck, T. N. (2023). Nurturing curiosity and creativity in primary school classrooms. *Teaching and Teacher Education, 135*, 104356.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

Taghizadeh, M., & Basirat, M. (2022). Investigating pre-service EFL teachers' attitudes and challenges of online teaching. *Computer Assisted Language Learning*, 1–38.

Zhang, W. (2022). The Role of Technology-Based Education and Teacher Professional Development in English as a Foreign Language Classes. *Frontiers in Psychology, 13*, 910315.

如何利用實境解謎遊戲激發後設認知以提升價值教育

How to use Alternate reality game to stimulate metacognition to enhance value education

李浩然

香港聖公會聖馬利亞堂莫慶堯中學公民、經濟與社會科科主任

香港教育大學課程與教學系兼任講師

lhy@smcc.edu.hk

【摘要】本文章主要探討在實境解謎遊戲中，後設認知在解謎過程中的認知機制，並說明在言情況下，學生在遊戲中的後設認知過程，塑造價值觀的的可能性。本文探討會以哲學家 John McDowell 的價值理論為基礎，說明價值概念是由事件/環境通過內省的過程而形成，在此，我以為 McDowell 所指的內省過程也就是後設認知 Metacognition 過程，文章中亦會以心理學家 Borkowski, J. G 的後設認知模式印證其可能性。同時，文章亦能說明以實境解謎的遊戲式學習，能達致有效的價值教育的可能。

關鍵字：價值觀; 次性; 後設認知; 實境解謎遊戲

Abstract: This article mainly discusses the cognitive mechanism of metacognition in the process of solving puzzles in Alternate reality game and explains the possibility of students' metacognitive processes in the game shaping values under the circumstances. This article's discussion will be based on the value theory of philosopher John McDowell, which explains that the concept of value is formed by events/environments through the process of introspection. Here, I think the introspection process McDowell refers to is the metacognition process. The article will also use the metacognitive model of psychologist Borkowski, J.G. to prove its possibility. At the same time, the article can also illustrate the possibility of effective value education through game-based learning that uses real-life puzzle solving.

Keywords: values, secondary nature, metacognition, Alternate reality game

1. 前言

是次探討主要提出價值觀的塑造與獲得，與後設認知有密切的關係，而實境解謎遊戲，既是通過與實際環境的互動達到學習效果，同時解謎的過程，離不開需要後設認知的運用，才能解決謎題，因為學生的價值教育可以利用實境解謎遊戲有效地建立正向的價值觀的可能性。因此，是次不會以量化的研究作探討方向。

2. McDowell 的價值感知模式

我們的價值觀，例如幫助別人是一項良好的行為是怎樣學會？又或者，我們看到這種行為時，怎樣學會判斷這是一項好行為呢？McDowell 將價值觀的獲得，以我們獲得顏色的概念為例：一件事物具備甚麼顏色，是獨立於我們對該事物的經驗；但要分辨所經驗到的顏色，則我們必須具備顏色的概念。例如一件

東西是否有紅色的性質，是獨立於我們經驗這東西，縱使沒有經驗該東西，它所具備的顏色性質也不會消失。可是，我們判斷該東西是否為紅色，是我們的視覺神經系統在正常的環境，與物體表面的質地與光作出不同程度的反射互動所產生的。

同理，一項價值觀除了我們具備相關的概念外，還包括客觀的環境或事實。即價值觀是在正常的情況或條件之下被所引起的。因此，價值觀並不能抽象地，或獨立於事實，我們需要在相關的環境與事實互動下才能形成及獲得。

例如：如果我們經驗到有一位年青人攙扶一位行動不便的老人過馬路，並能解到這是一件美德的事情，那麼，除了要存在該事實〔有一位年青人攙扶一位行動不便的老人過馬路〕外，我們也需要具備的相關的概念〔例如：年青人、老人、過馬路和美德等〕，才能認識到年青人做了一件美德的行為。

然而到此，我們可能會問：我們如何從經驗到一位年青人攙扶一位行動不便的老人過馬路這行為看到這是美德的行為呢？McDowell 認為一項行為具有的在價值性質，可以把它理解為一種「意向性質」(dispositional property)。何謂意向性質呢？這種性質只有在特定的條件與環境互動之下才會出現。就好像鹽巴可溶解這種性質正常來說，並不會直接在鹽巴中找到，但當鹽巴被放在水中時，可溶解的性質才出現，而「可溶解」是便是鹽巴的「意向性質」。同埋，McDowell 認為一項事實如果具有價值的性質也是像一種次性，必須通過某種「條件」才能被經驗到。而我以為這「條件」可以用後設認知來說明為何我們能經驗到一項具有

價值的行為，或價值觀的形成。

3.Borkowski J. G.的後設認知模式

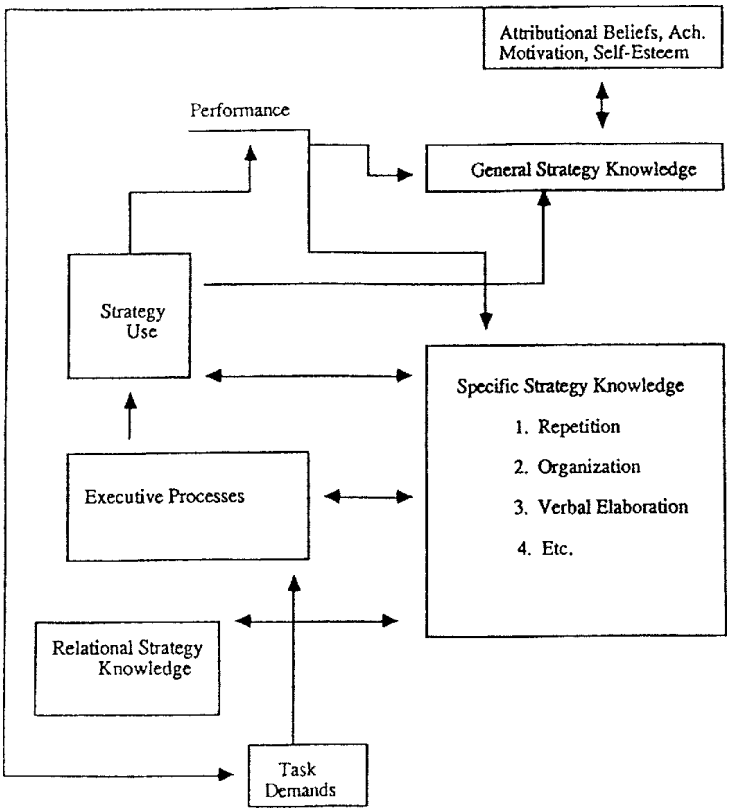
根據 Swartz 及 Perkins 定義，後設認知 Metacognition 指的是對認知的認知一種個人控制及引導心智歷程的現象。[1]即個人對自己的認知歷程的掌握、控制、支配、監督與評鑑的一種知識，也就是認知你的認知，或對你的思考進行思考，或反思；例如，當你經驗到一位年青人攙扶一位行動不便的老人過馬路這行為，你會告訴自己我正看到一位年青人攙扶一位行動不便的老人過馬路，並且你可能會告訴自己對這行為的感受，以及給予自己這行為的評價，並向自己說明有這些評價的理由等。因此之故，我們可以看到從一項事實到一項事實的價值判斷，必須通過這個認知歷程，相關的價值性質才會出現。

根據心理學家 John Flavell 的說法，「後設認知」可以分為「元知識」及「元經驗」兩部份。後設認知是一種個人反省自我認知過程、成果和自我調適的能力。具體的說，經由認知思維從事求知活動時，自己既能明確了解他所學知識的性質與內容，而且也能了解如何進一步支配知識，以解決問題。[2]這種歷程，除了我們可以用於對認知的事實進行價值判斷，我們也可用於學習價值概念的策略，讓我們瞭解到自己的經驗同時，透過控制自主的思想模式，把經驗與價值概念之間產生關連，從而達至建立個人價值觀的學習方法。

我們可以根據心理學 Borkowski, J. G.所提出的後設認知模式 ，看到我們以後設認知從經驗獲得價值觀的可能性。在以下圖中，提及的特定策略知識、執行控制及策略運用，是價值觀形成的重要因素。

[1] Burke, 1999, p.128; as quoted from Barell, 1992, p.258

[2] Flavell, J. H.(1976). Metacognitive Aspects of Problem Solving. In Resnick L. B. (Ed), The nature of intelligence, Hillsdale, New Jersey: Lawrence Erlbaum.



Borkowski, J. G.所提出的後設認知模式

特定策略知識 Specific strategy knowledge 指與某一策略相連的知識，例如：價值概念。執行控制式 Executive processes 指對認知過程的控制，策略的選擇及監控 Cognitive monitoring，也就是說，我們經驗某項事實時，如何將價值概念作關連。最後，策略運用 Strategy Use 指運用特定策略知識與執行工作，即在價值認知時，如何將價值概念整合在經驗事實的策略。

由此可見，後設認知是價值觀的形成有著關鍵性的能力，如果讓學生提升對價值認知的後設認知能力，學生不但可以準確認識正向的價值概念，從而能夠作出準確的價值判斷。而我以為實境解謎遊戲是一種理想而有效的活動，幫助學生提升價值認知的後設認知能力。

4. 實境解謎遊戲、解謎與學習

實境遊戲是指運用現實世界的實際環境作為部分遊戲情境的一種遊戲形式，例如：公園、名勝古蹟、街道等。當遊戲是一個發生在商場的情境，而玩家也是

在商場裡面，該玩家就是一個參與實境遊戲的人，他必須在商場的指定地點進行遊戲，完成不同的遊戲劇情與任務。這些遊戲與任務，通常都會以一些謎題或解碼方式與環境互動。過程中也必須運用不同的關鍵能力，例如：邏輯推理、資訊收集和分析，或辨識能力等。以得到通關的提示。解謎遊戲具有統整複合技能、運用多元的策略等特色，讓玩家身歷其境體驗挑戰[3]。

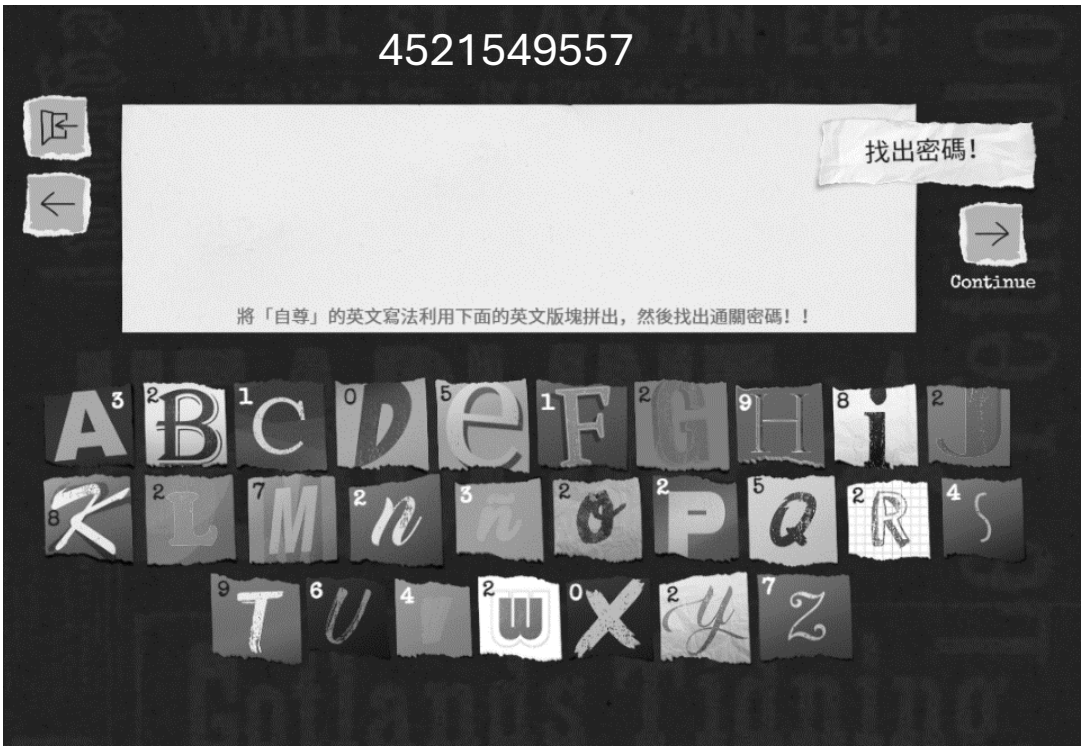
在設計上，實境遊戲以解謎入劇情等設計，使玩家可以沉浸於其中。其沉浸的特點，主要特徵是：在活動過程中完全的集中精神；二為從活動經歷享受，此經驗所帶來的效果則是會讓使用者比較注重過程而非結果[4]。因此學生在實境遊戲的過程中，可以專注於關鍵能力的過程，也即後設認知的思考過程。

由於實境遊戲著利用身體的感知（embodied cognition），美國哲學家 Mark Johnson 認為以身體感知的觀點，他把身體視為一種具有感官動能（sensorimotor）的有認知功能，意義是身體與環境互動之認知過程所衍生的（鄧育仁、趙綺芳、余舜德，2011）。而衍生的過程，我們都能夠利用 Borkowski, J. G.所提出的後設認知模式作出說明。即按 Borkowski 的後設認知模式，將身體在環境中的感知結果結合腦部演釋分析，調整認知策略，產生判斷認知的結果。謎題能幫助認知者專注、幫助以及協助的一種模式，使得在認知過科中提升後設認知中的分析能力，而解謎的過程中，認知者也得運用先備知識(特定策略知識)幫助解開謎題。不同類型的謎題可以視作為不同的學習鷹架，引導學生在把謎題的答案找出，如此也等於學生達到學習目標。

例如：學生認識「自尊」的英文拼法，我們可以讓他們先想提示一組數字及數字所對應的字母，讓他們把數字與字母之間的關係找出，從而得到通關答案。在解謎的過程中，學生運用後設認知將情境---通關要求、特定的策略知識---數字及字母，以及認知策略---排列，整個學習要求都在解謎的過程中。

[3] Markus Wiemker, Errol Elumir, Adam Clare (2015). Escape Room Games: "Can you transform an unpleasant situation into a pleasant one?" Retrieved from https://www.researchgate.net/publication/348870975_Escape_Room_Games_Can_you_transform_an_unpleasant_situation_into_a_pleasant_one

[4] Ghani, A.J., & Deshpande, P.S. (1994). Task characteristics and the experience of optimal flow in human-computer interaction. The Journal of Psychology, 128(4), 381-391.



轉碼式謎題

5.實境解謎遊戲與價值教育

由以上我們可以看到，解謎的題型能夠恰當地設計，便能有效地改善後設認知的能力，學習更準確的知識或概念。同理，如果按 McDowell 的價值認知模式，我們可以作類比能夠以實境遊戲中，通過解謎的活動有效培養學生正向的價值觀。

利用實境遊戲的好處，在於學生能夠有實的環境，強化後設認知能力。由於價值觀是從現實生活中實現出來，所以在真實的環境對學生連結事件與價值概念時更有效。例如：下圖是在香港某社區的負責管理公營房屋的機構，學生通過實境遊戲平台及手機的定位裝置，了解該機構在社區的位置以及機構的功能，進一步推論該機構在社區的作用，找出/實現對市民有何價值處。



實境遊戲平台

利用實境解謎的方式，除了能讓學生在解謎過程中，幫助掌握事實與價值之間的關連外，也能讓學生在解謎過程中更專注於觀察環境。教師在設計題型時，更能將想要學生注意的考察點變成解謎的內容，使學生更清楚知道題目所帶出的內容與價值的關連性。而以實境作為學生解謎的考察點，可以減低他們後設認知的操作過程，減低認知負荷。

6.總結

價值教育有別於其他科目的教學方法，一般的科目，我們可以通過課本，或者不同的學與教策略，完成教學目標。例如：學生學習數學可以在與現實生活沒有關聯下，掌握數學概念與知識。比如，學習國文科，雖然文章內容可以與生活有關聯，但學生在掌握文章意境時，也不必與現實生活互動，也能掌握當中的意義。

然而，建立價值觀必須與現實情況結合，才能掌握事件中的價值意涵。另一方面，事實本身的價值性質，必須在當下與事實真實互動，後設認知的認知模式，把事實中的價值實現出來，如此，學生才能準確掌握事實中的價值意涵。

參考文獻

Burke, (1999), p.128; as quoted from Barell, (1992). p.258

Borkowski, J. G., Estrada, M. T., Milstead, M., & Hale, C. A. (1989). General problem-solving skills: Relation Between Meta-cognition and strategic processing. Learning Disability Quarterly, 12(winter), 57-70.

Flavell, J.H.(1976). Metacognitive Aspects of problem solving. In Resnick L.B.(Ed), The nature of intelligence. Hillsdale, New Jersey: Lawrence Erl-baum

Ghani, A.J., & Deshpande, P.S. (1994). Task characteristics and the experience of optimal flow in human-computer interaction. The Journal of Psychology, 128(4), 381-391.

John Locke, (1975). An Essay concerning Human Understanding, Book II, Chapter VIII, §10, §1, §15.

John McDowell, (1994). Mind and World, Harvard University Press,

John McDowell, (1987). 'Values and Secondary Qualities', in Mind, Value, &, Reality, Harvard University Press, 1998.

John McDowell, (1987). 'Projection and Truth in Ethics', in Mind, Value, &, Reality, Harvard University Press, 1998.

Markus Wiemker, Errol Elumir, Adam Clare (2015). Escape Room Games: "Can you transform an unpleasant situation into a pleasant one?" Retrieved from https://www.researchgate.net/publication/348870975_Escape_Room_Games_Can_you_transform_an_unpleasant_situation_into_a_pleasant_one

遊戲式學習應用於培育 21 世紀關鍵能力趨勢 – 應用文獻計量分析之初探

Game-based Learning for Cultivating the 21st Century Skills: A Preliminary Study on the Research Trends Using Bibliometrics Analysis

王舒民
中國文化大學資訊管理學系
wsm6@faculty.pccu.edu.tw

【**摘要**】近年來遊戲式學習被廣用於各種教學情境中。其被認為可有效促進學習動機。此外，透過模擬真實情境與問題，遊戲能使得學習者在以單人或多人協作完成挑戰與任務時，加深知識理解，並培養如：問題解決能力、批判思考與創造力、溝通、協作等被認為是面對 21 世紀多變複雜環境必備的關鍵能力。由此，本研究應用書目計量分析法，蒐集分析過去 30 年來遊戲學習應用於培養關鍵能力之相關文獻計 1,222 篇，以瞭解相關領域之研究趨勢與現況，俾能做為後續更為深入探究之起點，同時可供未來相關遊戲式學習應用於培養關鍵能力之研究做為參考。

【**關鍵字**】遊戲式學習；21 世紀關鍵能力；書目計量分析；文獻回顧

Abstract: *Game-based learning has been widely used in various educational settings. It is believed to effectively enhance learning motivation. Moreover, by simulating real problems, games enable learners to deepen their understanding of knowledge and cultivate key competences such as problem-solving, critical thinking, creativity, communication, and collaboration through completing challenges and tasks individually or collaboratively. These skills are considered essential for the 21st century. Therefore, this study employed bibliometric analysis to collect and analyze literature on the application of game-based learning in cultivating key competences over the past 30 years. This is done to understand the research trends and status in related fields, serving as a starting point for further in-depth exploration and providing a reference for future studies on the application of game-based learning in cultivating key skills.*

Keywords: Game-based Learning, 21st Century skills, bibliometrics, literature review

1. 前言

遊戲式學習為近年來備受關注的研究主題，其被廣用於各式樣的教學情境中。遊戲式學習透過設計針對特定學科或能力培養之遊戲提昇了教學場域之教學策略的豐富多樣性 (Hou, 2023; Wang et al., 2020)。針對遊戲式學習對於促進學習動機，學習投入，乃至於學習成效之優勢，歷年來亦已有許多學者提出相關的整理與見解 (Romero et al., 2014)。

另一方面，隨著時代推展，現今的環境變得更為動態與複雜，近年來常可見人們使用 VUCA 時代來總結當今環境的多變（Volatility）、不確定（Uncertainty）、複雜（Complexity），以及模糊（Ambiguity）。面對這樣的時代，過去學者提出了 21 世紀關鍵能力的概念，其包括了批判思考、創造力、資訊科技素養、溝通協作、社會文化技能等能力，期能培養學習者適應快速變化的技術、知識和工作需求。過去研究並據以設計各樣的教學策略、工作、框架來做為學習上的支持與引導 (González-Pérez & Ramírez-Montoya, 2022)。

綜合前述概念，遊戲式學習可作為一個有力的教學策略，用於培養學生快速適應未來的關鍵能力 (Qian & Clark, 2016)。本研究旨在蒐整分析過去相關研究文獻，初探現有的研究趨勢、重要的貢獻者、重要關鍵能力等。鑑於即有研究成果之浩瀚，本研究擬使用書目計量分析法（bibliometrics）進行分析，提供量化的文獻分析結果，做為後續研究方向此起點，期能符應本工作坊之主題，並做為未來深入探究與交流之基礎。

具體而言，本研究之研究目的即為應用書目計量分析法探討遊戲式學習應用於培養 21 世紀關鍵能力之研究現況與趨勢，並據以進行討論與探討未來研究方向。

2. 研究方法

本研究擬使用書目計量分析法進行研究分析。書目計量分析式使用量化的方式，用於分析較大規模與複雜的文獻關連、主題趨勢。主要的分析方式包括：引用分析、共被引分析和共現詞分析等，來探索文獻間的關聯性、知識結構、主題發展和作者關係等的應用。透過書目計量方法，有助於幫助研究者理解學科的現狀、趨勢、主要研究來源與議題等。

本研究使用 Web of Science 進行資料蒐集，並使用以下關鍵字組合進行文獻搜尋，搜尋欄位包括：主題、摘要、關鍵字等。隨後篩選搜尋結果為 Articles 及 Early Access 類別。本次搜尋，共得 1,222 文章，發表年份涵蓋 1994 – 2024 年 3 月間。其中，關鍵能力部份，除使用”21st Century skills”外，另參考 González-Pérez & Ramírez-Montoya (2022)所進行的文獻回顧中定義之關鍵能力，如：ICT literacy, social skills, communication 等關鍵字進行查詢。

> (“serious game*” OR “game-based learning” OR “educational game*” OR “MMORPG*” OR “board game*” OR “computer game*” OR “mobile game*”) AND (“learning” OR “education”) AND (“21st-century skills” OR "Communication" OR "Collaboration" OR "ICT literacy" OR "Social skills" OR "Creativity" OR "Critical thinking" OR "Problem solving" OR "Learning to learn" OR "Self-direction" OR "Flexibility and adaptability" OR "Risk taking" OR "Conflict management" OR "Sense of initiative and entrepreneurship")) AND (DT==("ARTICLE" OR "EARLY ACCESS"))

本研究主要使用 R 語言做為分析工具，使用套件為 bibliometrix 4.1.4 版 (Aria & Cuccurullo, 2017)。部份關鍵字比對與整理則另行撰寫 Python 程式語言進行處理。

3. 資料分析結果

本研究使用 Web of Science 進行資料蒐集，並使用以下關鍵字組合進行文獻搜尋，搜尋欄位包括：主題、摘要、關鍵字等。隨後篩選文件類別（Document type）為 Articles 及 Early Access 類別，以排除研討會文章、專書、專書章節等文件類型。使用前述條件搜尋後，共得 1,222 篇研究文章，發表年份涵蓋 1994 – 2024 年 3 月間。其它相關基本資料包括：1,222 篇文章共有計 3,965 名作者，平均每篇研究作者數為 3.67 位，而單一作者之研究則有 134 篇。此外，1,222 篇文章共引用了 47,120 篇參考文獻。

圖一顯示了各年度之研究文章發表篇數，並自 2006 年起，便每年均有相關研究發表，更於 2019 年時呈現明顯地躍進，由 2018 年的 43 篇增長至 2019 年的 77 篇，並逐年創下新高，直至 2023 年，且於 2024 年第一季便已有 29 篇相關文獻發表，年平均成長率達 14.08%，可見各界對於此一研究主題之關注度。

表一呈現了遊戲式學習應用與關鍵能力培育相關研究領域中的主要貢獻者。由表中的資訊，可觀察到前五名的主要貢獻者便有四名來自台灣的研究學者，也可看出台灣學者在此一遊戲式學習領域之投入與貢獻。表一中之加權平均為考量多名作者共同發表之影響，其計算

方式為一篇文章若有兩位作者時，則各計 0.5 篇，為三名作者時，各計 0.333 篇，並依此類推。經考量加權分數後，即可發現排序有所異動，亦可提供做為觀察研究貢獻之參考。

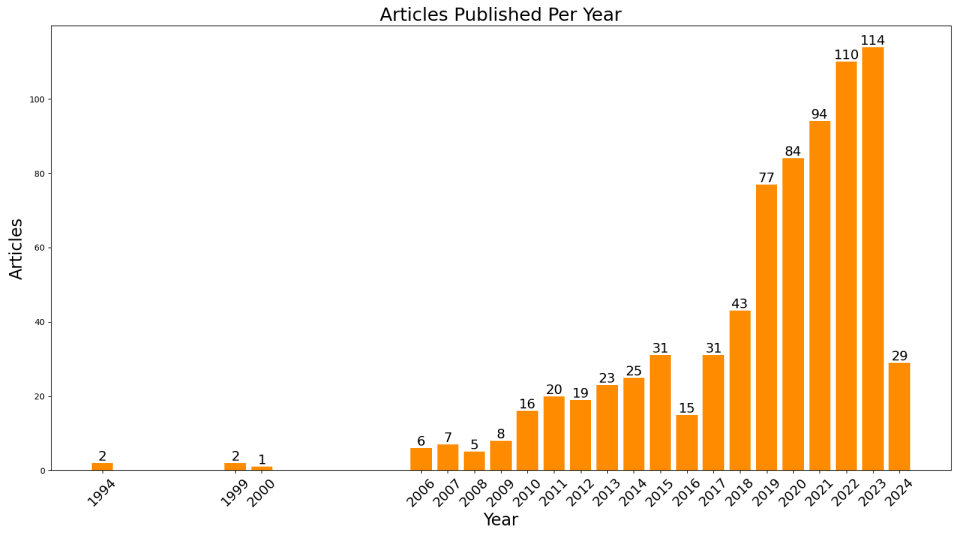


圖 1 各年度發表文章篇數

表 1 作者發表篇數及依作者數加權平均後之篇數

排序	作者	文章篇數	加權平均*	加權後排序
1	Hwang, GJ	17	5.57	2
2	Ke, FF	13	6.28	1
3	Hou, HT	11	4.00	3
4	Yeh, YC	10	3.67	4
5	Chang, CY	7	2.25	5

*: 加權平均為依文章作者數加權後，重新計算個人發表篇數

表二則就文章發表出處與篇數進行統計。由表二可看出最主要的發表出處為 Computers & Education，共歷年累積篇數遠超過排名第二後各期刊，顯見其於此領域之重要性與影響力。表二亦可提供未來相關學者在進行文獻蒐集與評估著作發表之期刊選擇時之參考，。

由於共現詞分析之結果多集中於較共同性的主題詞彙，例如：education, computer games, gamification, engagement 等。較未能提供有意義的資訊，故本研究另就 González-Pérez & Ramírez-Montoya (2022)所定義之關鍵能力於出現各文章摘要中之詞數進行統計。其出現次數如下表 3 所示。由表中可觀察到在遊戲式學習相關研究文獻中經常被提及的關鍵能力包括：communication, problem solving, collaboration, creativity, critical thinking 等，而 social skills 則接在後方。由表 3 亦可一窺相關研究對於各關鍵能力之關注程度。

表 2 文章發表出處及累積發表數

排序	文章發表出處（期刊名稱）	文章篇數
1	Computers & Education	80
2	Interactive Learning Environments	34
3	Journal of Educational Computing Research	28
4	British Journal of Educational Technology	26
5	Etr&D-Educational Technology Research and Development	26

表 3 關鍵能力次數統計

排序	關鍵能力	次數	排序	關鍵能力	次數
1	Communication	632	9	Creative thinking	25
2	Problem solving	561	10	Conflict management	12
3	Collaboration	332	11	Adaptability	11
4	Creativity	286	12	Learning to learn	4
5	Critical thinking	116	13	Risk taking	4
6	Social skills	62	14	ICT literacy	1
7	Flexibility	36	15	Self-direction	0
8	21st century skills	35	--	--	--

4. 結論

本研究應用書目計量分析法，分析過去 1994 年至 2024 年三月所發表之遊戲式學習應用培育 21 世紀關鍵能力之研究。本研究所蒐集之文獻來自 Web of Science 資料庫，該資料庫主要蒐集收錄於 SCI, ESCI, SSCI 索引之研究論文，前述索引廣被學術界所認可，其嚴謹性與代表性應具一定品質。研究共蒐集 1,222 篇相關文章，主要的分析結果如前節所述。

本研究之研究成果提供了遊戲式學習應用於培育關鍵能力之研究一個重要的視角。未來研究可持續應用本研究之研究方法持續拓展文獻蒐集範圍，以增加對於相關研究現況與趨勢之探討與理解。此外，亦可針對特定關鍵能力進行更深入的文獻分析，特別是使用質性的系統化文獻回顧研究法來取得更為深入的見解。

最後，由於本研究所分析之文獻主要為期刊文章，未納入研討會或學位論文共同探討。此外，量化文獻分析雖可以較大規模進行探討，但亦有未能就特定主題再做深入探討之研究方法上之限制，此外，在分析上亦可能因規模而有部份雜訊。在解讀本研究之研究成果時宜納入前述限制為考量。

參考文獻

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959-975.

González-Pérez, L. I., & Ramírez-Montoya, M. S. (2022). Components of Education 4.0 in 21st Century Skills Frameworks: Systematic Review. *Sustainability*, *14*(3), 1493.

Hou, H.-T. (2023). Diverse Development and Future Challenges of Game-Based Learning and Gamified Teaching Research. *Education Sciences*, *13*(4), 337.

Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, *63*, 50–58.

Romero, M., Usart, M., & Ott, M. (2014). Can Serious Games Contribute to Developing and Sustaining 21st Century Skills. *Games and Culture*, *10*(2), 148–177.

Wang, Y.-L., Hou, H.-T., & Tsai, C.-C. (2020). A systematic literature review of the impacts of digital games designed for older adults. *Educational Gerontology*, *46*(1), 1–17.

開發《食物忍者》互動式遊戲繪本以提升國小學生之創造力思維

Developing the Interactive Game-based Picture Book "Food Ninja" to Enhance Creative Thinking in Elementary School Students

藍文君[1]、莫德駿[2]、湯梓辰[3*]
[123*]國立臺灣科技大學應用科技研究所
[3*]jttang0@mail.ntust.edu.tw

【摘要】 本研究旨在探討互動式數位繪本教學對學生創造力的影響。近年來，創造力在全球教育領域受到極大的重視，因此，本研究基於創造力的定義以及繪本的特性，精心編制了符合研究目的且具有教育功能的《食物忍者》。未來，本研究將對互動式數位繪本進行實驗計劃，以提供更完整且清晰的探討。
【關鍵詞】 創造力；互動式學習；數位繪本

Abstract: *The purpose of this study is to explore the impact of interactive digital picture book instruction on students' creativity. In recent years, creativity has been highly valued in education globally. Therefore, based on the definition of creativity and the characteristics of picture books, this study has carefully developed the educational and purposeful "Food Ninja." In the future, experimental plans will be conducted on interactive digital picture books to provide a more comprehensive and clear exploration.*
Keywords: Creativity, Interactive Learning, Digital Picture Books.

1. 前言

學者松居直(1995)指出，豐富的想像力和創造力可以透過直接或間接的體驗獲得，接觸體驗越多，想像力和創作力就越豐富。繪本恰好能夠提供這些難得的經驗。周文敏(2005)也認為，繪本的內容豐富、題材廣泛，故事內容或圖畫表現均可成為孩子無限延伸的教材，以培養孩子創造思考的認知和情意層面的重要關鍵。

近年來，科技深深影響我們的生活習慣和學習模式。在科技的發展和興起下，紙本繪本逐漸轉變為數位繪本(Roskos, Burstein, You, Brueck, & O'Brien, 2011)。數位繪本與其他媒材繪本不同，它能透過繪本引導讓閱讀者與故事互動，在視覺、聽覺、觸覺上提供豐富回饋。透過繪本故事中的提問和延伸互動，有助於訓練閱讀者的專注力、認知力和創造力，同時提升讀者的閱讀與學習動機，使閱讀變得更有趣(賴玉釵，2018)。

綜上所述，在閱讀數位轉型的時代，數位繪本更能夠延伸繪本對於創造力思考教學的優勢，進而引起學習動機並增加學習成效。因此，本研究將進一步設計數位繪本，以及相應的創造力思考教學，為未來數位繪本內容設計和創造力思考教學提供相關實驗的研究工具。

2. 研究設計

本研究期望探討互動式數位繪本對國小六年級學生創造力學習成效的影響。預計使用實驗觀察法，隨機選取的六年級學生分為實驗組和對照組。實驗組將閱讀具有點選功能和特效的互動式數位繪本，而對照組則閱讀僅具有翻頁功能的數位繪本。並透過創造力思考測驗、半結構式訪談進行相互驗證，其中前後測問卷分別採成對樣本 t 檢定比較兩組間的差異以及獨立樣本 t 檢定比較同組內前後測的進步幅度，而半結構訪談則提供質性資料，以全面了解這種繪本對學生創造力的影響。未來實驗架構圖如圖 1 所示。

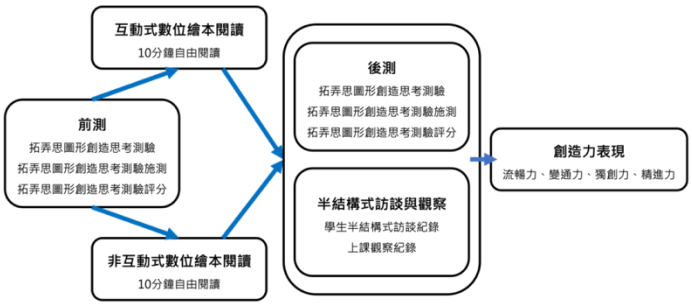


圖 1 未來研究架構圖

3. 研究工具

3.1. 前後測問卷

前後測問卷透過自身教學現場經驗、創造力相關理論以及吳靜吉(1981)所編撰的拓弄思圖形創造思考測驗(甲式)，所編製而成。測驗可分為四個子面向（流暢力、變通力、獨創力及精進力）的相關表現。

3.2. 互動式數位繪本設計

《食物忍者》為自編創意繪本，旨在透過食物忍者變身成各種食物，巧妙融入日常生活場景，以啟發孩子們的創意思考和視覺觀察力。本繪本考量到食物是孩子們日常生活中最為熟悉且有趣的元素之一，透過這樣的背景，預期能夠引起他們對於不同事物的好奇心，同時培養其創意思維。詳見圖 2、圖 3。



圖 2 繪本頁面任務圖

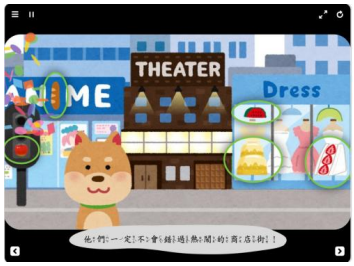


圖 3 繪本互動情況

3.3. 半結構式訪談

筆者透過自身教學現場經驗以及文獻回顧編製半結構式訪談表，用以了解學生的學習狀態，並與圖形創造力思考問卷之結果相互驗證。訪談分為三個面向，分別為過去經驗、活動偏好以及學習自評，試圖讓研究對象自我表述。

4. 研究預期結果

本研究透過前後測問卷以及半結構式訪談相互佐證並了解受測者之行為表現，預期結果有三：數位繪本對於創造力教學成效具有正向影響；數位繪本的互動行為對於創造力教學成效具有正向影響；數位繪本互動設計能提供學生良好的學習體驗。

參考文獻

吳靜吉(1981)。**拓弄思圖形創造思考測驗（甲式）指導及研究手冊**。遠流出版

周文敏(2005)。「創造性圖畫書教學」對國小學童創造力與繪畫表現之研究〔未出版之碩士論文〕。國立中山大學教育研究所。

松居直(1995)。幸福的種子。台英出版。

賴玉釵(2018)。閱聽人跨媒介詮釋圖像敘事之美感傳播芻議：以繪本、iPad 互動電子繪本為例。中華傳播學刊，33，181-215。

Roskos, K., Burstein K., Byeong-Keun You, Brueck J. (2011). A Formative Study of an E-book Instructional Model in Early Literacy. *Creative Education*, Vol. 2, pp. 10-17.

探討後疫情幼兒情緒學習在數位繪本之應用

Explore the Application of Post-Pandemic Children's Emotional Learning in "Digital Picture books

湯梓辰[1*]，趙彥俞[2]
[1*2]國立臺灣科技大學應用科技研究所
[1**] jttang0@mail.ntust.edu.tw

【摘要】後疫情代的影響，幼兒情緒學習面臨到極大的考驗，情緒學習體驗除了少了人與人面對面的互動機會，也缺乏了人因隱藏在口罩下的情緒辨識，因而使用科技輔助在幼兒情緒學習上將會是一個解決方針，而使用數位繪本被認為可以提升情緒方面的認知。本文以開發情緒教育的數位繪本為主，以繪本中的故事情緒引導幼兒學習情緒，未來本研究將以縱向追縱研究定期追蹤訪談家長、老師、以及觀察幼兒使用下的行為表徵，以作為數位學習設計者、幼教學者、幼兒園老師、家長參考。

【關鍵词】 後疫情時代；數位繪本；幼兒情緒學習；縱向追蹤研究

Abstract: *Affected by the post-epidemic era, children's emotional learning faces great challenges. In addition to the lack of face-to-face interaction opportunities, the emotional learning experience also lacks human emotion recognition hidden under masks. Therefore, technology is used to assist children's emotional learning. Learning will be a solution, and digital picture books are believed to improve emotional awareness. This article focuses on developing digital picture books for emotional education and uses the story emotions in the picture books to guide children in learning emotions. In the future, this study will use longitudinal tracking research to regularly follow up interviews with parents and teachers and observe children's behavioral representations when using them as a basis for digital Reference for learning designers, early childhood education scholars, kindergarten teachers, and parents.*

Keywords: post-pandemic era, children's emotional learning, digital-picture books, longitudinal research.

1. 前言

幼兒情緒學習通常透過繪本來呈現。繪本的故事情節提供孩子情境模擬的機會，讓他們透過與故事人物的共鳴，建立情緒經驗（湯梓辰，2019）。在後疫情時期，幼兒的情緒學習面臨了巨大挑戰。由于面對面的互動機會減少，以及口罩遮掩了人們的情緒表達，幼兒的情緒認知能力受到了影響。因此，利用科技輔助幼兒情緒學習被視為一個解決方案。數位繪本被認為可以增強幼兒對情緒方面的認識，從而促進其情緒發展（Ruba & Pollak，2020）。因此，本文以開發數位情緒繪本學習方式，未來進而探討和分析幼兒的情緒學習。

2. 數位繪本在情緒學習的應用

在台灣，幼兒情緒學習通常透過繪本來呈現。繪本的故事情節提供孩子情境模擬的機會，讓他們透過與故事人物的共鳴，建立情緒經驗，進而反思自己的行為。過去傳統的繪本面臨一些挑戰，情緒學習通常需要大人的引導，且家長時間有限或者是家長不知道如何引導孩子(湯梓辰，2019)。

數位繪本的興起為兒童提供了多重感官的體驗。根據林美玲（2014）的研究，數位繪本在教育上有著豐富的價值。相較於傳統繪本，數位繪本結合了文字、插畫和多媒體互動，擴展了兒童閱讀的多元體驗。此外，數位繪本具有互動性、多媒體性以及融合了遊戲元素的特點，營造了一個豐富有趣的學習環境，有助於提升兒童的閱讀動機和興趣。易於理解的介面

設計也有助於兒童突破文字閱讀的限制，進一步深化對閱讀的理解和涵養。此外，數位繪本還具有讓孩子獨立閱讀的特性，這有助於培養他們的自主學習能力和獨立思考能力。

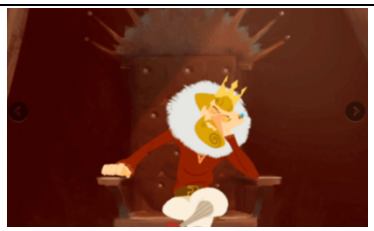
綜以上分析，數位繪本會幫助情緒學習上的應用。因此，本研究將進一步拓展繪本在情緒學習中的重要性，開發數位繪本的相關內容。

3. 開發數位情緒繪本

本研究開發幼兒情緒學習的數位繪本，以醜小鴨及國王的新衣為故事主軸，引導幼兒進行故事欣賞，並透過遊戲機制回答問題及回顧自己的作答來引發幼兒辨識情緒在不同故事角色上的意義。表 1 是本研究所開發的數位情緒繪本綱要。

表 1

幼兒數位情緒繪本綱要

故事關卡分享	醜小鴨	國王的新衣
		

數位繪本之設計概念圖初稿	應用在情緒學習的意義	應用在情緒學習的意義
1.故事欣賞。 2.以小遊戲引導玩家回答問題。 3.回顧自己的問題回答。	✓ 探討每個人都有自己的特色，應該以同理心去探討每個人的特色。 ✓ 幫助孩子瞭解自我的特色，培養孩子對自我的自信心。 ✓ 培養孩子同理每個人的不同差異，也能理解當他人受到排擠可能有的情緒。	✓ 我們無論什麼事都要自己親自去體驗，不能盲目的相信別人，對於自己的懷疑、不舒服的情緒也是要正面的去面對它、了解它。 ✓ 透過每個人說謊讚美國王新衣很漂亮，讓孩子知道為什麼有些人會隱藏情緒面的事實？引導孩子思考人的情緒是可能被隱藏著。

4. 未來研究規劃

情緒學習是需要通過經驗法則來學習認識自己與他人的情緒，是一種較難短時間觀察到的內化的學習歷程 (Lau et al, 2012)。幼兒的成長需要以熟悉幼兒的密切照顧者的角度定期的進行觀察，方可以給予準確的觀察意義 (何懿璉, 2017)。綜此，本計畫將來擬透過六個月的縱向追蹤研究，邀請幼兒密切照顧者協助本案的進行，進而探討孩童對於情緒學習歷程的成長。

參考文獻

林美玲 (2014)。運用電子繪本進行情緒教育的教學成效之研究－以國小二年級學生為例。萬能科技大學 資訊管理研究所 碩士論文。

何懿璉 (2017). 幼兒「日記畫」自發表現題材與形式之縱向追蹤研究。*Semantic Scholar*。DOI:10.6344/THE.NTUE.AD.001.2017.A10

湯梓辰（2019）。*進入孩子的思維世界: 幼兒遊戲探究*。索引數位出版社。https://doi.org/10.978.98698272/18

Ruba AL, Pollak SD (2020) Children' s emotion inferences from masked faces: Implications for social interactions during COVID-19. *PLoS ONE* 15(12): e0243708. https://doi.org/10.1371/journal.pone.0243708

Wang, L., Lee, H., & Ju, D. Y. (2019). Impact of digital content on young children's reading interest and concentration for books. *Behaviour & Information Technology, 38*(1), 1-8.

Lau, P. S. Y. & Wu, F. K. Y. (2012). Emotional Competence as a Positive Youth Development Construct: A Conceptual Review. *The Scientific World Journal.* 2012: 975189. doi:10.1100/2012/975189.

《ANesT 最強蟻巢》AR 教育桌遊設計與開發

Design and Development of Augmented Reality Educational Board Game - ANesT

徐美恩[1]，陳怡瑄[1]，周文己[1]，鄭夢慈[1*]
[1]國立彰化師範大生物學系
[*]mtcheng@cc.ncue.edu.tw

【摘要】教育桌遊可以有效培養與促進學生的概念學習，因此我們以螞蟻與喜蟻動物的互動關係以及螞蟻的特殊行為作為主題內容，設計了一款教育桌遊《ANesT 最強蟻巢》。透過遊戲機制與規則設計，使學習者能夠在遊戲過程中學習相關科學概念。同時，搭配擴增實境技術發展 AR 卡牌，採用動態的方式呈現螞蟻與喜蟻動物的互動模式，讓學習者在遊玩的過程中更清楚看到喜蟻動物與螞蟻的互動行為。

【關鍵字】螞蟻；喜蟻動物；桌遊；擴增實境

Abstract: *Education board games (EBG) can effectively cultivate and promote students' concept learning. Therefore, we developed an EBG, ANesT, whose content focuses on the interactive relationship between ants and myrmecophiles, as well as the unique behaviors of ants. Through playing the game, its game mechanics and rule design engage learners in the gameplay process to understand relevant scientific concepts. Coupled with augmented reality (AR) technology, AR cards are developed to illustrate the dynamic interaction patterns between ants and myrmecophiles, allowing learners to clearly understand the interaction behaviors between ants and their ant friends while playing the game.*

Keywords: ants, myrmecophiles, board game, augmented reality

1. 前言

「教育桌遊」是指將知識概念結合至桌上遊戲中，期望玩家在遊玩的過程當中學習相關的教學概念，目前運用於各領域教學中。而研究也顯示教育桌遊除了對教學有所助益外，對於學生的課堂參與度、學習動機以及學習興趣都有很大的幫助(盧秀琴、林毓哲，2018; 林長信、施如齡，2022)。市面上已有多款教育桌遊在進行販售，內容囊括多項學科主題(鄭秉漢、李文獻、張俊彥，2019)，其中也包含了生物學科。然而，在此學科內容下昆蟲行為仍是桌遊中較少涉及的題材，僅有少數關於昆蟲的特徵辨識與記憶的遊戲。《ANesT 最強蟻巢》將螞蟻與喜蟻動物的互動模式以及螞蟻有趣的行為融入至遊戲機制，由於互動方式通常不易直接觀察，此次桌遊開發結合擴增實境的技術，將螞蟻與喜蟻動物的互動關係以動態的方式呈現於卡牌中，讓玩家可以更清楚的看到互動行為，這樣的設計不僅令遊戲更具互動與趣味性，也協助玩家更深入了解遊戲中的科學概念。本論文呈現《ANesT 最強蟻巢》的遊戲機制與設計。

2. 《ANesT 最強蟻巢》概念內容

《ANesT 最強蟻巢》桌遊開發挑選世界上總數最多的昆蟲「螞蟻」為主角，依據鄭秉漢等人(2019)提出的模型化科學桌遊的設計理念進行開發。雖然螞蟻在日常生活中相當常見，但不同種類的螞蟻會展現出各自獨特的行為，其中許多特殊行為對一般大眾而言是未曾見過的。雖然蚜蟲、介殼蟲與螞蟻之間的互利共生關係較為人所知，但實際上蟻巢中還有許多其

他鮮少見的物種與螞蟻共生。遊戲將螞蟻的獨特行為以及與螞蟻互動的物種設計成卡牌，讓玩家在遊戲中探索螞蟻與喜蟻動物之間的互動關係和行為，同時透過遊戲機制理解這些行為與互動關係所產生的影響。

3. 《ANesT 最強蟻巢》遊戲機制說明

遊戲機制是根據 BoardGameGeek 所歸納的當前桌遊機制而來，採用其中三種，包括行動配額(Action point)、手牌管理(Hand management) 與擲骰子(Dice rolling)。《ANesT 最強蟻巢》遊玩人數可 3-6 人，遊戲時間約 30-40 分鐘，預設之玩家年齡為各年齡層皆可。

3.1. 卡牌設計

遊戲包含三種卡牌：朋友、動作和取食事件(圖 1)。朋友卡牌將喜蟻動物與螞蟻的互動行為設計為卡牌內容，並搭配 AR 圖像辨識，玩家透過手機掃描卡牌圖片，展示該物種與螞蟻互動的動畫，包括介殼蟲、蟻蟋、蚤蠅...等 9 種物種。動作卡牌將不同蟻種的特殊行為設計為卡牌內容，共 7 種。取食事件卡牌則是呈現螞蟻覓食時會遇到的 11 種突發事件。



圖 1 《ANesT 最強蟻巢》卡牌設計

3.2. 規則設定

每位玩家控制一蟻巢，起始配置 4 隻工蟻、4 顆卵並抽取 4 張卡牌做為手牌。遊戲採回合制，當輪到自己的回合時，玩家須先執行蟻巢狀態效果，接下來可選擇兩個動作執行(抽取卡牌、取食、生殖)，最後玩家可以決定是否出牌。若玩家蟻巢中卵數為 0 則遊戲失敗，檯面上的食物消耗殆盡遊戲結束，蟻巢存活的玩家可計算分數，高分者勝。

4. 結論

目前，《ANesT 最強蟻巢》開發設計階段已完成，並於 112 年下學期高中選修課程中試玩，檢測遊戲平衡性、難度與 AR 使用的流暢性，根據教師及學生的反饋進行遊戲規則調整，接下來開始進行小量生產。在實品製作完成後，我們將結合《ANesT 最強蟻巢》教學活動實際應用於以螞蟻為主題的高中選修課程當中，透過成就測驗以及學習投入自評量表，評估教學活動學習成效。同時，也會尋找合適的廠商進行生產執行商品化流程，使有興趣的一般民眾都有機會體驗到這款桌遊。

參考文獻

林長信、施如齡(2022)。史地文化複合式策略桌遊之設計與成效評估。**數位學習科技期刊**，**14**(4)，1-23。

鄭秉漢、李文獻、張俊彥(2019)。模型化科學桌遊。**科學教育月刊**，**419**，20-38。

盧秀琴、林毓哲(2018)。國小教師開發「昆蟲桌遊教具」以增強學童「沈浸經驗，科學過程技能」。**課程與教學**，**21**(1)，105-131。

核心素养视域下皮影戏游戏化教学融入初中语文课堂实施初探

A pilot study of Gamified Teaching of Shadow Puppetry within the Perspective of Core Literacy Integrated into Junior High Chinese Language Classes

郭泽宇[1] 林昕悦[1] 唐伟[2] 江丰光[1*]
[1]上海交通大学教育学院
[2]上海市闵行区教育学院附属友爱实验中学
* fkchiang@sjtu.edu.cn

【摘要】 本研究以初中语文名著导读知识为基础，结合游戏化教学理论，设计以《西游记》“三调芭蕉扇”为主题的皮影游戏；并在中国上海市某初中展开教学实践，分析皮影戏游戏化教学对学生核心素养培养的影响。研究发现通过皮影戏游戏化教学的初探，学生在文化自信、语言运用、思维能力和审美创造上均有了一定程度上的的提升。

【关键词】 游戏化；核心素养；皮影戏；西游记；名著导读

Abstract: *This study is based on the knowledge of Chinese classics in junior high school, combined with the theory of gamification teaching, to design a shadow play game with the theme of "three tones of banana fan" in "Journey to the West"; In addition, the authors carried out teaching practice in a middle school in Shanghai, China, and analyzed the influence of gamification teaching of shadow puppetry on the cultivation of students' core literacy. It is found that students' cultural confidence, language application, thinking ability and aesthetic creation have been improved to a certain extent through the gamification of shadow play teaching.*

Keywords: Gamification, Core Literacy, Shadow Play, Journey to the West, Masterpieces Guided Reading

1. 绪论

在新课程改革的大背景下，传统语文教学方法的不足逐渐显露，特别是"应试教育"导致学生过分依赖机械记忆，从而限制了他们的批判性思维和创造力的发展（王发秀，2022）。研究指出，创新教育模式能够有效培养学生的批判性思维、创新思维和终身学习能力（杜洪娜，2014），这与《义务教育语文课程标准》（2022 年版）对学生核心素养的培养要求相契合，为改进传统教学方法提供了新的方向。教育界开始采用游戏化教学课堂（申煜，2017）等新教学模式。目前，中华优秀传统文化的继承和创新面临了极大的挑战。而游戏化教学为中华优秀传统文化和语文课堂的融合提供了桥梁（Chen et al.,2023）。皮影戏作为中华优秀传统文化的代表集各学科知识为一体，能够很好的与游戏化教学中游戏设计元素性质相匹配。将皮影戏融入初中语文名著导读课堂为培养学生的核心素养提供了有力支持。

2. 课程设计

皮影戏演绎《西游记》“三调芭蕉扇”主题情节，借助皮影戏设计语文课堂游戏化教学。游戏任务为运用小说“三段式”分析法进行情节整合和人物分析，联系皮影戏表演情境适度拓展；游戏积分和排行榜通过任务完成度和学生互评打分体现，打分内容包括：剧本设计（30%）、皮影戏再现故事情节（45%）、场景还原（25%），分数最高者获得胜利。学生通过完成游戏目标的过程，掌握《西游记》中相关文学知识，理解情节内容和人物形象，运用皮影戏与名著结合

的方式，帮助学生更生动形象理解故事情节、分析人物形象，提升语文素养、培养文化自信。图一、图二分别为皮影戏《西游记》彩排与“三调芭蕉扇”课堂展示。



图 1 皮影戏《西游记》彩排



图 2 “三调芭蕉扇”课堂设计展示

3. 课程实践与评价

本研究在中国上海市某中学开展。研究对象为七年级某班 32 人，本研究采用皮影戏的教学方式，共 2 课时 80 分钟，具体内容分别为游戏问答、剧本创作和课堂表演实践。第 1 课时首先由教师做知识引入（5-8min）；其次教师开展闯关游戏挑战（10min），通过《西游记》快问快答和故事接龙两个游戏，考察学生对《西游记》的理解和掌握。为完成游戏分析“三段式”情节分析任务，学生自主进行剧本创作活动，教师指导协助完成剧本设计（25-30min）。第 2 课时各个小组展示皮影戏表演（25-30min），表演结束后学生依据教师提供的评价标准对每组的表演进行打分（10-15min），涉及故事情节的准确传达、角色表现的生动性、创意运用以及团队合作等多个方面计算各组的游戏积分。根据课程设计出《取经故事会》“本场最佳”评分表，在《取经故事会》的评分体系里，确保学生同伴评价与活动目标分数分配一致，如剧本设计占 30%。"本场最佳评分表"应细分情节创编与现场演绎，明确评价标准。教师需阐释评价细则，引导公正评价，确保反映学生真实表现。通过问卷，评价学生对传统文化了解、态度变化和传承意愿，如"我为文化遗产骄傲"，量化文化自信，评估教学效果。

4. 总结

采用皮影戏演绎《西游记》“三调芭蕉扇”片段的教学法与 2022 年版义务教育语文课程标准（教育部，2022）的核心素养要求相吻合，为名著教学带来创新。此方法在文化自信、语言运用、思维能力和审美创造方面提供新思路。《西游记》的丰富文化内涵与皮影戏的非物质文化遗产特性相结合，让学生深入了解中华传统文化，增强文化自信与认同，有助于文化遗产的传承。皮影戏的创作与表演锻炼学生的口语表达和沟通技巧，同时增强团队协作。跨学科知识的运用，如光影原理和文学分析，激发学生的创新思维和逻辑思维。皮影戏的综合性艺术形式，融合文学、美术、音乐和戏剧，培养学生的审美和创造力。

参考文献

中华人民共和国教育部（2022）。义务教育语文课程标准（2022 年版）。北京：北京师范大学出版社。

王发秀（2022）。**中学语文说明文教学的现状及改进策略研究。硕士学位论文。**宁夏大学。

申煜（2017）。中小学游戏化教学研究综述。**上海教育科研（6）**，23-28。

杜洪娜（2014）。加快高校教学管理创新的重要性及途径。**学园(7)**，29-30.

Chen. C，Nurullizam，J.，& Mao，Y. C.（2023）.The study on the effects of gamified interactive e-books on students’ learning achievements and motivation in a Chinese character learning flipped classroom. *Frontiers in Psychology*（14），1-14.

開發以學習人工智慧路徑演算法之 TAMGPA 學習教育模式初探

Development of TAMGPA Learning Education Model for Learning Artificial Intelligence Path Algorithm

朱志明[1*]，李品諺[2]，
[1, 2] 宜蘭大學 資訊工程學系
[*] cmchu817@gmail.com

【**摘要**】本研究旨在以四軸飛行器做為教學工具，設計一種學習人工智慧中的 Dijkstra 和最佳路徑演算法的學習模式，教授學生四軸飛行器的程式編程與操作。本系統由三個部分組成：1.Game: 學習 Dijkstra 及最佳路徑演算法之教育桌遊; 2. Practice: 以四軸飛行器進行程式編程及數據蒐集；以及 3.Analysis: 將所蒐集到的數據進行分析，並以視覺化來呈現結果。因此，本研究將這三個部分簡稱為 TAMGPA（Teaching Aids and Models of Game, Practice and Analysis）。

【**關鍵字**】人工智慧；教育桌遊；四軸飛行器

Abstract: *This study aims to use a quadcopter as a teaching tool to design a learning mode to learn Dijkstra and Optimal Path Algorithm in Artificial Intelligence, and to teach students the programming and operation of quadcopters. The system consists of three parts: 1. Game: an educational board game for learning Dijkstra and optimal path algorithms; 2. Practice: programming and data collection with a quadcopter; and 3. Analysis: analysing the collected data and presenting the results visually. Therefore, this study refers to these three components as TAMGPA (Teaching Aids and Models of Game, Practice and Analysis).*

Keywords: artificial intelligence; educational board games; quadcopter

1. 前言

現今社會因為技術的快速發展對於人才的知識需求越來越高，使學生面臨著巨大的挑戰，而人工智慧是一個改變現代社會的新技術（Makridakis, 2017）。人工智慧極大地提高了製造和服務系統的性能，對於生活的各個領域產生巨大的影響，被廣泛用於科學、工程、商業、醫學等各個領域（Pannu, 2015），因此本研究希望從教育上著手，藉由讓學生在孩童期間建立良好基礎使其加深對於人工智慧的認知與學習樂趣。

2. 文獻探討

人工智慧在教育領域的應用正在興起，對於研究人員和從業人員來說都是一個新興領域。教育桌遊是指能夠在桌面上或是某個可以多人面對面進行的平臺遊戲，其內容包含策略、益智、推理與角色扮演，這類遊戲更注重思維方式、語言表達能力與情商的鍛鍊，過程中遊戲機制可以充滿樂趣與挑戰性使玩家沉浸於遊戲之中，也可以因為難度不同提高玩家的動力，最終產生心流經驗（侯惠澤，2017）。本研究有別於傳統的講述法，透過桌遊使參與者能在遊戲情境中愉快的認識校園，提供一個豐富的介面與有趣的互動性，並因為有共同目標以團隊合作，來增進團體凝聚力（Kotarski, Piljek, Pranjić, Grlj, G., & Kasać, 2021）。

3. 研究方法

本研究自行開發之人工智慧路徑演算法之 TAMGPA 學習模型系統主要分成三個部分：路徑演算法之教育桌遊與擴增實境 APP、四軸飛行器情境模擬之資料蒐集、資料分析與視覺化圖形。教育桌遊採用實體的方式進行設計與製作，並應用擴增實境的 APP 結合遊玩，主要遊玩方式是透過骰子進行移動，規劃從起點到各項目標地點，最後抵達終點的路徑，應用 APP 對各項目標地點進行掃描並回答問題，來蒐集無人機的各種零件，當所有零件蒐集完並抵達終點及遊戲結束。

4. 結果與討論

為了探討不同組別之學生在實驗後，學習成效是否有顯著差異存在因此使用成對樣本 t 檢定進行分析，表 1 和表 2 分別為前測與後測之描述性統計。由表 2 可以得知，兩組在前測與後測有明顯差異存在，實驗組 p <.001 達顯著差異；控制組 p <.05 達顯著差異，表示實驗組在後測時皆有顯著優於前測。從表 1、2 可知，經由本次的實驗後，TAMGPA 學習模型系統使學生的學習成效得到顯著的提升，而且學生是有得到收穫的。

表 1. 實驗組前後測之描述性統計量表

	個數	平均數	標準差
實驗組前測	27	38.58	16.37
實驗組後測	27	63.63	16.70

表 4-2 實驗組前後測之成對樣本 *t* 檢定

	相關性檢定		成對樣本 *t* 檢定		
	相關	顯著性	*t*	自由度	顯著性(雙尾)
實驗組前測–後測	.718	.000	-10.486***	26	.000

*p <.05；**p <.01***p <.001

5. 結論與建議

本研究期望學生透過閱讀教科書、參與教育桌遊和回答擴增實境卡牌問題等多元化學習方式，引導學生透過實踐遊戲化應用，深入瞭解人工智慧的基本原理與應用情境，同時藉由四軸飛行器收集視覺畫數據資料，探索人工智慧在實際場景中的應用。未來可進一步使用準實驗法，比較兩組以上學生的學習成效。

參考文獻

侯惠澤（2017）。以認知理論設計素養導向述你教育遊戲：「微翻轉遊戲式學習模式」在教學現場的推廣與研究。教育研究月刊，282，26-42。

Kotarski, D., Piljek, P., Pranjić, M., Grlj, C. G., & Kasać, J. (2021). A modular multirotor unmanned aerial vehicle design approach for development of an engineering education platform. Sensors, 21(8), 2737.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. Futures, 90, 46-60.

Pannu, A. (2015). Artificial intelligence and its application in different areas. Artificial Intelligence, 4(10), 79-84.

針對視障生的實體程式積木感測板之開發
The Develop of Tangible Programming Sensor Board for the Visually Impaired Students

孔崇旭[1]，陳俊憲[1*]，郭承諺[2]，湯志炫[3]
[1]國立台中教育大學資訊工程研究所
[2]國立台灣大學重點科技學院積體電路設計與自動化
[3]國立台北科技大學電子工程學系
[*] BCS108110@gm.ntcu.edu.tw

【摘要】 程式設計課程用以訓練學生的邏輯思考方式，在許多教育現場的實踐中已被證實能夠顯著有效的提升數理學習的能力。圖形化方塊介面環境讓學生學習資訊課程，不僅能幫助學習運算思維，更能夠在解題過程對抽象概念產生正成效結果。然而，基於輸入輸出、與回饋機制產生的系統性障礙，使視障生無法操作圖形化方塊介面以學習程式設計，是教學現場面臨的一大挑戰。要符合視障生學習需求，使能與一般學生同步學習資訊領域課程的無障礙輔助教具與系統卻付之闕如。本研究嘗試研發可以與圖形化方塊介面產生互動能力的視障輔具，包含利用磁力感測方法偵測植入在積木裡的磁鐵，以辨識積木代表的指令功能，以及開發實體程式積木感測板，及可與產生互動的虛實整合圖形化程式系統。本研究尚進行中，已完成實體積木系統與磁感測器陣列部分，並有一位視障生就完成部分回饋意見。
【关键词】 運算思維；遊戲式學習；視覺化程式語言；實體程式積木；視障程式學習

Abstract: *The programming course has been used to train students in logical thinking, and its effectiveness in significantly improving mathematical learning abilities has been proven in many educational settings. The visual block-based environment for students learning information courses, not only aids in learning computational thinking but also yields positive results in grasping abstract concepts during problem-solving processes. However, due to the systematic barriers created by the input-output and feedback mechanisms, visually impaired students face a major challenge in controlling and communicating with visual block-based software or hardware when learning programming. This obstacle poses a major challenge in educational settings. Currently, there is a lack of accessible aids and systems tailored to meet the learning needs of visually impaired students, enabling them to synchronize with their peers in the information technology curriculum. This study aims to develop an assistive device for the visually impaired, making it possible for them to interact with Visual block-based interface. It utilizes magnetic sensing to detect magnets embedded within the blocks, facilitating recognition of the functions represented by the blocks. The research focuses on the development of a tangible programming block sensor board and an integrated visual programming system capable of interactive virtual and real-world integration. This study is still ongoing, and the Tangible block-based System and Magnetic Sensor Array parts have been completed, and feedback has been obtained from a visually impaired individual.*

Keywords: Computational Thinking, Game based Learning, Visual Programming Language, Tangible Programming Blocks, Visually Impaired Programming Learning

1. 前言

學習或工作的過程，選擇適當的資訊工具，以創造性的思考邏輯進行分析與歸納，找出可行的方法，解決面對的問題與挑戰，是運算思維的體現。運算思維是當前資訊教育的顯學，世界各國競相挹注龐大資源，從小學基礎教育扎根，延伸到高中階段，期望培養出具備競爭力的下一代(Su & Yang, 2023)。拆解問題、型態辨別、抽象概念、演算設計是運算思維的四種核心能力(Henderson, Cortina, & Wing, 2007; Lodi & Martini, 2021)，也是各教育階段要求能夠具備的素養，如何能夠養成運算思維核心能力，程式設計毫無疑問是有效的訓練方法(Melro, Tarling, Fujita, & Kleine Staarman, 2023)。不過程式設計跨入的門檻相對高，尤其是學習各種不同的程式語言，本科生按部就班學習尚且不易精熟，更何況是未經專業訓練的小學、國中、高中（後稱 K-12）階段學生(Kazemitabaar et al., 2023)。

研究顯示學生運用基於圖形化方塊介面的工具進行程式設計，比用文本方式進行程式設計更為容易(Saygıner & Tüzün, 2023)。學者透過量表以具體指標數據評估，發現用圖形方塊人機介面環境，讓中學階段學生學習資訊課程，不僅能幫助學生學習運算思維，更能夠在解題過程對抽象概念產生正成效結果(Çakiroğlu & Çevik, 2022)，圖形化方塊介面 Blockly、Scratch 與 Scratch.Jr，能夠輕易地設計一些多媒體課程，暨包含數學邏輯、寫作學習與故事性，甚至能納入家長作為輔助教學角色，協助教師在非傳統學習環境教導小學生運算思維(Smith & Burrow, 2023; Tawfik, Payne, & Olney, 2024)，因此圖形化介面訓練運算思維是非常好用的工具。

新課綱資訊領域課程融入運算思維於程式設計教育，對特殊教育身心障礙學生的學習，產生相當大的衝擊，其中面向相對弱勢的視障學生族群挑戰尤其巨大。視覺障礙學生運用專用電腦軟體，轉譯文本式檔案為語音及點字訊息，透過聆聽報讀與觸摸點字顯示器產生的點字資訊，來學習課業知識。視障生無法操作圖像式介面，因此圖形化方塊介面就無法適用於他們。但圖形化方塊介面以

其相對較低門檻，已普遍運用於 K-12 階段學生資訊領域課程，以培養其運算思維能力。所以基於視障生學習程式設計需求，發展實體化教具，使可以與圖形化方塊介面產生互動是必須的。

如前所敘述可以得知，近年程式設計教育可用於訓練學生的數理邏輯，進而提升學生的數理能力，因此在新課綱中成為資訊領域課程的教學重點。然而，所有的教學方法都是對於生理功能正常的學生，這些教具對於視障學生而言是不友善的，嚴重影響視障學生的受教權與教學資源的公平性。因此，有必要開發針對視障生這個少數群體的特殊教育之教具，用以解決學校教師對於視障生在程式教育的問題。本研究嘗試研發可以與圖形化方塊介面產生互動能力的視障輔具，即利用磁力感測方法偵測植入在積木裡的磁鐵，以辨識積木代表的指令功能，開發實體程式積木感測板與虛實整合圖形化程式系統。

2. 相關研究

運算思維在 SIGCSE(Henderson et al., 2007)被提出迄今，資訊教育研究者與第一線教育人員，其關注的方向皆圍繞在(1) 如何將運算思維融入與整合到 K-12 課程(Su & Yang, 2023)（2）培養與加強資訊本科學生抽象概念的思維能力(Çakiroğlu & Çevik, 2022) (3)基於非資訊本科學生的學習需求發展適當的課程以使具備運算思維的素養(Alfaro-Ponce, Patiño, & Sanabria-Z, 2023)。程式設計為養成運算思維概念與實作提供了必要的機制，也被認為是發展新技能的理想媒介(Basogain, Olabe, Olabe, & Rico, 2018; Tikva & Tambouris, 2021)，更有研究指出程式設計是數位時代學生必修的素養，且獲得的觀念影響其學習方法、理念與成果(Chou, Tang, & Tsai, 2021)。

史都華與貝克(Stewart & Baek, 2023)回顧 2010 至 2020 年文獻，發現圖形化方塊介面的程式設計軟體被大量運用在 K-12 階段學生，杜歐特蘭(Dúo-Terrón, 2023)分析 2012 至 2020 年發表的文章，結果顯示圖形化方塊介面軟體 Scratch 與教育相關的科學文獻，有漸進式的關聯性。吳等用視覺化程式語言 Scratch 作為運算思維教學與評估工具，結果顯示 Scratch 可以協助教師診斷學生的個別弱點，並可提供適時的介入，學生可以運用運算思維的核心步驟完成任務並解決問題(Wu, Lin, Wang, & Huang, 2023)。蔡等(Tsai, Chen, Hsieh, Chuang, & Lin, 2023)運用圖形化方塊介面於程式設計課程，研究高等教育中非資訊專業大學生學習程式設計的自我效能感與程式設計概念，結果顯示對於提高自我效能感與程式設計概念的表現有所幫助。

烏特拉斯與龐泰利(Utreras & Pontelli, 2023)研究視覺障礙學生面臨的挑戰，歸納出輸入媒體（如鍵盤和滑鼠）、輸出媒體（如顯示器）與回饋機制（如視覺資訊）等障礙，因為大多數現有輔助科技與這些基於運算思維開發出的介面環境不相容，這可能會影響其學習程式設計概念的能力。托帕利等(Topalli & Cagiltay, 2018; Utreras & Pontelli, 2023)研究證明圖形化方塊介面環境，可以改善學生在學習資訊科學入門課程的學習過程。但由於嚴重依賴滑鼠輸入，缺乏螢幕報讀及語音輸出，使得視覺障礙學生無法存取圖形化方塊介面環境(Milne & Ladner, 2019)，因此開發適用於視覺障礙學習者的程式設計環境，使獲得平等的學習機會是有必要的(Shaer & Hornecker, 2010)。

3. 系統設計與展示

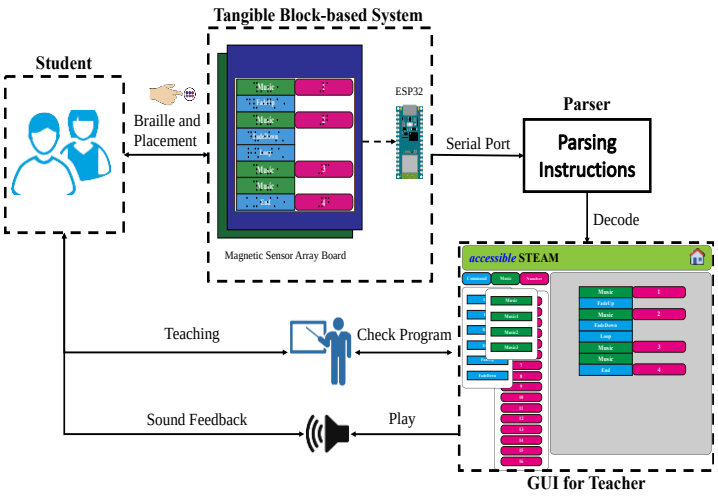


圖 1 系統架構

本研究嘗試設計開發適用於視覺障礙學習者的程式設計環境，研發可以與圖形化方塊介面產生互動能力的視障輔具，即利用磁力感測方法偵測植入在積木裡的磁鐵，以辨識積木代表的指令功能，開發實體程式積木感測板與虛實整合圖形化程式系統。在這個應用於視覺障礙學生之程式學習的系統中，系統架構圖如圖一所示，包含三個主要的組件：實體積木系統（Tangible Block-based System）與磁感測器陣列（Magnetic Sensor Array）是第一個主要組件，可以讓視障學生進行實體程式設計，以移動實體積木（上方以點字說明積木名稱）編排出老師要求的題目，同時由磁感測器陣列檢測所排列的積木為何種類型。解析器（Parser）是第二個主要組件，與磁感測器陣列的微控

制器進行通訊，解析由編譯器實作的程式積木指令碼，通過讀取磁感測器陣列的排列，以 8 位元格式解析每一行程式碼，並分析語法的正確性。視覺化系統 (GUI for Teacher) 為第三主要組件，提供老師觀察學生所排列的程式積木是否正確，同時也能夠將解析器解碼後的程式碼進行執行的動作。在本研究中由於是面向視障生學習程式撰寫的輔助學習，因此所執行的程式碼會以語音和音樂的形式進行播放，讓視障學生能夠了解自己所設計的程式運行狀況。

3.1 主要元件之功能：

3.1.1 實體積木系統 (Tangible Block-base System)

包含 Magnetic Sensor Array 和 Block-Based Pad 與微控制器組成。Block-Based Pad 可以讓視障學生將程式積木放入進行編排，而 Magnetic Sensor Array 則是放置於 Block-Based Pad 之下，用來感測上方所編排的程式積木類型。微控制器則是使用 Espressif Systems Inc.公司的 ESP32-Pico，用於掃描 Magnetic Sensor Array 的狀態，並將結果通過 serial port 傳輸到 Parser 進行解析。

3.1.2. 解析器 (Parser)

用於將 Magnetic Sensor Array 的感測結果進行解碼，並分析其語法，並將解析的結果呈現在 GUI 和執行。

3.1.3. 視覺化介面 (GUI for Teacher)

視覺化介面能夠呈現與視障學生所編排的實體程式積木相同的可視化程式積木，老師能夠藉由 GUI 即時觀察視障學生的程式編寫是否正確，並且給予回饋，同時這個 GUI 也能執行程式積木的解碼結果，以語音和音樂的形式進行播放回饋給視障學生，能夠方便視障學生進行程式積木的程式設計驗證。

3.2 系統操作及範例展示

本節來介紹我們所設計的系統的操作及範例展示。實體積木系統的 Block-Based Pad 包含儲存積木的工作區(Workspace)與工具箱(Toolbox) 二部分，如圖 2 及圖 3 所示。學生需要從工具箱取得表面印有點字的指定積木，依課程範例排列出要求的結果，後點按感測板上的開始鈕以執行，並透過電腦報讀聽取執行結果。

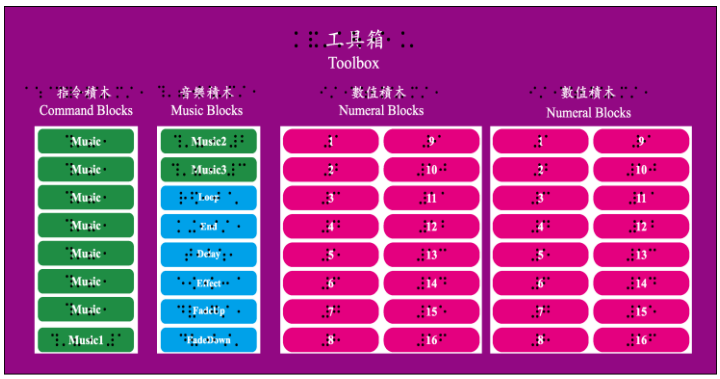


圖 2 工具箱



圖 3 工作區

此外我們目前也已經設計出 3 組教學範例，可讓視障生在依照教學範例練習後，學習程式設計中之循序概念及迴圈的概念，同時也學會聲音的基本處理概念。圖 4 為我們其中一個教學範例。此範例的目標為 (1)依序播放兩段音樂，(2)然後重複 3 次播放另外兩段音樂，(3) 接續依序播放兩段音樂。



圖 4 教學範例與說明

範例題目　目標：

1. 依序播放音樂 1,2、　2. 接續重複 3 次播放音樂 4,5 後結束　3. 接續播放音樂 1,2

步驟說明：

1. 依序取 2 個音樂積木(Music)放置在感測感左側命令積木區域,後放置數值積木(Number)1,2 在其右側數值積木區
2. 繼續拿取迴圈積木 (Loop)放置在左側命令積木區,後放置數值積木(Number) 3

3. 繼續拿取 2 個音樂積木(Music)放置在感測板左側命令積木區,後放置數值積木(Number) 4,5 在其右側數值積木區
4. 繼續拿取結束積木 (End)放置在左側命令積木區
5. 繼續拿取 2 個音樂積木(Music)放置在感測板左側命令積木區,後放置數值積木(Number) 1,2 在其右側數值積木區
6. 按感測板上開始鈕執行

4. 目前成果與未來展望

本研究擬開發針對視障生學習程式設計的實體積木感測板，目前發展出的實體積木系統(Tangible Block-base System)，我們邀請了 1 位高中階段的視障學生進行測試，並有以下意見回饋。認為(1)從工具箱取得（Toolbox）範例題目要求的積木，點字字體較小，需要多次確認才可辨認出要取得的積木。(2) 工作區（Workspace）與磁感測器陣列（Magnetic Sensor Array）組合的測試，對命令積木(Command Block)與數值積木（Numeral Block）的辨識率隨放置位置的調整可獲得較滿意的結果。(3)積木上可增加點字記號，以便用來辨識正反面。(4)命令積木、音樂積木與數值積木雖然都有圓角，但不容易分辨積木種類，建議可以增加凸凹形狀以作為種類與方向區別。後續我們要挑戰解析器（Parser）與視覺化系統 (GUI for Teacher)部分的研發，這包含虛實互動的介面與如何解析對應的積木指令。最後我們將針對學會範例後的視障學生進行成果測驗，與系統的易用性評估也是擬進行的工作。未來亦擬將目前 8cm x 8cm 磁感測陣列範圍擴充成為較大版面 (如 16cm x 24cm)，如此可進行更多指令功能的實體積木辨識，達到擴充整體系統功能目標。

參考文獻

Alfaro-Ponce, B., Patiño, A., & Sanabria-Z, J. (2023). Components of computational thinking in citizen science games and its contribution to reasoning for complexity through digital game-based learning: A framework proposal. *Cogent Education*, *10*(1), 2191751. doi:https://doi.org/10.1080/2331186X.2023.2191751

Basogain, X., Olabe, M. A., Olabe, J. C., & Rico, M. J. (2018). Computational thinking in pre-university blended learning classrooms. *Computers in Human Behavior*, *80*, 412-419. doi:https://doi.org/10.1016/j.chb.2017.04.058

Çakiroğlu, Ü., & Çevik, İ. (2022). A framework for measuring abstraction as a sub-skill of computational thinking in block-based programming environments. *Education and information technologies*, *27*(7), 9455-9484. doi:https://doi.org/10.1007/s10639-022-11019-2

Chou, T.-L., Tang, K.-Y., & Tsai, C.-C. (2021). A Phenomenographic Analysis of College Students' conceptions of and approaches to Programming Learning: Insights from a comparison of Computer Science and Non-computer Science Contexts. *Journal of Educational Computing Research*, *59*(7), 1370-1400. doi:https://doi.org/10.1177/0735633121995595

Dúo-Terrón, P. (2023). Analysis of scratch software in scientific production for 20 years: programming in education to develop computational thinking and steam disciplines. *Education Sciences*, *13*(4), 404. Retrieved from https://doi.org/10.3390/educsci13040404

Henderson, P. B., Cortina, T. J., & Wing, J. M. (2007). *Computational thinking.* Paper presented at the Proceedings of the 38th SIGCSE technical symposium on Computer science education.

Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). *Studying the effect of ai code generators on supporting novice learners in introductory programming*. Paper presented at the Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.

Lodi, M., & Martini, S. (2021). Computational thinking, between Papert and Wing. *Science & education*, *30*(4), 883-908. doi:https://doi.org/10.1007/s11191-021-00202-5

Melro, A., Tarling, G., Fujita, T., & Kleine Staarman, J. (2023). What else can be learned when coding? A configurative literature review of learning opportunities through computational thinking. *Journal of Educational Computing Research*, *61*(4), 901-924. doi:10.1177/07356331221133822

Milne, L. R., & Ladner, R. E. (2019). *Position: Accessible block-based programming: Why and how.* Paper presented at the 2019 IEEE Blocks and Beyond Workshop (B&B).

Saygıner, Ş., & Tüzün, H. (2023). The effects of block-based visual and text-based programming training on students' achievement, logical thinking skills, and motivation. *Journal of Computer Assisted Learning*, *39*(2), 644-658. doi:10.1111/jcal.12771

Shaer, O., & Hornecker, E. (2010). Tangible user interfaces: past, present, and future directions. *Foundations and Trends® in Human–Computer Interaction*, *3*(1–2), 4-137. doi:http://dx.doi.org/10.1561/1100000026

Smith, S., & Burrow, L. E. (2023). Programming multimedia stories in Scratch to integrate computational thinking and writing with elementary students. *Journal of Mathematics Education*, *9*(2), 119-131.

Stewart, W., & Baek, K. (2023). Analyzing computational thinking studies in Scratch programming: A review of elementary education literature. *International Journal of Computer Science Education in Schools*, *6*(1), 35-58. doi:10.21585/ijcses.v6i1.156

Su, J., & Yang, W. (2023). A systematic review of integrating computational thinking in early childhood education. *Computers and Education Open*, *4*, 100122. doi:10.1016/j.caeo.2023.100122

Tawfik, A. A., Payne, L., & Olney, A. M. (2024). Scaffolding Computational Thinking Through Block Coding: A Learner Experience Design Study. *Technology, Knowledge and Learning*, *29*(1), 21-43. doi:https://doi.org/10.1007/s10758-022-09636-4

Tikva, C., & Tambouris, E. (2021). Mapping computational thinking through programming in K-12 education: A conceptual model based on a systematic literature Review. *Computers & Education*, *162*, 104083.

Topalli, D., & Cagiltay, N. E. (2018). Improving programming skills in engineering education through problem-based game projects with Scratch. *Computers & Education*, *120*, 64-74. doi:https://doi.org/10.1016/j.compedu.2018.01.011

Tsai, C.-Y., Chen, Y.-A., Hsieh, F.-P., Chuang, M.-H., & Lin, C.-L. (2023). Effects of a Programming Course Using the GAME Model on Undergraduates' Self-Efficacy and Basic Programming Concepts. *Journal of Educational Computing Research*, 07356331231206071. doi:https://doi.org/10.1177/07356331231206071

Utreras, E., & Pontelli, E. (2023). Introductory programming and young learners with visual disabilities: a review. *Universal access in the information society*, *22*(1), 169-184. doi:https://doi.org/10.1007/s10209-021-00830-2

Wu, T.-T., Lin, C.-J., Wang, S.-C., & Huang, Y.-M. (2023). Tracking visual programming language-based learning progress for computational thinking education. *Sustainability*, *15*(3), 1983. doi:https://doi.org/10.3390/su15031983

第二外語對話練習平台系統開發

Second Language Dialogue Practice Platform System Development

洪霈羽*，施彥安，張立杰

中央大學 學習與教學研究所

* *u0843014@gmail.com

【摘要】本研究以生成式 AI 技術為基礎開發第二外語對話練習平台 Pocketor，為了提升學生學習第二外語時的口說機會，同時結合語音辨識、合成及學生模型之技術，學生及教師均可透過 AI 智慧生成且編輯教材。旨在減輕教師備課及教學壓力，並為學生提供適性化對話教材。

【关键词】生成式 AI；第二語言學習；聊天機器人

Abstract: *In this study, a second language speaking learning platform called Pocketor was developed by integrating the latest AI technologies, including generative AI, text-to-speech, speech-to-text, and student modeling. Pocketor enables both students and teachers to generate their own conversation materials with the support of generative AI. Moreover, it allows students to practice language speaking and engage in conversations with a chatbot. Ultimately, Pocketor aims to offer adaptive conversation materials, provide opportunities for students to enhance their second language speaking skills, and reduce teachers' preparation efforts for instruction.*

Keywords: generative AI, second language speaking learning, chatbot

1. 前言

會話對於語言學習之重要性，需考慮並統合第二語言的學習資源，設計利於學習者的雙語環境（陳浩然，2021），但在教學環境中所使用之教材往往卻需要投入較高的時間及人力成本（Akkakoson，2016）。近年 AI 技術的進步，ChatGPT 等生成式 AI 技術的投入使聊天學習機器人有了更大的發展空間，甚至能生成符合現實情境且多元的對話環境（Jeon 等人，2023）。本文便以此為前提，開發名為 Pocketor 之對話機器人。透過生成式 AI 及學生模型的技術，克服以往開發教材所面臨的高成本問題，並且針對學生的學習狀況及學習需求提供適性化之對話教材。

2. Pocketor 系統介紹

本研究團隊開發之 Pocketor(https://pocketor.cocoing.info/)之系統包含多種語言對話教材（目前版本包含英語、繁體中文、日文、韓文及法文），旨在提升學生練習口說第二外語之機會，及降低教師在備課與教學上的負擔。

學生在進行會話時，從教材獲取至整體的對話練習共會經過四個步驟，如圖一所示：選擇對話、對話及生詞確認、對話練習及結果評估。透過生成式 AI 及學生模型選擇適合學生程度之教材後，在對話及生詞確認階段讓學生預先理解自己將進行的對話內容，並圈選出自

己欲學習之生詞。對話練習則是結合 Google 的語音生成及辨識功能，建立學生與聊天機器人身臨其境的第二外語對話學習。最後藉由 AI 結果評估功能，提供適切於學生之評量評語，以實現個人化學習的良好第二外語學習體驗。



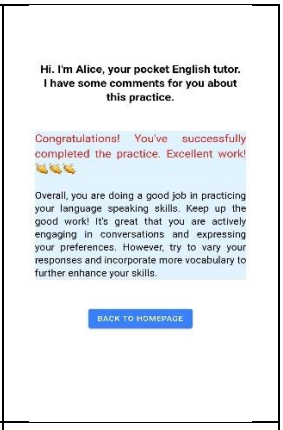
步驟一：選擇對話



步驟二：對話及生詞確認



步驟三：對話練習



步驟四：結果評估

圖 1 對話互動步驟圖

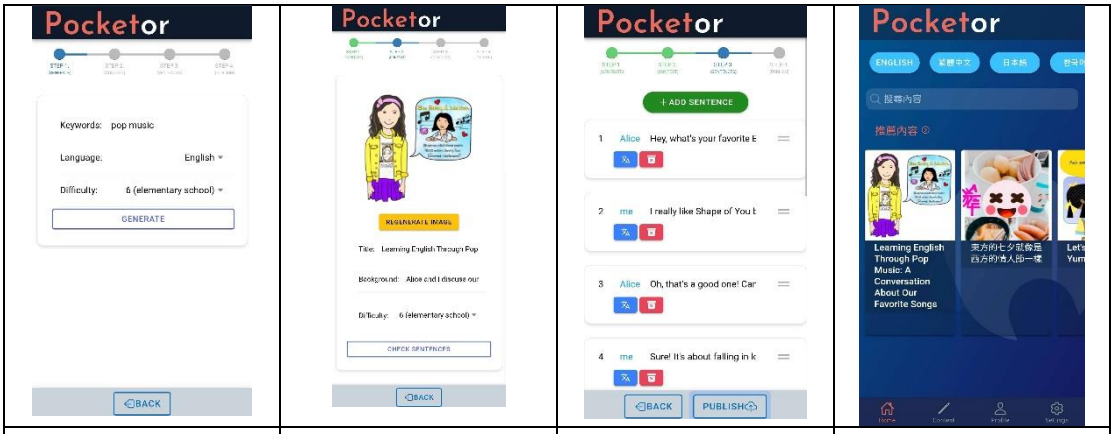
另外，使用者在使用 Pocketor 時應進行帳號註冊，以建立個人化的學生模型。當使用者登入並進行對話練習後，系統會將練習過的對話及學習到的生詞記錄於帳號中，形成個人化的學生模型。未來也預計以此模型發展，從難易度及喜好方面推薦學生合適的教材。

3. 內容編輯範例

教師在進行教材內容編輯時可以選擇使用 AI 智慧生成對話，或是自行新增自編教材。

以 AI 生成教材為例（如圖 2），使用者只需要輸入欲產生教材的題目及關鍵字，並且按照小學、國中及高中三個階段，從 12 個年級難度中選擇欲產生的教材難度，便可以透過 OpenAI 之技術即時生成第二語言對話教材，使用者也可以依照自己喜好進行內容編輯，並進行教材之公開分享。而教師及教材編輯者在進行使用時也能以 AI 生成之教材為基礎，進一步改寫內容。

以圖 2 內容為例，當使用者想要生成主題為流行音樂(pop music)的英語對話，只需要在第一步驟輸入關鍵字流行音樂，在第二步驟就會生成相對應的對話主題及圖片，由於圖片是由 AI 根據關鍵字生成，若出現不滿意的結果也可以重新生成。而步驟三便會根據關鍵字提供 8 至 12 句的對話，使用者可以在檢視對話的同時進行文字編輯以及新增生詞。例如：由於此對話與流行音樂相關，故新增 lyrics, powerful 等與歌曲及歌手特質相關的詞彙。或是將簡單或是艱深的詞彙進行替換，將 powerful 改寫成 mighty，進一步調整對話的難度。

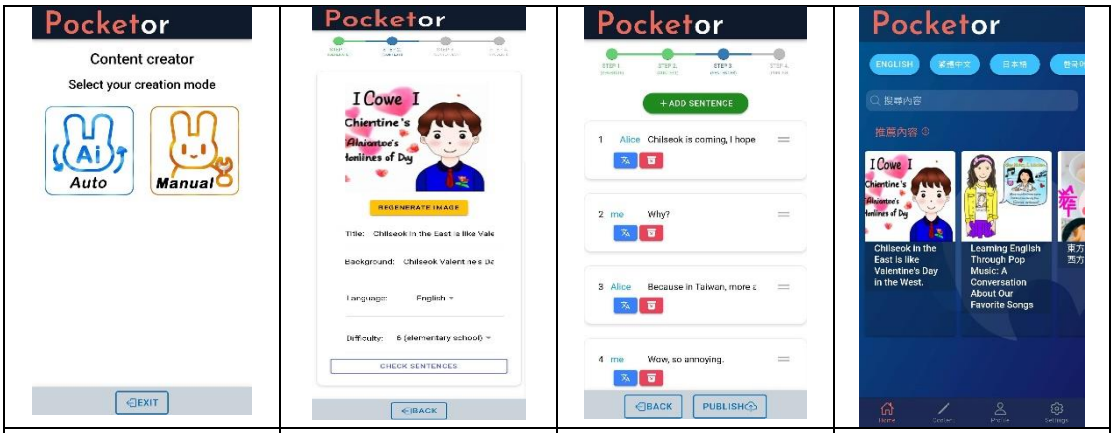


步驟一： 輸入關鍵字	步驟二： 生成圖片	步驟三： 編輯對話內容	步驟四： 教材上傳

圖 2 AI 智慧生成對話教材步驟圖

手動新增教材之功能（如圖 3），是為了希望利用 Pocketor 之既有功能的教師、教材編輯者得以匯入自身所編纂之教材所設計。使用者先是在新增教材的頁面中選擇手動模式(Manual)進入介面，在第二步驟中輸入關鍵字及課程標題後，透過 OpenAI 技術為既有教材生成的圖片，新增生詞及編輯教材後即可上傳至平台。

以圖 3 內容為例，當使用者已經有以情人節(Valentain's Day)的對話教材，他依舊可以從步驟二輸入關鍵字生成適合該教材的圖片，並在步驟三自行輸入使用者事先準備好的對話內容並標註生詞（例如：由於此對話與情人節相關，故新增 romantic, emotion 等形容浪漫行為的詞彙。）



步驟一：選擇 教材新增模式	步驟二： 生成圖片	步驟三： 編輯對話內容	步驟四： 教材上傳

圖 3 教師手動新增對話教材步驟圖

4. 結論

本文以 AI 對話機器人 Pocketor 之系統開發作為介紹之重點，學習者透過四個步驟的對話練習，獲取生成式 AI 所提供的個人化對話教材，進行對話練習並輔以 AI 結果評估。而透過 AI 生成協助教材編輯，使用者按照自己喜好或是以不同的難度、題材進一步設計適合自己的教材，同時也希望就此降低教師的備課及教學負擔。這是聊天機器人結合語言教學所邁

進的一大步，透過新技術的投入，可望提高學習者在語言口說上的能力，並為學習者帶來新的學習體驗。

目前系統已經完成雛形的開發，後續將進行一系列的評估，以了解這一系統對於學生的學習的影響。具體評估方式如下──未來預計將此系統融入外語教學之課堂，蒐集使用情形並依照能力進行分組前後測，以蒐集高低成就學生使用本系統之情形，並統計系統應用於課堂的可行性、實際效果以及其價值。

5. 參考文獻

陳浩然（2021）。有效營造雙語環境，提升中小學學生英語力。台灣經濟論衡。*19*(3)。26-35.

Akkakoson, S. (2016). Speaking anxiety in English conversation classrooms among Thai students. *Malaysian Journal of Learning and Instruction, 13*(1), 63-82.

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 104898.

A Research Proposal for a Dynamic Assessment-based Robot-Assisted Language Teaching Approach on Elementary School Students' English Learning Performance and Learning Experience

Shu-Yun Chien*
Department of Applied Foreign Languages, National Taipei University of Business, Taiwan
* victoriachien.taiwan@gmail.com

Abstract: *This study aims to investigate the impact of dynamic assessment-based robot-assisted language teaching approach on sixth-grade students' English listening and comprehension, learning motivation, and learning experiences. The research will employ an experimental design, recruiting two classes of sixth-grade students, one serving as the experimental group and the other as the control group. The experimental group will receive dynamic assessment-based robot-assisted language teaching approach, while the control group will use conventional robot-assisted language teaching approach. Data on students' English listening and comprehension abilities, learning motivation, and learning experiences will be collected and analyzed. It is anticipated that the results will demonstrate a significant impact of dynamic assessment-based robot-assisted language teaching approach on improving English learning outcomes for sixth-grade students, as well as positively influencing learning motivation and experience. This study contributes to a deeper understanding of the effects of dynamic assessment-based robot-assisted language teaching on elementary school students and provides empirical evidence to support the development and application of future educational technologies.*
Keywords: dynamic assessment, robot, language teaching, elementary school, motivation

1. Introduction

In recent years, there has been a growing global emphasis on English language instruction, driven in part by ongoing trends in internationalization. To achieve more effective teaching outcomes, various emerging teaching methods have gradually gained traction. These include integrating English learning into everyday life through the use of videos and news (Temyanikova et al., 2019), employing flipped classrooms to enhance students' English learning performance (Hsieh et al., 2017), and utilizing technology and robots to supplant traditional teaching methods (Engwall & Lopes, 2022), among others.

Past studies have highlighted the critical importance of listening in English learning, as it is fundamental to tasks ranging from learning letter pronunciation for beginners to comprehending English and engaging in conversation. As English proficiency improves, listening comprehension and speaking skills become increasingly indispensable. Therefore, it is widely acknowledged that listening is paramount in English learning. For instance, previous experiments using BBC programs for teaching have focused on improving subjects' listening abilities, yielding two significant conclusions: firstly, developing strong listening skills aids in speaking proficiency, and secondly, interactive learning is essential for efficient listening instruction.

Taiwan currently faces two primary challenges in English listening instruction. Firstly, there's a prevalence of "exam-oriented teaching," where students primarily focus on listening practice for exams. Classroom instruction often revolves around teachers delivering content followed by students completing exercises and reviewing listening materials. This approach limits students' listening practice to the confines of prescribed teaching materials, hindering their ability to apply their skills in real-life situations. Secondly, traditional teaching methods often fail to effectively stimulate students' motivation to learn. However, research suggests that integrating various technologies can enhance learning effectiveness and motivation (Bahari, 2023).

Recognizing the limitations of traditional teaching methods in English listening instruction, scholars have increasingly explored the use of robots to supplement and enhance teaching practices (Engwall & Lopes, 2022; Lee & Lee, 2022). Many of these endeavors have shown positive impacts on improving learning motivation and experience (Bahari, 2023).

To this end, this study proposes a dynamic assessment-based robot-assisted language teaching approach and investigates its effects on elementary school students' English listening and comprehension abilities, learning motivation, and learning experiences. The research questions are as follows:

1. Are elementary school students learning with the dynamic assessment-based robot-assisted language teaching approach significantly better in English listening and comprehension performance compared to students using conventional robot-assisted language teaching approach?
2. Are elementary school students learning with the dynamic assessment-based robot-assisted language teaching approach significantly better in learning motivation compared to students using conventional robot-assisted language teaching approach?

3. How do the learning experiences of elementary school students differ between those utilizing the dynamic assessment-based robot-assisted language teaching approach and those employing conventional robot-assisted language teaching approach?

2. Literature Review

2.1 Technologies assists learning English

In the process of learning a foreign language, learners often begin with enthusiasm and a sense of accomplishment as they engage with new inputs and handle simple content. However, as the language difficulty increases and novelty diminishes, learners may experience a decline in motivation and engagement, leading to fatigue. In today's technologically advanced era, the use of various technological products has been explored to help learners maintain motivation and efficiently master a foreign language, including English.

Scholars have investigated the effectiveness of various technological products in enhancing language proficiency and maintaining motivation during the language learning process. Examples include using applications to assist English learning (Li & Bonk, 2023), utilizing online platforms and applications for English proficiency enhancement (Tkachenok & Tumskiy, 2019), integrating virtual reality (VR) into language learning (Nicolaidou et al., 2023), and employing English teaching robots to enhance learning effectiveness and motivation (Peng & Porter, 2022), among others.

Research on applications for assisting learning (Wohlwend, 2015) suggests that colorful pictures and images can capture the attention of children, aiding in knowledge absorption and maintaining interest and motivation through various media such as sound, images, videos, and text. Peng and Porter's research (2022) demonstrated significant improvements in learning performance among participants assisted by robots, with positive responses regarding learning attitudes and motivation through questionnaire feedback. Additionally, experiments utilizing robots to assist young children in learning English (Mazzoni & Benvenuti, 2015) showed higher English proficiency test scores and increased engagement, interest, and sense of achievement compared to traditional learning methods.

However, technology-assisted English learning may encounter challenges such as a lack of essential infrastructure and technological resources, which can negatively impact learning (Bahari, 2023). Furthermore, Deng and Yu (2022) notes that while VR initially offers novelty, overuse can diminish its effectiveness in enhancing learning interest. Therefore, in addition to utilizing technological products, incorporating additional teaching strategies such as collaborative learning, group competitions, and self-directed learning can significantly enhance the effectiveness of technology-assisted learning (Li & Bonk, 2023).

2.2 Intelligent Robots

In the realm of education, intelligent robots are defined as machine devices equipped with autonomous learning and interaction capabilities. These robots have the ability to sense students' needs and provide personalized and adaptable educational services tailored to individual learning stages and student characteristics. The integration of intelligent robots in education aims to enhance teaching effectiveness, enabling students to learn English more efficiently through robot-assisted methods, and effectively address the challenges of modern education. As robot technology continues to advance, their application in education is becoming increasingly prevalent (Lane et al., 2015).

Intelligent robots offer numerous applications in education. They can analyze students' individual differences and deliver personalized teaching by providing tailored learning materials and instructional methods, thereby promoting more effective learning approaches. Additionally, intelligent robots can automate assessment processes, evaluating students' progress and understanding in real-time, and offering immediate feedback to assist teachers in refining teaching strategies (Shute et al., 1996). Serving as interactive learning assistants, robots can engage with students, answer queries, and provide explanations, fostering interactive and engaging learning environments (Graesser et al., 2004). Unlike conventional classes, robots can be seamlessly integrated into educational games, offering challenging and entertaining learning experiences through gamified approaches, thus stimulating students' motivation to learn (Hamari et al., 2014).

However, intelligent robots encounter various challenges in the educational landscape, spanning both technological and pedagogical domains. Technical constraints limit robots in understanding human language, adapting to diverse learning styles, and delivering truly personalized instruction (Holstein et al., 2017). Additionally, educators require training to effectively integrate intelligent robot technology into teaching practices, necessitating adaptation to new technologies and acquiring relevant skills to leverage these tools optimally (Clark et al., 2019).

2.3 Dynamic Assessment

Dynamic assessment, a model evaluating learners' performance throughout the learning process, provides insights into both their acquired knowledge and their ability to apply it (Vaezi et al., 2014). Among various models, the dynamic assessment interventionist model stands out. This model entails conducting a pre-test to ascertain participants' current levels, followed by tailored interventions to enhance their learning performance. Finally, a post-test evaluates participants' levels and performances, providing insights into their learning potential.

While dynamic assessment is increasingly integrated into teaching practices, it faces challenges and limitations. Factors such as learner anxiety, motivation, and interaction with evaluators may influence assessment outcomes (Lidz, 1991). Additionally, the design of dynamic assessment tools and procedures requires careful planning, and the reliability and trustworthiness of assessment tools may be compromised by evaluators' subjective judgments (Haywood & Lidz,

2007). Moreover, dynamic assessment may demand more time and resources, posing challenges, especially when implemented on a large scale (Grigorenko & Sternberg, 2001).

Our research aims to address these challenges by combining dynamic assessment with intelligent robots to observe sixth-grade elementary school students' English listening and understanding, learning attitudes and motivation, and learning experiences.

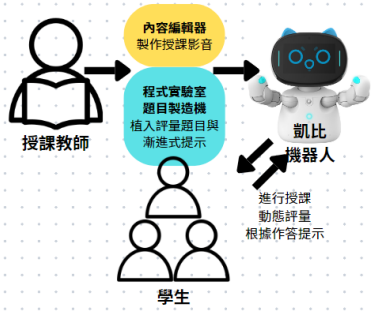
3. Dynamic Assessment-Based Robot-Assisted Language Teaching Approach

The dynamic learning model used in this study belongs to the Interventionist Model, which means that assessment will be based on each student's current abilities and provide progressive prompts accordingly. The process of dynamic assessment robot learning model is as follows: Teaching, Assessment, Prompting, Reassessment.

Firstly, the robot will present course materials through storytelling, conduct teaching, and after teaching, the robot will assess the students' level by asking questions related to the course content. If the students answer correctly, they will proceed to the next question. Otherwise, the robot will provide progressive prompts, gradually transitioning from subtle to explicit, guiding the students step by step to understand and familiarize themselves with the course content. Image 1 illustrates the relationship between the teaching instructor, robot, and students.

Figure 1

Dynamic assessment-based robot-assisted language teaching approach



4. Experiment Design

4.1 Participants

The study participants are sixty sixth-grade students from two classes in an elementary school in northern Taiwan. Each class consists of thirty students. This research adopts a quasi-experimental design, with the two classes divided into an experimental group and a control group. The experimental group utilizes the dynamic assessment-based robot-assisted language teaching approach, while the control group implements the conventional robot-assisted language teaching approach.

4.2 Experimental Procedure

The experimental procedure spans four weeks. In the first week, all participants undergo a pre-test and complete the pre-questionnaire. During the second and third weeks, instructional activities are conducted according to groupings, with the experimental group employing the dynamic assessment robot-assisted language teaching approach, while the control group adopts the conventional robot-assisted language teaching approach. Following the completion of instructional activities, in the fourth week, post-tests, a post-questionnaire, and interviews are conducted. Twelve participants, randomly selected from each group, are interviewed, with each interview lasting 10 minutes, to gain further insights into learners' perceptions of the instructional design.

4.3 Research Tool

The English listening and comprehension test consists of pre-test and post-test papers, which are prepared by an experienced teacher based on the curriculum content and key points. Each test evaluates learners' abilities before and after experimental activities with 25 multiple-choice questions, totaling 100 points. The pre-test ensures both learner groups have equal abilities, while the post-test assesses improvements after learning in different modes.

The learning motivation scale used in this study is adapted from the intrinsic motivation scale developed by Yin et al. (2021). The scale consists of five dimensions: "interest and enjoyment," "tension and pressure," "perceived choice," "perceived ability," and "perceived value." The questionnaire consists of 25 items and uses a 5-point Likert scale, with 1 indicating "strongly disagree" and 5 indicating "strongly agree." The intrinsic motivation scale is administered once before and once after the learning activities in this study to assess whether using different learning modes can engage students in the learning process and enhance their interest and motivation in learning.

The interview questions used in this study are modified based on the interview outline proposed by Hwang et al. (2009). The interview questions consist of seven items. For example, "How is this teaching method different from the way you used to learn English listening?" and "Overall, what do you think are the advantages of this learning method?"

5. Expected Results

This study anticipates that elementary school students utilizing dynamic assessment-based robot-assisted language teaching approach will demonstrate significantly higher levels of English listening and comprehension performance, as well as learning motivation, compared to students solely using conventional robot-assisted language teaching approach. Additionally, students learning with the proposed approach are expected to have positive learning experiences overall.

References

Bahari, A. (2023). Affordances and challenges of technology-assisted language learning for motivation: A systematic review. *Interactive Learning Environments*, *31*(9), 5853-5873.

Chen Hsieh, J. S., Wu, W. C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, *30*(1-2), 1-21.

Clark, D. B., Nelson, B., Chang, H. Y., Martinez-Garza, M., Slack, K., & D'Angelo, C. M. (2019). Scaffolding in the digital wild west: Supporting middle school science teachers in blended learning environments. *Journal of Science Education and Technology, 28*(1), 48-60

Deng, X., & Yu, Z. (2022). A systematic review on the influence of virtual reality on language learning outcomes. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *12*(1), 1-18.

Engwall, O., & Lopes, J. (2022). Interaction and collaboration in robot-assisted language learning for adults. *Computer Assisted Language Learning*, *35*(5-6), 1273-1309.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. Behavior Research Methods, Instruments, & Computers, 36(2), 180-193

Grigorenko, E. L., & Sternberg, R. J. (Eds.). (2001). Dynamic testing: The nature and measurement of learning potential. Cambridge University Press

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?—a literature review of empirical studies on gamification. 2014 47th Hawaii international conference on system sciences, 3025-3034

Haywood, H. C., & Lidz, C. S. (2007). Dynamic assessment in practice: Clinical and educational applications. Cambridge University Press

Chen Hsieh, J. S., Wu, W. C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, *30*(1-2), 1-21.

Holstein, K., McLaren, B. M., & Aleven, V. (2017). Improving robustness and transparency in educational technologies through targeted evaluations. *Journal of Educational Data Mining, 9*(1), 1-21

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

Lane, K. L., Royer, D. J., Messenger, M. L., Common, E. A., Ennis, R. P., & Swogger, E. D. (2015). Empowering teachers with low-intensity strategies to support academic engagement: Implementation and effects of instructional choice for elementary students in inclusive settings. *Education and Treatment of Children*, *38*(4), 473-504.

Lee, H., & Lee, J. H. (2022). The effects of robot-assisted language learning: A meta-analysis. *Educational Research Review*, *35*, 100425.

Li, Z., & Bonk, C. J. (2023). Self-directed language learning with Duolingo in an out-of-class context. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2023.2206874

Lidz, C. S. (1991). Dynamic assessment: An interactional approach to evaluating learning potential. New York: Guilford Press

Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning English using socio-cognitive conflict. *Journal of Educational Technology & Society*, *18*(4), 474-485.

Nicolaidou, I., Pissas, P., & Boglou, D. (2023). Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: A quasi-experiment. *Interactive Learning Environments*, *31*(4), 2001-2015.

Peng, K., & Porter, M. A. (2022). A majority-vote model on multiplex networks with community structure. *arXiv preprint arXiv:2206.13416*.

Shute, V. J., & Psotka, J. (1996). Intelligent tutoring systems: Past, present, and future. In S. P. Lajoie & S. J. Derry (Eds.), Computers as cognitive tools (pp. 289-308). Lawrence Erlbaum Associate

Temyanikova, E., Stolyarova, E., & Fedotova, M. (2019). The value of news programme videos as a tool in the English language teaching. In *EDULEARN19 Proceedings* (pp. 3163-3167). IATED.

Tumskiy, S., & Tkachenok, K. (2019). Teaching English to advanced-level learners at an international university. In *INTED2019 Proceedings* (pp. 3505-3510). IATED.

Vaezi, R., Tabatabaei, S., & Bakhtiarvand, M. (2014). A comparative study of speech-acts in the textbooks by native and non-native speakers: A pragmatic analysis of New Interchange Series vs. locally-made EFL textbooks. *Theory and Practice in Language Studies*, *4*(1), 167.

Wohlwend, K. E. (2015). One screen, many fingers: Young children's collaborative literacy play with digital puppetry apps and touchscreen technologies. Theory Into Practice, 54(2), 154-162.

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

ADDIE 模型指导下翻转课堂与 VR 技术融合的教学设计探究

Exploration of Instructional Design Integrating Flipped Classroom with VR Technology Based on the ADDIE Model

梁楚渝
福建师范大学
leung_choryu@163.com

【摘要】 随着教育信息化进程的深入推进，传统的教育模式正在经历一场深刻的转型。为响应我国教育现代化战略目标，翻转课堂作为一种以学生为中心、强调自主学习与教师引导相结合的新型教学模式，以其鲜明的个性化教学优势，日益赢得教育界的广泛关注与实践推广。本文以 ADDIE 模型为理论框架，从教学过程的三个阶段及教学设计的五个步骤出发，系统地探讨了翻转课堂与 VR 技术相融合的教学设计理论基础和实施路径，旨在为教育工作者提供一种结构化、可操作的翻转课堂教学模式，助力提升教学质量与学习成效。
【关键词】 ADDIE 模型；翻转课堂；教学设计；VR 技术；个性化教学

Abstract: *The deepening of educational informatization is driving a profound shift in traditional teaching models. To align with China's strategic vision for educational modernization, the flipped classroom, characterized by student-centeredness and balanced autonomy with teacher guidance, gains increasing recognition and adoption for its strong potential in personalized learning. Grounded in the ADDIE framework, this paper systematically investigates the theoretical underpinnings and practical pathways of instructional design that integrates flipped classroom pedagogy with VR technology, considering the three phases of the teaching process and the five stages of instructional design. Its aim is to offer educators a structured, actionable template for implementing flipped classrooms, ultimately enhancing teaching effectiveness and student achievement.*

Keywords: ADDIE Model, Flipped Classroom, Instructional Design, VR Technology, Individualized Teaching

1.引言

《中国教育现代化 2035 年》报告指出“加快推进教育现代化、建设教育强国、办好人民满意的教育”为我国未来信息化教育导向。随着中国大学 MOOC、网易云课堂等平台的拓展，翻转课堂作为一种信息化教育新模式，凭借其“以学生为中心”和“个性化教学”两大特点，深受广大学习者青睐。同时，虚拟现实技术（Virtual Reality Technology，简称 VR）是集成构想性、交互性与沉浸性于一体的技术，利用其反复交互特性与翻转课堂结合这一信息化发展新思路展现出强化个性化教学的能力，其中哈佛大学“River City”项目及南洋理工大学虚拟校园（张振刚, 2023）均为 VR 技术在教育领域应用的典范。

此外，ADDIE 模型已被广泛应用并证明在构建在线教学设计方案时的有效性和成熟度，多篇学术研究均证实了这一模型在该领域的坚实基础（Dong, 2021）。翻转课堂这一新型教学模式面临众多挑战，使用 ADDIE 模型指导的教学设计能够弥补自主学习能力要求高、评价体系不适应等缺陷。为了助力“VR+教育”的发展，本研究拟以 ADDIE 模型为理论框架，从教学设计全流程视角系统地剖析 VR 技术与翻转课堂相结合的构建机制与运行逻辑。

2.相关研究及现状

2.1. 翻转课堂相关研究及现状

虽然至今仍无确凿证据证实翻转课堂明显优于传统课堂模式，但已有实验数据显示前者在一定程度能有效增进课堂效率（Buhl-Wiggers et al., 2023）。有研究指出，K-12 教育推广实施需应对的关键挑战在于运营层面的资源配置。而高等教育领域，学者则认为存在“对如何设计与支持探究式学习方法的不足”（Lundin et al., 2018）。

2.2. VR 技术相关研究及现状

随着 VR 虚拟技术多元化展现日益普及，“VR+教育”融合发展被视为重要趋势。尽管有实验发现，VR 环境下陈述性知识获取并无显著优势，但此环境下大脑更易产生与现实场景相似且稳定的神经活动模式（Petukhov et al., 2020），且提升学生创造力作用显著，故可借助 VR 技术构建多样的学习情境，加深教学中知识的理解。加入 VR 技术元素的翻转课堂实践中，具身认知理论得以充分体现（胡艺龄等, 2021），因而本教学设计也鼓励学习者结合学习内容调动多种感官和肢体动作，以促进认知理解和知识内化。

2.3. ADDIE 模型相关研究

ADDIE 模型（如图 1）是由分析、设计、开发、实施和评价五个步骤所构成的系统化教学设计模型。它主要包括学习目标的制定、学习策略的运用以及学习评估的实施三大方面的内容（程豪, 2018），将抽象的教学设计流程可视化，并有效地实现了课程设计的规范化。

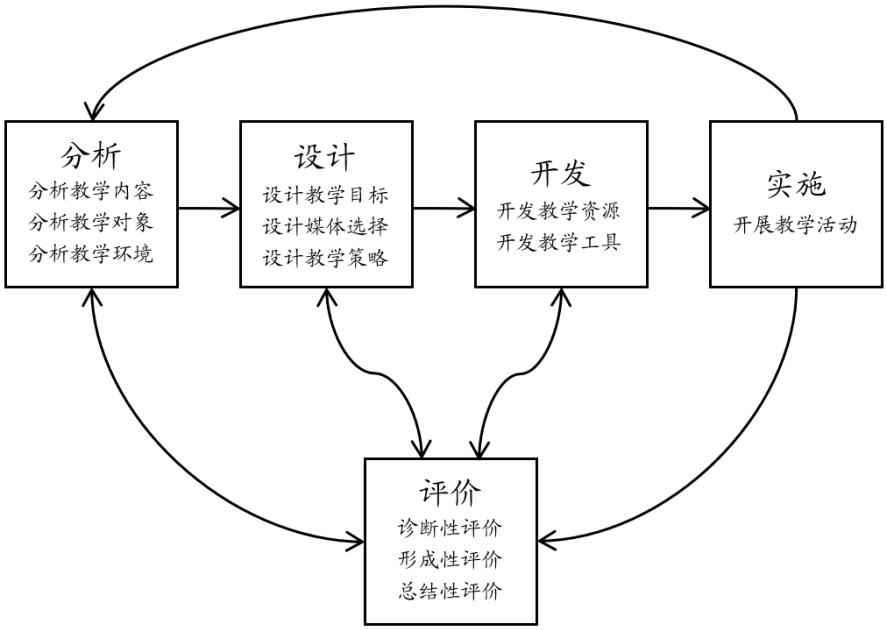


图 1.ADDIE 模型的五个操作步骤及内容

ADDIE 模型作为迭代型教学设计路径，设计者根据过程中的形成性评价返回之前的任一步骤。此特性不仅能适时调整设计思路、优化教学策略，还利于打造高效的教学方案。基于 ADDIE 模型的翻转课堂设计中各个环节皆为精细化、迭代化的设计过程，这确保了翻转课堂能够在有限教学时间内实现高参与度与个性化教学，从而实质提升教学质量和学习效果。

3.教学设计过程

本教学设计框架（如图 2）包括课前准备、课中学习和课后巩固三个阶段。各阶段根据 ADDIE 模型又分为分析、设计、开发、实施和评价五个步骤，其中评价贯穿每一个步骤。出于对原理性设计策略的考虑，在实施步骤完成后将再次回溯分析步骤进行迭代调整，全过程要求设计者统筹兼顾。

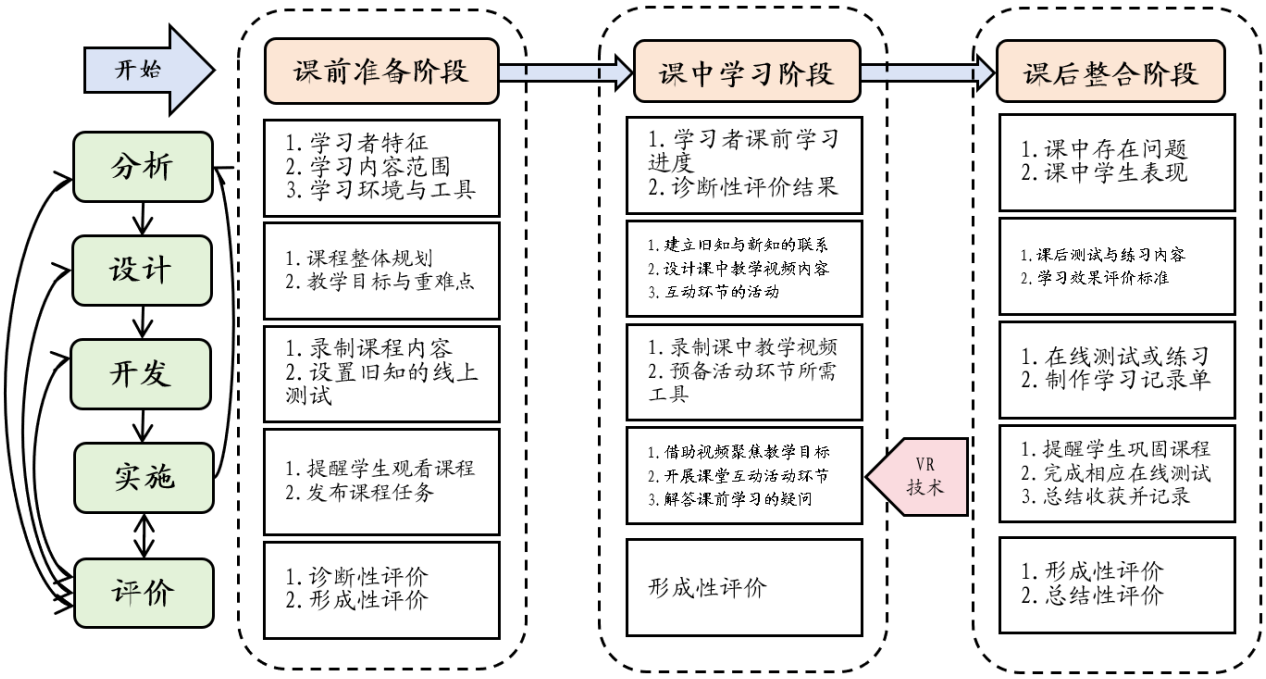


图 2.ADDIE 模型引导下翻转教学与 VR 技术融合的教学设计框架图

另外，本教学设计采用奥苏贝尔的“先行组织者”理论为指导策略。教学策略的选择可依教学需求而变，力求与各类教学情境相适配（Al-Samarraie et al., 2018）。借助建构主义学习理论建立新旧知间的联系，通过强化实现新知与技能掌握的教学目标。新知意义建构中，教学设计者结合学习者高级心理机能特性，突出课程重难点在学习过程中的地位。

3.1. 课前准备阶段

本阶段分析环节除剖析教学内容、资源等内部要素外，还涉及媒体选择、内容铺排及学习者特征等多元因素的考虑（Idsardi et al., 2023）。教学过程各环节均需在设计步骤中确立。实施环节则要求学习者在自适应环境中自主完成以下三个任务：第一，线上回顾旧识；第二，预习新知并建立联系；第三，在平台上积极讨论并提出关键问题。设计者在开发时借 VR 创设符合教学目标的情境，同时甄选有针对性的网络平台，确保 VR 数据与其他数据在此汇集。

3.2. 课中学习阶段

此阶段分析中，教师需要对学生自主学习情况展开评估。随后设计并开发 VR 技术以不同方式融入课堂，尤其是高风险或复杂的教学活动，如体验化学反应、穿越历史或地理探险。除 VR 头盔等硬件外，可辅以摄像头、传感器装置，但使用限 20 分钟内。根据先行组织者策略，教师于课程实施中开展多样化活动（如表 1）以激活旧识并新知接纳搭建桥梁。此外，教师解答课前疑问，普遍性或深度问题特设答疑，确保困惑在课堂得到有效回应。

表 1.活动形式、条件及方式选择参考

课中活动形式	方式	条件	作用
教师讲解	2D 视频展示		唤起旧知，为新知搭建支架
	VR 体验		
师生互动	VR 体验		分析新知
生生活动	小组讨论	简单的学习场景	运用新知
	音、视频资料		
	VR 辅助互动	复杂或存在危险的学习场景	
教师答疑	口头讲述	场景简单且有难度的困惑	理解问题并综合新知
	VR 辅助讲解	场景复杂且有难度的困惑	

3.3. 课后整合阶段

分析环节基于前期学习与 VR 数据，教师调整现阶段难度，设计并开发个性化在线测练。评价时，教师以全局视角审视教学全程并挖掘信息，利用 SPSS 深度分析学生数据及在线问卷收集的关于新难易度、课程满意度等意见，得出整体评估结果后进行教学成效验证。

3.4. 教学设计评价方法

对于整个教学设计的评估，本研究以两个数据源验证教学模式及内容可行性。一是内容效度指数量表（CVI），常应用于量化测评课程材料内容效度，且已被证实有效验证所提教学设计内容；另一是反思期刊意见表，用该表记录师生意见，使迭代教学设计更贴合教学目标与学生需求。此外，除随机抽取个别使用者进行访谈外，还需制作李克特量表针对平台使用的便利性、使用者对 VR 技术与课堂的结合度等问题设置五个等级，根据实际进行量化评估。

4.总结与展望

新时代下的课堂需要更科学和智能化，基于 ADDIE 模型的翻转课堂作为新兴教学模式，加之与 VR 技术的深度融合更能展现学习者主体性地位，推进个性化教学，为教育领域注入新活力。构建有效教学设计有三个关键阶段——构建开发、科学验证及实用价值论证。其中，前者已在此研究中得到理论依据支持，而后两者则为后续实践指引了研究方向。本研究是对翻转课堂与 VR 技术融合的一次浅探，但对于如何将此教学模式效益最大化、提高教学模式信度以及培养具体学科使用者的信息素养等问题，仍亟待未来更深层次的探究与发掘。

参考文献：

张振刚 & 程琳媛.(2023).高等教育数字化转型的动因、实践与启示——以新加坡南洋理工大学为例.*科技管理研究*(16),96-106.

胡艺龄,聂静,张天琦 & 吴忭.(2021).具身认知视域下 VR 技术赋能实验教学的效果探究.*现代远程教育研究*(05),94-102.

程豪.(2018).我国中小学综合实践活动课程开发模式研究——基于 ADDIE 课程教学模型.*当代教育与文化*(02),56-62.doi:10.13749/j.cnki.cn62-1202/g4.2018.02.010.

Al-Samarraie, H., & Saeed, N. (2018). A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, 77–91.

Buhl-Wiggers, J., la Cour, L., Franck, M.S.et al.(2023). Investigating effects of teachers in flipped classroom: a randomized controlled trial study of classroom level heterogeneity. *Int J Educ Technol High Educ* 20, 26 . https://doi.org/10.1186/s41239-023-00396-4.

Dong, H. (2021). Adapting during the pandemic: a case study of using the rapid prototyping instructional system design model to create online instructional content. *Journal of Academic Librarianship*, 47(3), 102356. https://doi.org/10.1016/j.acalib.2021.102356.

Idsardi, R., Friedly, I., Mancinelli, J. et al. (2023). Outcomes of Early Adopters Implementing the Flipped Classroom Approach in Undergraduate STEM Courses. *J Sci Educ Technol* 32, 655–670. https://doi.org/10.1007/s10956-023-10066-9.

Lundin, M., Bergviken Rensfeldt, A., Hillman, T. et al. (2018). Higher education dominance and siloed knowledge a systematic review of flipped classroom research. *Int J Educ Technol High Educ* 15, 20. https://doi.org/10.118.

Petukhov, I. V., Glazyrin, A. E., & Gorokhov, A. V. et al.(2020). Being Present in a Real or Virtual World: AEEG Study[J]. *International Journal of Medical Informatics*, 136:103977.

以體驗式學習的虛擬實境美術館增強學生的藝術鑑賞能力

Enhancing Students' Art Appreciation Abilities through an Experiential Learning Virtual Reality Art Gallery

邱敏棋[1*]，夏綠荷[2]，黃國禎[3]
[1]嶺東科技大學資訊管理系
[2]勤益科技大學體育室
[3]台中教育大學教育資訊與測驗統計研究所
及
台灣科技大學數位學習與教育研究所
[*]minky413@gmail.com

【摘要】 藝術鑑賞活動通常受到時間、距離和經費的侷限，且傳統教學方式主要依賴教科書、投影片或影片傳達資訊。然而，這種方式難以讓學生真正感受到藝術作品的美感及培養批判性思考。為此，本研究提出結合體驗式學習和虛擬實境技術的學習方法(Experiential learning Spherical Video-Based Virtual Reality, EL-SVVR)，旨在引導學生更深入地體驗和欣賞世界各地知名的藝術作品。本研究透過準實驗設計，開發了一套 EL-SVVR 學習系統，並對 45 名選修藝術課程的大學生進行了測試。這些學生分為兩組，其中一組作為實驗組使用 EL-SVVR 方法學習，另一組則作為控制組使用傳統的技術支持學習方法(CTL)。結果表明與 CTL 方法相比，EL-SVVR 學習方法增強了學生的學習成就、學習動機、自我效能、批判性思考與數位繪畫創作表現。

【關鍵字】 藝術課程；藝術鑑賞；體驗式學習理論；虛擬實境

Abstract: *Art appreciation activities are often limited by time, distance, and budget, and traditional teaching methods mainly rely on textbooks, slides, or videos to convey information. However, these methods may not effectively enable students to truly feel the beauty of art and develop critical thinking. To address this, this study proposes a learning method combining experiential learning with Virtual Reality technology (Experiential learning Spherical Video-Based Virtual Reality, EL-SVVR), aiming to guide students to experience and appreciate world-renowned art works more deeply. Through a quasi-experimental design, an EL-SVVR learning system was developed and tested on 45 university students enrolled in an art course. These students were divided into two groups, one using the EL-SVVR method and the other using traditional technology-supported learning methods (CTL). The results show that compared to the CTL method, the EL-SVVR learning method enhanced students' learning achievements, motivation, self-efficacy, critical thinking, and digital painting performance.*

Keywords: art course, artwork appreciation, experiential learning theory, virtual reality

1. 研究背景與動機

藝術鑑賞在教育領域扮演著關鍵角色，它不僅讓學生體會藝術之美，還有助於他們了解創作過程並積累美學知識(Conte et al., 2021)。但是，由於時間、地點和財政的限制，傳統藝術教育模式常常使學生難以親身體驗藝術作品(Hu Au & Lee, 2017)。為解決這一問題，研究者開始利用虛擬實境(VR)技術，為學生提供更加逼真的藝術學習環境。這種方法已經被廣泛應用於多個教育領域，並且被證實能夠顯著提升學習效果(Huang et al., 2020)。VR 技術的應用不僅豐富了藝術教育的內容，也有助於學生在多方面的學習和能力提升(Cheng & Tsai, 2019)。在藝術教育中，通過使用創新的學習工具來輔助學生發展批判性思考、創造性思考和反思能力，可以更有效地展示藝術作品的價值(Jónsdóttir, 2019)。研究表明，為了培養學生的批判性思維，提供真實的觀察和體驗機會是非常重要的(Bustami et al., 2018)。最近的研究發現，利用環景影像虛擬實境系統可以為學生提供真實的學習體驗，進而提升他們的自我效能感、學習成就和學習動機(Huang et al., 2020)。然而，將 VR 技術應用於藝術領域仍然是一個新興且待開發的領域。因此，本研究採用了基於體驗學習的環景影像虛擬實境(EL-SVVR)方法，讓學生體驗美術館中的藝術作品。因此，本研究採用基於體驗學習的環景影像虛擬實境(EL-SVVR)方法，引導學生體驗美術館中的藝術作品。為分析 EL-SVVR 方法對學生學習的影響，本研究提出以下研究問題：

(1)與(CTL)方法相比，EL-SVVR 方法能否增進學習成就？

(2)與 CTL 方法相比，EL-SVVR 方法能否提升學生的學習動機？

(3)與 CTL 方法相比，EL-SVVR 方法能否提升學生的自我效能感？

(4)與 CTL 方法相比，EL-SVVR 方法能否提升學生的批判思考傾向？

(5)與 CTL 方法相比，EL-SVVR 方法能否提升學生的數位繪畫創作表現？

2. 體驗式學習

Kolb 於 1984 年提出的體驗式學習理論，主張學習者能通過親身經歷和深入反思，將日常經驗轉化為知識，進而促進個人學習和成長。這一理論涵蓋四個基本階段：具體經驗、觀察與反思、抽象概念化和主動實踐。這些階段構成了一個廣泛被認可和應用的體驗式學習模式(Morris, 2020; Seaman et al., 2017)。體驗式學習不僅讓學習者在具體的環境中獲得豐富的學習體驗， 同時也鼓勵他們對個人經驗進行深入反思，以提高學習成效(Guthrie & Jones, 2012)。這種學習方式強調實踐和反思的重要性，幫助學習者建立穩固的知識基礎，並通過實際操作來驗證所學的知識(McMullan & Cahoon，1979)。體驗式學習已被證實是一種有效的教學方法，並已在教育領域中得到成功應用(Huang et al., 2016; Kolb, 1984; Murrell & Claxton, 1987)。例如：Cheng 和 Tsai(2019)結合了體驗式學習和 VR 技術，應用於社會學習，有效提高了學習者的學習動機和積極性。Kwon(2019)在沉浸式 VR 環境中實施體驗式科學學習，加強了學生的學習效果和互動性。Alrehaili 和 Al Osman(2019)利用 VR 遊戲讓學生體驗觀察蜜蜂的生態行為，從而增強學習的參與性和動機。本研究採用基於體驗學習的環景影像虛擬實境(EL-SVVR)方法，引導學習者遵循這四個學習階段，目的是提升他們的學習成就和學習動機。

3. 虛擬實境

虛擬實境(VR)技術透過頭戴式顯示器(HMD)建立三維的虛擬環境，為用戶提供了逼真的視覺與聽覺體驗，從而營造出強烈的臨場感和沉浸感(Biocca & Delaney, 1995; Radianti et al., 2020)。這項技術使用者能夠超越時空限制，享受即時的觀察和深入的沉浸體驗(Brooks, 1999)。

研究表明，將 VR 技術融入教育可以增強學生的體驗學習，從而提高教學質量(Wohlgenannt et al., 2020)。例如:Baceviciute 等人(2021)結合了 VR 與傳統學習材料，成功提升了學生的學業成就；Kin 等人(2018)在醫療保健教育中利用 VR，增強了學生的信心和臨床技能；Le 等人（2015）則利用 VR 進行建築安全檢查培訓，提高了學生的安全意識和知識。這些研究表明，VR 能有效提升學生的專業技能、知識建構和自我效能(Bakshi et al., 2020; Calvert & Abadia, 2020)。與傳統 VR 相比，環景影像虛擬實境 (SVVR)具有更低的技術門檻和成本，同時也能帶來良好的學習效果(Hosseini & Swaminathan, 2016; Huang et al., 2020; Wu et al., 2021)。科技輔助的體驗已成為提升藝術教育的有效手段(Bates, 1992; Gong, 2021; Kim & Lee, 2021)。由於實體美術館展覽受到時間和地點的限制(Hartley, 2015; Ness, 2020)，教師常通過教材或影片進行藝術教學，但這些方法往往難以提供身臨其境的體驗(Cavazza et al., 2005; Huang et al., 2020)。因此，本研究提出了結合 SVVR 技術的 EL-SVVR 學習模式，目的是通過互動學習環境提高學生的學習動機、自我效能和藝術欣賞能力。

4. 實驗設計

4.1. 參與者

在本研究中，共有 45 名年齡介於 22 至 24 歲的大學生參與。這些學生選擇了藝術設計課程作為他們的選修科目，且均為無設計經驗的初學者。這些學生分佈在兩個不同的班級中。其中一個班級，包含 23 名學生，被選定為實驗組，並在課程中應用了 EL-SVVR 學習方法。而另一個班級，由 22 名學生組成，作為控制組，使用了傳統的教學方法(CTL)，兩個班級由同一位老師授課。

4.2. 實驗流程

實驗流程如圖 1。首先，所有學生參加由教師主持的 120 分鐘藝術基礎知識課程介紹。這之後，所有學生進行 60 分鐘的前測，這包括學習動機、自我效能和批判性思考能力的評估，並填寫了相關的前問卷。在實驗階段，實驗組的學生使用移動裝置連接到 EL-SVVR 系統，在虛擬的美術館環境中進行了 80 分鐘的體驗式學習，並完成了一系列的互動任務。與此同時，控制組的學生則接受了傳統的教學方法，主要包括簡報和影片學習，並在教師的指導下進行了數位繪畫創作活動。最終，兩組學生均進行了學習成就的後測和後問卷調查。

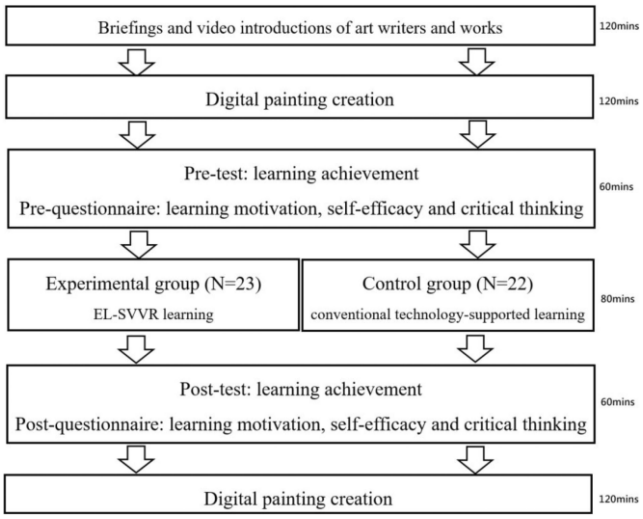


圖 1、實驗過程

4.3. 研究工具

在本研究中，學習成就的測評使用的前測和後測問卷是由兩位擁有超過十年藝術設計教學經驗的老師所設計的。這些問卷專注於評估不同學習領域的成就，包括前測和後測各有 22 個選擇題，總分為 100 分。該測試的目的是評估學生對於未來主義、歐普藝術、極簡主義、硬邊繪畫、構成主義、新造型主義等多種藝術風格的鑑賞能力。前測的 KR20 值為 0.74，後測的 KR20 值為 0.70。學習動機問卷包含六項指標，分為內在動機和外在動機各三項（Pintrich et al., 1991），其 Cronbach 的 alpha 值為 0.86。自我效能問卷則是基於 Pintrich 等人(1991)的設計，共包含八項指標，Cronbach 的 alpha 值為 0.89。批判思考能力的問卷則是參考 Sarıgöz（2012）的問卷設計，共有十個題項，Cronbach's alpha 值為 0.90。對於學生在數位繪畫創作中的表現評分，使用的標準是基於 Lai 和 Hwang(2015)的研究，並由兩位具有十年以上設計教學經驗的老師進行修改。評分的維度包括圖案、空間、色彩和主題。前測的 Kendall's ω 值為 0.804，後測的 Kendall's ω 值為 0.913，顯示評分者間的一致性較高。因此，最終學生在數位繪畫創作方面的成績是根據這兩位老師的評分平均值來評定。

5. 結果

5.1. 學習成就表現

經過迴歸係數同質性檢驗後，結果顯示 F 值沒有達到顯著性差異($F = 0.005$，$p = 0.94 > 0.05$)，未違反同質性之假定，可進行 ANCOVA 分析。如表 1 所示，排除前測成績後，實驗組(調整後平均值=88.67)的學習成就優於控制組(調整後平均值=75.51)，兩組間有顯著差異(F= 7.72，$p = 0.008 < 0.05$) ，效果量較大($\eta^2 = 0.155$)。

表 1、學習成就之共變數分析摘要表

變異數	組別	人數	平均數	標準差	調整後平均數	標準誤	F	η^2
學習成就	實驗組	23	87.17	13.21	86.67	2.80	7.72*	0.155
	控制組	22	75.00	15.50	75.51	2.86		

*p <0.05

5.2. 學習動機

經過迴歸係數同質性檢驗後，結果顯示 F 值沒有達到顯著性差異(內在學習動機：$F = 0.16$；$p = 0.86 > 0.05$)和(外在學習動機：$F = 1.65$；$p = 0.20 > 0.05$)，未違反同質性之假定，可進行 ANCOVA 分析。如表 2 所示，排除前問卷分數後，實驗組的內在學習動機和外在學習動機均顯著高於控制組($F = 5.86$，$p = 0.020 < 0.05$；$F = 4.67$，$p = 0.03 < 0.05$) 。內在學習動機，調整後實驗組的平均值為 4.39，控制組的平均值為 3.80。 就外在學習動機而言，調整後的實驗組平均值為 4.33，控制組為 3.80。此外，效應量 (η^2)內在學習動機為 0.122，外在學習動機為 0.100，表示效應大小為中等。

表 2、學習動機之共變數分析摘要表

變異數	組別	人數	平均數	標準差	調整後平均數	標準誤	F	η^2
內在學習動機	實驗組	23	4.44	0.61	4.39	0.16	5.86*	0.122
	控制組	22	3.74	0.98	3.80	0.17		
外在學習動機	實驗組	23	4.39	0.89	4.33	0.17	4.67*	0.100
	控制組	22	3.73	0.99	3.80	0.17		

*p <0.05

5.3. 自我效能

經過迴歸係數同質性檢驗後，結果顯示 F 值沒有達到顯著性差異(F= 0.005，p = 0.94 > 0.05)。因此，進行 ANCOVA 分析，如表 3 所示，實驗組的自我效能顯著高於控制組(F = 7.89，p = 0.020 < 0.05)。調整後的實驗組平均值為 4.23，控制組為 3.62，效果量較大(η^2 = 0.158)。

表 3、自我效能之共變數分析摘要表

變異數	組別	人數	平均數	標準差	調整後平均數	標準誤	F	η^2
自我效能	實驗組	23	4.25	0.70	4.23	0.15	7.89*	0.158
	控制組	22	3.59	0.83	3.62	0.15		

*p <0.05

5.4. 批判思考傾向

經過迴歸係數同質性檢驗後，結果顯示 F 值沒有達到顯著性差異(F = 1.03，p = 0.316 , p > 0.05) 。因此，進行 ANCOVA 分析， 如表 4 所示，實驗組的批判思考傾向顯著高於控制組(F = 9.79，p = 0.007 < 0.05)。調整後的實驗組平均值為 4.25，控制組為 3.67，效果量較大(η^2 = 0.189)。兩組學習者在批判思考傾向部分存在顯著差異。

表 4、批判思考傾向之共變數分析摘要表

變異數	組別	人數	平均數	標準差	調整後平均數	標準誤	F	η^2
批判思考傾向	實驗組	23	4.28	0.71	4.25	0.13	9.79*	0.189
	控制組	22	3.64	0.82	3.67	0.13		

*p <0.05

5.5. 數位繪畫創作的表現分析

經過迴歸係數同質性檢驗後，結果顯示 F 值沒有達到顯著性差異(圖案, F = 2.735, p = 0.106 > 0.05; 空間, F = 1.333, p = 0.255 > 0.05; 顏色, F = 0.428, p = 0.517 > 0.05; 主題, F = 0.106, p = 0.746 > 0.05)。因此，進行 ANCOVA 分析， 如表 5 所示，實驗組的圖案、空間和主題顯著高於控制組(圖案，F = 4.747，p = 0.035 < 0.05；空間，F = 4.169，p = 0.047 < 0.05；主題，F = 4.722，p = 0.035 < 0.055；主題，F = 4.722，p = 0.035 < 0.055；主題，F = 4.722，p = 0.035 < 0.055；主題，F = 4.722，p = 0.035 < 0.055；主題，F = 4.722，p = 0.035 < 0.055；主題，F = 4.722，p = 0.035 < 0.055)。另一方面，顏色的維度無顯著差異(F = 3.539，p = 0.067 > 0.05)。調整後實驗組平均值在 2.52 以上，控制組在 2.10 以上，效果量較大。結果顯示 EL-SVVR 學習方法有利於學生掌握藝術類別的特徵。

表 5、學習者繪畫作品成績之共變數分析摘要表

變異數	組別	人數	平均數	標準差	調整後平均數	標準誤	F	η^2
圖案	實驗組	23	3.00	0.00	3.00	0.55	4.747*	0.102
	控制組	22	2.82	0.39	2.83	0.57		
空間	實驗組	23	2.52	0.67	2.54	0.17	4.169*	0.090
	控制組	22	2.05	0.84	2.30	0.17		
顏色	實驗組	23	2.52	0.67	2.52	0.16	3.539	0.078
	控制組	22	2.09	0.81	2.10	0.16		
主題	實驗組	23	2.78	0.42	2.74	0.10	4.722*	0.101
	控制組	22	2.36	0.49	2.41	0.10		

$^{*}p<0.05$

6. 結論與建議

本研究提出環景影像虛擬實境(SVVR)系統，並證明利用體驗學習環景影像虛擬實境(EL-SVVR)方法能夠有效提高學生在學習成就、學習動機、批判性思考、自我效能以及數位繪畫創作方面的表現。研究中發現，透過參與 EL-SVVR 活動，學生的藝術品鑑賞能力得到顯著提升，且在學習成就上超越了控制組。SVVR 系統提供的 360 度虛擬環境，有助於學生探索各種學習要素，從而促進他們的學業表現。EL-SVVR 的學習模式將課堂知識與情境學習相結合，使學習者能夠更深入地參與學習過程，並提高了他們的內在和外在學習動機。該方法對提升學生的批判性思考能力也有顯著效果。此外，相比控制組，實驗組學生的自我效能感顯著提高，研究表明當學生在 VR 環境中表現良好時，他們的自我效能感也會隨之增強(Chen & Hsu, 2020)。在數位繪畫創作方面，實驗組在圖案、空間、主題等方面的表現明顯優於控制組，這可能是因為 VR 技術提高了學生的學習意願和專業知識，從而影響了他們的創作。研究的一個限制是 VR 場景主要基於維基百科上排名世界參觀人數最多的美術館。因此，建議未來的研究可以擴展到更多元的場景。此外，未來的研究也可以探討不同學習階段的學生在 SVVR 環境中的表現。

參考文獻

Alrehaili, E. A., & Al Osman, H. (2019). A virtual reality role-playing serious game for experiential learning. *Interactive Learning Environments*, https://doi.org/10.1080/10494820.2019.1703008

Baceviciute, S., Terkildsen, T., & Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. *Computers & Education*, *164*,104122.

Bates, J. (1992). Virtual reality, art, and entertainment. *Presence: Teleoperators & Virtual Environments*, *1*(1), 133-138.

Bakshi, S. K., Lin, S. R., Ting, D. S. W., Chiang, M. F., & Chodosh, J. (2020). The era of artificial intelligence and virtual reality: transforming surgical education in ophthalmology. *British Journal of Ophthalmology*, 1-4. http://dx.doi.org/10.1136/bjophthalmol-2020-316845

Brooks, F. P. (1999). What's real about virtual reality? *IEEE Computer graphics and applications*, *19*(6), 16-27.

Bustami, Y., Syafruddin, D., & Afriani, R. (2018). The implementation of contextual learning to enhance biology students' critical thinking skills. *Jurnal Pendidikan IPA Indonesia*,*7*(4), 451-457.

Calvert, J., & Abadia, R. (2020). Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Computers & Education*, *159*,

104005.

Cavazza, M., Lugrin, J. L., Hartley, S., Le Renard, M., Nandi, A., Jacobson, J., & Crooks, S. (2005). Intelligent virtual environments for virtual reality art. *Computers & Graphics*, *29*(6), 852-861.

Chen, Y. L., & Hsu, C. C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers & Education*, *154*, 103910.

Cheng, K. H., & Tsai, C. C. (2019). A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors. *Computers & Education*, *140*, 103600.

Conte, E., Habowski, A. C., de Almeida Piedade, J., & Milbradt, C. (2021). Art-education and its developments in pedagogical training. *Acta Scientiarum*, *43*, e47923.

Gong, Y. (2021). Application of virtual reality teaching method and artificial intelligence technology in digital media art creation. *Ecological Informatics*, *63*, 101304.

Gong, Y. (2021). Application of virtual reality teaching method and artificial intelligence technology in digital media art creation. *Ecological Informatics*, *63*, 101304.

Guthrie, K. L., & Jones, T. B. (2012). Teaching and learning: Using experiential learning and reflection for leadership education. *New directions for student services*, *2012*(140), 53-63.

Hartley, J. A. (2015). *Museums and the digital public space: Researching digital engagement practice at the Whitworth Art Gallery*. The University of Manchester.

Hosseini, M., & Swaminathan, V. (Eds.). (2016, December). *Adaptive 360 VR video streaming: Divide and conquer*. Paper presented at the International Symposium on Multimedia (ISM). San Jose, CA.

Hu Au, E., & Lee, J. J. (2017). Virtual reality in education: a tool for learning in the experience age. *International Journal of Innovation in Education*, *4*(4), 215-226.

Huang, H. L., Hwang, G. J., & Chang, C. Y. (2020). Learning to be a writer: A spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses. *British Journal of Educational Technology*, *51*(4), 1386-1405.

Huang, T. C., Chen, C. C., & Chou, Y. W. (2016). Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education*, *96*, 72-82.

Jónsdóttir, Á. B. (2019). Critical thinking and community engagement through artistic actions. *International Journal of Art & Design Education*, *38*(3), 700-709.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Kim, Y., & Lee, H. (2021). Falling in Love with Virtual Reality Art: A New Perspective on 3D Immersive Virtual Reality for Future Sustaining Art Consumption. *International Journal of Human–Computer Interaction*. https://doi.org/10.1080/10447318.2021.1944534

King, D., Tee, S., Falconer, L., Angell, C., Holley, D., & Mills, A. (2018). Virtual health education: Scaling practice to transform student learning. *Nurse Education Today, 71*, 7-9.

Kwon, C. (2019). Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virtual Reality, 23*(1), 101-118.

Lai, C. L., & Hwang, G. J. (2015). An interactive peer-assessment criteria development approach to improving students' art design performance using handheld devices. *Computers & Education*, *85*, 149-159.

McMullan, W. E., & Cahoon, A. (1979). Integrating abstract conceptualizing with experiential learning. *Academy of Management Review, 4*(3), 453-458.

Morris, T. H. (2020). Experiential learning–a systematic review and revision of Kolb's model. *Interactive Learning Environments*, *28*(8), 1064-1077.

Murrell, P. H., & Claxton, C. S. (1987). Experiential learning theory as a guide for effective teaching. *Counselor education and supervision, 27*(1), 4-14.

Ness, C. (2020). 'Dressed to the Nines', Birmingham Museum and Art Gallery, 7 December 2019-4 September 2020. *Textile History*, *51*(2), 254-257.

Pintrich, P. R., Smith, D. A. F., Garcia, T., and McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. MI: National Center for Research to Improve Postsecondary Teaching and Learning. (ERIC Document Reproduction Service No. ED 338122)

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Sarıgöz, O. (2012). Assessment of the high school students' critical thinking skills. *Procedia-Social and Behavioral Sciences*, *46*, 5315-5319.

Seaman, J., Brown, M., & Quay, J. (2017). The evolution of experiential learning theory: Tracing lines of research in the JEE. *Journal of Experiential Education, 40*(4), NP1-NP21.

Wohlgenannt, I., Simons, A., & Stieglitz, S. (2020). Virtual reality. *Business Information Systems Engineering, 62*(5), 455-461.

混成式教學於大學體育課對動機信念和合作能力影響之探究

Explore the Impact of Motivation Beliefs and Teamwork Cultivation via Integrates Blended Teaching and Collaborative Learning into University Physical Education Classes

陳偉瑀[1]，陶淑瑗[2*]，張家榮[3]
[1]德明財經科技大學體育室
[2]德明財經科技大資訊科技系
[3]元智大學資訊傳播學系
[*]tsy@takming.edu.tw

【摘要】本研究以大學體育課且導入混成式教學和合作學習策略於學生的學習。為檢驗該教學實踐場域對於學習者的影響，以一年級羽球單元學生為研究對象，透過學生在動機信念、學習策略、和合作能力之回饋。研究結果證實跳脫傳統體育課程設計，有效提升學習者的動機信念與學習策略目標。透過組員間的互動，和期末團體競賽的共同目標，讓組員間學習如何進行管理與協商。換言之，不僅讓學習者於體育課程精進運動技能，且學習者在此課程建構的氛圍中，強化其動機信念和堅定了學習策略。該研究有助於教授體育課程之教師，跳脫傳統式教學的另一種課程設計。

【关键词】混成式教學；動機信念；學習策略；合作能力；體育課

Abstract: *This study takes university physical education classes and integrates blended teaching and collaborative learning into students' learning. In order to evaluate the effects of teaching practice field on learners, first-year badminton unit students were used as the research participates, and feedback was obtained through participates' motivational beliefs, learning strategies, and teamwork cultivation. The results reveal that those pedagogies could effectively improve learners' motivational beliefs and learning strategy goals. Through the interaction between team members and the common goal of the final team competition, team members learn how to manage and negotiate. In other words, not only do learners improve their motor skills in physical education courses, but also they strengthen the motivational beliefs and learning strategies under the atmosphere constructed by this course. This research wouldl help teachers who teach physical education courses to design another course that breaks away from traditional teaching.*

Keywords: blended teaching, motivation beliefs, learning strategies, teamwork cultivation, physical education

1. 前言

在體育課教學中，傳統教學往往以教師為中心，透過口述和親自示範、和提醒動作要領進行授課；學生則模仿教師的動作，且持續不斷練習以學習運動技能。也因此，教學現場的授課教師須關注每位學生在學習過程中產生的千奇百怪的身體協調，並且給予持續回饋與調整；但是藉由這樣不確定的動態教學過程，即使學生大腦懂得怎麼運用肢體模仿動作卻不代表該生已經學會相關技能。隨著科技輔助教學之普及，礙於體育的刻板印象往往是基於大肌肉為主的身體活動，體育課還是以傳統授課方式居多；而另一個讓體育教師怯於導入科技到教學現場的主因，則是對於科技產品結合授課規劃的陌生。在 2019 年末新冠肺炎發生，2020 年疫情嚴重性持續衝擊全球，不僅影響民生、百業、和經濟，學校之授課也影響巨大。從幼兒園到高教體系，不得不停止面對面教學，轉而為線上授課，以避免人群各種面對面接觸進

而導致非必要的傳播病毒。對於體育課程而言，也受到這波影響，體育教師勢必結合資訊科技以進行線上教學，才能順應該波疫情而讓體育課的衝擊降低到最小。

資訊科技輔助教學，實已行之有年且有其成效。Alavi 與 Leidner(2001)的研究強調，整合資訊科技於教學現場，學習者能在資訊科技建構的學習環境和知識內容媒介，進行知識或經驗的傳遞與交換，而師生間也能進行同步或非同步的溝通與學習。對於體育課程而言，以多媒體媒材形式呈現體育技能，是資訊科技輔助體育課教學很重要的一環。例如以多媒體資訊科技輔助教學呈現於遠距課程的設計，或是實體課程中扮演技能學習的重要推手以精進學生技能學習(林怡君，2018；劉與正，2017；李昆霖，2016；陳怡良、李昆霖，2017）。藉由多媒體教材教學不僅讓傳授知識的範圍變寬廣，且可協助學生對於不易掌控和理解的抽象知識或動作演練，以視覺形式展現，教師不再限於單純的口述講課，而能充分藉助資訊科技媒介進行教學，引領學生於視覺化、生動與具象、且不受時空限制的學習世界。此種技能類的學習，符合 Fitts 與 Posner(1967)提出的「動作學習模型」，強調運動技能可透過三階段來學會動作技巧；分別是認知（Cognitive）、聯結（Associative）與自動（Autonomous）階段，對應時間軸則是初學、練習與熟悉的三層(Chan，2020)。也因為如此，多媒體融入體育課程，在妥適導入教學策略，不僅能提升其學習成效，且有助於深化學習者的學習動機(黃相瑋、沈淑鳳、陳儷今與蘇榮立，2020)。而夏綠荷與林彥男(2018)研究，則是基於數位學習與體育教學之文獻回顧，歸納出發展數位化體育教學之建議如下：一、參考運動學習三個階段的特性與歷程，將有利於學習成效的提升；二、結合科技與不同的學習策略，能更貼近技能學習理論，拓廣輔助學習的範圍；三、無須過度害怕技術門檻，只要結合良好的教學策略，依然可以創造出理想的數位學習環境。而 Hung、Young 與 Lin（2018）的研究，則是以 iPad 即時錄影和即時回放之功能，發展在羽球技能學習的教學模式；研究結果顯示，該教學模式輔助教學的實驗組，其羽球技能與學習動機皆顯著優於傳統教學法之控制組。

體育教師若能認知到運動技能的學習對於初學者而言，從無到有到精熟乃須經過反複練習，如何有更積極的教學策略讓學生克服學習過程伴隨挫折，且讓學習者有很堅定的動機信念，將有助於學生提升其運動技能。基於前述脈絡，本研究將建構系列化羽球基本動作多媒體數位教材，結合傳統口述與動作示範教學法，且導入線上數位教材，於教學過程輔助課堂教學，以順應時代變遷與潮流，期能改變教學現場的學習氛圍具彈性與趣味化，讓學習者有強烈的學習動機，進而習得運動技能。

此外，根據聯合國教科文組織(UNESCO)、歐盟(EU)或國際經濟合作發展組織(OECD)，把團隊合作能力的培養，視為高等教育改革的重點(吳明烈，2010)。大學生具備團隊合作能力或領導能力，乃為各大學畢業核心指標能力之一。尤有甚之，在職場中團隊合作能力是排序第 2 高的需求能力(行政院青輔會調查，2011)；團隊合作能力強調工作中所需的協調溝通、互助完成組織共同任務的能力。也因此，若能在學生學習階段即導入合作學習，有助於培育學生該方面的素養。合作學習的教學模式基本上有四個程序：教師引導活動、小組互動討論學習活動、小組報告與師生討論、評量與表揚等；教師於單元學習時導入合作學習，能強化學生學習的認知過程並提高學習成效(Johnson & Johnson, 1992)。設計良好的合作學習讓成員能以團體目標為優先，藉由成員間特殊的技能，不計較利益相互支援貢獻所長，也藉由溝通批評討論，進而完成團隊任務(Parker,1985; Sallis, 2014)。

基於前述的相關研究，本研究認為在大學體育課程中，由教師錄製羽球技能多媒體數位教材且放置於線上學習平台，且規劃合作學習策略於課堂學習活動，將合作學習、多媒體教材與羽球課程學習目標進行整合，應該值得進一步探討對於學生的動機信念、學習策略、和合作能力遷移的改變。為達到本研究之目的且檢驗其可行性，本研究將導入前述課程設計於大一學生體育課羽球項目之學習，且擬定研究問題如下：

- 學習者在動機信念的遷移表現為何？
- 學習者在學習策略的遷移表現為何？
- 學習者在合作能力的遷移表現為何？

2. 研究方法

本研究以支援大學一年級體育課程羽球項目，授課教師進行混成式教學，且在學習過程讓學生分組，以進行合作學習。以下分別就研究對象和活動設計、研究工具、資料收集和分析進行詳細說明。

2.1. 研究對象與活動設計

研究對象為台灣北部某科技大學一年級修習體育課羽球項目，共 102 位學生修讀，但是有 6 位學生於中途退選，沒有完成該課程。配合該課程之羽球項目共進行 8 週之教學實驗，前三週尚在加退選作業，故主要以讓學生清楚該課程各學習單元和對應之週次活動。在確定修課名單後，將學生進行分組，以 8-10 人為一組，該學生數之規劃乃基於羽球項目最後週進行的分組團體循環賽。第八週進行分組團體循環賽，讓學生體驗以選手角色為團體爭取較優成績的經驗；該團體賽每場比賽共有 5 點，分別為男子雙打與女子雙打各 2 點、和混雙 1 點。

該課程之教學場域採混成式：一為數位學習教學平台（後續簡稱 e 學院平台），係指教師在課前即根據授課單元錄製對應之教材影片，且上傳至該校的 e 學院平台（如圖 1 所示）；二為面對面的教學現場，課堂上教師進行動作示範、口語解說、和立即回饋。採用該混成式教學，將有利於學生在練習時，隨時可透過手機去 e 學院平台，反覆觀看教材影片（如圖 2 所示），藉由該影片呈現教師的正確動作，即時修正與調整自身的錯誤姿勢，讓基本技能更臻正確。



圖 1 e 學院平台的教材影片

研究對象於前三週將填寫三份問卷，而在活動結束後的第八週再次填寫此三份問卷。藉由學前和學習後問卷的填寫，以檢視本研究所設計的教學活動對於學習者在動機信念、學習策略、和合作能力的覺知反應。



圖 2 學生以手機到 e 學院平台，觀看教材影片，以修正姿勢

2.2. 研究工具

本研究所使用的工具，分別為：體育課動機信念量表、體育課學習策略量表、和合作能力量表，其運用方法說明如下。

- 動機信念量表
  本研究工具採用陳美榕與陳偉瑀（2007）所編製的「體育課動機信念量表」（共 20 題），做為瞭解大專生在體育課學習過程中，用來維持學習活動與達成目標動機與信念的方式。分為四個向度：「內在價值」、「考試焦慮」、「自我期許」與「學習目標」。採 Likert 5 點量表計分，其解釋總變異量達 65.73%。在信度考驗方面，四個向度分別介於 0.77 - 0.89 之間，重測信度介於 0.73-0.86 之間。
- 學習策略量表
  本研究工具採用陳偉瑀與陳美榕（2007）所編製的「體育課學習策略量表」（共 19 題），做為瞭解學生在學習過程中，為促進學習結果，達成學習目標，所使用的各種有順序性的學習方法、技術、技巧或內在的心理歷程。分為四個向度「考試準備」、「情緒支持」、「資源管理」和「動機控制」。採 Likert 五點量尺計分，其解釋總變異量達 63.20%。在信度方面，四個向度分別介於 0.79 - 0.82 之間，重測信度介於 0.72-0.85 之間。
- 團隊合作能力量表
  本研究工具採用楊文輝與吳致維（2010）所編製的「合作能力量表」（共 21 題），作為瞭解學生在體育課的學習過程中，是否能有效提升團隊合作能力。分為四個向度「認同感與榮譽感」、「管理與協商」、「表達與溝通」和「反思與評鑑」。採 Likert 五點量尺計分。在信度方面，總體量表的 Cronbach' s α 係數為 .981，四個向度分別介於 0.87 - 0.96 之間。

2.3. 資料收集與分析

本研究資料收集來自於動機信念、學習策略、和合作能力等三項問卷的填寫。研究對象於課程進行的前三週前間填寫一次，且於第八週再填寫一次；而前三週填寫的問卷，將做為期初學生在此三項的覺知反映，稱之為期初覺知，單元末填寫的則稱之為期末覺知。如此的調查填寫安排，主要是希望學習者在經歷過八週的混成式學習羽球課程後，自身在動機信念、學習策略、和合作體驗等三項目的內在遷移感受。期初每份問卷各回收 102 份，期末問卷則由於 6 位學生退選，每份問卷各回收 96 份，有效問卷為 96 份。

根據收集到的問卷，進行相關統計分析；根據期初覺知和期末覺知進行成對樣本 t 檢定，以呈現在動機信念、學習策略、和合作能力三方面，學習者的遷移改變。

3. 研究結果與討論

3.1. 學習者在動機信念的遷移表現

從表 1 觀察到研究對象在動機信念面開學時到 8 週學習單元結束時，他們在「內在價值」($t=2.32, p=.002$)、「考試焦慮」($t=5.71, p<.001$)、「自我期許」($t=3.68, p<.001$)和「學習目標」($t=4.13, p<.001$)此四向度的覺知遷移均達到顯著。本研究推測主要是因為課程設計採混成式，在課堂上教師的授課、和隨時隨地可查看線上數位教學影片，有助於他們在學習時縱使遇到困難，也能從平台上取得支援。

在「考試焦慮」該向度，由期初的考試焦慮，而經過幾週後的課程，學習者不僅是課堂上教師的示範，且能夠於 e 學院平台即時取得期需要的學習資源，在改善與掌握技能後，對於後續的考試也不再那麼焦慮。而「內在價值」、「自我期許」和「學習目標」三個向度達到顯著提升，本研究推測一則是混成式教學給學習者有困難時的有效支援；此外以合作學習進行分組，在團隊組員互相砥礪之下，也漸漸地建立明確的學習目標，進而原本恐懼羽球技

能的學生，也能在同儕影響力之下對自我挑戰的信心提升（內在價值），進而更能期許正向的自我學習效果。

表1 羽球課程設計在動機信念之成對樣本 t 檢定

向度		Mean	SD	t
內在價值	期初	3.98	.78	2.32**
	期末	4.20	.80	
考試焦慮	期初	3.95	.65	5.71***
	期末	3.13	.69	
自我期許	期初	3.36	.92	3.68***
	期末	3.76	.96	
學習目標	期初	3.62	.82	4.13***
	期末	4.04	.73	

*p<.05；**p<.01；***p<.001

3.2. 學習者在學習策略的遷移表現

從表 2 觀察到研究對象在學習策略面開學時到 8 週學習單元結束時，他們在「考試準備」(*t=3.65, p<.001*)、「情緒管理」(*t=2.97, p=.004*)和「資源管理」(*t=3.60, p<.001*)此三向度的覺知遷移均達到顯著，而「動機管理」(*t=1.93, p=.056*)該項也趨向於顯著。本研究推測主要是因為課程設計採混成式，對於學習者而言，他們在課堂上觀看過教師的示範動作後，若是練習時有疑慮，除了可以求助教師即時的指導外，也可以上 e 學院平台觀看相關單元的多媒體數位教材。如此一來，縱使無法馬上獲得教師的指導，也能夠在學習平台上取得學習資源。所以在考試準備該向度，從一開學到第 8 週，研究對象更有信心來面對學習期間的考試準備。

表2 羽球課程設計在學習策略之成對樣本 t 檢定

向度		Mean	SD	t
考試準備	期初	3.82	0.67	3.65***
	期末	4.14	0.77	
情緒管理	期初	3.95	0.65	2.97**
	期末	4.20	0.69	
資源管理	期初	3.91	0.71	3.60***
	期末	4.25	0.69	
動機控制	期初	4.10	0.60	1.93
	期末	4.27	0.77	

*p<.05；**p<.01；***p<.001

此外，該羽球學習課程，除了每週兩小時的授課學習外，學習者還能夠隨時隨地的上 e 學院平台查看多媒體教材。相較於沒有該學習平台輔助的課程，對於學習者而言更能夠依照自身便利的時間，排定適合的學習時段。換言之，對學習者相對而言，學習資源的使用更加有彈性，應該是導致於資源管理面的覺知反映顯著提升的主要原因。而由於資源管理和考試準備兩方面，對於學習者而言有較穩定之表現，反應在情緒管理面，相較於課程開始之際和進行八週後的學習，學習者也能掌控到情緒方面的應對。

3.3. 學習者在合作能力的遷移表現

表 3 顯示研究對象於合作能力該項目「認同感與榮譽感」(*t=2.06, p=.042*)、「管理與協商」(*t=2.79, p=.007*)和「表達與溝通」(*t=2.48, p=.015*)三個向度均達到顯著。本研究推測，在學習活動開始之際在此三個向度雖有較高的覺知表現，而在學習活動結束時依然有顯著的增

加原因，應該是隨著小組間學習活動進行，和期末的團體競賽接近的影響，讓學習者以小組為單位，進行各項單元學習、觀摩、和配打，且在學習過程中，也以取得團體競賽較優成績為目標，進行多次每個打點成員的調整與配合；以上種種的行為進而提升了學習者在此三個向度的覺知表現。

此外，在「管理與協商」面為合作能力的四個向度中是最顯著的。本研究推測是期末辦理的分組團體循環賽，有各打點組員規劃與分配對應於團體目標（勝率）的考量，因此在八週課程期間，隨著每位組員技能方面的精進與純熟，彼此間會不斷的進行協商和調整各打點安排。在此過程中，有助於提升研究對象在團隊合作時，覺知到管理與協商的重要性。而在「反思與評鑑」(*t=1.60, p=.113*)該向度雖未達顯著，本研究推測為學習者的期初反應已經是四個向度中最高，換言之就是已經有較高度肯定，而在活動進行的過程中，要再達到顯著提升已經有其限制所在。

表3 羽球課程設計在合作能力之成對樣本 t 檢定

向度		Mean	SD	t
認同感與榮譽感	期初	4.09	.63	2.06*
	期末	4.23	.65	
管理與協商	期初	4.03	.60	2.79**
	期末	4.22	.63	
表達與溝通	期初	4.10	.62	2.48*
	期末	4.27	.64	
反思與評鑑	期初	4.13	.67	1.60
	期末	4.27	.74	

*p<.05；**p<.01；***p<.001

4. 結論與建議

本研究以體育課程的進行混成式和合作學習教學，且探討這樣的課程設計對於學習者在動機信念、學習策略、和合作能力等覺知遷移的影響。研究結果證實，該項課程單元的設計，跳脫傳統體育課程的設計，且達到提升學習者的動機信念與學習策略目標。此外，透過小組組員間的互動，和期末團體競賽的共同目標，讓組員間有機會學習如何進行管理與協商。如此一來，不僅讓學習者於體育課程精進運動技能，也讓學習者在此課程建構的氛圍中，強化其動機信念且堅定了學習策略。該研究有助於教授體育課程之教師，跳脫傳統式教學的另一種課程設計。

本研究現階段僅就研究對象的主觀覺知反應進行分析，後續研究有下列建議事項：

- 大學體育課為男女學生一起上課，可就性別進行群類分析。
- 加入學習成效面和這些主觀反應進行關聯性比對，如此可更清楚學習成效和這三類彼此間的關聯。

參考文獻

陳美榕、陳偉瑀（2007）。大學生體育課動機信念之分析研究。載於**臺北體育學院（主編），2007 年臺北體院運動教育國際學術研討會論文集**（頁 85-95）。臺北市：臺北體育學院。

陳偉瑀、陳美榕（2007）。大學生對體育課學習策略之研究。**長榮大學學報，11(2)**，81-93。

陳怡良、李昆霖（2017）。資訊科技融入體育教學對學習動機及成效之影響－以慢動作錄影融入排球低手發球為例。**管理資訊計算，6(1)**，171-186。

黃相瑋、沈淑鳳、陳儷今、蘇榮立（2020）。資訊科技在大學體育課程之應用。**中原體育學報，(14)**，95-106。

李昆霖（2016）。**資訊科技融入體育教學對習動機及成效之影響－以慢動作錄影融入排球低手發為例**。〔未出版碩士論文〕。樹德科技大學資訊工程研究所。

林怡君（2018）。**智慧型手機結合虛擬實境輔助國中籃球教學之行動研究**。〔未出版碩士論文〕。國立台灣師範大學體育學系。

劉與正（2017）。**多媒體輔助教學對學習者自我覺察、自我監控及學習成效影響之研究**。〔未出版碩士論文〕。國立體育大學運動科學研究所。

夏綠荷、林彥男（2018）。數位化的體育教學。**社會服務與休閒產業研究，(7)**，1-9。

楊文輝、吳致維（2010）。培養合作能力的教學模式。**網路社會學通訊期刊，88**，33。資料引自：http://society.nhu.edu.tw/e-j/88/index.htm。

Alavi, M., & Leidner, D. E. (2001). Research commentary: Technology-mediated learning—A call for greater depth and breadth of research. *Information systems research*, 12(1), 1-10.

Fitts, P.M., & Posner, M.I. (1967). *Human performance.* Brooks/Cole.

George C. (2020). 怎麼更有效的學會「動作」？運動技能的三階段學習模型。擷取自：怎麼更有效的學會「動作」？運動技能的三階段學習模型. 我們正處在學習動作的哪個階段？怎麼樣可以幫助我們更好的進入下一個過程？或是我們要… | by George Chan | moodwide | Medium

Hung, H. C., Young Shelley, S. C., & Lin, K. C. (2018). Exploring the effects of integrating the iPad to improve students' motivation and badminton skills: a WISER model for physical education. *Technology, Pedagogy and Education*, 27 (3), 265-278.

Johnson, D. W., & Johnson, R. T. (1992). Positive interdependence: Key to effective cooperation. *Interaction in cooperative groups: The theoretical anatomy of group learning*, 174-199.

Parker，R.E.(1985).Small-group cooperative learning--improving academic, social gains in the classroom. *Nass Bulletin*, 69(479),48-57.

Sallis, E. (2014). *Total quality management in education*. Routledge.

提問引導同儕回饋策略對新詩創作線上學習社群互動參與的影響

Effect of a Question-promoted Peer Feedback Approach on Social Interactions and Engagement in an Online Poetry Writing Learning Community

梁心怡

義守大學創意整合設計學系

cindy@isu.edu.tw

【摘要】新詩創作學習社群多透過舉辦線上或實體讀詩會活動讓學習社群成員可直接討論交流彼此創作的想法，改善新詩創作表現。然而，由於活動時間限制，可能無法讓所有參與者均參與討論交流。為改善參與度不均的問題，本研究提出一提問引導同儕回饋策略輔助線上學習社群活動進行，分析此策略對線上學習社群互動模式與參與度的影響，並透過團體訪談了解線上學習社群成員對此策略的觀感評價。結果顯示此策略可讓更多成員有機會分享自己的意見，然而活動時間長度影響線上學習社群互動模式和成員參與度。

【關鍵詞】 學習社群；新詩創作；同儕回饋；參與度；社群互動

Abstract: In poetry writing learning communities, conducting online or offline learning community activities is the most common way to enable learners to directly discuss their poems and improve their writing performance. However, it is less possible for all learners to share their comments in a limited time of learning community activity. For this reason, this study proposed a question-promoted peer feedback approach to encourage engagement. The effect on social interactions and engagement in an online poetry writing learning community is investigated, and learners' perceptions are also explored through a group interview. The results show that this approach could encourage more learners to share their comments in a limited time. However, the length of time affects social interactions and engagement in online learning communities.

Keywords: learning community, poetry writing, peer feedback, engagement, social interaction

1. 前言

學習社群(Learning Community)是指一群具有共同學習興趣或學習目標的成員所組成，透過持續分享與互動，可改善成員對該領域專業的知識與技能(Liu et al., 2010)。透過部落格、社群媒體或線上討論區等數位媒體可提供對新詩創作有興趣的創作者一交流互動空間，讓創作者有機會在公開場合發表自己的詩作，接觸並認識其他創作者，從他人回饋中反思學習，改善其新詩創作表現(Novakovich, 2016)。此外，創作者也可透過參與學習社群的活動，如定期組織的讀詩會活動，與其他創作者即時互動，透過討論獲得更多創作的啟發和改善建議，並透過社群成員間的情感支持，改善持續創作動機與精進其創作表現(Naji et al., 2019)。

過去由於疫情影響，許多實體讀詩會活動改為線上會議方式舉行。無論透過實體或線上讀詩會活動，均可提供創作者一即時互動的管道讓創作者與其他創作者深入交流，並讓新詩初學者有機會與資深創作者或知名詩人對談，獲得創作的實質幫助，然而由於活動時間限制，可能無法讓所有參與成員對作品陳述意見，因此唯恐讀詩會成為資深創作者或詩人老師的「一言堂」，難以讓創作者多方獲取其他不同的意見回饋(Naji et al., 2019)。除了參與度不均

的問題外，過去研究也發現創作者不一定能在讀詩會活動獲得有用的回饋：一則是因顧及創作者的感受，參與者多給予支持性的情感回饋，缺乏實質性的批判意見，因此不利創作者改善其創作表現(Guasch et al., 2019)；二則是大多數初學者由於缺乏相關知識與經驗，較難發覺他人詩作的問題，因此難以給予深入的回饋建議(Resien et al., 2020)。過去研究發現許多社群成員傾向保持沉默，較少發表自己的意見，這更加劇多方收集不同參與者回饋的困難(Novakovich, 2016)。

為促進線上學習社群中的成員主動參與即時互動討論，在有限的活動時間內多方收集不同參與者的回饋意見，過去研究發現適切同儕回饋策略輔助可促進學習社群的參與度並改善成員間互動交流，幫助創作者統整想法，改善創作表現(López-Pellisa et al., 2021; Lin, 2019)。其中，提問(Questioning)可引導參與者檢視詩作，改善批判思考，幫助參與者反思，給與作者更多認知及後設認知的回饋 (Guasch et al., 2019)。除了提問策略，W-S-Q(Watch-Summary-Question)也是一種常被使用的同儕回饋策略，透過引導學習者觀看影片(W)，摘要註記觀看重點(S)，提出自己的反思問題(Q)，可引導學習者有效反思學習，改善學習成效(Hsia et al., 2022; Hsia & Hwang, 2020)。然而，過去研究同儕回饋策略大多應用於非同步互動之情境，著重於分析對學習成效之影響，較少研究探究應用於同步互動情境中的影響，因而仍需進一步研究了解同儕回饋策略對學習社群互動行為及學習成效的影響(Li et al., 2020)。

綜上所述，本研究提出一提問引導同儕回饋策略，希望透過結合提問引導與 W-S-Q 策略，可有效改善線上學習社群的即時互動討論交流，促成更多社群成員間的雙向互動，藉由導入提問引導同儕回饋策略，引導參與者透過閱讀他人作品、註記詩作閱讀心得與想法、提問、以及與他人的討論互動中，提供創作者更多實質的批評分析與創作建議，並藉此反思自身創作的不足，提升創意寫作能力。期許透過此方法可改善線上學習社群中的社群討論互動，引導參與者獲得更多不同的意見，讓創作者可從他人的質疑和對自身作品的辯白中，助益提升其個人創作表現。本研究為評估提問引導同儕回饋策略是否改善新詩創作線上學習社群成員的參與討論，收集分析該學習社群的討論互動對話及表單回饋資料，其研究問題如下所述：

1. 提問引導同儕回饋策略如何影響線上學習社群互動？
2. 線上學習社群成員如何評價提問引導同儕回饋策略？

2. 線上學習社群

2.1 學習情境

本研究以臺灣一個新詩創作社團的線上學習社群為研究對象。該社團為跨校學生社團，成員主要為臺灣各大專院校及高中在學學生，以學校所在地區將社團成員分為北區、中區、南區、東區四大區，每月均辦理分區讀詩會活動，社長擔任活動主持人，邀請欲參加之學習社群成員繳交一首自己的詩作，由分區總召負責彙整投稿詩作，約定時間聚會，交流討論彼此的詩作。本次研究以該學習社群所辦理的一場線上讀詩會活動為分析之學習情境，該活動透過 google meet 線上會議辦理，活動前一天分區總召將彙整所有投稿詩作，發佈於線上學習社群提供給參與成員參考。當天活動由主持人管控整體進行流程，分為兩大階段(圖 1)：

(1) 與談詩人導讀講座：

邀請具有豐富得獎與出版經驗的詩人老師進行詩選導讀，由與談詩人主導，與主持人共同引導參與者共同閱讀名家詩作，了解如何解構與營造詩意，創建個人風格美學，並與參與者討論分享彼此的詩觀與創作理念，講座時間約 1 小時。

(2) 成員創作分享與回饋：

由主持人管控討論進行流程，分區總召依照繳交作品順序請作者朗讀自己的創作，而後主持人將請有意願發言的參與者將闡述自己的意見，或點名特定參與者發言，與作者和其他參與者對談討論彼此對此作品的想法，而後由與談詩人進行最後講評。本次活動總計有 14 首詩作參與討論，每首詩作討論時間約15至30分鐘，整體討論時間約7小時(如圖1所示)。

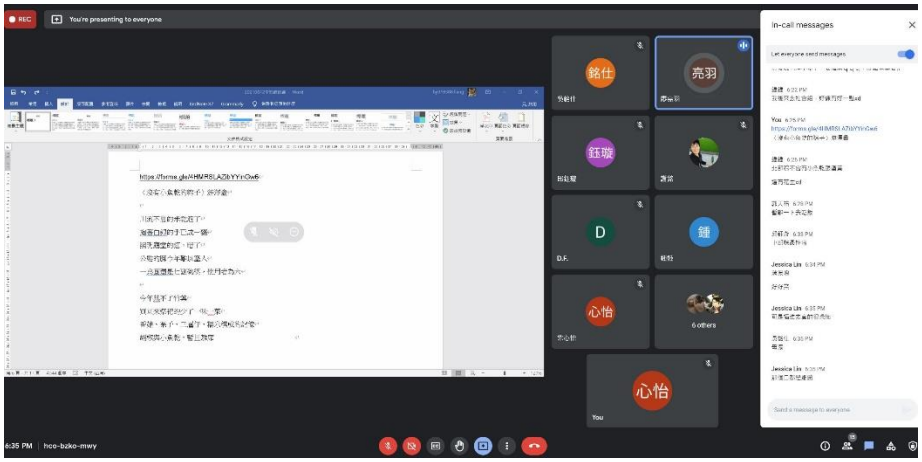


圖 1 、線上學習社群活動學習情境

2.2 提問引導同儕回饋策略設計

為鼓勵參與社群成員均能踴躍參與討論，本次活動應用提問引導同儕回饋策略鼓勵參與者在討論詩作時分享自己的意見，每首詩均提供一專屬回饋表單，在討論該詩作前，研究者將提供此回饋表單給所有參與者。討論該詩作時，參與者可與作者對談討論，或是在閱讀詩作後填寫回饋表單給予回饋。回饋表單參考 Yamaç et al. (2020)所編撰的評分尺規，透過提問協助參與者從語意表達、節奏音韻、修辭技法、創意表現四方面檢視詩作，給予 1 分(差)、2 分(佳)、3 分(優)的評分，其提問引導檢核項目如表 1 所示：

表 1、提問引導檢核項目

面向	提問檢核項目
語意表達	1. 作品是否內在邏輯清晰，可傳達作者之情懷與思想？ 2. 作品中是否有線索模糊或邏輯上的斷裂? 3. 作品中是否含有晦澀用語或專有名詞不易讓讀者理解詩意?
節奏音韻	1. 作品是否結構精緻？ 2. 作品是否營造良好適切的節奏音樂性？
修辭技法	1. 作品用語是否包含需要再精煉的語言或不精確的用詞？ 2. 作品中是否包含錯字？
創意表現	1 作品是否具有強烈明顯的個人風格？ 2 作品是否包含創新性的主題、想法、創作手法或佳句？

參與者完成評分後，可上傳自己閱讀該詩作的註記圖片，或寫下建議、問題或意見提供給作者參考(圖 2)。



圖 2、提問引導同儕回饋單

作者也可填寫表單自我評估，點選「觀看得分」可觀看得分評語，點選「其他回覆」可瀏覽所有參與者給予的評分與回饋意見。作者可主動回應回饋內容，或是參與者針對回饋內容作補充論述。若參與者無其他意見欲討論，主持人請與談詩人講述總評語建議，並在討論結束前，主持人提醒作者檢視回饋單意見，以利針對想回應的意見或問題做最後回覆與討論，整體進行流程如圖 3 所示：



圖 3 、提問引導同儕回饋策略流程

3. 實驗設計

3.1 實驗對象

本研究分析參與線上讀詩會活動的學習社群互動行為，共有 21 位成員參與本實驗，其中 2 位為高三學生、13 位為大專院校及研究所學生；6 位為已畢業之社會人士，包含 4 位已畢業之成員、主持人、與談詩人。其中 9 位有校內或國內文學獎得獎經驗，其餘 12 位無得獎經驗：參與男性 11 人，女性 10 人。

3.2 實驗流程與研究工具

活動開始前，研究者先向學習社群成員解說本次活動流程，示範與解說提問引導同儕回饋策略使用方式，而後請參與成員填寫基本資料問卷，以利了解參與者相關資訊，如：年齡、性別、學校，以及新詩創作學習經驗，如：參與學習社群活動經驗、新詩創作及出版經驗、新詩閱讀習慣以及得獎經驗。而後由與談詩人進行約 1 小時的詩選導讀講座，講座結束後開始第二階段活動，由主持人主持討論活動，分區總召請參與者依序朗讀自己的創作，主持人及其他參與者與作者對談討論詩作，最後由與談詩人進行總講評回饋。在每首詩討論前，研究者將提供此首作品的回饋表單連結給所有參與者，討論過程中，參與者可使用提問引導同儕回饋策略共同檢視詩作及補充回饋意見，作者可與主持人以及參與者討論詩作，或瀏覽回饋單意見回應或補充論述，無論作者是否回應回饋單內容，最後主持人都會請作者再次檢視回饋單意見做最後的補充論述討論，討論結束前並請與談詩人做總結評語回饋。整體活動時間共約 8 小時，活動結束後進行約 30 分鐘的群體訪談，請參與者就本次使用提問引導同儕回饋策略講述學習觀感，評估是否有助於改善社群學習成效，並舉例說明提問引導同儕回饋策略的優缺點，以及改善建議與想法，以利參考訪談回饋對實驗數據分析結果做進一步之詮釋與理解。

3.3 實驗數據收集分析

本研究收集分析線上學習社群成員之互動行為，其分析數據來源為第二階段活動之討論對話紀錄及提問引導同儕回饋單所收集之回饋內容，以利了解參與者參與情況，從參與互動人員、參與人數、發言次數與回饋數量等社群互動行為分析紀錄，探究提問引導同儕回饋策略對線上學習社群互動的影響：(1) 活動討論對話紀錄：活動全程錄影，而後透過 Whisper AI 工具轉錄其語音為文字內容，並經由人工校對參與者說話之對話內容後進行統計分析，以參與者每次發言之每則對話內容為一次發言為分析單位進行分析，分析參與者發言次數和參與活躍階段；(2)回饋單內容：提取參與者撰寫之意見作為分析編碼內容，以參與者所提交的每則回饋內容為分析單位進行分析，分析參與者發表回饋次數及參與活躍階段。

參與活躍階段以每首討論的詩作為單位，從討論和回饋分析參與者的社群互動參與。此外，將每首作品討論對話紀錄過程編碼分析線上學習社群使用該策略的討論互動模式及討論時間，其編碼依據參與者主動提出回饋單意見討論的時間點切割為不同階段，編碼表如表 2 所示：

表 2、線上學習社群討論階段編碼表

編碼	階段	對話範例
S	開始討論詩作	下一位請良念一下她的作品
F	討論回饋單	我有看到回饋單的意見，我想問一下
C	停止討論回饋單	應該沒有什麼要討論的了
E	結束詩作討論	好，我們謝謝良

訪談內容使用內容分析，依據相似性及關連分類編碼對話內容，以利了解線上學習社群對此策略之觀感評價。

4. 結果與討論

本研究共收集 14 首作品之討論對話紀錄及其回饋單內容作為分析數據來源，總計收集 21 位線上學習社群參與者共 7 小時之 10128 句對話，共 1808 筆發言次數之對話紀錄，以及 14 張回饋單共 98 筆回饋內容。每首詩作的討論時間及參與統計結果如表 3 所示：

表 3、線上學習社群討論時間及參與行為描述性統計結果

詩作編號	討論時間(分:秒)		發言次數		回饋單意見
	線上對談	討論回饋單	線上對談	回饋單討論	
1	14:15	03:55	61	20	11
2	16:41	01:20	53	2	16
3	12:33	05:21	49	7	9
4	25:25	02:32	67	3	9
5	26:56	03:42	151	8	9
6	17:08	00:00	91	0	9
7	25:02	00:24	96	4	8
8	22:38	00:00	169	0	5
9	31:42	00:00	134	0	5
10	38:56	00:00	182	0	5
11	22:21	00:00	75	0	4
12	33:52	00:00	151	0	7
13	42:28	00:00	139	0	4
14	49:36	05:41	140	14	7

進一步分析線上學習社群使用提問引導同儕回饋策略的社群互動模式，活動前半場之特群互動模式大多為 S→F→C→S→E，討論詩作時由主持人與作者及參與者對談到一段落後，才會檢視回饋單意見做進一步討論，而後由與談詩人與作者對談進行總結回饋；然而活動後半場除了最後一首有參與者主動討論回饋單意見，其他多未就回饋單意見再做延伸討論(S→E)。

分析參與者參與討論之互動情況，由圖 4 可看出主持人及與談詩人發言次數最多，除主持人、與談詩人和少數參與者全程參與討論互動，有較多的發言次數，多數參與者僅參與部分詩作討論互動，而未繳交詩作的參與者(15-19)多為旁聽，較少發表自己的意見。以活動時間分析社群互動參與情形，Wilcoxon Sign Rank Test 結果顯示活動後半場的社群互動參與(平均 4.24，標準差 4.15)低於活動前半場(平均 6.14，標準差 3.98)，趨近顯著差異(p=.052)。許多參與者討論完自己的詩作後就不再參與互動，分析參與者互動的方式，發現主持人與談詩人和未繳交詩作的參與者傾向直接發言討論交流，多數參與者並無顯著偏好之社群互動方式，大多均參與討論發言和給予文字回饋互動(p>.05)。

由圖 4 結果可發現，每首詩參與討論人數有限，很難讓所有參與成員發表自己的意見，訪談意見回饋說明參與者(8,14)認為提問引導同儕回饋策略有助於在參與者較多時，幫助作者可收集到更多不同的回饋意見;與談詩人補充說明除此之外，透過閱讀回饋可幫助作者再次統整所得到的資訊，幫助作者反思得到更多啟發。然而提問評分設計有較多爭議：主持人認為評分回饋制度可能不適合線上學習社群的使用情境，因許多同學分享詩作的目的不在於參與文學獎或投稿，而是藉由創作抒發心聲，希望透過創作找到知己，此外也提到初學者評分的困難，擔心較嚴厲的批評會打擊初學者創作的信心與動機，建議以鼓勵參與者發表自己意見即可。從社群互動參與分析結果也可看出旁聽成員較偏好以討論形式互動，參與者(16)在訪談回饋中解釋因有具名評分的設計，不好意思填寫回饋。與談詩人則提到文學獎評分標準的分歧和詩觀的多元特性，因此可能需要更謹慎檢視評分標準的設計，避免侷限或誤導參

與者創作。此外，提問引導同儕回饋策略使用時機可提早至活動開始前，讓參與者有更充裕時間填寫回饋單提供意見。

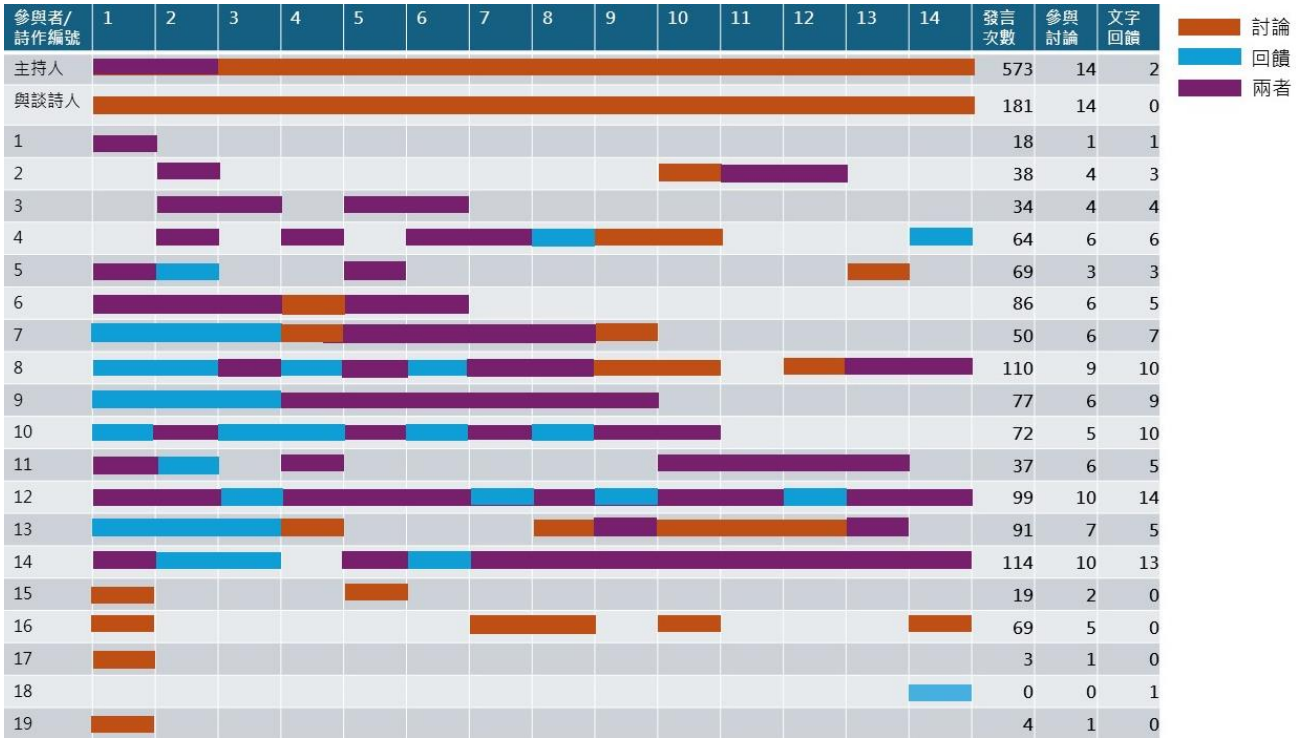


圖 4、線上學習社群成員參與過程與互動方式

5. 討論與結論

本研究提出一提問引導同儕回饋策略輔助線上學習社群進行新詩創作討論分享，結果顯示透過填寫回饋單可收集到更多參與者的回饋，作者也可透過檢視回饋單，進一步補充論述或與參與者從各面向更深入討論詩作。然而，活動時間影響社群互動模式及參與度，對活動後半場發表詩作的參與者較不公平，因此建議每次活動討論時間可依據不同社群規模與活動類型適度調整，或是透過事前收集回饋、限制無關發言或每次發言時間等方式改善討論效率。

研究結果也指出參與者對評分回饋設計的顧慮，呼應 Guasch et al. (2019)研究結果發現參與者可能因情感顧慮因此較少提出批評性回饋。公開場合的具名討論回饋儘管有助於學習社群增進情感交流，改善創作者的自信心，但相較於匿名回饋或一對一的私密回饋而言，不利於收集建設性的批判意見(Zhang et al., 2022)。因此學者建議可結合具名回饋與匿名回饋互動方式，讓參與者可免除顧慮暢所欲言，以利增進討論深度(Liu et al., 2021)。

本研究初步探究提問引導同儕回饋策略對新詩創作線上學習社群互動參與的影響，其結果可提供教育相關研究者參考，以利了解線上學習社群互動與學習方式。然而，受限於學科差異，本研究結果可能不適用於科學或其他學科，關於學習社群之特性差異之影響仍待進一步研究。為更深入了解學習社群如何透過社群互動共享知識與學習，未來可結合量化研究方法評估學習成效，或透過行為序列分析、社會網絡分析、知識網絡分析等統計分析方法，進一步了解學習社群的社群互動行為及知識建構過程，以利深入了解影響學習社群學習的相關因素。此外，未來研究可就具名與匿名回饋等不同回饋方式進一步研究對社群學習之影響，也可針對學習者之個別差異，如先備知識和新詩創作成就等因素探討社群學習之差異，以利提供更適切的學習輔助工具或策略，改善線上學習社群的學習成效。

參考文獻

Guasch, T., Espasa, A., & Martinez-Melo, M. (2019). The art of questioning in online learning environments: the potentialities of feedback in writing. *Assessment & Evaluation in Higher Education*, *44*(1), 111-123.

Hsia, L.-H., Hwang, G.-J., & Lin, C.-J. (2022). A WSQ-based flipped learning approach to improving students' dance performance through reflection and effort promotion. *Interactive learning environments*, *30*(2), 229-244.

Hsia, L. H., & Hwang, G. J. (2020). From reflective thinking to learning engagement awareness: A reflective thinking promoting approach to improve students' dance performance, self-efficacy and task load in flipped learning. *British Journal of Educational Technology*, *51*(6), 2461-2477.

López-Pellisa, T., Rotger, N., & Rodríguez-Gallego, F. (2021). Collaborative writing at work: Peer feedback in a blended learning environment. *Education and Information Technologies*, *26*(1), 1293-1310.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education*, *45*(2), 193-211.

Lin, C.-J. (2019). An online peer assessment approach to supporting mind-mapping flipped learning activities for college English writing courses. *Journal of Computers in Education*, *6*(3), 385-415.

Liu, F., Du, J., Zhou, D. Q., & Huang, B. (2021). Exploiting the potential of peer feedback: The combined use of face-to-face feedback and e-feedback in doctoral writing groups. *Assessing Writing*, *47*, 100482.

Liu, I.-F., Chen, M. C., Sun, Y. S., Wible, D., & Kuo, C.-H. (2010). Extending the TAM model to explore the factors that affect intention to use an online learning community. *Computers & Education*, *54*(2), 600-610.

Naji, J., Subramaniam, G., White, G., Naji, J., Subramaniam, G., & White, G. (2019). Literature and reading communities. *New Approaches to Literature for Language Learning*, 129-148.

Novakovich, J. (2016). Fostering critical thinking and reflection through blog-mediated peer feedback. *Journal of Computer Assisted Learning*, *32*(1), 16-30.

Resien, C., Sitompul, H., & Situmorang, J. (2020). The effect of blended learning strategy and creative thinking of students on the results of learning information and communication technology by controlling prior knowledge. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *3*(2), 879-893.

Yamaç, A., Öztürk, E., & Mutlu, N. (2020). Effect of digital writing instruction with tablets on primary school students' writing performance and writing knowledge. *Computers & Education*, *157*, 103981.

Zhang, M., He, Q., Du, J., Liu, F., & Huang, B. (2022). Learners' perceived advantages and social-affective dispositions toward online peer feedback in academic writing. *Frontiers in Psychology*, *13*, 973478.

Using Chatbots to Assist English Learners in Reading Strategies

Chiou-Hui Chou[1*], Si-Yu Liu[2]
[1] National Tsing Hua University
[2] National Tsing Hua University
[*] chou@mx.nthu.edu.tw

Abstract: *Reading comprehension strategies are important for English readers to understand the content of a text. In response to the development trend of mobile applications and social media, this study built a chatbot that can fill in text and guide students to develop reading comprehension strategies through the free platforms of LINE and Google. The researchers collected the background data of the chatbot to record the students' reading process and analyzed the feasibility of using chatbots to assist students in EFL reading comprehension strategies. This study employed a mixed research method to collect data. Six Chinese folktales and idiomatic stories translated into English were used as the text in this study. Nine fifth graders in an elementary school in Taiwan participated in this study. Analyzing the students' pre-test and post-test results, the learning process records of the chatbot backend, and the students' questionnaire feedback, the researchers found that the chatbot-assisted reading project had a positive impact on the students' reading enhancement. Suggestions and implications were made for future research.*

Keywords: chatbot, EFL, elementary students, English reading, reading comprehension strategies

1. Introduction

Reading is an active process that requires a lot of practice and skills. Reading comprehension strategies are important tools that proficient readers use to solve the difficulties they encounter in a text. Learning reading comprehension strategies can be an effective way to improve reading for second language learners. With the development of computer- and mobile-assisted language learning (CALL and MALL), chatbots are becoming increasingly popular in the English classroom. Many chatbots that can serve as learning partners, teaching agents, and learning assistants have been developed as language learning tool (Lokman & Ameedeen, 2019). Despite the positive impact that chatbots can have on learning and living, there is still a lack of research related to EFL elementary school students in reading comprehension.

To follow the technology trend, the researchers decided to use a LINE chatbot as an instructional aid. Through LINE and Googles free platform, pre -designed reading comprehension lessons were imported into the LINE chatbot, which can be used to guide young English learners to think about and acquire reading strategies during the reading process. By analyzing the effectiveness of the chatbots in actual lessons, the researchers analyzed the feasibility of chatbots as a reading teaching assistant tool.

2. Literature Review

Chatbots are programs that are developed to hold conversations with humans. Due to the rapid growth of chatbots, researchers have used chatbots in language classrooms. According to Fryer and Carpenter (2006), students preferred to communicate with chatbots rather than talk to their classmates or teachers. Compared to traditional e-learning systems, chatbots are more interactive. One of the reasons for the popularity of chatbots is the increased efficiency of the learning process, as they can increase user and student satisfaction by speeding up the response time and can answer or clarify any question. It allows teachers to avoid answering repetitive questions that can be easily answered by chatbots (Mageira et al., 2022). Chatbots can simply provide question-and-answer functions between students and the system. Chatbots can also create a friendly learning environment for students, which can effectively reduce stress and increase willingness to language study. Moreover, chatbots are excellent data analysis tools for teachers (Mageira et al., 2022).

In this study, the researchers referred to the reading comprehension strategies and instructional frameworks mentioned by Moreillon (2007) to design the research framework. The seven reading comprehension strategies employed in this study include questions of details, predicting and inferring, using sensory images, determining main ideas, summarizing, activating background information, and making judgments.

Recently, chatbots in reading context have gained growing interest, which assume various roles such as tutor, companion, mentor, facilitator, and tutee (e.g., Xu et al., 2021). Chatbot designs in the research address the need to improve reading comprehension, engagement, and affective states. Selected studies on chatbot-assisted EFL reading instruction that matched this study for children are reviewed (Cheng et al., 2024; Chu, 2021; Liu, 2022; Liu et al., 2024; Xu, 2021; Xu, 2022). Their main themes for elementary students were exploring the role of chatbots in engaging students in reading and increasing their interest in reading (e.g., Liu, 2022; Liu et al., 2024). Still, the role of chatbots in improving students reading comprehension strategies was not explored. Thus, this study set out to explore how chatbots can assist develop EFL elementary students' reading comprehension strategies via reading English stories.

3. Methodology

The participants in this study were nine fifth-graders in Taiwan. The research tool is the chatbot built by the researchers. The chatbot links the LINE official account manager with the backend of Google Drive and the researchers imported the designed stories into the LINE chatbot through the backend form of Google Drive. Students communicated with the chatbot and learned English reading comprehension strategies by using the LINE app. The researchers obtained the students' learning records through the records of Google Drive and analyzed them.

Six English texts adapted from Chinese folktales and idiomatic stories were used in this study. Considering the students' English level, they were written by the researchers. Folktales included *The Snail Girl* and *The Story of Sun Moon Lake.* Idiomatic tales included *Mr. Ye Loves Dragon*, *Guard A Tree-Stump and Wait for Rabbits*, *Pulling Up Crops to Make Them Grow Faster*, and *Adding Feet to A Snake*. The researchers designed comprehension questions for each story. Each story includes seven questions and contains at least four different types of comprehension strategies based on the content of the stories. There are 6 questions for details, 5 for predicting and inferring, 6 for using sensory images, 5 for main ideas, 7 for summarizing, 5 for activating background information and 10 for making judgments. The format of the research process is based on a question-and-answer dialogue, in which the chatbot responds with feedback, and students need to write their answers. Most of the questions are open-ended, except for those related to detailed information about the story. The design of one sample text and the interface of the chatbot is illustrated in Figure 1.

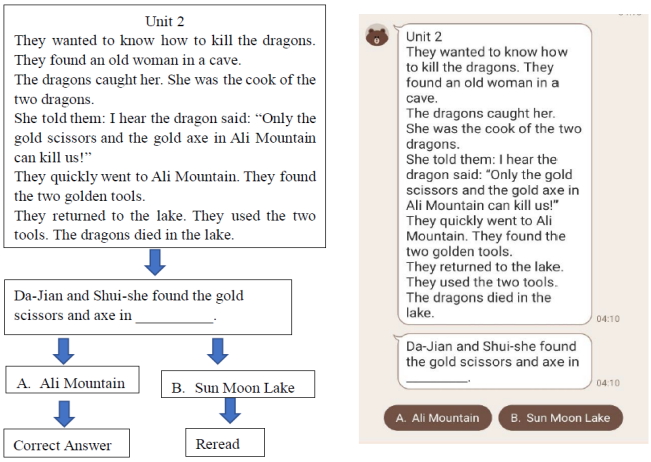


Figure 1. The design of the text and the interface of the chatbot.

This study lasted for five weeks. During the first week, the students took a pre-test of a reading text, *The Snail Girl*. The pre-test was in a paper-pencil format. They could answer the questions in either Chinese or English based on their comprehension of the article because this study mainly looked into the students' reading comprehension ability. During

the second week, the students' English teacher used *Mr. Ye Loved Dragons* as an example to teach them the process and method of using the LINE chatbot. During Week 3 and Week 4, participants started to work independently in the chatbot reading session. The teacher did not intervene in the learning process, but only provided some help on the use of the chatbot. Four stories were arranged for students to read during the two weeks in different days. During Week 5, students took a post-test, same as the pre-test. A feedback questionnaire was used to collect their opinions and attitudes about this study. The questionnaire consisted of fifteen questions that explored students' perceptions and attitudes toward reading comprehension strategies learning with chatbots. The responses were categorized into five levels, in this order as strongly agree (5), agree (4), neutral (3), disagree (2), and strongly disagree (1).

4. Results and Discussion

Comparing students' pre and post-tests, the researchers found that some students whose answers were incorrect in the pre-test corrected their errors in the post-test. Some students corrected previous answers that were not relevant to the topic in the posttest. Students' responses were more relevant to the theme of the story comparing to the pre-test. In addition, some students' responses changed from phrases to complete sentences. The details of their responses were observed. Moreover, when confronted with multiple questions, most students showed the phenomenon of listing point responses in the post-test, a feature that did not appear in the pre-test. Furthermore, while the majority of the students would answer in Chinese, a few students attempted to answer in full English on the post-test.

In regard to the students' learning with the chatbot project, the results of the questionnaire are shown in Table 1. Only two items, Item 5 and 11, did not show a mean above 4. Basically, this chatbot project has shown a positive impact on developing students' reading comprehension strategies and on learning English.

Table 1. Results of the questionnaire.

Questions	Means
1. I find the chatbot English reading program very interesting.	4.3
2. I like the chatbot English reading program.	4.0
3. I think the chatbot English reading program helps me to learn English.	4.2
4. I think my English reading has improved through the chatbot English reading program.	4.3
5. I feel that I enjoy learning English more after taking the chatbot English reading program.	3.7
6. I have looked at the illustrations in the articles of the chatbots carefully.	4.1
7. I think reading the illustrations can help me understand the articles.	4.0
8. During the reading process, my life experience can help me understand the story.	4.3
9. I stop and think about the story during the reading process.	4.4
10 During reading, I read forward and backward to help me understand sentences I don't understand.	4.1
11. During the reading process, I guess what will happen next in the story.	3.7
12. During the reading process, I think about what the story is trying to teach me.	4.1
13. During the reading process, I evaluate the characters in the story (e.g., I think about what they did well and what they didn't do well).	4.1
14 After reading, I can tell when and where the story took place and what happened in the story.	4.1
15. After reading the story, I usually think about the problems the characters had and how they could solve them.	4.2

Based on the pre- and post-tests as well as the questionnaire feedbacks, firstly, it can be seen that using chatbots as an assistant is effective in facilitating students' reading comprehension strategy development. The effectiveness is shown by the fact that more information in the story can be elicited and understood, capturing more details about the story is increased, and the theme of the story can be better understood and judged. Students indicated that they would start to stop and think about the storyline during the reading process and to connect the story content to their life experiences. They would make connections and activate knowledge that they have learned before. After reading a story, the students indicated they would think about the problems that the characters encountered and consider methods to solve their problems. This study demonstrated that the students comprehended the story better after this chatbot-guided conversation program. This result was not surprising given that the vast education literature documenting the added value of guided

conversation (e.g., Xu, et al., 2021). Secondly, students showed relatively positive attitudes towards the chatbot reading lessons. They affirmed the interesting nature of the project and said it had a positive effect on their English reading. This study confirms chatbot learning activity can significantly enhance students' reading interest and engagement (Liu et al., 2024) as well as develop reading comprehension strategies.

However, based on the questionnaire data, it can be seen that students were less inclined to actively think about the next development of the story during the reading process. The main reason might be that they already had a certain understanding of the idiomatic stories. The second reason might be that the short length of the stories did not leave room for reflection. Therefore, in the future, researchers may include longer texts for students to learn. One thing we would like to note is that in this study the chatbot English learning experience did not correlate to their motivation to learn English. In the future, researchers might explore some more aspects in learning English and motivation.

5. Conclusion

This study has demonstrated that a chatbot as a reading partner is adequate for developing EFL learners' reading comprehension strategies. Students have positive attitudes towards chatbot reading lessons. They affirmed the interesting nature of the project and confirmed it had a positive effect on their English reading. The results of this study add to the literature of chatbot partner reading in facilitating English learners' reading comprehension strategy development.

However, the researchers also encountered difficulties, which was the application of LINE accounts. During the preparation stage, the researchers needed to apply the LINE account for the participants because they were too young to have their own LINE accounts. The application of the account relied on a cell phone number, which was equivalent to increasing the pressure of time and money for the preparation of this project. Future research might also take this aspect into consideration when implementing similar projects.

Reference

Cheng, X., Yin, L., Lin, C., Shi, Z., Zheng, H., Zhu, L., Liu, X., Chen, K., & Dong, R. (2024). Chatbot dialogic reading boosts comprehension for Chinese kindergarteners with higher language skills. *Journal of Experimental Child Psychology*, *240*, 1-10.

Chu, S.-Y., & Min, D. (2021). Development of an AI chatbot-based teaching model for English picture book retelling activities. *Modern English Education, 22*(4), 37-50.

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology, 10*(3), 8-14.

Liu, C.-C., Chen, W.-J., Lo, F.-y., Chang, C.-H., & Lin, H.-M. (2024). Teachable Q&A agent: The effect of chatbot training by students on reading interest and engagement. *Journal of Educational Computing Research*, *0*(0), 1-33.

Liu, C.-C., Liao, M.-G., Chang, C.-H., & Lin, H.-M. (2022). An analysis of children's interaction with an AI chatbot and its impact on their interest in reading. *Computers & Education*, *189*, 1-16.

Lokman, A. S., & Ameedeen, M. A. (2019). Modern chatbot systems: A technical review. *Proceedings of the Future Technologies Conference (FTC), 2*, 1012-1023.

Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. *Applied Sciences-Basel*, *12*(7), 1-6.

Mohamed Haggag, H. (2021). Implementing artificial intelligence-powered tools in enhancing TOEFL-ITP test takers' reading and writing skills. *Journal of the Faculty of Education.*, *32*(128), 129-156.

Moreillon, J. (2007). *Collaborative strategies for teaching reading comprehension*. American Library Association.

Xu, Y., Aubele, J., Vigil, V., Bustamante, A. S., Kim, Y. S., & Warschauer, M. (2022). Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement. *Child Development*, *93*(2), 149-167.

Xu, Y., Wang, K., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing Childrens reading with a conversational agent vs. a human partner. *Computers & Education*, *161*, 2-17.

Students' Capacities to Provide Feedback in L2 Writing After Receiving Teacher and Peer Feedback

Peiling Ren[1], Shuya Wang[2], Jiaojiao Ji[2], Yong Wu[1*]
[1 2] School of Humanities, Beijing University of Posts and Telecommunications
[*] wuyong@bupt.edu.cn

Abstract: *Student feedback literacy is a key quality involved in effective learning. The existing studies mainly focusing on students' uptake of feedback varied by different feedback sources, few have tested the benefits of providing feedback in L2 writing. The present study investigated student feedback literacy through how students provide feedback under the pattern of teacher feedback and online peer feedback. 44 college students in an English writing class taking part in 16-week practices of English essay writing, finished a pre-test and post-test of providing feedback on essay revision. The results show: (1) The feedback provided by teacher feedback group and peer feedback group are generally similar in terms of feedback type, but differ in focus and feature; (2) After feedback exercises, the feedback-providing capacities of the two groups are almost consistent; (3) The feedback provided by the peer feedback group is relatively more detailed.*

Keywords: student feedback literacy; feedback-providing capacities; providing feedback; peer feedback

1. Introduction

Second language (L2) writing is an essential skill for college English majors. Due to a shortage of English writing courses in college and a large number of students, it is difficult for students to receive timely, thorough, and detailed feedback from teachers. Peer feedback has entered the perspective of educators, and in-depth researches have been conducted on whether peer feedback can solve the current teaching dilemma (Fan & Xu, 2020; Cheng et al., 2023). By comparing the effects of online peer feedback and traditional teacher feedback on students' capacities to provide feedback, this study aims to provide references for teachers to develop new teaching models, improve student-centered L2 writing classrooms, and cultivate student feedback literacy through cooperative learning.

2. Literature Review

2.1. Student Feedback Literacy and Feedback Provision Capacities

Student feedback literacy is the understanding, capability, and inclination needed to comprehend and use information to improve work or learning strategies (Carless & Boud, 2018). With the emergence of peer feedback, it was redefined as students' capacities and attitudes in providing and receiving feedback (Dong et al., 2023). Nevertheless, existing research has mainly focused on how students utilize feedback (Wood, 2021; Chong, 2020), paying little attention to students' capacities in providing feedback. Some scholars started to emphasize the importance of feedback provision capacities (Wu & Schunn, 2023; Zong et al., 2023). Ying et al. (2023) investigated features and quality of feedback provided and received by high and low L2 proficiency students, and the results indicated that providing feedback (both in quantity and quality) is more conducive to revision than receiving feedback.

2.2. Teacher Feedback and Peer Feedback on Student Feedback Literacy

Some scholars began to focus on how different feedback sources enhance student feedback literacy. Han & Xu (2020) explored how undergraduate students' feedback literacy affected their engagement and uptake of teacher

feedback through a research. Zhang et al. (2023) focused on 3 first-year undergraduate students to explore how peer feedback affects student feedback literacy, which revealed that these students gained significant improvements in their cognitive, behavioral, and emotional abilities after receiving peer feedback training. The current studies focused more on the aspect of implementing feedback, but fewer on providing feedback from different sources.

Therefore, this study aims to explore students' capacities to provide feedback in terms of the type, focus, and feature of feedback after receiving teacher and peer feedback. The research questions are as follows: 1)What are the similarities and differences between the two groups of students in providing feedback in the two tests? 2) How did the students' capacities to provide feedback change in the post-test compared to the pretest?

3. Research Design

3.1. Research Context and Participants

This study was conducted in a public English writing class at a university in Beijing with 44 undergraduate English majors as participants. The participants were randomly divided into 2 groups, the teacher feedback group (N=20) and the peer feedback group (N=24).

There were 6 writing tasks in this writing class, and the process of each writing task was as follows: each student in the peer feedback group submitted the first draft to a peer-review system (Peerceptiv), gave feedback on three anonymous essays, and received three anonymous comments based on different dimensions of essays; students in the teacher feedback group submitted their first drafts to the teacher, who gave them feedback. Then, students of both groups revised their essays and submitted the revised drafts according to the received comments and then submit their final drafts. Besides, students attended a pre-test and a post-test at the middle and end of the semester respectively. In each test, the teacher distributed a sample essay written by one student from the previous session, which was selected from a certain task. Both groups need to provide comments on the same sample.

3.2. Data Collection and Analysis

The data collected included written feedback of the two groups in the pre- and post-tests.

We segmented each piece of collected feedback of the 2 tests, ensuring that each feedback segment referred to only one question. All of the coding work was finished by 4 researchers to ensure the consistency and reliability of all data. Then the segmented comments were coded according to feedback type (praise, summary, criticism). And all criticisms were further coded in terms of feedback focus (content feedback and form feedback) and feedback feature (with or without explanations and suggestions). Feedback focuses refer to whether the criticism focuses on the structural or the linguistic level of an essay. The reliabilities were all above 0.6 (Kappa feedback type = 0.87***, Kappa feedback feature = 0.67**, Kappa feedback focus = 0.83***).

4. Results and Discussion

4.1. The Similarities and Differences Between the Two Groups of Students in Providing Feedback in the Two tests

In order to explore the similarities and differences of the two groups, we counted the number of comments provided by them in the 2 tests on the three dimensions of feedback type, focus, and feature.

In the pre-test (see Table 1), the feedback types of the 2 groups were almost the same, with the criticisms provided higher than the other two types, which manifests the whole feedback literacy is sound (Wu & Schunn, 2023). In terms of feedback focus, teacher feedback group focused more on the content of the essay while peer feedback group paid more attention to the form. When it comes the feature, both groups tended to include suggestions in their feedback,

although the peer feedback group provided more suggestions. Besides, the amount of comments with explanations were relatively small.

Comments provided by the 2 groups in the post-test were also analyzed (see Table 2). The two groups remained consistent in their tendency to provide more criticism in terms of the type of feedback provided. Besides, both groups focused more on the content compared to the form of the essays, which is different from the finding of Fan and Xu (2020) and the percentage of comments that included explanations was essentially the same for both groups. Nevertheless, the peer feedback group provided a greater percentage of suggestions in criticism than did the teacher feedback group.

Table 1. Feedback provided in the pre-test.

Classification		Teacher feedback group		Peer feedback group	
		n / N	P	n / N	P
Type	Praise	58 / 303	19.14%	71 / 410	17.32%
	Summary	9 / 303	2.97%	8 / 410	1.95%
	Criticism	236 / 303	77.89%	331 / 410	80.73%
Focus	Form	92 / 236	38.98%	201 / 331	60.73%
	Content	144 / 236	61.02%	130 / 331	39.27%
Feature	Explanation	36 / 236	15.25%	75 / 331	22.66%
	Suggestion	155 / 236	65.68%	273 / 331	82.48%

P: The proportion of comments for each dimension on its corresponding base.

Table 2. Feedback provided in the post-test.

Classification		Teacher feedback group		Peer feedback group	
		n / N	P	n / N	P
Type	Praise	28 / 265	10.57%	44 / 359	12.26%
	Summary	11 / 265	4.15%	15 / 359	4.18%
	Criticism	226 / 265	85.28%	300 / 359	83.57%
Focus	Form	89 / 226	39.38%	95 / 300	31.67%
	Content	137 / 226	60.62%	205 / 300	68.33%
Feature	Explanation	94 / 226	41.59%	111 / 300	37.00%
	Suggestion	147 / 226	65.04%	249 / 300	83.00%

4.2. Changes in Students' Capacities to Provide Feedback Between the Two Tests

After experiencing almost half of the semester, for the teacher feedback group, neither the feedback type nor focus has changed, with the only variation being in the feature of feedback provided. This group included more explanations in their comments to illustrate the issues in the essays. Similarly, the peer feedback group has not shown any changes in

feedback type but has undergone certain changes in two other dimensions. One is the feedback focus, shifting from initially focusing more on the form of essays to paying more attention to the content (with a larger proportion of content-related comments in the post-test). The other is the feedback feature, where more explanations were added when providing criticisms compared to the pre-test. Overall, both groups have shown improvements in their ability to provide feedback between the two tests, with the changes being a little more pronounced in the peer feedback group.

5. Conclusion

In summary, students' capacities to provide feedback tends to be similar after experiencing the two feedback modes. Both groups of students tended to provide criticisms throughout and demonstrated a focus on the content aspect of the essays. They consciously provided more suggestions and also incorporated explanations appropriately, which indicates that the feedback provided by both groups are specific.

A conclusion can be drawn that peer feedback can be employed as a complement to teacher feedback in L2 teaching, as students are able to provide good quality, detailed comments with peer feedback training. However, research methods such as interviews will be utilized in future studies to further analyze the reasons behind.

Acknowledgements

This research is supported by the 2023 Beijing University of Posts and Telecommunications Undergraduate Innovation and Entrepreneurship Training Programs named "A Study of the Effects of Teacher Feedback and Peer Feedback on Students' L2 Writing Proficiency" (Project No.: 202309013) and "A Study of Feedback Literacy Assessment Framework for College EFL Learners" (Project No.: 202309008).

References

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325.

Cheng, X., Liu, Y., & Wang, C. (2023). Understanding student engagement with teacher and peer feedback in L2 writing. *System, 119*, 103176.

Chong, S. W. (2020). Reconsidering student feedback literacy from an ecological perspective. *Assessment & Evaluation in Higher Education, 46*(1), 92–104.

Dong, Z., Ying, G., & Schunn, C. D. (2023c). Assessing students' peer feedback literacy in writing: scale development and validation. *Assessment & Evaluation in Higher Education, 48*(8), 1103–1118.

Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing, 50*, 100775.

Han, Y., & Xu, Y. (2019). The development of student feedback literacy: the influences of teacher feedback on peer feedback. *Assessment & Evaluation in Higher Education, 45*(5), 680–696.

Wood, J. M. (2021). Making peer feedback work: the contribution of technology-mediated dialogic peer feedback to feedback uptake and literacy. *Assessment & Evaluation in Higher Education, 47*(3), 327–346.

Wu, Y., & Schunn, C. D. (2023). Passive, active, and constructive engagement with peer feedback: A revised model of learning from peer feedback. *Contemporary Educational Psychology, 73*, 102160.

Ying, G., An, Q., & Schunn, C. D. (2023). The bilateral benefits of providing and receiving peer feedback in academic writing across varying L2 proficiency. *Studies in Educational Evaluation, 77*, 101252.

Zong, Z., Schunn, C. D., & Wang, Y. (2023). When do students provide more peer feedback? The roles of performance and prior feedback experiences. *Instructional Science, 51*(6), 977–1003.

Zhang, F., Min, H., He, P., Chen, S., & Ren, S. (2023b). Understanding EFL students' feedback literacy development in academic writing: A longitudinal case study. *Assessing Writing, 58*, 100700.

個人化音樂指揮機器人學習系統之開發

Development of a Personalized Music Conducting Robot Learning System

黃國豪[1*]，劉晉均[2]，陳宥瑋[2]，黃暄涵[3]
[1]國立雲林科技大學智慧數據科學研究所
[2]國立雲林科技大學資訊管理系人工智慧技優專班
[3]國立臺南藝術大學中國音樂學系
[*]ghhwang0424@gmail.com

【摘要】 音樂指揮最主要的作用是詮釋作曲者創作的樂曲，藉由整合及統一眾多演奏者的演奏，在觀眾面前完整的展示。然而，在學校教育資源有限的情況下，音樂指揮教育的教材與方式相對缺乏創新與互動性。本研究旨在開發一套「個人化音樂指揮機器人學習系統」，藉由國小學童喜歡機器人的特性，除了能讓學習者可以在沒有指揮教師的情況下以機器人引導學習者進行個人化的音樂指揮動作的練習，透過機器人的攝影鏡頭辨識學習者的動作，即時評估其正確性並給予回饋，以增進學習動機並提高學習成效。
【關鍵詞】 個人化；音樂；指揮；機器人；學習系統

Abstract: *The main function of music conducting is to interpret the music composed by the composer and to present it to the audience by integrating and unifying the performances of many players. However, with the limited educational resources in schools, the teaching materials and methods of music conducting education lack innovation and interactivity. This study aims to develop a "Personalized Music Conducting Robot Learning System", which takes advantage of the fact that children in elementary schools like robots. In addition to allowing learners to practice personalized music conducting without a conductor, this system also allows learners to recognize their movements through the robot's camera, and to evaluate their accuracy and give feedback in order to increase motivation and improve learning effectiveness.*

Keywords: personalization, music, conducting, robot, learning system

1. 前言

音樂指揮最主要的作用是詮釋作曲者創作的樂曲，藉由整合及統一眾多演奏者的演奏，在觀眾面前完整的展示（Schuller, 1998）。指揮教育在學童的音樂學習中扮演著至關重要的角色。指揮不僅是音樂表達的一種形式，讓學童學習指揮更是培養學生音樂感、節奏感和團隊合作精神的重要途徑。指揮教育能有助於學生對音樂作品的深入理解，也讓他們的音樂表達和創造力有進一步的發展。Sun（2022）指出，學習音樂會對學生心理健康和學業表現有好處，強調自信和自尊在學習過程中會逐漸提升。研究顯示，音樂課程能讓學生心情更好，學習成績也會提升。因此，讓學童認識基本的指揮姿勢與原理對學習音樂是很有幫助的。

傳統音樂教學往往以教師為中心，學生被動接受指導。Corbalán、Pérez-Echeverría、Pozo 與 Casas-Mas（2019）提到，傳統的指揮教學主要是讓學生模仿老師的動作，這樣的教學模式可能無法充分激發學生的學習興趣，特別是對於音樂指揮這樣需要豐富情感表達和動作技巧的學科。在上課時，有教師可以示範及立即回饋指正學習者的錯誤姿勢，但是學習者回家練

習時雖然可以利用觀看教學影片複習教師的示範，但自己的指揮動作是否正確就無法得到立即的回饋及修正。

近年來，科技的迅速發展帶動了教育領域的創新，而機器人技術的進步為教育帶來嶄新的可能性。沈秋宏（2021）指出未來教育時期，資訊科技是教育現代化的基礎，也是教育改革的重要途徑。隨著教育機器人技術的日益成熟，教育也面臨挑戰。特別是在音樂教育領域，指揮是學習樂器和音樂理論的關鍵元素之一。杜倩慧（2017）指出教師們普遍認為，傳統的音樂教學模式往往需要專業的音樂老師，而且學習的效果常受到學生動機和參與度的影響。Colwell 和 Hewitt（2010）指出，傳統教學模式已經難以讓現代的小孩子對指揮感興趣，而音樂教師也很難對所有的學生一對一指導其指揮的動作。Salgian、Agina 與 Nakra（2014）則指出，機器人可以不知疲倦地指揮同一樂曲，讓學生受益。

綜合上述，本研究開發了一套「個人化音樂指揮機器人學習系統」，藉由國小學童喜歡機器人的特性，除了能讓學習者可以在沒有指揮教師的情況下以機器人引導學習者進行指揮動作的練習，還能透過機器人的攝影鏡頭辨識學習者的動作，即時評估其正確性並給予回饋，以增進學習動機並提高學習成效。

2. 文獻回顧

2.1. 音樂指揮教學之類型及問題

在過去 Corbalán、Pérez-Echeverría、Pozo 與 Casas-Mas（2019）曾提到一些指揮的風格與學習理念，他們將其分為三種主要類型：

(1)傳統型（Traditional Profile, P1A）: 這類型的指揮主要依靠直接教學，主要是讓學生模仿老師的做法。他們較少使用複雜的教學過程，且不太注重樂譜處理。這種方法不太講究創新或變通，就是按部就班地教。

(2)專注閱讀型（Focus on Reading Profile, P1B）: 這類型的指揮著重於樂譜，較少提到身體的表現，但會幫助學生發展一些思考和理解音樂的能力。在一些複雜的過程中相較傳統型指揮優秀。

(3)專注學習與表現型（Focus on Learning and Representation Profile, P2）: 這類型的指揮使用較多複雜的過程，鼓勵學生自己去探索和理解。這種方法已經不只是教學生唱歌，還要教他們怎麼自己學習和表達。

綜合上述，本研究整理相關文獻之指揮教學類型如表 1 所示。傳統教學方法雖然有真實的情境，但有時間與地點之限制，也無法有效的提升學生學習的興趣。

表 1、指揮教學類型比較表

教學類型	優點	缺點
傳統型	提供清晰的指導和技巧	缺乏創造性
專注閱讀型	培養學生的批判性思考和自我評估能力，也能幫助學生更好地理解和解釋樂譜	缺乏實際指揮經驗
專注學習與表現型	鼓勵學生自主探索和表達，促進創造性思維	需要更多時間和資源

雖然現在可以非常容易的在網路上找到跟指揮相關的教學影片或是文章，也會在日常生活中看到一些簡單的指揮，像是學校樂團指揮或是社團的合唱團指揮等，但是音樂指揮不是

短時間接觸就可以達到期望的成效。因此，若能利用新興科技（如:姿態辨識）讓學生在課堂或課餘能主動練習，將可能提升學生的學習動機與成效。

2.2. 教育機器人技術之發展

隨著科技及人工智慧迅速的發展，機器人能夠獨立執行的任務也越來越複雜，並且被使用於許多領域，其中也包含應用於教育。教育機器人顛覆過去看著螢幕的數位教學方式，改用智能機器人的各種動作、表情和語音指令，充分發揮聲、光、色、及時境動態的教學效果，以達到互動的教學目的。李鴻宗（2016）提到體感遊戲如 Wii、Kinect 等開始盛行後，其可透過無線的遊戲裝置配合肢體動作如揮手、跑、跳、爬等，讓人於遊戲中有身歷其境的感覺，因具有結合視覺、動覺及聽覺之優勢，愈來愈多教育研究者將體感遊戲應用於教學中，透過遊戲式學習的方式，來引起學生主動學習的興趣。然而應用的對象年齡層大多為國小、高中以及成人。教育機器人已成為當今人機互動的重要工具，吳彥彰（2012）指出，在現今社會中，人們有更多的機會接觸到機器人的功能，像是生活中常見的智慧型服務機器人以及救援機器人。因此人機互動在當今成為一個重要的研究課題。

在過去的研究中發現，機器人教師能快速根據學生的學習情況調整適性化需求，使學生能得到個人化的學習指導，此外，與人類教師相比，機器人教師更能吸引學生的注意力，提高學習參與度以及學習成效，Wang、Sang、Huang、Li 和 Guo（2023）指出使用教育機器人的教學方法比傳統的講課式教學更能促進學生的學習成果，並提到機器人能創造有趣的學習活動和實踐經驗，從而激發學生的好奇心和享受學習的環境。亦有學者指出，在幼兒時期，利用機器人輔助學習有益於兒童早期識字和語言學習（Crompton, Gregory, & Burke, 2018; Neumann, 2020）。Alves-Oliveira（2017）的研究納入 52 名兒童，探討兒童對於機器人角色的看法，研究中發現雖然已與兒童說明機器人的角色為教師，但是兒童較傾向認為機器人像朋友，經過兩個月互動，兒童認為機器人的角色像同學。針對台灣學前幼兒第一次接觸智能機器人的反應之觀察研究中提出，幼兒傾向於以類似於人類的方式與智能機器人進行互動，並且智能機器人能吸引幼兒的注意力，引發學習的興趣。

目前市場上有著許多教育機器人產品，例如：沛博科技 Perobot 的 Pepper 機器人、Toyota（2016）的 Kirobo Mini 機器人、華碩電腦股份有限公司的 Zenbo 機器人及 Nuwa (n.d.)創造的 Kebbi 機器人等。其中沛博科技 Perobot 的 Pepper 機器人擁有人性化的外觀和情感識別能力，可以更容易地吸引學生的注意力並建立情感聯繫。它可以提供定制的教育內容，例如語言學習、科學實驗等，通過互動和遊戲方式讓學習變得更有趣（Perobot, n.d.）。Toyota 公司的 Kirobo Mini 機器人則可在語言學習方面發揮作用，它小巧可愛，適合與孩童互動。它可以幫助孩子練習發音、聽力和對話技能，同時提供有趣的故事和遊戲，以促進學習興趣的培養（Toyota, 2016）。華碩公司的 Zenbo 則具有多功能性，可以擔任教育助手的角色，提供筆記、語音搜索、教育應用程式等功能，幫助學生更有效地學習和組織資訊。它還可以與家庭娛樂和家庭自動化系統整合，提供更全面的家庭學習體驗（華碩電腦股份有限公司, n.d.）。劉紘誠（2021）指出 Zenbo 智慧機器人進行國小 AI 單元課程教學確實可以達到與一般教師相同的學習成效，再進一步分析得知男生對於不同的教學方法學習成績有顯著差異，導入 Zenbo 進行教學男生可以確實提升其學習成績。但是 Zenbo 機器人沒有手，並不適合成為指揮教學的機器人。而 Nuwa 創造的 Kebbi 機器人系列，是目前市場上較為知名的學習型陪伴機器人產品之一。它結合了 AI 技術，能夠進行圖像和語音辨識，並且通過互動遊戲和教育軟體提供學習支持，如 STEAM 教育、英語學習等，是一款多用途的陪伴和教育機器人。近年 Kebbi 已逐漸被應用於許多研究，教育環境中，像是林柏恩（2023）曾使用 Kebbi 機器人來幫助唐氏症兒童在溝通表達能力上的學習，並使用跨受試者多基準線設計的個案研究方法來探討其成效。Kebbi 可以發揮科技所帶來的高互動性與聲光效果，並且有豐富表情及可愛聲音跟孩童進行近距離

的互動，也可開發充滿趣味及遊戲的教學程式，激發小朋友的學習動機，以達到快速融合共學環境的目的。林柏恩（2023）指出認知障礙兒童對於傳統的教育方法無法保持長時間的專注度，但對於 Kebbi 機器人則表現出高度的興趣。

Kebbi 機器人在語音互動和肢體語言識別方面的能力，為學習和教學提供了豐富的互動體驗。這使得它在音樂教育，特別是指揮家辨識系統的開發上，展現出巨大的潛力。因此，本研究將以 Kebbi 機器人為基礎，開發一套「個人化音樂指揮機器人學習系統」，期待其能提高國小高年級學童學習音樂指揮的學習動機與學習成效。

3. 系統設計與開發

3.1. 音樂指揮教材內容

本研究將採用 Kebbi 機器人為基礎，開發一套「指揮機器人學習系統」，自行設計軟體，進行 ConductIT（2020）中提到最基本的 5 個指揮學習單元：

(1) 一拍子：這種指揮模式通常用於非常快的節奏，如 1/4。這種指揮風格主要是一系列下拍伴隨單一反彈。
(2) 二拍子：主要包括下拍和上拍，重點放在第一拍的強烈重音，第二拍較輕。
(3) 三拍子：常用於 3/4。三拍子的指揮風格通常形成一個三角形，包括下拍、右移和上拍。
(4) 四拍子：常見於 4/4 拍子。這種指揮模式涵蓋下拍、左移、右移和上拍。
(5) 六拍子：在 6/8 拍子中，指揮通常分為兩個大拍，每個大拍含三個較小的節奏單位。這種指揮模式強調每個大拍的開始，特別是第一和第四小節。通常以下拍開始第一大拍，接著進行寬廣的圓弧動作來指揮接下來的兩個小節。第二大拍從第四小節開始，也是以下拍為起點，然後用相似的圓弧動作來完成第五和第六小節。

為了讓國小學童能在歡樂的氣氛中學習指揮，本研究特別挑選了受兒童喜愛且熟悉的卡通歌曲為練習的教材內容，詳細資料整理如表 2 所示。

表 2、音樂指揮教材內容

單元	曲名	長度
一拍子	崖上的波妞主題曲	1 分 22 秒
二拍子	哆啦 A 夢主題曲	0 分 52 秒
三拍子	神隱少女主題曲	1 分 53 秒
四拍子	名偵探柯南主題曲	1 分 20 秒
六拍子	We Wish You a Merry Christmas	2 分 00 秒

3.2. 開發環境及相關技術

本研究採用 Android Studio 為開發工具，使用 Java 程式語言，如圖 1 左所示，開發環境包含 Kebbi 模擬器，如圖 1 右所示。

人體偵測利用 OpenPose 進行處理，系統通過攝像頭捕捉學習者的指揮手勢，並利用內建鏡頭從中提取出影像，以便更精確地追踪和分析手部動作。

Kebbi 機器人的肢體及臉部表情控制功能，透過內建圖形化程式編輯工具，來定義機器人的動作和行為以及表情控制，創建動態的表情變化，以達到模仿指揮時的面部表情。

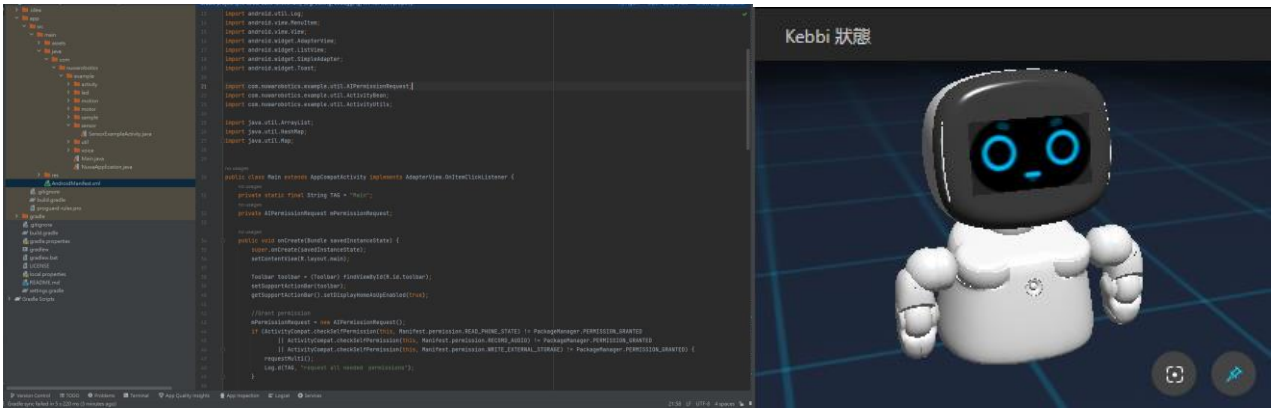


圖 1、Android Studio 與 Kebbi 模擬器

3.3. 示範指揮功能

本研究透過 Kebbi 機器人的肢體及臉部表情控制功能，能摸擬一位真正的指揮時的表情和動作示範 5 首不同節奏之樂曲。可以讓學習者自由指定要觀看示範的樂曲。透過內建圖形化程式編輯工具，來定義機器人的動作和行為以及表情控制，創建動態的表情變化，以達到模仿指揮時的面部表情。

3.4. 個人化即時回饋功能

另外本研究透過 Kebbi 機器人的攝像頭和處理能力來分析和回應學習者的指揮動作，在這個系統中，人體偵測利用 OpenPose 進行處理。系統通過攝像頭捕捉學習者的指揮手勢，並利用內建鏡頭從中提取出影像，以便更精確地追踪和分析手部動作。並利用所提到的手勢追踪和分割技術，將連續的指揮手勢劃分為個別的動作，以便更準確地分析和回應。同時，系統還會利用動態時間扭曲（DTW）分類器來比對學習者的指揮動作與預先記錄的標準手勢模板，從而判斷其準確度和節奏。透過這種方法，系統能夠有效地檢測學習者的手勢，從而提供即時的錯誤反饋和指導。也會分析學習者目前的學習狀況診斷學習者的指揮弱點，主動推薦學習者應該要再觀看的示範樂曲。

本系統的回饋內容將依據學生指揮的動作及拍子，進行不同的回饋及分數，如表 3 所示。例如學生指揮的拍子早於或晚於指定時間至指定位置，表示學生未在正確的時間點上執行正確動作，系通將顯示出錯誤的時間點為哪一個小節，使學生了解自己在哪一小節的拍子出現錯誤或是不穩定，進而改正及反覆練習；若是指揮的動作不正確，如手肘過低或過高、手腕過低或過高及指揮拍子的位置方位錯誤等，系統將顯示其標準姿勢與學生的姿勢差異，並給予針對該次練習需要改進的講解及回饋。

另外，本系統會記錄每一位學習者的學習歷程，並整理 5 個單元學習者所得的分數做比較，使學習者可以在練習一段時間後，有效了解自己的強項及弱項，並且針對各個單元給予的意見分析及回饋，進行調整，以利於提升學習者的學習成效。

表 3、節奏及動作判斷標準及回饋分析

	判斷標準	扣分	回饋或動作分析
節奏時間錯誤	早於指定時間至指定位置 晚於指定時間至指定位置	-2	顯示錯誤之小節並給予正確時間標示
動作位置錯誤	手肘過低或過高	-1	顯示正確手肘之位置並給予意見回饋 如:手肘位置過低需提高手肘位置
	手腕過低或過高	-1	顯示正確手腕之位置並給予意見回饋 如:手腕位置過低需提高手腕位置
	指揮拍子的位置及方位錯誤	-3	顯示正確之位置方位並給予意見回饋 如:四拍中第三拍位置應於由左至右

3.5. 操作流程與介面

本系統採用語音互動的方式讓學童進行自主學習，操作流程圖如圖 2 所示。首先學習者必須進行登入的動作，讓機器人知道現在的學習是誰。接著機器人會分析出這位學習者目前的學習進度及狀況，提供學習者適當的學習建議，如建議新學習者從第 1 單元開始；舊學習者則會依據已經學習過的記錄，判斷目前需要加強的單元，建議學習者選擇。學習者選擇好要學習的單元後，系統會進行遊戲規則的說明，但學習者如果已經清楚規則，是可以略過這一步的。接下來機器人會跟著音樂進行該單元的指揮示範，然後要求學習者跟著音樂做一次。接著機器人會分析學習者的動作並提出建設性的回饋修正意見，最後會顯示分數及排名。新的學習資料滙整至系統後，再回到前面的個人化對話互動，重新提出接下來的學習建議。

每位學童練習的紀錄都會存放在 Kebbi 機器人內建資料夾內，當每位學童在機器人上登入後即可從資料夾內讀取學童歷史紀錄，學童可從系統上得知較弱的單元進行加強學習，針對學生的弱點進行有針對性的指導和練習，以提高其學習效果。

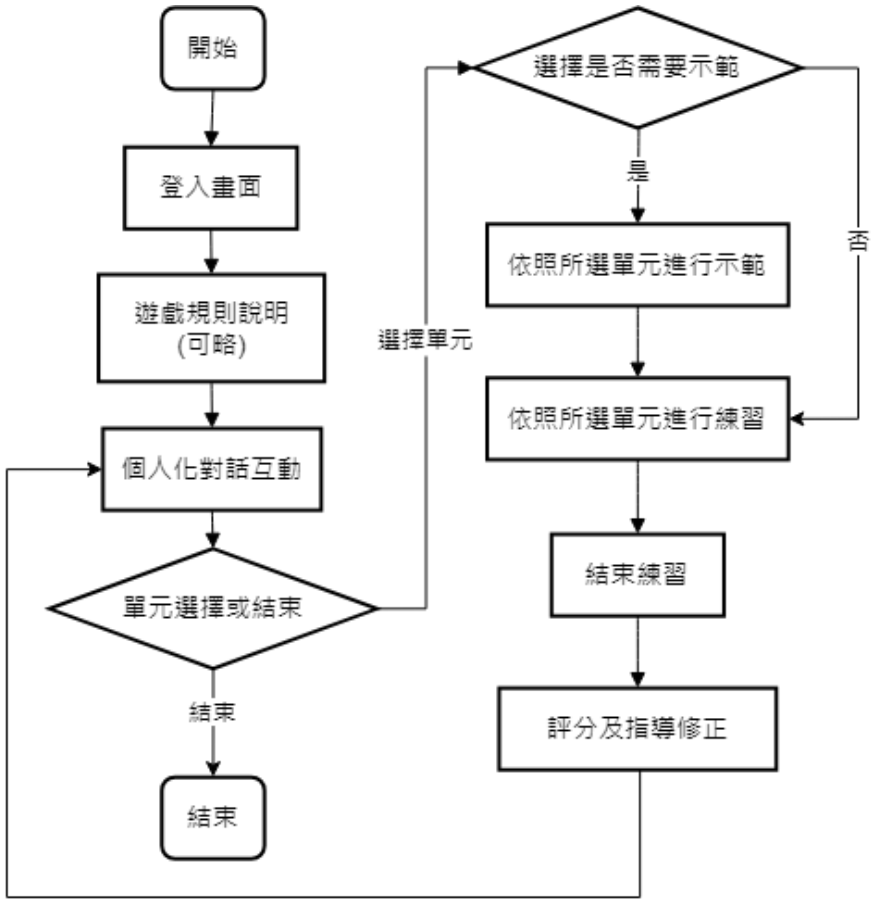


圖 2、系統操作流程圖

系統的互動過程及介面如圖 3 所示。從登入畫面登入後機器人會串接資料夾紀錄每位學童的學習狀況，依照學童的學習進度及成效進行個人化的對話互動(左上)，若為首次使用的學童則會建議從第一單元開始，若非首次使用的學童，會建議選擇較弱的單元進行學習，接著讓學生選擇五種單元中的其中一種單元（右上）。為了讓學生更快了解遊戲規則及增加系統的可用性，在選擇單元後，Kebbi 機器人將說明遊戲規則（中左及中右）。在遊戲進行中，使用者先跟著音樂來觀看機器人手部動作來學習指揮家動作。機器人教學結束後，畫面將進入倒數 10 秒，倒數結束後，畫面顯示該歌曲的樂譜並撥放歌曲，輪到使用者來做學習成效檢

測，機器人監看使用者動作是否正確，在檢測途中 Kebbi 機器人會給予提示及回饋（左下），結束後，將針對使用者的動作做出評分，並針對錯誤的手勢提出教學指導修正提高學習成效，學習結束後給予鼓勵提高學童學習意願（右下）。



圖 3、系統與學習者的互動過程與介面設計

4. 結論與未來展望

本研究開發了一套「個人化音樂指揮機器人學習系統」，藉由國小學童喜歡機器人的特性，除了能讓學習者可以在沒有指揮教師的情況下以機器人引導學習者進行個人化的音樂指揮動作的練習，透過機器人的攝影鏡頭辨識學習者的動作，即時評估其正確性並給予回饋，以增進學習動機並提高學習成效。

未來本研究將先進行系統的可用性評估，以確定系統的成熟度。接著將再進行正式的教學實驗，以確定系統的有效性。期望對國小學童的指揮教育能有所貢獻。

誌謝

本研究感謝國科會經費之補助，計畫編號：MOST 111-2410-H-24-004-MY2。

參考文獻

吳彥彰（2012）。**結合多感測器的手勢辨識於人機互動及智慧型服務機器人之應用**（碩士論文）。取自華藝線上圖書館系統。(系統編號 U0001-2707201210011100)

李鴻宗（2016）。**遊戲式學習結合體感操作之認知負荷評估-以全球暖化的科學迷思為例**（碩士論文）。取自華藝線上圖書館系統。（系統編號 U0015-2908201611272800）

杜倩慧（2017）。**高中職音樂教師教學信念、困境及其因應策略**（未出版博士論文）。東吳大學，台北市。

沈秋宏（2021）。教育機器人在未來教育中的應用與挑戰。**學校行政，136**，137-163.

林柏恩（2023）。**以智慧型機器人輔助教學認知障礙兒童學習成效評估**（碩士論文）。取自華藝線上圖書館系統。（系統編號 U0017-2211202315350967）doi:10.6840/cycu202301619

華碩電腦股份有限公司（n.d.）。Zenbo. Retrieved from https://zenbo.asus.com/tw/

劉紘誠（2021）。**Zenbo 智慧機器人導入國小 AI 課程教學之研究**（碩士論文）。取自華藝線上圖書館系統。（系統編號 U0002-1007202108485600）doi:10.6846/TKU.2021.00226

Colwell, R., & Hewitt, M. (2010). The teaching of instrumental music (4th ed.). *Routledge.* https://doi.org/10.4324/9781315665016

ConductIT. (2020). Basic beat patterns. Retrieved from https://conductit.eu/study-room/technique/technique-1/basic-beat-patterns/

Corbalán, M., Pérez-Echeverría, M. P., Pozo, J.-I., & Casas-Mas, A. (2019). Choral conductors to stage! What kind of learning do they claim to promote during choir rehearsal? *International Journal of Music Education*, *37*(1), 91-106. https://doi.org/10.1177/0255761418800515

Crompton, H., Gregory, K., & Burke, D. (2018). Humanoid robots supporting children's learning in an early childhood setting. *British Journal of Educational Technology*, *49*(5), 911-927.

Neumann, M. M. (2020). Social robots and young children's early language and literacy learning. *Early Childhood Education Journal*, *48*(2), 157-170.

Nuwa. (n.d.). Retrieved from https://www.nuwarobotics.com/zh-hant/product/

Perobot. (n.d.). Retrieved from http://www.perobot.com.tw/

Salgian, A., Agina, L., & Nakra, T. M. (2014). Teaching robots to conduct: Automatic extraction of conducting information from sheet music. In *Proceedings of the International Computer Music Conference (ICMC).*

Schuller, G. (1998). *The compleat conductor*. Oxford University Press.

Sun, J. (2022). Exploring the impact of music education on the psychological and academic outcomes of students: mediating role of self-efficacy and self-esteem. *Frontiers in Psychology, 13,* Article 841204. https://doi.org/10.3389/fpsyg.2022.841204

Toyota (2016). Toyota launches 'Kirobo Mini' companion robot. Retrieved from https://startuplatte.com/2016/11/22/toyota_kirobo_mini_companion_robot/

Wang, K., Sang, G.-Y., Huang, L.-Z., Li, S.-H., & Guo, J.-W. (2023). The effectiveness of educational robots in improving learning outcomes: A meta-analysis. *Sustainability*, *15*(5), 4637. https://doi.org/10.3390/su15054637

先備知識對遊戲化程式設計翻轉教室之影響

Impact of Prior Knowledge on Flipped Classrooms of Gamified Programming Courses

黃國豪
國立雲林科技大學智慧數據科學研究所
ghhwang0424@gmail.com

【摘要】本研究將 Kahoot 遊戲與闖關實作解題遊戲帶入程式設計課程的翻轉教室中，並進一步探討先備知識對此種教學方式的影響。研究對象為臺灣中部某科技大學資訊管理系的 28 位技優專班大一學生。在經歷一學期的教學實驗後，高、低先備知識的學習者在內、外在動機、自我效能、Kahoot 遊戲滿意度及課堂闖關實作遊戲滿意度均沒有顯著的差異。但是高先備知識者在闖關遊戲實作的成績表現顯著優於低先備知識者。而低先備知識者在觀看影片的成績顯著優於高先備知識者，也就是低先備知識者在這種教學模式下比高先備知識者更願意花時間觀看教學影片和回答其中的問題。

【關鍵詞】 先備知識；遊戲化；程式設計課程；翻轉教室

Abstract: *This study intergrated Kahoot and level-breaking practical games into programming flipped classroom. This study also explored the impact of prior knowledge on this teaching method. The research subjects were 28 first-year students in the Department of Information Management of a university of science and technology in Central Taiwan. After a semester of teaching experiments, there were no significant differences in intrinsic and extrinsic motivation, self-efficacy, Kahoot game satisfaction, and level-breaking practical game satisfaction between learners with high and low prior knowledge. However, those with high prior knowledge performed significantly better than those with low prior knowledge in the scores of the level-breaking game. The performance of those with low prior knowledge in watching videos is significantly better than those with high prior knowledge.*

Keywords: prior knowledge, gamification, programming course, flipped classroom

1. 前言

許多學者均認為學習程式設計很重要，但是要學好程式設計並不容易，因為初學者往往需要磨練多年才有機會成為程式設計的專家（何昱穎、張智凱、劉寶鈞，2010；黃國豪、陳碧茵、曾薇方、黃筱惠，2013；Winslow，1996）。Husain、Tarannum 和 Patil（2013）指出，如果程式設計課程的教學時數不足，常造成學生只能花時間在理解程式基礎語法，而缺乏實際操作和練習的機會。Dalbey 和 Linn（1985）研究顯示，教材會直接影響學生學習程式設計的成效。使用適合學生程度的程式設計教材，可以提高學生對程式設計的理解與學習動機。學生喜歡在實例中學習，而不是在長篇大論的語法中學習。因此，面對程式設計的初學者，除了基本語法的描述外，更多淺顯易懂的應用範例及更多的實作機會是必要的。為了節省上課時教師講解的時間，程式設計課程採用翻轉教室的模式是不錯的選擇。

翻轉教室的概念是由 Bergmann 與 Sams（2012）所提出來，上課方式為讓學生利用課前的時間先自主學習教師事先錄製的教學影片，再利用上課時間讓學生討論、提問或進行實作應用練習，這樣的上課方式除了更有彈性，也可以增加師生及同儕互動的機會。也就是說，翻轉教室的優點是讓學生可以依照自己的時間、地點和進度來學習（Bergmann & Sams,

2012）。有程式設計基礎的同學可以跳過已經會的教學影片，而完全不會或學習有困難的同學則可以多看幾次不懂的部份。因此，翻轉教室的教學模式可以改善先備知識差異太大所造成的教學進度問題（McLaughlin et al., 2013）。然而，如何讓學生在課前願意認真觀看教學影片及上課時願意認真參與實作解題的練習，導入適當的課堂活動是重要的，其中遊戲化學習就是一個不錯的方式。

遊戲化學習（gamified learning）是採用遊戲化的方式來進行學習（Roussou, 2004; Smith & Pellegrini, 2008）。它是目前相當流行的教學方式。Kiesler、Kraut、Koedinger、Aleven 與 Mclaren（2011）指出當今學校面臨著學生學習動機低落的問題，而遊戲化學習就提供了一個改善的機會，它有機會可以幫助學校解決這些難題。Dicheva、Dichev、Agre 與 Angelova（2015）則對遊戲化學習進行了文獻回顧整理，他們認為有很多關於在教育中使用遊戲化的研究都僅介紹了一些遊戲機制，並對這些遊戲機制的有效性進行了實證研究。但在真實學習環境中使用，長期進行遊戲化學習的實證研究仍不多。也就是大部份的研究都沒有進行長期的評估，這使得對這些研究的結果進行分析困難，也難以推測其成功或負面結果的一般原因。雖然大多數研究者都認為，遊戲化如果設計和使用得當，有潛力改善學習，但他們認為遊戲化學習還是缺乏大量正式課程的驗證。González、Torres 與 Ramon（2023）指出，與傳統教學方法相比，實施翻轉教室和遊戲化策略可以積極影響學生的參與、動機和學業成績，從而提高教學品質和滿意度，並顯著提高學業成績。因此，本研究於每週的課堂中會導入 Kahoot 的遊戲搶答機制，讓學生在歡樂的氣氛中複習教學影片的內容，接著再以闖關實作遊戲讓學生加強程式設計的實際解題能力。採用 Kahoot 的遊戲搶答機制是因為在 Zhang 與 Yu（2021）分的研究中指出，無論是在傳統課堂或是翻轉教室的學習環境，只要搭配良好的課程規劃，Kahoot 的遊戲搶答機制可以改善學生的學習成效，加強學生和教師之間的課堂互動，以及學生之間的課外合作。而採用闖關實作遊戲在課堂中，則是希望藉由利用遊戲化的課堂實際要求學生多做程式的實作練習可以讓學生在做中學，慢慢建構自己的程式能力。

此外，有許多研究（陳明溥，2007；Aman, 1992；Clark, 1992；Carlson & Wright, 1993；Mayer & Sims, 1994）認為，程式設計的教學必須要考量學生先備知識的差異，提供適合的教學法才能達到較好的教學成效。因此，為了確認先備知識對遊戲化翻轉教學程式設計課程的影響，本研究除了將 Kahoot 的遊戲搶答機制與闖關實作解題遊戲帶入程式設計課程的翻轉教室中，並將進一步探討先備知識對此種教學方式的影響。本研究之研究問題具體陳述如下：

(1) 在遊戲化程式設計翻轉教室中，先備知識是否對學生的內在動機、外在動機及自我效能造成顯著的影響？
(2) 在遊戲化程式設計翻轉教室中，先備知識是否對學生的 Kahoot 遊戲搶答機制滿意度、闖關實作遊戲滿意度、課堂出席率造成顯著的影響？
(3) 在遊戲化程式設計翻轉教室中，先備知識是否對學生的闖關實作遊戲成績、觀看教學影片成績、期末考實作成績及學期總成績造成顯著的影響？

2. 研究方法

2.1. 研究架構

本研究探討先備知識對程式設計課程融入 Kahoot 遊戲搶答機制及闖關實作遊戲之影響。自變項為先備知識，包括高先備知識及低先備知識；依變項為內在動機、外在動機、自我效能、Kahoot 遊戲搶答機制滿意度、闖關實作遊戲滿意度、課堂出席、闖關實作遊戲成績、觀看教學影片成績、期末考實作成績及學期總成績；控制變項為教師、教學內容、實驗時間。其研究架構如圖 1 所示。

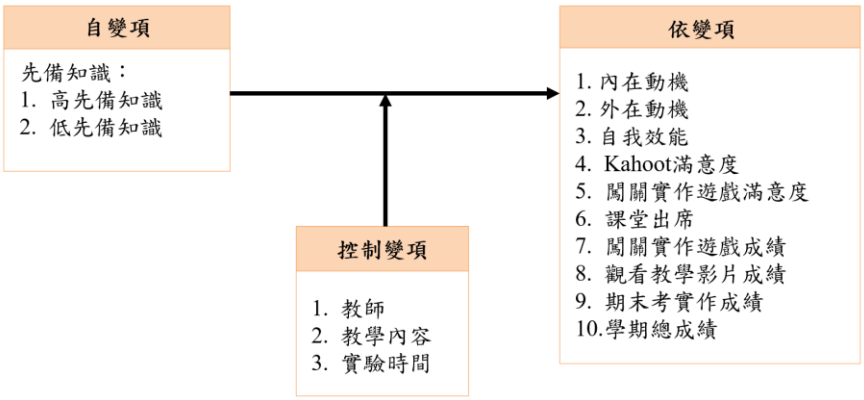


圖 1、研究架構圖

2.2. 研究對象

本研究之研究對象為臺灣中部某科技大學資訊管理系的 28 位技優專班大一學生，其中高先備知識學習者 12 人，低先備知識學習者 16 人；男生 20 人，女生 8 人。他們有些是具有程式設計基礎的學習者，但有些是第一次接觸程式設計課程。

2.3. 研究工具

本研究之研究工具除了包含 Kahoot 線上即時遊戲互動系統及闖關實作遊戲使用之 Dev C++整合型程式開發環境外，還包括課前觀看影片的 TronClass 數位學習管理平台及內在動機問卷、外在動機問卷、自我效能問卷、Kahoot 遊戲搶答機制滿意度問卷、課堂闖關實作遊戲滿意度問卷與前後測試卷。其中學習動機問卷是採用 Wang 與 Chen（2010）修改自 Pintrich、Smith、Garcia 與 McKeachie （1991）所提出的問卷，包括內在動機問卷 3 題及外在動機問卷 3 題，共 6 題，如表 1 所示。自我效能問卷為參考自 Pintrich、Smith、Garcia 與 McKeachie（1991）所提出之問卷，共有 8 題，如表 2 所示。Kahoot 遊戲搶答機制滿意度問卷為自編問卷，共 3 題，如表 3 所示。闖關實作遊戲滿意度問卷亦為自編問卷，共 3 題，如表 4 所示。所有問卷皆採用李克特 5 點量表（Likert, 1932），其中 1 表示「非常不同意」，5 表示「非常同意」。所有問卷皆經由兩位具數位學習研究資歷超過十年之學者進行編修，符合專家效度。前測試卷為 20 題 C++語法的基本觀念選擇題，包含變數、資料型態、運算式、條件控制及迴圈等。後測（期末考）為 10 題實作題，每題 2 分共 20 分。範圍還包含陣列、指標、函數、物件等。

表 1 學習動機問卷內容

構面	題號	問題
內在動機	1	在本課程中，我比較喜歡有挑戰性的教材，因為這樣我可以學到新的事物。
	2	在本課程中，我比較喜歡能引起我好奇心的教材，即使困難也無所謂。
	3	如果可以，我會選擇能學到東西的課程，即使分數不高也無所謂。
外在動機	4	在本課程中得到好成績，對我來說是最滿足的事情。
	5	如果可以，我希望能在本課程中得到比大多數學生好的成績。
	6	我希望在本課程中能有好的表現，因為在家人、朋友、老師或其他人面前展現我的能力是很重要的。

表 2 自我效能問卷內容

題號	問題
1	我相信我可以在本課程中得到優異的成績。
2	我確信我能理解本課程中最困難的部分。
3	我有自信能理解本課程所教授的基本觀念。

4	我有自信能理解本課程中老師所教最複雜的部分。
5	我有自信能在本課程的作業和測驗上表現優異。
6	我預期能學好本課程。
7	我確信能精通本課程所教授的技能。
8	考量本課程的難度、老師和我的能力，我覺得我可以學好本課程。

表 3 Kahoot 遊戲搶答機制滿意度問卷

題號	問題
1	我喜歡課堂 Kahoot 遊戲活動。
2	我覺得課堂 Kahoot 遊戲活動很有趣。
3	我覺得課堂 Kahoot 遊戲活動對我是有幫助的。

表 4 課堂闖關實作遊戲滿意度問卷

題號	問題
1	我喜歡課堂闖關實作遊戲活動。
2	我覺得課堂闖關實作遊戲活動很有趣 。
3	我覺得課堂闖關遊戲活動對我是有幫助的 。

2.4. 實驗流程

本研究之實驗流程如圖 2 所示。首先進行 50 分鐘的前測及前內在動機問卷、前外在動機問卷、前自我效能問卷的填寫，接著進行 14 週的遊戲化翻轉教學，每週三節課共 150 分鐘。其中第一節課的前 30 分鐘先進行 Kahoot 互動遊戲，以確認學生是否有認真觀看教學影片。接著進行 120 分鐘的闖關實作遊戲，以確認學生是否會活用程式語法，學生必須在時間內完成 6 題程式設計實作題，若是沒有完成，則在下一週將繼續原來的關卡，直到完成前一個關卡才能進入下一個關卡。隨著教學進度的進行，每一週會多開放一個關卡讓學生闖關。在學期末時進行 150 分鐘的期末考實作測驗。最後，學生須填寫後內在動機問卷、後外在動機問卷、後自我效能問卷、課堂 Kahoot 遊戲搶答機制滿意度問卷、課堂闖關實作遊戲滿意度問卷。

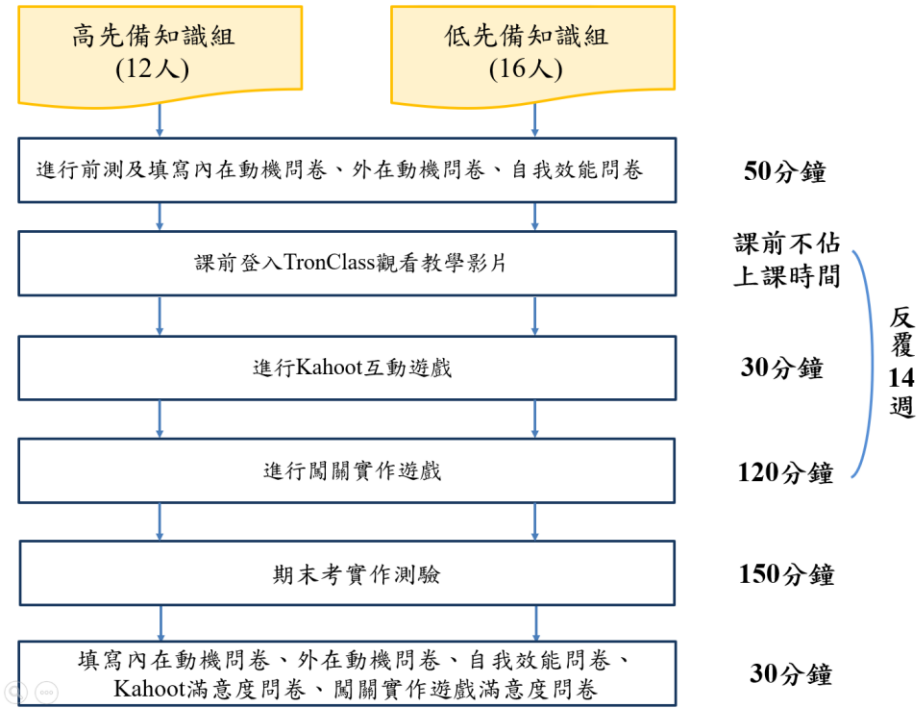


圖 2、實驗流程圖

2.5. 資料分析

本研究採用 SPSS 19 統計套裝軟體進行分析。先使用信度分析來確定問卷的可信度，接著利用成對樣本 *t* 檢定確認學習前與學習後學生的內在動機、外在動機與自我效能是否有顯著差異。接著將學生依照先備知識分為高先備知識組與低先備知識組，再利用獨立樣本 *t* 檢定比較兩組學生在期末的內在動機、外在動機、自我效能、Kahoot 滿意度、闖關實作遊戲滿意度、課堂出席、闖關實作遊戲成績、觀看教學影片成績、期末考實作成績以及學期總成績是否有顯著差異。

3. 結果與討論

3.1. 信度分析結果

為了確認本研究問卷調查是可信的，本研究使用 SPSS 統計套裝軟體進行信度分析。結果發現，除了前內在動機構面的 Cronbach's alpha 值為 0.610 外，其餘構面均大於 0.7（介於 0.704 至 0.923 之間）。此結果顯示，本研究問卷調查的信度大致上是可被接受的。詳細的信度分析結果如表 5 所示。

表 5、信度分析

構面	題數	Cronbach's alpha
前內在動機	3	.610
前外在動機	3	.789
前自我效能	8	.902
後內在動機	3	.704
後外在動機	3	.919
後自我效能	8	.923
Kahoot 遊戲滿意度	3	.915
闖關遊戲滿意度	3	.891

3.2. 成對樣本 t 檢定結果

為了確認學生在學習前與學習後的內在動機、外在動機與自我效能是否有顯著差異，本研究對 28 位學生的相關資料進行了成對樣本 *t* 檢定。結果發現，前後內在動機、外在動機與自我效能均無顯著差異（$p > .05$），此結果與 Cuervo-Cely、Restrepo-Calle 與 Ramírez-Echeverry (2022)的研究結果類似。雖然一般認為遊戲化學習相對於傳統教學較能提高學生的學習動機，但是對於程式設計這類較難的課程，並不一定能有顯著的提升，因為課程的難度會帶給學生一些挫折感，間接影響了學習動機，尤其是內在動機（黃國豪，2021）。詳細的成對樣本 *t* 檢定結果如表 6 所示。

表 6、成對樣本 t 檢定

構面	人數	前平均	前標準差	後平均	後標準差	*t*	*p*
內在動機	28	3.762	0.647	3.548	0.837	1.236	.227
外在動機	28	3.988	0.751	3.810	1.105	0.965	.343
自我效能	28	3.813	0.583	3.804	0.720	0.084	.934

3.3. 高、低先備知識之獨立樣本 t 檢定結果

為了確認高、低先備知識者對內在動機、外在動機、自我效能、Kahoot 滿意度、闖關遊戲滿意度、課堂出席成績、闖關遊戲的實作成績、觀看影片完成度、期末考實作成績以及學期總平均是否有顯著差異，本研究針對 12 位高先備知識者以及 16 位低先備知識者進行獨立樣本 t 檢定。結果顯示，除了闖關遊戲以及觀看影片的構面有顯著差異（$p < 0.05$），其餘構面則均無顯著差異（$p > 0.05$）。高先備知識者在闖關遊戲構面大於低先備知識者，表示高先備知識者對闖關遊戲實作表現較好，此發現與 Phungoen 等人（2020）的研究結果類似。而低先備知識者在觀看影片構面大於高先備知識者，表示低先備知識者較認真觀看影片，此發現與 Hong（2024）的研究結果類似。詳細的高、低先備知識之獨立樣本 t 檢定結果如表 7 所示。

表 7、高、低先備知識之獨立樣本 t 檢定

構面	先備知識	個數	平均數	標準差	*t*	*p*
內在動機	高	12	3.611	.583	0.342	.735
	低	16	3.500	1.004		
外在動機	高	12	3.750	1.147	-0.242	.810
	低	16	3.854	1.109		
自我效能	高	12	3.813	.847	0.056	.956
	低	16	3.797	.637		
Kahoot 遊戲滿意度	高	12	3.389	1.127	-0.836	.411
	低	16	3.688	.765		
闖關遊戲滿意度	高	12	3.222	1.095	-0.985	.334
	低	16	3.583	.848		
課堂出席	高	12	17.199	1.546	-0.912	.376
	低	16	17.639	.739		
闖關實作遊戲成績	高	12	39.841	.550	2.385*	.027
	低	16	38.958	1.338		
觀看教學影片成績	高	12	19.570	.620	-2.275*	.043
	低	16	19.980	.081		
期末考實作成績	高	12	20.000	.000	1.789	.094
	低	16	18.625	3.074		
學期總成績	高	12	96.610	1.963	1.039	.308
	低	16	95.203	4.357		

*$p < .05$

4. 結論與建議

4.1. 結論

本研究將 Kahoot 遊戲與闖關實作解題遊戲帶入程式設計課程的翻轉教室中，並進一步探討先備知識對此種教學方式的影響。研究對象為臺灣中部某科技大學資訊管理系的 28 位技優專班大一學生。在經歷一學期的教學實驗後，本研究發現，學生在學習前與學習後的內在動機、外在動機與自我效能並沒有顯著差異。高、低先備知識的學習者在內、外在動機、自我效能、Kahoot 遊戲滿意度及課堂闖關實作遊戲滿意度均沒有顯著的差異。但是高先備知識者在闖關遊戲實作的成績表現顯著優於低先備知識者。而低先備知識者在觀看影片的成績顯著

優於高先備知識者，也就是低先備知識者在這種教學模式下比高先備知識者更願意花時間觀看教學影片和回答其中的問題。其原因應該是高先備知識者對於程式設計的概念較為全面，而低先備知識者需一步一步學習程式設計。對於低先備知識者觀看影片是相當重要的，但對於高先備知識者而言，部份影片內容相對較為枯燥無聊。

4.2. 建議

本次研究的前內在動機是唯一信度分析低於 0.7 的構面，原因可能為題目敘述模糊，導致學生無法精確回答，未來進行類似研究時可以再修改題內容，以提高問卷的信度。本次的教學實驗雖然是一學期的長時間實驗，但是樣本數兩組一共只有 28 人，相對較少。建議未來可以增加總樣本數到 60 人以上，以增強結果的顯著性和可信度。另外，也可以考慮分析自我效能及性別對程式設計課程融入 Kahoot 及闖關實作遊戲之影響。

誌謝

本研究感謝教育部教學實踐研究計畫經費之補助，計畫編號：PEE1121126。

參考文獻

何昱穎、張智凱、劉寶鈞（2010）。程式設計課程之學習焦慮降低與學習動機維持－以 Scratch 為補救教學工具，**數位學習科技期刊**，**2**（1），11-32。

陳明溥（2007）。程式語言課程之教學模式與學習工具對初學者學習成效與學習態度之影響。**師大學報**，**52**，1-21。

黃國豪（2021）。遊戲化同儕教導於翻轉教室全校性程式設計課程之行動研究。2021 台灣教育傳播暨科技學會年會國際學術研討會 TAECT(2021)，國立臺北護理健康大學，台北市。

黃國豪、陳碧茵、曾薇方、黃筱惠（2013）。程式設計課程應用悅趣化遊戲競爭策略於學習動機與學習成效之影響，**數位學習科技期刊**，**5**（4），27-43。

Aman, J. (1992). Gender and attitude toward computers. In C. D. Martin & E. Murchie-Beyma (Eds.), *In search of gender free paradigms for computer education*, (pp. 33-46). Eugene, OR: International Society of Technology Education. https://www.amazon.com/Search-Gender-Paradigms-Computer-Education/dp/1564840093

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Washington, DC: International Society for Technology in Education.

Carlson, R. & Wright, D. (1993). Computer anxiety and communication apprehension: Relationship and introductory college course effects. *Journal of Educational Computing Research*, *9*(3), 329-338.

Clark, R. E. (1992). Facilitating domain-general problem solving: Computers, cognitive processes and instruction. In E. D. Corte, M. C. Lin, H. Mandl, and L. Verschaffel (Eds.), *Learning environment & problem solving* (pp. 265-285). New York: Springer-Verlag.

Cuervo-Cely, K. D., Restrepo-Calle, F., & Ramírez-Echeverry, J. J. (2022). Effect of gamification on the motivation of computer programming students. *Journal of Information Technology Education: Research*, *21*, 1-23.

Dalbey, J., & Linn, M. C. (1985). The demands and requirements of computer programming: A literature review. *Journal of Educational Computing Research*, *1*(3), 253-274.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society*, *18*(3), 75-88.

González, I. P., Torres, A. J., & Ramon, M. M. (2023). The effect of a combination of flipped classroom and gamification on university student's perceived teaching quality, subject satisfaction and academic performance. *Retos-Nuevas Tendencias En Educacion Fisica Deporte Y Recreacion*, (50), 403-407.

Hong, B. (2024). Who watched pre/post-lecture tutorial videos? Does flipped learning help beginners in economics? *International Review of Economics Education, 45*, 100283

Husain, M., Tarannum, N., & Patil, N. (2013, December). Teaching programming course elective: A new teaching and learning experience. *In 2013 IEEE International Conference in MOOC, Innovation and Technology in Education (MITE)* (pp. 275-279). IEEE.

Kiesler, S., Kraut, R. E., Koedinger, K. R., Aleven, V., & Mclaren, B. M. (2011). Gamification in education: What, how, why bother?. *Academic Exchange Quarterly*, *15*(2), 1-5.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 1-55.

Mayer, R. R. & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389-401.

McLaughlin, J. E., Griffin, L. M., Esserman, D. A., Davidson, C. A., Glatt, D. M., Roth, M. T.,…Mumper, R. J. (2013). Pharmacy student engagement, performance, and perception in a flipped satellite classroom. *American Journal of Pharmaceutical Education*, *77*(9).

Phungoen, P., Promto, S., Chanthawatthanarak, S., Maneepong, S., Apiratwarakul, K., Kotruchin, P., & Mitsungnern, T. (2020). Precourse preparation using a serious smartphone game on advanced life support knowledge and skills: randomized controlled trial. *Journal of Medical Internet Research*, *22*(3), e16987.

Pintrich, P. R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ). MI: National Center for Research to Improve Postsecondary Teaching and Learning. (ERIC Document Reproduction Service No. ED 338122)

Roussou, M. (2004). Learning by doing and learning through play: an exploration of interactivity in virtual environments for children. *Computers in Entertainment (CIE)*, *2*(1), 10-10.

Smith, P. K., & Pellegrini, A. (2008). Learning through play. *Encyclopedia on Early Childhood Development*, *24*(8), 61.

Wang, L. C., & Chen, M. P. (2010). The effects of game strategy and preference- matching on flow experience and programming performance in game- based learning. *Innovations in Education and Teaching International*, *47*(1), 39-52.

Winslow, L. E. (1996). Programming pedagogy: A psychological overview. *SIGCSE Bulletin*, *28*, 17-22.

Zhang, Q., & Yu, Z. (2021). A literature review on the influence of Kahoot on learning outcomes, interaction, and collaboration. *Education and Information Technologies*, *26*(4), 4507-4535.

網路反思提示下小組友誼對高中生科學探究歷程反思的影響

The Impact of Peer Friendship on High School Students' Reflective Processes in Science Inquiry under Online Reflective Prompts

邱瓊慧
國立臺灣師範大學資訊教育研究所
cchui@ntnu.edu.tw

【摘要】 本研究的目的在探究高中學生於網絡反思提示輔助下，進行小組科學探究活動時的反思歷程，以及小組內部友誼關係對其反思品質的可能影響。研究資料來自 51 個自選小組、共 104 位綜合高中學生。他們完成了六週有平版及軟體輔助、以力與運動為題的科學探究活動，並在網路反思提示下對探究過程進行合作反思。研究顯示學生的反思品質隨週次與活動性質而波動，且組內友誼對自主探究活動（第四至六週）時的反思品質有顯著的正面影響。針對強化科學探究歷程的反思，本研究突顯了小組內友誼的重要性，為未來教學策略和學習環境設計提供有價值的參考。

【關鍵字】 科學探究；小組友誼；高中生；網路反思提示

Abstract: *This study examined high school students' reflection processes during group scientific inquiry, facilitated by online prompts, and assessed how within-group friendships affect reflection quality. Data from 104 students in 51 self-selected groups were analyzed. Over six weeks, students, supported by tablets and apps, engaged in inquiry activities and group reflections. Results showed that reflection quality varied over time and activity type, with friendships within groups notably enhancing reflections in self-directed activities during weeks four to six. The study underscores friendships' role in enriching scientific inquiry reflections, offering insights for future educational strategies and learning environment planning.*

Keywords: scientific inquiry, peer friendship, high school students, online reflective prompts

1. 序論

各國科學課程改革普遍重視探究教學的重要性。美國「國家科學教育標準（National Science Education Standards, NSES）」（National Research Council [NRC], 1996）以及「探究與國家科學教育標準（Inquiry and The National Science Education Standards）」（NRC, 2000）皆強調「探究」是科學學習的核心，認為科學探究是科學素養不可或缺的一部分。科學探究是科學家透過多種途徑對自然界進行瞭解與發展知識的方法，經常包括確認問題、尋求解答、與執行實驗等歷程。隨著手機與平板等行動裝置的普及，與其可裝載的程式越來越豐富且容易使用後，這些歷程將可能產生許多改變。

科學探究重視「知道什麼」、「為何知道」以及「如何知道」等思考過程（NRC, 2000）。在這一過程中，學生的反思至關重要，因為反思能夠促使學生從自身經驗中學習，從而深入理解科學探究的多個層面，包括科學方法、實驗設計、數據分析、結果解釋，以及科學溝通等等（Loh et al., 2001），學生因此不僅能夠進行深層的學習，還能加強對科學概念的理解，同時培養批判性思維與問題解決能力。為促進學生的學習反思，可以參採後設認知文獻建議的方法，例如，針對任務的計畫、監控或評估等面向，提供明確的提示，或是依據學習活動

屬性提供結構化的引導問題（Kramarski & Mevarech, 2003; Zion et al., 2005），這些策略都對促進反思有相當的效果。

合作互動也是科學探究中至為重要的部分，「2061 計畫」呼籲，學生應該要像科學家探究科學一樣，經常以小組合作的方式，分享程序和意義、辯證發現結果、並評估研究進展，共同追求對於知識的了解或發現（American Association for the Advancement of Science [AAAS], 1990）。這可幫助學生從中獲得成就感，同時得到別人的回饋、建設性的評論、或更進一步的澄清。隨著電腦和網路技術的進步，相關平台的應用更可支持這樣的合作，使學生能夠即時交流、共享資源，進而增強合作的效率和深度。

分組是執行合作學習重要的第一步，一般採隨機分派、教師指派、或學生自選。隨機分派對每個學生或小組能維持公平性；教師分組通常會依學生能力、專長或其他背景分派；學生自選則是學生有選擇自己組員的自由。雖然許多分組研究普遍針對同質與異質組合做探討，近來的研究卻也發現，在不同的同儕分配規則中，同儕組成的變化對於表現差異的貢獻微小；相對的，自由選擇對於小組表現則會有直接的影響（Kiessling et al., 2022）。此外，許多調查發現青年學生傾向自選小組，且特別喜歡選擇自己朋友，而這樣的友情小組雖能順利溝通與討論（Nhan & Nhan, 2019），卻也容易花時間談論與任務無關的話題（Mozaffari, 2017）。

儘管合作學習與反思在科學探究教學中的重要性普遍得到認同，但如何透過網路工具來加強學生的反思過程，以及小組內友誼如何具體影響反思品質，學術界尚未有足夠研究。尤其是學生在組隊過程中自選朋友的情形，這種選擇如何影響他們的學習和反思品質，更是一個值得深究的問題。

2. 研究目的與問題

為促進高中學生透過小組合作進行科學探究，本研究在探究過程中引入了網路反思提示。本研究旨在探討於網路反思提示支持的科學探究學習環境中，高中生如何進行過程反思，亦即這些反思的表現和品質，同時，基於友情因素可能促進高中學生間的互相支持與合作，本研究特別關注小組內友誼對於促進反思品質的潛在作用。具體而言，本研究在回答下列兩個問題：

1. 在網路反思提示支持的科學探究學習中，高中學生於探究過程的反思表現如何？
2. 在網路反思提示支持的科學探究學習中，小組內的友誼是否影響高中生的反思品質？

3. 研究方法

3.1. 參與者

本研究的科學探究實驗課程在一市立普通型完全中等學校實施，有 4 個「探究與實作」班級、共 137 名 11 年級學生註冊，其中 104 名學生（30 名男生和 74 名女生，年齡介於 16 至 17 歲之間）同意提供本研究所需資料。這所學校位於非都市區域，包含初中及高中階段學生，高中學生的學業表現在全國學生中屬於中等水平。參與學生對本次學習活動與安排的態度普遍積極，在使用電腦或平板進行協作或探究的經驗則相當有限。

3.2. 活動與設置

由於參與者參與科學探究的經驗有限，不容易僅只透過一個完全開放的問題，便任由他們自行去收集、解釋資料、或形成知識。本研究因此採引導式探究方法，透過案例探索、實作練習、自主探究等程序，幫助參與者藉由經歷引導獲得的經驗（第一至三週），逐步實踐自己的科學探究（第四至六週），並完成科學探究成果。學生以 2-3 人自選組員的方式於學校實驗室使用平版電腦（iPad）進行此活動。

3.3. 網路反思提示

本研究透過 Trello 的卡片功能提供反思提示並要求學生記下反思結果。反思提示參考 Kramarski 和 Mevarech（2003）與 Zion et al.（2005）的提問引導，並考量不同探究階段的屬性做設計，區分成知識與執行兩個面向，共 18 個提示。學生平均每週會有三個提示。除第六週外，一提示於課程一開始時提供，另二個則於課程結束時呈現。此外，本研究引導小組內成員以 Think-pair-share 協作策略進行互動以回應反思提示、進行小組的反思。

3.4. 數據收集

資料來源主要包括兩部分：分別是小組於活動過程的反思品質，與學生選組員時考量友誼的程度。

小組反思品質乃根據學生在 Trello 卡片上對每個反思提示的回應來評定。兩位受過訓練的評分員根據回應的品質給予 0 至 2 分的評分，這意味著每個階段每組學生的得分範圍是 0 至 6 分。評分者信度通過 Cohen's kappa 係數計算，結果為 0.805。

學生自選組員時考慮到友誼的程度，乃透過學生對於陳述「我選擇我的組員是因為我們是好朋友」的回應來評估。學生根據自己選組的情況選擇：完全符合、大部分符合、大部分不符合、或完全不符合，本研究則分別以 4、3、2、1 的分數進行後續分析。

4. 結果與結論

4.1. 過程反思表現

表 1 呈現小組學生反思品質的描述性統計資料與混合線性模式分析結果。固定效應模式估計結果顯示，時間的主要效應達到顯著水準，$F(5, 245.45) = 12.03, p < .001$，且隨著週次與活動屬性變化，小組反思品質不盡相同。針對前導訓練活動，包括第一週前置準備、第二週經驗探索、第三週互動示範，若以前一週為基準，學生在第二週有明顯提升，第三週因學生需有較多的探索，反思品質明顯下降。自第四週起，學生開始進行自主探究，他們的反思表現下降趨緩（未達顯著），但在進入第五周後，反思表現復顯著提升。

表 1

學生小組於各週次之反思品質和變化

	反思品質		Linear mixed model fit by REML						
	M	*SD*	**Effect**	*Estimate*	*SE*	*95% CI*	*df*	*t*	*p*
Week1	3.40	1.64	(Intercept)	4.135	0.139	[3.86, 4.41]	49.1	29.67	< .001
Week2	4.71	1.02	Week1 - 2	-1.229	0.202	[-1.63, -0.83]	245.7	-6.07	< .001
Week3	4.12	1.42	Week 2 - 3	0.507	0.202	[0.11, 0.90]	245.7	2.50	0.013
Week4	3.74	1.15	Week 3 - 4	0.389	0.200	[-0.003, 0.78]	245.1	1.94	0.053
Week5	4.28	1.47	Week 4 - 5	-0.534	0.201	[-0.93, -0.14]	245.5	-2.65	0.008
Week6	4.66	1.25	Week 5 - 6	-0.380	0.202	[-0.78, 0.02]	245.1	-1.88	0.061

4.2. 組內友誼影響

表 2 呈現學生小組於四至六週進行自主科學探究活動時的反思表現，混合線性模式分析結果顯示，組內的友誼關係對學生的反思歷程有顯著的影響，$F(1, 48.56) = 4.40, p = .041$。

表 2

組內友誼對於自主探究學習時之反思品質的影響

Effect	*Estimate*	*SE*	*95% CI*	*df*	*t*	*p*
(Intercept)	4.218	0.134	[3.95, 4.48]	48.6	31.50	< .001

Week 4 - 5	-0.529	0.206	[-0.93, -0.12]	98.7	-2.56	0.012
Week 5 - 6	-0.380	0.207	[-0.79, 0.03]	98.0	-1.84	0.070
友誼關係	0.555	0.265	[0.04, 1.07]	48.6	2.10	0.041

4.2. 結論與建議

研究結果顯示，隨著時間的推移，學生小組的反思品質經歷了顯著的變化。在初期的前導訓練活動中，學生的反思品質在第二週明顯提高，但在第三週時下降，這可能是由於學生在探索活動中遇到了挑戰。然而，隨著學生從第四週開始進行自主探究，他們的反思能力顯示出逐步改善的趨勢，並在第五週達到顯著提升。此外，友誼關係對於促進學生在反思歷程中的正面效應不容忽視，這表示，在小組組成時同儕的良好關係，可能有利於促進積極的互動，因而能提升學生的反思品質。本研究除揭示任務屬性與友誼因素如何影響學生合作反思過程外，也提示教學設計者需要重視學生在小組學習過程中的感情連結，並提供持續和階段性的指導，以幫助學生更有效地從科學探究活動中學習。

致謝

感謝行政院國家科學委員會對本研究經費的補助，編號為 MOST 109-2511-H-003 -012 -MY3。

參考文獻

American Association for the Advancement of Science [AAAS]. (1990). *Science for all Americans, Project 2061*. New York, NY: Oxford University Press.

Kiessling, L., Radbruch, J., & Schaube, S. (2022). Self-selection of peers and performance. *Management Science, 68*(11), 8184–8201. https://doi.org/10.1287/mnsc.2021.4267

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American educational research journal, 40*(1), 281-310.

Loh, B., Reiser, B. J., Radinsky, J., Edelson, D. C., Gomez, L. M., & Marshall, S. (2001). Developing reflective inquiry practices: A case study of software, the teacher, and students. *Designing for Science: Implications from Everyday, Classroom, and Professional Settings*, 279-323.

Mozaffari, S. H. (2017). Comparing student-selected and teacher-assigned pairs on collaborative writing. *Language Teaching Research, 21*(4), 496-516.

National Research Council [NRC]. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council [NRC]. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press. doi: 10.17226/9596.

Nhan, H., & Nhan, T. A. (2019). Different grouping strategies for cooperative learning in English majored seniors and juniors at Can Tho University, Vietnam. *Education Sciences, 9*. https://eric.ed.gov/?id=EJ1211904

Zion, M., Michalsky, T., & Mevarech, Z. R. (2005). The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. *International Journal of Science Education, 27*(8), 957-983.

Revitalizing Mathematical Proficiency through ScaffoldiaMyMaths: Empowered by Adaptive Cognitive Scaffoldings

Daner Sun[1*], Zhizi Zheng[2], Lan Yang[3], Kee-lee Chou[4], Chee-Kit Looi[5], Zhixuan Song[6]
[1,2,6]Department of Mathematics and Information Technology, The Education University of Hong Kong
[4]Department of Asian and Policy Studies, The Education University of Hong Kong
[5]Department of Curriculum and Instruction, The Education University of Hong Kong
[*] dsun@eduhk.hk

Abstract: *Amidst the increasing reliance on online learning, understanding students' perspectives and addressing the challenges they face is crucial. This system, namely, ScaffoldiaMyMaths, has been designed and developed to facilitate underprivileged and lower-ability children's maths learning in and out of the classroom. ScaffoldiaMyMaths integrates context-specific and dynamic scaffoldings, offering personalized support to students. The system features dynamic virtual manipulative representations, adaptive cognitive and metacognitive scaffoldings, explicit procedural and conceptual supports, and AI-based responsive assessment.*
Keywords: learning system, maths education, adaptive scaffoldings

1. Introduction

Numerous studies have shed light on the advantages of utilizing scaffolding-based mathematics tools for students with mathematics learning difficulties (MLD) (Chin & Fu, 2021; Lei et al., 2020). For instance, Hwang et al. (2015) demonstrated how a learning system integrating symbolic and graphical representations significantly boosted learning achievement, attitudes, and motivation in fraction-related problem-solving among students. Similarly, Shin and Bryant (2017) showcased the effectiveness of Fun Fraction, a multi-component computer-assisted instructional program, in enhancing word problem-solving skills with fractions among students with MLD. Despite these strides, notable research gaps persist. While past studies have explored diverse forms of scaffolding changes in mathematics learning, such as adding support or maintaining the status quo, there remains a dearth of investigation into the effects of fading scaffolding, especially concerning students with learning difficulties. This gap underscores the imperative for further exploration into the intricate dynamics of scaffolding within educational contexts. This paper demonstrates a maths learning system, namely, ScaffoldiaMyMaths. The platform is not solely focused on addressing academic challenges; it also provides an extensive array of resources and tools explicitly tailored to assist students in refining their self-directed learning skills, thereby empowering them to take charge of their educational journey. Nurturing self-directed learning abilities enables students to channel their focus more effectively toward academic pursuits, thereby enhancing the overall efficacy of the learning management process.

2. The Structure and features of ScaffoldiaMyMaths

Building upon the existing design, ScaffoldiaMyMaths, a comprehensive web-based learning management system, is currently being developed to facilitate primary mathematics learning for students in Hong Kong. This innovative platform integrates context-specific and dynamic scaffolding principles, a feature rarely seen in current learning environments. Within ScaffoldiaMyMaths, instructional lessons are structured into five sets of fraction problems, each comprising ten tasks with consistent underlying structures but varying surface characteristics. To support students in problem-solving, scaffoldings are designed for each task. Key scaffoldings include explanation prompts (EP) in the form of worked-out examples and problems, following the framework proposed by Atkinson and Renkl (2003). These EPs consist of four fixed solution steps, visually presented and tailored to the nature of each task. A demonstration of these scaffoldings in

action can be accessed through the provided link, along with Figures 1.a and 1.b illustrating the interface of adaptive scaffolding in ScaffoldiaMyMaths.

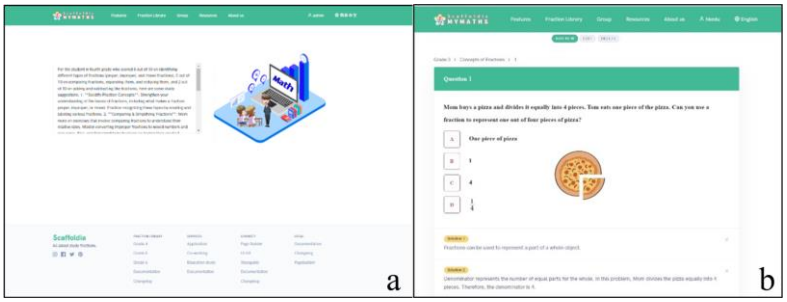


Figure 1. The interface of adaptive scaffolding in ScaffodiaMyMaths

The scaffolding operates across four distinct conditions, providing varying levels of support to learners throughout the learning process. These scaffoldings are designed within the system to optimize student engagement and comprehension. The system designed four types of scaffoldings: (a) Fixed backward fading scaffolding condition, (b) Fixed fading and adding scaffolding condition, (c) Performance-based fading scaffolding, and (d) Performance-based fading and adding scaffolding to help a diversity of students to breakthrough the difficulties of maths.

3. Conclusion

ScaffoldiaMyMaths learning environment boasts several key features. Firstly, it incorporates dynamic virtual manipulative representations to alleviate learners' cognitive burden, thus relieving them from the task of maintaining abstract concepts. Secondly, the system integrates dynamic cognitive and metacognitive scaffoldings, drawing on relevant studies to accommodate individual differences and address specific needs. This includes adjusting the frequency and nature of scaffoldings, such as fading, adding, and fading/adding supports, through a dynamic assessment of students' abilities. Thirdly, procedural and conceptual scaffoldings are provided explicitly to assist students in identifying the operational aspects and essential knowledge of problems. Moreover, ongoing system development focuses on enhancing the functions of learning management and learning analytics, enabling smoother teacher-student interaction, assessment, and monitoring through AI technology. Finally, the pedagogical design of ScaffoldiaMyMaths learning activities, complemented by a theoretical framework of technology integration, aims to better facilitate teachers and students in utilizing the system effectively and pedagogically.

Acknowledgments

The research is funded by General Research Fund: 18609221 Preparation for Post Pandemic Society: Providing Technology-Supported Adaptive Scaffoldings for Students with Mathematics Learning Difficulties in Hong Kong Primary Schools.

References

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of educational psychology*, *95*(4), 774.

Chin, K. E., & Fu, S. H. (2021). Exploring the Implementation of an Intervention for a Pupil with Mathematical Learning Difficulties: A Case Study. *Journal on Mathematics Education*, *12*(3), 531-546.

Hwang, G. J., Lai, C. L., & Wang, S. Y. (2015). Seamless flipped learning: a mobile technology-enhanced flipped classroom with effective learning strategies. *Journal of computers in education*, *2*, 449-473.

Lei, Q., Xin, Y. P., Morita-Mullaney, T., & Tzur, R. (2020). Instructional Scaffolds in Mathematics Instruction for English Learners With Learning Disabilities: An Exploratory Case Study. *Learning Disabilities: A Contemporary Journal*, *18*(1), 123-144.

Shin, M., & Bryant, D. P. (2017). Improving the fraction word problem solving of students with mathematics learning disabilities: Interactive computer application. *Remedial and Special Education*, *38*(2), 76-86.

A Game design and development for Enhancing Self-Efficacy

Yu, Bai[1], Wenhao Wang[2], Fuzheng Zhao[3], Chengjiu Yin[4*]

[1] Graduate School of Information Science and Electrical Engineering, Kyushu University

[2] Graduate School of System Informatics, Kobe University

[3] Education Technology Center, Jilin University

[4] Research Institute for Information Technology, Kyushu University

[*] yin.chengjiu.247@m.kyushu-u.ac.jp

Abstract: *Recognizing the significant impact of low self-efficacy on learning motivation, dropout rates, and academic achievement, it is imperative to address this challenge in instructional design. Game-based learning, which seamlessly embeds educational content into interactive game environments, presents a compelling avenue for intervention. This research delves into the realm of addressing modes learning within operating system courses, serving as a focal point for examining the efficacy of game-based learning in bolstering self-efficacy. Through the development of a side-scrolling video game tailored to this purpose, the study aims to evaluate its effectiveness in mitigating the effects of low self-efficacy.*

Keywords: Low self-efficacy, addressing modes learning, a side-scrolling video game

1. Introduction

The persistent issue of low self-efficacy in learning endeavors has attracted considerable scholarly interest due to its detrimental effects on academic performance and psychological well-being. Varying levels of self-efficacy influence learning pressure, with appropriate levels aiding in mitigating stress, improving efficiency, and sustaining interest (Bandura et al., 1996). Conversely, low self-efficacy leads to heightened pressure, avoidance behavior, and increased dropout rates. In computer-related courses, self-efficacy is particularly crucial, as challenging concepts often lead to diminished motivation (Pajares, 2002). Educational technology interventions, including mobile learning, online platforms, personalized learning, and virtual reality, aim to address self-efficacy challenges by adapting to learners' needs and providing immersive experiences (Howard, 2003). Game-based learning is also highlighted as a promising method for enhancing self-efficacy by fostering intrinsic motivation and creating supportive learning environments (Tobias, Fletcher, & Wind, 2014). This study focuses on applying gamified learning to improve self-efficacy in operating system courses, exemplifying the potential of game-based approaches in addressing self-efficacy issues in specific academic domains.

2. Game design and development

The gaming system comprises three essential modules: the UI module, responsible for navigation, content display, and basic controls; the Physics module, which governs the interaction of objects based on their physical attributes, such as the force exerted when a character jumps; and the Scene module, which manages the terrain, levels, and characters within the game. These modules were developed using the Unity game engine. Initially, the scene module was created using Unity's physics, graphics, and audio systems. Subsequently, game characters were incorporated into the scene, with their behavior and game logic implemented through Unity scripts. Finally, the uGUI was utilized to design the game's interaction interface, including various visual effects.

2.1. Implement of learning content in the game design

Correspondence between data locations and game maps. The intricate process of data storage and processing involves the CPU, memory, and auxiliary memory. Direct access to memory data contrasts with the indirect access

required for hard disk data, facilitated by memory. Both memory and hard disk storage rely on data blocks, with memory employing paging and segmentation tables. Game maps align with data locations, mirroring memory locations and integrating data knowledge. Bytes within the game's memory map represent data storage, enhanced by page and segment tables. Gates between game maps symbolize transitions to different data storage locations. This synergy between computer science and game design enriches the educational experience. Correspondence between proper nouns and the design of game items. Game design complexity arises in translating OS proper nouns into game elements. Game clues guide players through addressing processes, starting with logical addresses and progressing to clues embedded in page or segment tables. Game props, like doors representing data blocks, immerse players in data search realms. This relationship demands understanding and unravels addressing mode complexities, transcending traditional boundaries in gamified learning. Correspondence between addressing operations and game actions. Addressing operations in OS are gamified into game clues and props, guiding players through operational processes. Initiating addressing operations from the task board leads players through gathering essential clues and navigating immersive maps. Each step mirrors addressing mode learning, aligning with specific learning points, and culminates in solidifying mastery of addressing mode concepts. This sequential process offers an educational journey into OS intricacies within the gamified learning experience.

2.2. Learning strategy incorporated in the design of game modules

The educational system incorporates three distinct modules: the mission module, designed to guide students through addressing mode sequences; the guide module, providing instructions on game operation; and the feedback module, responsible for delivering task and character position feedback. Within the mission module, six missions represent various addressing modes, each with tailored steps. Participants acquire clues, calculate physical addresses, retrieve data, and interact with the environment. The feedback module seamlessly integrates game operations with learning instructions, offering comprehensive guidance that gradually diminishes as participants progress, promoting independent mastery. The guide module ensures participants remain informed about mission status, delivering timely updates on game location and clues. Mission location information aids in understanding mission progress, while participant location information enhances comprehension of game clues and props. This module empowers participants to effectively master their objectives by providing real-time insights and up-to-date information. The strategic integration of mission and participant location information enhances the overall effectiveness of the game-based learning experience.

3. Conclusions

The game was developed with the addressing mode of the operating system course as the focal learning content. This integration occurred through three key approaches: utilizing the hard disk, memory, and CPU as game maps; transforming technical terms into game elements; and gamifying operational behaviors from the addressing mode. Additionally, the game features three modules incorporating learning strategies: the mission module for providing cues, the feedback module for offering tips, and the guidance module, which serves a similar function to providing tips. These modules were designed to enhance students' engagement and understanding of addressing modes within the context of the game.

References

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child development*, 67(3), 1206-1222.

Heward, W.L. (2003). *Exceptional Children (7th edn).* Upper Saddle River, NJ: Merrill-Prentice Hall

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into practice*, 41(2), 116-125.

Tobias, S., Fletcher, J. D., & Wind, A. P. (2014). Game-based learning. *Handbook of research on educational communications and technology,* 485-503.

Embedding Student Teachers' Feedback Literacy Development in Online Peer Assessment Activities

Ying Zhan [1*], Daner Sun[2], Jan Christian C Gube[1], Ngok Cheng Chan[1], Kam Wing Chan[1], Zhi Hong Wan[1] and Junjun Chen[3]

[1] Department of Curriculum & Instruction, The Education University of Hong Kong

[2] Department of Mathematics and Information Technology, The Education University of Hong Kong

[3]Department of Education Policy and Leadership

[*] zhanying@eduhk.hk

Abstract: *Student feedback literacy has been recently regarded as a crucial condition for deciding the power of feedback over learning. It has also been highlighted as a core one of graduate attributes that supports students' future careers and life-long learning. However, the cultivation of student feedback literacy is a challenging task for university teachers. This teaching project aims to develop student teachers' feedback literacy via online peer assessment through three phases for two rounds in three educational courses. This quasi-experiment involves 240 student teachers in three experimental class groups (~120 students) and three control class groups (~120 students). A pre-and-post survey on feedback literacy, individual interviews and log data are to be collected to discern whether online peer assessment exerts impacts the development of student teachers' feedback literacy and summarize the lessons of this assessment innovation embedded in course teaching.*

Keywords: student teachers, feedback literacy, online peer assessment, teacher education

1. Introduction

Feedback is always perceived as a crucial element in higher education which contributes to the enhancement of student learning worldwide (Winstone et al., 2021). However, there exists a paradox on feedback: why is feedback theoretically powerful yet apparently ineffective in classrooms (Hattie & Clarke, 2019). Discontent with feedback practice prevails in higher education contexts. More and more researchers have realized that the power of feedback over learning cannot be realized without students' active engagement with feedback (Wood, 2021). Students' agentic engagement with feedback requires them to be equipped with adequate feedback literacy, which includes their competencies and dispositions (Carless & Boud, 2018). In addition, student feedback literacy has long-term impact on their future career and life-long learning (Winstone, Balloo & Carless, 2022).

Cultivation of student feedback literacy is a challenging task for university teachers. They are facing the challenges of integrating training of student feedback literacy into the crowded curricula and finding effective methods to conduct such training. To meet the first challenge, the development of feedback literacy can be integrated within disciplinary curricula as opposing to using a 'bolt-on' approach to save course time and space (Winstone, Balloo & Carless, 2022). For the second challenge, researchers have proposed potential ways to holistically enhance student feedback literacy in higher education. Carless and Boud (2018) suggested that the strategies of peer assessment and analysing exemplars could be used to foster the development of feedback-literate students. Malecka, Boud and Carless (2020) recommended four strategies namely feedback requests, self-assessment, peer review and e-portfolios to develop student feedback literacy. Among these strategies, peer assessment has been applied to authentic teaching settings and has been initially demonstrated as a powerful approach (Ketonen, Nieminen, & Hähkiöniemi, 2020).

Peer assessment could be regarded as a powerful way to cultivate student feedback literacy. On one hand, when students take on the role of assessors, they are expected to analyze and assess their peers' work, identifying both their strengths and weaknesses based on predetermined criteria. This process provides valuable opportunities for students to deepen their understanding of discipline-specific knowledge and criteria (Tai et al., 2018). Meanwhile, as students assume the role of assessors, their motivation to actively participate in the assessment process and their responsibility for evaluating their peers' work will increase. On the other hand, when students act as assessees, they carefully examine and reflect upon the feedback given by their peers, utilising it to improve their own work (Li, 2017). This provides valuable opportunities for students to enhance their evaluative and reflective skills as they make informed judgments based on the feedback received from their peers. In all, peer assessment not only enhances students' knowledge and skills to engage in feedback but also improves their positive tendency towards feedback which is known as student feedback literacy.

The potential of peer assessment in promoting feedback literacy has been enhanced by technology (Carless & Boud, 2018). In addition to convenience, speed of delivery and traceability, the biggest affordance of technology is to provide more efficient and meaningful interaction between students when they are engaged in peer assessment process (Wood, 2021; Zhan et al, 2022). Interaction between peers is crucial for student feedback literacy development because it provides opportunities of questioning, clarifying, reflecting and monitoring to understand and use feedback for improvement (Malecka et al., 2020; Schilings et al., 2020).

2. Research questions

This project attempts to address two following research questions:

- To what extent and in what aspects does online peer assessment improve feedback literacy of student teachers?
- How is online peer assessment designed and implemented to improve feedback literacy of student teachers?

3. Design of online peer assessment activities

Enlightened by the latest literature, this project proposes that online peer assessment is integrated meaningfully and cumulatively into three educational courses to promote student teachers' feedback literacy. Online peer assessment will be designed in a structured way to facilitate a productive dialogue around peer feedback in online environment. Three stages of online peer assessment will be conducted for two rounds (see Figure 1).

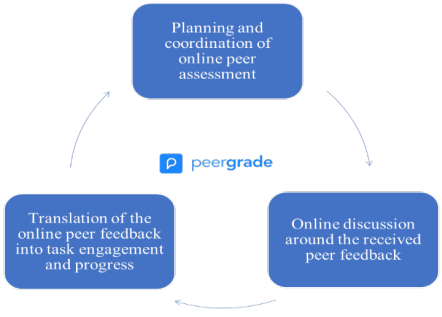


Figure 1. The three-stage process of online peer assessment adopted in this project

In planning and coordination phase, training is regarded as a powerful design element across cultures and should be conducted at the very beginning. In training, a blended method of teacher instruction and student discussion are adopted to train the participants to better understand the provided assessment criteria which build up the base for fruitful dialogic peer feedback in the second phase. In the discussion phase, the participants try to understand the received feedback by reflecting on it and discussing it with their peers. This is the key stage because during discussion, the participants have chances to better understand different perspectives towards assessment plans, justify their perspectives with more evidence and reasons and accordingly make wise decision for refinement in the next stage. In the translation phase, the

participants are provided with revision opportunities. The revision is very important since it is the decision made by the participants after discussing with peers about the received comments.

4. Procedure of the project and data collection

This quasi-experiment involves 240 student teachers in three experimental class groups (~120 students) and three control class groups (~120 students). At the beginning of experiment, experimental groups attend a training session to familiarise themselves with the platform and preparing for online peer assessment activities. In addition, experimental and control groups are invited to answer the survey questions on student feedback literacy. The feedback literacy scale contains six dimensions, which include students' competencies of a) eliciting, b) processing and c) enacting feedback and the students' disposition in terms of d) appreciation of feedback, e) readiness to engage and f) commitment to change.

The experiment groups do online peer assessment when they learn the relevant topics of assessment in the course. In total, they do online peer assessment for two rounds through 3 stages as mentioned before. Each round of online peer assessment activity lasts for two weeks. The control groups are taught without online peer assessment activities.

At the end of the experiment, experimental and control groups are invited again to answer the survey questions on student feedback literacy. 6 students with different extents of development on student feedback literacy (high, middle and low level) from each experimental group are approached and invited to participate in individual interviews. In the individual interviews, the participants share their experiences of online peer assessment and gains in feedback literacy, reveal their perceived facilitating and inhibiting factors and suggest ways to sharpen online peer assessment for student feedback literacy development. In addition, log data concerning students' participation in online peer assessment, design work, and communication between peers are collected to triangulate survey and interview data concerning the effects of online peer assessment on student teachers' feedback literacy development.

5. Innovation of the project

The project accumulates exemplary experiences and resources of online peer assessment to promote student feedback literacy in the following innovative ways:

First, online peer assessment has been seldom tried out to promote student feedback literacy in spite of its potential. In addition, peer assessment is a social activity and situated in specific contexts. Therefore, the design of online peer assessment needs to suit the Chinese context and Chinese students, which will make it different and context-specific.

Second, this is an attempt to integrate the cultivation of student feedback literacy into the educational courses without requiring extra course time and space. The integration should align with course goals and content, which requires teaching wisdom and careful design.

Third, different from other studies which have claimed the effectiveness of peer assessment on student feedback literacy in a qualitative way, this study uses the scale of student feedback literacy developed by the first author to make a more accurate estimation of the effects of online peer assessment in a statistical way.

6. Pedagogical implications

The assessment innovation generated by this project enriches and promotes university teachers' practice in online classes. 4 ready-for-use assessment tasks for online peer assessment are to be shared with course teachers. They can make further adaptations and integrate them into their own courses. They can use online peer assessment to expand their repertoire of pedagogical strategies in an online environment and make their teaching more student-centred and beneficial for undergraduates and postgraduates. Meanwhile, online peer assessment also reduces teachers' workload of providing feedback to individual students. For the teachers who teach other courses, although they cannot directly use the 4 ready-for-use assessment tasks, they can refer to the effective design principles and implementation guidance of online

peer assessment activities identified by this project and then integrate their course content into the concrete assessment tasks. Thus, this project indirectly generates pedagogical impacts on other course teachers.

Furthermore, because the participants in this projects are student teachers, they have chances to propagate such feedback practice into their future classrooms by observing their teachers and reflecting on their experiences. This will generate pedagogical implications for schools.

Acknowledgements

This project has been funded by Teaching Development Fund (T0262) in the Education University of Hong Kong and Early Career Scheme (No.: 28609721) and General Research Fund (No.: 18605223) funded by the University Grants Committee in Hong Kong

References

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315-1325. https://doi-org.ezproxy.eduhk.hk/10.1080/02602938.2018.1463354

Hattie, J. & Clarke, S. (2018). *Visible learning: Feedback* (1st ed.). Routledge. https://doi-org.ezproxy.eduhk.hk/10.4324/9780429485480

Ketonen, L., Nieminen, P., & Hähkiöniemi, M. (2020). The development of secondary students' feedback literacy: Peer assessment as an intervention. *The Journal of Educational Research, 113*(6), 407-417. https://doi-org.ezproxy.eduhk.hk/10.1080/00220671.2020.1835794

Li, L. (2017). The role of anonymity in peer assessment. *Assessment & Evaluation in Higher Education, 42*(4), 645-656. https://doi-org.ezproxy.eduhk.hk/10.1080/02602938.2016.1174766

Malecka, B., Boud, D., & Carless, D. (2020). Eliciting, processing and enacting feedback: mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, *27*(7), 908–922. https://doi-org.ezproxy.eduhk.hk/10.1080/13562517.2020.1754784

Schillings, M., Roebertsen, H., Savelberg, H., Whittingham, J., & Dolmans, D. (2020). Peer-to-peer dialogue about teachers' written feedback enhances students' understanding on how to improve writing skills. *Educational Studies, 46*(6), 693-707. https://doi-org.ezproxy.eduhk.hk/10.1080/03055698.2019.1651692

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education, 76*, 467-481. https://doi-org.ezproxy.eduhk.hk/10.1007/s10734-017-0220-3

Winstone, N., Boud, D., Dawson, P., & Heron, M. (2022). From feedback-as-information to feedback-as-process: a linguistic analysis of the feedback literature. *Assessment & Evaluation in Higher Education, 47*(2), 213-230. https://doi-org.ezproxy.eduhk.hk/10.1080/02602938.2021.1902467

Winstone, N. E., Balloo, K., & Carless, D. (2022). Discipline-specific feedback literacies: A framework for curriculum design. *Higher Education, 83*, 57-77. https://doi-org.ezproxy.eduhk.hk/10.1007/s10734-020-00632-0

Wood, J. (2021). A dialogic technology-mediated model of feedback uptake and literacy. *Assessment & Evaluation in Higher Education, 46*(8), 1173-1190. https://doi-org.ezproxy.eduhk.hk/10.1080/02602938.2020.1852174

Zhan, Y., Wan, Z. H. & Sun, D. (2022). Online formative peer feedback in Chinese contexts at the tertiary Level: A critical review on its design, impacts and influencing factors. *Computers & Education*, Article 104341 https://doi.org/10.1016/j.compedu.2021.104341

基於情境感知與機器學習引導機制之遊戲式學習模式對學習者學習成效之影響

Effects of a game-based learning approach utilizing context awareness and machine learning guided mechanisms on learners' learning performance

陳君銘[1*]，葉鎮源[1]
[1]國立自然科學博物館營運典藏與資訊組
[*] jackychen.academic@gmail.com

【摘要】 近年來，隨著博物館在教育領域逐漸受到重視，遊戲中結合有效的學習引導，將更能提升學習者的學習成效與動機。然而，在博物館導覽時間有限的情況下，通常沒有辦法給予學習者足夠的時間或引導，這也造成學習過程中重要資訊無法被掌握，進而顯著影響學習者整體的學習表現。為此，本研究提出了基於情境感知與機器學習引導機制之遊戲式學習模式，強化學習者對博物館展示內容的記憶與理解。實驗結果顯示，本研究提出的結合情境感知與機器學習引導機制之遊戲式學習模式能有效提升學生的學習動機和學習成效。
【關鍵字】 情境感知；機器學習技術；博物館學習；自主學習

Abstract: *With the increasing importance of museums in the field of education, the learning performance and motivation can be enhanced through the combination of games with effective learning guidance and techniques. However, learners are normally not provided with sufficient time and guidance due to the limitations of time, such that much important information is abandoned, significantly influencing a learner's overall learning performance. To cope with the problem, this study proposes an integrating context-awareness and machine learning into a game-based learning approach that enhances learners' understanding and memory of museum exhibits. The results of the study indicate that the proposed game-based learning approach can effectively improve students' learning motivation and performance.*

Keywords: context-awareness, machine learning, museum learning, self-directed learning

1. 前言

近年來博物館的教育功能已普遍受到認同與重視，博物館提供了一個支持學習者主動和自發性的學習環境 (Hu & Hwang, 2023; Noble, 2021)。相關研究探討了博物館如何促進參與式學習、遊戲式學習和社會文化意識的培養，並顯示博物館學習對於發展學習者各方面的能力也有正面的影響。(Lin, Tang, & Tu, 2023; Shahab, Mohtar, Ghazali, Rauschnabel, & Geipel, 2023)。然而，在傳統單向的博物館導覽學習中，缺乏互動性與參與性，無法有效引導學習者探索、形成概念或應用概念解決問題。同時在缺乏足夠的學習引導與輔助的情況下，往往影響學習者的學習表現。因此，有效的學習引導及工具，對於在博物館學習的學習者來說是必要的 (Chen, Zhang, Chen, & Fan, 2023; Karaduman, Alan, & Yiğit, 2023; Tatlı, Çelenk, & Altınışık, 2023; Xu, Dai, Shen, & Yao, 2023)。

因此，本研究以情境感知與機器學習為基礎，發展了一種遊戲式學習模式，並建立相應的行動導覽學習環境。透過故事情境引導與任務模式，引導學習者進行統整、歸納與組織概念知識。同時將機器學習決策樹技術融入到數位遊戲式學習模式中，進而幫助學習者掌握概念知識及促進學習過程中的知識整合，期望讓學習者在博物館導覽學習過程不只增進學習興趣，更能提升學習者的學習表現。

2. 基於情境感知與機器學習引導機制之行動遊戲式學習系統

為探討基於情境感知與機器學習引導機制之遊戲式學習模式對學習成效的影響，本研究以環境生態作為遊戲學習內容，引導學習者在遊戲過程中，學習環境生態相關知識，並於每關結束後進行學習評量。藉由學習者與實際環境及學習系統間的互動，逐步建構出對環境生態的知識。本系統架構包含：

(1) 情境感知遊戲化模組：提供遊戲中結合即時定位、路線導引、遊戲情境與遊戲獎懲機制等功能，可主動偵測學習者位置並導引至各學習目標進行學習活動。

(2) 機器學習模組：透過遊戲中行為編碼，包括探索、學習、回答、正確/錯誤答題等學習行為，整理出可能的行為特徵納入決策樹建置模式中，並藉由決策樹 C4.5 建構模式評估決策樹中各子節點最適合的分類屬性，透過決策樹剪枝與交叉驗證，進而更適切的瞭解學習者對學習目標的觀察及內容知識的掌握，再透過更深入發掘學習者理解的真實情形，進而有效診斷出學習者問題點，引導相關的提示及補充教材。

(3) 測驗評量及成績管理模組：記錄遊戲過程中的測驗與學習成果。

本研究透過遊戲式漸進的關卡任務引導，讓學習者進行系統化學習及評量。同時，藉由情境感知與機器學習引導機制，有效地管理及掌握學習者個別的學習狀態，進而針對學習者較弱的知識環節，提供適當的引導策略、提示及補充教材。

圖 1

學習系統介面及流程

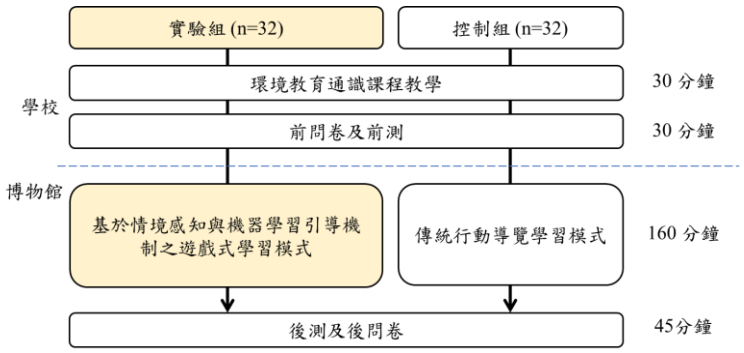


3. 實驗設計

本研究對象為某大學二年級通識課程的學生。實驗組採用基於情境感知與機器學習引導機制之遊戲式學習模式，控制組採用傳統行動導覽學習，具備與實驗組相同的學習教材、評量內容及行動導覽服務，但無導入情境感知與機器學習引導機制，引導學生至各學習點觀察學習目標及進行學習活動與回饋。本實驗流程如圖 2 所示，首先進行環境教育通識課程教學後，再進行前問卷及前測。其中實驗組使用情境感知與機器學習引導機制之遊戲式學習模式，進行學習活動，控制組則使用傳統行動導覽學習模式進行博物館學習。結束後，進行後測及後問卷。本研究學習動機問卷則採用 Wang 及 Chen (2010)所開發的李克特 5 點動機量表，信度 Cronbach's alpha 值為 0.79。認知負荷問卷則採用 Hwang、Yang 及 Wang (2013)所開發的認知負荷 5 點量表，其問卷包含 5 題心智負荷及 3 題心智努力面向，對應的 Cronbach's alpha 分別為 0.88 和 0.84。

圖 2

實驗流程圖



4. 實驗結果

4.1.學習成就分析

為了評估本研究提出「基於情境感知與機器學習引導機制之遊戲式學習模式」對於學習成就是否有顯著影響，本研究針對進行兩組前測成績 t 檢定分析。實驗結果顯示控制組在前測平均數及標準差分別為 58.44 和 5.14；實驗組則為 59.38 和 6.57。t 檢定結果(t = -0.635, p > 0.05)說明活動前控制組與實驗組之間沒有明顯的差異。為了比較實驗組和控制組學生的學習成就，本研究採用單因子共變數分析(ANCOVA)方法評估兩組學習者的測驗成績，並以前測成績為共變量、後測成績為因變數進行分析。透過兩組成績進行變異數同質性檢定與組內迴歸係數同質性檢定，驗證是否符合共變數分析之基本假設。實驗結果發現迴歸係數同質性檢定分析得到顯著性為.677 > .05 亦未達顯著水準，表示兩組迴歸線之斜率相同；變異數同質性檢定得到顯著性為.055（F = 3.825） > .05，未達顯著水準，表示兩組成績之變異數具有同質性，符合共變數分析之基本假設。接著進行共變數分析。實驗結果顯示後測調整平均數及標準誤在實驗組則為 69.56 和 1.26，控制組為 65.76 和 1.26。結果(F = 4.57, p < 0.05)發現兩組在學習成就上有顯著差異，這也表示「基於情境感知與機器學習引導機制之遊戲式學習模式」的實驗組，比使用「傳統行動導覽學習模式」的控制組更有助於提升學習成就。

4.2. 認知負荷

本研究採用 Hwang、Yang 及 Wang (2013)發展的認知負荷 5 點量表評估不同學習方法是否對學習者的認知負荷產生不同的影響 (5 題心智負荷及 3 題心智努力面向， Cronbach's alpha 分別為 0.88 和 0.84)。實驗結果顯示實驗組的平均數及標準差分別為 2.44 和 0.56；控制組則為 2.87 和 0.34。t 檢定結果(t = -3.99, p > 0.01) 顯示兩組心智負荷面向達顯著差異，這也表示基於情境感知與機器學習引導機制之遊戲式學習模式，有助於降低學習者的認知負荷。

4.2.學習動機分析

本研究使用成對樣本 t 檢定分析評估學習者經歷不同學習模式前後的學習動機是否有所差異。實驗結果顯示控制組「傳統行動導覽學習模式」學習前後的學習動機沒有顯著差異(t = -0.44, p > 0.05)，而實驗組在學習後的學習動機顯著優於學習前的表現(t = -5.36, p < 0.001)；其結果顯示本研究「基於情境感知與機器學習引導機制之遊戲式學習模式」，透過引導學生觀察真實的學習目標及在學習過程給予適當的知識學習引導，協助學習者釐清學習目標之間的差異，將有助於提升學習者對於學科的學習動機。

5. 結論

本研究提出一種基於情境感知與機器學習引導機制之遊戲式學習模式，探討使用新的學習模式對博物館導覽學習成效上的影響，並以實驗評估其有效性。實驗結果發現本研究提出的方法與傳統行動式行動導覽學習方法相比，能顯著提升學習者學習動機與成效。然而，本研究因研究對象及時間等因素限制了樣本數量，未來可進一步研究探討對其他年齡層的影響，並進一步結合其他機器學習模式或人工智慧技術，以建立適性之學習診斷模型，提供更細微的學習回饋，作為未來相關領域學科教育及研究之參考。

參考文獻

Chen, G., Zhang, Y., Chen, N. S., & Fan, Z. (2023). Context-aware ubiquitous learning in science museum with ibeacon technology. In Learning, Design, and Technology: *An International Compendium of Theory, Research, Practice, and Policy* (pp. 3369-3392). Cham: Springer International Publishing.

Hu, Y., & Hwang, G. J. (2023). Promoting students' higher order thinking in virtual museum contexts: A self-adapted mobile concept mapping-based problem posing approach. *Education and Information Technologies*, 1-25.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education, 69*, 121-130.

Karaduman, H., Alan, Ü., & Yiğit, E. Ö. (2023). Beyond "do not touch": the experience of a three-dimensional printed artifacts museum as an alternative to traditional museums for visitors who are blind and partially sighted. *Universal Access in the Information Society, 22*(3), 811-824.

Lin, C. J., Tang, K. Y., & Tu, Y. F. (2023). Advancements and research topics of museum-based mobile learning: a review of journal publications from 2008 to 2019. *Interactive Learning Environments, 31*(8), 4729-4752.

Noble, K. (2021). Challenges and opportunities: Creative approaches to museum and gallery learning during the pandemic. *International Journal of Art & Design Education, 40*(4), 676–689.

Shahab, H., Mohtar, M., Ghazali, E., Rauschnabel, P. A., & Geipel, A. (2023). Virtual reality in museums: does it promote visitor enjoyment and learning?. *International Journal of Human–Computer Interaction, 39*(18), 3586-3603.

Tatlı, Z., Çelenk, G., & Altınışık, D. (2023). Analysis of virtual museums in terms of design and perception of presence. *Education and Information Technologies, 28*(7), 8945-8973.

Wang, L. C., & Chen, M. P. (2010). The effects of game strategy and preference-matching on flow experience and programming performance in game-based learning. *Innovations in Education and Teaching International, 47*(1), 39-52.

Xu, W., Dai, T. T., Shen, Z. Y., & Yao, Y. J. (2023). Effects of technology application on museum learning: a meta-analysis of 42 studies published between 2011 and 2021. *Interactive Learning Environments, 31*(7), 4589-4604.

How to Design a Competitive Serious Game to Improve University Students' Concept of RGB Mixing

Hsi-Hsun Yang
Department of Digital Media Design, National Yunlin University of Science and Technology, Yunlin, Taiwan
jimmy@yuntech.edu.tw

Abstract: *This study developed a competitive serious game called "Rainbow Juice Shop" aimed at enhancing players' understanding of the RGB color mixing concept during gameplay. We invited 26 students from the Design College of a University in central Taiwan to evaluate the serious game. They were divided into 13 pairs, with each pair playing for 18 minutes and then completing a User Engagement Scale (UES) questionnaire. The main goal was to investigate their engagement with "Rainbow Juice Shop." The results showed that most participants experienced high levels of attention and engagement with the game. However, the game's usability and aesthetics were identified as areas needing improvement. The novelty of this research lies in its evaluation of a competitive serious game designed to improve the RGB color mixing concept, achieving educational outcomes through entertainment.*

Keywords: competition, game design and production, game evaluation, serious games

1. Introduction

Video games originated for entertainment, featuring diverse and freely designed content. Unlike video games, serious games were created for education and training. Hence, they might sometimes be less fun in comparison (den Haan et al., 2018; Laamarti et al., 2014). The intention behind serious games is not to be unnecessarily solemn or rule-bound but to extend the impact of gaming further. Besides retaining gameplay elements, serious games incorporate additional objectives, such as applications in education, military training, industry, scientific training, healthcare, and public safety (Burke et al., 2009). Serious games can play the role of educating through entertainment or even entertaining through education (Plewe & Fursich, 2018). There are three aspects through which serious games can affect learning: motivation, knowledge acquisition, and attitude change (Klimmt, 2009). Therefore, serious games can facilitate the acquisition of new knowledge and skills during gameplay, enhancing learners' willingness to learn through high interactivity and precise objectives (Girard et al., 2013). Computer games' high interactivity, entertainment value, personalized design, and visual presentation can enhance skills, thinking, innovation, attention, memory, and teamwork (Paraskeva et al., 2010; Chen & Chiu, 2016). Coupled with the high acceptance of games, integrating interactive features, fun, immersion, and dynamic game elements with educational content significantly benefits training or educational processes (Van Eck, 2006).

To guide players toward the goals within the game, it is crucial to clearly define the objectives and inform players of the rules they must follow; otherwise, players may quickly become lost. The design of serious games requires even greater rigor to ensure clarity over educational objectives and emphasize the purpose of learning. Moreover, interaction and feedback are vital characteristics of game design. With feedback, players can understand the rules and structure of the game world. Thus, special care and attention are needed to provide immediate interaction and feedback, which can influence players' performance in the game. Even rewards can encourage players to maintain a positive attitude (Sweetser & Wyeth, 2005).

Song et al. (2013) found in their game research that when sports games are divided into competitive and non-competitive versions, it affects the choice of high-competitive players and low-competitive players, with the former preferring the competitive version and the latter opting for the non-competitive one. The mindset of players can influence the success or failure of the educational or training objectives of serious games, but this is not necessarily a fault in the design content of serious games. Therefore, it is crucial to distinguish game characteristics carefully. Another example is the research by Cagiltay et al. (2015), which concluded that providing a competitive environment in

serious games significantly improves learning motivation and outcomes. Thus, it is a correct choice for game designers to integrate competitive elements into the games they create.

Competition is the act of people vying for a common goal, and according to social interdependence theory, competition and cooperation define the cognitive mindset of an environment (Deutsch, 1949, 1973). Thus, creating a competitive environment prompts competitive responses, while cooperative situations foster cooperative interactions. Cooperation and competition have distinct characteristics and produce different outcomes. Competition is seen as a positive element of learning activities, enhancing students' attention, interest in learning, and motivation (Lin et al., 2010), and it stimulates individual learning potential and efficiency (Cagiltay et al., 2015). Competitive learning might stimulate students' extrinsic motivation, allowing them to focus more energy on the current learning task (Van Eck & Dempsey, 2002). The findings of Chen et al. (2018) also reaffirm the characteristics of competitive games; they can enhance learning outcomes, and competitive game-based learning can improve the learning efficiency of learners acquiring procedural knowledge (Pareto et al., 2012).

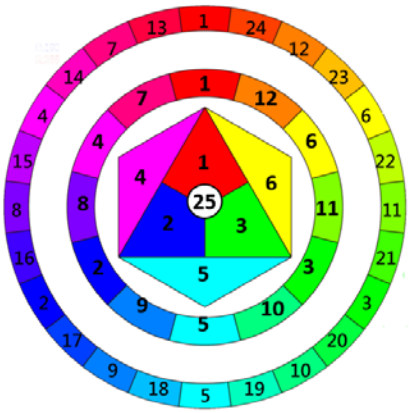
Peng et al. (2013) found that competitive games bring more enjoyment, a viewpoint that the author also supports. Players in competition can yield more positive outcomes and feel excitement, making it an effective way to motivate individual learning and play (Ebner & Holzinger, 2007; Cheng et al., 2013). Therefore, it seems reasonable to consider competitive elements to explain the fun in games. Chen and Chang (2020) note that the content of competitive games not only shapes the structure of the game (Tekinbas & Zimmerman, 2003) but, more importantly, reflects societal dynamics (Garcia et al., 2013). A classic example is the activation of player rankings by Cagiltay et al. (2015), stimulating higher motivation. This case is consistent with the findings of Hwang and Chang (2016), indicating that leaderboards can significantly enhance students' interest and attitude while reducing their cognitive load. Such results suggest that competitive modes can decrease players' anxiety, giving them more control, curiosity, and intrinsic interest (Chen & Chang, 2020).

In this study, we utilized an example from the Unity teaching website Code Monkey (2022) named Kitchen Chaos. After obtaining authorization from the website's administrators, the example was adapted into a serious game called "Rainbow Juice Shop." The educational theme of this study is the RGB color mixing concept. We integrated the production of 19 different colored juices into the operation of the Rainbow Juice Shop, moving beyond the traditional explanations of the three-color principle and color mixing theory. By creating 19 distinct colored juices within the serious game, we aim to generate heightened situational interest in the topic.

2. Game Design - Rainbow Juice Shop

The "Rainbow Juice Shop" game is based on the concept of color mixing, encompassing the principles of primary, secondary, and composite colors, featuring 25 different color mixing issues, as shown in Figure 1. The three primary colors can be mixed in various ratios to produce many colors. Adding red, green, and blue light in different proportions can synthesize multiple colors using an additive color model. Mixing two of the three primary colors can form secondary colors such as yellow, magenta, and cyan, and blending multiple colors can create tertiary or quaternary colors. Initially, 25 colors were proposed, but they were removed due to six of the colors needing to be more similar or complex. Therefore, the game can produce 19 different colored juices. Of course, this might only partially match the actual shades produced by mixing natural fruits since the game is designed for educational training on color mixing. However, through serious gaming, students' interest in color-mixing topics can be stimulated.

Figure 1.
The 25-Color structure of mixing primary colors.



"Rainbow Juice Shop" is a two-player competitive serious game. The gameplay process includes selecting fruits, cutting or peeling them, crushing them in a juicer, filling them into cups, and completing the fruit juice preparation. Each time players choose a fruit, the system records both players' numbers and the system time. The game randomly assigns 19 different color-mixing tasks, allowing players to try mixing the three primary colors in various proportions to produce fruit juice colors. The system records both players' fruit recipes for later analysis. When players complete and submit a juice with the correct color mix, the system is given the next task or moves to a more challenging level until the countdown ends or the game is over.

In addition to art design and programming, planning a game requires meticulous conceptual design, which includes a clear description of the game's internal economy. The internal economy involves collecting, consuming, and exchanging elements, not just referring to money or resources but also abstract concepts like stamina, prestige, and health that are part of the game economy. Using Machinations developed by Adams and Dormans (2012), we mapped out the entire internal economy of "Rainbow Juice Shop." Due to the complexity of Figure 2, different colors are used to correspond to the areas and directions of resource flow within the game. Red and blue lines represent the flow of personal resources and action orientation of the two players, respectively. Orange lines indicate the game's objectives and functional tools, while purple, yellow, and green areas represent different mechanism operations.

Green Area: Time Mechanism

Both players start simultaneously, and the competition's outcome is determined by the points each player accumulates. Every step within the game consumes varying amounts of time and resources. The game's timing is represented by symbols composed of equilateral triangles and circles.

Yellow Area: Order and Scoring Mechanism

In the game, both players will receive randomly assigned orders, each requiring the mixing of different colored fruits to produce the color specified in the order through a juicer. When a player successfully mixes the correct juice color and delivers it to the customer, a judgment mechanism composed of dice and diamonds is triggered, corresponding to the evaluation of fulfilled and unfulfilled conditions. If the order's requirements are met, the player's score will increase, and the order resource pool trigger will activate, adding one to the total number of correct orders. When a player's score reaches the set limit, a trigger made up of two square symbols is activated, indicating the completion of the individual's mission.

Red and Blue Lines: Learning and Game Mechanics

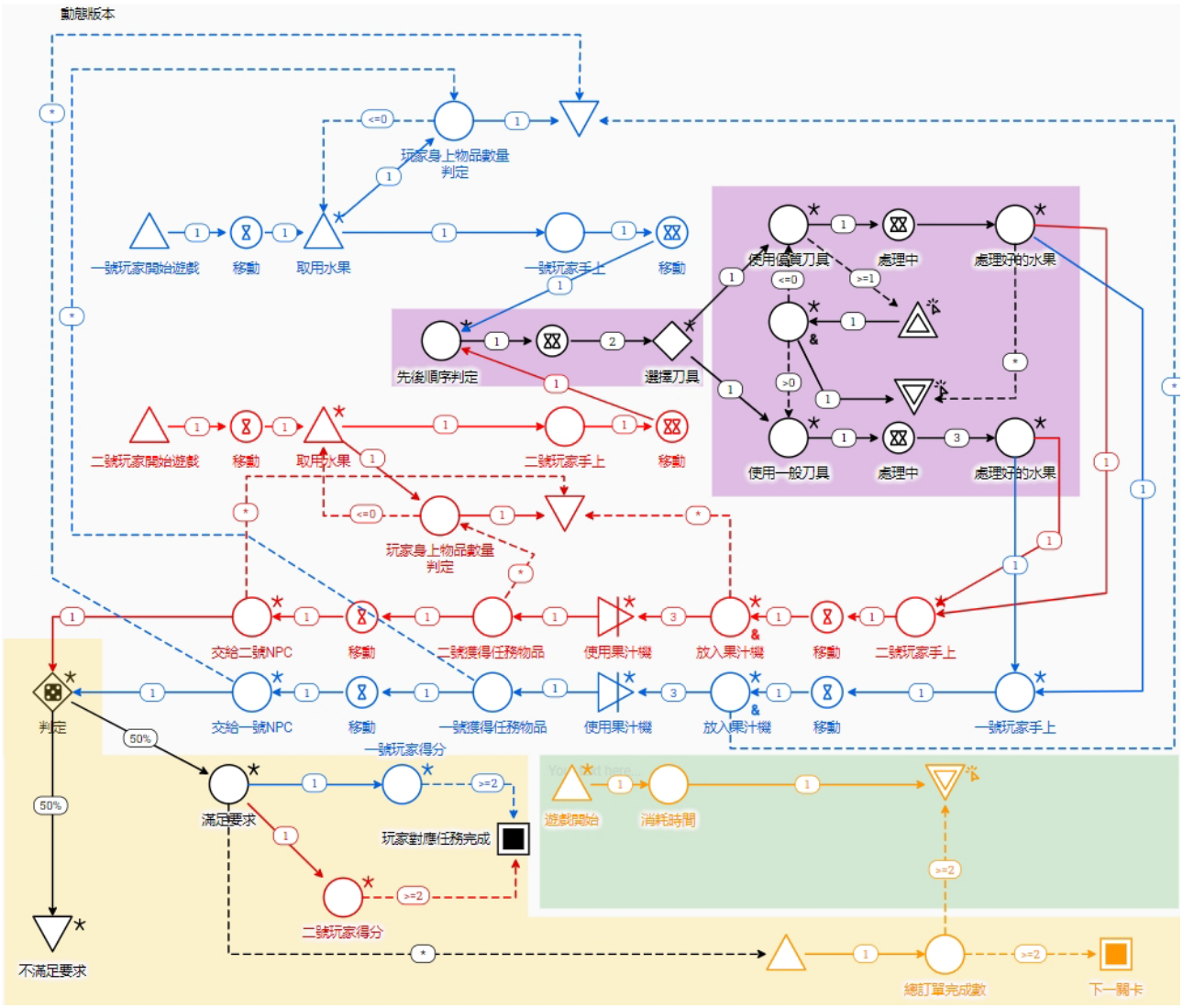
The blue and red lines represent the two players' resource flow and action orientation, respectively. Interactions are often indicated by triggers with a double outline, marking the material acquisition process with double hourglass symbols to signify that only one action can occur at a time throughout the duration. The number following the hourglass symbol indicates the number of turns to wait. Players must first use fruits as the starting point of their game actions. Since each player can only hold one fruit or task item at a time, they must finish processing a fruit and add it to the juicer before handling the next one. Therefore, Figure 2 includes many dashed lines and judgment conditions to trigger constraints. Each player manages their fruit materials independently during the processing, but tools like knives and the juicer are shared. If one player is using them, the other must wait.

Purple Area: Key Competitive Mechanism

The purple area is the critical zone for knowledge acquisition and strategic application. Players must identify the required primary colors for orders faster than their opponent to seize knives and juicers quickly. Securing these tools earlier can accelerate the completion of their tasks. In Machinations, we have presupposed that players will experience waiting times as both vie for the knives and juicers.

Figure 2.

The internal economy of "Rainbow Juice Shop" illustrated using Machinations.



Finally, we introduce the operation interface of "Rainbow Juice Shop," as shown in Figure 3.

(a). Juice Orders: Displays the fruit juice color required by the order.

(b). Order Transaction Area: After completing the fruit juice color blending, it is sent here for transactions.

(c). Countdown Timer: Shows the current countdown time.

(d). Fruit Area: The place where primary colored fruits (apples, grapes, and avocados) are placed.

(e). Juicer: A tool for crushing and blending various fruits.

(f). Juicer Warning Light: The juicer can only be used by one person at a time, and a red light indicates that it is currently in use.

(g). Chopping Board and Knife: Tools for cutting fruits.

(h). Chopping Board and Knife Warning Light: The knife can only be used by one person at a time, and a red light indicates that it is currently in use.

(i). Trash Bin: For disposing of fruit peels, pips, and other residues.

Figure 3.

"Rainbow Juice Shop" interface explanation.



3. Methodology

This study recruited 26 students from the Design College in central Taiwan, paired into two groups. The participants ages ranged from 20 to 25 (M = 22.12). The participants' educational backgrounds included undergraduate degrees (34.62%) and graduate degrees (65.38%), with 11 males and 15 females. The evaluation process involved participants first watching a game instruction video and reading a one-page guide. After ensuring all participants had no questions, the researcher proceeded with the game operation. The game was divided into four levels, with operation times of 3, 4, 5, and 6 minutes, totaling 18 minutes. After completing the game, participants completed a questionnaire that included basic information, the User Engagement Scale (UES), and an open-ended question about their gaming experience. Each participant received a reward of 100 NT dollars.

The User Engagement Scale (UES) is a tool for measuring self-assessed engagement used across various digital domains. This study's UES is based on the original framework by O'Brien and Toms (2010), with refinements from O'Brien and Toms (2013) and O'Brien et al. (2018), and further modifications for video game contexts, drawing from Bjørner (2021) and Wiebe et al. (2014). The final revised UES employs a 5-point Likert scale of 18 items across six constructs: focused attention, involvement, novelty, endurability, aesthetics, and perceived usability. The consistency of the items within each construct of the UES can be determined through reliability analysis (Table 1). The Cronbach's alpha values for the six constructs range from .744 to .952, indicating a satisfactory level of reliability for all items on the scale.

4. Results

The participants' evaluation results are shown in Table 1, from which the average scores and standard deviations for each construct can be deduced. The students playing "Rainbow Juice Shop" rated Focused Attention between "Agree" and "Strongly Agree" (4.487 points), showing that they were thoroughly engaged in the game. Among the feedback was a student's comment, "I found it really fun! It's rare for a game to give such a sense of achievement."

Perceived Usability was influenced by the inclusion of reverse-scored items, which could inevitably impact the accuracy of the responses, resulting in scores between "Agree" and "Neutral" (3.779 points). According to Chou's (2019) Octalysis Framework, gamification design aims to evoke epic meaning and accomplishment by seizing every possible opportunity for success and failure. Therefore, every action must be undertaken so as not totake advantage of every chance of obtaining the best experience. Feedback from students included remarks such as, "Looking at the colors for so long made me feel colorblind!" "There should be more differentiation between colors to make them easier to distinguish," and "It would be more user-friendly if there were hints after a certain number of errors."

Regarding the game's Aesthetics, the consensus fell between "Agree" and "Neutral" (3.872 points). Overall, participants rated aesthetics lower, which may be attributed to all 26 participants being design students and holding higher aesthetic standards. Feedback from students included suggestions like, "The main character could have more personality" and "The UI design could be made more distinct for better recognition."

The game's ability to maintain students' interest, scoring between "agree" and "strongly agree" (4.218 points) on Endurance, indicates students' acceptance of such serious games. The game's competitive nature ignites their fighting spirit, mirroring Peng et al. (2013) observations that competitive games offer more fun. Some students noted, "Competing with friends within a set time frame is thrilling, and the excitement increases with each level."

Regarding Novelty, the score fell between "agree" and "strongly agree" (4.231 points). This aligns with Chou's (2019) Octalysis Framework, where curiosity plays a significant role. A game incorporating many novel and unpredictable elements captures people's desire to continue seeking the final answers. Students' feedback also reflects this situation, "Level four is challenging, especially for players who are not very familiar with colors. It might be difficult, requiring step-by-step testing and memorization, but it also offers an opportunity to experience and learn."

The game's Involvement was scored between "agree" and "strongly agree" (4.615 points). Involvement is described as the desire to continue playing the game (Schønau-Fog & Bjørner, 2012). It is a prerequisite for other concepts such as engagement, flow, enjoyment, and immersion (Bjørner, 2021). The fun in serious games is complex, blending multiple factors that further engage players and encourage continued play. One student commented, "The thinking time is short, and sometimes I can't figure out how to mix colors. Seeing the opponent continuously completing tasks can be quite frustrating."

Table 1.

Reliability Analysis and Descriptive Statistics

Scale	No Items	Mean	S.D.	Crohbach's α
Focused attention	3	4.487	0.834	.952
Perceived usability	4	3.779	0.642	.791
Aesthetics	3	3.872	0.998	.938
Endurability	3	4.218	0.822	.822
Novelty	3	4.231	0.842	.798
Involvement	2	4.615	0.397	.744

When designing serious games, keeping the gameplay mechanics from being overshadowed by the correct educational message is crucial. This highlights the importance of considering the players and the environment when designing and evaluating serious games. Some students have provided feedback such as, "There might be a need to explain what RGB mixing colors are beforehand." and "I'm a bit colorblind; I can't distinguish well between green and light green; how am I supposed to mix colors?" These responses tell us that more relevant information needs to be provided to facilitate smoother gameplay.

5. Conclusions

From the UES scale and open-ended questions, it was found that participants in the "Rainbow Juice Shop" game had a very high level of Focused Attention. The game's pace and music create a relaxed atmosphere, and although it is competitive, the majority found it engaging and challenging. While the author feels that the Aesthetics aspect was quite good, this construct scored the second lowest, possibly because the participants were design school students with stricter aesthetic standards. Therefore, besides making minor adjustments afterward, involving students from non-design faculties in the evaluation could also be beneficial. The most successful aspect of the game seems to be Involvement, as evidenced by the tense atmosphere revealed by participants during the game, with some even adopting competitive postures and strategies.

In conclusion, "Rainbow Juice Shop" still requires improvements after evaluation, including adding hints to facilitate learning for players. Furthermore, this serious game fully presents the educational purpose, allowing participants to learn in an environment where education is blended with fun. Therefore, under effectiveness and stability, "Rainbow Juice Shop" can be expanded as a tool for future research on involvement and learning behaviors, showing great potential. Future studies will also include physiological instruments (such as eye trackers) to supplement the analysis of learning behaviors in serious games.

Acknowledgements

The research was supported by the National Science and Technology Council of Taiwan, and the Contract Numbers are MSTC 112-2410-H-224-026.

References

Adams, E., & Dormans, J. (2012). *Game mechanics: advanced game design*. New Riders.

Bjørner, T. (2021, September 9-11). *How can a serious game be designed to provide engagement with and awareness of the plastic crisis as part of UN's SDGs*. The Conference on Information Technology for Social Good (pp. 157-162).

Burke, J. W., McNeill, M. D. J., Charles, D. K., Morrow, P. J., Crosbie, J. H., & McDonough, S. M. (2009). Optimising engagement for stroke rehabilitation using serious games. *The Visual Computer*, *25*, 1085-1099. https://doi.org/10.1007/s00371-009-0387-4

Cagiltay, N. E., Ozcelik, E., & Ozcelik, N. S. (2015). The effect of competition on learning in games. *Computers & Education*, *87*, 35-41. https://doi.org/10.1016/j.compedu.2015.04.001

Chen, C. H., & Chiu, C. H. (2016). Employing intergroup competition in multitouch design-based learning to foster student engagement, learning achievement, and creativity. *Computers & Education, 103*, 99–113. https://doi.org/10.1016/j.compedu.2016.09.007

Chen, C. H., Law, V., & Chen, W. Y. (2018). The effects of peer competition-based science learning game on secondary students' performance, achievement goals, and perceived ability. *Interactive Learning Environments*, *26*(2), 235-244. https://doi.org/10.1080/10494820.2017.1300776

Chen, S. Y., & Chang, Y. M. (2020). The impacts of real competition and virtual competition in digital game-based learning. *Computers in Human Behavior*, *104*, 106171. https://doi.org/10.1016/j.chb.2019.106171

Cheng, Y. M., Lou, S. J., Kuo, S. H., & Shih, R. C. (2013). Investigating elementary school students' technology acceptance by applying digital game-based learning to environmental education. *Australasian Journal of Educational Technology*, *29*(1). 96-110. https://doi.org/10.14742/ajet.65

Chou, Y. K. (2019). *Actionable gamification: Beyond points, badges, and leaderboards*. Packt Publishing Ltd.

Code Monkey(2022) https://unitycodemonkey.com/about.php

den Haan, R. J., & van der Voort, M. C. (2018). On evaluating social learning outcomes of serious games to collaboratively address sustainability problems: A literature review. *Sustainability*, *10*(12), 4529. https://doi.org/10.3390/su10124529

Deutsch, M. (1949). A theory of co-operation and competition. *Human Relations*, 2(2), 129–152. http://doi.org/10.1177/001872674900200204.

Ebner, M., & Holzinger, A. (2007). Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & education*, *49*(3), 873-890. https://doi.org/10.1016/j.compedu.2005.11.026

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on psychological science*, *8*(6), 634-650. http://doi.org/10.1177/1745691613504114.

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of computer assisted learning*, 29(3), 207-219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Hwang, G. J., & Chang, S. C. (2016). Effects of a peer competition-based mobile learning approach on students' affective domain exhibition in social studies courses. *British Journal of Educational Technology, 47*(6), 1217–1231. https://doi.org/10.1111/bjet.12303

Klimmt, C. (2009). Serious games and social change: Why they (should) work. In Serious Games (pp. 270-292). Routledge.

Laamarti, F., Eid, M., & El Saddik, A. (2014). An overview of serious games. *International Journal of Computer Games Technology, 2014*, 358152. https://doi.org/10.1155/2014/358152

Lin, K. C., Wu, K. C., & Wang, T. B. (2010, Auguest 4-6) *Using competitive digital game-based learning to improve learning motivation*. 2010 IET International Conference on Frontier Computing. Theory, Technologies and Applications (pp. 128-132). https://doi.org/10.1049/cp.2010.0549

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies, 112*, 28-39. ttps://doi.org/10.1016/j.ijhcs.2018.01.004

O'Brien, H. L., & Toms, E. G. (2013). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing & Management, 49*(5), 1092-1107. https://doi.org/10.1016/j.ipm.2012.08.005

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology, 61*(1), 50-69. https://doi.org/10.1002/asi.21229

Paraskeva, F., Mysirlaki, S., & Papagianni, A. (2010). Multiplayer online games as educational tools: Facing new challenges in learning. *Computers & Education, 54*(2), 498-505. ttps://doi.org/10.1016/j.compedu.2009.09.001

Pareto, L., Haake, M., Lindström, P., Sjödén, B., & Gulz, A. (2012). A teachable-agent-based game affording collaboration and competition: Evaluating math comprehension and motivation. *Educational Technology Research and Development, 60*, 723–751. https://doi.org/10.1007/s11423-012-9246-5

Peng, W., Crouse, J. C., & Lin, J. H. (2013). Using active video games for physical activity promotion: a systematic review of the current state of research. *Health education & behavior, 40*(2), 171-192. https://doi.org/10.1177/1090198112444956

Plewe, C., & Fürsich, E. (2018). Are newsgames better journalism? Empathy, information and representation in games on refugees and migrants. *Journalism Studies, 19*(16), 2470-2487. https://doi.org/10.1080/1461670X.2017.1351884

Schønau-Fog, H., & Bjørner, T. (2012). "Sure, I would like to continue" a method for mapping the experience of engagement in video games. *Bulletin of Science, Technology & Society, 32*(5), 405-412. https://doi.org/10.1177/0270467612469068

Song, H., Kim, J., Tenzek, K. E., & Lee, K. M. (2013). The effects of competition and competitiveness upon intrinsic motivation in exergames. *Computers in Human Behavior, 29*(4), 1702-1708. https://doi.org/10.1016/j.chb.2013.01.042

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE), 3*(3), 1-24. https://doi.org/10.1145/1077246.1077253

Tekinbas, K. S., & Zimmerman, E. (2003). *Rules of play: Game design fundamentals*. MIT press.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review, 41*(2), 16-18.

Van Eck, R., & Dempsey, J. V. (2002). The Effect of competition and contextualized advisement on the transfer of mathematics skills in a computer-based instructional simulation game. *Educational Technology Research & Development, 50*(3), 23-41. https://doi.org/10.1007/BF02505023

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*, 123-132. https://doi.org/10.1016/j.chb.2013.12.001

基於概念圖的元宇宙探究式學習模式在地理課程學習成效之影響

Effects a Concept mapping-based strategy on students' Metaverse Inquiry-based learning effectiveness of a Geography curriculum

曾詩倩[1*]，朱蕙君[2]
[1, 2]東吳大學資訊管理學系, 台灣 台北
[*] jane900411@gmail.com

【摘要】 地理教育目的是培養學生的環境素養，然而，多維度的地理知識對學生來說具有挑戰性，傳統教學方法難以有效建構知識網。因此本研究提出基於概念圖的元宇宙探究式學習模式，幫助學生提升學習成效和環境素養。為了調查這種方法對學生的影響，在地理課程進行實驗，實驗組採用基於概念圖的元宇宙探究式學習模式，控制組則採用一般元宇宙探究式學習模式。評量標準包括學生的學業成績、學習動機和環境素養等，並進一步進行深度訪談，了解學生的學習歷程。

【关键词】 元宇宙；概念圖；探究式學習；環境素養

Abstract: *The aim of geography education is to foster students' environmental literacy. However, the multifaceted nature of geography knowledge poses challenges, and conventional teaching methods may not effectively establish knowledge frameworks. Thus, this study proposes a concept mapping-based strategy on metaverse inquiry-based learning to enhance students' learning efficacy and environmental literacy. An experiment was conducted in a geography course to assess the impact of a concept mapping-based strategy on metaverse inquiry-based learning model compared to a control group using traditional metaverse inquiry-based learning. Evaluation criteria encompassed academic performance, learning motivation and environmental literacy. In-depth interviews were conducted to glean insights into students' learning experiences.*

Keywords: Metaverse, Concept Map, Inquiry-Based Learning, environmental literacy

1. 緒論

隨著時代變遷，改變傳統的教學模式，學習不再以教師為主導，產生更多以學生為主導者的教學模式，例如探究式學習、專題導向學習和自主學習等教學模式，為學生提供更適合其學習的場域，實現因材施教。其中在探究式學習，老師的角色變成引導者，不再直接傳授知識，而是引導學習者在自主觀察、實驗、討論、反思中，嘗試彙整所學並建構知識網(Spronken-Smith et al., 2012)，進而提升學習者的學習成效和拓展高階思維能力。過往研究也應用於地理課程中，讓學生透過老師引導的課題，蒐集資料，彙整知識並實驗討論地理脈絡，提升學生的理解和學習成效(Brumann et al., 2019)。

近年來，資訊科技在教育領域蓬勃發展，元宇宙也成為教學應用的一大選擇，有別於VR/AR多了共享的特質，讓學習者可以與他人互動學習，創造更多元的學習機會。過往已應用於情境式英語(Guo & Gao, 2022)、考古(Getchell et al., 2010)、肺癌手術訓練(Koo, 2021)等研究，也有研究將實作科學探究式學習導入元宇宙中，提供在身歷其境、高互動和可協作的學習環境，學習者在假設、提問和測試等體驗，提升成效(Chen et al., 2023)。

在地理教育中，地理知識的複雜度，使有效建構和整合空間知識和概念知識成為提升成績的關鍵因素。然而，傳統教學常把概念知識和空間知識拆開學習，使學生難以建構有效的知識網，無法將知識融會貫通，更難以促進學生的環境素養。因此，需導入適當的教學策略和框架引導學習(Barzilai & Blau, 2014)。

在教學領域中，概念圖幫助學習者複習和系統化的組織內容，被視為一種有效的知識建構工具(Hwang, Lin, et al., 2023)，被應用於探究式學習中，增強學生的認知、成績，作為協作環境中的共識，對活動學習具有積極影響(Gijlers & de Jong, 2013)。

儘管如此，導入概念圖到元宇宙探究式學習模式的教育研究仍有限。因此，本研究旨在基於概念圖的元宇宙探究式學習模式中，透過概念圖引導學生的知識建構，並利用元宇宙增加沉浸感提升動機，進而提高學生的學習成效、學習動機和環境素養。

為了評估基於概念圖的元宇宙探究式學習模式對學生學習成效的影響，本研究提出以下研究問題：

- 與一般元宇宙探究式學習模式相比，基於概念圖的元宇宙探究式學習模式的學習方法能否提升學生的學習成就？
- 與一般元宇宙探究式學習模式相比，基於概念圖的元宇宙探究式學習模式的學習方法能否提升學生的學習動機？
- 與一般元宇宙探究式學習模式相比，基於概念圖的元宇宙探究式學習模式的學習方法對學生的環境素養差異？

2. 文獻探討

2.1. 探究式學習

探究式學習被視為一種提升學習者學習成效的教學方法，以學生主導學習，老師作為引導的角色，讓學生自主針對課題進行觀察、提問、實驗、討論和思考的流程，彙整其所學並找出解決方法、建構一套知識網，進而提升成效(Spronken-Smith et al., 2012)。

過往研究表明，探究式學習讓學生在地理課程中，有效建構知識網絡、拓展思維和主動學習，更可以提升學生的環境素養和態度(Brumann et al., 2019)。Drápela (2022)在地質遺址教學中導入探究式學習，透過逐步實施和評估學生的探究學習計畫，進而有效識別地質遺址。

2.2. 元宇宙在教育領域的應用

元宇宙被視為一種“電腦生成的宇宙”，它使人們不單單是觀察而已，更沉浸於體驗中，因為其共享的特質，創造新的互動學習場域(Hwang, Tu, et al., 2023)。給予學生更具吸引力的互動性、沉浸感和動手操作的體驗(Park & Kim, 2022)，並培養團隊合作和溝通技巧等高階思維能力。

在教育領域中，元宇宙在過往研究也被用於情境式英語(Guo & Gao, 2022)、肺癌手術訓練(Koo, 2021)和實作科學探究式學習(Chen et al., 2023)，提升學生的成效、高階思維能力。此外，學生對於元宇宙導入教學中抱持正向態度(Hwang, Tu, et al., 2023)。整體而言，元宇宙的應用展現出教育不同的前景，為學習環境帶來別於過往的共享性和互動性。

2.3. 概念圖

概念圖是一種有效並被廣泛採用的學習策略，將知識視覺化並呈現，利用節點表示概念，節點間的鏈結表示知識概念間的關係(Novak & Gowin, 1984)。可以使抽象概念具體化，提高正確的知識建構，進而提升學習成績、高階思維能力等(Hwang, Lin, et al., 2023)。在增強學生對複雜多面向的知識理解也具有巨大潛力，尤其是社會科學教育，將概念圖導入遊戲式學習，可以幫助學生整合應用空間知識和概念知識，也可更確定學習方向和主動學習(Hwang, Lin, et

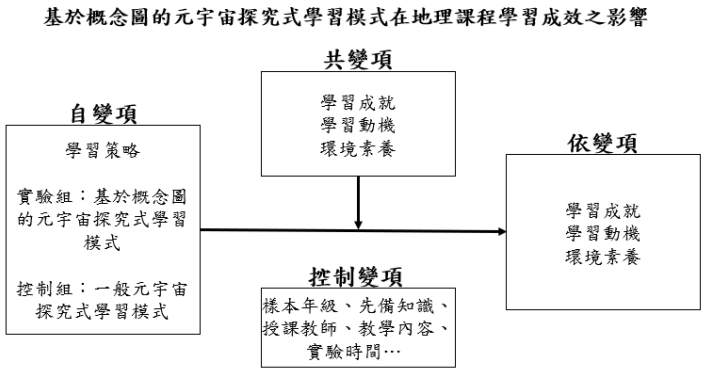
al., 2023)。概念圖對協作探究式學習環境中的共識、活動和學習產生積極影響(Gijlers & de Jong, 2013)。

3. 研究方法

3.1. 研究架構

圖 1.

研究架構圖

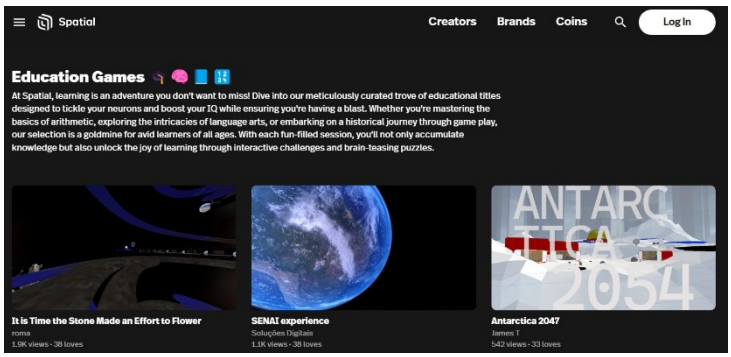


3.2. 學習模式設計-基於概念圖的元宇宙探究式學習模式

本研究的實驗組(EG) 採用基於概念圖的元宇宙探究式學習模式，利用 Spatial 作為教學平台，Spatial 是一款用全像投影（holographic）自由創建屬於自己的虛擬世界之平台，學習者透過行動裝置，登入帳號後即可開始使用，如圖 2 所示。

圖 2.

Spatial 系統介面



為讓學生更好操作並探索學習，老師在事先設好學習範圍的概念圖等教材內容，並導入元宇宙中。課堂上，學生透過小組為單位進入到元宇宙學習(如圖 3 所示)。

圖 3.

元宇宙學習參考頁面



透過隨機分組的方式，至多四人，以小組的形式共同在一個元宇宙，協力探索元宇宙空間內的人事物並蒐集知識，老師會引導學生團隊分工去協力完成任務。元宇宙中的課題也參考 Hwang, Lin, et al. (2023)和 Gijlers and de Jong (2013)在概念圖和探究式學習的設計改良，學生透過老師匯入到元宇宙中的概念圖(如圖 4 所示)釐清知識網絡，透過蒐集的知識，小組討論完成擴增概念圖，並共同完成元宇宙中的問答，如概念圖填空(如圖 5 所示)。概念圖的填空也會基於老師的概念圖進行挖空作為考題。學生在進行回答問題時，有任何回答錯誤或不理解，系統也會即時給予提示，協助學生主動尋找相關知識點補充學習。

以高中一年級課綱之東亞文化圈的日本和台灣地理做為課程主題，幫助從台灣、日本為出發點，探討兩者在地形、氣候、洋流、水資源等自然資源、人文和能源運用差異。再延續到近年頻頻被討論的時事議題，如核能、核廢水、工業污染、水資源不足問題。進而去了解兩國在地理上的差異，政策的規劃等，建構出更詳細的概念圖和知識脈絡，如圖 4 所示。議題舉例如核能，引導學生探討核能的需求、效能、過往資訊、其對環境的影響、排放核廢水對環境和人的影響、為何世界反彈聲音大，最後反觀台灣的核能發電，會有甚麼影響和未來發展。

圖 4.

課程內容的概念圖教材參考頁面

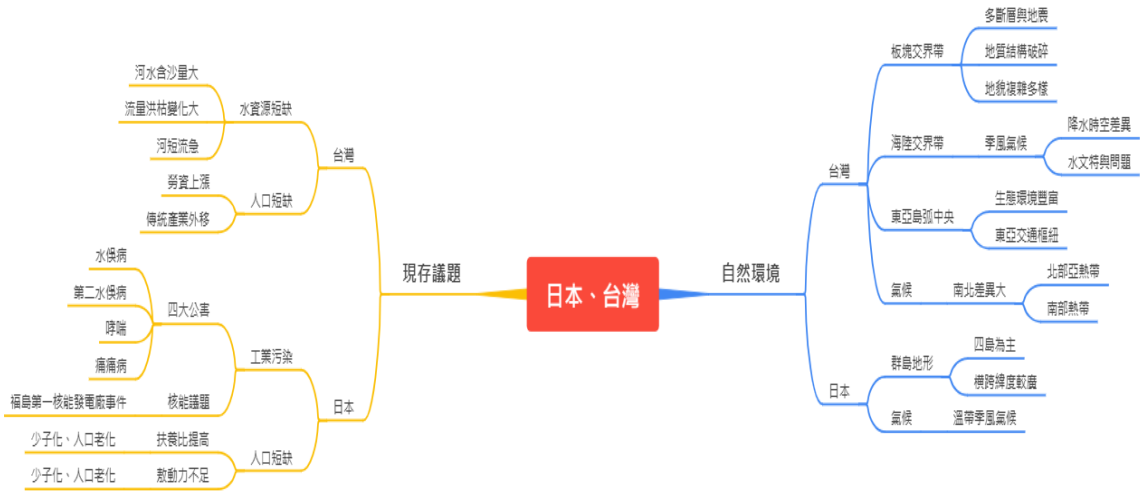
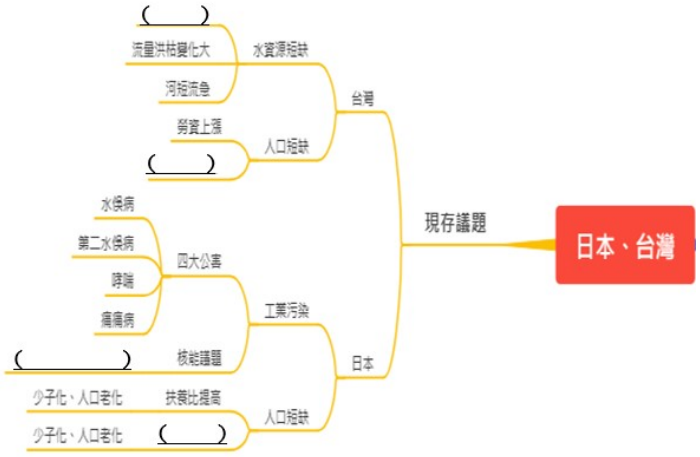


圖 5.

課程內容的概念圖填空參考頁面



4. 研究設計

4.1. 實驗對象

某高中一年級的數個班級，以一年級東亞文化圈和台灣地理為課程內容，採準實驗方法，實驗組(EG) 採用基於概念圖的元宇宙探究式學習模式，控制組(CG) 採用一般元宇宙探究式學習模式。

4.2. 實驗流程

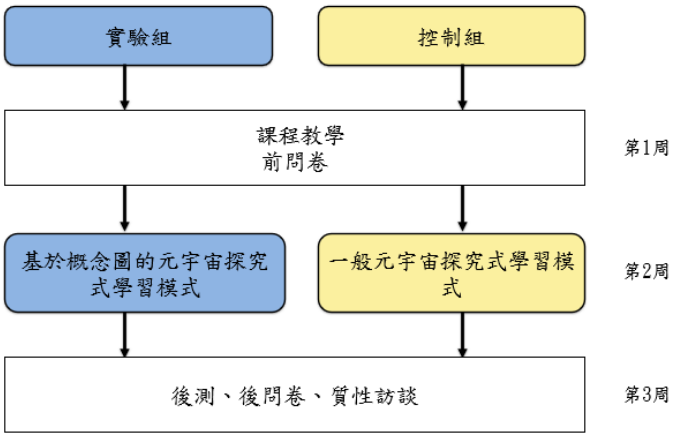
第一周，授課教師會給予該單元的基礎知識概念、架構等內容，並教導學習如何運用或使用概念圖和如何在元宇宙的學習環境進行操作。兩組同學都會進行學習前測和前問卷。

第二周，實驗組採用基於概念圖的元宇宙探究式學習模式進行學習和測驗，控制組則採用一般的元宇宙探究式學習模式。 兩組的差異就在於元宇宙的學習環境中是否有概念圖輔助學習，兩組學習的內容、測試皆為相同，透過老師引導小組合作完成學習。

第三周，兩組同學都會進行學習後測和後問卷調查，此外各小組會進行 10 分鐘的焦點小組訪談。

圖 6.

實驗流程圖



4.3. 研究工具

本研究所採用的研究工具為學習成績測驗、問卷調查和質性訪談。學業成績測驗由前測和後測組成，問卷由學習動機和環境素養問卷組成。

前測採用該班的期中考地理成績，後測則由兩位以上專家共同編製，由單選題、多選題和題組等方式組成，用於了解學生個人學習前後的差異，和不同學習模式的差異。

學習動機問卷改編自 Wang and Chen (2010)的研究，旨在調查學習者在學習活動前後學習語言的動機是否增強，由內在動機和外在動機兩個向度組成，參考問卷的 Cronbach α 係數分別為 0.76 和 0.80，各三個題項，共 6 題，採用李克特 5 點量表。

內在動機向度的一個問題是「在這門課程中，我更喜歡真正挑戰我的課程教材，這樣我就可以學習新東西。」；外在動機向度的一個問題是「我相信我在這門課上會取得優異的成績」。

環境素養問卷改編自 Liang et al. (2018)的問卷，為了解學習者活動前後環境素養的變化，分為「環境意識和敏感性」和「環境行為技能」兩個向度，各七題項，共 14 題，採用李克特 5 點量表。

「環境意識和敏感性」參考問卷的 Cronbach α 係數為 0.87，問題如「我可以與其他學生討論環境問題，以澄清因果關係」和「地球的資源是有限的，所以我會珍惜資源」等；而「環境行為技能」參考問卷的 Cronbach α 係數為 0.90，問題如「我可以反思自己對社會和環境影響的行為」和「我有責任影響我周圍的人，提高環保意識」等。

質性訪談的七個開放式問題改編自 Hwang et al. (2009)，每組 10 分鐘焦點小組訪談，會根據教學模式、小組和錄音時間進行錄音和轉錄。問題如「你認為基於概念圖的元宇宙探究式學習模式的優點是什麼？」以及「你從這種方法中學到最多的是什麼？」。訪談內容由兩位經驗豐富的老師進行分析編碼。

5. 預期結果

本研究期望學生透過基於概念圖的元宇宙探究式學習方法，為學生提供更互動和吸引力的體驗，激發學生對學習的興趣，從而幫助他們更有效地組織知識，提升學習效果和環境素養，給予不同的學習體驗，讓學生認知到生活中的環境議題，並促進學生成為一位具有責任感的公民。

參考文獻

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65-79. https://doi.org/10.1016/j.compedu.2013.08.003

Brumann, S., Ohl, U., & Schackert, C. (2019). Researching climate change in their own backyard—inquiry-based learning as a promising approach for senior class students. *Climate Change and the Role of Education*, 71-86. https://doi.org/10.1007/978-3-030-32898-6_5

Chen, X., Zou, D., Xie, H., & Wang, F. L. (2023). Metaverse in education: contributors, cooperations, and research themes. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/TLT.2023.3277952

Drápela, E. (2022). Assessing the educational potential of geosites: introducing a method using inquiry-based learning. *Resources*, *11*(11), 101. https://doi.org/10.3390/resources11110101

Getchell, K., Oliver, I., Miller, A., & Allison, C. (2010). Metaverses as a platform for game based learning. 2010 24th IEEE International Conference on Advanced Information Networking and Applications,

Gijlers, H., & de Jong, T. (2013). Using concept maps to facilitate collaborative simulation-based inquiry learning. *Journal of the learning sciences*, *22*(3), 340-374. https://doi.org/10.1080/10508406.2012.748664

Guo, H., & Gao, W. (2022). Metaverse-powered experiential situational English-teaching design: an emotion-based analysis method. *Frontiers in Psychology*, *13*, 859159. https://doi.org/10.3389/fpsyg.2022.859159

Hwang, G.-J., Lin, Y.-C., & Lin, H.-C. (2023). Associating spatial knowledge with concept maps to facilitate learning in digital gaming contexts. *Educational technology research and development*, *71*(6), 2221-2241. https://doi.org/10.1007/s11423-023-10274-5

Hwang, G.-J., Tu, Y.-F., & Chu, H.-C. (2023). Conceptions of the metaverse in higher education: A draw-a-picture analysis and surveys to investigate the perceptions of students with different motivation levels. *Computers & Education*, *203*, 104868. https://doi.org/10.1016/j.compedu.2023.104868

Koo, H. (2021). Training in lung cancer surgery through the metaverse, including extended reality, in the smart operating room of Seoul National University Bundang Hospital, Korea. *Journal of educational evaluation for health professions*, *18*. https://doi.org/10.3352/jeehp.2021.18.33

Liang, S.-W., Fang, W.-T., Yeh, S.-C., Liu, S.-Y., Tsai, H.-M., Chou, J.-Y., & Ng, E. (2018). A nationwide survey evaluating the environmental literacy of undergraduate students in Taiwan. *Sustainability*, *10*(6), 1730. https://doi.org/10.3390/su10061730

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. cambridge University press.

Park, S., & Kim, S. (2022). Identifying world types to deliver gameful experiences for sustainable learning in the metaverse. *Sustainability*, *14*(3), 1361. https://doi.org/10.3390/su14031361

Spronken-Smith, R., Walker, R., Batchelor, J., O'Steen, B., & Angelo, T. (2012). Evaluating student perceptions of learning processes and intended learning outcomes under inquiry approaches. *Assessment & Evaluation in Higher Education*, *37*(1), 57-72. https://doi.org/10.1080/02602938.2010.496531

Wang, L. C., & Chen, M. P. (2010). The effects of game strategy and preference-matching on flow experience and programming performance in game-based learning. *Innovations in Education and Teaching International*, *47*(1), 39-52. https://doi.org/10.1080/14703290903525838

A Proposal on Designing a Scaffolded Flipped Classroom Environment to Improve Students' Engagement: a Design-based Research

Hantong, Zhao[1*], Qiyun, Wang[2]
[1 2] National Institute of Education, Nanyang Technological University, Singapore
[*] NIE23.ZH1034@e.ntu.edu.sg

Abstract: *This proposal focuses on integrating scaffolded learning in the pre-class stage of the flipped classroom model to enhance students' engagement. Existing literature underscores the benefits of the flipped classroom model, yet a gap exists in understanding how scaffolded learning, especially in a blended learning environment, can specifically improve the pre-class stage, which is crucial for overall learning outcomes of flipped classroom model. To fill the gap, a design-based research will be planned to implement in a Singaporean primary school with grade six students and optimizes this design through 2-3 cycles of iteration. This design aims to build a blended flipped learning environment,where pre-class activities are completed asynchronously online and scaffoldings are provided to increase student engagement. This research aims to offer insights for researchers and educational practitioners, contributing to effective strategies for optimizing the pre-class stage learning performance, thus enhancing effectiveness of the flipped classroom model.*

Keywords: Flipped Classroom, Scaffolding, Engagement, Blended Learning

1. Introduction

The flipped classroom (FC) approach is a technology-supported pedagogical model consisting of two essential elements: (1) computer-based individual instruction via video lectures outside the classroom and (2) interactive group learning activities conducted inside the classroom (Bishop & Verleger, 2013). Although as a student-centered learning method, the FC model can better exert students' self-efficacy and help students achieve better learning results, some studies also have pointed out that there are still some challenges in the current FC format, such as the low engagement of learners in pre-class stage because of the extra burden of pre-class tasks and prompt guidance students need when they study at home (Akçayır & Akçayır, 2018), and there is also a lack of scientific ways to clearly understand the effectiveness of pre-class learning (Larson & Linnell, 2023).

Additionally, although the FC model has been proven to positively impact learners' learning satisfaction and academic performance in various educational scenarios, including foreign language learning domain, studies based on foreign language learning mainly focus on English as Foreign Language (EFL) area.Wang (2018) proposed that researchers should expand the advantages of the FC model, exploring beyond its focus solely on English learning.. Additionally, as a language with a rapid increase in learners (Wright, 2018), Chinese learning deserves more attention,which deserve more attention. Therefore, this study will be conducted in a primary school in Singapore where students are learning Chinese as a foreign language (CFL), and this research aims to answer the following research questions: 1) How can scaffolding improve learners' engagement at the pre-class stage of FC approach ? 2) What are the students' attitude and perception towards scaffolded asynchronous online learning environment in FC approach ?

2. Conceptual Framework

As discussed above, the critical reason hindering the maximization of the FC model's advantages is the low engagement and learning performance in pre-class activities (Morcom, 2015). This is primarily due to pre-class tasks exceeding learners' Zone of Proximal Development (ZPD), leading to frustration and anxiety without timely support

(Morcom, 2015). Thus, optimizing pre-class design for effective guidance within the ZPD is crucial for enhancing engagement (Athanases & Oliveira, 2014). Vygotsky's sociocultural theory emphasizes social interaction and emotions in learning, suggesting the involvement of those with greater experience to guide learners in their ZPD (Morcom, 2015). Scaffolding, providing temporary support within the ZPD, is essential for learners' development , and in online learning, technology-enabled scaffolding in aids learners in completing challenging tasks, boosting motivation and confidence, and this support encourages increased engagement and success in learning activities (Athanases & Oliveira, 2014).

2.1. Scaffolding and Engagement

Wood, Bruner, and Ross (1976) defined scaffolding as a process that involves offering assistance to learners in sustaining their direction.Technology-enhanced scaffolding is vital for promoting students' success. Belland, Kim, and Hannafin (2013) describe technology-enhanced scaffolding as a software program employing strategies like posing questions, providing modeling, and offering tools for data manipulation to enrich students' comprehension and metacognitive abilities,which is vital for supporting students in online learning, especially asynchronous learning.

Four types of technology-enhanced scaffolding were introduced by Hannafin, Kim, and Hannafin (1999) and Kim and Hannafin (2011): procedural, conceptual, metacognitive, and strategic scaffolding. Procedural scaffolding, including directive, anticipatory, and resource scaffolding, helps students navigate course features and tools, reducing negative emotions like anxiety (Stavredes, 2011). Conceptual scaffolding aids in grasping key concepts by assessing understanding, identifying misconceptions, and clarifying core ideas (Stavredes, 2011). Strategic scaffolding addresses issues encountered during the learning process, guiding students in developing cognitive skills through metacognitive scaffolding (Hannafin et al., 2013).

In the blended learning context, engagement is categorized into behavioral, emotional, and cognitive aspects (Fredricks et al., 2004). In language learning, behavioral engagement involves participation quality, while cognitive engagement pertains to attention, reflection, and self-regulation, and affective engagement includes positive and negative emotions, while social engagement involves cooperation (Svalberg, 2009). These engagement aspects interact dynamically (Hewett et al., 2019), where scaffolding can influence one aspect and potentially impact others.

3. Research Design

3.1. Context and participants

The FC learning environment is the research focus. Hwang et al. (2015) divide it into three stages: pre-class, in-class, and post-class. Scaffolding in the pre-class stage is crucial for overall FC model success (Lee & Choi, 2018).

A primary school in Singapore will be selected, with a Chinese classroom chosen. Research in K-12 on the FC model and Chinese as a second language (CSL) learning is limited (Lo & Hew, 2017). Language learning, especially listening and speaking, requires practice beyond the classroom. The pre-class stage in a flipped classroom can enhance these skills, saving class time for application and improving learning outcomes.

Singapore's multi-ethnic environment emphasizes Chinese learning, yet daily life predominantly uses English. FC aims to enhance independent learning, contextual understanding, and engagement in Chinese classes (Lo & Hew, 2017).

The selected primary school is experienced in FC, preventing bias from participant unfamiliarity. Participants are students in Grade 6, and are 30 participants in total. These students are randomly selected from all students in Grade 6.All the participants are proficient in using mobile devices and online learning systems.

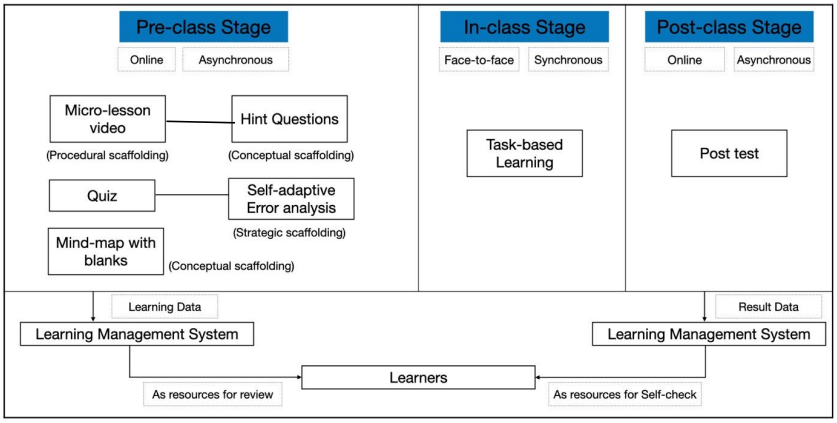
The study focuses on minor curriculum units, each lasting 5 days. Students engage in asynchronous pre-class tasks for the first three days, synchronous in-class learning on the fourth day, and asynchronous post-tests on the fifth day

3.2. Method (Design-based Research)

Plomp (2013) defined design-based research (DBR) as a suitable research approach for crafting solutions to intricate issues in educational practice or for the validation and development of theories related to learning processes and learning environments. This research will choose the DBR method to guide the design and development of the scaffolding supported FC environment.

This DBR aims to develop a flipped classroom with scaffolding to enhance learning engagement. Learners complete online pre-class activities asynchronously, divided into three 20-minute slices before face-to-face class. Data collection will use a Learning Management System (LMS), with each student and teacher having separate accounts. Learning for each unit spans one week, with teachers uploading materials and exercises on the LMS.The detailed pre-class learning activites include: A 5-minute micro-lesson video introducing unit objectives and methods, with guiding and quick-answer questions as procedural scaffolds, and hint questions for conceptual scaffolds. A self-adaptive error analysis quiz provides strategic scaffolding, with wrong answers recorded for review. A mind-map with blanks aids conceptual understanding. This learning process in this scaffolded flipped classroom environment will be shown in the following Figure 1 :

Figure 1. The learning process in the scaffolded FC environment



4. Evaluation

This proposal is designed to test if integrating scaffolding in the pre-class stage in a FC learning environment can 1) improve learners' engagement at the pre-class stage, and 2) the students' attitude and perception towards this design are will also be explored. To explore these research questions, both qualitative and quantitative data will be collected and analyzed. The data of students' learning behavior data, including the duration time, numbers of posts and the correct rate of quiz will be collected and analyzed to explore the impact of the scaffolding in this design on the learning engagement and performance of the pre-class stage. A semi-structure interview and survey filling will be conducted for students at the end of each cycle in this DBR. The data analysis of the interview and questionnaire will be used to explore the effectiveness of this design and the views of learners on this design.

5. Concluding Remarks

While our paper presents a research proposal, it's important to note that we're currently at the proposal stage. We've outlined the problem, research objectives, and proposed solutions but haven't conducted detailed empirical research or practical implementation. The aim is to garner interest from academia and practitioners, laying groundwork for future exploration.Our future plan involves empirical research to validate and assess the effectiveness of our framework. Further data collection and analysis are crucial to confirm real-world impact. In conclusion, we hope this proposal will stimulate further research and discussions, offering new solutions for flipped classrooms and learning challenges.

References

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers Education,* 126, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

Athanases, S. Z., & de Oliveira, L. C. (2014). Scaffolding versus routine support for Latina/O Youth in an urban school. *Journal of Literacy Research*, *46*(2), 263–299. https://doi.org/10.1177/1086296x14535328

Bishop, J., & Verleger, M. (2013). The flipped classroom: A survey of the research. 2013 ASEE Annual Conference Exposition Proceedings. https://doi.org/10.18260/1-2--22585

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74* (1), 59–109. https://doi.org/10.3102/00346543074001059

Hwang, G.-J., Lai, C.-L., & Wang, S.-Y. (2015). Seamless flipped learning: A mobile technology-enhanced flipped classroom with Effective Learning Strategies. *Journal of Computers in Education*, *2*(4), 449–473. https://doi.org/10.1007/s40692-015-0043-0

Hannafin, M., Land, S., & Oliver, K. (2013: 115140). Open learning environments: Foundations, methods, and models. In C. M. Reigeluth (Ed.), *Instructional-design theories and models* (pp. 115-140). New York: Routledge.(2013). https://doi.org/10.4324/9781410603784

Hewett, S., Becker, K., & Bish, A. (2019). Blended workplace learning: The value of human interaction. *Education Training*, *61*(1), 2–16. https://doi.org/10.1108/et-01-2017-0004

Lo, C. K., & Hew, K. F. (2017). A critical review of flipped classroom challenges in K-12 education: Possible solutions and recommendations for future research. *Research and Practice in Technology Enhanced Learning*, *12*(1). https://doi.org/10.1186/s41039-016-0044-2

Larson, M. P., & Linnell, J. (2023). Are students coming to class prepared? the importance of pre-class learning in a flipped classroom. *Issues in Accounting Education, 38*(3), 183–205. https://doi.org/10.2308/issues-2021-099

Lee, J., & Choi, H. (2018). Rethinking the flipped learning Pre-class: Its influence on the success of flipped learning and related factors. *British Journal of Educational Technology, 50*(2), 934–945. https://doi.org/10.1111/bjet.12618

Morcom, V. (2015). Scaffolding social and emotional learning within ‘shared affective spaces’ to reduce bullying: A sociocultural perspective. *Learning, Culture and Social Interaction, 6*,77-86. http://doi.org/10.1016/j.lcsi.2015.04.002

Svalberg, A. M.-L. (2009). Engagement with language: Interrogating a construct. *Language Awareness, 18*(3–4), 242–258. https://doi.org/10.1080/09658410903197264

Wang, J., An, N., & Wright, C. (2018). Enhancing beginner learners’ oral proficiency in a flipped Chinese foreign language classroom. *Computer Assisted Language Learning*, *31*(5–6), 490–521. https://doi.org/10.1080/09588221.2017.1417872

Wright, C. (2018). Effects of time and task on L2 Mandarin chinese language development during study abroad. *The Routledge Handbook of Study Abroad Research and Practice*, 165–180. https://doi.org/10.4324/9781315639970-11

使用 Line 社群與 ChatGPT 的翻轉學習來翻轉高中生對英文寫作的學習成效與影響

Using Line OpenChat and ChatGPT in Flipped Learning to Reverse the Learning Effectiveness and Impact of High School Students' English Writing

林芷均*，陳鏗任，戴于眞
國立陽明交通大學教育研究所
* linmaggie888.hs12@nycu.edu.tw

【摘要】對於考生來說，英文寫作是大考最後的得分關鍵，對於英文寫作技巧是透過練習與熟悉不同寫作架構去累積習得。然而在台灣的中等教育，以英文作為外語的學習者 EFL（English as a foreign language)通常缺乏實踐和應用的機會，且使用時可能感到不安。因此本研究選擇英文寫作課程進行翻轉學習，結合 Line OpenChat 通訊軟體與 ChatGPT 人工智慧，來解決學生對於英文寫作方面的焦慮與無從下筆的窘況。總結來說透過翻轉學習，不僅掌握了寫作技巧，還學會根據自己的情況和能力來掌控學習進度，實現了 108 課綱所強調的以學生為中心，提升了學生的自主學習和後設認知學習能力。
【關鍵詞】英文作為外語；英文寫作；翻轉學習；通訊軟體；人工智慧

Abstract: *For examinees, English writing is the key to scoring well on major exams. Mastering English writing skills involves practicing and becoming familiar with different writing structures. However, the secondary education in Taiwan, English as a Foreign Language (EFL) learners often lack practice opportunities and apply, leading to feelings of unease when using English. Therefore, this study aims for flipped learning in English writing courses, integrating the Line OpenChat communication platform and ChatGPT artificial intelligence to address students' anxieties and difficulties in English writing. Through flipped learning, students not only acquire writing skills but also learn to manage their learning progress, fulfilling the student-centered approach emphasized in the 108 Curriculum Guidelines and enhancing students' autonomy and metacognitive learning abilities.*
Keywords: Second Language Acquisition, English writing, Flipped Learning, Communication platform, Artificial Intelligence

1. 前言

學習寫作的過程中不盡然所有人都是順利無礙的，可能會因為心理或不同原因導致產生焦慮或寫作理解問題出現，像是字彙使用、面臨未學過的新文體架構有所卻步、以往在英文作文得分不高、作文當中使用的文法易出錯等等。因此對於學生們如何好好將英文應用在英文作文，並在課餘時間提升他們對於英文的使用習慣。然而在學習英文作文的方式，採用翻轉學習法來執行，放學之後在一個適合學習的環境下，觀看教師已錄製完成的影片或教師指定的網路上影音資源，使用有網路的任何載具來進行上課學習，並依學習者所需可以記下筆記或練習寫作。

台灣目前政府宣布啟動雙語教育實施，但實際層面的考量之下，使用與練習英文的機會仍少，針對英文寫作來進行授課的課程也不足。(National Commission on Writing, 2003)說明在美國的學校當中，寫作為連結所有知識點，帶來知識的力量、學生也可以對於自學自控有自信。基本學習三個 R 中(reading, writing and arithmetic 閱讀、寫作與算數)，傳統強調閱讀與算數，而寫作部分常常被忽略。近年來有所改變，韓國為了「翻轉下一代的人生」，投資教育不手軟，得知考題的論述題配分提高，也增加小學生補「論述」的流行(李岳霞，2014) ，如此一來可發現寫作課程為全世界的教育重要課題。

從個人的英文教學經驗中發現，學校教學時間非常有限，對於如何寫好一篇作文，教導在直接下筆之前，可以先擬好一個簡要大綱再去延伸思考，這樣也能提升學生對於敘述與文字表達的邏輯思考能力。在 Bloom(1956)的認知領域:知識、理解、應用、分析、評鑑、創造的六層次中，讓學習者能運用所學到的知識並理解之後，應用概念後往更高層次的創造出自己的創作。

於是需在有限的教學時間下，來做出一些改變，要求學生於到校上課之前，先學習好知識層面的內容，而到校之後在課堂時間中，教師不再為學習的中心來扮演學習設計師、學習策進者的角色，給於學生討論的主題與方式，「專題式學習」(Project-Based Learning，簡稱 PBL)分組讓學生進行討論，來腦力激盪出更多不同的想法，鼓勵學習者的認知參與並需要運用高階思維能力，讓學生對於知識更結構化與有脈絡性。

2. 研究問題、研究範圍與限制

*2.1.*研究問題

Q1.使用課前影片學習並於下堂課的課堂討論加上課後有 LINE 社群匿名討論與應己所需搭配使用人工智慧生成式轉換器 ChatGPT 對比傳統教學方式，在寫作的表現差異為何?

Q2.使用課前影片學習並於下堂課的課堂討論對比傳統教學方式，在寫作的表現差異為何?

Q3.使用課前影片學習並於下堂課的課堂討論加上課後有 Line 社群匿名討論與應己所需搭配使用人工智慧生成式轉換器 ChatGPT 對比使用課前影片學習並於下堂課的課堂討論，在寫作的表現差異為何?

Q4.翻轉教學的學習方式是否能夠提升學生的學習成效、學習興趣、學生的學業表現，降低學生對於第二語言習得的焦慮感?

2.2. 質性訪談問題

因為研究對象是未成年的台灣高中生，因此在訪談的過程中，要避免使用上對下的口吻或是過度為學生的回答做指引，一方面會擔心學生會因為我方訪問者是成年人，對學生來說是長輩，可能會遇到的困難是學生在回答上會有所顧忌或不完全表達出真正的想法或看法；另一方面也期望學生在回答時，可以盡量地表達不用需要太多訪談的指引或是提詞，以免學生開始按照訪問者的觀點出發。如此一來，在訪談的過程中，去營造一個輕鬆不緊張、嚴謹但不嚴肅的訪問環境來克服出現的訪談困難，以對話的方式來進行，讓受訪者補充更多細節。訪問所提出的問題順序性也是很重要的，前幾個問題盡量以寒暄的方式來展開訪問，基本的問題來拉開訪問的序幕，將整體訪談分為幾個部分，每個部份是一個小主題，每組主題有 4-5 個小問題來設計題目。

設定以下三大區塊主題:A.使用翻轉課程之前 B.使用翻轉課程中 C.使用翻轉課程後，將英文寫作翻轉課程的前中後來進行訪問。

A. 使用翻轉課程之前，詢問受訪者:

1. 傳統在教學句構與文法練習時，對於在練習這個層面上，會有什麼想法?
2. 老師是如何教學英文寫作的?
3. 在自行練習英文寫作時，有遇到什麼困難嗎?
4. 英文段考考寫作時，會花多少時間去完成呢?
5. 完成英文寫作會很焦慮嗎? 為什麼，可以多說一些嗎?

B.使用翻轉課程中，詢問受訪者:

1. 在家裡觀看英文寫作的教學影片過程中，能夠掌握影片中的概念?
2. 如果聽不懂影片內容的片段，會如何解決? 尋求什麼幫助?
3. 是否對於影片內容與寫作感到興趣? 為什麼?
4. 看完影片有自己練習寫一篇嗎?且分享於 Line 社群中嗎?
5. 在觀看過程中有什麼感受，思緒順暢?注意力?程度?忘記時間在流逝?

C.使用翻轉課程後，詢問受訪者:

1. 透過翻轉課程，對英文寫作的自信度是否提升?為什麼?

2. 回到教室之後，對於老師設計在課堂中的活動有什麼想法?哪一些活動最感興趣且印象深刻？

3. 對於在家裡使用 ChatGPT 3.5 與在學校使用 ChatGPT4.0 有感受到什麼差別?都如何發問的呢？

4. 透過不同主題性的寫作練習與影片解說觀看後，是否對英文寫作更加的熟悉呢?以後面臨大考或是學測還會焦慮嗎?

5. 傳統英文寫作與翻轉學習的英文寫作課終極二選一? 為什麼會這麼選?

在訪問的最後時可以用正面的方式進行收尾，對受訪者表達感謝與尊重。訪問者與受訪者需要隨時保持平等關係，不有所偏見與預設，結束之後，自己需要立即整理重點與當下所收穫到的，之後回後聽錄音看看倘問當下受訪者回答的聲調等細節，且在訪問之前先對受訪者說明錄音的必要性與存在。

*2.3.*研究範圍

本研究主要在探討新竹市地區高中生採用翻轉教學學習英文寫作，會有不同於傳統式教學法的影響:

(一)研究地區

本研究以新竹市地區的公立高中為研究範圍。

(二)研究對象

本研究係以新竹市地區的公立高中生為對象，總共三個班級。

(三)研究內容

本研究以英文作文融入翻轉學習與使用通訊軟體，來探討「學生自學」、「師生互動」、「同儕溝通」來探討學習效果與減少學習焦慮。

(四)研究方法

本問卷採前後寫作測驗與學習感受質性訪談法，依據這些測驗的結果與調查所得資料來進行統計分析與討論。

*2.4.*研究限制與挑戰

由於選定的實驗對象為公立高中生，可能會因為必須符合課綱的排課要求，壓縮寫作教學的時數，但若恰當與授課教師討論，事實上翻轉學習可能幫助公立高中生獲得較充分的寫作練習。選定目標為公立學校，如果是私立學校或雙語學校的學生，也許學生本身的寫作能力已達到精熟程度，或學校本身以安排充分節數加強學生能力，因此翻轉學習未必仍有幫助。此外，對於翻轉學習的英文寫作課，因為實驗對象為未成年的學生，需要在學期開始之前，讓家長同意此實驗與同意納入實驗數據當中，皆為此研究的限制與挑戰。

3. 研究內容

3.1. 翻轉學習的理念與應用

翻轉學習(flipped learning)有別於傳統式教學法，Jacob (2013)說明翻轉學習包含「主動學習」、「同儕學習」、「合作學習」、「問題解決能力」等理論，以下圖一為顯示此理論與其關係。符合十二年國教的「以學生為中心」(student-centered instruction)核心概念的翻轉學習，期望能提升學生的自主學習，有彈性的善用時間與學習環境，而學習者在下一堂課之前，先藉由預先錄製完成的影片進行自學之後，到教室後再進行相關的議題討論，如此一來，學習者可以好好地善加利用所學知識於課堂當中，並於回到課堂時可以藉由與同儕之間的討論完成作業或報告。也根據課綱中三面九向的自主行動、溝通互動、社會參，目標成為一個終身學習者，希望培養出能夠自發、互動、共好的學習者。從 2006 年起 Khan Academy 提供

各領域的免費線上學習教材、互動影片與習題，累積豐富的學習資料庫，開啟「線上學習」的時代。因此台灣也出現大規模開放線上課堂 MOOCs (Massive Open Online Course)，與可汗學院同為免費公開課程的數位學習平台 ewant 育網、learnmode 學習吧、因才網、均一教育等做為應用，可以依據平台功能去選擇來應用。現在部分大學納入 MOOC 平台有收費的學分課程，修讀完後獲得微碩士(MicroMasters)；然而，為符合課綱中的自主學習，有些高中學校要求學生需要修讀 MOOC 平台中的課程，像是 ewant 育網平台上，有開設一個專區提供高中生自主學習來修課。

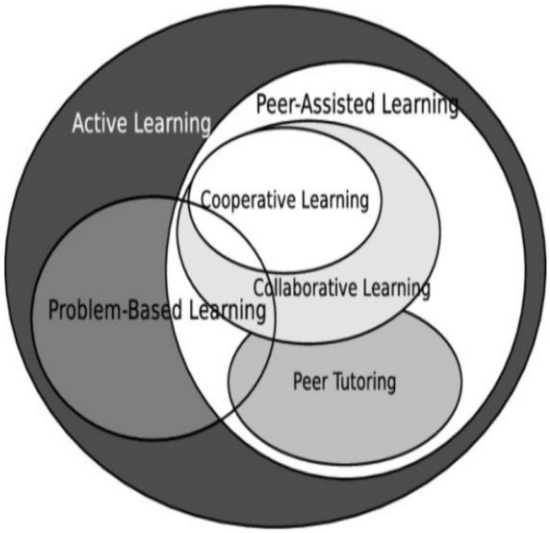


圖 1 學生為中心的學習理論和方法之維恩圖(Jacob， 2013)

根據 Bloom 的教育目標分類:記憶、理解、應用、分析、評鑑、創造，從低階層到高階層的傳統教學法翻轉變成以下圖二以 Bloom 認知領域教學目標解讀翻轉教學成效(劉怡甫，2013)，將記憶、理解低階層的部分由學生在課堂前完成，而應用、分析、評鑑與創造高階層的思維能力則是在課堂中老師的引導與發問，進行小組同儕討論與反思來完成，讓學生可以進行更高層次的思考。

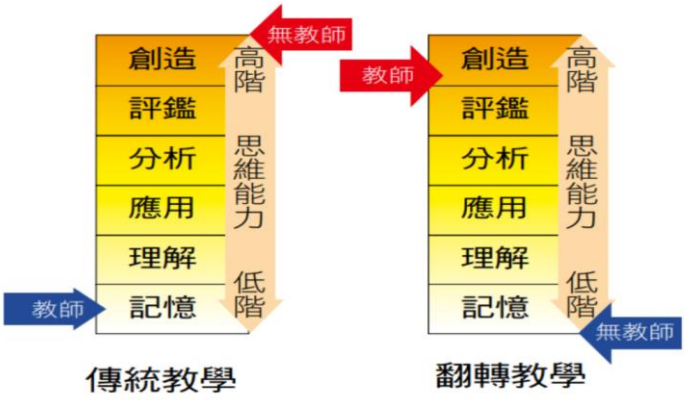


圖 2 以 Bloom 認知領域教學目標解讀翻轉教學成效(劉怡甫，2013)

許多研究指出翻轉學習的有效性與提升學生的學習效能、學習態度、學習動機等。Wu 與 Huang (2007)探討九年級學生以教師學習為中心與以學生為中心之差異，以台灣兩個科學班的學生來參與此研究，研究結果表明，兩方法皆促進學生對知識改念的理解，並提供學生對科學有不同的學習機會，但相較之下，以學生為中心的方法降低了學習焦慮程度、提升學生的自信心與增加積極正向的情緒與態度。Hsieh(2017)指出，學生成為學習的中心，而老師的角色為觀察者，在學生討論的過程當中去觀察學生的困難與錯誤，去給予適當的幫助與提

供意見回饋他們對於概念的釐清，如此一來，他們對學習內容不會不清楚，這樣能減少學習者對學習的焦慮。

3.2. 外語學習焦慮

以英文為外語的學習者在學習過程中產生「學習焦慮(learning anxiety)」可能影響原有的學業表現與學習成效。Dalkilic (2001)根據研究結果表明有許多因素而導致的，如想引人注目、缺乏自信、個性害羞、缺乏相關知識和對他人的過高期望。因此，有些學生在課堂中保持沉默不語、越來越沒自信導致焦慮與害怕的情緒產生。

MacIntyre & Gardner(1994)指出「外語學習焦慮」是在語言習得的認知過程的輸入、處理與輸出，換言之在學習外語的聽、說、讀、寫，會產生出焦慮與擔心的感受。周啟葶等人（2009）研究指出焦慮對於外語學習認知過程傷害，會干擾學習、與造成訊息接收障礙的情緒。因此李安悌（2009）提出建立學習者的自信、讓學習者了解自身的長處；也促進合作學習，避免過度競爭，發展學習的安定感而不焦慮。雖然「焦慮」所帶來的大多負面的影響，但存在心理學大師羅洛・梅（朱侃如譯，2004）提出「焦慮是我們的良師，人因為既有現又自由，既受限又無限，所以是焦慮的。自由與有限並存的弔詭情境，使人生而焦慮。」焦慮是無所不在的，深入發掘原因，以建設性態度面對並解決。因此對於教師來說，如何降低學生對英文寫作而產生的焦慮感為是首要的關鍵，去思考如何運用翻轉學習到英文的作文課程中，降低學生對於英文寫作的焦慮感為主要目標。

3.3. 通訊軟體與人工智慧為輔助工具融入教學

資訊時代的到來使得高科技的蓬勃快速發展，不同功能的軟體日新月異，社交媒體與通訊軟體也不斷推陳出新，更出現了以大型語言模型為架構的人工智慧聊天機器人，五花八門的輔助教學工具的出現，如何善加運用到教學中是值得深思的。楊凱鳴、陳驊、劉擎華（2013）認為透過行動化科技工具的輔助能貼近學生生活，引起學生學習興趣、打破傳統教學方式，也讓學習成果提升。此外，學生也可以透過使用 AI，在與大數據的問答對話中，讓作文整體架構更加釐清，善加練習、對寫作不焦慮後，面臨在無法使用電子設備的大型考試中，更有信心與把握去揮灑自己的文筆、應對自如，對於如何減少焦慮感與寫作能力的提升，能夠獲取更高的分數是極其重要的。

現今許多學者也應用於教育層面中，讓學習者可以不受時空限制來學習與溝通互動，黃敬恆(2017)線上溝通工具如 Line、Google Hangouts 以及 Google Group 除了為師生溝通的工具，也對促進合作學習有幫助。蔡孟學(2022)表示當教師使用師生 Line 群組時，會有班級經營效果，顯然師生 LINE 群組已然轉化為一種「線上的班級」。論及人工智慧（Artificial Intelligence, AI），因此李安悌（2009）提出建立學習者的自信、讓學習者了解自身的長處；也促進合作學習，避免過度競爭，發展學習的安定感而不焦慮。張靜淑(2023)比喻人工智慧應用在教育層面為「虛擬助教」提供學生即時回應，也能夠分析學生的學習模式、答題習慣和互動行為。如部分平台上有 AI 作為課程助教可以解答學生的基礎問題，幫助克服學習上的障礙。雖然方便與高效率，但終究是輔助工具，真人的實際教學與情感是不被取代的。

然而，人工智慧應用不斷演進與發展，教師、政府相關機構須攜手合作，接受人工智慧帶來挑戰和未來性，確保有效、合乎道德地融入，藉由善用科技數位工具、減少學習者的學習焦慮與困難，並提升解決問題的能力。蔡至誠、羅希哲(2014)說明學習者從被動學習轉為主動學習，在科技工具的輔助之下，可以更方便吸收知識，也更了解學習上的疑問與弱點，透過檢視自己來改變學習的方向，以獲得學習成就感為目標。

4. 未來工作與預期效果

研究實驗對象為公立三個班的高中生分成三組，第一實驗組為課前影片學習並於下堂課的課堂討論、第二實驗組課前影片學習並於下堂課的課堂討論加上課後有 Line 社群匿名討

論與應己所需搭配使用 ChatGPT；控制組為在教室當中進行傳統授課模式上課；學生皆需進行前後測驗與填寫學習感受問卷，以下表一為描述此堂課的課前、課中與課後的活動，圖三為研究的自變項、依變項與中介變項。預期結果顯示大多學生有進步，兩組實驗組進步的幅度更大，而第二實驗組進步最多，且大幅提升師生互動與同儕交流的程度。此外對於翻轉學習英文寫作課，實驗對象為未成年的學生，需在學期之前讓家長同意此實驗與同意納入實驗數據當中。最後，對於如何達到成功的「翻轉學習」方式，需花費大量的時間與詳細規劃。不僅是教師教學的轉變、學習者的學習態度與自控能力、積極的態度與合作構成主要的影響因素需要互相配合，這有趣、有效的課堂模式才能達成好的教學效果。且羅志仲(2014)表示只要教師有心嘗試，不論任何程度的學生都能收效，不同的教學方法經過調整都能適用的。

表 1 翻轉學習的流程

課前活動	● 觀看指定課程影片 ● 完成指定教材內容的練習
課中活動	● 小組依主題進行討論 ● 老師領導 Q&A ● 小組辯論競賽
課後活動	● 使用 Line 社群發表想法與疑點解惑 ● 行練習該週寫作文體並繳交於下週上課前繳交給授課老師

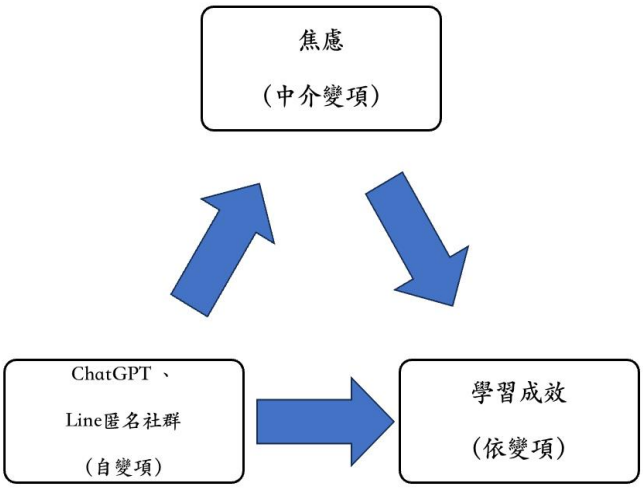


圖 3 本研究的自變項、依變項、中介變項

再者，課堂討論的規劃中，依據課前影片的進度選定主題，讓學生分組後進行討論，每堂課可以使用不同的方式去討論，像是心智圖、寫作漢堡圖、甚至是辯論比賽讓學生腦力激盪想出更多不同的方向與概念。而選定使用課後討論通訊軟體是 Line 社群的原因是有些在班上害羞的同學，可能擔心在課堂、具名發言會引來關注，因此 Line 社群提供匿名的聊天室，學習者可以對於課前影片內容不清楚的地方立即詢問、針對課堂討論內容，有新的或是其他想法可以分享、自行回家寫作練習或筆記，可以在這裡公開分享，大家也可以給遇見譯與評論。

最後對於學習感受度問卷的設計中，會於學期結束前給受試者來進行填寫，詢問學習者對於翻轉學習的態度與想法，對此學習方式的趣味性、同儕與師生之間的互動性、自身對於自主學習的安排與想法、對於培養獨立思考的能力幫助的程度維和、整體課堂的氣氛、課後在 Line 社群上的倘論參與度等題目來了解學生經過此學期的翻轉英文寫作課程的想法與感受。選項採用 Likert 五點亮表計分(1 分表示「此感受程度非常低」、5 分表示「此感受程度非常高」)，分數越高，代表對於此翻轉英文寫作課的滿意度、接受度越高。

參考文獻

朱侃如(譯) （2004）。焦慮的意義(原作者:R. Mayr)。**台北市:立緒出版社**。（原著出版年：1950）

李安悌（2009）。動機導向、學習壓力、學 習策略及學習焦慮之相關研究 —以 技職院校管理學院學生為例。**人力資源管理學報，9**(1)，45-67。doi:10.6147/ JHRM.2009.0901.03

李岳霞 （2014）。韓國小學生瘋補「論述」。**親子天下，56**，40-41。

李佳容（2018）。參與式翻轉教學的挑戰與未來。**臺灣教育評論月刊，7**(8)，31-34。https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-201808-201808230011-201808230011-31-34

周啟葶、程玉秀、宋秋美（2009）。英語自我效能、英語學習焦慮與英語學習成就 之相關研究—以臺北市高中生為例。**教育與心理研究，32**(2)，81-111。

孫瑞、夢瑞森、文萱（2015）。翻轉課堂教學模式在對外漢語教學中的應用。**語言教學與研究， 3** ，34-39 。

許佩玲、易正義（2016）。臉書社群有助於翻轉教學。**亞東學報，36**，63-81。https://www.airitilibrary.com/Article/Detail?DocID=18133754-201612-201710200017-201710200017-63-81

陳美子（2022）。國中英語教師改善英語教學成效之行動研究。**學校行政，141**，93-126。https://doi.org/10.6423%2fHHHC.202209_(141).0005

陳燕秋（2021）。現代科技融入翻轉教學於華語課堂之建置和運用。**華文世界，128**，70-96。https://www.airitilibrary.com/Article/Detail?DocID=10170855-202112-202201180013-202201180013-70-96

張靜淑（2023）。人工智慧在國中小教學現場之運用。**台灣教育，744**，73-80 https://www.airitilibrary.com/Article/Detail?DocID=18166482-N202312190007-00008

黃能富（2015）。磨課師（MOOCs）與師博課（SPOCs）協同授課之翻轉教學法。**教育脈動，1**，101-110。https://www.airitilibrary.com/Article/Detail?DocID=P20160218002-201503-201602180003-201602180003-101-110

黃敬恆（2017）。社群媒體與教學平台對輔助學習成效之研究〔碩士論文，淡江大學〕。華藝線上圖書館。https://doi.org/10.6846/TKU.2017.00157

曾璧光、陳美蓮、宋修德（2021）。人工智慧在中等技職教育之實踐芻議。**台灣教育，730**，65-74 。 https://www.airitilibrary.com/Article/Detail?DocID=18166482-202108-202108190010-202108190010-65-74

楊凱鳴、陳驊、劉擎華（2013）。應用 QR Code 於國小自然科輔導教學－以認識校園動植物為例。**文化創意產業研究學報，3**(1)，25-32。https://doi.org/10.6639/JCCIR.201303_3(1).0003

葉丙成（2015）。BTS 翻轉教室 -3 馴化篇 。取自 https://flipedu.parenting.com.tw/blog-detail?id=482

葉丙成（2015）。如何確保翻轉教學的成功？BTS 翻轉教學法。**中等教育，66**(2)，30-43。https://doi.org/10.6249/SE.2015.66.2.03

溫富榮、李錫鑫（2018）。一所高中翻轉教學紀實與分享。**臺灣教育評論月刊，7**(8)，62-67。https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-201808-201808230011-201808230011-62-67

賓靜蓀（2013）。專訪可汗學院創辦人 薩曼・可汗：創造不用怕丟臉的學習世界。**親子天下雜誌，46**。https://www.parenting.com.tw/article/5049500
蔡至誠、羅希哲（2014）。QR Code 應用於高職數學之教學歷程。**中等教育，65**(2)，110-121。https://doi.org/10.6249/SE.2014.65.2.07
蔡孟學（2022）。從 Goffman 戲劇論探析師生 LINE 群組。**臺灣教育評論月刊，11**(1)，93-102。https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-202201-202201030031-202201030031-93-102
劉怡甫（2013）。翻轉課堂—落實學生為中心與提升就業力的教改良方。**評鑑雙月刊**，41，31-34。https://www.airitilibrary.com/Article/Detail?DocID=a0000268-201301-201301150016-201301150016-31-34
謝傳崇、李孟雪（2021）。台灣國民小學教師翻轉教學之研究－TALIS 的分析與驗證。**師資培育與教師專業發展期刊，14**(1)，91-113。https://doi.org/10.3966/207136492021041401004
顏國樑(2018)。透過十二年國民基本教育課綱分析翻轉教學之實踐。**臺灣教育評論月刊，7**（8），41-48。
羅志仲（2014）。翻轉教室翻轉學習。**師友月刊，563**，20-24。https://doi.org/10.6437/EM.201405_(563).0005
Awidi, I. T., & Paynter, M. (2019). The impact of a flipped classroom approach on student learning 121 experience. **Computers & Education, 128**, 269-283.
Dalkilic, N. (2001). The role of foreign language classroom anxiety in English speaking courses. **ENSTITU DERGISI, 8**(8), 70-82.
Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. **Journal of Computer-Mediated Communication, 12**(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x
Hsu, L. (2017). Explorations in Flipped Classroom and College Student's English Learning Anxiety. **Journal of National Taichung University of Science and Technology, 4**(1), 127-143. https://doi.org/10.6902/JNTUST.201712_4(1).0006
Huang, B., Hew, K. F., & Lo, C. K. (2019). Investigating the effects of gamification-enhanced flipped 122 learning on undergraduate students' behavioral and cognitive engagement. **Interactive Learning Environments, 27**(8), 1106-1126.
Jacob, B. （2013）. The flipped classroom: A survey of the research. **120th ASEE Annual Conference & Exposition**
Lin, C. L. (2021). Examining Teachers' Acceptance of Flipped Learning by Using Technology Acceptance Model. **International Journal on Digital Learning Technology, 13**(4), 27-61. https://doi.org/10.3966%2f2071260X2021101304002
Lin, C. J., Hwang, G. J., Fu, Q. K., & Chen, J. F. (2018). A Flipped Contextual Game Based Learning Approach to Enhancing EFL Students' English Business Writing Performance and Reflective Behaviors. **Educational Technology & Society, 21**(3), 117-131. https://www.airitilibrary.com/Article/Detail?DocID=P20221223002-N202302130013-00011
MacIntyre, P. (1994). The subtle effects of language anxiety on cognitive processing in the second language. **Language Learning.**
Peng, T. L. (2022). Improving EFL College Students' Higher-Order Thinking Skills Through Flipped Instruction. **International Journal on Digital Learning Technology, 14**(2), 139-169. https://doi.org/10.53106/2071260X2022041402006
Magrath, C.P. et al.(2003).The neglected "R". **The College Board.**
State education about Curriculum Guidelines of 12-Year Basic Education in Taiwan adopts a "student-centered" approach.
Wu, H., & Huang, Y.-L. （2007）. Ninth-grade student engagement in teacher-centered and technology-enhanced learning environments. **Science Education, 91**（5）, 727-749.

結合雙層次測驗之適性輔助學習模式對於國小數學學習之問題解決與後設認知能力之影響

The Effects of an Adaptive Assisted Learning Model Integrating Two-tier Test on Students' Problem-Solving and Metacognitive Abilities in Elementary Mathematics Education

蘇芳玉[1], 王俊傑[1*]
[1]國立屏東大學教育學系
* jie809@gmail.com

【摘要】 台灣近年來結合「因材網數位適性學習平台」(因材網)（Tawian Adaptive Learning Platform, TALP）推動自主學習。學生使用因材網學習時，需閱讀影音教材、完成動態評量與測驗，但結果不一定能反映學生學習情況。為解決此問題，本研究採用 Zuvio 設計雙層次測驗輔助機制，協助學生反思個人答題根據，增加思考和釐清迷思觀念之能力。研究採準實驗設計，實驗組以導入雙層次測驗之適性化學習模式進行學習，控制組以一般適性化學習模式進行學習，以國小六年級數學為例，探討此學習模式對學生學習成效、自主學習能力、問題解決能力與後設認知能力之影響。

【关键词】 雙層次測驗、自主學習、適性教學、問題解決、後設認知

Abstract: *In Taiwan, the Tawian Adaptive Learning Platform(TALP) has been used to promote independent learning in recent years. When students use TALP to learn, they need to watch audio and complete dynamic assessments, but the results do not reflect their learning. To solve this problem, this study design a two-level test to assist, and to increase their ability to think and clarify their myths. The study adopted a quasi-experimental design, with the experimental group learning in the adaptive learning mode of the two-level test and the control group learning in the general adaptive learning mode. Taking mathematics in the sixth grade as an example, the study investigated the effects of this learning mode on students' learning outcomes, problem solving and metacognitive abilities.*

Keywords: two-tier test, self-regulated learning, problem-solving, metacognitive abilities

1.緒論.

1.1. 研究背景與動機

目前教育的主流之一，是希望訓練學生成為具有自主學習能力的學習者，因此，台灣近年推動因材網及自主學習輔助計畫，藉由具有適性教學輔助功能的學習平台，輔導學生成為具有自學能力的學習者。這裡提到的「因材網數位適性學習平台」(簡稱因材網)，是具有適性教學輔助學習功能的學習平台，因有著概念星空圖的設計，因材網具備非常優異的診斷性評量功能，能在快速的檢測中，更有效率地知道學生的學習弱點，進而建議學生個人化的學習路徑，因此能快速有效地改善學生學習與教師教學效能。因材網結合了 AI 技術，針對學生學習的不同進程，設定了「數位檢核點」，當影片播放至影片數位檢核點時，影片會暫停，並由視窗中跳出問題。學生需進行作答，正確後才能繼續觀看影片，接續學習進度。此外，因材網也能在教學影片中附加「線上隨堂考」，教師透過觀看學生自學階段的數據資料，了解學生的問題，確認是否有普遍出現的問題或個別的學習狀況，再進行教學設計的安排與輔

導機制。藉此提升教學的個別化與教學成效。(因材網網址： adaptive-learning.moe.edu.tw/)。張靖敏等人(2022)指出因材網能夠激發學生學習動機，並提高自主學習的效果。

然而，使用因材網學習，仍有些教學現場的問題亟待解決。例如劉昌鑫(2022)指出，不同學習成就之國中學生在自學階段的投入程度仍有顯著的差異。而學生在完成自學階段的任務後，從目前的教學機制不一定能確認學生是真的理解作答的答案正確與否的原因。因此，我們在本研究中，嘗試結合雙層次測驗機制，希望可以有效理解學生作答正確與否的概念，並加以輔導學生，達到個人化學習的成效。同時，探討結合適性化學習輔導平台學習結合雙層次測驗機制後，是否能讓學生專注於問題解決與瞭解自己的學習歷程，因此同步探討問題解決能力與對學生後設認知之影響。以下為本研究的研究問題：

(1) 相較於「一般適性化學習模式」，「導入雙層次測驗之適性化學習模式」能否提升國小高年級學童之數學科學習成效。
(2) 相較於「一般適性教學輔助學習模式」，「導入雙層次測驗之適性化學習模式」，能否提升國小高年級學童之數學科個人自我效能。
(3) 相較於「一般適性化學習模式」，「導入雙層次測驗之適性化學習模式」，能否提升國小高年級學童之問題解決能力。
(4) 相較於「一般適性化學習模式」，「導入雙層次測驗之適性化學習模式」，能否提升國小高年級學童之後設認知能力。

2.文獻回顧.

2.1 適性化教學輔助數學學習

「因材網數位適性學習平台」(簡稱因材網)目前在臺灣的國小教育上，已普遍將導入數學科學習扶助課程，研究顯示學生在認知學習及情意學習的表現上有顯著提升（陳佳琳、張美珍，2023)；除了將因材網用於學習扶助，蔡宗倫、陳美如（2022）研究發現因材網融入數學科教學，對於學童的自學能力及數學學習成就皆有所進步。已有許多學者研究提出，運用 AI 數位輔助學習平台融入數學科教學，能有效提升學生之學習成效(陳佳琳、張美珍，2022)；然而，陳昱宏、鄭宇哲(2023)的研究結果指出，於教學中加入因材網互動式學習，對於學習扶助生學習數學之學習成效並沒有顯著提升，卻可以改變學習習性。Agha 與 Ata (2016)聲稱學業成績會受到學習習性及後設認知影響，教師的引導、鼓勵和激勵是改善學生學習習慣和表現的顯著特徵，且高後設認知的學生會積極參與學習歷程。因此，期望設計一個讓學生可以監控自己後設認知的評量方式，使學生改變學習習慣與提升學習成就。

2.2 雙層次測驗

依據前段敘述，雙層次測驗(Two-tier test) 已被多位學者證實可以有效診斷學生的先備知識、誤解或替代概念(Yang, Hwang, Yang, & Hwang, 2015)。雙層次測驗包含一組兩層多選題，第一層評估學生對現象的描述性或事實性知識，第二層探討學生在第一層做出選擇的原因。雙層次測驗深入探索學生對事實知識的解釋，使教師或研究人員不僅能夠了解學生可能存在的錯誤觀念，還可以評估觀念背後的推理或深入理解(Chu et al., 2010)。

3.研究方法.

3.1 研究對象

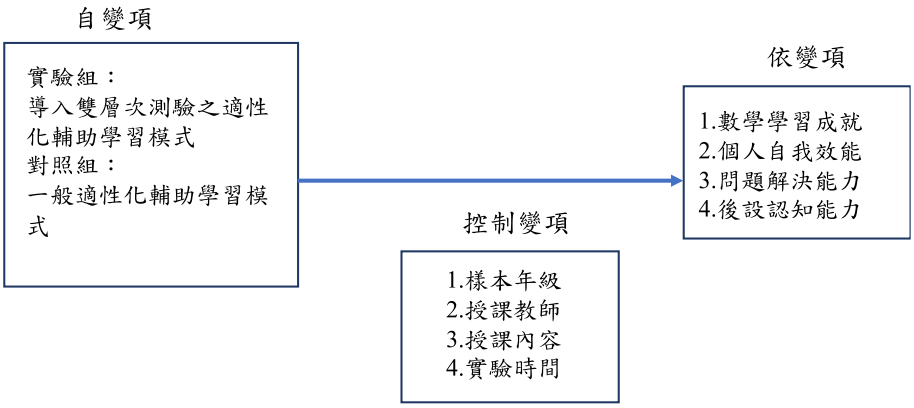
本研究以 112 學年度臺南市某國小六年級的兩個班級為研究對象，每班 28 位學生，共 56 位，研究者任教之班級為實驗組，另一個班級則為控制組。該校總班級數為 48 班，規模為大型學校。本研究時間將進行將近 10 週的實驗教學，每週有四節數學課，每節 40 分鐘，共計進行 40 節課，教學者為研究者本人。

3.2 研究設計

本研究採用準實驗研究來探討導入雙層次測驗之適性化學習模式對學生自主學習的實施影響。研究架構圖如圖 1：

圖 1

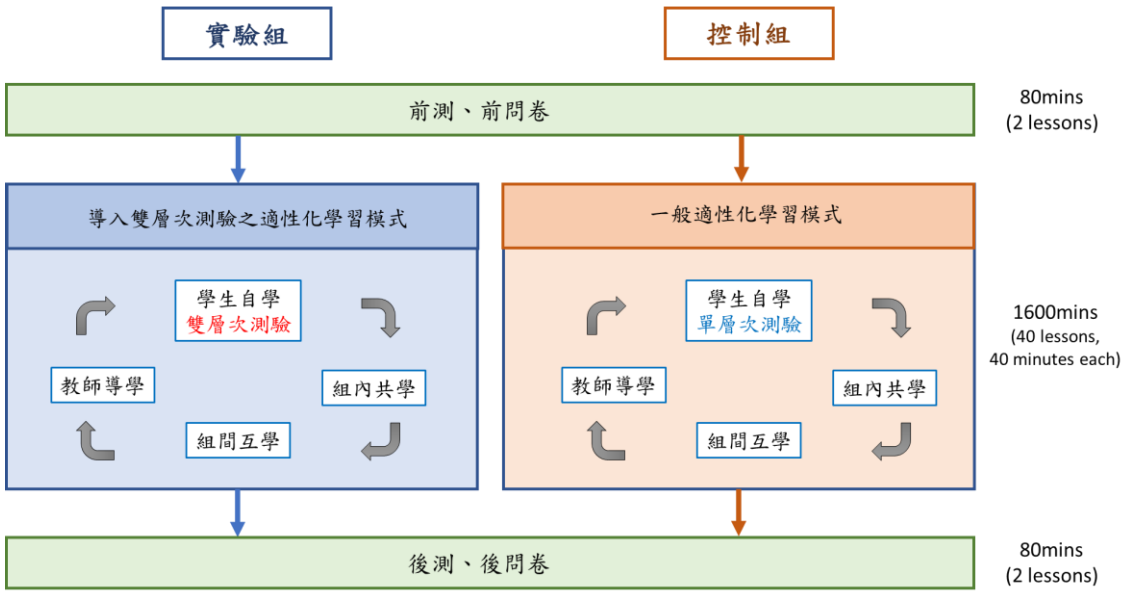
研究架構圖



實驗擬共教學七個單元，分別是長條圖與折線圖、小數的除法及圓周長與扇形弧長、比與比值與正比、縮放圖與比例尺、規律問題及速率。前測的採用工具為 112 年臺南市五年級數學學力測驗試題，後測的採用工具為本校 112 學年度六年級數學領域第二次定期評量。除此之外，在實驗教學進行之前，及教學實驗結束之後，讓兩組學生填寫前問卷與後問卷，觀察兩組學生在回饋問卷上的感受是否有差異，並將其意見納入修正之參考。學習活動流程圖如圖 2：

圖 2

學習活動流程圖



3.3 數位教材設計

將研究對象分為實驗組及對照組，兩組在數學課教學皆融入一般適性化學習模式，也就是因材網之數位教材，及採用自主學習之教學設計，進行學生自學、組內共學、組間互學及教師導學，探討學生對於問題解決與後設認知能力之影響，唯有實驗組學生在學習後導入雙

層次測驗。

以因材網數學領域來說，每一則影片都是一組知識結構，每當精熟一組知識結構的學習，就會在星空圖上增加一顆綠色「精熟」的節點；若是在該知識結構學習上未達成精熟標準，即會增加一顆橘色的「待補救」節點；假如該知識結構未進行測驗過，會顯示一顆灰色的「未測驗」節點。如圖 3 所示。

圖 3

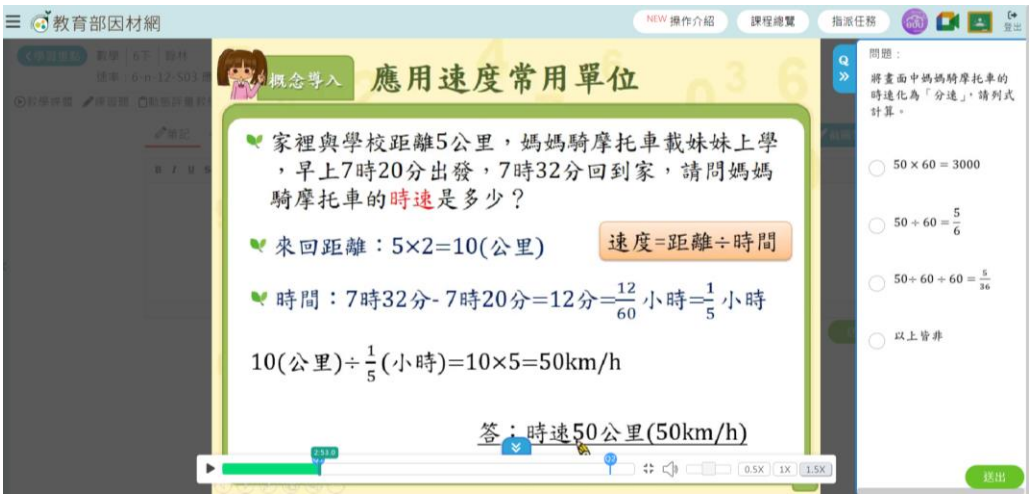
知識結構星空圖



在學生自學階段，兩組學生除了預習數學課本內容，也需登入因材網觀看任務影片。影片時間軸當中有兩個節點，每個節點會設計一題動態評量，針對學習概念進行即時診斷。如圖 4：

圖 4

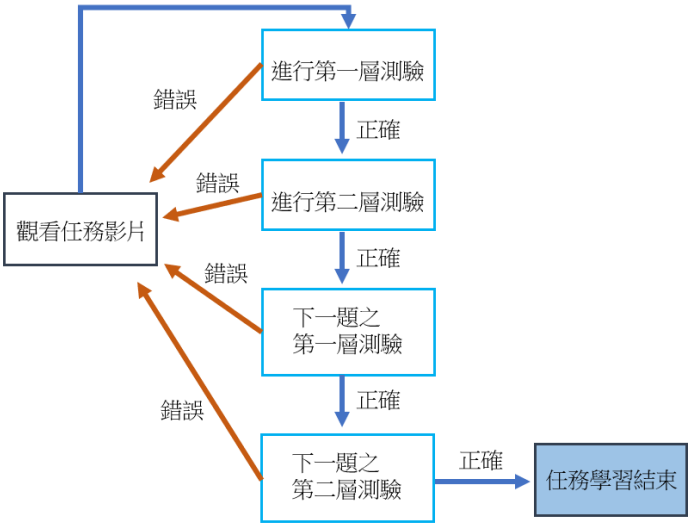
動態評量頁面



若答對會有正向鼓勵；答錯則會顯示正確答案之選項。完成動態評量後，對照組填寫紙本測驗題，若有迷思概念需標註；實驗組則填寫以 Zuvio 平台設計之雙層次測驗，雙層次測驗流程如圖 5 所示。

圖 5

雙層次測驗流程圖



以「速率」這個單元在 Zuvio 平台上之雙層次測驗，第一層設計 4 選 1 單選題；第二層則設計問答題，可以文字輸入，或是用拍照方式，讓學生將思考及計算歷程記錄下來。Zuvio 學生端作答頁面如圖 6：

圖 6

Zuvio 學生端作答頁面



學生完成作答後，教師端可從後台了解學生作答情形，包括答對率統計、答對人數及答對人名。Zuvio 教師端檢討頁面如圖 7：

圖 7

Zuvio 教師端檢討頁面



3.4 研究工具

教學前，為了瞭解學生基本能力，以 112 年度縣市國小學生學習能力檢測試題(以下簡稱 112 年度國小學力檢測)之五年級數學科為前測的研究工具，並填寫前問卷(個人自我效能、問題解決與後設認知能力)；後測的採用工具為本校 112 學年度六年級數學領域第二次定期評量，搭配後問卷，包含問題解決能力及後設認知察覺量表(Lai & Hwang, 2014)，觀察兩組學生在回饋問卷上的感受是否有差異。

3.4.1 前測

以國立臺中教育大學測驗統計與適性學習研究中心設計之 112 年度國小學力檢測為研究工具，因本研究之實驗設計為六年級上學期，學生未學習到六年級下學期之單元，不適合以六年級數學科為研究工具，而是以 112 年度國小學力檢測之五年級數學科為研究工具。試題包含 25 題 4 選 1 單選題(每題 4 分)，合計 100 分。

3.4.2 後測

第二次定期評量由兩位六年級導師共同設計，再交由四位六年級導師審題，包含 5 題是非題(每題 2 分)，10 題選擇題(每題 2 分)，13 題填空題(每題 2 分)，8 題計算題(每題 3 分)，5 題應用題(每題 4 分)，合計 100 分。

3.4.3 問題解決能力及後設認知察覺量表

表 1

問題解決能力及後設認知察覺量表

問題解決能力及後設認知察覺量表				
問題解決傾向 Problem-solving tendency				
題目	非常同意	同意	不同意	非常不同意
1.我相信我有能力解決我所遇到的問題。 I believe that I have the ability to solve the problems I encounter.				
2.我相信我可以靠自己解決問題。 I believe that I can solve problems on my own.				
3.我經歷過解決我所遇到的問題。				

I have experiences of solving the problems I encounter.				
4.當遇到問題時，我願意面對並處理。 When encountering problems, I am willing to face and deal with them.				
5.我不會逃避我所遇到的問題。 I will not escape from the problems I encounter.				
6.我總會盡我最大的能力來解決所遇到的問題。 I always try my best to solve the problems I encounter.				
後設認知察覺傾向 Meta-Cognition Tendency				
題目	非常同意	同意	不同意	非常不同意
1.我會定時檢視自己是否有達到我的目標。 I ask myself periodically if I am meeting my goal.				
2.我會定時檢視，幫助自己了解重要的關係性。 I periodically review to help me understand important relationships.				
3.我自己會定期檢視我的認知。 I find myself pausing regularly to check my comprehension.				
4.當我完成一個任務時，我會檢視自己達到預訂目標的程度。 I ask myself how well I accomplished my goals once I'm finished.				
5.一旦我完成一個任務，我會詢問自己學到的是否足夠。 I ask myself if I learned as much as I could have once I finish a task.				

4.預期的研究結果

本研究預期探討相較於「結合 AI 互動式學習平台之學習模式」,「結合雙層次自主學習機制與 AI 互動式學習平台之學習模式」能否有效提升國小高年級學童之數學科學習成效、個人自我效能，以及問題解決能力、後設認知能力等高層次思考能力。

參考文獻

林美如(2022)。一所偏鄉小學因應遠距教學之問題與對策。**台灣教育評論月刊**，**11(5)**，178-181。
教育部(2021)。十二年國民基本教育課程綱要。2021 年 2 月，取自
https://www.naer.edu.tw/upload/1/16/doc/288/%E5%8D%81%E4%BA%8C%E5%B9%B4%E5%9C%8B%E6%9

5%99%E8%AA%B2%E7%A8%8B%E7%B6%B1%E8%A6%81%E7%B8%BD%E7%B6%B1.pdf。

教育部(2019)。促進親師溝通、老師變輕鬆、學生喜歡用 因材網教育界引發風潮，能有效提升學習與教學效能，2019 年 7 月 24 日，取自 https://www.edu.tw/News_Content.aspx?n=9E7AC85F1954DDA8&s=6AA55A85D10DF9F8

陳佳琳、張美珍（2023 年 11 月）。**數位科技運用在偏鄉國民小學扶助教學數學科學習成效影響之研究－以因材網為例**。載於社團法人台灣工程教育與管理學會（主編），2023 第 19 屆科技教育研究與發展學術研討會 2023 第 12 屆工程、技術與 STEM 教育研討會論文集（588-603 頁）。台灣：社團法人台灣工程教育與管理學會。

陳昱宏、鄭宇哲（2023）。因材網互動式學習模式融入數學學習扶助教學之研究。**國際數位媒體設計學刊**，**15(2)**，49-61。

張靖敏、張道宜、劉雅萍、陳美瑩（2022）。因材網：教育部人工智慧知識結構分析診斷及適性學習平台之析論。**台灣教育研究期刊**，**3(1)**，313-331。

蔡宗倫、陳美如（2022）。運用適性教學輔助平臺融入數學教學提升自我調整學習能力與學習成效之研究。**中等教育**，**73(2)**，35-54。

劉昌鑫（2022）。**國中普通生及資優生自主學習之調查研究**。國立臺灣師範大學特殊教育學系碩士學位論文，取自華藝線上圖書館系統，臺北，未出版。

賴暄頤（2010）。全方位課程設計之初探。**特教論壇**，**(8)**，43-55。

Agha, S., & Ata, u. R. (2016). Learning habits as factors influencing academic performance in medical students. *Pakistan Journal of Psychology, 47*(2).

Chu, H.-C., Hwang, G.-J., Tsai, C.-C., & Tseng, J. C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers and Education, 55*(4), 1618-1627. https://doi.org/10.1016/j.compedu.2010.07.004

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

Yang, T., Hwang, G., Yang, S. J. H., & Hwang, G. (2015). A two-tier test-based approach to improving students' computer-programming skills in a web-based learning environment. *Journal of Educational Technology & Society, 18*(1), 198-210.

探討不同 VR 學習模式對國小學生地震防災學習成效與情境興趣之影響

Effects of Different VR Learning Models on Elementary School Students' Learning Performance and Situational Interest in the Context of Earthquake Disaster Prevention

陳振遠[1]，孫之元[2*]
[1]桃園市立幸福國民小學
[1,2]國立陽明交通大學 教育研究所
[*] jerrysun@nycu.edu.tw

【摘要】 近年，VR 虛擬實境科技已經開始應用於教育領域中，藉由 VR 的擬真環境、沉浸性與互動功能，能改善過去防災教育的缺點，且頭戴式顯示器 VR 系統更有機會提高學生的學習成效與避難技能，因此本研究探討頭戴式 VR 學習模式與一般電腦桌面式 VR 學習模式對於國小高年級學生地震防災課程學習之影響。研究結果顯示，使用頭戴式 VR 學習模式的學生，有更好的學習成效，以及更高的情境興趣，因此在地震防災課程中導入頭戴式 VR 裝置，對於學生知識與技能的學習上有正面的影響，能提升學生對於教材的理解與參與度，故對於國小教育現場，頭戴式 VR 裝置有應用與推廣之價值。

【關鍵字】 虛擬實境；頭戴式顯示器；情境興趣；地震防災教育

Abstract: *In recent years, Virtual Reality (VR) technology has started to be applied in the field of education. Through the realistic environment, immersion, and interactive features of VR, it can address past shortcomings in disaster preparedness education. Additionally, VR systems with head-mounted displays have the potential to enhance students' learning outcomes and disaster preparedness skills. Therefore, this study investigates the impact of head-mounted VR learning mode versus traditional desktop VR learning mode on earthquake disaster preparedness education for elementary school students. The results show that students using the head-mounted VR learning mode achieve better learning outcomes and exhibit higher situational interest. Therefore, integrating head-mounted VR devices into earthquake disaster preparedness courses has a positive impact on students' acquisition of knowledge and skills.*

Keywords: virtual reality, head-mounted display, situational interest, earthquake disaster prevention education

1. 研究動機

近年，虛擬實境（virtual reality, VR）科技與教材逐步發展成熟，其設備外觀的精簡與功能提升，使 VR 科技開始廣泛應用於教育與職業訓練等領域（Soliman et al., 2021）。在臺灣教育方面，教育部積極推動「112-113 年 5G 新科技學習示範學校計畫」，協助各級學校採購與應用 VR 設備進行創新教學，引導教師藉由虛擬實境跨越時空限制，提供學生多元環境與沉浸式的學習體驗，提升不同領域之學習興趣與自主學習能力（教育部，2023）。

臺灣地處太平洋地震帶，且過去發生許多重大災害，因此學生需在各教育階段學習災害成因、認識風險管理與進行救難演練實作等，強化未來面對重大災害之應變與自我照護能力（國家教育研究院，2019），但 Feng 等人（2020）指出目前防災教育課程多採用紙本教材、多媒體欣賞或實地演練等方式進行，然而學生對此活動的參與度不高，且一般課程也難營造真實場景讓學習者進行體驗與學習，而上述之困境能藉由 VR 設備與教材來解決，虛擬實境設備能提供使用者身歷其境的感受，將 VR 導入防災教育能給予學生真實與安全的訓練，有效提升地震應變能力（Shu et al., 2019），因此 VR 設備導入地震防災教育有其價值與必要性，

但過去研究多以成人或中、高等教育為主（Hsu et al., 2013；; Shu et al., 2019; Feng et al., 2018; Wibisono, 2019），故有必要探討 VR 設備實踐於國小教育現場之效益，以展現 5G 新科技之學習價值。

綜上所述，本研究設計一套以 VR 設備與教材為基礎之地震防災課程，導入虛擬實境一體機頭盔與 VR 遊戲化教材，比較不同形式之 VR 環境對國小生在學習經驗上之差異，也藉由創新科技的模擬動畫與地震真實情境營造，讓學生進行高度沉浸與互動之學習體驗，以了解 VR 對學生學習成效與情境興趣之影響，期待學生經由課程提升地震相關知識與熟練避難保護措施。

2. 文獻探討

2.1. VR 應用於防災教育

虛擬實境是只由電腦或是資訊科技建構出的 3D 立體模擬環境，並具有互動功能、使用者操作介面與沉浸感體驗之數位系統（Pesce, 1995; Koolivand et al., 2024）。將 VR 導入教學中，學生藉由與虛擬物件互動與探索新環境，形成獨特的實作經驗與主動構建知識，並對於學習能有更高程度的參與，提升對教材的理解度以及學習興趣（Marougkas et al., 2024）。

VR 技術已能模擬真實生活情景並具有各種互動與操作功能，其具有提供多感官刺激、呈現重複與關鍵場景、有效降低大型活動演練成本等特性，更能提升災害演練與危險應變訓練之效益，因此 VR 設備已逐漸應用於各種災害預防演練之課程（Shen et al., 2019; Paszkiewicz et al., 2021）。但 VR 防災教育課程的主題多以火災避難為主，對於地震演練相關內容較少，且參與者多以大學生與成人為主（Hsu et al., 2013; Shu et al., 2019; Wibisono, 2019; Feng et al., 2020），但兒童在面臨災害時相較於成人更容易受傷與無助，因此我們更應關注兒童對於地震避難之學習（Feng et al., 2018），藉由 VR 的沉浸與互動功能提升兒童相關知識與技能（Jiang & Li, 2022）。

2.2. 不同 VR 學習模式對參與者之影響

VR 虛擬實境裝置從桌上型電腦發展為頭戴式顯示器（Head Mounted Displays, HMD），過去因技術與成本考量，學校採用之 VR 學習系統多以電腦螢幕呈現 3D 立體虛擬環境，並採用鍵盤與滑鼠操控進行互動（Makransky & Petersen, 2019），此種形式之 VR 學習模式藉由視覺畫面呈現與互動性提供臨場感，形成與傳統多媒體不同之學習方式，給予學生真實的學習經驗，能有效提升學習者的學習興趣、動機與參與度，以及在醫學、數學與建築安全等不同領域的學習成效（Huang et al., 2010; Xu & Ke, 2016; Pham et al., 2018）。

現今科技進步，VR 的頭戴式顯示器已逐步降低體積與成本，並具有更多的互動功能（Soliman et al., 2021），與一般桌上型 VR 系統相比，頭戴式 VR 提供更真實的第一人稱視角，並採用手把操控提升參與者的臨場感與即時互動性，具有更好的沉浸體驗（Bower & Sturman, 2015; Marougkas et al., 2024）。而針對大學與高中生的研究指出，頭戴式 VR 系統能塑造更真實的環境，具有較高沉浸感與即時互動性，能提升學生在歷史與 STEM 等領域的學習表現（Calvert & Abadia, 2020; Johnson-Glenberg et al., 2020），且對於地震防災教育學習也有較佳的學習自我效能與成效（Shu et al., 2019），但目前缺乏不同 VR 學習模式對於國小生在地震防災學習影響之研究，故本研究將對此進行探討，以了解頭戴式 VR 系統對教學現場之效益。

2.3. 情境興趣與 VR 相關研究

興趣能影響個人行為，而由環境引起的情境興趣有助於提升學習表現（Yu et al., 2019），Hidi（2006）指出教師能透過不同的教學方法或教材，建構特殊情境引發學生的情境興趣，增進學生在課堂的注意力、互動以及對學習的好奇心，使其學習更多知識。在各式科技輔助學

習中，提高學生的情境興趣有助於線上學習的參與度（Sun & Rueda, 2012），對於 AR/VR 等學習系統，其比一般的數位系統更能提升學生的情境興趣，進而提高學習動機與課堂互動，並培養學生對於不同領域知識的長期學習興趣（Ainley, 2007; Yu et al., 2019; Lin et al., 2021）。

情境興趣分為「探索意圖、即時享受、新穎性、注意需求、挑戰性和整體興趣」六個面向（Chen et al., 1999），其可用來探索學生者使用創新科技的學習經驗，以了解學習者對於 AR/VR 等學習設備或教材的興趣來源，並評估系統的學習效益（Yu et al., 2019; Lin et al., 2021），因此本研究藉由情境興趣探國小學生使用不同 VR 學習模式之感受，以提供未來實踐之建議。

3. 研究設計

3.1. 實驗設計與研究參與者

本研究為了解頭戴式 VR 學習模式與桌面式 VR 學習模式對國小生在地震防災課程學習之影響，以及其情境興趣之差異，故採用準實驗研究法，分為實驗組（使用頭戴式 VR 設備進行學習）與對照組（使用桌上型電腦式 VR 設備進行學習），學習教材採用教育大市集提供之 VR 教材軟體，其分為桌上型電腦 VR 版本與頭戴式 VR 版本，兩種版本學習內容皆一致。

本次研究對象為臺灣北部某國小六年級學生（平均年齡 12 歲），其實驗組 50 人（兩個班），對照組 48 人（兩個班）。實驗組學生已有頭戴式 VR 設備操作與進行學習之經驗，對於手把操作與 VR 教材互動進行學習並無障礙；對照組學生熟悉桌上型電腦與該版本 VR 教材的操作。

圖 1

實驗組-頭戴式 VR 學習模式



圖 2

對照組-桌面式 VR 學習模式



3.2. 研究工具

3.2.1.頭戴式 VR 設備與學習教材

本次採用頭戴式 VR 設備為 Meta Quest 2，內建高解析度的立體成像螢幕，以及採用無線通訊手把，讓使用者無需串聯其他設備，提高移動性與沉浸感。在設備佩戴上，Meta Quest 2 可穿戴眼鏡後再配戴頭盔進行體驗，對近視的參與者提供便利性（Meta, 2022）。Meta Quest 2 與其他 VR 設備相比具有較佳的新手引導以及可用性，其手把設計更符合人體工學（Han & Soares, 2023），因此對國小生使用操作上較為容易。

本次使用之 VR 教材軟體來自教育部建置之教育大市集，由臺灣藝術大學李俊逸教授團隊設計的「逃出吧！地震小尖兵」，內容包含認識地震成因與特性、避難防災模擬以及實作演練，藉由 VR 建構真實情境，提升學生對於地震的認識與避難防災技能（教育大市集，2021）。

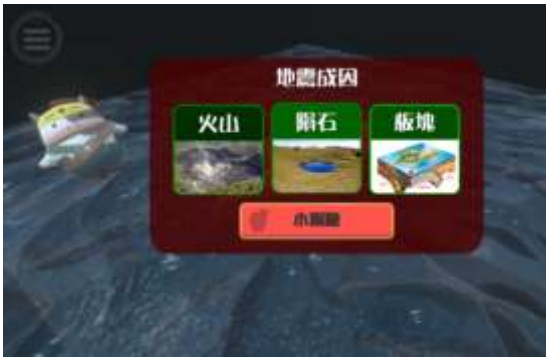
圖 3

頭戴式 VR 設備- Meta Quest 2



圖 4

教材學習畫面-逃出吧！地震小尖兵



3.2.2. 地震知能測驗卷

本研究為了解學生使用不同 VR 學習模式後的學習成效，依自然領域課程與 VR 教材內容進行題目設計，並邀請兩位 10 年以上國小自然領域教學經驗之教師協助審題，最終挑選出符合本研究目的之題目共 20 題，組成滿分 100 分之測驗卷，並於實驗前後進行施測。

在地震知能前測中，兩組的平均數和標準差如下：實驗組平均值為 40.10，標準差為 13.11；對照組平均值為 40.31，標準差為 17.46。根據統計，兩組前測之 Levene 檢定有顯著差異（$p < .05$），其 t 檢定結果，頭戴式 VR 學習模式組和桌面式 VR 學習模式之無顯著差異（$t(87.18) = -0.07$，$p > .05$），故兩組無先備經驗上之差異，也不影響後續實驗結果。

3.2.3. 情境興趣量表

本研究採用之「情境興趣量表」是由 Lin、Yu、Sun 與 Jong（2021）依據 Chen、Darst 與 Pangrazi（1999）提出之量表編製而成，再由研究者依研究情境進行題目敘述調整。本問卷由 6 個子構面組成，共 24 題，每個子構面有 4 個題敘述，其包含「探索意願」、「即時享受」、「新穎性」、「注意需求」、「挑戰性」和「整體興趣」。量表採用 6 點李克特量表，「1」代表「強烈不同意」，「6」代表「強烈同意」。情境興趣量表之 Cronbach's α 為 0.94。每個子構面之 Cronbach's α 值分別為 0.87、0.90、0.85、0.93、0.96 和 0.93，整份量表具有高度一致性。

4. 研究結果與討論

4.1. 不同 VR 學習模式之學習成效分析

首先，分析兩組學生使用 VR 學習後之成效，由表 1 得知兩組採用不同 VR 設備進行地震防災課程的學習，皆有顯著進步（實驗組 t（49）= 8.15，$p < .05$；對照組 t（47）= 7.44，$p < .05$），可得知將頭戴式 VR 或桌面式 VR 導入教學中，皆能有效提升學生的學習效果（Makransky & Petersen, 2019）。

表 1

兩組前後測之成對樣本 t 檢定

		前測		後測			
	n	*M*	*SD*	*M*	*SD*	*t*	*df*
實驗組 頭戴式 VR 學習模式	50	40.10	13.11	64.50	16.45	**8.15***	49
對照組 桌面式 VR 學習模式	48	40.31	17.46	56.67	18.99	**7.11***	47

*$p < .05$

再者，分析兩組學生差異，根據表 2 的獨立樣本 *t* 檢定結果顯示，進行頭戴式 VR 學習模式的學生在地震防災學習上有更為顯著的好表現（t（96）= 2.19，p < .05），藉由頭戴式 VR 提供更身臨其境的互動學習經驗，提高參與者對教材的理解程度（Marougkas et al., 2024），且與一般桌面式 VR 模式相比，頭戴式 VR 呈現更真實的環境與互動，更能培育學生的地震與相關領域的防災知能（Wibisono, 2019；Feng et al., 2020）。

表 2

兩組後測之獨立樣本 t 檢定

	實驗組（n = 50）頭戴式 VR 學習模式		對照組（n = 48）桌面式 VR 學習模式			
變項	*M*	*SD*	*M*	*SD*	*t*	*df*
單元後測成績	64.50	16.45	56.67	18.99	**2.19***	96

*p < .05

4.2. 不同 VR 學習模式之情境興趣分析

本研究分析兩組於情境興趣各面向之差異，其結果如表3所示。情境興趣和所有子構面的Levene檢定中除挑戰性（ $p > .05$）外，其它結果皆為顯著（ $p < .05$）。根據*t*檢定結果，整份情境興趣量表有顯著差異（ t（71.71） = 2.65，p < .05），表示學生使用頭戴式VR學習模式有更高的情境興趣，這也與過去研究呼應，頭戴式VR對於學生有更高的吸引力，能讓學生產生更多互動與課堂參與，進而提高防災避難的學習表現（Parong & Mayer, 2018；Feng et al., 2020）。

進一步分析各子構面，結果顯示探索意圖（ t（90.06） = 1.41，p = .16）、即時享受（ t（76.24） = 1.51，p = . 13），注意需求（ t（84.29） = 1.70, p = .09） 和挑戰性（ t（96） = 1.14, p = .26）四個子構面，兩組表現並無顯著差異，而新穎性（ t（76.76） = 2.12, p < .05）與整體興趣（ t（71.86） = 3.27, p < .05）呈現顯著差異，其中新穎性能呼應Lin等（2021）之研究，頭戴式VR比一般數位學習導覽系統能讓學習者感到新奇有趣，進而產生更大的學習興趣。與過去研究不同在於兩組挑戰性並無顯著差異，可能原因在於學生已熟悉相關設備操作與使用（Lin et al., 2020）。另外，在整體興趣部分，實驗組認為使用頭戴式VR能提供更多的樂趣，也產生更好的學習效果（Shu et al., 2019）。

表 3

兩組情境興趣之獨立樣本 t 檢定

	實驗組（n = 50）頭戴式 VR 學習模式		對照組（n = 48）桌面式 VR 學習模式			
變項	*M*	*SD*	*M*	*SD*	*t*	*df*
情境興趣	116.86	10.83	108.19	20.02	**2.65***	71.71
探索意圖	21.58	3.23	20.54	4.02	1.41	90.06
即時享受	21.70	2.59	20.60	4.32	1.51	76.24
新穎性	21.92	2.41	20.51	3.99	**2.12***	76.76
注意需求	21.14	3.67	19.58	5.19	1.71	84.29
挑戰性	8.08	3.98	7.29	2.72	1.14	96
整體興趣	22.44	2.82	19.67	5.18	**3.27***	71.86

*p < .05

4.3. 訪談與質性資料

本研究藉由觀察與訪談深入了解學生對於頭戴式 VR 學習模式之看法，學生給予以下回饋：

- *使用 VR 來學習，會有身歷其境的感覺，這種教學十分有趣*
- *覺得身如其境*
- *能自己親眼看到地震運動與感受搖晃*
- *平常都是翻開課本上課，現在可以實體玩跟體驗*
- *有震撼的畫面體驗*
- *跟角色有互動，而且可以親自實作，讓我感到很快樂*

學生之回饋呼應本研究對於情境興趣之分析結果，學生藉由頭戴式 VR 系統的互動性與沉浸性獲得更多樂趣與不同的學習體驗，也藉由互動實作實踐地震防災技能，展現出更好得學習效果。

5. 結論

本研究探討頭戴式 VR 學習模式與桌面式 VR 學習模式對國小生在地震防災課程學習之影響。藉由課程實踐中發現，頭戴式 VR 的高度沉浸與互動性，讓國小學生產生更高的情境興趣，並能有效提高學生在該主題學習上的表現，因此頭戴式 VR 裝置在防災避難等強調實作經驗的課程上有應用與推廣之價值。

本研究主要以國小高年級學生為研究對象，並將課程主軸限縮於地震防災，未來可增加不同年級學生與課程主題進行比較分析，以擴展頭戴式 VR 設備的教育價值。

誌謝

本研究感謝國科會(MOST 111-2410-H-A49-018-MY4 和 NSTC 112-2410-H-A49-019-MY3）與教育部（112-113 年 5G 新科技學習示範學校）計畫之支持，以及國立陽明交通大學互動學習科技與動機研究室（https://iltm.lab.nycu.edu.tw/）成員之協助，謹此致謝。

參考文獻

國家教育研究院（2019）。**十二年國民基本教育課程綱要-議題融入說明手冊**。

教育部（2023，3 月 25 日）。**5G 智慧共享教學沉浸式學習體驗「教育部 112-113 年 5G 新科技學習示範學校計畫」全面啟動**。。https://www.edu.tw/News_Content.aspx?n=9E7AC85F1954DDA8&s=A65C46E37F430ED3

教育大市集（2021，11 月 24 日）。**逃出吧！地震小尖兵**。VR 教材開發計畫。https://market.cloud.edu.tw/resources/web/1806342

Ainley, M. (2007). Being and feeling interested: Transient state, mood, and disposition. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 147–163). Elsevier Academic Press. https://doi.org/10.1016/B978-012372545-5/50010-1

Bower, M., & Sturman, D. (2015). What are the educational affordances of wearable technologies? *Computers & Education, 88*, 343-353. https://doi.org/10.1016/j.compedu.2015.07.013

Calvert J, & Abadia, R. (2020) Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Computers & Education, 159*(6), Article 104005. https://doi.org/10.1016/j.compedu.2020.104005

Chen, A., Darst, P. W., & Pangrazi, R. P. (1999). What constitutes situational interest? Validating a construct in physical education. *Measurement in Physical Education and Exercise Science, 3*(3), 157–180. https://doi.org/10.1207/s15327841mpee0303_3

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education, 127*, 252-266. https://doi.org/10.1016/j.compedu.2018.09.002

Feng, Z., González, V. A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & CabreraGuerrero, G. (2020). An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Advanced Engineering Informatics, 45*, Article 101118. https://doi.org/10.1016/j.aei.2020.101118

Han, Y., Zhang, Y., & Soares, M. M. (2023). The usability assessment of Meta Quest2 and recommendations for improving. In Marcus, A., Rosenzweig, E., & Soares, M.M. (Eds.), *Design, User Experience, and Usability* (pp. 184-193). Springer, Cham. https://doi.org/10.1007/978-3-031-35702-2_13

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review,1*(2), 69–82. https://doi.org/10.1023/A: 1016667621114

Hsu, E. B., Li, Y., Bayram, J. D., Levinson, D., Yang, S., & Monahan, C. (2013). State of virtual reality based disaster preparedness and response training. *PLoS Currents,* Article *5*.

Huang, H. M., Rauch, U., & Liaw, S. S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education, 55*(3), 1171–1182. https://doi.org/10.1016/j.compedu.2010.05.014

Jiang, J., & Li, X. (2022). A research on wisdom classroom teaching supported by virtual reality. In Hong, W., & Weng, Y. (Eds.), *Computer Science and Education* (pp. 575-584). Springer. https://doi.org/10.1007/978-981-99-2446-2_52

Johnson-Glenberg, M. C., Ly, V., Su, M., Zavala, R. N., Bartolomeo, H., & Kalina, E. (2020, June 21-25). *Embodied agentic STEM education: Effects of 3D VR compared to 2D PC*. [Conference presentation]. 2020 6th International Conference of the Immersive Learning Research Network Online, United States. https://doi.org/10.23919/iLRN47897.2020.9155155

Koolivand, H., Shooreshi, M. M., Safari-Faramani, R., Borji, M., Mansoory, M. S., Moradpoor, H., Bahrami, M., & Azizi, S. M. (2024). Comparison of the effectiveness of virtual reality-based education and conventional teaching methods in dental education: A systematic review. *BMC Medical Education, 24*(1), Article 8. https://doi.org/10.1186/s12909-023-04954-2

Lin, H. C.-S., Yu, S.-J., Sun, J. C.-Y., & Jong, M. S. Y. (2021). Engaging university students in a library guide through wearable spherical video-based virtual reality: Effects on situational interest and cognitive load. *Interactive Learning Environments, 29*(8), 1272–1287. https://doi.org/10.1080/10494820.2019.1624579

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education, 134*, 15–30. https://doi.org/10.1016/j.compedu.2019.02.002

Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2024). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. *Multimedia Tools and Applications, 83*(6), 18185-18233.

Meta(2022). *Meta – VR Headsets, Smart Displays and AR Glasses.* https://www.meta.com/tw/en/

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241

Paszkiewicz, A., Salach, M., Dymora, P., Bolanowski, M., Budzik, G., & Kubiak, P. (2021). Methodology of implementing virtual reality in education for industry 4.0. *Sustainability, 13*(9), Article 5049. https://doi.org/10.3390/su13095049

Pesce, M. (1995). *VRML browsing and building cyberspace*. New Riders Publishing.

Pham, H. C., Dao, N., Pedro, A., Le, Q. T., Hussain, R., Cho, S., & Park, C. S. I. K. (2018). Virtual field trip for mobile construction safety education using 360-degree panoramic virtual reality. *International Journal of Engineering Education*, *34*(4), 1174-1191.

Shen, H., Zhang, J., Yang, B., & Jia, B. (2019). Development of an educational virtual reality training system for marine engineers. *Computer Applications in Engineering Education, 27*(3), 580–602. https://doi.org/10.1002/cae.22099

Shu, Y., Huang, Y. Z., Chang, S. H., & Chen, M. Y. (2019). Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Reality*, *23*, 437-446. https://doi.org/10.1007/s10055-018-0376-x

Soliman, M., Pesyridis, A., Dalaymani-Zad, D., Gronfula, M., & Kourmpetis, M. (2021). The application of virtual reality in engineering education. *Applied Sciences*, *11*(6), Article 2879. https://doi.org/10.3390/app11062879

Sun, J. C. Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British journal of educational technology*, *43*(2), 191-204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

Wibisono, R. A. (2019). Self-evacuation drills by mobile virtual reality application to enhance earthquake preparedness. *Procedia Computer Science*, *157*, 247-254. https://doi.org/10.1016/j.procs.2019.08.16

Xu, X., & Ke, F. (2016). Designing a virtual-reality-based, gamelike math learning environment. *American Journal of Distance Education*, *30*(1), 27-38. https://doi.org/10.1080/08923647.2016.1119621

Yu, S. J., Sun, J. C. Y., & Chen, O. T. C. (2019). Effect of AR-based online wearable guides on university students' situational interest and learning performance. *Universal Access in the Information Society*, *18*, 287-299. https://doi.org/10.1007/s10209-017-0591-3

生成式 AI 聊天機器人結合自我調節學習理論應用於大專院校學習者之實作專題課程

Implementation of a Generative AI Chatbot Integrated with Self-regulated Learning Theories for University Students in Project-based Courses

徐唯芝[1]，孫之元[2*]
[1, 2] 國立陽明交通大學教育研究所
* jerrysun@nycu.edu.tw

【摘要】 本研究探討生成式 AI 聊天機器人結合自我調節學習理論應用於大專院校學習者之音訊編輯實作專題課程，對其自我調節與學習投入之影響，研究參與者為 55 位大專院校學習者，隨機分派至生成式 AI 聊天機器人結合自我調節學習理論組，共 28 位，以及生成式 AI 聊天機器人學習組，共 27 位。研究經過三週的教學實驗後發現，控制前測差異後，在學習成效的各成果變量上並未觀察到顯著的組間差異。然而，實驗組在課程完成率和後測成績上呈正向成長，未來的研究可以探索如何進一步優化這種教學結合方式。研究建議應針對學習者的個別差異，進行更細緻的教學設計。

【關鍵詞】 聊天機器人；生成式 AI；自我調節學習；實作專題課程；大學生

Abstract: *This study explored the impact of integrating generative AI chatbots with self-regulated learning theory in college-level audio editing courses. Fifty-five students were divided into two groups for a three-week trial. Results showed no significant differences in learning outcomes after controlling for pre-test scores. However, the experimental group demonstrated slight improvements in course completion and post-test scores, indicating potential benefits of the AI-enhanced educational approach. Future research should consider individual learner differences for personalized instruction.*

Keywords: Chatbot, Generative AI, Self-Regulated Learning, Project-Based Learning , University Education

1. 前言

生成式 AI 聊天機器人的出現，象徵著教育技術的轉型時代，成為一個複雜的人工智能技術與教學理論交會的節點。AI 在教育中的潛力探索已不是新話題；然而，結合自我調節學習理論則為生成式 AI 在教育上的應用帶來能促進其發展的效益。在近年生成式 AI 聊天機器人的研究中，不僅局限於專業領域，教育系統的結構中廣泛應用，這些應用的廣泛也反映了朝向適性化學習環境的轉變，學習者不再是信息的被動接收者，而是他們積極參與者，其中，自我調節學習是一個賦予學習者自覺監控和控制他們學習過程的理論框架，對於學生之適性化學習歷程至關重要。大專院校之實作專題課程（Project-Based Learning, PBL），在教學上經常需要教師個人化的輔導學生實作上所遇到的難題，當學生在沒有教學者在場的情境，若以生成式 AI 聊天機器人作為其自主學習之輔助工具，學習者在實作問題中尋求輔助與解答，並獲得生成式 AI 聊天機器人連續的反饋和改進循環，與自我調節學習（self-regulated learning, SRL）的基本原則相呼應，並在此過程中，期望除了提升學習成效外，亦能促進學習者之學習投入度與自我調節學習能力。

2. 文獻回顧

2.1. 生成式 AI 聊天機器人

根據相關文獻的綜述，生成式AI聊天機器人在教育領域的應用越來越廣泛。研究指出，AI聊天機器人在傳統教學中可以提供個性化的教學支持，並對提升學習成就、學習樂趣以及降低學習焦慮具有顯著效果（Hsu et al., 2023）。此外，聊天機器人也能滿足醫護教育中特定需求，提供個人化學習服務，並促進臨床情境中的學習，這對於促進專業知識的獲得與自主學習能力的提升尤為重要（莊宇慧等，2021）。現代的聊天機器人融合了專家決策機制，能更好地輔導學生在學習過程中建構知識，這種基於專家決策的聊天機器人（EDM-chatbot）與傳統的提供資訊型聊天機器人（C-chatbot）相比，能顯著提高學生的學習成效和學習享受，同時減少學習焦慮（Hsu et al., 2023）。這些研究反映出生成式AI聊天機器人在教育領域中的發展，已從單純的信息提供者轉變為決策支持者。生成式AI聊天機器人的運用已顯示其對促進成人學習者學業成就與自我效能有正面影響（Abbas et al., 2022）。本研究針對大學生自主學習所設計之教材，亦能幫助學生在沒有教師的輔導下完成實作專題作品。

2.2. 自我調節學習

本研究探討了生成式 AI 聊天機器人如何輔助大專院校學習者於線上課程中進行自我調節學習（SRL），並專注於學習目標的設定過程。過去研究發現，透過與以 SMART 目標設定策略為基礎的聊天機器人互動，學生能夠更清晰地界定學習目標並提高對設定學習目標的自覺性（Du et al., 2021），在該研究中學生反映，聊天機器人活動不僅幫助他們明確學習目標，同時也提高了他們在學習開始前設定學習目標的意識。學生期望未來的聊天機器人能夠在他們的學習過程中提供更多支持。例如，透過提醒學生重要的事項和最新的學習任務，幫助他們根據課程進度調整學習計劃；提供更多的學術資源，協助學生在學習過程中擁有更完善的學習輔助。根據學生的期待，未來的聊天機器人應該能夠在整個學習過程中提供持續的支持，包括跟蹤學習進度、提供學習資源和即時回饋，根據自我調節學習理論之機制鼓勵學生反思和自我評估，將聊天機器人結合自我調節學習理論運用於大專院校學習者的實作專題課程中，這對於學生學習動機與持續參與至關重要（Abbas et al., 2022）。

3. 研究方法

3.1. 研究對象

本研究採準實驗研究法，以 55 位大專院校相同科系之學習者作為學習對象，受試者皆未修習過音訊編輯相關課程之經驗，將受試者隨機分派為實驗組，學習過程中使用生成式 AI 聊天機器人結合自我調節機制之教材輔助學習與專題實作，共 28 人，其中男性 8 人，女性 20 人，其餘 27 人則為控制組，學習過程中使用生成式 AI 聊天機器人教材輔助學習與專題實作，其中男性 7 人，女性 20 人。

3.2. 研究流程

該課程教授十六週的時間，第一週時兩組學生均須填寫音訊編輯測驗卷，在最後三週時將進行課程專題實作，學習者需要在實作過程中自主學習，並於此三週期間開放使用生成式 AI 聊天機器人作為輔助教材以及專題實作之問題解決工具，使用教材學習三週後，兩組學生均須填寫學習投入量表、自我調節量表以及音訊編輯後測驗卷。

3.3. 研究工具

3.3.1. 學習投入量表

本「學習投入度量表」是修改自 Sun 與 Rueda（2012），該篇研究參照 Fredricks 等（2005）的量表進行編製。採用 Likert 式 5 點量尺（5 為非常同意；4 為同意；3 為普通；2 為不同意；1 為非常不同意。兩組學習者均於課後填寫該問卷。

3.3.2. 自我調節量表

本「自我調節量表」是修改自楊惠婷等（2023），該篇研究參照 Barnard 等（2009）的量表改編而成。採用 Likert 式 6 點量尺（6 為非常同意；5 為同意；4 為有點同意；3 為有點不同意；2 為不同意；1 為非常不同意）。兩組學習者均於課後填寫該問卷。

3.3.3. 音訊編輯測驗卷

本「音訊編輯測驗卷」是參考「跟 Adobe 徹底研究 Audition CC」一書（陳芳智，2016）之內容並由課堂教學者編寫測驗卷題目製作而成，測驗卷共 25 題，均為選擇題，滿分為 100 分。

3.3.4. 生成式 AI 聊天機器人

在本研究中，研究者以ChatGPT作為生成式AI之自然語言處理能力，結合LINE聊天機器人，旨在探索生成式AI聊天機器人結合自我調節學習機制在大學生實作課程之應用潛力。開發過程始於在LINE平台註冊開發者帳號，隨後創建一個聊天機器人並獲取相應的Channel ID、Channel Secret以及Access Token。使用Python語言開發了機器人的服務器端應用，負責解析來自LINE平台的Webhook請求並生成回饋訊息。為了實現機器人的自然語言處理功能，需要將OpenAI提供的ChatGPT API整合至開發的機器人系統中，確保機器人能夠將收到的用戶指令傳遞給ChatGPT並將其回應傳回給學習者。下圖1為實驗組（生成式AI聊天機器人結合自我調節學習機制）與控制組（生成式AI聊天機器人）之介面示意圖。

圖 1

介面示意圖



3.4. 資料分析分法

本研究採用 R 語言進行資料分析。目標是透過單因子共變異數分析（Analysis of Covariance, ANCOVA）探討音訊編輯前測分數及在實驗組與控制組中對於五項學習成果變量，包含學期成績、音訊編輯後測得分、課程完成率、自我調節學習、學習投入度，計算其各組之間平均分數是否存在顯著差異，以回答研究的主要問題。

4. 研究結果

兩組之自我調節學習、學習投入、課程完成率、學期成績以及音訊編輯測驗之描述性統計顯示，偏態與峰度皆介於正負 3 之間，各組皆呈現常態分布。接者，透過 Bartlett 檢驗不同組別的變異數是否相同，也就是同質性檢驗，亦均符合同質性假設。故能進行後續統計分析。首先，進行組內回歸係數同質性假設，結果顯示五個成果變量皆不拒絕同質性假設，故可進行 ANCOVA 分析，分析結果顯示，在控制了音訊編輯前測分數的情況下，在學期成績、音訊

編輯後測、課程完成率、自我調節學習以及學習投入這五個成果變量中，實驗組與控制組均無顯著差異，將群組排除前測差異影響，調整後的變項平均數如表 1 所示，進一步將各組均值結果呈現於圖 2 中，可以看出雖然在統計上並無顯著差異，但課程完成率與後測成績實驗組略高於控制組的傾向。

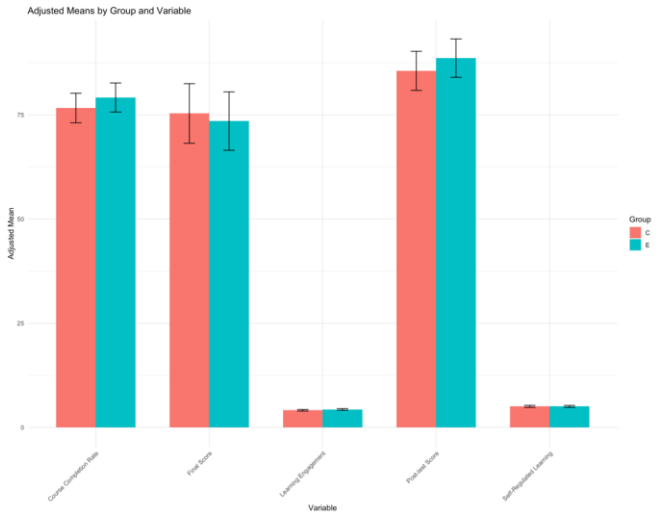
表 1

群組調整後的變項平均數

變項	群組	調整後平均數	標準誤	95%信任區間下限	95%信任區間上限
學期成績	控制組（C）	75.33	3.56	68.18	82.47
	實驗組（E）	73.51	3.50	66.49	80.52
音訊編輯測驗卷	控制組（C）	85.54	2.33	80.86	90.23
	實驗組（E）	88.58	2.29	83.98	93.18
課程完成率	控制組（C）	76.63	1.77	73.08	80.18
	實驗組（E）	79.15	1.74	75.67	82.64
自我調節學習	控制組（C）	5.06	0.12	4.81	5.30
	實驗組（E）	5.05	0.12	4.81	5.29
學習投入	控制組（C）	4.13	0.10	3.93	4.32
	實驗組（E）	4.31	0.10	4.12	4.51

圖 2

組別之間均值差異



5. 結論與建議

在研究結果中顯示，在控制了前測差異後，在成果變量中均無顯著差異，但考慮到實驗組在某些學習成果上呈現輕微的正面趨勢，未來研究可進一步探討如何優化結合 AI 聊天機器人的教學策略，以增強其對學習成效的正面影響，而針對學習者的個別差異，如先備知識，應在教學研究中被考量與設計，以便更精準地對學習策略進行適性化調整。本研究進行時間為三週，未來進行相關研究，建議進行長期追蹤研究，以觀察 AI 聊天機器人結合自我調節學習理論對學習者長期學習成效的影響。本研究亦發現，雖然聊天機器人在提供資訊和答疑方面表現突出，但在與學生建立更深層次互動方面仍有進步空間。因此，未來的策略應包括改

進聊天機器人的互動性，讓其更貼近學生的需求，並結合教學策略來增強其在自我調節學習中的作用（Crawford et al,. 2024）。

誌謝

本研究感謝國科會(MOST 111-2410-H-A49-018-MY4 和 NSTC 112-2410-H-A49-019-MY3)之支持。

參考文獻

莊宇慧、陳怡婷和郭倩琳（2021）。 聊天機器人於臨床護理教育的設計與運用。**護理雜誌，68**（6），19-024。https://doi.org/10.6224/JN.202112_68(6).04

陳芳智（2016）。**跟 Adobe 徹底研究 Audition CC(CD-ROM)**。上奇資訊。

楊惠婷、孫之元和林毓晏（2023）。數位教材回饋及信號設計對奈米科技學習自我效能與自我調節之影響。**數位學習科技期刊，15**（3），81-114。https://doi.org/10.53106/2071260X2023071503004

Abbas, N., Whitfield, J., Atwell, E., Bowman, H., Pickard, T., & Walker, A. (2022). Online chat and chatbots to enhance mature student engagement in higher education. *International Journal of Lifelong Education, 41*(3), 308–326. https://doi.org/10.1080/02601370.2022.2066213

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O. & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *Internet and Higher Education, 12*(1), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005

Crawford, J., Allen, K. A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education*, 1–15. https://doi.org/10.1080/03075079.2024.2326956

Du, J. H., Huang, W. J., Hew, K. F., & Ieee. (2021, Dec 05-08). *Supporting students goal setting process using chatbot implementation in a fully online course.* Paper presented at the IEEE International Conference on Engineering, Technology and Education (IEEE TALE), Cent China Normal Univ, Wuhan, PEOPLES R CHINA.

Fredricks, J. A., Blumenfeld, P., Friedel, J. & Paris, A. (2005). School engagement. In K. A. Moore & L. H. Lippman (Eds), *What do children need to flourish? Conceptualizing and measuring indicators of positive development* (pp. 305–321). Springer Science and Business Media.

Hsu, T.-C., Huang, H.-L., Hwang, G.-J., & Chen, M.-S. (2023). Effects of Incorporating an Expert Decision-making Mechanism into Chatbots on Students' Achievement, Enjoyment, and Anxiety. *Educational Technology & Society, 26*(1), 218–231. https://doi.org/10.30191/ETS.202301_26(1).0016

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology, 43*(2), 191–204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

AI 輔助之問題導向自主學習對於職前教師口語表達學習之影響

The Effects of AI-Assisted Problem-based self-directed learning model on Preservice Teachers' Oral Expression Education

王俊傑
國立屏東大學教育學系
jie809@mail.nptu.edu.tw

【摘要】 在問題導向式學習(Problem-based Learning)模式中，學生大多藉由合作或協作方式完成教師所賦予之任務。期間必須花費大量心力於較低階之蒐集、統整資訊上，若能藉由 AI (Artificial Intelligence)的加入，以 ChatGPT(Chat Generative Pre-trained Transformer）作為學習同伴，分擔較低階之任務，而將心力聚焦於推理和預測之類高階的思維，應能有效增進相關學習成效。因此，本研究提出結合 AI 輔助之問題導向自主學習模式，探討 AI 的加入，對於職前教師於中文口語表達課程相關學習成效之影響。同時，也將蒐集學生使用 AI 之學習行為模式，進行歸納與分析，以作為 AI 融入教學設計之參考。

【關鍵詞】問題導向式學習、AI 人工智慧、ChatGPT、口語表達、職前教師

Abstract: *In Problem-based Learning(PBL)mode, students mostly cooperate or collaborate to accomplish the tasks given by the teacher. During this period, much effort must be spent on lower-order information collection and consolidation. Suppose AI(Artificial Intelligence)can be added, such as using ChatGPT as a learning companion to share the lower-order tasks and focusing on higher-order thinking such as inference and prediction. In that case, it should be able to improve the related learning effectiveness effectively. Therefore, this study proposes an AI-assisted problem-based self-directed learning model to investigate the effects of AI on pre-service teachers' learning outcomes in the Chinese oral expression course. Meanwhile, we will also collect and analyze students' learning behaviors for reference in the design of AI-assisted teaching.*

Keywords: Problem-based Learning, AI, ChatGPT, Oral Expression, Pre-Service Teacher

1. 前言

1.1. 研究背景與動機

問題導向式學習(Problem-based Learning)模式已被廣泛運用於教學上並獲致良好的成效。在 PBL 模式中，學生大多藉由合作或協作方式完成教師所賦予之任務。然而不論是合作或協作模式，學生皆需要大量蒐集、統整資訊以界定學習或解決問題。在傳統 PBL 模式之下，學生需要花費大多數的時間在此階段，若能將此類較低階的活動交由 AI (Artificial Intelligence)協助，而將時間聚焦於推理和預測等高階思維上，是否能增進學生協同學習、解決問題的效率，進而提升整體學習成效？

Ruiz-Rojas(Ruiz-Rojas et al., 2023)等人指出 ChatGPT 是教師最常使用的 AI 工具，使用率達到 95.2%在教育用途上，教師可以將其作為虛擬助理、回答學生問題或提供個人化支持。因此，本研究希望在 PBL 模式的教學之中，透過以 AI 為學習同伴之情境進行 PBL 學習，探討學生學習行為之改變，以及相關學習成效之差異。

1.2. 研究目的與問題

相較於「傳統 PBL 自主學習模式」,「AI 輔助之 P 自主學習模式」能否提升職前教師之中文口語表達學習成效、學習動機、自我效能、課堂參與度；以及能否降低其中文口語表達

之口說焦慮，並探究在此學習模式下，學生運用 AI 之學習行為。

2. 文獻回顧

2.1. 問題導向式學習(Problem-based Learning)

Vernon 和 Blake(1993)針對 PBL 的回顧性研究指出，學界普遍認為 PBL 優於傳統教學法。Prince(2004)的回顧性研究則指出 PBL 能使醫學院學生產生積極的學習態度、提高學習保留率、圖書館的使用率、教科書閱讀率、課堂出席率，甚至是更加理解學習的意義並非僅止於回憶學習內容。同時，Norman 和 Schmidt(2000)則點出而在實際執行 PBL 的教學時，若是能使用分組學習與合作學習架構(cooperative learning structures)，則效果會更顯著。

值得留意的是，PBL 在某些情況下容易導致負面效果(Prince, 2004)，例如沒有專業的指導者，或是讓學生自訂進度(selfpaced)與自主學習(self-directed learning)都容易對於學習產生負面影響，例如 Albanese 和 Mitchell(1993)指出，實施 PBL 學習的學生，其基礎科學測驗成績都低於傳統教學的學生。因此，Prince 的建議是教師要謹慎處理要求學生自我指導之程度。其次，相關研究多是針對醫學與資優教育，了解技能較差的學習者如何達成學習目標也是需要研究的重點之一(Hmelo-Silver, 2004)。

至於結合 PBL 模式用於口語表達學習之研究，Lin(2015)在國小的英語字彙學習中採用了 PBL 教學模式，研究者指出雖然在字彙測驗中，兩組並沒有顯著差異，但在寫作與口語表達上，採用 PBL 的組別皆能有較佳的表現。諸如實驗組使用較多教材以外的字彙、超出 2000 字之字彙水準的比率較高，在自我報告中表示獲得足夠會話練習機會等等。整體而言，研究者認為 PBL 提供了正向的人際溝通效應、形成更持續的學習氛圍，以及營造了具鷹架支持的學習環境。Zou 等學者(Zou et al., 2023)在針對環景影像虛擬實境 (Spherical video-based virtual reality, SVVR)用於語言教育之回顧性研究中提到，SVVR 結合 PBL 鷹架策略是有效的方式。

2.2. AI (Artificial Intelligence)作為學習同伴

Xie (Xie, Hwang, & Wong, 2021)等學者指出，隨著深度學習(Deep Learning)技術在 AI 領域的快速發展，AI 在教育領域扮演的角色也隨之改變，未來可能擔任智慧教師(intelligent tutor)、智慧學生(intelligent tutee)、智慧學伴(intelligent learning partner)或決策顧問(policy making advisor)等角色。Hwang 等學者(Hwang et al., 2020)也指出，讓 AI 擔任學習同伴的優點在於可以幫助學習者有效率的蒐集和分析資訊，這能讓學習者專注於推理和預測之類高階的思維。Mollick 與 Mollick (2022)則針對 ChatGPT，主張 AI 聊天機器人能夠帶來三種新的學習模式：提升學習遷移(improving transfer)、打破深度解釋之假象(breaking the illusion of explanatory depth)，以及能批判性的評價所獲得之解釋 (critically evaluate explanations)。而在語言學習中，學習遷移向來是重要的議題；另外，在學習第一語言時，學生也容易產生深度解釋之假象。本研究即試圖探究，在 PBL 模式中結合 AI 學伴，對於學生之學習遷移與打破深度解釋假象之影響。

有關 AI 做為學習同伴之相關研究，Liang 等學者(Liang et al., 2024)在結合了 AR 遊戲(Alternate Reality Games, ARGs)的博物館學習課程中，加入了聊天機器人(Chatbot)作為學習中的學習夥伴。由於傳統的糾正回饋訊息無法滿足學生的個別需求，而 AI 聊天機器人正可以支援解決個別問題。因此，結果顯示實驗組的學生在後設認知意識、情感投入和行為投入上皆有顯著提高，也能促進學生的自主學習(self-regulated learning)和雙循環學習(Double-Loop Learning)。

3. 研究方法

3.1. 研究對象

本研究以臺灣南部某大學教育系一年級兩個班級作為研究對象，兩班各約 45 位學生，皆為修習師資培訓學程之學生，平均年齡約 18 歲。隨機擇取一班為實驗組，採用 AI 輔助之問題導向自主學習模式；另一班為控制組，採用傳統問題導向自主學習模式。

3.2. 實施課程內容

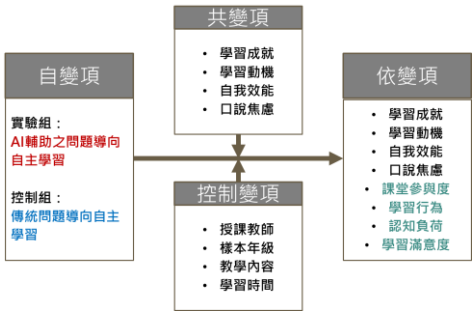
實驗預計配合國音及口語表達課程進行，共計 16 週，每週 2 小時。課程內容包括中文基礎發音知識(聲、韻、調)與口語表達(說故事、朗讀、演講、教學)。

3.3. 研究設計

本研究採準實驗研究法，主要探討在 PBL 學習模式之下，導入 AI 輔助機制對於學生學習中文口語表達之影響。同時也蒐集課堂參與度、學習行為、認知負荷、學習滿意度等資訊，藉以分析在此教學模式下，高低成就學生之學習行為與感受差異。

圖 1

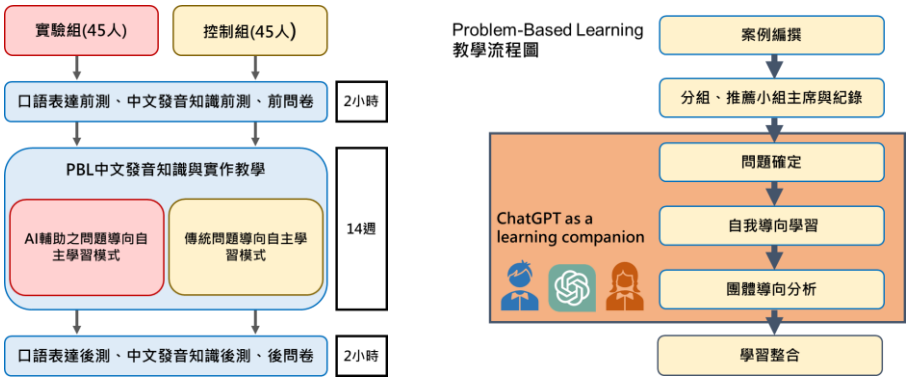
研究架構圖



本課程教學融入 PBL 教學模式，以問題開啟教學，讓學生藉由理解問題，界定其學習需求、統整資訊，既而將所學應用於解決一開始的問題。兩組不同之處在於實驗組在進行專題設計、活動規劃、小組討論、反思修正等學習階段時，將加入 ChatGPT 作為學習同伴。實驗組也將由教師帶領，根據 Bsharat(2023)等學者所指出之提示語(prompt)鷹架，與 ChatGPT 進行對話。控制組則維持傳統 PBL 策略，不刻意引導使用 AI 輔助，而是由學生自行決定如何蒐集資訊、設計問題與進行討論和反思。

圖 2、圖 3

實驗流程圖、教學流程圖



3.4. 研究工具

(1) 學習前、後測驗：國音學常識與知識測驗，皆由同一位資深授課教師命題。

(2) 前、後問卷：學習動機、自我效能、口說焦慮、課堂參與度、認知負荷、學習滿意度。

(3) 口語表達能力：根據朗讀聲韻正音檢核表，針對學生口說表達進行發音準確度之分析。

(4) AI 使用行為資料蒐集：蒐集學生運用 ChatGPT 之相關紀錄，分析學生使用 AI 輔助之情況。

4. 預期實驗結果

預期結合 AI 輔助之問題導向學習模式，學生中文口語表達之學習成效、學習動機、自我效能、課堂參與度皆能有所提升，口說焦慮能有所下降。另外也將觀察並分析學生運用 AI 輔助之學習行為，作為未來進一步進行 AI 介入之研究與教學參考。

參考文獻

Bsharat, S. M., Myrzakhan, A., & Shen, Z. (2024). Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4. *ArXiv:2312.16171*. https://doi.org/https://doi.org/10.48550/arXiv.2312.16171

Hwang, G., Xie, H., Wah, B. W., & Ga sevic, D. (2020). Vision, Challenges, Roles and Research Issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100001. https://doi.org/https://doi.org/10.1016/j.caeai.2020.100001

Hmelo-silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, *16*(3), 235–266.

Liang, H., Hwang, G., Hsu, T., & Yeh, J. (2024). Effect of an AI-Based Chatbot on Students' Learning Performance in Alternate Reality Game-Based Museum Learning. *British Journal Of Educational Technology*. https://doi.org/DOI10.1111/bjet.13448

Lin, L. (2015). The Impact of Problem-Based Learning on Chinese-Speaking Elementary School Students' English Vocabulary Learning and Use. *System*, *55*, 30–42. https://doi.org/https://doi.org/10.1016/j.system.2015.08.004

Mollick, E. R., & Mollick, L. (2022). New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments. *SSRN Electronic Journal*

Norman, G. R. (2000). Effectiveness of Problem-Based Learning Curricula: Theory, Practice and Paper Darts. *Medical Education*, *34*, 721–728. https://doi.org/https://doi.org/10.1046/j.1365-2923.2000.00749.x

PBLWorks. *Gold Standard PBL: Essential Project Design Elements A research-informed model for improving, calibrating, and assessing your practice.* PBLWorks. https://www.pblworks.org/what-is-pbl/gold-standard-project-design

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Ruiz-Rojas, L.I.; Acosta-Vargas, P.; De-Moreta-Llovet, J.; Gonzalez-Rodriguez, M. Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix. *Sustainability 2023, 15*, 11524. https://doi.org/10.3390/ su151511524

Savery, J. R. (2006). Overview of Problem-Based Learning: Definitions and Distinctions. *Interdisciplinary Journal of Problem-Based Learning*, *1*(1), 9–20. https://doi.org/10.7771/1541-5015.1002

Vernon, D. T. A., & Blake, R. L. (1993). Does Problem-Based Learning Work? A Meta-Analysis of Evaluative Research. *Academic Medicine*, *68*(7), 550–563

Xie, H., Hwang, G., & Wong, T. (2021). Editorial Note: From Conventional AI to Modern AI in Education: Reexamining AI and Analytic Techniques for Teaching and Learning. *Educational Technology & Society*, *24*(3), 85–88. https:// doi:10.30191/ETS.202207_24(3).0006

Zou, D., Jong, M. S. Y., Huang, X., Cheng, G., Hwang, G. J., & Jiang, M. Y. C. (2023). A systematic review of SVVR in language education in terms of the ADDIE model. *Interactive Learning Environments*, 1–26. https://doi.org/10.1080/10494820.2023.2277747

運用互動式視覺化工具促進人工智慧概念的理解：從自我調整學習的觀點

Using Interactive Visualization Tools to Enhance Understanding of Artificial Intelligence Concepts: A Perspective from Self-Regulated Learning

鄭年亨*
臺北醫學大學通識教育中心
* hercycheng@tmu.edu.tw

【摘要】 本研究旨在探討互動式視覺化工具對於醫學領域學生在學習人工智慧概念時自我調整學習的影響。本研究為使非資訊領域學習者能理解人工智慧概念，採用互動式視覺化工具引導學生觀察機器學習模型的效果。研究結果顯示，學生的內在目標導向和自我效能感有顯著提升，而測驗焦慮顯著降低。

【關鍵詞】 互動式視覺化；自我調整學習；人工智慧教育

Abstract: *This study aimed to investigate the impact of interactive visualization tools on self-regulated learning among medical students when they learn concepts of artificial intelligence. To help non-information learners understand AI concepts, interactive visualization tools were utilized to guide students in observing the effects of machine learning models. The results of the study showed significant improvements in students' intrinsic goal orientation and self-efficacy, along with a significant decrease in test anxiety.*

Keywords: Interactive visualization, self-regulated learning, artificial learning education

1. 前言

近年來人工智慧已經深入影響學生的學習與生活，可以預期當學生進入工作或研究場域時，人工智慧可能已經產生新型態的應用，對職場有本質上的變革。例如未來的醫學相關工作和研究將會大量運用人工智慧技術，促使醫學領域的學生培養跨領域能力，思考如何將新技術應用於醫學實踐中。換言之，非資訊領域學生即使無須編寫人工智慧程式，也需要理解人工智慧的科學原理，此為當前時代新的學習需求。

然而，對非資訊領域學生而言，這樣跨領域學習需求充滿挑戰。一方面，在學生端，非資訊領域學生可能當前尚未意識到人工智慧的重要性。特別是醫學領域學生常專注於專業的學習，一般而言對科技領域缺乏興趣，而難以連結人工智慧。另一方面，在課程端，人工智慧的抽象性也是一個挑戰；由於人工智慧的本質是數學，在不使用程式教學的情況下，學生很難探究其原理。若僅是讓學生學習使用人工智慧工具，學生雖能感受到人工智慧的便利性，但也僅止於訝異，難以理解其概念。

本研究目的在於發展一個以概念為導向的人工智慧學習模式，此模式將以非資訊領域學生為對象，不使用數學與程式的教學方式，使非資訊領域學生能理解人工智慧的原理。先前學者們建議，為了幫助學生更好地理解科技概念，可以採用互動式視覺化（interactive visualization）工具（周保男、薛曉琳，2017）。透過過人工智慧模型的視覺化，將抽象的科技概念轉化為直觀的圖像，並以互動式操作，使學生更容易理解與探索複雜的概念，從而提高他們的學習興趣和動機。為了解此模式對於非資訊領域學生是否有意願學習人工智慧，本論文也藉由一學期在台灣北部某大學的實證研究，探究此模式對於醫學領域學生的自我調整學習動機面向的變化。

2. 相關研究

互動式視覺化工具在各個 STEM 教育領域中被廣泛應用，包括數學（黃秋霞、楊志強，2020）、物理（Lontschar, Pietroszek, Humer, & Eckhardt, 2020）、化學（Fatemah, Rasool, & Habib, 2020）等。透過這些工具，學生能夠以更主動的方式探索學習環境，並具有更多的控制權。他們可以通過試誤法來嘗試解決問題，並利用策略找出最佳解決方案。這樣的學習過程不僅能夠提高學生對科學概念的理解，而且即使在失敗的探索中，學生仍然可以建構意義。由於人工智慧的概念通常涉及數學公式，對非資訊領域的學生來說理解其功效可能會有困難。互動式視覺化工具可以作為數學原理的替代方案，讓學生透過參數的調整直觀地觀察模型的效果，這對非資訊領域的學生來說尤其適用。

目前，許多資訊領域的研究者和工程師已經開發了各種人工智慧的互動式視覺化工具。例如 Fuchs 等學者（2019）開發了稱為 EduClust 的互動式視覺化工具，該工具主要模擬集群演算法的效果，透過動畫演示演算法的每個步驟，藉此讓學生探索不同集群參數對資料集的影響，同時可以查看偽程式碼（pseudocode）來理解技術細節。又如 Zhou 等學者（2020）開發了 AISpace 此套互動式視覺化工具，該工具目前提供多種人工智慧演算法的工具，包括搜尋、規劃、推理等。此工具直接以程式碼方式執行，讓學生觀察模擬動畫來理解人工智慧概念。雖然該研究的自評學習成效獲得正面結果，但誠如該學者所言，該研究仍然需要長期教學實踐研究來獲得更嚴謹的證據。

3. 學習活動設計

本研究所採用的學習活動如圖 1 所示。首先，每個機器學習的單元從應用實例開始，引入相關的人工智慧案例，讓學生意識到人工智慧在醫學或生活的實際應用，希望能讓學生連結其生活經驗，藉此引發學生的興趣。在簡短講解基本概念後，實施互動式視覺化（interactive visualization）工具的操作，並以學習單引導學生探索機器學習的數學模型。學生無需編寫程式或理解數學定理，只要通過資料輸入和參數設定，觀察模型的效果，就能理解機器學習的概念。最後，在期末專題中，學生被要求選擇他們感興趣的真實問題，並如同資料科學家一樣，運用人工智慧工具收集實際資料來訓練模型，以解決問題。

圖 1

學習活動設計

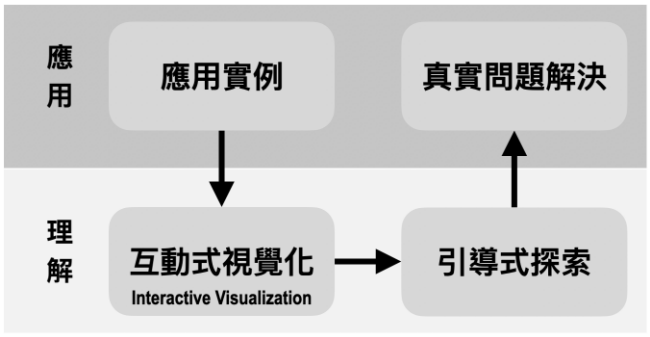
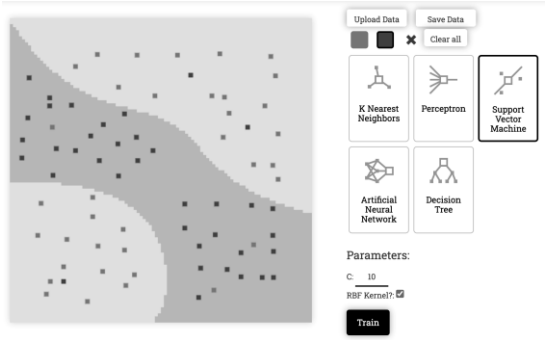


圖 2

教學工具



在機器學習的分類問題單元，本研究採用「機器學習遊樂場」此一免費線上學習資源（https://ml-playground.com/，請參見圖 2）。該平台提供了五種常見的分類演算法：K-鄰近演算法、感知器、支持向量機、類神經網路和決策樹。學生可以透過滑鼠點擊左側的資料區域，輸入橘色和紫色的訓練資料，每個資料點的 X 和 Y 座標即為該資料點的特徵。此外，學生還可以添加噪音資料，以測試不同的人工智慧演算法的穩健性。每個演算法都有不同的超參數可供調整，學生可以通過調整這些超參數來探究其對分類效果的影響。此平台提供了互

動式視覺化的學習環境，讓學生無需數學基礎知識與程式開發經驗，就能直觀地體驗不同機器學習演算法的模型效果。除此之外，過去很難理解的過度配適（overfitting，又譯過擬合）現象，也能透過觀察的方式了解該現象發生的原因。

4. 研究方法與結果

4.1. 研究問題

本研究擬從自我調整學習的角度探討互動式視覺化工具是否能提升醫學領域學生的學習動力。由於科學探究活動十分依賴於自主學習，學生在此類活動中需要有充足的學習目標，認同學習內容的重要性，也相信努力能夠影響結果，並有足夠的信心，才能夠完成學習任務，因此自我調整學習能夠解釋學生運用互動式視覺化工具探究機器學習概念的學習情況。

4.2. 研究對象

本研究為單組前後測設計，研究對象為台灣北部大學人工智慧導論修課學生，共 141 位大學生，其主修皆為醫學相關領域科系，無資訊相關科系。由於該課程為必修課，學生並非按自己興趣選修。

4.3. 研究工具

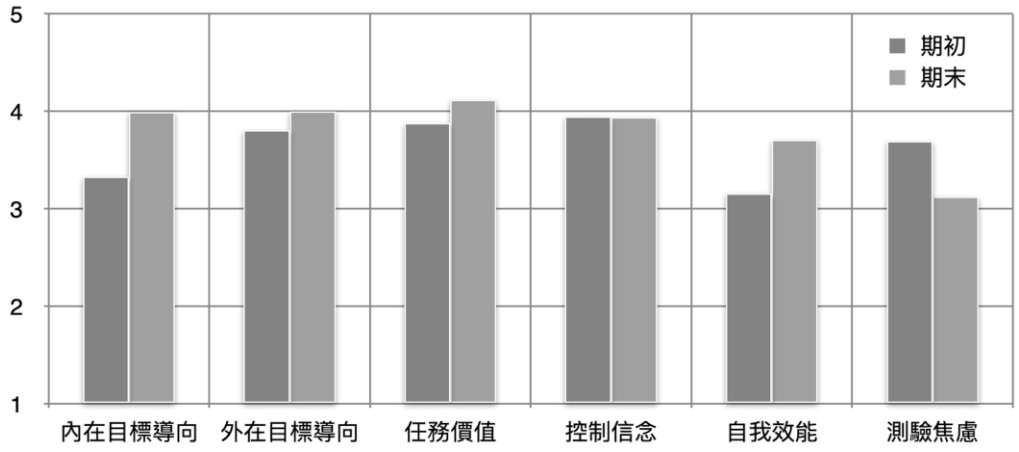
本研究採用 Pintrich 等人（1991）的學習激勵策略量表（Motivated Strategies for Learning Questionnaire, MSLQ）之動機分量表。該量表共 31 題，分為六個構面，分別為內在目標導向（4 題）、外在目標導向（4 題）、任務價值（6 題）、控制信念（4 題）、自我效能（8 題）與測驗焦慮（5 題）。此量表被廣泛運用在教育研究上，信效度良好。

4.4. 研究結果

研究結果如圖 3 所示。t 檢驗顯示內在目標導向（t=4.64, p<0.001）與自我效能（t=3.79, p<0.001）有顯著提升，測驗焦慮顯著降低（t=-3.29, p<0.01），其餘構面無顯著變化。首先，內在目標導向顯著提升顯示學生在學習過程中不再只追求好成績，也開始為了自我成長而學習人工智慧。再者，自我效能顯著提升可能是因為互動式視覺化工具能夠從簡單的互動中即時地給予學生正確的回饋，進而提高了自我效能感。最後，測驗焦慮顯著降低可能是因為互動式視覺化工具有助於學生更好地理解考試內容，從而減輕了焦慮感。

圖 3

研究結果



5. 結語

本研究旨在探討互動式視覺化工具對於醫學領域學生學習人工智慧概念時的自我調整學習的影響。課程透過引入機器學習概念的應用實例，以及利用互動式視覺化工具進行模擬科

學探究活動，學生能夠在不需要數學或程式背景知識的情況下，直觀地理解機器學習模型的效果。此外，透過期末專題的實際應用，學生被鼓勵運用所學知識解決真實世界的問題，從而將理論與實踐相結合。研究結果顯示，互動式視覺化工具的使用對於學生的內在目標導向、自我效能感和測驗焦慮都產生了顯著的影響。學生不僅開始關注自我成長，而非僅僅追求成績，還對自己對課程內容的理解展現出更高的自信心，並且不在擔心考試的結果。這些結果強調了運用互動式視覺化工具的科學探究活動有助於提升學生的自我調整學習的動機。

醫學與研究環境在可預見的未來將面臨許多人工智慧技術的應用，這些技術將改變醫學診斷、治療和研究的方式。因此，醫學領域的學生需要深入理解人工智慧的原理，以便能夠妥善地應用人工智慧產品。這種對人工智慧的理解將使醫學領域的學生能夠更有效地與資訊專家合作，共同研發下一代應用於醫學領域的人工智慧技術。透過與資訊專家的合作，他們可以更好地理解醫學問題與人工智慧技術之間的關聯，並共同開發解決方案。這也是我們推動醫學領域人工智慧教育下一階段的目標。

致謝

本研究受教育部「教學實踐研究計畫」經費補助（計畫編號 PGE1110468），特此致謝。

參考文獻

周保男、薛曉琳（2017）。互動式視覺科技導入國中 STEM 教育之研究-以 GeoGebra 輔助數學教學為例。**工程與科技教育學術研討會論文集**（頁 295-307）。社團法人台灣工程教育與管理學會。

黃秋霞、楊志強（2020）。國小數學科互動式體積 AR 輔助學習軟體應用之探討。**特殊教育發展期刊，70**，17-34。

Fatemah, A., Rasool, S., & Habib, U. (2020). Interactive 3D visualization of chemical structure diagrams embedded in text to aid spatial learning process of students. *Journal of Chemical Education, 97*(4), 992-1000.

Fuchs, J., Isenberg, P., Bezerianos, A., Miller, M., & Keim, D. (2019). Educlust: A visualization application for teaching clustering algorithms. In *Eurographics 2019: 40th Annual Conference of the European Association for Computer Graphics* (pp. 1-8). Eurographics Association.

Lontschar, S., Pietroszek, K., Humer, I., & Eckhardt, C. (2020). An Immersive and Interactive Visualization of Gravitational Waves. In *2020 6th International Conference of the Immersive Learning Research Network (iLRN)* (pp. 178-184). IEEE.

Pintrich, P. R., Smith, D. A. F., Duncan, T., & Mckeachie, W. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. The University of Michigan.

Zhou, C., Kuang, D., Liu, J., Yang, H., Zhang, Z., Mackworth, A., & Poole, D. (2020, April). AISpace2: an interactive visualization tool for learning and teaching artificial intelligence. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 09, pp. 13436-13443). AAAI Press.

小学生配对编程脚本设计与应用研究：基于有效失败视角

Research on the Design and Application of Pair Programming Scripts for Primary School Students: Based on the Perspective of Productive Failure

王靖*，李会 ，陈小汾
江南大学 江苏“互联网+教育”研究基地，江苏无锡 214122
Leehui.jnu@qq.com

[摘要] 当前小学配对编程面临着无法从常常遭遇的失败中获取有效的学习经验，而小学生在配对协作上的不成熟又加剧了这种问题的发生。文章从有效失败视角出发，围绕“如何基于有效失败设计面向小学生的配对编程脚本,，以提升其编程学习效果” 的核心问题，采用基于设计的研究方法开展三轮教学实践，得到了有效失败视角下的小学生配对编程脚本设计原则及策略，原则为引导学生利用原有知识生成多样化的问题解决方案、给予学生验证方案并探索失败原因的机会、引导学生对比多样化的方案以理解新旧概念的差异、促进学生在迁移应用中加深对概念的理解。对比分析三轮教学实践可知，上述原则指导下的脚本能够丰富学生的协作对话，促进学生触发有效失败发生的四个关键条件，加深其对目标概念的理解，提升学生的程序质量及态度。

[关键词] 小学生配对编程；有效失败；协作脚本

Abstract: *Currently, primary school pairing programming faces the challenge of not being able to gain effective learning experiences from frequent failures, and the immaturity of primary school students in pairing collaboration exacerbates this problem. Starting from the perspective of productive failure, the article revolves around the core issue of "how to design paired programming scripts for elementary school students based on productive failure and generate a paired programming learning platform to improve their programming learning effectiveness." Using a design based research method, three rounds of teaching practice are carried out, and the principles and strategies for designing paired programming scripts for elementary school students from the perspective of productive failure are obtained, The principles are to guide students to use their existing knowledge to generate diverse problem-solving solutions, provide students with opportunities to verify solutions and explore the reasons for failure, guide students to compare diverse solutions to understand the differences between new and old concepts, and promote students to deepen their understanding of concepts in transfer applications. By comparing and analyzing the three rounds of teaching practice, it can be seen that scripts guided by the above principles can enrich students' collaborative dialogue, promote the four key conditions for triggering productive failures, deepen their understanding of the target concept, and improve the quality and attitude of students' programs.*

Keywords: pair programming at primary school, productive failure, collaborative scripts

1.前言

编程教育正在受到世界各国的普遍重视，其中配对编程是编程教育中的重要教学组织形式。但当前小学配对编程面临的现实问题是，无法从常常遭遇的失败中获取有效的学习经验，而配对协作是促进编程学习高效开展的重要因素，配对协作本该促进学生对失败原因的深入思考与知识整合，并持续探索完善已有的解决方案，但这种高效协作在小学配对编程中并不能自然发生。有效失败理论指导下的脚本为学生提供了高效协作的机会，能够帮助学生

从失败中获取学习经验，加深其对后续所学新知识的理解。但鲜有研究将有效失败和协作脚本结合解决这一现实问题。基于以上背景，本研究的核心研究问题是“如何基于有效失败设计面向小学生的配对编程脚本，以提升其编程学习效果？”，分为以下三个子问题：

第一，如何生成有效失败视角下的小学生配对编程脚本设计原则和策略？

第二，如何依据上述原则和策略，设计对应脚本和配对编程学习平台，并进行迭代改进？

第三，在应用上述脚本后，小学生配对编程学习效果呈现出怎样的变化？

2.文献综述

已有研究者根据短期表现和长期学习的不等价性将学习现象分为四类，有效成功、有效失败、无效成功、无效失败，如表 1 所示。其中有效成功和有效失败均指向了长期学习的目的，即学生真正理解了目标概念并能准确应用于实际问题。有效成功和有效失败虽然指向了同一目的，但实现路径却有所不同。有效成功是通过提供支架、工具等保证学生成功解决问题并实现学习的发生 。相对于有效成功，有效失败是指学生在学习目标概念（即一节课中将要学习的新概念，简称“目标概念”）之前，尝试采用先验知识解决问题，但由于缺乏必要的概念支持导致问题解决的失败，然而这种失败却能够促使学习者激活更多的先验知识去解决问题，进而生成多种不成熟的解决方案，在这种失败方案的对比中，学生逐渐发现问题本质及目标概念的必要性，这为后续正式学习目标概念做好了认知准备和动机准备。因此在该过程中，虽然学生没有生成标准、正确的解决方案，但学习是真正发生了的。

表 1 学习现象的含义及案例表

学习现象	含义	案例
无效失败（Unproductive Failure）	教学实施后，学生短时间的成绩表现不佳，而且学习并未真正发生，长期学习效果也不佳。	纯粹的发现式学习（discovery learning），学生自始至终没有接受任何指导，完全靠自己解决问题学习目标概念。
有效失败（Productive Failure）	通过设计学习情境，学生短时间的成绩表现不佳，但从长期来看学习是发生了的。	为学生设计复杂的问题情境，学生生成了多种不成熟的解决方案，但激活了先验知识，为后来学习做了准备（Kapur, 2012）。
无效成功（Unproductive Success）	教学实施后，学生短时间内的成绩表现很优秀，但学习并未真正发生，因此对长期学习无益，是一种学习假象。	通过大量练习和死记硬的学习，只是依靠短期记忆提升成绩，但没有深刻理解。如，单词学习的场景（Miller, 1987）[1]
有效成功（Productive Success）	教学实施后，学生短时间的成绩表现很优秀，并且学习是真正发生了的，因此长期学习效果也很好。	建构主义者常常使基于问题的学习方法（PBL）和引导式探究方法，引入问题解决活动支架，以引发学习，随着活动的进行，支架逐渐撤去。

有效失败能够帮助学生从失败中获取有效的学习经验，并正视失败。该理论鼓励“先做后学”、“先难后易”，即分为两个阶段。第一个阶段先让学习者尝试协作解决复杂、劣构的新问题，学生在该阶段不断地尝试构建、修改问题解决方案，但最终都面临失败，即无法解决问题，在该过程中学生逐渐感知到自己与同伴已有认知的局限性，并认识到解决该问题所需知识的关键特征是什么及其必要性。第二阶段由教师介入并展开教学，首先教师带领学生分析

并总结全班同学呈现出的解决方案的失败原因，而后学生带着修正失败的动力，跟随教师学习目标概念并解决原来的问题，此时学生将再次体会目标概念的必要性及应用场景[2]。其中目标概念是指本节课将要学习的新概念，同时也是解决该问题的必要概念。在这两个阶段中，有效失败的作用机制在于，小组在第一阶段会产生失败方案，这种失败刺激着学生通过协作讨论，调动更多先验知识来解决该问题，多种失败方案的对比让学习者认识到目标概念的关键特征及必要性，为第二阶段的学习做好了准备，因此在第二阶段的时候，学习者能够基于教师讲授的新知识及标准解决方案，实现对目标概念的深刻理解及个体意义的建构，同时对失败的价值有了新的认识。在编程中，目标概念所对应的是以编程概念为核心的算法设计。

协作脚本是一种旨在通过构建学习者之间的互动来促进协作学习的活动程序。大量学者设计了多种类型的协作脚本为学生提供协作支持，如角色脚本、同伴反馈脚本，以探究协作脚本的支持机制。小学配对编程需要这种协作支持，通过设计配对小组的互动结构与流程，实现高效协作，促进知识的共享与共建。

因此，本研究将基于有效失败的视角为小学生配对编程设计协作脚本，以期帮助学生通过高效协作，充分参与到有效失败过程中，从失败中汲取经验，实现编程水平的提升。

3.基于有效失败视角的小学生配对编程脚本设计原则与策略初步构建

为了指导小学生配对编程，研究依据有效失败的发生阶段及其核心机制，提炼出适用于小学配对编程的四个关键条件，如表 2 所示。

表 2 有效失败发生的四个关键条件

关键条件	详细描述
激活先验知识	学生通过协作充分调用了其现有的知识和技能，仍然无法有效解决问题。
自我认知重构	学生通过解决问题的失败，刺激其探究失败原因，模糊意识到自己所欠缺的知识和技能所具备的特征，实现自我认知重构，这个重构的本质是为即将学习的目标概念做好认知图式上的充分准备。
新旧知识整合	学生通过对失败过程的充分分析（班级范围内）和反思来实现“激活先验知识”与“自我认知重构”的联通，并将新旧知识进行充分整合。
完善认知图式	学生通过在新场景中迁移应用，进一步理解新旧知识差异，完善认知图式。

基于此,研究者构建了基于有效失败视角的小学生配对编程脚本设计原则与策略(见图 1)，以提升学生的配对编程学习效果。其中第二行激活先验知识、自我认知重构、新旧知识整合、完善认知图式是有效失败理论指导小学生配对编程的关键条件，对应这四个关键条件，有四项设计原则，即引导学生利用原有知识生成多样化的问题解决方案、给予学生验证方案并探索失败原因的机会、引导学生对比多样化的方案以理解新旧概念的差异、促进学生在迁移应用中加深对概念的理解。而每项设计原则下主要包括三个方面的设计策略，即任务设定、角色设定、事件序列，其中下图中的事件序列主要呈现了关键的支持。



图 1 基于有效失败视角的小学生配对编程脚本设计原则与策略

研究根据上一章节构建的小学生配对编程脚本设计原则与策略,提出了基于有效失败视角的小学生配对编程脚本设计框架，见表 3，并将其部署至配对编程学习平台，如图 2 所示。

表 3 第一轮次基于有效失败视角的小学生配对编程脚本设计框架

活动	任务	角色	事件序列
协作讨论使用turtle解决某个绘画问题	生成不成熟方案阶段（小组）	基于原有认知协商问题解决方案者、观点评审员	1. 小组内两位同学分别思考解决方案，写在纸上。随后两位同学交换想法，通过讨论最终形成解决方案。首先提出方案的人作为“观点提出者”，详细阐述自己的观点；另一位同学作为“观点评审员”，对同伴的观点提出质疑或支持，最后二人将完善后的想法写在平台中。 2. 协作论证支持：如有论证上的困惑，可以点击“如何反驳/维护观点”获取论证支持。
	探索失败原因阶段（小组）	基于协商驾驶员、结果编程导航员实现方案	1. 两位同学分别担任驾驶员和导航员。其中驾驶员将在平台键入代码，导航员则提供编码建议，检查并指出代码中的错误。 2. 协作论证支持：同上。 3. 角色切换支持：如有角色交换的矛盾，可点击按钮“和同伴交换角色时出现矛盾怎么办”
		针对方案失败之处尝试查找原因者、观点评审员	1. 两位同学讨论方案失败的原因或缺陷。首先提出观点的人作为观点提出者，另一位同学作为观点评审员，小组最终总结出至少 1 条失败原因，二人记录所有失败原因或缺陷记录。在完成上述操作后，可以选择是否补充新的解决方案。 2. 协作论证支持：同上。 3. 失败原因探究支持：在平台设置用于整理失败原因的文本框。
	整合失败原因阶段（班级）	对比失败方案反思已有认知不足者、观点评审员	1. 教师带领全班讨论对比典型失败方案。典型小组展示方案，其他小组评价并提出质疑，最终总结出失败原因。随后二位同学将其他方案的失败原因或缺点补充在平台中。 2. 协作论证支持：同上。 3. 总结失败原因支持：在平台设置用于整理不成熟方案及失败原因的文本框
		聚敛生成概念并应用解决原来问题	1. 两位同学各自在 python 平台模仿标准方案，编写代码。
	迁移应用探究阶段（小组）	迁移应用概念至新问题 观点质疑者、观点维护者	1. 双方分别在平台键入代码解决新问题。书写完成后，向同伴口述自己的思路，同伴作为观点质疑者对其方案提出质疑，并写在失误单上。撰写代码的同学作为观点维护者解释并维护个人观点。 2. 失败原因探究支持：在平台设置用于整理失败原因的文本框。



图 2 配对编程学习平台界面 1

4.研究设计

4.1 研究对象

研究者面向华东地区某小学的 14 名四年级学生，在编程社团课中开展实证研究。 这部分学生是主动选择这门编程社团课的，因此对该课程的情感倾向是一致的。在本学期编程学习之前，在前一学期已经在课堂上初步接触过编程。

4.3 研究思路

研究遵循设计研究“设计—实施—评价”的循环，如图 3 所示。对基于有效失败视角的小学生配对编程脚本设计原则和策略进行迭代优化。研究根据学生所持先验知识确定了每一轮次的学习主题，结合该轮次的脚本设计原则和策略，生成了配对编程脚本设计框架。为了监测学生使用脚本的情况，研究将框架部署到了配对编程学习平台。接下来实施教学，并进行评价。研究者在每一轮评价反思时，均依据四个角度的数据：一是小组在迁移应用目标概念的结果，数据来源为迁移应用探究任务中的程序方案；二是小组配对编程过程中的互动过程，数据来源为对话音频；三是学生的情感变化，数据来源为学生填写的失败态度问卷和协作态度问卷；四是学习者体验，数据来源为对学习者的访谈和课堂观察。据此优化下一轮次的脚本设计原则和策略，开展新一轮次的教学实践。



图 3 研究思路

4.5 数据分析方法

本研究针对收集到的数据， 主要采用统计分析法、滞后序列分析法，对配对编程学习过程和结果以及学生的态度变化进行分析。其中滞后序列分析采用的是 Hopcan 等人(2022)的配对编程活动中的协作行为编码表，笔者进行了适度改编，主要包括计划、贡献、期待、称赞、社会交互五个部分，此处的社会交互指的是与课堂无关话语。在使用该表对学生的协作话语编码时，两位研究者选取了其中 20%的内容进行编码，经过修正，最终评分者一致性信度达到 97.41%，笔者对剩余的协作话语数据进行了编码。

5.小学生配对编程脚本设计原则与策略的迭代

5.1. 首轮教学实践实施与评价

根据第一轮次教学实践中课堂内容的目标概念“向列表添加新元素”，研究设计了真实的问题情境“定制方格纸”。研究对教学实施中收集到的数据进行分析，有以下几个发现：第一，在编程程序质量方面，学生的整体水平处于中等偏下的水平（程序质量分数的描述性统计表见表 4）。第二，在编程学习过程方面，研究从基于有效失败过程的编程学习过程、配对编程中的协作行为模式进行分析，发现学生在触发有效失败发生的关键条件上有所欠缺，存在诸如无法探知方案失败的真正原因、缺乏描述程序失败原因的能力等问题；而在配对编程中的协作行为模式上，出现了四种模式：互帮互助的积极协作模式，以及先分享者方案优先执行、

争夺角色、浅层论证的消极协作模式。第三，在情感态度方面，学生对待失败的态度以及协作的态度是积极的，但部分学生存在对团队失败状态的感知不到位的问题。除此之外，研究者在课堂观察中发现，学生在进行任务二“编程实现方案”时，就已经意识到方案中存在的问题，并进行了程序的改进，即学生在进行任务二的同时，也完成了任务三“探究失败原因”。

表 4 第一轮教学实践程序质量分数描述性统计

变量名	最大值	最小值	平均值	标准差	中位数	方差	峰度	偏度	变异系数（CV）
程序质量分数	9	1	4.571	2.533	4.5	6.418	-0.731	0.3	0.554
程序完整性	2	1	1.357	0.497	1	0.247	-1.838	0.67	0.366
程序技术性	5	0	2.357	1.646	2	2.709	-0.823	0.295	0.698
程序效果实现	2	0	0.857	0.77	1	0.593	-1.123	0.264	0.899

5.2. 第二轮设计、实施与评价

5.2.1. 设计优化

针对设计原则层面的问题，研究将第一项原则修改为“引导学生激活多样化的先验知识，做好认知准备”。针对设计策略层面的问题，第一，研究结合上述修改后的设计原则，修改了第一项设计原则对应的设计策略，主要是对事件序列进行了更新，包括方案共享支持、先验知识清单和协作论证支持。方案共享支持是用于加强学生分享自身方案意识的策略支持；先验知识清单是指为学生提供已经学习过的所有知识的清单，以促进学生激活与目标概念相关的更多先验知识；协作论证支持则是在原有的协作论证支持的基础上进行了完善，在“如何维护观点”中增设接受对方的观点时，还要说出接受的理由，以此表明自己真正认同对方的观点。第二，研究修改了第二项设计原则对应的设计策略，将两个协作任务进行合并，修改为基于编程方案撰写并完善程序。另外，研究将事件序列修改为协作论证支持、角色支持、失败原因探究支持，优化后的协作论证支持与原则一中的协作论证支持保持一致，角色支持则明确了确定角色和交换角色的规则，失败原因探究支持则给出了失败原因的描述示例。第三，研究将第三项设计原则对应的设计策略进行了修改，将事件序列中的支持修改为协作论证支持、总结失败原因支持，其中协作论证支持同第一项原则中的策略，总结失败原因支持主要体现在两个方面，一是要求学生在教师带领全班同学分享方案并探讨失败原因时，学生要立即将该失败方案总结到文本框中，二是在原有总结失败原因支持的基础上添加了如何描述失败原因的示例。第四，研究对第四项设计原则对应的设计策略进行了更新，增添了事件序列中的支持，即失败原因探究支持，设置失败原因的描述示例。

5.2.2. 实施与评价

根据第二轮次教学实践中课堂内容的目标概念“向列表的特定位置添加新元素”，研究设计了真实的问题情境“绘制节日花环”。研究对教学实施中收集到的数据进行分析，有以下几个发现：第一，在编程程序质量方面，相较于第一轮次有少许提升，但差异不显著，见表 5。第二，在编程学习过程上，研究从基于有效失败过程的编程学习过程、配对编程中的协作行为模式进行分析，发现学生在触发有效失败发生的关键条件上有所改变，例如，学生在激活先验知识、自我认知重构这两个条件方面有所改善，但触发新旧知识整合、完善认知图式这两个条件方面还有很大不足；在配对编程中的协作行为模式上，出现了四种模式，包括浅层论证的协作模式、以推动进程为中心的协作模式、争夺角色的协作模式、激励共进的协作模式，其中浅层论证等协作模式与第一轮次中的不同，其产生原因多是更新后的脚本加重了学生的认知负荷。第三，在情感态度方面，学生对待失败的态度和对待协作的态度是相对积极的，但对于协作中的角色分配及角色职责方面还不是很清晰，甚至产生了分歧。

表 5 第一轮次与第二轮次的程序质量分数配对样本 T 检验

配对变量	平均值±标准差	t	df	**P**	Cohen's d

	配对 1	配对 2	配对差值（配对 1-配对 2）				
第 1 阶段-第 2 阶段	4.571±2.533	5.071±2.586	-0.5±-0.053	-0.809	13	**0.433**	0.216

注：***、**、*分别代表 1%、5%、10%的显著性水平

5.3. 第三轮设计、实施与评价

5.3.1. 设计优化

研究针对第二轮次中的问题，对脚本设计原则与策略进行了优化。第一，研究修改了第一项设计原则对应的设计策略，主要对协作论证支持进行了修改。将协作论证支持分成“如何反驳观点”、“如何表示同意”、“如何维护观点”三个提示，并对每个提示做出解释和示例。为了减轻学生的认知负荷，研究基于第 12 课时的内容做了示例，并且在上课之初由教师和一位学生按照该示例展示科学论证的过程。第二，研究修改了第二项设计原则对应的设计策略，首先是对协作论证支持的优化，与第一项原则中的协作论证支持保持一致。其次是对角色支持的优化，包括角色切换机制的设定和角色职责的进一步明确。对于角色切换机制，每过 10min 教师提醒可以切换角色，只要小组内有一人想要切换角色，必须立即切换，否则只能等到下一次提醒才能切换；对于角色职责，要求由导航员将找出的错误及时记录在平台上，而驾驶员不再撰写编程失败的原因。第三，研究将第三项设计原则对应的设计策略进行了修改，在对比失败方案这一任务中，第二轮次的脚本要求学生在班级共享方案的同时，跟随记录其不成熟方案及失败原因到平台上，结果表明学生还不具备边听边整理并将已分享内容和当前内容整合起来的能力，需要教师介入支持。因此在本轮次教学实践中，研究将这一脚本事件修改为“在每个小组展示不成熟方案之后，教师在黑板上将该小组的方案及失败原因简写下来，同时学生完成当前失败方案与已有失败方案的整合”，由教师带领学生整理出当前的方案，学生完成所有不成熟方案的整合。另外，第三项原则对应下的协作论证支持也得到了更新。

5.3.2. 实施与评价

根据第二轮次教学实践中课堂内容的目标概念“值为列表的字典写法与读取”，研究设计了真实的问题情境“制作趣味校园卡”。研究对教学实施中收集到的数据进行分析，有以下几个发现：第一，在编程程序质量方面，相较于第二轮次有少许提升，但差异不显著，与第一轮次有显著性差异，见表 6、表 7。第二，在编程学习过程上，研究从基于有效失败过程的编程学习过程、配对编程中的协作行为模式进行分析，发现学生在触发有效失败发生的关键条件上有所改善，但还有少数学生因为协作等方面的问题无法触发这些关键条件，特别是在自我认知重构这一条件上；在配对编程中的协作行为模式上，研究总结出了三种模式，包括循环论证的协作行为模式、以推动进程为中心的协作模式、切换角色时矛盾频发的协作行为模式，其中切换角色所引发的矛盾对协作对话的质量影响最大。第三，在情感态度方面，多数学生对待失败的态度和对待协作的态度是相对积极的，但仍有部分小组因为角色矛盾引发的一系列问题而持有消极的失败态度和协作态度。因此，未来还需要进一步优化该脚本。

表 6 第二轮次和第三轮次的程序质量分数配对样本 T 检验

配对变量	平均值±标准差			t	df	**P**	Cohen's d
	配对 1	配对 2	配对差值（配对 1-配对 2）				
第 2 轮次-第 3 轮次	5.071±2.586	5.929±2.165	-0.857±0.421	-1.158	13	**0.268**	0.309

注：***、**、*分别代表 1%、5%、10%的显著性水平

表 7 第一轮次和第三轮次的程序质量分数配对样本 T 检验

配对变量	平均值±标准差			t	df	**P**	Cohen's d
	配对 1	配对 2	配对差值（配对 1-配对 2）				

第 1 轮次-第 3 轮次	4.571±2.533	5.929±2.165	-1.357±0.368	-2.267	13	**0.041****	0.606

注：***、**、*分别代表 1%、5%、10%的显著性水平

6.结论与不足

经过三轮迭代，研究最终得出的结论如下：第一，生成有效失败视角下的小学生配对编程脚本设计原则和策略，需要剖析有效失败发生过程，提取触发有效失败发生的四个关键条件，即激活先验知识、自我认知重构、新旧知识整合、完善认知图式。分别对应脚本的四项设计原则。第二，经过迭代改进，研究生成了基于有效失败视角的小学生配对编程脚本设计原则，即引导学生利用原有知识生成多样化的问题解决方案、给予学生验证方案并探索失败原因的机会、引导学生对比多样化的方案以理解新旧概念的差异、促进学生在迁移应用中加深对概念的理解。第三，上述配对编程脚本的应用能够提升学生的程序质量，在学习过程上能够帮助学生触发有效失败发生的四个关键条件，加深其对目标概念的理解；同时能够丰富学生的协作对话，促进其形成循环论证、以推动进程为中心的协作模式；能够帮助学生认识到失败的价值，以及高效协作的优势，改变其对遭遇失败、协作的态度。

本研究还存在一些不足，一是在第三轮次教学实践后，没有基于对脚本的评价和建议生成新的脚本，并投入实践进行检验；二是研究对象的范围和实施情境具有局限性，没有面向除四年级以外的其他年级开展实践研究，也没有在除 Python 语言以外的其他语言课程中开展教学实践；三是研究时长具有局限性，研究仅在一个学期开展，难以分析脚本的长期效应，未来还需要延长研究时长，探究脚本的长期影响。

参考文献

曹鹭.有效失败与知识迁移:理论、机制与原则[J]. 开放教育研究, 2021, 27(3): 4-14.
郭耕汝,贺鹏洋,黄琳淇等.协作角色脚本支持下“边缘人”认知网络分析[J].教学研究,2023,46(05):1-8.
刘敏,汪琼.结对编程:中小学编程教育的首选教学组织形式[J].现代教育技术,2022(3):102-109.
Miller G, Gildea P. How children learn words[J]. Scientific American, 1987, 257: 94 - 99.
Chen, W., et al.(2023). Students' appropriation of collaboration script in a networked class: An exploratory study. Technology, Pedagogy and Education, 1–19.
Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. Educational Psychologist, 48(1), 56–66.
Hopcan, S., Polat, E., & Albayrak, E. (2022). Collaborative Behavior Patterns of Students in Programming Instruction. Journal of Educational Computing Research, 60(4), 1035–1062.
Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. Journal of the Learning Sciences, 21(1), 45–83.
Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning, 2(2–3), 211–224.
Sbaraglia, M., Lodi, M., & Martini, S. (2021). A Necessity-Driven Ride on the Abstraction Rollercoaster of CS1 Programming. Informatics in Education, 20(4), 641–682.
Shao, J., Cheng, L., Wang, Y., Li, K., & Li, Y. (2023). How peer feedback with regulation scripts contributes to the development of critical thinking in dialogues: Strengthening cognitive and affective feedback content. Interactive Learning Environments, 1–20.

Personalized Learning Supported by Learning Analytics in Smart Education: A Systematic Literature Review

Jing Ning[1], Binger Zhang[2],
Zhiqiang Ma[3*]
[1] School of Design, Jiangnan University
[2] Faculty of Humanities, Jiangnan University
[3] Faculty of Humanities, Jiangnan University
[*] mzq1213@jiangnan.edu.cn

Abstract: *With the rapid development of information technology, the educational profession is progressively seeking more intelligent, flexible, and personalized learning models. As a major technology in smart education, learning analytics provides the opportunity to get in-depth insights into individual diversities such as learning habits, interests, educational background, and cognitive capabilities. The objective of this paper is to comprehensively explore personalized learning guided by learning analytics in smart education through an in-depth and systematic literature review. The integration of learning analytics and personalized learning has been shown to be a revolutionary force in education, promoting a student-center approach that customizes instructional strategies to individual demands, preferences, and learning styles. This is evident when considering the major findings and developing trends. By seamlessly integrating learning analytics technologies into intelligent learning environments, instructors may leverage big data and artificial intelligence to analyze student behaviors, preferences, and performance trends at a more granular level.*
Keywords: personalized learning, learning analytics, systematic literature review

1. Introduction

In the landscape of rapid development in intelligent education, personalized learning is becoming a central topic in the innovative education domain (Kustitskaya et al., 2023). In the current wave of intelligent education, learning analytics, as a cutting-edge technology, is ushering in a new era for education. With continuous innovations in information technology, there is a growing demand in the education field for more sophisticated, adaptable, and individualized learning. Learning analytics, as a data-driven educational practice and research method, plays a crucial role in personalized learning. Learning analytics collects, processes, and analyzes data generated during the learning process, revealing learners' behavior patterns, learning needs, and learning difficulties, and providing educators with real-time personalized feedback and support (Doroudi, 2024).

In recent years, with the development of artificial intelligence, researchers have utilized learning analytics technology to achieve the effective collection and processing of data generated during the learning process, providing vital support for the realization of personalized learning. Learning analytics technology is one of the key technologies supporting the implementation of personalized learning (Abrahamson et al., 2022). Currently, with the continuous development of technologies such as machine learning, data mining, and artificial intelligence, researchers have proposed a series of advanced learning analytics methods and algorithms, including deep learning, reinforcement learning, and swarm intelligence. The application of these technologies enables a more precise and comprehensive analysis of learners' learning behaviors and characteristics, thus providing technical support for the implementation of personalized learning (del Pilar Gonzalez & Chiappe, 2024).

There are empirical research and relevant review studies on "personalized learning" and "learning analytics" recently, with the passage of time, new research findings and innovations continue to emerge, particularly in the field of technology. In recent years, the rapid development of artificial intelligence and machine learning technologies has provided new possibilities and opportunities for personalized learning and learning analytics. The application of these new technologies has made personalized learning and learning analytics more feasible. Previous review articles may have certain limitations and constraints. For instance, prior review articles may have failed to cover the latest research findings or may not have fully considered different research designs or application scenarios. Furthermore, existing review articles lack an objective evaluation and comprehensive summary of the overall research findings. For example, some review articles focus on how learning analytics supports personalized learning, while others summarize the current applications of learning analytics. However, they lack an in-depth analysis of the advantages brought about by learning analytics as a technological support for personalized learning. Therefore, writing a systematic literature review can objectively assess existing research findings and provide a more comprehensive and accurate understanding of the various benefits and outcomes of learning analytics as technological support for personalized learning.

1.1 Personalized Learning

Personalized learning, as an educational strategy, adapts teaching delivery, content, and pacing to accommodate the diverse needs, interests, and abilities of individual students. The intrinsic advantages of this approach are evident in its capacity to cater to the multifaceted requirements of learners, thereby enhancing their engagement and academic performance (Al-Badi & Khan, 2022). By tailoring learning experiences according to individual learning styles and preferences, personalized learning fosters deeper subject understanding, promotes intrinsic motivation, and enhances student achievement by recognizing and addressing individual differences. Furthermore, personalized learning cultivates learners' awareness of the self-directed learning process, facilitating their self-inquiry and independent learning. It is noteworthy that personalized learning holds the promise of reducing educational disparities by providing targeted support to students in need of additional assistance or facing learning challenges (Shorfuzzaman et al., 2019).

The effective of implementing personalized learning requires a comprehensive approach encompassing several key components: (1) meticulous data collection about individual students' progress levels, strengths, and weaknesses across different subject domains, (2) the adaptation of learning materials and assessments to meet the unique needs of each student, and (3) the provision of customized feedback to guide their learning trajectory (Chen & Li, 2021). However, the current educational landscape not only features individual assessment but also emphasizes collaborative collective projects, underscoring the urgent need for educators to provide personalized guidance across various teaching modalities (Fachola et al., 2023). Consequently, the exigencies of personalized learning impose significant temporal and human resource demands. Regrettably, empirical evidence suggests a prevailing disparity between the proportion of teachers' work hours allocated to classroom instruction—averaging 46%—and the requisite resources for individualized student monitoring and instructional personalization. As a result, educators may encounter challenges in autonomously monitoring and discerning the behavioral patterns, aptitudes, and preferences of each student (Mathrani et al., 2021).

1.2 Learning Analytics

In recent years, with the development of information technology and artificial intelligence, the application of learning analytics in personalized learning has continuously deepened and expanded, achieving a series of remarkable results. In fact, learning analysis of personal learning behavior not only focuses on learners' surface behaviors, such as clicks and browsing but also emphasizes the deeper learning processes and cognitive activities (Al-Badi & Khan, 2022). In recent years, researchers have begun to integrate learning behavior analysis with theories from cognitive processes and strategy selection during learning. For instance, some studies have utilized eye-tracking technology and psychological experimental methods to investigate learners' attention allocation and information processing strategies during reading comprehension and problem-solving processes, thereby revealing learners' cognitive behavior patterns and learning strategy selections (Cavalcanti et al., 2021).

In addition, current research has utilized the analysis of learners' learning behaviors and social network data to construct models of learners' behavioral patterns and social relationship networks, thereby achieving more precise and personalized learning path recommendations (Schumacher, Klasen & Ifenthaler, 2019). This data-driven approach to learning path recommendation provides educators with more flexible and effective personalized learning support, promoting learners' motivation and engagement. Meanwhile, some studies identify learners' cognitive difficulties and error patterns by analyzing their learning behaviors and process data, thereby providing educators with personalized learning feedback and support (Shorfuzzaman et al., 2019). This form of learning effectiveness assessment, based on learning behavior analysis and cognitive diagnostics, offers educators a more comprehensive and accurate profile of learners, facilitating the implementation and development of personalized learning (Yang et al., 2021).

1.3 Paper Structure and Research Questions

This paper was structured in four primary phases: 1) problem definition; 2) creation of criteria; 3) data collection and analysis; and 4) discussion. There are significant knowledge gaps between how learning analytics support personalized learning in practice, especially learning analytics as a technique to identify learners' learning needs, as well as how learning analytics can be applied in personalized learning under practical teaching situations, and what would be the outcomes or benefits of using learning analytics as a technique through personalized learning. Hence, the following research questions:

RQ1: How is personalized learning effectively supported by learning analytics in published empirical studies?

RQ2: What benefits does learning analytics bring as technological support for personalized learning in terms of practical applications?

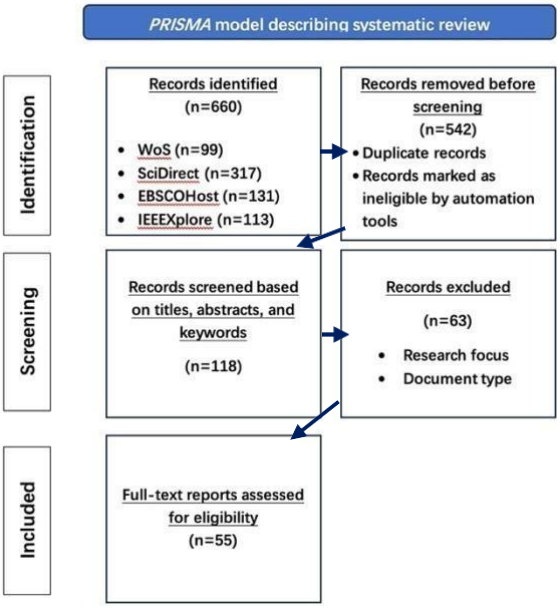
2. Research Methodology

This paper has adopted the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach to ensure the greatest levels of scientific rigor and transparency in the systematic literature review. This study

has enhanced reproducibility and comparability by adhering to the PRISMA guidelines and pertinent steps, such as literature search, screening, data extraction, and quality assessment. Which has effectively reduced the impact of subjective judgments and increased the credibility of the research results.

In order to ensure that all pertinent publications in the field are included, this study has chosen the following well- known and credible databases for the literature review: Science Direct, EBSCOhost, IEEE Xplore, and Web of Science. These databases have consistently yielded the most pertinent journals for the systematic review. The Google Scholar H- index has been used to choose articles for this paper, with an emphasis on Q1 and Q2 journals, from the following categories: "education", "education technology", "higher education", and "teaching and teacher education". This selection has aimed to control the scope of the papers while ensuring a sufficiently broad overview. The influence of the chosen journals is validated using SJR (Scimage journal Rank). The time frame has been filtered to include papers from 2015 to 2024, reviewing papers on "personalized learning", "individualized learning", "customized learning", "learning analytics", "learning analysis", and the combination of these terminologies as these have been a popular issue among academics and policymakers. The language has been filtered to include only English papers that involve technology integration and are research articles published in one of the listed peer-reviewed scientific journals to control the search scope.

Figure 1.
Flowchart of the literature search process



3. Results

3.1 How is learning analytics effectively applied in personalized learning?

3.1.1 Learning Analytics Frameworks and Models

Among this reviewed literature, the categories of "learning analytics framework and models" applied to support personalized learning can be summarized. Researchers have proposed a series of concepts and techniques for collecting, analyzing, and interpreting learning data (Kustitskaya et al., 2023). These frameworks aim to assist educators and learners in understanding patterns, trends, and relationships in the learning process, thereby providing personalized support and advice for individual learners (Vieira, Parsons & Byrd, 2018). Learning analytics dashboards are one essential tool in this regard, capable of aggregating learning data from various sources and providing deep insights and predictions through methods such as data mining, machine learning, and statistical analysis (Cavalcanti et al., 2021). Additionally, learning recommendation algorithms can recommend suitable learning resources and pathways based on learners' interests, behaviors, and learning histories. Overall, these frameworks and methods provide important tools and approaches for educational practice, helping to optimize the learning process and outcomes (Zhou et al., 2018).

This category encompasses research on learning analytics frameworks and models, including but not limited to the proposal of a comprehensive and modular platform concept framework, validation of the effectiveness and accuracy of learning analytics, the use of the Allocation Error Matrix (AEM), the Adaptive Learning Objects (ALOs) model, and the proposal of the Big Data Visualization-Assisted Multimodal Feedback Framework (BDVMFF) (Diego, Carlos & Jose, 2019). These studies aim to leverage data analysis and model construction to provide a deeper understanding of the learning process and environment and explore how personalized learning support can be provided through learning analytics (Diego, Carlos & Jose, 2019).

3.1.2 Multimodal Learning Analytics (MMLA):

Multimodal learning analytics is a comprehensive approach that integrates data from various sources to gain a holistic understanding of learners' learning processes and performances in different environments. By collecting and analyzing students' cognitive, affective, behavioral, and physiological data, researchers can assess learners' learning states and needs more accurately (Emerson et al., 2020). Furthermore, leveraging computational algorithms and artificial intelligence technologies allows for the extraction and analysis of multimodal data such as video, audio, and text, revealing underlying patterns and trends in learners' learning processes. Multimodal learning analytics offers new perspectives and methods for

personalized learning and teaching, aiding in the optimization of learning environments and instructional strategies (Abrahamson et al., 2022).

This category encompasses research on multimodal learning analytics, including the use of gameplay, eye-tracking, facial expressions, and other data to predict students' learning performance and interests, as well as principled approaches to collecting, analyzing, and presenting various data in online and offline learning environments. These studies aim to deepen the understanding of cognitive, affective, behavioral, and physiological processes in the learning process through multimodal data analysis, guiding personalized learning support (Abrahamson et al., 2022; Emerson et al., 2020).

3.1.3 Learning Analytics Applications and Tools:

Personalized learning and adaptive systems aim to provide tailored learning support and resources based on the characteristics and needs of learners. The Adaptive Learning Objects model allows learners to autonomously select learning paths and resources based on individual interests and learning history. Learning recommendation algorithms leverage learners' data and behaviors to recommend suitable learning resources and pathways (Klašnja-Milićević, Ivanović & Stantić, 2020). Through big data analysis and machine learning technologies, educators can better understand learners' needs and behaviors, thereby optimizing the teaching and learning processes. Personalized learning and adaptive systems offer a flexible, personalized approach to learning, contributing to improved learning outcomes and experiences (Sedrakyan et al., 2020).

This category encompasses research on the application and tools of learning analytics, including the use of commercial applications and network analysis tools to create predictive models, and how teachers track student activities by processing digital traces left by online interactions. These studies aim to utilize existing tools and technologies to identify potential at-risk students, develop personalized learning resources, simulate collective learning community behaviors, and more.

3.1.4 Educational Technology and Tools:

Educational technologies and tools play a crucial role in learning analytics, supporting personalized learning and instructional optimization through the collection and analysis of learning data. Commercial applications and network analysis tools offer rich functionalities and tools for collecting, processing, and analyzing learning data (Han et al., 2021). Learning management systems and e-learning platforms provide centralized management and delivery of learning resources, facilitating the implementation of learning analytics and personalized learning support. The application of educational technologies and tools provides convenience and support for learners and educators, promoting personalized learning and instructional optimization (Cui et al., 2021).

This section includes an overview of commercial applications and network analysis tools used for learning analytics, such as SPSS, Stata, Nvivo, Google Analytics, and others (Zacharis, 2015). Additionally, it covers topics such as learning recommendation algorithms, learning path recommendation, the use of ePortfolios, and the development of automated feedback systems (Pospíšilová & Rohlíková, 2023). These studies aim to provide personalized support and feedback to students through personalized approaches and technologies, thereby optimizing their learning outcomes and teacher efficiency.

3.1.5 Philosophical and Theoretical Frameworks of Learning Analytics:

The philosophical and theoretical frameworks of learning analytics explore the foundations and objectives of learning analytics research, as well as the philosophical and methodological differences with related fields. These frameworks help researchers and practitioners better understand the nature and purpose of learning analytics, promoting collaboration and development within the field of learning analytics (Muslim, Chatti & Schroeder, 2022). Furthermore, philosophical and theoretical frameworks also emphasize the role of learning analytics in supporting personalized learning, including aspects such as data collection, prediction and modeling, feedback, and assessment (Berding et al., 2022).

This category encompasses research on the philosophical and theoretical frameworks of learning analytics, including the design of theoretical frameworks, understanding philosophical perspectives, and the support of learning analytics for personalized learning. These studies aim to better understand learning analytics and its related fields from philosophical and theoretical perspectives, and to promote collaboration and learning within the field.

3.1.6 Methods and Techniques in Learning Analytics:

Learning analytics methods and technologies provide a rich array of tools and techniques for collecting, analyzing, and interpreting learning data, supporting personalized learning and instructional optimization. These methods and technologies include text mining, feature extraction and clustering, and semantic and syntactic analysis algorithms, which can be utilized to generate personalized feedback and customized learning support (Shorfuzzaman et al., 2019). Furthermore, learning analytics technologies play a significant role in the development of automated feedback systems, providing learners with personalized feedback and support through the analysis of learning data. Learning analytics methods and technologies offer important tools and approaches for educational research and practice, aiding in the optimization of learning environments and instructional strategies, and enhancing learners' learning outcomes and experiences (Doroudi, 2024).

This category encompasses research on learning analytics methods and technologies, including text mining, longest common subsequence, feature extraction and clustering, and semantic and syntactic analysis algorithms, among others. These methods and technologies aim to generate personalized feedback and assist in customized learning to improve students' learning outcomes and teachers' efficiency.

3.2 What are the practical benefits or outcomes of learning analytics technically supporting personalized learning?

3.2.1 Enhancement of Personalized Feedback

Dispositional Learning Analytics (DLA) is an emerging approach that incorporates personality trait data, particularly mindset data, into the realm of learning analytics. This approach aims to enhance the predictive capabilities of learning analytics applications and improve the accuracy of personalized feedback provided to learners. The integration of personality traits into learning analytics enables educators to gain a deeper understanding of students' learning characteristics, thus facilitating more personalized and effective feedback (Aparicio-Fernández et al., 2023). DLA utilizes data related to students' mindsets, which encompass their beliefs, attitudes, and perceptions about learning. By analyzing this data alongside traditional learning analytics metrics such as performance data and engagement levels, DLA aims to uncover patterns and insights that

contribute to a more comprehensive understanding of students' learning behaviors and outcomes. The personalized feedback generated through DLA takes into account not only students' academic performance but also their mindset characteristics (Aparicio-Fernández et al., 2023). This feedback is tailored to address individual students' strengths, weaknesses, and learning preferences, thereby providing more relevant and impactful guidance.

The research findings suggest that DLA can benefit both educators and learners in several ways. For educators, DLA offers valuable insights into students' learning dispositions, allowing them to tailor their teaching approaches and interventions accordingly. Educators can identify potential barriers to learning, such as fixed mindsets or negative attitudes towards certain subjects, and implement strategies to address these challenges effectively (Karaoglan Yilmaz & Yilmaz, 2020). For learners, DLA facilitates a more personalized and supportive learning environment. By receiving feedback that acknowledges their unique mindset characteristics, students are empowered to develop a growth mindset and adopt effective learning strategies. Personalized feedback based on DLA can boost students' confidence, motivation, and self- efficacy, ultimately leading to improved learning outcomes (Tempelaar, Rienties & Nguyen, 2021).

In summary, Dispositional Learning Analytics (DLA) represents a promising approach to learning analytics that considers students' personality traits, particularly their mindsets, in the analysis and provision of personalized feedback. By integrating mindset data into the learning analytics process, DLA enhances the predictive capabilities of analytics applications and enables educators to provide more targeted and impactful support to learners. This approach has the potential to contribute significantly to the enhancement of teaching and learning practices in various educational settings.

3.2.2 Improvement in Practical Teaching and Learning

Learning analytics plays a crucial role in instructional improvement, as educators can gain deeper insights into students' learning needs and behavior patterns by analyzing their learning data and behaviors, thus adjusting teaching strategies to provide more effective learning support. This process not only helps improve students' academic performance but also enhances teaching effectiveness and educational quality. Firstly, learning analytics can assist educators in gaining a more comprehensive understanding of students' learning needs and challenges. By analyzing students' learning data, such as learning progress, assignment completion, and exam scores, educators can identify the problems and difficulties students encounter during the learning process and intervene and support them promptly (Wu et al., 2024). For instance, if a student performs poorly in a certain topic, educators can provide additional learning resources or personalized tutoring to help them overcome difficulties and improve learning outcomes. Secondly, learning analytics can help educators better utilize teaching resources and optimize the teaching process. By analyzing students' learning data, educators can understand which teaching resources and methods are more popular among students and which may need adjustment or improvement (Yilmaz et al., 2022). For example, if the utilization rate of a teaching resource is low, educators can consider optimizing or replacing it to enhance student engagement and learning effectiveness.

Moreover, learning analytics can assist educators in evaluating teaching effectiveness and making more accurate instructional decisions. By analyzing students' learning data and behaviors, educators can understand the impact of different teaching strategies and methods on students' academic performance, thus adjusting teaching strategies to improve teaching effectiveness (del Pilar Gonzalez & Chiappe, 2024). For example, by comparing the performance of different teaching methods in terms of student learning outcomes and satisfaction, educators can determine which method is more suitable for the current teaching environment and student population, achieving better teaching results.

In summary, learning analytics is of great significance in instructional improvement. By analyzing students' learning data and behaviors, educators can gain a more comprehensive understanding of students' learning needs and behavior patterns, thus adjusting teaching strategies to provide more effective learning support. This process not only helps improve students' academic performance but also enhances teaching effectiveness and educational quality, promoting students' overall development and growth.

3.2.3 Reinforcing Relationship between Learning Analytics and Personalized Learning

The relationship between learning analytics and personalized learning design is mutually reinforcing, providing important support for improving educational quality and student learning outcomes. Firstly, learning analytics provides essential data support for personalized learning design. By analyzing students' learning data and behaviors, educators can gain a more comprehensive understanding of students' learning needs and characteristics, thereby designing corresponding learning tasks and resources according to students' personalized needs (Lluch Molins & Cano García, 2023). For example, based on students' learning progress and interests, educators can design learning tasks of different difficulty levels and content to meet the diverse learning needs of students. Secondly, personalized learning design provides important guidance and application scenarios for learning analytics. Through personalized learning design, educators can design more diverse and student-centered learning tasks and resources, thereby providing learning analytics with more diverse and rich learning data and behaviors (Yang et al., 2021). These data and behaviors can provide learning analytics with more comprehensive and accurate data support, thereby enhancing the effectiveness and quality of learning analytics.

Finally, the mutual promotion between learning analytics and personalized learning design contributes to the optimization of the education process, improving educational quality and student learning outcomes (Yang et al., 2021). By combining learning analytics and personalized learning design, educators can gain a more comprehensive understanding of students' learning needs and behavior patterns, thereby adjusting teaching strategies to provide more effective learning support. This process not only helps improve students' academic performance but also enhances educational quality and student learning experiences, promoting students' overall development and growth (Lluch Molins & Cano García, 2023).

In conclusion, there is a mutually reinforcing relationship between learning analytics and personalized learning design, both of which provide important support for improving educational quality and student learning outcomes. By combining learning analytics and personalized learning design, educators can gain a better understanding of students' learning needs and behavior patterns, adjust teaching strategies to provide more effective learning support, and promote students' overall development and growth.

4. Discussion

The advancement of technology and data plays a pivotal role in learning analytics research. The rapid progress in information technology and data science has provided a robust foundation for the collection, storage, and processing of educational data. The widespread use of technologies such as learning management systems, online learning platforms, and mobile learning applications provides researchers with abundant sources of learning data (Zhang, Basham & Yang, 2020). This progress in technology and data has provided valuable research subjects and analytical tools for learning analytics, laying a solid foundation for a deeper understanding of the learning process. Moreover, the rise of personalized learning has sparked interest in learning analytics research. Traditional teaching models often fail to meet the personalized needs of diverse learners, making personalized learning a hot topic in today's education sector (Namoun & Alshanqiti, 2020). As a method capable of quantifying and evaluating the learning process, learning analytics offers strong support for personalized learning. Educators and learners alike seek to understand learners' characteristics and needs through learning analytics, aiming to provide personalized learning support and guidance accordingly. Furthermore, the impetus for educational reform has also driven the development of learning analytics research (Rumawatine, 2023). In today's society, there is increasing attention to education quality and effectiveness, making educational reform a critical agenda for governments and educational institutions worldwide. As a method capable of quantifying and evaluating the learning process, learning analytics provides robust support for educational reform. Policymakers and education decision-makers hope to optimize the teaching process and enhance student learning outcomes and satisfaction through learning analytics (Kabudi, Pappas & Olsen, 2021). Additionally, the rise of interdisciplinary collaboration has brought new opportunities and challenges to learning analytics research. Learning analytics is an interdisciplinary research field that involves knowledge and methods from various disciplines such as education, psychology, and computer science. Interdisciplinary collaboration promotes the cross-fertilization and innovative development of learning analytics research (Tsai, Perrotta & Gašević, 2020). Experts and researchers from different disciplines collaborate to explore the theory and practice of learning analytics, providing strong support for the in-depth exploration and expansion of learning analytics research.

The results of learning analytics research provide researchers with a range of important insights that have guiding implications for future research and practice. Firstly, personalized learning and instructional optimization are among the core objectives of learning analytics research. Researchers need to delve into the theories and methods of personalized learning, developing personalized learning models and tools applicable to various learning environments and disciplines (Lim et al., 2021). Through personalized learning, diverse learner needs can be better met, enhancing their learning outcomes and satisfaction. Secondly, the application scope and potential of learning analytics are vast, encompassing various aspects of education. Researchers should focus on the application of learning analytics in different educational stages and contexts, exploring its role in classroom teaching, online learning, mobile learning, and other areas (Pishtari et al., 2020). By integrating learning analytics technologies with educational practices, optimization of the teaching process and improvement of learning outcomes can be facilitated. Moreover, interdisciplinary collaboration is a key factor driving the development of learning analytics research. Researchers should actively promote collaboration and communication across different disciplinary fields, jointly exploring the theories and methods of learning analytics. Interdisciplinary collaboration helps broaden the perspective of learning analytics research and promotes innovation in theory and practice (Kew & Tasir, 2022). Additionally, researchers should pay attention to the application of learning analytics in educational policies and practices. Policymakers and education decision-makers need to understand the potential value and prospects of learning analytics, promoting the widespread application of learning analytics technologies in the field of education. By collaborating with policymakers and practitioners, researchers can translate the results of learning analytics research into practical applications, driving educational reform and enhancing educational quality (Viberg, Khalil & Baars, 2020, March).

The results of learning analytics research provide valuable insights for instructional designers, which are of significant importance in guiding instructional design and practice. Firstly, personalized learning is one of the core objectives of instructional design. Through learning analytics technology, instructional designers can better understand learners' personalized needs and learning characteristics, thereby designing instructional content and strategies that better meet learners' needs. Personalized learning design can enhance learners' motivation and engagement, promoting improvement in their learning outcomes (Herodotou et al., 2020). Secondly, learning analytics technology can provide real-time instructional feedback and support for instructional designers. By analyzing learning data, instructional designers can promptly identify learners' learning difficulties and issues, and accordingly adjust instructional content and methods to provide targeted instructional support and guidance. This personalized instructional feedback helps learners better understand and grasp learning content effectively. Additionally, learning analytics technology can assist instructional designers in better evaluating instructional effectiveness and learning outcomes (Kem, 2022). Through the analysis of learning data, instructional designers can understand learners' learning progress and achievements, and evaluate the effectiveness of instructional methods, and the quality of learning outcomes. This data-driven evaluation method can help instructional designers adjust instructional strategies promptly, optimize the teaching process, and improve instructional effectiveness. Moreover, learning analytics technology can provide guidance and support for instructional designers. By analyzing learning data and exploring learning patterns, instructional designers can understand learners' learning preferences and habits, and accordingly design instructional content and activities that align with learners' needs and interests (Ouyang et al., 2023). This personalized instructional design helps stimulate learners' interest and enthusiasm for learning, enhancing their learning outcomes and satisfaction.

5. Conclusion

In conclusion, according to the literature review and the results mentioned before, the integration of learning analytics into personalized learning represents a significant advancement in educational practices, with the potential to revolutionize teaching and learning experiences. By harnessing commercial applications and network analysis tools, educators can collect,

analyze, and report data generated during the learning journey, enabling a deeper understanding of student behaviors and needs. The benefits of learning analytics in personalized learning are manifold. Through personalized recommendations and instructional feedback, learning analytics helps identify learning gaps, improve learning behaviors, and enhance student engagement and performance. Moreover, it fosters a comprehensive understanding of students' learning habits and backgrounds, enabling the provision of effective pathways in personalized learning environments. Furthermore, learning analytics empowers educators to better meet student needs and enhance teaching and learning effectiveness. By analyzing student data and learning behaviors, teachers can tailor instruction and support, ultimately leading to improved educational outcomes and student experiences. However, despite its potential to enhance teaching and learning practices, the integration of learning analytics into personalized learning faces challenges such as privacy concerns and the oversight of qualitative indicators.

To be more specific, learning analytics is a powerful methodological approach that leverages data and technology to assess, analyze, and optimize the learning process. Its integration into personalized learning holds promise for improving educational outcomes and fostering positive learning experiences. Moving forward, addressing challenges and ensuring ethical practices will be crucial in maximizing the potential of learning analytics to transform education and enhance student success. As researchers and practitioners continue to explore and refine the application of learning analytics in personalized learning, the field stands poised to make significant strides in advancing educational practices and outcomes.

References

Abrahamson, D., Worsley, M., Pardos, Z. A., & Ou, L. (2022). Learning analytics of embodied design: Enhancing synergy. *International Journal of Child-Computer Interaction*, *32*, 100409.

Al-Badi, A., & Khan, A. (2022). Perceptions of Learners and Instructors towards Artificial Intelligence in Personalized Learning. *Procedia Computer Science*, *201*, 445-451.

Aparicio-Fernández, C., Torner, M. E., Cañada-Soriano, M., & Vivancos, J. L. (2023). Analysis of the energy performance strategies in a historical building used as a music school. *Developments in the Built Environment*, *15*, 100195.

Berding, F., Riebenbauer, E., Stütz, S., Jahncke, H., Slopinski, A., & Rebmann, K. (2022). Performance and Configuration of Artificial Intelligence in Business Education Learning Analytics Applications. A Content Analysis-Based Approach.

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, 100027.

Chen, K. Z., & Li, S. C. (2021). Sequential, typological, and academic dynamics of self-regulated learners: Learning analytics of an undergraduate chemistry online course. *Computers and Education: Artificial Intelligence*, *2*, 100024.

Cui, Y., Song, X., Hu, Q., Li, Y., Shanthini, A., & Vadivel, T. (2021). Big data visualization using multimodal feedback in education. *Computers & Electrical Engineering*, *96*, 107544.

del Pilar Gonzalez, N. A., & Chiappe, A. (2024). Learning analytics and personalization of learning: a review. *Ensaio: Avaliação e Políticas Públicas em Educação*, *32*, e0244234.

Diego, M., Carlos, G., & Jose, A. (2019). Adaptive learning objects in the context of eco-connectivist communities using learning analytics. *Heliyon*, *5*(11).

Doroudi, S. (2024). On the paradigms of learning analytics: Machine learning meets epistemology. *Computers and Education: Artificial Intelligence*, *6*, 100192.

Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. (2020). Multimodal learning analytics for game-based learning. *British Journal of Educational Technology*, *51*(5), 1505-1526.

Fachola, C., Tornaría, A., Bermolen, P., Capdehourat, G., Etcheverry, L., & Fariello, M. I. (2023). Federated Learning for Data Analytics in Education. *Data*, *8*(2), 43.

Han, J., Kim, K. H., Rhee, W., & Cho, Y. H. (2021). Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation. *Computers & Education*, *163*, 104041.

Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., & Rienties, B. (2020). How can predictive learning analytics and motivational interventions increase student retention and enhance administrative support in distance education?. *Journal of Learning Analytics*, *7*(2), 72-83.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, *2*, 100017.

Karaoglan Yilmaz, F. G., & Yilmaz, R. (2020). Student opinions about personalized recommendation and feedback based on learning analytics. *Technology, knowledge and learning*, *25*, 753-768.

Kem, D. (2022). Personalised and adaptive learning: Emerging learning platforms in the era of digital and smart learning. *International Journal of Social Science and Human Research*, *5*(2), 385-391.

Kew, S. N., & Tasir, Z. (2022). Learning analytics in online learning environment: A systematic review on the focuses and the types of student-related analytics data. *Technology, Knowledge and Learning*, *27*(2), 405-427.

Klašnja-Milićević, A., Ivanović, M., & Stantić, B. (2020). Designing personalized learning environments—The role of learning analytics. *Vietnam Journal of Computer Science*, *7*(03), 231-250.

Kustitskaya, T. A., Esin, R. V., Kytmanov, A. A., & Zykova, T. V. (2023). Designing an Education Database in a Higher Education Institution for the Data-Driven Management of the Educational Process. *Education Sciences*, *13*(9), 947.

Lim, L. A., Dawson, S., Gašević, D., Joksimović, S., Pardo, A., Fudge, A., & Gentili, S. (2021). Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses. *Assessment & Evaluation in Higher Education*, *46*(3), 339-359.

Lluch Molins, L., & Cano García, E. (2023). How to Embed SRL in Online Learning Settings? Design through Learning Analytics and Personalized Learning Design in Moodle. Journal of New Approaches in Educational Research, 12(1), 120-

138.
Mathrani, A., Susnjak, T., Ramaswami, G., & Barczak, A. (2021). Perspectives on the challenges of generalizability, transparency and ethics in predictive learning analytics. *Computers and Education Open*, *2*, 100060.
Muslim, A., Chatti, M. A., & Schroeder, U. (2022). Supporting Indicator Personalization and Platform Extensibility in Open Learning Analytics. *Technology, Knowledge and Learning*, *27*(2), 429-448.
Namoun, A., & Alshanqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, *11*(1), 237.
Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*, *20*(1), 4.
Pishtari, G., Rodríguez-Triana, M. J., Sarmiento-Márquez, E. M., Pérez-Sanagustín, M., Ruiz-Calleja, A., Santos, P., ... &Väljataga, T. (2020). Learning design and learning analytics in mobile and ubiquitous learning: A systematic review. *British Journal of Educational Technology*, *51*(4), 1078-1100.
Pospíšilová, L., & Rohlíková, L. (2023). Reforming higher education with ePortfolio implementation, enhanced by learning analytics. *Computers in Human Behavior*, *138*, 107449.
Rumawatine, Z. (2023). The Effect Of Personal Learning Models On Self-Confidence And Learning Outcomes To Play Soccer In Extracurricular Men's Soccer. *JIM: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, *8*(2), 864-873.
Schumacher, C., Klasen, D., & Ifenthaler, D. (2019). Implementation of a learning analytics system in a productive higher education environment. In *Emerging trends in learning analytics* (pp. 177-199). Brill.
Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, *107*, 105512.
Shorfuzzaman, M., Hossain, M. S., Nazir, A., Muhammad, G., & Alamri, A. (2019). Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment. *Computers in Human behavior*, *92*, 578-588.
Tempelaar, D., Rienties, B., & Nguyen, Q. (2021, September). Dispositional learning analytics for supporting individualized learning feedback. In *Frontiers in Education* (Vol. 6, p. 703773). Frontiers.
Tsai, Y. S., Perrotta, C., & Gašević, D. (2020). Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics. *Assessment & Evaluation in Higher Education*, *45*(4), 554-567.
Viberg, O., Khalil, M., & Baars, M. (2020, March). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 524-533).
Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, *122*, 119-135.
Wu, J. Y., Liao, C. H., Tsai, C. C., & Kwok, O. M. (2024). Using learning analytics with temporal modeling to uncover the interplay of before-class video viewing engagement, motivation, and performance in an active learning context. *Computers & Education*, *212*, 104975.
Yang, S. J., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, *2*, 100008.
Yilmaz, R., Yurdugül, H., Yilmaz, F. G. K., Şahin, M., Sulak, S., Aydin, F., ... & Ömer, O. R. A. L. (2022). Smart MOOC integrated with intelligent tutoring: A system architecture and framework model proposal. *Computers and Education: Artificial Intelligence*, *3*, 100092.
Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, *27*, 44-53.
Zhang, L., Basham, J. D., & Yang, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educational research review*, *31*, 100339.
Zhou, Y., Huang, C., Hu, Q., Zhu, J., & Tang, Y. (2018). Personalized learning full-path recommendation model based on LSTM neural networks. *Information sciences*, *444*, 135-152.

CSCL 视域下元认知支架对群体感知的促进作用研究

Research on the Facilitating Effect of Metacognitive Scaffolding on Group Awareness

李红慧[1*]，曹兰英[1]，张思[1]
[1]华中师范大学人工智能教育学部
[*]edulihonghui@163.com[1]

【摘要】 计算机支持的协作学习场景中群体感知有助于任务完成和知识建构，但学习者仍需要额外的支持以帮助其将群体感知转化为可操作的策略。元认知支架能以显性化干预的方式为学习者参与复杂的问题解决过程提供帮助，然而少有研究关注元认知支架对群体感知的促进作用。基于此，本研究聚焦 CSCL 视域下元认知支架对群体感知的影响，通过内容分析、滞后序列分析和认知网络分析等方法共同探索元认知支架对群体感知的促进作用。研究结果表明，元认知支架有助于学习者在群体感知支持的协作学习环境中调节自身的协作行为，优化小组协作问题解决的行为模式，同时元认知支架的设计需适应学习者协作问题解决过程的阶段性发展。研究为后续设计与应用元认知支架和群体感知提供理论和实践参考。

【关键词】 CSCL；元认知支架；群体感知；滞后序列分析；认知网络分析

Abstract: *Group awareness tools in computer-supported collaborative learning (CSCL) settings can facilitate task completion and knowledge construction, but learners still need additional support to help them turn group awareness information into actionable strategies. Metacognitive scaffolding provides a powerful aid for learners to engage in complex problem-solving processes in the form of explicit interventions. However, few empirical studies have focused on the facilitating effect of metacognitive scaffolding on group awareness tools. Based on this, this study focuses on the effect of metacognitive scaffolding on group awareness in the CSCL perspective, and jointly explores the facilitating effect of metacognitive scaffolding on group awareness through content analysis, lag sequence analysis, and epistemic network analysis. The findings suggested that metacognitive scaffolding helped learners to regulate their own collaborative behaviors and optimize the behavioral patterns of group collaborative problem solving in collaborative learning environments supported by group awareness tools, while metacognitive scaffolding needed to be designed to adapt to the stage-by-stage development of learners' collaborative problem solving process. The study provides theoretical and practical references for the subsequent design and application of metacognitive scaffolding and group awareness tools.*

Keywords: CSCL; metacognitive scaffolding; group awareness; lag sequence analysis; epistemic network analysis

1.前言

成功的协作学习离不开成员间的相互感知(Strauß & Rummel, 2021)。群体感知工具（Group Awareness Tool, GAT）以可视化方式将小组成员间难以直接观测到的认知、行为和社会群体感知及时准确地呈现给彼此，为学习者激发元认知意识以调节自身或同伴的协作行为提供了指引(张思等, 2023)。然而大部分学习者很难直接将获取到的群体感知转化为可操作的行为策略，这要求学习者具有较高的元认知调节水平(Järvelä 等, 2015)。元认知支架通过外部干预的

[1] 国家自然科学基金面上项目“面向大规模在线教育的学习者协作会话能力评估模型及干预机制研究”（No. 62077016 ）

途径为学习者有效解决协作问题提供了良好的支持和帮助(陈锋娟等, 2023)。相关研究表明元认知支架有助于学习者理解与运用群体感知信息，帮助学习者维系有效的协作交流，促进协作效果(Phielix 等, 2011)。然而以往研究多关注群体感知和元认知支架的单独作用，并且元认知支架对群体感知的影响机制多停留在理论层面，较少从实证视角关注元认知支架对群体感知的促进作用机制。关注元认知支架对群体感知的促进作用将有助于教师深入理解元认知支架的作用机制,为后续群体感知和元认知支架的设计与应用提供参考。因此,本研究聚焦 CSCL 视域下元认知支架对群体感知的影响，设计了群体感知工具支持的在线协作学习活动，采用准实验研究范式，通过内容分析、滞后序列分析和认知网络分析等方法共同探索元认知支架对群体感知的促进作用。

2.文献综述

2.1.CSCL 视域下的群体感知 成功的协作学习需要建立在成员间的相互感知上(Strauß & Rummel, 2021)。小组成员通过了解小组整体或小组其他成员个体的信息(被称之为群体感知),进而调节自身的元认知状态，并采取相应的行为促使有效协作的发生(张思等, 2023)。然而在时空分隔的计算机支持的协作学习环境下，小组成员很难自发地直接获取到小组整体或同伴个体的信息，因而增加了小组成员间协作的努力与困难程度，这也凸显了技术对群体感知支持的重要性(陈向东等, 2021; Ollesch 等, 2021)。群体感知工具通过为学习者提供小组整体/同伴间难以直接观测到的信息,减少群体成员间协调的努力,提高了协作学习的效率(梁妙 & 郑兰琴, 2012; 张思等, 2023)。目前相关研究已经证实 CSCL 环境下群体感知对学习者的协作过程会产生积极影响。例如，Engelmann 等(Engelmann & Hesse, 2010)通过向学习者呈现组内各成员绘制的知识概念图，帮助学习者及时比较自身与同伴的知识信息差距，提升了同伴间协作解决问题的效率和质量。Sangin 等(Sangin 等, 2011)基于自主开发的知识感知工具（Knowledge Awareness Tool，KAT）向学习者展示组内各成员知识掌握情况的信息,为学习者提供客观的知识线索，促进了同伴间有效的知识协商和沟通。由此可知，CSCL 视域下的群体感知有助于提升学习者的协作效果，促进学习者有效的协作解决问题。尽管相关研究已经证实在在线协作学习过程中群体感知的积极影响(Engelmann & Hesse, 2010; Sangin 等, 2011)，但学习者仍较难在群体感知工具支持的在线协作学习中受益，主要原因在于他们缺乏根据显示的同伴信息做出有效行为策略的能力(Strauß & Rummel, 2021)。这启示我们可通过进一步提供适当的元认知支架为学习者的协作学习过程提供帮助，激发他们有效应用群体感知调节协作过程，提升协作问题解决的能力。

2.2.元认知支架对群体感知的促进作用 元认知是一种高阶认知能力，对实现深度而有意义的学习具有重要作用(Garrison & Akyol, 2015)。美国学者 Flavell(Flavell, 1979)将元认知定义为个体对认知的监控和控制，既包括个体对自身认知的监控和控制，同时也包含个体对社会环境中其他人认知的监控和控制。元认知支架基于元认知的概念，通过外部干预的方式支持学习者反思评估当前学习中存在的问题和与期望目标间的差距，找到可解决问题或实现目标的策略(陈锋娟等, 2023; Kim & Hannafin, 2011)。元认知支架有助于学习者在协作过程中调节和评估，缩小学习者现实与目标间的知识差距，为学习者有效参与复杂的问题解决过程提供有力帮助(陈锋娟等, 2023)。CSCL 环境下学习者在应用群体感知以协作解决问题的过程中需要元认知支架的支持。群体感知为学习者了解与监控自身和同伴当前的协作状态提供了及时可靠的客观信息数据,元认知支架则有助于将群体感知转化为有效解决协作问题的可操作性策略,促进学习者有效解决协作问题，从而提升协作的效果。目前相关学者已开始关注 CSCL 环境下元认知支架对群体感知的影响(Siegel, 2012)。例如，Phielix 等(Phielix 等, 2011)在向学习者

提供同伴反馈评价的信息基础上进一步向他们提供了反思性的元认知支架，帮助学习者反思过去、现在小组存在的问题，进而为小组成员进一步规划后续计划和实施措施提供可操作的策略。该研究结果表明元认知支架能够促进学习者对群体感知的应用，并有效提升小组成员在协作问题解决过程中的认知和社会表现。然而，目前有关元认知支架对群体感知的促进作用的实证研究较少，元认知支架对群体感知的促进作用机制仍需进一步探讨，亟需更多实证研究探明元认知支架对群体感知的促进作用机制。因此，研究采用准实验的研究范式，通过对比有无元认知支架来探究 CSCL 视域下元认知支架对群体感知的影响。本研究的研究问题具体包括以下三点：①有无元认知支架支持，小组的协作问题解决行为体现怎样的频数差异特征？②元认知支架对小组协作问题解决行为序列会产生怎样的影响？③元认知支架对小组协作问题解决网络模式和发展过程会产生怎样的影响？

3.研究设计

3.1.研究背景及研究对象 在 2021-2022 学年秋季课程“信息技术教育应用”中，本研究以华中地区 H 大学 30 名师范专业的师范生为研究对象，开展了为期 8 周的在线协作学习活动。授课内容为信息化教学理论知识及应用实践。课程采取混合式教学方式授课，线上学生需在 QQ 群和小雅云平台完成在线协作学习任务；线下学生需参与课程学习并于学期末汇报展示小组作品。在线协作学习活动为 3 人一组，成员依据 PISA（2005）协作问题解决测试和先验知识测试结果进行分组。研究对象共分成 10 组，随机分成实验组和对照组两个大组，每个大组分别由 5 个小组构成。在协作学习过程中，学生需协同完成两项小组任务（教学设计和教学课件），在每项小组任务中，分别包含了三阶段（组内初稿——组内修改稿——组内终稿）。

3.2.实验设计 3.2.1.实验流程设计 在正式活动开始前，教师通过先验知识问卷和协作问题解决能力测试对学生进行分组，随后实验组和对照组两组学生同时开展 6 周的群体感知工具支持的在线协作学习活动，其中群体感知工具由助教将学生外显的群体感知呈现到实验组和对照组各自的小组群中。在每周的协作学习活动中，助教为实验组提供了元认知支架，对照组则不予提供。在活动结束后，实验组和对照组学生需要填写群体感知工具体验问卷，同时实验组学生需额外填写元认知支架体验问卷。

3.2.2.群体感知工具设计 在小组协作过程中，学生了解当前其他小组成员对任务的理解信息有助于提高小组协作问题解决的效果(梁妙 & 郑兰琴, 2012)。据此，研究从认知维度出发为学习者提供群体感知工具——个人思路图。个人思路图旨在通过可视化小组各成员的任务理解思路（也就是本研究中的群体感知），增进同伴对彼此想法的理解，进而促进同伴间协作解决问题的成效。群体感知工具的设计思路是：通过让组内各成员针对当前阶段任务先自行理解与思考，并将个人思路进行可视化以帮助组内各成员及时感知成员间的知识信息和逻辑结构，为后续小组协作讨论，完成协作任务提供有益支持。实验组和对照组的所有学生在每阶段任务发布后先行完成个人思路图的绘制，为避免组内成员存在划水抄袭现象，同时也为了保证各位成员能够充分发挥主观能动性并积极思考，各成员绘制完成个人思路图后由助教统一收集并发至小组群中供成员间针对个人的思路开展讨论。

3.2.3.元认知支架设计 为探索元认知支架对群体感知的影响，研究人员每周向实验组提供元认知支架。元认知支架旨在支持学生将个人思路图呈现的小组任务理解思路转化为可操作的协作行为策略。在本研究中，元认知支架具体设计思路为：根据课程中协作任务的要求和作品的评估标准（TPACK 量规），元认知支架以问题方式收集学习者对自身知识掌握程度的评价（例如：我对本组教学设计中所包含的学科内容的理解程度为……），以及对当前小组协作互动状态的评价（例如：我对当前小组的协作互动过程的满意度为……），并由研究

人员将这些信息转化为雷达图呈现给对应小组，帮助该小组成员及时感知当前小组知识薄弱的维度以及社会交互中存在的问题，并结合教师反馈的自适应建议，进行及时监控和反思，调整小组协作策略。基于上述设计思路，本研究中元认知支架具体实现功能是指在每周小组讨论开始前，学生根据个人思路图上显示的任务理解思路，结合教师反馈的自适应建议，反思当前小组在知识理解和社会交互维度上存在的问题，从而在后续小组讨论中明晰讨论内容和讨论方式。

3.3.编码框架 本研究采用 Von Davier 等(Von Davier 等, 2017)人提出的协作问题解决编码框架。该编码框架旨在通过分析学生的话语数据来探究学生在协作过程中表现的问题解决技能。该框架包含 4 个一级维度（共享、协商、调节和沟通）和 33 个子维度。共享维度用来表征学生通过分享资源和信息进行协作对话的活动，如分享任务要求等；协商维度用来表征同伴间针对彼此观点的异同点进行协商讨论的活动，如针对同伴观点表达不认可等；调节维度用来表征协调小组解决问题过程的活动，如表达对小组协作表现的满意等；交流维度用来表征团队的沟通氛围，如使用社交话语打招呼等。该编码框架的 4 个一级维度相互独立且并列，共同刻画了学生协作问题解决活动的全景图。

3.4.数据收集与分析 研究收集了所有实验组和对照组的在线协作会话数据，经数据清洗和整合后共有 10124 条有效数据，其中实验组有 3786 条数据，对照组有 6339 条数据。在数据编码阶段，两位编码人员先进行编码训练，在保证两人对编码框架理解趋于一致后，随机抽取 15%的文本数据进行预编码(Zhang 等, 2022)，编码单元为一条帖子，若出现一条帖子包含多个编码维度则将其进行拆分以匹配到对应维度,最终两位编码人员的编码一致性系数(Cohen's Kappa)为 0.896，具有良好的一致性。随后两位编码人员分开完成剩余会话文本的编码工作。同时，为进一步了解学生对群体感知工具和元认知支架的使用体验，课程结束后研究人员分别向实验组和对照组学生发放有关群体感知工具的使用体验问卷，其中实验组额外发放了有关元认知支架的使用体验问卷。该问卷基于 Davis(Davis 等, 1989)提出的技术接受程度调查和李艳燕等(Li 等, 2021)提出的学生对群体感知工具接受程度的调查问卷改编而成，包含有用性、易用性、满意度和乐用性等四个维度，问卷结果显示实验组和对照组学生对群体感知工具的使用体验均持积极正向的态度（$M_{实}$=4.24、$M_{对}$=3.97），实验组成员对元认知支架的使用体验也持积极态度（$M_{实}$=4.07）。

针对第一个研究问题，研究采用统计分析计算实验组和对照组协作问题解决行为的比例和频数，并通过卡方检验计算两组在各个维度上的差异，从而检验元认知支架对小组协作问题解决行为的影响。针对第二个研究问题，研究采用滞后序列分析分别对实验组和对照组的协作问题解决行为序列进行分析，探索两组的协作问题解决行为转化路径的差异。针对第三个研究问题，研究采取认知网络分析法（ENA，https://app.epistemicnetwork.org）分析实验组和对照组在整体上以及在各个阶段上的协作问题解决网络模式及其发展过程。

4.研究结果

4.1.元认知支架对协作问题解决行为频数和比例的影响 实验组在共享、协商、调节和交流四个维度上的频数占比分别为 5.57%、33.29%、23.36%、37.78%，而对照组在四个维度上的频数占比分别为 4.51%、30.97%、22.97%、41.54%。由描述性统计可知，交流维度是实验组和对照组的学习者在协作解决问题过程中占比最高的特征，这也体现出交流行为在学习者在线协作解决问题过程中的主导作用。值得注意的是，加入元认知支架的实验组在共享、协商与调节行为的占比明显高于只提供群体感知工具的对照组，这也反映出元认知支架有助于学习

者表达自身观点和协调小组任务过程。卡方检验的结果表明，实验组和对照组学习者的协作问题解决行为在协商、调节和交流三个维度上的占比差异具有统计显著性差异($p<0.001$)。

4.2.元认知支架对协作问题解决行为序列的影响 依据滞后序列分析的计算规则，残差参数（Z-scores）大于正 1.96 表明行为转换具有统计显著性意义，因而分别绘制了实验组和对照组的协作问题解决行为序列转换图（如图 1 和图 2 所示），其中各个节点代表协作问题解决的各个行为，连线代表两个行为之间转化频次的显著性值，箭头所指的方向即为行为转换的方向，连线的数值为显著性的残差参数（Z-scores）。Z 值越大连线越粗，行为序列的显著性越强。数据分析的结果表明，实验组和对照组的协作问题解决行为序列存在明显差异。具体表现为，实验组具有显著性意义的协作问题解决行为序列为：S→S、N→N、R→R、C→C、S→R，而对照组具有显著性意义的协作问题解决行为序列为：S→S、N→N、R→R、C→C。

4.3.元认知支架对协作问题解决认知网络模式与发展过程的影响 4.3.1. 元认知支架对协作问题解决认知网络模式的影响 实验组和对照组的协作问题解决保持相似的模式，即两组在“交流-调节”、“交流-协商”、“协商-调节”建立较强的连接，且通过曼-惠特尼非参数检验可知两组间的协作问题解决模式不存在显著性差异（$p=0.55$）。通过网络图的连线可以进一步发现两组间仍存在着细微的差异。其中，对照组在“协商-调节”、“协商-交流”的连接强度（连接系数为 0.49、0.64）略高于实验组（连接系数为 0.45、0.61）。在“交流-调节”上，实验组（连接系数为 0.58）的连接系数略高于对照组（连接系数为 0.54）。这一结果表明相较于对照组，元认知支架支持的实验组在协作问题解决的过程中更倾向于调节小组的协作过程以完成任务，而对照组更倾向于通过协商观点的方式去完成任务。

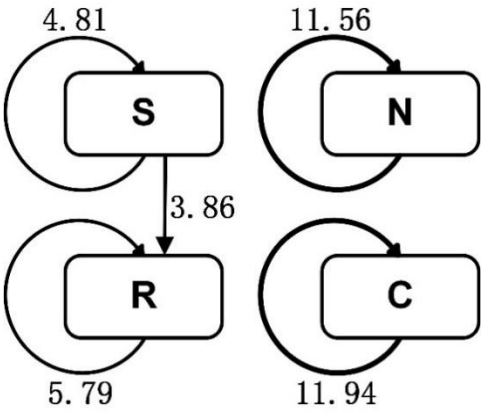


图 1 实验组协作问题解决行为转化图

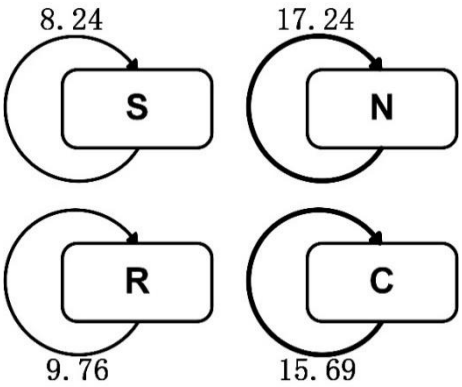


图 2 对照组协作问题解决行为转化图

4.3.2.元认知支架对协作问题解决网络发展过程的影响 为深入理解元认知支架对小组协作问题解决网络发展过程的影响，研究进一步利用认知网络分析对实验组和对照组在不同阶段下的协作问题解决过程展开动态、精细化分析。如图 3 所示，实验组和对照组在协作问题解决过程中的认知网络存在明显的阶段性差异，其中对照组在协作过程中保持着相似的协作问题解决模式，即在交流、协商与调节三个维度上建立了较强而稳定的联系，同时也与共享维度保持有连线。而实验组在协作活动前期与对照组的协作问题解决模式类似，即在交流、调节与协商三个维度上建立强而稳定的联系，同时与共享维度保持有联系，而在协作活动后期，实验组学生在调节维度的连线强度明显减弱，同时在协商维度的连线强度明显增强。

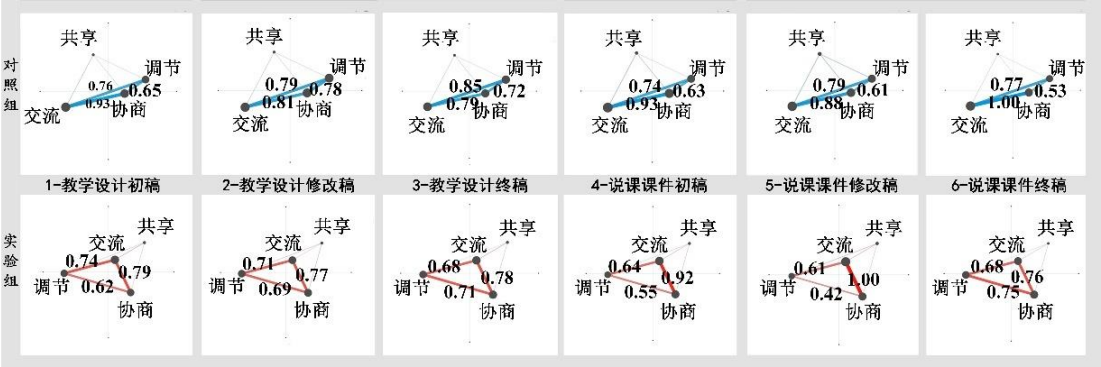


图 3　实验组和对照组各阶段协作问题解决过程的认知网络图

5.结论与讨论

5.1.元认知支架有助于激发小组成员表达自身观点，调节小组协作过程 统计分析的结果表明，两组学习者在交流维度上的行为比例最高，这与在线协作学习情境相契合，同伴间良好的沟通是小组成员协同解决问题，完成小组任务的基础(Zhang 等, 2022)。提供元认知支架的实验组在协商和调节维度上的行为比率显著高于对照组，而 ENA 的结果也证实了元认知支架有助于小组成员调节协作过程。这一结果表明，在群体感知工具支持下的在线协作学习活动中，元认知支架有助于小组成员采用调节行为策略引导小组协商讨论，完成小组任务。已有研究发现，学习者个体较难自发激活元认知意识，而通过有效的外部干预将有助于促进个体的反思和调节，以学习者自身行为数据为支撑的元认知支架有助于激发个体学习者的元认知反思，进而转化为可调节自身学习行为的策略(毛刚, 2018)。本研究进一步证实了在群体感知工具支持的在线协作学习活动中，元认知支架仍能发挥其优势，促进协作小组成员表达自身观点，优化个体和集体的学习行为。因此，在群体感知工具支持的在线协作学习活动中，教师和教学设计者可通过使用元认知支架，如问题引导式自我报告、学习分析技术支持的元认知反馈等方式为学习者搭建有效的线索，激发学习者针对性地反思并调节协作学习活动过程，从而协作解决问题。

5.2.元认知支架有效优化了小组协作问题解决过程，提高协作效率 滞后序列分析的结果表明，在群体感知工具支持的在线协作学习活动中，学习者能够及时了解小组/同伴的协作状态，并相互共享观点，交流协作，而元认知支架的加入有助于学习者进一步优化协作问题解决过程，小组成员在协作问题解决的过程中不再仅仅通过彼此分享、回答与任务有关信息的单一途径解决问题，而是会采取多线程方式寻找有效的协作策略，调节小组协作讨论过程，共同完成任务。出现这一结果的原因可能是无元认知支架支持的学习者在实际协作过程中较少获悉和监控小组的协作状态，因而很难灵活、及时地调整小组协作策略，解决协作过程中遇到的问题。已有研究表明元认知支架能够促进小组成员间的社会元认知互动，参与更多知识共建的协作活动，同时元认知支架能够引导学习者及时监控与评估，提升调节策略的准确性(陈锋娟等, 2023; Molenaar 等, 2014)。这与本研究得出的结论一致，即元认知支架能够有效优化小组的协作问题解决过程。本研究还进一步明确了元认知支架优化学习者协作问题解决行为路径的方向：共享观点指向协作调节。因此，教师和教学设计者在设计在线协作学习活动时，可为学习者提供当前协作阶段的元认知反馈，让学习者能够及时根据该阶段存在的问题做出调整，促进整个协作过程的高效开展。

5.3.元认知支架的设计需适应学习者协作问题解决过程的阶段性发展 ENA 的结果表明对照组的协作问题解决网络发展模式是相对稳定的，而加入元认知支架的实验组，其协作问题解决网络的发展模式是变化的，该变化主要出现在协作活动的后期。出现这一结果的原因可能是在协作活动后期实验组的学习者对知识的掌握趋于一致，元认知支架收集与反馈给学习者的元认知信息结果也基本相同，因而未能产生更多的指导性作用，且增加了学习者的外在认知负荷。因而在协作活动后期，实验组的协作问题解决发展模式产生了变化。此外，在课程结束后的访谈中，实验组学习者也表示元认知支架在协作前期比较有用，而在协作后期的作用不明显。因此，教师和教学设计者在设计元认知支架支持的多阶段协作学习活动时，有必要适应学习者协作问题解决过程的阶段性发展，特别是在协作活动后期可考虑撤离元认知支架或采取不同方式收集与呈现元认知信息，从而促进学习者更有效的协作。

6.结语

研究面向群体感知工具支持的在线协作学习情境，以元认知支架为干预手段并将其应用于在线协作学习活动中以促进小组有效的协作解决问题，为后续设计与应用元认知支架于群体感知工具支持的在线协作学习情境提供了理论和实践参考。研究结果表明在群体感知工具支持的在线协作学习环境下，元认知支架有助于激发小组成员表达自身观点，调节小组协作过程；同时元认知支架有效优化了小组协作问题解决过程，提高协作效率；元认知支架的设计需适应学习者协作问题解决过程的阶段性发展。本研究也存在一定的不足。首先，研究未单独设置对照组验证群体感知工具对小组协作问题解决模式的影响，群体感知工具和元认知支架对小组协作问题解决过程的独立和整合影响有待后续进一步探索。第二，研究仅采用了单一类型的元认知支架进行比较研究，未来也将进一步丰富元认知支架的类型，通过对比不同类型的元认知支架对群体感知的影响，以丰富元认知支架和群体感知的理论和实践成果。第三， 研究中数据收集的类别主要为学习者的协作会话文本数据，数据类型较为单一，对学习者协作问题解决模式的刻画仍需通过多模态数据进一步扩展。

参考文献

陈锋娟, 刘清堂, 张思, & 高倩倩. (2023). 面向小组互评的集体元认知支架模型及实证研究. 电化教育研究, *44*, 94-100+109. https://doi.org/10.13811/j.cnki.eer.2023.04.013

陈向东, 胡优立, & 张蕾. (2021). 共享调节学习中的群体感知:作用与机制. 远程教育杂志, *39*, 34–44. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2021.05.004

梁妙 & 郑兰琴. (2012). 支持协作学习的觉知工具:研究现状总结与思考. 远程教育杂志, *30*, 30–39. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2012.04.005

毛刚. (2018). 学习分析作为元认知反思支架的效能研究. 电化教育研究, *39*, 22–27. https://doi.org/10.13811/j.cnki.eer.2018.09.003

张思, 李红慧, 惠柠, 郭桐羽, 李秀晗, & 刘清堂. (2023). 多模态视域下的群体感知：内涵、功能与实现路径. 电化教育研究, *44*, 20–28. https://doi.org/10.13811/j.cnki.eer.2023.05.003

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, *35*(8), 982–1003.

Engelmann, T., & Hesse, F. W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, *5*(3), 299–319.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.

Garrison, D. R., & Akyol, Z. (2015). Toward the development of a metacognition construct for communities of inquiry. *The Internet and Higher Education*, *24*, 66–71.

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., Koivuniemi, M., & Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*, 125–142.

Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education*, *56*(2), 403–417.

Li, Y., Li, X., Zhang, Y., & Li, X. (2021). The effects of a group awareness tool on knowledge construction in computer-supported collaborative learning. *British Journal of Educational Technology*, *52*(3), 1178–1196.

Molenaar, I., Sleegers, P., & Van Boxtel, C. (2014). *Metacognitive scaffolding during collaborative learning: A promising combination. Metacognition and Learning, 9 (3), 309–332*.

Ollesch, L., Heimbuch, S., & Bodemer, D. (2021). Improving learning and writing outcomes: Influence of cognitive and behavioral group awareness tools in wikis. *International Journal of Computer-Supported Collaborative Learning*, *16*(2), 225–259.

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in human behavior*, *27*(3), 1087–1102.

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in human behavior*, *27*(3), 1059–1067.

Siegel, M. A. (2012). Filling in the distance between us: Group metacognition during problem solving in a secondary education course. *Journal of Science Education and Technology*, *21*, 325–341.

Strauß, S., & Rummel, N. (2021). Promoting regulation of equal participation in online collaboration by combining a group awareness tool and adaptive prompts. But does it even matter? *International Journal of Computer-Supported Collaborative Learning*, *16*(1), 67–104.

von Davier, A. A., Hao, J., Liu, L., & Kyllonen, P. (2017). Interdisciplinary research agenda in support of assessment of collaborative problem solving: Lessons learned from developing a collaborative science assessment prototype. *Computers in Human Behavior*, *76*, 631–640.

Zhang, S., Gao, Q., Sun, M., Cai, Z., Li, H., Tang, Y., & Liu, Q. (2022). Understanding student teachers' collaborative problem solving: Insights from an epistemic network analysis (ENA). *Computers & Education*, *183*, 104485.

基于反馈理论的课堂教学资源智能推荐框架设计

Design of an Intelligent Recommendation Framework for Classroom Teaching Resources Based on Feedback Theory

刘欣鑫[1]，钟林容[1*]，段金菊[2]
[1.2]西南大学教育学部
[*]linrongzhong032@163.com

【摘要】 人工智能时代，如何让教师精准找到符合学生学情的教学资源，是推进教育高质量发展的关键。本研究借助知识图谱和知识追踪等技术基础，以 Hattie 和 Timperley 的反馈模型为理论指导，依据课堂教学流程，设计课堂教学资源智能推荐框架。该框架包含目标反馈、过程反馈、结果反馈三个阶段，根据不同的反馈目的，提出相应的实践路径，以实现资源的智能化推荐。该框架能够很好地适用于课堂教学情境下的资源推荐，实现教学资源的选择与学习需求高度匹配，提升学生的知识境界与能力境界。
【关键词】 反馈理论；教学资源；智能推荐；框架设计

Abstract: *In the era of artificial intelligence, how to enable teachers to accurately find teaching resources that align with students' learning situations is the key to promoting high-quality education development. This study, leveraging technologies such as knowledge graphs and knowledge tracing, design a framework for intelligent recommendation of classroom teaching resources based on the feedback model proposed by Hattie and Timperley and aligned with classroom teaching processes. This framework comprises three stages: target feedback, process feedback, and result feedback. Depending on the specific feedback objectives, corresponding practical paths are proposed to achieve intelligent resource recommendations. This framework is well-suited for resource recommendations in classroom teaching contexts, enabling a high degree of alignment between the selection of teaching resources and learning needs, thereby enhancing students' knowledge and ability levels.*

Keywords: feedback theory;, teaching resources, intelligent recommendation, design framework

1. 引言

中共中央、中国国务院印发的《中国教育现代化 2035》中明确提出，要构建数字化、个性化、终身化的教育体系，加快建设学习型社会。这一目标的实现离不开教学资源的支持，资源作为信息化教育教学创新发展的核心要素，其发展方向在一定程度上决定了未来教育发展的格局与质量（柯超清等人，2021）。近年来，在国家相关政策支持与国家教育资源公共服务体系持续建设背景下，我国课堂教学资源在数量上已经有较大规模（陈明选等人，2022）。但是在长期的教育实践中，由于学生在个体上存在显著的差异性，他们的学习效果参差不齐，且学习状态也会随着时间和情境的变化而波动（刘海鸥等人，2019），导致教育工作者很难全面、深入地了解学生所需要的教育资源。因此，教育工作者难以在海量的教学资源中针对学生的学习需求进行教学资源的精准选择。

随着现代信息技术的快速发展，人工智能逐渐赋能于教育教学的各个方面，基于知识图谱技术、大数据平台的资源智能推荐系统等就能够有效解决教学资源与学生实际情况适应性不

足的问题。目前，借助大数据分析和人工智能等技术，深度挖掘学习者的学习数据，精准构建学习者模型，以满足学习者动态、模糊和个性化的资源服务需求为出发点，为其提供高质量、可扩展的学习资源服务是大数据教育应用研究的重要方向（张进良和叶求财，2019）。本研究利用 Hattie 和 Timperley 的反馈理论，梳理课堂教学环境下资源推荐的流程和方法，借助相关智能技术的支持，设计教学资源智能推荐框架，以期为相关资源服务机构和研究者提供借鉴和参考。

2. 课堂教学资源推荐的研究现状

目前，国内外学者对教学资源智能推荐进行了广泛而深入的研究。这些研究主要集中在以下几个方面：

第一，基于学习兴趣或学习风格的个性化资源推荐模型构建与优化。孙雨生等通过系统分析国内大数据赋能的智慧推荐资源管理技术，发现大数据赋能的用户兴趣建模技术多研究用户兴趣模型表示方法、初始化、进化，以提高用户兴趣建模性能、计算效能，支持智慧推荐服务全景、动态、精准感知用户兴趣（孙雨生等人，2023）。姜强整合学习风格和认知水平，构建个性化本体学习资源推荐模型，从推断学习风格及认知水平两个层面加以整合实现个性化本体学习资源推荐（姜强等人，2010）。

第二，学习资源的推荐算法研究。传统的推荐算法主要有协同过滤算法、基于内容推荐的算法以及混合推荐（Adomavicius & Tuzhilin, 2005）。例如，Yang En-yue 等提出了一种新颖算法，通过使用不同类型的用户反馈，以隐私感知的方式实现异构协同过滤（HCF）(Yang En-yue et al., 2021)；Trirat P 等在内容推荐中引入深度神经网络并提供学生问答推荐器，利用学习者的学习行为数据来提高推荐的精度（Trirat P et al., 2020）；Zhang W 等将协同过滤和视频基因组合在一起形成混合推荐（Zhang W et al., 2019）。随着人工智能的发展，研究者们又开发了基于知识图谱技术的学习资源推荐模型，以解决资源推荐中存在的数据稀疏和冷启动问题（李春英等人，2024）。

第三，场景化资源推荐策略。安璐等人分析公众文化认知特点并证实了公众文化认知存在局限性，从认知增强的角度提出国潮文化资源推荐策略，以帮助其获得更全面、更深层次的文化信息（安璐等人，2023）。刘东邻全面分析了高校科研用户的个性化科研资源服务需求，设计了融合推荐系统架构，最终实现了基于资源热度、项目内容相似度、相似用户协同过滤的 Top-N 推荐（刘东邻，2022）。杨莉娜等人基于 TPACK 框架构建了学科教师教研能力发展模型，并在此基础上对学科教师的教研资源需求进行分析，构建 TPACK 框架下的教研资源库和教研资源智能推荐模型（杨莉娜等人，2021）。

第四，技术对课堂教学资源智能推荐的支持。资源推荐所使用到的技术主要包括知识图谱技术、知识追踪和学习路径规划。知识图谱技术是将现实世界中的实体以网状形式关联起来，可以很好的挖掘现实世界中实体与实体之间的关系(胡泽文等人，2013)。知识追踪（Knowledge Tracing, KT）是实现自适应教育及个性化教育的有效实施途径，它可以通过过学生的历史答题情况，发掘学生的隐藏知识状态，精准评估学习者的知识状态，并根据此状态预测学生下一步的答题情况（闫秋艳等人，2024）。学习路径是学习者在学习过程中所经历的学习内容和学习活动的有序序列，学习者在经历该序列的过程中实现基本知识的学习、方法体系的掌握、问题解决与任务的完成，从而提升相应的能力（朱天宇，2018）。

上述研究为教学资源的智能化推荐奠定了一定的基础，提供了很好的思路和建议。但现有研究中提出的个性化资源推荐模型多适用于在线教育平台，针对课堂环境下的教学资源推荐策略涉及较少。现有推荐系统往往只是简单地将资源推送给用户，缺乏与课堂教学活动的有效融合。在推荐资源时，没有考虑到学生在课堂上的实时反馈和需求变化，这导致推荐的资

源可能无法与课堂教学活动形成有效的互补和支撑。因此，为了打造高质量课堂，很有必要利用现有智能技术设计适合课堂教学环境下的资源智能推荐框架。

3. 课堂教学资源智能推荐框架

3.1. Hattie 和 Timperley 的反馈模型

反馈一词是从工业领域引入到教学中的，其原始定义是指来自系统输出的信息，并被反馈到系统中（Lipnevich, A. A et al., 2021）。学者们一直在研究向学生提供有关他们在任务中表现情况的信息如何影响他们的学习，多年来，反馈的定义和理论不断发展，学者们不断积累证据，证明反馈在学生学习中的关键作用，开发了许多反馈模型来描述反馈的过程和机制。例如：Narciss 和 Huth 构建了一个交互式双反馈循环模型，该模型表明了两个反馈回路的相互作用过程和因素，这些反馈回路可以解释各种各样的反馈该模型指出当学生收到反馈时，解释学生反应的不仅仅是反馈消息的特征，相反，它是一个互动的过程，在这个过程中，学生和学生的特征创造了一种特定类型的反馈处理(Narciss, 2008; Narciss et al., 2004) 。Lipnevich 等人构建了一个互动反馈模型，该模型强调反馈总是在上下文中收到的，相同类型的反馈将根据班级、学术领域或任务的后果性质进行不同的处理（Lipnevich et al., 2016）。

Hattie 和 Timperley 的模型是迄今为止引用最多的模型，它对反馈的理论和实践都产生了巨大的影响(Lipnevich et al., 2021)。Hattie 和 Timperley 的模型中，反馈是代理人（例如，老师、同伴、书籍、父母、自我、经验）提供的关于一个人的表现或理解方面的信息(Hattie et al., 2007)。该模型认为反馈应有助于缩小预期目标与当前绩效之间的差距，为此，Hattie 和 Timperley 提出了不同的方法使学生和教师可以缩小这种差距。为了使反馈更有效，它应该回答三个问题，每个问题都代表一种反馈类型：我应该去哪里？我该如何到达目的地？接下来该怎么做？在本研究中将三个问题解释为：目标反馈、过程反馈与结果反馈。在目标反馈阶段，需要给学生提供需要完成的目标，即要告知学生他们需要达到什么水平，以便他们能够知道和评估自己的努力。在过程反馈阶段，教师需要向学生提供与目标相关的一些信息，比如预期标准、先前的成功或者失败经历。在结果反馈阶段，教师需要给学生提供一个结果，这个结果可能是更多的信息、更多的任务或者更多的期望，以便学生通过自我调节或者使用更多的策略和努力来实现目标。

3.2. 课堂教学资源智能推荐框架设计

在教育与教学领域，及时且精准的反馈是衡量学习者表现的关键，它为缩小既定学习目标与实际成就之间的差异提供了至关重要的信息。教育研究者和实践者能够利用这些反馈信息来调整和提升教学策略及方法，进而助力学习者更有效地开展学习活动。当反馈信息有效时，它不仅能揭示当前数字化学习资源整合设计中存在的不足，而且还能基于学习者的真实使用体验，将这些宝贵的“反馈意见”转化为对数字化学习资源体系在组织结构和表征形式上的实质性改进。这样的优化不仅加强了教师对教学资源组织和选择的能力，还更好地满足了学习者多样化的需求。因此，确保反馈信息的有效性并将其转化为资源整合的改进措施，是实现教学资源动态整合并促进教育效果的重要途径。

课堂是资源的载体，资源只有运用到实际教学中才能发挥它的真正价值。课堂要素主要包括学生、教师、教学目标、教学内容、教学活动、教学评价等，并在此基础上结合具体教学情境明确课堂各要素的具体内容（谢幼如等人，2016）。课堂教学流程主要包括教学目标、教学过程和教学评价。遵循动态性、有效性和智能性等原则，依据反馈理论模型，结合课堂课堂要素和课堂教学流程，本研究构建了课堂教学资源的智能推荐模型（如图 1 所示）。课堂教学资源智能推荐模型的重点是动态、精准地为教师课堂教学提供适应性服务，通过挖掘学生学习等多模态数据，预测学习者的资源需求，结合学科学习资源特性，实现实时、动态、

精准推荐资源。该模型由目标反馈阶段、过程反馈阶段和结果反馈阶段三个环节组成，在三个环节中分别推荐备课资源、生成性资源和巩固性资源。

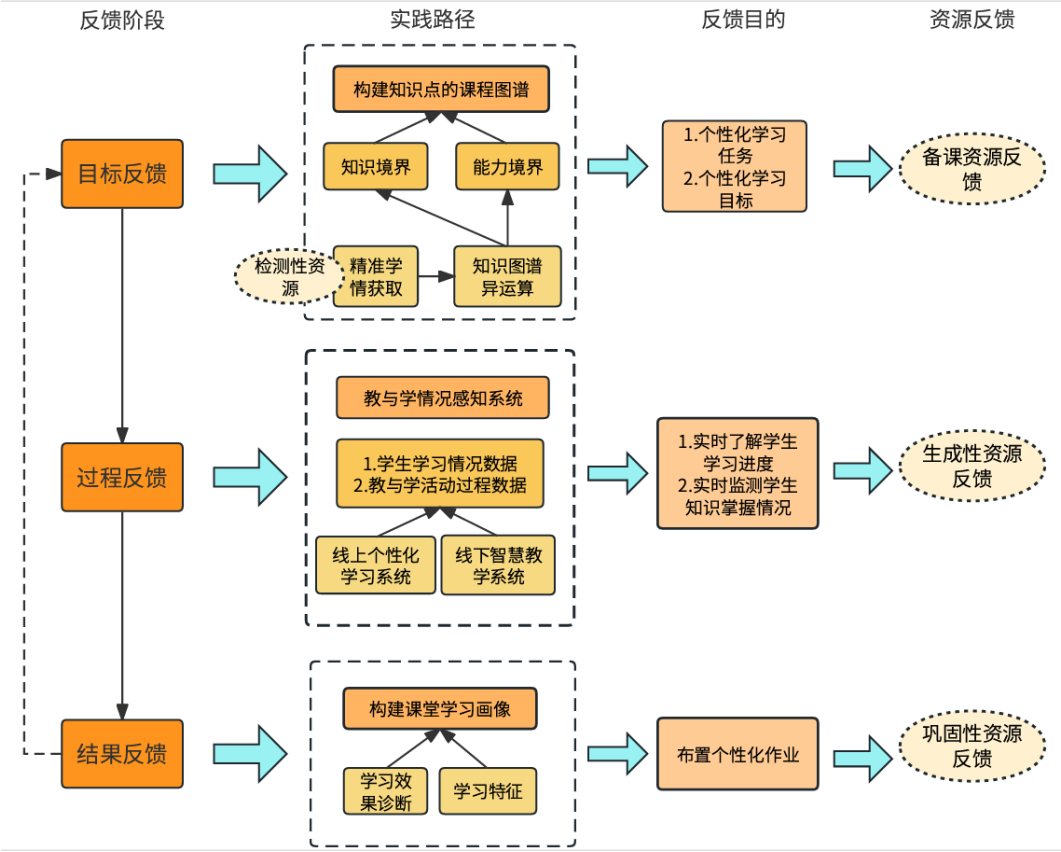


图 1 课堂教学资源智能推荐框架

3.2.1. 目标反馈阶段

在目标反馈阶段可以为学生制定个性化的学习任务和个性化的学习目标，并且为教师反馈备课资源，主要通过三个环节实现。首先需要知道学生需要掌握的知识与能力，即构建目标状态知识图谱，然后利用检测性资源精准获取学生学情，了解学生知识与能力的掌握程度，即构建学生课前状态的知识图谱，最后通过对目标状态知识图谱和课前状态知识图谱的异运算，得到学生的知识境界和能力境界，构建知识点的课程图谱（钟绍春等人，2020）。在这个过程中，知识境界包括知识点的记忆类内核问题与知识掌握程度间的关系，知识点的理解类内核问题及与知识掌握程度间的关系；能力境界是指应用、分析、综合、评价和创造五个层次的内核问题/任务与能力形成程度间的关系（谢幼如等人，2016）。能力境界涉及的五个层次分别为：应用、分析、综合、评价、创新。应用是指能够直接应用知识解决问题的能力；分析是指要对问题做剖析和转化后才能够应用知识解决问题的能力；综合则是指要对问题做剖析、分解和转化后才能够应用知识解决问题的能力；评价是指学生能够对所学知识和所形成的能力做系统梳理总结、判定自己实际水平的能力；创造则是指能够触类旁通归纳出新知识的能力。学生的知识境界和能力境界共同构成了知识点的课程图谱，即个性化学习任务和个性化学习目标。

3.2.2. 过程反馈阶段

在过程反馈阶段，利用线上学习系统和线下智慧教学系统收集学生的学习情况和教与学活动过程中的数据，对这些数据进行分析，可以实时了解学生学习进度和检测学生知识掌握情况。根据系统分析出的学生学习情况，教师在接收反馈智慧，可以实时调整资源内容以及资源应用策略等，此时，系统为教师提供生成性资源反馈。

4.2.3. 结果反馈阶段

在结果反馈阶段为学生设计个性化作业，提供巩固性资源反馈，通过两个环节实现。首先对学生学习效果进行诊断，获取结果性数据，在中小学教育阶段，学生的学习结果数据主要是由课堂练习、作业和考试完成情况，以及各类与课程相关的专题学习活动参与情况等数据构成。然后结合学生的学习特征，构建学生的课堂学习画像，课堂学习画像是指学生在课堂学习过程中所形成的问题解决能力及相应的知识体系，通过课堂学习形成的创新思维能力与综合素养，以及课堂学习中所表现出来的课程特有的认知能力和兴趣偏好等方面的特征。构建学生的课堂学习画像可以知道学生对于本堂课的学习还有哪些不足，据此推荐个性化作业。

4. 课堂教学资源智能推荐策略的优势分析

4.1. 课堂流程适配性

现有的资源推荐模型虽然在线上教育领域已经取得了一定的成果，但对于课堂环境下的应用还存在较大的局限性。当前，众多教师仍需依赖于个人经验来判断学生的学情，并据此选择相应的教学资源，这样的方式往往缺乏精准性和个性化，导致教学资源与学生实际学情之间存在脱节。本研究依据教师教学的环节，从教学目标、教学过程以及教学评价三个核心阶段出发，设计了课堂教学资源智能推荐框架。这一框架不仅能够确保教学资源的选择与教师的教学计划、学生的学习进度以及学习需求高度匹配，还可以为人机协同的智能教学系统开发提供理论参考。

4.2. 资源推荐精准性

现有的资源推荐模型大多采用求同思维，属于半个性化，个性化程度仍显不足。这种求同思维往往导致推荐结果过于局限，使用户陷入“信息茧房”的困境。本研究采用求异思维，对学习者学习行为的大数据进行深入分析，准确获取学生学情，并借助知识图谱技术，明确学生当前状态与目标状态之间的差距，通过精准的资源推荐提升学生的知识境界与能力境界。这种基于求异思维的资源推荐策略，不仅能够拓宽学生的学习视野，让他们接触到更多元化的知识和信息，还能够激发他们的学习兴趣和动力，促进全面发展。

5. 课堂教学资源智能推荐框架的典型应用场景

5.1. 基于教学资源智能推荐的备课系统

随着数字化教育的快速发展，各类教学资源不断涌现，如何高效地整合与利用海量教学资源，为教师提供个性化的备课支持，成为了亟待解决的问题。目前，为了克服教师个人备课的局限性，提高课堂教学质量，不少学校开展了集体备课，同时，部分教学教研平台也开始尝试开设集体备课功能。例如，希沃信鸽推出的集体备课功能支持多人协同操作，并自动归档备课资料，生成校本资源库，从而助力学校打造个性化题库。教师在筛选习题时，可根据题型与难度进行精准选择，同时，系统会显示已布置习题的平均正确率等信息，有助于教师判断题目难度及学生掌握情况。然而，现有备课系统大多注重教师之间的协作与教学资源的提供，缺乏教学资源智能推荐的支持，使得教师在筛选和利用教学资源时仍面临诸多困难。在线教育平台可汗学院就构建了知识地图来表示知识点之间的先后依赖关系，教师通过知识点之间的先后序关系以及学习材料的对应关系，可以获得学生的知识掌握情况，从而选择符合学生学情的教学资源。知识图谱作为一种体量巨大、在知识点与知识点以及知识点与教学资源之间建立连接的语义网络，能够在资源的个性化服务等方面发挥重要的基础性作用（LEO Y et al., 2006）。本研究提出的课堂教学资源智能推荐框架旨在通过检测获取群体学生学习大数据，并根据教与学路径的需要，帮助教师科学、精准地选择合适的教学资源。这一框架能够有效避免教师过度依赖个人经验，提升备课质量与教学效果。

5.2. 基于教学资源智能推荐的人机协同智能教学系统

智能教学系统是一种先进的计算机教学系统（张蓉菲等人，2019），它集合了人工智能、计算机科学、认知科学、思维科学、教育学和心理学等多个学科，为学生提供高度个性化和智能化的学习体验，能根据学生的需求和喜好调整学习内容和知识表示(孟亚玲等人，2020)。陈凯泉等通过对文献的综合分析，指出智能教学系统正朝着多模态学习分析、适应性反馈、人机协同等方向不断拓展（陈凯泉等人，2019）。人机协同重点关注教师如何与智能教学系统协同，为学生提供智能化的精准教学。依据本研究所提出的课堂教学资源智能推荐框架，人机协同的智能教学系统应该为教师提供精准化的教学目标、设计精准化的教学资源，同时进行精准化教学干预，满足学生个性化学习需求。首先机器能够全程检测课前、课内、课后的学习，负责易量化学生数据的采集、处理、画像建模，并以可视化方式呈现给教师，教师综合考虑机器反馈数据精准分析学情；其次机器要标签化处理与提取数字资源，建立教学资源与教学目标序列、学习需求映射关系并加以匹配，向学生、教师自动推送，教师依据学情选择适切的教学资源；最后机器能够依据学生课堂学习画像为学生推荐个性化作业，帮助教师设计分层作业，同时实时自动一对一答疑，减轻教师负担。

5.3. 基于教学资源智能推荐的个性化学习系统

实现个性化教育是新时代对未来教育的美好理想和普遍追求，也是智慧教育所追求的教育目标（李艳燕等人，2019）。美国知名在线教育平台 Knewton 基于学科知识图谱，在个性化学习应用中取得了较好的效果。该平台通过学科知识图谱定位学生困难类型，主要考虑了四种逻辑关系，即包含、评测、指导、先序（万海鹏和汪丹，2016），利用这些关系，将不同学科、不同学段、不同知识点以及所有的学习资源链接在一起，形成了一个巨大的知识地图，平台在诊断学生的认知水平后，为其推送个性化的学习资源。斯坦福大学教授创建的大型网络公开在线课程平台 Coursera，在学习者浏览某一课程时，平台将显示其他用户对该课程的评价，同时推荐对这门课程感兴趣的人同样感兴趣的其他课程。国内的科大讯飞基于学生学情、学科教学内容等数据，构建教育领域知识图谱，教师通过预设教学重点，使用知识点图谱引导学生学习，提升了学习效率，达到使学习者自主学习的目标(董永峰等人，2023)。上述案例都在一定程度上实现了个性化学习资源的推荐，但缺乏对学生学习后的成果进行及时有效的反馈和评价，并且在推荐学习资源时，平台仅聚焦于知识点的学习，忽视了学生能力的发展。本研究提出的教学资源智能推荐框架注重对学生学习成果的全面评估与反馈，旨在提供更加精准、全面的个性化学习支持。

参考文献

叶韵（2020）。融合学习情绪的学习者知识状态建模及在中学生在线学习中的应用。浙江工业大学。

刘海鸥、刘旭和姚苏梅等（2019）。基于大数据深度画像的个性化学习精准服务研究。**图书馆学研究**，**15**，68-74。

刘冬邻（2022）。高校科研资源的个性化融合推荐。**重庆大学学报**，**45(07)**，122-130。

安璐、陈苗苗和郑雅静（2023）。面向认知增强的国潮文化资源推荐策略。**图书情报知识**，**40(03)**，25-37+51。

孙雨生、范颖和祝博（2023）。国内大数据赋能的智慧推荐资源管理技术研究进展。**农业图书情报**，**12**，4-17。

闫秋艳、孙浩和司雨晴等（2024）。融合遗忘机制的多模态知识追踪模型。**计算机科学**，**03(24)**，1-12。

张进良和叶求财（2019）。大数据视阈下学习资源智能推荐模型构建。**湖南科技大学学报(社会科学版)**，**22(04)**，178-184。
张蓉菲、赵磊磊和李玥泓等（2019）。国外教育人工智能研究主题及趋势分析——基于 Web of Science 文献关键词的可视化分析。**现代教育技术**，**29(12)**，5-12。
李春英、武毓琦和汤志康等（2024）。融合知识图谱的学习者个性化学习资源推荐。小型微型计算机系统，**45(02)**，285-292。
李艳燕、张香玲和李新等（2019）。面向智慧教育的学科知识图谱构建与创新应用。**电化教育研究**，**40(08)**，60-69。
杨丽娜、陈玲和张雪等（2021）。基于 TPACK 框架的精准教研资源智能推荐研究与实践。**中国电化教育**，**02**，43-50。
陈明选、来智玲和蔡慧英（2022）。我国基础教育数字资源及服务:现状、问题与对策。**中国远程教育**，**06**，11-20+76。
陈凯泉、张春雪和吴玥玥等（2019）。教育人工智能(EAI)中的多模态学习分析、适应性反馈及人机协同。**远程教育杂志**，**37(05)**，24-34。
钟绍春、钟卓和张琢（2020）。如何构建智慧课堂。**电化教育研究**，**41(10)**，15-21+28。
唐烨伟、茹丽娜和范佳荣等（2019）。基于学习者画像建模的个性化学习路径规划研究。**电化教育研究**，**40(10)**，53-60。
姜强、赵蔚和杜欣等（2010）。基于用户模型的个性化本体学习资源推荐研究。**中国电化教育**，**05**，106-111。
胡泽文、孙建军和武夷山（2013）。国内知识图谱应用研究综述。**图书情报工作**，**57(03)**，131-137+84。
柯清超、林健和马秀芳等（2021）。教育新基建时代数字教育资源的建设方向与发展路径。**电化教育研究**，**42(11)**，48-54。
闾汉原、申麟和漆美（2011）。基于"态度"的知识追踪模型及集成技术。**徐州师范大学学报（自然科学版）**，**29(4)**，54-57。
万海鹏和汪丹（2016）。 基于大数据的牛顿平台自适应学习机制分析——"教育大数据研究与实践专栏"之关键技术篇。**现代教育技术**，**26(5)**，5-11。
孟亚玲、武帅和魏继宗（2020）。人工智能教育研究的现状、热点与趋势——基于 1979～2019 年 1043 篇人工智能教育文献的数据分析。**现代教育技术**，**30(03)**，120-123。
董永峰、王雅琮和董瑶等（2023）。在线学习资源推荐综述。**计算机应用**，**43(06)**，1655-1663。
谢幼如、夏婉和邱艺等（2023）。数字化转型赋能高质量课堂：逻辑要义、实践向度与典型样态。**中国电化教育**，**09**，50-58。
Adomavicius, G.,Tuzhilin, A. (2005). Toward the next generation of recom-mender systems: a survey of the state-of-theart and possible extensions. *lEEE Transactionson Knowledge and Data Engineering, 17(6),* 734-749.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77(1),* 81-112.
HUANG, Z. et al. (2019). Exploring multi-objective exercise recommendations in online education systems. *Proceedings of the 28th ACM International Conference on Information and Knowledge Managemen*t, New York: ACM, 1261-1270.
LEO, Y., UREN, V., & MOTTA, E. (2006). SemSearch; a search engine for the semantic web. *Lecture notes in computer science,* (10), 238- 245.

Lipnevich, A. A., Berg, D., & Smith, J. K. (2016). Toward a Model of Student Response to Feedback. *Human Factors And Social Conditions In Assessment*,169-185.

Lipnevich, A. A., Murano, D., Krannich, M., & Goetz, T. (2021). Should I Grade or Should I Comment: Links Among Feedback, Emotions, and Performance. *Learn and Individual Differences* ,vol.89(p.102020).

Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *In Frontiers in Education*, Vol. 6 (p. 720195).

LIU, Y. P. et al.(2016). Collaborative learning team formation: a cognitive modeling perspective. *Proceedings of the 21st International Conference on Database Systems for Advanced Applications*, Dallas, Apr 16-19.

Narciss, S. (2008). Feedback Strategies for Interactive Learning Tasks. *Handbook of Research on Educational Communications and Technology, 3rd Edn*, 125–143.

Narciss, S., & Huth, K. (2004). How to Design Informative Tutoring Feedback for Multi-media Learning. *Instructional Design for Multimedia Learning*, 181–195.

Trirat, P., Noree, S., & Yi, M. Y. (2020). Intelligent MOOC: intelligent online learningframework for MOOC platforms. *13th International Conference on Educational Data Mining*,682-685.

Yang En-yue, Huang Yun-feng & Liang Feng et al. (2021). FCMF : federatedcollective matrix factorization for heterogeneous collaborative filte. ring. Knowledge-based Systems, 220, 106946.1-10694610.

Zhang, W., Wang, Y., & Chen, H. et al. (2019). An efficient personalized video recommendation algorithm based on mixed mode. *lEEE Inter-national Conferences on Ubiquitous Computing and Communications(lUCC) and Data Science and Computational Intelligence(DSCl) and Smart Computing, Networking and Services (SmartCNS)*, 367-373.

基于人机协同的学生提问交互活动设计

Design of interactive student questioning activities based on human-machine collaboration

付玉娜*，赵赫璇，马国威，雅古，区文诗，孙伟华，王妍莉
西北民族大学
217619566@qq.com

【摘要】 随着信息技术的飞速发展，人机协同在教育领域中的应用日益广泛，作为一种新型的教育模式，逐渐成为研究和实践的热点。本研究基于批判性思维培养的政策需求和我国高等教育领域学生提问的现状及挑战，提出了基于人机协同的学生提问交互活动设计方案，并通过实际案例分析进行了详细的设计和阐述。初步实践结果表明，该活动能够显著提高学生的提问积极性，促进批判性思维的发展，为教育信息化背景下的教学理论体系提供了新的实践路径和思路。

【关键词】 学生提问；人机协同；交互活动设计；批判性思维

Abstract: *With the rapid development of information technology, the application of human-machine collaboration in the field of education is becoming increasingly widespread. As a new type of education model, it has gradually become a hot topic in research and practice. This study is based on the policy needs of cultivating critical thinking and the current situation and challenges of student questioning in the field of higher education in China. A design scheme for student questioning interaction activities based on human-machine collaboration is proposed, and detailed design and explanation are carried out through practical case analysis. The preliminary practical results indicate that this activity can significantly improve students' questioning enthusiasm, promote the development of critical thinking, and provide new practical paths and ideas for the teaching theory system under the background of educational informatization.*

Keywords: Student questioning; Human machine collaboration; Interactive activity design; Critical thinking

本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学本科教学质量提升项目省级一流课程《信息技术教学法》（项目编号：2022SJYLKC-37）资助；受“中央高校基本科研业务费专项资金”(项目编号：31920240127)资助。

1. 前言

2007 年美国 21 世纪技能联盟在“21 世纪技能框架体系”的学习与创新 4C 技能中旗帜鲜明地提出了批判性思维的关键能力（彭正梅和邓莉，2017），各个国家在建构和发展面向 21 世纪的核心素养框架以及面向未来的学校教育变革图景的过程中，也都将批判性思维作为不可或缺的关键素养，并形成了特有的理论框架。因此学生需要具备批判性思维以适应充满挑战的世界，教育应关注培养学生的批判性思维。已有研究发现学生提问和批判性思维水平两者之间紧密相连，保罗·弗雷尔将“问题提出教育”作为强调批判性思维的隐喻（张玲、宋乃庆和蔡金法，2019），纳尔顿和李发现培养学生问题提出能力的过程可以促进其批判性思维的发展(Nardone, C. F., & Lee, R. G. , 2010)。由此在一定程度上反映出问题提出活动不仅需要批判性思维的参与，还可作为批判性思维的发展工具。

而课堂提问的价值正是促进学生批判性思维的发展(Stenton, A. , 2011)，因此将学生作为课堂提问的主体是一种很好的教育理念，以课堂提问价值理性主导的课堂教学会使学生真正

成为课堂提问的主体并表现出对学生主动提出问题的理解(Rop, C. J. , 2002)。但研究发现当代大学生提问能力存在问题意识薄弱、缺乏质疑精神、知识建构能力不强、批判性思维欠缺、元认知监控能力下降、提问积极性降低等问题（龚一鸣，2022），在数字化环境下，虽然在线学习平台和学习社区为学生提供了提问和互动的空间，但仍存在缺乏有效反馈和指导、信息过载和学生筛选困难等问题（Liu, D., & Chen, B. , 2020)，因此我国高等教育领域需要创新性地设计和改进学生提问交互活动，以提升其效果和质量。

2. 理论基础

建构主义教学理念强调学生通过与环境的互动，基于已有知识构建新的理解，在这种理念下，课堂提问不仅是教师向学生提问，也包括学生向教师和其他学生提问，从而促进学生的主动学习和知识建构（吴青平，2015）。

人机协同理论强调学生和机器的分工和有机协同，提升学习的效率和效能，促进学生智慧和机器智能的共同增长，实现人机协同的教育智慧创生，完成超越人类智慧和机器智能的复杂任务，促进学生的知识建构、认知发展、思维提升和智慧养成（王一岩、刘淇和郑永和2024）。

3. 学生提问交互活动设计

本研究在学生提问活动中融入人机协同理论，为学生提供了互相学习和给予反馈的机会，学生可以从协作讨论和人机互动的过程中，获得从不同背景和角度看待同一问题的解答，这种多样性可以帮助学生获得更全面的视角，从而发展更为全面和深入的批判性思维。

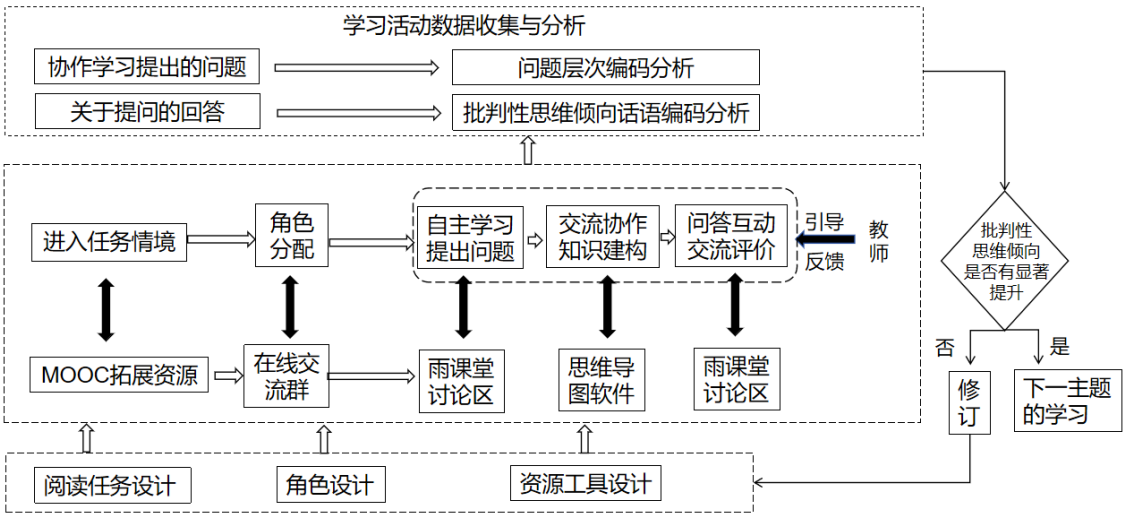


图 1 基于人机协同的学生提问交互活动设计

3.1. 任务及问题设计

本研究基于教育技术学本科专业中理论性较强的课程，任务特点是基本概念的梳理、分析与建构，通过 MOOC、智能问答系统、在线交流平台和在线协作工具等实现人机协同技术的工具，辅助学习者进行提问学习活动，教师通过 MOOC 和雨课堂发布指导性问题，引导学生利用智能问答系统思考讨论，学生通过 MOOC 平台与学习资源进行交互，通过在线交流平台或在线协作工具与同伴讨论问题，利用智能问答系统“智谱清言”辅助自主思考和解惑。

3.2. 角色设计与划分

角色类型包括自发角色和脚本角色（Palincsar, A. S., & Herrenkohl, L. R. , 2002)，应用于在线讨论的脚本化角色可分为功能性角色和认知性角色两类，对应的学习成果预期分别是解决问题和学术研讨，

学术研讨类任务注重思维过程，认知性角色是指在协作探究思维化成果的过程中，每个角色所承担的思考处理信息、认知参与等内隐性职责（王智颖、翟芸和吴娟，2021），强调学习者是否能够从协作的过程中获得思维方法的提升以及合作能力的培养（郭晓枫，2013）。本研究基于理论性较强的课程进行设计，注重对学生批判性思维和提问水平的提升，因此采用郭晓枫对认知性角色的划分，包括启动者(Starter)、支持者(Supporter)、提问者(Questioner)、挑战者(Challenger)、推理者(Arguer)和控制者(Timer)6 种角色（郭晓枫，2013）。在整个协作过程中各成员角色不变，根据实际分组人数，组内商定一人多角或一角多人，一个角色的扮演者可以不止一人，也即是说对于有些出现频率较高的角色可以分由多人承担。并且自己和他人的角色身份公开于在线交流群中，学生在每个人基本的协作分工任务基础上理解并承担自己的角色，可以参考提示语发言。

角色(roles)	功能描述(function description)	提示语（scripts or prompts)
启动者 (Starter)	对提出的任务进行初步分析，在协作过程中不断地提出观点，引领他人参与讨论，启发集体智慧，在一定的时候做出总结，促进协作地进一步发展	首先，我认为…；我们可以首先确保…；让我们开门见山；让我多说几句；从这个故事中，很明显；让我们现在解决问题；
支持者 (Supporter)	支持他人提出的观点，给予肯定地回应，并积极寻找相关的依据来支撑，注重思维的承接性	我同意，因为…；这是正确的；我明白你在说什么；一个例子是…；我读过…；
推理者 (Arguer)	承接他人的会话内容做补充说明，对论点、论据做出深入的分析，给出令人信服的推理过程，帮助他人进行深入地理解，注重严密的推理和思考的逻辑性	我认为两者都对，因为…；总结一下…；通过讨论，我们可以…；所以你的意思是…；这听起来很棒，因为…；
提问者 (Questioner)	积极回应他人，总是在寻找问题，以推动协作会话地继续进行，注重批判性看待问题	为什么；你说的是什么意思；你能说更多吗；你怎么认为；关于这事有其他看法吗；你在哪里看到的；你在哪里听到的；
质疑者 (Challenger)	总是对他人提出的观点进行反对或是大胆质疑，有理有据地“唱反调”的同时启发他人对某一观点进行更为批判性地思考，防止众人的盲目信从	我不同意，因为…；你为什么这么说；请说明原因；有证据吗；但我们能信吗；我们难道不需要更多证据吗；一个反对的论点是…；另一种观点可能是；我觉得有些不同；事实是这样吗；我不同意，因为；我不太确定；
控制者 （Timer）	在整个协作会话中起协调作用，控制讨论的节奏，合理分配讨论时间，在恰当的时候做出总结，避免讨论跑题或是在某一小问题上停留过久，当观点不同的人发生争执时能够及时解决，避免影响整体讨论进度	你愿意吗；我们可以吗；很抱歉；好的我们继续吧；

图 2 认知性角色设计

3.3. 资源工具的设计

3.3.1. 雨课堂

在自主学习和提出问题环节，协作小组在自主学习和协作交流的过程中，针对学习资源思考产生的问题经小组讨论可发布在雨课堂讨论区，其他小组可即时看到其他组提出的问题，并从中任选 1-2 个其他组的问题准备回答；在问答互动和交流评价环节，各小组成员可通过综合与智能AI问答、同伴交流、自主学习的结果，形成对问题的解答，在各小组对抽选的问题进行回答后可进行对问题的评价，雨课堂讨论区的点赞功能支持学生间的互评，教师也可通过点赞和口头评价。

3.3.2. 思维导图软件

在交流协作和知识建构环节，各小组成员在自主学习MOOC资源后通过在线交流群和思维导图软件进行协作交流，梳理相关概念特征，完成知识建构，交流不同的看法。

3.4. 活动评价设计

教师在活动过程中收集学生提出的问题，并利用基于布鲁姆教育目标分类的提问水平评价量表（王陆和彭玏，2020）对问题水平进行编码，并将答案依据批判性思维倾向测量表进行编码，在活动前后依据彭美慈修订版的批判性思维倾向测量表（CTDI-CV）进行批判性思维倾向问卷测试（彭美慈和汪国成等，2004）。

表 1 基于布鲁姆教育目标分类的提问水平评价量表

问题水平及层次		表征	举例
低水平问题	知识性水平	简单对所学知识进行回忆、复述	……的概念是？ 什么是……？ ……是否是……？
	理解性水平	在复述基础上，用自己的语言表达出来	用自己的话叙述……？ ……的物理意义是？ 怎么理解……？ ……和……的异同是？
	运用性水平	在情景中运用相应的概念、观点、公式等，情景要加以变换	举例说明……？ ……在……是如何运用的？ 怎样利用……公式解决……问题？
高水平问题	分析性水平	用所学知识，分析事实、结论、价值，并对其之间的关联加以思考、论证	……为什么……？ ……与……的关系是什么？ 为什么……具备……的价值？
	综合性水平	抽象、概括所学的内容，综合输出	如果……会……？ 根据……你能……？ 运用……单元知识，你能解决……问题么？
	评价性水平	对知识、材料加以科学价值判断	……的科学价值体现在哪里？ 你对……知识（公式）有什么看法？ 你认为……如何？为什么？

表 2 批判性思维倾向测量表

批判性思维特质	定 义
寻找真相	对寻找知识抱着真诚和客观的态度。若找出的答案与个人原有的观点不相符，甚至与个人信念背驰，或影响自身利益，也在所不计
开放思想	对不同的意见采取宽容的态度，防范个人偏见的可能
分析能力	能鉴定问题所在，以理由和证据去理解症结和预计后果
系统化能力	有组织，有目标地去努力处理问题
批判思维的自信心	对自己的理性分析能力有把握
求知欲	对知识好奇和热衷，并尝试学习和理解，就算这些知识的实用价值并不是直接明显
认知成熟度	审慎地作出判断、或暂不下判断、或修改已有判断。有警觉性地去接受多种解决问题的方法。即使在欠缺全面知识的情况下，也能明白一个即使是权宜的决定有时总是需要的

4. 教学活动案例

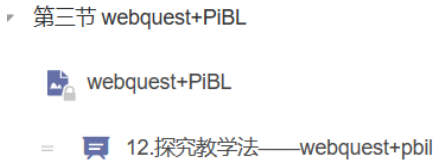
本案例的教学内容为《信息技术教学法》课程中“信息化教学方法”模块下的 Webquest 和 PBL 教学法，课时为 2 学时，有智慧教室的教学环境支持。

表 3 基于人机协同的学生提问交互活动案例

学习主题	**信息化教学方法-Webquest 和 PBL 教学法（2 学时）**
学习目标	理解 Webquest 和 PBL 教学法的内涵；领会 Webquest 和 PBL 教学法的本质；应用 Webquest 和 PBL 教学法进行教学设计；
学习重点	理解 Webquest 和 PBL 教学法的内涵并与探究式教学法作分析区分是学生学习的重点
学习难点	学习难点在于在理解的基础上如何应用 Webquest 和 PBL 教学法进行教学设计
教学设计思路	(本单元教学设计思路，包括： ①教学内容：教学内容包括探究式教学法的两个典型案例：

Webquest 和 PBL 教学法；

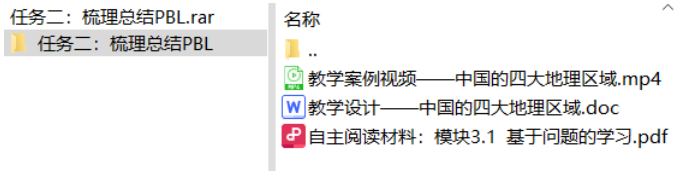
②教学资源：为学生提供了教学课件、自主阅读材料以及雨课堂学习平台等教学资源；



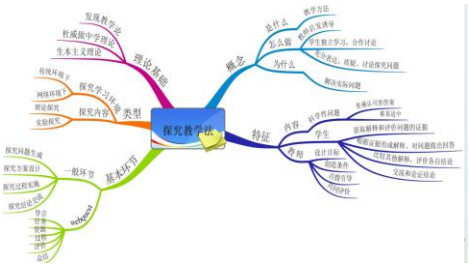
③教学流程及方法：混合式教学，利用雨课堂随机点名和弹幕等功能促进交流互动，激发学生学习兴趣；

1. 课前回顾（15min）：梳理回顾情境教学法和探究式教学法，便于学生理解 Webquest 和 PBL 教学法

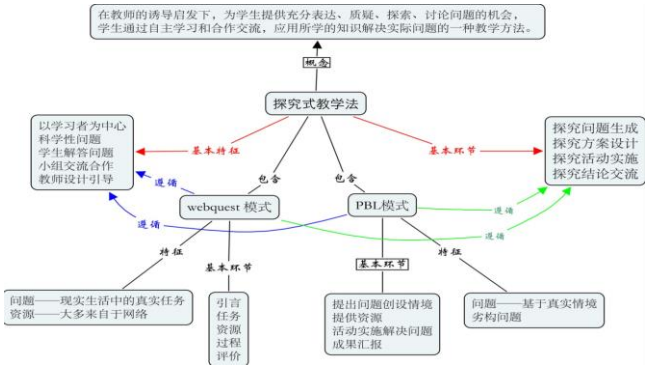
2. 自主协作，提出问题（30min）：小组协作研读 Webquest 和 PBL 来深入学习探究式教学，学生在自学过程中独立思考提出疑惑的相关问题，组内讨论自行解答一些问题，无法解答或有争议的问题可作为小组问题向其他组进行提问；



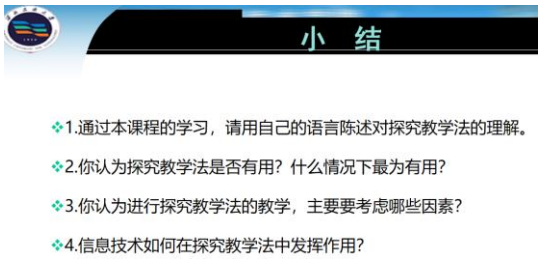
3. 汇报评价，回答反馈（20min）：总结梳理 Webquest 和 PBL 教学法的概念、特征、理论基础，以及和探究式教学的异同，全面梳理知识框架，提问组向其他组提问问题，随机抽选或自愿成为回答组，结合智能问答系统在组内展开讨论并组织答案，提问组并根据回答进行反馈评价；



4. 教师讲授（15min）：总结讲授 Webquest 和 PBL 教学的设计；



5. 思考问题（10min）：通过上节课关于探究式教学法的概念

<table>
<tr><td></td><td colspan="3">学习和本节课对探究式教学法两种典型模式的案例学习，以及三者间联系导图的梳理，在此基础上延申出关于探究教学法的几个开放性的思考问题，激活学生思维，提高学生参与度，促进知识的内化；
 

④教学评价：课后通过自测巩固反馈，通过自学评测及时评价学生的自学效果。</td></tr>
<tr><td rowspan="5">教学过程</td><td>教学环节</td><td>学习活动</td><td>设计意图/教师支持</td></tr>
<tr><td>课前回顾</td><td>回顾情境教学法和探究式教学法</td><td>梳理其他信息技术教学法，便于学生理解 Webquest 和 PBL 教学法</td></tr>
<tr><td>自主协作提出问题</td><td>以小组为单位，任意从 Webquest 和 PBL 教学法中选择一种教学方法进行讨论</td><td>研读探究教学的两个典型案例 Webquest 和 PBL 来深入学习探究式教学，提出问题</td></tr>
<tr><td>汇报评价回答反馈</td><td>总结梳理 Webquest 和 PBL 教学法的概念、特征、理论基础，以及和探究式教学的异同</td><td>全面梳理 Webquest 和 PBL 教学的知识框架，引发学生思考，组间提问并回答反馈</td></tr>
<tr><td>教师讲授</td><td>总结讲授 Webquest 和 PBL 教学的设计</td><td>在学生充分自学和思考讨论的基础上进行总结讲述，利于学生深度学习</td></tr>
<tr><td></td><td>思考问题</td><td>思考并回答有关探究式教学的一些问题</td><td>通过探究式教学法的概念学习和两种典型模式的案例学习，以及三者间联系导图的梳理，在此基础上延申出关于探究教学法的几个开放性的思考问题，激活学生思维，提高学生参与度，促进知识的内化</td></tr>
<tr><td>课后作业与评价</td><td colspan="3">基于学习任务单要求学生完成指定 MOOC 的学习；</td></tr>
<tr><td>总结与反思</td><td colspan="3">结合其他信息技术教学法的学习，便于学生理解不同教学方法的内涵特征，对学生的分析总结能力提出了要求</td></tr>
</table>

5. 总结与反思

本研究针对我国高等教育领域学生提问的现状及挑战，提出了一种基于人机协同的学生提问交互活动设计方案。通过实际案例分析，我们发现该活动能够显著提高学生的提问积极性，促进批判性思维的发展。这为教育信息化背景下的教学理论体系提供了新的实践路径和思路。虽然初步实践结果表明该活动具有一定的效果，但样本量有限，未来研究可以进一步扩大样本量，探讨人机协同的学生提问交互活动在不同教育阶段、不同学科领域的适用性和效果，以提高其普适性和推广价值。

参考文献

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

王陆和彭玏（2020）。2015—2019 年中小学课堂高阶问题特征图谱。电化教育研究，41(10):65-72113。

王一岩、刘淇和郑永和（2024）。人机协同学习：实践逻辑与典型模式。开放教育研究，（30），65~72。

朱伶俐和刘黄玲子（2006）。CSCL 实时交互效果影响因素的实证研究。开放教育研究，（5），51~55。

张玲、宋乃庆和蔡金法（2019）。问题提出：基本蕴涵与教育价值。中国电化教育，(12)，31-39。

吴青平。建构主义教学理念下的课堂提问研究。广州大学，2015。

俞树煜、聂胜欣和王国华（2016）。促进批判性思维发展的在线学习活动设计：一项基于设计的研究。电化教育研究，37(10)，49-56。

郭晓枫（2013）。计算机支持协作学习的角色设计及实证研究。华东师范大学。

龚一鸣（2022）。大学生问题能力培养与提升研究。哈尔滨师范大学，000331。

彭正梅和邓莉（2017）。迈向教育改革的核心:培养作为 21 世纪技能核心的批判性思维技能。教育发展研究，37(24)，57-63。

彭美慈和汪国成等（2004）。批判性思维能力测量表的信效度测试研究。中华护理杂志，（9）：644~647。

Liu, D., & Chen, B. (2020). Fostering student engagement in online discussions: A review of empirical research. Computers & Education, 103, 104-117.

Nardone, C. F., & Lee, R. G. (2010). Critical inquiry across the disciplines: Strategies for student-generated problem posing. College Teaching, 59(1), 13-22.

Palincsar, A. S., & Herrenkohl, L. R. (2002). Designing collaborative learning contexts. Theory into practice, 41(1), 26-32.

Rop, C. J. (2002). The meaning of student inquiry questions: A teacher's beliefs and responses. International Journal of science education, 24(7), 717-736.

Stenton, A. (2011). How could a dialogic approach to teaching transform students' learning? A discussion with reference to medical education. Higher Education Research Network Journal, 4, 15-22.

素养本位下生成式人工智能在小学科学探究实践中的应用研究

Research on the Application of Generative Artificial Intelligence in Elementary Science Inquiry Practice under the Literacy-Based Education Paradigm

杨宁，杨锦，包正委
福建师范大学
1367233652@qq.com

【摘要】 智能时代下，科技与社会深度交融，科学素养已成为现代公民必备的关键素质。科学教育肩负着提升全民科学素养的重任，其中，科学探究实践被认为是支撑学生科学素养形成的主要路径。然而，在传统教育课堂中，“科学探究”、“科学实践”等理念的落实一直饱受诟病。本研究立足于既有文献并结合实际课堂观察，梳理出当前我国小学科学教育课堂所面临的若干问题。提出将生成式人工智能技术融入科学教学，围绕教学目标、教学过程、技术支持、教学环境四个方面构建了课堂教学模型，希望能为基础教育课堂中科学探究实践教学的优化提供参考。

【关键词】 生成式人工智能；探究实践；科学素养；科学探究；小学科学

Abstract: *In the age of intelligent tech, where science and society are inseparable, scientific literacy is essential for modern citizens. Science education must enhance this literacy, with hands-on inquiry central to its development in students. Yet, traditional classrooms struggle to effectively implement such approaches. This study, combining literature review and classroom observation, identifies issues in Chinese primary science education. It proposes integrating generative AI into teaching, structuring a model around goals, process, tech support, and environment, to guide the optimization of inquiry-based, practical science instruction in elementary education.*

Keywords: *Generative AI, Inquiry-based Practice, Scientific Literacy, Science Inquiry, Elementary Science*

1.前言

《义务教育科学课程标准（2022 年版）》（以下简称“新课标”）强调科学素养是学生发展的重要目标，将科学教育要培养的学科核心素养分解为“科学观念、科学思维、探究实践、态度责任”四个部分，并提出要在小学科学教育阶段“加强探究实践”（中华人民共和国教育部，2022）。从原先强调的“科学探究”过渡到学科核心素养中的“探究实践”，《新课标》指出，小学科学不仅要培养学生的科学探究能力，更要在此基础上结合技术与工程实践来培养适应社会发展的“小科学家”“小工程师”，克服传统教育中知识建构与实践活动相脱离的短板。然而，在传统教育课堂中，与培养学生科学素养相关的“科学探究”“实验教学”“科学实践”等教学理念的落实一直饱受诟病，其中不乏出现“背实验”“看实验”“假问题”等教育现象。这不仅阻碍了学生科学素养的培养，更对提升全民科学素质发出了重大挑战。在科技迅猛发展的新时代，将 ChatGPT、虚拟现实等新的人工智能技术引入基础教育课堂，以人工智能赋能科学教育可能会成为解决这些教育难题的一个契机。

2.科学素养是现代社会公民必备的关键素质

早在20世纪50年代就出现了科学素养这个概念。虽然对“科学素养是什么”这个问题，学界一直未达成共识，但是对具有科学素养的人的特征表述上大概都涵盖了以下几个维度：理解科学知识和科学过程（Garfield E，1988）、科学态度（Haney R，1964）、科学探究的本质（BouJaoude S，2002）、科学地解释现象和阐述数据进行论证（OECD，2015）、科学身份的认同感（OECD，2025）等。

科学素养是国民素养的重要组成部分，是现代社会公民适应科技时代应该具备的关键素质。我国一直高度重视提升全民科学素养，并将青少年的科学素养培育作为重点任务。《国务院关于印发全民科学素质行动规划纲要（2021-2035年）的通知》中明确指出，要建设世界科技强国和社会主义现代化强国，需要科学素质建设担当更加重要的时代使命，我国2025年公民具备科学素质的比例要达到15%，2035年要达到25%（中华人民共和国国务院，2021）。2023年5月，教育部办公厅印发《基础教育课程教学改革深化行动方案》中，将“科学素养提升行动”列为本次课程教学改革的重点任务之一，贯彻落实教育、科技、人才“三位一体”战略要求，把三者协同的链条起点前置到基础教育阶段（郑永和等人，2023）。科学素养是科学教育的起点和归宿，而小学科学作为科学教育的开始，在促进青少年科学素养的养成中起着重要作用，基础教育阶段是青少年科学素养养成的关键时期。因此，我国的科学课程标准始终都是围绕着科学素养来制定的，在2001年颁布的《全日制义务教育科学（3～6年级）课程标准（实验稿）》中指出小学科学课程要以“培养学生的科学素养”为宗旨，将科学课程的总目标指向科学素养（杨晓梦，2023）。《义务教育小学科学课程标准（2017年）》进一步强调了小学科学课程中培养学生科学素养的重要性，指出学生在了解必要的科学知识和技术的基础上要具备运用其解决实际问题的能力（中华人民共和国教育部，2017）。而2022年颁布的新课标其突出特点之一，就是对学科核心素养进行了界定，指出小学阶段的科学素养培养应该围绕着“科学观念、科学思维、探究实践、态度责任”四个维度落实。

3.科学探究实践是支撑科学素养形成的主要路径

3.1. 从“科学探究”到“科学探究实践”的转变

科学活动的本质在于探究，小学科学是一门基于探究的学科。新课标当中提出的“探究实践”是对以往“科学探究”的进一步完善与发展，在原本探究的基础上增加了工程技术元素，强调让学生在“做中学”“学中思”。这实际上是响应了国际科学教育界中倡导的从“科学探究”向“科学实践”转型的号召。这里面提到的实践，并不是对科学探究的完全推翻重建，而是对课堂中僵化的科学探究机制作出的完善，是一种基于科学探究，又高于科学探究的一系列认知的、社会的行为活动（林兆星，2020）。科学探究实践的方式和过程是多元化的，没有固定的模式，但完整的过程一般包括情境创设、提出问题、作出假设、制定计划、收集证据、处理信息、得出结论、表达交流、反思评价、应用迁移等环节（胡卫平等人，2021），在真实的教学中并不要求一定要经历所有的环节。与传统的科学探究相比，探究实践更关注的是探究过程中的具体活动，其起点更强调真实的问题情景，活动的各个环节要求具有连续性，系统性和逻辑性（张红霞等人，2023），致力于改变传统课堂上探究活动与知识建构相脱离的倾向（林兆星，2020）。这与目前课标中提倡的大单元教学理念不谋而合，从某种程度上来说，基于项目的大单元教学或许是小学科学课堂进行有效科学探究实践的重要途径。

3.2. 科学素养的发展寓于科学探究实践活动之中

在国内外的科学教育变革研究中，“科学素养”、“科学实践”和“基于探究的科学教学”这些词往往都是一起出现的。在以“探究”为核心的小学科学课堂中，其科学探究实践无疑是支持学生科学素养培养的有力手段。“对科学领域相关知识、观念的认识与掌握是落实课程标准的前提基础，而探究实践是学生形成科学素养的主要途径（Hsiao C et al.,2019）。”

有研究者通过对国内外科学素养研究的关键词进行聚类分析发现，国内外学者提出了科学探究、实验教学、社会性科学议题、科学本质教学、科学史教学等众多旨在提升学生科学素养的教学方法策略，其中最为重要的就是科学探究和实验教学（郝琦蕾等人，2023），而两者在落实到课堂中的时候往往又是分不开的。对于在 PISA 测试中科学素养成绩常年稳居前列的芬兰学生来说，芬兰赫尔辛基大学教育科学学院雅里·拉沃宁（JariLavonen）教授指出其关键就是在教学的过程中重视实践性、探究性学习（钱文丹等人，2023）。学生科学素养的提升离不开课堂中科学探究实践活动的开展，以培养学生科学素养为导向的教学，其定位就是要让学生通过发生在真实情境下的探究促进新旧知识的融合，在动手的实践活动中进行知识建构，最终促进学生科学素养的提升。

4. 当前小学科学探究实践活动中的教学困境与契机

在小学科学教育领域，“探究实践”占据着重要的地位，然而，在实际教学情境中，这一理念的贯彻与落实仍旧面临一系列亟待解决的问题。本研究通过整理既有文献资料并结合课堂实地观察，将目前小学科学面临的探究实践难题总结为以下几类：探究问题选择困难、探究过程形式机械、资源与环境配置有限、课堂互动不足、教师定位不明与能力欠缺、学生探究水平不足、探究实践局限于课堂以及课堂评估机制不完善。具体分析如表 1。

表 1 小学科学探究实践课堂现存问题

维度	阐述说明
探究问题	①学生缺乏问题意识，无法提出问题；②探究问题由教师提供，“提出问题”这一重要环节被简单处理，忽视了学生的提问能力培养；③学生分不清“简单问题”与值得探究的“科学问题”，无法进行选择；
探究过程	①把科学探究作为科学学习的唯一方式，认为探究就是做实验的刻板印象；②科学探究教学中的各个环节和要素被僵化为一系列机械、固定的程序（林兆星，2020），课堂上的科学探究实践活动按照程序进行；③“探究”与“实践”相分离，实践活动与知识建构缺乏目标一致性；
资源配置	①提供的资源不足，无法支持学生进行有效的科学探究实践（硬件资源和软件资源）；②资源缺乏多样性，仅局限于文字或者图片；③无法搭建一个公平、自由、开放的科学探究实践环境；④新技术（人工智能、虚拟现实等）的引入欠缺，停留在旧的技术支持教学阶段（交互式电子白板、PPT 等）；
课堂互动	①师生互动：受限于大班教学，教师无法及时关注到每一个学生，师生互动不足导致学生产生迷思概念或学习积极性受挫；②生生互动：小组讨论常常是组长主导的活动，缺乏质疑与争论，更多是为了寻求一个权威的答案（张红霞等人，2023）；
教师水平	①教师的知识基础不够：对于学生超出预设的发现和问题无法及时给出反馈，甚至选择忽视；②教师的课堂把控能力欠缺：在指导学生进行科学探究实践的过程中，往往出现两种极端，一种是教师过多引导导致许多环节缺乏学生思考，另一种是教师缺乏引导导致学生科学探究实践难以进行下去（应向东，2006）。
学生能力	小学阶段，尤其是低年级学段的学生探究行为多是不完善的，再加之教师引导不到位，其探究实践活动往往因为年龄小、知识储备少、动手能力差、思维不成熟等导致学生“心有余，而力不足”。而高年级学段的学生也容易陷入浅层探究的泥泞里面；
局限课堂	探究活动局限于课堂，课本开放性不够。当前大多数科学探究学习仍是以课堂为主阵地，活动资源局限于课本，学生的思维活动被限定于教材和课堂的狭小空间，教学资源与环境相对封闭（陈明选等人，2020）。
教学评估	课堂评价体系主要侧重于理论知识的考核，而对探究过程的深度、能力提升等方面的评估相对匮乏。此外，单一的评价方式往往不能够反映学生的实际科学素养和进步情况。

科技浪潮席卷全球，各种新的技术手段层出不穷，推动着教育数字化变革稳步发展。面对小学科学课堂中这些亟待解决的教育难题，新兴的人工智能技术将成为一股强大的赋能力量。

5. 生成式人工智能是赋能小学科学探究实践教学的强大力量

5.1. 生成式人工智能介绍

2022 年 11 月 30 日，美国人工智能研究实验室 OpenAI 推出最新自然语言处理技术支持的人工智能聊天机器人——ChatGPT（Chat Generative Pre-trained Transformer），是一款基于 Transformer 架构的生成式预训练语言模型（杨宗凯等人，2023）。它可以根据用户输入的问题或指令，自动生成连贯、有逻辑且高度类人化的文本回应，凭借其强大的自然语言生成能力，能够为用户提供自然流畅的互动体验。以 ChatGPT 为代表的生成式人工智能，正在全球范围内引领着一场深刻的科技革命，给基础教育变革带来了机遇，同时也发出了挑战。然而，结合我国具体的教育情况来看，受到地域限制，将 ChatGPT 引入课堂仍旧是个难以解决的问题。从本土化发展来看，使用国内研发的其他可替代的生成式人工智能模型（比如讯飞星火认知大模型、百度文心一言和通义千问等平台）来辅助科学课堂，可能是目前更可行的方案。

5.2. 生成式人工智能的对话反馈特征

生成式人工智能的两个主要功能分别是模式识别和智能生成（苗逢春，2023），这两者相辅相成，共同构建起一个既能精准响应用户需求，又能深度理解并适应对话情境的智能化交互体系。与传统的技术支持教学手段相比，生成式人工智能最突出的优势就是它能够在课堂中与学生进行类人化的对话互动。它不仅可以根据用户的指令实时生成内容，还可以基于对话的上下文演绎推理，进行内容迭代和个性化的反馈。将其应用于小学科学课堂教育中，其对话反馈具有明显的六个特征：“丰富性、实时性、迭代性、包容性、精准性以及可定制性。”具体分析见表 2。

表 2 生成式人工智能的对话反馈特征

特征	
丰富性	能够基于大量训练数据生成丰富多样的内容，例如可以根据用户的指令生成文字、视频、图片、代码、诗歌等；
实时性	实时互动反馈，能够快速响应指令并生成相关内容，能够识别用户的语气或情绪变化，动态调整生成内容以满足需求；
迭代性	模型本身可以根据不断学习进行迭代改进，其次对话中可以根据用户的反馈进一步补充或者完善前面的生成内容，以逐步贴近用户期望；
包容性	塑造包容、开放的对话环境，能够理解和再现多元文化背景下的表达方式，能够生成满足不同文化、语言和社会群体特点的内容；
精准性	基于目标导向原则，当有明确任务要求或目标约束时，能够根据指令准确捕捉到用户的需求并生成对应内容；
可定制性	在与用户交互的过程中，可以根据上下文实时调整模型状态或更新生成策略以适应用户偏好，可以通过对对话内容的推理演绎为用户推送个性化资源；

5.3. 生成式人工智能适用于小学科学探究实践教学

生成式人工智能以其高适应性、强互动性与个性化能力，契合小学科学探究教学中激发学生兴趣、促进深度理解、支持协作学习及培养创新能力等需求。从教学对象来看，小学生正处于具体运算向形式运算转变的阶段，具有极强的好奇心同时又受到自身能力水平的制约而难以有效开展探究，生成式人工智能能够根据学生的理解水平调整反馈话语，为学生提供符合认知能力发展的支架支持；从教学过程来看，生成式人工智能能够围绕“探究前准备、探究过程引导、探究后反思”三个环节有效地补充并强化探究式科学教育中的教学活动，帮助学生主动生成探究问题并为促成探究实践活动顺利开展提供多样化的资源支持；从教学目标来看，小学科学教育的最终目标是培养学生的科学素养，生成式人工智能通过提供适应性的

支持显著强化了小学科学课堂中的探究过程，在强体验中促进了学生提问能力、科学表达能力、问题解决能力、批判性思维等高阶能力的培养。

5.4. 生成式人工智能在小学科学探究实践教学中的角色定位

生成式人工智能进入小学科学课堂，其教师的部分角色可由生成式人工智能替代，通过整理已有研究和对课堂观察数据进行定性分析，确定了生成式人工智能应用于小学科学课堂探究实践中的四种角色定位（对话者、资源提供者、教学助理以及评价助手）。

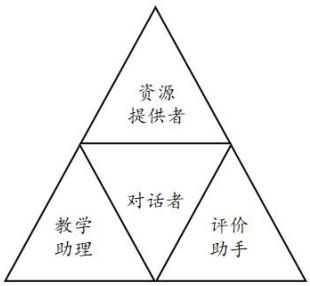


图 1 生成式人工智能在小学科学探究实践课堂中的角色

5.4.1. 对话者：与学习者进行自然语言交互并提供即时反馈

以“对话互动”为核心的生成式人工智能，凭借着其出色的语义理解和生成能力重塑人机交互的边界。在与学习者的交互过程中，它通过对学习者的指令进行识别和解析，理解学习者的问题、需求或困惑，及时生成有针对性的回答和解释。小学科学探究实践课堂当中，生成式人工智能作为“对话者”可以发生关键作用。首先，“对话者”能够实时地回应学生关于科学概念、现象、实验步骤等疑问，并为其提供精准、多元且符合小学生认知水平的解释，确保学生在探究过程中及时获得所需的知识支持；其次，“对话者”在与学生的对话可以对其进行探究式引导，帮助学生细化探究主题，确定科学的探究问题，以及帮助他们提出合理的假设并设计实验方案。最后，通过与“对话者”的对话，学生可以培养自己的提问能力和科学表达能力，同时，通过对生成内容的判断、对比和归纳总结也可以培养学生的逻辑思维和批判性思维。

5.4.2. 资源提供者：根据学习者的需求提供个性化的学习资源

生成式人工智能凭借对海量数据的深度学习与训练，具备提供多元化教学资源的能力。在小学科学探究实践课堂当中，生成式人工智能扮演着“资源提供者”的角色，可以为学生和教师提供丰富的、与探究主题紧密相关的学习材料和符合学生个性化需求的定制化学习资源。例如，学生在进行科学探究之前，“资源提供者”可以为学生提供关于探究项目的背景资料，帮助学生对其快速熟悉并提出探究问题。其次，新课标中明确强调要在小学科学中进行跨学科学习，“资源提供者”可以通过提供跨学科资源帮助学习者进行跨学科知识的联结，促进学科的交叉融合，摆脱局限于课本和课堂的探究。

5.4.3. 教学助理：协助教师进行教学管理与设计

生成式人工智能引入小学科学课堂，推动教学模式从传统的“师—生”二元结构转变为“师—生—机”三元结构，课堂教学从“以教为中心”转变为“以学为中心”，教师在课堂中的角色功能被重塑和更新（杨宗凯等人，2023）。“人师—机师”协同作用的双师课堂成为目前生成式人工智能赋能教育的趋势，生成式人工智能引入小学科学课堂可以弥补部分教师水平不足的问题。作为帮助教师承担部分教学责任的“教学助理”，生成式人工智能可以在小组讨论中充当“智能导师”的角色，为学习者进行及时的答疑解惑。其次，也可以通过提供学习支架的形式帮助学习者搭建知识框架。例如，项目结束后学习者在撰写反思报告的时候，“智能导师”可以为学习者提供支架，提示学习者从不同方面反思自己的学习过程。其次，

“教学助理”也可以辅助人类教师对课堂进行调控和管理，例如通过回收学习者的数据进行分析之后使用智能分组功能根据实际需求进行同质分组或异质分组，提升学生协作学习的效率。值得注意的是，在人机协同的课堂当中，人工智能不是教师的潜在替代品，教师在课堂中作为促进者和引导者的积极角色不能被忽略（Bower M，2019）。

5.4.4. 评价助手：帮助学生进行自主监控学习的助手

为了引导学生实现由知识到素养的进阶，新课标强调要进行“教—学—评”一体化的单元整体设计教学。在以培养学生科学素养为最高目标的小学科学探究实践教学中，教师应该立足“学习化评价”的理念，把评价也作为一种学习，重视过程性评价、表现性评价的运用，坚固多元评价主体，多维度的评价方法，让评价从判断目标达成度的单一功能转向成为学生自主监控学习的重要手段，生成式人工智能可以作为学生学习化评价的助手、数据来源而不仅仅是对学生学习成果的评判者（宋萑等人，2023）。学生通过参与评价活动，如自我评估、同伴互评、反思日志等，能够识别自己的知识盲点和技能短板，从而主动调整学习策略和计划。而作为学生学习评价助手的生成式人工智能可以实时收集、分析学生的数据，为学生生成个性化的学习分析报告，帮助学生进行元认知层面的思考，提高其学习效果。

6.基于生成式人工智能的小学科学探究实践目标与过程模型

基于对生成式人工智能在小学科学课堂中的角色定位和对科学探究实践教学的目标与过程（胡卫平等人，2021）解析，从教学目标、教学过程、技术支持、教学环境四个方面构建基于生成式人工智能的小学科学探究实践目标与过程模型，如图 2 所示。

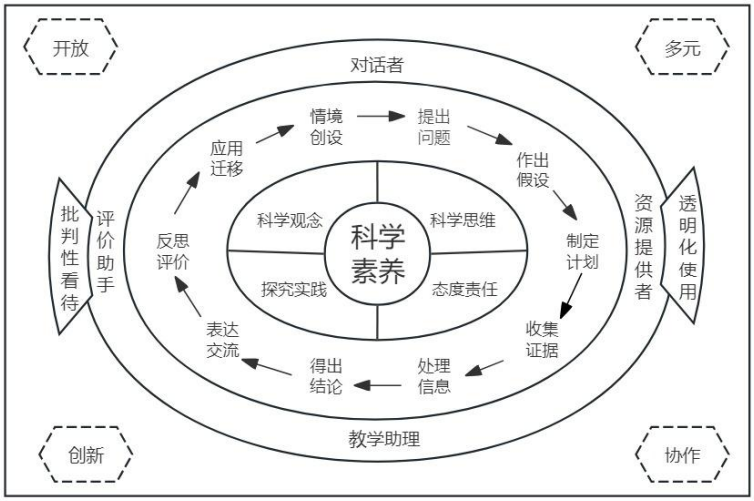


图 2 基于生成式人工智能的小学科学探究实践目标与过程模型

6.1. 小学科学的最终目标是培养学生的科学素养

科学素养是小学科学教育的出发点和归宿，小学科学课程是一门以培养学生科学素养为宗旨的科学启蒙课程。新课标当中，将基础教育阶段学生要培养的学科核心素养分解为“科学观念、科学思维、探究实践、态度责任”四个部分，这其实是对科学素养培养如何在课堂上落实进行了细化。在小学科学教育阶段，学生应该要通过学习科学知识形成初步的科学观念，通过比较与分析、抽象与概括、归纳与演绎等形成初步的科学思维，通过像科学家一样亲历科学探究的一般过程形成初步的探究实践能力，保持对自然现象的好奇心和探究热情，具备正确的价值观和社会责任感（中华人民共和国教育部，2022）。

6.2. 探究实践是支持学生科学素养培养的有效途径

基于小学科学学科属性解析，探究实践无疑是支持学生科学素养培养的有力手段。科学探究实践的方式和过程是多元化的，没有固定的模式，本研究通过对已有研究进行比较分析，从相关性和完整度进行考量最终遵循了胡卫平教授提出的“情景创设—提出问题—作出假设—制定计划—收集证据—处理信息—得出结论—表达交流—反思评价—应用迁移”的科学探究基本过程进行搭建（胡卫平等人，2021），教师需要注意在实际的科学探究实践课堂当中应当根据需要进行调整，并不要求学生必须亲历每一个环节。

6.3. 生成式人工智能是支持学生开展探究实践的强有力工具

生成式人工智能作为当前智能技术发展的前沿领域，将其引入小学科学课堂，为学生的探究实践提供了强有力的支持，可以尝试解决当前小学科学课堂当中确定探究问题困难、资源配置不足、课堂互动缺乏、师生探究水平欠佳、教学评估相对单一等一系列的问题。在人机协同的课堂模式中，生成式人工智能同时充当“智能伙伴”和“智能导师”的角色，不但能够帮助教师监控和管理课堂而且能够帮助学生调控和改进自己的探究实践学习，能够有效提升师生探究水平。它凭借着其强大的语义理解和内容生成能力，能够为学生提供丰富的实践材料，帮助学生提出科学的探究问题，同时弥补探究课堂资源配置不足的问题。其次，团队协作作为科学探究实践的核心机制，“互动”是高效推进探究实践活动的关键要素，生成式人工智能的引入不仅可以弥补传统课堂探究中学生互动不足的问题还可以拓展互动的渠道和形式，基于生成式人工智能的多向互动将有效提高探究实践的质量。此外，评价作为课堂“教—学—评”一体化里面的重要环节，生成式人工智能不仅可以帮助教师生成、修正评价量表和测试题，而且可以作为“机师”参与评价丰富评价的形式，同时也可以为学生生成个性化的评价反馈报告，帮助学生监控和调节自己的学习。生成式人工智能进入小学科学课堂当中，通过促进小学科学探究实践的高效开展，将有效促进小学生科学素养的全面提升。

6.4. 开放、多元、创新、协作的探究环境是生成式人工智能介入课堂后的预期效果

生成式人工智能介入小学科学课堂之后，营造的应该是一个开放、多元、创新、协作的探究环境。开放性体现的是资源的开放与共享，学生可以借助生成式人工智能自由的获取资源。多元性是指其互动的包容和自由，鼓励学生提出多样化的观点，尊重每一个学生的个性差异。生成式人工智能鼓励并支持学生新的思维方式，通过启发引导、提供支架等形式激发学生的灵感，帮助学生提出创新性的观点。此外，生成式人工智能介入到小组探究中，能够帮助学生进行科学的任务分配、协调任务进度，促进学生的协作学习。在这样的课堂环境当中，教师更多的是承担一个引导者和监控者的角色，有两条主要的原则，一是要让学生批判性的看待生成式人工智能生成的内容，即要明晰生成式人工智能生成的内容并不都是正确的；二是要让学生透明化的使用生成式人工智能，即在探究实践中哪个环节使用了生成式人工智能需要记录和说明。

7. 结语

生成式人工智能引发的教育变革已经开始，本研究构思了生成式人工智能在小学科学探究实践教学中的应用路径，将生成式人工智能在课堂中的角色定位成“对话者、资源提供者、教学助理和评价助手”，并从教学目标，教学过程，技术支持和教学环境四个维度构建出基于生成式人工智能的小学科学探究实践目标与过程模型，为生成式人工智能在基础教育中的应用提供借鉴。后续，研究团队将继续研究生成式人工智能在小学科学课堂中的应用，并结合实际的课堂实践对模型进行改进和修正，力求构建一个生成式人工智能在小学科学探究实践课堂中各个环节的功能和角色模型，并对教师在人机协同视域下的功能角色定位做明晰。

参考文献

中华人民共和国教育部.(2022).义务教育科学课程标准（2022 年版）.北京师范大学出版社.
中华人民共和国国务院.(2021).国务院关于印发全民科学素质行动规划纲要(2021—2035 年）的通知.中华人民共和国国务院公报.
中华人民共和国教育部.(2017).教育部关于印发《义务教育小学科学课程标准》的通知.今日教育.
宋萑,林敏.(2023).ChatGPT/生成式人工智能时代下教师的工作变革：机遇、挑战与应对.华东师范大学学报(教育科学版)(07),78-90.
陈明选,陈金艳.(2020).基于互联网思维的小学科学探究活动设计.中国电化教育(10),97-105.
张红霞,郁波.(2023).从“探究”到“实践”:科学教育的国际转向与本土应对.教育研究(7),66-80.
应向东.(2006).“科学探究”教学的哲学思考.课程·教材·教法(5),64-68.
郑永和,杨宣洋,彭禹,王佳宁.(2023).深化基础教育课程教学改革⑤提升中小学生科学素养的内涵要义与实践路径.人民教育(20),61-65.
杨宗凯,王俊,吴砥等.(2023).ChatGPT/生成式人工智能对教育的影响探析及应对策略.华东师范大学学报(教育科学版)(07),26-35.
杨晓梦.(2023).新课标视域下中小学科学教育的发展方向与推进路径.中小学管理(6),30-33.
林兆星.(2020).科学教学如何实现从科学探究到科学实践的转型.课程·教材·教法(3),65-69.
苗逢春.(2023).生成式人工智能技术原理及其教育适用性考证.现代教育技术(11),5-18.
郝琦蕾,万雅馨,杨小微,(2023).国内外科学素养研究的进展与反思.中国教育科学（中英文）(6),122-137.
胡卫平,郭习佩,季鑫,等.(2021).思维型科学探究教学的理论建构.课程.教材.教法(06),123-129.
钱文丹,刘春荣.(2023).芬兰激发小学生科学兴趣的重要经验.人民教育(8),72-76.
Bower,M.(2019).Technology-mediated learning theory. British Journal of Educational Technology(3),1035–1048.https://doi.org/10.1111/bjet.12771.
BouJaoude,S.(2002).Balance of science literacy themes in science curricula： The case of Lebanon[J].International Journal of Science Education（24）,139-156.
Garfield,E.(1988).What is science literacy and why is it important .Current comments(11),251-257.
Haney,R.E.(1964).The development of scientific attitudes.The Science Teacher（12）,33-35.
OECD.(2015).PISA 2015 Draft Science Framework.OECD Publishing.
OECD.(2023).PISA 2025 Draft Science Framework.OECD Publishing.
Hsiao-Ching She& Huann-shyang Lin&Li-Yu Huang.(2019).Reflections on and Implications of the Performance of Taiwanese Students in the Programme for International Student Assessment 2015 (PISA 2015): The Role of Epistemic Beliefs about Science in Scientific Literacy.Journal of Research in Science Teaching(10),1312.

智慧学习环境中指向创新素养提升的协作探究活动设计与实践研究

——以《莓农匠心大比拼》课程为例

Design and Practice of Collaborative Inquiry Activities in Smart Learning Environments for Promoting Innovation Competencies

——Take the Course "Strawberry Growing Competition" for Example

李梦兴[1]，李琳[1]，夏伟婕[2]，刘妍[1*]
[1]上海交通大学教育学院
[2]张江高科实验小学
[*] flyliuyan0707@sjtu.edu.cn

【摘要】 科学教育是培养学生创新素养的主阵地。在科学教育中鼓励学生协作学习，能有效调动学生积极性，促进其主动探究。新兴智能技术赋能学习环境，推动学习方式变革，如何让其更好融入协作探究以实现学生的创新素养提升是亟待解决的问题。基于此，文章构建了基于智能技术活动空间的协作探究活动模式，并在课堂实践中验证了该模式的应用效果。实践表明，基于智能技术的活动空间能更好融合真实、虚拟环境，拓展学生思维，提升学生学习活跃度。在该空间中进行协作学习，能有效促进学生主动探究，在实践中提升其创新素养。
【关键词】 智慧学习环境；智能技术；创新素养；协作探究

Abstract: *Science education is the main ground for cultivating students' innovation competencies. Encouraging collaborative learning in science education can effectively mobilize students' enthusiasm and promote their active inquiry. How to make intelligent technology better integrate collaborative inquiry to realize the improvement of students' innovative literacy has been widely concerned in the field of education. The article constructs a collaborative inquiry activity model based on intelligent technology activity space, and verifies the effect in classroom practice. Practice shows that the activity space based on intelligent technology can better integrate the real and virtual environments, expand students' thinking, and stimulate students' learning activity; collaborative inquiry in this space can effectively promote students' active inquiry and improve their innovation competencies in practice.*
Keywords: smart learning environment. intelligent technology, innovation competencies, collaborative inquiry
基金项目：本文系中国教育学会 2022 年度教育科研重点规划课题“核心素养导向的跨学科创新课程设计与诊断研究”（课题编号：202200301401A）的阶段性研究成果。

1. 引言

党的二十大报告指出，要“全面提高人才自主培养质量，着力造就拔尖创新人才”。2023 年 2 月，习近平总书记在中共中央政治局第三次集体学习时进一步强调，要“做好科学教育加法，激发青少年好奇心、想象力、探求欲，培育具备科学家潜质、愿意献身科学研究事业的青少年群体。”科学教育是科技创新后备人才培养的主阵地，以高质量科学教育促进青少年创新素养提升刻不容缓。当前，以信息技术、人工智能为代表的智能技术快速发展，为科学课程提供多样化的学习环境支持，不断改变着科学教育的组织形式。VR、AR、生成式人工智

能等数字技术进课堂，不仅有效提升了学生思维活跃性，更为其主动探究提供了更大可能性（冯婷婷等,2024）。但是，技术的盲目推广应用也导致学生出现了低效、无趣等学习问题与孤独、焦虑等情感困境（高奇琦,2023;赵鑫等,2022）。因此，有必要改变当前的技术支持模式，探索更能激发学生学习兴趣、促进其主动探究的新方式。

协作探究式学习作为一种有效的学习模式，鼓励学习者在集体学习环境中通过观察、实验、思考、讨论等方法主动探究、协商反思，实现知识建构与认知水平提升(Vandenberg et al.,2021)。借助智能技术，学生能在已有平台中与同伴互动、与教师交流，在集体环境中实现从学习孤岛到学习群岛的转变，从个体探究到组内、组间协同探究的转变，从而不断激发学生学习兴趣，促进其持续实践。然而，当前协作学习仍面临缺少真实探究环境支持、技术融入协作探究形式固化等问题（王洋,沈书生,2023）。如何真正用好技术服务协作探究成为亟待解决的问题。

技术不应局限于平台搭建，协同探究不应受制于技术而被动发展。新时代的科学教育更应借助智能技术构建人机交互的智慧空间，让学生在智慧空间中与伙伴协作、与机器协同，于主动探究中实现知识浸润、能力进步、素养提升。因此，本研究尝试设计依托于智能技术的协作探究活动模式，并在课程实践中检验该模式的实际效果，助力学生创新素养提升。

2. 文献综述

实践创新是《中国学生发展核心素养总体框架》中六项素养之一，指学生在日常生活中遇到问题或者面对挑战时表现出的实践能力和创新意识（林崇德,2016）。有学者将创新素养的关键指标概括为问题解决、劳动意识和技术运用三个方面（黄四林等,2021）。在基础教育领域，中小学生的创新素养包含创新人格、创新思维和创新实践三个要素，培养创新素养应该从创新实践入手，全面融合创新人格的塑造与创新思维的训练（甘秋玲,2020）。

教学实践中，有学者提出思维型课堂有助于培养青少年儿童的创造力，其中包括自主探究与合作交流两个重要环节（胡卫平,2016）。合作交流基于协作学习，在这一策略下，教师将学生分成不同的学习小组，组内成员之间通过相互帮助、相互合作、资源共享获取知识和技能，达到教学目标（黄荣怀,2001）。研究表明，协作学习能够对学生的学业、社会关系、学习动机以及个人发展有积极的影响（Roseth et.al,2008）。有研究者探究了小学机器人课堂的协作学习活动设计与实践，围绕协作学习的五个基本要素搭建了机器人课堂协作学习活动框架并进行行动研究，发现协作学习活动实践后，学生在积极互赖、个人与小组责任、促进性互动、合作技能与小组评价五方面均有正向提升，体现出协作学习在创新素养培育方面的促进作用（缪茜,2020）。协作学习模式培养创新素养的实施始于良好的学习活动设计，包括学习任务的真实性、整体性、协作互助性与整合创新性。在此基础上创设任务情境，分组开展协作活动，并辅以智能技术的支持（肖兵,2013）。

以协作学习的形式培养学生的创新素养常见于科学教育中。科学教育是以自然科学内容为主的，传授科学知识体系、科学研究过程和方法，培养科学认知方式的教育教学活动（郑永和等,2023）。狭义的科学教育是指在中小学阶段通过物理、化学、生物等自然学科课程渠道，关于培养科学知识、思想、方法和精神的整体性教育（裴新宁,2022）。以探索创新思维、能力、人才的培育方法、策略与路径为目标，科学教育的研究通常从课程内容、教学方式、教学资源展开（胡勇等,2023）。义务教育阶段，课程内容作为科学教育的重要载体，在知识维度上体现整合与拓展的形态，实践中常常表现为跨学科教学（刘新阳,2023）。有学者将小学科学课程加入跨学科内容，以学科的核心概念的交叉点为抓手，将跨学科融入教学目标、教学支架、教学内容、教学活动与教学评价，提高了学生的创新思维能力与学生的课堂参与度（余康妮,2024）。学会思考是学生创新素养培育的重要目标，基于此，科学教育的实施关注学生独立

思考、自主知识建构与探究问题的能力（魏锐,2020）。因此科学教育中常见项目式教学方式，以问题为导向，激励学生运用高阶思维解决问题，并且通过与他人合作的学习方式解决问题（侯肖等,2016）。站在技术赋能的视角，人工智能和虚拟现实技术的不断进步，智能平台的搭建、智慧资源的共享以及数据驱动的教学评价与反馈得以实现，科学教育服务的供给得到了保障（王阿习,2023）。有学者依据智能课堂的理论模型研究，提出了小学科学智慧课堂动态生成过程，智能技术贯穿教学设计、教学活动和教学评价（杨玉佩,2019）。智能手机、平板等智能终端赋能科学课堂，能有效实现教学资源快速获取、师生课堂双向互动以及学生数据实时分析等功能，同时突破时间与空间的限制，打造混合式教学课堂，有利于实现学生创新素养的全面提升。

综上所述，协作学习是以学生为主体的学习活动方式，在与智能技术结合下，能够为学生提供方便舒适的学习环境和个性化服务，促进学生的自主知识建构与能力提升。将协作学习置于科学教育这一具体的课堂中，不失为一种培养学生创新素养的路径。本研究在智慧环境学习背景下，聚焦科学课堂探究协作学习方式对学生创新素养的培育。

3. 模式构建与活动设计

本研究提出的基于智能技术活动空间的协作探究活动模式以协作知识建构理论为指导，依托于智能技术搭建的活动空间实现人人协作、人机协同，强调在人机互动中实现知识的主动学习与拓展，在人人协作中实现知识的综合与升级，最终于协作知识建构中提升创新素养。

3.1. 理论基础

3.1.1. 群体动力理论

群体动力理论最早由 Kurt Levin 提出，认为个体之间、个体与群体之间以及群体与外部环境之间相互依赖、互相影响，从而导致个体的行为表现与其独处时不尽相同（Lewin,1936）。为确保个体在具有错综复杂关系的群体中保持生命力并实现可持续性发展，良好的群体内聚性是关键。而群体内聚性受群体凝聚力、群体驱动力、群体耗散力三大动力因素综合影响，需从群体领导者、个体成员、群体目标、群体结构和共同环境等方面进行考虑（舒杭,王帆,2016; 徐晶晶等,2022）。基于此，本研究构建了协作探究活动模式的基础框架。

3.1.2. ACT-R 认知理论

Anderson（2000）提出 ACT-R 认知理论（Adaptive Control of Thought-Rational），将知识分为陈述性知识和程序性知识：前者指向“是什么”，常表现为个体知道并可以表达的内容；后者指向“如何做”，通常依赖具体情境，经由特定问题解决条件来获取，并进一步形成认知技能。该理论强调，陈述性知识与程序性知识能促进个体认知能力的提升，学习即是通过问题解决来实现二者的获得与迁移。具体而言，首先，陈述性知识进入记忆系统并不断积累，为程序性知识生成做铺垫；其次，学习者以问题解决为导向构建具有不同层次子目标的等级目标结构，伴随新知产生不断激活其中的子目标，促进知识程序化；随后，学习者基于前人经验与常见方法初步探索问题解决方案；最后，学习者将所学知识模块化，综合问题情景优化解决方案，生成解决特定问题的程序性知识，从而提高问题解决速度，实现能力提升（辛自强,林崇德,2002）。本研究将以ACT-R认知理论为依据，开展协作探究活动模式的流程设计。

3.1.3. 协作知识建构理论

Scardamalia 与 Bereiter 在 20 世纪 80 年代末提出协作知识建构理论，强调知识建构需要交互、学生观点和社区公共知识的生成性发展三个核心要素（万海鹏等,2021）。该理论认为社会、认知等多方面因素共同影响知识构建过程，在此过程中学习者主要行为包括分享、比较、讨论、检验以及新知识构建等阶段(Gunawardena et al.,1997)。本研究将围绕该理论中的三大核心要素，丰富协作探究活动模式的内涵。

3.2. 模式构建：基于智能技术活动空间的协作探究活动模式

智慧学习环境中指向创新素养提升的协作探究活动不只是简单的知识传递，更是学习者应用数字技术自主探索并发现问题、组内合作与组间探讨逐步实现知识的深层加工的过程。由此，本研究设计了基于智能技术活动空间的协作探究活动模式，如图 1 所示。该模型以智能技术构建的活动空间为基础，通过自主探究与组内外协作探究融合的方式实现能力的逐级提升。下面将以能力发展为主线对该模式进行说明：在初始阶段，学习者明确学习任务，借助 VR、AR、MR 等交互技术进行沉浸式自主探索，不断吸收陈述性基础知识；同时，在教师的引导下加入协作小组，为后续实践探究做准备。在具体实施阶段，学习者以小组形式提出问题并将其进行拆解，以此为导向在智能学习地图中通过点触、滑动等方式进行组内探究，实现新知学习；同时，学习者在实验室中寻找证据，强化新知识的理解与记忆。因各小组操作逻辑、点击范围不尽相同，智能学习地图产生的新知具有一定差异性，有利于小组成员的个性化探索。在能力深化阶段，学习者能在线上社区实时讨论，与组内成员、组件成员共享所学，或提出问题共同解决，并结合真实环境中的探究实验实现各项子问题的解决，实现新知的迁移应用。在能力提升阶段，学习者将开展组间分享讨论，通过整合观点形成共同理解。在该过程中，数字技术可借助捕捉到的过程性数据及时反馈，评价学习者实际学习效果。在实践过程中，各环节可以迭代提升，促进学生的高阶知识建构与创新素养提升。



图 1 基于智能技术活动空间的协作探究活动模式

3.3. 活动设计：《莓农匠心大比拼》项目

为验证上述协作探究活动模式的有效性，研究设计了以科学课程为核心的跨学科项目《莓农匠心大比拼》，如图 2 所示。该项目整体时长为 4 个月，分为“协作准备：草莓知识知多少”、“组内协作探究：草莓糖分检测”、“组内协作探究：草莓比拼种植试验”、“组内协作探究和组间分享：草莓展示会”、“组间分享：总结与优化”五个阶段。在项目过程中，学生可在由 VR、虚拟实验平台、在线社区等技术共同构建的智慧学习环境中实现沉浸式自主探究，借助组内协作活动和组间分享活动促进知识的深入理解与迁移应用，进而提升创新素养。

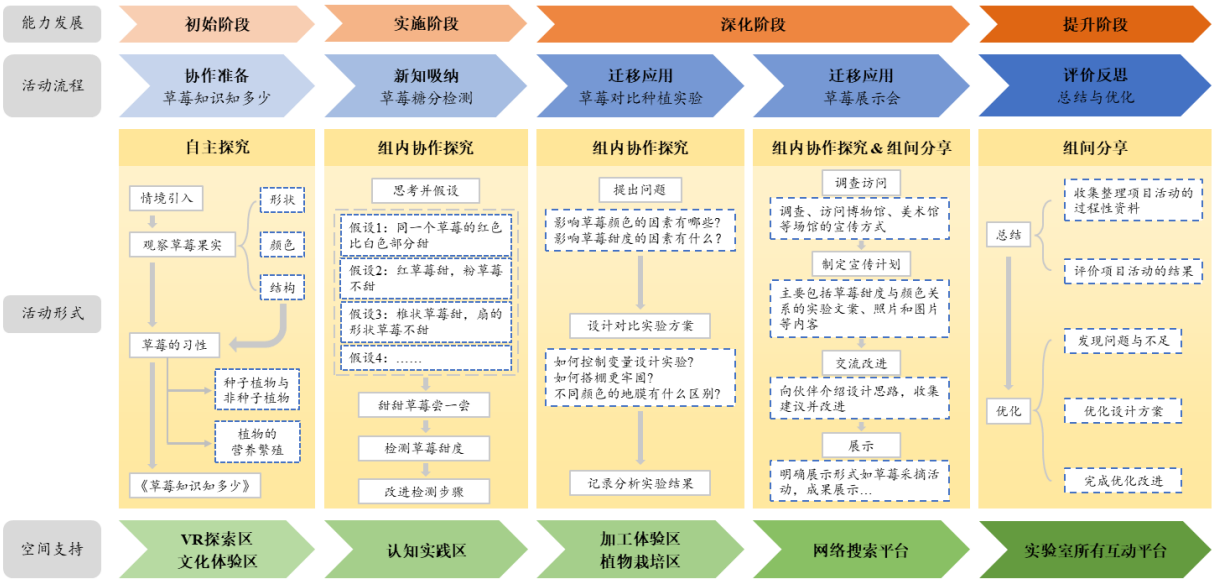


图 2《莓农匠心大比拼》活动设计

4. 应用实践与总结反思

4.1. 实践过程

本协作探究活动在上海市某小学五年级开展，共有 31 名学生参与此次项目，如图 3、图 4 所示。学生通过 VR 探索区、文化体验区、认知实践区、加工体验区、植物栽培区五大场景的实践完成探究。在第一阶段“草莓知识知多少”中，学生在 VR 探索区走进教师提前创设的情境中，通过观察草莓果实了解草莓的习性；同时，在文化体验区了解草莓的前世今生，增强对其的全方位认识。第二阶段“草莓糖分检测”主要在认知实践区完成，学生将在小组内头脑风暴，针对草莓糖分影响因素提出假设，并在虚拟实验平台中完成虚拟实验以获得基础认知，随后在实验室中检测糖分并通过组内探究改进检测步骤。第三阶段“草莓对比种植实验”在加工体验区和植物栽培区完成，学生需要在在线社区中分享已有知识与发现，并在学习别组知识后针对影响草莓颜色、甜度的因素与小组成员共同设计对比实验方案，随后完成实验并记录分析实验结果。在第四阶段“草莓展示会”中，学生将借助网络调研可行的宣传方式，并通过组内思维碰撞的方式确定最优的成果展示计划，完成小组成果展示，此过程主要在普通教室完成。在第五阶段“总结与优化”中，学生可以在实验室中借助智能技术获取过程性学习数据，通过组间分享讨论的方式互相学习，共同进步，并在汲取他人经验后反思自我，发现不足并通过优化设计方案、升级实验等方式完成优化改进。



图 3 实践过程（1）

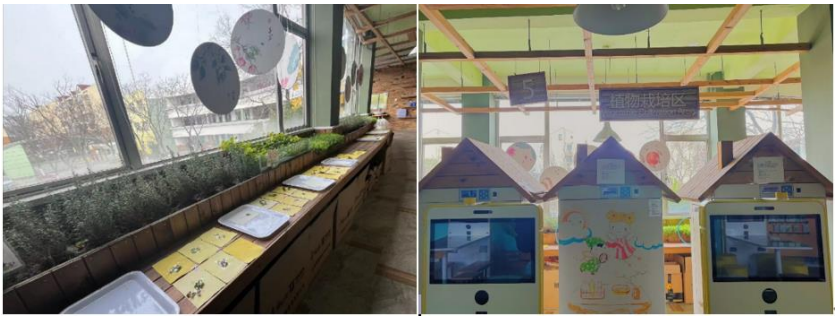


图 4 实践过程（2）

4.2. 评价分析

为验证模式有效性，本研究采取量化、质性相结合的方式进行综合评估，判断学生创新素养提升情况。

在量化评估方面，研究设计了创新素养评估问卷，该问卷改编自甘秋玲等人（2020）的创新素养模型，从创新思维、创新实践、创新人格三方面评价学生的创新素养。其中，创新思维维度包含发散、重组等创新观点体现和质疑、批判、检验等创新行为表现，创新实践维度包括观察、实验、测量等实践方法和提出问题、作出假设、制定计划等实践行为表征，创新人格维度包括好奇、开放、主动等特征。通过整体活动开展前与结束后两次测量，评估学生的创新素养变化情况，结果如图 5 的雷达图所示。雷达图显示，依据基于智能技术活动空间的协作探究活动模式设计的协作探究课程能有效提升学生的创新素养，主要表现为创新实践和创新思维方面的提升。在活动过程中，VR、虚拟实验平台等技术能有效激发学生好奇心，鼓励其以问题解决为导向进行主动探索。如在活动第二阶段，学生通过虚拟现实技术了解草莓习性后，能更好对项目核心问题进行定义与发散思考，进而敢于挑战传统，提出草莓糖分影响因素的各种假设。

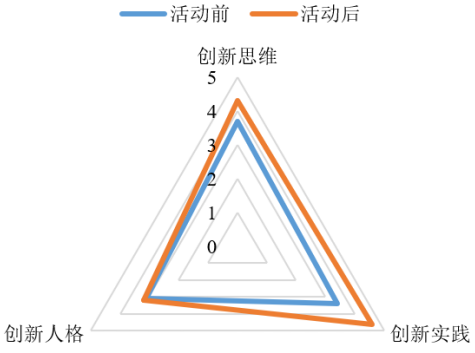


图 5 创新素养量化评估

在质性评估方面，研究根据创新素养模型制定了小组学习单支持的评价标准，从学习单的个性化设计水平、内容创新性、思考深入度三方面对学生的创新素养进行评估。其中，个性化设计水平指在设计学习单时融入个人想法和创意的程度，例如“学生能通过自己的语言和表达方式让学习单内容具有小组风格和特色”、“学生能在小组学习单中展现多样化的构图和色彩搭配”；内容创新性包含方案制定新颖程度、跨学科综合程度两个板块，例如“学生能通过小组协作提出解决问题的独特方案”、“学生能综合多学科知识构建实现方案的优秀方法”；思考深入度包含问题探索和自我反思两个板块，例如“学生能在学习单中提出更具探索性的问题”、“学生能在学习单中识别所遇学习困难并提出改进策略”。随后，研究收集了活动各个阶段结束时的小组学习单（如图 6、图 7 所示），并根据该评价标准进行分析。多样化的学习单表明，通过协作探究，各小组积极进行了所学知识综合以提出个性化的实验方案，并在后续实验过程中不断迭代优化，最后生成了各具特色的总结反思报告。在此过程中，学生展现出了较好的学习单个性化设计水平、一定的内容创新性和思考深入度，其创新实践能力得到

了有效提升。

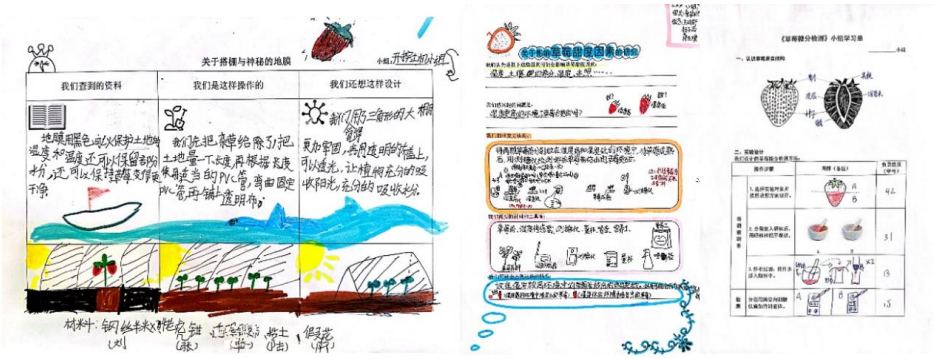


图 6 部分学习单展示（1）

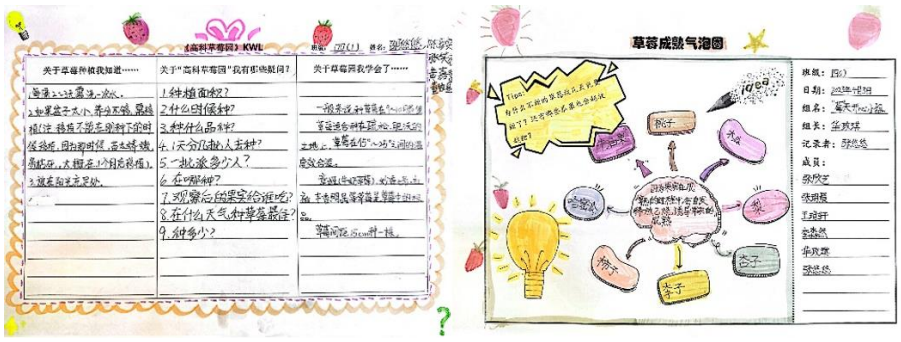


图 7 部分学习单展示（2）

4.3. 总结反思

实践表明，基于智能技术的活动空间能更好融合真实、虚拟环境，拓展学生思维，激发学生学习活跃度。在该空间中进行协作学习能有效促进学生主动探究，在实践中提升其创新素养。未来，研究还将继续优化基于智能技术活动空间的协作探究活动模式，让智能技术更好服务协作探究，助力创新素养的有效提升。

参考文献

冯婷婷、刘德建、黄璐璐、曹培杰和曾海军(2024)。数字教育：应用、共享、创新——2024 世界数字教育大会综述。**中国电化教育(03)**,20-36。

甘秋玲、白新文、刘坚、魏锐、马利红、徐冠兴、刘妍和康翠萍(2020)。创新素养:21 世纪核心素养 5C 模型之三。**华东师范大学学报(教育科学版)(02)**,57-70。

高奇琦(2023)。大模型技术引发的教育革命与"空心人"风险：从洛克出发的思考。**学习与探索(12)**,33-40。

侯肖和胡久华(2016)。在常规课堂教学中实施项目式学习——以化学教学为例。**教育学报(04)**,39-44。

胡卫平(2016)。**中国创造力研究进展报告**。西安:陕西师范大学出版总社。

胡勇和刘丹旸(2023)。大学与中小学协同推进科学教育的实践探索：基于经合组织创新学习环境的探讨。**中国教师(07)**,39-43。

黄荣怀和刘黄玲子(2001)。协作学习的系统观。**现代教育技术(01)**,30-34+41-76。

黄四林、张叶、莫雷、张文新、李红和林崇德(2021)。核心素养框架下创新素养的关键指标。**北京师范大学学报(社会科学版)(02)**,27-36。
林崇德(2016)。中国学生发展核心素养:深入回答“立什么德、树什么人”。**人民教育(19)**,14-16。
刘新阳(2023)。学校科学教育变革的四个指向。**人民教育(07)**,25-28。
缪茜(2020)。**小学机器人课堂合作学习活动设计与实践**(硕士学位论文,陕西师范大学)。
裴新宁(2022)。重新思考科学教育的若干概念与实施途径。**中国教育学刊(10)**,19-24。
舒杭和王帆(2016)。群体动力学视角下的 MOOC 本质及其教学转变。**现代远距离教育(01)**,13-19。
万海鹏、余胜泉和王琦(2021)。连接式建构：知识建构研究的新取向。**电化教育研究(10)**,12-18+24。
王阿习(2023)。技术赋能科学教育服务供给路径与实施建议。**现代教育技术(08)**,12-18。
王洋和沈书生(2023)，面向协同知识建构的 AR 协作探究式教学模型设计与应用。**现代教育技术(02)**,110-118。
魏锐(2020)。科学教育如何促进学生“学会思考”。**科学与社会(03)**,22-31。
肖兵和黄烟波(2013)。未来课堂协作学习模式研究。**现代教育技术(12)**,20-25。
辛自强和林崇德(2002)。认知负荷与认知技能和图式获得的关系及其教学意义。**华东师范大学学报(教育科学版)(04)**,55-60+77。
徐晶晶、胡卫平和逯行(2022)。在线协同学习的群体动力理论模型、案例设计与实现策略。**中国电化教育(03)**,81-89。
杨玉佩和彭虹斌(2019)。小学科学智慧课堂动态生成过程探究。**教学与管理(14)**,49-50。
余康妮(2024)。**基于科学思维培养的小学科学跨学科教学设计与实践研究**(硕士学位论文,山东师范大学)。
赵鑫、吕寒雪和吴涛(2022)。从“情知分离”到“情知共生”：在线学习变革的情感哲学审思。**中国电化教育(12)**,53-60。
郑永和、苏洵、谢涌和余舒雯(2023)。全面落实做好科学教育加法 构建大科学教育新格局。**人民教育(19)**,12-16。
Glaser, R., & Anderson, J. R. (2000). *Implications of the ACT–R Learning Theory: No Magic Bullets*. In Advances in Instructional Psychology, Volume 5 (p. Advances in Instructional Psychology, Volume 5, 2000). United Kingdom: Taylor & Francis Group.
Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of educational computing research, 17(4)*, 397-431.
Lewin, K. (2013). *Principles of topological psychology*. New York: McGraw-Hill Book Company.
Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: the effects of cooperative, competitive, and individualistic goal structures. *Psychological bulletin, 134(2)*, 223.
Vandenberg, J., Zakaria, Z., Tsan, J., Iwanski, A., Lynch, C., Boyer, K. E., & Wiebe, E. (2021). Prompting collaborative and exploratory discourse: An epistemic network analysis study. *International Journal of Computer-Supported Collaborative Learning, 16(3)*, 339-366.

在线协作工具支持的六顶思考帽对学生自我效能感与群体效能感的影响

Effects of Online Collaborative Tool-Assisted Six Thinking Hats on Students' Self-Efficacy and Collective Efficacy

黄益品[1]，郑晓丽[1*]，潘盈盈[1]，涂芸芳[1]
[1]温州大学 教育技术学系
[*] tilly222@163.com

【摘要】 在教育领域，六顶思考帽作为一种创新的思维工具被广泛应用于提高学生的批判性思维和问题解决能力。同时，六顶思考帽也是一种有效的协作学习策略，有助于提高小组参与度和引导集体思考。为了检验六顶思考帽的有效性，本研究将其引入师范生信息化教学活动的协作学习和设计中，来探讨其对学生自我效能感和群体效能感的影响。实验结果表明，与在线协作工具支持的一般协作学习相比，在线协作工具支持的六顶思考帽能显著提升学生的群体效能感，但不能显著提升自我效能感。本研究为未来的相关研究提供了一定的实践借鉴。

【关键词】 在线协作工具；六顶思考帽；自我效能感；群体效能感

Abstract: *In the field of education, the six thinking hats have been widely used as an innovative thinking tool to enhance students' critical thinking and problem solving skills. Additionally, the six thinking hats are also an effective collaborative learning strategy that helps to increase group participation and lead to collective thinking. In order to examine the effectiveness of the six thinking hats, this study introduced them into the collaborative learning and design of informational instructional activities to explore their effects on students' self-efficacy and collective efficacy. The experimental results revealed that compared with the online collaborative tool-supported conventional collaborative learning strategy, the online collaborative tool-supported six thinking hats significantly enhanced collective efficacy, but not self-efficacy. This study provides some practical guidance for future related research.*

Keywords: Online collaborative tool, six thinking hats, self-efficacy, collective efficacy

1. 引言

在过去的几十年，教育领域一直关注协作学习(Collaborative Learning)和计算机支持的协作学习（Computer-Supported Collaborative Learning, CSCL）的研究和实践。许多研究者发现，协作学习和 CSCL 对促进学生学习有着积极的效果（Hernández-Sellés et al., 2019）。然而，也有许多研究者指出，并非所有的协作学习都是有效的(Kirschner et al., 2009)。例如，在小组协作讨论过程中，学生参与的广度不够、小组讨论的组织效率低、讨论偏离主题或不深入等，这些因素都可能影响解决问题的效率（张育桂，2009）。因此，在小组讨论过程中为学生提供有效的学习干预或支架，可以促进学生认知和思维的发展，提高协作问题解决学习的效率（蔡慧英，顾小清，2017）。

众多研究表明，将六顶思考帽融入课堂教学中是在教学过程中促进协作问题解决的一种有效方法。六顶思考帽主要强调“平行思维”，为学生提供了一个全面思考问题的思维模式，

这能帮助均衡提高团队成员的集体创造力，还能激发协作学伴的积极思考和沟通，促进对问题的批判分析，从而发现解决问题的最佳方案（Kivunja, 2015）。

从现有研究来看，六顶思考帽在师范教育中的研究甚少，其教学效果还需要更多的实证支持。因此，本研究将六顶思考帽应用到信息化教学活动的协作学习和设计中，旨在探讨六顶思考帽对学生自我效能感和群体效能感的影响。本研究就以下的问题展开了实证研究。

1. 在线协作工具支持的六顶思考帽是否比在线协作工具支持的一般协作学习更能显著提高学生的自我效能感？

2. 在线协作工具支持的六顶思考帽是否比在线协作工具支持的一般协作学习更能显著提高学生的群体效能感？

2. 文献综述

2.1. 六顶思考帽

六顶思考帽由英国著名学者爱德华·德·博诺（Edward·D·Bono）于 1985 年提出，是一种思维训练模式（De Bono, 1985）。六顶颜色的帽子（白帽、红帽、黑帽、黄帽、绿帽和蓝帽）代表六种不同的思维状态或方式，换戴帽子则象征着转变为一种新的思维状态（表 1）。六顶思考帽强调“平行思维”，即让所有参与者在同一时间思考一件事。这种思维模式避免参与者将时间浪费在互相争执上，也避免小组讨论散漫。在这个过程中，所有的参与者都能调动各项思维能力，进行全面思考，这使小组的无意义的争论变成集思广益的创造（杨秀芹，陈一凡，2017）。

六顶思考帽已经在各个领域得到广泛应用，尤其是在商业和医学领域。近年来，国内外研究者也将六顶思考帽应用于教育教学领域中，旨在提高学习者的课堂参与度，帮助学生学会团队协作思考，培养学生的创造性思维、批判性思维和问题解决能力(Göçmen & Coşkun; 2022; Kivunja, 2015; Vernon & Hocking, 2014; 刘淳和蔡兰珍, 2015; 王路路等人, 2023)。另外，Smith 和 Cook (2012)在基于问题的学习过程中让大学生使用六顶思考帽，发现学生的学习动机和学习成绩都有所提高。

Gupta 和 Bhattacharya（2015）提出可以将六顶思维帽子运用在课堂活动中，如解决问题、讨论主题、探讨备选方案、头脑风暴等。同时，研究者也提出使用六顶思考帽的两种方式，一种是偶尔使用（Occasional Use），一种是系统使用（Systematic Use）。前者意味着在对话中，某一参与者建议使用其中一顶帽子，对话继续进行，持续使用两到三分钟。这种偶尔使用帽子的方式可以使某一参与者要求一种特定的思维方式或建议转换思维方式。后者指的是按照一定的顺序一个接一个地使用帽子。系统使用帽子包括演变顺序（Evolving Sequence）和预设顺序（Preset Sequence）。通过演变顺序，引导者先引导参与者选择第一顶帽子，然后选择下一顶帽子，依此类推。而预设顺序则要求在讨论开始定好帽子的顺序，这是为了让参与者能在某个主题上进行富有成效的思考。为了提升教学效果，也可以采用最优组合使用、循环使用、嵌套使用等方式来应用六顶思考帽。例如，王路路等人（2023）提出基于六顶思考帽的 PBL 教学策略模型，入项探索环节，蓝帽；知识与能力建构环节，白帽；合作探究环节，白帽和绿帽；形成与修订成果环节，黄帽和黑帽；出项环节，蓝帽和白帽；反思环节，红帽、黄帽和黑帽。研究发现基于六顶思考帽的 PBL 教学模式比传统的教学教学模式更能显著提高小学生的问题解决能力和协作编程能力。

表 1

六顶思考帽

帽子颜色	思维模式	含义	举例
白色思考帽	中立、客观	戴上白色思考帽，人们的思考关注客观事实和数据。	我已掌握的信息有哪些。
红色思考帽	直觉、情感	戴上红色思考帽，人们可以表现自己的情绪，还可以表达直觉、感受、预感等。	我的直觉是什么。
黑色思考帽	谨慎、批判	戴上黑色思考帽，人们可以否定、怀疑、质疑，合乎逻辑地对实物进行批判，尽情发表反面地意见，找出逻辑上的错误。	我对此事的否定意见是什么。
黄色思考帽	价值、肯定	戴上黄色思考帽，人们从正面思考问题，表达乐观的、积极的、建设的观点。	我对此事的正面看法是什么。
绿色思考帽	创意、求异	绿色思考帽寓意创造力和想象力。它代表创造性思维、头脑风暴、求异思维等。	除此之外，还有什么。
蓝色思考帽	管理、控制	蓝色思考帽负责控制各种思考帽的使用顺序，它规划和管理整个思考过程，并负责做出结论。	我的结论是什么。

2.2. 自我效能感和群体效能感

自我效能感（Self-efficacy）最早由心理学家班杜拉（Bandura, 1993）提出，指的是人们对自己实现指定目标或取得成果能力的自信程度。众多研究发现，自我效能感对学业成绩有非常重要的影响（Bandura，1996）。Vygotsky(1978）指出，知识是在社会环境中通过与他人的互动过程建构起来的。因此，群体效能感（Collective efficacy）是协作学习活动中需要考虑的一个重要因素。根据 Bandura（1997）的定义，群体效能感是指一个群体对其实现既定目标的能力的共同信念。换句话说，群体效能感关注的是群体作为一个整体的表现能力。许多研究证实，群体效能感与协作绩效有着重要关系（Chen & Hwang, 2019）。

对于师范生而言，学会将信息技术与教学策略相结合以此来设计有效的课堂活动以促进学生的学习至关重要。因此，本研究将六顶思考帽应用在师范生信息化教学活动的协作学习和设计中，来探讨其对学生的自我效能感和群体效能感的影响，为今后相关研究提供方法借鉴。

3. 在线协作工具支持的六顶思考帽设计

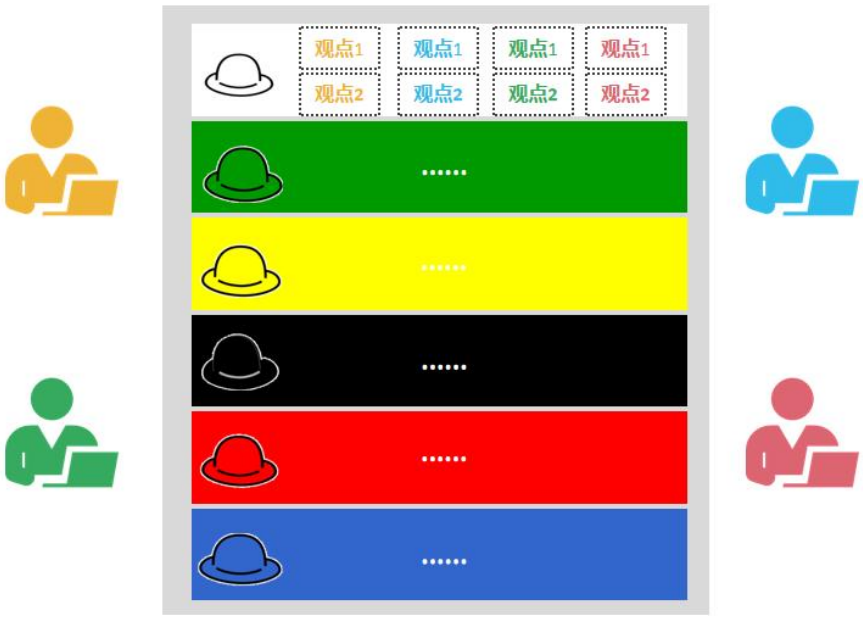
本研究中，学生通过教师的讲解和各种教学活动案例，学习当前热门的 4 个信息化教学策略（如计算机支持的协作学习、概念图、游戏化学习和故事叙述策略）。在深入学习这些教学策略后，3 至 4 人构成一个小组，利用在线协作工具支持的六顶思考帽进行小组讨论并记录关键点。最后，小组协作撰写教学活动方案。

本研究借助腾讯文档在线协作工具，将文档预先设置为六个色块，每个色块代表一种帽子。在小组讨论环节，小组成员将每个人观点记录在相应的色块内。换句话说，学习者每次只戴一顶帽子,在几分钟内将思维引向特定的重点领域,然后在腾讯文档上做笔记。依据 Gupta

和 Bhattacharya（2015）建议的顺序，见图 1，实验组的学生使用帽子的顺序为白帽（获取信息）、绿帽（提出想法和建议）、黄帽（评估每个备选方案的优点）、黑帽（评估每个备选方案的缺点）、红帽（评估此时的感受）和蓝帽（总结陈述、做出决策）。

图 1

六顶思考帽使用顺序示意图



4. 研究方法

4.1. 研究对象

本研究的参与对象为中国某高校修读《现代教育技术应用》课程的师范生 71 名。实验组共有 34 名学生，使用在线协作工具支持的六顶思考帽。对照组共有 37 学生，学习活动中有在线协作工具的支持，但不使用六顶思考帽。3 至 4 人构成的小组进行协作学习。该门课程由同一位教师授课。

4.2. 实验流程

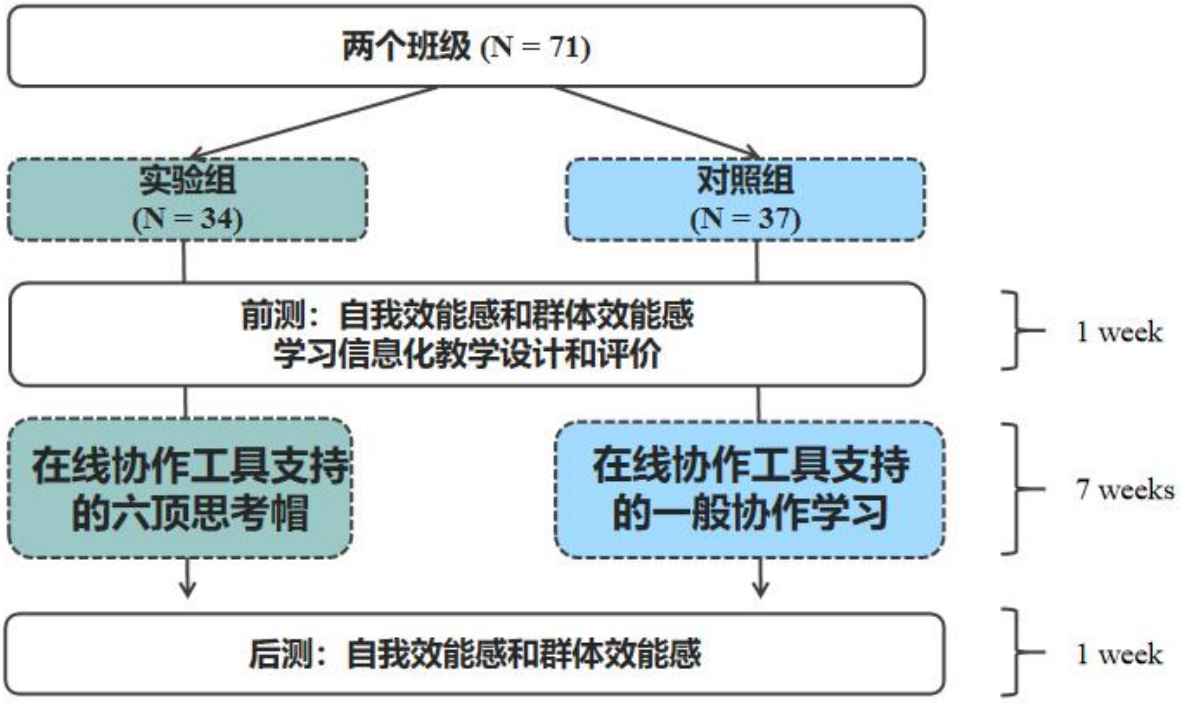
在正式教学实验前一周，两组学生完成自我效能感和群体效能感的问卷。同时，学生学习信息化教学设计和评价，实验组学生被培训如何使用在线协作工具支持的六顶思考帽，而对照组仅被培训如何使用在线协作工具。随后的 7 周课程，学生进行 4 个信息化教学策略的学习以及小组讨论和撰写教学活动方案。该协作任务的目的是鼓励师范生能够将信息化教学策略融入自己的学科教学活动中。教学实验结束后，学生完成自我效能感和群体效能感的问卷。实验流程见图 2。

4.3. 研究工具

本研究采用的测量工具包括自我效能感和群体效能感问卷。自我效能感问卷采用的是 Pintrich et al.(1991)提出的测量工具，包括 8 个题目，例如：“我确信能精准本课程所教授的技能”。群体效能感采用的是 Wang and Lin (2017)提出来的工具，也包括 8 个题目，例如“我相信我们小组能在这门课程的学习任务中取得优异的成绩”。每题从“完全不像我”到“极为像我”分别记为 1 到 7 分。自我效能感的前后问卷信度分别为 0.95，0.97。群体效能感的前后问卷信度均为 0.97。

图 2

实验流程图



5. 研究结果

5.1. 在线协作工具支持的六顶思考帽对自我效能感的影响

首先，对实验组（$M = 40.53$）和对照组($M = 37.89$)的自我效能感进行了单因素方差分析，见表 2，发现两组的自我效能感水平没有显著差异（$F = 1.42, p = .24 > .05$）。因此，可知两组学生的自我效能感相当。

在确定回归同质性的假设成立时（$F = 1.35, p = .25 > .05$），实施了单因素协方差分析。如表 3 所示，实验组和对照组在自我效能感后测水平上没有显著差异($F = .32$，$p = .57 > .05$)。

表 2

自我效能感前测单因素分析

组别	*N*	*M*	*SD*	*F*	*p*
实验组	34	40.53	9.41	1.42	.24
对照组	37	37.89	9.22		

表 3

实验组和对照组的自我效能感后测整体分析

组别	*N*	*M*	*SD*	*Adjust M*	*SE*	*F*	*p*
实验组	34	41.79	9.02	40.96	1.27	.32	.574
对照组	37	39.19	9.45	39.96	1.22		

5.2. 在线协作工具支持的六顶思考帽对群体效能感的影响

为了探究六顶思考帽对群体效能感的影响，先对实验组（$M = 41.35$）和对照组($M = 39.68$)的群体效能感进行了单因素方差分析，见表 4，发现两组的群体效能感水平没有显著差异（$F = .57, p = .45 > .05$）。因此，可知两组学生的初始群体效能感水平相当。

在确定回归同质性的假设成立时（$F = 1.29$，$p = .26 > .05$），实施了单因素协方差分析。如表 5 所示，实验组和对照组在群体效能感后测水平上有显著差异（$F = 5.61$，$p = .021 < .05$, $\eta^2 = .076$）。也就是说，使用六顶思考帽的小组在群体效能感上显著高于不使用六顶思考帽的小组。

表 4

群体效能感前测单因素分析

组别	*N*	*M*	*SD*	*F*	*p*
实验组	34	41.35	8.05	.57	.45
对照组	37	39.68	10.46		

表 5

实验组和对照组的群体效能感后测整体分析

组别	*N*	*M*	*SD*	*Adjust M*	*SE*	*F*	*p*	η^2
实验组	34	45.85	8.67	45.54	1.44	5.61*	.021	.076
对照组	37	40.51	9.26	40.81	1.38			

*$p < .05$

6. 讨论

本研究以师范生为对象，探究了在线协作工具支持的六顶思考帽（实验组）和在线协作工具支持的一般协作学习（对照组）对学生自我效能感和群体效能感的影响。研究结果表明了实验组的学生在自我效能感上没有显著高于对照组的学生，而在群体效能感上显著高于对照组的学生。

实验组和对照组在自我效能感上没有显著差异，这表明了六顶思考帽并不能提升学生的自我效能感，即学生个体完成信息化教学活动设计的自信程度。这一发现意味着，六顶思考帽并不一定会带来更好的学习成果。六顶帽子的使用顺序可能会造成学习结果的不同（Gupta & Bhattacharya, 2015）。这个结果也可能是因为两组的学生在协作学习和设计信息化教学活动时，主要集中在如何协作完成学习任务上。

在线协作工具支持的六顶思考帽比在线协作工具支持的一般协作学习更能显著提高学生的群体效能感。也就是说，六顶思考帽增强了他们与同伴协作学习的信心和期望。这个研究和以往的研究结果类似。如 Chen 和 Hwang（2019）发现在课堂学习活动中，集体问题探究翻转课堂教学法比传统翻转课堂教学法更能提高学生的群体效能感。Sung et al. (2016)发现提出问题策略的移动协作学习比一般的移动协作学习更能显著提升学生的群体效能感。

7. 结论

本研究将在线协作工具支持的六顶思考帽应用在课堂教学中，并发现了这种方法能提升学生的群体效能感，这一发现在一定程度上预测了该方法可能会提高学生的团队绩效。然而，本研究也存在一定的局限性：只探讨了六顶思考帽方法中的一种固定顺序。根据先前研究，六顶思考帽的使用顺序或者使用组合具有灵活性。未来的研究还需要进一步探讨在某类学习活动中，不同的六顶思考帽使用顺序或组合的有效性。

参考文献

蔡慧英和顾小清（2017）．协作问题解决学习中干预有效性的影响因素研究．**电化教育研究，38（6）**, 103-110．https://doi.org/10.13811/j.cnki.eer.2017.06.016

刘淳和蔡兰珍(2015)．戴上帽子去思考——六顶思考帽在英语课堂中的应用．**海外英语,(18),** 66-67.

王路路、黄渊博、蒋淑琪和郑晓丽(2023).六顶思考帽对学生问题解决能力与协作编程能力的影响——以小学 Scratch 编程 PBL 教学为例.**开放学习研究**, **28(05)**, 22-30.

杨秀芹和陈一凡(2017). 基于“六顶思考帽”的讨论式教学组织与操作框架. **当代教育科学, (1)**, 80-83+87.

张育桂（2009）．六顶思考帽在讨论式教学法中的应用．**青海师专学报, 29(4)**,93-95．

Bandura, A.(1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148. https://doi.org/10.1207/s15326985ep2802_3

Bandura, A. (1996). Failures in self-regulation: energy depletion or selective disengagement?. *Psychological Inquiry, 7*(1), 20-25. https://doi.org/10.1207/s15327965pli0701_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Chen, P. Y., & Hwang, G. J. (2019). An IRS-facilitated collective issue-quest approach to enhancing students' learning achievement, self-regulation and collective efficacy in flipped classrooms. *British Journal of Educational Technology, 50*(4), 1996-2013. https://doi.org/10.1111/bjet.12690

De Bono, E. (1985). *Six thinking hats*. Penguin Books.

Göçmen, Ö., & Coşkun, H. (2022). Do De Bono's green hat and green-red combination increase creativity in brainstorming on individuals and dyads?. *Thinking Skills and Creativity, 46*, Article 101185. https://doi.org/10.1016/j.tsc.2022.101185

Gupta, T., & Bhattacharya, S. B. (2016). Fostering thinking skills through six thinking hats in education. *India JQ Refer. J. Dialogues Educ, 4*(4), 54-74.

Hernández-Sellés, N., Muñoz-Carril, P. C., & González-Sanmamed, M. (2019). Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers & Education, 138*, 1-12. https://doi.org/10.1016/j.compedu.2019.04.012

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*(2), 306-314. https://doi.org/10.1016/j.chb.2008.12.008

Kivunja, C. (2015). Using De Bono's six thinking hats model to teach critical thinking and problem solving skills essential for success in the 21st century economy. *Creative Education, 6*(3), 380-391. https://doi.org/10.4236/ce.2015.63037

Pintrich, P.R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning.

Smith, M., & Cook, K. (2012). Attendance and achievement in problem based learning: The value of scaffolding. *Interdisciplinary Journal of Problem-based Learning, 6*(1), 129-152. https://doi.org/10.7771/1541-5015.1315

Sung, H. Y., Hwang, G. J., & Chang, Y. C. (2016). Development of a mobile learning system based on a collaborative problem-posing strategy. *Interactive Learning Environments, 24*(3), 456-471. https://doi.org/10.1080/10494820.2013.867889

Vernon, D., & Hocking, I. (2014). Thinking hats and good men: Structured techniques in a problem construction task. *Thinking Skills and Creativity, 14,* 41-46. https://doi.org/10.1016/j.tsc.2014.07.001

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Harvard University Press.

Wang, S. L., & Lin, S. S. J. (2007). The effects of group composition of self-efficacy and collective efficacy on computer-supported collaborative learning. *Computers in Human Behavior, 23*(5), 2256-2268. https://doi.org/10.1016/j.chb.2006.03.005

新课标背景下生成式人工智能在小学科学教学评价中的功能与角色

Functions and Roles of GenAI in New Curriculum Standards Oriented Learning Evaluationsof Primary School Science Educations

杨宁，李晓璠，包正委
福建师范大学
qsx20230305@student.fjnu.edu.cn

【摘要】 《义务教育科学课程标准(2022 年版)》的颁布，标志着小学科学教学迈入了以学科核心素养为导向的新阶段。新课标对小学科学的教学评价提出了新要求：评价内容需整合科学核心素养要素；评价方式应注重学生的学习进阶发展过程；评价结果的反馈需要加强实时性和个性化。在此基础上，该研究深入分析小学科学新课标内涵，探究生成式人工智能在小学科学教学评价内容的设计、方式的创新、结果的反馈等方面的创新功能，明确其在教学评价中的设计辅助者、学习成效评估员和个性化助手的角色定位，实现技术赋能教育教学的高质量发展。

【关键词】 新课标；生成式人工智能；教学评价；小学科学

Abstract: *The Compulsory Education Science Curriculum Standards (2022 Edition) marks a new stage in primary school science teaching, with an emphasis on the core literacy of the subject. This new standard requires primary school teaching evaluation to: integrate scientific literacy into the evaluation content, examine students' learning progression development process, and provide real-time, personalised feedback. This study delves into the connotations of the new standard and investigates the innovative capabilities of generative AI in designing assessment content, applying assessment methods and providing feedback on assessment results, thus acting as a design tool, learning assessor and personalised assistant for assessment. This aims to facilitate high-quality primary science teaching.*

Keywords: new curriculum standards, generative artificial intelligence, evaluation of teaching and learning, primary science lessons

1. 前言

为了进一步规范和完善科学学科的教育体系，促进学生科学核心素养的形成，教育部颁布了《义务教育科学课程标准(2022 年版)》(以下简称“新课标”)。新课标明确指出构建科学核心素养导向的综合评价体系需要充分利用信息技术来提高评价的科学性、专业性和客观性(中华人民共和国教育部，2022)。应用信息技术优化教学评价已成为必然趋势。

新课标对新时代小学科学教学评价提出了新的要求：一是应紧密围绕科学课程目标以及学业质量标准，以培养学生的科学核心素养为导向，提升其探究能力、实践能力和创新思维；二是遵从多元化、过程性、发展性的原则，强调评价主体多元、方法多样、内容全面，充分发挥学校、教师、学生等多主体参与评价的积极性；三是关注个体差异，充分认识到每个学生在认知能力、兴趣特长、学习方式等方面的独特性，并在此基础上进行精准评估；四是创新应用人工智能等现代教育技术工具，设计适合项目式教学的评价方式，充分发挥评价的导向功能、诊断功能和教学改进功能。

以 ChatGPT 为代表的生成式人工智能的出现意味着人工智能技术应用的重大突破，它通

过模拟人类智能进行机器创造和逻辑推理分析，已经重塑了新时代信息资料获取的方式(朱永新和杨帆，2023)。师生能够与它进行互动对话，以此实现各项功能，并在教学评价环节中实时、动态且精准地收集与处理数据信息。本研究在深入分析小学科学新课标内涵以及其对教学评价提出的要求的基础上，试图探究生成式人工智能在小学科学教学评价中的功能和角色定位，助力更加科学化、专业化和高效能的小学科学教学评价体系的形成。

2.新课标背景下小学科学教学评价面临的挑战

新课标明确了小学科学教学中探究性学习、实践操作和创新能力培养的重要性，强调通过科学课程，学生需要明白科学知识如何获得与应用，以及科学是如何通过一系列的概念帮助我们进一步理解我们周围的世界。这对小学科学课程的教学评价提出了全新的审视，为教学评价带来了新的挑战。

2.1 评价内容需整合科学概念和探究实践能力

新课标新增了素养导向的学业质量要求，意味着教学评价内容不再仅局限于书本知识的掌握程度，还应该注重考察学生的综合素质与科学核心素养。核心素养是学生在接受相应学段教育过程中，逐步形成的适应个人终生发展和社会发展需要的必备品格与关键能力，是知识、技能、态度等多方面的结合体（林崇德，2017）。其中，科学核心素养既包含科学学科概念和跨学科概念，又包括将科学知识和技能应用于实践来解决问题等方面的能力，具有高度整合性的特征。然而，这种内涵的复杂性和整合性导致了教学评价的困境，传统的评价在内容设计上常将科学概念知识和科学实践技能分割开来，这无法满足科学核心素养的全方位评价需求，也与当前教育测评理论脱节。

因此，鉴于新课标的要求，小学科学教学评价应当紧密围绕科学核心素养的构成要素，结合信息技术工具，从科学观念、科学思维、科学探究实践、跨学科概念以及态度责任等方面开展整合设计，力求通过有限的评价次数，精准捕捉到学生在多个层面的表现和发展状况。

2.2 评价方式应注重学生在科学学习过程中的进阶发展

新课标确定了学生在义务教育阶段应该培养的一系列科学核心素养，并强调学生理解科学概念间的联系、发展科学实践能力以及深入领悟核心素养是一个不断进阶发展的过程(梁东红和张军霞，2023)。而科学核心素养又具有内隐性和抽象性的特征，学生的真实理解情况和能力发展水平需要外显后才能被准确测评。因此，传统教学评价仅聚焦于学习结果和知识掌握程度，忽视学习过程中素养水平动态演进的评价方式已经难以适应当前需求。

遵循新课标的指引，教学评价的方式应转向过程性评价，注重学习的进阶属性，揭示不完全的理解与实践过程。针对核心素养的评价则可以采用表现性评价和情境性评价方式，一方面考察学生在完成表现性任务时产出作品的质量，以此衡量其知识运用与创新思维能力；另一方面，在创设的真实科学情境中评估学生应用科学知识解决实际问题的表现(崔允漷和郭洪瑞，2021)，从而综合评判学生在科学实践过程中学习的理解情况和能力水平的进阶发展情况。

2.3 评价结果需强调实时个性化反馈对学习的改进作用

新课标指出教学评价的核心目的不应仅仅停留在精确评估学生学习情况的层面，更重要的是借助评价结果的反馈机制，帮助学生发现自身对科学概念的理解和科学实践方案设计等方面的优势与不足(肖龙海等人，2023)。这一过程旨在激励学生优化科学探究实践的方法，逐步深化对科学概念及跨学科概念的认知层次，从而实现个体成长的全面性与充分性。然而，传统教学评价往往在教学阶段结束后才将评价结果集中反馈给全体学生，存在两方面的局限：一方面是反馈时间的滞后性会导致学习问题的识别和解决滞后，降低了学习改进的效果；另一方面是反馈内容无法提供针对性强、契合度高的指导建议，忽视学生的个性化需求。

鉴于此，满足新课标要求的教学评价需关注并致力于实现评价结果反馈的提效工作，特别是在强化反馈的实时性与个性化这两个关键维度上，以充分发挥其对科学实践过程的精准改进作用。例如在小学科学教学过程中，教师需要持续密切关注学生的学习动态，分析每个学生或学习小组的学习成果，如实验报告、学习任务单和实验设计方案等，并在课堂上及时给予个性化评价反馈，帮助学生在课堂学习的过程中改进科学实践过程，提升科学学习的成效。

3. 生成式人工智能在小学科学教学评价中的创新功能

2022 年 11 月，OpenAI 发布了自然语言处理技术驱动的智能聊天机器人 ChatGPT，是一种生成式预训练语言模型，它能够接受大量人工标注的文本数据的强化学习训练，能够理解书面提示信息并自动生成连贯、自然的文本(OpenAI,2022)。在新课标的背景下，探讨人工智能在教育领域所能承担的角色及功能，以及如何为教学提供支持以应对新的挑战，已经成为值得研究的新课题。

3.1 从“人工创造”转向“交互生成”的评价内容设计

在传统的人工创造模式下，教师是教学评价设计的唯一主导者，根据科学素养培养目标、教学内容和学生特点进行设计，主要依靠教师的个人经验和专业知识，设计过程往往较为封闭和独立，设计的评价内容也经常出现知识和技能分割开的问题。尽管新课标要求将科学概念、跨学科思维和科学实践能力等科学核心素养的标准有机整合到教学评价的内容中，但这无疑对教师构成了一种新的挑战。教师不仅需要深入理解各个科学领域的核心概念，还要将跨学科的知识巧妙融入到评价内容中，并确保教学评价能反映出学生的思维和能力情况，这对教师来说实在有难度。

表 1. 生成式人工智能设计测试题功能测试

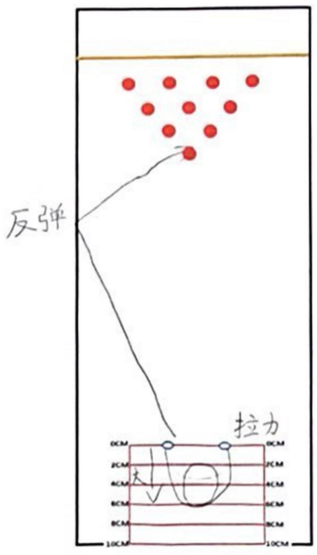
提问	你是一名小学科学教师，拥有丰富的教学经验与专业的学科知识。你有能力依据科学核心素养的目标将科学概念、科学思维与科学实践技能进行系统整合。现在五年级学生已经完成了“微生物与细胞”这一内容的学习，请你设计一份能够有效测评学生在以上几个方面水平的测试题。
设计整合性测试题	一、选择题 1.下列哪种生物不属于微生物？ A.细菌 B.真菌 C.病毒 D.蚯蚓 2.细胞是构成生命体的基本单位，以下哪个不是细胞的主要组成部分？ A.细胞膜 B.细胞核 C.液泡 D.地壳 二、填空题 3.微生物在生态系统中扮演着重要的角色，如某些微生物可以分解_____，参与物质循环。4.细胞通过_____进行生长和分裂，以增加细胞数量。 三、判断题 5.所有的微生物都是单细胞生物。（ ） 四、简答题 6.描述一下显微镜的正确使用步骤，并解释为何要按照这些步骤操作。 7.设计一个简单的实验来检测环境中的微生物数量，描述实验步骤并预测实验结果。同时，分析影响微生物数量的因素有哪些？ 这份测试题，旨在考察学生对微生物与细胞基本概念的理解，科学思维的应用（如推理、判断），以及科学实践技能的掌握（如显微镜使用、实验设计等）。

然而，生成式人工智能具有强大的智能分析和创造生成功能，为应对上述挑战提供了极具潜力的解决方案。生成式人工智能为教师提供了交互生成的教学评价设计模式。在这一模式下，评价内容的设计不仅依赖于教师的经验，还充分利用了信息技术和数据分析手段，使得设计过程更加开放、多元和高效，设计的评价内容也更具有整合性，契合新课标的要求。例如，教师可以与生成式人工智能互动生成综合测试题。目前，生成式人工智能可以通过其语料库中的内容，为测试题设计提供详细的构思与组织（贺樑等人，2023）。教师可以向它提出评价内容的设计要求，它就能自主生成包含科学概念、科学思维、实践技能等方面的测试题目，如表 1 所示，并帮助教师规划不同评价内容在测试题目中所占的权重。此外，教师还可以向生成式人工智能提交已经设计好的评价形式，让其通过自然语言处理和机器学习技术进行智能分析，指出教师设计的评价中包含的评价内容是否符合新课标要求并给出修改建议。

3.2 从“传统工具”转向“技术赋能”的评价方式创新

新课标背景下的教学评价着重关注学生在学习过程中展现的知识理解深度和能力水平的进阶发展变化，尤其重视运用过程性评价、表现性评价和情境性评价等多种评价方式来揭示学生的核心素养水平随着学习进程的持续动态变化。然而，此类评价方式针对的是非结构化形式的学习材料和行为表现情况，若教师沿用传统工具对其进行评判往往会出现评价结果不准确、操作难度较大等问题。

生成式人工智能在分析复杂形式数据的方面具有一定优势，可以为教师开展过程性评价和表现性评价提供技术工具。教师可以利用其逻辑分析能力针对不同学生的作业、论文或试卷等各类实质性学习材料进行多维度分析(卢宇等人，2023)。如图 1 所示，通过分析学生提交的学习任务单，生成他在学习风格、科学思维、科学实践能力以及团队精神等方面的情况，如其是倾向于独立思考还是合作探究、是善于理论构建还是实践操作、能否从不同角度审视问题并提出独到见解等。此外，教师还可以借助生成式人工智能的创造生成功能来创设科学实际情境，通过让学生解决情境问题、角色扮演等不同方法来测评学生的知识与能力变化(杨宗凯，2020)，进而增强教学评价的客观性、公正性和有效性。以“简单电路”这一单元内容为例，在完成教学活动之后，生成式人工智能可以创设电路问题情境，将抽象的知识点转化为生动具体的实践应用，如表 2 所示，并通过与学生开展对话交流来测评学生科学核心素养的水平，如自主学习、批判性思维、创新思维和问题解决能力等方面。



提问	这是小学二年级科学课中“推与拉”单元的学习任务单，该单元的教学目标是：学生能够使用前后左右等描述物体所处的位置、知道推力和拉力是常见的力、知道力可以使物体的形状发生改变。你是一名经验丰富的科学教师，请你对完成这份任务单的学生进行评价并给出评语。
评价结果	优点: 1. 学生能够清晰地识别出推力和拉力，并将其标记在图中。 2. 学生能够运用前后左右等方位词描述物体的位置，说明他们掌握了基本的空间认知技能。 3. 图中的线条和符号表明学生试图表达力的方向和大小，显示出一定的抽象思维能力和图形解释能力。 改进之处: 1. 尽管学生尝试了用文字和图画相结合的方式来表达概念，但整体布局显得有些杂乱无章。建议引导学生更加系统化地组织信息，比如使用表格或流程图等形式来呈现内容。 2. 在描述反作用力时，学生仅标注了箭头而没有明确指出其来源或作用对象。为了更准确地表达物理原理，鼓励他们在以后的学习中注意细节处理，确保每个元素都有对应的解释。

图 1 生成式人工智能的智能分析功能测试

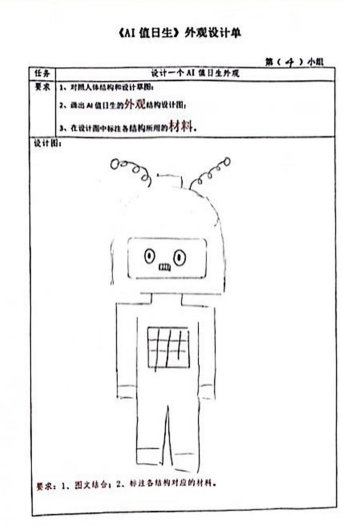
表 2. 生成式人工智能的创设情境功能测试

提问	我是一名四年级的学生，我刚学完“简单电路”这一单元，你能通过与我对话的方式创设一个情境，考察我学习的情况吗？
生成问题情境	情境：今天，你拿出最喜欢的电动玩具小汽车玩耍，按下开关后发现原本应该飞速旋转的车轮却纹丝不动。你决定当个小工程师，找出问题所在。 1. 首先，你可以告诉我，“简单电路”是由哪些基本部分组成的吗？ 2. 为了检查小汽车的问题，你会首先检查什么？（提示：电池、开关、导线和电动机在电路中的作用是什么？） 3. 如果你发现电池电量充足，开关也正常开启，那么接下来你会如何检测电路是否畅通呢？（提示：如何使用一个小灯泡或测电表进行检测？） 4. 假如电路是通的，但小汽车的电动机还是不转，这可能是什么原因？（可以从电流、电压或者电动机内部结构等方面考虑）

3.3 从“教师提供”转向“机器提供”的评价结果反馈

新课标指出小学科学教学评价应提高结果反馈的效率，着力增强评价反馈的即时性和针对性。但是，在真实的科学教学课堂中，教师往往难以做到对每位学生的全程深度关注，而要实现对每一位学生即时、个性化的反馈将会耗费大量时间，在现实中是难以实现的。

因此对于需要在课堂情境中及时反馈的教学评价可以借助生成式人工智能的支持，利用它的智能分析功能，实现针对学生技能掌握程度或知识学习水平的一对一自动判别和个性化分析反馈(苗逢春，2023)，形成覆盖每个学生的精准反馈，满足学生的个性化需求。例如在小学科学课堂中的项目设计环节，学生可以将自己的设计方案通过拍照或上传文档的形式提交给生成式人工智能，它就可以给出与设计方案匹配的针对性建议，如图 2 所示，帮助学生在课堂中完成学习的进阶迭代。此外，生成式人工智能还能连续记录并分析学生在不同阶段的评价结果，形成详尽的评价记录并反馈给教师，便于教师追踪学生的学习成长过程，精准掌握每位学生的个体化学习状况。这样一来，教学评价真正实现了从“一刀切”到“量体裁衣”的转变，确保每一位学生都能得到与其实际需求相匹配的精准反馈与指导，从而有效促进他们的全面发展和提升。



提问	我是一名五年级的学生，这是我在“人体系统”单元中画的模仿人类打扫卫生的“AI值日生”的设计图，请你给我提一些改进建议。
提出改进意见	1. 功能性：考虑增加更多的功能来帮助机器人更好地完成任务。例如，可以添加一个吸尘器或扫帚来清理地面，或者一个喷雾器来消毒表面。 2. 可操作性：确保机器人的各个部分都可以轻松地移动和操纵。你可以尝试在手臂、腿部等部位添加关节，以便于机器人能够更灵活地执行任务。 3. 美观度：考虑使用不同的颜色和图案来装饰机器人，使其更加吸引人。你也可以尝试为机器人添加一些细节，如眼睛、嘴巴等，以使它更具个性。 4. 实用性：思考如何让机器人更加实用。例如，它可以配备一个储物箱来存放清洁工具，或者有一个显示屏来显示当前的任务状态。

图 2 生成式人工智能的改进功能测试

虽然生成式人工智能可以为师生提供具体的教学评价结果，但其更多的是为评价提供技术

性支持。在与人工智能进行协作评价的过程中，生成信息的可靠性和有用性仍然需要师生依据专业资源进行改正和优化(Westphal M et al.,2023)。因此，生成式人工智能在小学科学教学评价中的角色定位应当是辅助性的。

4. 生成式人工智能在小学科学教学评价中的角色定位

教学评价的合理设计和安排是进一步促进教学的前提条件，但传统的评价方式已经不能再满足多元评价学生的需求。而面对当前科技飞速发展的时代背景，生成式人工智能有望在教育领域得到广泛应用。因此，结合新课标新要求，充分发挥人工智能在教学评价中的创新功能，明确其在小学科学教学评价中的角色定位是至关重要的。本研究基于新课标对教学评价提出的挑战，结合生成式人工智能已有的相关研究，从教学评价的设计、评价方式的创新和评价结果的反馈三个方面明确了生成式人工智能在教学评价中的角色，分别是设计辅助者、学习成效评估员和个性化助手，如图 3 所示，以探讨人工智能在小学科学教学评价中的应用可能。

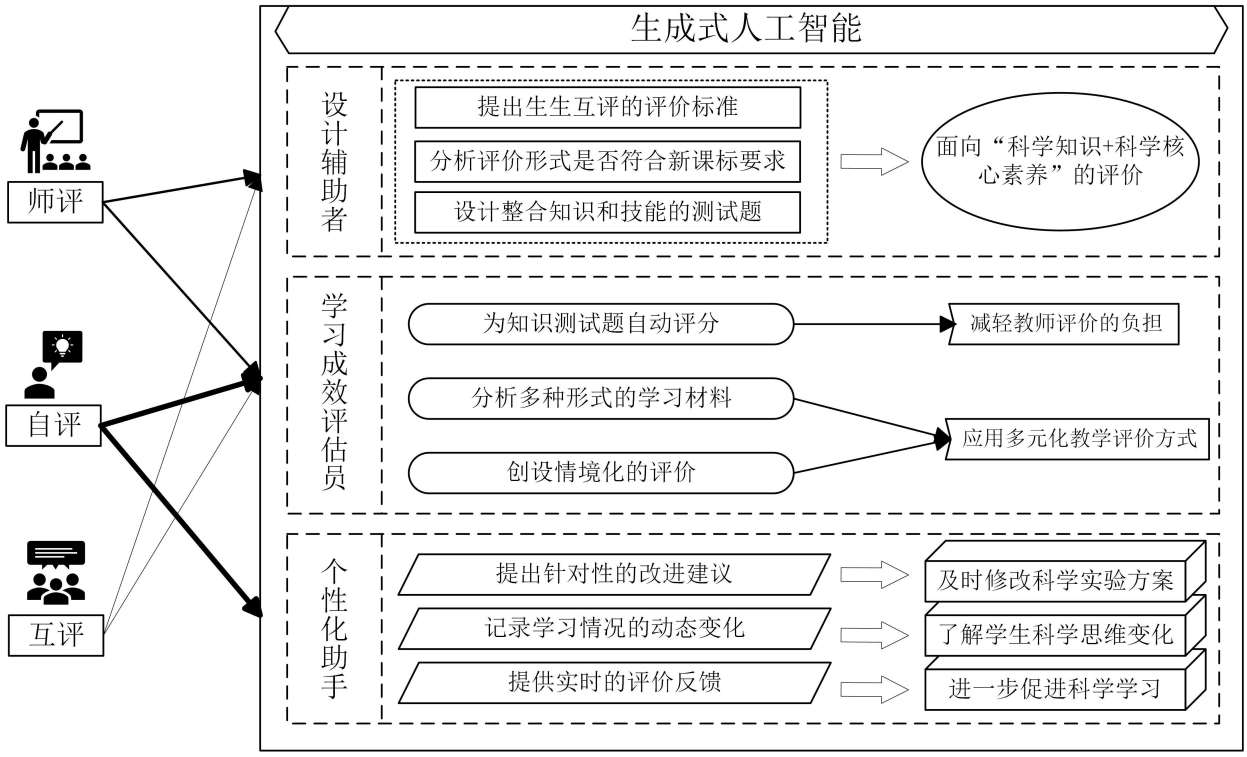


图 3 生成式人工智能在小学科学教学评价中的角色

4.1 设计辅助者：为设计教学评价提供互动性技术支持

科学的方案和操作指南既可以为行动提供有效指导，也可以确保行动取得良好效果。对教学评价而言，制定科学、有效、可行的评价方案直接影响评价工作是否有序推进、是否可以达到预期效果。生成式人工智能拥有强大的创造力，可以为设计教学评价提供丰富的想法，而教师则可以依据专业的学科知识和教学经验对这些想法进行筛选和修改，设计促进学生科学知识进步和科学核心素养提高的教学评价。

在设计评价标准的过程中，生成式人工智能可以为教师评价和生生互评两个环节提供支持。在教师评价环节，教师可以向人工智能提交教学目标及评价内容，并根据它生成的评价标准进行二次完善。在生生互评环节，学生则可以向人工智能详细描述评价的情境或提交需要评价的资料，进而询问互评的标准并经过小组或集体讨论修改后应用。在设计评价方式的过程中，生成式人工智能可以依据评价内容提供不同的支持。评价知识内容时，人工智能可以形成测试题或问卷的评价形式。评价素养内容时，它则可以创设出几种相关的评价情境。例如

评估学生在经历科学教学活动后，对“微生物的繁殖条件”这一内容的掌握程度，人工智能就可以扮演一个需要延长草莓保存时间的同学，通过对话的形式考察学生在复杂情境下做出决策和处理实际问题的能力(宋萑和林敏，2023)。

此外，师生应该按照评价的需求，而不是依据固定的评价模式来应用生成式人工智能作为教学评价的设计辅助者提供的功能。

4.2 学习成效评估员：为多元评价方式提供智能化工具辅助

生成式人工智能凭借对多种数据的高效处理以及多维度的精细评价功能，能够以学习成效评估员的身份为学生的学习效果提供全方位、多层次的深度解析与精准诊断，提高教学评价的客观性、公正性与准确性。

面对各类测试题与问卷，尤其是拥有明确答案的题目，生成式人工智能能够精确匹配预设的答案库，自动执行评分操作，助力教师实现批改与评分工作的自动化。这一方式既有效缓解了教师在评价环节的压力，又同步提升了评价过程的公正性和工作效率。当涉及到开放性答案的评判时，生成式人工智能同样大有作为。它能从思维、能力水平及知识掌握深度等多个维度，对学生的回答进行多角度的分析，输出类似于人类评分员的诊断报告(KiM S etal.,2019)，为师评、学生自评以及互评提供了具有价值的参考。而对于评价量表及其他学习数据，生成式人工智能可以将其转化为直观可视的评价结果，使得复杂的评估信息得以清晰呈现，帮助教师快速把握学习情况。

4.3 个性化助手：为实时精准的评价反馈提供数字化手段支撑

新课标强调构建一种能够满足各类学生多元化需求、激发其内在潜能、促进全面发展的个性化学习模式。因此，教学评价体系的反馈机制需由聚焦集体转向关注每个独立的学生个体。在这种背景下，生成式人工智能个性化助手的身份应运而生，它能够为每位学生精准提供个性化的学习反馈。

首先，生成式人工智能可以从改进学习的角度，基于对话数据和学习数据深度分析并优化学生的学习过程，为他们提供针对性且合理有效的指导。例如，利用生成式人工智能评价学生的课后作业等学习材料，不仅能指出学生在某个知识点理解上的模糊之处，还能深入剖析其在学习策略、解决问题的能力、高阶思维等方面的短板。其次，生成式人工智能能够记录与学生多轮交流中的文本表达和数据分析结果，为学生提供持续的针对性学习追踪服务(杨宗凯等人，2023)，挖掘学生的科学思维、科学知识、科学素养等方面在学习过程中的动态变化。最后，生成式人工智能可以根据动态的学习记录，为学生提供个性化的自评方法或标准，帮助学生识别自身的优势与瓶颈，明确后续努力的方向，同时避免了主观因素导致的评价偏差。

5. 结语

作为智能时代的突破性技术，生成式人工智能正改变教学场景中的人机交互与协同模式，能够以设计辅助者、个性化助手及学习成效评估员的角色为教学评价的转变带来机遇，有助于应对新课标背景下小学科学教学评价面临的挑战。研究团队将继续探索在小学科学的教学评价中应用生成式人工智能的实践，力图打造一个兼顾科学性、精确性、公平性的智慧化评价体系。

参考文献

杨宗凯. (2020). 利用信息技术促进教育教学评价改革创新. 人民教育(21),30-32.

林崇德. (2017). 21 世纪学生发展核心素养研究. 北京师范大学出版社.

苗逢春. (2023). 生成式人工智能技术原理及其教育适用性考证. 现代教育技术(11),5-18.

宋萑, 林敏. (2023). ChatGPT/生成式人工智能时代下教师的工作变革：机遇、挑战与应对. 华东师范大学学报(教育科学版) (07),78-90.

朱永新, 杨帆. (2023). ChatGPT/生成式人工智能与教育创新：机遇、挑战以及未来. 华东师范大学学报(教育科学版) (07),1-14.

梁东红, 张军霞. (2023). 基于学业质量的小学科学课程设计研究. 课程. 教材. 教法 (10),126-132.

崔允漷, 郭洪瑞. (2021). 试论我国学科课程标准在新课程时期的发展. 全球教育展望(09),3-14.

杨宗凯, 王俊, 吴砥, 陈旭. (2023). ChatGPT/生成式人工智能对教育的影响探析及应对策略. 华东师范大学学报(教育科学版) (07),26-35.

中华人民共和国教育部. (2022). 义务教育科学课程标准(2022 年版). 北京师范大学出版社.

卢宇, 余京蕾, 陈鹏鹤, 李沐云. (2023). 生成式人工智能的教育应用与展望——以 ChatGPT 系统为例. 中国远程教育(04),24-31+51.

肖龙海, 陆叶丰, 龚待忠. (2023). 双线融合课堂的表现性评价研究: 技术整合路径与实践模式建构. 课程. 教材. 教法(01),58-65.

贺樑, 应振宇, 王英英, 孙文琪. (2023).教育中的 ChatGPT: 教学能力诊断研究. 华东师范大学学报(教育科学版) (07),162-176.

Kim, S., Park, J., & Lee, H. (2019). Automated essay scoring using a deep learning model.Journal of Educational Technology Development and Exchange,2(1), 1-17.

OpenAl.(2022).ChatGPT:Optimizing language models for dialogue.[2022-11-30].Retrieved from https://openai.com/blog/ChatGPT/.

Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. B., & Rafaeli, A. (2023). Decision control and explanations in human-AI collaboration: Improving user perceptions andcompliance.Computers in Human Behavior,144,107714.

面向项目化学习的表现性评价设计：基本思路与具体路径

Design performance assessment for project-based learning: basic model and specific path

沈诗淼[1*]，马志强[1]

[1]江南大学 江苏“互联网+教育”基地

[*] 1594349812@qq.com

【摘要】 素养是当前我国人才培养的重要目标，项目化学习是发展素养的重要方式，而评价是推动项目化学习由知识目标转向素养目标的重要抓手。项目化学习中指向高阶认知的特性决定了评价要能全过程、全方位关注学生在情境中重构整合知识的综合表现。表现性评价作为一种指向复杂、高水平能力的过程性评价方法，既能通过过程性反馈有效引领项目推进，又能持续提供可靠证据以刻画学生素养水平变化，但目前在课堂中应用该方法还存在一些困境。本研究旨在从逆向设计的设计思路切入，通过目标-方式-任务的评价路径贯穿项目全程，形成基本模型与具体路径。

【关键词】 表现性评价；项目化学习；模型；路径

Abstract: *Literacy is an important goal of talent cultivation in China, project-based learning is an important way to develop literacy, and evaluation is an important hand to promote project-based learning from knowledge goals to literacy goals. The characteristics of project-based learning pointing to higher-order cognition determine that the evaluation should be able to pay attention to the whole process, all-round attention to the comprehensive performance of students' reconstruction and integration of knowledge in the context. As a process evaluation method pointing to complex, high-level competence, performance assessment can effectively lead project advancement through process feedback and continuously provide reliable evidence to portray changes in students' literacy levels, but there are still some dilemmas in applying the method in the classroom at present. This study aims to start from the design idea of reverse design, and form a basic model and specific path through the whole program through the evaluation path of goal-method-task.*

Keywords: performance assessment, project-based learning, model, path

1.问题提出

素养是当前我国人才培养体系的重要目标。项目化学习（“PBL”-Project Based Learning或“PL”-Project Learning）作为一种代替以讲授为主的传统教学法的教学方式，受到欧美等发达国家的青睐，也为我国落实以核心素养为导向的课堂教学提供了契机。随着我国教育部发布《义务教育课程方案和课程标准（2022 年版）》，项目化学习逐渐走进基础教育各阶段素养导向下的课堂教学，受到了教育研究者和一线教师的重点关注。

相比于常规教学，项目化学习的教学目标、教学内容、教学形式等都发生了巨大改变，评价作为教学的指挥棒，理应发生对应的调整。而且项目化中的学习评价是推动项目化教学由知识目标转向素养目标的重要抓手，也是推进项目化学习高质量发展、学生素养发展过程中不可或缺的一环（余明华，2021）。项目化学习中指向学生认知、能力、情感等多元且复杂目标的特性决定了其评价要能够全过程、全方位关注学生在真实情境中整合知识、技能、情感等的综合表现，也要能够同时兼顾行为表现和项目成果（夏雪梅，2021）。表现性评价作为一种指向复杂的、高水平能力（周文叶，2022）的过程性评价方法，既能够通过过程性的指导和反馈，有效引领项目推进，促进学生素养发展，又能够依靠多元化的评价主体、评价方式及评价工具持续提供丰富可靠的评估证据，以刻画学生多方面能力的过程性变化，逐渐成为项目化学习评价中的焦点。

但表现性评价在项目化学习中可行但未畅行的原因在于：一方面，一线教师对于项目化学习中科学有效的设计和实施表现性评价的方法还不甚了解，面向项目化学习的表现性评价设计和开展过程不够清晰（周亮，2023）；另一方面，表现性评价在实际应用中常常流于形式，存在“新瓶装旧酒”的问题（蔡文艺，2022），如评价内容只关注项目成果、评价主体也只限制于教师、评价滞后于教学环节（评价即时性不够）（周青，2023）、与教学的整合性不

够等，这些都会给教师和学生带来比较重的评价负担，导致表现性评价在项目化学习中的应用效果差强人意。但在新课标积极鼓励项目化教学、核心素养导向驱动课堂评价变革的大背景下，面向项目化学习的表现性评价亟待开展。鉴于此，本研究尝试讨论如何将表现性评价科学有效地应用于项目化学习中，实现教学评一体化，以达到培养学生核心素养的目标。

2.表现性评价应用于项目化学习的适切性

2.1.项目化学习与表现性评价的内涵与特征分析

项目化学习是一种从现实生活中的真实问题出发，引导学生积极建构、共同探究、合作解决问题，形成公开可见的成果，以促进学生核心素养发展的学习模式。根据学者夏雪梅对项目化学习的定义和理解，可凝练出以下五个要素特征：（1）挑战情境引入：与传统教学不同，项目化学习的开展离不开近乎真实的、具有挑战性的情境，挑战性的情境是学生进入项目化学习、激活调动学生知识技能的必要条件，为情境问题的引出提供基础空间；（2）驱动问题引发：驱动性、开放性的问题是学生进行高质量项目化学习的必要条件，与学生真实生活有关的驱动性问题能够激发学生的学习兴趣，有利于深度学习的发生；（3）连续任务引导：连续持久的任务活动是保证学生深度思考、积极建构的重要条件；（4）项目成果公开：学生学习成效的可视化则表现在项目表现和成果公开中，反映了学生对情境问题的理解，体现学生重构整合知识来完成任务的能力水平；（5）评价活动支持：项目化学习评价在其中起到的作用不仅仅是在项目过程中刻画并描述学生的能力水平，更重要的是能够引导学生朝向更接近评价目标的方向发展。

表现性评价是一种在尽量合乎真实的情境下，运用评分规则对学生完成任务的学习表现和学习成果进行判断的评价方法，评价的往往是复杂、高层次、高水平的能力（周文叶，2021），其中包含三个基本要素：表现性目标、表现性任务、评分规则。依据其定义和要素，表现性

评价的设计和实施应该具备以下几个重要特征：（1）情境性：表现性评价为学生创造了在真实情境中解决问题的机会，只有在真实情境中学会的能力才是真实的能力，而通过应试做题习得的能力只适用于测验，无法迁移到新的情境之中；（2）表现性：由于评价任务往往来源于真实生活场景，是近乎真实且复杂的，而且指向高水平目标的评价任务往往需要学生积极主动地花时间、精力来建构任务方案，在这个过程中学生会产生丰富多样、连续层级的学习表现（柴唤友，2024）；（3）一致性：表现性评价中目标-任务-评分规则其实也反映着课堂教学的结构，一定程度上有机融合教与学，驱动教与学的一致性（向浩，2023）；（4）多元性：由于表现性评价往往检测的是复杂且高层次的能力，对应着评价目标也往往是多元化的，如包括但不限于知识技能、思维能力、情感态度、综合素养等等，评价目标的多元化又决定着评价手段和工具的多元性，不同的评价目标应该匹配不同的评价手段、工具，如学科知识目标可采用纸笔测验来评价，而思维能力这种具有连续层级的评价目标则该采用评价量规。此外，表现性评价方法还秉持多元参与的理念，主张学生也参与到评价中去，有利于提高学生对高层次表现的理解（周青，2023）。

2.2.表现性评价应用于项目化学习的适切性

根据上述对项目化学习和表现性评价内涵与特征的分析，表现性评价用于项目化学习评价的适切性体现在以下三个方面：（1）评价目标的一致性：项目化学习强调在项目实践中激活惰性知识，实现跨情境的迁移，指向的是学生在不同情境中创造性解决问题的能力，而表现性评价适合检测这种高水平、高层次能力；（2）评价过程的统一性：项目化学习是由情境问题引发一系列项目活动的建构过程，表现性评价要求学生在近乎真实的情境下主动投入、积极建构生成学习表现，这两者都包含情境、问题、实践、高阶认知等关键要素（杨宏，2023），在课堂评价实施过程中表现出高度统一性。此外，项目化学习是一种学习方式，表现性评价

是一种评价方法，两者都能够有机融合于课堂教学中，有助于实现减负提质、教学评一体化，这也是教学评价的整体变革路向；（3）评价开展的可行性：一方面在于项目化学习不同于以讲课-考试为主的传统教学模式，它是一种以完成任务和解决问题为主要形式的学习，学生在完成任务展示作品的过程中会产生很多可观测的行为表现和学习成果，而表现性评价的关键则是关注学生能“做什么”，显然这为表现性评价提供了良好的数据支撑和评价空间。另一方面，表现性评价体系能通过开放的评价任务、进阶的评分标准、多元的评价工具来引领项目有效进行、引发学生素养表现，这为收集学习者在项目化学习中的学习表现和成果提供客观证据和有效手段（安富海，2022）；（4）多元融合的适切性：项目化学习常常用于多学科或跨学科场景中，在评价上也往往包含多个评价目标和评价主体，这与表现性评价要素多元、层次丰富、操作灵活的特性相匹配。综上所述，表现性评价应用于项目化学习中具备一定的适切性和可拓展性。

3.项目化学习中表现性评价设计的基本思路与具体路径

3.1.设计思路：证据导向的逆向设计

项目化学习是一种在近乎真实情境下以完成项目任务和解决问题为主的建构性学习，不同于传统的“听课-练习-考试”学习模式，学生在项目进程中会产生很多可观测的行为和证据，而表现性评价方法可以为其提供丰富的手段和工具，以获得指向目标水平的表现性证据。而证据导向的逆向设计理念可以为设计面向项目化学习的表现性评价提供有效思路（桑国元，2023；安富海，2022；夏雪梅，2021）。本研究则基于证据导向的逆向设计理念将项目化学习与表现性评价联系在一起，并构建了以下的设计思路：（1）理念层：证据导向的逆向设计主张以评估员的视角来设计学习，包括确定预期结果、明确评估证据、设计学习体验和教学三个阶段（威金斯，2017）。其价值在于一方面能够分别指导项目化学习和表现性评价的

设计，另一方面又能够明确项目化学习设计和表现性评价设计之间的内在联系，进而整合教学与评价，破解评价滞后于教学的难题，达到教学评一体化的协同增质效果；（2）评价层：明确预期结果一方面是明确学生要达到的学习目标（即表现目标），另一方面是要确定表现性评价目标不同水平下的预设表现（即表现性标准）；其次，确定评估证据则是要找到学生在实际课堂中达到不同目标水平的具象化体现即明确用什么样的方式才可获得学生的表现性数据；而设计学习体验和教学则对应着设计引发学生体现目标水平的表现性评价任务；（3）项目层：具体到项目化学习中，证据导向的逆向设计能为项目化学习设计提供有效指导，其中预期结果对应着项目目标，学习体验对应着项目情境和任务，评估证据则对应着可观察的项目成果和表现，同时这也是表现性评价贯穿于项目化学习设计的具体表现。

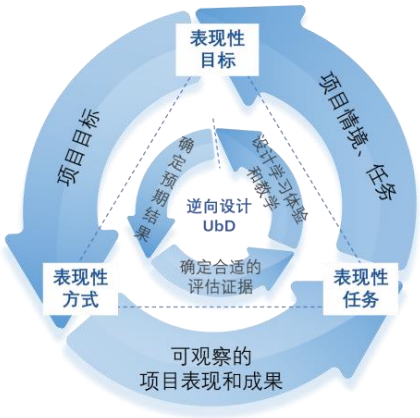


图 1 面向项目化学习的表现性评价设计思路

3.2.评价路径：贯穿评价要素的项目评价设计

基于证据导向的逆向设计思路，可将面向项目化学习的评价设计流程归纳为以下三个环节：设计表现性目标、设计表现性方式、设计表现性任务。

3.2.1.明确指向项目预期的表现性目标

表现性评价目标的设计是面向项目化学习的表现性评价设计的起点，也指向逆向设计的起点——想获得什么样的预期结果，具体到项目化学习中，则对应着项目化学习中学生要达到

的预期表现。项目化学习是一种在复杂情境中解决一系列问题的建构性学习方式，这决定了项目化学习的评价目标也应该是多元且复杂的，不仅仅要涵盖指向学科知识技能的认知目标，这些学科知识是建构复杂技能不可或缺的一部分，还应该包括情感态度、合作沟通等非认知目标。此外，项目化学习还强调深层次的概念理解和问题理解，这其实是指向素养这种综合性的、高层次的成就目标，也是学生在不同情境下运用综合技能进行问题解决的重要体现（林崇德，2017；夏雪梅，2021）。



图 2 指向项目预期的表现性目标

3.2.2.确定形成评估证据的表现性方式

在明确了学生要达到什么预期目标之后，应该要基于预期目标确定对应的表现性评价方式，这决定着如何获取学生完成项目过程中的表现性数据。此外，表现性评价方式是用于刻画学生反映目标水平的行为表现或学习成果，以获得能体现学生是否达到预期结果的评估证据，是表现性目标（抽象）映射到学生行为表现（具象）的直接体现。不同评价目标下对应的表现性评价方式、手段也是有所不同的，其适用范围也有所不同。周纯（2023）曾基于量规提出三类表现性方式：标准型评价量规、分析特征型评价量规（将学生实践方案设计或实践表现等根据不同的维度进行评估）、整体型评价量规（对实践方案设计、实践表现或实践作品等生成一个整体的等级或分值）。而国外研究者 Wiggins（1998）等人根据任务类型界定了三种评价方式：测验考试、学术讨论（描述一个主题或情境，使学生参与讨论）、表现性任

务和方案策划。詹泽慧等人（2022）则根据学生表现形式提出了行为类、作品类、工作纸类、试题类、量表类五类表现性方式。

那么本研究则基于证据导向的逆向设计思路，从项目化学习中能反映表现性目标的学生表现切入，学生在项目化学习中会通过意义协商、回答问题、反复实践、汇报成果、自我反思（胡红珍，2023）等外在行为表现出内在的知识、技能、情感乃至素养水平，基于这些可能会出现的反映目标水平的行为表现，可将项目化学习中表现性评价方式归纳为五类：纸笔测验类、量规量表类、报告汇报类和档案袋类。（1）纸笔测验类：学生在完成项目任务的过程中会调用一些学科知识技能，一些学科知识技能也是学生在项目化学习过程中必须要达到的基本目标，而检测学生是否掌握了必要的学科知识可以采用纸笔测验的方式，通过学生在良构或劣构问题上的作答反应来判断学生是否掌握基础的知识技能；（2）报告表单类：一些目标表现是需要通过学生自主完成才能够体现，如在项目开始前学生需要填写 KWH 表单工具（张灿永，2022）明确自身的知识经验、项目进程中学生会进行一些项目工作记录（詹泽慧，2022）、以及项目完成后学生需要完成反思单（杨宏，2023）总结经验等，这时评价不仅仅是完成目标-证据的闭环，还发挥了一些引导、调节的功能，帮助学生找到努力的方向，提高自身的元认知水平；（3）行为表现类：学生在项目化学习过程中出现的一些外在行为，都是可以成为检测内在目标是否达到的证据的，如学生之间意义协商的会话互动行为、学生实践过程中的调试修订行为、公开汇报时的语言表达、项目结束时的反思汇报等等，通过一定的手段或工具观察收集这些学生表现可将其变为指向目标的有力证据；（4）作品成果类：作品成果是项目化学习的重要特征，也是表现性评价的重要内涵体现，是反映学生能做什么的重要形式，对作品成果的评价也可以视为项目化学习中的终结性评价，从学生完成作品成果的程度来综合判断学生达到项目目标的综合水平。如对学生反映学生计算思维水平的编程作

品进行评价、或对学生巡线机器人作品的全面评价等等；（5）档案袋类：档案是一种可以全面记录学生学习过程中学习表现的工具，既可从多维度上横向看学生的综合发展，也可从时间维度上纵向看学生的增值发展，是项目化学习中实现全面评价的重要评价方式。

这几种表现性评价方式能够同时兼顾学生在项目化学习中的过程性表现和最终的项目成果，足以满足项目化学习评价中全方位、全过程的评价需求。在表现性评价方式的基础上，设计合适的评价工具是将这些学生表现转化成评估证据的重要环节。量规（Rubric）作为一种具体罗列评价维度与评分标准的评分工具（陈素平，2020），常常用于项目化学习和表现性评价中，一方面其能通过呈现连续层级的表现性标准避免给出单一孤立的分数，另一方面也是为了告知学生努力的方向，给学生提供支架，引导学生朝着预期目标继续进步，发挥以评促学的重大效用。

表 1 五种表现性评价方式

评价方式	学生表现	适用范围	评价者	评价工具
纸笔测验类	作答反应	适用反映学生对知识技能的掌握程度	教师	评分表
报告表单类	文稿图纸	适用于反映学生的认知状态、认知过程、情感态度等	学习者自身	量规
作品成果类	项目成果	适用于反映学生在项目学习中的综合表现情况	教师、同伴或自身	量规
行为表现类	外在行为	适用于通过观察记录来反映学生知识、技能、情感等方面的情况	教师、同伴	量规
档案类	任何可观察的数据	适用于全面记录学生全过程表现，但较费时费力，目前电子档案比较常见	教师	评价档案

3.2.3.设计贯穿项目全程的表现性任务

表现性评价任务是表现性评价设计的核心，是引发学生表现出目标水平的必要条件，同时也是逆向设计中学习体验设计和项目化学习实施的具体表现。有研究者表明完整的项目就是由一个个相互承接、逐级递进的任务构成（胡嘉康，2023），表现性评价任务作为项目化学习中的重要组成部分，不仅仅起着引发学生学习表现的评价作用，也是驱动引导项目进行的一个个进阶支架。

从表现性任务的纵向维度来看，表现性评价设计的任务环节与项目化学习中的实践活动是相互依赖、互相联系的。因此，表现性任务的设计要能够体现项目化学习的结构和进程，同时也应该满足表现性评价任务的特征属性。项目化学习本质上是一个以项目问题解决为导向，需要调动学生知识、技能、情感态度等多方面能力的学习建构过程，夏雪梅老师（2021）曾提出项目化学习的六个阶段：入项、知识与能力建构、探索与形成成果、评论与修订、公开成果以及反思与迁移。由于这六个阶段都是由持续的学习实践活动构成的（夏雪梅，2021），那么为了更加具体地表征项目化学习的任务环节，本研究将结合夏雪梅老师提出的五种项目实践类型，凝练出项目化学习中的表现性任务环节：情境问题理解、知识能力建构、项目成果形成、成果公开与修订、反思与迁移。由于项目的完成往往依赖于个体与个体之间的共同理解、相互沟通。因此，其中每个环节中都应该包括社会性的实践活动，也有研究表明协作是与项目化学习经常关联出现的词汇（周亮，2023），团队协作在项目化学习中至关重要，可提高项目化学习的效果和效率（CHANG L C，2010）。

从表现性任务的横向维度来看，构成项目的表现性任务设计应该给学生呈现一个具有挑战性、可能性的待解决问题，学生为了一个确定的对象开发具体的产品或作出相应的表现，同时应该被提前告知评估指标和表现标准。那么依据威金斯等人（2021）开发的 GRASPS 模

型可凝练出表现性任务设计的六个元素——Goal（目标）、Role（角色）、Audience（对象或受众）、Situation（情境或处境）、Performance/Product（产品）、Standards（提供成功标准）。（1）目标（goal）：指向项目目标，或者说在这个项目里“我们要解决什么问题？”，任务目标的描述应该清晰、具体，使学生能读懂、理解；（2）角色（role）：要明确学生担任的角色是什么？学生将自己代入角色后，会主要从角色的视角去思考，从而完成任务、解决问题，有更强的真实感、代入感；（3）对象或受众（audience）：要确定项目任务面向的客户或服务对象是谁？学生可在角色和受众之间进行换位思考，从客户或服务对象的视角去反思、改进、优化任务，引发学生的自我调节与持续成长；（4）情境或处境（situation）：任务的情境一定程度上决定了任务乃至项目的复杂性，往往蕴含着项目要解决什么问题、现在具备什么条件等等；（5）产品（performance/product）：要让学生了解任务完成需要交付的产品或成果是哪些，既包括个人成果，也包含团队成果，其中产品应该具备可视化、可衡量、可评估的特性；（6）成功标准（standards）：任务是否成功应该提前告知学生具体的标准，或者为学生提供相关的案例供学生参考。

目前基础教育阶段各学科教学往往通过项目的形式开展，发展的是学生运用学科、跨学科等知识与技能解决问题的综合能力，评价是教学实践中的重要环节。表现性评价作为一种嵌入课堂教学的评价方法，为评价项目化学习中学生多方面的综合能力指明了方向。本研究后续将基于上述思路和基本路径开展基础教育阶段的课堂实践，以期获得良好的评价效果。

参考文献

Grant W，Jay M． 追求理解的教学设计(第二版)[M]．闫寒冰，宋雪莲，赖平，等，译．上海:华东师范大学出版社，2017: 77．

安富海. (2022). 项目化学习的实践困境及改进策略研究. 上海师范大学学报 (哲学社会科学版)，*51*(4), 119–125.

蔡文艺. (2022). 项目化学习中表现性评价的运用研究. 上海教育科研, *2022*(12), 62–66.

柴唤友, 郑勤华, 胡天慧, & 王怀波. (2024). 基于信息技术的表现性评价：概念解析、构成要素及分类框架. 中国电化教育, *2024*(2), 7.

陈素平. (2020). 让评价支持和促进学习——美国 STEM 项目学习评价的特点与启示. 上海教育科研, *4*, 18–22.

胡红珍 & 夏惠贤. (2023). 第三空间视角下项目式学习的主要特征、学习序列与教师行动. 全球教育展望, *52*(12), 120–128.

胡嘉康 & 田莉. (2023b). 指向教学变革的项目式学习设计：价值定位、形态厘清与实践策略. 现代教育科学, *4*, 125–131, 138. 林崇德. (2017). 构建中国化的学生发展核心素养. 当代教育家, *2*, 10–13.

桑国元, 叶碧欣, 黄嘉莉, & 罗颖. (2023). 构建指向中国学生发展核心素养的项目式学习标准模型. 中国远程教育, *43*(6), 49–55.

夏雪梅. (2022). 指向核心素养的项目化学习评价. 中国教育学刊, *9*, 50–57.

向浩, & 李萍. (2023). 表现性评价: “教”与“学”方式转变关键着力点. 语文建设, *3*, 50–55.

杨宏. (2023). 量规:表现性评价应用于学科项目化学习的指南针. 基础教育课程, *5*, 55–64.

杨宏. (2023). 让评价为学习增值——表现性评价嵌入项目化学习的校本实践. 基础教育课程, *19*, 66–72.

詹泽慧, 姚佳静, 吴倩意, & 黄秉刚. (2022). 人工智能课程中表现性评价的设计与应用. 现代教育技术, *32*(5), 32–41.

张灿永. (2022). 初中科学课程跨学科项目化模式的建构、实施与测评. 教学与管理, *10*, 56–58.

周亮 & 陈明选. (2023). 国际教育技术领域近十年项目化学习研究分析. 软件导刊, *22*(8), 221–229.

周青, & 田建荣. (2023). 基于核心素养发展的表现性评价设计. 考试研究, *19*(4), 1–9.

周文叶, & 毛玮洁. (2022). 表现性评价：促进素养养成. 全球教育展望, *51*(5), 94–105.

CHANG L C，LEE G C. A team-teaching model for practicing project-based learning in high school：collaboration between computer and subject teachers［J］. Computers & Education，2010，55（3）：961-969.

混合式教学中同伴互评对共享调节能力的影响作用研究

Research on the influence of peer review on shared adjustment ability in blended teaching

赵赫璇*，马国威 ，付玉娜，孙伟华，区文诗，王妍莉
西北民族大学教育科学与技术学院
*3466270662@qq.com

【摘要】：共享调节是当前协作学习的重要研究方向，是学生学习质量的重要表征；而同伴互评可以促进学生认知、元认知的发展，具有促进深度学习的潜能。基于此，文章首先构建了促进共享调节能力的同伴互评框架，随后以“现代教育技术”课程为例设计了同伴互评活动，最后从共享任务分析、共享计划、共享阐述、共享监控和策略应用五个维度研究了混合式教学中同伴互评对共享调节能力的促进作用，并探究了同伴互评对共享调节能力的影响机制。文章发现：混合式教学中的同伴互评促进了学生的共享调节水平。文章探索了共享调节能力的发生机制，可以更好地指导同伴互评活动，并为促进共享调节能力的同伴互评实践活动提供参考。

【关键词】：同伴互评；共享调节能力；混合式教学

Abstract: *Sharing regulation is an important research direction of collaborative learning and an important representation of students ' learning quality. Peer review can promote the development of students ' cognition and metacognition, and has the potential to promote deep learning. Based on this, this paper first constructs a peer review framework to promote shared adjustment ability, then designs peer review activities with the course of ' Modern Educational Technology ' as an example, and finally studies the role of peer review in promoting shared adjustment ability in blended teaching from five dimensions : shared task analysis, shared plan, shared elaboration, shared monitoring and strategy application, and explores the influence mechanism of peer review on shared adjustment ability. It is found that peer review in blended teaching promotes students ' level of shared regulation. This paper explores the mechanism of shared regulation ability, which can better guide peer review activities and provide reference for peer review practice activities that promote shared regulation ability.*

Keywords: *Peer review ; shared regulation ability ; blended teaching*

本文为教育部人文社科规划基金项目“基于共享调节的大学生群体认知生成路径及优化机制研究”（项目编号：23YJA880053）阶段性研究成果；受西北民族大学本科教学质量提升项目省级一流课程《信息技术教学法》（项目编号：2022SJYLKC-37）资助；受“中央高校基本科研业务费专项资金”(项目编号：31920240127)资助。

1、引言

同伴互评是在线协作学习中的一种重要活动，它允许学习者通过相互评估来共同建设和完善知识，培养高阶思维和能力，如批判性思维和评估技能，这些都是 21 世纪不可或缺的技能（马宁等，2022）。此外，同伴互评还有助于促进学习者之间的交流和理解，提高学习投入度，并为群体创新活动提供支持（李红霞等，2021）。然而，有效的同伴互评并不是一件简单的事情，它需要参与者在评估过程中投入足够的注意力和精力，以确保评估的准确性和有效性（李红霞等，2020）。

在进行同伴互评的过程中，学习者需要围绕评价目标，通过集体性任务规划、共享监督、相互评估和团队管理等调节行为，解决和应对评估过程中可能出现的认知冲突、评价标准不一致、评价深度不足等问题（马志强等，2016）。例如，有研究者开发了一种基于网络的同伴互评系统，以帮助学习者在评估过程中更好地理解和应用评价标准，从而提高评价质量（Alemdag&Yildirim，2022）。

尽管同伴互评在在线协作学习中具有重要作用，但目前对其影响共享调节学习行为的研究仍然不足。为此，本研究通过实施为期 9 周的准实验，采用 SPSS 分析和滞后序列分析等方法，旨在探讨在线协作学习中同伴互评活动的影响，以及如何设计和实施有效的同伴互评活动，以提高在线协作学习的效果。

2、文献综述与研究问题

2.1 混合式教学环境中的共享调节学习研究

混合式教学环境下，共享调节对学生的协作学习效果是具有促进效果的，例如：Tsai 等人通过他们的研究指出，网络媒介作为一种重要的工具，可以显著提升学生在共享调节学习过程中的学习绩效（Tsai et al.,2018）。共享调节学习是一种学习方式，它强调学习者通过共同制定和遵循协作规则和标准，共同完成学习任务。该研究表明，通过网络媒介进行的共享调节学习，可以帮助学生更好地理解和应用知识，提高学习效果（Tsai et al.,2018）。Li 的研究进一步证明了共享调节在学习过程中的积极作用。在基于游戏的合作学习环境中，共享调节对学生的积极学习态度及学习绩效有促进作用。这表明，在游戏化的学习环境中，通过共享调节学习，学生可以更好地参与学习，提高学习态度和绩效（Li et al., 2023）。

2.2 同伴互评对共享调节行为的影响

共享调节行为是指在学习过程中，学生共同制定目标、监控学习活动和进度和调整策略并及时反思的行为。近年来，同伴互评在教育领域受到了越来越多的关注。同伴互评能够提高学生的元认知能力。在同伴互评过程中，学生需要审视他人的学习过程和成果，从而激发对自己的思考，明确自己的学习目标。通过对比分析，学生可以更深入地了解自己的优势和不足，提高对学习过程的监控和调整能力。为了提高同伴互评的效果，教师需要引导学生明确评价标准，培训评价

技能，确保评价的客观性和公正性。同时，教师还要关注同伴互评过程中的情感因素，营造轻松、和谐的学习氛围，使学生在互评过程中能够真诚地给予反馈，勇于接受他人的批评和建议。通过同伴互评，学生可以提高元认知能力、问题解决能力，促进情感交流和人际沟通能力。

2.3 研究问题

本研究通过实施一项为期 9 周的准实验，对同伴互评如何影响共享调节学习行为进行研究，具体研究问题如下：（1）同伴互评对共享调节水平的影响是怎样的？（2）实验班小组前后测的共享调节行为有怎样的差异？

3、研究设计与实施

3.1 研究对象

本研究以西北地区 X 大学 2021 级教育技术学专业的两个班级，共 78 名本科生为研究对象，并在“现代教育技术”课程中开展了同伴互评活动。此课程阐述了现代教育技术课程在中学教育的地位和作用，是教育技术学专业必修课程。课程内容主要包括现代教育技术概述、信息化资源的获取和处理、教学设计、信息化教学资源的设计和开发、信息化教学环境。课程开展的环境为面授和在线教学环境：面对面教学环境为网络互动教室；在线学习环境为“超星学习通”平台，此平台具有组织管理、互动交流、开展教学活动等功能，并且能够记录学生的学习进度和交流、讨论、答疑、作品等活动数据。

3.2 研究工具

3.2.1 共享调节水平量表

共享调节水平量表选自罗淳（2019）编制的量表。该问卷共 21 道正向题，包含共享认知、共享元认知、任务冲突和共享情感动机四个一级维度（罗淳，2019）。问卷采用里克特五点计分法:“完全不同意”到“完全同意”。

3.2.2 共享调节行为编码

为了深入分析对象在网络环境下的共享调节行为模式，本研究应用了滞后序列分析法和编码框架，该框架专注于捕捉协作学习中的关键调节活动。正如表 1 所展示的，该编码框架聚焦于五个核心活动领域：共享任务分析、共享计划、共享阐述、共享监控以及策略运用。通过这一方法，两位研究人员独立地对多个在线学习小组的互动数据进行了编码。随后，他们使用滞后序列分析技术对这些数据进行了系统的分析。

表 1 编码框架

行为类型	编码	描述
共享任务分析	A	讨论在线协作组任务，并进行学习任务分工
共享计划	P	对在线协作组或其他成员表达任务计划的观点，包括为完成任务而应采取的下一步行动

共享阐述	E	阐述自己的想法，使在线协作组成员能够了解自己的想法
共享监控	M	管理在线协作学习时间，或适当地对在线协作组给出“指令”，或协商对话以解决冲突，并达成相互理解
策略运用	S	检验自己提出的主意，或采纳在线协作组成员提出的策略建议

3.2.3 自我调节水平量表

自我调节水平采用 Pintrich 等 1991 年编制的 MSLQ 量表，包括元认知自我调节和努力调节两个维度（Pintrich，et al., 1991）。问卷采用里克特五点计分进行评分，1-5 分别代表“完全不符合”-“完全符合”，且问卷具有良好的信效度。

3.3 研究流程

研究实施环境是基于泛雅超星学习通网络教学平台，参与本次研究的学生通过此平台参与在线协作学习活动。第 1 周，学生接受了超星学习通平台的使用培训，且了解现代教育技术概述，并填写在线自我调节水平量表和共享调节水平量表；随后，根据学生的自我调节水平得分，且以此为依据进行异质性分组。第二至八周，学生分别围绕四个在线协作学习任务开展协作活动，主要包括面授课堂中的确定项目主题，完成项目设计，开发作品初稿，开发作品定稿。第 9 周，学生再次填写在线共享调节水平量表，研究者对学生的在线学习数据进行采集、清洗，以及行为编码与分析。

4、数据分析与结果

4.1 同伴互评对共享调节水平的影响

4.1.1 问卷设计

调查问卷是本研究收集数据的主要工具和方法。问卷采用封闭式题型，问卷结构主要包括问卷前言，向调查者说明本研究的研究背景，调查目的以及调查内容；第二部分为实验者的基本信息调查；第三部分为实验者自我调节水平量表、共享调节水平量表，采用 Likert Scales 五级量表，由“非常不同意”到“非常同意”。

4.1.2 问卷回收

于西北某大学发放 80 份电子问卷，回收 80 份，回收率 100%，剔除 3 份无效问卷，获得有效问卷 77 份，有效率为 96.25%。

4.1.3 实验班前后测共享调节水平对比

表 2 实验班配对样本统计

	平均值	个案数	标准 偏差	标准 误差平均值

前测共享调节	3.4862	39	.44806	.07175
后测共享调节	3.9248	39	.37712	.06039

表 3 实验班配对样本检验

	配对差值							
	平均值	标准偏差	标准误差平均值	差值 95%置信区间		T	自由度	Sig（双尾）
				下限	上限			
前测共享调节-后测共享调节	-.43860	.59176	.09476	-.63042	-.24677	-4.629	38	.000

通过实验班配对样本统计我们可以发现在共享调节能力方面有了显著的提升。从 3.4862 提升至 3.9248，表明了个体在团队工作中的资源整合和协调能力有所增强。此外，共享调节能力的提高也可能与个体的社交技能、沟通能力和团队意识的发展有关。这为个体和社会的和谐发展提供了积极的支持。

根据实验班配对样本检验我们可以得知，双尾检验的 Sig.值均为 0.000，小于 0.05 的常用显著性水平，我们可以得出以下结论：实验干预对学生共享调节水平的提升效应是显著的。这意味着通过实验干预，学生的共享调节能力得到了显著的提高。

4.1.4 对照班前后测共享调节水平变化

表 4 对照班配对样本统计

	平均值	个案数	标准 偏差	标准 误差平均值
前测共享调节	3.4260	38	.58394	.09473
后测共享调节	3.6680	38	.56218	.09120

表 5 对照班配对样本检验

	配对差值							
	平均值	标准偏差	标准误差平均值	差值 95%置信区间		T	自由度	Sig（双尾）
				下限	上限			
前测共享调节-后测共	-.24196	.53970	.08755	-.41935	-.06456	-2.764	37	.009

享调节

根据实验班和对照班配对样本统计对比可知，实验班和对照班在后测共享调节能力上均有所提高，但实验班的提升幅度更大。实验班共享调节能力的提高表明实验干预对提升学生的协作能力和资源整合有显著效果，而对照班共享调节能力的提升可能仅仅反映了时间的推移或自然成长，其提升幅度不如实验班显著。

根据对照配对样本检验我们可以得知，双尾检验的 Sig.值均为 0.000，小于 0.05 的常用显著性水平，我们可以得出以下结论：无干预对学生共享调节水平的提升效应也是同样显著的。

4.2 小组前后测的共享调节行为分析

4.2.1 整理小组前后测各个阶段频次结果

小组前测 A（共享任务分析，35）;P（共享计划，7）;E（共享阐述，22）;M（共享监控，2）;S（策略运用，2）。

小组后测 A（共享任务分析，26）;P（共享计划，23）;E（共享阐述，97）;M（共享监控，24）;S（策略运用，17）；

从统计数据可以看出，小组在后测阶段在共享计划、共享阐述、共享监控和策略运用这些方面的表现有了显著提高，这表明他们在这些领域的沟通和协作技巧得到了加强。这种进步很大程度上得益于成员们在实际操作中积累的经验，随着任务的不断实施，他们学会了如何更有效地协调和规划，以提高工作效率。然而，值得注意的是，在共享任务分析这一项上，小组的后测表现相较前测有所下降。这种变化可能是由几个因素造成的。首先，小组在任务分配和角色确认上可能遇到了一些问题，这影响了他们在分析任务时共享信息和观点的能力。其次，小组成员可能还未完全掌握任务分析的方法和技巧，这导致了分析能力的下降。最后，小组在后测可能过于专注于其他方面的表现，如共享计划和共享阐述，而忽略了任务分析的重要性，这可能导致他们在这一方面的表现不如前测。

4.2.2 共享调节课堂前后测行为模式分析

依据滞后序列分析的原理，当 Z 值>1.96 时，就认为该行为序列存在统计意义上的显著性（p<0.05）。本研究依据小组前后测变化总过程的残差值，绘制了两组的共享调节课堂的变化五阶段行为模式图，如图 1、2。



图 1 小组前测行为模式图

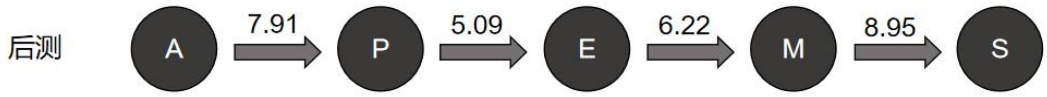


图 2 小组后测行为模式图

根据行为模式图，我们可以得出以下结论：

1、在前测阶段，小组成员在完成任务分析后，很少能顺利地进入到计划阶段，这一转变的困难可能源于缺少清晰的指导或协调。计划阶段对后续的阐述阶段几乎没有影响，这表明小组成员在制定计划时未能进行充分的沟通和解释。阐述阶段对监控阶段的影响也较小，这是因为监控机制不明确或成员对监控阶段不够重视。最后，监控阶段对策略运用的影响也较小，这是因为成员在监控后的策略应用能力和意识不足。

2、在后测阶段，情况有了显著的改善。任务分析对计划阶段的影响增强，成员在完成任务分析后能更有效地制定计划，这是由于他们通过实践和学习提高了计划能力。计划阶段对阐述阶段的影响也增强，成员在计划完成后能更清晰、更充分地解释他们的计划，这是由于他们对计划的理解更明确了，沟通也更为有效。阐述阶段对监控阶段的影响同样增强，成员在解释计划后能更积极、更有效地进行监控，这是由于他们对监控过程的重视程度提高了。最后，监控阶段对策略运用的影响也增强，成员在监控后能更有效地应用策略，这是由于他们的监控能力和对策略运用的理解都有所提升。

5、研究结论与建议

5.1 结论

5.1.1 同伴互评对共享调节水平有促进作用

首先，根据实验班和对照班配对样本统计对比可知，实验班和对照班在后测共享调节能力上均有所提高。这表明，无论是否实施同伴互评，共享调节能力都会在教学过程中得到提升。然而，实验班的提升幅度更大，这表明同伴互评对共享调节水平的提升具有显著的促进作用。同伴互评能够促进共享调节能力的提升，原因如下：一方面，同伴互评过程中，学生需要互相交流、讨论，这有助于培养学生的人际沟通能力和团队协作精神（唐晓丽，2018）。在交流过程中，学生可以相互启发、取长补短，从而提高自身的共享调节能力。

5.1.2 实验班小组前后测的共享调节行为存在差异

经过对实验班和对照班在前测和后测的共享调节能力的对比分析，我们发现实验班在后测的共享调节能力上有了显著的提升，尤其是在共享计划、共享阐述、共享监控和策略运用方面。然而，在共享任务分析方面，实验班的表现却有所下降。这一结果表明，同伴互评教学策略在提高学生的共享调节能力方面具有显著效果，但需要注意对任务分析方面的培养。因此，在今后的教学中，应将同伴互

评与其他教学方法相结合，以全面提高学生的共享调节能力（刘怡然，2023）。

5.2 建议

5.2.1 充分利用同伴互评提高共享调节能力

同伴互评是一种有效的教学策略，能够在学生之间建立积极的互动和合作，提高共享调节能力。为了更好地利用同伴互评提高学生的共享调节水平，可以注意加强同伴互评的培训和指导。同伴互评并非简单的相互评价，而是需要学生具备一定的评价能力和批判性思维（张涛等，2022）。因此，教师应提供充分的培训和指导，教会学生如何进行客观、公正的评价，并强调评价的重要性。

5.2.2 关注共享任务分析能力的培养

共享任务分析是小组协作学习中的重要环节，对于提高学生的共享调节能力具有重要作用。为了更好地关注共享任务分析能力的培养，可以注意明确任务分析的要求和指导。确保学生能够更好地进行任务分析，教师应为学生提供明确、详细的任务分析指导，帮助学生掌握任务分析的方法和技巧。

参考文献

Alemdag, E., & Yildirim, Z. (2022). Design and development of an online formative peer assessment environment with instructional scaffolds. Educational technology research and development, 70(4), 1359-1389.

Li, J., Lin, Y., Sun, M., & Shadiev, R. (2023). Socially shared regulation of learning in game-based collaborative learning environments promotes algorithmic thinking, learning participation and positive learning attitudes. Interactive Learning Environments, 31(3), 1715-1726.

Pintrich, P. R., Smith, D. A. F., Teresa Duncan, & McKeachie, Wilbert J. (1991). A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ). Retrieved August 29, 2016, from http:// file:///D:/Google%20Download/MSLQ%20manual.pdf

Tsai, C. W., Shen, P. D., Chiang, I. C., Chen, W. Y., & Chen, Y. F. (2018). Exploring the effects of web-mediated socially-shared regulation of learning and experience-based learning on improving students' learning. Interactive Learning Environments, 26(6), 815-826.

马宁,路瑶,郭佳惠等.评价支架对教师在线同伴互评质量的影响研究[J].电化教育研究,2022,43(02):34-41.

马志强,龙琴琴,王靖等.促进知识表征的在线协作评价机制研究[J].远程教育杂志,2016,34(04):93-101.

刘怡然.高校英语写作教学法回顾与同伴互评研究——评《英语写作教学与研究》[J].中国高校科技,2023(04):109.DOI:10.16209/j.cnki.cust.2023.04.009.

罗淳. 协作学习中共享任务理解的研究[D].华东师范大学,2019.

李红霞,赵呈领,蒋志辉.匿名与量规对基于 SPOC 的混合式同伴互评投入度的影响[J].现代教育技术,2020,30(10):20-27.

张涛,张思,高倩倩等.促进在线同伴互评中学习者批判性思维发展研究[J].电化教育研究,2022,43(06):53-60.DOI:10.13811/j.cnki.eer.2022.06.008.

唐晓丽.“互联网+”背景下初中英语写作教学中同伴互评的应用研究[J].现代职业教育,2018(22):4.

生成式 AI 支持的人机协同论证：基于两种提示设计的实证比较

Generative AI-enabled human-computer collaborative argumentation: comparison of two types of prompt design

崔鑫[1]，尤欣雅[2]，马志强[2]
[1]江南大学设计学院
[2]江南大学教育学院 江苏“互联网+教育”研究基地

【摘要】 传统协同论证情境中，学生难以自发进行高质量的论证性对话。提示作为引发高质量对话的灵活干预，常被应用于计算机支持的协同论证情境中。ChatGPT 类生成式人工智能的出现，具备与学生进行人机协同论证的潜能，然而学生需要输入恰当的提示才能得到满足自己需求的回应。这意味着在生成式人工智能支持的人机协同论证情境中，提示设计成为人机协同论证的核心。基于此，该研究以生成式 AI 支持的人机协同论证活动为情境，以引导论证性对话的提示设计原则为基础，结合人机对话的特点，设计修辞性提示与辩证性提示，比较不同提示下人机协同论证活动对学生论证话语模式的影响。研究发现，修辞性提示下的人机协同论证活动使得学习者注重从证据到主张的推理过程，且能够积极地反驳他人的观点，为自己的观点进行辩护；辩证性提示下的人机协同论证活动使得学习者更加注重证据来源的可靠性，且能够在审查彼此证据的的基础上达成共识。
【关键词】 协同论证；生成式 AI；人机协同论证；提示语工程；ChatGPT

Abstract: *Students have difficulty spontaneously engaging in high-quality argumentative dialog during collaborative argumentation. Prompts, as flexible interventions for triggering high-quality conversations, are often used in computer-supported collaborative argumentation contexts. The emergence of generative AIs has the potential to engage in human-computer collaborative argumentation with students, however, students need to input appropriate prompts in order to receive a response that meets their needs. This implies that prompt design becomes the core of human-computer collaborative argumentation in generative AI-supported human-computer collaborative argumentation contexts. Based on this, this study takes the generative AI-supported human-computer collaborative argumentation activity as a context, takes the principles of prompt design for guided argumentative dialogues as the basis, and designs rhetorical prompts and dialectical prompts in combination with the characteristics of the human-computer dialogues to compare the effects of the human-computer collaborative argumentation activity on students' argumentative discourse patterns under different prompts. The study found that the human-computer collaborative argumentation activity with rhetorical prompts made the learners focus on the reasoning process from evidence to claim, and they were able to actively refute others' views and defend their own views; the human-computer collaborative argumentation activity with dialectical prompts made the learners pay more attention to the reliability of the evidence sources, and they were able to reach a consensus on the basis of reviewing each other's evidence.*

Keywords: collaborative argumentation, GAI, human-computer collaborative argumentation, prompt**,** ChatGPT

1. 研究背景与文献综述

协同论证是指两个或多个成员针对同一个议题提出各自的论点，通过比较、评估和综合不同的论点以达成共识的过程，有利于学习者调动更多的认知资源，发展高阶认知能力，如推理、批判、反证等(Rapanta & Felton,2022)。判断与整合多方的论点往往需要高质量的论证

性对话(Harney et al., 2015) 。而由于论证任务的复杂与开放性，学生难以自发进行高质量的对话，因此在协同论证研究与实践中，常运用提示来引发论证主体之间的交互，从而提升论证质量，达成共识。提示常表现为引导性问题、线索、句子开头等多种形式。例如，提示学生检查彼此证据的来源、提示学生进一步论证自己的主张等(Webb,1995) 。实证研究表明，提示能够提升协同论证的质量 与论证性文章写作的深度 。之后，研究者开始探讨如何设计提示，设计何种类型的提示能提升论证质量，从而指导协同论证的教学实践。Malogianni 等人(2021)将现有协同论证中的提示划分为修辞性论证提示（rhetorical argumentative prompt）与辨证性论证提示（dialectical argumentative prompt）。修辞性论证提示能够引导学生在论证时从利益相关者的立场进行更深入的思考，目的是说服他人。辨证性论证提示鼓励学生在论证过程中考虑话题的两面性，思考与自己立场相反的观点。研究发现，在有关商业问题的议论文写作上，两种类型的提示都能引发学生使用验证与阐释等论证策略。然而，少有研究探讨两种提示设计的依据与原则，也未能在更广泛的情境中比较两种提示的作用。

随着计算机支持的协作学习(CSCL)领域的发展，提示被嵌入到各种计算机环境中，如知识建构平台、对话代理等。对话代理是能够运用自然语言与人类用户进行交互的一类技术的统称，包括实体和非实体代理，例如基于文本的聊天机器人、基于语音的虚拟助手等。在协同论证实践中，对话代理的研究方向之一是通过人机对话引导和支持学习者思考与讨论论证问题。在对话代理支持的协同论证中，提示表现为学习者的提问，即学习者通过提问与对话代理进行讨论，对话代理选择或生成最相关或相反的论点与学习者保持连贯与活跃的对话。此类代理的开发大多基于检索模型，通过检索算法从预设的论点库中检索论点(Guo et al., 2023)。缺陷是如果学生的提问超出论点库的范围，对话代理将无法响应，这导致人机对话脱离了社会情境，且难以对学习者的需求进行个性化的支持，可能使得学生对对话代理失去信任(Rheu et al., 2020)，影响协同论证的效果。

生成式人工智能的出现，使得更自然、更个性化的人机对话成为可能。以 ChatGPT 为代表的对话代理从包含上万亿互联网文本中进行预训练后，再从人机对话中获取反馈，并根据反馈进行微调以提高性能。这使得 ChatGPT 几乎可以理解用户的任何提问，并提供连贯且有逻辑的回答(Chan & Hu, 2023)，并且界面简洁、免费且易于使用。已有研究证明，生成式人工智能支持的人机协同论证能够激发学生的想法，拓宽学生的论证视角(Su et al., 2023)。然而，由于生成式人工智能的生成性、开放性与复杂性，在与 ChatGPT 交互时，用户需要输入恰当的提示才能得到满足自己需求的回应。提示的内容、视角和顺序都将影响 ChatGPT 生成的预测性内容。因此，提示语工程成为人工智能的热门子领域（Clark, 2023）。提示语工程是用户有组织地将提问想法转化为一组指令集的过程，目标是使得 ChatGPT 输出满足用户期望的回应，以实现有效的人机对话。这意味着在生成式人工智能支持的人机协同论证情境中，提示设计成为人机协同论证的核心。现有教育领域提示语设计研究取得了初步进展，如设计框架 CLEAR (Lo, 2023)和设计原则(Hwang & Chen, 2023)等。这些研究关注更广泛的教育情境下提示的语言与结构。而在协同论证领域中，已有研究探索了在学生论证结束后，如何设计提示引导 ChatGPT 对学生论证内容的评价与反馈，但尚未有研究关注人机协同论证情境中的提示设计。

基于此，本研究将以人机协同论证活动为情境，以引导论证性对话的提示设计原则为基础，结合人机对话的特点，设计修辞性提示与辨证性提示，比较不同提示下人机协同论证活动对学生论证话语模式的影响。研究问题如下：

（1）参与不同提示条件下（无提示、修辞性论证提示、辨证性论证提示）人机协同论证活动的学生小组，其论证话语的结构性模式有何不同？

（2）参与不同提示条件下（无提示、修辞性论证提示、辩证性论证提示）人机协同论证活动的学生小组，其论证话语的社会性模式有何不同？

2. 生成式 AI 支持的人机协同论证活动设计

本研究依据 Mark (2023)提出的生成式人工智能支持学习与认知发展的指导框架和 Linn 等人(2006)的知识整合框架，构建了结合生成式人工智能教育应用和协同论证原则的人机协同论证活动。其中，生成式人工智能支持学习与认知发展的指导框架包含五个层次：理解生成式人工智能技术工具的能力和限制(即理解)、操作生成式人工智能技术工具(即接入)、通过提示语与生成式人工智能技术工具互动(即提示)、验证响应的准确性(即证实)，以及将信息融入任务(即整合)。知识整合框架涵盖观点生成、添加、辨别和反思。两框架的整合构成了人机协同论证活动的核心结构，如图 1 所示。

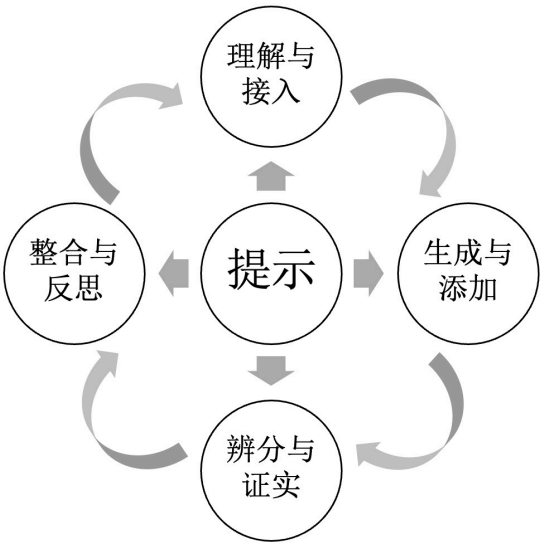


图 1 生成式人工智能支持的人机协同论证活动框架

理解与接入阶段：在活动开始前，教师向学生介绍生成式 AI 技术工具的功能和局限，并对 ChatGPT3.5 与提示语工程进行培训。

生成与添加阶段：教师阐述学科中的关键问题及其争议，提出论证主题。活动开始后，小组成员运用在线异步沟通工具企业微信通过文字交流进行协同论证。这种交流模式为所有学习者提供了足够的时间来构思论点和收集论据，避免即时交流中可能出现的错误和冲动回应，并且能够留下详细记录，便于后续的回顾与反思。在协同论证过程中，小组成员首先基于讨论主题独立思考，明确自己的立场和论点，并将其表达在企业微信小组群聊中。接着，通过提示与 ChatGPT 进行对话，引入更多的思考角度。

辨分与证实阶段：每位小组成员结合自身的知识经验或外部资料，对同伴和 ChatGPT 的观点进行评估、追问和解释。通过提示来谨慎判断 ChatGPT 响应内容的来源与质量，通过横向比较多个数据源来验证内容的准确性。

整合与反思阶段：每位小组成员使用在线协作白板工具，如小画桌，利用不同颜色的“便利贴”来简化表示 Toulmin 的论证结构（Stegmann et al., 2012），并以论证地图的形式将协同论证活动的结果进行可视化展示。在下次课上，每个小组现场展示论证地图，以确保对论证话题的理解。最后，小组成员对人机协同论证过程进行反思，尤其需要反思 ChatGPT 回复的内容是否能支持自己的论证，进一步理解生成式人工智能在支持人机协同论证方面的优势与缺陷。

3. 人机协同论证活动中的两种提示设计

本研究还强调提示语的设计对于促进有效的人机协同论证的重要性，从而突出其在对话互动中的中介作用。因此，本研究将重点放在提示的设计上。

修辞性论证提示的设计依据来源于社会建构主义与教育界的提示工程研究。社会建构主义强调知识和社会现实是多元和相对的，与社会和文化环境紧密相关(Lock& Strong, 2010) 。不同角色和身份的个体和群体可能受到其文化、社会背景、利益的影响，具有不同的认知风格、经验和专业知识，对某一领域问题的有不同的理解。通过考虑不同利益相关者的立场，与其进行协商与对话，可以建立更全面、多元、包容性的论证，在决策中考虑多方利益，创造更符合社会需求的解决方案，从而深刻地理解争议性主题与建构知识(Erduran& Jiménez-Aleixandre,2008) 。教育领域的提示工程研究发现，当要求 ChatGPT 讨论一个话题时，指定特定的角色或视角可能会有所帮助(Atlas, 2023). ChatGPT 可以模拟不同的对话角色，生成的输出将与该特定角色对提出的问题的典型反应保持一致(Hwang & Chen, 2023)。因此，本研究设计的观点生成提示将引导 ChatGPT 模拟与论证话题相关的利益相关者，为学习者提供与不同角色与身份的人的交流体验，从而刺激学生者从多个视角考虑论点。针对论证话题“你认为 AI 绘画是不是艺术，为什么？”本研究所设计的修辞性论证提示如表 1 所示。

表 1 修辞性论证提示设计

	观点生成提示
1	假如你是一名原创画家，请用专业口吻向我解释 AI 绘画是艺术（划线处可以替换你的观点，如： AI 绘画不是艺术）
2	假如你是一名研发 AI 绘画技术公司的员工，请用专业口吻向我解释……
3	假如你是一名绘画收藏家，请用专业口吻向我解释……
4	假如你是一名消费者，请用专业口吻向我解释……
5	假如你是一名意识形态与伦理研究者，请用专业口吻向我解释……
6	假如你是一名学习绘画的学生，请用专业口吻向我解释……
7	假如你是一名美术老师，请用专业口吻向我解释……
8	假如你是一名（其他你能想到的想要 AI 扮演的角色），请用专业口吻向我解释……

辩证性论证提示的设计依据为批判性问题模型(CQMAA)(Nussbaum & Edwards, 2011)。批包括支持论证的九个关键问题，如表 2 所示。学生运用这些关键问题思考其他论证主体的论点，可以判断其论证强度与说服力，从而帮助学生考虑是否接受他人立场并加强不同论点的整合。

表 2 批判性问题模型

维度	批判性问题
结构	论点中有主张吗？主张有证据支持吗？有证据到主张的推理吗？
意义	论点中的术语/概念是什么意思？能举例解释吗？
连贯的推理	推理的过程是否有可遵循的路径？是否有遗漏的步骤？
证据	论点有证据支持吗？证据是否可靠？
理由	理由中有不真实或不正确的吗？
准确性	该论点是否与公认的科学原理有关？
可选择性	这一证据是否有其他的解释或结论？能排除其他或相互竞争的解释或结论吗？

完整性	论证中是否有薄弱的地方？有没有遗漏的部分？
权衡	论证方案是否符合实际？

依据研究情境，本研究设计的辩证性论证提示如表 3 所示。

表 3 辩证性论证提示设计

	观点批判提示
1	我认为 AI 绘画是艺术（划线处可以替换你的观点）有哪些证据可以支持这一观点？
2	你提供的证据是否有参考来源？
3	你提供的证据是否有遗漏？
4	如果我不同意你的观点，你可以进一步解释你提供的证据吗？
5	请用具体例子来解释你的证据如何支持你的观点。
6	从以上证据到论点的推论过程是否正确？
7	我认为 AI 绘画是艺术（划线处可以替换你的观点），请反驳我的观点。
8	我认为 AI 绘画是艺术（划线处可以替换你的观点），请解释这一观点成立的限制条件是什么。

4. 人机协同论证中两种提示设计的实证比较

本研究运用准实验研究方法比较不同提示下人机协同论证活动对学生论证话语模式的影响。其中，对照组学生基于人机协同论证活动框架，运用 ChatGPT 进行自由的协同论证；实验组 1 学生基于人机协同论证活动框架，运用修辞性提示与 ChatGPT 进行协同论证；实验组 2 学生基于人机协同论证活动框架，运用辩证性提示与 ChatGPT 进行协同论证。实验结束后，比较三组论证对话在结构性与社会性维度上的差异。

4.1. 研究对象

研究对象是来自中国江苏某大学的教育学相关专业一年级研究生，共计 93 名。学生平均年龄为 23 岁，专业背景涉及多个学科，包括小学教育、心理健康教育、现代教育技术、语文教学、英语教学、音乐教学以及教育技术学。66 名学生参加了信息化教育理论与实践课程，7 名学生参加了教育技术学基本理论课程。两门课程均在 2023 年秋季学期开课，课时相同，由同一名教师授课。两门课程的教学内容有一部分是相同的，即要求学生通过协同论证感知与反思信息化教学实践中的争议性问题所有学生自愿组成 4-5 人的小组。其中，信息化教育理论与实践课程的 66 名学生被随机分配于对照组与实验组 1，其中对照组共 33 人，有 7 个小组；实验组 1 共 33 人，有 8 个小组。教育技术学基本理论课程的 28 名学生被分配于实验组 2，共 6 个小组。

4.2. 实验流程

本研究通过一个 7 周、共 29 小时的准实验来评估不同提示下人机协同论证活动的效果。图 2 为实验流程。在培训阶段，教师向学生介绍人机协同论证的流程，并提供关于如何使用 ChatGPT3.5 和在线协作白板工具小画桌的培训。在人机协同论证阶段，两个实验组和一个对照组在不同提示条件下进行人机协同论证活动。这一阶段包括教师讲授论证话题相关的背景知识、小组成员进行人机协同论证、以及论证地图的绘制，每轮活动持续 5 小时，共进行三轮。论证话题包括：“AI 绘画是否属于艺术？”、“虚拟场景与现实场景哪个更有利于学习？”以及“在中小学阶段使用个性化学习推荐系统的利弊”。在展示与点评阶段，每轮活动结束后，小组代表展示其论证地图，随后教师对展示进行点评。

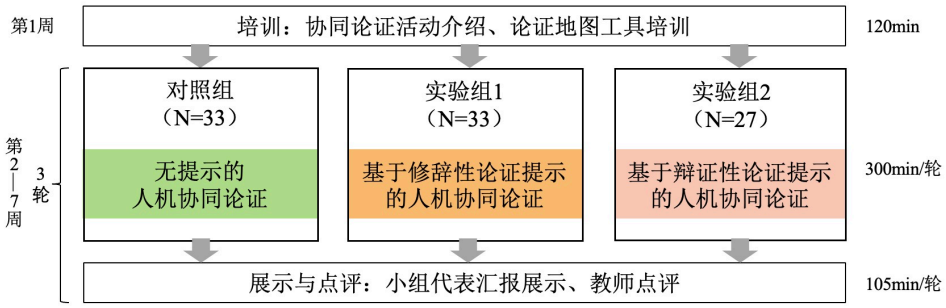


图 2 实验流程

4.3. 研究工具

本研究依据论证话语编码框架对学生的论证对话进行编码。该框架包含结构性与社会性两个维度，改编自(Darmawansah et al., 2022; Su et al., 2021; Chen et al., 2016) 。结构性维度包含主张（C）、证据（E）、推理（R）、反驳（CA）四个要素，每个要素下有三个由低到高的层级。如在主张要素中，低层级编码为 CC，即复述原材料或其他论证主体的主张；中层级编码为 CE，即简单提出自己的主张；高层级编码为 CD，即深入分析理由去支持自己的主张或在新证据的基础上对该主张进行改编。社会性维度包含 11 个要素，分别是提出自己的论点（P）、解释拓展自己的论点（E）、修改自己的观点（R）、询问他人的观点（A）、支持他人的观点（SP）、归纳总结他人的观点（SG）、指出他人论点的不足（C）、检查他人论据的有效性与可信度（CK）、说服他人接受自己的论点（D）、被他人激发出新的论点（SA）、将不同的论点整合为新的论点（I）。

4.4. 数据收集与分析

研究收集了每个小组企业微信的对话记录，共 1629 条。观察数据发现，每一条对话框中的文字内容都很丰富，最长的一条对话框包含 7 条句子。这可能是因为相比于面对面口头论证，文字通讯工具为所有学习者提供了充足的时间来构思论点、收集论据，一条对话框体现了学生完整的思考与表达，其中涉及多种论证要素，适合分析其共现关系。因此，分析者将一条对话框视作一个编码单元。依据 ENA 的编码表模板，按照编码框架标注其中出现的论证要素。由两名了解协同论证的研究者进行独立编码，当出现不确定的地方协商解决。编码结束后，运用 ENA 分析两组结构性要素与社会性要素的认知网络及其差异。

4.5. 实验结果

4.5.1. 结构性维度

三组认知网络分析如图 3 所示，对照组高层次主张、中高层次推理的共现关系较强（CD-RA-RE）。实验组 1 的高层次主张、中高层次推理共现关系较强（CD-RA-RE），且强于对照组。实验组 2 的高层次主张、中层次证据（CD-ES-RE）的共现关系明显强于前两组。

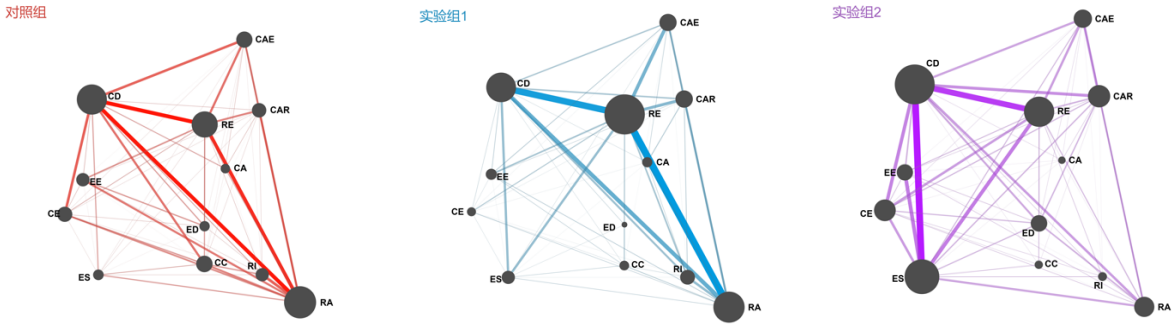


图 3 三个条件组的结构性论证要素认知网络

这说明，相比于无提示条件，修辞性论证提示下的人机协同论证活动使得学生能够提出较高质量的主张，并详细论证证据与主张之间的联系，有力地从证据推论到结论。而在辩证性论证提示下，学生能够从可靠来源中选择相关的证据支持自己的复杂主张。

4.5.1. 社会性维度

三组的认知网络分析如图 4 所示，对照组的提出论点-解释论点（P-E）共现强度较高，而实验组 1 在提出论点-解释论点-指出他人论点的不足-说服他人接受自己的论点（P-E-C-D）的共现关系更明显；实验组 2 在提出论点-解释论点-检查证据有效性-支持(P-E-A-CK-SP)的共现强度较高。

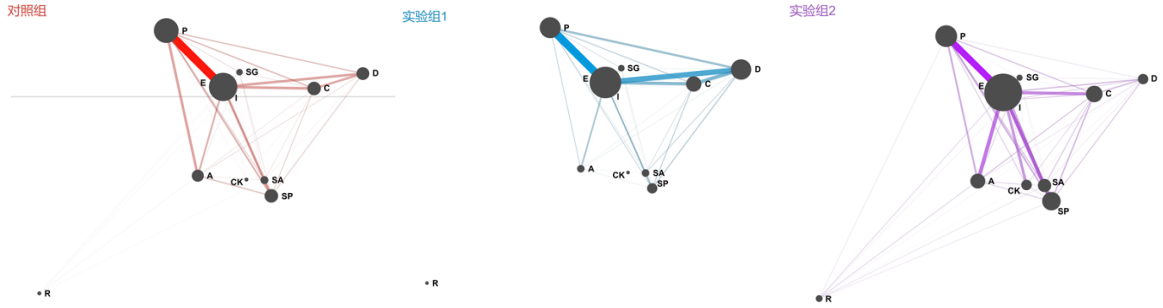


图 4 三个条件组的社会性论证要素认知网络

这说明，相比于无提示条件，在基于修辞性论证提示的人机协同论证活动中，学生除了论点论据的方式阐述自己的观点外，也会积极地反驳他人观点，来加强自身论证的合理性。而在辩证性论证提示下，会更加注重检查证据的有效性与可信度，来审查彼此提供的证据的真实性，并在此基础上支持、整合多方的论点。

5. 结语

本研究聚焦协同论证这一重要的教育场景，以引导论证性对话的提示设计原则为基础，结合人机对话的特点，设计修辞性提示与辨证性提示，比较不同提示下人机协同论证活动对学生论证话语模式的影响。研究发现，修辞性提示下的人机协同论证活动使得学习者注重从证据到主张的推理过程，且能够积极地反驳他人的观点，为自己的观点进行辩护；辩证性提示下的人机协同论证活动使得学习者更加注重证据来源的可靠性，且能够在审查彼此证据的的基础上达成共识。

本研究设计的提示只引发了一问一答式的单向互动，学生更期待与生成式 AI 进行双向的、多轮次的沟通，实现真正意义上的人机协同论证，未来研究可以继续关注引发双向互动的提示设计，让提示真正成为人机交互的中介。此外，本研究的实验周期有限，因此建议未来的研究在更长的时间内验证融合生成式 AI 的协同论证活动的机制与作用。

参考文献

Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI.* Retrieved from https://digitalcommons.uri.edu/cba_facpubs/548.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1), 1-18.

Chen, Y. C., Park, S., & Hand, B. (2016). Examining the use of talk and writing for students' development of scientific conceptual knowledge through constructing and critiquing arguments. Cognition and Instruction, 34(2), 100-147.

Clark,P.A. (2023, February 22). AI's rise generates new job title: Prompt engineer. Axios. https://www.axios.com/ 2023/02/22/chatgpt-promptengineers-ai-job.

Darmawansah, D., Lin, C. J., & Hwang, G. J. (2022). Empowering the collective reflection-based argumentation mapping strategy to enhance students' argumentative speaking. Computers & Education, 184, 104516.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). Argumentation in science education. Perspectives from classroom-Based Research. Dordre-cht: Springer.

GRAILE AI(2023). Conference Opening and Keynote: Embracing a culture of innovation in the Era of AI[EB/OL]. [2023-10-23]. https://www.youtube.com/watch?v=C-CQqcsZvgA.

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. Computers & Education, 104862.

Harney, O. M., Hogan, M. J., Broome, B., Hall, T., Ryan, C. (2015) Investigating the effects of prompts on argumentation style consensus and perceived efficacy in collaborative learning. International Journal of Computer-Supported Collaborative Learning, 10(4), 367–394.

Hwang, G. J., & Chen, N.S. (2023). Editorial Position Paper: Exploring the Potential of Generative Artificial Intelligence in Education: Applications, Challenges, and Future Research Directions. Educational Technology & Society, 26(2):1-14.

Linn, M. C., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. Science, 313(5790), 1049-1050.

Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. The Journal of Academic Librarianship, 49(4), 102720.

Malogianni, C., Luo, T., Stefaniak, J., & Eckhoff, A. (2021). An exploration of the relationship between argumentative prompts and depth to elicit alternative positions in ill-structured problem solving. Educational Technology Research and Development, 69, 2353-2375.

Nussbaum, E. M., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. Journal of the Learning Sciences, 20(3), 443–488.

Rapanta, C., & Felton, M. K. (2022). Learning to argue through dialogue: A review of instructional approaches. Educational Psychology Review, 1-33.

Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. International Journal of Human–Computer Interaction, 37(1), 81-96.

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. Instructional Science, 40, 297-323.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. Assessing Writing, 57, 100752.

Su, Y., Liu, K., Lai, C., & Jin, T. (2021). The progression of collaborative argumentation among English learners: A qualitative study. System, 98, 102471.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes and outcomes. Educational Evaluation and Policy Analysis, 17(2), 239–261.

四维共筑：普及创客、人工智能教育之平阳行动

One Point Four Sides: Pingyang Action for Popularizing Creativity and Artificial Intelligence Education

谢贤晓 1*
[1] 平阳县教师发展中心
14405612@qq.com

【摘要】 温州市教育局 2015 年发文《温州市中小学创客教育实施方案》和 2020 年温州市教育局发布了《温州市中小学推进人工智能教育实施方案》，我县高度重视本项工作，区域整体推进，多部门通力合作，为全面普及创客、人工智能教育进行整体规划，本文通过四大模块：基础化、多元化的课程研发；专题化、系列化的师资培训；结实性、通用性的硬件、平台选择；实用性、多功能化的实验室创建分享平阳行动。

【关键词】创客；人工智能；区域；教学实践

Abstract： *Wenzhou Municipal Bureau of Education in 2015 issued the "Wenzhou City, primary and secondary schools to implement the program of creativity education" and 2020 Wenzhou Municipal Bureau of Education issued the "Wenzhou City, primary and secondary schools to promote the implementation of artificial intelligence education," the county attaches great importance to the work of this item, the region as a whole to promote the overall cooperation of multiple departments for the full popularization of creativity, artificial intelligence education for the overall planning, this paper through the four major modules: foundational, diversified Curriculum research and development; thematic, series of teacher training; sturdy, universal hardware, platform selection; practical, multi-functional labs created to share the Pingyang action.*

Keywords： *Creativity; Artificial Intelligence; Regional; Teaching Practices*

1. 前言

近年来，国务院和教育部相继发布《新一代人工智能发展规划》和《教育信息化 2.0 行动计划》等重磅文件，在 2024 年十四届全国人大二次会议上的政府工作报告中提出要“大力发展数字教育、推进教育现代化、建设高素质专业化教师队伍和教育强国建设”，同时提出人工智能+行动，大力发展人工智能已成为了国家战略，人工智能已经成为推动社会进步的重要力量。熊璋教授在甘肃的报告《建设信息科技课程，推动人工智能赋能》中也指出人工智能赋能教育新生态时不我待，中国的教育要掌握在自己的手中，我们要做出有中国特色的教育，作个性化的教育，作公平教育。

温州市教育局先后发布了《温州市中小学创客教育实施方案》和 2020 年发布了《温州市中小学推进人工智能教育实施方案》，我县早于 2015 年已经开始实施创客教育，与 2015 年成功承办了温州市首期创客研讨会。在新趋势下，从创客教育的普及过渡到人工智能教育的普及还面临着许多问题，由于平阳县是浙江省山区 26 县之一，资金紧缺；新技术新平台新硬件更新换代频繁，急需匹配的课程作为教师教学的脚手架；教师的专业能力有待提高；适合的教学环境建设等。基于此，在本文中对区域如何通过课程、师资、硬件和实验室四个模块的十年行动进行探讨。

2. 多部门联动 整体推进

2.1. 研发基础化、多元化课程

新技术的推陈出新，开源硬件和编程软件的更新换代，要普及创客、人工智能教育，使创客、人工智能教育真正全面扎实有效落地于中小学课堂，需要为教师们提供简单易用的课程，作为他们实施教学的“脚手架“。我县从 2015 年开始组建团队，研发系列创客、STEAM 和人工智能课程，并通过开源的方式分享给全国一线教师。截止至今，区域已经完成了多个开源课程并推广，出版的分别是：《Arduino 创意机器人——基于 Mixly》、《Mind+和掌控板互

动创意设计》、《Arduino 创意机器人——基于 Mind+》。已经网络开源分享的有《趣味物联网作品设计》和《物联网和科学探究创意实验》，每个课程都开发了各种课程资源，教学设计、微课程和程序全部采用 CC-BY 协议开源，方便一线教师开展教学。

2021 年，区域小学信息科技、创客教研员担任“全国中小学人工智能教育培育区（校）培育计划”课程开发专家组组长，组建了一支来自全国各地一线教师团队，研发了中小学人工智能微课程，目前共研发完成中小学人工智能微课程 15 个。2021 年，区域小学信息科技、创客教研员担任清华大学出版社《信息科技》教材小学段副主编，撰写五、六年级的开源硬件、物联网和人工智能共 4 册，清华版教材六年级上下两册人工智能书籍，已经与 2022 年上半年正式出版发行。

研发的课程从单一到多样性，从创客课程的研发到 STEAM 课程、再到人工智能课程的研发，参与研发的教师从单一的信息科技老师到多学科老师的融合。每个课程的研发都经历了前期、中期和后期。前期：选定物美价廉的硬件和通用的编程软件，对接硬件公司，挑选硬件和罗列设备清单，供课程研发使用；做好课程定位和撰写课程纲要，完成团队人员分工；中期完成课程电子稿、微视频、教学设计和源代码，针对编程软件中出现的 bug 和需要增加的功能模块，积极主动地与软件研发组的工程师们进行沟通；后期通过十几所试点校的一个学期教学不断完善课程及相关内容。



图 1 已经出版的相关书籍

2.2. 做好专题化、系列化的师资培训

为提高全县小学信息科技教师的教学能力和专业能力，组织系列、专题的相关培训，培训采用线上线下双管齐下的方式进行，一方面通过专家讲座与实践操作，了解创客、人工智能概念，学习最前沿成果，拓展学员视野，学会使用开源硬件和软件进行造物。另一方面，通过课堂教学观摩、案例分享等实践活动，提高教师的教学设计能力。每学期组织一次这样的专题培训，此项培训已经坚持了十年左右，从最早的 Scratch 到最新的新一代人工智能浦育平台，我们组织了将近 20 期的专题培训。从最初的基于硬件和编程平台的专题培训到融合了学科教材教学的专题培训，下图是其中一次一次基于教材的人工智能专题培训的课程表：

专题	课程（专题）名称	学时	授课教师	授课形式
理论学习提升	《核心素养背景下的单元整体教学设计与实践》	4	冯京和	专题讲座
	《信息技术单元设计与实施》	4	黄菊敏	专题讲座
技术操作指导	《面向新一代人工智能技术的 AI 科创活动》	4	谢作如	专题讲座 实践操作
	《基于开源硬件的人工智能课程开发和实践》	4	谢贤晓	专题讲座 实践操作
名师团队展示	《基于项目化学习的单元活动设计》	4	黄莹莹	专题发言
	《物联网项目实践——测温仪》		谢珍视	课堂观摩
	《垃圾分类机器人》		温小敏	课堂观摩
	《有趣的班级管理员》		屠盈盈	课堂观摩
学员展示交流	《初识人工智能》教材解读与实施建议	4	郑书琴	专题发言
	试课展示《走进人工智能》		许一炜	课堂观摩
	课例展示《文字识别》		吴宁静	课堂观摩
	课例展示《人脸识别》		钟云晴	课堂观摩

图 2 平阳县小学信息科技人工智能模块的单元整体教学设计与实践专题培训课程表

组织学员开展线上研讨活动。线上的前期准备工作是将学员分成四个小组，引导每个小组认真研读教材六年级上册第二单元《初识人工智能》、六年级下册第一单元《万物互联》、第二单元《与机器人零距离》进行单元整体教学设计，确定项目主题。县名师单独组成一个团队，针对教材六年级上册第二单元《初识人工智能》进行单元整体教学设计并展示。

线下的三天集中培训。通过专家讲座，深入了解人工智能；通过市教研员分享，知道单元整体教学设计的一般思路和方法；通过县名师团队的单元整体教学设计的呈现，供学员学习。最后是培训学员的成果展示，每个小组根据市综合学科单元整体教学设计评审，提供的单元设计框架撰写教学设计；通过专家评选，从四组中选择一组进行半天的课堂展示，小组成员选出一名做教材单元整体解读和实施建议，小组成员通过展示课、说课、试课、教学设计等方式展示，选出一名学进行单元整体教学设计的专题发言。

与此同时，我县也阶段性的组织面向教师的创客马拉松大赛，以学校为单位，多学科老师参与，组建队伍，进行为其 6 小时的现场创作。这种活动能展现参与者的科学素养、创新精神和实践能力，三位评委全程参与各组的指导，确保每支队伍都能够在规定时间内完成作品的创作，进而提高老师的创客素养。



图 3 创客马拉松现场照片

2.3. 选用结实性、通用性的硬件和平台

普及创客、人工智能教育需要为各学校配置相应的硬件，县装备中心联合县教师发展中心，根据教学需要罗列硬件清单。从 2015 年开始，我们联系硬件公司，从最初选定硬件设备用于教学到现在能够根据教学需要，让公司设计硬件设备。我县选用通用性强的 Arduino 板、掌控版、Microbit 板、行空板等开源硬件，激光切割选择雷宇和 Makerblock 的 Xtool，在确定学校硬件清单前，都会先亲自使用，根据其结实性和通用性再进行选定。确保每个学校在开展创客、人工智能教育时有充足的设备来支撑教学，为了更好的普及创客、人工智能教学，我们联合公司设计基于开源硬件的一体机设备，便于大班教学，让每个孩子都能够体验、创造。



图 4 基于掌控板的掌控小车

图 4 基于掌控板的掌控小车是在研发清华版小学信息科技教材时，根据课程需要，联合

盛思公司研发的一体机设备，便于在常规信息科技课堂中，让孩子们进行创作。2022 年新课标颁布，信息科技新课标中，增加了实验教学，为了更好的展开实验教学，我县联合 DF 公司，正在研发基于行空板的一体机设备，计划在 2024 年的暑假将此设备纳入人工智能教育的购买清单，为今后的实验教学提供更好的硬件保障。

除了硬件选择外，我县在平台的选择上，也是与时俱进，从最初的易乐谷（海归画图），通用性很强的 Scratch、Mind+、mPython，再到浦育平台。我县不仅做通识培训，而且在课程研发的过程中，多次对接平台研发人员，针对教学需要和平台存在的 BUG 进行沟通，完善了平台，从界面到指令，从菜单规划到功能增加，我们一直在坚持。

2.4. 创建实用性、多功能化的实验室

为了让学生更好的进行体验、创作，需要在各校建立创客空间、人工智能实验室、信息科技实验室等实验室，实验室需要拥有先进的设备和资源，以满足各类实验需求，包括计算机、服务器、软件开发工具等硬件与软件设备。由于我县是浙江省山区 26 县之一，经费紧张，各校在创建实验室时，选择了资金投入少、实用性强、能满足多学科教学需要的实验室。部分学校由于教室紧张，没有教室可以用于实验室创建，建议学校给机房增加一些便于硬件摆放的架子，在机房展开创客、人工智能教学。



图 5 某校在机房展开掌控板教学

3. 多平台搭建 促师生成长

为进一步提升教师的信息素养，促进教师专业成长，提高教师综合素质，加强教师队伍建设，我县为一线教师提供了各级各类平台。带领团队多次参加全国中小学 STEAM 教育大会的工作坊工作，我县作为清华版教材《人工智能》试教学点，共有昆阳实验小学、鳌江一小等 10 所学校参与试点其中。平阳县昆阳二小等 5 所学校被评为全国中小学人工智能教育培育校。带领团队申报省、市课题，省课题《基于物联网的 STEAM 课程研发与实践》《基于项目化的人工智能课程研发与实践》结题，《“小学 Arduino 创意机器人”拓展课程建设的研究》和《基于“App Inventor”的中小学创客课程建设研究》获市一二等奖。

我县教师积极带领学生参加创客、人工智能各项评比，在 2023 年全国青少年人工智能创新实践活动中，我县 4 所学校入选全国青少年人工智能创新教育学校，21 各作品被评委优秀作品。通过现场激烈的比拼，我县学生在 AI 交互设计荣获五星卓越和四星卓越称号。

4. 结语

通过十年努力，我县创客、人工智能教育得到了多方认可，温州日报发文《研训和课程开发，让平阳创客教育领跑全市》，温州教育发布《研发课程，培育教师！平阳的创客教育之路》和《基于掌控板的项目智造案例—以“校园智能广播播报系统”为例》。在刚结束的第十一届中小学 STEAM 教育大会，大会揭晓 2023 年度创客教育创新人物评选结果，我县信息科技、创客教研员谢老师凭借其在创客教育领域的杰出贡献和影响力，成功获得提名。颁奖词写道：“她是一名行动力很强的教研员，积极带动身边的信息科技教师，打造了一支全国最强大的创客团队，她所开发的 Arduino 创意机器人等系列课程，影响了很多人。”这段颁奖词充分肯定了我县创客、人工智能教育。

展望未来，平阳县将秉持这种精神，进一步深化人工智能教育，推动青少年人工智能教育的普及开展，致力于培育更多具备创新能力和信息素养的未来科技英才。

参考文献

[1]温州市教育局.市教育局发布《温州市中小学推进人工智能教育实施方案》[DB/OL].[2022-03-09].https://www.wenzhou.gov.cn/art/2022/3/9/art_1217833_59154015.html

[2]梁森山,王阳,刘军,等.寻味人工智能教育:祛魅之下的实践与思考[J].中国信息技术教育,2018[1](02):4-11.

[3]李永智.新一代人工智能技术对教育的影响[J].教育参考,2023(04):2-3.

[4]张珊珊,杜晓敏,张安然.中小学开展人工智能教育的挑战、重点和策略[J].中国电化教育.2020(11):67-72,96.

专业标准视域下的中小学人工智能教育教师培训课程设计

Design of artificial intelligence education teacher training course in primary and secondary schools under the perspective of professional standard

夏正仁[1]
[1]温州市教师教育院

【摘要】国家提出在中小学阶段设置人工智能教育相关课程的要求，各地却面临缺乏相应师资的难题。对学科教师进行专题培训，被认为是主要的解决途径。但当前的师资培训大多不符教师专业标准的要求，存在过于注重专业知识传授而忽略专业能力与专业精神培育的问题。认为应该关注培训课程设计的科学性，要按照中小学人工智能教育课程实施与教师自身专业发展的实际需求来确定培训目标与内容，要参照教师专业标准来划分培训的内容模块，要依照教师核心素养形成路径来设计培训的活动与方式。
【关键词】人工智能教育；教师专业标准；教师核心素养；培训课程设计

Abstract: *The state puts forward the requirement of setting up courses related to AI education in primary and secondary schools, but all places are faced with the problem of lack of corresponding teachers. Special training of subject teachers is considered to be the main solution. However, most of the current teacher training does not meet the requirements of teachers' professional standards, and there are some problems of paying too much attention to the teaching of professional knowledge while ignoring the cultivation of professional ability and professional spirit. Think should focus on the scientific nature of the training course design, according to the artificial intelligence education curriculum implementation and the actual needs of teachers 'professional development to determine the training objectives and content, to refer to teachers' professional standards to divide the content of the training module, according to the teacher core literacy formation path to design training activities and ways.*
Keywords: Artificial intelligence education, teacher professional standard, teacher core quality, training course design

1. 引言

《新一代人工智能发展规划（国发〔2017〕35 号文件）》提出“实施全民智能教育项目，在中小学阶段设置人工智能相关课程，逐步推广编程教育”要求。《高中信息技术课程标准（2017 版）》将其列为选择性必修课程，并在必修课程中开设通识性教育。《义务教育信息科技课程标准（2022 版）》则将其列为六大逻辑主线之一，并将教学安排在第四学段。另外《教育部关于加强和改进中小学实验教学的意见（教基〔2019〕16 号文件）》提出“注重加强实验教学与多学科融合教育、编程教育、创客教育、人工智能教育、社会实践等有机融合，有条件的地区可以开发地方课程和校本课程”任务。

综上所述，在中小学校开设人工智能教育，主要有两大目的。一是提升全体学生的智能素养，以适应智能社会的发展；二是激发部分学生的学习兴趣，来为国家培育潜在专业人才。与前者对应的课程，主要有高中信息技术的必修课程和初中信息科技课程；与后者对应的课程，主要是高中信息技术的选择性必修课程与包含融合教育在内的各类中小学拓展课程。由于绝大部分学科教师缺乏人工智能专业知识，通过培训途径来提升专业认知，就自然成为相关部门的主要选择。但当前的培训课程设计，大多不符教师专业标准的要求，存在偏重专业知识传授而忽略教师专业能力与专业精神培育的问题。

教师专业标准是指教育部颁发的《中小学幼儿园教师专业标准（试行）》（以下简称专业标准），它“是国家对合格教师的基本专业要求，是引领教师

专业发展的基本准则，也是教师培训的重要依据”。培训课程设计则认为是根据培训需求与培训目标，来对课程内容、培训方法与活动安排等方面进行综合设计的过程。

本文从专业标准视角出发，来探讨中小学人工智能教育学科教师培训课程设计的科学性，这对确保培训质量、促进教师专业能力发展与推动中小学人工智能教育普及等具有积极的意义。

2. 相关研究现状

与培训课程设计相关文献除了专业标准之外，还有教育部《关于大力推行中小学教师培训学分管理的指导意见（教师〔2016〕12 号）》（以下简称指导意见）、教师核心素养观、课程设计论、成人学习理论与中国教育科学研究院领头研制的《中小学人工智能教育教师能力标准》（以下简称能力标准）等。它们的重要观点与指导作用，综述如下。

指导意见提出“分层提供教师培训课程，强化培训内容的针对性和系统性”“根据教师年度发展和周期性发展需求，进行递进式设计，推动教师持续成长”。该文件为培训的长期规划与分层设计，提供了政策层面的支持。

专业标准将教师专业划分为“专业理念与师德”“专业知识”与“专业能力”这三个领域。“专业理念与师德”包含职业理解与认识、对学生的态度与行为、教育教学的态度与行为、个人修养与行为；“专业知识”包含教育知识、学科知识、学科教学知识、通识性知识；“专业能力”包含教学设计、教学实施、班级管理与教育活动、教育教学管理与评价、沟通合作、反思与发展。该文件可为培训内容选择及其模块划分提供了参照，至于各领域中教师已具备或者是弱相关的内容，则可以根据实际情况进行去留。

教师核心素养还没有统一的定义，但已被认为是影响教师专业成长的关键。杨志成认为“是教师在职前学习和职后实践与培训过程中逐步形成的，满足教育教学和学生发展需要的必备品格和关键能力”。本文则在其基础上，再结合学生核心素养的概念，认为它由正确价值观念、必备品格与关键能力三个维度组成，各维度素养形成的具体路径如下图 1 所示。它可为培训活动及其方式设计，提供理论方面的指导。

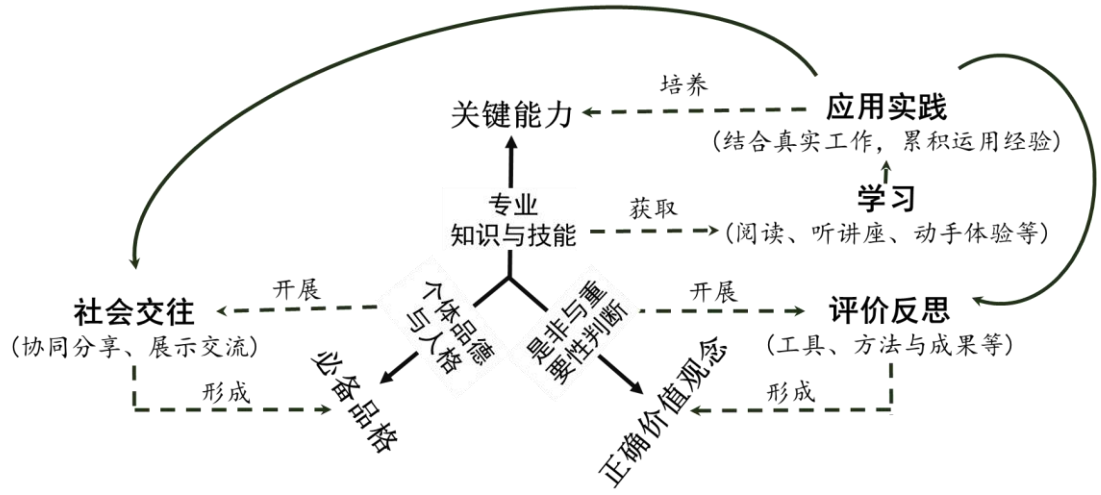


图 1 教师核心素养及其形成路径

课程设计论认为“几乎所有的课程设计，都可归为三种模式：学科为中心、

学生为中心、问题为中心”。因此可为培训课程设计，提供模式方面的指导。其中专业知识领域的设计，较适合选用学科为中心模式；专业理念与师德领域的设计，较适合选用学生为中心模式；专业能力领域的设计，较适合选用问题为中心模式。

成人学习理论认为成人学习具有“能够自我指导学习、丰富的经验、以问题为中心及内在动机驱动”等特征。这为选择以解决具体的教育教学问题作为培训主题与选用任务驱动作为培训模式，提供了理论的依据。

能力标准对学科教师提出“人工智能理解与意识、基本知识、基本技能、问题解决、教学实践、伦理与安全”六大要求，并详细描述了具体的内涵。因此该文件可用作对学员专业能力的评价工具，以帮助学员进行自我测评。

3. 教师培训课程的设计

学科教师所面临的专业挑战，主要有人工智能科技的迭代、教育教学环境的创建升级、拓展课程的开发、育人理念的更新与教育教学方式的变化等方面。因此培训课程设计，要考虑需求的针对性、内容的全面性与活动的科学性。

3.1. 按照实际需求设计培训目标与内容

培训需求是目标制定与内容筛选的重要依据。教师参与培训的动机，主要分为内外两种不同因素。外因是为了满足国家提出普及人工智能教育的工作要求，内因则是为了提升自身专业水平而发自内心的学习需要。因此培训存在以下两种基本需求：一是面向全体学科教师，为实现普及人工智能教育而开展的培训，目标是要求学员具备胜任高中信息技术或初中信息科技课程教学即可，内容与国家课程紧密相关，通常以学科的基本概念或原理、教学设计与活动实施为主；二是面向部分骨干教师，为满足他们深入研究人工智能教育而开展的培训，目标是除了要求胜任拓展课程教学之外，可能还要具备校本课程开发等方面的能力，内容通常紧跟学科发展前沿，并注重对专业能力的引领提升。

3.2. 参照专业标准划分培训的内容模块

培训内容设计要涵盖学科知识、学科教学与教师发展。因此可以参照专业标准所提倡的三个领域，即将培训内容划分为专业精神、专业知识与专业能力这三个模块。其中专业精神主要体现在学科教师对开展人工智能教育而应该持有的工作态度与师德表现；专业知识主要是指人工智能专业的知识与技能、开展人工智能教育教学所需的知识与技能；专业能力则主要体现在运用人工智能专业知识与技能来解决真实日常生活问题的能力、开展人工智能教育教学及其研究的工作能力。

3.3. 依照素养的形成路径设计培训活动

学员习得专业知识、发展专业能力与形成专业精神，都是通过相应培训活动才得以实现。对此可以借鉴核心素养形成路径，来设计规划培训活动，如下图 2 所示。通过阅读或听讲方式来获取专业知识，通过动手体验方式来掌握专业技能；通过真实开发与实践检验等途径，来帮助学员累积知识技能在真实复杂情境中的灵活运用经验，进而发展成为专业能力。通过对所开发的作品、运用的方法或工具等进行评价反思，进行经验总结与成果提炼并进行推广交流，进而激发专业精神。

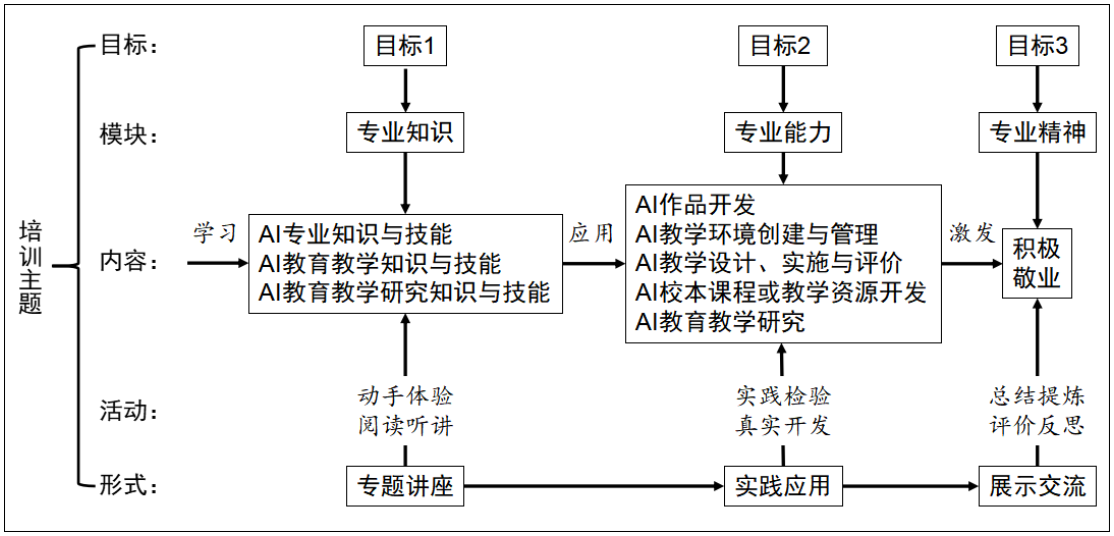


图 2 培训目标、内容模块与活动形式之间的关系

4. 培训成效与问题反思

培训成效通常要经历一段时间的酝酿，才会显现它的真实价值。温州市教师教育院自 2017 年 7 月起，就陆续对市域内的学科教师开展人工智能教育系列专题培训。培训起初是缘于创客教育的创新需要，后期则是基于国家课程的实施要求。

4.1. 应用成效分析

通过对温州市中小学人工智能教育情况调研和对近几年来学科优质课、展示课、教学设计与论文评比等活动的观察，本文认为自从实施基于教师专业标准的培训之后，主要取得以下两方面的成效。

一是成为全国最早开展中小学人工智能教育地区之一。主要原因在于前瞻性的视野与专业培训的引领，让学科教师有了正确的专业发展方向与成长途径，进而带动温州市中小学人工智能教育的开展。因为培训较早解决了“教什么”与“怎么教”的问题，所以就有了早期以创客教育为主的校本拓展课程，以及后来国家课程实施储备一定的师资力量。

二是触发了学科教师的关注与相关教育主管部门的重视。主要表现在各级各类学科教研活动中，人工智能教育一直是较为热门的主题；人工智能教学实验室开始成为一些学校的新基建项目；相关教育主管部门，后来推出了配套的区域推进政策。这是因为培训告知了“当前缺什么”问题，所以就带来了一波的研究热、创建热与推进热。

4.2. 存在问题与反思

基于专业标准的教师培训课程设计，也会存在“理想与现实脱节”问题，即与实际发生的课程会有一定的出入。问题产生的主要原因与优化思考，整理归纳如下。

一是受限于上课专家。首先在人工智能教育发展初期阶段，普遍存在专家的稀缺难题。其次专家上课大多各自为战，很少顾及他人，也没深入理解课程设计者的初衷。因此要加强课程设计者与专家的沟通，要为他们提供彼此交流渠道，以确保培训的连贯与衔接。

二是受限于培训环境。学员学习人工智能专业知识，也需要算力、算法与数据等学科要素的支持。但在当前起步阶段，一般师训机构难以提供这样的环

境。因此要加强与相关企业的合作，借用他们力量与条件来提升学员的学习感受，进而提升学科教师教训的质量。

5. 结束语

探讨培训课程设计这个话题，旨在交流师资培训的科学性与规范性，进而促进学科教师的专业发展，推动中小学人工智能教育的普及。但由于自身能力不足，以及存在对教师专业标准的理解偏差，文中想法难免会有不妥之处，敬请大家的批评与指正。

参考文献

教育部.普通高中信息技术课程标准（2017 年版）[M].**人民教育出版社**,2018.

中华人民共和国教育部.义务教育信息科技课程标准《2022 年版)[S].**北京:师范大学出版社**.2022

侯丽娜. (2015). 教师专业标准的研究现状与实施建议. **大连教育学院学报**, 31(1), 9-11.

杨志成. (2017). 中国中小学教师发展核心素养体系建构研究. **Teacher Development Research**, 1(1).

程明喜. (2017). 诺尔斯成人学习理论及对我国中小学教师培训的启示. 成人教育, 37(12), 75-78.

戚业国, & 刘光亮. (1995). 当代西方大学课程设计模式评析. **上海高教研究**, (4), 69-72.

校地合作背景下中小学人工智能教育实施路径探索与实践

——以温州大学附属茶山实验中学为例

Exploration and Practice of Artificial Intelligence Education in Primary and Secondary Schools under the Background of School-local Cooperation

-- Take Chashan Experimental High School Affiliated to Wenzhou University as an Example

郭晓戈[1*]，涂芸芳[1]，方建文[1]，胡开胜[1]，梁晨[1]
[1]温州大学教师教育学院
[*]1342777584@qq.com

【摘要】 人工智能技术正在快速发展，对教育领域产生了重大影响。中小学人工智能教育是培养未来人才的重要举措，需要校地合作，构建协同育人的实施路径。本文分析了中小学人工智能教育存在的问题，旨在探索高校和地方中小学校合作的模式，以便让人工智能教育在中小学更好地落地，推进人工智能教育的普及和发展。
【关键词】 人工智能教育；校地合作；中小学

Abstract： *Artificial intelligence technology is developing rapidly and has a significant impact on the education sector. Artificial intelligence education in primary and secondary schools is an important measure to train future talents, which requires school-local cooperation to build a implementation path of collaborative education. This paper analyzes the problems existing in artificial intelligence education in primary and secondary schools, aiming to explore the mode of cooperation between universities and local primary and secondary schools, so as to better implement artificial intelligence education in primary and secondary schools and promote the popularization and development of artificial intelligence education.*
Keywords： Artificial intelligence education, School-local cooperation, Primary and secondary schools

1. 引言

人工智能作为引领未来的新兴战略性技术，已成为国际竞争和产业变革的焦点。新技术的变革和竞争主要是人才的竞争，人工智能的发展离不开优秀人才的支撑。在我国，中小学开展人工智能教育已经成为了教育改革的重要方向之一。2017 年，国务院印发《新一代人工智能发展规划》，明确提出"在中小学开设人工智能课程"（国务院，2017）；2018 年 4 月教育部印发《高等学校人工智能创新行动计划》，提出“构建人工智能多层次教育体系，在中小学阶段引入人工智能普及教育”（教育部, 2018）；2022 年，教育部印发的《义务教育信息科技课程标准(2022 年版)》提出要将人工智能作为课程内容（中华人民共和国教育部，2022)。在这些政策的推动下，人工智能教育逐渐融入中小学课程体系，为学生们开辟了更广阔的学习领域。

在中小学开展人工智能教育，一方面可以培养学生的人工智能素养，使学生理解智能时代的各种新技术、新手段、新应用、新发展，更好地适应社会发展的需要；另一方面可以识别发现潜质学生，做好拔尖创新人才的早期培养，可以让学生认识人工智能技术带来的前所未有的兆战，看到自己的差距，培养勇于追求、不断创新的进取精神和使命感。

然而，目前人工智能教育在地方中小学的普及程度还远远不够，面对这一新兴领域，地方中小学在课程设置、教学实施和师资培训等方面也都存在诸多挑战，而高校则具备丰富的教育资源、师资培训能力和研究经验。因此，本研究旨在探索高校和地方中小学校合作的模式，以便让人工智能教育在中小学更好地落地，推进人工智能教育的普及和发展。

2. 中小学人工智能教育现状

随着人工智能技术在企业界取得成功，引发了社会对中小学人工智能教育关注。人工智能技术引领了当下产业变革，国家战略和教育政策文件相继出台。信息科技学科新课标也提出了中小学阶段开展人工智能教育的相关要求。然而，温州大学人工智能教育研究团队尽管有这些要求，但真正落地执行的力度仍然不够，大多数中小学的人工智能教育实施存在不全面、不充分的问题。

2.1. 课程设置

课程设置方面存在诸多挑战。目前中小学人工智能教育的课程设置相对薄弱，缺乏系统性和实践性。教育部门在课程设计方面缺乏统一规划和指导，学校普遍缺乏人工智能教育专业化的教学团队和课程资源，教学内容和方法相对单一，难以满足学生多样化的学习需求。

2.2. 师资力量

师资力量也是一个亟待解决的问题。大多数教师缺乏人工智能领域的专业知识和教学经验，难以胜任人工智能教育的教学任务。教师队伍中缺乏专业人才和培训机会，限制了人工智能教育的深入开展。

2.3. 教学资源

教学资源方面存在着明显的不均衡现象。一些学校缺乏先进的教学设备和技术支持，导致人工智能教育的质量和水平参差不齐。尤其是在偏远地区和经济欠发达地区，教学资源的匮乏程度更加显著，制约了人工智能教育的普及和发展。

2.4. 学生 AI 素养评估

当前中小学阶段对于人工智能素养的评价指标相对匮乏，这主要是因为人工智能领域的快速发展导致了教育与评价体系的滞后。传统的学科评价指标难以全面覆盖人工智能素养的各个方面，因此需要通过与高校等合作，借鉴其在人工智能领域的先进经验，构建适用于中小学生的评价指标体系。

3. 温州大学附属茶山实验中学人工智能教育实施路径探索和实践

学者认为，在高校与地方中小学的合作下，双方可以实现互利共赢。在这种校地合作背景下，探索中小学人工智能教育的实施路径与实践非常重要。高校可以通过借助地方中小学的实践平台，将理论知识转化为实际技能，从而提高教育的针对性和实效性。同时，地方中小学也可以通过高校的资源共享和专业指导，进一步提升自身的教育质量和教学效果。这种合作模式不仅有助于推动中学人工智能教育的普及和发展，也有助于实现教育资源的优化配置。鉴于过去的实践，主要的实施路径包括以下几个方面：

3.1. 建立校地合作机制

校地合作是教育领域的一项重要机制，旨在促进学校与地方社区之间的紧密合作，共同推动教育事业的发展。近年来，相关政策和倡议强调了加强校地合作的必要性，以促进教育改革与创新，提高教育质量和效益。在校地合作的框架下，学校与地方中小学开展合作，为人工智能教育带来了显著优势。中小学通过与高校合作，能够获取专业支持和资源优势，包括先进的教学理念、方法和资源。尤其是在人工智能教育领域，高校拥有丰富的学术研究成果和专业教学团队，能够为中小学提供前沿的知识和实践指导。通过与中小学的合作，高校能够将学术研究成果转化为实际教学内容，并将先进的教学理念和方法应用于实践中。这有助于提升高校的社会影响力和声誉，同时为高校教师提供了教学实践和经验积累的机会，促进了教学水平的提升。因此，校地合作被视为一种有效的途径，推动中小学人工智能教育的发展。这不仅有助于提升中小学教育水平，还为高校的教学和科研工作带来了新的发展机遇。

为了推动新一代人工智能教育在温州的中小学中落地生根并茁壮发展，温州大学人工智

能教育研究团队着手建立了校地合作机制。首先，成立了温州大学与茶山实验中学校地合作领导小组，由学校领导、教师代表和地方政府教育部门组成，负责统筹协调合作事务。其次，温州大学人工智能教育研究团队制定了详细的校地合作协议，明确了合作内容、方式、期限和责任，重点围绕人工智能教育的课程开发、师资培训、实践活动等方面展开合作。最后，搭建了温州大学与茶山实验中学校地合作平台，提供在线资源共享、课程交流和教学指导的便利，以促进双方的互动与合作(如图 1)。



图 1 温州大学与茶山实验中学授牌仪式

3.2. 开设人工智能课程

在中小学教育领域，人工智能课程体系的建设与实施一直存在着诸多挑战。中小学缺乏完整的人工智能课程体系，同时也缺乏与之配套的案例资源等教学支持。相较之下，高校的研究团队拥有丰富的学术资源和科研成果，具备设计完整的人工智能课程体系的能力。然而，这些课程体系往往缺乏与实践相结合的机会，这就导致了理论与实践之间的脱节。

为了解决这一问题，校地合作成为了一种有效的解决途径。温州大学的人工智能教育研究团队通过专业的人工智能培训和研究学习，整合梳理了丰富的教学资源，开发了面向中小学的人工智能课程（包括通识课程和社团课程）。这些课程不仅在理论层面上严谨完整，而且在实践操作方面也得到了充分考虑。通过与温州大学附属茶山中学的合作，研究团队能够为中小学提供定制化的人工智能课程，满足学校的实际需求，课程如图 2。同时，学校教师和学生也能够借助这一合作平台获得更多的实践机会和指导，从而提升课程的实效性和可操作性。通过这种校地合作的模式，高校研究团队的学术资源得以与中小学的实践需求相结合，为人工智能课程的建设与实施提供了有力的支持。

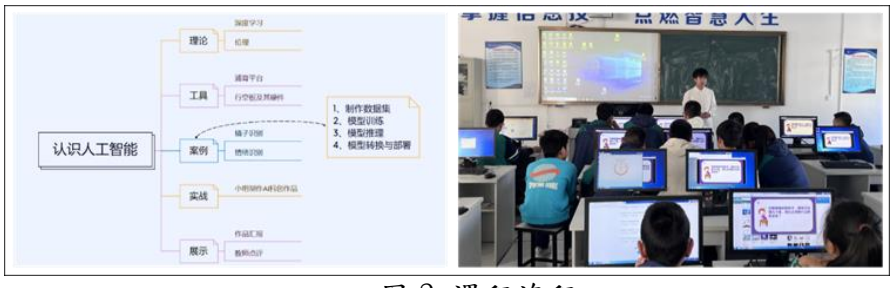


图 2 课程流程

3.3. 支持教学实践

在中小学人工智能教育中，涉及到深度学习和机器学习等底层逻辑的专业课程，学习难度常常较高。为了有效地应对这一挑战，温州大学人工智能教育研究团队采用了项目化学习和小组式课堂教学方法。这两种教学方式相辅相成，使学生能够在理论学习和实践操作中更好地理解和应用所学知识。

项目化学习注重将抽象的理论知识与具体的实践项目相结合。通过与中小学的合作，学校借助自身在学术研究方面的优势，弥补了在实践内容方面的不足，为学生提供了更丰富的

实践机会。这种校地合作模式下展开的项目教学，使学生能够在实践操作中更好地融会贯通，提升实际操作能力。

小组式教学方法则强调学生之间的合作与交流。教师将学生分组，并布置具有指导性的学习任务，以案例或应用为指导，帮助学生更生动地理解和掌握知识内容。在小组学习过程中，学生通过自由思考和充分交流，不仅提升了解决实际问题的能力，还培养了综合素养。

在这种校地合作的模式下，温州大学人工智能教育研究团队展开了更加合理的教学，使学生在理论学习与实践操作相结合中，能够更好地融会贯通，并提升实际操作能力。

举例来说，在进行“橘子识别”项目教学时，教师首先展示了相关的 AI 科创作品，激发了学生的学习兴趣并引起了他们的好奇心。接下来，教师将项目分解为深度学习的原理，包括数据集的制作、模型训练、推理以及模型的转换与部署，深入剖析了每个环节的原理。随后，学生在师生指导下动手操作每个环节，运用课堂中学到的应用知识进行编程操作，如图 3。项目完成后，教师和学生一同对程序设计的可行性进行了检测，并探讨了具体的可行性参数。在项目检验的过程中，教师对学生的设计进行了点评，并针对实际工作中的操作技巧进行了讲解，从而帮助学生提升实践经验和操作水平，实现了应用技术水平的提高。



图 3 上课状况

3.4. 促进教师成长（在职教师（茶山中学老师）和师范生）

在人工智能教育领域，大多数中小学教师缺乏必要的专业知识和教学经验，这阻碍了他们有效地开展人工智能教育的任务。教师队伍中的专业人才匮乏以及培训机会的不足，限制了人工智能教育的深入推进。然而，高校研究团队在这方面拥有丰富的理论知识和研究经验，具备领先的人工智能教育理念和实践成果，可组织并开展专门针对中小学教师的培训活动，以帮助他们更好地理解人工智能教育的核心概念和教学方法。

为了弥补茶山实验中学教师和师范生在人工智能教育方面的不足，温州大学人工智能教育研究团队开展了专门的师资培训课程，如图 4。这一课程由温州大学的教育学院和人工智能专业教师共同设计和授课，旨在为茶山实验中学的教师提供系统的培训和指导。通过与温州大学的合作，茶山实验中学的教师能够及时更新知识和提升能力，为人工智能教育的发展提供了可靠的师资支持。



图 4 教师培训

3.5. 组织社团活动

为了培养学生的竞赛能力和实践经验，温州大学人工智能教育研究团队收集组织了中小学人工智能领域的竞赛活动，鼓励学生积极参加，并为学生提供相关的技术培训和比赛准备指导。同时，温州大学人工智能教育研究团队提供比赛资源支持，包括技术设备、资料信息等，确保学生能够顺利参与竞赛活动并取得优异成绩。温州大学人工智能教育研究团队致力于激发学生的学习兴趣和创新能力，为其未来的发展奠定坚实基础。

3.6. 建立 AI 素养评估体系

人工智能素养是当今 21 世纪学习者发展中的重要议题，涵盖了人工智能知识、技能、态度和价值观的培养。本研究旨在通过采用 Ng 等人（2023）提出的基于 ABCE 框架的评估工具，来评估中学生的人工智能素养，如图 5。此评估工具分为四个分量表，根据 ABCE 框架的定义，分别从情感、行为、认知和道德维度来衡量学生的素养水平。具体而言：情感学习包括四个因素，用于衡量学生的主观感受，包括内在动机、自我效能感、职业兴趣以及学习人工智能的信心。行为学习包括两个因素，涉及学生的行为承诺和建立合作关系，以在人工智能环境中实现学习目标。认知学习包括三个因素，评估学生的知识和技能成就，从低级的了解和理解人工智能，到中级的使用和应用人工智能，再到高阶的评估和创建人工智能的思维技能。道德学习包括七个方面，涉及学生对人工智能的积极心态，包括可靠性、安全性、隐私、责任、透明度、意识以及为社会公益带来的积极影响。

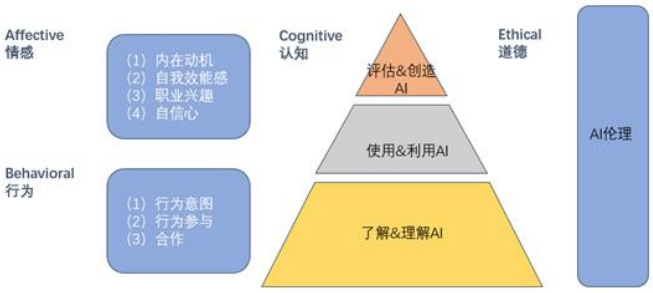


图 5　ABCE 框架

4. 结语

在校地合作背景下，本研究以温州大学附属茶山实验中学为例，探索了中小学人工智能教育的实施路径。通过建立校地合作机制、开设人工智能课程、支持教学实践、促进教师成长、组织学生竞赛和建立 AI 素养评估标准等举措，取得了不错的成效，为推动人工智能教育的发展提供了重要经验。然而，实践时间尚短，后续研究会继续推进，期待更多经验跟大家分享。

参考文献

国务院.**国务院关于印发新一代人工智能发展规划的通知**[EB/0L].(2017-07-20)[2023-09-23].https://www.gov.cn/zhengce/content/2017-07/20/content 5211996.htm.

教育部.**教育部关于印发《高等学校人工智能创新行动计划》的通知**[EB/0L].[2018-06-29.www.moe.gov.cn/srcsite/A16/s70621201804120180410 332722.html

中华人民共和国教育部.**义务教育信息科技课程标准《2022 年版)**[S].北京:师范大学出版社.2022

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2023). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*.

信息科技核心素养框架下人工智能课程设计与策略研究

——以《语音识别技术》为例

Research on Artificial Intelligence Curriculum Design and Strategies under the Framework of IT Core Literacy

--Taking *Speech Recognition Technology* as an Example

王千姿[1*]，陆盈莹[1]
[1]温州大学 教师教育学院
*wqz2316108432@163.com

【摘要】在数字化时代，信息技术已经成为社会发展的重要引擎，而人工智能作为信息技术的前沿领域更是引领着未来的创新和发展。本研究聚焦于信息科技核心素养框架下的人工智能课程设计，响应国家政策号召，对接《义务教育信息科技课程标准（2022 年版）》和美国《CSTA K-12 计算机科学课程标准（2017 年版）》，以八年级信息科技课程《语音识别技术》为例，探讨了课程设计策略与实施方案，旨在助力我国中小学阶段人工智能教育的高质量发展，切实提升学生的数字素养和创新能力，为其应对未来智能化社会挑战奠定坚实基础。
【关键词】人工智能课程；核心素养；信息科技；语音识别

Abstract: *In the digital era, information technology (IT) has become an important engine of social development, and artificial intelligence (AI), as a cutting-edge field of IT, is leading the future innovation and development. This study focuses on the artificial intelligence curriculum design under the framework of IT core literacy, responds to the national policy call, buttresses the Compulsory Education IT Curriculum Standards (2022 Edition) and the U.S. CSTA K-12 Computer Science Curriculum Standards (2017 Edition), and explores the curriculum design strategy and implementation plan by taking the eighth-grade IT course "Speech Recognition Technology" as an example. It aims to help the high-quality development of artificial intelligence education at the primary and secondary school levels in China, effectively improve students' digital literacy and innovation ability, and lay a solid foundation for them to meet the challenges of an intelligent society in the future.*

Keywords: Artificial intelligence curriculum, Core literacy, Information technology, Speech recognition

1.引言

在当前信息化社会，人工智能已成为各个领域的热点话题，它不仅引领着科技创新的浪潮，更在教育领域中不断焕发生机与活力。2017 年 7 月，国务院发布第一份有关人工智能行业的系统部署文件《关于印发新一代人工智能发展规划的通知》，作为我国人工智能领域的顶层设计，明确了人工智能发展的战略目标、重点任务和保障措施，为人工智能技术的快速发展奠定了坚实基础。2018 年 4 月，教育部印发《高等学校人工智能创新行动计划》，提出“构建人工智能多层次教育体系，在中小学阶段引入人工智能普及教育。2022 年，教育部印发《义务教育信息科技课程标准（2022 年版）》，以数据、算法、网络、信息处理、信息

安全、人工智能为信息科技学科逻辑主线，注重培养学生的核心素养。今年全国两会期间，民进中央提出《关于积极推动人工智能赋能教育变革的提案》，建议推动人工智能技术在教育领域的应用。一系列政策的颁布，为我国中小学人工智能课程的开展指明了方向，人工智能正成为中小学阶段信息科技课程重要的教学内容之一。在此背景下，本研究围绕信息科技核心素养对人工智能课程进行设计，探究人工智能课程在设计过程中的关键点，在实践中探寻初中阶段人工智能课程的实施策略。

2.课程设计背景与理念

2.1. 信息科技核心素养

在 2022 年发布的《义务教育信息科技课程标准（2022 年版）》（以下简称“新课标”）中，人工智能的相关内容作为独立模块依托信息科技课程的开展在学校落地实施(欧等, 2022)，推动学生数字素养与技能逐步提升。在新课标中，人工智能的学习目标依据信息科技的学科核心素养进行制定，第三学段（5-6 年级）和第四学段（7-9 年级）是幼小衔接阶段，以此为例进行梳理（见表 1）。

表 1　第三学段与第四学段的核心素养要求

核心素养	学段	
	第三学段（5-6 年级）	第四学段（7-9 年级）
信息意识	1.感受应用信息科技获取与处理信息的优势 2.有意识地选用信息技术工具处理信息	1.了解人工智能对信息社会发展的作用 2.具有自主动手解决问题、掌握核心技术的意识
计算思维	1.通过生活实例，了解算法的特征和效率 2.能够将问题分解成可处理的子问题，了解反馈对系统优化的作用	1.通过案例分析，理解人工智能。 2.合理选用人工智能，比较使用人工智能和不适用人工智能处理同类问题效果的异同
数字化学习与创新	1.有意识地将算法应用与数字化学习过程中，适应在线学习环境 2.利用恰当的数字设备规划方案、描述创作步骤	1.具备利用信息科技进行自主学习和合作学习的能力 2.主动利用数字设备开展创新实践活动
信息社会责任	1.认识到算法对解决生活和学习中的问题的重要性 2.认识到自主可控技术对保障网络安全和数据安全的重要性	1.通过体验人工智能应用场景，了解人工智能带来的伦理与安全挑战 2.合理地与人工智能开展互动，增强自我判断意识和责任感

从表中可以看出两个学段的课程目标相互联系、相互促进，从整体性的角度去设计教学内容，学生经过第三学段对算法的学习与运用，认识到自主可控技术的重要性和必要性，衔接第四学段对人工智能的理解和分析，让学生更加深刻地体会到人工智能与智慧社会的关联，以及发展人工智能应遵循的伦理道德规范。因此，在根据第四学段的人工智能课程设计教学时，应注意创造真实情境，深化学生对知识的原理认识，引导学生利用信息科技手段去探索解决生活中实际问题的有效方法(熊璋等, 2022)。

2.2. 《CSTA K-12 计算机科学标准》

2017 年，美国计算机科学教师协会（CSTA）为了有效推进计算机科学教育的持续发展，开发并发布了《CSTA K-12 计算机科学课程标准（2017 年修订版）》（以下简称为《CSTA 课程标准》），这是美国当前基础教育计算机科学课程的国家标准。它明确指出，学生要能够使用人工智能与人类玩游戏或解决实际问题时发展计算思维（Computer Science

Teachers Association[CSTA], 2017），提出了计算机科学教育的具体内容，描述了一整套核心学习目标，明确计算机科学这门学科应该掌握的知识与技能，其中包括 5 个核心概念和 7 个核心实践内容。随着大数据、机器学习、云计算等信息技术在教育领域的深入运用，世界各国都在加强人工智能人才的培养。为顺应人工智能时代的发展，该标准还提出培养学生的计算思维与数字信息素养的课程目标。

而计算思维不仅是信息科技学科重要的核心素养之一，还是人工智能课程中问题解决必备的核心素养(沈书生 & 翁子凌, 2023)。为了适应时代的变化，该计算机标准最大的特点就是从原来的“具体标准”转变为“素养标准”，更是强调了计算思维在所有学科中的应用，包括对于解决问题、建立科学知识体系和知识再创造的重要作用。

《CSTA 课程标准》共有 120 条，提到的 5 大核心概念：（1）计算机系统（CS）（2）网络及互联网（NI）（3）数据与分析（DA）（4）算法与编程（AP）（5）计算机的影响（IC）。据统计，“算法和编程”在 K-12 每个等级中的占比是最大的，是计算机科学课程的重点，也是培养学生计算思维的关键，将此概念引入人工智能课程设计中有利于塑造学生的人工智能认知。

《CSTA 课程标准》中的七大核心实践如下图所示：

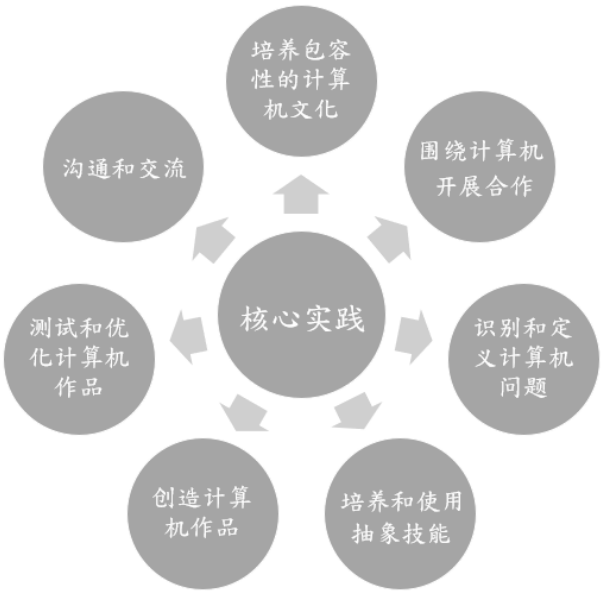


图 1 CSTA K-12 计算机科学七大核心实践

这七大核心实践对人工智能课程的开展有重大的指导意义：首先，培养包容性的计算机文化。在人工智能领域也需要尊重多元文化和不同观点，这样才能创造出更具有包容性和创新性的人工智能应用；其次，围绕计算机开展合作。人工智能课程不仅仅是学习技术，更是培养协作能力的过程，人工智能项目往往需要多人合作才能完成；识别和定义计算机问题、培养和使用抽象技能，这些都是 AI 学习中必不可少的。人工智能的核心就是要能准确识别问题，并抽象出解决方案；创造计算机作品。对人工智能课程来说，学生需要动手实践，创造出真正的 AI 作品，进而更好地掌握相关知识和技能；测试和优化计算机作品则是保证创造出来的 AI 作品质量和性能的关键；最后是沟通与交流。无论是人工智能领域还是其它领域都是非常重要的。

因此，教师在设计人工课程时可以参照标准中对于不同等级的要求，并结合实际教学环境

选择核心概念所对应的知识内容，将七大核心实践融入教学活动中。《CSTA 课程标准》对我国的人工智能课程开发产生比较广泛的影响，我国教育部也随之逐步更新相应的课程标准，本研究也将此作为参考进行核心素养下人工智能课程设计。

3.课程设计依据与策略

《中小学人工智能课程指南》提出，初中阶段的人工智能课程在设计的侧重点上注重初步学习基本概念和基本原理(中小学人工智能课程指南课题组 & 江波, 2023)，在体验的基础上感受到人工智能的魅力，让学习者进一步思考人工智能如何解决问题。新的信息科技课程标准强调素质教育、强调大概念、主题式、体验性、综合化，不同于以往重复操作、简单编程、知识点罗列的教学方式(熊 & 李, 2023)，因此本研究以信息科技学科核心素养、CSTA K-12 计算机科学标准实践维度为设计依据，选择浙教版八年级下册第二单元第五课《语音识别技术》作为教学内容进行人工智能课程设计。

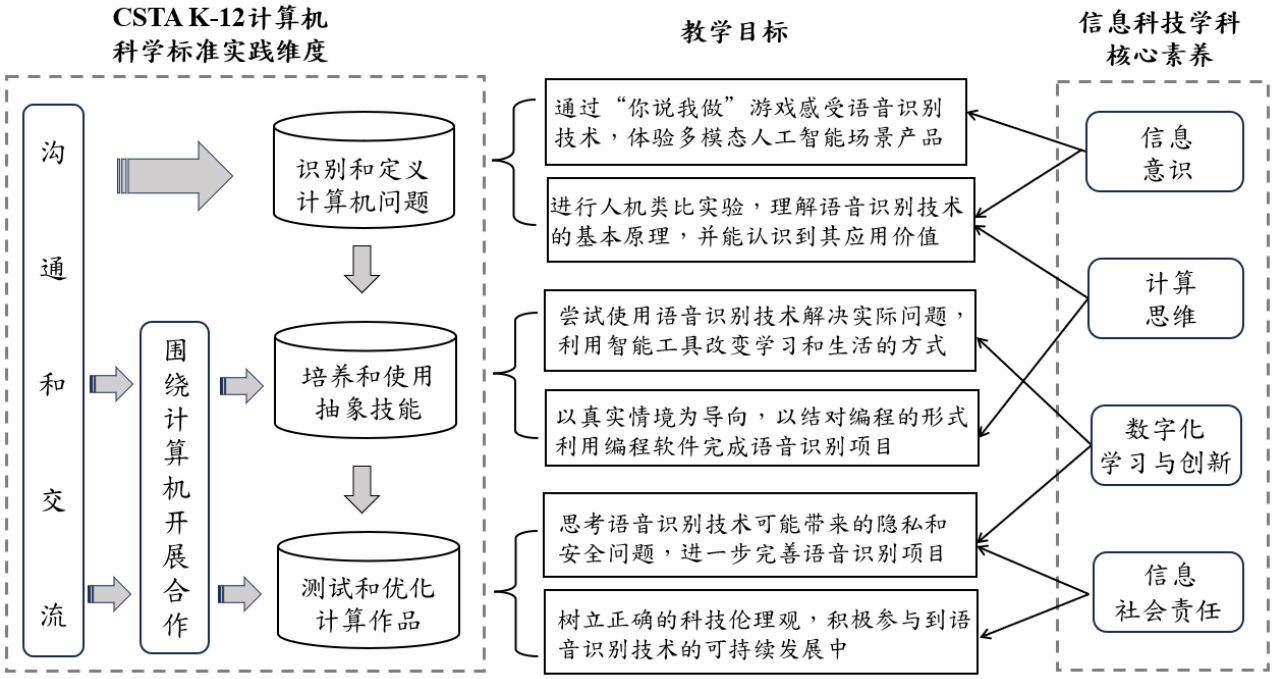


图 2 课程目标与核心素养图

《语音识别技术》根据《CSTA 课程标准》中识别和定义计算机问题、培养和使用抽象技能、测试和优化计算作品、围绕计算机展开合作以及沟通和交流五个实践维度进行教学设计，教学目标与实践维度相对应，沟通和交流这一实践维度贯穿于整个教学流程之中，在任务驱动中锻炼学生围绕计算机展开合作的能力。在教学活动实施的过程中尊重学生的主体地位，让学生在亲身体验中了解语音识别技术，理解语音识别的工作流程，循序渐进地引导学生探究语音识别技术的基本原理，在项目化实践的过程中逐步落实对学生核心素养的培养要求，从而培养学生交叉学科的创新思维，进而思考如何利用人工智能解决问题(谢忠新等, 2019)。

4.课程设计流程与实施

本研究围绕八年级的信息科技课展开设计，课题是《语音识别技术》，属于单元教学《智

能技术初体验》中的第一节课，既是在了解第一单元《人工智能在身边》之后体验感较强的一节课，也是作为人工智能技术入门的第一节课，有着激发学生对人工智能兴趣的重要作用。本节课的教学设计环节如下：

4.1. 识别和定义计算机问题

环节一：创设情境，游戏导入

在教学实践中，情境创设是至关重要的。教师在教师端打包发送“你说我做”小游戏程序包至学生端，以便学生通过语音指令控制游戏角色的移动。在游戏操作过程中，让学生进行思考：游戏角色是如何识别语音指令并做出相应反应的？因此，本部分中将以生活中人工智能在语音识别领域的实际应用作为本节课的导入环节，使学生得以体验语音识别技术的通用性，从而认识到计算机在解决现实生活问题的潜力。同时，这一举措也将激发学生对语音识别技术的求知欲，为后续教学环节展开做好铺垫。

环节二：对比学习，了解原理

在此环节中，教师首先会准备一段发音标准、长度适宜、语速较慢的名人名言朗诵，内容简单易懂，学生在任务单中记录相应的听写内容，同时教师利用讯飞开放平台上的实时语音听写功能与学生同步进行听写，记录人机各自听写的完成情况。随后教师再准备另一段发音标准，长度较长、语速较快的诗词朗诵，内容同样简单易懂。人机再次同时完成听写任务。教师根据两次的听写情况，对比人机之间的准确率与流畅度，引导学生思考：在听写内容难度增加的情况下，人类与机器之间完成听写任务的差异。学生以小组形式讨论并总结人类和机器进行语音识别的具体流程，以流程图形式展示分享。教师进一步总结人机在语音识别流程中的相似之处，讲解语音识别的基本流程，并简单介绍影响语音识别技术准确率的因素，如远场、噪音、背景音、口音、吞音等。

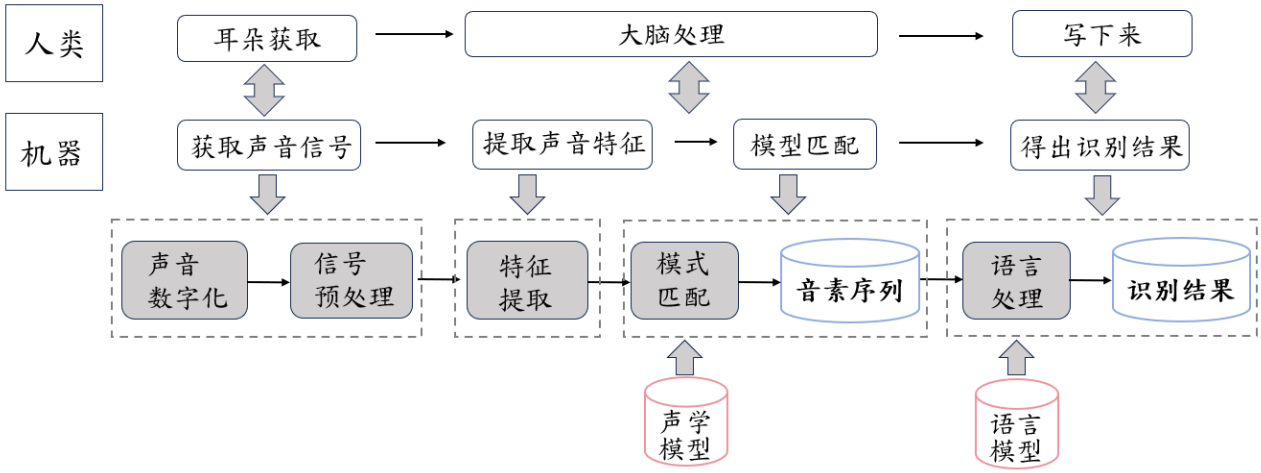


图 3　人机语音识别流程

通过人机同步听写环节的设计，学生将参与手工模拟 AI 算法的活动(方 & 黄，2020)，使用纸笔和流程图分析机器在语音识别中的基本原理。这种原理可视化的方式有利于培养学生将现实问题分解成可控的小问题的实践能力和逻辑思维，符合信息意识中理解人工智能对生活、社会影响的素养导向。

4.2. 培养和使用抽象技能

环节三：程序设计，掌握技术

学生在了解语音识别技术的基本原理之后，将依据学习任务单上的提示，以结对编程（Pair Programming）的形式完成游戏模块搭建。学习任务单包括基本任务和提升任务，其

中基本任务回归到导入环节中的“你说我做”小游戏，让学生在实践操作中认识每个控件和模块的作用和意义，体验语音识别模块的神奇之处，激发学生探究信息技术的兴趣。此外，学有余力的同学可以完成提升任务，在基本任务的基础上对原有作品进行升级改造，进一步实现如跟读、点歌等其他的复杂指令，使游戏角色更加丰富完整。

表 2 “结对编程”学习任务单

“结对编程”学习任务单		
基本任务	完成游戏模块搭建并运行，实现角色能识别并执行简单的动作指令，如左转、后退等	
	提示	1、根据语音识别流程搭建模块：听到-识别-执行
		2、涉及模块类型：事件动作、控制、声音
提升任务	创新升级作品，实现角色能识别并执行其它复杂的指令，如跟读、点歌等	
	提示	跟读即将识别结果录音并播放、点歌即识别结果并播放相应音乐

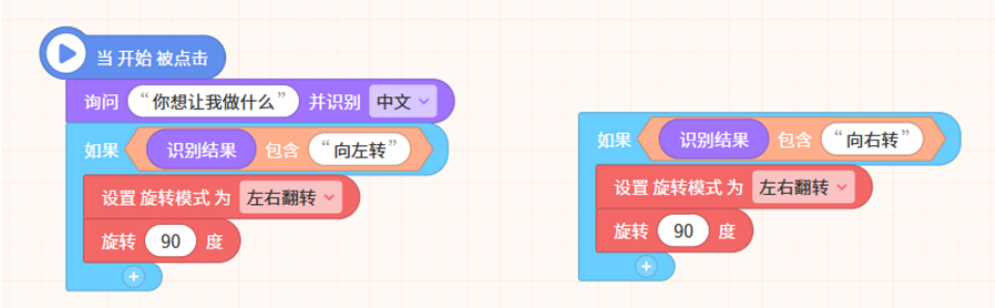


图 4 编程设计基本任务模块

在学生有了识别模式、分解问题、使用抽象思维和利用已有资源的经验之后，应该注重发展学生的抽象能力，而搭建游戏模块的实践操作就能创造机会让他们开发通用模块，减少复杂度，并能使其适用于更加广泛的问题。学生以结对编程的形式完成编程设计，即强调与同伴在沟通交流中围绕计算机开展合作，教师在这一过程中需要注意不同个体适合的协作对话模式(Tan 等，2024)，因势利导培养学生的计算思维和数字化学习能力。 同时，考虑到学生的个体差异，本环节中设置了不同难度的任务，便于开展分层教学。编程基础较好的同学可以适时引导其创建更加复杂的、且更符合现实生活的模型，来理解和评估可能产生的结果，以此达到个性化学习的教学目的。

4.3. 测试和优化计算作品

环节四：作品展示，优化创新

各小组经过合作探究完成编程作品后，将游戏作品源文件上交至学习平台，教师选取在编程设计过程中有代表性的作品进行展示，小组成员说明设计思路和理念以及分工安排，其他同学同时在学生端完成教师发送的作品评价表，并尝试提出优化建议。教师鼓励小组之间在学习平台上积极交流和评价，接纳来自老师和同学的可行性意见，以此为基础进行完善和优化游戏作品。

表 3　编程作品评价表

作品评价表				
作品名称		评价主体	教师	
			组别	
评价要素		评价标准	分值	得分
作品成果	规范性	程序准确、完整	10	
	技术性	运行流畅、不卡顿	20	
	创意性	作品富有创意、有多个创新点	30	
展示交流	汇报表现	表达清晰、回答问题具有针对性	20	
课堂表现	合作交流	能互帮互助、相互答疑	10	
	学习态度	积极回答问题、任务完成速度快	10	
总分			100	

测试和优化的过程是一个反复迭代的过程，学生应该在这个过程中学会发现问题、修复问题以及优化作品。在反复评估和优化计算机作品的过程中，学生对比“预期结果”和“实际结果"，使用策略和技巧去识别和修复程序错误，以此提高作品的性能和可靠性，培养其数字化学习与创新的能力。而在评价环节的实施，聚焦学生的核心素养以及高阶能力发展水平，及时检验学生的学习效果，让学生对人工智能的进一步学习抱有期待。

环节五：归纳总结，拓展延伸

在上一环节结束后学生思考本课收获，教师进行补充的同时让学生发散思维讨论语音识别技术在实际应用时的潜在风险，包括隐私数据泄露和语音误识别等问题，结合学生日常生活中发现的问题，迁移新知，产生联想。在实际生活情境中引导学生意识到承担信息社会责任的必要性，对于培养学生的技术伦理意识和社会责任感具有重要意义，符合信息意识中使用信息科技解决学习和生活中问题的导向。由此，学生将能够更加全面地理解和评估技术应用的影响。

5.人工智能课程设计的建议

5.1. 融入不插电活动

尽管目前关于发展计算思维是通过插电活动还是不插电活动尚未有定论，一方面有研究表明，在经过短期教育干预后，插电与不插电的活动都可以显著提高幼儿的计算思维，实验组与对照组之间并不存在显著差异(Zhang 等，2024)，另一方面也有研究采用 Meta 分析方法定量分析不插电编程活动对 K-9 年级学生计算思维的影响，结果表明不插电编程对 K-9 年级学生的计算思维能够产生积极影响(Lin 等，2023)，但是考虑到插电活动在人工智能课程中表现出的无屏幕、阈值低、认知负荷低、成本低等优势，因而在进行人工智能课程设计时，可以适当加入不插电的活动，如本研究中基于人机听写环节让学生思考画出人机进行语音识别的具体流程，类似的活动设计可以减少教师在开展人工智能相关教学活动时的束缚，同时学生通过不插电活动可视化展现其复杂和抽象的计算思维，直观地类比人工智能技术原理的实现过程，教师也因此可以通过观察学生的编程行为来了解其实际的掌握程度并实施干预。

5.2. 完善课程评价体系

收集评估证据开展多元评价，是将学生学习表现外显表征为人工智能意识、技术应用能力、实践创新思维和信息社会责任方面的素养达成，是开展课程评价的关键(熊 & 李, 2023)（马等，2023)。人工智能课程的目标之一就是学生核心素养的培养，因此在人工智能课程评价的设计上，评价内容要丰富化、方式多样化、主体多元化、工具科学化。

评价内容丰富化可以从课程内容的多样性、实践项目丰富性以及课程创新前瞻性等维度来考量。评价方式多样化可以结合人工智能技术的优势，设计多种评价形式，注重评价的客观性和精确性，全面评估学生的学习成果和能力。要实现评价主体多元化首先应该鼓励学生针对所创作的课程作品进行自评和互评，以全面反映其学习体验和成长情况，培养其自我反思与评价能力；其次，教师评价也必不可缺，其专业视角有助于精准评估学生的学习效果，并为人工智能课程的改进提供方向；家长作为重要的利益相关者，其评价可以从家庭教育的角度提供补充；此外，还可以引入行业专家和企业代表的评价，能从实际应用的角度考虑课程的前瞻性和实用性。因此，多元化的主体评价可以为人工课程优化和发展提供有力支撑。人工智能课程的评价工具可以充分利用人工智能技术使其更为科学和高效，综合使用自动化评价系统、语义分析技术及机器人辅助评价，同时，结合量表、量规等传统评价工具全面评估学生知识掌握与核心素养。借助科学化的评价工具，能更精确地评估人工智能课程的教学质量与学生的学习成效。

5.3. 注重跨学科融合

目前新兴的学校人工智能教学实践侧重于认知人工智能内容、工具、科学技术和思维(Chiu, 2021)，在义务教育阶段主要依靠信息科技课程来落地人工智能教育仍存在诸多局限性，包括重点弱化、主线模糊、衔接不畅等问题(钟 & 詹, 2022)，而人工智能教育本身也是一种跨学科教育，与数学、心理学、计算机科学等学科密切相关，具有包容性、综合性和实用性的特点，因此在进行人工智能课程设计时，可以通过跨学科教育的形式培养学生的人工智能素养。在学科交叉设计的环节中，如本研究中学生完成语文学科的听写内容类比机器进行语音识别的过程，一方面学生在真实的互动体验中复习巩固了语文的古诗词内容，另一方面又能深刻地感知语音识别技术的基本原理，在学科融合的过程中学生内在的知识结构体系相互联系，形成更加严密的逻辑框架，从而培养学生的关键能力，为他们适应未来的工作环境奠定坚实的基础，成为智能时代积极的参与者。

6.结语

本研究面向中小学教育阶段，首先简单梳理了国内关于人工智能课程的相关政策，明确人工智能教育在教育领域的总体定位；接着从我国信息科技课程标准和美国计算机科学课程标准对人工智能课程开展提出的指导意见出发，并以此为设计理念选取具有代表性的人工智能课程《语音识别技术》作为设计案例，提出了三条适用于人工智能课程设计的可行策略。研究强调在人工智能课程设计中以 CSTA K-12 中的五个实践维度为导向，将信息意识、计算思维、数字化学习与创新、信息社会责任核心素养贯穿其中，通过情景导入、类比学习、动手实践等教学环节，引导学生理解和应用人工智能技术，解决实际问题。同时还提倡在人工智能课程中结合不插电活动，多元化评价体系，注重跨学科融合，培育学生核心素养，赋能学生终身发展。

尽管当前已有部分地区率先进行了中小学人工智能教育教学的探索与实践，但由于缺乏统一的课程标准、学科体系与资源庞杂、各地实际教学环境的差异等现实问题导致人工智能课程难以推行，本研究呼吁更多教育工作者和研究者共同参与，以期形成更为完善的课程体系与实施策略。

参考文献

马昊学, 郑苏 & 张玲.(2023).面向核心素养的中小学人工智能课程设计研究.**中国信息技术教育**(02), 83-87.

方圆媛，& 黄旭光.（2020）. 中小学人工智能教育：学什么，怎么教——来自“美国 K-12 人工智能教育行动”的启示. **中国电化教育**，10，32－39.

沈书生 & 翁子凌.（2023）. 人工智能课程中促进计算思维培养的场景设计研究. **远程教育杂志**，41(6)，94－103. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2023.06.010

钟柏昌，& 詹泽慧.（2022）. 人工智能教育的顶层设计：共识、差异与问题——基于 4 套标准文件的内容分析. **现代远程教育研究**，34(4)，29－40.

谢忠新，曹杨璐，& 李盈.（2019）. 中小学人工智能课程内容设计探究. **中国电化教育**，4, 17－22.

熊璋，赵健，陆海丰，& 欧阳元新.（2022）. 义务教育阶段信息科技课程的时代性与科学性——《义务教育信息科技课程标准(2022 年版)》解读. **教师教育学报**，9(4)，63－69. https://doi.org/10.13718/j.cnki.jsjy.2022.04.008

熊璋，& 李正福.（2023）. 义务教育阶段信息科技课程建设路径研究. **中国电化教育**，1，127－132.

欧阳嘉煜，王宇，& 汪琼.（2022）. 美国 K-12 阶段人工智能课程大概念与课例设计解读. **现代教育技术**，32(12)，13－22.

中小学人工智能课程指南课题组 & 江波.（2023）. 中小学人工智能课程指南. **华东师范大学学报（教育科学版）**，41(3)，121－134. https://doi.org/10.16382/j.cnki.1000-5560.2023.03.013

CSTA Standards Task Force.（2017）CSTA K－12 Computer Science Standards: Revised 2017, ACM CSTA.

Chiu, T. K. F. (2021). A Holistic Approach to the Design of Artificial Intelligence (AI) Education for K-12 Schools. *TechTrends*, 65(5), 796－807. https://doi.org/10.1007/s11528-021-00637-1

Lin, Y., Liao, H., Weng, S., & Dong, W. (2023). Comparing the effects of plugged-in and unplugged activities on computational thinking development in young children. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12181-x

Tan, J., Wu, L., & Ma, S. (2024). Collaborative dialogue patterns of pair programming and their impact on programming self‐efficacy and coding performance. *British Journal of Educational Technology*, 55(3), 1060－1081. https://doi.org/10.1111/bjet.13412

Zhang, Y., Liang, Y., Tian, X., & Yu, X. (2024). The effects of unplugged programming activities on K-9 students’ computational thinking: Meta-analysis. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-023-10339-5

生成式人工智能赋能精准教学的实践路径与框架设计

Practice paths and framework design of Generative artificial intelligence enabling precision teaching

廖龙燕[1*]，任吉坚[1]
[1]温州大学 教师教育学院
[*]405910606@qq.com

【摘要】 近年来，精准教学和生成式人工智能受到了教育工作者的广泛关注，生成式人工智能凭借其优秀的数据分析及内容生成等能力为精准教学的实践提供了新的可能性。本文旨在探讨生成式人工智能如何赋能于精准教学，以促进教育个性化学习的实现。本文围绕两个核心目标展开：第一，设计生成式人工智能赋能精准教学的实践路径，旨在明确生成式人工智能在精准教学中的具体应用步骤，确保技术有效集成与应用；第二，基于实践路径系统设计实践框架，该框架旨在指导实践者如何系统利用生成式人工智能技术，使其作用于精准教学的实践当中。本文为教育领域提供了新的视角和实践指导，希望生成式人工智能在推动精准教学实施、提升教学质量与效率方面发挥更关键的作用。

【关键词】 生成式人工智能；精准教学；个性化学习；实践框架

Abstract: *In recent years, precision teaching and generative artificial intelligence have attracted widespread attention from educators. Generative AI, with its excellent capabilities in data analysis and content generation, offers new possibilities for the practice of precision teaching. This paper aims to explore how generative AI can empower precision teaching to promote the realization of personalized learning in education. This paper revolves around two core objectives: first, to design a practical path for empowering precision teaching with generative AI, aiming to clarify the specific application steps of generative AI in precision teaching to ensure the effective integration and application of technology; second, to systematically design a practical framework based on the practical path. This framework is intended to guide practitioners on how to systematically utilize generative AI technology in the practice of precision teaching. This paper provides a new perspective and practical guidance for the field of education, hoping that generative AI will play a more crucial role in promoting the implementation of precision teaching and improving the quality and efficiency of teaching.*

Keywords: Generative artificial intelligence; Precision teaching; Personalized learning; Practice framework

1. 引言

长期来看，教育界一直在努力解决规模化教育与个性化需求之间的矛盾，展现了积极的应对态度，正如中共中央、国务院印发的《中国教育现代化 2035》规划中提到，要“利用现代技术加快推动人才培养模式改革，实现规模化教育与个性化培养的有机结合”(叶新东 & 刘泽民, 2024)。精准教学正成为解决这一问题的关键，其通过精确测量学习过程，来追踪学生的学习表现，并进行科学、精确的教学决策的初衷受到了教育工作者的关注与青睐(万力勇, 黄志芳, & 黄焕，2019)，但精准教学在实践过程中依然面临着对学习者学习需求、状态、路径的分析不足等诸多困境(郝建江 & 郭炯, 2022)。

近年来，生成式人工智能技术的突破为精准教学提供了新的动力和可能性。通过模拟人类思维方式生成文本、图像、音频和视频等内容的能力，生成式人工智能技术在个性化内容创制、学习过程优化及实时反馈提供等方面展现出巨大潜力，可以有效地解决当前精准教学过程中个性化内容生成和个性化评估等问题。例如，赵元成等人提出，生成式人工智能根据用户的不同性格特点和个性化需求，为其适配精准化、人性化的内容(赵元成 & 何叶, 2024)。但尽管生成式人工智能在教育领域掀起了一场巨大革命，预示着教学方法和学习体验的根本

变革，但要充分激发这项技术的巨大潜力，并确保其在实现精准教育目标上发挥最大效能，仍然是一个待解的复杂问题。我们面临的挑战不仅仅是技术如何使用，更在于如何合理利用生成式人工智能的强大能力精准匹配教育需求，以推动个性化学习的深度发展。

为此，本文着手于两个关键的工作：首先，本文构建了一条清晰的生成式人工智能赋能精准教学的实践路径，旨在勾画出生成式人工智能作用于精准教学的具体步骤，以确保技术的有效集成与应用。紧接着，基于实践路径，我们进一步发展出一个综合的实践框架，该框架不仅指导实践者系统地将生成式人工智能技术融入精准教学的具体操作中，同时也为教育政策制定者提供决策支持，促使教育技术创新与教育公平的双重进步。这一方法论框架旨在为教育领域的未来探索提供新的视角，助力生成式人工智能技术在推进教育现代化进程中发挥其应有的价值。

2. 关键概念界定

2.1. 精准教学

精准教学(Precision Teaching)起源于 20 世纪 60 年代，由美国奥格登·林斯利（Ogden Lindsley）基于斯金纳的行为学习理论提出，用于追踪小学生的学习表现和支持数据决策，后来发展为用于评估任意给定的教学方法有效性的框架，可评估任何学科、任何学段的教学(段沙 & 周怡, 2017)。学术界的不同学者对精准教学持有不同的观点，Lindsley 认为，精准教学是根据标准变速图表上学生持续的自我监控表现频率的变化而形成的教育决策(Lindsley, 1992)。目前国内祝智庭教授团队对精准教学概念有着深刻的理解，他们认为，精准教学的开展需要做到目标、问题和干预三个方面的精准(彭红超 & 祝智庭, 2016)。他们从课堂教学数据的记录、分析和决策等三个要素总结了精准教学过程中进行教学决策的方法，并指出精准教学的价值和意义在于支持每一位学习者对所有知识点或技能的掌握(潘巧明 & 赵静华, 2019)。从中可见，精准教学不仅延续了早期对教学决策和学习评价功能的关注，而且将“科学取向”的教育思想借由技术应用渗透到教学活动的全过程，通过对教学数据的挖掘和多元分析，实现利用技术和数据辅助教师更精准地教、学生更精准地学。

2.2. 生成式人工智能

联合国教科文组织于 2023 年 9 月发布《生成式人工智能教育与研究应用指南》从人工智能对人类思维表征符号系统的模拟角度，对“生成式人工智能(Generative Artificial Intelligence，以下简称GAI)”进行明确的定义：GAI是根据人类借助思维符号表征系统表达的提示(Prompts)自动生成内容的人工智能技术(苗逢春, 2023)。

这种技术的核心在于使用算法模仿输入数据的分布，从而产生与原始数据类似但又不完全相同的新数据。GAI 的应用非常广泛，从自动撰写新闻报道、创作诗歌、故事，到生成逼真的图像、音乐作品、甚至是仿真语音，都体现了其创造力的巨大潜力。这些技术背后最常见的算法模型包括生成对抗网（GANs）和变分自编码器（VAEs），以及最近非常流行的大型语言模型（如 GPT 系列）。

3. 生成式人工智能赋能精准教学的实践路径

随着技术的快速发展和教学需求的日益复杂化，传统教学方式已经难以满足个体化学习的需求，这使得寻找能够实现高度个性化教学内容和方法的新路径变得尤为重要(陈雄辉, 谷紫阳, 覃以凤, 马宗兵, & 龙娜, 2022)。精准教学和 GAI 出现让教育工作者看到了新的突破点，但面临的问题是如何发挥 GAI 的最大价值，使之作用到精准教学的教育实践。任红杰指出，基于大数据的精准教学的生成需遵循学生画像、教学目标、教学过程和教学辅导四个方面的基本路径(任红杰, 2017)；郭利明等人指出，精准教学应从教学方法、教学模式、教学活动、教学策略和教学评价五个方面打造高效课堂(郭利明, 杨现民, & 张瑶, 2021)。综合前人所述，本文认为 GAI 以其强大的数据处理和内容生成能力，为精准教学的实践提供了可行的道路，在实现教育内容的个性化推送、教学活动的多样化设计、学习路径的个性化生成、学习

疑问的情境性解答以及学习评估的动态性反馈方面有着巨大潜力。

接下来将详细阐述 GAI 赋能精准教学的实践路径能够实现的基础及其路径效果的介绍，为构建具体的实践框架提供理论基础和实践指导。

3.1. 路径的实现基础

生成式人工智能系统 ChatGPT 于 2022 年 11 月由美国开放人工智能研究中心正式发布，且迅速成为教育领域关注和讨论的焦点。以 ChatGPT 为典型代表的 GAI 技术，在文本内容创造及上下文情境解析等领域展现了杰出的性能，这有可能促成从教育理念到教学实践的根本性变革(卢宇, 余京蕾, 陈鹏鹤,& 李沐云, 2023)。此外，它为长期以来由技术限制所束缚的教学方法，如启发式教学和个性化反馈，提供了新的可能性，开启了智能教育应用的新篇章。综合来看，与传统的弱人工智能相比，以 ChatGPT 等大模型为代表的 GAI 正在向强人工智能迈出坚实的一步，为赋能精准教学的实践路径打下了坚实的基础。以 ChatGPT 等大模型为代表的 GAI 有如下显著的优点和趋势：

3.1.1. 模型训练方面

大型预训练模型和迁移学习技术只需标记少量数据就可实现微调预训练，很多情况下无需开发人员从头开始训练，这有利于节约大量的时间成本(唐亚汇, 2021)。同时，预训练模型通过数据集增加和模型优化，可不断提升应用场景多样性和反应准确性。理论上，整合全球所有可获取的文本、图片等数据，模型能自动掌握知识，无需人工干预(毕文轩, 2023)。特定任务的微调提升了模型的灵活性和适应性，使其能够高效服务于多任务、多场景，让 GAI 模型成为多样化内容生产的高效“流水线”，推动数字内容生产进入新阶段。

3.1.2. 内容生成方面

ChatGPT 具备十分强大且多样的内容生成能力。ChatGPT 在创造性工作如编写儿童读物、学术论文和诗歌等方面开始替代人类，并通过上下文语境的对话实现与人类类似的流畅交流，体现生成逻辑的流畅性(李海峰 & 王炜, 2023)；同时具备生成包括文本、图片和视频在内的多模态内容的能力，突显生成技能的多样性；具备道德意识，能进行连续性追问、承认错误、质疑不当前提和拒绝不适当请求，反映生成思想的道德性(Kulkarni, 2023)；随着模型参数超过特定临界值，展现出独特的创造能力，其大模型参数达到 15 亿个，表明生成模型的巨大性(Wei et al., 2022)。

3.1.3. 模型应用方面

在算法、数据和算力的综合推动下，GAI 的应用已迅速展开。GAI 的应用发展主要依托于算法、数据和算力的进步，分为三个主要层面：首先，有实力的企业通过 API 或开源方式推出基础生成算法模型和通用预训练大模型；其次，对这些大模型进行专业训练和调整，打造适应教育、金融、传媒、游戏等行业需求的定制化小模型；最后，在此基础上开发面向个人和企业的终端应用产品，如聊天机器人和图像生成软件等(李颖婷, 2023)。

GAI 的以上特性不仅促进了教育内容和方法的创新，还为实现个性化、高效率和公平的精准教学提供技术支持，同时也为学生提供更加丰富、互动和个性化的学习体验。

3.2. 路径的具体效果

3.2.1. 教学内容精准推送

GAI 针对用户的独特需求生成个性化内容，包括文本、图片、视频和音频等形式的学习材料，同时分析学生的学习数据，包括知识掌握水平、学习风格和兴趣，然后基于这些数据自动生成个性化的教学内容。这意味着每个学生可以获得专门为他们量身定做的学习材料，从而提高学习效率和兴趣。一个典型的例子是采用 GAI 技术的自适应学习系统，它依赖于学习者模型、教学方法模型和学科知识模型三种核心模型，以确保学习者能够接触到与自身能力匹配的知识内容(白雪梅 & 郭日发, 2024)。

3.2.2. 教学活动精准设计

GAI 通过其先进的数据处理和分析能力，设计高度个性化和多样性的教学活动。这种技术能够根据学生的学习习惯、兴趣和理解速度生成适宜的学习材料，如定制化的习题、教学

视频和互动式模拟活动。同时实时调整教学内容和难度，确保学习路径与学生的实际进展相适应，最大化学习效率。例如，当学生的思维能力有待提高时，以 ChatGPT 为代表的 GAI 为学生开启一场“苏格拉底式对话”的教学活动提供了技术条件(周洪宇 & 常顺利, 2023)，教师可以在教学过程中让学生针对具体问题与 GAI 系统展开“苏格拉底式对话”，这有利于学生在轻松愉悦的氛围下去独立解决问题，锻炼思维能力。

3.2.3. 学习路径精准生成

GAI 通过深入分析学生的学习风格和偏好，精准设计出完全个性化的学习路径。它根据每位学生的具体需要调整教学内容，创建或优化学习材料以匹配学生的知识水平和兴趣点。通过这种方式，GAI 制定既富有挑战性又贴合学生个人发展的目标。例如，在文本写作中，根据写作需求，用户要求 GAI 提供个性化的精准资源推荐和学习路径规划服务，例如提供写作理论、注意事项、文体特征和范文等，为写作活动形成支架(汪靖, 米尔外提·卡马勒江, & 杨玉芹, 2023)。

3.2.4. 学习疑问精准解答

在精准教学的实施中，向学生提供定制化的解答和指导是至关重要的一环。GAI 技术的使用能减轻教师答疑的负担，同时提供一对一定制化答疑服务。这意味着学生在学习过程中遇到问题时能得到针对性的、个性化的解决方案。例如，在教授编程语言如 Python 时，它首先解释变量的概念，接着逐步介绍函数、循环等更高级的概念。通过这些方式，GAI 成为了学生学习过程中的强大辅助工具，帮助学生克服学习障碍，提高学习效率。

3.2.5. 学情评估精准反馈

以 ChatGPT 为典型代表的 GAI 技术，能够实时提供个性化反馈，帮助学习者即时了解自己对知识点的掌握程度。及时反馈对于学生深入理解学习内容、识别并改进学习中的不足至关重要。更进一步，当 AI 通过分析学生过往的互动数据来理解其学习需求和习惯时，它能够不断优化反馈的质量，更加精准地满足每位学生的个性化学习需求。通过这种深度学习和个性化适应的过程，不仅提升了学生学习效率，还极大地增强了学习体验。例如，ChatGPT-4 能让学习者迅速创建多样化的角色扮演场景，通过在提示中引入变量，用户可以轻松切换语言和背景，实现实时互动。

4. 生成式人工智能赋能精准教学的实践框架

本文在前文已经阐述 GAI 赋能精准教学的实现路径，介绍了路径的实现基础和具体效果，但路径该如何在教学全过程中实施，仍然是有待解决的问题。为此，本文将详细介绍 GAI 赋能精准教学的实践框架，为 GAI 的实践提供具体的指导。

4.1. 设计思路

4.1.1. 打造整合 GAI 的教育数据平台

GAI 能在创造文本、图像和音频内容方面展现了卓越的能力，然而，若要在教育领域实现精准化教学，仅凭其单一能力显然还不能达到既定的要求。精准教学依赖于对每位学生的个性化需求和能力的深入理解，进而提供量身定制的学习体验。实现这一目标，需要数据平台的支持，该平台能够收集学生的行为模式、成绩、偏好以及反馈，并将这些数据提供给 GAI，以便有效地产出适宜的学习资源。数据平台的动态追踪功能确保所提供的教学内容与学生的实际水平和需求同步更新。更重要的是，数据平台能整合多源数据，为 GAI 构建一个全方位的数据视图，促进对学生学习过程的全面理解。通过持续记录学生的表现并评价教学成果，数据平台可建立起一个闭环反馈系统，以不断优化教学质量和效果。

4.1.2. 基于实践路径的教学全流程设计

基于 GAI 打造的教学全流程框架旨在深度实现个性化教育的理念，通过 GAI 在精准推送教学内容、设计多样化教学活动、生成个性化学习路径、提供情境性学习解答，以及进行动态学情评估与反馈的五大实践路径，充分挖掘每位学生的独特需求和潜能。此流程设计核心在于借助 GAI 定制化地为每位学生设计最适宜的学习材料和路径。这种方法不仅极大提升了

学习的效率和动机，还能够根据学生的实时学习进展灵活调整教学计划和难度，确保教学策略始终与学生的个人发展需求保持同步。进一步地，这一教学全流程框架通过提供实时、针对性的反馈和支持，加强了学生问题解决能力和自主学习能力，促进了学生全面发展。同时，它还通过精细化的学习数据分析，不仅让教师能够深入理解学生的学习状态，还能够持续优化教学方法和内容，推动教育资源高效利用和教学质量持续提升。

4.2. 具体框架

根据以上阐述的实践框架的设计思路，本文设计了 GAI 赋能精准教学的实践框架，旨在打造 GAI 赋能的教学全流程，实施个性化、精准化的课堂教学，如图 1 所示。其核心是围绕整合 GAI 的教育数据平台去实施教学中的一系列步骤，包含教学内容的精准推送、教学活动的精准设计、学习路径的精准生成、学习疑问的精准解答和学情评估的精准反馈。上一章节已论述这五条实践路径为教育教学带来的效果，接下将围绕这五个实践路径如何实施具体介绍。

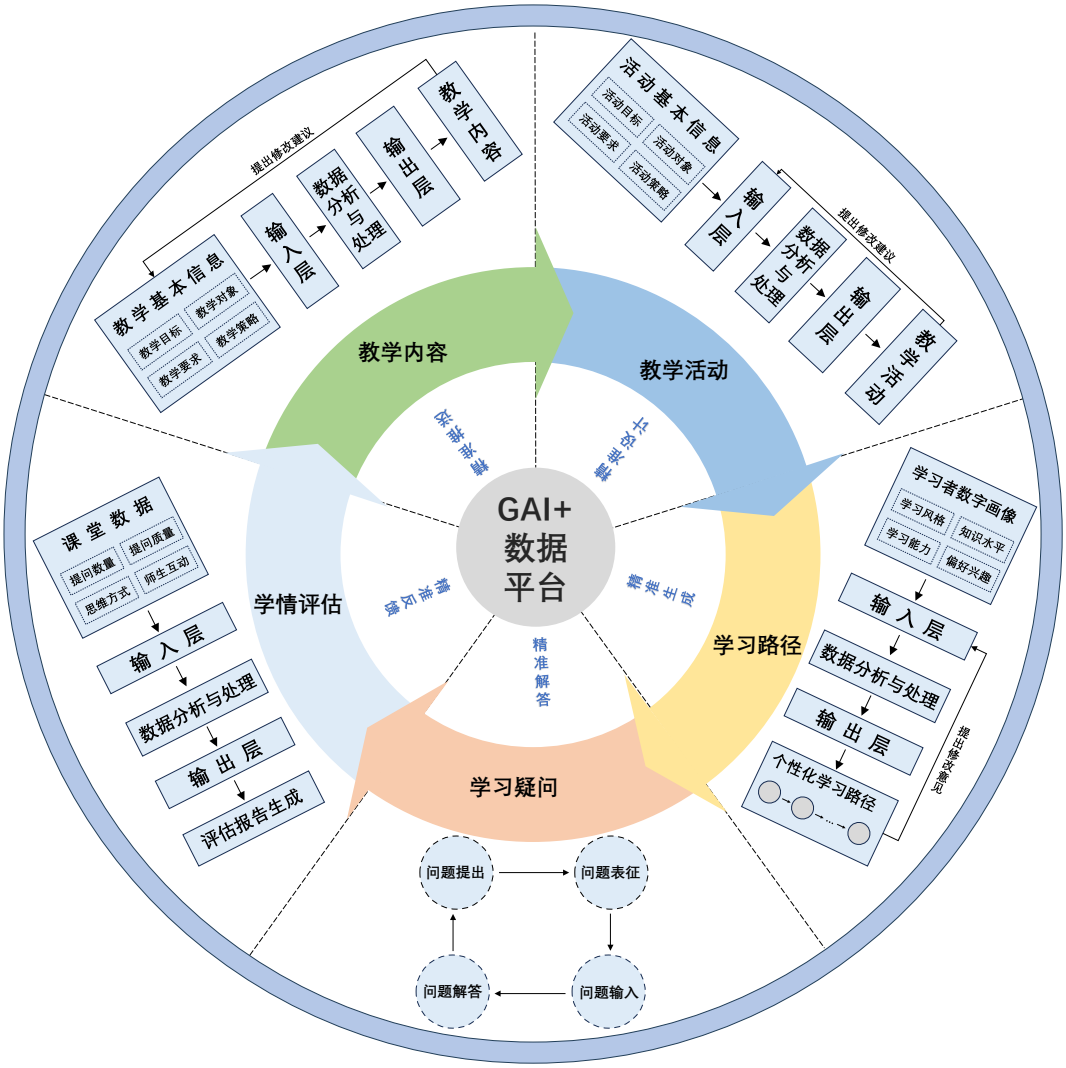


图 1 生成式人工智能赋能精准教学的实践框架

4.2.1. 生成式人工智能赋能教学内容精准推送

在当代教育实践中，精确推送定制化教学内容的追求旨在使每位学生均能获得与其个性化需求最匹配的教育资源，此目标的达成依托于 GAI 技术与教育数据平台的深度融合与应用。张一春等人也鼓励教师在教学中探索和使用生成式人工智能技术，为丰富教学内容提供技术手段、打造高效课堂(张一春, 钟秋菊, & 任屹远, 2024)。

首先，教育工作者需明确教学内容的基本信息，涵盖教学目标、对象、要求及策略等关

键维度，确保教育内容具备明确的方向性和目标性。上述教学基础信息被详细输入至教育数据平台，并开始个性化教学内容的生成过程。在此阶段，GAI 通过对教师提供的教学基础信息精准构建起初步的教学内容。随后经数据分析与处理，教学内容将展现于教师面前。最后，教师基于对教学内容的仔细审读，可能会根据实际教学情况提出调整意见，最终形成符合教学目标且深刻贴合学生个性化需求的教学资源。

4.2.2. 生成式人工智能赋能教学活动精准设计

在现代教育体系中，精准设计教学活动的目的是使教学活动能达到既定的教学目标，提升课堂的学习效果。此目标的实现依赖于 GAI 技术与教育数据平台的融合应用，体现了教育技术领域的创新与进步。

首先，教育工作者应对教学活动基础要素有一个明确的规划，包括活动目标、对象、要求及策略等，这些活动基本信息随后被输入至教育数据平台。随后经过数据处理与分析，数据平台将生成的教学活动呈现给教师。最后，教师根据初步结果进行审查，并可能根据实际情况和学生反馈提出调整意见。数据平台依据这些反馈再次进行数据处理和活动内容生成，最终形成符合教师预想的教学活动设计方案。

4.2.3. 生成式人工智能赋能学习路径精准生成

在当代教育实践中，精准生成个性化学习路径有助于丰富了学生的学习体验，提高学习效率和成效。每位学生因而能得到一条基于其个人特征定制的学习路径，预示着个性化学习发展的新趋势。例如汪靖等人指出，用户可在写作方面对 GAI 提出写作路径需求，为写作活动提供支架，助力写作效率的提升(汪靖 et al., 2023)。

首先，借助教育数据平台，学习者可以清晰明确自己的学习风格，了解自己数字画像，这包含了自己的学习风格、学习能力、学习水平和兴趣爱好等信息，随后，当学生请求生成个性化学习路径时，其数字画像即成为 GAI 进行数据分析与处理的关键依据，最后，学习者再根据生成的学习路径学习时的学习体验，可以给 GAI 提出评估和修改的建议，及时优化自己的学习路径，使其更具备个性化特征。

4.2.4. 生成式人工智能赋能学习疑问精准解答

在当代课堂实践过程中，精准且及时解答学生在学习过程中所遇到的疑问，有望提升学习过程的效率及教学互动的质量。

首先，学生在教师的引导下对所面临的学习难题进行精准解答与准确表征，学生将问题的具体描述输入教育数据平台，随后该平台通过对问题进行深入的分析和处理，输出有指导意义的答案，该答案可能正确也可能错误，故也要求学习者秉持批判对待人工智能技术的态度去解决学习疑问。

4.2.5. 生成式人工智能赋能学情评估精准反馈

在当代课堂实施过程中，实施精确的学情评估反馈机制是提升教育质量和优化学生学习成效的关键。GAI 的介入为构建一个能够实现高度个性化和精确度的学情评估体系提供了技术基础。

首先，通过对教育数据平台所收集到的课堂行为数据进行分析与处理，包含但不限于学生的提问频次、提问内容的深度、逻辑思维的展现及师生互动的质量与频率，随后当用户输入生成学情评估报告的指令后，GAI 会输出一份学情评估报告。该报告详尽反映了学生课堂行为的全貌，从提问的积极性、思维逻辑到社交互动，同时对学生在特定知识点的掌握程度进行精准评估，明确指出了学生的学习优势及需进一步加强的方面。

5. 结语

本文深入探讨了生成式人工智能如何赋能于精准教学，推进教育个性化学习的实现。面对教育领域日益增长的复杂性与个性化需求，传统教学已突显局限性，亟需借助新兴技术进行革新。生成式人工智能以其卓越的数据分析及内容生成能力，为精准教学提供了一系列解决方案，包括教育内容精准推送、教学活动精准设计、学习路径精准生成、学习疑问精准解

答以及学情评估精准反馈。

本文通过建立生成式人工智能赋能精准教学的实践路径与实践框架，旨在指导教育实践者如何有效地整合并应用生成式人工智能技术于精准教学中，同时为教育政策制定者提供了决策参考，以促进教育公平的实现。尽管生成式人工智能在赋能精准教学方面展示了巨大的潜能，但实践中也遭遇了诸多挑战，包括智能技术与教学实践的融合不足、教师在数据解读与干预实施方面的能力不足和数据隐私与安全性问题。为应对这些挑战，需采取多方面的措施，如优化生成式人工智能与教学实践的整合策略、提升教师的数据素养和加强数据隐私保护措施。

总之，该论文为理解与实施生成式人工智能在教育领域的应用提供一定程度的见解与实践指导。展望未来，随着技术的不断进步与教育实践的逐步深入，生成式人工智能将在推动教育公平、提升教学品质与效率方面扮演更为关键的角色。

参考文献

万力勇, 黄志芳, & 黄焕. (2019). 大数据驱动的精准教学:操作框架与实施路径. 现代教育技术, *29*(1), 31–37.

白雪梅 & 郭日发. (2024). 生成式人工智能何以赋能学习、能力与评价？. 现代教育技术, *34*(1), 55–63.

叶新东 & 刘泽民. (2024). 基于多模态大模型的精准教学支持体系构建研究. 远程教育杂志, *42*(1), 84–93. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2024.01.010

卢宇, 余京蕾, 陈鹏鹤, & 李沐云. (2023). 生成式人工智能的教育应用与展望——以 ChatGPT 系统为例. 中国远程教育, *43*(4), 24–31, 51. https://doi.org/10.13541/j.cnki.chinade.20230301.001

任红杰. (2017). 基于大数据的精准教学:生成路径与实现条件. 黑龙江高教研究, (9), 165–168.

毕文轩. (2023). 生成式人工智能对教育行业的挑战与回应——以 ChatGPT 为分析对象. 江苏高教, (8), 13–22. https://doi.org/10.13236/j.cnki.jshe.2023.08.002

李海峰 & 王炜. (2023). 生成式人工智能时代的学生作业设计与评价. 开放教育研究, *29*(3), 31–39. https://doi.org/10.13966/j.cnki.kfjyyj.2023.03.003

李颖婷. (2023). 生成式人工智能给图书馆带来的机遇、挑战及应对策略. 图书与情报, (2), 42–48.

张一春, 钟秋菊, & 任屹远. (2024). 高校教学信息化创新发展的核心内容与实践进路——基于教育数字化转型的 TASH 视角. 电化教育研究, *45*(2), 71–76, 83. https://doi.org/10.13811/j.cnki.eer.2024.02.010

汪靖, 米尔外提·卡马勒江, & 杨玉芹. (2023). 人机共生的复合脑：基于生成式人工智能辅助写作教学的应用发展及模式创新. 远程教育杂志, *41*(4), 37–44. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2023.04.004

陈雄辉, 谷紫阳, 覃以凤, 马宗兵, & 龙娜. (2022). 新时代教育信息化人文价值的实现路径. 中国电化教育, (9), 24–29, 83.

段沙 & 周怡. (2017). 精准教学文献综述. 英语教师, *17*(24), 64–70.

周洪宇 & 常顺利. (2023). 生成式人工智能嵌入高等教育的未来图景、潜在风险及其治理. 现代教育管理, (11), 1–12. https://doi.org/10.16697/j.1674-5485.2023.11.001

苗逢春. (2023). 生成式人工智能技术原理及其教育适用性考证. 现代教育技术, *33*(11), 5–18.

赵元成 & 何叶. (2024). 生成式人工智能模式下的文化传播：变革、风险与治理. 理论导刊, (3), 118–124.

郭利明, 杨现民, & 张瑶. (2021). 数据驱动的精准教学五维支持服务框架设计与实践研究. 电化教育研究, *42*(4), 85–92. https://doi.org/10.13811/j.cnki.eer.2021.04.012

郝建江 & 郭炯. (2022). 智能技术赋能精准教学的实现逻辑. 电化教育研究, *43*(6), 122–128. https://doi.org/10.13811/j.cnki.eer.2022.06.016
唐亚汇. (2021). 大规模预训练模型对 AI 应用与普及的影响. 上海信息化, (9), 52–55.
彭红超 & 祝智庭. (2016). 面向智慧学习的精准教学活动生成性设计. 电化教育研究, *37*(8), 53–62. https://doi.org/10.13811/j.cnki.eer.2016.08.008
潘巧明 & 赵静华. (2019). 区域精准教学改革实践的探索与研究——以丽水市精准教学改革实践为例. 电化教育研究, *40*(12), 108–114. https://doi.org/10.13811/j.cnki.eer.2019.12.015
Kulkarni, S. V. (2023). 11. Future of Technology ChatGPT: Optimizing Language Models for Dialogue. *ICP Monogram on Digital Technology in Clinical Medicine*, 62.
Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis*, *25*(1), 51.
Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., … Metzler, D. (2022). Emergent abilities of large language models. *arXiv Preprint arXiv:2206.07682*.

基于人工智能教育机器人的中学实验安全课程设计与开发

Design and Development of Experimental Safety Course in Middle School Based on Artificial Intelligence Education Robot

林晓婧[1]，连丽雯[1]，吴仙勇[1*]
[1]福建师范大学教育学院
*wuxianyong1999@163.com

【摘要】 实验室环境是中学教学和科学研究的重要场所。然而传统实验室安全课程教育过程中不仅学生学习积极性不高，教师也需要承担大量教学工作。随着人工智能的不断发展，人工智能教育机器人（AIER）不仅可以增强学习者交互体验，还能够降低教师工作负荷。但技术并不是教育的万能药，长时间使用机器人进行学习会降低学习者学习兴趣。因此本研究将游戏化策略融入 AIER 中来设计一节中学实验安全课程，以探究突破传统中学实验安全教育所面临困境的方式。

【关键词】 人工智能教育机器人；游戏化；实验室安全

Abstract*: Laboratory environment is an important place for middle school teaching and scientific research. However, in the process of traditional laboratory safety course education, not only students are not enthusiastic , but also teachers need to undertake too much teaching work. With the development of artificial intelligence, artificial intelligence educational robot (AIER) can not only enhance the interactive experience of learners, but also reduce the workload of teachers. However, technology is not a panacea for education, and using robots for a long time will reduce learners' interest in learning. Therefore, this study integrates gamification strategy into AIER to design a middle school experimental safety course, in order to explore ways to break through the difficulties faced by traditional experimental safety education.*

Keywords: Artificial intelligence, Education robot, Gamification, Laboratory safety

1.引言

人工智能(Artificial Intelligence, AI)的发展为教育领域带来了全新视野。具有一定智能水平的人工智能教育机器人（Artificial Intelligence Education Robot, AIER）在教育过程中不仅可以作为教师助手答疑解惑，还可以作为学习同伴协助管理时间和任务、参与或引导互动以及在学习过程中为学习者提供情感交互与具象社交（黄荣怀等人, 2023）。以 AIER 为载体的学习方式可以减少学生学习时的孤独感、提高沟通技巧、增加学生学习兴趣。然而，没有对 AIER 进行合理设计和应用将难以发挥其优势。相比于虚拟机器人，充当教学辅助工具、智能教师、学习同伴等角色的实体机器人，在与学习者交互过程中能更好与学习者交互，并强化学习者积极学习行为和情感结果(Belpaeme et al., 2018)。

实验室作为中学教学和科学探究的重要场所，实验室安全事故频发。然而，传统教学方式受时间和地点的限制，导致学习者对安全教育提不起兴趣。甚至，加重教师的工作负担(Viitaharju et al., 2021)。而教育机器人在优化学习者学习体验，提高教师教学效率上发挥积极作用（单俊豪等人, 2019）。但有研究表明，学习者长时间使用机器人进行练习不仅不能提高学习者学习的兴趣，还会适得其反(Fryer et al., 2017)。游戏化作为增加学习者学习兴趣的有效

途径，可将其元素纳入非游戏化的教学环境中以提高学生学习投入度(Hanus & Fox, 2015)。且游戏化元素中的成就解锁是学习者学习过程中保持学习积极性的有效途径(Groening & Binnewies, 2019)。

游戏化的教学策略已经在多门学科教学中表明其有效性，但很少研究将游戏化策略应用于AIER在实验室安全领域的开发中。为了提高学习者学习动机、减少教师工作量，本研究尝试将AIER与游戏化元素相结合，开发一节中学实验室安全课程。并拟通过准实验研究来验证本研究所提方法的有效性，即相较于传统的教学方式而言，本研究所提方法是否能够帮助学习者更好地掌握实验室安全教育知识，提高其学习动机。与此同时，本研究拟通过半结构化访谈的方式来进一步了解学习者学习感受。

2.基于游戏化AIER系统的结构

本研究关于游戏化的实验室安全系统包括学习内容模块、互动练习模块、游戏化学习模块、学习材料显示模块和积分排行模块，以及语音识别数据库、图像识别数据库、学习材料数据库、游戏内容数据库和积分排行数据库，如图1所示。学习者两人一组进入学习，学习内容模块为学习者提供闯关前的先备知识。学习材料显示模块可以将知识以视频、语音交互、图画和文本的方式传递给学习者。互动练习模块为学习者提供答题环境。当学习者完成相应知识内容的学习后进入互动练习模块进行竞争答题或协作答题。游戏化模块增加学习的趣味性，在组内两位学习者竞争答题过程中对每个学习者答题情况进行计分后分出胜负，并根据学习者答题情况授予相应徽章。学习者在学习过程中积分排行模块自动记录每组学习者在每一道题上所获得的积分以及在每个关卡结束后每组学习者所获得的总积分。

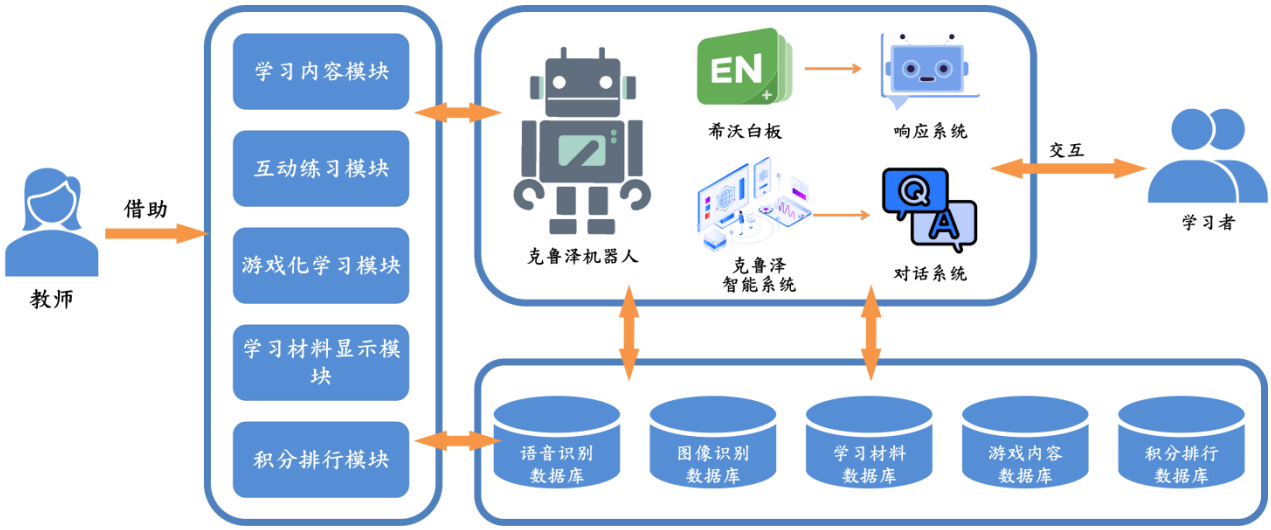


图1 系统结构图

AIER的智能对话系统除了具有语音交互、人脸识别、人脸跟随、声源定位转向等功能外，还可以借助躯体与学习者互动。当AIER利用人脸识别检测到人脸时，便由待机模式转换为唤醒模式，并播报对应欢迎语和展示欢迎动作。AIER的语音交互功能可以快速识别学习者语言，并通过自然语言处理系统识别精准和模糊关键词，再根据关键词搜索系统内开发者拟定的知识数据库显示相应的视频、图片、文本、语音和肢体动作。AIER的定制问答界面将学习内容分为消防安全、电气安全、实验室安全、警示教育和通识安全五个问题类别。为了方便学习者使用自己所习惯的方式提问，每个问题类别中包含15-20条精确问答和4-6条模糊词条。

3.基于游戏化 AIER 系统的内容与框架

游戏化策略作为教育创新的驱动力之一，希望通过引入游戏元素激发学生学习动机。为提升游戏化学习的有效性，Huang 和 Hew(2018)将自我决定理论、心流理论、目标设定理论、社会比较理论和行为强化理论进行融合，进一步提出了 GAFCC 动机理论模型，为游戏化策略与技术的融合提供了理论框架。该模型将促进动机体验的核心要素进行概述，并把动机需求总结为五个关键要素：目标(Goal)、访问(Access)、反馈(Feedback)、挑战(Challenges)和协作(Collaboration)。基于该模型来设计教学流程，可以为学习者提供个性化的挑战和任务，使学习者更加投入到学习过程中。如图 2 所示，积分、闯关、竞争等游戏化元素可以将动机需求的目标、访问、反馈、挑战、协作五个基本要素具体化。并且，每组学习者在学习过程中可以将游戏化元素中的积分和徽章作为学习目标，不断解锁新的关卡。与此同时，AIER 也会根据学习者在游戏化的学习过程中的表现给予即时反馈。

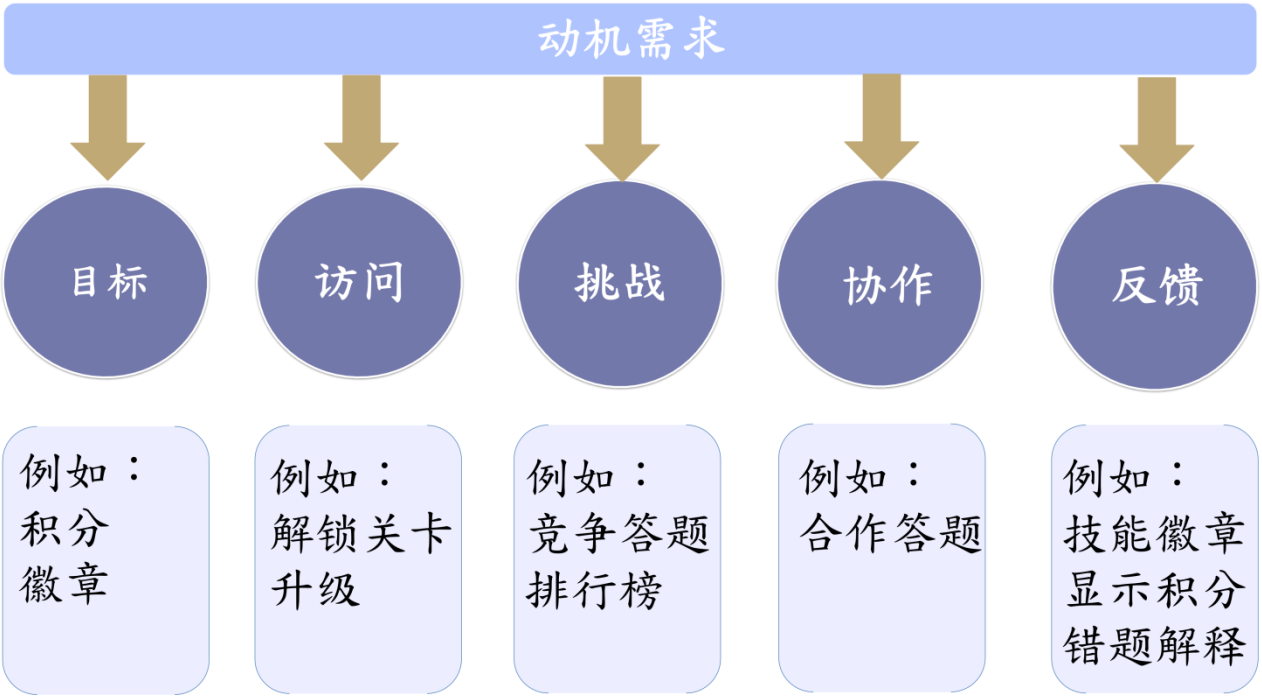


图 2 GAFCC 模型的游戏化设计

学习者在学习完 AIER 中以故事化方式呈现的基础知识后便开启闯关任务。在闯关开始前，两位学习者讨论并设置自己想要获得的目标积分，在设置目标结束后，介绍积分积累与徽章颁发的规则。学习者初始积分为 0，在答对一道题目时，积分加 10，答错一道题目时系统暂停答题并及时给予相应知识点反馈，而反馈一次积分将减少 5 分。每个关卡有一次与同伴进行竞争答题的机会，竞争答题过程系统分别记录两位学习者答题情况并判断胜负。不同题目涉及不同知识点不同，学习者在学习过程中所错误题目可能不同，所获得的反馈也不同。此外，学习者在答题过程中可以借助 AIRE 的语音交互功能向系统寻求帮助或者与同伴相互探讨以进行合作答题。在学习者答题过程中，系统不仅记录每组学习者每一道题所获得的积分，也记录每一组学习者在每一个关卡的总积分，并将积分获得情况上传数据库。学习者在学习过程中还可以借助进度条了解当前答题进度，及时调整答题状态。当完成一个关卡时，系统将显示该组学习者分数、在当前积分数据库中排行以及目标积分与实际积分的差值。如若学习者对该阶段学习内容存在疑问，可以选择打开数据库了解具体答题情况。该模块的呈现，可以帮助学习者反思学习过程，进而调整自身闯关状态。基于游戏化的 AIER 系统设置五个

关卡，五个关卡分别对应五枚徽章，五枚徽章分别为电气安全徽章、实验室安全徽章、警示教育徽章、消防安全徽章、通识安全教育徽章。



图 3 学习界面

4.基于游戏化 AIER 系统的课程设计以及评估流程

基于游戏化的 AIER 学习系统借助 Cruzr Business Intelligence System 和 Seewo APP 进行中学实验安全课程开发。开发过程遵循以下五阶段，如图 4 所示。

4.1 设置目标

该系统开发主要用以培养学习者自身安全意识、教授实验室安全守则、防止实验室安全事故发生。并且，该系统在实际教学应用过程中需要克服传统学习方式下学习者学习参与感不高、教师工作量大的困境。因此，本研究将长期目标分解为若干个短期目标，长期目标的实现依托各个短期目标的完成。课程设计分为基础知识学习与游戏化闯关两部分，学习者在完成基础知识学习这一短期目标后便进入游戏化闯关。游戏化闯关包含五个关卡，学习者在每个关卡开始前设置短期目标，例如获得的目标积分。通关后系统颁发徽章以激励学学者继续学习完成下一阶段的短期目标。而学习者在获得五枚徽章后便顺利通关实现长期目标。

4.2 制定计划

为了实现教学目标，在课程正式开发前，选择合适的人工智能技术和游戏化元素应用于课程内容的开发。且在计划制定过程中也应考虑中学生学习倾向、个性偏好，并利用学习者学习兴趣制定课程内容。例如，课程开发前，向学习者分发问卷以了解学习者对学习材料的偏好、学习过程中遇到的困难以及所感兴趣的游戏化元素。了解学习者需求后，将知识内容细分为电气安全、实验室安全、警示教育、消防安全、通识安全教育五大模块，将教学环节划分为基础知识传授与知识应用。基础知识传授主要通过机器人界面以故事化的形式向学习者呈现文本、图片、视频等材料。而知识应用模块借助机器人的交互面板向学习者呈现知识点对应选择题、填空题以及连线题等题型。

4.3 课程教学

4.3.1 课程内容与 GAFCC 模型的结合

将游戏化中竞争、挑战、故事化等元素与教学计划阶段制定的教学内容相融合，并在 AIER 上具体化。(1)目标：本研究采用积分、徽章的游戏化元素实现 GAFCC 中的目标模块，学习者在学习过程中以获得目标积分和相应的荣誉徽章为目的，不断激励自身进一步学习；(2)访问：该课程设计将知识内容划分为五大关卡并为学习者提供合适的挑战过程，即学习者通过解锁关卡，由易到难的学习新知识；(3)挑战：每个关卡内部不仅设置同伴间竞争答题培养个体学习者竞争学习意识，还设置排行榜以进行组间竞争；(4)协作：学习者两人一组通过与机器人进行交互以学习、应用知识。且学习过程中同伴间可以互相探讨、协商。当学习者对知识点存在疑惑时，可以借助 AIER 的语音交互功能向机器人寻求帮助。协作学习过程不仅培养学习者合作意识，还可以通过讨论发展深层思维能力；(5)反馈：当学习者借助交互面板完成答题后，系统自动对比后台正确答案，借助面板、肢体语言以及语音交互及时给予学习者反馈。例如，答错后机器人展示出悲伤的表情和动作，并模仿教师肢体动作向学习者解释该题错误原因。

4.3.2 课程安排

第一步，开启机器人，机器人向学习者介绍本节课的课程总目标与课程内容后，学习者点击开始学习进入界面。第二步，两位学习者共同接受关卡任务、了解积分规则、讨论并设置目标积分后边进入学习界面。第三步，机器人向学习者呈现文本、图片、视频等学习材料，学习者在学习过程中不仅可以相互交流讨论，还可自主向机器人提问，以获得所需学习资料。第四步，两位学习者在完成基础知识后便开启合作答题，合作答题过程中遇到任何问题可通过语音向机器人寻求帮助。题目类型包括选择题、填空题、连线题等以满足不同学习者偏好。并且，答题过程系统可以根据学习者答题情况给予及时反馈。第五步，两位学习者进行合作答题后便进行竞争答题，竞争答题过程不仅考验学习者答题准确率，也考验学习者答题效率。引入竞争元素的竞争答题激发学习者竞争意识，促使学习者在有限时间内迅速而准确地回答问题。第六步，完成答题后，系统告知学习者所获积分及目标积分与实际积分直接地差距，并向学习者颁发该关卡对应徽章。

4.4 监控设计

课程开发、教学过程紧密围绕已设定的安全教育目标，防止教学过程偏离安全教育方向。且安全教育设计过程以学习者为主体，围绕安全问题设计活动内容与任务，以确保课程能够有效辅助教学和提高学生实验室安全意识。本研究在系统设计过程中导入游戏化学习管理系统实时监控学习者学习进度和学习目标的完成情况。学习者在学习学习过程中，也可根据答题界面中的进度栏监控自身学习状况，以便调整学习状态。该系统所具备的数据分析功能不仅可支持教师对学习者学习活动的实时监控，还可对学习者学习表现进行评估。

4.5 评估课程

将开发好的教育机器人教学系统部署到实际的教学环境中，进行试点运行。在试点运行期间，收集反馈信息，评估课程实际教学效果，包括学生的学习效果、学习动机以及通过这种方式进行学习的具体感受。

4.5.1 评估方式

本研究拟采取定量分析和定性分析的两种评估方式对教学效果进行检验。首先，分别在实验开始的前后，测验两组学习者关于实验室安全知识的掌握情况，以及对实验室安全教育课程的学习动机，学习者被要求在 40 分钟的时间内完成相对应的测试题(包含试卷和量表)。通过数据处理来了解游戏化 AI 教育机器人对学习者的学习效果以及学习动机的影响。其次，本研究拟通过对学习者进行访谈，以深入理解学习者在学习过程中的体验与感受。

4.5.2 测量工具

通过“实验室安全知识前后测量试卷”来检测学习者对知识的掌握情况；前测量试卷的填写是在学习实验之前进行，目的是评估学习者已有知识，即他们对实验室安全知识的掌握情况，并且了解不同组别受试者之间是否存在差异，确保每组的同质性。测试卷根据学校实验室安全教育考试的试题数据库开发的，确保所选知识内容的前、后测试能够充分评估学生的学习成绩。而后测试卷题型和分值与前测试卷相似，以本节课的课程内容和知识为基础并在实验结束后进行填写，以评估学习者在学习活动中对知识的掌握情况。

借由“学习动机量表”来检测学习者在学习过程中的学习动机；学习动机量表由 Wang 和 Chen 在 Pintrich 等人的基础改编而来(Pintrich, 1991; Wang & Chen, 2010)。这项测试包括 6 个项目，采用李克特 5 点量表，从 1 到 5 同意程度依次递增。测量维度分为内在动机和外在动机。在课程前后进行测量，观察学生在实验前后学习动机的变化，克朗巴赫 α 值为 0.79。例如，“在本课程中我比较喜欢有挑战性的教材，因为这样我可以学习到新的事物”和“在本课程中得到好成绩，对我来说是最满足的事情”

通过半结构化访谈的方式来了解学习者在学习过程中的感受，访谈提纲改编自 Hwang 等人设计的半结构性访谈提纲(Hwang et al., 2009)。该访谈有七个开放性问题例如“这种游戏化的 AI 教育机器人与你以前学过的课程有什么不同?这种学习方法的整体优势是什么？”等。本研究随机从不同组别中抽取等量的学习者作为访谈对象，了解他们在相应学习环境中的想法和感受，访谈内容通过录音的方式进行记录。

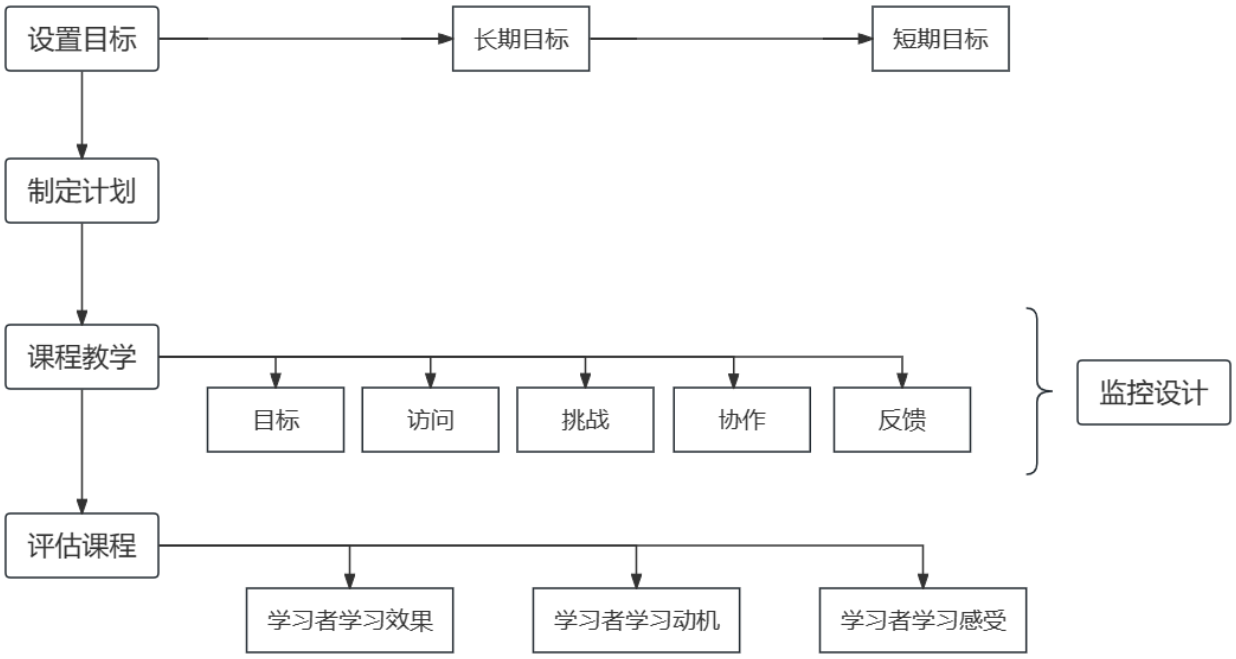


图 4 课程设计流程图

5.建议与展望

基于游戏化的 AIER 系统为中学实验室安全教育提供了新的可能性。通过 AIER 辅助的中学实验安全教育，不仅可以增加学生的实践操作能力，还可以利用 AI 技术获得可视化的交互体验。因此，本研究基于本次系统设计提出以下建议：充分利用人工智能技术，发挥游戏化策略优势。当前社会，人工智能正从影响教育变革的外在推力转变为引发教育变革的内生动力（胡艺龄和陈彦君,2024）。而新技术的发展也为游戏化在教学中的应用提供了新的机遇和

挑战。在中学实验安全课程设计过程中，应充分将 AIER 功能与游戏化元素相结合，最大化提高教学成效。

探索人机协同教育应用新模式。以 AIER 为代表的人工智能具有高速运算能力、大数据处理能力和模式识别能力。而教师具有人工智能所不具备的创新思维、复杂决策能力和价值判断能力（詹海宝等人, 2023）。因此，人机协同的方式在课程教学中具有巨大潜力。然而，在实际应用过程中应考虑学习活动的多样性、学习者认知过程的科学性等要素以设计出符合教学规律的教学过程（卢宇等人, 2020）。

加入个性化学习路径设计。基于知识图谱和机器学习构建的学习者模型可以为学习者提供个性化学习方式（卢宇等人, 2020）。因此，建议在游戏化人工智能教育机器人中加入更多个性化学习方式，以满足学习者学习需求和偏好，进而提高学习参与度。

参考文献

卢宇, 薛天琪, 陈鹏鹤, & 余胜泉. (2020). 智能教育机器人系统构建及关键技术——以“智慧学伴”机器人为例. 开放教育研究, *26*(2), 83–91. https://doi.org/10.13966/j.cnki.kfjyyj.2020.02.008

单俊豪, 宫玲玲, 李玉, & 闫寒冰. (2019). 教育机器人对学生学习成果的影响——基于 49 篇实验或准实验研究论文的元分析. 中国电化教育, *5*, 76–83.

胡艺龄, & 陈彦君. (2024). 人机协作下的共享心智模型研究. 现代教育技术, *34*(1), 64–72.

黄荣怀, 刘德建, 提利利阿罕默德, 张国良, 陈莺, & 王欢欢. (2023). 人机协同教学:基于虚拟化身、数字孪生和教育机器人场景的路径设计. 开放教育研究, *29*(6), 4–14. https://doi.org/10.13966/j.cnki.kfjyyj.2023.06.001

詹海宝, 孙江, & 郭梦圆. (2023). 基于人机协同的网络文化安全治理新模式. 传媒, *15*, 75–77.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, *75*, 461–468. https://doi.org/10.1016/j.chb.2017.05.045

Groening, C., & Binnewies, C. (2019). “Achievement unlocked!”—The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior*, *97*(November 2018), 151–166. https://doi.org/10.1016/j.chb.2019.02.026

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers and Education*, *80*, 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education*, *125*, 254–272. https://doi.org/10.1016/j.compedu.2018.06.018

Hwang, G. J., Chang, S. C., Chen, P. Y., & Chen, X. Y. (2018). Effects of integrating an active learning-promoting mechanism into location-based real-world learning environments on students' learning performances and behaviors. *Educational Technology Research and Development*, *66*(2), 451–474. https://doi.org/10.1007/s11423-017-9567-5

Hwang, G.-J., Yang, T.-C., Tsai, C.-C., & Yang, S. J. H. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & Education*, *53*(2), 402–413. https://doi.org/10.1016/j.compedu.2009.02.016

Pintrich, P. R., & Others, A. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*.

Viitaharju, P., Yliniemi, K., Nieminen, M., & Karttunen, A. J. (2021). Learning experiences from digital laboratory safety training. *Education for Chemical Engineers*, *34*, 87–93. https://doi.org/10.1016/j.ece.2020.11.009

Wang, L., & Chen, M. (2010). The effects of game strategy and preference-matching on flow experience and programming performance in game-based learning. *Innovations in Education and Teaching International*, *47*(1), 39–52. https://doi.org/10.1080/14703290903525838

基于 Citespace 的作文自动评价系统研究的可视化分析与概述

Visual Analysis and Overview of the Research of Automated Writing Evaluation Based on CiteSpace

Zheng Kunhan[1*]，Huang Biyun[2]，Xie Shiyi[3]
[123]School of Education, City University of Macau
[*] kunhan.cheng@gmail.com

【摘要】 随着科技的发展进步，计算机技术已被广泛应用于语言测试领域，而过程教学法的兴起和形成性评价理论的发展，也使得自动评分这一新型的反馈方式正被越来越多地运用到主观题测试中。目前，作文自动评价系统的研究与应用也逐渐成为国内外计算机辅助写作教学研究领域的热点。本文运用 Citespace 软件对国内学者关于作文自动评价系统研究、作文自动评价系统在现实教学情境中的应用和对学生之影响三个方面进行可视化分析和概述,并探讨其未来发展方向，希冀为相关学者和一线教师提供帮助和参考。
【关键词】 作文自动评价系统；Citespace；可视化分析；外语教学

Abstract: *With the development and progress of science and technology, computer technology has been broadly used in the field of language testing. And the rise of process method and the development of formative assessment theory also make automated evaluation, a new type of feedback method, is being used more and more in subjective tests. At present, the research and application of automated writing evaluation system has also gradually become a hot spot in the research field of computer-aided writing teaching at home and abroad. In this research, we use Citespace to conduct visual analysis and overview of the research of domestic scholars on the automated writing evaluation system, the application of automated writing evaluation system in the real teaching situation and the impact on students, and to explore the direction of its future development in the hope that it will provide help and reference for the relevant scholars and teachers.*

Keywords: Automated Writing Evaluation, Citespace, visual analysis, foreign language teaching

1. 作文自动评价系统之介绍与发展

1.1. 作文自动评价系统之介绍

作文自动评价系统 Automated Writing Evaluation（AWE）是一种利用信息自动化手段对学生作文进行客观评判与打分的工具（Shermis＆Burstein, 2013）。AWE 的诞生最初是为了减轻教师的评鉴工作量，并提升对中英文写作评鉴上的客观性。目前，得益于 AWE 在为学生生成个性化的反馈和建议功能方面的发展，其已发展成为教育工作者和学习者于课堂内外使用的常用工具，与此同时，它也是一种利用当前社会的写作理论和学习理论进行创新性评价的重要手段。

目前的主流 AWE 系统主要是通过运用计算语言学方法来提取语言元素作为评价特征，从而模拟人工评分。这一过程包括评分标准的收集、文本预处理、提取特征、构建评分模型和生成分数、反馈及个性化建议这几个环节，如图 1 所示。

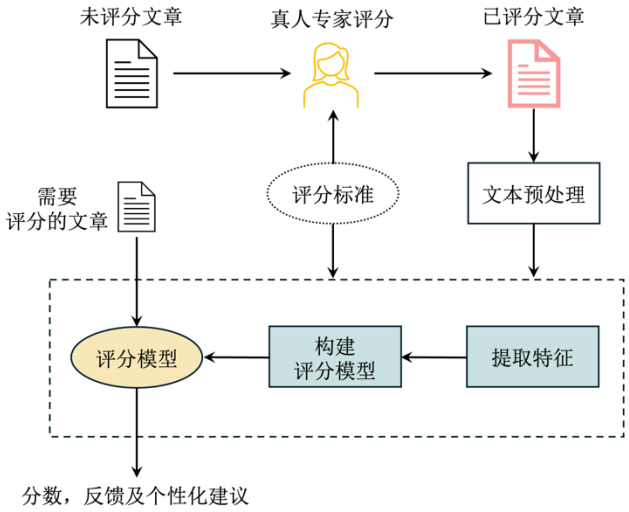


图 1 作文自动评价系统之工作流程

1.2. 作文自动评价系统之发展

作文自动评价系统起源于美国，纵观其发展脉络，可概括为以下三个关键阶段：

第一阶段起始于二十世纪六十年代，美国杜克大学 Ellis Page（1968）等人所联合研发的 Project Essay Grander（PEG），成为全球首个作文自动评分系统。PEG 可对一篇文章的可测量特征如长度以及平均句长进行多元回归分析，然后基于人工评判的大量作文语料库建立评分模型。

随着二十世纪九十年代的到来，作文自动评价系统迎来了第二阶段的发展。在这一阶段，得益于计算机与人工智能技术的迅猛发展，涌现出了基于自然语言处理、人工智能以及潜伏语义技术的作文自动评价系统。同 PEG 类似，这些系统也是基于人工评判的大量作文语料库建立评分模型，但是包含更多有关词法、句法以及语篇内容和结构的元素，分析更加复杂。

进入二十一世纪，作文自动评价系统步入了第三个发展阶段。在第二阶段的研发基础和技术积累之上，衍生出了 Holt Online Essay Scoring、Criterion 等综合评价系统。这些系统不仅具备了自动评分功能，还能提供个性化反馈及英语写作辅助，进一步丰富了作文自动评价系统的功能与应用场景。

目前，国外的主流作文自动评价系统包括：E-rater，Criterion，My Access!，Project Essay Grade 和 Intelligent Essay Assessor 等。而国内主流的作文自动评价系统有大规模考试英语作文自动评价系统（EFL Essay Evaluator）、句酷批改网、冰果英语智能作文评阅系统、TRP 教学资源平台、易改等。

2. 国内学者关于作文自动评价系统之研究

本文对作文自动评价系统的相关文献回顾与分析之数据均来自中国知网（CNKI）论文总库，在 CNKI 中选择“高级检索”类型，选择“主题”检索，检索条件为“作文自动评价”、“AWE”、“AEE”、“自动评价”等与作文自动评价系统相关的关键词进行搜索，截止到 2024 年 3 月 28 日，共检索出 119 篇相关文献,通过手工筛选，剔除会议、报纸、图书及相关性较低的文章后，共得到 110 篇论文。其中，研究作文自动评价系统之原理和信效度的论文共 55 篇；使用作文自动评价系统在实际教学情境中进行具体应用的论文共 29 篇；从学生角度出发、以学生为主体，探讨学生在使用作文自动评价系统后的反馈的论文共 26 篇，如图 2 所示。

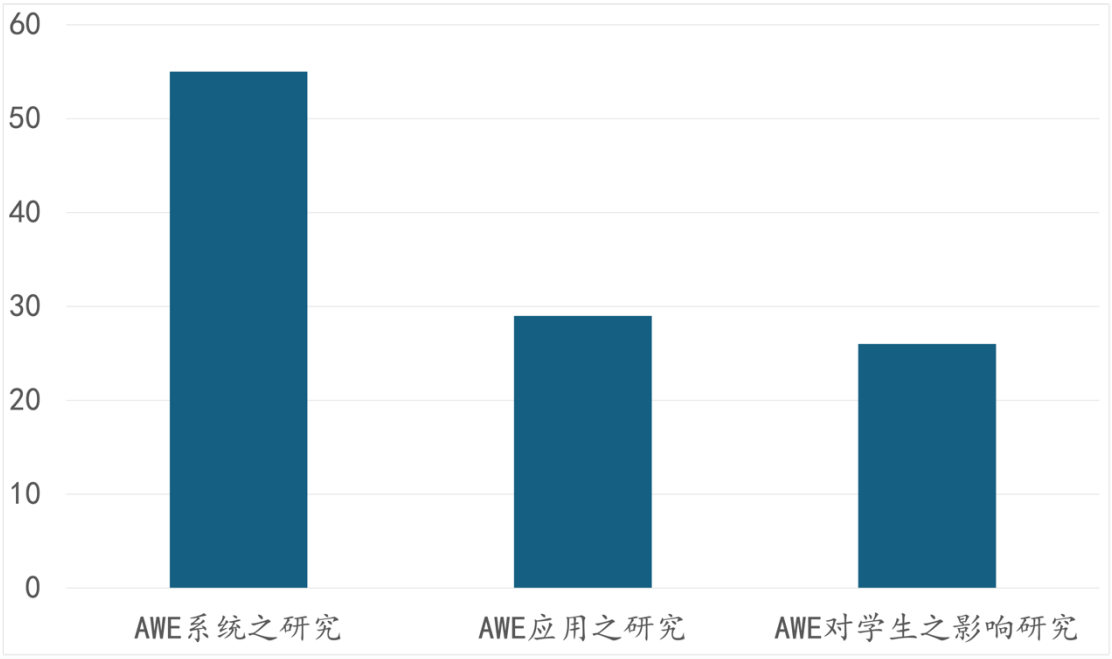


图 2 国内学者关于作文自动评价系统论文研究之分类

关于作文自动评分系统的研究指的是国内学者对作文自动评价系统之原理和信效度的相关研究。我们将介绍作文自动评价系统之原理及有效性相关的论文导入 CiteSpace 6.3.R1 软件，以“Keyword”为节点，关键词出现频率的阈值（Threshold）设置为 5，用 Citespace 进行可视化运算，运算后得到由 145 个节点和 388 条连线构成的关键词共现知识图谱（如图 3 所示）。关键词使用频率越高，则字体越大。在文献收集过程中，由于选取“作文自动评分系统”作为检索词，故剔除在关键词共现知识图谱中的“作文自动评分”“自动评分”“自动作文评分”等不属于分析范畴的关键词。进一步分析得到词频排名前 3 的关键词，分别为自然语言处理、信度和效度。

具体来看，张荔（2017）从测量性、归纳性和外推性三方面对批改网的效度进行验证。结果显示,批改网的人机评分同一分数档内的完全一致性以及完全加相邻分数档一致性与国外同类 AWE 系统基本相似,人机评分显著相关,说明其具有一定的可靠性。陈小芳（2017）借助句法复杂度分析器（L2SCA）对学生提交的初稿与终稿文本的分析,从句法复杂度的角度探讨了自动写作评分系统对学生作文输出的影响。通过 T 检验进一步证实自动写作评分系统会促使学生产出更为复杂的句子,尤其在字数和从属从句的运用上更为明显。高健民（2021）考察了系统自动评分与人工评分的一致性及系统自动评分与文本特征指标的相关性后,发现批改网自动评分结果仅能有效反映学生作文的词汇复杂度和流畅度。张国强和何芳（2022）使用 L2 Lexical Complexity Analyzer、L2 Syntactic Complexity Analyzer 及 Coh-Metrix 软件分析了 826 篇不同类型写作任务的大学英语六级应试作文，探究英语自动评分系统对不同类型写作任务评阅的信度和效度。作文评阅过程中,系统更关注写作规范、高级词汇、长句、复杂句及其句式的变化等,而严重淡化句法错误、语篇意义连贯性等文本特征,这无疑会对作文评阅的质量造成不利影响。综合而言，我国部分学者在对待作文自动评价系统时更加关注其有效性、可信度和稳定性。

近年来，随着技术的不断进步和应用场景的不断拓展，我国学者对作文自动评价系统的研究也在不断深入和拓展。张锐捷等人（2019）研究与实现的英语作文自动评分系统，可进行提取作文内容等一系列特征进行综合考评,使用自然语言处理、文本特征提取、潜在语法分析等技术,实现英语作文自动评分功能。此系统还可通过添加自定义的停用词词典,而扩展运用到更多的场景。李俊娴（2020）所设计与实现的基于多特征融合和机器学习理论的英语机考作

文自动评分系统。在硬件设计方面设计了控制电路实现对评分系统的全局控制。软件设计方面采用多特征融合技术跟踪英文词汇,并计算目标词汇和评分标准词汇之间的相似度;利用机器学习中的逻辑回归算法,设计作文自动评分逻辑。夏林中等人（2020）研究了一种基于新型神经网络结构的自动作文评分模型。实验结果表明,对比其他基准模型(如双向 LSTM 模型和 SKIPFLOW-LSTM 模型等),基于注意力机制的双层 LSTM 模型所获二次加权 kappa 系数平均值最好。

由此可见，我国学者的研究也正进一步地对作文自动评价系统进行改进和优化，且希冀探索一种更加本地化、更加适用中国国情的应用场景。

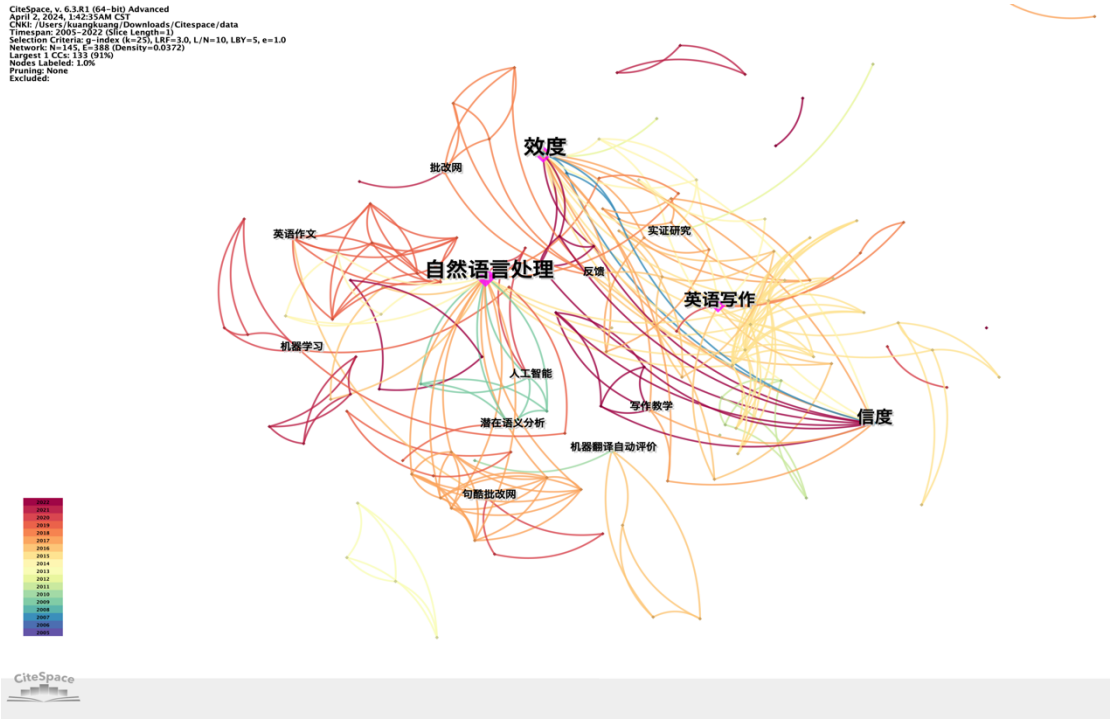


图 3 国内学者关于作文自动评价系统之研究的关键词共现知识图谱

3. 国内学者关于作文自动评价系统之应用研究

作文自动评价系统之应用研究指的是我国学者对作文自动评价系统在实际教学情景中的具体应用的相关研究。同样的，将使用作文自动评价系统进行应用研究类的论文导入到 CiteSpace 软件进行可视化运算后，得到关键词共现知识图谱（如图 4 所示）。进一步分析得到词频排名前 3 的关键词，分别是反馈、过程写作和批改网。

具体来看，王婷和王芹（2016）的研究以句库批改网为例,抽取新疆农业大学外国语学院 2013 级英语专业学生中不同分数的作文,从中找出作文中存在的差异。研究结果表明,作文分数的高低受词汇及句子的影响,同时发现,在题目的要求中没有限制字数时,分数与字数存在正比关系。唐芳等人（2017）同样以句酷批改网平台为例，以应用中的实际例子说明该自动评分系统的一些优缺点,然后提出如何有效利用自动评分系统进行大学英语写作教学。柯育强（2017）通过研究大学英语作文评分难题,结合文本聚类理论,完善其应用,将聚集不到一起差异性较大的可能跑题作文交由老师评判,可大大减少教师工作力量,提高作文评判的准确性和效率,对大学英语作文教学具有良好的促进作用。宋毅宁（2019）依托本土化的自动作文评分系统批改网,构建了大学英语数字化写作平台,并对其应用效果进行了实证研究。问卷调查和访谈结果初步证实了数字化写作平台介入写作教学的可行性和有效性,研究者针对平台存在的缺陷和不足进一步提出了合理化建议。

以上研究所得出的结论也为我们带来了若干启示：首先，在使用作文自动系统辅助教学的初期阶段，教师应积极引导并帮助学生建立正确的作文错误修改观念。其次，为了未来教学方式的转变，教师和学生应充分利用计算机资源，将其作为教学的有力辅助。再次，教师在使用自动作文评分系统时，应根据学生的语言水平进行有针对性的选择。最后，为了促进学生的自主学习，教师应为学生提供多元化的英语写作学习平台，并鼓励学生积极利用这些平台进行自我学习和提升。

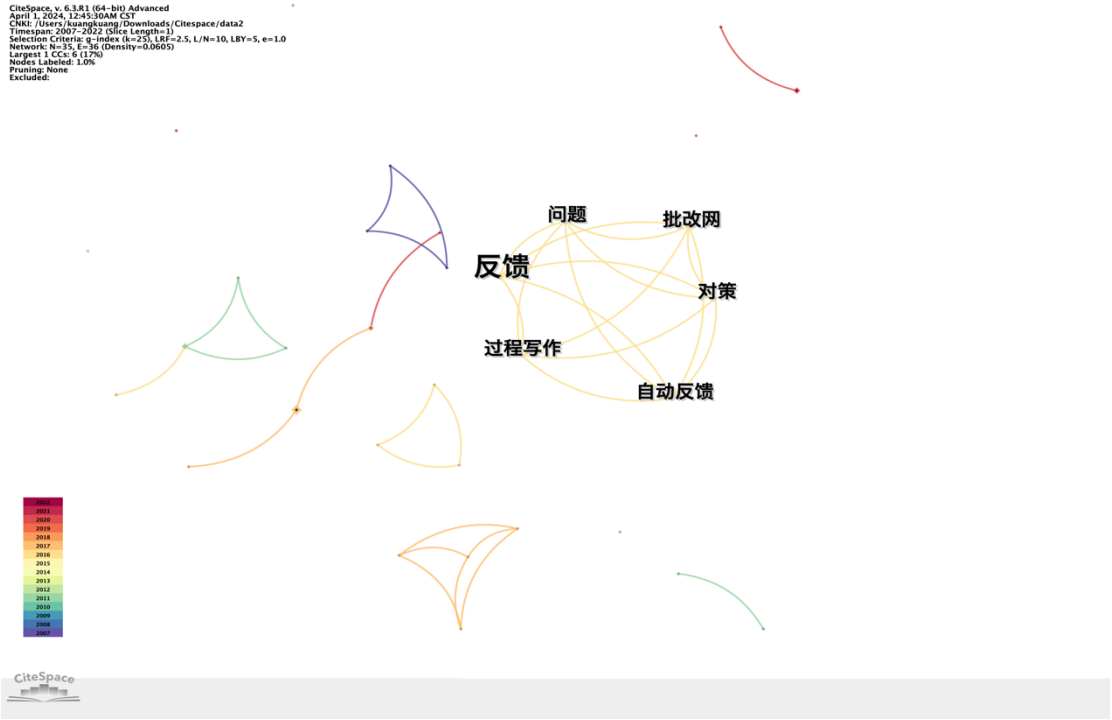


图 4 国内学者关于作文自动评价系统应用之研究的关键词共现知识图谱

4. 国内学者关于作文自动评价系统对学生之影响研究

作文自动评价系统对学生之影响的研究指的是我国学者以学生为主体，探讨作文自动评价系统的应用成效之相关研究。随着作文自动评价系统的广泛应用，国内学者关于该系统对学生之影响的研究也在不断深入。同上，将探讨作文自动评价系统对学生之影响的研究的论文导入到 CiteSpace 软件进行可视化运算后，得到关键词共现知识图谱（如图 5 所示）。进一步分析得到词频排名前 3 的关键词，分别是英语写作、大学生和写作成绩。

具体来看，黄礼珍（2017）对将近 1000 名大学生使用句酷批改网的情况进行调查。调查发现，批改网可以降低写作焦虑感,增强写作信心;批改网批改及时,有助于学生及时了解自己的错误并纠正;大部分学生认真对待批改网的写作。但是批改网也存在局限性，例如无法准确判别作文的内容和连贯性;有时按提示修改,成绩反而变低等。曹婷(2018)以句酷批改网为例，探讨自动作文评分系统反馈对不同水平学生写作句法复杂度的影响。研究结果显示,自动作文评阅系统反馈对二语学习者的写作句法复杂度有积极影响,且不同语言水平的学生在句法复杂度指标上存在显著差异,主要体现在子句平均长度、句子平均长度、T 单位平均长度、句子中的子句数量、T 单位的子句数量、T 单位的复杂 T 单位数量、子句中从句数量、T 单位中从句数量、字句中并列短语数量、T 单位中并列短语数量、T 单位中复合名词短语数量和动词短语数量等方面。宋毅宁（2019）通过实证研究探讨以批改网为代表的本土化自动作文评分系统对大学生写作能力的影响。教学实验结果表明,该系统不仅能从整体上提高大学生的写作水平,还能对除篇章结构之外的三个分项维度产生积极影响,尤其是词汇和句子的分项成绩提高极为显著,而其影响程度与大学生的英语水平呈负相关。

综上所述，尽管作文自动评价系统在线上反馈方面表现出一定的优势 ，但目前在外语写作教学过程中，关于作文自动评价系统线上反馈对于提升教学效果的具体作用，尚未得出明确结论。

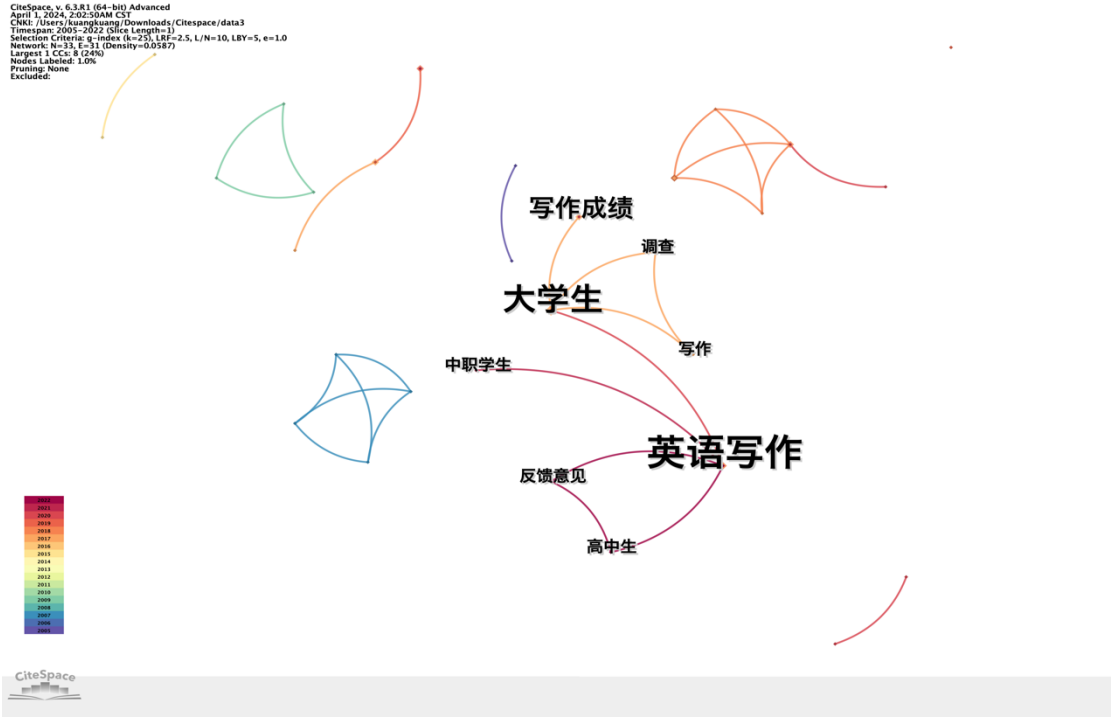


图 5 国内学者关于作文自动评价系统应用之研究的关键词共现知识图谱

5. 结语

本文运用 Citespace 软件对国内学者关于作文自动评价系统研究、作文自动评价系统在现实教学情境中的应用和对学生之影响三个方面进行可视化分析和概述。研究发现，我国学者在较早期的研究阶段，对待作文自动评价系统时更加聚焦于其有效性、可信度和稳定性等方面，对作文自动评价系统的接受度也有所不同。我国学者的研究也正进一步地对作文自动评价系统进行改进和优化，且希冀探索一种更加本地化、更加适用中国独特教育环境的应用场景。在作文自动评价系统的应用方面，鉴于作文自动评价系统技术和设计上仍具有局限性，教师在用其进行辅助教学时，应保持审慎的态度，作文自动评价系统的实施需要经过深思熟虑的教学设计，学会根据不同学生语言水平有选择的使用自动作文评分系统，为学生提供多样化的英语写作学习平台,激发学生自主学习的热情，培养他们的创新能力和批判性思维。与此同时，教师和学生要学会充分利用计算机资源,为未来教学方式的转变奠定坚实基础。但是，目前针对外语教学过程中基于作文自动评价系统的线上反馈是否能够有效促进外语教学尚未形成定论，未来研究应进一步探讨基于作文自动评价系统的写作教学在线反馈的最佳模式，以充分发挥该系统在外语教学中的潜力。

参考文献：

陈小芳.(2017).自动写作评分系统对作文句法复杂度的影响.*黑龙江工业学院学报(综合版)*(11),122-125. https://doi:10.16792/j.cnki.1672-6758.2017.11.024.

高健民.(2021).批改网英语作文自动评分系统评分质量研究.*哈尔滨学院学报*(07),102-105. https://doi:CNKI:SUN:HEBS.0.2021-07-028.

黄礼珍.(2017).对大学生使用作文自动评分系统写作的调查.*江苏外语教学研究*(04),18-21. https://doi:CNKI:SUN:JSWY.0.2017-04-005.

张国强 & 何芳.(2022).英语作文自动评分系统的信度和效度研究——基于不同类型写作任务文本量化特征分析. *外语测试与教学*(01),44-56. https://doi:CNKI:SUN:WYCJ.0.2022-01-006.

张荔.(2017).自动作文评分系统测量性、归纳性和外推性效度研究. *外语与翻译*(03),64-71+98. https://doi:10.19502/j.cnki.2095-9648.2017.03.011.

张锐捷,宋万里,刘琪,范圆圆,姜戴锐 & 葛艺芃.(2019).英语作文自动评分系统研究与实现. *现代信息科技*(04),27-29. https://doi:CNKI:SUN:XDXK.0.2019-04-011.

李俊娴.(2021).基于多特征融合和机器学习理论的英语机考作文自动评分系统. *微型电脑应用*(10),157-160. https://doi:CNKI:SUN:WXDY.0.2021-10-047.

夏林中,罗德安,刘俊,管明祥,张振久 & 龚爱平.(2020).基于注意力机制的双层 LSTM 自动作文评分系统. *深圳大学学报(理工版)*(06),559-566.

王婷 & 王芹.(2016).自动评分系统对高低分作文的反馈——以新疆农业大学外国语学院英语专业学生为例. *现代交际*(09),142-143. https://doi:CNKI:SUN:XKJJ.0.2016-09-106.

杨蔚.(2018).英语作文自动评改系统的可靠性研究——以句酷批改网为例. *湖南科技学院学报*(09),130-132. https://doi:10.16336/j.cnki.cn43-1459/z.2018.09.028.

柯育强.(2017).大学英语作文自动评分系统中文本聚类的应用. *电子技术与软件工程*(05),205.

夏林中,罗德安,刘俊,管明祥,张振久 & 龚爱平.(2020).基于注意力机制的双层 LSTM 自动作文评分系统. *深圳大学学报(理工版)*(06),559-566.

宋毅宁.(2019).基于自动作文评分系统的数字化写作平台的构建及应用效果探究. *英语广场*(10),92-94. https://doi:10.16723/j.cnki.yygc.2019.10.042.

曹婷.(2018).自动作文评分系统反馈对不同水平学生写作句法复杂度的影响——以句酷批改网为例. *考试与评价(大学英语教研版)*(05),94-101. https://doi:10.16830/j.cnki.22-1387/g4.2018.05.021.

Page, E. B. (1968). The use of the computer in analyzing student essays. *International Review of Education*, *14*(2), 210–225. https://doi.org/10.1007/bf01419938

Shermis, M. D., Burstein, J., & Bursky, S. A. (2013). Introduction to automated essay evaluation. *Handbook of Automated Essay Evaluation*. https://doi.org/10.4324/9780203122761.ch1

Supporting Students' Artificial Intelligence Literacy Development by Promoting Authentic Learning Experience Using Programmable Robots

Yunsi Tina Ma [1*], Chi-Fuk Henry So [2]
[1,2] Centre for Learning, Teaching and Technology, [2] Department of Mathematics and Information Technology
The Education University of Hong Kong
[*] yma@eduhk.hk

Abstract: *In the rapidly evolving landscape of artificial intelligence (AI), there is a pressing need for learners to acquire AI literacy. A clear understanding of fundamental AI concepts and principles is critical for leveraging this technology to enhance employability and productivity and to equip individuals with the necessary skills to identify and solve real-world problems in a digital context. Promoting programming with the assistance of tangible robots has proven to be an effective way to introduce abstract concepts to learners. The current study aims to support the AI literacy development of pre-service teachers from various subject disciplines by providing programmable robotic activities in an authentic learning environment. Furthermore, the study aims to develop students' collaboration, communication, and problem-solving skills through workshops and teaching practices. The research participants will be required to attend training workshops before conducting teaching practice in local primary or junior secondary schools. Guided by the authentic learning approach, a training plan and learning and teaching resources have been developed to implement the intervention. A set of instruments will be used to evaluate its effectiveness and the participants' perceptions of acquiring AI literacy through the authentic learning experience.*

Keywords: *artificial intelligence literacy, authentic learning, primary and secondary schools, pre-service teachers, problem-solving skills*

1. Introduction

The rapid advancement of artificial intelligence (AI) has made it increasingly important for learners to acquire AI literacy. A comprehensive understanding of the fundamental concepts and underlying principles of AI is needed for individuals to effectively utilise the technology to enhance their employability, work productivity, and problem-solving abilities within a digital context. Meanwhile, researchers and educators have begun to explore the use of AI to enhance learning, teaching, and assessment. Therefore, teachers must be equipped with AI literacy to be able to apply it in their teaching practice.

The use of tangible robots to promote programming has proven to be an effective strategy for introducing abstract concepts to learners. The present study aims to facilitate the development of AI literacy among pre-service teachers from diverse subject disciplines by providing them with programmable robotic activities to create an authentic learning environment. The study also seeks to foster the enhancement of students' collaborative skills, communication skills, and problem-solving abilities through interactive workshops and teaching practices. The research participants will be required to attend training workshops before engaging in teaching practice at local primary or junior secondary schools.

Guided by an authentic learning approach, a comprehensive training plan will be developed, along with relevant learning and teaching resources, to implement the intervention successfully. A set of evaluation tools will be designed to evaluate the effectiveness of the intervention and gauge the participants' perceptions of the acquisition of AI literacy and a series of generic skills within the authentic learning experience.

2. Research aim and Objectives

The aim of the current research is to develop a set of learning and teaching resources guided by an authentic learning pedagogical approach. Students who complete the workshops and are equipped with the relevant skill sets will be allowed to adapt and adopt the teaching materials and co-teach workshops on related topics for local primary or junior secondary school students with their peers.

The specific objectives of the study are to:

- Develop pre-service teachers' AI literacy and digital competency by providing authentic learning experiences in robotics education through workshops and teaching practice;
- Enhance pre-service teachers' employability and boost their confidence by providing teaching practice and encouraging reflective learning; and
- Cultivate students' collaboration, communication, and problem-solving skills through workshops and teaching practice.

The study addresses three research questions guided by the research aim and objectives.

- RQ 1: How does providing authentic learning experiences in robotics education through workshops and teaching practices contribute to the development of pre-service teachers' AI literacy and digital competency?
- RQ 2: How does the inclusion of teaching practices and reflective learning opportunities enhance the employability and confidence of pre-service teachers?
- RQ 3: To what extent do workshops and teaching practice focused on robotics education cultivate students' collaboration, communication, and problem-solving skills?

3. Literature Review

3.1 Authentic Learning

Authentic learning takes place when learners are situated in a social context in which they can gain relevant real-life experience in specific fields and subject areas (Lombardi & Oblinger, 2007; Stein et al., 2004). Emerging technologies are opening up new possibilities for creating authentic learning environments for students eager to gain first-hand experience before participating in real-world workplace practice (Herrington et al., 2014). Research has revealed that authentic learning can be achieved through problem-based activities (Lombardi & Oblinger, 2007). Studies have also found that working in groups in an authentic learning environment helps students to improve their communication and collaboration skills (Ornellas et al., 2019; Singer et al., 2020). Providing an authentic learning context for students enables them to apply and reflect on their knowledge application in the workplace (Hwang et al., 2022). Many efforts have been made to promote authentic teacher education (Huang et al., 2023; Radović et al., 2021; Theelen et al., 2022).

3.2 Artificial Intelligence Literacy Education

As AI has penetrated all of our daily lives, it is important for people of all educational levels to acquire AI literacy. Therefore, AI literacy education is of great value and importance. While children should learn about the potential of AI at an early age, adults in higher education and beyond should have at least a basic understanding of AI (i.e., AI literacy) to interact effectively with the technology. Based on a scoping literature review, Laupichler et al. (2022) argued that the definition of AI literacy education for adults needs to be further refined. Kong et al. (2022) developed a course to introduce university students to basic AI concepts that effectively lower the barrier to entry for AI literacy for students from a non-computer science background. A framework comprising cognitive, metacognitive, affective, and social dimensions was subsequently proposed as a holistic framework for AI literacy (Kong et al., 2024).

3.3 Artificial Intelligence Education through Programmable Robots

Programmable robots have been widely used in K-12 computational thinking and STEM education (Kong & Wang, 2019; Yang et al., 2022). Studies have indicated that AI robots have the potential to design educational programmes for the K-12 school sector or professional teacher training (Chu et al., 2022; Yang, 2022). The current study will apply programmable robots, such as the AlphAI robot developed by Learning Robots (Figure 1), to conduct workshops and teach fundamental AI concepts to the participating students and guide them to enhance the learning and teaching resources for conducting teaching practice in local primary and secondary schools.

The AlphAI robot is a programmable robotic vehicle that can be trained to navigate obstacles. Learning to program the robot can help students to understand AI concepts, such as supervised and unsupervised learning and deep learning (Martin et al., 2023). The AlphAI robot is equipped with multiple sensors, including a wide-angle camera, ultrasound, and an infrared sensor, allowing students to train a model using different data types. A graphical user interface (GUI) accompanying the robot depicts the neural network that controls the robot's training. Students can visualise the neural network as the robot is trained with supervised or reinforcement learning. This essentially "opens the black box" of deep learning, and helps students understand the reciprocal relationship between the robot's performance during and after training and the components of the network architecture, such as the nodes, weights, and inputs. Martin et al. (2023) found that the AlphAI robot helped primary students to understand the basics of AI and developed their persistence and mindset in learning through trial-and-error, thus empowering them to use AI and learning in general.

Learning Robots has also partnered with the association that developed the Thymio robot to integrate the AlphAI GUI with Thymio (Figure 2). Unlike the AlphAI robot, the Thymio robot only has infrared sensors (Riedo et al., 2012; Mondada et al., 2017). The GUI allows students to control the Thymio robot through a neural network, and again visualise how the architecture of the neural network changes as they train the robot with supervised or reinforcement learning.

In collaboration with industrial partners, the current project will design and develop learning and teaching resources and organise workshops to instruct students on how to use the programmable robots to support AI literacy learning. Students who complete the workshops and are equipped with related skill sets will be allowed to adapt and adopt the teaching materials and co-teach workshops on related topics for local primary or junior secondary school students with their peers of the students who complete the workshops.

Figure 1. The AlphAI Robots

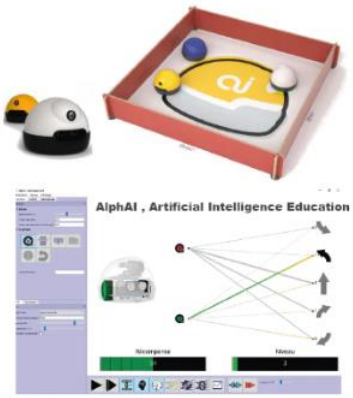
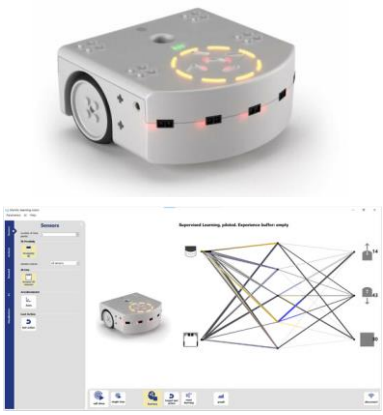


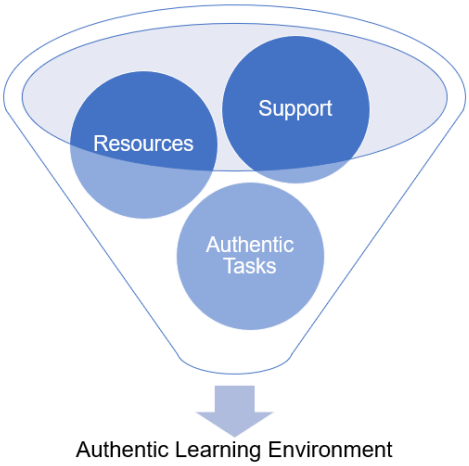
Figure 2. The AlphAI GUI with Thymio



4. The Authentic Learning Environment

In the current study, the authentic learning activities comprise training workshops and teaching practices. The study aims to support AI literacy development among pre-service teachers from various subject disciplines by providing an authentic learning environment that incorporates programmable robotic activities. Furthermore, the study targets the development of students' collaboration skills, communication skills, and problem-solving skills through the workshops and teaching practices. The research participants will be required to attend training workshops before conducting teaching practice in local primary or junior secondary schools. Guided by the authentic learning approach, the current study proposes an authentic learning environment with resources, support, and authentic tasks as input (Figure 3). A training plan with learning and teaching resources will be developed to implement the intervention, and a set of instruments will be designed to evaluate the effectiveness of the intervention and the participants' perceptions of acquiring AI literacy through the authentic learning experience.

Figure 3. The Authentic Learning Environment

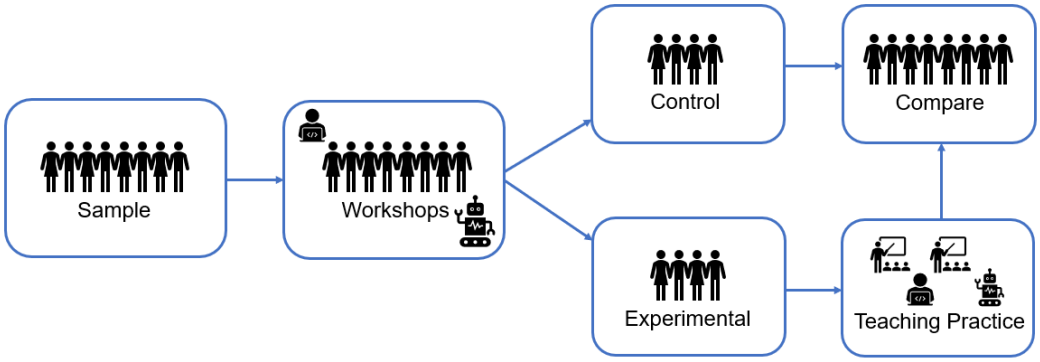


5. Methods

5.1 Research Design

The study will involve approximately 320 university students over two years. All of the participants must attend two 2-hour workshops. In the first 2-hour session, students will be instructed to utilise the programmable robots to learn fundamental AI concepts. In the second 2-hour session, the emphasis will be on formulating and enhancing the learning and teaching resources to conduct teaching practice in primary and junior secondary schools. After the workshops, the students will be divided into a control group and an experimental group. The students in the experimental group will conduct teaching practice in the local primary or junior secondary schools. The research design is presented in Figure 4.

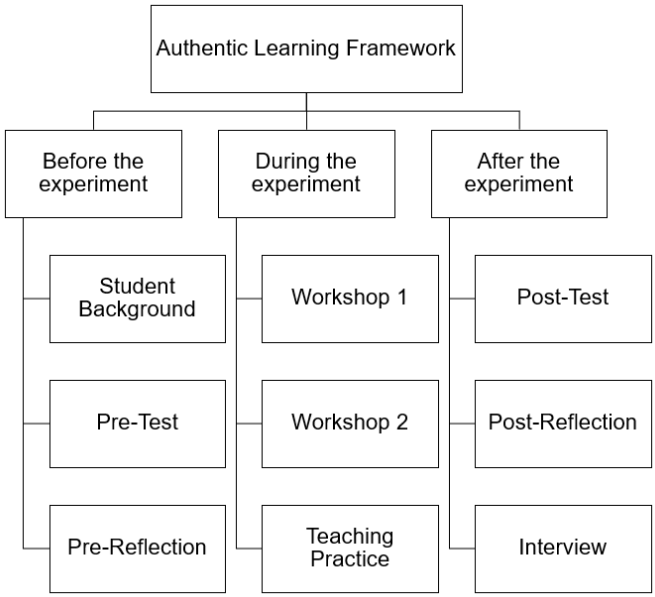
Figure 4. The Research Design



5.2. Data Collection and Data Analysis

To address the three RQs, qualitative and quantitative data will be collected through (1) pre- and post-workshop surveys, (2) reflective writing, and (3) focus group interviews. Pre- and post-workshop surveys will be conducted in the same venue where the workshops will be held. The reflective writing will be guided by three open-ended questions. Focus group interviews will be conducted to collect in-depth feedback from the students. Figure 5 sets out the experimental and data collection procedure.

Figure 5. Experiment and data collection procedure.



5.3. Development of the Survey Instruments

A set of survey instruments (Table 1) has been developed by adapting various instruments developed by previous studies (Herrington & Oliver, 2000; Keinänen et al., 2018; Pu et al., 2016; Tezer et al., 2018). Each survey will take approximately 15 minutes to complete. Responses to the survey items will be given on a 5-point Likert scale. After data collection, a Cronbach's alpha reliability test will be conducted to measure the reliability of the authentic learning scale using the pre-test data. An exploratory factor analysis will be conducted to identify the key factors.

Table 1. The Authentic Learning Scale

Items (5-point Likert scale)
Problem-Solving Skills and Innovation
1. I can produce alternative solutions to real problems
2. I can present real-life problems
3. I can improve my problem-solving skills
4. I can be motivated
5. I can relate the knowledge I acquired to real life
6. I can find opportunities to express myself to others
7. I can apply my knowledge to different subject fields
8. I can relate new information to my old knowledge
9. I can take different roles in learning environments
10. I can easily and openly attend to learning activities
11. I suggest ideas for others to explore how the job should be done
12. I suggest new ideas for solving problems
13. I suggest new practical solutions to reach a goal
14. With my competence, I help the group achieve its goals
15. I offer new ideas to make them available to others
16. I can steer the group towards the goal
Collaboration
17. I can construct information by cooperating with other learners
18. I can create social bonds with people around me (friend, teacher, specialist, etc.)
19. I can define myself as a member of the group I am in
20. I can work in harmony with my fellow group members
21. I can acquire advantages like brainstorming and discussing ideas in collaborative environments
22. I can keep up with the new technological advancements by using social networks (YouTube, Wikipedia, Facebook, etc.)
23. I can keep up with the new technological advancements by using online communication tools (WhatsApp, Zoom, Skype, Gmail, etc.)
24. I can solve the problems more easily when I work with other learners
25. I can reflect on what I have learned both individually and collaboratively with activities
Authentic Contexts and Activities
26. I can learn better by interacting in real life
27. I can understand how the information is used in real life
28. I can apply the skills and knowledge I have acquired in real life
29. I can face complex tasks
30. I can come up with solutions to complex tasks
31. I can acquire rich learning experiences
32. Through the activities, I am able to analyze and critically consider what I have learned, which helps me understand and organise my knowledge

33. Through the activities, outstanding students receive recognition for their performance, which helps motivate me to perform better
34. I feel that through each activity, I am able to ascertain areas in which I need further personal development and learning.
35. Each activity allows everyone to participate together, which allows students to provide each other with recommendations.

Table 2 shows the alignment among the RQs that will be formulated, the issues that will be addressed, the nature of the issues, the instruments used to collect data, and the corresponding data analysis approach.

Table 2. Issues to be Explored, Instrument to be Deployed, and Statistical Methods to be Adopted

RQs	Issues	Instrument/ Data Sources	Analysis
RQ 1: How does providing authentic learning experiences in robotics education through workshops and teaching practices contribute to the development of pre-service teachers' AI literacy and digital competency?	The effects of authentic learning on AI literacy and digital competency development	Pre- and post-workshop surveys, reflective writing, focus group interview	Descriptive statistics, analysis of covariance (ANCOVA)
RQ 2: How does the inclusion of teaching practices and reflective learning opportunities enhance the employability and confidence of pre-service teachers?	The effects of authentic learning on pre-service teachers' employability and confidence	Reflection, focus group interview	Descriptive statistics, thematic analysis
RQ 3: To what extent do workshops and teaching practice focused on robotics education cultivate students' collaboration, communication, and problem-solving skills?	The effects of authentic learning on generic skills	Pre- and post-workshop surveys, reflection, focus group interview	Descriptive statistics, ANCOVA, thematic analysis

Acknowledgements

The project was supported by the Fund for Innovative Technology-in-Education provided by the University Grants Committee, Hong Kong. Ethical approval was granted by the Human Research Ethics Committee of The Education University of Hong Kong with the reference number [2023-2024-0139].

References

Aynas, N., & Aslan, M. (2021). The effects of authentic learning practices on academic success in science courses. *Educational Research Quarterly*, *44*(4), 3–30.

Chu, S. T., Hwang, G. J., & Tu, Y. F. (2022). Artificial intelligence-based robots in education: A systematic review of selected SSCI publications. *Computers and Education: Artificial Intelligence*, *3*, 100091.

Herrington, J., Reeves, T. C., & Oliver, R. (2014). Authentic learning environments. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 401–412). Springer. https://doi.org/10.1007/978-1-4614-3185-5_32

Huang, Y., Richter, E., Kleickmann, T., & Richter, D. (2023). Comparing video and virtual reality as tools for fostering interest and self-efficacy in classroom management: Results of a pre-registered experiment. *British Journal of Educational Technology*, *54*(2), 467–488.

Hwang, W. Y., Luthfi, M. I., Hariyanti, U., & Wardani, R. (2022). Evaluation of fraction learning in authentic context using Ubiquitous Fraction App. *Education and Information Technologies*, 1–25. https://doi.org/10.1007/s10639-022-11453-2

Keinänen, M., Ursin, J., & Nissinen, K. (2018). How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments. *Studies in Educational Evaluation*, *58*, 30–36. https://doi.org/10.1016/j.stueduc.2018.05.007

Kong, S. C., Cheung, W. M. Y., & Zhang, G. (2022). Evaluating artificial intelligence literacy courses for fostering conceptual learning, literacy and empowerment in university students: Refocusing to conceptual building. *Computers in Human Behavior Reports*, *7*, 100223. https://doi.org/10.1016/j.chbr.2022.100223

Kong, S. C., Cheung, M. Y. W., & Tsang, O. (2024). Developing an artificial intelligence literacy framework: Evaluation of a literacy course for senior secondary students using a project-based learning approach. *Computers and Education: Artificial Intelligence*, 100214. https://doi.org/10.1016/j.caeai.2024.100214

Kong, S. C., & Wang, Y. Q. (2019). Nurture interest-driven creators in programmable robotics education: An empirical investigation in primary school settings. *Research and Practice in Technology Enhanced Learning*, *14*, 20. https://doi.org/10.1186/s41039-019-0116-1

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, *3*, 100101.

Lombardi, M. M., & Oblinger, D. G. (2007). Authentic learning for the 21st century: An Overview. *Educause Learning Initiative*, *1*, 1–12.

Martin, M., Chevalier, M., Burton, S., Bonvin, G., Besançon, M., & Deneux, T. (2023). Effects of introducing a learning robot on the metacognitive knowledge of students aged 8-11. In Balogh, R., Obdržálek, D., Christoforou, E. (Eds), *Robotics in Education. RiE 2023. Lecture Notes in Networks and Systems, Vol. 747.* Springer. https://doi.org/10.1007/978-3-031-38454-7_15

Mondada, F., Bonani, M., Riedo, F., Briod, M., Pereyre, L., Retornaz, P., & Magnenat, S. Bringing robotics to formal education: The Thymio open-source hardware robot. *IEEE Robotics and Automation Magazine*, *24*(1), 77–85. https://doi.org/10.1109/MRA.2016.2636372

Ornellas, A., Falkner, K., & Edman Stålbrandt, E. (2019). Enhancing graduates' employability skills through authentic learning approaches. *Higher Education, Skills and Work-Based Learning*, *9*(1), 107–120. https://doi.org/10.1108/heswbl-04-2018-0049

Pu, Y. H., Wu, T. T., Chiu, P. S., & Huang, Y. M. (2016). The design and implementation of authentic learning with mobile technology in vocational nursing practice course. *British Journal of Educational Technology*, *47*(3), 494–509. https://doi.org/10.1111/bjet.12443

Radović, S., Hummel, H. G., & Vermeulen, M. (2021). The challenge of designing 'more' experiential learning in higher education programs in the field of teacher education: A systematic review study. *International Journal of Lifelong Education*, *40*(5–6), 545–560. https://doi.org/10.1080/02601370.2021.1994664

Riedo, F., Rétornaz, P., Bergeron, L., Nyffeler, N., & Mondada, F. (2012). A two year informal learning experience using the thymio robot. In *Advances in Autonomous Mini Robots: Proceedings of the 6-th AMiRE Symposium* (pp. 37–48). Springer.

Singer, A., Montgomery, G., & Schmoll, S. (2020). How to foster the formation of STEM identity: Studying diversity in an authentic learning environment. *International Journal of STEM Education*, *7*(1), 1–12. https://doi.org/10.1186/s40594-020-00254-z

Stein, S. J., Isaacs, G., & Andrews, T. (2004). Incorporating authentic learning experiences within a university course. *Studies in Higher Education*, *29*(2), 239–258.

Tezer, M., Yildiz, E. P., & Uzunboylu, H. (2018). Online authentic learning self-efficacy: S scale development. *Quality & Quantity*, *52*(Suppl 1), 639–649.

Theelen, H., van den Beemt, A., & Brok, P. D. (2022). Enhancing authentic learning experiences in teacher education through 360-degree videos and theoretical lectures: Reducing preservice teachers' anxiety. *European Journal of Teacher Education*, *45*(2), 230–249.

Yang, W. (2022). Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, *3*, 100061.

Yang, W., Ng, D. T. K., & Gao, H. (2022). Robot programming versus block play in early childhood education: Effects on computational thinking, sequencing ability, and self-regulation. *British Journal of Educational Technology*, *53*(6), 1817–1841. https://doi.org/10.1111/bjet.13215

OpenHydra：一个本地化部署的开源人工智能教学平台

OpenHydra: A locally deployed open source AI teaching platform

谢作如 [1*]
[1] 温州科技高级中学，上海人工智能实验室
[*]8334180@qq.com

【摘要】 在中小学普及人工智能教育不仅需要合适的工具和课程，还需要本地的人工智能教学平台。各种人工智能实验的开展，都需要借助特定的软硬件环境支持。OpenHydra 则是在这一背景下推出的国产开源项目，用于搭建一个本地的人工智能教学平台。本文分析了中小学人工智能教学平台的核心功能，并介绍在一台算力服务器上快速搭建 OpenHydra 的过程，最后对 OpenHydra 项目提出了新的规划。

【关键词】 人工智能教育；OpenHydra；算力分割；算力虚拟化

Abstract: *Universal AI education in primary and secondary schools requires not only appropriate tools and courses, but also local AI teaching platforms. The development of various artificial intelligence experiments needs to be supported by specific hardware and software environments. OpenHydra It is a domestic open source project launched in this context to build a local AI teaching platform. This paper analyzes the core functions of the AI teaching platform for primary and secondary schools, introduces the process of quickly building the OpenHydra on a computing power server, and finally puts forward a new plan for the OpenHydra project.*

Keywords: Artificial intelligence education, OpenHydra; computing power segmentation, computing power virtualization

1. 前言

当数据、算法和算力被公认为人工智能时代的三驾马车，就意味着生物启发式方法（搭建深度神经网络并在数据中获取智能），已经成为人工智能最重要的研究范式。中小学人工智能教育需要密切围绕“模型”开展，不能仅仅停留在体验和应用模型层面，还要尝试收集数据并训练、部署模型，在亲历“智能从无到有”的流程中理解模型。令人惊喜的是，随着以 Keras、XEdu 为代表的人工智能学习和开发工具的出现，学生只要借助少量代码甚至不写代码都能训练出人工智能模型。如果拥有合适的数据，完全可以做到解决各种真实问题。

2. 中小学人工智能教学平台的核心功能分析

大部分人工智能学习和从业者，往往是借助商业云算力来开展各种实验或者工作。与高校不同的是，中小学人工智能教育往往在课堂上完成，对教学时间的把控要求比较高。而商业云算力并没有提供适合中小学的容器（学习环境），用户需要从零开始配置环境。最麻烦的事，做深度学习实验时还需要在服务器和本地之间传输大量的训练数据。因此商业云算力的方案对中小学来说并不适合，有条件的学校需要本地搭建内置算力的教学平台。而这一平台除了提供教学资源和学生管理外，最核心的功能是算力分割和学习环境管理。

2.1. 算力分割管理

虽然借助 XEdu 一键安装包，在任何一台 win7 以上电脑上都可以开展人工智能教学，但是训练模型仅限于 CPU 环境，做对算力要求不高的机器学习和全连接神经网络的实验。至于基于卷积神经的图像分类和目标识别实验，没有支持 CUDA（NVIDIA 推出的通用并行计算

架构）的 GPU 设备，一节课很难完成教学任务。从性价比和维护难度看，买一台算力服务器比为每一台学生电脑配置 GPU 的做法要更加合适。

需要强调的是，人工智能教育并非是针对少数人的高端课程，大部分学校的人工智能课是面向一个班级开展。因此这个用于教学的算力服务器需要支持 30-40 人左右的并发量。一般来说，一块 24G 的显卡一般可以切分为 4 份来使用。这样一计算，支持一个班级至少需要 2-4 块显卡。即便如此，40 人的班级还要分组，让 2-3 人使用 1 个 GPU 容器，毕竟 GPU 设备太贵了。

2.2. 学习环境管理

这里的“学习环境”，指的是人工智能的开发和训练环境。众所周知，搭建环境能劝退 90%以上想学习人工智能的人。因此用 Docker 容器预先安装好各种相关开发框架、工具和 Python 库，成为一种通用的选择，绝大多数的商业云算力平台都选择了内置 Docker 容器技术。所谓 Docker 容器类似一台无桌面系统的虚拟机，借助交互 API，即可在浏览器上编写代码，在远程容器上运行，再返回到浏览器上呈现。这个过程对学生用户来说是透明的，跟本地操作没有区别。

目前最适合中小学生学习人工智能的环境首选 XEdu。XEdu 实际上是由 MMEdu、BaseML、BaseNN 和 XEduhub 等一系列 Python 库组成，涵盖了 Pytorch 和 sklearn 等必备的人工智能学习框架。除此之外，Keras 和 PaddlePaddle 也值得作为学习环境，完成一些特定的实验。为了避免 Python 库的版本冲突，这一平台最好能支持多个容器。

3. 用 OpenHydra 搭建人工智能教学平台的过程

OpenHydra 是由全国师范生人工智能教育共同体、亓绚科技联合发起的一个开源项目，旨在为中小学 AI 教育提供一个开箱即用的学习与实践平台。OpenHydra 基于 K8S（全称 Kubernetes，一个用于管理容器的开源平台）开发的系统，默认内置了最新版本的 XEdu。

OpenHydra 的开源地址：https://github.com/openhydra/

3.1. 服务器软件准备

对于熟悉算力管理的高手来说，可以在 K8s 系统上直接安装 OpenHydra 的插件。对中小学来说推荐使用编译好的 OpenHydra 镜像文件来安装。只要将编译好的 ISO 文件烧录到 U 盘上，然后用这个 U 盘系统引导电脑启动，即可一步一步完成 OpenHydra 系统的安装。

3.2. 启动服务器并设置 IP 地址

全新安装 OpenHydra 实际上包含了安装 Ubuntu（版本为 server 20.04.06）和 OpenHydra 两个环境。安装过程类似对一台新电脑安装系统，关键操作是设置 IP 地址、指定磁盘和用户名。从系统引导开始计时，安装时间预计 20-30 分钟。需要强调的是，磁盘的数据要及时备份，安装过程中会初始化磁盘。

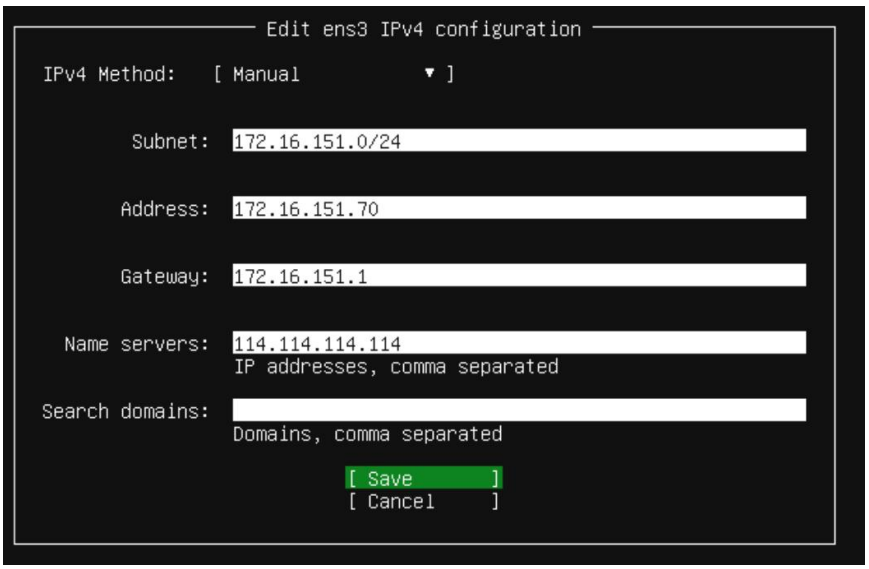


图 1 安装过程的信息设置

3.3. 测试教学平台

系统安装完成后，重启电脑后在显示器上看到如图所示的信息，说明一切准备就绪，不需要再做任何操作，就可以开始上课了。

```
$ journalctl -u maas -f

# 输出
# 如果安装完成那么我们应该可以看到 maas.service: Succeeded 的输出
-- Logs begin at Sun 2023-12-10 10:58:29 UTC. --
...
Dec 10 11:29:14 test systemd[1]: maas.service: Succeeded.
Dec 10 11:29:14 test systemd[1]: Finished maas service.
```

图 2 系统完成后的信息提示

用其他电脑打开浏览器，输入服务器地址和 30001 端口，如“http://192.168.3.233:30001”，就能看到 OpnHydra 的登录页面，如图 3 所示。

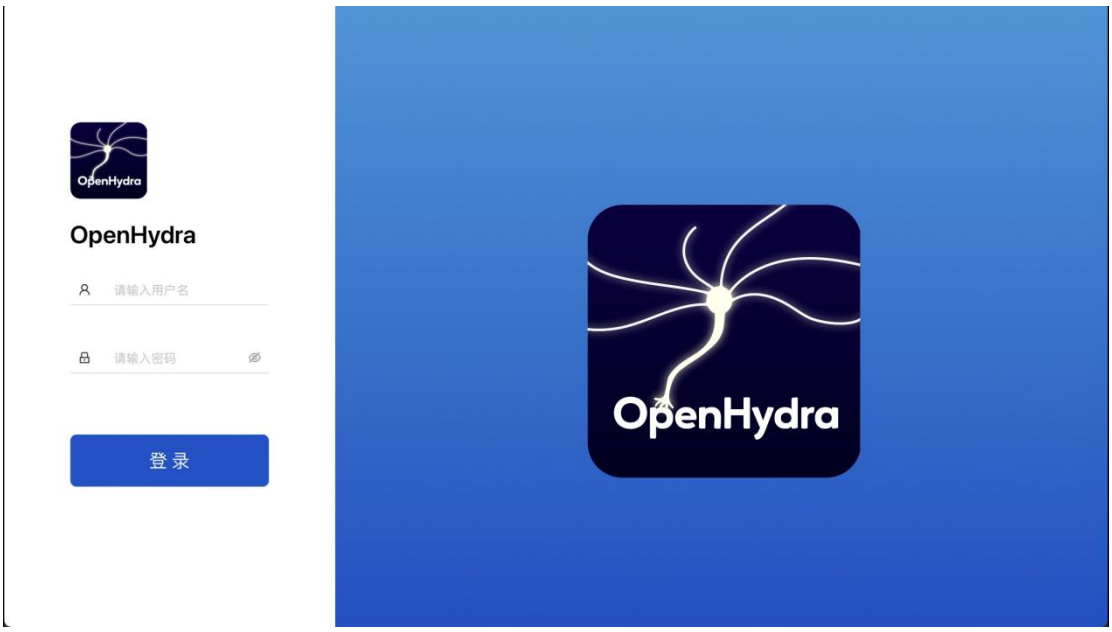


图 3 OpnHydra 的登录页面

教师在上课之前，首先需要配置学生账号。“设备管理”页面将列出所有的用户登录和使用情况，如图 4 所示。

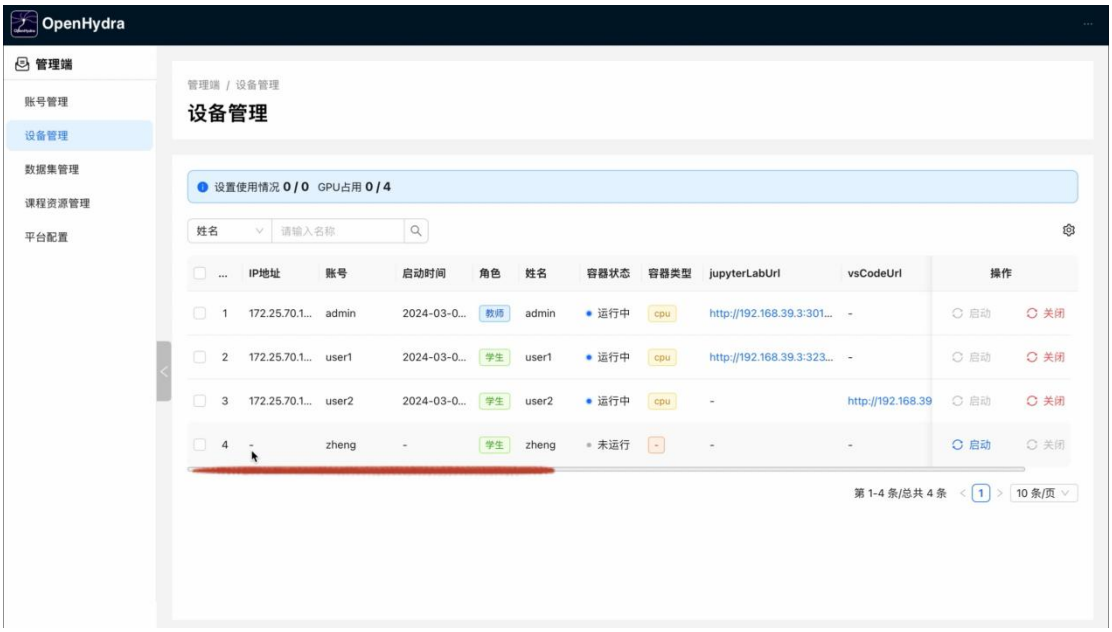


图 4 用户登录和使用情况

对学生而言，在 OpenHydra 上能看到所有支持的容器，并选择熟悉的方式打开实验环境。如选择 jupyter 形式，则自动启动 jupyter 页面，具体过程跟使用 XEdu 一键安装包 OpenInnolab（浦育平台）几乎是一致的。

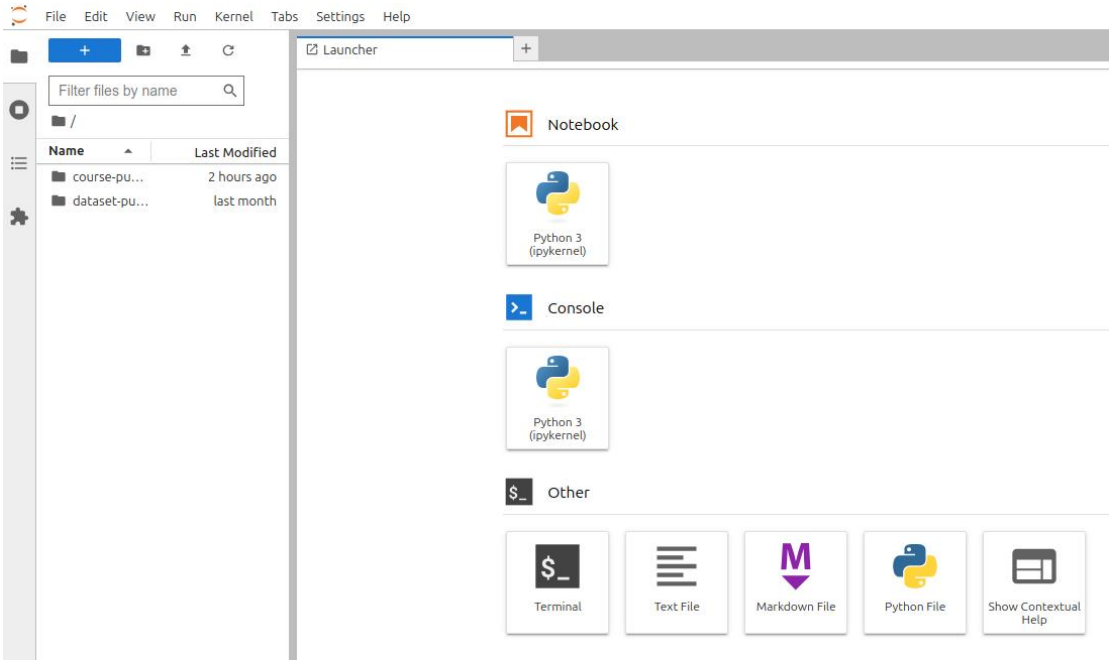


图 5 选择熟悉的方式打开实验环境

4. OpenHydra 的后续功能规划

据大量用户的测试反馈，OpenHydra 的安装和配置都足够简单，界面也非常友好，很好地解决了算力分割的问题。但因为项目刚刚发布，还需要进一步完善。目前 OpenHydra 社区汇聚了大量开源的 AI 开发工具和 AI 教学实践项目，尤其期待多方力量的加入后，这个系统能实现更多的功能，比如：

4.1. 支持更多的容器环境

如图所示，目前 OpenHydra 的内置容器只有 XEdu 标准版，还需要加入 Keras、PaddlePaddle 的容器。即使 XEdu 版本也应该有 IOAI 版本（AIO 奥林匹克专用）、清华大学教材版本、浙江教育出版社教材版本等。甚至一些用户数量较大的课程也可以封装为一个独立的版本。这些容器可以由管理员或者教师来手动更新，自由选择。

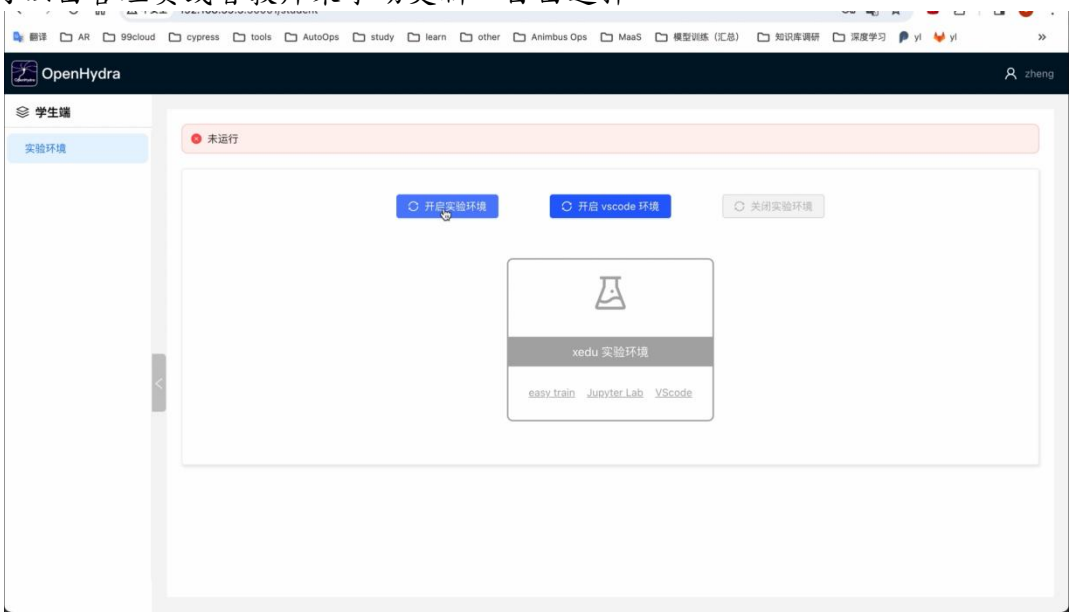


图 6 OpenHydra 的内置容器

4.2. 支持在线更新课程资源

目前 OpenHydra 团队提供了课程模板，按照一定规范即可为这个项目提供标准课程。课程压缩为 zip 格式，即可导入到 OpenHydra 平台。但是人工智能的课程往往要会内置特定的数据集，而深度学习的数据集都很多，少的数百张图片，多的要数万张。那这些课程资源存放在哪里？如何更新？如果不设计在线增量更新的方式，肯定会限制后续的发展。

课程分享专用地址：https://github.com/openhydra/open-hydra-course

4.3. 支持更多算力设备

众所周知，GPU 设备是最昂贵的。目前一些国产的通用 GPU 设备也兼容 CUDA，如沐曦、天数智芯等。因为中小学的人工智能实验不会追求最新的技术，往往会选择一个最常用的计算框架版本，不需要实时更新。只要做好一次适配，就能用较长一段时间。这个适配工作意义很大，但需要算力企业的支持。

4.4. 支持各种教学插件开发

中小学一旦普及人工智能教育，各种需求就自然而然发生。如用户通用认证、数据标注、作业评价、模型在线部署服务和特定的教学活动支持等，这些功能应该以插件的形式吸引有能力的企业创客和教师开发。比如上海人工智能实验室和中科协发起的“与 AI 共创未来”活动中有个“算法挑战赛”，以“天梯排行”的形式定时更新代码的推理成绩，深受学生喜欢。这一活动就很适合做成一个插件，让学生挑战起来。

4.5. 支持各类输入和输出设备

人工智能研究类人的智能，即使再聪明，也始终是抽象的。为吸引中小学生，人工智能往往也结合各种输入和输出的设备，使其具象化。输入设备如麦克风、摄像头和各种传感器，输出设备如各种开源硬件。OpenHydra 也要考虑让容器中的代码跟本地代码一样支持麦克风和摄像头，还要能够驱动开源硬件，让中小学的人工智能教育更加丰富多彩。

5. 总结

全国师范生人工智能教育共同体主任委员钟柏昌教授提出，要开展中大型深度学习模型的训练和部署，除了利用云服务器外，未来开展大规模常态化人工智能课程教学，离不开学校本地服务器的算力资源，而这种算力资源要同时满足几十上百个模型的同步训练，又离不开算力资源的合理分割。OpenHydra 项目的出现非常及时，不仅可以用于中小学，也适用于一些规模较小的高校实验室使用。人工智能的发展离不开教育，而教育需要更好的工具，我们对 OpenHydra 项目的发展充满了期待。

参考文献

谢作如.借助商业服务破解 AI 科创活动的算力难题[J].**中国信息技术教育**,2023,(07):91-93.

钟柏昌 余峻展 谢作如.中小学人工智能课程需要何种智能硬件?——现状分析与发展方向[J].**远程教育杂志**,2024(01):74-83.

基于 OpenInnolab 平台的初中人工智能校本课程设计及其教学实践探索

School-based curriculum design of junior middle school based on artificial intelligence based on OpenInnolab platform

陈婧[1*]，孟晓培[1]，应利波[2]
[1]温州大学 教师教育学院
[2]温州大学电气与电子工程学院
[*]2594552593@qq.com

【摘要】人工智能作为新一轮技术革命的引擎，各国纷纷将其纳入各级教育体系，以推动多领域创新发展。中国高中阶段已布局人工智能教育，但在初中层面，课程体系尚显薄弱，缺乏统一标准与教材，导致实践性教学不足。为弥补这一短板，亟需构建实践导向的人工智能课程，使学生通过动手操作深入理解并运用 AI 技术。本研究借助 OpenInnoLab 平台，开发了一套融合前沿知识与实践应用的初中人工智能校本课程，并已成功应用于教学实践中。本研究的校本课程设计范例，对弥补当前初中阶段人工智能教育实践性不足的问题做出了实质贡献。

【关键词】OpenInnolab 平台；初中人工智能课程；校本课程；教学实践

Abstract: *As the engine of a new round of technological revolution, artificial intelligence has been incorporated by various countries into education systems at all levels to promote innovation and development in various fields. AI education has been arranged in high schools in China, but at the middle school level, the curriculum system is still weak and lack of unified standards and teaching materials, which leads to the lack of practical teaching. In order to make up for this shortcoming, it is urgent to build practice-oriented AI courses, so that students can deeply understand and use AI technology through hands-on operations. With the help of OpenInnoLab platform, this research has developed a set of junior middle school AI curriculum integrating cutting-edge knowledge and practical application, which has been successfully applied in teaching practice. The example of school-based curriculum design in this study makes substantial contributions to making up for the lack of practicality of AI education in junior middle school.*

Keywords: OpenInnoLab Platform**,** Junior Middle School Artificial Intelligence Curriculum，School-Based Curriculum，Teaching Practice

1. 前言

人工智能（AI）已经成为引领新一轮技术革命和产业创新的核心驱动力，对于各国在教育、医疗、交通、农业和工业等多个领域的发展起到了推动作用。为了加速人工智能技术在不同领域的探索和应用，各国纷纷制定了人工智能发展规划，将其纳入高等教育甚至中小学教育的教学课程中。中国教育部早在 2017 年新版高中信息技术课程标准中就已布局人工智能教育，将其融入必修与选择性必修课程，覆盖数百万计的学生。

然而，目前中国初级中学的人工智能相关课程仍相对匮乏，缺乏统一、完善的人工智能课程标准和教材，使得人工智能教育仍处于从零到一的探索阶段。在人工智能热潮的影响下，一些 STEAM、创客、开源硬件、机器人和编程等内容被纳入人工智能课程中，然而，这些课

程往往仅限于书本上的理论知识，缺乏实际应用与实践的机会。学生只是停留在对知识的被动接收，而不能真正体验到人工智能的应用过程。因此，人工智能教育需要更多强调实践性，使学生能够通过实际操作来深入理解和运用人工智能技术，而不仅仅是停留在理论层面。在建设人工智能课程的过程中，更加注重实践和体验，将理论知识与实际运用相结合，将为学生提供更加全面和深入的学习体验，促使其更好地适应未来社会对人工智能应用的需求。

浦育青少年人工智能开放创新平台 OpenInnoLab 致力于为从事 AI 教育的教师和对 AI 科创感兴趣的学生提供前沿、多元、易用、连通的学习创作工具。本研究将 OpenInnoLab 融入初中人工智能课程体系，精选理论与实践紧密结合的教学内容，进而设计出一套基于该平台的初中人工智能校本课程，并已付诸实际教学应用。

2. OpenInnolab 平台介绍

OpenInnoLab 青少年人工智能开放创新平台致力于服务 AI 教育者和学生，提供前沿、多元化、易用且具备互联功能的学习创作工具。平台结合丰富课程与实践案例，全面涵盖用户管理、课程研发、自主学习、作业管理、在线实验等教学环节，并集成大数据分析能力，助力青少年 AI 教育普及与实践。此外，平台搭建了完整的科创竞赛系统和智能教育门户，衔接各类 AI 创新工具与教育资源，推广 AI 教材与科研成果。尤其值得注意的是，平台提供的 AI 教育创新实践工具集，涵括多种云端开发工具，支持青少年从数据标注至模型部署全过程的 AI 实践，有效降低了 AI 学习门槛。

3. 基于 OpenInnoLab 平台的初中人工智能校本课程设计

依托 OpenInnolab 平台的特性，本研究针对初中阶段设计了一套基于 OpenInnolab 的人工智能校本课程，并全面探讨了课程建设中的目标定位、内容架构、教学模式及评价机制设计。

3.1. 课程目标定位

依据《义务教育信息科技课程标准（2022 年版）》并结合 OpenInnoLab 平台的特性，本研究为校本人工智能课程确立了如下总目标与分目标。总体而言，课程旨在培养学生具备应对智能社会所需的各项品格与关键能力。具体分目标如下：在信息意识层面，学生需理解人工智能对信息社会发展的贡献，养成自主解决问题并掌握核心技术的意识。在计算思维层面，引导学生在实践探究过程中，运用计算思维解析并内化人工智能的基本原理与实现方法。在数字化学习与创新能力培养方面，期待学生能充分利用 OpenInnoLab 平台展开实践创新，熟练运用数字设备促进学习，精通自主与合作学习策略，高效检索所需资源；在信息社会责任层面，要求学生在网络与物联网环境中实施安全防护，尊重隐私，理解自主可控对国家安全与未来发展的重要性。

3.2. 课程内容架构

本研究遵循信息科技课程标准，以人工智能技术的原理与应用为核心，依托 OpenInnolab 平台，精心设计了一套理论与实践紧密结合的课程体系。整个课程共包含 16 个课时，分为三个主要篇章，包括理论篇、案例篇、实践篇，以确保学生全面了解人工智能的核心技术，同时将其融入日常生活场景，使学习更加具体和贴近实际，课程内容大纲见表 1。

理论篇着重阐述人工智能伦理要义，强调伦理原则在技术进步中的关键地位，培养对社会影响的敏感认知。在理论篇中，课程还特别增设了工具篇，旨在引导学生熟练掌握 OpenInnoLab 平台的各项功能与操作，以便学生在后续学习与实践中能有效地运用这一平台进行人工智能项目的开发与实践。案例篇则主要选取橘子识别与情绪识别实例，深入剖析应用实例，使学生直观感受人工智能在日常生活中的应用与技术基础。实践篇通过项目化学习，

让学生自主选择并实践人工智能项目，锻炼自主学习、问题解决及创新实践能力。最终，学生通过作品展示分享学习成果，展示其人工智能实践成果。

整个课程紧密贴合学生的生活经验，以项目为核心，通过逐步深化的学习过程，培养学生的实际应用能力和创新思维。教师在其中扮演引导者的角色，促使学生在探索中成长，为未来的人工智能应用奠定坚实基础。

表 1 基于 OpenInnoLab 的初中人工智能校本课程内容架构大纲

课时	内容大纲
第一课时	理论篇——人工智能伦理
第二课时	工具篇——认识OpenInnolab平台
第三课时	案例篇——智能家居（积木式编程）
第四课时	案例篇——橘子数据集制作
第五课时	案例篇——橘子识别模型训练
第六课时	案例篇——橘子识别模型推理和部署到硬件
第七课时	案例篇——情绪识别数据集制作
第八课时	案例篇——情绪识别模型训练
第九课时	案例篇——情绪识别模型推理和部署到硬件
第十课时	实战篇——选题
第十一课时	实战篇——数据集制作
第十二课时	实战篇——模型训练和改进
第十三课时	实战篇——模型推理和部署到硬件
第十四课时	作品展示汇报
第十五课时	作品展示汇报
第十六课时	作品点评和教师反馈

3.3. 基于 6E 教学法的教学模式

6E 教学法主要包括六个过程：引入（engage）、探究（explore）、解释（explain）、工程（engineer）、丰富（enrich）、评价（evaluate），在基于 OpenInnolab 平台的初中人工智能校本课程中采用基于 6E 教学法的教学模式设计人工智能课程，旨在全面激发学生兴趣，培养其主动探究与解决问题的能力。

3.4. 评价机制设计

评价的核心目标在于激发学生对人工智能技术学习与实践的热情，助推其人工智能素养全面发展。为此，本研究采取过程性与总结性评价双管齐下的机制，旨在优化人工智能课程教学过程与学习成效。

过程性评价主要体现在对学生在三个主要篇章学习过程中表现出的主动参与、知识理解、技能掌握、问题解决、小组协作、创新思维等方面的持续观察、记录与反馈，根据过程性评价教师及时调整教学策略，学生及时查缺补漏以促进其全面发展。总结性评价集中体现在实战篇中，学生以小组形式创作小组作品，并进行作品展示与汇报，其间进行自我评价、小组间互评及教师专业点评打分，旨在全面考察学生在实战环节的知识应用、团队协作、创新思维及成果呈现等多维度能力。

4. 基于 OpenInnoLab 平台的初中人工智能校本课程实践案例

4.1. 实践对象

校本课程开发后，本研究在温州大学附属茶山实验中学人工智能科创社团课程进行了教学实施，实施对象为七年级部分学生，课程安排为每周一课时。

4.2. 实践案例

整个学习流程如图 1，下面以人工智能教育探究主题“情绪识别小助手”为例，阐述具体的流程。

在引入（engage）环节，通过设定真实情境问题，引起学生的学习兴趣，促使学生运用已有的知识解决问题，确定学习目标和任务。学生面临的学习压力与情绪困扰催生了对具备情绪识别与管理功能工具的需求，期望借此提升情绪自控力、准确洞察他人情绪以促进人际交往与合作。在上述真实情境问题背景下，学生基于对人工智能作品的认知与兴趣，决定研发“情绪识别小助手”项目以应对所面临的挑战。

在探究（explore）环节，学生围绕真实的问题情境，通过自主以及教师引导的探究，小组内部展开讨论，参考既有人工智能作品，初步构想项目设计方案，涵盖数据集制作、模型训练、推理及部署等核心环节。教师激励学生勇于创新，设计独具特色的“情绪识别小助手”方案。

在解释（explain）环节，各小组成员生动展示并详述其设计方案，凸显其在情绪识别领域的独特性、创新性与可行性，可能辅以图表、模型等可视化工具以清晰传达设计理念。此环节不仅局限于小组内部交流，更鼓励组间互评，让各组得以汲取他组智慧，激发创新灵感，同时通过提问与建议推动跨组深度交流与合作。教师在此过程中起到关键指导作用，通过严谨的评价与针对性反馈，助力学生精益求精地完善“情绪识别小助手”设计方案。教师适时抛出深度问题，引导学生深入思考设计的核心问题，持续激发其优化方案的动力。

在工程环节（engineer），根据优化后的“情绪识别小助手”设计方案，学生动手将其转化为实体模型，并通过反复试错与调整，持续提升原型性能与功能完备度。

在丰富（enrich）环节，当学生在开发“情绪识别小助手”过程中遭遇知识与技能瓶颈时，教师引导其回溯理论基础，拓宽相关概念理解，构建新知识在项目中的应用框架，旨在提升学生知识迁移与问题解决能力。此时，教师还会指导学生对已积累的实践经验进行深度反思，并鼓励其积极进行项目拓展，以进一步丰富“情绪识别小助手”的功能与应用场景。

在评价（evaluate）环节，借助过程性评价与总结性评价，结合教师的专业意见、其他小组的建设性反馈以及组内自我审视，对“情绪识别小助手”项目进行全面而深入的优化。

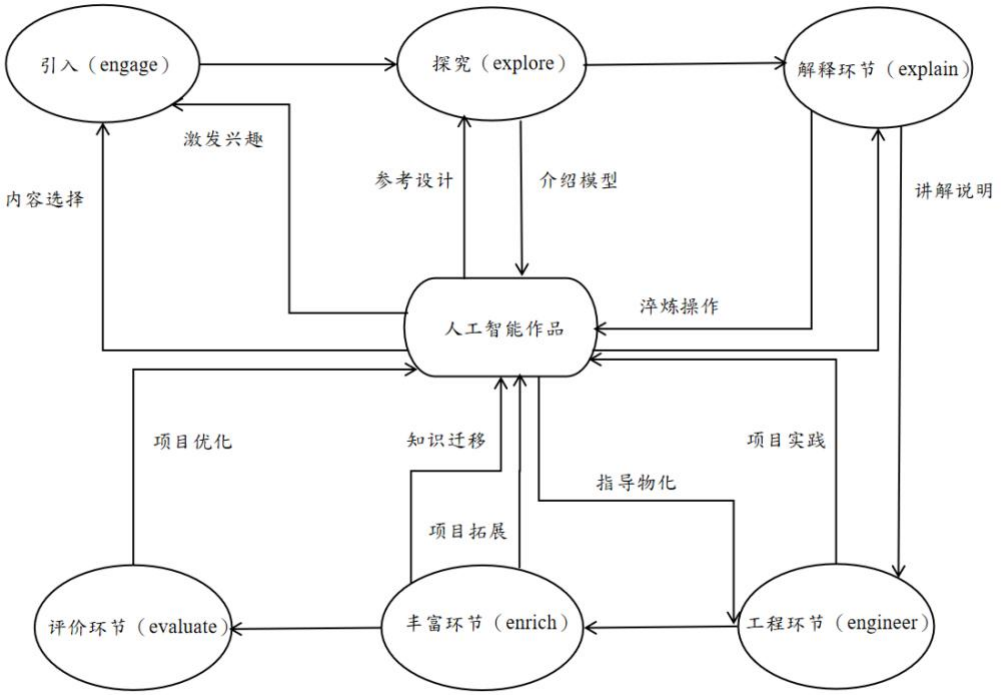


图 1 学习流程图

5. 基于 OpenInnoLab 平台的初中人工智能校本课程实践展示

在初中人工智能科创社团活动中，学员们凭借 OpenInnolab 平台的助力，展现出卓越的创新意识和技术执行力。他们不仅系统学习并牢固掌握人工智能的基础知识，还能熟练地将所学知识应用于实践活动，创造出一批极具吸引力的作品。课程结束时，如图 1 到图 6 所示，每个小组学员们满怀热情地向教师和其他小组学员展示了自己的学习成就。



图 2 校服识别小助手



图 3 路口神探



图 4 牙齿分辨器



图 5 猫窝里的奇幻世界



图 6 花草识别小助手



图 7 情绪识别小助手



图 8 校园小型作品展览——学员作品互动解说现场

午餐间隙，社团举办了一场校园小型作品展览，迅速吸引了全校广大同学前来体验和互动，近距离地感受人工智能的魅力。如图 7 所示，参与展出的学员们十分热心地解说各自作品的功能和创新点，生动展示了该社团课程将理论知识与实践相结合的教学理念与成果。

6. 结语

本研究在《义务教育信息科技课程标准（2022 年版）》的指导下，充分利用了 OpenInnoLab 这一人工智能开放创新平台，成功开发了一套植根于校本情境的初中人工智能课程。这套课程精心构筑了总体目标与阶段性目标体系，紧密围绕人工智能核心技术脉络，以项目化学习为载体，巧妙编排为 16 个富有挑战与探索性的课时单元。

在教学实践中，学生们循序渐进地经历了数据采集、参数调试、模型训练、推理验证直至模型部署的完整生命周期，亲手揭开人工智能技术运作的神秘面纱。这一过程不仅使他们在深度学习原理的探究中实现自主知识建构，更促使他们运用所学，将训练成熟的模型迁移至行空板等硬件设备，创造出颇具实用价值的人工智能作品。

然而，受限于笔者实战经验与专业知识，课程在内容选择、资源整合、活动设计等方面存有局限。面对这些问题，我们将结合反馈，着力跨学科融合与数字化评价，不断优化课程，旨在提升其科学性、趣味性与实效性，培育未来所需复合型 AI 人才，推动义务教育阶段信息科技教育高质量发展。

参考文献

[1]Yang, Y., Qin, J., Lei, J., & Liu, Y. (2023). Research Status and Challenges on the Sustainable Development of Artificial Intelligence Courses from a Global Perspective. *Sustainability*, *15*(12), 9335.

[2]Ilkka, T. (2018). *The impact of artificial intelligence on learning, teaching, and education*. European Union.

[3]Trustworthy Artificial Intelligence (AI) in Education: Promises and Challenges.

[4]王波. (2020). STEM 教育中 6E 学习模式的教学初探. **生物学教学**, 45(1), 71.

高中人工智能教育实验室建设

——来自温州科技高级中学的经验

Construction of Artificial Intelligence Education Laboratory in High Schools

-- Experiences from Wenzhou Science and Technology Senior High School

陈恩慈[1]，林淼焱[1]
[1]温州科技高级中学
1090984675@qq.com,657894692@qq.com

【摘要】 本研究聚焦于高中人工智能教育实验室的建设实践，以温州科技高级中学为例，探讨了在当前教育背景下，如何有效推进人工智能教育的实施。研究基于国家政策支持和教育需求，通过问卷调查分析了教师对实验室建设的期望，并提出了“五+一”发展目标，涵盖师资、课程、学生、成果和平台建设，以及大语言模型的应用。研究详细阐述了实验室建设的具体目标、课程体系构建、设备配置和算力服务器的重要性。最终，总结了实验室建设的关键启示和对未来教育变革的展望。

【关键词】 人工智能教育；实验室建设；课程体系；算力服务器；教育创新

Abstract： *This study focuses on the construction and practice of artificial intelligence (AI) education laboratories in high schools, taking Wenzhou High School of Sci. & Tech. as a case study to explore effective strategies for implementing AI education within the current educational context. Supported by national policies and educational demands, the research analyzes teachers' expectations for laboratory construction through a questionnaire survey and proposes a "5+1" development strategy, covering five dimensions—faculty), curriculum, students, outcomes, and platform—and the application of large language models. The study elaborates on the specific goals of laboratory construction, curriculum development, equipment configuration, and the significance of computational power servers. Finally, it summarizes key insights from the laboratory construction and provides a perspective on future educational transformations.*

Keywords： *AI Education; Laboratory Construction; Curriculum Development; Computational Power; Educational Innovation*

1. 背景

人工智能快速发展对经济社会等方面产生着深远而重大的影响。中国高度重视人工智能的发展，并将其视为推动经济社会转型的关键技术。《新一代人工智能发展规划》明确强调了人工智能教育的重要性，并提出了在中小学阶段普及人工智能知识的要求。此外，《数字中国建设整体布局规划》也强调了数字技术在教育中的应用，推动教育资源的数字化、网络化和智能化，为人工智能教育的实施提供了政策支持。

高中阶段是基础教育的关键时期，也是开展人工智能教育的良好时机。尽管全国各地的教育部门和学校积极推动人工智能教育实践并取得一定进展，但现有教育资源难以满足当前高中学生对人工智能知识的渴求。新一代人工智能教育亟需一个支持深度学习和实践的先进学习环境。因此，建设高中人工智能实验室成为当务之急。人工智能实验室是一个学生实践和探索的平台，不仅可以培养学生的编程，数据分析和机器学习等技术，而且还能激发创新思维和培养问题解决能力。

2. 研究现状

在国外，人工智能教育的研究和实践主要集中在高等教育和 K-12 教育阶段。这些研究关注学生对人工智能的认知，包括核心能力和基于教与学的设计考虑。高等教育领域的研究表明，人工智能在适应不同类型学习者需求、提供定制化反馈、开发评估工具以及预测学业成

功等方面具有潜力。而 K-12 教育中，培养学生的人工智能素养是重点。此外，一些国家的教育机构和企业，如麻省理工学院和谷歌，提供了丰富的人工智能教育资源和平台，包括学习单元、教学工具和在线课程。在国内，人工智能教育同样受到了高度重视。国务院发布的《新一代人工智能发展规划》强调了人工智能教育的重要性，并提出了具体的实施目标和要求。国内已经出版了多种人工智能教材，涵盖了人工智能的基本概念、核心算法、典型应用等内容。

尽管国内外对于人工智能实验室的研究已经有一些了，但是大多集中在理论框架、教材研究的层面。在义务教育阶段以及高中阶段，人工智能实验室及课程建设关联的研究相对较少。高中人工智能实验室的建设主要集中在个别学校的探索和实践中，部分学校的实验室建设以机器人、竞赛类代码编程为主，没有算力，不谈模型，是较为传统的“信息技术机房”，而非“人工智能实验室”。温州科技高级中学作为一所以科技为特色的高中，有必要建设一个高中人工智能实验室，以适应新一代人工智能教育的需要。通过整合各方面资源开发一套与高中课程标准相适应的人工智能课程体系，并配置各种软硬件设施，达成一个可复制、可推广的高中人工智能实验室建设模式。



图 1 温州科技高级中学人工智能科创中心

3. 实验室建设

3.1. 问卷调差结果

2024 年 2 月，上海人工智能实验室发起了一项针对“如何筹建学校的人工智能实验室？”的问卷调查，得到了一线信息技术和科创教师的广泛支持。这份问卷反映了教师们对于人工智能实验室建设的多样化需求和期望，以及在资源、师资、课程开发等方面的具体挑战。教师们普遍认为实验室应具备多功能性，既能服务于广泛学生群体，也能满足特定兴趣学生的需求，并且强调了实验室在教学和学生创作中的重要性。同时，问卷也揭示了在筹建过程中可能遇到的困难，如资金限制、师资不足、课程缺乏等。

3.2. 实验室建设目标

根据问卷调查结果结合温州科技高级中学学校自身的办学理念，我们了确定了“五+一”发展目标，“五”指的是五大维度——师资、课程、学生、成果、平台，“一”指的是一个辅助——大语言模型。

在温科高的人工智能教育发展蓝图中，五大发展维度构成了坚实的基石。通过构建一支融合多学科专长的师资队伍，开发创新的人工智能课程体系，培养具备科学精神和实践能力的学生。一个辅助代表利用人工智能大预言模型辅助师生教学，实现教育教学成果的优化与

辅助，在此合力之下，将本校搭建为一个有地区影响力的科创平台，共同推动人工智能教育的全面发展。

3.2.1. 五大维度均衡发展，铸造人工智能体系基石

①师资：打造以人工智能科创教育为特色，由多个学科教师组成，国内顶级科创专家引领的STEAM 科创队伍；定期组织教师参与人工智能相关的培训和研讨会，全方位培养科创教育青年骨干力量，在国内各类科创活动中展示学校特色项目的教学成果，在各种教学期刊中发表文章，并打磨课程参与各种评比和展示，定期出版课程配套教材或者教学指南。

②课程：开发以人工智能教育为核心，以真实世界前沿科技为主要学习内容，以跨学科解决真实问题为方向，以项目化学习为形式的科创教育课程体系；将相对成熟的课程教材、实践手册结集出版。

③学生：设立“科高未来讲坛”，邀请不同领域专家开展面向学生的系列讲座，培养全体学生的科学精神，提升科学素养；提高科创类社团的占比，结合人工智能通识课，培养一批具备人工智能科创能力的未来人才，在各类科创比赛中获得佳绩。

④成果：定期整理教育教学成果与师生科创项目，2025 年开始每年编印一册《温科高人工智能科创项目集锦》。通过校内外展览、研讨会等形式展示优秀的学生项目和教师教学案例。

⑤平台：形成辐射作用，定期开放人工智能科创中心，供小学、初中学生参观体验，并联合市科协、市教育局等单位组织公益夏令营活动；承办中国科协青少年科技中心、上海人工智能实验室、中国电子学会等单位组织的全国性人工智能科创活动，推广学校的科创教学成果。



图 2 五大维度

3.2.2. 重点打造校内大语言模型，助力常识推理与知识归纳

大语言模型是人工智能大模型是指具有庞大参数规模、复杂结构和强大学习能力的人工智能模型。这些模型通常基于深度学习技术，由数十亿甚至数万亿个参数组成，可以通过在大规模数据上进行预训练和微调来学习和推理。[2]

除了调用他人训练好的模型，我们还可以造一个“自己的大模型”。在文本问答领域，常见的开源大语言模型框架有 InternLM，Llama，ChatGLM 等，要想让大语言模型能够回答特定领域的问题，我们可以将上述框架作为基底进行微调，也就是使用进行用各领域内的数据集进行训练，将其回答内容往对应领域对齐。假设将数据集替换为程序语言数据集，我们能得到类似的精准回复的效果。下图为我校教师积极研究编程大语言模型辅助编程教学的思考。



图 3 使用大模型进行编程辅助 开发校内编程平台与各类课例

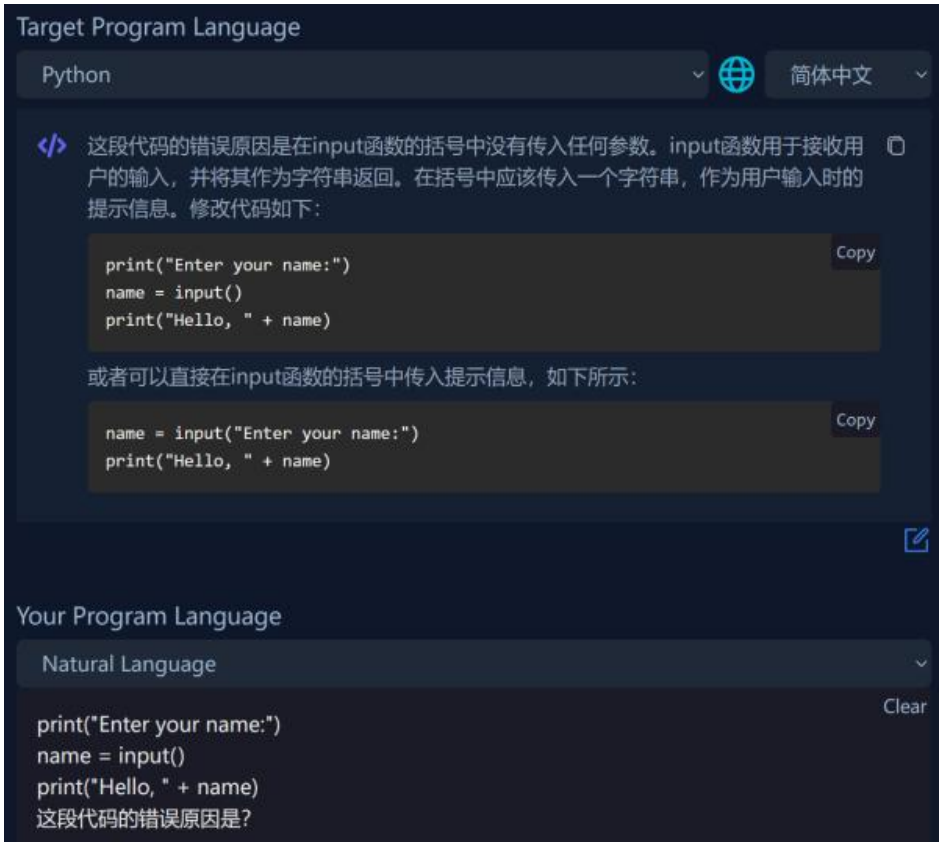


图 4 在基础课堂教学中 引导学生通过编程辅助工具及时纠错



图 5 在通识课中 利用 llama 部署 Qwen1.8B 大模型完成 MV 项目中歌词的创作

3.2.3. 校内多模态人工智能应用开发

在实现大语言模型的合理应用后，我们还可以将大语言模型和智能语音，计算机视觉结合起来 ，结合校内社团开发各类应用。

①智能客服与虚拟助手：

结合大语言模型的自然语言处理能力，智能语音的语音识别和合成技术，以及计算机视觉的图像理解能力，可以创建一个能够通过语音交互、理解和回应用户问题的智能客服系统。计算机视觉可以用来分析用户的情绪或行为，从而提供更加个性化的服务。

②内容创作与媒体：

在内容创作领域，大语言模型可以生成文本内容，计算机视觉可以分析和编辑图像和视频，智能语音可以用于配音和语音解说。这样的组合可以大大提高内容创作的效率和质量，创造出更加丰富和互动的媒体体验。

③智能家居与物联网：

在智能家居系统中，计算机视觉可以用于监控家庭安全和识别用户行为，智能语音可以用于接收和执行用户的语音命令，而大语言模型可以用于理解和预测用户的需求，提供更加智能和便捷的家居管理。

3.3. 课程体系建设

《中小学人工智能课程指南》指出，中小学人工智能课程是一门旨在全面提升学生人工智能素养，帮助学生掌握人工智能技术基础知识与技能，增强智能意识，激发智能思维，提升智能应用与创造能力，培养智能社会责任的基础课程。[1]课程需要内容充分考虑到中小学生的认知水平和理解能力，通过具体的情境体验和动手操作来促进知识的掌握。

结合《中小学人工智能课程指南》，在课程设计中，笔者所在学校开设了普适性课程、社团类课程、竞赛类课程旨在通过多元化的课程设置，满足不同学生群体的需求，同时激发和培养学生对人工智能领域的兴趣和能力。

本研究以普及教育为出发点，面向所有学生，旨在培养学生的人工智能核心能力，选择与学生日常生活和学习紧密相关的人工智能应用场景作为教学案例。通过这样的教学方法，学生能够在亲身体验和实践的过程中，逐步感知、理解和掌握人工智能的基本概念、技术和操作方法。在通识课程的基础上，鼓励学生去探索人工智能在日常生活中的应用潜力，激发他们提出创新的问题解决方案，并培养他们批判性地评估人工智能的能力，从而培育他们的问题解决技能、创新意识和批判性思维，以适应不断发展的智能社会。

3.3.1. 面向全体，开设 AI 通识课程

我们将面向全体学生，开出人工智能通识课程，不讲技术原理，只讲应用情况，向学生科普对人工智能的发展历史、基本概念、生活中的应用场景、常见工具平台等，能够从跨学科的视角思考人工智能伦理、治理以及人工智能与社会发展之间的关系。让学生早一点、深入一点了解人工智能及其带来的社会变革，对人工智能有基本认识，培养学生的未来视野。该课程以校内科创团队授课为主，不定期邀请校外专家开设讲座，主要讲授人工智能前沿动态、人工智能科普知识，帮助学生丰富科学素养，提升科学认知，了解 AI 时代的社会动态，培养适应智能社会发展的未来公民。

3.3.2. 面向部分，开设 AI 兴趣课程

目前已经组建了科创类社团“深度学习和无人车”“天空探险家”“智联网研究所”，在高一下学期、高二，我们还将组建其它人工智能科创社团，逐步开出无人机、智能车视觉识别、物联网开源硬件、人工智能大模型应用等课程。这些课程将在真实环境中探究实践，熟悉设备操作；利用深度学习模型，完成语言处理、图像识别、语音识别等任务（制作一个基于图像识别的垃圾分类系统、制作一个基于图像识别坐姿矫正系统、基于基于自然语言理解诗词推荐系统、利用 Ollama 玩转本地大模型）；让学生以智慧校园作为人工智能实践场所，结合智慧校园的各种场景（智能农场、停车场车辆接送、访客预约、安全隐患预警），引领学生运用人工智能解决现实问题，把奇思妙想、兴趣爱好变为现实，培养学生科学探究活动的兴趣和能力，培养创新意识和创新素养。

3.3.3. 满足个性，开设 AI 竞赛课程

从科创类社团中挑选对科创感兴趣、有潜质的学生进行深度培训，让学生进一步受到科学精神的感染与熏陶，拓展知识技能，培养他们的硬件操控和工程能力，参加全国中小学信息技术创新与实践大赛（NOC）、全国青少年人工智能创新实践交流展示会、长三角青少年人工智能奥林匹克挑战赛、青少年无人机大赛、IOAI 国际人工智能奥林匹克等竞赛活动。



图 6 温州科技高级中学人工智能课程体系

3.4. 实验室设备配置

基于人工智能课程体系以及人工智能科创中心总体规划，我们计划采购以下设备，

3.4.1. 硬件

核心器材：

①核心服务器 1 台，提供算力用于模型训练，实现模型训练本地化。

②拿铁熊猫 30 套，安装 Linux/Windows 双操作系统。

③数据采集外设若干（罗技 C270 摄像头、esp32 摄像头、耳机带麦克风）。

④行空板、掌控板等开源硬件若干。

⑤个人工作站 2 台，配置高性能的显卡。

⑥一体机一台。

服务器机箱机:2U
CPU:2 颗英特尔® 至强® 银牌 4216
内存:2*32GB RDIMM,3200MT/s,双列
硬盘:1*480GB SSD SATA 读取密集型 6Gbps 512 2.5 英寸热插拔 AG 硬盘
硬盘:2*4T10K RPM SAS 12Gbps 512n 2.5 英寸热插拔硬盘
阵列卡:PERC H330 RAID 控制器,
适配器电源:双,热插拔,冗余电源(1+1)
服务器机箱机:2U
CPU:2 颗英特尔® 至强® 银牌 4216
内存:8*32GB RDIMM,3200MT/s,双列
硬盘:1*480GB SSD SATA 读取密集型 6Gbps 512 2.5 英寸热插拔 AG 硬盘
阵列卡:PERC H730P RAID 控制器,2GBNV 缓存,适配器
网卡 1:英特尔 X520 双端口 10GbE SFP+适配器,PCIe 半高
电源:双.热插拔,冗余电源(1+1).1100 瓦
GPU:2 块 NVIDIA® Teslarm T4 16GB,被动式,单宽,全高 GPU
电源:双,热插拔,冗余电源(1+1)

服务器配置详情

表 1 其余器材清单（部分）：

序号	设备	型号	数量
1	3D 打印机	拓竹 X1-Carbon Combo	2
2	桌面级激光切割机	xTool P2	1
3	大型激光切割机	雷宇 Nova35	1
4	行空板	DF Robot	22
5	掌控板	DF Robot	50
6	传感器套件	DF Robot	22
7	掌控宝	盛思 2 代	30
8	掌控板扩展板	盛思	30
9	AI 下棋机器人	商汤 元萝卜 SenseRobot	2
10	AI 视觉套件	商汤 商小鸣	10
11	编程无人机	大疆 Rmbomaster TT	16
12	编程机器人	大疆 Rmbomaster EP	2
13	智能小车	幻尔 AiNova	10
14	MakeX Challenge 赛事套装	MakeX	2

3.4.2. 软件

OpenXLabEdu 简称 XEdu ，是为中小学 AI 教育设计的一套完整的学习工具。核心工具为计算机视觉库 MMEdu，加上神经网络库 BaseNN 和传统机器学习库 BaseML，覆盖了中小学可能涉及到 AI 技术的所有领域。我们会在每一台台式电脑和拿铁熊猫上安装 XEdu 的离线包，实现 XEdu 的本地化。



图 7 XEdu 的特点

与企业合作，利用高性能专用工作站与算力服务器，基于开源人工智能教学一体化平台Open-Hydra，搭建校内算力服务中心，满足人工智能日常化教学中的各类需求，如算力、编辑器、用户管理等。



图 8 OpenHydra 界面

3.4.3. 为什么要配置算力服务器

数据、算法和算力被公认为推动人工智能发展的三驾马车。正如华为创始人任正非先生所强调的，大算力是第四次工业革命的基石。在人工智能领域，算力是执行复杂算法和处理大量数据的关键。没有足够的计算能力，就无法实现深度学习、机器学习和其他 AI 技术的发展。在做一个真实的人工智能项目的时候，经常需要自定义模型和训练数据集，以达到解决具体问题的精度要求。算法是开源的，数据是搜集的，而算力是需要购买的，算力的强弱直接影响到模型训练的效率和质量。相较于 CPU，使用 GPU 练模型速度快了很多。在图像分类任务中，使用 GPU 进行模型训练比 CPU 快了约十倍，显著缩短了训练时间。在高中学业压力本来就比较大的情况下，提供一台算力服务器能帮助他们快速训练模型，提高效率。在经费不足的情况下，我们也可以使用浦育平台提供的 GPU 云算力进行模型训练。同时算力服务器也有机串联起

笔者相信，随着第四次工业革命的兴起，算力服务器正逐渐成为人工智能实验室不可或缺的核心设施。

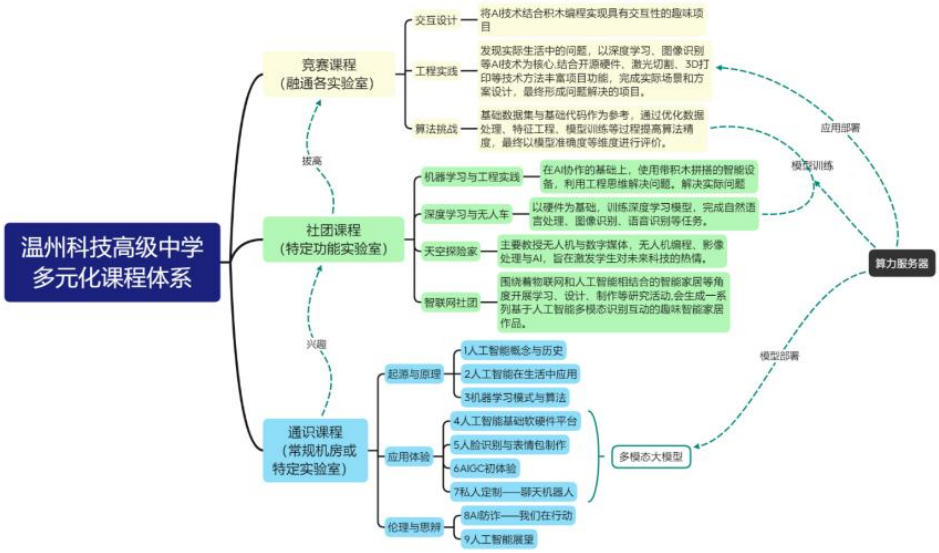


图 9 算力服务器在温州科技高中课程体系中的运用

4. 实验室建设的启示和经验

在温州科技高级中学高中人工智能教育实验室的建设实践中，我们获得了一系列宝贵的启示和经验：

（一）整合政策与规划先行——国家政策的引导和长远规划是实验室建设成功的基石。实验室的构想与实施应与国家教育战略紧密对接，确保项目的方向性和可持续性。

（二）深入的需求调研——通过问卷调查和深入分析，我们认识到了解教师和学生的实际需求对于实验室的功能设计和课程开发至关重要。这种需求导向的方法确保了实验室能够满足用户的期望，并有效支持教学和学习活动。

（三）算力的核心地位——算力是人工智能教育的心脏，它不仅是实现人工智能教育的关键基础设施，更是提高教学效率、支持学生进行深入学习和实践的重要工具。其配置水平直接影响到实验室的教学和研究能力。投资高性能的算力资源，尤其是 GPU 服务器，对于执行复杂的 AI 算法和数据处理至关重要。

（四）多功能与灵活设计——实验室的设计应具备多功能性和灵活性，以适应多样化的教学和研究需求。这样的空间不仅支持传统的教学活动，还能够鼓励创新和跨学科合作。

（五）师资队伍建设：优秀的师资队伍是实验室成功的关键。通过定期的专业培训和学术交流，教师能够不断提升自己的专业技能，并更有效地指导学生进行人工智能的学习和研究。

（六）课程的持续迭代——随着人工智能技术的快速发展，实验室的课程体系需要不断更新和优化。持续的课程迭代确保了教学内容的时效性和适用性，满足了学生不断变化的学习需求。

（七）合作与资源共享——与企业和高校的合作为实验室带来了丰富的资源和技术，这种合作模式不仅增强了实验室的综合实力，也为学生提供了接触前沿技术的机会；通过众筹课程的方式让更多的学校和老师参与课程的开发共同开发课程，汇聚更多的教育资源，促进资源的共享。

（八）成果的展示与交流：通过参加或举办活动和比赛，使实验室的成果得以展示和交流，这不仅提升了学校的声誉，也激发了学生和教师的参与热情。

高中人工智能教育实验室的建设是一个综合性系统性的工程，涉及政策支持、需求调研、

设计规划、资源配置、师资培养、合作交流、课程开发、资金管理等多个方面的考虑和精心规划，通过吸取这些经验和启示，让未来的实验室建设将更加高效、实用，更好地服务于人工智能教育的发展。

5. 结语与展望

在本研究中，我们深入探讨了高中人工智能教育实验室的建设与发展，旨在为高中阶段的学生提供一个实践和探索人工智能技术的平台。

随着人工智能技术的不断进步和教育模式的创新，展望未来，我们将继续探索和完善高中人工智能教育实验室的建设方案，我们期待人工智能实验室能够继续发挥其在教育前沿的作用，不仅是学生学习人工智能知识和技能的平台，更是激发学生创新思维、实践能力和跨学科协作精神的摇篮，成为推动教育变革、培养未来创新人才的重要力量。在此，我诚挚地邀请更多的教育同仁加入到这一伟大的事业中来，让我们携手并进，共同书写人工智能教育的新篇章。

参考文献

中小学人工智能课程指南课题组. (2023). 中小学人工智能课程指南. 华东师范大学学报（教育科学版）, 41(3), 121-134. https://doi.org/10.16382/j.cnki.1000-5560.2023.03.013

中国人工智能学会. (2023). 2023 中国人工智能白皮书-深度学习. https://www.caai.cn/index.php?s=/home/article/detail/id/3156.html

借势·借域·借力

——基于浦育平台在农村小学开展新一代人工智能教育的实践

Borrow power·Borrow territor·Borrow power

—— The practice of developing a new generation of artificial intelligence education in rural primary schools based on the Openinnolab

陈慧娴[1*]
[1]平阳县昆阳镇第六小学
[*]38917141@qq.com

【摘要】 温州市教育局计划推动人工智能与基础教育深度融合，实现全市中小学人工智能课程的全面普及，以深度学习技术为代表的“新一代人工智能教育”被提出。然而，由于农村小学设备紧缺、教学环境落后等原因，实现人工智能教育的全覆盖面临困难，浦育平台的使用可以弥补这一缺陷。本文通过探讨在平阳县昆阳镇第六小学这一农村区域试点小学使用浦育平台，开展新一代人工智能教育的教学实践，旨在解决农村地区人工智能教育的难题。

【关键词】 深度学习技术；新一代人工智能教育；农村小学；浦育平台；教学实践

Abstract: *Wenzhou Education Bureau plans to promote the deep integration of artificial intelligence and basic education, achieve the full popularization of artificial intelligence courses in primary and secondary schools in the city, and the "new generation of artificial intelligence education" represented by deep learning technology is proposed. However, due to the shortage of equipment in rural primary schools, backward teaching environment and other reasons, it is difficult to achieve full coverage of artificial intelligence education, and the use of Openinnolab can make up for this defect. This paper discusses the use of Openinnolab in the pilot primary school of Kunyang Town Sixth Primary School in Pingyang County to carry out the teaching practice of a new generation of artificial intelligence education, aiming to solve the difficult problem of artificial intelligence education in rural areas.*

Keywords: Deep Learning Technology, New Generation Of Artificial Intelligence Education, Rural Primary Schools, Openinnolab,Teaching Practice

1. 前言

随着人工智能技术的快速发展和广泛应用、STEM 教育的重视以及未来工作岗位的变化和人才需求的新挑战，我们正推动着人工智能教育的发展和普及。2017 年，国务院印发了《新一代人工智能发展规划》[1]（以下简称《规划》），其中指出“人工智能发展进入新阶段”，这也预示着中小学的人工智能教育将逐步转变成以深度学习技术为代表的“新一代人工智能教育”。2022 年温州市教育局发布了《温州市中小学推进人工智能教育实施方案》[2]（以下简称《实施方案》），提出“将人工智能纳入温州中小学课程体系中，全市中小学将于 2025 年实现人工智能课程全面普及”。

在新趋势下，农村小学将面临许多困难，包括地处偏远、资金紧缺、缺乏适合的人工智能教学环境、相关设备和师资等。为了解决这些问题，我们可以在信息科技课堂上使用摄像头和耳麦这 2 个简单的硬件设备，并借助上海人工智能创新中心开发的浦育平台，让学生能

够体验和创作人工智能作品，从而让新一代人工智能教育触手可及。

基于此，在本文中对平阳县昆阳镇第六小学这一农村小学，在信息科技课堂上使用浦育平台，开展新一代人工智能教育的教学实践进行探讨。

2. 借教育之势——何为新一代人工智能教育？

2.1. 教育政策出台

《规划》指出“人工智能发展进入新阶段”，即为新一代人工智能。对于教育者而言，在国家教育政策的助推下，首先需要了解的就是新一代人工智能技术的含义以及与传统人工智能教育的区别。新一代人工智能教育更注重培养学生对于深度学习技术的学习与应用能力，学生将具备更强的创新与实践能力，能够运用掌握的人工智能技术来解决实际问题，并能够创造出更多的人工智能应用场景。例如，使用浦育平台的“AI 体验”模块，通过关键词的输入，就能形成一幅与关键词相关联的图画（图 1），无需再从海量的网络资源中寻找合适的图片。

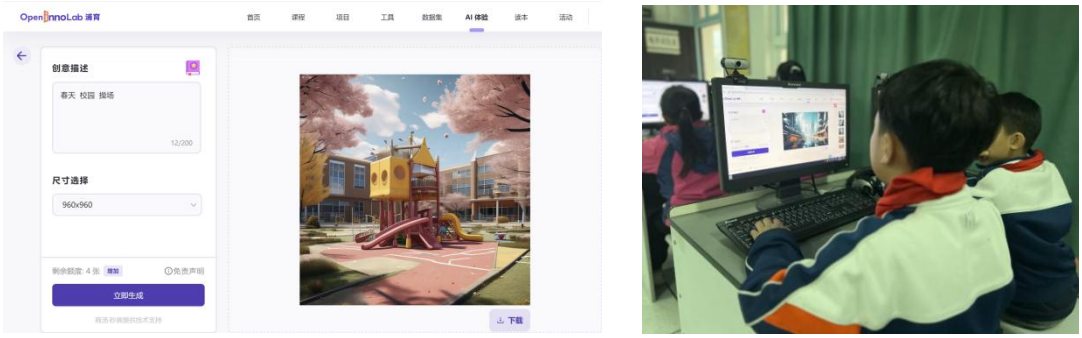


图 1 应用浦育平台“AI 体验”模块

2.2. 目标与内容

新一代人工智能教育的宏观目标是培养具备应对未来人工智能时代发展需求的人才。具体而言，该目标包括至少两个方面：一是培养适应人工智能时代的人才。例如，未来的知识学习将更多由类似 ChatGPT、“文心一言”等大语言模型辅助完成，更注重学生基于知识学习的素养发展。二是培养能够研发人工智能的人才[3]。例如，降低学习和研发人工智能的门槛，将冗长的深度学习技术代码简化封装成小学生可理解的图形化编程（图 2），让深度学习技术可视化。



图 2 浦育平台“创意积木工坊”图形化编程工具

人工智能的三大核心要素是数据、算法和算力[4]。为迎接未来人工智能时代，新一代人工智能教育将涵盖人工智能基础知识、编程技能、数据科学、深度学习技术等多方面内容，以支持这三大核心要素的发展。

未来的人工智能时代将是“人机共智”的现代社会，学生将能够掌握以深度学习技术为核心的新一代人工智能技术。新一代人工智能教育将为此做出积极贡献，培养创新思维和实践能力，拓宽知识面，增强竞争力，从而更好地适应未来社会的发展需求。这将为实现人与智能科技的融合提供坚实基础，推动“人机共智”时代的到来，为现代社会发展注入新的活力和动力。

3. 借试点之域——为何在农村小学开展？

3.1. 必要性与重要性

根据《实施方案》，全市中小学将于 2025 年全面普及人工智能课程，即使是地处偏远、条件相对落后的农村小学也不例外。然而，农村信息科技教师面临着教育资源匮乏的挑战，开设人工智能课程存在难度。不过，互联网为农村教师提供了搭建学习桥梁的机会，帮助他们获得必要的教育资源，了解前沿的人工智能技术、提高人工智能教学水平，为在农村小学开展人工智能教育提供了有力保障。对于学生而言，作为教育的主体，学习新一代人工智能教育能够帮助他们更好地适应未来的发展趋势，培养创新思维、提高解决问题的能力和团队合作精神。

3.2. 软硬件设备支持

为了顺利开展新一代人工智能教育，学校需要配备能够支持创作人工智能项目的软硬件设备。然而，市面上的硬件设备价格高昂，这使得让每位学生都能自主创作人工智能项目作品变得困难。同时，编程软件的繁琐登录方式也不利于在学校开展人工智能项目的编程创作。鉴于这些问题，作为农村区域试点小学的平阳县昆阳镇第六小学，学校为唯一的一个机房配备了 48 个摄像头和耳麦，并选择了国家支持、免费对外开放、登录方式简单的浦育平台（图 3）。通过降低设备购置成本，学校为每位学生都提供了进行人工智能作品自主创作的可能性。



图 3 机房设备配备情况

4. 借平台之力——如何在农村小学实践？

4.1. 课时保障与课程安排

平阳县昆阳镇第六小学为保障学生学习新一代人工智能技术，教师结合学生的编程和人工智能的学习情况，为不同年级的学生开设《浦育之旅》社团课，分别是：每周三下午两节四、五年级社团课（26 人）和每周五下午一节三年级社团课（19 人），并且为每位学生创建浦育平台个人独立账号，在“创意积木工坊”模块实现人工智能作品的自主创作。

尽管学生们都没有人工智能编程学习的基础，但是两个社团课班级中的学生在信息科技学习基础、学习水平和学习能力方面存在较大差异。为了照顾不同学生的需求，教师进行了综合考虑，并制定了不同的教学计划表（图 4）。

浦育之旅－－四、五年级
教学计划表 2023学年第2学期

周次	教学内容	课时
2	《初识浦育平台》	2课时
3-4	《指尖上的钢琴家》	4课时
5-6	《智能电视》	4课时
7	《百变表情包》	2课时
8	《遇见甲骨文》	2课时
9-10	《校园植物大搜索》	4课时
10（11）-12	《校园智能门禁》	4课时
13-14	《音乐指挥家》	4课时
15-16	《一起来学八段锦》	4课时
17-18	现场任务挑战	4课时

浦育之旅－－三年级
教学计划表 2023学年第2学期

周次	教学内容	课时
2	《初识浦育平台》	1课时
3-4	《指尖上的舞蹈家》	2课时
5-6	《智能电视》	2课时
7-8	《红领巾检测仪》	2课时
9-10	《你是我的眼》	2课时
12（11）	《芝麻开门》	2课时
12-13	《百变魔术师》	2课时
14-15	《广播体操检查器》	2课时
16-18	现场任务挑战	3课时

图 4 《浦育之旅》社团课教学计划表

4.2. 《智能电视》项目实践

以《浦育之旅——三年级》教学计划表中的《智能电视》人工智能项目为例，该项目是借助摄像头，使用浦育平台“创意积木工坊”模块的“AI 训练（手势分类）”扩展模块开展以深度学习为代表的新一代人工智能教育的项目案例。项目通过手势识别切换电视节目，体验基于 AI 训练（手势分类）的人机交互方式。

●数据——训练数据的采集、整理和数据模型的训练

学生打开 AI 训练（手势分类）训练器，为两类手势命名，并采集一定数量的这两类手势的样本数据，再对采集的数据集进行训练，生成能够对这两类手势进行分类的模型（图 5）。

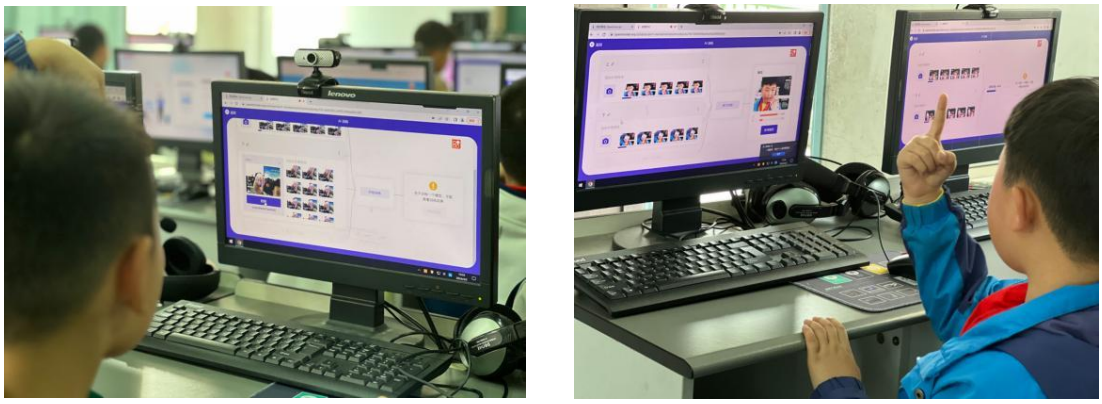


图 5 数据模型训练

●（人工智能）算力——目标结果输出的计算能力

模型训练受到模型运行服务的计算机系统能力的影响（图 6）。将学生采集添加的两类手势样本数据形成的数据集，输出生成两类手势分类模型的处理能力体现了人工智能算力。

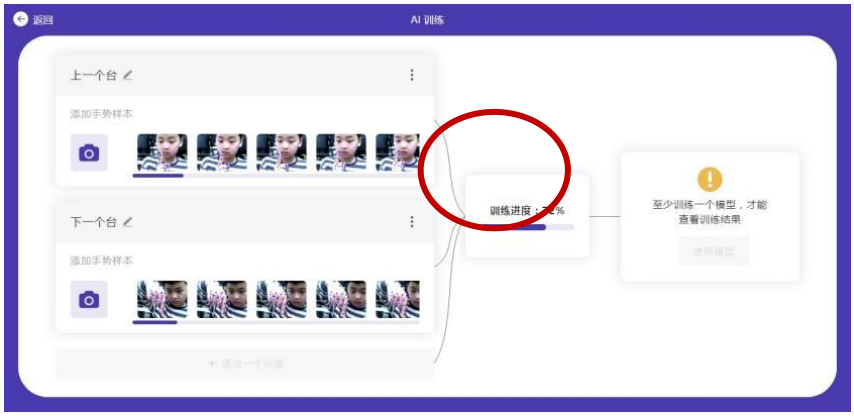


图 6“创意积木工坊”模块人工智能算力的体现

●算法——解决问题的方法与步骤

使用已训练好的能够对两类手势进行分类的模型，通过持续对当前帧进行预测，当预测结果的置信度达到一定要求时，确定预测结果的标签，从而实现上下手势识别并切换不同的电视节目（图 7）。

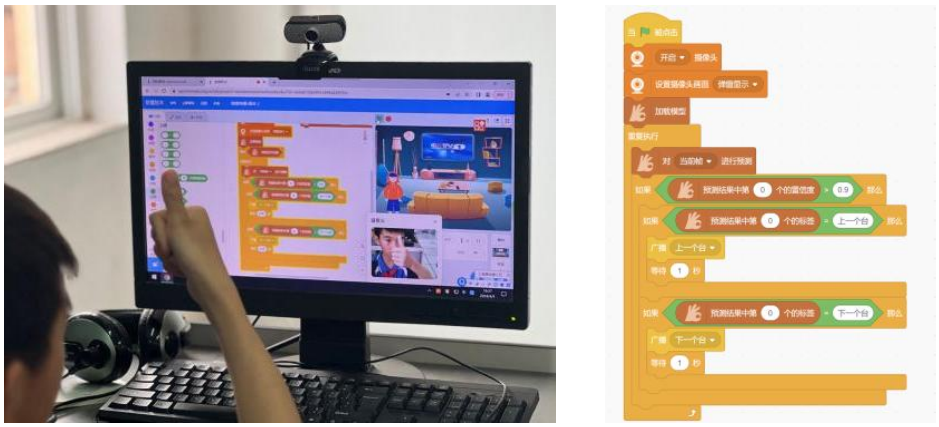


图 7 设计并测试算法

学生通过参与实践《智能电视》项目，亲身体验了数据采集、数据集制作、手势分类模型训练以及程序编写的全过程，使他们在简化的学习过程中了解并学习了深度学习技术。不仅是 AI 训练（手势分类）扩展模块，还有其他能够通过 AI 训练进行分类的扩展模块(图 8)，所有拓展模块的指令都是相似的，只是 AI 分类训练的类型不同而已。所以只要拥有足够丰富的数据，借助算力和开源算法，训练 AI 模型不过是流程化的操作[5]，小学生也能用人工智能来解决问题。

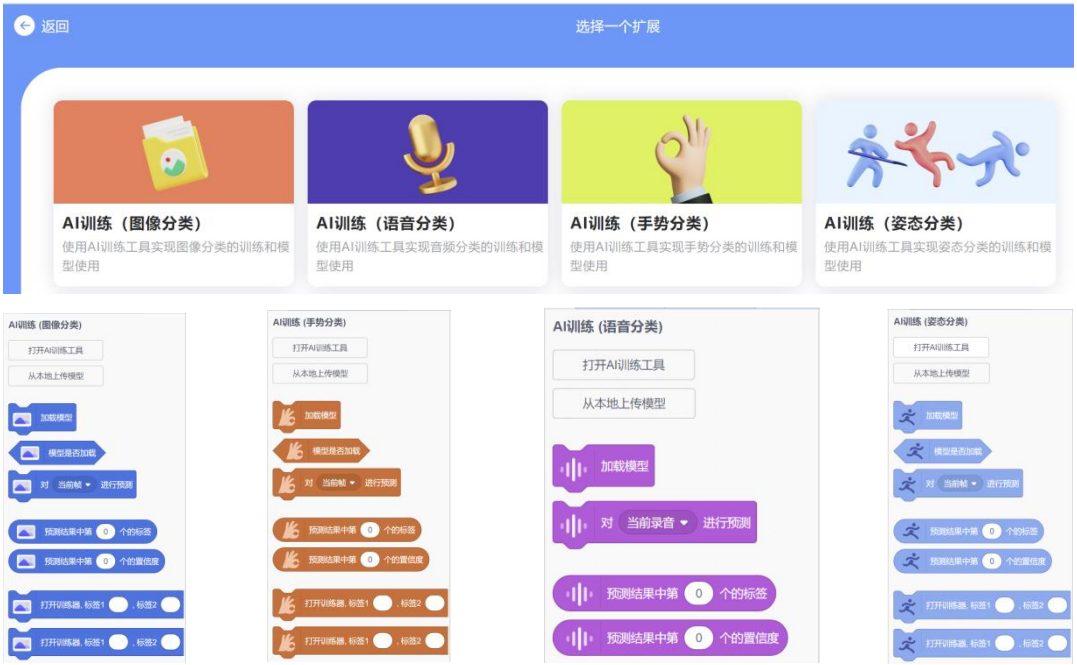


图 8 AI 分类训练扩展模块

5. 见探索之效

平阳县昆阳镇第六小学的 4 名学生参加了 2023 年全国青少年人工智能创新实践活动，荣获“优秀作品”荣誉。其中，两位学生的作品《三月歌飞》入围了在上海举行的交流展示会（图 9），并在现场任务挑战活动中表现出色，荣获了 AI 交互设计“四星卓越”创新作品奖。他们

的指导老师也被评为“优秀指导师”，学校则被评为“全国青少年人工智能创新教育特色学校”。





图 9 入围交流展示会

作为农村区域试点小学的平阳县昆阳镇第六小学，学校的老师与学生积极投入实践，不断探索着通往新一代人工智能教育的创新之路，通过开展教学活动和参与竞赛项目，共同提升知识水平，拓宽认知视野。这种全面的教育方式不仅培养了学生们的创新思维和实践能力，也为学校的教学模式注入了新的活力和动力。

6. 结语

在新一代人工智能的国家战略和温州市推动人工智能教育的大趋势下，作为地域偏远、设备紧缺的农村小学，我们需要使用成本低廉的摄像头和耳麦等设备，以及方便易用、开放性高的浦育平台的“创意积木工坊”模块，创作出体现深度学习技术的新一代人工智能创新作品。因此，农村小学开展新一代人工智能教育并非是遥不可及的梦想，而是近在眼前的未来，让学生能够更好地适应人工智能时代的社会发展需求。

参考文献

国务院.国务院关于印发新一代人工智能发展规划的通知[DB/OL].[2017-07-20].https://www.[1]gov.cn/zhengce/content/2017-07/20/content_5211996.html

温州市教育局.市教育局发布《温州市中小学推进人工智能教育实施方案》[DB/OL].[2022-[1]03-09].https://www.wenzhou.gov.cn/art/2022/3/9/art_1217833_59154015.html

梁森山,王阳,刘军,等.寻味人工智能教育:祛魅之下的实践与思考[J].**中国信息技术教育**,2018[1](02):4-11.

李永智.新一代人工智能技术对教育的影响[J].**教育参考**,2023(04):2-3.

谢作如,方建文.让学生真正理解数据、算法和算力——谈新一代人工智能教育教什么和怎么教[J].**中小学信息技术教育**,2023(06):8-10.

人工智能领航山区教育创新

——以校本课程开发为例

Artificial Intelligence Leading Educational Innovation in Mountainous Areas: A Case Study of School-Based Curriculum Development

张艳超[1*]，徐昊宇[1]，赵杰[1]
[1]丽水学院
*zychero1979@163.com

【摘要】 我国《义务教育课程方案（2022 年版）》规定地方课程要充分利用地方特色教育资源，校本课程要发挥特色教育教学资源优势。人工智能为山区教育带来了前所未有的发展机遇，山区学校可以突破传统师资条件的限制，借助人工智能语言大模型，生成山区学校校本课程选题、目标、内容、实施、评价等五个开发环节的学习支架，将山区各类资源整合融入课程建设，成为人工智能时代专业高效的创课者。“人工智能+”山区学校校本课程开发，有助于提升山区学校教育的质量和水平，有效促进学生个性化发展。

【关键词】 山区学校；校本课程；人工智能；大语言模型；开发

Abstract： *The "Compulsory Education Curriculum Program (2022 Edition)" in China stipulates that local curricula should fully utilize local characteristic educational resources, and school-based curricula should leverage the advantages of distinctive educational and teaching resources. Artificial Intelligence (AI) has brought unprecedented development opportunities to education in mountainous areas. Schools in these regions can break through the limitations of traditional teacher qualifications by utilizing large AI language models to generate learning scaffolding for five key stages of school-based curriculum development: topic selection, objectives, content, implementation, and evaluation. This integration allows for the incorporation of various local resources into curriculum construction, transforming schools into professional and efficient curriculum creators in the age of AI. The combination of AI and the development of school-based curricula in mountainous area schools helps to enhance the quality and level of education, effectively promoting the personalized development of students.*

Keywords： *Mountainous area schools; School-based curriculum; Artificial intelligence; Large language models; Development*

1. 引言

1.1. 背景介绍

发展山区教育对于实现教育高质量发展和教育公平具有极其重要的意义。在 2024 年全国教育工作会议中，教育服务高质量发展的重要性被再次强调，提出了构建与中国式现代化相匹配的高质量教育体系的目标。在这一体系中，山区教育占据着重要地位，其改善和提升直接关系到教育公平的实现和教育质量的整体提升。在中华人民共和国第十四届全国人民代表大会第二次会议记者会中，教育部部长怀进鹏指出：“把人工智能技术深入到教育教学和管理全过程、全环节，来研究它的有效性、适应性，未来我们将致力于培养一大批具备数字素养的教师，让学生青年一代更加主动地学，让教师更加创造性地教。”当前，人工智能语言大模型技术飞速发展，从最初的模糊概念到如今的广泛应用，其强大的数据处理能力和创新性的内容生成方式正在改变众多行业的面貌。

在山区教育领域，由于地理位置偏远、教学技术落后以及信息流通不畅等因素，山区学校面临着师资力量不足、教学质量不高、学生发展受限等多重困境，但山区学校也拥有其独特的课程资源优势。例如丰富的自然资源、独特的文化背景以及理想的环境教育实践平台，都为山区学校提供了宝贵的课程资源。如果能够充分发挥这些优势，结合人工智能语言大模型技术，开发具有地方特色的校本课程，将极大地丰富山区学校的教育内容，提升其教育质量，促进学生全面发展。但是山区学校的老师们对人工智能技术发展了解偏少，往往不知道如何开发合适的校本课程，难以形成具有地方特色和教育价值的课程体系，这成为了制约山区教育快速发展的重要因素。我们需要关注到山区教育的独特性和困境，提升一线教师数字素养，

推动更多的山区学校利用人工智能语言大模型技术辅助校本课程开发。因此深入研究和探讨如何利用人工智能语言大模型技术推动山区学校校本课程的开发显得尤为重要。

1.2. 研究意义

本文为山区学校校本课程开发提供了新的方法、模式和手段，有助于提升山区学校课程开发的质量、效率和水平。通过人工智能语言大模型的运用，赋能教师数字素养提升，充分发挥山区特色，培育学生山区情怀，促进教育资源的均衡分配，缩小城乡教育差距，减轻教师的工作压力及非教学任务的课业负担。

1.3. 研究目的与问题

本文旨在探索“人工智能语言大模型”如何领航山区教育，特别是对校本课程开发模式的创新。深入探索其在山区校本课程开发中的支架作用，提出有效的课程开发策略。通过对语言大模型技术的整合与利用，我们期望为山区学校提供一套具有针对性的校本课程开发的模式以及校本课程开发的案例模板，以满足山区教师开发适应地区、学校、学生发展的校本课程，促进学生的全面发展。

2. 国内外相关文献综述

校本课程这一概念最早出现在 20 世纪 70 年代的英国和美国等发达国家，以改变长周期课程开发模式，使课程更适应社会和经济发展的需求。校本课程开发涉及学校自主规划学生学习内容及其实施，包括课程设计、实施与考核。萨巴尔进一步将校本课程开发定义为狭义上的单课程设计与实施，以及广义上的学校整体课程开发过程。托马斯提出校本课程开发的四步骤方案，斯基尔贝克则在此基础上增加目标分析，形成五阶段模型。上世纪 90 年代初，我国内地开始逐渐引入“校本课程开发”的概念与理论，也开始进行探索和发展校本课程。2001 年新课程改革不仅推动了我国课程管理体制的改革，还确立了国家、地方、学校的三级课程管理制度，并且明确提出学校应该“结合本校的传统和优势、学生的兴趣和需要，开发或选用适合本校的课程”。由此，“校本课程开发”不仅成为课程改革的一项重要实践活动，也日渐成为课程专家和一线教师普遍关注的研究课题。而校本课程因其灵活性、个性化教育、教师专业发展、学校特色和教育改革的需要，持续受到教育界的重视。2022 年教育部出台的新课程标准（2022 年版）再次将校本课程作为本次改革的重点与国家课程和地方课程相结合，反映了各校的办学宗旨、学生的特别需要和学校的资源优势。国内学者鲁艳（1999）认为校本课程实质上是一个以学校为基地进行课程开发的开放民主的决策过程，即校长、教师、课程专家、学生以及家长和社区人士共同参与学校课程计划的制定、实施和评价活动。郑金洲教授《走向校本》对此作了如下阐释：所谓校本，一是为了学校，二是在学校中，三是基于学校。宋林飞认为，国家课程校本化实施建立在尊重国家课程意志基础上，是连接生活世界与书本世界，依循课程形态转化规律与学习规律，密切关注校情、学情、班情，调适与创生课程，让学生得到适宜发展的学习活动。从育人价值统整、内容结构统整、课程实施统整三个方面进行国家课程的校本化实施。李臣之关注开发主体和方式，李介认为校本课程开发有助于提升学生实践能力，需因材施教、因地制宜。陈伶俐指出特色课程对学校品牌形象的重要性。冯恩旭指出“将本土文化引进学校主流教育，做到国家课程校本化，校本课程特色化，在课程开发中发展学生核心素养实践上”部分地区学校已有成果。崔允漷强调挖掘地方文化资源开发校本课程，弥补传统课堂的局限。吴刚平在《校本课程开发 20 年》一文中论述了校本课程发展存在区域不平衡等问题，仍需继续深化研究和改进探索。陈迪喜,王江丽提到校本课程编写过程包括开发、评估、实践。孙渊,赵媛媛认为转变校本课程开发的研究视角，从学校和教师转向学生，关注学生发展核心素养的培养以及学科核心素养的落实，无疑会是未来“校本课程开发”研究的一个重要方向。

大语言模型(Large Language Model，LLM）技术已经在教育领域探索应用。许多研究者和实践者通过语言大模型进行教育资源的整合与优化，为学生提供更加个性化、精准的学习支持（Smith et al., 2023）。同时，国外学者也在不断探索语言大模型在课程设计、教学评价等方面的应用。通过BERT模型，教育评估者可以更加客观、准确地评价学生的语言能力和思维水平，为教育决策提供有力支持。特别是 OpenAI 公司在 2022 年 11 月发布 ChatGPT之后,更是引起了全世界的广泛关注。GPT-4 模型拥有强大的文本生成能力，可以辅助教师进行教案编写、课程设计等工作，GPT-4 还能根据学生的作业和考试答案进行自动批改和评分，大大减轻了教师的工作负担。同时，GPT-4 还可以为学生提供个性化的学习建议，帮助他们更好地掌握知识。从 ChatGPT 到 Sora，人工智能技术多模态发展日新月异,不断刷新着人类对生成式人工智能的认识。国内文心一言、讯飞星火、通义千问等人工智能语言大模型的快速发展，为我国教育数字化、智能化的快速发展提供了有利的技术支撑。大模型在落地应用过程中也面临着技术和社会伦理层面的问题，需要各国在生成式人工智能的管理方面进行规范与治理。目前，OpenAI、微软、Duolingo与冰岛政府等均有合作，迅速建立并不断拓展产业应用生态，以推动教育的创新与发展。越来越多的国内外学者认为，语言大模型作为人工智能的底层逻辑架构，在技术上降低了课程开发的门槛和难度，减少了教师的开发工作量，提升了校本课程开发质量和效率，有利于各地因地制宜的进行适合学生发展和地方特色的校本课程开发。

3. 基于人工智能的校本课程开发模式

3.1. 基于语言大模型的校本课程开发模式

随着云计算和网络技术的发展，即使是地处偏远的山区学校也能够通过互联网学习和使用人工智能语言大模型，让山区学校教师通过技术赋能降低校本课程开发的难度。其强大的计算和数据处理能力可以为教育资源匮乏的山区学校提供个性化和高质量的教学内容。其处理自然语言方面的能力，使其可以根据山区的具体文化和语言环境，开发出更加贴近学生生活实际的课程内容，增强学生的学习兴趣和参与度。随着大模型的不断学习和进步，它能够适应教育需求的变化，为学生提供最新的知识和信息。这也将极大地推动教育公平，帮助山区学生打开知识的大门，迈向更加广阔的未来。基于斯基尔贝克提出的校本课程开发五阶段模型，本文设计的面向山区学校校本课程开发的流程，如图 1 所示。主要包括五个环节：一是校本课程选题开发；二是校本课程目标开发；三是校本课程内容开发；四是校本课程实施开发；五是校本课程评价开发。五个环节相互关联、相互促进，共同推动着校本课程的迭代开发与实践。课程开发团队与人工智能语言大模型持续互动交流，共同完成课程建设。

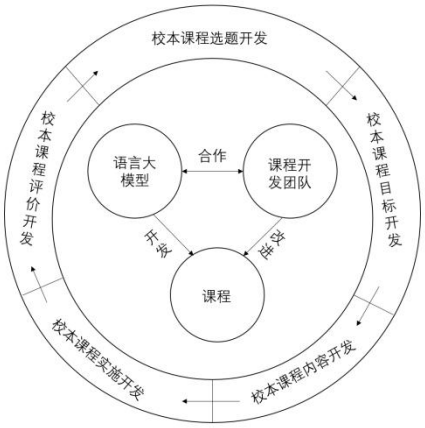


图 1 基于语言大模型（LLM）的校本课程开发模式

3.1.1. 基于 LLM 的课程选题开发

课程开发团队需要明确课程开发的选题方向，选题应当遵循以下 5 个原则：第一是政策性原则，以确保课程内容符合国家教育方针和政策导向；第二要体现校本性原则，紧密结合山区学校的实际情况和特色；第三要考虑适应性原则，确保课程内容适合山区学生的年龄特点和认知水平；第四是特色性原则，使课程内容能够凸显山区学校的独特性和创新性；第五是可行性原则，则强调课程开发的可操作性和实效性。借助人工智能语言大模型，我们可以根据这五个原则开发校本课程选题评价量表，根据山区学校的实际情境提供多个校本课程开发选题，并对选题进行综合评价，帮助教师选择和完善适合山区学校校本课程建设的选题。

3.1.2. 基于 LLM 的课程目标开发

确定了选题方向后，我们进入课程目标的制定阶段。课程目标应当明确、具体，能够指导后续的课程内容选择和教学方法设计。同时，课程目标应与学生的实际需求和发展方向相契合，确保课程的价值和意义。借助人工智能语言大模型，可以根据课程选题，自动生成课程三维目标，即知识与技能、过程与方法、情感态度与价值观。

3.1.3. 基于 LLM 的课程内容开发

在课程内容的选择上，我们要充分考虑山区学校的实际情况和资源优势，挖掘具有地方特色的教育资源，丰富课程内容。在内容的组织上，要注重知识的逻辑性和系统性，确保学生能够循序渐进地掌握知识和技能。我们也需要设计合适的教学方法与手段，教学方法的选择应根据课程内容和学生的实际情况进行灵活调整，注重激发学生的学习兴趣和主动性。在文本内容方面，我们可以借助文心一言、讯飞星火等人工智能语言大模型生成所需要的课程资料；在音频内容方面，可以借助 Suno、SongR、SkyMusic 实现对中英文音乐的创作；在图片内容方面，文心一言、讯飞星火等语言大模型的插件功能可一键生成；在视频内容方面，Sora 和剪映均可实现文本、图片等信息一键生成视频，而且讯飞星火等语言大模型还可以实现绘本教育资源的制作。因此，借助这些生成式人工智能和语言大模型，如表 1 所示，可以快速生成校本课程开发所需的各种多媒体资源。

表 1 可以用于校本课程内容开发的“智能”工具

文本生成	音频生成	图片生成	视频生成	绘本生成
ChatGPT Copilot 文心一言、讯飞星火、通义千问、盘古、Kimi 等	Jukebox AudioLM WaveGrad MelGAN Vocoder 系列 Suno SkyMusic	DALL-E 2 Stable Diffusion DeepDream GAN 系列（如 ProGAN、StyleGAN 等） BigSleep 讯飞星火等	Gen-V VideoGPT MoCoGAN TGAN Sora 新壹视频大模型 剪映	讯飞星火、imagestory 等

3.1.4. 基于 LLM 的课程实施开发

接下来是课程实施阶段，我们要关注教学过程的监控与管理。包括制定详细的教学计划，确保教学进度和质量；加强师生之间的沟通与互动，及时解决教学过程中出现的问题；以及开展丰富多样的教学活动，促进学生的全面发展。借助人工智能语言大模型，我们可以开发

校本课程实施方案量规表，找出课程实施的不足和存在的问题，帮助教师更好地将事前规划的课程付诸实际教育行动。

3.1.5. 基于LLM的课程评价开发

最后我们要对课程开发进行评价与反思。通过收集和分析学生的学习成果、教师的教学反馈以及社会各界的意见建议，借助人工智能语言大模型生成测评试卷，对课程开发的效果进行全面评价。同时，我们还要总结课程开发过程中的经验和教训，为今后的课程开发工作提供借鉴和参考。

人工智能语言大模型分为通用和专用两种形式。面对信息技术基础参差不齐的教师群体时，本研究提出了一种双轨并行的课程开发流程。对于信息技术运用能力较薄弱的教师，基于通用人工智能语言大模型，降低校本课程开发的技术门槛并提升课程的规范性，以便更加有效地协助课程开发团队构建适合山区学校的校本课程；而对于那些信息技术运用能力较为突出，且已经有一定通用语言大模型操作经验的教师，通过运用专用语言大模型进行课程开发，进一步提升校本课程开发的专业性、科学性和适用性，以期达到更高水平的山区学校特色校本课程开发。

3.2. 基于通用语言大模型的校本课程开发流程

基于语言大模型(LLM)的校本课程开发模式流程图(如图 1 所示)，整个开发起始于“校本课程选题开发”，选题过程需紧密围绕学校的办学特色、教育理念以及学生的实际需求，确保课程内容既具有前瞻性适应未来，又贴近学生生活适应当下。选题的开发过程，是利用人工智能语言大模型进行挑选合适选题的人机合作过程。随后，进入“校本课程目标开发”阶段，目标的设定，是一个基于选题内容和课程理念，进行合理组合的内容，也是语言大模型能辅助课程开发团队，短时间内较为专业的进行校本课程理论与实际结合，为课程内容的选择与组织提供了依据。课程目标开发过程中，需要教师与大模型多次交流迭代，充分考虑学生的认知特点、兴趣需求以及社会发展的需要，确保课程目标的科学性与前瞻性。

此外，“校本课程内容开发”和“校本课程实施开发”两个分支体现了语言大模型、课程开发团队对当地相对应的课程开发资源的整理利用。内容开发充分的利用了语言大模型的知识库优势，能够形成系统丰富的课程素材和各种形式的课程内容。同时大语言模型对实施阶段的开发也比较全面，可以结合教师能力，学生实际，课程理论，文化传统，政策导向等多维度情况进行充分的思考，给出合理的课程实施方案。这两个分支相互交织、相互促进，共同推动着校本课程的深入开发与实施。校本课程内容开发注重课程内容的系统性与连贯性，强调知识的深度与广度以及与实际生活的联系；校本课程实施开发关注课程的实际运作过程，包括教学方法的选择、教学手段的运用以及课程管理的优化等。

在“校本课程评价开发”这一阶段中，主要是对课程实施效果的反馈与调整，通过对教学过程的观察、学生反馈的收集以及课程成果的评估，不断优化课程内容与教学方法。评价开发强调过程性与结果性的结合，既关注学生在课程学习中的过程体验，也重视课程实施后的实际成效。课程整个开发过程，是一个不断循环、不断完善的过程。通过选题、目标设定、内容与实施、评价开发等环节的相互衔接与配合，推动着校本课程开发的不断深入与完善。

3.3. 基于专用语言大模型的校本课程开发流程

专用大模型的使用是校本课程开发流程中的进阶环节，相较于通用大模型的使用，专用大模型可以实现一定行业范围内的更为专业的问题回答，以更加专业性、学术性、适切性的生成内容满足教师的课程开发需要。教师可以利用平时积累的工作场所知识建立专用知识库，让语言大模型在专用知识库范围内精准搜索，并利用专用大模型机制，优先搜索知识库中涉及的知识，知识库中未涉及的知识内容则通过通用大模型回答，分层的知识库系统架构如图

2 所示，有效解决了通用大模型“广而不专”、专用大模型“专而不广”的问题。

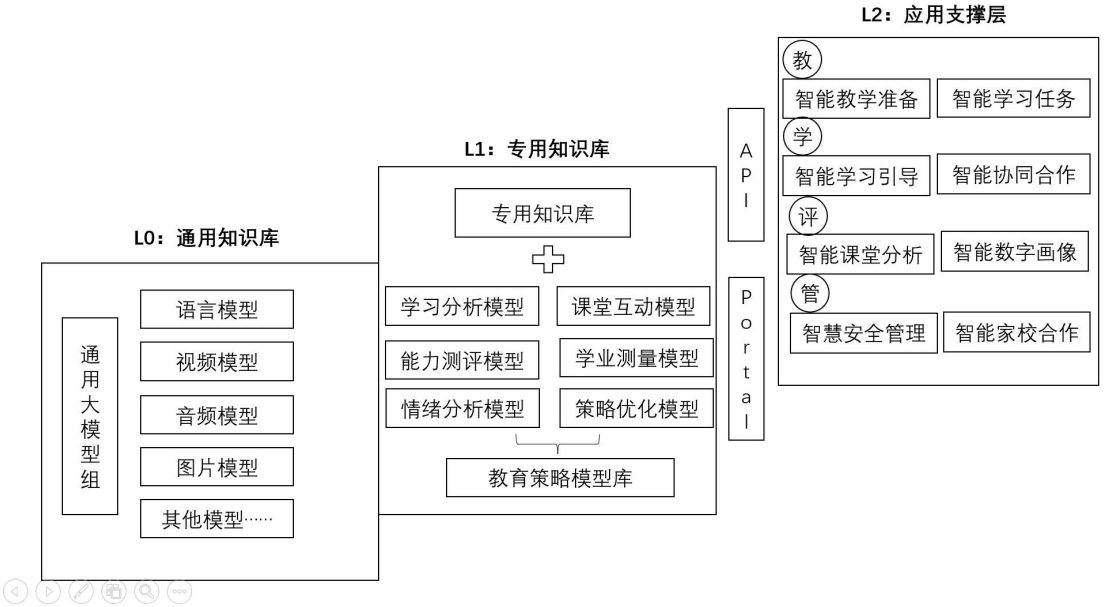


图 2 分层的知识库系统架构

分层的知识库系统架构参考了之江实验室大模型分类，旨在通过整合不同层次的信息来提高数据处理和应用服务的质量。该系统由三个主要层次组成：通用知识库（L0 层）、专用知识库（L1 层）和应用支撑层（L2 层）。[20]L0 层：通用知识库 这一层包含了广泛的基础信息和数据，为系统提供了必要的语言模型、常规任务型对话和问答型对话的基础设施。此外，它还支持视频、音频、图片等多模态内容的生成。L1 层：专用知识库 在这一层，通用知识库的信息被进一步整合，并与特定领域的专业信息相结合。这一过程涉及到学习分析、调查与分析等多个模块，以形成适用于高质量对话任务、问答系统和教育辅导解决方案的专有知识。L2 层：应用支撑层 应用接口层利用 API 将专门知识进行整合和处理，以适应不同的应用场景，如智能客服和智能助手。API 与 Portal 则作为连接应用接口层和最终用户的桥梁，实现信息和服务的有效传递。该系统架构展示了从通用知识到专用应用的逐层深入，每一层都在为下一层提供支持和资源，共同构建起一个高效、可靠的知识和服务提供平台。这种架构不仅能够提升信息处理的质量，还能够为最终用户提供更加精准和个性化的服务。

L0 层通用知识库的语言大模型，如文心一言，提供了一个广泛的信息和数据集合，这些信息和数据是不特定于任何单一领域的。它适用于大多数教师，特别是那些没有深入掌握语言大模型基础知识的教师，因为它提供了一个基本的、易于接入的知识平台，可以支持日常的教学和研究活动，而不需要用户具备深厚的技术背景。相比之下，L1 层的专用知识库，以华云 AI 数字人为例，是为了满足特定领域需求而设计的。它不仅整合了通用知识库的信息，还包括了特定领域的深度知识和数据。这就需要专用语言大模型建构与训练，对校本课程理论、新课程标准、优秀校本课程开发案例资源等进行学习，并结合通用模型中各校所处的社会经济、文化、地缘等具体背景，综合考量学生需求和教师能力的基础上，人机互动进行智能化的山区校本课程开发的框架、思路、理论、内容等指导，帮助教师形成适合学生年龄特征、地域特色、核心素养培育的山区学校校本课程选题开发、目标开发、内容开发、实施开发、评价开发的全过程，并给予多套开发方案供教师选择。因此，L1 层更适合那些已经具备一定语言大模型基础知识的教师使用，他们可以利用这些专业知识来进行更高级的教学和研究工作，例如开发专门的课程内容或进行深入的学术研究。总的来说，L0 层的用户群体主要

是面向对人工智能不太熟悉的教师，而 L1 层的用户群体则是面向需要更专业、更定制化知识服务的教师。这种架构设计保障了不同层次的教师能够根据自身的需求和技术能力，开发合适教学需要的资源和工具。

4. 基于人工智能的山区学校校本课程开发示例

4.1. 校本课程开发各环节量规表设计

山区学校校本课程的开发主要包括选题、目标、内容、实施、评价五个环节。限于篇幅，本文将对利用语言大模型进行选题开发的环节进行详细介绍。在选题开发过程中，首先，我们将从“语言大模型”这个起点开始，这是整个流程的基石。登录语言大模型并建立新的对话，这是开始学习的第一步。接下来，我们需要收集初始信息，这些信息将为后续的学习提供重要的参考。之后，他们将再次建立新的对话将收集到信息进行发送，这进一步推进了学习的进程。在学习过程中，我们需要输入关于学校的初步信息，并讨论地域特点和学生特点，这些信息有助于人工智能语言大模型更好地理解该校的背景和环境。随后，采用合适的选题开发的原则，令语言大模型进行学习1。同时，基于这些原则制定量规表，如表 2 所示。并根据这些原则和学校的实际情况来设计初始量表，这为课程选题提供了一个明确的评分框架。

表 2 校本课程选题评价量规表

原则	内容	赋分
政策性原则	选题紧密围绕国家教育政策，体现教育发展方向。	10 分
	选题符合地方教育政策要求，有助于推动地方教育发展。	10 分
校本性原则	选题与学校办学特色相结合，能够凸显学校的教育理念和特色。	10 分
	选题能够充分利用学校的教学资源和师资力量，实现资源共享和优势互补。	10 分
适应性原则	选题适应学生的年龄特点和认知水平，能够激发学生的学习兴趣和积极性。	10 分
	选题适应社会发展需求，能够培养学生的综合素质和创新能力。	10 分
特色性原则	选题具有创新性，能够体现课程教学的独特性和新颖性。	10 分
	选题在教学方法、教学内容等方面具有明显特色，有助于提升课程质量。	10 分
可行性原则	选题具有明确的实施计划和步骤，能够确保课程教学的顺利进行。	10 分
	选题所需的教学资源、教学条件等能够得到保障，具有较高可行性。	10 分
总分		100 分

4.2. 基于通用语言大模型的校本课程开发

在量表设计完成后，课程开发团队将对量表进行头脑风暴，这有助于发现量表中的问题并进行优化。进入修订阶段，教师们将修改后的量表让人工智能大模型进行润色学习，提升学习效果。最后，在生成阶段，基于量表生成多个选题方案，并对已有的选题进行评分反馈。

这将帮助教师们筛选出最佳的学习方案，为他们的决策提供有力的支持。通过这个系统化的流程，教师们将能够在一个有指导性的环境中学习，提升他们的学习效果和决策能力。具体过程如图 3 所示：

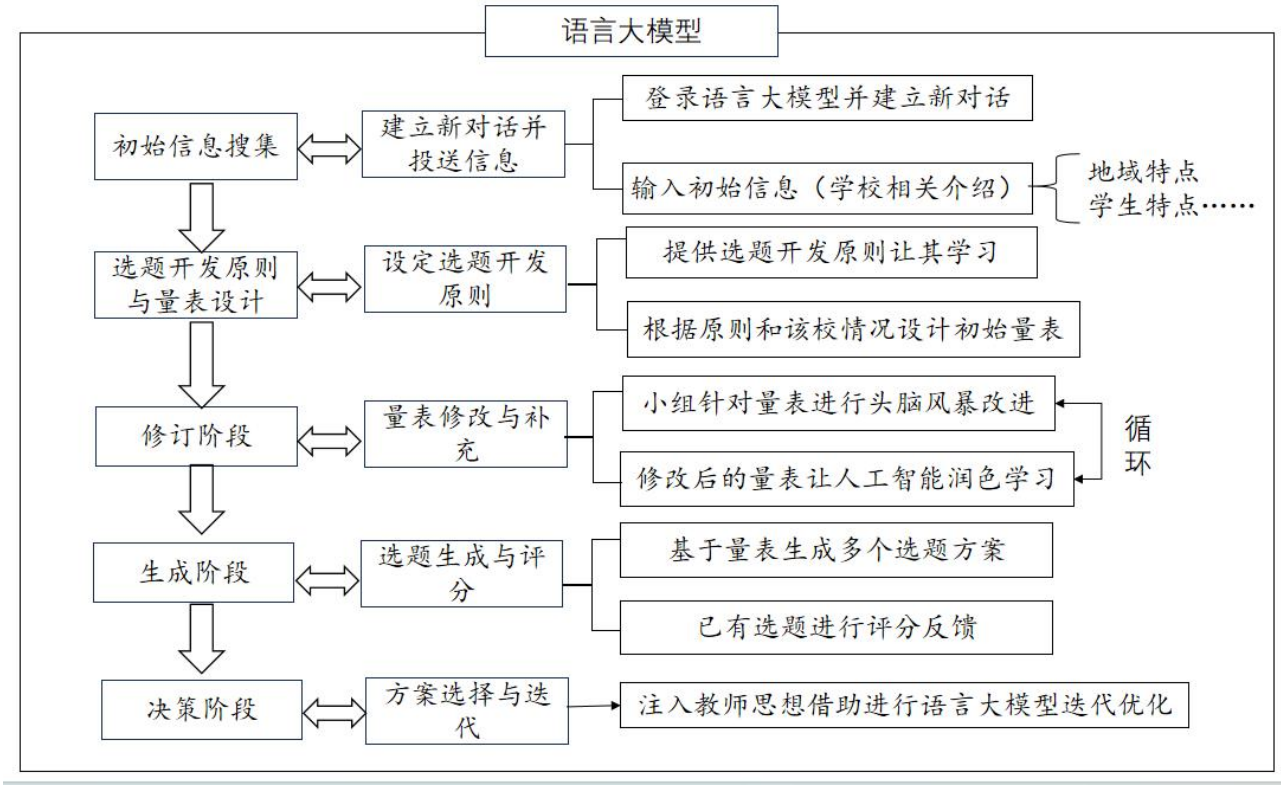


图 3 语言大模型促进校本课选题开发流程图

4.3. 基于专用语言大模型的校本课程迭代开发

如果教师想利用专用大模型进行更高水平的课程开发，则需在通用大模型基础上对所用领域知识进行汇总，并建立私有的相关课程资源知识库，便于专用大模型调用知识库相关知识，进行精准辅助课程开发。以华为公司开发的“华为 AI 数字人”大模型为例，具体如下图 4 所示。

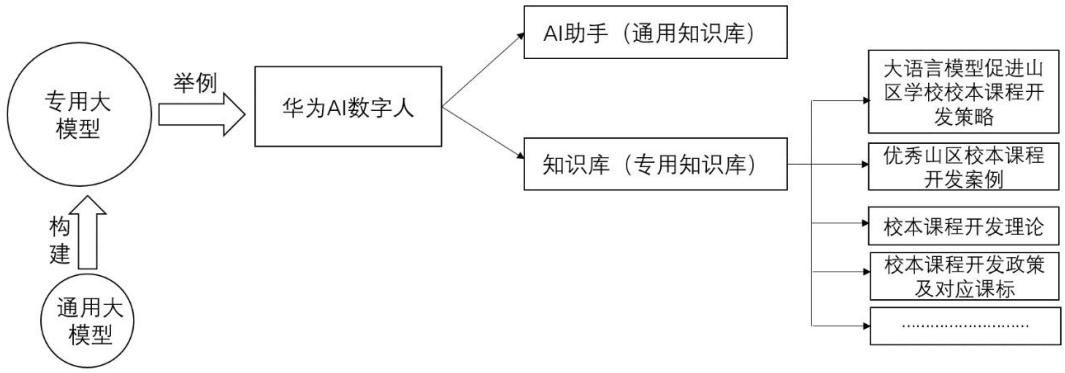


图 4 专用大模型促进校本课选题开发流程图

本图描绘了从广泛的知识集合到定制化模型的转换过程，其中 AI 助手（通用知识库）和华为 AI 数字人进行了融合式应用。在此过程的起始阶段，即“通用大模型”，对广泛的知识和

资源进行集合，并为后续的转化提供了基础。AI 助手（通用知识库）介入，作为转化过程中的关键媒介，通过筛选和定制化，将通用知识转化为满足特定需求的专业知识。在“专用大模型”阶段，AI 助手根据特定的需求或情景，从专用知识库中提取出相关领域的专业知识，或者提供针对性的解决方案。这一专业化和定制化的过程确保了资源的精准性和适用性。华为 AI 数字人具有重要作用，它与用户互动，提供必要的指导和建议，使用户能够更直观、更便捷地利用 AI 助手和专用知识库。在整个过程中，制定有效策略的重要性得到了强调，包括校本课程开发策略、知识库使用策略等，以确保资源的有效整合和利用，课程开发顺利进行。

5. 结论与建议

本文探讨了如何通过整合人工智能技术，助力山区学校充分利用自身资源和文化特色，开发出具有地方特色的校本课程。这不仅丰富了学生的学习体验，也深化了他们对山区文化的理解和热爱，培养了他们的山区情怀。同时，语言大模型的应用还极大地减轻了教师的负担，提高了教学效率，使得教育资源得到更为合理的分配和利用。当然，我们也应清醒地认识到，人工智能技术在山区教育的应用仍面临着诸多挑战：如技术更新换代的快速性，要求山区教育工作者不断学习和适应新技术；如何确保技术在教育中的合理应用，避免过度依赖技术而忽视教育的本质，这也是需要深入思考的问题。我们建议一方面应继续加大对山区教育信息化的投入，为山区学校提供更多的技术支持和资源保障；另一方面也应加强对山区教育工作者的培训，提升他们的信息素养和技术应用能力，使他们能够更好地利用人工智能技术推动山区教育的创新和发展。为此，我们先后开发了《人工智能与未来教育》《“人工智能+”山区学校校本课程开发》等数字化课程，致力于持续提升高校师范生和山区教师的数字素养和校本课程开发能力，使之成为人工智能时代专业高效的创课者。

参考文献

李欢欢.基于核心素养的高中生物校本课程开发研究——以大连市为例[D].辽宁师范大学, 2018:3.

Tomas I D. A decision framework for school-based decision making [J].The Australian Science Teachers Journal,1978,24(2): 63-68.

Skilbeck M. School-based curriculum development [M].London: Harper &Row, 1984:127.

教育部．基础教育课程改革纲要(试行)[J]．人民教育，2001(1) : 17 - 19．

鲁艳．校本课程：概念必须正确理解[J]．教育发展研究，1999，(12)：19-23

杨妍. 中职语文课程校本化研究[D].南京师范大学, 2017.

宋林飞.国家课程校本化实施的系统理解与整合行动[J].上海教育科研,2019(04):63-67.

王远彬,曾曼丽.课程统整:国家基础课程校本化实施新路径[J].中国教育学刊,2020,No.322(02):105.

李臣之.校本课程开发的三个基本问题[J].课程·教材·教法,2012,32(5):8-14.

李介.国外校本课程开发模式带给我们的启示[J].教育理论与实践,2010(9): 18-20.

陈伶俐.基于核心素养的初中科学校本课程开发研究[D].上海师范大学.2018:5.

尹少淳.核心素养大家谈[M].湖南:湖南美术出版社,2021.296

吴刚平,陈华,徐晨盈,赵晓雨.校本课程开发 20 年[J].全球教育展望,2021,50(12):3-18

陈迪喜,王江丽. 校本课程开发背景中的教师专业成长[J].新校园, 2022(11):58-60.

孙渊,赵媛媛.我国“校本课程开发”研究的热点与趋势——基于 CNKI(1999 ~ 2019)的知识图谱分析[J].当代教育理论与实践,2021,13(01):48-53.DOI:10.13582/j.cnki.1674-5884.2021.01.009.

OpenAI.GPT-4 Technical Report [EB/OL].https://arxiv.org/ abs/2303.08774v1,2023-12-01.

冯志伟.从 ChatGPT 到 Sora 发展中的术语问题[J].中国科技术语,2024,26(02):88-92.

Li Y . Specifics of Regulatory and Legal Regulation of Generative Artificial Intelligence in the UK, USA, EU and China [J]. Pravo. Zhurnal Vysshei shkoly ekonomiki, 2023, (3): 245-267.

郭全中,朱燕.广布局、重应用：生成式大语言模型的新进展[J].新闻爱好者,2023(08):21-25.DOI:10.16017/j.cnki.xwahz.2023.08.001.

中国教育科学研究院数字教育研究所、之江实验室智能教育研究中心.重构教育图景：教育专用大模型研究报告[R].2023.12.

AIGC 在幼儿家庭教育中的应用与思考

——基于微信公众号“扑满的杂记团”的案例分析

Application and Reflection of AIGC in Early Childhood Family Education — A Case Study Based on the WeChat Public Account "Puman's Chronicles"

周纲[1]，黄今来[2]，黄鲁凡[3]
1 温州市政协
2 温州市机关一幼
3 温州市瓯海区人民法院
79814885@qq.com

【摘要】 本研究以“扑满的杂记团”微信公众号为案例，运用参与式观察、自传式研究和反思性研究方法，结合问卷调查结果，探讨了 AIGC 技术在幼儿家庭教育中的应用，重点分析了 AIGC 技术在日记配图、宝宝学诗和科学实验等场景应用实例，展示了 AIGC 通过个性化内容生成、可视化概念理解、互动式学习体验，促进幼儿的认知和情感发展的作用。研究表明，AIGC 技术在激发幼儿学习乐趣、促进家庭教育亲子互动、加强家园共育等方面具有显著优势。文章最后探讨了幼儿学习应用 AIGC 中对技术依赖、隐私安全等风险隐患，提出了合理应用策略，并对未来发展趋势进行了思考展望。
【关键词】 人工智能；AIGC；幼儿家庭教育；教育创新

Abstract: This study uses the WeChat public account "Puman's Chronicles" as a case study and employs participant observation, autoethnography, and reflexive research methods, combined with the results of a questionnaire survey, to explore the application of AIGC technology in early childhood family education. It focuses on analyzing the application of AIGC technology in scenarios such as diary illustration, children's poetry learning, and scientific experiments, demonstrating the role of AIGC in promoting children's cognitive and emotional development through personalized content generation, visual concept understanding, and interactive learning experiences. The research indicates that AIGC technology has significant advantages in stimulating children's interest in learning, facilitating parent-child interaction in family education, and strengthening home-school co-education. The paper concludes by discussing the risks and hidden dangers of technological dependency and privacy security in the application of AIGC in children's learning, proposes strategies for rational application, and provides thoughts and prospects for future development trends.

Keywords: Artificial Intelligence;AIGC; Early Childhood Family Education;Educational Innovation

1. 前言

AIGC（Artificial Intelligence Generated Content）技术，即人工智能生成内容，是指利用人工智能算法自动创造文本、图像、音频和视频等内容的技术。作为一种教育创新工具，AIGC 也开始应用到幼儿家庭教育中，赋能幼儿学习与发展。

本研究聚焦的研究对象：一位幼儿园儿童“扑满”和他的微信公众号“扑满的杂记团”（简称“公众号”）。他从中班开始，与父母共同将口述日记结合 AIGC 技术，写成个人公众号，记录自己的成长、表达自己的思考。2023 年 3 月开通一年多来发表了 60 多篇故事，呈现了一个幼儿园儿童学习应用 AIGC 的生动案例。

本研究采用社会学“参与式观察”（Participant Observation）、“自传式研究”（Autoethnography）及“反思性研究”（Reflexive Research）”方法，结合问卷调查统计分析，将亲身参与的公众号亲子创作及 AIGC 应用过程体验，作为实证研究案例，进行个人生活故事叙述和反思，探讨 AIGC 技术如何融入幼儿家庭教育。

该研究的意义在于，有助于发现 AIGC 技术在幼儿教育中的潜力，有助于思考如何将先进技术与传统教育相结合，支持和促进幼儿的认知和情感发展，也能为 AIGC 在幼儿家园共育的未来教育技术研究，提供一定启发。

2. 案例分析

2.1. 缘起

2022 年，AI 生成画作《太空歌剧院》在美国人类艺术比赛中夺冠，引发激烈争议，也带来广泛关注，AIGC 各类应用软件也由此迅速“飞入寻常百姓家”。

本案例主角扑满，是个 6 岁男孩，两周岁半才开始学会说话，同时也开始进行口述日记。看着自己的话，变成本子上的字，能锻炼语言表达能力，也有助认字识字。扑满爱听故事，特别是家长自编的《牛虎大战》童话，里面化入了三十六计和成语故事。后来，扑满开始用自己的理解与想象，给爸爸妈妈讲述他的牛虎大战故事。他期待将来牛虎大战能攒出一本书。在口录手绘了十几集后，5 岁中班时，妈妈给他开了个公众号，“扑满的杂记团”，既连载牛虎大战，也发他讲的幼儿园故事和生活趣事。

公众号的开通，让口述日记从私密亲子交流，变成了共同面对“他者”世界，出现了两个问题：一个是如何让读者“悦读”，纯文字和手绘配图太过单调、简陋，要找到合适、连贯的配图很难，而且涉及版权；另一个是如何让自己信息安全，实拍手写日记和真人出镜，容易暴露过多个人隐私信息。一个偶然的机会，我们开始试着用 AIGC 来优化公众号，从牛虎大战第四集起，尝试 AI 文生图，再到图生图、图生视频，语音转文字，制作 AI 音乐等 AIGC 技术，扑满和爸爸妈妈一起学、一起做，AI 成了孩子有趣的玩伴、耐心的老师。

2023 年 3 月开通公众号一年多来，陆续发布了 60 多篇原创内容，90%以上用到 AIGC，虽然点击量不高，但作为学习工具，激发了孩子的兴趣和想象，也验证了幼儿的 AIGC 操作能力和家庭教育的应用潜力。

2.2. 应用

“杂记团”AI 应用工作流：聊天对话录音/录像——语音转文字（通义听悟等）——整理文稿（KIMI 等）——AI 文生图（无界 AI、腾讯智影等）——AI 图生视频（无界 AI、PixVerse 等）——AI 音乐（KIMI 生成歌词、SUNO 生成歌曲等）——串联成篇/片（135 编辑器、剪映等）——发布公众号图文或小视频。

上述工作流实际往往采用部分即可，无界 AI 文生图、KIMI 问答较为常用。在应用场景上，更多用于日记配图、宝宝学诗和科学实验。

2.2.1. 应用场景一：日记配图——AI 让故事更好看

最初，我们用微信小程序腾讯智影文生图，虽然生成的图片，还是动物形象，有的还是五条腿、两条尾巴的“老虎”，但总比自己手绘好看。扑满也很兴奋，用他的手机一起画。后

来发现无界 AI，可以更好生成可爱的动物卡通形象，穿着中国服装，更符合三十六计故事场景，还能形成风格相对统一、情节基本连贯的画面，如《牛虎大战第十四集·空城计》[1]。

图生图随之而来。在《一起画个空间站》里，幼儿园举办“咻——太空，我来啦”活动，让孩子和爸爸妈妈一起研究空间站的结构[2]。扑满画了一张空间站，再用 AI 图生图，打印出来带到幼儿园，老师很惊讶，“ai 还画得挺好！”同学也很惊讶“这是你自己画得吗，这和真正的空间站差不多呀！”

图生视频也是如此。在《有心弗怕路远，海燕归巢可待》，我们用 PixVerse 软件，让象征海外留学生回归的轮船动起来，让误解生气的妈妈，变成女神大力扭开可乐瓶的动作（形容要把爸爸头拧下来），有动图，更生动[3]。

AI 语音转文字功能，让我们能利用碎片时间，做口述日记。有时候是在旅途中，像《扑满的西游记》[4]，是一家人过年出行中的许多段录音，识别整理而来；有时候是在吃饭、哄睡中，像《扑满解诗·王安石<梅花>》[5]，就是在被窝里轻声聊天录音整理。“没时间坐下来手写，小家伙想到，我们可以把牛虎大战录音下来，回去整理啊！用微信语音输入，转化文字，是个好办法”（爸爸的朋友圈）

有了语音，有了画面，自然而然进入小视频。《童趣雪花》里，我们用剪影，对原声录音和 AI 图片，做成小视频[6]。为了认字，有时特意把图文公众号，让扑满读一遍录音上传，让不识字小朋友也能听能看，如图文版的《肚子疼》，做成《月光守护者》小视频；有时还用 AI 字幕识别，配上中英文字幕，顺带学下英语[7]。

AIGC 生成图像、音频和视频的玩法，让幼儿口述故事有声有色、妙趣横生，更让口述笔录的传统方式，变成录音整理的现代方法，成为口述日记的常态。

2.2.2. 应用场景二：宝宝学诗——AI 让古诗变好玩

幼儿背诗，最怕文字古奥、晦涩枯燥，我们更倾向读读亲切、有趣的儿童诗，如泰戈尔《新月集·12 点钟（扑满读诗）》，录音里扑满边读边，笑乐不可支[8]。

但 AI“陪读”后，扑满竟然喜欢上了古诗词，在《元日》背诵作业中，我们用 AI 配图、KIMI 解答，从题目“元日”，到每一句，特别是“屠苏”“曈曈”“新桃”“旧符”这些古汉语，找童言童语解释，找现代对应实物，还延伸了解作者生平和王安石变法[9]。老师看到了，留言“看完这篇，我的第一想法，是给小朋友看，从小朋友的角度给大家解释通俗易懂”。老师让扑满上台一边讲解，一边播放，同学们说，“原来王安石经历了这样一个故事，才写了元日啊！你这些内容都是哪里来的，我也想看看！”

和 AI 一起读古诗，让画面注解文字，补齐想象。扑满说，“如果用 AI 文生图的话，你有了图片，再去背这首诗，就会很好记了，如果你记不得下一句的诗，那你就想想自己画的图片，就更好记了。我知道他经历的故事以后，我觉得他写的诗也更好背了。”

还有一种更好玩的玩法，魔改“歪诗”[10]。在背诵杜牧《清明》中，“借问酒家何处有”，扑满想到现在我们问路找店，都用导航、美团，改出了一首“歪诗”《清明》：“清明时节雨纷纷，上坟路上人挤人。借问酒家何处有，美团一搜三里屯。”扑满觉得，“歪诗也不错啊，可以让古诗变得很有趣，不然，很多小朋友可能都不喜欢背古诗，包括我也是啊，歪诗就不一样了，哈哈哈哈！”

在《跟着节气读古诗·春夜喜雨》中，扑满用宝宝被狗咬的故事，解构了原诗，“好雨知时节，好狗不挡道。随风潜入夜，吓我一大跳。野径云俱黑，狗叫我也叫。晓看红湿处，昨晚被狗咬。”在幼儿园上台讲解正诗后，扑满也念了这首歪诗，回来告诉我们，“同学们说，这是什么诗啊？春夜喜雨，哈哈哈哈，再念一遍吧，求求你了，再念一遍吧！他们都很想我念，我上厕所他们都跟出来，再念一遍。”

在《青玉案·元夕（扑满解诗）》中，我们用了作者“辛弃疾”与“星期几”的谐音梗，还用爸爸元宵节带宝宝看花灯，把宝宝搞丢了，最后灯火阑珊处父子重逢的情景，作出了自己的解读。

对于歪诗的好处，扑满理解，“如果有歪诗，你就觉得读诗很好玩，如果世界上没有歪诗，只有正诗的话，大家可能觉得，哎呀，只有正诗，这么没意思，还是不背诗了吧。如果这样的话，那诗人都亏了。”

慢慢地，我们解锁了更多玩法。在学习《题惠崇春江晚景》[11]中，用题画诗所描述景象，尝试 AI 复活失落前年的这幅古画，还讨论了吃货苏东坡在诗中藏了一道宋朝菜谱“芦蒿炖河豚”。在贺知章《咏柳》KIMI 解读中，有参考链接提到，全诗没有一个“柳”字，却没有一个字不是写柳，“不着一字，尽得风流”，于是我们用 AI 画出身披绿丝绦的宋代美女图片，把题目《咏柳》留到最后念出，就变成了“诗里小姐姐是谁”的谜语[12]。

AI 赋能，“换个姿势背古诗”，让死记硬背变成深度阅读体验，更容易让幼儿爱上古诗。《牛虎大战》故事日渐淡出，读诗解诗成了公众号常客。

2.2.3. 应用场景三：科学探秘——AI 让研学更神奇

AI 赋能之前，爸爸妈妈协助扑满准备幼儿园科学小实验，是小红书搜方案、淘宝买道具、一起想台词，就像《科学实验——桔子的秘密》[1]里，做完实验，“很费桔子，很费爹”。

用上 AI 后，会直接给出一系列实验方案、原理解释和解说词。一次，老师要我们准备一个干冰实验，我们和扑满用 KIMI 提问，KIMI 马上给出答案和讲解词。追问下，又给出了五个更好方案。妈妈找了个干冰龙卷风的实验方案，和扑满一起做道具，扑满还把讲解词变成童话故事，再加上 ai 绘图，做成了一期《干冰侠大战狂风怪》[13]，还用 SUNO 作了一首曲子，作为主题曲和片尾曲。这个实验用于幼儿园一次考察接待，受到好评后，又在幼儿园“童享直播间”进行了网络直播。

校园之外，科学探索有着更丰富的话题与场景，我们尝试把平凡的生活过成一场有趣的探险。比如扑满对电梯有特别的痴迷，我们对着绘本《电梯的秘密》，到上海实地寻访中国最早的电梯（和平饭店）、世界最快的电梯（中心大厦）等等。最近，在 AI 帮助下，我们一起学习滑轮原理，尝试给幼儿园设计一个“不用电的电梯”，方便同学们搬运换洗被子。

AI 接触多了，不仅是探秘的引路老师，AI 本身也成为探秘对象。扑满对 AI 兴趣日浓，参观浙南云谷体验 AI 科技，参加心耕国际 AI 音乐工坊公益课程，旁听张文宇教授 AI 科普讲座[2]、温州大学 AIGC 俱乐部成立大会[3]、温州理工学院“中国（温州）人工智能产业峰会”、温州市实验中学 AI 中小学人工智能课程研讨等，体验了有香味的 AI 电视机，看到了各类 AI 工业应用，听到了 AI 改变未来的预言，从操作能力、未来图景、道德伦理等方面，都有了更深的感受。

扑满向 AI 提出各种问题，寻求 AI 的回答帮助，也对他和 AI 的关系有了自己的思考。在《第一次听大人讲座》中，他向主讲教授追问“AI 到底安不安全”的“人类幼崽”之问；在看到温大学生 AI 热情，他产生了“学会了以后，作业都不用写了，直接让 AI 写？”的困惑；在 AI 帮助写运动员代表发言后，他反问“我们让 AI 写我的运动员代表发言，它又不是运动员，这样子对吗？”为了更准确地问出提示词，他兴致勃勃地阅读《儿童艺术百科全书》《少儿趣味音乐史》等绘本，为了让练习成果用于公众号配乐、配诗，他也更用心地练钢琴、背古诗，AI 正在让他的学习更丰富、更有趣，也深刻地改变了他的学习方式。

除了以上应用场景，AI 也在进入扑满的日常生活。家里破壁机显示 E2 故障，没有说明书，扑满就用 AI 寻找答案，一起照着给出的视频链接解决了问题。几次经验后发现，AI 的生活问答，比网络搜索更便捷、准确，“万事不决问百度”转向了“问 AI”。

2.3. 结论

本案例研究证明，AIGC 在幼儿家庭教育中的应用，能在幼儿学习、亲子互动和家园共育发挥积极作用。近期我们以微信问卷星做了一起《幼儿园小朋友 AIGC 应用问卷》（简称《问卷》），62 名受访者（幼儿园家长占 40%）的回答，也较好地支持了这一结论。

2.3.1. AIGC 有助于幼儿学习激发乐趣

传统的教育技术多为书本教育、单向输出，受限于教材和师资，而 AIGC 技术可以突破这些限制，为幼儿提供丰富多样的学习体验，其核心优势在于能按幼儿学习需求和兴趣，生成个性化、可视化、互动式的学习内容。例如，通过 AIGC 生成的动画和视频，幼儿可以直观地理解学习中遇到的深奥、抽象的概念，如干冰的“升华”。这种视觉化、动态化学习方式，加深了幼儿的知识理解，尤其人机对话，先通过语言描述转化为 AI 图像，“看图理解”后，再修正、重绘，达到更深理解。这是一个深化理解的过程，也是一个激发想象的过程，更能将自己心中的故事，创作为电子绘本，以自媒体展示公众、收获反馈。日益成熟的语音输入法，足以让幼儿直接参与 AI 互动。在 AIGC 交互式学习中，幼儿不再是被动接受知识的容器，而是成为积极参与和创造知识的主体，更好地激发对传统学习内容的好奇心、探索欲和快乐感。

正如《问卷》回答里，74%认为幼儿教育中“需要或很需要”“在幼儿教育中使用 AIGC”，需要的原因里（多选），最集中的选项就是认为“AI 可以让小朋友学习更有趣，也更有兴趣”，占 63%，其次是认为“AIGC 可以帮助小朋友更好理解学习内容”，占 56%。AIGC 学习赋能的显著效用，得到了多数人的认可。

2.3.2. AIGC 有助于家庭教育亲子互动

本案例幼儿口述日记，从笔录拍照到 AI 赋能，让亲子创作的美好时光生活记录，更加有声有色，也减轻了家长前期“备课”和后期制作负担，能有更多亲子陪伴时间，能更关注孩子的个性化需求和全面发展。面对这项方兴未艾、快速迭代的新技术，家长和幼儿在同一起跑线上，一起参加讲座研学，一起点亮新技能，一起创作有趣作品，一起感受成果喜悦，家长的阅历通识，孩子的奇思妙想，相映成趣，在共同创作 AIGC 公众号里，把平常生活化为童话之旅。

《问卷》结果显示，虽然受访者表示挺了解 AI 的仅占 45%，但高达 82%表示支持和很支持“孩子学习使用 AIGC”，37%“和孩子一起学习和使用 AIGC”，更有 81%期待或很期待“和孩子一起参加 AIGC 研学、参观、体验、讲座”，对此表示不作考虑的仅有 3%。可见，多数人对于 AIGC 促进亲子教育互动有共鸣、有期待。

2.3.3. AIGC 有助于家园共育

幼儿园作业经过 AIGC 赋能，学习材料搜集更便捷，学习路径设计更丰富，学习活动互动更出彩，家园共育更有效。孩子将 AIGC 生成的古诗画面、背景故事、趣味改编，用公众号文章固定下来，在课堂上用孩子的视角、孩子的语言，向同学讲述，与同学分享。孩子将幼儿园里的趣事写入公众号、加入 AIGC，让家长倾听了解，让同学回味分享，如《幼儿园介绍自己名字》[4]《有魔法的贴贴纸》[5]《我是小砾》[6]等幼儿园系列，拉近了家庭教育和校园

教育。老师很乐意让孩子上台讲解，孩子在讲解中获得了更多成就感，家长与老师的交流配合也更为密切。

《问卷》结果中，对于“AIGC 应用对家庭和幼儿园的家园共育，会带来帮助”的看法，70%表示很同意或同意，不同意和很不同意仅占 8%，显示了绝大多数受访者对 AIGC 促进家庭和幼儿园共同培育作用的认可。

3. 思考展望

《问卷》结果显示，受访者中 77%表示自己想或很想学习 AIGC，82%支持和很支持“孩子学习使用 AIGC”，更有 48%认为“幼儿园需要设置小朋友学习 AIGC 课程”。AIGC 从幼儿抓起，已有一定程度的社会共识和推广普及需要。

但对幼儿使用 AIGC 的风险和害处，受访者也表示了许多忧虑。最普遍的担心，依次为“看屏幕太多，容易近视”（79%）、“容易沉迷手机小视频和游戏"（61%），“容易滥用 AIGC 完成作用"（60%）;最强烈的担心，依次为“容易滥用 AIGC 完成作业”（37%）“小朋友不应该太早接触使用手机电脑”（32%），“容易沉迷手机小视频和游戏（32%）”。此外，受访者对使用 AIGC 容易混淆真实和虚拟、接触网络陌生人容易上当受骗表示了忧虑，均占 47%。虽然这些问题，更多来自于 AIGC 使用所涉工具，而并非其本身，但是对幼儿健康的负面影响，对作业“取巧”的担忧顾虑，乃至对幼儿认知形成的潜在风险，都是公众所关注的问题。

3.1. AI 是科技赋能，更需人性光辉

AI 的聪明、博学、高效，让我们惊叹甚至困惑。《运动员代表，代表什么？》中，扑满对 6 秒钟 AI 一下子写了 5 篇发言稿，非常震惊，进而纠结“让 AI 写我的运动员代表发言，这样子对吗？它又不是运动员。”张文宇教授讲座对 AI 与人性关系的深刻见解，开启了我们的思考——“我们应该去发挥蕴含着人类灵感与人性光辉的创新能力，把自己进化得更像人”，“天才是 1%的灵感加上 99%的汗水，人工智能可以代替人类 99%的汗水，但那 1%的人类灵感和人性光辉，是机器取代不了的”。这让我们对“滥用 AI 完成作业”的普遍困惑，有了自己的回答：“让机器回答，不能让我们变得更聪明，还会变懒变笨。AI 也是这样，不是要它的答案，是让它帮我们一起找更好的答案。”这更让我们警觉，在陪伴孩子中，要更注重一起发现人类的灵感，一起追求人性的光辉，一起多陪伴老人、多体验传统节日，多看代表人类美好的作品，像陪奶奶过年的感受写成一首《年味是什么》，像下雪了妈妈喊冷时让妈妈摸摸他很烫很烫的耳朵就不冷了的温馨画面，像夜间急诊里发现和感谢这个城市的“月光守护者”们。这些体验和感动，是远比公众号、比 AIGC，更让我们珍惜和欣喜，更能让孩子学会与世间美好相伴。

3.2. AI 是智能学伴，更要智慧同行

在学习 AI 技术中，我们深感 AI 进化之快，从一开始傻瓜式“胡说八道”回答、六指畸形图片的初级阶段，到秒生歌曲、堪比大片的文生视频的技术大爆发，间隔时间越来越短。一开始，我们在网上盲目搜索操作技能，去寻找系统课程，后来发现，实地体验、社群分享，才是更管用的问道之途。家长和扑满一起参加中国 AIGC 产业联盟、温大 AIGC 俱乐部、AIGC 音乐体验工作坊等社群，在群友分享中收获了很多 AI 技能。特别是谷雨、方建文、倪考梦等 AI 领域专家，对我们认识 AI、应用 AI，都带来非常大的帮助。

在这些活动和交往中，孩子更看到了更广阔的 AI 世界，比如，香味电视机、AI 设计服装、AI 装修设计等，都激发了孩子对 AI 强烈的兴趣，甚至为此更主动地学习艺术知识，更自觉地自我管理时间。为了写出更准的提示词，扑满更愿意去读百科全书绘本，为了让 AI 公众号有新学会钢琴曲配乐、新背会古诗制片，他对传统教育内容学习更有兴趣、更有动力。特别是 AIGC 使用中，更多学会了提问的能力、追问的能力、分辨是不是“胡说八道”的能力，这在未来教育中将日益重要。

3.3. AI 是学习工具，更是未来教育

AIGC 需要借助手机、微信和电脑，有不少人反对幼儿过早接触，我们也经常心生顾虑，会不会影响视力，会不会产生手机依赖，会不会影响正常学习。孩子自己也在思考这个问题，也在注意这些问题，还带着家长学会手机自律。

接触 AI 日久，我们感觉，AI 正在深刻的改变我们的学习、生活，乃至思维方式，在可见可感的不久将来，AI 还将系统结构、全面重塑我们的生产和工作，这将是一场颠覆性的变革，正如张文宇教授所说，AI 是新质生产力的引擎，一次人类的伟大革命，其意义甚至可比人类发明用火。

如何享受 AI 时代红利，如何面对 AI 时代冲击，我们相信，人与 AI 的人机对话的能力，与 AI 同行的能力，将会成为未来主人翁这一代人的核心竞争力。关心孩子成长，要像我们这代人重视外语能力一样，把这种人机对话语言能力培养，作为一门基础语言，让 AI 不仅是孩子的学习工具，更成为学习本身。

参考文献：

扑满. (n.d.-b). *牛虎大战第十四集·空城计*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/YxeNUP3U36b76R2zkUCiBA

扑满. (n.d.-c). *一起画个空间站*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/xNt89aePa7dOvfZKx9PR4A

有心弗怕路远，海燕归巢可待，祝温州海外学子龙行龘龘前程朤朤！. (2021). 微信公众平台. https://mp.weixin.qq.com/s/tdBh_U0LeYIOqOU54qtrLg

扑满. (n.d.-c). *扑满的西游记*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/uIbZqPz_MVbakx8MukcDJg

扑满解诗·王安石《梅花》. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/fIEB-7c30R5c_fXX3s-cbA

童趣雪花. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/RUP_gkg82eG1iVK8z1g3ow

月光守护者. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/KtHUYkho1r3HWjF86xG7pw

新月集·12 点钟（扑满读诗）. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/AqpPAgy4RBUSKT2jDtlfZg

扑满. (n.d.-a). *【探秘】元日，原来古代的爆竹、屠苏酒、桃符！怪兽竟然这么害怕？*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/uDRTqqosq9Ed81MemR1KYg

扑满. (n.d.-a). *【揭秘】歪打正着背古诗：AI 领你解码杜牧《清明》背后的秘密*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/hSKecTXmqp_Qm5bPoQEMpQ

《题惠崇春江晚景》，诗入画，文生图. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/IzzaCz25lflvRbKYjAF4lQ

诗里藏着谜语，小朋友们快来猜猜吧. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/moo8Oq-eSEnCdjmbZBTiDA

科学实验——桔子的秘密. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/_9eB7GSvAzXuZJmb-QBscA

干冰大战狂风怪. (n.d.). Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/H5n0us2_x3-94uirVfGzbg

扑满. (n.d.-c). *电梯探秘之旅*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/X2qN0QJ3rN6x_Gbv00eD_A

扑满. (n.d.-g). *探索元宇宙 【惊呆了！香味电视机，居然能闻到花的香味！】*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/l1ITmc_OfPeHNlvVtXv50w

扑满. (n.d.-d). *第一次听大人的讲座*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/VlqVvTaRXjhXz1f17n1fMg

扑满. (n.d.-a). *【惊爆！】机器换人之后，人工智能将代替白领精英？2024 年 AIGC 俱乐部灵魂拷问：未来职场风云巨变，你上车了吗？*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/i53_E4z5St1-QAoYJZPh8A

扑满. (n.d.-k). *幼儿园介绍自己名字*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/ddX8Xza-Bwe4qO7w7SO0bw

扑满. (2021). *有魔法的贴贴纸*. 微信公众平台. https://mp.weixin.qq.com/s/n8DyfFFh8Lnyd9LvVA2ZAw

扑满. (n.d.-j). *我是小砾*. Weixin Official Accounts Platform. Retrieved April 23, 2024, from https://mp.weixin.qq.com/s/lD6DctLrqkZs127F22lBRg

Epistemic Agency for STEM Learning through Design-based Integrative Approach

Jialiang LIU, The University of Hong Kong, ljliang@connect.hku.hk
Carol K.K. Chan, The University of Hong Kong, ckkchan@hku.hk
Jianhua Zhao, Southern University of Science and Technology, zhaojh@sustech.edu.cn
Jing Kong, FoShan University, kongjing_2022@fosu.edu.cn

Abstract: *This study was to develop a design-based scaffold that would enhance students' epistemic agency by utilizing documentation and reflection to foster their STEM inquiry. A quasi-experimental design was adopted in a Chinese college to investigate the effects of the design-based scaffold on enacting students' epistemic agency. The experimental class (n = 25) used the design-based scaffold, whereas the comparison class (n = 25) conducted the same knowledge-building activities without the scaffold support. Quantitative findings suggest that students from the experimental class created more innovative artifacts and engaged in higher-level epistemic moves than those from the comparison class. Qualitative data analysis illustrated the collaborative trajectory with the scaffold in the knowledge building environment. Implications of epistemic agency for STEM learning are discussed.*

1. Introduction

Knowledge creation practices emerge as a dynamic process that addresses constantly evolving challenges and uncertainties (Cetina, 2000; Paavola & Hakkarainen, 2005; Zhang et al., 2022). To engage students in real knowledge creation activities, science education is situated within an educational philosophy that places great emphasis on authentic science practices (National Research Council, 2012; NGSS Lead States, 2013). To nurture such authentic practices in the science context, educators need to enable students to assume epistemic agency (Odden et al., 2023; Zhang et al., 2022). Some efforts were dedicated to design social configuration that promote students' agency in leading and regulating their own collaborative learning (González-Howard & McNeill, 2020; Ko & Krist, 2019; Miller et al., 2018; Tao & Zhang, 2021; Zhang et al., 2022).

STEM learning can be treated as knowledge creation activities that "learning takes place through collaborative activities whose aim is to create new knowledge through work on shared objects" (Damşa et al., 2010, P146). Students were expected to assume epistemic agency to deal with the uncertainty and diverse ideas in STEM learning. However, a significant portion of STEM education in higher education still adheres to traditional instructional models that restrict students' agency over their own learning (Apkarian et al., 2021; Lombardi et al., 2021; Viskupic et al., 2019). Furthermore, there has been a notable lack of empirical research investigating the cultivation of epistemic agency in STEM learning (Odden et al., 2023). Thus, we are interested in exploring the ways that design-based scaffolds can be leveraged to create more opportunities for epistemic agency in higher education STEM.

2. Epistemic agency for STEM learning with design-based scaffold

In this study, we drew on the Knowledge Building pedagogy to examine the design-based scaffold that effectively fostered students' epistemic agency in STEM activities. Knowledge Building (KB), is an educational approach that involves learners continually work on ideas improvement and pushing the boundaries of knowledge through collective efforts in a community (Bereiter, 2002; Chan & van Aalst, 2018; Scardamalia & Bereiter, 1994, 2014a)). Knowledge Forum (KF), a computer-supported collaborative learning environment, was customized for KB integrating theory, pedagogy and technology (Scardamalia, 2004; Scardamalia & Bereiter, 2014b). Knowledge building environment supports multimodality, such as text, images, and videos, allowing students to visualize and document their trajectory of STEM inquiry. KB involve a serious of principles and dynamics that encourage students work progressively on idea improvement and knowledge advancement (Scardamalia, 2002). Epistemic agency, as a fundamental principle of Knowledge Building, entails students taking personal and collective responsibility for guiding the trajectory of sustained inquiry and the knowledge advancement (Scardamalia, 2002; Scardamalia & Bereiter, 1991) and can be defined as "students being positioned with, perceiving, and acting on, opportunities to shape the knowledge building work in their classroom community"

(Miller et al., 2018, P6). To facilitate students' epistemic agency in STEM learning, we deployed design-based scaffold to help students document, reflect, and manage their process of STEM inquiry.

The design-based scaffold included two parts, "problem-scoping" scaffold and "iteration" scaffold. The "problem-scoping" scaffold were developed to help students in understanding the boundaries of problems and identifying problems constrains and goals. The scaffolds included the following prompts: "What is the core problem?", "What information is needed to explain the problem?", "What are the constraints?", and "What ideas do you have for problem solving". The "iteration" scaffold was provided to students for iterative refinement including "What works have been developed so far (text, images, and videos)?", "Does it meet the constraints of the core problems?", "What are the problems with the current work?", "What measures can be taken to improve the work?", and "What are your additions and refinements in response to others' ideas?".

First, the design-based scaffold leveraged the documentation of STEM inquiry. As students engaged in STEM activities, such as scoping authentic problems and fabricating artifacts, they increasingly worked on deeper understanding and better design with the facilitation of the scaffolds, including framing constrains of the core problems and current achievement regarding artifacts. In addition, the documentation of STEM inquiry provided students with materials for reflection, and process-oriented reflection facilitated students in identifying problems, developing measures for improvement, adding new ideas to the design, and transforming them into better products. Therefore, the design-based scaffold offered a social configuration that enacted students' epistemic agency in the process of STEM inquiry. Students assumed epistemic agency through constructing group mechanism, defining the core problems, figuring out the solution, and designing the artifacts. The mechanism arbitrated the communication and discourse moves as they proposed new ideas and found problems or deficiencies in artifacts throughout the inquiry process. By engaging reflective activities on documentation, students could monitor, adjust, and plan their inquiry, which guided and adapted the trajectory of collaborative learning.

We see great opportunity in the scaffold to help students engage deeper interaction, high-level discourse moves, and more innovative artifacts in a knowledge building community. Therefore, the study adopted a comparative method to investigate design-based scaffolds for leveraging students' epistemic agency in STEM learning. Three questions were asked: (1) Did students in the experimental class created more innovative artifacts? (2) What epistemic moves were enacted by the experimental class? (3) How did student engage in collaborative trajectory with the design-based scaffolds?

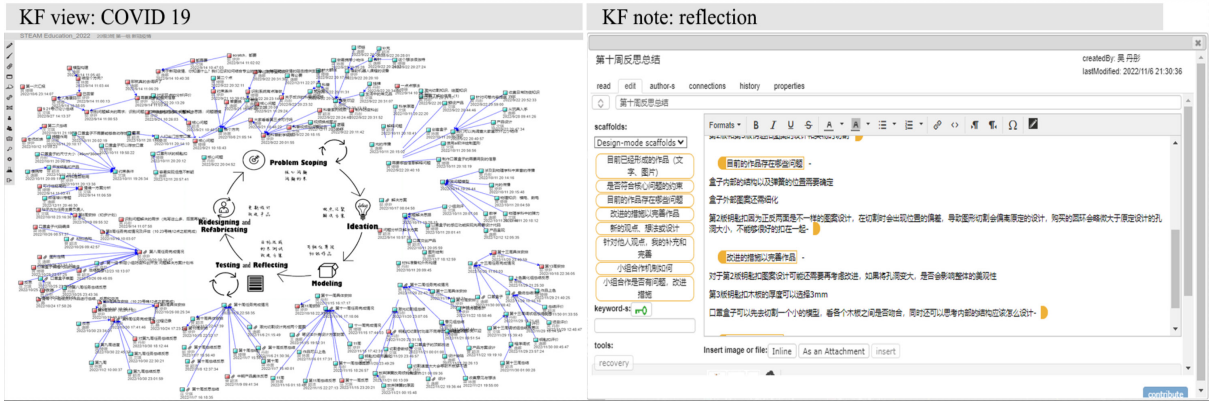
3. Methods

The study involved a quasi-experimental design that conducted in two undergraduate classes at a university in South China, including 25 students in each class. Both classes studied the same course, "STEM Education" which lasted for 16 weeks with two lessons per week and were taught by the same teacher who held a doctoral degree in educational technology.

A knowledge-building environment with specific scaffolds for problem solving and documentation was introduced to support students' STEM learning (Figure 1). In phase I, students were encouraged to propose diverse ideas with respect to the topics "Pandemic," "Rocket," "Bridge," and "Culture" and were divided into four groups based on their interests. In phase II, students first formulated and scoped authentic problems and engaged in artifacts production. The "problem-scoping" scaffolds were used to help students in understanding the boundaries of problems and identifying problems constrains and goals. Then, they started to figure out the solution to the real problems and fabricated the artifacts. The "iteration" scaffold helped students document the process of artifacts production, and more importantly, reflected on the documentation to identify error and deficiencies, check the meeting of constraints, goal, and attainment, and develop strategies for improvement, add new ideas to the design, and transform the products into better ones. In phase III, students collectively reflected on artifacts and STEM understanding that they gained in STEM learning.

For the comparison class, students engaged in STEM learning following same procedure in a regular KB environment without "problem-scoping" scaffolds and iteration scaffolds.

Figure 1
KF View and Notes



4.Data sources

To answer research questions, research data, such as ATK indices and discourse on KF, artifacts, and classroom observation, were collected and analyzed to illustrate the effect that KB with design-based scaffolds for students' epistemic in STEM learning. Students' artefacts rated in terms of 10-point Likert scale Likert scale, including resolution, style, and novelty (Table 1) that O'Quin and Besemer (2006) framed as a measurement for creative work. Knowledge Forum notes were examined for discourse moves reflecting knowledge-building processes based on a coding scheme commonly used in other KB studies (Table 2).

Table 1

Coding Scheme for Students' Artifacts

Dimension	Description
Resolution	Functionality, relates to logical, useful, value and understandable factors in the artifacts.
Style	Presentation to customers, consist of organic, well-crafted, and elegant indicators in the artifacts.
Novelty	Creativity, such as new materials, new processes, new concepts, and other elements of newness in the artifacts, incudes originality and surprise.

Adapted from O'Quin and Besemer (2006) and Hong et al. (2019).

Table 2

Coding Scheme for Content Analysis of Discourse Moves

Codes	Sub-codes	Description (*examples*)
Questioning	Fact-seeking	Questions about definition of the terms or seeking factual information *(e.g., what is plan of inquiry this week?)*
	Explanation-seeking	Questions about seeking open-ended responses with explanation *(e.g., How is modern technology used to reproduce traditional wedding clothes?)*
	Sustained inquiry	Asking further questions based upon previous notes or ideas to deepen the discussion *(e.g., As we defined the core problem, what's the plan to solve it?)*
Theorizing	Simple claim	Simple (dis)agree or repeat a statement *(e. g., yes, I agree.)*
	Proposing an explanation	Proposing a theory that explains certain phenomena for the first time *(e.g., I think technology can help students understand experimental principles better.)*
	Supporting an explanation	Enriching a theory with elaboration and new information *(e.g., you're right, if you can use digital technology to present the process of the experiment.)*
	Improving an explanation	Supporting an already existing theory another student proposed and providing a justification *(e.g., for specific technology, we can use 3D software to project the wedding clothes)*
Community	Connection	Introducing information from websites, papers, cases, other material as evidence to strengthen ideas and generate questions, or reference a note that they or others created *(e.g., the following link is about how to protect the*

		following four types of people psychologically during the pandemic)
	Lending support	Proposing to move to a new stage of inquiry and sustain the willingness to learn *(e.g., as we have finished the current phase of the plan already, I think we need to move forward to fabricate the design into an artifact.)*
	Problem detecting	Finding problems existed in the current work *(e.g., The hole is too small, the ring can't go through, keep fine-tuning to find the most suitable parameters.)*
	Synthesizing notes	Coordinating and integrating collective ideas from previous discussion with a reflection *(e.g., Referencing some key notes in the discussion, "our core problems are as follows: protect and promote Cantonese by using short videos and targeting college students in our school.")*

Adapted from Hakkarainen (2003) and (Fu et al., 2016).

5.Finding

RQ1: Effects on Artefacts in STEM learning

We examined students' final artifacts that made with facilitation of STEM design-based scaffolds in the knowledge-building environment. There was total 8 artifacts, 4 in each class. Table 3 illustrates Mean and SD of students' artifacts across the two classes. A one-way ANOVA test suggest statistic significances in functionality ($F(1, 198) = 7.740$, $p < .05$), value ($F(1, 198) = 6.512$, $p < .05$), comprehensibility ($F(1, 198) = 5.411$, $p < .05$), and well-craftedness ($F(1, 198) = 4.398$, $p < .05$), indicating that with facilitation of design-based scaffolds, students enacted higher-level epistemic agency and created more innovative artifacts.

Table 3

Means and SDs of Students' Artifacts Across the Two Classes.

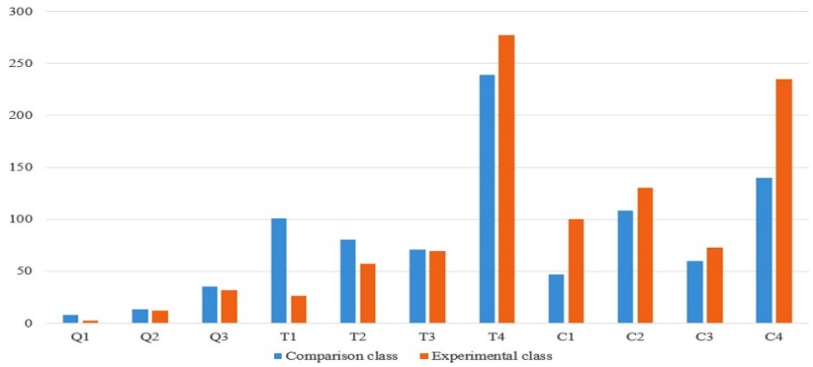
Dimension	Sub-dimension	Experimental class	Comparison class
Novelty	Originality	8.49/1.059	8.21/1.274
	Surprise	8.30/1.202	7.92/1.515
Resolution	Functionality	8.16/1.229	7.62/1.503
	Usefulness	8.19/1.116	7.85/1.445
	Value	8.25/1.149	7.76/1.538
	Comprehensibility	8.47/.937	8.08/1.390
Style	Organic qualities	8.38/1.270	8.24/1.256
	Well-craftedness	8.66/.956	8.32/1.309
	Elegance	8.53/1.141	8.12/1.493

RQ2: Difference on Epistemic Moves

To test the effect of the design-based scaffolds on students' knowledge-building discourse, MANOVA tests was conducted to analyze the difference between the two classes. The results showed that comparison class made significantly more simple claims, $F(1, 48) = 11.159$, $p < .05$, while experimental class had conducted significantly more connection $F(1, 48) = 6.441$, $p < .05$ and synthesizing, $F(1, 48) = 24.067$, $p < .01$. This analysis (Figure 2) demonstrated that students in experimental class engaged higher-level epistemic moves (connection and synthesizing) compared to comparison class (simple claims).

Figure 2

Comparison of Epistemic Moves between the Two Classes



Questioning: Q1: Fact-seeking, Q2: Explanation-seeking, Q3: Sustained inquiry; Theorizing: T1: Simple claim, T2: Proposing an explanation, T3: Supporting an explanation, T4: Improving an existing explanation; Community: C1: Connection, C2: Lending support, C3: Problems detecting, C4: Synthesizing notes)

RQ3: Epistemic Agency for Collaborative Knowledge-Building Practices with Design-Based Scaffolds

To investigate the way that a design-based scaffolds leveraged students' epistemic agency in STEM learning, six themes with a specific case, "Bridges," are introduced to characterize the processes that collaborative STEM inquiry in knowledge-building environment. Teacher's observation and students' KF notes were the data source for analysis.

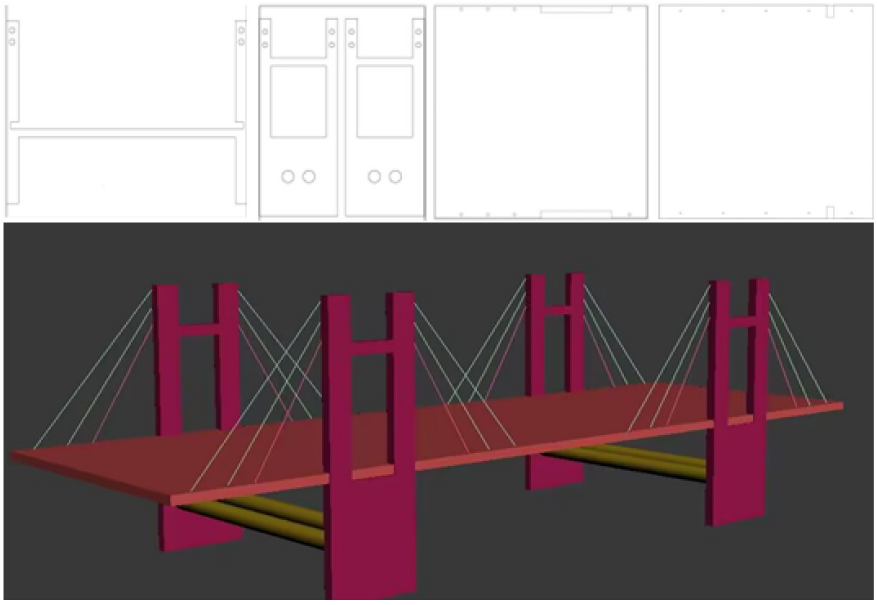
Theme 1: Proposing diverse ideas and forming groups. Students assumed epistemic agency to post the initial ideas that they were interested in (*S2: What are the types of bridges? What materials are needed for construction?*) and were divided into different groups based upon their interests.

Theme 2: Scoping the core problems with problem scoping scaffolds. A set of problem-scoping scaffolds were used to facilitate students to scope the authentic problems. The data suggest that students actively engaged in a discourse inquiry on problems scoping, including searching information, discussion, constrains identifying, and ideas for problems solving (*S2: [What are the core problems?] How is the bridge load-bearing? S1: [Constraints?] What are the factors that affect the load-bearing capacity, such as size and material, that need to be considered? How much tonnage can be supported?*). Meanwhile, students used the scaffolds as guidance to summarize and reflect their achievement in form of e-portfolios.

Theme 3: Ideation and Sketching. As the authentic problems were posed, students were guided by the common objectives and began to generate ideas on how to solve the problems. The iteration scaffolds supported them to develop ideas for problem solving and display the main ideas into sketches taking constrains and goals into consideration (Figure 3). Sketching required students think systematically about the components and forms that build connection and annotation between sketches and disciplinary knowledge as well. To achieve these accomplishments, students formed group interactive mechanism, created design sketches, and fabricated the artifacts. (*S4: To achieve our goals in the experiment, there was a reasonable division of labor with a designated note-taker, a modeling analyst, and a draughtsman. Each group member actively contributed their personal opinions, raised any remaining doubts, and other members provided active clarification. Any areas of uncertainty were subsequently documented and resolved*.)

Figure 3

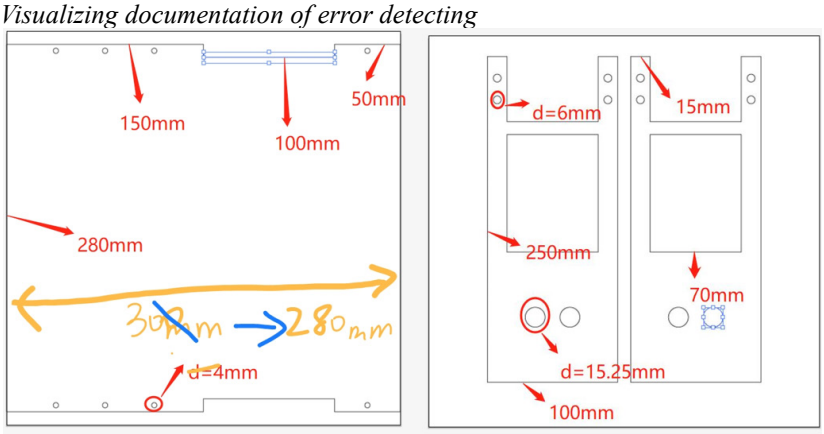
Initial sketches of bridge structures



Theme 4: Modeling and transforming design into products. With initial design, students moved to modeling stage. The KB environment supported any forms of representation (words, pictures, and videos), allowing students to document the process of modeling. They iteratively optimized their products with the support of iteration scaffolds though reflecting on the documentation of modeling process. The strength of reflecting on process documentation were problems and deficiencies detecting and increasing the efficiency and feasibility of design. Some examples included (*S6: ["What are the problems with the current work"] Due to certain deviations in the dimensions of the cut-out parts, there are certain gaps between the parts, the slots and holes. S3: [Measures for improvement] Based on the previous records of sketches, the following modifications can be made, see the picture (Figure 4).*). The reflection on documentation helped students monitored and adjusted the process of collaborative design through shaping and reshaping their evolving inquiry directions based on the emergent problems and new ideas. (*S3: The problems acquired from reflection on documentation and new ideas that group members actively proposed greatly contributed to the design of the bridge.*)

Figure 4

Visualizing documentation of error detecting



Theme 5 Reflecting and improving for final artifacts. The iteration scaffold enacted students' epistemic agency for better design through reflecting on their products and being aware of problems and deficiencies. Some examples included (*S3: [What are the problems with the current work] - revise the design and explore what is appropriate in the mortise and tenon construction? S4: [Measures for improvement] Choose a triangular shape for a more stable mortise and tenon construction.*). Testing the products' functionality and feasibility whether they meet the constrains and goals. Through testing and reflection, students analyzed their making progress, found strengths and weaknesses, and redesign and reconstruct their products applying disciplinary knowledge. With the

facilitation of the scaffold, students found the problems, new measures, and directions that guided their collaborative process for artifacts refinement. Eventually, they reached a balance and finalized the products (Figure 5).

Figure 5
Artifact from the "Bridges" group



Theme 6: Collective reflection on the trajectory of collaborative design. At end of STEM inquiry, students collectively reflected their experiences of design and fabrication and what they had learned about disciplinary knowledge, and STEM understanding, etc. *S5: Through the bridge building, I had become more proficient in the use of the laser engraving machine.* After STEM learning, students recognized that "*S6: STEM education can truly develop our creative and complex capabilities, including problem-solving, critical thinking, communication and collaboration, regulated learning and lifelong learning*" and "*S4: It is necessary for everyone to assume corresponding responsibilities, collaborate and divide tasks, and exert their subjective agency in order to complete the project*".

6. Scientific and scholarly significance of the study

This study highlighted the design-based scaffold enacting students' epistemic agency for STEM inquiry in knowledge building environment. The quantitative and qualitative evidence advocated that the designed KB environment with the design-based scaffold was conducive to students' epistemic agency in terms of artifacts production and epistemic moves. The study enriched the empirical evidence that scaffolds can be a facilitator for students' epistemic agency in higher education STEM programs.

References

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E., & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PloS one*, *16*(2), e0247544-e0247544. https://doi.org/10.1371/journal.pone.0247544

Bereiter, C. (2002). *Education and mind in the knowledge age*. Lawrence Erlbaum Associates. https://doi.org/d9sw

Cetina, K. K. (2000). Objectual practice. In K. K. Cetina, E. von Savigny, & T. R. Schatzki (Eds.), *The Practice Turn in Contemporary Theory* (pp. 184-197). Routledge. https://doi.org/10.4324/9780203977453-22

Chan, C. K. K., & van Aalst, J. (2018). Knowledge building: Theory, design, and analysis. In F. Fischer, C. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 295-307). New York: Routledge. https://doi.org/d9t9

Damşa, C. I., Kirschner, P. A., Andriessen, J. E. B., Erkens, G., & Sins, P. H. M. (2010). Shared Epistemic Agency: An Empirical Study of an Emergent Construct. *The Journal of the Learning Sciences*, *19*(2), 143-186. https://doi.org/10.1080/10508401003708381

Fu, E. L. F., van Aalst, J., & Chan, C. K. K. (2016). Toward a classification of discourse patterns in asynchronous online discussions. *Intern. J. Comput.-Support. Collab. Learn*, *11*(4), 441-478. https://doi.org/10.1007/s11412-016-9245-3

González-Howard, M., & McNeill, K. L. (2020). Acting with epistemic agency: Characterizing student critique during argumentation discussions. *Science education (Salem, Mass.)*, *104*(6), 953-982. https://doi.org/10.1002/sce.21592

Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching*, *40*(10), 1072-1088. https://doi.org/dw6p8x

Hong, H.-Y., Lin, P.-Y., Chen, B., & Chen, N. (2019). Integrated STEM Learning in an Idea-centered Knowledge-building Environment. *Asia-Pacific Edu Res*, *28*(1), 63-76. https://doi.org/10.1007/s40299-018-0409-y

Ko, M. L. M., & Krist, C. (2019). Opening up curricula to redistribute epistemic agency: A framework for supporting science teaching. *Science education (Salem, Mass.)*, *103*(4), 979-1010. https://doi.org/10.1002/sce.21511

Lombardi, D., Shipley, T. F., Bailey, J. M., Bretones, P. S., Prather, E. E., Ballen, C. J., Knight, J. K., Smith, M. K., Stowe, R. L., Cooper, M. M., Prince, M., Atit, K., Uttal, D. H., LaDue, N. D., McNeal, P. M., Ryker, K., St. John, K., van der Hoeven Kraft, K. J., & Docktor, J. L. (2021). The Curious Construct of Active Learning. *Psychological science in the public interest*, *22*(1), 8-43. https://doi.org/10.1177/1529100620973974

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053-1075. https://doi.org/10.1002/tea.21459

O'Quin, K., & Besemer, S. P. (2006). Using the Creative Product Semantic Scale as a Metric for Results-Oriented Business. *Creativity and Innovation Management*, *15*(1), 24-44. https://doi.org/http://dx.doi.org/10.1111/j.1467-8691.2006.00367.x

Odden, T. O. B., Silvia, D. W., & Malthe-Sørenssen, A. (2023). Using computational essays to foster disciplinary epistemic agency in undergraduate science. *Journal of Research in Science Teaching*, *60*(5), 937-977. https://doi.org/10.1002/tea.21821

Paavola, S., & Hakkarainen, K. (2005). The Knowledge Creation Metaphor – An Emergent Epistemological Approach to Learning. *Sci Educ*, *14*(6), 535-557. https://doi.org/10.1007/s11191-004-5157-0

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a Knowledge Society* (Vol. 97, pp. 67-98). Open Court. https://bit.ly/3013EQ8

Scardamalia, M. (2004). CSILE/Knowledge Forum. In A. Kovalchick & K. Dawson (Eds.), *Education and technology: An encyclopedia* (pp. 183-192). ABC-CLIO. https://bit.ly/303RJkB

Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences*, *1*(1), 37-68. https://doi.org/fqvz5q

Scardamalia, M., & Bereiter, C. (1994). Computer Support for Knowledge-Building Communities. *The Journal of the Learning Sciences*, *3*(3), 265-283. https://doi.org/10.1207/s15327809jls0303_3

Scardamalia, M., & Bereiter, C. (2014a). Knowledge Building and Knowledge Creation: Theory, Pedagogy, and Technology. In (pp. 397-417): Cambridge University Press.

Scardamalia, M., & Bereiter, C. (2014b). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397-417). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.025

Tao, D., & Zhang, J. (2021). Agency to transform: how did a grade 5 community co-configure dynamic knowledge building practices in a yearlong science inquiry? *International Journal of Computer-Supported Collaborative Learning*, *16*(3), 403-434. https://doi.org/10.1007/s11412-021-09353-7

Viskupic, K., Ryker, K., Teasdale, R., Manduca, C., Iverson, E., Farthing, D., Bruckner, M. Z., & McFadden, R. (2019). Classroom Observations Indicate the Positive Impacts of Discipline-Based Professional Development. *Journal for STEM education research.*, *2*(2), 201-228. https://doi.org/10.1007/s41979-019-00015-w

Zhang, J., Tian, Y., Yuan, G., & Tao, D. (2022). Epistemic agency for costructuring expansive knowledge-building practices. *Science education (Salem, Mass.)*, *106*(4), 890-923. https://doi.org/10.1002/sce.21717

小学知识建构课堂中协作认识代理的模式研究

Understanding Primary Students' Shared Epistemic Agency in Knowledge Building Processes

张晓洁[1, 2]，陈桂涓[1]，冯诗惠[1]，赵建华[3*]
[1]香港大学 教育学院
[2]南方科技大学 高等教育研究中心
[3]南方科技大学 未来教育研究中心
*zhaojh@sustech.edu.cn

【摘要】 知识建构鼓励学生参与高水平的集体知识建构，包括提出他们关心的问题，与同龄人交换概念资源，以及在彼此的想法基础上建构共同体知识。在集体知识建构中，学生要共同承担促进集体知识进步的责任，即协作认识代理（Shared Epistemic Agency）。本研究旨在通过有向网络分析（Ordered Network Analysis，ONA），对知识建构环境下学生的协作认识代理的模式提供更深入的见解。研究结果表明，高表现水平的小组会在讨论过程中更加主动地承担改进现有观点、总结不同观点、以及管理讨论进度的责任。这些协作认识代理在讨论过程中出现的时机也许是促进共同体知识发展的关键。

【关键词】 知识建构；协作认识代理；学习分析；有向网络分析

Abstract: *Knowledge Building encourages students to engage in high-level collective knowledge construction, including proposing their questions, exchanging information with peers, and building community knowledge based on each other's ideas. In collective knowledge building, students should take the collectively responsibility of promoting collective knowledge progress, namely the Shared Epistemic Agency. The aim of this study is to provide deeper insights into the patterns of shared epistemic agency in a knowledge construction environment through ONA network analysis. The research results indicate that high performing groups are more proactive in taking responsibility for improving existing ideas, summarizing different perspectives, and managing discussion progress during the discussion process. The timing of these shared epistemic agency may be key to promoting the development of community knowledge.*

Keywords: Knowledge building; Shared epistemic agency; Learning analysis; Ordered network analysis

1.前言

协作学习和探究能力对21世纪的学生来说非常重要。随着技术的发展，计算机支持的协作学习（CSCL）为学生提供了各种学习环境，让他们随时随地进行协作探究。作为CSCL领域的一种基于原则的教学法，知识构建（Knowledge Building，KB）旨在促进知识创造社区的发展（Scardamalia 和 Bereiter，2014）。KB鼓励学生参与高水平的集体知识建构，包括提出他们关心的问题，与同龄人交换观点和信息，以及在彼此观点的基础上建构集体知识（Scardamalia 和 Bereiter，2014）。在集体知识建构中，学生要共同承担促进集体知识进步的责任，即协作认识代理（Shared Epistemic Agency）。在知识建构中，协作认识代理是指学生对自己的学习拥有责任并为观点发展做出贡献的能力（Scardamalia 和 Bereiter，2014）。高水平的认识代理对知识创造至关重要，知识创造需要学生发现复杂的真实问题，不断改变解决这些问题的策略，并在集体知识中取得进步。具有高水平认识代理的学生可以参与高级认知行动，提高对现有知识的共同理解，并产生和应用创新观点（Damşa 等，2010）。

Damşa 等（2010）将学生的协作认识代理分为两个维度。第一个维度，即认知维度，包括识别知识差距、促进共同体知识、协商不同想法、综合不同想法和产生新想法等活动。第

二个维度，元认知维度，包括以元认知为中心的活动。这包括规划未来的活动和制定联合计划，监测实现目标的进展，协调与目标相关的活动，以及反思和规范合作活动。研究表明，协作认识代理在成功的合作探究和知识建构中起着至关重要的作用 (Chen 等，2015；Yang 等，2024)。然而很少有研究关注小学生在知识建构过程中的协作认识代理。因此，本研究旨在探究小学知识建构课堂中，不同水平的协作认识代理。

KB 强调随着时间的推移而发生的集体知识改进。因此有必要对共同体知识的集体构建过程和学生协作认识代理的时序发展进行研究。认知网络分析（Epistemic Agency Analysis，ENA）可以在特定的时间窗口内分析在线论坛对话中个人或群体之间的联系，并允许研究者对社会认知行为之间的时间共现模式进行建模，可视化这些共现，并在不同群体或不同活动之间进行统计比较（Shaffer 等，2009）。已有研究对知识建构过程中的学生对话进行 ENA 分析（Hod 等, 2020；Yang，2023）。作为 ENA 的扩展，研究人员引入有向网络分析（Ordered Network Analysis，ONA）来解释研究现象的成分之间的关联方向（Tan 等，2022）。ONA 不仅可以捕捉元素之间的连接，也可以在统计和视觉上表示这些连接的强度和方向（Tan 等，2022）。因此，本研究旨在通过 ONA 网络分析，对知识建构环境下学生的协作认识代理的模式提供更深入的见解。本研究的研究问题为：不同水平的知识建构小组，协作认知代理的模式有什么不同？

2.研究方法

2.1.研究环境和参与者

本研究是在广东省的一所小学进行的。参与者为 28 名小学五年级的学生，这些学生参加了学校的课后延时服务课程。他们在为期 8 周，每周 2 节课（每节课 40 分钟）的课程中体验科学探究过程，探索超出课程标准的关于“光”的科学知识。

2.2.教学设计

学生参与的知识建构过程应用了基于先前 KB 研究（Yang，2022）修改的三阶段教学设计。第一阶段，学生参与在班级水平上的线上讨论，提出他们关于“光”的初步问题和多样性的观点（比如“光在微观上是沿曲线传播的”,“光有什么作用和危害”，“光和电之间有什么关系”）。第二阶段，学生分为 9 个兴趣小组，以小组为单位参与线上讨论探究活动，针对不同的子主题展开深入的讨论，写出实验计划并开展实验探究。第三阶段，在班级共同体中，学生整合不同子主题的知识并总结反思整个过程。

表 1

协作认识代理编码表

编码	**定义**
问题	提出关于讨论内容的问题
想法	对当前的讨论内容提出自己的观点和想法
新问题	在现有想法的基础上产生的新问题
信息	在网络上检索到的相关信息
改进	改进现有的想法
管理	管理讨论过程
总结	总结不同的想法

2.3.数据来源

为了探究不同水平小组的协作认识代理模式的不同，本研究选取第二阶段小组探究的数据作为数据来源。学生在讨论过程中产生的小组实验计划，以及实验探究结束完成的小组实

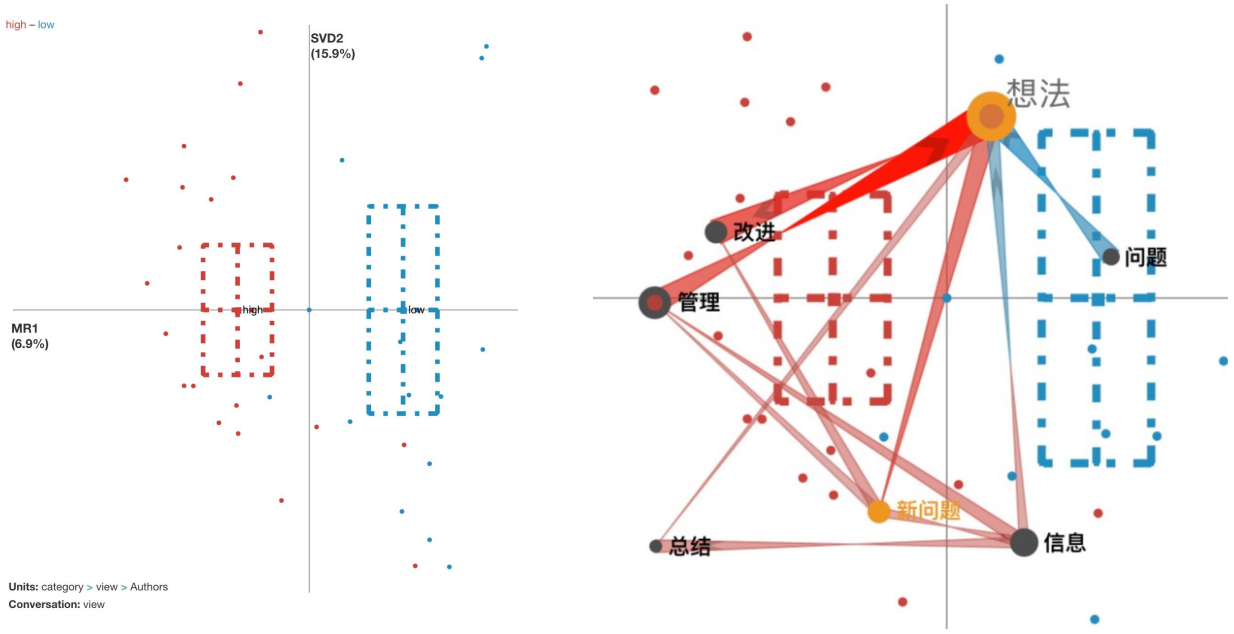
验报告被作为学生小组作品。根据这些学生作品，9 个兴趣小组被分为高表现水平的小组，以及低表现水平的小组。学生在小组讨论过程中产生的 203 条笔记被用来分析学生的协作认识代理。

2.4.数据分析

首先，学生的 203 条在线讨论笔记被编码为不同的认识代理类别。认识代理编码表改编自以往知识建构研究中的知识对话编码表（van Aalst，2009；Yang，2023），如表 1 所示。编码后的数据使用 ONA 网络工具进行分析（http://app.epistemicnetwork.org），考察学生认识代理之间的有向联系。

3.研究结果

3.1.高低表现小组之间的差异



（1）高表现小组（红色）与低表现小组（蓝色）ONA 网络图重心的区别

（2）高表现小组（红色）与低表现小组（蓝色）ONA 网络图的比较

图 1

高低表现水平小组的认识代理 ONA 比较

ONA 分析首先比较了高低表现水平小组的差异，如图 1（1）所示，Mann-Whitney 检验显示，在α= 0.05 水平上，高表现水平的小组（Mdn = -0.16，N = 21）与低表现水平的小组（Mdn = 0.23，N = 16，U = 316.50，p = 0.00，r = -0.88）在统计学上有显著差异。即高低表现小组的 ONA 网络分析存在显著性差异。

3.2.高低表现组认识代理模式的不同

图 1（2）显示了不同表现小组在参与小组讨论的过程中，协作认识代理模式的不同。首先，高表现水平的小组有“想法→改进”的认识代理模式，这说明在发表初始的想法后，高表现水平小组的学生会更加主动地承担起“改进现有的想法和观点”的责任，以促进集体知识的发展。比如在讨论“光的衍射”实验时，一位同学提出用带孔的圆盘来做实验，另一位同学紧接着提出更好的观点“可以用刀片代替”。其次，高表现水平的小组呈现出明显的“想法→总结”和“信息→总结”的认识代理模式，这表明这些同学更容易承担起总结不同的想法以促进知识

进步的责任。最后，高表现水平小组的“管理→想法”与“想法→管理”模式，表明高表现组的学生会在发表想法参与讨论的过程中，主动承担起管理讨论内容的责任。相反，低表现水平小组的学生则没有意识到改进现有的想法、总结不同的观点、以及管理讨论过程的责任。他们更多的在发表重复的观点和问题，以及简单的复制从网络上检索来的信息。

4.讨论与总结

本研究从 ONA 的角度对学生在知识建构对话中承担的协作认识代理提供了详细的见解。研究结果表明，高表现水平的小组会在讨论过程中更加主动地承担改进现有观点、总结不同观点、以及管理讨论进度的责任。这些协作认识代理在讨论过程中出现的时机也许是促进共同体知识发展的关键。下一步研究将关注于如何利用 ONA 分析的结果支持学生的知识建构反思过程，以促进低表现水平小组的协作认识代理发展。

基金项目

国家社会科学基金（BCA200090）

参考文献

Chen, B., Scardamalia, M., & Bereiter, C. (2015). Advancing knowledge-building discourse through judgments of promising ideas. *International Journal of Computer-Supported Collaborative Learning*, *10*(4), 345–366. https://doi.org/10.1007/s11412-015-9225-z

Damşa, C. I., Kirschner, P. A., Andriessen, J. E. B., Erkens, G., & Sins, P. H. M. (2010). Shared Epistemic Agency: An Empirical Study of an Emergent Construct. *Journal of the Learning Sciences*, *19*(2), 143–186. https://doi.org/10.1080/10508401003708381

Hod, Y., Katz, S., & Eagan, B. (2020). Refining qualitative ethnographies using Epistemic Network Analysis: A study of socioemotional learning dimensions in a Humanistic Knowledge Building Community. *Computers & Education*, *156*, 103943. https://doi.org/gg66d4

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (eds), *Cambridge handbook of the learning sciences* (pp. 397–417). Cambridge University Press.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. *International Journal of Learning and Media*, *1*(2), 33–53. https://doi.org/10.1162/ijlm.2009.0013

Tan, Y., Ruis, A.R., Marquart, C., Cai, Z., Knowles, M.A., Shaffer, D.W. (2023). Ordered Network Analysis. In: Damşa, C., Barany, A. (eds) *Advances in Quantitative Ethnography. ICQE 2022. Communications in Computer and Information Science*, vol 1785. Springer, Cham. https://doi-org.eproxy.lib.hku.hk/10.1007/978-3-031-31726-2_8

van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning*, *4*(3), 259–287. https://doi.org/10.1007/s11412-009-9069-5

Yang, Y., Feng, X., Zhu, G., & Xie, K. (2024). Effects and mechanisms of analytics-assisted reflective assessment in fostering undergraduates' collective epistemic agency in computer-supported collaborative inquiry. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12915

How Can Children Advance Collective Knowledge? An Exploratory Study Using the CICR Approach in Primary One Science

Xueqi Feng[1*], Jianhua Zhao[1], Yuqin Yang[2], Carol K.K. Chan[3]
[1] Center for Future Education Research, Southern University of Science and Technology
[2] Faculty of Artificial Intelligence in Education Central China Normal University
[3] Faculty of education, The University of Hong Kong
[*] fengxq@sustech.edu.cn

Abstract: *This study investigates the advancement of collective knowledge among first-grade children engaged in authentic scientific practices through the Culturation-Ideation-Conceptualization-Reflection (CICR) approach. Utilizing qualitative analyses of classroom discourse, student models, and observational data, this study explores how synergistic interactions between scientific inquiry, scientific thinking, and scientific communication can enhance both model building and idea improvement. This research demonstrates the effectiveness of the CICR approach in fostering students' engagement in authentic scientific practices and advancing collective knowledge.*
Keywords: collective knowledge advancement, authentic scientific practices, CICR, primary science

1. Introduction

In the realm of scientific research, genuine scientists continuously advance collective knowledge within their fields (Schwartz & Lederman, 2008). Authentic scientific practices, defined as the various approaches or strategies that scientists use to advance collective knowledge (NGSS Lead States, 2013), are crucial to this process. However, emulating such practices in science education presents significant challenges. Traditional pedagogical methods often overemphasize scientific procedures outlined in textbook laboratory exercises, leading students to merely follow a series of instructions to observe scientific phenomena (Osborne, 2014). These authentic scientific practices encompass not only scientific inquiry, such as problem identification, but also critical aspects of scientific thinking like modeling, and multi-level scientific communication.

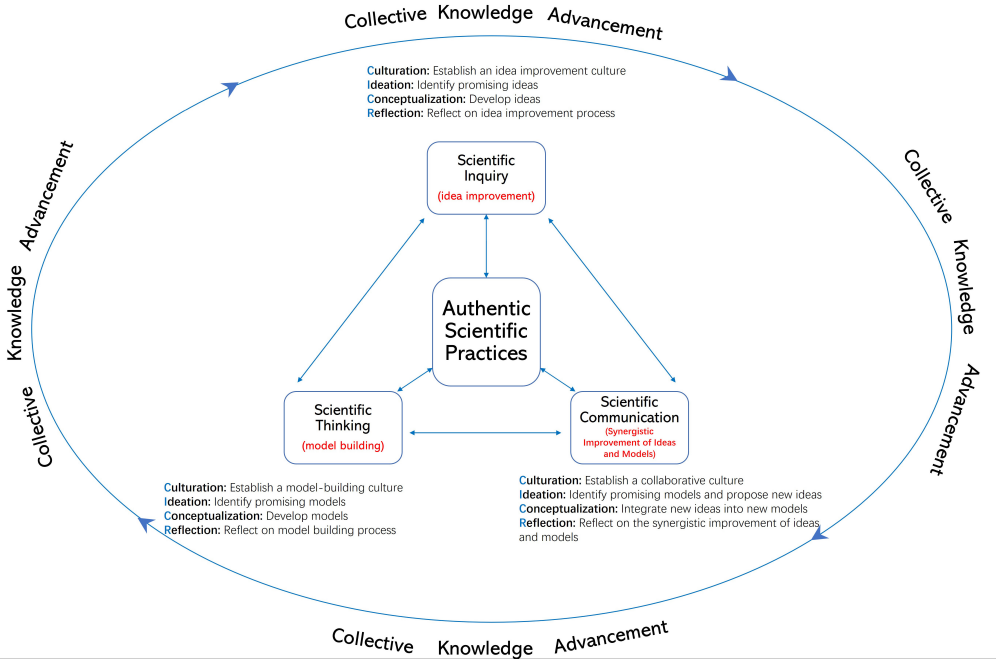
Scientific inquiry begins with the identification of problems, a process that includes hypothesis formulation, evidence gathering, and conclusion drawing, and can be viewed as a continuous activity of idea improvement (Furtak, 2006). In terms of scientific thinking, model building serves as an effective method for developing these thoughts, as it integrates and connects disparate ideas within the community and facilitates ongoing development (Clement, 2008). The processes of idea improvement and model building are inherently complementary, necessitating robust multi-level scientific communication among community members (Nielsen, 2013). Therefore, integrating scientific inquiry (the process of idea improvement), scientific thinking (the process of model building), and scientific communication (the synergistic processes of developing ideas and models) into the design of science education programs—thereby fostering student participation in authentic scientific practices and advancing collective knowledge—poses a significant challenge.

Creating authentic environments that mirror the practices of the discipline's culture is a long-standing focus of the learning sciences (Danish & Gresalfi, 2018). *Knowledge building*, which is a major research theme in the learning sciences, emphasizes the creation of scientific communities (Scardamalia & Bereiter, 2014). In knowledge-building research, numerous studies have enhanced idea improvement by fostering a culture conducive to knowledge building (*Culturation*), pinpointing promising ideas (*Ideation*), transcending collective notions (*Conceptualization*), and implementing reflective assessments (*Reflection*) (Chen et al., 2015; van Aalst & Chan, 2007; van Aalst & Truong, 2011;

Yang et al., 2020). In addition, various research groups have delved into the intricacies of model building. For instance, Lin and Chan (2018) investigated how models could facilitate the development of ideas, whereas Solima and Feng (2024) focused on how the iterative refinement of students' ideas could enhance models. Despite these contributions, there remains a lack of comprehensive exploration into how synergistic interactions between idea improvement and model building can collectively propel knowledge forward. To address this gap, the present research introduces the *Culturation-Ideation-Conceptualization-Reflection* (CICP) approach. This framework is specifically designed to amalgamate scientific inquiry, scientific thinking, and scientific communication, thereby creating a synergistic effect that promotes both model building and idea improvement, ultimately augmenting collective knowledge (See Figure 1). Research questions are: (1) How was collective knowledge advanced? And (2) How did children engage in authentic scientific practices to advance collective knowledge?

Figure 1

Conceptual Framework and Embedded Design of the CICP Approach



2. Methods

2.1. Research Context and Participants

This study was conducted in a primary school located in China, involving a cohort of 46 first-grade students, all aged between six and seven years. Over a period of ten weeks, the children participated in a weekly 45-minute session focused on a collective inquiry into *Seed Germination*. The sessions were facilitated by the first author, who is an experienced teacher specializing in knowledge-building pedagogies.

2.2. Data Source and Data Analysis

Data were collected from student models, classroom discourse videos, observations, and teacher's slides. We employed qualitative analysis, beginning with identifying the model-building processes from the students' models. We then analyzed how students' models and ideas synergistically developed, using evidence from the teacher's slides and classroom videos. Finally, we assessed the advancement of collective knowledge and explored themes related to students' engagement in authentic scientific practices.

3. Results

3.1. Collective Knowledge Advancement

In the field of *Seed Germination*, students initially posited imaginative ideas, suggesting seeds possess "teeth" for nutrient absorption and a "launcher" for emerging from the soil. As the deepened practices, students identified key anatomical parts of seeds such as the radicle, plumule, hypocotyl, and cotyledons. Afterward, the collective knowledge evolved: the radicle develops into roots, the plumule forms shoots, and the cotyledons, while acting as nutrient sources, do not regenerate after detachment. This transition from creative hypotheses to empirically supported scientific explanations illustrates the significant advancement of collective knowledge.

3.2. Children's Engagement in Authentic Scientific Practices

Theme 1: Cultivating a Culture of Collective Knowledge Advancement (Culturation)

Students demonstrated understanding of science through models, recognizing that science originates from questions, requires experimentation, and generates knowledge (Figure 2(a)). Consequently, the teacher facilitated classroom dialogues fostering a culture where students actively propose questions and articulate explanations through models.

Theme 2: Identifying and Refining Promising Ideas (Ideation)

Focusing on the topic of *Seed Germination*, students construct models to explore the question, "Why do trees grow and bear fruits?" (Figure 2(b)). Motivated by this collective curiosity, they continued to use models for further explanation. Upon analyzing the model in Figure 2(c), students noted that seed germination involves sprouting upwards and rooting downwards, prompting the collective inquiry, "How do seeds germinate and develop roots?"

Theme 3: Synergistic Improvement of Collective Ideas and Models (Conceptualization)

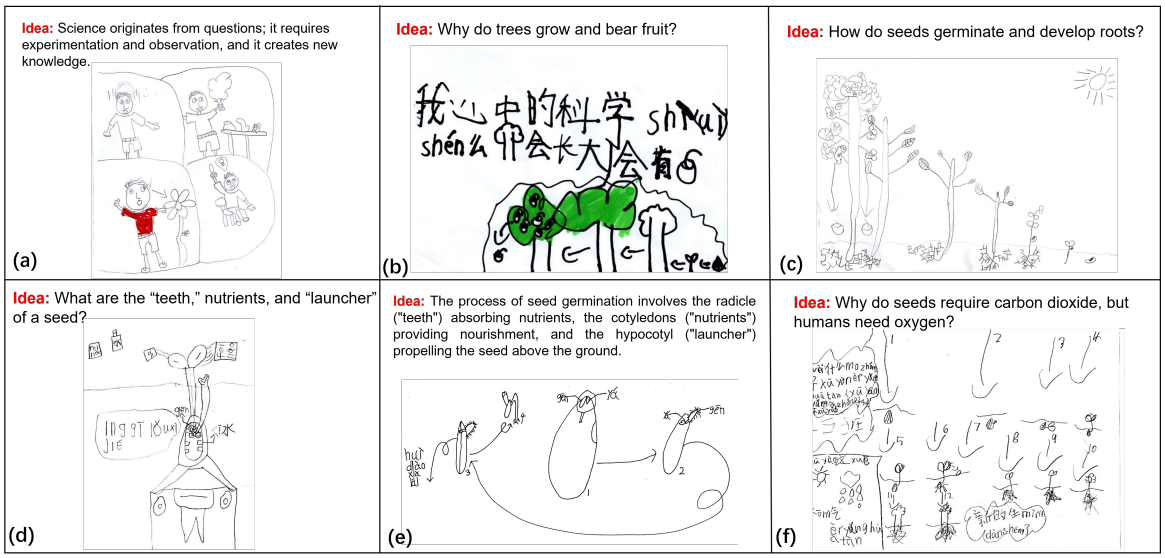
Using their creativity as shown in Figure 2(d), students envisioned the seed as having "teeth" for absorbing nutrients, "nutrients" for sustenance, and a "launcher" for emergence above the ground. By dissecting seeds, they identified key components such as the radicle, plumule, hypocotyl, and cotyledons. Through experiments and observations on seed germination detailed in Figure 2(e), students refined collective knowledge: the radicle (teeth) absorbs nutrients, the cotyledons (nutrients) provide nourishment, and the hypocotyl (launcher) drives the emergence of the seed above the soil.

Theme 4: *Reflect on the Learning Process and Formulate New Problems (Reflection)*

Students reflected on the learning process, examining how model building and idea improvement developed synergistically. They created new models and posed new research questions, such as "Why do seeds require carbon dioxide but humans need oxygen?" (see Figure 3(f)).

Figure 2

Synergistic Improvement of Collective Ideas and Models



4. Conclusion

The findings from this exploratory study underscore the efficacy of the CICR approach in advancing collective knowledge among young learners. Through engaging first-grade students in the authentic scientific practices of scientific inquiry, thinking, and communication, significant progression was observed from initial creative imaginations to detailed, scientifically validated explanations concerning *Seed Germination*. The four themes—Culturation, Ideation, Conceptualization, and Reflection—effectively facilitated this transition, fostering a dynamic learning environment where students not only engaged with scientific ideas but also actively participated in the creation and refinement of collective knowledge. This study contributes to the field of education by illustrating how integrating elements of authentic scientific practice into early science education can advance collective knowledge advancement.

Acknowledgements

This study is supported by the National Social Science Fund of China (BCA200090), Guangdong Planning Office of Philosophy and Social Science (GD23CJY19), and Shenzhen Planning Office of Educational Science (ybzz21010).

References

Chen, B., Scardamalia, M., & Bereiter, C. (2015). Advancing knowledge-building discourse through judgments of promising ideas. *International Journal of Computer-Supported Collaborative Learning*, *10*(4), 345-366.

Clement, J. J. (2008). *Creative Model Construction in Scientists and Students*. Springer.

Danish, J. A., & Gresalfi, M. (2018). Cognitive and sociocultural perspectives on learning: tensions and synergy in the learning sciences. In *International Handbook of the Learning Sciences* (pp. 34-43). Routledge.

Furtak, E. M. (2006). The problem with answers: An exploration of guided scientific inquiry teaching. *Science Education*, *90*(3), 453-467.

Lin, F., & Chan, C. K. K. (2018). Promoting elementary students' epistemology of science through computer-supported knowledge-building discourse and epistemic reflection. *International Journal of Science Education*, *40*(6), 668-687.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. The National Academies Press.

Nielsen, K. H. (2013). Scientific communication and the nature of science. *Science & Education*, *22*(9), 2067-2086.

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25*(2), 177-196.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 397-417). Cambridge University Press.

Schwartz, R., & Lederman, N. (2008). What scientists say: Scientists' views of nature of science and relation to science context. *International Journal of Science Education*, *30*(6), 727-771.

Solima, D., & Feng, X. (2024). A Knowledge Building-Modeling Approach to Scientific Inquiry. In *Learning as a cornerstone of healing, resilience, and community, International Conference of the Learning Sciences (ICLS) 2024*. International Society of the Learning Sciences.

van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences*, *16*(2), 175-220.

van Aalst, J., & Truong, M. S. (2011). Promoting knowledge creation discourse in an Asian primary five classroom: Results from an inquiry into life cycles. *International Journal of Science Education*, *33*(4), 487-515.

Yang, Y., Chen, Q., Yu, Y., Feng, X., & van Aalst, J. (2020). Collective reflective assessment for shared epistemic agency by undergraduates in knowledge building. *British Journal of Educational Technology*.

核心素养背景下知识建构促进小学生科学模型建构的设计研究

Knowledge Building Designs for Model Building in Primary Science under the Context of Core Literacy

姜羽*，吴少文
深圳市龙华区松和小学
*787690487@qq.com

【摘要】《义务教育科学课程标准（2022 年版）》明确了核心素养的四个维度：科学观念、科学思维、探究实践和态度责任，而课标中首次明确地将模型建构作为科学思维的重要组成部分。因此，提升学生的核心素养，关键在于培养他们的科学思维能力。本文将通过融合知识建构理论与数据分析技术，在学生观点发展的基础上，改进模型建构教学过程，提高模型教学的有效性，以此促进学生的思维发展，培养学生的核心素养。

【关键词】核心素养；模型建构；知识建构理论；科学思维

Abstract: *The "Compulsory Education Science Curriculum Standards (2022 Edition)" have clearly defined four dimensions of core literacy: scientific concepts, scientific thinking, inquiry practice, and attitude responsibility. Cultivating students' scientific thinking is of great significance for promoting students' core literacy. For the first time, the curriculum standards explicitly include model building as an important component of scientific thinking core literacy. This paper will, based on the development of students' ideas, integrate knowledge building theory with big data technology to improve the teaching process of model building and enhance the effectiveness of model learning. This approach aims to promote the development of students' thinking and cultivate their core literacy.*

Keywords: Knowledge Building, Elementary School Science, teaching strategies, Science literacy

1.引言

《义务教育小学科学课程标准（2022 年版）》提出，核心素养包含科学观念、科学思维、探究实践、态度责任等方面。科学思维是从科学的视角对客观事物的本质属性、内在规律及相互关系的认识方式。模型建构作为科学思维的重要内容，体现在以经验事实为基础，对客观事物进行抽象和概括，进而建构模型。

当前，很多科学教学中的“科学探究”，流于形式，只是机械僵硬的实验流程或手工制作的活动课，难以培养学生的模型建构能力，也无法有效培养学生的科学思维。模型建构教学离不开思维的培养，是培养学科学生思维的有效载体。本文将知识建构理论融入模型建构教学中，通过学生提出问题、聚焦研究问题、解决问题等一系列活动，可视化学生在《地球科学》主题学习中核心观点（思维）的变化情况，让学生通过模型建构的具体方法来理解知识的本质。学生在不断优化科学模型建构的学习中培养科学思维，促进核心素养发展。

2.知识建构理论

知识建构（Knowledge Building），旨在通过推进知识创新共同体的概念来促进教育的根本变革，这与科学家所从事的“真实的科学实践”是一致的。Scardamalia (2002)提出了基于知识创新共同体的 12 条知识原则：如“追求知识，自主自力”原则，强调学生主动学习，主理自己的学习过程，对自己的学习负责

任；“不断钻研，完善观点”原则，强调学生力臻观点完善，寻求理论建构的持续过程；“共同承担，知识无限”原则强调学生应主动承担集体认知责任，共同促进共同体知识的持续发展。为了支持知识建构理论与原则，知识建构研究者们开发了基于在线讨论的网络平台“知识论坛”（Knowledge Forum），学生们在知识论坛上能够发表观点、共同建构解释、并持续推动共同体知识的发展。

学生在知识论坛中发表观点，如何在大量观点中有效提取有价值的研究问题和确定研究方向呢？借助知识论坛平台数据分析软件（Knowledge Building Discourse Explorer），建立笔记内容中关键词之间的网络联系，形成关键词网络，促进学生在反思，不断挖掘更新的观点、建构模型，实现真正的科学实践。

3.基于知识建构理论的模型建构教学设计

学生与教师一同参与了为期五周十节课的《地球科学》知识建构共同体活动。在这个过程中，学生根据人工智能技术所呈现的共同体中不断变化的观点，对“地球科学模型”进行了“动态调整”。知识论坛视窗如图 1a 所示：学生在网络平台知识论坛（Knowledge Forum）上共建科学历程：“科学团队：团队协作→科学群体：科学交流→科学领域：科学的进步与发展”；在人工智能数据挖掘软件 KBDeX 的辅助下，持续不断地提出问题、建立解释、寻找证据、得出结论、总结他人成果、反思自身研究历程，并在原的研究基础上，开展更具有研究前景的新研究。

教学设计思路具体如图 1b 所示：围绕模型建构教学，教学设计以知识建构理论为教学模式，学生可以在知识论坛上发表观点，建构观点，以此推进集体知识进步；借助人工智能软件 KBDeX，进行数据挖掘，并以词频、关键词图谱等形式可视化学生观点聚焦的核心词汇，促进“地球科学系统”模型建构。通过建构模型解释问题，充分理解“火山、海啸地震等”自然现象背后的本质原因，从而起到促进学生科学思维形成，培养学生核心素养的目标。



图 1a:知识论坛视窗

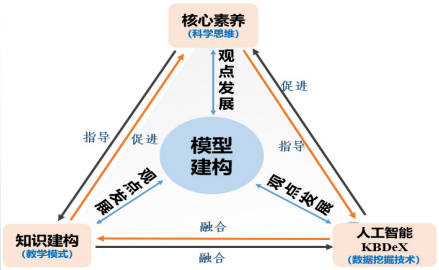


图 1b:教学设计思路

3.1 创设情境，借助支架提出问题

知识建构课堂强调学生在真实情境中提出问题。课程开始“地震、火山喷发和海啸”等真实自然现象的视频为切入点，鼓励学生使用“我的想法”、“我的问题”、“新的信息”等支架在知识论坛上提出问题，支架详见图 2a:

借助支架，学生在知识论坛上提出观点和问题，如“为什地震会频繁出现”“为什么岩浆是热的？”“为什么会有地壳运动”等 19 个问题，并通过联系日常生活经验以及从互联网上搜集的权威信息对这些问题做出解释。例如，S2 同学在回答 S1 同学“在哪些地方容易引发地震时”这个问题时，S2 同学提出“中国容易发生地震的地方：四川、河北、河南、山东、内蒙古等地”，S3 学生回复到：“为什么这些地方更容易引发地震呢？”，S4 同学回答：“应该是这个地方有很

多山，地壳运动很强烈”，在知识建构环境中，学生实现了课后的线上互动与观点表达，以进一步建构集体知识，如图 2b：



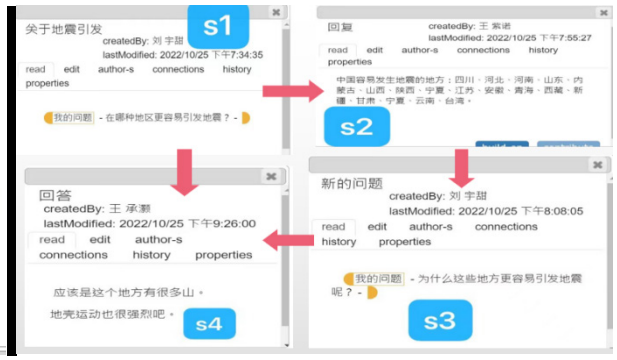


图 2a：问题支架　　图 2b：学生论坛观点呈现

3.2 识别具有前景的研究问题，通过科学模型解释问题

科学实践始于具有研究前景的研究问题。根据学生论坛的三类观点，将其分为三个研究方向：“地震”、“火山爆发”、“海啸”。学生根据感兴趣的研究问题分小组研究：通过“毛巾模拟地震运动”“橡皮泥模拟火山喷发运动”等模拟实验，再次探究了“地震”、“火山爆发”、“海啸”三个现象的成因，在现有主题的基础上不断完善观点，并建立科学模型来解释三个现象的成因，学生的模型建构历程如下图 3（以一个研究小组为例展示）。学生从对自然现象的认识，上升到对现象背后的本质认识：“地震”、“火山爆发”、“海啸”三个现象的成因与“地壳运动”有关。并且能够在科学模型中借助箭头、关键词来描述重要信息，并建立模型中元素的联系。在此过程中加深对科学知识的认识，同时培养学生模型建构的科学思维。

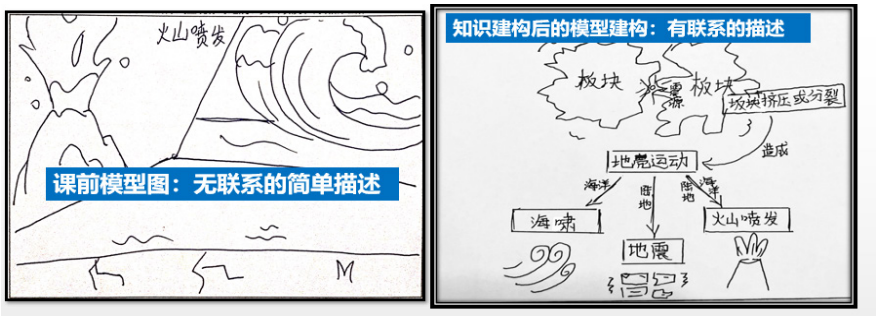


图 3：小组关于三个自然现象的模型解释

3.3 数据挖掘，改进“地球科学系统模型 ”

回顾与反思是科学实践中的重要环节。因此，学生在“科学团队的团队协作”追踪他们的研究问题，并在论坛中持续发表、建构观点，推进集体知识的升华。聚焦研究问题，帮助学生在海量的信息中，挑选有价值、能够解决问题的信息显得尤为重要。借助人工智能技术 KBDeX 数据分析软件对学生在知识论坛中的笔记进行分析。建立论坛中所有笔记的联系，并分析出在学习过程中学生关注的问题、讨论的问题之间的联系以及学生讨论高频率词汇。通过 KBDeX 软件分析的关键词网络，可以清晰的发现：学生讨论的中心词汇是：“能量”、“火山”“地震”等词，以此关键词网络作为学生重要的反思学习、深入认识“地震”、“火山爆发”、“海啸”三个现象的成因与“板块运动”“能量”的关系的重要资料，关键词网

络如图 4a。

结合 KBDeX 关键词网络与学生论坛笔记，引导学生讨论“好的模型的特点”，并多次修改模如图 4b。结果表明，通过论坛笔记与关键词图谱，学生能够有效修改模型，这对科学核心素养的培养是至关重要的。

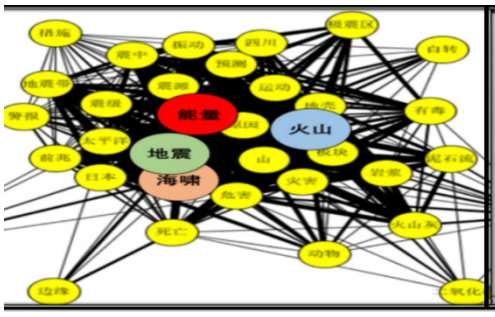


图 4a：KBDeX 关键词网络



图 4b：改进后“地球科学系统”模型

4.结论

文章基于知识论坛学生观点的变化发展，借助科学模型这个思维可视化的工具来帮助学生理解“地震、火山喷发、海啸”三个自然现象的成因。在模型建构教学中，应用大数据分析软件 KBDeX 对学生观点的处理，形成关键词图谱，让学生参与并迭代修改、解释“地震、火山喷发及海啸”成因的模型，是本教学设计中推进学生“地球科学系统”模型建构的关键。也是辅助小学科学模型建构教学过程的重要工具。在后续的教学中，教师还需要持续深入的实践，形成适合不同年级、不同学习内容的知识建构理论与模型建构融合的教学方法，为学生科学思维的发展提供抓手，从而促进学生的核心素养形成。

参考文献

Bereiter, C., Scardamalia, M., Cassells, C., & Hewitt, J. (1997). Postmodernism, knowledge building, and elementary science. *The Elementary School Journal*, *97*(4), 329-340. https://doi.org/dg498w

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. The National Academies Press. https://doi.org/drgp

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a Knowledge Society* (Vol. 97, pp. 67-98). Open Court. https://bit.ly/3013EQ8

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397-417). Cambridge University Press. https://doi.org/d9tv

贺子聪，Z. C.，& 高晓燕，X. Y. (2021). 月相相关知识及教学策略. 科教导刊,36(24), 36-39.

钱星星.(2020).小学科学的模型建构：以“我们来‘造’环形山”一课为例.教育观察,9(43),126-128.

教育部.(2022).义务教育小学科学课程标准.https://www.moe.gov.cn/

基于会话代理促进学习者反思性思维的框架构建研究

Constructing a Reflective Thinking Promotion Framework Based on Conversational Agents

缪静敏 * 林婉芸
广州大学 教育学院（师范学院）
*mjm@pku.edu.cn

【摘要】 近年来，生成式人工智能的迅速发展显著提高了会话代理的能力，并扩大了其在教育领域的应用。反思性思维是一种旨在解决问题或得出结论的，主动、持续、有序与深入的思考方式，对提升学生的问题解决能力与批判性思维有重要作用。会话代理能力的提升为基于人机对话来促进学习者的反思性思维提供了契机。为了实现上述目标，本文构建了基于会话代理促进学习者反思性思维的框架(RTPF-CA)。该框架根植于杜威的反思性思维理论，系统呈现了会话代理促进学习者反思性思维各个阶段的路径与策略。最后，本文结合一个探索性的案例探讨了应用上述框架的方式、挑战及解决策略。

【关键词】 会话代理；反思性思维；生成式人工智能；高阶思维；对话学习

Abstract: *In recent years, the rapid development of generative artificial intelligence has significantly improved the capabilities of conversational agents and expanded their application in education. Reflective thinking is a proactive, continuous, organized, and in-depth mode of thinking aimed at solving problems or reaching conclusions. Developing reflective thinking is crucial for improving students' problem-solving skills and enhancing their critical thinking. The improved capabilities of conversational agents have opened opportunities to enhance learners' reflective thinking via human-computer dialogue. In pursuit of this objective, this paper introduces a systematic framework named the Reflective Thinking Promotion Framework Based on Conversational Agents (RTPF-CA). This framework, grounded in Dewey's theory of reflective thinking, systematically outlines how conversational agents can facilitate each stage of learners' reflective thinking. Finally, the paper illustrates the approach to implementing RTPF-CA through an exploratory case study, while also discussing associated challenges and their solutions.*

Keywords: Conversational Agent, Reflective Thinking, Generative Artificial Intelligence, Higher Order Thinking, Dialogue-based Learning

1. 引言

反思性思维（Reflective Thinking）是一种旨在解决问题和厘清困惑的思维方式，要求个体能够围绕问题情境进行主动、连续与系统的思考，深入探究问题的本质和解决问题的可能途径，具有“深思熟虑”的特征。在教育实践中，反思性思维在促进学习者深层学习，提升问题解决能力，发展批判性思维和提升学业表现方面发挥着重要作用（Chamdani et al.，2022；Ghanizadeh, 2017; Nhat et al., 2023; Kholid et al., 2022）。这也使得如何培养学生的反思性思维成为实践者与研究者关注的重点。围绕上述问题，研究人员开发了多种促进反思性思维的策略，包括撰写反思性日志，参与反思性对话，获取同伴反馈，开展探究性学习等（Tsingos et al., 2015）。

尽管上述策略在促进反思性思维方面取得成效，但由于缺乏支持性的环境，一些对时间与人力成本有较高要求的策略很难在实践中大范围的实施。近年来，生成式人工智能支持下的会话代理（Conversational Agent, CA）的发展，为促进学习者的反思性思维提供了的新的思路。学习者可以通过与会话代理对话获得促进反思性思维的必要指导。为此，本文尝试构建了基于会话代理促进反思性思维的系统框架，并基于一个探究性的案例探讨了上述框架应用于实践时可以采取的方式，以及面临的挑战与应对策略。

2. 反思性思维的过程与特征

有关反思性思维的系统论述最早来自杜威（2005）。杜威指出，人具有多种思维方式，反思性思维是其中比较好的思维方式。反思性思维旨在得出结论，是一种主动的、持续的、有序的、基于证据推理的思考过程。它既不是想入非非的零散意识流，也不是对问题不假思索的直接反应，它要求个体对问题进行审慎的思考，以获得对问题成熟与周全的理解。从思维过程来看，反思性思维包含暗示、理智化、假设、推理和用行动验证这五个阶段。其中，具有反思性思维的个体会在暗示阶段对各种可能的暗示保持开放与敏感，个体已有的丰富经验能够为多种暗示的出现提供可能。在理智化阶段，个体会对初步的暗示进行更深入的思考，以清晰地界定问题，形成更加具体与有形的思考对象。在假设阶段，个体提出多个假设与方案作为对问题的回应，这些假设与方案基于已有信息和逻辑推理得来，对不同的可能性开放，并且是可以被测试和验证的。在推理阶段，个体需要更深入分析此前阶段提出假设与方案的内部逻辑结构，并考虑其可能导致的具体结果。相关问题包括假设基础是否合理，是否考虑了多种可能，逻辑是否连贯，是否有证据支持，是否需要调整假设等等。在行动验证阶段，个体需要通过采取实际行动来测试假设与方案的有效性。如何准确地记录结果，反思行动效果，以及基于获得的新知来调整假设与方案是这一阶段的关键。此外，验证过程可能会指向新的暗示并引发新一轮的反思性思维循环。总而言之，杜威对反思性思维的描述完整呈现了反思性思维的主要过程，并强调了反思性思维区别于其他思维的关键特质。

除了杜威之外，其他学者也从不同视角对反思性思维进行了深入的讨论。其中，舍恩从专业实践特征的角度强调了反思性思维具有的动态特性（2007）。面对专业实践中不确定的、独特的，甚至是存在冲突的情境，个体需要通过即时反思与动态调整以更好地在“低洼湿地”中解决问题。由此，反思不仅包括对行动反思，也包括行动中的反思，即反思不仅仅是一个事后的评估过程，还是一个在实践中即时发生、能够促进学习和适应的动态过程。其次，Mezirow（1991）从成人教育和变革性学习的视角出发，强调了反思性思维的批判性特质。他认为，反思性思维可以促进个体自我意识的增强，通过批判性地审视自己的假设和信念的准确性，促使个体认知框架的改变。这种过程能够使人获得新的理解世界的视角，进而促进个人和职业层面的深刻成长。

尽管不同学者对反思性思维的探讨各有侧重，但他们对于反思性思维的描述仍然展现出了许多共通之处。其中被普遍认可观点包括：1）反思性思维来源于对疑难情境的思考，旨在解决问题或得出结论；2）反思性思维是一个持续与动态的过程，能够提高个体行为的适应性和灵活性；3）反思性思维具有批判性，为避免潜在偏见与盲点，个体需要批判性地审视已有假设、信念与行动模式；4） 反思性思维致力于达成对事物深刻的理解，它要求个体深入探究和掌握更加根本的原理与概念。

3. 反思性思维与社会互动

反思性思维虽然发生在个体层面，但是社会互动对激发反思性思维的过程具有关键影响。杜威、舍恩与麦基罗等人在讨论反思性思维的特质时，都强调了社会互动在促进反思性思维

过程中的重要作用。杜威（2005）特别强调了交流在反思和学习过程中的作用，他认为教育应该鼓励学生在真实的社会情境中学习，通过实际的社会互动来促进思维的发展。舍恩（2009）则在他的反思性实践理论中提出，专业人士通过在实践中的反思行动和对行动的反思来增强其职业能力，而这一过程往往需要与同行或导师的交流和合作。Mezirow（1991）的变革性学习理论进一步强调了对话和批判性反思在成人教育中的重要性。他认为，个体通过反思自我和社会的预设，可以促进认知结构的根本变革，而这种变革往往需要在社会互动和对话中才能实现。

社会互动对反思性思维的影响在知识共建的相关研究中也有所体现。在这些研究中，知识共建不仅会激发与支持反思性思维，反思性思维也会反过来促进知识共建达到更高层次。Garrison 等人（2000）构建的探究性共同体（Community of Inquiry , CoI）框架正是以杜威关于实践探究与反思性思维的论述作为理论根基。该框架包含认知临场感、社会临场感、教学临场感三个要素及各要素之间的关系构成。其中，认知临场感阐述了知识共建的认知过程，涵盖了触发事件、探索、整合和解决四个关键阶段。在触发事件阶段，特定的问题或挑战激发了学习者的兴趣和好奇心。进入探索阶段，学习者通过广泛的讨论、信息和资源的收集，深化对复杂概念和情境的理解。整合阶段要求学习者融合获取的信息、不同观点和资源，以形成一种更加连贯和整体的认识。解决阶段则是将前期获得的知识和洞察应用到具体的问题或情境中，努力针对最初的问题或挑战找到解决方案。此外，Guanawardena 等人（1997）提供的社会互动分析模型（Interaction Analysis Model， IAM），虽然没有直接指向反思性思维，但其中所描述的知识共建过程与 Garrison 等人存在相似之处，能够在某种程度上体现反思性思维特征。该模型包含五个阶段：1）信息的分享/比较；2）发现和探索不一致或不协调的观念、概念或陈述；3）意义的协商/知识共同建构；4）测试与调整共同建构或整合的方案。5）新建构意义的达成/应用。在上述知识共建过程中，学习者也需要识别各类问题、疑难、困惑与矛盾，对各类信息与观点保持开放，通过共同推理与协商求解问题解决方案，并最终通过知识应用与验证获得清晰一贯的理解。总而言之，上述模型框架均能够体现反思性思维与社会交互的促进关系。

4. 生成式人工智能支持的会话代理

会话代理是一种智能系统，它可以通过自然语言与人类用户互动来完成特定任务，典型会话代理如聊天机器人、虚拟个人助理、社交机器人等。在教育领域，会话代理取得了广泛的关注，并被视为支持个性化学习与提高教学效率的关键工具。以会话代理的代表性应用聊天机器人为例，自 19 世纪 60 年代世界上第一款聊天机器人 Eliza 诞生以来，聊天机器人系统经历了多次技术更迭，其智能化程度越来越高，并在教育领域得到广泛应用。Okonkwo 等人（2021）在其综述研究中指出，聊天机器人能够为教育提供多方面的支持，特别是在教学与学习方面，聊天机器人可以用于提供教学内容、回答学生问题、提供个性化帮助、解决疑问，以及提供标准化测试的准备材料等。Labadze（2023）在回顾 67 篇相关文献后指出，对于学生而言，聊天机器人能够提供作业和学习辅导，丰富个性化学习体验，并促进包括写作、问题解决、小组讨论在内的各类技能发展；对教师来说，聊天机器人不仅能够作为辅助工具节省时间，还能够丰富和完善教学方法。由此可见，以聊天机器人为代表的会话代理已经在教育领域得到了多样化的应用。

近年来，随着生成式人工智能技术的迅速发展，会话代理在理解和处理自然语言方面取得了显著突破。以 ChatGPT 为典型代表的会话代理系统能够通过较为准确、丰富与多样的回答，实现与用户之间的自然和流畅的对话。底层技术的发展极大拓展了聊天机器人在教育领域应用的深度与广度，吸引了众多研究者和实践者的关注。特别是在角色扮演和对话学习这

一类的复杂任务中，会话代理能够通过扮演不同角色，在人机协同知识构建过程中激发、引导乃至促进学习者的思考，并最终实现更高层次的认知目标。Chen（2023）等人通过基准测试系统评估了角色扮演会话代理的社会交互能力。评估结果发现，虽然角色扮演会话代理在群体水平上的社交互动能力仍有待提高，但是在个体水平上表现出了令人满意的性能，包括对角色描述的自我意识、对环境的情感感知和长期对话记忆。通过扮演各种各样的角色，会话代理能够在与学习者共同进行知识构建的过程中发挥激发、引导或促进的作用，帮助学习者更有效地完成任务。Webb（2023）就在医学教育项目中利用 ChatGPT 创建情境案例和模拟急诊室患者的角色，辅助医学生发展他们的沟通技巧。结果显示，会话代理可以创建现实的任务场景，模拟医生-病人的交流，并评估个别用户的表现。总之，会话代理性能的提升为其在教育场景中扮演具备复杂对话能力的特定角色奠定了基础。这也使得社会性互动和知识构建的过程不仅能在人与人之间发生，也能在人与机器之间进行。

与此同时，定制智能体平台的出现（如 OpenAI 的 GPTs，智谱清言智能体），大大降低了构建专属会话代理的技术门槛，开发者可以直接通过输入精心设计的问题或指令来创建专属的会话代理来完成特定类型的任务。会话代理能力的提升与定制门槛的降低，激发了研究人员面向各类教育场景开发聊天机器人与变革学习过程的热情。在上述背景下，如何基于会话代理构建专属聊天机器人，并基于上述代理开展高质量的对话学习成为当下研究关注的热点。

5. 基于会话代理的反思性思维促进路径

虽然会话代理对话能力的提升为促进学习者的反思性思维提供了技术条件，但要提高其促进效果，并实现更高层次的认知目标——如促进学习者的反思性思维，仍然需要对代理功能及其使用方式进行系统化的设计。为此，本文基于杜威有关反思性思维过程的描述，尝试构建了基于会话代理的反思性思维促进框架（Reflective Thinking Promotion Framework Based on Conversational Agents，RTPF-CA）。具体而言，在社会交互过程中，会话代理可以通过引导五类反思来促进学习者的反思性思维，包括对观察的反思、对问题的反思、对视角的反思、对论据的反思和对行动的反思。其中，每一类反思都对应着反思性思维的特定阶段，具体如图 1 所示。

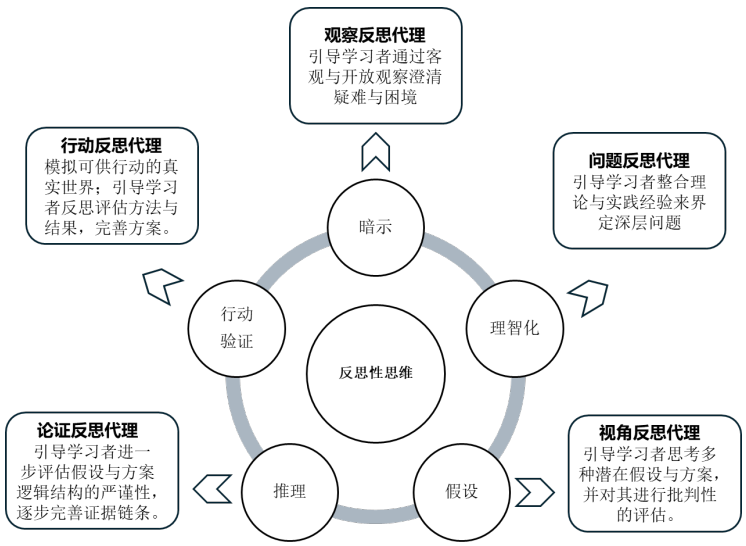


图 1 基于会话代理的反思性思维促进框架（RTPF-CA）

（1）对观察的反思

过于匆忙地接受最先出现的暗示，而没有充分探究事件的全貌，是个体在反思性思维的暗示阶段经常会出现的问题。为此，会话代理可以通过引导学习者从多个角度观察眼前的景

象，获取更加全面的信息，并思考各种可能的情况，以实现开放性的探索。换言之，会话代理在该阶段关注的问题在于学习者观察到了什么，以及如何通过引导性问题与补充性信息引导学习者观察到多种可能性。该阶段的人机对话的主题在于学习者如何根据观察来澄清困惑或冲突现象。

（2）对问题的反思

在反思性思维的理智化阶段，学习者需深入理解现象本质并清晰界定问题，这是解决问题的基础。该阶段面临的主要挑战则是问题界定不够清晰、具体或过于狭窄。理论知识与实际经验不足，以及先入为主的思维模式是导致上述现象出现的重要原因。为此，会话代理可以通过提供概念性和经验性的支持，引导学习者触及核心概念与原理，从而更有效地识别问题。这种方法不仅帮助他们建立坚实的知识基础，还促进了对实际情境的理解，使他们能够更准确地界定和分析问题。该阶段人机对话的主题在于学习者如何基于观察到的现象来界定问题的类型。

（3）对视角的反思

在反思性思维的假设阶段，最常见的问题是过早地固定在某个假设上而忽视了其他假设与方案的可能性，基于不完全或者错误的信息生成假设与方案，以及提供了不可测试与验证的假设与方案。考虑到每个假设与方案的背后都隐藏着看待问题的特定视角，会话代理可以通过多视角的提问，让学习者一方面关注到存在多种类型的假设与方案，另一方面意识到需要对假设进行批判性的评估，以确保在假设与方案有充分且准确的信息作为支持，并且可以被测试和验证。该阶段人机对话的主题在于学习者基于何种视角来分析问题，并提出可以被测试的假设。

（4）对论证的反思

在推理阶段，个体需要进一步评估此前阶段提出假设与方案的内部逻辑结构。在该阶段常见的问题是假设基础不够合理、忽视替代性解释或假设、逻辑推理不够连贯，以及缺乏证据支持等。为此，会话代理可以通过引导学习者思考替代性的假设与方案，识别论证中存在的逻辑错误，寻求支持假设更充分的论据（如数据、研究或者案例），以及系统评估其假设论证过程，从而增强论证的深度与广度，并形成完善的证据链条。该阶段人机对话的主题在于学习者如何进一步论证与优化其提出的假设与方案。

（5）对行动的反思

在行动验证阶段，个体需要在行动中验证或者调整假设。在该阶段的常见问题是，观察方法存在问题，数据收集不够充分，行动过程未能进行有效的监控与调整，以及对行动效果评估缺乏批判性。此外，在课堂教学场景下，学习者还面临的另一问题是行动验证阶段往往因为条件限制很难实施，对行动的反思也因此经常遭到忽视。为此，会话代理一方面可以模拟真实的问题情境，为学习者提供实施方案、收集数据与评估效果的机会；另一方面也可以引导学习者反思与调整观察方法，合理评估行动效果并调整行动方案。该阶段人机会话的主题在于学习者是否能够灵活与有效地实施行动方案，并充分收集数据来合理评估方案实施效果，从中调整与完善方案。

6. 基于会话代理的反思性思维促进框架应用案例

基于 RTPF-CA 框架，研究选取本科生毕业论文指导作为案例场景，探索其应用方式。选取本科生毕业论文指导作为案例场景的原因，主要包括以下两个方面：其一，本科生毕业论文指导的重点，本身就在于通过复杂问题求解过程来训练学生的学术思维能力，其中就包含了反思性思维的成分；其二，本科生毕业论文指导还面临着规模化指导的困境。本科生与研究生相比，学生规模大，研究经验少，指导周期短。如何在上述情况下，为本科毕业生提供

充分的指导，使其能够高质量的达到学术训练的要求，成为了不少高校教师面临的难题。综上，研究关注如何基于 RTPF-CA 框架，通过提示工程开发能够支持本科生毕业论文指导的会话代理。其中，研究选取 Open AI 的 GPTs 功能来实现提示工程的开发。

目前，研究还处于初期探索阶段，关注的问题在于如何应用 RTPF-CA 框架中对观察的反思构建会话代理，引导学习者在研究初期澄清其关注的实践现象，从中挖掘潜在的问题。主要的做法是，让会话代理向学习者按照特定序列向学习者提出问题，并在学生完成所有的回答之后，根据对完整对话的评估，向学习者提供反馈建议。在提示工程开发方面，研究采取了如图 2 所示的路径。

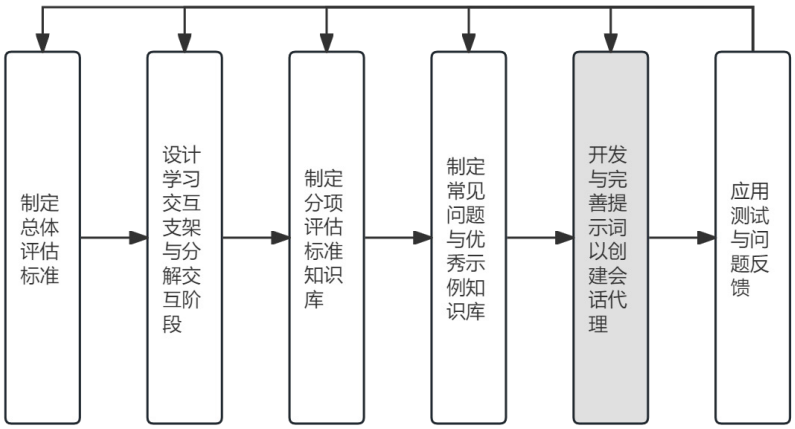


图 2 会话代理提示工程开发流程

在上述路径中，研究首先明确了当前阶段学生表现的评估标准，为后续设计与开发提供总体指引。在此基础上，研究进一步设计交互支架，明确学生与会话代理的交互方式。测试过程中发现，为了能够在现有的技术范围内更好地促进学生的思维过程，需要分解会话代理与学生的交互阶段，通过有针对性地优化每一个阶段的促进效果来提升整体效果。认知过程与交互阶段的分解也带来了评估标准的分解。研究针对每个阶段制定的明确的评估标准，并形成标准知识库供会话代理在评估与反馈时使用。为了能够提升会话代理的效果，使其能够更好的实现教学意图，在提供评估标准知识库之后，研究进一步构建了常见问题与优秀示例知识库，帮助会话代理更有针对性地识别学生特定阶段出现的问题。在确定交互方式、评估标准知识库和典型案例知识库之后，研究制定了提示词方案，并基于提示词构建了会话代理。此后，研究通过小规模测试来评估会话代理的使用效果，并根据测试反馈的问题调整了提示工程开发相应阶段的产出，以此实现提示词的优化。

在此过程中，RTPF-CA 框架应用于指导评估标准的制定、交互方式的设计以及各类示例的创建。其中，在交互方式的设计上，研究确定了以问题序列来引导学习者参与的方法。有关问题类型与问题序列的设计，其重点在于如何让学习者通过开放、持续与深入的思考，逐步理清自己关注的研究现象。问题类型及序列如下表所示。

表 1 问题类型与序列

序列号	问题类型
1	有什么现象引起了你的好奇或困惑？这个现象的疑难之处在于哪里？
2	为什么你会关注这个现象？ 有什么特别的原因吗？
3	你认为你的个人经验会如何影响你对这个现象的探索？
4	有其他人观察到类似的现象吗？他们是如何解释这种现象的？他们的解释还存在什么不足？

5	你打算从哪个问题出发来探索这个现象，以弥补现有解释的不足？
6	你觉得你探索这个现象的价值和意义是什么？
7	你认为你还需要补充哪方面的信息来更好地描述清楚这个现象？

根据上述路径，研究基于开发并测试了名为 Thesis Mentor 的会话代理。学习者与 Thesis Mentor 对话示例如图 3 所示。在该对话示例中，会话代理向学生抛出了第一个问题，并通过提出评估标准来明确对答案的期待。在学生提供答案之后，会话代理识别学生答案中存在的问题，并提供了针对性的反馈与建议。同时，会话代理还允许学生就反馈与建议不理解的地方进一步提问，以确保学生能够清晰地理解。当学生表示能够理解之后，会话代理就会向其抛出下一个问题。

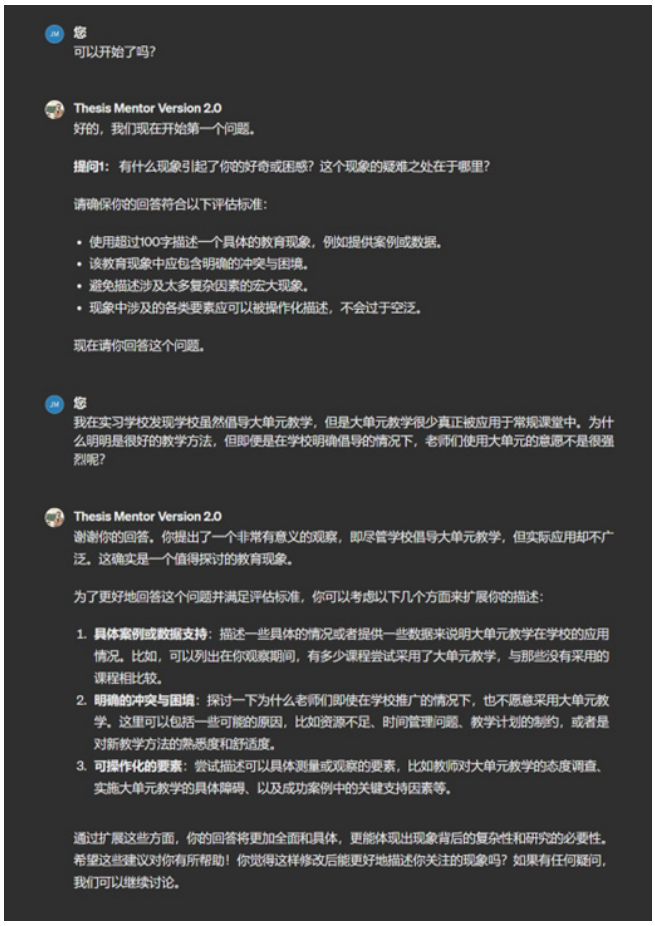


图 3　Thesis Mentor 与学习者交互过程示例

目前，Thesis Mentor 的功能主要为引导学习者澄清自身所关注的研究现象，对 RTPF-CA 框架的应用也主要聚焦于对观察的反思部分。随着后续研究的推进，Thesis Mentor 的功能将会逐渐完善，对 RTPF-CA 框架的应用也将更加全面。

7. 应用挑战与对策

在基于提示工程方法，将 RTPF-CA 框架整合到 Thesis Mentor 系统中时，研究人员面临的挑战主要体现在三个方面。其一，复杂认知过程的开放性和不确定性，与提示工程对执行过程与标准的确定性之间产生了冲突。为此，研究采用的应对策略是对复杂认知过程进行了适当拆解，形成一系列的交互阶段，在 Thesis Mentor 系统中表现为会话代理提出的问题类型及其序列。其二，会话代理所基于的底层模型有其自身的知识与经验储备，这些储备既有可能促进会话代理的表现，也有可能干扰其偏离系统开发者本身的意图。就像与他人沟通时，

对方的已有经验会影响其对他人观点的理解。为此，即便对交互过程进行了分解，仍然需要对每一个子过程进行精细化的设计，在 Thesis Mentor 系统中表现为提供分项评估标准、常见问题示例与优秀示例。其三，学习者缺乏与会话代理交互的技巧会抑制会话代理的输出质量，使其作用无法得到有效发挥。就像在日常毕业论文场景的指导下，学生缺乏与导师的沟通技巧，也会导致沟通障碍。为此，可以通过交互方式的设计，为学习者提供相关支架，引导其更好地与会话代理沟通，在 Thesis Mentor 系统中具体表现为，会话代理在提问时就向学习者明确回答的预期形式，并在提供反馈后引导学习者就尚未理解的反馈与其进一步沟通。

此外，Thesis Mentor 目前面向的是本科生毕业论文指导场景，侧重于思维能力的训练，研究领域知识库的缺乏对功能的影响目前还没有特别突出。但如果将该系统应用于研究生毕业论文指导时，缺乏足够的研究领域知识库，可能会限制系统功能的发挥。正如人类导师也会因为缺乏相关领域的知识而无法很好地指导学生。构建相应的知识库或调用外部的插件可能是解决问题的方式，这也是研究后续探索的方向。

参考文献

约翰·杜威(2005)。 **我们怎样思维·经验与教育**（姜文闵译)。北京：人民教育出版社。

唐纳德·A·舍恩(2007)。**反映的实践者：专业工作者如何在行动中思考**（夏林清译）。北京：教育科学出版社。

Chen, H., Chen, H., Yan, M., Xu, W., Gao, X., Shen, W., Quan, X., Li, C., Zhang, J., Huang, F., & Zhou, J. (2024). RoleInteract: Evaluating the Social Interaction of Role-Playing Agents. *arXiv preprint arXiv:2403.13679.*

Chamdani, M., Yusuf, F. A., Salimi, M., & Fajari, L. E. W. (2022). Meta-Analysis Study: The Relationship between Reflective Thinking and Learning Achievement. Journal on Efficiency and Responsibility in Education and Science, 15(3), 181-188.

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, 74, 101-114.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17(4),* 397–431.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2(2-3)*, 87-105.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers & Education: Artificial Intelligence, 2(100033)*. https://doi.org/10.1016/j.caeai.2021.100033

Kholid, M. N., Sa'Dijah, C., Hidayanto, E., & Permadi, H. (2022). Students' reflective thinking pattern changes and characteristics of problem solving. Reflective Practice, 23(3), 319-341.

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20, 56. https://doi.org/10.1186/s41239-023-00426-1

Tsingos, C., Bosnic-Anticevich, S., & Smith, L. (2015). Learning styles and approaches: can reflective strategies encourage deep learning?. *Currents in Pharmacy Teaching and Learning*, 7(4), 492-504.

Webb, J. J. (2023). Proof of Concept: Using ChatGPT to Teach Emergency Physicians How to Break Bad News. *Cureus*, 15(5), e38755. https://doi.org/10.7759/cureus.38755

Nhat, T. N. M., & Van Le, T. (2023). Effects of Reflective Thinking on Deep Learning in Theoretical Linguistics Classes. *International Journal of Instruction*, 16(4).

Empowering Teachers to be Learning Designers: A Transformative Model to Teacher Professional Learning

Kun Liu[1], Jianhua Zhao[2], Nancy Law[3*]

[1,2] Department of Education, The University of Hong Kong

[3] Southern University of Science and Technology

nlaw@hku.hk

Abstract: This paper discusses challenges in promoting pedagogical innovation among university teachers and proposed a model and explore of its impact on teachers' understanding of learning design, design practice. As part of an ongoing Design-based research project, preliminary analysis suggests that the process of teachers' change is characterized by an ongoing and sustained development cycle, driven by evolving beliefs and enhanced understanding, reinforced through the iterative cycles of inquiries. The iterative nature of this cycle fosters a dynamic environment that supports continuous growth, enabling teachers to become more effective designers of learning experiences and environments.

Keywords: Learning Design, Pedagogical Innovation, Higher Education, Learning Design Tringle

1. Introduction

Currently, higher education has faces challenges, such as the diversifying student needs and expectations, rising expectations about graduate capabilities, intensification of pressures on teaching staff, and accelerating technological innovation (Goodyear, 2015; Rapanta et al., 2020). In response to this， there is a call and growing trend for higher-education teaching innovation and widespread implementation of technology-enhanced learning (TEL) and student-centered approaches to cater to diversity in learning needs and empower learners' capacities for the 21st century., This also requires teachers to reconceptualize themselves as learning designers - *"not to transmit knowledge to a passive recipient, but to structure the learner's engagement with the knowledge, practicing the high-level cognitive skills that enable them to make that knowledge their own" (Laurillard, 2008, p. 527)*.

Learning Design is an emerging field which involves teachers as professional practitioners, learning sciences researchers, as well as instructional designers field aim to responding to the call for teachers to re-define their role as learning designer (Law, 2017) and create learning environments and learning experiences aligned to pedagogical approaches and learning objectives to improve practice (Mor, Craft & Maina, 2015). However, it's challenging to support in-service higher education teachers to re-conceptualize their role and in transitioning from traditional, content-driven instruction to designing learning experiences and learning environments prepares for 21st-century needs.

First, pedagogical innovation challenges teachers' long-lasting values and beliefs about how learning take place, which represents a significant challenge to teachers with deep-rooted educational philosophies and the status of existing design practice (Law, 2008). Higher education teachers are disciplinary experts with a strong emphasis on research, with limited trainings and experiences in development of pedagogical skills. As a result, many educators mirror their personal educational experiences to approach learning design, mostly lecturing rather than student-centered pedagogies. To facilitate this shift, teachers need to be provided with immersive experiences to examine and question the underpinning beliefs that guide their teaching and to develop a comprehensive understanding of learning design principles.

While transforming beliefs is a vital initial step, it does not automatically translate into changes in educational practice. Operationalizing a pedagogical philosophy into concert design is a complex process, which involves design considerations at multilevel and interactive processes which require back and forth alignment check among design the different design components (Law & Liang, 2020). Ensuring the alignment and coherence across various levels and elements of course design (e.g. learning objectives (LOs), learning activities, and assessments) is critical for the success of LD strategies, but this high level of complex and reflective practice requires advanced metacognitive skills, which also presents a considerable challenge (Law, 2008). To achieve this, teachers need a robust conceptual tool to guide the multi-level decision-making process in the design of curriculum units and practical guidance to help them critically

analyze and refine their designs to ensure that the overall flow and structure of these design are aligned with appropriate pedagogical design principles to achieve the targeted learning outcome goals (Chen, Liang &Law, 2022).

Other challenges for many higher education teachers are that they work individually in teaching practices, with limited opportunities for sharing and collaboration. Implementing innovation can evoke fears and anxieties, potentially hampering a teacher's confidence and willingness to implement new pedagogy. So teachers need a supportive community for idea exchange and foster continuous collective growth and sustained innovation (Sancar, Atal, & Deryakulu, 2021).

Considering these challenges, there is a pressing need for a transformative professional learning model that support in-service higher education teachers to be learning designers for pedagogical innovation (Bennett et al., 2022). This model should not only facilitate the creation of learning environments and experiences to promote shifts in individual understandings of learning design but also provides guidance to support teachers in systematically operating their evolving understandings through iterative cycles of inquiry to improve practice and sustaining pedagogical innovation.

The objective of this study is to propose and implement a transformative teacher professional learning model and investigate the extent to which this model can support higher education teachers in shifting their understandings of learning design and improving their course redesign practices for effective pedagogical innovation.

2. Methodology

This study was conducted in at a mainland China university. This paper is part of an ongoing Design Based Project where in-service teachers interested in and had experiences redesigning courses were purposively recruited. An in-depth DBR methodology to explore how teachers engage in inquiry into their own learning designs and implementation processes. DBR's iterative cycles of design, implementation, analysis, and redesign particularly fit the objective of this study. Participating teachers were selected from the former study aiming at understanding university teachers' understandings of learning design and how they go about design practices.

The core design elements of the model is guided by two design principles: 1) Learner-centered collaborative inquiry of practice 2) Iterative cycles of inquiry supported by the Learning Design Tringle (LDT) (Law & Liang, 2020) with going support. The design principles guide the design of core intervention elements, including collaborative workshops and the LDT to support individual teachers' inquiry process. The Learning Design Triangle is a framework to guide the multi-level decision-making process in the design of curriculum units. It ensures that the overall flow and structure of these units adhere to appropriate pedagogical design principles and the disciplinary area of the curriculum to achieve the targeted learning outcome goals. Central to the LDT are three pivotal elements: Learning Outcomes, Disciplinary Practices and Pedagogical Approaches. The LDT mandates a structured, systematic approach to course design, starting with explicit learning outcomes, linking them to genuine disciplinary practices, choosing to fit pedagogical strategies, and mapping out a comprehensive educational journey for students. Law and Liang（2020）proposed a 7-step procedural guide for transforming abstract pedagogical thinking into concrete learning contexts:

Figure 1. Virtual representations of the 7-step LDT framework

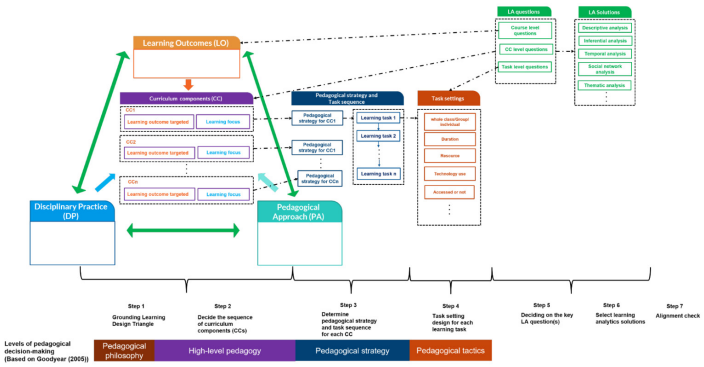


Figure 3. A visual representation of the 7-step LDT framework for guiding the LD decision making process.

1. Learner-Centered Collaborative Workshop Designed Under the Guidance of Learning Design Principles.

The collaborative workshops provide a platform for teachers with shared vision and goals to improve course design through engaging in collective problem-solving, collective inquiry into innovative pedagogical approaches, and sharing of experiences with a professional network of innovators. The intervention includes a series of workshops, each with specific themes and objectives. The first workshop was designed by the research team based on the teachers' interview data collected in the previous study. The other two workshop themes are proposed by the teachers themselves (Workshop 2: What is student-centered learning; Workshop 3: Visualizing thinking process). The timing of these workshops is planned monthly as advised by the teachers to allow for reflection and implementation between sessions. Due to pandemic, the monthly workshop shifts to small groups.

2. Individual Teacher's iterative cycles of inquiry into design practice supported by the LDT

We offered all the teachers additional support after participating the workshop 1. Among the teachers, two teachers showed significant interest in enhancing their courses and actively sought individual support for applying the learning design theory and make use of LDT to analyze, reflect and refine their design. Ongoing support (co-planning, lesson observations and post-class debriefing sessions) was available upon teachers' request, which aims to support them to reflect their teaching strategies and inquiry into their teaching and improve their courses.

All the workshops and individual follow-up and support sessions are also audio recorded and transcribed. Course materials and lesson observation notes are also collected. The data was analyzed using content analysis.

3. Preliminary results

The analysis of the results indicates that the initial workshop promotes teachers to challenge their underlying beliefs and develop understandings of learning design. Initially, teachers mostly adopted a teacher-centered, content-driven approach focused primarily on lecturing and knowledge transmission. However, after participating in the first workshop, they began to adopt a new perspective on learning design and start to consider students.

Teacher C: This training was incredibly useful. Having participated in various trainings before, from sessions led by top experts to local workshops, this one felt different—it wasn't just a training but a collaborative creation. Such cases are beneficial for any grade level. The core value he presented was in demonstrating a complete lesson or a series of lessons, allowing us to first experience the pitfalls before understanding the design. This approach is immensely instructive for us. The sharing of this case study significantly guides our overall educational design and the specific teaching activities that follow.

Teacher A：The most significant change I've made recently is to start prioritizing students' understanding. Previously, I taught based on my own ideas, assuming that if I thought they could understand, then they should. However, I've come to realize that they might not have understood at all. Now, I'm gradually adjusting my lesson plans based on students' feedback.

Teacher B：I've started to reflect on what I really want my students to achieve. It's not just about what I want to teach them, but what they need to learn. Based on this realization, I've decided to completely overhaul the course.

This shift encourages teachers to use LDT to critically examine their existing design practices, questioning their underlying beliefs, and intentionally exploring pedagogical experiments aimed at transitioning towards a student-centered approach.

Teacher B: After my next class, I'll discuss with you how it went. Next week, I'll explain using this model. I'll share how I designed the course, what my model is, and what my objectives are. I'll discuss how I divided the course into different sessions to achieve these objectives, and how each module within these sessions is structured in layers.

The quotes illustrate how the teacher make use of LDT as a common language and conceptual tool to analyze and planning the redesign. As teachers advance through this iterative inquiry journey, facilitated by LDT and enriched by sharing experiences and receiving feedback within their community, they develop a collective and sustained inquiry into understanding students, visualizing students' cognitive processes.

Teacher C： What I've learned today is to center more on the students, but not in the traditional sense where students are left to their own devices. Rather, it's about intentionally letting students take the lead, guided by our educational objectives. This approach allows them to actively engage and take ownership of their learning, which is directed towards achieving specific goals we set.

Spiral development in teachers' understandings leads them to prioritize developing students' thinking and disciplinary skills, and to re-conceptualize themselves as learning designers and design learning environments that are student-centered, self-motivated, active, reflective, and collaborative to achieve the changing learning outcomes.

Teacher A：This semester, I've made changes by analyzing from the students' perspective. I found some parts unnecessary and asked myself, "What is the purpose of your design? If it doesn't overcome this hurdle, just pass it." Previously, I didn't understand the principles behind my choices; they were unclear. Now, my principle is to center around students' needs—whether they really need this particular content. Previously, I was concerned about the completeness of knowledge, whether the content was comprehensive and logically correct. Now, I'm trying to minimize my subjective views to make it more objective. Today, I see the teacher as a designer, designing an environment that showcases my students' progress in knowledge, skills, and non-technical thinking. The design aims to visualize the learning path and the students' capabilities.

The process of teachers' change is characterized by an ongoing and sustained development cycle, driven by evolving beliefs and enhanced understanding, reinforced by emotional resilience and increased motivation cultivated through the iterative cycles of inquiries. The iterative nature of this cycle fosters a dynamic environment that supports continuous growth, enabling teachers to become more effective designers of learning experiences and environments.

4. Conclusion

This ongoing project discusses challenges in promoting pedagogical innovation among university teachers and proposed a model and explore of its impact on teachers' understanding of learning design, design practice. The study how teachers develop their understanding of LD and implementation of pedagogical innovation in iterative cycles of design process, which provide valuable insights into effective Teacher Professional learning for pedagogical innovation. More in-depth analysis from ongoing work will explore the mechanisms of how collaborative inquiry, iterative cycles of practice, and the LDT framework in facilitating this transition.

References

Bennett, S., Agostinho, S., Lockyer, L., Jones, J., & Gigliotti, A. (2022). From learning design to teacher design practice: researching how teachers design for technology integration. In Handbook of Digital Higher Education (pp. 149-160). Edward Elgar Publishing.

Chen, D., Liang, L., & Law, N. (2022). Navigating multilevel challenges in learning design: an investigation of novice designer teams' learning trajectory. International Journal of Mobile Learning and Organisation, 16(3), 377-403.

Goodyear, P. (2015). Teaching as design. Harvard review of higher education, 2(2), 27-50.

Laurillard, D. (2008). Technology enhanced learning as a tool for pedagogical innovation. Journal of Philosophy of Education, 42(3-4), 521-533.

Law, N. (2008). Teacher learning beyond knowledge for pedagogical innovations with ICT. International handbook of information technology in primary and secondary education, 425-434.

Law, N. (2017). Instructional design and learning design. In The sciences of learning and instructional design (pp. 186-201). Routledge.

Law, N., & Liang, L. (2020). A multilevel framework and method for learning analytics integrated learning design. Journal of Learning Analytics, 7(3), 98-117.

Nguyen, G. N., & Bower, M. (2018). Novice teacher technology-enhanced learning design practices: The case of the silent pedagogy. British Journal of Educational Technology, 49(6), 1027-1043.

Sancar, R., Atal, D., & Deryakulu, D. (2021). A new framework for teachers' professional development. Teaching and teacher education, 101, 103305.

促进知识建构的信息科技课堂学习空间设计研究——基于发生认识论视域

A Study on the Design of Learning Spaces for Promoting Knowledge Building in Information Technology Classrooms — Based on the Perspective of Genetic Epistemological

丁添[1*]，唐璇[2]
西北师范大学 教育技术学院
[*]2581116193@qq.com

【摘要】 数字化时代呼唤学习空间的重组。信息科技课程学习多以项目式学习方式为主，而传统课堂空间难以有效促进学生知识建构。该文基于发生认识论设计了信息科技课堂学习空间布局结构与功能要求。物理空间主要由组内讨论演示区、创意作品实践区、集中分享展示区构成，以“云电脑”和“云端空间”构成的虚拟学习环境具备数字资源仓库、学生数据画像、智能孪生助手等关键要素，虚实融合空间保障学生认知加工与认知外化，以期辅助促进学生知识建构过程。

【关键词】 发生认识论；知识建构；学习空间；信息科技课程

Abstract: *The digital era demands a reorganization of learning space. As information technology classrooms primarily adopt project-based learning, and traditional classroom spaces are difficult to effectively promote student knowledge construction. The paper designs the layout structure and functional requirements of the information technology classroom learning space based on the theory of epistemology. The physical space mainly consists of a group discussion and presentation area, a creative work practice area, and a centralized sharing and display area. Virtual learning environments composed of "cloud computers" and "cloud spaces" have key elements such as digital resource repositories, student data portraits, and intelligent twin assistants. The integration of real and virtual space guarantees students' cognitive processing and cognitive externalization, with a view to assisting in facilitating the process of knowledge construction.*

Keywords: Genetic Epistemological, Knowledge Building, Learning Spaces, Information Technology Courses

1.引言

学习空间是教与学发生的关键场所，2022 年教育部发布义务教育阶段信息科技课程标准，提出“学生需要利用各类数字设备与团队成员合作解决学习问题；教师应提供数字化、虚拟化、仿真化的学习环境”（中华人民共和国教育部，2022）。学习环境促进学生认知发生与知识建构，新型学习空间支撑着学生信息科技核心素养的培养，学生不断更新的学习特征和需求也助推着学习环境的重塑（李艳燕，2023）。而当前学校信息科技课堂教学环境大多以“排排坐”的形式布局，学生座位之间呈现“左右相连、前后间隔”的空间布局，此种布局形式难以支持学生开展项目式学习、协作式学习，学生知识建构过程受阻，因而无法有效促进学习的发生。另外，还存在学习空间的应用与教学割裂、教与学以物理空间为主等问题，并且新课程标准要求学生学习需要大量的虚拟仿真环境，这对于传统机房计算机的算力提出了巨大挑战，传统课堂学习空间亟待向新型学习空间转变。基于此，为了有效培养学生信息科技课程核心素养，促进学生知识建构，本研究以发生认识论作为理论基础设计信息科技课堂学习空间结构与功能，以期为信息科技课堂教学环境优化提供思路与理论借鉴。

2.文献综述

2.1. 学习空间研究现状

学习空间是学生学习发生的主要场域。有研究认为学习空间是指空间的存在状态，从构成要素来看，学习空间由物理场所、虚拟场所、工具性客体和对象性客体四大要素构成（景玉慧和沈书生，2021）；从物质形态来看，学习空间通常由物理空间与虚拟空间构成（Beery et al., 2013）；从空间设计遵循的理念来看，学习空间需要激发学生学习兴趣，强调师生间、生生间的交互(Bines & Jamieson, 2013)。从促进知识建构的角度来看，空间需要支持学习者协作学习中的资源分析、方案分析、问题分析与解决、学习反思等过程（李海峰和王炜，2020）。

对于学习环境的设计应该充分满足学生的心理发展需求，学生要在学习环境中与多种要素充分互动，以促进其自身的发展。学习空间的设计需要考虑学习者、教师、学习内容、教学设施与技术之间的交互作用(Yazzie-Mintz, 2009)。随着智能技术的发展，学习空应具有开放性、智能性、多样性和协同性等特征(杨霞和范蔚，2023)，空间设计也要考虑布局规划、物理环境和服务、设施陈列以及智能技术在空间中发挥的支撑作用（许亚锋和高红英，2018）。也有研究提出学习环境的设计要紧紧围绕未来社会人的发展，并提出主要面对的三大挑战，即人机交互、学习反思、创新思维（塔卫刚，2018）。

在虚拟学习环境方面，以智能化、个性化、动态化、泛在化为主要特征（赵一婷等人，2022）的网络学习空间为学生学习打造了沉浸式学习环境。在智能技术的支持下，学生个体与虚拟学习环境进行真实交互，并在交互中完成知识意义的建构过程。虚拟学习空间应具备的主要功能有知识呈现、模拟训练、环境体验，学习者在虚拟学习空间中打破知识呈现的空间立体局限，在近似真实的环境中训练操作技能（孙志伟等人，2019）。

2.2. 发生认识论相关研究现状

发生认识论是皮亚杰提出的有关于人类心理机制的理论（皮亚杰，1981），认为人类的认识起源于主体与客体之间的交互作用，并从生物与心理双重维度揭示了人类认识发生的规律（景玉慧等人，2023）。基于发生认识论，多位研究者探讨了以学生为中心的学习空间构建、学生思维评价（张沿沿等人，2020）、素养内涵及培养（陈羽洁等人，2021）等内容。课堂学习环境是学生核心素养培养的重要场所，学生知识建构过程离不开新型学习空间的支持。因此信息科技课堂环境的设计需要将形成核心素养过程中的内在认知放在首要位置，关注学生认知发生机制、认知特点（景玉慧等人，2023），以及促进学习发生必需的学习活动，进而设计开展学习活动所需的学习空间，探析学生认识发生与学习空间的内在关系（景玉慧等人，2022），让信息科技的学习空间更好更适切的提供个性化学习支持。

3.发生认识论视角下的信息科技课堂学习空间设计

以发生认识论作为信息科技学习空间设计的理论基础，有学者提出学习空间具有“居所”和“转运”两种属性，其中居所属性强调学生认知发生需要在物理和虚拟环境中，转运属性强调学习空间需要给学生提供交互、展示的环境，以支持学生学习过程中内在信息的输入、加工与输出，包括人类主体之间的交互，学生学业作品的展示等（景玉慧等人，2023）。信息科技课程学习需要学生之间不断沟通交流，进行知识建构，共同解决新问题，新型学习空间的建设更应有利于学生开展信息科技主题式、跨学科探索与实践，并能够支持学生动手操作、深度参与、设计制作、系统搭建等活动（杨晓哲和刘昕，2022）。基于对现有文献的梳理和学习空间设计要素的分析，本研究对信息科技学习空间要素进行逐一分析设计，并说明学习空间如何助力学生认知发生与知识建构。

3.1 新型物理学习空间设计

3.1.1 多边形座位空间布局：促进认知建构

信息科技学习不仅需要合作式的动手实践，更需要小组深入讨论交流，这些均需要合理的座位空间布局提供支持。学生座位布局设计应能满足学生小组间交流讨论的需求，学生需要在进行持续的面对面交流和反思过程中建构和确认意义（叶映华和尹艳梅，2019），并涌现出新的有价值的观念或知识（胡金艳等人，2024）。以多边形结构为特征的学生座位布局充分发挥了学生的主体地位（如图 1 所示），能够支持以小组为单位的合作学习、项目式学习等，这种去领导化的空间布局结构，强调人人平等的地位，学生通过近距离的沟通交流，合作学习，分享认知，即时传递知识（杨俊锋等人，2013）。

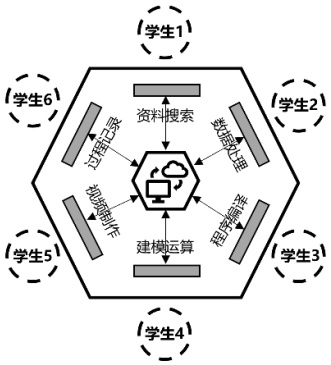


图 1 学生小组座位布局

在问题解决过程中，小组内部的交流互动、人工制品制作等活动，将分布在个体内部、个体之间的认知进行扩展（张立新和秦丹，2018），促进学生个体内部认知结构的完善。学生在环形小组合作中，对陌生任务情境的理解与认知并不相同，在协作交流过程中不同认知的碰撞产生认知冲突，通过不断的资料搜索输入外部信息，在讨论和反思中重组转变认知，最终以设计方案、作品制作等形式输出个体与群体的认知结果，实现意义建构。环形物理空间的设置能够让组内任何成员有与他人面对面交流的机会，有效促进了认知冲突、认知重组的过程（胡卫平等人，2023），提升认知效率，如图 2 所示。

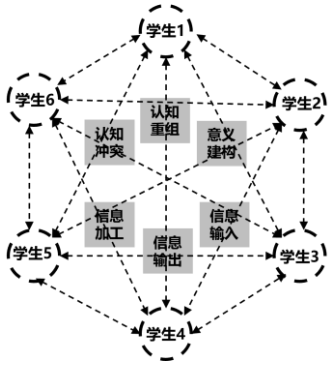


图 2 小组空间布局下的认知发生

在多边形座位布局的基础上，学生信息科技课程中使用的计算机也需同步更新。传统机房中使用由独立主机、显示器、键鼠设备组成的硬件系统，这需要专业的技术维护人员定期进行硬件设备与软件系统的维护和升级，使用时间一旦过长，系统卡顿、病毒入侵等问题随之而来。新型学习空间中需充分利用云计算、云存储技术，采用应用端集约化方案，学生用机仅有一台显示设备，所有的计算机运算、数据存储、安全防护、均在云端完成，这样的“云”电脑具有即开即用、集中管理、高效运维、便捷安全、互联互通的技术优势，能够大幅降低学校教师工作成本，提升教学效率，同时也保障了学生信息科技学习的顺利开展。

在信息科技课堂中发挥物理空间布局优势与技术优势，学生在项目式学习中剖析讨论情境问题，设计探究方案，分工协同完成探究任务，“云”电脑根据学习任务需求与难度等级将算力、网络、存储资源等弹性分配给每位学生，学习小组的资料搜索、数据处理、程序编译、

建模运算、视频制作、过程记录等分工得到了最大限度保障，实现资源利用效率的最大化，保障学习效果最优化。

3.1.2. 物理空间功能分区：保障认知发生

学校信息科技教学空间由多个功能分区组成，除了上述学生学习小组式布局外，信息科技学习空间中还应包括组内讨论演示区、创意作品实践区、集中分享展示区等（如图 3 所示）。

组内讨论演示区为学生小组设置“随想随写”的墙壁白板区域，学生在解决项目问题时，将讨论交流过程随笔记录下来，弥补计算机记录过程过慢的缺点，讨论演示区也是学生认知过程外显化的一种有效方式，记录下来的讨论过程也便于后期整理，并向全班学生进行展示，以出声思维的的方式介绍问题解决的路径。

创意作品实践区为学生提供 3D 打印机器、人工智能套件等设备，支持学习小组实践创造创意作品，这些创意人工制品的形成是学生进行知识建构的外显行为表现（万恒和高辛宇，2023）。学生在信息科技课程中经过讨论、设计、理论建模等系列活动，汇聚小组学生共同智慧，最终形成作品设计方案。如在“互联智能设计”跨学科主题的学习中，学生可剪辑宣传视频、设计智能无人机、制作智能交通工具等多种创作形式实现创意想法。

集中分享展示区设置稍高于教室地面的分享站台，这样的布局设置更容易让学生在分享时产生自信的情感。同时能够将其他学生的注意力聚焦于分享者。这是学生对自身认知个性化输出、诊断及优化学习结果的主要学习区域（景玉慧等人，2023）。学生以小组为单位对设计方案、人工制品等信息科技作品进行分享，将小组的分析过程、解决思路、设计亮点向其他人展示出来，帮助梳理自身认知的内在结构，也可以建立认知的逻辑（沈书生，2020）。对于其他学生而言，在分享展示区域与其他学习小组成果进行对比，以弥补自身的学习缺陷，提升自身观点。通过生生评价，提出具有创新性、操作性的评语，在这样的交互中完善自身的内部认知结构。

通过上述对空间布局及其作用的剖析，将学生认知发生的过程辅助区域、结果外显区域、认知传播区域等多种物理区域集成于同一学习空间中，减少了教师固定教学位置的主导性，同时充分发挥学生个性特征。相较于传统“排排坐”的空间布局，此类空间布局能够为学生提供更为充足和灵活的学习活动空间，有效调节学习活动的视觉交替节奏，学生身体在不同的空间位置（中心或边缘）也有助于引导协商共识、评价反馈等社会互动（仇晓春，2023），从而促进知识建构的过程。

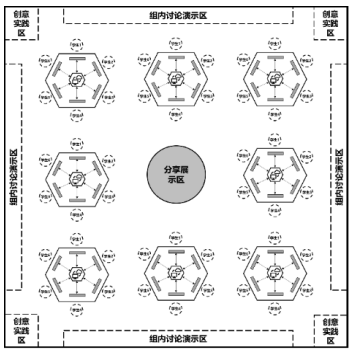


图 3 教室区域空间布局

3.2 新型虚拟学习空间设计

在人工智能技术的快速发展下，学习空间内涵进一步扩展，虚拟空间与实体空间相互融合，共同构成了学生认识发生的关键场所。虚拟学习空间搭载人工智能、大数据、云计算等智能技术，具备资源立体呈现、智慧问答交互、学情实时分析等关键功能，对学生学习全过程予以记录与支持。下面对虚拟学习空间的数字资源、数据画像、智能助手等关键要素及对认识发生的支持作用进行逐一分析。

数字资源仓库：信息科技课程学习具有技术操作性强、知识内容新、概念抽象等特点，这就要求学习资源应以立体化、动态化、新颖化的方式进行呈现。虚拟学习空间内嵌大量数字资源，使用虚拟现实、增强现实、混合现实等技术，为学生信息科技学习打造元宇宙空间，将数字资源立体呈现给学生，增强学生与各类资源的交互性，扩展学生认知渠道，扩大学生的认知场域。

学生数据画像：学生信息科技学习的发生离不开虚实融合的学习空间，学生学习情况是教师调整教学的依据，也是教学效果的直接体现。虚拟学习空间具备学习行为分析、学习成绩统计、学习情感分析、任务作品评估等功能，将学生信息科技学习全过程的情况记录下来，以数据精准刻画学生学习画像。具身认知理论认为外显行为能够反映内在认知（焦彩珍，2020），通过对学生信息科技学习行为的记录和分析，将内隐的认知过程客观呈现出来，帮助学生及时调整学习方式，并为教师提供教学改进的依据。

智能孪生助手：认知发生离不开主客体、多主体之间的交互。在虚拟空间中，学生以数字身份与其他学生进行交流，通过“获取信息—知识建构—知识内化—外化表征”进行认知交互，从而实现意义建构（刘玲，2021）。同时虚拟学习空间搭载生成式人工智能技术，以虚拟学伴的身份与学生进行智能言语对话，帮助学生突破认知局限，解决学生的认知冲突。此类智能学伴具有知识体量大、信息科技能力强、反馈迅速的特点，能够很好支持信息科技学习中的主体探究。

3.3 虚实融合空间促进知识建构

物理学习空间与虚拟学习空间的不同功能均能有效促进学生知识建构，两种环境的深度融合体现出新型学习空间的无缝衔接化、功能互补化、工具丰富化、数据流转化，如图 4 所示。首先，虚实融合的学习空间给学生提供了无缝学习的机会与场域，学生学习不再受到时间与空间的限制，能够实现物理空间学习过程的虚拟化、共享化，以及虚拟空间学习结果的实物化。其次，虚拟与物理空间相辅相成，相互支持，虚拟空间提供大量技术功能与数字资源，给物理空间的小组探究提供“原材料”，在物理空间中学生将这些“原材料”进行加工整理，设计探究，深度开展创新创造活动。再次，虚实融合空间提供不同种类的协同知识建构工具，如知识表征工具、思维外化工具（思维导图等）、效能提升工具（协同编辑文档等）、整合反思工具等，物理空间工具支持学生灵活调整学习节奏，帮助组内组间的共享协商（马宁等人，2017）；虚拟空间工具即时记录学生学习过程（沈书生，2017），将物理工具同步在虚拟空间中，允许学习社区成员参考借鉴，促进知识流转。最后，学生学习过程以数据形式在两种空间中不断循环流转（杨现民和李新，2024），提升教育数据的服务能力，构成数据驱动的知识建构空间。在人工智能技术的加持下，学生物理空间学习行为得以观测记录，通过多模态数据分析转化为一张张虚拟学习画像，建立起数据化的学习档案，存储在虚拟空间中供师生随时调用。这些虚拟数据代表学生学习的每个过程，反应学习的困难问题，师生可应用数据开展物理空间中的教与学。

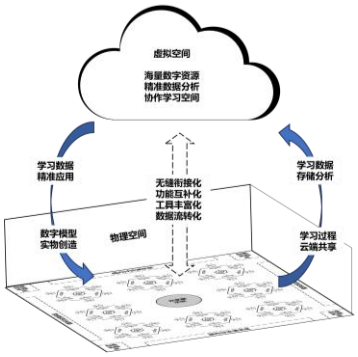


图 4 虚实融合学习空间

4.信息科技课程学习空间应用实践——以乡村学校为例

信息科技课程新型学习空间为学生知识建构提供了全方位的保障。对于乡村学校而言，信息科技课程学习存在物理空间陈旧、虚拟空间缺乏、技术支持不到位、教与学模式传统等现实问题，导致学生知识建构过程受阻。本研究以 Z 省乡村学校为例，具体分析虚实融合学习空间对信息科技教学与学生知识建构的支撑效果。

4.1 实践案例背景

“云机房”作为本研究新型学习空间的典型应用案例，能够很好体现出虚实融合的空间特点。其物理空间设置多边形座位布局、云电脑桌面（仅有物理显示器，系统运算存储均在企业云端），虚拟空间搭载云电脑系统，配有多种数字化资源以及智能学习功能（如图 5 所示）。在全国多个省市乡村学校均有所应用，Z 省乡村学校借助科技企业力量引入并建设“云机房”，现已运行近 2 年时间，能够较好体现出新型学习空间对知识建构的促进作用。该校依托“云机房”环境开展人工智能技术支持的信息科技课堂教学。本研究通过对课堂教学过程观察、教师和学生访谈、学生作品分析多种方式分析新型学习空间对学生认知发生与知识建构的支持效果。



图 5 乡村学校信息科技教学空间

4.2 应用实践案例分析

以“云端+地面”为特征的虚实融合空间为乡村学校的信息科技课程教学提供了全方位的支持。通过课堂观察发现，在“小型系统模拟”等主题学习中教师借助双重空间设计实施了多样化的教学活动。教师将学生随机分为多个学习小组，利用云电脑中的资源将主题场景立体呈现给学生，为学生创设符合认知规律的学习情境，并提出具体的学习主题任务。在该主题学习情境中，学生小组对具体生活场景需求与任务要求进行分析，通过讨论交流逐步明确具体场景中所需的开关系统功能，将需求与系统功能关系同步记录在组内讨论演示区，组内学生利用“云”电脑即时记录讨论过程，并上传至云端存储。教师在过程中通过过程干预、针对性指导、提供探究支架等方式，辅助学生进行认知加工。通过合作探究活动，小组将最终的模拟开关系统设计方案在创意作品实践区域进行组装实现。

在“三维建模”主题的探究学习中，学生开展了交流讨论、合作探究、作品制作等活动，这些学习活动蕴含了学生认知发生的全过程。模型构建与实物制作、作品分享展示是将认知过程外显化的方式。作品制作活动中学生利用虚拟学习空间中配备的建模软件建立模型，物理教室创意实践区中的人工智能套件、3D 打印机器，将设计构建的 3D 模型进行制作组装（如图 6 所示）。学生小组选择代表在分享展示区域向其他学生展示、讲述设计理念、设计过程等，小组内其他学生利用虚拟空间的展示共享功能，将记录的认知过程、作品细节图片、视频等同步传送到每位学生的电脑桌面，辅助分享认知。

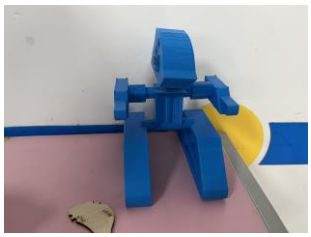



图 6 学生虚拟建模+模型打印

通过对教师访谈与学生访谈的分析发现，教师表示“新的学习空间改变了过去信息科技教学割裂的情况，特别是多边形的座位布局更容易开展协作交流活动，学生更喜欢进行讨论和创作”。在协作讨论方面，学生同样表示“小组合作更加方便”“变化更多的是学习前后对物联网（概念）的认知不一样了，很多同学都能提出不同的观点”。能够看出师生均认为新型学习空间的使用有助于协作学习的开展，学生知识建构过程更加便捷有效。

5.结语

本文设计的虚实融合新型学习空间给学生信息科技课程学习提供了环境支撑，物理学习空间的布局、陈设，虚拟空间的功能、技术等均需指向学生学习中的认知发生过程，帮助促进学生知识建构过程。从学生学习认知的角度为学习环境设计研究提供了微观视角的解读与探索，同时也为一线学校建设新型学习空间提供了理论参考。

参考文献

万恒 & 高辛宇. (2023). 项目化学习中的教师：角色认知与胜任力要素. 教师教育研究 (02), 63-68+91.

马宁,吴焕庆 & 崔京菁.(2017). 以协同知识建构为核心的教师混合式研训模型研究.教师教育研究(03),31-38.

中华人民共和国教育部.(2022). 教育部关于印发义务教育课程方案和课程标准(2022 年版)的通知 .http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html,2022-04-06.

仇晓春.(2023). 涌现的学习活动——ACAD 学习空间设计框架评介. 远程教育杂志(04),70-78.

皮亚杰. (1981). 发生认识论原理[M]. 王宪钿, 张梅玲, 刘静和, 等译. 北京: 商务印书馆.

叶映华 & 尹艳梅. (2019). 大学生批判性思维的认知特点及培养策略探析——基于小组合作探究的实证研究. 教育发展研究 (11), 66-74.

李艳燕. (2023). 科技赋能教育创新：学习环境的变革与升级. 人民论坛·学术前沿 (20), 73-79.

李海峰 & 王炜.(2020). 实习教师实践性知识生成的混合学习模式探索——基于五环建构法. 现代远距离教育(04),10-18.

孙志伟, 李小平, 张琳, 姜丽萍 & 毛旭. (2019). 虚拟现实技术下的学习空间扩展研究. 电化教育研究 (07), 76-83.

许亚锋 & 高红英. (2018). 面向人工智能时代的学习空间变革研究. 远程教育杂志 (01), 48-60.

刘玲.(2021). 网络学习空间中分布式认知建构模式研究. 教育实践与研究(C)(03),12-15.

杨霞 & 范蔚. (2023). 数字时代学习空间的边界重构与权力回归——以人的发展为分析视角. 中国电化教育 (11), 18-27.

张沿沿, 冯友梅, 顾建军 & 李艺. (2020). 从知识结构与思维结构看思维评价——基于皮亚杰发生认识论知识观的演绎. 电化教育研究 (06), 33-38.

陈羽洁, 张义兵 & 李艺. (2021). 素养是什么？——基于皮亚杰发生认识论知识观的演绎.电化教育研究 (01), 35-41.
杨晓哲 & 刘昕. (2022). 面向数字素养的义务教育信息科技课程. 全球教育展望 (06), 109-117.
杨俊锋, 黄荣怀 & 刘斌. (2013). 国外学习空间研究述评. 中国电化教育 (06), 15-20.
杨现民 & 李新. (2024). 数据充分赋能教育数字化转型的内在逻辑与实践路向. 中国教育学刊(02),9-15.
张立新 & 秦丹. (2018). 分布式认知视角下个人网络学习空间中有效学习的保障路径研究. 电化教育研究 (01), 55-60.
沈书生. (2020). 学习空间：学习发生的中介物. 电化教育研究 (08), 19-25+42.
沈书生.(2017). 从教学结构到学习结构:智慧学习设计方法取向. 电化教育研究 (08), 99-104.
赵一婷, 钟绍春 & 唐烨伟. (2022). 技术赋能视角下网络学习空间生态研究——内涵、要素与架构. 中国电化教育 (10), 126-133.
胡卫平,徐晶晶,皮忠玲 & 张阳.(2023). 智能时代的教育变革:思维型教学理论引领“技术赋能教学”.远程教育杂志(06),3-9.
胡金艳,杨霞,高兴启,等.知识建构社区中多主体模拟支持的观点涌现模型构建：复杂性科学的视角[J].电化教育研究,2024,45(01):100-107.
景玉慧 & 沈书生. (2021). 理解学习空间：概念内涵、本质属性与结构要素. 电化教育研究(04), 5-11.
塔卫刚. (2018). 学习科学视野下学习环境设计研究. 现代教育技术 (06), 5-10.
景玉慧, 王成梁, 沈书生 & 李浩君. (2023). 高质量学习空间设计：视角、理路与实践——以初中阶段为例. 现代远距离教育 (04), 53-60.
景玉慧, 沈书生 & 李浩君. (2022). 如何设计“以学习为中心”的学习空间？——以小学阶段为例. 现代教育技术 (02), 54-62.
焦彩珍. (2020). 具身认知理论的教学论意义. 西北师大学报(社会科学版) (04), 36-44.
Beery, T. A., Shell, D., Gillespie, G., & Werdman, E. (2013). The impact of learning space on teaching behaviors. Nurse education in practice, 13(5), 382-387.
Bines, J. E., & Jamieson, P. (2013). Designing new collaborative learning spaces in clinical environments: experiences from a children’s hospital in Australia. Journal of interprofessional care, 27(sup2), 63-68.
Yazzie-Mintz, E. (2009). Engaging the voices of students: A report on the 2007 & 2008 High School Survey of Student Engagement. National Association for College Admission Counseling.

数智时代基于知识建构的教师数字素养提升逻辑与实践

王贤灿[1,2]
[1]西北师范大学教育技术学院
[2]深圳市南山区松坪学校
627336458@qq.com

【摘要】 数智时代教师数字素养的提升对于教育数字化转型与教育现代化具有重要意义。教师数字素养提升现实中面临着认知偏差、尚未有成熟模式以及配套机制不健全等问题。基于知识建构的教师数字素养提升行动依据激发兴趣、合作建构、分享成果、评价反馈等环节，建立研训逻辑，打造教师数字素养提升模式，充分利用数字技术，进行教师深度研修闭环设计与实践，实现教师数字素养的提升，促进教师专业发展。

【关键词】 知识建构；教师数字素养；提升逻辑；研训实践

Abstract: *The improvement of digital literacy among teachers in the digital age is of great significance for the digital transformation and modernization of education. In reality, the improvement of digital literacy for teachers faces problems such as cognitive bias, immature models, and inadequate supporting mechanisms. The action of enhancing teacher digital literacy based on knowledge building is based on stimulating interest, collaborative construction, sharing results, and evaluating feedback. It establishes a research and training logic, creates a model for improving teacher digital literacy, fully utilizes digital technology, designs and practices a closed-loop system for deep teacher training, allowing teachers to achieve easy and deep learning, improve their digital literacy, and promote their professional development.*

Keywords: Knowledge building; Teachers' digital literacy; Promotion logic; Research and training practice

1. 前言

教师是教育者，教师的素养直接影响到学生的成长。随着信息技术的不断进步，特别是大数据、云计算、人工智能等技术的发展，人类已进入数智时代。数智时代的科技发展日新月异，生产力不断提升，人们的生产、方式、学习、思维方式发生巨大的变化，对教师素养提出了新的要求。2022 年 11 月 30 日，教育部发布了“《教师数字素养》标准”(JY/T0646-2022)，从数字化意识、数字技术知识与技能、数字化应用、数字社会责任、专业发展等五个维度描述了数智时代应具备的数字素养。知识建构是个体在某一特定社区中互相协作、共同参与某种有目的的活动（如学习任务、问题解决等），最终形成某种观念、理论或假设等智慧产品（赵建华，2007）。对数智时代的教师个体来说，在物理空间和网络空间基于知识建构发展数字素养尤为重要。

2. 教师数字素养培育的现实审思

2.1. 对数字素养培育存在认知偏差

《教师数字素养》标准是“扎实推进国家教育数字化战略行动，完善教育信息化标准体系”（中华人民共和国教育部，2022）的重要内容。现实工作中，部分教师对数字素养培育的必要性和迫切性持有疑虑或抵触情绪，甚至认为花费时间提升数字素养的投入产出比偏低，这些认知差异影响了数字素养培育行动的推进。

2.2. 数字素养培育模式尚未建立

“专业发展”是《教师数字素养》标准中列明的五大维度之一。专业化要求教师按照专业的标准行事，遵守外界订立的专业行规，接受专家的临床指导和考核评价（叶澜等人，2001）。在教师数字素养培育行动中，需选择系统、科学又有用、有趣、有效的培育模式。现今，教师数字素养培育主要采用培训讲座、研讨会等方式，缺乏系统设计，以致不能帮广大教师迅速进入数智时代的新角色，缺乏系统的理论学习，不能适应数字时代教与学的变化，在教育

教学实践中缺乏有效抓手和方法，不能真正落地开展素养导向的教学。

2.3. 数字素养培育的配套保障机制不健全

2022 年 3 月 28 日，国家智慧教育公共服务平台正式开通。为广大教师和学生提供了丰富的教育教学资源。但在资源配套与保障方面，存在虚拟仿真项目、智能学习模块缺失等情况。在现实中，省、市、区、学校都建有大量的资源平台，数字资源的治理效能不高，客观上的多种平台同时存在，造成了教师选择困难，难以找到值得学习与借鉴的优秀案例资源。同时，在网络接入、资费投入等方面的保障机制也不健全。

3. 基于知识建构的教师数字素养提升逻辑

数智化发展体现社会和经济向新范式的根本转变，带来产业组织模式、现代基础设施体系、科技人才培育体系、社会发展治理模式等的革新与重构（赵璐和李振国，2021）。

3.1. 激发兴趣：指向教育教学中的实际问题

教育必须因时而变，教师需要转变观念，以适应时代的需要。凯里尤基(Kariuki)在总结了许多学者的观点后认为，教师专业发展是一种教师为了改进其教学知识和技能的过程（Kariuki, P. K.，2009）。

数智时代，为了提升教师的数字素养，需充分利用 AR、VR、MR 技术，以吸引教师的兴趣，便于教师建立新旧知识之间的联系，激发教师对数智时代技术在教育中应用的关注。观念的产生常常同真实的、联系密切的事物相关（赵建华，2007）。在教师数字素养培育过程中，充分利用线上线下学习在时间、空间上的优势，把数字素养提升行动与教师教育教学中的问题解决尝试结合起来，设计混合式的专业发展学习与实践活动，既包括线上的课程学习、网络研讨会和在线协作，也包括线下的研讨会、工作坊和实地观摩公开课等，以满足教师不同的学习偏好和时间安排，提供灵活和多样化的学习机会，以保持教师的学习热情，激发教师利用数智技术改善教学的探索欲望和求知欲，实现全方位、灵活和个性化的专业发展。充分利用元宇宙、数字孪生等技术模拟真实的教学环境，让教师在虚拟空间中，面对教育教育中的各种教学环节、突发事件等实际问题进行教学演练、策略调整和创新尝试，以便教师提升教学应对能力。

3.2. 合作建构：基于线上线下融通的交流场域

在明确具体学习目标的情况下，建设线上线下融通的教师数字素养学习交流场域，打造教师学习空间，为教师提供一个共享经验、资源和知识的平台。在这样的社区中，教师可以相互交流、互相学习，共同建构知识，促进专业发展。与传统学习空间相比，网络学习空间在物质、社会和精神三个维度上展现出更加明显的联通与融合特征（陈丽和沈欣忆，2024）。

线上线下融通的交流场域，教师通过与同事、教育专家或其他领域的专业人士合作学习，共同探讨教学实践、研究成果和教育理论，通过互动和讨论来共同建构知识，助力教师获得新的思路和方法，促进教师个体和集体的学习和发展。

陈向明教授认为：教师的知识结构分为“理论性知识”和“实践性知识”。“教师实践性知识”是教师真正信奉的，并在其教育教学实践中实际使用和(或)表现出来的对教育教学的认识。”，由六个方面的内容构成:教育信念、自我知识、人际知识、情境知识策略性知识和批判反思知识（陈向明，2003）。所以，在教师数字素养提升行动中，除了加强教师理论性知识的学习，也需加强教师实践性知识的学习。

线上线下融通的交流场域具有特有的文化背景和价值观，影响着教师个体的认知和学习方式，帮助教师形成正确的教育信念，塑造个体的认知模式和思维方式，影响教师对数字素养提升重要性的理解，帮助个体解决问题和改进学习策略。教师个体通过参与数字素养学习空间的打造，分享反馈和建议，建立良好的人际关系和社会网络，增强对数字素养提升的积极性，促进教师个体专业发展。

3.3. 分享成果：数字技术支持下全方位展示

在教师数字素养培育过程中，教师通过培训、实践和自我提升创造出多种类型的成果，这些成果既可以反映教师数字素养的成长过程，也是他们在教学实践中应用数字技术能力的具体体现。

这些成果形式多种多样，多多益善。可以是数字化教学设计方案，或是自主开发或改编的数字教学资源，如微课视频、动画、图表、音频解说、互动模拟实验等；或是记录教师在实际教学中应用数字技术的案例报告，包含具体操作步骤、学生反馈及教学效果分析等内容；或是关于数字技术在教学中应用的实践反思文章；或是基于实证研究探讨数字素养提升对学生学习成效影响的学术论文。或是教师个人在课程设计、教学实施、学生评价等方面运用数字技术的全过程记录和成果集合。也可以是数字化教学评估工具，如开发或使用的在线测试、学习分析工具，用于对学生学习进度和能力的实时监控与评价。还可以是与其他教师合作完成的项目，例如共创的数字教材、资源共享库、线上研讨会议记录等。以及参加各类数字素养提升培训后获得的认证或结业证书，以证明教师在某一领域内的数字技术能力水平。

通过这些成果的展示与分享，体现了教师数字素养的提升过程，也为同行交流、教学改革以及自身专业发展提供了有力的支撑材料。

可通过国家、区域教育云平台搭建数字化分享展示平台，充分利用云存储空间或者专门的项目展示场景，方便教师上传、展示各类数字素养提升行动成果。在成果分享过程中，结合论坛、博客、微博、QQ 群、微信群等社交网络工具，方便教师交流心得。定期组织线上成果发布会、研讨会等活动，促进面对面和远程相结合的讨论交流。

借助数字技术手段全方位展示和讨论教师数字素养培养成果，不仅可以激发教师的学习积极性和创造性，也有利于构建一个开放、互动、持续发展的教学资源库。优秀成果可以被长期保存在网络平台上，形成学习资源库，供后续参考学习，同时也为教师数字素养的持续提升积累宝贵财富。

通过数字素养提升行动的成果展示与分享，加强教师个体与外界信息的互动、内部思考和个体经验的整合来建构知识。

3.4. 评价反馈：共性与个性的深度刻画与持续改进

数字技术应用于教育教学中，具有直观、灵活、实时、立体等特点。数据在教育数字化转型过程中承担“动力引擎”的重要作用，能够有效提升教育解释力、教育诊断力、教育预测力、教育决策力以及教育监督力，驱动数字化教育创新发展（杨现民和李新，2024）。

利用线下线上融通的学习平台，采集教师学习数据，为教师提供数字素养学习与实践反思的支持。教师通过学习平台制定个性化的学习计划，反思自己的教学实践和专业经验，参与自我评估和实践改进，与其他教师进行知识协同建构，不断提升教师个体与群体的教学能力和专业素养。

在教师数字素养培育行动中，利用嵌入式的物联网、网络学习平台及人工智能技术，对教师的教学实践、教育背景、培训记录、教学质量评价、科研成果、学生反应、网络教学活动等多维度数据进行收集与分析，构建教师数字素养培育画像模型，揭示教师在不同维度的优势、特长、不足及其发展趋势，精确地评估教师的专业水平和发展潜力，帮助教师发现自己的教学偏好、挑战和成长点，针对性地制定教师职业发展规划和培训方案，助力实现精细化、个性化的教师评价与反馈机制，以实现教师数字素养发展共性与个性的深度刻画与持续发展。

4. 基于知识建构的教师数字素养提升实践案例

手机在人们日常生活中承担着越来越重要的角色。为了让手机更好的为教育教学服务，学校以《手机在教学中的创新应用》为主题，从教学实际出发，通过提供灵活的研修形式、实用的工具内容、真实的课例资源、线上线下跟进式评价指导，成功的实施了一次基于知识建构的教师数字素养提升行动。

4.1. 提升行动准备阶段：激发兴趣，资源保障

4.1.1. 调研需求

通过学校管理层与教师进行面对面沟通、全员线上调研等方式，了解当前手机在教学中的使用频率、主要用途及遇到的问题。

4.1.2. 确定目标

数字化意识目标：具备积极拥抱数字化教学的意识和态度；主动探索和实践手机在教学中的创新应用。

数字技术知识与技能目标：熟悉手机在教学中的应用方法和技巧；具备基本的数据分析和处理能力，能够利用手机收集、整理和分析学生的学习数据，为教学决策提供依据。

数字化应用目标：实施基于手机的教学活动和任务，创设互动、个性化的学习环境；有效利用手机应用和资源；能够利用手机进行教学效果的实时反馈和评估。

数字社会责任目标：正确使用手机进行教学，避免不良信息的干扰和侵害；关注学生的数字素养培养，引导学生树立正确的数字使用观念，培养学生的自律意识和责任感。

4.1.3. 开发平台

利用市教育局提供的云平台，自主开发主题为“手机在教学中的创新应用”的教师数字素养提升行动项目平台。该平台的基本功能有：一是在线研修。实现课程教学直播、录播，手持设备、台式电脑均可学习。二是班级研讨。三是资源推送。实现数字化资源的主动推送。四是指导反馈。实现任务驱动，专家引领、跟进、指导。

4.1.4. 归集资源

一是整理案例。搜集国内外手机教学应用的成功案例，分析教师合理使用手机提升备课效率、改善教学效果方面的积极作用。二是设计课程。建立包括技术操作、案例分析、教学设计等模块的课程体系。三是资源分类。准备相关的视频教程、PPT 课件、操作指南等数字化教学资源，并将这些资源上传到数字素养提升行动项目平台。四是场地准备。选择宽敞明亮、设备齐全的教室作为培训场地，确保每位参训教师都有足够的空间进行操作。利用人工智能、物联网等技术搭建微格教室实现数据无感记录、采集、分析与反馈。

4.2. 提升行动实施阶段：自主合作，充分建构

4.2.1. 明确目标

一是提出问题：“设想一下，如果在您的教学中融入手机技术，可以在哪几个环节实现哪些优化？”二是明确目标。确定具体的数字素养提升目标：提高参训教师在教学中熟练使用手机进行备课、上课、辅导、反馈的能力，设计出一个利用手机优化教学的案例。三是公布培训手册。介绍培训的背景、目的和安排，激发参训教师的学习兴趣和热情。明确纪律。强调培训纪律和要求，确保培训过程中的秩序和效果。

4.2.2. 自主建构

组织参训教师进行手机教学应用的实践操作，学习与操作各类手机应用于教学的工具，如手机语音识记、辅助备课软件、智能测评系统等。引导教师利用学习平台学习手机应用于教学的理论基础、应用原则和方法技巧。鼓励参训教师提问和讨论，特别是利用案例进行剖析。

4.2.3. 合作建构

教师按照学科背景自愿分组。通过“手机在教学中的创新应用”的教师数字素养提升行动项目平台，针对教学备课、上课、练习、辅导、反馈等环节中的某一问题确定一个项目任务，进行研究，共同设计一个手机在教学环节中的创新应用案例。

4.2.4. 互动交流

安排小组讨论和分享环节，鼓励参训教师之间分享手机教学的经验和心得。每组教师代表进行现场展示和讲解，演示如何在真实的课堂教学中创新应用手机改进教学。激发大家的学习热情和创新意识。每组教师对教师个体学习情况和其它小组的案例展开自评与它评。

4.2.5. 答疑解惑

在“手机在教学中的创新应用”的教师数字素养提升行动项目平台，为每位教师配备一个智能学伴，掌握教师的学习与实践情况，及时解答参训教师的疑惑和困惑，给教师提供实时的、必要的学习支持与反馈服务。

4.3. 提升行动总结阶段：线上线下，分享成果

4.3.1 成果展示

组织参训教师进行手机在教学中的创新应用成果展示，利用线上线下空间展示教师数字素养提升行动成果和创新实践。评选出优秀作品和表现突出的参训教师，进行表彰和奖励，激励大家继续努力。

4.3.2. 整理归档

教师整理培训过程中产生的相关资料和成果，包括教师的学习行为数据、教学课件、操作指南、优秀作品等，按要求上传到“手机在教学中的创新应用”的教师数字素养提升行动项目平台指定位置，为后续的提升行动积累丰富的数据资源。

4.4. 提升行动升华阶段：评价反馈，拓展升华

4.4.1 反思提高

教师参考“手机在教学中的创新应用”的教师数字素养提升行动项目平台的数据，根据自己在整个数字素养提升行动中的情况，填写个人学习反馈表，评估自己在掌握人工智能技术方面的进步和未来应用计划。

4.4.2. 总结反馈

根据“手机在教学中的创新应用”的教师数字素养提升行动项目平台的数据分析结果，与参训教师进行沟通，了解教师个体利用手机开展教学创新的情况，对每个小组的实践活动给予实时的反馈和建议。根据在参与提升行动的教师中进行访谈与调查，对整个提升过程进行总结和反馈，分析行动实践的优点和不足，为后续的数字素养提升行动提供经验借鉴。

4.4.3. 拓展升华

给出进一步探索教师数字素养提升长效机制的问题：除手机外，如何利用其它数字工具，将数字素养与教师的日常工作更加紧密结合，以及如何利用人工智能、大数据等技术来促进教师的个性化和终身学习，拓展前面的研究成果，为后续研究升华打下基础。

在本次主题为“手机在教学中的创新应用”教师研修活动中，成功开展了基于知识建构的数字素养提升行动。教师依据工作节奏自控时间实现轻松学习；通过提升项目平台实时掌握学习情况实现自适应学习；通过跨学科跨班组交流实现共同建构；达成了本次教师数字素养提升行动既定的目标。

5. 结语

数字时代基于知识建构的教师数字素养提升行动，强调教师作为学习者在数字素养提升过程中的主体性，通过建立线上线下融通的教师数字素养提升项目学习平台，基于知识建构，依照激发兴趣、合作建构、分享成果、评价反馈等工作逻辑，完善配套保障机制，促进共同体知识的增长。

教师数字素养的提升是一个长期且持续的过程，需要教育行政部门、学校以及教师个人共同努力，以适应并推动教育信息化的深入发展。教育行政部门做好政策制定与政策实施督导，为基于知识建构的教师数字素养提升行动提供充足的算力、通讯、资金等方面支持。学校做好教师数字素养研训的落实与反馈支持，建立线上线下融通的学习空间，促进教师实现知识合作建构，提升教师的获得感、幸福感。教师加强数字素养提升的使命感和紧迫感，扎实学习数字技术，努力创新，不断提升利用数字技术实施教学改进与创新的能力，实现教师专业水平的不断提升。

参考文献

赵建华。(2007)。知识建构的原理与方法。**《电化教育研究》,(5)**, 14。
教育部。(2022)。 关于发布《教师数字素养》教育行业标准的通知。 检索自 http://www.moe.gov.cn/srcsite/A16/s3342/202302/t20230214_1044634.html
叶澜等。(2001)。 **教师角色与教师发展新探**。 北京: 教育科学出版社。
赵璐和李振国。从数字化到数智化：经济社会发展范式的新跃进。检索自 https://www.cas.cn/zjs/202111/t20211129_4816191.shtml
Kariukic(2009). Professional development for 21st century teachers: Effective classroom management . Alabama: Jackson State University.
赵建华。 (2007)。 知识建构的原理与方法。 **《电化教育研究》, (5),** 11。
陈丽和沈欣忆。(2024)。终身学习数字化转型的哲学观与创新方向。**《远程教育杂志》, 42(1)**, 12。
陈向明。 (2003)。 实践性知识: 教师专业发展的知识基础。**《北京大学教育评论》**, **1(1),** 105－113。
杨现民和李新。(2024)。数据充分赋能教育数字化转型的内在逻辑与实践路向。**中国教育学刊(02)**,9-15。

基于知识建构的素养课堂教学策略研究

——以一年级小学语文《影子》为例

肖利猛[1,2*]，黄芳[1]，彭家益[3]，肖霞洁[3]，冯雪琦[3]
[1]华中科技大学教育科学研究院
[2]南方科技大学附属教育集团
[3]南方科技大学教育集团（南山）附属实验学校
*979387090@qq.com

【摘要】 教师课堂教学设计关系教育教学和人才培养的质量，优化教学策略是提升教学质量的有效措施，也是教育高质量发展的必然要求。在新课标核心素养背景下，基于知识建构的理论，以小学语文一年级上册课文《影子》一课作为案例分析，提出“创设真实情境—鼓励分享观点—加强知识发展—注重学习评价”四个方面的素养课堂教学策略，探讨提升课堂教学的效果。

【关键词】 核心素养；知识建构；教学策略

Abstract: *Classroom teaching design of teachers is related to the quality of education teaching and talent training, and optimizing teaching strategies is an effective measure to improve teaching quality, and it is also an inevitable requirement for the high-quality development of education. Under the background of the core literacy of the new curriculum standard, based on the theory of knowledge building, this paper takes the lesson "Shadow" in the first volume of the first grade of primary school Chinese of the People's Education Edition as a case study, and puts forward four literacy teaching strategies: "creating real situations, encouraging the sharing of views, strengthening knowledge development, and paying attention to learning evaluation", and discusses the effect of improving classroom teaching.*

Keywords: core literacy, knowledge building, teaching strategies

1. 引言

在核心素养导向的背景下，课堂教学质量受到进一步的重视和提升。课堂教学是教师的核心任务，也是学校教育教学的主阵地，影响学生的培养与发展，关乎国家和民族的未来。目前，我国小学阶段的语文课堂教学存在一定的改善空间，距离打造素养课堂有一定的差距，需要进一步加强教学策略，提高课堂教学质量，提升学生核心素养。

弗兰克·费舍尔等学者认为“教学是促进学习的最重要的文化成就之一”。著名人类学家玛格丽特·米德提出在教学过程中起决定作用的不是教学内容，而是信息传递的方式和社会环境。知识是由个人和社会建构的，知识建构和协调学习是教学的主要功能。Marlene Scardamalia 等学者课题组通过 30 年对于知识建构学习模式进行的对比研究发现，在知识建构的过程中，学生的认知水平、交流协助能力和素养有一定的提高。因此，基于知识建构理论的课堂教学是提高学生学科素养的一种有效途径。本文借助知识建构理论分析课堂教学，总结提出相关的教学策略，打造学生素养课堂。

2. 理论框架

2.1. 核心素养

新课标提出以核心素养为导向，结合生活情境加强课程一体化设计，凸显学生的主体作用，注重实现“教-学-评”一体化，构建校内与校外相结合的育人机制。语文课程标准提出学生要建构的核心素养包括文化自信、语言运用、思维能力和审美创造四个方面，教师要在识字写字、阅读鉴赏、表达交流和梳理与探究的过程中提升学生语文核心素养，依托学习任务群创设学习情境，拓展教学资源，探索教学方式的改革，支持学生自主、合作、探究学习。小学教师要不断完善课堂教学设计和方法，提高学生核心素养。

2.2. 知识建构

知识建构被认为是一种深层的建构主义，以发展社区内的公共知识为目标，学生是承担社区知识创造责任的积极认知者（张义兵等，2012）。知识建构强调建立在先前知识、协作、脚手架、真实学习情境的基础上扩展学生的认知（任友群等，2020）。知识建构不同于单一的学习，是在创建和修正公共知识的过程中促进个人的学习（赵建华，2007）。关于知识建构的研究逐渐从释义进入到影响方面，Marlene Scardamalia 等学者认为需要将学校改造为知识建构共同体，知识建构的教学应该更加关注问题观点而非知识，关注建构过程而非结果，关注共同体而非个人知识，强调开放互动性和广泛渗透性。多项研究发现基于知识建构教学模式和 12 条原则的学生学习效果更佳，学生的词汇、拼写等基本技能和思维技能、问题解决能力、合作交流能力有显著提高（张义兵等，2012）。知识建构能够促进核心素养背景下的课堂教学，因为它有助于实现从“知识教学”向“素养教学”的转变，这与当前教育改革的方向相契合。因此，打造素养课堂，提升学生的核心素养需要充分结合知识建构理论进行课堂教学。

3. 《影子》课程教学设计案例分析

张义兵等学者认为分科教学是是实施基于知识建构理论教学活动的一大困难，主张整合为综合课程或者在教学中实现学科综合。以小学语文一年级上册课文《影子》部分教学设计为例，该设计结合知识建构原则进行教学设计，同时进行跨学科统整，将语文与科学等学科相结合，提升学生核心素养。具体设计如下：

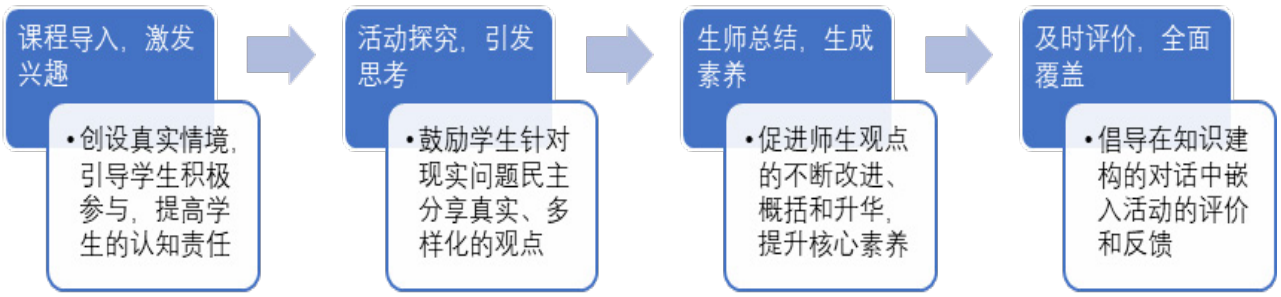


图 1 基于知识建构的课堂教学设计流程图

3.1. 课程导入，激发兴趣

知识建构理论强调现实的问题和真实的观点，教师需要选择与学生生活相关的、真实的社会现象且与教学内容和目标相对应的内容，增强内容的现实性、针对性和多样性，强化在经验中学习，构建新知识与学生已有经验的重新链接。

《影子》课程导入通过“看影子猜动物”的游戏调动学生兴趣，引发学生主动参与到课文中。具体设计如下：

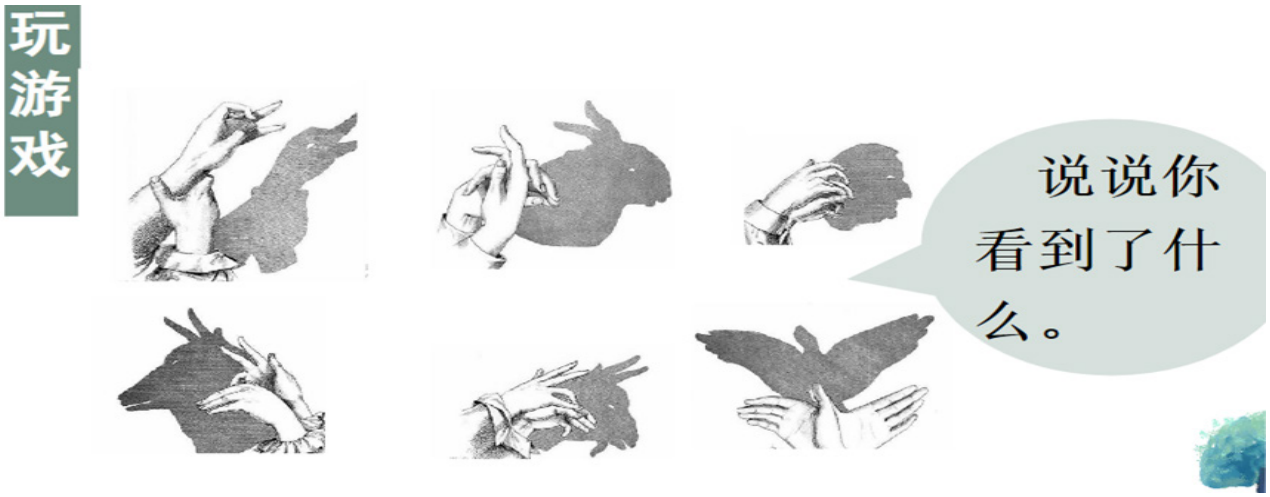


图 2 《影子》课程导入设计图

课前导入部分选取了现实生活中有趣的光影现象为切入口，通过“看影子猜动物”的课前游戏调动起学生兴趣，引导学生主动参与到课文中。接着通过各种形式的朗读，让学生感受到影子的有趣，大大激发了学生学习的快乐和热情。

3.2. 活动探究，引发思考

知识建构理论以观点为中心，主张学生积极分享各种观点，通过探究不断去概括、改进和升华观点，倡导自主、合作和探究学习，在课堂教学设计和讲授中，要突出问题导向，引导学生发现问题、分析问题和解决问题，同时提高学生的动手能力和参与度，通过设置形式多样的实验、讨论、表演等互动形式，让学生在实践中去解决课程的目标和重难点。

《影子》通过进行课堂小实验让学生自主探究和合作学习。教师拿出实验道具，实验工具主要有小人和电筒。假定把电筒当作太阳，小人当作被阳光照射的人，手拿电筒，可以按照不同角度和方向照射小人，也可以调整小人和电筒的距离。同时教师也可以邀请学生上台亲手实验，激发孩子的学习兴趣。教师在有趣的实验中不断引导孩子进行发现，要求学生在实验中观察影子所发生的变化，询问学生为什么会有影子及影子的特点。

3.3. 生师总结，生成素养

知识建构理论以社区知识发展为目标，强调对等的知识发展，要求学生承担责任成为积极的认知者，在社区知识创造中实现个人的认知发展。要求改变传统的教师教授型授课，强调坚持教师的价值引导与学生的主体建构相统一。在具体的探究活动和学习任务开始前要提前向学生设定问题，在完成探究活动后要及时引导学生进行自我归纳和总结，进行知识到能力到价值观的升华，通过不断的点拨、提醒和强调引导学生生成学科核心素养。

《影子》一课在引导学生认识了解基本的字词和结构的同时，还引发学生们思考和想象，进行能力的提升，生成素养。通过让学生们学着说一说，参考“影子就像小黑狗”比喻进行比喻，询问学生弯弯的月亮就像？飘落的树叶就像？圆圆的太阳就像？提升学生的语言运用和审美创造素养。

3.4. 即时评价，全面覆盖

知识建构 12 条原则把评价也作为重要部分，要将评价嵌入教学过程中。教学评价能够及时反应教学的成效，也能提醒教师不断调整完善教学设计。新课标也强调“教-学-评”一体化设计，要求教师在课堂上及时进行评价，便于了解学生的学习成效。《影子》课

堂评价设计了两项活动，一是结合学生的座位让学生说出自己座位的前后左右同学分别是谁？二是提问学生在生活中哪些地方见过影子，在纸上画出生活中的光源和影子。

4. 基于知识建构建构的素养课堂教学策略

结合《影子》的课堂教学设计分析，基于知识建构的素养课堂应该包括以下四个方面的教学策略：

4.1. 创设真实情境，建立学生参与班级知识创造的责任

意义的学习发生在真实的学习情境和任务中，课堂教学要创设真实的情境，加强课程内容与学生切身实际的联系，吸引学生的注意力，唤起学生的好奇心，充分激发学习的兴趣和动机，以便建立学生参与班级知识创造的责任。

4.2. 鼓励分享观点，引导学生做积极的认知者

通过具体的活动和问题可以引导学生在整个课程进行思考与分享，增强问题的现实性、针对性和多样性，引导学生做积极的认知者，民主地分享多样化的观点。

4.3. 加强知识发展，促进学生的自主平等构建

在设置问题了解学生已有经验和认知的基础上，针对存在冲突或者矛盾的内容，引导学生进行自主探究，倡导师生对等的知识发展，在实验和分享中不断构建和更新班级的知识，同时提高学生的个人学习。

4.4. 注重学习评价，关联学生的生活实际

注重课堂互动，通过分组交流和分享等注重学习的交互性，加强学习的评价。学习的评价同时也是为了进一步促进知识的建构，通过与学生生活实际关联的评价促进知识的迁移和素养的生成。

5. 总结

基于知识建构的素养课堂教学策略研究旨在促进学生核心素养的提升。结合知识建构的相关理论和原则，教师课堂教学可以从创设真实情境，建立学生参与班级知识创造的责任；鼓励分享观点，引导学生做积极的认知者；加强知识发展，促进学生的自主平等构建；注重学习评价，关联学生的生活实际四个方面进行探索和完善，在实践中进一步优化教学策略，提高课堂教学效果，提升学生核心素养。

参考文献

中华人民共和国教育部。**义务教育课程方案（2022 年版）**。北京师范大学出版社。
中华人民共和国教育部。**义务教育语文课程标准（2022 年版）**。北京师范大学出版社。
任友群、赵建华、孔晶和尚俊杰(2020)。 国际学习科学研究的现状、核心领域与发展趋势——2018 版《国际学习科学手册》之解析。**远程教育杂志， 38(1)**， 18–27。
张义兵、陈伯栋、Marlene Scardamalia 和 Carl Bereier(2012)。从浅层建构走向深层建构——知识建构理论的发展及其在中国的应用分析。**电化教育研究(09)**，5-12。
赵建华 (2007)。知识建构的原理与方法。**电化教育研究(05)**，9-15+29。
高文、徐斌艳和吴刚(2008)。**建构主义教育研究**。教育科学出版社。
费舍尔、赵建华、尚俊杰和蒋银健.(2022)。**国际学习科学手册 = International handbook of the learning sciences**。华东师范大学出版社。
Marlene Scardamalia、张建伟和孙燕青(2005)。知识建构共同体及其支撑环境。**现代教育技术(03)**，5-13。

人工智能技术支持下的协同知识建构教学模式研究

Research on Collaborative Knowledge Construction Teaching Model Supported by Artificial Intelligence Technology

杨文静*，赵建华
南方科技大学
* kjywj0911@163.com

【摘要】“人工智能+教育”时代的到来，正在引领教育教学理念、模式、方法、过程和技术向更高效、个性化和智能化的方向迈进。人工智能赋能教育，为教育赋予了新的活力和可能性。协同知识建构理论已在日常教学中得到广泛应用，其核心目标在于培养学习者的高阶思维能力。本研究基于人工智能技术的支持，构建协同知识建构教学模式，该模式包括前期准备、个体知识建构、组内协同知识建构和组间协同知识建构四个核心环节，旨在推动协同知识建构过程的优化与升级，进而提升学生的创新思维能力、自主学习和深度学习能力，激发学生的学习兴趣和潜能等。以期促进学习成果的创造和元认知的发展，进而形成共同学习的良性循环，推动教育教学质量的全面提升。
【关键词】 人工智能技术；协同知识建构；教学模式

Abstract: *The arrival of the era of "artificial “intelligence + education” is leading educational and teaching concepts, models, methods, processes and technologies to move in a more efficient, personalized and intelligent direction. Artificial intelligence empowers education and gives education new vitality and possibilities. The theory of collaborative knowledge construction has been widely used in daily teaching, and its core goal is to cultivate learners' higher-order thinking abilities. This study builds a collaborative knowledge construction teaching model based on the support of artificial intelligence technology. This model includes four core links: preliminary preparation, individual knowledge construction, intra-group collaborative knowledge construction, and inter-group collaborative knowledge construction. It aims to promote the collaborative knowledge construction process. Optimize and upgrade, thereby improving students' innovative thinking abilities, independent learning and in-depth learning abilities, and stimulating students' learning interests and potential. In order to promote the creation of learning results and the development of metacognition, thereby forming a virtuous cycle of joint learning and promoting the overall improvement of the quality of education and teaching.*
Keywords: Artificial intelligence technology, collaborative knowledge construction, teaching model

1. 问题提出

在当代教育领域，随着信息技术的飞速发展，人工智能（AI）技术的应用已经成为推动教学创新和提高教育质量的关键因素。特别是在协同知识建构的教学模式中，人工智能技术提供了前所未有的支持和可能性，促进了教与学的深度融合和优化。为培养新时代社会所需人才，《教育信息化 2.0 行动计划》明确指出，

应强化以能力为先的人才培养理念；推动教育理念更新、模式变革、体系重构（教育部， 2018）。2017 年，我国颁布的《新一代人工智能发展规划》中明确提出要发展“智能教育”，利用智能技术加快推动人才培养模式、教学方法改革，构建包含智能学习、交互式学习的新型教育体系（国务院， 2017）。协同知识建构教学模式侧重于学习者在社会互动过程中通过集体智慧共同创造知识，强调知识的社会性和建构性。近年来，多种人工智能工具和平台已经被开发和应用于教育实践中，这些工具不仅能够提供个性化学习路径，还能促进学习者之间的协作和知识共享。尽管人工智能技术在教育中的潜力巨大，但如何有效整合人工智能技术与协同知识建构教学模式，以及这种整合对教学效果的影响，仍然是需要深入研究的问题。本研究将深入分析人工智能技术与协同知识建构教学模式的结合点，以期揭示人工智能技术如何助力构建更为高效、互动和个性化的学习环境，进而推动学习者的主动学习、批判性思维和创造力的发展。

2. 研究现状

2.1. 协同知识建构

Scardamalia 认为知识建构是对社区有价值的观点和思想的产生和不断改进的过程，其目的是通过社区中大于个体贡献之和的集体贡献来实现的（Scardamalia, 1996)。在此基础上，野中郁次郎提出知识转化 SECI 模型，将知识建构的过程分为个人知识建构与社会知识的协作建构两个维度。Gerry Stahl 等认为，个体独立的学习并不能使承担的任务或问题得到解决时，就需要通过协作进行知识意义的共同建构，也就开始了学习的社会化过程（Stahl, 2018)。协作知识建构就是在此基础上提出的，是指个体在特定的组织中互相协作、共同参与某种有目的的活动，最终形成某种观点、思想、方法等智慧产品的过程（谢幼如， 2010)。协同知识建构强调知识是学习共同体在学习社区的社会交互中不断产生和发展的（杜岩岩，2020），师生、生生、学生和学习资源之间，通过交互进行知识共建。任剑锋认为协同知识建构有不同层次，浅层协同知识构建在信息共享条件下即可发生，但较高层次的协同建构发生在“实时共做”层面，（任剑锋， 2021），换而言之，如果没有教学策略的推动，即使引入智能技术，并将学习者置于共同交流的空间，高质量的协同知识建构也很难实现。Vandenberg 等认为，学习者通过在集体环境中分享自主建构的知识，并通过与其他学习者和教师进行协商和反思，进而实现新知识的建构和个体认知水平的提升（Vandenberg, 2021）。周平红等在对学生协同知识建构的社会认知演化轨迹研究中发现，与边缘组学生相比，核心组形成了高层次认知网络结构（周平红, 2021）。马海华认为，协同知识建构是师生之间、生生之间、学生和学习资源之间，通过深度交互和思维的深度加工进行知识共建，通过最终形成的群体智慧化成果实现创造性学习和深度学习，从而促进学习者知识迁移能力、复杂问题解决能力等高阶思维的发展（马海华， 2023)。协同知识建构强调的是先由学习群体公平、自由地进行思想与观念的交流，在“产生冲突—解释与澄清—冲突解决”的过程中达成一致，形成学习群体认可的公共知识，再由群体中学习者个人将这些公共知识内化为个体知识。学生通过彼此的协作和会话，可以更好地帮助他们实现对知识的理解和意义建构，有助于学习者的深度学习。

2.2. 协同知识建构教学模式

目前关于协同知识建构模式的教学流程，国外学者 Gunawardena 基于学习者交互层面，将协同知识建构过程分为共享和比较信息、发现和分析观点间的差异、

通过提出新建议共同建构知识、成员间达成共识并应用新知识四个基本阶段（Gunawardena, 1997)。Stahl 则从社会建构主义对学习的理解出发，提出了包括个人建构和社会知识建构的双循环模式的 11 个阶段（Stahl, 2018)。Beers 等人认为知识建构包括外化、内化、协商和集成四个阶段（Beers, 2005)。国内学者谢幼如在基于网络学习环境研究中提出“共享—论证—协商—创作—反思”的协同知识建构过程模型（谢幼如，2008)。赵海霞从问题解决学习目标出发，认为协同知识建构包括“学习共同体创建、基于认知冲突的协作知识建构和基于整合”三个阶段（赵海霞，2013）。李海峰等基于“场”和自组织理论，提出“问题识别、问题分析、形成策略、方案生成、评估实践”的在线协作知识建构模型（李海锋，2018）。通过这些模式可以看出，协同知识建构旨在调动学习者作为学习主体的积极性，强调通过学习者之间的协作和知识共建来实现深度学习，并促进学习者知识迁移能力、问题解决能力和合作精神。

3. 人工智能技术支持下的协同知识建构模式

人工智能与教育教学的紧密结合，不仅重塑教学环境、更新教学理念、调整师生关系、优化教学方式、改进评价方式等，更将推动教学模式的根本性转变。人工智能教育注重基于学习知识开展高级认知思维的过程，能有效促进学生深度学习并提升创新能力（丁美荣，2021）。随着人工智能技术的进一步发展，如何利用自然语言处理、视觉分析以及序列模式分析、聚类分析等智能算法辅助赋能课堂诊断与洞察能力成为研究热点（吴南中，2019），如基于教学切片法、应用人工智能技术实现对课堂的教学事件识别与教学阶段划分等(孙众，2021)，结合人工智能技术改变课堂环境和教学方式，为教学模式的转变提供了解释性和指向性建议。往常知识建构分析方法多为人工采集数据并进行处理分析，数据量小，分析结果反馈滞后，难以实现对学习过程的实时管理；随着教学环境逐渐智能化，获取到的学习数据不断增加，进一步加大助学者准确分析、诊断学习活动的难度。此外，现有的一些基于特定理论视角的分析方法，其计算产生的分析结果对协同知识建构过程特征的解释还有不足。因此，人工智能技术的快速发展，为破解上述问题提供了可能。在人工智能技术的助力下，协同知识建构教学模式的核心环节涵盖前期准备、个体知识建构、组内协同知识建构以及组间协同知识建构（如图 1 所示）。前期准备阶段，教师发挥主导作用，致力于为学生构建一个智能化的在线教学环境。进入个体知识建构阶段，学生的自主学习和深度学习能力得到重点培养。紧接着的组内协同知识建构阶段，是对已建构知识的内化过程，学生在此阶段表达自己的观点，通过与他人交流、辩论、最终达成小组共识，并形成作品。至于组间协同知识建构阶段，则将小组作品进行展示，通过与其他小组做对比，并进行交流论证，最后对作品进行评价反思，修改完善。值得一提的是，在后三个环节中，均设有评价环节，其目的在于通过持续的评价过程，发现问题并寻求解决方案，从而确保学生更好地掌握知识。

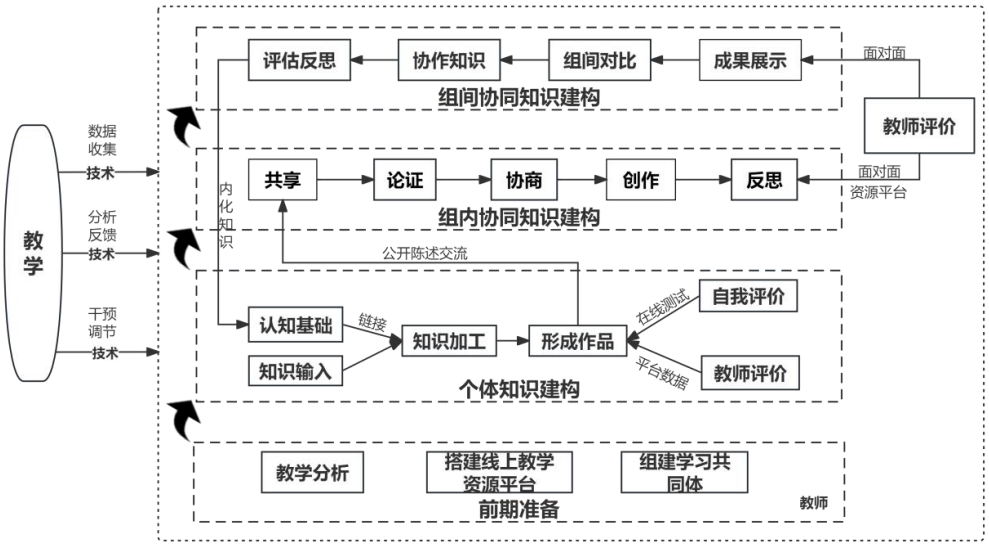


图 1 人工智能技术支持的协同知识建构教学模式

3.1. 前期准备

1. 教学分析

在利用人工智能技术推动协同知识建构的教学过程中，一方面教师需具备坚实的知识建构理论基础，深入理解和把握协同知识建构的核心教学原则，教师还应展现出对新教学模式和风格的快速适应能力，确保教学质量与效果的持续提升。另一方面，教师需要制定一套新的、适应于目标学生的教学策略，能够很好的应用人工智能技术等教学工具，帮助学生开展高质量的教学活动。从知识建构的角度出发，教学模式应着重于推动学生的“深度学习”，这要求学生在学习过程中进行深入地探索和研究。因此，教师在设计教学活动时，不能仅仅局限于课程教材，而是要根据课程目标，整合各种教学资源，为学生提供更多探究性学习的时间和空间。打破时间和空间的限制，教师可以充分发挥实践的优势，融入跨学科知识，进而促进学生创新学习的建构。在开始教学活动之前，为了保障教学质量，教师需要先了解学生的学习动机、学习特点以及学习能力等方面的情况。在掌握了学生的学习特征后，教师才能进入下一个教学准备阶段。

2. 搭建线上教学资源平台

教师作为学习环境和学习资源的构建者，在协同知识建构教学模式的框架下，首要任务是为学生的知识建构过程构建一个丰富多样的教学资源平台。这一平台需能够提供智能化的学习环境，以支持学生的自主学习和深度学习。此外，教师还需构建完善的知识网络，并为学生提供参与式的技术工具，以引导并促进学生在线上和线下环境中有效展开学习活动。这些举措将共同助力学生在协同知识建构的过程中实现全面的发展和提升。为贴合学生的认知水平和满足学生个性化学习需求，教师可以借助人工智能技术，以图片、声音、虚拟实验视频等多种形式呈现学习资源。此外，还需提供导学、课件、拓展资料、习题、测验等配套资源，以确保学生能够得到充分进行个性化学习体验。

随着“文心一言”“华为云”和“字节跳动”等生成式人工智能平台的相继推出，我们有必要考虑将教学资源平台与这些先进的生成式人工智能技术相融合。这种融合将为学生带来更加个性化且高效的学习体验。传统教学资源平台主要依赖人工整理和上传资料，而生成式人工智能则能自动从互联网上抓取、整理与教

学密切相关的内容。此举不仅能显著提升教学资源的更新速度，更能确保资源的准确性和时效性。通过深度分析学生的学习行为和兴趣偏好，生成式人工智能平台能为学生推荐适合他们的学习资源和课程，进而提升学习效率并激发他们的学习兴趣和动力。此外，平台还能智能分析学生的学习数据，为教师提供精准的教学反馈和建议，使教师能更有针对性地调整教学策略和方法，从而提高教学效果。教学资源平台与各类教学软件的融合，将为教育领域带来更加智能化、个性化的解决方案，推动教育领域的数字化转型和创新发展。

3. 组建学习共同体

在构建学习共同体时，分组策略的选择至关重要。在教学实践中，常见的分组方式包括随机分组、异质分组和同质分组。学习小组的设立有助于形成紧密的学习团体，促进深度交流，为协同知识构建提供组织基础。在人工智能技术的支持下，我们可以根据学生的认知水平、知识储备、性格特点及性别等因素，进行更为精准的分组。异质分组能够促进组内互补，增强组间竞争公平性，但也可能导致低水平学生边缘化。而同质分组则能增强学生的归属感和认同感，促进合作学习，减少边缘化现象，但可能加剧组间竞争差异。为了提升协同知识建构的效果，教师在分组时应灵活多变，根据具体学习任务选择合适的策略。例如，对于复杂程度较高的学习任务，可以适当增加小组规模；对于个性化较强的学习任务，建议采用同质分组，并考虑社交关系和性别平衡。对于综合性学习任务，如文献调研、社会调查、技术评估、文稿撰写、沟通协调和现场汇报等，推荐采用异质分组，并明确任务分工，以充分发挥每个成员的优势，提升他们的价值感和成就感。

3.2. 个体知识建构

在此阶段，教师作为一个引导者，可利用已构建的教学资源平台的相关功能，结合人工智能技术对收集的学生学习数据进行深入分析。根据分析结果，教师应有效地指导和监督学生，以促进其学习质量的进一步提升。协同知识建构教学模式强调以多样化的学习任务为导向，教师需设计具有一定挑战性的学习任务，这些任务的难度应略高于学生的现有知识水平，例如，练习闯关游戏等。同时，任务应具备开放性和探究性，以促进师生之间和学生之间的协作与交流。学生应从教师提供的在线学习资源平台和自身的认知基础上汲取知识，整理要点，明确逻辑关系，并对所学知识进行深入加工。在此过程中，学生应能够发现问题，并通过再次查阅资料来寻找解决问题的方法。

3.3. 组内知识建构

协同知识建构在团队中占据重要地位，它涵盖资料共享、观点阐述、质疑论证、协商创作以及评价反思这几个核心环节。首先，小组学生依据各自的分工，将搜集的资料及关于问题的见解在小组内部进行分享。在此过程中，学生可以利用思维导图、概念图等工具，将内隐的知识外显化，以便组内的交流和理解。由于学生的知识背景和认知结构存在差异，对于同一问题，会有不同的见解，这种差异为组内质疑和论证提供了可能。通过相互质疑和深入论证，组内成员间的认知冲突得以激发，促使学生进行更加深入地思考。这一过程不仅有助于修正和完善个人观点，还能推动深度学习的实现。经过充分质疑和论证后，组内成员对问题形成初步共识，并将这一共识通过可视化的方式呈现为作品。这一作品不仅体现了团队协同知识建构的成果，也为后续的组间协同知识建构奠定了基础。最后，进行反思，反思包括对整个合作与分工过程的回顾与总结，以及对个人知识建构策略和水平的深入剖析，旨在学生能够在后续的协作学习中更加高效地完成任务，

不断提升团队的整体效能。

3.4. 组间知识建构

组间协同知识建构进一步促进了连接的扩展和增强，主要包括成果展示、组间对比、协作知识、评估反思环节。展示阶段是观点汇聚、共享理解的过程，小组依次展示组内协同知识建构的成果和协作学习过程。该过程旨在让同学和教师了解每个小组的知识建构过程及其成果，以便学生汲取优秀成果和经验，实现知识的共享。在小组展示后，不同小组之间进行深入的交流和辩论，从而形成集体范围的组间协同知识建构。此外，教师也会提出相关问题，引发学生的思考和交流。在这一环节中，小组成员需要充分调动自己的思维，结合已有的知识背景进行有力的论证，积极参与辩论，为集体的知识建构贡献自己的智慧和见解。经过组间的质疑和辩论，各小组将不断完善自己的观点。若通过协商可以形成共同理解时，则达成了共识。若无法达成共识，此时就需要教师发挥主导作用，进行必要的引导和协调。此外，只靠学习者的协同知识建构有时难以达到对某些重难点知识的高水平理解和掌握，因此有必要发挥教师的主导作用来及时讲解。为了确保学生的学习效果，教师应根据学生在协同知识建构过程中的表现进行形成性评价，及时识别学生尚未掌握的知识点，并进行有针对性地补充讲解。这将有助于学生更全面、更深入地理解和掌握所学内容。

3.5. 评价反思

在协同知识建构的过程中，评价与反馈占据着至关重要的地位。为确保评价方法的严谨性与有效性，须考虑以下几方面。第一，评价需具备多维度（态度、能力等）和过程性，涵盖从课前到课后的全方位考量。在课前阶段，应对学生的知识水平和能力基础进行评估；同时，考察学生在小组内协同知识建构过程中的贡献比例和结果。进入课堂环节，应重点评估学生在课堂活动中的参与质量以及小组间协同知识建构的质量。课后阶段，则需对学生的自主学习情况和深入学习能力进行考量。这些方面的综合评估将构成评价方法的核心要素，确保评价结果的全面性和准确性。二是评价的多主体性。评价可以通过教师、学生、督导专家等多种主体结合开展，如组内协同知识建构的评价可以通过学生互评和教师评价来实施，课后评价可以包含教师评价和学生自评。三是定性评价和定量评价。定性评价即教师或学生对小组成员的协作表现、建构成果等进行直接的经验性的观察和分析，并做出口头或文本形式的定性结论和反思总结；定量评价可通过对照评价标准进行量化计分来实施。其次，教师根据评价方法，选择能够支撑评价活动的人工智能技术工具和平台。教师依托人工智能技术工具进行多维度学情数据统计（如学情统计数据、学习行为数据等），并对数据进行有效提取和分析，及时总结基于各平台的学情大数据，给出诊断性评价。最后，评价结果的反馈同样至关重要。我们应通过个人反馈、组内反馈、组间反馈、平台反馈等多种形式提供详细的学情反馈给学生。这不仅能够帮助学生反思自己的学习过程与协同知识建构策略，还能帮助他们发现问题并提出改进措施。同时，这种反馈也有助于学生审视自己的认知策略与认知过程，促进自主学习元认知的发展。

4. 结语

人工智能进入学校、课堂，培养新时代数字素养人才已经成为各类学校教育发展的新趋势，传统教师角色发生改变，教师的育人角色更加凸显，现今，人机协同的教学将会把学生的创新能力和自主学习能力放在首位。教师要不断提升教育理论水平，同时要掌握新一代智能技术和教育教学深度融合创新的方法，为教

学效果的提升提供方法支撑。本文通过对协同知识建构相关理论的研究和解读，再结合人工智能的技术特点，构建了多视角、技术增强的协同知识建构教学模式，旨在增强学习者的知识理解、创新思维和实践能力。该模式能够有效地结合人工智能技术和协同知识建构理论，为学习者提供更加个性化、高效、系统的学习体验。未来，我们将继续深入研究和探索该模式的应用和发展，为教育领域的创新和发展贡献更多的力量。

参考文献

国务院．(2017)．*新一代人工智能发展规划的通知*．http://www.gov.cn/zhengce/content/2017-07/20/ content_5211996.htm.

教育部．(2021) *教育部关于印发《教育信息化 2.0 行动计划》的通知*．http://www.moe.gov.cn/srcsite/ A16/s3342/201804/t20180425_334188.html,2018-04-18.

丁美荣，王同聚．(2018)．人工智能教学中“知识建构、STEM、创客”三位一体教学模型的设计与应用．*电化教育研究*，42(04)：108-114.

杜岩岩，和黄庆双．(2020)．在线深度学习的发生机理与促进策略．*中国高教研究*，322(6)：58-63.

李海峰，王炜．(2021)．面向问题解决的在线协作知识建构．*电化教育研究*，2018,(1)：36-41.

马海华，邓淼磊，和董恒．(2023)．基于联通主义的协同知识建构教学模式研究与实践．*河南工业大学学报(社会科学版)*，39(05)：104-111.

彭绍东．(2015)．混合式协作学习中知识建构的三循环模型研究．*中国电化教育*，(9)：39-47.

任剑锋．(2021)．促进高层次协同知识建构的CSCL可视化支持系统研究．*首都师范大学学报(社会科学版)*，262(5)：56-64.

孙众，吕恺悦，和施智平．(2021)．TESTII 框架：人工智能支持课堂教学分析的发展走向．*电化教育研究*，42(2)：33-39,77.

韦怡彤，王继新，和丁茹．(2019)．混合式学习环境下深度学习导向的协同知识建构模式研究：以《教育技术学导论》课程为例．*中国电化教育*，392(9)：128-134.

吴南中，夏海鹰，和张岩．(2019)．信息技术推动教育形态变革的逻辑、形式、内容与路径．*中国电化教育*，(11)：24-33.

谢幼如，宋乃庆．(2008)．基于网络的协作知识建构及其共同体的分析研究．*电化教育研究*，(4)：38-42.

谢幼如，宋乃庆，和刘鸣．(2010)．基于问题的网络课堂协作知识建构模式．*电化教育研究*，(1)：38-46.

赵海霞．(2013)．网络环境下基于问题的协作知识建构设计与实践——以大学生“结构化学”课程教学改革为例．*中国电化教育*，(1)：100-105.

周平红，周洪茜，和张屹．(2021)．深度学习视域下学习者协同知识建构历程的社会认知网络特征分析．*电化教育研究*，42(9)：99-107.

Beers, P. J., Boshuizen, H. P. E., Kirschner, P. A., & Gijselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. Computers in Human Behavior,21(4), 623-643.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. Journal of educational computing research, 17(4), 397-431.

Scardamalia, M., & Beretier, C. (1996). Computer support for knowledge-building communities. In T.

Koschmann (Ed.), *CSCL:Theory and Practice of an Emerging Para-digm*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Stahl, G. (2013, April). A model of collaborative knowledge-building. In *International conference of the learning sciences* (pp. 70-77). Psychology Press.

Vandenberg, J., Zakaria, Z., Tsan, J., Iwanski, A., Lynch, C., Boyer, K. E., & Wiebe, E. (2021). Prompting collaborative and exploratory discourse: An epistemic network analysis study. International Journal of Computer-Supported Collaborative Learning, 16(3), 339-366.

College of Design Students' Intention to Use Generative AI Tools: The Role of Intrinsic and Extrinsic Motivation

Hsi-Hsun Yang
Department of Digital Media Design, National Yunlin University of Science and Technology, Yunlin, Taiwan
jimmy@yuntech.edu.tw

Abstract: *This study proposes a hypothesis model that combines the Technology Acceptance Model (TAM) and motivation theory to explore the behavioral intentions of design college students in using generative AI (GenAI) tools. A questionnaire survey method was employed, randomly sampling college and university students in Taiwan, resulting in 214 valid samples. The analysis showed that the model explained 50.5% of the variance in Behavioral Intention to Use (BIU), confirming that both Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) are significant predictors of BIU. Moreover, intrinsic motivation (IM) was found to have a more substantial influence on PEOU and PU than extrinsic motivation (EM). This exciting discovery holds significant breakthroughs for both academic researchers and practitioners.*

Keywords: GenAI, intrinsic motivation, extrinsic motivation, TAM

1. Introduction

The year 2023 has been dubbed the "Dawn of Generative AI," with GenAI capable of creating various data types, such as images, videos, audio, text, and 3D models. It also has the potential to transform higher education teaching methods, assist in scientific research, and influence organizational management. Among the most notable examples of GenAI is ChatGPT, which has rapidly increased in just two years, along with other AI tools. The impact of these tools has been mixed; a study by Eloundou et al. (2023) analyzed a database of 1,016 professions and found that approximately 80% of the U.S. workforce is at risk of experiencing significant impacts, around 19%, due to the introduction of AI tools, especially ChatGPT. Meanwhile, data from the McKinsey Global Institute (MGI) suggests that by 2030, 3% to 14% of the global workforce may need to acquire new skills or change occupations (Manyika et al., 2017). Concerns about job replacement due to the threat of AI technology to employment rights are increasing.

Several theories have been introduced to effectively explain the behavioral patterns of users willing to use specific technologies, among which the most famous is the Technology Acceptance Model (TAM) by Davis (1989). Davis et al. (1989) and Hung et al. (2005) highlighted that intention determines the extent to which users utilize technology, which is PU and PEOU influence. Scherer and Teo (2019) noted that in some studies, PU is a stronger predictor than PEOU; however, the opposite was found in other studies. Similarly, research has shown a positive and significant relationship between PU and behavioral intention (BI), while other studies have not found evidence to support this relationship.

Model (TAM), this study incorporates a crucial predictor in learning - motivation theory (Donnermann et al., 2021), attempting to observe the influences of EM and IM. IM refers to the motivation to engage in an activity due to the internal satisfaction it brings. It is a significant force affecting learning, adaptation, and capability, making it a powerful source of behavioral motivation (Deci & Ryan, 1985). Conversely, EM is driven by engaging in activities that lead to valuable outcomes, such as improved job performance, salary increases, and promotions.

The TAM is used because, compared to other acceptance models, it offers comprehensiveness, explanatory power, and applicability to consumer environments (Verdegem & De Marez, 2011). Furthermore, intrinsic and extrinsic motivations are two critical factors in students' use of information technology (Venkatesh, 1999). Therefore, this study combines the two theories to expand the explanation and construction of the technology acceptance model from various social and psychological perspectives.

2. Methodology

This study employs a cross-sectional survey method targeting students from colleges of design across Taiwan, including undergraduates, master's, and doctoral students who have experience using GenAI for design-related tasks. In addition to ChatGPT, generative AI tools like image generators Midjourney, Stable Diffusion, DALL-E 3, Imagen, and video generators such as Runway and Lumen5 are threatening the job security of (future) designers. The sample was randomly selected, and the survey was distributed via social media platforms (e.g., Facebook and Line). The survey was open for responses from January 8, 2024, for one week. After collecting the questionnaires, a total of 253 samples were gathered. Following a review to exclude non-student participants, 214 valid questionnaires remained. The analysis was conducted using IBM SPSS Statistics 25 and SmartPLS 4.

This study primarily utilized a structured questionnaire as its research tool, based on the research dimensions. The questionnaire content included (1) demographic variables such as gender, age, years of experience, educational level, and frequency of using AI tools; (2) nineteen questions designed around perceived ease of use (PEOU), perceived usefulness (PU), intrinsic motivation (IM), extrinsic motivation (EM), and behavioral intention to use (BIU) (Deci & Ryan, 1985; Deci et al., 2001; Venkatesh et al., 2012; Duong et al., 2023), to understand each student's intention to use GenAI tools. A Likert 7-point scale was used for scoring.

3. Results

Based on the questionnaire data analysis, the distribution by frequency showed that there were 144 female participants (67.3%) and 70 male participants (32.7%). Regarding age, 94 people (43.9%) were aged 18-20, 85 people (39.7%) were aged 21-25, 16 people (7.5%) were aged 26-30, and 19 people (8.9%) were aged 31 and above. In terms of educational level, there were 163 undergraduate students (76.2%) and 51 postgraduate students (master's or doctoral) (23.8%). Regarding the frequency of using GenAI tools per week, 117 people (54.7%) used them 1-3 times, 57 people (26.6%) used them 4-6 times, 15 people (7.0%) used them 7-10 times, and 25 people (11.6%) used them more than 11 times.

A total of 214 valid questionnaires were collected. First, the normality test of the collected samples showed the highest average score was 5.76, and the lowest was 3.71, with kurtosis ranging from -0.935 to 1.470 and skewness from -0.898 to 0.080. Cronbach's alpha ranged from .700 to .871, and the Average Variance Extracted (AVE) from .624 to .721, all within reasonable ranges, indicating that the data for these five dimensions can be considered normally distributed. Furthermore, the Variance Inflation Factor (VIF) values ranged between 1.472 and 2.532, all below the threshold of 5 (Hair et al., 2011), indicating that the collinearity issue within the constructs of the study's framework is not severe. Therefore, collinearity will not adversely affect the estimation of path coefficients in this study's model.

Table 1.

Regression coefficient.

	IV→DV	Original sample	T statistics	Result	R^2	f^2	95%CI LL	95%CI UL
H1	IM -> PEOU	0.487	6.452	Supported	0.352	0.170	0.338	0.637
H3	EM -> PEOU	0.135	1.597	Unsupported		0.013	-0.037	0.298
H2	IM -> PU	0.422	6.546	Supported		0.232	0.290	0.541
H4	EM -> PU	0.388	6.141	Supported	0.697	0.227	0.268	0.517
H5	PEOU -> PU	0.128	2.607	Supported		0.035	0.033	0.225
H6	PEOU -> BIU	0.136	2.251	Supported	0.505	0.025	0.021	0.254
H7	PU -> BIU	0.625	1.133	Supported		0.536	0.514	0.73

The results of the path analysis in Table 1 confirm the significance between constructs; for example, intrinsic motivation (IM) significantly impacts Perceived Ease of Use (PEOU) with β=0.487, t=6.452, while extrinsic motivation (EM) does not have a significant effect on PEOU with β=0.135, t=1.597. Their combined values explain 35.2% of the

variance in PEOU. IM (β=0.422, t=6.542), EM (β=0.388, t=6.141), and PEOU (β=0.128, t=2.607) significantly influence Perceived Usefulness (PU), and their combined values explain 69.7% of the variance in PU. PEOU (β=0.136, t=2.251) and PU (β=0.625, t=1.133) significantly affect Behavioral Intention to Use (BIU), and their combined values explain 50.5% of the variance in BIU, as shown in Figure 1 and Table 2.

Figure 1.

SEM Model.

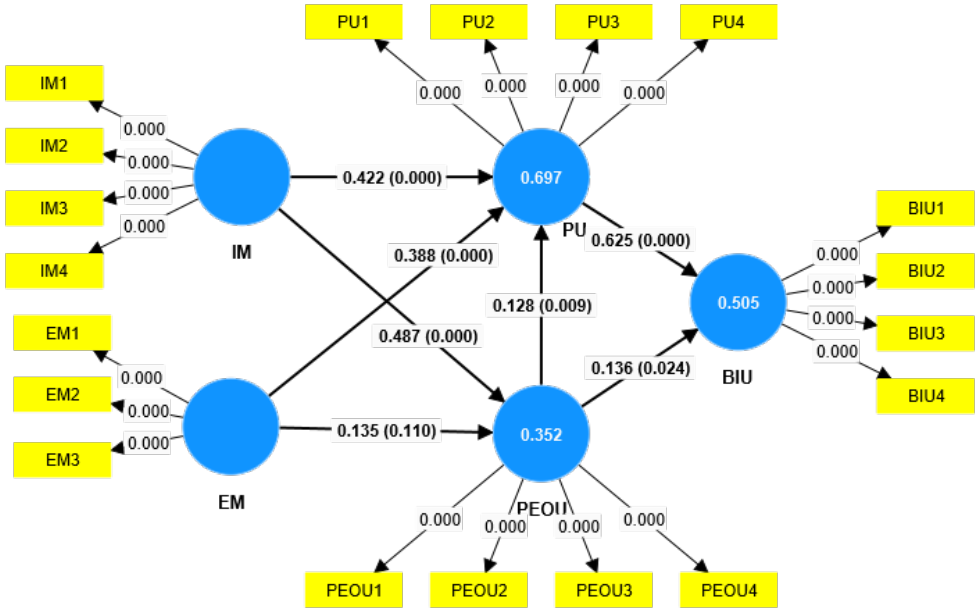


Table 2.

Mediating effects.

IV	Mediator	DV	Direct effect	Indirect effect	Total	Mediation Analysis
IM	PEOU	BIU	0.487*** (6.452)	0.066* (1.981)	0.369*** (7.918)	Complementary Partial mediation
	PU		0.422*** (6.546)	0.263*** (5.425)	0.369*** (7.918)	Complementary Partial mediation
EM	PEOU	BIU	0.135 (1.597)	0.018 (1.191)	0.272*** (5.717)	No effect
	PU		0.388*** (6.141)	0.243*** (5.289)	0.272*** (5.717)	Complementary Partial mediation
PEOU	PU	BIU	0.128* (2.607)	0.080* (2.543)	0.216*** (3.803)	Complementary Partial mediation

Note: $^{*}p < 0.05$, $^{***}p < 0.001$

4. Discussion and conclusion

Chapter 3's statistical analysis indicates that PEOU and PU significantly and positively impact the intention to use GenAI tools. Moreover, IM significantly and positively affects PEOU and PU, and EM influences PU. The study identified three mediating effects: PEOU and PU mediate the relationship between IM and BIU, PU mediates the relationship between IM and BIU, and PU mediates the relationship between PEOU and BIU. The overall structural model demonstrates a good fit, with six hypotheses supported and one not supported, and for the endogenous latent variables BIU, PEOU, and PU, the R2 values reached 50.5%, 35.2%, and 69.7%, respectively. This shows that the study's research model effectively explains the variance of these variables.

The results of this study demonstrate that Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) significantly and positively affect design college students' intention to use artificial intelligence tools. PEOU and PU have been identified as critical factors influencing behavioral intention by Davis et al. (1989) and Hung et al. (2005). This implies that when design professionals believe that AI tools are effective in completing tasks and are convenient, their willingness to use them increases. Additionally, the mediating effects in Table 2 reveal that the impact of intrinsic motivation (IM) is more significant than that of extrinsic motivation (EM). This suggests that efforts to reduce learning anxiety might help improve students' perceptions of artificial intelligence through intrinsic motivation (for example, including engaging elements), thereby enhancing learning intentions (Wang et al., 2022).

This study makes three main contributions. First, the empirical results provide additional evidence clarifying the relationship between design college students and AI tools under integrating the TAM and motivation theory. Second, the study confirms the importance of PEOU and PU for technology acceptance and usage intentions, further revealing how motivation influences BIU through PEOU and PU. This highlights the necessity of considering motivational factors in technology acceptance research. Third, this research offers a deeper understanding of the acceptance and usage behavior of artificial intelligence technology, aiding in the theoretical comprehension of how design college students form behavioral intentions towards adopting emerging technologies.

References

Hung, S. Y., Liang, T. P., & Chang, C. M. (2005). A meta-analysis of empirical research using TAM. *Journal of Information Management*, *12*(4), 211-234.

Scherer, R., & Teo, T. (2019). Unpacking teachers' intentions to integrate technology: A meta-analysis. *Educational Research Review*, *27*, 90-109

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former eastern bloc country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin*, *27*(8), 930–942. https://doi.org/10.1177/014616720127800

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Duong, C. D., Vu, T. N., & Ngo, T. V. N. (2023). Applying a modified technology acceptance model to explain higher education students' usage of ChatGPT: A serial multiple mediation model with knowledge sharing as a moderator. *The International Journal of Management Education*, *21*(3), 100883. https://doi.org/10.1016/j.ijme.2023.100883

Wang, Y. M., Wei, C. L., Lin, H. H., Wang, S. C., & Wang, Y. S. (2022). What drives students' AI learning behavior: A perspective of AI anxiety. *Interactive Learning Environments*, 1-17.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. arXiv preprint arXiv:2303.10130.

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). A future that works: AI, automation, employment, and productivity. McKinsey Global Institute Research, Tech. Rep, 60, 1-135.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

Donnermann, M., Lein, M., Messingschlager, T., Riedmann, A., Schaper, P., Steinhaeusser, S., & Lugrin, B. (2021). Social robots and gamification for technology supported learning: An empirical study on engagement and motivation. *Computers in Human Behavior*, *121*, 106792

Venkatesh, V. (1999). Creation of favorable user perceptions: Exploring the role of intrinsic motivation. *MIS quarterly*, 239-260.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*(2), 109-134.

A Systematic Review of AI-supported Public Speaking Instruction in the Past Decade

Ping Lu, Chunping Zheng*
Beijing University of Posts and Telecommunications
* zhengchunping@bupt.edu.cn

Abstract: *Artificial intelligence refers to the science and engineering of making intelligent machines, especially intelligent computer programs, capable of performing tasks that typically require human intelligence. It has been applied in public speaking instruction to improve learners' public speaking ability, engagement and alleviate public speaking anxiety. This review study investigates 20 empirical studies in Web of Science database related to AI-supported public speaking instruction from 2014 to 2023. A coding framework was formulated for conducting a systematic literature review, and content analysis explored the general research trends, research approaches, research foci, the applied AI technologies and benefits and challenges of AI-supported public speaking instruction. According to the reviewed studies, many countries have conducted research in this field and higher education is the typical research setting. Researchers preferred mixed research approach to address the research questions. The main research foci include learners' psychological characteristics while using AI technologies or tools, performance testing of AI technologies or tools, affordance of AI technologies for public speaking ability and affordance of AI technologies for non-linguistic knowledge and skills. Based on our research findings, we conclude that AI technologies has the potential to enhance learners' public speaking in various aspects. However, there are still some challenges to be addressed in order to effectively implement the AI technologies in public speaking instruction.*

Keywords: public speaking instruction, AI technologies, systematic review

1. Introduction

With the rapid advancements and innovations of technologies, the pervasive adoption of artificial intelligence (AI) in the realm of education has demonstrated an escalating trend, specifically within the domain of public speaking instruction. Artificial intelligence (AI) refers to a field of study that focuses on the use of machines that behave and think in either an ideal manner or in a similar manner to humans (Russell & Norvig, 2016). These machines continuously adapt to their surroundings, resembling humans in their ability to accomplish tasks. Initially, AI was introduced to streamline and enhance the completion of tasks traditionally performed by humans. Consequently, individuals were able to optimize their time and deploy resources towards other pursuits (Kim, Merrill, & Collins, 2021). Public speaking is one of the typical activities of oral expression, reflecting the learners' complex oral expression ability. Different form other language abilities, public speaking requires students to comprehensively apply language skills of listening, speaking, reading and writing, and poses higher requirements for students' language proficiency, cross-cultural ability, critical thinking ability and research ability (Jie & Jin, 2022). Therefore, it is necessary to pay attention to the cultivation of learners' public speaking ability. In today's rapidly evolving technological era, AI has the potential to revolutionize the way of education. By harnessing the power of AI, public speaking instruction can become more personalized, interactive and efficient, contributing to improved learning outcomes for students. Several review articles have addressed the affordance, opportunities and challenges related to the application of AI technologies in public speaking instruction and it has been shown to be effective to enhance learners' engagement and motivation during conducting oral presentations (Barrett & Liu, 2016).

The current study reviewed 20 empirical studies related to AI-supported public speaking instruction from 2014 to 2023, retrieved from Web of Science database. Content analysis was employed to figure out the general research trends, research approaches, research foci and benefits and challenges associated with AI-supported public speaking instruction. To fulfil the research purpose, the following four research questions (RQs) were addressed:

RQ1: What are the general research trends (including yearly publications, target languages, research sites and settings) of the reviewed empirical research?

RQ2: What are the research methodologies (including research approaches, data collection methods and data analysis methods) employed in the reviewed publications?

RQ3: What kinds of AI technologies have been applied in the reviewed studies?

RQ4: What are the research foci of the reviewed publications?

RQ5: What are the potential benefits and challenges of AI-supported public speaking instruction?

2. Methodology

2.1. Literature Research

The author manually screened the studies retrieved from Web of Science between 2014 and 2023 based on the titles, abstracts and keywords. Two sets of keywords were used for literature screening: (1) artificial intelligence, AI; (2) public speaking, oral presentation, presentation. Articles containing information related to these sets of keywords were identified for further analysis. Moreover, we only retained original research articles and excluded book reviews, commentaries and editorial materials. A total of 20 empirical studies were eventually retained based on their relevance to our research purpose.

2.2. Coding Scheme

The coding scheme used in this study is depicted in Figure 1.

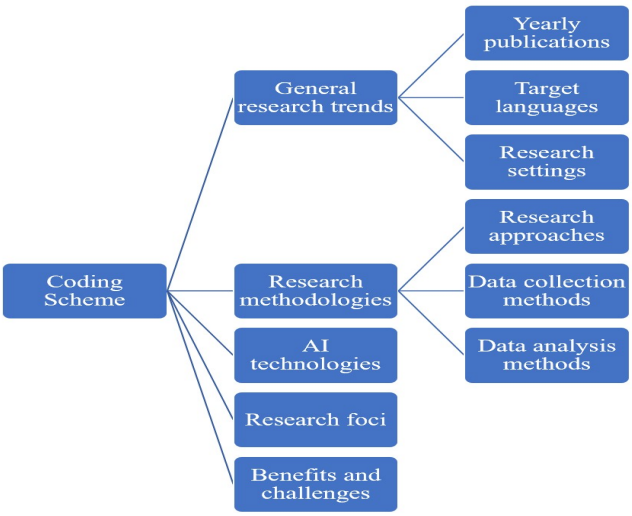


Figure 1. Coding Scheme

3. Results

3.1. The General Research Trends

3.1.1 Number of empirical studies published from 2014 to 2023

As depicted by Figure 2, 20 empirical studies were published from 2014 to 2023. Research on the application of artificial intelligence technologies in public speaking instruction exhibited a fluctuating state from 2014 to 2023, which corresponded to the undulating development of AI technology in the past 10 years. As AI begins to show their unique advantages in the field of public speaking instruction, and AI technologies continue to optimize and upgrade, there is no doubt that more empirical researches will be conducted in the future.

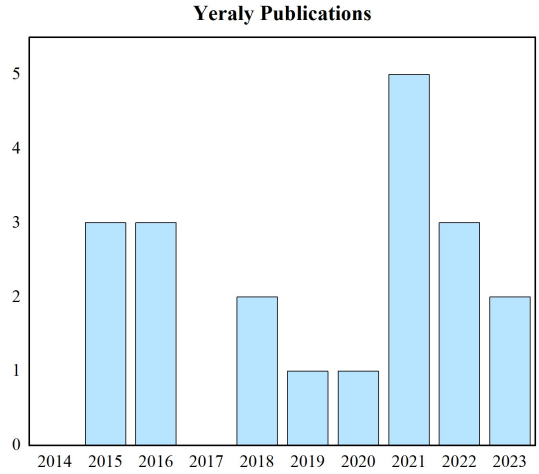


Figure 2. Number of empirical studies published from 2014 to 2023

3.1.2 Target languages studied in the reviewed publications from 2014 to 2023

As exhibited in Figure 3, most of the research did not explicitly indicate the types of target languages involved and English is the primary target language among the literature that has clearly specified the target language, which is closely related to the fact that English is a universal language or lingua franca.

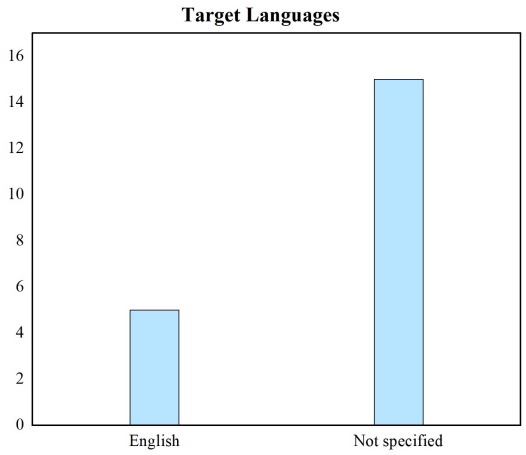


Figure 3. Target languages studied in the reviewed publications from 2014 to 2023

3.1.3 Research sites and settings of the studies from 2014 to 2023

Based on the author's geographical distribution and the research sites for conducting the research, it is found that Chinese and American researchers have been relatively active in this field from 2014 to 2023. In addition, it is evident that numerous countries around the world pay attention to the application of artificial intelligence technologies in public speaking instruction. As shown in Figure 5, most of the empirical studies were conducted in higher education settings whereas there are only 2 studies in secondary education settings. The reason for the above-mentioned phenomenon may be that the application of artificial intelligence technology in higher education has high controllability.

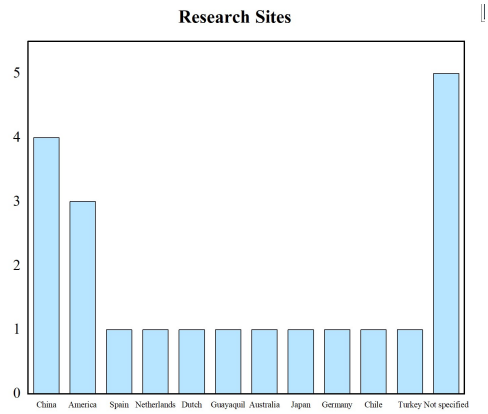


Figure 4. Research sites of the studies from 2014 to 2023

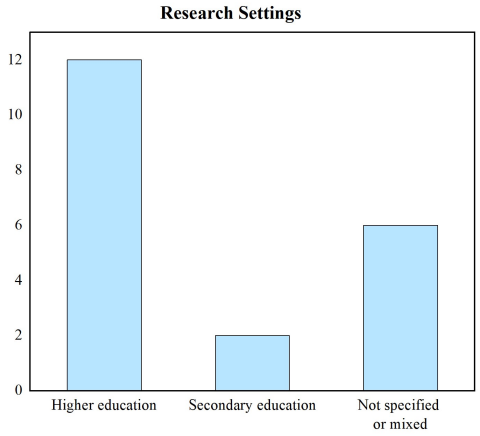


Figure 5. Research settings of the studies from 2014 to 2023

3.2. Research Methodologies in the Reviewed Publications

3.2.1 Research approaches

As shown in Figure 6, 11 studies of the reviewed publications adopted mixed research, accounting for a large part of the whole studies. There are 9 studies conducted quantitative research. A possible reason for the wide use of mixed research approach is that it may allow researchers to explore the comprehensive process of chatbot-assisted language learning. As shown in Table 1, among mixed research, explanatory sequential mixed methods design accounts for a large portion and among quantitative research, experimental design is the most frequently adopted design.

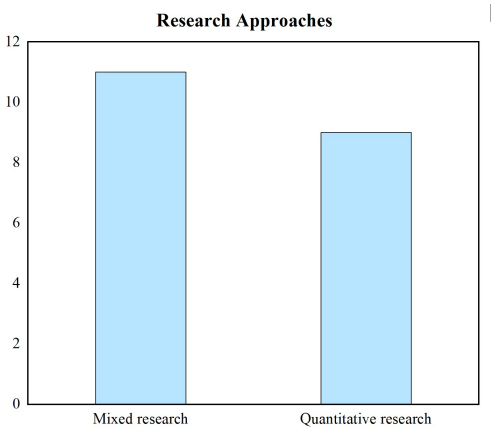


Figure 6. Research Approaches in the reviewed publications

Table 1. Research approaches and research designs of the reviewed studies (2014 – 2023).

Research approaches	Research designs	Number of studies
Mixed research	Explanatory sequential mixed methods design	8
	Convergent mixed methods design	3
Quantitative research	Experimental design	3
	Pre-Posttest Design	2
	Validation research design	1
	Between-subject design	1
	Others	2

3.2.2 Data collection methods

As exhibited in Figure 7, the reviewed studies employed a various kind of data, among which the most frequently used data collection method was survey or questionnaire (45%). Moreover, both focus group interview and testing account for 12.5. Furthermore, corpus, video/audio taped materials, comparative groups, student-generated content, log files and private document contribute to 7.5%, 5%, 2.5%, 2.5%, 2.5%, and 2.5% respectively.

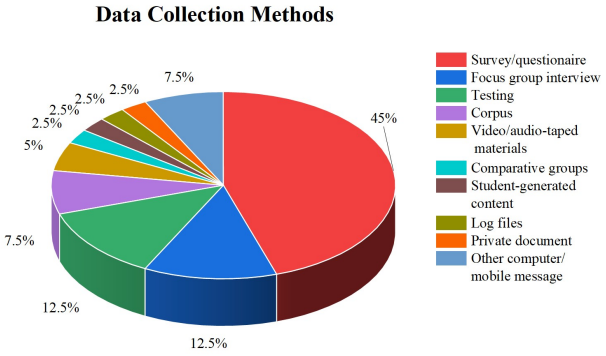


Figure 7. Data collection methods in the reviewed publications

3.2.3 Data analysis methods

As described in Figure 8 and 9, it can be found that both quantitative and qualitative data analysis methods were applied. As for quantitative data analysis methods, descriptive statistics, t-test and correlation account for a considerable portion, with the percentage of 26.4%, 22.6% and 17% respectively. Among qualitive data analysis methods, 75% reviewed studies employed corpus-based content analysis, 8.3% adopted descriptive summary, 8.3% used thematic analysis and 8.3% utilized interpretive analysis. Besides, most of the reviewed studies jointly employed quantative and qualitative data analysis methods, which corresponded to the above-mentioned conclusion that the most frequently used research approach was mixed research.

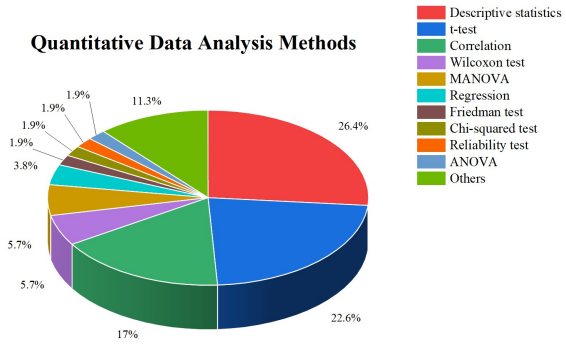


Figure 8. Quantitative data analysis methods in the reviewed publications

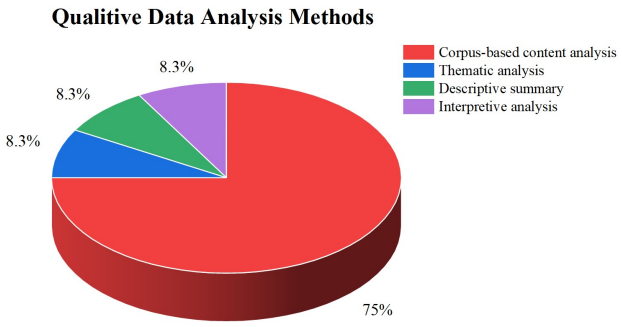


Figure 9. Qualitive data analysis methods in the reviewed publications

3.3. AI Technologies Applied in the Reviewed Publications

As described in Figure 10, machine learning, automatic speech recognition, natural language processing and computer vision are the most popular AI technologies in the reviewed studies, accounting for 19.6%, 19.6%, 17.6% and 11.8% respectively. Apart from the aforementioned technologies, deep learning, social signal processing, depth sensing, intelligent robot, multimodal learning analytics and behavior tree are also employed in the research related to public speaking instruction in the past 20 years.

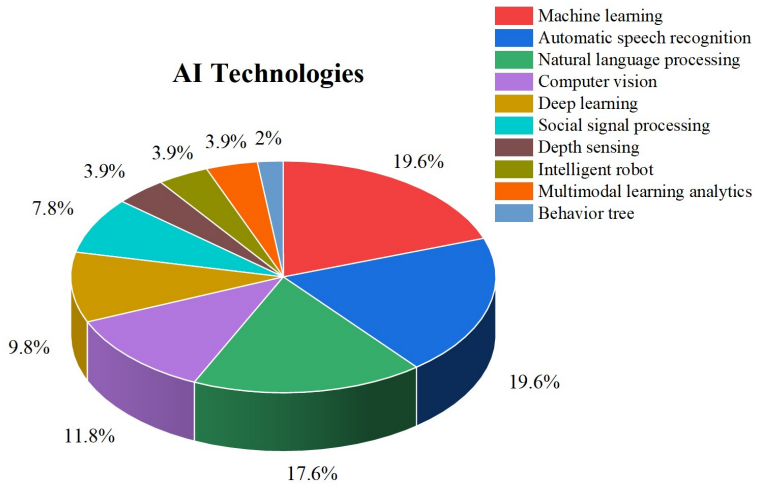


Figure 10. AI technologies applied in the reviewed publications

3.4. Research Foci in the Reviewed Studies

As illustrated in Table 2, a total of four research foci related to chatbot-assisted language learning were extracted, including learners' psychological characteristics while using AI technologies/tolls (13 studies), performance testing of AI technologies/tools (11 studies), affordance of AI technologies for public speaking ability (12 studies), affordance of AI technologies for non-linguistic knowledge and skills (1 studies).

Because most of the reviewed publication have more than one research focus, the total number is more than 20. For instance, Chen et al. (2022) focused on testing the reliability of this AI-assisted platform as well as students' views and responses on it. Chen Hsieh and Lee (2021) paid attention on how could the AI technologies influence learners; performance on public speaking and learners' emotions, grit, perceptions, motivation, effectiveness, engagement and

satisfaction during the process of utilizing AI technologies or tools. Chen (2022) aimed to assess the degree of learners' public speaking anxiety and whether their oral presentation improved or not while using AI technologies/tools.

Table 2. Research foci of the reviewed studies from 2014 to 2023.

Research foci	Number
Learners' psychological characteristics while using AI technologies/tools	13
Affordance of AI technologies for public speaking ability	12
Performance testing of AI technologies/tools	11
Affordance of AI technologies for non-linguistic knowledge and skills	1

3.5 Potential Benefits and Challenges of AI for Public Speaking Instruction

According to the reviewed publications, artificial intelligence technology has indeed greatly improved public speaking instruction in many aspects, such as increasing the learners' motivation to practice, improving the ability of learners to identify their own mistakes in real time, enhancing the learners' performance (Schneider, 2016).

However, there are still some problems and challenges. Firstly, the use of AI technologies to evaluate oral presentations has its constraints and cannot solely be trusted as an assessment tool. Communication is intricate and highly complex. To align more closely with human ratings and complement human assessors, the training platform must incorporate additional AI tools. These tools should enhance the evaluation of crucial concepts like fluency, accuracy, and facial expressions. Furthermore, exploring the possibility of broadening AI evaluation to cover other key aspects, such as eye contact and helpful signposting, is crucial. Future efforts must also refine the criteria for evaluating presentations, thus enhancing the reliability of AI assessments. The potential of combining artificial and human intelligence in tandem should be continuously explored to find solutions that enrich learning experiences and improve outcomes (Chen et al., 2022). Secondly, another significant problem that remains unexplored concerns the long-term usage of AI technologies and its impact on learning outcomes (Wyeld et al., 2021). Short-term research findings sometimes do not fully illustrate certain issues, and it is only through long-term practical application that more convincing results can be obtained. Therefore, undertaking solid and long-term empirical research is not only necessary but also urgent.

4. Conclusion

This study reviewed 20 empirical investigations on AI-supported public speaking instruction internationally from 2014 to 2023. A coding scheme was formulated from five aspects including the general research trends, research approaches, research foci, AI technologies being applied and benefits and challenges of AI-supported public speaking instruction in the past 10 years. The application of artificial intelligence in public speaking instruction context can improve learners' public speaking capacity, enhance their learning engagement and motivation and alleviate their learning anxiety, thus promoting learning effectiveness. However, there are still some challenges faced with AI-assisted public speaking instruction, such as their limited incapabilities. Therefore, it still calls on unremitting efforts to more effectively implement AI technologies in oral presentation context.

Acknowledgements

This research is funded by the National Natural Science Foundation of China (Grand No. 62107005, awarded to Dr. Chunping ZHENG).

References

Barrett, N.E., & Liu, G. (2016). Global Trends and Research Aims for English Academic Oral Presentations. *Review of Educational Research, 86*, 1227 - 1271.

Chen Hsieh, J., & Lee, J.S. (2021). Digital storytelling outcomes, emotions, grit, and perceptions among EFL middle school learners: robot-assisted versus PowerPoint-assisted presentations. *Computer Assisted Language Learning, 36*, 1088 - 1115.

Chen, J., Lai, P., Chan, A., Man, V., & Chan, C. (2022). AI-Assisted Enhancement of Student Presentation Skills: Challenges and Opportunities. *Sustainability*.

Chen, Y. (2022). Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning*.

Kim, J., Merrill, K., & Collins, C. (2021). AI as a friend or assistant: The mediating role of perceived usefulness in social AI vs. functional AI. *Telematics Informatics, 64*, 101694.

Russell, S. J., & Norvig, P. (2016). Artificial intelligence: A modern approach. *Pearson Education Limited*.

Schneider, J., Börner, D., Rosmalen, P.V., & Specht, M.M. (2016). Can You Help Me with My Pitch? Studying a Tool for Real-Time Automated Feedback. *IEEE Transactions on Learning Technologies, 9*, 318-327.

Wyeld, T.G., Jiranantanagorn, P., Shen, H., Liao, K., & Bednarz, T.P. (2021). Understanding the effects of real-time sentiment analysis and morale visualisation in backchannel systems: A case study. *Int. J. Hum. Comput. Stud., 145*, 102524.

揭薇 &金艳. (2022).《中国英语能力等级量表》口头表达典型活动和语言特征量表研究. 外语界, (06), 2-10.

基于人工智能技术的教学过程优化思考

Thoughts on Optimization of Teaching Process Based on Artificial Intelligence Technology

康健*
西北师范大学
18393418337@163.com

[摘要]文章通过可行性分析，阐述了人工智能技术与教学过程融合的可能性，指出了人工智能在提高教学效率和质量方面的积极作用。随后，提出了一个基于人工智能技术的教学过程模式，展示了如何利用这些技术来设计和实施更加智能化、高效的教学。文章最后部分深入讨论了使用人工智能技术进行教学改革时所引发的三个主要思考：伦理、技术软硬件完善的必要性以及教师的职前职后培养与培训。在伦理部分，主要讨论了数据安全、技术依赖、伦理决策的问题。在技术完善层面，指出了持续改进人工智能系统的软硬件配置，以满足教学需求的必要性。最后，在教师培养与培训方面，文章强调了需要为师范生和在职教师提供关于人工智能技术的知识和技能培训，以确保他们能有效利用这些技术来优化教学过程。

【关键词】人工智能技术 教学过程 教学模式

1. 问题提出

随着人工智能（AI）技术的飞速发展，时代对教育教学的质量和方式提出了新的要求和挑战。AI 的出现为教育领域带来了前所未有的机遇，使得个性化学习、资源优化分配以及学习效率的提高成为可能，时代的发展倒逼着教学过程发生改变（焦建利，2023）。2024 年 3 月，时任教育部部长怀进鹏表示，将培养一大批具有数字素养的教师，把人工智能技术深入到教育教学和管理全过程、全环节，让学生更加主动地学，让教师更加创造性地教（教育部，2024）。在有关时代和政策的背景下，专家学者围绕着人工智能赋能教育做了大量的研究，刘革平指出在以人工智能为核心的信息技术驱动下，教育领域正经历着深刻的变革。这些变化主要体现在六个方面：教育资源数字化、教育平台智慧化、教学模式混合化、学习形态个性化、教育评价精准化以及教育治理科学化（刘革平和胡翰林，2024）。吴砥等人的研究指出智能技术自身的发展进步及其在评价中日益广泛深入的应用，将助推结果评价更加科学、过程评价更加智慧、增值评价更加便捷、综合评价更加完善，并给出智能技术赋能教育评价改革的实践路径建议（吴砥等人，2023）。张秀梅等人指出未来教学模式趋势表现为三个方面：面向核心素养，借由深度学习，教学模式走向人工智能驱动的智能化路径。同时指出教学模式更加强调基于技术的教学目标、环境、过程、主体、评价的一体化融（张秀梅等人，2023）。人工智能技术的持续发展已然成为推动教育变革的重要力量，人工智能赋能教育作为一种必然趋势正塑造着教育发展的新范式和新形态（郑庆华，2024）。

但与此同时，我们需要保持清醒的头脑，重新审视技术与教育教学过程的融合，思考深度分析人工智能技术与教学过程融合的的内在逻辑问题。基于此，本研究提出以下问题：人工智能技术优化教学过程可行性？如果可行，人工智能技术与教学过程融合的模式是怎样的？最后针对构建的模式，引发伦理，技术发展，教师培养与培训三个方面的思考。

2. 人工智能技术优化教学过程的可行性

2.1. 人工智能技术与学习内在逻辑的一致性

在教育领域，传统信息技术的引入旨在促进教学效率和学习体验，但实际效果却并不理想。一方面，技术发展本身存在瓶颈，语音识别、自动翻译、智能阅卷、自然语言理解等领域亟待有质的突破（刘燕和胡凡刚，2022），另一方面以技术植入为导向的技术赋能教学在教育实践过程中并没有从根本上提升教学质量、改进教学效果。问题的根源在于技术应用与学习内在逻辑之间的矛盾（张晓荣，2009）。人类的学习过程是基于对日常感知的外在信息积极处理和自主建构，指在个体现有的认知框架基础上获取生活经验，构建对客观世界的主观理解。学习不仅仅是知识和技能的获取，也不只是行为的改变，而是应该包括在知识经验的基础上培养良好的行为习惯，形成高尚的道德品质，并充分发挥个体的潜能和价值。

传统技术在教学的过程中的使用往往仅作为内容展示方式的改变或进行单一的数据分析，难以真正融入教学和学习过程中，促进教与学的效果。相比传统技术，人工智能（AI）技术的本质在于模仿、扩展和增强人类智能，通过计算机系统执行原本需要人类智能才能完成的任务，包括学习、理解、推理、规划和语言交流等。AI 技术旨在创造能够自主学习、适应并执行复杂任务的系统，从而扩展人类的认知和操作能力（邓晓芒，2022）。因此，AI 技术与人类学习的内在逻辑之间存在更高的一致性，与传统信息技术相比，人工智能技术在促进教学过程中展现出更加符合人类学习需求的特性，理论上更能与学习的内在逻辑相协调，从而有望在优化教学过程中发挥更好的作用。

2.2. 人工智能技术促进教学过程的功能梳理

人工智能（AI）技术应用于教学过程是可行的，而且在近年来已经展示了显著的潜力和成效。本文通过文献梳理发现，人工智能可以通过多种方式赋能教学，包括学情分析（冯敬杰和冯敬舜，2024）、智能辅导系统（管佳等人，2022）、课堂教学评价（谷雨和郝春东，2024）等。以下是几个关键领域，展示了 AI 在教育中的应用和潜力，以智能辅导系统为例，生成式人工智能可以与学生以对话的形式完成学习辅导，开展自学。

表 1 人工智能技术促进教学过程的功能梳理

AI 技术功能	描述
学情分析	学习需求，学习特征分析，自动批改作业和考试，分析学习趋势
智能辅导系统	提供实时反馈和解答，模拟一对一辅导的体验
课堂教学评价	评估学生的语言、情感、参与度和情绪状态，为教师提供教学策略反馈

3. 基于人工智能技术的教学过程模式

3.1 模式构建原则

3.1.1 技术使用的必要性、可行性、适切性原则

基于人工智能技术开展教学离不开技术的使用，但是不是所有的教学过程都

需要应用技术，在使用技术过程中需要遵循技术使用的必要性，可行性、适切性原则。必要性是指并非所有的教学过程都要使用信息技术，教材本身蕴含的知识，方法才是教育教学的核心，技术只是辅助教师教学、学生学习的工具，因此在教学活动的设计过程中，不可本末倒置过于追求技术的使用。所谓可行性就是指教学环境中是否存在使用相应技术的条件，例如，不能在只具备电子白板教室设计虚拟交互体验类教学活动。适切性是指选择合适的技术使用，不是所有的信息技术都可以设计运用到实际教学过程中，例如数学运算过程，在学生初步接触相应运算时不能直接利用技术给出相应结果，这样不利于学生对相应知识，方法过程的掌握。

3.1.2 学生主体原则

学生主体原则是指在课堂教学过程中以学生为中心，教师引导展开教学的原则。传统课堂教学中一般以教师讲授为主，学生被动的接受，很难调动学生主动学习的积极性。而教育真正需要培养的是学生学习和发展的能力，这就要求课堂教学需要发生改变，从以教师讲授为主转成以学生自学为中心的教学，让学生通过教学活动主动探索、自主建构的学习方式来学习知识和解决问题，真正让学生成为教学活动的主体，在这个过程中，教师需要引导和督促学生开展教学活动，突出学生在课堂学习中的主体性与主观能动性。教师在设计教学活动的时候需要充分考虑到资源对学生学习的支撑以及学生对于教学活动的可参与性，便于学生在教学活动中主动完成意义构建。

3.1.3 数据驱动决策原则

基于人工智能的教学过程模式的构建原则之一，即数据驱动决策，强调了利用大数据和机器学习算法来指导教学策略和个性化学习路径的重要性。在这一原则下，学习者的每一次互动、回答、甚至是停顿时间都被视为宝贵的数据源，用于深入理解学习者的需求、偏好和挑战。通过对这些数据的收集、分析和应用，教育技术可以不断调整和优化其教学内容、难度和方法，确保每位学习者都能在最适合自己的节奏和方式下进行学习。

此外，数据驱动的决策不仅仅局限于改善个体学习体验，还包括对整个教学系统的改进。这意味着，通过分析来自广泛用户群体的数据，可以识别出教学过程中的普遍性问题和机会，进而引导课程设计和教学方法的系统性调整。例如，如果数据显示大多数学习者在某个特定概念上遇到困难，系统可以自动引入额外的教学材料和练习，或者提示教师对该部分内容进行重点解释。

数据驱动决策的实施要求高度的技术专业性，包括数据采集的准确性、分析方法的先进性以及对结果的合理解释。同时，这种方法还需要一个强大的伦理框架来确保数据的安全、隐私和公平使用，防止数据滥用和偏见的产生。最终，通过将这一原则融入人工智能教学模式的构建中，可以极大地提升教育系统的效率和效果，实现真正意义上的个性化学习，帮助每位学习者实现其最大潜能。

3.2 教学模式构建

在分析了人工智能技术促进教学过程的可行性，以及梳理了可用于促进教学过程的人工智能技术之后，考虑教学过程构建的基本原则，本研究基于建构主义理论，构建了基于人工智能的教学过程模式，如图 1 所示：

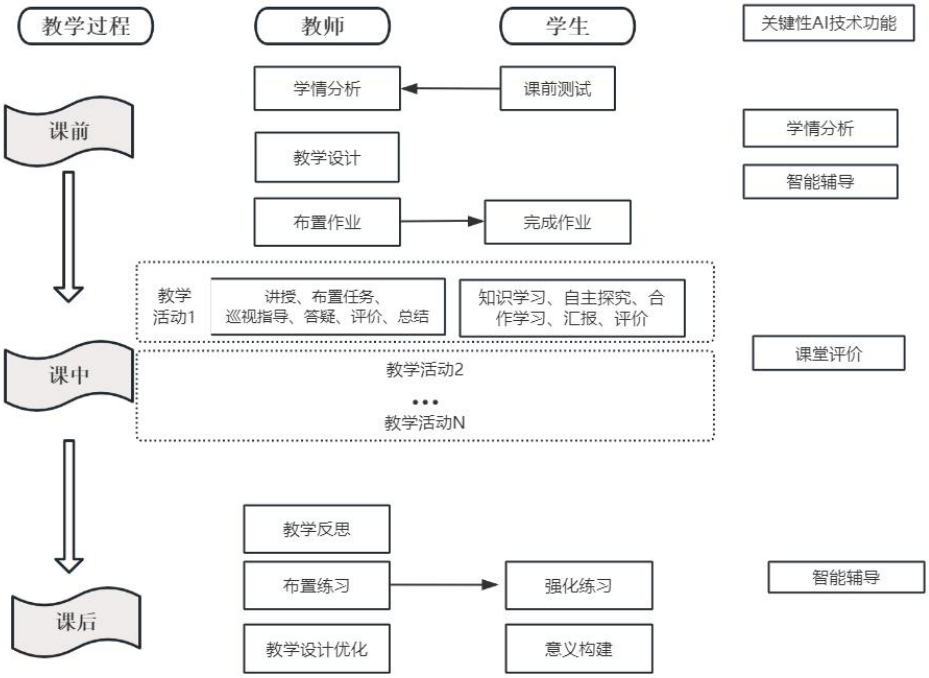


图 1 基于人工智能技术的教学过程

如图所示：整个教学过程模式分为三个阶段，在课前，教师利用平台（含学情分析功能）技术了解学生学情，基于学生学情在课程设计技术的支持下开展教学设计，布置作业。学生在完成课前测试的基础上利用智能辅导功能完成教师布置的作业，并依据教师布置的作业开展个性化学习。

课中开展教学活动，教师基于课程设计技术设计的教学方案开展教学，在教学活动中完成知识的讲授、布置任务、巡视指导、答疑、评价、总结。学生在这个过程中完成知识学习、自主探究、合作学习、汇报、相互评价。这一过程中主要的技术为课堂评价技术，利用人工智能技术对学生的语言、情感、参与度和情绪状态等进行评估，为教师提供教学策略反馈，以便教师更好的把握教学进度与教学质量，同时这一过程遵循以学生为主体的教学范式，教师在人工智能技术的支撑下帮助学生完成知识构建。

课后教师基于数据进行教学反思，并布置相应课后练习，优化下一节课的教学设计。学生进行针对性的强化练习，完成知识构建，这一过程中的主要技术有智能辅导、自动化评估、个性化学习、

3.3. 教学模式实施建议

3.3.1 知识分层教学

在人工智能技术赋能的教学过程中，知识分层教学方法根据布鲁姆的知识分类理论，将学习内容分为不同的层级，从而提高教学效率并针对性地解决学习中的难点。具体来说，这一方法建议将事实性知识的学习交由学生在家中通过智能辅导技术自主完成。这种方式不仅能够充分利用学生的自学时间，通过个性化的智能学习平台提供的资源和辅导，还能确保学生在基础知识上的坚实掌握。

接着，教师在课堂上主要集中精力于概念性知识和程序性知识的讲解和讨论。概念性知识通常包括了对事物的分类、术语的定义等，它们是理解更高层次知识的基础；而程序性知识则关注于“如何做”，即方法、技能和策略的使用。通过在课堂上重点讲解这两类知识，教师能够引导学生构建深层次的意义，激发学生

的思考和探究，从而实现以学生为中心的教学目标。

此外，通过在课堂上实行评测，教师可以及时了解学生对事实性知识的掌握程度，及时调整教学策略，确保学生能够在理解的基础上进一步构建知识体系。这样的教学模式不仅提高了课堂教学的效率，而且通过智能技术和分层教学的结合，更能精准地解决学习过程中的重难点，促进学生全面而深入的学习。

3.3.2 有效利用智能技术开展教学

在人工智能技术赋能的教学过程中，有效利用智能技术不仅可以增强教学的针对性和有效性，还能实现对学生学情的精准把握、创新的教学设计以及高效的课堂评价。首先，通过智能技术，教师能够实时收集和分析学生的学习数据，包括学习进度、理解程度和学习习惯等，这有助于教师全面了解学生的学情，从而提供更加个性化的教学支持。例如，利用数据分析技术，教师可以识别学生的学习难点和兴趣点，进而调整教学内容和方法，以满足不同学生的学习需求。

其次，在教学设计方面，智能技术的应用使得教师能够根据学生的实际学情设计更具针对性的教学活动和资源。通过智能教学平台，教师可以轻松创建和管理富媒体教学资源，如交互式视频、模拟实验和在线测验等，这些资源可以极大地提高学生的学习兴趣和参与度。

最后，课堂智能评价是智能技术赋能教学过程中的另一重要应用。通过智能评价系统，教师不仅可以实时跟踪学生的学习表现，还可以根据系统提供的反馈及时调整教学策略。此外，智能评价系统还能为学生提供及时、个性化的反馈，帮助学生更好地理解自己的学习状态，从而促进自我调节学习。

综上所述，有效利用智能技术进行教学，能够让教师更好地分析学生学情、开展创新的教学设计并进行有效的课堂评价，这不仅可以提升教学质量，还能激发学生的学习兴趣，最终实现教学的高效促进。

3.3.3 以人为本开展教学

在人工智能技术赋能的教学过程中，以人为本的教学方法强调将学生的需求、能力和兴趣放在教学活动的中心位置，确保技术的应用增强而非替代教师与学生之间的互动。尽管人工智能可以提供个性化学习路径、智能评估和丰富的教学资源，但教育的核心仍然是人的发展和成长。因此，教师需要利用人工智能技术作为辅助工具，来更好地理解每位学生的独特需求，支持他们的个性化学习，并促进他们的全面发展。

以人为本的教学方法鼓励教师使用人工智能技术来进行学情分析，从而精准识别学生的学习风格、能力水平和知识空缺，以便设计更加符合学生需求的教学内容和方法。此外，通过智能系统提供的实时反馈和评估，教师可以及时调整教学策略，以更有效地支持学生的学习进程。

同时，以人为本的教学还意味着在利用人工智能技术时，应充分考虑学生的情感和社交需求。教育不仅仅是知识的传授，更是情感联系和社会技能发展的过程。因此，教师应当利用人工智能技术来创造更多的互动和合作学习机会，鼓励学生之间的讨论和交流，以促进学生的情感发展和社交能力的提高。

总之，以人为本的教学方法强调在人工智能技术支持下的教学应当聚焦于学生的全面发展，通过个性化学习的支持、情感和社交技能的培养，确保每位学生能在技术赋能的教学环境中实现其最大潜力。

4. 关于模式的思考

4.1 人工智能技术支持的教学软硬件开发

随着人工智能技术在教育领域的应用日益广泛，其在优化教学过程中的作用变得尤为显著。这种技术不仅能够提供个性化学习路径，还能通过智能分析帮助教师更好地理解学生的需求。然而，现有的技术还无法满足实际教学需求，要充分发挥人工智能在教育中的潜力，就需要开发适合教学需求的软硬件系统。

软件系统如智能辅导系统（软件）：这类系统能根据学生的学习进度和能力提供定制化的学习资源和练习，如Coursera和中国大学慕课等在线学习平台，它们利用算法来适配学生的学习路径，确保教学内容与学生的实际水平相匹配。智能问答机器人（软件）：例如，基于自然语言处理技术的Duolingo，能够实现与学生的互动式学习,通过模拟对话练习语言学习,提高学生的语言实践能力。

硬件系统如可穿戴设备：如智能手表和健康跟踪器，能够监测学生在学习过程中的生理反应（例如，心率和注意力水平），为教师提供关于学生学习状态的反馈，从而调整教学策略。虚拟现实（VR）和增强现实（AR）设备（硬件）：通过提供沉浸式学习体验，如Google Cardboard或Oculus Rift，可以让学生在虚拟环境中探索复杂的科学概念或历史事件，提高学习的互动性和趣味性。

4.2技术使用带来的伦理问题思考

将人工智能（AI）技术应用到教学过程中，确实会带来了一系列的伦理问题。这些问题触及学生的数据安全、技术依赖、伦理决策等多个方面。思考技术使用所引发的伦理问题，对于确保教育公平、透明和责任至关重要，

4.2.1数据安全问题

与所有技术应用相似，人工智能技术的使用同样面临着数据安全的风险。利用人工智能技术来收集和分析学生数据时,确保学生的个人信息安全是一个重要的问题。必须合理地收集数据，即只收集实现教育目的所必需的信息，避免过度收集可能侵犯隐私的数据。其次，数据收集的目的应该明确且透明，确保学生和家长理解数据如何被使用,并获得他们此外,数据的安全存储和处理也至关重要,其中包括使用高标准的数据加密技术来防止未经授权的访问和数据泄露。学校和教育机构应实施严格的数据安全政策，定期进行安全审查和更新，以应对不断变化的安全威胁。最后，还应建立适当的数据使用和保护机制，包括在数据处理结束后及时删除不再需要的信息，以保护学生的隐私权。这些措施共同构成了全面的数据保护策略，旨在利用人工智能技术的便利，保护学生的基本权利。。

4.2.2技术依赖问题

人工智能技术在教学过程中的应用正在不断扩展,带来了便利和创新。然而,这种技术的广泛使用也引发了技术依赖的问题。过度依赖技术可能会极大教育的核心实践，包括传统的教学方法与师生之间的直接互动。此外，这种依赖可能对学生综合能力的发展产生不利影响，尤其是批判性思维、人际交往以及解决复杂问题等能力对于学生的全面发展至关重要,需要在现实交流和波动的思辨中得到锻炼和提高。因此，教育者在整合人工智能技术时，应重点考虑如何保持技术工具的辅助性，避免其取代传统教学的核心要素。同时，应加强对学生关键技能的培养，确保他们在技术便利的同时，也能在人际互动和复杂思维挑战中成长。这种平衡对于培养未来社会所需的全面发展的人才是必需的。

4.2.3伦理决策问题

将人工智能技术应用到教学中时，教育者面临决策伦理问题。教育者应遵循以学生利益为先的伦理指导原则，确保技术应用符合教育目的。教育机构需建立伦理框架，考虑教育公平性，增强人工智能系统的透明度和可解释性。此外，教育者需接受专业培训，理解人工智能技术及伦理标准，机构应提供必要的资源和

工具。建议设立伦理委员会，定期审查人工智能使用情况，确保符合教育伦理标准，有效支撑教育目标。这些措施有助于维护教育伦理的高标准，同时提升教学质量。

4.3 教师培养和培训的思考

为适应人工智能时代教育教学的变化，教师作为教学的重要参与人员，需要做出改变，以适应时代需求，教师的改变可以从职前-师范生阶段，职后-教师两个阶段分开阐述

4.3.1 师范生培养

师范生培养主要从个性化数字学习空间构建、智能课程体系改革、智能测评、智能教育实训场景打造四方面入手。在信息技术迅猛发展的当下，构建师范生个性化数字化学习空间已成为教育转型升级的关键。这一空间集成了高速网络、云服务、智慧教室等先进技术设施和丰富的数字资源，如电子图书馆、在线课程、模拟教室等，为师范生提供了一个以学生为本、交互性强、资源丰富的网络化学习环境。这样的环境不仅支持师范生通过多样化的学习方式提升学习和教学能力，还通过人工智能管理平台实现对培养过程的精准监控和高效协调，使教育路径更加透明化和个性化。

为了适应职前教师教育的发展，智能课程体系的构建至关重要。通过更新和优化课程资源，引入最新的教育理念和人工智能技术，师范生课程体系得以丰富和多元化，帮助师范生掌握人工智能的相关知识及应用方式。同时，通过融合线上选修和线下必修的模式，结合传统教室学习和在线平台资源，师范生能够享受到更加灵活、高效和个性化的学习体验。

在专业发展方面，建立智能测评工具对师范生的能力培养至关重要。基于教育课程知识图谱与能力要求，智能数据采集设备能够对师范生的培养全过程进行跟踪和分析，形成详细的数据信息链，从而为师范生提供精准的个性化反馈和指导。

最后，随着人工智能技术的发展，为师范生打造多样化的智能教育实训场景变得可行。这不仅能够提供更丰富的实践机会，还能通过智能训练和诊断平台帮助师范生精准掌握自身的学习情况，有效提升教学法指导服务的质量和效能。通过打造个性化的数字化学习空间、智能化的课程体系、科学合理的评价制度以及多样化的实训场景，可以极大地提升师范生的学习效率和教学能力，为未来的人工智能教育职业生涯奠定坚实的基础。

4.3.2 教师培训

在人工智能技术广泛应用到教学过程之后，在职教师的培训是必要的，只有教师掌握人工智能技术的使用，学会处理人工智能技术使用带来的伦理问题，人工智能技术才能真正的为教学过程服务。

对于人工智能技术和教学方法创新的相关培训，培训应关注于如何将 AI 技术融入现有的教学方法中，包括创建更加个性化和参与性的学习经验；培训不应只注重技术培训，而应该让教师学会使用技术开展教学，如利用人工智能生成教学设计，利用人工智能进行教学评价等等；同时，也应关注如何利用人工智能帮助教师增效减负，减少教师在非授课事件上的精力消耗，提升教师的工作效率进而提升教学效果。

加强伦理问题解决的相关培训；人工智能与教学的融合不可避免，这个过程必定会带来一些伦理问题，而教师作为一线教育的掌舵人，必须学会处理这些问题，才能更好的保障人工智能技术与教学融合的实际效果，因此需要依据或者制

定相关标准，对一线教师展开培训，帮助教师解决在实际教学过程中遇到的伦理问题。

参考文献：

焦建利(2023)。ChatGPT 助推学校教育数字化转型——人工智能时代学什么与怎么教。中国远程教育,(04),16-23。https://doi.org/10.13541/j.cnki.chinade.20230226.001

教育部十四届全国人大二次会议举行民生主题记者会.[EB/OL]。http://www.moe.gov.cn/jyb_xwfb/s5147/202403/t20240310_1119456.html

刘革平和胡翰林(2024)。技术赋能教育强国建设。 西南大学学报(社会科学版),(02),168-180。https://doi.org/10.13718/j.cnki.xdsk.2024.02.014

吴砥、郭庆、吴龙凯和程浩 (2023)。智能技术赋能教育评价改革。开放教育研究,(04),4-10。https://doi.org/10.13966/j.cnki.kfjyyj.2023.04.001

张秀梅、赵明仁和陆春萍(2023).技术赋能的中小学教学模式创生路径研究——政策、理论、成果、特点与趋势。中国电化教育,(08), 32-40。

郑庆华(2024)。人工智能赋能创建未来教育新格局。中国高教研究,(03),1-7。https://doi.org/10.16298/j.cnki.1004-3667.2024.03.01

刘燕和胡凡刚(2022)。智能技术之于教师的隐忧省视与消解路径。现代教育技术,(02),37-43。

张晓荣(2009)。试论人类学习的本质与机器学习。学术论坛,(01),53-55。https://doi.org/10.16524/j.45-1002.2009.01.029

邓晓芒(2022)。人工智能的本质。山东社会科学,(12),39-46。https://doi.org/10.14112/j.cnki.37-1053/c.2022.12.002

冯敬杰和冯敬舜(2024)。智慧课堂教学模式的研究与构建。中国现代教育装备,2024(01), 37-39+42。https://doi.org/10.13492/j.cnki.cmee.2024.01.023

管佳、韩婷芷和徐国兴(2022)。人工智能技术赋能我国高等教育拔尖人才培养。中国电化教育, (10), 97-101。

谷雨和郝春东(2024)。智慧课堂助力语文课堂教学评价的路径探究。黑龙江教育(教育与教学), (S1), 61-63。

人工智能赋能书院制高校学生管理：内涵、应用及创新路径

Artificial Intelligence Empower Residential College Student Management: Connotation, application and path[1]

于越[1]，王令军[2]，陈劲松[3*]
[1]西北师范大学教育技术学院
[2]清华大学教育研究院
[3]南方科技大学人文科学中心
* chenjs@sustech.edu.cn

【摘要】 拔尖创新人才是国家实现高水平科技自立自强的攻坚力量，书院制培养是我国进行拔尖创新人才培养的重要探索和举措之一。目前全国已有 200 多所高校正在探索书院制人才培养模式，但书院制学生管理中始终存在两条线管理职责不清、书院导师育人效果不佳、辅导员对学生关注不足三重困境，人工智能赋能为优化和提升书院制高校学生管理提供了新的机遇。通过剖析人工智能赋能书院制高校学生管理的内涵和意义，重点阐述人工智能赋能书院制高校学生管理的应用，包括在智慧校园和智慧学工系统、人机融合治理模式、书院社区管理、心理健康监测、学生评价、学生就业、舆情风险治理等方面；并从“以管理为中心”向“以师生为中心”的转变、提升师生数字素养与数字胜任力、建设人工智能治理共同体、促进智能产品与治理场景适配、建立风险防范与监管机制五个方面出发，探索了人工智能赋能书院制高校学生管理的创新路径。

【关键词】 人工智能；书院制高校；学生管理；创新路径

Abstract: *Top innovative students are the key force for the country to achieve high level of scientific and technological self-reliance, and the residential college mode is one of the important explorations and measures to cultivate top innovative students in our country. At present, more than 200 colleges and universities nationwide are exploring the residential college mode, but there are always three difficulties in the college system student management, such as two-line management, poor education effect of college tutors, and insufficient attention to students by counselors. Artificial intelligence can provide new opportunities to optimize and improve the residential college student management. This paper analyzes the connotation and significance of student management enabled by artificial intelligence in residential college. Then, the article focuses on the application of artificial intelligence-enabled residential college student management, including smart campus system and smart student-work system, man-machine integration governance model, college community management, mental health monitoring, student evaluation, student employment, public opinion risk management. Finally, the paper explores the path of student management in residential colleges and universities enabled by artificial intelligence from five aspects: the transformation from "management-centered" to "teacher-student centered", the improvement of digital literacy and digital competency of teachers and students, the construction of artificial intelligence governance community, the promotion of the adaptation of intelligent products to governance scenarios, and the establishment of risk prevention and supervision mechanisms.*

Keywords: artificial intelligence, residential college, college student management, innovation path

1. 引言

1 本文系深圳市教育学会 2023 年度教育科研课题《三全育人背景下书院制高校学困生帮扶对策研究》（YB2023008）阶段成果。

党的二十大报告提出，要加快建设教育强国、科技强国、人才强国，坚持为党育人、为国育才，全面提高人才自主培养质量，着力造就拔尖创新人才，聚天下英才而用之（习近平，2022）。拔尖创新人才是国家实现高水平科技自立自强的攻坚力量，书院制是一种跟素质教育、通识教育相关的人才培养模式，是我国进行拔尖创新人才培养的重要探索和有力举措之一。但是目前书院制存在两条线管理职责不清、书院导师育人效果不佳、辅导员对学生关注不足的问题。习近平总书记指出，教育数字化是开辟教育发展新赛道和塑造教育发展新优势的重要突破口（2023）。人工智能具有高效性、可靠性、个性化服务、自主学习的特点，为书院制高校解决学生管理中存在的问题提供了新的思路。明晰人工智能赋能书院制高校学生管理的应用并寻求创新路径，不仅是政策所指，也是现实所需。

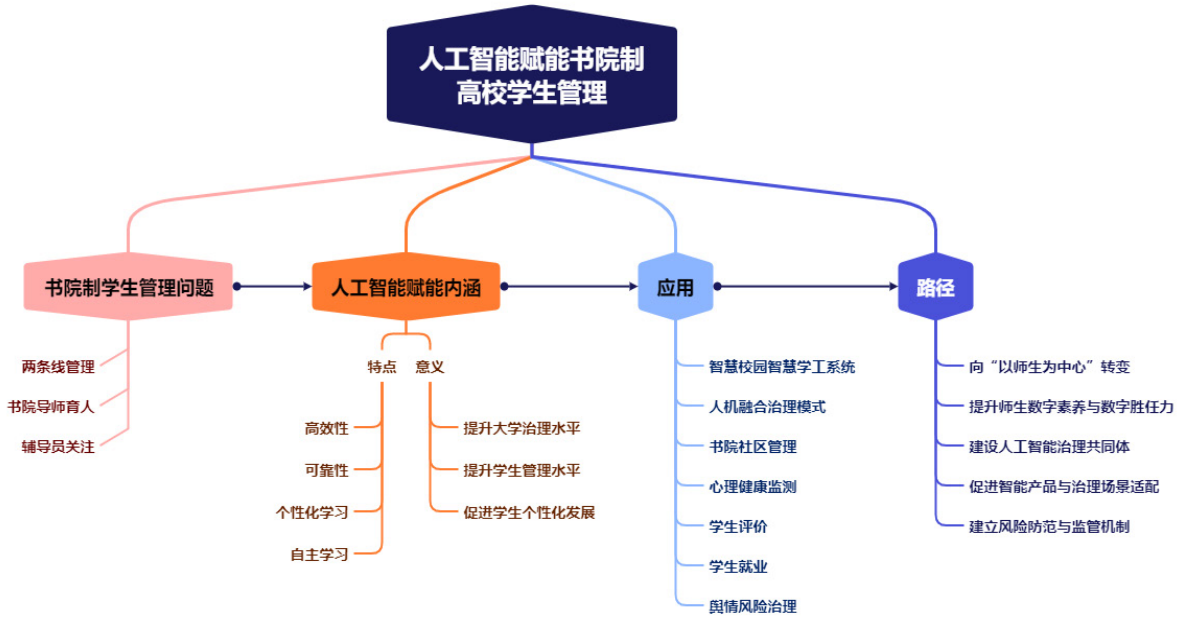


图 1 *人工智能赋能书院制高校学生管理框架图*

2. 书院制高校学生管理困难及人工智能赋能的内涵

2.1. 国内书院制高校发展历程及分类

在全球化时代，知识经济快速发展，拔尖创新人才的培养尤为关键，成为衡量一个国家综合竞争力的核心指标。为了提升教育质量、推动学生全面发展，国内一些大学以学生住宿社区为基础，借鉴国外大学的住宿学院经验，结合中国古代书院的教育观念，实行了书院制。2005 年复旦大学、西安交通大学为适应大类招生的需要，相继实行书院制管理。目前全国已有 200 多所高校探索书院制人才培养模式（李凤亮，2023）。李禾等指出，国内高校书院制改革的路径大致可以分为四类。一是囊括全员的大型住宿式书院，比如西安交通大学把本科生整体纳入八大书院，书院跟专业学院是并行的。二是精英书院，主要进行拔尖人才培养，比如北京大学元培学院、清华大学博雅书院、中国海洋大学行远书院。这种书院规模小，学生要经历竞争性的筛选才能进入，书院有相对独立的培养方案和专业建制。三是特色型书院，主要是针对某个群体进行培育，或因高校要推行某种独特的教育理念而成立的书院。比如潍坊医学院乐道书院和济世书院。四是阶段性的教育管理型书院，也就是新生书院，把本科教育的前期阶段拿出来，集中管理，进行通识教育，高年级在管理上回归专业院系（2023）。比如南方科技大学致仁、致诚、致新、树仁、树德、树礼六大书院。书院制基于整合学习与

生活的理念，通过发挥社区育人功能，加强课外通识教育，推行导师制、混合住宿等，构建课内外经验整合的教育模式，形成对专业、课堂教学的重要补充。

2.2. 书院制高校学生管理面临的困境

高校学生管理是高校工作的重要内容之一，是维护学校正常教学科研秩序、保证学生健康成长的基础性工作。但目前书院学生培养目标的高标准与实际学生管理的低效率，是当前书院制高校学生管理的一对主要矛盾。具体表现为三点：

2.2.1. "两条线"管理职责不清

高校内部教学管理条线与学生工作管理条线彼此分开，形成"两条线"管理体制。产生这一问题的根本原因，在于高校没有改变传统教学口和学工口严重分离的育人与管理模式，并对两者之间的关联与分工存在较大理解偏差。当两条线严重分割时，两者之间基本上是你做你的、我做我的，各自争着要做出自己的工作业绩（叶志明，2022）。

2.2.2. 书院导师育人效果不佳

书院导师面临的教学、科研压力较大，虽然学校把书院导师在书院中培养学生的情况纳入了考核指标，也对优秀书院导师进行了一定的奖励，但对书院导师的评价仍主要取决于导师在教学、科研上的成绩，故导师投入在书院中培养学生的精力有限，与学生的互动和指导较少。

2.2.3. 辅导员对学生关注不足

辅导员在愈发复杂的大学生管理工作中经常超负荷运转，存在"心有余而力不足"的窘境。尤其在书院社区管理、书院特色活动等工作中，过多的琐碎性事务让辅导员总是处于被动的接受状态，而不是积极的应对状态。尤其当众多任务"迫在眉睫"时，辅导员易产生紧张急躁的心理，本应该慢慢来的育人过程变得"急功近利"，导致育人效果达不到预期目标（祁叶达等，2024）。

尽管书院制高校投入了巨大的人力、物力、财力到学生管理中，但上述三个问题依旧存在。书院制高校亟须通过更智能、更高效、更精细的方式，依托人工智能、大数据相关的理念和技术来解决学生管理中的问题。

3. 人工智能赋能书院制高校学生管理的内涵与意义

人工智能是一个结合计算机科学和强大数据集来解决问题的领域，包含机器学习和深度学习的子领域，目前应用包括语音识别、客户服务、计算机视觉、推荐引擎等（IBM，2023）。人工智能具有高效性、可靠性、个性化服务、自主学习的特点。高效性，即人工智能可以在短时间内处理大量的数据和任务，提高学生管理的效率；可靠性，即相较于人类，人工智能可以更快速、更准确地执行任务，不受疲劳、情绪等因素的影响；个性化服务，即人工智能可以通过大量的学生数据，提供个性化服务和推荐，提高师生体验和满意度；人工智能的自主学习，即人工智能可以通过机器学习和深度学习等技术，自主地学习和优化学生管理应用模型，提高模型的精度和效果。人工智能赋能书院制高校学生管理，就是利用人工智能技术的高效性、科学性、个性化服务、自主学习等优势，做好数据的挖掘、分享和应用，打通学生管理各相关部门，减轻书院导师、辅导员日常工作压力，提高学生管理水平，影响学生管理与服务理念，最终促进书院制高校治理体系与治理能力的现代化发展。从高校、学生管理者和学生自身出发，人工智能赋能书院制高校学生管理有三方面意义。

3.1. 提升大学治理水平

从高校角度出发，人工智能赋能高校学生管理，可以提升大学治理水平。借助人工智能技术优势，构建多元参与的教育治理新机制，形成多元主体、大规模意见交互，提升科学决策、民主决策水平。强调用数据说话、事实说话，摒弃主观主义、经验主义，更好地把握教

育变革与发展的客观规律。通过跨领域、全样本的数据采集与分析，利用数据的观点辅助教育决策，着力推进智能算法与数据驱动的教育治理模式，以保证教育决策和学生管理决策的客观性和科学性（管培俊，2021）。

3.2. 提升学生管理水平

从学生管理者出发，人工智能赋能高校学生管理，可以提升学生管理水平。在学生管理过程中，注重充分发挥数据的价值，推动教育服务创新。从书院导师、辅导员、学生等不同服务对象的实际需求出发，利用数据共享，推动信息系统间的协同联动，创新优化教育服务。还可通过全量数据的共享交换，精简办事流程、缩短办理周期，做到“让数据多跑路、师生少跑路”，减轻导师、辅导员的工作负担，切实提升师生的获得感、幸福感和满足感。

3.3. 促进学生个性化发展

从学生自身角度出发，人工智能赋能高校学生管理，可以促进学生个性化发展。探索基于大数据的用户行为分析，为广大学生提供个性化的教育服务，促进学生的个性化成长，提升教师的教育教学水平。

4. 人工智能赋能书院制高校学生管理的应用

目前，国内很多高校已经在人工智能赋能大学治理与学生管理方面进行了积极探索，书院制高校可以吸收借鉴先进经验，结合自身高校实际，形成特色的“人工智能+书院”管理模式和治理文化。

4.1. 智慧校园和智慧学工系统

在高校治理的过程中，治理主体的决策水平直接关系到最终的治理效果，而决策水平的高低则取决于决策时收集到的信息数量、质量，以及对信息处理的程度（申国昌等，2023）。人工智能所具备的各种算法系统为书院制高校学生管理治理决策的科学化提供了有力的技术支撑。在学校层面，建设智慧校园，技术快速且高效地对海量数据进行深度挖掘与动态分析，保证决策信息的丰富性。例如，“智慧北航”搭建了一套由存储基础设施、数据采集系统、数据交换平台、智能分析系统和智能决策系统等组成的虚拟化数据中心架构，对有关教学、科研、实验、档案、图书、管理、服务、人才、国际合作等方面的多源数据进行深度挖掘和处理，将数据作为高校治理的重要支撑保障（徐青山等，2016）。在学生工作层面，搭建智慧学工系统，打通智慧校园系统，做到数据共享共通。比如，中国民航大学智慧学工系统涵盖了公寓管理、综合测评、奖学金、荣誉称号、违纪处分、学困生管理、助学金管理、勤工助学、困难补助、学费减免、助学贷款、医保管理、晚查寝、请销假、离返校管理、心理咨询、就业管理等各类模块，实现对学生管理业务的全覆盖（王艺超等，2023）。

4.2. 人机融合治理模式

人机融合的治理模式将人与人工智能的优势充分结合，是高校治理智能化发展的重要表征。人工智能在信息录入、信息分析、规律总结、结果展示等方面能够模拟人的行为，形成智慧化的治理模式，从而拓宽高校治理的数字化边界，实现 24 小时全天候的在线服务，为满足师生个性化的发展需求提供有力的技术支撑。例如，中央财经大学在高校治理中通过多种智能终端设备进行人与物的信息交换，实现教学活动检测、设备运行监控、无感知考勤、校园安全监控、声光电等生活环境监测，人工智能技术为多重业务场景提供了应用基础（王芳，2022）。

4.3. 书院社区管理

书院社区是学生居住、生活和思想交流的重要场所，是开展学生思想政治教育的重要阵地，也是校园精神文明建设的窗口。人类凭借经验积累与形象思维，擅长做出不精确的、定性的决策，而人工智能可以凭借极快的速度完成准确的、定量的计算。智能技术的嵌入应用

使高校治理中大量包括身份辨别、安全保卫、文字记录、图像识别、决策选择等传统上基于诸多人力的工作由智能系统代替，压缩了冗长的科层链条和漫长的手续流程，推动了高校治理中人力资源的优化配置。比如，山东建筑大学学生工作部公寓服务中心运用 AI 技术在全校 16 栋学生公寓建立了人脸识别门禁系统，通过智能设备和管理系统的深度协作，构建“人防”+“物防”+“技防”三级联动机制，有效控制外来人员随意进入，保障学生人身及财产安全（张敬如等，2023）。

4.4. 心理健康监测

学生心理健康问题来源于各方面综合因素，仅仅单方面做出努力，难以根本上解决学生心理问题，效果大打折扣。运用大数据、人工智能等新一代信息技术推动学生心理健康智能监测的信息化、精准化，融合人才培养质量和学生安全保障，以数据驱动高校大学生心理健康安全智能监测，实现高效育人，助力高校解决学生心理健康教育工作的难题。南方科技大学凭借丰富的算法模型和建设经验，结合多维心理测评建立人工智能算法模型，对学困生的心理行为进行智能监测与针对性训练，做到自动识别潜在风险点，如针对学习成绩、学生作息、行为轨迹、社交等维度，将“事后分析”变为“事前预测预警”。

4.5. 学生评价

《深化新时代教育评价改革总体方案》提出要“创新评价工具，利用人工智能、大数据等现代信息技术，探索开展学生各年级学习情况全过程纵向评价、德智体美劳全要素横向评价”。目前学生评价工作透明度不高、公正性不强，根源在于未充分利用互联网思维、技术和模式提升评价质量。通过“区块链+AI”助力高校学生评价，利用区块链保存数据，实现可信的数据共享，利用 AI 的安全计算平台产生更智能的安全空间，以区块链和 AI 为抓手，设计一种网络架构，实现学生管办评分离，构建信任的协同创新模式，克服传统教育评价信息模态单一、缺乏过程性数据、反馈滞后的缺点，建立科学的教育评价提供了全方位支持（刘宗妹，2020）。

4.6. 学生就业

传统就业工作的重心相对集中在末端的就业季，育人作用体现不够充分，大学生的职业成长也普遍缺乏规划。人工智能服务于高校毕业生的就业指导工作，借助智能系统对学生的专业能力、职业意向、薪资要求等信息进行深度挖掘和动态追踪，生成毕业生的就业画像，再经由招聘信息流的混排算法，依托关系链匹配并及时公布、推送供需信息，实现海量就业信息下求职和招聘的精准化配对，提升高校就业指导工作的实际成效。比如，中国人民大学推出“五融一体”的智慧职业发展中心平台，建立了四个方向的胜任力模型，设置了六阶成长系统，开发 AI 面试模拟、智能简历分析等助力学生就业（杨菲菲，2024）。

4.7. 舆情风险治理

智能技术能够实时传递信息并模拟现场情况，帮助高校治理主体实现科学决策和远程指挥。例如，在高校舆情风险治理方面，大数据、人工智能技术的数据挖掘、交叉检验、多样本之间统计分析，有助于探索热点舆情的关联原因、内在规律及发展趋势，进而提升新时期高校舆情治理的品质和效率，为校园和谐稳定创设健康的生态环境（姚翼源等，2023）。

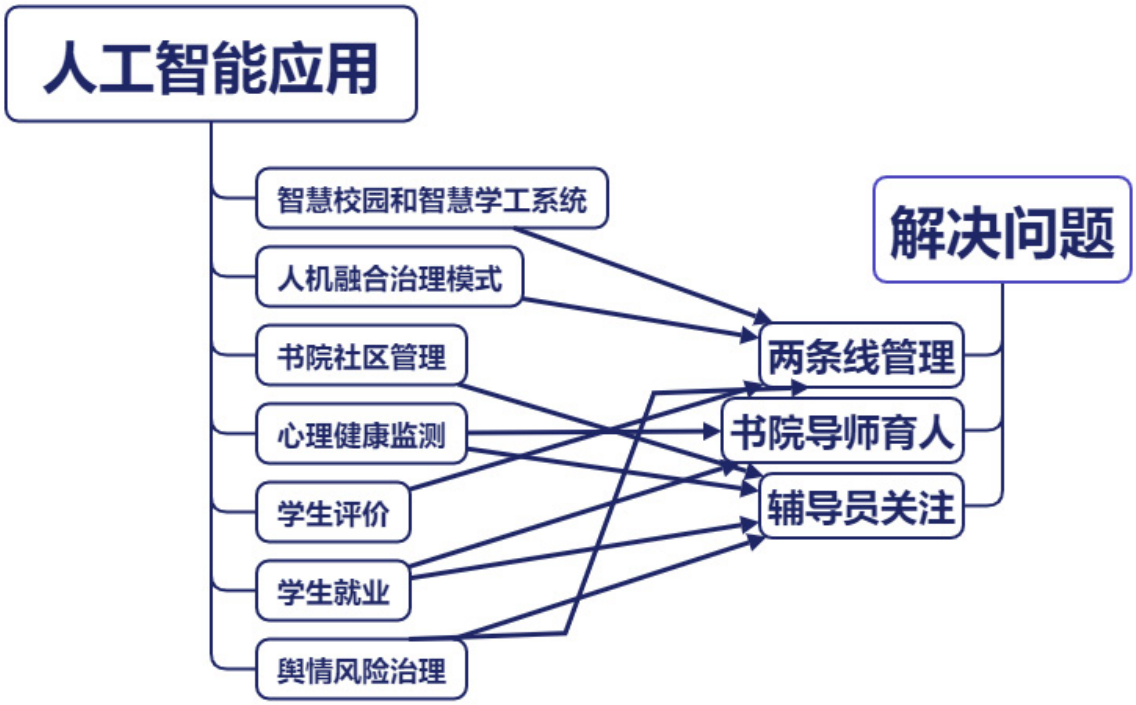


图 2 *人工智能管理应用与解决书院制高校学生管理问题一览*

5. 人工智能赋能书院制高校学生管理的创新路径

人工智能技术的渗入改变了高校治理的生态，并在客观上对高校的治理环境与模式提出了新的诉求。因此，面对人工智能带来的机遇和挑战，应当从高校治理理念、治理体系、治理能力、治理模式、治理制度等方面着手，消除阻碍人工智能技术嵌入应用的不利因素，为推动高校治理的智能化发展营造良好的环境。

5.1. 从“以管理为中心”向“以师生为中心”转变

科学技术必须与人文精神相融合，独特的人文精神，是人类区别于人工智能的宝贵价值（张道正，2019）。在人工智能赋能书院制高校学生管理的过程中，关注人工智能效益最大化的同时也要时刻警惕人文精神的流失。在人工智能赋能书院制高校治理的过程中，过度强调智慧化与标准化，容易忽视以人为本的教育内核，这与书院制通识教育和综合素养的培养相违背，应当构建以人为中心、以师生为中心、具备人文精神的高校智能化治理系统。

5.2 提升师生数字素养与数字胜任力

面对教育高质量发展的迫切需求和以人工智能为代表的新一代信息技术变革的挑战，必须把数字素养与胜任力作为师生的必备素养。2022 年底，教育部发布《教师数字素养》标准，对教师数字素养概念作了阐释，构建了包含数字化意识、数字技术知识与技能、数字化应用、数字社会责任、专业发展 5 个一级维度，以及 13 个二级维度和 33 个三级维度的素养框架。同样，应该重视学生数字素养与数字胜任力的提升。北京师范大学发布了《青少年人工智能素养核心框架》，认为青少年 AI 素养框架应该包含青少年 AI 认知，青少年 AI 能力，以及青少年 AI 安全与伦理三个方面（贾红，2024）。

5.3 建设人工智能治理共同体

建立有效的信息共享机制和反馈机制，减少学生管理部门之间因信息不对称和数据失真等因素对人工智能赋能书院制高校学生造成不利影响。搭建和优化书院制高校学生管理云平台，打破主体间的数据信息壁垒，推动数据信息全面流通，从而强化多元主体互联互通的学

生管理格局。建设人工智能治理共同体，建立跨部门、跨院系的智能技术共享与协同机制，全面统筹、协调和推进学生管理的智能化推进工作。

5.4 促进智能产品与治理场景适配

在推动高校治理智能化发展的实践过程中，智能产品与治理场景不匹配、不适应的现象屡屡出现，严重阻碍了人工智能技术在高校治理中应用的深度和广度。因此，高校治理能力的提高必须着眼于科技层面的突破，增强人工智能与高校治理的适配性，打造优质的高校治理生态（申国昌等，2023）。不同书院制高校在地理位置、书院文化、书院组织形式等方面存在差异，要根据不同的学生管理需求，从书院制学校的独特性出发，不断优化或者不断研发创新适配性更强的智能产品，以适应书院制高校学生管理的个性化发展。

5.5 建立风险防范与监管机制

随着人工智能与书院制高校学生管理的深度融合，诸多不可预测的风险逐渐浮现，主要包括数据信息泄露引发的伦理冲突、技术滥用带来的安全威胁以及算法管理缺陷导致的安全隐患。因此，在人工智能技术的选择与应用中应始终保持警惕，注意做好风险防范，建立反馈和监管机制，需要进一步加强技术防护和网络安全队伍建设，提升基础防护能力，全面跟踪高校治理实践中人工智能产品的使用情况，并根据评价结果做出针对性调整，确保在保障数据安全、算法安全的基础上提高治理效率。

6. 结语

教育部部长怀进鹏强调，人工智能是把“金钥匙”，不仅影响未来的教育，也影响教育的未来，这里有机遇也有挑战，并提出要把人工智能技术深入到教育教学和管理全过程、全环节（2024）。书院制是中国高等教育的有益探索，但目前面临两条线管理职责不清、导师育人效果不佳、辅导员对学生关注不足的三重困境。通过人工智能在智慧校园和智慧学工系统、人机融合治理模式、书院社区管理、心理健康监测、学生评价、学生就业、舆情风险治理七个方面应用于书院制高校学生管理，为破解书院制高校学生管理难题提供了新的思路。通过“以管理为中心”向“以师生为中心”的转变、提升师生数字素养与数字胜任力、建设人工智能治理共同体、促进智能产品与治理场景适配、建立风险防范与监管机制，书院制高校学生管理质量与水平将大大提高，人工智能赋能书院制高校学生管理大有可为。

参考文献

习近平(2022). 高举中国特色社会主义伟大旗帜为全面建设社会主义现代化国家而团结奋斗. 人民日报,11568,1.

习近平(2023). 习近平在中共中央政治局第五次集体学习时强调 加快建设教育强国为中华民族伟大复兴提供有力支撑. 新华社. http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/202305/t20230529_1061907.html

王艺超、李佳为(2023). 聚焦智慧民航 | “人工智能+民航”探索正在升级. 中国民航报,11,5.

叶志明(2022). 四大难题困扰高校书院制探索. 中国科学报,9,7.

申国昌和申慧宁(2023). 人工智能赋能高校治理现代化：内涵、愿景与路径.现代教育技术,10,5-13.

刘宗妹(2020). “区块链+AI”助力高校学生评价体系研究.中国教育信息化,21,38-42.

祁叶达和徐小强(2024). 人工智能赋能高校辅导员工作高质量发展探析.学校党建与思想教育,1,18.
李凤亮(2023). 现代书院制：高校正在探索的人才培养模式.学习时报 6,6.
李禾和李诏宇(2023). 书院制：变革高校人才培养固有模式.科技日报,12453,6.
怀进鹏(2024). 民生主题记者会上，教育部部长怀进鹏答问全记录.教育部新闻办. https://mp.weixin.qq.com/s/m0glaNjqAIqT0EMcVbx8Dg
张芳(2022) 高校智慧校园的建设实践及创新应用——以中央财经大学为例. 现代教育技术,4,113-125.
张敬如和李静雯(2023). 山东建筑大学：基于 AI 技术构建高校学生公寓管理新模式.全国高校思想政治工作网. https://www.sizhengwang.cn/a/dxjy_xyaq/230504/1493750.shtml
张道正(2019). 郝平：科学技术必须与人文精神相融合.北京大学新闻网. https://news.pku.edu.cn/mtbdnew/de6563249f164928a8d8415e8ef536e5.htm
杨菲菲(2024). 人民大学开发智慧职业发展中心平台,用 AI 技术助力学生就业.新京报,3,5.
姚翼源和阮一帆(2023). 数智技术嵌入高校舆情风险治理的新变化及其实现路径.西南民族大学学报(人文社会科学版),1,139-145.
徐青山、张建华和杨立华(2016). 高校智慧校园建设的顶层设计及实践应用——以“智慧北航”为例.现代教育技术,12,112-118.
贾红(2023). 北师大发布《青少年人工智能素养核心框架》《生成式人工智能未成年人保护和发展评估指标体系》.北京师范大学网站. https://news.bnu.edu.cn/zx/xzdt/2a277af29bd14905b8667e6b3b59d5ef.htm
管培俊(2021). 人工智能时代的高等教育治理.中国教育在线. https://news.eol.cn/xueshu/hui/202105/t20210527_2114968.shtml
IBM(2023). 什么是人工智能 (AI)？IBM 官网. https://www.ibm.com/cn-zh/topics/artificial-intelligence

机器评估与人工评估视角下不同学业投入学生创造性思维特征探究

Exploring Characteristics of Creative Thinking in Students with Different Levels of Academic Engagement: Perspectives from Machine and Human Assessment

赵亚南，毛云庆，普琳洁，唐天奇，沙景荣*
西北民族大学
* shajr818@163.com

【摘要】 借助大语言模型的发展浪潮，我们进行了一项实证探索，着重研究学生在同伴反馈中创造性思维认知的研究过程。基于“人工评估”和“机器评估”两种评估方式，从认知网络分析角度对中文提问方式下的反馈效果比较发现：（1）ChatGPT 的反馈效果在不同认知行为方面呈现不同水平；（2）相较于人工评估，机器评估反馈态度较为中立；(3) 机器评估的结果更为丰富。为了最大化 ChatGPT 在提升学术评论分析能力方面的优势，教师需要重新审视现有的教学评价方式，增强数据规范意识，严格把控机器生成的反馈质量。

【关键词】 同伴反馈；创造性思维；ChatGPT；认知网络分析；教学评价

Abstract: *With the development of large language models, we conducted an empirical exploration focusing on the research process of students' creative thinking cognition in peer evaluation. Based on the two coding methods of "human assessment" and "machine assessment", from the perspective of cognitive network analysis, the feedback effect of Chinese questioning mode is compared: (1) the feedback effect of ChatGPT presents different levels in different cognitive behaviors; (2) compared with human assessment, the feedback attitude of machine assessment is more neutral; (3) the results of machine assessment are richer. In order to maximize the advantages of ChatGPT in improving the ability of academic review and analysis, teachers need to re-examine the existing teaching evaluation methods, enhance the awareness of data standardization, and strictly control the quality of machine-generated feedback.*

Keywords: peer evaluation, creative thinking, ChatGPT, cognitive network analysis, teaching evaluation

[基金项目] 甘肃省优秀研究生创新之星项目“教育数字化转型赋能增值评价模型构建与应用”（2023CXZX-190）。

1. 前言

在当今教育领域，同伴反馈作为一种重要的学习方式，被广泛应用于促进学生的学习和发展，然而传统的基于同伴反馈的教学评价往往滞后且质量难以保证，容易出现评价要点遗漏以及反馈单一，片面等问题。在这种背景下，如何有效评估学生在同伴反馈过程中的创造性思维特征，成为了教育研究和实践中的一项重要的挑战。

随着人工智能的迅速发展，生成式人工智能（AIGC）聊天工具如 ChatGPT 为教育领域的数字化转型注入了新的活力。ChatGPT 采用了独特的训练方式，将提示学习与人类反馈相结合，不仅能够深入理解自然语言，还能产生类人的反应，展现出高度的智能与互动性（Lund & Wang, 2023）；它主要基于对输入文本上下文的深入解析，通过精准捕捉前文的信息，推断当前语句的含义，并基于此来生成相应的回复；这种基于自然语言的对话能力使其在教学评价与反馈中具备良好的应用潜力。

鉴于此，本研究尝试在同伴反馈过程中引入 ChatGPT 等人工智能工具对学生创造性思维进行评估，通过采用认知网络分析方法，将不同学业投入水平小组的文本内容进行人工评估和机器评估，以量化表征学习者思维过程并比较评估效果；我们将重点关注 ChatGPT 在教学评价中的应用潜能，以及其对教学质量和效率的影响，并为将 AIGC 技术和工具融入未来的教育教学提供思路和参考。

2. 文献综述

2.1. 生成式人工智能在教学评价中的应用

ChatGPT 等智能技术的引入为教育教学带来了深刻的变革动力，其在推动个性化学习、深化形成性评估以及推动教育教学研究方面凸显出巨大潜力；ChatGPT 可以作为虚拟导师，回答学生的问题并辅助他们进行在线独立学习，帮助学生评估他们的同伴反馈内容，从而使学生能够学会改进他们的反馈(Jia, et.al., 2021)。在减少教师负担方面，教师可以利用 ChatGPT 对学生的作业进行反馈(Qadir, 2023)，还可以用它来辅助教学例如创建各种形式的测试(Zhai, X., 2022)。

国内目前利用 ChatGPT 的研究主要集中于教育评价、支持一线各科教师教学等方面，例如陈思文等（陈思文，孔亚琪和刘宇，2023)对 ChatGPT 评价内容与教师评价内容进行对比，结果表明 ChatGPT 的评价内容情感更为积极，其评价内容与教师评价的相似度高达 75.21%以上，已具备实际应用价值与潜力；王丽等（王丽、李艳和陈新亚，2023）使用提问设计原则，比较发现 ChatGPT 在良好的提问设计下更有助于生成质量较好的反馈。

尽管 ChatGPT 等智能技术在教学评价等领域展现出了显著的潜力，但我们仍需要正视其当前的局限性。具体来说，由于其推理能力和准确性尚未达到完美水平（Borji, 2023），这在一定程度上制约了其在教学评价中的表现；例如它在没有充足的上下文的支持下，可能会无法理解会话（Kalla & Smith, 2023），使得用户很难理解或解释它的响应，所以它暂时还无法为学生提供有意义的反馈（King & ChatGPT, 2023）；此外，ChatGPT 在教育中需要解决隐私、公平、非歧视和透明度(Mhlanga, 2023)等问题，这对 ChatGPT 在教育中的广泛应用提出了挑战。

2.2. 从认知网络角度分析学生创造性思维发展

学习是学习者主动形成认知结构、积极构建知识体系的过程(肖少北，2001)，传统的从总体、静态角度进行评估的学习分析方法具有一定局限，认知网络特征分析法是一种协作学习分析的方法、可以对学习者在认知活动中产生的交互数据进行认知层面的分析，可用于表征和分析深度协作学习的过程（Shaffer, Collier, & Ruis, 2016）。在教育领域，认知网络分析法被广泛应用于学习分析研究中，其独特的可视化表征能力可揭示被研究者的认知差异，将其思维发展过程与认知变化规律可视化，深度追踪和分析被研究者的思维架构和认知发展轨迹(丁继红，2019)。随着自然语言处理的发展，已有研究通过自然语言处理的认知网络自动分析，来提供自适应反馈(Schulz, Meyer, & Gurevych, 2019)。

因此，本研究从认知网络角度分析学生的创造性思维发展，有利于明晰学生在同伴反馈过程中的认知发展特征，有利于教师们在开展深度教学时提供参照，并推动个性化的教学干预以及智慧化全面化评价体系的构建。

2.3. ChatGPT 在文本分析方面的应用

自然语言处理可以胜任数字学习环境中的自动化反馈场景；并且最新一代的大型语言模型（LLM）——ChatGPT 和 GPT-4（OpenAI, 2023），使得自然语言处理在许多领域中得到了新的应用，例如：文本的自动评估、自动论文评分（Zhang & Litman, 2021）等。自然语言处

理的总体目标是根据一个或多个评估指标，精确地根据文本输入来预测目标，该目标包括广泛的任务类型，从文本分类的表面级任务到文本生成等深度语义驱动的任务。

至今，尚未构建出基于自然语言处理的同伴反馈支持的整体框架，这主要是因为同伴反馈数据类型的多样性和教育领域相关任务的复杂性带来了一系列亟待解决的挑战，这些挑战包括数据的稀缺性、结果的可解释性、对潜在偏见的敏感性，以及随着数据隐私保护意识的提升而日益增长的隐私保护需求，这些问题不仅限制了同伴反馈在教学评价中的深入应用，也为自然语言处理技术的发展提出了新的要求。

3. 研究方法与数据分析

3.1. 研究对象

本研究选取西北某高校远程教育应用课程大三学生的在线课堂讨论来进行分析；考虑到学习者们的认知建构过程与其学业投入水平有关，因此对学生们的学业投入水平进行了前测（Fredricks, Blumenfeld, & Paris, 2004），问卷全部收回，有效率为 100%，以测试结果分值的高低为依据，对学生按照 4 人为一组进行组间异质、组内同质分组，随机选取高、低等学生各一组，以他们在课程前中后期的三次在线讨论中认知发展为研究对象进行深入研究，看不同学业投入水平的学生针对课堂主题下的讨论，在人工评估和机器评估下的认知差异。本研究设计流程见图 1。

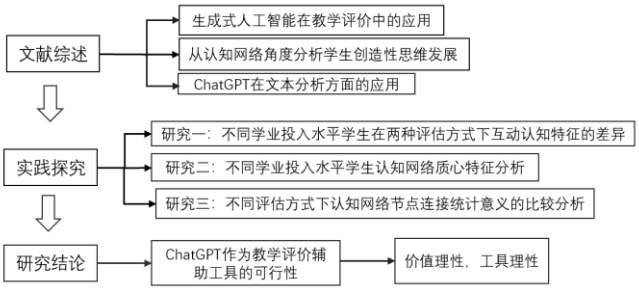


图 1 研究设计流程图

3.2. 评估框架

对学生的讨论文本按照 Sun M（Sun, et.al., 2022）等人的创造性思维框架进行评估(见表 1)。

表 1 评估框架

学习方式所在维度	评估格式	同伴反馈具体维度	解释
发散性思维	A	联想	将概念、物体或者情境联系起来
	B	分解	将概念、物体或情境分为更丰富的细节
	C	调整结合	合并、调整概念、物体或情境
观点的产生	D	新观点	从一个以前没有提到过的新角度产生一个想法
	E	新观点的建构	修改、提炼或扩展以前的想法来发展一个新想法
元认知	F	反思管理	管理和反思这个过程
社会交流	G	细化阐释	用例子、类比、修理或提供细节来解释一个想法
	H	提问	问问题以获取信息或进一步阐述
	I	直接回答	直接回答问题而不详细说明
	J	同意	对一个想法的积极评价（如：接受、支持等）
	K	不同意	对一个想法的负面评价（如：怀疑、拒绝等）
	L	争论	用一个明确的理由对一个观点进行适当性或价值性的辩论
其他	M	无关主题	与给定任务无关的对话
情绪	N	正向	正面的情绪词语或表情
	O	负向	负面的情绪词语或表情

3.3. 数据收集与整理

本次研究共收集了学生们课程前中后期三次主题讨论中共 1126 条讨论数据，数据来源于远程教育应用课程在线 QQ 小组讨论，从中剔除系统通知消息等与课程无相关性的内容，得到的有效评语为 1049 条，分别对数据进行人工评估以及机器评估。人工评估的具体操作流程如下：两位研究者进行合作，对学生们的讨论内容参照创造性思维框架的评价维度进行背对背

标注；随后，研究者们比对标注结果，对不一致的标注进行协商并达成共识，经过一致性计算，五位研究者的标注一致性达到 0.85，表明信度良好。ChatGPT 评估的具体操作程序如下：研究者首先将评估框架以及学生的文本讨论内容以固定格式一起发送给 ChatGPT，逐一获得首次反馈信息，并将其录入 Excel 文档，获得 ChatGPT 对文本讨论的评估分类；其次对不同学业投入水平学生在两种评估方式下的互动特征差异进行分析；之后，研究者使用在线建模平台 ENA Webkit 对两个小组的人工与机器评估结果分别进行认知网络分析，观察两种评估方式在认知发展路径方面的差异。

4. 研究结果分析

4.1. 不同学业投入水平学生在两种评估方式下互动认知特征的差异

本研究通过卡方检验，比较不同评估方式下两个小组在发散性思维、观点的产生、元认知、社会交流、其他及情绪共六个维度上的具体差异，以对比高、低学业投入水平小组互动中认知特征在两种评估方式下的差异，见表 2、表 3。

表 2 高等学业投入水平互动认知特征卡方检验分析结果

行为	评估	code(%)		总计	χ^2	ρ
		human-coded	machine-coded			
发散性思维	A	2(12.50)	17(33.33)	19(28.36)	9.536	0.008**
	B	1(6.25)	15(29.41)	16(23.88)		
	C	13(81.25)	19(37.25)	32(47.76)		
观点的创造	D	20(50.00)	8(38.10)	28(45.90)	0.786	0.375
	E	20(50.00)	13(61.90)	33(54.10)		
社会交流	G	19(24.68)	15(15.96)	34(19.88)	16.438	0.006**
	H	20(25.97)	30(31.91)	50(29.24)		
	I	17(22.08)	8(8.51)	25(14.62)		
	J	19(24.68)	36(38.30)	55(32.16)		
	K	0(0.00)	5(5.32)	5(2.92)		
	L	2(2.60)	0(0.00)	2(1.17)		
情绪	N	78(97.50)	70(95.89)	148(96.73)	0.011	0.917
	O	2(2.50)	3(4.11)	5(3.27)		
总计		271	272	543		

* p<0.05 ** p<0.01 *** p<0.001

表 3 低等学业投入水平互动认知特征卡方检验分析结果

行为	评估	code(%)		总计	χ^2	ρ
		human-coded	machine-coded			
观点的创造	D	13(72.22)	12(70.59)	25(71.43)	1.083	1.000
	E	5(27.78)	5(29.41)	10(28.57)		
社会交流	G	0(0.00)	1(2.56)	1(1.85)	9.743	0.083
	H	13(86.67)	20(51.28)	33(61.11)		
	I	2(13.33)	2(5.13)	4(7.41)		
	J	0(0.00)	5(12.82)	5(9.26)		
	J	0(0.00)	4(10.26)	4(7.41)		
	L	0(0.00)	7(17.95)	7(12.96)		
情绪	N	11(91.67)	8(100.00)	19(95.00)	0.000	1.000
	O	1(8.33)	0(0.00)	1(5.00)		
总计		271	272	543		

* p<0.05 ** p<0.01 *** p<0.001

其中，高等学业投入水平小组在发散性思维和社会交流维度的两种评估方式呈现出较大差异，且差异显著，整体来看，机器评估的方式，对评语中联想、分解、同意、争论行为的评估识别更多；在低等学业投入水平小组，两种评估方式得到的结果无显著差异，间接说明机器评估一定程度上的可行性，但也与低等小组学生在反馈时的评语内容及类型有关。描述性统计分析结果表明，观点的产生以及情绪维度在高、低等学业投入水平小组中占比均较高，

说明这两种行为是同伴反馈过程中的重要特征。相较于低等学业投入水平小组，高等学业投入水平小组学生在发散性思维上的认知行为占比较多，说明高等学业投入水平小组的学生在互动过程中更善于进行联想、分解、调整结合行为。

4.2. 不同学业投入水平学生认知网络质心特征分析

为探究学生的文本讨论内容在两种评估方式下认知网络的发展特征差异，本研究将学生们的讨论按照课程讨论进度划分为不同阶段，并以活动阶段和学生个体为单元展开分析。

质心是认知网络结构的缩影，代表其网络模型连接结构的特定排列。本研究针对人工评估方式与机器评估方式，分别对高低等投入水平小组学生讨论的网络质心情况绘制了认知网络质心图，如图 2（低等组）、图 3（高等组）所示。每一个图中的圆点对应该组学生的创造性思维认知要素，正方形表示特定评估方式下所有学生认知网络结构的平均质心，虚线框表示95%置信区间。其中橙色部分表示学生在课程前期讨论的认知网络表现情况，绿色的部分为课程中期讨论的最终认知呈现，粉色部分为课程后期学生讨论的认知网络表现情况。

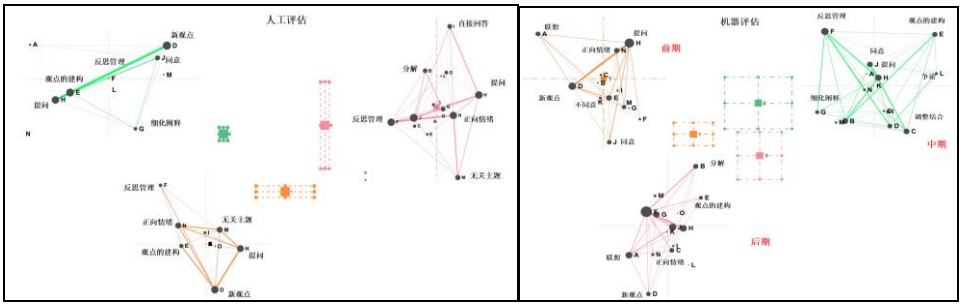


图 2 低等学业投入水平学生在两种评估方式下的认知网络结构差异

可以看出：首先，低水平组在前中后三次讨论中的质心位置明显分离，说明学生们在讨论过程中认知发生了差异；其次，在不同评估方式下其三次讨论的平均网络质心位置发生了较大变化，学生在两种评估方式下认知网络结构存在直观的差异，例如在中期讨论中，人工评估的方式下其平均网络质心位于Y轴左侧，但在机器评估的方式下位于Y轴右侧，正是由于机器评估中有了更多其他节点之间的连接，尤其是反思管理、观点的建构等边缘节点与其他节点的连接使得其质心位于Y轴右侧，X轴高处。从节点之间的连接来看，低水平组在人工评估方式下的认知网络结构图更集中于提问–新观点–新观点这些行为的建构，而在机器评估方式下的节点更加丰富，认知网络结构图更复杂，呈现出更多认知要素与其它节点的强连接。

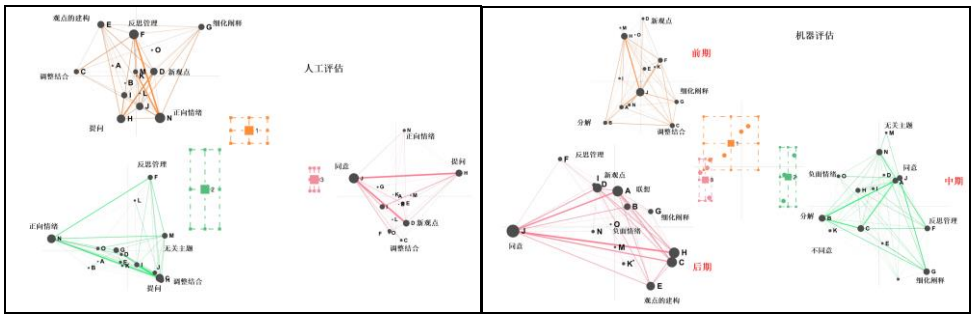


图 3 高等学业投入水平学生在两种评估方式下的认知网络结构差异

对高等学业投入水平小组进行分析发现，其三次讨论的质心明显分离，说明学生在讨论过程中认知发展也有较大差异。其认知网络结构图相较于低等较为完善，但在后期讨论（粉色）中的人工评估和机器评估差异较大，具体表现为机器评估的节点更为丰富，相比较于人工评估中仅有的提问–同意，新观点–同意之间的强连接，在机器评估中呈现出更多其他认知要素例如联想、细化阐释、分解等与同意节点之间的连接，这与讨论文本内容的数据类型、数据质量、人工评估的相对主观性均有关系。

4.3. 不同评估方式下认知网络节点连接统计意义的比较分析

为了从统计意义上更直接地观察两种评估方式下的效果差异，本研究以低等学业投入水平小组为例，计算了两种评估方式下学生在第一次互评过程中的认知网络连线系数值，如图 4 所示。

图 4 低等学业投入水平人工评估和机器评估的认知网络连线系数值

节点连接	人工评估	机器评估	节点连接	人工评估	机器评估	节点连接	人工评估	机器评估
新观点-提问	0.37	0.39	提问-无关主题	0.28	0.16	联想-新观点的建构	-	0.15
新观点-不同意	-	0.02	提问-正向情绪	0.26	0.27	联想-无关主题	-	0.10
新观点-反思管理	0.22	0.05	提问-负向情绪	0.11	-	联想-正向情绪	-	0.03
新观点-无关主题	0.20	0.16	提问-直接回答	0.13	0.04	联想-不同意	-	0.05
新观点-新观点的建构	0.32	0.22	反思管理-正向情绪	0.07	-	联想-提问	-	0.15
新观点-直接回答	0.07	0.04	反思管理-无关主题	0.06	0.02	联想-调整结合	-	0.05
新观点-正向情绪	0.27	0.16	反思管理-直接回答	0.02	0.02	联想-同意	-	0.09
新观点-同意	-	0.08	反思管理-提问	0.08	0.15	不同意-正向情绪	-	0.03
新观点的建构-提问	0.08	0.20	反思管理-不同意	-	0.02	调整结合-不同意	-	0.03
新观点的建构-反思管理	0.12	0.04	细化阐释-提问	-	0.05	调整结合-反思管理	-	0.02
新观点的建构-正向情绪	0.10	0.07	细化阐释-同意	-	0.02	调整结合-新观点	-	0.06
新观点的建构-无关主题	0.14	0.05	直接回答-正向情绪	0.11	0.10	分解-不同意	-	0.08
新观点的建构-细化阐释	-	0.04	直接回答-无关主题	0.14	0.02	同意-正向情绪	-	0.07
新观点的建构-不同意	-	0.05	直接回答-同意	-	0.02	同意-无关主题	-	0.03
新观点的建构-同意	-	0.10	调整结合-同意	-	0.02	同意-不同意	-	0.08
无关主题-正向情绪	0.30	0.15	调整结合-提问	-	0.03			
联想-新观点	-	0.14	调整结合-新观点的建构	-	0.02			

通过上图中数值及空值对比，可知机器评估方式比人工评估反馈的节点连线种类更多，说明其评估的结果更详细丰富，这也是造成二者产生较大差异的原因；同时存在较多的节点连接，例如新观点–提问，新观点–新观点的建构，新观点–直接回答，提问–正向情绪等，连线系数值几乎一致（相关性一致），说明在以人工评估为基准时，机器评估的可行性，但仍不可忽略存在差异的方面，比如存在人工或者机器评估方式下超过 0.20 的连接，但另一种评估方式节点之间并非强关联的现象，例如新观点–反思管理，新观点–无关主题，新观点–正向情绪，说明机器在对此类评语进行评价反馈时，无法较好地对学生的行为进行评估，这也间接印证了机器评估与人工评估依然无法保持较高一致性的问题。

5. 研究结论与启示

5.1. 研究结论

通过比较两种评估方式对学生文本讨论内容的评价与反馈效果，得出以下结论：

第一，ChatGPT 对学生文本讨论内容的反馈效果在不同的语句类型中呈现不同的质量。具体表现为，正向情绪、无关主题、新观点创造行为识别与人工评估相对一致，当学生的认知行为处于发散性思维时，ChatGPT 较难正确识别，其原因可能有 ChatGPT 不是为了教育内容文本分类而创生的，其次还会受到数据内容类型、质量等的影响等，可通过优化提问方式，进行追问等手段优化。

第二，ChatGPT 因其基于大语言模型构建的优势，对文本具有更强的理解能力，能够通过理解文本内容的上下文，从中提取评价要点，在评价过程中表现出良好的潜力，在本研究中表现为识别效果更细腻，评价效果比人工评估更为丰富；但也可以看出，ChatGPT 更擅长对学生所表达的句意明确的文本内容进行划分并给出反馈。此外，观点的产生这一行为以及情绪维度在高、低等学业投入水平小组中占比均较高，说明这两种行为是同伴反馈过程中的重要特征，并且高等学业投入水平小组的学生在互动过程中更善于进行发散性思维相关的行为。

5.2. 研究启示

基于以上研究发现，ChatGPT 等 AIGC 技术和工具在教学评价与反馈方面展现出了巨大的潜力，本研究得出以下两方面的启示：

首先，从效率与成本的角度来看，ChatGPT 编码因其自动化和高速处理的能力，显著提升了编码工作的效率，降低了人工成本；这提示我们，在处理大规模文本数据时，采用机器编码可以极大地减少人力投入，提高工作效率。

其次，从评估准确性来看，ChatGPT 底层认知逻辑并非完全等同于人类认知，而是通过统计建模来实现的，并不是基于对真实世界的经验和理解，它被多次证明在反馈时会存在错误（Hosseini, Rasmussen, & Resnik, 2023），因此教师需要提升自身的数字化教学能力，扮演好确保 ChatGPT 融入教育评价“把关人”的角色，使 AIGC 等智能工具能够提供更为准确的生成反馈信息。通过对比分析，我们也可以更深入地理解机器评估在哪些情况下可能优于或劣于人工评估，从而为实际应用提供更准确的指导。

最后，我们要保持价值理性和工具理性之间的平衡，避免过度依赖技术，从而引发学习惰性等风险。综上所述，尽管科技赋能教育评价仍然处于初级阶段，但随着技术的不断进步，ChatGPT 等智能工具在教育领域的应用前景广阔，未来研究可以进一步探讨机器评估与人工评估的优势融合方式等。

参考文献

丁继红.（2019）. 深度学习中的学习者认知网络和动机策略分析——旨向深度学习的 U 型翻转教学效果研究. 远程教育杂志（06），32-40.

王丽、李艳、陈新亚和徐翎衲.（2023）. ChatGPT 支持的学生论证内容评价与反馈——基于两种提问设计的实证比较. 现代远程教育研究（04），83-91.

闫志明、朱友良和刘方媛.（2022）. 新一代信息技术支撑的教育评价:价值诉求、现实问题与建设进路. 现代教育技术（11），34-41.

肖少北.（2001）. 布鲁纳的认知——发现学习理论与教学改革. 外国中小学教育（05），38-41.

陈思文、孔亚琪和刘宇.（2023）. 基于生成式人工智能的学业评价应用研究——以 ChatGPT 为例. 软件工程（10），27-31.

林郁鹏.（2023）. 基于 MacBERT__RCNN 的反思文本自动化分类系统设计与实现（硕士学位论文，西南大学）.

Borji, A. (2023). A categorical archive of chatgpt failures. *arXiv preprint arXiv:2302.03494*.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

Hosseini, M., Rasmussen, L. M., & Resnik, D. B. (2023). Using AI to write scholarly publications. *Accountability in research*, 1-9.

Jia, Q., Cui, J., Xiao, Y., Liu, C., Rashid, P., & Gehringer, E. F. (2021). All-in-one: Multi-task learning bert models for evaluating peer assessments. *arXiv preprint arXiv:2110.03895*.

Kalla, D., & Smith, N. (2023). Study and analysis of chat GPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, *8*(3).

King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and molecular bioengineering*, *16*(1), 1-2.

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries?. *Library hi tech news*, *40*(3), 26-29.

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. In *FinTech and Artificial Intelligence for Sustainable Development: The Role of Smart Technologies in Achieving Development Goals* (pp. 387-409). Cham: Springer Nature Switzerland.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional science*, *37*, 375-401.

Nisar, S., & Aslam, M. S. (2023). Is ChatGPT a good tool for T&CM students in studying pharmacology?. *Available at SSRN 4324310*.

OpenAI. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Qadir, J. (2023, May). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE.

Schulz, C., Meyer, C. M., & Gurevych, I. (2019, July). Challenges in the automatic analysis of students' diagnostic reasoning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 6974-6981.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, *3*(3), 9-45.

Sun, M., Wang, M., Wegerif, R., & Peng, J. (2022). How do students generate ideas together in scientific creativity tasks through computer-based mind mapping?. *Computers & Education*, *176*, 104359.

Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*.

Zhang, H., & Litman, D. (2021). Essay Quality Signals as Weak Supervision for Source-Based Essay Scoring. *Grantee Submission*.

基于知识图谱可视化表征的形成性反馈设计

Formative feedback design based on visual representation of knowledge graph

张灵琳*，马志强
江南大学 江苏“互联网+教育”研究基地
* 972312018@qq. com

【摘要】 形成性反馈是改善学习效果的重要策略。但反馈的有效性在教与学实践中往往难以保证。为适应数字化教育时代下的学生个性化需求，有必要寻找使反馈价值最大化的途径。知识图谱因其在可视化表征、信息承载、知识建构方面的优越性，为形成性反馈优化与改进带来了新的发展契机。因此，本研究梳理了形成性反馈设计过程与方法，构建了基于知识图谱可视化表征的形成性反馈设计模型，并具体设计了个体与群体两类可视化反馈图谱，用以表征学习者在学习进展、认知发展、知识达标、学习路径方面的反馈信息，为相关研究提供了新的视角和参考。

【关键词】知识图谱；可视化表征；形成性反馈；教学反馈

Abstract: *Formative feedback is a critical strategy for enhancing learning outcomes. However, the effectiveness of feedback is often challenging to guarantee in the context of teaching and learning practices. To accommodate the personalized needs of students in the digital education era, it is imperative to seek avenues that maximize the value of feedback. The knowledge graph, with its superiority in visual representation, information encapsulation, and knowledge construction, presents new opportunities for the optimization and enhancement of formative feedback. Consequently, this study has systematically organized the design process and methods of formative feedback, developed a design model for formative feedback based on the visual representation of the knowledge graph, and specifically designed two types of visual feedback maps - individual and group - to represent feedback information regarding learners' learning progress, cognitive development, knowledge acquisition, and learning trajectories. This study offers new perspectives and references for related research.*

Keywords: Knowledge graph, Visual representation, Formative feedback, Teaching feedback

1. 引言

反馈作为形成性评价的核心组成部分，在改进教与学方面发挥着越来越重要的作用。评价中的形成性反馈，是指向学习者传递的旨在改变其思维或行为以提高学习效果的信息，已被证明能够缩小学习者当前理解或表现与预期目标之间的差距，降低认知负荷并改善学习效果。要在实际教学中提供有效的反馈，不仅需要考虑反馈的时机与形式，也要关注到学生的个体差异。很多调查发现，教师提供有效反馈的频率非常低，经常偏离有效反馈原则，更倾向于使用笼统的、宽泛的表扬，缺乏实质性内容（Voerman et al. , 2012）。目前，教师主导的传统文本型反馈方式仍占主流，难以真实反映学生在知识习得、认知发展过程中的微观细节表征，容易存在信息延时、个性化缺失、过程性数据匮乏、直观性不强等问题，并且已然无法精准地对一些使用新技术的教学模式进行反馈，如大规模开放在线课程（MOOC）以及翻转课堂等，亟需从技术与策略的角度对优化智能学习环境下的形成性反馈评价进行更深入的探索。

随着数字化阅读时代的到来，技术驱动的反馈评价方式逐渐兴起，尤其是可视化表征技术与工具的介入，人们有了更多机会直观获取关于学生过程性表现和评价的数据，更精确地分析学习过程。其中，知识图谱（Knowledge Graph）作为可视化表征的重要技术分支，能够

基于图形思维将知识元素、属性和关系等信息进行关联并统一建模，已在教育教学领域逐渐释放其价值。基于知识粒度的图谱不仅能结构化地呈现知识的复杂建构过程，还能以语义图示的方式给予教师及学习者即时性、形成性的反馈与评价，能够有效解决当前复杂学习情境下的反馈问题。基于此，本研究尝试从知识图谱的可视化表征视角出发，探讨其对于优化和改进形成性反馈评价的设计与应用，以对相关研究提供一定的理论依据与实践参考。

2. 文献综述

2.1. 形成性反馈

起初，反馈被用于教育领域时，常被理解为“告诉”的同义词，指由教师流向学习者的单向信息传递，学习者通常被动接受信息，难于对自身学习过程与能力形成真实感知。进而，在各大学习理论流派从心理学范式的视角深入探索学习的发生机制后，反馈的内涵也随之得以深化。行为主义心理学用“反馈”解释行为的结果如何影响未来行为发生的可能性。在斯金纳提出的程序教学理论中，及时反馈被视为基本原则之一，强调多次及时的反馈信息对于反应强化的重要性。基于认知主义视角看，只有缩小实际水平与目标水平之间差距的信息才是真正的反馈（Ramaprasad, 1983）。Sadler 认为知识与技能的提升往往发生在有反馈循环的支持性环境中，并提出了反馈的三个阶段：1）明确学习要达到的目的和标准；2）要把目前的发展阶段和目标阶段进行对照；3）要有针对性地弥合两者间的空白。这将反馈与形成性评价的理念联系起来，凸显了反馈的形成性取向，即通过对学生表现的持续性评价，得到有助于改进和优化后续学习的信息（伍绍杨，2021）。

在此逻辑下，大批学者从不同角度对形成性反馈进行了多方面的研究，包括反馈设计要素、作用过程等。Ilgen 等人将反馈源、反馈信息和反馈接收者视为反馈的三个元素，并在此基础上构建了一个四阶段的接收者反馈过程模式，包括感知反馈、接受反馈、预期反馈和计划反馈。Bangert 等人提出学习者接受反馈的五阶段模型，将学习者置于接受反馈的中心位置，其认知状态经历了初试状态、激活检索策略、即时反应、评估与调整五个阶段。Hattie 在大量实证研究基础上提出了反馈模型，将反馈定义为由某一主体（如教师、同伴、互联网等）提供的关于某人的表现或理解等方面的信息，认为有效反馈必须回答三个重要问题：我要去哪里？我进展如何？下一步去哪里？分别称为“正馈”、“后馈”和“前馈”，说明了反馈内容的设计应包含目标、过程进展与改进策略三个方面。

综上，已有研究基于理念与模型探讨了反馈的内容以及使反馈产生作用的关键要素，为反馈的设计与开展提供了基本环节的参考，但鲜少关注如何基于学习者特征呈现更适合他们的反馈，并反映和区分不同学生在知识习得、认知发展过程中的微观细节表征。尤其在各种新型学习媒体与资源在课程教学中广泛应用的大时代背景下，学习者在学习中的过程性数据更为复杂多元，如何以更好的方式表征多样化的反馈信息并更加直观形象地传递给学习者仍值得深入研究。

2.2. 基于可视化表征的形成性反馈

在教与学的实践中，为将形成性反馈的效用最大化往往需要考虑更复杂的维度，包括反馈的时机、互动结构及媒介（伍绍杨，2021）。同样地，Narciss 等人认为应在考虑教学目标和学习者特征的基础上适时调整反馈信息的内容、功能和呈现格式，以最大限度地提高反馈的信息价值。在传统的课堂中，反馈常以口头或书面语言的方式传递，教师往往不能全面顾及所有学生，且容易忽视个性化需求。随着信息技术的发展，一大批自动化反馈工具涌现，如智能问答机器人、自动化考试平台等，能够基于人机交互原理为学习者的学习参与（如答题、测试等）提出及时的反馈意见。但就目前而言，大多数自动化反馈系统提供的反馈信息往往停留在任务层面，如识别错误等，缺乏过程和自我调节层面的反馈，即关于策略或如何改进的提示等，难以促进学习者的深层学习（Keuning, 2018）。

为寻求更加高效的信息传达，以语义图示和符号实现可视化知识表征的相关技术兴起，

并逐渐融入到教育反馈与评价中。表征是指探讨符号如何表现“意义”，可视化表征是指以视觉符号为中介进行的意义交流，是借助于技术实现的外部表征。现有的智能化教学平台中常使用统计图表、图片等各种形式化图示来表征和分析学习内容与过程。常见的如用折线图呈现学习者在一段时间内的在线学习趋势与变化情况，用堆积条形图可视化展示多项错因的频率和比例关系等。如吴东平等人基于 P&P 智慧课堂平台，借助学生画像、表情标签等视觉要素对学生的课堂纪律、学习表现和学习效果等开展形成性评价，并即时反馈给学生；黄昌勤等人探讨通过雷达图、标签云、仪表盘等动态可视化方式呈现学习者云空间的行为关联数据，为他们提供及时反馈、监督与指导。已有研究表明，学习的发生通常伴随着知识结构的生成，知识的外在呈现方式深刻地影响着学习者对知识的认知、理解。现常用的图示工具虽能够在一定程度上反映出学习者阶段性的学习行为序列及效果，却不能详细呈现生成性的知识脉络。而知识图谱作为一种揭示领域知识、概念、关键内容等实体之间关系的结构化语义分析方法，本身就融合了知识表征、可视化、知识建模的多种理念，在提供形成性反馈信息、辅助形成深度理性化的知识体系方面有着巨大发展潜力。

总之，形成性反馈评价的设计随着技术与教育的深度融合发展呈现出新的研究范式，在内容呈现、流程实施方面都具有不同特点。目前，国内外的相关研究较少顾及到学习者反馈信息的个体差异化，也未重视到学习者知识建构过程与完整体系脉络的反馈。

2.3. 教育知识图谱的反馈功能

知识图谱作为一种复杂数据分析和决策支持的有效手段，在教育领域也深受重视。已有研究表明，知识图谱在教育大数据智能化处理、教学资源语义化聚合、适应性学习诊断方面具有重要应用前景（李振，2019）。关联主义学习理论指出，知识是由事实、概念、命题和原则等要素通过关系构成的网络，而学习的本质也就是增强学习者对该网络的认知，建构学习者个体的认知结构（Joseph, 2016）。基于这一视角，教育知识图谱在表达教学过程各要素及有教育意义的认知关系的同时，还叠加了学习者的知识掌握状态，助力构建学习认知地图（余胜泉，2019）。此外，教育知识图谱以知识空间理论、认知诊断理论为基础，以统计测量、学习分析、数据可视化、数据挖掘等智能技术为手段，能够赋能自适应和差异化的测评过程，为表征目标达成度、认知及知识维度提供了粒度载体，根据知识间蕴含的前驱和后继关系呈现学习者认知风格、知识状态等特征，并提供个性化学习路径推送（李振，2019）。这与有效反馈所提到的三大要素：当前状态、预期状态及缩小两种状态间的差距（Hattie，2007），具有高度的弥合性。

综合上述研究成果，本研究认为，知识图谱作为一种组织和表征知识的工具，反映了知识点的先后顺序及认知依赖关系，提供了较为完善和科学的学习认知表征体系，基于知识图谱设计的形成性反馈符合知识的认知规律与教学的基本逻辑，为解决现有形成性反馈信息延时、个性化缺失、过程性数据匮乏、直观性不强的问题提供了契机。基于此，本研究尝试以知识图谱的可视化表征为立足点，探讨其对于优化和改进形成性反馈评价的设计与应用，以对相关研究提供一定的理论依据与实践参考。

3. 知识图谱可视化表征驱动的形成性反馈设计

基于多种反馈类型的大量研究，学者 Narciss 和 Huth 曾明确提出一个形成性反馈设计概念框架，指出在反馈设计时应考虑的三个主要元素：（1）反馈的内容信息，如学习状态、所达到的学习水平、与学习标准的距离等，或与主题、任务、错误或解决方案相关的附加信息；（2）反馈的功能，如促进信息处理的认知、正确反应的强化或学习动机的唤起等；（3）反馈的呈现，如呈现的频率、时间、模式、数量、形式等。国内学者李振等人所提出的教育知识图谱技术框架中，阐述了图谱构建的基本步骤及技术路线，主要包括知识本体构建、命名实体识别、实体关系挖掘、知识融合四个方面。综合知识图谱构建的基本流程，本研究认为，知识本体构建与命名实体识别是描述知识体系框架、构建图谱的基础，主要是提取与识别教

学资源与内容中的实体对象，如知识点、学习活动等，涉及到了反馈的内容信息层面；实体关系挖掘是判断实体间存在的语义关联关系，将知识组织成具有一定映射关系的网络，如认知层次维度、认知前驱后继关系等，涉及到了反馈的功能层面；而知识融合是对来自多个数据源的实体信息进行融合组织并生成最终知识图谱的过程，涉及到了反馈的呈现层面。

基于上述两个框架可知，教育知识图谱的技术框架与形成性反馈的设计理念之间具有一定的逻辑关联，为此本研究将以图谱构建技术流程为主线，融入形成性反馈设计理念的基本要求，利用实体、关系、组织分别表征反馈中的内容信息、功能关联、呈现形式，构建知识图谱驱动形成性反馈的三重可视化表征模型，如图 1 所示。

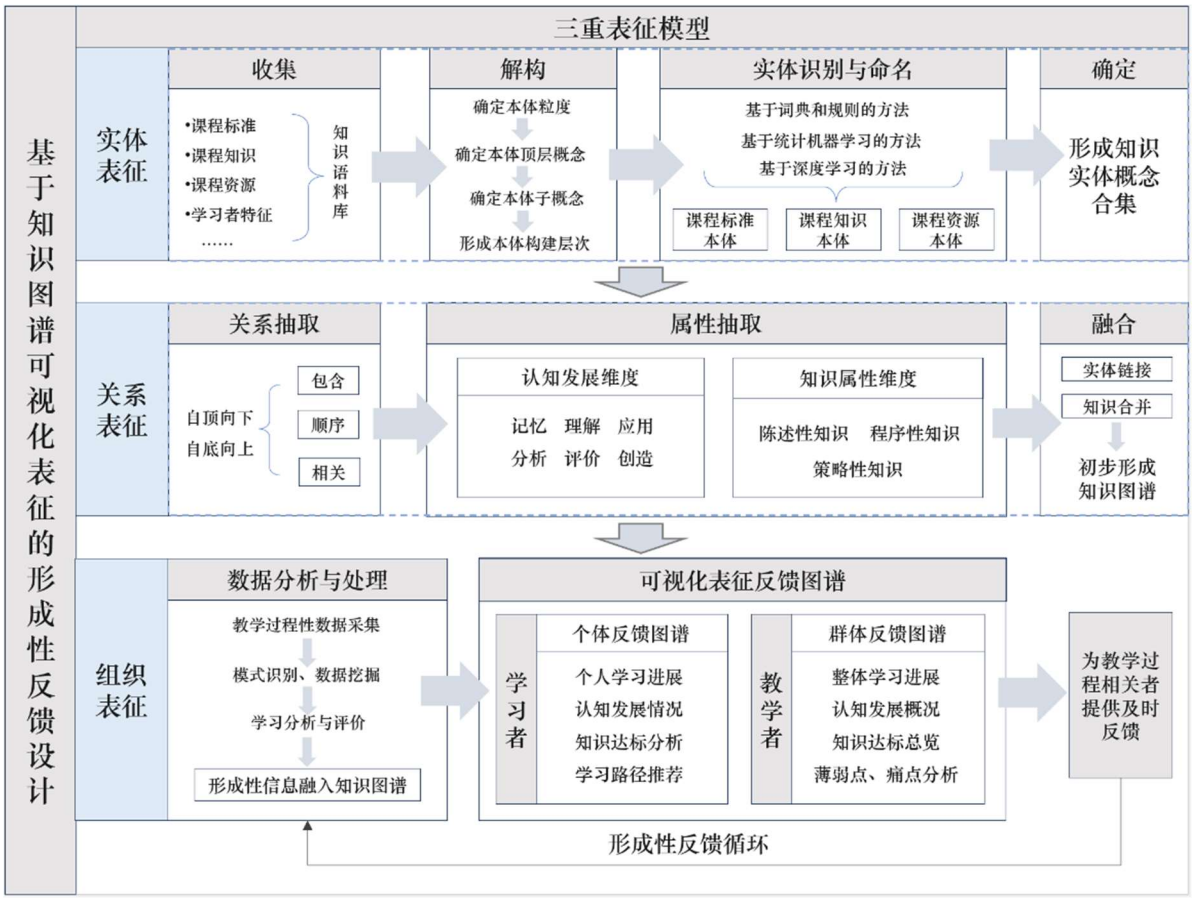


图 1 基于知识图谱可视化表征的形成性反馈设计模型

3.1. 实体表征：课程知识图谱实体挖掘

作为图谱最底层的模式结构，实体的挖掘是形成可视化表征的首要步骤。本研究结合教育知识图谱建构的一般性流程，将实体挖掘分为收集、解构、实体识别与命名三个主要环节。学习过程所涉及的内容繁多，包含了不同层次的知识内容以及不同形式的教学媒体，为从文本、图像、音频、视频中提取出统一性质的实体信息，首先需要将教学所涵盖的所有信息收集并合成知识语料库，例如学科课程标准、知识体系、教学资源等等。接着，根据教学实际需求、学科性质等确定本体划分的粒度，遵循“从整体到部分、从抽象到具体”的设计原则确定本体的顶层概念与建构层次，对知识语料进一步解构。一种教育领域常见的知识本体构建层次如图 2 所示，包括课程标准本体、课程知识本体、课程资源本体三个子本体。

在完成内容解构的基础上，借助现有的实体识别与命名技术如基于词典和规则的方法、基于统计机器学习的方法和基于深度学习的方法，获得精确度高、独立性强的知识实体并形成概念合集，用于后续知识图谱的关系组织与建构。

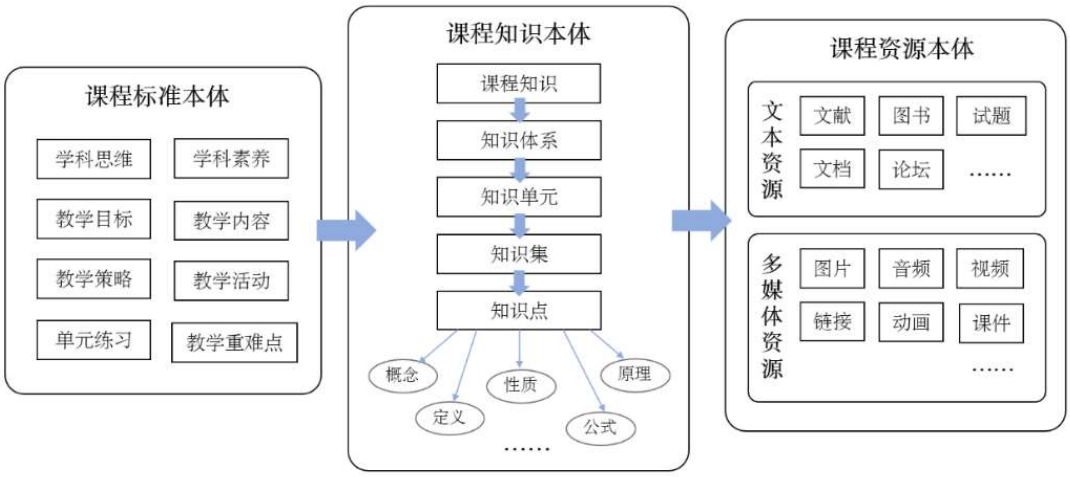


图 2 一种知识本体构建层次

3.2. 关系表征：知识映射关系网络构建

文本语料经过实体抽取所得到的是一系列离散的命名实体，为了构成网状知识结构，还需要从相关语料中提取出实体之间的关联关系，这一步骤称为关系抽取。本研究对本体的关系抽取主要包括知识点间关系抽取、知识点属性归类两个环节。知识点间的关系定义是关系抽取的前提条件。常见关系有包含、顺序、相关等，本研究采用自顶向下和自底向上相结合的构建方式，将所识别的知识实体与关系相融合，表达出“实体-关系-实体”的语义关系组。Nelson 等研究者认为，反馈是一种为了让学习者达到设定的目标，而告知学习者的表现情况，以此来激励和强化学习者的学习策略。因此，反馈的信息应与预期目标之间有所关联。本研究依据知识目标分类的规则来表征实体的关系与属性，以此反馈已有知识水平和学习内容之间存在的认知差异，主要有两种属性抽取规则，一是结合布鲁姆对认知领域的目标分类，对知识点进行认知维度的划分，从低到高分别对应记忆、理解、应用、分析、评价和创造六个层级；二是基于现代心理学对知识内容本身的分类，将其分为递进性回答“是什么-怎么做-如何学习”的知识，即陈述性知识、程序性知识和策略性知识。在此基础上完成对实体关系的抽取，并结合实体链接、知识合并技术生成一个相对完整的知识图谱，使得非结构化的学习内容通过结构化、知识化的加工，成为有语义关联关系的知识对象，为学习者提供可重组、可融合、可创生的知识体系。

3.3. 组织表征：形成性反馈设计与呈现

知识可视化表征的实质是将被称为认知知识“制品”的个体知识以“图解手段”表示出来，形成直接作用于人类感官的显性物理知识的外在表现形式，从而促进知识的传播和创新（赵国庆，2009）。基于此，本研究在组织表征阶段，首先需要采集和处理来自教学过程的各类数据，通过模式识别、数据挖掘、学习分析等将学习者阶段性、细颗粒度的学习表现数据转化成形成性信息，并融入到先前构建的知识图谱之中。由于教学过程同时涉及教师与学习者的共同参与，反馈信息的呈现也应顾及双方（de Kleijn，2021）。因此，在组织表征的过程中，知识图谱作为可视化表征的反馈“制品”，主要分为两种类型：面向学习者的个体反馈图谱以及面向教学者的群体反馈图谱。

在个体反馈图谱中，学习者可以获得个人学习进展、认知发展情况、知识达标分析以及学习路径推荐的反馈信息，在智能学习平台或系统中，可以作为四类型视图标签集中在同一知识图谱中，如图 3 所示。

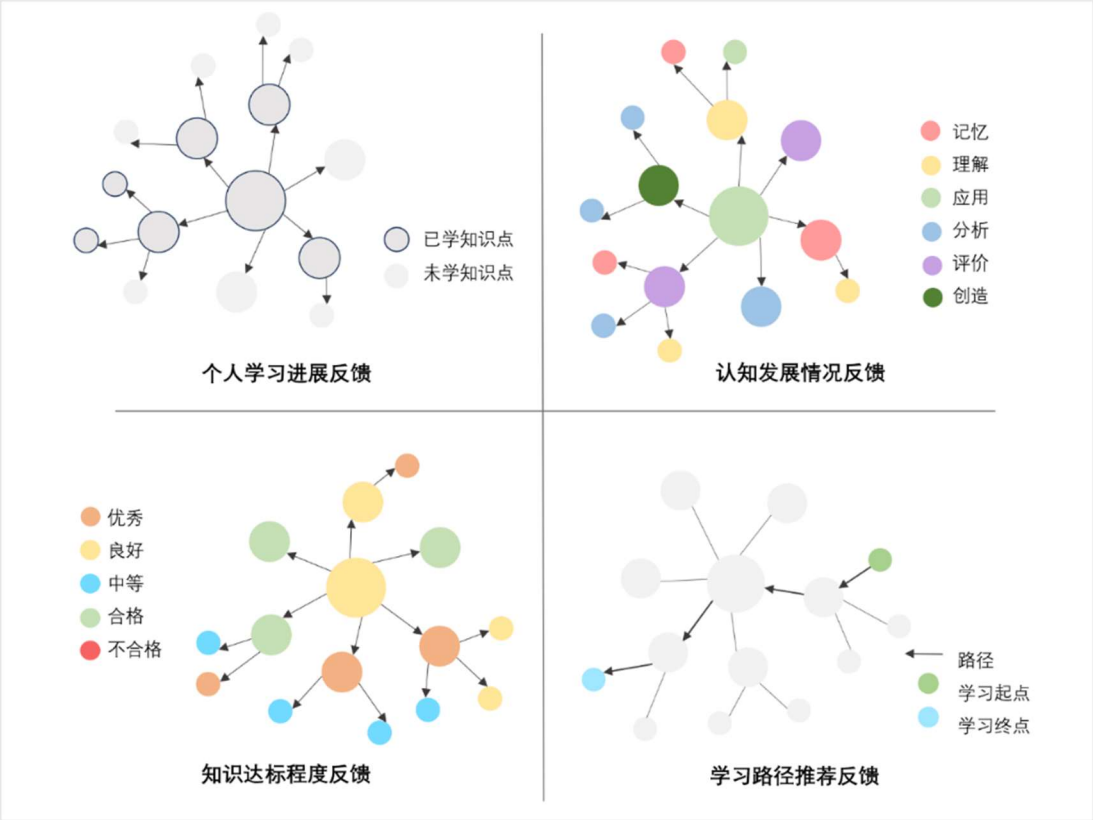


图 3 个体反馈图谱四类型视图

（1）个人学习进展反馈。每位个体学习者在课程开始前都有属于自己的知识图谱，映射了整体课程的知识体系。随着课程学习的开展，学习者的过程性进展信息被融合到图谱之中，并借助可视化的视觉表达如节点颜色、形状、样式等进行明显区分，在同一图谱中分层次地表示已学与未学知识点，类似于地图中的“已打卡点”，使学习者对自身学习进展以及当前表现和预期目标之间的差距有一个清晰的认识和理解，以便随时调整学习步调，明确学习方向，激活学习动机，灵活有效地完成学习目标。

（2）认知发展情况反馈。随着新课标改革的不断推进，课程标准与教学目标在“以人为本”的学生观指导下设置得更为客观具体。每一课程中的不同章节、每一章节中的不同知识点往往涉及到不同的认知发展维度，对应了不同程度的掌握要求。在传统课堂中，课程教学目标常以文本形式简单呈现，很难引起学习者的感知与重视。在知识图谱的可视化表征支持下，课程知识粒度与目标发展属性维度能够很好地结合，学习者可以直观了解到每一知识点所对应的认知维度与层次，有助于高效筛选与统计，能够以此更好地制定目标与计划。

（3）知识达标程度反馈。教育教学中常采用测试、作业、考试等对学习者阶段性学习成果进行检验与评价。为检测学习者的综合素养，教学者在设置考核题目时，通常尽可能地涉及到包含所有认知维度的知识点，这也就是说，同一份考核试题中的知识点很可能碎片化地分散在教材或课件之中，在没有明确指引的情况下，学习者很难精准定位到自己的薄弱点并进行及时的纠错与复习。为解决这一困境，本研究基于知识图谱将所有题目与知识体系关联起来，一方面利用知识点粒度表征学习者的具体达标情况，如按照课程考核五级制分为优秀、良好、中等、合格、不合格五类指标，分别对应不同的视觉样式；另一方面则是借助图谱提供了知识溯源的导航，便于学习者快速定位到知识点所在单元，获取相关的补充性学习材料，高效进行查漏补缺。

（4）学习路径推荐反馈。由于遗传、环境、身心发展规律等等，每个学生在发展速度、

认知水平、认知方式及个人优势等领域都存在或多或少的差异，为此反馈应遵循因材施教的原则，根据学生的个体差异提供针对性的反馈信息，并且能够切实帮助学生改善自身问题，包括指导性信息以及提升建议等等。在综合前三项反馈信息的基础上，追踪学习者真实学习感知状态与知识建构过程，结合当前表现与预期目标的差距、知识达标程度、薄弱点等重要信息，形成针对性的个性化学习路径，并在原有图谱中进行表征与反馈，基于学习者个性化学习需求进行推荐，使自适应学习、精准教学具备了实现的可能。

群体反馈图谱的设计以及呈现与个体类似，不同点在于其主要是面向教学者呈现的学习群体宏观学情反馈，汇集个体学习者过程性数据，将学习进展、认知发展、知识达标等情况提升到以班级为单位的宏观统计角度，分析整体学生的普遍薄弱点、痛点，呈现平均及统计指标，如完成率、掌握率等。教学者在图谱的可视化表征下，能够快速了解班级的整体学习情况，基于知识脉络体系精准定位学生学习短板以及薄弱知识点，同时及时调整教学目标、内容、过程以及方法，有针对性地干预学习者的学习进程，从而优化学习路径，促进学、教、评的有机结合。而学习者与教学者在接收反馈信息的基础上也会产生新的过程性数据，并不断融入到已有的知识图谱中，持续性生成新的反馈信息，建立起形成性反馈循环，促进了生成性的知识建构与教学优化，有助于学生及教师两方面参与者知识与技能的提升。

4. 结论与展望

随着信息技术的发展，教育教学对形成性反馈提出了更高的要求，反馈从机械化、单一性向个性化、适应性发展。在数字化阅读的时代背景下，具有强可视化表征优势的知识图谱与形成性反馈的结合具有了可行性与适切性。本研究提出了基于知识图谱可视化表征的形成性反馈设计模型，以实体、关系、组织三重表征模型分别对应设计形成性反馈的内容信息、功能意义、呈现形式，并在组织表征阶段具体呈现了面向两类学习主体的反馈图谱——个体反馈图谱和群体反馈图谱，分别介绍关于学习进展、认知发展、知识达标、学习路径等方面反馈信息的表征方式，以期基于学生过程性数据呈现有针对性的实时反馈，精准干预学生学习进程，并基于图谱智能推荐个性化学习路径，进而最大化发挥反馈优化学习过程、提高学习效果的价值。

在未来，本研究将基于该模型开展真实教学情境中的实践应用，以验证其对于提高学习效果、改进教学过程的科学性和有效性，并结合具体的教与学过程以及用户反馈意见，进一步优化模型的设计及反馈图谱的构建与应用，从而为教学过程相关者提供更优质更高效的形成性反馈，使反馈真正有效地促进学习。

参考文献

伍绍杨 & 彭正梅. (2021). 迈向更有效的反馈:哈蒂“可见的学习”的模式. 开放教育研究 (04), 27-40.

李振, 周东岱 & 王勇. (2019). “人工智能+”视域下的教育知识图谱:内涵、技术框架与应用研究. 远程教育杂志 (04), 42-53.

李振 & 周东岱. (2022). 基于学科知识图谱的智能化认知诊断评估方法. 现代教育技术 (11), 118-126.

董艳. (2020). 学生反馈素养论纲:内涵、模型与发展. 开放教育研究 (05), 26-39.

戴静, 顾小清 & 江波. (2022). 殊途同归：认知诊断与知识追踪——两种主流学习者知识状态建模方法的比较. 现代教育技术 (04), 88-98.

王海英 & 吴爽. (2020). 形成性评价视域下反馈在中小学教学管理中的运用. 现代教育管理 (03), 103-109.

詹泽慧, 李通德, 邹萱萱 & 李梁. (2023). 智慧课堂教学交互事理图谱的构建与实证：融合知识与事理逻辑. 电化教育研究 (09), 77-85.

刘梦君, 尹加琪, 杨兵 & 周雨虹. (2022). 基于智能交互的形成性反馈研究. 现代教育技术 (06), 62-71.

沈红叶, 肖婉, 季一木 & 刘尚东. (2023). 教育知识图谱的类型、应用及挑战. 软件导刊 (10), 237-243.

余胜泉 & 李晓庆. (2019). 区域性教育大数据总体架构与应用模型. 中国电化教育 (01), 18-27.

黄昌勤, 朱宁, 黄琼浩 & 韩中美. (2019). 支持个性化学习的行为大数据可视化研究. 开放教育研究 (02), 53-64.

顾小清 & 权国龙. (2014). 以语义图示实现可视化知识表征与建模的研究综述. 电化教育研究 (05), 45-52.

John Hattie & Helen Timperley.(2007).The Power of Feedback.Review of Educational Research(1),81-112.

de Kleijn R A M. (2021) Supporting student and teacher feedback literacy: an instructional model for student feedback processes. Assessment & Evaluation in Higher Education,: 1-15.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. Instructional design for multimedia learning, 181195.

Rizun, M. (2019). Knowledge Graph Application in Education: a Literature Review. Acta Universitatis Lodziensis. Folia Oeconomica, 7–19. https://doi.org/10.18778/0208-6018.342.01

Shute Valerie J..(2007).FOCUS ON FORMATIVE FEEDBACK.ETS Research Report Series(1),i-47.

AI 绘画以图解文的中职诗词教学策略研究

Research on the Poetry Teaching Strategy in Secondary Vocational Schools with the Text-to-image Function of AI Painting

陈志雄
集美工业学校 基础教学系
243567513@qq.com

【摘要】 AI 绘画技术具有应用于诗词教学的适切性。AI 绘画以图解文的中职诗词教学策略是在直观教学原理、语图互文理论、认知图式理论、格式塔理论等理论指导下，将 AI 绘画技术与诗词教学深度融合的创新性教学策略，包含言图互仿、图嵌于言、图丰于言、言图互间四种策略。阶段性实验证明，该教学策略能调动学生学习参与积极性，实现学生与 AI 绘画工具高频互动，推动学生的诗词学习由浅层理解向深度领悟转变。
【关键词】 AI 绘画；诗词教学；语图互文理论；认知图式理论；格式塔理论

Abstract: *AI painting is the application of artificial intelligence technology in the field of art. AI painting technology has the function of interpreting texts with images in the field of language teaching and it has the appropriateness of being applied to poetry teaching. Chinese teachers in secondary vocational schools can give full play to the advantages of AI painting tools such as text analysis, text visualization, graphic interaction and so on to innovate the teaching strategies of poetry, stimulate students' interest and enthusiasm in learning poetry and improve their Chinese core literacy and artistic creativity. Under the guidance of intuitional instruction, language-image intertextuality theory, cognitive schema theory, Gestalt theory and other theories, the poetry teaching strategy with the text-to-image function of AI painting is an innovative teaching strategy that deeply integrates AI painting technology with poetry teaching. It includes four strategies: understanding the whole text with corresponding images, perceiving elements in the text with images, enriching the understanding of the text with images, comprehending the text through mutual complementation of text and image. The phased experiment proves that the poetry teaching strategy with the text-to-image function of AI painting can mobilize students' enthusiasm for learning participation, realize the high-frequency interaction between students and AI painting tools, and promote the transformation of students' poetry learning from shallow understanding to deep comprehension.*
Key words: AI painting,poetry teaching,language-image intertextuality theory,cognitive schema theory,Gestalt theory

1. 引言

AI 绘画是人工智能技术在艺术领域的应用，AI 绘画生成工具能分析使用者输入的提示文本，实现“以文生画”，创作不同风格的绘画（刘洋，2023）。2015 年，美国的罗比·巴拉特开始借助生成式对抗网络（GAN 算法）创作 AI 图像，迅速掀起“AI+艺术”的浪潮。2018 年，AI 绘画《贝拉米家族的埃德蒙德·贝拉米》在佳士得高价成交。2022 年，AI 绘画技术突飞

猛进，百度 AI 绘画平台“文心一格”续画陆小曼未完成的山水画《夏日山居图》引发广泛关注（王嘉，2022）。AI 绘画工具可以借助诗词的关键字词作为提示文本，快速生成 AI 画作，助力学生深度学习诗词。2022 年 11 月，教育部发布《教师数字素养标准》指出“教师适当利用数字技术获取、加工、使用、管理和评价数字信息和资源，发现、分析和解决教育教学问题，优化、创新和变革教育教学活动所具有的意识、能力和责任（中华人民共和国教育部，2022）。”中职语文教师可以发挥 AI 绘画工具以图解文的教学优势，创新中职诗词教学策略。

2. 中职诗词教学现状分析

2.1. 中职诗词教学实施的要求

根据《中职语文课程标准（2020 年版）》（以下简称《课标》），中职语文课程基础模块的专题 4 为古诗文选读，该专题要求学生体会诗词的精神内涵、审美追求与文化价值，教师引导学生在阅读作品的过程中，学习运用评点方法，记录自己的感受和见解，并就作品涉及的文化现象与同学展开交流和讨论，联系生活经验，表达自己的看法。此外，《课标》要求教师提高信息素养，适应以互联网、大数据、人工智能为代表的新一代信息技术的发展趋势（中华人民共和国教育部，2020）。

2.2. 中职学生诗词学习面临的挑战

2.2.1. 对诗词大意的理解不清晰

中职学生对古代汉语词汇、句式等语言基础知识比较生疏，且文本解读时想象能力有限，对诗词内容的理解存在困难，无法清晰地解读出诗词中描述的内容，对诗词所表述的大意理解不充分，甚至由于诗词中关键性词汇的理解错误，丧失对诗词的学习兴趣。

2.2.2. 对诗词意象的分析不深入

中职学生的诗词意象理解与表达技巧比较匮乏，对诗词常见意象的积累较少，容易停留在对诗词意象表层意思的理解，对诗词意象的分析不深入，忽视物象与作者创作意图之间的密切联系。

2.2.3. 对诗词意境的领会不准确

中职学生由于积累的中华传统文化背景知识有限，对诗词所创造出的环境比较陌生，难以凭借对创作者人生阅历的认知，准确领会诗词的意境。

2.2.4. 对诗词叙事的时空关联认识不全面

中职学生具象思维较强，能结合具体生活体验，想象诗词所描绘的空间性画面，生成感性认识。但是，中职学生逻辑思维尚不成熟，难以对诗词内容的时间性变化做出推断，无法自主建立诗词叙事的时间与空间关联。

2.3. 现有中职诗词图画教学策略不足

2.3.1. 使用静态图画教学，图画工具使用频率较低

现有中职诗词教学主要使用课本插图、电子图片等素材均属于静态图画，往往在单次使用后就失去其教学价值，未能与学生产生双向交互，教学使用频率较低。

2.3.2. 运用图画解说诗词，图画工具使用方式比较单一

图画除了作为教师辅助解析文本语言的工具之外，还能作为学生学习成果的审美性、思想性表达工具，展示与检验学生对诗词的学习成效。教师单独运用图画解说诗词，无法展示学生诗词学习成果，忽视了图画资源在学生“学”的呈现。

2.3.3. 教学活动浮于表层，学习深度不足

教师运用图画开展的诗词教学活动，较为注重理解与分析诗词的文本大意、创作意象，

缺少对诗词文化背景知识、时空叙事等的深度解读，使学生难以形成可迁移的诗词鉴赏能力。

3. AI 绘画以图解文助力诗词教学

3.1. AI 绘画技术的应用现状

随着人工智能技术迅速发展，AI 绘画技术在数字艺术、虚拟形象设计、数字媒体等领域得到广泛应用，帮助艺术家实现数字化创作和艺术表达。AI 绘画技术可以帮助设计师和开发者制作出更加逼真、鲜活的虚拟形象。AI 绘画技术可以为数字媒体提供更加精彩、有趣、丰富的内容。

3.2. AI 绘画技术在语言教学领域具有以图解文功能

AI 绘画技术已经具有文本分析、文本可视化、图文交互等功能优势。在语言教学领域，教师可以赋予 AI 绘画技术以图解文功能。语言文本经过 AI 绘画平台的文本分析、可视化处理，被智能化绘制成图像，人们可以借助 AI 绘画平台实现图像、文本之间的双向互动（见图 1）。

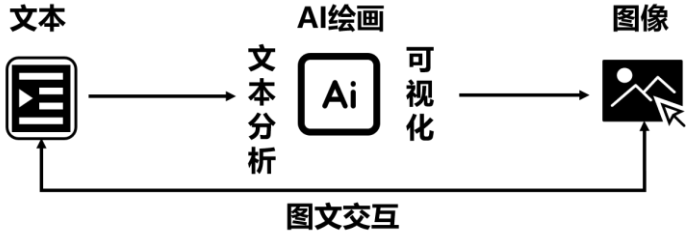


图 1 AI 绘画技术在语言教学领域的以图解文功能

3.2.1. 文本分析功能

AI 绘画技术的文本分析功能是使用人工智能技术分析文本数据，提取信息，帮助绘画软件自动生成画作或进行相关的处理。在语言教学中，文本分析功能可以识别语言文本的关键词、情感和主题等，根据这些特征信息生成相对应的绘画元素，如颜色、线条、形态等。

3.2.2. 文本可视化功能

AI 绘画技术的文本可视化功能是利用人工智能技术将语言文本中的特征信息转化为可视化的图像或动态图像。在语言教学中，文本可视化功能使原本抽象、晦涩的语言内容以图像形式展现出来，调动学习者直觉感官，辅助学习语言文本。

3.2.3. 图文交互功能

AI 绘画技术的图文交互功能是利用人工智能技术与绘画软件进行信息交互，实现对绘画作品修改、调整和创作等，获取更加灵活与个性化的绘画体验。在语言教学中，教师利用AI 绘画技术分析学习者的绘画行为，并在学习者绘制的过程中提供实时反馈和建议，帮助学习者更加精准、全面地掌握语言学习内容，方便学习者自主、快速和准确地完成绘画作品。

3.3. AI 绘画技术应用于诗词教学的适切性

1）AI 绘画技术可以帮助学生更直观地理解、呈现诗词文本内容。

2）AI 绘画工具可以吸引学生参与语言创造性活动。学

3）AI 绘画工具可以生成学生的语言与艺术审美能力，培养学生的实践操作能力。

4. AI 绘画以图解文的中职诗词教学的理据及策略

4.1. 理论基础

4.1.1. 直观教学原理

夸美纽斯的《大教学论》提出“一切知识都是从感官开始的”（王天一，1985），强调直观教学对于学习者认识抽象性知识的重要性，教学应充分调动学习者的感官知觉，立足学习者已有经验，使学习者通过各种感知形成感性认识与直接经验。AI 绘画技术能将使用者的语言文本进行智能化分析，将其可视化为艺术绘画，对于语言的直观化教学具有辅助作用。学生借助 AI 绘画技术，将自己对诗词的文本解读生成绘画，再观察 AI 绘画反思对诗词语言理解，继续修正和完善对诗词的解读。

4.1.2. 语图互文理论

赵宪章提出“语言和图像是人类社会有史以来最基本的两种表意符号”（赵宪章，2012），虽然两者存在不对称性，但是语言与图像可以互仿，图像对语言的模仿属于顺势模仿，语言对图像的模仿属于逆势模仿（赵宪章，2011）。余俊宏等提出“语图互文意即语言和图像之间的互文性关系”（余俊宏等，2022），提出语图互文视角下语言与图像的四种关系：语图互仿、图溢言外、言大于图、语图互间（见图 2）（余俊宏等，2022），指出了语图互文理论对于语文教育的指导意义。



（a）语图互仿　（b）图溢言外　（c）言大于图　（d）语图互间

图 2 语图互文视角下语言与图像的关系

语图互文理论比较清晰地阐释了语言与图像的关系。其中，语图互仿是图像与语言的模仿，两者表达的范围基本相当；图溢言外是图像表达范围大于语言所指；言大于图则是语言表达范围大于图像；语图互间是语言与图像在表达上的相互补充，达到时间与空间合一。

中职诗词教学离不开对语言与画面的阐释，教师可以基于语图互文视角，结合中职学生具象思维较强的学情，设计教学策略。

4.1.3. 认知图式理论

1932 年，心理学家巴特利特（Bartlett）正式把图式理论引进认知领域。认知心理学家鲁墨哈特（Rumelhart）认为图式是一种可以交互作用的认知知识结构，它是认知能力基本要素。哈德逊（Hudson G.）运用图式理论研究外语阅读。马笑霞提出将图式理论应用到语言教学中，辅助学生建立认知结构体系（马笑霞，1997）。依托认知图式理论，建构诗词的常见意象图式，贴近学生认知思维发展特点，契合《课标》要求，有助于建立学生良好的诗词认知结构。

4.1.4. 格式塔理论

格式塔理论提出感觉是“先感知到整体，而后才注意到构成整体的各个部分”（库尔特·考夫卡，1997），整体的内容大于各部分之和。格式塔理论的系统论强调整体功能大于各部分功能之和，完形压强论强调人们认识过程会自主对事物的缺陷进行补充“完形”，使其成为完满的整体。教师利用系统论引导学生建构对诗词的整体认知，利用完形压强论指导学生赏析诗词的不确定性与模糊性。

4.2. 教学策略

4.2.1. 言图互仿：绘整体诗意

在言图互仿教学策略中（见图 3），教师引导学生浏览诗词，寻找诗词描写的所有物象，借助思维导图工具，分类梳理物象，建立比较系统的画面元素，撰写诗词整体解读的语言文

本。学生使用 AI 绘画工具以人工智能技术分析语言文本，智能绘制诗词整体画面。学生在教师启发下比较 AI 绘画生成的诗词整体画面与诗词的整体解读文本，借助图像强化对诗词语言整体内容的认识，继续修改诗词解读语言，调整、补充和完善 AI 绘画作品。

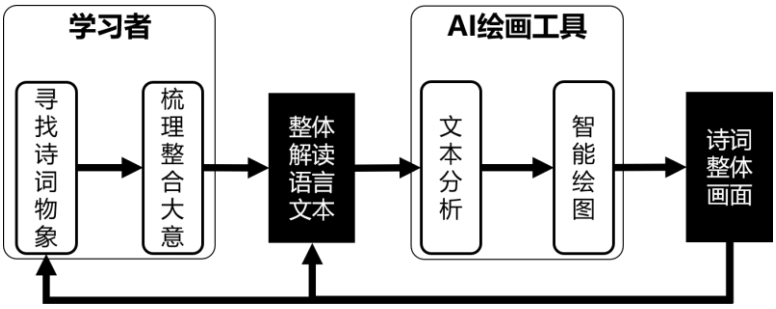


图 3 言图互仿的教学策略示意图

4.2.2. 图嵌于言：描诗歌意象

图嵌于言的教学策略中（见图 4），学生寻找诗词中具有代表性的意象，借助诗词意象图式分析意象，撰写意象解读语言文本。学生在 AI 绘画工具输入语言文本进行智能化文本分析，完善原有的诗词整体画面细节，使诗词意象嵌入画面，实现诗词构图的完形，使画面的情景描绘更加细致、生动。教师启发学生对比诗词意象画面、意象解读语言文本，强化对诗词意象的理解。学生再次完成对诗词内容的图示分析，重新描绘原有画作中的意象元素。

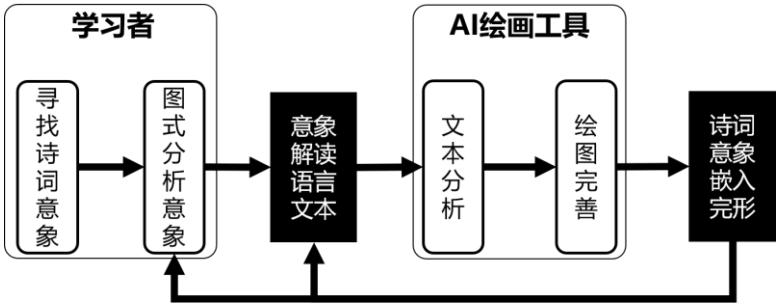


图 4 图嵌于言的教学策略示意图

4.2.3. 图丰于言：琢创作意境

图丰于言的教学策略（见图5）需要学生在学习诗词过程中保持强烈的求知兴趣，使用课内外诗词学习资源充实诗词的背景知识，学习诗词的创作背景，思考、领悟诗词创作意境，撰写意境解读的语言文本。学生借助 AI 绘画工具，对诗词创作意境进行完形，使诗词画面具有图像之外的深层意境感。教师指导学生比较诗词创作意境的画面、意境解读的语言文本，强化对诗词意境的感悟，不断雕琢 AI 画作的意境表达形式。

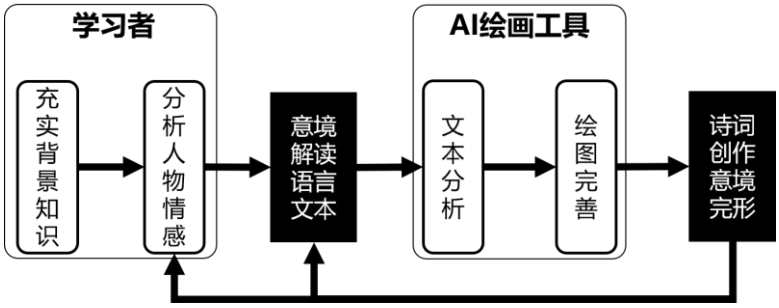


图 5 图丰于言的教学策略示意图

4.2.4. 言图互间：融时空叙事

言图互间的教学策略中（见图 6），学生尝试以叙事视角赏析诗词，将诗词作为事件进行文本解读，在教师引导与提示下寻找诗词中具有叙事特征的语言，分析诗词的叙事线索，集

体交流研讨并梳理诗词的叙事脉络，撰写诗词时间叙事的语言文本，使用 AI 绘画工具创作时空融合的叙事画面。教师指导学生再次研讨诗词叙事线索、叙事脉络，增强诗词绘画的时空叙事内容，使画作的叙事层次更加丰富。

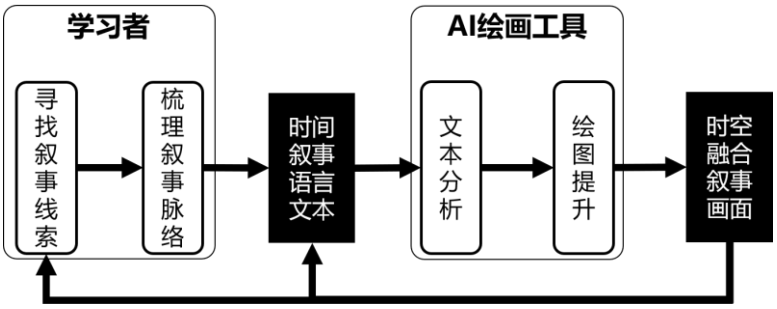


图 6 言图互间的教学策略示意图

4.3. 应用列举

4.3.1. 言图互仿教学策略应用列举

言图互仿教学策略应用以高教版《语文（基础模块)》上册的诗词《无衣》教学为例进行阐释。诗歌《无衣》出自《诗经·秦风》，是一首秦地将士慷慨激昂的战歌，主要描述了秦地军队在物资匮乏的艰难环境下，团结一心、奔赴前线的雄壮场景，具有突出的写实风格。教师运用言图互仿教学策略教授诗歌《无衣》时，设计了 2 个学习任务。任务 1：浏览诗歌，寻找诗歌描写的关于战争的物象，借助思维导图梳理诗歌整体图画内容。任务 2：借助 AI 绘画工具绘制诗歌《无衣》的全景画面，观察并修改 AI 画作，使画面内容更接近诗歌内容。

4.3.2. 图嵌于言教学策略应用列举

图嵌于言教学策略应用以高教版《语文（基础模块)》上册的诗词《采薇》教学为例进行阐释。诗歌《采薇》选自《诗经·小雅》，描写了戍边战士生活艰苦，日夜思归的场景。诗歌以“薇菜”作为重要意象，透过薇菜不同季节的生长变化来暗示战士戍边时光流逝，进一步点明诗歌中戍边战士因常年离家在外，思念家人却不得归家的主题。教师运用图嵌于言教学策略教授诗歌《采薇》时，设计三个学习任务。任务 1：阅读诗歌，寻找诗歌中描写“薇菜”的语句，初步认识诗歌里“薇菜”的特征。任务 2：回顾所学刻画游子的诗歌，思考“薇菜”激发了戍边战士怎样的思绪。任务 3：借助 AI 绘画工具在绘制诗歌《采薇》里戍边战士采摘“薇菜”的画面，对照“薇菜”意象分析内容，修改 AI 画作。

4.3.3. 图丰于言教学策略应用列举

图嵌于言教学策略应用以高教版《语文（基础模块)》上册的诗词《沁园春·长沙》教学为例进行阐释。《沁园春·长沙》是由毛泽东创作的爱国词作，作品的意境高远、气势磅礴，表达了青年毛泽东以天下为己任豪情壮志，以及对民族和国家前途的乐观主义精神。教师运用图丰于言教学策略指导学生解读《沁园春·长沙》中的毛泽东形象，设计了 3 个任务。任务 1：自主学习《沁园春·长沙》相关创作背景知识及作者相关文化背景知识，思考毛泽东创作这首词时的心境。任务 2：分析词作表达的人物情感，感悟主人公面对湘江秋景，追忆往昔“同学少年”报效国家的奋斗激情。任务 3：使用 AI 绘画工具绘制出体现诗词创作意境的人物形象，参照对人物情感的理解和感受修改 AI 画作，使 AI 画作实现情景交融。

4.3.4. 言图互间教学策略应用列举

图嵌于言教学策略应用以高教版《语文（基础模块)》上册的诗词《国殇》教学为例进行阐释。诗歌《国殇》是一首追悼为国牺牲将士的挽歌。这首诗歌描绘了一场楚国将士抵抗秦军的悲壮战役，经过激烈厮杀，全体楚国将士为国捐躯，既有基于时间的战斗过程叙事，也有依于空间的战斗场面叙述。教师运用言图互间教学策略教授《国殇》时，设计了 3 个任务。

任务 1：阅读诗歌，寻找诗歌中描述战役经过的关键语句，确定诗歌的叙事线索。任务 2：结合诗歌叙事线索、诗歌创作的文化背景知识，借助思维导图梳理《国殇》的叙事脉络。任务 3：使用 AI 绘画工具按照诗歌时间叙事语言文本绘制出融合时空叙事的诗歌画面，继续寻找诗歌叙事线索，再次润色，以达成时空叙事巧妙融合。

5. 评价和反思

5.1. 效果评价

5.1.1. 评价维度

为全面客观地评价 AI 绘画以图解文的中职诗词教学策略应用成效，研究团队按照教师和学生两个主体设计教学评价内容，分学习行为、学习成果两个基本评价维度。学习行为测评主要关注学生的诗词学习活动参与频率。学习成果测评依照《课标》中“学业质量”的“学业质量水平1”描述（中华人民共和国教育部，2020）设计为：1）整体感知文本思想内容；2）正确把握思想感情；3）健康的审美意识；4）整体感受作品中的形象；5）学习中华优秀传统文化的兴趣；6）梳理文化常识、文化现象，增加文化积累。

5.1.2. 评价结果

为验证 AI 绘画以图解文的中职诗词教学策略实施成效，研究团队以同一授课教师、统一教学进度的三个平行班级（150 人）作为研究对象，完成一个学期 6 次诗词课程的教学实验，通过在线学习平台获取实验统计数据（见图 7、图 8），分析学习行为、学习结果测评数据，得到以下结论。

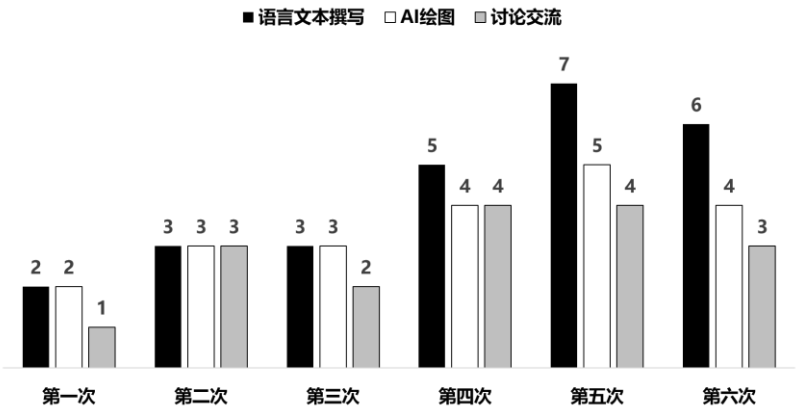


图 7 AI 绘画以图解文的中职诗词学习行为测评

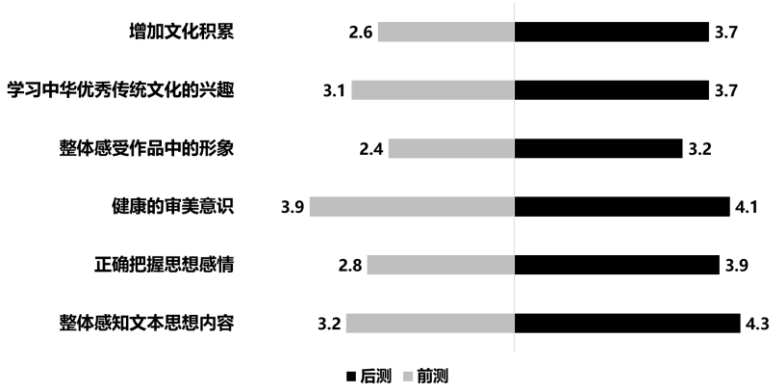


图 8 AI 绘画以图解文的中职诗词学习成果测评

1）AI 绘画以图解文的中职诗词教学策略调动了学生参与积极性。如图 7，学生诗词解读语言文本撰写、AI 绘图、讨论交流的参与频率总体呈上升的趋势，学生的学习参与度显著提

高。

2）AI 绘画与语言文本撰写两种行为呈现同步推进趋势，学生与 AI 绘画工具互动频繁。如图 7，学生 AI 绘画与语言文本撰写参与频率接近，两种学习行为存在较强的正相关性，表明学生与 AI 绘画工具之间存在高频率的互动行为。

3）AI 绘画以图解文的中职诗词教学策略实施后，学生对诗词的学习经历了从浅层理解到深度领悟的转变，诗词鉴赏能力得以增强，达成中职语文课程学业质量培养目标。如图 8，学生在“整体感知文本思想内容”“正确把握思想感情”“整体感受作品中的形象”“增加文化积累”等能力明显提升。

5.2. 设计反思

AI 绘画以图解文的中职诗词教学策略在具体实施过程中也面临着许多新问题、新挑战。

1）即使 AI 绘画技术可以为中职学生的诗词具象化学习提供帮助，在实际应用中也需要注意人工智能技术存在一定的限制性。例如，AI 绘画在对学生提供的语言文本进行分析时，可能无法充分理解诗人和学生的文化背景、阅读经验、创作目的等因素，无法完全代替老师和学生之间的交流和体验。

2）在诗歌教学中应当有意识地将 AI 绘画技术与传统教学相结合，让学生充分感受到技术带来的益处的同时，也应该高度重视学习者作为人的创造主体性。教师需要关注学生对诗词的个性化学习和表达，培养学生学习语文课程的创新和思考能力。

6. 展望

AI 绘画技术在教育领域的应用将日益广泛，AI 绘画以图解文的中职诗词教学策略作为一种融合了人工智能、艺术、文化和教育的创新教学模式，在未来的发展中将有广泛的应用与发展前景。该教学策略可以帮助学生更加深入地理解诗词内容，提高其诗词鉴赏能力，AI 绘画技术为诗词教学增添了更多视觉表达和趣味性。此外，AI 绘画技术与诗词教学相结合的方式，为学生带来多元文化体验和艺术创作灵感。AI 绘画技术与使用者之间如何有机协同、密切配合，如何在 AI 技术使用过程中有效地保护和发挥人的创造性等问题，亟待越来越多的教育研究者继续在实践中探索与破解。

参考文献

刘洋。(2023-01-04)。AI 绘画与古代诗词“花”之意象表现。《中国书画报》,003。

王嘉。(2022-04-21)。AI 绘画是人类情感的表达吗？。《成都日报》,008。

中华人民共和国教育部。(2022)。教师数字素养:JY/T 0646-2022。北京:中华人民共和国教育部。

中华人民共和国教育部。(2020)。中等职业学校语文课程标准。北京:高等教育出版社。

王天一。(1985)。外国教育史(下册)。北京：北京师范大学出版社。

赵宪章。(2012)。语图符号的实指和虚指——文学与图像关系新论。《文学评论》, (02),88-98。

赵宪章。(2011)。语图互仿的顺势与逆势——文学与图像关系新论。《中国社会科学》, (03), 170-184。

余俊宏&李丹丹。(2022)。语图互文视角下图像的教育价值论析——以统编小学语文教材插图为例。《南昌师范学院学报》, (03),128-130。

马笑霞。(1997)。图式理论与语文阅读教学。《教育研究》, (05), 66-70。

库尔特·考夫卡,黎炜译。(1997)。格式塔心理学原理(上)。杭州：浙江教育出版社。

基于知识建构原则的“人工智能+教育”实践探索——以一年级跨学科课程《可持续发展目标 11：可持续城市和社区》为例

Exploring the Principle-Based Knowledge-Building Practices of Artificial Intelligence Education - A Case Study from Grade One Interdisciplinary Course Sustainable Development Goal 11: Sustainable Cities and Communities"

林洪全，冯雪琦*，李燕莹，胡秀，李娜，吴珍科，周辉，李慧哲

南方科技大学附属实验学校
* fengxq@sustech.edu.cn

【摘要】人工智能时代随着生成式人工智能技术的演进，推进了教学模式从“师-生”二元结构转向“师-机-生”三元结构。而大语言模型本身具备丰富知识库的特征，也将让以“集体知识”为重点导向的知识建构理论产生新的理论原则与实践探究。本研究在知识建构原则的基础上，结合人工智能教育的实践需要，提出“动态迭代，反映心声”、“智能创新，引领未来”、“协作共建，智慧互鉴”三条原则，以促进人工智能与知识建构理论的双向赋能，并以“可持续城市和社区”为主题进行基于知识建构原则的人工智能教育实践探索。

【关键词】 知识建构理论；跨学科课程；可持续发展目标；人工智能；

Abstract: In the era of artificial intelligence, with the evolution of, the changes in the field of education have been pushed forward, and the teaching mode has been shifted from the "teacher-student" to the "teacher-machine-student," while the richer knowledge base brought by the big language model will definitely make the "community knowledge" oriented Knowledge Building theory generate new theoretical principles and practical investigations. Based on the knowledge building principles and the practical needs of artificial intelligence , this study proposes three principles: "Dynamic Iteration, Reflecting the Voice of the Heart", "Intelligent Innovation, Leading the Future", and "Collaboration and Mutual Learning". We propose three principles to promote the two-way empowerment of artificial intelligence and the theory of knowledge building, and carry out a practical exploration based on the principles of knowledge building with the theme of "sustainable cities and communities".

Keywords: Knowledge Building, Interdisciplinary curriculum, Sustainable Development Goal, AI

1. 引言

《2017 年中国人工智能行业白皮书》指出人工智能技术是构建信息流的关键技术（陈凯泉，2018），2024 年教育部部长怀进鹏在世界数字教育大会上提出要促进智能技术与教育教学（AI for education）的融合，可采取师生自由创造，学校自主建设等方式开发课程。而随着生成式人工智能（AIGC）带来的颠覆性应用，如 Chatgpt 和 Sora，推动教学模式从“师-生”二元结构转向“师-机-生”三元结构，推 进教学内容从人工生产转向智能生产（杨宗凯，2023）。知识建构理论强调每个人都是知识的贡献者且推动着观点的升级。但知识建构理论下促进集体知识的发展是核心原则，知识建构作为一种教育方法，从根本上说是一种观点建构的挑战；它是学生集体承担认知责任，共同推动集体知识发展的过程（Bereiter,2014）。如何发展知识建构原则促进人工智能教育的实践应用，是本

研究想要解决的问题。因此本研究以联合国制定的全球发展目标 11:可持续城市和社区为主题，以校园文化这一学生熟悉的社区为实践落脚点，探讨如何基于知识建构理论，提出促进人工智能教育的设计原则，从而推动“校园文化”这一集体知识的持续发展。

2. 概念框架

2.1 基于原则的知识建构教学法

知识建构理论（Knowledge Building）是一种关于知识创造和社区共建的教育理论，学习者对共同探究的领域发表观点并质疑延伸，从而得到观点的升级（Scardamalia，2005）。其有 12 条基本原则作为教学指南、技术设计规范和评估现有做法的依据，包括“真实的问题和想法”“多样化的观点”“观点的升华”“学生是积极的认知者”“知识建构对话”等（Scardamalia，2010），这些原则推动着让课程培养学生成为善于解决问题的工作者，也推动着学校成为知识建构的共同体。在传统的“师-生”二元结构下很难构成 KB 中的观点改进行为的追踪与分析（胡金艳，2021）。因此，在人工智能时代背景下，如何更有效应用知识建构原则，以实现师-生-机共同促进集体知识的发展，成为现阶段的难题

2.2 人工智能教育实践

人工智能教育包括利用人工智能赋能的教育和以人工智能为学习内容的教育（彭绍东，2021）。赋能任何参与教育环节的对象，以形成教育的智能生态。但人工智能与教育分属不同的领域，这样的跨界融合，需要教育理念的指导。目前人工智能在应用到教育实践方面，仍缺少科学完善的应用指南、创新实施方法和行动措施（卢宇等，2023），在课程上，应着重探讨人工智能与教育主体的互动关系与互动模式（崔金贵等，2023），以上困难都将削弱了人工智能与教育融合的适配度。因此，亟需应用教育理念来指导“师-机-生”三元结构的教学模式。

3. 原则设计

3.1 原则 1：动态迭代，反映心声

知识建构主义理论以观点为中心,不强调任务和活动为中心;坚持以 12 条知识建构原则为教学设计的基础（张义兵，2012）。这种建构不是一蹴而就的，而是在争论中，合作沟通中建立起来的（Improvable Ideas），这是一个民主的知识社区。老师需要在这个过程中根据学生的最近发展区来制造认知冲突，借助支架的力量来帮助学生进行观点的改进和升华。这也是融合 KB 理论和 AI 技术来推进 KB 教学模式的重要原则一：动态迭代，反映心声。我们的课程是在学生建构的知识基础上进行迭代，根据学生构成的知识社区的个性化来推动课程的进展。

3.2 原则 2：智能创新 引领未来

遵循“协同的认知责任”和“无处不在的知识建构”的基本原则，本次跨学科主题课程设计的原则二是智能创新，引领未来。跨学科要求学生掌握不同领域的基本知识，因此观点建构和对话就发生在不限定的场合和科目。学生在人工智能技术的帮助下，有更大的创造力想象力，借助 AI 的技术和学生的元对话来促进 big idea 的生成，这是遵行 12 条基本原则中的此外知识建构的课堂要求教师更加注重学生的过程性评价，关注学生每个观点的变化过程、小组每个阶段的交流分享，而借助人工智能技术的辅助，可以让学生的多样性的观点更好的呈现也可以利用人工智能技术来完善自己观点。

3.3 原则 3：协作共建，智慧互鉴

KB 理论与众不同之处在于它强调推动社区知识的发展，学生需要发挥认知的能动性推进观点的发展（Scardamalia,2021）。基于“民主化”知识原则的指引，本次课程设计的原则三是协作共建，智慧互鉴。课堂话语是促进学生参与课堂互动的核心要素之一，而现有的课堂话语体系多是围绕教师展开的“师生对话”或是被少数“话霸”主导的“生生对话”，缺乏全员互动的“广度”与探究的“深度”（张义兵，2020）。本原则强调协作共建，这是巨大的挑战，需要老师经过设计提供支架，帮助集体进行反思，重视他们的批评反馈或想法，从而获得更深层次的理解。在原则三的指导下，本次实践为学生的观点升华（rise above）提供了卡片支撑，确保每个人都能发言补充，每个人都能互相讨论建构集体的知识。

4. 基于知识建构原则的人工智能教育实践与结果

4.1 原则 1:动态迭代，反映心声——学生观点发展导向的动态课程设计

在原则 1 的指导下，学生从“关于学校最印象深刻的”的话题下发起讨论，从学生完成的任务单来看，“操场”“班级”“校园”是最频繁发生的场所，而这些都是来源与校园文化的建设，因此学生发现了真实的问题“校园文化的建设可以为学习生活带来快乐，我们如何去建设更好的校园文化？”学生由此开始更深入的讨论，“给每一个场合设计独一无二的 logo”，以获得更多的校园认同感，为发展对于共同体有价值的思想。在更广泛的讨论和合作后，学生的观点迭代到“校园建筑”，老师就“想象与实际的 STEAM 塔”让学生探究并呈现观点，知识社区形成“点、线、面、体”更丰富的观点，之后观点升级到“稳固性与对称面”，师生利用棍棒小球进行心声的呈现，搭建稳固的体。

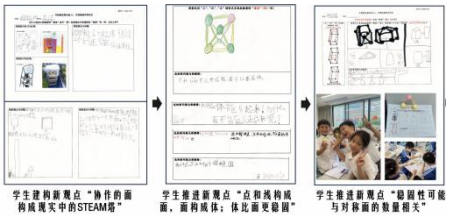


图 1 学生观点的发展历程

4.2 原则 2：智能创新 引领未来——生成式人工智能驱动的创新思维培养

我们鼓励学生利用生成式人工智能技术，突破传统思维的局限。在课程实施上，与人工智能的结合体现在 1、生成式 AI 辅助下的设计原则总结，借助生成式 AI 形成校园不同角落的 LOGO，并通过元对话建构出设计原则：发挥想象力，借助形状、标志物、符号、色彩呈现观点。通过与 AI 的对话，描述，将自己的观点与 AI 进行碰撞，构建出新的知识，丰富了知识社区，提高了质量，让学生寻找乐趣的同时完成了更高级的知识建构。（如图 2 所示）



图 2 学生利用人工智能完善作品设计

2、知识建构理论指导下的教学模式会产生大量的观点，但是观点的收集是为了促进观点的升级以及建构新概念（rise above），利用 WORD CLOUD 来支持的抽象概念推理，学生比较“想象中的 STEAM 塔与实际的 STEAM 塔”，并

借助 WORD CLOUD 开展课堂元对话，促进学生对抽象概念的推理。

4.3 原则 3：协作共建，智慧互鉴——知识建构卡片赋能的群体协作

课程强调集体智慧的力量，在原则 3 的指导下，课程必然经过观点分享的过程。在就“如何通过想象力来呈现观点”的讨论上，教师提供支架——知识建构卡片，卡片正面有学生发挥想象力创作的校园文化 logo，卡片反面是每个同学所要表达的观点的问题，学生就 logo 来表达“形状”“标志物”“符号”“色彩”是什么，它们所呈现的观点以及作者是如何发挥想象力来创作这个校园 logo 的。在建构活动中，每个学生都能领到属于自己的一个问题，借助知识建构卡片，让学生在开展小组活动时“有事做”“共同做”“建构着做”。



图 3 知识建构卡片设计

5. 结语

本文整合知识建构原则，提出三条原则来指导“师-机-生”三元结构的融合人工智能的知识建构教学模式。通过师机生促进“校园文化”集体知识的发展。在教学方法上利用学习单、人工智能技术、知识建构卡片来支持“社区知识”的建立，验证了三条原则的准确性和可行性。同时本研究既是对人工智能赋能教育的实践探索，也是使用知识建构理论来完善人工智能对教育领域的渗透的尝试。

参考文献

崔金贵 & 马莹莹.(2023).我国人工智能教育研究进展与展望. 高校教育管理(06),31-39.

彭绍东.(2021).人工智能教育的含义界定与原理挖掘. 中国电化教育(06),49-59.

卢宇,余京蕾,陈鹏鹤 & 李沐云.(2023).生成式人工智能的教育应用与展望——以 ChatGPT 系统为例.中国远程教育(04),24-31+51.

杨宗凯,王俊,吴砥 & 陈旭.(2023).ChatGPT/生成式人工智能对教育的影响探析及应对策略. 华东师范大学学报(教育科学版)(07),26-35.

陈凯泉,何瑶 & 仲国强.(2018).人工智能视域下的信息素养内涵转型及AI教育目标定位——兼论基础教育阶段 AI 课程与教学实施路径. 远程教育杂志(01),61-71.

张义兵,陈伯栋和 Marlene Scardamalia 等（2012）。从浅层建构走向深层建构——知识建构理论的发展及其在中国的应用分析。电化教育研究 33(09):5-12。

殷常鸿,张义兵 & 王晴燕.(2020).运用知识建构圈促进学生课堂深度互动研究. 中国电化教育(02),102-108+125

胡金艳,蒋纪平,陈羽洁 & 张义兵.(2021).知识建构社区中观点改进的机理研究：知识进化的视角。电化教育研究(05),47-54.

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

Scardamalia, M., & Bereiter, C. (2010). A brief history of knowledge building. Canadian Journal of Learning and Technology/La revue canadienne de l’ apprentissage et de la technologie, 36(1).

Bereiter, C., & Scardamalia, M. (2014). Knowledge building and knowledge creation: One concept, two hills to climb. In Knowledge creation in education (pp. 35-52). Singapore: Springer Singapore.

GCCCE 2024

德彰天下

学贯中西

ISBN: 978-6-26974-785-6



全球華人計算機教育應用學會