学习科学与计算机支持协作学习

移动、泛在与情境化学习

悦趣化学习、教育游戏与数字玩具

高等教育与成人学习的技术应用、教师专业发展

科技增强语言与人文学科学习

人工智能教育应用、智慧学习环境

学习分析与学习评估

STEM与创客教育

教育技术创新、政策与实践

融合前沿科技：成就新形态的学习者

Embracing Frontier Technology:

Cultivating the New Paradigm of Learners

第二十九届全球华人计算机教育应用大会

GCCCE 2025

工作坊论文集

Workshop Proceedings

中国 无锡 | 江南大学

中国 香港 | 香港理工大学

Jiangnan University，Wuxi, China

The Hong Kong Polytechnic University，Hong Kong, China

2025年5月24日至28日

24th - 28th May 2025

Copyright 2025 Global Chinese Conference on Computers in Education
All right reserved

ISBN: 978-626-97478-7-0
Publisher：Global Chinese Society for Computers in Education

第 29 届全球华人计算机教育应用大会

The 29th Global Chinese Conference on Computers in Education

GCCCE 2025 工作坊论文集

GCCCE 2025 Workshop Proceedings

主编 Editors

贾积有 北京大学
Jia Jiyou, Peking University
黄国豪 台湾云林科技大学
HWANG Gwo-Haur, Yunlin University of Science and Technology, Taiwan
黄凌云 香港教育大学
Huang Lingyun, The Education University of Hong Kong, HongKong
孟青泉 首都师范大学
Meng Qingquan, Capital Normal University
孙丹儿 香港教育大学
SUN Daner, The Education University of Hong Kong
尚俊杰 北京大学
SHANG Junjie, Peking University
温 韞 南洋理工大学
WEN Yun, Nanyang Technological University
吴声毅 台湾清华大学
WU Sheng-yi, Tsing Hua University
江绍祥 香港教育大学
KONG Siu Cheung, The Education University of Hong Kong
马志强 江南大学
MA Zhiqiang, Jiangnan University
胡祥恩 香港理工大学
HU Xiangen, The Hong Kong Polytechnic University, Hong Kong

编辑与校对 Editing and Proofreading:

刘靖熙 香港教育大学
LIU Jingxi, Jasper, The Education University of Hong Kong
许志文 香港教育大学
XU Zhiwen, Wynna, The Education University of Hong Kong
郑之姿 香港教育大学
ZHENG Zhizi, The Education University of Hong Kong
王 晨 香港教育大学
WANG Chen, The Education University of Hong Kong
贾凤麟 香港理工大学
JIA Fenglin, The Hong Kong Polytechnic University
王 娟 香港教育大学
WANG Juan, The Education University of Hong Kong
田沛瑶 香港大学
TIAN Peiyao, The University of Hong Kong

封面设计 Cover Page Design:

顾燕 江南大学
Gu Yan, Jiangnan University

目录

一、序言 Message from the Organiser 1
二、大会组织 Conference Organisation 2
GCCCE 2025 Workshop 1-10 4
結合生成式 AI 擬人諮詢顧問回饋機制的情境式中醫藥教育遊戲的初步評估 5
融入增强现实和游戏化策略的跨学科主题学习活动设计与实施研究 12
生成式人工智能支持的写作指导对学生学习意愿的影响机制研究 21
基於時間序列概念圖之機器人數位說故事學習模式對於學生學習成效、批判思考能力與團體自我效能之影響 30
The Effect of a Two-Tier-Test Approach Integrated with a Tablet IRS on University Students' Self-Efficacy and Attention 39
以擴增實境結合桌上遊戲建置協同學習系統探討遊程規劃之成效差異 44
結合「全班鷹架式合作學習」與「個人化即時回饋學習平臺」對於國小高年級學童書法學習之學習成效分析 49
因材網 e 度應用於數學文字題之成效分析-以小學五年級補救教學學童為例 57
基於雙層次測驗為基礎之博物館環景導覽對學習者學習動機與成效之影響 65
探討專業學習社群對國小教師在虛擬實境融入教學之教學自我效能與 TPACK 的影響 69
ChatGPT 模擬 OSCE 護病溝通中學習感知與溝通表現之性別差異與相關性分析 78
PAPAMAMA-TOMO: A Chatbot-Based Support System for Foreign Guardians to Navigate Japanese School Culture 86
以「雙媒材鷹架」促進學生故事創作：桌遊與生成式 AI 91
建立「整合歸納式與演繹式的學習環境」來提升英語文法的學習動機 95
年齡對國小學童使用虛擬實境交通安全學習系統之影響 105
教出好品德：運用 Zuvio 即時反饋系統輔助探討大學生感恩學習效果於多元創新教學實踐課程之研究 113
基于代理人基模型与大语言模型人机协作生成式 AI 桌游设计与模拟：以《儒行天下：孔子周游列国》为例 121

導入即時回饋系統於遊戲化節奏訓練對學習者學習動機與自我效能之影響 130
支持個人化學習的沉浸式 VR 學習環境：四季古代詩詞追尋之旅 134
生成式 AI 輔助軟體系統架構工具之設計與開發 139
運用生成式 AI 工具輔助學習軟體系統架構之學習成效分析與探討 148
應用線上協作學習於新詩創作課程對自我反思、洞察及學習態度的影響 156
從性別差異的角度探討使用數位遊戲來學習同義字之影響 164
結合翻轉教學與遊戲式學習提升中學生人工智慧素養 173
专业成长视角下的中小学教师人工智能素养测评方案研制 182
具身认知视角下 AI 机器人编程课程的创新设计与课堂实践 186
双师一线：基于数字课程的人工智能实践教学探索 195
智能时代的信息甄别：生成式人工智能对批判性思维影响的系统性文献综述 201
面向"AI4S"的高中人工智能课程开发与实施 210
人工智能应用于 STEM 教育以培养计算思维的系统性文献综述 219
优选·精研·迭代：探索基于浦育平台新一代人工智能教育的乡村小学智慧校园管理新路径 227
融合范例学习与多智能体的高中 STEM 课程设计与实践 236
一项基于 AI 的多语同步口语平台在新加坡小学的应用探索 244
AIGC 课程设计策略与实践路径——以“诗歌集创作”项目为例 253
三阶四维：基于 ADDIE 模型的人工智能数字课程设计研究 259
在第二外语学习中大语言模型应用的系统文献综述 265
Multisensory Learning Method in Plant Science for Elementary School Students with Visual Impairments 279
結合行動科技脈絡線索鷹架與競標機制的桌上遊戲以培訓採購策略思維之初探 282
The Impact of Junior High School Science Teachers' Professional Development on Beliefs and Practices of AI in Assessment 285
學習信念、機器人程式設計態度對於運算思維、創造力與合作學習的影響研究 295
STEM 積木建構活動之學生策略運用類型之文獻分析 304
生成式人工智能赋能科普图像及动画创作—以航天科普动画“落火”为例 307

結合擴增實境與 STEAM 教育之幼兒創傷知情互動軟布書設計與實作 316
針對完整主題的系列教具和新的論述 320
PuppyCodeReview: An AI-based Code Review System 328
沉浸式虛擬實境學習導向教育元宇宙平台開發：從設計思維到教學應用 331
信息技术赋能科学素养评测：发展脉络、研究热点与未来趋势 336
初探虛擬實境教材沉浸程度之影響：以 VR 動畫專注觀察為例 345
多智能体协作下的人智教研共生体建构与实践模式研究 350
多智能体系统如何塑造未来教师？——AI-TPACK 素养提升的实验研究 359
促进小学生计算思维的均衡发展：一种跨区域协作的无屏幕机器人编程教育模式 368
小学机器人教育中具身学习模式构建与案例开发 376
哪一种学习支架更有效？学习支架对不同元认知水平学生学习成绩和计算思维的影响研究 384
基于隐喻的机器人编程方法：促进幼儿的计算思维和积极学习行为 392
運用 VEX 機器人學習编程：提升 K-12 教育中的計算思維與解決問題能力 402
基于 STEAM 的幼儿园大班实体编程课程开发：活动设计与推进路径的实证研究 411
小学信息科技体验式学习活动设计与实践——以《算法的特征》一课为例 417
游戏课程化视域下幼儿编程教师专业发展的实践探索 425
A Practical Exploration of Early Childhood Programming Teachers’ Professional Development in the Perspective of Game Curricularization 425
项目编程游戏：幼儿计算思维的阶梯式培养 430
Project Programming Game: Step-by-Step Development of Computational Thinking in Young Children 430
玛塔挑战坊：幼儿园跨学科编程课程的实施路径探索 438
Matatalab Challenge Program: Exploring Implementation Pathways for 438
Interdisciplinary Programming in Kindergartens 438
基于 GenAI 的结对编程策略对初中生计算思维态度和编程自我效能感的影响研究 443
A study of the effects of GenAI-based pair programming strategy on middle school students' attitudes toward computational thinking and programming self-efficacy 443

家長參與對兒童编程學習影響的研究綜述與展望 454
未来教育视域下“智慧•玩创”学习空间新样态的构建与实施 463
Construction and Implementation of the New Model of "Smart Play and Creation" Learning Spaces from the Perspective of Future Education 463
数字驱动，智启未来——可持续发展理念下的幼儿园数智+评价创新实践 469
Data-Driven Enlightenment: Innovation of Intelligent Assessment in Preschool Education within Sustainable Development Framework 469
计算思维导向下大班游戏化编程教学的实践研究——以"机器人夺宝"科学规则游戏活动为例 478
“编”玩边学，“程”现童趣——数智 AI 赋能幼儿园编程启蒙课堂的实践研究 485
计算思维在幼儿艺术教育中的创新实践 490
Innovative Practices of Computational Thinking in Early Childhood Art Education 490
Google Translate 圖像報讀功能對讀寫障礙的閱讀理解之學習成效 497
The Learning Effectiveness of the Google Translate Image Reading-Aloud on Reading Comprehension for Students with Dyslexia Function 497
2015-2024 年职业教育数字化研究：基于 CiteSpace 的可视化分析 505
A Study on Digitalization of Vocational Education from 2015 to 2024: Visual Analysis Based on CiteSpace 505
Generative Artificial Intelligence in Language Learning: A Review of Selected Journal Publications from 2023 to 2024 513
以 ChatGPT 創作故事對國二生學習動機與學習體驗心理因素之相關研究 518
The Study on the Relationship Between Using ChatGPT to Create Stories and the Learning Motivation and Psychological Factors of Learning Experience in 8th Grade Students 518
我国教育元宇宙研究现状与未来趋势——基于 Citespace 可视化图谱分析 527
Investigating the Relationships among Virtual Reality Features, Technology Acceptance and Learning Performance in VR-assisted Language Learning 534
数字赋能终身学习的研究热点与阶段特征——基于中国知网 2014-2024 年的文献可视化分析 539

校園行政工作者運用生成式 AI 輔助行政業務之行為意向關係研究－以《國小行政工作者》為例 548
創意知識工作者之創造力需求及工作投入: 以內在動機與自我效能為中介變項 557
A Study on the Impact of Embodiment in AR Learning on Learners' Technology Acceptance 565
智能时代下 GenAI 的使用对学生批判性思维影响研究——基于 UTAUT 模型的使用行为机制解释与认知负荷的调节作用 572
联通主义学习群体协同知识创生过程中的概念网络演化规律* 581
数字化背景下职业教育教师教学能力现状及提升策略研究 590
国家中小学智慧教育平台的应用意愿与行为的影响因素研究 598
大語言模型辅助近体诗创作的实践与反思：基于 Deepseek 之人机交互分析 607
AI 智能技术赋能高等教育过程性评价体系构建 616
面向小学生信息社会责任培养的体验式教育游戏框架设计与应用研究 625
生成式人工智能赋能课堂教学的模式和实践路径——基于扎根理论的研究 634
数字化转型背景下的循证教研新范式：模式构建与实践探索 644
师范生数字化教学能力的影响因素及培养策略研究 652
人工智能教育应用的伦理问题及其治理对策 661
教育数字化转型背景下校长信息化教学领导力影响因素研究 667
智慧教室环境下促进语文课堂互动的教学策略研究 676
协作知识建构对小学生认识论信念的促进研究 685
变革型领导与智能教学评价系统的有效整合——基于组织学习理论的多案例研究 690
人工智能赋能高等教育课程教学创新研究进展综述 698
高校教师数字化教学胜任力核心要素研究 703
Mechanisms of Collective Knowledge Advancement in Knowledge-Building Communities: A Perspective from Epistemic Emotion 709
教师成为学习设计师的实践探索：专家指导与 AI 协同下 K-12 STEM 课程的五轮迭代研究 719
元宇宙视域下中国传统文化传播形态与教育应用研究 724
流程自動化之多層級教育關係人行動建議智慧儀表板：以台灣教育大數據為例 732

探索人工智能增强的 STEAM 课程中小学生的计算思维过程：一项滞后序列分析方法 740
人智协同的跨学科知识创生双链螺旋模型研究 748
GAI 时代提示词效能感和教学自我效能感对职前教师 TPACK 水平的影响研究 753
翻转课堂下不同平均绩点学生的自主学习路径分析：K-means 聚类与过程挖掘 762
生成式人脸隐私保护赋能安全学习分析 771
基于生成式人工智能的集体反思中教师合作对话模式的认知网络分析 780
机器遗忘赋能安全智能教育应用实践 789
基于视频分析的职前科学教师集体反思：一项基于设计的研究 798
六顶思考帽助力的协同反思对促进师范生教学设计学习效果的影响研究 807
大学人工智能通识课程体系构建:现状与实施建议 814
Research on the Construction of Assessment Indicators for College Students' Learning Engagement in Blended Learning Environment 823
教育数字化转型下可信数据空间的构建路径 832
大模型智能体评估教学实施新方法：使用教学设计达成度 838
Exploring the effects of AI feedback on affective-motivational outcomes in language learning: A Meta-analysis 845
人智协同视角下 AIGC 对协作学习中学生批判性思维能力的影响研究 856
基于 ICAP 框架：生成式人工智能对职前教师教学设计能力提升的实证研究 863
社交批注对学生在同步在线学习中参与度的影响 869
基于调节学习的多智能体支持的协同论证设计研究 878
从“局外人”到“小玩家”：幼儿在线媒体沉迷向具身游戏参与的转化引导与支持策略研究——一项长周期个案探索 887
Research on the Design of Multi-agent Education Architecture for Innovative Thinking 899
LLM-RAG 系统的设计与教学实践应用 909
基於小學語文課程的多模態學習分析方法研究 918
聚焦创造性问题解决：跨学科项目化学习实施策略分析——以“许你一个不迷路的校园”两轮迭代为例 928

基于多智能体的科学问题解决能力评测模型自动构建方法.....936

一、序言 Message from the Organiser

全球华人计算机教育应用大会 (Global Chinese Conference on Computers in Education，简称 GCCCE) 是由全球华人计算机教育应用学会主办的国际性学术会议。第 29 届全球华人计算机教育应用大会（GCCCE 2025）于 2025 年 5 月 24 日至 28 日由江南大学和香港理工大学共同举办。

GCCCE 2025 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。

本届大会共设立 10 个工作坊，最终收录论文 140 篇。许多工作坊安排了专家讲座与特邀报告。每一个工作坊都包含两个地区以上的学者参与，内容丰富并积极推进了优质的交流。这 10 个工作坊的主题分别叙述如下：

工作坊 1：创新互动回馈科技提升学习动机
工作坊 2：电脑支援个人化与合作学习
工作坊 3：K-12 中的人工智能教育与课堂创新实践
工作坊 4：科学及跨领域教育的实作创新：发展与研究
工作坊 5：生成式人工智能支持的多智能体教育应用研究
工作坊 6：迎接未来的挑战：智能时代儿童计算思维与编程教育
工作坊 7：迎接「元宇宙」的世代，如何融入新科技于教学
工作坊 8：智能时代的知识与教学创新
工作坊 9：学习投入与学习行为建模
工作坊 10：教育智能体与学习分析研究

本届工作坊筹备历程能有效进展，由衷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！

贾积有 北京大学
工作坊执行主席

二、大会组织 Conference Organisation

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)
江南大学
Jiangnan University
香港理工大学
The Hong Kong Polytechnic University

承办单位 Host

江南大学人文学院
School of Humanities Jiangnan University
江苏“互联网+教育”研究基地
Jiangsu Research Center of “Internet Plus Education”, Jiangnan University
香港理工大学高等教育研究及发展院
Institute for Higher Education Research and Development, The Hong Kong Polytechnic University

大会主席 Conference Chair

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

大会顾问 Conference Consultants

陈德怀 台湾中央大学
CHAN Tak-Wai, Central University, Taiwan
朱蕙君 台湾东吴大学
CHU Hui-Chun, Soochow University, Taiwan
黄龙翔 南洋理工大学
WONG Lung Hsiang, Nanyang Technological University, Singapore
刘革平 西南大学
LIU Geping, Southwest University

国际议程协调委员会 International Program Coordination Committee

主席 Chair

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

副主席 Co-Chairs

尚俊杰 北京大学
SHANG Junjie, Peking University
温 韞 南洋理工大学
WEN Yun, Nanyang Technological University, Singapore
吴声毅 台湾清华大学
WU Sheng-yi, Tsing Hua University, Taiwan

在地组织委员会 Local Organizational Committee

主席 Chairs

马志强 江南大学
MA Zhiqiang, Jiangnan University
胡祥恩 香港理工大学
HU Xiangen, The Hong Kong Polytechnic University, Hong Kong

在地委员会成员 Members of the Local Organizational Committee

大会统筹领导组 (Conference Coordination Leadership Team)

王 晖 (Wang Hui) 伍红林 (Wu Honglin) 陈明选 (Chen Mingxuan)
马志强 (Ma Zhiqiang) 刘向永 (Liu Xiangyong)

保障组 (Guarantee Team)

刘向永 (Liu Xiangyong) 钱逸舟 (Qian Yizhou) 严大虎 (Yan Dahu)
权国龙 (Quan Guolong) 田 娜 (Tian Na) 舒 杭 (Shu Hang)

会场组 (Venue Team)

王志军 (Wang Zhijun) 牟智佳 (Mou Zhijia) 王 萌 (Wang Meng)
高 红 (Gao Hong) 刘文辉 (Liu Wenhui)

材宣组 (Material Promotion Team)

王 靖 (Wang Jing) 张红英 (Zhang Hongying) 顾 燕 (Gu Yan)
姚佳佳 (Yao Jiajia) 朱芳芳 (Zhu Fangfang)

联络组 (Contact Team)

蔡慧英 (Cai Huiying) 晋欣泉 (Jin Xinquan)

指导与支持组 (Guidance and Support Team)

张春梅 (Zhang Chunmei) 徐 璐 (Xu Lu) 王小根 (Wang Xiaogen)
孙 力 (Sun Li) 张 群 (Zhang Qun) 陈根岩 (Chen Genyan)
魏建新 (Wei Jianxin) 刘学勤 (Liu Xueqin) 徐 跃 (Xu Yue)

国际和在地对接小组 (International and Local Coordination Team)

香港教育大学 (The Education University of Hong Kong)

刘靖熙 (LIU Jingxi, Jasper) 许志文 (XU Zhiwen, Wynna) 郑之姿 (ZHENG Zhizi)
王 晨 (WANG Chen) 王 娟 (WANG Juan) 王晓兰 (WANG Xiaolan)

香港理工大学 (The Hong Kong Polytechnic University)

李蕙綸 (Lee Wai Lun, Aka)

工作坊提案及组织委员会 (Workshop)

主席 (Executive Chair)
贾积有 北京大学
Jia Jiyou, Peking University

副主席 (Co-Chairs)
黄国豪 台湾云林科技大学
HWANG Gwo-Haur, Yunlin University of Science and Technology, Taiwan
黄凌云 香港教育大学
Huang Lingyun, The Education University of Hong Kong, HongKong
孟青泉 首都师范大学
Meng Qingquan, Capital Normal University

GCCCE 2025 Workshop 1-10

結合生成式 AI 擬人諮詢顧問回饋機制的情境式中醫藥教育遊戲的初步評估

Preliminary Evaluation of a Situational Traditional Chinese Medicine Educational Game Combined with a Generative AI Anthropomorphic Consultant Feedback Mechanism

簡志忠[1]，王瑛玫[2]，賴冠名[2]，侯惠澤[1*]
[1]臺灣科技大學應用科技研究所
[2]臺北榮民總醫院新竹分院
[*] hthou@mail.ntust.edu.tw

【摘要】中藥複雜成分與人體系統功能之間的非線性與多維度相互作用極為精密，增加了研究的挑戰性。在以臨床應用為核心的中醫藥創新與傳承過程中，師徒制在學術醫學領域發揮重要作用，透過提供指導、資源與專業陪伴，能有效提升受訓者的自主性與熟練度。此外，透過擬真情境的設計，讓學習者在實際解決問題中獲得經驗，可促進學習遷移。本研究以生成式 AI 為即時互動的擬真師傅，設計一款情境教育遊戲，幫助新進藥劑師理解中藥方劑的應用。初步結果顯示，該遊戲顯著提升了學習者的學習心流、降低認知負荷，並提高了遊戲的接受度。

【關鍵詞】 中藥方劑；遊戲式學習；情境擬真；生成式 AI

Abstract: *The nonlinear and multidimensional interactions between the complex components of traditional Chinese medicine (TCM) and human physiological systems are highly intricate, adding to the challenges of research. In the process of innovation and inheritance of TCM with a focus on clinical application, the mentorship model plays a significant role in the field of academic medicine. By providing guidance, resources, and professional companionship, it effectively enhances trainees' autonomy and proficiency. Moreover, designing realistic scenarios allows learners to gain practical experience in solving real-world problems, thereby facilitating transfer of learning. This study utilizes generative AI as an interactive virtual mentor to develop a scenario-based educational game aimed at assisting novice pharmacists in understanding the application of TCM formulas. Preliminary results show that the game significantly improves learners' flow, reduces cognitive load, and increases game acceptance.*

Keywords: Chinese Medicine Formulas, Game-based learning, Situational simulation, Generative AI

1. 前言

現代科技推動中醫藥研究發展，但中醫藥複雜成分與人體系統的非線性、多維相互作用影響，增大了研究難度，以臨床應用為導向的中醫藥學創新與傳承成為核心議題(Li et al., 2023)。Li 等 (2008)研究提到，中醫藥中一些草藥組合穩定且明顯有效，因此被記錄下來並作為配方代代流傳下來，這導致了超過 30 萬個已知配方的產生，為中醫臨床治療奠定了基礎。然而，中藥複方藥理研究面臨巨大的挑戰，每個複方需要分析 100 種或更多種化合物（也稱為成分）的組合(Zhao et al., 2010)。因此，草藥和方劑相關的複雜藥理機制學習對於中醫藥領域的人員來說是一項艱鉅的任務(Li et al., 2022)。同時，中醫藥教學面臨如何在傳統有限課程中向學生傳授基礎知識並激發學習熱情的挑戰(Zhu et al., 2022)。因此，幫助新進藥劑師理解中藥方劑的應用及提升學習效率，為本研究的研究動機。

Sezgin 等 (2020) 研究指出，採用遊戲式學習活動可以搭配多種學習理論，並提供問題解決和決策的任務探索。可以較容易地吸引學習者的注意力並對激發學習動機有積極影響 (Chang et al., 2017)。另外，Sezgin 與 Yüzer (2020) 研究也提到，遊戲式學習透過情境化設計可對學習者的參與動機和對遊戲的有效性產生直接影響。Rakhit 等 (2024)研究提到，學術醫學中的師徒制通過指導、資源和陪伴，能有效提升學習者的自主性與專業熟練度。另外，Tapingkae 等 (2020)研究發現，達到學習遷移的關鍵在於擬真情境與現實場景的相似性，這能增加解決實際問題的練習機會。

有鑑於此，本研究設計一款以生成式 AI 作為即時互動擬真師傅，協助新進藥劑師在探訪患者的任務過程中，達到師徒機制隨時討論及引導成效，以提升理解中藥方劑應用之情境擬真教育遊戲。Chien 等 (in press)研究指出，有限制資料來源的 AI 工具應該是較具潛力做為學習鷹架的運用方式，學習者可依據生成式 AI 檢索後反饋的資訊脈絡理解程度來調整自己的檢索節奏及檢索內容，提升學習控制感，更感受到即時反饋互動式鷹架之有用性。本研究同時初探學習者的學習心流、認知負荷以及遊戲接受度之心理狀態。

2. 研究方法

2.1. 研究對象與程序

本研究參與者為台灣 15 名具有合格藥劑證照之藥劑師。採自由於網路報名。參與者在施測前均未曾以 GAI 互動機制認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)、進行教學活動(30 分鐘)，後測問卷施測評量 (15 分鐘)。

2.2. 教學遊戲設計

本研究設計一套擬真情境互動劇本教育數位遊戲《榮藥使命》。本研究採用 ThingLink 平台建構情境擬真場景，場景中各線索如圖 1 所示。

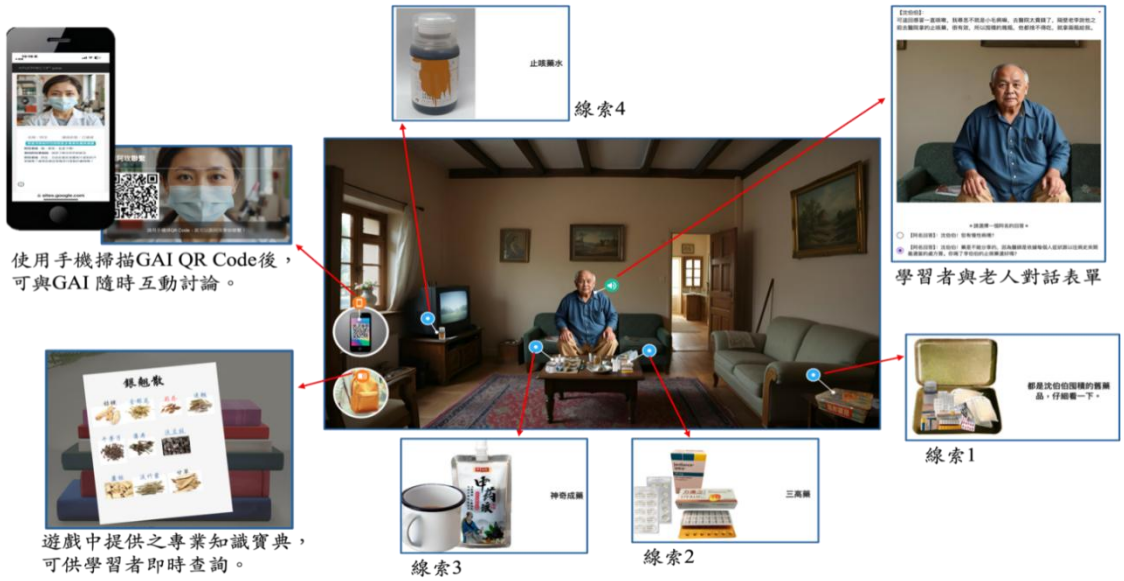


圖 1 戲畫面

遊戲任務是讓學習者化身一位新進藥劑師，奉命探訪偏鄉獨居老人。學習者依據老人所描述身體的不適症狀及透過平台探索老人生活相關訊息，再透過詢問 GAI (資深前輩)理解臨床中醫藥相關藥理經驗，其中 GAI 即時互動鷹架的引導特色，如圖 2 所示。同時，過程中

也可以查詢遊戲所提供的『中藥材寶典』，來推論老人身體不適可能為何種中藥材影響，進而推估來路不明的方劑成分。最後學習者需提出更適合案例老人的方劑，即可達成遊戲任務。學習者達成任務後，立即可獲的任務星等及解析，其目的透過與 GAI 即時互動及反思達到師徒制學習成效。

本研究所架構之 GAI 即時互動鷹架機制，學習者可先透過手機掃描 QR Code 後，即可在手機上以語音輸入方式與 GAI 互動，而 GAI 以文字方式反饋，此目的可達到互動紀錄，以利學習者隨時查詢反思。

圖 2 GAI 即時互動鷹架的引導特色

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 與 Li (2014)翻譯為中文版本。共有兩大維度，分別為心流前提以及心流經驗，共計 22 題。此問卷採用李克特氏五點量表，內部一致性 Cronbach's α=.937，顯示該量表具有極高的信度。

3.2. 遊戲接受度量表

為了瞭解學習者對學習方式接受度之差異，本研究修正 Davis (1989) 所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，共計 12 題。此問卷採用李克特氏五點量表，內部一致性 Cronbach's α=.931，顯示該量表具有極高的信度。

3.3. 認知負荷量表

為了探討學習者在整個學習歷程的認知負荷差異，本量表以 Leppink 等 (2013)的認知負荷量表為基礎，並參考 Klepsch 等 (2017) 所提出的以差異化的認知負荷量表改編而成，其方

式可靠且有效地測量。量表分為三大維度共 8 題，包含內在認知負荷、外在認知負荷 、增生認知負荷。此問卷採用李克特氏五點量表，內部一致性 Cronbach's α=.767，顯示該量表具有極高的信度。

3.4. 活動心得問卷

本問卷採用 5 個開放式問題，每題分別皆以量化分析及質性方式進行，以瞭解學習者對學習歷程的心得與建議回饋。此問卷採用李克特氏五點量表，內部一致性 Cronbach's α=.891，顯示該量表具有極高的信度。

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 = 4.30)、心流前提(M = 4.33)、心流經驗(M = 4.27)皆顯著高於量表中位數(即 3)。Chen 與 Syu (2024) 研究提出，透過提供數位角色扮演遊戲來強化擬真情境的主要特徵，從而增強學習者在學習過程中的臨場感和心流。另有研究發現在體驗式的擬真情境學習環境，可以增強或可能取代傳統的學習方法，使得學習者更專注學習 (Smith & Ericson, 2009)。

在遊戲心得方面：整體遊戲回饋(M = 4.63)、遊戲有用性(M = 4.68)、遊戲易用性(M = 4.49)皆顯著高於 3。Liu 等 (2018) 研究指出，具有可用性、感知易用性和認知觀點改變的情境線索，可導致持續的表現，這種表現可以由在情境中遊戲的心流體驗和滿意度顯現出來。另外， Wang et al. (2021)研究指出，在情境學習中，學習者的知識可通過基於具體場景的實踐思維和活動獲得，而且情境式學習是一個強大而實用的框架，可以突顯隱性知識和技能、活動能力 (Frade & Da Rocha Falcão, 2008)、自我效能(Donaldson et al., 2020)，對決策的理解 (Walker & Zeidler, 2007)，以及興趣和動機。

在認知負荷方面：外在認知負荷(M =1.80) 顯著低於 3，同時增生認知負荷(M =4.07) 顯著高於 3。Schnotz 與 Kürschner (2007)研究指出，促進有效深入理解的一種方法是減少外在認知負荷，通過減少外部認知負荷而釋放的工作記憶容量，可用於產生增生認知負荷，這將更有可能激發深度理解。Kalyuga (2011)也指出，教學設計需修改簡化學習目標可能涉及的學習程序或概念，以適當的教材呈現方式來提高增生認知負荷。

另外，在透過與 GAI 的互動對話內容情境之問卷上，皆顯著高於 3。Chien 等 (in press) 研究指出，GAI 具有提供動態和互動式即時回饋的潛力。Zuo 等 (2022)也指出，在一個遊戲化學習的研究中，可以嘗試設計一款線上情境化教育遊戲，並使用 GAI 鷹架來輔助學習。另外，學習者在問卷中反饋『AI 回覆訊息不會不知所云，甚至會給提示，讓學習玩家可以更身歷其境』、『AI 會即時回覆資料庫中的內容，且能反覆提問，會比跟真人討論時更容易集中在問題上找答案』、『互動對話內容情境更加針對問題產生答案』、『很擬真，比較像真人對話回應我的疑惑』。由此推論，本研究所設計的 GAI 擬人諮詢顧問即時互動認知鷹架，不僅具有專業引導成效，同時擁有類似真人關懷互動，提升情境式師徒制學習成效。

表 1 單一樣本 Wilcoxon 符號等級檢定摘要表

	(N=15)			
	M	*SD*	*Z*	*Sig.*
整體心流	4.30	0.55	3.41***	<.001
心流前提	4.33	0.55	3.42***	<.001
心流經驗	4.27	0.57	3.30***	<.001
整體遊戲回饋	4.63	0.48	3.45***	<.001

有用性	4.68	0.47	3.49***	<.001
易用性	4.49	0.65	3.47***	<.001
內在認知負荷	3.17	0.88	0.86	0.389
外在認知負荷	1.80	0.55	-3.43***	<.001
增生認知負荷	4.07	0.64	3.43***	<.001
透過與 GAI 的互動對話內容情境:				
讓我感受到對話中的引導。	4.53	0.64	3.42***	<.001
比日常工作中真人面對面的討論藥材解析時，讓我更沒壓力。	4.67	0.62	3.49***	<.001
比日常工作中真人面對面的討論藥材解析時，更能集中思維。	4.33	0.62	3.40***	<.001
比起一般運用 ChatGPT 的對話更加擬真。	3.67	0.90	2.27*	0.023
遊戲結束之後，讓我更清楚自己在中藥方劑的解析能力。	4.47	0.52	3.51***	<.001

* p <0.05, *** p <0.001

5. 結論

《榮藥使命》是本研究設計一套藥劑師培訓『理解中藥方劑』的擬真情境互動劇本教育數位遊戲。綜合以上數據分析，顯示本研究以新進藥劑師探訪老人情境擬真體驗式學習，以導入生成式 AI 擬人諮詢顧問即時互動機制，在學習過程中可以讓學習者維持高度學習心流、並降低外在認知負荷，且提高增生認知負荷。本研究結果推論擬真情境結合生成式 AI 擬人諮詢顧問互動機制的遊戲設計，可達到師徒制學習，是適合複雜的中醫藥學學習，應有助於學習投入及成效。未來研究可以再增加更多的樣本量，並探討鷹架導向的 GAI 擬人諮詢顧問互動機制，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

Chang, C. C., Liang, C., Chou, P. N., & Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior*, *71*, 218–227. https://doi.org/10.1016/j.chb.2017.01.031

Chen, C. H., & Syu, J. Y. (2024). Effects of integrating a role-playing game into a virtual reality-based learning approach on students' perceptions of immersion, self-efficacy, learning motivation and achievements. *British Journal of Educational Technology*, *55*(5), 2339–2356. https://doi.org/10.1111/bjet.13436

Chien, C. C., Chan, H. Y., & Hou, H. T. (in press). Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction. *Journal of Research on Technology in Education*, 1–20. https://doi.org/10.1080/15391523.2024.2338085

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Donaldson, T., Fore, G. A., Filippelli, G. M., & Hess, J. L. (2020). A systematic review of the literature on situated learning in the geosciences: beyond the classroom. *International Journal of Science Education*, *42*(5), 722–743. https://doi.org/10.1080/09500693.2020.1727060

Frade, C., & da Rocha Falcão, J. T. (2008). Exploring Connections Between Tacit Knowing And Situated Learning Perspectives In The Context Of Mathematics Education. *New Directions for Situated Cognition in Mathematics Education*, *45*, 205–231. https://doi.org/10.1007/978-0-387-71579-7_10

Hou, H. T., and Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior*, *30*, 29–38. https://doi.org/10.1016/j.chb.2013.07.052

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, *23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7

Kiili, K. (2006). Evaluations of an Experiential Gaming Model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *2*(2), 187–201. https://jyx.jyu.fi/handle/123456789/20195

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology, 8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Li, D., Guo, H., Niu, L., Yin, Q., Zhang, Y., & Zhuang, P. (2023). Clinical value-oriented research paradigm about inheritance and innovation development of TCM dominant diseases. *Chinese Herbal Medicines*, *15*(4), 476–484. https://doi.org/10.1016/j.chmed.2023.09.002

Li, W.-F., Jiang, J.-G., & Chen, J. (2008). Chinese medicine and its modernization demands. *Archives of Medical Research*, *39*(2), 246–251. https://doi.org/10.1016/j.arcmed.2007.09.011

Li, Y., Lin, Z., Wang, Y., Ju, S., Wu, H., Jin, H., Ma, S., & Zhang, B. (2022). Unraveling the mystery of efficacy in Chinese medicine formula: New approaches and technologies for research on pharmacodynamic substances. *Arabian Journal of Chemistry*, *15*(11), 104302–104302. https://doi.org/10.1016/j.arabjc.2022.104302

Liu, Y., Liu, D., Yuan, Y., & Archer, N. (2018). Examining situational continuous mobile game play behavior from the perspectives of diversion and flow experience. *Information Technology & People*, *31*(4), 948–965. https://doi.org/10.1108/itp-02-2016-0042

Rakhit, S., Fiorentino, M. N., Alvarado, F. A., Eastham, S. C., Gondek, S. P., Patel, M. B., & Streams, J. R. (2024). Mentorship in Surgery: Best Practices for Mentor–Mentee Relationships. *Current Surgery Reports*, *12*(4), 58–66. https://doi.org/10.1007/s40137-024-00390-3

Schnotz, W., & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, *19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Sezgin, S., & Yüzer, T. V. (2020). Analysing adaptive gamification design principles for online courses. *Behaviour & Information Technology*, *41*(3), 1–17. https://doi.org/10.1080/0144929x.2020.1817559

Sezgin, S., Ve Oyun, O., Öğrenme, T., Dijital, Y., Tipolojileri, O., Bir, Makale, M.-S., Öz, B., & Makalesi, A. (2020). Digital player typologies in gamification and game-based learning approaches: a meta-synthesis. *Bartın University Journal of Faculty of Education*, *9*(1), 49–68.https://doi.org/10.14686/buefad.610524

Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual Reality*, *13*(2), 87–99. https://doi.org/10.1007/s10055-009-0113-6

Tapingkae, P., Panjaburee, P., Hwang, G.-J., & Srisawasdi, N. (2020). Effects of a formative assessment-based contextual gaming approach on students' digital citizenship behaviours, learning motivations, and perceptions. *Computers & Education*, *159*, 103998. https://doi.org/10.1016/j.compedu.2020.103998

Walker, K. A., & Zeidler, D. L. (2007). Promoting Discourse about Socioscientific Issues through Scaffolded Inquiry. *International Journal of Science Education*, *29*(11), 1387–1410. https://doi.org/10.1080/09500690601068095

Wang, W., Coutras, C., & Zhu, M. (2021). Empowering computing students with proficiency in robotics via situated learning. *Smart Learning Environments*, *8*(1). https://doi.org/10.1186/s40561-021-00167-6

Zhao, J., Jiang, P., & Zhang, W. (2010). Molecular networks for the study of TCM Pharmacology. *Briefings in Bioinformatics*, *11*(4), 417–430. https://doi.org/10.1093/bib/bbp063

Zhu, Y., Wu, Y., Zhang, H., Zhao, Y., Ren, Y., Hu, W., Li, P., Wang, H., Wang, Y., Zhang, B., Wang, Y., Yuan, H., Zhang, J., Li, W., Zhao, Y., Zhao, J., Shi, Y., Shu, Y., Zhang, X., & Sun, Y. (2022). A national survey on how to improve traditional Chinese medicine learning internationally: Perceptions from both teachers and students. *Integrative Medicine Research*, *11*(4), 100895–100895. https://doi.org/10.1016/j.imr.2022.100895

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

融入增强现实和游戏化策略的跨学科主题学习活动设计与实施研究

Research on the Design and Implementation of Interdisciplinary Theme Learning Activities Incorporating Augmented Reality and Gamification Strategy

顾双玲[1] 张慕华[1*]
[1]首都师范大学初等教育学院
*zhangmuhua@cnu.edu.cn

【摘要】 游戏化学习作为一种新兴学习方式，目前已经在教育领域得到广泛应用。随着教育数字化转型政策的推进，增强现实、虚拟现实等技术成为支持游戏化学习开展的有效手段。然而当下小学跨学科主题学习融入增强现实技术和游戏化策略的案例较少。本研究采用准实验研究法，在北京市某小学开展《探索太阳系》主题学习活动，融入增强现实技术和游戏化策略，旨在探究其对小学生跨学科主题学习的影响。研究发现，在小学跨学科主题学习中融入增强现实技术和游戏化策略，能够给学生提供积极的学习体验，并有效提升学生的学习成效。

【关键词】 增强现实；游戏化学习；学习体验；学习成效；跨学科主题学习

Abstract: *Game-based learning, as an emerging learning method, has been widely applied in the education sector. With the promotion of digital transformation, technologies such as augmented reality (AR) and virtual reality have become effective means to support the development of game-based learning. However, there are relatively few cases of integrating AR and gamification strategies into interdisciplinary theme learning in primary schools at present. This study adopts a quasi experimental research method to carry out the theme learning activity of "Exploring the Solar System", aiming to explore its impact on students. Research has found that incorporating AR and gamification strategies into interdisciplinary theme learning in primary schools can provide students with positive learning experiences and effectively improve their learning effects.*

Keywords: Augmented Reality, Game-based learning, Learning experience, Learning effect, Interdisciplinary thematic learning

1.引言

游戏化学习作为一种将游戏或游戏元素、游戏设计和游戏理念应用于非游戏情境中的教学方法，具有娱乐性、游戏性、目标性和规则性等性质，能够为学习者提供积极、有趣和有效的学习体验，在真实的情境中培养学习者的问题解决能力和创造力，从而提升学习者的学习效果（曹语苍，2023）。随着新一轮科技革命和产业变革的深入发展，数字技术越来越深刻地影响着人类思维方式与生产生活方式。学校教育应当重视人工智能技术与信息素养教育的有机结合，发挥主观能动性与创造性来营造传统教学不易实现的情境，从而调动学习者的积极性。增强现实（AR）技术具备虚实结合、实时互动等特点，有助于创设沉浸式情境，能够让学习的发生不再拘泥于传统的学校范畴，还可以发生在不断变换的碎片化时空场景之中，使得游戏化学习的场景更加贴合学习主题（武法提等，2018）。同时，有研究表明基于

增强现实的学习模式会吸引和激励学习者学习并增强他们的信心，有助于提高学习者的学习动机和参与度（Sümer et al.,2024）。因此，使用游戏化学习策略和 AR 技术可以为学习者提供丰富的学习体验，通过营造真实有趣的学习氛围以提升学习成效。

跨学科主题学习是构建我国教育学自主知识体系的重要概念，其本质在于通过整合至少两门课程，围绕特定主题或问题开展综合性学习活动，以培养学生的核心素养（崔允漷等，2023）。跨学科主题学习具有学科性、关联性、实践性、探究性和多样性等特征（张涛，2024），可以有效改变当前小学教育中分科教学各自为政、互不关联的状态，让学习者达到最佳学习效果，培养自主解决问题的能力。同时，《义务教育课程方案（2022 年版）》明确要求设置“跨学科主题”学习活动，占本学科总课时不少于 10%，强化学科间的相互关联，增强实践性，以带动课程综合化实施。在小学跨学科主题学习活动中使用游戏化学习策略，有助于打破学科壁垒，引导学习者使用批判性思维看待问题，学会整合信息并构建知识体系，最终将理论知识应用于实际情境中。由此可见，游戏化学习的性质与跨学科主题学习特征具有一致性，使用游戏化策略可以有效解决跨学科主题学习存在的局限性和挑战。然而，小学教师没有意识到游戏化学习策略与跨学科主题学习的适配性，尚未在跨学科主题学习活动中广泛使用游戏化学习策略。

基于此背景，本研究试图设计一项跨学科主题学习活动，使用增强现实技术开展游戏化学习，通过观察学习者活动表现、收集过程性数据等方法，以质性和量化结合的方式来评估学习者的学习体验和学习成效。因此，本研究确定如下研究问题：

（1）融入增强现实和游戏化策略的跨学科主题学习活动中，学生的学习体验如何？

（2）融入增强现实和游戏化策略的跨学科主题学习活动中，学生的学习成效如何？

（3）在设计和实施融入增强现实技术和游戏化策略的跨学科主题学习时，有哪些可行的策略？

2.研究设计

2.1 研究对象与实施情境

本研究在中国北京市某小学开展教学，研究对象为六年级 23 名选修科学天文选修课的学生。活动开展形式为线下，共持续三周，每周活动时长为 60 分钟。活动开展思路借鉴美国生物学课程研究（BSCS）采用建构主义教学模式所开发的 5E 教学模式（李艳晖等，2024），大致分为：教师导入吸引兴趣；学生小组合作探究体验；教师解释原理，学生分享所学；知识迁移，学生制作模型作品；作品展示，多元评价，教师总结五个阶段。实施流程图见图 1。

图 1 实施流程图

2.2 研究工具

本研究用于评估学习者学习体验和学习成效的工具主要有：太阳系知识前测与后测题目、游戏化学习接受度量表、访谈、模型作品评价量规。

太阳系知识前测问卷共 5 道主观题，用于了解学生对太阳系知识的认知水平；后测问卷分为两部分，第一部分为客观选择题，第二部分为主观填空题。客观选择题主要针对太阳系八大行星和地球内部结构的特点编制，考察学生的掌握程度。主观填空题围绕太阳系八大行星的名称与距离太阳的远近进行考察。前测与后测问卷满分均为 40 分。

游戏化学习接受度量表基于技术接受度量表和增强现实技术应用态度问卷进行编制（Davis,1989;Küçük et al.,2014），包含感知有用性、感知易用性、满意度与使用意愿四个维度，所参考量表问卷 Cronbach's alpha 值为 0.936，信度较好。问卷采用李克特五级量表形式，共 15 道题目，均为态度倾向型，分为非常同意、比较同意、中立、不同意、非常不同意五个等级。该量表于活动结束之后填写。

访谈于活动及问卷填写后进行，根据学生的活动表现和体验状态选取部分同学进行。访谈内容与问卷类似，主要希望了解学生在跨学科主题学习中体验增强现实技术支持的游戏化学习的感受与想法，了解其对于 AR 的技术接受度、以及在体验学习中遇到的问题等。

为了了解学习者能否将在游戏化学习过程中学习到的知识内容进行迁移，活动设置了制作太阳系八大行星模型环节，并制定作品评价量规用于评定学生的作品完成情况。量规从科学、数学、美术、道德四个学科维度进行制定，满分为 40 分。量规与评价标准见表 1。

表 1 太阳系八大模型评价量规与评分标准

评价维度	**评价内容**	**评分标准**		
		0 分	**1-3 分**	**4-5 分**
科学(A)	A1.行星外观符合科学事实；	模型成品与太阳系八大行星无关。	能够制作八大行星内容但存在错误。	能够制作完整准确的八大行星模型。
	A2.能够基于科学事实进行创新。	无创新行为。	有创新行为但不符合科学事实。	能够基于科学事实进行创新创作。
数学(B)	B1.行星半径、体积符合真实比例； B2.行星日距恰当。	模型比例完全错误。 日距完全不符合科学事实。	模型比例部分正确。 日距较为恰当。	模型比例完全正确。 日距完全恰当。
美术(C)	C1.模型作品外观整洁美观； C2.模型作品能够体现审美意识。	外观肮脏凌乱。 无审美意识体现。	外观较为整洁美观。 审美意识体现	外观非常整洁美观。 审美意识体现充分。
道德(D)	D1.制作过程中不浪费实验材料； D2.制作结束后保持桌面整洁。	过度浪费实验材料。 未收拾桌面。	具有节约意识但存在改进空间。 有收拾桌面意识但清理不到位。	能够做到完全不浪费实验材料。 能够主动将桌面收拾整洁。

2.3 研究方法与数据收集分析

研究采用准实验研究方法，运用描述性统计分析方法和内容分析法，通过太阳系知识问卷前后测情况与模型作品评价学习者的学习成效；依据游戏化学习接受度问卷数据和访谈数据分析学习者学习体验。

太阳系知识前测、后测问卷及游戏化学习接受度量表问卷各发放 23 份，回收 23 份，有效问卷 23 份。实验结束后，使用 SPSS 27.0 以及 RStudio 1.5.57 对实验数据进行分析。

3.学科融合与游戏化策略框架设计

本研究中的跨学科主题学习活动以小学科学、数学、美术、信息科技和道德与法治五个学科为基础，结合学科知识围绕活动主题选取相应教学内容，并针对不同教学内容进行游戏化策略设计，具体内容如表 2 所示。

表 2 跨学科学习与游戏化策略融合框架

学科	教学内容	游戏化策略设计
科学	1.知晓太阳系中八大行星的名称，准确描述其在太阳系中的位置； 2.了解八大行星的外观特点、物理属性、日距。	1.*宇宙探险家*：结合任务单要求，使用 AR 设备尽可能多地搜集与行星相关的内容，并将信息填写在任务单相应位置； 2.*行星排位赛*：根据行星外观特点辨认行星，并按照其在太阳系中的位置进行准确摆放。
数学	1.依据科学事实，对行星半径、体积、日距等信息进行排序； 2.根据科学事实，结合实物模型材料限制，合理缩小行星体积、日距，能够准确在模型作品中表现行星的体积关系和日距远近。	1.*太阳系小判官*：在规定时间内，依据行星卡片上的信息，按照体积、半径或日距等信息进行准确排序； 2.*先思后行我在行*：拿到实物模型材料后，对材料进行合理划分，运用比例尺等工具初步拟定各行星的体积，在瓦楞纸板上设计摆放的位置。
美术	1.依据科学事实，选用与行星外观特点匹配的彩泥、颜料等手工材料进行模型制作； 2.能够在尊重科学事实的基础上，在模型作品中进行创意创新，并体现审美意识。	*指尖上的太阳系*： （1）能够制作以太阳为中心、由八颗行星组成的太阳系模型，准确呈现太阳系大致结构； （2）尽可能还原行星各自最显著的外观特点，按比例呈现体积、半径、日距等关系； （3）在作品中体现审美意识与创意创新想法。
信息科技	1.掌握 AR 技术的使用方法，能够利用设备与互联网搜寻所需信息； 2.能够使用计算机、手机或 Pad 设备上传、查看本小组的作品，具备使用网络进行交流的能力。	*作品巧思听我讲*：模型作品制作完成后，拍照上传至班级社群，每小组派一名代表结合图片介绍本组作品，并着重讲解作品中蕴含的创新想法。班级内进行自评、组间互评、教师评价活动，综合评定模型作品完成情况。
道德与法治	1.在制作模型作品时，能够具备勤俭节约意识，做到不浪费实验材料； 2.课程结束后，能够保持座位周围的整洁，自主收拾实验材料。	*我是环保小达人*：以小组为单位，将剩余材料进行整理、归还，共同收拾桌面及座位周边卫生，并开展评比活动，对照优秀组检查本组情况进行改进。

4.研究结果

4.1 学生知识技能提升明显

本研究依据太阳系知识前测和后测问卷及太阳系八大行星模型制作水平来衡量学生的学习成效。

太阳系知识前测和后测分数分析结果如图 2 所示。Mann-Whitney U 检验显示，实验前后班级的测验成绩有显著变化（Z=-2.561，p=0.01），后测的平均成绩（均值=27.87）高于前测的平均成绩（均值=22.59）。图 2 中的误差棒为标准差（SD），实验后的标准差（SD=4.42）低于实验前（SD=10.69），说明实验后班级的测验成绩更加稳定，学生间成绩差异减小。

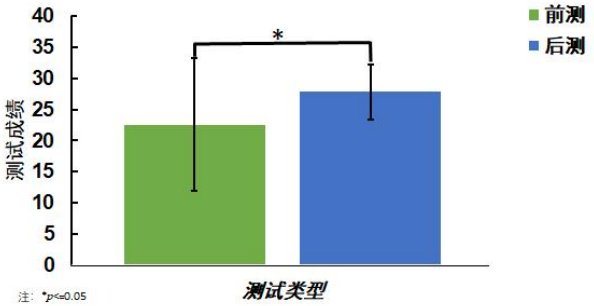


图 2 太阳系知识问卷前后测分数柱状图

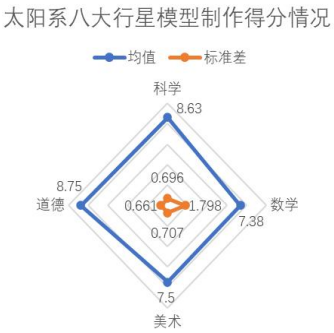


图 3 太阳系八大行星模型制作得分情况

学习者以小组合作的形式在第二课时和第三课时完成太阳系八大行星模型制作，共上交模型作品 8 份。依据评价量规对作品进行评价，评价结果如图 3 所示。学习者在道德角度践行节约意识和保持桌面整洁方面得分最高。其次是在科学角度依据科学事实模拟制作行星方面有良好表现，均值 8.63 分，说明学习者在学习行星外观、特点等内容时留下了较为深刻的印象，并且能够在制作活动中将所学知识进行迁移运用。学习者在美术角度和数学角度得分稍低，但是在作品中能够体现学习者在游戏化学习中所学内容。虽然没有按比例摆放行星位置，未能体现行星间体积关系和日距，但是学生上交的作品表明学习者能够结合在游戏化学习中体验到的场景进行自主创作，可以看出学习者已经对八大行星具备了一定知识储备。至于得分较低的原因，可能是因为跨学科主题学习活动中游戏化策略设计偏向于检验学习者的学习成效和已有知识经验在新情境中的调用，没有将相应知识与技能纳入游戏环节中，忽略学生先前知识储备差异。以至于有部分学生在制作模型时，不了解如何按比例缩小行星半径、体积或日距，仅将注意力和精力放在还原行星外观特点和进行创意创新角度，导致数学维度平均得分较低。

因此，本研究认为通过学习者制作的八大行星模型作品，可以表明其在体验了增强现实技术支持的游戏化学习后，对太阳系相关科学知识掌握程度有所提升，能够将所学理论知识

迁移到模型制作当中，创作出还原度较高的创意物化作品，具有明显的学习成效。

综上所述，班级后测成绩与前测成绩相比有明显提高，说明在小学跨学科主题学习中使用游戏化学习策略有助于学生的知识获得；学生模型制作各维度得分较高，说明学习者在接受游戏化学习策略后已经具备将理论知识进行迁移的技能。因此，本研究认为在小学跨学科学习中使用游戏化学习策略可以有效提升学习者的知识技能。

4.2 学生学习体验积极

学习者于活动结束后填写游戏化学习接受度量表问卷,反馈结果如图 4 所示。

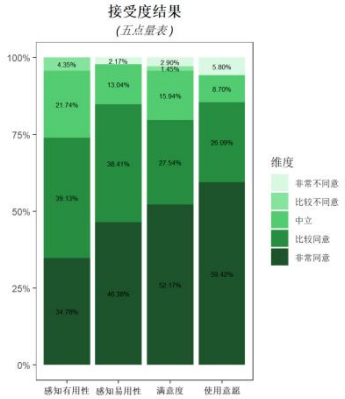


图 4 游戏化学习接受度百分比堆积柱状图（N=23）

由图 4 可知，73.91%的学生在游戏化学习有用性维度选择同意和非常同意选项，认可 AR 技术在活动过程中发挥的作用，并在访谈中表示“*能自己操控，想看哪一部分就可以自己去动*”，认为可操纵性能够帮助其细致观察天体关系；84.79%的学生在易用性维度选择同意和非常同意选项，认为 AR 游戏容易上手，使用起来较为便捷，但也有少数学生在访谈中表示“*有一些键和按钮用不明白*”，这说明教师需要在学生使用前对游戏进行机制和规则的讲解；79.71%的学生在满意度维度选择同意和非常同意选项，说明学生满意在跨学科主题学习中使用游戏化学习，并在访谈中表示“*期待在其它的科目学习中也能使用融入 AR 技术的游戏化学习*”，侧面印证了在跨学科主题学习中融入 AR 技术和游戏化策略的成功；85.51%的学生在使用意愿维度选择同意和非常同意选项，所占比例最高，说明学生对于在跨学科主题学习中使用 AR 技术和游戏化学习具有强烈的使用意愿，在访谈中表示“*AR 技术能让我自己去控制它，我想看哪儿就可以点哪儿*”，表明 AR 技术的特性在提升学生使用意愿方面发挥的作用较为明显，对于提升学习者的学习体验具有一定帮助。

综上，游戏化学习接受度量表在各个维度均具有较高得分，说明学生认可 AR 支持的游戏化学习用于跨学科主题学习的价值，对游戏化学习的接受度较高，并在学习过程获得了积极的学习体验。

5.教学设计与实施策略

为了增强 AR 技术和游戏化策略在跨学科主题学习中的可用性和有效性，现将结合实践经验提出相关策略。

5.1 围绕教学内容设立脚手架，引导学生带着问题开展探索

活动开展过程中，学习者存在缺少“抓手”这一问题。游戏化环节的设置，增添了竞赛韵味，但是导致了学习者盲目搜寻信息以量取胜，缺少信息整合的环节。游戏化策略设计给学习者呈现了较为丰富、开放的学习内容，学习者很有可能无法抓住学习的重点，发生漫无目的地学习行为，从而削弱学习成效。游戏化学习理论认为学习的核心是问题的发现与核心

问题的引领（庄惠芬，2020）。因此，教师在进行教学设计时应结合教学目标设计能够激发学习者思考的环节或问题，引导学习者在开始体验前明确本节课需要达成的目标和完成的任务。教师在课前聚焦于课程重难点内容，基于学习者的认知规律来精心设计学习单，有助于发挥脚手架作用，激发学生的自主学习意识，为学生的学习指明方向（张明华，2020）。对于跨学科主题学习而言，可以围绕学科主题设计学习单，也可以围绕学习内容进行设计，需要注意的是需要综合考虑纳入学习单的任务数量，避免增加学习者的认知负荷，从而降低学习成效。

5.2 善用游戏技巧与游戏情境，以学习动机促进学习成效的提升

从学生的心理动机出发，掌握并调动其内在的求知欲是游戏化教学成功的关键之一，教师需要站在学习者视角思考问题，关注学习者的认知和心理需要（严加平，2018）。因此，教师需要在进行教学设计前充分了解学情，确保设计的教学内容是学生尚未掌握的、感兴趣的，并综合考虑学生之间的认知差异，在教学过程中体现难度差异，力求让每位学习者都能有所收获。对于较为枯燥、抽象的学习内容，教师应思索其与游戏化学习之间的关系，创设真实贴切的情境，帮助学习者在游戏化学习与知识内容间建立练习，从而促进知识学习的内化。同时，教师可以利用穿插游戏环节的形式刺激学生的学习兴趣，激发学习动机。但需要注意游戏的难度设置应该适中，避免过于简单无法激发学生调动高阶思维进行深入思考，或过于困难而导致学生感到挫败。在跨学科主题学习中，教师应注意帮助学生建立游戏间与学科间的联系，尽可能以“进阶”、“升级”等形式将游戏环节串联，让学习者获得连贯的体验感，有助于建立系统的知识体系。

5.3 关注学习者课堂生成，构建游戏化学习反馈系统

为了了解学习者的学习情况，教师应关注学习者的课堂生成，并将其作为调整教学过程的重要依据。同时，作为学习过程中的重要组成部分，反馈是学生改善学习体验的关键要素（陈鹏等，2024）。在游戏化学习反馈系统中，学生的表现和答题结果通常会得到系统的即时反馈，以便学生及时了解自己的学习进展和不足之处。教师应积极利用反馈激发学生学习兴趣，帮助其了解学习进度，并在课堂的不同环节充分利用游戏化学习反馈。在课前，教师可以将游戏化学习反馈系统作为预习热身工具，通过设置趣味性的任务和挑战，提高学生学习积极性。在课中，教师可以将其作为课堂互动工具，活跃课堂氛围，提高学生的参与度。在课后，教师还可以将其作为单元测验的工具，利用游戏化元素设计练习题或挑战任务，检测学生对知识的掌握情况，及时调整教学策略（王瑾等，2024）。此外，教师还需要根据学习者的课堂表现和知识习得情况，及时调整反馈系统，以保证游戏化学习反馈系统对学习成绩的积极影响，并持学生的学习兴趣与参与度。

5.4 重视学习者游戏经验差异，善用新兴技术辅助教学

伴随着网络的发展和智能移动终端的普及，游戏化策略在提高在线教育和网络学习效果中的作用备受关注，游戏化学习成为实现高质量在线教育的重要解决方案（关京，2021）。已有研究表明，游戏经验差异在学生群体中普遍存在，部分学习者可能由于自身游戏经验的不足导致无法适应游戏化学习环境，进而阻碍学习体验与成效（王鉴羽等，2024）。因此，教师需要预先了解学习者的游戏经验情况与对新兴技术的了解程度，以便开展个性化教学与辅导。其次，教师应该考虑不同学习类型学生的不同内在需要，设计具有竞争性、合作性等不同类型的活动，匹配不同的奖励方式，充分满足不同年龄阶段学生的特点，激发不同学习者的兴趣（Kapp,2016）。同时，教师应贯彻以学习者为中心的原则，在充分了解学情的基础上因地制宜选择游戏化策略。此外，虚拟现实技术、大语言模型等新兴技术的迅速发展，

有助于给学习者提供更具有沉浸感的体验学习，教师应选取恰当的新兴技术辅助教学。

6.结论与展望

在信息化时代，游戏化学习具有激发学生学习动机、发展学生的认知能力以及促进学生积极参与等优势，逐渐成为教育者关注的新兴领域。本研究采用准实验研究法，在小学跨学科主题学习中融入增强现实技术与游戏化策略，评估其对学习者学习体验和学习成效的影响。研究结果显示，在小学跨学科主题学习中使用游戏化策略可以有效提升学习者的学习成效，并给学习者带来积极的学习体验。同时，研究针对实践中存在的问题进行反思，结合实践经验总结在跨学科主题学习中更有效融入增强现实技术和游戏化学习的策略，对于教育领域的学者和教师具有一定借鉴意义。此外，本次跨学科主题学习实践活动也存在一定的局限，例如进行游戏化学习干预的周期较短且缺少控制组设计。在未来的教学实践中，会设计更长周期的干预研究，考虑加入控制组来进行对照，以进一步评估融入增强现实技术和游戏化策略在跨学科主题学习中的应用价值，并探究游戏化学习支持的小学跨学科系列课程。同时，在前测设计中忽略了学习者初始学习能力对测试结果产生的影响，未来将在实验设计中考虑这一因素所发挥的影响作用，进而在分析结果时能够更加精准地看到前测对后测产生的影响。

参考文献

王鉴羽和王雪(2024)。学习者游戏经验差异影响游戏化学习的机制与对策——基于大理论视角的研究。电化教育研究，45 (04), 59-66。https://doi:10.13811/j.cnki.eer.2024.04.009.

王瑾、胡玥和范文翔(2024)。游戏化学习反馈系统真的能提升学生成绩和改善情感态度吗？——基于 37 项 Kahoot!相关实验和准实验研究的元分析。现代教育技术，34 (12), 76-85。

关京(2021)。游戏化学习对于教育的启示。中小学心理健康教育，(26), 4-7。

庄惠芬(2020)。小学数学游戏化学习策略的建构。教学与管理，(32), 58-59。

李艳晖、刘莉妹、石瑛和赵晓燕(2023)。科学探究环境下 5E 教学模式的应用。中学生物教学，(08), 19-22。

张明华(2020)。数学主题式学习单设计策略。教学与管理，(26), 58-59。

张涛(2024)。跨学科主题学习的概念意蕴、实践问题及改进路径。现代基础教育研究，55 (03), 109-114。

严加平(2018)。什么内容适合采用游戏化方式来学习——基于对游戏本质、知识分类及儿童学习理解的思考。上海教育科研，(08),23-28。https://doi:10.16194/j.cnki.31-1059/g4.2018.08.007.

陈鹏和白佳鹭(2024)。学习反馈影响学生学习投入的元分析研究。黑龙江高教研究，42 (08), 139-146。https://doi:10.19903/j.cnki.cn23-1074/g.2024.08.018.

武法提、黄石华和殷宝媛(2018)。场景化:学习服务设计的新思路。电化教育研究，39 (12), 63-69。https://doi:10.13811/j.cnki.eer.2018.12.009.

曹语苍(2023)。游戏化学习：国际实践与成果。现代教育，(09), 12-17+21。

崔允漷和郭洪瑞(2023)。跨学科主题学习：课程话语自主建构的一种尝试。教育研究，44 (10), 44-53。

Davis, F. D. . (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Mis Quarterly, 13*(3), 319-340.

Kapp, K. M. . (2016). Choose Your Level : Using Games and Gamification to Create Personalized Instruction. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handbook on personalized learning for states, districts, and schools* (pp. 131 – 143). Philadelphia, PA:Temple University, Center on Innovations in Learning.

Küçük, Sevda, Yılmaz, Rabia M, Baydaş, Özlem, & Göktaş, Yüksel. (2014). Augmented reality applications attitude scale in secondary schools: validity and reliability study. *Education and Science, 39*(176), 383-392.

Sümer, M., & Vaněček, D. (2024). A systematic review of virtual and augmented realities in higher education: Trends and issues. *Innovations in Education and Teaching International*, 1-12. https://doi.org/10.1080/14703297.2024.2382854

生成式人工智能支持的写作指导对学生学习意愿的影响机制研究

Research on the Impact Mechanism of Writing Instruction Supported by Generative Artificial Intelligence on Students' Learning Willingness

徐明芮[1*]，李建伟[2]，李凌宇[3]
[1, 2]北京邮电大学人文学院
[3]北京建筑大学智能科学与技术学院
[*] xumingrui2023@bupt.edu.cn

【摘要】 随着生成式人工智能的发展，其在写作教育中的应用受到广泛关注。本文旨在构建并验证理论模型，探讨生成式人工智能支持的写作指导如何影响学生的学习意愿。基于对中国某高校学生的问卷调查数据，采用结构方程模型分析，研究重点考察了期望绩效、期望努力、感知偶然性与感知拟人性对学习意愿的影响机制。结果表明，生成式人工智能能够提高学生的学习意愿。其中，期望绩效和期望努力在感知偶然性与学习意愿之间起到中介作用；而感知拟人性对期望努力与学习意愿具有负向调节作用。基于此，研究提出相关教育启示，为人机协同学习模式提供参考。

【关键词】 生成式人工智能；写作；学习意愿；影响机制

Abstract: *With the rapid development of generative AI, its role in writing education has garnered widespread attention. This study constructs and validates a theoretical model to examine how AI-supported writing instruction affects students' learning willingness. Using survey data from Chinese university students analyzed via structural equation modeling, this study investigates the impact of performance expectancy, effort expectancy, perceived randomness, and perceived anthropomorphism on learning willingness. The findings reveal that generative AI significantly enhances learning willingness, with performance and effort expectancies mediating the effect of perceived randomness, while perceived anthropomorphism negatively moderates the relationship between effort expectancy and learning willingness. Based on these findings, the study proposes relevant educational implications and offers insights for the development of human-machine collaborative learning models.*

Keywords: generative artificial intelligence, writing, learning willingness, impact mechanism

1.前言

在高等教育中，写作不仅是一项基本技能，更是培养学生批判性思维和创新能力的重要载体（Bai et al., 2024）。在中国的教育体系中，写作指导被视为提高学生综合素养的关键，对于提升学生写作水平与学习动力至关重要。然而，目前市面上如 Grammarly、Ginger 等写作批改工具主要聚焦于语法、拼写等基础错误的纠正，忽视了对文章结构、语篇连贯性以及情感表达等更高阶写作技能的指导（Godwin et al., 2024）。这不仅制约了学生构建逻辑清晰、内容丰富文章的能力，也在一定程度上削弱其对写作学习的兴趣和积极性。

生成式人工智能（Generative Artificial Intelligence, GAI）的不断发展和应用为写作指导注入了新的活力。借助个性化反馈、智能提示和情境模拟等功能（Pesovski et al., 2024），GAI

可以提供更为全面和精准的写作支持，不仅弥补了传统工具在高阶写作技能指导上的短板，也为激发学生的学习动机提供了全新的可能性（赵寰宇和唐文娟，2024）。

尽管已有研究探讨了 GAI 在写作教学中的应用，但关于其如何通过内在机制影响学生学习意愿的研究仍然较为匮乏。因此，本文旨在构建理论模型，探讨 GAI 支持的写作指导对学生学习意愿的影响机制，从而为高等教育写作教学的实践与理论发展提供新的视角和实证依据。

2.文献综述

近年来，GAI 在教育领域的应用日益广泛，其在写作中的应用正逐渐引起学界关注。以 ChatGPT 为代表的对话型人工智能，凭借其强大的自然语言处理能力，正逐步渗透到教学实践中。具体而言，GAI 通过对大量文本数据和专业知识的学习，已在作文教学（夏静芳和赖伊玲，2025）、写作辅导（Chan, 2025）以及作文评估（黄子芹等人，2025）等多个方面展现出较好的应用成效。此外，GAI 能够根据学生个性化需求提供定制化学习资源和实时反馈，显著提升学生的学习体验和教学效果（Pesovski et al., 2024）。

然而，尽管 GAI 在写作领域展现出巨大潜力，关于其在写作教育中系统部署与评估的研究仍较为稀缺，尤其是在探讨如何通过改善写作指导进而提升学生学习意愿的内在机制方面的探讨尚不充分。部分研究尝试借鉴人工智能技术接受模型（如 UTAUT 及其扩展模型），探讨技术易用性、绩效预期和社交影响等因素在技术采纳中的作用（Gansser & Reich, 2021）。已有研究显示，这些模型在部分教育领域成功解释了学生对新技术的接受与使用行为（Faqih & Jaradat, 2021），但将其应用于写作指导领域尚处于初步探索阶段。

因此，本研究旨在构建理论模型，探讨 GAI 支持的写作指导对学生学习意愿的影响机制，从而为高等教育写作教学提供新的视角和实证依据。

3. 研究问题与假设

本研究聚焦于写作教育，旨在探讨 GAI 支持的写作指导如何影响学生的学习意愿。基于前期文献综述，本研究构建了一个包含五个关键构念的理论模型：感知偶然性、感知拟人性、期望绩效、期望努力与学习意愿（见图 1）。

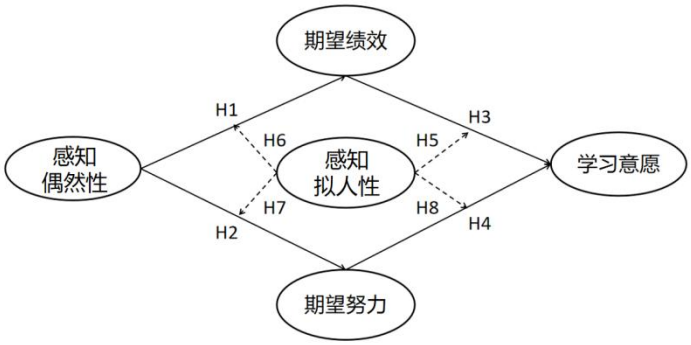


图 1 假设模型

在该模型中，感知偶然性指学习者在与系统互动过程中，对系统所提供的超出预期且具有偶然性回答的感知程度（Go & Sundar, 2019）；感知人性性用以衡量学习者对系统类人化程度的感知（Chung et al., 2023）。期望绩效反映学习者对人机协同学习系统有效性的期望，包括更高的学习成效和更少时间投入（Zhai et al., 2017）；期望努力则指学习者对系统易用

性及操作便捷性的感知（Gansser & Reich, 2021）；学习意愿则定义为学习者主动使用教育技术进行学习的意愿。

3.1. 感知偶然性与期望绩效和期望努力的关系

期望违背理论为感知偶然性对人机协同学习影响的解释提供了理论支持。该理论认为，期望之外的变化能引发生理唤醒反应，从而增强个体的兴趣和注意力（梁哲浩和陈涛，2023）。随着智能代理和聊天机器人等技术的广泛应用，偶然性在人机交互中发挥了重要作用，促使学习者参与多轮双向交流，从而形成主动互动模式（Go & Sundar, 2019）。

基于该理论，当学习者与 GAI 支持的写作指导系统互动时，感知到的偶然性可能引发不同于预期的反馈。这种反馈可能激发学习者的兴趣，增强其注意力，从而更容易投入到学习中。因此，提出以下假设：

H1：在 GAI 支持的写作指导中，感知偶然性正向影响学习者的期望绩效。

H2：在 GAI 支持的写作指导中，感知偶然性正向影响学习者的期望努力。

3.2. 期望绩效和期望努力的中介效应

根据人工智能技术接受模型（Gansser & Reich, 2021），人工智能教育产品的易用性和对学习绩效的提升作用是决定学习意愿的关键因素。GAI 的发展为教育教学创造了一个新的数字环境，系统的易用性和对学习效果的提升作用直接影响学生的学习意愿。在这一环境下，学习者更加关注系统的易用性，并期望通过使用该系统提高学习绩效。因此，提出以下假设：

H3: 在 GAI 支持的写作指导中，期望绩效在感知偶然性与学习意愿之间起到中介作用。

H4: 在 GAI 支持的写作指导中，期望努力在感知偶然性与学习意愿之间起到中介作用。

3.3. 感知拟人性的调节机制

根据拟人化理论，拟人化互动能够模拟真实的教学情境，满足学习者对社交互动的基本需求（Baumeister & Leary, 1995）。GAI 支持的写作指导系统通过模拟不同教学风格的教师，为学习者创造了更加类人化的学习互动体验。随着人工智能技术的不断进步，这种类人化的人机互动模式不仅提高了学习者的参与度和动力，还通过提供个性化反馈、情感支持等方式进一步增强了学习动机和学习效果。因此提出以下假设：

H5：在 GAI 支持的写作指导中，感知拟人性正向调节期望绩效与学习意愿间的关系

H6：在 GAI 支持的写作指导中，感知拟人性正向调节感知偶然性与期望绩效间的关系。

H7：在 GAI 支持的写作指导中，感知拟人性正向调节感知偶然性与期望努力间的关系。

H8：在 GAI 支持的写作指导中，感知拟人性正向调节期望努力与学习意愿间的关系。

4. 研究设计

4.1. 研究对象

本研究的研究对象为中国某高校的学生，共计 30 人，其中包含 18 名男生和 12 名女生。

4.2. 研究工具

4.2.1. GAI 支持的写作指导系统

本研究使用团队自研的 GAI 支持的写作指导系统。该系统融合 GAI 技术，通过知识库和提示工程技术的整合，为学生提供个性化且即时的写作指导与反馈。系统的设计与实现如下：

（1）知识库的构建：系统通过构建写作知识库，为学生提供丰富的写作素材与知识支持。

（2）提示工程技术：系统采用分阶段的多层次提示策略，为学生提供精准、个性化的写作指导与反馈。首先，系统模拟资深写作专家和教育工作者，根据作文内容和评分标准，利用零样本提示技术生成初步写作框架，帮助学生迅速构建写作思路；随后，系统运用思维链

提示技术将写作任务拆分为一系列启发性问题，引导学生逐步深化思考；在输出最终反馈前，系统通过自我一致性提示技术并行生成多个建议，并筛选出逻辑性和有效性最佳的方案，确保反馈的准确性和针对性。最终，在学生完成作文后，系统结合评分标准对作文进行综合评价，并提供改进建议。

4.2.2. 调查问卷

研究问卷采用李克特五点量表（1-5，表示“非常不同意”至“非常同意”），分数越高表示对应维度的程度越强。其中，学习意愿量表参考人机协同学习意愿的心理需求研究量表（Jahn et al., 2021）；期望绩效量表依据学习者对翻转课堂模式的感知价值量表（Zhai et al., 2017）编纂；期望努力量表参考人工智能接受度研究中的期望绩效量表（Gansser & Reich, 2021）设计；感知拟人性量表参考人机协同场景下社交机器人的人性化程度量表（Chung et al., 2023）；感知偶然性量表依据 Go & Sundar（2019）的感知偶发性量表编制。

4.3. 研究流程

研究主要经历了三个阶段：前测、教学实施、后测。设计流程如图 2 所示。

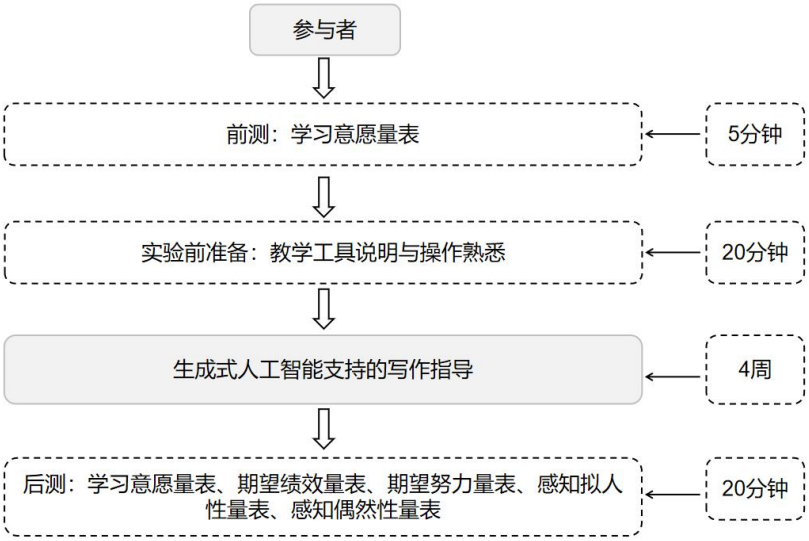


图 2 研究流程

（1）前测：这一阶段要求所有参与者在实验开始前完成学习意愿问卷，耗时 5 分钟。

（2）实验前准备：在正式实验开始之前，对本研究中使用的 GAI 支持的写作指导系统进行简要培训，介绍注册登录、获取反馈、修改、提交作文等基本操作，耗时约 10 分钟。随后给予参与者 10 分钟的自主探索和练习时间，期间研究者将现场提供技术支持与答疑。

（3）教学实施：教学周期为 4 周，学生使用团队自研的 GAI 支持的写作指导系统进行写作的学习与练习。

（4）后测：教学结束后，所有参与者需完成后测问卷，包括学习意愿量表、期望绩效量表、期望努力量表、感知拟人性量表、感知偶然性量表，共计 20 分钟。

4.4. 数据分析方法

本研究使用 SPSS v26.0 和 SmartPLS v4.0 作为数据处理与分析的主要工具。

（1）数据预处理：使用 SPSS 进行初始数据清理和预处理,剔除具有明显规律性且答题时间过短的问卷。

（2）信效度分析：信度方面，利用 Crobach’s 系数评估各量表的内部一致性；效度方面，使用 Kaiser-Meyer-Olkin（KMO）值和 Bartlett 球形度检验评估数据是否适合进行因子分析。

（3）描述性统计分析：使用 SPSS 的描述性统计分析对前测和后测中的学习意愿进行平均值及标准差的描述性统计，以了解数据的基本分布特征。

（4）独立样本 t 检验：采用独立样本 t 检验比较指导前后学生在学习意愿上的差异。

（5）结构模型方程分析：利用 SmartPLS 分析期望绩效、期望努力、感知拟人性、感知偶然性与学习意愿之间的关系。

5.研究结果

5.1. 信效度分析

本节对学习意愿量表、期望绩效量表、期望努力量表、感知拟人性量表与感知偶然性量表的信度与效度进行了检验，结果见表 1 和表 2。

各量表的 Cronbach's alpha 均超过 0.8，表明量表具有极高的内部一致性。量表的 KMO 值均超过 0.6，且 Bartlett 球形度检验的 p 值显著（$p < 0.05$），验证了因子分析的适用性。这些结果验证了量表的结构合理性，增强了研究的可信度和普遍性。

表 1 量表的信度分析结果

量表	Cronbach’s alpha	题项
学习意愿量表	0.901	4
期望绩效量表	0.836	4
期望努力量表	0.847	4
感知拟人性量表	0.883	4
感知偶然性量表	0.879	4

表 2 量表的效度分析结果

量表	KMO	Bartlett 球形度检验		
		近似卡方	自由度	显著性
期望绩效量表	0.674	39.741	6	.000
期望努力量表	0.668	27.961	6	.000
感知拟人性量表	0.787	88.689	6	.000
感知偶然性量表	0.722	69.225	6	.000
学习意愿量表	0.764	53.628	6	.000

5.2. 学习意愿

本研究采用描述性统计分析和独立样本 t 检验，对研究对象在学习意愿的前测与后测得分进行了比较，结果如表 3 所示。分析发现，学生在整体学习意愿得分上有所提升（从前测的 3.66 提高到后测的 4.34）。进一步分析发现，GAI 支持的写作指导对学生学习意愿的影响显著（$p<0.05$）。这些结果表明，GAI 支持的写作指导可以显著提高学生学习意愿。

表 3 学习意愿前后测得分情况

	前测	后测	T	P
学习意愿	3.66±0.287	4.34±0.190	-8.170	.000

5.3. 结构模型分析

本研究利用 SmartPLS 建立偏最小二乘结构方程模型，分析学生对于 GAI 支持的的写作指导的学习意愿，运用包含 3000 个子样本和 95%置信区间的偏差矫正自举法评估全样本（N=30）结构模型假设关系，结果见表 4。

结果显示，感知偶然性对期望绩效（$\beta=0.285$，$p<0.05$）、期望努力（$\beta=0.288$，$p<0.05$）存在显著正向影响，表明 H1、H2 成立。期望绩效（$\beta=0.425$， $p<0.05$）与期望努力

（β=0.230，p<0.05）对学习意愿存在显著正向影响，表明 H3、H4 成立。这说明感知偶然性对学习意用的作用受期望绩效与期望努力的中介影响。

让人意外的是，感知拟人性在感知偶然性与期望绩效、感知偶然性与期望努力、期望绩效与学习意愿中的调节作用均不显著（p>0.05），表明 H5、H6、H7 不成立。但感知拟人性在期望努力和学习意愿之间（β=-0.160，p<0.05）存在负向调节作用，表明 H8 成立。

表 4 假设系数、显著性检验及假设判断结果

假设	系数	标准偏差	P	95%置信区间	假设判断
H1	0.285	0.154	0.014	[0.179，0.391]	成立
H2	0.288	0.157	0.039	[0.179，0.397]	成立
H3	0.425	0.063	0.018	[0.318，0.532]	成立
H4	0.230	0.149	0.035	[0.133，0.427]	成立
H5	0.089	0.065	0.472	[−0.038，0.216]	不成立
H6	0.075	0.064	0.163	[0.042，0.392]	不成立
H7	0.090	0.068	0.259	[−0.041，0.221]	不成立
H8	−0.160	0.057	0.003	[−0.048，-0.272]	成立

6.讨论

6.1. GAI 对学习意愿的影响

研究结果表明，GAI 支持的写作指导能显著提升学生的学习意愿。这一发现与翟雪松等（2023）的研究结果一致。

在 GAI 支持的教学环境中，学生可以获得更具个性化、即时化的学习支持（Pesovski et al., 2024），帮助他们克服写作瓶颈，这是促进学生学习的重要动力。同时，GAI 能促进学生的自主反思，鼓励他们根据反馈调整学习策略（Noy & Zhang, 2023）。这种个性化、动态的学习体验增强了学生的内在动机（Bai et al., 2024），提升了他们的学习意愿。

6.2. 感知偶然性通过期望绩效和期望努力中介学习意愿

研究结果表明，学习者的感知偶然性在增强学习意愿上具有显著的影响，并通过期望绩效和期望努力起到中介作用。

首先，期望绩效在提高学习效率方面发挥着关键作用。学习者期望在有限的时间内获取更多信息（Zhai et al., 2017）。在信息爆炸的时代，传统的知识搜寻显得较为低效。GAI 通过交互式的信息检索和应用，精准地满足了学习者对高效学习的需求。因此，GAI 不仅增加了知识获取的偶然性，还通过高效的知识管理和整合策略显著提升了学生的学习意愿，并加强了他们对学习效率的感知。

其次，期望努力在这一过程中也展现了显著的中介效应。期望努力反映了学习者对系统易用性的感知。当学生感知到系统的操作简便且功能互补时，他们更有可能继续使用该系统进行学习（Gansser & Reich, 2021）。通过提升系统的易用性和功能整合，GAI 能有效降低学习者在使用过程中的心理负担，进一步增强其学习意愿，激发他们主动参与学习的动力。

因此，学习者的感知偶然性通过期望绩效和期望努力的中介作用，促进了学生学习意愿的提高。这一机制表明，GAI 在提升学习效率和易用性方面的优势，能够显著激发学生的学习动机，帮助他们实现更高效的学习体验。

6.3. 感知拟人性的负向调节机制

研究结果表明，学习者对GAI的拟人化感知仅负向调节期望努力与学习意愿之间的关系，对其他关系的调节作用并不明显。

对于感知拟人性负向调节期望努力与学习意愿，这一发现与“恐怖谷效应”的理论（Mori et al., 2012）一致,即当系统的类人化程度超过一定临界点后，可能引发学习者的不适感，从而降低其对系统易用性的正向预期。因此，增强虚拟教师的类人化程度反而负向调节期望努力和学习意愿的关系。

此外，在当前教学环境中，学生可通过丰富的线下互动获得必要的情感支持，因而对数字系统的类人化需求相对较低（翟雪松等人，2023）。这可能解释了为何拟人化在其他调节路径中未能展现出显著作用。未来研究应进一步探讨拟人化设计的最优程度，以平衡系统功能性与情感互动需求。

7.结论

7.1.研究结果与建议

7.1.1.研究结果

本研究聚焦于写作教育，探讨GAI对学生学习意愿的影响机制。研究结果表明，GAI能显著提高学习者的学习意愿。而且，期望绩效和期望努力是感知偶然性与学习意愿之间的中介变量；而感知拟人性对期望努力与学习意愿具有负向调节作用。

7.1.2.建议

基于此，对教育工作者提出以下建议：

（1）GAI技术在教学中的整合

教育工作者应考虑将GAI技术整合到课堂教学中。GAI能够根据学生的个体需求提供个性化和即时化的反馈，帮助学生在面对挑战时保持专注，激发内在动机，从而有效提升学习意愿和学习体验。

（2）促进AI与教师协同的教学模式

尽管GAI能提供类人化的学习支持，但过度的拟人化可能会降低学习者学习意愿。因此，教师在教育过程中仍然具有不可替代的重要作用。教育工作者应积极探索AI与教师有效协同的教学模式，以促进学生学习。

7.2.研究不足与改进

本研究在探讨GAI支持的写作指导对学生学习意愿的影响机制方面取得了一定成果，但仍存在一定的局限性。首先，本研究作为一项探索性研究，样本规模较小且来源局限于特定地区，这可能影响结果的普适性，无法全面代表不同背景学生的情况。其次，本研究主要关注短期教学效果，未对长期效果进行深入探讨。未来的研究应扩大样本规模和范围，探索不同教育背景下的效果，并对长期影响进行更深入的分析。

参考文献

蔡芬、贾枭和沈文钦(2025)。生成式人工智能在我国研究生学术写作中的应用现状及其影响。中国高教研究,1,75-82。

梁哲浩和陈涛(2023)。人工智能+公共服务中的决策困境——基于民众的感知视角。社会政策研究,1,118-136。

翟雪松、楚肖燕、焦丽珍、童兆平和李艳(2023)。基于“生成式人工智能+元宇宙”的人机协同学习模式研究。开放教育研究,29(5)，26-36。

赵寰宇和唐文娟(2024)。GenAI 赋能国际中文教育写作教学的可行性与潜在应用。长春大学学报,12,65-69。

Ahmad, K., Iqbal, W., El-Hassan, A., Qadir, J., Benhaddou, D., Ayyash, M., & Al-Fuqaha, A.I.(2024). Data-driven artificial intelligence in education: A comprehensive review. *IEEE Transactions on Learning Technologies,17*,12-31.

Allen, M., Naeem, U., & Gill, S. S. (2024). Q-Module-Bot: A generative AI-based question and answer bot for module teaching support. *IEEE Transactions on Education, 67(5)*,793-802.

Bai, S., Gonda, D.E., & Hew, K.F. (2024). Write-curate-verify: A case study of leveraging generative AI for scenario writing in scenario-based learning. *IEEE Transactions on Learning Technologies,17*,1313-1324.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin,117(3)*,497-529.

Cascella, M., Semeraro, F., Montomoli, J., Bellini, V., Piazza, O., & Bignami, E. (2024). The breakthrough of large language models release for medical applications: 1-year timeline and perspectives. *Journal of Medical Systems,48(1)*,22.

Chung，H.，Kang，H.，& Jun，S（2023). Verbal anthropomorphism design of social robots：Investigating users' privacy perception.*Computers in Human Behavior，142*,107640.

Cooper, M.M., & Klymkowsky, M.W. (2024). Let us not squander the affordances of LLMs for the sake of expedience: Using retrieval augmented generative AI chatbots to support and evaluate student reasoning. *Journal of Chemical Education,101(11)*, 4847-4856.

Gansser，O. A.，& Reich，C. S.（2021）. A new acceptance model for artificial intelligence with extensions to UTAUT2： An empirical study in three segments of application. *Technology in Society,65*,101535.

Go，E.，& Sundar，S. S.(2019). Humanizing chatbots： The effects of visual，identity and conversational cues on humanness perceptions. *Computers in Human Behavior,97*, 304-316.

Jahn, K., Kordyaka, B., Machulska, A., Eiler, T.J., Grünewald, A., Klucken, T., Brück, R., Gethmann, C.F., & Niehaves, B.(2021). Individualized gamification elements: The impact of avatar and feedback design on reuse intention. *Computers in Human Behavior,119*,106702.

Li, B., Bonk, C.J., Wang, C., & Kou, X. (2024). Reconceptualizing self-directed learning in the era of generative AI: An exploratory analysis of language learning. *IEEE Transactions on Learning Technologies,17*,1515-1529.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine,19(2*),98-100.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence.*Science,381(6654)*,187-192.

Pesovski, I., Santos, R., Henriques, R., & Trajkovik, V. (2024). Generative AI for customizable learning experiences. *Sustainability,16(7)*,3034.

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2016). Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology,37(1*),8-25.

Zhai, X., Gu, J., Liu, H., Liang, J., & Tsai, C. (2017). An experiential learning perspective on students' satisfaction model in a flipped classroom context. *Journal of Educational Technology & Society,20*,198-210.

基於時間序列概念圖之機器人數位說故事學習模式對於學生學習成效、批判思考能力與團體自我效能之影響

The effects of time sequence-oriented concept mapping for robot-based digital storytelling on students' learning effectiveness, critical thinking, and group self-efficacy

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】 隨著健康意識的提升，如何透過學校教育幫助青少年了解「預防癌症的飲食知識」已成為公共衛生領域的重要課題。過往護理課程透過機器人說故事學習模式來提升學習成效，但若無良好的學習引導策略，會導致不容易建立良好的知識脈絡。時間序列概念圖是一種以時間軸來組織知識和相關事件的心智工具，可用來作為機器人說故事的故事腳本設計大綱，讓故事情節可以在防癌飲食故事中按照時間序列逐漸發展推演。因此本研究在高中資訊課程中，設計一套基於時間序列概念圖之機器人數位說故事模式，幫助學生提升學習成效、批判思考和團體自我效能。

【关键词】 時間序列概念圖、機器人數位說故事、防癌飲食、自律學習、批判思考

Abstract: *In recent years, there has been a notable increase in cancer incidence, with the affected population becoming progressively younger. Consequently, there is a growing need to educate young students about the relationship between dietary habits and chronic diseases. In this study, a robot digital storytelling model based on a time sequence-oriented concept mapping was designed in the high school information course. This model is designed to facilitate the integration of knowledge networks and narrative scripts, allowing the storyline to unfold progressively according to temporal sequences. It is anticipated that this approach will enhance learning effectiveness, promote critical thinking, and strengthen group self-efficacy.*

Keywords: time sequence-oriented concept mapping, robot-based digital storytelling, cancer-preventing diet, self-regulated learning, critical thinking

1. 緒論

隨著健康意識的提升，如何透過有效的教學方式促進青少年對正確防癌飲食知識的理解與實踐，便成為健康教育領域的重要課題(Al-Hosni et al., 2024)。過往研究將自律學習(Self-regulated learning)用於健康護理課程中，讓學生根據任務分析並制定相對應的護理解決方案，最終評估自身的學習成果，從而提升學生的學習成效(Hohenhaus & Provost, 2024)。而機器人輔助數位說故事學習策略也被應用於住院兒童照護中(Chang et al., 2023)，不僅提升患者對嚴重疾病醫療過程的認知，還減輕焦慮感，但諸多研究表明在設計數位說故事時，需要學習鷹架或工具來幫助學生組織所學的知識概念，以有效提升學習成效。

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)。在 Chu 等人(2015)的研究中，基於時間序列概念圖用於歷史

課程中，透過事件時間的前後順序來編排彙整成完整的歷史知識脈絡，促進學生更容易理解事件的前因後果，進而有效提升學習成效。

因此，本研究設計一套基於時間序列概念圖之機器人數位說故事學習模式，並在結合自律學習和一般數位說故事的 PWUDE(計劃 Plan、撰寫故事 Write story、使用機器人 Use kebbi showtime、展示評估 Display and Evaluation)學習策略下，引導學生運用時間序列概念圖完成防癌飲食議題的故事設計，期望能提高學生的學習成就、溝通傾向、合作學習能力、批判思考與團體自我效能，並降低溝通焦慮。

2. 文獻探討

2.1. 機器人數位說故事

機器人數位說故事學習策略，是一種將知識概念設計成數位故事，根據故事設計機器人的動作、行為，讓學習者在機器人上進行播放學習知識的方法，提供更多元的體驗式學習，透過機器人提升注意力、趣味性、互動性和積極的學習體驗 (Chen et al., 2022)，在 Chang 等人 (2023) 學者的研究中，讓專業培訓師在糖尿病教案設計中，運用機器人數位說故事學習模式設計疾病的教學內容，提高學習者的學習成效、學習態度、批判思考能力和學習滿意度。在過往研究中，表明機器人數位說故事可以幫助學生在溝通傾向和溝通焦慮中產生正向影響 (Liang & Hwang, 2023)，但在使用機器人進行互動學習時，若缺乏相關支持，學生在學習過程中會面臨組織知識等的困難，表現出害怕恐懼，導致學習成效不如人意(Chen Hsieh, 2021)。

2.2. 時間序列概念圖

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 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，在 Chu 等人(2015)的研究中，基於時間序列概念圖用於歷史課程中，可以幫助學生以事件發生的時間順序來理解知識間的前因後果，有效幫助學生建立知識脈絡，從而提升學習成效。而在過往研究中，也表明概念圖在協作活動中可以提升學生的自我效能 (Chen, 2020)。

3. 研究方法

3.1. 學習模式設計-基於時間序列概念圖之機器人數位說故事學習模式

本研究的實驗組(EG)使用基於時間序列概念圖之機器人數位說故事學習模式，控制組則是一般機器人數位說故事學習模式。機器人數位說故事的學習平台使用的是女媧科技平台和該公司的凱比機器人作為機器人數位說故事的教學平台和機器人，本研究主要採用「程式實驗室」與「內容編輯器」兩項核心模組，引導學生進行機器人腳本撰寫與情節設計，以實現故事的互動呈現與跨模態整合，如圖 1 所示。

圖 1 系統介面-左為內容編輯器介面，右為程式實驗室介面

本研究參考 Lee (2014)和 Chen 等人(2024)學者在自律調節學習單的相關設計，並結合 Chen Hsieh 和 Lee (2023)的數位說故事的教學策略，設計成本研究學習模式流程和引導式學習策略，稱為 PWUDE 學習策略，包含四個階段：計畫(Plan)、撰寫故事(Write story)、使用機器人(Use Kebbi robot 平台)、展示(Display)與評估(Evaluation)。根據 Chen et al.(2015)所提出之理論架構，時間序列概念圖能強化學習者在敘事構思過程中之知識彙整與轉化決策，特別適用於具有結構性與創作性的數位說故事任務。為確保教學歷程的可控性與比較一致性，實驗組與控制組皆遵循 PWUDE 模式進行學習活動，差異僅在「撰寫故事」與「使用機器人」階段，實驗組額外運用時間序列概念圖支援其知識結構建構，如圖 2 所示。

圖 2 PWUDE 策略說明圖

首先，「計劃(Plan)」階段中，教師在課堂內給予學生關於防癌飲食知識和生活的教材內容，要求小組進行工作分配、規劃目標排程和故事基礎設定。在課外時間需自行與小組進行討論和蒐集防癌飲食、防癌生活習慣資料，實驗組利用概念圖彙整知識脈絡，控制組則採用文字敘述、條列摘要或表格式方式進行資料整理，未使用概念圖作為學習工具。

本研究使用 GitMind 線上平台，讓學習者以小組為單位進行協作繪製概念圖。當學生於課堂中學習完授課教師提供的防癌飲食教材內容後，可以上網尋找延伸資訊以及探索學習活動所需的知識、概念內容，透過小組討論和繪製概念圖來幫助知識整合和將概念關聯性視覺化。如圖 3 所示。

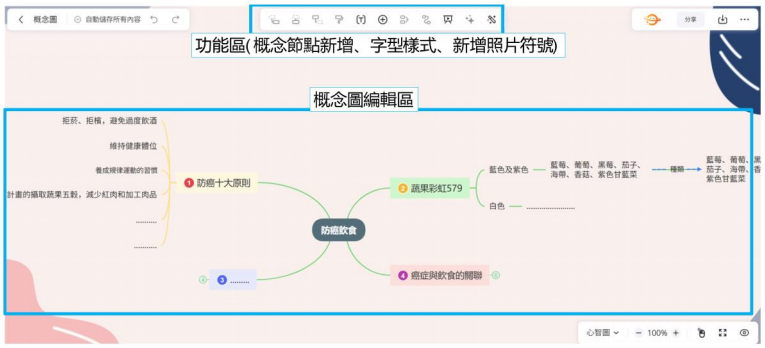


圖 3 GitMind 編輯頁面

接著「撰寫故事(Write story)」階段，當小組成員協作完成概念圖後，會指導學習者根據概念圖知識在學習活動中設計故事內容，小組將故事每段情節會運用的概念知識、角色、事件等內容彙整成概念圖。再透過線條連結不同情節來表示每段情節的先後順序，最終完成時間序列概念圖，如圖 4 所示。教師會引導學生透過 AI 工具生成出所需的場景圖片、語音旁白等，如 Tensor.Art、TTSMAKER。

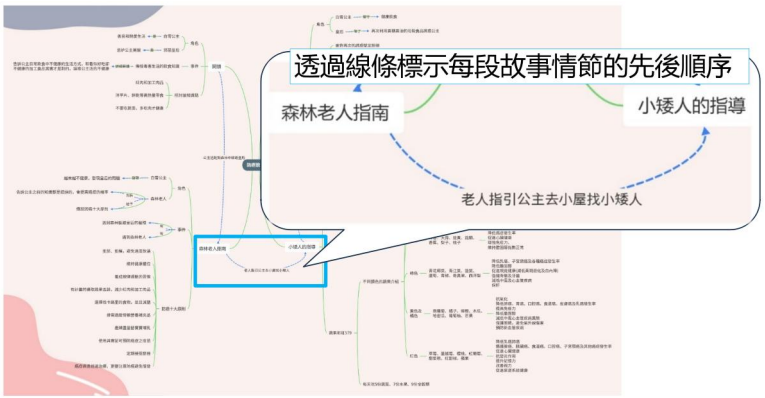


圖 4 時間序列概念圖示意圖

此外，學生必須在學習單上記錄自己或小組在在該階段完成了甚麼任務、費時多久和進度描述等，如圖 5 所示。而控制組差別自在於不會進行時間序列概念圖的製作，而是創作防癌飲食的故事後，便設計場景圖片、語音旁白即可。

在執行任務時，請每位成員實時回報進度，紀錄任務到監控表中，發現有成員在執行任務時遇到問題也可以立即給予協助。

➢ 「珍愛」的監控表：

你們完成了甚麼？ (例如：完成腳本設計)	花費多少時間(幾天、幾小時)	哪些事情幫助/影響了你？		
		學習地點	使用那些學習工具	干擾因素

圖 5 PWUDE 學習單之任務監控表

「使用機器人(Use kebbi showtime)」階段，學生在這一階段會進行機器人數位說故事設計任務，會請學習小組到女媧科技系統的內容編輯器、程式實驗室模組進行故事編輯，透過教師的引導，按照語音→圖片→機器人動作的順序進行建置故事專案。該階段學生同樣必須記錄自己或小組完成了甚麼任務、費時多久和進度描述等，前文的圖 5 所示。

最後的「展示評估(Display and Evaluation)」階段，每個小組會在課堂上進行主題成果報告，報告後會由教師指導回饋、組間互評，組間互評標準參考自 Mokhtar 等人 (2011)說故事呈現的評估標準並由授課教師制定而成。學生不僅必須為不同小組評鑑分數，同時更要求於提供相關的評論及建議，評分標準如圖 6 所示。控制組則不使用時間序列概念圖，其概念圖評分標準調整為評量防癌知識脈絡之清晰性與組織結構。

Rating Dimension	1	3	5
故事內容正確性	包含大量不正確的防癌內容（錯字、不正確的概念或描述）	包含少量不正確的防癌內容（錯字、不正確的概念或描述）	呈現的防癌飲食內容完全正確
故事內容完整性和豐富度	防癌飲食的內容並未呈現在機器人故事中	防癌飲食的內容少量出現在機器人故事中	防癌飲食的內容完整並豐富呈現在機器人故事中
機器人腳本設計	腳本設計不佳，對象設定不佳，內容過於雜亂	腳本設計不夠清晰，對象設定不夠明確，內容有點雜亂不容理解	腳本設計清晰，對象設定明確，內容清楚明瞭容易理解
概念圖設計	簡報內容未使用概念圖	概念圖少量融入簡報中	完整融入概念圖簡報內容中
簡報講述	沒有和觀眾有互動，講述能力可再加強	講述能力不錯，還可以增加更多與同學的互動	講述生動，自然，有助於我瞭解簡報內容，與觀眾有互動
簡報製作	簡報製作和設計不太能順利讓觀眾理解報告內容	簡報製作與設計，能理解報告內容	簡報製作精美，容易了解簡報內容

圖 6 PWUDE 學習單之組間互評規準表

主題報告後學生撰寫個人反思日誌， 參考 Chen 等人 (2024)學者的反思表進行改編，問題共九題，例如：「經過本活動的學習，您認為自己是否有學好防癌飲食？」，要求學生反

思、檢討和分析個人或小組缺失、不足之處，加以修正，重新審視學習計畫，修改不足之處，並於最終報告前完成作品。

4. 研究設計

4.1. 實驗對象

本研究的實驗對象為台北市某高中一年級和三年級的兩個班級，以健康與護理課程第一章「了解預防慢性病和傳染病」的防治癌症作為主題範圍，採準實驗方法，實驗組為 19 名，採用基於時間序列概念圖之機器人數位說故事學習模式，而控制組為 19 名，採用一般機器人數位說故事模式。

4.2. 實驗流程

第一周，授課教師會講解四週課程規畫，給予防癌飲食相關的基礎知識概念、架構等內容，並教導學生如何運用或使用概念圖和機器人平台的操作教學。兩組同學都會在教學前進行學習前測和前問卷。

第二周至第四周，實驗組採用基於時間序列概念圖之機器人數位說故事學習模式進行學習，控制組則採用一般機器人數位說故事學習模式。 兩組的差異就在於學習活動中是否有時間序列概念圖輔助學習，兩組學習的內容、測試皆為相同，老師會根據 PWUDE 學習策略來引導小組合作完成數位故事設計活動。

第五周，兩組學生會進行最終成果報告和同儕互評，讓每位同學為其他組別進行評分和給予建議，以及教師給予每組學生的發表內容建議與回饋。兩組同學都會在成果展示後進行學習後測和後問卷調查，此外各小組會隨機抽選一名學生進行 10 分鐘的焦點小組訪談。

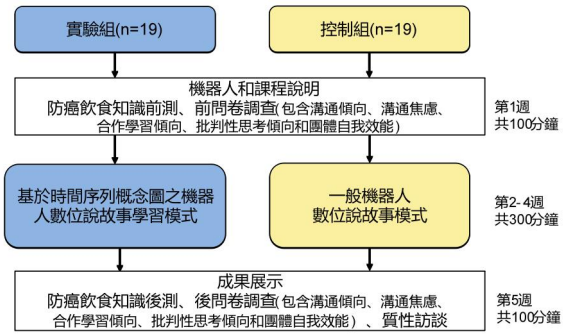


圖 7 實驗流程圖

4.3. 研究工具

本研究所採用的研究工具為學習成就-防癌飲食知識測驗、問卷調查和質性訪談，問卷由溝通焦慮、溝通傾向、合作學習傾向、批判思考傾向和團體自我效能問卷組成。

前測、後測皆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，滿分皆為 100 分。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。

溝通焦慮問卷改編自 AlcCroskey 和 Richmond (1992)的研究，旨在調查學習者在學習活動前後衡量學生在小組學習活動中是否可以自然地進行溝通交流，抑或是表現出緊張害怕，問卷的 Cronbach α係數為 0.84，共 6 題，採用李克特 5 點量表。

溝通傾向問卷改編自 Lai 和 Hwang (2014)的研究，旨在衡量學習者在團體學習活動中是否會關注他人的感受，有較高的溝通傾向，也就代表學習者能尊重包容且友善地與組員、團體進行交流，問卷的 Cronbach α係數為 0.92，共 7 題，採用李克特 5 點量表。

合作學習傾向問卷改編自 Hwang 等人 (2011)的研究，旨在衡量學習者在團體學習活動中是否願意接納同儕的意見和觀點，問卷的 Cronbach α係數為 0.89，共 7 題，採用李克特 5 點量表。

批判思考傾向問卷改編自 Lin 等人 (2019)學者在 Chai 等人 (2015)基礎上修訂的，旨在衡量學習者對於學習的內容是否會用不同的觀點去理解和分析，問卷的 Cronbach α係數為 0.86，共 6 題，採用李克特 5 點量表。

團體自我效能問卷改編自 Wang 和 Lin (2007)學者修訂的，旨在衡量團體成員對於團體能力的判斷，團體自我效能高低會影響團體的學習成效、學習行為、目標的制定以及團體的努力程度，問卷的 Cronbach α係數為 0.89，共 8 題，採用李克特 5 點量表。

質性訪談的七個開放式問題改編自 Hwang et al. (2009)，會針對兩班每個小組中隨機抽選一人進行 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「你認為基於時間序列概念圖之機器人數位說故事學習模式的優點是什麼？」，訪談內容由兩位經驗豐富的老師進行分析編碼。

5. 研究結果

兩組之學習成就、溝通焦慮、溝通傾向、合作學習傾向、批判思考傾向以及團體自我效能之描述性統計顯示，各組皆呈現常態分布。並以獨立樣本 t 檢定來檢驗，結果顯示兩組學生在實驗前對於各項無顯著差異。接者，檢驗不同項目的同質性檢驗，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，結果顯示六個項目變量皆不拒絕同質性假設，故可進行 ANCOVA 分析。六個項目的共變數分析如表 1 所示，可以看出除了溝通焦慮外，其他項目在統計上皆有顯著差異。

表 1 各項目的共變數分析摘要

變項	Group	N	Mean	SD	Adjusted Mean	Adjusted SD	F	η^2
溝通焦慮	實驗組	19	15.53	3.96	15.49	1.15	0.27	
	控制組	19	16.32	5.77	16.35	1.15		
學習成效	實驗組	19	86.53	11.64	86.55	3.40	6.42*	0.16
	控制組	19	74.37	17.04	74.35	3.40		
溝通傾向	實驗組	19	3.73	0.74	3.77	0.14	4.51*	0.11
	控制組	19	3.40	0.64	3.35	0.14		
合作學習傾向	實驗組	19	3.96	0.66	3.95	0.13	7.90**	0.18
	控制組	19	3.43	0.46	3.43	0.13		
批判思考傾向	實驗組	19	3.92	0.64	3.93	0.12	4.58*	0.12
	控制組	19	3.56	0.40	3.56	0.12		
團體自我效能	實驗組	19	3.92	0.64	3.90	0.12	11.13**	0.24
	控制組	19	3.32	0.40	3.34	0.12		

此外，訪談施測為學習活動結束後當日，安排學習者進行開放性問題質性訪談，訪談的目的為了解本次實驗學習活動的學習者對於學習過程的感受、意見與想法。表 2 是根據 Strauss's (1987)的理論基礎進行分析的訪談結果摘要。即按照三種策略對記錄進行分析和編碼，以確定編碼結果：開放式編碼、軸向編碼和選擇性編碼。在編碼過程中，兩名編碼員反覆閱讀訪談資料並為其命名，還分享和討論各自的編碼結果，直到達成一致性。

表 2 訪談編碼結果

主題	代碼	提及次數	
		實 驗	控 制

		組	組
學習過程 (Learning Process)	學習動機（Learning Motivation, LM）	5	5
	學習方式變化（Change in Learning Approach, CLA）	3	3
	新穎性體驗（Novelty Experience, NE）	3	0
	知識建構（Knowledge Construction, KC）	6	1
學習成效 (Learning Outcomes)	技能學習（Skill Acquisition）	6	5
	知識獲得（Knowledge Gain）	4	5
	學習單應用（Use of Worksheets）	5	5
	團隊合作（Teamwork）	6	2
學習挑戰 (Learning Challenges)	時間限制（Time Constraints）	6	2
	技術挑戰（Technical Difficulties）	3	3
	教師引導需求（Need for Teacher Guidance）	3	3
	團隊合作問題（Team Collaboration Issues）	0	1
	學習壓力（Learning Pressure）	1	4
創意表達 (Creative Expression)	創意內容設計（Creative Content Design）	4	5
	創意技術應用（Creative Technical Application）	6	4

訪談結果顯示，實驗組普遍認為基於時間序列概念圖之機器人數位說故事學習模式較能幫助其提升知識建構、團隊合作兩面向的成效。而控制組普遍認為一般機器人數位說故事學習模式讓其專注於故事的創意內容設計。

6. 結論與建議

研究指出，概念圖能連結新舊知識，有助釐清複雜內容，不僅激發邏輯思維，亦促進創意與有意義的學習(Dong et al., 2021)。若結合鷹架策略，可協助學生選擇合適知識並減輕任務負擔，有效完成專題故事。本研究發現，實驗組在防癌飲食知識的理解上顯著優於控制組，呼應過去研究指出心智圖與概念圖能提升學習成效與批判思維，幫助統整與理解複雜知識(Yang et al., 2025; Chiu & Hwang, 2024)。然而，在溝通焦慮上，兩組無顯著差異。研究者認為，PWUDE 學習單促進了小組討論與專題進展，使溝通焦慮未因學習策略差異而上升。建議未來教材設計應依學科特性選擇合適資料呈現方式，並結合多元策略，提升解題能力、減少挫折，有助學生掌握知識建構並提升成效(Chu et al., 2015)。也可進一步探討該模式對創造力與高階思維的影響。本研究也存在一些限制，包括實驗對象單一、樣本量較少、指導老師的介入影響以及研究持續時間較短皆可能影響學生的學習成效，建議未來研究可以針對這些限制加以研究。

參考文獻

Al-Hosni, K., Sumri, H. A., Chan, M. F., & Al-Azri, M. (2024). Introduction of Cancer Education Program in Secondary School Curriculum: Perceptions of Education Teachers and Policymakers in Oman. *Journal of Cancer Education*, 1-10. https:/doi.org/10.1007/s13187-024-02457-2

AlcCroskey, J. C., & Richmond, V. P. (1992). Communication apprehension and small group communication.

Chang, C. C., Hwang, G. J., & Chen, K. F. (2023). Fostering professional trainers with robot-based digital storytelling: A brainstorming, selection, forming and evaluation model for training

guidance. *Computers & Education*, *202*, 104834. https://doi.org/10.1016/j.compedu.2023.104834
Chang, C.-Y., Hwang, G.-J., Chou, Y.-L., Xu, Z.-Y., & Jen, H.-J. (2023). Effects of robot-assisted digital storytelling on hospitalized children's communication during the COVID-19 pandemic. *Educational technology research and development*, *71*(3), 793-805. https:/doi.org/10.1007/s11423-023-10209-0
Chen Hsieh, J., & Lee, J. S. (2023). Digital storytelling outcomes, emotions, grit, and perceptions among EFL middle school learners: Robot-assisted versus PowerPoint-assisted presentations. *Computer Assisted Language Learning*, *36*(5-6), 1088-1115. https:/doi.org/10.1080/09588221.2021.1969410
Chen Hsieh, J. (2021). Digital storytelling outcomes and emotional experience among middle school EFL learners: Robot-assisted versus PowerPoint-assisted mode. *TESOL Quarterly*, *55*(3), 994-1010. https://doi.org/10.1002/tesq.3043
Chen, Y.-C., Hwang, G.-J., & Lai, C.-L. (2024). Motivating students to become self-regulatory learners: a gamified mobile self-regulated learning approach. *Education and Information Technologies*, 1-24. https:/doi.org/10.1007/s10639-024-12462-z
Chen, Y. L., Hsu, C. C., Lin, C. Y., & Hsu, H. H. (2022). Robot-assisted language learning: Integrating artificial intelligence and virtual reality into English tour guide practice. *Education Sciences*, *12*(7), 437. https://doi.org/10.3390/educsci12070437
Chen, X. (2020). Pre-service Teachers' Self-efficacy of Interdisciplinary Team Teaching through the Use of Collaborative Concept Map. *International Journal of Technology in Teaching and Learning*, *15*(2), 76-94.
Chiu, M.-C., & Hwang, G.-J. (2024). Enhancing students' critical thinking and creative thinking: An integrated mind mapping and robot-based learning approach. *Education and Information Technologies*, 1-34. https:/doi.org/10.1007/s10639-024-12752-6
Chu, H. C., Yang, K. H., & Chen, J. H. (2015). A time sequence-oriented concept map approach to developing educational computer games for history courses. *Interactive Learning Environments, 23(2), 212-229.* https:/doi.org/10.1080/10494820.2014.979208
Dong, Y., Zhu, S., & Li, W. (2021). Promoting sustainable creativity: An empirical study on the application of mind mapping tools in graphic design education. *Sustainability*, *13*(10), 5373. https://doi.org/10.3390/su13105373
Hohenhaus, A. E., & Provost, D. C. (2024). A Unique Spectrum of Care Tool Provides a Self-Regulated Learning Opportunity and Facilitates Client Communication. *Journal of Veterinary Medical Education*, e20230144. https://doi.org/10.3138/jvme-2023-0144
Hwang, G.-J., Chen, H.-C., Hsu, C.-Y., & Hwang, G.-H. (2023). Effects of a graphic organizer-based two-tier test approach on students' learning achievement and behaviors in spherical video-based virtual learning contexts. *Computers & Education*, *198*, 104757. https://doi.org/10.1016/j.compedu.2023.104757
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., Yang, T.-C., Tsai, C.-C., & Yang, S. J. (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
Hwang, G. J., Shi, Y. R., & Chu, H. C. (2011). A concept map approach to developing collaborative Mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology*, *42*(5), 778-789. https://doi.org/10.1111/j.1467-8535.2010.01102.x
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*(3-4), 276-291. https://doi.org/10.1504/IJMLO.2014.067029
Lee, A. (2014). *Socially shared regulation in computer-supported collaborative learning*. Rutgers The State University of New Jersey, School of Graduate Studies.
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. https://doi.org/10.1016/j.compedu.2020.103899
Liang, J. C., & Hwang, G. J. (2023). A robot-based digital storytelling approach to enhancing EFL learners' multimodal storytelling ability and narrative engagement. *Computers & Education*, *201*, 104827.
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. https://doi.org/10.1016/j.compedu.2019.05.014
Mokhtar, N. H., Halim, M. F. A., & Kamarulzaman, S. Z. S. (2011). The effectiveness of storytelling in enhancing communicative skills. *Procedia-Social and Behavioral Sciences*, *18*, 163-169. https://doi.org/10.1016/j.sbspro.2011.05.024
Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. cambridge University press.
Wang, S.-L., & Lin, S. S. (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
Yang, K. H., Chu, H. C., Hwang, G. J., & Liu, T. J. (2025). A progressive concept map-based digital gaming approach for mathematics courses. *Educational technology research and development*, 1-31. https://doi.org/10.1007/s11423-025-10461-6

The Effect of a Two-Tier-Test Approach Integrated with a Tablet IRS on University Students' Self-Efficacy and Attention

Jerry C. Y. Sun [1], Yen-Ling Chen [2], Li-An Li [3]
[1.2.3] Institute of Education, National Yang Ming Chiao Tung University, Taiwan
[*] jerrysun@nycu.edu.tw

Abstract: *Effective feedback is crucial for promoting learner engagement and improving cognitive outcomes, yet its design and delivery remain challenging. This study investigated the effects of integrating a two-tier testing approach with elaborated feedback (EF) into a tablet-based interactive response system (IRS) on university students' self-efficacy and focused attention. A quasi-experimental design with 74 participants was conducted to compare different feedback strategies. While no statistically significant improvement was observed in self-efficacy, students who received elaborated feedback demonstrated higher mean scores in focused attention. The findings suggest that promoting active reasoning through feedback design may enhance learner engagement in technology-supported environments.*

Keywords: Tablet PCs, Interactive Response System, Two-tier test, Self-efficacy, Focused attention

1. Introduction

The integration of Interactive Response Systems (IRS) into classroom teaching has been widely recognized for its ability to enhance student engagement and interaction. By utilizing the Interactive Response System (IRS), students can receive immediately feedback and information, which not only enhances their learning experience but also improves their academic performance (Fu et al., 2025). A two-tier test is a diagnostic assessment method designed to evaluate the correctness of a student's response, and their reasoning behind it as well (Hwang et al., 2023). This study investigates the impact of integrating two-tier polling with elaborated feedback on university students' self-efficacy and attention. By incorporating structured feedback into IRS, this study examines its role in enhancing learning confidence and sustaining attention in classroom settings. Accordingly, the study addresses the following research questions:

1. Does the two-tier polling strategy integrated with feedback significantly improve learners' self-efficacy?
2. Does the two-tier polling strategy integrated with feedback significantly improve learners' attention?

2. Literature review

2.1. Interactive Response System (IRS) and Two-tier test

Interactive Response Systems (IRS) provide immediate feedback, fostering an interactive learning environment while enhancing students' knowledge recall and attention (Muis et al., 2015). Effective feedback mechanisms allow learners to assess their understanding, adjust their strategies, and improve learning efficiency (Lang & Betsch, 2024). This study categorizes feedback into knowledge of correct response (KCR) and elaborated feedback (EF). KCR informs learners whether

their answers are correct, while EF provides detailed explanations, helping students refine their knowledge and correct misconceptions. Integrating both types of feedback into IRS enhances learning engagement and cognitive processing.The two-tier test further supports learning by assessing both declarative knowledge in the first tier and explanatory reasoning in the second (Hwang et al., 2023). This format enables instructors to identify misconceptions and provide targeted interventions. When combined with IRS, the two-tier test encourages deeper reflection and conceptual understanding (Yang et al., 2015). This study explores the impact of integrating IRS with two-tier testing on students' self-efficacy and attention, aiming to foster a more engaged and cognitively stimulating classroom experience.

2.2. Self-Efficacy and Attention

Self-efficacy, the belief in one's ability to complete learning tasks, significantly influences motivation and persistence (Bandura, 1977). Helping learners to set short-term and specific learning objectives, providing explicit rubrics that can help learners succeed will all improve their self-efficacy (Schunk & Mullen, 2012). It is shaped by external factors, personal competence, and prior experiences. Fortner-Wood et al. (2013) found that IRS polling activities reduce learning anxiety and promote a more interactive and engaging learning environment, thereby improving self-efficacy. Attention is another crucial factor in learning engagement, as higher sustained attention helps students stay focused, minimize distractions, and improve comprehension. Building on these insights, this study explores how integrating two-tier polling with elaborated feedback in an IRS-based learning environment enhances students' self-efficacy and attention, contributing to a more interactive and cognitively engaging learning experience.

3. Method

3.1. Participants and Experimental design

The research subjects were 74 participants of two classes: 49 from the "Introduction to Networks" class and 25 from the "Information Security" class at a university in Taiwan. There were 29 male students (39.2%), and 45 female students (60.8%). This quasi-experimental study divided students into three groups: Group A received correct answer feedback (n=25), Group B took a two-tier test with correctness feedback (n=27), and Group C took a two-tier test with elaborated feedback. Before the experiment, participants completed a pretest on self-efficacy (n=22). The intervention consisted of two 50-minute sessions over one week. Afterward, participants completed post-tests on self-efficacy and attention. During class, all groups used tablets for polling activities. Group A received single-choice questions with correctness feedback. Group B received two-tier test questions with correctness feedback, while Group C received additional explanations and reasoning. Students worked in groups of four, discussing answers but responding individually on tablets.

3.2. Instruments

Self-efficacy was measured using an adapted version of the Self-Efficacy for Learning and Performance subscale from the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al., 1993), consisting of eight items rated on a 6-point Likert scale (α = .68), indicating acceptable reliability (Nunnally, 1967). Attention was assessed using a translated version of the Focused Attention Scale (Christakou et al., 2012), which included ten items—three of which were reverse-

coded—rated on an 8-point Likert scale (α = .81), showing good internal consistency (Nunnally, 1978). The tablet-based IRS employed in this study was the Interactive Feedback-Based In-Class Teaching System (iFIT) (Sung et al., 2016). Feedback varied across groups: Group A received knowledge of correct response (KCR); Group B completed a two-tier test with KCR and response summaries; and Group C completed a two-tier test with both KCR and elaborated feedback (EF), which provided detailed explanations to reinforce conceptual understanding. Figure 1 illustrates the tablet interface used in each condition.

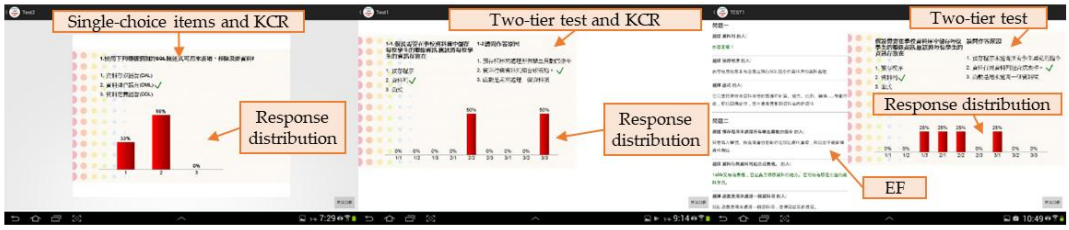


Fig.1 Examples of the tablet interface for each group

4. Result

After accounting for pretest scores, the differences in posttest self-efficacy among the groups were not statistically significant ($F = 2.67$, $p = .08$, $\eta^2 = .07$). Although both the elaborated feedback (EF) and correctness feedback groups had higher adjusted means than the correct answer group, the overall group effect ***did*** not reach significance. This outcome suggests that while two-tier polling may reflect a positive trend in enhancing students' self-efficacy, the differences observed were not strong enough to support a clear conclusion under the current study conditions. For focused attention, Levene's test indicated unequal variances across groups ($F = 7.46$, $p = .001$). As a result, a Brown-Forsythe test was conducted and showed a statistically significant overall group effect ($F = 3.40$, $p = .042$). However, the follow-up Dunnett T3 comparisons did not reveal any significant pairwise differences. Although the mean score for the elaborated feedback group (M = 46.15) was descriptively higher than that of the correctness feedback group (M = 42.15), the difference was not statistically significant. These findings imply that while elaborated feedback may be linked to better attention outcomes, high variability and limited sample size may have reduced the ability to detect significant differences.

5. Discussion and conclusion

Although This study found that, contrary to earlier findings suggesting that two-tier testing combined with elaborated feedback (EF) promotes learning confidence and cognitive engagement (Yang et al., 2015), no statistically significant improvements in self-efficacy were observed. A plausible explanation concerns the short duration of intervention. Yang et al. (2017) emphasized that reflective thinking and the development of self-beliefs rely on sustained and repeated exposure to diagnostic feedback. In the present case, the limited time frame may not have allowed sufficient opportunities for meaningful internalization, especially given that self-efficacy typically grows through mastery experiences, deliberate practice, and self-regulated reflection (Bandura, 1977). Beyond the issue of duration, the way learners engaged with the feedback may have further

constrained its impact. Gajos and Mamykina (2022) demonstrated that when learners are simply presented with correct answers and explanations, they often engage passively, completing tasks without activating the deeper cognitive processes needed for long-term learning. In this study, despite the provision of detailed feedback, the absence of active reasoning tasks may have limited the opportunities for students to reconstruct their understanding or to critically evaluate their own thinking. Passive acceptance of information facilitates surface-level task completion but rarely promotes integration, reflection, or the strengthening of self-efficacy over time. Future designs should aim to require learners to articulate their reasoning processes, predict outcomes, or self-explain responses after receiving feedback. Embedding such reasoning-based activities can promote deeper cognitive engagement, facilitate knowledge construction, and ultimately support more enduring gains in self-efficacy.

Acknowledgements

This research was supported by the National Science and Technology Council (formerly the Ministry of Science and Technology) in Taiwan through Grant numbers NSTC 112-2410-H-A49-019-MY3 and MOST 111-2410-H-A49-018-MY4. We would like to thank NYCU's ILTM (Interactive Learning Technology and Motivation, see: http://ILTM.lab.nycu.edu.tw) Lab members and the students for participating in the study, as well as the reviewers for providing valuable comments.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191-215.

Christakou, A., Zervas, Y., Psychountaki, M., & Stavrou, N. A. (2012). Development and validation of the Attention Questionnaire of Rehabilitated Athletes Returning to Competition. *Psychology, Health & Medicine*, *17*(4), 499-510.

Fortner-Wood, C., Armistead, L., Marchand, A., & Morris, F. B. (2013). The effects of student response systems on student learning and attitudes in undergraduate psychology courses. *Teaching of psychology*, *40*(1), 26-30.

Fu, Y. L., Hsu, H. P., Guo, J. L., Huang, C. M., & Yang, C. L. (2025). Development and evaluation of a simulation activity on preoperative patient care: Cross-sectional study applying the National League for Nursing/Jeffries Simulation Theory. *Nurse Education in Practice*, *83*, 7, Article 104282. https://doi.org/10.1016/j.nepr.2025.104282

Gajos, K. Z., & Mamykina, L. (2022). Do people engage cognitively with AI? Impact of AI assistance on incidental learning. Proceedings of the 27th International Conference on Intelligent User Interfaces,

Hwang, G. J., Chen, H. C., Hsu, C. Y., & Hwang, G. H. (2023). Effects of a graphic organizer-based two-tier test approach on students? learning achievement and behaviors in spherical video-based virtual learning contexts. *Computers & Education*, *198*, 17, Article 104757. https://doi.org/10.1016/j.compedu.2023.104757

Lang, A. N., & Betsch, T. (2024). Children Learn Simple, Adaptive Decision Strategies from Probabilistic Feedback. *Journal of Cognition and Development*, *25*(1), 148-173. https://doi.org/10.1080/15248372.2023.2248257

Muis, K. R., Ranellucci, J., Trevors, G., & Duffy, M. C. (2015). The effects of technology-mediated immediate feedback on kindergarten students' attitudes, emotions, engagement and learning outcomes during literacy skills development. *Learning and Instruction*, *38*(2015), 1-13.

Nunnally, J. C. (1967). *Psychometric theory* (2nd ed.). McGraw-Hill. Input this URL in a browser to get JPEG cover art data. http://firstsearch.oclc.org/WebZ/DCARead?standardNoType=1&standardNo=007047849X:srcdbname=worldcat:fromExternal=true&sessionid=0

Input this URL in a browser to get HTML table of contents data. http://firstsearch.oclc.org/WebZ/DECRead?standardNoType=1&standardNo=007047849X&sessionid=0&srcdbname=worldcat&key=605cc821e3dec85ac2abc83fa9e855f283a991f15f486fa730c5162e5aa6cbbc&ectype=TOC

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill. Input this URL in a browser to get JPEG cover art data. http://firstsearch.oclc.org/WebZ/DCARead?standardNoType=1&standardNo=007047849X:srcdbname=worldcat:fromExternal=true&sessionid=0

Input this URL in a browser to get HTML table of contents data. http://firstsearch.oclc.org/WebZ/DECRead?standardNoType=1&standardNo=007047849X&sessionid=0&srcdbname=worldcat&key=605cc821e3dec85ac2abc83fa9e855f283a991f15f486fa730c5162e5aa6cbbc&ectype=TOC

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement (EPM)*, *53*(3), 801-813.

Schunk, D. H., & Mullen, C. A. (2012). Self-efficacy as an engaged learner. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 219-235). Springer.

Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, *94*(2016), 252-275.

Yang, T.-C., Chen, S. Y., & Hwang, G.-J. (2015). The influences of a two-tier test strategy on student learning: A lag sequential analysis approach. *Computers & Education*, *82*(2015), 366-377.

Yang, T. C., Fu, H. T., Hwang, G. J., & Yang, S. J. H. (2017). Development of an interactive mathematics learning system based on a two-tier test diagnostic and guiding strategy. *Australasian Journal of Educational Technology*, *33*(1), 62-80. https://doi.org/10.14742/ajet.2154

以擴增實境結合桌上遊戲建置協同學習系統探討遊程規劃之成效差異

Using Abletop games Combined with AR to Build a Collaborative Learning system to Explore the Effectiveness of Itinerary planning

林佳陞[1]，洪浚祐[2*]，林芝育[3]，徐瑋葶[4]
[1]銘傳大學休閒遊憩管理學系
[2]國立陽明交通大學科技管理研究所
[3]國立嘉義大學行銷與觀光管理學系行銷管理碩士班
[4]銘傳大學觀光事業管理學系碩士班
[*] gyo.cyh@gmail.com

【摘要】 科技為學習帶來全新感受，而遊戲能打造出趣味途徑，遊程規劃作為觀光領域重要核心能力，內容涵蓋行程設計、資源配置與決策思維。目前教學上卻仍以傳統教學為主，為解決教學模式單一之問題。本研究建置一套擴增實境結合桌上遊戲的協同學習系統，探討其對遊程規劃學習者學習成效與遊戲化體驗之影響。研究結果顯示，擴增實境不僅能刺激學習者的感官互動、強化悅趣性與成就感外，更有助促進知識內化與創造，提升學習成效。足以可見若能打造多元教學方式讓教學有別以往便能更勝以往，希冀此結果能為觀光旅遊領域的教學提供新的參考與發展方向。
【關鍵字】 擴增實境；桌上遊戲；協同學習；遊戲化體驗；學習成效

Abstract: *Technology brings a new dimension to learning, while games create engaging learning pathways. Travel itinerary planning is a key competency in tourism, encompassing itinerary design, resource allocation, and decision-making. However, traditional teaching methods still dominate, limiting diverse learning approaches. This study develops a collaborative learning system integrating augmented reality (AR) with board games to explore its impact on learning outcomes and gamified experiences. Results indicate that AR enhances sensory interaction, enjoyment, and achievement while promoting knowledge internalization and creativity, ultimately improving learning effectiveness. These findings suggest that diversifying teaching methods can enhance learning experiences. This research hopes to provide valuable insights for tourism education.*
Keywords: Augmented reality, Collaborative learning, gamified experience, learning effectiveness, Tabletop game

1. 前言

旅遊業因其對經濟貢獻與大量就業機會的創造，被視為全球關注的重要產業之一 (Hu et al., 2021)，其中，遊程規劃更是推動觀光發展的重要關鍵，透過有效的規劃可將觀光景點串聯，提升旅遊體驗與產業效益。為提升遊程規劃教育之專業品質，中華民國遊程規劃設計協會積極與臺灣各地高中職及大專院校合作，推動遊程規劃與設計相關專業課程，並設立證照考取制度與競賽。然而，由於目前遊程規劃課程仍主要以傳統講授方式，教學模式相對單一，若能導入更多創新元素將有助於增加學習多樣性與成效。

桌上遊戲作為一種學習工具，在教育領域展現出潛在價值 (Poole et al., 2022)。傳統桌上遊戲主要依賴卡片、骰子等靜態實體道具，隨著技術進步，數位螢幕與智慧遊戲版已逐步融入桌上遊戲設計，以提升互動性與沉浸感，此外隨著行動裝置普及，促使擴增實境成為桌上遊戲應用的一部分，並進一步強化遊戲體驗 (Coolen et al., 2023)。Vassigh 等人 (2020)則指出擴增實境結合協同學習可促進學習者積極參與和互動。綜上所述，本研究旨在將擴增實境技術導入遊程規劃教學中，並使用學習導向之桌上遊戲創建出協同學習的學習環境，使學習者能於學習過程中更深入理解遊程規劃之內容，達到小組間能共同討論以安排出最合適的遊程。

2. 文獻探討

旅遊往往因對環境的不確定性進而引發擔憂，然而透過擴增實境預先了解資訊與展示真實地貌之特性，可提升遊客對目的地資訊之理解，並正向影響其未來選擇旅遊目的地之行為 (Ahmad et al., 2023)。此外，擴增實境不僅創新教材呈現形式，亦促進學習者的主動參與和互動，在提升學習成效上更是顯著優於其他教學工具 (Cai et al., 2021)。同時，擴增實境更有利於團隊溝通互動、知識保留與強化協同學習體驗 (Rani et al., 2023)。

然而，根據自我決定理論，其內心需求能透過遊戲化體驗獲得滿足，使學習者自主性與成就感提升並更專注於學習活動，高度提升學習成效 (Luarn et al., 2023)。桌上遊戲本質亦屬遊戲的一種，而遊戲已被廣泛認為是有效的學習工具之一，遊戲元素可增強參與者的情感反應，進而提升學習動機與吸收能力，對於鞏固學習記憶具有積極的影響，因此在教育環境中展現了極大的應用潛力 (Ninaus et al., 2019)。

3. 研究方法

3.1. 研究架構

本研究冀望建置一套擴增實境桌上遊戲卡牌學習系統，讓學習者即便在無教師指導的情況下，亦能自主習得遊程規劃相關知識，同時，進一步探討學習者在有無擴增實境的環境下學習成效與遊戲化體驗之差異，進而了解作為輔助學習的擴增實境與卡牌結合對學習者體驗之影響，綜合以上，本研究建構之架構圖如下圖 1 所示。

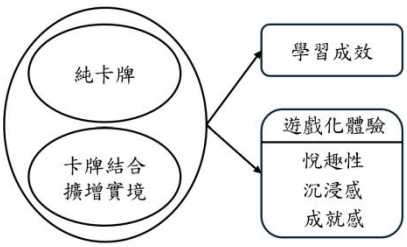


圖 1 研究架構圖

圖 2 實驗流程圖

3.2. 研究對象與實驗流程

本研究之參與對象為大專院校含以上，對遊程規劃相關知識與擴增實境具有興趣者之學習者，總樣本數為 120 位。實驗開始前，研究以隨機分配原則，依照協同學習方式按三人一組分配至控制組與實驗組。隨後，學習者將接受實驗說明與簽署實驗同意書，並根據組別不同進行相關教學，此階段之內容與流程對兩組學習者均一致。待學習者具獨立操作卡牌與系統之能力後，即進入學習體驗環節進行遊程規劃教學，並展開自主學習。教學結束後，學習

者將根據表單與課程測驗進行桌上遊戲結果評分，同時填寫包含學習成效與遊戲化體驗之後測問卷，至此完成整體實驗流程，本研究之實驗流程圖如上圖 2 所示。

3.2. 實驗與研究工具

本研究學習系統建置可分為卡牌設計與擴增實境設計兩個部分，前者主要以 Adobe Illustrator 進行卡牌的外框、繪畫與著色，後者則以 Makar 為核心，利用 Adobe Illustrator 與 Blender 進行模型繪製，搭配 TTSMaker 錄製語音說明輔助學習。研究採用遊程規劃學會考題之一的澎湖作為主要教學內容，卡牌與系統呈現如下圖 3 與圖 4 所示。

圖 3 卡牌示意圖

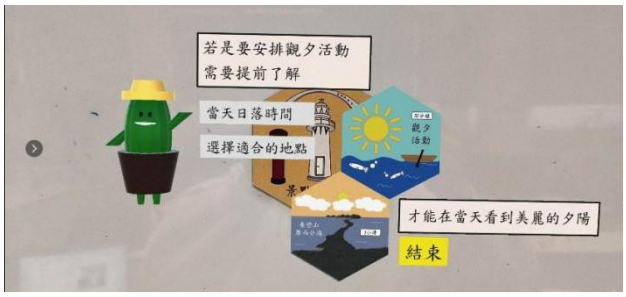


圖 4 系統呈現畫面

為探討學習者學習成效與遊戲化體驗程度，本研究採 Snell 與 Lau (2020)提出之學習成效量表，並根據本研究之目的與主題進行修訂，共計四題，例如：「我能夠應用此實驗所學到的知識來處理複雜的議題。」遊戲化體驗程度則參考 Högberg 等人 (2019)所提出之遊戲化體驗量表，原始量表計有七個子構念，分別為挑戰感、競爭性、引導、社會經歷、悅趣性、沉浸感以及成就感。鑒於本研究之實驗設計，未涉自我挑戰與進行比賽，故刪除挑戰感與競爭性，同時研究採用自主學習模式進行桌上遊戲，無須引導者介入，故刪除引導；而本研究與社會相關經驗無關，故刪除社會經歷。最終，保留悅趣性、沉浸感與成就感，保留之題項分別為六題、三題、四題，並根據研究目的與主題進行修訂，例如：「遊程規劃桌遊給我一種我需要實現目標的感覺。」整題量表項目皆採用李克特五點尺度（1 代表非常不同意，5 代表非常同意）進行衡量，並經具多年遊程規劃教學經驗之專家與資訊教育領域專家進行專家效度檢驗，同時重新檢測學習成效、悅趣性、沉浸感與成就感之 Cronbach's α值分別為 0.866、0.888、0.744、0.823，顯示量表具有良好的一致性。

4. 研究結果

本研究將所蒐集之資料進行整理，剔除未填答正確或異常之問卷後，最終有樣本數為 97 份，為驗證不同學習方式對學習成效與遊戲化體驗之影響，進一步透過成對樣本 t 檢定進行考驗。分析結果顯示兩組學習者在學習成效上具有顯著差異 ($t = -2.687, p < .01$)。遊戲化體驗方面則除了沉浸感並未達顯著差異 ($t = -0.622, p > .05$)外，其餘悅趣性 ($t = -2.672, p < .01$)與成就感 ($t = -2.607, p < .05$)皆具有顯著差異。由上述結果表明，桌上遊戲結合擴增實境之學習方式方能有效提升學習成效與遊戲化體驗，特別是在學習過程中的愉悅與成就感，卻無法有效提升學習者之沉浸感。整體分析結果如下表 1 所示。

表 1 學習成效、遊戲化體驗之成對樣本 t 檢定分析結果

因素	純卡牌組 N=48		卡牌結合擴增實境組 N=50		t
	平均數	自由度	平均數	自由度	
學習成效	3.63	96	3.93	96	-2.687**
悅趣性	4.12	96	4.41	96	-2.672**

沉浸感	3.13	96	3.23	96	-0.622
成就感	3.76	96	4.10	96	-2.607*

註：*$p < .05$、**$p < .01$（N=測驗人數）

5. 結論與研究限制

本研究旨在探討擴增實境導入桌上遊戲，對學習者學習成效與遊戲化體驗之差異，結果表明，除遊戲化體驗之沉浸感未達顯著差異外，其餘悅趣性、成就感與學習成效皆存有顯著差異，本研究根據受測者回饋得知，系統施測採用平板電腦進行，實驗過程中容易受到外部環境與討論需求影響，難以完全沉浸於擴增實境學習情境中。此外，本研究加入互動遊戲與即時回饋元素，如夜釣小管遊戲，學習者可直接點擊螢幕互動並即時看到分數，提升成就感之餘更緩解學習的枯燥感。同時系統結合圖片、3D 模型、語音及按鈕等多模態元素，不僅提升了視覺效果，更允許學習者依進度反覆操作，加深理解與吸收。本研究獲得研究結果驗證擴增實境結合桌上遊戲於遊程規劃教學具有價值，未來可針對不同的樣本類型、互動元素或學習時長與歷程進一步了解學習成效與遊戲化體驗之差異。此外，本研究期望為觀光與教育領域提供創新教學模式，促進更多元且具實效之學習方式，以提升學習品質與產業人才培育。

參考文獻

Ahmad, H., Butt, A., & Muzaffar, A. (2023). Travel before you actually travel with augmented reality–role of augmented reality in future destination. *Current Issues in Tourism*, *26*(17), 2845-2862. https://doi.org/10.1080/13683500.2022.2101436

Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J. C. (2021). Effects of learning physics using Augmented Reality on students' self-efficacy and conceptions of learning. *British Journal of Educational Technology*, *52*(1), 235-251. https://doi.org/10.1111/bjet.13020

Coolen, K., Groen, J., Lu, H. X., Winkelman, N., & Van Der Spek, E. D. (2023, November). Augmented Adventures: Using Different Perspectives to Design Novel Tabletop Game Experiences with Augmented Reality. In *International Conference on Entertainment Computing* (pp. 251-260). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-8248-6_21

Högberg, J., Hamari, J., & Wästlund, E. (2019). Gameful Experience Questionnaire (GAMEFULQUEST): an instrument for measuring the perceived gamefulness of system use. *User modeling and user-adapted interaction*, *29*(3), 619-660. https://doi.org/10.1007/s11257-019-09223-w

Hu, I. L., Chang, C. C., & Lin, Y. H. (2021). Using big data and social network analysis for cultural tourism planning in Hakka villages. *Tourism and Hospitality Research*, *21*(1), 99-114. https://doi.org/10.1177/1467358420957061

Poole, F. J., Clarke-Midura, J., Rasmussen, M., Shehzad, U., & Lee, V. R. (2022). Tabletop games designed to promote computational thinking. *Computer Science Education*, *32*(4), 449-475. https://doi.org/10.1080/08993408.2021.1947642

Rani, S., Mazumdar, S., & Gupta, M. (2023, December). Marker-Based Augmented Reality Application in Education Domain. In *International Conference on Machine Learning, Image*

Processing, Network Security and Data Sciences (pp. 98-109).Cham: Springer Nature Switzerland.

Snell, R. S., & Lau, K. H. (2020). The development of a service-learning outcomes measurement scale (S-LOMS). *Metropolitan Universities*, *31*(1), 44-77. https://doi.org/10.18060/23258

Vassigh, S., Davis, D., Behzadan, A. H., Mostafavi, A., Rashid, K., Alhaffar, H., ... & Gallardo, G. (2020). Teaching building sciences in immersive environments: A prototype design, implementation, and assessment. *International Journal of Construction Education and Research*, *16*(3), 180-196. https://doi.org/10.1080/15578771.2018.1525445

Luarn, P., Chen, C. C., & Chiu, Y. P. (2023). Enhancing intrinsic learning motivation through gamification: a self-determination theory perspective. *The International Journal of Information and Learning Technology*, *40*(5), 413-424. https://doi.org/10.1108/IJILT-07-2022-0145

結合「全班鷹架式合作學習」與「個人化即時回饋學習平臺」對於國小高年級學童書法學習之學習成效分析

An Analysis of the Learning Outcomes of Combining 'Whole-Class Scaffolding Cooperative Learning' and 'Personalized Instant Feedback Learning Platforms' in Calligraphy Learning for Upper Elementary School Students

林采瑩[1*],王俊傑[1]
[1]屏東大學教育學系
[*]linty0424@gmail.com, jie809@mail.nptu.edu.tw

【摘要】隨著科技進步，在寫字教學上已能透過 AI 數位學習平臺給予學習者個人化的回饋。但即便能獲得個人化回饋，個別學習者的學習表現仍有差異。因此本研究擬探討結合「全班鷹架式合作學習」在國小書法教學的應用，是否能提升學生的學習表現，並分析對學生的相關影響。研究採準實驗設計，實驗組學生以「結合全班鷹架與個人化回饋之合作學習模式」進行學習，控制組則採用「結合個人化回饋之合作學習模式」進行學習。本研究以國小五年級學童為實驗對象，探究實驗組學生在學習成效、學習動機、自我效能、合作學習參與度及問題解決傾向等各方面的表現。

【關鍵字】合作學習；全班鷹架式合作學習；個人化即時回饋；書法教學；寫字教學

Abstract: *With the advancement of technology, AI-powered digital learning platforms can now provide learners with personalized feedback in handwriting instruction. However, despite receiving personalized feedback, individual learners still exhibit differences in learning performance. Therefore, this study explores the application of Whole-Class Scaffolded Cooperative Learning in elementary school calligraphy instruction to determine whether it enhances students' learning performance and analyzes its effects on students.This study adopts a quasi-experimental design, where the experimental group engages in learning through a cooperative learning model integrating Whole-Class Scaffolding and Personalized Feedback, while the control group follows a cooperative learning model with Personalized Feedback only. The participants are fifth-grade elementary school students, and the study examines their performance in various aspects, including learning outcomes, learning motivation, self-efficacy, participation in cooperative learning, and problem-solving tendencies.*

Keywords: Cooperative Learning, Whole-Class Scaffolded Cooperative Learning, Personalized Immediate Feedback, Chinese Calligraphy Instruction, Calligraphy Instruction

1. 緒論

1.1. 研究背景動機

在數位科技發達的時代，手機打字與語音輸入已成為主流，但書寫能力依然是語文素養的重要基石，不僅影響學習者的學習過程，更關係到學生的認知發展、記憶力與表達能力。

而研究者於實際教學現場中發現，許多學生因缺乏練習機會與學習動機，導致書寫的筆畫與結構掌握不佳，甚至影響學習信心與表現，隨著科技進步與數位工具的廣泛應用，學習者對於傳統手寫的專注度與投入程度逐漸下降。歷年來的課程綱要，都明確強調工整書寫的重要性，即使在資訊發達的時代，書寫技能仍是教學與考試中不可或缺的一環。除此之外，基於中文字形來源與構造的獨特性，寫字教學也向來是中文學習中重要的一環。尤其是中文初階學習者，例如國小學童、外籍人士，書寫漢字都是一個很重要的學習成果的輸出方式。至於進階的中文學習者，更可以將目標推向美感或藝術性的展現。

然而，隨著資訊科技的普及，教師應強化媒體素養，以有效融入教學實踐。而媒體素養的培養，更能有效地提升學生的書寫能力，在教育課程中融入資訊媒體的訓練，可以有效提升學生的學習能力，故教師須同時培養學生傳統及資訊的書寫能力(Hobbs & Frost, 2003；Levin & Arafeh, 2007)。相對於閱讀、寫作、或是聆聽與口語表達，目前此領域之相關研究較少，根據目前既有的資料顯示，臺灣僅有 16 篇與「寫字教學」相關的碩博士論文；而根據中國知網（CNKI）資料庫，大陸則僅有 17 篇與「寫字」相關的博士論文，而且這些研究主要聚焦於書寫困難和科技融入等議題。因此，本研究試圖針對數位輔助中文書寫進行研究，以彌補此一研究上之缺憾。

中文寫字的傳統教學方式，多由教師示範，學習者仿寫、反覆練習以進行學習。在個人的寫字練習上，學習者常遇到的問題在於個人練習時，難以獲得個人化的即時引導與回饋，學生在教師的教學過程中，若缺乏有效的教學策略，容易感到學習過程枯燥無味，進而影響學習動機，長期下來，可能會形成學習成效不佳的惡性循環。目前，許多研究明確指出透過 AI 數位學習平臺，已能解決學習動機下降、教材枯燥、學生缺乏互動，以及未能即時獲得回饋等問題，並且能有效協助學習者理解與掌握書寫的結構與技巧（絲凱郁、張介英，2022）。平臺能展示標準的筆順和字形結構，學生在個人化練習時，若筆順、筆畫有誤，皆能獲得即時的回饋與引導。然而學習者仍常有著對於回饋訊息吸收程度不同(例如對於字體部件、位置的觀察能力落差)、學習動機落差等情況。有些學生可能對回饋並不感興趣或缺乏動力去改正錯誤，這使得即便提供了即時回饋，也無法顯著提升學習成效。

因此本研究提出了結合「全班鷹架式合作學習」(Whole-Class Scaffolding, WCS)與「個人化即時回饋學習平臺」的數位學習模式，期望能解決上述問題。合作學習已被證實能產生許多正向的學習成效(Chen et al., 2018；Pyle et al., 2017)，也有不少學者指出合作學習在實際施行上可能遇到的困難，例如學生參與度不均、學生能力落差過大、小組討論失焦等問題。本研究更透過「學會一起學習（Learning to Learn Together, L2L2）」的對話機制，使學生透過結構化的合作學習模式，讓學生不僅是「一起寫字」，而是「一起討論與評估字形」，補足僅使用「個人化即時回饋學習平臺」可能造成因個人能力而有學習落差之情況。並且期望在提升書法學習成效之外，進而對於學習動機、自我效能、合作傾向與問題解決傾向等面向皆能有正向之影響。

1.2. 研究問題

本研究採準實驗研究之不等組前後測設計，將學習者分為：採用「結合全班鷹架式合作學習與個人化即時回饋的數位學習模式」之實驗組，與採用「傳統合作學習的數位學習模式」之控制組。本研究欲探討之研究問題如下：

(1) 相較於「結合個人化回饋之合作學習模式」，「結合全班鷹架與個人化回饋之合作學習模式」能否提升國小高年級學童書法學習之學習成效。

(2) 相較於「結合個人化回饋之合作學習模式」，「結合全班鷹架與個人化回饋之合作學習模式」能否提升國小高年級學童書法學習之學習動機。
(3) 相較於「結合個人化回饋之合作學習模式」，「結合全班鷹架與個人化回饋之合作學習模式」能否提升國小高年級學童書法學習之自我效能。
(4) 相較於「結合個人化回饋之合作學習模式」，「結合全班鷹架與個人化回饋之合作學習模式」能否提升國小高年級學童書法學習之合作學習傾向。
(5) 相較於「結合個人化回饋之合作學習模式」，「結合全班鷹架與個人化回饋之合作學習模式」能否提升國小高年級學童書法學習之問題解決傾向。

2.文獻回顧

2.1. 寫字教學

小學生的書寫能力發展對於學習相當重要，然而約有 10%至 30%的學生在書寫過程中遇到困難，這可能影響學業表現與自信。許多傳統的寫字教學方法，如反覆抄寫，雖然能夠提升書寫熟練度，但可能降低學生的學習動機(Li-Tsang et al., 2023)。傳統的寫字教學多採用「重複練習法」。而 Li-Tsang et al.(2023)認為這種方法可能導致學生對寫字感到乏味，影響學習興趣，相較之下，透過遊戲、視覺-動作整合訓練及小組活動，能有效改善學生的書寫速度、字體結構。除了傳統的寫字教學方法，書法訓練也和學習者書寫能力高度相關，書法訓練以漢字的筆畫、結構及書寫規範為基礎，再透過正確的筆畫運用與字形結構安排，幫助學生增強書寫的規範性與流暢度，進而提升其書寫技巧。褚家辰（2023）認為學習者透過書法來理解漢字的結構與起源，可以避免死記硬背，並能更有效理解和記住漢字。絲凱郁與張介英（2022）設計了 APP（E 筆）結合書法課程的創新教學模式，除了增強學生的學習動機，更提升學習成效與資訊科技素養，故未來可把 E 筆做為書法課程延伸應用之工具。。

2.2. 合作學習法

根據 Chen et al.(2018)的研究指出，合作學習能有效提升學生的學習動機和學業成就，也強調學生間的互動，對於共同建構知識相當重要，而在合作學習環境中使用適當的科技工具，可以提升學生發展高階認知能力。根據 Pyle et al.(2017)的研究，合作學習中的同儕介入策略（PMIs）對學術成就有顯著正向影響，尤其是在詞彙和理解能力方面。教師在設計合作學習活動時，應重視學生間的互動，故教師在分組時要有策略，考慮學生的能力差異，使學習成效最大化。

張清濱（2020）提到合作學習法為學習者互相合作，藉由小組共同討論的形式，讓學習速度較快的學生教導學習速度較慢的學生，發揮團隊精神，而教師則擔任諮詢者的角色。汪慧玲與沈佳生（2013）認為合作學習法相較於傳統教學法，較能提升學生的學習成效，除了師生互動增加外，學習態度也更加積極。王金國（2005）更提出運用合作學習法於國語課中，除了能增強合作學習的知能與信念，更同時也增進了國語科教學的能力。

2.2.1. 全班鷹架式合作學習

Walshaw 與 Anthony(2008)強調教師應該通過引導學生進行解釋、預測、辯論等活動來促進深度學習，促使學生在課堂上進行有意義的對話。而教學的成功更取決於教師如何有效地運用教學策略，並促進學生的深度學習與對話，故教師應重視課堂對話的重要性，並清楚如何透過良好的教學策略，促進學生的學習與社會發展。而有研究結果指出，教師在全班教學中提供適當的鷹架學習（Whole-Class Scaffolding, WCS），能夠有效促進學生的數學問題解決（Mathematical Problem Solving, MPS）(Abdu, 2015)。

特別是當教師透過運用電腦支持的合作學習（Computer-Supported Collaborative Learning, CSCL）時，不僅能促進學生間的互動與合作。透過這些科技平臺，教師更可運用策略示範、問題分解與學習如何一起學習（Learning to Learn Together, L2L2）的對話，使學生透過小組合作發展認知能力，提升問題解決技巧(Abdu, 2013; Smit et al., 2013)。在這樣的環境下，教師不僅是知識的傳遞者，更是合作學習的促進者。教師可以設計問題解決情境，鼓勵學生在小組內進行討論、分析問題並共同尋找解決方案，這樣的學習過程不僅有助於學生個人的認知發展，也能促進他們在團隊合作中的學習技巧。

3.研究方法

3.1. 研究對象

本研究之對象為臺灣南部某國小五年級學生。人數共 41 人，其中男生 22 人，女生 19 人，隨機擇定為實驗組與控制組。實驗組採用「結合全班鷹架與個人化回饋之合作學習模式」，而控制組則採用「結合個人化回饋之合作學習模式」。

3.2. 研究設計

本研究將採用準實驗設計，探討「全班鷹架式合作學習（Whole-Class Scaffolding Cooperative Learning）」與個人化即時回饋學習平臺對於提升學生書法學習成效、合作學習傾向及學習動機的影響。研究設計將採用前測與後測，並隨機將學生分為實驗組與控制組進行比較分析。研究架構如圖 1 所示：

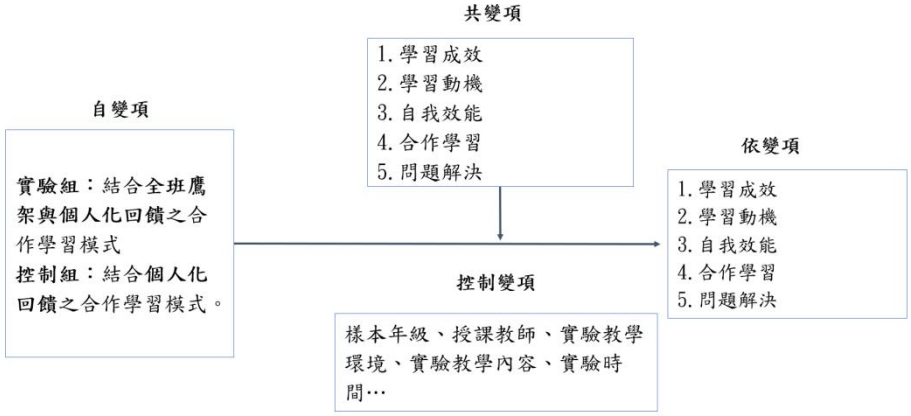


圖 1 *研究架構圖*

在實驗組的教學設計中，核心的策略是「學會一起學習（L2L2）」對話機制。這種對話機制將貫穿三個階段：前期討論（Preliminary L2L2 Talk）、常規討論（Routine L2L2 Talk）與總結討論（Summative L2L2 Talk）。學生將在小組內討論字形、筆順及結構等書法問題，並互相給予反饋，從而實現深度學習與合作。

為了增強學生在學習過程中的即時回饋，實驗組將使用數位工具來進行「筆順比對測試」之個人化即時回饋，這些工具可以幫助學生即時修正書寫錯誤，進而提高學習效率。而研究設計中的另一個關鍵是「分階段學習模式」，透過將書法練習分為三個階段—「拆解筆畫」、「部首組合」、「完整書寫」，幫助學生有條理地逐步學習書寫技巧，此模式不僅能增強學生對書寫細節的專注度，還能培養他們在書法練習中的耐心。學習活動流程圖如圖 2 所示：

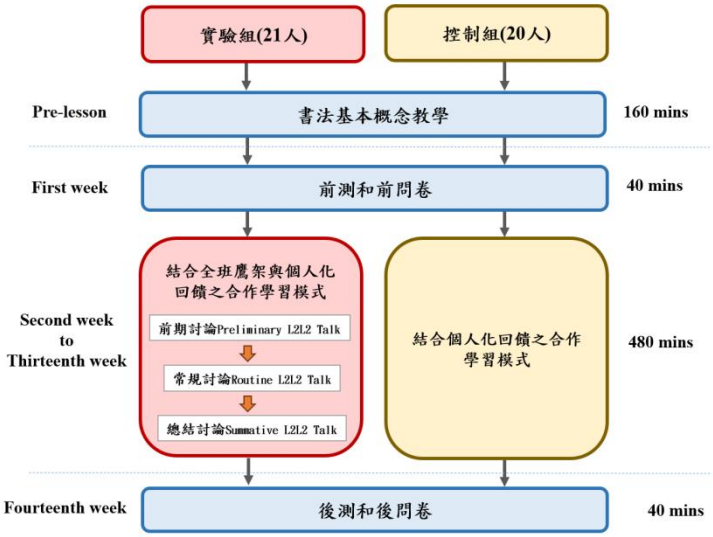


圖 2 學習活動流程圖

3.3. 數位教材設計

此研究將設計並應用一系列數位教材和工具，以促進學生的書法學習並提供即時反饋，為了實現筆順比對，研究將利用數位平臺及 E 筆，允許學生在其設備上即時比對自己書寫的字形與筆順。將結合視覺化功能，展示標準的筆順和字形結構，學生可以在這些比對資料的幫助下，發現和修正書寫中的錯誤。而這些數位工具會提供反饋機制，讓學生能夠在書寫過程中獲得即時的改進建議。為了提升學生的合作學習效果，利用線上合作平臺讓學生在進行 L2L2 對話，可以共享書法作品，並進行集體討論與回饋。接著，學生可以拍攝自己的書法過程，並利用數位標記工具對照他人的作品，提出建議或發現錯誤，不僅能促進學生之間的互動，還能加強學生對字形結構的理解。

3.4. E 筆系統

為了加強學生的字形學習，本研究將利用數位工具進行生字教學。學生可以在 E 筆平臺上學習新字，並通過操作界面進行筆畫練習及字形結構比對。平臺將顯示正確的字形結構，並提供引導，協助學生逐步掌握正確書寫技巧（筆順錯誤修正與步驟回溯功能）。學生可以隨時檢查自己的書寫成果，並在即時反饋的指導下進行修正，從而有效提升生字學習的效果。

運用數位學習平臺及 APP（E 筆）促進學生的書法學習並提供即時回饋。為了實現筆順比對，學生可以在設備上即時比對自己書寫的字形與標準筆順。E 筆除了能展示標準的筆順和字形結構，更結合評語功能，使學生能夠透過比對資料，發現並修正書寫中的錯誤。該系統提供即時反饋機制，讓學生在書寫過程中獲得改進建議，進而提高書寫準確度。

3.5. 研究工具

3.5.1. 學習成效

本研究參考 2024 年臺灣語文競賽寫字比賽之命題及評分標準，以評估學習者的寫字能力。評分標準包含筆法、結構與章法、以及正確性與書寫速度三個面向。為確保評分的公正性與信效度，本研究採用專家評分法，由 2 位具備書法評審經驗的專家進行評定，並計算平均分數作為最終得分。整體評分標準參照全國語文競賽，期望能有效測量學生的書寫能力，並作為後續教學成效分析的基礎。

後測評量採用書法期末評量，根據學生學習進度、課程目標及全國語文競賽評分標準共同編製，以確保評量題目符合學習內容並具適切性。為提升評量的專業性與公信力，題目編

製完成後再交由二位具備專業書法背景的專家進行審核，確認評量內容在字體選擇、難易度及評分標準方面皆符合書法教學目標，並確保試題的公平性與信效度。最後本研究將根據此評量結果，與前測成績進行比較分析，以探討實驗介入對學生書寫能力的影響，進一步檢驗本研究之教學模式是否有效提升學生的書法表現。

3.5.2. 學習動機

採用內在動機量表（Intrinsic Motivation Inventory, IMI）作為測量學生學習動機的工具。該量表最初由 McAuley et al.(1989)發展，後經 Yin et al.(2021)根據數位學習情境進行修訂與應用。本研究依據 Yin et al.(2021)的版本，根據研究需求適度調整題目內容，以確保適用於本研究的教學情境。內在動機量表涵蓋五個主要構面，包括興趣與樂趣、緊張與壓力、感知選擇、感知能力以及感知價值，共 25 題，並採用 5 點李克特量表（Likert scale, 1=非常不同意，7=非常同意）進行測量，而本研究將於學習前後進行施測，以比較不同教學模式對學生內在動機的影響。

3.5.3. 自我效能

採用群體自我效能量表（Collective Self-Efficacy Scale）測量學生在合作學習環境中的自我效能。該量表源自 Pintrich et al.（1991）所發展的「學習動機與策略問卷（Motivated Strategies for Learning Questionnaire, MSLQ）」，並由 Wang 與 Lin(2007)進一步修訂，以適用於電腦輔助合作學習（Computer-Supported Collaborative Learning, CSCL）情境。本研究依據 Wang 與 Lin(2007)的版本，並根據研究需求適度調整，以確保適用於本研究的教學環境。本量表包含八個題目，採用 5 點李克特量表（Likert scale, 1=非常不同意，5=非常同意），主要用於測量學生對於小組合作學習的信心與預期成效，而本研究將於學習前後進行施測，以探討不同教學模式對於學生群體自我效能的影響。

3.5.4. 合作學習傾向及問題解決

為探討不同學習模式對學生合作學習與問題解決能力的影響，採用 Lai 與 Hwang(2014)量表，包含「合作學習傾向」與「問題解決傾向」，透過 5 點李克特量表評估學生在不同教學模式下的行為與認知變化。合作學習傾向測量學生在小組學習中的合作態度與行為，包括對團隊成員的信任、合作信心、溝通品質及任務分工合理性。本研究將比較在實驗組與控制組這兩種合作學習模式下，對學生合作能力的影響。而問題解決傾向評估學生解決問題的信心與態度，包含面對挑戰的積極性、解決經驗及不逃避問題的程度。並透過前後測分析不同學習模式對學生問題解決能力的影響，最後透過此量表分析不同學習模式對學生合作學習行為及問題解決能力的影響。

4.預期研究結果

本研究將比較全班鷹架式合作學習與傳統合作學習模式對國小高年級學童書法學習的影響，預期實驗組在學習成效、學習動機、自我效能、合作學習傾向及問題解決能力方面皆能有優於控制組之表現。

參考文獻

王金國（2005）。共同學習法之教學設計及其在國小國語科之應用。*屏東師院學報*，22，103–130。

汪慧玲、沈佳生（2013）。合作學習教學策略對大專學生之學習成效與學習態度之影響：以兒童發展評量與輔導課程某單元為例。*臺中教育大學學報：教育類*，27 卷 1 期（2013/06），57-76。

張清濱（2020）。*教學原理與實務*（第二版）。五南。

絲凱郁、張介英（2022）。數位 e 筆融入書法教學實踐。*北市大語文學報*，第 27 期，51-68 頁。臺北市立大學中國語文學系。

褚家辰（2023）。*書法藝術及其在華語文教學中應用之研究*(博士論文)。東吳大學。

Abdu, R., Schwarz, B., & Mavrikis, M. (2015). Whole-class scaffolding for learning to solve mathematics problems together in a computer-supported environment. *ZDM Mathematics Education, 47*(7), 1163–1178.

Abdu, R. (2013). Peer scaffold in mathematical problem solving. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *Proceedings of the computer supported collaborative learning conference* (pp. 2–10). Wisconsin: Madison.

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The Role of Collaboration, Computer Use, Learning Environments, and Supporting Strategies in CSCL: A Meta-Analysis. *Review of Educational Research*, *88*(6), 799–843. http://www.jstor.org/stable/45277266

Hobbs, R., & Frost, R. (2003). Measuring the acquisition of media-literacy skills. *Reading Research Quarterly, 38*(3), 330–355. https://doi.org/10.1598/RRQ.38.3.3

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.

Levin, D., & Arafeh, S. (2007). The digital disconnect: the widening gap between internet savvy students and their schools. American Institutes for Research for the Pew Internet & American Life Project. Washington, DC: Pew Charitable Trust. http://www.pewinternet.org/~/media//Files/Reports/2002/PIP_Schools_Internet_Report.pdf.pdf

Li-Tsang, C. W. P., Li, T. M. H., Yang, C. N., Cheung, P. P. P., Au, K. Y., Chan, Y. P., Cheung, K. Y., Ho, K. H., Kwok, K. W., & Leung, H. W. H. (2023). Evaluation of a group-based sensorimotor intervention programme to improve Chinese handwriting of primary school students. *Heliyon, 9*(1), e12554. https://doi.org/10.1016/j.heliyon.2022.e12554

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*(1), 48–58.

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).

Pyle, D., Pyle, N., Lignugaris/Kraft, B., Duran, L., & Akers, J. (2017). Academic Effects of Peer-Mediated Interventions With English Language Learners: A Research Synthesis. *Review of Educational Research*, *87*(1), 103–133. http://www.jstor.org/stable/44667650

Smit, J., Van Eerde, H. A. A., & Bakker, A. (2013). A conceptualization of whole-class scaffolding. *British Educational Research Journal, 39*(5), 817–834.

Walshaw, M., & Anthony, G. (2008). The Teacher's Role in Classroom Discourse: A Review of Recent Research into Mathematics Classrooms. *Review of Educational Research*, *78*(3), 516–551. http://www.jstor.org/stable/40071136

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.

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.

因材網 e 度應用於數學文字題之成效分析-以小學五年級補救教學學童為例

An Analysis of the Effectiveness of Using the Generative Artificial Intelligence on Mathematical Word Problems: A Case Study of Fifth-Grade Remedial Instruction Students

胡振猷[1]，林秋斌[1]，楊肅健[2*]
[1]台灣清華大學 學習科學與科技研究所
[2]台灣清華大學 跨院國際博士班學位學程
*sujiann@gmail.com

【摘要】數學文字題包含數學運算與閱讀理解，對於低學習成就學生而言，往往是學習上的一大挑戰。本研究融入生成式 AI 因材網 e 度工具，透過其蘇格拉底提問式教學法，引導學生理解題目並培養數學解題能力。研究對象為某國小五年級補救教學學生，採單組前後測設計，前 9 週由教師用蘇格拉底提問式教學法進行數學文字題教學，後 9 週則加入因材網 e 度進行數學文字題學習。研究結果顯示，因材網 e 度能夠提升國小五年級補救教學學生於數學文字題上的學習成就，其整體學習態度皆有提升，特別在學習焦慮方面，顯示學生在解數學文字題時的焦慮感顯著降低。
【關鍵字】因材網；補救教學；數學文字題；蘇格拉底提問式教學法

Abstract: *Mathematical word problems involve both mathematical operations and reading comprehension, which often pose a significant challenge for students with low academic achievement. This study integrates generative AI (Adaptive Learning Platform, ALP) into instruction, using its Socratic questioning approach to guide students in understanding problems and developing problem-solving skills in mathematics. The participants were fifth-grade remedial students from an elementary school. A one-group pretest-posttest design was adopted: during the first 9 weeks, instruction on mathematical word problems was delivered by the teacher using the Socratic questioning method; in the following 9 weeks, the generative AI (ALP) was incorporated into the learning process. The results showed that the use of generative AI (ALP) effectively enhanced students' achievement in solving mathematical word problems. Overall, students' learning attitudes improved, with a notable reduction in learning anxiety, indicating that their anxiety when solving mathematical word problems significantly decreased.*
Keywords: Generative AI, Mathematical Word Problems, Remedial Instruction, Socratic Method

1. 緒論

數學文字題又稱數學應用問題，是以文字及數字描述的一種題型。數學文字題能夠將日常生活情境融入問題中，因此學生較能理解問題情境，是數學學習中重要的一環(涂金堂，2007)。美國國家數學教師協會 NCTM (2000)指出，數學教育的核心目標之一為「問題解決」，期望學生能利用數學概念，解決現實世界中的問題。現今談及數學素養著重與現實生活的連結，強調數學的應用性與實用性(郭明田、溫嫩純，2021)。數學素養不僅關乎計算能力，更注重在不同情境中靈活運用數學知識(OECD, 2018)，而 PISA 測驗的題型中多以數學文字題呈現，便是為了評估學生在生活中運用數學的能力。研究者認為，數學文字題不僅是課程內容和教學中的核心要素，更是通往素養導向學習的關鍵，然而低成就學生常常因為多

重障礙，致造成學習成績低落。本研究主要針對數學低成就學生在解題數學文字題上所遇到之困難，特別關注在學生理解問題的能力，並希望透過生成式 AI-因材網 e 度的使用，幫助學生解題。本研究欲探討的研究問題為：(1)運用生成式 AI-因材網 e 度的使用是否能提升學生學習成效?(2) 生成式 AI-因材網 e 度的教學模式，學生對於數學的學習態度是否帶來正向的影響?

2. 文獻探討

2.1. 蘇格拉底提問式教學法

蘇格拉底提問式教學法又稱「詰問法」、「產婆法」，是一種透過問題對話，逐步引導學生思考的教學方法。這種方法主要強調教師不直接提供答案，而是透過連續性的問題，幫助學生思考，從而建構知識。蘇格拉底提問式教學適用於醫學、法律、數學等多種學科，可以促進學生批判思考和問題解決能力(Wilberding, 2019)。透過蘇格拉底提問式教學法，能夠促進學生的高階思維能力、提升自主學習與自我改進能力、增強溝通與表達能力，也符合現代社會中，透過對話過程，學生能培養傾聽欣賞別人不同的意見、接納不同意見和觀點的態度(Dalim et al., 2022; Delic & Becirovic, 2016; Lam, 2011)。多項研究亦指出，蘇格拉底提問式教學法有助於提升閱讀理解能力(Acim, 2018; Mahfuddin, 2019; Mahmud & Tryana, 2023)，因此本項研究採用蘇格拉底提問式教學法於補救教學學生的數學文字題學習。

2.2. 生成式 AI 教學應用

近年來，生成式 AI 如雨後春筍般出現，正在逐漸推動教育的變革和發展。生成式 AI 指的是能夠生成文本、圖像、音訊等內容的人工智慧技術。在生成式 AI 平台中，數學文字題的解題處理能力方面展現了一定的應用潛力，是目前教育新興的應用場景。以 ChatGPT 而言，其在數學文字題方面，具有良好靈活性，使用者將數學文字題輸入後，可以與 ChatGPT 進行互動，並逐步引導解題，而後提供答案，有助於學生自主學習。透過生成式 AI 的生成鷹架輔助學生，有助於提高學生學習效果，並改善學生語言與認知能力(Singh et al., 2024)。不過，ChatGPT 在引導使用者解決問題的過程中，容易出現計算錯誤或是理解錯誤，根據 Shakarian et al. (2023)的研究指出，ChatGPT 在處理文字題時，其表現會因是否要展示解題過程而有顯著變化。當需要展示解題過程時，ChatGPT 的錯誤率為 20%，而不需要展示時則高達 84%，因此在引導解題過程中，使用者需多次自行確認與修正，才能得到正確的結果。

因材網打造了一款生成式 AI 學伴- e 度，其以對話引導形式為學生提供一對一的教學及輔導，提出個性化問答以引導學生深度學習，並能協助教師制定課程計劃和批改作業，為每一位學生定制學習方案。e 度能根據學生的回答和需求，動態調整教學內容和策略，從而更有效地幫助學生理解題意，並提升他們解決文字題的能力，這種互動式的學習方式，不僅可以提高學生的數學成績，還能增強他們的學習信心和動機。

2.3. AI 提問類型

提問是我們與 AI 溝通的橋樑，我們所提供的「提示詞(Prompt)」是否明確、具體且有邏輯，關乎 AI 的回應品質，有效的提問能夠顯著提升 AI 的回應質量。Liu et al. (2023)將使用者與 AI 的對話類型分成六大類：(1)搜尋查詢(Search Query)：直接輸入問題，以獲得答案。(2)漏斗式對話(Funneling Conversations)：先提出較大方向的問題，再根據 AI 的回應補充問題。(3)探索性對話(Exploring Conversations)：使用者與 AI 進行開放式對話，旨在探討某個主題或領域。(4)雕刻式對話(Chiseling Conversations)：使用者根據 AI 的回覆，反覆修改問題，以獲得更符合需求的回應。(5)擴展性對話(Expanding Conversations)：根據 AI 的回答，提問要求進一步的細節與內容。(6)精確性對話(Pinpointing Conversations)：提問內容非常具體，要求明確答案。本研究參考此分類架構，並透過分析學生使用 e 度的歷程數據，進一步歸納出五種主要的學生提問類型如表 5 所示。

Ahn et al. (2024)亦將使用者回應 AI 的態度分為四大類型：(1)無回應(No Verbal Input or Response)：使用者未回應 AI，保持沈默。(2)關鍵字回應(Keywords)：使用者使用單一或少數關鍵字回應 AI。(3)自然對話(Natural and Casual)：使用者使用完整句子回應 AI。(4)命令式回應(Giving a Demand)：使用者直接使用命令句回應 AI。本研究參考此分類架構，並透過分析學生使用 e 度的歷程數據，進一步歸納出四種主要的學生回應類型如表 6 所示。

3. 研究方法

3.1. 研究流程

本研究採用單組前後測設計法，研究時程共十八週，旨在探討生成式 AI-因材網 e 度對五年級補救教學學生於數學文字題的學習成效及學習態度之影響。研究對象為經「國民小學及國民中學學生學習扶助科技化評量」教學未通過的五名五年級數學低成就學生，前九週教師以蘇格拉底提問式進行數學文字題教學，並進行文字題前測及後測，以了解其成效並分析學生錯誤類型。後九週則使用因材網 e 度進行文字題教學，期間進行前後測，並加入數學文字題學習態度量表施測，以探討學生在學習態度上的變化。實驗結束後，進行學生訪談，透過質性分析深入了解學生使用 e 度的情形。

3.2. 研究工具

本研究教材內容採用因材網(https://adl.edu.tw/)，網站上 e 度連結了 ChatGPT，學生可以打字或語音方式輸入數學文字題，e 度將透過蘇格拉底對話方式，逐步引導學生思考與解題。學生在使用 e 度的過程中，所有的對話紀錄都會被自動保存下來，教師可到後台即時查看學生的學習歷程，掌握使用狀況及其解題過程。而第一、二次前後測，試卷乃參考學習扶助科技化平台歷年的測驗卷，搜羅出與數學文字題相關題目，並將其數字稍作修改，經過預試篩選成最終試題各 10 題二張，內容具一致性及可靠性。數學學習態度量表乃參考徐揚智(2020)「數學科領域學習態度量表」和康雅芳(2006)「國小高年級數學態度量表」進行改編，量表分為數學自信、數學有用性、學習動機、學習焦慮四個面向，各分量表的 Cronbach's α係數依序為 0.89、0.9、0.91、0.88，解釋量分別為 56.57%、53.45%、52.29%、63.34%，每個面向五題，共二十題，並採李克特四點量表計分。

4. 結果分析

4.1. 前九週教師採蘇格拉底提問式教學之分析

研究者在經由五位學生前測結果，將數學錯誤類型分成數學概念不足、計算錯誤、不理解題意、學習態度所致四大類，經過蘇格拉底提問式教學法，後測之後再分析其數學錯誤類型，其結果如表 1 所示。前後測分析比較，清楚的數學概念有所提升；計算錯誤則比率一樣；反而不理解題意和學習態度不佳比率增多。

表 1 第一次前/後測錯題分析

試題內容/學生	甲生		乙生		丙生		丁生		戊生	
	前測	後測	前測	後測	前測	後測	前測	後測	前測	後測
正方形面積計算	A	o	A	A	A	A	A	A	A	o
同分母分數整數倍計算	o	o	A	C	o	o	o	o	A	B
較大位數除法直式計算	o	o	o	o	D	C	o	C	o	B
較大位數乘法直式計算	B	o	D	B	B	o	C	o	o	o
同分母分數減法計算	o	o	A	A	A	C	o	D	B	A

較大位數乘法直式計算	o	o	o	D	B	o	o	o	o	B
較大位數除法直式計算	?	o	o	D	B	D	o	D	o	o
長方形面積計算	A	o	o	o	A	A	o	o	o	o
同分母分數減法計算	o	o	A	D	A	A	D	o	o	A
同分母分數減法計算	o	o	B	o	B	D	D	C	B	o
總分	60	100	40	30	10	30	60	50	60	50

註：A 數學概念不足、B 計算錯誤、C 不理解題意、D 學習態度所致、o 代表通過、?為無法判定

4.2. 後九週 e 度融入進行文字題教學之分析

研究者分析五位學生前測結果，將數學錯誤類型分成數學概念不足、計算錯誤、不理解題意、學習態度所致四大類，經過 e 度融入進行文字題教學，後測之後再分析其數學錯誤類型，其結果如表 2 所示。比較分析結果，整體而言，所有學生的錯誤類型皆有減少，特別是歸責於不理解題意的錯誤明顯下降，顯示學生在使用 e 度後，不僅分數有所提升，在題意理解、學習態度、計算能力與數學概念上皆有正向改變。

表 2 第二次前/後測錯題分析

試題內容/學生	甲生		乙生		丙生		丁生		戊生	
	前測	後測	前測	後測	前測	後測	前測	後測	前測	後測
連除兩步驟	o	C	C	C	C	o	o	C	C	C
乘法結合律	A	o	C	o	B	o	o	o	B	o
單位複名數計算	o	o	B	o	B	o	o	o	o	o
單位複名數計算	A	o	C	B	o	o	A	A	D	o
單位複名數計算	A	o	D	B	A	B	A	o	D	B
單位複名數計算	o	o	C	o	B	o	o	o	D	o
單位複名數計算	D	o	D	o	B	B	A	o	A	o
單位複名數計算	C	o	C	o	D	o	B	o	C	o
單位複名數計算	o	o	C	o	o	o	C	o	o	o
乘法結合律	o	o	o	o	C	B	C	B	o	o
總分	50	90	0	70	20	70	40	50	30	80

註：A 數學概念不足、B 計算錯誤、C 不理解題意、D 學習態度所致、o 代表通過

4.3. 學習成效

第一次和第二次前後測進行無母數統計分析，其結果如表 3 所示，前九週教師口頭蘇格拉底式對話進行數學文字題教學，學生前後測的學習成效並無顯著差異；而後九週生成式 AI-因材網 e 度對於補救教學學生在數學文字題上的學習成效，其 z 值等於-2.032，p=.042，達顯著差異。

表 3 學習成效無母數統計分析

	第一次前後測					第二次前後測			
項目	人數	平均數	標準差	z 值	*p* 值	平均數	標準差	z 值	*p* 值
前測	5	46	21.91	-.412	.680	28	19.24	-2.032	.042*
後測	5	52	28.64			72	14.83		

註：*p<.05

4.4. 學習態度

學習態度調查分成數學自信、數學有用性、學習動機與學習焦慮四個面向，每個面向各有五題，共計二十題，其描述性統計及無母數統計結果如表 4 所示，在數學自信方面，前測平均分數為 2.64 分，屬於中等水平，部分學生仍在數學文字題上較無信心，然而，後測分數提升至 3.12 分，學生在使用 e 度後，對數學文字題的信心有所增加。此外，根據訪談結果，受訪學生都認為，在使用 e 度後，信心提升很多，不再排斥解數學問題。在數學有用性方面，前測平均分數為 2.88 分，後測分數則提升至 3.32 分，結果顯示學生在使用 e 度後，更能認為數學文字題是有用處的，並在生活中具有應用價值。在學習動機方面，前測時平均分數為 2.68 分，屬中等水平。後測分數提升至 3.36 分，結果顯示學生在使用 e 度後，對數學文字題的興趣有所提升，更願意主動練習。另外，訪談結果更進一步支持此發現，受訪學生表示，在遇到不會的題目時，他們會選擇使用 e 度進行解題，並會主動進一步練習。在學習焦慮方面，前測時的平均分數為 2.72 分，後測下降至 1.16 分，結果顯示，學生在解數學文字題時的焦慮感明顯降低，並達顯著差異。透過訪談分析發現，受訪學生對使用 e 度的體驗持正向態度，認為 e 度在引導過程中更有耐心，使學習氛圍更為輕鬆愉快，進而減少學習焦慮，與前述量化結果相呼應。整體而言，學生使用 e 度學習後，在數學自信、數學有用性、學習動機皆有所提升，尤以學習焦慮在前測與後測之間呈現顯著差異(p=.042<.05)，學習焦慮明顯降低，對於學習態度帶來正向影響。

表 4 學習態度量表統計分析

面向		題數	平均數	標準差	題平均數	z	p
數學自信	前測	5	13.2	2.59	2.64	-1.461	.144
	後測	5	15.6	0.89	3.12		
數學有用性	前測	5	14.4	1.34	2.88	-1.633	.102
	後測	5	16.6	1.95	3.32		
學習動機	前測	5	13.4	3.05	2.68	-1.604	.109
	後測	5	16.8	2.49	3.36		
學習焦慮	前測	5	13.6	3.78	2.72	-2.032	.042*
	後測	5	5.8	0.84	1.16		
量表總分	前測	20	54.6	0.80	2.73		
	後測	20	54.8	0.80	2.74		

註：*p<.05

4.5. e 度提問類型分析

本研究參考 Liu et al. (2023)將使用者與 AI 的對話類型分成直接提問型、探索提問型、嘗試提問型、主動學習型提問、非學習型提問五類，紀錄了五位學生在使用 e 度過程中的提問次數與類型分佈如表 5 所示，學生在使用 e 度的過程中，以直接提問型比例最高，所有的學生主要提問方式皆是直接輸入問題讓 e 度引導解題。探索型提問及嘗試型提問比例相近，主動學習型提問的比例較低，非學習型提問最少，顯示他們在使用 e 度時，很少出現與學習無關的提問。

表 5 學生提問類型表

學生	直接型	探索型	嘗試型	主動學習型	非學習型	總計
甲生	21(41.8%)	7(16.3%)	8(18.6%)	4(9.3%)	3(7.0%)	43(100%)
乙生	21(44.7%)	8(17.0%)	9(19.1%)	5(10.6%)	4(8.5%)	47(100%)
丙生	18(46.2%)	8(20.5%)	6(15.4%)	4(10.3%)	3(7.7%)	39(100%)
丁生	17(41.5%)	8(19.5%)	7(17.1%)	6(14.6%)	3(7.3%)	41(100%)
戊生	16(40.0%)	9(22.5%)	6(15.0%)	6(15.0%)	3(7.5%)	40(100%)

4.6. 學生回應類型分析

學生在輸入問題後，仔細閱讀 e 度回應的自然對話，本研究紀錄了五位學生在使用過程中的回應次數與類型分佈如表 6 所示，學生在回應 e 度時，以積極型回應比例最高（74.4%~90.2%），所有的學生主要回應方式都以積極型為主，顯見他們在使用 e 度時，多數情形下能主動參與、接受引導並作出相對回應。被動回應型（4.9%~11.6%）與忽略回應型（4.9%~14.0%）比例則相對較低，學生在使用 e 度時偶爾回應方式較為被動，或是偶爾會直接忽略 e 度的引導，未作出回應。困惑型回應的比例最低（0.0%~4.3%），大部分學生在與 e 度互動時，都能理解 e 度的引導，僅有乙生（4.3%）與戊生（2.5%）偶爾會對 e 度表現出困惑。

從 e 度的回應面來看，其能夠以「直接給答案」、「引導思考」、「階段回應」、「錯誤診斷」多種回應方式，幫助學生循序漸進地理解題目並解出答案。然而，在少數情況下，e 度仍有可能出現「回答錯誤」的狀況，略影響學生學習效果。整體而言，e 度在學生學習數學文字題的過程中發揮一定的正向影響，學生能夠透過 e 度的引導解題並修正錯誤，並能主動使用 e 度來詢問不會的數學題。然而，學生的提問方式與回應態度可能會影響其學習效果，因此，未來得進一步探討如何幫著學生設計良好提問，以達更有效的學習。

表 6 學生回應類型表

學生	積極型	被動型	忽略型	困惑型	總計
甲生	32(74.4%)	5(11.6%)	6(14.0%)	0(0.0%)	43(100%)
乙生	38(80.9%)	4(8.5%)	3(6.4%)	2(4.3%)	47(100%)
丙生	33(84.6%)	3(7.7%)	3(7.7%)	0(0.0%)	39(100%)
丁生	37(90.2%)	2(4.9%)	2(4.9%)	0(0.0%)	41(100%)
戊生	31(77.5%)	4(10.0%)	4(10.0%)	1(2.5%)	40(100%)

5. 結論

5.1. 學習成就

學生在使用 e 度進行數學文字題學習後，其平均分數明顯提升，並達顯著差異，表示 e 度能夠幫助補救教學學生提升數學文字題的理解與解題能力，進而提升學習成效。在錯題分析，結果顯示學生在後測中的錯誤類型明顯減少，在「不理解題意」、「數學概念不足」、「計算錯誤」與「學習態度所致」方面皆少於前測，其中，「不理解題意」的錯誤減少最多，顯示 e 度的引導能夠幫助學生理解題意及解題，進而提升學習成效。

5.2. 學習態度

本研究透過數學文字題態度量表，從數學自信、數學有用性、學習動機和學習焦慮四面向，探討 e 度是否能影響學生在數學文字題上的學習態度。研究結果顯示，學生在使用 e 度進行學習後，整體學習態度皆有提升，特別在學習焦慮方面，達顯著差異，顯示學生在解數學文字題時的焦慮感顯著降低。

參考文獻

涂金堂(2007)。國小學生數學文字題問題結構與數學解題表現之相關研究。屏東教育大學學報，26，101-136。

徐揚智(2020)。應用數位遊戲式學習於國小六年級數學低成就學生數學課後複習學習成效之研究 —以 PaGamO 數位遊戲式平台為例［未出版碩士論文］。國立清華大學。

康雅芳(2006)。國小五年級學童認知風格、數學態度與數學解題表現之關係研究［未出版碩士論文］。國立高雄師範大學。

郭明田、溫嫩純(2021)。國中數學素養導向教學設計與學習成效之行動研究。臺灣教育評論月刊，10(10)，196-228。

Acim, R. (2018). The Socratic method of instruction: An experience with a reading comprehension course. *Journal of Educational Research and Practice, 8*(1), 41–53. https://doi.org/10.5590/JERAP.2018.08.1.04

Ahn, S.-Y., Kim, S., Lee, E.-Y., Lee, H., Song, E.-Y., Song, J., & Won, J. (2024). *How do AI and human users interact? Positioning of AI and human users in customer service*. Text & Talk. https://doi.org/10.1515/text-2023-0116

Dalim, S. F., Ishak, A. S., & Hamzah, L. M. (2022). Promoting students' critical thinking through Socratic method: Views and challenges. *Asian Journal of University Education, 18*(4), 1034-1047. https://doi.org/10.24191/ajue.v18i4.20012

Delic, H., & Senad, B. (2016). Socratic method as an approach to teaching. *European Researcher, 111*(10), 511-517. https://doi.org/10.13187/er.2016.111.511

Lam, F. (2011). *The Socratic method as an approach to learning and its benefits* [Unpublished honors thesis]. Dietrich College of Humanities and Social Sciences, Carnegie Mellon University.

Liu, F., Budiu, R., Zhang, A., & Cionca, E. (2023). *ChatGPT, Bard, or Bing Chat? Differences among 3 generative-AI bots*. Nielsen Norman Group. https://www.nngroup.com/articles/ ai-bot-comparison/

Mahfuddin, M. H. (2019). *The effect of Socratic questioning on students' reading comprehension of analytical exposition text: A quasi-experimental study at eleventh grade of SMA PGRI 56 Ciputat* [Undergraduate thesis]. Syarif Hidayatullah State Islamic University.

Mahmud, L. H., & Tryana. (2023). Promoting reading comprehension by using Socratic questioning. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra, 9*(1). https://doi.org/10.30605/onoma.v9i1.2221

NCTM (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

OECD (2018). *Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework.* https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/global-competence/Handbook-PISA-2018-Global-Competence.pdf

Shakarian, P., Koyyalamudi, A., Ngu, N., & Mareedu, L. (2023). An independent evaluation of ChatGPT on mathematical word problems (MWP). *CEUR Workshop Proceedings*, 3433.

Singh, Y., Singh, D. P., Chander, N., Singh, Y. P., & Singh, P. R. (2024). Generative AI for enhancing education and skill development. *International Journal of Innovative Science and Research Technology*, 883–888. https://doi.org/10.38124/ijisrt/ijisrt24nov431

Wilberding, E. (2019). *Socratic methods in the classroom: Encouraging critical thinking and problem solving through dialogue*. Prufrock Press.

基於雙層次測驗為基礎之博物館環景導覽對學習者學習動機與成效之影響

Effects of incorporating a two-tier assessment-based guiding mechanism into museum panorama tours on learners' learning motivations and achievements

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】隨著 COVID-19 疫情的影響，博物館環景虛擬導覽已成為推廣博物館教育與學習體驗的新趨勢，但現有的環景導覽大多缺乏有效的評量機制，難以掌握學習者的學習狀況，導致無法提供即時且適切的回饋與指導。本研究整合雙層次測驗機制與 720 度環景技術，開發了一套基於雙層次測驗之博物館環景導覽學習系統，以探究學習者對博物館展品內容的認知理解程度，並即時提供回饋與補救引導。研究結果顯示，此系統能有效提升學習者的學習動機與學習成效且不會顯著增加學習者的認知負荷。
【關鍵字】博物館學習；雙層次測驗；環景導覽；學習成就

Abstract: *With the impact of the COVID-19 pandemic, virtual panoramic tours have emerged as a key strategy for promoting museum education and enriching learning experiences. However, most existing virtual tour systems generally lack effective assessment mechanisms, making it difficult to accurately diagnose learners' understanding and thus failing to provide timely and appropriate guidance and feedback. Therefore, this study proposes an interactive virtual tour system integrating a two-tier assessment mechanism with 720-degree panoramic technology. The developed system aims to explore learners' cognitive understanding of museum exhibit contents and provides immediate feedback and remedial guidance. Experimental results indicate that this two-tier assessment-based virtual panoramic tour system can effectively enhance learners' motivation and learning performance without significantly increasing their cognitive load.*
Keywords: museum learning, two-tier assessment, virtual panoramic tour, learning achievement

1. 前言

近年來，COVID-19 疫情的爆發與延續，促使全球各博物館加速數位化發展，虛擬環景導覽技術因而成為博物館推動遠距學習與提供創新學習體驗的重要趨勢 (Lee, Yoon, & Jung, 2023)。環景技術能打破實體參觀的時間與空間限制，使博物館教育更具彈性與可及性，也增加了使用者對博物館資訊與展品的接觸機會 (Chen & Yang, 2022; Li, Zheng, Watanable & Ochiai, 2024)。然而，儘管環景虛擬導覽已逐漸普及，多數系統仍停留在展示層面，缺乏有效的學習診斷與評量機制，導致博物館教育難以即時掌握學習者的真實理解狀態，導致學習成效的提升有限 (Wang, Zhang, & Yang, 2022)。因此，如何透過適切的評量方法，診斷並改善虛擬導覽中的學習成效，已成為當前博物館教育領域的重要研究議題。為解決此問題，本研究嘗試將雙層次測驗與 720 度環景技術結合，建構一套兼具學習診斷與互動回饋功能的博物館虛擬導覽系統。此系統不僅讓學習者能更深入地互動和理解博物館的展品資訊，更透過即時的診斷機制，幫助學習者修正錯誤概念並加深學習印象，從而提升整體的學習效益。

2. 基於雙層次測驗之環景導覽學習系統

本研究以自然生態為學習主題，透過任務導向的學習設計，引導學習者透過與虛擬環景及學習系統的互動，逐步建構對自然生態的整體理解，並於每項目標完成後實施學習評量。本系統架構包含：

(1) 720 度環景互動模組：結合 3D 空間掃描與虛擬實境建構技術，實現實體展覽場域之數位孿生轉化，進一步營造沉浸式遠距教學環境。此技術應用有效突破時空的限制，提升觀眾對展覽內容的理解深度與遠距互動參與度，讓學習者即使未親臨現場，亦能獲得貼近實境的學習體驗。虛擬展場模組提供可自主探索的數位展覽情境，學習者可於沉浸式空間中自由移動，並透過點擊互動機制即時獲取展品資訊。

(2) 雙層次測驗評量模組：本系統在導覽過程中導入雙層次評量機制，以掌握學習者的學習歷程。第一層問題用以瞭解學習者對展品內容知識的掌握，第二層問題則針對其答題理由進行探究學習者解題的想法，進一步剖析其背後可能的迷思概念。

(3) 即時回饋與引導模組：根據雙層次測驗模組之診斷結果，提供學習者引導性提示與補救教材，以協助釐清錯誤概念，深化對展品知識之理解，進而提升整體學習成效與知識內化程度。

圖 1 *學習系統介面及流程*

3. 實驗設計

本實驗以某大學通識課程 60 名學生為研究對象，隨機分派為實驗組與控制組。實驗組採用基於雙層次測驗之環景導覽學習模式，控制組採用傳統環景導覽學習，兩組的任務與學習內容相同。在實驗開始之前，所有學生已學過自然生態通識課程的基本知識與概念。之後，進行前測與填寫前測問卷，以確認兩組在先備知識上的一致性。在學習活動開始時，學生接受了 20 分鐘的環景導覽系統操作教學，以確認實驗組與控制組對學習系統有一定程度的了解。在正式學習過程中，每位學生使用平板進行 100 分鐘的學習活動。活動結束後，兩組學生花費 40 分鐘完成後測與後測問卷，以檢測他們是否真正理解知識內容。本研究學習動機問卷取自 Wang 及 Chen (2010)所開發的五點式李克特動機量表。其問卷包含各 3 題內在動機與外在動機，對應的 Cronbach's α分別為 0.65 和 0.68。認知負荷問卷則採用自 Hwang 等人 (2013)之研究，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 α分別為 0.87 和 0.86。

4. 實驗結果

4.1. 學習成就分析

為了評估實驗組與控制組學生在學習成就上是否有顯著差異，本研究採用單因子共變數分析 (ANCOVA)，並以前測成績為共變項、後測成績為依變項進行分析。透過兩組成績進行 Levene 變異數同質檢定來探討兩組學生在先備知識上是否有顯著差異，其結果顯示未違反同質性假設($F = 1.61, p > 0.05$)。此外，迴歸係數同質性檢定結果($F = 2.87, p > 0.05$)也未達顯著水準，代表兩組迴歸線之斜率相同，兩組學生的先備知識相近，符合共變數分析之基本假設，可進行 ANCOVA 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 68.85 和 1.32，控制組為 64.15 和 1.32。實驗結果發現兩組在學習成就上有顯著差異($F = 5.88, p < 0.05$)，這也表示「基於雙層次測驗為基礎之環景導覽學習模式」的實驗組，比使用「傳統環景導覽學習模式」的控制組更有助於提升學生的學習成就，使他們能更好地理解學科知識。

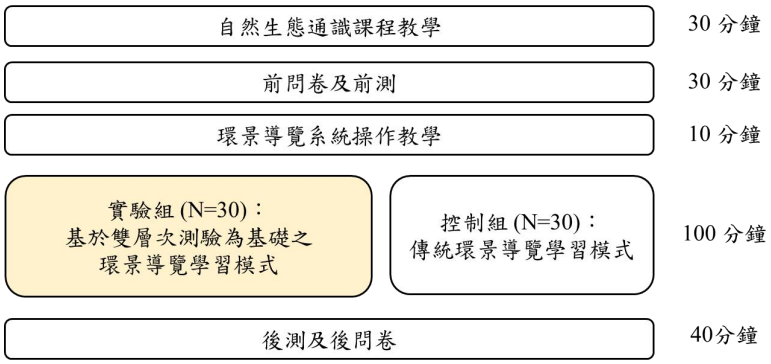


圖 2 實驗流程圖

4.2. 認知負荷

為了探討實驗組學生與控制組學生在認知負荷上的差異，本研究採用了獨立樣本 t 檢定。在心智負荷方面，實驗結果顯示兩組之間無顯著差異($t = 1.29, p > 0.05$)。實驗組與控制組的平均值分別為 2.53 與 2.37，顯示「基於雙層次測驗為基礎之環景導覽學習」與「傳統環景導覽學習」的學習內容皆未引起過多的心理負荷；另一方面，在心智努力方面，兩組也無顯著差異($t = 0.72, p > 0.05$)。實驗組與控制組的平均值分別為 2.53 與 2.43，說明「基於雙層次測驗為基礎之環景導覽學習模式」與「傳統環景導覽學習模式」皆未造成學習者過度的心理努力。

4.3. 學習動機分析

為了探討實驗組學生在學習動機方面是否與控制組學生有顯著差異，本研究採用單因子共變數分析(ANCOVA)。首先，Levene 變異數同質檢定結果顯示變異數同質性假設未達顯著水準($F = 2.35, p > 0.05$)，表示兩組變異數相同。此外，迴歸係數同質性檢定結果($F = 2.83, p > 0.05$)也未達顯著水準，顯示兩組的迴歸斜率一致，符合變異數同質性之假設，可進行 ANCOVA 分析。實驗結果顯示實驗組學生的學習動機顯著高於控制組 ($F = 4.67, p < 0.05$)，表示「基於雙層次測驗為基礎之環景導覽學習模式」能有效提升學生的學習動機，使他們更加投入學習並提升學習興趣。

5. 結論

本研究結合雙層次測驗與 720 度環景技術，建構一套基於雙層次測驗為基礎之博物館環景導覽學習模式，探討使用新的學習模式對學習者動機與成效的影響。實驗結果顯示採用該系統的實驗組學生在學習動機與學習成就上皆顯著優於傳統環景導覽的控制組，顯示該系統有助於提升學習者對展品知識的理解與學習投入。此外，兩組在認知負荷與心智努力方面並無顯著差異，證明此本研究提出的導覽學習模式在不增加額外負擔的前提下，能有效促進學習。然而，本研究對象及樣本數具有一定侷限，建議未來可擴大至不同年齡層或教育階段，以提升研究的適用性。其次，本研究以自然生態為導覽主題，尚未驗證系統於其他學科領域的適用性與成效。未來可針對歷史、人文、科技等主題進行跨領域應用與比較分析。同時，亦可結合 AI 推薦機制或學習歷程數據分析，以優化引導策略，進一步強化學習效果與系統智能化程度。

參考文獻

Chang, C. Y., Hsu, K. Y., & Lin, Y. H. (2023). Diagnostic effects of two-tier assessments in science education: A meta-analysis. *Educational Psychology Review, 35*(1), 243-267.

Chen, Y. T., & Yang, K. C. (2022). Exploring immersive experiences: The impact of 720-degree virtual reality on museum visitors' engagement and learning outcomes. *Computers & Education, 192*, 104629.

Lee, H., Yoon, H., & Jung, Y. (2023). Virtual reality museum tours during COVID-19: Perspectives on technological innovation, visitor experiences, and digital cultural heritage. *Journal of Cultural Heritage, 60*, 143-152.

Li, Y., Hwang, G. J., & Chen, N. S. (2022). A two-tier diagnostic learning approach to improving students' conceptions and motivations in a contextual mobile learning environment. *Journal of Computer Assisted Learning, 38*(2), 465-478.

Li, J., Zheng, X., Watanabe, I., & Ochiai, Y. (2024). A systematic review of digital transformation technologies in museum exhibition. *Computers in Human Behavior, 108407*.

Hwang, G. J., Lin, A. C., & Chang, S. C. (2023). Effects of incorporating a two-tier assessment-based guiding mechanism into contextual games on students' learning achievement and behaviors in horticultural education. *Interactive Learning Environments, 32*(10), 6845-6867.

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.

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.

Wang, X., Zhang, J., & Yang, T. (2022). Evaluation and improvement strategies for digital museum exhibitions under the context of COVID-19. *Education and Information Technologies, 27*(6), 8355-8375.

探討專業學習社群對國小教師在虛擬實境融入教學之教學自我效能與 TPACK 的影響

The Impact of Professional Learning Communities on Elementary School Teachers' Teaching Self-Efficacy and TPACK in Integrating Virtual Reality into Instruction

陳振遠[1]，林筠修[2]，孫之元[3*]
[1]桃園市立幸福國民小學
[2]桃園市立新埔國民小學
[1,3]陽明交通大學 教育研究所
[*] jerrysun@nycu.edu.tw

【摘要】 虛擬實境 (VR) 科技已成為逐步應用於教育領域，其身歷其境的特性能提升學生的學習表現，但對於新科技的導入，我們也需培訓教師具備相關能力，使其能更妥善應用 VR 科技進行教學。因此本研究建構以 TPACK 為基礎的虛擬實境教學專業學習社群，提供多元的增能研習與專業成長活動，希冀提升教師的專業知識與自我效能。研究結果顯示，參與為期一學年的虛擬實境教學專業學習社群後，25 位國小教師的 TPACK 與教學自我效能均有顯著提升，且社群活動能有效促進教學實務發展，因此本研究之發現能進一步擴展虛擬實境科技融入教學的實踐價值。

【關鍵字】 虛擬實境；專業學習社群；TPACK；教學自我效能

Abstract: *Virtual Reality (VR) has been increasingly used in education due to its immersive features that can enhance student learning. However, teachers need proper training to effectively apply VR in teaching. This study established a TPACK-based professional learning community that offered various workshops to enhance teachers' knowledge and self-efficacy. The results showed that 25 elementary school teachers experienced significant improvements in both areas, and the activities effectively supported their teaching practices. These findings highlight the value of integrating VR into education.*

Keywords: virtual reality, professional learning community, TPACK, teaching self-efficacy

1. 研究動機與目的

隨著科技進步，虛擬實境（Virtual Reality, VR）的設備與教材已廣泛運用於教育領域，而隨著新科技融入課堂教學的比例增高，更需進一步探索教師如何有效的結合科技、學科內容與教學法來增進學生的學習效益（Kapici & Akcay, 2023）。此外，我們也需提供有效的措施來提升教師的科技教學專業知能，使其更能妥善運用創新科技進行教學（Fabian et al., 2024）。

教師是教學成功的重要因素之一，但目前對於教師如何有效利用虛擬實境進行教學，以及相關能力培育之研究甚少（Alalwan et al., 2020），且也需進一步以教師角度去探討國小現

場使用虛擬實境科技的優勢與挑戰，才能提升現場老師的科技融入態度與信心，使其更正向與積極的將 VR 融入各領域教學中（Walstra et al., 2023）。

科技教學內容知識 TPACK（Technological Pedagogical Content Knowledge）架構，其涵蓋教師在教學中的科技整合能力與相關教學知能等面向，並能協助教師的專業發展（Sun et al., 2023）。教師專業發展影響著學生學習成效與整體教育品質，近期發展的專業學習社群模式藉由提供同儕支持、對話、分享與實踐的機會，形成另一種有效促進教師專業發展的策略（孫志麟，2010）。因此本研究以 TPACK 架構為基礎，成立虛擬實境教學專業學習社群，提供參與教師多元的增能研習與專業成長活動，期待提升參與教師對於虛擬實境科技融入教學的自我效能與 TPACK。

根據上述背景與目的，本研究進行以下問題之探究。

1) 虛擬實境教學專業學習社群是否能顯著提升國小教師的 TPACK?
2) 虛擬實境教學專業學習社群是否能顯著提升國小教師的教學自我效能感?

2. 文獻探討

2.1. 虛擬實境應用於國小教學上的優勢與挑戰

虛擬實境（VR）能模擬現實世界，並讓使用者與各種物件互動且不受時間與空間之限制，營造出高度沉浸感的學習環境（Coban et al., 2022），VR 能帶給師生不同的教學資源與學習體驗（Pellas et al., 2021），並提升不同教育階段學生的各項表現與學科實際應用能力（Hung et al., 2023; Sun et al., 2023）。但研究也指出，為使創新科技能於課堂中進行有效利用，需深入了解與解決科技導入教學的各項因素與阻力，例如：師生對新科技的看法與期待、過去經驗、使用動機與科技使用之門檻等（Alfalah, 2018），才能發揮創新科技的教育價值。

過去研究指出，教師們多數認同虛擬實境（VR）對於國小學生在學習上的幫助（Alalwan et al., 2020），但對於教師如何利用 VR 教學並無有效的支持，導致教師會因為對新科技的恐懼，或無法正確操作教學使用等因素，產生該科技對學習上的負向效果（Alfalah, 2018; Pellas et al., 2021）。其他研究也發現，教師在課堂中使用 VR 科技面臨到諸多挑戰與障礙，使其無法進行有效的 VR 融入教學，例如：教師個人的科技能力不足（Alalwan et al., 2020）；無法獲得有用的 VR 教材（Walstra et al., 2023）；科技設備導入教學的複雜性（Cheng, 2024）；教師需要額外的設備、軟體學習與備課時間（Alfalah, 2018）；VR 科技如何適當的融入學科與進行教學（Fransson et al., 2020）；以及缺乏教學實踐的機會（Alalwan et al., 2020）。因此 VR 教學的專業發展極為重要，學校也更應提供相關課程與培訓活動，並鼓勵教師參與各種科技實務訓練，培養相關的數位教學能力，提升教師 VR 教學實踐的知識與技能，並進一步促進學生的學習表現（Alalwan et al., 2020; Cheng, 2024; Fransson et al., 2020; Walstra et al., 2023）。

2.2. 專業學習社群與教師專業成長

科技與資訊已在課堂中占有重要地位，藉由科技融入能提升教師教學與學生學習成果，因此我們需提供相關的專業發展計畫，進一步協助教師改善課堂中的科技融入策略，提升教學知識與技能（Bowman et al., 2022）。但過去教師專業發展計畫多以個人為主，其無法顧及學校組織層面的發展，導致整體發展成效不彰（Admiraal et al., 2021; 孫志麟，2010），而為解決此困境，近期推廣的專業學習社群（Professional Learning Community, PLC）成為新的教師專業學習與發展策略，當教師參與成功的專業學習社群，能有效改善其教學成效並對學生表現有正向影響，進而提升學校整體教育發展（Antinluoma et al., 2021）。

社群是指一群具有相同目標與承諾的人，透過相互協助與知識分享來達成共同目標，進而形成擁有共同經驗與價值觀的群體，而教師專業學習社群進一步強調教育人員在群體中的學習，促進其教學專業成長，並在分享與實踐中進一步精進教學（孫志麟，2010）。此外，教師專業學習社群提供同儕支持與觀摩學習的機會，讓教師能於合作中提升教師自我效能與進行教學創新（劉倚禔、吳勁甫，2017；孫志麟，2010；翁暄睿、陳慧綺，2020）。

本研究彙整過去研究，發現成功的專業學習社群具以下幾有個特徵：1）建構共同目標；2）提供專業能力的培訓活動；3）社群中的教師有對話與合作的機會；4）成員能針對教學實務進行分享與討論（Admiraal et al., 2021; Antinluoma et al., 2021; 孫志麟，2010; 翁暄睿、陳慧綺，2020），因此本研究將以上述特徵為基礎，成立「虛擬實境教學專業學習社群」，期待參與社群的教師能共同成長，提升教師專業與效能。

2.3. TPACK 與教學自我效能對教師之影響

過去教師的資訊與科技等專業能力發展或培訓計畫，其著重於操作技能的培養，並不強調讓教師學習科技融入教學與課堂營造的策略，因此後續發展出 TPACK 架構，藉由「科技教學內容知識」（Technological Pedagogical Content Knowledge, TPACK）框架能更深入了解教師在課堂中使用科技的狀況，並協助他們將科技有效地融入教學中（Kapici & Akcay, 2023）。TPACK 架構能有系統的探討不同領域教師對於新科技的知識、教材內容、教學策略之理解，以及如何整合於課堂中形成有效的教學模式（黃品齊等人，2020）。

過去研究發現，以 TPACK 為主軸的教師專業發展計畫，能有效促進各教育階段教師在資訊與科技領域中的教學能力，其包含各項教學知識、教案設計能力、科技實作技能與教學自我效能等態度與能力的提升（Sun et al., 2023）。而針對虛擬實境等新型學習科技，將 VR 納入 TPACK 架構中，其能協助教師挑選合適的虛擬實境教材與有效的教學策略，並設計出符合學習目標之教案，進一步促進學生的學習成效（Shin & Kim, 2024）。因此提供 TPACK 架構的 VR 專業成長計畫，能幫助教師有效整合 VR 科技於課堂中（Khukalenko et al., 2022）。但研究也指出，對於教師 TPACK 的發展需要長期的實踐過程，才能真實了解 VR 在國小教育現場的教學與學習效益（Kim & Yoon, 2024）。

另一方面，教學自我效能是教師能否成功將科技導入教學的重要關鍵，其影響著教師實際展現出的教學知識與行為（Sun et al., 2023; Yildiz Durak et al., 2023），且也與教師在課堂中使用科技進行教學的信念有關（Kapici & Akcay, 2023）。研究指出，參與教師專業學習社群能有效提升教學自我效能（孫志麟，2010），當教師具有高度的自我效能，能正向影響 TPACK 表現與更願意主動進行科技融入教學（黃品齊等人，2020），並進一步提升學生學習成效（Yildiz Durak et al., 2023）。

3. 研究方法

3.1 研究設計與研究參與者

根據研究目的，為了解國小教師參與教師專業學習社群後的 TPACK 與教學自我效能提升狀況，本研究採用單組前後測設計，於台灣北部某國小建立以 TPACK 為基礎之虛擬實境教學專業成長社群，並有 25 位國小教師自願參與。專業學習社群時間為期一學年，25 位教師皆完整參與所有活動並進行相關量表前後測，故相關數據皆具有效性。

3.2. 研究工具

本研究為了解參與專業學習社群之教師在 TPACK 與教學自我效能上的變化，採用「TPACK 量表」與「科技教學自我效能量表」兩項工具進行評估與分析。

「TPACK 量表」來源為黃品齊等人（2020）改編之量表，並邀請兩位資深國小主任協助修改字句，以符合本研究之實驗情境所需。TPACK 量表為李克特氏 6 點量表，共分為三個面向，其分別為 TCK（科技學科知識），Cronbach's α值為.88；TPK（科技教學知識），Cronbach's α值為.90；TPCK（科技教學學科知識），Cronbach's α值為.92。整體量表 Cronbach's α值為.95，具有高度一致性。

為探討教師的教學自我效能，本研究使用「教學自我效能量表」進行評估，採用之量表參考黃品齊等人（2020）改編之量表再進行字句調整。教學自我效能量表採用 6 點量表形式，共分為三個面向，其分別為「教學成效預期」，Cronbach's α值為.82；「教學知識效能」，Cronbach's α值為.95；「科技融入教學效能」，Cronbach's α值為.94。整體量表 Cronbach's α值為.96，具有高度一致性。

本研究為進一步了解教師真實感受，採用社群討論觀察、開放式問卷與訪談收集質性資料，以供研究結果佐證。

3.3 虛擬實境教學專業成長社群活動規劃

本研究為能有效提升教師之 TPACK 與教學自我效能，因此專業學習社群活動根據研究工具之面向內涵進行設計與規劃（如表 1 所示）。學習社群每月皆有 1 至 2 場活動（如圖 1-4 所示），提供以增能學習為主的活動，也讓社群成員有同儕回饋、分享、合作與展示個人成果的機會，並藉由教學實踐提升成員的專業發展（Admiraal et al., 2021）。

表 1 專業學習社群活動面向

面向	社群專業發展活動
TCK 科技學科知識	● 虛擬實境設備體驗與操作工作坊 ● 虛擬實境教材軟體設計實作
TPK 科技教學知識	● 虛擬實境教學實務經驗分享研習 ● 科技領域專家入校輔導會議
TPCK 科技教學學科知識	● 虛擬實境跨領域教案設計研討會議 ● 虛擬實境融入教學之公開備、觀、議課活動
TPACK 教學自我效能	● 虛擬實境教學實踐 ● 實體例行討論會議 ● 網路社群即時溝通、回饋與分享

圖 1 虛擬實境設備體驗與操作工作坊

圖 2 虛擬實境教學實務經驗分享研習

圖 3

VR 融入教學之公開備、觀、議課活動

圖 4

虛擬實境教學實踐

4. 研究結果與討論

4.1 TPACK 成效分析

本研究採用成對樣本 *t* 檢定來探討教師在參與專業學習社群後的 TPACK 改變狀況，其結果如表 2 所示。分析結果顯示，歷經一學年的學習社群參與，25 位教師在整體 TPACK 上皆有顯著提升，並在各面向的專業知識也有顯著成長。本次結果與 Alfalah （2018）及 Khukalenko 等人（2022）的研究一致，當提供 VR 專業發展培訓，能有效提升教師的科技整合、應用與教學能力，並讓老師們更有意願進行科技教學實踐。

表 2 TPACK 量表之成對樣本 t 檢定結果

面向	配對	人數	平均數	標準差	*t* 值
TCK 科技學科知識	後測問卷得分	25	5.00	0.60	**2.56***
	前測問卷得分	25	4.59	0.91	
TPK 科技教學知識	後測問卷得分	25	4.85	0.72	**2.70***
	前測問卷得分	25	4.41	0.98	
TPCK 科技教學學科知識	後測問卷得分	25	4.80	0.88	**3.26****
	前測問卷得分	25	4.40	1.10	
整體 TPACK	後測問卷得分	25	4.88	0.69	**3.00****
	前測問卷得分	25	4.40	0.96	

$^{**}p < .01$

針對學習社群提供支持的部份，參與教師表示「***有專業夥伴的協助，自己能熟悉 VR 操作，並學到更多課程設計的專業知識-教師 05***」；「***初期適應有困難，但經由社群的多場研習活動，自己已經能有效的利用 VR 來進行教學-教師 25***」，因此提供專業的科技培訓以及促進教師間的同儕合作，能有效協助老師跨越 VR 技術障礙（Alfalah, 2018; Fransson et al., 2020），也進一步提升 TPACK 教學專業來滿足不同學習者的需求（Kim & Yoon, 2024），展現新科技在教育上的價值（Cheng, 2024）。

4.2 教學自我效能分析

本研究採用成對樣本 t 檢定來探討教師在參與專業學習社群後的教學自我效能改變狀況，其結果如表 3 所示。分析結果顯示，參與學習社群之教師在教學自我效能上皆有顯著提升，

其結果與 Sun 等人(2023)及 Kapici、Akcay (2023)的研究一致，當提供教師學習與科技實作的機會，對教師自我效能有積極正向的影響。

表 3 教學自我效能量表之成對樣本 t 檢定結果

面向	配對	人數	平均數	標準差	*t* 值
教學成效預期	後測問卷得分	25	4.98	0.66	**2.52***
	前測問卷得分	25	4.62	0.97	
教學知識效能	後測問卷得分	25	4.86	0.93	**3.37****
	前測問卷得分	25	4.39	1.11	
科技融入教學效能	後測問卷得分	25	4.60	0.86	**3.92****
	前測問卷得分	25	3.97	1.09	
整體教學自我效能	後測問卷得分	25	4.81	0.77	**3.90****
	前測問卷得分	25	4.33	1.00	

*p <.05，**p <.01

彙整教師訪談結果發現，「***有工作坊與夥伴的協助，我能妥善解決課堂上的設備問題-教師 08***」；「***藉由 VR 頭盔，可以帶給學生更高的學習興趣與動機-教師 11***」；「***我在課堂中能利用 VR 設備幫助學生理解學科知識-教師 18***」，因此針對教師需求建立之虛擬實境專業學習社群能有效提升教師的科技融入教學能力與態度（Antinluoma et al., 2021; Bowman et al., 2022）。

表 4 TPACK 與教學自我效能之 Pearson 相關分析表

	TPACK	教學自我效能
TPACK	1	
教學自我效能	**.829***** .000	1

***p <.001

最後分析自我效能與 TPACK 之關聯，本研究採用 Pearson 相關分析(如表 4 所示)，結果顯示兩變項相關係數為.829（p <.001），表示教師的 TPACK 與教學自我效能具有高度正相關，此與過去研究一致，將 TPACK 架構與教師自我效能納入教師專業發展計畫，其有助於教學專業成長，並形成相互正向之影響（Yildiz Durak et al., 2023）。

5. 結論與建議

有效教學的重要關鍵之一是教師的專業能力，提升教師的科技教學專業知能，能使其更妥善運用虛擬實境等創新科技來增進學生的學習成效。本研究建構以 TPACK 為基礎的虛擬實境教學專業學習社群，提供多元的增能研習與專業成長活動，協助教師們克服 VR 技術障礙，增進課程設計等專業知識，並更有效地利用 VR 進行教學實踐。研究結果也顯示，參與為期一學年的虛擬實境教學專業學習社群後，25 位國小教師的整體 TPACK 及教學自我效能均有顯著提升。因此本研究提出之虛擬實境教學專業學習社群，其能有效促進國小教師在虛擬實境融入教學方面的專業發展，提升教學信心與能力。此結果對於未來推動教師專業發展，以及促進虛擬實境在國小教育現場的有效應用具有重要的實務意義。

而未來的研究可以朝向更深入地探討教師背景、擴大研究樣本、進行更細緻的分析，以及設立對照組等方面努力，以更全面理解專業學習社群在促進教師虛擬實境融入教學能力發展中的作用。

致謝

本研究感謝國家科學及技術委員會（NSTC 112-2410-H-A49-019-MY3 和 MOST 111-2410-H-A49-018-MY4）與教育部（114 年 5G 新科技學習示範學校）計畫之支持，以及感謝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

劉倚禔、吳勁甫（2017）。教師專業學習社群與教師教學效能關係之後設分析。**教育科學期刊, 16**（1），157-169。

孫志麟（2010）。專業學習社群：促進教師專業發展的平台。**學校行政，69**，138-158。https://doi.org/10.6423/hhhc.201009.0138

翁暄睿、陳慧綺（2020）。國民小學教師專業學習社群對教師教學創新影響之研究：以正向心理資本為中介變項。**師資培育與教師專業發展期刊，13**（3），113-138。https://doi.org/10.3966/207136492020121303005

黃品齊、葉錦樹、簡桂彬、梁至中（2020）。體育教師教學自我效能與科技教學內容知識關係之探討。**數位學習科技期刊，12**（1），83-108。https://doi.org/10.3966/2071260x2020011201004

Admiraal, W., Wouter, S., Loes, D. J., Yolande, E., & and Sligte, H. (2021). Schools as professional learning communities: what can schools do to support professional development of their teachers? *Professional Development in Education, 47*(4), 684-698. https://doi.org/10.1080/19415257.2019.1665573

Alalwan, N., Cheng, L., Al-Samarraie, H., Yousef, R., Ibrahim Alzahrani, A., & Sarsam, S. M. (2020). Challenges and prospects of virtual reality and augmented reality utilization among primary school teachers: A developing country perspective. *Studies in Educational Evaluation, 66*, 100876. https://doi.org/10.1016/j.stueduc.2020.100876

Alfalah, S. F. M. (2018). Perceptions toward adopting virtual reality as a teaching aid in information technology. *Education and Information Technologies, 23*(6), 2633-2653. https://doi.org/10.1007/s10639-018-9734-2

Antinluoma, M., Ilomäki, L., & Toom, A. (2021). Practices of Professional Learning Communities. *Frontiers in Education, 6*, 617613. https://doi.org/10.3389/feduc.2021.617613

Bowman, M. A., W., V. V., Zilu, J., & and Xie, K. (2022). Teachers' exposure to professional development and the quality of their instructional technology use: The mediating role of teachers' value and ability beliefs. *Journal of Research on Technology in Education, 54*(2), 188-204. https://doi.org/10.1080/15391523.2020.1830895

Cheng, K.-H. (2024). Development of an immersive virtual reality system for learning about plants in primary education: evaluation of teachers' perceptions and learners' flow experiences and

learning attitudes. *Educational technology research and development, 72*(2), 845-867. https://doi.org/10.1007/s11423-023-10300-6

Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review, 36*, 100452. https://doi.org/10.1016/j.edurev.2022.100452

Fabian, A., Backfisch, I., Kirchner, K., & Lachner, A. (2024). A systematic review and meta-analysis on TPACK-based interventions from a perspective of knowledge integration. *Computers and Education Open, 7*, 100200. https://doi.org/10.1016/j.caeo.2024.100200

Fransson, G., Holmberg, J., & Westelius, C. (2020). The challenges of using head mounted virtual reality in K-12 schools from a teacher perspective. *Education and Information Technologies, 25*(4), 3383-3404. https://doi.org/10.1007/s10639-020-10119-1

Hung, C. Y., Lin, Y. T., Yu, S. J., & Sun, J. C.-Y. (2023). Effects of AR- and VR-based wearables in teaching English: The application of an ARCS model-based learning design to improve elementary school students' learning motivation and performance. *Journal of Computer Assisted Learning, 39*, 1510-1527. https://doi.org/10.1111/jcal.12814

Kapici, H. O., & Akcay, H. (2023). Improving student teachers' TPACK self-efficacy through lesson planning practice in the virtual platform. *Educational Studies, 49*(1), 76-98. https://doi.org/10.1080/03055698.2020.1835610

Khukalenko, I. S., Kaplan-Rakowski, R., An, Y., & Iushina, V. D. (2022). Teachers' perceptions of using virtual reality technology in classrooms: A large-scale survey. *Education and Information Technologies, 27*(8), 11591-11613. https://doi.org/10.1007/s10639-022-11061-0

Kim, J.-U., & Yoon, H.-G. (2024,). Exploring pre-service elementary teachers' technological pedagogical content knowledge in developing and using virtual reality materials reflecting spatial thinking in astronomy class. *Journal of Science Education and Technology*, 1-17. https://doi.org/10.1007/s10956-024-10157-1

Pellas, N., Mystakidis, S., & Kazanidis, I. (2021). Immersive virtual reality in K-12 and higher education: A systematic review of the last decade scientific literature. *Virtual Reality, 25*(3), 835-861. https://doi.org/10.1007/s10055-020-00489-9

Shin, Y. C., & Kim, C. (2024). Pedagogical competence analysis based on the TPACK model: Focus on VR-based survival swimming instructors. *Education Sciences, 14*(5), 460. https://doi.org/10.3390/educsci14050460

Sun, J., Ma, H., Zeng, Y., Han, D., & Jin, Y. (2023). Promoting the AI teaching competency of K-12 computer science teachers: A TPACK-based professional development approach. *Education and Information Technologies, 28*(2), 1509-1533. https://doi.org/10.1007/s10639-022-11256-5

Sun, J. C.-Y., Ye, S.-L., Yu, S.-J., & Chiu, T. K. F. (2023). Effects of wearable hybrid AR/VR learning material on high school students' situational interest, engagement, and learning performance: The case of a physics laboratory learning environment. *Journal of Science Education and Technology, 32*(1), 1-12. https://doi.org/10.1007/s10956-022-10001-4

Walstra, K., Cronje, J., & Vandeyar, T. (2023,). A review of virtual reality from primary school teachers' perspectives. *Electronic Journal of e-Learning, 22*(3), 1-11. https://doi.org/10.34190/ejel.21.4.3060

Yildiz Durak, H., Atman Uslu, N., Canbazoğlu Bilici, S., & Güler, B. (2023). Examining the predictors of TPACK for integrated STEM: Science teaching self-efficacy, computational thinking, and design thinking. *Education and Information Technologies, 28*(7), 7927-7954. https://doi.org/10.1007/s10639-022-11505-7

ChatGPT 模擬 OSCE 護病溝通中學習感知與溝通表現之性別差異與相關性分析

Gender Differences and Correlations Between Learning Perceptions and Communication Performance in ChatGPT-Simulated OSCE Nurse–Patient Communication

楊馥宇[1*]，吳美玲[2]
[1]長庚科技大學臨床技能中心
[2]長庚科技大學護理系
[*] fyyang@mail.cgust.edu.tw

【摘要】本研究探討護生性別在 ChatGPT 模擬 OSCE 護病溝通之學習感知與溝通表現之差異與關聯。28 位護生（男 12，女 16）參與 ChatGPT 模擬問診與衛教，完成學習感知問卷，依 OSCE 評量指標評分溝通。採用 Mann-Whitney U 檢定性別差異，Spearman 檢定變項相關，並以 Z 分數比較學生學習感知與溝通表現落差。結果顯示男護生學習感知顯著高於女護生，但學習感知與溝通表現無顯著相關，男護生傾向高估自身表現，女護生則相對一致，顯示性別可能影響 AI 模擬學習感知與評估一致性；對話數與溝通得分具正相關，無性別差異。未來生成式 AI 輔助學習應納入性別與對話參與指標，確保教學回饋與評估準確性。
【關鍵字】護病溝通；學習感知；客觀結構式臨床評量；性別差異；ChatGPT

Abstract: *This exploratory study examined gender differences in nursing students' learning perceptions and communication performance during ChatGPT-simulated OSCE communication. Twenty-eight students (12 male, 16 female) participated, completing OSCE evaluations and perception questionnaires. Mann–Whitney U tests showed that male students had significantly higher perception scores than female students ($p < 0.05$), but Spearman's rho found no correlation between perceptions and performance. Z-scores revealed that male students tended to overestimate their performance, whereas female students' learning perceptions were more aligned with their actual scores. Additionally, the number of dialogue turns was positively correlated with communication scores, though neither variable showed significant gender differences. Despite the small sample size, these findings suggest that gender may influence perceptions in AI-assisted learning environments, supporting the inclusion of gender and engagement metrics in AI educational design.*
Keywords: Nurse–Patient Communication, Learning Perception, OSCE, Gender Differences, ChatGPT

1.前言

護病溝通能力被廣泛認為是影響醫療照護品質的關鍵要素。具備熟練溝通技巧的護理人員，能更有效地提供醫療服務，減輕患者的焦慮、疼痛與疾病症狀，並於護理過程中建立患者對醫療體系的信任(Siokal et al., 2023)。因此，護理人員需持續精進其溝通能力，根據患者的年齡、情緒狀態、生理條件、語言與文化背景、社經地位及溝通環境等因素，調整溝通策略，採取同理、專業與個別化的訊息傳遞方式，進而提升患者對醫療資訊的理解與接受度，

增進其滿意度(Naz et al., 2024 ; Siokal et al., 2023)、治療依從性(Kerr et al., 2022)，並改善生理與功能狀況(Nummer et al., 2024)。相對地，若溝通技巧不足，將可能增加醫療錯誤風險，對病人造成負面後果。

然而，實務中護病溝通常面臨多重障礙，其原因包括護理人員溝通訓練不足、專業知識不全、同理心與自信心缺乏、對患者態度消極，以及語言與文化隔閡等因素，皆可能影響其傾聽與回應能力，進而削弱溝通成效(Alshalawi et al., 2025)。因此，護理教育機構有責任在教育階段強化學生之溝通技巧與態度，協助其建立積極正向的護病關係，以因應臨床現場之實際需求。

近年研究亦指出，護病溝通技巧可能存在性別差異。護理人員的性別會影響其與患者之間的互動方式，這可能與性別刻板印象、溝通風格差異或文化期望有關(Alharazi et al., 2025)。Graf 等人(2017)指出，女性護理系學生在溝通技巧表現上顯著優於男性護理系學生，女性對溝通訓練態度更積極，且溝通方式較具體；此外，女性在解釋與澄清、提問技巧，以及正向回饋等面向的得分也顯著較高(Kalati et al., 2025)。然而，性別是否必然影響溝通表現，尤其是在人工智慧模擬 OSCE （客觀結構式臨床評量）情境中，仍有待進一步驗證。

一般而言，護病溝通可區分為語言、非語言與副語言三種形式。本研究聚焦於語言溝通的表現，亦即護理學生在 OSCE 評量中與 ChatGPT 模擬病人互動時，與任務相關的語言對話 次數與得分。需要注意的是，若對話內容未涵蓋評分指標中所規範之任務項目，則即使互動頻繁，亦不代表其溝通表現良好。反之，若對話內容過度集中於特定任務，忽略其他評分項 目，也可能導致整體分數偏低。換言之，對話次數與溝通得分並不一定呈現正向關係，其間是否存在性別差異，亦值得探討。

目前，多數台灣護理教育機構採用 Harden 與 Gleeson (1979)所提出的 OSCE 模式，作為評估學生臨床技能與畢業門檻的依據。OSCE 強調標準化病人、情境模擬與結構化評量，旨在提升護學生的臨床信心與專業能力(Onwudiegwu, 2018)。惟傳統 OSCE 仰賴人工病人與大量人力資源，成本與訓練時間高昂。隨著生成式人工智慧(Generative AI)技術快速發展，包含 ChatGPT 在內的大型語言模型(LLMs)逐漸被導入醫護臨床與教育場域中，作為教學與學習的輔助工具(Pan & Ni, 2024)。

ChatGPT 具備自然語言互動、即時回饋、低成本與高度可調整等優勢，能有效支援護理學生進行臨床情境模擬訓練，包括病史採集、病歷記錄、問診、衛教、交班與臨床決策等(Amin & Alanzi, 2024 ; Saad et al., 2025 ; Huang et al., 2024 ; Tsang, 2023)。於 OSCE 應用層面，亦可協助教師快速設計案例、標準化情境與評分機制，有效降低評估成本(Misra & Suresh, 2024)。透過 ChatGPT 取代標準化病人，可讓學生進行反覆練習，緩解考試壓力，並可即時提供溝通表現的評估與回饋，進而提升其自我反思與溝通能力。然而，目前針對 AI 模擬 OSCE 在實際溝通訓練中之成效與性別影響，仍缺乏實證研究。

另有研究指出，性別可能影響學習者的溝通風格、科技接受度與自我效能評估。性別不僅可能影響對科技工具的接受態度，也會形塑其學習經驗與知覺學習成效。Pan 與 Ni (2024) 調查來自五個醫護相關專業、不同學年的 1,718 名學生，結果發現男性對大型語言模型的理解與信任度顯著高於女性，並較可能認為其益處大於風險。此一差異亦可能影響護理學生於 ChatGPT 模擬 OSCE 中的態度與自我評估，導致學習感知與實際表現之間出現落差。然而，目前針對此類學習模式中性別差異所造成之主觀與客觀表現落差之探討，仍付之闕如。

綜上所述，儘管現有文獻已初步揭示 ChatGPT 於模擬 OSCE 護病溝通訓練的應用潛力，惟針對性別差異對於學習感知與實際溝通表現間關係之影響，實證研究仍屬有限。有鑑於此，

本研究擬進一步探究性別對 AI 模擬 OSCE 護病溝通學習歷程之影響，並聚焦於學習感知、互動頻率與溝通表現之間的關聯性，本研究欲探討以下三項研究問題：

一、不同性別的護理學生在 AI 模擬 OSCE 護病溝通中，於 OSCE 溝通得分、對話次數與學習感知是否存在顯著差異？

二、OSCE 溝通得分、對話次數與學習感知三者之間，是否具有顯著相關性？

三、護理學生的學習感知是否能真實反映其實際的 OSCE 溝通表現？

2.方法

2.1. 受試者與測試環境

本研究受試者為臺灣某大學護理系三、四年級學生共 28 名（男性 12 名，女性 16 名），平均年齡為 20.96 歲，採課堂宣導方式招募，並以自願參與為原則。所有受試者於研究開始前，均已簽署研究參與同意書。其中 21 名學生曾於研究前使用過 ChatGPT，另有 7 名從未使用。所有受試者皆接受過基本護病溝通訓練，並具備 OSCE 評量與臨床實習經驗。

測試環境設於一間安靜研究室內，學生透過筆記型電腦以語音轉文字方式與 ChatGPT 互動，螢幕為 30 吋之外接顯示器，作為模擬對話視覺介面。測試進行時僅有一位受試者與一位研究人員在場，以避免他人干擾或造成訊息洩漏，確保資料的客觀性與真實性。

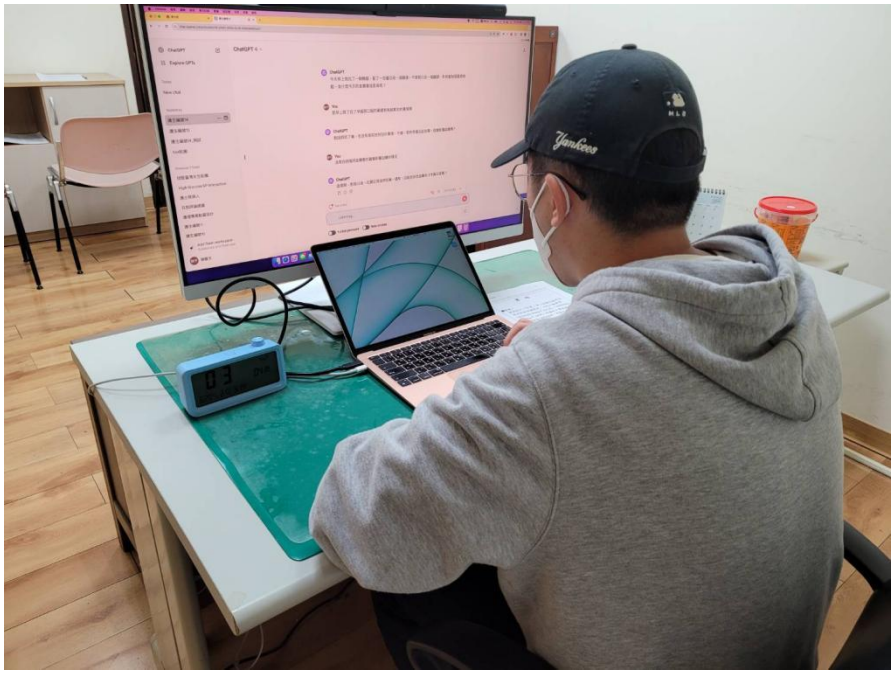


圖 1 ChatGPT 模擬 OSCE 護病溝通測試環境

2.2. 研究流程

本研究為個案研究，使用 OpenAI 開發的 ChatGPT-4 模型，模擬 OSCE 護病溝通與評分，藉由觀察、記錄和分析個別護生實際測試 ChatGPT 模擬 OSCE 問診評估與衛教之內容和行為，瞭解不同性別的護生使用 ChatGPT 進行高齡高血糖患者的護病溝通表現和學習感知。

研究分三階段進行：(1)設計提示語階段：利用具有專家信、效度的 OSCE 教案開發護病溝通的提示語，經過數十次測試，確認 ChatGPT 可以正確模擬 OSCE 中的標準病人，與護生溝通對話，確保護生練習問診和衛教之回應內容，皆符合正確病例情境訓練資料的範圍。(2)專家評估階段：請護理專家實際與 ChatGPT 進行問診與衛教，修正不符合實體 OSCE 過

程的互動說法與行為。(3)護生實測階段：護生實際進行 OSCE 護病溝通，瞭解 ChatGPT 模擬成效，如評分、回饋準確性和護生自評測試過程的學習感知。

正式評估流程為：(1)研究者說明本實驗目的和作法。(2)護生填寫受試者同意書。(3)系統操作說明：研究者教導護生如何運用語音轉文字與 ChatGPT 互動，並讓護生實際練習。(4)護生用一分鐘時間詳讀情境與考生任務。(5) ChatGPT 模擬 OSCE 護病溝通實作：護生在八分鐘時間內與 ChatGPT 模擬老年高血糖患者進行血糖、用藥、飲食與運動問診評估與衛教。(6) 護生填寫自編「ChatGPT 學習感知問卷」。

2.3. 操作工具及提示語設定

由於護理教育相當重視專業術語和特定疾病判斷標準的正確性，因此，在護生與 ChatGPT 模擬 OSCE 護病溝通時，必須確保 ChatGPT 能根據 OSCE 臨床情境劇本回應護生，並能依照事先提示語設定的評分標準，正確評估護生的 OSCE 溝通能力。本研究在 ChatGPT 的指令中，輸入評分標準和回應模式，例如當護生說「請評分」時，ChatGPT 就要根據預設評分標準，將前面問診與衛教的對話逐一檢查，提供護生各項評分結果與原因，以及當護生問到飲食或運動的問題，ChatGPT 模擬的病人必須主動追問相關問題。

研究者會先告知 ChatGPT 接下來對話，請它扮演老年糖尿病患者，接著把病患的病史貼 在對話中，請 ChatGPT 根據這些資料回答，最後再放一張由 ChatGPT 繪製的老年女性糖尿病 患者圖片，讓護生在一開始和 ChatGPT 對話時，會先看到那張病患圖片，幫助護生投入 OSCE 溝通情境中。

2.4. 資料蒐集與分析

本研究所蒐集之資料包括：ChatGPT 所記錄的護理學生與模擬 OSCE 病人之完整對話歷程，以及學生填寫之學習感知問卷。所有統計分析皆使用 IBM SPSS Statistics 28 版進行處理。在溝通表現方面，本研究採用兩項指標進行評估：其一為護生與 ChatGPT 模擬病人互動 的總對話次數；其二為語言溝通能力評分，依據黃湘萍等人(2017)所建構之 OSCE 案例「老年高血糖病患血糖控制評估與指導」之標準化溝通評量準則進行評分。該量表原始總分為 28 分，考量本研究採用之 ChatGPT 無法進行非語言表達，本研究刪除其中一項非語言評分指 標，調整後滿分為 26 分。原評量工具具良好信效度，整體評分者一致性指標 ICC 為.98，評分者信度達 100% ,原始 Cronbach's a 為.54。

本研究邀請兩位具臨床與教學經驗之護理專家，分別針對護生的語言表達、同理展現與內容正確性進行獨立評分，針對評分不一致之項目進行討論修正，最終互評者信度 Cronbach's a 達.96，具高度一致性。

學習感知部分則以自編問卷進行評估，共包含 20 題，採李克特氏五點量表作答，分數愈高代表自我學習成效感知愈佳。其中 8 題為對 ChatGPT 學習工具之態度評價，12 題為「學習成效感知」構面，後者之內部一致性信度為 Cronbach's a = .94，整體問卷信度為.95。研究分析中使用之學習感知分數，係取上述 12 題之平均分數。

在統計分析方面，本研究採用 Mann-Whitney U 檢定比較不同性別護生在對話次數、OSCE 溝通分數與學習感知三項變項之差異，並計算 Cohen's r 作為效果量，評估性別影響之實質程 度。為進一步分析變項間之關聯性，採用 Spearman 等級相關分析進行檢驗，顯著水準設為 α =.05。此外，為探討學生的主觀學習感知是否與其實際表現相符，將學習感知與 OSCE 得分進行 Z 標準化處理（Z-score)，並計算其差距值（Z 感知-Z 表現）。Z 值為正，代表學生自我感知高於實際表現，反之，則表示自評偏低。最後，再以 Mann-Whitney U 檢定

分析不同性別學生在 Z 落差分數上的顯著性差異，以評估性別對學習感知與實際表現一致性的潛在影響。

3.結果

本研究探討 ChatGPT 模擬 OSCE 護病溝通時，護生的性別對於對話數、OSCE 溝通分數與學習感知之差異與關聯。

3.1. 性別在學習感知、對話數與溝通分數比較

表 1 的 Mann–Whitney U 檢定結果顯示，男護生在 OSCE 溝通分數（M = 12.00, SD = 4.41）上略低於女護生（M = 12.44, SD = 2.83），男護生在對話數（M = 13.17, SD = 6.04）亦略低於女護生 M = 14.50, SD = 5.14），但 OSCE 溝通表現（U = 91，p>.05）與對話數（U = 84.5，p >.05），皆無顯著性別差異。

問卷中，與學習感知相關題分數（M = 4.35, SD = 0.54）。Mann–Whitney U 檢定結果顯示，男護生在學習感知分數上（M = 4.59, SD = 0.55）顯著高於女護生（M = 4.16, SD = 0.46），U = 41.5，p = .011<.05，Cohen's r=0.482，具中等效果量，結果顯示性別可能影響學生對 AI 模擬學習經驗的主觀評價學習感知。

表1 Mann–Whitney U 檢定性別在學習感知、對話數與OSCE溝通分數比較

項目	性別	數字	平均等級	等級總和	U	Z	顯著性
OSCE 溝通分數	男	12	14.08	169	91.0	-0.234	0.815
	女	16	14.81	237			
對話數	男	12	13.54	162.5	84.5	-0.535	0.593
	女	16	15.22	243.5			
問卷分數	男	12	19.04	228.5	41.5	-2.552	0.011*
	女	16	11.09	177.5			

3.2. 學習感知、對話數與溝通分數之關聯性分析

本研究採用 Spearman 等級分析學習感知、對話數與 OSCE 溝通分數之相關性，結果顯示對話數與 OSCE 溝通分數呈現顯著正相關（r=0.4, p<.05），溝通時對話數較多者，可能 OSCE 溝通分數亦較佳。相對地，護生在 AI 模護護病溝通的學習感知、對話數與 OSCE 溝通分數之間則無顯著相關（p>.05）。

3.3. 感知與表現 Z 分數落差之性別差異

為進一步探討學生在學習感知和實際 OSCE 溝通分數，是否會因不同性別而有所差異，本研究將學習感知與 OSCE 得分進行 Z 標準化後，計算其 Z 差距值（Z 感知-Z 表現）。結果顯示男護生整體在「主觀學習感知」的 Z 分數高於實際 OSCE 溝通分數（Z = -2.365，p =.018<.05)，而女護生之主觀學習感知與實際表現則相對一致，統計檢定顯示兩者落差具有顯著性別差異 p = 0.043，代表男護生相較於女護生，傾向高估自身 OSCE 溝通表現，顯示性別對於感知與實際表現落差具有統計顯著性。

4.討論

本研究運用 ChatGPT 模擬病人情境，進行 OSCE 護病溝通測試，並由具備臨床經驗之護 理專家依據標準化 OSCE 評分準則進行評量，目的在於探討生成式 AI 導入後，不同性別護生 在學習感知、對話次數及實際溝通能力之異同與關聯性。

研究結果顯示，男、女護生在 OSCE 溝通分數與對話數方面，雖然平均表現上女性略優於男性，但未達統計顯著差異，且效果量屬極小或小幅度。此結果意涵，在生成式 AI 模擬護 病溝通任務中，性別並非影響實際溝通表現的主要因素。此一發現與過去部分研究所指出之 性別影響溝通技巧的觀點(如 Kalati et al., 2025)有所不同，或可說明在標準化 AI 對話環境 中，溝通表現受到性別以外因素，如任務理解、專業知識的影響可能更為顯著。

相較於實際表現，學習感知得分則呈現明顯的性別差異。本研究發現男性護生在學習感知上的自我評估顯著高於女性，且具有中等效果量，顯示性別可能影響學生對 AI 輔助學習的主觀感受。進一步以 Z 標準化進行「學習感知與實際表現一致性」分析時，也發現男護生整體傾向高估自身表現，其主觀感知與實際 OSCE 分數存在顯著落差，反觀女性護生則在主觀 評估與客觀表現間呈現相對一致的趨勢。此發現呼應了先前關於性別在科技使用與學習自我 效能上差異的相關研究(Pan & Ni, 2024)，亦與 Kirkpatrick & Kirkpatrick (2007)對於學習成效評量層次的觀點一致，指出僅以學習者主觀反應評估學習成效可能有所偏差。

上述結果亦突顯教育評量過程中，性別公平性所應關注之問題。若 AI 教學系統設計者、教師或研究者僅依據問卷回饋判斷學生學習效果，可能高估男性學生的學習成果而低估女性學生的真實表現，進而產生不公平的教學回饋或評估結論。因此，建議未來在設計生成式 AI 作為護理教育溝通訓練工具時，應納入性別作為學習分析與教學調整的參考變項，並結合多 元資料來源，如問卷、實作成績、反思紀錄等，以提升教學回饋的準確性與評量的信效度。

整體而言，本研究結果指出生成式 AI 具備模擬護病溝通的潛力，不同性別學生在實際表 現上差異不大，但其在學習感知上的認知偏差，值得教育者與系統設計者關注。此發現對於未來 AI 應用於護理教育溝通能力培訓與學習回饋系統的設計，具有理論與實務的參考價值。

5.限制和建議

本研究具備初步探索價值，惟仍存在數量限制。首先，研究樣本僅來自臺灣一所科技大學之護理系學生，且樣本規模有限（n = 28)，因此研究結果在統計推論與外部效度上仍具侷限性。其次，儘管本研究揭示性別在學習感知層面上具有顯著差異，但此結果是否能擴及不同背景、年齡層或跨文化樣本，仍需進一步驗證。

因此，建議未來研究應擴大樣本數，以提升研究之代表性與結果的普遍性，並納入更多控制變項，如學生先備知識或溝通風格，進一步釐清影響學習感知與表現差異的潛在因子，及結合質性訪談資料，深入瞭解學生對生成式 AI 輔助學習的實際經驗，與性別角色預期對評估的影響。本研究提供有關 AI 模擬 OSCE 護病溝通應用之初步實證基礎，期盼能作為未來護理教育整合人工智慧科技與強化性別敏感教學設計的重要參考依據。

6. 結論

本研究旨在探討護理學生於 ChatGPT 模擬 OSCE 護病溝通情境中的性別差異，聚焦於學 習感知、對話次數與溝通能力分數三項指標之差異與關聯性。透過分析護生與生成式 AI

進行 問診與衛教任務過程中之對話紀錄、OSCE 溝通評分與學習感知問卷結果，發現若干具啟發 性的現象與趨勢。

研究結果顯示，男護生在學習感知之自我評估得分顯著高於女護生，然此感知與其實際 OSCE 溝通表現間並無顯著相關。進一步以 Z 分數分析學習感知與實際表現的一致性，亦發現男護生傾向高估自身能力，女性護生則呈現較高的一致性表現，顯示性別可能影響主觀學習感知與客觀學習表現之落差。相對而言，對話次數與 OSCE 溝通分數之間存在顯著正相關，代表對話次數可能有助於提升溝通任務達成表現，但此關係未因性別而產生顯著差異。

綜合而言，本研究初步指出在 AI 模擬 OSCE 溝通情境中，性別可能影響學生的主觀感知 與評估一致性，惟對實際溝通能力與互動參與之影響有限。此一結果對未來 AI 輔助學習系統 之設計與回饋機制具有實務啟示，建議教學者在進行生成式 AI 應用教學評估時，應將性別變 項與對話次數納入參考指標，以提升評估準確性與教學公平性。

致謝

本研究承蒙臺灣國家科學及技術委員會（NSTC）專題研究計畫「NSTC-113-2410-H-255-002-」和長庚科技大學臨床技能中心(Grant No. ZRRPF3N0081)之補助，謹此致謝。

參考文獻

黃湘萍、趙莉芬、王瑜欣、劉英妹、倪麗芬、簡淑慧（2017）。建構及檢測護理客觀結構式臨床技能測驗考題之信效度、鑑別度及難易度。**護理雜誌，64**(6)，67-76。https://doi.org/10.6224/JN.000084

Alharazi, R. M., Abdulrahim, R. J., Mazuzah, A. H., Almutairi, R. M., Almutary, H., & Alhofaian, A. (2025). Barriers and factors affecting nursing communication when providing patient care in jeddah. *Clinics and Practice, 15*(1), 19. doi:https://doi.org/10.3390/clinpract15010019

Alshalawi, S., Aljedaani, S., Fintyanh, D., Salim, S., Fairaq, W., & Refaei, S. (2025). The Effect of Nurse-Patient Communication on Patient Satisfaction in the Emergency Department. *Journal of Nursing & Midwifery Sciences, 12*(1). https://doi.org/10.5812/jnms-158740.

Al-Kalaldeh, M., Amro, N., Qtait, M., & Alwawi, A. (2023). Barriers to effective nurse-patient communication in the emergency department. *Emergency Nurse, 31*(5).

Amin, H. A. & Alanzi, T. M. (2024) Utilization of Artificial Intelligence (AI) in Healthcare Decision-Making Processes: Perceptions of Caregivers in Saudi Arabia. *Cureus, 16*(8), e67584. doi:10.7759/cureus.67584

Chang, C. Y., Yang, C. L., Jen, H. J., Ogata, H., & Hwang, G. H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology & Society, 27*(1), 215-230.

Harden, R. M., & Gleeson, F. A. (1979). Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical education, 13*(1), 39-54.

Huang T, Hsieh P, Chang Y. (2024).Performance Comparison of Junior Residents and ChatGPT in the Objective Structured Clinical Examination (OSCE) for Medical History Taking and Documentation of Medical Records: Development and Usability Study. JMIR Med Educ, DOI: 10.2196/59902

Kalati, Seyed Hossein, Ahmadzadeh Tori, Neda1, Kalati, Javaneh, Norouzi Parashkouh, Nastaran, Karimi, Hengameh （2025）. The effect of blended educational methods in communication skills in nursing with students' interactions with patients: A quasi-experimental study. Journal of Education and Health Promotion 14(1), 1-8. DOI: 10.4103/jehp.jehp_1479_23

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2007). Implementing the four levels: A practical guide for effective evaluation of training programs. Berrett-Koehler Publishers.

Lin, H. L., Liao, L. L., Wang, Y. N., & Chang, L. C. (2024). Attitude and utilization of ChatGPT among registered nurses: A cross-sectional study. International Nursing Review. https://doi.org/10.1111/inr.13012

Misra, S. M., & Suresh, S. (2024). Artificial Intelligence and Objective Structured Clinical Examinations: Using ChatGPT to Revolutionize Clinical Skills Assessment in Medical Education. *Journal of Medical Education and Curricular Development, 11*. https://doi.org/10.1177/23821205241263475

Naz, R., Khaleel, S., & Naz, S. (2024). Evaluation of nurse-patient communication and its impact on satisfaction. *Biological and Clinical Sciences Research Journal, 2024*(1), 936. https://doi.org/10.54112/bcsrj.v2024i1.936

Onwudiegwu, U. (2018). OSCE: Design, development and deployment. *Journal of the West African College of Surgeons*, *8*(1), 1–22.

Pan, G., & Ni, J. (2024). A cross sectional investigation of ChatGPT-like large language models application among medical students in China. *BMC Medical Education, 24*(1), 908. https://doi.org/10.1186/s12909-024-05871-8

Saad, O., Saban, M., Kerner, E., & Levin, C. (2025). Augmenting Community Nursing Practice With Generative AI: A Formative Study of Diagnostic Synergies Using Simulation-Based Clinical Cases. Journal of Primary Care00000000000000 & Community Health, 16, 21501319251326663.

Sibiya, M. N. (2018). Effective Communication in Nursing. Nursing, 19, 20-34. doi: 10.5772/intechopen.74995

Siokal, B., Amiruddin, R., Abdullah, T., Thamrin, Y., Palutturi, S., Ibrahim, E., ... & Mallongi, A. (2023). The influence of effective nurse communication application on patient satisfaction: a Literature Review. Pharmacognosy Journal, 15(3), 479-483. doi:10.5530/pj.2023.15.105

Tsang, R. (2023). Practical applications of ChatGPT in undergraduate medical education. *Journal of Medical Education and Curricular Development*, *10*, 23821205231178449.

PAPAMAMA-TOMO: A Chatbot-Based Support System for Foreign Guardians to Navigate Japanese School Culture

Xiaoyan LI [1*], Atsushi SHIMADA[2],
[1] Center for the Promotion of Interdisciplinary Education and Innovation, Kyushu University, Japan
[2] Faculty of Information Science and Electrical Engineering, Kyushu University
[*] lixiaoyan@kyoso.kyushu-u.ac.jp

Abstract: *This study addresses the challenges faced by foreign guardians in Japan in understanding Japanese elementary school culture. Building on previous efforts using educational games, we introduce a chatbot system named PAPAMAMA-TOMO. This chatbot offers interactive, real-time support tailored to the linguistic and cultural needs of foreign guardians. Instead of passively reading school handouts, guardians can actively ask questions about school events, required items, and behavioral expectations. The chatbot integrates insights from educational prediction technologies and conversational AI, offering context-sensitive and inclusive assistance. The goal is to reduce misunderstandings and cognitive stress, thereby supporting smoother parental involvement in school life.*
Keywords: chatbot, parental support, Japanese school culture, hidden curriculum, natural language interface

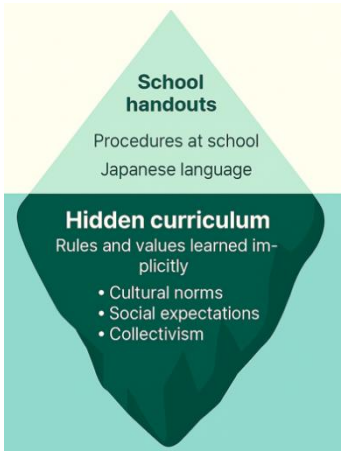
1. Introduction

The name “PAPAMAMA-TOMO” is derived from the Japanese term “mamotomo” (ママ友), which refers to friendships formed among mothers through their children’s schooling. By intentionally including both “papa” and “mama”, the name emphasizes inclusivity, encouraging all guardians—regardless of gender—to engage with the chatbot and school culture. This culturally embedded naming convention reinforces the system’s goal of providing friendly, accessible, and community-oriented support for navigating Japanese school life. To support foreign guardians in understanding the implicit norms and expectations in Japanese schools, we developed PAPAMAMA-TOMO, a chatbot that provides personalized and real-time assistance. Japanese elementary schools host culturally embedded events such as sports days and entrance ceremonies. While these events promote community ties, they often present barriers to non-native guardians due to unspoken cultural expectations and language complexities. Our approach builds on this foundation by offering a more accessible and conversational interface through a chatbot, enhancing immediacy and inclusiveness.

2. Background and Related Work

Japanese schools are known for their unique school culture, often described metaphorically as an iceberg (Figure 1), where the visible tip represents formal communication like school handouts, while a much larger portion—the implicit values, norms, and social expectations—remains hidden beneath the surface. In Japanese elementary schools, the primary mode of communication between schools and guardians is through printed handouts. These documents convey explicit information, representing the visible part of school culture. However, the so-called “hidden curriculum” (Jackson,

1968) lies beneath the surface, encompassing a range of unspoken rules and cultural values that students and guardians internalize through school life. These include egalitarianism, group orientation, motivation-driven instruction, and behavioral norms. In particular, the hidden curriculum plays a critical role in fostering students' implicit understanding of social behaviors, communication styles, and group dynamics within Japanese schools. These invisible norms, which are often unfamiliar to foreign guardians, are crucial for successful parental engagement and student adjustment.



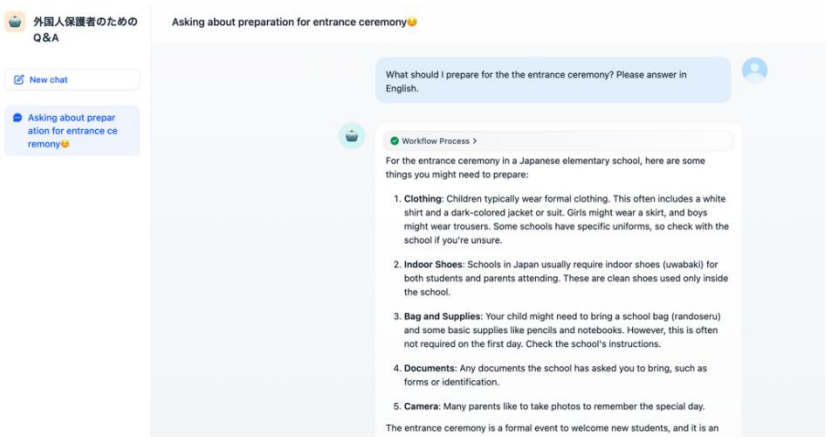


Fig. Iceberg model of Japanese school culture Fig.2 Screenshot of PAPAMAMATOMO Chatbot

For foreign guardians unfamiliar with Japanese schooling and with limited Japanese language skills, understanding school handouts is more than a linguistic task—it is an essential cultural skill (Li & Honda, 2015). Even guardians with conversational fluency in Japanese may struggle with these implicit norms, leading in some cases to disengagement or withdrawal from the school community. This study therefore addresses two central questions: (1) What cultural elements contribute to the difficulty in understanding the hidden curriculum? and (2) How can AI-based technologies support foreign guardians in comprehending these cultural dimensions? Japanese school events are key to student and parent integration but contain a "hidden curriculum" (Li, 2017; Li, 2023) that foreign guardians may find difficult to interpret. The lack of culturally adaptive educational tools exacerbates this challenge. Meanwhile, advances in conversational AI, including large language models, have shown potential for personalized support (Chen *et al*., 2024). In predictive learning, Murata *et al*. (2022) demonstrated how time-series analysis and personalization can anticipate learner needs. PAPAMAMA-TOMO incorporates these insights into a single, user-centered support system.

3. System Design: PAPAMAMA-TOMO

PAPAMAMA-TOMO is a chatbot developed to help foreign guardians navigate Japanese elementary school life. It serves as a virtual assistant that can respond to natural language questions about school events, expectations, and required items. Key design principles include accessibility, personalization, and cultural sensitivity.

Main Features include the following:

- Multilingual dialogue (Japanese, English)
- Cultural interpretation support (explains behavioral expectations)
- Visual and audio aids for key terms (planned)
- 24/7 accessibility via web interface

Furthermore, each system feature, such as real-time analytics and early dropout prediction, is designed to directly support the goal of timely intervention and personalized cultural navigation assistance for guardians unfamiliar with Japanese school systems. For example, we ask “What do I need to prepare for the sports day?”Chatbot: “For undoukai (sports day), please bring a lunch box, towel, hat, and indoor shoes. Would you like to hear how to say these items in Japanese?” (Figure 2). Murata *et al.* (2022) showed the importance of adaptive systems in predicting learning needs. Chen *et al.* (2024) found that strategic interaction with chat-based AI enhances comprehension. PAPAMAMA-TOMO adopts both approaches: personalization from behavioral input, and conversational scaffolding to aid understanding.

4. Database Construction and E-Learning Implementation

To support the chatbot system, we also focus on the development of a school culture database. This effort begins with an analysis of the “hidden curriculum” as embedded in school newsletters and documents. By utilizing an existing corpus (www.lixiaoyan.jp), we conduct both quantitative and qualitative analyses to uncover cultural elements such as egalitarianism, motivation-oriented education, learning attitudes, and group dynamics. These elements, along with all major school events (table 1), are systematically organized into a database that supports gamified and chatbot-based learning.

Table 1 Major annual events in Japanese elementary schools

April–July	**September–December**	**January–March**
Opening ceremony	Opening ceremony	Opening ceremony
Entrance ceremony	Disaster drill	Calligraphy contest
Physical fitness test	Open class observation	Off-campus learning
Spring excursion	Autumn excursion	Open class observation
Open class observation	Learning presentation	Graduation ceremony
Sports day	Traffic safety class	Closing ceremony
Pool opening	General cleaning	
Closing ceremony	Closing ceremony	

Furthermore, the system includes the following E-learning implementation goals:

1) Integrating Generative AI into the Chatbot
   Generative AI capabilities enable foreign guardians to engage in self-directed learning through open-ended conversations. For example, when a user expresses curiosity about a cultural event, the chatbot dynamically generates context-sensitive explanations, encouraging deeper exploration based on user interests.
2) Employing Real-Time Analytics for Immediate Support
   To address RQ2 regarding AI-based support for cultural comprehension, PAPAMAMA-TOMO employs real-time analytics. The system continuously monitors user interactions to detect immediate challenges, such as confusion about event preparation (e.g., missing required items)

or frequent inquiries about behavioral expectations. Based on these detected patterns, the chatbot provides context-sensitive prompts, such as checklists for school event preparations or simplified explanations of social norms. This real-time feedback loop enables guardians to receive personalized guidance at the point of need, reducing cognitive stress and promoting active engagement.

3) Data Collection and Iterative Improvement
   Aggregated user interaction data are analyzed to identify common misunderstandings and recurring cultural challenges. Insights from these analyses are used to iteratively refine both the chatbot's scenario database and its interaction strategies, ensuring that the system evolves to better meet the needs of a diverse user base over time.
4) Applying Early Dropout Prediction for Sustained Engagement
   In addition, PAPAMAMA-TOMO applies early dropout prediction techniques to interaction logs to support timely interventions. Indicators such as abrupt cessation of dialogue, prolonged inactivity, and repeated linguistic expressions of confusion (e.g., frequent use of phrases like "I don't understand" or "I'm not sure what to do") are modeled as disengagement risks. When such signs are detected based solely on textual interaction patterns, the system proactively prompts users with encouragement messages, offers simplified re-engagement options, or suggests connecting with human support if needed. This approach, aligned with RQ2, aims to sustain guardians' motivation and prevent feelings of isolation during their cultural adaptation process.

5. Discussion and Conclusion

PAPAMAMA-TOMO offers an inclusive, practical solution for foreign guardians by transforming one-way information delivery into two-way communication. It empowers users to clarify uncertainties about school culture, supports multilingual understanding, and adapts to diverse learning styles. Although empirical testing is scheduled for future research, the current system design draws upon well-established theories in educational AI and cross-cultural communication. Nonetheless, certain limitations remain, including the current restriction to two languages and potential scalability issues when expanding to support a wider variety of cultural backgrounds and educational contexts.

Acknowledgements

This work was supported by the SENTAN-Q Program, Kyushu University, and JSPS KAKENHI Grant Number 25K04342 (PI: LI Xiaoyan).

References

Chen, L., Li, G., Ma, B., Tang, C., Okubo, F., & Shimada, A. (2024). How do strategies for using ChatGPT affect knowledge comprehension? In S. Smith & T. Johnson (Eds.), Proceedings of the 25th International Conference on Artificial Intelligence in Education (AIED 2024) (Vol. 1, pp. 151–162). Springer. https://doi.org/xxxxxxxxxxxx

Jackson, P. W. (1968). Life in classrooms. New York: Holt, Rinehart and Winston.

Li, X. (2017). Japanese language support for foreign parental guardians: An analysis of elementary school handouts and “school culture vocabularies.” Bulletin of the Graduate School of Integrated Sciences for Global Society, Kyushu University, 24(2), 1–12.

Li, X. (2023). Communication with foreign guardians: Completing school newsletters. Tokyo: Kuroshio Publishing.

Li, X., & Honda, H. (2015). Language barriers and cultural adaptation: A study of foreign guardians raising children in Japan. In Proceedings of the Japan Society of Multicultural Education, 19, 45–52.

Murata, R., Okubo, F., Minematsu, T., Taniguchi, Y., & Shimada, A. (2022). Recurrent neural network-FitNets: Improving early prediction of student performance by time-series knowledge distillation. Journal of Educational Computing Research. https://doi.org/10.1177/073563312211297

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. Educational Psychologist, 50(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

以「雙媒材鷹架」促進學生故事創作：桌遊與生成式 AI

Promoting Students' Story Creation through Dual-Media Scaffolding: Integration of Board Games and Generative AI

陳志洪
國立臺灣師範大學 資訊教育研究所
zhchen@ntnu.edu.tw

【摘要】為了提升學生的故事創作能力和動機，在過去的研究已經有許多相關教育理論和應用方法。然而，從文獻回顧中可以發現，不論是著重於「自我創造」或是「素材重混」的工具或平台，大多屬於故事延伸發展，即「從 1 到 100」的促進發展，較少著重於「從 0 到 1」點子概念的發想。因此，本研究提出一個「雙媒材鷹架」的策略，結合「紙卡媒材」和「生成式 AI」，先以桌遊卡牌提供故事骨架，再以生成式 AI 擴展細節。此「雙媒材鷹架」的策略可以幫助學習活動設計和規劃，並落實於教學現場的課程實踐。
【關鍵字】故事創作；動機；媒材設計；教育桌遊；生成式 AI

Abstract: *In order to enhance students' story creation abilities and motivation, many educational theories and applied methods have been proposed in previous research. However, most methods—whether focused on "self-creation" or "remixing existing materials"—primarily support the extended development of stories (the "from 1 to 100" process). There has been relatively little emphasis on the initial ideation of stories (the "from 0 to 1" process). Therefore, this study proposes a "dual-media scaffolding" strategy that combines physical card-based media with digital media. The approach first uses tabletop story cards to provide a basic framework, and then employs generative AI to expand on the story details. This "dual-media scaffolding" strategy can be integrated into curriculum design and implemented as a structured learning activity.*
Keywords: story creation, motivation, instructional material, educational broad game, generative artificial intelligence

1. 前言

為了提升學生的故事創作能力和動機，在過去的研究中，已經出現不少相關教育理論和應用方法的探討。在教育理論部份，鷹架支架（scaffolding）」是一種常見的輔助策略，可以視為一種後設認知的指導方式，協助學生將特定策略應用於適當的學習情境中（Hattie, Biggs, & Purdie, 1996）。此外，「近側發展區（Zone of Proximal Development, ZPD）」理論則是另一種建構的思維（Vygotsky, 1978；Wood, Bruner, & Ross, 1976），透過提供學生適當的支持，以提升故事產出的效率和品質。

從上述教育相關理論可知，故事結構提供了故事創作的整體觀點，讓學生能有系統地檢視與組織自己的故事。因此，一個發展完善的故事結構對於提升故事品質是有幫助的。基於這樣的想法，許多研究進一步提出不同形式的輔助創作工具，包括：(1)「自我創造」工具。有些研究著重於學生的創意，例如在 KidPad 系統（Hourcade, Bederson, & Druin, 2004）與 Crazy Brush 系統（Lee et al., 2010），學生自行繪製角色、物品與場景，創造自己的故事，

透過「創而學（learning by making）」的經驗，幫助學生提升創作產出與興趣。(2)「素材重混（remix）」工具。有些研究則提供「預設」的故事物件（如圖片、聲音、動畫）來輔助創作，例如 FaTe2（Garzotto & Forfori, 2006）、StoryMat（Cassell & Ryokai, 1999）。在這些環境中，學生的創作表現與創意可顯著提升。

然而，雖然這些工具有助於學生的創作品質、興趣和態度，但多數的設計是屬於「從 1 到 100」的促進故事發展，仍然少了「從 0 到 1」的點子創意發想。因此，本研究提出一個「雙媒材鷹架」系統，整合兩種媒材（桌遊和生成式 AI），促進學生的故事點子與創作。具體而言，前者提供基礎的故事素材和結構，並透過桌遊卡牌的「取材」和「組織」設計，讓學生在操作中產出故事，以解決「從 0 到 1」的點子概念生成。後者則進一步透過生成式 AI 的對話，持續「擴展」故事篇幅和「修飾」故事內容，以達到「從 1 到 100」的完整故事創作。本研究包含三個研究子問題：（1）「雙媒材鷹架」是否有助於促進學生故事創作之創意？（2）「雙媒材鷹架」是否有助於提升學生故事創作之品質？（3）「雙媒材鷹架」是否有助於培養學生故事創作之興趣？

2. 「雙媒材鷹架」設計

「桌遊卡牌」和「生成式 AI」這兩種學習媒材，一種是紙卡媒材，另一種是數位媒材，兩者看似差異頗大、毫不相關，但正是因為這樣的差異特性，才能讓兩者互補，整合成更為完整的故事創作鷹架、學習流程，如下圖一所示。

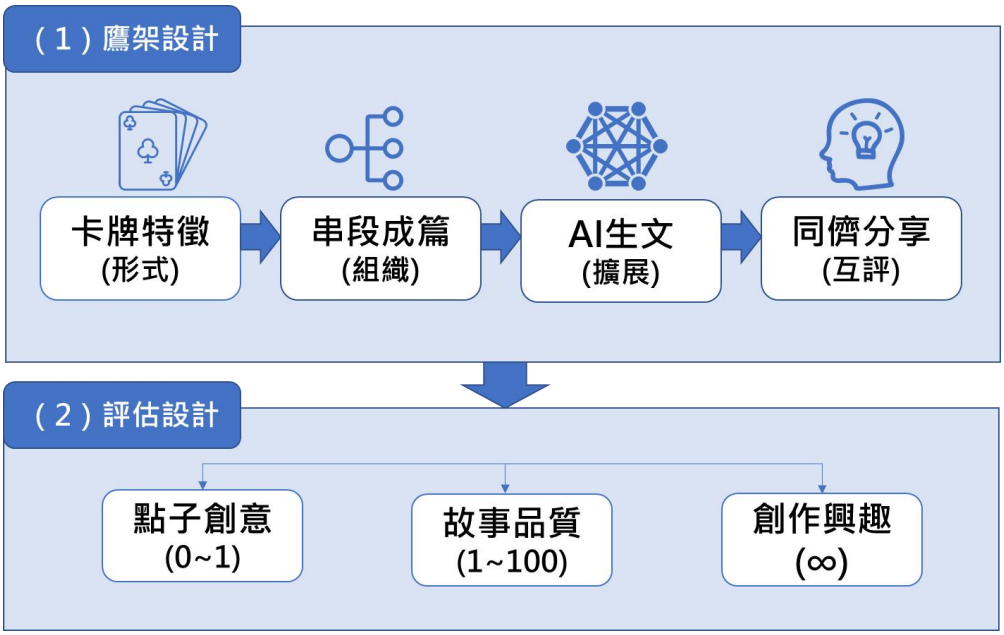


圖 1 故事鷹架與評估設計

首先，「桌遊卡牌」具有引導性和操作性，透過卡牌特徵（如版刑、顏色、符號），可以提供形式鷹架（如開始、發展、結局），再透過卡牌排列（如排序、接龍、分類），串段成篇，形成一篇基本故事的雛形，激發學生的故事創意。本研究以「鬼話連篇」桌遊為例子，該桌遊包含六大經典故事的情節，如白蛇傳、西遊記、聊齋誌異、搜神記等，以這些經典情節為素材，提供熟悉的取材內容，幫助學方快速產生一故事的原型。

其次，「生成式 AI」則將上述故事雛形，透過提示詞指令，進一步加深加廣，在內容上擴展成為具有更完整的故事。本研究基於 ChatGPT 的基礎功能，設計一款 GPTs，以卡牌

故事內容為知識庫，讓學生依序輸入卡牌標題之後，即可以產出完整的故事，做為參考。最後，在「同儕分享」活動中，提供學生同儕交流、互相分享、互評的機會，並帶來滿足感。

(1) 卡牌特徵：「鬼話連篇」桌遊包含六大經典故事的情節，並以不同顏色的卡牌，呈現不同的故事階段，例如：綠色為開始卡、黃色為發展卡、紅色為結局卡。

(2) 串段成篇：除了以顏色和版面設計認識故事的基本元素之外，另外基於三幕式故式架構，透過 6 張卡牌排列（1 張開始卡、4 張發展卡、1 張結局卡），可構成一個完成故事。

(3)AI 生文：以卡牌故事內容為知識庫的 GPTs，提供快速動態產生故事範文，可供學生編故事和說故事之參考。依序將卡牌標題輸入 GAI 中，透過客製化的調整，可以提供學生更多細節，篇幅更長的擴展故事。

(4) 同儕交流：最後以小組為單位，提供「說故事」舞台，讓學生各自扮演「說書人」角色，向同組的學生陳述自己的創作故事。過程中，每位學生也練習當「評審」，鑑賞和評論作品的優缺點。透過兩輪式票選，每一組先選出組內的最佳故事，再針對每組的代表作品，票選出全班的最佳作品。

圖 2 以桌遊卡牌顏色區分故事階段

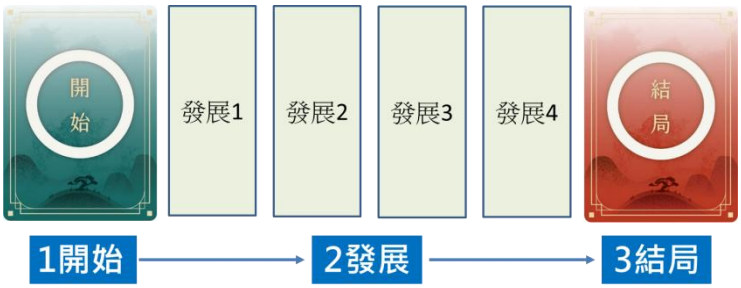


圖 3 排列桌遊卡牌形成故事結構

圖 4 以生成式 AI 擴展文章

圖 5 與同儕分享創作的故事

參考文獻

Cassell, J., & Ryokai, K. (1999). StoryMat: A play space for collaborative storytelling. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 272–273). New York, NY: ACM Press.

Garzotto, F., & Forfori, M. (2006). FaTe2: Storytelling edutainment experiences in 2D and 3D collaborative spaces. In *Proceedings of the 5th International Conference on Interaction Design and Children* (pp. 113–116). New York, NY: ACM Press.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*(2), 99–136.

Hourcade, J. P., Bederson, B. B., & Druin, A. (2004). Building KidPad: An application for children's collaborative storytelling. *Software: Practice and Experience, 34*(9), 895–914.

Lee, Y. C., Liao, C. C. Y., & Chan, T. W. (2010). Crazy Brush: Designing the scribbles environment to improve children's interest in writing. In B. Chang et al. (Eds.), *Proceedings of the 18th International Conference on Computers in Education* (pp. 73–75). Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100.

建立「整合歸納式與演繹式的學習環境」來提升英語文法的學習動機

Establishing an integrated inductive and deductive learning environment to enhance learners' motivation in English grammar learning

梁駿佑[1]，陳攸華[2]
[1]台灣中央大學 網路學習科技研究所
*s113524@g.ncu.edu.tw
* sherry@cl.ncu.edu.tw

【摘要】本研究結合電子閱讀和數位繪本來建立「整合歸納式與演繹式的學習環境」，另外也以實證研究探討認知風格對於使用此學習環境的影響。結果顯示多數學習者動機有所提升。整體型與序列型學習者對於此學習環境的正面觀感相似，然而序列型學習者有較多的負面觀感，特別在於結構、格式與導航細節等方面。本研究的貢獻不僅在於藉著利用電子閱讀和數位繪本來發展「整合歸納式與演繹式的學習環境」；而且根據結果發展一個框架，此框架可用於開發個人化之「整合歸納式與演繹式的學習環境」，以兼顧不同認知風格學習者的差異性，進而提升學習動機。
【關鍵字】認知風格；歸納式學習；演繹式學習；英語繪本

Abstract: *This study integrates digital e-reading and a pictorial e-book to construct an "integrated inductive and deductive learning environment (IIDLE)". Additionally, we examined the impacts of cognitive styles on the use of this learning environment. The results indicated that most learners' motivation was increased. Holists and Serialists exhibited similar positive perceptions for the IIDLE; however, Serialists had more negative feedback, especially structure, format, and navigation details. This study's contributions lie not only in the integration of digital e-reading and a pictorial e-book to create the IIDLE but also in proposing a framework based on the findings. This framework can be utilized to design personalized IIDLEs, which can satisfy the needs of diverse cognitive styles, thus further enhancing their learning motivation.*
Keywords: Cognitive Styles, Inductive Learning, Deductive Learning, Pictorial E-book

1. 前言

語言是人類溝通的重要工具，而文法則是構成語言的核心結構之一(Rossiter, 2021)。因此，在學習英語的過程必須了解英語文法(Dalil & Harrizi, 2013)，英語文法指的是英語中用以組織字詞、形成句子、表達語意的規則與結構，包括時態、語態、詞性、句型等多個層面。學習英語文法的其中一個重要性在於準確地表達、理解與組織語言(Hans & Hans, 2017)。另一個重要性則和溝通能力密切相關。對於非英語母語的學習者而言，扎實的文法基礎有助於提升英語的各項技能，進一步提升整體語言表現(Lin, 2008)。

儘管英語文法有上述的重要性，但是學習英語文法被認為既困難又缺乏吸引力(Ajaj, 2022)，包括文法規則過於抽象，缺少實際應用的情境(Al-Mekhlafi & Nagaratnam, 2011)，以

及因大量記憶的需求，而產生認知負荷(Apolonio, 2021)。這些因素導致學習者產生挫折感，進而影響學習動機與學習成效(Anjomshoa & Sadighi, 2015)。

在另一方面，歸納式學習(Inductive Learning)與演繹式學習(Deductive Learning）在英語文法的習得過程中扮演重要角色(Noveria, 2021)。歸納式學習強調學習者自主探索，他們透過觀察各種語言現象，進而推導出規則(Rungwaraphong, 2021)。例如，學習者可能會先閱讀一篇英語文章或分析多個英語句子，然後藉著歸納發現文法規則。相對地，演繹式學習則是由教師先明確講解規則，學習者再運用這些規則來解決語言問題(Obeidat & Alomari, 2020)。舉例來說，教師可能會直接傳授某項文法概念，提供相關範例，然後要求學生進行英語文法練習。

這兩種學習英語文法的方式各有不同的優缺點，歸納式學習能提升學生的參與感與自主學習能力，但通常需要較長的時間，而演繹式學習則較具效率，但學習成果可能較容易遺忘(Xin, 2012)。因此，若能將兩種方法相互結合，便能在學習過程中取長補短，提升語言學習的整體效果。有鑑於此，本研究藉著數位學習工具來整合歸納式學習和演繹式學習，以幫助學習者增進獲取英語文法知識的動機。

更明確的說，本研究結合電子閱讀和數位繪本來建立「整合歸納式與演繹式的學習環境」，電子閱讀則是用於提供英語文章，讓學習者可以透過閱讀英語文章來主動發掘英語文法規則，而數位繪本則是站在雙碼理論的基礎上，同時透過文字與圖像說明英語文法規則。換言之，電子閱讀則是用來執行歸納式學習，而數位繪本乃是在實施演繹式學習。

電子閱讀可以提供多種功能，例如註釋，超連結和書籤，易於讓學習者展現出較密集的互動(Jou et al., 2016)。而數位繪本在語言學習的過程中，同時利用文字與圖像給予回饋，可以提升正向效果(Sino, 2024)。也就是說，這兩種數位學習工具提供了極佳的互動性與回饋性。所以本研究整合此兩種數位學習工具，使得「整合歸納式與演繹式的學習環境」包含了多項元素；同時學習者也有多項特質，因為個別差異性存在於每一個學習者之間。

數位學習領域的研究者探討各類不同的個別差異性，包括先備知識(Chen & Hsiao, 2024)、性別差異(Chen & Lin, 2024)、學業情緒(Cheng et al., 2023)和認知風格(Chen & Chen, 2025)等。其中認知風格不僅決定了學習者如何處理資訊，還影響他們組織資訊方式的偏好(Lugli et al., 2017)。認知風格有很多面向，學者發現序列型學習者與整體型學習者之差異會影響英語學習，例如 Chen 與 Tseng（2021）設計了一款結合即時回饋與多重鷹架策略的數位英語自我評量系統（SEEL），以支援大學生在動詞時態、被動語態與關係子句等文法概念上的學習。研究結果指出，整體型學習者不僅在後測成績表現優異，學習進步幅度也高於序列型學習者，且展現出較高的彈性與適應力。

之後，Chang & Yang（2023）開發了一款結合不同鷹架設計的數位遊戲式學習平台，用來學習英語中的關係代名詞與關係子句等文法概念。實驗結果發現，整體型學習者在接受具彈性與探索性的軟式鷹架教學下，表現出較高的學習成效、較低的認知負荷與較正向的學習情緒，而序列型學習者則在結構清晰、步驟導向的硬式鷹架教學中學習表現最佳。Iku-Silan 等人（2023）開發了一款具決策導引功能的聊天機器人學習平台，用於國中生進行結合英語語言與原住民傳統醫學知識的跨領域學習任務，他們提出結合導引式結構與即時回饋之設計對序列型學習者特別有利，而整體型學習者則可能需要更多情緒支持與降低認知負荷的策略。

上述的這些研究顯示出認知風格在影響數位英語平台之使用歷程與學習成效上扮演關鍵角色。因此本研究的研究問題乃是「認知風格如何影響學習者回應『整合歸納式與演繹式的學習環境』?」，這也是本研究所探討的研究問題。綜言之，本研究的貢獻重點有兩個面向，

其中一個面向乃是藉著電子閱讀和數位繪本來建立「整合歸納式與演繹式的學習環境」，以提升學習者對於學習英語文法的動機，另一個面向則是透過實證研究來深入了解認知風格對於使用「整合歸納式與演繹式的學習環境」之影響。

2. 整合歸納式與演繹式的學習環境

本研究開發了一個「整合歸納式與演繹式的學習環境」，也就是兼併使用歸納式學習和演繹式學習的方式，來幫助學生學習英語文法，詳述如下:

- 歸納式學習—提供生活中常見的 20 篇英語文章，文章選用範圍依據生活時事新聞，例如：「Welcome to AI Land! 實現 AI 大未來」，允許學習者依據興趣自由選擇想要觀看文章閱讀，選擇文章之後，就可以進入到單篇文章學習頁面，此頁面除了提供文章外，還另外新增文法按鈕，當學習者點擊此按鈕之後，此篇文章的文法重點則被黃色標示出來(圖 1)，這些黃色標示用來幫助學習者去自我歸納英語文法規則，而在歸納英語文法規則的過程中，學習者可以查閱涵蓋圖像與文字的英語文法說明(圖 2)，以便讓他們不僅能夠學習如何進行歸納式學習，而又可以獲得即時的鷹架輔助。
- 演繹式學習—用繪本的方式呈現 31 個英語文法概念(圖 3)，例如關係代名詞、分詞構句和被動語態等。每一個英語文法概念的陳述都包含三個部分：(1)圖像-以視覺化的情境圖解來說明英語文法概念，以增強學習者的認知與記憶效率(2)文字-用文字及例句給予每一個文法概念額外的解釋，以使得學習者能獲得更深入的解析(3)測驗-以五個四選一的選擇題讓學生進行測驗，而且他們可以立即獲知測驗之後的結果，亦即是使用不同顏色的標示讓他們知道自己的答案是否正確，以便能夠自我評量所理解的文法概念是否正確。提供上述三個部分的理念在於讓學習者掌握控制權，也就是他們不僅可以選擇其中任何一個部分來與繪本互動，也可以用導航圖標來查閱他們有興趣的文法概念，其目的在於提升互動性，進而可以使他們有更高的學習動機。

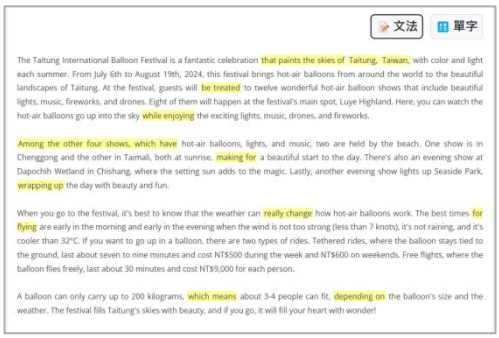


圖 1　時事文章文法重點

圖 2 單篇文章學習頁面

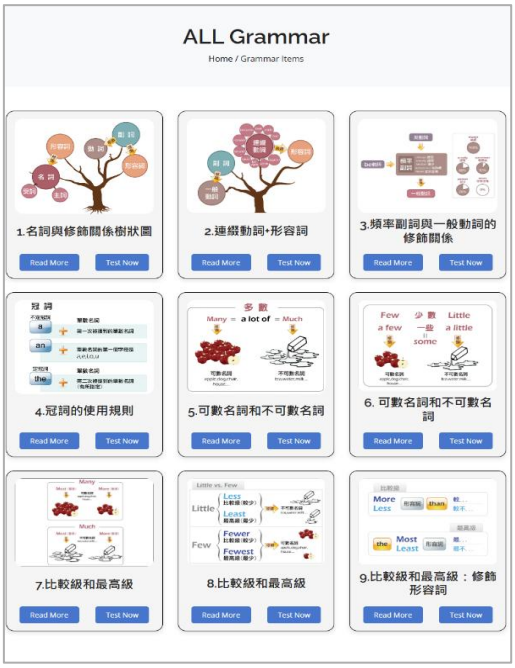


圖 3 演繹式文法繪本

綜上所述，此整合性學習環境具有如下的特色：

■ 整合歸納式與演繹式學習法

本研究融合了歸納式與演繹式兩種學習理論，前者透過電子閱讀的方式，讓學習者藉由閱讀英語文章自主探索並歸納文法規則，後者則運用數位繪本進行明確講解，輔以範例與引導式練習。透過此種整合方式，兼顧學習的參與性與效率性，進一步提升學習成效與動機。

■ 雙碼理論

如近年實證研究亦所證實，雙碼理論（Dual-Code Theory）(Paivio, 1991)可以顯著提升語言學習的效果(Wong & Samudra, 2021)。所以此學習環境依據 Paivio 的雙碼理論，在設計數位繪本時，同時呈現文字與圖像提示。透過圖像與文字資訊同步呈現能有效強化學習效果，因此數位繪本以同步呈現的方式，幫助學習者掌握文法概念與語意內容，提升理解與記憶表現。

■ 靈活與個別化的學習路徑

如 Raj 與 Renumol（2021）在其文獻回顧中指出個人化學習可提升學習成效與滿意度。故此學習環境使用動態連結統整學習資源，並且提供多層次學習路徑，因此學習者可以自由選擇觀看任一篇的英語文章或特定的英語文法單元，也就是允許他們可以依據興趣制訂學習策略，來符合不同學習者的需求與學習目標 (Shemshack et al., 2021)，以實現個人化的學習體驗。

3.研究方法

3.1. 研究工具

本研究使用以下三個實驗工具

■ 學習偏好認知風格問卷（Study Preferences Questionnaire, SPQ）：此工具由 Ford（1985）所開發，用以辨識學習者的認知風格。先前研究(Ku et al., 2016)指出，SPQ 在區分學習者認知風格上相對容易，且具備適當的信度（Cronbach's α = 0.67）。因此，本研究使用

SPQ 辨識學習者的認知風格，也就是整體型或序列型。SPQ 共有 17 個題目，每題包含兩個描述，分別是整體型描述和序列型描述。若選擇整體型描述的次數較多，則歸類為整體型學習者，反之，若選擇序列型描述的次數較多，則歸為序列型學習者。

- 開放式問卷：為了讓學習者能用自己的語言表達對於使用「整合歸納式與演繹式的學習環境」的觀感，本研究設計一份開放式問卷。此份開放式問卷共分為五大項目—(1) 探討學習者在互動過程的前後動機是否有所改變；(2)分析整體設計的優點和缺點；(3)分析互動操作的優點和缺點；(4)分析繪本格式與學習介面的優點和缺點；(5)分析顏色選取的優點和缺點。涵蓋此五大項目之目的在於分析學習者實際使用歷程中的動機變化、主觀感受與具體建議。
- 前測與後測：前測包含 10 個句子，每題皆設有四個空格，每一格 2.5 分，滿分為 100 分，學習者需從所提供的詞彙表中選擇適當的詞填入空格，使句子語意完整且文法正確。所有學習者的測驗表現將作為前測成績並進行紀錄。關於後測，其題型、題目數量皆與前測相同，藉以觀察受測者學習成效的變化。

3.2. 研究流程

本研究的參與對象為 12 名台灣北部某所大學之研究生，參與者皆具有基礎英文知識並具有基礎的電腦操作能力。首先，學習者會先填寫 SPQ，最終結果為六名整體型和六名序列型學習者。然後再以前測來評估每位參與者的英語文法能力。之後，每位受測者則開始與「整合歸納式與演繹式的學習環境」互動，互動時間並沒有限制，當互動結束之後，參與者則會進行後測，藉以觀察受測者在與「整合歸納式與演繹式的學習環境」互動前後的差異，最後每位參與者需用 Google 表單填寫開放式問卷，以表達他們的使用觀感。

4.結果

本研究採用整合性的資料收集和分析方式，也就是兼顧質性與量化資料，在質性資料方面，則是以開放性問卷來收集學習者對於使用「整合歸納式與演繹式的學習環境」的動機與觀感，然後進一步歸納出學習者在正向與負向評價中的具體表現與特徵，在量化資料方面，則是先分析開放性問卷回覆的關鍵字，然後再統計整體型學習者和序列型學習者於各關鍵字出現的次數，以便能夠區別不同認知風格學習者在介面評價中的相似性與相異性。不論是質性資料或是量化資料，皆是在探討「整合歸納式與演繹式的學習環境」的設計是否能促進學習動機及滿足不同認知風格學習者的多元需求。

4.1 繪本使動機由低轉高

本研究的結果顯示，六位受試者在過去學習英文時皆沒有很高的學習動機，造成他們沒有學習動機的原因包括「文法結構太多太繁雜」，「學習英文太困難和無聊會想放棄」等。然而在使用英語文法繪本之後，其中有五位受試者的學習動機有明顯改變，提升動機的原因包括「利用繪本的圖文並茂方式幫助快速理解」，「整理出文法的精華可以易於查閱」。而再細部分析整體型學習者和序列型學習者對於問卷的回應，發現兩種認知風格的學習者呈現許多相似的正面觀感，例如:「操作按鈕時會給予回饋」，「操作直覺，可以很快找到想要的東西」。這些相似的正面觀感有可能是另外一個原因來解釋為甚麼學習者的動機可以由低轉高。

簡言之，此「整合歸納式與演繹式的學習環境」的設計不僅以雙碼理論為基礎，而且也提供回饋性的操作的方式。關於前者，此「整合歸納式與演繹式的學習環境」以圖像呈現英語文法知識，再輔以文字的說明，使得學習者能夠從多重的學習管道獲得英語文法知識。關

於後者，此「整合歸納式與演繹式的學習環境」提供互動式的操作方式，可以使學習者得到充分的回饋。上述的兩項設計元素不僅促使學習者能夠更快速地掌握文法概念，也進而提升了參與度與學習動機。

4.2 認知風格影響負面觀感

如章節 4.1 所顯示，整體型學習者和序列型學習者具有相似的正面觀感，此相似性在「量」與「質」兩方面皆有出現，在量的方面，整體型學習者和序列型學習者之正面觀感的數量無明顯差異（整體型：N=28，序列型：N=29），在質的方面，兩方均提出了相似的正面觀點，如「頁面清楚易懂能快速找到需要的內容」、「操作直覺，可以很快找到我想要的東西」以及「自由操作，容易上手」，此發現意味著「整合歸納式與演繹式的學習環境」的優點可以被兩種認知風格的學習者發覺。

另一方面，整體型學習者和序列型學習者卻具有相異的負面觀感。更明確的說，序列型學習者提出的負面觀感多於整體型學習者（整體型：N=22，序列型：N=26），而在細部檢查序列型學習者所提出的負面觀感之後，發現可以區分為四大類:(1)文法資訊的擷取 (N=8)—例如，內容排序單一，缺少有紀錄指出哪些文法是曾經看過的，(2)內容呈現的方式(N=5)—例如，字體不夠大，文章頁面的頁碼不夠粗，(3)介面元素的設計(N=10)—例如，排版或顏色無法任意調整，頁尾選取按鈕不夠大，(4) 操作流程的控制(N=2)—例如，網站的操作不夠順暢，操作步驟太多。

上述的發現與 Chen & Chen (2025)所得到的結果相似，也就是序列型學習者在與數位學習系統互動時較容易遇到困難，此可能因為序列型學習者屬於線性思考，而數位學習系統多屬於非線性的設計，因此會遇到比較多的障礙 (Chen & Tseng, 2021)。數位學習系統需慎重考慮此項發現，才能兼顧整體型學習者和序列型學習者之需求，以使得本研究所設計的「整合歸納式與演繹式的學習環境」或其他數位學習系統能夠提高兩種認知風格的學習者之學習動機。

4.3 整體與序列有不同重點

除了正面觀感與負面觀感之外，兩種認知風格的學習者也有不同的重視面向，整體型學習者側重於整體美感，也就是對於介面視覺設計回饋與系統自由度有較多的評價(N=17)，例如，「特別的功能都有加上比較顯眼的顏色」與「點選按鈕有特效反應可以增加學習樂趣」。另外他們也關心流暢度(N=19)，特別是操作的便利度與學習流程的自由度，例如，「可以直接選擇測驗很方便」和「先解釋文法再提供例句的順序很符合我學習的習慣」。

相反地，序列型學習者則注重邏輯架構(N=22)，多數序列型學習者強調結構清晰與進度追蹤的重要性，例如他們提到「沒有學習記錄，會有點迷路」、「不同的內容，用了不同區塊，容易閱讀」等觀點，此可能因為序列型學習者仰賴現有教材內容的架構，而非自己重新組織資訊，來產生一個適合自己的學習路徑 (Hsieh & Chen, 2016)。同時，序列型學習者也需要較多的顯著線索(N=12)，例如「文字可以增添不同色彩來提升吸引力」、「例句學習可以在多提供一些範例」。這些發現與 Ku 等人(2016)研究之結果一致，也就是序列型學習者需要較多的輔助支援。

上述的發現與兩種認知風格的特質相呼應，亦即是整體型學習者著重於整體視覺效果和自由選擇學習路徑，而序列型學習者看重是否可以得到結構化的輔助支援。基於這些發現，數位學習系統的設計需提出能滿足不同認知風格學習者偏好之解決方案。例如，透過鮮明的視覺元素與直觀的導航設計來增強整體型學習者之體驗，而又提供清晰的邏輯結構與輔助工

具來幫助序列型學習者建立自己的知識架構。這些解決方案的終極目標乃是期待整體型學習者與序列型學習者之學習動機皆能提升。

4.4 操作與格式各為優缺點

在細部檢查學習者對於開放式問卷的細部回應之後，發現「整合歸納式與演繹式的學習環境」的優點多集中於操作層面(N=30)，如:「每個操作都是知道會達成哪些目的」、「按鈕方便操作，有回饋感」、「方形字卡在滑過去時都有給反應回饋」等。而對於這些優點的回應大部分是來自於整體型學習者(N=18)，而非序列型學習者(N=12)。相反的，「整合歸納式與演繹式的學習環境」的缺點出現於格式層面(N=35)，例如:「切換頁的導航按需放大避免誤觸其它按鈕」、「可以利用深色網底表示目前頁面」等，而對於這些缺點的回應大部分是來自於序列型學習者(N=21)，而非整體型學習者(N=14)。

這些發現不僅再一次顯示出認知風格會影響學習者對於使用「整合歸納式與演繹式的學習環境」的觀感，並且也呼應章節 4.2 的結果，也就是序列型學習者有比較多的負面觀感，因此他們能夠對於「整合歸納式與演繹式的學習環境」指出較多的缺點，而這些缺點很大一部分出現在格式層面，格式層面的缺點大多屬於細節設計的部分。序列型學習者強調程序上的細節，而不是整體架構，這也許就是為甚麼他們可以指出許多格式上的缺點(Wang et al., 2018)，例如用粗體標示解釋、例句和頁碼，自行調整顏色和內容順序等。

有鑒於此，需要一個解決方案來克服序列型學習者所指出的缺點，以降低他們的負面觀感，而客製化也許就是其中一個解決方案，客製化將控制權給予學習者，使他們能夠自行修改內容呈現、格式佈局和導航功能(Treiblmaier et al., 2004)，因此將客製化整合至「整合歸納式與演繹式的學習環境」將可允許序列型學習者按照自己的細部需求調整格式，以減少負面情緒，進而提高學習動機。

5.結論

本研究結合電子閱讀和數位繪本來建立「整合歸納式與演繹式的學習環境」，探討學習者的學習動機變化與使用觀感，並從認知風格的角度比較整體型與序列型學習者的回應差異。實證研究結果顯示「整合歸納式與演繹式的學習環境」可以提升學習英語文法的動機，主要是因為此學習環境具有圖文並茂的設計與即時的回饋，有助於理解英語文法。除了提高動機之外，兩種認知風格的學習者對此學習環境皆有一些正面觀感，如操作直覺、內容清晰。然而序列型學習者卻也呈現了一些負面觀感，包括結構邏輯、格式與導航設計。

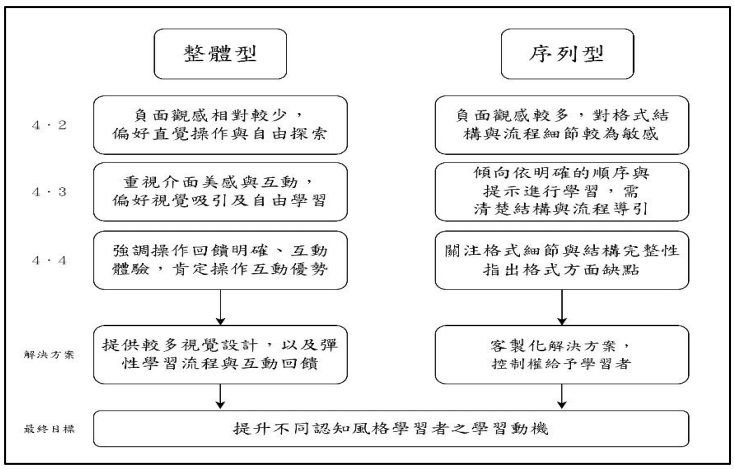


圖 4 研究結果的框架

呈上所述，本研究的貢獻不僅在於技術面，也出現在應用面。關於技術面，本研究藉著利用電子閱讀和數位繪本來發展「整合歸納式與演繹式的學習環境」，使學習者可以根據自己的喜好，來採用適當的方式學習英語文法，以提高他們的學習動機和參與度。關於應用面，本研究根據實證探討的結果發展了一個框架(圖 4)，此框架敘述整體型學習者與序列型學習者的觀感差異，故可以用於開發個人化之「整合歸納式與演繹式的學習環境」。

雖然有上述的多面向貢獻，但是本研究仍存在限制，例如樣本數量有限，且未涵蓋其它個別差異變項，如性別、語言程度等。未來研究可進一步擴大樣本範圍，並納入更多個別差異變項，使此種整合歸納式與演繹式的學習環境更能滿足每一位學習者的需求。

參考文獻

Ajaj, I. E. (2022). Investigating the difficulties of learning English grammar and suggested methods to overcome them. *Journal of Tikrit University for Humanities*, *29*(6), 45-58. https://doi.org/10.25130/jtuh.29.6.2022.24

Al-Mekhlafi, A., & Nagaratnam, R. P. (2011). Difficulties in teaching and learning grammar in an EFL context. *International Journal of instruction*, *4*(2).

Anjomshoa, L., & Sadighi, F. (2015). The importance of motivation in second language acquisition. *International Journal on Studies in English Language and Literature (IJSELL)*, *3*(2), 126-137.

Apolonio, J. A. (2021). Needs Analysis, Challenges, and Teaching Strategies in English Grammar Learning. *Online Submission*.

Chang, C.-C., & Yang, S.-T. (2023). Interactive effects of scaffolding digital game-based learning and cognitive style on adult learners' emotion, cognitive load and learning performance. *International Journal of Educational Technology in Higher Education*, *20*(1), 16. https://doi.org/10.1186/s41239-023-00385-7

Chen, S. Y., & Chen, C.-Y. (2025). An investigation of the impacts of digital entertaining English logic learning: Holists vs. serialists. *International Journal of Human–Computer Interaction*, *41*(2), 906-918. https://doi.org/10.1080/10447318.2024.2306435

Chen, S. Y., & Hsiao, T.-C. (2024). The effects of incorporating an electronic book into digital game-based learning: A prior knowledge perspective. *International Journal of Human–Computer Interaction*, *40*(19), 6015-6026. https://doi.org/10.1080/10447318.2023.2247595

Chen, S. Y., & Lin, Y.-Z. (2024). An investigation of a customizable entertaining animated e-book: A gender difference perspective. *International Journal of Human–Computer Interaction*, 1-12. https://doi.org/10.1080/10447318.2024.2364467

Chen, S. Y., & Tseng, Y.-F. (2021). The impacts of scaffolding e-assessment English learning: A cognitive style perspective. *Computer Assisted Language Learning*, *34*(8), 1105-1127. https://doi.org/10.1080/09588221.2019.1661853

Cheng, C.-Y., Chen, J.-M., & Chen, S. Y. (2023). The influences of academic emotion and prior knowledge in the context of online tests. *Interactive Learning Environments*, *31*(5), 3257-3272. https://doi.org/10.1080/10494820.2021.1924793

Dalil, Z., & Harrizi, M. (2013). The importance of grammar in second language teaching. *Unpublished Manuscript). Universite Hassan II, Casablanca, Morocco*.

Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology*, *16*(1), 65-77. https://doi.org/10.1111/j.1467-8535.1985.tb00483.x

Hans, A., & Hans, E. (2017). Role of grammar in communication–writing skills. *International Journal of English Language, Literature and Humanities*, *5*(1), 39-50. https://doi.org/10.24113/ijellh.v5.issue1.42

Hsieh, C.-W., & Chen, S. Y. (2016). A cognitive style perspective to handheld devices: customization vs. personalization. *International Review of Research in Open and Distributed Learning*, *17*(1), 1-22. https://doi.org/10.19173/irrodl.v17i1.2168

Iku-Silan, A., Hwang, G.-J., & Chen, C.-H. (2023). Decision-guided chatbots and cognitive styles in interdisciplinary learning. *Computers & education*, *201*, 104812. https://doi.org/10.1016/j.compedu.2023.104812

Jou, M., Tennyson, R. D., Wang, J., & Huang, S.-Y. (2016). A study on the usability of E-books and APP in engineering courses: A case study on mechanical drawing. *Computers & education*, *92*, 181-193.

Ku, O., Hou, C.-C., & Chen, S. Y. (2016). Incorporating customization and personalization into game-based learning: A cognitive style perspective. *Computers in Human Behavior*, *65*, 359-368.

Lin, L.-L. (2008). The Role of Grammar Teaching in Writing in Second Language Acquisition. *Online Submission*.

Lugli, L., Ragni, M., Piccardi, L., & Nori, R. (2017). Hypermedia navigation: Differences between spatial cognitive styles. *Computers in Human Behavior*, *66*, 191-200. https://doi.org/10.1016/j.chb.2016.09.038

Noveria, A. (2021). The effects of inductive-deductive grammar instruction on students' grammatical accuracy. *ELS Journal on Interdisciplinary Studies in Humanities*, *4*(3), 316-321. https://doi.org/10.34050/elsjish.v4i3.18123

Obeidat, M. M., & Alomari, M. d. A. (2020). The Effect of Inductive and Deductive Teaching on EFL Undergraduates' Achievement in Grammar at the Hashemite University in Jordan. *International journal of Higher education*, *9*(2), 280-288. https://doi.org/10.5430/ijhe.v9n2p280

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255. https://doi.org/10.1037/h0084295

Raj, N. S., & Renumol, V. (2022). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. *Journal of Computers in Education*, *9*(1), 113-148.

Rossiter, A. (2021). The Importance of Grammar. *Online Submission*.

Rungwaraphong, P. (2021). Inductive Teaching Approaches in Business English Writing in an EFL Context: Paper-Based and Product-Based Instructions. *English Language Teaching Educational Journal*, *4*(1), 34-48.

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education*, *8*(4), 485-503.

Sino, R. (2024). The Impact Of Visual Support On Efl Learners' Vocabulary Acquisition When Reading: Självständigt arbete på grundnivå (yrkesexamen), 10 poäng/15 hp. In.

Treiblmaier, H., Madlberger, M., Knotzer, N., & Pollach, I. (2004). Evaluating personalization and customization from an ethical point of view: An empirical study. 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the,

Wang, J.-H., Chang, L.-P., & Chen, S. Y. (2018). Effects of cognitive styles on web-based learning: Desktop computers versus mobile devices. *Journal of Educational Computing Research*, *56*(5), 750-769.

Wong, K. M., & Samudra, P. G. (2021). L2 vocabulary learning from educational media: Extending dual-coding theory to dual-language learners. *Computer Assisted Language Learning*, *34*(8), 1182-1204.

Xin, J. (2012). The effect of integrating inductive approach and deductive approach with multimedia assistance into acquisition of subjunctive mood. *Sino-US English Teaching*, *9*(9), 1510-1515.

年齡對國小學童使用虛擬實境交通安全學習系統之影響

Impact of Age on the Use of Virtual Reality Traffic Safety Learning Systems by Primary School Children

楊雅婷[1]，賴冠廷[1]，周孟慶[2]，沈彤[2]，黃國豪[2]，楊晰勛[1]，李佳蓉[3*]，陳品誌[4]
[1]國立雲林科技大學 數位媒體設計系
[2]國立雲林科技大學 智慧數據科學研究所
[3]國立雲林科技大學 通識教育中心
[4]國立雲林科技大學 資訊管理系人工智慧技優專班
[*] leecr@yuntech.edu.tw

【摘要】本研究開發了一套高互動沉浸式虛擬實境交通安全學習系統，並探討年齡對國小學童之學習成效、學習滿意度、暈眩感及臨場感的影響。實驗對象為臺灣中部的兩間國小共53位學生，以學生年齡區分為6至9歲及10至12歲兩個組別進行比較。結果發現，年齡對國小學童學習交通安全的學習成效及學習滿意度有顯著影響，10至12歲學童的學習滿意度及學習成效顯著優於6至9歲的學童，而臨場感與暈眩感則沒有顯著的差異。另外也發現，年齡與學習滿意度及學習成效均呈現正相關的關係。虛擬實境的臨場感也和滿意度間呈現正相關的關係。暈眩感則和其它構面都沒有相關性，平均值也不高，在可接受的範圍。
【關鍵字】年齡；國小學童；虛擬實境；臨場感；暈眩感

Abstract: *This study used a self-designed virtual reality traffic safety learning system to explore the effects of learning effectiveness, learning satisfaction, dizziness and sense of presence among primary school children of different ages. A total of 53 students from two primary schools in the central region were divided into two groups, aged 6 to 9 and 10 to 12 years old. It can be seen from the results that this teaching material has a significant impact on the learning effectiveness and learning satisfaction of primary school children, and it is found that age does affect learning satisfaction and learning effectiveness, and the sense of presence of virtual reality will also indirectly affect learning satisfaction and maintain a low sense of dizziness.*
Keywords: Age, primary school children, Virtual reality, a sense of presence, Dizziness

1. 前言

當前虛擬實境（VR）技術的發展引起了教育領域的極大關注。對國小學童來說，VR系統的使用效果與他們的年齡密切相關（Digennaro & Visocchi, 2024），虛擬實境逐漸融入教學現場的時代讓兒童和青少年的學習者身分有所改變，學習者以角色扮演與沉浸式體驗學習特定主題或時事，強調以身體素養建構中各種重要要素的發展，也能藉由沉浸感增加平面座標的感官反應（Cuturi et al., 2023），虛擬實境技術更是可以於不同類別的教育現場發揮一定效用，像是Mariscal Vivas 等人（2020）研究如何提高救護員的實際操作技術，建立沉浸式虛擬實境技術讓實驗組學生模擬於交通緊急狀況進行急救訓練體驗，並設計一般虛擬實境桌面技術讓對照組學生在實驗室中模擬事故，比較是否加入沉浸式體驗，對救護員對活動的

滿意度和工具適用性為何，結果發現在虛擬實境融入沉浸式體驗的學習效果更好，這與 Salameh 等人（2024）提高護理教育中滿意度、自信心和表現的干預措施的研究結果有相似之處，若在護理師培訓課程中適當加入 VR 模擬技術，能為護理學生提供了一個安全且有效的教學環境，又或者 Alsswey 等人（2024）將 VR 重力草圖工具應用於建築設計的大學課程，提升學生對課程教材的可用性和適用性滿意度，比起傳統草圖工具的大學課程，使用 VR 重力草圖工具的學生更能提升自我效能與學習策略，增加建築設計學生的模擬訓練效果。由上述可知，若將 VR 模擬技術與傳統教學活動結合，除了可以減少教材準備不便、實作課程體驗效果有限及臨床課程難以控制等問題產生，更能利用 VR 技術模擬更多現實中較不易發生的特殊事件與場景，改善實做所面臨的體驗限制問題，故本研究為自行設計之國小學童虛擬實境交通安全學習系統中，增加沉浸式體驗與沉浸式評量，讓學童在學習國內交通安全的課題中能換為思考，理解行人與駕駛者皆應注意的行車安全問題，並提升道路防範意識，行人與駕駛者的用路安全。

不同年齡對 VR 系統的教學感受差異甚遠，例如 Cao 等人（2024）在少數民族聚集區的兩所學校，分別找 73 名小學四年級學生和 86 名國中二年級學生，比較不同年齡對虛擬實境科技教育的體驗差異，結果發現小學生會因為沉浸感與想像力增加學習成效，中小學生皆對提升問題解決能力與自我效能感有顯著影響，且同時擁有極高的科技接受度，不會因為新穎科技融入教學現場而影響學生學習行為，反而提升學生的素養表現。即便是職業類科的學生也會因擴增實境融入沉浸體驗，提升機械工程專業的技術，Conesa 等人（2024）的研究在車床實際操作課程上加入沈浸式擴增實境，除提升學生的學習成效外，同時減少學生在此類課程的危險程度。雖然沉浸式虛擬實境是加強車床操作訓練的寶貴工具，提供引人入勝且真實的體驗，但部分使用者提出有輕微頭暈等身體不適的症狀，暈眩症等問題在 Cao 等人（2024）對國小學生的沉浸虛擬實境的體驗也有發現，暈眩感的出現是直接或間接影響學習成效或學習情緒的重大問題之一，在 Lonne 等人（2023）的研究就提到，增加沉浸感的擴增實境，能正面提高培訓課程的臨場感與學習成效，也會因為沉浸感的提升而增加正面情緒的產生，例如幸福感等情感，但暈眩感的出現會降低幸福感的產生，進而影響使用沉浸式擴增實境的情感效果，針對上述提及的暈眩問題，應於設計沉浸感虛擬實境的同時，注意避免或減少暈眩感等不良問題的產生，故本研究為了檢測所設計之國小學童虛擬實境交通安全學習系統的暈眩感受狀況，將實驗對象分為 10-12 歲及 6-9 歲兩個年齡層進行比較，理解暈眩感是否會因年齡差異而有所影響。

本研究以交通部（2022）的交通安全教育指引手冊，行人用路安全、大車行駛危機及道路交通設施與通行路權為學習內容，設計了一套具備相關知識且結合沉浸式虛擬實境的國小學童虛擬實境交通安全學習系統，並探討下列研究問題：

(1).不同年齡層的國小學童使用 VR 交通安全學習系統的學習滿意度是否有顯著差異？

(2).不同年齡層的國小學童使用 VR 交通安全學習系統的學習成效是否有顯著差異？

(3).不同年齡層的國小學童使用 VR 交通安全學習系統的臨場感是否有顯著差異？

(4).不同年齡層的國小學童使用 VR 交通安全學習系統的暈眩感是否有顯著差異？

(5).小學學童的年齡對於使用本研究所設計之交通安全 VR 系統時學習滿意度、學習成效、臨場感及暈眩感之間的關係為何？

2.系统介绍

2.1.1 學習目標

本研究系統為國小學童所設計的虛擬實境交通安全學習系統，設計原理是將交通安全教學與肢體行為進行結合，並利用角色扮演與虛擬故事模擬實際道路上可能發生的一切事故與危機，學生操作如圖 1 (A)部分所示，以此達成以下的學習目標：

(1).理解交通規則的重要性：將現實交通規則轉化為虛擬故事情境，透過實際體驗事件，讓學童能深刻理解規則並應用於生活中。

(2).體驗駕駛員視角的挑戰：讓學童以駕駛員的角度模擬公車視線盲區及車道突發狀況，除學習常見號誌的識別與應用外，讓學童注意大車視線盲區的重要性，以便在現實生活中學會避免此類危險發生。

(3).提升反應能力：在體驗環節中增加部分突發情境，例如行人闖入車道或急煞等情況，訓練學童如何在發生緊急事故時能快速且正確地應對，以此了解事故的危險性。

(4).強化行人安全行為：讓學童透過虛擬實境中模擬行人違規穿越馬路等事件，指導學童了解該如何安全並正確穿越馬路，並學習避開危險區域，學會解讀交通標誌和標線，建立正確的行人用路安全意識。

2.2 硬體設備

本研究系統之硬體設備為虛擬實境頭戴顯示器，也就是 Meta Quest 2，讓學童能夠透過雙眼清晰觀察交通標誌、地面標示及其他細節，以動作追蹤裝置，也就是 Oculus Touch 控制器，以雙手控制把手實現手勢追蹤功能，並透過內建喇叭重現實生活會聽到的車輛喇叭聲及急煞聲，增加學童在體驗時的方向感知，硬體設備如圖 1 (B)部分所示。

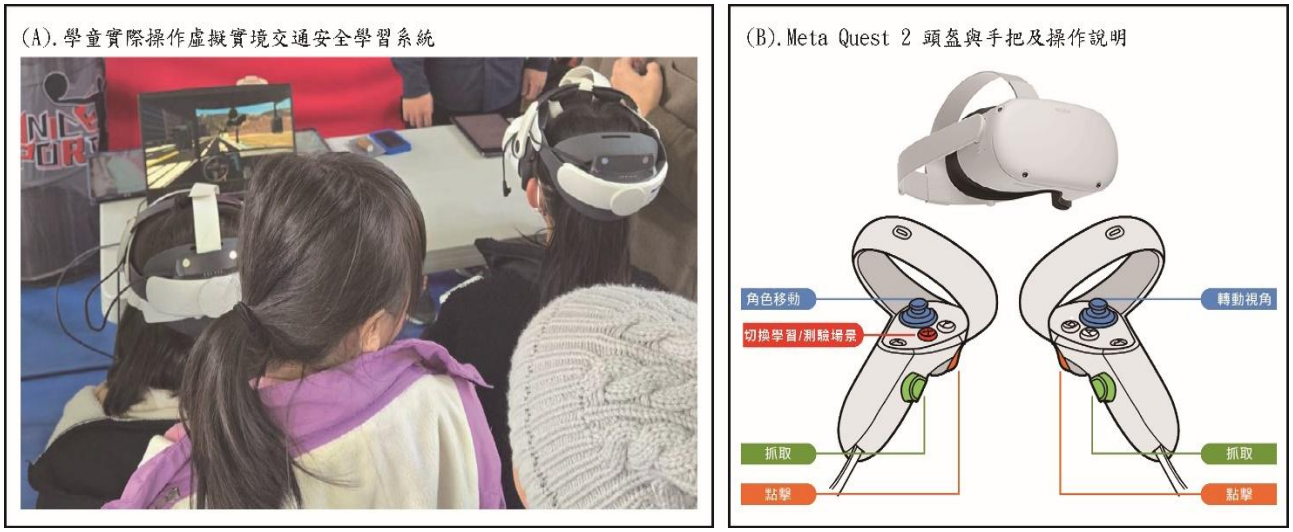


圖 1 學生操作虛擬實境交通安全學習系統及 Meta Quest 2 頭盔與手把的操作說明

2.3 场景设计

本研究系統之場景設計以針對不同的學習目標設計三個模擬場景進行定制化開發，分別為都市公車駕駛場景、高速公路駕駛場景及綜合評量密室逃脫場景，場景設計如圖 2 所示，以下針對各式場景與學習重點逐一說明：

(1).都市公車駕駛場景：從公車駕駛員視角行駛於虛擬都市場景，讓學童了解公車內部結構及駕駛視野，特別是駕駛座周邊的盲區情況，以及模擬駕駛轉彎時的實時交通狀況，指導學童在道路與轉彎時的盲區範圍，以及理解自我保護與行駛的應對策略為何。

(2).高速公路駕駛場景：從公車駕駛員視角行駛於高速公路場景，透過高度還原高速行車場景，結合開窗、強風、高速行駛與急煞等挑戰，提升學童認識高速公路上見的號誌與標線規則，並提升在高速行駛的正確應對概念。

(3).綜合評量密室逃脫場景：將虛擬都市場景與高速公路場景結合密室逃脫規則，讓學童在遊戲化情境評量場景進行答題，並參考核心素養的概念，將虛擬與實際案例結合，讓學童在體驗模擬事故時檢視所學的知識並牢記，以預防現實生活可能遭遇的危機。

圖 2 都市公車駕駛場景、高速公路駕駛場景及綜合評量密室逃脫場景畫面

3.研究方法

3.1 研究對象

本研究之研究對象與資料分析方法，研究對象為中部地區某國小共 53 位學童，其中男生為 34 人，女生為 19 人，年齡介於 6 至 12 歲之間，其中 6-9 歲為 18 人，10-12 歲為 35 人。

3.2 實驗流程

本研究之實驗流程如圖 3 所示，整個實驗共進行約 25 分鐘，分為 6～9 歲及 10～12 歲兩組，包含操作講解 5 分鐘、VR 體驗 15 分鐘，以及填寫問卷表單 5 分鐘。問卷內容涵蓋學習滿意度（評估對實驗的整體感受）、臨場感（測量 VR 體驗的沉浸程度與真實感受）、暈眩感（了解體驗過程中是否有不適反應）及交通安全測驗（檢驗對交通安全知識的掌握）。

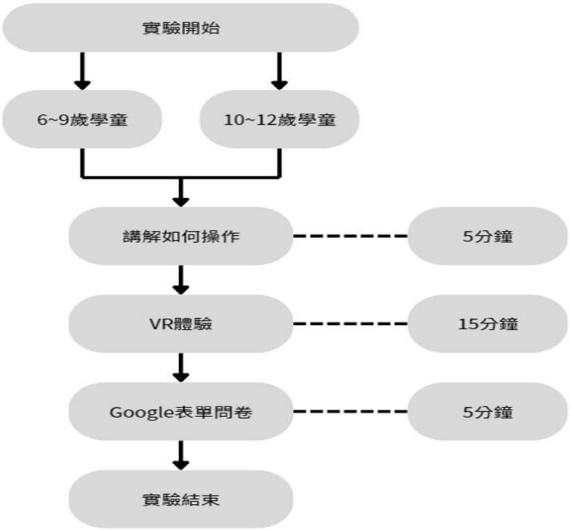


圖 3 實驗流程圖

3.3 研究工具

本研究之研究工具分為一套虛擬實境交通安全學習系統，一份有關學習滿意度、臨場感與暈眩感問卷，以及交通安全學習評量，其內容如下所示：系統為自行設計之虛擬實境交通

安全學習系統，以操作虛擬實境頭戴顯示器 Meta Quest 2，進行虛擬城市與高速公路場景體驗公車在現實常見之道路號誌與規範，學習發生突發狀況與事故時應如何預防，以行為體驗建立道路駕駛與行人用路安全意識，虛擬實境交通安全學習系統工具詳細說明請詳閱第三章系統介紹；學習滿意度問卷參考至 Chu 等人（2010）所提出的學習滿意度量表，此量表一共十題，臨場感問卷參考至 Hwang 等人（2022）所提出的臨場感量表，此量表一共五題，暈眩感問卷參考至陳昕昊（2024）所提出的暈眩症量表，此量表一共五題，三份問卷皆以李克特五點量表所示，其中 5 分為「非常同意」，1 分為「非常不同意」；交通安全學習評量參考至交通部（2022）的交通安全教育指引手冊，第二章節 6-17 歲兒童及少年交通事故概況的用路人角色與交通事故說明，並依據交通安全教育課程架構規劃進行文字題目設計，因應教育部推廣提升核心素養能力，本研究參考教育部（2021）十二年國教課程綱要總綱，在國小社會領域的「校園安全」與「家鄉運輸」兩個單元中結合情境式評量，讓學童以情境題目替代文字題目，加深交通安全的重要性與生活應用能力。

3.4 資料分析

本研究採用獨立樣本 *t* 檢定檢測不同年齡層的國小學童學習滿意度、學習成效、臨場感與暈眩感是否有所差異，以皮爾森積差相關分析年齡與學習滿意度、學習成效、臨場感與暈眩感是否有正負相關性。

4.結果與討論

本研究之研究結果如下，會先以信度分析檢查學習滿意度、臨場感與暈眩感問卷的信度是否可信，再以獨立樣本 *t* 檢定分別分析年齡是否分別對學習滿意度、學習成效、臨場感與暈眩感有所差異，最後再以皮爾森積差相關分析年齡與學習滿意度、學習成效、臨場感與暈眩感的相關程度。

4.1 學習滿意度、臨場感與暈眩感問卷信度分析結果

信度分析的結果如表 1 所示，學習滿意度、臨場感與暈眩感的 Cronbach's Alpha 值均大於 0.8，顯示本研究的信度相當良好。整體而言學生的學習滿意度(*M=4.645*)與臨場感(*M=4.121*)分數不錯，且暈眩感(*M=2.464*)的感受不高，不過學生間個別差異較大(*SD=1.414*)。

表 1 實驗對象之描述性統計

構面	題數	Cronbach's Alpha 值	平均數	標準差
學習滿意度	10	0.891	4.645	0.582
臨場感	5	0.810	4.121	1.019
暈眩感	5	0.918	2.464	1.414

4.2 不同年齡層國小學童的學習滿意度結果

採用獨立樣本 *t* 檢定確認不同年齡層的國小學童對使用虛擬實境交通安全學習系統之學習滿意度差異，結果如表 2 所示。從結果可以發現 10-12 歲的學童在學習滿意度顯著高於 6-9 歲的學童（$t = 2.382, p < .05$），顯示此系統對 10-12 歲的學童在學習交通安全的效果更好，針對此結果的推測是年紀稍大的學童在使用科技的熟練度與接受度較高，但也能因年紀稍小的學童人數較少所導致，未來可再增加低年齡層的實驗對象以驗證其差異。

表 2 不同年齡層的國小學童學習滿意度結果

構面	年齡	樣本數	平均數	標準差	*t*
學習滿意度	10-12 歲	35	4.809	0.309	2.382*

	6-9 歲	18	4.328	0.827

*$p < .05$

4.3 不同年齡層的國小學童的學習成效結果

採用獨立樣本 *t* 檢定確認不同年齡層的國小學童對使用虛擬實境交通安全學習系統之學習成效差異，結果如表 3 所示。從結果可以發現 10-12 歲的學童在學習成效顯著高於 6-9 歲的學童（$t = 3.203, p < .01$），10-12 歲的學童之平均數高達 82.29 分，遠遠高於 6-9 歲的學童的 66.67 分，顯示此系統對 10-12 歲的學童在學習交通安全成績更好，針對此結果推測不排除年紀稍大的學童可能比年齡稍小的學童有較多交通安全的先備知識，故未來可再增加前測進行排除以驗證其差異。

表 3 不同年齡層的國小學童學習成效結果

構面	年齡	樣本數	平均數	標準差	*t*
學習成效	10-12 歲	35	82.29	16.104	3.203**
	6-9 歲	18	66.67	18.150	

**$p < .01$

4.4.不同年齡層的國小學童的臨場感結果

採用獨立樣本 *t* 檢定確認不同年齡層的國小學童對使用虛擬實境交通安全學習系統之臨場感差異，結果如表 4 所示。從結果發現不同年齡並未影響臨場感的差異（$t = 0.844, p > .05$），針對此結果的推測是可能年齡劃分較模糊且人數差異過大所導致，但不論是哪個年齡層的學童都有較高的平均數與較低的標準差，故未來可調整實驗對象年齡的劃分並增加人數以驗證其差異。

表 4 不同年齡層的國小學童臨場感結果

構面	年齡	樣本數	平均數	標準差	*t*
臨場感	10-12 歲	35	4.206	0.924	0.844
	6-9 歲	18	3.956	1.193	

4.5.不同年齡層的國小學童的暈眩感結果

採用獨立樣本 *t* 檢定確認不同年齡層的國小學童對使用虛擬實境交通安全學習系統之暈眩感差異，結果如表 5 所示。從結果發現不同年齡並未影響暈眩感的差異（$t = -1.172, p > .05$），針對此結果的推測是可能年齡劃分較模糊且人數差異過大所導致，僅有 6-9 歲學童的平均數 2.511，可能是不同年紀感受的體驗時長與接受度有所差異，也可能是不同年齡的實驗對象人數有所差距導致差距不明顯，故未來可調整實驗對象年齡的劃分並增加人數以驗證其差異。

表 5 不同年齡層的國小學童暈眩感結果

構面	年齡	樣本數	平均數	標準差	*t*
暈眩感	10-12 歲	35	2.440	1.507	-1.172
	6-9 歲	18	2.511	1.252	

4.6.小學學童的年齡對學習滿意度、學習成效、臨場感及暈眩感的關係結果

針對年齡對學習滿意度、學習成效、臨場感及暈眩感進行皮爾森積差相關分析，結果如表 6 所示。分析結果指出年齡對學習滿意度（$r = 0.296, p<.05$）和學習成效有顯著正相關（$r = 0.398, p<.01$），且學生對臨場感的感受會正面影響學習滿意度（$r = 0.282, p < .05$）。根據

上述分析結果推斷，年齡對臨場感與暈眩感的相關性不高，且發現暈眩感對年齡、學習滿意度、學習成效、臨場感皆無任何影響，由此推斷暈眩感對使用此系統的學習效果毫無關聯。

表 6 國小學童的年齡對學習滿意度、學習成效、臨場感及暈眩感之關係結果

相關性	年齡	學習滿意度	臨場感	暈眩感	學習成效
年齡	1				
學習滿意度	0.296*	1			
臨場感	0.170	0.282*	1		
暈眩感	-0.014	-0.150	0.223	1	
學習成效	0.398**	0.232	0.105	-0.173	1

**$p < .01$; *$p < .05$

5.結論與建議

根據研究結果得知，10-12 歲學童在使用本研究系統為國小學童所設計的虛擬實境交通安全學習系統後的學習成效與學習滿意度顯著高於 6-9 歲學童，與 Ou 等人（2021）、Cao 等人（2024）及 Alazmi 與 Alemtairy（2024）等的研究結果有些許相關，皆有極高的學習滿意度與學習成效，但在臨場感與暈眩感兩個構面並無顯著差異，而在國小學童的年齡對學習滿意度、學習成效、臨場感及暈眩感之相關結果可以發現，學童的年齡會直接且正向的影響虛擬實境教學的學習滿意度與學習成效，但臨場感與暈眩感兩個構面並不受年齡差異有所影響，這與 Cao 等人（2024）及 Chin 與 Wang（2024）的研究有所差異，本研究也發現學童對虛擬實境的臨場感會直接且正向影響學習滿意度，在 Mariscal Vivas 等人（2020）及 Chen 與 Liao（2022）的研究有類似結果，表示場景的細緻程度對使用者有一定的影響力，根據上述發現的結果，推測其結果產生有以下幾點原因，其一是為實驗對象的年齡區間較廣且人數不均，故導致整個結果不顯著，其二為進行課程規劃時，並未考量不同年齡在操作上消耗與接受的時常不同，導致實際實驗時發現時間會些許影響每位學童體驗感受，進而影響學習效果不佳。

所以應針對上述問題，本研究針對國小學童的虛擬實境交通安全學習系統之未來研究後續提出方向。未來可將實驗對象拆分成國小高年級組、中年級組及低年級組，進行實驗組與對照組的長期研究，探討不同年級在使用虛擬實境交通安全學習系統的差異為何，同時增加認知負荷或交通安全知識前測等，探討是否能降低國小學童對交通安全的認知負荷，以及提升學童對交通安全的學習成效與認知等差異。

誌謝

本研究感謝臺灣教育部「113 年虛擬實境教學應用教材開發與教學實施計畫」之經費支持，計畫編號為：113-V12。

參考文獻

交通部 (2022)。交通安全教育指引手冊。取自 https://168.motc.gov.tw/theme/teach/post/2209071540400

陳昕昊 (2024)。虛擬療癒庭園之漫遊體驗—高齡者之空間行為與健康效益評估。國立臺灣大學園藝暨景觀學系學位論文。臺北市，臺灣。

教育部（2021）。十二年國教課程綱要總綱。取自 https://www.naer.edu.tw/PageSyllabus?fid=52

Alazmi, H. S., & Alemtairy, G. M. (2024). The effects of immersive virtual reality field trips upon student academic achievement, cognitive load, and multimodal presence in a social studies educational context. *Education and Information Technologies*, 1-23.

Alsswey, A., Malak, M. Z., & El-Qirem, F. A. (2024). Effect of virtual reality on perceptions of usability, suitability, satisfaction, and self-efficacy among architecture and design university students. *Architectural Science Review*, 1–9. https://doi.org/10.1080/00038628.2024.2338212

Cao, S., Chu, J., Zhang, Z., & Liu, L. (2024). The effectiveness of VR environment on primary and secondary school students' learning performance in science courses. *Interactive Learning Environments*, 1-17. https://doi.org/10.1080/10494820.2024.2312921

Chen, H. L., & Liao, Y. C. (2022). Effects of panoramic image virtual reality on the workplace English learning performance of vocational high school students. *Journal of Educational Computing Research*, *59*(8), 1601-1622. https://doi.org/10.1177/0735633121999851

Chin, K. Y., & Wang, C. S. (2024). The effectiveness of a VR-based mobile learning system for university students to learn geological knowledge. *Interactive Learning Environments*, *32*(8), 4281-4296. https://doi.org/10.1080/10494820.2023.2196790

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. Computers & Education, 55(4), 1618-1627. https://doi.org/10.1016/j.compedu.2010.07.004

Conesa, J., Martínez, A., Mula, F., & Contero, M. (2024). Learning by doing in vr: A user-centric evaluation of lathe operation training. *Electronics*, *13*(13), 2549. https://doi.org/10.3390/electronics13132549

Cuturi, L. F., Cooney, S., Cappagli, G., Newell, F. N., & Gori, M. (2023). Primary Schoolers' Response to a Multisensory Serious Game on Cartesian Plane Coordinates in Immersive Virtual Reality. *Cyberpsychology, Behavior, and Social Networking*, *26*(8), 648-656. https://doi.org/10.1089/cyber.2022.0209

Digennaro, S., & Visocchi, A. (2024). Nurturing Body Literacy: Transforming Education in the Virtual Reality Era to Shape Children's Identities and Redefine Educator Roles. *Education Sciences, 14*(3), 267. https://doi.org/10.3390/educsci14030267

Hwang, G. J., Chang, C.C., & Chien, S.Y. (2022). A motivational model-based virtual reality approach to prompting learners' sense of presence, learning achievements, and higher-order thinking in professional safety training. *British Journal of Educational Technology*, *53*(5), 1343-1360. https://doi.org/10.1111/bjet.13196

Lonne, T. F., Karlsen, H. R., Langvik, E., & Saksvik-Lehouillier, I. (2023). The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study. *Heliyon*, *9*(6). https://doi.org/10.1016/j.heliyon.2023.e17196

Mariscal Vivas, G., Jiménez García, E., Vivas Urías, M. D., Redondo Duarte, S., & Moreno Pérez, S. (2020). Virtual reality simulation-based learning.

Ou, K. L., Chu, S. T., & Tarng, W. (2021). Development of a virtual wetland ecological system using VR 360 panoramic technology for environmental education. *Land*, *10*(8), 829. https://doi.org/10.3390/land10080829

教出好品德：運用 Zuvio 即時反饋系統輔助探討大學生感恩學習效果於多元創新教學實踐課程之研究

Cultivating Good Characters: Applying the Interactive Response System Zuvio to Explore the Effects of Gratitude Learning Among University Students in a Diverse and Innovative Teaching Practice Course

劉雅文
國立臺北商業大學體育室
* yaya0007@ntub.edu.tw

【摘要】 本研究目的在探討運用 Zuvio 即時反饋系統輔助探討大學生感恩學習效果於多元創新教學實踐課程之感恩知覺效益。研究對象為選修高齡者運動指導課程的大學生，共 58 人（男 10 人，女 48 人）。本研究以準實驗法 (quasi-experimental design)，招募自願參與課程的大專校院學生進行研究，使用共變數分析，分別用於實驗組與控制組前測與後測是否有顯著的不同。結論：運用 Zuvio 即時反饋系統中輔助大學生感恩學習，大學生對於感恩的知覺，具有正向發展的成效。建議未來課程可以運用 Zuvio 即時反饋系統來輔助課程進行，對於提升學生學習成效具有良好效益。

【關鍵字】 創新體育課、課堂互動、品格教育、感恩、正向心理學

Abstract: *The purpose of this study is to explore the application of Zuvio interactive response system to assist in exploring the gratitude learning effect of college students in the gratitude perception benefits of a divers innovative teaching practice course. The research subjects were 58 college students (10 males and 48 females) who took a course on the complete exercise guide for older adults. This study adopted a quasi-experimental design and recruited college students who voluntarily participated in the course. the analysis of covariance was used to determine whether there were significant differences between the pre-test and post-test between the experimental group and the control group. Conclusion: The application of the Zuvio interactive response system assists college students in their gratitude learning and has a positive effect on the perception of gratitude. It is suggested that future courses apply Zuvio interactive response system to assist in the course, which will be of great benefit in improving students' learning outcomes.*

Keywords: innovative physical education, classroom interaction, character education, grstitude, positive psychology

1.前言

品德教育一直為學校教育重要的一環（吳清山、林天祐，2005），培養具有良好品德的大學生，更是我們身為大學老師們的共同目標，期許我們培養的學生，能具備良善的人格與擁有良好的良好美德。品德教育可以是多元養成（林欣穎，2016），可以把品德教育融入社會研究、運動健康與生活藝術的課程中（成群豪，2020）；也可以身教與言教，因為在潛移默化之中，教導學生良好的行為與態度；境教也是很重要的環節，老師教學設計融入適當的

情境安排，讓學生可以正向發展與學習；在體育課中，更應該引導學生蘊含在運動之中的道德與品格，進而內化於其心與行為之中，帶給學生心靈富裕（陳盈淳、胡凱揚，2024）。從體育教育培養有品德的競爭力，潛移默化的融入課程，跟得上潮流變化，並實踐運動精神於生活中。

品德教育的根本應從「心」出發（林欣穎，2016），如果大學生擁有一顆溫暖善感的心，就可以發自內心感受自然萬物所帶來的美好，使得性格與思想得以提升，品德教育為教育之根本，身為大學老師應該想辦法把品德融入課程，培育出有品德的良善大學生。

感恩 (gratitude) 是一切善行美德的泉源（林貴芬，2015），也是正向心理學中，一個是很重要的概念 (Tsang, 2002)，正是我們大專院校學生需要的正面道德力量，也對大學生有良好的生活助益 (何家齊，2007；Kirschman et al., 2009; Magyar-Moe, 2009; Seligman et al., 2005)。感恩指的是一種人格特質與善意舉動，可以是從他人處收到一份禮物、協助、讚賞或好意，以及助人行為等正向經驗有關聯，人們會感到溫暖的感覺，這就是感恩或謝意 (林志哲，2012)，我們使用感恩策略 (劉雅文，2015)，在 Wood 等 (2010) 指出感恩也是能夠預測主觀感恩或是幸福感的重要因子，習慣感恩的人，越能感受到生活中的美好事物，自然容易快樂，在面對生活壓力與挫折時得以身心平衡。感恩的人會產生較高的正向情緒，進而促進幸福感 (吳相儀等，2017 ; Fredrickson, 2004; Kashdan et al., 2006) 。研究上也發現正向引導，才會讓學生變得幸福、更專注與更有創造力 (Fredrickson, 2001)。藉由感恩心的培養可建立良善的品德，不僅使國民能夠具有健全人格，亦能達到德育的目標，更能擁有更完善的人際互動，促進人類社會的美好與完善。

由於科技的進步，數位技術已滲透到我們大部分的現代生活，特別是數位化與體育活動的互動融合已有看到許多重要的發展，數位科技已經成為青少年的生活的一部分，成為他們的體育活動知識和行為產生強大的影響 (Meier et al, 2023)。學校應該接受學生的數位生活，並教導如何安全與負責任的使用科技 (Ohler, 2012)，使數位科技與課程可以結合運用，輔助課程教學，讓學生學習更可以達到學習成效。尤其教師們為設計更為優秀的教學教案，數位學習工具已被應用於教室環境中，就像 Zuvio 即時反饋系統，它對學生們傳遞了正向積極態度，且確實調解了學生與教師，以及同儕之間的互動（陳瑜芬、賴虹霖，2023），並提升學生的學習動力（林凱胤，2014）。本課程為吸引學生學習的興趣，特別運用 Zuvio 即時反饋系統來輔助課程的執行，來增加與學生的互動，因為，隨著科技的創新進步，大學老師也需要與時俱進，將數位科技運用融入課程，讓學生喜歡上課，同時也可以提升學生學習成效。

作者在課前，會先做適當的 Zuvio 互動平台問題的建置設計，並讓學生願意主動回饋、也充分地表達或澄清自己的觀點與看法，彼此分享意見，並在教學時搭配線上點名系統的應用，身為大學老師，總是需要隨時提醒同學要擁有良善的好品格，從點名 GPS 定位開始，便是品德教育培養的開始，也是品德情境的設計，為自己負責，培養誠實與自律，不會偷點不誠實，當然，教師在平時也可以了解學生的出席狀況，學生能為自己的出缺席與學習狀況有所警惕，在 Zuvio 即時反饋系統中的練習與回答，教師學期末時，作為學生總成績參考的依據（陳瑜芬、賴虹霖，2023）。

Zuvio 即時反饋系統，比較不方便的是教師若是在非有電腦的教學現場，手機的操作使用會比較不方便，無法臨時加題或修改，必須要教師在課前先操作好題目與分組，比如課程題庫的題目建置，手機的教師版介面，無法新加題目的建置，只能單選題的即時即答，比較可惜；研究者本身為體育老師，所以教學現場大部分是在沒有電腦的配備環境下教學，除非

自帶筆記型電腦；不過，也是不太方便。因此，作者都會事先於課前先做好各項課程題庫的建置，以方便課程使用，讓課程進行順利.

本研究是教育部教學實踐研究計劃，搭配「高齡者運動指導」課程，教師在課程中設計多元創新課程，也加入感恩的練習以及運用 Zuvio 即時反饋系統輔助課程，做一個完美的搭配使用，讓課程更為吸引同學的注意，也提升學生的學習興趣，同時教師可以從 Zuvio 即時了解學生的學習情形、感受與想法，課程也可以根據學生的學習情形，做一個適度的調整。

因此，本研究目的在探討運用 Zuvio 即時反饋系統輔助探討大學生感恩學習效果於多元創新教學實踐課程之感恩知覺效益。讓大學課程融入 Zuvio 即時反饋系統，讓課程也更為活潑，教師同時與時俱進跟上數位科技的潮流，培養學生良好品德，擁有感恩心，也提升學習學習興趣與學生學習效能。

2.研究方法

主要研究目的是進行實驗組與對照組,是否有所差異。

(1)受試者

本研究對象為本課程學生共 58 人，為實驗組學生共 28 人（男 7 人，女 21 人），平均年齡為 18.50(±.96)；控制組為 30 位（男生 3 人，女生 27 人），平均年齡 17.10 歲(±.30)，為一般課程學生。採用共變數分析(analysis of covariance, ANCOVA)，進行兩組實驗處理效果的比較。

(2)研究工具

以「大學生感恩量表」(林志哲、葉玉珠，2011) 共 26 題，本量表為六點量表，是以人類在遭遇逆境或面臨困頓的正負面情緒狀態的題目為發展，包括感謝他人、感念恩典、知足惜福、感謝逆境與珍視當下為感恩的整體表現得分量表；總量表的 Cronbach's α 係數為.93，具有良好的信度與效度。

(3)研究架構

高齡者運動指導	一般課程
⬇	⇩
實驗組 人數 28 人	控制組 人數 30 人
⬇	⇩
準實驗法 感恩量表	準實驗法 感恩量表
⬇	⇩
課程安排： 1.數位軟體教材融入 2.【青銀共學】北商大創意樂齡課程 3.陪伴長者人文關懷 4.完成阿公阿嬤上大學的夢想 5.感恩練習 6.反思 7.成果發表	課程安排： 1.體育課程 2.期中術科考試 3.期末術科考試 4.報告繳交
⬇	⇩
Zuvio 同儕評量與回饋	Zuvio 點名

(4)資料分析

進行單因子變異數分析，得知實驗組與控制組是否有顯著差異。

3.研究輔助工具

本研究採用 Zuvio 即時反饋系統來增進教學者與學生之間即時性的互動與討論，讓 Zuvio 成為教師課堂的好幫手，成功輔助課程，也可以立即性的蒐集學生課堂資料，了解學生上課吸收專業知識的程度，以及是否融入課程當中，Zuvio 其主要的功能有備題、互動與統計等功能。課程中老師可以先將要出的課題，先存放於 Zuvio 題庫中，等課程進行一個階段後，適時的給同學做答題目；也給予題目後，讓同學寫好後，可以拍照上傳至系統中，也方便將資料完整地保存下來以及成果分享資訊；也可以讓學生上傳資料與報告。Zuvio 即時反饋系統對於老師教學非常地有幫助，也幫助同學可以將心裡的話勇敢的在 Zuvio 寫下來。

4.研究實踐課程

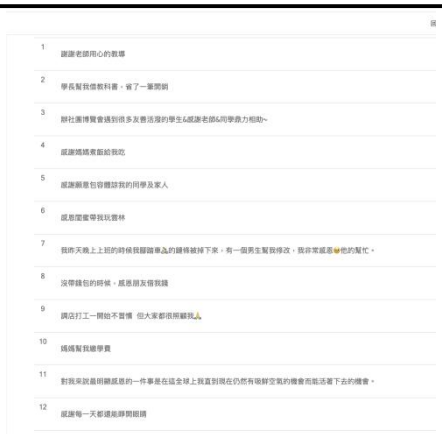
(1) 18 週的課程設計

本研究共 18 週的課程設計，搭配運用 Zuvio 即時反饋系統輔助課程，我們從第一週選修加入學生名單確認，介紹學生使用 Zuvio 系統，手機下載 APP 操作、練習回答、拍照上傳開始，讓同學開始熟悉操作運用，讓課程銜接沒有違和，一切都很進行地順暢。我們共使用了點名簽到（GPS 定位）、課前課程公告、課前課程題庫、學生私訊回饋、學生修課/分組名單設定、課前課程題庫、隨機抽點、同儕互評、統計資料下載等．如下表所示：

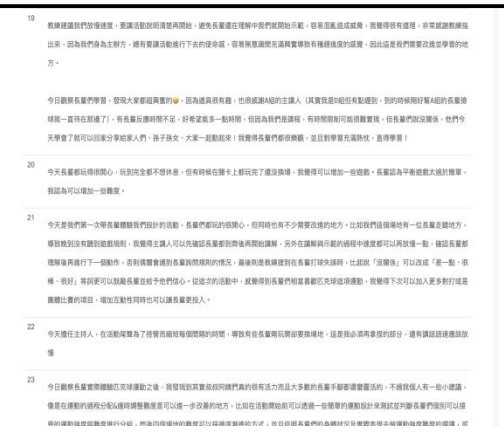
週次	內容	Zuvio 輔助
第一週	【認識課程的意義與介紹】，我們可以為社會貢獻什麼，為什麼我們要服務高齡者，與 SDGs3 健康與福祉、SDG4 優質教育以及 SDG5 性別平等之關聯	課程加入 Zuvio 即時反饋系統，介紹學習 Zuvio 使用與練習、點名簽到（GPS 定位）、課前課程公告、課前課程題庫
第二週	【認識老人久坐對健康壽命的影響、運動對老年人的好處】，讓學生更能理解為什麼老年人需要運動	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習、學生私訊回饋、學生修課/分組名單設定、課前課程題庫
第三週	【影音剪接課程規劃】，培養學生除了自我專業之外，也可以增進自己拍攝、運用與剪接的能力，培養自己的軟實力	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習、學生私訊回饋、學生修課/分組名單設定、課前課程題庫
第四週	【高齡者檢測與評估項目】認識與了解高齡者檢測項目與進行評估的方法	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習學生私訊回饋、課前課程題庫
第五週	【高齡者暖身設計與發想】要怎麼設計出一堂適合高年級的課程，結合健康與安全的設計思考之下，讓每位長	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習學生私訊回饋、課前課程題庫

	輩都適合運動的設計	
第六週	【高齡者營養與健康】如何吃出健康與美味，也是我們要注意與結合的地方	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習學生私訊回饋、課前課程題庫
第七週	【高齡者體驗】讓學生穿上高齡者體驗裝，體驗老年人的種種不方面，更培養學生的同理心	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題練習學生私訊回饋、課前課程題庫
第八週	【北商大創意樂齡活力樂齡大學_青銀共學】任務工作會議	點名簽到（GPS 定位）、課程題庫問答題練習、學生私訊回饋、課前課程題庫
第九週	【北商大創意樂齡活力樂齡大學_青銀共學_功夫小子】高齡者反應+協調	點名簽到（GPS 定位）、課程題庫問答題練習、學生私訊回饋、課前課程題庫
第十週	【北商大創意樂齡活力樂齡大學_青銀共學_遇見北極光】高齡者平衡+認知	點名簽到（GPS 定位）、課程題庫問答題練習學生私訊回饋、課前課程題庫
第十一週	【北商大創意樂齡活力樂齡大學_青銀共學_力拔山河】高齡者肌力+爆發	點名簽到（GPS 定位）、課程題庫問答題練習學生私訊回饋（拍照上傳）、課前課程題庫
第十二週	【北商大創意樂齡活力樂齡大學_青銀共學_龍的船人】高齡者協調+心肺	點名簽到（GPS 定位）、課程題庫問答題練習學生私訊回饋(拍照上傳)、課前課程題庫
第十三週	【北商大創意樂齡活力樂齡大學_青銀共學_健步齊飛】高齡者心肺+敏捷	點名簽到（GPS 定位）、課程題庫問答題練習學生私訊回饋（拍照上傳）、課前課程題庫
第十四週	【北商大創意樂齡活力樂齡大學_青銀共學_一起練舞功】高齡者柔軟+協調	點名簽到(GPS 定位)、課程題庫問答題練習學生私訊回饋（拍照上傳）、課前課程題庫
第十五週	【反思】課程的引導與設計，從服務內容、感受反應或問題挫折、學習成長、省思展望。期望學生從一連串的專業課程到場域設計參與活動，真正落實 SDGs 的永續經營理念。	點名簽到（GPS 定位）、隨機抽點、課程題庫問答題、課程公告（反思單上傳作業）、課前課程題庫
第十六週	期末成果發表，邀請專家學者進行講評與建議	點名簽到（GPS 定位）、同儕互評、課前課程題庫
第十七週	線上成果展示_Facebook 社群展示成果與心得	點名簽到（GPS 定位）、隨機抽點、課程公告(上傳作業)、統計資料下載
第十八週	期末_整體總評與回饋	點名簽到（GPS 定位）、統計資料下載

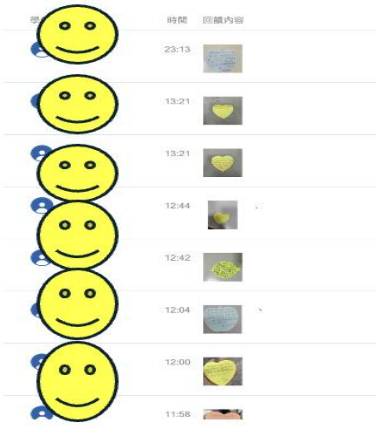
(2) 學生 Zuvio 回饋部分截圖展現



學生感恩寫作



學生反思回饋



學生感恩作品上傳

學生感恩練習

5.研究結果

問卷量化經實驗組與對照組前後測，本研究對象為本課程學生共 58 人，為實驗組學生共 28 人（男 7 人，女 21 人），平均年齡為 18.50(±.96)；控制組為 30 位（男生 3 人，女生 27 人），平均年齡 17.10 歲 (±.30)，為一般課程學生。採用共變數分析(analysis of covariance, ANCOVA)，進行兩組實驗處理效果的比較。

(1)感恩知覺分析

表 1 為感恩量表組內迴歸係數同質性考驗結果(組別*前測成績欄列之數據)，組別自變項與共變項前測成績之交互作用項考驗的 F 值為.068；顯著性考驗的 p 值=.796>.05，未達顯著水準，表示各組內的共變數與依變項的線性關係具有一致性。因此，繼續進行共變數分析。

表 1 組內迴歸係數同質性檢定摘要表

變異來源	SS	df	MS	F
組間×共變數	9.960	1	9.960	.068n.s
誤差	7938.937	54	147.017	

表 2 為運用 Zuvio 即時反饋系統在「感恩知覺」效果之共變數分析摘要表，可以知道組間效果的考驗則達顯著水準，F(1,55)= 2.379，P<.01，表示因為感恩策略融入課程，影響學生可以產生有感恩的知覺，在事後比較的結果則指出實驗組的感恩知覺的確大於控制組的同學。

表 2 運用 Zuvio 即時反饋系統在「感恩知覺」效果之共變數分析摘要表

變異來源	SS	df	MS	F	淨η^2	事後比較
共變數	16535.944	1	16535.944	114.415***	.675	I>II
組間	343.780	1	343.780	2.379	.041	
組內	7938.937	55	144.525			
全體	24828.621	57				

P<.05 ***p<.05

6.結論與建議

本研究結論為運用 Zuvio 即時反饋系統中輔助大學生感恩學習，大學生對於感恩的知覺，具有正向發展的成效。目前尚未有數位科技與品德教育的研究性論文，本研究運用 Zuvio 及時反饋系統輔助課程，將是一個創新的開始，也是一個良善的開端。品德教育是教育部一直重視的議題，也是西方教育的基礎，數位時代開啟了品格教育的新紀元，其直接目的是應對數位生活方式帶來的機會和挑戰，真正的品格教育計劃建立在社會的定義的價值觀之上，在數位科技的不斷進入到我們的社會之中，我們還需要面對的品德教育還很多，包括：網路行為、數位道德和其他令人困惑的問題 (Ohler, 2012)。身為大學老師要面對未來的核心目標之一，是培養下一代能夠在體育活動中熟練運用數位技術，同時以建設性與批判性的角度看待潛力與風險 (Meier et al, 2023)，並適時地融入品德教育，潛藏在課程中，用非正式課程的教學，讓學生習得良善的品德，應用於生活中。因此，建議未來課程可以運用 Zuvio 即時反饋系統來輔助課程進行，對於提升學生學習成效具有良好效益。

7.未來研究方向

品德教育應多元養成，尤其應該善用數位科技，未來的課程改革應該以更全面的方式解決數位化和體育問題，使學生能夠在數位領域批判性地駕馭體育運動和運動(Meier et al, 2023)，以及教師開發感恩數位教材等上傳平台，以供使用與閱讀，或是架設互動式網站，進行促進感恩對談與交流．希望善用數位科技融入大學體育課，讓學生潛移默化，學習專業與品德，培養出擁有良善好品格的優質大學生。

參考文獻

成群豪（2020）。大學生的品德教育的取向與方案。台灣教育評論月刊，9(1)，87-94。

何家齊（2007）。二十四德—孩子一生的品格計畫。台北市：旭智文化。

吳清山、林天祐（2005）。品德教育。教育資料與研究雙月刊，64，150。

吳相儀、辛昱融、陳琬云、簡晉龍、鍾昆原（2017）。感恩課程對大學生感恩、幸福感與復原力影響之研究。教育心理學報，49(1)，23-41。

林志哲（2012）。「感恩量表」之中文化修訂。教育與心理研究，35(4)，103-134。

林欣穎（2016）。孩子的品德教育，不能等！談品德教育的多元養成。中等教育，67(3)，117-119。

林貴芬（2015）。學習感恩，留住幸福的青鳥－感恩教育推行策略。學校行政，98，188-206。

陳盈淳、胡凱揚（2024）。從品德教育融入體育課程教學之論述。品質與效率期刊，2(2)，19-28。

林志哲、葉玉珠（2011）。「大學生感恩量表」之發展。測驗學刊，58，147-178。
林凱胤（2014）。即時回饋機制對學生學習專注力影響之研究。科學教育學刊，22(1)，87-107。
陳瑜芬、賴虹霖（2023）。疫情之下大學生使用即時回饋系統的現況調查。數位學習科技期刊，15(4)，25-53。
劉雅文（2015）。感恩策略融入大專校院選修舞蹈課對學生感恩知行效果之影響歷程 (未出版博士論文)。國立體育大學，桃園市。
Fredrickson, B. L. (2004). Gratitude, like other positive emotions, broadens and builds. In R. A. Emmons & M. E. McCullough (Eds.), *The Psychology of Gratitude*, 145-166.
Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218-226.
Kashdan, T. B., Uswatte, G., & Julian, T. (2006). Gratitude and hedonic and eudaimonic well-being in Vietnam war veterans. *Behaviour Research and Therapy, 44*, 177–199.
Kirschman, K. J. B., Johnson, R. J., Bender, J. A., & Roberts, M. C. (2009). *Positive psychology for children and adolescents: Development, prevention, and promotion*. In S. J. Lopez & C. R.
Magyar-Moe, J. L. (2009). *Therapist's guide to positive psychological interventions*. San Diego, CA: *Academic Press.*
Meier, S., Poweleit, A., Nigg, C. R., & Ruin, S. (2023). Talking 'bout digitality. An analysis of pe curricula in german-speaking countries. *Current Issues in Sport Science, 8*(3), 1–18.
Ohler, J. (2012). Digital citizenship means character education for the digital age. *Education Digest, 77*(8), 14–17.
Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist*, 60, 410–421.
Tsang, J. A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology, 82*(1), 112-127.
Wood, A. M., Froh, J. J., & Geraghty, A. W. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical Psychology Review, 30*(7), *890*-905.

基于代理人基模型与大语言模型人机协作生成式 AI 桌游设计与模拟：以《儒行天下：孔子周游列国》为例

Collaborative Human-AI Generative Board Game Design and Simulation Using Agent-Based Modeling and Large Language Models: A Case Study of Confucian Odyssey: The Journeys of Confucius

陈弘正[1]，林玉惠[2*]，李威[2]
[1]黄冈师范学院 机电与智能制造学院
[2]黄冈师范学院 教育学院
[*] linyuhui2023@gmail.com

【摘要】 本研究以《儒行天下：孔子周游列国》桌游为例，探索大型语言模型（LLMs）与代理人基模型（ABM）在人机协作生成式桌游设计与模拟中的应用。通过采用 Deepseek 模型与 P-T-R 迭代工作流，完成儒家文化与策略机制的融合设计。游戏构建多元胜利路径，涵盖文化传承、政治扩张与道德修养。ABM 模拟揭示了资源动态、道德抉择与玩家互动对局势演变的影响，验证了机制的平衡性与教育潜力。研究进一步通过模拟分析优化了资源系统与规则设计。结果表明，LLMs 能有效提升复杂桌游的概念生成与内容建构效率，ABM 则提供量化验证支持。本研究展示了结合生成式 AI 与系统模拟的新方法，为历史教育与游戏化学习的整合提供了实践参考。

【关键词】 大语言模型（LLMs）；Deepseek 模型；代理人基模型（ABM）；桌游设计与模拟；人机协作与游戏化学习

Abstract: *This study explores the application of Large Language Models (LLMs) and Agent-Based Modeling (ABM) in collaborative human-AI generative board game design and simulation, using Confucian Odyssey: The Journeys of Confucius as a case study. Using the Deepseek model and a P-T-R iterative workflow, the research integrates Confucian cultural elements with strategic gameplay. The board game features multiple victory paths that include cultural transmission, political expansion, and moral cultivation. ABM simulations reveal the influence of resource dynamics, moral choices, and player interactions on the evolution of the game, validating the balance and educational value of the game. Based on simulation analysis, the resource system and rules were further optimized. The results demonstrate that LLMs enhance the efficiency of concept generation and content development, while ABM provides quantitative validation. This study presents a new approach that combines generative AI and system modeling, offering practical insights into the integration of historical education with game-based learning.*

Keywords: Large Language Models (LLMs), Deepseek Model, Agent-Based Modeling (ABM), Board Game Design and Simulation, Human-AI Collaboration and Game-Based Learning

1. 前言

近年来，桌游已超越纯娱乐功能，逐渐成为教育、历史传承和文化理解的重要工具。研究显示桌游能促进多领域学习，包括数学、科学、语言及社会情感发展，并能有效提升问题解决能力。特别值得注意的是，大富翁 (Monopoly) 等经典桌游不仅能改善学习成果，还能作为文化记忆的载体。采用游戏化学习 (Game-Based Learning) 思维设计的桌游，为学生提供能主动参与的学习环境，使抽象概念具体化，并促进面对面的社交互动。

近年来，大型语言模型（LLMs）的快速发展为内容创作、教育辅助等领域带来了革命性的变革（Tinterri et al., 2024）。而 LLMs 在自然语言理解、生成与推理方面的卓越能力，亦使其在游戏设计领域展现出巨大的应用潜力。Deepseek 作为一种先进的 LLMs，具备深度思考与复杂问题求解能力，为人机协作（Human-AI Collaboration）的创新设计模式提供了技术基础。

1.1. 《大富翁》在教育中的应用与游戏化学习实例

《大富翁》在经济与财商教育中被广泛应用，帮助学生理解市场交易、资产管理和资源分配（Estep, 2020）。研究发现，该游戏能够提升学生的财务规划能力，使其体验经济决策带来的长期影响（Vijayakumar & Kulkarni, 2020）。此外，在财务报表分析教学中，学生通过《大富翁》学习现金流、投资与债务管理，研究表明该方法能有效提高学习成效（Vijayakumar & Kulkarni, 2020）。在更广泛的游戏化学习应用中，桌游已被用于 STEM、历史和商业教育。例如，“啤酒配送游戏”被用于商学院供应链管理课程，让学生体验长鞭效应（Thompson & Badizadegan, 2015）。此外，物理教育研究发现，工程建造类型桌游能显著提高学生的概念理解和课堂参与度（Hoy, 2018）。在历史课堂中，定制桌游如“Policing the Sound”被用于教授历史档案局限性，结果显示该方法提高了学生的历史分析能力（Porter, 2021）。总体而言，游戏化学习结合桌游提供了一种更具互动性和实践性的教育模式，有助于提高学习效果和学生的批判性思维能力。

1.2. 大型语言模型赋能桌游设计的人机协作

大语言模型在桌游设计与游戏化学习中的应用正逐步发展。LLMs 在游戏中可以扮演多种角色，包括内容生成者、玩家助手、叙事复述者、游戏机制设计师等。例如，GPT-3 已成功用于自动设计高质量解谜游戏关卡（Gallotta et al., 2024）。在桌面角色扮演游戏（TTRPGs）中，LLMs 可动态生成故事情节和对话，增强沉浸感（Tinterri et al., 2024）。此外，LLMs 可作为游戏规则解释助手，使桌游更易学习（Gallotta et al., 2024）。在教育应用方面，研究表明 LLMs 能辅助教师调整桌游规则，以更好地契合教学目标。盲测实验发现，ChatGPT 提供的游戏个性化建议在某些方面优于人类专家（Tinterri et al., 2024）。未来，LLMs 或可实现游戏机制的实时调整和个性化学习反馈。然而，研究者指出，LLMs 在桌游中的应用仍需关注内容可靠性、伦理风险及认知负荷（Gallotta et al., 2024）。总体而言，LLMs 正推动桌游与教育游戏的发展，为个性化学习提供创新方案。

1.3. 代理人基模型分析策略桌游模拟

代理人基模型（Agent-Based Modeling, ABM）作为模拟复杂系统行为的方法，在策略桌游设计中展现出巨大潜力。ABM 结合游戏理论，为研究个体行为如何影响集体结果提供了全面视角（Sotomayor, 2020）。策略桌游通常涉及多个玩家之间的复杂互动，ABM 通过模拟自主智能体（agents）的决策过程，能够更精准地再现玩家在不同情境下的策略演化（Lustick et al., 2021）。研究指出，桌游可以被视为小型复杂适应系统，其中玩家的选择和

互动会导致涌现行为（emergent behaviors），使游戏的策略深度更具现实感（Sotomayor, 2020）。例如，在冷战主题策略桌游 Twilight Struggle 的设计中，玩家分别扮演美国或苏联，并基于各自的目标和行动规则进行决策，这种游戏机制与 ABM 中的智能体建模原则高度相符（Lustick et al., 2021）。此外，ABM 允许设计者利用模拟数据优化游戏规则，平衡不同策略的竞争性，提高游戏的公平性与挑战性（Lustick et al., 2021）。

近年来，ABM 被用于开发和优化策略桌游的机制。例如，Virtual Strategic Analysis and Forecasting Tool (V-SAFT) 作为 ABM 平台，可模拟真实世界的政治冲突，并在游戏设计中验证不同规则的有效性（Lustick et al., 2021）。这种方法不仅适用于传统的纸质桌游，也为数位与实体桌游的融合提供了新的可能性。通过 ABM，设计者可以测试玩家在不同规则环境下的行为模式，调整游戏参数，使其更符合现实策略情境（Lustick et al., 2021）。此外，ABM 还能在桌游的平衡性测试中发挥关键作用。研究表明，通过建立基于 ABM 的虚拟桌游环境，开发者能够在原型制作前优化游戏机制，减少人工测试的时间与成本（Sotomayor, 2020）。随着桌游市场的成熟，ABM 在游戏机制研究、平衡性优化及策略模拟等方面的应用将进一步拓展，为未来的策略游戏开发提供强有力的理论支持。

1.4. 研究动机与目标

本研究的主要目标包括：

- 设计一款基于儒家文化与历史情境的桌游：以孔子周游列国的历史事件与儒家思想为主题，设计游戏背景、角色、机制与配件，力求在游戏中再现儒家文化精神与历史氛围。
- 探索 LLMs 在桌游设计中的协作模式：运用 Deepseek 模型，探索人机协作的桌游设计工作流，分析 LLMs 在游戏概念生成、机制设计、内容创作等方面的作用与价值。
- 验证 ABM 在桌游模拟与优化中的应用：构建基于 ABM 的桌游模拟模型，进行多回合的模拟与数据分析，评估游戏的平衡性、策略深度与教育意义，并基于模拟结果进行优化。
- 探讨历史策略桌游的设计原则与教育启示：基于游戏设计与模拟的过程与结果，反思历史策略桌游的设计原则，并就如何提升游戏的教育性、趣味性与历史真实性提出建议。

2. 运用 Deepseek 进行生成式桌游设计人机协作工作流

本研究核心在于探索 LLMs 在复杂策略桌游设计中的辅助作用，特别是如何通过人机协作模式，将深厚的历史文化内涵与游戏机制有效融合。吾人采用一种基于迭代问答的工作流，利用 Deepseek 模型强大的自然语言理解和生成能力，使其作为一位智能协作者，与人类设计师共同完成《儒行天下：孔子周游列国》桌游的生成与设计。

此工作流可概括为标准的 P (Prompt, 使用者提问) -> T (Thinking, AI 深度思考) -> R (Response, AI 回答) 循环。每一次循环都是对桌游设计的某一个环节进行深入探讨和细化：

- **P (使用者提问)**: 人类设计师根据当前的設計進度、遇到的问题或希望细化的方向，向 Deepseek 提出具体的、结构化的设计请求或挑战。这些问题可能涵盖游戏背景概念化、机制设计、内容填充、平衡性调整等多个方面。
- **T (Deepseek 深度思考)**: Deepseek 接收到提问后，不仅进行简单的信息匹配，而是结合其内在知识库（包括通用的游戏设计原则、历史文化知识、语言理解能力）和之前对话中积累的上下文信息，对问题进行深度分析和推理。它会尝试理解提问背后的意图、识别潜在的冲突，并构思多种可能的解决方案。这个阶段是 Deepseek 作为智能协作者的关键所在，它能够进行跨领域的知识整合和概念转换。

- **R (Deepseek 回答)**: 根据上述的深度思考，Deepseek 得以生成详细的回答，并以使用者需求提供具体的设计建议、规则草案、内容片段、分析评估等。其回答具有结构清晰、逻辑严谨，并附带必要的解释说明的特点，可帮助人类设计师理解其思路和背后的原理。

通过反复进行 P->T->R 循环，人类设计师可以引导 Deepseek 逐步完善游戏设计，从宏观概念到微观细节，不断迭代和优化方案。这种模式充分发挥了人类的创意主导和方向控制能力，以及 AI 强大的信息处理和生成能力。

关于详细的游戏设计机制、模拟结果等人机协作过程的完整细节，因本桌游尚在申请专利中，相关细节无法在本文完整披露。因此，我们将通过概念性的描述和关键工作流片段示例，清晰地展现 Deepseek 如何被用于将孔子周游列国的历史文化背景转化为富有策略性和教育意义的桌游设计。

2.1. 游戏背景与目标的概念转化

游戏的起点是将孔子周游列国的历史背景和传播儒家思想的目标，转化为桌游的基本玩法和胜利条件。Deepseek 在此阶段扮演着“概念转换”的角色，帮助把历史元素映射到游戏框架上。

以下工作流示例说明如何架构桌游背景及目标概念化。

- **P01 (使用者提问)**: “我们想基于孔子周游列国設計一个教育性桌游。核心是传播思想和‘礼乐治国’。请参考大富翁模式，提供游戏目标和如何将历史背景融入地图和行动。”（此处我们提供大富翁的专利文档、孔子周遊列國的基本信息及具教育意義的孔子語錄文档作为 Deepseek 的参考框架）
- **T01 (Deepseek 深度思考)**: “好的，需要将历史背景和游戏机制结合。大富翁的核心是地产和资源，我们可以将‘占地’转化为‘建立思想据点’，‘资源’可以是支持传播的‘文化’或‘影响力’。‘礼乐治国’是抽象目标，需要分解为游戏中的具体胜利条件。可以考虑編纂经典、影响诸侯、提升道德等作为胜利目标。地图设计可基于孔子周游路线，事件格加入历史事件和语录。需要确保教育性和游戏性的平衡。”
- **R01 (Deepseek 回答)**: “基于孔子周游列国背景和您的想法，建议游戏目标包含：1. 文化传承（如完成经典編纂）；2. 思想传播（如建立據点）；3. 道德境界（如提升道德属性）。地图設計可采用孔子周游路线的環形路径，设置不同诸侯国作为據点格，以及事件格和困境格来体现旅途中遭遇的挑战与學習。这些设计将历史与游戏机制初步连接起来。”

上述工作流片段展示了 Deepseek 如何理解使用者提出的宏观概念和参考框架，并将其转化为游戏化的具体要素，为后续的设计奠定了基础。

2.2. 游戏架构与机制的具象化

在确定了游戏的基本概念（game concept）后，Deepseek 协助人类设计师将这些概念转化为具体的游戏架构和机制细节，包括棋盘、资源、卡牌和玩家行动等。它像一位“游戏机制设计师助手”，根据历史文化内涵填充游戏内容并构建规则。

首先，Deepseek 协助设计了游戏棋盘的布局，确定了不同类型的棋盘格子及其功能，这些格子（如代表诸侯国的據点格、事件格、困境格等）都融入了孔子周游列国的历史情境。它还帮助定义了游戏中的核心资源（如反映学习和实践的资源、代表影响力的资源等）以及玩家角色的属性，确保这些元素既符合儒家文化内涵，又能支持游戏策略。

接著，在核心资源设计方面，Deepseek 极富创意的提出了代表儒家学习和实践的“六艺资源”(礼、乐、射、御、书、数) 以及代表思想传播效果的“影响力”作为游戏的基础资源。

此外，Deepseek 还设计了反映个人修养的“道德属性”(仁、礼、智、勇) 作为高级资源，影响玩家行动的成功率和策略选择。最终，Deepseek 帮助定义了这些资源如何在游戏中获取、消耗以及它们与玩家道德属性之间的互动关系。

关于角色与弟子系统，Deepseek 则为玩家提供了不同的起始能力和发展方向。Deepseek 根据历史人物特点，为玩家角色（孔子及其主要弟子）设计了差异化的初始属性和专属技能，并在弟子系统中设计了招募和技能解锁机制。这些设计都力求与历史人物的性格和事迹相契合。

最后，在卡牌系统设计中，Deepseek 协助生成并细化了不同类型的卡牌内容，将孔子语录和历史事件转化为具有游戏效果的卡牌。它帮助构思如何将这些文化元素与游戏机制（如资源得失、属性判定、玩家互动等）相结合，从而增加游戏的教育性和策略深度。

本研究针对事件卡机制具象化的工作流片段举例如下：

- **P05 (使用者提问)**: "我们需要设计一些事件卡。请根据孔子语录和相关历史事件，设计卡牌效果，例如一些能让玩家获得资源或应对挑战的卡。"
- **T05 (Deepseek 深度思考)**: "好的，事件卡是连接语录和游戏的载体。需要从儒家经典中挑选有代表性的语句，并思考如何将其含义转化为游戏效果。例如，强调学习的语句可以对应获得知识资源，強調人际关系的语句可以对应玩家间的互动。同时要考慮平衡性，卡牌效果不能过于强大或弱小。每张卡都应包含原文和簡要解释，以提升教育性。"
- **R05 (Deepseek 回答)**: "以下是一些事件卡的设计思路示例：**‘学而时习之’**：可能效果是获得某种资源。**‘三人行，必有我师’**：可能效果是与其他玩家进行资源交换或学习技能。**‘陈蔡之厄’**：作为困境卡，可能需要玩家消耗资源或通过属性判定来应对挑战。这些卡牌将语录或历史事件与游戏机制联系起来，提供不同的游戏体验。"

上述片段展示了 Deepseek 如何将抽象的文化元素（语录和历史事件）转化为具体的游戏内容（事件卡），并构思其功能，以支持遊戲目标和策略。

2.3. 游戏规则与平衡性的迭代优化

良好的游戏设计其关键在于规则的清晰性和整体的平衡性（game balance）。Deepseek 在此阶段协助人类设计师细化规则，并根据游戏长期模拟结果或逻辑分析，提出游戏平衡性調整的建议。它像一位“规则校验员”，帮助填补规则漏洞并优化游戏流程。

使用者透过向 Deepseek 提问，使其協助制定了游戏的详细规则，包括玩家如何移动、如何触发和处理不同类型的格子事件、如何使用资源、如何招募和使用弟子、以及如何进行玩家互动和赢得游戏。通过多次 P->T->R 循环，规则的表述变得更加清晰，潜在的平衡性问题（如某种策略过于强大或弱势）被识别并提出了调整方向。

针对规则细化与平衡调整的工作流片段示例如下：

- **P10 (使用者提问)**: "请详细说明游戏的规则。同时，请检查一下目前的规则設計是否完备和平衡，有没有需要补充或调整的地方？"
- **T10 (Deepseek 深度思考)**: "好的，将整合所有已設計的要素，形成一套完整的规则草案。需要涵盖游戏准备、流程、行动、资源、胜利等各个环节。在此基础上，将对规则進行逻辑检查和初步的平衡性评估。可能需要关注的点包括：不同胜利路径的达成难度是否均衡，资源获取与消耗的速度是否合理，玩家互动规则是否清晰有效，特殊卡牌和技能效果是否会破坏平衡。将列出需要进一步细化或可能需要调整平衡性的项目。"
- **R10 (Deepseek 回答)**: “以下是整合后的游戏规则说明...（此处会列出游戏规则的各个部分）... 规则设计初步评估与待优化项： 目前规则框架已初步建立，涵盖了核心玩法。

待优化项包括：部分玩家互动（如结盟）的详细规则；特定卡牌效果的平衡性（需进一步测试）；游戏时长控制；以及不同胜利条件分数权重的最终确定。这些部分需要进一步细化和调整以确保游戏体验。"

上述工作流片段展示了 Deepseek 如何帮助人类设计师梳理和完善复杂的规则，并基于对整体游戏设计的理解，主动识别出需要进一步工作或潜在的平衡性问题，引导后续的优化方向。而通过上述 P->T->R 循环的人机协作工作流，Deepseek 模型在也展现了其作为智能协作者的价值：它能够理解复杂的设计需求，将历史文化内涵转化为游戏概念和机制，并通过迭代問答协助细化规则和评估平衡性，极大地提高了桌游设计的效率和质量。

3. 运用 Deepseek 进行《儒行天下：孔子周游列国》桌游模拟及优化

为进一步验证游戏设计的合理性与平衡性，本研究运用 Deepseek 模型进行了多回合的游戏模拟与优化。首先，基于 ABM 理论框架，构建了包含四位玩家代理人（Agent）的桌游模拟模型。模型中，每位玩家 Agent 根据其角色属性与策略偏好，自主进行决策与行动，游戏环境则模拟了棋盘地图、事件卡、资源系统等元素。

3.1. 基于代理人基模型的游戏模拟模型构建

本研究基于 ABM 的理论框架，构建了《儒行天下：孔子周游列国》桌游的模拟模型。模型的核心在于代理人的设计，模型中设计了四位玩家 Agent，分别对应孔子、子路、子贡、颜回四个角色。每位 Agent 的设计均源于其历史人物特质并考虑游戏策略性，各自拥有独特的目标函数、核心属性与行为规则，如表 1。

表 1 游戏代理人设计

代理人	目标函数	核心属性	行为规则
孔子(A)	最大化文化传承	仁、礼、智	优先占据据点 → 编纂经典 → 结盟抗风险
子路(B)	最大化政治据点	勇、资源获取速度	武力扩张 → 牺牲短期资源换取据点 → 忽视道德试炼
子贡(C)	最大化影响力与资源流动	礼、外交效率	交易结盟 → 跳跃占点 → 短期利益优先
颜回(D)	最大化道德属性总和	仁、智、困境豁免	避开冲突 → 专注试炼 → 利用规则漏洞

游戏环境则被构建为包含棋盘、资源、事件和规则引擎的 Agent 环境。棋盘包含各类格子，资源包括六艺和影响力，事件系统则包含论语、道德试炼和天命挑战卡牌。这个模型旨在捕捉个体玩家策略、玩家互动与环境随机性如何共同影响游戏全局演化动态。

3.2. 多回合游戏模拟与数据分析

本研究基于 ABM 模型进行了 20 回合約三個小時的游戏模拟，并在游戏模拟结束后对游戏历程进行详细复盘与局势消长分析，从回合阶段、玩家策略、资源消长、关键转折等方面，解析四位玩家的优劣势变化与最终胜败原因（表 2）。

表 2 游戏回合阶段与局势消长总结

回合阶段	局势特征	玩家策略	资源消长	关键转折
回合 1-5	初期分散，资源争夺	玩家 A 稳健开局，玩家 B 武力扩张，玩家 C 外交灵活，玩家 D 稳扎稳打	玩家 C 资源领先，玩家 A 稳步提升，玩家 B 经济压力显现，玩家 D 资源慢但道德提升	回合 3：玩家 C 天命卡占领宋国
回合 6-10	中期分化，	玩家 A 文化路线，玩	玩家 A 资源加速积	回合 8：玩家 D 编

	策略博弈	家 B 经济崩溃，玩家 C 资源控制，玩家 D 逼近道德胜利	累，玩家 C 资源持续增长，玩家 B 濒临淘汰，玩家 D 道德属性提升	纂《礼记》，道德优势初显
回合 11-15	后期决战，胜利冲刺	玩家 A 文化冲刺，玩家 B 无力回天，玩家 C 政治扩张，玩家 D 锁定道德胜局	玩家 A 资源透支，玩家 C 资源良好，玩家 B 濒临淘汰，玩家 D 道德属性达顶峰	回合 13：玩家 D 触发道德胜利
回合 16-20	终局竞逐，胜败分析	玩家 A 文化冲刺，玩家 C 政治扩张，玩家 D 坐等胜利	玩家 A 文化分数领先，玩家 C 政治分数提升，玩家 B 淘汰，玩家 D 道德分数达顶峰	回合 20：全局结算，玩家 C 政治胜利，玩家 A 文化第二

上述游戏模拟结果揭示多个在游戏动力及文化内涵上深富意义的特点，简述如下：

1. 多元胜利路径的验证：模拟結果显示，文化传承、政治扩张和道德修养这三种不同胜利路径在模拟中都有玩家能够推进，表明多元胜利设计是可行的。
2. 策略互动体现：充足回合的模拟展现了玩家间的结盟、竞争、交易等互动行为，这些互动显著影响了玩家的资源累积和策略执行。
3. 机制动态影响：事件卡和困境格的随机性以及某些规则的特殊效果（如对规则的修改及优化）为游戏带来了变数，影响了玩家的策略应对，但也暴露了潜在的平衡性风险。
4. 历史与哲学映照：模拟中的一些关键事件和玩家决策过程，在一定程度上再现了孔子周游列国时面临的挑战（如陈蔡绝粮）或儒家思想在乱世中的实践困境。例如，道德路线的玩家如何应对现实的压力，以及文化传播的艰难性。

3.3. 基于模拟结果的游戏优化

上述基于多回合模拟结果所提供的洞见提供了游戏优化调整方向的思路。Deepseek 在此阶段协助构思和细化了规则调整方案，以提升游戏的平衡性、策略深度和教育性。例如：

1. 平衡胜利条件：限制了某些强大规则的使用次数或增加了冷却时间，并可能调整了胜利所需的门槛，以避免单一策略的垄断（对应工作流 R19, R24 中实施的规则调整）。
2. 优化资源系统：调整了资源获取和消耗的比例，并可能引入了新的资源互助或贷款机制，以提高资源流向的灵活性和玩家的生存率（对应工作流 R19, R24 中实施的资源系统调整）。
3. 调整卡牌效果：对某些事件卡的效果进行了修改，可能增加了保底机制或调整了触发条件，以平衡随机性对游戏的影响（对应工作流 R19 中实施的事件卡调整）。
4. 强化互动机制：细化了玩家互动（如结盟、交易）的规则，增加了互动收益或懲罚，鼓励玩家间的策略性互动（对应工作流 R19 中实施的互动规则调整）。

4. 结论与建议

本研究通过运用 Deepseek 大语言模型与 ABM 理论，成功设计并初步验证了一款名为《儒行天下：孔子周游列国》的教育策略桌游。研究结果凸显了人机协作模式在复杂游戏设计领域的潜力，并揭示了游戏机制如何有效结合历史文化内涵，实现寓教于乐的目标。

4.1. 研究发现与贡献

1. 人机协作新范式：本研究探索并实践了一种迭代式的 P-T-R 人机协作工作流，证明了 LLMs（特别是 Deepseek）能够作为智能协作者，在游戏概念转化、机制设计和内容生成等方面展现丰富的设计及分析能力，为历史策略桌游的设计与开发提供了新思路。
2. 教育与策略融合验证：通过 ABM 模拟，我们验证了游戏設計能够支持多元胜利路径（文化、政治、道德），这与儒家“内圣外王”的思想相呼应。模拟结果显示，游戏機制能够反映资源管理、道德选择和外部环境对玩家策略的影响，使玩家在体验孔子周游列国历程的同时，思考儒家智慧在实践中的应用。
3. 游戏动态机制洞察：ABM 模拟揭示了游戏中的一些动态现象，如特定策略（如道德专注）可能带来的优势，以及过度擴張或资源失衡导致的风险，为游戏的平衡性提供了数据支持和理论基础，也为游戏进一步优化指引新的方向。

4.2. 游戏优化与未来方向

1. 游戏平衡性微调：基于多回合模拟数据的经验，可以进一步调整不同胜利路径的分数权重，以及事件卡和机制（如据点收益、资源转换）的效果，以确保更佳的策略平衡和玩家体验。
2. 玩家互动与随机性：在不过度增加游戏玩法难度前提下，进一步细化玩家间的互动规则，并优化事件卡和规则触发的随机性设计，以增加游戏的策略深度和重玩价值。
3. 游戏设计扩展应用：未来可以探索 LLMs 在游戏其他方面的应用，如动态生成事件情节、设计具有特定行为模式的 AI 对手，进一步提升游戏的智能性和沉浸感。

4.3. 结语

本研究通过《儒行天下：孔子周游列国》教育策略桌游的设计与模拟实践，成功探索了 LLMs 与 ABM 技术在教育游戏领域的整合应用，这与当前国际上 LLMs 在游戏设计创意辅助和 ABM 在复杂系统模拟分析的趋势高度契合。本研究构建的迭代式人机协作工作流，由 Deepseek 协助人类设计师进行游戏概念转化、机制构建和规则细化，展现了 LLMs 作为智能协作者的强大潜力；而 ABM 模拟则有效地评估了游戏平衡性、策略深度以及机制可行性。

更重要的是，这种 LLMs+ABM 整合模式为历史文化主题的教育桌游提供了创新范式，能够将抽象的历史哲学内涵转化为具象的游戏机制，并通过策略博弈促进学习者对知识的理解和思考。此外，为了更进一步验证本游戏所主张的教育成效，并将模拟所展示的潜力转化为可衡量的学习成果，下一步的研究将聚焦于开展针对目标学习者的实地游戏原型测试。通过收集问卷反馈、观察学习过程以及进行认知评估等方式，得以全面验证游戏在提升玩家历史理解和儒家思想认知方面的价值，从而完善设计并强化其教育影响力，更好地实现“寓教于乐”，在全球范围内推动跨文化理解。

参考文献

Estep, S. (2020, February 12). *Students Play Monopoly to Learn Key Finance Concepts*. https://haslam.utk.edu/news/students-play-monopoly-learn-key-finance-concepts/

Gallotta, R., Todd, G., Zammit, M., Earle, S., Liapis, A., Togelius, J., & Yannakakis, G. N. (2024). Large Language Models and Games: A Survey and Roadmap. *IEEE Transactions on Games*, 1–18. https://doi.org/10.1109/TG.2024.3461510

Hoy, B. (2018). Teaching History with Custom-Built Board Games. *Simulation & Gaming, 49*(2), 115–133. https://doi.org/10.1177/1046878118763624

Lustick, I., Garces, Mi., McCauley, T., & O'Mahen, P. (2015, September 3). Approaching ABM Virtualizations as Complex Games. *The American Political Science Association Annual Meeting*, San Francisco.

Porter, T. (2021, Spring). *"Playing with the Past": Bowdoin's Rael on Teaching History with Games* [Https://www.bowdoin.edu/news/2021/12/playing-with-the-past-bowdoins-rael-on-teaching-history-with-games.html].

Sotomayor, M., Pérez-Castrillo, D., & Castiglione, F. (2020). *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models*. Springer US. https://doi.org/10.1007/978-1-0716-0368-0

Thompson, K. M., & Badizadegan, N. D. (2015). Valuing Information in Complex Systems: An Integrated Analytical Approach to Achieve Optimal Performance in the Beer Distribution Game. *IEEE Access, 3*, 2677–2686. https://doi.org/10.1109/ACCESS.2015.2505730

Tinterri, A., Pelizzari, F., Vignoli, G., Palladino, F., & Di Padova, M. (2024). Towards AI-assisted Board Game-based Learning: Assessing LLMs in Game Personalisation. *European Conference on Games Based Learning, 18*(1), 1107–1115. https://doi.org/10.34190/ecgbl.18.1.3055

Vijayakumar Bharathi S., & Kulkarni, M. S. (2020). Competition in Monopoly: Teaching-Learning Process of Financial Statement Analysis to Information Technology Management Students. *International Journal of Information and Communication Technology Education, 16*(3), 70–91. https://doi.org/10.4018/IJICTE.2020070106

導入即時回饋系統於遊戲化節奏訓練對學習者學習動機與自我效能之影響

The impact of introducing real-time feedback system into gamified rhythm training on learners' learning motivation and self-efficacy

黃暄涵[1]，夏綠荷[2*]，黃國豪[3]
[1]國立臺南藝術大學 中國音樂學系
[2]國立勤益科技大學數位學習科技系
[3]國立雲林科技大學 智慧數據科學研究所
[*]shaer.holder0130@gmail.com

【摘要】 本研究採用真實驗研究法，實驗對象為臺灣中部某音樂教室樂器學習者。實驗組採用節奏訓練師APP進行遊戲化教學，控制組則採用節拍器遊戲化教學。結果顯示，該系統與傳統節拍器均無法顯著提升學生的學習動機與自我效能。但由訪談結果可知，相較於節拍器，該系統較能維持學生的專注力，減少練習時的疲勞感。而節拍器則較適合用於基礎節奏穩定訓練。顯見即時回饋技術在音樂教育中的應用，具有可行性與效益，但仍需根據學生需求進行調整與補充。
【關鍵字】即時回饋；遊戲化；節奏訓練；學習動機；自我效能

Abstract: *This project uses the true experimental research method, the experimental subjects are students who learn musical instruments in a music classroom in central Taiwan. The experimental group used the Rhythm Trainer APP for gamified teaching, and the control group used the traditional metronome. The results show that neither the system nor the traditional metronome can significantly improve students' learning motivation and self-efficacy. However, the interview results show that compared with traditional metronomes, this system can better maintain students' concentration, and reduce fatigue during practice. The metronome is more suitable for basic rhythm stability training. Therefore, this study confirms that the application of instant feedback technology in music education is effective, but still need to be adjusted and supplemented according to student needs.*
Keywords: immediate feedback, gamification, rhythm training, learning motivation, self-efficacy

1.緒論

節奏對於我們的生活、舞蹈、音樂和體育技能都是非常重要的，尤其是在音樂教育。然而在過去的節奏訓練中，最常見的問題是使用填鴨式節奏教學法，容易使學習者失去興趣，造成練習次數減少，學習效果也不佳 (鄭方靖，2012)。因此，為了激發學生的學習動機與興趣，不少老師會將教學遊戲化，包含增加獎勵制度、獎狀或禮物，以增加學生的學習動機。

另外，大部分的教師會採用節拍器對學生進行節奏訓練，因為節拍器的穩定性，有助於學生穩定其節奏。然而，它無法給予學習者即時性的回饋。當學習者獨自訓練節奏時，無法自我糾正，進而導致學習效率及自我效能降低。面對這個問題，許多學者建議利用即時回饋系統，來幫助學生在自主學習時，修正自己的錯誤，提升學習成效及自我效能。例如，陳育慈 (2022) 的研究顯示，透過行動科技的立即回饋機制，能有效提升音樂學習者的學習成效與內在動機。另外，郭志安與吳昭儀 (2020) 在探討使用線上即時反饋系統對學生的專注力與學習成效之影響時發現，學生的專注力有顯著的提升。顯見，即時回饋對學生學習的幫助性極

高，且能提升學生的熱忱。然而這些研究，都是應用在大班教學，且非音樂個別課程的學習。音樂個別課程的節奏訓練，是否會因為導入即時回饋系統而有所幫助，仍是一個值得探討的研究議題。因此，本研究將應用一款即時回饋節奏訓練系統，於遊戲化節奏訓練中。利用其即時回饋的特性，提供學生在家自主練習時，能檢視自己的練習過程。研究問題如下：

(1) 使用即時回饋節奏訓練系統於遊戲化節奏訓練中，是否可以顯著提升學生的學習動機及自我效能？

(2) 使用即時回饋節奏訓練系統與傳統節拍器於遊戲化節奏訓練中，其優缺點分別為何？

2.文獻回顧與探討

2.1 遊戲化節奏教學之困境

許多學者指出，在教導學生訓練節奏時存在著一些困境。例如：在訓練節奏或學習音樂的過程，缺乏有趣的課程規劃，只是不停地重複訓練相同內容 (林佩儒與柯志欣，2009)。這樣的教育模式較無法提升學習者對於音樂學習，或節奏訓練的熱忱及興致。為了改善學習者在反覆練習的過程中失去興趣、樂趣及動機，不少教學者推薦採用遊戲化教學模式。遊戲化是指將遊戲的元素，如點數、排行榜或等級關卡等，應用於非遊戲的各種教學活動中，以提高參與者投入活動的熱忱與效率 (Dahalan et al., 2024)。以藉此增加吸引力，激勵學習者持續投入學習。然而，在學習的過程中，回饋的獲得也十分重要。缺乏回饋往往會造成學習的挫折，降低自我效能，進而影響學習動機和學習成效 (Sembey et al., 2024)。因此，如果能搭配即時回饋系統，於遊戲化的學習環境當中，將會是更好的組合。

2.2 即時回饋增強學習之相關研究

即時回饋是指在學習的過程中，提供學生即時回饋以幫助其得知自己的學習狀況。許多研究發現，即時回饋對於增強學習者學習是有幫助的。例如，黃建翔 (2017) 指出，讓學習者立即收到回饋，可提升學習者在課堂上的參與感和互動性，加深學習者對於學習內容的了解程度，有強化學習興趣及投入程度的效益。顯見，在學生學習的過程中，提供即時回饋的適用性，是被接受且予以肯定的。因此，本研究將在遊戲化的音樂節奏訓練課程中，導入即時回饋系統。讓學習者在自主進行節奏練習時，能得到立即的回饋，以了解自己節奏的錯誤之處，並立即修正直到正確為止。

3.學習系統介紹

本研究所採用的即時回饋系統，為節奏訓練師 (Rhythm Trainer) APP。此系統是由 Demax & Guitar Tabs 所開發。它的運作原理是由系統播放節奏，學習者聆聽後必須在手機或平板的螢幕上，正確地敲出節奏。此系統可安裝在 iOS 和 Android 平台上使用。從最基礎的節拍開始，但隨著學習的深入，學習者會學習到越來越複雜的節奏，為演奏最複雜的音樂做好準備 (Clark, 2023) 。

此系統針對節奏訓練，主要分成以下四個項目：培養節奏感、視讀節奏記譜、根據音符準確地演奏節奏及用耳朵聽節奏中的錯誤。在使用者進入系統後，系統會先測試使用者的節奏程度，由最簡單的四分音符為一拍，一小節有四拍為第 1 個關卡，並慢慢增加節奏複雜程度。如圖 1，系統將難度分成不同的關卡。若使用者通過該關卡右方會顯示打勾，正在進行的關卡會顯示三角形符號，而未通過的關卡則顯示鎖頭符號，需等到使用者通關後，方能開啟後續關卡。

圖 1 學生學習過程

4. 研究方法

4.1 研究對象

研究對象為臺灣中部兩間音樂教室，共有 24 位樂器學習者。年齡介於 15 至 22 歲，學習樂器時間皆為 3 年以上，並將其分為實驗組 12 位及控制組 12 位。

4.2 研究工具

學習動機及自我效能參考自 Pintrich 等 (1991) 所提出的問卷。包括內在動機 3 題、外在動機 3 題，及自我效能 8 題。問卷均採用五點量表，即非常同意 (5 分) 至非常不同意 (1 分)。內在動機其中一題為「在節奏訓練課程中，我比較喜歡多元節奏的單元，即使困難也無所謂。」。外在動機其中一題為「在節奏訓練課程中得到好成績，對我來說是最滿足的事情。」。自我效能其中一題為「我相信我可以在節奏訓練課程中得到優異的成績。」所有構面的 Cronbach's Alpha 值介於 0.71 至 0.97 之間

4.3 實驗流程

本研究進行為期 4 週的教學實驗，第 1 週進行前測問卷調查 10 分鐘，接著進行節奏訓練第 1 關至第 5 關的教學 20 分鐘，並完成這 5 關的檢測動作。其中實驗組的學習者，由教師指導學習者使用即時回饋節奏訓練系統，並要求學習者回家每天自主練習至少 20 分鐘。控制組則由教師提供和實驗組一樣的節奏訓練樂譜，同樣要求學習者每天使用節拍器及紙本樂譜，自主練習至少 20 分鐘，並進行自主練習紀錄。接續的每週，教師會於課堂上檢測節奏練習成果 20 分鐘，以驗收課後自主練習成果。於第 4 週時，進行後測問卷調查及訪談各 10 分鐘。

5. 結果與討論

5.1 學習動機及自我效能

因實驗樣本數偏小，採無母數統計分析法的 Wilcoxon 符號等級檢定進行分析。兩組間的學習動機及自我效能並無顯著差異 ($p > .05$)，也就是兩種方法對學生學習節奏的學習動機，及自我效能影響都不大。

5.2 訪談結果分析

訪談結果顯示，即時回饋節奏訓練系統的主要優點包括：能幫助學生發現自己的節奏錯誤，且能讓學生更專注於練習。其缺點則為：需要一段時間適應系統的操作、系統偶爾會發生不穩定或延遲的問題。傳統節拍器的優點在於：能有效幫助學生維持穩定的節奏感、可以讓學生清楚判斷自己的拍子是否正確。其缺點包括：長時間練習後容易感到疲乏、對於有速度變化的曲子較難適用、部分學生認為節拍器聲音吵雜。

綜合上述，即時回饋節奏訓練系統的優勢在於能提供個人化的即時回饋，幫助學生發現錯誤、專注於練習，但需要一定的適應時間；而傳統節拍器則能穩定節奏感，但有容易疲乏

及並非所有曲子皆適用等問題。這顯示兩者各有優缺點，教師與學生可依需求選擇最適合的訓練方式。

6. 結論與建議

本研究結果顯示，即時回饋系統與傳統節拍器，均無法顯著提升學生的學習動機與自我效能。由訪談結果可知，相較於傳統節拍器，即時回饋系統較能維持學生的專注力，減少練習時的疲乏感。此研究結果證實提供即時回饋資訊，可提升學生持續學習的時間與成果 (Liu & Liao, 2025)。然而，該系統仍需適應期，且偶爾會發生技術問題，而傳統節拍器則較適合用於基礎節奏穩定訓練。因此，本研究證實即時回饋技術在音樂教育中的應用，具有可行性與效益，但仍需根據學生的需求進行調整與補充。

建議教師在教學過程中，靈活運用即時回饋節奏訓練系統與傳統節拍器，以發揮兩者的優勢。讓學生在穩定節奏的同時，也能透過即時回饋了解自己的錯誤，更專注地投入於練習。此外，未來研究可針對不同年齡層與音樂類型，進一步探討即時回饋技術對節奏學習的影響，以驗證其適用範圍與成效。同時，系統開發者應優化系統的穩定性，並增加更多個人化訓練模式，以滿足不同程度學生的需求。

誌謝

本研究感謝台灣國科會大專計畫的支持，計畫編號為：NSTC 113-2813-C-167-002-E 和 NSTC 113-2410-H-224-007-MY2。

參考文獻

林佩儒、柯志欣 (2009)。線上音樂遊戲對音樂學習與節奏感提升成效之研究。屏東教育大學學報-教育類，(33)，37-67。

郭志安、吳昭儀 (2020)。Kahoot!線上即時反饋系統對學生的專注力與學習成效之影響－以綜合高中經濟學課程為例。臺中教育大學學報：數理科技類，*34* (1)，21-37。

黃建翔 (2017)。淺談 IRS 即時反饋系統運用至大學課程教學之策略。臺灣教育評論月刊，6 (10)，81-87。

陳育慈（2022）。行動科技輔助動態評量教學對國小學童音樂節奏學習成效之影響。國立臺中教育大學碩士論文，臺中市。

鄭方靖 (2012)。當代五大音樂教學法。高雄市：復文圖書。

Clark B. (2023, March 10). 8 Great Rhythm Training Apps and Websites. Retrieved from https://www.musicianwave.com/rhythm-training-apps-websites/

Dahalan, F., Alias, N., & Shaharom, M. S. N. (2024). Gamification and game based learning for vocational education and training: A systematic literature review. *Education and Information Technologies, 29*(2), 1279-1317.

Liu, B., & Liao, Y. (2025). Integrating IBM Watson BEAT generative AI software into flute music learning: the impact of advanced AI tools on students' learning strategies. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13394-y

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.

Sembey, R., Hoda, R., & Grundy, J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software,* 111988.

支持個人化學習的沉浸式 VR 學習環境：四季古代詩詞追尋之旅

An Immersive VR Learning Environment for Personalized Learning:

A Journey to Explore Ancient Poetry through the Four Seasons

林雨辰[1]，黃子珈[2]，陳政煥[3*]，陳妤宣[4]，陳麗琪[4]
[1]清華大學 計算與建模科學研究所
[2]銘傳大學 人工智慧應用學系
[3]清華大學 學習科學與科技研究所
[4]亞洲大學 行動商務與多媒體應用學系
[*] PolarChen@mx.nthu.edu.tw

【摘要】 本研究設計以中國古典詩詞與二十四節氣為主題的沉浸式虛擬實境（immersive virtual reality）學習平台，運用 CoSpaces Edu 開發個人化的線上學習環境，提升學生在國文學科的學習動機與理解能力。平台結合自我調節學習（self-regulated learning）理論和遊戲式學習（game-based learning）的設計理念，透過鷹架提示與互動關卡，協助學生於探索中進行知識建構與反思。

【關鍵字】 沉浸式 VR；中國古典詩詞；個人化學習；自我調節學習；遊戲式學習

Abstract: *This study designed an immersive virtual reality learning platform centered on classical Chinese poetry and the twenty-four solar terms. Utilizing CoSpaces Edu, a personalized online learning environment was developed to enhance students' motivation and comprehension in Chinese language learning. The platform integrates self-regulated learning theory and game-based learning design principles, incorporating scaffolded prompts and interactive tasks to support knowledge construction and reflection through exploration.*

Keywords: immersive VR, classical Chinese poetry, personalized learning, self-regulated learning, game-based learning

1.前言

隨著數位科技的發展，虛擬實境（virtual reality, VR）與擴增實境等技術逐漸應用於教學現場，學生的學習環境不再侷限於傳統課堂教學。傳統課堂中，教師面對面授課往往缺乏互動性，難以引起學生學習興趣，且許多學科知識難以透過口述或黑板教學明確講授，造成學生理解困難。因此，線上學習平台崛起，透過圖像、影像、音訊或 3D 真實場景模擬，幫助學生有效理解與內化知識。若進一步結合遊戲式學習（game-based learning, GBL），有助於激發學生主動學習，將遊戲元素融合於 VR 技術中，可以有效提高學生的興趣與參與度（Lampropoulos & Kinshuk, 2024），讓學生在虛擬世界中自由探索環境，提升教學的趣味性與互動性，創造全新的學習體驗。

在個人化學習中，自我調節學習（self-regulated learning, SRL）至關重要，涵蓋「計劃」、「表現」、「反思」三大階段。一、計劃：針對學習內容與任務，搭配自我學習動機，訂定學習目標；二、表現：監控自身學習行為與過程，不斷思考、改善學習策略；三、反思：

評估自身學習狀況與結果，進一步改善學習成效。從初步的規劃，到朝目標執行，不斷的反思、評估學習效果，改變學習策略，學生在學習過程中重複經歷此三階段的循環，可有效提升自我調節能力，進而獨立建構知識與整合。而結合自我調節（SR）與虛擬實境（VR）技術的 SRVR 學習環境，應有助於提高學生參與度、SRL 能力及認知程度（Wang et al., 2025），因此，本研究欲以自我調節學習理論為基礎，設計一套沉浸式 VR 學習環境，透過古代詩詞與二十四節氣為教學主題，期望提高學生自我調節能力、知識建構與轉移以及整體的認知學習成效。

2.環境設計

2.1.開發環境選擇與教學需求分析

本研究使用 CoSpaces Edu（近期已更名為 Delightex）做為開發平台，該平台支援使用者自由開發 VR 學習環境，提供多人協作、物件編輯、程式撰寫等功能。使用者可操作內建素材，也可透過編寫程式碼設計角色與場景物件間的互動。除了使用電腦或筆電進入場景外，也能透過頭戴式裝置（head-mounted display, HMD）進行沉浸式體驗。多數研究認為，藉由頭戴式裝置實現的沉浸式 VR 相較桌上型 VR 帶來更高的學習效果（Matovu et al., 2023）。因此，本研究將聚焦在沉浸式 VR 的環境設計，期望藉由其真實感與臨場感提升學生的專注度、參與度及學習動機。

VR 技術已被運用於不同學科領域進行實際教學與探討，多數應用於自然科學領域（Liu et al., 2020），也有應用於小學生外語學習之案例（Chang et al., 2024）。上述研究顯示，VR 技術在不同學科中，一定程度的提升學生的學習成績、學習動機與興趣。然而，不同學科具有不同的教學目標與需求，VR 的應用方式與學習成效亦有所差異。在中小學的教學科目中，國文為主要學科之一，其教材內容包含大量書籍、詩詞、散文的理解與分析，相較其他學科，國文教材有較多艱深難懂的文字，且文意抽象，若缺乏具體情境引導，易導致學生無法有效吸收，產生學習焦慮。為了探討是否能藉由 VR 學習環境改善教學方式與提高學習樂趣，本研究以中學國文詩詞作為主題，結合二十四節氣，打造融合國學常識、詩詞文化與二十四節氣判讀的學習環境，讓學生能夠真正「看見」詩詞中的意境、「感受」節氣的變化。

2.2.場景規劃與教學設計

本研究以「春、夏、秋、冬」四季作為主軸進行場景規劃（如圖 1），分別以「鄉間書院」、「田園農莊」、「軍兵訓練營」、「雪景村落」為主題進行設計，結合四季風格的主題背景，搭配對應季節的詩詞作品及詩人介紹，學生透過第一人稱視角進行主題活動，感受每個季節獨特的自然景觀與風貌，逐步探索詩詞中描繪的景象，也藉此了解詩人的生平背景與背後的創作情境，發揮沉浸式 VR 的特性，讓學習者跨越時間與空間的限制，彷彿穿越至古代文人所處的環境，真實感受詩人藉由詩詞欲傳遞的情感與涵義，加深對詩詞內涵的理解與共鳴。

在線上教學平台中，提示設計是建構學習環境的關鍵。虛擬環境對於學生而言屬於新穎的學習體驗，若缺乏適當的引導，容易使學生對操作流程感到困惑，影響學習成效。本研究場景規劃目標為降低學生的認知負荷，增加學習的樂趣與流暢性，專注於虛擬世界中的學習內容，另一方面融合自我調節學習的三個階段設計，希望能提高學生自我調節能力與後設認知程度，三階段說明如下。一、計劃：在每個關卡開始皆設置字卡提示（如圖 2），詳細說明學生所需尋找的指定物件與待回答的問題，協助學生明確掌握學習目標並規畫行動步驟，且僅在作答完畢後，系統才會出現指定人物，探索場景尋找特定角色才能前往下一關，避免

因操作不熟悉而錯誤點擊或迷失方向，產生不必要的焦慮與干擾；二、表現：學習者可自由選擇「春、夏、秋、冬」四大場景的體驗順序，認識對應不同季節的詩詞與詩人介紹，以及學習節氣判讀，在每個關卡的問答設置環節，皆明確指示學生問題完成次數，讓學生清楚得知還有多少指定物件與問題尚未完成，幫助學生掌握學習進度，達到完整的知識建立；三、反思：每個場景皆設置五個指定物件與問題（如圖 3），問題的設置讓學生能夠檢驗自身是否有效理解學習內容與成功地進行知識整合，若學生在作答過程遇到困難，場景內設置了特定書本，點擊即可跳轉至知識補給站中（如圖 4）尋找問題相關線索。此設計鼓勵學生自由摸索，不僅提升學生的探索動機與自主學習能力，也有助於成功回答問題，並促進知識的理解與內化。

圖 1 選擇主題關卡，場景以四季作為主軸

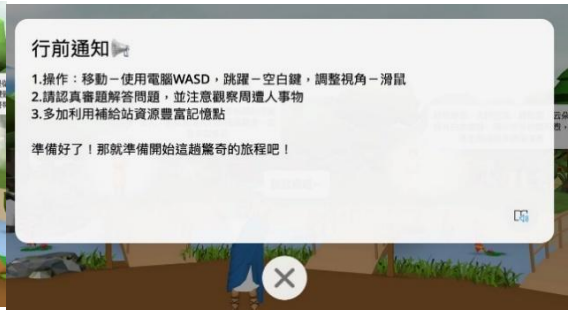


圖 2 關卡前提示字卡

圖 3 回答問題關卡環節

圖 4 知識補給站，輔助回答問題

3.預期結果

本研究針對中學生的國文教學場域，設計一套以文學詩詞為主軸的 VR 學習平台。藉由沉浸式 VR 的特色，建構出高度模擬真實場景的教學環境，突破傳統書本的限制，讓學生能夠身歷其境，進入詩詞所描繪的世界，親身體驗節氣變換所帶來的自然風貌與詩意景象。未來欲設計教學實驗，透過中學生的實際操作與體驗，得知本研究的線上學習環境是否能對學生學習表現、知識保留與轉移、認知能力和問題解決能力等帶來顯著的成效，以及對學生興趣、參與度、學習動機與認知負荷等會有多少程度的影響。希望藉由此平台引起學生的興趣，也有助於知識的理解與吸收，透過個人化的線上學習，學生可以依照自身節奏進行探索，進而促進後設認知發展與自我調節能力的培養。同時，期望得知與其他學科或領域的 VR 個人化學習，會帶來何種層面不同的效果，探討學習主題對學習成效的影響。

在 GBL 環境中，教學提示的設計是提升學習成效的關鍵因素。不論是針對操作流程的引導、平台環境的介紹或學習內容的補充，適當的提示能幫助學生迅速適應新的學習環境，

降低不必要的認知負荷，將注意力聚焦於課程內容。根據 Kao 等人的研究，在遊戲中提供教學提示的鷹架（scaffolding）可提升學生知識習得、敏感性與靈活性（Kao et al., 2017）。透過鷹架的設置，輔助學生達成自身當前難以習得的技能與能力，特別是標記關鍵特徵的鷹架，相較於提供詳細訊息的鷹架，更能顯著提高學生的學習成效。本研究參考上述理論，將提示嵌入不同學習場景中。在學生回答錯誤時，不直接給予正確答案，而是引導學生識別自身缺失之處，自行尋找線索、修正錯誤。使學生經由思考探索，逐步內化、建構自身的知識技能，增強個人化學習能力，也能將所學應用於不同的學習情境，保持持續性的自我學習能力。

根據 Zhao 等人的研究，基於自我調節的遊戲化 VR 平台有助於提升學生在英語課堂的成績、學習動機、學習態度、後設認知與自我調節能力（Zhao et al., 2024）。本研究以自我調節學習的理論框架為基礎，打造一個遊戲式的 VR 學習環境。目的在於引發學生的興趣與學習動機，促進其積極參與並主動探索。在學習過程中，學生需自行規劃學習路徑，並透過每個關卡前的提示字卡，掌握關卡流程與任務目標，進而擬定解題策略，若遇到困難或作答錯誤，也能跳轉場景尋找解題線索。透過問題檢驗，讓學生明確知道自身有哪些知識掌握不足。平台引導學生掌握「計劃、表現與反思」三階段，使學生不僅主動建構知識，更透過不斷的檢驗與反思，提升學習的有效性，進而促成更深層的知識理解與持續性的學習成長。

討論與建議

本研究以文學課程為學習主題，未來研究可延伸至科技、醫學、社會人文等其他學習領域，包含 STEAM（Science, Technology, Engineering, Arts and Mathematics）相關知識與能力，以及其他跨學科知識的建構與整合。探究在不同教學內容與教案的設計下，VR 技術或遊戲化學習對學生成績、興趣及學習動機的影響，是否在不同層次與學習面向上產生顯著差異。

可在環境中增添更多學習者與場景的互動與自由度，像是提供學生自由創作的空間，包含詩詞創作或以圖像、繪畫及空間設計等描繪詩詞內容與季節風景，讓學生能發揮虛擬環境中所學，形成更具體、富建設性的產出結果。另外，若搭配聲音的設計，在環境中融入對話情節、提示音或背景音樂等聽覺感受，使學習環境更具真實感與沉浸感，讓學生更能融入整個教學過程。

根據鷹架理論，設計學習環境應提供多元的支援與輔助，包含提示的設計，協助學生能更專注有效率的進行學習。隨著 AI 技術的發展，鷹架的建立也能結合 AI 的即時偵測功能，可對學生答題正確率、答題時間、錯誤次數、尋找線索時間、頻率等行為數據進行蒐集，更有效分析學生的學習行為，根據資料提供個人化的教學及指導，或針對學生需求給予不同的問題或提示，更快速且有效的提升學生學習成效，增進學生的自我調節，增進未來在其他學習領域，保持個人化學習、建立知識架構的能力。

參考文獻

Chang, H., Park, J., & Suh, J. (2024). Virtual reality as a pedagogical tool: An experimental study of English learner in lower elementary grades. *Education and Information Technologies, 29*(4), 4809–4842.

Kao, G. Y.-M., Chiang, C.-H., & Sun, C.-T. (2017). Customizing scaffolds for game-based learning in physics: Impacts on knowledge acquisition and game design creativity. *Computers & Education, 113*, 294–312.

Lampropoulos, G., & Kinshuk. (2024). Virtual reality and gamification in education: A systematic review. *Educational Technology Research and Development, 72*(3), 1691–1785.

Liu, R., Wang, L., Lei, J., Wang, Q., & Ren, Y. (2020). Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons. *British Journal of Educational Technology, 51*(6), 2034–2049, Article e13434.

Matovu, H., Ungu, D. A. K., Won, M., Tsai, C.-C., Treagust, D. F., Mocerino, M., & Tasker, R. (2023). Immersive virtual reality for science learning: Design, implementation, and evaluation. *Studies in Science Education, 59*(2), 205–244.

Wang, W.-S., Lee, H.-Y., Lin, C.-J., Li, P.-H., Huang, Y.-M., & Wu, T.-T. (2025). Enhancing students' learning outcomes in self-regulated virtual reality learning environment with learning aid mechanisms. *British Journal of Educational Technology, 56*(1), 366–387.

Zhao, J. H., Panjaburee, P., Hwang, G.-J., & Wongkia, W. (2024). Effects of a self-regulated-based gamified virtual reality system on students' English learning performance and affection. *Interactive Learning Environments, 32*(9), 5509–5536.

生成式 AI 輔助軟體系統架構工具之設計與開發

Design and Development of a Generative AI-Assisted Learning Software System Architecture Tool

孔崇旭，林峻劭*，陳姵予
國立臺中教育大學 資訊工程學系
* bcs113118@gm.ntcu.edu.tw

【摘要】隨著軟體系統日趨複雜，軟體工程透過標準化開發流程提升開發效率、確保軟體品質並降低維護成本，而其中需求規格的完整性與準確性對最終成品影響深遠。近年來，生成式人工智慧在軟體工程中的應用逐漸受到關注，其可輔助需求解析、加速程式碼生成，提升開發效率。然而，目前生成式人工智慧仍面臨程式碼品質參差、難以處理大型軟體架構等挑戰。因此，本研究提出一種結合程式函式、框架與自然語言需求規格的方法，透過函式級別的生成與架構優化，提升 AI 生成程式碼的準確性與穩定性，縮短開發週期，並增強開發人員在大型軟體系統中的規劃能力，以提升軟體工程的整體效能。
【關鍵字】生成式人工智慧；軟體工程；需求規格；程式框架；

Abstract: *As software systems become increasingly complex, software engineering employs standardized development processes to enhance efficiency, ensure software quality, and reduce maintenance costs. Among these, the completeness and accuracy of requirement specifications significantly impact the final product. In recent years, the application of generative artificial intelligence (AI) in software engineering has garnered growing attention, offering assistance in requirement analysis and accelerating code generation to improve development efficiency. However, current generative AI models face challenges such as inconsistent code quality and difficulties in handling large-scale software architectures. This study proposes an approach that integrates function-level code generation, software frameworks, and natural language requirement specifications. By optimizing function-level generation and architectural structuring, the proposed method enhances the accuracy and stability of AI-generated code, shortens development cycles, and strengthens developers' capabilities in planning large-scale software systems, ultimately improving the overall efficiency of software engineering.*
Keywords: Generative Artificial Intelligence, Software Engineering, Requirement Specification, Software Framework

1.前言

隨著軟體工程技術的發展，開發效率、軟體品質與長期維護成本已成為業界關注的核心議題。傳統軟體開發流程涵蓋需求分析、設計、開發與測試等階段，其中需求分析作為開發初期的關鍵步驟，不僅影響系統架構的合理性，更直接決定最終成品的品質。然而，由於需求描述的不確定性與開發過程中的人為誤差，軟體工程常面臨開發週期冗長、維護成本高昂以及需求變更所帶來的修改導致的大量時間與人力成本等挑戰。

在此背景下，生成式人工智慧（Generative Artificial Intelligence）的崛起為軟體工程領域帶來了新的可能性。生成式 AI 能夠快速生成程式碼、輔助理解軟體需求規格，並提升開

發人員的生產力，從而縮短開發週期並減少人力成本。然而，現有的生成式 AI 仍然面臨無法處理大型系統開發、程式碼品質不穩定等挑戰。因此，如何將生成式 AI 有效應用於軟體開發，提升生成程式碼的品質與可維護性，並降低開發風險，成為一項重要的研究課題。

為了解決這些問題，本研究提出一個基於程式函式、程式框架與自然語言函式規格的新穎方法，透過將程式碼生成單位細分為函式級別，並結合函式架構與明確定義的函式需求規格，來提升生成 AI 產出的穩定性與可用性。該方法不僅能夠有效防止 AI 產生幻覺（Hallucination），還能強化開發人員在大型軟體系統中的規劃能力，從而進一步縮短軟體開發週期，並提升最終產品的品質。

本研究的目標是通過結合生成式 AI 與需求規格，探索其在軟體開發領域的應用潛力，通過規劃大型軟體架構，填寫嚴謹且制式化的需求規格，最終達成自動化大型軟體系統的開發，藉此大幅縮短軟體開發流程，並維持最終成品的穩定性。

2.相關研究

2.1. 軟體系統架構的核心價值

軟體系統架構作為軟體開發的基礎框架，其核心在於如何將龐大的系統拆解為小型且可控的單元，並建立這些單元之間的協作關係。透過合理的架構設計，開發人員能夠有效降低系統的複雜性，使得整體開發與維護工作更加清晰可行。這種模組化的方法不僅提升了開發效率，也讓大型軟體專案變得更易於管理與擴展。

根據 (Bass et al., 2021)等人所述，優秀的軟體架構師應當具備以下關鍵能力：

1.設計具備高可讀性與高可維護性的架構，同時能夠解決複雜的技術挑戰

2.具備分析現有系統的能力，能夠找出影響效能、安全性等高層次問題的根本原因

3.熟悉團隊內的開發流程，並能夠清晰地向開發團隊傳達架構設計與實施要求，通常需要高度的技術專業能力

4.熟知多種開發技術，並能夠針對不同的專案需求選擇最適合的技術方案

2.2. 軟體系統架構對軟體品質的影響

(Garlan & Perry, 1995)兩人指出，軟體系統架構的核心關注點在於大型軟體系統的組織與整體結構，而非單一模組內的演算法或資料結構。優質的架構設計不僅能夠提高系統的可理解性、可重用性、可擴展性、一致性及可管理性，同時也能夠顯著降低開發時間與成本。此外，對個人而言，深入學習軟體架構能夠拓展技術視野，提升專業競爭力，對企業而言，具備架構設計能力的團隊能夠為組織帶來更高的技術價值與市場競爭優勢。

2.3. 生成式人工智慧的興起與多領域的應用

近年來，生成式人工智慧的發展突飛猛進，並廣泛滲透至人們的日常生活(Pelau et al., 2021)。這類技術的核心優勢在於模仿人類的思維模式，能夠流暢地進行自然語言互動、資訊查詢，甚至在某些情境下提供情感支持(Adiwardana et al., 2020; Arteaga et al., 2019; Dibitonto et al., 2018; Falala-Séchet et al., 2019; Schaaff et al., 2023) 。除了這些應用，生成式 AI 在程式設計領域的角色也日益顯著。它不僅可以擔任對話式助理來協助開發者解決技術問題，還能夠直接生成程式碼，為軟體開發帶來嶄新可能性(Loh, 2023; Mollick, 2022)。而伴隨著開發工具的演進，現今的整合開發環境（Integrated Development Environment, 以下簡稱 IDE）、單元測試框架以及效能分析工具已讓程式碼撰寫的門檻大幅降低，使得不論是資深開發者還是初學者，都能更輕鬆地進行程式開發(Alizadehsani et al., 2022)。然而，即便技術層面日益進步，軟體開發依舊是一項高度仰賴人力與時間的工作，如何有效減少開發成本並

提升效能仍是業界關注的重點(Bluemke & Malanowska, 2021; Butt et al., 2021; Jadhav et al., 2022)

2.4. 生成式 AI 在程式設計的優勢與限制

雖然生成式 AI 被視為提升開發效率的重要技術，它仍然存在許多限制。首先，AI 產出的程式碼並非總是正確或穩定的，其可靠性仍受限於訓練數據與演算法的能力。其次，AI 生成的程式碼在執行效率與可維護性 方面可能存在不足，這對於長期開發專案而言是一項潛在風險(Zhang et al., 2023)。因此，依賴 AI 進行輔助開發時，開發者仍需對 AI 產出的程式碼進行細緻的審查與測試，以確保其符合專案需求並達到可接受的品質標準。

根據 (Idrisov & Schlippe, 2024)兩人於 2024 年的研究結果，即便生成式 AI 在某些情況下能夠生成正確的解法，大多數情境下仍無法穩定產出高效且可維護的程式碼。這說明，AI 雖然能夠加快開發過程，但仍然無法完全取代人類的程式設計能力。即便如此，生成式 AI 作為輔助工具，仍能夠大幅加速開發週期，使開發人員能夠更迅速地建立雛形並進行後續優化。因此，未來的研究方向應聚焦於如何透過更精確的引導方式，提升 AI 生成程式碼的品質與適用性，使其在軟體開發領域發揮更大效益。

3.系統設計與展示

本研究結合生成式人工智慧，開發一個輔助軟體開發流程的系統。使用者繪製類別圖後，將產生出的程式框架上傳至系統上，並以單一個函式為單位，依據引導填寫函式的需求規格。當所有軟體系統中的函式都生成完後，使用者可以通過該系統將所有函式進行合併。如果在生成式人工智慧生成完成式碼後覺得不妥，可以重新撰寫需求規格並再次上傳，亦可在系統上直接通過程式編輯器進行修改。此外，本系統的需求規格與程式碼皆有版本控制功能，可以讓使用者推回至前任一版本。

3.1. 系統架構

系統架構設計圖如(圖 1)所示，以下進行分項說明：

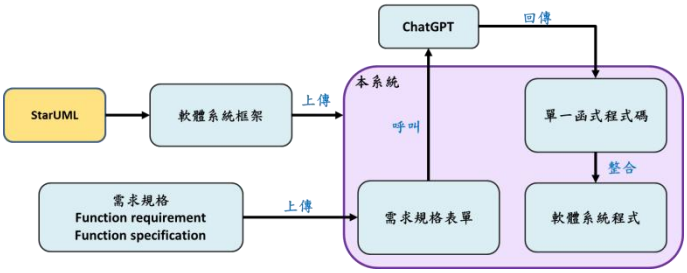


圖 1 系統架構設計圖

1.使用者通過 StarUML 繪製類別圖後，通過該軟體的插件生成無函式內容的程式框架

2.將軟體系統框架上傳於本系統

3.將對應函式的需求規格通過需求規格表單填寫至本系統

4.系統通過呼叫 ChatGPT API，透過軟體系統框架與需求規格生成單一的函式程式碼

5.通過本系統合併所有函式，最終完成軟體系統的開發

3.1.1 生成式 AI 模型版本與 API 相關設定

本系統使用 OpenAI gpt-4o 進行函式程式碼的生成，模型使用 OpenAI gpt-4o，並設定 temperature 為 0，用以減少幻覺並提升生成內容的一致性；max_tokens 設為上限值以避免截斷；其餘生成參數（如 frequency_penalty）皆設為 0，避免系統干預模型自然輸出。

為提升生成內容對程式架構與需求規格的符合度，系統在每次呼叫 API 時均設定 system 級別提示語（system prompt），明確指定模型角色為「具備專業經驗的 Java 資深開發人員」，並要求其依照使用者所提供之函式結構與需求內容進行函式本體的生成。此類提示語已透過研究團隊多輪測試與調整，能有效引導模型生成邏輯一致、可讀性高且可編譯的程式碼，亦為本系統提升生成穩定性之關鍵機制之一。

3.2. 系統展示

系統頁面圖如(圖 2)至(圖 5)所式，以下進行分項說明：

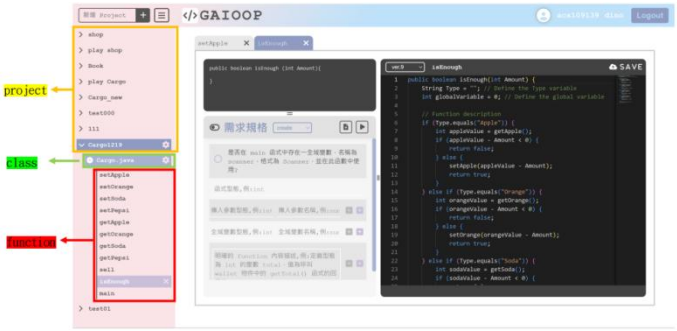


圖 2 系統首頁

(圖 2)所示為本研究所開發的系統的首頁，使用者可以在該頁面的左上角進行新專案(Project)的建立。在左側欄位中的物件(Class)選項，可以將產生完成的程式框架上傳。當上傳完成後，即可在左側欄位中看到該物件中所擁有的函式(Function)。

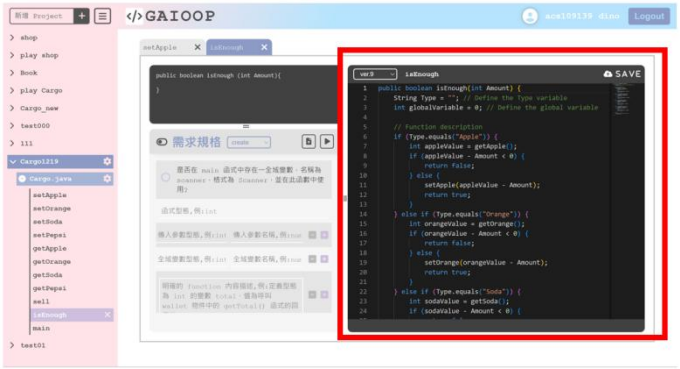


圖 3 程式碼編輯器

(圖 3)所示為本研究所開發的系統的其中一項功能，該功能容許使用者在系統上即時編輯程式碼，不論是生成式人工智慧所生成的，還是已經被使用者自行修改過的程式碼，都可以再進行修改，此外，該功能還可以顯示整個程式碼的縮圖(如紅色方框右上角)，容許使用者快速導覽單一函式中的整體程式碼，同時可以讓使用者通過滑鼠左鍵點擊快速跳轉至所點選的段落。

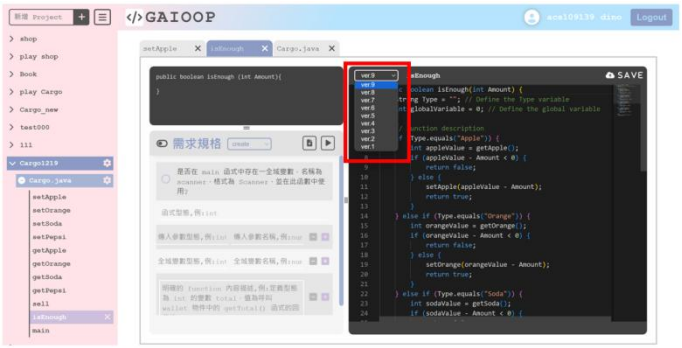


圖 4 版本控制功能

(圖 4)展示了本系統的版本控制功能，此功能建構於 MINIO 容器之上。使用者可在提交需求規格後、手動編輯程式碼後、選擇歷史版本後，或手動存檔後，將該版本的需求規格或程式碼設為最新版本，以確保版本管理的靈活性與可追溯性。

圖 5 需求規格填寫表單

(圖 5)呈現了本系統的需求規格表單，該表單對使用者的需求規格撰寫提供了明確的引導。透過標準化的填寫方式，使用者需針對 函式型態、傳入參數、同物件內的全域變數以及函式的具體內容描述進行規範化的撰寫，以確保需求規格的完整性與嚴謹性。

4.目前結果

本研究已完成輔助學習系統的功能建置，並於內部實驗環境下進行多輪測試，以確保需求規格撰寫、程式碼生成與整合流程的基本正確性與操作穩定性。為評估本系統在實際場域中的實用性與使用者接受度，本研究招募臺中某大學資訊工程學系碩士班與學士班學生共十五人，進行實測操作與問卷評估。

4.1. 系統使用成效評估

使用者測試內容包含：繪製 UML 類別圖、填寫函式需求規格、觸發 AI 生成對應函式程式碼，以及最終整合完整程式架構。統計結果顯示，多數使用者能在約兩小時內完成一套中型應用系統的開發，平均使用提示語（Prompt）次數為 50 次，遠低於使用一般網頁版 ChatGPT 需 100 次以上交互對話與 9 小時的作業時間。此結果證實本系統具備縮短開發流程與強化任務導向生成效率之潛力。

實驗中使用者共建置四種類型之應用系統：個人聯絡簿、圖書管理、工作任務管理與倉儲管理系統。系統自動生成的程式碼行數平均介於 1800 至 2300 行，內容涵蓋多個模組與物

件結構，顯示本系統具備支援中大型應用程式開發的能力，且模組化程度與語法正確性亦可接受。

4.2. 初步分析說明

為評估系統之易用性與使用者體驗，研究團隊對參與實驗的學生實施系統易用性量表的調查，其中包含 SUS(System usability scale)軟體易用性評估、5Es(Effectiveness, Efficiency, Engaging, Error tolerance, Easy to use)軟體評估與軟體淨推薦分數表(Net Promoter Score)評估，其評估結果統計如表 2、圖 6~圖 8 所示。統計結果分別說明如下：

系統易用性量表設計涵蓋學習性、效率、一致性、信心與使用意願等面向，最終經轉換得出 0~100 分的 SUS 總分，作為整體可用性評估指標。本次實驗共蒐集來自碩士班與學士班學生共 15 份有效問卷，經 SUS 計分後，其統計結果如下表 1 所示：

表 1 系統易用性量表分析

平均數（Mean）	標準差（SD）	最高分	最低分	樣本數（N）
82.0	19.21	100.0	27.5	15

根據(Bangor et al., 2009)等人所提出的 SUS 評等標準，SUS 平均得分為 82.0 時，對應之形容詞等級為「Excellent（優秀）」，其綜合評分區間約落於 80 至 90 分之間，平均為 85.5 分。此等級代表系統的整體可用性極佳，在使用者心中具備高度的正面評價，且屬於可用性測量中的上位百分群（約第 90 百分位）。

以下為 SUS 軟體易用性評估、5Es 軟體評估與軟體竟推薦分數表的詳細分析

4.2.1. SUS 軟體易用性評估(如圖 6 所示) 十題中的所有正向題的平均皆為 4.3 分，這證明了本系統對於使用者來說相當友善；「我想我需要有人幫助才能使用這個輔助學習系統。」與「我覺得這個輔助學習系統過於複雜。」在反向題中的分數較差，這顯示了本系統可能還不夠新使用者友善，在剛接觸平台時操作可能會不熟悉，需要加入更多介面的提示或操作引導，以進一步提升自主使用的流暢度。

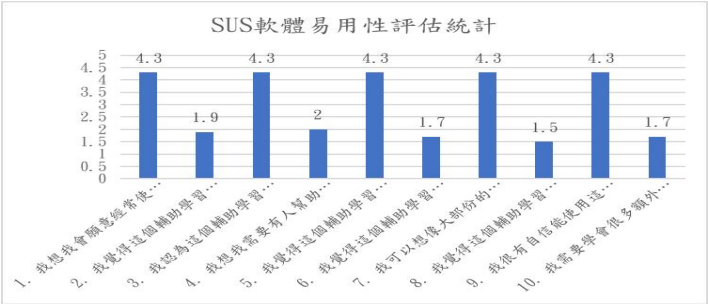


圖 6 SUS 軟體易用性評估統計

4.2.2. 5Es 軟體評估(如圖 7 所示) 本研究參考(Quesenbery, 2004)所提出的 5Es 使用性評估模型，從五個方面分析系統的可用性表現，分別為：有效性（Effectiveness）、效率性（Efficiency）、吸引力（Engaging）、容錯性（Error Tolerance）與易用性（Easy of use）。此模型強調不同使用情境下使用者需求的多樣性，並提供評估使用者互動體驗的多維觀點。

調查結果顯示，「我覺得這個輔助學習系統的畫面是簡潔的」獲得平均 4.5 分，代表系統在吸引力方面具有良好表現，使用者對其視覺設計與互動風格給予正面肯定。反之，「我無法接受輔助學習系統內出現許多錯誤」（負向題）僅獲得 2.0 分，反映出系統在容錯性方面仍有待改善，使用者在操作過程中仍可能遭遇未預期錯誤或回復不易的情況。

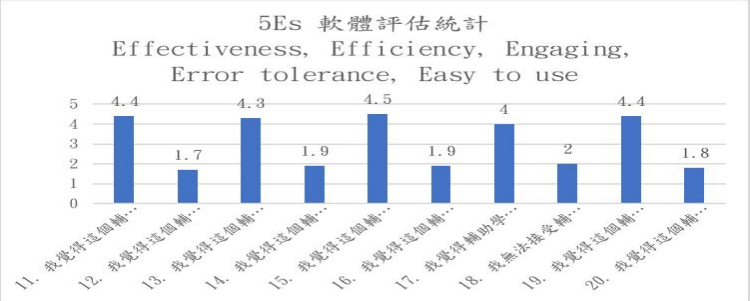


圖 7 5Es 軟體評估

4.2.3. 軟體淨推薦分數表(如圖 8 所示) 本研究依據(Reichheld, 2003)提出之 Net Promoter Score（NPS）模型，調查使用者對本系統的推薦意圖。根據 15 份有效問卷，有 7 人為推廣者（給予 9~10 分）、6 人為中立者（7~8 分）、2 人為貶抑者（0~6 分），計算得出 NPS 分數為 +33.3，顯示系統具備一定程度的使用者忠誠度與正向推薦傾向。雖未達企業級產品常見之+50 標準，但已展現良好發展潛力。

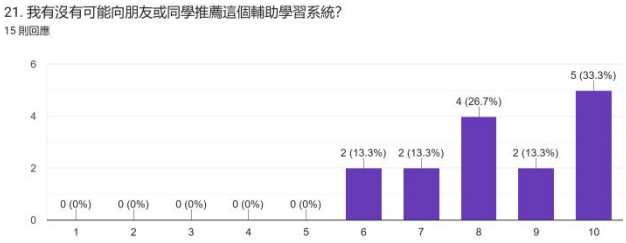


圖 8 軟體淨推薦分數表

4.3. 小結與問題反思

整體成果顯示，該系統在架構設計、生成式人工智慧整合以及任務導向操作流程等面向皆具高度成熟度，能有效協助使用者完成從需求規格撰寫到程式碼生成的完整開發流程。實測結果亦指出，其在開發效率與使用者滿意度上展現顯著優勢。

然而，儘管系統在大多數易用性指標中獲得正向評價，根據 SUS 量表中反向題項的回應結果顯示，部分初次接觸系統者在操作過程中出現輕微的不確定感。此現象可能源於使用者對需求規格填寫格式、操作邏輯或流程步驟尚未完全熟悉。

為進一步強化研究的說服力與成果的解釋能力，未來將拓展使用者研究的實驗方法，採用實驗組與對照組進行實驗，控制組將使用一般網頁版 ChatGPT 進行自然語言需求描述與程式碼生成，實驗組則使用本研究開發之系統進行相同任務操作。藉此可在開發時間、交互次數、錯誤率與程式碼品質等量化指標上進行對比分析，更精確地驗證本系統在引導結構化需求撰寫與任務導向生成流程方面的優勢。

5.未來展望

本研究已順利完成系統的基本建置，並驗證其於需求規格管理與程式碼生成方面的可行性。然而，目前系統尚缺乏自動化測試機制，需由開發者手動驗證生成的 Java 程式碼是否正確，當程式碼無法編譯或執行錯誤時，使用者需自行排除問題，這不僅降低了開發效率，也增加錯誤遺漏的風險。未來將針對自動化測試流程進行整合與優化，以提升系統的實用性與穩定性。此外，本研究也將聚焦於引入靜態程式碼分析、測試覆蓋率評估及人工智慧輔助的程式碼品質評分等技術，以作為補強生成碼可靠性之客觀佐證。

6.致謝

本研究承蒙教育部「教學實踐研究計畫」(計畫編號：PEE1137600)與國立臺中教育大學「高教深耕計畫」之經費支持與資源挹注，特此致謝。亦感謝本校師長與研究團隊在研究過程中的協助與建議，使本研究得以順利進行。

參考文獻

Adiwardana, D., Luong, M.-T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., & Lu, Y. (2020). Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*.

Alizadehsani, Z., Gomez, E. G., Ghaemi, H., González, S. R., Jordan, J., Fernández, A., & Pérez-Lancho, B. (2022). Modern integrated development environment (ides). Sustainable Smart Cities and Territories,

Arteaga, D., Arenas, J., Paz, F., Tupia, M., & Bruzza, M. (2019). Design of information system architecture for the recommendation of tourist sites in the city of Manta, Ecuador through a Chatbot. 2019 14th Iberian conference on information systems and technologies (CISTI),

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114-123.

Bass, L., Clements, P., & Kazman, R. (2021). *Software architecture in practice*. Addison-Wesley Professional.

Bluemke, I., & Malanowska, A. (2021). Software testing effort estimation and related problems: A systematic literature review. *ACM Computing Surveys (CSUR)*, *54*(3), 1-38.

Butt, S. A., Misra, S., Piñeres-Espitia, G., Ariza-Colpas, P., & Sharma, M. M. (2021). A cost estimating method for agile software development. Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part VII 21,

Dibitonto, M., Leszczynska, K., Tazzi, F., & Medaglia, C. M. (2018). Chatbot in a campus environment: design of LiSA, a virtual assistant to help students in their university life. Human-Computer Interaction. Interaction Technologies: 20th International Conference, HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part III 20,

Falala-Séchet, C., Antoine, L., Thiriez, I., & Bungener, C. (2019). Owlie: a chatbot that provides emotional support for coping with psychological difficulties. Proceedings of the 19th ACM international conference on intelligent virtual agents,

Garlan, D., & Perry, D. E. (1995). Introduction to the special issue on software architecture. *IEEE Trans. Software Eng.*, *21*(4), 269-274.

Idrisov, B., & Schlippe, T. (2024). Program code generation with generative ais. *Algorithms*, *17*(2), 62.

Jadhav, A., Kaur, M., & Akter, F. (2022). Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach. *Mathematical Problems in Engineering*, *2022*(1), 5782587.

Loh, E. (2023). ChatGPT and generative AI chatbots: challenges and opportunities for science, medicine and medical leaders. *BMJ leader*, leader-2023-000797.

Mollick, E. (2022, December 14). *ChatGPT is a tipping point for AI*. Harvard Business Review https://hbr.org/2022/12/chatgpt-is-a-tipping-point-for-ai

Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, *122*, 106855.

Quesenbery, W. (2004). Balancing the 5Es of usability. *Cutter IT Journal*, *17*(2), 4-11.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, *81*(12), 46-55.

Schaaff, K., Reinig, C., & Schlippe, T. (2023). Exploring ChatGPT's empathic abilities. 2023 11th international conference on affective computing and intelligent interaction (ACII),

Zhang, B., Liang, P., Zhou, X., Ahmad, A., & Waseem, M. (2023). Practices and challenges of using github copilot: An empirical study. *arXiv preprint arXiv:2303.08733*.

運用生成式 AI 工具輔助學習軟體系統架構之學習成效分析與探討

Analysis and discussion on learning effectiveness of using generative AI tools to assist learning software system architecture

孔崇旭，陳姵予 *，林峻劭
國立臺中教育大學 資訊工程學系
* bcs113117@gm.ntcu.edu.tw

【摘要】 本研究探討生成式人工智慧應用於物件導向軟體開發教學中的應用與學習成效，並開發一套結合生成式 AI 與統一建模語言(UML)的學習輔助平台。該平台允許學生以自然語言描述程式功能規格，並結合使用 UML 語言所繪製的軟體系統架構，自動生成所有模組功能並整合完整的軟體系統，以提升「軟體架構設計」的學習成效與系統規劃能力。研究採用實驗組與對照組設計，比較使用學習輔助平台與傳統學習方法的學生在軟體開發學習成效上的差異，並進行統計分析。結果顯示，學習輔助平台能顯著減少學生在開發過程中的時間成本，並提升中低成就學生對物件導向程式設計的理解與應用能力，相較於傳統方法更具學習成效。本研究證實生成式 AI 可作為程式設計教育的有效輔助工具，並為未來課程設計提供實證參考，進一步推動程式設計教學的創新與發展。

【關鍵字】 生成式人工智慧；輔助軟體系統開發；物件導向程式教學；統一建模語言；學習成效

Abstract: *Generative artificial intelligence has flourished in recent years, profoundly impacting various fields, including education. Despite this, research on programming education, particularly in object-oriented software system development, remains limited in effectively utilizing generative AI for instructional support. This study develops a teaching assistance platform integrating generative AI with Unified Modeling Language, enabling students to generate and integrate an entire system automatically. By combining natural language descriptions of program functions with code frameworks, this approach addresses the time-consuming nature of software development in educational settings and enhances students' ability to design complex software system architectures.*

Keywords: Generative Artificial Intelligence, Assisted Software System Development, Object-Oriented Programming Education, Unified Modeling Language, Learning Effectiveness

1. 前言

在軟體開發領域，傳統的開發流程通常包含需求分析、設計、開發、測試、上線等階段，這些過程往往需要大量人力與時間，導致開發週期冗長。此外，軟體開發的複雜性隨著系統規模的增長而提升，開發人員需要考慮多種架構設計模式、模組化開發方法以及效能優化策略，使得開發成本進一步提高。

近年來，生成式 AI 蓬勃發展，其能夠快速生成程式碼的特性，可以提高軟體開發的效率，使開發人員能夠更專注於架構設計與系統優化。儘管生成式 AI 十分便利，其在程式設計教學輔助方面仍存在一些限制，目前，生成式 AI 較難以建立較大且複雜的軟體系統，尤其在處理較複雜的操作流程（scenario）時，所生成的程式碼往往需要人工修改，且程式碼的品

質參差不齊。有時生成的程式碼語法簡單易懂，有時則過於複雜，使得初學者在沒有指導的情況下難以理解和應用。

物件導向程式設計的課程目標是培養軟體系統開發的高階人才，要能熟悉開發較大複雜的軟體系統，操作較複雜的操作流程，學生在學習物件導向程式設計時，往往面臨理解和應用複雜概念的挑戰。

如前所敘述可以得知，傳統的教學方式往往需要透過大量的理論講解與實作練習來培養學生的程式設計能力，但初學者在理解物件導向概念、設計模式以及大型系統架構時，常會因抽象程度過高而感到困難。此外，現有的學習資源雖然豐富，但缺乏能夠即時回饋並引導學生修正錯誤的工具，使得學習過程較為被動，進而影響學習成效，因此本論文開發一個運用生成式 AI 技術優點的系統，設計一個生成式 AI 物件導向軟體系統開發的輔助學習平台(Generative AI Object-Oriented System Development Platform，以下簡稱 GAIOOP)，用來協助軟體系統開發的教學，該系統能讓學生可以透過統一模型語言(Unified Modeling Language, UML) 以及自然語言以產生相對應之程式碼，輔助學生學習軟體系統開發和程式設計。

2. 相關研究

2.1. 生成式人工智慧

生成式人工智慧(Generative Artificial Intelligence, 簡稱生成式 AI)是一種強大的深度學習技術，通常是基於 RNN (Yu et al., 2019)或是(Vaswani et al., 2017)等人提出的 Transformer 等架構，訓練過程中，模型學習大量文字、圖片、影音或是其他類型的資料，從而能夠模仿人類生成文字、圖片、影音或其他多種形式的內容。這項技術已經在多個領域取得了重大突破，包括自然語言處理、圖像生成、音樂創作和程式碼生成等。目前常見的生成式 AI 模型有 GPT-4、LLaMA、Gemini、Grok 等。GPT-3(Generative Pre-trained Transformer) (Floridi & Chiriatti, 2020) 是由 OpenAI 所開發的第三代自我迴歸語言模型(Autoregressive model)，它使用深度學習來產生類似人類的文本，會將稱為提示語(prompt)的來源輸入開始產生單字、程式碼或其他資料序列，例如，它在機器翻譯中用於統計預測單字序列。2018 年 GPT 的第一次迭代使用了 1.1 億個學習參數(即神經網路在訓練過程中嘗試優化的值)，一年後，GPT-2 使用了 15 億個學習參數，如今，GPT-3 使用 1750 億個參數，它使用微軟 Azure 的 AI 超級電腦(Scott2020)進行訓練，其訓練費用非常昂貴，估計花費了 1200 萬美元。通過大量文本訓練，能製造出類似人類的文字，這有助於多種自然語言處理工作，如翻譯、摘要、問答和聊天應用。ChatGPT 能夠以對話方式理解並回應人類輸入，非常適合互動應用。此外，ChatGPT 可以根據不同資料集和任務進行微調，從而實現定制化，處理特定任務並產生更準確、更實用的輸出。LLaMA(Large Language Model Meta AI)是由 Meta AI 公司於 2023 年 2 月發布的大型語言模型集合，使用參數從 70 億到 650 億不等。LLaMA-13B 的效能超過了 GPT-3，且其體積卻小了 10 倍以上，LLaMA-65B 與其他大型語言模型相比也相當有競爭力。綜上所述，生成式人工智慧是一項強大且多用途的深度學習技術，其帶來了多領域的應用，如翻譯、問答系統和互動應用。生成式人工智慧的不斷演進和新型模型的出現，為各領域帶來更多可能性，並提供了更多元化的選擇，有助於實現更精確且實用的應用。

2.2. 生成式人工智慧在教學上的應用

目前已經有許多人在研究生成式 AI 在各種領域上的影響，人們普遍認為，從問題的自然語言描述自動合成原始碼可能會大大提高專業開發人員的生產力(Peng et al., 2023)，尤其是學生能夠輕鬆地自動生成程式碼並評估解決方案，這引起了對已有深入研究的入門程式設計

領域的關注(Becker & Quille, 2019)，像是(Baidoo-Anu & Ansah, 2023)探討了 ChatGPT 在教育領域上的影響，ChatGPT 作為一種強大的語言生成模型，具有出色的互動性，它能夠在教育環境中發揮多種作用，包括教學輔助、研究輔助以及評估等。這種技術的應用大大縮短了完成這些工作所需的時間，使教學和評估變得更加高效。然而，與其優勢相對應的是一些挑戰，ChatGPT 雖然能提供快速回答問題，但有時也可能生成錯誤答案或編造不存在的內容，這可能會對教學產生負面影響。在教育中，確保生成式人工智慧的安全並具有建設性至關重要，制定相應的指南和原則，確保這種技術的正確應用，避免可能的負面影響，這是一個迫切需要討論的問題。

(Cooper, 2023)針對 ChatGPT 在科學教育的應用上也做了一些探索性研究，在學生使用上，ChatGPT 在回答問題時通常會提供相關答案，但其可信度不足且有倫理問題，例如版權問題。在教學中，可利用 ChatGPT 生成 5Es表單以進行科學教學，但其輸出結果仍有待改善，且教師應對 AI 資源進行批判性評估，並根據情況調整，所以 AI 尚不能取代教師的專業知識。在研究中，ChatGPT 能協助作者修改句子、提升描述清晰度，但使用 AI 平台時需要更透明，也需要明確指導方針以確保科學知識的推進。

程式教育是現代教育體系中的重要組成部分，因為程式設計技能在當今數位化社會中的重要性越來越高。生成式 AI 技術的崛起為改進程式教學方式提供了新的機會和挑戰。知名的幾個基於大型語言模型的程式生成工具，包含 OpenAI 的 Codex、DeepMind 的 AlphaCode、Amazon 的 CodeWhisperer 等，都能自 6 動生成程式碼，從簡單的範例到複雜的演算法，這使得初學者能夠更容易地理解程式設計概念，並減少錯誤，同時節省了時間。(Denny et al., 2024; Sarsa et al., 2022)針對了 Codex 進行評估，並提出了利用生成式 AI 輔助學習程式設計的優點，其優點是能提供更多的練習資源，利用 Codex 生成的程式設計練習題目，約有 80% 的練習題是嶄新的，且與人工生成的練習題目相比，其數量也遠超過人工生成的，若是再加以修改，甚至能針對學生進行個性化生成，且 Codex 所生成的程式碼可以通過 CodeCombat 所制定的電腦科學一級(CS1)測驗，其成績表現在兩次考試中，得分皆是 15.7/20(78.5%)，若是使用 Codex 來輔助學生解決類似 CS1 問題的程式考試，普遍有使用 Codex 學生的成績，相較於未使用 Codex 的學生還好（(Finnie-Ansley et al., 2022)。

其他研究表明，這類工具在處理 Insertion Sort 或是 Tree Traversal 等常見的演算法問題時，其所生成的程式碼較正確有效(Dakhel et al., 2023)，也可以為資料結構和演算法等課程中的典型問題生成正確程式碼(Finnie-Ansley et al., 2023)。

2.3. 物件導向技術學習之困難

多年以來，物件導向程式設計課程的輟學率和不及格率居高不下，因此吸引許多研究者探討其成因與解決方法，根據(Tan et al., 2009)的研究，他們調查了 182 位大學生對於學習程式語言時所面臨的困難，他們在設計解決特定任務的程式、分割功能成各個程式、學習語法、找出程式錯誤等方面感到挑戰。此外，在程式語言的各種主題中，封裝、繼承和多型是他們最難以理解的部分，但這三部分恰好是 Java 中最為重要的概念。該研究也發現實作練習、小組練習和個人練習，都能有效增進學生的學習。提供範例程式和練習題目也被認為是最有幫助的學習材料。而且，導致學生成績不好的主要原因中，包含了練習範例過少和教學大綱過於重視理論知識等問題。

在教學中，評估學生的程式設計作業往往僅著眼於最終結果，而忽略了學生在解決問題過程中的發展和學習， (Hosseini et al., 2014)的研究強調了評估學生 Java 程式設計作業時應考慮到過程和結果的平衡。不僅僅關注最終成果，而是對學生在解決問題過程中的學習和進

步進行更全面的評估，這將有助於更好地指導學生的學習策略並提高其程式設計能力。最後的結論是，約有 77.63%的學生通常從簡單的程式開始，隨著概念的增加和測試案例的擴展，逐步提升程式的正確性，且能夠有效改善學習成效。

綜合以上觀點，我們了解到學生在學習物件導向程式語言時所遇到的困難，尤其是在 Java 的封裝、繼承和多型等重要概念的理解上。因此，未來的教學方法和資源提供應更加注重實作練習，提供更多的範例程式和練習題目，以幫助學生更好地理解並應用這些核心概念。通過改善教學方式，我們希望降低學生學習曲線的陡峭度，提高學生對於程式語言的理解和應用能力，從而有效提升學生的學習成效。透過對學生程式設計作業的評估的研究，我們意識到過程中的發展和學習對於程式設計能力的提升至關重要。而在我們開發的輔助學習平台中，我們特別強調了這一點。除了強調最終成果外，我們平台所提供的程式碼版本控制功能，使得學生能夠追踪其程式作業的演進過程。這使得學生能夠更好地理解和分析自己作業的每一個階段，從而更有效地提升其程式設計能力，而且通過觀察和評估學生程式碼在不同階段的變化，我們也能夠更全面地了解他們的學習路徑和程式設計思維模式，從而更有針對性地提供指導。

3. 研究目標與方法

3.1. 研究目標

本研究旨在探討生成式人工智慧（Generative AI）在輔助學習軟體系統開發中的應用，並評估其對學生學習成效的影響。具體目標如下：

3.1.1 開發生成式 AI 輔助學習平台

設計並實作一套結合生成式 AI 與統一建模語言的學習輔助平台（GAIOOP），幫助學生透過自然語言描述程式功能，並自動生成對應的軟體架構與程式碼，以提升學習效率與理解能力。

3.1.2. 比較 AI 輔助與傳統學習方法的成效

採用實驗組與對照組設計，分析使用 GAIOOP 學習的學生與採用傳統教學方法的學生，在學習成效、程式設計能力與系統開發能力上的表現差異，並透過統計分析驗證其有效性。

3.2. 研究對象

本研究以臺灣中部某大學資訊工程學系修習「物件導向程式設計」課程之學生，共 34 位，主要修習課程之學生為資訊工程學系二年級的學生。

3.3. 研究架構

本研究之研究架構圖如圖 1 所示，各項說明如下：

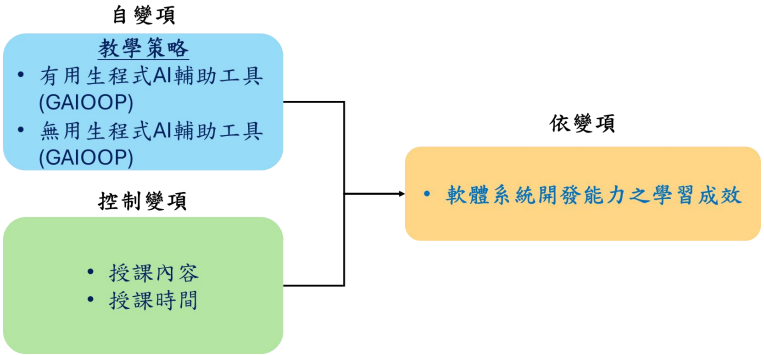


圖 1 實驗架構圖

3.3.1. 自變項 以「教學策略」為自變項，其中包含兩種策略的使用，分別為「有使用生成式 AI 輔助工具(GAIOOP)」及「無使用生成式 AI 輔助工具(GAIOOP)」。

(1)「有使用生成式 AI 輔助工具(GAIOOP)」：將學生學習時，將使用 GAIOOP 系統來輔助產生程式碼，在 GAIOOP 系統中會有結合 UML 的架構觀念及軟體規格，來產生透過 GAIOOP 系統與 ChatGPT 來溝通，並經過 GAIOOP 系統來做 Source Code 的產生來做系統的整合來產生完整的應用程式。

(2)「無使用生成式 AI 輔助工具(GAIOOP)」：將學生學習時，將自行使用繪製 UML 的架構觀念，學生依據軟體規格，自行撰寫程式碼，自行產生 Source Code，並自行來做系統整合產生完整的應用程式。

3.3.2. 依變項 研究中的依變項為「軟體系統開發能力之學習成效」，其中「軟體系統開發能力之學習成效」以開發軟體系統為主，分別使用前後測採上機實作評量，物件導向使用技術採上機實作評量。

3.3.3. 控制變項 為了減低自變項以外的其他變項對實驗效果產生影響，本研究在實驗過程中盡量掌控除了自變項之外的其餘教學條件，並使其盡可能相同。實驗進行時兩組之教學者皆由研究者親自擔任，兩組教學時間同為 16 節課，教學內容均以物件導向程式設計課程內容為主。

3.4. 研究流程

研究實施流程如圖 2 所示，各項說明如下：

1. 一般上課：包含基本物件導向能力技術的培養 7 週，期中上機考試 1 週。
2. 分組：依據教學方式分成 2 組，實驗組使用「GAIOOP 輔助教學系統」，對照組使用「一般 PPT 教學教材」。
3. 教學實驗：共 6 週，實驗組使用「GAIOOP 輔助教學系統」，對照組使用「一般 PPT 教學教材」，並實施 2 週的後測，並收集成績資料。

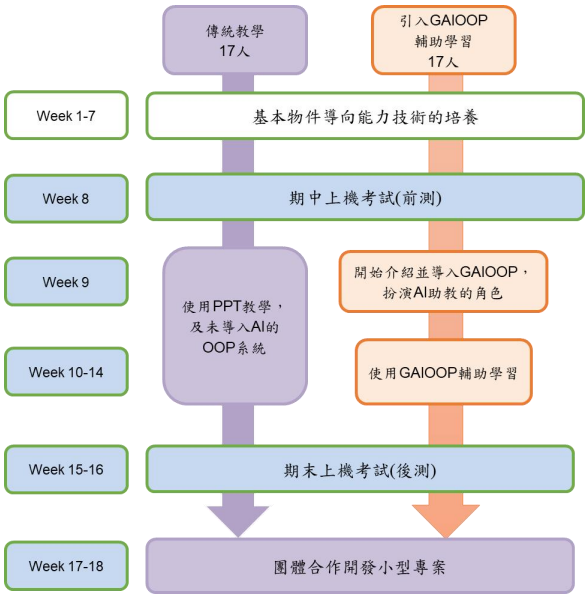


圖 2 研究實施流程圖

3.5. 研究工具與使用流程

3.5.1. GAIOOP 輔助教學平台系統功能 GAIOOP 輔助教學教學平台的系統功能有規格表單，用以引導學生填寫函式的細部需求功能規格；程式碼解析器用於學生上傳程式架構時，將多

個檔案與整串的程式架構解析承擔個函式；線上程式碼編輯器可以讓學生在線上即時修改生成的程式碼，或者式修改生成式 AI 生成的程式錯誤；版本控制可以讓學生倒回任一歷史版本的需求規格表單與程式碼；儲存容器負責分類與儲存學生上傳的程式框架、需求規格表單、完成生成後的完整軟體系統程式；Log 檔紀錄使得教師可以詳細分析每個學生與生成式 AI 的問答紀錄、使用的 token 數量與時間；自動程式碼合併功能則是負責將所有生成完成後的程式碼自動合併至該函式對應的物件。

3.5.2. 使用流程 本研究所開發的 GAIOOP 輔助教學平台使用流程如圖 3 所示，各項說明如下：

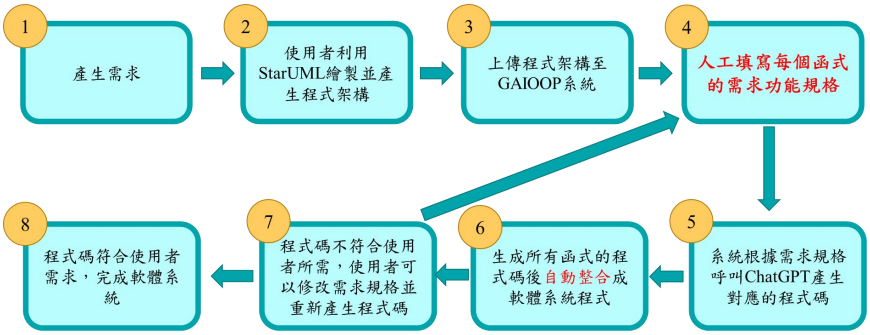


圖 3 使用流程圖

學生在產生需求後，先於 StarUML 上繪製類別圖(Class Diagram)，並通過 StarUML 中的 Java generate code 插件，將類別圖轉換成空白的程式架構。接著將程式架構上傳至 GAIOOP 輔助教學平台，並在 GAIOOP 輔助教學平台完成程式碼的分割後，以單個函示為單位依序填寫每個函式的需求功能規格，並在填寫完後送出需求規格，並等待生成式 AI 生成程式。

當所有函式的需求功能規格填寫完後，可以將各個函式自動整合成軟體系統程式。

4. 實驗結果分析

本研究旨在探討生成式 AI 輔助學習軟體系統（GAIOOP）是否能夠提升學生在軟體系統開發與程式設計方面的學習成效。為了達成此目標，我們依據學生的前測成績進行分組，並於學期末進行後測，以評估學習成效。後測採用期末上機評量的方式，包含五道難度相同的試題，學生將隨機抽取一道題目，需從無到有撰寫出一個完整的軟體系統，以進行上機實作考試。本研究採用共變異數分析（ANCOVA）進行分析來評估 GAIOOP 對學習成效的影響。以下為詳細的結果分析。

4.1. 共變異數分析（ANCOVA）

為確認 ANCOVA 分析使用前提是否成立，進行了迴歸係數同質性檢定。交互作用項（group × pretest）未達顯著水準（$p = 0.575 > 0.05$），表示實驗組與對照組之間的前測與後測之回歸斜率相近，符合迴歸係數同質性假設。因此可進一步使用 ANCOVA 比較兩組在後測成績的差異。表 1 為 ANCOVA 分析的結果。

分析顯示，在控制前測成績後的後測成績比較結果，實驗組後測平均數為 40.28，高於對照組的 24.69，調整後平均數分別為 41.37 與 25.17，顯示 GAIOOP 對學習成效有正向影響，但差異未達統計顯著性（$F = 2.606$，$p = 0.117 > 0.05$）。

表 1 ANCOVA 分析結果

變異數	組別	個數	平均數	調整後平均數	標準差	*F*	*p*
教學策略	實驗組	17	40.28	41.37	29.8	2.606	.117
	對照組	17	24.69	25.17	27.3		

因此，本研究將高分群分離，實驗組與對照組分別減去 4 個最高分的數據(17*0.75=12.75)，以此取得實驗數據中的後 75%數據，用來分析中低分組。首先進行迴歸係數同質性檢定。交互作用項（group × pretest）未達顯著水準（$p = 0.127 > 0.05$），符合迴歸係數同質性假設。因此可進一步使用 ANCOVA 比較兩組在後測成績的差異。表 2 為中低分群 ANCOVA 分析結果。

數據顯示，實驗組後測平均數為 27.4，高於對照組的 11.6，調整後平均數分別為 28.35 與 10.73，顯示 GAIOOP 對中低成就學生的學習成效具有正向影響，兩組間差異已達統計顯著（$F = 5.923$，$p = 0.0231 < 0.05$），相較於全體樣本的分析結果（$p = 0.117$），中低分組的 p 值顯著降低，顯示 GAIOOP 對中低分學生的影響更為明顯。

表 2 中低分組 ANCOVA 分析結果

變異數	組別	個數	平均數	調整後平均數	標準差	*F*	*p*
教學策略	實驗組	13	27.4	28.35	20.65	5.923	.0231
	對照組	13	11.6	10.73	14.34		

5. 結論與未來展望

本研究探討生成式人工智慧（Generative AI）在輔助學習軟體系統開發中的應用，並透過開發 GAIOOP 學習輔助平台，分析其對學生學習成效的影響。研究結果顯示，GAIOOP 在提升學生學習效率、改善物件導向程式設計開發能力方面具有正向效果。分析進一步顯示，移除高分群後，GAIOOP 對中低分程度學生的影響較為顯著，效果達顯著門檻，顯示其對中低分學生的助益較好，生成式 AI 輔助學習顯著提升了其軟體架構開發能力，相較於傳統教學方法更能降低學習門檻，幫助學生掌握軟體架構設計的核心概念。然而，分數較高的學生在使用 GAIOOP 時成效未顯著提升，可能是因為高分學生已具備較強的程式設計能力，生成的程式碼對其幫助有限，因此生成式 AI 在不同學習程度學生中的應用效果仍需進一步探討。

未來研究將聚焦於檢索增強生成（Retrieval-Augmented Generation, RAG）技術的應用，以提升 GAIOOP 生成程式碼的品質與準確性。RAG 技術可透過動態檢索相關學習資源與最佳實踐，輔助生成式 AI 在程式碼生成過程中進行自動校正，有效降低語法錯誤與邏輯缺陷，確保輸出內容的可讀性與可執行性。此外，GAIOOP 系統將新增學生使用行為紀錄功能，以支援後續的學習行為分析，也會加入系統之易用性量表、學習動機量表及學習態度量表進一步評估生成式 AI 輔助學習的效果。

6.誌謝

本研究承蒙教育部「教學實踐研究計畫」(計畫編號：PEE1137600)與國立臺中教育大學「高教深耕計畫」之經費支持與資源挹注，特此致謝。亦感謝本校師長與研究團隊在研究過程中的協助與建議，使本研究得以順利進行。

引用文獻

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52-62.

Becker, B. A., & Quille, K. (2019). 50 years of cs1 at sigcse: A review of the evolution of introductory programming education research. Proceedings of the 50th acm technical symposium on computer science education,

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of science education and technology*, *32*(3), 444-452.

Dakhel, A. M., Majdinasab, V., Nikanjam, A., Khomh, F., Desmarais, M. C., & Jiang, Z. M. J. (2023). Github copilot ai pair programmer: Asset or liability? *Journal of Systems and Software*, *203*, 111734.

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM*, *67*(2), 56-67.

Finnie-Ansley, J., Denny, P., Becker, B. A., Luxton-Reilly, A., & Prather, J. (2022). The robots are coming: Exploring the implications of openai codex on introductory programming. Proceedings of the 24th Australasian computing education conference,

Finnie-Ansley, J., Denny, P., Luxton-Reilly, A., Santos, E. A., Prather, J., & Becker, B. A. (2023). My ai wants to know if this will be on the exam: Testing openai's codex on cs2 programming exercises. Proceedings of the 25th Australasian Computing Education Conference,

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, *30*, 681-694.

Hosseini, R., Vihavainen, A., & Brusilovsky, P. (2014). Exploring problem solving paths in a Java programming course.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590*.

Sarsa, S., Denny, P., Hellas, A., & Leinonen, J. (2022). Automatic generation of programming exercises and code explanations using large language models. Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1,

Tan, P.-H., Ting, C.-Y., & Ling, S.-W. (2009). Learning difficulties in programming courses: undergraduates' perspective and perception. 2009 International Conference on Computer Technology and Development,

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270.

應用線上協作學習於新詩創作課程對自我反思、洞察及學習態度的影響

Effect of Online Collaborative Learning on Self-reflection, Insight, and Learning Attitude in Poetry Writing Courses

梁心怡[1*]，謝予騰[2]
[1]義守大學創意整合設計學系
[2]國立臺南大學國語文學系
[*] cindy@isu.edu.tw

【摘要】 新詩創作可幫助學生理解自我和他人的感受，然而，若缺乏適切學習輔助引導，學生往往難以有效組織想法，影響其創作效率及新詩創作學習表現。協作學習有效幫助學生創意發想，透過同儕的回饋改善學生新詩創作學習成效，因此，本研究應用線上協作學習模式於新詩創作課程，探討相較於一般個人創作學習模式，不同學習模式對自我反思、洞察及學習態度的影響。結果顯示兩種學習模式在自我反思和學習態度上並無顯著差異，然而線上協作學習模式可有效促進學生認知校準，促進自我反思及洞察。考量本研究的實驗介入時間較短，以及小樣本數之限制，實際成效差異有待未來研究進一步評估。
【關鍵字】 新詩學習；協作學習；自我反思；自我洞察；創意寫作

Abstract: *Poetry writing assists students in understanding their own and others' emotions. However, without appropriate learning support, students often struggle to effectively organize their thoughts, which in turn affects their creative efficiency and learning outcomes. Collaborative learning effectively supports students in generating creative ideas and enhancing their poetry writing performance through peer feedback. Therefore, this study applied an online collaborative learning approach to a modern poetry composition course, aiming to investigate the effects of different learning approaches on self-reflection, insight, and learning attitude. The results indicate that the online collaborative learning approach can effectively facilitate students' cognitive calibration, supporting the development of self-reflection and insight. However, considering the short duration of the intervention and the small sample size, further studies are needed to evaluate the long-term effects.*
Keywords: poetry learning, collaborative learning, self-reflection, insight, creative writing

1. 前言

新詩是現代白話文重要的語言實驗與創新，是文學體裁中一種創意寫作(Creative writing)的形式，透過精練的語言、符號與象徵結合語言修辭和聲韻節奏，運用想像力與表達創作者的情感與思想(Mahmud, 2017)。新詩在文學上佔有重要的地位，過去研究發現新詩可幫助學生理解自我和他人的感受，透過新詩傾聽自我的內在聲音，學習與自己溝通，促進建構健全人格與身心健康發展(Sigvardsson, 2020)。透過新詩學習可培養學生的創意思考，幫助學生反思自我與關注社會議題，因此過去常被用於品德教育，透過新詩描寫的各種議題幫助學生理解及反思道德與人性(Guttesen & Kristjánsson, 2022)；或是做為一種有效的心理工具，幫助學生建立自我認同，從相似經驗中自我療癒，與自我和解，重新認識自己(Eliopulos & Moffett,

2005; Pearson et al., 2022)。

儘管新詩在促進心理健康、人格發展、品德教育、培養人文素養等方面都具有重要價值，但許多學生對學習新詩抱持排斥態度，認為新詩是無趣難懂的，甚至害怕接觸新詩(Molderez et al., 2021)。負面的學習經驗大幅影響學生學習態度與動機，影響新詩學習成效與創作意願(Hughes, 2008)。因此，需要提供適切的輔助幫助學生理解詩意，改善新詩創作學習的觀感。

學生解讀語意困難的主要原因在於，新詩敘述的方式和日常語言用法不同，新詩透過創作者的想像力描述其情感經驗(Emotional experiences)，依循的非一般理性敘事邏輯，而是透過較私密的心理象徵與內在邏輯(Inner logic)，描述創作者主觀認知與感受的外在真實與現象(Dilthey, 1985)。過去研究指出自我反思與自我洞察可幫助學生理解自己的內在聲音，進而用文字形式書寫自身的情感經驗，透過新詩創作表達詩意(Rosenhan & Galloway, 2019)。自我反思與自我洞察均為理解自我及他人，維護心理健康和社會和諧人際關係的重要能力，需要適當引導與練習方能培養相關能力。過去講授新詩創作透過新詩鑑賞及創作，兩者相輔相成，引導學生先透過閱讀新詩，理解新詩文本結構及常見創作手法，而後透過創作檢驗應用所學成果(Jusslin & Höglund, 2021)。教師通常會引導學生進行創意發想，如：腦力激盪法，請學生就感興趣的主題聯想所有想到的詞彙與感受，而後進行新詩創作 (Wiseman, 2011)。然而，若缺乏適切的教學策略輔助，學生很難將抽象的想法與感受，透過適切的文字形式清楚表達(Li et al., 2023)。學生往往要花費大量時間創意發想及組織想法，創作效率低落，新詩創作成果也不盡理想 (Kangasharju et al., 2024)。

為改善新詩創作學習成效，過去研究發現協作學習可有效幫助學生創意發想，改善創作效率(Kangasharju et al., 2024)。透過與同儕討論的協作學習模式，可促進學生探索自我，促進自我反思，在與他人交流想法與感受的過程中，更加清楚認識自我及如何有效透過詩歌表達意涵 (Hurt, 2013)。此外，透過協作學習，也提供一個機會讓學生間可互相支持，共同解決創作過程中遇到的困難，改善學習觀感與態度(Woloshyn et al., 2022)。然而，目前關於新詩創作成效的研究有限，較少實證研究探討與評估不同教學方法在促進自我反思、洞察的學習成效(Moudatsou et al., 2020)。

綜上所述，本研究提出一線上協作學習模式輔助新詩創作學習，透過六周的新詩創作學習單元課程，希望讓學生認識新詩創作的概念，並透過新詩創作練習，逐步探索自我與練習自我表達，透過不斷與自我對話的過程，培養自我反思、洞察的能力，改善新詩創作學習態度。其研究問題如下所敘述：

1. 線上協作學習模式以及一般新詩創作學習，如何影響學生的自我反思？
2. 線上協作學習模式以及一般新詩創作學習，如何影響學生的自我洞察？
3. 透過線上協作學習模式，是否相較於一般新詩創作學習，可有效改善學生的學習態度？

2. 教學情境設計

本研究以臺灣南部兩所大學的通識國文課程為介入課程，規劃規劃共 6 週的新詩學習與創作課程，透過每週 2 節課授課(100 分鐘)及每週作業，引導大學一、二級學生學習新詩創作的基本概念，並透過線上協作學習平台分組共創，進行個人新詩創作。其課程設計分三階段：

(1) 新詩賞析介紹(2 週)：

課程介紹新詩之定義，透過詩人代表作品介紹，了解現代詩之多元可能，包括形式、語言、風格和不同時期詩人之創作特色等等，突破過去學生因考試而對現代詩產生的可能排斥

心理，進一步找到自己喜歡或可以接受的作品。

(2) 新詩創作練習(3 週):

認識新詩的基本概念後，第二階段為新詩創作練習，每週為一指定命題的新詩創作練習，透過課堂教師示範及分組共創活動，引導學生逐步進行意象發想，創作自己的新詩作品。新詩創作練習分為課堂分組共創活動和課後個人新詩創作兩部分，在課堂活動中，教師在講解完新詩創作的概念後，學生 4 到 5 人一組，分組進行意象發想共創，透過各組的 Miro 線上協作平台，進行與主題相關名詞、動詞、修飾語(形容詞或副詞)的聯想。在各組初步腦力激盪後，教師將引導學生分類與評估所想到的詞彙，挑選出較創新、適合的詞彙進行語句組合用以書寫意象，完成個人的新詩創作作品。如圖 1 所示。學生在課堂中分組完成意象發想共創後，可應用共同發想的意象詞彙表格，進行個人的新詩創作練習，並於下周上課前繳交個人新詩創作成果，以利課程評分。

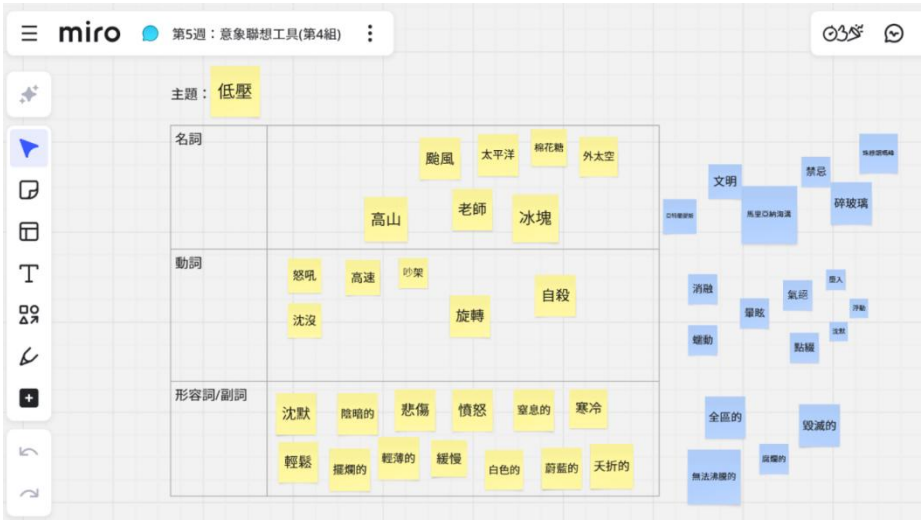


圖 1 學生分組使用線上協作平台共創進行創意發想

(3) 新詩文學獎(1 週):

最後一周模擬國內文學獎主流運作方式，評量每位學生之最後新詩創作成果。以同儕互評方式，請學生挑選三週練習作品中最滿意的一首繳交，於 Padlet 線上平台發佈作品 (圖 2)。評選分為兩階段：第一階段全班共同投票，選出前 6 名選作品，而後全班評比入選的 6 首作品名次，最終依照總得分評選出第一至第六名作品。

圖 2 學生於線上平台發表新詩創作作品

3. 實驗設計

3.2 *實驗對象*

本實驗研究對象為臺灣南部的兩所大專院校通識國文課程，參與實驗學生為大一至大二來自各學院學系的學生，以準實驗研究設計，將實驗對象依授課班級分為實驗組及控制組，進行新詩創作學習活動。排除未全程參與課程活動等因素之無效樣本，分析所得有效樣本共28位學生，進行後續結果分析：實驗組班級共18位學生使用線上協作學習模式進行新詩創作學習；控制組班級共10位學生使用一般新詩創作學習模式進行教學。

3.3 研究工具

自我反思與洞察問卷改編自 Silvia (2020)的自我反思與自我洞察量表，以5點量表評估參與學生使用不同學習策略進行新詩學習活動後，對自我反思與洞察的影響。自我反思量表(Cronbach's α=0.92)共12題，評估個人評估自我感知、情緒及行為的參與程度和需求，如：我經常檢視我當下的感受；自我洞察量表共8題(Cronbach's α=0.75)，評估個人覺察自我狀態的自我理解程度，如：我對自己做過的行為有清楚的想法認知。

學習態度問卷取自黃國禎學者研究團隊採用之學習態度問卷(Cronbach's α=0.91)，以5點量表共7個問項評估學生對該學習主題的學習態度(Hwang et al., 2013)。

除了自我反思與洞察問卷和學習態度問卷外，為了解實驗組學生對線上協作學習模式的觀感，另設計3題開放式問項，以利了解學生對自我反思、洞察、學習態度的觀感，用以解釋量化分析結果：

(1) 你覺得在現代詩創作單元學習後，自己在新詩創作手法或思考上，有什麼差異嗎？
(2) 你覺得透過現代詩創作學習可以獲得什麼？或是學習到什麼？有什麼改變？
(3) 整體而言，上完課後，你喜歡現代詩嗎？為什麼？

3.4 實驗流程

本研究第一週課程開始前，首先進行自我反思與洞察、學習態度前問卷(10分鐘)，而後進行新詩基礎概念介紹。在進行完前2周的新詩基本概念介紹後，第3週至第5週的課程，實驗組學生使用線上協作學習模式學習，在教師示範講解新詩創作概念後，分組透過 Miro 線上協作平台進行意象詞彙創意發想，而後應用其分組共創的意象詞彙，進行個人新詩創作練習；控制組學生使用一般新詩創作學習模式學習，教師在示範與講解新詩創作概念後，隨機請學生逐一上臺填寫所聯想到與主題相關的名詞、動詞及修飾語，而後請學生完成個人的意象詞彙聯想表格，並使用其表格中的詞彙進行個人新詩創作練習。總言之，實驗組使用線上協作學習模式進行創意發想，而後進行個人的新詩創作練習，控制組學生則使用一般個人學習模式進行創意發想與新詩創作練習。

學生進行完5週的課程後，於第6週上課前繳交一首新詩代表作品做為最終課程成果，第6週成果評估模擬目前文學獎評選進行方式，以同儕互評方式，評選出全班最好的6首新詩作品。並於講評結束後，請學生完成填寫自我反思與洞察、學習態度後問卷 (10分鐘)。

4. 結果

4.1 自我反思

本研究以共變數分析(ANCOVA)了解學生使用不同學習模式於新詩創作學習對自我反思的影響，在進行 ANCOVA 分析之前，共變數分析結果顯示符合同質性假設($F_{(1)}$=1.52，

p>.05)，可進行 ANCOVA 分析。結果顯示學生使用線上協作學習模式和一般個人學習模式，在自我反思無顯著差異(p>.05)，結果如表 1 所示。

表 1 自我反思共變數分析結果

組別	人數	前測		後測			F
		平均值	標準差	平均值	調整後平均	標準誤	
實驗組	18	4.09	0.79	4.04	3.92	0.11	0.91
控制組	10	3.65	0.40	3.53	3.75	0.14	

進一步分析不同學習模式對自我反思的變化，Mann-Whitney U 檢定分析結果顯示兩組學生在新詩創作學習後，自我反思的自我評分均略有下降，但兩組未達顯著差異(p>.05)，結果如表 2 所示。

表 2 自我反思差距 Mann-Whitney 檢定分析結果

組別	人數	平均值	標準差	U	z	p^2
實驗組	18	-0.05	0.37	83	-0.34	0.76
控制組	10	-0.12	0.59			

4.2 自我洞察

為了解不同學習模式對自我洞察的影響，本研究在進行 ANCOVA 分析前，先檢驗資料同質性，同質性變異數檢定結果顯示符合同質性假設(Levene's test= $F_{(1)}$=0.61，p>.05)，結果顯示控制組學生(調整後平均 3.94，標準誤 0.05)在自我洞察表現顯著高於實驗組學生(調整後平均 3.71，標準誤 0.04)，$F_{(1)}$=10.75，p<.05。進一步比較其自我洞察評分差距，表 3 顯示一般新詩創作學習的學生自評改善自我洞察表現，而使用線上協作學習模式的實驗組學生卻有下修自我洞察表現的傾向(p<.05)。

表 3 自我洞察的 Mann-Whitney U 檢定分析結果

項目	組別	人數	平均值	標準差	U	z	p^2
自我洞察前測	實驗組	18	4.07	0.50	15.00	-3.60	.000***
	控制組	10	3.37	0.30			
自我洞察後測	實驗組	18	3.93	0.50	35.00	-2.64	.007**
	控制組	10	3.53	0.21			
自我洞察差距	實驗組	18	-0.14	0.12	15.00	-3.60	.000***
	控制組	10	0.16	0.18			

**p<.01, ***p<.001

分析實驗組學生共 47 則回饋內容結果可知，大多數學生提到在線上協作學習可改變他們原有的思考方式(N=19)，從更多元的面向思考，如：有學生(E12)提到線上學習模式時描述：「這種思考方式很不一樣，很有突破性。」、「思考更多元。」(E01)。學生也肯定新詩創作學習對自我洞察的改善(N=27)，如有學生(E19)說明：「更加理解自己的感受，對於情感的表達更加順暢」，「可以更認真的去感受生活帶來的美好，悲傷，一切」(E04)。透

過新詩創作學習可讓學生「去深思自己的想法」(E17)，有機會檢視自己的想法與感受，透過文字表達與抒發情緒，「生活中一些無以名狀的想法跟情緒，都能融入字詞中」(E09)，「可以跳出不一樣的視角面對自己的情意」(E11)，這些回饋反映新詩創作學習可幫助學生深入自我探索，理解自我的想法、需求與情感，因而影響學生的自我洞察程度。

4.3 學習態度

為了解不同學習模式對學習態度的影響，本研究在進行 ANCOVA 分析前，先檢驗資料同質性，同質性變異數檢定結果顯示符合同質性假設($F_{(1)}$=0.00，p>.05)，可進行 ANCOVA 分析。結果如表 4 所示，兩組學生在學習態度上並無顯著差異(p<.05)。

表 4 學習態度共變數分析結果

組別	人數	前測		後測			F
		平均值	標準差	平均值	調整後平均	標準誤	0.003
實驗組	18	3.93	0.51	4.17	3.94	0.11	
控制組	10	3.30	0.40	3.53	3.95	0.16	

分析學生的回饋，多數學生(N=16)肯定線上協作學習模式對新詩創作學習的效益，但也有少數學生排斥此種學習方式(N=3)，認為線上協作學習模式比較麻煩且缺乏與人直接的互動，如：有學生說明：「用電子設備蠻沒溫度的」(E05)、「肯定有更不麻煩的方式」(E11)。

5. 討論及結論

本研究探討線上協作學習模式和一般個人學習模式應用於新詩創作課程，對自我反思、洞察和新詩創作學習態度的影響。結果顯示兩組學生在自我反思的表現無顯著差異，然而相較於前測而言，兩組學生在新詩創作學習後，均有微幅下修自我反思的自我評分。此外，關於自我洞察的成效影響，結果顯示控制組學生的自我洞察表現顯著優於實驗組，實驗組學生在線上協作學習後，有顯著下修自我洞察評分的現象。從實驗組學生的回饋結果可知，大多數學生均有意識到新詩創作學習對自我反思和自我洞察的影響，學生肯定線上協作學習模式對於探究及認識自我的正向價值，可幫助學生更深入了解自我，促進情感表達與認知。因此，量化分析結果與質化分析結果衝突的原因，可能反映了認知校準效應 (Cognitive Calibration)，說明學生在經過新詩創作學習後，對於自我認知評估的校正現象(Kruger & Dunning,1999)。學者以鄧寧-克魯格效應（Dunning-Kruger Effect）解釋，學生在未學習掌控一新學科知識或技能時，容易存在認知偏誤，造成對自己能力評價與實際表現情況不一致的情況(Dunning, 2011)。此現象可以用來解釋本研究結果：雖然學生有感於新詩創作學習對自我反思和洞察的助益，卻在自我評估中下修自我反思和洞察的程度。

研究結果顯示兩組學生均有下修自我反思程度的情況，且實驗組學生同時下修自我洞察程度，這可能暗示大多數學生在自我反思與洞察方面，有高估表現的傾向。考量錯誤的自我認知可能影響學習表現，而自我反思和洞察有助於認識自我，幫助學生有效自我調節，調節生活中遇到的壓力與挫折等情況，維持身心健康的狀態(Nesbit, 2012)。新詩創作學習是一種有效的學習方法，可用來促進學生培養自我反思與洞察。此外，本研究結果顯示線上協作學習模式相較於一般個人學習模式，可更有效輔助學生自我反思，增進自我覺察。這可能是因為在線上協作學習模式中提供學生與同儕交流與討論想法的機會，而來自同儕和教師的外在

回饋，可幫助學生有效進行認知校準，促進自我反思，增進自我覺察與理解 (Yan et al., 2023)。因此，線上協作學習模式相較於一般個人學習模式，對於自我洞察有顯著的認知校準效應。

儘管本研究結果顯示線上協作學習對自我反思、洞察的認知校準效應，結果卻未能提升學生的自我反思和洞察程度。這可能是由於自我反思和洞察需要長時間培養才能看出變化，而在學習初期，可能經歷新資訊的整合以及既有基模的破壞與重構，造成暫時的困惑和認知衝突，因此影響自我反思與洞察表現(Rumelhart & Norman, 1976)。因此，關於線上協作學習模式對於自我反思和洞察的影響，可能需要更長期的觀察，了解其成效。

針對第三個研究問題，結果顯示學生對於線上協作學習模式和一般個人學習模式，此兩種學習模式均有正向的學習態度，對於新詩創作的學習態度影響並無顯著差異。未來可能需要更長的介入時程，評估兩種學習模式對於新詩創作學習態度的影響。

本研究為探討自我反思及洞察提供參考實證依據，未來研究可探討不同學習模式和學習活動的成效差異。然而，考量到本研究應用於通識課程，在收案管理上有較大困難影響收案人數，研究結果可能難以適用於其他類似情境或不同學科，未來需納入更多樣本數，以利檢驗線上協作學習模式對新詩創作學習成效之影響。

誌謝

本研究感謝國科會專題研究計畫的支持，計畫編號為 NSTC 113-2410-H-214-001。

參考文獻

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in experimental social psychology* (Vol. 44, pp. 247-296). Academic Press.

Eliopulos, T. D., & Moffett, T. S. (2005). *The Everything Writing Poetry Book: A Practical Guide To Style, Structure, Form, And Expression*. Everything Books.

Guttesen, K., & Kristjánsson, K. (2022). Cultivating virtue through poetry: an exploration of the characterological features of poetry teaching. *Ethics and Education*, *17*(3), 277-293.

Hughes, J. (2008). The 'screen-size' art: Using digital media to perform poetry. *English in Education*, *42*(2), 148-164.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Mahmud, M. (2017). Teaching students to develop paragraphs by poetry writing. *International Journal of Language Education*, *1*(1), 37-50.

Molderez, I., Baraniuk, D., & Lambrechts, W. (2021). The role of poetry in promoting a sustainability mindset: Walter Benjamin as a guide toward a slow journey. *Frontiers in Sustainability*, *2*, 694317.

Nesbit, P. L. (2012). The role of self-reflection, emotional management of feedback, and self-regulation processes in self-directed leadership development. *Human Resource Development Review*, *11*(2), 203-226.

Pearson, M., Rennick-Egglestone, S., & Winship, G. (2022). The poetic wavelength–a narrative interview study exploring the potential of poetry to support meaning making and recovery following psychosis. *Psychosis*, 1-12.

Rosenhan, C., & Galloway, N. (2019). Creativity, self-reflection and subversion: Poetry writing for Global Englishes awareness raising. *System*, *84*, 1-13.

Rumelhart, D. E., & Norman, D. A. (1976). *Accretion, tuning and restructuring: Three modes of learning*. Center for Human Information Processing, Department of Psychology, University of California, San Diego.

Sigvardsson, A. (2020). Don't fear poetry! Secondary teachers' key strategies for engaging pupils with poetic texts. *Scandinavian Journal of Educational Research*, *64*(6), 953-966.

Silvia, P. J. (2022). The self-reflection and insight scale: Applying item response theory to craft an efficient short form. *Current Psychology*, *41*(12), 8635-8645.

Woloshyn, V., Obradovć-Ratković, S., Julien, K., Rebek, J.-L., & Sen, A. P. (2022). Breathing our way into mindful academic writing: A collaborative autoethnography of an online writing community. *Journal of Further and Higher Education*, *46*(8), 1135-1148.

Yan, Z., Panadero, E., Wang, X., & Zhan, Y. (2023). A systematic review on students' perceptions of self-assessment: usefulness and factors influencing implementation. *Educational Psychology Review*, *35*(3), 81.

從性別差異的角度探討使用數位遊戲來學習同義字之影響

The Impacts of Game-based English Synonyms Learning: A Gender Difference

范斐芃[1]，陳攸華[2]
中央大學網路學習科技研究所
*youmi24@g.ncu.edu.tw
* sherry@cl.ncu.edu.tw

【摘要】 學生在英文字彙學習中常因壓力而缺乏動機，為此本研究融合手眼協調與擴增策略，設計了英語詞彙學習遊戲，期望透過低壓力的互動提升學習動機。研究採用開放式問卷進行質化分析，探討不同性別對遊戲式學習系統的使用觀感。結果顯示，男女學習者皆重視按鈕設計與遊玩順暢度，但在其他面向呈現差異，女性偏好色彩、趣味性與操作引導，且需求較低；男性重視資訊清晰與即時回饋，展現目標導向的特徵，並肯定低壓力學習氛圍。本研究揭示遊戲式學習中性別觀感差異，也強調設計應兼顧美感與資訊、趣味性與功能性，以滿足多元學習需求。

【關鍵字】 性別差異；遊戲式學習；英語詞彙

Abstract: *Students lack motivation to learn English vocabulary. Accordingly, this study combines hand-eye coordination and expanding strategies to design a Game-based English Vocabulary Learning System (GEVLS) to enhance their motivation. This study used an open-ended questionnaire to qualitatively explore how gender differences influenced their perception for the GEVLS. The results showed males and females valued button design and smooth gameplay. However, females preferred color coordination, entertaining elements, and clear guidance, and showed lower demands; males focused on information clarity and immediate feedback, showing goal-oriented tendencies and appreciating low-pressure learning. This study exposed gender differences in perceptions for the GEVLS and highlighted the essence of balancing aesthetics and information, as well as entertainment and functionality, to meet diverse needs of learners.*

Keywords: gender differences, game-based learning, English vocabulary

1. 前言

英語正迅速成為學術溝通的通用語言(Alhasnawi, 2021)，廣泛應用於研究發表與學術會議中。這也顯示出培養學生的學術英語能力是相當必要的(Abidjanova, 2024)。此外具備良好的英語運用能力，也有助於學生參與非學術性的活動，例如日常溝通。簡言之，英語學習是學術活動或日常生活不可或缺的基礎。英語學習包含很多面向，其中詞彙佔有極其重要的地位(Susanto, 2017)。閱讀、聽力、口說、寫作和文法應用都有賴於足夠的英語詞彙作為基礎。

在另一方面，實證研究的結果也顯示出英語詞彙量的多寡會影響英語能力。例如，Maamuujav et al.(2021)分析青少年在撰寫論文時的文法與詞彙特徵，他們發現詞彙量的多寡會直接影響語言表達能力，尤其在學術寫作中更為顯著。同年，Enayat 和 Derakhshan(2021)探討了詞彙量對於第二語言（L2）口語能力的影響，其結果發現詞彙量不僅顯著影響整體口語表現，也與詞彙使用的多樣性密切相關。由此可見，持續累積詞彙量是提升整體英語能力的關鍵條件之一。

儘管詞彙學習對語言能力發展至關重要，學習者在實際學習過程中仍常面臨多項挑戰。其中一項挑戰就是學習焦慮。具體來說，學習焦慮被認為是影響外語學習最重要的情感因素之一(Wang & Zhang, 2021)。學習者可能因擔心記憶失敗或答錯而感到壓力，進而降低學習成效。而學習焦慮特別會出現在缺乏即時回饋時，學習者無法適時調整自己的學習策略(Wulantari et al., 2023)，進而對自身能力產生懷疑，甚至降低學習動機。針對學習焦慮，Fu 和 Li(2022)比較即時回饋和延遲糾正反饋對英語學習成效差異，他們發現即時回饋可以降低學習焦慮，因為即時回饋可有助於學習者釐清自身的學習盲點，進而能夠提升學習效果(Zuhaib et al., 2024)。

簡言之，即時回饋對英語學習有正向的影響，而遊戲式學習（GBL）則是一種可以提供即時回饋的機制(Rasti-Behbahani, 2021)，實證研究的結果也指出 GBL 具備降低學習焦慮的潛力。Chan 和 Lo(2021)探討遊戲化在英語課程中的效果，多數大學生和教師對在英語課堂中使用各種遊戲工具感到舒適，並認為這些工具有助於創造更具互動性的學習環境。此外，Ahmed et al.(2022) 進行了一項實驗，招募 58 名英文學習者，並將其分為 GBL 組與傳統教學組。透過外語學習焦慮量表與學習動機量表進行前測。經過一輪的學習後，再做同樣的測量表作為後測。結果顯示，採用 GBL 的學習者，其學習壓力明顯降低。之後，Fawaid et al.（2024）也驗證了此現象。將 70 名參與者平均分為實驗組和對照組。實驗組使用融合遊戲元素的互動式教科書，而對照組則接受傳統的課堂教學。研究結果發現，實驗組學生的學習焦慮顯著下降，尤其對於原本具有中度至高度焦慮的學習者影響更為明顯。

同時 GBL 也擁有多項有利於語言學習的優勢，包括提供「自由度」，促使他們願意重複練習與進行試錯(Pérez et al., 2022)，具備趣味性，可以提高學習動機與參與意願(Nadeem et al., 2023)，以及營造一個積極且具吸引力的學習環境(Vnucko & Klimova, 2023)，進而提升學習者的舒適度與信心(Wichaidit & Wichaidit, 2024)。綜上所述，GBL 被證實能有效提升英語學習的整體成效(Yang & Chen, 2023)。

由於這些優勢，GBL 廣被用於支援英語學習，然而個別差異性存在於每位學習者之間，而其中一個比較不會改變的個別差異性就是性別。有些學者認為性別不會影響學習者在基於遊戲的課堂中的積極性與參與度(Almusharraf et al., 2023)，而且對 GBL 的接受度也沒有顯著差異(Orak, 2025)。相反地，另外有些學者的研究結果顯示性別不僅會影響 GBL (Aguilar-Cruz et al., 2023)，而且也會在英語學習的效果有中所不同(Rahimi, 2022)。換句話說，性別對於 GBL 的影響尚未有定論，因此需要更多的實證研究來探討這個議題。

有鑑於此，本研究的貢獻在兩方面，其一是運用 GBL 設計一款英語詞彙學習遊戲，透過結合視覺化詞彙展示、即時回饋與互動機制來克服學習詞彙過程中的焦慮，並提升詞彙學習的整體成效。其二是以質化的實證研究來探討男性學習者和女性學習者對於使用此英語詞彙學習遊戲是否有觀感上的差異。他們的觀感差異可以用來提供設計者更具包容性的建議，以兼顧不同性別學習者的需求與偏好。

2. 英語詞彙學習遊戲

如前所述，GBL 有許多優勢，因此本研究利用這些優勢來發展一款英語詞彙學習遊戲，而此遊戲共有五大特色，詳述如下:

- 射擊式遊戲

本研究使用射擊遊戲讓學習者在低壓力的環境中學習同義字，增加英文字彙量。更具體的說，每一個任務用八個泡泡來代表八個選項，其中有兩個泡泡是正確的選項(圖 1)。學習

者需透過左右方向鍵控制飛機位置，並利用空白鍵發射子彈，使子彈能夠射中正確的泡泡，因此是否能準確發射子彈來射擊就是其中一個趣味性。而另外一個趣味性就是時間的壓力，因為每一回合的遊戲時間限制為五分鐘，學習者需在時限內射擊兩個選項，才算是完成該回合。

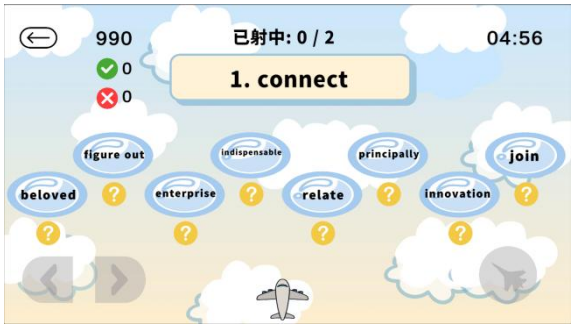


圖 1 射擊遊戲內容

圖 2 提示內容

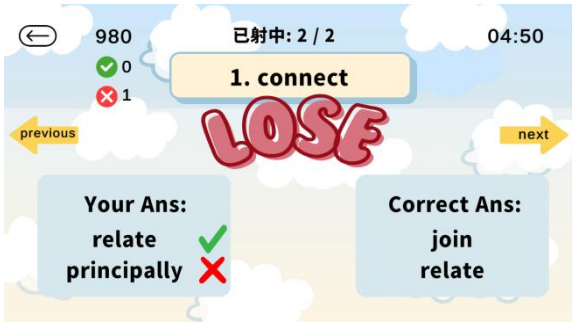


圖 3 即時回饋內容

圖 4 結算成績內容

- 反覆式練習

在每一個關卡中的題目會呈現一個英文單字，而學習者須在前述的射擊遊戲中擊中此單字的兩個同義字，而此遊戲共有 12 個關卡，因此學習者總共可以學會 36 個單字。在這 36 個單字之中，會隨機出現八個單字在任何一個關卡，因此同樣的單字會在不同關卡中重複出現，如此重複的出現，可以讓學習者獲得反覆練習這 36 個單字的機會，進而可以加強記憶與鞏固學習效果。

- 有價式鷹架

為降低遊戲過程的困難程度，進而提升其學習的信心與成效，此遊戲提供三種鷹架提示: (a) 翻譯提示--可讓學習者理解單字的基本意思，(b) 詞性提示--幫助認識單字在句中的功能，(c) 例句提示--幫助學習者掌握單字在實際情境中的應用(圖 2)。然而此三種提示屬於有價式鷹架，因為每使用一次提示訊息即扣除 10 分，而每一次的扣分則是從提示的原始分數(1000 分)來相減。使用此種有價式鷹架之目的是為了避免學習者過度依賴提示，並鼓勵他們來獨立思考和作答。

- 立即式回饋

學習者在進行每一個射擊泡泡的遊戲過程中，可以立即獲得三種回饋:(a) 當子彈發出時，遊戲會及時讓學習者知道現今的狀況，也就是會顯示已射中一個，已射中兩個或是沒有射中，(b) 在進行每一個關卡的過程，學習者會用綠色和紅色的標示來分別告知成功和失敗的關卡各有多少，(c) 在一個關卡結束之後，遊戲會顯示「WIN」或「LOSE」來讓學習者知曉是否成功過關以及正確的答案(圖 3)。

- 自我式競爭

完成全部 12 關卡之後，會進入 Final Score 的畫面，畫面中有兩個區塊，分數區和時間區(圖 4)。分數區包含任務分數、提示分數和對錯的個別題數。時間區則是有每一題作答所使用的時間。這些資訊不僅可以幫助學習者進行自我評量，以便能夠判別自己的優勢和劣勢，也可以做為自我競爭的基礎，亦即是當他們下一次再進入這個遊戲時可以比較每一次從遊戲所獲得的分數是否有所進步或退步，以便能進一步改進。

3. 研究方法

3.1. 開放式問卷

開放式問卷有以下三個優點:(a)能夠引導受訪者產出更真實且具深度的回答，相較於封閉式問卷更能反映受訪者的真實想法;(b)蒐集未被預設選項涵蓋的觀點，豐富研究中所欠缺的項目;(c)鼓勵自由表達，有助於提升資料的整體品質與詮釋力(Neuert et al., 2021)。

由於這些優點，本研究設計了一個開放試問卷，來蒐集學習者對於使用英語詞彙遊戲意見。內容包括四個面向，(a)介面設計的優點和缺點，(b)互動操作的優點和缺點，(c)內容格式的優點和缺點，(d)顏色選取的優點和缺點，再將學習者所填內容進行重新分類與彙整，藉以觀察性別差異對遊戲體驗觀感的影響。

3.2. 實驗流程

參與本研究的對象為 12 位台灣北部某大學的研究生，也就是六位女性學習者與六位男性學習者。參與者皆具有基礎英文知識，並具有基礎的電腦遊戲操作能力。整個流程分為兩個步驟，首先所有學習者皆透過電腦進行「英語詞彙學習遊戲」，題目總共 12 題，每題作答時間上限為三分鐘，完成所有關卡的最高時長為 36 分鐘。接著在結束後，學習者會填寫開放式問卷，表達他們對於英語詞彙遊戲的看法。

4. 研究結果

本研究以開放式問卷來探討性別如何影響學習者對於「英語詞彙學習遊戲」的使用觀感，包括介面設計的偏好、遊戲體驗的需求、操作方式的需求以及整體使用感受。然後再進一步根據所收集到的問卷進行質化分析。質化分析的結果顯示女性學習者和男性學習者在“按鈕設計”(女性: N=4,男性:N=5)和“遊玩順暢度”(女性: N=4,男性:N=4)展現出相似的觀感，關於前者，明確的按鈕互動提示是介面設計的共同要求。例如:“返回首頁的按鈕沒有變色”以及“題目叉叉沒有回饋”。此結果意味著所有學習者都傾向能藉著動態效果，來收到明確的操作回應。

關於後者，流暢且直覺的操作是男性學習者和女性學習者的共同需求。例如:“無法透過 Enter 或空白鍵進入下一題”，以及“例句要再按一次有點不順手(個人感想)”，反映出不論是哪一種性別的學習者都期待能以更少的動作完成更多操作。除了上述的相似觀感之外，男性學習者和女性學習者在使用體驗上也展現出不同的需求。如差異化教學 (Differentiated Instruction) (Tomlinson, 2017)所指出，教學環境應依照學生個別差異及需求來做彈性調整。因此本研究亦根據分析結果，提出具體的設計改善建議，詳述如下:

4.1 女性(顏色) vs. 男性(訊息)

女性學習者和男性學習者的其中一個相異觀感在於，前者看重視覺的設計，而後者則重視資訊的呈現。更明確的說，女性學習者在顏色面向上的關注顯著多於男性(女性：N=5，男性：N=1)，例如：“作答前跟作答後的返回，顏色不同，感覺可以用統一的顏色”和“進入

關卡的按鈕是綠色的，跟整體色系有落差”。這些意見顯示，女性學習者對色彩變化具有較高敏感度，並偏好整體視覺呈現統一與和諧的風格。

相較之下，男性學習者則在訊息面向表達出較多的需求(女性: N=2,男性:N=6)，例如：“提示型的標語，可以一致為英文或中文，像裡面有：已射中、ExitGame、遊戲說明”以及“最後結算畫面顯示秒數的區塊可以同時顯示該題答題情況”。這些被提出的需求反映出男性學習者傾向資訊結構的清晰性、功能設計的明確性與資訊傳遞的一致性。換句話說，儘管男女學習者皆關注介面的可用性與操作回饋，但前者偏好視覺設計的和諧與整體感，後者則傾向於語言與功能設計的一致與清晰。這樣的差異可能來自於女性男性的不同。女性可能色彩感知比較好(Panchal et al., 2013)，傾向注意介面整體的色彩協調與風格一致性，並可能會透過色彩變化來辨識系統狀態與操作回饋。相對地，男性可能對美感較不敏感，但在資訊的邏輯結構上則較為敏銳，傾向透過一致、明確的語言資訊中掌握操作邏輯與系統提示。

基於上述的結果，當從性別導向來提供實務建議時，介面設計時需要同時考慮顏色的設計與訊息的呈現，以能夠兼顧不同性別學習者的偏好。例如，在設計按鈕時，相同的功能使用相同的顏色，並搭配一致的文字標示，提升整體一致性。這樣的做法有助於兼顧女性學習者對美感的要求，以及男性學習者對邏輯與資訊一致性的需求。

4.2 女性(趣味性) vs. 男性(回饋性)

女性學習者與男性學習者在遊戲體驗上展現出不同的關注重點，更具體的說，女性學習者偏向情感層面，尤其是趣味性(N=4)，例如：“射擊的小動畫很可愛＂和“遊戲性豐富”。值得一提的是女性學習者提出關於趣味性的回應皆是正面的，意味著女性學習者對此系統的趣味性與視覺互動元素表達肯定，而男性學習者不僅沒有呈現出這種肯定，對於趣味性也沒有任何回應。相反的，男性學習者則關注學習面，也就是它的回饋性(女性: N=1,男性:N=3)。例如：“缺少單字發音回饋”和“答錯會有正確題示”，這些發現顯示出男性學習者期待能透過即時的提示，協助辨識正確的答案，並在錯誤後即時修正。

綜言之，女性學習者比較重視遊戲的趣味程度與視覺吸引力，而男性學習者則重視遊戲能否提供即時且具功能性的學習回饋。這種差異可能來自於女性較容易被情感氛圍所吸引(Fischer et al., 2018)，所以會對有趣性的反應較強，男性則傾向以目標導向的方式進行學習(Thao et al., 2023)，更重視資訊的回饋是否能幫助他們達到學習目標。所以在開發遊戲式學習時，應同時兼顧娛樂性與學習性，以回應不同性別學習者在遊戲互動中的需求差異。例如，在射擊泡泡的環節中，可透過設計具有動態效果的射擊畫面，來幫助女性學習者感受到趣味性，另一方面也提供即時回饋機制，讓男性學習者清楚感知當前的學習狀態。

4.3 女性(如何玩) vs. 男性(如何學)

女性學習者與男性學習者在操作方式呈現明顯差異，更具體的說，女性學習者在操作層面的「如何玩」有較多正面反饋(女性: N=6,男性:N=3)，例如: “首頁直接顯示要破關的關卡，一目瞭然~~”和“就算沒有看說明書也可以知道是什麼樣的遊戲”，反映女性學習者偏好能以簡單明確的方式呈現遊戲邏輯與操作步驟，降低學習門檻，而男性學習者則是更關注「如何學」(女性: N=3,男性:N=9)。例如: “提示的觀看例句，很明確讓人知道”與“一次可以了解三個相似單詞”，顯示男性學習者更在意學習內容的呈現。

換句話說，女性需要遊戲操作過程中的清楚引導，男性則更重視遊戲所能提供的學習內容。這樣的差異可能是兩者在遊戲經驗有所差異。女性可能較少接觸數位遊戲(Lopez-Fernandez et al., 2019)，對遊戲操作流程較不熟悉，因此依賴清楚的引導來幫助她們順利完成任務。相反的，男性學習者通常對遊戲操作較為熟悉，可是在英語學習上較女生困難

(Chen & Lin, 2024)，所以更關注學習資訊的呈現。綜合來看，遊戲式英語學習系統需同時兼顧簡單易懂的操作及學習資訊的整合，以符合不同性別學習者的使用需求。例如，在設計提示機制時，提供明確可辨識的按鈕，並在操作後呈現完整的學習內容，協助學習者理解學習目標與操作步驟。

4.4 女性(低需求) vs. 男性(低壓力)

當細部檢查女性對於問卷內容的回覆之後，發現她們填寫很多的“無”(女性：N=7，男性：N=1)，顯示出她們對於「英語詞彙學習遊戲」並沒有太多意見，此意味著她們呈現「低需求」的期待。相反地，男性卻有不少的回應，特別是他們對於「低壓力」學習環境有正面看法(N=2)，例如：“配合非常和諧，讓人使用起來不會有壓力”與“結算畫面的資訊，呈現得很整齊，資訊量大，但負荷不高”。這些回應顯示，男性學習者特別在意學習過程中的認知負荷。

總結以上，女性學習者和男性學習者都對沒有太多負面回應，女性學習者普遍未表達過多額外需求，而男性學習者則對低壓力的學習氛圍有正面看法。這樣的差異再一次與他們的遊戲經驗和語言能力有關。如前所述，女性較少接觸數位遊戲，因此在回應中較無其他需求。而男性學習者沒有太高的英語能力，所以希望在低壓力的環境學習，避免過多的認知負荷。因此，在設計遊戲式英語學習系統時，應該注意不要給予男性學習者過多的壓力。例如，在設計遊戲的成績結算畫面時，可將不同類型的成績（如正確題數、用時、分數）以區塊區分，並清楚標示呈現順序與分類，避免資訊過多造成認知負擔。

5. 結論

本研究探討性別差異如何影響學習者對於「英語詞彙學習遊戲」的使用觀感。結果顯示，男女學習者在「按鈕設計」與「遊玩順暢度」方面有相似需求，皆重視明確的操作回饋與直覺化的流程設計。然而，兩者亦展現出顯著差異：女性學習者偏好整體「色彩」協調與視覺「趣味性」，男性學習者則更重視「資訊呈現」的清晰性與「即時回饋」；操作上，女性關注「如何玩」的引導性，男性則偏向「如何學」的內容功能。此外，女性回饋呈現「低需求」特徵，男性則強調「低壓力」學習氛圍的重要性 (圖 5)。

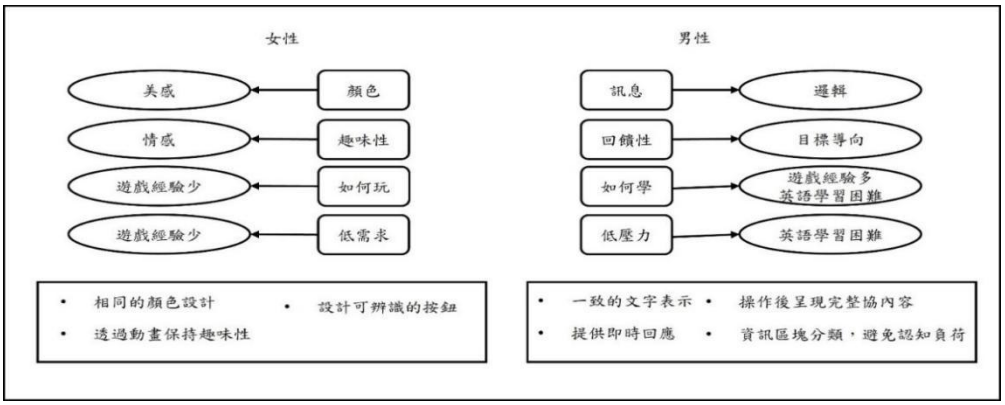


圖 5 結果架構圖

從以上結果來看，本研究在性別差異與遊戲式學習設計兩個面向皆有所貢獻。針對前者，本研究揭示不同性別在使用遊戲式學習時所展現的學習方式差異，進一步拓展對男女學習者需求差異的理解。換句話說，本研究為差異化教學提供性別導向設計的參考依據，有助於教師因應不同性別的偏好與需求，來進行教學設計，以提升學習成效與參與度。針對後者，本研究強調遊戲設計應考量性別差異的互動偏好，並為未來的個人化學習提供了設計參考。因此，設計其他遊戲式學習時可以參考這些意見，來滿足各個性別的學習偏好。

在另一方面，未來研究可進一步探討如何針對不同性別學習者設計更具適應性的遊戲式學習介面與回饋機制，特別是在提升使用者體驗與降低學習壓力方面，做出更具體的優化方案。雖然本研究有前述之多重貢獻，然而，本研究亦存在若干限制，例如樣本數較少且僅聚焦於性別差異，未能納入其他個體差異變項，如先備知識、認知風格等。未來研究期望擴大樣本規模，並納入更多元的學習者特徵，以建立更具代表性與應用性的設計建議。

參考文獻

Abidjanova, S. (2024). Highlighting How Lack of Preparation in Academic English Impacts Student Success in Global University Contexts. *Talqin va tadqiqotlar ilmiy-uslubiy jurnali*, *2*(56), 94-98.

Aguilar-Cruz, P. J., Wang, P., Xiang, Z., & Luo, H. (2023). Factors influencing game-based learning in the Colombian context: A mixed methods study. *Sustainability*, *15*(10), 7817. https://doi.org/10.3390/su15107817

Ahmed, A. A. A., Ampry, E. S., Komariah, A., Hassan, I., Thahir, I., Hussein Ali, M., Fawzi Faisal, A., & Zafarani, P. (2022). Investigating the Effect of Using Game-Based Learning on EFL Learners' Motivation and Anxiety. *Education Research International*, *2022*(1), 6503139. https://doi.org/10.1155/2022/6503139

Alhasnawi, S. (2021). English as an Academic Lingua Franca: Discourse hybridity and meaning multiplicity in an international Anglophone HE institution. *Journal of English as a Lingua Franca*, *10*(1), 31-58.

Almusharraf, N., Aljasser, M., Dalbani, H., & Alsheikh, D. (2023). Gender differences in utilizing a game-based approach within the EFL online classrooms. *Heliyon*, *9*(2). https://doi.org/10.1016/j.heliyon.2023.e13136

Chan, S., & Lo, N. (2021). *Impacts of Gamification in Remote Learning in Tertiary Education* Proceedings of the 7th International Conference on e-Society, e-Learning and e-Technologies, Portsmouth, United Kingdom. https://doi.org/10.1145/3477282.3477290

Chen, S. Y., & Lin, Y.-Z. (2024). An investigation of a customizable entertaining animated e-book: A gender difference perspective. *International Journal of Human–Computer Interaction*, 1-12. https://doi.org/10.1080/10447318.2024.2364467

Enayat, M. J., & Derakhshan, A. (2021). Vocabulary size and depth as predictors of second language speaking ability. *System*, *99*, 102521. https://doi.org/10.1016/j.system.2021.102521

Fischer, A. H., Kret, M. E., & Broekens, J. (2018). Gender differences in emotion perception and self-reported emotional intelligence: A test of the emotion sensitivity hypothesis. *PloS one*, *13*(1), e0190712. https://doi.org/10.1371/journal.pone.0190712

Fu, M., & Li, S. (2022). The effects of immediate and delayed corrective feedback on L2 development. *Studies in Second Language Acquisition*, *44*(1), 2-34. https://doi.org/10.1017/S0272263120000388

Lopez-Fernandez, O., Williams, A. J., & Kuss, D. J. (2019). Measuring female gaming: Gamer profile, predictors, prevalence, and characteristics from psychological and gender perspectives. *Frontiers in Psychology*, *10*, 898. https://doi.org/10.3389/fpsyg.2019.00898

Maamuujav, U., Olson, C. B., & Chung, H. (2021). Syntactic and lexical features of adolescent L2 students' academic writing. *Journal of Second Language Writing*, *53*, 100822. https://doi.org/10.1016/j.jslw.2021.100822

Nadeem, M., Oroszlanyova, M., & Farag, W. (2023). Effect of digital game-based learning on student engagement and motivation. *Computers*, *12*(9), 177. https://doi.org/10.3390/computers12090177

Neuert, C., Meitinger, K., Behr, D., & Schonlau, M. (2021). The use of open-ended questions in surveys. *Methods, data, analyses: a journal for quantitative methods and survey methodology (mda)*, *15*(1), 3-6.

Orak, S. D. (2025). EFL Students' Perception Of Gamification-Integrated English Language Education. *Pegem Journal of Education and Instruction*, *15*(1), 143-158. https://doi.org/10.47750/pegegog.15.01.14

Panchal, G. S., Mehta, A. S., Nair, G., Dani, J. K. S., Panchal, J. R., & Jadeja, J. (2013). A comparative study of color perception in young males and females. *Int. J. Basic Appl. Physiol*, *2*(1), 177-182.

Pérez, R. F., Chung, S., Zastavker, Y. V., Bennett, V., Abdoun, T., & Harteveld, C. (2022). How Can Game-Based Learning Affect Engineering Students' Confidence? 2022 IEEE Frontiers in Education Conference (FIE),

Rahimi, A. R. (2022). The role of online L2 motivational self system in predicting EFL learners' online achievements: The case of language massive open online course (LMOOC). *Journal of Teaching Persian to Speakers of Other Languages Print Issn*, *5394*, 2322. https://doi.org/10.30479/jtpsol.2023.17442.1594

Rasti-Behbahani, A. (2021). Why digital games can be advantageous in vocabulary learning. *Theory and practice in language studies*, *11*(2), 111-118. https://doi.org/10.17507/tpls.1102.01

Susanto, A. (2017). The teaching of vocabulary: A perspective. *Jurnal Kata*, *1*(2), 183. https://doi.org/10.22216/JK.V1I2.2136

Thao, L. T., Hai Yen, T. T., Anh Thi, N., Hoang Yen, P., Trut Thuy, P., Anh Thu, H. T., & Huong Tra, N. (2023). The impact of gender on English learning approaches for outgoing learners: Unveiling intriguing insights. *Cogent Arts & Humanities*, *10*(2), 2286737. https://doi.org/10.1080/23311983.2023.2286737

Tomlinson, C. A. (2017). Differentiated instruction. In *Fundamentals of gifted education* (pp. 279-292). Routledge.

Vnucko, G., & Klimova, B. (2023). Exploring the potential of digital game-based vocabulary learning: A systematic review. *Systems*, *11*(2), 57. https://doi.org/10.3390/systems11020057

Wang, X., & Zhang, W. (2021). Psychological anxiety of college Students' foreign language learning in online course. *Frontiers in Psychology*, *12*, 598992. https://doi.org/10.3389/fpsyg.2021.598992

Wichaidit, S., & Wichaidit, P. R. (2024). Beyond play: The interplay of analogy and enjoyment in game-based learning. *Journal of Pedagogical Research*, *8*(2), 276-295. https://doi.org/10.33902/JPR.202425578

Wulantari, N. P., Rachman, A., Sari, M. N., Uktolseja, L. J., & Rofi'i, A. (2023). The Role Of Gamification In English Language Teaching: A Literature. *Journal on Education*, *6*(01), 2847-2856.

Yang, K.-H., & Chen, H.-H. (2023). What increases learning retention: Employing the prediction-observation-explanation learning strategy in digital game-based learning. *Interactive Learning Environments*, *31*(6), 3898-3913. https://doi.org/10.1080/10494820.2021.1944219

Zuhaib, A., Katta, A., Prathima, T., & Prasuna, P. (2024). Level Up: Exploring the Dynamics of Student Achievement through Game-Based Learning. *Grenze International Journal of Engineering & Technology (GIJET)*, *10*.

結合翻轉教學與遊戲式學習提升中學生人工智慧素養

Integrating Flipped Learning and Game-Based Learning to Enhance Junior High School Students' AI Literacy

黃詩樺[1]，許庭嘉[2*]
[1,2]臺灣師範大學，科技應用與人力資源發展學系
[*] ckhsu@ntnu.edu.tw

【摘要】隨著人工智慧技術在生活中日益普及，學生需具備理解、應用與評估人工智慧技術與倫理道德之能力，然而現行國中缺乏人工智慧素養教育相關課程，學生常因抽象知識難以理解而降低學習動機，因此本研究設計結合翻轉教學與遊戲式學習之教學活動，引導學生透過課前影片學習人工智慧核心概念，課堂中透過桌上型遊戲具象化知識並搭配生成式 AI 實作與引導反思。本研究採準實驗設計，針對兩組國中學生進行前後測，分析其在人工智慧素養學習成效與態度之變化，期望提供中學人工智慧教育之實證依據與教學設計參考。
【關鍵字】翻轉教學；遊戲式學習；人工智慧素養

Abstract: *As AI technology becomes increasingly prevalent, students must develop skills to understand, apply, and evaluate AI and its ethics. However, current junior high curricula lack AI literacy education, and students often struggle with abstract concepts, leading to low motivation and limited engagement. This study designed a teaching approach combining flipped learning and game-based learning. Students learn AI fundamentals through pre-class videos and engage in tabletop games, generative AI tasks, and reflective activities in class. Using a quasi-experimental design, the study conducted pre- and post-tests on two student groups to analyze changes in AI literacy and attitudes, aiming to provide valuable empirical support and instructional references for AI education in secondary schools.*
Keywords: Flipped Learning, Game-Based Learning, AI Literacy

1. 前言

近年來，科技發展快速人工智慧從日常生活到工作都影響著我們的生活，人工智慧技術已經成為全球趨勢，面對未來社會對人工智慧的高度依賴，培養學生具備基本的人工智慧素養（AI Literacy），已成為教育現場的重要課題。學生除了需理解人工智慧的基本概念與應用，更應具備辨識、評估人工智慧系統的能力，並能以倫理觀點思考人工智慧技術帶來的社會影響（Long & Magerko, 2020），人工智慧知識涵蓋機器學習、演算法等對於國中生而言艱深且抽象的概念，若缺乏適切的學習策略，學生容易感到挫折而影響學習成效與動機，傳統講述式教學常難以引起學生的學習動機與參與度，因此，如何透過創新教學模式有效提升學生的人工智慧素養，成為當前教學改革的挑戰。

為提升學習效率，教學方式需從傳統講授轉向更具互動性與可理解性的策略，翻轉教學強調課前自主學習與課堂深化應用，有助於學生依自身步調掌握基礎知識，而遊戲式學習則透過設計具挑戰性與趣味性的學習情境，將原本艱澀抽象的知識具象化、簡單化，使學生能在輕鬆的氛圍中理解關鍵概念，進而提升學習興趣與投入（Ng et al., 2024），兩者結合可促

進學生在課前有效吸收人工智慧基礎知識，課堂中透過遊戲更能在情境中強化吸收並降低認知負荷，提升整體學習成效與動機，尤其在國中階段，學生正處於自主學習能力與對於資訊理解、批判性思考的發展關鍵期，採用多元互動的教學設計，有助於學生理解抽象 AI 概念並提升學習參與度。

然而，目前國中課程中對於人工智慧相關知識的融入仍相當有限，多數學校在課程設計上仍偏重基礎資訊科技與程式語言教學，對於機器學習、演算法原理與生成式 AI 的應用等較高層次內容缺乏系統性安排，使學生難以具備全面的人工智慧素養（Meng, 2021），師資與教學資源不足亦限制了人工智慧教育的發展，導致教師在教授人工智慧時缺乏教材與實務引導，學生也因此缺乏理解與應用人工智慧技術的實際機會，針對此現況，有學者指出應積極引導學生在真實情境中探索人工智慧技術的應用，並透過具體學習活動強化其概念理解與倫理道德，以應對生成式 AI 崛起帶來對於資訊的批判性思考與道德挑戰（Lee et al., 2021）。因此，若能設計結合生活經驗、具互動性與實作的教學模式，有助於學生理解抽象的人工智慧知識，並建立批判性思考與問題解決能力，翻轉教學與遊戲式學習的結合不僅能補足傳統教學的不足，更能透過遊戲具象化學習內容、降低認知負荷，進而提升學生的學習動機與參與度，為人工智慧素養教育帶來新的契機。

基於上述的背景及動機，可以了解到人工智慧素養在未來科技化的世代中，是一項不可或缺的能力，學生不僅要懂得如何使用人工智慧外，更需要具備批判性思考與解決問題的能力。然而，目前對於國中生透過結合翻轉教學與遊戲化學習進行人工智慧素養學習成效的研究較為不足，因此，本研究從國中生的生活角度為出發點，設計一系列的教學活動，活動包含帶領學生透過自行觀看相關知識點影片，以及透過桌上型遊戲具體化抽象的知識，了解人工智慧其是如何進行大型語言模型訓練以及培養學生如何正確的向生成式 AI 下提示詞，最後再透過實作活動讓學生更深入的理解其運作方式，並且進行活動總反思，更進一步提升學生人工智慧素養，欲用實驗設計的方法進行研究，探討結合翻轉教學與遊戲式學習對學生人工智慧素養之學習成效與態度的影響，並聚焦於以下研究問題：

一、透過翻轉教學結合遊戲式學習是否會提升中學生人工智慧素養學習成效及態度？

二、翻轉教學結合遊戲式學習比傳統教學是否會提升中學生人工智慧素養學習成效及態度？

2. 文獻探討

2.1. 翻轉教學

翻轉教學是一種以學生為中心的教學方法，將傳統教學模式重新設計的學習方式，學生在課前透過數教學影片自主學習學科內容，課堂時間則用於互動、討論和實作活動（Strelan et al., 2020），其核心理念是「課前理論學習，課堂強化應用」，強調學生的主動學習與參與度。根據 Zainuddin 等人（2019）的研究，翻轉教學具備以下特徵：(1)課前學習，學生透過觀看影片教材自主學習吸收新知識；(2)課堂活動，透過小組合作、案例分析及問題解決來深化知識；(3)教師角色轉變，由傳統知識傳授者變為學習引導者，引導學生探索與應用知識。Hwang 等人(2015)表示，傳統的教學方法往往忽略學生的自主性，翻轉教學法可以讓學習者在課堂外進行知識、理解比較低層次的學習，在課堂中進行高層次的思考。

根據 Strelan 等人（2020）的研究顯示，翻轉教學使用於中學階段會相較於其他教育階段更有學習成效，由於中學生的學習能力與自主性剛好處於最佳發展階段，且中學課程內容既需要概念理解，又需要大量練習與應用，翻轉教學能提供更多課堂互動與問題解決機會，

強化學習效果，尤其是若學科涉及大量的概念理解、案例分析、問題解決與應用，翻轉教學能夠提供更靈活的學習方式，提升學習成效。

近年來，也有許多學者開始探討翻轉教學與遊戲式學習的結合運用以提升學生學習興趣與參與度（Atanasova & Molnar, 2024），Tang（2023）研究探討遊戲式學習如何補足翻轉教學學生因為自主學習壓力較大而產生焦慮感，透過遊戲的互動性、回饋與挑戰機制，能有效減少學習焦慮，提高學生的學習投入度與體驗。

2.2. 遊戲式學習

遊戲式學習是一種結合遊戲元素與教育內容的學習方法，透過遊戲機制提升學生的學習動機與學習成效（Lei et al., 2022），其核心理念是「寓教於樂」，透過挑戰、即時回饋與互動機制，使學生能夠在沉浸式環境中學習知識（López-Fernández et al., 2021） 。根據 Lei 等人 2020 的研究發現，遊戲化學習具有以下特徵：(1)目標導向學習：遊戲內設計明確的學習目標，讓學生在玩遊戲的過程中獲取知識，確保知識內化與應用。(2)沉浸式學習體驗：透過故事情境與互動模擬，使學生更容易進入學習狀態。(3)遊戲互動性：提供即時回饋與挑戰的機制（例如：分數、徽章、關卡進度）增強學生學習投入度。

根據 Mikrouli 等人（2024）的研究，遊戲式學習在中學教育階段有顯著的學習成效，特別是在提升學生的學習動機與概念理解方面，該研究發現遊戲式學習能夠有效增強學生對於複雜概念的掌握，並促進知識的實際應用，尤其是在科學、數學與技術教育領域中，相較於傳統教學方式，遊戲式學習能夠透過互動機制與回饋，幫助學生更有效理解與應用知識，提升學習成效。

近年來，遊戲式學習的應用不僅提升了學生的學習興趣與學習成效，也開始與人工智慧教育相結合，以培養學生的人工智慧素養與數位能力，根據 Zhan 等人（2022）的研究，遊戲式學習已被廣泛應用於 AI 教育，並發展出多種遊戲類型，例如：解謎遊戲、推理策略遊戲、角色扮演遊戲和模擬遊戲等，幫助學生在互動式學習環境中理解人工智慧概念。

2.3. 人工智慧素養

人工智慧素養可以定義為一系列的能力，可以使個人能夠批判性地評估人工智慧技術並且與人工智慧有效溝通和合作並在網路上、家庭和工作場所使用人工智慧作為工具（Long & Magerko, 2020），隨著人工智慧在日常生活與教育領域的應用日益廣泛，培養人工智慧素養已成為現代教育的重要課題。Ng 等人（2021）將人工智慧素養劃分為：(1)認識與理解：學習 AI 基本概念與運作方式。(2)應用：將 AI 技術運用於不同情境。(3)評估與創造：批判性思考 AI 影響，並學習與 AI 合作。(4)AI 倫理：理解 AI 的公平性、透明度與社會影響。

根據 Yi（2021）的研究，AI 素養的培養應從學校教育開始，特別是在 K-12 階段，透過遊戲式學習、案例分析與專題導向學習，使學生在互動中理解 AI 概念，透過這些方式學生不僅能提升對於 AI 技術理解能力，還能培養批判思考，使其更好應對未來的社會挑戰。

近年來，學者們開始探討人工智慧素養如何影響學習成效與社會發展，Ng 等人（2021）的研究指出，人工智慧素養的教學應結合實作與倫理，幫助學生理解 AI 技術的應用與潛在風險，此外，人工智慧素養不僅應限於學生，還應推廣至一般社會大眾，以確保不同年齡層都能具備 AI 基礎知識與應用能力（Kasinidou, 2023）。人工智慧素養的發展對於數位時代的學習者至關重要，未來教育應積極融入人工智慧素養相關課程，並強調技術應用與道德判斷，使每個人能夠在 AI 主導的環境中靈活應對與創新。

3. 研究方法

3.1. 課程與教材設計

本研究透過自行研發的桌上型遊戲作為遊戲式學習教材，引導學生認識生成式 AI 的基本應用概念，進而提升學生的人工智慧素養。理論課程內容採用自編教材，涵蓋人工智慧、機器學習及人工智慧素養等主題，並於遊戲過程中鼓勵學生以小組形式進行合作討論，透過使用電子設備查詢資訊來解決問題，此外，課程也結合 ChatGPT、Leonardo AI 等生成式 AI 工具，協助學生從實作中體驗與理解生成式 AI 的運作原理，即使第一次接觸的學生都可以從 0 開始學習體驗。

3.1.1.遊戲式學習教材設計

本研究所使用的遊戲式學習教材是透過自行研發的桌上型遊戲，主要是聚焦於生成式 AI 應用，以提升國中學生的人工智慧素養為課程目標，遊戲過程採用任務導向的模式，學生需要針對提供一張黑白且模糊的圖片選擇合適的提示詞（Prompt）進行解碼，模擬人類應該如何精準的下指令給生成式 AI 工具，才能獲得較為正確以及符合期待的回答。遊戲進行時，學生以小組合作方式進行，學生之間透過討論以及推測從 42 張提示詞卡中挑選出最符合圖片特徵的提示詞，若選出越符合圖片的提示詞，可以提高圖片的準確率，另外設置準確率門檻卡，透過競爭式的遊戲，可以增加學生學習生成式 AI 的投入度以及興趣。

3.1.2.翻轉教學結合遊戲式學習課程設計

本研究為半日 3 小時的課程，課程設計對象為國一至國三學生，學習目標是希望學生在課堂中培養人工智慧素養的能力，能夠學習人工智慧、機器學習以及人工智慧素養基礎概念知識，並透過遊戲化進行學習以提升學生的學習興趣。

課程總共分成四個部分，第一部分為理論課程，翻轉教學組別均以觀看教師自製的課程影片搭配學習單進行自主學習生成式 AI 相關核心概念，接著進行第二部分遊戲式學習課程，使用教師設計之桌上型遊戲以小組合作方式進行，學生會隨時向 Chat GPT 進行提問來獲得答案，並和同組同學進行討論，進行更高層次的思考，第三個部分由教師引導學生使用 Leonardo AI 文生圖軟體，透過實作體驗生成式 AI 的運作原理以及應用技巧，最後的第四部分為引導反思課程，帶領學生回顧遊戲體驗並深化學習成果，加強學生對人工智慧素養的知識掌握。

3.2. 研究對象

本研究之研究對象為國中一至三年級學生，實驗地點位於桃園市大園區某國中，教學實驗為各 3 個小時的營隊，於 2025 年 1 月進行，參與此實驗翻轉教學組共 24 名，傳統教學組共 27 名國中一至三年級學生，在實驗活動開始前，所有學生皆沒有參與過透過翻轉教學與遊戲式學習的人工智慧素養課程。

3.3. 研究工具

本研究主要實驗法與問卷調查法的方式來進行，研究測量工具為前後測測驗卷以及量表問卷，學習成就前後測卷則使用自編的人工智慧素養評量測驗卷，量表包含人工智慧素養、學習行為態度量表。

3.3.1. 人工智慧素養學習成就測驗卷

本研究使用人工智慧素養評量測驗卷分別於課程開始前、課程結束後進行施測，該測驗為研究者內容涵蓋人工智慧基礎概念、機器學習、生成式 AI 及其應用，最後將由專家審核過後訂定。此測驗卷分為單選題及配合題兩種題型，單選題有 24 題，配合題有 1 題，總共 25 題，總計 100 分。

3.3.2. 人工智慧素養態度量表

為了測量學習者的人工智慧素養態度，本研究將此量表分為人工智慧素養、學習人工智慧自我效能、學習行為、生成式 AI 接受度四大面項，量表均採用李克特五點量表進行評估，範圍從 1 分代表「非常不同意」至 5 分代表「非常同意」。

面項一為人工智慧素養，採用 Ng et al. (2022)所發展的人工智慧素養量表，量表分為兩個子面向，共計 12 題。量表中的兩個子面向分別為(1)應用人工智慧，即是否會在生活中使用到人工智慧，此子面向 Cronbach's α = 0.93。(2)理解人工智慧，即能理解人工智慧的定義及優缺點，此子面向 Cronbach's α =0.87。面項二為學習人工智慧自我效能，採用 Ng et al. (2023)所發展的量表共計 5 題，Cronbach's α =0.85。面項三為學習行為，採用 Ng et al. (2023)所發展的量表分為三個子面向，共計 11 題。量表中的三個子面向分別為(1)內在動機，即學生對人工智慧學習的興趣與相關性，此面向 Cronbach's α = 0.77。(2)行為意圖，即評估學生在未來是否有持續學習人工智慧的規劃，此面向 Cronbach's α =0.84。(3)行為投入，即評估學生在學習人工智慧時的參與度，此面向 Cronbach's α =0.83。面項四為生成式 AI 接受度，採用 Yilmaz et al. (2023)所發展的量表分為兩個子面向，共計 12 題。量表中的兩個子面向分別為(1)表現期望，即學生對於生成式 AI 在生活與學習中的效能認知，此子面向 Cronbach's α = 0.96。(2)努力期望，即學生對生成式 AI 的易用性認知，此子面向 Cronbach's α =0.96。

3.4. 教學實驗設計

本研究採用準實驗研究設計，旨在探討翻轉教學結合遊戲式學習對於國中生人工智慧素養的影響，在課程開始之前會先進行填寫人工智慧素養能力測驗卷、人工智慧素養態度量表，檢測學習者在課程學習前的能力，結束後進行知識點講解，實驗組採用翻轉教學法透過觀看影片進行學習，而控制組則是使用傳統講述式教學法進行學習，接著進行生成式 AI 應用桌上型遊戲體驗，在遊戲過程中遇到問題，使用 Chat GPT 查詢，並且和組員共同相互討論，遊戲結束後進行軟體實作活動加深知識點印象，最後進行引導反思，教師使用 4F 引導反思法提出問題，讓學生回顧遊戲過程、軟體操作中的學習，課程最後透過後測檢測學生在人工智慧素養中的學習成效並進行人工智慧素養量表。而控制組則是使用傳統講授式教學法進行知識點教學，其餘流程及上課形式均與實驗組相同，實驗流程圖如圖 1 所示。

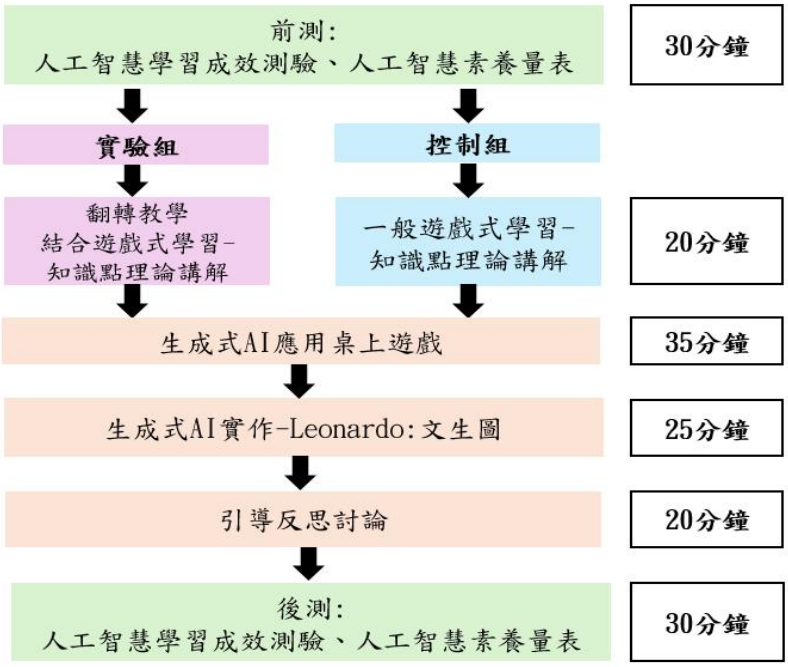


圖 1 實驗流程圖

4. 結果與討論

本研究為了解透過翻轉教學結合遊戲式學習是否可以提升學生人工智慧素養學習成效以及人工智慧素養態度，因此發展一套人工智慧素養桌遊，知識點部份透過學生自主觀看影片並撰寫學習單進行學習，遊戲式學習則是透過玩桌上型遊戲進行，最後會利用實作與反思為課堂進行知識聚焦，研究對象為國一至國三的學生，實驗組 24 名學生、控制組 27 名學生參與此實驗，由於樣本數量較少，因此檢定均使用無母數統計進行。

4.1. 人工智慧素養學習成效前後測評量檢定

使用翻轉教學結合遊戲式學習，雖然學生後測成績(M =48.92；SD =17.39)高於前測成績(M =44.08；SD =17.86)但經過無母數統計後顯示此差異未達統計上的顯著性(p=.384)，造成此結果的可能原因為，教學介入時間較短，學生對於人工智慧相關知識的吸收與理解需更長時間累積，短期內難以展現顯著學習成效。統計結果如下表 1。

表 1 人工智慧素養學習成效前後測評量檢定

	N	平均值	標準差	平均等級	z
前測	24	44.08	17.86	11.71	-.87
後測	24	48.92	17.39	10.06	

4.2. 人工智慧素養態度量表前後測檢定

使用翻轉教學結合遊戲式學習，學生後測成績(M =3.87；SD =0.84)顯著高於前測成績(M =3.35；SD =0.73)經過無母數統計後顯示具有統計上的顯著性(p=.00)，表示此教學模式對於學生人工智慧素養的學習態度具有正向影響。此結果可能來自於翻轉教學提升了學生的學習自主性，使其在課前能依自身節奏掌握知識，降低學習壓力，而遊戲式學習則透過互動性與趣味性，創造愉悅的學習氛圍，進一步提高學生對人工智慧學習的興趣與參與度（Ng et al., 2024）。統計結果如下表 2。

表 2 人工智慧素養態度量表前後測檢定

	N	平均值	標準差	平均等級	Z
前測	24	3.35	0.73	13.68	-3.71***
後測	24	3.87	0.84	4.00	

*>.001****

4.3. 實驗組與控制組人工智慧素養態度量表檢定

翻轉教學實驗組學生後測成績(M =48.92；SD =17.39)高於傳統教學控制組學生後測成績後測成績(M =40.07；SD =15.43)，但經過 Mann-Whitney 無母數統計後顯示此差異未達統計上的顯著性(p=.057)。此結果可能與學生對於新教學模式的適應期有關，翻轉教學與遊戲式學習屬於較為創新且主動參與的學習方式，部分學生在短期內尚未完全適應課前自主學習與課堂互動的節奏，導致學習態度提升幅度受限，統計結果顯示透過翻轉教學已經趨近於顯著之正向影響，若樣本數增加或教學時間延長，差異或可更為明顯。統計結果如下表 3。

表 3 實驗組與控制組人工智慧素養態度量表檢定

	N	平均值	標準差	平均等級	等級總合	U	Z
實驗組	24	48.92	17.39	30.21	725.00	223.00	-1.90
控制組	27	40.07	15.43	22.26	601.00		

4.4. 實驗組與控制組人工智慧素養態度量表檢定

翻轉教學實驗組學生後測成績(M =3.87；SD =0.84)高於傳統教學控制組學生後測成績後測成績(M =3.84；SD =0.78)，但經過無母數統計後顯示此差異未達統計上的顯著性(p=.74)。此結果可能與學生在前測時即已對人工智慧學習持正向態度有關，導致後測提升幅度有限，使兩組學生在態度量表的最終成績趨於一致，此現象可能涉及量表受限於「天花板效應」，即多數學生已接近量表分數上限，教學介入的影響難以顯現。統計結果如下表 4。

表 4 實驗組與控制組人工智慧素養態度量表檢定

	N	平均值	標準差	平均等級	等級總合	U	Z
實驗組	24	3.87	0.84	26.73	641.50	306.50	-0.33
控制組	27	3.84	0.78	25.35	684.50		

5.結論

本研究以翻轉教學結合遊戲式學習的方式，探討其對中學生人工智慧素養學習成效與人工智慧素養學習態度的影響，研究結果顯示，在學習成效方面，學生後測成績雖高於前測，且實驗組表現優於控制組，但均未達統計顯著性；在人工智慧素養態度方面，實驗組前後測有顯著提升，但與控制組相比則無顯著差異。

本研究結果與 Ng 等人（2024）研究類似，其指出遊戲式學習有助於提升學生學習動機與參與，但在短期內對於知識掌握的成效表現有限。因此，未來課程設計可延長教學時間，並增加評量頻次，以更全面掌握學習成效變化，並優化遊戲設計，提升其對學科知識內化的效果。雖量化結果未能完全支持教學成效顯著提升，但根據教學現場觀察，學生在遊戲活動中展現積極參與，能初步將人工智慧概念具象化，進而提升其對人工智慧學習的興趣與理解。在教育目標多元的背景下，學習成效應不僅侷限於成績提升，更應關注學習動機、態度與參與行為等情意層面的成長（Yu et al., 2020）。

此外，翻轉教學與遊戲式學習融合後的課程設計，對教師而言具有高度靈活性，但亦對教學策略與學生引導提出挑戰，教師需適時提供鷹架與反思引導，協助學生將學習活動轉化為知識內化，本研究亦存在一些限制，包括教學介入時間有限，學生對人工智慧知識的吸收與應用需長期累積，短期成效難以全面反映，因此建議延長實驗時間，讓學生在每個教學環節中能有充分時間進行深入學習與實作，以提升學習成效與知識內化程度；樣本集中於單一地區之國中學生，背景相對一致，結果不易推廣至不同年齡與學習情境，未來應持續探索多樣化教學模式對不同層面素養發展的影響，並發展適用於中學生的人工智慧教育資源與評量工具；未來也建議評量工具可以結合質性資料蒐集，例如：生成式 AI 工具的實際應用進行行為紀錄與反思分析，了解學生在互動過程中的學習歷程與挑戰，以深入掌握學生在翻轉教學與遊戲式學習中的認知與情意變化，進而支持學生全面性的素養養成。

致謝

本研究感謝國科會部分經費補助，計畫編號: NSTC 111-2410-H-003-168-MY3、NSTC 113-2628-H-003-002。

參考文獻

Atanasova, D., & Molnar, V. (2024). Flipped classroom, gamification and game-based learning in the traditional classroom and in non-formal education. In *2024 47th MIPRO ICT and Electronics Convention (MIPRO)* (pp. 436–440). IEEE.

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.

Kasinidou, M. (2023). AI literacy for all: A participatory approach. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 2* (pp. 628–629). https://doi.org/10.1145/3587103.3594135

Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021). Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 191–197). https://doi.org/10.1145/3408877.3432513

Lei, H., Chiu, M. M., Wang, D., Wang, C., & Xie, T. (2022). Effects of game-based learning on students' achievement in science: A meta-analysis. *Journal of Educational Computing Research, 60*(6), 1373–1398. https://doi.org/10.1177/07356331211064543

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). https://doi.org/10.1145/3313831.3376727

López-Fernández, D., Gordillo, A., Alarcón, P. P., & Tovar, E. (2021). Comparing traditional teaching and game-based learning using teacher-authored games on computer science education. *IEEE Transactions on Education, 64*(4), 367–378. https://doi.org/10.1109/TE.2021.3057849

Meng, Y. (2021). Research on AI education for primary and secondary schools in China. *Academic Journal of Psychological and Educational Processes*, *2*, 44–51. https://doi.org/10.23977/APPEP.2021.020109

Mikrouli, P., Tzafilkou, K., & Protogeros, N. (2024). Applications and learning outcomes of game-based learning in education. *International Educational Review*. https://doi.org/10.58693/ier.212

Ng, D., Xinyu, C., Leung, J., & Chu, S. (2024). Fostering students' AI literacy development through educational games: AI knowledge, affective and cognitive engagement. *Journal of Computer Assisted Learning, 40*, 2049–2064. https://doi.org/10.1111/jcal.13009

Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review, 30*, 100314. https://doi.org/10.1016/j.edurev.2020.100314

Tang, B. (2023). Combining game-based learning and flipped classroom—An analysis based on affordance theory. *Journal of Education, Humanities and Social Sciences, 22*. https://doi.org/10.54097/ehss.v22i.13350

Yi, Y. (2021). Establishing the concept of AI literacy. *JAHR*, *12*(2), 247–254. https://doi.org/10.21860/j.12.2.8

Yu, Z., Gao, M., & Wang, L. (2020). The effect of educational games on learning outcomes, student motivation, engagement and satisfaction. *Journal of Educational Computing Research, 59*(3), 522–546. https://doi.org/10.1177/0735633120969214

Zainuddin, Z., Haruna, H., Li, X., Zhang, Y., & Chu, S. K. W. (2019). A systematic review of flipped classroom empirical evidence from different fields: What are the gaps and future trends?

International Journal of Educational Technology in Higher Education, 16(1), 1–19. https://doi.org/10.1186/s41239-019-0171-0

Zhan, Z., Tong, Y., Lan, X., & Zhong, B. (2022). A systematic literature review of game-based learning in artificial intelligence education. *Interactive Learning Environments, 32*(3), 1137–1158. https://doi.org/10.1080/10494820.2022.2115077

专业成长视角下的中小学教师人工智能素养测评方案研制

Development of an Assessment Framework for K-12 Teachers' AI Literacy from the Perspective of Professional Growth

夏正仁，温州市教师教育院
林淼焱，温州科技高级中学
wzxzr@qq.com

【摘要】随着中小学人工智能教育的普及与人工智能学科的快速发展，学科教师不断面临专业成长的新需求，但却遭遇成长路径模糊与测评体系不完善等问题。本研究从学科教师的人工智能素养内涵出发，构建了“专业知识-专业能力-伦理态度”三维测评框架，旨在为学科教师提供清晰的专业成长导航图，为区域教师培训方案设计提供科学的决策依据。研究打破了传统素养测评偏重知识考核的模式，并推动教师从“技术应用者”向“教育研究者”的转型，从而为中小学人工智能学科教师的专业发展，提供了“测评-培训-实践”的一体化发展模式。

【关键词】中小学人工智能教育；学科教师；人工智能素养；教师专业成长；测评方案设计

Abstract: *With the popularization of AI education in primary and secondary schools and the rapid development of the AI discipline, subject teachers are facing new demands for professional growth. However, they often encounter challenges such as unclear growth pathways and underdeveloped assessment systems. This study, grounded in the connotation of AI literacy for subject teachers, constructs a three-dimensional evaluation framework encompassing professional knowledge, professional competence, and ethical attitudes. The goal is to provide teachers with a clear roadmap for professional development and offer a scientific basis for the design of regional teacher training programs. The study breaks away from the traditional knowledge-heavy assessment model and promotes a shift in teachers' roles from "technology users" to "educational researchers." In doing so, it proposes an integrated development model of assessment–training–practice to support the professional advancement of AI subject teachers in K-12 education.*

Keywords: K-12 AI Education，Subject Teachers，AI Literacy，Teacher Professional Development， Assessment Framework Design

1.引言

自从《新一代人工智能发展规划》提出“在中小学阶段设置人工智能相关课程”的要求以来，各地纷纷通过师资培训的方式来解决师资短缺和专业能力不足的问题。然而，随着人工智能学科的飞速发展，学科教师的专业成长难以跟上发展的步伐，成为新的挑战。教师作为人工智能教育的实施主体，是保障教学质量的关键。然而，要真正胜任这一角色，教师必须具备主动与人工智能学科同步发展的意识与能力。然而，许多教师对自身专业成长的正确路径缺乏清晰的认知。为此，可以构建一套中小学人工智能学科教师素养测评体系，为教师

提供清晰且可靠的专业能力发展“导航图”。该体系不仅能够帮助教师明确成长方向，还能引导他们从“技术学习者”向“教育研究者”顺利转型，从而全面提升中小学人工智能教育的整体质量。

2.相关研究状况

关于人工智能素养，当前还未有统一或权威的定义。学者 Burgsteiner 等（2024）指出，人工智能素养的基础是理解人工智能的基本原理与概念；而 Long 等（2024）进一步强调，是个体需具备批判性评估技术、人机协作沟通及伦理决策等能力。Ng 等（2024）提出四维度的模型：知识理解、技术应用、创新创造及伦理责任，凸显素养的全面性。国内学者则大多强调个体在智能时代需具备理解、应用和批判性评估人工智能综合能力。因此人工智能素养可以认为是个体应该具备的理解人工智能基本原理所需的专业知识、批判性地应用人工智能解决实践问题的专业能力及在解决问题过程中必须要遵循的伦理态度。

关于教师的人工智能素养，则是人工智能素养在教育领域的特殊延伸，其核心是为了促进教育技术与育人目标的有机融合。根据 2022 年教育部发布的《教师数字素养》标准，认为是教师必须掌握人工智能在教学中的应用策略，包括教育教学资源开发、数据分析及跨学科学习活动设计能力等。很多学者还对其做进一步的解构：一是对人工智能的认知，涵盖了对人工智能原理与教育应用场景的理解；二是解决教育教学实践问题的能力，包括人工智能工具的使用、项目化教学设计及学情分析等能力；三是指向更为高阶的思维，强调批判性思维与创新意识；四是应该遵守的伦理责任，主要涉及数据隐私保护、算法偏见规避及技术伦理教育等。

关于人工智能学科教师的人工智能素养，则鲜有学者进行研究论述，但可以认为是学科教师在人工智能学科领域及其教育教学活动中的专业化延伸。如关于人工智能学科及其教育教学的知识，关于人工智能学科教育教学的实践创新能力，在人工智能学科或其教育教学活动中必须要遵守的伦理道德意识等方面。

关于人工智能素养测评工具及其框架，国内外相关研究大多处于发展的起步阶段，可借鉴参考程度普遍不高。由于还存在“教师人工智能素养”“教师数字素养”“教师智能素养”等概念上的纠缠，导致测评目标定位大多较为模糊，并有权威性缺乏、信效度不足与内容覆盖不全等方面问题。如测评指标侧重人工智能领域的知识技能，对运用人工智能解决实践问题的专业能力及其所表现出伦理态度关注不够等问题。另外还缺乏基于教师专业发展阶段的差异化测评标准，难以满足学科教师对个性化的专业发展需求。

3.人工智能学科教师人工智能素养测评方案的设计

（1）测评目标

建立科学的人工智能学科教师人工智能素养发展坐标系，以确保实现对专业知识、专业能力与伦理态度的发展水平进行测评。通过测评数据的收集与分析，能够提供“学科教师人工智能测评分析报告”，从而为教师的自身专业发展规划路线及师训部门的师资专题培训提供科学的决策依据。

（2）测评原则

科学性原则。基于《义务教育信息科技课程标准（2022 版）》《普通高中信息技术课程标准（2017 版）》的学科教育教学理念及人工智能学科发展的趋势，构建出能够满足人工智能学科教育教学需求的“专业知识-专业能力-伦理态度”三维模型测评框架。

实践性原则。聚焦于人工智能实验、项目化学习等典型的教育教学活动场景，通过挑战真实的教育教学任务，来检测学科教师对专业知识的理解、在解决真实复杂问题过程中灵活运用程度及其所表现的伦理态度。

发展性原则。基于“成长档案袋”理念，长期追踪学科教师人工智能素养提升的轨迹，进而实现为促进学科教师专业发展画像、规划成长路径并提供个性化的发展资源。

（3）测评框架的构建

根据学科教师在发展“专业知识”“专业能力”“伦理态度”三个领域的一般性学习投入及人工智能学科自身发展对教师专业发展的引领性需求，构建了如下表 1 所示的各项指标与权重。

表 1 人工智能学科教师人工智能素养测评框架

维度	指标	考查	权重
专业知识	AI 学科认知	机器学习、数据科学等 AI 基础概念掌握程度	25%
	教育教学认知	实验教学、项目化学习及 AI 与学科融合教学法理解	20%
专业能力	AI 实验能力	各种验证机器学习的学科实验	20%
	AI 项目开发能力	AI 课程设计、跨学科项目等案例开发	15%
	AI 教育教学环境搭建	教学平台、实验教学环境及配套学习资源开发等	10%
伦理态度	技术伦理意识	算法偏见识别、数据隐私保护等	5%
	师德表现	AI 育人价值认知、技术风险规避等	5%

（4）测评工具的编制

专业知识维度的试题编制。在力求全面覆盖 AI 学科和教育教学知识，并遵循先易后难的难度梯度分配基础上，设置了 30 道以选择题与填空题为题型的测试题（其中包含 20%的情境化题目）。

专业能力维度的试题编制。设置 3 个情境任务，以问答题为主，内容囊括 AI 实验、AI 项目开发与 AI 教育教学环境搭建这三个领域。在评分标准方面，要求包含规范性（40%）、创新性（30%）与可行性（30%）这三个方面。

伦理态度维度的试题编制。融入专业能力维度的试题，以选择题形式出现并采用 Likert5 级量表并结合案例情境进行分析，以区分测试者采用的不同伦理态度标准。

4.测评方案的实施与分析报告的使用

（1）测评方案的实施

主要由承当组织培训任务的师训部门相关工作人员负责。先组织专家根据“专业知识-专业能力-伦理态度”这个测评框架进行命题，并确保命题过程的专业性与试题质量的可靠性；再在规划与实施培训活动前，组织区域内的中小学人工智能学科教师进行有效度的统一测评；最后将测试的结果交给专家进行数据分析并撰写相应的分析报告。

（2）分析报告的使用

分析报告分为个人分析报告与区域分析报告这两部分。个人分析报告，可根据需求直接提供给学员，以帮助其确认自身的专长短板并助力其实现专业成长。区域分析报告，则统一

交给各县市区的师训员，以为其制定区域的教师培训计划、确认每期的培训主题与内容、筛选适宜的学员及设计培训活动方式等提供科学的依据。

5.结束语

本研究构造了“专业知识-专业能力-伦理态度”三维测评体系，并实现从“知识考核”到“专业发展”的范式转变。未来还需通过纵向追踪，来研究验证方案的长期效度，并积极探索“测评-培训-实践”的一体化教师发展模式，从而为中小学人工智能学科教师的专业成长提供可持续性地支持。

参考文献

[1]Burgsteiner,C.,Doe,J.,& Smith,A.(2024).AI Literacy: A Framework for Understanding and Teaching Artificial Intelligence. Journal of Educational Technology,53(2),45–62.
[2]Long, Y.,Chen, L.,& Wang,M.(2024).Defining AI Literacy: A Systematic Review. Computers & Education,198,104682.
[3]Ng,W.H.,Lee,S.,& Zhang, L.(2024). AI Literacy for the Future: A Four-Perspective Model. International Journal of Human-Computer Studies,185,103124.
[4]中华人民共和国教育部.(2022). 教师数字素养（JY/T 0646-2022）[S]. 北京：北京师范大学出版社.
[5]Jenkins,A.,&Smith, B.(2023,April). AI Literacy Framework for Educators.In Proceedings of the CHI Conference on Human Factors in Computing Systems(pp. 1–12).Hamburg,Germany.
[6]王晓明，李华.(2024).教师人工智能素养测评工具本土化研究[J]. 中国电化教育，(6), 78–85.
[7]华东师范大学团队.(2024).教师人工智能素养发展报告[R].上海：华东师范大学出版社.

具身认知视角下 AI 机器人编程课程的创新设计与课堂实践

Innovative Design and Classroom Practice of AI Robot Programming Curriculum under Embodied Cognition Perspective

张翼菲[1*]，陈菱琦[2]，赵嘉华[3]
福建师范大学 教育学院， 福建福州 350000
[*]2297796905@qq.com

【摘要】 在人工智能时代，编程教育成为培养信息素养关键，但传统编程教学方式对低年级学生适配性不足。研究基于具身认知理论，采用准实验法对比机器人辅助与纸笔工具教学效果。结果显示，机器人支持的高具身教学显著提升学生学习成绩与批判反思能力，问题解决能力虽未达统计显著但呈积极趋势。研究表明，机器人具身化操作与即时反馈帮助学生直观理解编程逻辑，形成“动作-思维”深度联结。机器人作为具身认知实践载体，不仅降低认知负荷，更拓展理论在编程教育应用，为课程设计提供实证依据，建议未来延长实验周期深化验证。

【关键词】 人工智能；机器人；具身认知理论；编程课程

Abstract: *In the era of artificial intelligence, programming education plays a critical role in fostering digital literacy, yet traditional methods often lack suitability for younger students. Grounded in embodied cognition theory, this quasi-experimental study compares the effects of robot-assisted versus paper-based instruction. Findings reveal that high-embodiment teaching supported by robots significantly enhances students' academic performance and critical reflection skills, with problem-solving showing a positive trend. Robot interaction and real-time feedback promote intuitive understanding of programming logic through deep action-thought integration. Serving as a vehicle for embodied cognition, robots reduce cognitive load and broaden the application of theory in programming education. The study offers empirical evidence for curriculum design and calls for extended experimental durations for further validation.*

Keywords: Artificial Intelligence,Robotics, Embodied Cognition Theory,Programming Courses

1.绪论

1.1 研究背景

随着人工智能技术的迅猛发展，其在各个领域的广泛应用正深刻地改变着我们的生活和工作方式。教育领域也不例外，人工智能教育应运而生，成为全球基础教育改革的重要议题（Cheng & Wang, 2023）。自 2017 年我国《新一代人工智能发展规划》颁布以来，中小学编程教育逐步纳入课程体系，旨在通过编程学习培养学生的计算思维与创新能力，为未来的人工智能时代打下坚实基础（康建朝, 2021）。然而，传统编程教学多聚焦于屏幕端的代码编写与语法规则训练(孙立会, 2019)，对低年级学生的认知适配性存在显著局限。研究表明，处于具体运算阶段的小学生仍以直观思维为主导，其抽象逻辑能力尚未完善（李水明, 2014），面对传统编程教学中复杂的符号系统与离屏式学习模式，易产生理解障碍（王靖等, 2022）。

这种“从文字到文字”的教学范式，不仅割裂了知识的情境关联，更因过度依赖机械记忆导致学生思维僵化，难以实现编程思维的内化。

具身认知理论为破解这一困境提供了新思路。该理论强调认知是身体与环境动态交互的产物，主张通过身体动作、感知体验与物理操作促进抽象概念的理解(姜美玲, 2003)。这一理论与编程教育的融合，能够将抽象的算法逻辑转化为具象的物理行为，如通过肢体动作模拟程序流程或操控实体对象完成指令序列。在此背景下，“非计算机化”编程教育模式逐渐兴起，其通过有形技术（Tangible Technology）降低认知负荷，使学生在无屏环境中通过卡片、积木等工具建构编程思维。机器人作为具身认知的实践载体，进一步凸显了其介入的必要性。机器人通过可触控的硬件交互与即时反馈机制，能够将编程任务具象化为“动作-响应”的闭环系统。例如，学生通过指挥机器人完成迷宫导航任务(Gross & Veitch, 2013)，可直观观察指令序列与运动轨迹的因果关系，从而深化对循环、条件判断等逻辑的理解。研究证实，机器人辅助的编程活动不仅能提升学生的课堂参与度，还可通过跨学科项目实践促进计算思维与协作能力的协同发展(周明, 2017)。

综上，传统编程教育在低龄段的适配性不足与具身认知理论的应用潜力，共同指向机器人介入编程课程的必要性。本研究以具身认知理论为框架，探索机器人辅助的编程教学模式如何通过身体参与情境化任务设计，破解小学生编程学习中的抽象认知壁垒，为其思维进阶与素养发展提供实践路径。通过这一研究，我们期望为人工智能教育在低年级阶段的实施提供理论支持和实践指导，推动人工智能教育的普及与发展。基于此，本文提出三个研究问题：

（1）在编程课程中进行机器人支持下的具身认知教学对小学生的学习成绩有何影响？

（2）在编程课程中进行机器人支持下的具身认知教学对小学生问题解决能力有何影响？

（3）在编程课程中进行机器人支持下的具身认知教学对小学生批判反思能力有何影响？

1.2 研究目的与意义

本研究创新性地提出“具身认知理论指导下的机器人辅助编程课程”方案，基于具身认知理论，设计并实施机器人辅助教学，通过准实验方法对比机器人辅助学习（高具身认知组）与纸笔工具学习（低具身认知组）两种方式，检验其对学生学习效果的影响。研究将学习效果划分为学习成绩、问题解决能力与批判反思能力三个维度，通过实证研究具体呈现课程成效，并为优化小学编程教学提供依据。

本研究融合具身认知理论与机器人教育实践，构建“理论指导—实践应用—效果评估”模型，帮助学生突破编程学习中的抽象认知障碍。在理论层面，研究拓展了具身认知理论在编程教育中的应用边界，丰富相关理论成果，揭示具身认知与编程教学融合的内在机制，为后续研究提供理论支持。在实践层面，研究为编程教师提供新思路与方法，提升教学效率与质量，激发学生学习兴趣，培养创新精神与实践能力，增强信息素养，为其后续学习发展奠定基础。

2.概念界定及理论基础

2.1 概念界定

2.1.1 编程教育

编程教育是通过学习编程语言和技术，培养学生的计算思维、逻辑推理能力以及问题解决能力的一种教育形式(孙丹 & 李艳, 2019)。编程教育不仅仅是学习编程语言本身，而是通过编程活动训练儿童的思维方式，在提升理论思维水平的同时，构建出可迁移至日常生活的哲学方法论体系(孙立会, 2019)。这种从“技术工具”到“思维载体”的认知转变，恰与当下

教育理念革新形成共振。随着教育理念的深化，编程教育逐渐从单纯的技术教学向综合素质培养转变，强调通过编程学习提升学生的思维能力和解决实际问题的能力(姜美玲, 2003)。

2.1.2 批判反思能力

批判反思能力是指个体对自己或他人的观点、行为、决策等进行批判性思考和分析的能力（黄飞莺 & 周志毅, 2003）。在编程教育中，批判反思能力的培养尤为重要。批判反思能力体现为学生能够对自己编写的程序进行反思和评价，发现程序中的不足之处，并提出改进方案；同时，也能够对他人的编程思路和代码进行分析和借鉴，不断优化自己的编程方法和技能。通过编程教育，学生可以学会如何分析问题、设计解决方案、调试程序以及优化算法，这些过程都需要批判性思维的支持(周明, 2017)。

2.1.3 问题解决能力

问题解决能力是指个体运用认知过程和技能，通过一系列的思维活动，面对和解决实际情境中的问题的能力（赖小琴, 2009）。它强调的是在面对复杂情境时，能够综合运用多种知识和技能来解决问题的能力(袁维新, 2011)。在编程学习中，问题解决能力表现为学生能够理解编程任务的要求，运用编程知识和算法设计合理的程序逻辑，解决程序运行过程中出现的错误和问题，实现编程目标。

2.2 理论基础

2.2.1 具身认知理论

具身认知理论强调认知是身体与环境相互作用的结果(王靖等, 2022)，其核心特征在于：认知过程的实现方式和步骤由身体的物理属性决定，认知内容由身体提供，认知与身体、环境构成一体化系统（叶浩生, 2010）。该理论认为认知不仅是对大脑抽象符号的加工，更是通过身体动作、感知与环境互动共同建构的(叶浩生, 2014)。此外，具身认知还突出认知过程的动态性和情境性，强调认知活动对具体情境的依赖性和身体动态交互的不可或缺性(李志河等, 2018)。基于这种身心交融的认知机制，学习者通过具身参与——身体的动作、体验和操作，能够更深刻地理解和掌握抽象知识（孙立会, 2019）。

2.2.2 建构主义视角下的 PBL 教学模式

建构主义理论主张知识是个体在与环境互动中主动建构的产物，而非被动接受的信息（何克抗, 1997）。该理论源于皮亚杰的认知发展观和维果茨基的社会文化理论，它认为学习是一个社会性、情境性的过程，强调学习者通过与他人的互动和合作来构建知识。此外，建构主义还注重学习者的先前经验和认知结构，认为新知识的建构需要在已有经验的基础上进行，通过认知冲突和问题解决来实现（马淑风 & 杨向东, 2022）。教学实践中，建构主义理论提倡以学生为中心的教学模式，强调学生的主动参与和探究活动。教师应设计富有挑战性的任务，创设情境，激发学生的学习兴趣和动机，同时提供必要的支持和引导（何克抗, 1997）。

建构主义理论所强调的主动建构与情境认知特征，在问题式教学（PBL）模式中得到了具象化呈现，这种以问题驱动为核心的教学模式，本质上是建构主义理念在教学实践中的创新投射。问题式教学（Problem-Based Learning, PBL）是一种以问题为中心的教学模式，其核心在于通过真实、复杂的问题情境引导学生主动学习和建构知识(汤丰林 & 申继亮, 2005)。在该模式下，学生围绕真实情境中的问题展开学习，通过自主探究、合作学习等方式，运用所学知识解决问题(姜美玲, 2003)。PBL 教学模式强调学生的主体地位，注重培养学生的问题解决能力、自主学习能力和团队合作精神(陈丽虹等, 2013)。追溯理论根源，Barrows 模型作为 PBL 原始框架，提出"提出问题-自学解疑-重点讲授-总结归纳"四步流程（张帅 & 魏晓

燕, 2017）。后续研究进一步深化实践，将经典 PBL 细化为"问题提出、建立假设、资料收集、论证假设、总结反思"五阶段操作模型（车艳玲等, 2012），为教学实践提供了更具操作性的实施路径。

3.研究现状

当前小学生编程教育研究首要挑战在于学生具象思维与抽象编程逻辑之间的认知鸿沟（李庆武, 2022）。研究表明，小学生因受限于知识储备和思维发展阶段，难以直接理解编程语言及算法等抽象概念（辛跃武, 2020），传统教学中理论灌输与实践操作的割裂进一步加剧了这一矛盾。具身认知理论的兴起为破解该困境提供新视角，该理论突破传统认知科学框架，强调身体与环境互动对知识建构的决定性作用（孟洁 & 孙晶华, 2023）。教育研究者据此提出具身学习概念，通过生成式、情境式等学习范式实现认知与身体的深度耦合（杨子舟等, 2017），但现有研究多滞留哲学阐释层面，缺乏实证支撑（叶浩生, 2010;贾丽娜等，2016）。

机器人编程教育的实践探索则架起理论与应用之间的桥梁，Resnick 等人开发的 Scratch 编程语言与乐高机器人结合，通过可视化编程控制机器人动作，有效降低了编程门槛，提升学生的逻辑思维与动手能力（孙立会 & 周丹华, 2019）。Bers 等人提出“编程即游戏”理念，利用 Kibo 机器人引导低龄儿童通过物理模块拼接完成编程任务，研究表明，具身化的机器人操作显著增强学生对抽象编程概念的理解（沈慧敏 & 程秀兰, 2023）。深圳部分小学引入 Dash & Dot 机器人开展编程教学，学生通过编程指令操控机器人完成迷宫挑战，结果显示学生的算法设计能力显著提升（孙秀芝, 2019）。值得注意的是，尽管机器人编程教育已展现独特优势，但现有研究仍存在三大局限：具身认知机制与编程学习的关联模型尚未建立（周冠男 & 张栋科, 2023;丁继红等, 2024）,教学效果的跨学科评估体系缺失(钟伯昌, 2016)，以及低龄段长期干预研究近乎空白(沈慧敏 & 程秀兰, 2023;朱月姣 & 杨力, 2022;王学男等， 2019)。本研究拟基于贝芽机器人编程课程，构建“具身操作-认知发展”的动态映射模型，为深化该领域研究提供实证依据。

4.基于编程的人机交互场景下的教学模式

4.1 研究设计

本研究聚焦机器人在小学一年级编程课程中的应用效果，采用准实验法探究其对学生学习成绩、问题解决能力与批判反思能力的影响。结合当前小学编程教学困境，依据具身认知理论，设计高具身认知组（机器人辅助教学）与低具身认知组（纸笔工具教学）进行对比实验。实验以是否使用机器人为自变量，学习成绩、问题解决与批判反思能力为因变量。实验选取一年级两个班，共 40 人，分别实施一课时机器人辅助《初识编程》课程与传统教学模式，具体研究设计思路如图 1 所示。通过前测与后测的问卷调查（评估问题解决与批判反思倾向）及编程测试（考察知识掌握程度），运用统计软件分析两组差异。研究旨在验证机器人支持的具身认知对提升小学生编程学习效果的假设，为优化编程教学提供实证依据。

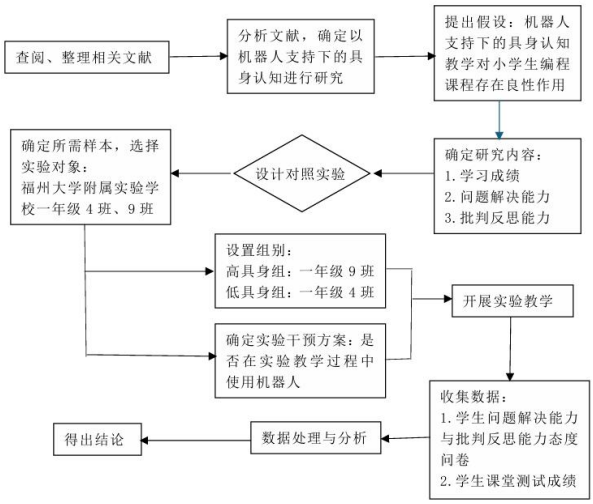


图 1 研究设计线路图

4.2 研究方法

本研究采用文献研究法、问卷调查及准实验法。首先，通过查阅国内外权威文献，梳理具身认知理论与编程教育的研究现状，构建理论基础。其次，设计问卷调查工具，包括学习成效测试及问题解决(Lai & Hwang, 2014)、批判反思倾向问卷（Chai et al., 2015;Lin et al., 2019），均采用李克特五点量表，信度分别为 0.78、0.71。实验前后分别对实验组与对照组进行测试与问卷发放，当场回收确保数据有效性。准实验法方面，以是否使用机器人辅助教学为干预变量，开展教学实验，收集测试成绩、问卷数据及课堂表现记录，通过量化分析探究机器人支持的具身认知对小学生编程学习的影响。

4.3 教学设计

为课程设计框架。教学过程分为五个阶段：首先通过动画故事创设“小贝”赴城堡舞会的任务情境，高具身组结合机器人与实物棋盘地图形成直观认知，低具身组则借助多媒体与 PPT 地图进行想象；随后分组分析路径与编程需求，高具身组扫描指令卡片观察机器人动作映射，低具身组通过讲解建立抽象认知；在实践探究中，高具身组使用实物拼图规划路径并调试机器人，体验即时反馈，低具身组在图纸上标注指令进行模拟；成果展示时，高具身组以机器人演示路线，低具身组用 PPT 讲解设计；最后教师从准确性、效率等维度评价，引导学生互评反思，高具身组强化具身转化，低具身组巩固抽象认知，具体流程如表 1 所示。

表 1 教学过程内容设计表

教学过程	高具身组	低具身组
1.创设情境，提出问题	使用机器人“小贝”播放动画+6×6 实物棋盘地图，建立直观联系	多媒体播放动画+PPT 呈现地图，依赖画面想象
2.组成小组，分析问题	观察实物地图+动作指令卡片扫码录入演示（动作指令模块与机器人动作映射）	观察纸质地图+教师讲解动作指令含义（抽象认知动作指令）
3.实践探究，解决问题	（1）实物拼图规划路线	（1）图纸画路线

	（2）编程记录板记录指令	（2）编程记录板记录指令
	（3）指令录入与即时调试	（3）路线标注指令与反复验证
4.展示成果，分享经验	（1）机器人运行演示+指令同步讲解	（1）PPT 画笔功能展示路线设计+指令同步讲解
	（2）问题解决经验分享	（2）问题解决经验分享
5.评价反馈，深化总结	（1）教师统一评价路线设计与指令程序	
	（2）学生自评、互评反思设计，提升编程认知	

5.实验结果

5.1 学习成绩

在学习成绩方面，以编程知识前测为协变量，组别为自变量，后测成绩为因变量进行协方差分析（ANCOVA）。实验前两组学生的编程知识水平无显著差异（F (1,38)=2.339，p=0.135>0.05），表明分组均衡。回归斜率同质性检验显示组别与前测交互作用不显著（F=2.127，p=0.153>0.05），满足协方差分析前提。分析结果显示，组别对后测成绩具有显著影响（F (1,36)=9.57，p=0.004<0.05），修正模型整体效应显著（F (3,36)=5.415，p=0.004）。高具身组后测平均成绩显著高于低具身组。单变量对比进一步验证组间差异的统计学意义（F=9.57，p=0.004），表明机器人支持的具身认知教学对提升学习成绩具有显著促进作用。

表 2 学习成绩协方差分析结果

组别	人数	平均值	标准误差	F	结果比较
高具身	20	84.84a	1.24	9.57**	(1)>(2)
低具身	20	79.41a	1.24		

*p <0.01

5.2 问题解决能力

在问题解决能力方面，以前测得分为协变量，组别为自变量，后测为因变量进行协方差分析（ANCOVA）。两组前测得分无显著差异（F (1,38)=2.955，p=0.094>0.05），表明基础相当，满足实验要求。回归斜率同质性检验结果显示组别与前测交互作用不显著（F=0.078，p=0.781>0.05），满足协方差分析前提。组别对后测影响未达到显著水平（F (1,36)=0.006，p=0.936>0.05），修正模型整体效应亦不显著（F (3,36)=1.518，p=0.227>0.05）。但单变量对比显示边缘显著趋势（F (1,37)=3.857，p=0.057），高具身组后测均值高于低具身组。结果表明，机器人支持的具身认知教学在问题解决能力提升方面未显著，但存在微弱正向趋势。

表 3 问题解决能力协方差分析结果

组别	人数	平均值	标准差	F	结果比较
高具身	20	4.09	0.15	3.857	(1)>(2)
低具身	20	3.64	0.15		

*p <0.05

5.3 批判反思能力

在批判反思能力方面，以前测为协变量，组别为自变量，后测为因变量进行协方差分析（ANCOVA）。两组前测得分无显著差异（F (1,38)=0.687，p=0.413>0.05），表明实验前基础相当，满足实验要求。回归斜率同质性检验显示组别与前测交互作用不显著（F=2.283，p=0.140>0.05），满足协方差分析前提。组别对后测影响接近显著（F (1,36)=3.530，p=0.068），修正模型整体效应达边缘显著（F (3,36)=2.659，p=0.063）。单变量对比显示，高具身组后测均值显著高于低具身组，差异达到统计学意义（F (1,37)=4.635，p=0.038<0.05）。结果表明，机器人支持下的具身认知教学对小学生批判反思能力提升具有显著促进作用。

表 4 批判反思能力协方差分析结果

组别	人数	平均值	标准差	F	结果比较
高具身	20	4.19	0.13	4.635*	(1)>(2)
低具身	20	3.79	0.13		

*p <0.05

6.研究结论

本研究通过准实验设计，验证了机器人支持的具身认知教学对小学生编程学习的促进作用。结果显示，高具身组在学习成绩（p=0.004）和批判反思能力（p=0.038）上表现出显著优势，契合具身认知理论强调的“动作-思维”认知路径。机器人将抽象概念具象化、提供即时反馈以及促进多模态参与，帮助学生建立“动作-指令-结果”的强关联，提升理解与记忆。高具身组在 PBL 框架下的协作编程与具身化展示，如路径演示，进一步强化了批判反思能力，形成“操作-感知-反思”的闭环。

尽管问题解决能力未达统计显著（p=0.936），高具身组后测均值略高，且边缘显著趋势（p=0.057）暗示具身教学对高阶思维的潜在促进。该结果或与实验局限有关：单课时干预难以充分培养复杂思维，问卷侧重主观态度，低具身组图纸标注亦具有一定训练效应。研究拓展具身认知理论在编程教育中的应用，并为教学实践提供操作范式：建议结合机器人工具与情境任务，运用 PBL 强化探究与反思。未来研究可延长干预周期、扩大样本，融合跨学科项目，深化对核心素养成效的验证。

参考文献

康建朝.(2021).芬兰中小学编程教育的缘起、实践路径与特征.电化教育研究,42(08),101-107+115.doi:10.13811/j.cnki.eer.2021.08.014.

孙立会.(2019).聚焦思维素养的儿童编程教育:概念、理路与目标.中国电化教育,(07),22-30.

李水明.(2014).把握学生认知规律,提高数学课堂有效性.新课程导学,(14),86.

孙丹 & 李艳.(2019).国内外青少年编程教育的发展现状、研究热点及启示——兼论智能时代我国编程教育的实施策略.远程教育杂志,37(03),47-60.doi:10.15881/j.cnki.cn33-1304/g4.2019.03.005.

姜美玲.(2003).基于问题的学习:一种可资借鉴的教学模式.全球教育展望,32(03),62-66.

黄飞莺,周志毅.(2003).师范生批判性反思能力的培养:意义·方法.浙江师范大学学报,(04),111-114.

周明.(2017).基于计算思维培养的中小学编程教育校本课程开发与实践.中小学信息技术教育,(03),61-65.
赖小琴.(2009).PISA 中问题解决能力评价的特点及启示.教育测量与评价(理论版),(05),9-11+15.doi:10.16518/j.cnki.emae.2009.05.012.
袁维新.(2011).国外关于问题解决的研究及其教学意义.心理科学,34(03),636-641.doi:10.16719/j.cnki.1671-6981.2011.03.004.
王靖,马志强,刘亚琴 & 杜鸿羽.(2022).面向计算思维的可视化编程活动设计与应用.现代教育技术,32(09),55-63.
叶浩生.(2010).具身认知:认知心理学的新取向.心理科学进展,18(05),705-710.
叶浩生.(2014).“具身”涵义的理论辨析.心理学报,46(07),1032-1042.
李志河,李鹏媛,周娜娜 & 刘芷秀.(2018).具身认知学习环境设计:特征、要素、应用及发展趋势.远程教育杂志,36(05),81-90.doi:10.15881/j.cnki.cn33-1304/g4.2018.05.011.
何克抗.(1997).建构主义的教学模式、教学方法与教学设计.北京师范大学学报(社会科学版),(05),74-81.
马淑风 & 杨向东.(2022).什么才是高阶思维？——以“新旧知识关系建立”为核心的高阶思维概念框架.华东师范大学学报(教育科学版),40(11),58-68.doi:10.16382/j.cnki.1000-5560.2022.11.005.
汤丰林,申继亮.(2005).基于问题的学习与我国的教育现实.比较教育研究,(01),73-77.
陈丽虹,周莉,吴清泉,等.PBL 教学模式效果评价及思考[J].中国远程教育,2013,(01):70-73.DOI:10.13541/j.cnki.chinade.2013.01.008.
张帅 & 魏晓燕.(2017).基于问题学习法在神经生物学教学中应用.神经解剖学杂志,33(04),508-510.doi:10.16557/j.cnki.1000-7547.2017.04.025.
车艳玲,尹志伟 & 栾金红.(2012).基于 PBL 教学法在中医院校超声影像教学中的探讨.中国中西医结合影像学杂志,10(06),570-571.
李庆武.(2022).编程教育课程在小学信息技术教学中的应用.亚太教育,(12),13-15.
辛跃武.(2020).少儿编程在小学信息技术教学中的探索.新课程导学,(17),87.
孟洁 & 孙晶华.(2023).具身认知理论在教育中的应用研究综述.白城师范学院学报,37(01),101-109.
杨子舟,史雪琳 & 荀关玉.(2017).从无身走向有身:具身学习探析.教育理论与实践,37(05),3-6.
贾丽娜,田良臣,王靖,马志强 & 周倩.(2016).具身教学的设计研究——基于身体参与的多通道整合视角.远程教育杂志,34(01),82-89.doi:10.15881/j.cnki.cn33-1304/g4.2016.01.010.
孙立会 & 周丹华.(2019).国际儿童编程教育研究现状与行动路径.开放教育研究,25(02),23-35.doi:10.13966/j.cnki.kfjyyj.2019.02.003.
沈慧敏 & 程秀兰.(2023).美国幼儿编程教育：Kibo 机器人课程介绍与借鉴.陕西学前师范学院学报,39(07),52-61.
孙秀芝.(2019).《机器人向导》教学案例.中国信息技术教育,(09),12-14.
周冠男 & 张栋科.(2023).基于深度学习的图形化编程教学活动模型构建与实证研究——以小学信息科技课程为例.中小学电教,(06),3-9.
丁继红,范志浩 & 刘华中.(2024).基于协作编程多模态数据的学习投入可视化及关联分析——理解行为、认知、社会、情感的交互关系及对学习的影响.远程教育杂志,42(04),40-49.doi:10.15881/j.cnki.cn33-1304/g4.2024.04.005.

钟柏昌.(2016).中小学机器人教育的核心理论研究——机器人教学模式的新分类.电化教育研究,37(12),87-92.doi:10.13811/j.cnki.eer.2016.12.012.

朱月姣 & 杨力.(2022).可编程教育机器人：一种儿童早期计算思维培养的新工具.中国教育技术装备,(23),59-60+64.

王学男,林众 & 朱慧.(2019).基于科学素养的机器人教育与人才培养——访清华大学人工智能研究院院长张钹院士.中国电化教育,(06),1-5+36.

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

Cheng, E. C. K., & Wang, T. (2023). Leading digital transformation and eliminating barriers for teachers to incorporate artificial intelligence in basic education in Hong Kong. Computers and Education: Artificial Intelligence, 5, 100171.

Gross, M. D., & Veitch, C. (2013). Beyond top down: Designing with cubelets. Tecnologias, Sociedade e Conhecimento, 1(1), 150-164.

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(3-4), 276-291.

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.

双师一线：基于数字课程的人工智能实践教学探索

Dual Teachers One Line: Exploration of Artificial Intelligence Practical Teaching Based on Digital Courses

冯京和[1*]，王晓艳[2]，张荣荣[3]
[1]温州市教育教学研究院
[2]温州市市府路小学
[3]温州市第二实验小学
* 349018676@qq.com

【摘要】 人工智能教育对未来创新性人才的培养意义深远。本文阐述基于数字课程开发的人工智能实践教学的独特价值，通过开发人工智能知识图谱数字课程，依托数字课程以“场景延学、原理探学、创新引学、OMO融学”的方式开展双师课堂，并以“寻AI、学AI、研AI、创AI、评AI”问题解决主线，贯通线上线下的人工智能实践教与学过程，让学生有体验、可探讨、有创新，为推动人工智能教育提供新的一种学教方式。
【关键词】 数字课程；人工智能教学；双师课堂；知识图谱；学教方式

Abstract: *Artificial intelligence (AI) education holds profound significance for the cultivation of future innovative talents. This paper expounds on the unique value of AI practical teaching based on the development of digital courses. By developing digital courses of AI knowledge graphs, and relying on these digital courses, dual-teacher classrooms are carried out in the forms of "scene-based extended learning, principle exploration learning, innovation-driven learning, and Online-Merge-Offline (OMO) integrated learning". Moreover, with the main thread of problem-solving, namely "seeking AI, learning AI, researching AI, creating AI, and evaluating AI", the practical teaching and learning process of AI both online and offline is connected. This enables students to have experiences, engage in discussions, and demonstrate innovation, thus providing a new teaching and learning approach for promoting AI education.*
Keywords:digital Course ,artificial Intelligence Teaching ,dual-teacher Classroom,nowledge graph,teaching and Learning Approach

1. 引言

2024年教育部印发《教育部部署加强中小学人工智能教育》，明确了加强中小学人工智能教育的总体要求，凸显了其在教育战略布局中的重要地位，对满足未来创新型人才培养需求意义深远。

人工智能作为《义务教育信息科技课程标准（2022年版）》（以下简称“课标”）六大逻辑主线之一，内容模块安排在9年级。但中小学人工智能教材及其配套资源的匮乏、人工智能师资的不足、学教方式的范式不明等诸多现实难题，使得人工智能创新教育的普及教育存在着很大的难度。

熊璋教授的中小学人工智能教育基本架构中提出要开拓创新中小学人工智能教育，要为中小学生营造人工智能学习场景和生活场景、提供学习和研讨人工智能的资源、助推他们探索和创新应用人工智能的尝试，构建适合他们综合性学习的的人工智能平台。[1]

本文基于人工智能知识图谱开发数字课程，构建“双师一线”的人工智能实践教学模式，为培养未来创新型人才助力。

2. 基于数字课程开展人工智能实践教学的价值

本文阐述的人工智能数字课程是一种课程形态，通过数字化材料（如视频）、互动性（如在线讨论、测验）以及技术整合来设计课程。人工智能数字课程在丰富性、灵活性以及线上名师的专业优势方面有独特的价值，突破了传统教学的时空限制与资源瓶颈，为学生、教师、家长和社区成员提供了丰富的学习资源、广阔的交流平台和有力的实践支持。

2.1.支持项目实施

高等师范教育中人工智能教学正处于快速发展阶段，其培养体系并不完善。当下，人工智能教育师资大多来自于信息科技教师群体，他们在指导学生开展人工智能项目时，仍面临诸多挑战。三阶数字课程丰富的项目案例资源为一线教师组织开展校园人工智能项目实施提供有力支持。

2.2.链接家校实践

新课标要求强化课程综合性和实践性，推动育人方式变革，着力发展学生核心素养。[2]人工智能课程具有综合性和实践性的特性，需要大量的实践，布局在线上的数字课程，孩子除了可以在校内课堂开展教学实践，还可以助力孩子自己在家里开展自主探究，反复迭代项目成果，提升学习自驱力，或者孩子和父母一起开展家庭场景实践，增进亲子互动的同时，营造良好的家校共育环境。

2.3.拓展社区研学

人工智能技术更新快、门槛高，社区内缺少系统的人工智能教育资源和专业指导，难以组织有效的研学活动。数字课程为拓展社区研学共同体创造了良好条件。社区内的学生、家长以及教育爱好者可借助数字课程平台，共同参与人工智能研学活动。这不仅能丰富社区文化生活，还能提升社区成员的整体科技素养，让人工智能教育在社区中得到更广泛的传播和发展，促进社区形成一个学习型的人工智能研学共同体。

2.4.加强互联网交流

学生在进行人工智能学习时，遇到专业的问题，还需要请教同学、教师、线上名师及各界人工智能爱好者。在数字课程搭建起来的互联网交流空间中，学生可以随时发布自己在人工智能学习中的疑问、创意想法，来自不同地区的人都能给予回应，从而拓宽思路。这种融通的互联网交流空间，能整合各方资源，加速知识的传播与创新，推动小学人工智能教学不断进步。

3. 双师一线：贯通实践，助力创新人才培养

[1]熊璋,张洁.瞄准教育强国战略目标 加强中小学人工智能教育[J].人民教育,2024(24).

[2] 教育部.义务教育信息科技课程标准（2022 年版）[M].北京:北京师范大学出版社,2022.

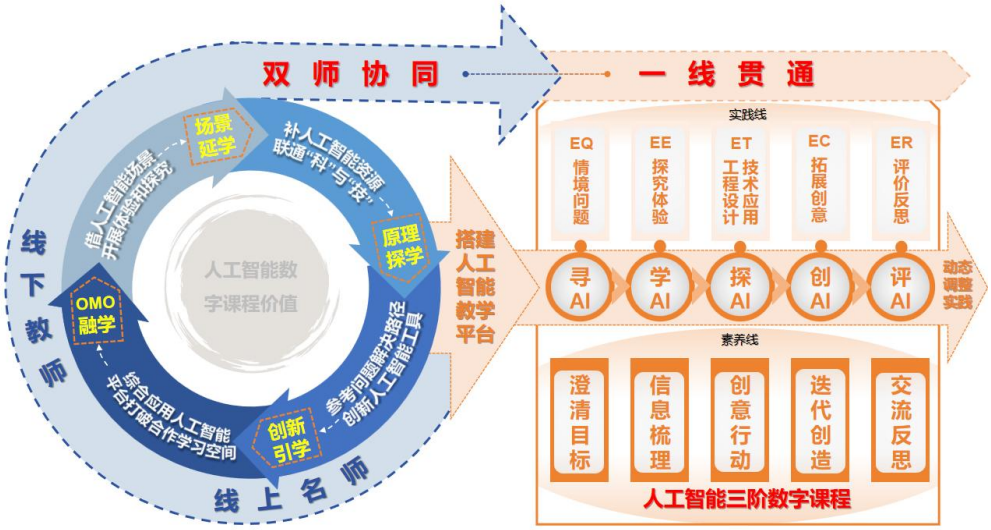


图 1 基于数字课程的人工智能“双师一线”教学模式图

3.1.线上名师：开发数字资源，完善人工智能知识图谱

黄蓉怀教授等人指出：基于课程标准与学科教学规律进行构建的教育知识图谱可以充分考虑学习知识点的先后顺序、同一知识点在不同学习阶段的不同要求，可以将碎片式、分散的教学资源与相关实体关联成一张巨大的语义网络。3[3]

因此，我们从课标的人工智能体系出发，将人工智能知识图谱化，开发人工智能基础性数字课程，可以助力师生实现个性化教学和学习路径规划。考虑到人工智能学习是以培养学生问题解决创新能力为导向，而项目化学习是落实核心素养的重要途径。所以，将人工智能项目的实践经典案例，转化为数字课程，开发了系列的微项目课程和跨学科课程，形成了中小学人工智能三阶课程体系。

从下表可以看到，基础性课程主要基于课标出发，聚焦起源发展、应用影响、识别原理、典型算法和其创新性的一些应用。微项目课程和跨学科课程更关注学生实践能力的提升，其中微项目课程以模型训练和智能识别为重点，展现学生在制作人工智能项目时会遇到的典型操作过程；跨学科课程则是关注智能技术与工程实践，更关注学生对项目问题的分解、分析和解决的流程。

**市中小学人工智能三阶课程体系			
基础性课程:起源发展、应用影响、识别原理、智能创作典型算法、无人驾驶、安全伦理、生成式人工智能		微项目课程 模型训练与智能识别	跨学科课程 智能技术与工程实践
小学 8课	01走近人工智能、02文字识别、03人脸识别、04语音识别、05语音合成、 06自然语言处理、07综合应用、08人工智能安全伦理	01植物动物果蔬识别 02人脸车牌身份识别 03温州方言识别 04语音智能服务 05手势姿势识别 ……不同场景	01《AI物联网农场》 02《AI稻草人》 03《无人机编队》 04《无人驾驶》 05《智能家居&芯片》 06《AI机器人》……
初中 14 + 12课	01走进人工智能、02人工智能应用、03人工智能技术基础、04人工智能的影响、 05语音识别技术、06语音合成技术、07图像识别技术、08人工智能中的算法、 09人工智能中的机器学习、10智能创意、11物联网数据的特点、 12智能物联系统的设计、13智能物联系统的硬件搭建、14智能物联系统的软件编写、 15智能物联系统的调试与完善、16智慧社会、17人工智能伦理、18人工智能安全、 19人工智能发展、20人工智能畅想、21预测原理探究、 22预测模型构建、23智能预测出行方式、 24认识无人机、25无人机飞行、26无人机创意飞行		
高中 18 + 4课	01人工智能的起源、02人工智能的现状和发展、03类脑计算、04逻辑推理、 05基于搜索的问题求解、06决策树、07回归分析、08贝叶斯分析、 09神经网络学习、10混合增强智能、11机器学习的起源与发展、 12机器学习的原理、分类与内涵、13人工神经网络与深度学习、 14“智能+X”推动社会进步、15自然语言理解：机器翻译、 16智能模拟：人机博弈、17智能控制：无人驾驶车系统、 18混合智能：脑机接口、19人工智能发展对社会的潜在影响、20概述、 21人工智能伦理、22人工智能安全		

[3] 刘超,黄荣怀等.基于知识图谱的新型教材建设与应用路径探索[J].中国大学教学,2023(08)

图 2 **市中小学人工智能三阶课程体系

3.2.双师协同：自选数字课程，满足教与学实践需求

3.2.1.场景延学：借数字课程的真实场景，开展体验和探究

线下教师在开展人工智能教学的过程中，根据教学需要选择线上名师提供的真实场景案例，组织学生利用数字课程开展体验和探究活动。例如，在讲解图像识别技术时，数字课程中的超市的自助收银系统识别商品、安防监控中的人脸识别等，让学生切实感受到人工智能的应用价值。在项目拓展的教学时，数字课程中的智能医疗影像诊断等，学生能真切地体验复杂的场景。大量真实且复杂的场景驱动学生的思维，从而锻炼学生从发现问题、分析问题到解决问题的完整能力链条，全面提升问题解决能力。

3.2.2.原理探学：借数字课程的原理动画，联通“科”与“技”

人工智能知识体系庞大且复杂，线下教师和学生常常很难理解技术背后的原理解析过程。线上名师通过数字课程，将复杂的人工智能概念、实验过程和原理技术以生动有趣的形式呈现出来。比如，在讲解人工智能算法时，利用动画演示算法的运行过程，将抽象的理论变得直观易懂。对于一些学校难以开展的实验，通过展示学生探究纪实视频，让学生即使没有实际设备也能了解实验步骤和结果。

线下教师在课堂教学中，根据学生的学习进度和困难，适时引入数字课程中的生动原理解析。这不仅补充了学生学习所需的人工智能资源，还帮助教师更好地备课和授课。双师课堂联通了“科”与“技”的难点落地。

3.2.3.创新引学：借数字课程的解决路径，创新实施和迭代

线上名师凭借丰富的教学经验和专业知识，在数字课程中展示大量的人工智能项目问题解决路径和案例。例如，在设计智能垃圾分类提醒器、智能校园巡逻机器人时，介绍如何借助机器分类等算法对数据进行学习，并依靠数据进行准确预测和分类等。

线下教师引导学生借鉴线上名师的问题解决分解路径，学生在创新路径的指引下，充分发挥自己的想象力和创造力，将所学知识应用到实际项目中。教师则在一旁给予指导和帮助，鼓励学生大胆尝试新的技术和方法。

3.2.4.OMO 融学：借数字课程的互动平台，打造泛在学习场域

OMO 学习是 Online-Merge-Offline 学习的简称。OMO 融学即借助线上线下各自的优势，集合数字课程和泛在学习空间，为创新引学、原理探学、场景延学打造学习场域。

线上名师将人工智能经典项目，布局到 UMU 互动平台上，在项目推进的过程中链接对应的“人工智能基础性数字课程”，组织学生开展线上的讨论、协作等活动，学生也可以根据线上活动指引，开展线下研学实践，并将自己的实践成果上传平台，和网络上的学习伙伴分享、交流。线下教师可以根据项目指引，指导学生在课前完成研学实践并上传平台，在课中组织学生开展面对面的交流和实践活动，对学生的线上学习成果进行检验和指导，在课后，引导学生开展拓展延伸，迭代创新。

这种泛在式学习空间和合作方式，让学生即时获得支持，助力自己自由开展项目研究。

3.3.一线贯通：搭建数字课程平台，融汇教与学实践过程

结合 5EX 教学模型，我们设计了“寻 AI、学 AI、探 AI、创 AI、评 AI”实践线，布局学习互动平台各环节，又依据每个环节中问题解决的需要融入数字课程支持学生创新素养的提升。本文将结合“校园情绪疗愈师的诞生”案例阐明一线贯通的过程，为师生开展线上线下人工智能教与学实践服务。

3.3.1.布局互动平台，构筑 AI 实践线

基于数字课程开展的学习实践活动的前提是要搭建互动学习平台。借助 UMU 平台中的课程创建功能，可以快速将课程分为五个章节，串联成实践线：

①“寻 AI”：帮助学生从生活中发现存在的问题，到找到问题解决方案的过程中，需要打通学生的基础认知和对 AI 技术的理解，寻找到合适的 AI 技术，形成解决方案的设计。

②“学 AI”：澄清目标后，学生要通过已有的信息，遵循 AI 技术解决方案的逻辑规则，找到唯一正确的答案或解决方案，即通过辐合思维将多种思路和各种信息梳理聚焦，从而确定项目探究时间轴。

③“探 AI”：将想法通物化成项目作品，进而实现解决真实的场景问题，这样从创想到创造的过程，是学生创新思维和创新实践集中表现的重要环节。

④“创 AI”：创新是一个不断调整的过程，更应该聚焦到学生迭代设计的过程引导中。

⑤“评 AI”：项目研究中的交流和反思可以促进学生思维从横纵两个方向深化思考，由此展现作品、听取多角度的意见就显得尤为重要。

如，《校园情绪疗愈师的诞生》一课，在“寻 AI”章节组织学生初步设计项目，确定校园情绪疗愈师的项目方向；在“学 AI”章节带领学生建构 AI 技术原理，并选择适合的 AI 技术开展实践，从而绘制出项目研究的时间轴；在“探 AI”章节中则参考“机器学习图像分类”的功能实现过程，完成项目物化；“创 AI”章节中，在原有物化基础上完善其他的功能；“评 AI”环节听取他人建议反思过程。

3.3.2.添加数字课程，指向创新素养

数字课程是指导学生创新问题解决方向的重要要素。在五个章节中添加“数字课程”，为学生创新素养实践的关键节点提供支持。

在“寻 AI”环节添加人工智能基础数字课程，引导学生观察是数字课程中的生活和技术应用场景，判断“能否借助该 AI 技术制作项目作品？”，从而澄清目标。人工智能数字课程聚合了不同 AI 技术原理讲解动画、实验视频等，将其添加到“学 AI”环节中，学生通过数字课程理解了 AI 技术背后的科学原理，梳理创新思维，确定了“怎么样的 AI 技术可以解决这个问题?”借由人工智能微项目和跨学科项目的数字课程，在“探 AI”章节中教师可以遴选出具有问题解决路径的内容，为学生的创意行动提供可行的问题解决思路，回答“如何物化实现人工智能项目成果？”。创新还在于对问题的不断迭代，由此，在“创 AI”和“评 AI”环节，借由数字课程中跨学科策略、迭代策略反思项目成果，回答“如何迭代”的问题。

如，在《校园情绪疗愈师的诞生》项目中，学生小组想要借助掌控板的机器学习识别故事卡片的内容，学习后识别的结果也是不错的。但当故事卡片放在“情绪树桩”的树桩框内去识别时效果就很差。后来才发现，框内光线差才导致识别率变差。怎么解决呢？学生又重新在数字课程中找到图像识别的多个解决策略，如直接改用 Mind 软件机器学习模块用计算机摄像头、沿用二哈摄像头但使用其中的二维码识别的办法。检测后换成了二维码，还加了线上同学提到的二哈摄像头加强光照的功能，识别率明显提高了。

3.3.3.依据互动反馈，调整活动环节

项目实践是一个动态调整的过程。在每个章节下的问卷、签到、提问、讨论、考试、作业、游戏等互动环节可以根据上一个环节学生反馈的内容删减和增加互动的内容，做到灵活灵活调整。动态调节的意义是尊重学生的实践结果，同时又体现了教师积极指导的作用。

如在学生项目迭代的过程中，发现学生上传的作品相似性比较高，由此利用 SCAMPER 思维法搭建了各小组的讨论环节。根据“情绪树桩”项目的成果，学生利用合并思维

（Combine）思考从其他组猜拳机器人中学到了石头剪刀布识别手势的玩法，借助代替思维（Substitute）用积极手势（加油、OK、正确等）手势替换，做出了“积极手势游戏”的功能。在交流中发现椴木板用特殊切割的方式可以做曲面，顺着改造思维（Adapt）想到用激光切割做了一个曲面的树桩侧面，求助老师对外观做了升级。所以，在这样的环节中，教师的教学策略可以随时发生变化，又让学生的主动性自然生成。

4. 结语

综上所述，“双师一线”模式通过数字课程贯穿教学各环节，展现出线上线下教师的各自的优势与价值，通过“寻、学、研、创、评”AI 的五环节搭建的人工智能学习平台，为学生提供了丰富的学习体验，促进了学生创新与实践能力的提升。未来，应进一步深化对该模式的研究与应用，不断完善数字课程资源，加强双师协作，持续推动小学人工智能教学向更高水平迈进。

参考文献

冯京和(2024).数字+课程体系:中小学人工智能课程实施推进的温州路线[J].中国科技教育
甘秋玲,白新文(2020).创新素养:21 世纪核心素养 5C 模型之三[J].华东师范大学学报
刘超,黄荣怀等(2023).基于知识图谱的新型教材建设与应用路径探索[J].中国大学教学
邱菡(2023).基于 OBE 的“七段式”混合式教学设计研究[J].大学教育
熊璋,张洁(2024).瞄准教育强国战略目标 加强中小学人工智能教育[J].人民教育

智能时代的信息甄别：生成式人工智能对批判性思维影响的系统性文献综述

Information Screening in the Age of Intelligence: A Systematic Literature Review of the Impact of Generative Artificial Intelligence on Critical Thinking

田婉萌[1]，张帝[2*]
[1 2]浙江师范大学教育学院
zhangdi0909@126.com

【摘要】 在智能时代背景下，生成式人工智能（AIGC）的迅速发展为教育领域带来了深远的影响，尤其对批判性思维（CR）能力的培养引起了学术界的广泛关注。本研究采用系统综述的方法，对 Web of Science (WOS) 和 Scopus 数据库中 2019 至 2024 年间的相关文献进行了深入分析，旨在探讨 AIGC 对批判性思维影响的研究现状。通过对关键词共现分析、样本群体、使用的 AIGC 工具和 AIGC 对 CR 的影响多个维度的系统梳理，本研究揭示了 AIGC 在促进与挑战批判性思维能力发展方面的双重作用，同时，指出了现有研究的不足并推荐了未来研究的潜在课题。

【关键词】 生成式人工智能；批判性思维；系统综述

Abstract:*In the context of the intelligent era, the rapid development of generative AI (AIGC) has profoundly impacted the education field, particularly in the cultivation of critical thinking (CR) skills. This study employed a systematic review method to analyze relevant literature from the Web of Science (WOS) and Scopus databases between 2019 and 2024, aiming to explore the current research status of AIGC's impact on CR. Through multi-dimensional analyses of keyword co-occurrence, sample groups, AIGC tools used, and their effects on CR, this study revealed the dual role of AIGC in both promoting and challenging CR development, identified gaps in existing research, and proposed potential topics for future research.*

Keywords: Generative artificial intelligence,critical thinking,systematic review

1.引言

随着人工智能技术的迅速发展，现在已经进入了一个由 AI 重塑人们学习、工作和生活方式的时代(Ahmad et al., 2023; Morandini et al., 2023)。在此背景下，生成式人工智能（AIGC）的崛起尤为引人关注，它不仅能够根据用户的个性化需求自动生成多种内容形式（如文本、图像和视频），还为教育领域带来了前所未有的机遇与挑战(Ma & Huo, 2023)。然而，随着生成式人工智能工具的应用，教育界对其影响的看法呈现出明显的两极化反应，AIGC 在为学生提供便捷的学习资源和即时解决方案的同时，也引发了对其对批判性思维（CR）能力产生影响的广泛关注(Rahman, 2023)。

批判性思维强调对信息的分析、评估和反思，是培养学生解决复杂问题和做出合理决策的基础(Facione, 2011)。然而，生成式人工智能通过提供即刻、广泛的信息和多样化的答案，可能使得学生在面对问题时过于依赖工具的自动生成内容。这种依赖性可能对批判性思维的培养产生不利影响，尤其是在学生可能会减少深入挖掘和质疑信息来源及其背后的逻辑假设

(Pennycook et al., 2020; Chan & Hu, 2023; Ding et al., 2023)。但是，生成式人工智能也并非完全负面，ChatGPT 也可以是增强批判性思维技能的宝贵工具，因为它允许从多个角度检查同一主题(Rudolph et al.，2023)。这种多元化分析的过程可以拓宽人们的视野，使学生能够从不同的角度思考问题，丰富思维过程(Mollick & Mollick, 2022)。因此，如何平衡生成式人工智能带来的信息多样性与学习者批判性思维培养，成为当前教育研究的重要议题。

为了考察 AIGC 对批判性思维的影响,本研究采用文献计量学映射分析和系统综述的方法,对 WOS 和 Scoups 数据库中发表的研究进行了分析,为此，本研究提出了以下研究问题：

（1）在 AIGC-CR 研究中，最常用的关键字是什么，关键字之间有什么关系？

（2）在 AIGC-CR 研究中，其样本群体和使用的 AIGC 工具有什么特点？

（3）生成式人工智能对批判性思维的影响是什么？

2.文献综述

2.1. 生成式人工智能

生成式人工智能是利用生成对抗网络、变分自编码器以及 Transformer 等技术，从输入的数据中捕捉特征，并生成高质量的内容（Li & Wu, 2022）。生成式人工智能的发展经历了几个重要阶段，早期的生成式模型主要基于简单的统计方法和规则系统，生成的内容较为有限(Facione, 2011)；随着深度学习技术的发展，生成的内容更加复杂和自然(Binns, 2022)；Transformer 架构的出现，进一步提升了生成式模型的性能，使内容更加流畅和连贯(Vaswani et al., 2017)。2022 年 11 月，OpenAI 发布了大规模预训练模型 ChatGPT，自推出以来已成为多个领域的变革力量(Zhang et al., 2023)。因此，如何平衡 AIGC 的便利性与批判性思维的培养，成为教育研究的重要议题。

2.2. 批判性思维

批判性思维（Critical Thinking, CR）是指在面对问题时，通过分析、评价、推理和反思等方式，对信息进行深度审视，以做出合理的判断和决策的能力(Facione, 2011)。培养批判性思维不仅是教育的主要目标之一，也是学生能够适应快速变化和复杂环境的关键能力(Halpern et al., 2021)。然而，在信息爆炸和知识更新迅速的时代，学生往往面临大量未经分析和筛选的信息，这使得批判性思维的培养变得尤为重要。随着生成式人工智能技术的发展，教育者开始关注如何在这种新技术环境下有效地促进学生的批判性思维能力。

2.3. 生成式人工智能对批判性思维的影响

有研究表明，生成式人工智能在教育中能够有效促进学生批判性思维的培养。AIGC 通过提供多元化的视角和解决方案，能够刺激学生在思考问题时从多个角度进行分析，这为批判性思维的发展提供了新的支持(Binns et al., 2022)。此外，AIGC 生成的内容可以帮助学生在解决复杂问题时形成更加全面的思考框架，这有助于培养学生的综合思维能力和批判性评估能力(Binns et al., 2022)。然而， Pennycook 等人(2020)指出，AIGC 提供的即时解答和信息可能会使学生过度依赖工具，而不进行深入的分析和批判。在这种情况下，学生可能会陷入表面理解的陷阱，缺乏对复杂问题背后深层次逻辑的深入剖析，这可能妨碍批判性思维的深度发展(Rahman et al., 2023)。因此,本研究进行了一项范围审查,探索并总结了 AIGC 对批判性思维的影响。

3.研究设计

3.1. 文献检索

为了回顾有关生成式人工智能对批判性思维影响的文献，研究人员检索了 2019-2024 年 Web of Science 和 Scoups 两个数据库中的论文。研究人员使用以下搜索术语及其同义词：("generative Artificial Intelligence" OR "chatbot" OR"Intelligent Tutor"OR"Intelligent Agent"OR"Virtual assistant" OR "ChatGPT" OR "Conversational AI" OR "Image Generation" OR "Intelligent Content Generation"OR"AIGC"OR "GAI" OR "Large Language Model" OR "Generative Adversarial Networks")AND("Critical Thinking" OR "Analytical Thinking"OR"Evaluating Thinking" OR "Judgment" OR "Reflective Thinking" OR "Discernment")。检索后首次得到 1864 篇论文，其中 41 篇是 AIGC 对 CR 影响的最终审查论文。

3.2. 文献筛选和纳入

为了保证文章研究结果的可靠性，制定了如下标准进行文献筛选:1)至少使用一种基于生成式人工智能的工具;2)研究对象为学生或教师；3)生成式人工智能需用于教育目的；4)文章需评估或分析批判性思维的任何维度。

首先在预选数据库中按照既定关键词对文献进行检索,共得出 1864 篇文章，通过过滤器被应用于自动排除非期刊文章、非英语文章和非教育&教育科研研究领域的文章,筛选后剩下 327 篇文章。接下来通过手段阅读标题和摘要对文章进行筛选，其中文献中有 222 篇文章基于以下 4 种原因被剔除。因此，有 105 篇文章被选中进行进一步的全文筛选，基于以上四种原因,105 篇文章中有 64 篇被删除。因此最后发现有 41 篇文章符合纳入标准。

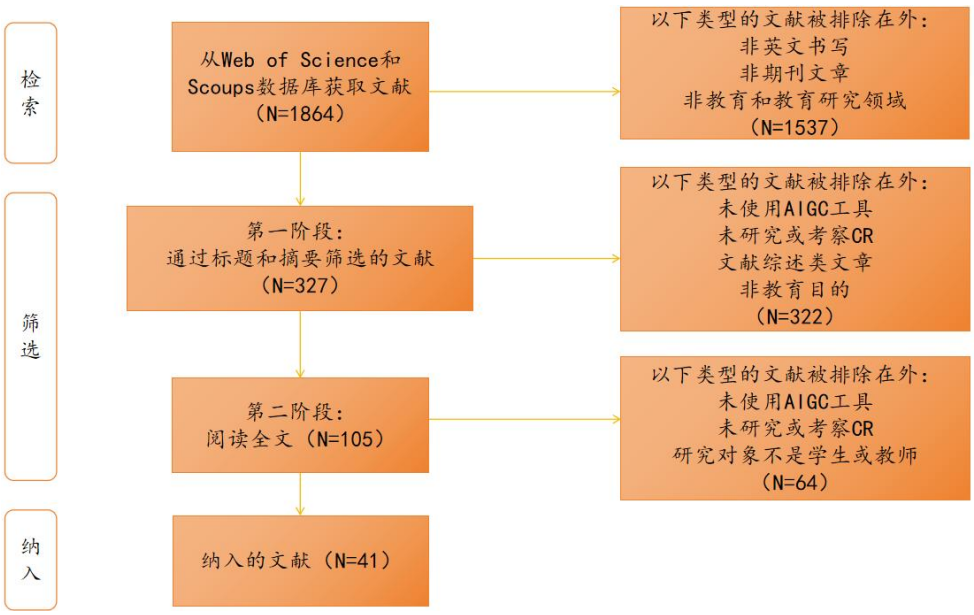


图 1 符合标准文章的检索过程

4.结果

4.1. 关键词

41 篇 AIGC 对 CR 影响的文章中共包含 110 个关键词，图 2 为 VOSviewer 生成的关键词聚类分析结果图。使用频率最高的关键词是“chatgpt”（f=13）、“artificial intelligence”（f=6）、“critical thinking”（f=5）、“higher education”（f=4）、“large language model”（f=2）、“artificial intelligence in education”（f=2）、“learning”（f=2）、“management education”（f=2）和“stem education”（f=2）。从图 1 也中可以看出，对 AIGC-CR 的研究集中在四个主要集群，即 chatgpt、人工智能、批判性思维和高等教育。

红色集群描述了 chatgpt 是 AIGC 的代表，即侧重于使用 chatgpt 应用于 STEM 领域以支持教与学。绿色集群描述了人工智能在教育中的跨学科应用，比如 Chang 等人(2024)探讨了将基于 CIDI 模型的 ChatGPT 系统融入护理与健康教育培训课程的应用模式及效果。蓝色集群描述了在本科教育（undergraduate students）中人工智能对学生批判性思维能力的测评和训练非常重要。紫色集群则侧重于高等教育，关注 AIGC 在高等教育中的应用及其产生的教学意义。

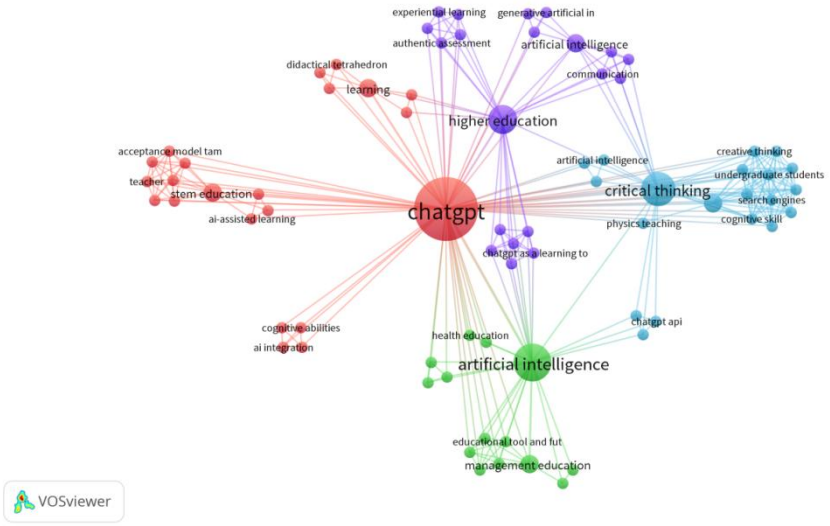


图 2 AIGC-CR 研究中最常用的关键词

4.2. 样本群体

图 3 中显示了生成式人工智能对批判性思维影响的研究中采用的样本群体。研究发现，高等教育阶段（85%）的学生是最常被采用的样本，其次是混合样本组（7%）和教师（5%），还发现在这些研究中采用中学生作为样本组的研究只有一项。

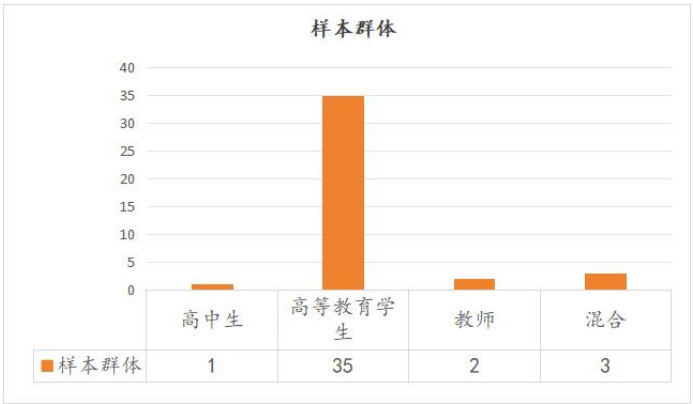


图 3 AIGC-CR 研究中的样本群体

4.3. 生成式人工智能工具

该研究的纳入标准之一是使用了 AIGC 工具。研究发现 41 篇论文中都使用了 AIGC 工具，如图 4，大多数（65.3%）都为 ChatGPT；其次 12 种（24.29%）的 AIGC 工具式研究人员选择的其它类型的工具；最后，有 5 种是由研究人员自建的工具，例如 InquiryGPT(Wang et al.,2024)和 GPT-Assisted Summarization Aid (GASA)(Lin,2024)等。

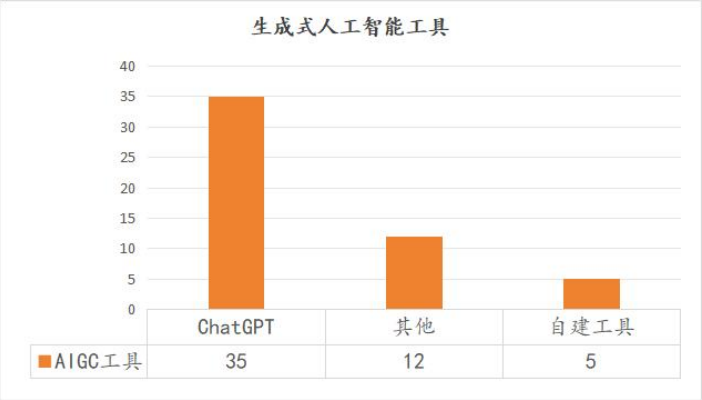


图 4 AIGC-CR 研究中使用的 AIGC 工具

4.4. 生成式人工智能对批判性思维的影响

在本研究中，探讨了 AIGC 对被研究者 CR 的影响，如图 5。需要说明的是，为了保证研究的客观性和科学性，所以本研究剔除了 5 篇仅使用问卷调查这种研究方法得出研究结果的文献。因此，共有 28 篇（77.78%）文章表明 AIGC 能够促进 CR 的提升，共有 6 篇（16.67%）文章表明 AIGC 不利于 CR 的提升，2 篇（77.78%）文章显示 AIGC 对 CR 没有影响。

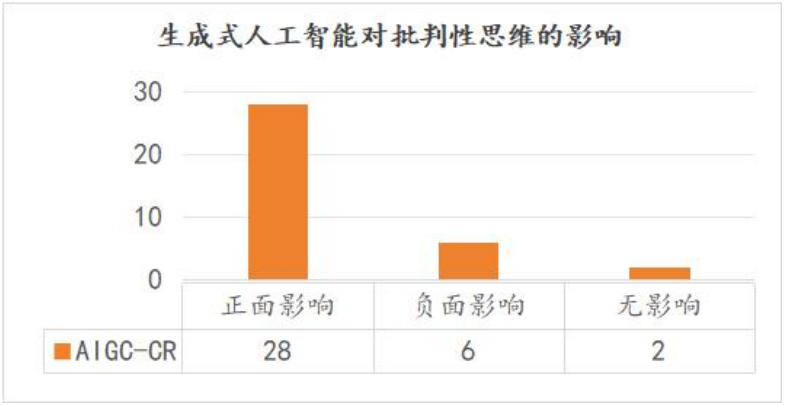


图 5 AIGC 对 CR 的影响

具体而言，AIGC 在不同研究中对 CR 的影响呈现出多样化的趋势。例如，Bitzenbauer 等人(2023)发现，教师引导学生通过对 ChatGPT 输出内容进行多步骤处理可以使学生意识到 ChatGPT 信息可能因提示不同而有差异，从而促使他们从多源核实信息，有助于培养高中生的批判性思维习惯；Darmawansah 等人(2024)发现，使用 ChatGPT 支持的协作论证方法的学生在批判性思维意识方面比使用传统技术基础的协作论证方法的学生有显著提高；Wu 等人(2024)的研究发现，实验组教师使用能够模拟同伴互动和提供结构化反馈的 ChatGPT 进行编程学习，并完成与每周主题相关的开放式项目时，会更有效促进批判性思维的发展。

表 1 生成式人工智能对批判性思维的影响评审论文列表

作者	使用的生成式人工智能工具	生成式人工智能对批判性思维的影响	探索的结果
Bitzenbauer（2023）	ChatGPT	+	教师引导学生通过对 ChatGPT 输出内容进行多步骤处理使学生意识到 ChatGPT 信息可能因提示不同而有差异，从而促使他们从多方面核实

			信息，有助于形成批判性思维的习惯
Darmawansah（2024）	ChatGPT	+	使用 ChatGPT 支持的协作论证方法的学生在批判性思维意识显著提高
Lin（2024）	ChatGPT	+	使用 PA-GPT 的实验组在批判性思维方面有显著提升
Dadan（2023）	ChatGPT	-	它可能导致学生过度依赖，减少独立思考和解决问题的能力，缺乏对信息来源和准确性的判断
Esmaeil	ChatGPT	-	学生意识到过度依赖 ChatGPT，这可能会削弱他们的学习和批判性思维能力
Liang（2024）	ChatGPT	/	学生在参加了该项测试后，其批判性思维技能没有发生显著变化

6 篇（16.67%）文章显示 AIGC 会对 CR 产生负面影响，Dasari 等人(2024)的研究表明，ChatGPT 等生成式人工智能在学生学习过程中，虽然能提供信息和帮助，但也存在一些问题影响学生批判性思维发展。比如，它可能导致学生过度依赖，减少独立思考和解决问题的能力，学生可能会不加批判地接受 ChatGPT 提供的信息，缺乏对信息来源和准确性的判断，从而对学生的批判性思维产生负面影响。Esmaeil 等人(2023)的研究发现，虽然 ChatGPT 能够提供信息和指导，帮助学生更快地完成写作任务，但学生也表达了对其准确性的担忧，并意识到过度依赖 ChatGPT 可能会削弱他们的学习和批判性思维能力。

只有两篇文章（5.56%）表明 AIGC 对 CR 没有影响，Liang 等人(2024)的研究发现，学生参与了基于 ChatGPT 的批判性英语学习任务后，其批判性思维技能没有显著变化。

5.讨论

本研究分析了 WOS 数据库中 2019 年至 2024 年间发表的 41 篇 AIGC 对 CR 影响的论文。许多研究表明,使用 AIGC 工具对提高学生的批判性思维能力有很大的潜力。从分析结果中，得出了以下发现和启示:

（1）对关键词进行聚类分析，发现 AIGC-CR 研究有四个聚类：即 chatgpt、人工智能、批判性思维和高等教育。这表明当前研究主要集中在这些领域，特别是 ChatGPT 在教育中的应用及其对批判性思维的影响。

（2）AIGC 对批判性思维影响的研究中最常采用的样本群体是高等教育学生，其次是混合和教师群体。另一方面，AIGC-CR 研究很少采用小学生和初高中生。这可能是由于小学生和初高中生在认知发展、信息获取能力等方面存在较大差异，使得在这些群体中进行批判性思维的影响研究更具挑战性。同时，AIGC 技术在基础教育阶段的应用尚未广泛推广，甚至在一些地方尚未得到充分的认可和规范。因此，小学生和初高中生作为样本群体的研究较少。此外，选择高等教育学生可能是因为他们具有较高的认知水平，能够更好地理解和使用 AIGC 工具进行批判性思维的训练和评估，以及大多数研究人员都在大学任教，因此他们从本系的学生那里收集数据更加方便。

（3）在此研究中最常用的生成式人工智能工具是“ChatGPT”，其次是研究人员选择的其它类型的 AIGC 工具，最后是自建工具。这可能是由于 ChatGPT 是当前最知名的 AIGC 工具之一，功能强大并拥有大量用户，研究者倾向于选择易被公众接受且有较多研究参考文献的工具。

同时，一些研究可能需要特定领域的 AIGC 工具，这些工具在特定任务上的表现可能优于通用工具，因此也有部分研究人员选择自建所适配的 AIGC 工具。

（4）在 AIGC 对 CR 的影响方面大多数研究表明是正面的，认为它能够促进学生的批判性思维能力、提高信息筛选的效率和准确性。这是因为批判性思维是培养学生分析和评估信息的关键能力。然而，也有一些研究发现，AIGC 可能对学生产生负面影响，特别是在过度依赖 AIGC 的情况下，可能削弱学生的自主思考和批判性分析能力。这种现象在某种程度上是合理的，因为过度依赖 AIGC 可能导致学生对信息的深度分析和独立判断能力的降低。

6.结论

总之，本文发现 AIGC 的进步使得研究人员对批判性思维进行了多样化的研究，基于研究结果和本文的讨论，对相关领域提出以下几点建议：

（1）未来的研究应不仅聚焦于高等教育学生，还应将中小学生纳入研究样本群体。这将有助于更***全面地理解 AIGC 对不同教育阶段学生批判性思维的影响，特别是在基础教育阶段，学生的认***知发展和信息获取能力存在较大差异，研究这些群体将更具挑战性但也具有价值。

（2）未来的研究可以开发如何在相应的课程中应用适配的生成式人工智能工具，不再局限于使用已有的生成式人工智能工具，进而使其对学习者越来越有吸引力，根据他们的年龄和发展的独特阶段特征量身定制工具，这将有益于进一步的提升学习者的批判性思维。

（3）未来的研究可以全面检查批判性思维的多个维度，包括分析、评估、推理和反思等。此外，研究应探讨 AIGC 如何在不同维度上促进或抑制批判性思维的发展。这将有助于更深入地理解 AIGC 对批判性思维不同方面的具体影响，从而为教育实践提供更有针对性的指导。

7.限制

这项研究有一些局限性需要注意。尽管从两大数据库中获取了 AIGC 对 CR 影响的相关研究，但最终纳入分析的文章数量仍然有限。这一现象可能源于 AIGC 对学习者 CR 影响的多元性和复杂性，它不仅涉及信息判断与评估能力的培养，还关乎学生的学习成就和其他高级思维技能。因此，探讨 AIGC 对 CR 影响的研究仍然具有挑战性，需要后续的研究进一步完善该领域。

参考文献

Ahmad, H., Hanandeh, R., Alazzawi, F., Al-Daradkah, A., ElDmrat, A., Ghaith, Y., & Darawsheh, S. (2023). The effects of big data, artificial intelligence, and business intelligence on e-learning and business performance: Evidence from Jordanian telecommunication firms. *International Journal of Data and Network Science*, *7*(1), 35-40.

Binns, A. (2022). AI in education: Opportunities and challenges. Educational Technology Research and Development, 70(3), 1234-1256.

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, *15*(3), ep430.

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), 43.

Chang, C. Y., Yang, C. L., Jen, H. J., Ogata, H., & Hwang, G. H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology & Society*, *27*(1), 215-230.

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G. J. (2024). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies*, 1-25.

Dasari, D., Hendriyanto, A., Sahara, S., Suryadi, D., Muhaimin, L. H., Chao, T., & Fitriana, L. (2024, January). ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning. In *Frontiers in Education*. Frontiers Media SA.https://doi.org/10.3389/feduc.2023.1295413

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education*, *20*(1), 63.

ESMAEIL, A. A. A., MAAKIP, I., MATANLUK, O. O., & MARSHALL, S. (2023). Understanding student perception regarding the use of ChatGPT in their argumentative writing: A qualitative inquiry. *Jurnal komunikasi: Malaysian journal of communication jilid*, *39*(4), 150-165.

Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight assessment*, *1*(1), 1-23.

Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology press.

Li, P. H., Lee, H. Y., Lin, C. J., Wang, W. S., & Huang, Y. M. (2024). InquiryGPT: Augmenting ChatGPT for Enhancing Inquiry-Based Learning in STEM Education. Journal of Educational Computing Research, 62(8), 2157-2186.

Li, Y., & Wu, J. (2022). Emerging trends in generative artificial intelligence: Definitions and applications. IEEE Transactions on Neural Networks and Learning Systems, 33(10), 4792-4803.

Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, *54*, 101645.

Lin, C. J., Lee, H. Y., Wang, W. S., Huang, Y. M., & Wu, T. T. (2024). Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid. Education and Information Technologies, 1-23.https://doi.org/10.1007/s10639-024-13072-5

Ma, X., & Huo, Y. (2023). Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework. *Technology in Society*, *75*, 102362.https://doi.org/10.1016/j.techsoc.2023.102362

Mollick, E. R., & Mollick, L. (2022). New modes of learning enabled by ai chatbots: Three methods and assignments. *Available at SSRN 4300783*.https://dx.doi.org/10.2139/ssrn.4300783

Morandini, S., Fraboni, F., De Angelis, M., Puzzo, G., Giusino, D., & Pietrantoni, L. (2023). The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations. *Informing Science*, *26*, 39-68.

Pennycook, G., Rand, D. G., & Thoma, D. A. (2020). Cognitive bias and the use of generative AI tools in education: A cautionary perspective. *Journal of Educational Psychology, 112*(2), 211-225.

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

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of applied learning and teaching*, *6*(1), 342-363.

Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.arXiv:1706.03762.

Wu, T. T., Lee, H. Y., Chen, P. H., Lin, C. J., & Huang, Y. M. (2025). Integrating peer assessment cycle into ChatGPT for STEM education: A randomised controlled trial on knowledge, skills, and attitudes enhancement. *Journal of Computer Assisted Learning*, *41*(1), e13085.

Zhang, C., Zhang, C., Zheng, S., Qiao, Y., Li, C., Zhang, M., ... & Hong, C. S. (2023). A complete survey on generative ai (aigc): Is chatgpt from gpt-4 to gpt-5 all you need?. *arXiv preprint arXiv:2303.11717*.

面向"AI4S"的高中人工智能课程开发与实施

Curriculum Development and Implementation for High School Artificial Intelligence Education: A Focus on "AI4S"

林淼焱[1]，陈恩慈[2]，谢作如[3]
[123]温州科技高级中学
wzlinmiaoyan@163.com

【摘要】 本文探讨了 AI4S 理念在高中生人工智能素养培育中的应用。通过开发并实施课程体系，加强中学生 AI 核心素养的培养。课程基于 UbD 逆向设计和跨学科融合理念，以生成式 AI、计算机视觉和自然语言处理等项目为载体，结合学科知识，引导学生从 AI 模型训练到硬件部署，制作解决生活问题的科创作品。
【关键词】 AI4S；人工智能素养；课程开发

Abstract: *This document explores the application of the AI4S (AI for Science) concept in cultivating high school students' artificial intelligence literacy. By developing and implementing a curriculum system, the study aims to strengthen students' core AI competencies. The curriculum is designed based on UbD backward design and interdisciplinary integration principles, using projects such as generative AI, computer vision, and natural language processing as carriers. It combines subject knowledge to guide students from AI model training to hardware deployment, enabling them to create innovative technological works that address real-life problems.*
Keywords: AI4S; Artificial Intelligence Literacy; Curriculum Development

1. 研究背景与意义

1.1. AI for Science

2024 年诺贝尔物理学奖和化学奖均与人工智能（Artificial Intelligence）相关，分别表彰了人工智能在科学研究中的应用以及其对科学认知的改变。对此，浙大教授吴飞(2025)评价道：AI 天然具备推动学科交叉的潜力，无论是从人工智能角度解决科学问题（AI for Science），还是从科学的角度优化人工智能（Science for AI），未来的重大突破都将源自于这种交叉领域的工作。

前述提及的 AI4S（AI for Science），是利用人工智能技术推动科学研究的学科研究方向，也被广泛视为一种革新科学研究的理念。

1.2. 中学生 AI 核心素养培育

人工智能素养已经经过了学界的充分研究。有学者将其界定为学生适应智能社会所需的知识、技能和态度的复合体，强调 AI 知识、技能和态度的综合培养（钟柏昌, 2023）。也有学者提出“应用、模型、算法”三个用 AI 解决问题的层次（谢作如, 2024）。综合来看，笔者认为人工智能素养来自于学生对人工智能工具的实践应用以及对人工智能模型的理解与创新，这一观点与学界的研究成果相契合。

中学人工智能教育实施主流路径包括机器人、AIGC 等方向，缺乏对技术原理、算法及跨学科结合的深度探讨，难以满足 AI4S 范式下对学生能力的要求。

因此，本研究以温州市某高中的学生为研究对象，基于当前高中人工智能教育内容、现状，以 AI4S 理念为指导开发并实施课程体系，加强中学生 AI 核心素养的培养，以适应未来科学研究和社会发展的需求。同时，通过课程实践检验 AI4S 理论在不同学科场景中的适配性，探索“AI＋学科”的融合式课程设计路径，为教育者提供可复制的创新实践模板，助力 AI 教育规模化推广。

1.3. 课程开发实践经验

笔者所在单位创设的"通识-社团-竞赛"三级课程框架，为本课程提供了衔接保障。

在通识课程阶段，学生每周都会了解到人工智能的基本概念和应用。这一阶段的课程设计注重基础知识的普及，让学生对 AI 有一个全面的认识。

与此同时，学校每周开设的社团课程中也有较高占比的科创社团，供学生深入学习 AI 模型的构建和应用。

针对能力突出或兴趣浓厚的学生，他们将会加入到科创团队，学习竞赛相关课程，开发开源硬件交互作品，或是不断优化算法以解决生活问题，已在各类竞赛中斩获佳绩。

2. 课程规划

2.1. 课程设计理念与目标

2.2.1. 设计理念

基于 UbD 逆向设计的课程理念：逆向设计是以目标为导向的教学方法，分为三个阶段，确定预期成果、确认评估证据、设计学习活动。人工智能作为一门刚刚进入中小学的学科，各部门的标准都在紧锣密鼓地制定中，教师更需要聚焦于关键的学习目标，避免教学内容的随意性和盲目性。而 UbD 逆向设计正有助于提高教学的针对性和实效性，帮助教师理清授课思路，更好地整合教学资源，创设真实的学习情境。

跨学科融合的课程视角：AI4S 理念与跨学科设计理念不谋而合，都致力于打破学科界限，促进知识的综合应用与创新。本课程将以计算机领域的项目为载体，结合学科知识，用 AI 的视角分析生活中的各类科学问题。

2.2.2. 课程目标

知识目标：掌握 AI 基础知识（如编程逻辑、机器学习流程、深度学习模型搭建）。

技能目标：通过实践操作，培养数据处理、模型训练与部署、硬件交互设计等能力。

情感目标：让学生认识到数据隐私的重要性，理解保护个人和集体数据隐私是每个人的责任，并在项目实践中主动采取措施保护数据安全。引导学生在开发和使用 AI 技术时，始终考虑其伦理和社会影响，避免技术滥用。

素养目标：培养学生的 AI 思维（如“数据决定模型效果”），增强解决实际问题的能力。

2.2.3. 课程配套软硬件

课程配套软硬件包括 OpenHydra 平台，行空板、摄像头、USB 小音箱、传感器、小舵机等。该平台能较好支持课堂活动开展，其两大特性能满足中学 AI 课程的教学需求。首先是支持 GPU 算力切分，允许多用户共享资源，降低硬件成本的同时满足深度学习需求，避免算力不足的窘境。其次，平台无需重复配置学生机环境，资源上传后，学生就能通过浏览器进行学习与编程实践，教师也能直接访问到学生的容器查看学习情况。

综上，本课以开源硬件为辅助，引导学生从 AI 模型训练到硬件部署，制作解决生活问题的科创作品。融入 AI4S 理念，结合化学、生物等学科实验场景，通过机器学习解决科学问题，并衔接相关赛事。

2.2. 课程体系

整体课程预期花费 20 课时，如果按每周一课时的课程安排，可以将课程拆分为普通学生（12 课时）+兴趣学生（8 课时）。模块安排见下图，每课详细信息见附录，现对各个模块的内容作简单介绍。

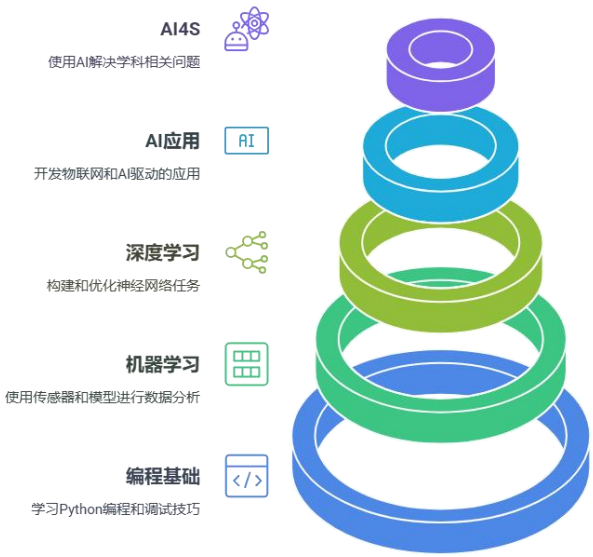


图 1 课程模块示意图

2.2.1. 大模型助我学编程

本模块旨在帮助初学者快速掌握编程基础，并依托大模型的引导完成编程实践。从 Python 语言基础入手，借助 XEdu 环境和预训练大模型，帮助学生掌握编程逻辑与调试技巧，涵盖数据类型、控制结构及智能体构建等内容，为学生后续学习奠定扎实基础。还涉及大模型与智能体，调用 XeduHub 库的预训练模型与 XEduLLM 框架构建智能体应用。

2.2.2. 机器学习

学生通过开源硬件连接传感器，训练经典机器学习模型。在这个过程中，学生将经历“获取数据-验证结论”的过程，例如利用线性回归和多项式回归分析实验数据，建立输出值与实际值的数学关系，并通过行空板实现模型推理与交互设计，强调数据驱动科学问题的解决思路。

2.2.3. 深度学习

本模块聚焦神经网络搭建与优化。每课的最后一个环节都为“测试模型推理效果，并转化为 ONNX 部署”，其中图像分类与目标检测任务需要本地或云算力支持。结合 BaseNN、MMEdu 等框架，从全连接网络到卷积神经网络逐步深入，完成手势分类、昆虫识别等任务，同时探讨模型参数量与推理速度的优化策略；详解 ImageNet 数据集格式，对比全连接与卷积网络的特征提取效率差异，并通过 ONNX 格式部署模型，引导学生了解跨平台轻量化部署的意义。

2.2.4. AI 应用开发

学生通过 PinPong 库（国产开源硬件接口框架）驱动硬件，使用代码控制物联网设备，将代码和模型部署到行空板，开发出具有实际应用价值的作品。本模块从读取传感器数据、

控制执行器开始，让学生能以较低门槛构建多模态应用，将图像模型分类结果输入大模型，通过语音合成模型转化为自然语音输出，实现“图像识别-语义生成-语音反馈”的端到端流程；学习实时人机交互，部署手势识别模型，解析手势类别并映射为舵机动作，强化学生对AI技术实际落地的理解。

2.2.5. AI4S

学生在这一模块需要具备一定的学科知识基础，才能将学科问题如递推数列参数求解、区分植物细胞壁结构差异抽象为回归、图像识别任务，培养AI辅助科学推理的能力，总结出“问题抽象-特征分析-模型选型-结果归因”的AI4S方法论。作为问题解决的有力工具，教师还可以引导学生了解嵌入技术与序列建模（循环神经网络与Transformer架构），并调用ModelScope平台上的预训练模型进行实践。

2.3. 思维链实践

虽然各路大模型都搭载了思维链（Chain of Thought）功能，但学生在使用AI辅助解决融入了学科知识的AI4S问题时，仍需主动参与和理解每个步骤的逻辑与原理，避免过度依赖于输出的答案。笔者根据学生反馈，总结出了解决问题的思维链如下。

表1 复杂问题下询问AI要点

核心步骤	示例
描述问题类型	一道力学题
列出所有已知信息	质量、速度、摩擦力
明确需求	求滑行10米后的速度

基于以上思考，学生对大模型的合理提问应该是：“已知质量3kg的物体以6m/s滑动，摩擦系数0.2，求滑行8米后的速度。请分三步：1.列出公式；2.计算摩擦力做的功；3.用动能定理求速度”，必要时上传题目图片或公式。

具体来说，AI4S方法论的核心在于将学科问题转化为AI任务，并通过AI技术解决实际问题。以下是操作流程及示例：

2.3.1 问题抽象

将学科问题抽象为AI可处理的形式。例如：

学科问题：求解递推数列的参数

AI任务：将问题转化为回归问题，预测数列的未知参数

2.3.2 特征分析

提取问题中的关键特征。例如：

递推数列的特征：初始条件、递推公式、已知项

2.3.3 模型选型

选择合适的AI模型。例如：

回归问题：选择线性回归模型

2.3.4 结果归因

分析模型输出结果与实际问题的关系。例如：

模型输出：预测的参数值

实际问题：验证预测值是否符合递推数列的规律

3. 经典课例分析

现从整体课程设计中选出一节课程的设计方案进行解析，该方案曾在上海奉贤区面向无机器学习基础的初中生开设，在课堂上有较高的任务完成度。课例主题为“看谁扔得准——机器学习初探”。教学资源包括 XEdu 项目中的 BaseML 工具，教学素材包括多项式回归示例代码、投掷数据集等。

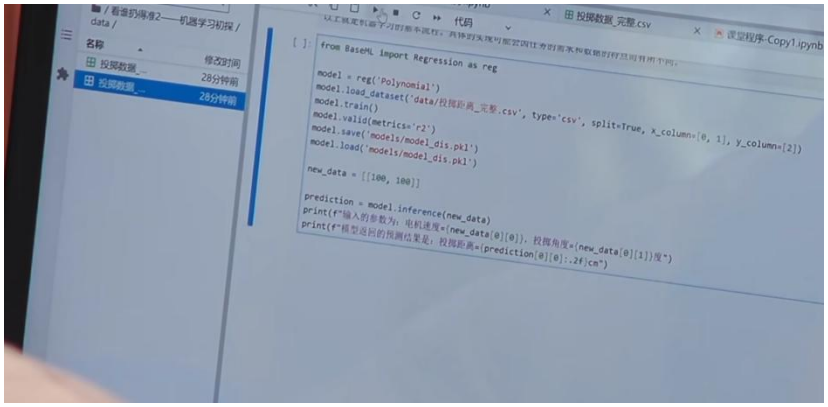


图 2 学生访问服务器进行在线编程

3.1. 确定预期成果

本节课的目标是让学生了解机器学习的基本流程，掌握多项式回归算法的作用及其适用范围。学生将亲历“整理数据—训练模型—推理模型”的全过程，理解 AI 将成为科学研究和技术创新的新工具。在知识层面强调机器学习“数据➡训练➡推理”流程及多项式回归适用于非线性关系预测的特点；技能层面借助 BaseML 工具实现四行代码完成模型训练与特征调整；素养层面则着重培养“数据决定模型效果”的 AI 思维，并通过机器人投掷测试等真实场景强化实践感知。

3.2. 确认评估证据

教学评估采用多维度证据体系，包括课堂实操评分（代码运行、模型准确率）、过程观察（调试参与度、小组协作）及态度反馈（反思日志深度）。特别设计了兴趣问卷（如“如何用 AI 解决其他问题”）和课堂问答（如“为何需同时输入角度与速度”）等形成性评价。预期学生能独立完成特征列调换等简单迁移任务，通过误差分析归因数据或算法问题，表达拓展 AI 应用场景的兴趣。

3.3. 设计学习活动

活动 1（导入）：观看“扔瓶子大赛”视频，讨论“如何让机器投掷更准”，引出数据驱动思想。

活动 2（探究）：运行教师提供的 BaseML 示例代码，观察数据与模型关系，填写关键参数表格，分析机器学习的一般步骤与流程。

活动 3（实践）：分组修改数据集训练模型，用 GenAI 辅助调试代码错误（如数据路径报错）。

活动 4（迁移）：给定新目标距离，调整特征数据并训练模型，观察训练效果。核心的逻辑如下图。

	特征		标签
	A	B	C
1	motor_speed	angle	distance
2	70	80	17
3	70	90	19
4	70	100	18
5	70	110	21.6
6	70	120	20.4
7	70	130	17.5
8	95	80	33
9	95	90	35.9
10	95	100	37.1
11	95	110	36.5
12	95	120	34.3
13	95	130	30.3
14	100	80	34

调整

	特征		标签
	A	B	C
1	angle	distance	motor_speed
2	80	17	70
3	90	19	70
4	100	18	70
5	110	21.6	70
6	120	20.4	70
7	130	17.5	70
8	80	33	95
9	90	35.9	95
10	100	37.1	95
11	110	36.5	95
12	120	34.3	95
13	130	30.3	95
14	80	34	100

图 3 调换数据集以解决新问题

4. 实践成果与辐射效应

由北京科学智能研究院等单位联合举办的 AI4S TEEN Challenge 系列赛涵盖数学、物理、化学、生物等领域。我校学生在生物细胞分类（区分若干类叶片组织图片）、求解正三棱锥的体积（给定几何条件，估计正三棱锥体积）、金属 X 在硝酸溶液中的反应（分析反应产物的比例变化，推断物质系数）等跨学科项目中，得到较高的识别准确率或预测精度，验证了 AI 技术驱动科学探究的可行性。

基于 AI4S 理念及部分课程资源，我校科创团队联合多机构发起“AI4T 教师培育计划”。活动吸引 1638 名教师报名，507 人进入实战环节，提交跨学科 AI 项目作品。培训直播累计观看 11752 次，辐射范围广泛，为中小学 AI 教育普及提供了可复制的模式。

图 4 基于本课理念的网络培训实践

方案中 OpenHydra 平台（本地部署成本＜5 万元）与行空板（硬件套件＜600 元）的组合方案，突破传统实验室建设的高成本壁垒，提供 GPU 算力支持模型训练，为县域中学提供可复制的技术路径。

5. 总结

课程实践过程中暴露的问题主要集中在技术平台与师资方面。尽管 OpenHydra 平台和开源硬件硬件组合为教学提供了很大便利，但其稳定性和易用性仍有提升空间。其次，许多教师缺乏系统培训，难以独立开展教学活动，我在实践中也深感自身知识的局限，迫切需要技术的提升与更多的交流机会。

AI 与基础学科的融合正成为教育的新趋势，文中各类经验将为落实"2030 年中小学 AI 教育普及"目标提供可操作的实践范本，让 AI 不再高不可攀，真正走入中学课堂。

引用文献

北京科学智能研究院. (2023). AI4S 科学智能赛事：中学生人工智能素养挑战赛 [在线资源]. Bohrium 竞赛平台. https://bohrium.dp.tech/competitions/8524835166

顾小清, 李睿, & 李世瑾. (2023). 人工智能教育现状如何？——中小学人工智能教育竞赛结果的证据回应. 开放教育研究, 29(6), 82-91.

王冬青, 汪婷婷, & 汤化涛. (2022). 我国中小学人工智能课程发展现状及策略研究. 教育信息技术, 2022(Z1), 3-8.

吴飞. (2025). DeepSeek：迈向全社会分享的普遍智能. 文汇报·上观新闻.

新华社. (2023). 科技部启动“人工智能驱动的科学研究”专项部署工作 [在线资源]. 中国政府网. https://www.gov.cn/xinwen/2023-03/27/content_5748495.htm

教育部办公厅. (2024). 关于加强中小学人工智能教育的通知 (教基厅函〔2024〕32 号) [政策文件].

周澍云. (2023). 高中生人工智能素养评价指标体系构建研究 [博士论文, 华东师范大学].

钟柏昌. (2023). 如何通达新一代人工智能教育的核心. 中小学数字化教学, 2023(11), 1.

钟柏昌, 刘晓凡, & 杨明欢. (2024). 何谓人工智能素养：本质、构成与评价体系. 华东师范大学学报（教育科学版), 2024(1), 71-84.

谢作如. (2024). 青少年需要怎样的 AI 通识课程. 中国信息技术教育, 2024(22), 1.

6. 附录

具体课程设计表

节次	名称	内容
1-1	初识 Python 和大模型	了解 Python 语言；安装 XEdu 并编写代码；面向大模型的报错查看与提问方法；
1-2	Python 数据处理	数据类型；数据容器（list、str、dict）和基本操作；
1-3	控制结构	常见分支结构写法；两种不同的循环结构；
1-4	从大模型到智能体	调用 XeduHub 上的模型；基于 XEduLLM 开发连接大模型的应用；
2-1	数据可视化	数据可视化之 Pandas、matplotlib 库；
2-2	线性回归：寻找数据背后的规律	采集并分析 LM35 温度模块的数据；训练线性回归模型预测真实温度；
2-3	多项式回归：看谁扔得准	使用投掷装置的数据训练多项式回归预测投掷距离；
2-4	机器学习模型的部署和交互设计 （根据硬件情况选上本课）	使用行空板+摄像头获取画面中的特征点信息（人脸、手势等）；将模型部署到行空板上，测试真实情境下的推理效果；
3-1	全连接神经网络和手势分类	用 BaseNN 搭建全连接神经网络；
3-2	MMEdu 和手势分类	了解 ImageNet 数据格式和制作； 用 MMEdu 的 LeNet 网络训练手势识别模型；
3-3	MobileNet 和昆虫分类（GPU）	用 MobileNet 训练手势识别模型；
3-4	SSD 和目标检测（GPU）	了解 coco 数据集格式和制作（标注）； 用 MMEdu 的 Det 模块 SSD 算法训练模型；
4-1	用 pinpong 驱动硬件	常见传感器的读取；执行器的控制（LED 和舵机）；
4-2	语音讲解小助手	训练一个图像分类模型，结合大模型、语音合成模型播报结果；
4-3	人机互动装置	用 ONNX 模型识别手势并控制舵机；
5-1	使用大语言模型辅助理解学科问题	基于大语言模型的思维链建构；用回归方法解决 AI4S 例题之“递推数列的参数求解”与“来自未知星球的金属元素的还原反应求解”
5-2	AI4S 情境问题解决	使用卷积神经网络解决问题“区分三类植物细胞”与“寻找暗物质晕的踪迹”；总结解题的一般流程
5-3	嵌入技术和向量比较实验	了解嵌入技术；验证向量的加减实验；

5-4	序列处理（一）：循环神经网络和自动作诗机（GPU）	了解循环神经网络的原理和搭建方法；用 BaseNN 搭建循环神经网络并训练自动作诗机模型；
5-5	序列处理（二）：Transformer 架构与注意力机制（GPU）	了解主流模型分享网站如 hugging face，modelscope 等；调用各类预训练模型

人工智能应用于 STEM 教育以培养计算思维的系统性文献综述

A Systematic Literature Review of Artificial Intelligence Applied to STEM Education to Foster Computational Thinking

上官舒涛[1] 刘晨伟[2] 杨启凡[3]
福建师范大学 教育学院，福建福州 350000
sgst18335689913@163.com

【摘要】 本文是一篇系统性文献综述，旨在探讨人工智能在 STEM 教育中培养计算思维的应用现状与作用。通过检索中外数据库，筛选出 2014 年至 2024 年间的 12 篇相关文献进行分析。研究发现，人工智能在 STEM 教育中应用广泛，其中以自主开发的工具为主，机器人和大语言模型应用较少。大部分研究表明，人工智能技术有效促进学生的计算思维，但过度支持可能削弱学生的自主思考能力。国内外研究均以小学生为主要样本群体，国内实证研究较少。研究建议加强人工智能工具研发，探索与 STEM 教育的深度融合，关注学生个体差异，并加大国内实证研究投入，以推动教育的发展。

【关键词】 人工智能；STEM 教育；计算思维；文献综述

Abstract: *This paper is a systematic literature review aiming to explore the current status and role of the application of artificial intelligence in cultivating computational thinking in STEM education. By searching Chinese and foreign databases, 12 relevant literatures between 2014 and 2024 were screened and analyzed. It was found that AI is widely used in STEM education, which is dominated by self-developed tools, with less application of robotics and large language modeling. Most of the studies showed that AI technology effectively promotes students' computational thinking, but over-support may weaken students' ability to think autonomously. Both domestic and international studies take primary school students as the main sample group, and there are fewer domestic empirical studies. The study suggests strengthening the research and development of AI tools, exploring the deep integration with STEM education, focusing on individual student differences, and increasing the investment in domestic empirical research to promote the development of education.*

Keywords: Artificial intelligence, STEM education, Computational thinking, Literature review

1.引言

人工智能新时代来临，对创新科技与人才培养提出了新的战略要求。为全面推进教学改革，切实提高育人质量，教育部于 2022 年推出的新版义务教育课程方案将“加强课程综合，注重关联”作为其中基本原则(教育部，2022)。作为跨学科教育的代表，STEM 教育在培养关联不同学科主题学习中具有极大潜力(崔嘉欣等人，2024)。并且，基于现阶段国家创新人才培养需求，STEM 人才的培养规模与质量将直接关系到国家科技创新能力的提升和高水平科技自立自强的实现(郑庆华，2025)。因此，STEM 教育愈加重要。

随着先进技术和云计算工具的出现，未来社会需要大量具备人机协同能力的高水平人才，计算思维成为人人必备的能力(王学男 & 李永智，2024)。特别是，STEM 教育中必须培养计

算思维技能以适应当下教育变革(Swaid, 2015)。并且，计算思维与问题解决能力的培养是STEM教育的核心目标，将计算思维融入STEM教育体系，为培养具备技术应用与创新能力的新一代学习者提供了有效路径(Wahab et al., 2021)。将计算思维融入STEM教育对于21世纪新时代学生的教育具有十分重要的意义，然而目前STEM教育中明确关注计算思维的研究十分稀缺(Li et al., 2020)。推动计算思维与STEM教育深度融合的进程应成为未来研究的重要议题(Wang et al., 2022)。

本研究聚焦于人工智能时代STEM教育中计算思维培养的整体现状，基于综述法进行文献分析，旨在探讨人工智能技术在STEM教育中的应用对培养计算思维的作用。基于此，本研究提出以下研究问题：

(1) STEM教育中培养计算思维的人工智能技术应用情况如何？

(2)人工智能技术应用于STEM教育对于培养计算思维有什么作用？

(3)国内外人工智能技术应用于STEM教育以培养计算思维的研究有什么相似和区别？

2.概念界定

2.1. STEM教育

STEM代表科学(Science)、技术(Technology)、工程(Engineering)和数学(Mathematics)的集合(Martín-Páez et al., 2019)。综合国内外对于STEM教育的定义主要分为三类：首先是对于四门课程的简单综合；另一种是跨学科的教学方式，主要强调学科间的融合；第三种是指STEM理念或者素养，帮助学生理解科学(Moore et al., 2014)。然而，STEM教育并不是科学、技术、工程和数学教育的简单叠加,而是要将四门学科内容组合形成有机整体,以更好地培养学生的创新精神与实践能力(余胜泉 & 胡翔，2015)。

2.2. 计算思维

计算思维的概念最早由周以真教授提出：通过运用计算机科学的基础概念，将复杂问题转化为可解形式（如通过分解、嵌入、转换或模拟等方法），从而实现问题解决、系统设计及人类行为理解的思维范式(Wing, 2006)。虽然目前对计算思维的定义众说纷纭，但大多数学者考虑的因素有许多的相似之处(Ezeamuzie & Leung, 2022)。例如国际教育技术学会(ISTE)对于计算思维的操作性定义，即计算思维由分解、模式识别、抽象和算法四部分能力组成(Prottsman, 2022)。2022年，我国教育部发布的《义务教育信息科技课程标准（2022年版）》中对计算思维的阐述为：计算思维是指个体运用计算机科学领域的思想方法，在问题解决过程中涉及的抽象、分解、建模、算法设计等思维活动(教育部，2022)。综合各方面定义，本研究认为计算思维是学生运用计算机领域的思想方法解决问题的一种思维方式。

3.研究设计

研究综述法SALSA模型，即检索 (Search) 、评估 (AppraisaL) 、综合 (Synthesis) 和分析(Analysis)是指通过对某一领域文献进行检索和编码，根据条件筛选综合与研究主题相关的文献并进行数据分析，最终形成研究结果的文献分析方法。本研究基于研究综述法进行系统性文献综述，以探究人工智能技术应用于STEM教育以培养计算思维的研究现状(刘超洋 & 陈冲，2019)。

本研究外文数据库选择Web of Science，检索类型设置为“期刊”，检索关键词为“Artificial Intelligence”、“STEM education”及“computational thinking”等多种组合，检索时间为2014年至今10年内的相关文献；中文数据库选用中国知网，检索类型为“学术期

刊”，关键词由“人工智能”、“STEM 教育”及“计算思维”等各种组合或全部，检索时间为 2014 年至今。

文献筛选标准为：一是研究主题需要满足至少有一种以上的“人工智能技术”；二是“STEM 教育”需要以四门学科内容组合形成的教育模式，并非其中单一学科；三是“计算思维”的定义应满足我国义务教育标准的狭义定义，即“在问题解决过程中涉及的抽象、分解、建模、算法设计等思维活动”。值得注意的是，“STEAM 教育”（添加“艺术(Art)”）也将会被纳入研究范围。

从中文数据库中依次添加“人工智能”、“STEM 教育”及“计算思维”关键词检索后共有 49 篇文章，其中学术期刊共 11 篇，筛选后满足条件仅剩 3 篇。从外文数据库中依次添加“Artificial Intelligence”、“STEM education”及“Computational thinking”关键词检索后共有 23 篇文章，其中学术期刊共 15 篇，筛选后满足条件仅剩 9 篇。如图 1。

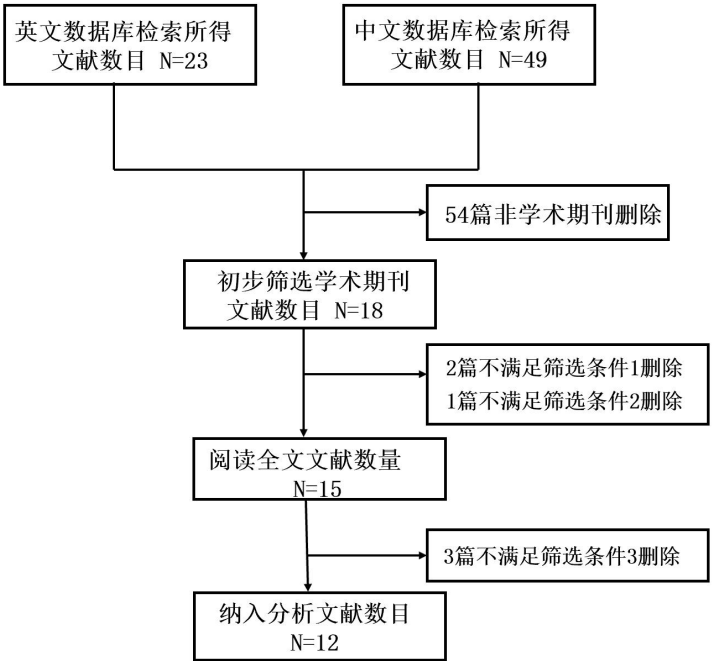


图 1 符合标准的文献筛选过程

4.研究结果

4.1. 人工智能工具的应用情况

图 2 表明人工智能应用于 STEM 教育以培养计算思维的研究中的人工智能工具应用情况。本研究所筛选的文章中均有使用人工智能工具。人工智能工具分类主要按照“机器人”“大语言模型”（例如 ChatGPT）与其他三个类别划分。其中机器人（16%）与大语言模型（25%）占比较小，大多研究采用其他自主开发的人工智能工具（58%）。

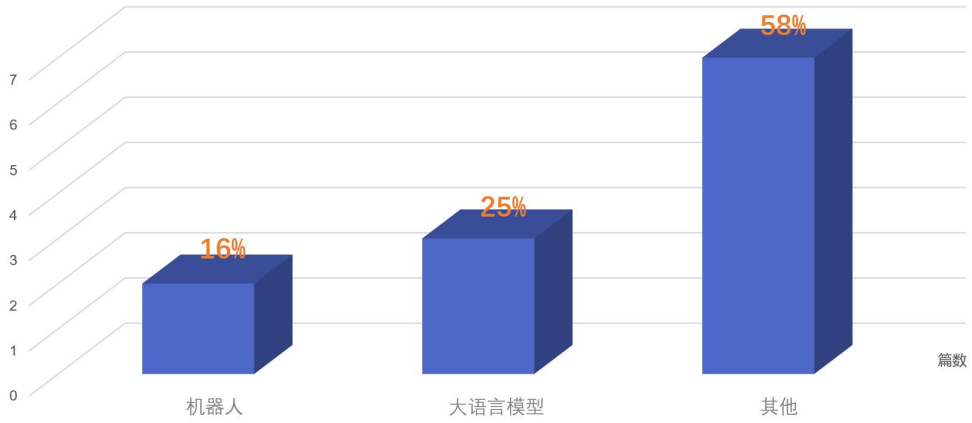


图 2 人工智能工具的应用情况

4.2.对计算思维的影响

图 3 表明人工智能应用于 STEM 教育以培养计算思维的研究中对计算思维的影响情况。为了保证研究的客观性和科学性，本研究已筛除仅采用问卷收集数据的文献。在本文所筛选的 12 篇文献中，统计所得共有 11 篇文献（92%）表明将人工智能工具应用于 STEM 课堂能够促进学生的计算思维发展。Naya-Varela 等人（2023）表明像 Nao 和 Pepper 等机器人，通过编程实现与学生的互动，协助教学，助力学生发展问题解决和分析能力，提升学习 STEAM 学科的积极性，并促进学生计算思维发展。Guo 等人（2025）研究发现将语言大模型融入课堂之中能够进一步给予学生学生支持，从而提高学生动手能力，发展计算思维。Khuda 等人（2024）发现在项目式学习中整合学科知识，结合 AI 辅助教学工具能够进一步培养学生的计算思维能力。除此之外，Wu 等人(2022)发现将人工智能技术应用于课堂若提供过多的学习支持，则会导致学生缺乏思考的过程，从而导致计算思维无法得到发展。因此，综上所述适当的是支持对于培养计算思维是必要的。

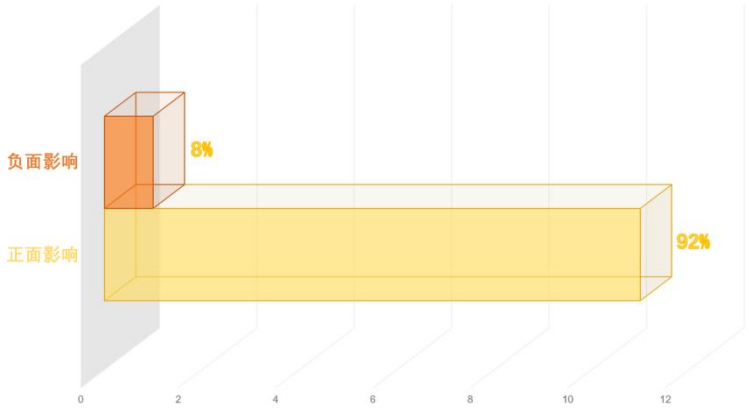


图 3 对计算思维的影响

4.3. 国内外样本群体与外文献来源分类

如图 4 表明，本此筛选文献中有 3 篇来自中文数据库，9 篇来自外文数据库。如图 5 表明了人工智能应用于 STEM 教育以培养计算思维的研究中的采样群体分布。研究发现小学阶段（66%）的学生群体最常被采样。初中以及高中以上的群体（16%）相当。值得注意的是

在高中以上的群体中存在一组高等教育阶段本科生的采样群体。

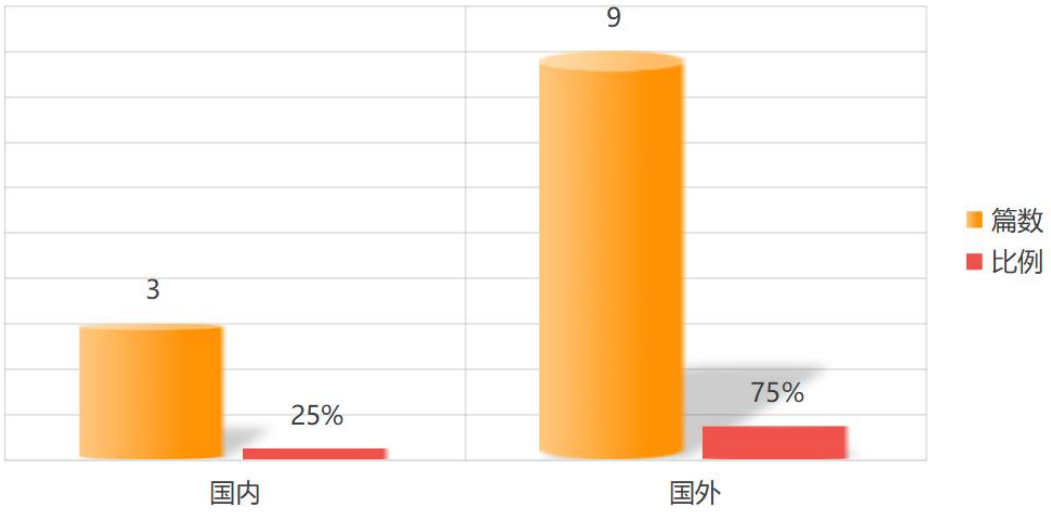


图 4 样本来源分布

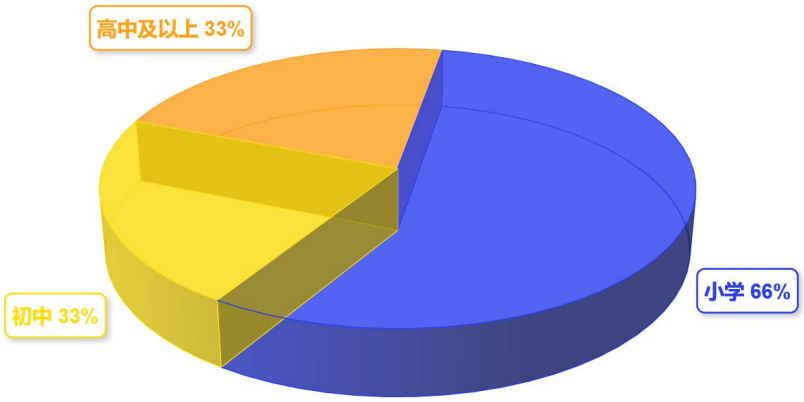


图 5 样本年龄段分布

5.讨论

本研究分析了 Web of Science 数据库与中国知网数据库中 2014 年到至今发表的 12 篇人工智能应用于 STEM 教育以培养计算思维的论文，大部分研究表明将人工智能技术应用于 STEM 课堂能够进一步促进学生的计算思维发展，所以人工智能技术存在巨大的潜力。从分析结果中，得到了以下的发现与启示：

（1）从人工智能工具的应用情况分析，人工智能工具主要还是以其他类型工具为主，机器人与大语言模型位列其次。从学生认知发展分析，这可能是由于机器人和语言模型的使用需要一定的基础和能力，在本次研究中大部分研究样本为小学生群体，这可能不太适用。同时，针对小学生思维特点，构建趣味化情境教学模型，选择简单易操作、符合小学生认知水平的其他工具（例如结合 Scratch 的人工智能工具），更符合其思维发展特点(Jiang & Li, 2021)。从技术操作层次分析，对于小学低年级学生来说，理解和操作机器人的编程指令存在困难，难以达到预期教学效果(Žanko et al.,2022)。同时，从技术本身分析，机器人和语言

模型在教育应用中存在一些局限性。机器人的成本较高，包括购买、维护和编程学习成本等，这限制了其在教育中的广泛应用。

(2)从培养计算思维的角度分析，大多文章表明人工智能对计算思维培养的影响是起到积极作用。从技术角度出发，人工智能技术为 STEM 教育提供了丰富的实践情境，有助于学生更好地理解和应用 STEM 知识，让学生在实际操作中锻炼计算思维技能(Akramova et al., 2024)。从学生发展角度出发，基于人工智能技术的课堂中，学生大多面临各种复杂的任务和挑战，学生有解决任务的需求，在不断尝试和解决问题的过程中，学生学会从不同角度思考问题，优化解决方案，从而提升了计算思维水平。

(3)从国内外样本群体与来源分析。在人工智能技术应用于 STEM 课堂中以促进计算思维的国内外研究中，从不同点来看，国外对于这类实证研究的相关研究较多，而国内相关研究较少。从共同点来看，国内外大部分均选用小学生样本群体，其次是中学生及以上学段，对于高等教育阶段的研究涉及并不多。综合数据分析，从学生发展阶段来看，这可能是因为小学阶段是学生思维快速发展的关键时期，具有很强的可塑性。同时，小学阶段的学生好奇心强，对新鲜事物充满探索欲望在这一阶段引入人工智能技术开展 STEM 教育，能够进一步提高学生的学习积极性，对于培养学生的计算思维等关键能力具有重要意义(Liu et al., 2023)。从研究实施的角度来看，小学相对集中的课程安排和相对统一的教学环境便于进行研究样本的选取和实验的开展。从文献来源角度分析，可见国内关于人工智能技术应用于 STEM 课程促进计算思维发展的实证类研究较少，国外相关内容的实证内研究较多，可能是因为国外在人工智能和教育融合领域的研究起步较早，积累了更丰富的研究经验和成果，同时，国外一些发达国家在教育资源投入上较为充足，为开展实证研究提供了良好的条件。

6.结论

总之，本文发现将人工智能技术的进步使得研究人员对计算思维发展进行了多样化的研究，基于研究结果和本文的讨论，对相关领域提出以下几点建议：

（1）对于人工智能工具应用角度，应优化机器人、大语言模型等系统界面，使其更符合学生认知发展阶段，降低操作难度，提高学生的学习体验。同时，人工智能技术应更贴合 STEM 教育的实际需求，加强对人工智能工具的研发和改进。

（2）对于人工智能应用课堂培养计算思维角度，未来的研究需进一步探索人工智能技术与 STEM 教育的深度融合促进计算思维方式。同时，应探索如何更好地引导学生利用模型进行自主学习和探究，提高学习效果。此外，关注学生差异，包括学生的认知水平、学习风格和兴趣爱好等方面差异，未来研究应重视这些差异，实现个性化教育

（3）相比于国外，国内在人工智能应用于 STEM 教育培养计算思维的实证研究相对较少，需要加大投入，开展长期跟踪研究，深入分析人工智能技术应用的效果和影响因素。通过实证研究，为教育实践提供更具科学性和可靠性的依据，推动国内人工智能教育的发展。

参考文献

教育部. (2022).义务教育课程方案（2022 年版）[EB/OL]..https://www.gov.cn/zhengce/zhengceku/2022-04/21/content_5686535.htm

崔嘉欣,蒋雨辰,李静 & 冯秀梅. (2024).STEM 视域下师范生跨学科教学能力培养研究——基于部属师范大学物理专业人才培养方案的分析.物理与工程,1-8.

郑庆华.(2025).人工智能赋能 STEM 教育创新发展：认识与实践.中国高教研究,(01),1-7.doi:10.

16298/j.cnki.1004-3667.2025.01.01.
王学男 & 李永智.(2024).人工智能与教育变革.电化教育研究,45(08),13-21.doi:10.13811/j.cnki.eer.2024.08.002.
余胜泉 & 胡翔. (2015). STEM 教育理念与跨学科整合模式. *开放教育研究, 21* (04), 13-22. doi:10.13966/j.cnki.kfjyyj.2015.04.002.
为未来社会培养具有数字素养与技能的人才——义务教育信息科技课程标准（2022 年版）解读. (2022).基础教育课程,(10),67-73.
刘超洋 & 陈冲.(2019).国外教育研究综述法的源起、方法论基础和基本步骤.高教探索,(08),111-118.
Akramova, G., Ma'murov, B., Akramova, S., Qo'ldoshev, R., & Shodmonova, A. (2024). Methods of using STEAM technologies in the development of pupils' computational thinking. In *E3S Web of Conferences* (Vol. 538, p. 05034). EDP Sciences.
Ezeamuzie, N. O., & Leung, J. S. (2022). Computational thinking through an empirical lens: A systematic review of literature. *Journal of Educational Computing Research*, *60*(2), 481-511.
Jiang, B., & Li, Z. (2021). Effect of Scratch on computational thinking skills of Chinese primary school students. *Journal of Computers in Education*, *8*(4), 505-525.
Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2020). On computational thinking and STEM education. *Journal for STEM Education Research*, *3*, 147-166.
Liu, X., Wang, X., Xu, K., & Hu, X. (2023). Effect of reverse engineering pedagogy on primary school students' computational thinking skills in STEM learning activities. *Journal of Intelligence*, *11*(2), 36.
Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, *103*(4), 799-822.
Moore, T. J., Stohlmann, M. S., Wang, H. H., Tank, K. M., Glancy, A. W., & Roehrig, G. H. (2014). Implementation and integration of engineering in K-12 STEM education. In *Engineering in pre-college settings: Synthesizing research, policy, and practices* (pp. 35-60). Purdue University Press.
Prottsman, K. (2022). *Computational thinking meets student learning: Extending the ISTE standards*. International Society for Technology in Education.
Swaid, S. I. (2015). Bringing computational thinking to STEM education. *Procedia Manufacturing*, *3*, 3657-3662.
Wahab, N. A., Talib, O., Razali, F., & Kamarudin, N. (2021). The big why of implementing computational thinking In STEM education: A systematic literature review. Malaysian Journal of Social Sciences and Humanities (MJSSH), 6(3), 272-289.
Wang, C., Shen, J., & Chao, J. (2022). Integrating computational thinking in STEM education: A literature review. International Journal of Science and Mathematics Education, 20(8), 1949-1972.
Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.
Naya-Varela, M., Guerreiro-Santalla, S., Baamonde, T., & Bellas, F. (2023). Robobo smartcity: An autonomous driving model for computational intelligence learning through educational robotics

. IEEE Transactions on Learning Technologies, 16(4), 543-559.
Guo, Q., Zhen, J., Wu, F., He, Y., & Qiao, C. (2025). Can Students Make STEM Progress With the Large Language Models (LLMs)? An Empirical Study of LLMs Integration Within Middle School Science and Engineering Practice. Journal of Educational Computing Research, 07356331241312365.
Khuda, I. E., Ahmed, S., & Ateya, A. A. (2024). STEM-Based Bayesian Computational Leaning Model-BCLM for Effective Learning of Bayesian Statistics. IEEE Access.
Wu, S. Y., & Yang, K. K. (2022). The effectiveness of teacher support for students’ learning of artificial intelligence popular science activities. Frontiers in Psychology, 13, 868623.
Žanko, Ž., Mladenović, M., & Krpan, D. (2022). Analysis of school students' misconceptions about basic programming concepts. Journal of Computer Assisted Learning, 38(3), 719-730.

优选·精研·迭代：探索基于浦育平台新一代人工智能教育的乡村小学智慧校园管理新路径

Optimization, In-depth Research, and Iteration: Exploring New Paths for Smart Campus Management in Rural Primary Schools Based on the New Generation of Artificial Intelligence Education on the Pu Yu Platform

陈慧娴[1] 温小敏[2]
[1] 平阳县昆阳镇第六小学
[2] 平阳县昆阳镇第七小学*
38917141@qq.com * 392766525@qq.com

【摘要】 随着新一代人工智能技术的迅猛发展，教育领域正经历着深刻的变革。本文聚焦于乡村小学智慧校园管理的创新实践，以基于浦育平台的《红领巾检查小助手》课程项目为例，探讨如何通过项目式融合课程设计与实施，促进人工智能教育与校园管理的深度融合，解决乡村小学在设备管理、教学环境及学生行为规范等方面的实际问题。本文通过优选编程平台、深入研究学习课程以及迭代项目功能这三大核心实践策略，旨在平阳县昆阳镇第六小学为试点，为乡村小学探索一条在智慧校园管理中应用新一代人工智能教育的新途径，从而提升整体的管理水平和教育成效。

【关键词】 乡村小学；智慧校园管理；浦育平台；新一代人工智能教育

Abstract: *With the rapid development of the new generation of artificial intelligence technology, the education sector is undergoing profound changes. This article focuses on the innovative practice of smart campus management in rural primary schools, taking the "Red Scarf Inspection Assistant" course project based on the PuYu platform as an example. It explores how to promote the deep integration of artificial intelligence education and campus management through the design and implementation of project-based integrated courses, and solve practical problems in rural primary schools such as equipment management, teaching environment, and student behavior norms. Through three core practical strategies: selecting an optimal programming platform, in-depth research on learning courses, and iterative project function development, this article aims to pilot at the Sixth Primary School of Kunyang Town, Pingyang County, and explore a new path for the application of new-generation artificial intelligence education in smart campus management in rural primary schools, thereby enhancing the overall management level and educational effectiveness.*

Keywords: Rural Primary School; Smart Campus Management; PuYu Platform; New Generation Artificial Intelligence Education

1.前言

在基础教育阶段，学生的日常行为规范管理是德育工作的重要组成部分。然而，传统的管理方式在应对大规模学生群体时往往效率低下，尤其是在乡村小学，资源有限，管理难度

更大。根据《温州市中小学推进人工智能教育实施方案》(国务院，2017)，温州市教育局推动的人工智能与基础教育深度融合计划，为解决这一问题提供了新的视角。本文以此为背景，结合乡村小学实际情况，笔者设计并实施了一项基于浦育平台的新一代人工智能教育课程项目——《红领巾检查小助手》，旨在通过技术创新优化校园管理，同时提升学生的信息素养和创新能力。

2.以实践应用为依托

在笔者所在的乡村小学，值日生每天早晨在学生入校时会在校门口检查他们是否佩戴红领巾和穿着校服。通过观察笔者发现，很多学生经常会忘记佩戴红领巾。当值日生发现有学生未佩戴红领巾时，他们需要逐一登记这些学生的信息。随着时间的推移，值日生向笔者反映，在学生入校高峰时段，人流量较大时，容易出现遗漏的情况。记录工作也会变得繁琐，有些学生甚至故意提供错误的姓名和班级，给值日生的统计工作带来了困难，导致统计未佩戴红领巾学生名单的准确性受到影响。笔者身为信息科技教师，以“能否引导学生借助新一代人工智能技术来解决遇到的问题？”这一真实问题为指引，开启了《红领巾检查小助手》课程项目的设计与实践。

2.1.优选编程平台，推动深度学习体验

在编程平台的选择上，笔者考虑了多种编程软件，如 Mind+、腾讯扣叮等。在考虑过程中，她主要关注三个关键因素：1.适配机房配套设备。2.操作简便的编程平台。3.全程学习和体验 AI 技术。基于这些考虑，笔者最终选择了浦育平台的“创意积木工坊”。

首先，学校的机房配备 45 个摄像头和耳麦（图 1），能够借助浦育平台的“创意积木工坊”进行多类型的人工智能项目创作。

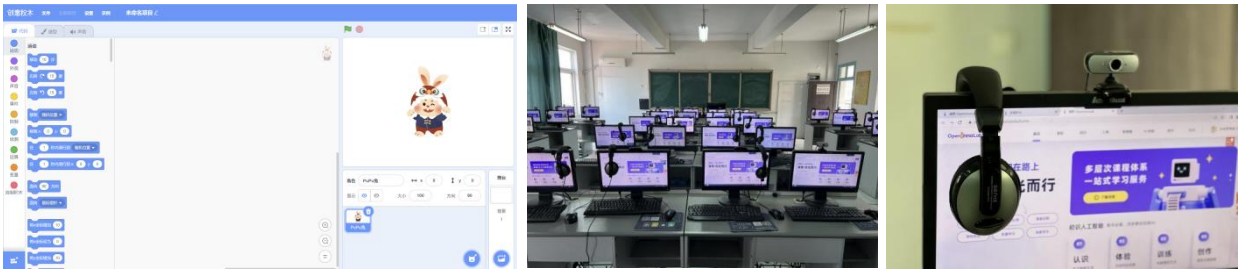


图 1 浦育平台“创意积木工坊”界面和学校机房配备 45 个耳麦和摄像头

其次，“浦育平台”允许创建群组，为学生建立独立账号，使他们可以在线进行编程创作并分享项目，同时“创意积木工坊”采用图形化的编程方式，更适合小学段学生进行编程创作。

最后，学生在使用浦育平台的“创意积木工坊”时，能够学习并亲身体验数据采集、数据集制作、分类模型训练以及程序编写的整个过程，无需依赖第三方平台。

综合以上分析，笔者选择浦育平台的“创意积木工坊”开展《红领巾检查小助手》课程的课堂教学，使学生能够在简化的学习流程中了解、学习和体验深度学习技术。

2.2.精研学习课程，推动解决实际问题

2.2.1.搭建课程内容框架

根据学生的信息科技学习水平及认知发展阶段，该课程项目更倾向于面向三四年级的社团课学生。尽管这两个年级的学生在信息科技课上尚未接触编程，但他们对电脑的基本操作掌握得还是不错的，具备学习该课程所需的基本能力。为此，笔者精心设计了包含 6 个课时

的学习任务（图 2）。首先，通过生动有趣的项目分析环节，激发学生对即将展开项目的兴趣与好奇心，并帮助他们明确学习目标与预期成果。随后，利用浦育平台的“创意积木工坊”编程学习工具，结合摄像头设备，引导学生学习图像分类与列表存储的知识，以此为基础完成项目实现。最后，组织项目展示环节，让学生在相互体验与评价中进一步优化和完善项目。

图 2《红领巾检查小助手》课程项目思维导图

学生通过浦育平台的“创意积木工坊”，借助“AI 训练（图像分类）”模块，学习课程并设计制作《红领巾检查小助手》编程项目。该项目解决了值日生检查学生是否佩戴红领巾时可能出现的遗漏和记录繁琐的问题。在课程项目创作的过程中，学生成为了校园的主人，学会了使用智慧、科学的方式来解决校园管理中的实际问题。学习课程能够培养学生解决复杂问题的实践能力和创新能力(王艳芳等人，2018）。

2.2.2.实践课程具体内容

●项目分析

由于项目情境源自学生在校期间的日常行为规范，学生在遇到“值日生”困惑时能够快速产生共鸣，从而迅速进入项目学习状态。笔者顺势提出课程的驱动性问题，并引导学生展开头脑风暴，将他们的功能设想记录在导学单上。学生分享他们的想法（图 3），笔者指导他们概括《红领巾检查小助手》的功能效果，并展示课程项目的思维导图，以帮助学生理清项目创作思路。

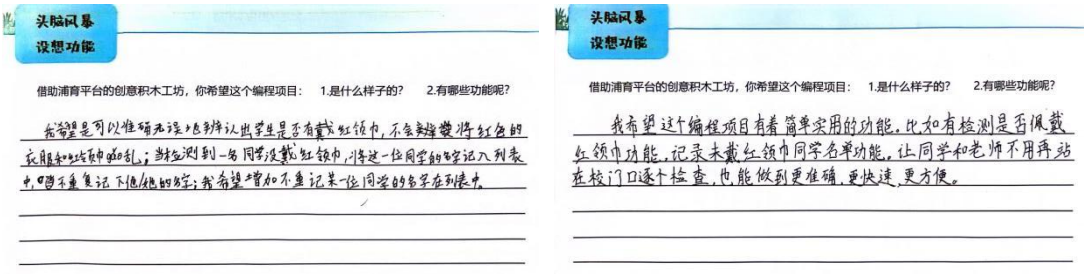


图 3《项目分析》课时部分学生头脑风暴填写情况

●项目实现：图像分类

学生通过回顾上节课梳理的思维导图以及头脑风暴后确定的两大必要的作品项目功能，即：1.区分学生是否佩戴红领巾。2.自动获取未佩戴红领巾的学生姓名。本课的核心问题聚焦于：如何区分学生是否佩戴红领巾？为此，笔者设计了 4 个教学环节，每个环节都围绕一个活动任务展开，以引导学生深入探讨和解决本课的核心问题。其中，笔者重点突破以下内

容：AI 图像分类模型的训练方法、程序编写实现，以及变量的定义与应用，最后查看图像分类结果。

①AI 训练（图像分类）

在该环节中，学生先是通过观看视频了解何为图像分类，图像分类的一般过程。笔者会继续引导学生思考：若要区分 1 位学生是否佩戴红领巾，需要分几类？紧接着，学生通过自主探索知道了通过“扩展”按钮➡添加“AI 训练（图像分类）”模块，在教师的指导下学会单击“打开 AI 训练工具”进行模型训练。最后，学生完成任务二：使用“图像分类”模块，根据分类标签命名要求，分别训练两位同学，戴红领巾和没戴红领巾的情况（图 4）。

图 4《项目实现：图像分类》课时 AI 训练图像分类教学过程

②编写程序

学生通过填空流程图的方式，确定使用图像分类模型区分学生是否佩戴红领巾的算法设计。过程中，笔者引导学生理解置信度，帮助学生更好地完成流程图，梳理作品的创作思路，完成任务三：如果学生佩戴红领巾，提示：请进校园如果学生未佩戴红领巾，提示：姓名没戴红领巾。

③认识变量，查看结果

学生在调试程序时，会出现无法显示输出结果的情况。笔者通过“每天吃掉两个糖果盒子里的糖果”的情境，让学生去感知变量是：会变化的量（或者是：数值），并拓展延伸出：常量。通过这样的方式，引导学生更好地理解：何为变量？何为常量？以及变量的组成部分。她再指导学生完成预测结果揭秘的任务：通过新建“预测结果”变量，查看通过图像分类，预测出的图像类型。学生就能够根据识别结果，修改程序。

学生借助本课的导学评价单，学习掌握角色与背景的添加，训练与使用图像分类模型，认识与应用变量，多分支结构的使用等相关知识，编写程序解决“如何区分学生是否佩戴红领巾？”的核心问题。

●项目实现：列表存储

本课的核心问题聚焦于：如何自动获取未佩戴红领巾的学生名单？学生通过再次回顾作品项目的功能设想，笔者引导学生发散思维，寻找解决问题的方法。为解决本课的核心问题，设计了 3 个教学环节，同样每个环节都围绕一个活动任务展开。

①认识新建列表

与变量一样，学生对于列表也是初次接触，非常陌生。笔者借助“对抽屉的操作”模拟列表的替换、删除与插入等基本操作，帮助学生熟悉列表，明白列表是：能存放许多变量的一种容器。可以存储或者获取容器中任意一个变量。学生会感知到：变量与列表是存在联系的。同时，教授学生列表的组成部分，再指导学生完成任务一：新建“没戴红领巾学生名单”列表，用于存放没有佩戴红领巾的学生名单。

②编写程序

学生观察列表的各个指令，思考自动获取未佩戴红领巾的学生名单的方法。再由笔者引导学生借助呈现的思维导图，梳理创作思路：①先通过判断语句：如果……那么……和“运算”指令类里的关键词查找“……包含……？”指令相结合，找到未佩戴红领巾的学生。②然后使用“运算”指令类中的“……的第……个字符”和“连接……和……”指令从未佩戴红领巾的预测结果中自动获取学生姓名，通过列表的“将……放入……”指令将自动获取的学生姓名存放在“没戴红领巾学生名单”列表中。学生再明确编程思路后，完成任务二：实现自动获取预测结果中没有佩戴红领巾的学生名单（图 5）。

图 5《项目实现：列表存储》课时编写程序教学过程

③添加声音，优化程序

考虑到《红领巾检查小助手》的应用场景是在学生入校期间,所以项目可以增加语音播报功能。这样一来,当检测到学生佩戴红领巾时,就可以通过语音直接提示“请进校园”；当检测到学生未佩戴红领巾时,则可以语音播报“XXX，没戴红领巾”。这种直接的语音提醒,不仅能引起学生的注意,也更加贴近实际应用场景,为学生提供及时有效的提示，提高学生的自律意识。

学生借助本课的导学评价单，学习掌握认识与应用列表，增加声音指令等相关知识，编写程序解决“如何自动获取未佩戴红领巾的学生名单？”的核心问题。

●项目展示

本课是课程项目的最后一节课，学生充分体验项目，对自己的作品进行评价。并且将自己创作的项目向所有同学进行展示。学生在体验、评价与展示项目后，思考：1.项目还可以增加哪些功能？ 2.项目还可以解决哪些问题？ 将想法填写在导学单中，根据填写的设想，尝试修改程序，实现项目的迭代升级（图 6）。

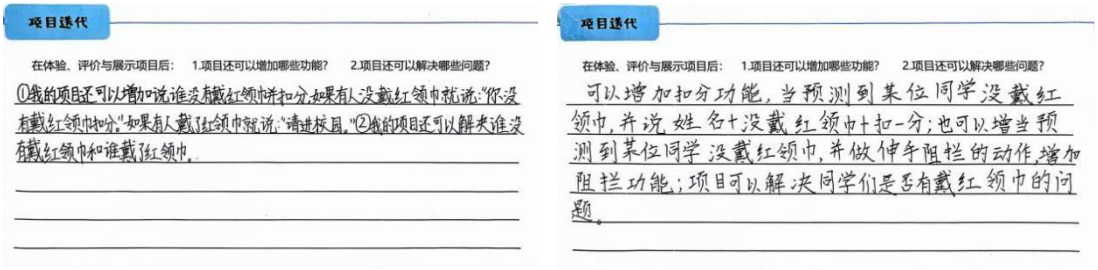


图 6《项目展示》课时部分学生项目迭代填写情况

2.2.3.整合课程学习资源

笔者在《红领巾检查小助手》课程项目教学圆满结束后，并未止步，而是倾注了大量宝贵的时间与精力，深入细致地整合并优化了该课程的全部学习资源，课程资源涵盖了详尽的教学设计，帮助学生和教师更好地理解课程框架与教学目标；包含了生动直观的教学 PPT，

以图文并茂的形式展现课程重点与难点；此外，还提供了完整的程序源代码，供学生参考与学习，以便他们更深入地理解编程逻辑与实践应用；同时，导学评价单等辅助材料也一应俱全，帮助学生自我检测学习进度与成果；以及动画微课视频，帮助学生突破时空限制，随时巩固知识点，并通过动态演示强化理解（图 7）。这一整套丰富、多元的学习资源，可以极大地提升学生的学习体验，满足不同学习风格和需求，助力他们更全面地掌握课程知识，提升信息科技的综合实践能力。

图 7《红领巾检查小助手》课程学习资源

2.3.迭代项目功能，推动现实场景应用

针对《红领巾检查小助手》课程项目，笔者认为还有很多可以拓展优化，迭代延伸的方向：1.项目本身：应用在校园中，加减分，音效，硬件设备等功能。2.日常行为规范：增加是否穿着校服的检查、卫生检查、学生离校班牌管理等功能。3.AI 校园管理：校园安全监控、学习辅助应用、学生考勤管理、校园器材设施管理、学生行为管理等方面。

为了实现项目落地，笔者进一步扩展《红领巾检查小助手》项目，通过 Mind+支持模型推理的功能和简单的设备搭建，该项目成功应用于校园现实场景（图 8），同样有效解决了值日生在检查学生是否佩戴红领巾时可能出现的遗漏和记录繁琐的问题。

图 8《红领巾检查小助手》项目校园实际应用场景

具体操作流程如下：笔者首先通过浦育平台的在线可视化工具训练模型，获得模型文件；接着，利用浦育的在线算力将该模型文件转换为.onnx 格式的通用模型；之后，在 Mind+的行空板上，使用 xedu-hub 图形化扩展进行镜像推理，从而得到图像分类的结果。最终，《红领巾检查小助手》项目在 Mind+环境下，结合 XeduHub 和行空板，通过摄像头实现图像的实时识别，将图像分类结果显示在行空板上，并终端输出“未佩戴红领巾的学生名单”（图 9）。

图 9《红领巾检查小助手》项目校园实际应用场景功能演示

3.以探索成效为动力

3.1.技术融合改变校园管理方式

良好的行为规范是一个孩子成长成才的基石，也是学校德育教育的基础（张小玉等人，2021）。学校的值日生对学生红领巾佩戴情况进行检查，是规范小学生日常行为的一种极为基础的校园管理方式。然而，传统的管理方式主要依赖值日生的眼睛观察和纸笔记录，这种方式很容易出现遗漏或记录繁琐的问题。当传统的管理方式融合了新一代人工智能技术，笔者设计出了《红领巾检查小助手》项目并实现了项目在地化，它能够有效地提高学校红领巾检查的效率，并增强学生的使命感和自律意识。

3.2.技术融合助力课程资源推广

笔者有幸在谢作如名师工作室举办的《中小学 AI 开源课程众筹计划》（第六期）活动中，进行了题为《AI 校园管理——红领巾检查小助手》课程的线上分享（图 10）。分享中，笔者聚焦于乡村小学在新一代人工智能教育方面的课程资源和实践经验，该内容引发了来自各学校教师的支持与鼓励，也引起了其他乡村小学教师的共鸣和思考。课程资源的推广打破了地域和时间的限制(UNESCO，2021)，使得更多师生能够从中受益，尤其是在乡村小学等教育资源相对匮乏的地区，这种推广显得尤为重要。

《AI校园管理——红领巾检查小助手》资料

图 10《红领巾检查小助手》课程项目线上分享

3.3.技术融合提升师生综合能力

在《红领巾检查小助手》课程项目实施后，无论是教师（图 11）还是学生（图 12）都收获颇丰，在各方面均取得了显著的成就和荣誉，提升了师生在信息科技方面的综合能力。

序号	学生姓名	获奖项目	奖次	颁发单位
1	优秀指导师	2023 年全国青少年人工智能创新实践活动	全国优秀指导师	中国科协青少年科技中心
2	论文、专题发言	《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》	第 28 届全国华人计算机教育应用大会	第 28 届全国华人计算机教育应用大会
3	公开课	《新一代人工智能教育——AI 图像分类模型训练与应用》	市公开课	温州市教育教学研究院
4	经验交流	《人工智能在农村小学的教学实践与探索》	市经验交流	温州市教育教学研究院
5	精品课例资源	《英语单词认读器》	市三等奖	温州市教育教学研究院
6	优秀指导师	2024 年平阳县师生信息素养提升实践活动	县优秀指导师	平阳县教育局
7	专题发言	《人工智能核心素养在课堂中的探索与实践》	县专题发言	平阳县教师发展中心
8	案例	《探索农村小学智慧校园管理新路径——基于浦育平台《红领巾检查小助手》的新一代人工智能教育实践》	市二等奖	温州市教育教学研究院
9	案例	《探索农村小学智慧校园管理新路径——基于浦育平台《红领巾检查小助手》的新一代人工智能教育实践》	县一等奖	平阳县教育局办公室
10	讲座	《面向小学的 AI 入门课程开发》	市讲座	温州市教师教育院

图 11 教师获奖情况汇总

表 12 学生获奖情况汇总

序号	学生姓名	获奖项目	奖次	颁发单位
1	谢语晨、谢尚硕	AI 交互设计	全国优秀作品	中国科协青少年科技中心
2	胡梦婷	AI 艺术生成	全国优秀作品	中国科协青少年科技中心
3	徐珲鸽	AI 艺术生成	全国优秀作品	中国科协青少年科技中心
4	谢语晨、谢尚硕	AI 交互设计	全国“四星卓越”创新作品	中国科协青少年科技中心
5	高凌轩、邱俨晞	创意编程（专项）	县一等奖	平阳县教育局
6	朱欣欣、邓羽辰	创意编程（专项）	县一等奖	平阳县教育局
7	高凌轩、邱俨晞	创意编程 项目	市三等奖	温州市教育技术中心

4.结语

笔者以学校值日生遇到的真实问题解决为导向（谢作如，2024），顺应人工智能教育新形势，设计并实施了《红领巾检查小助手》课程项目。平阳县昆阳镇第六小学作为试点学校的实践表明，该项目不仅有效解决了学校日常管理中的实际问题，还显著提升了学生对人工智能技术的理解与应用能力，激发了他们对科技创新的兴趣。该项目的成功实施也为乡村小学在资源有限条件下探索人工智能教育的新模式提供了可行路径。未来，笔者将继续深化人工智能与教育的融合，不断拓展项目的应用场景，为乡村小学的智慧校园管理注入更多活力与创新。

参考文献

温州日报.（2022）.市教育局发布《温州市中小学推进人工智能教育实施方案》.温州市人民政府.https://www.wenzhou.gov.cn/art/2022/3/9/art_1217833_59154015.html

王艳芳,李正福,祁荣宾,等.（2018）.项目驱动，破解中小学人工智能教育难题.基础教育课程,(09):28-32.

张小玉,李德功.（2021）.乡村小学生日常生活守则和行为规范养成教育策略研究.科幻画

报,(01):61-62.
UNESCO.(2021).AI and education: Guidance for policy-makers.United Nations Educational,Scientific and Cultural Organization.https://unesdoc.unesco.org/ark:/48223/pf0000376709
谢作如.（2024）.用人工智能解决真实问题的技术地图.中国信息技术教育,(01):85-87.

融合范例学习与多智能体的高中 STEM 课程设计与实践

Design and Practice of High School STEM Curriculum Integrating Few-Shot Learning and Multi-Agent Systems

孟晓培[1]，池家义[1]，涂芸芳[1*]
[1]温州大学 教师教育学院
[*]sandy0692@gmail.com

【摘要】近年来，STEM 教育在全球教育改革中日益受到关注，其主要强调跨学科知识与现实问题解决。问题解决能力的培养需要程序性知识与概念性知识的结合。尽管基于范例的学习已被证实为有效的教学策略，但学生在实践中往往面临范例资源匮乏和反馈滞后的挑战。多智能体系统凭借协作性、自适应性和分布式优势，能够为学生提供持续且个性化的学习支持。基于此，本研究设计并开发了一套融合范例学习与多智能体系统的 STEM 课程，并在高中课堂中成功应用。该课程设计为推动高中阶段人工智能与 STEM 课程深度融合提供了实践示范与参考。

【关键词】STEM 教育；范例学习；多智能体；人工智能工具

Abstract: *In recent years, STEM education has received increasing attention in global education reform, emphasizing the integration of interdisciplinary knowledge and the ability to solve real-world problems. The development of problem-solving skills requires a combination of procedural and conceptual knowledge. Although example-based learning has been proven to be an effective instructional strategy, students often face challenges such as limited access to example resources and delayed feedback during practice. Multi-agent systems, with their collaborative, adaptive, and distributed characteristics, can provide continuous and personalized learning support for students. Based on this, the present study designed and developed a STEM curriculum integrating example-based learning and multi-agent systems, which has been successfully implemented in high school classrooms. This curriculum design offers practical guidance and a reference for promoting the deep integration of artificial intelligence and STEM education at the K–12 level.*

Keywords: STEM education; Example-based learning; Multi-agent systems; AI tools

1.前言

STEM 教育在近年来的全球教育改革中逐渐崭露头角。其目标是通过结合科学、技术、工程和数学等学科及其实际应用，帮助学生解决现实世界中的复杂问题(Schnittka et al., 2016)。最近，多项研究已证明了 STEM 教育的巨大潜力(A dedokun et al., 2013; Sahin et al., 2015) 。此外，问题解决任务是许多学科学习活动的重要组成部分，尤其是在 STEM 领域。解决问题需要掌握两类关键知识：程序性知识，即操作步骤的执行方式；以及概念性知识，即操作背后的原理(Giang et al., 2024)。

基于范例的学习是一种观察学习的形式。近年来，范例学习在非正式和正式学习环境中具有广泛应用，促进了其在不同教育场景中的普及(Hoogerheide& Roelle, 2020)。研究表明，基于范例的学习是一种高效的知识获取方式，同时也是习得新技能的有效教学策略，通过学习范例，学生能够不仅掌握“如何做”，还能够理解“为什么这样做”以及“如何解决类似

问题”(Dyer et al., 2015)。在 STEM 教育中，许多学生难以独立找到并有效应用范例资源，未能及时获得反馈可能导致他们失去学习兴趣或自我效能感下降，这对教师提出了更高的要求，特别是在如何提供个性化反馈、灵活调整教学策略以及满足学生多样化学习需求方面(Lesseig et al., 2016; Ramey& Stevens, 2023)。

作为引导和支持学生持续学习的有效工具，智能体能够在学生遇到困难时提供及时的个性化反馈，帮助他们克服学习障碍(Kim& Baylor, 2016)。然而，单一智能体系统在应对任务规模和复杂度增加时面临着扩展性问题，且在任务分解、管理和适应动态环境方面存在瓶颈，缺乏足够的专业化和灵活性(Ishibashi& Nishimura, 2024; Jiang et al., 2024)。多智能体系统(Multi-Agent Systems, MAS)是由多个相互作用的智能体组成的计算系统，每个智能体都具有独立且清晰的内部结构，能够在系统中独立运作并实现协同工作(Hong et al., 2023)。根据已有的研究，多智能体凭借其协作、分布式和自适应等特点，通过信息共享和任务分工，有助于提升大语言模型的思考与自我反思能力，从而为学生提供更精准的指导和支持，并帮助学习者培养独立完成任务和自我思考的能力(Puntambekar, 2022)。因此，本研究结合范例学习和多智能体策略，开发了针对高中生 STEM 学习的课程，并在一所高中开展了实践。

2.融合范例学习与多智能体系统的高中 STEM 课程设计

2.1.mPython 编程平台

mPython 是一款强大的编程平台，集成硬件编程、Python 学习和 JupyterLab 环境，支持一键安装，方便快速搭建开发环境。它帮助学生在 microPython 和 Python 之间平滑过渡，提供丰富示例程序，助力初学者理解编程逻辑。自带掌控板仿真功能，让学生无需实体硬件即可在线运行和调试代码。双屏联动教学模式实时同步代码与运行界面，帮助学生直观理解程序结构和效果，使学习更加高效。

2.1.多智能体开发与部署的 coze 平台

扣子是新一代 AI 应用开发平台，专为 AI 应用（涵盖多智能体）开发者打造，提供从应用开发、测评、监控到多渠道发布的全方位、一站式服务。本研究依托 coze 平台，设计并开发了一套专为 STEM 课程教学打造的多智能体，如图 1 所示，并邀请了部分高中生进行实际体验与测试。

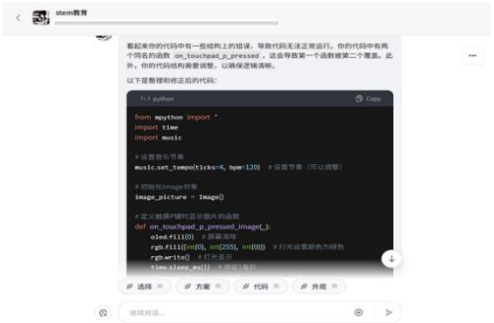


图 1 STEM 教育多智能体

2.3.课程目标定位

本课程融合范例学习与多智能体系统，面向高中阶段的基于设计的 STEM 学习，紧密契合《普通高中信息技术课程标准（2017 年版 2020 年修订）》中关于学科核心素养与课程目标的要求。课程由教师在课前提供高质量的范例，帮助学生建立初步的问题认知和设计框架，培养学生主动获取信息、分析信息价值和提炼关键信息的能力，发展信息意识。学生在观看范例后，以小组为单位开展作品设计与 mPython 编程实践、手工制作，过程中自主界定问题、

抽象建模、优化设计，并通过编程在掌控板上实现功能，全面促进计算思维的发展。在整个实施阶段，学生通过与多智能体系统互动，获得持续的问题指导、创意提示和反馈建议，进一步激发创新思维，合理选择数字化工具，高效组织团队协作，培养数字化学习与创新能力。整个学习过程鼓励学生在解决实际问题中遵循信息伦理规范，理解信息技术对社会与环境的影响，增强信息安全意识和社会责任感。通过“范例引导—自主设计—多智能体支持—作品实现”这一完整路径，学生不仅掌握数据、算法和信息系统知识，还能在真实场景中锻炼系统化解决问题的能力，形成多元视角和创新意识，成为具有计算思维、创新能力和社会责任感的数字化时代合格公民。

2.4. 课程内容架构

本研究根据五步 STEM 课程设计框架(Hsu& Fang, 2019)构建了一套融合范例学习与多智能体系统的高中 STEM 课程内容架构，如图 2 所示，以期通过系统性、创新性的课程设计，培养学生的核心能力，激发其对 STEM 领域的深度兴趣与探索精神。

在课程设计的起始阶段，精准识别 STEM 核心能力是构建整个课程体系的逻辑起点。本课程聚焦于问题解决能力、反思能力以及设计美学等关键能力的培育，这些能力不仅契合 STEM 教育的内在要求，更是学生在未来多元学习情境与职业发展路径中不可或缺的素养根基。通过对这些核心能力的明确定义与细致剖析，为后续课程内容的开发与实施提供了精准的指向标，确保每一步设计都紧密围绕能力培养目标展开。

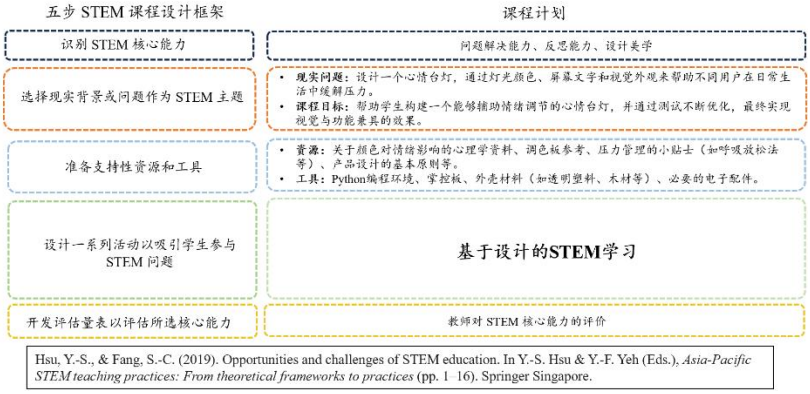


图 2 五步 STEM 课程设计框架

选择现实背景或问题作为 STEM 主题是激发学生内在学习动机的关键策略。现实的背景不仅让学生直观感受到 STEM 知识的实用性，更能在解决实际问题的过程中让学生体验到学习的成就感与价值感，进而增强学习的主动性和积极性，培养其将所学知识应用于现实情境的能力。基于此，整个课程选取了三个现实背景或者问题，并进行了系统的课程计划，整个课程计划的制定遵循连贯性与进阶性的原则，精心规划各个学习阶段的内容与目标，使知识体系与技能训练呈现出由浅入深、循序渐进的态势。如表 1 所示，整个课程共包含 7 个课时，分为四个主要模块，每个模块都结合了不同的范例学习类型，以促进学生的深度理解和应用能力。

表 1 融合范例学习与多智能体系统的高中 STEM 课程课程内容架构大纲

课时	内容大纲	范例类型
第一课时	编程环境入门与掌控板初体验	
第二、三课时	心情台灯	工作范例(Worked Example)
第四、五课时	智能陪伴装置-为独居老人设计	完成范例(Completion Example)
第六、七课时	电子班级名牌	综合活动(Integrative Activity)

第一课时是编程环境入门与掌控板初体验，这一基础模块旨在让学生熟悉编程工具和硬件设备的基本操作，为后续的项目实践奠定坚实的基础。学生将通过实际操作，了解编程环境的界面布局、基本语法和调试技巧，同时掌握掌控板的连接、配置和简单控制方法。

第二、三课时聚焦于设计一个心情台灯，通过灯光颜色、屏幕文字和视觉外观帮助用户缓解压力。这一模块采用工作范例(Worked Example)形式，详细展示从问题定义到原型制作的完整过程。工作范例是基于范例学习中最常见的一种形式，通常由详细的解题过程和步骤组成。它通过展示解决问题的完整过程，帮助学生理解如何从问题的初始状态到达最终解决方案。在这一过程中，学生不仅能看到问题的解答，还能了解每一步的推理过程，包括具体的提示、策略和知识点(Dyer et al., 2015; Hoogerheide& Roelle, 2020)。例如，在设计心情台灯的过程中，教师会逐步展示如何分析用户需求、选择合适的硬件组件、编写控制程序以及测试和优化设计。通过这种方式，学生可以清晰地看到每一个步骤如何执行，从而更好地理解和掌握相关的 STEM 原理和方法。

第四、五课时转向智能陪伴装置的设计，专为独居老人打造，涉及对老年人生活需求的深入理解与相应功能的开发。这一部分采用完成范例(Completion Example)形式，在提供部分解决方案的基础上，鼓励学生主动思考和应用所学知识来完成设计。完成范例的主要特点是在部分解题过程上留白，要求学生根据已给出的部分信息完成剩余的解题步骤。这种范例形式能够激发学生的主动性和创造性，帮助他们在实践中巩固和深化对所学内容的理解(Hoogerheide& Roelle, 2020)。在智能陪伴装置的设计中，教师会先提供一个部分完成的设计方案，例如已经设计好的基本框架或部分功能模块，然后让学生根据这个基础，进一步完善设计，添加更多功能或优化现有功能。这种方式不仅提高了学习的互动性，还能够帮助学生逐步建立问题解决的信心和能力。

第六、七课时则围绕电子班级名牌展开，涵盖了班级文化展示、信息传递等功能的设计与实现。这一模块作为综合活动(Integrative Activity)，要求学生综合运用前几课时学到的知识和技能，解决更为复杂的问题，培养其跨学科思维和知识整合能力。这种综合性的项目不仅能够提高学生的综合思维和实际应用能力，还能培养他们的团队合作精神和项目管理能力(Renkl, 2014)。

在准备支持性资源和工具方面，本课程为学生提供了丰富多元的学习素材与实践工具。涵盖教育、心理等多方面的理论资源，为学生的设计构思提供坚实的理论支撑；而 Python 编程环境、掌控板、多样化的外壳材料及电子配件等实践工具，则为学生的创意实践搭建了广阔的平台，使其能够将理论设想转化为实际可操作的作品，切实提升动手能力与创新思维，实现知识的迁移与应用。

设计一系列活动以吸引学生参与 STEM 问题是课程实施中的关键环节。通过编程实践、模型搭建、团队协作等多元化的活动形式，营造生动活泼且富有挑战的学习氛围。本课程选择按照基于设计的 STEM 学习模式作为参考。在基于设计的 STEM 学习中，学生需经历完整的六步流程，如图 3 所示：从明确问题定义，精准把握用户需求与使用场景；到大胆进行方案构思，绘制设计草图，提出多种创新性解决方案；再到细致入微的设计规划，确定具体的制作步骤、材料选用与项目时间表；随后投入制作原型，借助编程环境与掌控板等工具将设计变为现实；之后进行严格的测试与优化，根据实际使用反馈不断改进产品性能；直至最终的展示与分享，向同伴与教师呈现自己的创意成果。这一过程融合了反思和协作。

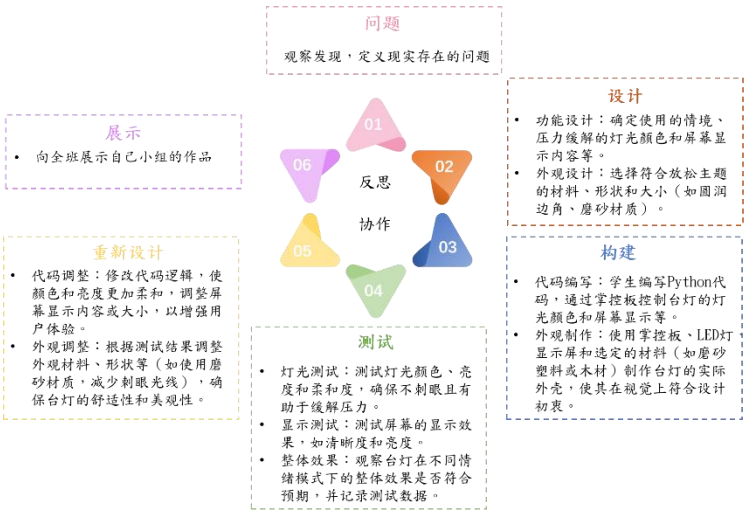


图 3 基于设计的 STEM 学习模式

本课程多维度、精细化评价学生 STEM 核心能力，真实反映学习成果与能力发展水平。教师持续观察学生表现，结合评估量表进行个性化评价，提供针对性反馈与指导。

3.融合范例学习与多智能体系统的高中 STEM 课程实践案例展示

课程开发完成后，本研究在温州艺术学校高一年级掌控板课程中开展教学实践，每周一课时。以下以第二、三课时“心情台灯”为例，具体阐述教学实施过程，如表 2 所示。

表 2 心情台灯教学流程实施过程

阶段	教师	学生	多智能体
课前准备	• 确定课程目标与学生需求，选择情绪调节主题 • 资源准备	• 阅读关于颜色对情绪的影响资料	• 提供关于颜色心理学、产品设计的基础知识
情境导入	• 介绍多智能体的概念，并阐明其在实现课程目标中的作用 • 导入“心情台灯”问题情境	• 进行头脑风暴，提出设计方案的初步想法	• 提示学生不同情境下心情台灯的功能需求
设计阶段	• 播放范例视频，展示设计框架与实施过程 • 解决课堂问题	• 设计台灯的功能、外观与情境应用	• 提供创意建议与设计灵感，协助思考功能与外观
构建阶段	• 范例演示如何编写代码并进行硬件连接	• 编写代码、调试程序，并开始构建台灯外壳	• 提供代码示例、解答编程问题，支持调试过程
测试阶段	• 指导学生进行测试，确保灯光效果与显示效果符合设计要求，并提供必要的调整建议。	• 进行灯光颜色、亮度等测试，并记录数据	• 分析测试结果并建议优化
重新设计阶段	• 引导学生根据测试结果优化设计	• 调整代码与外观，确保符合设计目标	• 提供优化建议，推荐改进方案
展示反思阶段	• 组织展示与总结，引导学生思考应用前景	• 展示成果，反思项目中的挑战与改进	• 提供反思模板，帮助学生总结设计经验与收获

在课前准备阶段，教师首先识别本项目所需发展的 STEM 核心能力，包括问题解决能力、设计思维、反思能力、计算思维与美学素养。随后，教师选择贴近学生生活的现实问题作为课程主题——设计一个心情台灯，旨在通过灯光颜色、屏幕文字和视觉外观帮助不同用户在日常生活中缓解压力。

为支持学生顺利完成任务，教师精心准备了支持性资源与工具。资源包括：颜色对情绪影响的心理学资料、调色板参考图、压力管理技巧（如呼吸放松法）、产品设计基本原则及范例视频。工具方面则配备了 Python 编程环境、掌控板、外壳材料（如透明塑料、木材等）及必要电子配件。

在课堂实施阶段，教师设计了基于设计学习的完整活动路径，并借助多智能体系统和范例教学策略贯穿全过程提供动态支持。

首先在情境导入阶段，教师通过现实案例引导学生思考“如何设计一款帮助调节情绪的台灯”，并开展小组头脑风暴。多智能体系统此时提供启发性提示和关联案例，帮助学生扩展思路。学生在此阶段完成了若干反思性和创造性练习，例如“我们如何借助多智能体实现更好的功能构思？”

在设计阶段，教师播放自主录制的范例视频，详细展示心情台灯的情境设计、功能设计与外观设计案例，如图 4 所示。同时，学生结合范例视频和首次项目概念图，完成学习单中的自我解释任务。随后，学生借助多智能体可以产生创意想法的功能，在小组讨论中提出了与现实需求高度相关的台灯设计方案，包括不同情境下灯光模式、动态显示文字内容及外观造型创意。

图 4 学生进行心情台灯设计范例视频学习

构建阶段，教师再次播放范例视频，讲解三个主要功能模块的代码结构、逻辑关系及实现步骤。学生在小组内分工协作，依据概念图和示例代码完成程序编写、代码调试与掌控板部署。同时，进行台灯外观加工制作。此阶段多智能体系统为学生提供代码智能讲解、调试提示和材料选用建议，帮助学生在遇到困难时及时获得指导与及时性反馈，如图 5 所示。在此过程中，反思与协作持续嵌入，学生通过学习单记录阶段性问题与调整策略，团队内相互协助解决技术难题。

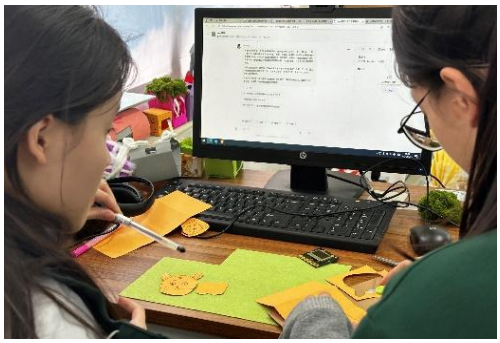


图 5 学生在根据多智能体的建议制作心情台灯外观

测试阶段，学生小组协作完成对成品的灯光测试（颜色、亮度、柔和度）、屏幕显示测试（清晰度、亮度）、整体效果评估（不同情绪模式下的感官体验），并通过数据记录表收集反馈。多智能体系统此时提供数据分析辅助与改进建议。

在重新设计阶段，学生根据测试数据与反思记录，对代码逻辑进行调整，如图 6 所示，使颜色与亮度表现更加柔和；同时根据用户体验优化屏幕内容显示、调整字体大小。外观方面，学生根据测试结论改进材料与造型，例如调整外形线条以增加美观度与舒适性。

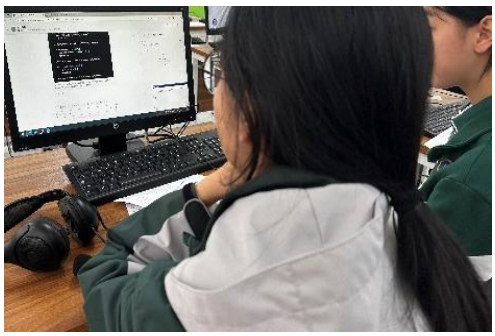


图 6 学生在询问多智能体继续优化代码

最终在展示阶段，各小组向全班展示设计成果，如图 7 所示，接受同伴与教师的质询。教师引导学生结合多智能体提供的历史记录与反馈报告，反思整个设计过程中的学习收获与能力提升，并鼓励学生思考如何将所学迁移到更广泛的社会与企业应用中。

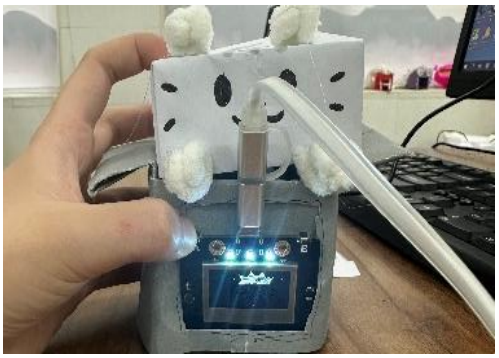


图 7 学生的心情台灯作品

整个项目实践过程中，教师通过范例学习引导设计思路，多智能体系统提供智能支持与个性化提示，反思与协作作为贯穿始终的过程性元素被嵌入每一阶段，支持学生在不断循环的“设计—实施—测试—再设计”路径中逐步实现创新性成果。

4.结语

本研究在《普通高中信息技术课程标准（2017 年版 2020 年修订）》的引领下，融合了范例学习与多智能体系统，创新性地为高中生设计了一套 STEM 课程。此课程通过 mPython 编程平台与 coze 多智能体开发平台，精心构建了七个循序渐进的课时，涵盖从编程入门到复杂项目设计的全过程。学生不仅掌握了 STEM 核心知识，还在真实情境中锻炼了问题解决与创新能力。然而，课程实施中仍存在资源分配、隐私等方面的挑战。未来，我们将根据反馈持续优化课程，强化跨学科融合，以期培育更多具备计算思维与创新能力的数字化时代公民。

参考文献

Adedokun, O. A., Bessenbacher, A. B., Parker, L. C., Kirkham, L. L., & Burgess, W. D. (2013). Research skills and STEM undergraduate research students' aspirations for research careers: Mediating effects of research self-efficacy. *Journal of Research in Science teaching*, *50*(8), 940-951.

Dyer, J. O., Hudon, A., Montpetit-Tourangeau, K., Charlin, B., Mamede, S., & Van Gog, T. (2015). Example-based learning: comparing the effects of additionally providing three different integrative learning activities on physiotherapy intervention knowledge. *BMC medical education, 15*, 1-16.

Giang, N. T. C., Anh, N. T. Q., Dao, T. T., Tuan, P. A., Linh, C. T. H., & Chau, P. T. H. (2024). A systematic review of problem-solving skill development for students in STEM

education. *International Journal of Learning, Teaching and Educational Research, 23*(5), 1-20.
Hong, S., Zheng, X., Chen, J., Cheng, Y., Wang, J., Zhang, C., ... & Wu, C. (2023). Metagpt: Meta programming for multi-agent collaborative framework. *arXiv preprint arXiv:2308.00352, 3*(4), 6.
Hoogerheide, V., & Roelle, J. (2020). Example-based learning: New theoretical perspectives and use-inspired advances to a contemporary instructional approach. *Applied Cognitive Psychology, 34*(4), 787-792.
Hsu, Y.-S., & Fang, S.-C. (2019). Opportunities and challenges of STEM education. In Y.-S. Hsu & Y.-F. Yeh (Eds.), Asia-Pacific STEM teaching practices: From theoretical frameworks to practices (pp. 1–16). Springer Singapore.
Ishibashi, Y., & Nishimura, Y. (2024). Self-organized agents: A llm multi-agent framework toward ultra large-scale code generation and optimization. *arXiv preprint arXiv:2404.02183.*
Jiang, Y. H., Li, R., Zhou, Y., Qi, C., Hu, H., Wei, Y., ... & Wu, Y. (2024). AI Agent for Education: von Neumann Multi-Agent System Framework. *arXiv preprint arXiv:2501.00083.*
Kim, Y., & Baylor, A. L. (2016). based design of pedagogical agent roles: A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education, 26*, 160-169.
Lesseig, K., Nelson, T. H., Slavit, D., & Seidel, R. A. (2016). Supporting middle school teachers' implementation of STEM design challenges. *School Science and Mathematics, 116*(4), 177-188.
Puntambekar, S. (2022). Distributed scaffolding: Scaffolding students in classroom environments. *Educational Psychology Review, 34*(1), 451-472.
Ramey, K. E., & Stevens, R. (2023). Dilemmas Experienced by Teachers in Adapting to the Role of Facilitator in the STEAM Classroom. *Teaching and teacher education, 133*, 104271.
Renkl, A. (2014). Toward an instructionally oriented theory of example‐based learning. *Cognitive science, 38*(1), 1-37.
Sahin, A., Gulacar, O., & Stuessy, C. (2015). High school students' perceptions of the effects of international science Olympiad on their STEM career aspirations and twenty-first century skill development. *Research in Science Education, 45*, 785-805.
Schnittka, C. G., Evans, M. A., Won, S. G., & Drape, T. A. (2016). After-school spaces: Looking for learning in all the right places. Research in Science Education, 46, 389-412.

一项基于 AI 的多语同步口语平台在新加坡小学的应用探索

“Learn”: An AI-Based Synchronous Multilingual Oral Learning Platform for Primary Schools in Singapore

胡月宝，林庆荣
南洋理工大学国立教育学院亚洲语文与文化学部
guatpoh.aw@nie.edu.sg
khengyong.lim@nie.edu.sg

【摘要】 随着人工智能，尤其是机器学习在教育领域的快速发展，智能教材的设计逐渐转向以学习者为中心的个性化方向。“乐言”是由新加坡理工大学与南洋理工大学联合开发的图画驱动 AI 口语学习平台，专为小学低年级学生设计，支持华语、马来语和淡米尔语的多语同步学习。平台融合自然语言处理、语音识别与个性化反馈机制，提供涵盖发音、词汇与口语对话的图像引导式任务练习。本文围绕“多语同步开发”理念，介绍平台的构建路径与三语协同机制，分析其在语言对等设计、智能反馈与跨语适配方面的关键策略，并结合试点应用，探讨系统面临的挑战与优化方向。

【关键词】 教育人工智能；机器学习；自然语言处理；多语学习；语音交互

Abstract: *LEARN is a picture-based AI oral learning platform co-developed by the Singapore Institute of Technology and Nanyang Technological University. Designed for lower primary students, the platform supports synchronous multilingual learning in Chinese, Malay, and Tamil. It integrates natural language processing, speech recognition, and personalized feedback to provide contextualized oral practice through image-guided tasks. Centered on the concept of multilingual synchronous development, this study outlines the platform's design process and collaborative framework across languages. Key features such as content alignment, intelligent feedback, and cross-linguistic adaptation are examined. Preliminary pilot use highlights both the platform's effectiveness and areas for improvement, offering insights into AI-driven approaches to oral language learning in multilingual primary education settings.*

Keywords: AI in education, machine learning, natural language processing, multilingual learning, speech interaction

1. 研究背景与问题陈述

近年来，人工智能技术，特别是机器学习与自然语言处理（NLP），在教育领域的应用迅速扩展，推动了智能教材与个性化学习工具的发展（Chen et al., 2020；Holmes et al., 2019）。深度学习模型在语音识别、语言生成和语义理解方面的突破，使得构建互动性强、响应灵敏的语言学习平台成为可能（Zawacki-Richter et al., 2019）。语言模型（如 Transformer 架构）和语音识别系统的不断进化，为提升儿童语言输入质量和学习反馈的即时性提供了技术基础（VanLehn, 2006）。在低龄儿童语言教育场景中，机器学习的自适应能力尤其适合处理初学者多变的语言输入，提供可调整的学习内容与反馈路径（Chen et al.,

2020）。同时，语音转文本（ASR）、语义理解（NLU）与对话生成（NLG）等关键技术的结合，使 AI 系统能够理解学生表达、识别语言偏误，并提供即时响应，从而构建具备互动、诊断与支持功能的智能会话环境（VanLehn, 2006）。机器学习技术，尤其是具备自适应与迁移能力的语言模型，能够针对本地学生的语言背景设计定制化的输入输出路径，从而弥补语言环境的不足（Zawacki-Richter et al., 2019）。通过对大量新加坡本地化语料（如校本语言表达、口语练习答案、生活化用语）的训练，AI 系统可以模仿真实语境中的语言使用方式，增强语言输入的文化关联性与情境真实感。与此同时，智能反馈机制也能根据学生具体表达方式给予语言风格与语法层面的微调建议，有效提升他们在非母语语境下的表达信心与准确度（Holmes et al., 2019）。

基于此，本研究提出核心问题：如何借助融合语音识别与自然语言理解的人工智能技术，构建一个支持多语同步开发的口语学习平台，服务于新加坡小学一至二年级学生的母语学习需求，推动跨语口语教材的智能化转型与本地化适配。

2. 项目目标

本研究历经两年，在新加坡理工大学智能机器学习团队与南洋理工大学国立教育学院语言教育团队（NIE）的通力合作下，开发了一个图画驱动、任务导向、互动性强的 AI 语言学习平台，旨在提供新加坡小学一至二年级学生在日常生活中使用母语（华语、马来语、淡米尔语）进行简短对话的自主习得机会。该平台整合多模态互动机制（听、说、看、反馈），并将其应用于仿说练习、词汇学习与智能对话任务中。同时，系统设计了即时反馈模块，帮助学生提升自主学习能力与自我监控意识，为儿童语言习得提供个性化、沉浸式的技术支持。

3. 平台设计与功能

3.1. 文化协同与图像内容的统一设计

平台的学习内容由贴近日常生活的主题图画所驱动，围绕个人、家庭与学校等常见情境精心绘制，风格活泼、生动，富有趣味。图画内容体现新加坡多元文化的社会生活，并根据华族、马来族与印度族三大族群的文化背景进行适当调整，旨在增强学生的跨文化认同与情境理解。开发过程中，图像内容须兼顾各族文化的共通性与差异性。例如，原设定的生日宴会场景中出现的“烧猪”图因不适用于马来族而被提出修改；最终，为确保三语版本内容统一，平台将相关图像统一调整为冷盘食物。

3.2. 仿说练习（基本口语词库 Wordbank）

每幅图配套句子，学生反复聆听录音并模仿发音。系统提供三次录音机会，并通过语音识别显示错误词语。每轮结束后，自动总结学生表现，展示“好词”与“多练习”词语。请参考下列图像 1 至 3。

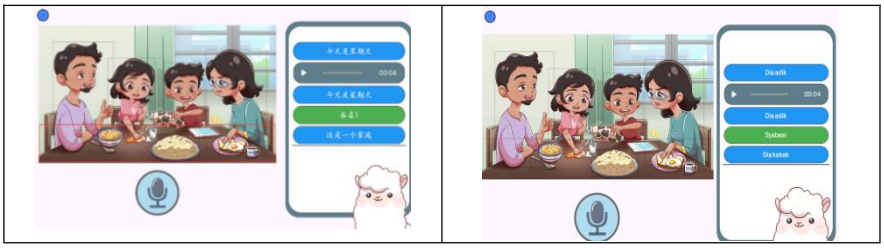


图 1 成功答题系统会以赞语回应，并自动切换到下一题。以上为华语和马来语平台的示例。

图 2 读错的字词，系统会亮出红字，显示答错。以上为华语、马来语和淡米尔平台的示例。

图 3 在“好词”与“多练习”词汇表里，每种语言都有独立的二表，每个类别的总结上限为六个。若所有词语都能正确读取，“多练习”词汇表会亮出赞语‘恭喜！’，类似下列马来语图示；若全部读错，“好词”词汇表会亮出类似下列淡米尔语图示的鼓励语。

3.3. 多选择词语练习（MCQ）

系统自动播放题干与选项音频，学生听后选择正确词语。每题提供三次作答机会：答对即亮出赞语；前两次答错将给予鼓励语引导重试；第三次答错则提供正确答案及语音反馈。答题结束后，系统将根据整体表现给予相应的鼓励反馈。请参考下列图像 4 至 7。

图 4 每图设有四题，每题共有四个选项。

图 5 答对时，系统会以赞语弹窗回应，并让用户点选下一题。

图 6 前两次若选择错误，系统会亮示鼓励语让学生继续尝试。

图 7 第三次答错，系统就会亮示鼓励语，并自动展示正确答案。

为帮助用户更专注内容理解，系统会随机排列每道题的选项，避免用户通过记住选项的位置或颜色机械作答，影响练习效果。

完成四道题目后，系统会以星形标记正确答题数，并按各语提供表 1 的赞语或鼓励语。学生点击“可以/OK/சரி (OK)”键后，系统会返回题目选择主页。请参考下列图像 8。

表 1 各语的赞语或鼓励语。

	华语	马来语	淡米尔语
四题全对， 四星：	恭喜！	Syabas! 好样的！	வாழ்த்துகள் 恭喜
答对三题， 三星：	真棒！	Bagus! 很好！	நல்ல முயற்சி 不错的尝试
答对一、两题， 一、二星：	太好了，继续努力！	Jangan putus asa! 不要放弃！	உன்னால்/ உங்களால் முடியும் 你能做到
四题全错， 无星：	再努力，你可以的！	Ingin mencuba lagi? 想再试一次吗？	மீண்டும் முயற்சி செய்யலாமா 我们可以再试一次吗？

图 8 四题全对时系统的赞语弹窗。

3.4. 智能对话（Chatbot）

基于图画与任务情境，学生与 AI 展开简短会话（约 7-13 轮）。采用生成式 AI 技术，依据学生回应动态调整对话路径。

“乐言”系统的智能对话在展示新对话时，首先播放相关句子。学生听完句子后，如果不清楚，可以点击对应按键重复播放。确认听清楚后，学生点击麦克风按钮进行答复。完成后，学生可点击录音回放检查是否满意。如同仿说练习（Word Bank），如果不满意，学生可以继续录制，直到达到理想效果再点击发送。如图所示，该学生录制了四次，并最终选择最满意的一次发送。收到答复信息后，系统自动生成回复，开展下一个对话。目前，不论答对与否，系统都统一回应，“正确的答案是……”，如表 2 和图 9 所展示。

表 2 各语开展对话的范例节选。

	华语	马来语	淡米尔语
系统提问：	“欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？”	Selamat datang! Mari kita lihat budak lelaki yang memegang buku dalam gambar. Apakah budak lelaki itu sedang lakukan? （译文：欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？）	வரவேற்கிறேன்! படத்தில் ஒரு புத்தகம் பிடித்திக் கொண்டிருக்கும் ஒரு சிறுவனை நோக்குக. சிறுவன் என்ன செய்கிறான் （译文：欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？）
用户选择最佳答复，点送录音内容：	看书。	Baca buku. （译文：读书。）	புத்தகம் படிக்கிறார் （译文：他在读书）
系统回应：	正确的答案是“看书。”。男孩在书上做什么？”	Jawapan yang betul adalah: ‘Baca buku’. Apa yang budak itu lakukan dengan buku itu? （译文：正确的答案是“读书。”。男孩在书上做什么？”）	சரியான பதில்: 'புத்தகம் படிக்கிறார்'.. அந்த சிறுவன் புத்தகத்துடன் என்ன செய்கிறான்? （译文：正确的答案是“他在读书”。男孩在书上做什么？）

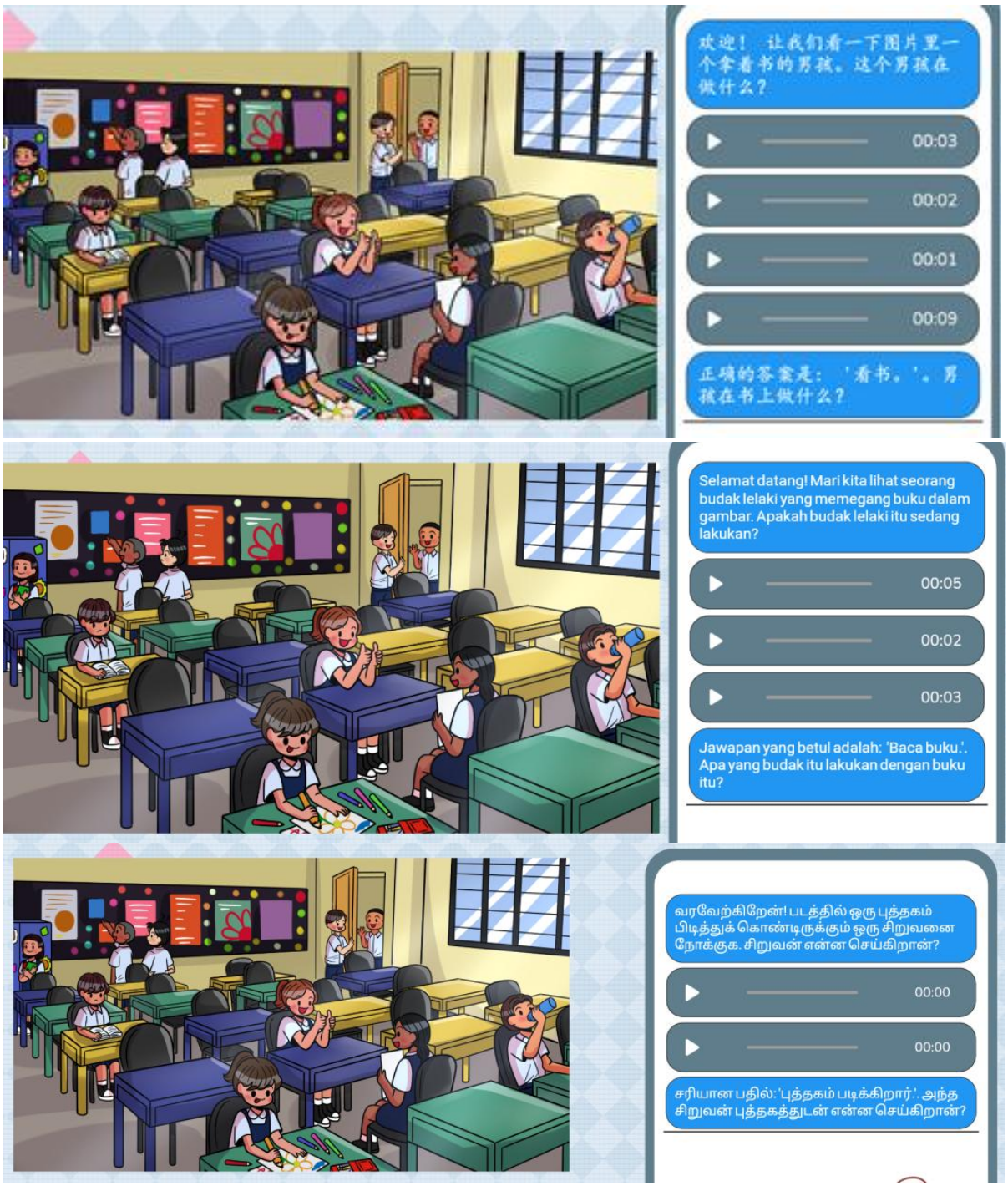


图 9 学生可重复录音，并择优提交。

三个语言平台在界面布局与使用功能上保持高度一致，体现统一的设计框架，其核心语言学习辅助功能均以语音播放为主。在此基础上，平台根据各语言的语音特点、教学进度安排，以及日常用语和文化表达习惯的差异，进行内容层面的本地化适配，实现“结构统一、内容因语而异”的三语同步设计理念。表 3 列举了三种语言在内容呈现上的部分差异示例：

表 3 三语内容呈现差异概览

	华文	马来文	淡米尔文
平台语言辅助差异	通过拼音注音作为支持，帮助学生绕过汉字识读障碍，专注于口语	无	无

	华文	马来文	淡米尔文
	练习本身，提升语言输出的流畅度。		
发音	属于声调语言，有四个声调，声调是第二语言习得的难点。	音节较简单，没有字调问题。	包含丰富的辅音－元音组合，音节结构相对复杂；语音系统的学习难度明显高于华语和马来语。
语系	属于汉藏语系，广泛使用于中华地区及全球华人社区。	属于南岛语系，广泛使用于马来西亚、新加坡、印度尼西亚等地。	属于达罗毗荼语系，主要用于印度南部和斯里兰卡。

4. “乐言”——智能机器学习的特色与创新点

“乐言”平台融合了机器学习、自然语言处理与教育科技设计的优势，体现出智能语言学习系统的关键特色与创新路径，主要包括以下四点：

4.2 智能对话系统

基于自然语言处理技术，平台能够模拟日常情境中的语言交流，实现与学生进行多轮互动式对话，提升语言输出的真实性、频率与表达质量。

4.3 个性化反馈机制

系统结合语音识别与学习数据分析，对学生的发音、语调、词汇准确度等维度进行即时诊断，提供针对性的改进建议，有效支持差异化教学需求。

4.4 可视化学习体验

以图像引导语言输入与输出，通过图画描述任务降低认知负担，同时激发学生对场景语言的理解力与表达兴趣。

4.5 自主学习支持

平台具备高度可调节性，学习者可依据自身节奏自由选择学习内容与练习频率，强化学生的学习主控感与自我监控能力。

5. 实践反思与系统改进

在新加坡“英语主导、母语传承”的双语教育政策框架下，学生除修读英语，并以英语为各个课程的教学媒介语外，亦须选择一门族群母语（华语、马来语或淡米尔语）作为第二语言。由于母语课程紧密承载各族群的语言与文化传统，不同语言在语用、表达方式及内容组织上存在因此显著差异。这使得在统一课程目标与教学任务下，如何实现多语教学内容的等效设计，成为课程开发与资源建设中亟待解决的关键难题。

本项目在技术团队与三大母语教育专家（华文、马来文、淡米尔文）的多轮协商、测试与迭代基础上，遵循“异中存同、同中尊异”的设计原则，尊重各语言的语法逻辑与学习规律，推动平台内容的等效性与本土化适配。在协作过程中，以英语作为跨语种的沟通媒介，协助多语言专家与技术专家团队达成共识，实现内容的同步开发。

新加坡这一独特的多语教育生态为本研究提出关键议题：如何借助融合语音识别与自然语言理解的人工智能技术，开发一套支持多语同步推进的口语学习平台，从而推动语言教材

的智能化转型与跨语种应用。作为新加坡首次以三语同步开发的 AI 口语学习系统，“乐言”不仅回应本地教育政策的多语需求，也为国际社会提供了跨语言智能教材开发的实证范式。

作为第一代智能口语学习平台，“乐言”在实践过程中积累了宝贵经验，也暴露出一些需优化的系统问题。本项目已完成初步开发，并在真实教育场景中进行了应用测试。

首先，第一代系统在反馈机制方面仍有待完善。目前的设计在部分任务中未能提供清晰的“正误”判断，容易造成学生对答案准确性产生疑惑，影响学习信心。特别是在开放性问题中，现有系统无法提供任何回应，而是直接跳转到下一个话题。这种 AI 回应的模糊性与单一反馈方式，使学生难以理解其表达是否得当。新工大目前正尝试引入新的语言模型来改善回应机制，具体进展尚不明确。此外，淡米尔语的问题也最为明显，原因有待查验。

其次，在对话交互系统中，存在跳题、回应逻辑不连贯的问题。这影响了语言任务的连贯性与真实感，削弱了学生的沉浸体验。未来版本应进一步优化对话策略的上下文保持能力，提高 AI 在连续对话中的表现一致性。

此外，系统对于不同表达方式的识别度还需提升，例如学生用词与参考答案不同但语义相近时，系统反馈仍显僵化，未能体现智能教学的包容性与灵活性。

随着自然语言处理技术与生成式 AI 的持续进化，这些问题在未来版本中有望得到解决。目前团队已针对上述问题制定功能更新路线图，逐步完善反馈精度、对话连贯性与语义识别能力，以进一步提升学习成效与用户体验。

6. 结论

“乐言”平台展现了人工智能在跨语言教学领域的创新潜能，尤其在多语学习环境中构建“智能一机多语”的应用框架，回应了新加坡作为双语教育国家对语言平等与教育公平的现实需求。平台在设计之初即采用跨语同步开发策略，确保教学任务、图像场景与交互机制在华语、马来语和淡米尔语三种母语中实现等效表达，既尊重语言差异，又促进内容协同。

平台融合图画驱动、任务导向与个性化反馈机制，为学生构建出一个具有高度互动性、即时应答与语言迁移支持能力的学习空间。通过人机对话模拟日常交流情境，即便学生处于母语输入相对匮乏的家庭语境中，亦能借助 AI 系统获得持续、结构化的语言练习机会，从而提升其口语流利度、语用准确性与表达信心。实际应用结果显示，系统有效拓展了学生接触母语的时间与频率，也为跨语智能教材的开发提供了实践路径与可推广模型。

展望未来，随着自然语言处理与语音技术的持续优化，以及虚拟现实（VR）与增强现实（AR）在教育场景中的逐步普及，“乐言”平台有望发展为一个集智能问答、情境互动与虚拟角色扮演为一体的沉浸式语言学习系统。它不仅能进一步提升学习兴趣与表现，还将为人工智能在语言教育领域的深层融合提供实践范式与研究借鉴。“乐言”平台展示了人工智能在小学母语教育中的创新应用。通过图画驱动、任务导向、反馈增强的语言学习系统，学生在家庭中获得更多互动机会，提升口语能力。未来结合 VR/AR 等沉浸式技术，“乐言”将为语言学习者提供更丰富、真实、有趣的学习体验，为 AI 教育实践提供重要参考路径。

鸣谢：

本文研究成果源自新加坡人工智能署（AI Singapore）资助项目 **LEARN: Language automated Evaluation by generating Answers/questions from caRtooNs**（项目编号：AISG-SIT 12/22 SA）。特此致谢。

参考文献

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). *Application and theory gaps during the rise of artificial intelligence in education. Computers and Education: Artificial Intelligence, 1, 100002*. https://doi.org/10.1016/j.caeai.2020.100002

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston: Center for Curriculum Redesign.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). *Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 1–27*. https://doi.org/10.1186/s41239-019-0171-0

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, 16(3), 227–265.

AIGC 课程设计策略与实践路径——以“诗歌集创作”项目为例

Curriculum Design Strategies and Implementation Pathways for AI-Generated Content (AIGC): A Case Study of the 'Poetry Anthology Creation' Project

谢贤晓 1，郑唐怡轩 2，
平阳县教师发展中心
平阳县昆阳镇新城小学*
14405612@qq.com

【摘要】随着生成式人工智能（AIGC）技术在中小学教育中的深度融合，如何引导学生正确认知并合理使用 AIGC 工具成为 AIGC 教育的关键问题。本文以“诗歌集创作”项目为例，从主题、认知、过程、伦理四个方向提出 AIGC 课程设计策略，通过真实问题驱动课程设计，重构学生对 AI 的工具化认知，引导学生从技术操作转向创新迁移，并融入伦理反思以强化社会责任意识，探索 AIGC 教育“可用”且“善用”的实践路径。实践表明，该策略能有效提升学生的创造性思维与问题解决能力，同时为中小学人工智能课程与 AIGC 教育提供了可操作的参考框架。

【关键词】生成式人工智能；AIGC 教育；课程设计；中小学人工智能

Abstract : *With the deepening integration of generative artificial intelligence (AI-Generated Content, AIGC) in K-12 education, guiding students to develop accurate cognitive frameworks and responsible utilization of AIGC tools has become a pivotal challenge in AIGC pedagogy. This study employs the "Poetry Anthology Creation" project as a case study to propose a multidimensional curriculum design strategy encompassing thematic contextualization, cognitive restructuring, process optimization, and ethical integration. By anchoring instruction in authentic problem-driven scenarios, the framework systematically reshapes students' instrumental cognition of AI, transitions their focus from technical operation to innovative application, and embeds ethical reflection to strengthen social responsibility. The research explores practical pathways for achieving both technological accessibility ("usability") and ethical utility ("responsible use") in AIGC education. Empirical findings demonstrate that this approach significantly enhances students' creative thinking and problem-solving competencies while establishing an actionable reference framework for integrating AIGC into K-12 artificial intelligence curricula.*

Keywords: Generative AI ,AIGC Education, Course design,K-12 Artificial Intelligence Education

1.前言

生成式人工智能（Artificial Intelligence Generated Content，简称 AIGC）是一项可以利用大型语言模型和深度学习算法，模拟人类的创造性思维过程，生成文本、图像、视频、音频等多类型内容的技术（柯清超等人，2024）。近两年来，以 ChatGPT、DeepSeek 为代表的生成式人工智能迅速迭代，其使用场景已与中小学教育深度融合。有研究表明这项技术虽然能够为学生提供有效的学习辅助，但若缺乏正确引导，也可能会导致学生过度依赖 AI，丧失独立思考能力和创造性思维（沈苑等人，2024）。所以，如何让学生从人工智能课程中对

AIGC 技术树立正确的认知，学会正确地使用 AIGC 工具是小学人工智能课程建设的重要议题（周玲等人，2024）。基于上述背景，笔者结合自身开发的 AIGC 课程的经验，以“诗歌集创作”微项目为例，从主题、认知、过程、伦理四个方向提出课程设计策略，旨在探索生成式 AI 教育“可用”且“善用”的实践路径，为中小学人工智能课程建设提供参考。

2.主题设计：以真实需求驱动课程目标

AIGC 课程需避免“为技术而技术”的误区，目前市面上有部分 AIGC 课程以单纯的“提示词”的撰写为主要教学内容，通过了解提示词的组成结构学习生成图片，或者通过简单的提示词进行创意创作，虽能快速产出视觉成果，却割裂了技术与真实需求的关联，导致学生陷入“生成即掌握”的认知偏差，既难以理解技术解决实际问题的逻辑，也无法形成可持续的创新迁移能力。

因此，AIGC 作为一种实用工具，课程设计需遵循“从问题中来，到问题中去”的循环逻辑，基于学生认知发展规律与技术真实应用场景，以生活中用 AIGC 技术解决的典型问题作为课程驱动。通过分析学生年龄段的特定困境（如艺术创作能力不足、文化理解局限等），提炼可被 AIGC 技术赋能的真实任务场景，再基于学生创作路径中可能会遇到的真实问题规划任务链，引导学生从工具探索走向创新应用，最终实现技术认知与问题解决能力的同步提升。

例如，笔者通过观察校园内的学生活动发现，在语文四年级下册第三单元综合性学习“轻叩诗歌的大门”中，学生们学生需自主编撰诗歌集并设计插图，但由于绘画能力有限，插图效果不理想（如图 1）。

图 1 学生手绘诗歌集插图

所以笔者以“诗歌集创作”作为 AIGC 课程的项目主题，项目课时设计如表 1 所示。让学生使用 AI 工具为诗歌集“定制”图片。课程采用“探索工具→技术体验→拓展创作→深度思考”的实践链条，每个课时都紧扣前一步的实际需求进行延伸，引导学生从个人需求出发，通过探索 AIGC 工具的多样化应用方式解决真实存在的问题，逐步实现从操作到应用的能力提升，最终完成个性化的项目作品。

表 1 “诗歌集创作”项目课时内容设计

课时	阶段	真实需求	能力目标	具体任务
1	认知	如何完成一幅优秀的诗歌插图？	概念认知与应用探索	了解 AIGC 技术，检索 AIGC 工具，并能够基于任务选择合适的图像生成工具。
2	体验	如何使用图像生成工具？	提示词撰写与观念重构	学会提示词的撰写逻辑，尝试先进行图像的构思再用提示词描述，形成 AIGC 技术是工

课时	阶段	真实需求	能力目标	具体任务
				具的观念，认识 AI 工具中常见功能键并学会使用。
3	创作	图像生成效果不佳如何解决？	工具筛选与二次创作	总结图像生成具体的问题，学会针对具体的问题寻找和利用合适的图像生成工具实现图像的修改和作品的创新。
4	解析	AI 为什么无法生成我们想要的图像？	原理解析与伦理思考	理解 AI 生成图像的原理，思考目前 AIGC 技术存在的伦理问题。
5	拓展	AIGC 技术在生活中有哪些应用？	拓展应用与技术迁移	学会自主探索、了解更多 AIGC 技术，可以综合利用多种生成工具升级作品。了解 AIGC 工具在其他领域的应用。

3.观念重构：从“AI 创作”到“AI 工具”的认知转型

AIGC 课程设计的一大核心目标是培养学生对 AI 技术的正确认知，以输入关键词获取 AI 生成内容的 AIGC 相关课程忽视了引导学生进行深度思考，容易导致学生形成“AI 可以替代人类创作”的认知偏差。事实上，AI 并非创作的替代者，而是一种辅助工具，其生成的内容本质上是由人类通过精确控制提示词和不断调整来实现的。因此，解构 AI 生成内容的本质，并强调人类意图在算法输出中的决定性作用，在课程设计中应使学生首先明确自己的创作目标，再借助 AI 工具实现创意，从而完成从“AI 可以代替我们实现创意”到“我们可以使用 AI 更好地实现创意”的观念转变。

笔者建议通过“认知解构-实践探索-思维内化”三个步骤逐步实现学生的认知重构。首先，通过讲解解构 AI 生成内容的本质，明确 AI 的输出效果完全依赖于人类的输入和引导，强调人类意图对算法输出的决定性作用。其次，建立人类与 AI 生成技术的协作关系，明确分工逻辑：人类负责创意构思与价值判断，AI 则承担素材生成与构思实现的任务。最后，通过多次生成图像的尝试，巩固“AI 是用来使用的”系统思维，并认识到人类与 AI 在想象力与生产力上的互补性。这一过程不仅帮助学生理解 AI 的工具属性，还能培养他们在创作中的主体意识。

以笔者设计的“诗歌集创作”项目为例，在第二课时“体验”环节中，首先请学生欣赏“未来城市”的 AI 作品，对其中反复出现的元素（如飞行汽车），引导学生思考思考“AI 的想象力从何而来”，初步建立“AI 技术受人类影响”的认知。随后，要求要求学生在使用图像生成工具之前先设计简单的诗歌配图的草图，基于草图撰写提示词（如图 2），并通过调整提示词直观感受到 AI 生成的内容完全取决于自己的输入，强调图像创作的主体是人。同时鼓励他们使用多种工具完成图像的设计，例如借助大语言模型优化提示词，进一步加深对 AI“工具”属性的理解，最终实现从“AI 创作”到“AI 工具”的观念重构。

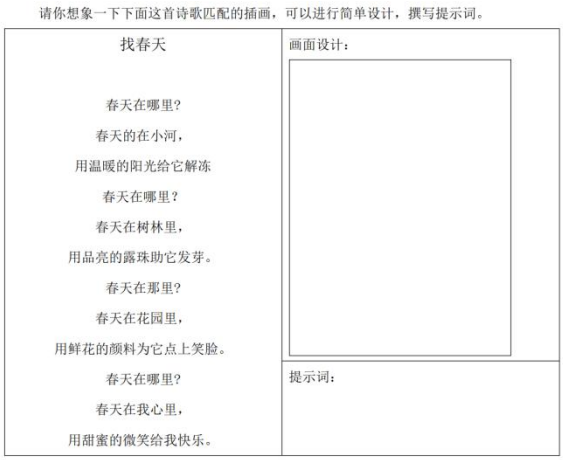


图 2 学生基于草图撰写提示词的辅助单示例

4.过程实施：分层适配与个性化探索路径

因 AIGC 技术在多领域的广泛应用性，AIGC 课程的实施过程也应体现自由性和个性化。笔者认为在课程设计中应构建“发现问题-工具延伸-分层适配”的学习路径：当学生在使用技术工具时遇到生成效果与预期不符的情况，教师不必急于提供标准答案，而是引导他们主动思考背后的原因——是工具本身的限制，还是操作方式需要优化？例如，当 AI 生成的图像细节不够精准时，是提示词描述出了问题还是需要其他辅助工具进行后期完善。在此过程中，教师可提供多种解决方案的参考方向和工具链，但将选择权交给学生，鼓励他们根据自身兴趣和能力组合使用不同工具。这种自主探索的经历不仅能让学生更深刻地理解技术工具的边界，还能培养他们灵活应对挑战的思维习惯。当他们在解决具体问题的过程中积累了经验，便能将这些方法迁移到其他场景——比如将图像优化经验转化为视频剪辑中的创意技巧。这种从实践到迁移的学习循环，既保留了技术应用的自由度，又让每个学生都能在适合自身节奏的探索中，实现从基础操作到创造性解决问题的有效跃升。

例如在“诗歌集创作”的项目实践中，学生生成的图像常常出现人物手指变形、场景衔接生硬等问题，即使反复修改文字描述，依然难以达到理想效果。针对这一情况，笔者要求学生记录文字描述的调整过程，并总结出工具难以精确控制的环节，例如物体数量、位置关系等，从而帮助学生更清晰地理解技术的局限性，并为后续优化提供依据。再在课堂上提供多种辅助工具（如表 2）或让学生自主搜索其他工具，学生可根据具体问题选择适配工具组合使用——想修改画面局部就用编辑工具，需要补充细节就用实时生成工具，让技术真正服务于创作想法。

表 2 辅助工具示例集

工具	功能	实现效果
图像编辑	可以使用框选或点选对局部细节通过提示词进行修改或消除。	

实时创作	可以根据绘制的内容进行实时 AI 填充和美化。	
风格迁移	可以根据提供的风格图片对原始图片进行风格化。或直接参考图片风格生成新图片。	
AI 扩图	可以根据原图进行扩图。	

教学实践表明，这样的实践路径的设计能有效提升学生解决问题的能力。初期学生普遍执着于反复调整文字描述，经过工具协作的启发后，多数学生开始主动组合使用两至三种工具完善作品。在创作中，有的学生综合使用了图像编辑和 AI 扩图分别生成了更符合配图需求的插画（如图 3），有的用风格迁移工具将现代风格的图像转为水墨画风（如图 4）。这种从发现问题到自主解决问题的转变，正是技术创新教育希望培养的核心能力。

图 3 学生项目成果 1

图 4 学生项目成果 2

了解了图像生成工具的细化应用，还可以进一步了解其他分类的 AIGC 技术的广泛应用，例如利用 AIGC 生成的广告视频案例。也可以鼓励学生将诗歌集转化为动态视频作品，或用音乐生成工具为作品配乐。操作类似的 AIGC 工具的使用既能让学生及时产生知识迁移，也可以让每个学生都能找到最适合自己创意的个性化表达方式，实现从工具操作到创意表达的认知跃迁。

5.伦理渗透：技术向善的价值引领

掌握 AIGC 技术应用能力的同时，学生也应了解 AIGC 技术的使用原则和风险。当前 AI 生成技术已引发多重伦理争议，包括版权归属模糊、虚假信息泛滥以及责任界定不清等问题。这些争议不仅揭示了技术滥用的潜在危害，更凸显了在 AI 教育中融入伦理维度的紧迫性。在联合国教科文组织《人工智能伦理建议书》中就提到，青少年 AI 教育必须融入伦理维度，这要求课程设计可针对不同的课程开展模式和不同的 AIGC 课程的设计也应该适当增加伦理内容。

笔者认为针对不同学段学生认知特点，可采用不同的活动设计。面向低年级学生，通过具象化案例研讨建立伦理感知。例如，以“萝卜快跑”交通事故探讨责任归属问题，引导学生思考“责任是由开发者、使用者还是 AI 系统本身承担？”；围绕“虚拟癌症患者骗捐案”分析伪造内容的社会危害，帮助学生掌握虚假信息识别方法并思考技术使用边界。对于高年级学生，则可开展进阶探讨会或辩论会，例如围绕“AI 创作艺术作品是否应享有版权？应该属于谁？”等论题，请学生自主收集观点论据进行讨论；制定《AI 创作伦理守则》，从“禁止恶意伪造他人形象”、“标注 AI 生成内容”等具体条款入手，将抽象伦理转化为可执行的行为规范。课程设计还应包含从“AI 如何工作”到“AI 为何出错”的原理解析，最终导向“AI 应如何负责”，落实到如何在网络上安全、负责任地分享和发布内容，培养学生的信息社会责任。

6.结语

本研究以“诗歌集创作”项目为实践载体，验证了生成式人工智能教育在中小学阶段的可行性与有效性。本项目通过真实问题驱动，使学生不仅掌握了 AIGC 工具的使用方法，还实现了从“依赖技术创作”到“技术辅助创作”的认知重构，通过工具探索与伦理反思环节的设计，进一步培养了学生的创造性思维与社会责任意识。在课程设计策略上形成“问题→观念→能力→责任”闭环。相信在未来，AIGC 技术将持续迭代，其教育应用场景也将更加多元。本论文为 AIGC 教育提供了可参考的设计范式，但如何量化素养培育的长效影响，仍需后续实践探索与理论完善。

参考文献

王同聚.中小学人工智能课程教育实践策略的设计与实施[J].现代教育技术,2024,34(12):95-104.
周玲,王烽.生成式人工智能的教育启示：让每个人成为他自己[J].中国电化教育,2023,(05):9-14.
沈苑,房斯萌,柳晨晨,等.生成式人工智能教育应用治理：案例与反思[J].开放教育研究,2024,30(06):39-47.DOI:10.13966/j.cnki.kfjyyj.2024.06.005
柯清超,米桥伟,鲍婷婷.生成式人工智能在基础教育领域的应用：机遇、风险与对策[J].现代教育技术,2024,34(09):5-13.
胡佳健,汪维富,钟志贤.国内外生成式人工智能教育现状分析与热点探究[J].数字教育,2024,10(06):84-92.

三阶四维：基于 ADDIE 模型的人工智能数字课程设计研究

Three Levels and Four Dimensions: A Study on AI Digital Course Design Based on the ADDIE Model

冯京和[1*]，林雅[2]
[1]温州市教育教学研究院
[2]温州市实验中学
* 349018676@qq.com

【摘要】 本文以 ADDIE 模型为框架，探索人工智能教育中数字课程的设计路径。通过《人工智能中的算法》（基础课）和《AI 校园植物识别》（微项目课）的设计分析，提出了三阶课程体系与四维内容结构的系统化设计方法。研究展示了数字课程如何融入生活化、可视化、交互性和系统化特征，提升学生核心素养，并为人工智能课程创新提供理论与实践参考。

【关键词】 数字课程；ADDIE 模型；人工智能；核心素养

Abstract: *This article explores the design path of digital courses in artificial intelligence education within the framework of the ADDIE model. Through the design analysis of "Algorithms in Artificial Intelligence" (basic course) and "AI Campus Plant Recognition" (micro-project course), a systematic design method featuring a three-tier curriculum system and a four-dimensional content structure is proposed. The study demonstrates how digital courses can integrate features such as life-oriented, visual, interactive, and systematic characteristics, enhance students' core competencies, and provide theoretical and practical references for AI curriculum innovation.*

Keywords: Digital course, Addie model, artificial intelligence, core literacy

在教育数字化转型背景下，《义务教育信息科技课程标准（2022 年版）》提出培养学生信息意识、计算思维和数字创新能力的目标。数字课程作为技术与教学融合的产物，突破传统教学局限，成为课程改革的关键。本文基于 ADDIE 模型，通过两节课的设计案例，探讨如何系统化设计数字课程，以提升教学质量并推动教育创新。

1.数字课程的内涵与特征

数字课程并非教学内容的简单数字化，而是以学习者为中心，整合技术、内容与教学策略的动态生态系统[4]。它通过多媒体与交互手段，将抽象知识转化为可感知的体验，赋予学习过程更大的灵活性与深度[5]。其主要特征包括[6]：生活化（将技术概念嵌入真实场景）；可视化（通过多媒体直观呈现复杂原理）；交互性（强调任务驱动与实践参与）；系统化（构建逻辑清晰的认知流程）。这些特征使数字课程在人工智能教育中独具价值。它不仅能帮助学生理解算法、人工智能等抽象概念，还通过技术赋能，为课程创新注入新的活力。

[4] 张伟.数字课程设计与开发研究[J].教育技术研究,2023,45(2):12-18.
[5] Jonassen D. Enhancing Learning Through Technology: The Essence of Digital Courses[J]. Journal of Educational Technology,2021,38(4):25-32.
[6] 王静. 数字课程的特性与优势分析[J]. 现代教育技术, 2022, 32(5): 40-46.

2.基于 ADDIE 模型的数字课程设计路径

ADDIE 模型（分析、设计、开发、实施、评估）为数字课程设计提供科学指导。通过初中人工智能课程《人工智能中的算法》和小学人工智能课程《AI 校园植物识别》两节课的设计过程，阐述如何运用这一模型打造科学、有效的课程内容。

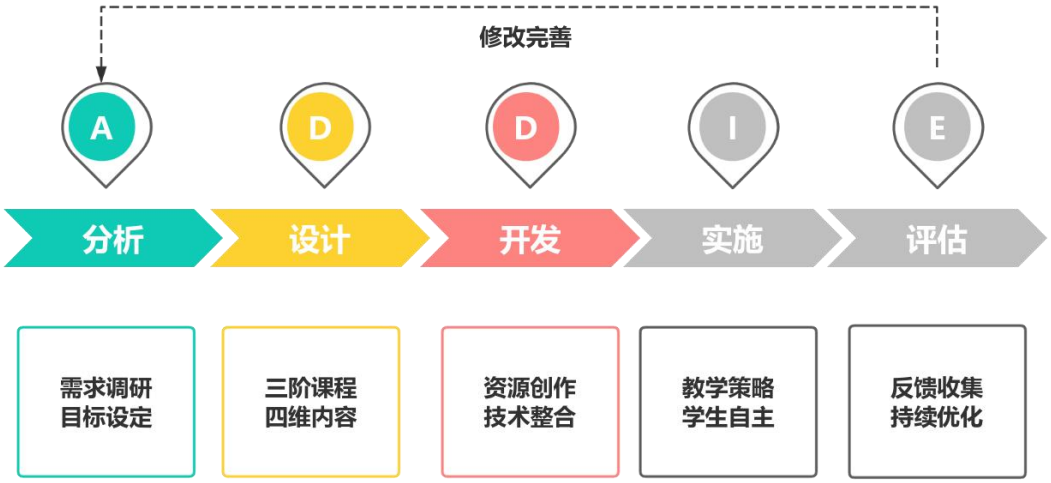


图 1 数字课程的设计框架

2.1.分析：需求明确与目标定位

在 ADDIE 模型中，分析阶段是数字课程设计的基础，通过系统分析学习者需求、课程目标和教学环境，为后续设计提供科学依据。这一阶段明确“为何设计”的问题，并为课程内容的针对性和层次性奠定基础。

2.1.1.需求调研与问题识别

通过需求调研与问题识别，明确课程设计的根本目的。比如，《人工智能中的算法》基础课的调研发现，初中生普遍觉得算法抽象难懂，尤其是“数据如何驱动决策”的逻辑不易把握，反映了理解算法原理的难点。《AI 校园植物识别》则源自学生想解决“校园植物难认”的实际问题，体现他们对技术应用和真实场景的关注。两节课的调研都聚焦学生实际困难，为设计奠定了针对性基础。

2.1.2.目标设定与层次定位

基于调研，课程目标得以明确并分层定位，确保内容的层次性和科学性。《人工智能中的算法》目标是通过具体情境和任务，帮学生理解推理、搜索和预测算法的原理，同时对接新课标计算思维要求，夯实基础并激发兴趣。《AI 校园植物识别》则通过实践驱动的微项目课，让学生掌握机器学习基础流程，培养信息意识和创新能力。两节课的分析阶段都注重目标与需求的契合，确保设计贴合学生实际并符合课程标准，为后续开发提供支撑。

两节课的分析阶段都强调需求的针对性和目标的层次性，确保设计方向既实用又符合标准，为内容开发打下坚实基础。

2.2.设计：体系构建与结构规划

在 ADDIE 模型中，设计阶段是课程从需求分析走向内容落地的重要环节，其核心在于规划课程的整体框架、内容结构和教学策略，确保课程科学有效且激发学生潜能。本研究提出了三阶课程体系和四维内容结构，通过《人工智能中的算法》和《AI 校园植物识别》两节课的设计，展示了理论框架的具体应用。

2.2.1.三阶课程的应用

三阶课程体系借鉴维果茨基“最近发展区”理论，分为基础性课程、微项目课程和跨学科课程，分别面向全体学生、兴趣小组和拔尖学生。《人工智能中的算法》作为基础性课程，聚焦算法原理的普及与计算思维的培养，通过直观情境降低抽象概念的学习难度。《AI 校园植物识别》定位为微项目课程，强调机器学习技术的实践应用和创新能力的拓展，为有兴趣的学生提供更深层次的探索空间。这种分层设计兼顾知识的普适性与个性化发展需求。

图 2 数字课程三阶课程体系

2.2.2.四维内容的构建

四维内容结构包括情境导入、原理探究、实践演练和总结自评，构建从问题驱动到知识内化的完整认知闭环。

表 1 四维数字课程

	内容描述	
第一部分片头	包含包含“温州市人工智能数字课程”字样、微课课题、制作者姓名、制作者单位等	
第二部分出镜	制作者全身出镜并介绍课程，例如“大家好，欢迎来到……（课题名称），本课的目标为……（建议设定 1～2 个学习目标），将通过……（学习方式），了解 / 掌握 / 知道……（学习内容）”	
第三部分学习过程	（1）情境引入	学生视角，聚焦生活中的真实情境，提出一个需要解决的问题
	（2）原理探究	提炼本课相关概念或者原理、知识点，以生动有趣的方式呈现
	（3）案例解析	(可以从案例分析、探究活动、体验应用三者中选一个) 按照课本内容进行体验或者实践。建议在画面中出现实景应用，比如无人机课程，可以呈现动手实践与无人机飞行效果；又比如物联网课程，可以呈现动手编程及实际效果
	（4）总结自评	回归本源，梳理本课知识，呈现本课的板书内容；提供评价量表，学生自主评价；进行适当拓展延伸，以拓宽知识面，引导学生迁移应用
第四部分片尾	呈现“感谢观看”等字幕信息	

以《人工智能中的算法》为例，课程设计围绕“如何让图书馆更智能”展开：①情境导入：以学生熟悉的图书馆系统为切入点，提出“如何推荐合适的书籍”问题，激发学生对算法的探究兴趣；②原理探究：通过渐进式图表和动态演示，分解推理、搜索和预测算法的工作原理，帮助学生将抽象概念转化为可理解的逻辑；③实践演练：学生在模拟的数字图书馆系统中输入关键词，观察算法如何排序和推荐书籍，直观感受算法的高效性；④总结自评：通过开放性问题“算法还能优化哪些生活场景”，引导学生反思所学内容的应用价值。通过生活化情境和可视化手段降低了学习门槛，增强了学生的参与感和成就感。

《AI 校园植物识别》作为微项目课程，则更侧重实践性与创新性，其设计围绕“如何用 AI 认识校园植物”展开：①情境导入：以“校园植物无标识”的现实问题为驱动，激发解决实际问题的兴趣；②原理探究：用动画展示机器学习的数据要求（如多角度、多样本），引导学生理解“数据质量决定模型效果”的核心逻辑；③实践演练：将项目分解为数据采集（拍摄植物照片）、数据整理（标注分类）、模型训练和应用开发四个任务，学生在操作中逐步掌握技术流程；④总结自评：通过用户反馈，学生测试识别应用的准确性并提出改进建议。通过问题驱动和递进式实践，引领学生在真实问题解决中提升信息意识和创造力。

设计过程中，四项原则贯穿始终：生活化增强关联性，可视化降低认知难度，交互性激发参与，系统化保障内容深度。这些原则为课程设计的科学性与创新性提供了坚实支撑。

2.3.开发：资源设计与技术整合

在 ADDIE 模型中，开发阶段是将设计理念转化为教学资源的关键一步。通过多媒体内容、互动元素和实践任务的创作，我们将课程目标和教学策略具象化，为实施打下基础，确保数字课程既满足教学需求，又能激发学生潜能。

2.3.1 多媒体资源的创作

在《人工智能中的算法》课中，我们用生动的动画和直观的图表让抽象算法变得易懂。例如，模拟图书馆推荐书籍的动画展示了推理、搜索和预测算法的应用；树形图和曲线图拆解了算法步骤。互动任务则让学生在模拟系统中操作，增加学习的趣味和实效。对于《AI 校园植物识别》，我们制作了分步视频教程，带学生从数据采集走到模型训练，强调数据质量对模型效果的重要性。学生通过拍摄校园植物、训练分类模型，体验“做中学”，提升信息意识和创新能力。

2.3.2 技术工具的集成与应用

UMU 平台是教学技术的核心，整合了动画、视频和互动任务，提供流畅的学习体验。它支持多媒体展示和实践任务的执行，确保教学顺利推进。平台的数据分析功能还能让教师实时跟踪学生进度和效果，灵活调整策略，提升教学质量和效率。

通过这些资源创作和技术应用，我们将设计理念落地为教学资源，确保课程科学、系统且具创新性，充分激发学生潜能。

2.4.实施：理念落地与教学支持

在 ADDIE 模型中，实施阶段是将设计理念与资源应用于实践的关键环节，通过支持体系确保课程目标实现。这一阶段强调教学设计落地，通过科学策略和资源整合，支持教师有效教学，促进学生深度参与与能力提升。

2.4.1.教学策略与情境创设

对于《人工智能中的算法》，教学从播放图书馆推荐系统动画开始，引导学生进入“如何让借书更智能”的情境，激发学习兴趣。《AI 校园植物识别》则通过视频介绍项目目标

（如开发校园植物识别应用），以“校园植物识别”的现实问题为驱动，点燃学生的探索热情。这些策略通过生活化情境和可视化资源，将抽象概念转化为直观体验，实现了从情境导入到实践探索的学习旅程。

2.4.2.教学策略与情境创设

在实施过程中，教师通过适时引导和资源支持，助力学生自主学习。《人工智能中的算法》中，教师利用互动任务指导学生输入数据（如书名），观察算法运行效果，并组织小组讨论，让学生分享“哪种算法更适合推荐”的见解。《AI校园植物识别》以项目为核心，学生分组完成从拍摄植物照片到训练模型的实践任务，每一步可通过反馈工具检查成果（如照片质量），教师则在关键节点（如模型训练出错时）提供指导，而非全程主导。

通过分步任务和灵活的资源整合，两节课在不同场景下确保了学生自主性与教学实效性的平衡，培养了解决问题的能力。

2.5.评估：效果分析与设计优化

评估阶段是ADDIE模型中检验设计效果并推动持续改进的关键步骤，其目标是通过系统化的反馈收集与分析，确保课程的科学性与实用性。

2.5.1.反馈收集与问题诊断

《人工智能中的算法》课程通过问卷调查收集学生反馈，发现动画资源直观易懂，但互动任务的挑战性不足，未能充分激发学生的计算思维。教师观察也显示，学生对算法在实际场景中的应用理解较浅。《AI校园植物识别》课程的评估则以学生作品为依据，发现多数小组的植物识别模型准确率偏低，原因在于学生对模型训练流程理解不足，尤其不了解数据预处理对模型效果的影响。学生反馈指出，视频教程内容清晰但缺乏直观的解释性演示。这些反馈和问题诊断为后续的优化提供了针对性方向。

2.5.2.优化策略与迭代设计

针对反馈与诊断，我们优化了《人工智能中的算法》和《AI校园植物识别》课程。《人工智能中的算法》课程增加了“预测借书趋势”的复杂案例，要求学生分析一周借阅数据并设计推荐策略，提升了任务挑战性，深化了算法理解，符合培养计算思维的目标。《AI校园植物识别》课程在视频中加入动画展示数据清洗和参数调整，并要求学生对比清洗前后的模型效果，解决了理解障碍，强化了技术能力，符合培养创新意识的目标。

评估形成“反馈—调整—优化”闭环，确保资源与教学设计持续优化。我们将AI伦理融入课程，如在《人工智能中的算法》增设“算法公平性讨论”，引导学生思考推荐系统中的偏见；在《AI校园植物识别》项目强调数据的隐私性。这种伦理维度不仅体现课程科学性，而且为培养具伦理素养的AI创新人才奠基。

3.评估：研究结论与启示

基于ADDIE模型，以《人工智能中的算法》和《AI校园植物识别》为例，本文系统探索了数字课程设计路径。研究提出了三阶课程体系（基础性课程、微项目课程、跨学科课程）与四维内容结构（情境导入、原理探究、实践演练、总结自评），并将生活化、可视化、交互性和系统化原则贯穿于课程设计的全过程，有效提升了学生的信息意识、计算思维与创新能力。数字课程的设计不仅是技术与教学的融合，更是教育理念的革新。通过科学的设计与持续的优化，数字课程将为培养未来社会所需的创新型人才提供有力支撑。

参考文献

冯京和.(2024).数字+课程体系:中小学人工智能课程实施推进的温州路线[J].中国科技教育,(08):25-29.

赵明华,李晓东.(2024).数字课程中的交互设计与学习体验优化研究[J].教育研究.45(4),98-105.

周明,张艳.(2024).信息科技教育中的精准教学模式研究——基于数字课程的实践探索[J].课程·教材·教法.43(2),67-74.

张伟,刘明.(2023).数字时代信息科技教育的创新发展与实践探索[J].电化教育研究.44(8),76-83.

李明,陈光.(2023).基于 ADDIE 模型的数字课程开发路径与实践策略[J].现代教育技术.33(12),45-52.

王静,李华.(2023).信息科技教育中数字课程的特征与应用研究[J].中国教育技术.42(5),89-96.

在第二外语学习中大语言模型应用的系统文献综述

A Systematic Literature Review of Large Language Models in EFL

魏宏[1*]，廖龙燕[1]
[1]温州大学教师教育学院
* 3563984610@qq.com

【摘要】 随着技术发展推动教育变革，大语言模型在语言学习领域展现出新活力。研究发现，ChatGPT 等大语言模型在教育中的应用已较为广泛，但集中在高等教育中，在 K-12 教育领域的研究仍然有限。本文探讨了 ChatGPT 等大语言模型在 K-12 教育中的实施现状和潜在作用。通过系统文献综述法，本研究基于 PRISMA 框架，对 2022 年至 2025 年的相关实证研究进行了文献筛选和分析，旨在探讨大语言模型支持下的第二外语学习的 AI 角色类型、教育领域、教育阶段、教师角色，为未来研究提供参考，助力第二外语教育发展。
【关键词】 第二外语学习；大语言模型；K-12 教育；系统文献综述

Abstract: *As tech advances reshape education, large language models (LLMs) are becoming increasingly popular in language learning. Research shows LLMs like ChatGPT are widely used in education, mainly in higher education, with limited K-12 application. This paper examines their implementation and potential impact in K-12 education. Using a PRISMA - based systematic literature review, this study analyzed 2022-2025 empirical studies. It focuses on AI roles in LLM - supported second foreign language learning, covering educational domains, stages, and teacher roles. The findings offer insights for future research and practical suggestions for LLM integration in this field, advancing it and providing a reference for educators and researchers.*
Keywords: EFL, Large language models, K-12 education, Systematic literature review

1. 引言

随着技术的飞速发展，人工智能（AI）技术正深刻地改变着教育领域。大语言模型（Large Language Models, LLMs）作为人工智能的重要分支，凭借其强大的自然语言处理能力，为教育带来了新的机遇和挑战。在人工智能研究领域，OpenAI 的 GPT 系列模型已处于前沿地位，其通过处理海量文本数据所学知识，为众多任务提供了先进解决方案，教育领域便是其潜在的重要应用场景之一（Heidt, 2023）。近年来，LLMs 如 ChatGPT 等在教育中的应用逐渐增多，其在个性化学习、智能辅导、自动化作业批改等方面展现出巨大潜力，显著提升了教学效率和学习体验（卢宇, 余京蕾, 陈鹏鹤, & 李沐云, 2023）。有学者提出，ChatGPT 将会对教育教学理念、教学模式、教育科学研究范式、人才培养等方面带来巨大冲击（夏立新，2023）。其优势在于能为学生的论文、作业提供支持，促进个性化学习，同时辅助教师的教育教学，生成教学反馈与评价，助力教学创新（AlAfnan, et al，2023）。然而，它也存在一些潜在风险，例如回答问题的准确性不足、伦理与安全问题、知识剽窃和学术不端行为等（王天恩，2023）。与此同时，公众对 ChatGPT＋教育的态度是可以提供教育资源，有助于教育发展(刘天丽, 杨现民, 李康康, & 糜泽稼, 2023)。这些研究大多证明 ChatGPT 在教育以及学习方面会产生积极影响。本文筛选文献，旨在关注以 ChatGPT 为首的大语言模型在不同语言技能、不同学段学生中的实施现状。回顾文献综述(Jaeho Jeon, 2023)，该综述通

过文献计量学分析语音识别聊天机器人在语言学习领域的研究趋势等，发现重点集中在高等教育应用上，指出了 K-12 教育领域的研究缺口。因此，有必要在 K-12 教育领域对第二外语学习中使用以 ChatGPT 为主的大语言模型的研究现状进行全面审查。

以 ChatGPT 为首的基于大语言模型搭建而成的聊天机器人在 EFL(English as a foreign language learning,英语作为二外的学习)中的研究已经渐渐形成体系。例如，已有研究探讨了语言学习者如何应对与聊天机器人交流时出现的障碍（Moussalli & Cardoso，2020），与聊天机器人的互动如何影响他们的学习成果（Dizon，2020），以及使用聊天机器人对诸如交流意愿、动机和焦虑等情感变量的影响（Chen, et al，2020；Jeon，2022；Tai & Chen，2020）。总体而言，这些研究证实了使用具备语音识别功能的聊天机器人对语言学习的益处。此外，随着各种由大型语言模型驱动的、具有语音识别功能的聊天机器人在社会中日益普及，它们可能会吸引越来越多的学术界的关注。随着大语言模型与学生和 EFL 之间进行协同，我们需要理解它在协作中的具体角色，以及协作的情况如何演变，为今后更好地使用工具提供指导。于是本文提出研究问题，旨在揭示大语言模型在 EFL 中的角色及协作情况。

根据系统评价和元分析(PRISMA)框架的首选报告项目，对一系列研究进行了检查，以回答以下研究问题:

• RQ1: 在 K-12 教育中，大语言模型支持下的第二外语学习的协作有哪些类型（即大语言模型在协作中的角色）？协作的情况如何？

• RQ2: 在 K-12 教育中，大语言模型应用在哪个教育阶段和哪个语言技能，是如何实施的？实施的结果如何？

• RQ3: 在 K-12 教育中，教师在大语言模型应用到语言学习的角色是什么样的？态度如何？对未来发展有什么建议？

通过这一系统综述，本文期望为大语言模型在 K-12 教育的 EFL 中的应用提供全面的了解，并为未来的研究和实践提供有价值的指导。

2. 文献综述

近年来，随着人工智能和机器学习技术的迅猛发展，像 OpenAI 的 GPT 系列这样的大型语言模型已成为该领域的重要构成部分。这些模型在训练过程中能够“学习”到大量的信息，从而可以执行更复杂的任务，例如回答问题、写作以及编程等。它们通常运用深度学习技术，针对海量的文本数据进行训练，旨在生成与训练数据相似的新文本，或者理解和回答有关训练数据的问题（吴河江 & 吴砥, 2024）。凭借其多功能性，大型语言模型目前在不同领域都发挥着一定的作用，尤其在第二外语教育领域。

在教育领域的最新趋势中，大型语言模型正逐渐成为各教育层次的重要工具，尤其在低年级阶段，比如小学，其应用已变得更加普遍和关键。在小学五年级中，配备 AI 技术的聊天机器人“Sister Fish”被设计为能够理解 157 本书的基本内容，模仿人类同伴的互动方式，与小学生进行交流以提高学生阅读兴趣（Liu,2022）。在小学六年级的英语教育中，教师将大型语言模型（LLMs）整合到外语写作教学中，对学生的写作表现、自我调节学习策略使用以及写作动机产生了积极影响（Liu ,2024）。

随着大语言模型在教育领域的深入应用，其影响力正从传统学科向新兴教育领域快速扩展，不仅赋能传统教学，还为创新教育模式注入了独特价值。这种技术的广泛应用不仅推动了教育手段的革新，也为学生的学习体验和教学效果带来了深远的积极影响。一方面，生成式人工智能（GAI）和自动语音识别（ASR）技术为外语学习带来了新的工具，尤其是 GAI 聊天机器人。比如用 GAI 聊天机器人 CoolE Bot 对与小学英语学习者进行口语交流。这些技术能够模拟真实的语言环境，提供个性化的学习体验，并给予实时反馈，推动了教育手段的革新（Tai，et al, 2024）。另一方面，ChatGPT 作为一种非目标导向的聊天机器人，能够显

著支持 EFL 学习者的语言学习过程，尤其是在写作、阅读、语法和词汇方面；也能给学生带来非同寻常且新奇的学习体验，以提高教学效果（Nguyen Hoang,et al,2024）。

在英语作为第二外语的学习过程中，学习者需要掌握与外语相关的知识和技能，而聊天机器人已被视为一种重要的语言学习工具（Lin & Mubarok, 2021; Yang et al,2022）。然而，由于课堂时间和资源的限制，教师需要为所有学习者提供及时且有针对性的反馈是一项具有挑战性的任务（Luo, 2016; Wang et al., 2022）。对于外语非母语学习者来说，聊天机器人提供了便捷的学习机会（Jeon，2024），使他们可以不受时间或地域的限制参与学习（(Chen et al.,2020）。学习者可以在模拟场景中与聊天机器人交谈，获得高效、多样化和个性化的学习体验，且不受同伴压力的影响（Moussalli & Cardoso，2020）。

总之，大语言模型（尤其是 OpenAI 的 GPT 系列）正在教育领域释放出前所未有的潜力。这些模型通过深度学习技术处理海量文本数据，能够精准适配不同教育层次和语言技能的需求。在低年级教育中，它们被广泛应用于提高学生阅读能力的练习活动，并显著提升学生的外语写作能力。与此同时，在口语等语言技能的教学设计中，大语言模型展现了卓越的优化能力，为教学方法注入了创新活力。此外，这些模型不仅极大地改善了学生的学习体验，还通过个性化指导显著提升了学生的批判性思维能力。对于教师而言，它们提供了强大的教学支持工具，帮助教师更高效地设计课程、评估学生表现，并实现因材施教。总的来说，大语言模型在教育领域的应用正释放出巨大的潜力，展现出无限广阔的应用前景。随着技术的不断进步，它们有望成为未来教育生态系统中不可或缺的核心驱动力。

3. 研究方法

系统文献综述法最初定义为“对医疗保健干预效果做出的明确的、可重复的和最新的总结”（Pollock & Berge, 2018），主要应用于医学领域，后逐渐广泛应用于管理研究中。为此，国际研究者制定了 PRISMA 声明（Preferred Reporting Items for Systematic Reviews and Meta-Analyses），以规范系统综述文章的标准。本文使用系统文献综述法，对 2022 至 2025 年关于大语言模型在 K-12 教育 EFL 中应用的文献进行统计分析，并使用 VOSviewer 软件对文献数据进行可视化分析，使研究结果更加直观。

4. 研究过程

在确定研究问题后，本研究遵循以下步骤进行：选择关键词、检索关键词、初步筛选文献、深入分析文献、最终得出研究结果。在文献检索过程中，我们严格遵循 PRISMA（系统综述和荟萃分析的首选报告项目）程序。2021 年国际知名专家团队联合发布了 PRISMA 2020 版声明并更新了条目清单，进一步规范了系统文献综述法的研究流程。初步检索后，我们首先通过审查文献的标题和摘要，筛选出可能符合纳入标准的研究。接下来，对这些初步筛选的文献进行更深入的全文审查，以确定它们是否完全符合系统综述的纳入标准。符合标准的文献将被选中进行详细分析。在整个过程中，研究人员需要准确记录并报告每个步骤中检索、筛选、评估和纳入的研究数量，以及排除研究的具体原因。最后，研究人员将针对之前提出的研究问题进行深入讨论和总结。

4.1. 检索关键词

在制定检索策略时，输入尽可能多的关键术语至关重要，这能避免忽略重要文章。本研究采用布尔搜索策略确定和检索关键词，使用“AND”和“OR”运算符有效检索文献，检索前先对关键词分类，再按类别扩展更多相关关键词，以确保全面检索到所需文章（Wright, Brand, Dunn, & Spindler, 2007）。根据研究的主题，我们将关键词分为“EFL”、“LLMs”和“K-12 education”三大类。在“EFL”这一类中，我们增添了“AI-assisted language learning”、“Language learning”、 “Second language”、“Foreign language”、“EFL”、“ESL”、“Language

acquisition”和“Language teaching”这几个关键词，较为全面地涵盖了第二外语学习和教育。在“ChatGPT”这一类中，我们选定了“ChatGPT”、“AIGC”、“Generative AI”、“artificial intelligence”、“OpenAi”、“Large language models”、“LLM”、“LLMs”和“chatbot”，包含了与ChatGPT有关的各种概念和相关描述。在“K-12 education”这一类中，我们选取了“K12 education”、“K-12”、“primary education”、“middle school”、“high school”、“kindergarten”，囊括了K-12教育的所有学段。在每个类别中，关键词之间用“OR”进行组合，两个类别的关键词之间用“AND”连接。本研究在Web of science核心合集数据库中进行检索，考虑到OpenAi是在2022年开发ChatGPT，我们新增了“出版时间”这一限制条件，使论文的时间在2022~2025年之间，进一步保证论文检索的严谨性。论文检索的条件限制如下图所示。

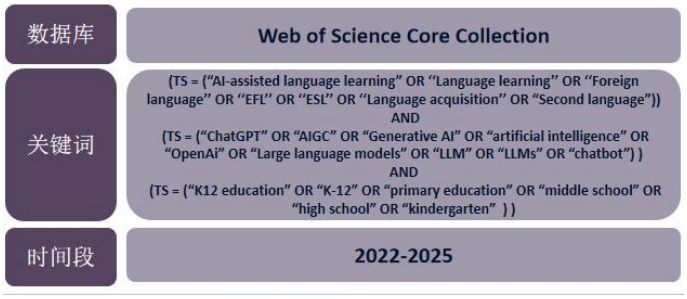


图1 论文检索条件

4.2. 筛选文献

确定检索范围后，共有625篇文献入选，接下来需要将入选的文献进行筛选。在初步筛选中，研究人员排除了420篇文献，剩下205篇文献需要进一步排查。在二次筛选的过程中，研究人员排除了未明确提及大语言模型、EFL和K-12教育的文献（n=181）。最后，研究人员对剩下的24篇文章进行了第三次筛查，通过精度文献与仔细评估，继续排除了3篇论文，其中2篇无法检索，1篇未发表。通过严格的筛查流程（见图2），最终21篇文献被纳入系统综述的数据库中，为后面的数据分析奠定基础，选取和排除的标准如表1。

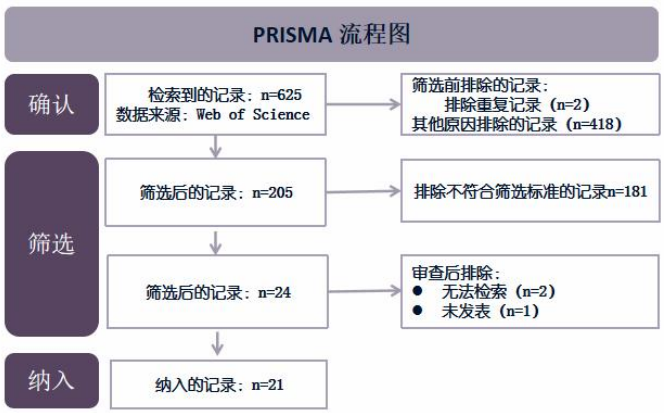


图2 PRISMA流程图

表1 选取和排除标准

选取标准	排除标准
文献内容与K-12中的学段相关	文献内容不包含K-12任何学段（即幼儿园、小学、初中、高中）
文献内容有明确提到大语言模型	文献内容并未提及大语言模型
文献内容包含EFL	文献内容不包含EFL
文献发表时间于2022年及以后	文献发表时间在2022年以前

4.3. 分析文献

确定好文献的数据库后，研究人员使用 VOSviewer 进行关键词的聚类分析。VOSviewer 是由荷兰莱顿大学开发的一款文献计量分析软件，也是一个知识图谱可视化工具。它支持标签视图、密度视图、聚类视图和分散视图四种浏览方式，帮助研究者更直观的进行文献分析。文献分析的过程按照“导入数据——清洗数据——生成可视化图谱”的步骤进行。其中在清洗数据的环节，研究人员需要剔除重复表达的数据以保证知识图谱的简洁性与准确性。在生成可视化图谱环节，研究人员需要将同类的关键词整理在同一个聚类，作为研究结果总结的前期准备工作。

5. 研究结果

5.1. 文献基本信息分析

关于文献的分布，符合标准的文献呈现出多样化的地理分布。主要集中在亚洲、欧洲和美洲。这些地区包括英国、美国、中国、韩国、新加坡等。这种分布显示了研究的国际化特征，表明相应的研究话题在全球范围内受到重视。特别值得注意的是，最后筛选出的 21 篇文章全部是实证研究，尤其是那些运用实验法的研究。然而，相比之下，使用访谈法进行的实证研究在入选文献中相对缺乏。

关于文献的研究方法，大部分使用了混合的研究方法，除了做定量的数据分析之外，还针对某些研究问题进行了定性分析。这种方法的优势在于，它能够结合定量数据的客观性和定性数据的深度见解。单独使用定量或定性方法的文献数量也相当，研究者能够在统计验证和深入理解之间取得平衡，从而获得更全面的研究结果。

关于文献的内容，涵盖协作类型、教学水平、语言技能和教师角色四个主要维度。在协作类型方面，ChatGPT 主要作为助教（66.7%，n=14）和学习伙伴（33.3%，n=7）参与教学活动，且作为助教的研究占比显著高于作为学习伙伴的研究。从教学水平来看，文献中讨论的学习者的教育水平涵盖了高中、初中、小学。初等教育（n=7，33.3%）比中等教育（n=12，57.1%）所占的比例略少，除此之外，部分研究所代表的教育水平为初等教育和中等教育混合，统一概括为 K-12，表明中等教育阶段是研究的重点领域。在语言技能分布上，写作技能占比最高（38.1%，n=8），其次是口语（28.5%，n=6）和综合技能（23.8%，n=5），词汇和阅读技能仅各占 4.8%（n=1），显示出写作技能的研究显著多于其他语言技能。至于教师角色，其分布呈现多样化，教师辅助角色占比最高（52.4%，n=11），其次是无教师参与（23.8%，n=5）和教师主导（23.8%，n=5），表明教师在大语言模型相关教学中的辅助性作用受到较多关注。总体而言，研究在协作类型、教学水平、语言技能和教师角色方面均呈现出一定的集中性和不平衡性，为后续研究提供了重要的参考方向。如下表所示。

表 2 文献基本信息表

分类	详细内容	样本量（%）	对应文献（见附录）
协作类型	作为助教	14（66.7%）	ID1 、 ID5 、 ID6 、 ID8 、 ID11、ID12、ID13、ID14、 ID15、ID16、ID18、ID19、 ID20、ID21
	作为学习伙伴	7（33.3%）	ID2 、 ID3 、 ID4 、 ID7 、 ID9、ID10、ID17、

教学水平	初等教育	7（33.3%）	ID1、ID8、ID9、ID11、ID15、ID16、ID17、
	中等教育	12（57.1%）	ID4、ID5、ID6、ID7、ID10、ID12、ID13、ID14、ID18、ID19、ID20、ID21
	K-12 教育	2（9.6%）	ID2、ID3、
语言技能	综合技能	5（23.8%）	ID2、ID12、ID13、ID15、ID20
	口语	6（28.5%）	ID3、ID4、ID9、ID16、ID17、ID19、
	写作	8（38.1%）	ID1、ID5、ID6、ID7、ID8、ID10、ID18、ID21
	词汇	1（4.8）	ID11
	阅读	1（4.8%）	ID14
教师角色	教师主导	5（23.8%）	ID1、ID8、ID11、ID13、ID17
	教师辅助	11（52.4%）	ID2、ID3、ID4、ID5、ID6、ID9、ID10、ID14、ID15、ID18、ID21
	无教师参与	5（23.8%）	ID7、ID12、ID16、ID19、ID20

5.2. 文献关键词分析

该关键词的聚类分析图谱系统整合了大语言模型与 K-12 阶段 EFL 教育研究的核心议题，通过技术工具、认知理论、行为机制及教学实践四维框架，揭示了人工智能时代语言教育生态的复杂关联。技术工具维度中，“automated writing evaluation”与“artificial intelligence chatbot corrective feedback”等关键词突显以 ChatGPT 为代表的生成式 AI 在 EFL 教学中的功能定位，其自动化评估与即时反馈机制已实现从辅助工具到教学流程深度整合的范式转变。值得注意的是，“ELIZA”与当代模型的共现不仅映射对话系统技术的历时性演进，更暗示教育技术需在继承性创新中实现场景适配。认知理论维度通过“chinese dual code theory”与“language learning information technology”的关联，阐释了双编码理论的多模态认知机制如何为增强现实语言教学提供科学基础——该理论主张语言符号与非语言表征的协同加工，恰与“augmented reality english”的具身学习设计形成理论-技术闭环。行为机制维度中，“motivation”、“acceptance”等心理变量与“determinants”、“agents”等系统要素的交织，揭示技术落地本质上是多主体参与的复杂适应过程。研究表明，AI 工具效能不仅受算法准确性的单维度制约，更需考量教师技术采纳意愿、学生认知负荷阈值以及课堂技术生态的协同效应。教学效果维度的“accuracy”与“interference”构成辩证张力：大语言模型虽能通过海量语料训练提升语言形式准确性，但其过度纠错可能抑制学习者的容错阈值，而母语负迁移与技术依赖更可能消解“creative writing”所需的元语言意识，这种矛盾凸显 AI 工具需在精准性与教育性之间建立动态平衡。

值得注意的是，图谱的网络化特征在"applications in subject areas language classroom experiences"节点得到集中体现，该枢纽通过连接技术实践、理论框架与教学情境，构建起"技术具身化-理论操作化-场景适应性"的三元结构。这提示教育技术研发需超越简单的工具移植，转而追求教学场景驱动的技术再语境化（recontextualization）过程。

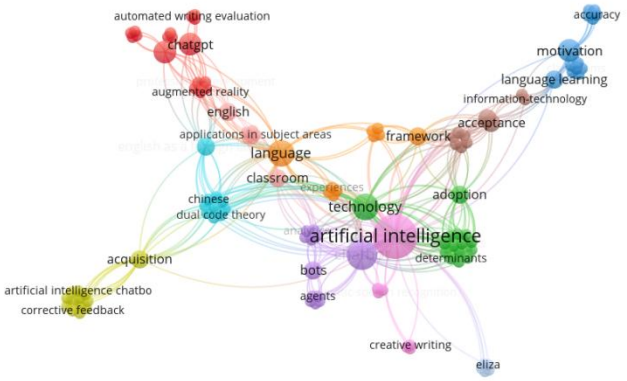


图 3 关键词聚类图谱

密度可视化图谱通过色阶梯度映射了大语言模型在 K-12 阶段 EFL 教育研究的领域聚焦度，其中“automated writing evaluation”、“ChatGPT”及“augmented reality”构成了深色核心区，表明自动化写作评估、生成式 AI 及增强现实技术已成为当前研究的热点。围绕核心区的“corrective feedback”（纠错反馈）、“dual code theory”（双编码理论）与“agents”（智能代理）等次级概念，揭示了技术应用与认知理论深度融合的特征——例如，双编码理论为 AR 英语教学的多模态交互提供了认知框架，而智能代理系统则支撑了个性化反馈的实践落地。图谱的扩展趋势体现在“language classroom experiences”、“information-technology”的色域过渡上，暗示研究重心正从工具效能验证转向技术生态构建，如信息技术如何重塑语言课堂的师生互动范式。值得注意的是，“analytical intelligence chatbot”（分析型智能对话系统）与“determinants”（技术采纳动因）作为新兴节点，其浅色分布反映了前沿探索方向：前者指向超越基础对话功能的认知分析能力开发，后者则关注教师技术接受度的多变量影响机制。而边缘区域的“Eliza”（早期对话系统）与“creative writing”（创意写作）形成鲜明对比，前者作为技术演化的历史锚点，后者则暴露出现有研究对 AI 工具可能抑制语言创造力的隐忧，此类话题的稀疏分布恰凸显了技术伦理研究的不足。

在应用场景维度，核心技术在“subject areas”（学科整合）与“classroom”（课堂教学）中已展现结构化渗透，但在“acquisition”（语言习得机制）与“motivation”（学习动机维持）等深层教育过程的研究密度相对较低。特别是跨文化语境下的技术适配问题（如“chinese”节点的孤立性），以及 AI 反馈对元语言意识发展的长周期影响，尚未形成系统化研究网络。未来亟需在“human-AI collaborative framework”（人机协同框架）构建、“cognitive load optimization”（认知负荷优化）及“ethical boundary design”（伦理边界设计）等方向深化探索，这些空白领域或将决定大语言模型能否真正实现从"教学工具"到"教育范式"的质变跨越。

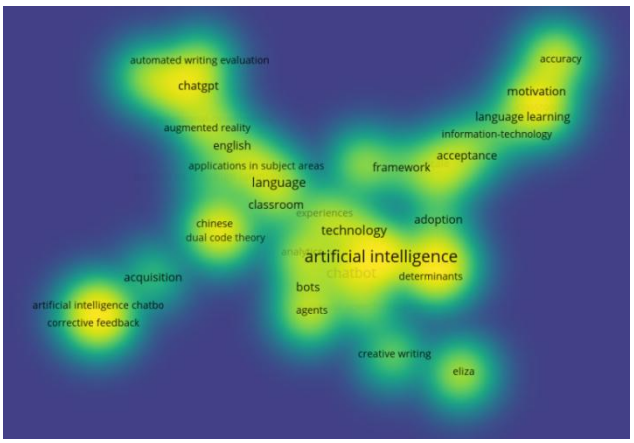


图 4 密度可视化图谱

6. 讨论

6.1. RQ1: 在 K-12 教育中，大语言模型支持下的第二外语学习的协作有哪些类型（即大语言模型在协作中的角色）？协作的情况如何?

在 K-12 教育中，大语言模型（LLMs）的协作类型主要分为两类，其中以助教角色为主以学习伙伴为辅。一方面，大语言模型在协作过程中以助教身份提供支持。例如，在写作训练中，大语言模型作为辅助工具，嵌入认知学术语言学习方法（CALLA，Cognitive Academic Language Learning Approach）框架，通过个性化支持、实时反馈和适应性交互，针对学习者的错误提供纠正性重铸反馈，帮助学习者练习和掌握英语致使移动结构，从而辅助语言学习并支持教师完成教学任务。然而，这种助教角色并未完全替代教师的主导地位，也不同于平等交流的学习伙伴角色，而是以辅助身份协助学生进行写作练习（Liu.et al, 2024）。此外，在词汇学习中，基于人工智能的物体检测翻译（AI-based ODT）应用程序以助教身份为学生提供词汇学习支持，包括识别物体、翻译和提供发音等，但并不主导教学过程，而是辅助学生完成词汇学习任务（Liu & Chen，2023）在阅读练习中，聊天机器人（ChatGPT）通过提示工程技术（prompt engineering）协助教师自动生成英语阅读理解问题，减轻教师负担，以助教身份支持教学活动（Lee.et al,2024）此外，“AI 教练”（AI Coach）作为一种虚拟代理，通过个性化反馈帮助学生进行口语和听力练习、学习词汇以及纠正发音问题，从而培养学生的语言综合技能（Wang.et al,2023）。另一方面，大语言模型以学习伙伴的身份陪伴学生学习。它能够与学生互动交流，提供学习资源、协助完成学习任务，并满足日常交流需求，从而支持学生的自主学习。例如，学生在使用 ChatGPT 时，将其视为能够提供帮助和支持的伙伴（Nguyen.et al,2024）。此外，聊天机器人“Ellie”被设计为与学生进行英语对话的伙伴，通过鼓励学生参与交流，辅助提升其语言能力，承担学习伙伴的角色，以提高学生的口语交际能力（Yang.et al,2022）。

总之，大语言模型作为一种强大的人机协作工具，在以上两个个方面发挥着积极的作用。作为助教，可以提供个性化支持、实时反馈和适应性交互，帮助学生掌握语言结构，辅助教师完成教学任务，但不替代教师的主导地位；作为学习伙伴，通过互动交流和资源支持，帮助学生完成学习任务，促进自主学习。这种双重协作模式表明，大语言模型在 K-12 教育中既可以作为助教提供教学支持，也可以作为学习伙伴促进学生的自主学习。

这些研究发现揭示了大语言模型在 K-12 教育中的 EFL 作为协作工具的多样化应用潜力。其多重协作类型为教育创新提供了多维视角，推动了个性化学习的发展，并促进了学生的全面成长。综合来看，大语言模型在 K-12 教育中的 EFL 应用不仅丰富了教学实践，还预示着教育领域的持续创新与进步。

6.2. RQ2: 在 K-12 教育中，大语言模型应用在哪个教育阶段和哪个语言技能，是如何实施的？实施的结果如何？

大语言模型在 EFL 领域应用多样化，涉及不同教育阶段以及不同语言技能，包括：在高中生口语培训方面，聊天机器人“Ellie”通过提供对话机会和任务导向的交流来支持韩国高中生的口语技能提升（Kim.et al,2022）；在高中写作领域，基于 ChatGPT 的聊天机器人（KR-Automated Essay Scoring System）主要用于辅助教师对学生的英语作文进行评分和提供反馈，通过分析学生作文的任务完成情况、内容、结构和语言使用等方面，给与学生及时、个性化的反馈（Shin1 & Lee，2024;Luo.et al,2025）；在小学口语训练中，Dialogflow 开发的聊天机器人与学生进行结构化、重复性对话，提供反馈，帮助学生练习英语口语，在练习过程中如同伙伴一样与学生互动交流，通过互动式口语测试等方式评估学生的口语能力，重点关注口语技能的提升（Jeon & Lee,2024）。

研究主要集中在中等教育阶段和写作领域，而初等教育甚至幼儿园教育的应用则相对较少，尤其是在词汇和阅读能力培养方面，仍存在较大的空白。这些发现不仅展示了大语言模型在教育领域的应用潜力，也揭示了其在特定场景中的局限性。尽管研究覆盖了中小学教育阶段，但幼儿园阶段的应用似乎较为欠缺。这可能是因为幼儿园教育更注重亲身体验和人际交往，对机器学习工具的需求相对较低。然而，大语言模型的应用并不仅限于学生，还涉及教育者。大语言模型在辅助教育者设计课程和开发评估工具方面具有显著潜力(Cooper, 2023)。

在实施过程中，大语言模型在提升学生技能、情感和学业表现方面展现出显著的积极作用。具体而言，在技能方面，基于人工智能的英语学习支持系统通过分析文本和视频内容生成个性化学习材料，帮助学生高效学习英语词汇；同时，系统提供的学习计划和自主学习功能促使学生深入阅读和理解英语文本，而测试模式及内容整理分享活动则进一步锻炼了学生的写作与表达能力，最终全面提升其语言综合应用能力（Lee.et al,2023）。情感层面，人工智能交互支架（AIIS）凭借动态反馈与个性化支持，助力学生设定学习目标并激发学习兴趣，从而显著增强其内在学习动力。这种动机提升与积极情感体验紧密相连，表明 AIIS 能够有效改善学生在非正式学习环境中的情感状态（Wang.et al,2025）。在学业表现上，聊天机器人写作练习对学生的写作能力提升效果显著。实验数据显示，实验组学生的后测写作成绩明显优于对照组，充分证明了聊天机器人在写作教学中的应用价值。此外，实验组学生普遍反馈，聊天机器人有助于复习课堂内容、提高写作技能，且在使用过程中减少了对犯错的焦虑感，表现出更高的学习舒适度（Kwon.et al,2023）。总之，大语言模型可以通过个性化学习支持、情感激励和互动练习，显著提升学生技能、情感和学业表现，助力教育创新与效率提升。

6.3. RQ3: 在 K-12 教育中，教师在大语言模型应用到语言学习的角色是什么样的？态度如何？对未来发展有什么建议？

教师在大语言模型应用到语言学习的角色是主要是辅助角色，还有部分是教师主导或者无教师参与。比如教师为学生介绍 ChatGPT、教授提示工程技能、提供指导和翻译等，但学生是在教师引导下自主与聊天机器人互动完成写作任务，给与学生学习自主权的同时减轻教师负担（Woo.et al,2024;Tai & Chen,2024）。同时，人工智能（AI）在英语教学中的应用，AI 的角色是作为助教，教师是作为教学的主导者。教师在教学过程中决定是否使用 AI 以及如何使用 AI 来辅助教学或者教授学生零样本提示方法和数字写作技能等，引导学生利用工具写作。更有甚者无教师参与，比如人工智能辅助语言学习（AILL）为学生提供如自动评分、反馈、个性化指导等帮助，辅助学生进行英语学习，再如 PengTalk 的语音识别聊天机器人帮助学生练习发音、纠正错误，辅助学生进行英语学习，但并不主导教学过程。且无教师参与，由此让学生自主学习、实现真正的教育放权（An.et al,2023）。

大语言模型对 K-12 教育中的 EFL 的影响是复杂多样的。研究表明，教师对 ChatGPT 的

认知程度和使用态度仍存在差异。部分教师对 AI 工具持欢迎态度，认为其能减轻教学管理和写作批改的工作量；但也有教师持谨慎态度，对 AI 工具的能力有所怀疑（Luo.et al,2025）。此外，教师们普遍认为 ChatGPT 在生成问题时表现出一定的有效性，但也存在一些问题。例如，ChatGPT 生成的“是/否”问题因句子结构复杂，可能分散学生注意力，导致学生花费更多时间理解问题结构而非内容。因此，教师建议在使用 ChatGPT 生成问题时提供更详细的提示和指导，并对生成的问题进行验证和调整，特别是在复杂任务和推断性问题上，以提高问题的准确性和相关性（Lee.et al,2024）。还有的建议教师要了解学生向生成式 AI 工具提问的目的，从而提供有针对性的指导和支持，帮助学生更好地利用 AI 工具进行写作（Woo.et al,2024）。

总体而言，教师和学生对大语言模型在 EFL 中的应用持复杂态度。对于其辅助教学设计、提升学习体验和学习动机的潜力，师生普遍持积极看法，并希望通过进一步的学习与培训，更好地将其作为一种学习工具加以利用。然而，在信息准确性等问题上，师生则表现出审慎甚至负面的态度，认为大语言模型可能对学习环境构成潜在风险。因此，大语言模型在教育中的应用亟需进一步规范，未来的研究应填补其在实证研究中的空白，以形成对大语言模型在 EFL 中更全面、更深入的认识。

7. 总结

本研究通过对 2022 年至 2025 年间发表的同行评议文章进行系统回顾和元分析，深入探讨了大语言模型在 K-12 教育阶段英语作为外语教学中的应用效果。研究结果显示，大语言模型在显著提升学习者英语学习成效方面发挥了重要作用，不仅有效促进了学习者的语言技能发展，还在提升其心理和情感幸福感方面展现出积极影响。这些发现为大语言模型在 EFL 教育中的应用提供了有力的实证支持，表明其在优化教学方法和学习体验方面具有巨大潜力。

本次研究揭示了大语言模型在 K-12 教育中 EFL 应用的不足，主要体现在对幼儿教育等阶段及某些语言技能研究的缺失，以及对技术潜在风险关注的不足。未来研究需聚焦于技术在这些薄弱领域的应用，并重视教育公平和学术诚信。同时，技术开发应围绕教育核心价值，如依学习理论开发多模态学习工具，利用增强现实技术打造沉浸式学习环境。还需构建全面的技术伦理框架，平衡效率与人文关怀，如设计反馈敏感度调节器，通过边缘计算缩小数字鸿沟。此外，跨学科研究方法创新至关重要，建议采用设计性研究范式，在真实教育场景中迭代验证 AI 工具，并构建教师-AI 协同教学模型，重新定义教师角色。总之，只有技术逻辑与教育逻辑深度融合，人工智能才能真正推动 EFL 教育革新。

参考文献

夏立新.ChatGPT 对教育的多重变[J].国家教育行政学院学报,2023,(3):9-12.

刘天丽, 杨现民, 李康康, & 糜泽稼. (2023). 社会公众如何看待“ChatGPT+教育”？——基于国内主流开放社区的文本分析. 现代教育技术, 33(10), 14－23.

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.

王天恩.ChatGPT 的特性、教育意义及其问题应对[J].思想理论教育,2023,(4):19-25.

吴河江, 吴砥. (2024). 生成式人工智能教育应用：发展历史、国际态势与未来展望. 比较教育研究, 2024(6), 13-23.

AlAfnan M A, Dishari S, Jovic M, et al. ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses[J]. Journal of Artificial Intelligence and Technology, 2023,(2):60-68.

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling

English teachers' behavioral intention to use artificial intelligence in middle schools. Education and Information Technologies, 28, 5187–5208.

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling students' perceptions of artificial intelligence assisted language learning. Computer Assisted Language Learning.

Chen, H. H. J., Yang, C. T. Y., & Lai, K. K. W. (2020). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. Interactive Learning Environments. Advance online publication.

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

Cooper, G. (2023b). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. Journal of Science Education and Technology, 32(3), 444–452.

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. Language Learning & Technology, 24(1), 16–26. https://doi.org/10125/44705

Heidt, A. (2023). 'Arms race with automation': professors fret about AI-generated coursework. Nature.

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. Education and Information Technologies, 28(12), 15873–15892.

Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. Computer Assisted Language Learning. Advance online publication.

Jeon, J., Lee, S., & Coronel-Molina, S. M. (2024). Rethinking AI: bias in speech-recognition chatbots for ELT. ELT Journal, 78(4), 435–445.

Jeon, J., & Lee, S. (2024). The impact of a chatbot-assisted flipped approach on EFL learner interaction.Educational Technology & Society, 27(4), 218-234.

Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. Language Learning & Technology, 26(1), 1–18.

Kim, R. (2024). Effects of learner uptake following automatic corrective recast from Artificial Intelligence chatbots on the learning of English caused motion construction. Language Learning & Technology, 28(2), 109–133.

Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool.Language Learning & Technology, 27(1), 1–19.

Lee, D., & Kim, H. (2023). Development research on an AI English learning support system to facilitate learner-generated context-based learning. Education Technology & Development, 71(6), 629–666.

Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in English education. Education and Information Technologies, 29, 11483–11515.

Lin, C. J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. Educational Technology & Society, 24(4), 16–35.

Liu, C. C., Liao, M. G., Chang, C. H., & Lin, H. M. (2022). An analysis of children'interaction with an AI chatbot and its impact on their interest in reading. Computers & Education, 189, 104576.

Liu, P.-L., & Chen, C.-J. (2023). Using an AI-Based Object Detection Translation Application for English Vocabulary Learning.Educational Technology & Society, 26(3), 5-20.

Liu, Z. M., Hwang, G. J., Chen, C. Q., Chen, X. D., & Ye, X. D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. Computer Assisted Language Learning, 1-25.

Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for

efficient classroom instruction. Computer Assisted Language Learning, 29(3), 451476.

Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learners ’ argumentative writing.Education and Information Technologies.

Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? Computer Assisted Language Learning, 33(8), 865–890.

Pollock, A., & Berge, E. (2018). How to do a systematic review. International Journal of Stroke, 13(2), 138–156.

Shin, D., & Lee, J. H. (2024). Exploratory study on the potential of ChatGPT as a rater of second language writing. Education and Information Technologies, 29, 24735–24757.

Tai, T. Y., & Chen, H. H. J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. Interactive Learning Environments. Advance online publication.

Tai, T. Y., & Chen, H. H. J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. Computers & Education, 220, 105112.

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. System,103528.

Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students ’ speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English.Interactive Learning Environments.

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis.Computers & Education, 194, 104703.

Wang, X., Pang, H., Wallace, M. P., Wang, Q., & Chen, W. (2022). Learners' perceived AI presences in AI-supported language learning: A study of AI as a humanized agent from community of inquiry. Computer Assisted Language Learning, 37(4), 814–840.

Woo, D. J., Susanto, H., Yeung, C. H., Guo, K., & Fung, A. K. Y. (2024). Exploring AI-Generated text in student writing: How does AI help? Language Learning & Technology, 28(2), 183–209.

Woo, D. J., Wang, D., Susanto, H., & Wallace, Y. (2024). Teaching EFL students to write with ChatGPT: Students ’ motivation to learn, cognitive load, and satisfaction.Education and Information Technologies, 29, 24963 – 24990.

Wright, R. W., Brand, R. A., Dunn, W., & Spindler, K. P. (2007). How to Write a Systematic Review: Clinical Orthopaedics and Related Research, 455, 23–29.

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. ReCALL, 34(3), 327–343.

附录

ID	出版年份	研究方法	文献名
1	2024	定量研究法	Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation
2	2024	混合研究法	ChatGPT as a tool for self-learning English

			among EFL learners: A multi-methods study
3	2022	定量研究法	Implementation of an AI chatbot as an English conversation partner in EFL speaking classes
4	2022	定量研究法	Design principles and architecture of a second language learning chatbot
5	2024	定量研究法	Effects of learner uptake following automatic corrective recast from Artificial Intelligence chatbots on the learning of English caused-motion construction
6	2024	定量研究法	Exploratory study on the potential of ChatGPT as a rater of second language writing
7	2024	定性研究法	Exploring EFL students' prompt engineering in human-AI story writing: an activity theory perspective
8	2023	混合研究法	The application of chatbot as an L2 writing practice tool
9	2024	混合研究法	Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions
10	2024	混合研究方法	Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process
11	2023	定量研究法	Using an AI-Based Object DetectionTranslation Application for English Vocabulary Learning
12	2023	定量研究法	Modeling students' perceptions of artificial intelligence assisted language learning
13	2023		

		定量研究方	Modeling English teachers' behavioral intention to use artificial intelligence in middle schools
14	2024		
		混合研究法	Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education
15	2023		
		混合研究法	What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis
16	2024		
		混合研究法	Rethinking AI: bias in speechrecognition chatbots for ELT
17	2024		
		混合研究法	The impact of a chatbot-assisted flipped approach on EFL learner interaction
18	2024		
19	2025	混合研究法	Exploring AI-Generated text in student writing: How does AI help?
		混合研究法	The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English
20	2023		
		混合研究法	Development research on an AI English learning support system to facilitate learner-generated-context-based learning
21	2025		
		定量研究法	The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing

Multisensory Learning Method in Plant Science for Elementary School Students with Visual Impairments

Chu-Yu, Cheng[1*], Chia-Hui, Feng[2], Chi-Hua, Wu[3]
[1] Southern Taiwan University of Science and Technology
[2] Southern Taiwan University of Science and Technology
[3] Southern Taiwan University of Science and Technology
[*] qa6@stust.edu.tw

Abstract: *This paper presents a multisensory teaching approach for elementary-level plant science education designed for visually impaired students. Following Universal Design principles, we developed learning materials that combine tactile 3D models, varied textures, audio guidance, and visual elements to create an inclusive environment benefiting both visually impaired and sighted students. Our approach addresses practical challenges in teaching plant biology, including seasonal limitations and extended timeframes required for observation. By integrating VRAK (Visual, Reading/writing, Auditory, Kinesthetic) learning strategies, we make abstract concepts tangible and accessible. We discuss the theoretical framework, design process, implementation, and contributing to knowledge on inclusive educational design.*

Keywords: universal design, multisensory learning, visual impairment, inclusive education, plant science

1. Introduction

Educational technology has revolutionized learning by visualizing abstract scientific phenomena, with recent innovations like metaverse concepts providing students unprecedented access to "invisible" worlds. However, this emphasis on visual learning and accelerated digital content can marginalize "tactile awakening" and "slow learning" approaches. As a result, students with special needs, especially those with visual impairments or learning disabilities, often find themselves at a considerable disadvantage when trying to access and engage with educational materials. Metatla et al. (2018) noted that while "visuals" represent one of technology's greatest educational contributions, properly leveraging technology's characteristics can provide more appropriate teaching tools and equitable assessment methods to achieve educational equality. This paper describes a project that employs a multisensory approach to develop teaching aids for the "Plant Body" unit in science curricula for 8-9 year old students. Teaching plant biology in late autumn creates practical challenges, as most plants are entering dormancy during this season, and growing specimens from seeds to mature plants requires more than a month—often extending into winter when conditions are less favorable. Appropriate teaching aids can therefore significantly enhance students' understanding of plant biology concepts despite seasonal limitations.

2. Theoretical Framework

Our approach integrates Universal Design principles with Body Tactile Learning strategies as recommended by Professor Tseng, a special education expert on visual impairment. Burgstahler (2005) defined Universal Design for learning as "the design of instructional materials and activities that make learning goals achievable by individuals with wide differences in their abilities to see, hear, speak, move, read, write, understand English, attend, organize, engage, and remember."

The project employs VRAK (Visual, Reading/writing, Auditory, Kinesthetic) learning theory across populations to develop inclusive teaching aids that integrate knowledge, affective, and skill-based objectives. This approach overcomes limitations of season, weather, and insufficient campus facilities while providing teachers with educational materials that support both itinerant teachers for visually impaired students and frontline inclusive education teachers.

3. Design and Development

The project builds upon previous teaching strategies while adding "kinesthetic tactile" elements through multi-material models with positioning cognition for model assembly. Key technologies employed include:

1. **Enhanced Tactile Materials:** Rich surface textures with pronounced height differentiation and three-dimensional elevated printing using multiple materials.

Figure1
Example of tactile illustration

2. **Disassembly 3D Models:** Modular plant components that can be assembled and disassembled to understand structural relationships.

Figure2
3D Models

3. **Complete Audio Guidance:** Voice narration of all content to support independent learning.

Figure3
Interactive device

4. Preliminary Findings

Initial classroom testing indicates positive reception from both teachers and students. Visually impaired students demonstrate enhanced comprehension of plant structures through the tactile models, while sighted students benefit from the multisensory approach by developing deeper understanding than traditional visual-only methods.

The audio guidance component proves particularly valuable for independent exploration, allowing visually impaired students greater autonomy in their learning process. Teachers report increased engagement across the entire classroom and note that the physical manipulation of plant components improves retention of structural knowledge.

5. Conclusion

Our project demonstrates how Universal Design principles can be effectively applied to natural science education in inclusive settings. The multisensory approach not only accommodates the needs of visually impaired students but enhances learning for all students through engaging, hands-on experiences. Future development will include expanded subject coverage and longitudinal studies to assess long-term learning outcomes. This project contributes to the growing body of knowledge on accessible educational design and demonstrates the value of cross-disciplinary collaboration in developing inclusive teaching methodologies.

Acknowledgements

We gratefully acknowledge the funding support from the National Science and Technology Council (NSTC) of Taiwan under grant NSTC 112-2410-H-218-007-MY2. We also thank all the collaborators, teachers, and students who participated in this project and provided valuable feedback.

References

Burgstahler, S. (2005). Universal design of instruction: Definition, principles, and examples. DO-IT, University of Washington.

Metatla, O., Serrano, M., Jouffrais, C., Thieme, A., Kane, S., Branham, S., Brulé, É., & Bennett, C. (2018). Inclusive education technologies: Emerging opportunities for people with visual impairments. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

結合行動科技脈絡線索鷹架與競標機制的桌上遊戲以培訓採購策略思維之初探

A Preliminary Study on a Board Game Integrating Mobile Technology Contextual Clue Scaffolding and a Bidding Mechanism to Train Procurement Strategy Thinking

李承泰[1]，黃玉如[2]，杜乙珊[2]，王筱喬[2]，李子鈺[2]，侯惠澤[1*]
[1]臺灣科技大學應用科技研究所
[2]基隆市八斗高級中學
[*] hthou@mail.ntust.edu.tw

【摘要】 本研究設計一款結合行動科技脈絡線索鷹架與競標機制的桌上遊戲遊以培訓採購策略思維。遊戲背景設定於基隆崁仔頂漁市場，遊戲元件包含角色卡與魚卡。玩家需以行動科技閱讀與分析卡片的脈絡線索資訊，確定正確的採購目標，同時權衡資金運用與制定整體最佳採購策略。本研究採用單一組前後測設計，對象為台灣高中二年級 24 名學生。研究結果顯示學生一定程度投入參與在遊戲活動中，且此競標式的遊戲並未造成學習者過度的焦慮和負荷。

【關鍵字】 行動科技；脈絡線索鷹架；競標機制；採購策略思維；桌上遊戲

Abstract: *This study designed a board game that integrates mobile technology contextual clue scaffolding and a bidding mechanism to train procurement strategy thinking. The game components including role cards and fish cards. Players use mobile technology to read and analyze contextual clue information on the cards, identify the correct procurement targets, and develop the optimal overall procurement strategy. This study adopted a single-group pretest-posttest design with 24 second-year high school students in Taiwan as participants. The results indicate that students were engaged in the game to a certain extent, and the bidding-based game did not cause excessive anxiety or cognitive load for the learners.*

Keywords: mobile technology, contextual clue scaffolding, bidding, procurement strategy thinking, board games

1. 前言

競標機制廣泛應用於市場交易與決策制定（Klemperer, 2002），個體需根據市場資訊與對手行為調整決策，以提高採購效益與競爭優勢。策略思維在此過程中扮演關鍵角色，特別是在需要即時決策與資源權衡的情境下（Camerer, 2011）。Lin and Hou（2024）指出，透過適當的情境設計與學習鷹架，學習者可在模擬環境中發展策略思維。

本研究設計一款結合行動科技脈絡線索鷹架與競標機制的桌遊「崁仔頂競標」，以培訓採購策略思維。遊戲故事背景設定在基隆崁仔頂漁市場，此市場的特色是買家以競標方式由出價最高者購得漁獲。遊戲配件包含 5 張角色卡和 9 張不同魚種的魚卡。角色卡正面有採購資金限制條件，卡片背面有一個 QR code，學生使用行動科技掃瞄 code 可顯示此角色的故事情境、採購任務與競標資金等線索資訊。魚卡正面有競標底價與線索，卡片背面有一個 QR code，內容為連結到角色卡採購目標的線索。這兩種在遊戲卡片上以行動科技建購的脈絡線索鷹架將用來幫助玩家分析與連結線索，進而思考採購與權衡資金運用的策略。

此遊戲為 4 位玩家一同競標，每人資金相同，1 位輔導員擔任莊家。每張角色卡的採購任務都對應到各 1 張最佳（競標買到得 10 分）、次佳（8 分）第三佳（6 分）魚卡，另外 6 張皆為不適合魚卡（0 分）。玩家第一回合同時扮演第 1 張角色卡（如：家庭主婦），閱讀角色卡和魚卡後，玩家各自確定採購目標與制定採購策略。莊家首先隨機出示第 1 張魚卡（如：鯖魚），詢問是否有玩家要競標。如有 2 位以上玩家欲購買同一張魚卡就出價競標，出價最高玩家購得魚卡。如只有 1 位玩家競標，則直接以底價購買。第一回合競標結束後，莊家公佈各玩家所得分數。整個遊戲結束後，玩家的資金剩越多，分數最高，則為贏家。

在競標過程中，玩家可使用行動科技脈絡線索鷹架，隨時調整自己的採購策略。例如：最佳魚卡目前競標價格過高，退而求其次採購次佳合魚卡。另一方面，如果玩家一味以最高價格競標魚卡，所剩資金相對就越少，這將會影響到後面角色卡的遊戲任務。在「崁仔頂競標」的競標遊戲機制中，玩家需分析每個角色的採購任務，不只要確定正確的採購目標，還要同時權衡資金運用與思考整體最佳採購策略，以在競爭的情境中成功完整遊戲任務。

圖 1 綠色角色卡與白色魚卡所建構的行動科技脈絡線索鷹架，並藉此串接卡片文字線索與載具呈現的情境脈絡線索，促進多重線索的策略思維能力

2. 研究方法與結果

本研究採用單一組前後測設計，初步探討學生的心流（Killi, 2006）、活動焦慮(Krashen, 1987)與認知負荷（Hwang et al., 2013）。研究對象為台灣高二學生共 24 人。此三份量表的信度（Cronbach's alpha）分別為 0.839、0.945 與 0.913，所有量表均採用李克特五點量表。遊戲活動時間為 30 分鐘，隨後進行三份問卷的填寫 15 分鐘。

研究結果的心流平均數為 3.76（標準差 0.56），表示學生一定程度的投入在遊戲活動中。活動焦慮與認知負荷平均數分別為 1.63（標準差 0.82）與 2.24（標準差 0.85），此結果顯示學生此競標機制的遊戲活動中並沒有感受到過度的焦慮，也沒有造成他們太大的負荷。

3. 討論

本研究設計一款結合行動科技脈絡線索鷹架與競標機制的桌上遊戲遊以培訓採購策略思維，初步研究結果顯示遊戲對心流、活動焦慮和認知負荷有正向影響。未來本研究將進行準實驗設計，調查此遊戲對學採購策略思維的有效性。此外，使用序列分析法探討學生在競標遊戲式學習活動的行為模式，將有助於了解學生在競爭機制桌遊的認知思考歷程與互動。

參考文獻

Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton university press.

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.

Klemperer, P. (2002). What really matters in auction design. *Journal of Economic Perspectives*, *16*(1), 169–189.

Krashen, S. D. (1987). *Principles and practice in second language acquisition*. Prentice-Hall.

Lin, Y. C., & Hou, H. T. (2024). Designing a strategic analysis and planning skills training board game using mobile technology and a dual-scaffolding mechanism. *The Asia-Pacific Education Researcher*, *33*(2), 431-445.

Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology*, *2*(2), 187-201.

The Impact of Junior High School Science Teachers' Professional Development on Beliefs and Practices of AI in Assessment

Kuan-Ying Chen[1], Yu-Che Hsieh[2], Chi-Shan Yu[3], Ping-Lin Fan[4], Ping-Han Cheng[5*,]
[125] Department of Science Education, National Taipei University of Education
[34] Department of Digital Technology Design, National Taipei University of Education
[*] phcheng@mail.ntue.edu.tw

Abstract: *International competency-based assessment has become a key indicator for evaluating the effectiveness of educational policies. PISA 2025 emphasizes decision-making and problem-solving, while AI offers potential to enhance teaching and science assessment. This study examined junior high school science teachers' beliefs about AI in assessment and their willingness to adopt it through a professional development program. Thirty science teachers participated in a three-hour workshop. Research* results *indicated a significant increase in teachers' recognition of AI as a learning tool (AILT), demonstrating the effectiveness of the workshop in deepening teachers' understanding and application of AI in science assessment. The findings affirm the potential of AI in science education assessment and offer valuable insights for future educational policy and professional development initiatives.*

Keywords: Teacher Professional Development, AI in Assessment, Teacher Beliefs, PISA 2025, Science Assessment

1. Introduction

In the context of science education, "assessment" typically refers to the process of collecting, analyzing, and interpreting data on students' performance during their learning journey. Its primary purpose is to evaluate students' comprehension and learning outcomes, as well as the effectiveness of teaching strategies and instructional methods (Kruit et al., 2020). With the increasing globalization and diversification of education, numerous international comparative studies have been developed to examine the effectiveness and impact of educational policies across countries. For example, the Trends in International Mathematics and Science Study (TIMSS) focuses on content knowledge, reasoning, and application, while the Progress in International Reading Literacy Study (PIRLS) tracks global trends in reading comprehension. Both assessments are administered by the International Association for the Evaluation of Educational Achievement (IEA). In contrast to curriculum-based evaluations, the Programme for International Student Assessment (PISA), conducted by the Organisation for Economic Co-operation and Development (OECD), is held every three years and targets 15-year-old students. It assesses competencies in mathematics, reading, and science literacy, with the primary domain rotating with each cycle. In response to evolving educational priorities, PISA has incorporated competency-based assessments that go beyond national curricula, aiming to reflect the outcomes of educational development and policy implementation on a global scale (Kim, 2021; Xu, 2023).

In PISA 2025, scientific literacy will again be the primary domain, following the previous focus in PISA 2015. Compared to the earlier framework, the 2025 assessment expands its scope to not

only include constructing scientific inquiry and interpreting data and evidence, but also newly incorporates the ability to research, evaluate, and apply scientific information in decision-making and action. Additionally, a new component—Learning in the Digital World (LDW)—has been introduced. These structural updates reflect broader shifts in the educational landscape and signal evolving international expectations for both student competencies and related educational policies. PISA 2025 places particular emphasis on students' data analysis, critical thinking, and problem-solving skills. In this context, artificial intelligence holds significant potential for application in the design and implementation of assessments (Deta et al., 2024; Wang et al., 2024; White et al., 2023).

The rapid advancement of artificial intelligence is reshaping the educational landscape, particularly in how teaching and assessment are approached. AI technologies are increasingly being applied to educational assessment through adaptive learning systems, intelligent tutoring, and learning analytics (Dogan et al., 2023; Ejjami, 2024; Hamal et al., 2022). However, in the context of science education, many teachers remain unfamiliar with AI tools and are skeptical about their accuracy (Choi et al., 2022; Gurer & Akkaya, 2021). Teachers' beliefs play a pivotal role in shaping their willingness and attitudes toward adopting AI in practice. Furthermore, inconsistencies in how AI is applied in assessments, along with misalignment with the principles outlined in science curricula, often require teachers to invest additional preparation time. These factors highlight the challenges and barriers to integrating AI into assessment practices (Herold et al., 2024; Holmes, 2019; Williamson, 2020). Ultimately, providing opportunities for professional development is key to enabling science teachers to effectively adopt and implement AI in assessment (Celik, 2023; Nazaretsky et al., 2022; Wang et al., 2023).

1.1. Research Objectives

1. To investigate the challenges and issues junior high school science teachers face when integrating AI into assessment.
2. To analyze the relationship between science teachers' beliefs and their willingness to adopt AI in assessment.
3. To design and implement a professional development workshop aimed at enhancing teachers' application of AI in assessment, and to examine its impact on teacher beliefs and practices.

1.2. Research Questions

1. What are the main challenges faced by junior high school science teachers in applying AI to assessment?
2. How do science teachers' beliefs influence their willingness to adopt AI in assessment?
3. How can professional development workshops enhance science teachers' understanding and use of AI in assessment?

2. Literature Review

2.1. The Relationship Between International Assessments and AI-Based Assessment

The PISA 2025 framework for scientific literacy emphasizes students' ability to investigate, evaluate, and apply scientific information in decision-making and action. This focus aligns with the timely integration of AI technologies into assessment, which is actively redefining both its methods and scope. Through automated scoring, adaptive testing, data analytics, and innovative item design, AI enhances assessment efficiency, captures student interest, and offers personalized learning feedback based on individual performance levels. Moreover, AI-driven simulations and interactive formats provide more authentic representations of students' capabilities in solving complex problems (Hao et al., 2024; Heeg & Avraamidou, 2023).

Despite these advantages, the application of AI in science education still faces challenges. Teachers' acceptance of such technologies, along with the appropriateness of their timing and context of use, can significantly influence their willingness to adopt AI-based assessments. In summary, AI introduces innovative potential to international assessments, yet its successful implementation depends on the development of supporting educational resources and infrastructure (Azari et al., 2024; Yılmaz, 2024).

2.2. Applications of AI Technology in Science Education Assessment

Artificial intelligence is considered a subfield of computer science within the broader domain of information and communication technology (ICT). Its integration into science education involves technologies such as neural networks, intelligent tutoring systems, deep learning, speech and image recognition, and natural language processing (Colchester et al., 2017; Lu et al., 2018). AI is applied in educational contexts through scenario simulations, automated feedback, adaptive learning, and chatbot-based tutoring systems (Heeg & Avraamidou, 2023; Mahligawati et al., 2023; Vierhauser et al., 2024). As such, both AI and ICT can be regarded as functional tools for teaching and assessment.

Studies on instructional practices suggest that teachers' ICT competencies and pedagogical beliefs have a measurable impact on students' scientific reasoning and data interpretation skills. While ICT proficiency alone does not directly influence academic performance, pedagogical beliefs aligned with inquiry-based instruction and the implementation of constructivist professional development activities can strengthen teachers' sense of efficacy in using AI for assessment. In other words, science teachers' engagement in targeted professional development enhances their ability to implement constructivist teaching approaches supported by AI tools in the classroom (Alt, 2018; Hu et al., 2021; Wilson, 2013).

2.3. The Relationship Between Teacher Beliefs and Professional Development Activities

Teacher professional development is a vital means of enhancing both teaching and learning. Teachers' attitudes and stances toward new instructional methods or strategies significantly influence the effectiveness of such activities. Factors such as the goals of professional development, the format of participation, collaboration among teachers, appropriate duration, and content relevance all affect whether teachers can successfully integrate what they have learned into classroom practice. Consequently, teachers' pedagogical beliefs shape both how they engage in

professional development and how they apply new knowledge in their teaching (Hubers et al., 2020; Smith & Lindsay, 2016; You et al., 2024).

The integration of AI-based assessment tools into professional development can further support teachers through instructional material analysis, classroom interaction, and reflective practice. For instance, AI tools can document teaching processes, analyze classroom interaction patterns, and generate formative assessment resources such as visuals, videos, audio, and contextual scenarios tailored to science instruction. These tools also offer specific suggestions and guidance to promote reflection and the refinement of teaching strategies (Baysal & Mutlu, 2021; Lumpe et al., 2012).

3. Research Methodology

This study adopted a one-group pretest-posttest design to examine the effects of a professional development workshop on junior high school science teachers. The workshop was designed to enhance teachers' understanding and application of AI in assessment and to provide practical guidance for implementation. The study focused on analyzing changes in teachers' beliefs and attitudes before and after the activity, with particular emphasis on how AI-based assessment influences their willingness to adopt such technologies and how it may strengthen their capacity to apply AI in science education assessment.

3.1. Participants

The participants of this study were 30 junior high school science teachers from northern Taiwan, with an equal gender distribution. In terms of age, 2 teachers (6.67%) were under 29, 13 (43%) were between 30 and 39, 14 (46%) were between 40 and 49, and 1 (3.33%) was aged 50–59. Regarding teaching experience, 5 teachers (16.6%) had less than 5 years of experience, 12 (40%) had 5 to 10 years, 10 (33.3%) had 10 to 15 years, and 3 (10%) had more than 15 years. As for school type, 18 teachers (60%) taught in urban schools and 12 (40%) in non-urban schools. In terms of school size, 5 teachers (16.6%) worked in schools with fewer than 12 classes, 14 (46.6%) in schools with 12 to 29 classes, and 11 (36.6%) in schools with more than 30 classes.

3.2. Process of Professional Development Activities

The workshop spanned a total of three hours and was structured into three sequential stages. The primary objectives were to help teachers understand the differences between the scientific literacy focus of PISA 2025 and previous assessment frameworks, enhance their confidence and comprehension of AI applications in assessment, and demonstrate practical AI-supported assessment scenarios in science education.

In the first stage (30 minutes), participants completed a pretest and were introduced to the core competencies of PISA 2025 scientific literacy, including the new dimension of Learning in the Digital World (LDW). Case examples were presented, building upon Taiwan's prior assessment experiences and data to contextualize the upcoming framework.

The second stage (90 minutes) addressed the newly added competencies in PISA 2025—specifically, investigating, evaluating, and applying scientific information for decision-making and action, along with LDW components. Teachers were prompted with the guiding question, "How can

we implement this?" to stimulate engagement. Various AI-supported assessment methods and case studies were shared, including data visualization, simulations, modeling, personalized learning, language-based learning, and virtual environments. A hands-on demonstration using Google Colab and Python guided teachers through generating scientific data visualizations (e.g., temperature or environmental data). Participants modified sample Python code to analyze real datasets (such as CO_2 emissions), then created new visual outputs and discussed their scientific implications.

In the final stage (60 minutes), teachers applied the introduced tools and strategies to design their own science teaching case. They collaborated in groups to develop and refine their ideas, uploaded their lesson plans for peer sharing, and completed a posttest to evaluate changes in understanding and attitudes following the workshop.

3.3. Measurement Tools

The pretest and posttest questionnaires used in this study contained identical items and consisted of three sections: the Teacher Epistemic Beliefs Scale (TEBS), the Teacher Pedagogical Beliefs Scale (TPBS), and the AI for Assessment Scale. A total of 28 items were included, all rated on a 5-point Likert scale, where 1 indicated "strongly disagree" and 5 indicated "strongly agree."

3.3.1. Teacher Epistemic Beliefs Scale (TEBS)

This scale comprises two dimensions: Authority and Expert Knowledge (AEK) and Certainty of Knowledge (CK). The AEK subscale, with a Cronbach's α of .80, includes six items (e.g., "I have no doubt about what experts say"), while the CK subscale, with a Cronbach's α of .61, includes three items (e.g., "If scientists work hard enough, they can find the truth to nearly all problems"). This scale, adapted from Chai et al. (2010), was used to examine science teachers' epistemic beliefs related to the use of AI in assessment.

3.3.2. Teacher Pedagogical Beliefs Scale (TPBS)

The TPBS includes two dimensions: Constructivist Pedagogical Beliefs (CPB) and Traditional Pedagogical Beliefs (TPB). Developed by Chai & Teo (2008), this scale was designed to assess teachers' instructional beliefs in the context of AI-supported assessment. The CPB subscale, with a Cronbach's α of .88, consists of five items (e.g., "Students' ideas are important and should be taken seriously"), while the TPB subscale, with a Cronbach's α of .84, also contains five items (e.g., "Learning mainly occurs through practice and repetition").

3.3.3. AI for Assessment Scale (AIAS)

This scale includes two dimensions: AI as an Information Tool (AIIT) and AI as a Learning Tool (AILT). The AIIT subscale includes five items and has a Cronbach's α of .83, while the AILT subscale includes four items with a Cronbach's α of .84. Adapted from Tondeur et al. (2007), this instrument aimed to explore whether science teachers perceive AI as primarily an information-processing tool or a tool for enhancing student learning within the context of assessment.

4. Results

The pretest and posttest scales each comprised six dimensions, with response patterns summarized in Table 1. All subscales demonstrated acceptable internal consistency, with Cronbach's α values exceeding 0.7. Notably, both AIIT (AI as an Information Tool) and AILT (AI

as a Learning Tool) received mean scores above 4 in both pretest and posttest. Post-workshop feedback further revealed a high level of teacher willingness to apply AI in science assessment, with a mean score of 4.1 (SD = 0.57).

Table 1. Response Patterns Across Six Dimensions in Pretest and Posttest Questionnaires.

Pretest	Mean	SD	Skewness	Kurtosis	No. of items
Teacher Epistemic Beliefs Scale (TEBS)					
AEK	2.43	.52	.49	.74	6
CK	2.81	.80	-.12	-.96	3
Teacher Pedagogical Beliefs Scale (TPBS)					
CPB	4.23	.57	-.52	-.33	5
TPB	3.39	.64	-.01	.59	5
AI for Assessment Scale (AIAS)					
AIIT	4.06	.54	-.34	.07	5
AILT	4.01	.56	-.05	-.20	4
Posttest	**Mean**	**SD**	**Skewness**	**Kurtosis**	**No. of items**
Teacher Epistemic Beliefs Scale (TEBS)					
AEK	2.28	.51	.72	2.04	6
CK	2.7	.83	.04	-.97	3
Teacher Pedagogical Beliefs Scale (TPBS)					
CPB	4.34	.49	-.45	.26	5
TPB	3.41	.78	-.01	.49	5
AI for Assessment Scale (AIAS)					
AIIT	4.18	.73	-1.75	4.68	5
AILT	4.36	.62	-.61	-.68	4

Table 2. Paired Sample T-Test Results.

Category	Mean (SD)		df	t	p
	Pretest	Posttest			
Teacher Epistemic Beliefs Scale (TEBS)					
AEK	2.43 (.52)	2.28 (.51)	29	3.07	.005*
CK	2.81 (.80)	2.7 (.83)	29	1.22	.231
Teacher Pedagogical Beliefs Scale (TPBS)					
CPB	4.23 (.57)	4.34 (.49)	29	-1.43	.165
TPB	3.39 (.64)	3.41 (.78)	29	-.21	.835

Table 2 (continued).

Category	Mean (SD)		df	t	p
	Pretest	Posttest			
AI for Assessment Scale (AIAS)					
AIIT	4.06 (.54)	4.18 (.73)	29	-.90	.375
AILT	4.01 (.56)	4.36 (.62)	29	-3.99	.000*

According to the paired sample t-test results (see Table 2), several key differences were observed between the pretest and posttest scores. The AEK (Authority and Expert Knowledge) dimension showed a significant decline in the posttest, indicating a reduced level of trust in expert knowledge after participating in the AI-focused professional development. In contrast, the AILT (AI as a Learning Tool) dimension showed a significant increase, suggesting that the workshop effectively strengthened teachers' beliefs in AI as a tool for supporting student learning.

No statistically significant differences were found in the CK (Certainty of Knowledge), CPB (Constructivist Pedagogical Beliefs), TPB (Traditional Pedagogical Beliefs), or AIIT (AI as an Information Tool) dimensions. This suggests that teachers' views in these areas remained relatively stable, possibly due to the influence of long-standing beliefs or the limited duration of the intervention.

5. Conclusion

This study set out to explore the challenges and concerns faced by junior high school science teachers in integrating AI into assessment practices, to analyze how their pedagogical beliefs influence their willingness to adopt AI-based assessment tools, and to enhance their application competencies through a professional development program framed around the evolving focus of PISA 2025 scientific literacy. The findings confirm that well-structured professional development can meaningfully influence teachers' beliefs and attitudes (De Vries et al., 2014; Sari et al., 2020; Vries et al., 2013), with particularly notable gains in their recognition of AI as a learning tool (AILT). Moreover, although the shift in perceptions of AI as an information tool (AIIT) was not statistically significant, there was observable growth, underscoring the potential of such interventions to expand teachers' understanding and practical use of AI in science assessment.

However, a significant decline in trust toward authority and expert knowledge (AEK) was observed post-intervention, indicating ongoing skepticism about the accuracy and reliability of AI technologies (Nazaretsky et al., 2022; Nazaretsky et al., 2021). Minor, non-significant increases in constructivist (CPB) and traditional pedagogical beliefs (TPB) suggest that deeply rooted teaching beliefs may require longer-term engagement to shift meaningfully (Meschede et al., 2017; Valcke, 2010). Overall, this study affirms the transformative potential of AI in science assessment and highlights the critical role of teacher professional development in supporting its effective adoption. At the same time, it brings to light persistent challenges—such as limited familiarity with AI tools, high preparation demands, and misalignment with existing curricular goals—that must be addressed to ensure sustainable integration of AI technologies in classroom assessment practice.

Acknowledgements

This work was supported by the National Science and Technology Council (NSTC) in Taiwan [grant number NSTC 113-2410-H-152 -041 -MY2].

References

Alt, D. (2018). Science teachers' conceptions of teaching and learning, ICT efficacy, ICT professional development and ICT practices enacted in their classrooms. *Teaching and Teacher Education*.

Azari, A., Rochintaniawati, D., & Agustin, R. (2024). Can Web-Based AI Be Implemented in the Middle School Science Classroom? A Critical Review. *Jurnal Inovasi Pendidikan IPA*. https://doi.org/10.21831/jipi.v10i2.75364.

Baysal, Y., & Mutlu, F. (2021). The Effect of Professional Development Programs on Teachers' Science Teaching Self Efficacy Beliefs: A Meta-Analysis Study. *International Journal of Contemporary Educational Research*.

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468.

Chai, C. S., & Teo, T. (2008). Confirmatory Factor Analysis of the Conception for Teaching and Learning Questionnaire (CTLQ). *The Asia-Pacific Education Researcher, 17*(2), 1-1.

Chai, C. S., Deng, F., Wong, B., & Qian, Y. (2010). South China education majors' epistemological beliefs and their conceptions of the nature of science. *The Asia-Pacific Education Researcher*.

Choi, S., Jang, Y., & Kim, H. (2022). Influence of Pedagogical Beliefs and Perceived Trust on Teachers' Acceptance of Educational Artificial Intelligence Tools. *International Journal of Human–Computer Interaction*, 39, 910 - 922.

Colchester, K., Hagras, H., Alghazzawi, D., & Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *Journal of Artificial Intelligence and Soft Computing Research*, *7*(1), 47-64.

De Vries, S., Van De Grift, W., & Jansen, E. (2014). How teachers' beliefs about learning and teaching relate to their continuing professional development. *Teachers and Teaching*, 20, 338 - 357.

Deta, U., Ayun, S., Laila, L., Prahani, B., & Suprapto, N. (2024). PISA science framework 2018 vs 2025 and its impact in physics education: Literature review. *Momentum: Physics Education Journal*.

Dogan, M., Dogan, T., & Bozkurt, A. (2023). The Use of Artificial Intelligence (AI) in Online Learning and Distance Education Processes: A Systematic Review of Empirical Studies. *Applied Sciences*.

Ejjami, R. (2024). The Future of Learning: AI-Based Curriculum Development. *International Journal For Multidisciplinary Research*. https://doi.org/10.36948/ijfmr.2024.v06i04.24441.

Gurer, M., & Akkaya, R. (2021). The influence of pedagogical beliefs on technology acceptance: a structural equation modeling study of pre-service mathematics teachers. *Journal of Mathematics Teacher Education*, 25, 479 - 495.

Hao, M., Wang, Y., & Peng, J. (2024). Empirical Research on AI Technology-Supported Precision Teaching in High School Science Subjects. *Applied Sciences*. https://doi.org/10.3390/app14177544.

Heeg, D., & Avraamidou, L. (2023). The use of Artificial intelligence in school science: a systematic literature review. *Educational Media International*, 60, 125 - 150. https://doi.org/10.1080/09523987.2023.2264990.

Herold, D., Murphy, G., & Buck, C. (2024). *Preparing the next generation: Integrating generative AI in entrepreneurship education for graduate capability development.*

Holmes, W. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. *Center for Curriculum Redesign*.

Hu, D., Yuan, B., Luo, J., & Wang, M. (2021). A review of empirical research on ICT applications in teacher professional development and teaching practice. *Knowledge Management & E-Learning: An International Journal*.

Hubers, M., D.Endedijk, M., & Van Veen, K. (2020). Effective characteristics of professional development programs for science and technology education. *Professional Development in Education*, 48, 827 - 846.

Kim, J. (2021). 'In numbers we trust': Statistical data as governing technologies in the era of student achievement and school accountability. *Educational Philosophy and Theory*, 54, 1442 - 1452.

Kruit, P., Oostdam, R., van den Berg, E., & Schuitema, J. (2020). Performance assessment as a diagnostic tool for science teachers. *Research in Science Education*, *50*(3), 1093-1117.https://doi.org/10.1007/s11165-018-9724-9.

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain intelligence: go beyond artificial intelligence. *Mobile Networks and Applications*, *23*, 368-375.

Lumpe, A., Czerniak, C., Haney, J., & Beltyukova, S. (2012). Beliefs about Teaching Science: The relationship between elementary teachers' participation in professional development and student achievement. *International Journal of Science Education*, 34, 153 - 166. https://doi.org/10.1080/09500693.2010.551222.

Mahligawati, F., Allanas, E., Butarbutar, M., & Nordin, N. (2023). Artificial intelligence in Physics Education: a comprehensive literature review. *Journal of Physics: Conference Series*, 2596.

Meschede, N., Fiebranz, A., Möller, K., & Steffensky, M. (2017). Teachers' professional vision, pedagogical content knowledge and beliefs: On its relation and differences between pre-service and in-service teachers. *Teaching and Teacher Education*, 66, 158-170. https://doi.org/10.1016/J.TATE.2017.04.010.

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British journal of educational technology*, *53*(4), 914-931.

Nazaretsky, T., Cukurova, M., & Alexandron, G. (2021). An Instrument for Measuring Teachers' Trust in AI-Based Educational Technology. LAK22: 12th International Learning Analytics and Knowledge Conference.

Sari, S., Nurkamto, J., & Rochsantiningsih, D. (2020). Teacher Professional Development: The Story of English Teacher's Beliefs and Practices. *ELS Journal on Interdisciplinary Studies in Humanities*.

Smith, K., & Lindsay, S. (2016). Building Future Directions for Teacher Learning in Science Education. *Research in Science Education*, 46, 243-261. https://doi.org/10.1007/S11165-015-9510-X.

Tondeur, J., Van Braak, J., & Valcke, M. (2007). Towards a typology of computer use in primary education. *Journal of computer assisted learning, 23*(3), 197-206.

Valcke, M., Sang, G., Rots, I., & Hermans, R. (2010). Taking Prospective Teachers' Beliefs into Account in Teacher Education. , 622-628. https://doi.org/10.1016/B978-0-08-044894-7.00668-0.

Vierhauser, M., Groher, I., Antensteiner, T., & Sauerwein, C. (2024). Towards Integrating Emerging AI Applications in SE Education. *2024 36th International Conference on Software Engineering Education and Training (CSEE&T)*, 1-5. https://doi.org/10.1109/CSEET62301.2024.10663045.

Vries, S., Grift, W., & Jansen, E. (2013). Teachers' beliefs and continuing professional development. *Journal of Educational Administration*, 51, 213-231. https://doi.org/10.1108/09578231311304715.

Wang, S., Li, J., & Yuan, Y. (2024). The power of convergence: STEM education in the era of artificial intelligence. In *Disciplinary and Interdisciplinary Education in STEM: Changes and Innovations* (pp. 63-80). Cham: Springer Nature Switzerland.

Wang, X., Li, L., Tan, S. C., Yang, L., & Lei, J. (2023). Preparing for AI-enhanced education: Conceptualizing and empirically examining teachers' AI readiness. *Computers in Human Behavior*, *146*, 107798.

White, P. J., Ardoin, N. M., Eames, C., & Monroe, M. C. (2023). Agency in the Anthropocene: Supporting document to the PISA 2025 Science Framework.

Williamson, B. (2020). New digital laboratories of experimental knowledge production: Artificial intelligence and education research. *London Review of Education*, *18*(2), 209-220.

Xu, Z. (2023). PISA's Impact on China's education policy. *Journal of Education, Humanities and Social Sciences*.

Yılmaz, Ö. (2024). Personalised learning and artificial intelligence in science education: current state and future perspectives. *Educational Technology Quarterly*. https://doi.org/10.55056/etq.744.

You, H., Park, S., Hong, M., & Warren, A. (2024). Unveiling effectiveness: A meta-analysis of professional development programs in science education. *Journal of Research in Science Teaching*.

學習信念、機器人程式設計態度對於運算思維、創造力與合作學習的影響研究

A Study on the Impact of Learning Beliefs and Attitudes Toward Robotics Programming on Computational Thinking, Creativity, and Collaborative Learning Ability

歐陽芳泉[1]，周縈綺[2*]
[1]臺北商業大學數位多媒體設計系
[2]臺北商業大學創意設計與經營研究所
[*] cyc8730@ntub.edu.tw

【摘要】 隨著程式設計與運算思維（Computational Thinking, CT）在教育領域的重要性提升，機器人程式設計已成為培養學生邏輯思維與問題解決能力的關鍵。然而，學生學習動機與信念可能影響其學習態度與成效。本研究基於期望價值理論（EVT），探討學習信念（效用價值、興趣價值、成就價值與自我效能）對教育機器人程式設計態度的影響，並分析二者如何影響學生的運算思維、創造力與合作學習。本研究採用問卷調查法及實作評估，透過 spss21.0 以進行迴歸分析。結果顯示，學習信念正向影響機器人程式設計態度，進而提升學生的運算思維及合作學習。機器人程式設計態度正向影響運算思維、創造力及合作學習。
【關鍵字】 教育機器人；運算思維；期望與價值理論

Abstract: *With the increasing importance of programming and computational thinking (CT) in education, robotics programming has become a key tool in cultivating students' logical thinking and problem-solving abilities. However, students' learning motivation and beliefs may influence their learning attitudes and outcomes. This study, based on the Expectancy-Value Theory (EVT), explores how learning beliefs (utility value, interest value, achievement value, and self-efficacy) affect attitudes toward robotics programming, and further analyzes how these attitudes influence students' computational thinking, creativity, and collaborative learning abilities. The study employs a questionnaire survey and practical assessment, with regression analysis conducted using SPSS 21.0. The results indicate that learning beliefs positively influence attitudes toward robotics programming, which in turn enhances students' computational thinking and collaborative learning abilities. Attitudes toward robotics programming positively influence computational thinking, creativity, and collaborative learning abilities.*
Keywords: Educational robotics, Computational Thinking, Expectancy-Value Theory

1. 研究背景與動機

隨著數位科技與人工智慧快速發展，運算思維（Computational Thinking, CT）已被視為 21 世紀核心素養之一，廣泛應用於程式設計、資料分析與跨領域問題解決（Wing, 2017）。CT 涵蓋問題分解、模式辨識與演算法三大核心能力，是數位素養與邏輯推理的基礎（Brennan & Resnick, 2012；Grover & Pea, 2018）。因此，許多國家已將 CT 納入中小學與高等教育課程，並透過教育機器人課程協助學生以實作方式應用程式語言於具體情境中，深化對運算思維概念的理解（Tengler et al., 2022；Alonso-García et al., 2024）。

然而，對於學生而言，抽象的程式語言常造成學習上的困難與挫折，進而影響其學習動機與學習成效。為此，教育現場導入教育機器人程式設計課程，透過具體、可操作的學習媒介，降低學生學習門檻，並提升其學習興趣與實作能力（Angel-Fernandez & Vincze, 2018）。此外，教育機器人課程也常結合專題式學習（Project-Based Learning, PBL）與合作學習的策略，讓學生在解決實務問題的過程中，發展運算思維、創造力與團隊合作能力（Grobelna & Plawiak-Mowna, 2024）。

在探討學生學習表現的相關研究中，學習信念是關鍵的影響因素。根據期望價值理論（Expectancy-Value Theory, EVT），學生的學習行為會受到自身對學習成就的期望與學習價值認知所影響（Wigfield & Eccles, 2000）。其中，效用價值、興趣價值、成就價值與自我效能構成學生的學習信念，這些因素直接影響其學習表現（Bayanova et al., 2023；Eccles & Wigfield, 2024）。過去研究指出，若學生對學習任務抱持正面信念，將有助於其投入學習表現（Eccles & Wigfield, 2024）。但目前在教育機器人課程中，針對學生學習信念、態度與學習成效間關係的研究仍屬有限，尤其缺乏以 CT、創造力與合作學習能力為指標進行系統性探討的實證研究。因此，本研究欲針對此研究之缺口，深入探討學習信念與機器人程式設計態度及學習表現之間的關聯性。具體研究目的如下：

1、探討學生對學習的信念（效用價值、興趣價值、成就價值與自我效能）是否影響其對機器人程式設計的態度。

2、分析學習信念與機器人程式設計態度對學生運算思維能力之影響。

3、探討學習信念與機器人程式設計態度是否能提升創意思考能力與合作學習。

2. 文獻探討

2.1 教育機器人

教育機器人 (Educational Robotics, ER) 已廣泛應用於教育現場，能提升學生的學習態度、運算思維、創造力與合作學習能力 (Angel-Fernandez & Vincze, 2018)。學生操作機器人與程式，可培養實作能力與學習動機。研究指出，學生進行機器人程式設計時，學習信念 (如課程實用性或趣味性) 會影響其學習態度 (Román-Graván et al., 2020)。當學生理解課程與生活的關聯時，他們的學習動機與運算思維能力更有可能得到提升。

此外，機器人課程多透過 PBL 與協作，學生需整合創意與邏輯解決問題，並合作完成任務，能培養創造力與合作能力 (Rapti & Sapounidis, 2024)。Noh 與 Lee (2020) 指出，機器人有助提升學生的創意問題解決能力與創新應變力。同時，合作學習是機器人程式設計中的關鍵要素。Demetroulis 等 (2023) 發現，小組設計任務能提升社交互動與合作技巧。綜上所述，ER 透過整合程式設計與跨域合作學習，能顯著提升學生的學習態度、運算思維、創造力、合作學習、實作能力與學習動機。

2.2 運算思維

運算思維 (Computational Thinking, CT) 是一種系統性解決問題的方式，包含分解 (Decomposition)、模式辨識（(Pattern Recognition）與演算法（Algorithm) 等核心能力 (Brennan & Resnick, 2012)，是現代學生必備的基礎素養。 教育中，培養學生運算思維是課程設計重點。研究發現，學生的學習信念 (如效用價值與自我效能) 會影響其 CT 任務表現 (Lee et al., 2024)。當學生相信程式設計與未來相關，且有信心完成任務時，學習動機與表現提升。 Lee 等 (2024) 的研究指出，學生的正向學習信念能夠預測他們的 CT 能力，自我效能感則是其中的中介因素。Alonso-García 等 (2024) 透過機器人教學發現，學生可將抽象概念

應用於操作，強化分析、組織與解決問題的能力。CT 為程式設計之核心要素，學習態度與信念亦為影響其發展之關鍵。

2.3 期望價值理論

期望價值理論 (Expectancy-Value Theory, EVT) 由 Eccles 等 (1983) 提出，是解釋學習動機的重要理論。該理論認為，個體對成功的期望 (Expectancy) 與學習價值 (Value) 會共同影響學習行為，學習價值進一步分為效用價值（Utility Value）、興趣價值（Interest Value）、成就價值（Attainment Value）與成本（Cost） (Wigfield & Eccles, 2000)。本研究探討學生的學習信念如何影響其使用機器人程式設計態度，進而影響運算思維、創造力與合作學習，這與 EVT 的核心概念高度相關。

近年來，EVT 在教育應用受關注，學生的學習信念與其學習表現之間存在顯著關聯 Bayanova 等 (2023) 指出，效用價值與興趣價值是提升 STEM 學習動機的關鍵。Torres 與 Inga (2025) 發現，結合程式設計與機器人有助提升學生興趣，增進學習參與。此外，EVT 可解釋合作學習動機。Avcı (2024) 指出，合作策略與機器人課程能提升學生投入與表現，顯示期望與價值影響團隊互動。EVT 提供學習動機與行為的理論架構，也是本研究分析學生信念、態度與學習成效關係的基礎。

3. 研究方法

3.1 研究架構

本研究探討學習信念、機器人程式設計態度、運算思維能力、創造力及合作學習之間的關聯，透過迴歸分析驗證各變數的影響關係，並據此提出教育應用上的建議。本研究模型如圖 1 所示。

本研究提出以下假設：

H1：學習信念對機器人程式設計態度具有正向影響。

H2：學習信念對運算思維能力（分解、模式辨別、演算法）具有正向影響。

H3：學習信念對創造力具有正向影響。

H4：學習信念對合作學習具有正向影響。

H5：機器人程式設計態度對運算思維能力（分解、模式辨別、演算法）具有正向影響。

H6：機器人程式設計態度對創造力具有正向影響。

H7：機器人程式設計態度對合作學習具有正向影響。

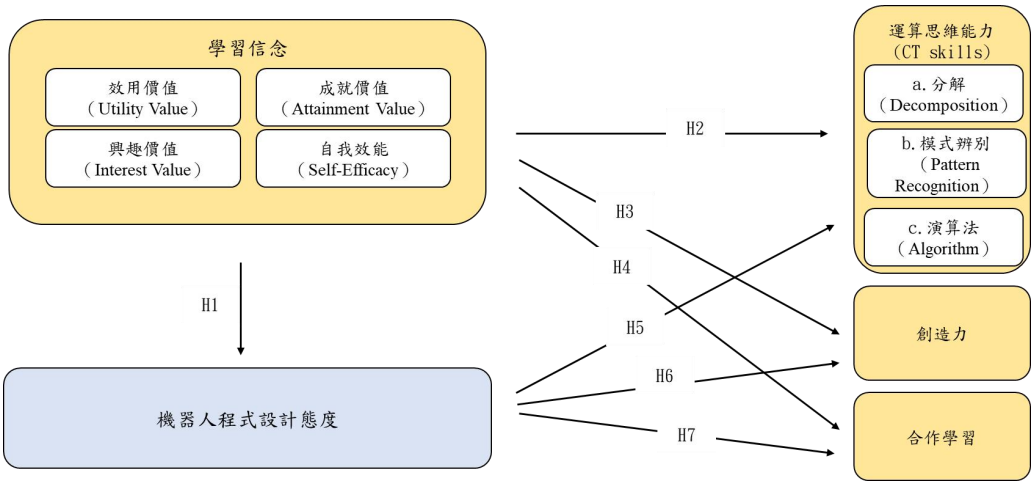


圖 1 本研究模型

3.2 教育機器人

本研究所使用之教育機器人教具為本研究團隊自行開發之 Mango X2 Starter 教育機器人，其使用樹莓派 Raspberry Pi Pico 為核心，可以支持 Python 程式的學習和 STEAM 教學使用，透過實作活動培養學生的電子電路與機電整合基本能力。Mango X2 Starter 的硬體設計以 Mango X2 開發板為核心，整合多種感測器與執行元件，如超音波感測器、紅外線感測器、光感測器、伺服馬達、直流馬達與 LED 燈等，能夠實現避障、循跡、環境感知與互動控制等功能。圖 2 為「Mango X2」開發板與「Mango X2 Starter」教育機器人。

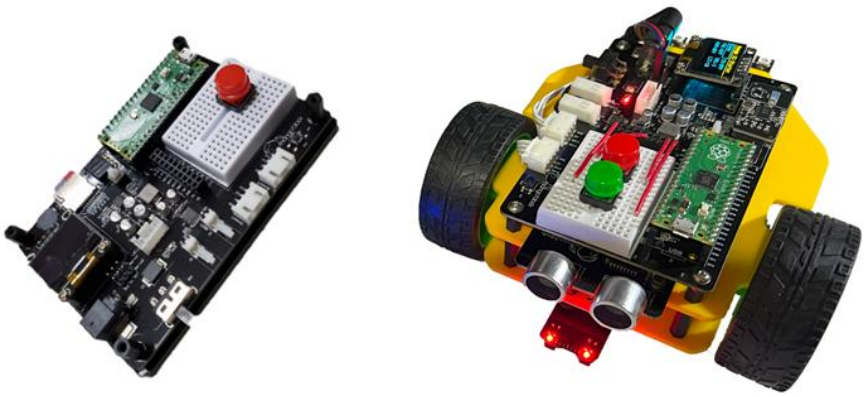


圖 2 本研究團隊研發之「Mango X2」開發板與「Mango X2 Starter」教育機器人

3.3 課程任務設計

為了促進學生們以合作學習模式，本課程採取小組合作的方式進行教學，每兩位學生分為一組，共同討論與解決課程中的任務。課程應用 Mango X2 Starter 作為學習工具，並使用 Thonny IDE 進行程式設計與編寫，使學生能夠透過實作來加深對程式設計的理解。

本研究設計了一系列專題任務，以促進學生運算思維的發展，並提升其邏輯推理與程式設計能力。課程任務涵蓋三個主要學習活動：(1) 問題分解：學生透過心智圖將複雜問題拆解為可管理的小問題，以增強分析與組織能力；(2) 模式辨識：學生使用流程圖建構程式邏輯，學習如何識別問題中的規律與模式；(3) 演算法設計與實作：學生透過 Python 程式設計，應用迴圈與條件判斷來解決問題。課程採用小組合作方式，鼓勵學生透過討論與交流，共同完成專題，並在學習過程中培養團隊合作與溝通能力。

4. 實驗設計

4.1 實驗流程

本研究的實驗對象為臺灣北部某大學一年級的學生，共 96 位，平均年齡為 19 歲。本實驗設計安排為 9 週進行，在第 1 週會對所有參與者進行課程說明，並同意參加課程實驗並填寫參與教學研究知情同意書，並進行 2 人為一組來分組學習；於第 2 週開始進行 MangoX2 Starter 教育機器人導入程式設計課程活動。在課程中，在第 5、9 週分派第 1 次和第 2 次課程任務，每位學生皆需要求完成任務（心智圖、流程圖繪製、程式設計編寫）。在完成為期 9 週的學習活動後，則對學生進行問卷（學習信念、機器人程式設計態度、創造力、合作學習）填寫，以瞭解各學生在經過學習後，對各應變項的影響。

4.2 測量工具

本研究調查問卷內容為「學習信念」、「機器人程式設計態度」「合作學習」及「創造力」四大構面。本研究採用 Likert 五點量表（1 表示非常不同意，5 表示非常同意），測量各構面之變項。在學習信念構面，基於期望價值理論（Expectancy-Value Model）（Eccles & Wigfield, 1995），認為期望（自我效能感）和價值（喜好、實用性、重要性和成本）是學習策略使用的關鍵因素。故本研究使用 Berger & Karabenick（2011）研究問卷中的效用價值（Utility Value）、興趣價值（Interest Value）、成就價值（Attainment Value）構面之問項，每個構面各自有 3 題問項，以及 Bhattacherjee & Premkumar（2004）自我效能（Self-Efficacy）4 題問項，共計 13 題作為學習信念構面之問項。在機器人程式設計態度（Robot programming attitude）構面，使用 Bhattacherjee & Premkumar（2004）研究問卷中的 3 題問項。在合作學習與創造力構面，使用 Korkmaz, et al. (2017) 研究問卷中的團隊合作與創造力構面之問項，各自有 4 題問項與 6 題問項。以上問卷均已經過研究之驗證，故具有良好的信度與效度。

在課程任務的學習成效部分，本研究基於運算思維核心要素的階段性評量尺規（Rubrics）。評量內容將專題實作細分為心智圖設計、流程圖繪製，以及程式設計等項目，確保評估過程具備系統性與一致性。如表 1 所示，透過評量尺規，學生的學習成效能以量化方式呈現，研究人員依據完成度和作品品質進行統計分析，以比較學生之間的學習差異。任務 Rubrics 有二位專家進行評分，並且在評分前先制定評分標準，於評分後進行差分檢核，透過這些措施，確保評量過程的公正性與準確性，使研究結果能有效反映學生的學習成效與任務表現。

表 1 任務評量設計

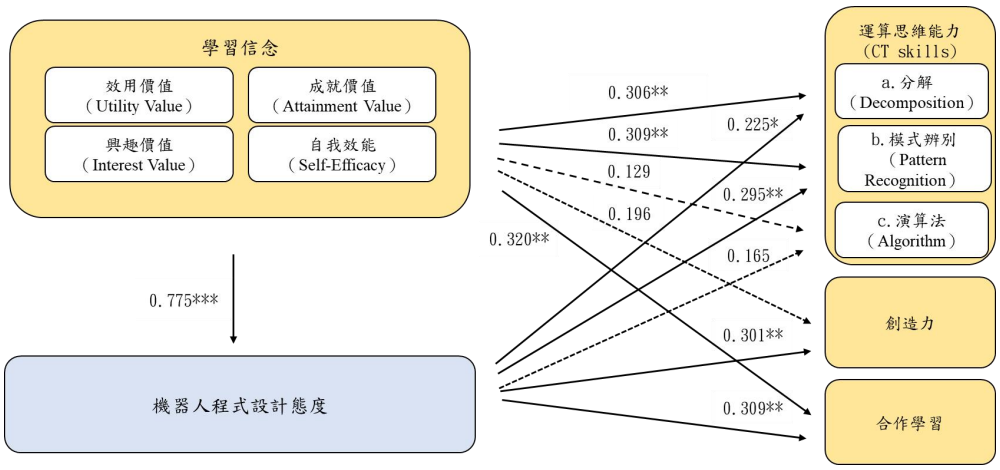
CT Skills	任務設計	觀察評估重點
分解	規劃機器人專題，畫出心智圖	1、心智圖包含的系統程式功能是否完整 2、問題的拆解是否適合
模式辨別	規劃機器人專題，畫出流程圖，完成程式的實作	1、流程圖的內容、順序正確性 2、流程圖與程式內容相符
演算法	規劃機器人專題，對於重複執行的功能進行 function 設計規劃。	1、程式功能錯誤 2、對於 function 的設計是否有達到可重複使用的功能

5.結果分析與討論

本研究依照回收的有效問卷(78 份)分析出學習信念與機器人程式設計態度、合作學習與創造力的影響。本問卷使用 Cronbach's α值來衡量問卷項目之間的一致性，並檢驗各個構面量表的內部一致性。根據 Guidford（1965）的說法，愈高的可信度的問項，其 α 係數大於 0.7，即表示可信度及穩定度愈高，介於 0.4~0.6 為中信度，反之 α 係數<0.35 則表示可信度愈低。本研究各題項的 Cronbach's α值皆大於 0.7，介於 0.796 與 0.924 之間，顯示大部分變數皆具有高信度值，也進而顯示大部分的變數具有高度的內部一致性，本研究問卷具有信度。

迴歸分析主要運用在一個或一個以上變數間的關聯性，找出依變項與自變項間的依賴狀況。再利用迴歸分析所有的標準迴歸係數，來判斷是否具有存在因果關係的推論。本研究透

過迴歸分析來檢驗學習信念與機器人程式設計的態度對運算思維技能、創造力與合作學習之影響，分析結果如圖 3 所示。



*：p<0.05，**：p<0.01，***：p<0.001

圖 3 本研究之分析結果

研究結果顯示，學習信念對機器人程式設計態度具有顯著正向影響，說明當學生認為學習有價值（如效用價值、興趣價值、成就價值）且具備較高的自我效能時，他們更積極投入機器人程式設計學習。此外，學習信念對問題分解能力與模式辨別能力具有顯著影響，表示學生如果對學習抱持正向態度，他們在拆解問題與發現規則更為投入並具有較好的成效（Lee et al. 2024)。然而，學習信念對演算法能力未達顯著影響（Grover & Pea 2018)，可能是因為演算法與撰寫程式實作能力相關，學生仍需要透過更多學習任務練習、實際操作才能有顯著的提升成效。

在機器人程式設計態度對於運算思維能力的影響部份，結果顯示學生對機器人程式設計的態度對於問題分解能力、模式辨別能力呈現顯著正向關聯，表示當學生對機器人程式設計抱持正向態度時，他們更能有效拆解問題並辨識程式規則模式（Noh & Lee ,2020）。然而，機器人程式設計態度對演算法能力的影響未達顯著水準，可能的原因亦同上所述，學生對於程式設計的技能提升可能需要更長期、更多的邏輯推理與程式技能的訓練。

另一方面，本研究亦發現學習信念對合作學習具有顯著影響，表示當學生相信學習有價值，且對自己的學習能力有信心時，他們更願意與同儕合作來解決問題（Avcı ,2024）。這顯示提升學生的學習信念有助於促進合作學習。此外，機器人程式設計態度對創造力與合作學習均有顯著影響，表示當學生對機器人程式設計抱持正向態度時，他們更能發揮創意並願意與同儕合作學習（Rapti & Sapounidis,2024）。這可能在機器人程式設計中，促使學生可以自由地探索和嘗試不同的方法來解決問題，並與同儕交流想法。

6.結論與建議

本研究探討學習信念、機器人程式設計態度，以及運算思維能力、創造力與合作學習之間的關聯。研究結果顯示，學生的學習信念對其機器人程式設計態度，以及運算思維能力中的分解與模式辨識能力，具有顯著的正向影響，此結果與期望價值理論相符，強調了學習信念在學習動機上的重要性。然而，學習信念對於演算法能力的影響相對不明顯，顯示演算法

學習需要透過更多的教學輔助與訓練。此外，學習信念對合作學習有顯著影響，指出學習態度能促進合作，此亦呼應先前研究指出合作學習能提升學習動機的研究。

基於上述結果，本研究提出以下建議：為提升運算思維學習成效，教師應致力於提升學生的學習信念，以增強其學習態度。教師可透過專題為導向的實務案例，使學生認知其效用價值，並設計挑戰性活動，建立自我效能感與成就價值，提升學習投入。同時，在演算法教學方面，除圖形化學習外，亦應融入文字程式範例，幫助學生理解演算法邏輯；提供適當的教學資源以強化學習者的程式/演算法技能。此外，教師可設計小組專題與競賽，鼓勵學生合作激發創意，以提升學生學習之成效。

參考文獻

Alonso-García, S., Fuentes, A. V. R., Navas-Parejo, M. R., & Victoria-Maldonado, J. J. (2024). Enhancing computational thinking in early childhood education with educational robotics: A meta-analysis. Heliyon, 10(13). https://doi.org/10.1016/j.heliyon.2024.e33249

Angel-Fernandez, J. M., & Vincze, M. (2018). Towards a formal definition of educational robotics. In P. Zech & J. Piater (Eds.), Proceedings of the Austrian Robotics Workshop 2018 (Conference series). Innsbruck University Press. https://doi.org/10.15203/3187-22-1

Avcı, F. (2024). Effects of robotic module-supported experimental activities with the cooperative learning method on student achievement, 21st-century skills, and students' opinions. Journal of Computer Assisted Learning, 40(5), 2325–2338. https://doi.org/10.1111/jcal.13030

Bayanova, A. R., Orekhovskaya, N. A., Sokolova, N. L., Shaleeva, E. F., Knyazeva, S. A., & Budkevich, R. L. (2023). Exploring the role of motivation in STEM education: A systematic review. EURASIA Journal of Mathematics, Science and Technology Education, 19(4), em2250. https://doi.org/10.29333/ejmste/13086

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. Learning and Instruction, 21(3), 416–428. https://doi.org/10.1016/j.learninstruc.2010.06.002

Bhattacherjee, A., & Premkumar, G. (2004). Understanding changes in belief and attitude toward technology usage: A theoretical model and longitudinal test. MIS Quarterly, 28(2), 229–254. https://doi.org/10.2307/25148634

Brennan, K., & Resnick, M. (2012, April). New frameworks for studying and assessing the development of computational thinking. In Proceedings of the 2012 Annual Meeting of the American Educational Research Association (p. 25). Vancouver, Canada.https://scratched.gse.harvard.edu/ct/files/AERA2012.pdf

Demetroulis, E. A., Theodoropoulos, A., Wallace, M., Poulopoulos, V., & Antoniou, A. (2023). Collaboration skills in educational robotics: A methodological approach—Results from two case studies in primary schools. Education Sciences, 13(5), 468. https://doi.org/10.3390/educsci13050468

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motives (pp. 75–146). Freeman.

https://osit.nv.gov/uploadedFiles/ositnvgov/Content/Meetings/STEM/2016/Expectancies%20Values%20and%20Academic%20Behaviors.pdf

Eccles, J. S., & Wigfield, A. (2020). From expectancy–value theory to situated expectancy–value theory: A developmental, social cognitive, and sociocultural perspective. Contemporary Educational Psychology, 61, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eccles, J. S., & Wigfield, A. (2024). The development, testing, and refinement of Eccles, Wigfield, and colleagues' situated expectancy-value model of achievement performance and choice. Educational Psychology Review, 36(2), Article 51. https://doi.org/10.1007/s10648-024-09888-9

Grobelna, I., & Pławiak-Mowna, A. (2024). Motivation and engagement of students: A case study of automatics and robotics projects. Electronics, 13(20), 3997. https://doi.org/10.3390/electronics13203997

Grover, S., & Pea, R. (2018). Computational thinking: A competency whose time has come. In S. Sentance, E. Barendsen, & C. Schulte (Eds.), Computer science education: Perspectives on teaching and learning in school (Vol. 19, Issue 1, pp. 19–38). Springer. https://doi.org/10.5040/9781350057142.ch-003

Korkmaz, Ö., Çakır, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). Computers in Human Behavior, 72, 558–569. https://doi.org/10.1016/j.chb.2017.01.005

Lee, S. W. Y., Liang, J. C., Hsu, C. Y., & Tsai, M. J. (2024). Students' beliefs about computer programming predict their computational thinking and computer programming self-efficacy. Interactive Learning Environments, 32(8), 4088–4108. https://doi.org/10.1080/10494820.2023.2194929

Noh, J., & Lee, J. (2020). Effects of robotics programming on the computational thinking and creativity of elementary school students. Educational Technology Research and Development, 68(1), 463–484. https://doi.org/10.1007/s11423-019-09708-w

Qu, J. R., & Fok, P. K. (2022). Cultivating students' computational thinking through student–robot interactions in robotics education. International Journal of Technology and Design Education, 32(4), 1983–2002. https://doi.org/10.1007/s10798-021-09677-3

Rapti, S., & Sapounidis, T. (2024). "Critical thinking, Communication, Collaboration, Creativity in kindergarten with Educational Robotics": A scoping review (2012–2023). Computers & Education, 210, 104968. https://doi.org/10.1016/j.compedu.2023.104968

Román-Graván, P., Hervás-Gómez, C., & Ballesteros-Regaña, C. (2020). Educational robotics and computational thinking: A didactic experience of innovation at the university level. In M. León-Urrutia, E. Vázquez Cano, N. Fair, & E. López Meneses (Eds.), Trends and good practices in research and teaching: A Spanish-English collaboration (pp. 157–172). Octaedro. https://doi.org/10.36006/16184-09

Torres, I., & Inga, E. (2025). Fostering STEM skills through programming and robotics for motivation and cognitive development in secondary education. Information, 16(2), 96. https://doi.org/10.3390/info16020096

Tengler, K., Kastner-Hauler, O., Sabitzer, B., & Lavicza, Z. (2022). The effect of robotics-based storytelling activities on primary school students' computational thinking. Education Sciences, 12(1), Article 10. https://doi.org/10.3390/educsci12010010

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. Contemporary Educational Psychology, 25(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wing, J. M. (2017). Computational thinking's influence on research and education for all. Italian Journal of Educational Technology, 25(2), 7–14. https://doi.org/10.17471/2499-4324/922

STEM 積木建構活動之學生策略運用類型之文獻分析

A Literature Analysis of Student Strategy Types in STEM Block Construction Activities

吳聲毅 [1*]
[1] 國立清華大學師資培育中心
[*] digschool@gmail.com

【摘要】 STEM 教育強調跨培養學生的科學探究能力與工程設計思維。其中，積木活動作為工程教育的重要工具，透過積木搭建過程，學生得以實踐結構設計與問題解決。然而，現有研究較少深入探討國小學生在積木建構活動中實際採取的策略。為了協助教師設計更符合學生需求的 STEM 活動與學習評量工具，本研究以文獻分析法，歸納並整理國小學生積木建構活動中的策略類型。研究結果歸納出常見的七大類策略，包括：探索性策略、規劃性策略、模仿與適應策略、分解與組合策略、反思與修正策略、合作與分工策略，以及創意性策略。此可作為未來教師設計 STEM 積木活動的教學引導依據，並提供學生學習歷程紀錄的參考。
【關鍵字】 STEM 教育；積木；工程設計思維；問題解決策略

Abstract: *Building block activities are widely used in STEM education to cultivate students' engineering design thinking and problem-solving abilities. Through hands-on construction processes, students explore structural stability, design concepts, and creative solutions. However, few studies have investigated the specific strategies elementary school students employ during building block construction. This study conducted a literature review to classify and summarize these strategies into seven categories: exploratory strategies, planning strategies, imitative and adaptive strategies, decomposition and integration strategies, reflective and iterative strategies, collaborative and division of labor strategies, and creative strategies. The results provide a theoretical framework for future instructional design, learning process documentation, and assessment development in STEM education.*
Keywords: STEM education, block, strategies, engineering design thinking, problem-solving strategies

1.前言

在 STEM 教育的課程設計中，積木經常被作為工程設計與結構建構的訓練工具。積木是一種可重複組合、模組化的建構材料，通常由塑膠、木材、磁鐵或紙板等材質製成。它具有標準化的接合方式（例如凹凸扣合、磁力吸附、插槽設計），可輕易堆疊、拆解與重新組裝。在教育上，積木是一種動手操作的學習媒介，支援學生透過具體實作，進行抽象概念的理解與建構。它可以用來模擬真實世界的結構、機械系統或數學模型，是具象化思考過程的橋梁。這種以「做中學」為核心的學習方式，除了能提升學生的空間思維與創造力，更能培養他們的工程設計思維與團隊合作能力，符合 STEM 教育強調的跨領域素養與實作精神。

許多研究已證實積木建構活動對於提升學生的空間能力、結構概念理解、創造力及問題解決能力均具有正面影響。然而，現有文獻多聚焦於活動成效評估或教師的教學設計策略，較少深入探討國小學生在實際參與積木建構活動時所運用的策略與行為模式。特別是國小學生的認知發展、學習動機及問題解決歷程，可能與高年級學生或成人有顯著差異，因此，理解國小學生在積木搭建過程中的策略選擇與思維歷程，將有助於教師設計更符合學生認知發展與學習需求的活動引導策略與評量工具，進一步提升 STEM 活動的學習效果與學生參與度。

基於上述背景，本研究的主要目的在於探討國小學生於 STEM 積木建構活動中的策略運用類型與特徵。研究將透過文獻探討，歸納國小學生在積木搭建過程中可能使用的策略類型，並建立初步的策略分類架構。期望研究結果能為未來 STEM 積木活動的教學設計、學習歷程評量，以及學生工程思維發展提供理論基礎與實務參考。

2.文獻搜尋策略

本研究以 Scopus 為資料庫，條件為標題、摘要與關鍵字包含 STEM、Engineering 和 Building Blocks((TITLE-ABS-KEY (stem) AND TITLE-ABS-KEY (engineering) AND TITLE-ABS-KEY (building AND blocks)))，共計有 407 篇。此研究將以這 407 篇文獻進行文獻分析。

3.學生策略運用類型

依據文獻分析的結果，本研究歸納出以下七種常見的類型：

1. 探索性策略 (Exploratory Strategies)

這種策略強調學生透過「做中學」，不預設任何固定的建構方法或最終成果。他們會以積木為媒介，進行隨機嘗試與排列組合，目標是熟悉材料特性與了解結構的可能性。過程中，容許試錯，甚至鼓勵學生進行無目的的探索，從探索中發現靈感或技術細節(Berland & Wilensky, 2015)。例子如：學生拿到一堆積木後，先拼接出各種不同的形狀，像是長橋、塔、斜坡等，看看積木的連接穩定性如何。

2. 規劃性策略 (Planning Strategies)

這種策略強調「先思考後動手」，特別適合目標清楚、結構複雜的任務。學生會花時間規劃整體藍圖、設計分工步驟，甚至預先評估哪些積木適合哪些位置，以提高建構效率與成功率(Goldman & Kabayadondo, 2016)。例子如：在設計積木橋之前，先畫一張橋的結構草圖，標出哪些部分需要特別加強。

3. 模仿與適應策略 (Imitative and Adaptive Strategies)

學生透過觀察他人、參考範例或閱讀說明書來學習技巧與思路，這是很重要的學習來源。不同的是，優秀的模仿者會根據自身條件進行適應，將原始設計改造成更符合自己目標或創意的作品(Hmelo-Silver & Barrows, 2006)。例子如：看到老師搭了一座橋，學生先照著老師的方法搭一座類似的橋，之後把部分橋墩改成可旋轉的機關，提升作品的趣味性。

4. 分解與組合策略 (Decomposition and Integration Strategies)

這種策略將大目標拆解成小單元，各個擊破，最後再把所有單元整合成完整作品。特別適合大型、複雜或功能多元的結構。這樣的拆解式思考不僅有助於降低難度，也可以方便多人協作(Jonassen, 2000)。例子如：建造一台積木機械手臂，先把「底座」、「關節」、「手掌夾具」三個單元分開設計，確認各部位功能正常後再組合起來。

5. 反思與修正策略 (Reflective and Iterative Strategies)

這個策略強調「邊做邊想」，學生隨時反思作品的穩定度、美觀度、功能性等，並根據反思結果立即修正，形成「建構-測試-修改」的循環迴路。這是一個培養工程思維的重要策略(Schön, 2017)。例子如：學生在搭建一座橋時，發現中段承重不足，於是立刻拆開，改用三角結構加強。

6. 合作與分工策略 (Collaborative and Division of Labor Strategies)

這種策略適合團隊合作的活動，強調明確分工、協商溝通與資源共享。每位成員負責特定區塊或功能，同時要保持整體設計的一致性與流暢性，並透過溝通調整彼此工作(Barron, 2000)。例子如：小組合作建造積木摩天輪，一人負責輪軸設計，一人負責座艙造型，一人負責底座結構，最後整合成一個完整摩天輪。

7. 創意性策略 (Creative Strategies)

這種策略強調打破框架與自由發想，學生不受既有範例或規則限制，盡情發揮想像力，甚至刻意挑戰傳統積木玩法，創造全新的設計或功能。這種策略尤其適合開放式任務或競賽型活動(Resnick & Rosenbaum, 2013)。例子如：學生不按積木說明書操作，自己設計一座「漂浮城市」，利用透明線或磁力裝置讓部分積木懸空。

4.結語

積木建構活動作為 STEM 教育中的重要實作工具，不僅讓學生透過操作體驗結構設計與工程挑戰，更是發展空間思維、創造力與問題解決能力的重要途徑。然而，學生在積木建構過程中所運用的策略歷程，過去研究較少深入探討。本研究透過文獻探討，歸納並整理國小學生於 STEM 積木建構活動中的策略運用類型，並建構七大策略分類框架，包括：探索性策略、規劃性策略、模仿與適應策略、分解與組合策略、反思與修正策略、合作與分工策略，以及創意性策略。

這些策略類型反映了國小學生在面對積木建構挑戰時，從探索、規劃、操作到反思修正的完整歷程，並展現出多元化的思維方式與問題解決模式。研究結果除了提供教師設計積木建構活動的教學引導參考，亦可作為學生學習歷程紀錄與策略發展評估的基礎工具。未來，建議進一步透過實證觀察與學生歷程檔案分析，驗證不同策略類型的使用頻率、策略轉換模式與其對學習成效的影響，逐步發展具備教育現場適用性的積木建構活動策略評量工具，為 STEM 教育中的工程設計教學與學習評量提供更完整的支持與參考。

參考文獻

Barron, B. (2000). Achieving Coordination in Collaborative Problem-Solving Groups. *The Journal of the Learning Sciences, 9*(4), 403-436.

Berland, M., & Wilensky, U. (2015). Constructionist Gaming: Understanding the Benefits of Making Games for Learning. *Educational Psychologist, 50*(4), 248-259.

Goldman, S., & Kabayadondo, Z. (2016). *Taking Design Thinking to School: How the Technology of Design Can Transform Teachers, Learners, and Classrooms*. Routledge.

Hmelo-Silver, C. E., & Barrows, H. S. (2006). Goals and Strategies of a Problem-Based Learning Facilitator. *Interdisciplinary Journal of Problem-Based Learning, 1*(1), 21-39.

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology Research and Development, 48*(4), 63-85.

Resnick, M., & Rosenbaum, E. (2013). Designing for Tinkerability. *Design, Make, Play: Growing the Next Generation of STEM Innovators*, 163-181.

Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge.

生成式人工智能赋能科普图像及动画创作—以航天科普动画“落火”为例

Empowering Science Animation with Generative AI: A Case Study of the Aerospace Science Animation “Landing on Mars”

戢逸雯[1]，陈弘正[2*]，马慧[1]
[1]黄冈师范学院 教育学院
[2]黄冈师范学院 机电与智能制造学院
[*] hungcheng@hgnu.edu.cn

【摘要】 本研究聚焦生成式人工智能在航天科普动画创作中的应用实践，以中国首次火星探测任务“天问一号”着陆过程“落火”为核心内容，系统探讨了 AIGC 技术在动画脚本生成、图像设计与动态视频合成等环节中的具体运用。研究通过收集权威资料与图像素材，借助 ChatGPT 辅助脚本撰写，结合 Midjourney 和可灵 AI 等平台进行图像与动画生成，构建出角色拟人化、画面卡通化、节奏专业化的科普动画作品。在后期处理方面，本研究整合剪辑、音效与配音流程，打造完整可用的航天主题科普短片。研究表明，AIGC 不仅能提升动画创作效率，也为科学传播提供了更具亲和力的表达路径。
【关键词】 生成式人工智能；科普动画；航天教育；图像生成；AI 脚本设计

Abstract: *This study explores the application of generative artificial intelligence (AIGC) in the creation of aerospace science animation. Focusing on the Mars landing stage of China’s Tianwen-1 mission, the project investigates the implementation of AIGC in scriptwriting, image generation, and video production. By collecting official visual materials and scientific data, the study utilizes ChatGPT to draft scripts, and applies platforms like Midjourney and Keling AI to generate stylized characters and animated scenes. The resulting animation adopts a cartoon aesthetic and humanized narration to enhance engagement. In post-production, editing, audio effects, and voice acting are integrated to complete a science communication short film. The findings suggest that AIGC significantly improves production efficiency while offering new creative possibilities for public science outreach.*
Keywords: Generative AI, Science Animation, Space Education, Image Generation, AI Scriptwriting

1.研究背景

传统的 3D 动画制作流程往往是一个高度复杂、耗时且依赖大量人力的过程。从 2D 概念设计、关键场景草图绘制，再到使用如 Blender、Maya、ZBrush 等建模软件创建三维角色与场景，接着进行骨骼绑定与角色 Rigging，之后才是动画制作、材质处理、光影调整，最终渲染成片。而每一个环节都需美术、动画、灯光、渲染等多个团队紧密配合，整个流程少则数周，多则数月乃至一年。

近年来，生成式人工智能（AIGC）技术在图像、文本与视频内容生成方面的突破，正在深刻变革动画创作与科学传播的方式（贺京华 & 姜皖，2024）。尤其是在动画领域，AIGC 不仅可在剧本生成、角色建模、场景设计等前期环节显著提升效率，更在后期特效合成与音效生成方面展现出强大能力，推动传统动画向“人机协同”模式转型（孙亮 & 沈超颖，2023；王硕等，2024）。

AIGC 的崛起为科普动画创作带来了新的可能性。通过提示词驱动生成图像与动画，不仅大幅降低了美术与建模门槛，还赋予内容生成更多的灵活性与风格多样性（黄高乐，2024）。已有研究指出，AIGC 可以有效辅助科学知识的“视觉转译”，提升受众的理解效率与接受意愿（朱颖祺，2024）。特别是在多模态模型与扩散模型（如 Midjourney、Stable Diffusion）持续发展下，AIGC 在“图文—动画—音效”全流程中的应用潜力日益凸显（祝智庭等，2023）。表 1 比较传统动画与 AIGC 动画制作流程的差异。

表 1 传统动画与 AIGC 动画制作流程比较

环节	传统动画制作	AIGC 动画制作
概念设计与草图阶段	原画师需手工绘制大量概念图、分镜脚本，反复修改耗时。	使用 Midjourney、Stable Diffusion 等工具输入文本描述，AI 可在几分钟内生成数十版方案，供创作者筛选或优化。
3D 建模与资产制作	角色/场景建模需艺术家在 Blender/Maya 中手动拓扑、雕刻。纹理绘制则依赖 Photoshop 或 Substance Painter 手工完成。	可灵 AI 一步到位——图生视频功能（整个过程只需 1-2 周）
角色绑定与动画	绑定师需手动设置骨骼权重、控制器。动画师需逐帧调整关键帧或使用动作捕捉。	
渲染与后期	高精度渲染依赖 CPU/GPU 长时间计算。特效合成需手动调整粒子、光照等参数。	
人力成本	多人团队	1 人

2 落火”航天科普动画内容设计与实现

本文的科普动画以“天问一号”的“落火”过程为核心内容，围绕中国首次火星探测任务中的离轨着陆阶段，设计了一部航天科普动画短片。本节将介绍动画创作的背景与脚本设计过程，为后续的图像生成与动画制作提供基础支持。

2.1. 火星探测器的结构组成

“天问一号”是中国首次实施的火星探测任务，包含一体化设计的环绕器与着陆巡视器。其中，环绕器负责在轨运行与中继通信，着陆巡视器则承担着陆与地表探测的关键任务。分述如下：

- 环绕器外部结构主要由太阳能帆板和高增益定向天线组成，具备独立轨道控制能力。在火星上空长期运行后，它为着陆器提供关键的定位数据与遥测支持。
- 着陆巡视器由背罩、大底、着陆平台和火星车构成，是“落火”过程中的主角。着陆平台搭载了推进系统与控制装置；火星车“祝融号”则是最终驶上火星地表的执行者。
- “祝融号”火星车高约 1.85 米，重达 240 公斤，拥有六轮设计。其顶部设有一对导航与地形识别相机，中间标有“火”字篆刻印章，富有中国文化色彩，火星车整体视觉上如一只展翼的蓝色蝴蝶。

2.2. “落火”过程九步详解

“落火”是指天问一号从火星轨道脱离并成功在火星地表软着陆的全过程。整个过程复杂精密，可细分为以下九个环节：

1. 分离准备与姿态调整：在轨飞行三个月后，探测器接受来自地球的降轨指令。减速降低轨道，18 分钟后进入撞向火星表面的飞行弹道。环绕器发动机前置减速，转 90 度侧向建立分离姿态。
2. 侧向分离与滑行阶段：着陆巡视器从环绕器侧向释放，中间有一段时间的飘飞以避免推进器尾喷对彼此飞行姿态产生干扰。
3. 进入火星大气层：着陆器以钝头朝前的姿态高速进入火星稀薄大气，约 125 公里高空开始“恐怖九分钟”。通过调整攻角和姿态，实现初步减速与稳定飞行。
4. 攻角配平与升力控制：火星大气层中，风的速度和方向，对落火的精度影响很大，着陆器下降配合背罩上部的小型配平翼，利用气动阻力稳定姿态，使气流对称以顺利开伞。中国团队首创配平翼设计，增强飞行稳定性。
5. 伞降控制：降落伞在超声速阶段弹出，之后大底抛开，着陆平台底部暴露。测速测距雷达启动，提供导航信息。速度下降后，着陆器姿态稳定。
6. 抛伞并启动发动机：在距地表 1.5 公里处，降落伞与背罩脱落，着陆平台主推力发动机点火，进入动力减速段，进一步降低下落速度。
7. 避障与成像决策：进入低空后，着陆器悬停于 100 米处，激光成像火星地貌，规避表面石块与坡度较大的区域，选择相对平坦的落点。
8. 缓速下降与冲击缓解：主发动机调节推力，使着陆器平稳下降。四条着陆腿内置的缓冲材料吸收下落冲击，确保设备安全。
9. 火星车驶离平台：约 12 分钟后，坡道展开，火星车的太阳能板依次张开。“祝融号”火星车通过自动程序完成解锁、启动、转向，1 小时内完成全部检测。约一周后，它成功驶离平台，开始火星探索旅程。

2.3. AI 辅助生成动画脚本

为增强科学准确性与视觉表现力，本研究采用生成式人工智能（如 ChatGPT）协助构建动画脚本。其应用流程如下：

1.科学资料结构化整理：收集自央视纪录片《你好！火星》第四集（朱宏展，2022）及哔哩哔哩 CNSA 中国航天文化官方账号发布的科普纪录片《火星来了》（万珂，2021）将技术信息提炼为关键信息片段。

2.编写提示词，生成脚本初稿：将结构化文本输入 ChatGPT，使用以下提示词引导脚本生成：“阅读以上文档，根据内容编写一份关于航天科普动画的脚本，脚本需包含画面描述、对应台词、时间轴，时长控制在 2-3 分钟，内容涵盖天问一号离轨着陆阶段，展现着陆巡视器如何降落在火星上，要求科学性与趣味性相结合。”

3.多轮调整，优化脚本逻辑：结合人工校审，进一步润色 AI 生成文本，确保内容衔接自然、语气生动，适用于青少年观众的科学传播。

2.4. 脚本示例

表 2 节选片段聚焦影片中节奏起伏最强烈的环节，结合关键场面、视觉节奏与音乐变化，展现动画在“落火”过程中的叙事高潮。

表 2 “落火”动画脚本示例

动画段落	时长	画面描述	台词 / 背景音
进入大气	5 秒	着陆器在接近声速状态下穿入火	背景音乐：纯音乐 2，增强太空

		星大气，背景橘红，画面开始剧烈抖动，火星表面轮廓逐渐显现	神秘感与张力（无对白）
配平翼展开	7 秒	背罩上弹出一个长方形小窗，配平翼缓缓伸展，通过改变气动布局配平飞行姿态	背景音乐：纯音乐 2 延续，旋律持续高张（无对白）
伞降控制	5 秒	舱顶弹出伞包，降落伞在高空中极速展开，气流卷动周围沙尘，减速效果显著	背景音乐：纯音乐 2，节奏迅速加强（无对白）
动力减速	5 秒	降落伞与背罩抛弃，平台上大推力发动机点火，喷出明亮蓝白火焰，开始动力减速	背景音乐：纯音乐 2 持续增强（无对白）
缓速下降	5 秒	成像完成，着陆器缓慢下降，姿态稳定，火星表面特写逐帧推进	背景音乐：纯音乐 2 稍作收敛，节奏转缓
避障机动	5 秒	着陆器漂移调整姿态，避开已抛出的组件并初步选定着陆区域，轻微机动转向	背景音乐：纯音乐 2 延续（无对白）
悬停成像	5 秒	距离地表约 100 米时，着陆器静止悬停，开启激光扫描火星地表，寻找合适着陆点	背景音乐：纯音乐 2 张力达到顶峰，激光扫描声效嵌入
着陆	7 秒	四条着陆腿展开，缓冲材料吸收冲击，尘土扬起，平台平稳落地，镜头下摇呈现着陆点全貌	背景音乐：纯音乐 2 转为收尾段落，气氛释然，庄重

3.航天科普动画“落火”AI 生成图像设计

本节聚焦于动画“落火”中静态图像设计的关键环节，探讨如何通过生成式人工智能工具，实现科学性与艺术性表现力兼具的视觉素材。图像设计均围绕“环绕器”、“着陆巡视器”与“火星车”三个主要角色及其任务场景展开，为后续动画动态生成与剪辑提供图像基础。

3.1. 图片素材收集及预处理

在开始 AI 图像生成之前，必须收集并处理具有参考价值的角色原始图像素材。这些图片主要来自航天科普宣传、纪录片截图以及官方发布的设备照片。

本研究首先通过 Photoshop 等图像处理软件对素材进行分割、抠图与调整，并根据拟人化需求，对角色图像进行局部改造与重构，确保轮廓不变形，细节完整：

- 着陆巡视器：使用套索工具更改局部色彩，补全顶部与按钮细节，并添加眼睛和嘴巴图层，使形象更具拟人感。
- 环绕器：替换绿色区域为黄色，添加太阳能板图层，并抠出整张角色图像，准备作为 AI 生成的“垫图”素材。
- 火星车：利用魔棒工具与橡皮擦分离车体与轮子部件，确保部件可以独立生成与组合。

这些前处理操作为 AI 模型后续生成提供清晰准确的形状与结构信息，有效提升图像输出的科学性与可控性。

3.2.提示词设计

本研究结合使用 OPS 提示词工作室（https://moonvy.com/apps/ops/）的提示词可视化编辑工具，支持提示词的分类、排序与词典引用，使提示词的组织更为清晰。使用流程包括（1）提炼角色与场景的关键词特征；（2）利用翻译软件将其转换为英文，校对语义表达；（3）在提示词工具中分类组合成风格、构图、细节类词组；及（4）输入至 Midjourney 并生成图像。

动画场景多为火星地表与太空背景，需准确还原火星地貌特征。本研究使用“可灵 AI”生成火星场景，并参考官方“天问一号着陆”画面，确保科学性与统一风格。

3.3. 图像细节刻划技巧

AI 图像生成有时会在细节部分偏离预期，尤其是在复杂结构、特殊角度或部件层次方面。为提升图像质量，本研究使用“以图生图”和“局部重绘”等技巧，在 Midjourney 平台中实现针对性修正。

3.1.1 以图生图（垫图）技巧

该方法借助一张已有的参考图（经过 Photoshop 处理），作为“视觉引导”，使 AI 图像更贴近设想中的构图与风格。重点是通过提示词强调需要补足或修改的部分，而非全部重新生成。例如：在环绕器原图上增加太阳能板时，使用以下提示词（图 1(a)－图 1(b)）：

There are three rectangular solar panels on each side, It is blue-white in color --cref 图片链接 --ar 16:9 --niji 6

此处的“--cref”参数引用了处理后的图像链接，确保生成图贴合原始结构。

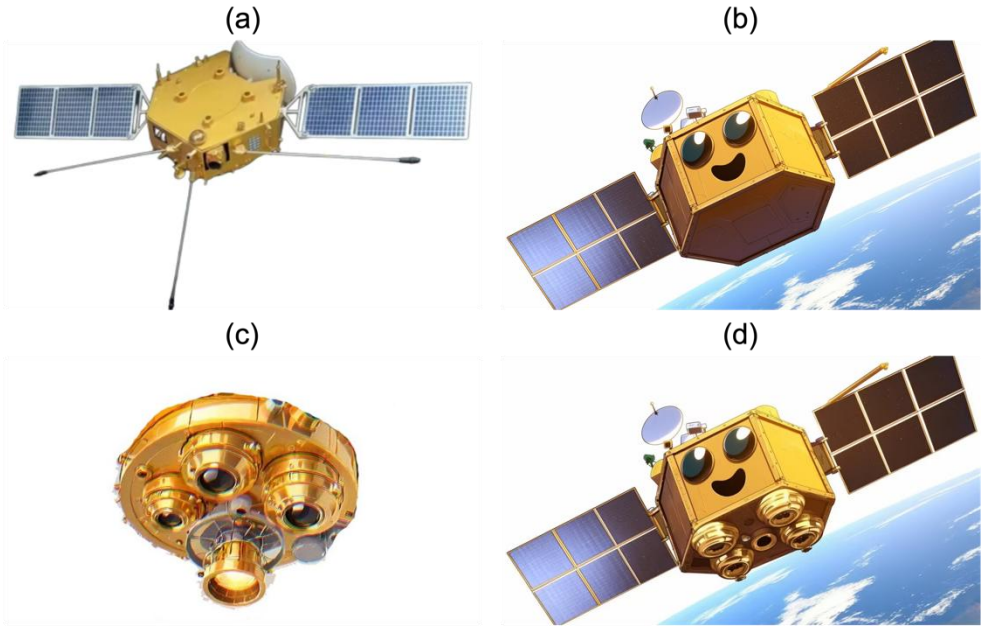


图 1 环绕器 AI 生图细节刻划設計。(a)-(b)為以图生图技巧，(c)-(d)為局部重绘技巧。

3.3.1. 局部重绘技巧

Midjourney 支持用户选定图像区域并仅重绘该部分。适用于当整体构图满意，但某处细节（如按钮、喷嘴）略有偏差的场景。重绘前输入新的提示词以精修局部。例如：

A golden-yellow conical ejector --cref 图片链接 --ar 16:9 --niji 6

这一策略被用于完善喷射器、燃料罐等部位（图 1(c)－1(d)）。

4.航天科普动画“落火”AI 生成动画设计

本节聚焦“落火”动画在生成过程中的关键技术路径，分为两个部分展开：其一为动画的生成过程，涵盖分镜头设计与 AI 视频生成；其二为后期处理，包括音效、剪辑与配音。核心内容聚焦于如何借助 AI 工具实现动态画面与科学叙事的有机结合。

4.1. 动画生成

4.1.1. 分镜头的图片设计

动画的第一步，是根据剧本内容制作分镜图像。此阶段决定了整体叙事节奏与画面结构，是 AI 生成动画的基础。本研究通过 Photoshop 完成 18 个动画片段的图像设计，部分片段只需首帧图像，复杂场景则需设计完整的首尾帧。操作步骤包括：

- 阅读脚本、分析画面：首先深入理解每个镜头的剧情与背景，参考国家航天局视频与纪录片截图，确保构图准确，细节科学。
- 角色与背景分离：利用 PS 工具将人物与场景抠图，重组为可用于动画合成的素材，为 AI 输入做好准备。
- 细节精绘：根据画面所需，设计如光线变化、气流流动、光束扫描等特效。尤其是“着陆器进入大气层”或“平台激光成像”等关键画面，这些动态效果直接影响后续 AI 渲染质量。

这些高质量图像为 AI 生成提供清晰输入，保证动画逻辑连贯、画面统一。

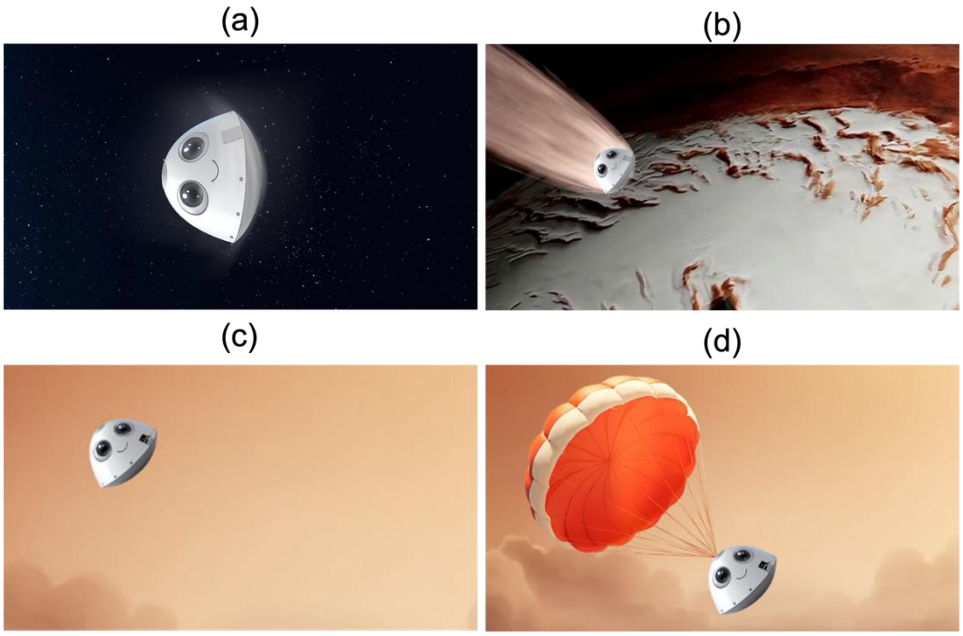


图 2 AI 生成动画首尾帧生成技巧，分别为“进入大气”的（a）首帧及（b）尾帧，以及“伞降控制”的（c）首帧及（d）尾帧。

4.1.2. AI 生成动画

分镜图像准备完毕后，进入 AI 动画生成阶段。本研究采用“可灵 AI”中的图生视频功能，根据输入的场景图像与文字描述，自动生成对应动态画面。此外，可灵 AI 动画生成依赖提示词模板“主体+运动，背景+运动”，需清晰描述主角的动作轨迹、背景的变化状态。为保持角色稳定性，还需附加“主体形状保持不变”等关键词。

本研究采用了两种关键技巧：

一、首尾帧生成：

适用于动作衔接、场景过渡等动态较明显的片段，能提升运动流畅性和画面自然度。例如：

- “进入大气”动画片段（参考图 2(a)、图 2(b)）：

提示词：“*主体在眨眼睛，做向下俯冲运动，主体其他部位保持不变，主体周围白色气流在动，背景星星在闪烁，背景也在运动。*”

- “伞降控制”动画片段（参考图 2(c)、图 2(d)）：

提示词：“*主体在做下降运动，后面天空也在变化，下降，主体外形保持不变，突然，降落伞打开。*”

通过这些提示词，动画不仅真实还原“落火”过程中的关键动作，还能维持角色在镜头中的一致性。

二、多图参考：

这个技巧用于火星地表环境中角色与背景融合困难的情境，尤其适用于后半段场景（如火星车与地形互动）。例如：

- “避障机动”动画片段（参考图 3(a), 3(b)及 3(c)）：

提示词：“*火星车乘坐着着陆平台一直在空中飞行，下方发射着白色的发射气和紫色的圆形探测波，扫描火星的地形。*”

- “悬停成像”动画片段（参考图 3(a), 3(b)及 3(d)）：

提示词：“*火星车乘坐着着陆平台在空中飞行，发射着绿色光线网格进行扫描火星的地形。*”

通过上述技巧，AI 能够较好地合成多图信息并渲染出光影效果与动态效果并存的场景，画面更加统一、生动。

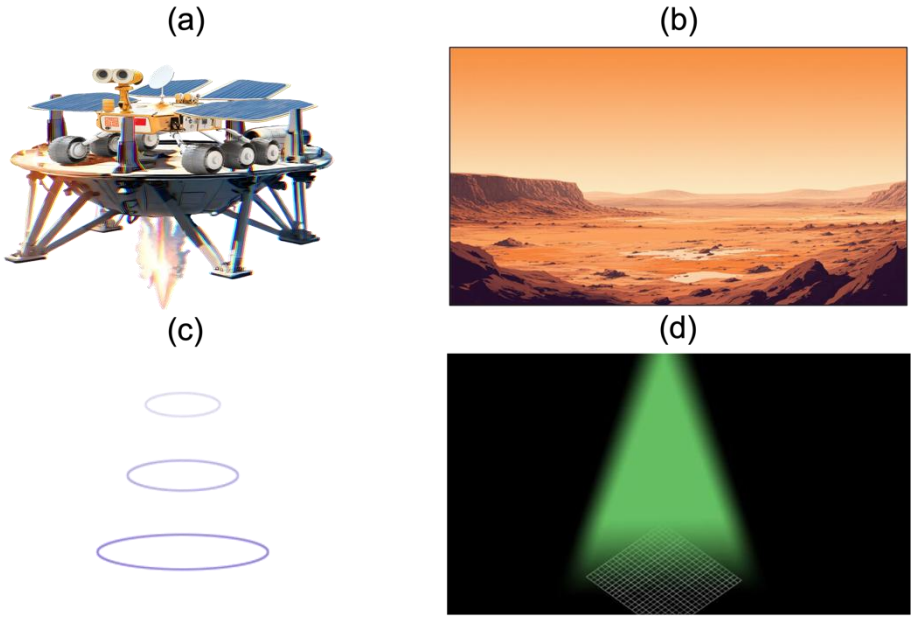


图 3 AI 生成动画多圖參考生成技巧，分别为避障机动”动画片段的(a)、(b)及(c)，以及悬停成像”动画片段的(a)、(b)及(d)。

4.2. 后期处理

后期处理是将画面、声音与叙事逻辑整合成完整作品的重要阶段。《落火》动画的后期制作主要围绕音效与配乐、画面剪辑以及角色配音展开，使用的核心工具包括剪映和 SUNO。通过这一系列处理，动画的表现力与科普传达效果得以显著提升。

在音效设计方面，本研究大量使用剪映内置的音效素材库，模拟如“冲入大气层”、“降落伞打开”、“喷气减速”与“激光扫描”等关键情节的声音效果。通过叠加不同音效并调整音量比例，使声效更具层次感与真实感。例如，表现“着陆平台”着陆时，选用了“沉闷的击中”配合“火星尘土飞扬”的音效，增强了视觉冲击力。

配乐方面则全部由 AI 音乐生成工具 SUNO 创作完成。该工具支持根据提示词生成与场景气氛相符的背景音乐。为适配不同动画段落，本研究为每个片段输入情绪描述词，例如紧张、轻松、科技感等，来引导音乐风格生成。紧张的着陆段落配以节奏强烈的悬疑音乐，而讲解段落则配以轻快明亮的旋律，从而使观众在观看时更容易沉浸其中。

完成音效与配乐后，动画进入剪辑阶段。本研究利用剪映平台对镜头衔接、节奏控制与视觉特效进行了整体编排。片头采用“摄影框+进度条”风格，营造纪录片氛围，明确动画主题。镜头之间通过推进、淡出等转场方式增强节奏流动性，结尾则通过角色合影与字幕打字机特效收束全片，营造完整的叙事闭环。

在配音方面，本动画的两位主要角色“陆陆”与“绕绕”采用不同风格加以区分。“陆陆”作为活泼角色，配音使用轻快语调，由手机录音完成，强调灵动与亲和力。“绕绕”则语调沉稳，后期通过剪映降调处理，展现其可靠、稳重的特质。所有配音录制后统一进行美化、降噪处理，确保语音清晰自然，贴合角色个性。

通过这一系列后期处理流程，动画不仅在画面上实现了动态流畅，更在声音与情绪传达上增强了叙事表现力，使科学知识在视觉与听觉的双重引导下更易于被理解与接受。

5. 结论与建议

本研究以“天问一号”着陆过程为蓝本，实践探索了生成式人工智能在科普动画创作中的多环节应用。从剧本生成到图像建构，再到动画制作与后期处理，整体流程融合了 AI 工具与人工审校，展现了教育技术在科学传播中的实际潜力。研究发现，AIGC 不仅能显著提升动画制作的效率，还能增强科普内容的趣味性与可视化表现，尤其适用于面向青少年群体的航天教育内容设计。

此外，本研究在操作层面总结出若干具有实用性的提示词设计与图像修正技巧，验证了“首尾帧生成”“以图生图”“局部重绘”等方法在动画生成中的适配效果。通过可灵 AI 与剪映等工具组合使用，最终形成一部具有完整传播功能的航天科普短片，展示了 AIGC 在教育传播场域中的创新应用价值。

然而，AIGC 在内容一致性、科学准确性与伦理规范等方面仍存在挑战。未来研究可进一步探索多模态协同生成机制、学科专家参与机制及 AI 输出审核机制，推动生成式人工智能在科学传播中的深入发展与规范使用，为新时代科普创作提供技术支持与方法参考。

参考文献

贺京华、姜皖 (2024)。AIGC 驱动下的动画创作：技术变革、融合路径与风险挑战. 北京印刷学院学报，**32**(5), 68–74.

孙亮、沈超颖 (2023)。赋能与重构：生成式人工智能浪潮下动画创作的新路径. 电影评介, (20), 6–11.

王硕、阎妍、李正风 (2024)。生成式人工智能赋能科学普及：技术机遇、伦理风险与应对策略. 科技传播，**19**(4), 5–13.
黄高乐. (2024)。生成式人工智能技术赋能科普内容创作. 科学教育与博物馆，**10**(1), 22–27.
朱颖祺. (2024)。生成式人工智能在科普传播中的应用研究. 数字通信世界 (11), 165–167.
祝智庭、戴岭、胡姣 (2023)。高意识生成式学习：AIGC 技术赋能的学习范式创新. 电化教育研究。**44**(6), 5–11.
朱宏展 (导演). (2022).《你好！火星》第 4 集 [Video recording]. https://shorturl.at/nG2Pl
万珂 (导演). (2021). 火星来了第三季 [Video recording].【全集！火星来了第三季】 https://shorturl.at/CAqJN

結合擴增實境與 STEAM 教育之幼兒創傷知情互動軟布書設計與實作

Design and Implementation of an AR-Integrated STEAM-Based Interactive Soft Book for Trauma-Informed Early Childhood Education

吳采綸[1]，李佳潔[1]，謝旻儕[1*]
[1]台南應用科技大學多媒體動畫系
[*] shiehminchai@gmail.com

【摘要】 本研究結合創傷知情設計、STEAM 互動式軟布書與擴增實境（AR）技術，以聯合國永續發展目標（SDGs）第三項「健康與福祉」為核心，鑒於童年創傷對身心發展的長期影響，以及傳統教育較缺乏心理支持機制，本研究透過可觸材質與 AR 搭配，提供兼具教育與療癒功能的互動軟布書，並參考孩子大腦發育過程，採用輔導三步驟作為中心發想並製作成互動式軟布書，透過擴增實境真實結合虛擬的元素，豐富化布書內容，再藉由元素結構讓兒童動手互動，並打造出「創傷知情」的氛圍，讓周圍的人們能提供更多支持與理解，幫助更多受害者走出陰影。

【關鍵字】 擴增實境；互動軟布書；STEAM 教育；創傷知情

Abstract: *This study integrates trauma-informed design, STEAM-based interactive soft cloth books, and augmented reality (AR) technology, centering on the United Nations Sustainable Development Goal (SDG) 3: "Good Health and Well-being." In light of the long-term impact of childhood trauma on physical and mental development, and the lack of psychological support mechanisms in traditional education, this research proposes an interactive soft cloth book that combines tactile materials with AR to offer both educational and therapeutic value. Drawing on the developmental stages of children's brains, the design is based on a three-step counseling model and is developed into an interactive format. Augmented reality enhances the content by integrating virtual elements into the physical book, allowing children to engage in hands-on interaction. Through this element-driven structure, a trauma-informed atmosphere is created, encouraging surrounding individuals to offer greater support and understanding, ultimately helping more survivors emerge from the shadows of trauma.*

Keywords: augmented reality, interactive soft cloth book, STEAM education, trauma-informed

1. 前言

在生活中，一般人在經歷較輕微的壓力或意外事件時，可能會產生像焦慮、失眠、注意力不集中、沮喪等心理反應，通常經過一段時間的調適，大多數人能夠恢復到正常的心理狀態。創傷後壓力疾患（Post-Trauma Stress Disorder, PTSD）是一種在經歷重大創傷事件後表現出異常創傷心理反應的狀態。包括：症狀持續的時間、強度和種類，並會影響到心理社會功能。在嚴重的情況下，可能還會併發其他身心疾病，如憂鬱症、物質濫用和人格分裂等問題，影響深遠，不容忽視（曾美鳳、陳信昭，2018）。

本研究希望經由互動式軟布書輔導受輔導的孩子如何面對創傷的目的。利用創傷知情設計，提供孩子安全（Safe）、平穩（Stable）、滋養（Nurturing）的生活環境，並透過創意

和藝術的力量來打開溝通和表達的管道，幫助有童年創傷的人尋找療癒和連結，並且促進社會中對於心理健康的關注和支持。對於孩子來說，互動立體布書通常更具吸引力，能夠引發好奇心，鼓勵孩子們對閱讀和探索感到興趣。

2. 文獻回顧與探討

2.1 兒童創傷前因後果與分析

創傷是要有暴力或是面臨生命威脅經驗，事實不然，創傷是非常主觀的感受，對於孩子來說，只要是不安全、無助的經驗都有可能造成壓力或創傷，創傷可能是單一事件、也可能是慢性長期的負向經驗感受，這些感受累積會影響兒童大腦，進一步衍生出情緒及行為問題（Wilson et al., 2013）。兒福聯盟以創傷知情的四要素來理解和回應孩子的自我破壞行為，如下表 1。

表 1 創傷知情四要素 4R (本表參考兒福聯盟並重新彙整)

創傷知情四要素 4R	理解問題點	解決方法
理解創傷(Realize)	了解創傷對大腦的傷害	提供創傷介入服務的專業人士應該了解創傷的廣泛存在和對個體的深遠影響，同時理解倖存者展現的行為狀態是他們對過去創傷經驗的一種適應性反應。
辨認創傷(Recognize)	辨別創傷的行為跡象	在倖存者和參與創傷介入服務體系的工作者身上，辨識出創傷症狀和相關的警訊。
利用創傷知識做回應(Respond)	提供支持性的回應和適當的幫助	在制定政策、流程以及實踐工作上，充分運用創傷相關的知能。這並不一定要求立即認同行為，而是可以從理解和認同情緒的角度出發。
防止再度受創(Resist Re-traumatization)	避免再次受到傷害	從察覺負面循環開始，正視並調整，以終止再度創傷。

2.2 調節連結再說理

此觀點是基於依附理論（Attachment Theory）的概念，該理論由英國心理學家 John Bowlby，在 1950 年所提出，強調嬰幼兒對於主要照顧者（通常是父母）的情感依附。這種關係對於嬰兒的安全感、情緒調節和發展起著重要的作用。父母的關愛、支持和保護能夠建立起一種安全基地，讓孩子有信心去探索世界（藍逸梅、李從業、陳嘉琦，2001）。此安全基地不僅體現在實際的身體接觸中，還包括了在腦中形成的安全感觀念，嬰幼兒會發展出一種對於主要依附對象的內在表示，包括：了該對象的各種感覺記憶，如：臉、聲音、味道、觸感等。這樣的內在表示成為一種「安全堡壘」，在面對壓力或焦慮時，嬰幼兒可以回憶起這些感覺，以此求得慰藉。若嬰幼兒在和照顧者關係中經常感到壓力、不安全感，諸如此類感受會進入他們右腦的長期記憶中，在內心對自己和父母產生一個失去功能的自己和他者。在面對和家長分離的情境下，受創的兒童很容易會表現出困惑、大發脾氣、無法被安撫；而當家長回來後，又可能同時表現出想親近和躲避之矛盾或行為不一致的混亂依附情況。混亂依附的兒童也會有一些特殊行為，例如：凍結、茫然或靜止、中斷的動作和表情（陳雅雪，2009）。

「治療」是在認知創傷形成後的一大課題，目前主流 3 種治療模式，分別從大腦發展、加入照護者共同面對，以及重建安全感，帶領兒童走上療癒之路（朱崇信，2021）。當孩子的神經覺感到危險、出現下層腦行為，這時候需要使用「由下往上」的方式處理，安撫孩子的下層腦。美國精神科醫師 Bruce Perry 提出面對孩子「由下往上」的三個步驟為調節

（Regulate）：先穩定呼吸與身體，讓身心從失控的狀態恢復平穩；同理（Relate）：當孩子感覺到被支持同理、感到安全，漸漸能不受情緒腦所掌控；再說理（Reason）：當情緒平穩後，便有機會開始思考，練習以理性腦表達自己的情緒和想法（留佩萱，2023）。

2.3 STEAM 教育的目標

美國率先提出以 STEAM 教育，培養動手做、發明、創新的下一代，後來再加入藝術（Art），成為 STEAM，期待學習更完整，周淑惠（2018）提出 STEAM 的四大特徵為：面對生活真實問題，以解決問題為目標；運用探究能力以求知、理解的「探究取向」；運用設計、製作與精進的「工程活動」；運用科學、數學與各類技術的「統整課程」。使學生能夠在解決問題的過程中綜合應用各種技能和知識。同時鼓勵學生主動學習，培養他們的好奇心和求知慾望。學生被引導思考、提問、解決問題，培養了批判性思維，具備應對未來挑戰的能力，塑造成為有創造力、有思考力、有人文素養的個體。

3. 實作與設計

本研究實作與設計兒童 STEAM 教育結合 AR 互動式軟布書並應用於兒童創傷知情，首先透過諮詢專業人士，針對創傷知情做相關的心理問題，蒐集相關資料以設計軟布書的內容。其中 AR 技術結合互動式軟布書的設計在引導使用者面對心理需求和問題，同時期望透過這樣的教材觸發內在的自癒力。由於軟布書對不同年齡層的兒童可能產生不同的影響，因此主要研究對象將以具有識字能力的兒童為基本要求，並評估其對軟布書內容的共鳴程度，以獲得使用者的心得回饋與深入討論。整個研究的步驟流程為下圖 1。

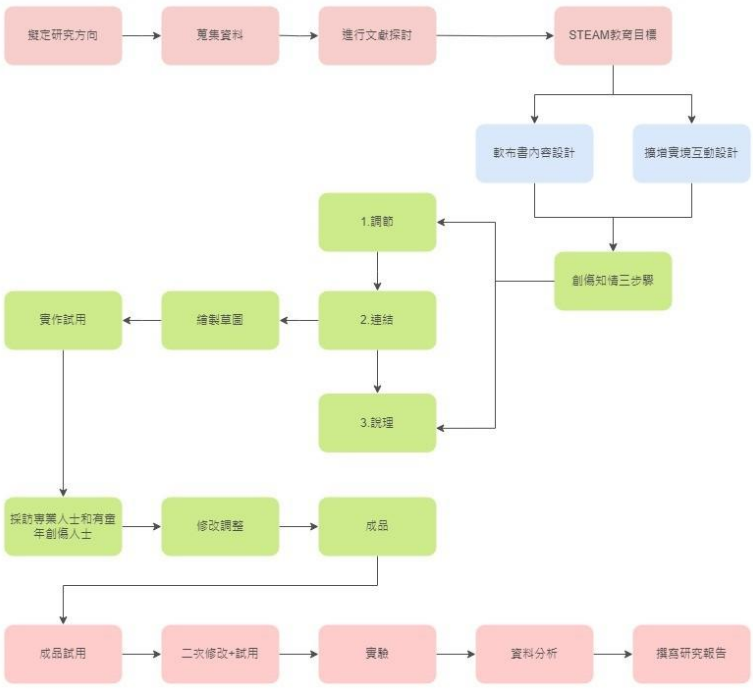


圖 1 步驟流程圖（本研究自行繪製）

本研究實作兒童 STEAM 教育結合 AR 互動式軟布書教材，軟布書的材質為秀士布，觸感滑順細膩，同時具有高度的韌性，容易清潔保養，印刷的顏色效果鮮明飽和，且具有無毒特性，並以柔軟的布書形式結合互動式配件，以實踐 STEAM 教育的理念。透過融合搭配，教材將在互動性與創造性上取得平衡，同時採用 AR 技術，使教具更豐富有趣。整體教材包

含三本軟布書，而軟布書的結構主要分為調節、連結再說理三部分。此設計能夠引導兒童在探索知識的過程中，藉由實際操作和互動來深化對 STEAM 領域的理解，如下圖 2 及圖 3。

圖 2 對著有 QR 碼的圖片做掃描

圖 3 出現 AR 互動動畫

4.結論與未來展望

本研究實作設計一款互動式軟布書，並提供給諮商心理師與輔導教師作為輔助工具，以達到推廣與實際應用的效果。透過專業領域的指導與結合，本研究所發展的軟布書能幫助使用者更深入理解創傷知情，提升對童年創傷的認識與關注。本作品融合 擴增實境技術，讓使用者能以輕鬆且療癒的方式獲得心理安慰與情緒支持，並透過直觀的互動體驗來理解抽象的情感。這不僅有助於情緒表達與自我認知，也提供了兼具趣味與療癒性的學習體驗。研究將進一步針對 STEAM 互動軟體書進行使用性評估，確保其設計符合不同年齡層的學習需求與操作習慣。同時，期望為對創傷知情相關議題較陌生的讀者提供學習途徑，讓更多人了解童年創傷對個人及社會的影響，進而減少其負面影響。透過本研究的成果，我們希望能推動心理健康教育的普及，提升社會對創傷知情的關注，並營造更友善且支持性的成長環境，幫助更多人獲得理解與療癒。

參考文獻

朱崇信（2021）。瞭望童心 1 - 了解創傷在大腦的痕跡::其他權益/議題:兒少權益網。取自：https://www.cylaw.org.tw/about/advocacy/11/595。

留佩萱（2019）。人本教育札記 365 期。是「上層腦」還是「下層腦」行為？取自：https://hef.org.tw/journal365-2/。

曾美鳳、陳信昭（2018）。驚魂未定之後──認識創傷後壓力症候群(兒童篇)。取自：https://psy-med.ncku.edu.tw/p/412-1128-16298.php?Lang=zh-tw。

周淑惠（2018）。具 STEM 精神之幼兒探究課程紀實：「一起創建遊戲樂園」主題。新北市：心理出版社。

陳雅雪（2009）。從依附關係淺談幼兒人際關係。家庭教育雙月刊，20，38-44。

藍逸梅、李從業、陳嘉琦（2001）。概念分析-親子依附。榮總護理，18(2)，125-131。

Wilson, C., Pence, D. M., & Conradi, L. (2013). Trauma-informed care. In *Encyclopedia of social work*.

針對完整主題的系列教具和新的論述

A Series of Teaching Aids and New Discourse on the Complete Topic

周鑑恒[1*]，曾瑞蓮[2]
[1*]萬能科技大學航空光機電系
[2]高雄科學工藝博物館
[*]chou0717@gmail.com

【摘要】 針對在教科書中或在常見科教活動中有瑕疵的某些重要主題，例如：光學成像、發電、熱機、風力發電、海浪發電、熱力學、流體力學等等主題(甚至微積分)，重新建構更完整、更實用的理論架構，開發一系列教具，以實驗闡釋此完整主題的各關鍵要點，並撰寫專書組織此完整主題並重實驗和理論的教材和教法，形成所謂「教學資源」。利用這些教學資源，使學習者能輕鬆但清晰地掌握整個主題，而能將所學到的知識應用於解決實際問題、或作進一步研究。本文僅以跨領域的「熱機教學資源」為例，說明教學資源的特點和內容。
【關鍵字】 教學資源；系列教具；實用知識；完整主題

Abstract: *Reconstructing a more complete and practical theoretical framework for certain important topics that are flawed in textbooks or common science education activities—such as optical imaging, power generation, heat engines, wind power, wave power, thermodynamics, fluid mechanics, and even calculus—developing a series of teaching aids to illustrate key points of each topic through experiments, and writing a specialized book that organizes the complete subject with an emphasis on both experiments and theory, to form what is called "teaching resources." By using these teaching resources, learners can easily yet clearly grasp the entire subject and apply the knowledge they have acquired to solve real-world problems or conduct further research. This article uses the interdisciplinary "teaching resources for heat engines" as an example to illustrate the features and content of such teaching resources.*
Keywords: Teaching resources, A series of teaching aids, Practical knowledge, Comprehensive topic

1.前言

大家都不太相信：教科書有不小的改進空間!教科書的瑕疵，常造成學習者難以了解某項主題。例如：在諸多教科書中發電只是法拉第定律的一個例題，這使得學習者誤以為只有法拉第定律涉及發電，而其他原理與發電的關係則可以忽略。又例如：在談到卡諾循環時，教科書並未強調所有的熱機(引擎或發動機)的原理都是一樣的。又例如：談到流體力學時，教科書常忽略了流體力學也是牛頓力學的一部分，絕對要完全遵守牛頓力學；並且還濫用伯努利方程式(Bernoulli's Equation)，台灣翻成伯努力「定律」，更容易誤導學習者。至於許多科教活動中，依循教科書的內容，也有不正確的迷思。

相較於單獨零星的科學內容，了解完整主題的所有科學原理，才方便應用此完整主題的科學原理，來解決實際問題。也有利於進一步學習更深入的科學。

我們創作的「系列教具」之功用在於，將整個完整主題的各個關鍵要點，用實驗加以闡釋，使學習者以眼見為憑的直觀方式，了解整個完整主題。即使一時倉促，來不及清楚了解所有內容，也能先將關鍵點演示出來，作為進一步學習的基礎。

更重要的是，創作直接了當、不多不少的完整主題概念圖和論述方式，這些新的概念圖如同更好的導覽圖，讓學生先了解全盤大概，再胸有成竹地探究細節，以達到深入淺出、以簡御繁的學習成效。

而統整的書籍則畫龍點睛、有系統地將系列教具和更新的理論架構組織起來，形成非常有用的教學資源。

限於篇幅，顯然無法分享上述各主題之所有教學資源，在本文中僅以「熱機」主題為例，說明教學資源的內容和特點。

2.熱機的教學資源

在許多教科書中熱力學章節中，都會談到卡諾引擎(Carnot Engine)，有關卡諾引擎的計算，最終推導出熵的概念($\mathrm{d}S = \mathrm{d}Q\ /\ T$)但教科書大多沒有強調「所有熱機的原理基本上是一致的」，使得學習者失去完整掌握熱機原理的機會，十分可惜。

新的論述如下：所有熱機的原理都是一樣的。主角都是工作流體，工作流體作為能量轉換的媒介，將人們所擁有的能源轉換成人們所需要之動力。在所有的引擎中工作流體都必須完成熱機循環，才能完成上述的任務。

所謂熱機循環就是工作流體低溫時被壓縮，壓縮之後被加熱，增加工作流體之內能，再之後工作流體消耗它的內能膨脹，對“某些物體”做功，最後工作流體降溫被壓縮，回到原始的狀態，於是就能將“熱”轉換成“功”。

圖 1 加熱透明希侖蒸汽機中的工作流體，使工作流體沸騰，並且(產生蒸氣)而劇烈膨脹，驅動希侖蒸汽機轉動，使得學習者可以完全了解：希侖蒸汽機中工作流體到底如何做功。

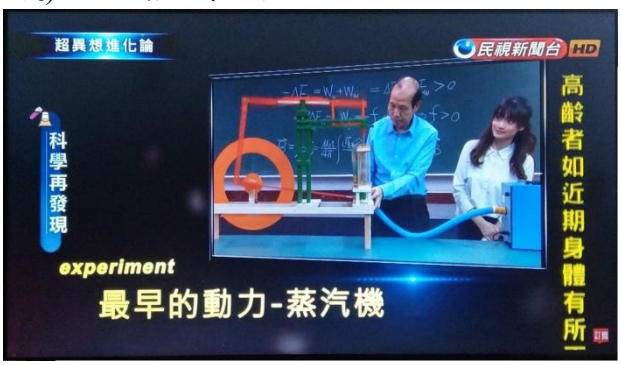
所謂不同之引擎是指：使用不同的工作流體、工作流體被使用的方式不同，用不同的設備和裝置，使工作流體完成熱機循環，並對不同的物體做功。

工作流體的 PV 圖是一封閉的曲線，中間的區域，代表工作流體膨脹並消耗內能所做的功，但並未說明工作流體在何處、對什麼東西做功?這是有趣之處!

系列教具的闡釋策略是：因為熱機教學資源的目的在於強調：所有的熱機基於相同的原理，熱機中的工作流體，會經過相同的熱機循環，將能源轉換成我們所須的動力。所以熱機教學資源的系列教具的設計策略是：研製各異其趣的熱機，並且引導學習者了解：各種熱機中的工作流體為何？工作流體又是如何被壓縮、被加熱？所使用的能源為何？……等等。在欣賞各個熱機設計的巧妙各有不同，別出心裁之餘，同時會發現所有熱機中的工作流體都經過非常相似的熱機循環。藉此達到跨物理、航空與機械等領域的教學目的。

首先創作一項透明的希侖蒸汽機，利用磁鐵吸住希侖蒸汽機上方尖錐狀的轉軸，利用重力與磁力相互抵銷，以致於尖錐狀的轉軸所受到的摩擦力矩極小，所以可以使透明玻璃球中的工作流體(也就是水，用過即拋)在沸騰時的蒸氣壓力並不大的狀況下，就足以產生能夠推動希侖蒸汽機轉動的蒸氣(見圖 1)。如此一來，學習者就可以完全了解：在透明的希侖蒸汽機中，工作流體被加熱並產生大量蒸氣，進而同時對工作流體本身以及透明的球狀玻璃容器做功，使得玻璃容器旋轉。

緊接著藉用吸塵器或鼓風機所造成的氣流，驅動各種古老的往復式蒸汽機(見圖 2)。因為通體透明，所以學習者可以看到設計蒸汽機各種巧思，從而了解蒸汽機的運轉原理。雖然是以吸塵器或者是鼓風機產生氣流來推動各種古老的往復式蒸汽機，但可以讓學習者了解，真實的蒸汽機中沸騰的水是如何推動活塞，使蒸汽機運轉。仔細推敲會發現作為工作流體的水(水蒸氣)也完成熱機循環。



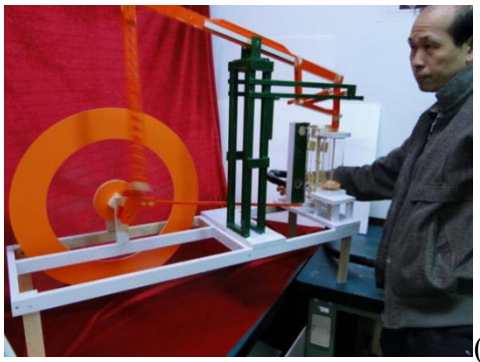
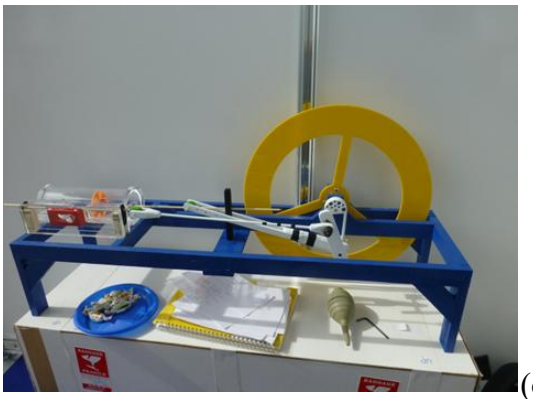
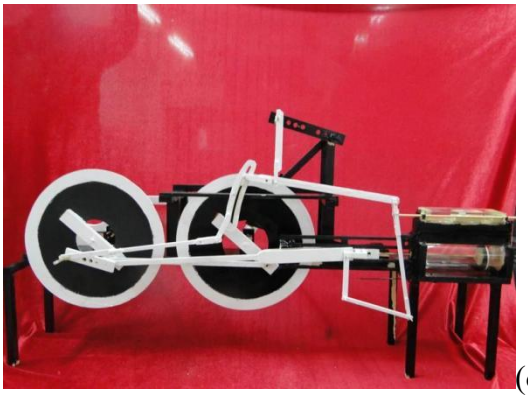
(a)

(b)

(c)

(d)

圖 2 完全透明的古老蒸汽機模型(a 與 b)。晚期較為先進的往復式蒸汽機(c)，這種蒸汽機最後裝設在蒸汽火車頭上，引領鐵路交通運輸近一百年(d)。

接著設計兩種用真的蒸氣驅動的蒸汽機模型(見圖 3)，這兩種蒸汽機模型也都是通體透明，所以學習者可以完全看到整個蒸汽機內部構造以及所有運動的零件。進而理解：的確是以透明三角錐瓶中的工作流體(水)作為工作流體，加熱使工作流體沸騰，變成蒸氣、體積膨脹，而對活塞做功。這樣就可以充分說明工作流體是各種引擎中的主角，用來進行能量轉換。

(a)

(b)

圖 3 兩具用真的蒸氣驅動的透明蒸汽機模型。

史特林引擎是一種奇特的外燃引擎(見圖 4)，可以使用多元的能源來驅動，所以也稱得上是一種綠色能源科技。史特林引擎中的工作流體，被壓縮、被加熱、膨脹、做功的過程，與其他發動機中的工作流體所經過的過程非常相似。但是，使工作流體進行熱機循環的方法各

有巧妙。所以演示史特林引擎可以強調，在不同的引擎中使用不同的方法和裝置，使工作流體經過類似的熱機循環。

教具DIY 周鑑恆動手教物理

(a)

(b)

圖 4 作者早期製作許多不同的史特林引擎(a)，有些組裝式史特林引擎甚至流行到國外(b)。

熱機主要是利用其工作流體進行能量轉換，將我們所擁有的能源，轉換成我們所需要的動力。太陽能加熱的史特林引擎，就可以用來特別強調這一點。圖 5a 跟圖 5b 所示，即為太陽能史特林引擎發電機。用反光碟反射陽光，加熱使特林引擎的溫差氣室，以維持溫差氣室上下兩面的溫度差即可運轉史特林引擎。

也可以利用菲涅爾透鏡(Fresnel lens)聚焦陽光，同樣能使史特林引擎轉發電機發電。

(a)

(b)

圖 5 簡單設計的太陽熱能史特林引擎，研製太陽能史特林引擎發電的經歷，曾經作為台灣科學月刊的封面故事(a)。用透明加熱的史特林引擎(b)。

史特林引擎也可以利用地熱來加熱，同樣可以產生動力。作者在台灣台北的北投地熱谷，曾經研製全台第一具示範史特林引擎地熱發電機。這說明熱機可以使用不同的能源加熱(見圖 6)。

(a)

(b)

圖 6 在台北北投地熱谷示範的史特林引擎地熱發電機(a)，與環境的景觀融合得非常和諧。這故事曾經在台灣的探索頻道(Discovery Channel)播出(b)。

為了吸引大眾與學習者的學習興趣，強調引擎輸出的動力可以用於發電或驅動車輛，作者特別研製了一輛史特林引擎載人車。用特製史特林引擎所產生的有限動力，驅動一輛木質的車輛，可以載重約 60 公斤以上的人。雖然行進的速度不快，但是，因為構造非常簡單(見圖 7)，頗能引起大眾的注意與好奇。

圖 7 史特林引擎載人車的特寫鏡頭，實際行駛的情形(右)。

利用共振的原理，構造已經十分簡單的史特林引擎，甚至可以進一步簡化，設計出所謂的自由活塞型史特林引擎(見圖 8)。美國太空總署，曾經長時間測試這類自由活塞型史特林引擎之可靠性，準備作為深太空探測的發電動力。

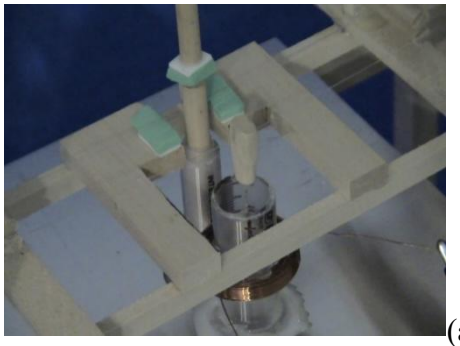
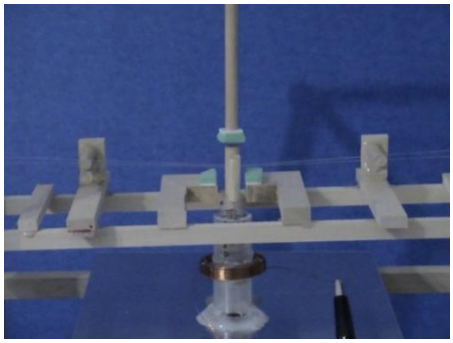

(a)

(b)

圖 8 利用共振原理研製出的自由活塞型史特林引擎(a)。其中移氣器(Displacer)以及活塞都利用彈力裝置，使兩者的震動頻率十分接近，而產生共振的效果，使得活塞以及移氣器的震動頻率相同，但又有必要的相位差。

嚴格來說，令人驚奇的熱聲效應(thermoacoustic)也是一種引擎的現象。乍看之下，令人匪夷所思，但仔細解釋則的確如此。在熱聲效應管中，只有一小部分的空氣能成為工作流體，會進行熱機循環，而能夠增強管中氣柱的基頻震動模式(fundamental normal mode)，所以可以發出很大的聲音，這種聲音甚至可以在管口造成氣流(見圖 9b)。其中熱聲堆的一端維持高溫，另外一端維持低溫(見圖 9a)，當空氣振動時，管中距管口適當距離的處一小段空氣，就可以被先壓縮再加熱，進行熱機循環，而能夠把熱能轉換成氣柱震動的波動能量。

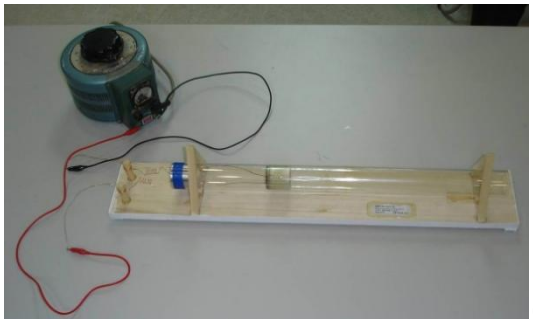
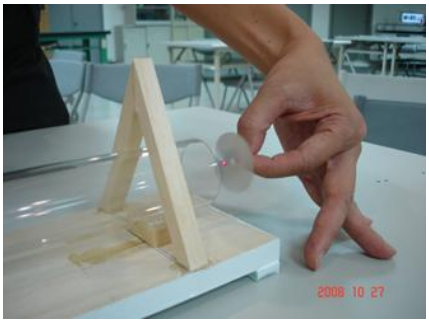

(a)

(b)

圖 9 整具熱聲管的特寫(a)，劇烈的震動會在管口引起氣流(b)。

有一種更為簡單的熱聲管：黎開管(Rijke tube)，管中只有一片被加熱至高溫的金屬網，金屬網距管口約 1/6 處，如果高溫的金屬網在管子下方，管子又垂直地面而引起微弱的熱對流，黎開管因熱聲效應(也是一種引擎的現象)，就會發出很大的聲音。

兩具長度不一樣的黎開管，甚至可以用來進行拍頻實驗。其實拍頻實驗，因為不容易控制兩具聲源的同調性，並不容易以現有的儀器進行。但是，以黎開管進行，拍頻的干涉現象非常顯著(見圖 10)。

圖 10 兩具長度不一樣的熱聲管所發出的聲音，因為干涉而會造成拍頻現象。

航空引擎是一種引人入勝的熱機，也是航空工程的核心領域，也遵守相同的熱機原理，其中的工作流體也經過熱機循環。雖然要製作渦噴(turbojet)或渦扇(turbofan)發動機並不容易，但作者研製一具渦軸發動機(見圖 11)，已能概略說明航空發動機的基本原理。仔細研究這具渦軸發動機(turboshaft)會發現，航空發動機的工作流體的確也進行所謂的熱機循環；特別是：航空發動機中的工作流體(空氣)，同時會對發動機本身以及工作流體做功。這個就會涉及到所謂的推進效率(Propulsive efficiency)等等問題，十分有趣。在實驗的過程當中，也可以讓學習者確認，發動機的工作流體的工作原理的確都非常相似，但是航空發動機之工作流體做功的對象，則更為複雜、耐人尋味。

圖 11 這具類似渦軸發動機的教具，當空氣被加熱後，先膨脹再因熱對流向上流動，驅動上方之渦輪轉動，與加熱時劇烈膨脹即高速流動的真實渦軸發動機有些差異。

火箭發動機也是熱機的一種，在發動機啟動後，它的工作流體才源源不斷地產生，並同時被加熱。作者為了以簡易的方式操作火箭連續發射，營造熱烈氣氛，將 1500cc 寶特瓶瓶蓋上鑽 7.2mm 的孔作為火箭，特別以酒精或甲醇作為燃料，根據氣溫可以控制裝在寶特瓶中與空氣混合的燃料之較佳比例。再用電極之電弧點火，利用電路之巧妙設計，一次可快速連續發射二、三十個酒精火箭，場面十分有趣(見圖 12)。

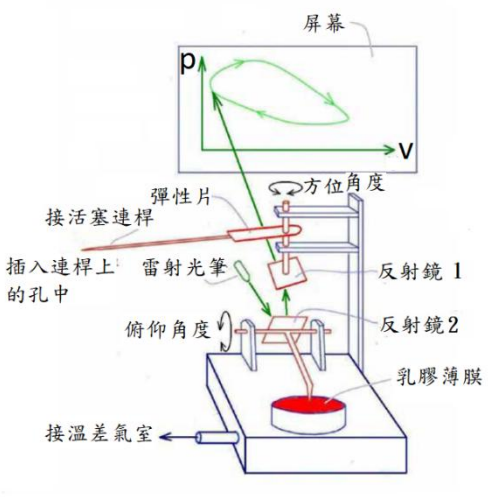
(a)

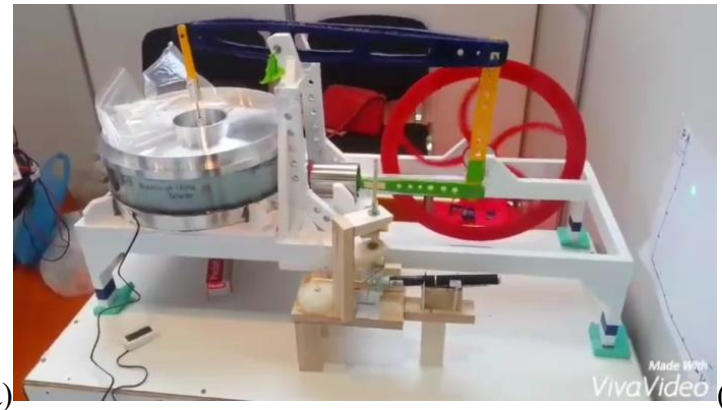
(b)

圖 12 並列在發射架上的多個醇類火箭(a)。多個醇類火箭發射的情形(b)。

因為“所有”熱機中的工作流體，都經過類似的熱機循環，所以工作流體的壓力與體積變化的過程，在壓力與體積圖中，會是一個順時間變化的封閉路徑，作者特別製作一具史特林引擎，能夠演示這項事實。

見圖 13(a)，用乳膠薄膜顯示引擎中工作流體之氣壓，並利用活塞連桿顯示活塞的位置，以便標示出工作流體的體積。再利用兩片被薄膜和活塞位置連動的鏡片，反射雷射光筆之光束，即能以屏幕上的光點顯示工作流體的 PV 圖，從圖上可以看出工作流體的 PV 圖的確為順時針的封閉曲線。



(a)


(b)

圖 13 可以目視、直觀體會的工作流體 PV 圖演示儀。

3.討論與推廣

針對值得改進的不同重要主題，進行改進與創新，形成重新整理的理論論述和配合的系列教具，所生成的教學資源，因主題不同，各「教學資源」具有不同的特色(詳見參考資料)。但基本上都是先建立更為簡潔、明晰、直接了當的原理論述體系，再研發針對原理中難點和要點的系列實驗和教具，並且使學習者能更直觀、更具象地了解完整的原理。

熱機教學資源所強調的理論論述，重點在於，所有熱機都基於相同的原理，工作流體是其中的關鍵角色。而系列教具和實驗則包括了各異其趣、甚至匪夷所思的各種熱機。在分析了各自獨特的巧思和創意之後，再歸納出所有發動機的工作流體都經過極其相似的熱機循環，突顯出以下幾項重點：(1)工作流體可能是水、空氣、二氧化碳、氫氣、氦氣或其他會發生相變的液體或氣體。(2)使用工作流體的方式可能是(a)用過即拋或(b)反覆循環使用同一工作流體。(3)加熱工作流體的方法可以是外燃或內燃。(4)壓縮工作流體的方法可能用活塞、高速旋轉的壓縮機葉片、震波(shock wave)等等。(5)工作流體消耗內能而膨脹，可能對不同的物體做功。

在組織此系列教具和理論論述的書籍(見文獻)中，更順便回答了下列有趣的問題：為什麼近代的民航發動機是大涵道比(旁通比)的渦扇發動機?為什麼渦噴發動機燃燒室的火焰不會向前噴射(因為壓力沒有增加)?為什麼壓力沒有增加，卻會受到推力?為什麼極音速飛行器要使用超(音速)燃(燒)衝壓發動機?

4.參考文獻

Benson, H. (1996). *University physics: Revised edition*. Wiley.

Eastop, T. D., & McConkey, A. (1993). *Applied thermodynamics for engineering technologists* (5th ed.). Longman Scientific & Technical.

Çengel, Y. A., & Boles, M. A. (2019). *Thermodynamics: An engineering approach* (9th ed.). McGraw-Hill Education.

PuppyCodeReview: An AI-based Code Review System

Chun-Hsiung Tseng [1*], Hao-Chiang Koong Lin[2], Chih-Wei Huang[3], Kai-Chun Hou[4], Jia-Rou Lin[1]

[1] Dept. of Electrical Engineering YuanZe Univ. Taoyuan, R.O.C.

[2] Dept. of Information and Learning Technology National Univ. of Tainan Tainan, R.O.C.

[3] Dept. of Psychology Fo Guang Univ. Yilan, R.O.C.

[4]Dept. of Digital Technology Design, National Taipei University, Taipei, R.O.C.

[*] lendle@saturn.yzu.edu.tw

Abstract: *This study proposes an artificial intelligence (AI)-based code review system, PuppyCodeReview, designed to streamline the code review process. While traditional peer code review is gaining importance in programming education, its effectiveness in university courses remains under question. PuppyCodeReview leverages AI technology to automate code reviews, providing suggestions across multiple aspects, including design, functionality, complexity, testing, naming, and commenting. The system also detects code smells, such as code duplication and long methods.*

Keywords: CodeReview, AI, Programming

1. Background

In emerging programming education methods, peer code review (Code Review / Peer Review[7]) has gained significant attention in recent years. Originally used as an agile development practice in the software industry, peer code review involves two distinct roles in the software development process: the coder (who writes the code) and the reviewer (who reviews the code). These roles can overlap or be interchanged. The primary objective of peer code review is to identify and rectify errors that were not detected during the initial development phase, thereby maintaining code quality. Many well-known international software companies, such as Google, implement internal peer code review practices to reduce coding errors[8]. Kubincová and Csicsolová attempted to introduce peer code review into high school programming courses. Their study found that 42% of students reported discovering more previously unnoticed coding errors, while 28% of students became more willing to spend time writing or modifying code compared to before (Kubincova & Csicsolova, 2018). However, Chong et al. have questioned the effectiveness of applying peer code review in university programming courses. They argue that due to university students' limited experience in software development, their ability to accurately identify coding errors remains highly uncertain (Chong et al., 2021). To remedy this, we proposed PuppyCodeReview, which is a AI-based code review system, to simplify the code review process.

2. Design and Implementation

[7] https://en.wikipedia.org/wiki/Code_review

[8] https://google.github.io/eng-practices/review/

The current implementation uses the Qwen2.5-Coder model to review submitted codes. The prompt below is used (some details are skipped to save space):

Please review the student's ${language} code. The purpose of this program is: ${objective}. First, let me know if the code is correct. Then, based on the provided guidelines, give review suggestions. Please respond in JSON format using the following structure:
... (output format omitted) ...Next, the student's code is: ${studentCode}.
The reference solution is: ${answer}. The review should consider the following aspects: Design, Functionality, Complexity, Tests, Naming, and Comments.Additionally, avoid the following code smells: Code Duplication, Long Methods, God Class, Too Many Parameters, Feature Envy, Inappropriate Intimacy, Refused Bequest, Lazy Class, Artificial Complexity, Overly Long Identifiers, Overly Short Identifiers, and Excessive Use of Literals.

The figure below demonstrates the output after reviewing a student's work:

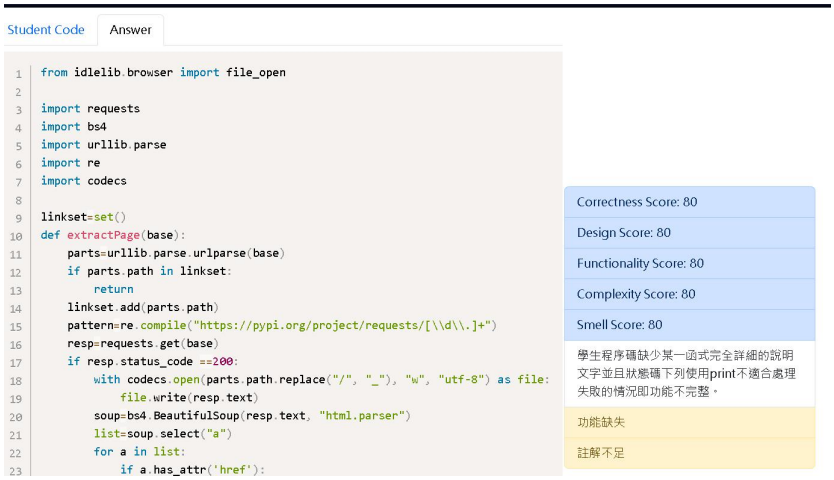


Figure 2 System Output

As shown in the figure, the student's code and the correct answer will be shown on the left hand side. On the right hand side, the score of each measurement will be shown and the system will generate some keywords for the code.

References

Chong, C. Y., Thongtanunam, P., & Tantithamthavorn, C. (2021). Assessing the Students' Understanding and their Mistakes in Code Review Checklists: An Experience Report of 1,791 Code Review Checklist Questions from 394 Students. *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*, 20 – 29. 2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET). https://doi.org/10.1109/ICSE-SEET52601.2021.00011

Kubincova, Z., & Csicsolova, I. (2018). Code Review in High School Programming. *2018 17th International Conference on Information Technology Based Higher Education and Training*

(ITHET), 1 – 4. 2018 17th International Conference on Information Technology Based Higher Education and Training (ITHET). https://doi.org/10.1109/ITHET.2018.8424617

沉浸式虛擬實境學習導向教育元宇宙平台開發：從設計思維到教學應用

Development of a Metaverse Platform for Immersive Virtual Learning-Oriented Education: From Design Thinking to Teaching and Learning Applications

王政弘[1*]，吳聲毅[2]，盧俐雯[3]，林知儀[4]
[1]高雄大學工藝與創意設計學系
[2]清華大學師資培育中心
[3,4]高雄大學數位內容設計研究中心
[*] wang101@nuk.edu.tw

【摘要】 COVID-19 疫情加速了遠距教學的普及，卻同時揭示現行數位學習環境在臨場感營造、互動參與與學習引導上的侷限。隨著元宇宙技術快速發展，教育領域積極探索其於沉浸式學習應用的可能性，期望結合虛擬實境（VR）與人工智慧（AI）等新興科技，建構具沉浸體驗與社會互動性的數位學習場域。本研究回應教師在教學現場的實務需求，整合設計思維與學習理論，開發出以 EWOVA（Engagement, Wondering, Organization, Validation, Application）學習行為歷程為核心的教育元宇宙平台。透過跨校教學實踐案例，說明 EWOVA 模式如何促進學生從動機引發到應用遷移的完整學習歷程。本研究亦提出具模組化與行為導向特色的平台建構架構，期能在科技發展下，邁向更具靈活性、自主性與素養導向應用的教育元宇宙平台。
【關鍵字】 教育元宇宙；EWOVA；沉浸式學習；教育創新；素養教育

Abstract: *The COVID-19 epidemic has accelerated the popularization of distance learning, but at the same time reveals the limitations of the current digital learning environment in terms of sense of presence creation, interactive participation and learning guidance. With the rapid development of meta-universe technology, the education field is actively exploring the possibility of its application in immersive learning, in the hope of combining emerging technologies such as Virtual Reality (VR) and Artificial Intelligence (AI), to construct a digital learning field with immersive experience and social interactivity. In this study, we responded to the practical needs of teachers in the teaching field, integrated design thinking and learning theories, and developed an educational meta-universe platform based on the EWOVA (Engagement, Wondering, Organization, Validation, Application) learning behavioral process. Based on the case studies, we illustrate how the EWOVA model facilitates the complete learning process from motivation to application. This study also proposes a modular and behavior-oriented platform architecture as a reference blueprint for the future development of the educational meta-universe, in the hope of leading digital teaching towards a new paradigm of greater flexibility, autonomy, and literacy orientation.*
Keywords: Metaverse, EWOVA, Immersive Learning, Educational Innovation, Literacy Education

1.前言

後疫情時代使遠距與混成教學成為全球教育常態，進一步突顯數位學習在臨場感、互動性與個人化引導方面的不足。城鄉教育資源落差、師資短缺與課程抽象化等結構性問題影響

學生的學習動機與成效。雖然多種線上學習平台迅速導入校園，但大多仍侷限於教材播放與視訊會議，難以提供具沉浸感與深度互動的學習經驗。

沉浸式學習（Immersive Learning）融合情境體驗、互動操作與角色扮演，能突破傳統教室限制，有助於抽象概念理解、高階思維訓練與社會情境實踐（Makransky & Petersen, 2021；Pagano, 2013；Parong & Mayer, 2018）。在此脈絡下，元宇宙（Metaverse）逐漸成為推動數位學習創新的新引擎，轉化學習場域為 3D 互動空間，發展出具多角色、跨時空社群互動特質的教育元宇宙（Lee et al., 2021）。

教育元宇宙結合虛擬實境（VR）、人工智慧（AI）、物聯網（IoT）等技術，能建構可共學、可評量、可適性化調整的學習空間，回應教學品質與教育公平的雙重訴求（Wang et al., 2022；Teng et al., 2022）。但有研究指出，沉浸感不必然提升學習效果，過度設計可能造成認知負荷（Makransky et al., 2019）。因此，沉浸式學習環境的設計需在技術呈現與教學目標間取得平衡（Wang，2025）。本研究聚焦於教育元宇宙在推動沉浸式學習轉型的潛力，並以 EWova 平台系統為實作範例，探討其設計邏輯與應用實踐對於教育現場問題的回應與貢獻。

2.文獻綜述

「情境學習」（Situated Learning）為沉浸式學習設計的重要理論基礎，主張知識應於真實或擬真的社會情境中建構，透過參與與實作產生意義（Brown et al., 1989；Lave & Wenger, 1991）。元宇宙作為具多模態與高度互動特性的虛擬場域，能有效促進社會化參與與情境互動（Sedlák et al., 2022），為實踐情境學習提供了新的可能。Dalgarno 與 Lee（2010）提出 3D 虛擬環境的五項學習可供性，包括空間導航、互動操作、情境任務、協作支援與概念視覺化，為沉浸式教學設計提供具體架構。

在 K-12 教育中，沉浸式學習已展現提升學習成效的潛力。Ng et al.（2023）指出，遊戲化與建構式任務有助於發展高階認知與團隊合作。Wang（2024）實作自然科 VR 教材，證實沉浸式學習有助學生理解抽象概念並提升主動探索行為。其他研究亦指出 VR 可減緩學習焦慮、促進知識建構（Concannon et al., 2020；Hu-Au & Okita, 2021），並提升參與與反思表現（Pirker et al., 2020；Holly et al., 2021）。

在學習行為設計層面，多項實證教學策略亦提供關鍵依據。PBL（Hmelo-Silver, 2004）促進學生於真實問題中建構知識，強調統整與應用；合作學習（Johnson & Johnson, 1999）則透過小組互動與任務分工增進參與與社會學習；POE 策略（White & Gunstone, 1992）聚焦科學探究與反思修正；SRL 理論（Zimmerman, 2002）強調自我監控與策略調整，以上策略皆對沉浸式學習中的歷程設計提供實踐啟示。在技術發展方面，元宇宙學習場域整合 VR、AI、5G 等技術，打造跨平台沉浸體驗（Andembubtob et al., 2023；Lee et al., 2021）。Wang et al.（2022）提出教育元宇宙生態系，強調數據驅動與即時互動；Teng et al.（2022）則以 UTAUT 模型驗證「績效期望」、「社會影響」與「沉浸感知」對學習採用意圖的顯著影響。

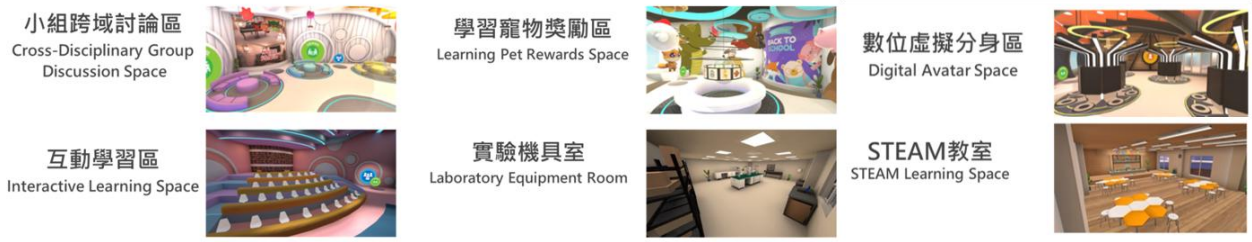
綜上所述，情境理論、教學策略與技術整合的多元觀點，共同構成本研究平台設計與學習行為模式（EWOVA）的理論基礎與發展脈絡。

3.EWova 教育元宇宙平台建置

本研究平台以雛形法（Prototyping Method）進行開發，採用 Unity 開發引擎，採立意取樣法針對臺灣具科技融入教學經驗的中小學與高中教師進行質性調查，共蒐集 205 份回饋。

研究運用內容分析法，歸納教師在教學應用與技術功能兩大面向的實際需求。教學層面顯示教師關注學生參與度不足、學習動機低落、缺乏反思與知識整合機制，以及學習成果難以遷移；技術層面則期待平台能支援歷程紀錄、協作互動、即時輔助與跨任務整合，以強化素養導向課程的教學支援。

為回應教學現場需求，本研究建構 EWOVA 學習行為模式作為平台設計核心，涵蓋參與（Engagement）、探究（Wondering）、組織（Organization）、驗證（Validation）與應用（Application）五大歷程，對應學生從情境導入到知識遷移的完整學習路徑。平台依據各階段行為建置對應模組功能，如「參與」整合課程管理、身分驗證與虛擬分身模組以營造情境；「探究」與「組織」階段結合 AI 輔助與歷程紀錄工具強化問題探索與概念建構；「驗證」與「應用」則透過互動模組、獎勵管理模組，支持歷程反思與成果應用。整體設計強調模組化、視覺化與彈性應用，並將模組功能置入相對應空間：小組討論區、互動學習區、數位分身區、學習獎勵區與實驗教室，作為沉浸式學習與教學實踐的重要場域。



圖一 EWOVA 教育元宇宙平台場域空間

4.教學應用成效

為檢證教育元宇宙平台在國小實際教學場域中對學習歷程與行為培養的成效，本研究選取臺灣南部三所國小進行應用觀察與分析，課程以社會領域結合在地文化議題，強化學生對生活場域的理解與參與意願。整體設計依據 EWOVA 行為歷程進行規劃，並結合平台核心模組，驗證其在沉浸式數位教學中的可行性與教學輔助效益。

三校採遠距同步共學形式進行課程，由教師透過 XR 直播於互動講台導入地方文化主題，提升學生沉浸參與感（Engagement）。學生分組進入文化教材場域，操作 3D 模型與虛擬互動工具進行議題探究與推論（Wondering）；在任務引導下記錄資料、運用模型召喚功能協助資料整理與概念統整（Organization）。小組分享時完成簡報發表與互動提問，藉由同儕互評與教師引導進行歷程反思與策略調整（Validation），最終於互動學習區提出應用構想，深化遷移與實踐能力（Application）。上述學習行為，教師亦可透過平台課程管理與歷程記錄模組追蹤學生表現與學習行為，展現 EWOVA 模式在素養導向教學中兼具彈性與實用性，能有效支援學生於各歷程中逐步建構與應用知識。

圖二 實際教學應用情形

5.結論

本研究整合設計思維、理論基礎與教師實務回饋，建構一套以學習行為為核心、設計驅動的教育元宇宙發展模式。EWOVA 模式將沉浸式學習歷程拆解為參與（Engagement）、探究（Wondering）、組織（Organization）、驗證（Validation）、應用（Application）五大行為環節，具體對應學生的學習動態與階段需求。此模式除提升學習參與與歷程反思，也提供教師課程設計與學習評估的可操作參照架構。實務應用顯示，EWOVA 能有效支援 PBL、POE、SRL 與合作學習等教學策略，實踐以學習者為中心的教學理念，展現高度彈性與擴展性。本研究亦提出 EWOVA 為數位學習場域中「行為驅動 × 技術支援 × 教學實踐」三維整合架構，彌補傳統元宇宙應用過度聚焦技術展示之不足。未來將持續深化歷程可視化、個別化支持與開放式學習空間，擘劃具實踐價值的教育元宇宙發展路徑。

參考文獻

Andembubtob, D. R., Keikhosrokiani, P., & Abdullah, N. L. (2023). A review on virtual reality and eduverse: History and technical perspectives. *Handbook of research on consumer behavioral analytics in metaverse and the adoption of a virtual world*, 89-111.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *1989*, *18*(1), 32-42.

Concannon, B. J., Esmail, S., & Roberts, M. R. (2020). Immersive virtual reality for the reduction of state anxiety in clinical interview exams: Prospective cohort study. *JMIR Serious Games*, *8*(3), e18313.

Dalgarno, B., & Lee, M. J. (2010). What are the learning affordances of 3-D virtual environments?. *British journal of educational technology*, *41*(1), 10-32.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn?. *Educational psychology review*, *16*, 235-266.

Holly, M., Pirker, J., Resch, S., Brettschuh, S., & Gütl, C. (2021). Designing VR experiences–expectations for teaching and learning in VR. *Educational Technology & Society*, *24*(2), 107-119.

Hu-Au, E., & Okita, S. (2021). Exploring differences in student learning and behavior between real-life and virtual reality chemistry laboratories. *Journal of Science Education and Technology*, *30*(6), 862-876.

Johnson, D. W., & Johnson, R. T. (1987). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Prentice-Hall, Inc.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Lee, L. H., Braud, T., Zhou, P. Y., Wang, L., Xu, D., Lin, Z., ... & Hui, P. (2024). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. *Foundations and trends® in human-computer interaction*, *18*(2–3), 100-337.

Makransky, G., & Petersen, G. B. (2023). The theory of immersive collaborative learning (TICOL). *Educational Psychology Review*, 35(4), 103.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction*, *60*, 225-236.

Manocha, R. (2024). Eduverse: Changing Paradigms in Fashion Education and Management. *Journal of Operations and Strategic Planning*, *7*(1), 80-104.

Ng, P. H., Chen, P. Q., Sin, Z. P., Jia, Y., Li, R. C., Baciu, G., ... & Li, Q. (2023, November). From classroom to metaverse: A study on gamified constructivist teaching in higher education. In *International conference on web-based learning* (pp. 92-106). Singapore: Springer Nature Singapore.

Pagano, K. O. (2013). *Immersive learning: Designing for authentic practice*. Association for Talent Development.

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785.

Pirker, J., Lesjak, I., Kopf, J., Kainz, A., & Dini, A. (2020). Immersive learning in real VR. In *Real VR–Immersive Digital Reality: How to import the real world into head-mounted immersive displays* (pp. 321-336). Cham: Springer International Publishing.

Sedlák, M., Šašinka, Č., Stachoň, Z., Chmelík, J., & Doležal, M. (2022). Collaborative and individual learning of geography in immersive virtual reality: An effectiveness study. *Plos one*, *17*(10), e0276267.

Teng, Z., Cai, Y., Gao, Y., Zhang, X., & Li, X. (2022). Factors affecting learners' adoption of an educational metaverse platform: An empirical study based on an extended UTAUT model. *Mobile Information Systems*, *2022*(1), 5479215.

Wang, C. H. (2024). Education in the metaverse: Developing virtual reality teaching materials for K–12 natural science. *Education and Information Technologies*, 1-22.

Wang, C. H. (2025). Construction of an educational metaverse shared learning
platform. *Impact*, *2025*(1), 18-20.

Wang, M., Yu, H., Bell, Z., & Chu, X. (2022). Constructing an edu-metaverse ecosystem: A new and innovative framework. *IEEE Transactions on Learning Technologies*, *15*(6), 685-696.

White, R., & Gunstone, R. (2014). *Probing understanding*. Routledge.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice*, *41*(2), 64-70.

信息技术赋能科学素养评测：发展脉络、研究热点与未来趋势

Information technology-enabled scientific literacy assessment: developmental history, research hotspots and future trends

吴国静[1]，王靖[2]
江南大学江苏“互联网+教育”研究基地
2994237780@qq.com

【摘要】 面对科学教育中科学素养评测科学化、高效化、个性化的需求，传统评测方式的短板愈加明显。因此，信息技术赋能的科学素养评测发展历程、现状和未来走向成了科学教育领域的关键议题。为此，对1950-2025年国内外科学素养评测实证研究成果进行系统性文献综述，讨论了科学素养评测方式的转变、热点和未来趋势。研究发现：科学素养评测方式经历了纸笔测试、计算机静态评测、基于计算机的人机交互三个阶段。研究焦点集中于基于多场景表现性评价和融合多模态数据的科学素养评测。最后，本文从积极融入生成式人工智能、关注本土创新等五个方面做出了展望。

【关键词】 科学素养评测；科学素养；信息技术；评测方式；科学教育

Abstract: *In the face of the demand for scientific, efficient and personalized scientific literacy assessment in science education, the shortcomings of traditional assessment methods have become more and more obvious. Therefore, the development history, current status and future direction of information technology-enabled scientific literacy assessment have become key issues in the field of science education. To this end, a systematic literature review of the empirical research results on scientific literacy assessment at home and abroad from 1950 to 2025 was conducted to discuss the transformation, hotspots and future trends of scientific literacy assessment methods. It is found that scientific literacy assessment methods have gone through three stages: paper-and-pencil test, computerized static assessment, and computer-based human-computer interaction. The research focuses on scientific literacy assessment based on multi-scenario expressive evaluation and integration of multi-modal data. Finally, the paper makes an outlook in five aspects, including active integration of generative artificial intelligence and attention to local innovation.*

Keywords: Science literacy assessment, scientific literacy, information technology, assessment methods, science education

1.引言

当前，建设科技强国，发展新质生产力已成为国家战略的核心命题，科学教育不仅承载着知识传递的功能，更肩负着培养科学素养、激发创造潜能、培育创新人才的重要使命。2022年我国教育部最新修订的《义务教育科学课程标准（2022年版）》为科学教育的开展提供了指导方针，并提出了将科学素养作为学生发展的重要目标。国务院2021年颁布的《全民科学素质行动规划纲要（2021-2035年）》提出“科学素质是国民素质的重要组成部分，是社会文明进步的基础”（国务院，2021）。对科学素养进行评测是检验科学教育效果不可或缺的环节。目前已有的评测主要是以书面形式为主，仅以文字和图片呈现给学生（陈

睿智等，2024），以分数体现学生的科学素质和能力情况。然而随着技术的发展，人工智能、大数据等为实现科学素养评测的动态化、个性化提供了契机。

在此背景下，本文提出以下研究问题：在已有的科学素养评测研究中，技术赋能的科学素养评测的经历了哪些发展阶段？当前新技术不断涌现，科学素养评测研究焦点集中在哪些方面？新兴技术与教育评测深度融合的背景下，科学素养评测将呈现怎样的发展趋势？为此，本文对科学教育中以技术支持的科学素养评测的发展进行分析与解读，通过分析科学素养评测的各阶段特点，总结以技术支持的科学素养评测研究的热点和未来发展趋势，以进一步把握科学素养评测的发展轨迹以及未来研究取向，为后续的科学素养评测的研究和实践提供有益参考。

2.科学素养概述

二战后，军事上形成了以美国和苏联为首两大阵营对峙的局面，科学研究团体认识到科学技术对于国防的重要价值，而科学技术的提升离不开公民对技术的理解和支持。美国国家科学基金会（National Science Foundation, NSF）在 1950 年明确表示“科学进步需要获得公众对科学教育和科学研究的支持”（AT Waterman，1960）。

1952 年，美国化学家、教育学家、哈佛大学前校长 J.B.柯南特第一个提出科学素养定义。柯南特指出，科学素养是一种明智地选择专家的能力，以及能够“聪明地与那些正在发展科学并应用科学的人交流的能力”（Richard，1998）。尽管没有对其意义与内涵进行深入研究，但是这从普通教育的层面对科学素养进行了定位，为后续的科学素养相关研究指明了方向。1958 年赫德在基础教育领域采用了科学素养这一概念，并将其解释为“理解科学且能把其应用在社会中”。可见，由于时代背景的需要，20 世纪五六十年代提倡的科学素养是科技精英人才、职业科学家所具备的科学知识、理解科学事业的能力。

60 年代中期，美国的佩勒（Pella）等人从六个方面对科学素养这一概念进行了综合性的概括（魏冰，2000）。佩勒等人的研究成果建立起了科学素养的基本框架，为科学素养的深入研究奠定了基础。

20 世纪 70 年代，经济全球萧条和科技进步带来的环境问题，国家重新关注个体的需求和发展，科技政策也将重点从国防转向社会服务。在此背景下，科学素养的内涵也进一步得到丰富。具备科学素养的人能够“理解科学、技术和社会经济发展其他方面的关系”（NST Association，1971）。到了 80 年代，科学素养不在局限于理论层面的讨论，而是作为教育目标明确提出“能够理解科学、技术、社会之间的相互影响并能在日常生活中运用这些知识做出决策”（TNST Association，1982），这也标志着科学—技术—社会教育正式拉开序幕。这一时期，“科学”和“技术”被置于同等地位，提升公民的科学素养不再是为了培养科学家，而是面向全体学生，实现“科学为大众”。

科学素养评测以明确科学素养的概念为基础，并在此基础上围绕科学素养的结构框架展开。1983 年 Miller 提出了科学素养三要素模型“（1）科学术语与概念的词汇量；（2）对科学过程的理解；（3）科学技术对个人和社会影响的了解”（Aikenhead G S & Ryan A G，1992）。这也是后来米勒公民科学素养三维理论模型的雏形。

20 世纪 90 年代末，人类迈入了信息时代，互联网、云计算、大数据等信息技术的蓬勃发展，短视频、公众号等新媒介不断融入公众的生活，打破了书籍、报纸和电视作为传播科学信息的唯一途径的局面。科技信息的获取、传播速率和内容产出在数字媒介技术迭代的驱动下便捷性不断提高，尽管公众接触低信度、伪科学的概率和风险增加了，但是媒体平台构建

的点赞、评论和转发的异步交互场景，已经深度重构和影响了公众对科学的态度和理解。在这高度信息化的社会背景下，科学素养内涵也不断拓宽。美国国家科学院联合多个组织机构编写的《科学素养：概念、情境与影响》强调了“面对信息洪流，科学素养需要的是整合并解释信息的能力，以及进行反思和评价的能力（凯瑟琳·E·斯诺，肯妮·A·迪布纳，2020）。大规模国际科学素养评测项目也在随着时代的变化完善评测内容，经济合作与发展组织（OECD）发布的 PISA2025 框架明确将“评估科学数据的可信度与缺陷”纳入核心能力维度。可见，随着科技的飞速发展，新兴事物不断涌现，日新月异的变革周期不断缩短，个体只有在掌握科学知识和技术的基础上独立批判地思考和发现问题、辩证地分析问题，才能做出合理决策来解决问题。

3.研究设计

3.1 研究方法

本研究采用系统性文献综述法（Preferred Reporting Items for Systematic Reviews and Meta-Analyses，简称 PRISMA），对国内外科学素养评测相关文献进行分析，以总结技术支持的科学素养评测发展脉络，精准把握当前该领域的研究热点，并基于现有研究对该领域未来发展进行展望。在明确研究问题的前提下，通过检索、筛选、合格和纳入 4 阶段对技术支持的科学素养评测的相关文献进行筛选。

3.2 文献获取

本文从 Web of Science 核心数据库中，使用数据库的高级检索功能以“Science Literacy Assessment or Scientific Literacy Assessment”为主题词进行检索，文献类型为“ARTICLE”，检索截止时间为 2025 年 3 月，共检索到科学素养评测相关文献 982 篇。

选取中国知网（CNKI）数据库的作为国内科学素养评测研究的数据库来源。以“科学素养评测”为主题，检索到论文共 59 篇。中外文数据库累计获取文献量 1041 篇。

3.3 文献筛选

为了保证文章的可靠性和质量，本研究制定了文献筛选标准，根据标准逐一对文献进行筛选。纳入与排除标准如下表所示。

表 1 文献筛选标准

编号	纳入标准	排除标准
C1	文献收录期刊是否是 SSCI、SCI、CSSCI 级别	非上述期刊
C2	研究主题是否为“科学素养评测”	研究主题不相关
C3	研究内容含有完整的研究设计、科学素养评测的工具或评测方法	无完整研究设计，不涉及科学素养评测工具或评测方法
C4	研究类型是否为实证研究	研究类型非实证

献的标题、关键词和摘要根据制定的文献筛选标准对文章进行初步筛选。其次，通过阅读全文对文献进行二次筛选。最后，利用“滚雪球”的方法检索纳入更多符合标准的文献，最终选出符合标准的文献共计 41 篇，文献筛选过程如图所示。

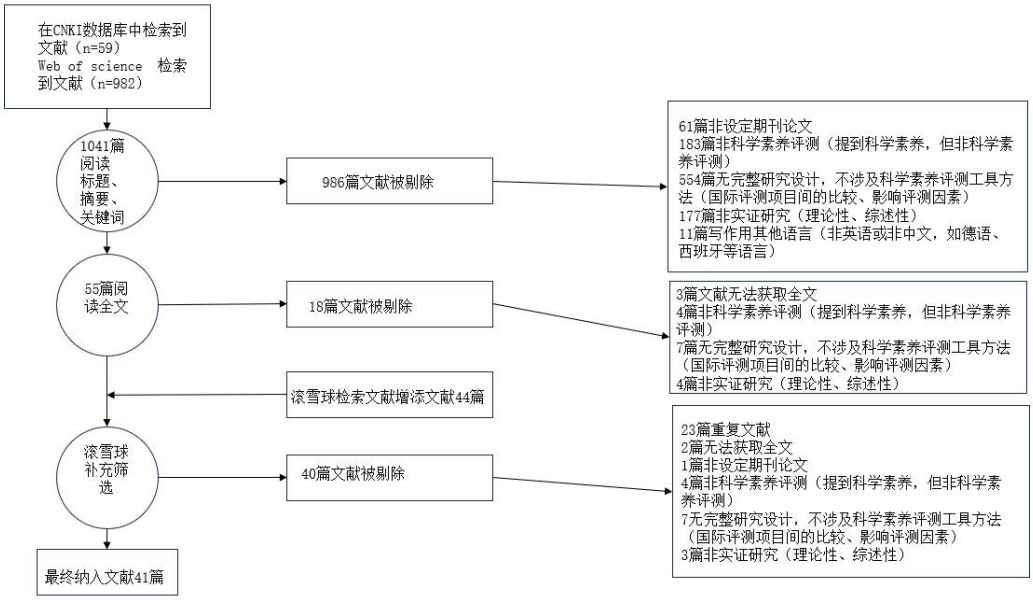


图 1 PRISMA 流程图

4.研究结果和讨论

4.1 技术支持的科学素养评测的历史演进

最早对科学素养进行评测可以追溯到 1957 美国开展的公众科学素养测评项目“公众对科学的理解”。这不久后教育评测领域就引入了计算机技术，但这之后的 20 年间，基于计算机的测试主要在课堂环境下用于学生的陈述性知识的回答判断（Shute, Valerie J. & Seyedahmad Rahimi，2017）。虽然引入计算机技术是教育评测领域上的一大创新性突破，但是该阶段并没有充分利用和发挥计算机的优势。试题类型只有选择题和填空题，计算机的反馈局限于作答结果的正误判断，被试与计算机之间的交互限制于简单的鼠标点击。

80 年代末 90 年代初，国际上陆续开展科学素养评测项目，并利用评测结果推动科学教育改革和科技政策制定。这期间，科学素养评测主要是通过静态纸笔测验的形式进行。尽管纸笔测验便于实施，但纸笔测验存在着一系列问题，不仅难以评测学生的高阶思维，还脱离真实情境，不能评估出学生在真实情况下观察的能力（Shavelson, R. J. , Baxter, G. P. , & Pine, J, 1994）。

90 年代末起，研究人员 Baker 和 Mayer 等人开始使用计算机来评估更复杂的认知技能。科学教育领域也逐渐突破纸笔评测局限，引入了利用人与计算机交互开展评测的思想。美国教育测量服务中心的梅斯雷弗（Mislevy,R.j.）等研究者于 1997-1999 提出了“证据中心设计”（ECD）理论，该理论指导了美国国家教育进展评估（National Assessment of Educational Progress， NAEP）、国际数学和科学成就趋势研究（Trends in International Mathematics and Science Study, TIMSS ）和国际学生能力评估项目（Program for International Student Assessment, PISA）等科学素养大规模评测项目中的基于信息和通信技术（Information and Communication Technology，ICT）任务的设计。

自此，人机交互理念走进大规模科学素养评测项目。2006 年，PISA 项目首次尝试利用人机交互的形式开展科学素养评测（OECD,2010）。随后在 2009 年， NAEP 使用了操作性任务和计算机交互任务两种新题型评估学生的科学素养(NAEP,2023)。这个时候尽管人机交互试题投入科学素养评测的时间较短，但其自动化和高效性还是受到了众多研究者的密切关注。

Quellmalz 等人依据提炼的基于计算机模拟的科学评估原则，围绕初中生态系统开发了 12 道题，通过 1566 名学生完成评估，得出了互动评估比静态评估在区分学生的探究能力方面更有优势的结论(Quellmalz, E. S.，Davenport, J. L.，Timms, M. J.，Deboer, G. E.，Jordan, K. A.，& Huang, C. W. , et al,2013)。

21 世纪以后，人机交互开展科学素养评测的研究不断深入。吴心楷等人开发了基于多媒体的科学探究能力评估工具，情境化开发了 114 道科学人机交互试题，并通过 1066 名学生的参与检验该工具上人机交互试题的有效性(Kuo, C. Y.，Wu, H. K.，Jen, T. H. , & Hsu, Y. S. ,2015)。2015 年 PISA 首次正式在大规模的科学素养评测中使用人机交互试题，在随后的 PISA2018 和 PISA2022 继续沿用了该评测形式。2019 年 NEAP 全面引入人机交互的测评题。TIMSS 也从 2019 年将评测形式从纸笔过渡到利用计算机进行评测。

上述描绘的技术支持的科学素养评测历史演进关键节点如下图所示。

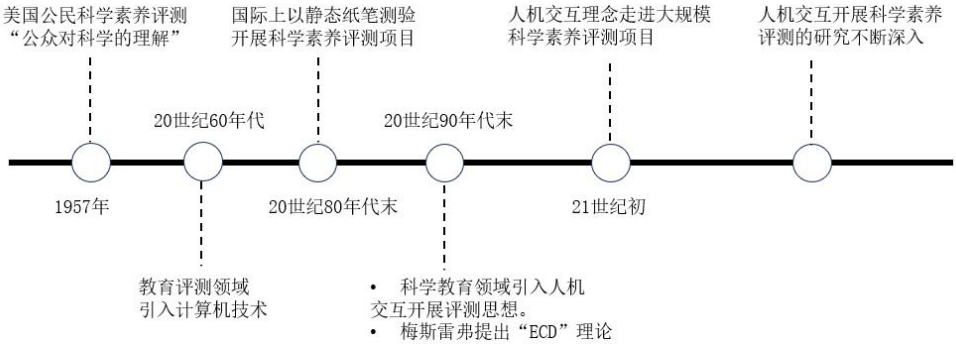


图 2 技术支持的科学素养评测历史关键节点

4.2 技术支持的科学素养评测研究热点分析

随着技术的发展，技术支持的科学素养评测正朝着个性化、智能化、更科学化评测的方向发展。在这一阶段，学者们不断探索和创新，尝试将多种技术和工具融入科学素养评测，期望提高科学素养评测的准确性和有效性。

“以证据为中心”的评测设计模式（evidence centered design, ECD）是系统性进行评价设计的模式，该模式已经成为教育评价领域主要的研究方向和应用领域之一。该模式明确指出教育评价本质上是基于学生的表现获取关于学生能力的证据进而对学生的知识、技能和成就进行推理（冯翠典，2012）。我国研究者在借鉴国际大规模评测项目和以证据为中心 ECD 模型理念的基础上，将科学素养评测方式从纸笔测验逐步过渡到计算机交互。陈慧珍等研究者借鉴 NAEP 计算机交互评测技术，围绕“温室效应与植物光合作用”设计并实施评测方案，施测结果显示，该评测技术能够充分考察学生的知识、能力素养等水平（陈慧珍等，2018）。刘枳衫、林娉婷同样设计并实施了基于计算机交互和动手实操任务的科学素养评测（刘枳衫，2016；林娉婷，2016）。除了基于计算机的交互，还有学者将传感器也融入科学素养评测中，以实现对学生的实操评测。庞谦竺在深入分析 PISA 和 NAEP 科学素养评估框架的基础上，设计了基于计算机和传感器的九年级科学素养评估方案（庞谦竺，2019）。周海涵同样将计算机平台、传感器技术融入到科学素养评测，设计了基于信息和通信技术（Information and communication technology，ICT）的“地球的运动”科学素养评测系统（周海涵，2022）。目前，PISA 已经采用自适应技术精准定位不同地区和群体的科学素养水平差异，动态匹配评测项目，满足个性化评测需求。也有研究者尝试将自适应技术的引入科学素养评测。通过在评测中根据学生的能力水平，动态调整评测项目的难度和内容，实现高效化、个性化科学素养评测(李小红等，2024)。

4.2.1 融合多模态数据的科学素养评测

在教育评测领域，充分借助数据的创新趋势来深化教育评价的趋势所在。随着人工智能技术、大数据、物联网等新兴一代技术的不断成熟，数据赋能于教育评价逐渐成为常态化。在此背景下，融入多模态的科学素养评测悄然兴起，并吸引了众多研究者进行尝试与应用。陈信宇通过对学生在科学素养评测中的概念任务和动手操作中引入眼动追踪技术，探究学生的注意力分配与认知差异，将客观的神经教育学数据融入科学素养评测（陈信宇，2020）。Andy Nguyen 等人借助生理信号、音频数据和视频数据用于探究光和声音物理概念协作学习中的社会调节过程，进而全面了解科学学习的过程，并利用多模态数据来调节、评估学习效果（Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J，2023）。

4.2.2 基于多场景表现性评价的科学素养评测

表现性评价能够反映学生习得知识并运用知识来解决问题，在这一过程中，重点评估学生的问题解决能力、高阶思维能力和实际应用能力，因此许多高学业成就的国家将其纳入课程和测试当中（张殷等，2024）。由上述的定义可知，表现性评价在评估批判性思维、创新能力等高阶能力方面具有巨大潜力。在科学素养评测上，表现性评价体现为进行实验设计、分析数据、得出结论，即完成一项实践探究，或者是解决实际问题，或者形成一个作品。具体来说有以下几个表现：设计情境化任务。学生在完成任务的过程中，需要读取相关资料，设计实验并解释现象。如 PISA2025 样题 1“温室效应”中通过阅读文字和图表，寻找支持“二氧化碳排放量上升会导致气温上升”，并解释哪些信息不支持该论点（张殷等，2024）。模拟科学实验环境的交互式科学任务。这些任务通常在基于计算机的虚拟环境中进行一系列实验操作，实现真实问题的解决。NEAP、PISA 和 TIMSS 目前都已经采用了基于计算机的交互式测评。动手实验任务。评测过程中学生需要利用科学仪器和材料完成实验任务，与此同时，研究者依据一定的标准进行评估（张殷等，2024）。

4.2.3 数据驱动的智能化科学素养评测

随着技术的发展，人工智能、大数据等新技术的应用不断深入，新技术赋能科学素养评测使得精细多元的评价方式和伴随式数据采集逐渐成为现实。刘邦奇等认为，智能技术与教育评价相融合，驱动数据采集过程化、分析多模化、评价分析反馈及时化，教育评价的数字化、智能化、专业化得到显著提升（刘邦奇等，2021）。在科学素养评测领域，不少学者借助人工智能技术，尝试实现智能化评测。

Wilson 等利用机器学习实现学生科学论证以及绘制的模型的自动评分，研究结果显示人类批改的分数和机器给出的分数具有高度一致性（CD Wilson 等人，2023；zhai、He 等人，2022）。孔艺和林静借助机器学习的自动评分工具 c-rater-ML，利用图尔敏模型分析逻辑结构分析学生科学论证内容，准确率达 90%（孔艺，林静，2021）；王晶莹等人借助模糊神经网络技术对影响学生科学素养的相关因素进行挖掘，根据研究结果对科学素养的培养提出建议（王晶莹等，2024）。

聚焦当下信息技术的发展，随着人工智能在教育领域的应用，自动化测评及多模态人机之间的协同与交互持续受到了研究者的高度关注。这一现状既是智能技术驱动科学素养评测从经验驱动向数据驱动转型的结果，也是人机协同评测模式在实际教育场景中的应用需求。

4.3 对未来我国科学素养评测研究的展望

运用新技术改进和优化测评方式已成为当前教育质量监测评价改革的重要趋势（李小红等，2024）。21 世纪以来，技术赋能科学素养评测涌现出的一系列研究热点与问题，是我

国教育评测和科学教育发展方向抉择的重要参照与影响因素。研究者应当面向未来教育，紧跟时代发展的脚步重新审视科学素养评估。

第一，打破学科融合，多方参与，建设评测共同体。高等院校继续发挥其在科学素养学术研究方面的优势，通过不同研究机构、研究者之间的沟通合作，确保以技术支持的科学素养评测研究质与量均衡发展。其次，增强各学科之间的融合，打破学科壁垒，从教育学、科学、计算机、传播学等多学科视角进行研究，实现高水平的科学素养评测创新。

第二，确保科学素养评测基础夯实并紧密追踪热点，确保科学素养评测满足时代需求。首先，科学素养评测相关的基本问题，如科学素养的概念，在新的时代背景和理论视角下需要进行反思和重构，增添时代赋予的意蕴，紧扣人才培养需求。其次，从文献关键词的整体分析来看，科学素养与核心素养之间的关系，以及落实到教学一线，二者如何作用与影响也需要进一步探讨。

第三，在与国内外已有宝贵经验对话与借鉴的同时，关注自身本土化创新。既要吸收引进国际技术支持的科学素养评测相关研究成果，又能立足于本土特色与国情，使得评测方式能够迁移与落地。在科学素养评测工具方面，在研究、参与国际评测项目的同时要尝试自主设计、开发和应用技术支持的科学素养评测方案与工具。尽管目前由于设备环境等条件，科学素养评测仍然是基于纸笔测验进行开展，但由于纸笔存在的仅聚焦于知识层面、无法记录作答过程等一系列不足，基于计算机的人机交互和数字化、智能化评测工具成为科学素养评测的重要途径。在设计与使用上述方案与工具的过程中，既要吸收 PISA、TIMSS 相关经验，如整个评测流程的设计理念、评测框架的梳理与构建、试题的编制、评测结果的分析方式等等。同时也要关注课标中学生的学业质量要求并在设置问题情境时积极融入中华传统优秀文化。

第四、积极融入生成式人工智能，推动评测方式更加智能化。以 ChatGPT 为代表的人工智能生成式内容（artificial intelligence generated content，AIGC）在教育领域受到了高度关注。与传统数字技术不同，生成式人工智能技术，是指具有文本、图片、音频、视频等内容生成能力的模型及相关技术。生成式人工智能主要赋能于评测资源整合、试题命制、评测实施与评测结果的分析反馈（国家互联网信息办公室等，2023）。AIGC 作为一种新兴技术，其用于科学素养评测尚处于积极探索阶段，虽然产出了一定的研究成果，但在实际评测中并未具体实施。因此，研究者在将 AIGC 应用于科学素养评测时，既要看到其广阔的应用前景，也要关注到应用中面临的诸多挑战，应当加强实验试点，积极探索将生成式人工智能融入科学素养评测的机制和路径，推动科学素养智能化评测体系的建立与完善。

参考文献

陈睿智,谢晓雨,罗莹.(2024).科学教育测评中人机交互试题的发展、特征与启示.中国考试,(08),79-88. https://doi.org/10.19360/j.cnki.11-3303/g4.2024.08.009

陈信宇.(2020).眼动技术在科学素养评测中的应用研究(硕士学位论文,东南大学).硕士 https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2020.001842doi:10.27014/d.cnki.gdnau.2020.001842.

陈慧珍,刘枳杉,柏毅.(2018).高中生物学素养评测方案的设计与实施.教育测量与评价,(08),49-55. https://doi.org/10.16518/j.cnki.emae.2018.08.008

冯翠典.(2012).“以证据为中心”的教育评价设计模式简介.上海教育科研,(08),12-16. https://doi.org/10.16194/j.cnki.31-1059/g4.2012.08.003
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,(19),12-20.2025-4-28 取自 https://www.gov.cn/gongbao/content/2021/content_5623051.htm
国家互联网信息办公室等（2023）.《生成式人工智能服务管理暂行办法》. 2025-4-28 取自 https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm
凯瑟琳·E·斯诺,肯妮·A·迪布纳(2020).*科学素养：概念、情境与影响*(裴新宁，正太年 译).北京：中国科学技术出版社.
孔艺,林静.(2021).基于机器学习的科学论证自动评分个案研究.中国考试,(08),61-68. https://doi.org/10.19360/j.cnki.11-3303/g4.2021.08.009
李小红,李珍,王克志.(2024).PISA 的新进展分析及其启示.教育科学,40(03),31-37. https://doi.org/10.3969/j.issn.1002-8064.2024.03.005
林娉婷.(2016).学生科学核心概念评测方案研究(硕士学位论文,东南大学).硕士 https://kns.cnki.net/kcms2/article/abstract?v=_dzes7vZToNws8h6lHVBnOF5GLcJTN8wFGQOQM6NUA_xrRVHFe1JhyUKIyVxXHwZIKrYcv1ODh9qfxwbGpuX9Y8d84uYQDqg3ruXGweb0REjBgjFGK8lAHVnb_lDvicqNtltBVN3uL2o03GIn_f69zucH520G-4qZ7KQ_Sgf6NCPSLx37k0Ifz4uizyyUxWlP9dsGOVMKWs=&uniplatform=NZKPT&language=CHS
刘邦奇,袁婷婷,纪玉超,刘碧莹 & 李岭.(2021).智能技术赋能教育评价：内涵、总体框架与实践路径.中国电化教育,(08),16-24. https://doi.org/10.3969/j.issn.1006-9860.2021.08.003
刘枳杉.(2016).学生科学素养评测方案研究(硕士学位论文,东南大学).硕士 https://kns.cnki.net/kcms2/article/abstract?v=_dzes7vZToP-F1vK6jAwCPouZVR87oL6QUXItKATeY7cZjU0GvvIUMZ-X_uN5v5JS9n7GJcofGBzd4ZsG69NEbvSgUcyRpQXZ9pNYCa0GQgNN12XSeA2kxb6-XRv4x4u5M450Yc_PFHdPTnqkHi6krOTfYmHcomX-uclieXWqdbwS5k_9lkWrHGFZZGxXTPfqymiTSznWec=&uniplatform=NZKPT&language=CHS
庞谦竺.(2019).九年级科学素养评测方案设计(硕士学位论文,东南大学).硕士 https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2019.003165doi:10.27014/d.cnki.gdnau.2019.003165.
魏冰.(2000).“科学素养”探析.比较教育研究,(S1),105-108.
周海函.(2022).基于 ICT 的科学素养评测方案开发(硕士学位论文,东南大学).硕士 https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2022.001532doi:10.27014/d.cnki.gdnau.2022.001532.
张殷,吴柳燕,罗星凯.(2024).聚焦科学高阶思维的科学表现性任务设计研究.天津市教科院学报,36(06),3-14. https://doi.org/10.3969/j.issn.1671-2277.2024.06.001
Association, N. S. T. . (2010). Principles of professionalism for science educators. national science teachers association position statement. National Science Teachers Association, 4. https://www.nsta.org/nstas-official-positions/principles-professionalism-science-educators
Association, T. N. S. T. . (1982). Science-technology-society : science education for the 1980s. The NSTA Position Statement. https://scholarworks.uni.edu/istj/vol20/iss2/11

AG Aikenhead . (2010). The development of a new instrument: 'views on science—technology—society' (vosts). Science Education, 76(5), 477-491.https://doi.org/10.1002/sce.3730760503
Kuo, C. Y. , Wu, H. K. , Jen, T. H. , & Hsu, Y. S. . (2015). Development and validation of a multimedia-based assessment of scientific inquiry abilities. International Journal of Science Education, 37(14), 2326-2357. https://doi.org/10.1080/09500693.2015.1078521
National Assessment of Educational Progress .(2012). Science in Action: Hands-On and Interactive Computer Tasks From the 2009 Science Assessment. https://doi.org/10.1007/978-94-007-2150-0_67
Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J. (2023). Examining socially shared regulation and shared physiological arousal events with multimodal learning analytics. British Journal of Educational Technology, 54(1), 293 - 312. https://doi.org/10.1111/bjet.13280
Organizationfor Economic Co-operation and Development .(2010). Pisa pisa computer - based assessment of student skills in science: (edition complète - isbn 9789264082038 - en angl. seulement). SourceOCDE Science et technologies de l'information, volume 2010, i-132(132). https://doi.org/10.1016/j.compedu.2020.103964
Quellmalz, E. S. , Davenport, J. L. , Timms, M. J. , Deboer, G. E. , Jordan, K. A. , & Huang, C. W. , et al. (2013). Next-generation environments for assessing and promoting complex science learning. Journal of Educational Psychology, 105(4), 1100-1114.https://doi.org/10.1037/a0032220
Shute, V. J. , & Rahimi, S. . (2017). Review of computer - based assessment for learning in elementary and secondary education. Journal of Computer Assisted Learning, 33(1), 1 - 19. https://doi.org/10.1111/jcal.12172
Shavelson, R. J. , Baxter, G. P. , & Pine, J. . (1991). Performance assessment in science. Applied Measurement in Education, 4(4), 347-362. https://doi.org/10.1016/S0191-491X(99)00021-8
Waterman, A. T. . (1960). National science foundation: a ten-year resume. Science, 131(3410), 1341-1354. http://doi.org/10.1126/science.131.3410.1341
Wilson, C. D. , Haudek, K. C. , Osborne, J. F. , Buck Bracey, Z. E. , Cheuk, T. , & Donovan, B. M. , et al. (2024). Using automated analysis to assess middle school students' competence with scientific argumentation. Journal of Research in Science Teaching, 61(1). http://doi.org/10.1002/tea.21864
Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. Journal of Research in Science Teaching, 59(10), 1765 - 1794. https://doi.org/10.1002/tea.21773

初探虛擬實境教材沉浸程度之影響：以 VR 動畫專注觀察為例

Investigating Immersion in VR-Based Learning: The Role of Focused Observation in Animated Experiences

侯愷均 [1*]，曾俊雄 [2]
[1] 國立臺北教育大學數位科技設計學系
[2] 元智大學電機工程系
[*] kaichu.hou@gmail.com

【摘要】 本研究旨在初步探討虛擬實境（Virtual Reality, VR）教材於大專院校設計科系學生學習體驗中所引發之沉浸感受程度。透過比較配戴 VR 頭盔與 2D 螢幕觀看相同 360 度動畫影片《Lutaw》的學習體驗。研究對象為大二與大三學生，實驗過程分兩階段，第一階段是專注力觀察測驗，學生需觀察影片內容細節並完成 10 題測驗題，完成此階段共 40 人；第二階段是更換體驗方式，從 2D 螢幕觀看改為配戴 VR 頭盔觀看，完整體驗此階段共 17 位。結果顯示，VR 組在沉浸感、臨場感、專注度與參與感等心理層面顯著優於 2D 組，然而測驗分數並無顯著差異。研究指出虛擬實境能有效提升學習沉浸體驗，而在評估成效的測驗階段上無論以何種方式體驗，並不會影響學習成效，為未來元宇宙教學應用提供參考依據。
【關鍵字】 虛擬實境；元宇宙；VR 動畫；VR 頭盔；專注觀察

Abstract: *This study aims to preliminarily explore the degree of immersion elicited by virtual reality (VR) educational materials among university students majoring in design-related fields. The research compares students' learning experiences while viewing the same 360-degree animated film Lutaw using either a VR headset or a traditional 2D screen. Participants were second- and third-year undergraduate students, and the experiment was conducted in two stages. The first stage involved a focused observation test, in which students watched the film and answered 10 multiple-choice questions based on its details. A total of 40 students completed this stage. In the second stage, students switched from 2D viewing to a VR headset experience; 17 students completed this phase. Results indicated that students in the VR condition reported significantly higher levels of immersion, presence, concentration, and engagement compared to those in the 2D condition. However, no significant differences were found in test scores between the two groups. The findings suggest that while VR can enhance the immersive quality of the learning experience, the mode of delivery—VR or 2D—does not appear to affect measurable learning outcomes. These insights offer valuable implications for future educational applications of VR in metaverse-based learning environments..*
Keywords: Virtual reality, Metaverse, , VR animation, HMD, focused observation

1. 研究背景與動機

隨著元宇宙 (Metaverse) 概念的興起，虛擬實境 (Virtual Reality, VR) 技術已日漸成為教育領域中具潛力的數位媒介之一。特別是在跨領域學科與設計相關教學中，具備高沉浸性、互動性與敘事張力的 VR 內容被視為新興的教學趨勢與學習模式。VR 不僅提供豐富的感官

刺激與空間臨場感，也可能提升學習者的學習動機、專注程度與情境參與度(Makransky &Lilleholt, 2018)。

然而，目前針對虛擬實境教材形式在高等教育中之實際成效，仍有值得探討的研究領域。尤其對於設計學門的學生而言，教材是否能有效支持其視覺理解、專注程度、情境感知與敘事掌握力的培養，將直接影響教學品質與學習成果。虛擬實境動畫結合影像敘事與沉浸體驗，若能被實證證明具教育價值，將為未來設計教育教材開發帶來新方向。為此，本研究以榮獲國際獎項的 360 度動畫短片《Lutaw》[9]作為實驗教材，透過 VR 頭盔與傳統 2D 螢幕兩種觀看情境，探討設計相關科系學生在沉浸感、臨場感、專注度與參與度等心理感受面向上的差異表現，並結合學習成效評量工具，以檢視虛擬實境動畫對學生學習經驗與投入程度之實際影響。本研究期望透過初步實證結果，補充 VR 動畫應用於高教教材設計的學術討論空缺，並提出教學與媒體選擇之實務建議。

2. 文獻探討

虛擬實境的教育應用近年來廣受關注，根據 Slater 與 Wilbur (1997) 所提出之沉浸感理論 (Sense of Presence)，使用者對虛擬環境的投入程度會直接影響其學習效果。Scholz 與 Duffy (2018) 指出，視覺沉浸度越高的 VR 內容越能強化學習動機與情境記憶。而根據 Makransky 與 Lilleholt (2018) 的實驗研究，沉浸式學習環境有助於學生提升專注力與情感投入，進而增進知識吸收。而隨著科技日新月異，新的載體誕生，可提供更具沉浸感的 VR 硬體頭盔，將在視覺上帶來新的體驗，本研究呼應上述文獻，嘗試以 360 度動畫實作驗證此沉浸經驗的可量化成效，並比較不同媒介呈現對感知層面的具體影響。

此外，在 VR 學習成效評估上，有學者指出雖 VR 能提升沉浸與興趣，但學習成效與傳統媒介相比未必顯著提升，需依內容與學習目標調整應用策略 (Parong &Mayer, 2018)。觀察學習是一種透過模仿他人行為來獲取知識的社會學習方式。心理學家班杜拉（Albert Bandura）的社會學習理論指出，觀察學習包含四個主要過程：注意 (Attention)、記憶 (Retention)、再現 (Reproduction) 和動機 (Motivation)。其中，注意力是第一步，學生必須專注於所觀察的行為或細節，才能有效吸收知識(Bandura &Jeffrey, 1973; Rumjaun &Narod, 2020)。有別於知識的背誦，本研究試著從專注觀察的層面進行學習成效上的探討。

3. 研究方法

本研究分由兩階段實驗所組成，第一階段為觀察測驗，採用準實驗設計法，實驗對象為某設計科系大二與大三學生共 40 位，為臺北市某大學某選修課程的學生，採自願制，無論是否參與實驗，皆與課程成績無關。實驗影片為得獎的《Lutaw》（圖 1），8 分鐘 360 度 3D 動畫虛擬實境影片，敘述貧困地區孩童以浮潛方式上學的真實故事。觀看方式共分兩種模式，分別為 2D 螢幕觀看，以及配戴 VR 頭盔進行體驗，頭盔型號是具有高度視覺包覆與高沉浸的 Meta Quest 3。實驗進行方式，第一階段隨機分兩組體驗（2D 螢幕與 VR 頭盔），體驗後進行觀察評測與填寫沉浸感受程度問卷量表；第二階段讓學生更換第二種體驗方式，僅需填寫沉浸感受量表，無需再次進行觀察評測（因為觀看影片與題目皆相同，只是答題順序為隨機）。並非所有學生皆參與第二階段，少部分學生對於 VR 頭盔感到暈眩，部分則是問卷量表沒有填寫完整（因採自願制，無強迫限制）而中斷實驗。在評量方式階段上，第一

[9] *Lutaw* 是一部 VR 動畫影片，由菲律賓希望之船(Yellow Boat of Hope)NGO 組織所發起，敘述兩姊弟透過創意找到解決上學的交通方式。

階段採用 10 題選擇題，作為學生是否正確答對影片中的細節；第二階段採用 5 階問卷量表，進行沉浸感(Agrawal et al., 2021)、臨場感(Lengyel et al., 2021)、專注度（本研究自製）與參與感（本研究自製）等四個構面進行心理主觀量測。

圖 1 Lutaw VR 動畫影片

4. 實驗結果

為探討不同動畫觀看體驗方式（2D 螢幕與 VR 頭盔）對 40 位學習者觀察測驗表現的影響，第一階段實驗，經過兩組的獨立樣本 t 檢定後，如表 1 所示。為檢驗受試者在不同觀看條件（VR 頭盔與 2D 螢幕）下的學習成效差異，本研究進行獨立樣本 t 檢定分析。2D 螢幕組（M = 85.22, SD = 16.48, N = 23）與 VR 頭盔組（M = 85.29, SD = 8.75, N = 17）之平均分數極為接近，t(38) = -0.017, p = .986 > .05，無達顯著水準。表示兩組學習者在細節觀察能力上表現相近，意謂著即使體驗方式不同（2D 或 VR），受測者在觀察測驗中的表現並未顯著改變。

表 1 兩組在專注觀察評測上的檢定

組別	N	平均數（M）	標準差（SD）	t 值	df	p 值（雙尾）	顯著與否	平均差異	95% 信賴區間
2D 螢幕組	23	85.22	16.48						
VR 頭盔組	17	85.29	8.75	-0.017	38	0.986	無顯著	-0.077	[-8.99, 8.83]

本研究第二階段實驗針對 2D 螢幕與 VR 頭盔體驗後的心理素質分析，此階段僅有 17 人完整歷經二種模式的體驗並填寫完整問卷，達一半以上受試數量，且問卷整體信度（Cronbach's Alpha）為 0.916，顯示量表具有極佳的一致性與可靠性。經由描述性統計與 t 檢定，結果顯示如下表 2：VR 頭盔組在所有心理感受項目（如「沉浸感」、「臨場感」、「專注程度」與「參與感」）均顯著高於螢幕觀看組（p < .05）。特別是在「臨場感」項目中 (t = -6.154, p < .001)，VR 觀看組表現尤為顯著。此結果支持虛擬實境裝置可有效提升學習者對教學媒材的沉浸與投入程度，並暗示教學媒體的感官包覆度可能為影響學習感受之關鍵因子。這項結果進一步支持虛擬實境在跨域與設計教育中的應用價值，並為教材設計提供實證基礎。

表 2 兩組體驗模式之檢定

感受項目	M (2D)	M (VR)	差異 (VR - 2D)	M 標準差（差異）	t 值	df	p 值	顯著性說明

專注度	4.94	5.65	0.71	1.16	-2.51	16	0.023	*
沉浸感	4.12	5.47	1.35	1.62	-3.45	16	0.003	*
臨場感	3.88	5.47	1.59	1.06	-6.15	16	0	**
參與度	4.29	5.29	1	1.37	-4.24	16	0.001	*

※ M 為平均數；$p < .05$ 為達顯著水準，$p < .01$ 為高度顯著水準，$p < .001$ 為極高度顯著水準

5. 結論

本研究旨在探討虛擬實境動畫教材對設計相關科系大學生在沉浸體驗與學習成效上的影響。透過《Lutaw》這部國際得獎之 360 度動畫短片，結合兩種觀看媒介（2D 螢幕與 VR 頭盔），以準實驗設計方式進行心理感受與學習成果的雙重評估。研究結果顯示，在沉浸感、臨場感、參與度與專注度等四個心理感受構面上，VR 頭盔觀看組顯著優於傳統 2D 螢幕組，顯示虛擬實境技術確實能在主觀體驗層面強化學習者之情境投入與感官包覆（$p < .05$）。此結果呼應 Slater 與 Wilbur (1997) 提出的沉浸感理論與 Makransky 與 Lilleholt (2018) 對沉浸式學習環境的正向影響之觀點。

然而，在學習成效方面（以影片細節專注觀察測驗得分為指標），兩組間並未呈現顯著差異（2D 螢幕組平均得分為 85.22，VR 頭盔組平均分為 85.29，$p=.986$，無顯著差異），顯示在短期知識掌握上，無論透過 2D 螢幕或是 VR 頭盔的動畫觀看體驗可能產生相近的認知成效。此發現與 Parong 與 Mayer (2018) 所指出之「VR 在學習興趣提升上具潛力，惟認知成效需視教學目標與情境策略調整」的觀點一致。

從社會學習理論 (Bandura & Jeffrey, 1973) 觀之，本研究設計以觀察影片細節為測驗基礎，與學習者在注意力與再現行為之間的連結相關。因此，雖 VR 能提升情境臨場感與心理沉浸，惟若教學評量仍著重於事實再現（如影片細節），則沉浸體驗未必直接轉換為答題成效。這提示未來在設計以 VR 為核心的教材時，應進一步思考如何將沉浸體驗轉化為學習表現的觸媒，例如導入互動式提問、情境角色扮演或反思性書寫等策略。

本研究初步實證 VR 動畫教材對學習體驗品質具有正向提升作用，特別在沉浸體驗與提升專注參與上具顯著成效，對設計科系學生而言更能呼應其視覺導向與敘事性思維需求。未來研究可延伸探討沉浸感與長期知識保留的關聯，並搭配腦波、眼動、行為紀錄等客觀量測工具，以建構更為全面之沉浸式學習評估模型。

參考文獻

Agrawal, S., Bech, S., Bærentsen, K., DeMoor, K., &Forchhammer, S. (2021). Method for Subjective Assessment of Immersion in Audiovisual Experiences. *Journal of the Audio Engineering Society*, *69*(9), 656–671. https://doi.org/10.17743/jaes.2021.0013

Bandura, A., &Jeffrey, R. W. (1973). Role of symbolic coding and rehearsal processes in observational learning. *Journal of Personality and Social Psychology*, *26*(1), 122–130. https://doi.org/10.1037/h0034205

Lengyel, A., Keczeli, D., Orosz, R., Bács, Z., Müller, A., Szőke, S., &Bába, É. B. (2021). Initial Validation of the Mindful Presence Scale: The Issue of the Construal Level of Scale Items. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.626084

Makransky, G., &Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

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

Rumjaun, A., &Narod, F. (2020). *Social Learning Theory—Albert Bandura* (pp. 85–99). https://doi.org/10.1007/978-3-030-43620-9_7

Scholz, J., &Duffy, K. (2018). We ARe at home: How augmented reality reshapes mobile marketing and consumer-brand relationships. *Journal of Retailing and Consumer Services*, *44*, 11–23. https://doi.org/10.1016/j.jretconser.2018.05.004

Slater, M., &Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

多智能体协作下的人智教研共生体建构与实践模式研究

Building and Implementing Model of a Human-AI Teacher Training Symbiotic System in Multi-Agent Collaborative Environments

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】 教师研修是促进教师专业发展的重要途径。然而传统教研模式存在效能低下与动态适应性不足等问题，难以满足教师深层次的专业发展需求。本研究通过解析教育多智能体系统的演进轨迹，基于共生理论提出涵盖知识创生、循证教研与共生实践的人智教研共生体理论模型，在基础上构建“共知—共论—共证—共创—共行”多智能体循证教研模式，实现教师隐性知识外显化、群体智慧协同化与实践证据动态化，为数智时代教研提供兼具理论深度与实践创新的人智协同教研新范式，助力教师专业发展与教学质量双重跃迁。
【关键词】 共生理论；教师研修；多智能体；人智教研共生体

Abstract: *Teaching Training is vital for teacher development, yet traditional models face inefficiency and rigidity. This study constructs a Human-AI Teaching Training Symbiotic System through multi-agent collaboration, grounded in symbiosis theory by analyzing educational MAS evolution. The model integrates knowledge co-creation, evidence-based research, and symbiotic practice, and proposes a five-phase framework—co-sensing, co-discussing, co-verifying, co-creating, and co-implementing—to externalize tacit knowledge, synergize collective wisdom, and dynamize practical evidence. This Human-AI collaborative paradigm offers theoretical-practical innovation for digital intelligence education, enhancing both teacher expertise and instructional quality.*
Keywords: symbiosis theory, teacher training, multi-agent system, human-AI teaching training symbiotic system

1. 引言

数智教育背景下，教师研修作为促进教师专业发展的重要途径，亟需突破传统模式的局限。当前教研普遍面临资源分配失衡、动态适应性不足、评价反馈滞后等困境，导致教研效能难以满足个性化与精准化需求（林敏等人，2023）。多智能体系统（Multi-Agent System，MAS）由多个智能体协同工作，其凭借分布式协作与自适应等特性，为解决教研中的复杂问题提供新路径（吴永和等人，2024）。然而，现有研究多聚焦单一智能体的技术实现，或囿于“人主导—机器辅助”的单向逻辑，未能充分释放人机协同的共生潜力（李海峰等人，2020）。

共生理论为重构人智关系提供了全新视角（方海光等人，2024），强调人与多智能体在教研过程中的动态性与协同演化性（李明蔚等人，2025）。其核心在于打破主客二元对立，通过“人-智”双向互塑机制，实现知识生产从个体孤立向群体共生的跃迁。如何基于共生理论构建人智教研共生体模型，推动教师知识的建构与外化，成为破解教研效能瓶颈的关键。

基于此，本研究在厘清教育领域多智能体系统演进路径的实质基础上，构建人智共生体理论模型，提出“人智共知、人智共创、人智共论、人智共行和人智共证”的循证教研模式，旨在为数智时代教研协同创新提供理论支撑与实践范式，助力教师专业发展与教育质量的双重跃迁，为新时代教研转型注入新动能。

2. 多智能体系统在教育领域的演进路径

当前，多智能体系统在教育领域广泛应用，已成为应对复杂问题的重要方法（吴永和等人，2024）。该系统源自分布式人工智能领域，展现出在解决复杂分布式任务时的卓越智能与灵活性。多智能体系统注重多个智能体间的协同合作，以共同实现目标或完成任务。随着人工智能技术历经计算、感知至认知阶段的演进，并受其理论基础的深刻影响，多智能体系统在教育领域展现出了独特的发展轨迹。

2.1. 行为主义视域下计算智能对个性化教研的技术实现

21 世纪初，智能教学代理作为计算智能在多智能体系统（MAS）的典型应用，在教育领域兴起，用以解决复杂问题（Sun et al.,2007），标志着 MAS 方法论从理论到实践的转变。基于计算智能的多智能体系统受到了行为主义心理学理论的深刻影响，尤其是斯金纳提出的程序教学理念（徐振国等人，2021）。该系统将教学过程抽象为预定义脚本操作（刘清堂等人，2023），将人类学习行为简化为机械式的操作模式，从而形成“人似机械”的隐喻（喻国明等人，2024）。这类系统通过“教师数字画像”标签体系实现个性化教研资源推荐（胡小勇等人，2019），虽在提升教研效率上取得进展，但受限于机械论世界观（陈万球等人，2024），系统缺乏情境感知能力，难以支持高阶思维培养（李海峰等人，2020）。尽管自适应技术部分缓解了路径僵化问题（郝祥军等人，2023），但其理论根基仍囿于行为主义与认知主义范式，亟待引入新视角以突破技术局限，实现教研的动态性与多元性跃迁。

2.2. 人本主义视域下感知智能对细粒度教研的多维支持

在智能感知技术发展中，学者们推动“机器拟人化”研究，旨在使教学智能体具备“人性化”特征，以服务人类为中心（俞鼎等人，2024）。教育智能体的拟人化不仅涉及外观和行为仿真（Wu et al.,2023），还包括情感、感知、欲望和意志的复杂心理模拟（Kühne et al.,2023）。随着虚拟数字人技术进步，虚拟教师依托文本、语音、手势等多模态技术增强社会文化功能（翟雪松等人，2023），实体教育机器人则通过具身交互提升教学临场感（黄荣怀等人，2023）。智能体借助视觉、听觉和触觉传感器与深度神经网络实现教研情境的细粒度感知（刘峰等人，2024），并结合边缘计算与 5G 网络支持实时数据分析（赵晓伟等人，2022），为教师提供数字素养评估、教学能力诊断等循证研训服务（刘邦奇等人，2023）。该技术范式通过多模态数据整合与情感认知分析，推动教研从传统的单模态、粗粒度模式转变为多模态、细粒度模式，但需进一步解决人智协同中人智关系与认知交互等深层问题。

2.3. 共生理论视域下认知智能对创生性教研的有利互补

后现代哲学家哈拉维（Donna Haraway）提出“赛博格”（Cyborg）的概念，揭示了人智交互的共生本质——人与智能体不再局限于主客二元对立，而是作为“同伴物种”共同构建世界（薛博文等人，2023；宋美杰，2024）。这种共生关系重塑了教育主体性，智能体通过认知智能的演进，不仅受人类价值观与需求驱动，同时反向塑造人类行为模式与社会结构，推动教育目标从“知识传递”向“知识创新”跃迁（余胜泉等人，2023）。基于语言大模型的多智能体系统通过模拟人类群体协作解决问题，共同参与规划、讨论和决策（Guo et al.,2024），从而促进知识创生，为教育领域的知识创新提供了新的可能性。在虚拟教研室中，智能体成为

协同教研的重要参与者（孟凡丽等人，2023），与人类教师共同推动教学知识的挖掘与创新。这种“人-智”协同共生的教研模式，引发教师教育目标、课程、模式、评估等的深刻变革，为教师专业发展的培训、教研和自主学习带来全新变化（林敏等人，2024）。此外，研究者们还在探索利用多智能体系统模拟多领域专家协同解决教育领域的复杂问题，并与传统单智能体对话模式进行对比，为未来教育领域的人智协同发展提供新的思路和实践经验（翟雪松等人，2024）。

3. 人智教研共生体的理论模型构建

3.1. 理论基础：共生理论

共生理论强调不同主体之间的相互依存与协同发展（吴磊等人，2024），这与在人智教研共同体中人类教师与智能体各自具有独特优势，需要通过合作实现优势互补的理念高度契合。在探讨人机协同及其在智力活动中的应用时，利克莱德（Licklider）于 1960 年提出了一个创新性的观点，即借鉴共生理论来构建人类与机器之间的互动模式。他认为，通过这种合作性的共生关系，可以显著提高处理复杂问题的智力效能。这一见解颠覆了传统认知，指出在智力活动中，人类并非唯一的主导者，而是与机器形成了一种认知上的共生伙伴关系（方海光等人，2024）。基于共生理论视角，人智教研共生体解析为：该共生体由“人-智”这一对共生单元构成，它们相互作用，共同担任教育教学的主体。在这一共生体系中，两个主体不断地进行物质、能量和信息的交流，遵循一种互惠共生的模式，共同推进知识的探索、创造与传播过程。因此，基于共生理论构建人智教研共生体的理论模型，以此来解析和指导实践。一方面，多智能体通过分析人智交互产生的数据，结合自身反馈机制，能够“学习”并优化其算法，进而提升智能水平，从而在服务提供方面更加精准高效。另一方面，人类学习者可以借助机器智能的输出，获得新的认知视角和解决问题的方法，进而扩展自己的知识边界和认知能力。这种互惠模式使得人类能够从多智能体的智能进化中受益，实现个人能力的提升。该模式下的“人-智”共生关系超越了传统的单向利用，形成了一种动态的、互为促进的智能协同模式（郝祥军等人，2023）。

3.2. 建构路径：数字循证教研

人智教研共生体的构建路径需依托大模型支持的数字循证教研来实现。这种模式强调将教育实践建立在经过验证的证据之上，教师不断地通过问题的澄清、证据的筛选、评估、验证和迭代，与同事和专家进行协商，从而突破和重组现有的经验（吴雨宸等人，2023）。教师在循证实践中面临挑战，如难以获取和利用证据、缺乏有效的反馈和评价机制，以及实践过程中指导和支持不足。而以大模型为代表的数字循证教研则进一步结合了经验与数字技术，强调以人的发展为核心，其基本结构由技术影响下的教研主体、教研证据和教研情境共同构成（张思等人，2024）。有学者通过实际案例研究，证明了该行动框架能显著提升师范生的循证实践能力和集体效能感（褚乐阳等人，2024）。因此，人智教研共生体的构建路径需要充分利用大模型支持的数字循证教研，以克服传统教研中的挑战，其核心目标是促进教师实践性知识的生成与外化，使得教师的实践性知识得以在行动中不断生成和显现，实现了教师专业发展和教学质量提升的双重目标（吴雨宸等人，2023）。

3.3. 目标指向：教师知识的建构与外化

在当前教育改革的大背景下，教师知识的建构与外化显得尤为重要。隐性知识与显性知识之间的动态转化机制为理解人智教研共生体中教师知识的生成、交流与融合过程提供了理论支持。已有研究指出，人机协同教研能够激发教师的教学反思，这是新教学理念和方法生

成的关键（蔡慧英等人，2023），数据驱动的教研活动被证明是教师知识精准生成与有效外化的有力工具（余胜泉等人，2019）。在协同知识建构领域，研究者们已经取得了一系列成果。例如，Harasim（1998）等学者强调了协作学习中的知识建构过程，包括主题探索、意义协商和观念形成等关键环节。Gunawardena（1997）的交互分析模型进一步细化了知识建构的五个阶段，为教师协同知识建构的教研提供清晰的指导路径。国内研究者（刘清堂等人，2022）在 Stahl 模型（Stahl，2000）的基础上，结合我国教育实际，构建了适用于教师工作坊的协同知识建构模型。此外，在探讨教师知识建构与外化的过程中，研究主要关注知识状态转化的机制。Nonaka 和 Takeuchi 提出的知识创造螺旋模型（Nonaka & Takeuchi，1995），揭示了隐性知识与显性知识之间的动态转化，涵盖了社会化、外化、组合和内化四个连续阶段。在此基础上，研究者进一步细化了知识转化的步骤，将其分为共感知识转化、概念知识生成、系统知识构建和操作知识内化四个阶段（李海峰等人，2020）。这一理论框架不仅揭示了知识转化的深层次逻辑，也为教师在人机协同教研环境中如何有效建构和外化知识指明了路径。

3.4. 理论模型：人智教研共生体模型

人智教研共生体模型是一个复杂而精细的人智关系网络，它深刻地揭示了人与智能体在教研过程中的共生与协同效应。模型主要包含三个层次：本质层——人智知识创生层次，路径层——人智循证教研层次，以及实践层——人智共生实践层次。这三个层次之间存在着密切的内在联系，共同构成了模型的完整框架。

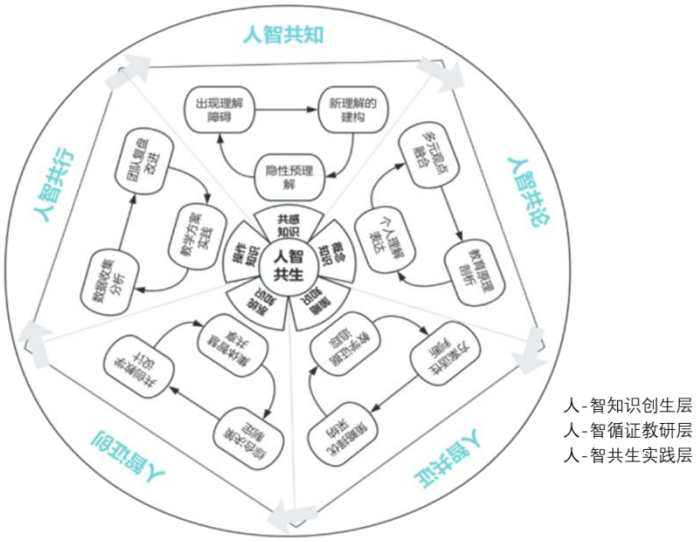


图 1 人智教研共生体模型

人智知识创生层次是模型的核心和本质。它关注的是人与智能体协同过程中教师知识的生成、交流与融合。通过共感知识、概念知识、策略知识、系统知识和操作知识五个环节的深化与拓展，人智知识创生层次为人智共生提供了坚实的理论基础。

人智循证教研层次是连接人智知识创生层次与人智共生实践层次的桥梁。它通过个人理解与建构、观点表达与剖析、方案探讨和评估、疑惑澄清和解决、计划实施和效果检验五个环节，将理论知识与实践活动紧密结合起来。这一层次不仅确保了教研活动的科学性和有效性，还为人智共生实践层次的实践提供了有力的支持。

人智共生实践层次是模型的实践层面，也是人智教研共生体最终追求的目标。它强调人与智能体在实践中的协同与合作，通过人智共知、人智共创、人智共论、人智共行和人智共

证五个环节，推动教研活动的创新与发展。这一层次不仅体现了人与智能体在教研过程中的共生关系，还展示了他们共同推动教育事业进步的巨大潜力。此外，模型还强调了两个主体——人和智能体的相互作用，以及一个核心机制——人智互惠机制。这两个要素共同构成了模型的灵魂，确保了模型在实践中的可行性和有效性。

4. 共生理论视角下多智能体循证教研模式构建

共生理论视角下的多智能体循证教研模式突破传统教研线性思维，基于“人-智”协同的多主体互惠机制，结合教研共生体结构和“人-智”角色分工，将教研过程解构为“人智共知、人智共论、人智共证、人智共创和人智共行”的“五共”链路，形成动态循环的证据生成与应用生态。模式通过“五共”递进实现教研效能跃升，为新时代教研转型提供兼具理论深度与实践张力的创新框架（如图 2 所示）。

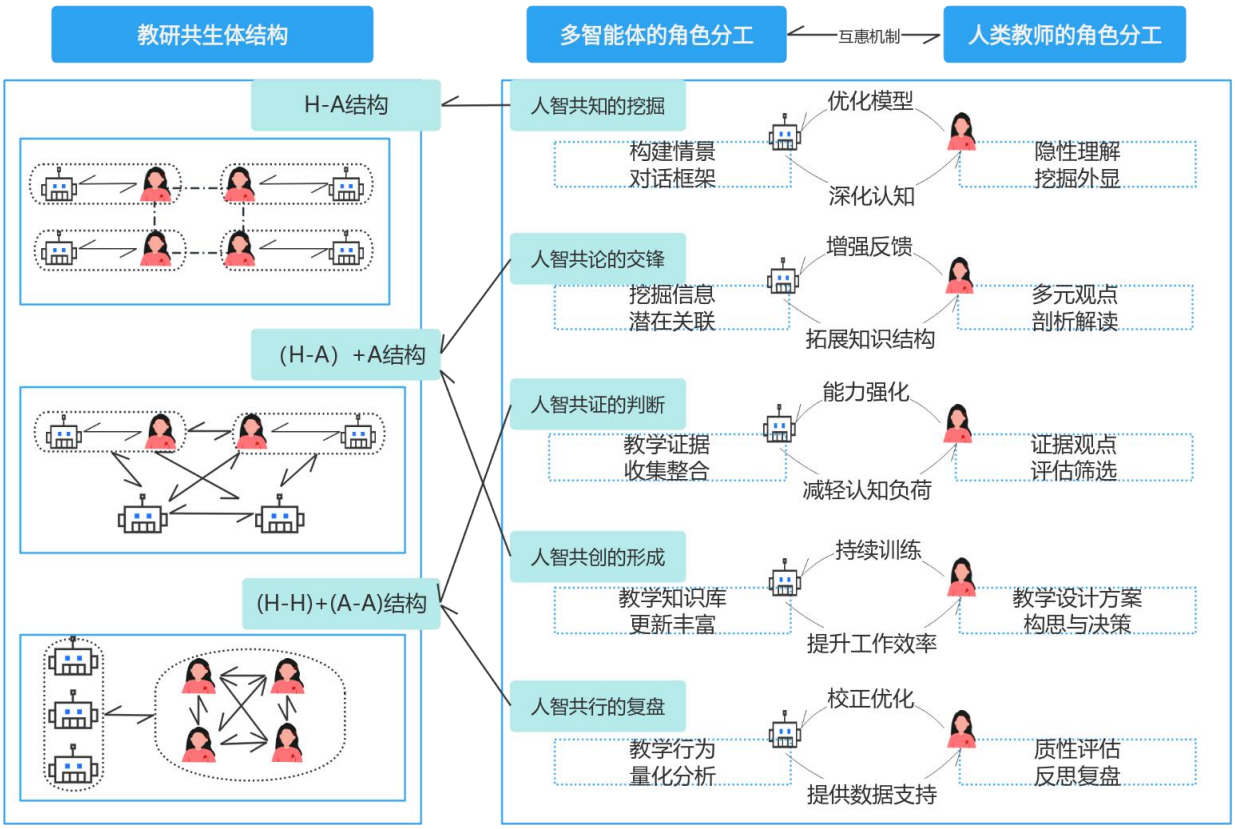


图 2 多智能体循证教研模式

4.1. 人智共知，教师个人隐性理解的挖掘

在人智共知阶段，首要环节聚焦于隐性认知的深度挖掘。在此过程中，智能体作为情境营造者，创造了一个支持性的学习环境，鼓励教师分享他们在教学过程中的直觉和经验。教师的角色转变为隐性知识的激活者，他们在智能体提供的情境中，开始意识到并尝试表达那些不易言说的教学见解（Guo et al.,2024）。随后，进程过渡到对话催化与灰色地带知识的构筑阶段。此时，智能体转变为对话的催化媒介，通过精心设计的提问策略与案例分享，有效触发教师的批判性思维与反思活动，促使他们主动承担起灰色知识构建者的角色（吴永和等人，2017）。这一转换旨在将先前隐性的认知元素转化为可探讨的灰色知识领域，鼓励教师尝试以语言为载体，对这些认知进行界定、阐述乃至优化，从而增强其表述性。最终，步入

反馈引导与隐性认知的精确定义阶段，智能体担任起反馈供给者的职责，通过细致入微的观察与即时有效的反馈机制，助力教师作为隐性认知的重塑者，深入剖析并明晰其隐性认知中的模糊边界与不确定性因素。

4.2. 人智共论，教师群体思想观点的交锋

初始阶段，智能体扮演着“教学诊断专家”的角色，通过自然语言处理，识别教师教学设计观点中的亮点与不足，并提出针对性的优化建议。帮助教师将个人较为抽象的教学构想转化为清晰、具体的教学方案，还通过智能体的算法优化，对教学内容进行精准提炼和重组，使得教学思考更具学术深度，同时也提升了教学交流的质量和效率。智能体在识别和比较教研组教师中多元教学观念的过程中，发挥了“观念桥梁”的作用（胡小勇等人，2019）。它通过对教师们的不同教学观点和策略进行综合分析，生成一份包含观点对比、优势融合建议以及潜在改进方向的综合性分析报告。教研组教师通过细致研读，能够更容易地找到共识点，引领教研讨论向更深层次推进，促进不同教学观念的相互借鉴与融合。最终，针对教学问题的深入探究和教育原理的阐释，智能体通过“议题提炼器”的功能，从教师群体的热烈讨论中捕捉并提炼出核心议题（徐振国等人，2021）。这一过程涉及智能体对讨论内容的深度理解和关键词汇的频繁出现分析，从而准确捕捉到教师们最关心的研究焦点。随后，智能体将这些核心议题转化为具体、可操作的教研问题，为教师的研究工作指明了方向。教研组教师则运用其深厚的专业知识和丰富的实践经验，对这些研究问题进行深度剖析，揭示出教育原理的深层含义，并据此制定出既符合教育规律又具创新性的教学策略或研究方案，确保了教研活动在理论与实践层面都能达到高度的深度与实效性。

4.3. 人智共证，教学证据迁移适用的判断

在方案探讨与评估流程中，人类教师与智能体的协同作业展现出高度的互补性和效率。智能体利用其卓越的数据处理能力，首先通过自动化的数据预处理步骤—包括清洗、分类及归纳—为人类教师提供了一个全面的数据支撑和参考框架。这样，人类教师便能够在此基础上，深入解析教学案例与实证材料，将他们的理论知识和实践经验转化为具有指导意义的信息和洞见。随着评估流程的深入，人类教师需全面考量教学方案的可行性、创新性、学生接纳度以及与教学目标的契合度等多个维度。在此关键阶段，智能体通过机器学习等技术进行动态监控和深度分析，基于收集到的评估数据，智能体能够自动生成多样化的教学方案备选，并提出针对性的改进建议。最终，在择优采纳与共识形成环节，智能体所提供的数据分析结果为人类教师的决策提供了客观且量化的依据。人类教师据此与团队成员展开深入讨论，通过细致的权衡和必要的妥协，综合各方意见和需求，确保了最终确定的教学方案既优化又具广泛共识。

4.4. 人智共创，教学方案共识决策的形成

在智能体与人类教师协作的教学模式中，智能体承担着收集、分析大量数据，识别教学中的模式和趋势，自动整理和更新知识库，以及维护共享平台的运行和信息安全等任务，而人类教师则提供专业见解和教学经验，参与讨论以提炼集体智慧，对智能体整理的知识进行审核和补充，确保内容的准确性和适用性。在综合意见与决策制定方面，智能体通过算法分析团队成员的意见，提供决策支持，预测不同决策的潜在结果，以及优化决策方案，而人类教师则负责最终的决策制定，利用智能体提供的分析结果，结合教育理论和实际教学经验，做出符合教学目标和价值的判断。在合作开发教学材料的过程中，智能体协助人类教师进行教学材料的初步编写和格式化，提供个性化学习内容的建议，以及自动化的课件制作和编辑

功能，而人类教师则负责教学材料的创意构思和内容把关，确保材料符合教学大纲和教育目标，同时融入人类教师的情感和教学艺术，提升教学材料的互动性和吸引力。

4.5. 人智共行，教学实践成效反馈的复盘

在实践检验阶段，智能体负责自动化监测教学活动的执行，确保新教学方案的一致性和标准化实施，并通过算法分析学生的互动模式，为教师提供学生学习行为的洞察，同时实时调整学习路径以适应学生的个性化需求；而人类教师则专注于观察学生的非言语反馈，如情感状态和参与度，根据智能体的反馈进行现场调整以优化教学互动，并评估智能体调整后的学习路径是否与教学目标一致。在数据收集与分析阶段，智能体执行大规模的数据收集任务，运用数据分析技术识别教学效果的关键指标，并输出数据报告为教师提供决策支持；人类教师则确定数据收集的重点领域，结合专业判断深入解读智能体提供的数据报告，并利用分析结果制定针对性的教学改进措施。在团队复盘与持续改进阶段，智能体提供一个结构化的复盘框架，记录和整理复盘会议内容，并根据复盘结果自动生成改进计划的初步草案；人类教师引导复盘会议，确保讨论内容覆盖教学所有关键方面，分享个人经验以促进团队知识共享和共识形成，最终基于团队讨论和智能体草案制定并执行具体的教学改进计划。通过这种互补的分工，智能体和人类教师共同推动教学质量的提升，实现教学活动的最优化。

5. 结论

本研究基于共生理论构建了“知识创生—循证教研—共生实践”三层框架的人智教研共生体理论模型，以及通过“共知—共论—共证—共创—共行”五阶段协同链路实现教师隐性知识外显化与群体智慧协同化的“多智能体循证教研模式”，为数智时代教研协同创新提供了兼具理论深度与实践创新的人智协同范式。本研究构建的理论模型的实践效度有待验证、模型场景适应性有待探讨以及实践落地的协同条件存在阈值效应等局限。未来需进一步聚焦人智协同教研的动态建构机制，修正完善理论模型；探索复杂教育场景（如乡村、跨学科教研）中多智能体系统的适应性优化路径；开发融合质性数据与量化指标的综合评价体系；推动多校联动的教研共同体生态建设，突破协同条件阈值，以实现人智协同教研的规模化落地与可持续发展。

参考文献

方海光,孔新梅,刘慧薇 & 王显闯.(2024).基于共生理论的人机协同教育主体合作博弈及其优化策略研究.电化教育研究,45(01),21-27.

刘清堂,曹天生,吴林静 & 郑欣欣.(2023).ChatGPT 类耦合教学代理：需求分析与教学应用.现代远距离教育,(06),3-11.

刘峰,刘雅旋,柴新宇,季昊涵 & 郑之行.(2024).智能教育中可计算感知技术：系统性综述.计算机科学,51(10),10-16.

吴永和,姜元昊,陈圆圆 & 张文轩.(2024).大语言模型支持的多智能体：技术路径、教育应用与未来展望.开放教育研究,30(05),63-75.

吴永和,刘博文 & 马晓玲.(2017).构筑“人工智能+教育”的生态系统.远程教育杂志,35(05),27-39.

吴磊,白青贺 & 范丽鹏.(2024).共生理论视域下在线协作学习机制与互动模式分析.现代远距离教育,(04),60-70.

吴雨宸,宋萑 & 徐兴子.(2023).循证教研：指向教师实践性知识生成的教研转型.教师教育研究,35(01),13-19.
陈万球 & 齐琦.(2024).人机关系的视差：智能时代的主体性重构.湖南师范大学社会科学学报,53(03),95-101.
李海峰 & 王炜.(2020).人机学习共生体——论后人工智能教育时代基本学习形态之构建.远程教育杂志,38(02),46-55.
李明蔚,张思,张津铭,张雯颖 & 童名文.(2025).人机协同视域下的教师研修:价值内涵、表征形态与实践进路.电化教育研究,(04),93-98+113.
张思 & 张津铭.(2024).数字循证教研:内涵、理念与推进路径.中国远程教育,44(09),54-64.
宋美杰.(2024).走出人类世：作为同伴物种的人工智能与跨物种交往.新闻界,(07),14-25.
余胜泉 & 汪凡淙.(2023).人工智能教育应用的认知外包陷阱及其跨越.电化教育研究,44(12),5-13.
余胜泉 & 王琦.(2019).“AI+教师”的协作路径发展分析.电化教育研究,40(04),14-22+29.
何皓怡,刘清堂,张思,关雪敏 & 覃伟华.(2022).在线协同研讨知识建构行为的自动分析研究.现代教育技术,32(05),93-101.
林敏,吴雨宸 & 宋萑.(2024).人工智能时代教师教育转型：理论立场、转型方式和潜在挑战.开放教育研究,30(04),28-36.
孟凡丽,马翔 & 王建虎.(2023).AIGC 视域下的虚拟教研室：概念特征、运行要素与建设进路.现代远距离教育,(04),14-21.
胡小勇 & 林梓柔.(2019).精准教研视域下的教师画像研究.电化教育研究,40(07),84-91.
郝祥军,张天琦 & 顾小清.(2023).智能时代的人机协同学习：形态、本质与发展.中国电化教育,(10),26-35.
俞鼎 & 李正风.(2024).生成式人工智能社会实验的伦理问题及治理.科学学研究,42(01),3-9.
徐振国,刘志,党同桐 & 孔玺.(2021).教育智能体的发展历程、应用现状与未来展望.电化教育研究,42(11),20-26+33.
黄荣怀,刘德建,阿罕默德·提利利,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(06),4-14.
喻国明,黄哲浩,唐子璇 & 梁禧.(2024).作为行动者的智能体：重构人类关系的未来范式——NBIC 会聚技术视角下的考察与分析.新闻界,(08),46-54+78.
褚乐阳,刘泽民,王浩 & 陈向东.(2024).大模型支持的教师循证实践：行动框架与案例应用.开放教育研究,30(04),91-103.
蔡慧英,卢琳萌 & 顾小清.(2023).人机协同教研会促进教师教学反思能力的发展吗?——基于课堂视频智能分析技术的实证研究.现代远距离教育,(01),40-49.
翟雪松,吴庭辉,李翠欣,仇婷婷 & 李艳.(2023).数字人教育应用的演进、趋势与挑战.现代远程教育研究,35(06),41-50.
翟雪松,季爽,焦丽珍,朱强 & 王丽英.(2024).基于多智能体的人机协同解决复杂学习问题实证研究.开放教育研究,30(03),63-73.
薛博文 & 冯建军.(2023).人类世到赛博格：后人类主义的教育反思与重构.现代大学教育,39(06),13-21.
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.

Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N. V., ... & Zhang, X. (2024). Large language model based multi-agents: A survey of progress and challenges. arXiv preprint arXiv:2402.01680.

Harasim, L. M. (1989). Online education: A new domain.Praeger Publish-ers.

Kühne, R., & Peter, J. (2023). Anthropomorphism in human‐robot interactions: a multidimensional conceptualization. Communication Theory, 33(1), 42-52.

Nonaka, I. & Takeuchi, H.(1995). The Knowledge Creating Firm.New York:Oxford University Press.

Stahl, G. (2000). A model of collaborative knowledge-building. MIT Press.

Sun, S., Joy, M., & Griffiths, N. (2007). The use of learning objects and learning styles in a multi-agent education system. Journal of Interactive Learning Research, 18(3), 381-398.

Wu, M., Wang, N., & Yuen, K. F. (2023). Deep versus superficial anthropomorphism: Exploring their effects on human trust in shared autonomous vehicles. Computers in Human Behavior, 141, 107614.

多智能体系统如何塑造未来教师？——AI-TPACK 素养提升的实验研究

How Do Multi-Agent Systems Shape Future Teachers? — An Experimental Study on AI-TPACK Competency Development

陈小娇[1]，陈思[1*]，王成梁[2]，王浩名[2]，
[1]浙江工业大学 教育学院
[2]华东师范大学 教育学部
[*] sichen@zjut.edu.cn

【摘要】 人工智能（GAI）正推动教师教育向智能化转型，但职前教师在教学设计中仍面临 AI 工具整合能力不足、教学方法选择受限、教学设计创新性不足等挑战。如何借助智能系统提升其 AI-TPACK 素养，已成为教师教育领域亟待解决的问题。本研究构建了一个基于多智能体系统（MAS）的教学设计支持工具，其中包括学科知识智能体（CK-Agent）、教学方法智能体（PK-Agent）和教学工具智能体（TK-Agent），协同支持教师的教学设计过程。实验招募 20 位职前教师，随机分为实验组（使用 MAS）和对照组（使用传统工具），要求他们完成教学设计任务，并基于 AI-TPACK 框架进行教学方案评估。此外，实验组在干预前后进行了 AI-TPACK 素养测评。研究结果表明，MAS 能有效提升职前教师的教学设计质量，并显著促进 AI-TPACK 素养的发展。本研究的贡献在于验证了 MAS 在教师教育中的可行性，拓展了 AI-TPACK 的应用场景，并为未来智能教师培训系统的设计提供了实证依据。
【关键词】 多智能体系统；AI-TPACK 素养；职前教师；教学设计；人工智能

Abstract: *Artificial intelligence (GAI) is driving the intelligent transformation of teacher education, but pre-service teachers still face challenges in integrating AI tools, selecting teaching methods, and fostering innovation in instructional design. Addressing how intelligent systems can enhance their AI-TPACK literacy has become a critical issue in teacher education. This study developed a multi-agent system (MAS)-based teaching design support tool, which includes subject knowledge agents (CK-Agent), teaching method agents (PK-Agent), and teaching tool agents (TK-Agent), working collaboratively to support the teaching design process. The experiment involved 20 pre-service teachers, randomly assigned to an experimental group (using MAS) and a control group (using traditional tools), who were tasked with completing teaching design tasks and evaluating their teaching plans based on the AI-TPACK framework. Additionally, the experimental group underwent pre- and post-intervention AI-TPACK literacy assessments. The results showed that MAS significantly improved the teaching design quality and promoted AI-TPACK literacy development among pre-service teachers. The contribution of this study lies in validating the feasibility of MAS in teacher education, expanding the application of AI-TPACK, and providing empirical evidence for the design of future intelligent teacher training systems.*
Keywords: Multi-Agent System; AI-TPACK Literacy; Pre-service Teachers; Instructional Design; Artificial Intelligence

1.前言

人工智能（AI）正以前所未有的速度变革教育生态，其在教师专业发展中的应用尤为关键。在智能教育环境下，教师不仅需要掌握学科内容（CK）和教学方法（PK），更需要熟练整合 AI 技术工具（TK），以提升教学设计质量，实现个性化与智能化教学（张群 & 田格格, 2023）。AI-TPACK 框架的提出正是对这一时代需求的回应，它在传统 TPACK 模型的基础上，强调 AI 技术在教学决策中的深度嵌入。然而，现有研究表明，职前教师在 AI 工具整合、教学方法选择及教学设计创新等方面仍存在显著不足，导致教学方案的适切性与智能化程度受限（王玮, 2020）。如何通过智能化手段提升职前教师的 AI-TPACK 素养，已成为当前教师教育领域亟待解决的核心问题。

多智能体系统（MAS）作为一种分布式人工智能架构，凭借其协同推理、自主决策和动态适应能力，在教育领域展现出广泛的应用前景(Hava & Babayiğit, 2025)。MAS 能够模拟复杂的教学设计过程，动态分析教师的教学需求，并提供针对性的智能支持。因此，将 MAS 应用于职前教师的教学设计训练，或能有效弥补其 AI-TPACK 素养的短板，助力未来教师适应智能教育环境的变革。

尽管 AI-TPACK 框架已为教师教育提供了理论指导，但如何在实践中提升职前教师的 AI-TPACK 素养仍然缺乏有效的教学支持工具(Karataş & Ataç, 2024)。本研究提出并构建了一种基于 MAS 的智能教学设计支持系统，旨在通过学科知识智能体（CK-Agent）、教学方法智能体（PK-Agent）和教学工具智能体（TK-Agent）的协同作用，优化职前教师的教学设计过程。本研究拟通过实验验证 MAS 在教师教育中的可行性，并探讨其对 AI-TPACK 素养提升的影响，为未来智能教师培训系统的构建提供实证支持。

RQ1：MAS 是否能有效提升职前教师的教学设计能力？

RQ2：使用 MAS 进行教学设计是否能促进职前教师的 AI-TPACK 素养发展？

RQ3：实验组（MAS 支持）与对照组（传统工具）在教学设计质量上是否存在显著差异？

2.文献综述

AI 在教师教育中的应用主要体现在智能导师系统、智能反馈机制及个性化学习支持等方面。例如，智能导师系统能够实时分析教师的教学行为，为其提供精准的教学建议（Park et al., 2023）；智能反馈机制则可通过大数据分析学生的学习轨迹，辅助教师优化教学策略（Chen et al., 2024）。此外，研究表明，AI 驱动的个性化学习支持能够有效提高教师的教学适配能力，使其在教学过程中更灵活地调整教学方法（Jin et al., 2024; Zhang et al., 2024）。然而，目前关于 AI 如何具体增强教师的 TPACK 素养，特别是 AI-TPACK 素养的研究仍较为有限。

MAS 由多个自主智能体（Agent）协同工作，能够模拟复杂的学习和教学环境，已在智能辅导、协作学习等领域得到了广泛应用。例如，智能辅导系统基于 MAS 架构，可提供动态的学习路径推荐（Johnson et al., 2014）；协作学习系统中的智能体能够调节小组互动模式，优化协同学习体验（Dillenbourg et al., 2011）。在教师教育领域，MAS 的应用仍处于探索阶段，但已有研究显示，MAS 能够有效支持教师的决策过程，提高其教学设计的精准性和创新性（Pugh et al., 2016）。本研究在此基础上，提出以 MAS 支持职前教师的教学设计，以弥补现有 AI-TPACK 素养培养方式的不足。

AI-TPACK 框架基于传统 TPACK 模型，扩展了技术知识（TK）的定义，将 AI 工具的应用纳入教师核心素养范畴（Mishra & Koehler, 2006）。该框架强调，教师不仅需要掌握学科知识（CK）和教学方法（PK），更需理解 AI 技术的教育价值，并能在教学设计中恰当地整合 AI 工具（TK）。现有研究表明，通过系统性的培训和实践，教师能够显著提升 AI-

TPACK 素养，从而优化其教学方案的智能化程度（Koehler et al., 2013）。然而，如何通过智能技术系统化地提升职前教师的 AI-TPACK 素养，仍是当前研究的空白。

3.研究方法

3.1 多智能体系统设计

本研究构建了一个基于 MAS 的教学设计支持工具，以提升职前教师的 AI-TPACK 素养。该系统由学科智能体（CK-Agent）、教学法智能体（PK-Agent）和教学工具智能体（TK-Agent）协同工作，通过智能化手段优化教学设计能力。其中，学科智能体基于自然语言处理（Natural Language Processing, NLP）技术，能够自动分析课程内容，提取核心知识点，并生成知识点图谱，以帮助教师快速理解学科内容结构并优化教学内容设计。教学法智能体通过集成教育学理论与人工智能算法，提供包括探究式、项目式和互动式在内的多种教学策略，以支持教师根据教学目标和学习者特点匹配适合的教学方法。教学工具智能体则依据课程需求和教学设计目标，推荐合适的 AI 工具、教育软件及数据分析工具，以促进 AI 在教学中的合理整合和应用。

为了提高多智能体系统的协同工作效率，本研究开发的 MAS 采用 Mixture of Experts（MoE）架构，利用强化学习的优化策略，实现智能体之间的动态任务分配。系统通过持续的学习和优化，为职前教师提供个性化的教学设计建议，并动态调整不同智能体的任务权重，以优化教学设计的整体效果。**图 1** 展示了本研究构建的多智能体协作架构。

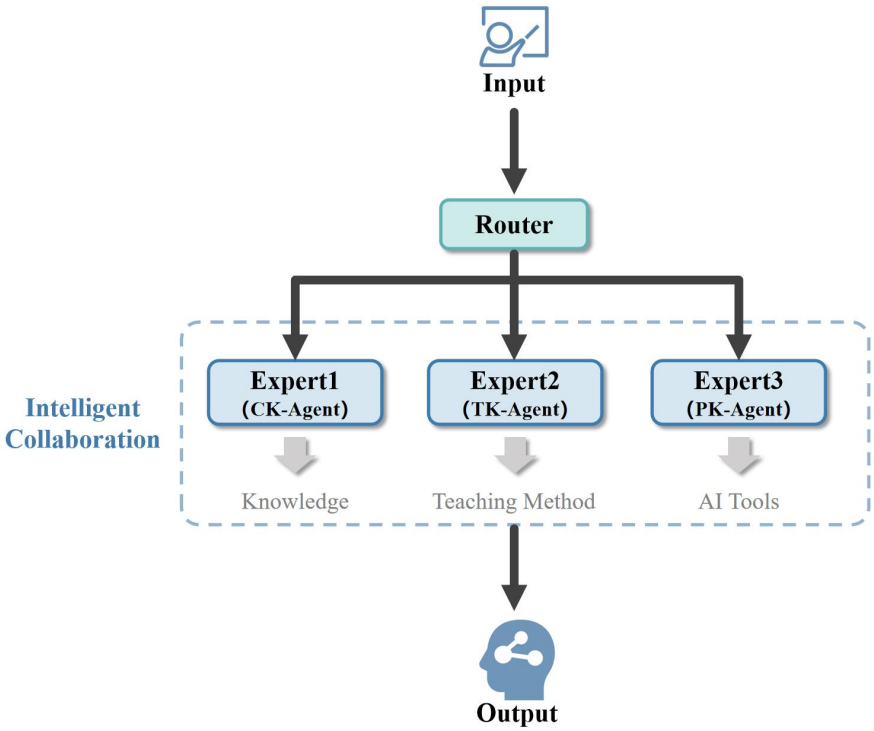


图 1 多智能体协作架构图

为了进一步提升 MAS 在不同教学情境下的泛化能力与个性化适应性，本研究引入元学习（Meta-Learning）机制，使智能体能够通过有限的训练样本迅速调整模型参数，以优化教学设计支持的精准度与个性化水平。

具体而言，MAS 采用 Model-Agnostic Meta-Learning（MAML）算法（Finn et al., 2017），以 CK-Agent、PK-Agent 和 TK-Agent 为元学习任务的优化对象。假设在多个教学任务 T_i 上进行元训练，每个任务包含一个数据集 $D_i=(X_i, Y_i)$，MAML 通过以下双层优化机制进行训练：

（1）内层优化（Inner Loop）：对于每个任务 Ti，智能体基于当前参数θ进行梯度下降更新：

$$\theta_i' = \theta - \alpha \nabla_\theta L_{Ti}(\theta)$$

其中，α为学习率，LTi 为任务 Ti 上的损失函数。

（2）外层优化（Outer Loop）：在多个任务上进行参数聚合，使模型具备快速适应新任务的能力：

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_i LT_i(\theta_i')$$

其中，β为元学习率（meta-learning rate）。

结合 MoE 框架，智能体可通过任务级专家分配不同权重，并通过强化学习机制进行持续优化，使其动态调整教学设计建议，提高 MAS 在 AI-TPACK 素养提升中的应用价值。如**图 2** 所示，该优化框架能够同时提升 MAS 的泛化能力、个性化适应性和动态优化能力，为未来的智能教师培训系统奠定基础。

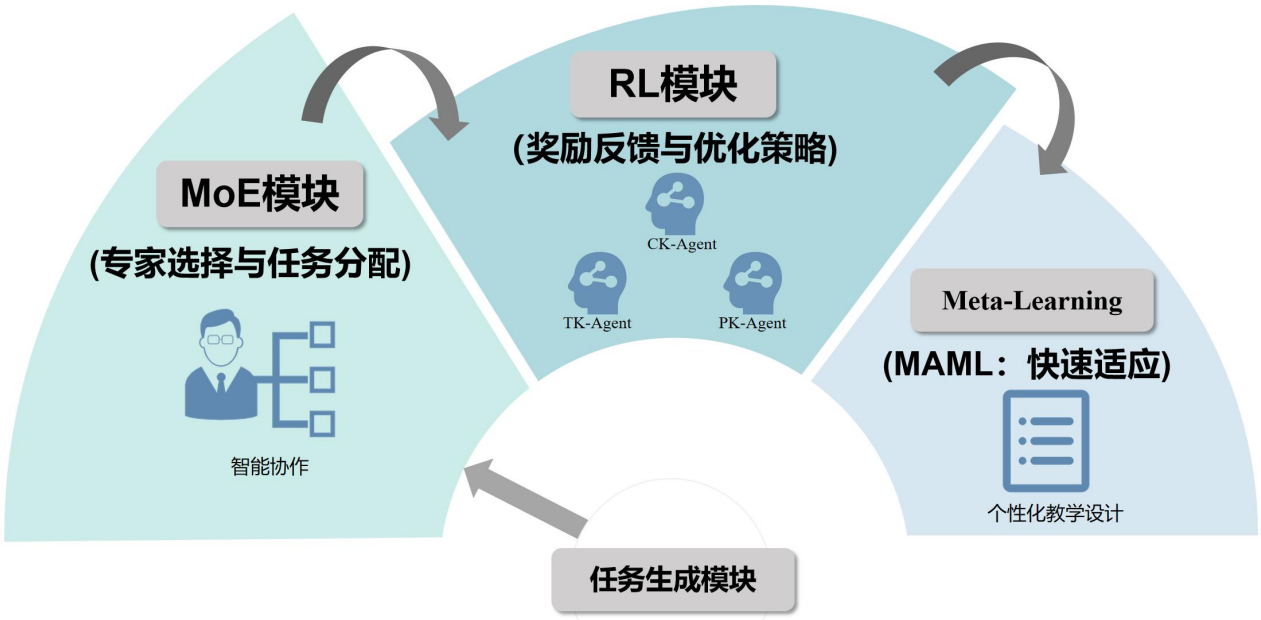


图 2 MoE + RL + Meta-Learning 协同架构

3.2 实验设计

本研究采用实验对照设计，以探究 MAS 对职前教师 AI-TPACK 素养的影响。研究对象为 20 名职前教师，随机分为实验组和对照组，每组各 10 人。实验组使用 MAS 进行教学设计，而对照组则使用传统教学设计工具（如 PPT、Word 等）。实验任务要求所有被试完成一项完整的教学设计方案，包括教学目标设定、知识点梳理、教学方法选择和教学工具整合等环节。通过这种方式，可以考察 MAS 在支持职前教师教学设计方面的有效性。

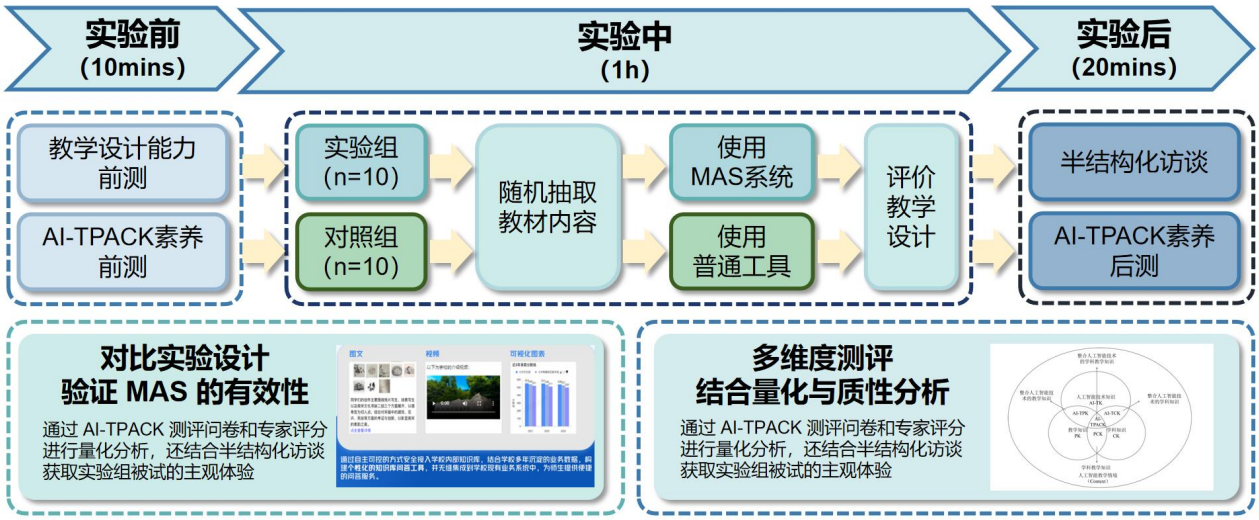


图 3 实验流程图

实验分为四个阶段进行，具体的实验流程图如**图 3** 所示。首先，所有被试在实验前需完成 AI-TPACK 素养测评问卷，以评估其基线水平。接着，在教学设计任务阶段，实验组借助 MAS 进行教学方案设计(交互界面如**图 4** 所示），而对照组则采用传统方法独立完成任务。教学设计任务完成后，所有被试需再次进行 AI-TPACK 素养测评，以评估实验干预对其素养提升的影响。此外，为了深入了解 MAS 在教学设计过程中的具体作用，实验组的被试需参与半结构化访谈，以收集其使用体验及对教学设计能力提升的反馈。最后，为了量化 MAS 对教学设计质量的影响，研究邀请三位教育技术领域的专家对两组被试提交的教学方案进行评分，评分标准涵盖知识点分析准确性、教学法选择合理性、AI 工具整合效果及创新性等维度。

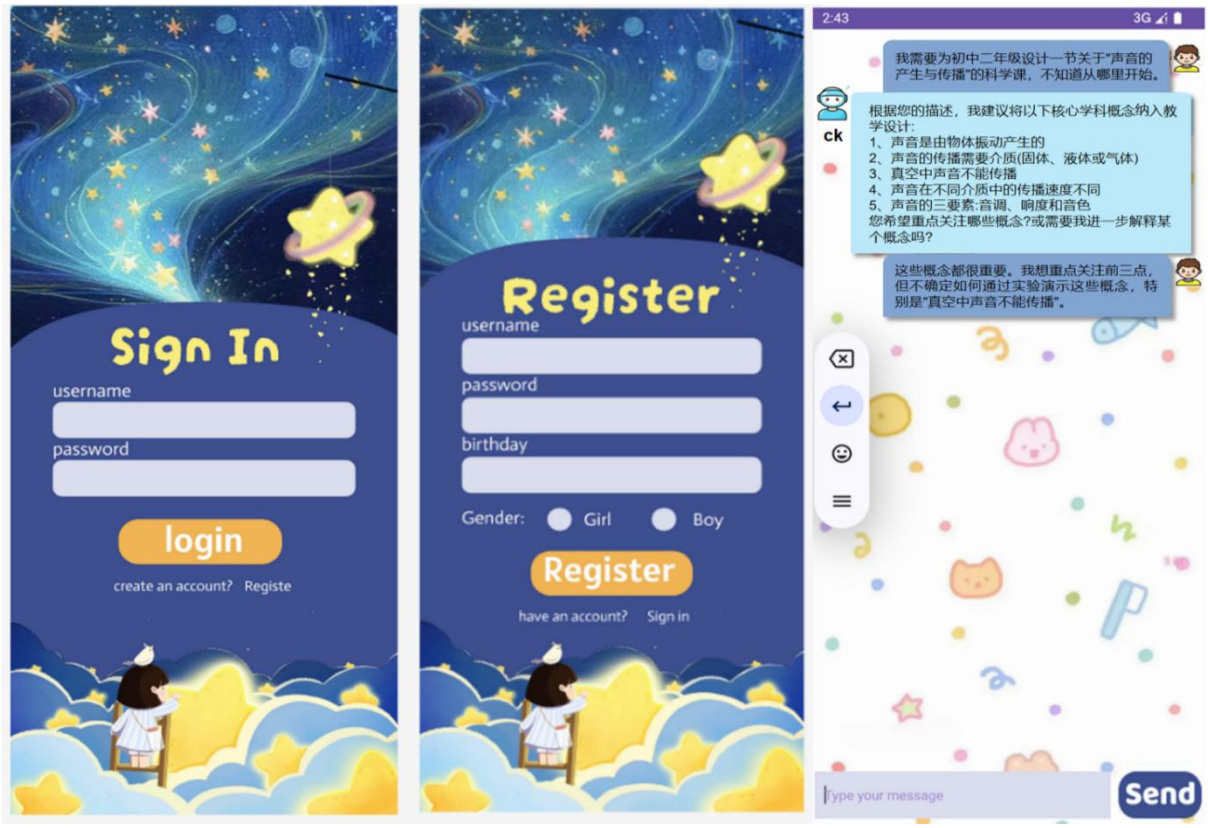


图 4 MAS 系统交互界面

3.3 评估方法

本研究采用定量与定性结合的方法，以综合评估 MAS 对职前教师教学设计质量和 AI-TPACK 素养的影响。首先，在教学设计质量的评估方面，由三位专家基于标准化评分量表对两组被试的教学方案进行评分，并计算实验组与对照组的平均评分差异。其次，在 AI-TPACK 素养的评估方面，研究采用 AI-TPACK 测评问卷，从技术知识（TK）、教学法知识（PK）、学科知识（CK）及其交叉领域（TCK、TPK、TPACK）六个维度测量职前教师的 AI-TPACK 素养。对实验组被试的问卷结果进行前测和后测比较，以分析 AI-TPACK 素养的发展情况。此外，实验组被试的访谈数据将采用主题分析（Thematic Analysis）方法，归纳其对 MAS 的使用体验及其对教学设计能力提升的看法。

4.结果与分析

本研究基于实验数据，对 MAS 在提升职前教师 AI-TPACK 素养和教学设计质量方面的有效性进行了分析。本部分主要包括教学设计质量评分的对比分析、AI-TPACK 素养的提升情况，以及访谈数据的主题分析，以探讨 MAS 在教师教育中的作用。

4.1 教学设计质量评分分析

为了评估 MAS 对职前教师教学设计质量的影响，研究邀请三位教育技术专家对实验组和对照组提交的教学方案进行了评分，评分维度包括知识点分析（CK）、教学方法选择（PK）、AI 工具整合（TK）、创新性（IN）及整体质量（OQ）。评分采用 5 分制（1=极差，5=优秀），两组的评分均取三位专家的平均值。**表 1** 显示了实验组和对照组在各评分维度上的均值（M）和标准差（SD）。

表 1 教学设计质量评分

组别	知识点分析（CK）	教学法选择 (PK)	AI 工具整合（TK）	创新性（IN）	整体质量（OQ）
实验组 (n=10)	4.32 ± 0.48	4.21 ± 0.52	4.45 ± 0.50	4.18 ± 0.47	4.29 ± 0.42
对照组 (n=10)	3.65 ± 0.55	3.72 ± 0.60	2.94 ± 0.58	3.21 ± 0.54	3.38 ± 0.50
t 值	3.24**	2.89**	5.21***	3.56**	4.12***
p 值	0.004	0.008	<0.001	0.003	<0.001

（注：**$p < 0.01$，***$p < 0.001$）

图 5 展示了实验组和对照组在教学设计质量评分上的对比，揭示了 MAS 在促进职前教师教学设计能力提升方面的显著作用。独立样本 t 检验结果显示，实验组在所有评分维度上均显著高于对照组（$p < 0.01$），尤其是在 AI 工具整合（$t = 5.21, p < 0.001$）和整体质量（$t = 4.12, p < 0.001$）方面差异最为显著。这表明，MAS 通过 CK-Agent、PK-Agent 和 TK-Agent 的协同支持，帮助职前教师更有效地整合 AI 工具，优化教学策略，并提升整体教学设计的创新性和质量。

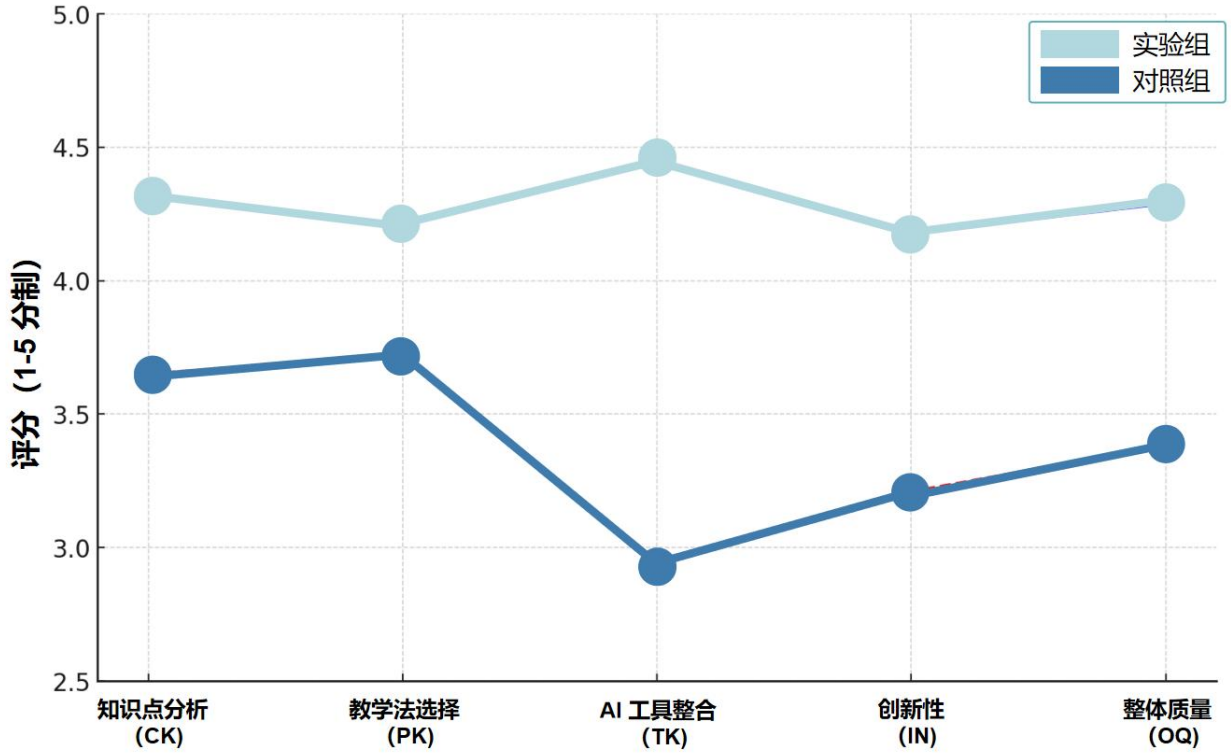


图 5 实验组与对照组的教学设计评分对比

这一结果验证了 MAS 在教师教育中的可行性，说明智能系统可以在教学设计过程中提供针对性的智能反馈，弥补职前教师在 AI 工具应用和教学创新方面的短板。特别是 TK-Agent 在 AI 工具整合方面的支持，使得实验组教师能够更加熟练地运用 AI 技术进行教学设计，符合 AI-TPACK 素养提升的核心目标。

4.2 AI-TPACK 素养提升分析

本研究采用 AI-TPACK 素养测评问卷，对实验组的职前教师在实验前（前测）和实验后（后测）的 AI-TPACK 素养进行了测量，以探讨 MAS 对教师 AI-TPACK 发展的影响。测评涵盖六个维度：技术知识（TK）、学科知识（CK）、教学法知识（PK）、技术与学科整合（TCK）、技术与教学法整合（TPK）以及整体 AI-TPACK 水平（TPACK）。**表 2** 展示了实验组前测与后测的均值（M）和标准差（SD），并采用配对样本 t 检验进行统计分析。

表 2 实验组 AI-TPACK 素养前后测对比

维度	前测（M±SD）	后测（M±SD）	t 值	p 值
技术知识	3.41 ± 0.62	4.38 ± 0.51	4.92***	<0.001
学科知识	4.02 ± 0.58	4.21 ± 0.56	1.32	0.212
教学法知识 (PK)	3.78 ± 0.59	4.32 ± 0.53	3.75**	0.003
技术-学科整合 (TCK)	3.54 ± 0.60	4.29 ± 0.52	4.45***	<0.001
技术-教学法整合 (TPK)	3.62 ± 0.65	4.42 ± 0.57	4.89***	<0.001
综合 AI-TPACK (TPACK)	3.58 ± 0.57	3.58 ± 0.57	5.12***	<0.001

（注：**p < 0.01，***p < 0.001）

配对样本 t 检验结果表明，实验组在技术知识（$t = 4.92, p < 0.001$）、教学法知识（$t = 3.75, p = 0.003$）、技术-学科整合（$t = 4.45, p < 0.001$）、技术-教学法整合（$t = 4.89, p < 0.001$）及综合 AI-TPACK 素养（$t = 5.12, p < 0.001$）方面均有显著提升，而学科知识（t =

1.32, p = 0.212）未出现显著变化。这说明 MAS 主要促进了职前教师在 AI 工具应用、教学方法选择及跨学科整合能力方面的提升，而对学科知识本身的影响较小。

4.3 访谈分析

为了进一步理解 MAS 在支持职前教师教学设计中的作用，本研究对实验组的 10 名被试进行了半结构化访谈，并采用主题分析法提取关键主题。

分析结果表明，受试者普遍认为 MAS 在教学工具推荐、教学方法指导和知识点结构化方面提供了有力支持。其中，80% 的被试表示 MAS 提供的 AI 工具推荐提高了他们的工具整合能力，70% 的被试认为系统生成的知识点图谱帮助他们更好地梳理课程内容，60% 的被试认为 PK-Agent 推荐的教学策略拓展了他们对不同教学方法的认知。然而，也有 40% 的被试表示，在面对较为复杂的教学情境时，MAS 的推荐策略仍然需要进一步优化，以提供更具针对性的个性化支持。

5.讨论与结论

5.1 讨论

基于 MAS 的教学设计支持工具为职前教师的 AI-TPACK 素养提升提供了深刻的推动力。学科知识智能体（CK-Agent）、教学法智能体（PK-Agent）与教学工具智能体（TK-Agent）在智能协作中形成了强大的支持框架，使教师能够更加有效地将技术与教学策略结合。在技术知识（TK）和技术-教学法整合（TPK）方面，MAS 展现出了显著的优势，帮助教师不仅更好地理解 AI 工具的应用，还提升了他们在教学设计中的创新性与科学性。尽管如此，学科内容知识的支持仍然相对薄弱，显示出未来 MAS 系统在学科领域深度挖掘的空间和需求。

MAS 不仅仅是在提升 AI-TPACK 素养方面发挥作用，它更激发了教师在教学设计中的创新潜力。智能推荐系统为教师提供了更加灵活、前瞻性的教学策略，使他们能够打破传统教学框架，采用更适合当前教育环境的方式。教师在使用 MAS 时反映，系统使他们能更自然地结合探究式学习、项目式学习等多样的教学方法，进而推动教学思维的变革。然而，部分教师也指出，在具体教学情境下，MAS 的智能推荐有时未能准确捕捉到情境的复杂性，缺乏足够的个性化反馈。这个问题揭示了 MAS 系统在情境感知与适应性方面的不足，要求未来的智能体系统在细致的个性化支持上做出改进。

技术的引入往往是复杂且带有挑战的，教师的接受度在 MAS 应用中的表现尤为关键。虽然 MAS 为教师提供了丰富的教学支持，但在实际使用中，部分教师因缺乏充分的技术背景和培训而感到困惑或抗拒。技术恐惧感和自我效能感的缺乏，往往导致教师未能完全拥抱系统的潜力。这一现象再次强调了技术融入教学过程的复杂性，提出了在未来研究中应加强教师技术接受度的培养和技术适应性支持，确保技术能够更好地服务于教师的个体需求和教学情境。

5.2 结论

本研究表明，基于 MAS 的教学设计支持工具在提升职前教师 AI-TPACK 素养和教学设计能力方面具有显著效果。MAS 通过智能体协作，为教师提供了跨学科的支持，推动了教学设计的创新，并帮助教师更有效地将 AI 工具与教学方法结合。然而，尽管 MAS 展现了强大的潜力，个性化支持和情境感知能力的提升仍是其发展的关键。未来的 MAS 系统需要在个性化适应性和情境分析能力上进行进一步的优化，同时，加强教师对新技术的接受度和自我认同，确保技术能够真正服务于教学实践。通过这些改进，MAS 将在教师培训和教育创新中发挥更大的作用，推动教育领域的智能化进程。

致谢

本研究受国家大学生创新创业训练项目计划(202310337023)资助。

参考文献

Chen, X. J., Hu, Z. B., & Wang, C. L. (2024). Empowering Education Development through AIGC: A Systematic Literature Review. Education and Information Technologies. 1-53. https://doi.org/10.1007/s10639-024-12549-7

Dillenbourg, P., et al. (2011). Technology for collaborative learning. The Oxford Handbook of Learning Sciences, 539-558.

Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. Proceedings of the 34th International Conference on Machine Learning (ICML 2017), 70, 1126–1135. PMLR.

Hava, K., Babayiğit, Ö. (2025).Exploring the relationship between teachers' competencies in AI-TPACK and digital proficiency. Educ Inf Technol 30, 3491–3508. https://doi.org/10.1007/s10639-024-12939-x

Jin, H., Lee, S., Shin, H., & Kim, J. (2024, May). Teach AI How to Code: Using Large Language Models as Teachable Agents for Programming Education. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-28).

Johnson, W. L., et al. (2014). The use of intelligent tutoring systems to teach mathematical problem solving. Educational Technology Research and Development, 62(3), 311-334.

Karataş, F., Ataç, B.A. (2024). When TPACK meets artificial intelligence: Analyzing TPACK and AI-TPACK components through structural equation modelling. Educ Inf Technol. https://doi.org/10.1007/s10639-024-13164-2

Koehler, M. J., et al. (2013). TPACK: A framework for teacher knowledge. In Handbook of Research on Educational Communications and Technology (pp. 101-111). Springer.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017-1054.

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

Pugh, K. J., et al. (2016). The role of agents in supporting teacher decision-making. Educational Technology Research and Development, 64(6), 1063-1079.

Zhang, Z., Zhang-Li, D., Yu, J., Gong, L., Zhou, J., Liu, Z., ... & Li, J. (2024). Simulating classroom education with llm-empowered agents. arXiv preprint arXiv:2406.19226.

王玮. (2020). 人工智能时代职前教师 TPACK 培养研究. 河南师范大学硕士论文.

张群, 田格格. (2023). AI-TPACK 理论视角下教师智能教育素养：模型构建及培养策略. 开放学习研究, 6.

促进小学生计算思维的均衡发展：一种跨区域协作的无屏幕机器人编程教育模式

Promoting Balanced Development of Computational Thinking in Elementary School Students: A Screen-Free Robotics Programming Education Model Based on Cross-Regional Collaboration

王寅玮[1*]，金凯莹[2]，应浩[3]
[1]宁海县银河小学
[2]宁海县银河小学
[3]宁海县桥头胡中心小学
[*] nhwyw@126.com

【摘要】 本研究聚焦于在城乡教育融合背景下，无屏编程机器人对小学生计算思维发展的影响。通过选取城区与城乡结合部两所小学一二年级学生开展 15 周对比实验，发现无屏编程机器人教学在提升学生计算思维方面效果显著，尤其在序列化思维和条件逻辑方面。且城乡结合部学生在模式识别能力提升上更为突出，同时跨校协作增强了学生的学习参与度。研究为义务教育低学段计算思维培养提供了城乡协同的创新路径，也为教育公平视角下的技术工具设计提供了实证依据。

【关键词】 无屏编程机器人；计算思维发展；城乡教育协作；跨区域学习共同体；教育公平

Abstract : *This study focuses on the impact of screen-free programming robots on the development of computational thinking among elementary school students, under the context of urban and rural education integration. By conducting a 15-week comparative experiment with first and second-grade students from two primary schools located in urban and peri-urban areas, it was found that teaching with screen-free programming robots is notably effective in enhancing students' computational thinking, particularly in sequential thinking and conditional logic. Students in peri-urban areas showed more significant improvement in pattern recognition abilities. Moreover, cross-school collaboration increased students' engagement in learning. This research provides an innovative approach for fostering computational thinking among younger students in compulsory education, promoting synergy between urban and rural areas, and offers empirical evidence for the design of technological tools aimed at educational equity.*

Keywords: Screen-free programming robots; development of computational thinking; urban-rural education collaboration; cross-regional learning community; educational equity

1.引言

1.1. 背景

当今人工智能与数字化转型的时代，计算思维（Computational Thinking, CT）已经成为 21 世纪核心素养的重要组成部分，也成为了基础教育改革的关键目标。自 2006 年提出“计

算思维是运用计算机科学基本概念解决问题、设计系统和理解人类行为的一种普适性思维”以来（Wing & J. M., 2006），其重要性得到了国际组织的广泛认可。经济合作与发展组织（OECD, 2020）在《面向 2030 的学习框架》中明确将 CT 列为“数字时代学生必备能力”，强调通过算法设计、模式识别等核心能力培养系统性逻辑思维。美国、英国、芬兰等国家已经将编程教育从中学下移至小学甚至幼儿园阶段，我国《义务教育信息科技课程标准（2022 年版）》也将 CT 列为学科核心素养，要求从小学低年级开始渗透“问题分解 - 抽象建模 - 算法实现”的思维训练。

然而，当前主流的编程工具大多依赖屏幕交互，这对低龄儿童的认知发展存在潜在的制约（张春艳和宋美霞，2023)。神经科学研究表明，低龄儿童（6 - 8 岁）前额叶皮层尚未成熟，过度依赖视觉符号操作可能导致认知负荷超载（Johnson, R., & Smith, T.,2024)）。例如，图形化编程中复杂的界面元素，像嵌套积木、多级菜单等，很容易分散儿童的注意力，而虚拟角色的抽象行为也难以与儿童的具身经验相联结。更关键的是，屏幕化工具削弱了“动手思考”的具身认知过程，儿童通过触觉反馈与物理空间操作建构逻辑的能力被显著抑制，这种“去身体化”的学习模式可能阻碍低龄儿童从具体经验向抽象思维的过渡。

与此同时，城乡教育资源存在着结构性差异，这导致计算思维教育呈现出“城区超前、乡村滞后”的失衡态势。据教育部统计，2022 年城市小学编程课程开设率达 78%，而乡村地区仅为 23%（教育部,2025)。这种差距不仅体现在硬件配置上，还反映在教学设计能力上。乡村教师常常因为缺乏培训，将编程课简化为“软件操作指导”，难以触及 CT 核心能力的培养。在教育公平的视角下，如何通过技术创新实现资源再分配，成为破解城乡教育失衡的关键命题。

1.2.研究目的与意义

本研究的目的在于探讨无屏编程机器人在跨区域学习共同体中对小学生计算思维发展的影响。通过选取城区与城乡结合部两所小学一二年级学生开展对比实验，我们希望了解无屏编程机器人教学是否能够提升学生的计算思维能力，以及这种教学方式在城乡不同区域的效果差异。

本研究具有重要的意义。一方面，为义务教育低学段计算思维培养提供了城乡协同的创新路径。通过跨区域学习共同体的构建，实现城乡教育资源的互补，让乡村学生也能享受到优质的编程教育资源，促进教育公平。另一方面，为教育公平视角下的技术工具设计提供了实证依据。研究结果可以为开发更适合低龄儿童的编程工具提供参考，推动儿童计算思维与编程教育的均衡发展。

2.实践对策

2.1.构建跨区域学习共同体

2.1.1.建立城乡合作机制

为了实现城乡教育资源的有效整合，促进小学生计算思维的发展，建立跨区域学习共同体的合作机制至关重要。首先，城区和城乡结合部的学校需要签订合作协议，明确双方的权利和义务。例如，城区学校可以提供专业的编程教师培训资源、先进的教学理念和优质的课程设计；城乡结合部学校则可以提供实践场地和学生参与实践的机会。

在合作过程中，要建立定期的沟通会议制度。双方学校的领导、教师可以每月举行一次线上或线下的会议，共同商讨教学计划、交流学生的学习情况和解决遇到的问题。此外，还

可以成立联合教学小组，由城区和城乡结合部的教师共同组成，共同备课、设计教学活动和评价方案。

2.1.2.设计跨校协作任务

设计适合跨区域学习共同体的跨校协作任务是激发学生学习兴趣和提升计算思维的关键。任务的设计要具有趣味性、挑战性和开放性。例如，可以设计一个“智能城市建设”的编程项目，让城区和城乡结合部的学生分组合作。城区学生可以利用其丰富的信息资源和先进的技术设备，负责设计智能城市的整体架构和功能模块；城乡结合部学生则可以根据当地的实际情况，提供关于城市建设的实践经验和创意想法，如设计适合乡村环境的智能农业设备。

在任务实施过程中，要为学生提供明确的任务目标、时间节点和评价标准。同时，要鼓励学生之间进行充分的沟通和协作，可以通过线上交流平台、视频会议等方式，让学生及时分享自己的进展和遇到的问题。

2.1.3.制定效度保障措施

①教学标准化：实验组的教师接受了玛塔机器人专项培训，共 12 学时，并且需要通过模拟授课考核才能上岗。同时，使用统一的教案，包括教学目标、任务脚本、评价标准等，教师每周还需要提交教学反思日志。

②数据三角验证：将 CTS 分数、机器人操作数据与作品评分进行交叉比对，确保研究结论的一致性。同时，随机抽取 30%的课堂录像，由双人独立编码，Kappa 值达到 0.85，保证了数据的可靠性。

③伦理控制：对照组学生在实验结束后，可以获得玛塔机器人的体验课程。所有数据都进行匿名处理，仅用于学术研究。

2.2.研究对象与分组

我们选择了某县城区小学（银河小学）和城乡结合部小学（桥头胡中心小学）的一、二年级学生，共 252 名学生参与研究。将他们分为实验组和对照组：

表 1 实验分组

实验组（126 名学生）	城区实验组（银河小学，63 名学生）	使用玛塔机器人，并参与跨校协作任务
	城乡结合部实验组（桥头胡中心小学，63 名学生）	同样使用玛塔机器人，但更侧重于硬件的调试和适应不同场景的操作
对照组（126 名学生）	城区对照组（银河小学，63 名学生）	上常规的信息技术课，比如文档编辑、网络搜索等，不涉及编程内容
	城乡结合部对照组（桥头胡中心小学，63 名学生）	课程内容与城区对照组相同，确保研究变量的唯一性

在研究开始前，通过计算思维量表（CTS）的前测（详见附件 1），确认四组学生在序列化思维、条件逻辑等关键维度上没有显著差异，保证了研究的公平性。

2.3.设备选择与课程架构

2.3.1.编程工具演进：从屏幕化到无屏具身交互

编程工具的设计直接影响着 CT 培养的效能。第一代屏幕化工具，如 ScratchJr，通过可视化积木降低了操作门槛，使得编程对于初学者来说更加容易上手。它的可视化界面让学生可以直观地看到编程的过程和结果，激发了学生的学习兴趣。然而，其二维界面难以支持空

间推理能力的发展。学生在操作 ScratchJr 时，更多地是在平面上进行编程操作，缺乏对三维空间的感知和理解。

近年来兴起的无屏编程工具，如 Matatalab、Cubetto，通过物理模块与机器人实体交互，更契合低龄儿童的具身认知需求（Matatalab,2025)。触觉反馈，如指令卡排列，帮助儿童将抽象逻辑“锚定”于物理操作。神经科学证据表明，无屏工具激活的体感皮层活动强度显著高于屏幕工具（fMRI, Smith et al., 2021）。例如，当儿童使用 Matatalab 的指令卡为机器人编程时，他们需要通过触摸、排列指令卡来实现机器人的移动，这种亲身的操作体验让他们能够更深刻地理解编程的逻辑。

为了更直观地比较屏幕化编程工具和无屏编程工具的特点，我们可以看以下表格：

表 2 屏幕化编程工具和无屏编程工具的特点

编程工具类型	优点	缺点
屏幕化编程工具（如 ScratchJr）	可视化界面，降低操作门槛，激发学习兴趣	二维界面，难以支持空间推理能力发展，削弱具身认知过程
无屏编程工具（如 Matatalab、Cubetto）	契合具身认知需求，通过触觉反馈促进逻辑理解	缺乏屏幕的直观展示，可能在信息呈现上不够丰富

基于上述分析，本研究主要使用玛塔编程机器人（Matatalab）作为教学工具。它的优势在于：

①不需要依赖屏幕，通过实体的指令卡来操作机器人，这样可以避免学生过度依赖屏幕；

②机器人有灯光、声音和动作等多种反馈方式，让学生能够更直观地看到自己的操作结果；

③支持蓝牙连接和云端数据同步，方便城市和乡村的学生进行协作编程。

2.3.2 课程实施递进：从基础化到综合实践应用

课程总共持续 15 周，每周安排 2 节课，每节课 40 分钟。课程分为三个阶段：

①基础操作阶段（第 1-4 周）：学生学习如何使用指令卡控制机器人，以及如何调试机器人的传感器，比如让机器人避开障碍物；

②协作建构阶段（第 5-11 周）：城市和乡村的学生组成小组，共同完成一些项目，比如设计城市到乡村的物资运输路线；

③综合应用阶段（第 12-15 周）：学生解决更复杂的实际问题，比如设计智能垃圾分类系统，这需要他们综合运用之前学到的序列、循环和条件逻辑等知识。

2.4.打造式代码新视界

2.4.1.基于具身认知理论的教学活动设计

根据具身认知理论，无屏编程教学应注重通过物理交互将抽象的编程逻辑转化为可触摸的具象经验。在教学活动设计中，可以采用以下方法：

首先，设计实物操作活动。例如，在教授编程中的序列概念时，可以让学生使用无屏编程机器人的指令卡，按照一定的顺序排列，控制机器人前进、后退、转弯等动作。通过这种亲身的操作体验，学生能够更直观地理解序列的含义。

其次，开展情境教学。创设与学生生活实际相关的情境，如“拯救小动物”的情境。学生需要通过编程控制机器人穿越障碍，找到被困的小动物。在这个情境中，学生不仅能够学习编程知识，还能培养解决实际问题的能力。

2.4.2 个性化教学与分层指导

每个学生的学习能力和兴趣爱好都存在差异，因此在无屏编程教学中要实施个性化教学和分层指导。教师可以在课程开始前对学生进行评估，了解他们的编程基础和学习能力。根据评估结果，将学生分为不同的层次，为每个层次的学生制定不同的教学目标和任务。

对于基础较弱的学生，教师可以提供更多的指导和帮助，从简单的编程指令开始教起，逐步引导他们掌握编程的基本技能。对于基础较好的学生，可以提供一些拓展性的任务，如让他们设计更复杂的程序或进行编程创意作品的制作。

同时，要关注学生的兴趣爱好，根据学生的兴趣设计教学内容。例如，如果学生对机器人足球感兴趣，教师可以设计与机器人足球相关的编程任务，让学生在自己感兴趣的领域中深入学习编程知识。

2.4.3.提升教师专业素养

①开展专项培训

为了确保无屏编程教学的质量，需要对教师进行专项培训。培训内容可以包括无屏编程工具的使用、具身认知理论在教学中的应用、跨区域学习共同体的组织与管理等方面。可以邀请相关领域的专家来校开展讲座和培训课程，让教师了解最新的编程教育理念和技术。同时，组织教师参加实践操作培训，让他们亲身体验无屏编程教学的过程，掌握教学技巧和方法。

②建立教师交流平台

建立教师交流平台可以促进教师之间的经验分享和合作。可以创建线上交流群，让城区和城乡结合部的教师在群里分享教学心得、遇到的问题和解决方案。此外，还可以定期组织教师教学研讨会，让教师们在研讨会上展示自己的教学成果，共同探讨教学中存在的问题和改进措施。

通过教师之间的交流和合作，能够不断提升教师的专业素养，为无屏编程教学的顺利开展提供有力的保障。

③完善教学评价体系

首先，传统的教学评价往往只注重学生的考试成绩，而在无屏编程教学中，需要建立多元化的评价指标。评价指标可以包括学生的计算思维能力、实践操作能力、团队协作能力、创新能力等方面。例如，在评价学生的计算思维能力时，可以通过观察学生在编程过程中对问题的分析、分解和解决能力来进行评估。在评价学生的实践操作能力时，可以看学生对无屏编程工具的熟练使用程度和完成编程任务的质量。在评价学生的团队协作能力时，可以观察学生在跨校协作任务中的沟通、合作和分工情况。

其次，要将过程性评价与终结性评价相结合，以更全面地反映学生的学习情况。过程性评价可以在教学过程中随时进行，通过课堂观察、学生作业、小组讨论等方式，及时了解学生的学习进展和存在的问题，并给予及时的反馈和指导。终结性评价则可以在课程结束后进行，通过考试、项目作品展示等方式，对学生的学习成果进行综合评价。将过程性评价和终结性评价的结果相结合，能够更客观、准确地评价学生的学习情况，为学生的学习和教师的教学提供有价值的参考。

表 3 教学评价

评价类型	评价方式	评价内容
过程性评价	课堂观察、学生作业、小组讨论	学习态度、知识掌握情况、团队协作表现

终结性评价	考试、项目作品展示	计算思维能力、实践操作能力、创新能力

3.实践结果分析

3.1.数据收集

为了全面评估玛塔机器人的教学效果，我们从以下几个方面收集数据：

①计算思维发展评估：使用修订版的计算思维量表（CTS），包含 15 个题目，信度系数为 0.81，分别在研究开始前和结束后进行测试，间隔 15 周。同时，玛塔机器人会自动记录学生的编程路径数据，比如指令卡的使用频率、调试次数等，这些数据能帮助我们分析学生的思维过程。

②课堂行为观察：记录城市和乡村学生结对讨论的次数，比如他们在一起讨论算法优化建议或者错误排查的情况，定义为每节课内持续时间≥30 秒的实质性互动。同时，通过课堂录像分析实验组学生操作机器人的时间占比，比如指令卡排列占 65%，传感器调试占 35%。

③项目作品分析：从逻辑严谨性和创新性两个维度对学生的项目作品进行评分。逻辑严谨性主要看算法是否能覆盖任务的所有场景，满分 5 分；创新性则看解决方案是否超出了教师预设的框架，满分 3 分。由三位老师独立评分，评分者间信度为 0.79，最后取平均值作为最终成绩。

3.2.数据分析

3.2.1.分析方法

基于计算思维量表对学生的计算思维能力进行评估，通过课堂观察了解学生的学习表现和参与度，以及对学生的作品进行分析以了解他们的思维过程和创新能力。具体的分析方法有如下两种：

①量化分析：采用双重差分法（DID）对比实验组和对照组的 CTS 进步差异，同时控制区域变量的影响；通过协方差分析（ANCOVA）检验城乡实验组学生的能力差距是否显著缩小；利用 Pearson 相关分析玛塔机器人操作时长与 CTS 得分之间的相关性，比如指令卡使用频率与序列化思维的关系。

②质性分析：基于玛塔机器人的编程日志，比如错误指令序列，绘制典型错误类型图谱；对跨校互动文本进行主题编码，比如“算法建议”“硬件经验分享”，提炼出高频协作策略。

3.2.2.分析结果

通过各种分析法结合分析发现，实验组（使用无屏编程）的 CTS 总分提升幅度明显高于对照组（没有编程干预）；城乡实验组学生的能力差距显著缩小，而对照组的差距则保持不变。具体来看：

①整体数据：实验组后测总分平均为 35.2 分，标准差为 4.1，相较于前测的 24.5 分（标准差 3.8），提升了 43.7%。而对照组仅从前测的 24.3 分（标准差 3.9）增长到后测的 25.8 分（标准差 4.0），增长了 6.3%，差异非常显著（$\beta=0.39$, $p<0.001$, 95% CI [0.28, 0.50]）。

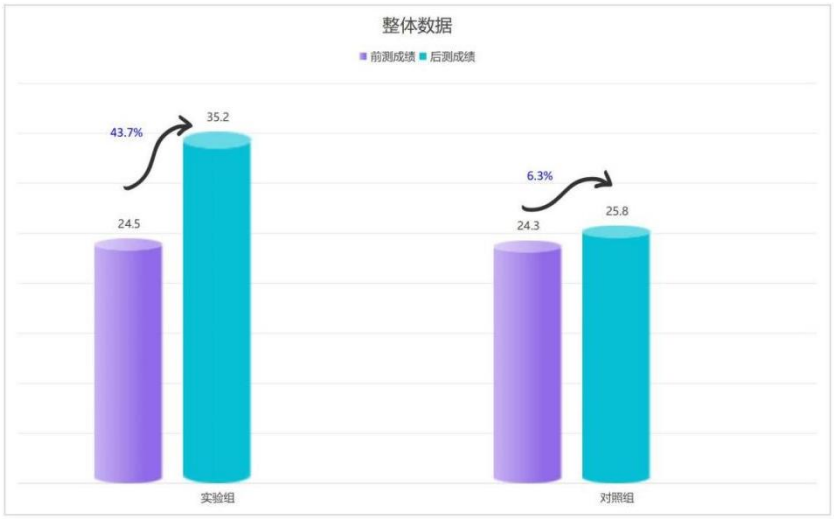


图 1 整体数据

③对比数据：整体表明城乡实验组学生都有进步，但在算法设计维度仍存在差异，城区学生的进步幅度（24.1%）大于城乡结合部学生（18.3%）。因为，城乡结合部的学生更关注硬件的适配，比如传感器的校准，这部分操作占到了 38%，而城区的学生则更侧重于算法的优化，比如路径的简化，这部分操作占到了 45%。而在模式识别维度，城乡结合部实验组学生的进步幅度（28.9%）却超过了城区学生（21.4%），组间差异缩小。

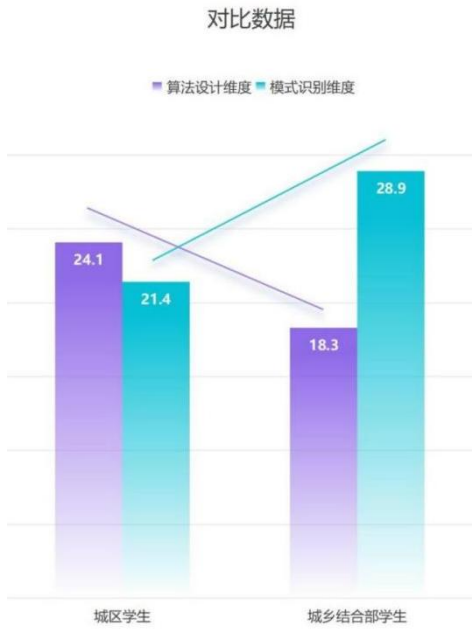


图 2 对比数据

4.实践启示

首先，本研究通过实证数据验证了无屏编程机器人在低龄儿童计算思维发展中的独特价值，无屏编程的触觉操作，比如指令卡排列、机器人路径调试，通过激活多感官通道，将抽象逻辑转化为可感知的具象经验。这表明触觉反馈能够作为低龄儿童抽象思维的“认知锚点”，帮助他们跨越从具体操作到符号化思维的认知鸿沟。

其次，跨区域协作对教育资源认知流动有着催化作用，并为教育公平研究提供了新的理论视角。城乡结合部学生在模式识别维度的追赶效应（28.9%）表明，跨区域协作能够通过认知资源共享缩小城乡能力差距。在城乡分工中，比如城市学生负责算法设计，乡村学生负责硬件调试，知识从资源丰富的方向薄弱的方向流动，同时乡村学生的实践经验也能反过来补充城市学生的理论框架。这揭示了技术赋能下“认知资源再分配”的新路径。

基于研究发现，本研究提出以下实践建议：

①课程普及：可在小学低年级中普及无屏编程机器人课程，优先适配教育资源薄弱地区。课程设计应遵循“实物操作先行、屏幕编程跟进”的渐进式路径，以降低低龄儿童的认知负荷。

②城乡结对学校的常态化协作：构建“双师课堂-任务分级-教研联盟”三位一体的协作机制。例如，城市学校提供结构化算法课程，乡村学校主导在地化问题解决实践，辅以月度联合教研活动共享教学案例；开发难度分级的跨校项目，搭建梯度化任务库，从基础的路径规划到复杂的环境模拟，适配城乡学生的能力差异；建立编程教育教研联盟，定期开展联合培训与案例分享，缩小师资能力差距。

③技术工具优化：开发支持多模态反馈与自适应难度调节的无屏编程工具，尤其需关注乡村学生的操作简易性需求。比如整合 AR 技术可视化抽象概念，循环次数投影，或者添加语音提示辅助低龄儿童理解复杂指令；根据学生操作表现动态调整任务难度，自动简化乡村学生的算法步骤；开发实时共享界面，支持城乡学生远程协同调试等。

5.总结

尽管本研究取得了重要发现，仍需在以下方向深化探索：

①实验周期对长效性结论的限制：15 周的干预周期只能捕捉计算思维的短期提升，无法验证其跨学科迁移效应，比如对数学问题解决能力的影响。未来研究需要延长追踪周期，并纳入 STEM 学科成绩等长效指标。

②家庭数字素养的潜在干扰：城乡家庭在技术接入与家长支持上的差异可能影响实验结果。后续研究需要采用多层线性模型（HLM）分离家庭与学校变量的作用。

未来也将围绕上述方向，进一步扩展研究范围，推动该领域的进一步发展。

参考文献

张春艳&宋美霞. (2023). 拒绝手机依赖, 让孩子脱“瘾”而出. 中学生博览, 33, 12－15.

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33－35.

OECD. (2020). Learning Compass 2030. OECD Publishing.

Smith, J., Brown, A., & Lee, C. (2021). fMRI study on somatosensory cortex activity during screenless programming. Journal of Educational Neuroscience, 12(4), 45－60.

Johnson, R., & Smith, T. (2024). The impact of screen interaction on prefrontal cortex development in early childhood. Developmental Neuroscience, 46(2), 78－90. ScratchJr. (2025). ScratchJr 官方指南. MIT Media Lab.

小学机器人教育中具身学习模式构建与案例开发

Construction of an Embodied Learning Model and Case Development in Elementary School Robotics Education

贾若文[1]，陈禹辰[2*]，董倩[1]，张新立[1]
1 温州大学教育学院
2 悉尼大学教育学院
* oakleychen@foxmail.com

【摘要】 机器人教育在 K-12 教育中受到广泛关注。现阶段中小学机器人多拘泥于传统的讲授式教学，侧重基础知识与技能的传授，缺乏合适的教学模式。有研究指出，在机器人课程中，学习者在学习抽象内容时面临挑战，如构建结构良好的机器人等。具身学习理论强调学习者的身体与环境积极交互，重视多感官参与，将其与机器人教学结合能帮助学生体悟机器人教学的过程，促进学习参与感与体验感。因此，本研究提出基于具身学习的计算机编程方法(EL-CP)，在**乐高编程机器人**课程中进行实验，旨在助力未来研究开发更有效的机器人教学方法和计算机编程活动设计。

【关键词】 具身学习；机器人教育；计算机编程；学习参与

Abstract: *Robotics education in K-12 is mostly traditional and lacks effective teaching models. Learners face challenges in understanding abstract concepts like building structured robots. Embodied learning, which emphasizes active interaction and multisensory participation, can enhance learning experience in robotics. This study proposes an Embodied Learning-based Computer Programming (EL-CP) approach, tested in the LEGO program robot, aiming to contribute to more effective robotics teaching and programming activity design.*

Keywords: Intelligent Pedagogical Agent, Self-efficacy, Creativity

1.引言

在人工智能 （AI）时代，机器人教育逐渐受到 K-12 教育的广泛关注，并已被纳入全球 K-12 学校课程 (Zhang et al., 2023)。以往的研究指出，机器人教育中常见的教学方法强调学生在学习过程中通过自主学习或合作探究来建构机器人知识和技能 (Cheng et al., 2013; Zhong et al., 2019)。然而，由于工程和技术领域的知识和技能有限，青少年学生可能难以准确地构建机器人。此外，他们无法快速、直接地将编程与机器人的动作联系起来，也无法编写计算机程序来驱动机器人。这些困难可能会影响学生在机器人课程中的学习效果(Sullivan & Bers, 2016)。

研究人员指出，教学方法对于机器人课程活动的设计和实施至关重要 (Zhang et al., 2023)。因此，有必要寻找一种合适的教学方法，帮助青少年学生克服这些困难。具身学习是具身认知理论的最新发展，强调学习者的身体与环境在认知过程中积极互动的重要性 (Shapiro, 2019; Zhong et al., 2021)。以往的研究已经验证了其在提高学习成绩、学习态度或课堂参与度方面的有效性 (Johnson-Glenberg et al.,2014)。因此，具身学习在支持青少年学生有效学习机

器人课程方面显示出巨大潜力。然而，将具身学习与机器人教学相结合，以促进学生机器人学习成绩的有效性却鲜为人知。因此，本研究提出了基于具身学习的计算机编程 (Embodied Learning-based Computer Programming, 即 EL-CP)方法，将其应用于乐高编程机器人课程，并以小学生为例，设计了相关教学案例，旨在为未来我国中小学机器人教学的开展提供有益参考。

2.机器人教育

机器人教育在 K-12 阶段日益普及 (Zhang et al., 2023; Zhong et al.,2019)。先前的研究发现，机器人教育对促进学生的学习和发展，提高学生自主学习的积极性起着至关重要的作用 (Bers et al., 2019)。

随着机器人技术的快速发展，越来越多教育机器人被应用于教学中，以促进特定领域知识和技能的发展 (Bers et al., 2019)。根据用户界面，教育机器人可分为有形机器人，虚拟机器人和混合机器人 (Sullivan & Bers, 2016; Jara et al., 2011)。在自动化和机器人课程中使用混合机器人能有效提高学生的机器人理论和知识水平以及实践能力。此外，研究表明，乐高机器人作为专为儿童设计的典型混合机器人已被广泛使用。在学习活动中融入基于乐高的作业可以激发学生的学习兴趣。为学生参与有效的机器人课程打下坚实的基础 (Lawhead et al., 2002; Chu et al., 2022)。因此，本研究采用乐高编程机器人套件作为学习工具，希望为学生提供更具吸引力和更有效的机器人教育体验。

3.具身学习

传统认知理论认为认知是一个计算过程 (Calvo & Gomila,2008)。随着学习科学的发展，第二代认知理论开始受到关注。具身认知作为该理论的重要组成部分，强调了大脑、身体和环境在认知过程中的紧密联系 (Shapiro, 2019)。

具身学习是一种源自具身认知理论的教育实践，强调学习者身体在学习过程中的积极参与的重要性 (Hung et al., 2018)。通过这种学习方式，学习者可以通过具体的感觉运动体验来构建抽象知识。因此，具身学习可以为教育人员提供一种富有创造性和灵活性的教学方法，在学习者的身体、认知和学习环境之间建立有意义的联系，促进他们构建特定的知识和技能 (Hung et al., 2018;)。

近年来，学者们根据教育情境提出了一些具身学习的分类法。例如，Melcer 和 Isbister (2016)开发了一个具身学习框架，包括身体性、对应性、协调性、转换性、映射性、游戏方式和环境七个类别。Skulmowski 和 Rey (2018)的具身学习分类法包括两个维度：身体参与(即低水平 vs. 高水平)和任务整合(身体活动是否以有意义的方式与学习任务相关，即偶然的 vs. 整合的)，此分类法为教育者提供了具身学习的全面概述，并被采纳为本研究的理论基础。

一些研究探索了使用机器人进行具身学习。例如，Zhong 等人 (2022)研究了具有不同程度具身性的可编程机器人工具对 67 名五年级学生学习布尔运算的影响。结果显示，高度具身化组在编程和期末考试中的表现明显优于中等程度具身化组。更重要的是，它可以提高学生的学习成绩、学习参与度和学习态度，并减轻他们的认知负荷。然而，很少有研究探索在机器人教学中融入具身学习以提高学生的机器人学习成绩。因此，本研究提出了基于具身学习的机器人编程教学方法，并将其应用于小学的乐高编程机器人课程。

4.小学机器人课程具身学习方法开发

具身学习理论强调学习者的知识学习、身体体验与环境耦合统一，注重学习者在自由开放的学习环境中展开多元互动，调动多感官通道的参与，有助于学习者的深度知识建构。由此可见，具身学习理论有望为改善我国中小学机器人教学中所存在的问题提供可行路径。因

此，本研究融合了具身学习的内涵、特征与设计理念，依据中小学生的身心特征，以项目式学习为依托，构建出具身视域下中小学机器人教学模式，包括四个主要阶段：介绍、知识和技能学习、小组工作、展示和评估（如图 1 所示），具体阐释如下：

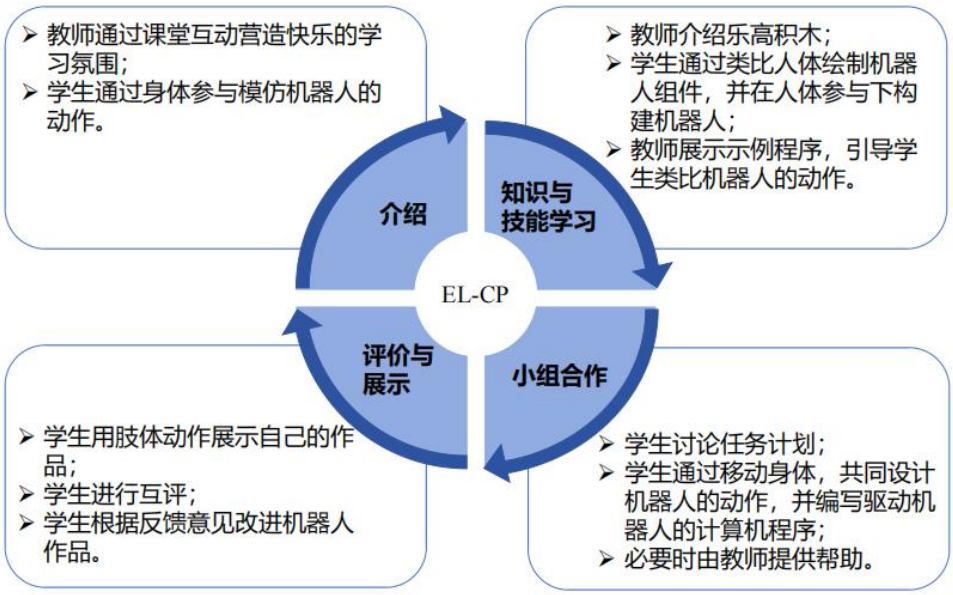


图 1 EL-CP 方法的学习过程

4.1. 导入阶段

介绍阶段是具身视域下中小学机器人教学的初始阶段。在此阶段，教师依据学习内容与学习者身心特征，借助多模态的教学媒体与资源，创设真实且趣味的教学情境；提出驱动问题，并引导学习者将情境与已有知识经验相联系；组织课堂互动，建立起愉快轻松的学习氛围并邀请学生通过身体参与模仿机器人的动作，旨在激发他们的学习兴趣。

4.2. 知识与技能学习

知识与技能学习是具身视域下中小学机器人教学的基础阶段。在此阶段，老师让学生用身体参与的方式，画出人体部件（如手臂、腿）的结构。然后，老师介绍与机器人部件相对应的乐高积木，并让学生类比人体部件的结构，画出机器人部件（如执行器、传感器、控制器）的结构。此外，老师引导学生找出它们之间的相似之处。旨在帮助学生通过有意义的身体动作和体验来理解和掌握机器人知识。接下来，学生需要利用积木、传感器、USB 线等搭建和连接机器人的各个部件。在搭建过程中，老师鼓励学生移动其对应的身体部件，帮助学生建立人体与机器人身体之间更深、更紧密的联系。搭建结束后，老师演示了驱动机器人的计算机程序示例，并指导学生根据所展示的程序类比机器人的动作，旨在帮助学生通过体会建立计算机编程与机器人动作之间的有意义的联系。

4.3. 方案设计与制作优化

小组合作在具身视域下中小学机器人教学中起着承前启后的作用。在此阶段，教师陈述任务要求后，学生与小组成员讨论任务计划。在此过程中，学生通过移动身体（如手臂、腿）来设计机器人的动作和驱动机器人的计算机程序。旨在将学生的身体经验与计算机编程建立有意义的联系，并帮助学生编写程序以准确地驱动机器人。根据计划，学生共同编写计算机程序来驱动机器人的运动，并优化工作直至完成。教师可以在必要时为学生提供帮助。

4.4. 展示与评价

展示与评价阶段的主要任务是通过展示与评价来检验学生学习成果。在此阶段，老师需要组织学生展示自己的机器人作品，并邀请其他同学进行点评和评价，学生可以在展示过程中与机器人一起跳舞和移动身体。展示结束后，学生需要根据反馈改进他们的机器人作品。

5.教学案例开发

本研究基于具身学习视域下中小学机器人教学模式，并立足于信息科技学科，针对中国东部某公立小学六年级学生开发了“跳舞机器人”的教学案例，根据小学生的知识接受程度及乐高 EV3 编程机器人的特点，将机器人与人体的特点联系起来，并根据机器人的各个部位设计不同的学习任务。具体活动流程阐述如下。

5.1. 教学内容与目标分析

“跳舞机器人”是针对浙教版小学信息技术六年级下册第二单元“与机器人零距离”设计的拓展性内容。本单元是小学信息科技人工智能模块的相关教学内容，是小学阶段机器人教学的尾声，也是五年级图形化编程学习内容的延续与深入。该课引入了乐高编程机器人，具体教学内容包括认识并了解乐高编程机器人的使用方法、组件、学会乐高编程机器人的搭建、通过基础的图像化编程使机器人运转起来。

该课的学习目标是：在介绍阶段，通过模仿机器人运动，了解机器人的构造及人与机器人的差异，激发学习兴趣与好奇心；在知识与技能学习阶段，了解乐高编程机器人零件的类型、作用与工作原理，并掌握控制机器人运行的图像化编程方法；在小组合作阶段，使用乐高编程机器人合作探索并完成“跳舞机器人”项目，以应用机器人知识与技能，培养创造力、问题解决与合作能力；在展示与评价阶段，通过展示、分享与评价，巩固与升华对机器人知识的理解，反思自身不足，提升学生的分享与表达能力。

5.2. 教学活动设计

5.2.1. 导入：和机器人一起跳舞

首先，教师播放了一段跳舞机器人的视频，并请学生模仿跳舞机器人的动作。接下来，教师重点讲解了“挥动手臂”这一学习任务。在播放机器人挥动手臂的视频后，教师邀请学生挥动手臂，感知和体验机器人手臂的动作。最后，所有学生都被要求思考两个问题：“组成机器人的部件有哪些？“和“人类是如何挥动手臂的？”图 2 显示学生观看视频并模仿跳舞机器人的动作。

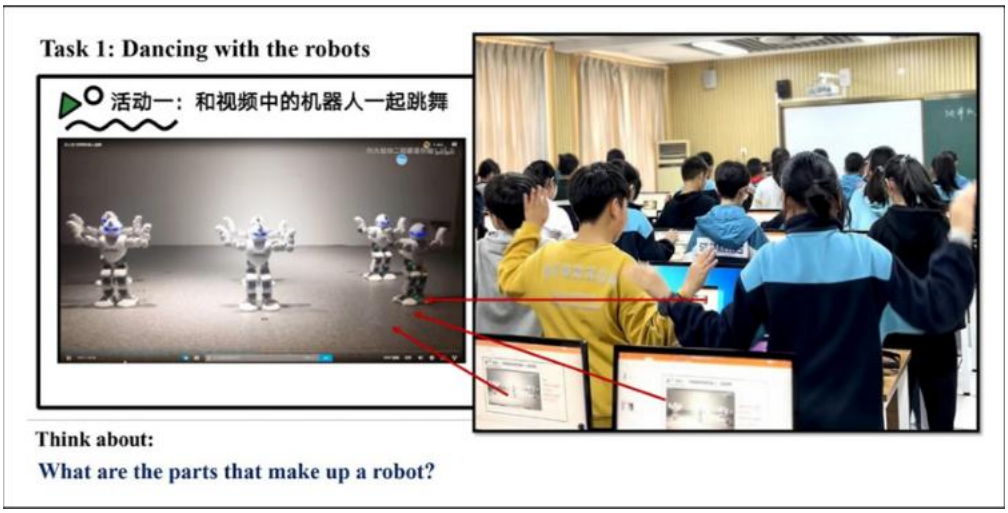


图 2 学生观看视频并模仿跳舞机器人的动作

5.2.2. 知识与技能学习：我与机器人的身体

首先，教师让学生挥舞并观察自己的手臂，在任务单上画出人类手臂的结构。然后，教师介绍了乐高的主要积木及其功能，以及一个与机器人手臂相对应的案例。然后，要求学生根据人的手臂在任务单上画出机器人手臂的草图（见图 3）。此外，教师还请学生找出人类手臂与机器人手臂之间的相似之处，帮助他们在具体的人类手臂与抽象的机器人手臂之间建立有意义的联系。随后，学生们根据自己的草图，用乐高积木搭建了机器人的手臂。在此过

程中，教师鼓励学生摆动手臂，帮助他们将人类手臂和机器人手臂更紧密地联系起来（见图 4）。搭建好机器人手臂后，教师在屏幕上演示了驱动机器人挥动手臂的示例程序。在此过程中，教师引导学生根据演示程序类比机械臂的摆动动作（见图 5）。

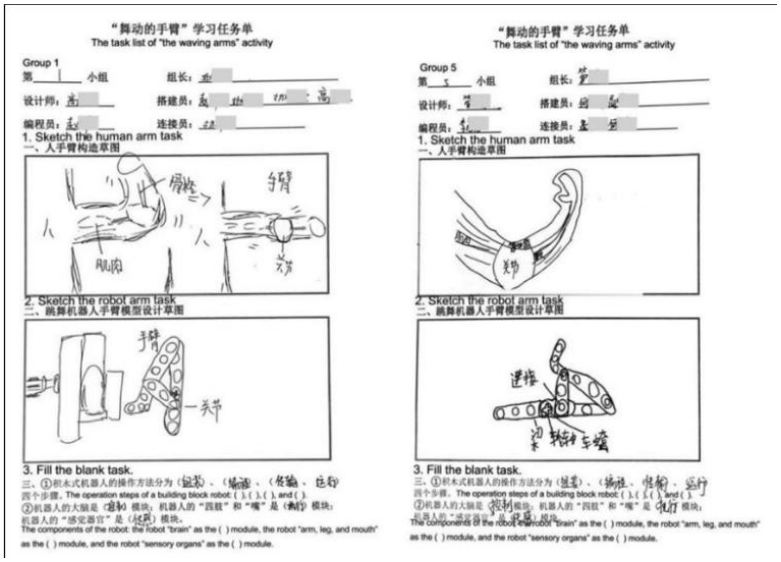


图 3 任务清单及两组示例回答

图 4 学生们一边摆动自己的手臂，一边制作机器人手臂

图 5 教师引导学生通过身体参与来完成机械臂动作

*5.2.3. 方案设计与制作优化：*我为机器人编舞

首先，教师提出任务要求，将学习任务单发给学生，请他们完成跳舞机器人挥动手臂的任务。随后，学生通过模仿机器人舞动手臂的肢体动作，编排和分解舞蹈动作。根据动作顺序，学生们描述了机器人手臂的动作，并绘制了跳舞机器人挥动手臂的编程流程图。图 6 是实验组学生制作的机器人编舞示例和编程流程图。根据编排方案和编程流程图，每个小组需要在电脑上打开名为 EV3 Classroom 的电脑编程软件，编写跳舞机器人挥动手臂的程序。图

7 展示了一组学生在 EV3 教室软件中创建的程序示例。然后，软件根据学生设计的电脑程序，生成并发送机器人动作指令，驱动机器人挥动手臂。学生们不断优化自己的作品，直至完成。

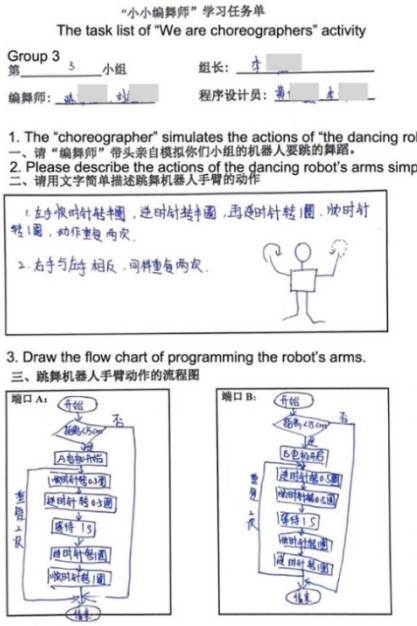


图 6 任务清单和一组学生的示范回答

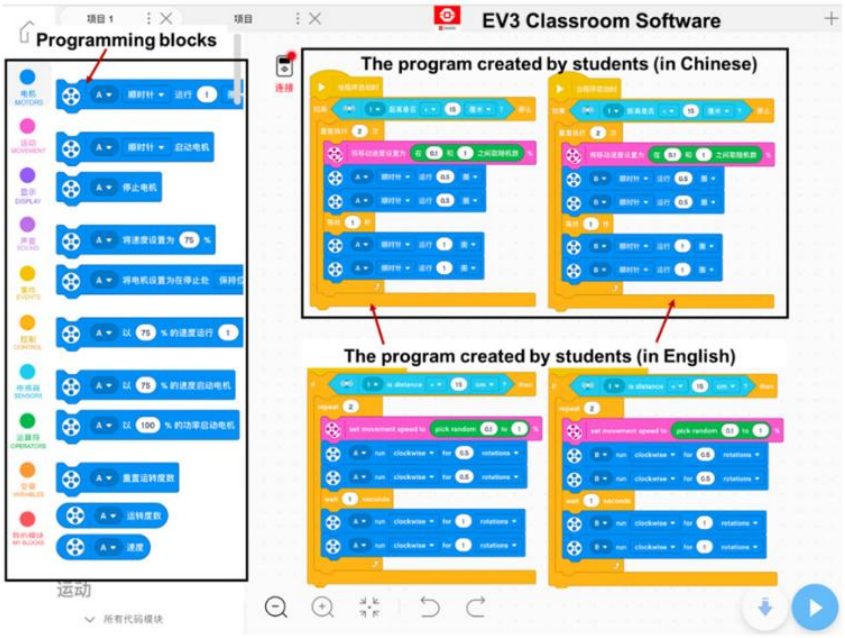


图 7 一组学生在 EV3 教室软件中创建的程序示例

5.2.4. 展示与评价：机器人舞蹈秀

学生们上台展示了他们的机器人挥动手臂作品。教师鼓励学生与机器人一起跳舞，展示他们的舞蹈作品（见图 8）。此外，在展示过程中，每个小组都要对其他小组的作品进行评价并提供反馈。学生可以根据反馈意见修改和优化自己的作品。

图 8 一组学生展示机器人并与之共舞的实例

6.结语

随着创客教育、STEAM 教育的不断推进，技术资源更加开源与低门槛，中小学机器人教育将刮起热潮。本研究构建的具身视域下的中小学机器人教学模式，不仅关注学习者对机器人基本知识与技能的掌握，同时强调学习者身体与环境交互的过程中动手操作、合作探究、创新创造，以体悟机器人教学的过程，旨在为未来我国中小学机器人教学设计与实践提供有益参考。未来研究者可以基于该模式设计针对不同学段的教学案例，展开相关实证研究。

参考文献

Bers,M.U.,Gonza lez-Gonz alez,C.,& Armas–Torres,M.B.(2019).Coding as a playground: Promoting positive learning experiences in childhood classrooms.*Computers & Education,138*(3),130–145.https://doi.org/10.1016/j.compedu.2019.04.013

Calvo,P.,& Gomila,T.(2008).*Handbook of cognitive science: An embodied approach.*Elsevier.

Cheng,C.C.,Huang,P.L.,& Huang,K.H.(2013).Cooperative learning in Lego robotics projects: Exploring the impacts of group formation on interaction and achievement.*Journal of Networks,8*(7),1529–1535.https://doi.org/10.4304/jnw.8.7.1529-1535

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.https://doi.org/10.1016/j.caeai.2022.100091

Hung,I.C.,Kinshuk,& Chen,N.S.(2018).Embodied interactive video lectures for improving learning comprehension and retention.*Computers & Education,117*,116–131.https://doi.org/10.1016/j.compedu.2017.10.005

Jara,C.A.,Candelas,F.A.,Puente,S.T.,& Torres,F.(2011).Hands-on experiences of undergraduate students in automatics and robotics using a virtual and remote laboratory.*Computers & Education,57*(4),2451–2461.https://doi.org/10.1016/j.compedu.2011.07.003

Johnson-Glenberg,M.C.,Birchfield,D.A.,Tolentino,L.,& Koziupa,T.(2014).Collaborative embodied learning in mixed reality motion-capture environments: Two science studies.*Journal of Educational Psychology,106*(1),86–104.https://doi.org/10.1037/a0034008

Lawhead,P.B.,Duncan,M.E.,Bland,C.G.,Goldweber,M.,Schep,M.,Barnes,D.J.,& Hollingsworth,R.G. (2002).A road map for teaching introductory programming using LEGO© mindstorms robots. *ACM SIGCSE Bulletin,35*(2),191–201.https://doi.org/10.1145/782941.783002

Melcer,E.F.,& Isbister,K.(2016).Bridging the physical divide: A design framework for embodied lea rning games and simulations.Proceedings of the 2016 CHI Conference extended abstracts on h uman factors in computing Systems.https://doi.org/10.1145/2851581.2892455

Skulmowski,A.,& Rey,G.D.(2018).Embodied learning: Introducing a taxonomy based on bodily en gagement and task integration.Cognitive Research: *Principles and Implications,3*(1),6.https://d oi.org/10.1186/s41235-018-0092-9

Sullivan,A.,& Bers,M.U.(2016).Robotics in the early childhood classroom: Learning outcomes fro m an 8-week robotics curriculum in pre-kindergarten through second grade.*International Journ al of Technology and Design Education,26*(1),3–20.https://doi.org/10.1007/s10798-015-9304-5

Zhang,X.,Chen,Y.,Bao,Y.,& Hu,L.(2023).Robot illustrated: Exploring elementary students' percepti ons of robots via the draw-a-robot test.*Journal of Research on Technology in Education,1–26*.h ttps://doi.org/10.1080/15391523.2023.2232058

Zhong,B.,& Wang,Y.(2019).Effects of roles assignment and learning styles on pair learning in robot ics education.*International Journal of Technology and Design Education,31*(1),41–59.https://d oi.org/10.1007/s10798-019-09536-2

Zhong,B.,Xia,L.,& Su,S.(2022).Effects of programming tools with different degrees of embodiment on learning Boolean operations.*Education and Information Technologies,27*(5),6211–6231.htt ps://doi.org/10.1007/s10639-021-10884-7

哪一种学习支架更有效？学习支架对不同元认知水平学生学习成绩和计算思维的影响研究

Which Scaffolding is More Effective? A Study of the Effect of Scaffolding on the Learning Achievement and Computational Thinking of Students with Different Levels of Metacognition

朱玉蝶[1], 李伟[1*], 林家苗[2]，刘城烨[3],黄佶艺[4], 李雯[1], 荣健康[1],
[1]温州大学教育学院
[2]海宁市盐仓小学
[3]香港中文大学课程与教学系
[4]温州市瓯海区瞿溪第一小学
[*] liweiwzu@wzu.edu.cn

【摘要】 本研究旨在探讨学习支架类型和元认知水平对学生学习成绩和计算思维的影响。温州市某小学四年级83名学生参与了为期六周的实验，被分配至实验1班（n＝41）和实验2班（n＝42）。实验1班的学生采用反思型学习支架进行干预，实验2班的学生则采用支持型学习支架进行干预。实验结果表明，在计算思维倾向方面，学习支架类型与元认知水平具有交互作用。对于低元认知水平学生而言，支持型学习支架比反思型学习支架更有利于提升他们的计算思维倾向。然而，学习支架类型和个体元认知水平之间的交互对学习成绩没有显著影响。

【关键词】 计算思维；学习支架；元认知

Abstract: *The purpose of this study was to investigate the effects of scaffolding types and metacognitive level on students' learning achievement and computational thinking. Eighty-three fourth-grade students in an elementary school in Wenzhou City participated in a six-week experiment and were assigned to Experimental 1 (n = 41) and Experimental 2 (n = 42) classes. Students in Experiment 1 class were intervened with the reflective scaffolding, while students in Experiment 2 class were intervened with the supportive scaffolding. The results of the experiment showed that the type of scaffolding interacted with the level of metacognition in terms of computational thinking and metacognitive tendencies. For students with low metacognitive levels, supportive scaffolding were more beneficial than reflective scaffolding in enhancing their computational thinking tendencies and metacognitive tendencies. However, the interaction between type of scaffolding and individual metacognitive level did not have a significant effect on learning achievement.*

Keywords: Computational Thinking, Scaffolding, Metacognition

1.前言

随着数字化时代的到来，计算思维被认为是21世纪青少年的一项重要思维技能(Bocconi & K., 2022)。计算思维指的是利用计算机科学的基本概念来解决问题、设计系统和理解人类行为的一系列思维活动(Wing, 2006)。研究者认为，计算思维是学生理解计算机入门课程的

先决条件(F Joseph Agbo, 2021)。作为成功解决问题的关键技能之一，计算思维与个人在日常生活中使用技术的能力有着密切关系(Kale et al., 2018)。

计算思维强调通过对任务的抽象和分解，将复杂的任务逐步自动化。然而，在实际执行任务时，学生往往难以选择合适的方式来描述问题，无法抽象问题，无法将一个不知道如何将一个大而复杂的问题分解成小而简单的步骤，这导致学生无法成功解决问题 (Rijke et al., 2018)。研究表明，如果没有适当的指导，学生在发展计算思维技能方面面临着重大挑战(Denner et al., 2012)。因此，研究者认为可以给学习者提供学习支架，帮助学习者将高难度任务分解，从而消除能力障碍的问题(Zhou et al., 2023)。研究者开发了一个学习者建模和自适应学习支架，结果证明，相较于未接受学习支架的学生而言，接受学习支架的学生能构建更准确的模型，对科学和计算思维概念表现出更好的理解(Basu et al., 2017)。此外，研究发现，在编程游戏中使用学习支架不仅可以提高学生的学习兴趣和动力，还可以帮助他们更好地掌握和应用计算思维(Tikva & Tambouris, 2022)。

早期的研究多关注学习支架的影响，但近期有更多的研究比较了不同支架类型的优势(Kim et al., 2021)。研究证明，不同类型的学习支架对学生的影响是不相同的(Atman Uslu et al., 2022)。例如，支持型支架可以通过提示的方法和机制来实现，即通过提供关于解决方案的可能路径的提示来帮助学生达成解决方案(Acovelli & Gamble, 1997)。反思型支架则可以帮助学习者回答“我做了什么？”或者“我下一步该怎么办？”，引导学生对自己的行为进行反思(Chen et al., 2023)。总之，研究显示，学习支架能显著促进学生计算思维的发展，并对学习成果产生正面效应。本研究中重点关注支持型学习支架和反思型学习支架，因为这两者分别与学生的领域知识和元认知思维密切相关(Kim & Lim, 2019)。

元认知作为一种深层次的认知活动，在人们的认知活动中起着监控、指导和调节的作用，它的发展水平直接制约着人的智力和思维的发展水平(秦利民 & Zhang, 2022)。研究表明，由于学生元认知水平不同，学生的解决问题的能力有所差异，从而会影响到学习成就的高低(Huang & Zheng, 2021)。元认知水平高的学生解决问题的能力会更高，而元认知水平低的学生在问题解决活动中可能更需要帮助。因此，针对不同元认知水平的学生应该提供不同的学习支架(Kim & Lim, 2019)。综上所述，本研究旨在探讨学习支架类型对不同元认知水平的学生的影响差异。研究问题如下：

(1) 学习支架类型对不同元认知水平学生的学习成绩是否有影响？

(2) 学习支架类型对不同元认知水平学生的计算思维倾向是否有影响？

2.实验设计

2.1. 研究对象

温州市某小学四年级 83 名学生参与了为期六周的实验，他们的平均年龄为 11-12 岁，被分配至两个班。其中一个班为实验 1 班（n＝41），另一个班为实验 2 班（n＝42）。学生信息均做匿名处理。

在实验班级内部，通过收集实验对象元认知倾向问卷数据进行处理，将实验对象根据元认知水平的高和低各分为两组。两个班由同一位老师授课，学习时间与课程内容相同。

2.2. 学习材料

本实验选取了浙江省义务教育教科书四年级信息技术教材中的部分内容作为主要的学习内容。在实验过程中，依据 Acovelli 和 Gamble (1997)的研究，支持型学习支架以提示和视觉材料的形式呈现，提供课堂知识内容和学习者需要接受的指导。反思型学习支架由探索性

问题、暗示、总结等部分组成，为学习者提供了元认知问题，帮助他们回顾学习内容和反思学习过程(Chen et al., 2023)。两种类型支架的差异如表 1 所示。在为期四周的学习过程中，指导者没有提供额外的支持来控制实验。每次完成支架的设计后，都交由具有 4 年教学经验的实验学校的教师进行检阅和修改。

表 1 支持型和反思型支架的差异比较

支架类型	关键特征	呈现方式	主要内容概括
支持型支架	提示	视觉资料	抽象理论解释
	指导	微视频	操作性知识步骤提示
			软件使用指导
反思型支架	问题	纸质资料	教学过程性反思
	反思	任务单	课堂总结性反思
			自我课堂评价反馈

根据支持型支架与反思型支架的差异，并结合学习内容，支架设计的具体内容如表 2 所示。

表 2 支持型和反思型支架的具体内容比较

课时	反思型支架	支持型支架
1 图文并茂美文章	在插入图片的过程中，你遇到了什么问题？ 在插入艺术字的过程中，你遇到了什么问题？ 本节课你学到了什么知识？ 请为你在本节课的表现做出评价。	提供插入图片的过程和方法的提示。 进行图片大小、位置调整的演示。 分享插入艺术字的方法视频。
2 巧用表格来排版	在插入表格的过程中，你遇到了什么问题？ 在美化表格的过程中，你遇到了什么问题？ 本节课你学到了什么知识？ 请为你在本节课的表现做出评价。	提示插入表格的过程。 分享调整表格大小、位置和颜色的视频。 演示添加文本框的步骤。 展示组合图形的要点与技巧。
3 采集生活点滴	在插入表格的过程中，你遇到了什么问题？ 在美化表格的过程中，你遇到了什么问题？ 本节课你学到了什么知识？ 请为你在本节课的表现做出评价。	展示耳麦的连接方法视频。 分享资料的分类技巧与方法。 提示录音的步骤与过程。
4 精彩的图片剪辑	在插入表格的过程中，你遇到了什么问题？ 在美化表格的过程中，你遇到了什么问题？ 本节课你学到了什么知识？ 请为你在本节课的表现做出评价。	提示保存图片的步骤与要点。 分享裁剪图片的过程、步骤和方法。

2.3. 研究工具

本研究的测量工具包括学习成绩前测试题、学习成绩后测试题、计算思维倾向问卷以及元认知倾向问卷。其中，学习成绩的前测、后测试题由经验丰富的教师结合学生所学知识自主出题，试卷完成后由该领域专家共同检阅通过，难度适合四年级学生的水平。前后测试题都包含 10 道选择题。每套题目中都含有两个反向思维题，每题十分，共计 100 分。前后测试题的组成结构大致相同，且难度相差不大。

计算思维倾向问卷改编自 Hwang and Li (2020)所开发的量表。它由 6 个题目组成，采用 5 点李克特量表评分方案（即：非常同意、同意、一般、不同意、非常不同意）。问卷的 Cronbach's alpha 值为 0.763。

元认知倾向问卷改编自 Lai and Hwang (2014)所开发的量表，包含 5 个条目，其 Cronbach's alpha 值为 0.757。问卷采用 5 点李克特量表进行评分，分值从 1 至 5，依次代表“完全不同意”、“不同意”、“中立”、“同意”和“完全同意”。

2.4. 实验流程

本实验共持续六周，每周一课时，每次课程 45 分钟。在第一周，学生需要完成前测和前问卷。在第二周至第五周，教师对学生们进行授课，其中，实验 1 班的学生采用反思型学习支架进行干预，实验 2 班的学生则采用支持型学习支架进行干预。在第六周，学生完成后测知识测验，并填写后测问卷。

2.5. 数据分析方法

在本次实验研究中，为了区分不同元认知水平的学生，根据元认知倾向前问卷评分，将学生分为高元认知组和低元认知组。实验 1 班得分高于中位数（3.8）为元认知水平高，低于的为元认知水平低；实验 2 班得分高于中位数（4.0）为元认知水平高，低于的为元认知水平低。

在剔除了异常情况的数据之后，收集到了实验 1 班高元认知水平组 16 个样本量，低元认知组 14 个样本量，1 组总计样本量为 30 个；实验 2 班高元认知水平组 15 个样本量，低元认知组 15 个样本量，2 组总计样本量为 30 个。两组共计样本量为 60 个。实验对象情况如表 3 所示。

表 3 实验对象情况调查表

支架班级	元认知分组	个案数
实验 1 班	低元认知	14
	高元认知	16
	总计	30
实验 2	低元认知	15
	高元认知	15
	总计	30
总计	低元认知	29
	高元认知	31
	总计	60

3.实验结果

3.1. 学习成绩

为了了解学习支架类型和元认知水平对学生学习成绩的影响，本研究采用双因素协方差的数据分析方法。将学生的前测成绩作为协变量，消除学习者在学习活动前原有知识水平的差异。自变量为学习支架类型和元认知水平，因变量为学生的学习成绩。

表 4 学生学习成绩的双因素协方差分析结果

变量	平方和	自由度	均方	*F*	*P*	η^2
前测知识总分	1023.897	1	1023.897	2.911	.094	.050
元认知分组	21.311	1	21.311	.061	.806	.001
班级	443.659	1	443.659	1.261	.266	.022
元认知分组 * 班级	314.317	1	314.317	.894	.349	.016
Error	19345.627	55	351.739			

双因素协方差分析结果如表 4 所示，在学习成绩方面，学习支架类型和元认知水平之间不存在显著交互作用($F = .894$，$p > 0.05$)，说明学习支架类型和元认知水平不会共同对学生学习成绩产生显著影响。此外，学习支架类型($F = 1.261$，$p > 0.05$)和元认知水平($F = 0.061$，$p > 0.05$)对学生的学习成绩也没有显著影响。

3.2. 计算思维倾向

为了了解学习支架类型和元认知水平对学生计算思维倾向发展的影响，本研究采用了双因素协方差分析方法。双因素协方差结果分析如表 5 所示，在计算思维倾向方面，支架类型和元认知水平之间存在显著的交互作用($F = 5.973$，$p < 0.05$)，两者之间的交互效应量为 0.098。

表 5 学生计算思维倾向的的双因素协方差分析结果

变量	平方和	自由度	均方	*F*	*P*	η^2
班级	3.179	1	3.179	14.720**	.000	.211
元认知分组	.012	1	.012	.056	.814	.001
班级 * 元认知分组	1.290	1	1.290	5.973*	.018	.098
Error	11.878	55	.216			

*$p<0.05$，**$p < 0.01$

进一步分析元认知水平对学生计算思维倾向的简单主效应，如表 6 所示。结果表明，实验 1 班($F = 3.658$，$p > 0.05$)和实验 2 班($F = 2.783$，$p > 0.05$)学生的元认知水平在计算思维倾向方面均没有显著性差异。

表 6 元认知水平对学生计算思维倾向的简单主效应分析结果

变量		平方和	自由度	均方	*F*	*P*	η^2
实验 1 班	组间	.819	1	.819	3.658	.064	.061
（反思型支架）	组内	6.042	27	.224			
	总计	437.417	30				
实验 2 班	组间	.588	1	.588	2.783	.129	.041
（支持型支架）	组内	5.709	27	.211			
	总计	561.083	30				

表 7 学习支架类型对学生计算思维倾向的简单主效应分析结果

变量		平方和	自由度	均方	*F*	*P*	η^2
低元认知水平	组间	4.247	1	4.247	18.527***	.000	.259

	组内	5.960	26	.229			
	总计	474.083	29				
高元认知水平	组间	.151	1	.151	.744	.314	.018
	组内	5.687	28	.203			
	总计	524.417	31				

***$p < 0.001$

表 7 所示为学习支架类型对学生计算思维倾向的简单主效应分析结果。低元认知水平的学生在学习过程中使用不同的学习支架，计算思维倾向发展有显著性差异($F = 18.527$，$p < 0.001$)。高元认知水平的学生在学习过程中使用不同的学习支架，计算思维倾向发展没有显著性差异($F = 0.744$，$p > 0.05$)。结果表明，相比元认知水平高的学生，学习支架更有利于元认知水平低的学生。

图 1 为学习支架类型与元认知水平对学生计算思维倾向影响的交互作用图。结果表明，采用支持型学习支架的学生比使用反思型学习支架的学生具有更高的计算思维倾向。此外，当使用支持型学习支架时，低元认知水平的学生的计算思维倾向发展明显高于高元认知水平的学生。

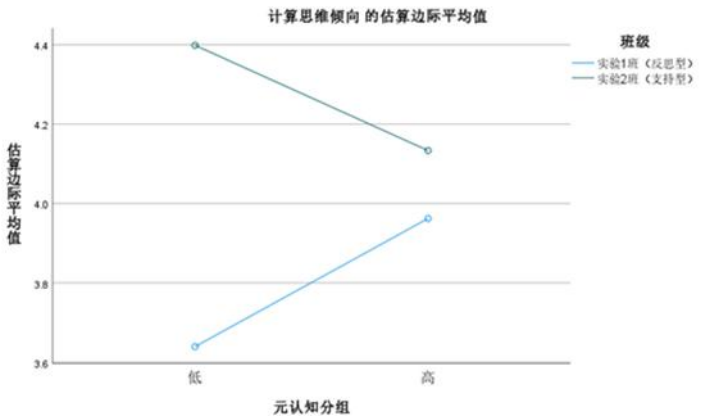


图 1 学习支架和元认知水平在计算思维倾向中的交互作用

4.讨论与结论

本研究考察了学习支架类型对不同元认知水平学生的学习成绩、计算思维倾向和元认知倾向的影响。实验结果表明，在计算思维倾向和元认知倾向方面，学习支架类型与元认知水平具有交互作用。对于低元认知水平学生而言，支持型学习支架比反思型学习支架更有利于提升他们的计算思维倾向和元认知倾向。然而，学习支架类型和个体元认知水平之间的交互对学习成绩没有显著影响。

在学习成绩方面，根据研究结果，学习成绩并未受到支架类型、元认知水平以及这两者相互作用的影响。这种现象可能源于实验持续时间较短和教学课时不足，这限制了支架对学生学习干预的效果。

在计算思维倾向的发展上，分析结果显示，支架类型与元认知水平之间存在显著的相互作用。具体来说，接受支持型学习支架的学生展现出比使用反思型学习支架的学生更强烈的计算思维倾向。以往研究发现，在空间能力自我效能方面，支持性学习支架比反思性学习支架更有效，这与本研究的发现相呼应(Atman Uslu et al., 2022)。因此，可以推断支持型学习支架对于促进学生计算思维倾向的发展起到了更为关键的作用。

此外，本研究也存在一些局限性。首先，本研究的样本总量较小，可能导致实验结果代表性弱，推荐后续研究可以扩大样本量。其次，本实验的实验周期较短，对于习惯传统模式教学的学生来说，支架的干预作用并不明显，使某些方面的实验结果并不显著。建议未来的研究者考虑延长实验时间，以增强支架对学生学习过程的干预效果。

参考文献

Acovelli, M., & Gamble, M. (1997). A Coaching Agent for Learners Using Multimedia Simulations. Educational Technology, 37, 44-48.

Atman Uslu, N., Yildiz Durak, H., & Ay, G. M. (2022). Comparing reflective and supportive scaffolding in 3D computer‐aided design course: Engineering students' metacognitive strategies, spatial ability self‐efficacy, and spatial anxiety. Computer Applications in Engineering Education, 30(5), 1454-1469.

Basu, S., Biswas, G., & Kinnebrew, J. S. (2017). Learner modeling for adaptive scaffolding in a Computational Thinking-based science learning environment. User Modeling and User-Adapted Interaction, 27(1), 5-53.

Bocconi, S., Chioccariello, A., Kampylis, P., Dagienė, V., Wastiau, P., Engelhardt, , & K., E., J., Horvath, M.A., Jasutė, E., Malagoli, C., Masiulionytė-Dagienė, V. and Stupurienė, G. (2022). Reviewing Computational Thinking in Compulsory Education.

Chen, C. H., Liu, T. K., & Huang, K. (2023). Scaffolding vocational high school students computational thinking with cognitive and metacognitive prompts in learning about programmable logic controllers. Journal of Research on Technology in Education, 55(3), 527-544.

Denner, J., Werner, L., & Ortiz, E. (2012). Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? Computers & Education, 58(1), 240-249.

F Joseph Agbo, S. S. O., J Suhonen，M Tukiainen. (2021). iThinkSmart: Immersive virtual reality mini games to facilitate students' computational thinking skills.

Huang, Y., & Zheng, X. (2021). Effects of Scaffolding Types and Individual Metacognition Levels on Learning Achievement in Online Collaborative Argumentation.

Hwang, G.-J., & Li, K.-C. (2020). Trends and strategies for conducting effective STEM reseGwo-Jen Hwang arch and applications: a mobile and ubiquitous learning perspective.

Kale, U., Akcaoglu, M., Cullen, T., Goh, D., Devine, L., Calvert, N., & Grise, K. (2018). Computational What? Relating Computational Thinking to Teaching. TechTrends, 62(6), 574-584.

Kim, J. Y., & Lim, K. Y. (2019). Promoting learning in online, ill-structured problem solving: The effects of scaffolding type and metacognition level. Computers & Education, 138, 116-129.

Kim, N. J., Vicentini, C. R., & Belland, B. R. (2021). Influence of Scaffolding on Information Literacy and Argumentation Skills in Virtual Field Trips and Problem-Based Learning for Scientific Problem Solving. International Journal of Science and Mathematics Education, 20(2), 215-236.

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.

Rijke, W. J., Bollen, L., Eysink, T. H. S., & Tolboom, J. L. J. (2018). Computational Thinking in Primary School: An Examination of Abstraction and Decomposition in Different Age Groups [Article]. Informatics in Education, 17(1), 77-92.
Tikva, C., & Tambouris, E. (2022). The effect of scaffolding programming games and attitudes towards programming on the development of Computational Thinking. Education and Information Technologies, 28(6), 6845-6867.
Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.
Zhou, Y., Chai, C. S., Li, X., Ma, C., Li, B., Yu, D., & Liang, J.-C. (2023). Application of Metacognitive Planning Scaffolding for the Cultivation of Computational Thinking. Journal of Educational Computing Research, 61(6), 1123-1142.
秦利民, & Zhang, L. J. (2022). 近 10 年(2011-2020)国际元认知研究动态:回顾与展望. 外语学刊 (3), 98-106.

基于隐喻的机器人编程方法：促进幼儿的计算思维和积极学习行为

A metaphor-based robot programming approach to facilitating young children's computational thinking and positive learning behaviors

董倩 1，陈禹辰 2，贾若文 1，张新立 1*
1 温州大学教育学院
2 悉尼大学教育学院
zhxl@wzu.edu.cn

【摘要】 在人工智能时代，培养幼儿的计算思维（CT）尤为重要。可编程机器人为幼儿提供了学习编程和培养计算思维的绝佳机会。然而，幼儿在学习抽象的计算思维概念时可能存在困难。因此，本研究提出了一种基于隐喻的机器人编程（MRP）方法，并探讨了其对幼儿 CT 和行为模式的影响。本实验共招募了 118 名 5-6 岁的儿童，实验组采用基于隐喻的机器人编程（MRP）方法，对照组采用传统的机器人编程（CRP）方法。结果显示，采用 MRP 方法可以促进学生的计算思维、学习成绩以及更积极的学习行为。因此，采用基于隐喻的机器人编程（MRP）方法有助于幼儿编程教学和 CT 发展。
【关键词】 教学/学习策略；改进课堂教学；交互式学习环境；隐喻；计算思维

Abstract: *In the era of artificial intelligence, cultivating computational thinking (CT) in young children is particularly important. Programmable robots provide an excellent opportunity for young children to learn programming and develop computational thinking. However, young children may encounter difficulties when learning abstract CT concepts. Therefore, this study proposes a metaphor-based robotics programming (MRP) approach and explores its impact on young children's CT and behavioral patterns. A total of 118 children aged 5-6 were recruited for this experiment; the experimental group adopted the MRP method, while the control group used traditional robotics programming (CRP). The results show that the MRP method can promote students' computational thinking, academic performance, and more positive learning behaviors. Thus, using a metaphor-based robotics programming (MRP) approach can contribute to young children's programming education and CT development.*
Keywords: Teaching/Learning Strategies, Improving Classroom Instruction, Interactive Learning Environments, Metaphor, Computational Thinking

1. 提出问题

在人工智能时代，计算思维（Computational Thinking，CT）作为一项数字能力受到了极大关注（Bers, 2019）。计算思维是一种思维能力，它利用计算机科学的基本概念来解决技术问题和现实生活中的问题（Wing，2006）。最近的研究表明，早期接触 CT 可以预测 STEM 学科的学习、执行功能、21 世纪能力、甚至是未来的职业道路（Sun et al., 2022；Bati,2021）。幼儿期是儿童成长和学习的关键阶段（Piaget, 1953），因此培养幼儿的 CT 至关重要。

编程被认为是促进学习者 CT 的最主要、最有效的手段（Bers，2018；Macrides et al., 2022）。可编程机器人作为一种无屏幕且适合儿童发展的手段，为低年级儿童提供了接触编程和 CT 的绝佳会（Critten et al., 2021）。然而，传统教学方法侧重于抽象概念介绍，未充分结合具体经验，导致幼儿难以掌握如算法、循环等抽象概念（Relkin et al.,2021）。

概念隐喻理论（CMT）指出，隐喻能让个体将相对抽象的经验领域与我们熟悉的更简单或更清晰的经验领域联系起来（Lakoff, 1993）。之前的研究已经验证了在语言、数学和生物等多学科教学中使用隐喻的有效性（Reinhardt, 2020; Yee, 2017），并在编程教学中显示潜力（P´erez-Marín et al., 2020）。

因此，本研究提出一种基于隐喻的机器人编程（MRP）方法来培养幼儿的 CT 能力。此外，本研究进一步探讨了 MRP 方法对幼儿行为模式的影响。我们假设，采用 MRP 方法的幼儿能更好地提高 CT 能力，并表现出更积极的学习行为。本研究的主要研究问题如下：

(1)与传统的机器人编程（CRP）方法相比，MRP 方法是否能更好地促进幼儿的 CT？

(2)MRP 方法比 CRP 方法更能促进幼儿的积极学习行为吗？

2.开发基于隐喻的机器人编程(MRP)方法

2.1. MRP 学习环境

本研究使用 Matatalab 编程套件作为学习环境(图 1)。Matatalab 是专为 4-9 岁儿童设计的无屏幕可编程机器人套件，包括一个可编程机器人(MatataBot)、指挥塔、控制板和编程块。Matatalab 编程套件可帮助儿童学习基本的编程和 CT 概念，探索机器人编程问题解决任务，并发展 CT(Fu et al., 2023; Yang, Ng, & Gao, 2022)。

图 1 Matatalab 编程工具包

2.2. 采用 MRP 方法的课程设置

机器人编程课程包括 10 个 40 分钟的活动，每周一个活动。孩子们被分成两人一组用一个 Matatalab 编程工具包。该课程名为“自然之旅”。该课程的主要学习目标是(1)让孩子们理解并掌握 CT 概念，包括硬件、软件、算法、模块化、调试等；(2)通过学习和练习，让孩子们掌握基本编程模块的符号表示法:前进、后退、左转、右转、数字、循环等；(3)让孩子们将不同的编程模块结合起来，运用基本的 CT 概念解决问题；(4)让孩子们体验机器人编程的乐趣。

2.3. MRP 方法的教学过程

MRP 方法的教学过程主要包括六个阶段:(1)根据目标域设计源域的事件和活动；(2)介绍源域的事件和活动； (3)通过源域解释目标域；(4)介绍与目标域相关的编程模块；(5)发布与目标域相关的机器人编程任务；（6）分享和反思(Bers, 2019)。

第一阶段是根据目标域设计源域的事件和活动。隐喻的基本要素是目标域、源域和映射，从源域到目标域的映射，就是把源域映射到目标域上，理解源域的结构来理解目标域(MacCormac, 1990; Saban, 2006)。在这一阶段，教师确定机器人编程学习所涉及的抽象 CT 概念(目标域)，并根据目标 CT 概念的特点设计适当的、以学习者为中心的源域事件和活动。源域的设计原则是应与儿童的日常生活情境紧密联系，他们已经学习并清楚理解的现有知识。本研究组建了一个专家小组，包括两名教育技术研究人员和两名经验丰富的幼儿 STEM 教师，以讨论和决定源域的事件和辅助活动。表 1 显示了本研究设计的目标域和源域概览。

表 1 本研究设计的目标域和源域概况

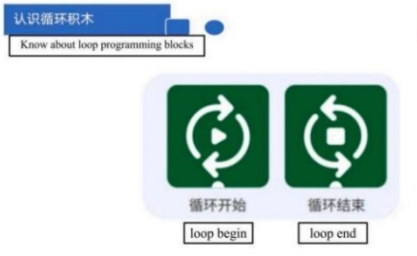
目标域	源域
算法/序列:通过一系列连续的步骤逻辑性地解决问题。	洗衣服的步骤:通过“洗衣服的步骤”这个比喻来解释算法/序列的概念:一个男孩在踢完足球后发现衣服弄脏了，遇到一个大问题。因此，他想洗脏衣服。他会按步骤洗吗?他应该一步一步地洗:脱掉衣服，打开洗衣机门，等待一段时间，拿到干净的衣服……就像这个男孩按照特定的步骤洗衣服一样，算法/序列是按照逻辑顺序一步一步地解决特定问题。
表征:理解符号代表概念、指令等。	交通灯和路标的指令被用作比喻来教授表示概念。当驾驶汽车遇到道路上的交通灯/路标时，交通灯/路标会向我们发出信号或指示。每个信号或指示代表一个特定的命令。我们需要识别交通灯/路标的命令，如停止、前进、左转和右转，然后遵循这些命令。同样，当我们遇到程序中的指令时，我们需要识别特定的命令或动作，并遵循它们。
循环:控制结构的一个相关概念，指的是重复执行同一指令多次，直到达到某个条件，有时等同于重复。	选择旋转木马作为比喻来解释循环的概念:当我们去主题公园时，旋转木马会一次又一次地旋转;例如，它不会停止，直到达到 20 次的条件。同样，循环会一次又一次地重复相同的命令，直到满足条件。
调试:识别错误的部分并尝试解决和修复它。	修理坏汽车的比喻被用来说明调试的概念。当汽车不能正常工作，我们需要找到坏的部分，思考原因，并尝试修理，直到它恢复正常。就像这样，调试是尝试找到错误并对其进行改进的过程，直到正确为止。

第二阶段是介绍源域的事件和活动。在这一阶段，教师通过介绍源域的事件和活动让幼儿回忆和感知与日常生活情境或熟知的知识相关的概念。然后，教师引导幼儿提取源域的特征，为接下来学习抽象的 CT 概念打下坚实的基础。以循环概念为例。首先，教师在屏幕上展示了旋转木马的图片，并播放了一段视频引导学生了解旋转木马的工作原理和特点(图 2)，并进行引导互动。在与孩子们互动后，教师绘声绘色地说道：“当我们在主题公园乘坐旋转木马时，它会反复旋转，直到达到 10 次的条件才会停止。”

图 2 教师向学生介绍旋转木马的事件和活动

第三阶段是通过源域解释目标 CT 概念。在这一阶段，教师引导孩子们注意到源域与目标 CT 概念之间的相似性，从而帮助他们更好地理解和掌握抽象概念(Peelle, 1983)。以循环概念为例。首先，教师向全班介绍了循环概念：“我们刚刚学习了旋转木马的工作原理。与旋转木马的工作原理一样，循环也是一遍又一遍地重复相同的指令，直到满足某个条件为止。”随后，教师引导幼儿注意旋转木马(源域)和循环(目标域)之间的相似之处。例如，一名儿童回答道：“旋转木马重复旋转，而循环也是重复执行一个动作。”

第四阶段涉及引入与目标领域相关的符号表征。每个编程块代表一个特定的指令。在这一阶段，教师向全班介绍与目标 CT 概念相关的编程指令/块，并结合案例解释其功能和用途。以循环为例。首先，教师向全班展示了循环程序块:“在 Matatalab 编程工具包中有相应的循环编程块:循环开始和循环结束(图 3)。”然后，教师通过案例解释了它们的功能和用途:“我们学过循环就是重复执行相同的指令，直到满足条件为止，你能认出这段代码中重复的指令吗?它们重复了多少次(图 4)?”孩子回答这个问题后，教师向全班解读了它们的功能和用法。



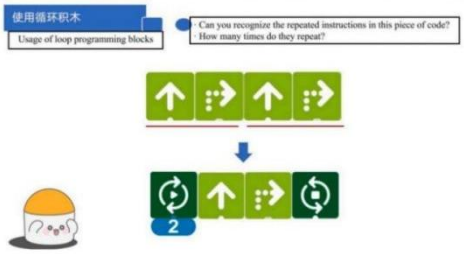


图 3 教师向学生介绍循环编程块 图 4 教师通过案例介绍编程块

第五阶段是针对目标域发布机器人编程任务。在这一阶段，教师针对 CT 概念的目标域发布机器人编程任务，并向儿童明确任务的要求和目标。孩子与同伴运用目标 CT 概念合作设计方案、编写、执行并调试程序，最终完成编程任务。教师观察孩子们的表现，并在必要时提供帮助。再以循环概念为例：教师发布了相应的编程问题解决任务:“ 玛塔四周转一转(图 5)。我们刚刚学习了循环，并将利用循环来解决机器人编程任务。请帮助玛塔实现目标。”

图 5 教师给学生布置机器人编程任务

第六阶段是分享与反思。在这一阶段，教师组织孩子们分享自己解决问题的方案和机器人编程作品，随后给予评价反馈并对本节课进行总结。接下来，孩子们对自己的方案和机器人编程作品进一步完善。

3.实验设计

3.1. 研究对象

本研究采用了准实验设计。参与者的纳入标准是他们以前没有相关的培训经验。研究从华东地区一所公立幼儿园自愿招募了 118 名 5-6 岁的儿童(平均=5.71 岁)。随机分配 60 名儿

童(32 名男孩和 28 名女孩)为实验组，58 名儿童(29 名男孩和 29 名女孩)为对照组。实验组采用 MRP 方法，对照组采用 CRP 方法。人口统计学变量(即年龄和性别)无明显差异，P>.05。

3.2. 实施保真度

在本研究中，我们努力确保高度的实施保真度:首先，两组均由同一位经验丰富、训练有素的教师授课。教师参加了三轮为期 2 小时的专业培训，以熟悉机器人编程工具、课程内容、隐喻教学过程等;其次，要求教师每周进行反思；最后，对两组的每节课进行录像，并由主要研究人员进行评估。

3.3. 测量工具

3.3.1.计算思维测试

为了评估的 CT 能力，我们改编了用于不插电的 CT 测试的工具——TechCheck-K。它是由 Rekin 和 Bers(2021 年)根据六个维度，即硬件/软件、算法、表征、模块化、控制结构和调试，专门为幼儿开发的 CT 测试工具，其 Cronbach’s α 为 0.78，达到了良好的内部信度。

3.3.2.儿童学习行为编码方案

本研究根据代表性的文献(Sun et al., 2021)，归纳出初步的儿童编程学习行为编码框架。接着对两组儿童的小规模样本进行了初步审查，检查是否捕捉到了所有行为特征。然后，专家小组就编码方案的适用性进行了讨论，并根据反馈意见对编码方案进行了完善。二位评分者之间的 Kappa 值为 0.82，表明一致性可以接受。表 2 显示了儿童编程学习行为的编码方案。

表 2 儿童编程学习行为的编码方案

代码	行为	说明
LT	听老师讲课	孩子们在课堂上听老师讲课。
IT	与教师互动	孩子们回答问题或向老师反馈。
DP	与同伴讨论	孩子们与伙伴讨论学习活动、编程和意见。
WP	编写程序	孩子们规划路线，摆放编程积木。
EP	执行程序	孩子们按下启动按钮，执行程序。
AH	寻求帮助	孩子们向老师或其他小组成员寻求帮助。
DE	调试	孩子们检查工具、编程块、步骤或路线，并尝试改进他们。
ET	扩展任务	孩子们在完成既定任务后，会自己尝试其他解决方案或者拓展任务。
SH	分享	孩子与教师或同伴分享他们的解决方案或任务。
IB	无关行为	孩子们聊天、玩耍或表现出其他不相关的行为。

3.4. 实验流程

在招募参与者后，儿童需要完成 15 分钟的 CT 前测。随后，在幼儿园科学教室开展了为期 10 周的机器人编程课程(每周 40 分钟)。实验组采用 MRP 方法，将隐喻方法融入机器人编程教学。对照组则采用 CRP 方法，侧重于传统的讲授的机器人编程教学。实验结束后，两组均立即进行了 CT 后测。

3.5. 数据分析

针对问题 1，我们进行了单因素方差分析(ANCOVA)，通过控制两组幼儿的 CT 前测得分，研究两组幼儿 CT 后测的差异。所有数据均使用 SPSS26.0 软件进行分析。

针对问题 2，研究人员对幼儿的课堂表现进行录像，并特意选取了后期幼儿对所采用的教学方法已经比较熟悉的两节活动课对两组幼儿的学习行为进行编码。当儿童做出特定行为时，研究人员会将其记录下来。研究人员通过 GSEQ 软件对儿童的学习行为进行了滞后序列

分析(LSA)。Z 值大于 1.96 代表关系显著。然后，绘制学习行为转换图来比较两组的行为模式。

4.结果分析

4.1. 儿童计算思维分析

Shapiro-Wilk 检验的结果表明，本研究收集的所有数据都近似于正态分布(p>.05)。

在控制 CT 前测得分的情况下，对儿童的后测得分进行了方差分析比较(表 3)。Levene 方差齐性检验的结果(F=1.39，p=0.242>0.05)显示无显著性水平，说明满足了组间方差齐性假设。因此，可以进行方差分析。方差分析结果显示，在控制了儿童测试前的 CT 分数后，儿童测试后的 CT 分数存在显著差异(F=40.29,p<0.001, η 2=0.26)。实验组儿童测试后 CT 分数的调整后均值和标准误差为 11.98 和 0.19，而对照组则分别为 10.33 和 0.18。研究发现，实验组的 CT 测试后得分高于对照组，这表明 MRP 组在 CT 方面的表现优于 CRP 组。

此外，还对 CT 的各个维度进行了方差分析。结果显示，两组学生在 CT 的四个维度上存在显著差异，即算法,序列(F=26.80，p<0.001， η 2=0.17)、表征(F=5.90,p<0.05, η 2=0.04)、控制结构(F=5.17,p<0.05, η 2=0.04)和调试(F=20.28,p<0.001, η 2=0.11)，这意味着 MRP 方法比 CRP 方法更有利于促进儿童学习算法/序列、表征、控制结构(包括循环)和调试。

表 3 实验组和对照组的测验后 CT 分数的方差分析结果

维度	组别	N	Mean	SD	Adjusted Mean	Std.error	F	η 2
整体	实验组	60	11.95	1.48	11.98	0.19	40.29***	0.26
	对照组	58	10.36	1.77	10.33	0.18		
硬件/软件	实验组	60	1.67	0.48	1.66	0.06	0.07	0.00
	对照组	58	1.64	0.48	1.64	0.06		
算法/排序	实验组	60	3.77	0.87	3.73	0.12	26.80***	0.17
	对照组	58	2.84	1.04	2.88	0.12		
表征	实验组	60	1.37	0.55	1.30	0.06	5.90*	0.04
	对照组	58	1.02	0.44	1.08	0.06		
模块化	实验组	60	1.52	0.50	1.54	0.06	0.56	0.00
	对照组	58	1.50	0.50	1.48	0.06		
控制结构	实验组	60	1.88	0.32	1.91	0.04	5.17*	0.04
	对照组	58	1.79	0.41	1.77	0.04		
调试	实验组	60	1.75	0.44	1.83	0.05	20.28***	0.11
	对照组	58	1.57	0.50	1.49	0.05		

4.2. 儿童学习行为模式分析

实验组共收集到 998 个事件代码，对照组收集到 968 个事件代码。在 MRP 实验组中，最常见的行为是 WP(即编写程序，15.6%)、DP(即与同伴讨论，15.2%)和 LT(即听老师讲课，14.0%)。对照组儿童的 IB(即无关行为，16.5%)、WP(即编写程序，12.4%)和 DP(即与同伴讨论，11.8%)表现较多。相比之下，对照组出现了更多的无关行为(如聊天分散注意力或玩耍)。

根据表 4 和表 5 所示的两组调整后残差表，MRP 实验组有 16 个显著序列，而对照组有 23 个显著行为序列。

两组重要行为序列的过渡矩阵图分别如图 6 和图 7 所示。黑线表示两组同时出现的重要序列，而红线表示只存在于实验组或对照组的重要序列。如图 6 所示，在教学过程中，实验组幼儿反复聆听教师的讲解并与教师互动(LT→IT→LT)，然后与同伴讨论(LT→DP)。在与同

伴讨论后，孩子们尝试编写程序，然后较少的孩子会寻求帮助(WP→AH)，而较多的孩子会执行程序(WP→EP)。当儿童完成执行程序后，一些儿童调试然后分享作品(EP→DE→SH)，而另一些儿童则直接分享作品(EP→SH)。实验组的孩子们反复分享和扩展学习任务(SH→ET→SH)。这表明，在实验组中使用隐喻学习编程有助于儿童掌握基本概念和技能，因为他们无需太多帮助就能完成编程作业，甚至能自发地扩展任务。

从图 7 中可以发现，对照组的儿童除了参与教学过程的儿童(LT→IT→LT)外，都会出现一些不相关的行为(IT→IB→LT)，这可能是因为有些儿童不能理解抽象的概念。在编写程序的环节中，大多数儿童喜欢与同伴反复讨论(WP→DP)，也有一些儿童会执行程序(WP→EP)或寻求帮助(WP→AH)。之后，他们往往会在执行程序(AH→EP→AH)或调试(AH→DE→AH)过程中反复请求他人帮助。这表明，对照组可能很难在没有他人帮助的情况下完成编程任务。此外，孩子们会在调试或扩展任务(DE→SH，ET→SH)后进行分享。此外，在扩展和分享任务后会做出一些不相关的行为(ET→IB，SH→IB)。

表 4 实验组调整后的残差值

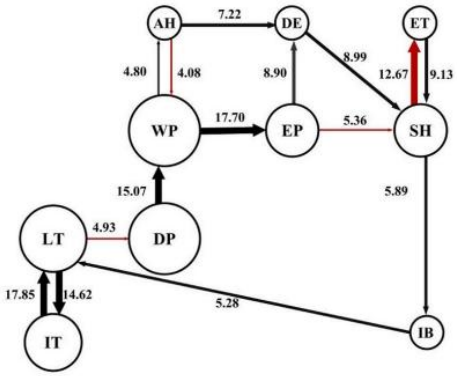
Given	LT	IT	DP	WP	EP	AH	DE	ET	SH	IB
LT	-5.05	14.62*	4.93*	-1.91	-4.26	-2.20	-2.55	-3.38	-4.35	1.23
IT	17.85*	-3.00	-2.70	-3.52	-2.79	-1.44	-1.67	-2.21	-2.85	-0.43
DP	-3.77	-3.41	-3.95	15.07*	-3.18	-1.64	-1.90	-2.02	0.81	0.88
WP	-3.88	-3.52	1.41	-3.80	17.70*	4.80*	-1.96	-2.60	-3.35	-2.88
EP	-3.08	-2.79	1.66	-3.28	-2.60	2.09*	8.90*	-0.90	5.36*	-0.68
AH	-1.59	-1.44	-0.94	4.08*	-1.34	-0.69	7.22*	-1.06	-1.37	-1.18
DE	-1.80	-1.63	-1.88	-1.91	-1.52	1.96	-0.91	-0.27	8.99*	0.37
ET	-1.42	-1.28	-0.67	-1.51	-1.20	-0.62	-0.72	-0.95	9.13*	-1.05
SH	-2.28	-2.07	-2.39	-2.43	-1.93	-0.99	-1.15	12.67*	-1.97	5.85*
IB	5.28*	-2.45	1.25	-1.98	-2.28	-1.18	-1.37	5.31	-1.28	-2.00

* Z > 1.96

表 5 对照组调整后的残差值

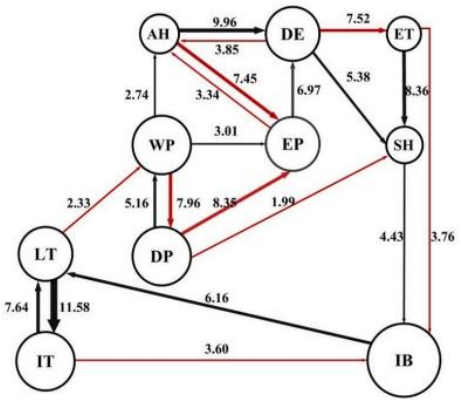
Given	LT	IT	DP	WP	EP	AH	DE	ET	SH	IB
LT	0.00	11.58*	-0.89	2.33*	-4.09	-3.29	-3.45	-2.99	-2.64	0.02
IT	7.64*	0.00	-3.09	0.03	-2.85	-2.29	-2.40	-2.08	-1.84	3.60*
DP	-3.21	-3.09	0.00	5.16*	8.53*	-2.32	-0.40	-2.10	1.99*	-3.87
WP	-3.32	-3.20	7.96*	0.00	3.01*	2.74*	-2.51	-2.17	-1.92	-1.27
EP	-2.96	-2.85	-0.69	-2.98	0.00	7.45*	6.97*	-1.94	-1.03	0.18
AH	-2.24	-2.16	0.60	-2.25	3.34*	0.00	9.96*	-1.46	-1.29	-2.70
DE	-2.32	-2.24	-2.26	-2.34	-1.51	3.85*	0.00	7.52*	5.38*	-0.96
ET	-1.62	-1.56	-1.58	-1.63	-1.45	-1.17	-1.23	0.00	8.36*	3.76*
SH	-1.35	-1.30	-1.31	-1.36	-1.21	-0.97	-1.02	4.08*	0.00	4.43*
IB	6.16*	-0.83	-0.52	-0.35	-3.23	-2.60	-2.72	4.97*	-1.50	0.00

* Z > 1.96



代码	行为
LT	听老师讲课
IT	与教师互动
DP	与同伴讨论
WP	编写程序
EP	执行程序
AH	寻求帮助
DE	调试
ET	扩展任务
SH	分享
IB	无关行为

图 6 实验组重要行为序列的过渡图



代码	行为
LT	听老师讲课
IT	与教师互动
DP	与同伴讨论
WP	编写程序
EP	执行程序
AH	寻求帮助
DE	调试
ET	扩展任务
SH	分享
IB	无关行为

图 7 对照组重要行为序列的过渡图

此外，实验组有四种行为模式达到了显著水平。LT→DP，即孩子们在听课后愿意与同伴讨论(LT→DP)。AH→WP 表示他们在寻求帮助后会继续编写程序。最理想的情况是 EP→SH 和 SH→ET:孩子们会在执行程序后转而分享他们的作品，而不是寻求帮助或调试，或者他们会在分享后扩展他们的学习任务。对照组中有 10 对达到显著水平的序列。LT→WP 代表孩子们先听课然后编写程序。WP→DP、DP→EP 和 DP→SH 表示儿童在编写程序后与同伴讨论，然后执行程序或进行分享。DE→ET 表示儿童调试并扩展任务。不过，也有一些不足之处:AH→EP 和 EP→AH 表示幼儿会反复寻求帮助并执行程序;IT→IB 和 ET→IB 表示幼儿在与教师互动或扩展任务后会做出与任务无关的行为。

5.讨论和结论

研究采用准实验设计，探讨了所提出的“隐喻式机器人教学法”对 CT 和行为模式的影响。结果显示，参加 MRP 组的儿童在 CT 方面的表现优于 CRP 组。此外，行为分析表明，提出的 MRP 方法可促进儿童取得优异的学习成绩和更积极的行为，从而帮助他们实现学习目标。

针对研究问题一，实验组在整体 CT 方面的表现优于对照组，并且在算法、序列、表示法、控制结构和调试方面的表现优于对照组。这表明与 CRP 方法相比， MRP 方法能更好地提高幼儿的 CT 能力。这一发现与 Pe'rez:Marin 等人(2020)的研究结果一致，即在编程学习中提供隐喻可以增强学习者的 CT 能力。从理论上讲，隐喻是一种有益的教学工具，能够将抽象的概念与儿童熟悉的生活经验联系起来，帮助学习者从已知概念迁移到未知概念，更好地促进学习者的学习成绩和学习表现。在本研究中教师通过旋转木马等生活实例来解释循环概念，使孩子们更容易理解并应用到机器人编程任务中，从而发展他们的 CT。皮亚杰的认知

发展理论指出，前运算阶段儿童的概念知识建构依赖于更具体、更熟悉的经验来建构抽象概念(Piaget, 1953)。因此，如果能在机器人编程教学中提供隐喻，将能有效促进幼儿 CT 的发展。

针对研究问题二，滞后序列分析显示，MRP 组的儿童与 CRP 组表现出不同的行为模式。实验组儿童表现出以下特点：(1)更专注于教学过程;(2)更投入地完成编程任务;(3)更多儿童在完成规定任务后投入到拓展学习任务中。这可能是因为在教授儿童机器人编程时，提供隐喻可以帮助教师建立一个抽象的 CT 概念与生动具体的内容之间的联系，使得儿童可以很好地感知、理解和建构相应的 CT 概念，并将其迁移和应用于完成机器人编程任务(P´erez-Marín et al., 2018, 2020)。此外，从更多儿童在完成规定任务后表现出主动扩展学习任务的积极行为来看，MRP 方法可以丰富儿童的编程学习体验，激发儿童的技能迁移能力和发散思维。至于 CRP 对照组，也可以得出三个主要结论:(1)对解决问题的任务更加茫然；(2)在调试程序时更多地寻求帮助；(3)课堂上表现出更多的无关行为，这与詹晓明等人(2022)的发现相似。这反映了传统机器人编程教育中存在的问题，即如果没有提供适当的教学方法，儿童可能难以理解和掌握基本概念，并试图寻求帮助或在学习任务中表现不佳。因此，所建议的基于隐喻的机器人编程方法可促进儿童取得优异的学习成绩和更积极的行为，从而帮助他们实现学习目标。

6.结语

本研究提出基于隐喻的机器人编程方法，并通过实验研究证实了其对 CT 和行为模式的积极影响，在编程和 CT 教育的教学方法方面做出了贡献。此外，本研究仍存在一些局限性。首先，本研究的参与者来自华东地区的一所幼儿园，这可能不具有普遍适用性。其次，本研究只实施了有限的干预，没有评估干预的长期效果。基于此，我们对后续研究提出了几点建议:扩大样本多样性和年龄范围；考虑个体差异(如性别、编程经验等)；如何提升自我调节和参与度，减少无关行为；继续开发有效教学法。

参考文献

Bati, K. (2022). A systematic literature review regarding computational thinking and programming in early childhood education. *Education and Information Technologies*, *27*(2), 2059-2082.

Beilin, H. (1971). Developmental stages and developmental processes. *Measurement and piaget*, 172-188.

Bers, M. U., González-González, C., & Armas–Torres, M. B. (2019). Coding as a playground: Promoting positive learning experiences in childhood classrooms. *Computers & Education*, *138*, 130-145.

Lakoff, G. (1993). The contemporary theory of metaphor.

Milner, W. W. (2010). A broken metaphor in Java. *ACM SIGCSE Bulletin*, *41*(4), 76-77.

Peelle, H. A. (1983). Computer metaphors: Approaches to computer literacy for educators. *Computers & Education*, *7*(2), 91-99.

Pérez-Marín, D., Hijón-Neira, R., Bacelo, A., & Pizarro, C. (2020). Can computational thinking be improved by using a methodology based on metaphors and scratch to teach computer programming to children?. *Computers in Human Behavior*, *105*, 105849.

Reinhardt, J. (2020). Metaphors for social media-enhanced foreign language teaching and learning. *Foreign Language Annals*, *53*(2), 234-242.

Relkin, E., de Ruiter, L. E., & Bers, M. U. (2021). Learning to code and the acquisition of computational thinking by young children. *Computers & education*, *169*, 104222.

Saban, A. (2006). Functions of metaphor in teaching and teacher education: A review essay. *Teaching education*, *17*(4), 299-315.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.

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.

運用 VEX 機器人學習编程：提升 K-12 教育中的計算思維與解決問題能力

Learning to Program with VEX Robotics: Enhancing Computational Thinking and Problem-Solving Skills in K-12 Education

Iek Chong Choi 1, Biyun Huang 1*, Qiqi Chen 1, Biaoxin Mai 2
1 School of Education, City University of Macau, Macao SAR, China
2 Honors Class of Faculty of Data Science, City University of Macau, Macao SAR, China
* byhuang@cityu.edu.mo

【摘要】 隨著科技的迅速發展所帶來的技術創新與數位化轉型的需求不斷提升，K-12 階段的資訊科技教育變得愈加重要。然而，對初學者而言，學習程式設計具有挑戰性，因為它需要邏輯應用與問題解決等複雜的認知能力。許多學生即使在入門課程中也感到困難。本研究參訪問了兩位中學階段的電腦和機器人教師，以瞭解了他們對於將 VEX 機器人運用于編程教學中的實踐經驗與看法。訪談結果表明，VEX 機器人，爲提升學生的計算思維及解決問題能力提供了有效途徑。而且，通過強調創新與互動的教學方式，有助於提升學生的學習動機與參與度。

【關鍵字】 K-12 教育；學習動機；計算思維；認知負荷；VEX 機器人

Abstract: *The importance of computer science education in K-12 has increased dramatically in response to the demand for technological innovation and digital transformation brought about by rapidly advancing computational technology. However, mastering programming can be challenging for beginners, requiring complex cognitive abilities such as logical application and problem-solving. Many students struggle, even in introductory courses. This study interviewed two K-12 STEM teachers to explore their practical experiences and perspectives on integrating VEX robots into programming instruction. The interview results indicate that VEX robots effectively enhance students' computational thinking and problem-solving abilities. Furthermore, the strategy promotes innovation and interaction, which increases students' learning motivation and engagement.*

Keywords: K-12 Education, Learning Motivation, Computational Thinking, Cognitive Load, VEX Robotics

1.前言

隨著未來科技發展驅動社會對技術創新與數位轉型需求的增加，K-12 年級階段的資訊科技教育重要性亦日益提升 (Chen et al., 2017)。21 世紀的來臨預示著計算思維（Computational Thinking, CT） 已成為一項關鍵能力，不僅在計算機科學領域中舉足輕重，在其他領域亦具深遠影響。CT 強調系統設計、問題解決能力與邏輯推理等核心技能的培養 (Wing, 2006)，構成了演算法與程式邏輯的基礎，對於開發高效能電腦應用程式至關重要 (Aho, 2012)。

然而，培養學生的計算思維能力是一項充滿挑戰的任務。研究指出，儘管學生能依照教師的示範撰寫程式，卻常無法理解程式結構背後的原理 (Lye & Koh, 2014)。當被要求獨立開

發程式時，學生往往難以正確完成，亦無法有效運用計算思維來解決問題。Guzdial (2015) 指出，學生在程式設計中所面臨的困難涵蓋錯誤觀念、練習不足及學科本身的複雜性。

初學者常覺得程式設計困難且複雜 (Sáez-López et al., 2016)。許多學生即使在入門課程中也面臨困難，因為掌握程式語言對他們而言具有挑戰性 (Mow, 2008)。同時，對於文字程式語言的文法理解及指令輸入，學生普遍感到困難，這些問題難以完全消除 (Wae-Shik, 2006)。傳統教學方法已無法有效應對多元學生的學習需求 (Papadakis & Kalogiannakis, 2018)。研究指出，互動式教育工具與圖形化程式語言能有效調節學習時的認知負荷，進而提升學習成效與學習動機 (W. C. Choi & Choi, 2024b)。

在程式設計教育中應用機器人，能將抽象概念以視覺化與實體化的方式呈現，有助於學生理解電腦科學原理。此外，學生在排除硬體故障與除錯程式碼過程中，能進一步培養問題解決能力。若結合團隊合作進行機器人編程專案，學生能透過交流、協作與互助，發揮各自的優勢 (Blanchard et al., 2010)。與機器人一起進行的活動會使課程更加愉快和有趣，因此，機器人成為培養積極態度的適當手段。此外，機器人常被視為發展多種能力的工具，如問題解決、創意思維、工程設計、科學探究與團隊合作等 (Altin & Pedaste, 2013; Darmawansah et al., 2023)。

然而，機器人教學在實際應用上仍面臨挑戰，特別是在教師與學生比例偏高的教室中，對實體設備與支援系統的需求較高，限制了其推廣與實施。這些挑戰突顯出，在真實教學環境中需尋找創新策略以維持教育品質。

為深入了解機器人教學對學生計算思維與問題解決能力的影響，本研究從教師的視角出發，進行訪談調查。透過蒐集教師與學生互動的第一手經驗，我們旨在獲得學生於課堂中與機器人互動的寶貴見解。藉由這些訪談，我們試圖揭示教師在培養學生計算思維與問題解決能力上的觀察、經驗與教學策略。此研究方法使我們能從教學第一線的獨特視角，探討影響學生學習動機的各種關鍵因素。本研究探討了以下研究問題：

1. 在中學程式設計課程中使用 VEX 機器人，如何影響學生的計算思考與問題解決能力？

從教師的角度來看，VEX 機器人的教學優點是什麼？

2.文獻探討

2.1. 計算思維

計算思維 (CT) 已逐漸被認定為 21 世紀的重要技能，強調系統設計、解決問題，以及藉由汲取電腦科學的基本概念來理解人類行為。此外，CT 與程式設計的表現有顯著的相關性 (W.-C. Choi et al., 2024)。此外，CT 涉及軟體工程師在撰寫電腦應用程式時所使用的問題解決技巧與技術 (Wing, 2006)。

Tang 等人 (2020) 將計算思維的定義分為兩大類別。第一類著重於技術層面，例如演算法設計、抽象化、分解等形式化的程式設計技能。第二類則強調更廣泛的能力，包括從事運算任務所需的技能、可應用於各種領域的一般性問題解決能力、批判性思維與學科專屬技能。

Selby 和 Woollard (2013) 進行了一項文獻回顧，並依據 Wing (2006) 的定義，提出計算思維歷程中的五個基本要素（見表 1），這些要素可應用於各種問題解決情境中，具高度普適性。

表 1 計算思維的五個基本構面與定義

構面	定義
抽象化	簡化複雜性以解決問題
分解	將問題拆解為較小且可管理的部分，以便逐步解決
演算法思維	以逐步程序的方式規劃解決方案
評價	辨識出最合適的問題解決方法
泛化	辨認出特定問題的解決模式，並將其套用至其他相似的情境中

此外，Brennan 與 Resnick's (2012) 所提出的計算思維框架，為在教育場域中教授這項關鍵素養提供了基礎途徑，使學生能在日益數位化的世界中有效學習與發展。該架構強調教學上的彈性與適應性，並圍繞三個核心方面進行設計：概念（Concepts）、實踐（Practices）與觀點（Perspectives）。這三大方面彼此協同運作，不僅能強化學生的計算思維能力，亦促進其批判性思考、問題解決與創造力，而這些能力皆具備高度的可轉移性，能應用於各種學科與實際生活情境之中 (Zhang & Nouri, 2019)。

2.2. 問題解決能力

計算思維（CT）是一種以電腦科學、系統設計與人類行為理解為基礎的問題解決模式 (Wing, 2006)。其核心不僅在於撰寫程式，更在於一種可被轉化為演算法與步驟的思考方式，進而解決複雜問題 (Aho, 2012)。因此，計算思維與問題解決能力密不可分，常被視為彼此高度相關的認知技能。

根據 Wing (2006)，在兒童教育早期，除了基本的讀寫與算術能力外，也應培養兒童理解資訊在解決問題過程中的角色，並學習如何運用邏輯推理與電腦邏輯技能。這些能力的養成，需要學生能將現實生活中的問題，轉化為結構化、可分析與可解決的問題模型，並藉由電腦科學的觀念與工具予以處理 (Yildiz Durak, 2020)。

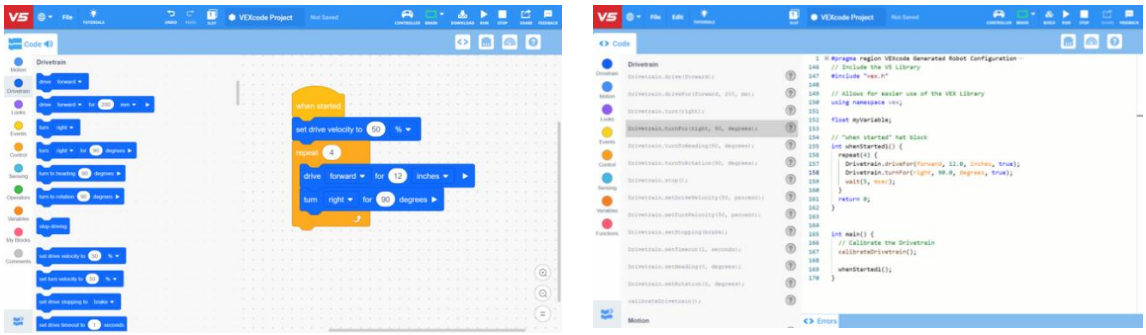
在程式設計學習中，僅具備撰寫程式碼的能力，並不足以有效解決實際問題。當程式設計訓練被視為一個涵蓋問題解決歷程的學習過程時，發展學生的問題解決能力便成為不可或缺的要素 (Gomes & Mendes, 2007)。

在此過程中，教育機器人提供了一個理想的平台，讓學生能以實體化的方式觀察其程式碼輸出對實際環境的影響，進一步強化反思性思考的誘發。例如，當學生觀察到機器人行為與預期不符時，會自發性地進行問題診斷與邏輯推演，這種即時回饋與可視化結果能大幅提高學生的動機與解題能力 (Sapounidis & Alimisis, 2020)。透過如機器人編程等實作活動，學生能夠在處理錯誤與修正的過程中，培養反思性的問題解決能力，將抽象概念轉化為實體操作與認知反饋，進一步強化其邏輯思維與策略應用能力 (Yildiz Durak, 2020)。

2.3. 教育機器人

將機器人融合進 STEM 教育的理念，源自建構主義學習理論，強調「動手做」的學習方式，主張學習應透過與環境的積極互動與參與來進行 (Papert, 2020)。近年的研究進一步驗證了這類教學方法的有效性，指出機器人學習活動能夠提升學生的計算思維能力 (Budiyanto et al., 2022)，也能增進對 STEM 科目的學習興趣 (Drakatos & Stompou, 2023)。機器人學習是一個極佳的 STEM 應用環境，有助於學生在實作中掌握跨領域知識 (Darmawansah et al., 2023)。此外，機器人也愈來愈被視為銜接非正式與正式 STEM 教育的橋樑。透過動手做的學習模式，不僅能提升學生的問題解決能力，也能有效提升參與感與學習動機 (Jäggle et al., 2020)。

在編程機器人套件中，VEX 機器人具備極高的機構設計彈性，使學生能自由探索並創造創新解決方案 (Liu, 2009)。VEX 機器人套件包含控制器、感測器、馬達、齒輪組與程式模組，提供一個完整的平台，讓學生投入於實作導向的 STEM 教育中。這種沉浸式的學習經驗能激發學習熱情，並在過程中帶來成就感與學習滿足。VEXcode 是一款設計給各年齡層學生使用的程式編寫環境，適用對象涵蓋從小學到高中的學生。 VEXcode 支援三種程式語言：圖形化（Block-based）、Python 以及 C++（見圖. 1）。圖形化語言強調視覺化編程，學生可透過拖拉與排列程式區塊的方式進行設計，無需撰寫複雜的文字程式碼，特別適合年齡較小的學生或初學者。此外，VEX 機器人競賽（見圖. 2）為學生提供了展現其機器人設計與編程能力的機會。這不僅能促進解決問題能力的發展，亦有助於培養領導力、溝通能力與批判性思考等軟實力，這些能力對未來從事 STEM 領域職涯至關重要 (Sullivan & Bers, 2019)。



(a) 圖形化語言　　(b) Python，C++

圖 1 VEXCode 編碼語言

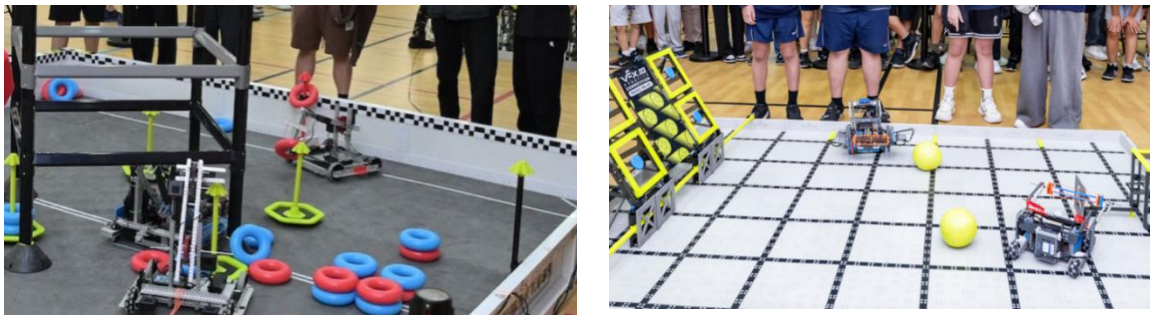


圖 2 VEX 機器人競賽

3.研究方法

本研究為一項質性研究，研究采訪了兩位教授電腦和機器人的教師, 以瞭解他們對將機器人運用于編程教學的看法。本文採用一對一的半結構式訪談 (Rubin & Rubin, 2011)。半結構式訪談通常以特定且預先設計的問題作為開端 (Strauss & Corbin, 1998)，但訪談過程中可隨著受訪者的回答自然延伸，涵蓋事先未預見的主題，具有彈性與深度幷重的特性。通過對兩位教師的訪談,我們瞭解到當 VEX 機器人運用到編程教學中時，實際產生的效果。

3.1.參與受訪者

我們選取了兩位在澳門一所國際中學有使用 VEX 教學經驗的機器人老師參與，並進行了一對一的訪談。其中一位是外籍教師，有 8 年教授電腦科學和機器人的經驗，另一位是本地教

師，有 4 年相同領域的教學經驗。他們在學校教授電腦科學和機器人，並參與機器人活動和比賽。為保障受訪者的隱私與研究倫理，本研究對兩位受訪教師皆採用化名處理。

3.2. 資料收集與分析

Selby 和 Woollard (2013) 解決問題的思考過程中的五個基本要素為基礎，開發了計算思考量表被採納為分析資料的指導框架。為了確保資料的準確性，訪談內容都經過錄音和逐字記錄 (Gibbs, 2018)。這種全面的方法讓我們能夠徹底檢視機器人教師所分享的細微差異和洞察力。此外，研究亦透過比對不同受訪教師之間的觀點，並與相關文獻進行交叉分析，以提升研究結果的客觀性與一致性 (Creswell & Poth, 2016)。

4.研究結果與討論

4.1. 透過機器人教育簡化程式設計概念

機器人教育的一項關鍵在於教導學生如何應用有序的步驟來組建可運作的機器人，如讓機器人移動並拆解為一連串可管理的順序步驟。透過這樣的歷程，學生也能理解如何有效地重複執行特定動作，還能學會如何優化這些步驟以提高效率。這對于設計能執行重複性任務的機器人來說至關重要，例如編寫程式讓機器人持續向前移動，直到偵測到障礙物為止。教師 B 有進一步描述,學生如何通過訓練機器人的過程來學習程式設計。

“在這里，您可以馬上訓練機器人循環，讓它向前、向後開動，還有……我們也是以類似的方式教他們變量和函數。”[教師_B]

教師表示這種任務分解和優化的過程鍛煉了學生的邏輯思維能力，也為他們未來在 STEM 領域的學習和職業發展奠定了堅實的基礎。

4.2. 將問題拆解為可管理的部分

學生在面對複雜的機器人專案時，若能將整體任務拆解為較小且可管理的部分，不僅能減少壓力，也能讓整個學習過程更有系統性，進而幫助學生保持專注與學習動機。Selby 與 Woollard (2013) 把這種將複雜任務拆解為較小部分的行為被稱為“任務分解”(Task Decomposition)。是一種在電腦科學和工程領域中廣泛使用的策略。通過任務分解，學生可以逐步攻克難題，逐步構建起複雜的機器人系統。這種策略也使教師能設定清楚且可達成的目標，讓學生在學習過程中能清楚追蹤進度並獲得階段性的成就感。

“以這種方式給孩子們設定小目標，讓他們實現，他們會不斷地收到。當你要求他們完成一項任務時，他們可以看到自己在進步。他們認為這是一個可以實現的目標。”[教師_B]

透過逐步處理每一階段任務，學生學會以有條理的方式應對挑戰，此過程同時培養他們在問題解決中的堅韌力與持續性 (Barak & Zadok, 2009)。

4.3. 實際操作中的編程技能應用

教育機器人結合了機械組件、感測器與編程平台，為學生提供一個能夠進行實驗、測試與修正程式的實作空間。教師 A 提到，學生在程式設計能力上的明顯進步，顯示出機器人強大的學習增能效益。類似的研究指出，這類互動式平台有助於循序漸進的培養關鍵的計算思維能力 (Durak et al., 2019)。

“這表明他們的編程技能有了提高。從最初搭建機器人開始，他們正逐漸賦予機器人生。”[教師_A]

4.4. 問題解決能力的培養

兩位教師皆強調，學生在實務問題解決能力方面有明顯的進步，特別是在處理連接問題、感測器與致動器故障等技術挑戰時表現出色。透過反覆測試與調整機器人系統的過程，學生

被迫面對錯誤與障礙，進而培養出批判性思維與有效的問題解決策略。在不斷嘗試解決技術問題的過程中，學生逐漸建立起對自己解決問題能力的信心，也表現出更高的堅持度與抗挫力，當面對未來挑戰時能更積極應對 (Zhong & Si, 2021)。

“他們在調試、理解錯誤信息、運用感覺-思考-行動循環以及通過閱讀被壓抑的信息來識別誤解方面做得更好。”[教師_B]

“最顯著的特點是解決問題的能力.....隨著機器人不斷適應和執行大量的程序來完成委托給它們的各種任務，它們自然會提高解決問題的能力。”[教師_A]

4.5.協作中的團體意識

透過將項目式學習（Project-Based Learning, PBL）融入機器人教學，學生能在有意義的專案中進行協作，共享想法、發揮個人長處，並互相支持。教師 B 認爲，這種學習模式鼓勵有效的溝通與衝突解決能力，這對於培養團隊合作精神至關重要。Lee 等人 (2015) 的研究也發現，在完成一個集體任務時，學生會在衝突中學會合作，提升衝突管理的能力。

“他們正在分配任務、承擔職責，並確保每個人都達成一致，因為他們已經學到了很多關於沖突管理的知識。”[教師_B]

4.6.多元程式語言提升操作靈活性

VEXcode 提供學生選擇多種程式語言實作的機會，有助於加深其程式理解力與靈活運用能力。VEXcode 支援圖形化 (block-based)、Python 和 C++ 等多種編程方式，能滿足不同程度學生的學習需求，從初學者到進階使用者皆可受益。

“最小的孩子使用 VEX 代碼平臺進行基於模塊的編碼。年齡較大的學生也使用 VEX 代碼，但他們的級別選擇是使用 Python 還是 c++。”[教師_B]

多元語言的設計使學生能夠依據自身節奏進行學習，同時體驗不同的編程環境。特別是積木式語言，透過視覺化介面來簡化程式學習流程，對初學者尤其友善。其拖曳式設計降低了因語法錯誤所產生的挫折感，讓新手沒有文字語法的負擔。隨著學生對程式設計的理解逐漸加深，他們便可以自然地從圖像過渡到文本語言，如 Python 或 C++方式，進一步拓展其程式設計技能。(W. C. Choi & Choi, 2024a; Hsu et al., 2017)。

4.7.遊戲化與競賽元素

遊戲化的學習結構透過設計明確的挑戰與目標，讓學生能專注於具體成果導向的學習過程。這種方式模擬真實世界的任務情境，使學習更加有趣且具情境相關性，進一步提高學生的參與度與學習動機。此外，在教育情境中導入機器人競賽等競爭性元素，也能激發學生的成就動機，促使他們更積極投入於程式設計與任務實作中。

“任何時候，孩子們都像機器人一樣急於完成某件事，好像這能改變什麼，而他們確實做到了。”[教師_B]

“組織一項活動，如小型錦標賽。每個人都想打敗另一個團隊。”[教師_A]

競賽元素的挑戰性和時間限製促使學生在實踐中不斷優化解決方案，從而提升了他們的應變能力和創新思維。這不僅增加了投入感，也有助於學生培養自我挑戰與持續精進的態度。

5.結論

將機器人融合進教育中，為學生提供了一個強大的學習平台。透過在學習歷程中融入計算思維、問題解決與合作能力，能有效培育學生面對數位時代複雜挑戰的能力。透過項目式導向學習 (Project-Based Learning)，學生獲得實作經驗，能夠將理論知識應用於實際任務，

縮短了抽象概念與真實應用之間的距離。而遊戲化與競賽元素的引入，更提升了學習的趣味性與參與度，使學習過程兼具教育性與娛樂性，進一步激勵學生勇於迎接挑戰並持續學習。VEXcode 這類結合視覺化程式環境的教學工具，有助於降低學習門檻，使抽象的計算概念更易於理解。同時,由於支持多種編程語言，VEX 機器人滿足了不同層次學生的學習需求，為 K-12 階段的計算思維教育提供了一種創新且有效的路徑。

參考文獻

Aho, A. V. (2012). Computation and computational thinking. The Computer Journal, 55(7), 832–835.

Altin, H., & Pedaste, M. (2013). Learning approaches to applying robotics in science education. Journal of Baltic Science Education, 12(3), 365.

Barak, M., & Zadok, Y. (2009). Robotics projects and learning concepts in science, technology and problem solving. International Journal of Technology and Design Education, 19(3), 289–307.

Blanchard, S., Freiman, V., & Lirrete-Pitre, N. (2010). Strategies used by elementary schoolchildren solving robotics-based complex tasks: Innovative potential of technology. Procedia-Social and Behavioral Sciences, 2(2), 2851–2857.

Brennan, K., & Resnick, M. (2012). Using artifact-based interviews to study the development of computational thinking in interactive media design. Annual American Educational Research Association Meeting, Vancouver, BC, Canada, 1–25.

Budiyanto, C. W., Fenyvesi, K., Lathifah, A., & Yuana, R. A. (2022). Computational Thinking Development: Benefiting from Educational Robotics in STEM Teaching. European Journal of Educational Research, 11(4), 1997–2012.

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. Computers & Education, 109, 162–175.

Choi, W. C., & Choi, I. C. (2024a). Exploring the Impact of Code. Org's Block-Based Coding Curriculum on Student Motivation in K-12 Education. 2024 12th International Conference on Information and Education Technology (ICIET), 93–97.

Choi, W. C., & Choi, I. C. (2024b). Investigating the Effect of the Serious Game CodeCombat on Cognitive Load in Python Programming Education. 2024 IEEE World Engineering Education Conference (EDUNINE), 1–6.

Choi, W.-C., Lam, C.-T., & Mendes, A. J. (2024). How Various Educational Features Influence Programming Performance in Primary School Education. 2024 IEEE Global Engineering Education Conference (EDUCON), 1–8.

Creswell, J. W., & Poth, C. N. (2016). Qualitative inquiry and research design: Choosing among five approaches. Sage publications. https://www.google.com/books?hl=zh-TW&lr=&id=DLbBDQAAQBAJ&oi=fnd&pg=PP1&dq=Qualitative+inquiry+and+research+design:+Choosing+among+five+approaches&ots=-ir4a7GPWs&sig=jBzPDCRzN7p8sDeOFg85cdSbnkc

Darmawansah, D., Hwang, G.-J., Chen, M.-R. A., & Liang, J.-C. (2023). Trends and research foci of robotics-based STEM education: A systematic review from diverse angles based on the technology-based learning model. International Journal of STEM Education, 10(1), 12.

Drakatos, N., & Stompou, E. (2023). The perspective of STEM education through the usage of Robotics. World Journal of Advanced Research and Reviews, 18(3), 901–913.

Durak, H. Y., Yilmaz, F. G. K., & Yilmaz, R. (2019). Computational thinking, programming self-efficacy, problem solving and experiences in the programming process conducted with robotic activities. Contemporary Educational Technology, 10(2), 173–197.

Gibbs, G. R. (2018). Analyzing qualitative data. Analyzing Qualitative Data, 1–232. https://doi.org/10.4135/9781526441867.n1

Gomes, A., & Mendes, A. J. (2007). Learning to program-difficulties and solutions. International Conference on Engineering Education–ICEE, 7, 1–5. https://www.ineer.org/Events/ICEE2007/papers/411.pdf

Guzdial, M. (2015). Learner-centered design of computing education: Research on computing for everyone. Morgan & Claypool Publishers.

Hsu, S.-Y., Lou, Y.-F., Jung, S.-Y., & Sun, C.-T. (2017). Shelves: A User-Defined Block Management Tool for Visual Programming Languages. In R. Bernhaupt, G. Dalvi, A. Joshi, D. K. Balkrishan, J. O'Neill, & M. Winckler (Eds.), Human-Computer Interaction – INTERACT 2017 (Vol. 10515, pp. 335–344). Springer International Publishing.

Jäggle, G., Lammer, L., Hieber, H., & Vincze, M. (2020). Technological Literacy Through Outreach with Educational Robotics. In M. Merdan, W. Lepuschitz, G. Koppensteiner, R. Balogh, & D. Obdržálek (Eds.), Robotics in Education (Vol. 1023, pp. 114–125). Springer International Publishing.

Lee, D., Huh, Y., & Reigeluth, C. M. (2015). Collaboration, intragroup conflict, and social skills in project-based learning. Instructional Science, 43(5), 561–590. https://doi.org/10.1007/s11251-015-9348-7

Liu, Y. (2009). From Handy Board To Vex: The Evolution Of A Junior Level Robotics Laboratory Course. 2009 Annual Conference & Exposition, 14.641. 1-14.641. 11.

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? Computers in Human Behavior, 41, 51–61. https://doi.org/10.1016/j.chb.2014.09.012

Mow, I. T. C. (2008). Issues and Difficulties in Teaching Novice Computer Programming. In M. Iskander (Ed.), Innovative Techniques in Instruction Technology, E-learning, E-assessment, and Education (pp. 199–204). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8739-4_36

Papadakis, S., & Kalogiannakis, M. (2018). Using gamification for supporting an introductory programming course. The case of classcraft in a secondary education classroom. Interactivity, Game Creation, Design, Learning, and Innovation: 6th International Conference, ArtsIT 2017, and Second International Conference, DLI 2017, Heraklion, Crete, Greece, October 30–31, 2017, Proceedings 6, 366–375.

Papert, S. A. (2020). Mindstorms: Children, computers, and powerful ideas. Basic books.

Rubin, H. J., & Rubin, I. S. (2011). Qualitative interviewing: The art of hearing data. sage.

Sáez-López, J.-M., Román-González, M., & Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary school: A two year case study using "Scratch" in five schools. Computers & Education, 97, 129–141.

Sapounidis, T., & Alimisis, D. (2020). Educational robotics for STEM: A review of technologies and some educational considerations. In Science and mathematics education for 21st century citizens: Challenges and ways forward (Vol. 2020, pp. 167–190). Nova Science Publishers.

Selby, C., & Woollard, J. (2013). Computational thinking: The developing definition.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques.

Sullivan, A., & Bers, M. U. (2019). VEX Robotics Competitions: Gender differences in student attitudes and experiences. Journal of Information Technology Education: Research, 18.

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. Computers & Education, 148, 103798. https://doi.org/10.1016/j.compedu.2019.103798

Wae-Shik, M. (2006). Analysis of Error Types occurring on Elementary School Student's Programming Learning. Journal of the Korea Society of Computer and Information, 11(2), 319–327.

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35.

Yildiz Durak, H. (2020). The Effects of Using Different Tools in Programming Teaching of Secondary School Students on Engagement, Computational Thinking and Reflective Thinking Skills for Problem Solving. Technology, Knowledge and Learning, 25(1), 179–195. https://doi.org/10.1007/s10758-018-9391-y

Zhang, L., & Nouri, J. (2019). A systematic review of learning computational thinking through Scratch in K-9. Computers & Education, 141, 103607.

Zhong, B., & Si, Q. (2021). Troubleshooting to Learn via Scaffolds: Effect on Students' Ability and Cognitive Load in a Robotics Course. Journal of Educational Computing Research, 59(1), 95–118.

基于 STEAM 的幼儿园大班实体编程课程开发：活动设计与推进路径的实证研究

Developing a STEAM-Based Tangible Programming Curriculum for Senior Kindergarten: An Empirical Study on Activity Design and Implementation Pathways.

赵丽婷，林丽云
浙江省文成县大峃镇龙川幼儿园
986430618@qq.com

【摘要】 大班幼儿实体编程启蒙课程需要基于 STEAM 教育理念，把握好大班幼儿的年龄特点，将课程目标定位于以问题解决为主要的目标导向，重过程轻结果，把握“启蒙”的度，保护幼儿的学习兴趣；课程实施的过程体现儿童中心，以问题为引领，引导幼儿在情境中体验，在游戏中探索；将编程与其他领域的内容融合其中；同时需要选择适宜的实物编程套件而非图形化编程系统作为编程学习工具。
【关键词】 实体编程启蒙课程;大班幼儿;STEAM 项目活动;推进路径

Abstract: *The kindergarten class's entity programming course needs to be based on the STEAM education concept, grasp the age characteristics of the children, and set the course to be problem-solving-oriented, focus on the process and not the result, grasp the "enlightenment" degree, and protect the children's interest in learning the process of the course implementation reflects the child-centered, with the problem as the lead, guide the children to experience in the situation, and explore in the game; integrate programming other domain content; At the same time, it is necessary to choose a suitable physical programming kit instead of a graphical programming system as a programming learning tool*
Keywords: Object-oriented programming introduction course, kindergarten children, STEAM project activities, advancement path

1.以 STEAM 教育理念为基础，理解幼儿编程启蒙教育的涵义

幼儿编程教育是一种编程启蒙教育，属于 STEAM 教育的第二种形态即智能化的 STEAM 教育。幼儿编程启蒙课程的建构，需要基于 STEAM 教育理念，既兼顾科学、技术、工程、数学，也关注和融入艺术；既传承 STEAM 教育的精髓，也创新运用智能化的教育载体；既关注 STEAM 中的各要素，更关注儿童的兴趣、经验和生活世界，关注儿童的游戏和游戏中的探索，让幼儿基于游戏化的情境，尝试解决现实生活或虚拟世界中的问题，以提升幼儿的逻辑思维、问题解决、合作共情和自我实现等学习能力与品质。（边霞.(2022).STEM 教育、STEAM 教育与幼儿编程教育.教育导刊,(05),5-10.doi:10.16215/j.cnki.cn44-1371/g4.2022.05.001.）

实体编程是一种特殊的编程形式，具有操作直接、编程简单、交互性强三个基本特征。将屏幕中的代码块实物化，封装实物积木块当中。这些实物模块被赋予不同的属性或功能，如前进、旋转、后退等。它们可以通过不同的拼接方式表达程序逻辑，并对指令做出响应。

因此，STEAM 理念下的幼儿实体编程启蒙课程指的是通过将屏幕中的代码块实物化，封装实物积木块当中，赋予不同的属性或功能，它们可以通过不同的拼接方式表达程序逻辑，并对指令做出响应的智能化 STEAM 教育实践过程。

2.以大班幼儿的年龄特点为抓手，分析实体编程启蒙课程的可行性

2.1. 幼儿生理发展的支持

2.1.1.神经的可塑性

神经可塑性是指的由于经验原因引起的大脑的结构改变。人在学习新的知识时，大脑神经元为适应环境会通过一个细胞激活其他细胞的特性来提高神经的交流能力。随着能力的加强，这些神经元群会成为一个整体共同行动。这证明了练习可以使神经细胞重组，幼儿在进行实体编程活动的过程中，神经会塑造成适合学习编程的发展状态。神经的可塑性和人的年龄有很大的关系，可塑性会随着年龄的增长不断降低，也就是说幼儿阶段拥有更多样化的潜能。因此需要重视学前教育阶段的实体编程教育，充分利用编程教学，促进幼儿大脑神经的发展。

2.1.2.幼儿发展的敏感期

儿童成长中的一些时间段内，他会只对周围环境中的某一项特征专注，从而不肯接受其他特质的事物，他还会无需特定的原因而对一些行为产生强烈的兴趣，兴高采烈地重复，直到突然爆发出来某种新的动机为止，这就是孩子的某个敏感期。在幼儿敏感期时，在其周围营造编程活动的环境，合理安排实体编程活动，抓住幼儿编程能力发展的关键期，会对幼儿产生潜移默化的影响，幼儿在年幼时将会对编程产生偏好，进而形成永久性的感受。

2.2. 实体编程为大班幼儿编程活动提供载体和内容

在当今社会，编程教育已慢慢成为教育界关注的热点之一，学前教育阶段作为幼儿成长的奠基阶段，需要重视起幼儿阶段的编程教育任务。实体编程活动可以让幼儿在不用过早接触电子屏幕的前提下，帮助幼儿学习编程知识与技能，让幼儿在探究使编程机器人按照程序行动的过程中，面对问题时可以理清思路、规划步骤，直到一步一步完成任务，形成编程思维。幼儿在锻炼编程思维的基础上提高逻辑思维能力，发展计算思维，以便未来可以更好地适应人工智能社会。

2.3. 大班幼儿的年龄特点

表 1 大班幼儿学情分析

大班幼儿学情分析	
思维方式	大班幼儿的主要思维方式是具体形象思维，幼儿虽然可以摆脱对实物与动作的依赖，但又必须借助事物的具体形象或表象进行事物感知。
心理特点	大班幼儿的兴趣比较广泛，好奇心与求知欲强烈且具一定模仿能力。但是幼儿的情绪波动比较大，自控力差。
个体差异	幼儿的性别、生活背景、知识掌握水平各不相同，逻辑思维与动手操作皆存在巨大的个体差异。

由上表可知，大班幼儿学习情况以思维方式、心理特点、个体差异三方面进行分析。即实体编程教学可以利用大班幼儿借助事物的具体形象感知学习的思维方式、好奇心与求知欲

强烈的心理特点以及性别、生活背景、知识掌握水平的个体差异为基础进行幼儿编程教学，大班幼儿的年龄特点已经可以进行幼儿实体编程的教学。

3.以“三点原则”为启发，定位大班幼儿实体编程启蒙课程目标

在幼儿园阶段开展编程教学活动要基于幼儿具备的学习能力和他们的身心发展特点来开展。因此，大班幼儿编程启蒙课程绝不是给幼儿上编程课，教授抽象的编程技能，而应该以游戏为基本活动形式，将目标重点落在幼儿的思维和逻辑教学活动上，培养幼儿问题解决和合作创新能力，激发幼儿的想象力、好奇心和兴趣，帮助幼儿学会创造性思考和做事。幼儿编程的核心思想是实现儿童的想法，让儿童给计算机编程，而不是计算机给儿童编程。任何编程教学活动模式都需要让幼儿感兴趣，并且实现跨学科融合。因此，在大班幼儿编程启蒙课程目标的把握中，应遵循三点原则：

3.1. 以问题解决为主要的目标导向

游戏任务应与幼儿生活经验相关联，为幼儿创设有意义的问题解决情境，在情境的创设中自然地激发幼儿问题解决的动机，从而进行更主动的探索。如在“龙川小导游”这一游戏活动中，启发幼儿思考如何安排玛塔机器人的旅游的路线才能如约地和小伙伴们一起参加龙川旅游节活动。在此过程中，幼儿能够非常容易地理解问题情境，进而思考如何使用指令积木操控玛塔机器人运行以解决这个问题。

图 1 STEAM 项目活动“我是龙川小导游”

3.2. 重过程轻结果

重视问题解决的过程，哪怕最后并没有完全解决问题，但是只要幼儿进行了积极地思考，也可以通过引导幼儿总结失败原因而使之得到思维的发展。有些大班幼儿编程启蒙课程过于结果导向，关注幼儿能否完成既定任务，幼儿一有偏差教师就急于纠正，有时反而影响了幼儿自主探索和试错的空间。其实，计算思维的一个重要方面就是“迭代”，即没有唯一答案，只有更优的策略。这也是幼儿编程启蒙的重要内容，即让幼儿明白问题解决是一个持续的过程，即使问题已经有了一个解决办法，也可以继续思考是不是能改进方案，找到更优解。

在《我是龙川小导游》的项目活动中，孩子们从小小导游初成长、小导游经验分享会、小导游上岗实习、设计龙川旅游景点图到最后龙川旅社的开张和小导游顺利地利用玛塔编程向他人介绍家乡龙床，他们遇到了各种各样的问题。比如在练习成为导游地阶段中，孩子们通过一次又一次地自愿上台试讲，不断地尝试选出更好地试讲方式。孩子们在问题地迭代中不断地去思考，去解决问题，这个过程是更重要的。

图 2 STEAM 项目活动“我是龙川小导游”

3.3. 把握“启蒙”的度，保护幼儿的学习兴趣

幼儿园阶段的幼儿因其年龄特点的影响，专注力、合作力、逻辑思维等能力的发展水平存在较大差异，因缺乏前期经验，对编程课程的学习内容也可能较为陌生。在引导幼儿进行编程教学活动时，应及时调整活动目标，保持幼儿的游戏兴趣。如在“龙川小导游”这一游戏活动中，如幼儿感觉玛塔机器人的路线太难，无法判断机器人的方位，教师可以降低难度，减少机器人需要的旅游地点。

编程游戏往往会对幼儿的逻辑思维能力提出挑战，如幼儿因发展水平较低而遇到困难，要从保护幼儿探究兴趣的角度做出引导，给幼儿更多时间，避免拔苗助长。归根结底，在幼儿园阶段开展编程启蒙活动，重在“启蒙”，即用有新鲜感的编程类游戏，结合幼儿已有经验，引发幼儿学习和探究的兴趣，体验逻辑思维的思考方法，为今后的学习打下基础。
在编辑玛塔的旅游路线时，有孩子提出了不同的编程方式，面对这个问题，教师没有马上给与他们答案，而是抽出更多的时间陪着这两个孩子去探索，验证编程方法的可行性，通过这种自主的探索方式，孩子们在接下去的活动中开始尝试用多种编程方式去解决问题，打破了孩子们思维的局限性。

4.以“四条路径”为引领，追踪大班幼儿实体编程启蒙课程实施过程

让我们回到大班幼儿编程活动中，关注大班幼儿在编程过程中遇到的问题，并通过互相讨论、补充和启发，以问题为导向，创设编程情景，选择合适的编程工具，结合多种领域丰富编程内容，进而解决幼儿编程过程中的问题，具体参考以下四条路径：

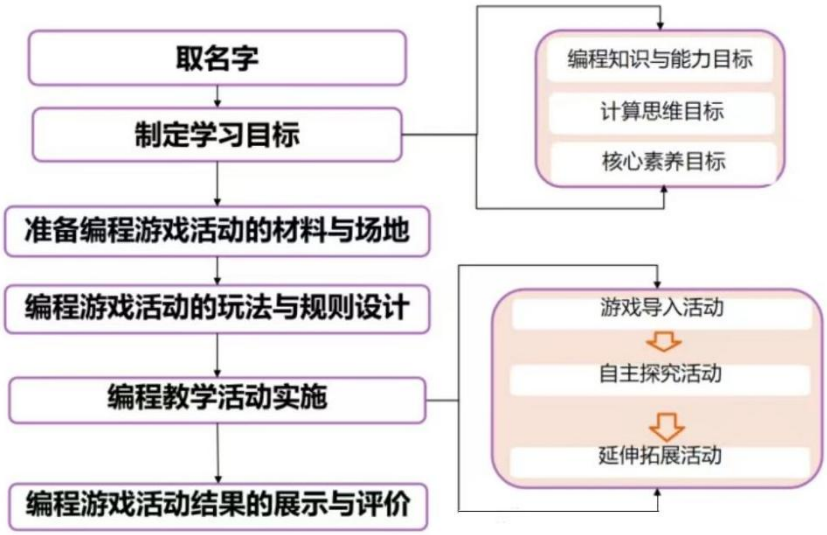


图 3 大班幼儿实体编程启蒙课程实施流程图

4.1. 问题驱动，引发编程思考

大班幼儿编程启蒙课程的实施过程应以问题解决为导向，以学习者为中心，注重幼儿在游戏化的学习情境中的自主学习和主动探究。如在“龙川小导游”这一游戏活动中，启发幼儿思考如何安排玛塔机器人的旅游的路线才能如约地和小伙伴们一起参加龙川旅游节活动。在此过程中，幼儿能够非常容易地理解问题情境，进而思考如何使用指令积木操控玛塔机器人运行以解决这个问题。

4.2. 情景创设，丰富编程内容

活动过程以幼儿在情境中的体验、观察与思考，在游戏或游戏状态下的探索、动手、验证、交流为主，教师提供适当的指导和支持，帮助幼儿通过自主探究和建构，达成各自在已有基础上的发展和进步。

作为一种对大班幼儿进行的编程启蒙教育，课程内容首先要注重与儿童已有经验、与现实生活场景的联系。如在“龙川小导游”这一编程启蒙活动的情境是因为幼儿园常常迎来从远地方来的客人老师，她们来到我们的家乡文成龙川，总是要问问龙川的风土人情，想知道孩子们的家乡有什么好玩的。孩子们总是非常热情地要给客人老师当小导游，向她们介绍龙川的文化和特色，但是如何当好家乡小导游呢，围绕着这个问题孩子们结合玛塔机器人编程技术以及对龙川的了解，为远方来的客人当导游。

4.3. 工具选择，立足编程条件

幼儿的编程启蒙需要合适的智能学习工具，相比于学龄阶段的儿童学习编程更多通过电脑屏幕，幼儿阶段的编程启蒙则更适宜实物学习，即将实体的物品作为幼儿直接的编程对象，通过实物指令而非电脑指令，控制物体执行操作。实物编程更为直观，更符合幼儿园阶段幼儿的学习特点，同时也对幼儿的视力发展更有好处。幼儿利用实物编程工具进行解决问题的游戏活动，能很好地促进其逻辑思维的发展，并为今后计算思维能力的发展奠定基础。如在“龙川小导游”这一游戏活动中，幼儿选用玛塔机器人作为编程工具，通过制作 3D 打印模型，创设了游玩龙川的游戏情景。

图 4 玛塔编程和 3D 打印工具

4.4. 领域结合，链接编程理念

作为 STEAM 教育的一种样态，大班幼儿编程教育还要体现学科或领域整合的特点，一般可采取将编程与数学、科学、艺术等多个学科相结合的方式组织课程，引导幼儿在探索性的编程活动中，学习和复习相关学科或领域的内容，解决相关学科或领域的问题。如在“龙川小导游”这一游戏活动中，孩子们在放假期间和爸爸妈妈一起参观龙川的景点，并集中梳理出的龙川的各个有名的景点，自主绘制出龙川的景点图。通过绘画龙川地图，为玛塔机器人创设游戏情境，将编程与艺术领域相结合。

图 5 幼儿自主绘制出龙川的景点

总之，适宜的早期儿童编程启蒙课程既是将编程和儿童的游戏相结合，让孩子们在游戏中探索有关编程教育的思维，有关 STEAM 的教育理念，让孩子在游戏中学习，在学习中游戏。

参考文献

边霞.(2022).STEM 教育、STEAM 教育与幼儿编程教育.教育导刊,(05),5-10.doi:10.16215/j.cnki.cn44-1371/g4.2022.05.001.

陈洁 & 陈丽.(2022).实物编程融合 STEAM 课例对儿童进行计算思维启蒙的探索.教育观察,11(03),60-63+74.doi:10.16070/j.cnki.cn45-1388/g4s.2022.03.014.

梁瑞萍.(2021).基于主题情景创设的儿童编程启蒙课程教学设计研究(硕士学位论文,海南师范大学).硕士 https://link.cnki.net/doi/10.27719/d.cnki.ghnsf.2021.000153doi:10.27719/d.cnki.ghnsf.2021.000153.

傅骞 & 章梦瑶.(2018).实体编程的教育应用与启示.现代教育技术,28(12),108-114.

王姣姣.(2019).儿童编程启蒙教育中的可视化设计应用探究(硕士学位论文,北京服装学院).硕 https://kns.cnki.net/kcms2/article/abstract?v=Jy1bRTva-nWPNqARnfr73b97RjK4sFC0hQYfGFA__wbv38vvyv0rVY7KjedLU_LWqKK6B3gq3Qkbgbg0WhtwMQ1ZzqQKJ7-WboGeQPhPj4UnhANakhao9jKvpoSKEkZukuySmLBegw50tBTCphG1SPWKiyv2Kuhb4uRQaciOrBxMxRBJyYz0SsFsbeJ_6z2QSKbsRIEFTDo=&uniplatform=NZKPT&language=CHS

小学信息科技体验式学习活动设计与实践——以《算法的特征》一课为例

Design and Practice of Experiential Learning Activities in Primary School Information Technology: A Case Study of the Lesson "Characteristics of Algorithms"

谢淑虹 1*，应浩 2，王寅玮 3
1 宁海县潘天寿小学
2 宁海县桥头胡中心小学
3 宁海县银河小学
*402194431@qq.com

【摘要】 体验式学习活动通过“亲历体验——交流分享——观察反思——总结提升——实践应用”螺旋上升的过程来促进学生综合素质和核心素养的生成。笔者以浙教版（2023）小学信息科技《算法的特征》这课为例进行活动设计与实践，以期为体验式学习研究提供理论和实践借鉴。
【关键词】 体验式学习;信息科技;算法教学;计算思维;学习活动

Abstract : *Experiential learning activities promote the development of students' comprehensive qualities and core competencies through a spiral ascending process of "personal experience - communication and sharing - observation and reflection - summary and improvement - practical application". Taking the lesson "Characteristics of Algorithms" from the Zhejiang Education Edition (2023) of Primary School Information Technology as an example, this paper designs and implements activities, aiming to provide theoretical and practical references for the research on experiential learning.*
Keywords : Experiential learning; Information technology; Algorithm teaching; Computational thinking; Learning activities

新课标明确指出，要“通过真实问题情境，引导学生在数字化学习过程中发展计算思维”。然而当前小学算法教学中普遍存在"重知识灌输轻体验建构"的困境，学生难以形成对算法本质的深度理解。基于此，本研究以体验式学习理论为指导，依托浙教版（2023）五年级《算法的特征》教学内容，通过构建“五阶螺旋”体验学习模型，设计序列化实践活动，探索在具身认知中培育计算思维的有效路径。

1.理据互契：体验学习与算法教学的融合根基

1.1 体验式学习理论

体验式学习理论是由库伯提出的，该理论认为，学生从他人建构的学习经验场中亲身体会，进而参与到学习交流与讨论中，发表看法，接着反思体验过程中感受到的知识，并总结形成内化的知识，最后将知识应用于实践的一个螺旋式循环过程。在小学信息科技课堂上实施的主要包括以下几个环节，如图 1 所示。

图 1 小学信息科技体验式学习活动的五个环节

体验式学习是一个以体验为基础的持续学习过程，与传统的学习方式不同。传统的学习方式如学习“银行”，教师单纯地将知识传授存放到学生这个仓库中。而体验式学习过程中，学生不是单纯地接受、容纳和记忆知识，每个人或多或少都带着一定的学习经验进入学习，培养学生质疑、思考、创新和实践的能力。

1.2 算法教学特征

《义务教育信息科技课程标准（2022 版）》明确提出“通过实例归纳算法特征”的素养目标，要求教学超越编程语法的表层训练，转向对算法本质属性的深度理解。然而，小学生正处于皮亚杰认知发展理论中的具体运算阶段，其思维依赖于实物操作与可视化表征，这与算法概念的抽象性形成显著矛盾。实践中，传统教学常陷入两极化困境：或过度依赖数学案例导致概念窄化，如仅用“鸡兔同笼”讲解算法，或机械模仿编程范例造成思维固化。因此，教学设计需在学科逻辑（算法特征的系统性）与认知规律（儿童思维的形象性）之间搭建桥梁。例如通过“游戏体验官”等情境，使学生在具身体验中自然提炼算法的普适性特征。

2.链构逻辑：“五阶体验链”的模型建构

笔者基于理论研究并结合实际教学经验，以学生为主体，提出了“五阶体验链”学习活动设计流程，如图 2 所示。

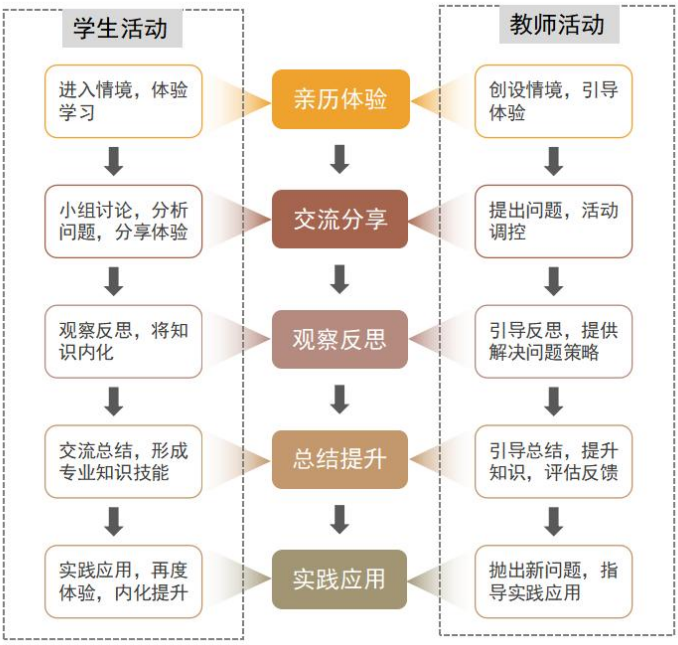


图 2 “五阶体验链”学习活动设计流程

2.1. 亲历体验

在小学信息科技课程中开展体验式学习活动，需要结合学生实际设计项目主题，教师需要创设贴近学生生活的真实情境，引导学生体验学习。在此环节教师作为一个引导者，要给予学生温暖和肯定的鼓励，营造一个可让学生亲自体验的技术环境。引导学生在体验式学习

活动中进入设计好的学习情境中去，在基于自身已有的学习经验的基础上参与体验学习，提高学习的主动性和积极性。

2.2.交流分享

基于上一步学生在体验的过程中发现问题，学生参与小组讨论探究，明确学习任务并进行分工学习，对问题进行分析，与同伴交流分享。教师作为促进者的角色，引导学生参与合作学习，并对活动的时间、交流方向等进行调控，帮助学生在合作学习中组织分享经验，引导学生在交流分享中反思学习。

2.3.观察反思

反思是体验式学习中将知识内化的一个重要环节，在经历了体验和与同伴交流分享经验后，学生需要在此环节中通过观察和反思将知识进行吸收和内化。在此环节中教师需要引导学生在活动中参与观察和进行学习反思，并提供问题解决的策略，帮助学生在反思和分析，引导学生思考，组织帮助学生将自己的思考与信息科技学科知识联系起来。

2.4.总结提升

活动任务完成之后，学生在老师的引导下尝试将解决问题的方案转化为专业知识技能，并将知识进行内化提升。在此过程中教师需要组织学生进行交流，并总结学生交流内容，进而提升知识应用。此时教师作为一个评估者的角色，引导和帮助学生在总结汇报中掌握信息科技知识和技能，并对其掌握的学习目标进行评估。

2.5.实践应用

实践应用是体验学习中将内化知识用于指导实践的过程，学生在此环节中学会运用理论知识进行实践活动，再次进行实践体验来内化知识。教师在此环节中再抛出类似问题，引导学生连接之前的体验过程中总结的知识经验去解决新的问题，再指导学生将该知识应用于更大范围的实际问题。

3.课例循证：体验式学习的实践验证路径

3.1 教材定位：算法思维的启蒙基点

本课是浙江省小学信息技术五年级上册第一单元《身边的算法》第5课，属于六大课程逻辑主线中的算法范畴，算法是计算思维的核心要素之一。课的教学内容是了解算法的三个特征：有穷性、确定性、有输出，并且能够根据算法的特征准确描述某一种行为的算法。通过本课学习，进一步提升学生对算法的理解，提高学生准确描述相关算法的能力，帮助学生进一步理解算法是通过明确、可执行的操作步骤描述的问题求解方案。本课承担着破除“算法=编程”迷思的关键作用，需通过具身体验帮助学生建立“算法是解决特定问题的步骤描述”这一本质认知，为后续学习条件判断、循环结构等抽象概念奠定基础。

3.2 学情分析：生活经验的算法化提取

本课学习对象为五年级的学生，算法场景从符合他们认知特点的具体问题出发，避免在问题的理解上造成障碍；通过之前的学习，学生具备使用自然语言和流程图描述算法的能力，前4课学习了身边的算法、自然语言描述算法、流程图描述算法、算法中的数据，为本课的学习奠定了基础。同时，他们善于合作，乐于表达，初步具备问题分析能力，但由于算法的特征较为抽象，学生理解起来会有一定难度。

3.3 核心目标：算法特征的具象化建构

①通过生活中的实例分析，知道算法的特征：算法的有穷性、算法的确定性和算法要有输出。

②通过不同算法的对比，领会算法的三大特征，理解算法是通过明确的、可执行的操作步骤的问题求解方案。

③通过完善水果茶的练习，理解算法的特征，有意识地将其应用于数字化学习过程中，体会到算法对解决生活和学习中的问题的重要性，形成用计算思维以及计算机程序解决实际问题的能力。

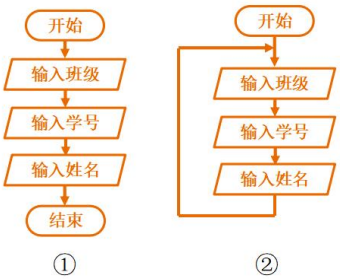
3.4 活动创生：五阶螺旋的体验实践

本案例基于体验式学习活动流程，设计了三个探索活动实例让学生参与体验，在体验过程中通过问题驱动、实例分析理解算法的特征，学会运用算法的特征检验算法的正确性。整体活动流程如图 3 所示：

图 3《算法的特征》一课体验式活动设计

实践案例围绕着体验“答题小程序”这一主线情境展开，旨在通过体验式学习活动引导学生学生对算法的理解，提高学生准确描述相关算法的能力，帮助学生进一步理解算法是通过明确、可执行的操作步骤描述的问题求解方案。在三个学习活动中学生依次经历了“亲历体验—交流分享—观察反思—总结升华—实践应用”的体验学习过程，设定贴近学生实际生活的学习主题更有利于激发学生已有的学习经验，引起学生的学习兴趣，在学习活动中学生在交流与讨论中提出问题、发表看法，在反思体验过程中慢慢感受到算法的三个特征，并总结形成内化的知识，最后将算法的特征应用于生活中其他算法上，形成一个螺旋上升的过程。具体体验活动设计如表 1、表 2、表 3 所示：

表 1 体验活动一：算法的特征——有穷性

	活动详情
亲历体验	抛出情境：六年级的学生设计了一款“答题小程序”，今天我们做为首批体验者，对这款小程序进行体验，提供宝贵意见。 尝试体验：学生打开桌面上的“答题小程序”进行体验。
交流分享	小组讨论：通过尝试，你认为“答题小程序”的流程图算法是右边哪一个？ 
观察反思	尝试反思：你觉得这个“答题小程序”能解决什么问题吗？如果能，请写出解决的问题，如果不能，请写出原因。

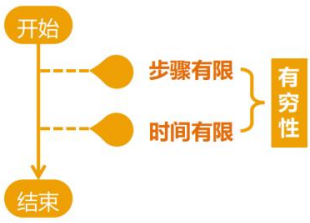
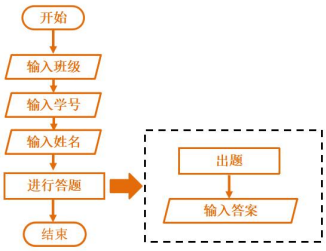
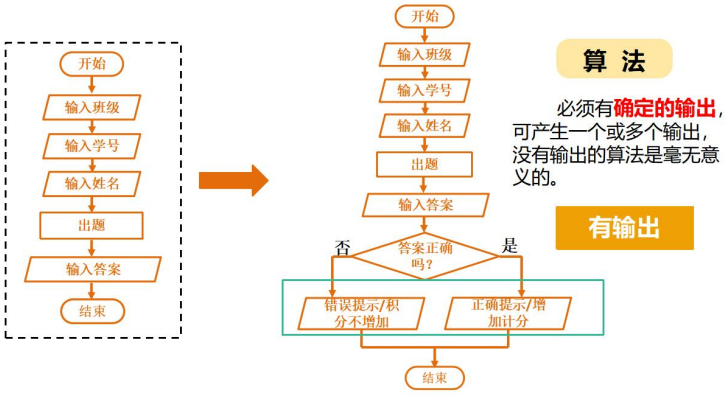
总结提升	回顾总结：算法是解决问题的方法及其步骤，如果解决问题的方法及步骤是无穷无尽的，那么这个算法就不能解决问题，即算法必须能在有限的时间内执行有限步骤之后终止。总结出算法的第一个特征——有穷性。 
实践应用	生活实践：比如煮水果茶的算法，若一直重复“洗水果”“切片”这个步骤，那么永远都煮不出茶来。

表 2 体验活动二：算法的特征——确定性

	活动详情
亲历体验	延续情境：为了让答题小程序能够答题，并能结束，体现算法的有穷性，请尝试修改流程图。 尝试体验：修改流程图。
交流分享	小组交流分享修改后的流程图，提出关键步骤：答题。
观察反思	提出问题：“答题”这一步，计算机知道怎么操作吗？“答题”的操作步骤具体应该是怎样的？ 教师引导：答题有很多种方式，比如闯关答题，pk 答题，积分答题等等，但不管哪种答题，首先系统得先出题，用户才能进行作答，对吧？ 
总结提升	小结：算法除了有穷性，算法中的每一步必须是确定且能有效执行的，而不能是模糊不清的。总结出算法的第二个特征：确定性。

实践应用	任务实践：根据算法的确定性特征，请你找出煮水果茶的算法中不合理的地方，并用方框圈出来。 步骤1：把水果洗净切片。 步骤2：放水果片若干。 步骤3：加入适量配料。 步骤4：加水。 步骤5：将水温煮到大概100摄氏度。 步骤6：再煮一会儿。

表 3 体验活动三：算法的特征——有输出

	活动详情
亲历体验	延续情境：六年级的学生接受了大家的意见进行了改进，请打开桌面上的“答题小程序 2”，再来看看这个程序的算法是否已经完善，达到你的预期？ 尝试体验：学生打开“答题小程序 2”进行体验。
交流分享	小组反馈交流：经过你的尝试，你有没有发现什么问题呢？（完成了答题之后没有任何反馈）
观察反思	思考：你希望得到怎样的反馈？如果是你来设计程序你会在反馈页面上怎么表达？（想知道这题是否正确，觉得反馈页面上可以显示对或错的文字、符号、还可以积分...）
总结提升	小结 1：算法必须有确定的输出，可产生一个或多个输出，没有输出的算法是毫无意义的。总结出算法的第三个特征——有输出。 小结 2：输出可以是数值、文字、图、等各种各样的数据。 
实践应用	联系生活应用：算法必须有输出，比如通过预约排号算法输出排队序号，通过导航算法输出多条备选路线，通过购买车票算法输出订单结果。 预约排号算法 输出 排队序号 导航算法 输出 多条备选路线 购买车票算法 输出 订单结果

4.体验进阶：算法教学的迭代反思与素养转向

通过《算法的特征》一课的体验式教学实践，笔者深刻认识到，从“知识传递”到“素养进阶”的转型需要突破传统教学范式，在实践反思中重构教学逻辑。本部分从实践成效、现存问题、改进路径三个维度展开，探寻契合儿童认知规律与时代需求的算法教学新样态。

4.1 具身体验激活计算思维的底层建构

通过“答题小程序”、“煮水果茶”等生活化情境的五阶体验，学生得以在操作中具象化理解算法的抽象特征。在答题流程无限循环的冲突体验中，学生通过对比流程图的有限与无限步骤，直观感知到“算法必须在有限步骤内终止的本质”；修改流程图时，学生发现有些模糊的表述无法被计算机执行，从而领悟“算法步骤需精确到机器可识别的操作”，将自然语言转化为结构化思维；在无反馈答题程序的缺陷分析中，学生主动设计多样化输出形式（文字、符号、积分），理解算法的价值在于解决问题并产生有效结果。这种“做中学”的方式，使计算思维的“抽象化、模型化、自动化”特征通过身体操作与认知加工深度融合，实现从“知识记忆”到“思维建构”的跨越。

4.2 个体差异与线性流程的适配困境

在体验式学习活动中，预设的“五阶体验链”线性流程（亲历体验➡交流分享➡观察反思➡总结提升➡实践应用）虽能保障教学逻辑的连贯性，却难以充分适配学生的差异化认知节奏。由于五年级学生在认知水平、技术操作能力上存在显著差异，在具体实践中也发现个别学生积极性不高导致没有完全跟上体验式学习活动的线性流程，学生在体验的过程中会存在过度依赖教师的现象；小组合作中，少数操作熟练的学生主导讨论，部分低起点学生沦为“旁观者”，影响了活动体验的积极性，未能实现“差异发展”的素养目标。

4.3 从固定五阶到动态适配的迭代优化

“五阶螺旋”模型虽构建了体验学习的基本框架，但其线性流程难以完全适配学生差异化的认知节奏与即时反馈。基于课例实践反思，本研究提出动态适配优化策略，通过弹性环节设计、螺旋迭代机制、差异化路径支持，使体验过程从“预设路径”转向“按需调节”。

4.3.1 弹性化环节嵌入

打破固定阶段划分，在交流分享、观察反思等关键节点增设动态调节接口。例如，当学生在答题小程序流程图修改中普遍对“确定性”特征理解模糊时，即时插入“二次具身体验”，例如让学生用实体卡片模拟“模糊步骤”与“精确步骤”的执行差异，通过对比操作强化对步骤可执行性的具象认知，避免因理解断层导致的素养发展停滞。

4.3.2 螺旋式迭代升级

将总结提升与实践应用重构为迭代优化循环，鼓励学生在解决新问题时反哺对原有算法特征的理解。如在“煮水果茶算法”实践中，学生发现放水果片若干 的模糊表述会导致结果不可控，这一反思被迁移到答题小程序的输出设计中，促使他们在“实践应用”阶段主动为反馈环节增加精确积分规则，形成“体验➡反思→修正→再体验”的深度循环，使算法特征的理解随问题复杂度提升而动态深化。

4.3.3 差异化路径支持

针对学生在活动中呈现的能力分层，可提供基础版与进阶版的双轨体验路径。例如，低起点学生可通过预设好的简单判断问题完成反思，如“步骤是否有限？是否每一步都明确？是否有输出结果？”，而能力较强的学生则可自主设计“算法改进方案”，还可设计开放任务，如挑战同时优化算法的有穷性与输出形式。这种动态适配机制，使五阶模型从“标准化流程”转化为“个性化脚手架”，确保不同认知水平的学生均能在最近发展区实现素养进阶。

从《算法的特征》一课的实践可知，体验式学习的价值不仅在于知识传授，更在于为学生打开一扇通往计算思维的大门。当算法教学真正扎根于儿童的生活经验，融入螺旋上升的深度体验，知识将转化为素养，技能将升华为能力。未来，需继续以新课标为指引，在理论与实践的双向互动中，构建更具包容性、生长性的信息科技课堂，让每个学生都能在数字时代自信地“用算法思维思考，用计算能力创造”。

参考文献

王国庆.(2023). 基于体验式学习理论，培养学生信息科技学科核心素养的教学实践.小学教学研究, 04, 41–43.

闫海红.(2023).新课标背景下小学信息科技项目式学习活动设计与实践.中国信息技术教育,13, 42–45.

游戏课程化视域下幼儿编程教师专业发展的实践探索

A Practical Exploration of Early Childhood Programming Teachers' Professional Development in the Perspective of Game Curricularization

张曦
温州大学附属实验幼儿园
390801850@qq.com

【摘要】 本文基于游戏课程化视角，探讨幼儿编程教育中教师专业发展的实践路径。针对当前幼儿园教师普遍存在的编程教育理念滞后、技术操作不熟及课程游戏性设计能力薄弱等问题，构建了“理念-技术-设计”三位一体的教师能力发展模型。通过分层晋级研训机制、名师示范课例研讨及专家主导的项目研究，系统提升教师在游戏化编程中的教学设计和技术应用能力。

【关键词】 幼儿编程教育；游戏课程化；教师专业发展

Abstract: *This study explores the practical path of teachers' professional development in early childhood programming education based on the perspective of game curricularization. Aiming at the problems of lagging behind in the concept of programming education, unfamiliarity with technical operation, and weak ability of curriculum game design, which are common among kindergarten teachers, we have constructed a three-in-one model of teachers' competence development, namely, "Concept-Technology-Design". Teachers' instructional design and technology application skills in gamified programming are systematically enhanced through a tiered progression research and training mechanism, master teacher demonstration lesson example seminars and expert-led project research.*

Keywords: STEAM, Project-based learning, Creativity, Science and Technology Innovation Labor

1.前言

在学前教育全科教师培养体系下，职前的教育阶段普遍存在通识性培养倾向，这直接导致了幼儿园教师的编程教育理念的迭代滞后。笔者的园所前期调研中发现，多数教师对图形化编程的认知尚处于工具论层面，将其机械解构为符号识别训练，而未能有效渗透游戏化情境中的隐性计算思维发展机制。同时，技术焦虑现象普遍存在于教师群体，对如 Scratch Jr 和可编程机器人等教具的操作熟练度明显不足。另外，教师普遍缺乏将抽象编程指令转化为符合幼儿认知水平的游戏化任务设计能力。基于幼儿园"游戏课程化"教育改革背景，本研究通过“理念-技术-设计”三位一体的幼师编程教育能力模型，推动编程教育从“技术训练”向“思维赋能”转型，提升教师的专业素养水平。

2.重构逻辑理念：游戏课程化与编程教育的理论耦合

在“游戏为基本活动”的学前教育改革背景下，幼儿园编程教育与游戏课程化的深度融合成为培养幼儿计算思维与创新能力的重要路径。游戏课程化理论（由王振宇教授团队提出）强调以幼儿自主游戏为起点生成课程，编程教育需从“预设指令训练”转向“游戏问题驱

动”，例如通过“拯救博物馆文物”“超级玛丽”“我上小学啦”等情境自然融入条件判断、循环等逻辑。这个过程要求教师具备“观察、分析、支持”的能力和素养，为幼儿在游戏化编程活动中的分解（任务拆解）→模式识别（规律发现）→抽象（算法设计）→评估（调试优化）等探索行为提供游戏性“脚手架”。教师们需要成为引导者和协助者，帮助孩子们在游戏的过程中自然而然地掌握编程的基本概念和技能。

在"游戏为基本活动"的学前教育改革浪潮中，幼儿园编程教育与游戏课程化的有机融合正开辟着培养幼儿计算思维与创新能力的创新实践路径。基于王振宇教授团队建构的游戏课程化理论框架，该范式主张以幼儿自发生成的游戏体验为课程生长点，将编程教育从机械的预设指令训练转向生动的游戏问题驱动。如在"拯救博物馆文物""超级玛丽闯关""我上小学啦"等主题情境中，巧妙融入编程思维的核心要素——条件判断、循环控制等逻辑概念。这种活动指导的转型要求教师以"观察-分析-支持"的专业素养，为幼儿在游戏化编程过程中展开的认知探索（任务拆解→规律辨识→算法构建→优化迭代）搭建游戏化认知支架。所以，教师角色应转变为幼儿编程教育活动中的引导者与协助者，让儿童在沉浸式游戏体验中水到渠成地习得编程基础能力与创新思维范式。

3. 创新实践路径：游戏课程化编程教育的教师研修

游戏课程化编程教育中，教师专业能力提升研训行动构建了三维赋能体系：其一实施领雁引领的分层递进式技术研训，其二开展名师领衔的示范课例深度研讨，其三推进专家主导的游戏项目创新攻关。依托梯度段位进阶培训体系与差异化任务驱动机制，系统提升教师的游戏课程化编程教育的专业理论素养和技术实操能力。通过专家的指导和项目合作，提高教师在编程教育中的游戏化设计研究能力。

3.1. 领雁+分层技术研训

3.1.1. 分层晋级机制

基于教师个体在兴趣取向、专业素养及学术背景上的多元差异，我们构建起精准匹配的"技术段位"评价体系，配套实施阶梯式成长方案。首先遴选优势型教师进行个性化培养方案锻造，打造具有辐射力的示范标杆。继而通过园本化分层晋级培训机制，引导教师在完成阶梯式成长任务的过程中，实现知行合一的能力晋级，由此形成动态化、内生性的分层研修生态。这种"晋级闯关"式培养模式持续激发教师突破职业舒适区的勇气，在自主进阶的良性循环中培育出强劲的内生成长动力。

表 1 幼儿园编程教师阶梯式晋级设计

阶段	能力特色	培训重点	晋级任务	任务类型	提交成果
初阶	将编程教育与幼儿游戏活动相结合	编程教育基本概念和工具；融入幼儿游戏方法；从幼儿兴趣出发设计游戏	设计一个简单的编程游戏，并提交操作视频；开展一次编程游戏活动，并提交活动总结报告	实践操作考核；活动总结报告	操作视频；活动总结报告
进阶	设计基于情境认知和动机主体性体验的游戏化编程课程	课程设计方法；编程工具多样化应用；提升幼儿编程思维	设计一个为期一周的游戏化编程课程，强调情境认知和动机体验；试讲设计的课程，并提交课堂观察记录	课程设计考核；实践反思	游戏化编程课程设计；课堂观察记录

高阶	具备本土化创新与研究能力	编程教育本土化创新；课题研究方法；推动编程教育整体发展	完成一个结合本土文化的游戏化编程课题研究报告；将设计的课程应用于实际教学，并提交应用效果报告	课题研究成果考核；课程推广	课题研究报告；应用效果报告

3.1.2. 任务分段设计

我们采用了任务驱动的方法，每个月都会发布新的晋级任务，教师们需要提交视频记录并参与互评。通过园内晋级考核，教师们可以获得“游戏课程化编程导师”的徽章，并将此作为绩效考核的加分项。这样的机制不仅激发了教师们的积极性，也促进了他们之间的交流与合作。

我们创新性引入任务驱动型成长模式，以月度为单位滚动发布进阶挑战任务，要求教师完整录制并提交实践视频，同时开展多维度同行互评。通过考核者将获得植入在“云慧玩”平台的"游戏课程化编程导师"徽章，并纳入个人绩效考评体系。

表 2 幼儿园编程教师分段任务设计

级别	考核内容	指向说明	任务范例	提交方式	预期成果
初阶示例 1	结合故事讲解编程概念	玛塔工具使用；幼儿故事与编程结合；激发幼儿编程兴趣	设计一款基于介绍编程概念的小故事，录制讲解视频	视频录制	故事文稿
初阶示例 2	利用游戏培养幼儿计算逻辑思维	逻辑思维训练游戏设计；玛塔逻辑模块应用	设计一个逻辑思维训练编程游戏，提交游戏演示视频；进行游戏，提交幼儿活动报告	游戏演示；幼儿活动观察报告	演示视频；观察报告

3.2. 名师+示范课例研讨

依托名师示范课的教学范式，进阶教师团队对计算思维的核心内涵展开多维解析，通过深度解构其构成要素，系统性重构了跨学科融合的游戏化编程教育课程体系。基于活动现场“翻转观察”的机制，教师团队以儿童认知发展规律为切入点，运用迭代式课例研发模式对活动方案进行动态优化，最终实现优质教育资源的体系化整合。这种创新实践不仅实现了教学品质的螺旋式上升，更深化了教师群体对游戏课程化编程教育的认知。

表 3 幼儿园编程活动课型设计

阶段	执教人	研讨预备	研讨重点	问题预设范例	游戏规则迭代
原型课	市级名师	《超级玛丽》示范课	打击乐音乐结构和游戏性设计	利用不同乐器特质设计音乐游戏规则	低音鼓对应跳跃动作，沙锤对应收集金币动作
重构课	进阶教师	重新设计教案	结合计算思维要素，融入循环等模块	融入循环模块	设计“重复巡逻”或“循环收集金币”的游戏规则
迭代课	高阶教师	收集幼儿试玩反馈	通过幼儿试玩反馈调整	调整游戏难度，增加“放大”“缩	增加音乐游戏中乐曲的“强弱变

3.3. 专家+游戏项目研究

为提升教师游戏课程化编程教育的研究能力，我们跟高校专家团队互动，通过个案研讨、过程性评估等方式，对幼儿计算思维能力的培养进行专业把脉。前期助推教师在编程教育中的游戏化设计实现从经验型向研究型转变。特别在高阶教师培养层面，通过与温州大学STEM研究中心建立协同创新机制，联合研究生团队开展实证研究。案例"小船启航"，通过系统规划跨学科游戏场景，落地执行幼儿计算思维培养策略，最终形成具有教师专业发展和幼儿计算思维提升的实践经验。

表 4 幼儿园项目化编程活动中的教师素养分析

幼儿计算思维要素	项目中的幼儿具体表现	教师编程教育专业能力	项目中教师的具体表现
抽象	幼儿通过观察和讨论，将“帮助保安爷爷打捞落叶”的问题抽象为设计一艘适合打捞落叶的小船。	课程设计能力	教师设计了从问题提出到实践操作的完整课程，包括抽象问题、分解任务、自动化实现和调试优化的环节。
分解	幼儿通过实验和探索，将问题分解为“什么样的材料适合做船”“如何让船动起来”“如何捕捞落叶”等子问题。	问题引导能力	教师通过提问（如“什么方法最省力？”）引导幼儿思考问题，并将问题抽象为设计小船的任务。
自动化	幼儿尝试为小船添加动力装置，并通过小组合作设计遥控装置，将解决方案转化为可操作的步骤。	技术支持能力	教师支持幼儿通过实验和探索（如材料测试、动力装置设计），并提供必要的工具和资源。
调试	幼儿在实验中不断调整材料和设计，例如发现不同的材料吸水性不同，最终选择中空塑料瓶作为船体材料。	反思与调整能力	教师在课程中不断反思幼儿的学习过程，调整活动设计（如从材料的试验到最终的选择决策）。

模式	幼儿通过观察和实验，总结出适合做船的材料特性（如木板、塑料瓶），并探索捕捞工具的设计模式。	评估与反馈能力	教师通过观察和记录幼儿的活动表现，及时给予反馈，并帮助幼儿总结经验（如材料选择和设计优化）。

4. 反思与展望

在游戏课程化编程教育教师专业发展实践中，幼儿园通过构建“分层研训+项目驱动+高校协同”的立体化培养体系。开发《E路博游》《国王与公主》等多个情景式编程课程，其中《无屏编程》项目获评温州大学年度标杆课题。幼儿园顺利通过“人工智能示范校”创建，辐射带动山海协作地区十余所幼儿园实现编程教育零突破，并在长三角园长大会、中央电教馆平台分享经验。但是，在提升教师的专业发展，幼儿园还应充分利用各种资源，优化和完善教师发展评价机制，实现教师专业能力与课程建设的双重突破。

参考文献

王振宇. (2017). 游戏的界限[J]. *幼儿教育*, (Z6), 13-15+20.

王振宇. (2022). 从活教育到活游戏: 幼儿学前教师教育用书 ptop 游戏课程化论丛. *上海交通大学出版社*.

项目编程游戏：幼儿计算思维的阶梯式培养

Project Programming Game: Step-by-Step Development of Computational Thinking in Young Children

谢凌凤，徐佳豪
温州大学附属实验幼儿园
1229184521@qq.com

【摘要】本研究针对幼儿编程教育中存在的公式化教学倾向与个性化指导困境，以无屏幕编程工具为载体，构建基于项目化学习的园本化编程游戏课程。基于周以真计算思维三维框架与构造主义学习理论，提出三阶行为发展模型。研究创新性实施环境分层策略与动态师幼互动模式，以“一丈高的世界”项目化编程游戏活动案例阐述了阶梯式培养计算思维的方案，让幼儿在“模仿➡调试➡重构➡创新”的思维跃迁中，实现计算思维核心能力的层级发展与跨领域整合。

【关键词】无屏编程；计算思维；园本课程；玛塔机器人；游戏化学习

Abstract: *This study addresses the dilemma of formulaic teaching tendency and personalized instruction in early childhood programming education, and constructs a garden-based programming game curriculum based on project-based learning by using screenless programming tools as a carrier. Based on Zhou Yizhen's three-dimensional framework of computational thinking and constructivist learning theory, a three-stage behavioral development model is proposed. The research innovatively implements the environment layering strategy and dynamic teacher-child interaction mode, and illustrates the step-by-step program of cultivating computational thinking with the case of the "ten-foot-high world" project-based programming game activity, so that young children can realize the core ability of computational thinking in the thinking leap of "imitation→debugging→reconstruction→innovation" and realize the core ability of computational thinking in the layer by layer. The program allows children to realize the hierarchical development and cross-disciplinary integration of the core ability of computational thinking in the process of "imitation → debugging → reconstruction → innovation".*

Keywords: Screenless Programming,Computational Thinking, Garden-based Curriculum,Marta Robotics, Gamification Learning

1.研究背景与理论基础

1.1.现实困境分析

1.1.1.园本化实践与公式化教学的取舍

幼儿阶段的编程活动很大程度上依赖外部教材参考，而这些活动中甚至存在直接沿用小学编程教学模式，以教师“教”、幼儿被动“看、听、操作”接受主，偏离了幼儿实际身心发展特点。为了形成符合幼儿认知特点的系统化课程体系，我们需要对外部教材进行园本化改造。从而避免活动内容与领域目标脱节、脱离幼儿生活经验的问题。

1.1.2.个性化指导与集体间协作的权衡

幼儿个体差异显著，在编程游戏开展过程中每位幼儿的想法与所遇到的挑战各不一样。例如，在“一丈高的世界”项目活动中，幼儿需调整建筑布局以适应机器人行动，但教师难以兼顾不同小组的进度与创意需求。此外，大班额环境下，如何协调个别幼儿或小组的个性化指导与集体协作间的平衡成为挑战。

1.2.研究理论基础

项目化活动强调以真实问题为导向，通过跨学科整合、实践探索和团队协作完成系统性任务，其核心特点包括情境化学习目标、分阶段任务拆解和创造性表达。幼儿编程游戏活动以构造主义学习理论为核心，其中西蒙·派珀特提出的构造主义强调，知识应通过与环境互动主动建构而非被动接受，需要通过环境互动与动手实践构建知识体系（如编程操作、项目设计）；将抽象概念（如循环、条件判断）转化为趣味性任务（如路线设计、角色操控），利用幼儿对游戏的天然兴趣提升学习动机，促进认知发展。

当两者相结合时，项目化活动可以为编程游戏提供结构化目标，而编程游戏通过趣味性任务赋予项目化活动低门槛的实践载体。基于周以真教授提出的计算思维三维框架（概念-实践-视角），在幼儿解决实际问题的过程中要注重培养幼儿的“抽象思维”、“分解思维”、“调试思维”因此要做到：

1.2.1.计算思维培养为核心导向

编程游戏活动的设计从单纯掌握编程能力转向以计算思维为核心的培养，强调抽象、分解、算法设计、调试等思维过程的训练。例如通过关卡设计逐步引导幼儿理解编程指令与逻辑的关联性，而非仅关注编程结果是否成功。

1.2.2.认知特点适配为核心要求

面向学龄前幼儿开展编程教育活动需要强调符合儿童认知特点的内容与活动形式适配。例如，玛塔机器人通过实物指令块实现无屏幕的编程降低认知门槛，避免屏幕依赖，符合幼儿具体形象的认知特点。同时，在课程设计上注重分阶段梯度培养，按照幼儿的发展阶段与认知特点设计不同难度梯度的任务和活动等，结合生活实际、融入科学、艺术等跨学科内容开展活动，幼儿更易于接受与学习。

1.2.3.游戏化与领域间融合并行

编程游戏活动从单一编程任务转向情境化、叙事化设计。例如，通过模拟城市建造或故事情节让机器人住进未来城市，幼儿在其中需要综合考虑城市的环境创设与机器人的宜居性条件满足，是一个融合社会领域、数学领域、艺术领域等多领域解决复杂任务的过程。

2.编程游戏阶梯式体系

2.1.三阶行为发展模型

2.1.1.感知与探索阶段

在幼儿初步接触编程活动阶段，我们注重培养幼儿的编程基础知识，让他们在简单的任务或情景中利用简单的材料、模块之间的组合关系感知潜在的逻辑并进行探索。比如通过使用积木、玩具等物理媒介，让孩子们了解基本的构建原则和拼装方法。同时，结合一些简单的编程游戏，帮助孩子们理解程序执行的概念和顺序结构，培养他们的逻辑思维和耐心细致的学习态度。

2.1.2.逻辑与操作阶段

当幼儿对基础的编程活动产生兴趣或萌生挑战的意图，我们通过设计一些相对复杂的积木拼装活动，如拼装小车、小机器人等，让孩子们掌握一些基本的机械知识和物理原理。此

外，通过引入稍微复杂的编程游戏，让孩子们学习如何编写简单的程序，培养他们的判断能力和问题解决能力。

2.1.3.调试与创造阶段

当孩子们不满足于给定的固定任务或要求时，我们鼓励孩子们开始创设自己的情景并尝试解决一些更复杂的任务。如在“一丈高的世界”这一综合性情景中，我们鼓励幼儿观察实际生活并利用身边的材料创设他们心中的未来城市，以激发孩子们的创造力和想象力。同时，通过引入更复杂的编程游戏与更丰富的编程模块，让孩子们学习如何编写更加“完整”和“复杂”的程序，让机器人在他们的“未来城市”中完成各种任务，从而培养他们的创新思维和技能。在操作和实践中，孩子们并不都能一次性完成目标任务，因而他们需要对自己的程序与操作进行回顾与审视，通过不断修改与调试排除错误的部分，达成任务目标。

3.教学实施路径

3.1.环境创设策略

3.1.1.基础感知区

玛塔机器人及配套基础运动模块、数字模块、4*4、5*5、6*6 的场景地图；设置可触摸指令墙（幼儿通过撕贴磁块组合程序）；投放可自由拆搭的积木块材料，如乐高积木、磁力片材料、阿基米德积木、插塑玩具、可拼接组装的小车材料等。

通过环境暗示，鼓励幼儿结合各种材料的特点进行组合与游戏，或操作玛塔机器人进行简单的任务如“路径规划类”：结合幼儿真实生活场景（如从教室到洗手间路线），使用编程块组合程序控制机器人行走；“模式识别类”：利用数字模块设计“找规律”游戏（如红蓝灯交替闪烁序列）。

3.1.2.情景操作场

在基础感知区的基础上可再增加更丰富的指令模块组合，如循环模块、函数模块等；投放磁力地图（可更换底板纸片），便于幼儿自由设计变换地图的形态与背景，鼓励幼儿创设他们感兴趣的场景。

3.1.3.创新实验站

在情景操作场的基础上再增加提供玛塔机器人模块化组件如控制器、灯光、声音、传感器指令编程块等，结合具体生动的情景创设有趣的任务。

3.2.师幼互动模式

3.2.1.引导策略的阶段性调整

1.感知期

示范拆解任务：如将“送包裹”分解为“取件-转向-送达”

手把手退位法：在幼儿出现程序编写错误时，教师通过手把手指尖指点提示找到错误处，逐步过渡到语言提示。

2.逻辑期

设计合作型任务：如双人协作控制机器人搬运积木，需设计路线、分配任务协同完成；

三分钟等待原则：给予幼儿一定独立思考时间，在幼儿出现困难时延迟指导或介入，培养自主探究能力；

亮点介绍与分享：当发现幼儿的程序或任务具有特色或亮点时，多给予幼儿分享与展示的机会，幼儿进行表达的过程也是其重新组织语言思考逻辑关系的过程，即促进其逻辑水平发展，又给予其正向的情绪价值支持。

设计“半成品”任务：提供部分正确程序，要求补充缺失模块（如缺少转向指令的路径程序）

开展“程序逆向工程”：展示机器人动作，让幼儿反推可能使用的指令组合。

3.创造期

提出开放问题：如怎样设计机器人行动的最优路线？当机器人狭路相逢时怎么继续通行？

开展跨区联动：将编程游戏活动与班级自主游戏进行联动，比如将各活动所产生的作品进行保存与物理联动；故事或任务主题进行联动，如在“一丈高的世界”案例中，幼儿规划道路时利用自主游戏时间设计合适的美工作品进入城市中。

4.以“一丈高的世界”项目化编程游戏活动为例

4.1.“一丈高的世界”案例简述

4.1.1.游戏缘起：高楼渐起

班级区域改造后，建构类材料与其他种类材料互通共享，

孩子们手下出现了许多有意思的建构作品，在越来越多同伴自发加入后，逐渐形成了微型城市群。

4.1.2.城市创想：观察与规划

为了保留保存创意作品同时激发幼儿兴趣，游戏场地从班级转移到了空间更富余、材料更丰富的科学工作室。这时候孩子们也提出了提议：共建“自己的城市”。

联系生活实际，孩子们发现他们的城市充满各式建筑却缺乏道路，通过对比真实城市地图，将复杂问题分解为“建筑搭建”与“道路规划”两部分。讨论分工后，一部分小朋友利用新材料设计多样化建筑（如带内部结构的房屋），而另一部分则铺设道路网络，初步构建城市框架。

4.1.3.玛塔住进城市：问题与解决

随着城市设施与环境逐渐完善，孩子们开始畅想未来城市，“未来城市需有机器人”，于是一致同意邀请玛塔机器人加入，但发现建筑比例、道路宽度不匹配。为了适配玛塔机器人的加入，未来城市亟需迭代改造。通过反复调整建筑布局、拓宽道路，拆除重建以适应玛塔行动需求，最终实现机器人在城市中自由穿行与生活。

4.1.4.更好的城市：任务与情景

机器人加入未来城市后，有了各种丰富的生活内容，如玛塔机器人通过标志旗执行运输任务（如运送货物至不同主题建筑）；幼儿为建筑赋予功能（动物馆、车站、旅店），玛塔与建筑随机触发新故事（如火车发车时运输物资），推动游戏情节动态发展，增强情境性与合作性。

4.2.大班幼儿学情分析：

4.2.1.认知特点分析

1.有意注意增强：大班幼儿的注意力稳定性显著提高，能持续专注于多步骤任务。在编程任务中，他们需先分解整体任务（城市改造）为建筑设计与路网布局，再结合需求规划改造方案，体现了目标导向的思维特点。

2.空间认知能力发展：地图布置和机器人操控需要空间方位判断（如起点、障碍物位置），符合其空间认知发展水平。

4.2.2.计算思维发展分析

1.分解思维：在设计搭建他们的“未来城市”过程中，孩子们需要对一系列综合性的较难问题进行拆解，如不适宜玛塔机器人的建筑尺寸重新改造需要先测量再改造再验证调试等。符合大班幼儿逐步发展的“问题分解能力”。幼儿能通过语言描述或操作演示展现拆分复杂问题的逻辑。

2.抽象思维：编程过程中需忽略地图上的干扰元素（如其他正在运行的机器人、其他岔路等），专注于当前任务目标，体现了抽象思维中的“过滤无关信息”能力。这与大班幼儿“能按类别记忆事物”的特点一致。

3.调试思维：幼儿在机器人路线错误时调整编程块顺序或方向，即通过试错优化解决方案。这种能力与幼儿抽象逻辑思维萌芽密切相关。

4.2.3.社会发展特点分析

1.同伴合作与分享意识增强：大班幼儿在小组合作设计路线、分享编程故事时，表现出较强的协作意愿和表达能力。活动中“小组合作设计挑战性路线”环节，能促进其分工协商、共同决策的社会技能。

2.规则内化与责任感提升：在复杂任务的分工与合作过程中，对轮流和顺序规则的遵守，反映其从“外在规则”向“自律”过渡的心理特点。活动中“玛塔机器人的工作和生活”的情境设计，也激发了幼儿的责任感与任务意识。

综上所述，大班幼儿在玛塔编程游戏活动中可以采取“逻辑期”与“创造期”的引导策略。

4.3.目标分层设计

1.基础操作（指令认知）：在“玛塔机器人”操作中，通过具象化指令（如前进箭头对应移动、转向图标对应方向变化），帮助幼儿建立“输入-反馈”的直观映射关系。例如，在“一丈高的世界”初始阶段，幼儿通过摆放实体编程模块控制机器人移动，同步理解“距离单位”与“步数”的数学对应关系，培养数学素养。

2.逻辑构建（逻辑与路径规划）：如在未来城市中设置多障碍场景，幼儿需分析障碍物分布，将路径拆解为“条件判断”（如“遇到障碍物则右转”）和“循环逻辑”（如“重复3次直行动作”）的组合模块，培养系统性思考能力。

3.创新应用（调试与优化）：开放任务中，幼儿需自主设计如设计最短运输路线。例如，通过对比不同路径的指令数量，引导幼儿发现“循环嵌套”可减少重复代码，体现算法优化思维。

4.4.任务难度递进

1.模板模仿：教师演示“穿越障碍物”任务时，幼儿复现基础指令序列（如“前进5步+右转90°”）。通过标准化操作掌握坐标系概念，为后续复杂任务奠定空间认知基础。

2.半开放情景：在“限时运输”任务中，引入时间变量（如“30秒内到达终点”）和多目标点。幼儿需调整代码执行效率，例如将分散指令整合为循环结构，同时学习优先级判断（如“先避开障碍再收集货物”）。

3.开放情景：“玛塔一日生活”项目中，幼儿分组设计完整叙事链（如“早晨起床→超市购物→回家充电”）。或设计小亮点（如“起床后必须完成播放起床音乐才能出发”），强化团队协作与系统设计能力。

4.5.项目化编程游戏的引导策略

4.5.1.内在需求转化游戏动力

兴趣驱动的情境创设：幼儿对“未来城市”的想象成为游戏核心动力。在创建未来城市时，幼儿不断进行设计、搭建、推翻、重建、重新设计等过程，在邀请玛塔机器人入住未来城市时，又要解决已有的城市创设与机器人生活适配的问题，因有着对这个“未来城市”更全面、完美的追求而促使着孩子们不断前行，在“设计-调试-实践”的循环中保持探索热情。

问题导向的持续挑战：城市是一个多元的综合的概念，既要考虑城市中各种建筑分布与各样的造型，又要考虑城市绿化与景观相间其中，如同真实的城市设计一样，设计总是伴随问题出现，问题又能因设计得到解决，例如当玛塔机器人因道路狭窄无法通行时，通过引导幼儿对比真实城市地图，发现“道路规划不合理”问题，激发重新设计路网的内在需求。

4.5.2.开放性问题激发创新

例如在“更好的城市”阶段，负责机器人编程的小朋友与责美化城市、设计城市环境的小朋友缺乏“互动的故事”，教师以玛塔机器人的一日生活为启发，引导幼儿关注机器人在城市中发生的机遇与故事，让玛塔在城市中像实际生活中一样有工作有娱乐游戏。

4.5.3.设计合作型任务

在任务过程中请两位或多位幼儿作为一组，通过合作设计循环路线，完成难度更高的挑战，如给予机器人运送超市货物或动物园饲料的工作，要求接力或分别行动完成任务。

4.5.4.嵌入式领域整合

科学领域：幼儿通过观察测量玛塔机器人的外形特征（重量与外观尺寸大小）感知重量与长度，结合城市中建筑的设计让玛塔机器人；真实走近各个建筑的内部，感知测量与度量单位的感念；

社会领域：在小组任务中，幼儿分工扮演“设计师”、“程序员”等角色，通过讨论达成共识，培养合作意识与规则责任感；

语言领域：在任务复盘环节，幼儿用“先…然后…最后…”句式描述编程步骤，锻炼逻辑表达能力。

艺术领域：幼儿利用预置音乐块创作“巡逻音乐”、“庆祝音乐”或为不同建筑物、道路设计辨别标识，将抽象分类概念转化为可视化符号。

4.5.5.逻辑重构与生活迁移

例如，在幼儿游戏过程中，鼓励幼儿进行实地考察参照，如观察幼儿园周边道路结构与车辆通行情况，发现“主干道需双车道宽度”进而对道路进行重新设计，在“城市干道”设置更宽的空间；如利用玛塔的载物功能，设计“货物运输”“快递配送”等任务，赋予建筑实际用途与真实可进入的建筑结构（如“超市”需接收货物、“医院”需配送药品、“动物园”需要运送饲料、“停车场”便于玛塔机器人中途休息等）为玛塔机器人创设工作需要，丰富游戏内容。

4.6.从问题解决到计算思维：

分解与调试思维：例如在“玛塔住进城市”活动中，教师通过提醒幼儿关注建筑与道路改造方案，引导幼儿将“城市改造”难题分解为“测量、建筑与道路改造、 重新布局”三步，帮助幼儿理解顺序逻辑，也将复杂问题拆解为一个个小问题，减低难度。

算法与调试思维：从“堆砌”到“规划”，初期幼儿搭建城市时，仅关注单个建筑的造型（如阿基米德积木黏合的各类造型建筑、纸杯木块组合的特殊造型等），随意摆放与组合，导致城市布局杂乱无章，部分地区热闹非凡，部分地区又冷冷清清，随意的城市布局也为玛塔机器人的入住带来了许多不便。通过组织“城市体检”活动，用玛塔机器人模拟邮差送信，暴露交通瘫痪问题。伴随着问题的出现，幼儿也开始关注到城市建筑布局与城市路网的问题，

开始对建筑与道路布局进行了改造，通过测量玛塔机器人的尺寸，调整建筑间距与道路宽度，加入磁吸地图便于小玛塔通行。

5.创新价值与反思

5.1.编程游戏项目化实践

5.1.1.从模仿到创造的思维跃迁：从“城市构建”到“系统优化”

在“一丈高的世界”案例中，幼儿的思维发展经历了“模仿➡调试➡重构➡创新”的跃迁：

模仿阶段：初期幼儿通过模仿真实城市地图设计建筑与道路，沿用简单搭建逻辑（如堆叠积木），表现为直接复现日常生活经验。

调试阶段：引入玛塔机器人后，因道路狭窄、建筑比例问题导致机器人行动受阻，幼儿自主发现问题（如“玛塔卡在建筑间无法转弯”），触发调试思维。通过测量机器人尺寸、重构路网布局（如拓宽道路至适配宽度），幼儿从“简单模仿”转向问题驱动的逻辑验证。

创新阶段：在解决机器人适配问题后，幼儿进一步设计“玛塔一日生活”（如运输任务、故事触发），赋予城市动态功能。例如，通过编程块设计“超市货物补给规则”，要求机器人按顺序完成“取货➡运输➡入库”任务链，体现复杂流程的逻辑抽象能力。这一过程展现了幼儿从“被动复制”到“主动设计”的思维跃迁，推动计算思维中的分解、抽象、调试能力层级发展。

5.1.2.跨领域整合的教学创新：从“单一任务”到“多元联结”

“一丈高的世界”案例深度融合多领域学习目标，形成编程游戏的复合型价值：

科学与工程整合：通过测量机器人尺寸（科学观察）与调整建筑间距（工程优化），幼儿将数学概念（如长度单位、几何空间）直接应用于实际问题，理解“测量与设计”的关联性。

社会与规则整合：在“玛塔快递配送”任务中，幼儿分工扮演“程序员”“物流员”等角色，通过协作制定规则（如“按顺序配送货物不得插队”），内化责任意识与社会规范，同时理解编程中的有序执行逻辑。

艺术与语言整合：幼儿为不同建筑设计标识（如用红色符号标记医院）、创作“巡逻音乐”（如用快节奏音乐对应紧急任务），将抽象指令转化为视听符号，同步提升具象表达能力；在任务复盘中，用“先规划路线，再调试代码”等连贯句式描述流程，强化逻辑性语言组织能力。

5.1.3.项目化活动与编程游戏结合的核心优势

目标驱动的渐进式学习：在“一丈高的世界”中，游戏任务嵌套于项目目标（如“设计我们的城市到让玛塔机器人住进未来城市”）下，促使幼儿能力发展：

初级阶段完成“机器人在城市中移动”的路径规划；中级阶段实现“多目标运输协同”（如同时完成超市送货与垃圾回收）；高级阶段拓展至“城市生态设计”（如玛塔机器人的一日生活），形成从单一指令到系统设计的逻辑进阶。

真实问题与游戏化反馈的闭环：项目化活动中的“城市体检”环节（用机器人模拟邮件配送暴露交通问题）将编程调试转化为具象的游戏挑战（如“拯救瘫痪的邮差”），通过即时验证（机器人成功通行）与失败反馈（机器人卡顿），帮助幼儿直观理解“代码逻辑”与“现实结果”的因果关系，激发持续探究动力。

低门槛与高拓展性的平衡：玛塔机器人通过实物编程模块降低操作门槛，同时借助开放任务（如“一丈高的世界”游戏案例），支持幼儿自由组合模块实现个性化叙事，在创造性应用中深化计算思维。

5.1.4.材料投放与收纳

玛塔编程块种类繁多，若一次性全部投放易分散幼儿注意力。实践中需采用“按需发放”策略，如仅提供当节课所需模块，课后分类回收。此外，机器人数量有限时，需设计轮换机制（如“机器人驾驶员”与“安全员”分工）维持课堂秩序。

5.2.反思

5.2.1.个案追踪缺失的局限性

难以捕捉个体发展差异性：5-6 岁幼儿的计算思维具有个别差异性，案例中对幼儿的观察主要依赖集体观察和活动前后对比，无法深入分析个体幼儿在课程中的动态发展轨迹。

弱化教师指导的针对性：教师若仅依赖集体观察，容易忽略幼儿行为背后的原因，导致“贴标签”式评价。例如，若某幼儿在编程游戏中频繁放弃任务，可能因逻辑思维薄弱或材料操作困难，但缺乏个案追踪记录时，教师易简单归因为“缺乏耐心”，错失针对性指导机会。

5.3.持续改进方向

1.制定分层观察计划：根据幼儿年龄和能力差异，设定不同追踪重点。例如，小班聚焦“顺序指令与问题分解”，大班侧重“逻辑概念与调试思维”。

2.连续性记录与分析：对重点幼儿进行定点观察（如每周 1 次），记录其在编程任务中的行为模式变化（如从模仿到创新）；留存幼儿编程作品（如路线设计图），结合时间轴分析其思维进阶。例如，对比同一幼儿在《包裹运送》单元初期与末期的路径规划复杂度。

参考文献

于颖, 周东岱, & 于伟. (2017). 计算思维的意蕴解析与结构建构. *现代教育技术*, *27*(5), 60-66.

周以真. (2021).计算思维教育新范式. *电化教育研究*.

张艳云. (2023).追随幼儿，互动共生——幼儿园游戏课程的生成构建研究. *盐城教育研究*, (2),45-50.

张新立, 朱萌, & 胡来林. (2019). 幼儿 STEM 教育现状的调查与分析. *中国教育信息化*, *25*(2), 35-38.

胡来林. (2023).幼儿编程教育中的具身认知实践. *学前教育研究*.

赵春香.(2022). 幼儿在前 课程随行——课程游戏化背景下幼儿园生成课程的实践与思考. *学前教育探索*, *18*(4),32-37.

孟鸿伟.(2024).面向数字化未来的“计算思维”. *中国教育信息化*, *30*(2),3-12.

Bers, M. U., Govind, M., & Relkin, E. (2022). Coding as another language: Computational thinking, robotics and literacy in first and second grade. In *Computational thinking in prek-5: empirical evidence for integration and future directions* (pp. 30-38).

Zhang, X., Chen, Y., Hu, L., Hwang, G. J., & Tu, Y. F. (2025). Developing preschool children's computational thinking and executive functions: unplugged vs. robot programming activities. *International Journal of STEM Education*, *12*(1), 1-18.

玛塔挑战坊：幼儿园跨学科编程课程的实施路径探索

Matatalab Challenge Program: Exploring Implementation Pathways for Interdisciplinary Programming in Kindergartens

徐佳豪，谢凌风
温州大学附属实验幼儿园
1229184521@qq.com

【摘要】 本研究以玛塔机器人（Matatalab）为核心工具，通过三轮递进式的课程实践，探索如何在幼儿园中开展跨学科编程教育。课程围绕《地球村》主题，从数学与编程的简单融合，逐步拓展到语言、艺术、科学等领域的整合，最终引导孩子解决真实的环保问题。研究发现，通过游戏化的任务设计和实体化的操作工具，幼儿不仅能轻松理解编程概念，还能在游戏中建立数学思维、语言表达能力和环保意识，同时提升幼儿的计算思维。研究结果为幼儿园教师提供了可复制的课程模板，让编程不再是冷冰冰的技术，而是孩子探索世界的趣味工具。

【关键词】 玛塔机器人；跨学科编程；幼儿园教育；计算思维

Abstract: *This study employs the Matatalab robot as a core tool to explore interdisciplinary programming education in kindergartens through a three-phase progressive curriculum. Centered on the *Earth Village* theme, the curriculum progresses from integrating mathematics and programming to incorporating language, art, and science, ultimately guiding children to solve real-world environmental issues. Findings indicate that game-based tasks and tangible programming tools enable preschoolers to grasp programming concepts effortlessly while cultivating mathematical thinking, linguistic expression, environmental awareness, and computational thinking. The research provides educators with a replicable curriculum framework, transforming programming into an engaging tool for children to explore the world.*

Keywords: Matatalab robot, interdisciplinary programming, kindergarten education, computational thinking

1.前言

当前趋势下，幼儿园课程改革中提出学科整合式教育，但在幼儿园实际的操作中，仍然存在“拼盘式”的整合问题。例如，在以探索春天这一季节为主题的课程中，虽然包含科学观察（植物生长）、数学测量（温度记录）、语言表达（春日诗歌）、艺术感知与创作（乐曲欣赏与绘画创作），但各个学科的内容仅机械并列，这导致了幼儿难以建立内在的知识联结。幼儿园的孩子正处于好奇心旺盛的年龄，他们对世界的探索充满热情，但传统的分科教学容易割裂知识之间的联系。孩子很难理解它们如何在生活中结合。而编程，尤其是实体化编程工具，恰好能成为连接不同学科的桥梁。以玛塔机器人为例，这个外形可爱的机器人不需要屏幕，孩子通过摆放实体编程指令块就能控制它行动。

举个例子：在一次“超市购物”任务中，孩子们需要用指令块让机器人从起点出发，经过三个货架（数学中的坐标概念），取回指定数量的商品（数数练习）。如果机器人中途撞

到障碍物（比如模拟的“水果堆”），孩子们要调整指令顺序（逻辑思维），并讨论如何避免“撞车”（团队协作）。这个过程中，数学、逻辑、沟通能力被自然地整合在一起。

2.玛塔挑战坊：把教室变成编程游乐园

为了让幼儿园编程教育更系统，我们设计了“玛塔挑战坊”——一个结合实体编程工具、游戏化任务和团队合作的特色活动场域。这里没有枯燥的代码，只有色彩鲜艳的指令块、可以自由拼搭的地图，以及孩子们充满创意的笑声。

2.1. 工具设计：像搭积木一样学编程

玛塔机器人的指令块设计非常“亲孩子”。比如：行动编程块：箭头明确指示左转、右转、前进、后退，孩子一摆弄就能理解；数字编程块：标有数字 2-5，控制机器人执行相对应命令的次数；循环编程块：一对标有环形的指令模块，代表重复指令动作。这些模块通过摆放在编程板上，孩子可以像拼积木一样组合指令。当按下执行键时，机器人会根据指令行动，成功时可以播放欢快的音乐，出错时则会发出声音提醒，这种即时反馈让孩子立刻知道哪里需要调整。

2.2. 任务设计：从简单到复杂的三阶挑战

课程任务分为三个难度级别，像游戏关卡一样激发孩子的挑战欲：

第一关：数学小能手

任务示例：“帮机器人快递员送包裹”

幼儿需要计算从中国（坐标 A3）到巴西（坐标 D4）的最短路径，避开海洋区域。在这个过程中，他们不知不觉学会了坐标定位和路径优化。

第二关：故事创作家

任务示例：“地球村文化之旅”

幼儿通过编程让机器人沿着大洲地图移动，每到一个国家就播放对应的音乐，并用彩色笔画出机器人的轨迹。同时他们需要为机器人设计“台词”，比如：“现在我们来到了法国，这里的埃菲尔铁塔有 300 米高！”

第三关：环保小卫士

任务示例：“地球保卫大作战”

幼儿利用玛塔传感器、感知编程块、色卡与纸模型，让机器人识别不同类型的垃圾，同时为不同类型的垃圾收集并设计分类路线。他们还用乐高积木给机器人加装“机械臂”，模拟夹起并移动垃圾的动作。

3.课堂实录：孩子们如何在“玩中学”

3.1. 场景 1：当数学遇上编程——一次“环球快递”任务

“老师！我们的机器人又掉进海里了！”5 岁的乐乐着急地喊道。原来，他的小组想让机器人从中国送快递到巴西，但设计的路径直穿太平洋。这个时候老师没有直接纠正，而是问：“你们觉得快递员是游泳过去快，还是绕道陆路更快呢？”孩子们看着地图争论起来：“绕路要走更多格子！”“但游泳会沉下去，快递就湿啦！”最终，他们决定让机器人向右绕开海洋。调整指令后，机器人顺利抵达终点，孩子们欢呼起来。 在这个过程中，幼儿不仅学会了坐标计算，还理解了“最优路径”的意义。

3.2. 场景 2：用编程讲故事——“非洲冒险“创作

在第二阶段的编程游戏活动中，5岁的朵朵所在的小组负责编排“非洲之旅”。他们给机器人设置了这样的指令：前进3步播放鼓声，左转遇到“大象”时停下，循环两次跳舞动作。当机器人随着音乐节奏扭动时，朵朵兴奋地解说：“看！机器人在跳非洲舞，它转圈的样子像不像大象甩鼻子？”在这里，编程成了幼儿表达创意的工具，语言能力和艺术思维同步提升。

3.3. 场景2：解决真实问题——设计“智能垃圾分类”

在第三阶段的任务中，孩子们参观了幼儿园附近的垃圾站，发现在幼儿园许多垃圾被错误地扔进可回收垃圾桶。回到挑战坊后，他们用玛塔机器人设计了一个“智能分类器”：当机器人检测到透明塑料瓶（蓝色卡片）时，会亮起蓝灯并运送到可回收区；如果是香蕉皮（黄色卡片），则亮起黄灯送到厨余垃圾区，在收集完成厨余垃圾后，还能让机器人用机械臂将厨余垃圾送到堆肥区。“以后我们的机器人可以帮助保洁阿姨！”孩子们骄傲地向老师展示成果。

4.跨学科编程教育的实施成效

4.1. 幼儿能力的全面提升

跨学科编程课程的实施，让孩子们在玩乐中收获了远超预期的成长。最明显的变化体现在三个方面：

逻辑思维显著增强：在“环球快递员”任务中，5岁的浩浩发现“绕开海洋比直线更快”，因为“直走会掉进水里，要重新开始”。这种对“效率”的理解，展现了初步的算法思维。课后统计显示，85%的幼儿能自主调整错误指令，而不再依赖老师直接纠正。

学科知识自然融合：5岁的朵朵在编程故事任务中，用“机器人转了3次弯”解释“三角形有三个角”，将数学概念与编程动作结合。教师观察发现，70%的孩子能在课堂讨论中主动关联两个以上学科知识。

协作能力大幅进步：混龄小组中，大班孩子担任“指令设计师”，中班孩子操作机器人，小班孩子负责装饰地图。这种分工让不同年龄的孩子都能参与，冲突率比传统课堂降低40%。

4.2. 教学模式的创新突破

跨学科编程课程实践不仅改变了孩子，也重塑了幼儿园的教学生态：

从“单科教学”到“主题探索”：过去数学课教数数、美术课画画，现在一次“地球村”挑战坊活动，就能融合坐标计算（数学）、文化故事（语言）、轨迹绘画（艺术）。教师在备课时不用再纠结这次的活动属于哪个领域的课程，孩子们的问题会自然引导课堂方向。

从“教师主导”到“儿童创生”：在“智能垃圾分类器”任务中，孩子们提出“用不同颜色的灯区分垃圾类型”，这个创意被纳入最终设计。在玛塔挑战坊活动中，幼儿提出的创意想法，也为活动创生提供了多种可能，让课程的生发与创生更加童趣。

4.3. 家园共育的新契机

跨学科编程课程的影响力延伸到了家庭：

亲子任务激发家庭互动：发放的“厨房清洁路线”编程包中，家长和孩子一起用指令贴纸规划扫地机器人路径。一位爸爸反馈：“孩子说我总把酱油瓶放在角落，机器人会撞到——没想到他连收纳都要用编程思维优化！”

家长教育观念转变：过去认为“编程就是敲代码”的家长，在开放日看到孩子用指令块创作的机器人运行轨迹后，感慨：“原来编程可以这么艺术！孩子说下次要用指令让机器人‘跳芭蕾舞’。”

4.4. 课程故事：从“旁观者”到“小领袖”的蜕变

5 岁的航航原本是班上最沉默的孩子。在第一次编程课上，他只在同伴的身边看，动手操作的次数很少。直到“设计公园路线”任务中，老师发现他悄悄用树枝在地面画出障碍物位置。当小组遇到“机器人总撞树”的问题时，航航突然指着自己的草图说：“这里多了一棵树！要往左边绕。”这个建议让小组成功通关，航航被推选为“地图顾问”。现在的他，会主动帮同伴分析指令错误：“你看，这里少了一个右转块，就像你上次少穿了一只袜子！”——用生活化的比喻解释编程逻辑，正是课程带来的惊喜改变。

5.创新实践建议

要让跨学科编程课程真正落地生根，需要从孩子的兴趣出发，结合幼儿园的实际条件，设计可操作的实施方案。以下是经过实践验证的四项核心策略，帮助教师轻松开展课程，同时激发孩子的学习热情：

5.1. 以生活为课堂，让编程与日常生活经验无缝衔接

孩子的学习始于对周围世界的观察。教师可以从他们熟悉的生活场景中提取任务主题，比如“帮妈妈取快递”或“整理玩具柜”。例如，在“超市购物”任务中，孩子们用编程规划机器人从入口到零食区的路线，途中需要绕过“水果堆”（障碍物），并计算需要拿几包饼干（数数练习）。这种设计让编程不再是抽象的技术，而是解决实际问题的工具。一位老师分享道：“当孩子发现编程能帮‘机器人快递员’更快完成任务时，他们主动要求多设计几个送货路线，甚至提议给机器人加上‘避雨功能’——这种自发的探索正是课程成功的关键。”

5.2. 打破年龄界限，用混龄协作点燃学习火花

不同年龄的孩子在同一个课堂中能碰撞出意想不到的创意。在实践中，我们让大班孩子担任“指令设计师”，负责规划复杂路径；中班孩子操作机器人，调试指令顺序；小班孩子则用贴纸、彩笔装饰任务地图。这种分工不仅让每个孩子都能参与，还促进了自然的知识传递。比如，在一次“公园导览”任务中，5 岁的大班孩子教 4 岁同伴：“转弯时要先放右转块，就像你骑车那样！”——用生活经验解释编程逻辑，让学习变得生动易懂。数据显示，混龄小组的合作效率比同龄组提高 35%，冲突率下降 40%。

5.3. 让亲子家庭成为编程教育的延伸课堂

跨学科编程教育不应局限于幼儿园。我们设计了“亲子编程工具包”，内含可粘贴的编程指令贴纸和家庭任务卡。例如“周末出游路线规划”任务，家长和孩子一起用贴纸在纸上标注景点顺序，拍照上传班级群分享。一位爸爸反馈：“孩子坚持要把动物园放在路线最后，因为‘看完老虎回家会做噩梦’——这让我意识到，编程不仅是逻辑训练，更是孩子表达内心想法的窗口。”此外，定期举办“家庭编程开放日”，让家长亲眼看到孩子如何用指令块创作轨迹画，彻底打破“编程=敲代码”的刻板印象。

5.4. 为教师赋能，用工具包化解备课压力

跨学科课程设计常让教师感到负担过重。为此，教师可以使用支持工具：学科联结模板：将数学符号、艺术元素与编程任务预设关联，教师只需“填空”即可生成教案；半成品任务卡库：包含多样的未完成的任务设计卡，例如“机器人快递员”任务卡已标注货架坐标，教师只需补充商品数量。这些备课工具的使用，也能让编程教育教师的备课变得更加轻松。

6.结论与展望

通过玛塔挑战坊的实践，孩子们在游戏中学会了逻辑思考、团队合作，更重要的是，他们开始用跨学科的视角观察世界。当 5 岁的浩浩骄傲地向妈妈展示他设计的“会画彩虹的机器人”时，当孩子们围在一起争论“怎样让垃圾分类更智能”时，我们看到了教育的另一种可能——编程不再是高深的技术，而是每个孩子都能掌握的探索工具。正如一位家长在反馈表中写的：“没想到我的孩子能说出‘编程就像搭积木，错了就重新来’——这种不怕失败、主动探索的精神，比学会多少知识更重要。”未来，我们将利用玛塔挑战坊的活动设计与开展继续优化跨学科编程活动设计，让更多幼儿园的孩子在欢声笑语中，成长为有逻辑、有创意、有责任感的小小世界公民。

参考文献

王素, 李芒(2020). 幼儿园 STEAM 教育中计算思维的培养路径.《学前教育研究》, 38(4), 45-52.

李泽湘. (2021). 幼儿计算思维的游戏化教学模式研究. 学前教育研究, 39(5), 18-25.

周以真.(2006). 计算思维.《中国计算机学会通讯》, 2(1), 27-29.

周以真. (2019). 计算思维：从概念到课程实施. 全球教育展望, 48(2), 25-32.

董奇(2018). 儿童早期计算思维发展的理论与实践.《教育研究》, 39(5), 112-120.

Bers, M. U.(2017). Coding as a Playground: Programming and Computational Thinking in the Early Childhood Classroom. Routledge.

Papert, S.(1980). Mindstorms: Children, Computers, and Powerful Ideas. Basic Books.

Resnick, M.(2017). Lifelong Kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play. MIT Press.

Wing, J. M.(2006). Computational Thinking. Communications of the ACM, 49(3), 33-35.

基于 GenAI 的结对编程策略对初中生计算思维态度和编程自我效能感的影响研究

A study of the effects of GenAI-based pair programming strategy on middle school students' attitudes toward computational thinking and programming self-efficacy

王千姿[1*]，孟晓培[1]
[1]温州大学 教育学院
*wqz2316108432@163.com

【摘要】 在数字化时代，计算思维已成为 K-12 阶段教育的重要议题。针对当前中小学编程教学面临的教学方法单一、学生编程信心不足等问题，本研究将 GenAI 技术结合结对编程策略，将温州市某学校七年级三个班的学生分为实验组 1（P-GenAI 编程学习方法）、实验组 2（C-GenAI 编程学习方法）和对照组（常规编程学习方法）。经过 10 周的教学实验，采用 Mann-Whitney U 检验对三组学生的计算思维态度和编程自我效能感进行了比较分析，结果显示，虽然三组学生在计算思维态度和编程自我效能感上均无显著差异，但实验组 1 的学生在各维度上的表现普遍优于实验组 2 和对照组，呈现出积极的趋势。基于此，研究提出了强化编程韧性、突破人机协作瓶颈以及构建多智能体协同生态等建议，以期为中小学编程教育提供科学依据和实践指导。
【关键词】 生成式人工智能；GenAI；结对编程；计算思维态度；编程自我效能感

Abstract: *This study explores the integration of GenAI technology with pair programming in K-12 education to address issues like limited teaching methods and students' lack of programming confidence. It involved three seventh-grade classes in Wenzhou, divided into three groups: Experimental Group 1 (P-GenAI programming), Experimental Group 2 (C-GenAI programming), and a control group (traditional programming). After a 10-week teaching experiment, the results showed no significant differences in computational thinking attitude or programming self-efficacy among the groups. However, Experimental Group 1 generally performed better in all aspects, suggesting positive trends. The study recommends enhancing programming resilience, overcoming human-computer collaboration challenges, and building a collaborative multi-intelligence ecosystem to improve programming education.*
Keywords: Generative artificial intelligence, GenAI, Pair programming, Computational thinking attitudes, Programming self-efficacy

1.引言

计算思维(Computational Thinking)作为数字时代公民必备的核心素养，已成为全球教育研究的重要议题，并逐渐向 K-12 阶段延伸(孟鸿伟, 2024)。近年来世界各国特别是美国、英国、新西兰等国家，已经将计算思维的培养纳入人才培养计划和课程体系(孙丹&李艳, 2019)。国内教育也在“双减”政策的推动下，积极推动中小学教育从单一的知识传授转向素养导向的全面发展，《义务教育信息科技课程标准（2022 年版）》中就明确提出计算思维是学科核心素养的关键要素。在诸多提升儿童计算思维的方法中，编程教育以其特有的潜能受到广泛关注(陆霞&董永权, 2024)，也逐步被列为我国中小学信息技术课程的核心内容。然而，当前中小学编程教学仍面临诸多挑战，如教学方法单一、学生编程信心不足以及教师忽视形成性评价(李彤彤 et al., 2022)，这些问题在一定程度上影响了编程教育的教学效果。随着生成

式人工智能(Generative Artificial Intelligence, 简称 GenAI)的快速发展，其在编程教育中的独特优势也逐渐突显。GenAI 不仅能提供即时的智能反馈，还能辅助学生进行代码理解、调试与优化，从而提升学习效果和编程体验(Yilmaz, 2023)。

在此背景下，本研究以 GenAI 支持的编程学习策略为切入点，设计并比较三种不同学习方法对学生计算思维态度与编程自我效能感的影响。研究选取温州市某中学七年级三个班共 104 名学生为研究对象，随机分为两个实验组和一个对照组。实验组 1 采用 GenAI 支持下融入结对编程策略的编程学习方法，简称为 P-GenAI(Pair Programming-Generative Artificial Intelligence)，即在学生两两结对的过程中，GenAI 作为智能伙伴提供辅助建议与代码反馈；实验组 2 采用 GenAI 支持下的常规编程学习方法，简称为 C-GenAI(Conventional－Generative Artificial Intelligence)，即学生个体在 GenAI 的智能引导下完成编程任务；对照组采用传统的编程学习方法，独立完成编程任务，简称为 C-C(Conventional－Conventional)。本研究通过对比三组学生在实验前后的表现变化，探讨不同编程学习策略对学生计算思维相关非认知变量的影响，以期为中小学编程教学的发展提供实践指导。

2.文献综述

2.1.计算思维态度与编程自我效能感的相关研究

计算思维的具体内涵和维度尚未完全统一。在诸多定义中，计算思维三维框架(Brennan & Resnick, 2012)基于可视化思维和过程思维，在教育实践中有较强的可操作性(Zhong et al., 2016)，而且计算概念，计算实践和计算观念也能很好地对照教学的三维目标(王旭卿,2014)。在计算思维三维框架的基础上，郁晓华等(2020)界定了计算观念（具备以计算的视角洞察与理解自身和周围世界的良好品质）的范畴。从这一定义出发，计算观念可以从计算思维态度和编程自我效能感两个方面来观测。其中，计算思维态度(Computational Thinking Attitudes)作为计算思维的非认知组成部分，已有研究通常强调个体在面对复杂问题时所表现出来的信心、坚持性和不确定性的容忍度等特质。Sun 等人(2022)在研究中使用编程态度量表来评估学生的编程态度，结果表明，学生的编程态度显著预测了他们的计算思维技能。此外，Zapata-Cáceres 等人(2022)在研究教师对计算思维的态度和自我评估时，强调了教师在教学中对计算思维重要性的认知和态度。自我效能感(Self-efficacy)的培养也被证实在促进编程学习中具有重要作用。Bandura(1997)提出，自我效能感是个体对自己完成某一任务能力的主观判断，直接影响其行为选择、努力程度与应对策略。自我效能感在计算机领域的应用时，其表现为计算机编程自我效能感(Programming Self-Efficacy)。编程自我效能感是指个人有能力在计算机领域发挥作用的信念。Ramalingam 等人(1998)将编程自我效能分为四个维度，即独立和毅力、复杂编程能力、自我调节能力和简单编程能力。已有研究表明，计算思维态度和自我效能感不仅对学生的编程学习成效具有显著预测作用，也对其计算思维的持续发展构成支撑。例如，Durak(2019)发现，在编程方面表现出色的学生既有较高的自我效能感水平，也有较高的计算思维技能。因此，对学生的计算思维态度和自我效能感进行探究，有利于使编程教学达到更理想的效果。

2.2.GenAI 在编程教育中的应用研究

近年来，随着预训练技术的发展和计算硬件的提升，生成式人工智能取得了突破性进展，为编程教育提供了新的可能性。GenAI 能够通过自然语言理解、代码生成、实时交互和智能反馈等方式辅助编程学习(Aljanabi, 2023)，提高学生的编程能力和学习效率(Piccolo et al., 2023)。例如，Yilma(2023)研究表明 ChatGPT 纳入的编程中，可以为学生提供关于编程主题

的清晰解释、编程示例以及应用程序，提高学生的认知能力，减少在编码任务上花费的时间；Piccolo 等人(2023)对来自生物信息学入门课程的 184 个编程练习进行了实验，结果表明，ChatGPT 第一次尝试就正确解答了 139 个练习(75.5%)。经过几次尝试和自然语言反馈之后，ChatGPT 的正确率达到了 97.3%；李秀(2024)结合布卢姆教育目标分类理论，设计了一套生成式人工智能应用于编程教学的“三学”策略，展示了生成式人工智能在编程教学中的教育价值。此外，GenAI 还能通过自动化调试功能，帮助学生发现和修正错误，提高其问题解决能力。尽管 GenAI 在编程教育中的应用展现出明显优势，但也有研究者担忧长期依赖 GenAI 可能会削弱学生的批判性思考、探索、验证和主动总结的能力(Yu, 2023)，容易产生“机械接受”的现象，使其在问题求解过程中过度依赖智能工具，从而影响深层次的知识建构。因此，在教学设计中，需要合理规范 GenAI 的使用方式，强调其作为辅助工具的角色，同时结合教师指导与策略支持，鼓励学生在使用 GenAI 工具的过程中主动思考与反思，从而实现计算思维与编程能力的同步发展。

2.3.结对编程中计算思维的相关研究

结对编程(Pair Programming)是一种合作学习策略，在编程教育中被广泛应用。该模式最早由 Williams 和 Kessler (2000)提出，指两名学生共同使用一台计算机，其中一人作为“驾驶员”(Driver)编写代码，另一人作为“观察员”(Navigator)进行代码审查和问题思考，并定期交换角色。结对编程的这种合作模式既能通过实时协作即时检测和纠正代码，从而减少缺陷并改进整体代码结构，也能在交流合作的过程中进行知识共享，提高学生的参与度。研究表明，与单独的编程任务相比，结对编程可以导致更高水平的学生满意度和减少的挫折感的报告(Wei et al., 2021)。在传统的教育环境中，结对编程的实施也面临诸多挑战。一是学生间的个体差异可能导致学习不均衡，影响合作质量。为此，研究者提出了多种配对策略，如基于性别(Choi K S, 2015)、学习风格(Demir, Ö. & Seferoglu, S. S, 2018)和伙伴关系(李彤彤 et al., 2022)等方式进行分组，以优化结对编程的教学效果。二是难以安排同伴之间的协作会话。尤其是在高等教育环境中，学生往往有不同的时间表和约定(Govender & Grayson, 2006)。考虑到传统结对编程存在的局限性，研究者开始将 AI 引入结对编程的实践中，探索人机协作的新范式。如 Fan 等人(2025)对 234 名学习 Java 网络应用程序开发课程的本科生进行了准实验研究，结果显示，虽然人工智能辅助结对编程提高了学习动机、减少了焦虑并改善了成绩，但它并不能完全与通过人-人结对编程实现的协作深度和社会存在相媲美。因此，如何在保证人工智能辅助结对编程优势的情况下，科学地提高学生在合作过程中的学习体验，是亟待解决的问题。

综上所述，计算思维态度与编程自我效能感作为计算思维的重要非认知维度，在提升学生编程学习动机与成效方面发挥着关键作用。GenAI 的发展为编程教学提供了便利的智能支持，通过自然语言交互与即时反馈降低了编程门槛，有助于增强学生的学习信心与持续投入。但同时也引发了对学生批判性思维削弱与过度依赖技术的担忧。将 GenAI 融入结对编程，有望在保留人际合作优势的基础上，拓展人机协作的深度与广度，提升学习适应性与支持度。当前针对该融合情境下学生非认知因素变化的实证研究较少，因此，深入探究 GenAI 支持下结对编程对学生非认知因素的影响，有助于拓展相关研究视角，丰富智能时代背景下编程教育的理论基础与实践路径。

3.研究设计

3.1.研究对象

本研究选取了温州市某学校七年级三个班的学生为研究对象，随机设置实验组和对照组。其中实验组 1 采用 GenAI 支持下融入结对编程策略的编程学习方法，简称为 P-GenAI 编程学习方法(Pair Programming-Generative Artificial Intelligence)；实验组 2 采用 GenAI 支持下的常规编程学习，简称为 C-GenAI 编程学习方法(Conventional－Generative Artificial Intelligence)；对照组采用的常规编程学习方法，简称为 C-C 编程学习方法(Conventional－Conventional)。本研究以初中生的计算思维态度和自我效能感为因变量。为了保证实证研究的科学性与严谨性，本研究三组的授课教师、教学内容、教学环境、教学时长和样本年级均相同。此外，在实验前，实验组 1、实验组 2 和对照组三个组的计算思维态度和自我效能感成绩均无显著差异。

表 1 研究对象统计

班级	组别	编程学习方法	N	男生	女生
七（5）班	实验组 1	P-GenAI	34	19	15
七（8）班	实验组 2	C-GenAI	33	18	15
七（7）班	对照组	C-C	37	23	14

3.2.研究工具

本研究采用计算思维态度量表由 Korkmaz 等人(2017)编制，共 8 道题，涵盖表达和创造力两个指标；编程自我效能感量表则引用 Volkan Kukul 等人(2017)编制的《计算机编程自我效能感量表》，共 20 道题，涉及独立性与毅力、复杂编程、自我调节及简单编程四个维度。为确保测量工具的可靠性，本研究对两个量表进行了信度检验，结果显示计算思维态度量表的 Cronbach's α系数为 0.875，编程自我效能感量表的 Cronbach's α系数为 0.860，表明两个量表均具有良好的内部一致性与测量稳定性。除此之外，研究者根据学生的课堂表现，在实验组中从不同性别、不同知识水平的学生中选取学生进行访谈，访谈大纲改编自 Fang et al.(2022)进一步为假设结论提供一定程度的参考。

3.3.教学过程

3.3.1.教学平台

目前，中小学中编程教学大多借助 Python、Scratch、编程猫等软件。文本编程工具如 Python、MaLT2 等在语法上更接近自然语言，易于阅读和维护(陆霞&董永权, 2024)；而图形化编程工具如 ScratchJr、Alice 等则只需要像摆积木一样拖放编程模块，就可以为作品赋予新的能力和实现更多创意玩法，在避开复杂语法的同时完美保留了编程思维。

图 1 Paracraft 软件用户编辑界面

实验教学选用的教学软件是 Paracraft 编程软件，用户编辑界面如图 1 所示。Paracraft 是一款免费开源的 3D 动画与编程创作软件，具有多语言编程环境，支持图形化、NPL、Python、C++等语言，软件中的图形化编程和文本编程还可以无缝切换，通过编写程序，能够控制动画，实现复杂的程序效果。Paracraft 相比其他编程软件创造了一个符合编程习得的环境，先建造、后动画、最后添加逻辑，学生学习兴趣浓厚；并且具备“多人联网”功能支持结对策略的开展（如图 2），适用于初中生的编程教学。

图 2 Paracraft 软件的“多人联网”功能

3.3.2.GenAI 学习工具

本研究选用的 GenAI 工具是 Keepwork，其深度融合 Paracraft 编程环境，并针对编程学习提供专业化的知识库支持，在一定程度上能弥补 GenAI 工具因通识性互联网数据训练而导致的知识偏差。具体来说，Keepwork 提供了多个版本(免费版、Turbo 和 pro)，满足不同用户的需求，能够在 Paracraft 编程环境中为学生实时解答疑惑，优化代码，提供更有针对性的指导和反馈。

图 3 Keepwork 平台主界面

3.3.3.教学计划与内容

实验周期共 10 周，分为准备阶段（2 周）、学习阶段（6 周）、后测阶段（2 周）。教学内容主要参考了《Paracraft 编程入门》（哈尔滨工业大学出版社）和《Paracraft 青少年 3D 动画编程入门》（清华大学出版社）两本教材，以及 Paracraft 编程软件内包含的课程、资源以及优秀作品等，根据学生现有知识水平，选取了五子棋和开心大闯关两个项目。五子棋项目涉及坐标定位、事件处理和游戏规则逻辑的编写，而飞行的小鸟项目则主要涉及物理模拟、碰撞检测和角色控制，让学生在掌握基本的编程技能同时提升逻辑思维和问题解决能力。

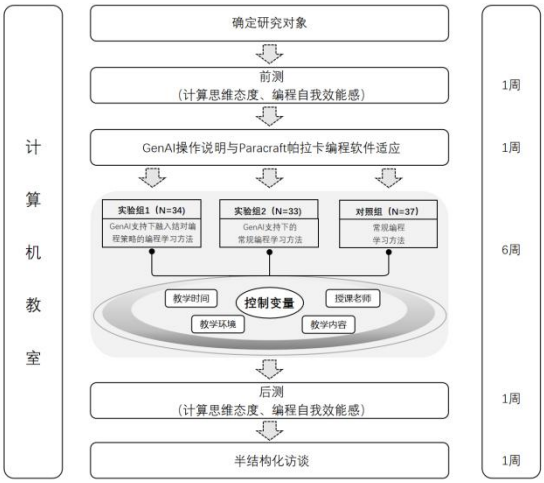


图 4 实验流程

4 研究结果与分析

4.1.量化数据分析

本研究对计算思维态度和编程自我效能感数据进行了 Shapiro-Wilk 检验，发现数据非正态分布，因此使用 Mann-Whitney U 检验来比较三组学生在计算思维态度和编程自我效能感上的差异。所有检验均使用 SPSS 26.0 和 jamovi2.3.28 软件完成。

表 2 为实验组 1 和实验组 2 Mann-Whitney U Test 的分析结果。计算思维态度后测的 Mann-Whitney U Test(U=366, p=.291>.05, Z=-1.063)和编程自我效能感后测的 Mann-Whitney U Test(U=386, p=.459>.05, Z=-.748)结果表明，实验组 1 和实验组 2 在计算思维态度和编程自我效能感后测上均不存在显著差异。此外，本研究还通过 Mann-Whitney U 检验分析了计算思维态度和编程自我效能感的各个维度。结果显示，计算思维态度和编程自我效能感各维度均无显著差异，但在中位数(Md)上实验组 1 均大于实验组 2。这些结果表明，实验组 1 和实验组 2 在计算观念各维度的差异上并不显著(p>0.05)，效应量较小，但实验组 1 在各维度上的表现存在普遍优于实验组 2 的趋势，该趋势并不排除偶然性因素的影响。

表 2 实验组 1 和实验组 2 计算观念水平后测分析结果

Variable	Group	*N*	Mean	*SD*	Md	U	*p*	*r*	*Z*
计算思维态度	实验组 1	34	3.68	1.11	4.00	366	0.291	0.160	-1.063
	实验组 2	33	3.32	1.10	3.13				
表达	实验组 1	34	3.76	1.13	4.00	356	0.228	0.182	-1.213
	实验组 2	33	3.34	1.19	3.00				
创造力	实验组 1	34	3.60	1.12	3.75	374	0.353	0.140	-.937
	实验组 2	33	3.30	1.11	3.25				
自我效能感	实验组 1	34	3.60	1.11	3.40	386	0.459	0.113	-.748
	实验组 2	33	3.32	1.03	3.20				
独立和毅力	实验组 1	34	3.66	1.17	3.60	376	0.363	0.137	-.917
	实验组 2	33	3.34	1.09	3.20				
复杂编程	实验组 1	34	3.68	1.12	3.50	364	0.273	0.163	-1.105
	实验组 2	33	3.30	1.01	3.00				
自我调节	实验组 1	34	3.53	1.14	3.33	384	0.434	0.118	-.790
	实验组 2	33	3.28	1.01	3.00				
简单编程	实验组 1	34	3.56	1.07	3.20	411	0.712	0.056	-.377
	实验组 2	33	3.35	1.16	3.20				

表 3 为实验组 1 和对照组 Mann-Whitney U Test 的分析结果。计算思维态度后测的 Mann-Whitney U Test(U=281, p=.248>.05, Z=-1.165)和编程自我效能感后测的 Mann-Whitney U Test(U=303, p=.454>.05, Z=-.759)结果表明，实验组 1 和对照组在计算思维态度和编程自我效能感后测上均不存在显著差异。此外，本研究还通过 Mann-Whitney U 检验分析了计算思维态度和编程自我效能感的各个维度。结果显示，计算思维态度和编程自我效能感各维度均

无显著差异，但在中位数(Md)上实验组 1 均大于对照组。这些结果表明，实验组 1 和对照组在计算观念各维度的差异上并不显著(p>0.05)，效应量较小，但实验组 1 在各维度上的表现存在普遍优于对照组的趋势，该趋势并不排除偶然性因素的影响。

表 3 实验组 1 和对照组计算观念水平后测分析结果

Variable	Group	*N*	Mean	*SD*	Md	U	*p*	*r*	*Z*
计算思维态度	实验组 1	34	3.68	1.11	4.00	281	0.248	0.187	-1.165
	对照组	37	3.21	1.24	3.25				
表达	实验组 1	34	3.76	1.13	4.00	272	0.189	0.212	-1.323
	对照组	37	3.24	1.23	3.25				
创造力	实验组 1	34	3.60	1.12	3.75	271	0.182	0.214	-1.343
	对照组	37	3.18	1.28	3.00				
自我效能感	实验组 1	34	3.60	1.11	3.40	303	0.454	0.122	-.759
	对照组	37	3.33	1.12	3.15				
独立和毅力	实验组 1	34	3.66	1.17	3.60	303	0.449	0.122	-.766
	对照组	37	3.39	1.28	3.40				
复杂编程	实验组 1	34	3.68	1.12	3.50	302	0.437	0.125	-.786
	对照组	37	3.35	1.24	3.25				
自我调节	实验组 1	34	3.53	1.14	3.33	288	0.301	0.167	-1.043
	对照组	37	3.20	1.15	3.20				
简单编程	实验组 1	34	3.56	1.07	3.20	313	0.566	0.093	-.582
	对照组	37	3.41	1.04	3.00				

表 4 为实验组 2 和对照组 Mann-Whitney U Test 的分析结果。计算思维态度后测的 Mann-Whitney U Test(U=313, *p*=.711>.05, Z=-.380)和编程自我效能感后测的 Mann-Whitney U Test(U=331, *p*=.971>.05, Z=-.046)结果表明，实验组 2 和对照组在计算思维态度和编程自我效能感后测上均不存在显著差异。此外，本研究还通过 Mann-Whitney U 检验分析了计算思维态度和编程自我效能感的各个维度。结果显示，计算思维态度和编程自我效能感各维度均无显著差异，但在中位数(Md)上实验组 2 的创造力、编程自我效能感、简单编程大于对照组。这些结果表明，实验组 2 和对照组在计算观念各维度的差异上并不显著(p>0.05)，且效应量较小。

表 4 实验组 2 和对照组计算观念水平后测分析结果

Variable	Group	*N*	Mean	*SD*	Md	U	*p*	*r*	*Z*
计算思维态度	实验组 2	33	3.32	3.13	1.10	313	0.711	0.062	-.380
	对照组	37	3.21	3.25	1.24				
表达	实验组 2	33	3.34	1.19	3.00	324	0.876	0.029	-.176
	对照组	37	3.24	1.23	3.25				
创造力	实验组 2	33	3.30	1.11	3.25	299	0.528	0.103	-.641
	对照组	37	3.18	1.28	3.00				
自我效能感	实验组 2	33	3.32	1.03	3.20	331	0.971	0.008	-.046
	对照组	37	3.33	1.12	3.15				
独立和毅力	实验组 2	33	3.34	1.09	3.20	327	0.904	0.021	-.130
	对照组	37	3.39	1.28	3.40				
复杂编程	实验组 2	33	3.30	1.01	3.00	325	0.881	0.026	-.160
	对照组	37	3.35	1.24	3.25				
自我调节	实验组 2	33	3.28	1.01	3.00	311	0.682	0.068	-.419
	对照组	37	3.20	1.15	3.00				
简单编程	实验组 2	33	3.35	1.16	3.20	331	0.963	0.009	-.056
	对照组	37	3.41	1.04	3.00				

4.2. 访谈结果分析

为了更加全面地了解不同编程策略对学生计算思维态度与编程自我效能感的影响，本研究在实验结束后随机访谈了实验组 1 与实验组 2 的部分学生，进一步挖掘学生在学习过程中的真实体验和主观感受。访谈结果显示，实验组 1(P-GenAI)中的多数学生认为，与同伴协作加上 GenAI 的辅助能够提升他们面对任务时的信心，减少编程任务带来的焦虑。一位学生提到：“当我不太懂的时候，搭档可以讲给我听，不懂的地方我们也可以一起问 AI，我觉得更轻松也更有动力。”这种编程协作模式在一定程度上增强了学生在学习过程中的参与感

和安全感，帮助他们更持久地投入到编程任务中，尽管在量化数据中未体现出显著差异，但访谈反馈反映出其对非认知因素的潜在正向作用。

相比之下，实验组 2(C-GenAI)的学生在学习过程中虽能得到 AI 的技术支持，但部分学生表示“有时候 AI 说的我还是看不太懂”，或“没有人讨论的话，我不太确定是不是理解对了”。这在一定程度上说明，单人使用 GenAI 虽具效率优势，但缺乏协同交流与思维碰撞，可能影响其在认知建构与态度养成方面的积极作用。

访谈数据与量化结果相辅相成，实验组 1 虽然未在统计意义上显著优于其他组，但在学生主观体验和非认知反馈中体现出更高的积极性与参与感。这一结果表明，在注重 AI 工具支持的同时，我们仍需关注人际协作的情感价值和交互作用，为学生提供更加和谐的智能学习环境。

5.研究总结与建议

本研究基于生成式人工智能支持下的结对编程教学，探讨其对初中生计算思维态度与编程自我效能感的影响。通过准实验设计比较三组学生在实验前后的变化趋势，结果表明：尽管实验组在量化数据上未表现出统计意义上的显著差异，但从中位数和访谈数据来看，实验组 1 在学生学习信心、协作参与和主动调试等方面表现出更积极的态度和感受。研究结果表明，GenAI 在辅助编程教学的同时，应关注学生非认知因素的变化，并在教学设计中寻求技术支持与人际交互的平衡。因此，在未来的教学实践中，应更加重视技术支持与人际互动的有机融合，兼顾学习效率与心理发展，以推动计算思维的深入发展。基于本研究的发现，提出以下三点建议：

5.1.强化编程韧性，筑牢计算思维的心理根基

编程韧性(Programming Resilience)是计算思维非认知维度培养的关键要素，强调学习者在编程过程中面对困难时的坚持与适应能力。王佑镁等(2023)基于“韧性”发展的 7C 能力结构模型与编程课程的实际特点，构建了编程韧性的“4C”能力结构模型，这为数字时代培养计算思维提供了新方向。在 GenAI 支持的结对编程教学中，学生在协作过程中可能因认知冲突和编程错误产生挫败感，而 GenAI 的即时反馈和自动补全虽然能够降低学习门槛，但若未能合理引导，可能导致学生过度依赖技术支持，从而削弱其独立解决问题的能力。因此，教学应从多个层面强化学生的编程韧性，帮助学生建立对挑战的积极认知。首先，采用任务递进式设计，让学生在挑战中不断适应编程逻辑，培养面对错误的耐挫力；其次，营造鼓励试错和反思的课堂氛围，要求学生记录调试过程，形成有效的问题解决策略；最后，结合项目式学习与竞赛机制，让学生在团队协作与竞争压力下不断提升自信和应变能力。通过以上策略，学生能够在真实的问题解决中逐步培养心理耐力，培养积极的学习态度与稳定的计算思维心理基础。

5.2.突破人机协作瓶颈，塑造批判性智能共创模式

GenAI 的融入让结对编程不再仅限于人与人之间协作，然而在人机协作过程中学生应该如何保持批判性思维，避免机械接受 AI 的输出内容，这成为当下智能时代亟待解决的关键问题。批判性思维强调学习者能够独立思考、分析和评估 AI 生成的代码，而非简单依赖 AI 提供的方案。在此基础上，可引入对弈性训练(Adversarial Training)模式(刘邦奇 et al., 2025)，让学生在代码编写过程中与 GenAI 进行“博弈”，即学生主动识别、挑战 AI 代码中的潜在漏洞，以提升代码优化与反思能力。此外，逆向工程(Reverse Engineering)作为计算思维训练的高级形式，能够帮助学生透过代码表象理解其底层逻辑(翟雪松 et al., 2024)。在教学设计上，

可提供 GenAI 生成的代码，让学生进行结构解析、错误修正与逻辑优化，同时鼓励学生基于已有代码进行拓展和重构。为增强教学实效，还可结合“代码挑战赛”或“人机对抗编程”模式，要求学生在限定时间内优化 GenAI 生成的代码，以此培养精确表达、逻辑推理和问题拆解能力。通过批判思维引导、对弈性训练及逆向工程策略的融合，学生不仅能够提升计算思维的精细化程度，同时也能在与 AI 的协作中培养更强的创造力与问题解决能力。

5.3.构建多智能体协同生态，拓展智能编程的教育边界

单一智能体的介入能够提升编程效率，但无法满足复杂学习场景的多样化需求，因此，构建多智能体生态成为增强人机协作效能的重要方向(翟雪松 et al., 2024)。当前的 GenAI 主要用于代码生成和错误检测，而未来的编程教育应进一步拓展其应用场景，使多个智能体在学习过程中各司其职、协同运作。例如，可引入三层智能体架构：第一层为代码生成智能体，如 ChatGPT，负责提供基础代码；第二层为优化与调试智能体，专注于检测语法错误、逻辑漏洞与性能优化；第三层为学习支持智能体，结合数据分析技术，为学生提供个性化学习建议。这种多智能体的协同生态模拟了真实的软件开发环境，使学生在任务驱动式学习过程中，学会合理调配不同智能体资源，提升任务拆解、调试优化及团队协作能力。此外，多智能体协作还能支持学习者的动态反馈机制，根据学生的学习数据调整任务难度，实现个性化编程教育。通过构建多智能体生态系统，不仅能够提升人机协作的整体效能，还能为学生提供更具交互性、适应性与智能化的学习体验，从而为未来智能编程教育的深度发展奠定基础。

参考文献

白雪梅 & 顾小清.(2019).K12 阶段学生计算思维评价工具构建与应用.中国电化教育,(10),83-90.

傅骞,解博超 & 郑娅峰.(2019).基于图形化工具的编程教学促进初中生计算思维发展的实证研究.电化教育研究,40(04),122-128.doi:10.13811/j.cnki.eer.2019.04.016.

李彤彤,郝晴,文雨,苏新玉,方玉华 & 刘佳旋.(2022).基于学习风格和伙伴关系的配对编程对小学生计算思维的影响研究.远程教育杂志,40(03),105-112.doi:10.15881/j.cnki.cn33-1304/g4.2022.03.007.

李锋,程亮 & 顾小清.(2022).计算思维学业评价的内容构建与方法设计——文献比较研究的视角.中国远程教育,(02),65-75+77.doi:10.13541/j.cnki.chinade.2022.02.008.

刘邦奇,聂小林,王亚飞,袁婷婷,赵子琪 & 张国强.(2025).生成式 AI 赋能教育：技术框架、应用场域及价值——2024 智能教育发展研究报告.中国电化教育,(03),61-70.

陆霞 & 董永权.(2024).面向计算思维培养的儿童编程教育研究综述.中国教育信息化,30(10),109-118.

孟鸿伟.(2024).面向数字化未来的“计算思维”.中国教育信息化,30(02),3-12.

任友群,隋丰蔚 & 李锋.(2016).数字土著何以可能?——也谈计算思维进入中小学信息技术教育的必要性和可能性.中国电化教育,(01),2-8.

孙丹 & 李艳.(2019).国内外青少年编程教育的发展现状、研究热点及启示——兼论智能时代我国编程教育的实施策略.远程教育杂志,37(03),47-60.doi:10.15881/j.cnki.cn33-1304/g4.2019.03.005.

王靖,马志强,刘亚琴 & 杜鸿羽.(2022).面向计算思维的可视化编程活动设计与应用.现代教育技术,32(09),55-63.

王佑镁,南希烜,李宁宇,尹以晴 & 柳晨晨.(2023).编程韧性：数字时代计算思维培养的新议题.现代教育技术,33(02),14-23.

郁晓华 & 王美玲.(2020).计算思维培养之路还有多远?——基于计算思维测评视角.开放教育研究,26(01),60-71.doi:10.13966/j.cnki.kfjyyj.2020.01.007.
岳彦龙,张学军 & 梁屿藩.(2022).人工智能教学如何培养高中生的计算思维?——基于人工智能案例驱动的 Python 编程教学的实证研究.基础教育,19(01),74-84.
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)[M].北京：北京师范大学出版社,2022.4-11.
张进宝，(2020).国际计算思维挑战赛试题集锦[M].北京：电子工业出版社.
翟雪松,张丽洁,夏亮亮,徐鑫 & 朱强.(2024).基于 GAI 的逆向工程教学思维在人机协作中的应用研究——以编程教育为例.电化教育研究,45(09),61-68.doi:10.13811/j.cnki.eer.2024.09.008.
Aljanabi, M. (2023). ChatGPT: Future directions and open possibilities. *Mesopotamian journal of Cybersecurity*, *2023*, 16-17.
Brennan, K., & Resnick, M. (2012, April). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada* (Vol. 1, p. 25).
Choi, K. S. (2015). A comparative analysis of different gender pair combinations in pair programming. *Behaviour & Information Technology*, 34(8), 825-837.
Demir, Ö., & Seferoglu, S. S. (2021). The effect of determining pair programming groups according to various individual difference variables on group compatibility, flow, and coding performance. *Journal of Educational Computing Research*, *59*(1), 41-70.
Durak,H,Y,Yilmaz,F.G,K.,&Yilmaz,R.(2019). Computational thinking,programming self-efficacy,problem solving and experiences in the programming process conducted withrobotic activities.*Contemporary Educational Technology*,10(2),173-197.
Fang, J. W., He, L. Y., Hwang, G. J., Zhu, X. W., Bian, C. N., & Fu, Q. K. (2023). A concept mapping-based self-regulated learning approach to promoting students' learning achievement and self-regulation in STEM activities. *Interactive Learning Environments*, *31*(10), 7159-7181.
Fan, G., Liu, D., Zhang, R., & Pan, L. (2025). The impact of AI-assisted pair programming on student motivation, programming anxiety, collaborative learning, and programming performance: a comparative study with traditional pair programming and individual approaches. *International Journal of STEM Education*, *12*(1), 16.
Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in human behavior*, *72*, 558-569.
Kukul, V., Gökçearslan, Ş., & Günbatar, M. S. (2017). Computer programming self-efficacy scale (CPSES) for secondary school students: Development, validation and reliability. *Eğitim Teknolojisi Kuram ve Uygulama*, *7*(1), 158-179.
Piccolo, S. R., Denny, P., Luxton-Reilly, A., Payne, S., & Ridge, P. G. (2023). Many bioinformatics programming tasks can be automated with ChatGPT. *arXiv preprint arXiv:2303.13528*.
Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.
Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence, 4,* 100147.

Yu, H. (2023). Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, *14*, 1181712.

家長參與對兒童编程學習影響的研究綜述與展望

Parental Involvement in Children's Programming Learning: A Review and Future Directions

Yitian Zhang[1]，Kunhan Zheng[1, 2]，Biyun Huang[1*]，Fangwen Guo[1]
[1]School of Education, City University of Macau, Macau SAR, China
[2]Guangzhou College of Applied Science and Technology
[*]byhuang@cityu.edu.mo

【摘要】 本文基於 15 篇國外實證研究，系統綜述家長參與對兒童編程學習的影響。研究發現，家長參與能顯著提升兒童的學習興趣、自我效能感與問題解決能力，但效果受家長知識水準、兒童認知特點和文化差異影響。研究同時指出，過度干預可能削弱兒童自主性。基於現有研究局限，本文提出以下展望：（1）追蹤家長參與長期效果；（2）**驗證家長信心與知識水平之交互機製**；（3）**深化兒童認知發展與編程工具適配性**；（4）開展跨文化比較。本研究為提升兒童編程學習效果提供了理論依據與實踐指導，助力兒童在數位化時代全面發展。
【關鍵字】 家長參與; 兒童編程教育; 親子互動;

Abstract: *This study reviewed 15 international empirical studies to explore parental involvement in children's programming learning. The findings indicate that parental involvement significantly enhances children's learning interest, self-efficacy, and problem-solving abilities. However, these effects are moderated by factors such as parental knowledge, children's cognitive development and cultural differences, while excessive intervention may hinder children's autonomy. Future research should focus on the following prospects: (1) tracking long-term effects, (2) verifying the interactions between parental self-efficacy and knowledge, (3) deepening the compatibility between children's cognitive development and programming tools, (4) conducting cross-cultural comparative studies. This research provides theoretical and practical insights for optimizing children's programming learning outcomes, supporting their holistic development in the digital age.*
Keywords: *parental involvement, children programming, parent-child interaction*

1.前言

近年來，人工智慧發展迅猛，未來的生活充滿了各種挑戰與不確定性。在數位化時代，編程能力已成為世界各國培養創新型人才的核心目標之一。2017 年，國務院印發《新一代人工智慧發展規劃》，明確提出在中小學階段推廣編程教育 (国务院, 2017-7-8)，標誌著編程教育從成人向青少年及兒童群體擴展。儘早讓兒童瞭解編程的程序和系統，對培養兒童的計算思維、成就感、應對挫折的能力以及合作交流能力大有裨益，同時能幫助兒童更好的適應未來。

以往關於編程的教學主要聚焦於教師和學生的互動，而對於家長和學生合作學習編程的研究甚少。家長參與作為兒童成長中的關鍵因素，對兒童在各領域的學習和發展有著深遠的影響。探究家長參與如何影響兒童的編程體驗對優化編程教育策略、提升兒童編程學習效果具有重要意義。不少研究表明編程教育不僅培養技術能力，還能促進兒童在情感、價值觀等非智力因素方面的發展 (Schleicher, 2020)。

鑒於國內缺乏相關研究，本文通過文獻綜述法系統梳理了 15 篇國外相關實證研究，旨在探討家長參與對兒童編程能力的影響，包括學習興趣、自我效能感、問題解決能力及情感和社會技能等方面，為未來教育實踐和研究提供理論依據與實踐指導。

2.家長參與兒童編程學習的研究內容分析

2.1 關鍵概念

2.1.1 家長參與

對於家長參與，至今還沒有一個完整統一的定義，但已被認為是改善兒童教育成果的重要因素。Borgonovi and Montt (2012) 將家長參與 （Parental Involvement） 定義為家長積極承諾花時間幫助他們孩子的學業和全面發展。Epstein (1995) 提出了家長參與的不同形式，包括家長參與育兒、與學校溝通、志願參加學校活動、支持在家學習、參與學校決策過程和參與社區資源等。Sylva et al. (2004)針對學前教育的長期跟蹤發現，家長在學生的整個教育過程中的持續參與為學生的成長提供了必要的支持，在學校教育的早期參與尤其有效。因此，本文主要研究家長參與兒童的編程教育，探討親子互動對兒童編程能力的影響，這種共同的學習經歷可以成為家庭共用的寶貴記憶，並可能激發孩子對技術領域的長期興趣。

2.1.2 兒童編程

兒童編程是為兒童設計的編程學習活動，通常採用可視化工具（如 Scratch、ScratchJr）或入門編程語言（如 Logo、Python），幫助兒童掌握基礎編程概念，如順序、條件語句和循環。其根本目的在於培養計算思維，包括分解問題、識別模式、抽象核心概念和設計算法。編程學習不僅能夠增強兒童的邏輯推理和創新思維能力，更能促進其問題解決、團隊協作以及持續學習等重要素養的發展(Papert, 1980)。編程活動可以應用於學校課程、工作坊或家庭學習，本文主要研究家長參與在兒童編程中的影響，幫助兒童在未來的學習和生活中更好地適應技術發展。

2.2 理論基礎

2.2.1 社會認知理論 (Social Cognitive Theory)

班杜拉的社會認知理論強調個體認知、行為和環境的相互作用，其中觀察學習使學習者通過模仿他人行為結果習得新技能，而自我效能感則影響個體行為選擇與努力程度 (Bandura, 2014)。觀察學習 （Observational learning） 是學習者通過觀察他人的行為及其結果來學習的過程 (Bandura et al., 1966)。在編程學習過程中，當家長表現出對編程的積極態度和信心時，孩子更有可能通過觀察學習模仿這種積極行為，從而對編程產生興趣和動力。這種「效能傳遞」在家庭學習環境中尤為顯著，家長的積極反饋能增強兒童的自我效能感知，形成正向學習迴圈。Daniel et al. (2016)的理論拓展了該理論的應用，他們發現，情境因素（如家庭環境編程）和個人因素共同影響學生的學業成績，並探討了兒童自我調節的學習行為在家庭參與與學生早期學業成績之間的作用。這表明社會認知理論在教育領域具有重要的指導意義，特別是在編程教育中，家長的態度和自我效能感對兒童的學習效果有著不可忽視的影響。

2.2.2 建構主義學習理論 (Constructivism Learning Theory)

建構主義學習理論強調，知識的獲取源於學習者與環境的積極互動和主動建構 (何克抗, 1997)。在親子編程活動中，兒童通過邊做邊學的方式，在與家長協作完成具體項目時實現深度學習。OurkidsCode 工作坊通過設置反思階段，引導親子回顧編程中遇到的問題、解決方法及合作感受。家長在此階段幫助兒童重新梳理整合知識，形成個人認知結構 (Lynch et al., 2019)。Papert (2020) 特彆強調，當兒童在真實專案中應用所學並反思過程時，知識建

構最為有效。工作坊設置的反思環節讓家長引導孩子回顧問題解決的過程，這種結構化支持既保障了探索的自主性，又避免 Sadka and Zuckerman (2017)指出的“無效探索”風險。因此需要家長平衡孩子的自主探索和及時指導，確保學習的有效性。

2.2.3 自我決定理論 (Self-Determination Theory)

自我決定理論探討個體在認識到個人需要和環境信息的基礎上，調節自己的行為並推動自身發展的內在動力。該理論認為個體要實現最佳功能，需要滿足自主需求、勝任需求和關係需求這三種基本心理需求 (Lerner & Grolnick, 2019)。自主需求強調個體對自己行為的掌控感，是內在核心動力的體現。在編程教育中，可以通過賦予選擇權、鼓勵探索性學習來支持兒童的需求，比如讓兒童自主選擇編程項目主題，而非強加固定任務; 允許兒童嘗試非傳統方法，即使這些方法可能導致錯誤。勝任需求的滿足源於兒童對自身能力的信心和完成任務的成就感，可以通過分層任務設計、即時反饋機制和錯誤調試支持來強化兒童的勝任需求。關係需求的滿足即兒童感受到與他人的連接，家長可以通過協作式學習、情感支持和社區互動來滿足這一需求。例如，家長與兒童共同參與編程專案，通過角色分工建立協作關係，通過分享作品和互相反饋強化歸屬感。

儘管社會認知理論與自我決定理論均支持家長參與的重要性，但兩者存在潛在衝突，家長通過高自我效能感樹立榜樣時需要避免壓制兒童的自主性，未來研究可探討如何協調這兩種理論。

3.研究設計

為獲取國外近二十年家長參與兒童編程的高品質實證研究文獻，選取 Web of Science 和 Scopus 兩大英文資料庫作為數據來源，確保覆蓋的全面性。文獻檢索時間設定範圍為 2000 年到 2025 年 1 月 31 日，使用資料庫的高級檢索功能，以「主題」為檢索項，以 “programming” “children programming education” “parent-child” “parental support” “parent involvement” 為關鍵詞進行檢索，共獲得 10799 篇文獻。為保證文獻分析結果的準確性和可靠性，基於研究問題初步檢索到的 10799 篇文獻制定了表 1 所示的文獻納入和排除標準。前六項標準是文獻分析中普遍採用的規範，用以保證樣本的品質; 第七項專門用於挑選實驗性實證研究，過濾掉那些雖然涉及大範圍調研但缺乏清晰研究目標、嚴格實驗步驟或合理方法的文章; 第八項則嚴格限定研究主題必須聚焦家長參與，排除那些僅把兒童編程作為單一因素的研究。本研究依據系統性文獻綜述與元分析（PRISMA）框架開展文獻研究(Liberati et al., 2009)，這套國際通用的系統評價方法需要詳細說明文獻的收集、篩選、採納與剔除的過程及理由，以此增強綜述結果的可靠性。基於該研究思路，本研究最終獲得符合條件的論文 15 篇，其中 Scopus 6 篇、Web of Science 9 篇。PRISMA 流程圖如圖一所示。

表 1 文獻納入/排除標准

序號	納入標準	排除標準
1	親子互動	非親子互動
2	實證研究	非實證研究
3	全文可獲取	全文不可獲取
4	文章是期刊論文	書稿、會議論文、報告等
5	文章至少包含三頁	少於三頁的海報、短論文或簡介等
6	文章標題只出現一次	重複性標題

7	研究包括明確的研究問題、研究方法以及研究結論	研究沒有呈現明確的研究問題、研究方法或研究結論
8	研究主題聚焦家長參與兒童編程教育	研究主題非家長參與兒童編程教育（如將兒童編程作為單一變數）

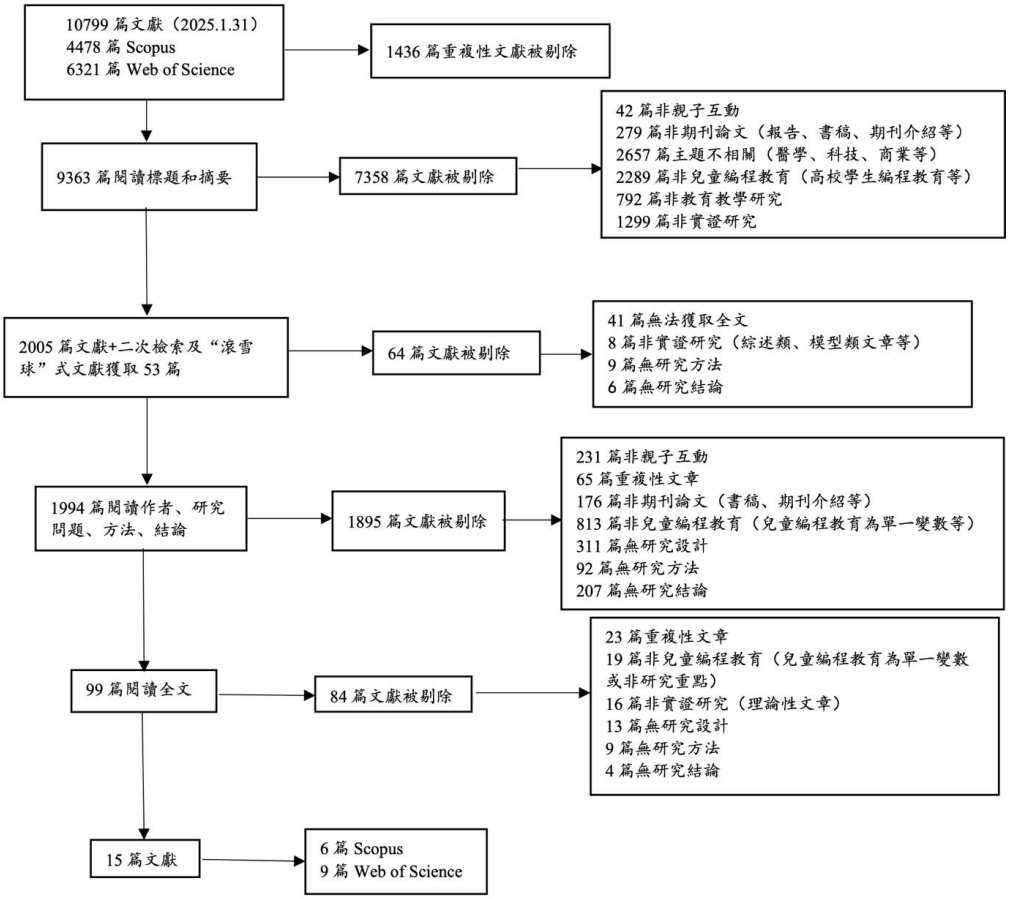


圖 1 PRISMA 流程圖

4.結果與討論

本研究根據對 15 篇國外實證文獻之系統分析，整理出家長參與對兒童編程學習影響的若干面向，涵蓋參與方式、影響因素與文化脈絡等層面。以下將依據分析結果，分項討論家長參與之具體樣態及其對兒童學習歷程與成效所產生之影響。

4.1 家長參與的影響

4.1.1 直接參與

直接參與通常指家長和孩子共同合作編程，輔助孩子完成編程內容。社會認知理論強調家長的示範行為促進兒童通過觀察學習內化編程思維，兒童通過模仿家長的積極態度與問題解決策略，逐步提升對編程的興趣與自我效能感(Bandura, 2014)。此外，家長的情感支持也屬於直接參與的一種，比如在一個面向中學生的編程夏令營中，孩子們將自己在編程營中創建的應用程式帶回家展示給父母，母親通過測試應用功能、討論編程過程和表達贊揚等方式參與孩子的學習，這是一個討論編程和計算機科學職業的媒介，這種共同行為促進了父母與

孩子之間的互動，增強了孩子對編程的興趣和信心 (Clarke-Midura et al., 2019)。然而，家長在直接參與的過程中需要平衡引導與賦權的張力。Lin and Liu (2012)的研究通過量化數據分析、統計驗證與文獻對比，發現親子在編程的學習過程中會自然而然形成一種特殊的形式，即父母和孩子分別扮演「評審者」和「驅動者」的角色。這種模式突破傳統的「家長主導」的單向傳遞，轉而強調雙向認知交換——孩子成為編程的主要推動者，負責程式的編寫與設計，父母從旁給予建議和評價，起到輔助和引導的作用，這種雙向的互動正是知識建構的過程，家長能夠在支持孩子學習的同時，尊重孩子的主體地位，促進孩子在編程學習中的全面發展。

4.1.2 間接參與

間接支持指家長向孩子提供學習編程的資源，包括學習環境、資金支持和軟體支持。家長為孩子報編程的興趣班，購買編程書籍和設備等，為孩子提供物質上的支持。此外，家長通過編程工作坊瞭解了編程的重要性，關注孩子編程的學習，營造支持性的家庭學習氛圍也是一種間接支持 (Clarke-Midura et al., 2019)。

然而，家長參與兒童編程教育雖有很多積極意義，但也存在「雙刃劍」效應。比如東亞家庭過度關注功利性目標（如競賽獲獎），這種傾向可能會削弱孩子對編程本身的興趣。因此，需要調整支持方式以平衡兒童的勝任需求與自主需求，實現兩者之間的平衡發展(孙丹 & 李艳, 2019)。此外， Fessakis et al. (2013)指出，兒童編程的案例多是短期研究，因此對於兒童長期使用編程環境的影響和學習成果的瞭解有限。Maruyama (2019)等多名研究者也發現家長參與對兒童編程能力具有短期積極影響，但長期效果尚不明確。鑒於這些研究的局限性，後續研究可結合學校教育和社區資源組織編程活動，持續提升家長編程知識與技能，觀察家長在長期參與孩子編程學習過程中的行為變化及對孩子的影響。家長參與路徑如圖二所示。

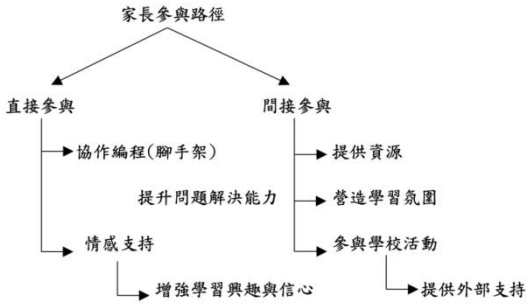


圖 2 家長參與路徑

4.2 影響親子編程合作的關鍵因素

4.2.1 家長之信心與知識水平

家長參與對兒童的教育影響顯著，這一點在眾多研究中得到了證實。但在編程這一專業性較強的領域，大多數家長因缺乏經驗，難以促進孩子的學習。社會認知理論的“效能傳遞”機制解釋了此矛盾：家長通過參與 OurKidsCode 工作坊提升數位技能後，其自我效能感顯著增強，通過行為示範（如協作編程）和積極反饋，推動兒童模仿並內化編程思維。進一步研究表明，家長的觀念與能力直接影響其參與效果(Lynch et al., 2019)。例如，香港學者 Kong et al. (2019)基於中國香港地區 524 對親子的調查數據，開發了首個家長編程教育觀念評估量表，發現家長對編程價值的認知與其支持策略顯著相關。此外，Bresnihan et al. (2021) 進一

步開發調查工具 PICS (Parental Involvement in Computer Science Survey Instrument)，通過量化分析揭示了家長的技術信心與編程經驗在兒童編程教育中的關鍵作用。

父母自身的編程知識水準也直接影響他們對兒童編程學習的效果。研究表明，高素養家長能通過結構化指導（如分解任務）提供說明，符合觀察學習中「榜樣示範」的要求; 而低素養家長可能因焦慮或過度干預，反而抑制兒童自主探索 (Maruyama, 2019)。同時，Clarke-Midura et al. (2019)的實證研究數據顯示家長信心和知識水平分別對兒童編程學習效果影響顯著，但兩個變量之間可能存在交互作用。為理清這一機制，未來研究可通過量化分析設計實驗，對比不同信心與知識組合的干預模式，系統測量兒童編程能力、自主性及長期興趣的變化。

4.2.2 兒童認知發展之特點

兒童編程學習的主要特點有具象性、遊戲性和遷移性，這些特點不僅反映了兒童認知發展的規律，也為家庭編程教育提供了重要的實踐指導 (王琳 et al., 2021)。首先，具象性是兒童編程的核心特徵之一，兒童的認知發展以具體形象思維為主，因此編程工具和學習內容需要具體化、可視化。比如 Scratch 等圖形化編程工具通過拖拽積木塊式的代碼塊來創建程式，將抽象的編程邏輯轉化為直觀的操作，體現了“做中學”的建構主義原則。其次，遊戲性指兒童的編程學習以遊戲貫穿，從角色定位到情境創設，兒童既是學習者又是遊戲角色，學習與遊戲過程深度融合。這種遊戲化學習不僅增強了兒童的學習興趣，還培養了他們的創造力和問題解決能力。第三，遷移性體現兒童編程學習的長期價值，建構主義重視知識的可遷移性與跨領域聯結。父母應當協助孩子將編程思維與跨學科內容有機融合，以此強化知識遷移的實際效果，提升兒童的認知、記憶、創造力和問題解決等能力(Fessakis et al., 2013)。基於建構主義的遷移性原則，未來可開發跨學科任務，探討家長如何引導兒童將編程思維遷移到其他學科領域（如語文、科學），提升兒童的綜合素養(Bers et al., 2014)。

4.2.3 文化與社會背景之差異

不同文化背景下家長參與編程協作的路徑存在顯著差異。以教育方向來說，西方國家較早將編程融入中小學教育，重視創造力的培養; 亞洲地區則更關注實際應用和考試加分，這種區別在跨國研究中很明顯。在非英語地區，語言和技術工具需要調整適應本地需求，西班牙及日本案例驗證方言適配能有效提升家長參與度 (Bers et al., 2019)。家庭分工方面，亞洲家庭里母親常負責督促孩子學習，歐美家庭則更傾向父母與孩子平等合作。社會觀念也影響技術接受度，中東地區對女性學程式限制較多，北歐則較開放，研究指出需要專門解決性別偏見問題 (Angeli & Valanides, 2020)。合作方式也有地區特色：愛爾蘭透過社區活動加強家庭互動，東亞多由學校主導，這顯示出文化中集體合作與個人探索的不同傾向 (Lynch et al., 2019)。未來需要設計彈性化策略，結合本地特色和國際經驗，同時打造多語言技術工具，讓不同背景的家庭都能更好地參與協作。

4.3. 研究省思與未來發展建議

經過梳理家長參與兒童編程的相關研究發現，國內研究在兒童編程研究領域內較少涉及家長參與的部分，國外則已積累較多家庭參與兒童編程的案例。基於現有研究的局限性，未來建議從以下方面深化探索：第一，進一步跟蹤家長參與對兒童編程影響之長期效果；第二，驗證家長信心與知識水平之交互機製；第三，深化兒童認知發展與編程工具之適配性研究；最後，開展跨文化比較研究。總之，家長參與在兒童編程教育中至關重要，需通過家庭、學校與社會的協同努力，構建更加包容和可持續的支持體系，助力兒童在數位化時代全面發展。

5.結論

本研究採取文獻綜述法，基於社會認知理論、建構主義學習理論與自我決定理論的動態交互共同構建了家長參與影響兒童編程學習的理論框架。研究發現，家長通過直接參與，即行為示範（如協作編程）和間接參與（如提供資源支持），能顯著提升兒童的學習興趣、自我效能感及問題解決能力。然而，這種影響存在邊界條件，社會認知的示範效應需要與自我決定的自主支持平衡，正如 Bresnihan et al. (2021) 指出，若家長的干預方式是控制型的或對孩子的潛力持負面看法，可能會對孩子的教育成果產生負面影響; 建構主義的協作探索需要家長根據兒童認知水平調整腳手架 (Scaffolding) 策略，這為理解家長參與的“雙刃劍”效應提供了新視角(Vygotsky, 1978)。當三種機制共同作用時，能最大化編程教育效果; 若偏重單一理論，比如僅強調社會認知而忽視自主性，則可能導致負面效果。研究同時揭示，兒童編程效果受家長信心與知識水平、兒童認知發展階段及文化背景的綜合影響，需在保障兒童自主性的前提下，平衡三者以優化家長參與的積極效應，實現數字時代的全人發展。

參考文獻

国务院. (2017 年 7 月 8 日). *国务院关于印发新一代人工智能发展规划的通知*. https://www.gov.cn/gongbao/content/2017/content_5216427.htm

何克抗 (1997)。建构主义的教学模式, 教学方法与教学设计. *北京师范大学学报: 社会科学版*, *5*(8), 100-120.

孙丹、李艳 (2019)。国内外青少年编程教育的发展现状、研究热点及启示——兼论智能时代我国编程教育的实施策略. *远程教育杂志*, *37*(03), 47-60. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2019.03.005

王琳、耿凤基、李艳 (2021)。编程学习与儿童认知发展关系的探讨. *应用心理学*, *27*(03), 262-271. https://kns.cnki.net/kcms2/article/abstract?v=4fayqqv3WFc1c8NY7vATGjVK-7ej3NyQ5Wc6tKBKtMs8iTBOmSJpEpPuzOuBUYZQXN42_I-zMFZk4gvvcAkgNbXuol05FK7rZdm13N2cBvIfAAPdxp1MIjVapZW61CDrgz7iEw6nCsnlT27qb2Zdlazx-xU-EN724TTQhIiL5aFGikENHOoEIWR_UEuPNvrrxUodKS6wzv4=&uniplatform=NZKPT&language=CHS

Angeli, C., & Valanides, N. (2020). Developing young children's computational thinking with educational robotics: An interaction effect between gender and scaffolding strategy. *Computers in human behavior*, *105*, 105954.

Bandura, A. (2014). Social cognitive theory of moral thought and action. In *Handbook of moral behavior and development* (pp. 45-103). Psychology Press.

Bandura, A., Grusec, J. E., & Menlove, F. L. (1966). Observational learning as a function of symbolization and incentive set. *Child development*, 499-506.

Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers & Education*, *72*, 145-157. https://doi.org/10.1016/j.compedu.2013.10.020

Bers, M. U., González-González, C., & Armas–Torres, M. B. (2019). Coding as a playground: Promoting positive learning experiences in childhood classrooms. *Computers & Education*, *138*, 130-145.

Borgonovi, F., & Montt, G. (2012). Parental involvement in selected PISA countries and economies.

Bresnihan, N., Bray, A., Fisher, L., Strong, G., Millwood, R., & Tangney, B. (2021). Parental Involvement in Computer Science Education and Computing Attitudes and Behaviours in the Home: Model and Scale Development. *ACM Transactions on Computing Education*, *21*(3), 1-24. https://doi.org/10.1145/3440890

Clarke-Midura, J., Sun, C., Pantic, K., Poole, F., & Allan, V. (2019). Using Informed Design in Informal Computer Science Programs to Increase Youths' Interest, Self-efficacy, and Perceptions of Parental Support. *ACM Transactions on Computing Education*, *19*(4), 1-24. https://doi.org/10.1145/3319445

Daniel, G. R., Wang, C., & Berthelsen, D. (2016). Early school-based parent involvement, children's self-regulated learning and academic achievement: An Australian longitudinal study. *Early Childhood Research Quarterly*, *36*, 168-177. https://doi.org/10.1016/j.ecresq.2015.12.016

Epstein, J. L. (1995). School/family/community partnerships. *Phi delta kappan*, *76*(9), 701.

Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study. *Computers & Education*, *63*, 87-97.

Kong, S.-C., Li, R. K.-Y., & Kwok, R. C.-W. (2019). Measuring parents' perceptions of programming education in P-12 schools: Scale development and validation. *Journal of Educational Computing Research*, *57*(5), 1260-1280.

Lerner, R. E., & Grolnick, W. S. (2019). Maternal involvement and children's academic motivation and achievement: The roles of maternal autonomy support and children's affect. *Motivation and Emotion*, *44*(3), 373-388. https://doi.org/10.1007/s11031-019-09813-6

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj*, *339*.

Lin, J. M.-C., & Liu, S.-F. (2012). An investigation into parent-child collaboration in learning computer programming. *Journal of Educational Technology & Society*, *15*(1), 162-173.

Lynch, Á., Millwood, R., Fisher, L., Strong, G., & Bresnihan, N. (2019). *OurKidsCode: Facilitating Families to Be Creative with Computing* Proceedings of the 11th International Conference on Computer Supported Education,

Maruyama, Y. (2019). An Investigation into the Effects of Programming Workshop Experiences on Parents' Concerns about Programming Education in Elementary School. *International Association for Development of the Information Society*.

Papert, S. (1980). *Children, computers, and powerful ideas* (Vol. 10). Harvester Eugene, OR, USA.

Papert, S. A. (2020). *Mindstorms: Children, computers, and powerful ideas*. Basic books.

Sadka, O., & Zuckerman, O. (2017). *From Parents to Mentors* Proceedings of the 2017 Conference on Interaction Design and Children,

Schleicher, A. (2020). 教育面向学生的未来, 而不是我们的过去. *华东师范大学学报 (教育科学版)*, *38*(5), 1.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., Smees, R., & Morahan, M. (2004). The effective provision of pre-school education (EPPE) project. *Children*, *29*(30), 31.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.

未来教育视域下“智慧•玩创”学习空间新样态的构建与实施

Construction and Implementation of the New Model of "Smart Play and Creation" Learning Spaces from the Perspective of Future Education

张 波[1] 王云燕[1*]
[1]浙江省杭州市滨江区钱塘山水幼儿园
[*]W_yunyan5991@163.com

【摘要】 未来教育视域下幼儿园育人空间亟待转型。构建“智慧·玩创”学习空间新样态，从空间构建、运行、评价三方面出发，遵循融合、个性、友好原则，致力培养幼儿自主学习能力为主旨，培养适应未来的核心素养为目标，打造智慧教育赋能、多元个性选择、儿童友好交互的学习空间。在空间中以自主驱动促进了幼儿学习的积极性，以亲身体验加强了幼儿与世界的深度连接，以空间交互激发了幼儿与同伴的积极交往，以数据循证扎实了幼儿发展的学习过程性，从而支持了幼儿的全面发展。
【关键词】 未来教育视域；“智慧·玩创”学习空间

Abstract: *In the context of future education, the educational spaces in kindergartens urgently need transformation. This paper explores the construction of a new model for "Smart Play and Creation" learning spaces, focusing on three aspects: space design, operation, and evaluation. Adhering to the principles of integration, personalization, and child-friendliness, the initiative aims to cultivate children's autonomous learning abilities and develop the core competencies required for the future. These spaces leverage smart education technologies, provide diverse personalized choices, and promote child-friendly interactions. Through self-directed learning, children's motivation is enhanced; hands-on experiences strengthen their connection to the world; interactive spaces stimulate positive peer communication; and evidence-based data supports and solidifies the learning process, ultimately facilitating children's overall development.*
Keywords: Future education perspective, "Smart Play and Creation" learning space

1.前言

未来教育是面向未来的教育。顾明远先生对未来教育的观点，即未来教育不是指若干年之后的教育，而是当下正在进行着的、为未来社会培养人才的教育。[i]未来是未发生的历史，未来是可预知的现在，未来教育是对美好教育的呼唤[ii]。近年来，教育研究人员一直在积极研究智能学习环境。截至 2021 年 3 月，谷歌学者和 Scopus 对关键词“智能学习环境”进行搜索，其中 80%以上的结果在 5 年内公布，几乎所有结果都涉及高等教育环境，包括高等教育、继续教育和开放教育。[iii]本研究中的未来教育，从培养目标上是未来社会所需要的人才为目标，儿童通过自适性学习、支持全面和个性发展。从培养方式上指向多元化、个性化、智能化、终身化的方式，从培养路径上是跨学科、跨时空、跨文化的学习路径；构成虚实交融的未来学习场景。[iv]然而，研究的“智慧·玩创”学习空间载体正式预期的未来学习场景，转变传统的育人观，是基于未来教育视域下幼儿园的学习空间新样态。以多元智能·丰盈素雅、包容开放·支持个性、儿童友好·情感交融为设计理念，兼顾学理基础与实践运行，凸

显空间特质与核心素养培养。根据园所资源构建的“科学探究”、“智慧阅读”、“艺术创想”、“友好生活”、“体感运动”、“自然探秘”六大“智慧•玩创”学习空间。各类空间均有指向儿童发展的核心素养，每个空间既相对独立又有一定交互性。其学习空间开放、自由、丰富、多元、智慧，符合幼儿自主学习发展的多元学习场景样态。各空间供幼儿自主选择、驱动幼儿自主学习为目的；通过空间的学习可以将幼儿自我认知与社会产生广泛联系，与幼儿、教师、各类智能游戏产生深度链接，建构其学习经验；同时，可以培养幼儿积极的社会交往能力；其所有幼儿的空间学习可通过大数据记录，利用数据询证支持幼儿的学习与发展。从构建与运行两方面进行空间学习场景、学习模式、学教方式、学习评价的新样态打造，创生致力于幼儿核心素养培养的空间学习育人样态。

基于“智慧•玩创”学习空间的特质，引导幼儿以多种身份参与“玩•创”，探索学习空间内容创建和幼儿学习方式，基于空间特质不同，幼儿运用多元的学习方式，通过IE双驱式学习、五感式学习、QE对话式学习、探究式等学习方式培养幼儿的综合素养，在实践中发展幼儿未来所需的关键能力，打造幼儿通向未来学习方式的空间。所以，本研究立足于创建“智慧•玩创”学习空间，探索未来教育视域下指向幼儿全面发展及个性成长的学习场变革，是必要且具有价值的。同时，为其他幼儿园提供通向未来教育之路的一种新范式。

2.未来教育视域下“智慧•玩创”学习空间的创建

随着学习理论和现代技术的发展，学习空间的内涵发生了很大的转变，翻转课堂、智慧教育、人工智能等逐渐取代了传统单一的学习环境。[v][vi]“智慧•玩创”学习空间是基于未来教育视域下幼儿园的学习空间新样态。创建原则坚持融合性、个性化、友好化的原则，通过数字大脑、智慧手环、虚拟现实等技术与空间融合，创设“数字大脑”智能学习系统，打造交互平台，形成智能评估反馈；为幼儿构建智能化、个性化的学习环境，通过空间融合提升幼儿的多元体验，构建幼儿智慧赋能的学习场。学习空间充分考虑孩子的个性和兴趣，为其提供充满探索机会和创造可能的环境。幼儿园的每一处空间、每一个活动场都应能与幼儿“对话”，满足幼儿自由创造的需求；满足幼儿个性化发展的需求，让空间成为幼儿体验探索的玩创园。空间应创设良好的互动氛围，建立积极的支持关系，让幼儿与空间产生有意义的对话，与同伴、教师产生友好互助的合作交流。[vii]空间创设尽可能满足幼儿需求、支持幼儿发起的活动、拓展学习并促进幼儿情绪情感的发展，让空间成为幼儿友好交互的理想国。

空间的创设通过四大步骤，玩家听证：发现儿童心声 了解学习需求；家园参访：调查规划空间，体现多元个性；山水邀约：交互创设空间，融合素养指向；现场循证: 优化运行空间，发展学习经验。期中，玩家听证是指向幼儿发起调查了解孩子的学习需求，采纳幼儿的创意，为园所空间维度与类型的创设奠定童声基础。家园参访是指园所组织家长和幼儿实地参访几大空间类型和整体空间立体模型，融合家长空间设计，保障儿童空间友好权力。山水邀约是指有兴趣的幼儿及家长共同参与环境创设，体现个性多元，主人翁意识。现场循证是指以循证的理论支撑为视角，以空间运行的实践现场为素材，以发展幼儿学习经验为目标，带领幼儿、教师、管理者等共同参与学习空间的自我评价与同伴互评，在运行中不断调整改进学习空间的空间布局、材料投放、幼儿学习方式等。

基于“完整儿童”发展的理念，结合《3-6岁儿童学习与发展与指南》的五大领域，重点培养幼儿自主学习能力同时关注幼儿的全面发展。[viii]所以，“智慧•玩创”学习空间结合室内与户外两大维度分别构建室内场景式学习空间和户外生态式学习空间。从空间模型建构、发展指向、具体空间构成方面构建六大类型学习空间。“智慧•玩创”六大学习空间在智慧

赋能、个性多元、儿童友好的构建原则支持下，六大空间聚焦科学、语言、艺术、社会、生活、健康各领域发展，每一类型学习空间的发展指向与幼儿各领域的学习均有所关联。[ix]

2.1 科学探究空间：指向幼儿的科学素养

幼儿在该空间学习的重要目标为：通过多样化内容和形式丰富科学认知、了解科学现象、学习科学探究方法；聚焦编程、木工、建构等活动，促进创造力、探究力、工程建构思维力发展；愿意主动与同伴、教师分享科学发现，遇到问题积极探索并尝试解决；在亲身操作和探究中感受科学现象的神奇、科学活动的乐趣。该空间有三大特点：智慧 AR 赋能、STEAM 探究体验、科学创造表达。

2.2 智慧阅读空间：指向幼儿的语言素养

幼儿在该空间学习的重要目标为：在温馨智慧的阅读空间中体验自主阅读、合作阅读、听说读写等多种阅读方式；学习正确的书籍阅读方法，发展阅读理解能力，锻炼大胆自信地完整表达；喜欢听故事，看图书；感受到自信大方表达的乐趣。该空间有三大特点：智慧点读融合；自主多元表达；智慧演绎衔接。

2.3 艺术创想空间：指向幼儿的艺术素养

幼儿在该空间学习的重要目标为：借助 AR 智慧沉浸设备等感受美、发现美、欣赏多样的艺术形式和作品；运用多种类型的材料丰富艺术创造体验和方式，促进艺术创造力与表达表现力；喜欢进行综合艺术活动并大胆表现自己，体验艺术创想的乐趣。该空间有三大特点：灵活创设平台；多元创意表现；家园社交互。

2.4 友好生活空间：指向幼儿的社会素养

幼儿在该空间学习的重要目标为：在多样化的生活材料操作中感受生活的气息和传统文化的魅力；在动手制作、品尝体验中，锻炼生活自理能力，体验生活的乐趣；理解关心他人、学会自我放松、沟通交流、分工合作方法，体验与人交往乐趣。该空间有三大特点：连结生活场景；多重角色体验；友好情绪浸润。

2.5 体感运动空间：指向幼儿的健康素养

幼儿在该空间学习的重要目标为：科技与运动智慧融合，激发运动兴趣，促进运动探索和坚持性；锻炼上下肢体协调、大肌肉动作等技能，具有一定的力量和耐力；乐于与同伴一起运动，相互协作挑战运动，享受运动的快乐。该空间有三大特点：多维空间运动；智能数据评价；合作竞赛交互。

2.6 自然探秘空间：指向幼儿的综合素养

幼儿在该空间学习的重要目标为：亲近和了解自然，萌发热爱、尊重和保护自然的意识和情感；聚焦自然探究，学习认知事物，在动手动脑中学习解决问题；和同伴开展自然游戏，交流沟通游戏规则，相互合作；大胆并积极地表达自己的需求，通过多元形式表征自己的发现。该空间有三大特点：真实环境营造；自由开放体验；多元综合探究。

3.未来教育视域下“智慧·玩创”学习空间的运行

在泛技术的支持下，高质量的学习空间越来越强调内部要素之间的互动，且这种互动是人、技术、环境、思维四要素相互作用的多维互动，[x]是促进学生经验生成以及教师专业发展的重要中介。[xi]现代技术的发展重构了幼儿的学习空间，它不仅转变了幼儿的学习方式，还转变了幼儿的思维方式。成人须重新审视这种空间的转换对幼儿学习的意义，通过重构幼儿学习空间及其学习方式来促进幼儿经验的发展进一步走向多元、深度和融合。[xii]

第一，“智慧·玩创”学习空间安排中，学习空间活动融入幼儿一日活动安排表中，并与班级教师明确具体开展要点，每周安排班级进行空间探索活动、中大班进行年段空间联动；将户外自由游戏与自然探秘空间和玩伴游戏空间活动融合，调整到下午时段，延长单次活动时间，给予支持幼儿更长时间游戏、探究；对于学习空间的使用分为三类形式：悦玩模式、联玩模式、随玩模式。悦玩模式：定点定时，一龄一班。幼儿定点空间，因学循序渐进；教师空间跟进，因学分组开展。 联玩模式：定点教师，跨龄融班。幼儿自选空间，因趣个性学习；教师定点空间，因需专业支持。随玩模式：时空不限，自选自适。空间一直处于敞开式开放状态。幼儿可在一日中自由时间进入学习空间自由活动。

第二，“智慧·玩创”学习空间中，幼儿采取 IE 双驱式、五感式、QE 对话式、探究式多样的学习方式。**IE 双驱式学习**是智能技术与情感浸润并驱。其核心理念在于将智能技术与情感教育相结合，形成一种双轮驱动的学习机制。这种模式不仅注重知识的传授和技能的培养，更重视情感的培养和个性的发展。“智慧·玩创”学习空间致力于构建一个集智能化、情感化于一体的新型学习共同体。幼儿能够在智能技术的辅助下，通过情感的引导，实现知识的增长和情感的丰富。通过“链接经验，唤起情感——智能运用，操作学习——解决难点，体验成功——升华情感，分享展示”路径，让空间学习有技术支持又有情感温度。如大班大班科学探究空间活动：大颗粒积木编程《升国旗》案例，教师首先从幼儿熟知的国旗经验导入，唤醒幼儿的爱国情感，激发参与活动的兴趣；其次使幼儿运用智能技术，在亲身体验和实际操作大颗粒和编程中创意建构、自主搭建；再次，通过讨论交流解决技术难点；最后展示结果，体验成就感的同时呼应爱国的情感主题，达到 IE 双驱式学习的有效成果。

五感式学习是一种全面调动幼儿感官参与的立体式学习法，它强调在游戏和生活情境中，通过视觉、听觉、触觉、味觉和嗅觉的协调运用，来激发幼儿的好奇心和探究能力。这种学习方式不仅能够提高幼儿的问题解决能力，还能培养他们对世界整体的感知和理解。

QE 对话式学习是一种以问题为驱动，以工程思维为导向的新型学习方式，它在"智慧·玩创"学习空间中得到了充分的应用。这种学习方式通过一系列精心设计的学习路径，鼓励幼儿在小组合作中探究真实问题，运用跨学科的知识和技能来创造性地解决问题，从而助推幼儿工程思维和创造思维的发展。[xiii]如：中班科学探究空间学习活动：自动浇水器中，靶向定位+设计分析：教师定位设计自动浇水器这一驱动性问题。鼓励幼儿通过绘画和搭建模型来表达自己的设计想法，引导他们分析方案的可行性和创新点。协同制作+测试更迭：幼儿在教师和同伴的协助下收集材料，制作浇水器并测试，观察其浇水效果，发现问题并调整，最后展示并分享自己的设计思路。通过交流幼儿能够从其他小组那里获得新的灵感和改进建议。

探究式学习是一种以实验操作和探究循环为基础的学习方式，它鼓励幼儿通过亲身实践和反复探索来获取经验。幼儿在“智慧·玩创”学习空间中提出猜想与假设，进行比较、观察和记录，归纳分析实验结果等，使自身获得的经验经历多次“循证”，更具深度。

第三，“智慧·玩创”学习空间运行中教师支持策略采用定学·助学·共学三种形式。期中个性化定学是定制“基于空间个性”的学期活动计划。每学期初，学习空间的专任教师会和课题研修小组基于空间个性，共同设计、商讨、制定本学期空间活动计划。融入本学期课程主题、班本化项目内容以及节日活动等，进行具体活动的统筹安排与规划，进行材料统整与支架准备。定制“基于幼儿个性”的活动观察方式。 教师会基于不同幼儿的学习特点、发展水平，定制独属于幼儿个人的个案记录。如对于善于表达表现的幼儿通过“现场倾听与支

持”进行对话式的观察；对于喜欢独处安静学习的幼儿通过“观察记录表”进行白描式的观察记录；对于乐于创造喜欢挑战的幼儿通过“智能语音笔”进行游戏式的智能观察。

三单式助学，为助力幼儿在六大学习空间中的自主驱动学习，本研究设计了 Why 好奇单、1+X 思考单、OK 成功单三种支持幼儿深度学习的支架，分别聚焦在发现幼儿学习兴趣、链接幼儿学习经验和丰盈幼儿学习体验。Why 好奇单主要是设计用于发现了解幼儿空间学习兴趣度、兴趣点和兴趣原因等。基于幼儿空间学习兴趣导向的三层深入式好奇，从空间好奇—活动好奇—问题好奇，逐层深入了解，维护幼儿好奇心，引导幼儿带着好奇持续性展开学习。要促进幼儿经验的深入发展，学习空间的建构就必须充分考虑各种学习影响因素的回应性，即当幼儿生发出新的兴趣点之后，学习空间要能有效支持幼儿的这一兴趣。只有将对物质的体验转换为精神的成长和意义的生成，幼儿的经验才有可能被称为深度发展的经验。[xiv]

“1+X”思考单，引导幼儿记录问题，辅助支持学习经验的链接与内化。**“1”**是指“问题式”思考单这一基础性探究载体，**“X”**代表对应不同经验问题类型下的思考单，包括“递进式”思考单、“拓展式”思考单和“整合式”思考单。幼儿选取单项或多项思考单灵活组合使用，进行思考观察、实验记录等，实现探究中“经验链”的形成，支持幼儿从浅层学习走向深度学习。

OK 成功单，一方面丰盈幼儿学习体验，增加幼儿学习探究成就感和自信心；一方面分析其游戏性、发展性，针对性地调整空间区域，动态追随幼儿发展。

携手式共学，教师通过与幼儿共同携手邀约活动、幼儿面对面对镜分享学习的方式携手式共学，从资源、活动、反思等方面共同促进幼儿在空间中的学习发展。

师幼携手邀约：贯彻儿童友好理念，教师与幼儿相互发起活动邀约，共同生发活动中的学习经验。一方面，教师基于幼儿好奇与兴趣进行下一次活动的邀约，通过师幼团讨、角色互换、空间对话、预约下次活动。另一方面，幼儿基于自己的探究兴趣主动与教师、同伴发起活动邀请，约定空间活动的内容，激发空间学习的内驱力。

幼幼携手对镜：镜面式的教学方式不止体现在教师与幼儿之间，更是融合在幼儿与幼儿的相互学习之中，为幼儿提供充分的同伴携手学习时空。教师通过引导幼儿通过大带小、小问大、同伴互夸、大家来发现等形式进行幼儿间的经验互学，发现身边同伴的闪光点，为同伴提出小建议，对自己提出下一次的加油目标等，幼儿在与同伴携手中，互相对镜，互相学习，互相成长。

4.未来教育视域下“智慧·玩创”学习空间的评价

幼儿园“数智+”评价体系及工具运用，对幼儿学习空间中的学习发展进行数智+评价变革。一方面，具体优化幼儿学习发展、教师支持、空间优化；[xv]一方面在运用中优化调试“数智+”评价工具及体系建设。

首先，评价技术革新：依托数智技术，采集空间学习数据。借助新技术构建园所“数智大脑”学习平台。通过摄像头、智慧手环等设备，将六大类学习空间环境、活动及幼儿学习表现进行家园线上线下平台打通，教师、家长、园方管理者能实时清楚地了解幼儿游戏、学习情况。通过刷环打卡，实时追踪定位空间活动频次及幼儿活动状态。通过后台大数据统计分析，发现幼儿自选空间活动的次数、兴趣偏好、活动持续时间等。利用视频魔方，了解幼儿空间活动中的专注性、持久性，优化空间配置、内容设计、评价方式。

其次，评价指标研新：优化数智内核，研制学习素养指标。围绕幼儿在空间活动中的“沟通表达、团队协作、解决问题、批判思维、创新意识”发展点，形成《幼儿空间学习的素养发展评定表》，优化数智＋评价的质性评价内核，[xvi]为教师评价提供易上手可操作的工具，也为接下来学习空间活动优化提供方向。

再者，评价载体融新：充实师幼评价，呈现幼儿学习图景。设计与运用“教师版空间档案”和“幼儿版空间档案”。教师版“空间档案”包括活动设计调整、活动观察与评价以及活动反思与优化三部分。幼儿空间档案包括空间回顾档案和空间学习档案。通过活动图记促进个性化支持，作品卡片优化互动性反馈。将教师、幼儿带入参与式评价体验中，实现对幼儿个性发展的有效支持。

最后，评价分析创新：可视空间学习，呈现幼儿数字画像。通过数智大脑 AI 剖析，形成幼儿空间学习素养发展画像。以雷达图呈现各能力发展是否均衡、长处与短板。以折线图呈现幼儿前后学习发展对比，明确幼儿近阶段成长的能力素养与下阶段的重点关注。通过 AI 检索，在线自动生发对应的支持策略和发展建议，分享给教师和家长关注优化。

参考文献

王伟宾, 张萌萌, 丁邦平. (2021). 试论未来教育的特点、理念及评价内容. 教学管理, (12), 20-22.

任莉. (2023). 学习空间重构：引导幼儿经验发展走向多元、深度与融合. 学前教育研究, (03), 91-93.

李季湄, 冯晓霞 (编). (2013). 3-6 岁儿童学习与发展指南解读. 人民教育出版社.

刘颖棣, 卢越. (2022). 国内学习空间文献综述——基于 CiteSpace 软件的可视化分析. 中小学电教, (03), 15-20.

陈卫东, 叶新东, 秦嘉悦, 等. (2011). 未来课堂——高互动学习空间. 中国电化教育, (08), 6-13.

徐梦影 (记录), 陈兴中, 向宴平 (整理). (2020). 未来教育的本质、目的、任务与发展趋势. 教育科学论坛, (12), 5-15.

杨雄, 杨锐. (2022). 基于 AR 技术的幼儿园课程评价路径. 学前教育研究, (05), 91-94.

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

姜强, 赵蔚, 王朋娇, 等. (2015). 基于大数据的个性化自适应在线学习分析模型及实现. 中国电化教育, (1), 85-92.

陶金玲. (2015). 儿童经验生成研究. 南通大学学报（社会科学版）, (11), 145-150.

温剑青. (2019). 发现 理解 支持——指向个性化教育支持的幼儿发展评价实践. 上海教育出版社.

虞易来. (2023). 一馆多玩：未来教育视角下幼儿园空间重构初探. 现代教育, (4), 16-19.

未来教育的本质、目的、任务与发展趋势. 教育科学论坛, 2020 年 12 期, 第 5-15 页.

Broderick, J. T., & Hong, S. B. (2023). 从儿童的兴趣到思维 [From children's interests to thoughts]. 中国轻工业出版社.

Cheung, S. K. S., Lam, F. K., Phusavat, K., & Yang, H. H. (2021). Shaping the future learning environments with smart elements: Challenges and opportunities. International Journal of Educational Technology in Higher Education, 18, 1-9.

数字驱动，智启未来——可持续发展理念下的幼儿园数智+评价创新实践

Data-Driven Enlightenment: Innovation of Intelligent Assessment in Preschool Education within Sustainable Development Framework

张波[1]，罗怡雯[1*]，黄璇[1]
1 杭州市滨江区钱塘山水幼儿园

【摘要】 在数字化转型与教育可持续发展深度融合的背景下，幼儿园教育评价亟需突破传统模式的桎梏。本文以杭州市Q园的数字化评价实践为例，通过整体构建“Q园大脑”智能系统、研发设计“数智+”评价体系，展开“数字化处理”和“嵌入式跟进”的创新实践。依托智慧手环、AI小智、视频魔方等工具载体，动态追踪生活、运动、游戏、学习全场景数据。同时通过“指标嵌入-智能分析-画像呈现-家园协同”嵌入式跟进评价，实现幼儿学习过程的全场景追踪、多主体参与及数据驱动的循证支持。实践表明，数智+评价不仅提升了教育评价的科学性与精准性，更通过动态反馈与生态化支持，实现从经验驱动到数据驱动的评价范式转型，推动幼儿的个性成长。

【关键词】 可持续发展1；数智+评价2；创新实践3

Abstract: *Amid digital transformation and educational sustainability integration, kindergarten evaluation demands paradigm innovation. Examining Hangzhou Q Kindergarten's "Q-Yuan Brain" system and "Digital Intelligence+" framework, this study implements digital processing and embedded evaluation using smart wristbands, AI tools, and multi-scenario tracking. The embedded process—criteria integration, intelligent analysis, profiling, and home-school collaboration—enables comprehensive learning monitoring and data-driven support. Results demonstrate enhanced assessment precision and a paradigm shift from experience-based to data-driven evaluation, promoting individualized child development through dynamic feedback.*

Keywords:Sustainable Development; Digital Intelligence+ Evaluation; Innovative Practices

1.前言

《2030年可持续发展议程》中强调教育应促进“全纳、公平、优质”，其核心在于：终身发展、生态协同、数据驱动。关注幼儿的全面、持续发展与评价，是学前教育工作者面临的重要任务。而随着“数字中国”战略的提出，大力发展数字教育已成为时代的趋势。幼儿园作为幼儿启蒙教育的场所，也在积极探索如何利用数字和大数据技术来创新幼儿学习与发展评价模式。Q园紧跟时代步伐，结合区域优势，积极开展“数智+”大数据评价实践，旨在通过现代化、数字化的评价方式，更全面、准确地了解幼儿的学习与发展状况，为其提供更加个性化、针对性的教育支持，以适应高质量园所发展的需求。

2.背景基础：立足园本的评价现状

2.1. 玩+课程的高质量发展求索

作为浙江省现代化幼儿园和数字教育创新学校，Q园主动拥抱数字教育，以人工智能时代为背景，以培养幼儿面向未来终身受用的品质为目标，营造促进技术融入的教育新生态。在“互联网+教育”背景下，让智慧教育赋能园所的“玩+”课程荣获浙江省精品课程，以丰富而有层次的多元活动为内容，积极融合、大胆创新。玩+课程发展五年至今，需要与时俱进，将数字技术赋能教育实践，在原有的评价基础上形成系统性评价，展开时代变迁下的迭代更新。

2.2. 幼儿发展的个性化成长需要

每个幼儿都是独一无二的个体，拥有不同的兴趣、能力和经验。在关注幼儿全面发展的同时，个性化成长需求日益凸显。因此，评价体系必须能够捕捉到这些差异，为每位幼儿提供量身定制的教育支持。应注重实施过程性评价，重视其成长发展轨迹；制定基于幼儿成长的多维度评价指标，覆盖幼儿发展的各个方面；采用形成性评价，避免单纯的横向比较，注重纵向发展。

2.3. 现有评价的局限性困点突破

反观以往的幼儿学习与发展评价，存在问题如下：碎片化与主观性：评价内容局限于知识掌握或行为表现，缺乏对学习过程的持续性追踪，教师依赖经验判断，科学性不足。数据缺失与反馈滞后：纸质档案难以记录动态过程，评价反馈滞后于幼儿发展需求，无法支持个性化教育决策。参与主体单一：家长、幼儿的参与度低，评价成为教师的“独角戏”，难以形成教育合力。生态性不足：评价未融入幼儿的生活场景，割裂了幼儿园、家庭与社区的协同作用，难以支撑可持续发展目标。

综上，将数字技术融入幼儿学习与发展评价，通过全场景数据采集、多维度分析与可视化呈现，为解决上述问题提供了技术路径，也为可持续发展理念的落地奠定了基础。

3.特质分析：数智+评价的适配融合

近年来，Q园一直在创新实践“数智+”评价体系，以可持续发展理念为支撑，以幼儿玩创素养发展为目标，将数字化评价与可持续的教育生态进行适配与融合，主要呈现以下特质：

3.1. 过程性追踪：循证成长全场景

通过智能系统、智慧手环等数字化工具、技术，实时记录幼儿在学习、游戏、运动、生活等各个场景中的行为轨迹与表现，形成连续性的发展档案，支持长期性评价。这些数据如同幼儿成长的“数字足迹”，不仅反映了他们的即时状态，更揭示了其成长轨迹与发展趋势。让教师能够更全面地了解幼儿，为其提供更加个性化的教育支持，同时也为家长的参与提供了科学依据，共同促进幼儿的全面发展。

3.2. 多主体协同：构建评价共同体

通过数字大脑平台，教师、家长、幼儿可以实时互动，交互式参与交流评价过程，形成评价共同体。家长可以了解幼儿在园的表现，提出自己的见解与建议；幼儿也能在评价中表达自己的感受与需求，成为评价的主体之一。这种多主体协同的评价模式，打破了传统评价中教师“独角戏”的局面，形成了教育合力，共同推动幼儿的成长与发展。

3.3. 生态化支持：实现联动多途径

在数智+评价中，通过数据整合与分析，打破幼儿园、家庭、社区之间的场景壁垒，实现园所管理、课程设计、家园共育等多方面的联动优化。通过分析幼儿在园的表现数据，教师可以及时调整课程设计，使其更加符合幼儿的兴趣与发展需求。

4.评价实践：数智+评价的创新路径

4.1. 系统构建：“数字大脑”智能系统

为更好地追踪幼儿个性化发展，开展精准化的学校管理，Q 园依托区域智慧教育平台优势，利用人工智能、大数据、物联网等先进技术，为幼儿学习发展及园所管理的全面推进构建一个后台支持服务系统，即“Q 园大脑”智能系统。系统中，打通学习空间、不同学段、各个班级、各个教师资源与评价学习平台，推动幼儿活动实施的数字化管理，更好地追踪幼儿个性化发展，形成面向幼儿全场景的数智化平台。

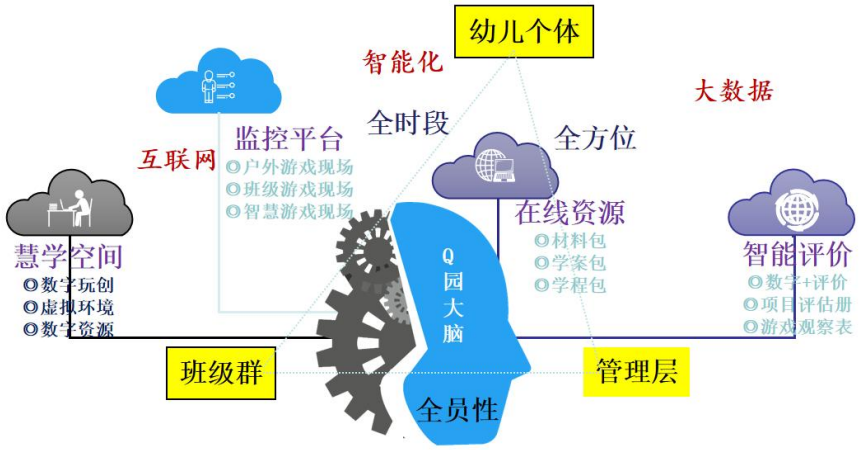


图 1 “Q 园大脑”智能系统构架

4.2. 研发设计：“数智+”评价体系

“Q 园大脑”智能系统中，“数智+”评价聚焦幼儿玩创学习与发展，基于智能手环与大数据平台，形成“一核 • 两径 • 三维 • 五面”评价模型，解决“为什么评？”“怎么评？”“谁来评？”“评什么？”的问题。一核：以幼儿玩创学习为核心，关注解决问题、迁移创新等高阶思维。两径：数字化处理（量化数据）与嵌入式跟进（质性分析）双轨并行。三维：教师、幼儿、家长三方协同评价，绘制幼儿立体发展画像。五面：聚焦“沟通表达-团队协作-解决问题-批判思维-创新意识”五维度，形成链式评价框架。

图 2 “数智+”评价模型

4.3. 实践运行：数字化与嵌入式二径齐驱

4.3.1 •. 数字化处理：循证支持高质数据

客观、系统、常态化的信息收集是保证评价科学性、有效性的重要基础[4]。在评价过程中，借助智慧手环、视频魔方等工具，捕捉幼儿在园生活学习等多样化成长量化数据，作为后续评价的重要依据。

以“智慧手环”循证追踪“全场景”

一方面，借助智慧手环，利用无感技术，实时、便捷、快速采集幼儿生活、运动、游戏、学习“全场景”综合信息。另一方面，通过手环所记录的幼儿大数据，经过大数据诊断、人工复查，形成对幼儿各方面的综合评价，个性化、精准地理解每位幼儿的发展规律，做到更有效的教育支持。

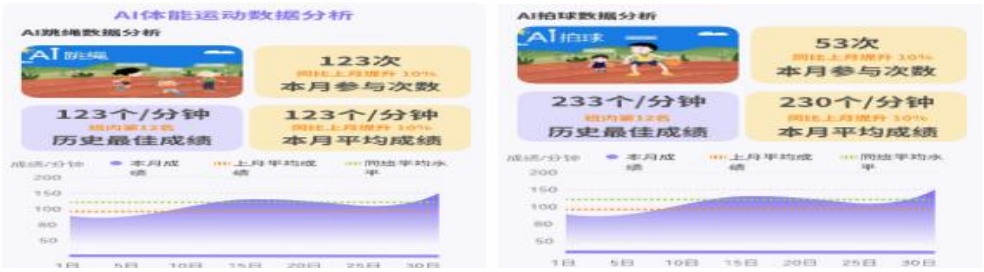
生活场景：包括入园、测温、喝水、洗手等生活活动的实时记录。入园时，手环与智能闸机相配合，通过人脸识别记录幼儿入园、健康码等数据，方便教师进行考勤与幼儿健康情况的追踪记录。手环自动记录幼儿测温数据。幼儿通过手环打卡，记录其喝水、洗手次数，及时进行健康预警和习惯引导，手环以收集生活习惯数据为前提，帮助教师实现幼儿的习惯引导。特别是之前疫情背景下帮助教师及时做好对应措施，提出相应建议，实现教师对幼儿健康习惯的评价与培养。

以喝水为例，在手环所对应的数据平台中，教师可查看一月内班级每个幼儿喝水次数的多少，通过直观的颜色区分，教师就能发现班级内喝水量较少的幼儿并及时提醒，进行正向鼓励。

图 3 幼儿喝水报告统计示例

运动场景：园所创设 AI 运动角，主要包含无感动态识别、智能计时计数、动作姿态捕捉、运动数据分析等四种功能，以 AI 视觉技术为依托，利用前端探测技术捕捉幼儿运动姿态，通过“Q 园大脑”大屏实时观测运动数据，有效实现幼儿体育运动的智能感知、智能监测，助力园所打造智慧运动课堂、开展常态化智能运动于一体的园所运动新生态。

针对幼儿发展关键期，研发对拍球、跳绳、开合跳等实现计数功能，实时呈现在数字大屏上，至多可供 4 个幼儿同时计数。运动结束实时进行评价激励，Q 园大脑后台实时统计运动结果，开展运动评价，进行月平均、学期平均统计，给出运动建议。

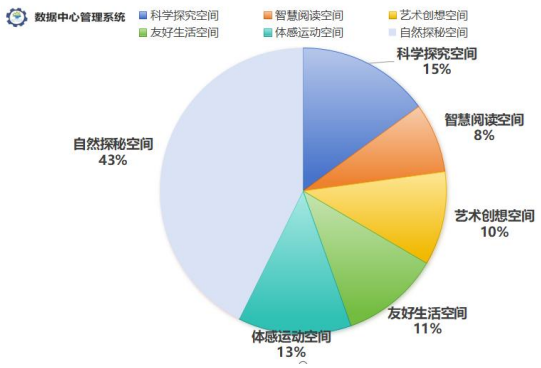


游戏场景：在班级区域游戏中，依托智慧手环，直接记录幼儿进出不同区域的频次与进行游戏的持续学习时间，并保存幼儿活动过程的视频材料，根据可视化数据帮助教师和家长精准分析幼儿游戏的情况，了解幼儿的兴趣和潜能。

以小三班乐乐小朋友为例：

- **数据显示：**2025 年 2 月，乐乐进入科学区 18 次，占总活动次数的 44%，美工区 12 次，占比 29%，语言区 11 次，占比 27%，而建构区和角色区均为 0 次。在科学区平均每次停留约为 30 分钟。
- **智能分析：**从频次和时间数据中，系统进行直接分析，乐乐对科学区的兴趣超过其他区角。且在科学区能较长时间保持专注，具有较高的坚持性。
- **游戏支持：兴趣拓展**：增加科学区的材料，如投放更多有趣的科学小实验套件，鼓励乐乐进一步探索科学，满足其好奇心。**兴趣迁移**：尝试将科学元素与建构、角色游戏结合，如开展关于身体科学实验的角色扮演，吸引乐乐参与，拓宽其兴趣领域。

学习场景：Q 园打造未来教育视域下的六大类玩创学习空间，分别是：智慧科探空间、智慧阅读空间、艺术创想空间、友好生活空间、体感运动空间、自然探秘空间。智能系统与园所六大类学习空间以及班级联通，通过智慧手环在学习空间中对幼儿学习行为的记录，教师能看到某幼儿一个月选择不同空间、材料的频次统计和每天的探究时长，也能看到班级整体幼儿对不同空间的选择频次占比，从而分析、了解班级幼儿在学习空间中的探究走向，及时调整空间创设和活动设计并给予支持，为幼儿玩创学习的深入推进提供指导。



如智慧阅读空间中的**阅读追踪：**幼儿通过手环打卡自主借阅图书，借助大数据分析，教师和家长可在幼儿自主记录的前提下发现幼儿的阅读习惯与兴趣取向，从而进行个性化教育和精准评价，或者对班级阅读活动进行调整。而通过阅读次数、时长等的记录，“Q 园大脑”系统平台也会给予幼儿相应的荣誉评价，激励幼儿坚持良好的阅读习惯。

图 4 幼儿六大学习空间选择比例图

以“AI 小智”重点赋能学教评

Q 园引进新型移动教学空间“AI 小智”，并将其嵌入“Q 园大脑”智能系统，辅助智慧手环，共同记录幼儿在园的学习情况。“AI 小智”突破了传统教学活动的空间限制，将互动投影、内容交互、AI 语音、图像识别、体感互动等技术与教学活动巧妙结合，任何场地都能瞬间化身为充满奇幻色彩的教学活动及运动空间。其内置丰富的主题活动内容，满足不同场景教学及活动需求。教师只需打开应用广场，一键选定所需的内容，即可直接投入使用，无需复杂的设置和调试。节省教师的时间和精力，也提高了教学活动的灵活性和多样性。

图 5 AI 小智嵌入智能系统的教育实践示例

以“视频魔方”聚焦分析研个例

园所班级区域、学习空间内放置了移动可调节的摄像头设备“视频魔方”，以 360 度、远近距离、立体式的摄像捕捉跟踪幼儿活动。通过智能系统中的素材分析、视频回放，对幼儿活动进行素材分析、经验再现，教师根据视频素材进行个例的针对性、聚点化评价研修，优化后续的活动组织与支持。

4.3.2. 嵌入式跟进：“四步走”支持幼儿个性生长

指标嵌入：5C 量表数字化呈现

Q 园聚焦幼儿玩创学习，围绕“沟通表达、团队协作、解决问题、批判思维、创新意识”五维度（5C 能力）评价指标，形成《幼儿玩创学习的素养发展评定表》，以量化+质性融合的方式，优化“数智+”评价工具。同时，将《幼儿玩创学习的素养发展评定表》嵌入数智系统，为教师的在线评价提供便捷操作和可视化理论支撑。

表 1 幼儿 5C 能力发展等级评定表

幼儿 5C 能力发展等级评定表

评价方法： 1. 根据幼儿的行为表现，勾选符合的评价指标，全部符合某水平的评价指标，方达到该水平；
2. 若幼儿的表现不完全符合水平三的评价指标，但接近水平一的评价指标，请先平阅水平非窗行为事例，补充并修正相应的评价指标，以此类推；
3. 平阅水平需要两一项水平符合一半及以上指标方可达到（如水平三有 3 个子指标，需符合 2 个子指标才能达到水平二）。

评价对象： **评价时间：** **评价者：**

发展关切点：沟通表达、团队协作、解决问题、批判思维、创新意识

发展关切点：沟通表达					
内涵解释：在与别人互动的过程中以恰当的方式倾听别人的意见，也表达自己的经历和想法。					
发展水平	水平一	水平二	水平三	水平四	水平五
评价指标	□不愿主动参加活动，极少主动表达自己的经历和想法 □表述事件或想法时语词混乱，难以清晰表达 □在与他人的交流中常出现游离状态，在理解他人的想法上存在困难		□能在他人的邀请下愿意说出自己的想法 □能基本表达自己的经历和想法，但语言的逻辑性和丰富性较弱 □不随意打断他人发言，但未能通过在语义、语调等主动尝试理解他人		□在公众场合能主动向他人表达自己经历和想法 □能以准确、恰当的语言有逻辑地表达自己，尝试获取他人的理解 □在他人表达时，能注视对方，认真倾听并尝试理解
行为示例	●幼儿选择独自去制作陀螺。 ●陀螺转起来后幼儿尝试向同伴分享，“转的陀螺，弄了一下不转，后来就转”，同伴表示不理解。 ●集体分享环节，幼儿东张西		●幼儿给制陀螺设计图，当同伴询问时，幼儿能主动表达设计想法。 ●制作陀螺时，幼儿激动地对同伴说，“成功了！”，同伴问其怎么做的，幼儿说，“先用这个弄一下，再把那个合起来，转一下。”		●陀螺制作完成后，幼儿主动向同伴介绍自己成功的经验。 ●制作成功后，幼儿按照制作步骤，将打孔、削棒、粘合、打磨等环节有逻辑、有顺序地介绍给大家。 ●当他人分享制作陀螺经验时，幼
	望，没有认真倾听同伴的发言。		●同小组的伙伴讲解完陀螺设计后开始分工，幼儿指着设计图说，“你说的是什么意思？”		儿注视对方，认真倾听，并向对方提出优化和装饰建议。
支持性策略	1. 在需要时，和幼儿一起玩，鼓励幼儿表达自己的想法； 2. 通过同伴或教师的表达为幼儿提供示范，更正或丰富幼儿的语言； 3. 通过眼神、动作提醒幼儿集		1. 用适宜的表情和言语回应幼儿的表达，对其表示赞赏和鼓励； 2. 在需要时，鼓励幼儿如何表达自己的想法提出建议或提供语言支架； 3. 认真倾听幼儿的想法，多询问“请问我说清楚了吗？”，鼓励幼儿复述		1. 鼓励并认可幼儿的表达，持续专注幼儿在表达中的情绪状态； 2. 鼓励幼儿活动后详细描述探究的过程，提供表征的支架，鼓励幼儿通过多种方式表达； 3. 鼓励幼儿对他人的想法提出问

智能剖析：AI 检索在线式分析

当教师对幼儿开展线上评价记录时，教师可结合嵌入的 5C 指标，通过触屏勾选、音视频插入、文字添加辅助等，线上评估幼儿的玩创学习。与此同时，“Q 园大脑”通过 AI 人工智能分析功能，系统自动化生成幼儿玩创学习的具体分析，通过 AI 自动检索、一键生成匹配适合幼儿当前发展的支持性策略和发展建议，教师在此基础上进行对校优化，节约记录时间的同时紧密结合指标，为教师评价幼儿提质增效，助力家园科学协调共育。

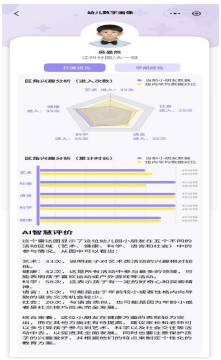
图 6 AI 检索在线分析示例

报告呈现：数字画像可视化生成

评价分析后，“Q 园大脑”智能系统会结合评价指标、AI 在线剖析，形成幼儿数字画像报告。从普适性发展和个性化发展两方面，呈现生活画像、运动画像、学习画像。不仅通过徽章形式进行在线评估表彰，还可借助多元图表形式，纵向呈现幼儿个人全面成长与个性发展，横向梳理同班幼儿、同年龄幼儿的整体发展，同时设置多个时间段，方便家长及教师查看一周、一月、一学期、一学年等不同时间段的统计分析数据，形成周度报告、月度报告、学期报告等，为后续提供有针对性地指导帮助。

徽章激励：徽章内容涉及出勤、运动、学习、午睡等方面，如早起小玩家、安睡小玩家、运动小玩家、科学小玩家、语言小玩家等，通过系统在线生成和教师评估的方式为幼儿颁发徽章，共同呈现幼儿在园生活的个性特长和成长激励。

图像分析：通过雷达图、折线图、条形图等多元形式呈现幼儿全场景下的个性玩创成长。如，以能力雷达图呈现各能力发展是否均衡、长处与短板；以折线图呈现幼儿前后学习发展对比，明确幼儿近阶段成长的能力素养与下阶段的重点关注；以条形图呈现各领域幼儿游戏时长、前后对比等，直观表明幼儿兴趣、坚持等学习品质的变化。



家园协同：成长云库生态式互动

“Q 园大脑”智能系统共研优化“成长云库”，设计“班级圈新闻”、“成长手册”、“5C 能力报告书”等，以“日-月-期”的时间机制，呈现幼儿每日动态消息、每月成长记录、每学期玩创能力综合分析，形成多维一体云档案。并同步跟进家园协同，形成生态式的家园互动。主要体现在：数据动态性：家长端实时查看幼儿生活、运动、游戏、学习各场景数据，数据动态生成，打破家园信息壁垒。内容全面性：整合健康、社会、语言、科学、艺术等多维度数据，数据完整全面，个性化凸显。交互融通性：家长可通过班级圈、成长手册等内容板块的评价讨论，记录反馈家中幼儿成长，教师依据数据和反馈信息，优化家园共育策略。

图 7 幼儿成长云库示例

5.评价成效：实现学习与发展的可持续

数智+评价实践中，Q 园形成了“数据驱动-精准施教-智能协同”的新型教育生态，展开“实证研究-迭代优化”的良性循环。

5.1. 幼儿发展：从“浅层学习”到“深度学习”

5.1.1. 多维活动开展中，幼儿玩创思维提升显著

教师聚焦幼儿评价中的兴趣与问题，展开项目活动、智慧游戏、主题节三大类活动实践，幼儿在玩中学、创中思的能力得到提升。经统计发现，同一班级幼儿提出问题的频次较之前增加了 42%，这一数据直观地反映了幼儿在探索与发现过程中的主动性与好奇心显著增强。同时，幼儿能够将所学经验进行迁移与创新，如中班幼儿项目活动《趣搭帐篷》中，从班级“搭帐篷”模型拓展至户外工程建构“搭露营帐篷”，展现了其迁移创新能力。

5.1.2. 多元学习支持下，幼儿学习品质优化明显

大数据支持下的数智+评价，同步为幼儿提供了更多元丰富的学习方式。在“虚拟增强现实”AR 技术支持下，通过场景模拟真实体验，展开 AR 动画创客、三 D 沉浸体验、AI 智能运动等活动体验，将静态事物转化为动态互动。通过智能技术与情感浸润并驱，感受技术带来的新型学习体验与经验增长。在多元学习方式的支持下，幼儿的专注力、兴趣度、好奇心得到进一步提升。智慧手环数据显示，专注时长超 30 分钟的幼儿占比从 25%提升至 68%。

5.2. 教师成长：从“经验驱动”到“循证支持”

5.2.1. 课程设计优化

基于区域游戏、学习空间的数据分析和成长画像图谱，进一步设计实践三类“STEAM 项目群”。通过“兴趣索引-项目孵化-经验迭代”的课程开发机制，提升课程设计及教育实践能力。

5.2.2. 评价素养提升

90%的教师能熟练运用数据平台分析幼儿行为，结合 AI 智能分析与指标支持，精准定位幼儿的发展水平与下阶段支持策略。同时，结合幼儿行为数据分析，实现因材施教，与家长实时互通，提升了家园共育成效。

5.2.3. 数字素养加强

Q 园建立了“智慧合伙人”研修机制，组建成“山水智库”，实现信息技术学习、设计与运用的初步融合。目前，基本所有教师已掌握虚拟仿真、增强现实等智能教学工具，能有效设计智慧教学活动，实现了跨空间情境化教学。

5.3. 园所治理：从“粗放管理”到“精准决策”

5.3.1. 打通了教育管理数据壁垒，确保园所管理一体化

“Q 园大脑”智能系统的建立，破除了“一园多平台”使用弊端，实现低成本覆盖集团园的需求；建立了数据驱动的教育管理机制，提升了教育治理水平；推动家园共育形成合力。同时，数据平台支持各分园区的薄弱项分析，实现师资培训与资源调配的精准化，共同促进教育共富，扩大了幼儿园高质量品牌影响力。

5.3.2. 形成了教育管理数链体系，推动教育决策精准化

以智能手环为载体，搭建了全场景下的数智化平台，记录下的幼儿数据形成了 Q 园教育管理的数据链，实现了教育数据伴随式采集，支撑管理扁平化和监测精细化。同时，数据大脑为各级教育决策者提供能直观监测所有辅助决策的幼儿纵向、横向发展数据，完善教育管理中枢决策系统的指挥调度功能，助力管理者科学决策，园所教育决策依据数据更为科学化、精准化。

6.反思与展望

可持续发展理念下的教育评价，既是科学命题，亦是教育情怀的体现。以数字化赋能评价创新，得以更贴近幼儿的真实成长，更精准地支持其个性发展。信息技术赋予幼儿园更多畅想未来教育的空间。未来需进一步：深化数据分析，避免技术异化；拓展应用场景，融入自然探究、社区实践等真实情境，强化评价的生态属性[6]。通过数字大脑，依托数智+评价，园所将持续探索基于大数据的精准评价与管理，不断拓展智慧化应用场景，建构幼儿个性化成长环境，用数字化关键动能，撬动“教育共富”的支点！

参考文献

United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development (pp. 3-6). United Nations.

Li, J. M.& Feng, X. X. (2013). 3-6 岁儿童学习与发展指南解读 (pp. 12-14). 人民教育出版社.

Wen, J. Q. (2019). 发现·理解·支持——指向个性化教育支持的幼儿发展评价实践 (pp. 26-36). 上海教育出版社.

He, T., & Guo, L. P. (2024). 信息技术赋能儿童发展性评价的价值体现及其发展方向. 幼儿教育（教育科学）, (4), 7-11.

Zhang, R. C., & Fang, L. (2025). AI 与儿童：人工智能设备在学前教育中的应用. 齐鲁师范学院学报, 40(1), 49-56.

Huang, R. H. (2023). 人工智能正加速教育变革：现实挑战与应对举措. 中国教育学刊, (6), 26-33.

计算思维导向下大班游戏化编程教学的实践研究——以"机器人夺宝"科学规则游戏活动为例

Practical Research on Gamified Programming Teaching in Large Classes under the Guidance of Computational Thinking — Taking the 'Robot Treasure Hunt' Science Rules Game Activity as an Example

沈瑶
杭州市滨江区物联网小镇幼儿园
shenyao2016@dingtalk.com

【摘要】 本研究通过构建编程游戏化学习（Game-based Learning，GBL）的教学模式，通过两轮行动研究验证游戏化编程活动对大班幼儿计算思维发展的促进作用。以"机器人夺宝"科学规则游戏为载体，设计包含分解策略、模式识别、抽象建模、算法调试等要素的沉浸式任务，形成游戏化编程的"四维游戏路径"，为幼儿计算思维启蒙提供实践范式。

【关键词】 计算思维；游戏化编程；机器人教育；科学活动

Abstract: *This study constructs the "GBL" (Game-based learning) teaching model and verifies the promoting effect of gamified programming activities on the development of computational thinking in large class children through two rounds of action research. Using the scientific rule game "Robot Treasure Hunt" as a carrier, design immersive tasks that include decomposition strategies, pattern recognition, abstract modeling, algorithm debugging, and other elements, forming a "four-dimensional game path" of gamified programming, providing a practical paradigm for the enlightenment of young children's computational thinking.*

Keywords: Computational Thinking，Gamified Programming，Robot Education，Scientific Activities

1.前言

1.1.研究缘起

在新时代教育改革的背景下，《义务教育信息科技课程标准》明确提出"依托真实情境培养计算思维"的核心要求，为学前阶段计算思维启蒙指明了方向[1]。然而当前学前教育编程教育实践仍存在显著偏差：一方面陷入"重操作轻思维"的机械化训练误区，将编程教学窄化为机器人操控技术学习，缺乏以从学前儿童年龄特点出发，以游戏化的方式激发幼儿学习兴趣或动机；另一方面受"重个体轻协作"的传统模式制约，忽视合作解决问题的社会性思维培养。与此同时，国际前沿理论的本土化进程亟待突破——Seymour Papert 提出的"建造主义"学习观与 Marina Bers 倡导的"编程即游乐场"理念虽具指导价值[11]，但在中国文化语境、幼儿园课程体系及 5-6 岁儿童认知特点间的适配转化仍缺乏系统研究，这一理论缺口导致教育实践难以形成科学可持续的实施路径。

1.2.核心概念

计算思维：一种运用计算机科学的基本概念，以解决问题、设计系统和理解人类行为的方式，分为计算思维概念、计算思维实践、计算思维观念。本文中所运用的是编程活动中学前儿童计算思维的关键发展指标[2]。游戏化编程教学：通过角色扮演、规则竞技等游戏规则活动，以实现编程思维的隐性传递。

2.GBL 教学模式构建

计算思维由计算思维概念、计算思维实践和计算思维观念三大内容构成。在学前儿童编程游戏领域，目前的编程教学在思维运用及团队合作方面存在一定欠缺。基于此，本文将运用计算思维实践内容，构建包含四个步骤的游戏化教学模式，旨在强化幼儿的计算思维能力。同时，将计算思维观念中的“主动性、坚持性、合作性与创造性”隐性融入这一模式[3]。通过创设趣味情境激发幼儿主动性参与，头脑风暴及游戏试错培养其坚持性，小组合作游戏强化合作性，而游戏中的自由创作与策略创新则锻炼创造性[4]。如此，让幼儿在规则游戏中逐步提升计算思维能力，为其后续发展奠定坚实基础。

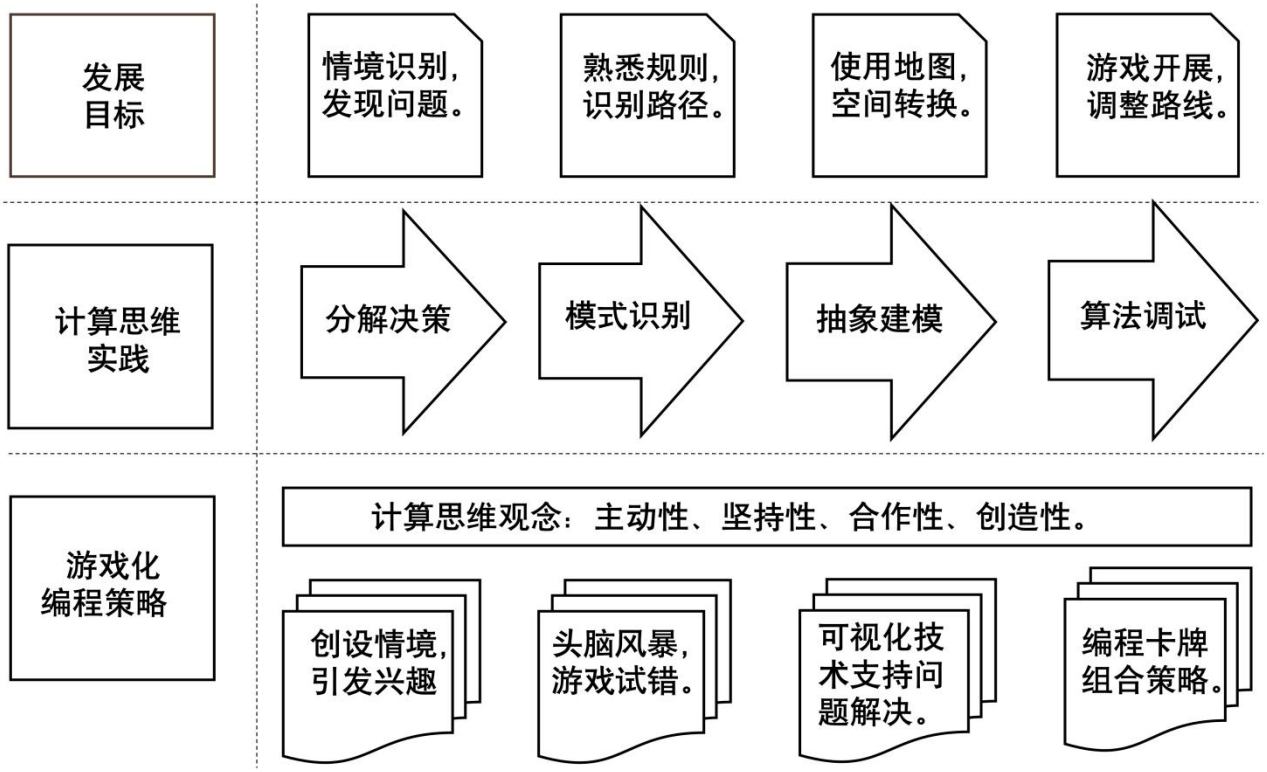


图 1 GBL 教学模式框架图

3.GBL 教学实施路径

3.1.情境识别，发现问题

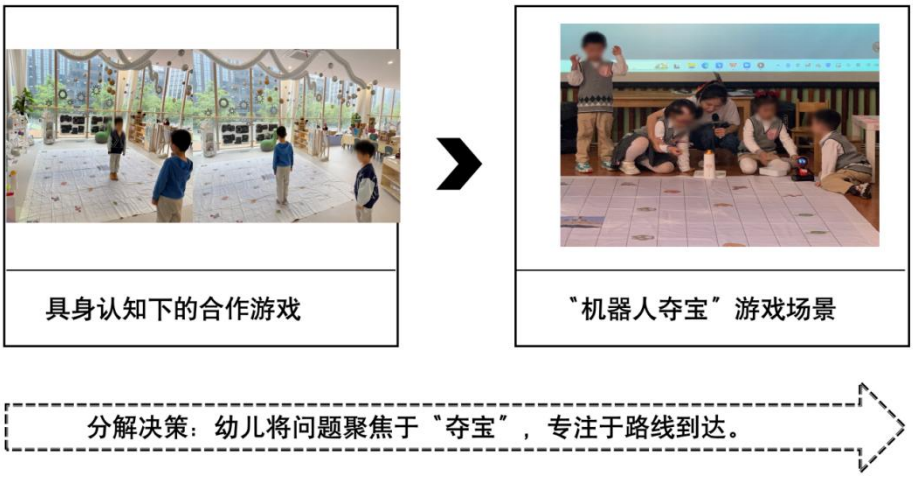


图 2 幼儿沉浸式游戏场景图

3.1.1.情境铺垫，深化趣味 幼儿对机器人编程游戏展现出显著的兴趣倾向，其核心在于机器人形象与编程逻辑的深度融合。

本次游戏设计的关键经验获取与核心探索方向，即围绕如何实现机器人与编程游戏的有机整合而展开。依据《3-6 岁儿童学习与发展指南》中强调的具身认知与情境体验原则，前期游戏阶段着重为幼儿营造多感官协同、具身实践及实物操作的沉浸式编程体验环境。

教师引导幼儿在大型感官垫上开展角色扮演式的机器人编程游戏。幼儿在分组活动中通过角色分工与互动，一方扮演“程序员”发出基于图形化编程指令（如前进、转弯、抓取等动作代码），另一方则化身“执行机器人”，根据指令在网格坐标系中精准移动特定物体，或沿预设轨迹行进。在进阶游戏环节，以“机器人夺宝”为主题的科学规则游戏将游戏情境深度拓展。教师运用叙事性讲述方式创设沉浸式冒险场景，将参与游戏的四组幼儿纳入竞争合作的多元互动框架[5]。

3.1.2.分解决策，发现问题 幼儿将游戏活动的情境进行分析与解剖，形成相应的聚焦问题，以此将所需要解决的问题具像化。

在“机器人夺宝”游戏中，教师与幼儿一起将游戏场景确定化，以此将编程的内容与方向指向于“到达终点”这一目标。

3.2.熟悉规则，识别路径

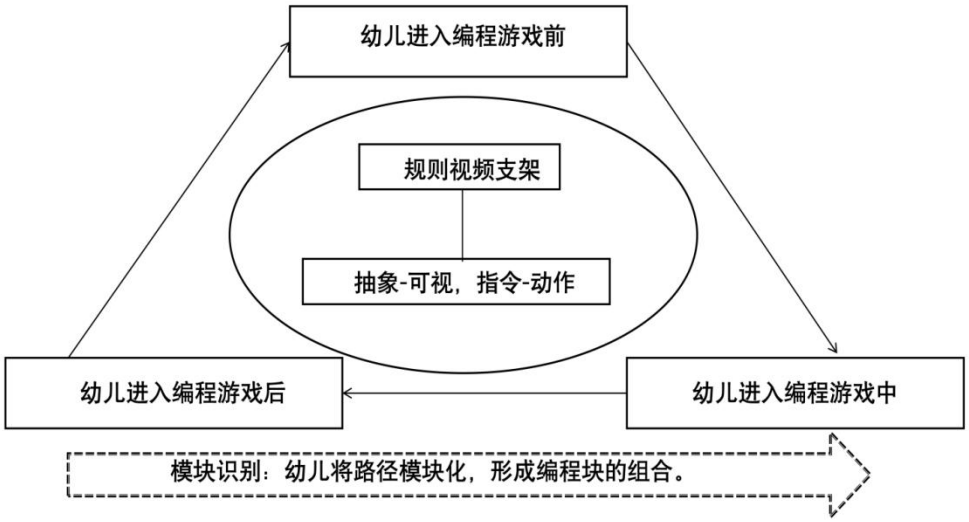


图 3 幼儿游戏化具象规则转换图

3.2.1.具象规则，直观了解

本次以实体编程机器人为媒介，教师秉持直观性教学原则与布鲁纳的“螺旋式课程”理念，运用信息技术手段将抽象的游戏规则进行可视化转译，制作成为动画小视频。

视频以幼儿即将深度互动的实体编程机器人——萤石为具象载体，通过生动的情境模拟与分步演示，将规则拆解为易理解的微单元[6]，如指令输入、信号反馈、动作执行等关键环节，帮助幼儿实现从具象操作到规则内化的平稳过渡。

在游戏导入阶段，教师引导幼儿观看视频，结合实时提问与互动讨论，激活幼儿的前运算思维，使其初步构建“指令—动作”对应关系的心理图式。同步将萤石机器人置于幼儿触手可及之处，鼓励其在观看过程中随时触摸、摆弄机器人，建立客体（机器人）与自我的联结[12]。

3.2.2.模块识别，认识路径

计算思维中的模块识别对于学前儿童来讲，主要是发现问题中重复出现的模式和规律，以此进行编程块的运用。

在进行机器人夺宝游戏的时候，幼儿识别到机器人从起点到终点的路径，形成模块化的方向识别与数量确定。

3.3.使用地图，空间转换

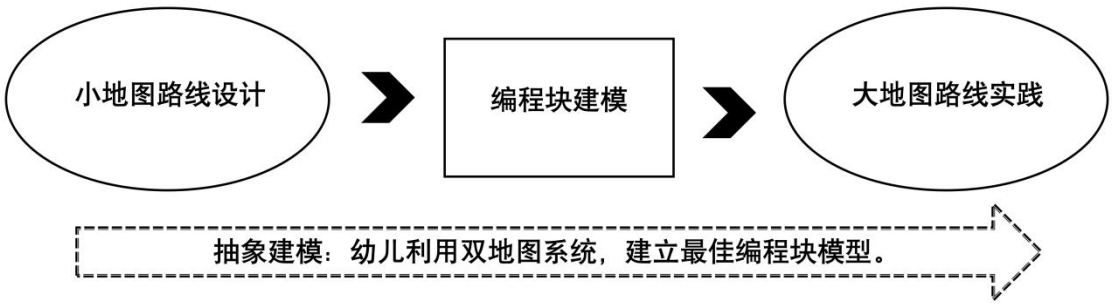


图 4 幼儿双地图模式运用图

3.3.1.大小地图，现二维空间

游戏中创新性地构建了二维空间认知架构，通过设计的双层地图系统实现。

该系统包含两套功能各异却又紧密关联的地图：其一为大型操作地图，作为幼儿开展实体“夺宝”活动的主场地，幼儿可在其上进行实时的机器人编程操控，将抽象编程指令转化为具体的机器人行动路径；其二为小型规划地图，其核心功能在于辅助幼儿进行前瞻性的路线规划与策略构思，可视为幼儿思维可视化的脚手架。幼儿首先在小地图上运用图形化编程块拼搭、组合，模拟构建虚拟行动轨迹[13]，这一过程实质是将编程逻辑与空间认知深度融合，待幼儿在小地图上完成初步的路线设计与编程块编排后，再将这套预设方案迁移到大地图，付诸于实体机器人实践操作，实现虚拟到现实、抽象到具象的无缝对接[7]。

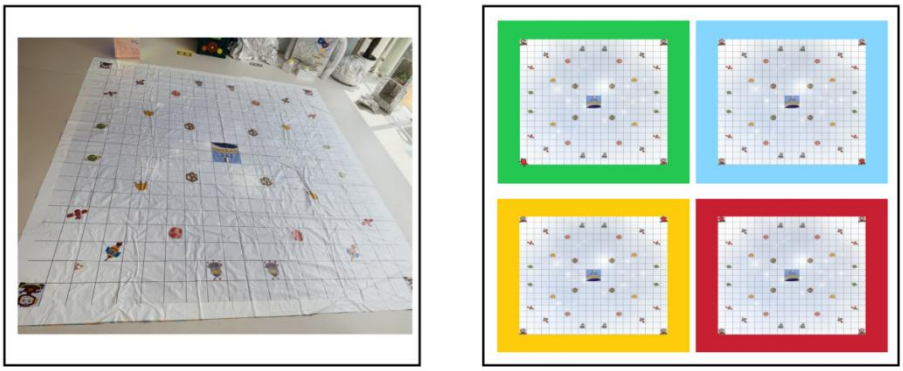


图 5 大地图 小地图

3.3.2.抽象建模，可视化支持

为助力幼儿在游戏中达成最优化任务目标——即精准且迅速地抵达终点，活动依托双地图系统，创设具象化、结构化的学习支架，融入抽象建模与可视化支持手段。

运用抽象建模思想，将“直行、后退、左转、右转、数量块和角度块”等离散的编程元素进行系统整合与模块化组合，构建虚拟行动轨迹。这一过程实质是将复杂路径问题分解并重构为可操作的编程模块，实现编程逻辑的可视化与具身化，帮助幼儿在脑海里形成清晰的行动方案。

3.4.游戏开展，调整路线

3.4.1.合作游戏，确定路线 幼儿以四人小组为单位展开合作。

这一设置旨在促进幼儿间的互动交流与协商，契合维果茨基的社会文化理论，强调同伴互动对认知发展的重要推动作用。游戏过程中，幼儿通过主客体调整，即不断协调自我认知与他人观点，运用协作对话这种社会性交互方式，共同商讨并确定组内的游戏路线。例如，在规划机器人夺宝路径时，幼儿们各自提出想法，通过轮流发言、角色扮演等协商策略，最终达成共识。

3.4.2.算法调试，组合策略 幼儿运用试错法对编码块进行系统性调试与优化。

这一过程具体表现为：幼儿依据初始假设，将“直行、转弯、前进、后退”等基础编码块进行组合拼搭，形成初步路线指令。随后，通过机器人执行指令，幼儿能即时观察行动结果与预期的偏差。基于观察反馈，幼儿运用批判性思维，对编码块组合进行调整、交换与重组，反复测试。

幼儿不仅逐步理解编程逻辑中的条件判断、循环结构等基础概念，还自然地将计算思维中的“优化”理念融入实践，以减少指令步骤、缩短机器人行进路径为目标，持续改进编程方案。例如，在多次试错后，幼儿发现通过调整角度块的参数，可使机器人以更优的转弯半径避开障碍物，实现更高效的路线规划[14]。

以下观察记录呈现了幼儿如何四人合作编程并预测行驶路径[8]。

A 幼儿开始尝试将编程块在编程板上进行组合。她先放置了一个“直行”块，然后是一个“右转”块。B 幼儿则在一旁用手指着编程板，建议 A 幼儿调整编程块的顺序。他拿起“右转”块，放在“直行”块前面。C 幼儿和 D 幼儿则在地图上模拟机器人的移动，验证编程块组合的效果。他们发现机器人按照当前编程块的组合无法避开障碍物，于是开始讨论如何调整。

语言表现：

A 幼儿：“这样不行，机器人会撞到障碍物。”

B 幼儿：“我们应该先把右转放在前面，然后再直行。”

C 幼儿：“对，这样它就能绕过那个星球了。”

D 幼儿：“我们再试试看。”

调整过程：

A 幼儿根据 B 幼儿的建议，调整了编程块的顺序：先“右转”，再“直行”。

B 幼儿则在编程板上增加了“左转”块，以确保机器人在绕过障碍物后能正确转向宝藏。

C 幼儿和 D 幼儿继续在地图上模拟机器人的移动，验证新的编程块组合。他们发现机器人现在可以顺利绕过障碍物，但还需要调整移动距离。

A 幼儿：“我们需要增加直行的距离，这样它才能到达宝藏。”

B 幼儿：“对，我再加一个直行块。”

C 幼儿：“这样应该可以了，我们再试一次。”

4.研究成效

4.1.幼儿计算思维能力提升

4.1.1.关键能力显著发展 通过“GBL”教学模式实施，幼儿在计算思维的关键能力维度取得明显进步。

在分解策略方面，面对“机器人夺宝”游戏中的复杂任务，幼儿能够将问题拆解为多个有序步骤。模式识别能力也得到有效锻炼，幼儿能敏锐察觉游戏情境中的规律。抽象建模能力稳步提升，幼儿借助小地图将具体路径抽象为图形化编程指令组合，再迁移到大地图实践，实现虚拟与现实的对接。此外，算法调试能力也得到培养，幼儿通过反复试错与调整，优化编程块组合，减少指令步骤、提高机器人行动效率[9]。

4.1.2.思维观念正向转变 幼儿在游戏过程中展现出更强的主动性、坚持性、合作性与创造性。

趣味情境激发幼儿主动探索编程玩法，遇到难题时，他们更愿意坚持尝试，通过头脑风暴与游戏试错寻找解决办法。小组合作游戏强化了合作意识，幼儿学会倾听他人想法、协商路线规划，共同攻克编程难题。同时，游戏中的自由创作空间鼓励幼儿创新策略，如设计独特的机器人行动路线、优化障碍规避算法等，多维度提升了计算思维观念水平[10]。

4.2.形成“GBL”游戏化编程教学模式

4.2.1.模式构建与验证 本研究成功构建“GBL”（Game-based learning，GBL）游戏化编程教学模式，并通过两轮行动研究验证了其有效性。

该模式以“机器人夺宝”科学规则游戏为载体，围绕计算思维核心要素，设计包含情境识别、熟悉规则、使用地图、游戏开展等四个步骤的“四维游戏路径”。从规划幼儿初步接触编程游戏，到运用模块识别等知识解决问题，再到利用大小地图实现空间转换与抽象建模[15]，最后通过合作游戏与算法调试达成最短路线目标，形成了完整且富有成效的教学流程。

4.2.2.实践应用与推广价值 在杭州市 W 园大班教学实践中，“GBL”模式取得显著成果。

幼儿参与度高，计算思维能力得到系统培养，为学前教育阶段编程教育提供了可借鉴的范式。该模式不仅契合国内教育标准与儿童认知特点，还弥补了国际前沿理论在本土实践中的空白，具有广泛的推广价值，能为其他幼儿园开展计算思维启蒙教育提供有效指导，推动学前教育领域编程教学的科学化发展。

参考文献

马宗兵(2020)。计算思维导向的小学编程教学策略研究。(硕士学位论文,华南师范大学)。

https://link.cnki.net/doi/10.27154/d.cnki.ghnsu.2020.000003doi:10.27154/d.cnki.ghnsu.2020.000003
刘微微(2024)。以《制作阳光运动计步器》课编程教学锻炼儿童计算思维能力。小学科学,(23),73-75。
高宏钰,王妍茜 & 申玮莉(2024)。编程活动中学前儿童计算思维的关键发展指标。福建教育,(43),20-22。
蒋云宵 & 高宏钰(2024)。器人编程和积木搭建对大班幼儿计算思维发展的影响研究。福教育,(43),23-25。
王程程,马曼妮,陆娴桉 & 李含灵秀(2024)。工智能赋能下儿童叙事语言能力与计算思维素养协同培养的活动模式探索。2024 计算思维与 STEM 教育研讨会暨 Bebras 中国社区年度工作会议论文集(pp.297-306)。首都师范大学学前教育学院; https://doi:10.26914/c.cnkihy.2024.010443.
张可 & 耿凤基(2024)。儿童计算思维发展的个体差异表现及其产生机制。杭州师范大学学报(自然科学版),23(03),225-232。https://doi:10.19926/j.cnki.issn.1674-232X.2024.03.222.
高宏钰,于开莲,蒋云宵 & 杨雨欣(2024)。无屏幕编程教育活动对学前儿童计算思维和创造性思维的影响研究。电化教育研究,45(04),96-103。 https://doi:10.13811/j.cnki.eer.2024.04.014.
高宏钰,蒋云宵,杨伟鹏,孙立会 & 宋高阳(2023)。学前儿童的计算思维发展情况及其影响因素——基于北京市 101 名 4～6 岁学前儿童的测试分析。现代教育技术,33(07),44-52。
朱月姣(2023)。面向儿童计算思维培养的可编程教育机器人设计与开发研究(硕士学位论文,佛山科学技术学院)。 https://link.cnki.net/doi/10.27960/d.cnki.gfskj.2023.000012doi:10.27960/d.cnki.gfskj.2023.000012.
孙怡然(2023)。基于故事的项目式儿童编程课程设计。中国信息技术教育,(09),91-94。
Kazimoglu, C. , Kiernan, M. , Bacon, L. , & Mackinnon, L. . (2012). A serious game for developing computational thinking and learning introductory computer programming. Procedia - Social and Behavioral Sciences, 47(1), 1991-1999.
Kazimoglu, C. , Kiernan, M. , Bacon, L. , & Mackinnon, L. M. . (2011). Understanding computational thinking before programming: developing guidelines for the design of games to learn introductory programming through game-play. International Journal of Game-Based Learning, 1(3), 30-52.
Tsalapatas, H. , Heidmann, O. , Alimisi, R. , & Houstis, E. . (2012). Game-based programming towards developing algorithmic thinking skills in primary education. Scientific Bulletin of the Petru Maior University of Tîrgu Mureș, 9 (XXVI)(1), 56-63.
Tsarava, K. , Moeller, K. , Pinkwart, N. , Butz, M. , & Ninaus, M. . (2017). Training Computational Thinking: Game-Based Unplugged and Plugged-in Activities in Primary School. 11th European Conference on Game-Based Learning ECGBL 2017.
Candrawati, E. , Sardjono, R. E. , Rohman, I. , & Prahani, B. K. . (2024). Implementation of computational thinking through game-based learning: a systematic literature review. KnE Social Sciences.

“编”玩边学，“程”现童趣——数智 AI 赋能幼儿园编程启蒙课堂的实践研究

赵丽婷，林丽云
浙江省文成县大峃镇龙川幼儿园
986430618@qq.com

【摘要】 本研究探讨数智 AI 技术在幼儿园编程启蒙教育中的应用，以中班"玛塔寻宝"活动为例，结合 Scratch Jr 软件开展实践研究。通过设计玛塔机器人实物编程与 Scratch Jr 虚拟编程相结合的双轨教学模式，构建游戏化学习任务，帮助幼儿掌握基础编程概念。研究发现：图形化界面显著降低认知难度，"虚实结合"方式有效提升学习兴趣和协作能力，AI 技术为教学评估提供支持。研究为幼儿园编程启蒙教育提供了可借鉴的实践方案，并就技术适配性提出优化建议。

【关键词】 数智 AI；编程启蒙；幼儿园；玛塔机器人；Scratch Jr

Abstract: *This study explores the application of digital AI technology in early childhood programming education, taking the "Mata Treasure Hunt" activity in the middle class as an example and carrying out practical research in combination with the Scratch Jr software. By designing a dual-track teaching model combining Mata robot physical programming and Scratch Jr virtual, game-based learning tasks are constructed to help children master basic programming concepts. The findings show that the graphical interface significantly reduces the cognitive difficulty, the "virtual and real combination" effectively enhances interest and collaborative ability, and AI technology provides support for teaching evaluation. The research provides a referenceable practice plan for early childhood programming education, and optimization suggestions are put forward the technical adaptability.*

Keywords: Digital AI, programming enlightenment, kindergarten, Mata robot, Scratch Jr

1.“编”写案例背景

社会环境中科学技术形式的不断创新，使教学领域出现更多信息教育技术应用于学科教育活动。尤其在学前教育时期，信息教育技术已经成为教学中必不可少的教育手段，加之幼儿园编程自身的趣味性和可操作性，其在教学中的应用，能在幼儿自己操作中凸显其主体地位；同时能全方位掌握幼儿活动前认知水平、活动中的参与程度、课后的拓展迁移情况；能极大地拓展幼儿的学习渠道和学习方式，提高其自主学习的能力。

同时，幼儿的心智发育还未成熟，他们的学习其实就是一个对环境和事物理解的过程。幼儿的早期教育应该着眼于发展他们的主动活动，让幼儿在与环境、事物的互动中获得学习与发展。蒙台梭利认为，儿童内在的潜能是在环境的刺激和帮助下发展起来的，是个体和环境之间相互作用的结果。现结合现代信息技术、图形化编程工具、幼儿实体编程等应用手段，为幼儿营造情景式学习环境，让幼儿在学习过程中始终保持兴奋、愉悦的状态，激发幼儿的学习兴趣和主动性，提高幼儿的科学素养。

接下来以中班编程活动《玛塔寻宝》中的实践情况进行分析与反思，从中总结出数字赋能课堂的有效手段。

2.“编”造探索过程

2.1. 设计意图

幼儿园教育指导刚要（试行）指出：从生活或媒体中幼儿熟悉的科技成果入手，引导幼儿感受科学技术对生活的影响，培养他们对科学的兴趣。图形化编程 APP 和实体编程工具都源于幼儿的生活，是目前较为成熟的幼儿编程科技成果，孩子们能够在操作中体验中提高自身的科学素养。同时，幼儿编程关注儿童的兴趣、经验和生活世界，关注儿童的游戏和游戏中的探索，让幼儿基于游戏化的情境，尝试解决现实生活或虚拟世界中的问题，以提升幼儿的逻辑思维、问题解决、合作共情和自我实现等学习能力与品质。因此，在幼儿阶段开展编程活动是迫切需要的，本节活动是玛塔编程启蒙学习的初始阶段，通过这节活动，我们将幼儿带入趣味的故事情境中，用不插电的编程活动帮助幼儿掌握机器人运行的规范指令，为今后操作使用玛塔机器人打好基础。

2.2. 活动目标

通过《玛塔寻宝》编程活动的开展，实现孩子三维目标和计算思维目标的同时发展，具体如下：

2.2.1. 三维学习目标

认知——能在玛塔编程游戏中了解“前进”、“后退”指令的含义。

技能——学会使用玛塔编程工具，并能设计简单的程序。

情感——愿意主动和他人交流，一起寻求问题解决方法。

2.2.2. 计算思维目标

拆解思维——理解机器人的程序指令是由一个一个单独的指令构成的，每发出一个指令，机器人只能做一个对应的动作。

模式识别——能够在不同的任务场号下总结出应用指令的规律，如当目标在机器人的后方，可以应用后退指令;而当目标在机器人前方，可以应用前进指令。

抽象思维——当玛塔机器人摆放方向和自己的正前方不一致时，能够排除掉干扰信息，正确判断玛塔机器人的前进和后退方向。

算法思维——能够在教师的引导下，一步步地发出指令，直到到达月标点拿到宝石。

2.3. 活动准备

教师材料准备:智能触控电子教学一体机 1 台、玛塔课程 PPT、玛塔机器人组建 1 套、编程块 1 套、空白地图 1 张、iPad 平板 1 个、智能手机 1 个。

幼儿材料准备:iPad 平板每组 1 个、玛塔机器人组建每组 1 套、编程块每组 1 套、空白地图每组 1 张。

本次教学活动利用具有播放音视频、电子白板、希沃交互等功能的智能触控电子教学一体机，应用希沃白板课件（内含有动画、音频、投屏等）、平板操作，实现了多教学系统融合。改善了教学环境，丰富了教学资源，创新了教学模式，提高了教学质量。使学科教学与智能化、数字化的实施有了条件保障和技术支持。

运用希沃白板动画功能能有效地化抽象为具体，把难以理解的内容用多媒体展现出来，激发幼儿的学习兴趣。利用音频，色彩变幻等信息技术达到幼儿视听结合的目的进行学习，让幼儿通过直观形象、生动的感官刺激最大限度地了解活动主题破解活动中的重难点。在活动中利用希沃授课助手以及平板自主操作，幼儿能够更加自主地进行知识学习，发挥探究潜能。投屏的交互达到良好的师幼互动、幼幼互动，人机互动。在有限地时间里全方位地获得

更多的信息，提高学习效率。

2.4. 活动过程

2.4.1. 利用信息技术巧设动画，激发幼儿的好奇心

出示玛塔家族“走”出来入场画面，引导幼儿观察画面，引导幼儿自主表达“它们是谁？”，再通过音频介绍玛塔家族的成员和功能，教师根据幼儿的观察与表达进行小结。

引导语：你们看，谁来了？这是玛塔的大家族，你们谁认识它们？

小结：我们把编程块放到编程板上，这个指令就可以通过控制塔传给玛塔，玛塔就可以在地图上行动起来了。

信息技术运用方法：活动开展前，利用信息技术出场动画，设置了一个走入与卡通语音介绍的效果。这样的设计能够先让幼儿进行思考表达，随着玛塔的入场与介绍吸引幼儿的兴趣，从而展开了对玛塔大家族的讨论，利用信息技术巧设动画，吸引幼儿兴趣，帮助幼儿回顾玛塔机器人的基本构成。

2.4.2. 利用信息技术多样组合，认识前进后退指令

幼儿观察玛塔的变化，教师引导提问，幼儿发现一个前进指令能让玛塔往前走一格，一个后退指令能让玛塔往后退一格，通过实际的观察与表达感知前进指令和后退指令的作用。

引导语：玛塔有很多的本领，它要给你们展示新的本领！玛塔展示了什么本领？它为什么会动呢？你们觉得它往前走应该用到哪块指令？两块前进指令要走几格？剩下的这一块指令是什么意思？玛塔接收到一块后退指令会怎么走呢？

小结：玛塔接收到这么一块前进指令就会往前走一格，接收到一块后退指令就会往后退一格。

信息技术运用方法：在此环节利用动画里的路线+音效实现幼儿发现编程指令和玛塔行动上的关联性，自然而然地引出了“前进指令和后退指令”这个本节活动的核心问题。可以说整个环节就是幼儿自主发现表达问题，真正突出了幼儿的主体地位，也为下一步的真正探究奠定了基础。

2.4.3. 利用电子设备交互操作，学习使用 Scratch jr

引导语：玛塔在一个藏宝洞里，藏宝洞里面有好多美丽的宝石，还有一只凶恶的老猫。玛塔要趁着老猫还在睡觉，赶紧采宝石了。玛塔需要从这个位置出发去采蓝宝石，玛塔要怎么走呢？我们需要用哪个指令？需要几块？我们还要把宝石运回到安全位置上，不然老猫醒了就完蛋了。玛塔要怎么走呢？我们需要用上哪个指令？需要几块？

A.使用一个前进指令，玛塔往前走一格，采取蓝宝石。

B.使用一个后退指令，玛塔往后退一格，回到安全位置。

幼儿学习使用 Scratch jr 软件中的开始、前进、后退指令模块，体验图形化编程工具的魅力，在操作的过程中师幼之间产生互动，进一步提升幼儿对 Scratch jr 软件的掌握。

信息技术运用方法：操作实验中，幼儿利用平板进行自主操作 Scratch jr 软件 APP，我用授课助手，把手机和电脑同步，把手机拍到的同步到大屏。不但教师可以时刻关注他们在圆形组合时的操作细节，幼儿自己也可以从大屏幕中看到自己及同伴操作的不同之处，从而能让幼儿有更好的创造思路与想法。

2.4.4. 利用编程路线难度升级，尝试多次使用指令

引导语：哇，又有一颗黄宝石出现了！请你们用 Scratch jr 自己摆一摆，记得要把宝石送回到安全位置上。

小结：原来采黄宝石需要用到两个前进指令和两个后退指令。

幼儿根据地图路线用 Scratch jr 软件进行编程，教师巡回指导，投屏进行分享交流。

信息技术运用方法：这一环节幼儿利用平板进行操作 Scratch jr 软件，探索多次使用前进和后退指令模块进行编程，教师通过投屏，展现幼儿的创作结果，幼儿可以相互学习，提升经验。

2.4.5. 利用实体玛塔编程工具，尝试完成编程任务

引导语：哇，又有一颗红宝石出现了！我给你们带来了玛塔机器人，请你们指挥玛塔去采红宝石，记得还要把宝石运到安全位置上！

小结：原来采红宝石需要三个前进指令和三个后退指令。

幼儿根据地图路线用实体玛塔编程工具进行编程，教师巡回指导，手机直播进行评价。

信息技术运用方法：教师使用智能手机终端对上述两类游戏活动进行实时直播，即时表扬鼓励有创意的想法，并将相关作品投送至大屏幕展示，便于每位幼儿观看细节、碰撞思维与互相启发。

2.4.6. 利用信息技术家园联动，经验拓展巩固学习

集体教学的时间是有限的，但是幼儿学习的创造力是无限的，同时在教学活动中孩子们是存在个体差异的，他们对知识的吸收能程度也是参差不齐的，因此幼儿可以利用课外的时间进行知识的巩固与学习。针对本节活动，在教学结束之后，制作了微课，并将微课运用信息技术分享在班级群中，孩子们在家里能根据需要继续游戏学习，幼儿将在家游戏的过程分享到群中，教师有针对性地进行指导，这样的互动学习有利于幼儿掌握所学知识，同时也可以让家长了解孩子在幼儿园的学习，达成家园联动。

3.“程”就反思展望

3.1. 活动流程

一节教学活动的完成是需要多维度思考、斟酌、打磨的过程，本次活动由选材——构思教学设计——调整教学环节——制作课件——修改完善课件——开展教学活动——反思总结系列环节环环相扣而成，在每一个环节都要“慎思”以达到满足教学活动需求为最终目标。编程图形化界面显著降低认知难度，"虚实结合"方式有效提升幼儿的学习兴趣和协作能力，AI 技术为教学评估提供支持。

3.2. 教师方面

3.2.1. 看见每一位，听见每一位

对于幼儿园一线教师来说，要努力学习现代信息技术，不断学习，与时俱进。利用各种信息技术技术，创设活动情景，优化丰富教学活动，培养幼儿的语言表达能力、想象力、创造力和动手操作能力。在教学活动中充分调动幼儿学习兴趣，激发幼儿探究欲望，平等尊重的看待每一位幼儿，让每一位幼儿的想法都能得到实现，让幼儿都有实际操作的机会。

3.2.2 要与时俱进，不断地学习

作为一线幼儿园教师，信息化教学活动设计与实施是一种挑战，更是一个机会。当前存在的实际问题如:幼儿教师信息化应用能力的不足，在平时容易忽视了自身信息素养的积累与提升，无法全面掌握教学技能等等，这些都是需要我们慢慢解决的。作为新时代的幼儿教师，要与时俱进，不断学习先进的教育理念。

3.3. 幼儿方面

促进幼儿对科学世界的好奇好问。经过教学活动设计和玛塔编程等活动的开展，幼儿的科学素养得到了具体且全面的发展。从幼儿的“科学兴趣与态度”了解到中班幼儿对玛塔有

一定兴趣，有探索的欲望。从“科学知识与内容”分析出中班幼儿学会了前进指令和后退指令。从“科学方法与能力”方面，可以通过幼儿的实际操作，掌握编程的组合与排列，提升幼儿的计算思维和实践能力。

3.4. 家园方面

本节活动通过信息化的形式与家园的联动，也增强了家园之间的联系，家长能更深入地了解幼儿学习的知识，幼儿根据自己的兴趣可以选择再操作。教师也能针对个别差异进行再引导，对于幼儿掌握不清晰的地方可以很好地进行云上沟通，通过云上沟通达到每位幼儿对知识都能全方面的掌握。

信息技术的发展是随时代的改变不断进步提升的，因此数字化的教育也是不断延续的，针对科学绘本类的研究必须严谨，教学之前，对于编程中概念性的名词与内容教师一定要精准地把握，只有教师完全理解才能准确地将这些知识传递给幼儿。在信息技术融入教学过程中要辩证地看待，在适合的教学环节使用适宜的信息技术，而不能为了使用信息技术而使用信息技术，要坚持信息技术始终是为教学而服务的。未来将继续坚持对“幼儿园教育数字化应用案例研究”进行深入探究。

参考文献

杨广英.(2024).现代信息技术助力幼儿园美术教育质量提升的策略探索——以大班创意漆线雕活动为例.陕西学前师范学院学报,40(01),115-117.

周梦雨.(2023).情境化、趣味化、实践化—在电子白板技术支持下有效开展小班幼儿科学活动的实践与探索——以小班科学活动《小帽子》为例.(eds.)2023 年第一届生活教育学术论坛论文集(pp.571-573).南京市江宁区滨河幼儿园;

谢浩.(2021).数字化教育技术在幼儿园教育教学中的应用.(eds.)华南教育信息化研究经验交流会 2021 论文汇编（十）(pp.283-285).昆山市淀山湖中心幼儿园;

吴湘青.(2019).数字化教育技术在幼儿园教育教学中的应用.教育信息化论坛,3(12),245-246.

计算思维在幼儿艺术教育中的创新实践

Innovative Practices of Computational Thinking in Early Childhood Art Education

虞哲夫[1*]，张曦[1]
温州大学附属实验幼儿园
532761158@qq.com

【摘要】 计算思维是21世纪儿童核心素养之一，其培养已成为幼儿园课程创新的重要方向。本文结合一线教学经验，探讨计算思维在幼儿艺术教育，特别是音乐律动课程中的实践应用。以“超级马里奥”音乐律动课程为例，分析计算思维核心概念如何在艺术活动中自然融入，并通过教学实践提升幼儿的计算思维能力。研究发现，音乐律动的节奏模式、动作编排与程序化思维存在高度契合点，能够帮助幼儿在游戏化学习中建立计算思维的初步概念。同时，本文总结了该类课程的实施经验，提出优化策略，并探讨其向其他艺术领域拓展的可能性，为幼儿园教师提供实践指导。

【关键词】 计算思维；音乐律动；艺术教育；课程创新

Abstract: *Computational thinking is a core competency for 21st-century children and a key focus in innovating kindergarten curricula. This paper explores how computational thinking can be integrated into early childhood art education, with a focus on music and movement. Taking the *Super Mario* music and movement course as an example, it demonstrates how core CT concepts—such as sequencing, patterns, and procedural thinking—can naturally align with rhythm and movement activities. The study shows that game-based learning in this context helps young children develop basic computational thinking skills. The paper also shares practical implementation experiences, offers strategies for course optimization, and discusses possibilities for extending this approach to other areas of art education, providing practical guidance for early childhood educators.*

Keywords: Computational Thinking, Music and Movement, Art Education, Curriculum Innovation

1.前言

近年来，计算思维作为一种重要的思维方式，逐渐受到基础教育阶段的关注（Wing, 2006）。尽管它最初主要应用于计算机科学领域，但其中的核心概念——如抽象、模式识别、逻辑推理和问题分解，其实在许多日常活动和学科学习中都能找到对应（Brennan & Resnick, 2012）。例如，在幼儿教育中，计算思维的培养并不意味着直接教授编程技术，而是通过游戏化、情境化 的方式，引导孩子建立基本的逻辑推理和问题解决能力。音乐律动课程天然具备节奏性、结构化和重复性，这些特性与计算思维中的模式识别和算法思维 有着有趣的契合点。例如，幼儿在学习一段舞蹈时，需要识别和记忆固定的节奏模式（类似循环结构）；在调整动作顺序时，需要运用逻辑推理 来优化编排；而在音乐游戏中寻找最佳解决方案的过程，更是计算思维的一种实践体现（周以真，2021）。本研究以幼儿园“超级马里奥”音乐律动课程为例，探索如何在艺术教育中自然融入计算思维，并分析这种融合模式的实际效果。首先，我们将讨论音乐律动与计算思维的内在关联，然后结合具体课堂案例，分析一线

教师的实践策略，最后探讨如何将这种教学模式拓展到其他艺术领域，并提出优化建议，以期为幼儿园教师提供更具操作性的教学思路。

2.幼儿音乐律动融合课程培养幼儿计算思维的可行性

在幼儿园课程中，音乐律动不仅是一种艺术表达形式，更是促进幼儿多维度发展的有效载体。其节奏的结构性、动作的模式化以及游戏化的体验，都为计算思维的培养提供了天然的优势（林怡君，2018）。结合一线教学实践，我们从以下两个方面探讨音乐律动课程如何促进幼儿计算思维的发展。

2.1 . 音乐律动内在的规律性与计算思维核心素养的契合

音乐律动的核心在于节奏和模式，而计算思维的关键是模式识别、逻辑推理和算法思维，这两者在结构上存在天然的联系（Brennan, K., & Resnick, 2012）。例如，在“超级马里奥”音乐律动课程中，幼儿通过学习和重复节奏型，可以直观感受到“循环”这一计算思维的重要概念。在实际教学中，我们尝试用更加直观的方式 帮助幼儿理解这一概念。例如，教师在课堂上使用 Scratch Jr. 编排简单的音乐节奏，并让孩子们观察其中的重复模式。相较于传统的节奏图谱，这种可视化方式更容易让幼儿理解“事项 × 次数”这一循环思维的核心（Resnicket al.,2009）。此外，音乐的层次感、节奏快慢变化，以及不同动作组合方式，也对应着计算思维中的抽象与分解能力。在音乐律动过程中，幼儿需要遵循一定的规则，如固定的节奏型、重复的旋律结构、先后顺序的动作安排等。这些规律不仅有助于幼儿理解音乐，也让他们在潜移默化中建立逻辑思维的框架。例如，在“超级马里奥”课程中，孩子们需要根据不同的音乐片段调整自己的律动动作，而这一过程本质上与计算机程序设计中的“条件判断”相似（周以真，2021）。当音乐变化时，幼儿需要快速调整动作模式，这种动态适应能力正是逻辑思维的一种体现。

2.2. 音乐律动游戏的趣味性与情境性能够激发幼儿计算思维的内在驱动

计算思维的培养并非单纯的知识传授，而是需要激发幼儿的学习兴趣。游戏化的音乐律动活动正好提供了一种生动有趣的方式，让孩子们在探索和挑战中自然接触计算思维。以“超级马里奥”课程为例，教师将音乐律动设计成“闯关挑战”，让幼儿在完成不同节奏任务后“通过关卡”。每个音乐片段都对应着一组特定的律动指令，例如：

（1）当音乐加快时，孩子们需要跟上节奏，完成快速踏步；

（2）当音乐减慢时，他们需要调整动作，模拟马里奥蹲伏前进；

（3）某些特殊旋律段落则需要做出跳跃等特定动作。

这一过程中，幼儿不仅沉浸在游戏的乐趣中，还在无形中练习了计算思维的核心概念，例如：

（1）顺序结构：按照音乐的节奏依次执行动作；

（2）条件判断：根据不同的音乐变化调整动作；

（3）循环结构：重复特定节奏模式并在多个回合中强化记忆。

此外，音乐律动本身带有情境感，如“超级马里奥”课程中，教师利用动画、故事、背景音乐等元素，将计算思维的训练嵌入艺术活动中，让幼儿在生动的体验中自然理解计算逻辑。这种方式区别于传统的指令式教学，更符合幼儿的认知特点，也更容易引起他们的兴趣和主动探索（Barsalou,2008）。

3.基于计算思维培养的幼儿园艺术融合课程设计

3.1 艺术实践课程中培养计算思维的理论依据

在面向大班幼儿的艺术实践课程中融入计算思维培养，需建立在对儿童认知发展阶段和学习方式的深刻理解之上。皮亚杰（Piaget，1972）的认知发展理论指出，5 至 6 岁儿童正处于“具体运算阶段”的早期，他们开始具备基本的逻辑推理能力，能理解事物间的因果关系和初步的分类、序列概念。此阶段儿童思维的显著特点是对具体操作和直观经验的依赖，而非抽象推理。因此，任何以培养“抽象的”计算思维为目标的教学，都必须通过具体、可感知、可操作的方式展开。在这一基础上，具身学习理论（Embodied Learning）为艺术课程中计算思维的渗透提供了理论支撑。该理论认为，身体运动与认知发展密切相关，学习不是仅发生在大脑中，而是通过全身感官与环境的互动过程建构起来的（Barsalou, 2008）。在音乐律动或视觉艺术等课程中，儿童通过身体动作与节奏、空间、形状等元素互动，有助于抽象概念（如顺序、循环、模式）的具体化与内化。这种“动中学”“玩中学”的方式，恰好契合大班幼儿通过具象经验发展逻辑思维的特点。此外，从信息加工角度出发，计算思维的培养并不等于“教孩子编程”，而是通过对问题的拆解、信息的排序、逻辑的推理等思维活动，帮助儿童形成良好的认知策略。而艺术课程特别是在音乐律动活动中，天然就包含大量的节奏重复、模式识别、动作规划等活动，能够在不知不觉中提升儿童的信息处理能力。因此，在艺术实践课程中开展计算思维的教学，不仅具备心理发展上的适宜性，也拥有理论层面上的支持。通过引导儿童在艺术活动中进行“操作性思维”和“身体化推理”，既尊重了儿童的发展规律，也为后续更高阶的抽象计算能力奠定了坚实基础。

3.2. 融合课程的目标设定

为了有效促进幼儿计算思维的培养，本课程以“计算思维启蒙”为核心目标，并结合艺术教育的特点，设定以下三维目标：

3.2.1. 计算观念

帮助幼儿形成计算思维的基本意识，例如通过音乐节奏感知模式变化(Grover & Pea,2013)。

3.2.2. 计算概念

让幼儿在游戏中理解基本的计算思维概念，如顺序、条件、循环等。例如，利用律动游戏设置“重复动作=循环指令”的概念，使幼儿在音乐实践中体验编程逻辑，通过舞蹈动作理解重复与循环（Brennan & Resnick, 2012）。

3.2.3. 计算实践

鼓励幼儿在艺术活动中尝试计算思维的方法，例如在音乐律动游戏中编排动作顺序、设计节奏模式，甚至运用 Scratch Jr 等工具进行可视化编程表达（Resnicket al.,2009）。

3.3. 融合课程的设计原则

在实际课程设计中，我们遵循以下四大原则，以确保计算思维的有效融入：

3.3.1. 幼儿发展性原则

课程内容需符合幼儿的认知特点和发展水平，计算思维的概念需要通过游戏化、直观化的方式呈现，如借助图示、身体律动、简单的编程模块等方式，让幼儿在互动中理解抽象概念（Piaget，1972）。

3.3.2. 趣味性原则

计算思维的培养需要建立在儿童喜爱的活动基础上，游戏化的体验能提高幼儿的学习兴趣，例如通过角色扮演、闯关挑战、音乐编排等方式增强趣味性（Hirsh-Pasek et al.，2009）。

3.3.3. 挑战性原则

课程应具备适度的挑战，让幼儿在尝试和探索中逐步建立计算思维。例如，在律动游戏中，教师可以设置不同难度的节奏任务，鼓励幼儿通过分析和调整完成挑战。

3.3.4. 渐进性原则

计算思维的培养需要遵循由浅入深的过程，课程应设计循序渐进的学习环节，例如先从简单的节奏模式入手，再逐步引入循环、条件等更复杂的思维概念(Grover & Pea,2013)。

3.4. 课程的实施与成效

在实际教学过程中，我们围绕“超级马里奥”主题音乐律动活动，尝试将计算思维元素自然融入到课程设计中，并通过幼儿的课堂表现和学习反馈评估其成效。本部分从课程的具体实施流程和教学成效两个方面进行梳理。

3.4.1. 课程实施：从律动游戏到计算思维培养

（1）课程导入：建立沉浸式游戏情境

为了激发幼儿的学习兴趣，我们采用“超级马里奥闯关”为主题，将课程设计为一场冒险游戏。教师首先通过播放《玛丽波尔卡》音乐，引导幼儿聆听音乐节奏，同时结合图片、动画等多种方式，让幼儿了解游戏的基本背景，如马里奥如何通过跳跃、躲藏、翻滚等动作完成闯关任务。

图 1

这一环节的核心目标是建立情境感，让幼儿在代入游戏角色的过程中，自然关注到动作与节奏的关系，为后续的计算思维训练做铺垫。

（2）探索环节：初步感知计算思维概念

教师引导幼儿观察音乐中的节奏模式（如动作轻重变化、重复循环等），并结合动作进行尝试。例如，在“跨越怪兽”环节，教师会引导幼儿发现，跨越相同类型障碍时，可以使用相同的动作模式（如“跳-跳-跳”），而这个模式的重复即对应了计算思维中的“循环”概念。

融合计算思维提问示例：

表 1

计算思维概念	教师提问示例
模式识别	“你发现这个音乐片段和刚才的有什么相同的地方？”
循环结构	“如果我们要再重复一次这个动作，

	该怎么做才不会太累？”
条件判断	“如果音乐变快了，你的动作需要发生什么变化？”
调试优化	“你的动作能不能调整的更简单一些？有什么更好的办法？”

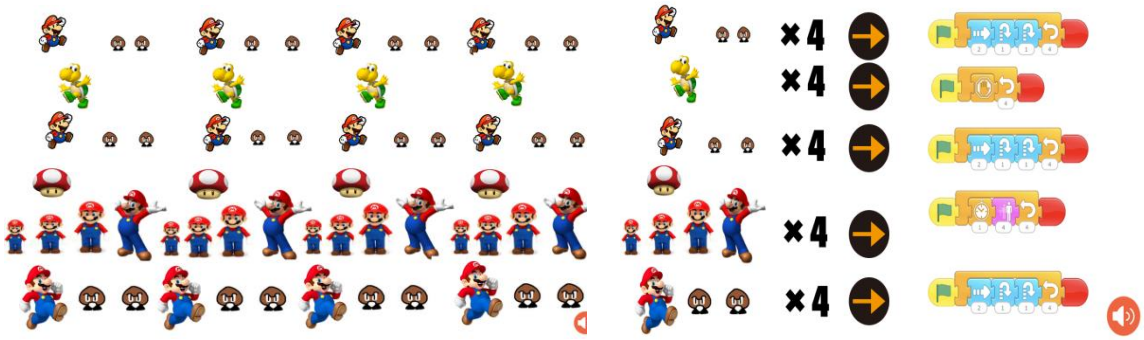


图 2 课件截图示例：由图谱转换为程序模块

此外，教师还鼓励幼儿自主思考：

1、如果遇到新的怪物，该如何调整动作？（调试与优化）

2、不同的段落节奏会影响动作的执行方式吗？（模式识别）

（3）自主创编环节：实践计算思维

在幼儿初步理解计算思维元素后，教师进一步引导他们自主设计新的闯关律动。例如，一组幼儿可能设计出“前进-跳跃-躲避”模式，另一组则可能尝试“躲避-跳跃-前进”模式。

在此过程中，教师通过 Scratch Jr 等工具，引导幼儿将自己的律动规则“编程”出来，帮助他们理解动作序列的逻辑性和可重复性，强化计算思维能力。例如：

1、让幼儿尝试使用“循环模块”优化动作组合，如“（踏步走×4+原地起跳×2）”代替“走-走-走-走-跳-跳”。

2、让幼儿通过条件判断调整规则，例如“如果音乐变快，就增加跳跃次数”。

（4）总结与回顾：反思与调整

在课程的最后，教师带领幼儿进行回顾总结，让他们分享自己在活动中的发现。例如：

1、“哪些地方可以用重复的动作完成？”（巩固循环概念）

2、“如果想要调整动作顺序，应该怎么做？”（培养调试与优化能力）

3、“你能设计一个新的闯关模式吗？”（鼓励创新与逻辑思维）

这一环节的核心目标是促进幼儿的思维迁移，让他们不仅在音乐律动活动中应用计算思维，同时也能在其他游戏或日常活动中运用类似的逻辑思考方式。

3.4.2 课程成效：幼儿计算思维能力的提升

在课程结束后，我们通过课堂观察、教师访谈和幼儿作品分析等方式，对幼儿的计算思维发展情况进行了评估，并发现以下几个主要成效：

（1）幼儿对“模式识别”与“循环”的理解显著增强

在初期活动中，幼儿往往是凭借直觉模仿教师的动作，而对动作背后的规律缺乏明确认知。然而，在课程实施后，大多数幼儿能够主动发现音乐节奏中的重复模式，并能够运用“循环”概念优化自己的律动设计。例如，在后续活动中，幼儿已经能够自发使用“（走×4-跳×2）”替代“走-走-走-走-跳-跳”这样的冗长动作序列，表现出计算思维的萌芽。

（2）幼儿在“调试与优化”方面的能力有所提升

在课程初期，幼儿在编排律动时常常不加调整地模仿教师给出的示范。但在后续活动中，我们观察到，许多幼儿开始主动调整自己的动作顺序。例如，有的幼儿发现“跳跃后立即拍手”比“跳跃-停顿-拍手”更流畅，于是主动修改动作设计。这一现象表明，幼儿已经初步具备了问题分析与优化的能力，这正是计算思维的重要组成部分。

4.总结

本课例的实施，不仅让幼儿在音乐律动活动中体验到了计算思维的基本概念，如模式识别、循环结构、问题分解与优化等，同时也激发了他们的创造力和团队协作能力。此外，教师的教学理念也得到了新的启发，对计算思维与艺术教育的融合有了更深入的理解。然而，我们也发现，计算思维的培养需要长期的积累，仅凭一次课程无法达到系统化培养的效果。因此，未来的教学中，我们将继续探索如何在日常活动中持续渗透计算思维训练，并通过动态调整教学策略，使课程更符合幼儿的发展特点，真正实现计算思维与艺术教育的深度融合。

参考文献

何婷婷. (2023). 幼儿园音乐律动课程对计算思维培养的影响研究. 音乐教育研究, 48(6), 67-75.

李泽湘. (2021). 幼儿计算思维的游戏化教学模式研究. 学前教育研究, 39(5), 18-25.

王芳. (2022). 计算思维视角下的幼儿园课程创新. 学前教育论坛, 44(1), 55-62.

周以真. (2019). 计算思维：从概念到课程实施. 全球教育展望, 48(2), 25-32.

周以真.（2021）.《幼儿计算思维能力发展的路径研究》. 华东师范大学硕士论文.

张伟. (2020). 人工智能背景下计算思维培养的路径探索. 中国教育技术学刊, 27(3), 33-40.

林怡君.（2018）.《从艺术课程中探索幼儿计算思维的萌芽：节奏与逻辑的融合路径》.《学前教育研究》，(6)，42－45 页.

林崇德（2009）。《发展心理学》。北京：人民教育出版社。

陈会昌、李虹（2019）。《儿童心理发展与教育》。北京师范大学出版社。

Brennan, K., & Resnick, M. (2012). Computational thinking in creative learning environments. International Journal of Computer Science Education, 7(3), 12-23.

Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617 – 645. Csikszentmihalyi, M. (2017). Flow: The psychology of optimal experience. China Renmin University Press.

Grover, S., & Pea, R. (2013). Computational thinking in K – 12: A review of the state of the field. Educational Researcher, 42(1), 38 – 43. Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). A mandate for playful learning in preschool: Presenting the evidence. Oxford University Press.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). A mandate for playful learning in preschool: Presenting the evidence. Oxford University Press.

Papert, S. (2015). Mindstorms: Children, computers, and powerful ideas. East China Normal University Press.

Piaget, J.（1972）. The Origins of Intelligence in Children. New York: International Universities Press.

Resnick, M. (2020). Lifelong kindergarten: Cultivating creativity through projects, passion, peers, and play. Posts & Telecom Press.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., ⋯ & Kafai, Y. (2009). Scratch: Programming for all. Communications of the ACM, 52(11), 60 – 67.

Google Translate 圖像報讀功能對讀寫障礙的閱讀理解之學習成效

The Learning Effectiveness of the Google Translate Image Reading-Aloud on Reading Comprehension for Students with Dyslexia Function

張羽萱[1,*]，洪榮昭[2]，蔡其瑞[3]
[1]國立臺師範大學 創造力發展碩士在職專班 學生
2 國立臺灣師範大學 工業教育學系 教授
3 國立臺灣師範大學 科技運用與人類資源發展學系 助理教授
[*] yushuan0313@gmail.com

【摘要】 本研究旨在探討 Google Translate 圖像報讀功能對於讀寫障礙學生在閱讀理解學習成效上的影響，並透過訪談分析其學習焦慮、自我效能及使用態度變化。本研究採用跨受試的單一 A-B-A'實驗教學設計，選取三位新北市國小四年級資源班的讀寫障礙學生，透過 Google Translate 圖像報讀功能進行國語文閱讀教學。結果顯示，該功能能提升學生閱讀理解能力，部分學生雖在維持期未持續進步，但整體學習成效仍達大效果量，展現其增強閱讀理解的潛力。

【關鍵字】 **Google Translate**；報讀；讀寫障礙；閱讀理解

Abstract: *This study aims to investigate the impact of the Google Translate image read-aloud function on reading comprehension outcomes for students with dyslexia. Through interviews, the study also examines changes in students' learning anxiety, self-efficacy, and attitudes toward using the tool. A multiple-baseline A-B-A' single-subject experimental design was adopted, involving three fourth-grade students with dyslexia from a resource class in New Taipei City. Mandarin reading instruction was conducted using the Google Translate image read-aloud function. The results indicate that the tool effectively improved students' reading comprehension. Although some students did not continue to improve during the maintenance phase, the overall learning outcomes still reached a large effect size, demonstrating the tool's potential to enhance reading comprehension.*

Keywords: Google Translate, image reading-aloud, dyslexia, reading comprehension

1.前言

隨著數位科技與人工智慧（AI）快速發展，電子化學習輔助工具已成為學習的重要支援工具。《身心障礙者權利公約》（Convention on the Rights of Persons with Disabilities, CRPD）(2006)中提出「通用設計」及「合理調整」原則，強調培養身心障礙者的自我倡議與決策能力。Wehmeyer 與 Schwartz（1997）也指出，具有認知或學習障礙的自我決策學生更有可能達到積極的成人結果(Field & Hoffman, 2002)，然而，學習障礙學生普遍面臨語文學習困難，其中最常見的挑戰即是識字量的不足(江毓鈞、楊曉玲，2012)，進而影響閱讀理解表現及學習動機。

根據國際讀寫障礙協會(The International Dyslexia Association,簡稱 IDA)，讀寫障礙的特徵在於單詞辨識的準確性和流暢性有困難，並且拼寫和解碼能力較差，這些困難可能衍生的

問題可能包括閱讀理解與減少閱讀經驗，以致妨礙字彙量成長與背景知識的增加(王瓊珠、洪儷瑜、陳秀芬，2007；Lyon et al., 2003）。由此可知，識字對閱讀確實有相當程度的影響(張長穎，2013)。

近年來的教育發展趨勢中，數位科技工具的應用扮演著關鍵角色。透過聲光效果輔助，數位工具能引導學生運用多感官學習，藉此有機會透過優勢的學習管道而有較佳的學習成效(謝慧如，2009)。其中，Google Translate 是一款功能多元、免費且有多種感官支援的應用程式，不僅提供文字輸入、語音、拍照、離線翻譯功能，亦可透過拍照辨識並報讀圖片中的文字，有助於提升閱讀理解(Wirantaka1 & Fijanah, 2021)。Kumnoed (2018)的研究發現，透過 Google Translate 的翻譯，學生能夠根據 Smith（1982）提出的閱讀理解層次，在字面理解層次上準確的理解段落內容。因此，學生近年來常利用 GT 來協助翻譯、寫作、學習新詞彙和閱讀理解（Wirantaka & Fijanah, 2021）。

基於上述背景與研究發現，本研究旨在探討 Google Translate 圖像報讀功能對於讀寫障礙學生在閱讀理解學習成效的影響，並透過訪談分析其學習焦慮感、自我效能感及使用態度的變化，以評估該工具在特殊教育中的的應用價值。

2.研究問題

基於上述研究目的，本研究之研究問題如下：

2.1. Google Translate 圖像報讀功能對讀寫障礙學生之閱讀理解表現是否具有立即成效？

2.2. Google Translate 圖像報讀功能對讀寫障礙學生之閱讀理解表現是否具有維持成效？

3.研究設計

3.1. 研究架構

本研究欲採用跨受試的單一 A-B-A’實驗教學研究設計，主要探討「Google Translate 圖像報讀功能對讀寫障礙學生之閱讀理解表現」，並依據 PIRLS 閱讀四層次設計閱讀測驗題目，進一步了解 Google Translate 圖像報讀功能對讀寫障礙學生之閱讀理解表現的立即成效與維持成效，且在教學後透過訪談了解研究參與者對於 Google Translate 圖像報讀功能的使用感受及該功能介入前後對閱讀感受的變化。

3.2. 研究程序與對象

本研究採跨受試的單一 A-B-A’實驗教學研究設計。在基線期(A)，先進行「2019 閱讀理解測驗」及三次自編 PIRLS 閱讀理解測驗，作為能力檢核的前測，再者，進入介入期(B)，實施為期三週的 Google Translate 圖像報讀功能教學，每週進行三次教學與評量，最後，在維持期(A’)，於教學結束一週後進行三次後測，並進行個別訪談，深入了解學習者對於 Google Translate 圖像報讀功能的使用感受及該功能介入前後對閱讀體驗的變化。

本研究之研究參與者為三位國小四年級讀寫障礙生，應符合以下條件：(1)新北市特殊教育學生鑑定及就學輔導委員會鑑定為學習障礙學童；(2)魏氏兒童智力量表之全量表分數顯示個案智力正常;(3)2019 閱讀理解測驗之測驗結果未達四年級切截標準。此外，排除條件包括：(1)無法依據教師指令配合；(2)挫折容忍度低而情緒失控。

3.3. 研究工具

本研究以「多感官學習」理論作為研究依據，採用「Google Translate 圖像報讀功能」、「四年級國語文教學教材」，並搭配自編的四層次閱讀測驗題作為研究工具，分別應用於前測、後測與追蹤三階段，以評估學生在閱讀理解方面的立即學習效果及維持學習效果。此外，

本研究亦設計了自編訪談自評表-「Google Translate 圖像報讀功能對閱讀理解學習成效之自評表」，以深入了解研究參與者對於 Google Translate 圖像報讀功能的使用經驗及該功能介入前後對閱讀感受的變化。研究工具說明如下：

3.3.1. Google Translate 圖像報讀功能

Google Translate 是由是由 Google 開發的一款多功能翻譯工具，其中的圖像報讀功能結合了光學字符辨識（OCR）技術，可將圖片中的文字轉換為數位文本，並進行即時翻譯與語音合成，讓使用者透過聆聽方式獲取翻譯結果。本研究充分運用了此功能的視覺與聽覺整合特性，透過照片選取操作，讓研究參與者能直接聆聽文本內容，進一步提升閱讀理解的便利性與效果。

3.3.2. 自編閱讀測驗評量工具

本研究考量研究參與者為讀寫障礙學生，且智力範圍落於 79~83，智力屬於中下的水平，因此選用難度偏易的閱讀文本作為教學素材，核心閱讀文本包括教育部編撰的四年級國語文教學教材及南一版四年級課文，閱讀內容涵蓋故事、記敘文等多種文體類型，並依據 PIRLS 四層次閱讀理解框架，自編設計測驗題目作為研究評量工具，閱讀測驗共計 10 題，包含直接提取（3 題）、直接推論（3 題）、詮釋與整合觀點及訊息（2 題）、評估與批判內容（2 題）。透過該評量工具，評估研究參與者在閱讀理解方面的學習成效。

3.3.3. Google Translate 圖像報讀功能對閱讀理解學習成效之自評表

為深入了解研究參與者對 Google Translate 圖像報讀功能的主觀感受，研究者設計了一份自評表，涵蓋「學習焦慮」、「學習興趣」、「閱讀主動性」、「自我效能感」及「Google Translate 使用態度」等五大構面。自評表採開放式問答形式，旨在蒐集質性資料，進一步分析學生在教學介入前後，於閱讀理解方面的感受與變化。

3.4. 資料處理與分析

本研究採用目視分析進行資料處理，並運用 C 統計及效果值分析，檢視基線期、介入期與維持期各階段內及階段間的變化情形，評估研究參與者使用 Google Translate 圖像報讀功能在閱讀理解學習成效上的差異。此外，本研究亦探討 Google Translate 圖像報讀功能對學習心態變化的影響，分析是否存在顯著差異。

4.研究結果

本研究旨在探討 Google Translate 圖像報讀功能對讀寫障礙學生閱讀理解表現的學習成效，以及教學介入後的維持成效。研究結果以目視分析、C 統計及效果值分析進行呈現，並透過行為反應學習曲線圖分別展示三位研究參與者的學習表現與變化情形。研究結果分述如下：

4.1. 研究參與者甲之目視分析與統計分析結果

由圖一可得知，研究參與者甲透過 Google Translate 圖像報讀功能對閱讀理解學習成效的基線期水準全距是 0-0，階段內水準變化為 0，階段內平均水準為 0，呈穩定的水平、等速趨勢(水準和趨勢穩定度皆為 100%)。由於研究參與者甲透過 Google Translate 圖像報讀功能對閱讀理解學習成效的階段內平均水準為 0，呈現穩定的未改善趨勢，是受到研究參與者甲無法獨立識字而影響，故進入介入期。

進入介入期後，透過 Google Translate 圖像報讀功能對閱讀理解學習成效的水準全距為 50-80，階段內水準變化為 0，階段內平均水準為 64.44，呈現不穩定上升、進步趨勢(水準和趨勢穩定度分別為 55.56%和 44.44%)。

從基線期至介入期，階段間水準變化為 60，平均水準變化為 64.44，趨向變化和效果為正向，重疊率為 0%；C 值為.740，z 值為 2.797，達.05 的顯著水準，此結果表示從基線期至介入期，閱讀理解學習成效有增加趨勢，它排除因機遇或誤差造成的結果，拒絕虛無假設而它為「真」時的誤差小於 5%。另計算效果量，由於欲介入期有明顯的趨勢，所以採用迴歸效果量，採「斜率改變之迴歸效果量」，檢定 Google Translate 圖像報讀功能與閱讀理解學習成效間關係強度後發現，效果量指標為f^2為.0061，超過 0.02 小效果量的值，但未達 0.15 中效果量的值，表示 Google Translate 圖像報讀功能對閱讀理解學習成效具有小的學習成效。由此可知，Google Translate 圖像報讀功能對研究參與者甲閱讀理解學習成效具有立即的學習成效，且達小效果量。

再者，從介入期至維持期，階段間水準變化為 0，平均水準變化為 2.23，趨向變化和效果為正向，重疊率為 100%；基線期和維持期階段間的 C 值為.725，z 值為 2.145，達.05 的顯著水準，此結果表示從介入期至維持期，閱讀理解學習成效有增加趨勢，它排除因機遇或誤差造成的結果，拒絕虛無假設而它為「真」時的誤差小於 5%。另計算維持期的效果量，採用標準化平均差異量，採「斜率改變之迴歸效果量」，檢定 Google Translate 圖像報讀功能與閱讀理解學習成效間關係強度後發現，標準化平均差異量 d 為 16.330，超過 0.80 大效果量的值，表示 Google Translate 圖像報讀功能對閱讀理解學習成效具有大的學習成效。由此可知，Google Translate 圖像報讀功能對研究參與者甲閱讀理解學習成效具有維持的學習成效，且達大效果量。

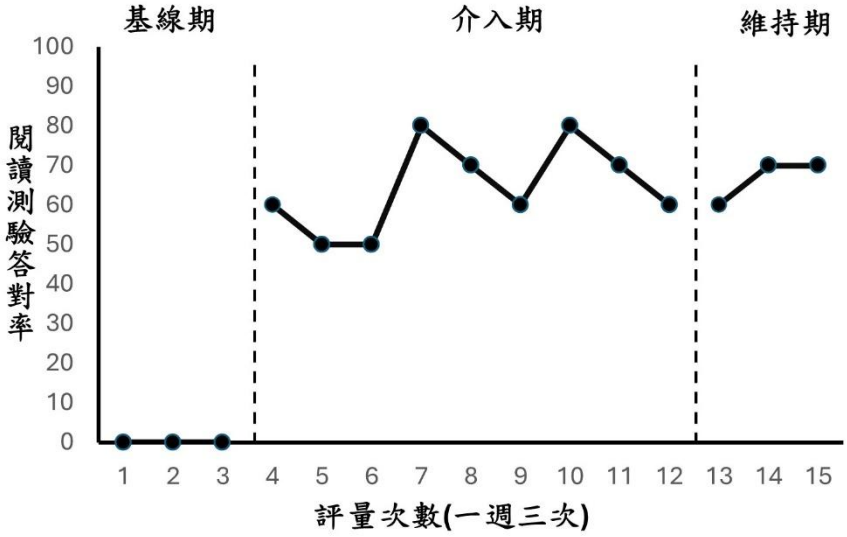


圖 1 Google Translate 圖像報讀功能對研究參與者甲閱讀理解學習成效之成效曲線圖

4.2. 研究參與者乙之目視分析與統計分析結果

由圖二可得知，研究參與者乙透過 Google Translate 圖像報讀功能對閱讀理解學習成效的基線期水準全距 50-70，階段內水準變化為 10，階段內平均水準為 60，呈穩定的上升趨勢(水準和趨勢穩定度分別為 33.33%和 66.67%)。由於研究參與者乙透過 Google Translate 圖像報讀功能對閱讀理解學習成效的階段內平均水準為 60，呈現不穩定的未改善趨勢，可能受到研究參與者乙閱讀時會跳行漏字及環境因素而影響，故進入介入期。

進入介入期後，透過 Google Translate 圖像報讀功能對閱讀理解學習成效的水準全距為 70-90，階段內水準變化為 10，階段內平均水準為 76.67，呈現穩定上升、進步趨勢(水準和趨勢穩定度皆為 88.89%)。

從基線期至介入期，階段間水準變化為 0，平均水準變化為 16.67，趨向變化和效果為正向，重疊率為 44.44%；C 值為.673，z 值為 2.547，達.05 的顯著水準，此結果表示從基線期至介入期，閱讀理解學習成效有增加趨勢，它排除因機遇或誤差造成的結果，拒絕虛無假設而它為「真」時的誤差小於 5%。另計算效果量，由於欲介入期有明顯的趨勢，所以採用迴歸效果量，採「斜率改變之迴歸效果量」，檢定 Google Translate 圖像報讀功能與閱讀理解學習成效間關係強度後發現，效果量指標為f^2為.00561，超過 0.02 小效果量的值，但未達 0.15 中效果量的值，表示 Google Translate 圖像報讀功能對閱讀理解學習成效具有小的學習成效。由此可知，Google Translate 圖像報讀功能對研究參與者乙閱讀理解學習成效具有立即的學習成效，且達小效果量。

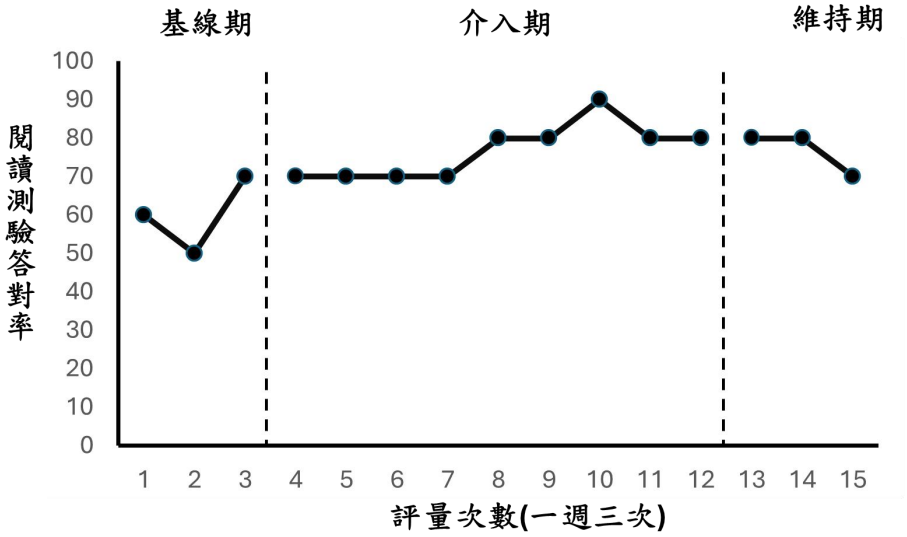


圖 2 Google Translate 圖像報讀功能對研究參與者乙閱讀理解學習成效之成效曲線圖

4.3. 研究參與者丙之目視分析與統計分析結果

由圖三可得知，研究參與者丙透過 Google Translate 圖像報讀功能對閱讀理解學習成效的基線期水準全距 40-70，階段內水準變化為 30，階段內平均水準為 60，呈穩定的上升趨勢(水準和趨勢穩定度分別為 0%和 66.67%)。由於研究參與者丙透過 Google Translate 圖像報讀功能對閱讀理解學習成效的階段內平均水準為 60，呈現不穩定上升趨勢，可能受到研究參與者丙閱讀時會跳行漏字及環境因素而影響，故進入介入期。

進入介入期後，透過 Google Translate 圖像報讀功能對閱讀理解學習成效的水準全距為 60-90，階段內水準變化為 30，階段內平均水準為 81.11，呈現穩定上升、進步趨勢(水準和趨勢穩定度皆為 77.78%)。

從基線期至介入期，階段間水準變化為-10，平均水準變化為 21.11，趨向變化和效果為正向，重疊率為 22.22%；C 值為.458，z 值為 1.732，達.05 的顯著水準，此結果表示從基線期至介入期，閱讀理解學習成效有增加趨勢，它排除因機遇或誤差造成的結果，拒絕虛無假設而它為「真」時的誤差小於 5%。另計算效果量，由於欲介入期有明顯的趨勢，所以採用迴歸效果量，採「斜率改變之迴歸效果量」，檢定 Google Translate 圖像報讀功能與閱讀理解學習成效間關係強度後發現，效果量指標為f^2為.4557，超過 0.35 大效果量的值，表示 Google Translate 圖像報讀功能對閱讀理解學習成效具有顯著的學習成效。由此可知，Google Translate 圖像報讀功能對研究參與者丙閱讀理解學習成效具有立即的學習成效，且達大效果量。

再者，從介入期至維持期，階段間水準變化為 0，平均水準變化為-1.11，趨向變化和效果為負向，重疊率為 100%；基線期和維持期階段間的 C 值為.464，z 值為 1.373，未達.05 的顯著水準，此結果表示從介入期至維持期，閱讀理解學習成效雖有下降趨勢，但無法排除因機遇或誤差造成的結果，未能拒絕虛無假設。另計算維持期的效果量，採用標準化平均差異量，採「斜率改變之迴歸效果量」，檢定 Google Translate 圖像報讀功能與閱讀理解學習成效間關係強度後發現，標準化平均差異量 d 為 1.155，超過 0.80 大效果量的值，表示 Google Translate 圖像報讀功能對閱讀理解學習成效具有大的學習成效。由此可知，Google Translate 圖像報讀功能對研究參與者丙閱讀理解學習成效在維持期雖然有下降趨勢，但學習成效仍達到大效果量。

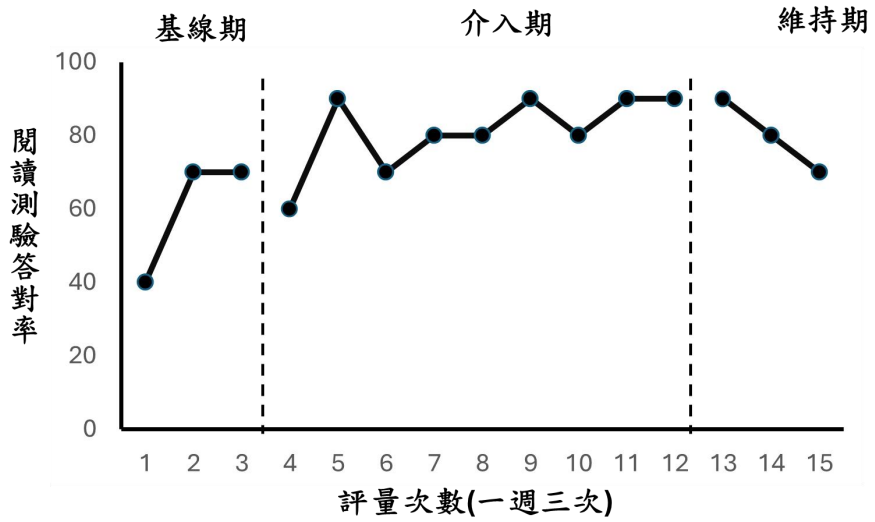


圖 3 Google Translate 圖像報讀功能對研究參與者丙閱讀理解學習成效之成效曲線圖

4.4. 額外發現

本研究除了量化資料的分析，亦採用「訪談」的方式深入了解三位研究參與者對於研究過程與研究工具使用之感受，再者，由於三位研究參與者皆為讀寫障礙的學生，因此表達方面較不流暢，過程中採「詢答」的方式，協助其回答問題。旨在探討研究參與者對於「Google Translate 圖像報讀功能」在閱讀理解學習上的感受變化與使用看法，分述如下：

4.4.1. Google Translate 圖像報讀功能對閱讀理解之學習焦慮的影響

三位研究參與者在學習焦慮感受上呈現不同程度的差異，研究參與者甲與乙表示，他們在閱讀測驗上都因為害怕答題錯誤而感到焦慮，但研究參與者丙則表示：「看不懂字時會感到困惑，但並不會覺得焦慮，因為閱讀理解錯誤沒關係，事後訂正理解就好。」這顯示研究參與者丙的學習焦慮感相較於研究參與者甲與乙較低。透過「Google Translate 圖像報讀功能」介入後，三位研究參與者一致認為，該功能能幫助解決識字量不足或部分不懂得字詞句所造成的閱讀理解困難，且此功能不僅能降低學習焦慮感，還促使他們更願意獨立進行課外閱讀，而非依賴他人的協助。綜合上述可得知，三位研究參與者對學習焦慮有下降的趨勢。

4.4.2. Google Translate 圖像報讀功能對閱讀理解之自我效能感的影響

三位研究參與者一致認為，Google Translate 的圖像報讀功能，幫助他們解決識字量不足或部分不懂得字詞句而導致閱讀理解困難的問題，進而提升了他們對閱讀的自信心。在訪談中，研究參與者也具體反映了這一點，首先，研究參與者甲表示：「Google 可以唸難的，我一下子就知道了。」，再者，研究參與者乙提到：「沒聽懂的可以重聽，看一次聽一次可以

重複確認訊息。」，最後，研究參與者丙則說：「它可以幫我讀我不會的，知道怎麼唸就可以幫助理解。」，上述回應顯示，三位研究參與者不僅能熟練使用該功能，還相信自己能藉助工具獨立解決閱讀問題。由此可知，三位研究參與者對於閱讀理解的自我效能感均有上升的趨勢。

4.4.3. 研究參與者對使用 Google Translate 圖像報讀功能的使用態度

在訪談時，三位研究參與者均認為 Google Translate 圖像報讀功能的操作流程簡單且便利，研究參與者們一致提到：「相機拍完照片後，進入 Google 翻譯尋找照片，並選取文字範圍就可以聽文章了。」，這表明該研究工具操作步驟簡易，且按鍵圖示淺顯易懂。此外，三位研究參與者均表示，若有認識閱讀困難的同儕時，他們會推薦該研究工具，因為該功能能幫助使用者透過聆聽更好的理解文章內容，並且能接觸到更多元的書籍。由此可知，三位研究參與者對於 Google Translate 圖像報讀功能皆持有正向的態度。

5.結論

本研究旨在探討 Google Translate 圖像報讀功能對四年級讀寫障礙學生在閱讀理解學習成效方面的影響，並深入分析研究參與者在使用該工具後，學習焦慮、自我效能感以及對工具使用感受與看法的變化。研究結果顯示，Google Translate 圖像報讀功能對研究參與者的閱讀理解表現與學習態度均產生了積極且正向的影響。

在介入期，所有研究參與者的閱讀理解能力均展現出立即成效，效果量範圍從小至大不等；在維持期，研究參與者甲的學習成效得以維持，而研究參與者乙與丙雖呈現下降趨勢，但學習成效仍達大效果量標準。透過訪談進一步了解，研究參與者對 Google Translate 圖像報讀功能給予高度肯定，認為該工具不僅有助於提升閱讀理解能力，還能增強學習的自我效能感與學習動機，同時促使他們願意主動的參與閱讀活動。

本研究建議結合更多元的數位輔助工具，進一步探討這些工具與不同學習特質學生之間的互動影響，從而驗證其在提升閱讀理解素養及教育適用性方面的整體效能。

參考文獻

王柏元、李芳媛、林香蘭、邱涵仕、陳燕琪、張仕麟、郭復齊、程臻寧、劉秀鳳、蔡嘉峯(2024)。身心障礙者權利公約(CRPD)教育人員宣導手冊：高中分冊。教育部。

王瓊珠、洪儷瑜、陳秀芬（2007）。低識字能力學生識字量發展之研究─馬太效應之可能表現。特殊教育研究學刊，3，1-16。https://doi.org/10.6172/BSE200709.3203001

江毓鈞、楊曉玲(2012)。識字教材結合互動式電子白板對學習障礙學童之教學應用。特殊教育學報，53，41-54。https://doi.org/10.7034/SEES.201206.0041

張長穎（2013）。圖解識字法對學習障礙學生之教學經驗分享。國小特殊教育，55，65-71。

謝慧如(2009)。圖卡教學與多感官教學對國小啟智班學生語彙學習成效之比較研究〔未出版之碩士論文〕。國立臺灣師範大學特殊教育學系在職專班。

Wirantaka, A., & Fijanah, M. S. (2021). Effective use of Google Translate in writing. *Proceedings of the International Conference on Sustainable Innovation Track Humanities Education and Social Sciences (ICSIHESS 2021)*, 15-23. https://doi.org/10.2991/assehr.k.211227.003

Kumnoed, P. (2018). An analysis of English reading comprehension and reading strategies through Google Translator of Loei Rajabhat University students. *The Asian Conference on Language Learning 2018, 3* (1), 49–62.

Field, S., & Hoffman, A. (2002). Lessons Learned From Implementing the Steps to Self-Determination Curriculum. *Remedial and Special Education*, *23* (2), 90-98. https://doi.org/10.1177/074193250202300205

2015-2024 年职业教育数字化研究：基于 CiteSpace 的可视化分析

A Study on Digitalization of Vocational Education from 2015 to 2024: Visual Analysis Based on CiteSpace

汪永健[1]，叶建宏[1,*]
[1]北京师范大学教育学部
202421010093@mail.bnu.edu.cn；* jianhong.ye@bnu.edu.cn

【摘要】 以中国知网（CNKI）作为文献搜索源，利用 Cite Space 系统梳理 2015-2024 年国内职业教育数字化的研究动态与演进趋势。研究发现：第一，文献数量受政策驱动显著，2022 年后发文量激增，研究焦点与国家教育数字化战略高度契合。第二，核心作者群体呈现“分散化”特征，机构合作网络密度低，高校与职业院校研究界限明显，理论构建与实践探索尚未形成协同效应。第三，关键词共现网络揭示“职业教育本体数字化转型”“技术驱动创新”“产教融合深化”三大研究主线，高频节点“数字技术”与“人工智能”凸显技术赋能特征，突现词演变反映研究重心从“信息化基建”向“战略系统重构”迁移。第四，研究主题聚焦基础理论、教学革新、“三融”发展及国际比较四大维度。未来需强化技术伦理与教育公平的交叉研究，构建本土化理论体系，并通过混合研究范式突破实证研究瓶颈，为职业教育数字化转型提供学理支撑与实践路径。

【关键词】职业教育数字化；可视化分析；文献计量

Abstract: *We used China National Knowledge Infrastructure (CNKI) as our literature search source and employed the CiteSpace system to analyze research dynamics and evolution trends in the digitalization of vocational education in China from 2015 to 2024. Our findings are as follows: First, the number of publications is significantly driven by policy. After 2022, the volume of articles surged and aligned closely with the national digital education strategy. Second, the core author group appears decentralized. Institutional collaboration networks show low density, and clear research boundaries exist between universities and vocational colleges. Theoretical development and practical exploration have not yet achieved synergy. Third, the keyword co-occurrence network reveals three major research themes: the digital transformation of vocational education, technology-driven innovation, and deepened industry-education integration. High-frequency nodes like “digital technology” and “artificial intelligence” highlight technology empowerment. The evolution of emergent keywords indicates a shift in focus from “information infrastructure” to “strategic system reconstruction.” Fourth，research themes concentrate on four dimensions: fundamental theory, teaching innovation, integration of industry, education, and research, and international comparisons.Future research should strengthen interdisciplinary studies on technological ethics and educational equity, develop localized theoretical frameworks, and overcome empirical research challenges through mixed-method approaches. These efforts will provide both theoretical support and practical pathways for the digital transformation of vocational education.*

Keywords: Digitalization of vocational education,Visual analysis,Bibliometrics

1.问题提出

在当代社会，数字化已成为推动生活、生产与创新的核心力量。随着人工智能、互联网、大数据等技术的突破，新技术通过重构教育时空场域、解构知识传播范式、重塑教学流程标准三重维度，打破了教育的传统组织形态（祝智庭、胡姣，2022），教育模式面临深刻转型。而教育数字化已上升为国家战略发展的关键领域，在全球数字竞争力格局加速重构的背景下，推进教育数字化转型既是实现教育公平的重要路径，更是培育未来人才核心素养的战略选择，直接关系到国家人才战略的实施成效。

作为国家教育数字化战略的关键组成部分，职业教育数字化转型具有特殊战略价值。中国中央国务院（2024）出台《教育强国建设规划纲要(2024－2035 年)》提出要“实施国家教育数字化战略。坚持应用导向、治理为基，推动集成化、智能化、国际化，建强用好国家智慧教育公共服务平台，建立横纵贯通、协同服务的数字教育体系。”职业教育作为产教融合的重要枢纽，其数字化进程尤为关键，既承担着培养产业数字化转型所需技术技能人才的职能，又为教育数字化改革提供了可借鉴的实践样本。

在此背景下，职业教育数字化应在继承普遍教育数字化特征的同时，更加聚焦其独特的实践性、行业性和就业导向性。具体而言，职业教育数字化转型是指在职业教育领域深度运用数字技术，以价值重塑为引领、以数据要素为驱动力，并以系统性变革为目标所展开的持续创新与变革过程（张慕文、祝士明，2023）。本研究借助 CiteSpace 对 2015-2024 年间国内关于“职业教育数字化”的核心期刊文献进行计量分析，把握中国职业教育数字化研究的图景变化和研究热点，为构建中国特色的职业教育数字化理论体系提供学理支撑。

2.研究设计

2.1. 数据来源

在 CNKI（中国学术期刊网络出版总库）文献数据库以“篇名=职业教育数字化”或者“主题=职业教育数字化、职业教育信息化、职业教育数智化、元宇宙”为关键词进行检索，查询年限规范为 2015—2024 年，期刊来源为 CSSCI 和北大核心期刊。共检索得 932 篇文献，通过统计工具剔除内容无关文献、重复文献以及综述、书评和会议报告中的相关文献， 最终得到 724 篇有效文献样本。

2.2. 研究方法

本研究将清洗后的 724 篇文献以“Refwork”格式导出，以“download_**.txt”命名后导入 CiteSpace 6.3.R1 进行格式转化、分析。时间跨度设置为 2015 年 1 月至 2024 年 12 月，时间切片设置为 1 年，其余参数保持默认不变，运用软件中的 Author（作者）、Institution（机构）、Keywords（关键词）等节点类型分析，直观呈现近十年来国内职业教育数字化研究领域的发展状况和热点问题，探索该领域发展的最新趋势。

3.研究结果

3.1. 论文数量与年度分布

近十年来，“职业教育数字化”研究文献整体呈现明显的上升态势，但增长幅度在不同年份间存在波动(见图 1)。研究发现，国家政策对该领域的影响显著。2015 年至 2021 年间关于“职业教育数字化”的研究总体上呈现缓慢增长态势，期间虽有波动，但并未出现显著的跃升。自 2022 年起，在国家宏观政策的强力驱动下，研究数量陡然上升，2023 年与 2024 年

的发文量分别达到 175 篇与 187 篇，合计占比 49.6%。这一转折与国家层面密集出台的教育数字化政策密切相关：2022 年 2 月国家开展教育数字化战略行动，2022 年 3 月教育部职业教育与成人教育司明确要求推动职业教育与继续教育的数字化升级。总之，职业教育数字化正逐渐成为职业教育教学改革及技术技能人才培养的核心推动力。

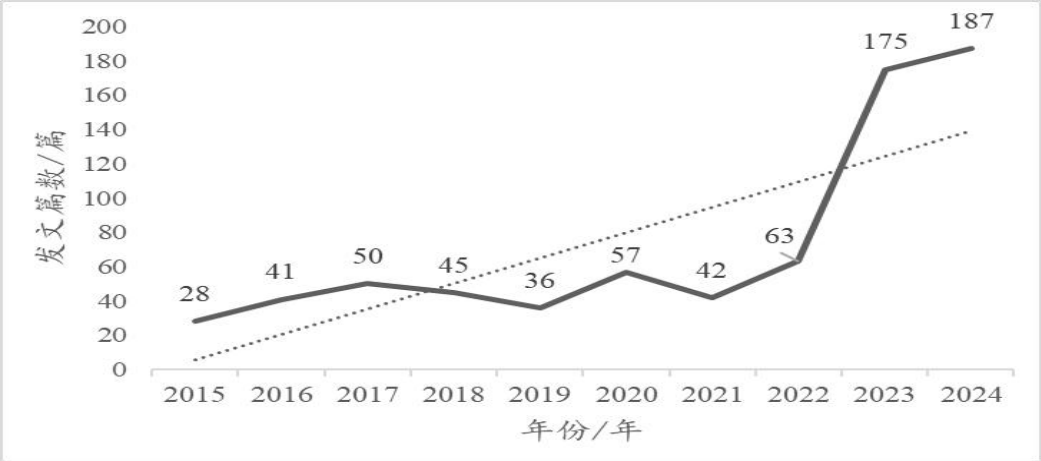


图 1“职业教育数字化”研究的论文数量与年度分布

3.2. 关键词共现分析

通过词频分析可确定职业数字化研究领域的热点关键词，如“职业教育”“数字化”“高职院校”“数字技术”等，并分析其重要性。图 2 显示共有节点 278 个，连线 384 条，网络聚类模块值 Q=0.5152>0.3,整体网络的聚类平均轮廓值 S=0.8860>0.5,聚类结构显著且主题间关联紧密。分析关键词共现可以发现：其一，职业教育与数字化双向互构是近十年的研究核心内容。高频关键词“职业教育”（频次 330，中心性 0.98）与“数字化”（频次 43，中心性 0.08）构成共现网络的枢纽节点，二者共同构成了研究领域的核心框架。职业教育作为数字化转型的实践载体，一方面，高频次“高职院校”“职业院校”等节点表明研究聚焦职业教育体系的数字化转型路径；另一方面，“数字技术”“信息化”等关键词凸显技术驱动下教学模式、资源供给与评价机制的创新需求。其二，“数字技术”“人工智能”“工业 4.0”等技术类术语构成独立子群，凸显出技术发展在职业教育数字化转型过程中的重要作用。

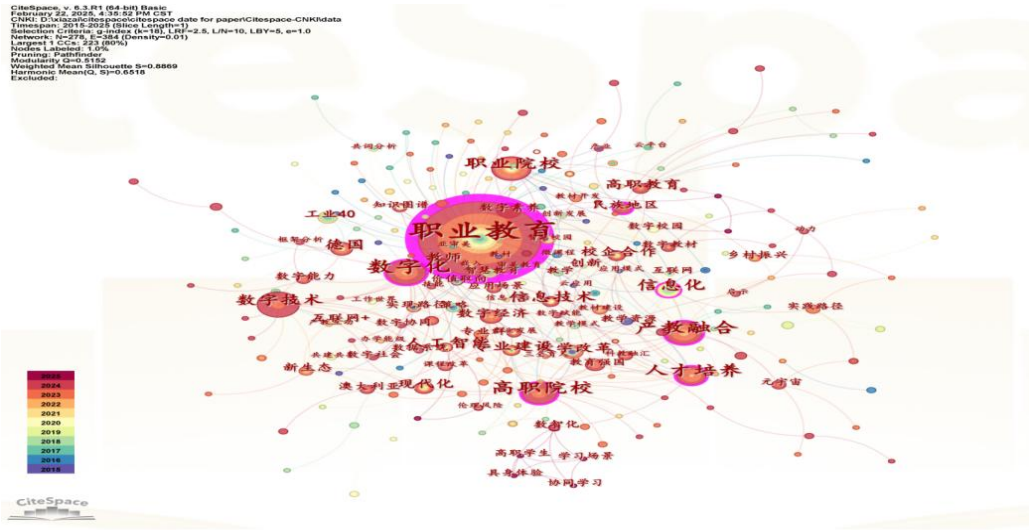


图 2 职业教育数字化研究的关键词共现知识图谱

3.3. 关键词中心性分析

在一定时间内，频次和中心度较高的关键词反映了研究者的广泛共识与关注点，从而能揭示出研究领域中的热点问题。近十年职业教育数字化研究出现频率较高的关键词有“职业教育”“数字化”“高职院校”“数字技术”“产教融合”“人才培养”等。基于表 2 的数据发现：“职业教育”作为核心节点，表明绝大多数研究围绕该主题展开辐射式研究，印证了已有研究始终将职业教育本体作为数字化转型的实践载体。“产教融合”作为连接理论与实践的关键枢纽，凸显出职业教育数字化研究的技术赋能特点和人才培养导向。“德国”作为唯一国际参照节点，其聚类与“双元制”“校企合作”形成稳定关联，反映我国职业教育数字化研究对德国经验的长期借鉴。

表 1 近十年“职业教育数字化”研究高频次和高中心性关键词

序号	关键词	频次	中心性
1	职业教育	330	0.98
2	数字化	43	0.08
3	高职院校	41	0.14
4	数字技术	36	0.02
5	职业院校	35	0.06
6	产教融合	35	0.1
7	人才培养	27	0.1
8	信息化	18	0.02
9	德国	16	0.01
10	信息技术	14	0.01
11	数字经济	13	0.01
12	数字素养	13	0.01
13	人工智能	13	0.05
14	互联网+	12	0.01
15	高职教育	11	0.07
16	教育强国	10	0.04
17	教学改革	9	0.03
18	现代化	8	0.02
19	工业 4.0	8	0.02
20	校企合作	7	0.02

3.4. 关键词聚类分析

借助 CiteSpace 的关键词聚类技术，研究者能够精准定位特定领域的研究热点及其演进趋势（段春雨、蔡建东，2016）。基于 CiteSpace 的关键词聚类知识图谱（见图 5），通过模块化算法提取出“人工智能”“信息化”“职业教育”“数字技术”“职业院校”“工业 4.0”“数字技术”“信息技术”“乡村振兴”“高职教育”“应用场景”“协同学习”11 个聚类主题，各类团之间相互交叉，反映出它们之间共性很强。首先，“职业教育”作为最大聚类（频次 330，中心性 0.98），其子群涵盖“高职院校”“职业院校”“人才培养”等节点，形成了“本体—载体—目标”的三层结构。其次，“技术”在数字化转型过程中发挥核心作用。核心聚类中“人工智能”“数字技术”“信息技术”“信息化”等均表明技术发展对于职业教育数字化的推动作用。职业教育的数字化转型不仅仅局限于技术改造，而是借助数字技术构建一个全新的职业教育生态体系（金波、郑永进，2023）。因此，如何认识技

术、改造技术、融入技术成为职业教育数字化研究的重要议题之一。另外，社会需求与政策导向明显。聚类图中乡村振兴成为单独聚类，产教融合聚类指向了校企合作、工学交替等场景化教学模式创新。

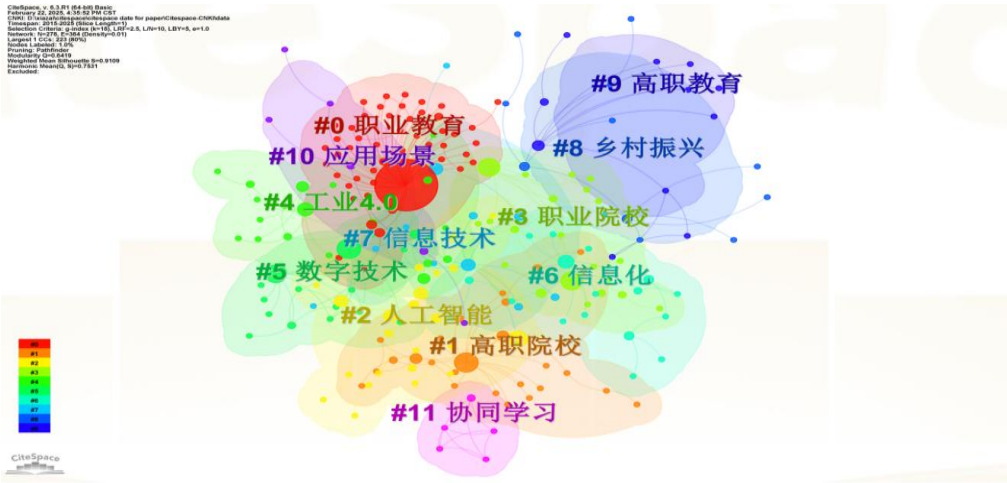


图 3 职业教育数字化的关键词聚类知识图谱

3.5. 关键词突现分析

近十年来的职业教育数字化研究前沿突变历程大致可以分为三个阶段（见图 4）。第一阶段为 2015-2019 年，此阶段突变词为“信息化”“互联网+”“教学”“信息技术”“云平台”等。这一阶段的关键词反映了职业教育数字化研究聚焦技术基础设施和教学信息化的初始特征。第二阶段为 2019-2022 年，此阶段突变词为“智慧教育”“人工智能”“创新发展”“产教融合”“乡村振兴”等。伴随人工智能技术的成熟和“职教 20 条”（2019）出台，研究重心转向技术驱动的教学模式创新及产教融合的深入发展。第三阶段为 2022-2024 年，此阶段突变词为“数字经济”“教育强国”“教材建设”。国家战略引领了职业教育数字化的研究方向，研究强调数字化转型的系统性与战略性。另外需要注意的是，2022 年后生成式 AI（如 ChatGPT）、元宇宙等技术已成为教育数字化的研究热点，但在突现词中未体现，这说明该领域研究具有一定滞后性。

Top 20 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2015 - 2025
信息化	2015	5.34	**2015**	2020	
互联网+	2015	3.8	**2015**	2021	
微课程	2015	1.8	**2015**	2016	
信息技术	2015	1.48	**2015**	2016	
教学	2015	1.47	**2015**	2018	
策略	2016	2.14	**2016**	2018	
价值取向	2016	1.45	**2016**	2018	
工业40	2017	3.31	**2017**	2018	
云平台	2017	1.75	**2017**	2018	
智慧教育	2018	2.35	**2018**	2019	
创新发展	2018	1.54	**2018**	2020	
人工智能	2018	3.34	**2019**	2020	
现代化	2019	1.38	**2019**	2021	
智慧校园	2020	2.42	**2020**	2022	
乡村振兴	2021	1.5	**2021**	2023	
产教融合	2017	1.46	**2021**	2023	
数字经济	2022	2.55	**2022**	2023	
新生态	2022	1.42	**2022**	2023	
教材建设	2022	1.42	**2022**	2023	
教育强国	2023	1.71	**2023**	2025	

图 4 职业教育数字化研究的关键词凸现分析图

4.研究主题与研究趋势分析

4.1.研究主题分析

结合文献内容，研究发现近十年来职业教育数字化的研究热点主要集中于职业教育数字化的基础理论、数字化背景下的职业教育教学教学、数字化背景下职业教育产教融合与职业教育数字化的国际比较研究四个方面。

4.1.1. 职业教育数字化的基础理论研究

职业教育数字化的基础理论研究是职业教育数字化研究的逻辑起点，已有研究从本体论与价值论的双重视角，探讨了数字化转型的深层内涵与价值意蕴。在职业教育数字化内涵方面，王敬杰（2022）指出职业教育数字化转型是一个超越传统教学框架的范式、系统和深刻的教育重构过程。邓小华（2023）认为职业教育数字化是立足于宏观职业教育系统的全方位创新，涉及教学范式、组织结构、业务流程、价值观念等多层面。在职业教育数字化价值方面，学者从技术现象学（赵磊磊、董艳，2024)、身体哲学（李延平，王景，2024）、伦理风险（焦晨东、黄巨臣，2023）等多学科视角出发对数字化转型进行理论反思。

4.1.2. 数字化背景下职业教育教学研究

教学活动作为教育体系中的核心环节，既是知识传授与能力培养的载体，更是实现教育目标的关键路径（朱德全、杨磊，2019）。在数字化转型背景下，研究重心逐渐从技术工具应用转向教育要素重构与教学范式创新。首先是教材的革新。教材从静态的知识载体向动态的知识生成转变，数字化技术赋予了职业教育教材动态生成性（骆琳等，2024），具体来说，即通过数字化技术实时过滤冗余信息，根据学习者认知状态生成个性化内容。其次是教法的创新。研究者关注到教学方法从单向传授到多维空间的转变，数字化教学不仅改变了职业教育的物质空间，更在社会空间维度改变了师生关系（潘海生、林旭，2023）。最后是教师的发展。生成式人工智能（如 ChatGPT））技术支撑的智慧教室、虚拟实训平台与混合式教学架构，通过创设沉浸式交互场景突破了时空限制，推动了项目化学习、情境化教学等新型教育模式的落地（徐国庆等，2023），因此生成式人工智能驱动的教学实践革新已成为职业教育研究的核心热点。

4.1.3. 数字化背景下职业教育产教融合研究

产教融合是建设现代职业教育体系的重要实践路径，伴随着技术进步和政策演进，产教融合成为职业教育数字化的热点。首先，产教融合是职业教育最为典型的场景，数字技术为职业教育产教融合提供了新的方式（邓会敏等，2023），有研究表明数字化转型是破解产教‘合而不融’之困的重要手段（邓小华，连智平，2024），通过数字化转型可以实现专业设置与招生环节的精准对接、培养环节的精准对接以及实现就业的精准匹配。数字化技术通过产业大数据分析平台与人才需求预测算法，实现专业设置与区域产业动态的智能匹配（宋亚峰、潘海生，2023）。总之，产教融合研究已是职业教育数字化的长期热点。

4.1.4 职业教育数字化的国际比较研究

国外职业教育数字化实践为我国转型升级提供了宝贵的经验和启示。焦晨东与黄巨臣（2022）研究指出美国数字化转型以学术力量为导向，从人才培养端推动转型，强调院校自主性和市场需求响应。陈琪（2024）研究中指出澳大利亚通过《澳大利亚劳动力数字技能框架》、《国家技能协议》等为职业教育数字化提供了制度供给。以德国为代表，针对中国职业教育数字化教学资源大而不强的现状，赵志群与黄方慧（2020）主张从德国职业教育数字化教学资源的设置中获得经验。综上所述，学者从政策与制度框架、技术应用与教学模式、

课程评价体系等多维度总结了世界主要国家在职业教育数字化转型中的经验，为我国职业教育数字化发展提供了重要借鉴。

4.2 研究趋势分析

4.2.1. 技术迭代驱动职业教育数字化研究深化

数字化转型不断催生新技术浪潮，如生成式人工智能、元宇宙、区块链等，这些技术正融入职业教育的各个环节。然而，现有研究对生成式 AI、元宇宙等新技术的探索明显滞后，未来研究应更聚焦于技术驱动下教学流程、教学资源配置以及评价机制的改进，探讨新技术如何赋能个性化学习、沉浸式实训和人机协同教学等前沿领域。与此同时，应重视技术应用中的伦理规范、数据安全以及教育公平问题，厘清数字技术与社会经济系统协同演化的内在逻辑，为职业教育改革提供理论与实践支撑。

4.2.2. 构建中国特色的职业教育数字化理论体系

目前，国内职业教育数字化研究在理论构建上多借鉴西方经验，如德国双元制模式，但其适用性与本土实践之间仍存在较大差异。未来应立足中国实际，构建起反映我国职业教育数字化转型内涵、价值及发展路径的自主理论体系。不仅要解释数字化对职业教育教学、人才培养和产教融合的深刻影响，还应对国家战略的需求做出理论回应。

4.2.3. 实证研究范式的多元化拓展

近十年来关于“职业教育数字化”研究中理论探讨和思辨研究占据主要地位，实证研究数量较少。未来的研究应加强以宏观政策、中观机构及微观教学实验为支撑的多层次实证研究设计，采用社会网络分析、纵向追踪、准实验设计等混合研究范式，构建数字化转型效果的量化评价模型，深入解析数字技术在不同应用场景下对技能形成、人才培养及产教融合的具体影响。同时，建立职业教育数字化案例库和相关指标体系，将有助于实现从经验总结向证据本位的理论突破。

5.结论

总之，近十年我国职业教育数字化研究呈现出政策引领、技术驱动与实践倒逼的鲜明特征。研究轨迹从数字化的局部探索转向整个教育生态的战略重构，映射出国家教育数字化战略的深化逻辑。未来研究需立足中国实际，在充分借鉴国际经验的基础上构建本土化理论体系，同时通过引入混合研究范式，强化宏观政策、中观机构与微观教学实践之间的协同作用。

基金资助

本研究获得“北京师范大学教育学一流学科培优项目自由申报课题”的支持。此项目的名称为“职业教育数字化转型的影响机制与发展路径研究”（YLXKPY-XSDW202408）。

参考文献

陈琪（2024）.澳大利亚职业教育数字技能人才培养的制度供给与实践路径.外国教育研究,51（03）,113-128.

邓会敏、白玲、李桂婷（2023）.数字化赋能职业教育高质量发展的耦合逻辑、作用机理与实践向度.教育与职业,2023（23）,22-28.http://dx.doi.org/10.13615/j.cnki.1004-3985.2023.23.008.

邓小华（2023）.职业教育数字化转型的理论逻辑与实践策略.电化教育研究,44（01）,48-53.

邓小华、连智平（2024）.数字技术重构职业教育场景：理论基础、表征样态与行动路径.重庆高教研究,12（01）,63-73.doi:10.15998/j.cnki.issn1673-8012.2024.01.006.

段春雨、蔡建东（2016）.国际泛在学习领域知识图谱研究.现代远程教育研究,2016（01）,85-95.

焦晨东、黄巨臣 （2023）.职业教育数字化转型中的伦理风险及其应对.中国职业技术教育,2023（21）,14-23.

焦晨东、黄巨臣（2022）.职业教育数字化转型的实践类型及其启示——来自美、德、澳三国的多案例研究.中国职业技术教育,2022（33）,11-21+29.

金波、郑永进（2023）.高质量发展背景下职业教育数字化转型实现路径研究.中国高教研究,2023（07）,97-102.

李延平、王景（2024）.从“离身”到“具身”：数字化时代职业教育身体回归的实践理路与国际借鉴.教育科学,2024（04）,89-96.

骆琳、张棉好、张国红、董可雷（2024）.职业教育产教融合型数字化教材开发研究——基于认知负荷的机理分析.中国职业技术教育,2024（32）,87-95.

潘海生、林旭 （2024）.数字技术赋能职业教育教学空间的协同建构.教育研究,2024（11）,110-124.

宋亚峰、潘海生（2023）.教育数字化背景下职业教育专业建设与产业发展谱系图构建研究.高等工程教育研究,2023（05）,137-143.

王敬杰（2022）.新时代职业教育数字化转型的内涵、困境与路径.职教论坛,2022（09）,5-12.

徐国庆、蔡金芳、姜蓓佳、李政、杨惠、郑杰（2023）.ChatGPT/生成式人工智能与未来职业教育.华东师范大学学报（教育科学版）,2023（07）,64-77.http://dx.doi.org/10.16382/j.cnki.1000-5560.2023.07.007.

张慕文、祝士明（2023）.职业教育数字化转型的内涵、逻辑与策略.现代教育管理,2023（03）,120-128.doi:10.16697/j.1674-5485.2023.03.013.

赵磊磊、董艳（2024）.职业教育数字化转型的技术现象学追问.教育研究,2024（08）,125-135.

赵志群、黄方慧（2020）.德国职业教育数字化教学资源的特点及其启示.中国电化教育,2020（10）,73-79.

中共中央 国务院（2024）.《教育强国建设规划纲要（2024－2035 年）》. https://www.gov.cn/zhengce/202501/content_6999913.htm

朱德全、杨磊（2019）.教学论发展 70 年：实践样态与逻辑路向.教育研究,40（09）,14-28.

祝智庭、胡姣（2022）.教育数字化转型的本质探析与研究展望.中国电化教育,2022（04）,1-8+25.

Generative Artificial Intelligence in Language Learning: A Review of Selected Journal Publications from 2023 to 2024

Xiaoshuang Zhang [1], Bowen Jing [1], Jiarong Chen [1], Lin Luan [1]*
[1] School of Humanities, Beijing University of Posts and Telecommunications, Beijing, China
* luanlin@bupt.edu.cn

Abstract: *Generative artificial intelligence (GenAI) based on large language models (LLMs) represented by ChatGPT-3.5 has ushered in promising tools for foreign language learning and provoked great interest in its application in language learning. This study reviews the empirical literature on the application of GenAI in language learning published on five high-impact journals between 2023 and 2024 since the launch of ChatGPT-3.5. Fourteen articles are selected and further analyzed from the perspective of research contexts, research methods and research findings. The review reveals that the current research on GenAI application in language learning mainly targets English and language skills including writing and speaking, focuses on higher education learners and employs mixed methods. Learners generally hold positive attitudes to GenAI-assisted language learning and benefit from GenAI application, while challenges also emerge in this context. Future studies should broaden the research scope by exploring other languages, language skills such as reading and listening, and educational contexts such as elementary education, meanwhile aiming to address these challenges. This study identifies limitations of current research on GenAI-assisted language learning and provides implications for future researchers and practitioners.*

Keywords: generative artificial intelligence (GenAI), language learning, literature review

1. Introduction

The rising pervasiveness of Artificial Intelligence (AI) has led applied linguists to combine it with language teaching and learning processes. As a new milestone of artificial intelligence, Generative Artificial Intelligence (GenAI) based on Large Language Models (LLMs), represented by ChatGPT-3.5 launched by OpenAI in November 2022, has attracted global attention, holding great implications for the field of foreign language teaching (Law, 2024). Through deep learning and massive text data, GenAI powered by LLMs acquires the ability of language understanding, information summary, content creation and language prediction (Jeon & Lee, 2023). The core of language acquisition lies in the learner exposure to and understanding of a large number of language inputs, as well as the construction and acquisition of language through the analysis of situational context (Krashen, 1981). The powerful ability of GenAI to understand and generate natural language based on LLMs has shown great potential in empowering foreign language teaching. As research on language learning assisted by GenAI has seen explosive growth, a systematic review of those empirical studies is needed. Despite that some review studies have been conducted on AI-assisted education, few reviews have focused specifically on the application of the newly-launched GenAI powered by LLMs in the field of second language learning and teaching. Therefore, this study aims to review studies on the application of GenAI powered by LLMs on second language learning published in top-tier journals. The research questions proposed in this study are as follows:

(1) What research contexts have been investigated and what methods have been employed in the selected studies?

(2) What research findings have been discovered in the selected studies?

2. Methodology

Considering that it is ChatGPT-3.5 that makes the education field realize the great potential of GenAI powered by LLMs, the related research starts from November 2022 in a real sense. Therefore, this review focuses on the studies published from 2023 to 2024 and follows the PRISMA guidelines. Following the study by Yang & Li (2024) for reviewing the application of ChatGPT for L2 learning, this paper focuses on five high-impact journals in the field of technology-enhanced language learning, namely, *Computer Assisted Language Learning*, *ReCALL*, *System*, *Language Learning and Technology*, and *Computers & Education*. The key words used in the searching of qualified articles in these targeted journals include: ("generative artificial intelligence" OR "large language model" OR "GenAI" OR "LLM" OR "ChatGPT") AND ("language learning" OR "language education" OR "language acquisition"). The initial search yielded 109 articles. The retrieved articles were screened on the title and abstract according to the following criteria: a) focused on the application of GenAI powered by LLMs in language learning; b) empirical studies; c) journal articles. The remaining articles were assessed in full text to confirm their eligibility and 14 studies were retained for further analysis. Data extraction focused on research context, participants, target language, target language skills, methods and findings. The extracted data were synthesized using a descriptive and thematic approach.

3. Results

All of the 14 selected studies target English as the foreign language to learn. As to the participants, as shown in Table 1, the majority of selected studies focused on higher education students including undergraduates of different grades. Only two studies (Z.-M. Liu et al., 2024; Wang et al., 2024) pay attention to elementary students. The other three have participants from informal language learning contexts, one investigating IELTS learners in Iran (Fathi & Rahimi, 2024) while the other two recruiting EFL learners from social media (G. L. Liu, et al., 2024; Tram et al., 2024). The first language of participants varies, including Chinese, Arabic, Korean, and Japanese, and most of these participants are Chinese-speakers. And the sample size of participants ranges from small to large according to their research focuses with different research methods. More studies targeting elementary and high school students are also needed, and the target language should be further expanded to other language except English.

Table 1 Overview of research articles.

Target language skills (number of articles)	**Author(s)**	**Participants**		
		Sample size	**Education background**	**First language**
Writing (6)	Chen Ziqi et al., 2024	45	Undergraduates	Mixed
	Z.-M. Liu et al., 2024	65	Sixth-grade students	Chinese
	Fathi & Rahimi, 2024	14	Mixed	Persian
	Liu et al., 2023	8	Undergraduates	Chinese
	Yan, 2024	117	Sophomore	Chinese

			undergraduates	
	Wang, 2024	79	Undergraduates	Japanese
Speaking (3)	Wang et al., 2024	99	Undergraduates	Chinese
	Tai & Chen, 2024	85	Sixth-grade students	Chinese
	Guan et al., 2024	48	Undergraduates	Chinese
General skills (5)	Manar Almanea, 2024	135	Undergraduates	Arabic
	G. L. Liu, et al., 2024	867	Mixed	Chinese
	Tram et al., 2024	334	Mixed	Vietnamese
	Liu & Reinders, 2024	24	First-year undergraduates	Chinese
	Lee et al., 2024	22	Undergraduates	Korean

The majority of selected studies employs mixed-method design (79%), which improves the reliability of results. Various methods are employed in these studies for data collection, ranging from questionnaire to audio-recording, and the majority adopts more than one type of them. The most popular methods include questionnaire, followed by interviews and essay/writing journals. For example, several studies investigated EFL learners' AI adoption experience and factors influencing their continuous usage (e.g. G. L. Liu, et al., 2024; Tram et al., 2024) through both questionnaire and interviews. Meanwhile, there is one study collecting data through observation during the writing process (Fathi & Rahimi, 2024) and one study acquiring speaking data through audio-recording (Tai & Chen, 2024).

As illustrated in Table 1, these selected articles focus on different language skills, which can be concluded into three categories, namely, writing, speaking and general skills (referring to the comprehensive language abilities integrating different aspects). Most studies investigate the effects of GenAI on language learners from the aspect of writing skills, such as the writing performance with the assistance of GenAI, the cognitive process in GenAI-assisted multimodal composing, and their attitudes to GenAI-generated corrective feedback. Studies targeting speaking skill mainly delve into whether GenAI influences speaking proficiency, speaking anxiety, communicative competence and willingness to communicate. Moreover, instead of targeting one specific language skill, nearly one-third of the selected articles explore the GenAI effects from the perspective of general language skills, mainly focusing on the overall attitudes to GenAI as a tool for language learning. There is a lack of studies focusing on the effects of GenAI on other language skills, such as reading and listening.

The research findings of the selected studies can be summarized into three themes. In terms of learner attitudes, learners exhibit a relatively high acceptance of GenAI-assisted language learning due to the perceived usefulness (Tram et al., 2024; G. L. Liu, et al., 2024). As to the impact of GenAI in language learning, it is revealed that the utilization of GenAI may benefit language learners in various aspects, including improving learning motivation and confidence (e.g., Z.-M. Liu et al., 2024), developing critical thinking and self-regulated strategies (Liu et al., 2023; Fathi & Rahimi, 2024), and reducing speaking anxiety (Wang et al., 2024). Meanwhile, challenges and concerns are also identified in the selected studies. While GenAI tools generally enhance language learning, the effect may vary due to internal factors. For example, Liu & Reinders (2024) highlighted the individual differences observed in learning motivation. The integration of GenAI in

English learning raises ethical concerns, such as academic dishonesty and over-reliance (Manar Almanea, 2024). GenAI tools occasionally generate unnatural or even inaccurate responses, negatively impacting the learning experience (Lee et al., 2024). The use of AI tools alone is insufficient to sustain long-term language learning outcomes and thus appropriate teacher guidance is urgently required (Guan et al., 2024; Wang, 2024).

4. Discussion and Conclusion

This study reviews the selected empirical studies on GenAI-assisted language learning from 2023 to 2024 published on five top-tier journals. With the PRISMA approach, detailed analysis was conducted on research contexts, research methods and findings. According to the results, studies on GenAI-assisted language learning are primarily English-targeted and focus on writing, speaking and general skills. Participants are mainly higher education students at campus and several studies recruit online informal English learners as participants. Most studies investigate employed mixed methods with diverse data collection methods. These studies revealed three types of findings, namely, language learners' general acceptance of GenAI-assisted language learning, the positive impact of GenAI tools on language learning, and concerns raised by GenAI application. It can be further proved that the application of GenAI in language learning is a promising research area with the potential to transform language education (Law, 2024). However, the current research also presents several limitations. Firstly, studies targeting non-English are quite limited and language skills such as reading and listening are ignored. Future research should enrich the target language and language skills. Furthermore, the selected studies mainly focus on higher education context and the GenAI application in the context of basic education and other languages needs more attention.

Acknowledgements

This paper was part of the project of "Research on Innovation and Application of Foreign Language Teaching Mode Empowered by Virtual Simulation Experiments in Colleges and Universities in the Context of New Liberal Arts" supported by 2024 Higher Education Research Project of China Association of Higher Education (Grant No. 24WY0303).

References

Chen Ziqi, Zhu Xinhua, Qi, L., & Wei, W. (2024). L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. *Computer Assisted Language Learning*, 1–21.

Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Computer Assisted Language Learning*, 1–40.

Guan, L., Zhang, E. Y., & Gu, M. M. (2024). Examining generative AI-mediated informal digital learning of English practices with social cognitive theory: a mixed-methods study. *ReCALL*, 1-17.

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28.

Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Pergamon Press.

Law, L. (2024). Application of Generative Artificial Intelligence (GenAI) in Language Teaching and Learning: A Scoping Literature Review. *Computers and Education Open*, *6*, 100174–100174.

Lee, J. H., Shin, D., & Hwang, Y. (2024). Investigating the capabilities of large language model-based task-oriented dialogue chatbots and the impact of "suggest replies" function from a learner's perspective. *System*, 103538–103538.

Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 1–29.

Liu, M. Y., Lawrence Jun Zhang, & Biebricher, C. (2023). Investigating students' cognitive processes in AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, 104977–104977.

Liu, M., & Reinders, H. (2024). Do AI chatbots impact motivation? Insights from a preliminary longitudinal study. *System*, 103544.

Liu, Z.-M., Hwang, G.-J., Chen, C.-Q., Chen, X.-D., & Ye, X.-D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1–25.

Manar Almanea. (2024). Instructors' and learners' perspectives on using ChatGPT in English as a foreign language courses and its effect on academic integrity. *Computer Assisted Language Learning*, 1–26.

Tai, T.-Y., & Howard Hao-Jan Chen. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions . *Computers & Education*, 105112–105112.

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, 103528.

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The Impact of Different Conversational Generative AI Chatbots on EFL Learners: an Analysis of Willingness to Communicate, Foreign Language Speaking Anxiety, and Self-perceived Communicative Competence. *System*, 103533–103533.

Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, 126, 103505.

Yan, D. (2024). Comparing individual vs. collaborative processing of ChatGPT-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, *28*(1), 1–19.

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, 124, 103351 – 103351.

以 ChatGPT 創作故事對國二生學習動機與學習體驗心理因素之相關研究

The Study on the Relationship Between Using ChatGPT to Create Stories and the Learning Motivation and Psychological Factors of Learning Experience in 8th Grade Students

蔡宜靜、洪榮昭、蔡其瑞
國立臺灣師範大學 創造力發展碩士在職專班
國立臺灣師範大學 工業教育學系
國立臺灣師範大學 科技應用與人力資源發展學系
yijingtsai1006@gmail.com

【摘要】 隨著第四次工業革名到來，運用 AI 工具於課程活動已成為熱門話題。本研究探討國二生使用 ChatGPT 創作故事，使用 ChatGPT 自我效能、操作 ChatGPT 挫折感、使用 ChatGPT 學習興趣、心流及持續體驗意圖之相關性。研究對象為研究者任教學校的國二生，共計 225 名，透過課程體驗活動與活動結束後問卷調查收集數據，以結構方程模式(SEM)進行假設驗證。研究結果顯示，使用 ChatGPT 自我效能顯著降低操作 ChatGPT 挫折感，並顯著提升使用 ChatGPT 學習興趣。此外，操作 ChatGPT 挫折感對心流有顯著的負向影響，而學習興趣對心流則有顯著的正向影響，心流則對持續體驗意圖呈現顯著的正向影響。
【關鍵字】自我效能；挫折感；學習興趣；心流；持續體驗意圖

Abstract: *With the advent of the Fourth Industrial Revolution, integrating AI tools into educational activities has become a popular topic. This study explores the use of ChatGPT by eighth-grade students to create stories, examining the relationships between ChatGPT self-efficacy, frustration in operating ChatGPT, interest in learning with ChatGPT, flow, and the intention to continue using the tool. Additionally, it analyzes how students' background variables regarding ChatGPT usage experience influence these constructs. The participants were 225 eighth-grade students from the school where the researcher teaches. Data were collected through course activity experiences and post-activity questionnaires. Structural equation modeling (SEM) was used to test the hypotheses, and independent sample t-tests were conducted to analyze whether different background variables significantly affected the constructs. The results showed that ChatGPT self-efficacy significantly reduced frustration in operating ChatGPT and significantly increased interest in learning with the tool. Furthermore, frustration in operating ChatGPT had a significant negative impact on flow, while learning interest had a significant positive impact on flow. Flow, in turn, showed a significant positive effect on the intention to continue using the tool. However, no significant differences were found in the five constructs among students with varying ChatGPT usage experiences.*
Keywords: self-efficacy, frustration, learning interest, flow, intention to continue experience

1.前言

1.1. 研究背景與動機

文字書寫的重要性從《十二年國教國語文課綱》的課程目標獲得肯定，語文是社會溝通與互動的媒介，也是文化的載體，學生運用恰當的文字語彙，抒發情感、表達意見（教育部，

2018），但是相比其他語言技能，學生寫作的態度消極，缺乏寫作的興趣與動機，甚至對寫作產生嚴重的焦慮 (Avci & Iseri, 2021)。

觀察自己目前任教的國二生，對於課堂寫作活動總是避之唯恐不及，寫作過程更是哀嚎聲連連，寫完能挑的佳作屈指可數。探究其原因，已有研究指出寫作強調主觀感受與創意表達，是較為複雜的心理歷程，而傳統的寫作教學模式，無法引起學生的寫作興趣，導致學生寫作內容貧乏、不知所云（彭柏緯、劉怡君，2020），國中生創造力的表現與寫作自我效能更有直接的關係 (Demir, 2013)，針對此一問題已有研究提出「創造思考作文教學策略」，透過多樣、多變的語文聯想遊戲，如：故事完成、感官應用法、類推比喻法等等，使學生的寫作意願與興趣能有所提升（劉佳玟，2007）。

2022 年起，隨著人工智慧 (Artificial Intelligece, AI) 技術的迅速發展，生成式人工智慧（Generative AI）的興起引起劇烈回響，其中矽谷公司 OpenAI 推出的 ChatGPT (Chat Generative Pre-trained Transformer)，在短短五天內就吸引了近百萬的用戶註冊，遠勝各大社群媒體竄流的速度（黃煒軒，2023）。更在短短數月，ChatGPT 以其高階的語言理解與無所不知的應答能力，不僅能自行生成新聞、小說、劇本等各種文案，更能即時回答各式千奇百怪的問題，同時引發各級教師，開始將 AI 應用於教育現場的熱烈討論（顏榮泉，2024）。然而，凡事一體兩面，不少人開始擔心 ChatGPT 快速生成文章、提供不一定完全正確的答案，對學生學習所帶來的衝擊。由此可見，ChatGPT 來勢洶洶，教育界勢必得正視它帶來的衝擊，籌謀對應之道，思考如何掌握此一改變的契機，認識它、善用它，讓它成為創新課程與優化教學的好工具，提升學習成效（張芬芬，2023）。

1.2. 研究目的

第四次工業革命已經到來，AI 技術正如火如荼的擴張，ChatGPT 已成為學生唾手可得的工具，探索和實證 ChatGPT 在故事創作教學的應用，將對文字書寫教學的創新產生深遠影響。本研究旨在探討國二生使用 ChatGPT 創作故事的學習經驗，基於上述的研究背景與動機，本研究將以結構方程模式分析法，探究國二生在使用 ChatGPT 創作故事的活動中，心流與持續體驗意圖的影響因素。本研究的情境為國二生使用 ChatGPT 創作故事時，學生使用 ChatGPT 自我效能和操作 ChatGPT 挫折感、使用 ChatGPT 學習興趣之相關性，進而研究與心流及持續體驗意圖間的相關性。通過研究分析，期望能為創新寫作教學提供實證基礎和實踐指南，推動 AI 技術在文字書寫教學中的創新應用。

2.文獻分析

2.1. 自我效能

自我效能(Self-Efficacy)是指個體對自身能夠成功完成特定行為或任務的信念 (Bandura, 1977)。它不僅影響行為的選擇，還決定了個人在執行任務時的努力程度、堅持時間以及面對失敗的方式 (Schwarzer & Fuchs, 1996)。根據班杜拉的社會認知理論，自我效能的來源主要包括個人成功的經驗、觀察他人成功的經驗、來自他人的言語鼓勵，以及個體的情緒與生理狀態。從事自己感興趣、喜歡的工作時，對自己的能力越有信心，就越能開發出不同且有效的方法，也就越能積極解決創新過程中遇到的困難和處理不確定的風險，從而產生更高的創造性表現(Sun, Hong, & Ye, 2022)。

本研究將自我效能義為「當學生使用 ChatGPT 創作故事時，擁有相信自己能順利完成任務的信心，在面對操作困難時，不易被焦慮、壓力的情緒干擾，這種信念將影響學生使用 ChatGPT 創作故事的興趣與動機，且願意繼續嘗試與 ChatGPT 有關的課程活動」。

2.1. 挫折感

挫折感是指個體在追求目標或完成任務的過程中，因為內外部障礙或阻力而無法達成期望時，所產生的負面情緒或心理狀態 (Lazarus, 1991)。這種情緒通常伴隨著失望、沮喪、不滿，甚至可能導致行為上的退縮或放棄，挫折感的程度取決於目標的重要性、達成的難易度以及個體對困難的容忍度。

本研究將挫折感定義為「當學生使用 ChatGPT 創作故事時遇到的困難，例如：問的問題沒有得到期望的回覆、不知道怎麼換句話說、擔心 ChatGPT 不理解所問的問題等等，這種負面的情緒會降低學生使用 ChatGPT 創作故事的意圖」。

2.2. 學習興趣

學習興趣是指個體對學習活動或學習內容所產生的積極情感與認知傾向，是驅動學習行為的重要內在動力，它能促使學習者對學習內容保持專注，並激發探索與求知的動機（教育部, 2018）。學習興趣的高低會直接影響個體的學習行為，因此在早期學習階段，學習興趣具有至關重要的作用 (Hidi, 2006)。

本研究將學習興趣定義為「當學生使用 ChatGPT 創作故事時，積極參與活動，且能自然而然的集中注意力，產生正向的情緒反應，這種正面的情緒將提高學生使用 ChatGPT 創作故事的動力」。

2.3. 心流

心流 (flow) 是指個體在進行某項活動時，進入高度專注與投入的心理狀態，在這種狀態下，個體會完全沉浸於活動中，忽略時間的流逝與外界的干擾，並感到愉悅與滿足 (Csikszentmihalyi, 1990)。在學習活動中，心流是學生學習過程中的一種理想心理狀態，能促使學生持續投入學習活動，並提升其學習興趣與成效。

本研究將心流定義為「當學生使用 ChatGPT 創作故事時，全神貫注地參與、心無旁騖地投入其中，使個體高度專注其中而不想停止，甚至忘記時間，產生扭曲的時間感，會想持續體驗與 ChatGPT 有關的活動」。

2.4. 持續意圖

持續意圖是指使用者在使用某項科技、系統或服務後，對於是否繼續使用的心理傾向和行為意願，此概念通常用於探討用戶在科技接受與使用情境中的行為，特別是對於資訊系統或教育工具的持續使用意願 (Bhattacherjee, 2001)。

本研究將持續意圖定義為「個體在不久的將來，仍願意繼續再次參與和 ChatGPT 有關的課程活動意願」。

3.研究方法

3.1 研究架構與假設

3.1.1 研究架構

依此設計研究架構如下圖所示：

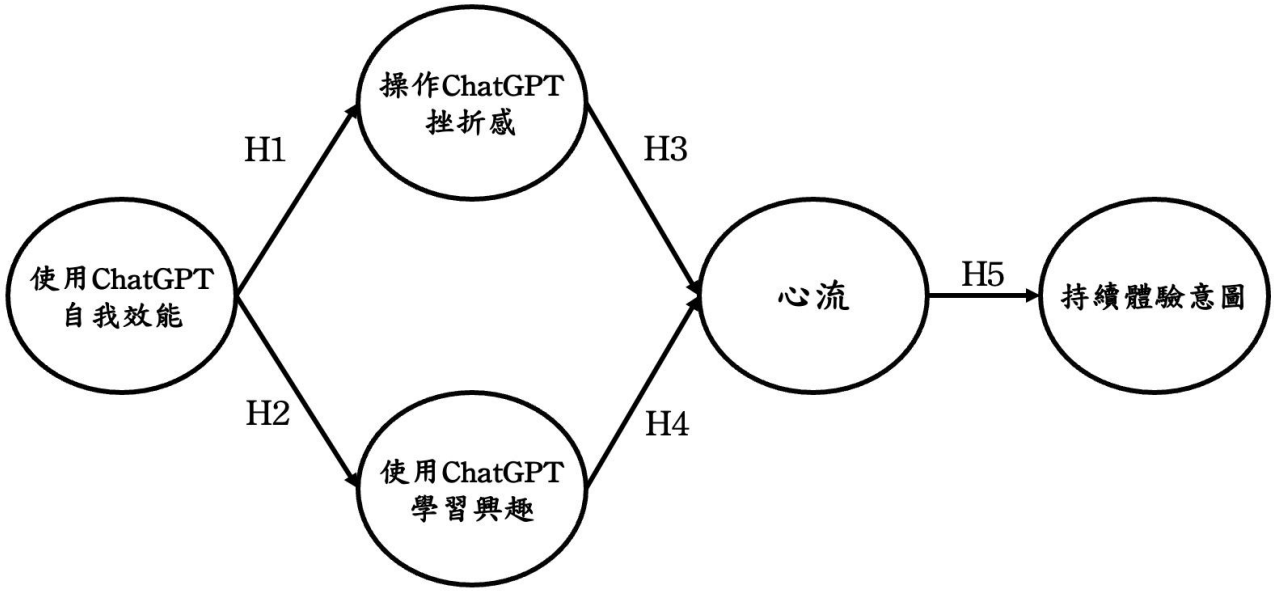


圖 1 研究架構圖

3.1.2 研究假設

本研究擬檢驗假設如下：

H1：使用 ChatGPT 自我效能對於操作 ChatGPT 挫折感具有顯著負相關。

H2：使用 ChatGPT 自我效能對於使用 ChatGPT 學習興趣具有顯著正相關。

H3：操作 ChatGPT 挫折感對於心流具有顯著負相關。

H4：使用 ChatGPT 學習興趣對於心流具有顯著正相關。

H5：心流對於持續體驗意圖具有顯著正相關。

3.2 研究參與者

本研究以參與 ChatGPT 創作故事之教學活動後的全體學生為研究樣本，皆為國二生，利用紙本問卷的方式收集樣本進行分析與了解。經回收後，一共獲取 225 份樣本，其中共有 219 份有效問卷。其中，曾經使用 ChatGPT 的學生共有 205 人（93.6%），不曾使用 ChatGPT 的學生共有 14 人（6.4%）。

3.3 研究工具

本研究採用問卷調查法，研究者建立了一份結構化的紙本問卷，蒐集「國二生參與以 ChatGPT 創作故事的學習活動，其學習動機與學習體驗心理因素之量表」，該量表採用李克特五點量表(Likert scale) 之原則設計，用以測量不同使用經驗的國二生，在體驗活動時使用 ChatGPT 的自我效能、挫折感、學習興趣、心流與持續意圖。

問卷共包含六個部分，分別為第一部分在課堂活動之前，有無使用 ChatGPT 的經驗、第二部分使用 ChatGPT 的自我效能、第三部分操作 ChatGPT 的挫折感、第四部分使用 ChatGPT 的學習興趣、第五部分使用 ChatGPT 創作故事的心流、第六部分參與其他 ChatGPT 創作故事活動的持續意圖。

3.4 研究程序

先按照學校規劃的課程進度教完〈張釋之執法〉一課的完整課文內容，確保學生對文本內容有基本理解，再以自己製作的簡報說明如何操作 ChatGPT 此一工具，和創作故事的精神與方法，啟發學生的創意思維。事先完成班級學生的分組，分組的方式是讓學生自行找組員，以一組五人為原則，電腦設備主要由學校的資源中心提供筆電，學生也可以帶自己在家習慣的筆電或平板，一組最少會有兩台電腦。課程活動的時間共有兩節課，課堂上讓學生大量使 ChatGPT 延伸創作與〈張釋之執法〉有關的故事，但不限定主題，學生可以將其創作

成有科幻精神、童話形式、改變角色狀況的故事等等，並說明老師會進行各組創作故事的評分，預告活動後會發下問卷做調查。

當學生參與完活動後，再發下原先設計好的紙本問卷，並在填答前，詳細說明研究的目的、問卷題數、大約作答的時間，並強調填答無對錯之分且不影響任何成績，學生可以放心的依序填答完問卷的六個部分，鼓勵學生誠實作答。問卷收回後，再進行資料的整理與分析，探討研究假設的各項內容。

4.研究發現

4.1. 問卷信效度

4.1.1 題項分析

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。其檢定須符合以下標準，方可視為符合檢定規範。分別為：卡方自由度比 (χ 2/df) 之數值須小於 5；RMSEA 之須數值小於 0.1；GFI 與 AGFI 之數值皆須大於 0.9；構面題項之因素負荷量 (factor loading, FL) 數值須大於 0.5，如若未達此數值則須將此題刪除 (Hair, Black, Babin, & Anderson, 2019)。經分析後得知，使用 ChatGPT 自我效能之題項從 5 題刪至 4 題；使用 ChatGPT 學習興趣之題項從 5 題刪至 4 題；心流之題項從 6 題刪至 4 題；持續體驗意圖之題項從 5 題刪至 4 題。本研究在項目分析中採用 t 檢定檢驗外部效度。Green 與 Salkind (2004) 指出如若結果顯示之 t 值超過 3 (p*** < .001)，亦即達到顯著之標準。經分析後可得知，本研究所有題項之數值介於 35.28 至 133.21 之間，表示是具有外部效度的，如表 2。

表 2 各構面之一階驗證性因素分析整合表

適配度	χ 2	df.	χ 2/df.	RMSEA	GFI	AGFI	FL	t-value
臨界值	---	---	<5	<0.10	>0.90	>0.90	>0.50	>3
使用 ChatGPT 自我效能	5.48	2	2.74	0.09	0.99	0.94	0.55-0.68	10.69-26.63
操作 ChatGPT 挫折感	24.35	14	1.74	0.06	0.97	0.94	0.52-0.69	-17.41--0.62
使用 ChatGPT 學習興趣	0.23	2	0.12	0.00	1.00	1.00	0.60-0.90	10.82-24.79
心流	2.75	2	1.38	0.00	1.00	1.00	0.21-0.71	0.52-12.29
持續體驗意圖	5.34	2	2.67	0.09	1.00	0.94	0.68-0.81	13.57-19.50

4.1.2 構面信度與效度分析

本研究在構面信效度分析中選擇採用 Cronbach's α to examine the internal consistency and composite reliability (CR) to test external consistency。Emerson (2019) 指出 Cronbach's α 值若高於 0.7 之標準，而吳明隆（2009）提及 Cronbach's α 0.5 或 0.6 也是可接受的範圍，皆可視為具有較好的內部一致性。CR 值高 0.7 之標準，方可視為具有良好的信度。In testing convergent validity of construct, H0air et al.(2019) also suggested factor loading (FL) and averaging variance extracted (AVE) should be high than 0.5。經分析後得知，本研究 Cronbach's α 值介於 0.57 至 0.84 之間；CR 值介於 0.55 至 0.87 之間；FL 值介於 0.21 至 0.90 之間；AVE 值介於 0.26 至 0.63 之間，如表 3。

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
使用 ChatGPT 自我效能	3.94	0.61	0.70	0.70	0.55-0.68	0.37

操作 ChatGPT 挫折感	2.34	0.69	0.79	0.80	0.52-0.69	0.36
使用 ChatGPT 學習興趣	4.15	0.74	0.84	0.87	0.60-0.90	0.63
心流	3.52	0.69	0.57	0.55	0.21-0.71	0.26
持續體驗意圖	4.04	0.76	0.82	0.83	0.68-0.81	0.54

4.1.3 構面區別效度

區別效度是以判定每一構面之 AVE 須大於各成對構面間之相關係數的平方值必須大於 Pearson 相關分析的絕對值 (Fornell and Larcker, 1981)。本研究各構面相關係數介於-0.541~0.74 之間，小於對角線個別構面之 AVE 平方根，如表 4。

表 4 各潛在變數之相關係數

構面	項目數	相關係數				
		A	B	C	D	E
A 使用 ChatGPT 自我效能	4	0.61				
B 操作 ChatGPT 挫折感	7	.54	0.51			
C 使用 ChatGPT 學習興趣	4	.43	.329	0.79		
D 心流	4	.42	.30	.58	0.63	
E 持續體驗意圖	4	.48	.27	.71	.53	0.74

註 1：對角反白的部份，其值為平均解釋變異量（AVE 平方根）；非對角線之其他數值為各構相關。

4.2 Model Fit Analysis

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。本研究經分析得知，χ2/df.為 1.685；RMSEA = 0.06；GFI = 0.87；AGFI =0.84；NFI =0.82；TLI = 0.91；CFI = 0.92；IFI = 0.92；RFI = 0.8；PNFI = 0.73；PGFI =0.71，由此可知此模型的適配指標整體良好。

4.3 路徑分析

本研究五個假設經分析得知，使用 ChatGPT 自我效能對於操作 ChatGPT 挫折感達到負向顯著的影響(β =-.73 , *t* =-5.53***)；使用 ChatGPT 自我效能對使用 ChatGPT 學習興趣達到正向顯著的影響(β =.53, *t* =4.92***)；操作 ChatGPT 挫折感對於心流達到負向顯著的影響(β =-.18, *t* = -2.71**)；使用 ChatGPT 學習興趣對於心流達到正向顯著的影響(β =.84, *t* =9.67***)；心流對於持續體驗意圖達到正向顯著的影響(β =.88, *t* =5.389***，其中*$p<.05$, **$p<.01$, ***$p<.001$。

Bido 等人(2014) 提出 Cohen 之 f 2 意指經由涵蓋和消釋模型的構面所得到的數值，是一項較為罕見卻極為有效的應用，可用來計算出效果量，其計算方式會透過解釋力 (R2) 得出，算式為 $f^2 = R^2 / (1 - R^2)$，且 Cohen 等人 (2007) 指出當 f^2 大於 0.02 時為低效果量；當 f^2 大於 0.15 時為中效果量；當 f^2 大於 0.35 時為高效果量，經分析後可知本研究的構面具備高效果量。

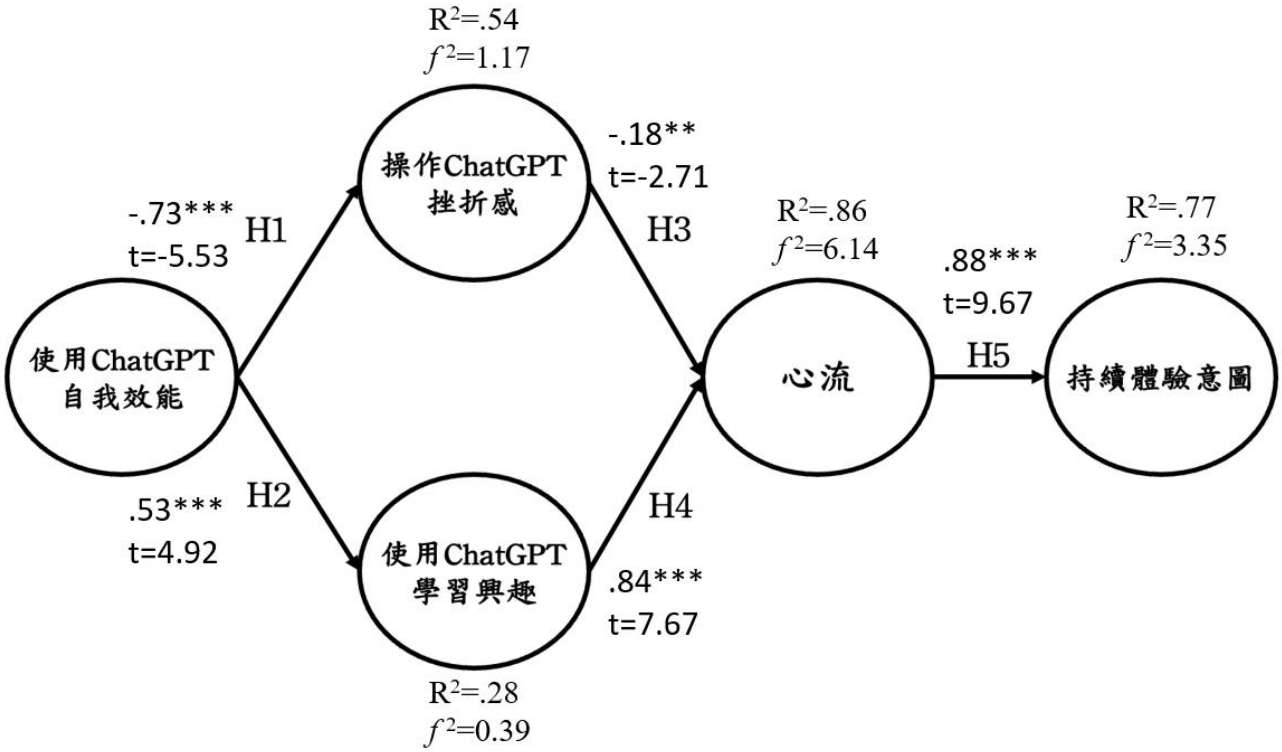


圖 2 路徑分析圖

5.研究討論

本研究探究「使用 ChatGPT 自我效能」、「操作 ChatGPT 挫折感」、「使用 ChatGPT 學習興趣」、「心流」、「持續體驗意圖」構面間的影響關係，從路徑分析檢測的結果來看，使用 ChatGPT 自我效能和操作 ChatGPT 挫折感呈顯著負相關；使用 ChatGPT 自我效能和使用 ChatGPT 學習興趣呈顯著正相關；操作 ChatGPT 挫折感和心流呈負相關；使用 ChatGPT 學習興趣和心流呈顯著正相關；心流和持續體驗意圖呈顯著正相關，進一步討論如下：

5.1 假設驗證

從研究數據可以知道當國二生使用 ChatGPT 自我效能較高時，由於擁有強烈的學習動機，能勇於面對挑戰，即使遇到困難也相信自己能順利克服，並積極尋找解決辦法，而非輕易放棄，因此操作 ChatGPT 挫折感顯著降低、使用 ChatGPT 學習興趣顯著提升。

另一方面，操作 ChatGPT 挫折感所引發的負面情緒，如：沮喪、焦慮、不安等等，會降低使用者操作 ChatGPT 的專注度，使其難以完全投入到當前的任務中，較難進入心流的學習狀態，因此心流顯著降低。然而，當使用 ChatGPT 學習興趣較高時，會因使用者有較強烈的內在學習動機而更積極主動的參與任務，對課程活動保持高度的專注力與持續性，容易進入心流的學習狀態，因此心流顯著提升。

值得我們注意的是，心流是學習的理想狀態，當國二生在使用 ChatGPT 創作故事進入心流的狀態時，會感受到高度的專注，並完全沉浸於創作的過程中，且能從中獲得深刻的滿足感，因此持續體驗意圖顯著提升。

5.2 建議

綜合上述的研究結果，我們可以發現提升國二生的自我效能，不僅可以減少操作 ChatGPT 的挫折感、提升學習興趣，還能增強學習的心流體驗，進一步促進持續使用 ChatGPT 的意圖，對課程活動的進行與故事創作的成果皆有正面的影響。

因此在教育現場，身為教師的我們可以思考如何透過適當的引導與支持，提升學生的自我效能，如：設計漸進式的學習任務、拆解教學步驟到較小的單位、提供成功的範例與即時回饋等等。此外，將 ChatGPT 等 AI 工具融入課程教學活動中，如何減少學生操作時的挫折感，並激發其學習興趣，是老師們設計課程活動的關注重點。

然而，本研究的樣本範圍局限於國二學生，未來研究可以考慮擴展至不同的年齡層，以驗證結論的普適性。另外，背景變項目前只有使用經驗作為單一變項，未來研究可以考慮增加其他變項，如：學科興趣、學習風格、科技接受度等等，用以探討其他變項對各構面的影響力，探討使用經驗對各構面的影響力為何有限。

6.結論

本研究發現使用 ChatGPT 自我效能與操作 ChatGPT 挫折感之間呈顯著負相關，與使用 ChatGPT 學習興趣之間呈顯著正相關，這表明在課程活中學生使用 ChatGPT 的自我效能，會影響其學習的自信心與積極性。進一步的，本研究顯示操作 ChatGPT 的挫折感與心流之間呈顯著負相關，使用 ChatGPT 的學習興趣與心流之間呈顯著正相關，這意味著挫折感會中斷學生學習的投入程度，學習興趣卻能穩定學習時的專注度，最後，本研究說明了心流與持續體驗意圖之間呈顯著正相關，這說明了學習進入心流狀態的重要性。

綜上所述，本研究不僅擴展了國二生使用 ChatGPT 狀態的理解，也為未來的教學實踐提供了實證支持，這些結果提供了實質的見解，有助於設計更加精準和有效的故事創作教學策略，從而提升學習者在多方面成果的表現。

參考文獻

教育部（2018）。十二年國民基本教育課程綱要國語文領域。教育部。

吳明隆（2007）。SPSS 操作與應用: 問卷統計分析實務。五南圖書出版股份有限公司

張芬芬（2023）。老師如何善用 AI 工具：認識 ChatGPT。臺灣教育評論月刊，12(10)，69-80 。 https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-N202309290021-00012

彭柏緯、劉怡君（2020）。淺談國小國語文寫作教學之困境。臺灣教育評論月刊，9(7)，123-128 。 https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-202007-202007070024-202007070024-123-128

劉佳玟（2007）。創造思考作文教學法對國小五年級學童在寫作動機及寫作表現上的影響。國立屏東教育大學

顏榮泉（2024）。從認知處理觀點評論生成式 AI 對學習的影響。臺灣教育評論月刊，13(3)，144-153 。 https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-N202403020019-00027

Sun, Y., Hong, J. C., & Ye, J. H. (2022). The effects of employee's perceived intrinsic motivation on knowledge sharing and creative self-efficacy. Frontiers in Psychology, 12, 762994. https://doi.org/10.3389/fpsyg.2021.762994 (SSCI)

Avci, A. S., & Iseri, K. (2014). Yaratici yazma etkinliklerinin sekizinci sinif ogrencilerinin yazma egilimleri ve yazma kaygılarına etkisi. *International Journal of Language Academy, 2*(4), 152–169.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly, 25*(3), 351–370.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.
Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.
Schwarzer, R., & Fuchs, R. (1996). Self-efficacy and health behaviours. In M. Conner & P. Norman (Eds.), *Predicting health behaviour: Research and practice with social cognition models* (pp. 163–196). Open University Press.

我国教育元宇宙研究现状与未来趋势——基于 Citespace 可视化图谱分析

Research Status and Future Trends of Educational Metacosmos in China - Based on Citespace Visualization Mapping Analysis

陆盈莹[1]，王千姿[1*]
[1]温州大学 教育学院
*wqz2316108432@163.com

【摘要】 教育元宇宙作为新兴研究领域，其研究现状与趋势备受关注。利用文献计量方法和 Citespace 软件，对中国知网(CNKI)数据库中 2021 至 2024 年间的教育元宇宙研究文献进行可视化分析。研究结果显示，该领域发文量呈现爆发式增长，研究主题从技术驱动逐渐转向应用驱动，研究方法由宏观理论过渡到微观实践，重点从知识传授转向知识创造与协作学习。关键词共现网络分析揭示了研究热点的演变和结构变化，而作者与研究机构的共现分析指出了研究团队的合作模式和地理分布特征。未来趋势表明，教育元宇宙将更注重技术在教育实践中的应用，促进教育模式创新和教育公平。
【关键词】 教育元宇宙；Citespace；技术应用

Abstract: As an emerging research field, the research status and trend of educational meta-universe have attracted much attention. Using bibliometric methods and Citespace software, the educational meta-universe research literature in the China Knowledge Network (CNKI) database between 2021 and 2024 was visualized and analyzed. The results of the study show that the number of publications in this field has shown explosive growth, the research theme has gradually shifted from technology-driven to application-driven, the research methodology has transitioned from macro-theory to micro-practice, and the focus has shifted from knowledge transfer to knowledge creation and collaborative learning. The keyword co-occurrence network analysis reveals the evolution and structural changes of research hotspots, while the co-occurrence analysis of authors and research institutions points out the cooperation patterns and geographic distribution characteristics of research teams. Future trends suggest that the educational meta-universe will focus more on the application of technology in educational practices, promoting innovation in educational models and educational equity.
Keywords: educational meta-universe，Citespace，technology application

1. 前言

元宇宙的英文为 Metaverse，其由“meta”（意为“超越”）与“verse”（意为“宇宙”）组合而成,旨在表明元宇宙是一种虚拟与现实无缝链接、深度融合的数字世界，其能够产生超越现实世界的显著价值。元宇宙作为连接虚拟世界和真实世界的全新概念，近年来引发了全球范围内的关注和研究热潮（钱小龙 et al.,2023）。它将虚拟现实、人工智能、区块链等技术与教育相结合，为教育变革提供了新的可能性。在此背景下，元宇宙在教育领域应用的可行性成为研究者关注的热点，教育元宇宙的潜力值得深入挖掘和分析（蔡苏 et al.,2022）。

为了更深入地了解我国教育元宇宙研究的现状和未来趋势，本文运用文献计量方法，对获取的文献索引资料进行可视化分析，力求探索国内精准教学的研究脉络、跟踪其研究走向、分析现有研究状况，给出未来研究趋势建议，为我国精准教学的研究和实践提供一些参考。

2. 数据处理与来源

2.1. 数据来源

本研究将中国知网(CNKI) 中文数据库作为研究的数据来源。为保证文献的全面性和权威性，在 CNKI 检索时选择采用高级检索，联合检索关键词为“教育元宇宙”“元宇宙 教育应用”，获得相关期刊文献 823 篇（截至时间 2024 年 7 月 2 日），学位论文 14 篇，相关会议报道 25 篇。剔除会议、学术资讯、主题内容等不符合要求的文献后，共得到相关文献 169 篇，均来自于“北大核心”“中文社会科学引文索引（CSSCI）”“中文科学引文数据库（CSCD）”“中国人文社会科学期刊综合评价数据库（AMI）”等国内权威数据库。以上述 169 篇中文文献为研究样本，对国内教育元宇宙研究展开可视化分析。

2.2. 研究工具

Citespace 是由美国德雷克塞尔(Drexel) 大学华裔学者陈超美教授基于 Java 语言开发的可视化软件，主要基于共引分析理论和寻径网络算法等，对特定领域文献(集合) 进行计量，能够探寻学科领域演化的关键路径及转折点，预测知识发展的新态势（陈悦 et al.,2015）。基于此，本研究利用 Citespace(6.3. R1 版本)绘制可视化知识图谱，对国内有关教育元宇宙的研究成果进行整理分析，梳理出国内在此方面的研究现状和未来趋势。

2.3. 数据处理

本研究主要采用知识图谱分析、聚类分析等方法对文献进行定量分析。使用 Citespace 进行文献分析前，将节点类型(Node Types) 选择“关键词(Keyword) ”，连线阈值数据对象强度设置为夹角余弦距离(Cosine) 类型，知识图谱网络的修剪方式则选择寻径网络算法(Pathfinder) 。这种算法的优点之一是在进行路径选择时，能够对所有数据之间所有可能的两点路径进行筛选，并保留最强的连接（马鑫&王芳，2023）。

3. 数据分析与讨论

3.1. 关键词共现网络分析

在“教育元宇宙”关键词共现分析方面，Citespace 能够实现对被引文献的关键词或者名词短语自动抽取，根据被引文献的共性关系生成聚类(Cluster) 标识，从而清楚地得到研究聚焦点，一个联系相对紧密的独立研究领域通常可以被视为一个聚类。具体参数设置如下: 文献出版“时间”为“2021—2024”; 时间切片(Time Slicing) 单位为 1 年，“Term Source”为“Title”“Abstract”“Author Keywords”“Keywords Plus”(即将标题、摘要、作者、关键词设为抓取源) ; “Node Types”为“Keyword”(节点类型为关键词)，得到的聚类图形如图 1 所示。

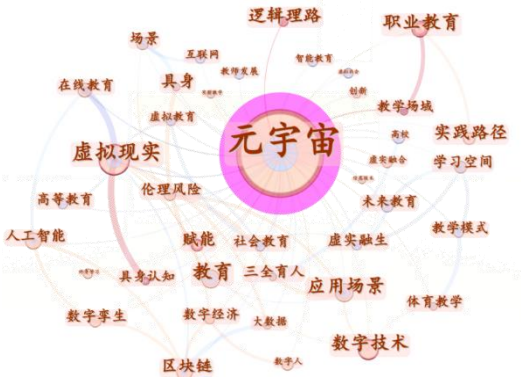


图 1 “教育元宇宙”关键词共现分析

对 CNKI 中文数据库中的 169 篇文献的关键词进行分析，在知识图谱中以圆圈来表示，节点的大小表示该关键词出现的频次，频次越高，节点越大。由图 1 可以看出，元宇宙、虚拟现实、数字技术、应用场景、教育、在线教育、职业教育、区块链、人工智能、赋能、学习空间、伦理风险、实践路径、未来教育、数字孪生、场景、数字人、高校(按节点大小排序) 为出现频次最高的 18 个关键词，说明这些领域是近年来教育元宇宙研究的热点问题。

中介中心性是一个用以测量节点在网络中地位重要性的指标，中介中心性可以显示领域的结构和动态本质。表 1 为教育元宇宙关键词共现频次和中介中心性。

表 1 教育元宇宙关键词共现频次和中介中心性

（以共现频次降序排列，仅列出前 18 位）

序号	频次	中介中心性	出现年份	关键词
1	117	1.59	2021	元宇宙
2	15	0.02	2021	虚拟现实
3	8	0.03	2023	数字技术
4	6	0.01	2022	应用场景
5	6	0.00	2022	教育
6	5	0.00	2021	在线教育
7	5	0.03	2023	职业教育
8	4	0.00	2022	区块链
9	4	0.00	2022	人工智能
10	3	0.00	2023	赋能
11	3	0.01	2022	学习空间
12	3	0.01	2022	伦理风险
13	3	0.03	2022	实践路径
14	3	0.01	2022	未来教育
15	3	0.06	2023	数字孪生
16	3	0.00	2022	场景
17	3	0.00	2023	数字人
18	3	0.00	2022	高校

3.2. 时间线视图

时间线视图的重点在于对聚类之间的关系以及某一聚类中文献的历史跨度进行勾画，可以展现出各聚类发展演变的时间跨度和研究进度，能更直观地显示研究热点的演变过程。以节点所属聚类为坐标纵轴，发表时间(2021—2024 年) 为坐标横轴，将各个节点合理分布在相应的位置上来生成时间线视图，如图 2 所示。

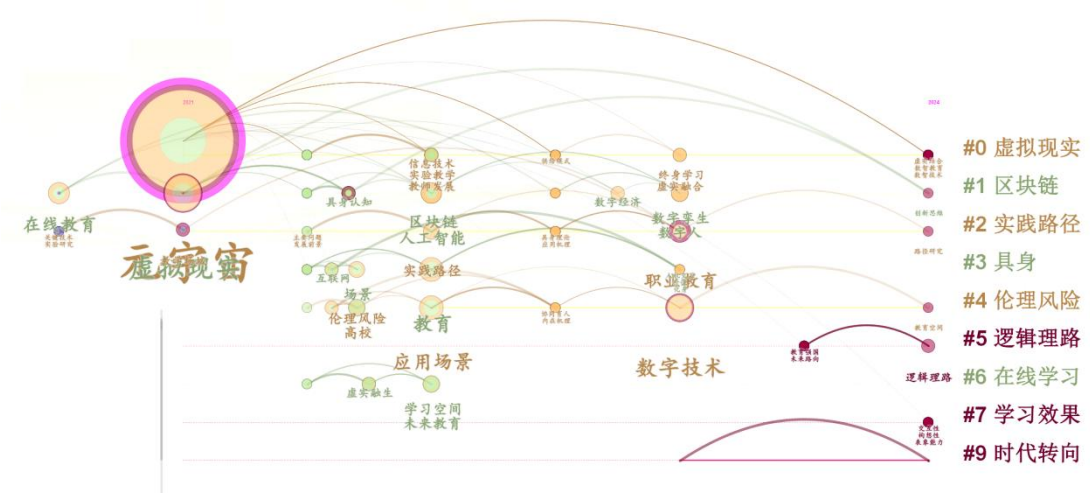


图 2 教育元宇宙时间线视图

图 2 中，元宇宙、虚拟现实、教学场域、在线教育等词的外周有其他颜色实线标粗，表示连接不同领域的关键枢纽，也代表了研究热点的发展情况和结构变化情况。另外，随着时间的发展，大数据、区块链、职业教育、数字技术、伦理风险等逐渐成为研究热点。由上分析，并结合信息技术时代发展特点，随着数字技术在教育领域的应用，当前学者们更加关注元宇宙的关键特征、核心技术与架构、应用实例与学习效果。

3.3. 作者共现分析

在 Citespace 绘制的可视化图谱中，节点的大小对应作者发文多寡，各个作者之间的连线则表示其合作关系的强度。在对某一学科或领域进行分析时，对相应的分析参数进行设置，Citespace 会自动筛选符合条件的作者并绘制关系图谱，如图 3 所示。

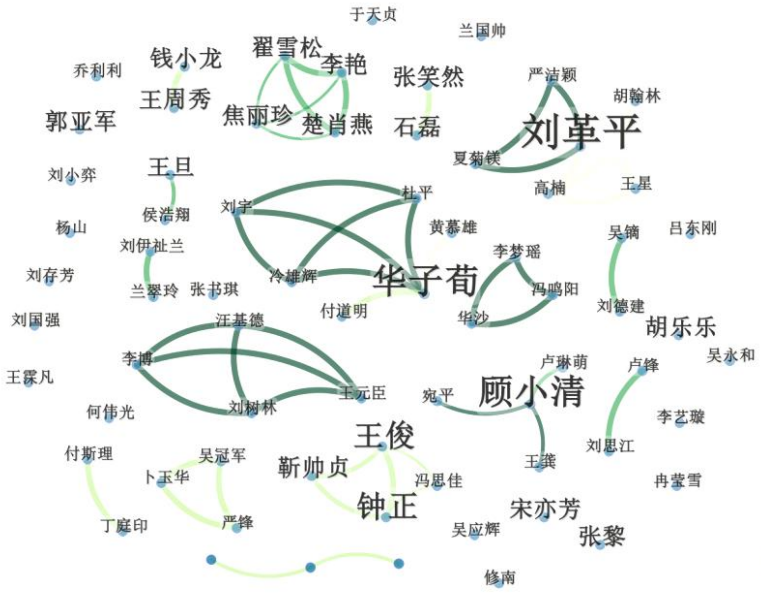


图 3 2021 年-2024 年教育元宇宙领域作者合作网络图谱

图 3 中出现次数最多的作者是刘革平（5 次），其次是顾小清（4 次）、华子荀（4 次），以及王俊、钟正(3 次) 。表 2 为教育元宇宙领域作者列表，出现频次较高的作者更加关注教育变革、教师专业发展、应用研究、实践困境等方面。从图 3 可得，知识图谱共 89 个节点，52 条连线。根据网络密度算法 $d(G)=\frac{2L}{N(N-1)}$ 可得，精准教学研究共被引文献的网络密度

为 0.0133，说明关于教育元宇宙研究的文献之间的相互引用较少，各个研究团队之间的互引关系较弱，教育元宇宙研究处于“部分集中、整体分散”状态。图 3 中包含了四个比较大的合作群体，他们大多为同一院校的教师、同一单位的同事或师生关系，而跨院校、跨学科合作的较少。

表 2 教育元宇宙领域作者列表

（以出现频次降序排列，仅列出前 19 位）

序号	频次	出现年份	作者
1	5	2021	刘革平
2	4	2021	华子荀
3	4	2023	顾小清
4	3	2022	王俊
5	3	2022	钟正
6	2	2022	宋亦芳
7	2	2022	张笑然
8	2	2023	张黎
9	2	2023	李艳
10	2	2022	杨现民
11	2	2023	楚肖燕
12	2	2023	焦丽珍
13	2	2022	王周秀
14	2	2022	石磊
15	2	2023	翟雪松
16	2	2022	胡乐乐
17	2	2022	郭亚军
18	2	2022	钱小龙
19	2	2022	靳帅贞

3.4. 研究机构共现分析

对相应的分析参数进行设置: 点击 Institude 选项，将阈值设定为 2，表示在教育元宇宙研究领域发文数量 2 篇及以上的机构，以此确定对该问题关注度较高的研究单位，如图 5 所示。从图 5 可以看出，我国重庆、上海、武汉等地区高校对于精准教学的研究较多，例如西南大学教育学部（9 次）、华东师范大学教育信息技术学系（8 次）、华中师范大学国家数字化学习工程技术研究中心（5 次）。总体而言，研究机构较为分散，多集中在中东部省份，西部省份较少。

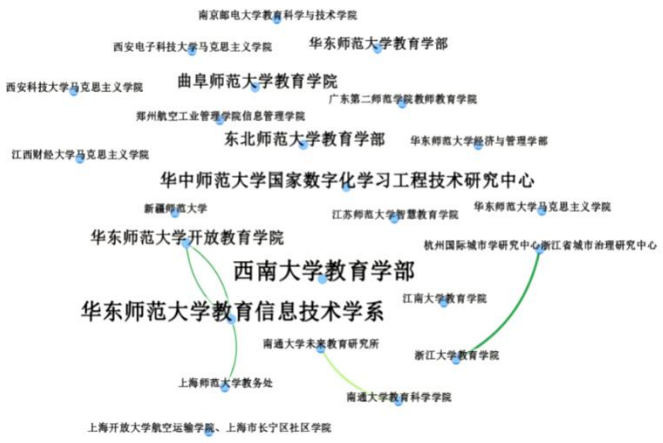


图 4 2021 年-2024 年教育元宇宙领域研究机构网络图谱

4. 我国教育元宇宙研究趋势

教育元宇宙作为一个新兴的研究领域，其发展趋势体现了学术界对于未来教育模式的探索和期待。基于 Citespace 可视化图谱分析，在分析我国教育元宇宙研究现状的基础上，我们可以从研究主题深化、研究方法创新和研究重点转移三个方面进一步展望我国教育元宇宙研究未来发展趋势：

4.1. 研究主题深化：从技术驱动到应用驱动

在教育元宇宙的初期研究阶段，技术驱动是主要的研究主题。早期研究者们集中关注于探讨虚拟现实（VR）、增强现实（AR）、区块链等前沿技术，探索技术在教育领域中实现的可行性和潜在效果（刘革平 et al.,2021）。然而，随着技术逐渐成熟，研究主题正从单纯的技术驱动向应用驱动转变（赵森 & 易红郡，2022）。未来的研究将更加关注这些技术在实际教育场景中的应用效果和实践问题，进一步探索技术融合（如 AR+VR 的个性化学习系统）与教育痛点结合，例如通过数字孪生技术解决高危专业（如医学、化工）的实训难题，或利用区块链构建去中心化的学分认证体系，促进终身学习。这既提升教学效果、又增强了学生的参与度和互动性。应用驱动的研究不仅关注技术本身，更关注技术在教育中的实际价值和应用前景。

4.2. 研究方法创新：从宏观理论到微观实践

早期的研究多采用文献综述与理论推演，以宏观理论为主，探讨元宇宙技术在教育领域的理论框架和发展前景。随着研究的深入，研究方法也在不断创新，逐渐从宏观理论转向微观实践。然而当前数据显示文献互引网络密度仅为 0.0133，这说明研究方法仍较孤立，因此未来的研究应倡导混合研究范式，注重量化分析和质性研究相结合，通过具体的教育实践和应用场景，探讨元宇宙技术在教育中的实际应用效果和存在的问题。例如，结合眼动追踪、脑电波等技术，客观评估元宇宙课堂的认知负荷与学习效率，再通过访谈与参与式观察，挖掘教师与学生在虚拟环境中的行为模式与情感需求。或者聚焦到某一群体，例如探索教师在教育元宇宙时代面临的挑战和素养需求（徐建 et al.,2022）。通过微观层面的研究，可以为元宇宙技术在教育中的应用提供更加实际和具体的指导。

4.3. 研究重点转移：从知识传授到知识创造

传统教育模式侧重于知识的传授，而元宇宙技术的引入将改变这一模式。关键词聚类显示，“协作学习”“数字人”等概念已形成独立研究群落，这标志研究重点将逐渐从单纯的知识传授转向生态化学习。这一转变可体现在三方面：第一是环境智能化，元宇宙技术提供了一个沉浸式的、互动性强的学习环境，可以激发学生的创造力和创新思维，促进他们在虚拟环境中进行探索和实验（翟雪松 et al.,2023），培养他们的实践能力和创新能力；第二是主体性强化，学习者通过虚拟化身参与知识共创，增强学习的主动性；第三是社会性联结，基于 DAO（去中心化自治组织）构建学习者社群，推动集体智慧沉淀，如元宇宙技术还可以促进教师和学生之间、学生与学生之间的协作学习，通过虚拟团队合作和项目式学习，增强学习的互动性和趣味性，从而提高学习效果。

5. 结语

通过对我国教育元宇宙研究现状与未来趋势的分析，可以看出教育元宇宙在我国教育领域的发展前景广阔。当前的研究主要集中在技术驱动阶段，但随着技术的不断成熟，研究逐

渐向应用驱动过渡，更加注重技术在教育中的实际应用效果。同时，研究方法也在不断创新，从宏观理论研究转向微观实践，强调通过实证研究和案例分析来探讨技术应用的具体问题和效果。此外，研究重点也在发生转变，从传统的知识传授逐步转向知识创造和协作学习，旨在提升学生的创新能力和实践能力。

未来，教育元宇宙的研究将继续深化，不仅要探索技术的前沿发展，还要关注其在教育实践中的应用价值。通过多样化的研究方法和丰富的应用场景，教育元宇宙将为教学模式的创新、教育资源的共享以及教育公平的实现提供新的路径和解决方案（顾小清 et al.,2024）。希望在不久的将来，教育元宇宙能够真正融入我国教育体系，发挥其独特的优势，为我国教育事业的发展做出更大的贡献。总之，我国教育元宇宙的研究虽已取得一定进展，但仍处于探索阶段。未来的研究应更加注重理论与实践的结合，推动技术创新与教育创新的深度融合，助力我国教育现代化的实现。

参考文献

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.doi:10.16192/j.cnki.1003-2053.2015.02.009.

蔡苏,焦新月 & 宋伯钧.(2022).打开教育的另一扇门——教育元宇宙的应用、挑战与展望.现代教育技术(01),16-26.

顾小清,宛平 & 祝智庭.(2024).元宇宙会是昙花一现吗？——“现象级”数智技术及其教育启示.现代教育技术(01),26-34.

刘革平,王星,高楠 & 胡翰林.(2021).从虚拟现实到元宇宙：在线教育的新方向.现代远程教育研究(06),12-22.

马鑫 & 王芳.(2023).元宇宙的概念、技术、应用与影响——一项系统性文献综述.图书情报工作(18),113-128.doi:10.13266/j.issn.0252-3116.2023.18.011.

钱小龙,张奕潇,宋子昀 & 黄蓓蓓.(2023).打开元宇宙学校之门：发端、现状与走向.现代教育技术(03),15-26.

王琳,陈泓舟 & 蔡玮.(2023).元宇宙中的教育：现状与革新的未来.华东师范大学学报(教育科学版)(11),38-51.doi:10.16382/j.cnki.1000-5560.2023.11.004.

徐建,王俊,钟正,张国良 & 冯思佳.(2022).教育元宇宙时代教师发展的挑战与应对.开放教育研究(03),51-56.doi:10.13966/j.cnki.kfjyyj.2022.03.005.

翟雪松,楚肖燕,顾建民,李艳 & 王会军.(2023).从知识共享到知识共创：教育元宇宙的去中心化知识观.华东师范大学学报(教育科学版)(11),27-37.doi:10.16382/j.cnki.1000-5560.2023.11.003.

赵森 & 易红郡.(2022).教育元宇宙:当前焦点、潜在主题与未来方向.教育学术月刊(06),3-11+18.doi:10.16477/j.cnki.issn1674-2311.2022.06.013.

Investigating the Relationships among Virtual Reality Features, Technology Acceptance and Learning Performance in VR-assisted Language Learning

Bowen Jing [1], Xiaoshuang Zhang [1], Jiarong, Chen [1], Lin Luan[1]*
[1] School of Humanities, Beijing University of Posts and Telecommunications, Beijing, China
* luanlin@bupt.edu.cn

Abstract: *Virtual reality (VR) technology characterized with real-time interactivity and strong sense of presence has emerged as an effective tool for language education. The importance of VR application in language learning has been widely acknowledged. However, the role of VR features in English as a foreign language (EFL) students' technology acceptance and learning outcomes remains under-investigated. To fill the gap, this research explored the intricate relationship among VR features (i.e., interactivity and presence), technology adoption (i.e., perceived ease of use, perceived usefulness and continuance intention) and students' learning performance in VR-assisted language learning based on the theoretical framework of technology acceptance model (TAM). Data collected from 254 undergraduates was analyzed through structural equation modelling. The findings demonstrated that interactivity and presence positively predicted EFL learners' perceived ease of use and perceived usefulness of the VR platform, which further facilitated continuance intention. Furthermore, students' continuance intention served as a pivotal antecedent of learning performance. The results encourage technological experts to promote the design of the VR platform and EFL teachers to unleash the full potential of VR properties in language courses.*

Keywords: virtual reality (VR), EFL learner, VR features, technology acceptance model (TAM), learning performance

1. Introduction

With the technology advancement in recent years, virtual reality (VR) has opened up a new avenue for language learning and teaching by providing an interactive, authentic and supportive environment (Man et al., 2024). Virtual reality is defined by Burdea and Coiffet (2003) as "a high-end user-computer interface which involves real-time simulation and interactions through multisensory channels". The situated interactive scenarios and strong feeling of presence embedded in VR-assisted instructional approach have been proved as essential in driving students' linguistic improvement (Luan et al., 2024). Considering the powerful affordance of VR in language learning, it is imperative to figure out how VR properties empower learning attitude and performance in an EFL setting.

2. Literature Review

2.1. VR Features

VR technology possesses three principal characteristics including immersion, interactivity and a sense of presence (Liu et al., 2023). Interactivity concerns the VR platform's capability of responding to learner inputs while presence describes a psychological sensation of being there in an unreal environment generated by VR systems (Makransky & Petersen, 2019). In a cost-effective and space-saving desktop VR environment, immersion is less influential in driving learning success

(Huang et al., 2016). This being the case, this research views interactivity and presence as decisive functional features of the desktop VR-assisted language learning context.

2.2. Technology Acceptance

Technology acceptance model (TAM) proposed by Davis (1989) provides a theoretical framework for educators to explain students' adoption of novel technologies in computer-assisted language learning (Alvi, 2024). Perceived ease of use (PEU), perceived usefulness (PU) and continuance intention (CI) comprise core elements of TAM (Huang & Liaw, 2018). In VR-assisted learning contexts, PEU entails the degree to which students believe that using VR platforms is free of effort; PU pertains to the degree to which students believe that applying the VR platforms will promote learning outcomes and CI involves students' readiness to use VR platforms in future learning activities (Man et al., 2024). Based on TAM, individuals' continuance intention to use an advanced technology is shaped by perceived ease of use and perceived usefulness of it (Davis, 1989). Additionally, students with higher continuance intention are more likely to find the academic benefit brought by VR technology, thus exhibiting enhanced learning performance (Guo et al., 2024).

2.3. VR Features, Technology Acceptance and Language Learning Performance

VR technology is capable of shaping students' learning attitude and performance due to its unique features of interactivity and presence (Liu et al., 2023). For instance, according to Huang et al.'s (2016) study, medical students would perceive the VR system as user-friendly when it enabled real-time and authentic interactions. In a similar vein, Alvi (2024)'s investigation revealed that a strong sense of presence provided by immersive VR learning environment was conducive to students' perceived ease of use as well as perceived usefulness, which stimulated their willingness to continue utilizing VR for English learning. Despite the close association between VR properties and technology adoption, how VR features contribute to students' technology acceptance and learning performance in a specific desktop VR-assisted EFL learning context keeps crucial yet under-investigated. In response, drawing on extant literature, we proposed hypotheses as follows:

H1 (a-b). VR features (i.e., interactivity and presence) positively predicts perceived ease of use.

H2 (a-b). VR features (i.e., interactivity and presence) positively predicts perceived usefulness.

H3. Perceived ease of use positively predicts continuance intention of VR-assisted language learning.

H4. Perceived usefulness positively predicts continuance intention of VR-assisted language learning.

H5. Continuance intention of VR-assisted language learning positively predicts learning performance.

3. Method

3.1. Research Context and Participants

The present study was conducted at a comprehensive university in mainland China with 258 undergraduates (75.2% male, aged from 17 to 21) involved. All of the participants attended to a 16-week English course delivered in a VR-assisted teaching approach. Specifically, students acquired English on a self-developed desktop VR platform named *Situated English in Virtual Reality*. This

VR platform is capable of simulating real-life scenarios at international airports with the purpose of enhancing EFL learners' English application and communicative competence.

3.2. Instruments

Revised from Luan et al.'s (2024) and Huang et al.'s (2016) well-developed survey, a VR-assisted language learning questionnaire was designed to gauged VR features (i.e., interactivity and presence) and technology acceptance factors. All of the items were presented in the form of five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) and translated into Chinese to avoid participants' misunderstanding. Additionally, students' learning performance was evaluated by the final-term examination which took the form of National College English test, a standardized exam to assess EFL learners' English proficiency in China.

3.3. Data Collection and Analysis

After removing four cases with missing value, a total of 254 valid questionnaires were collected. The data analysis procedure encompassed two steps. First, confirmatory factor analysis (CFA) was conducted with the assistance of SPSS and AMOS software in order to guarantee the validity and reliability of each variable. Second, structural equation model (SEM) was established and operated in AMOS to test the hypothesized relationship among VR features, technology acceptance components and learning performance in VR-assisted language learning.

4. Results

4.1. Reliability and Validity Analyses

CFA was performed to verify the reliability and validity of each construct. As shown in Table 1, the cronbach's α fluctuated between 0.91 and 0.96, indicating desirable internal consistency reliability of the adapted questionnaire. Furthermore, factor loading (FL), composite reliability (CR) and average variance extracted (AVE) were employed to measure the convergent validity of variables. The CFA results demonstrated that FL, CR and AVE of all the five constructs exceeded the benchmark of 0.50, 0.70 and 0.50 respectively, marking satisfactory convergent validity. In addition, the majority of correlation coefficients between latent variables were lower than the square root of AVE (see Table 1), thus showing acceptable discriminant validity.

Table 1. Construct reliability and validity.

Constructs	N	1	2	3	4	5	6	FL	α	CR	AVE
1. Interactivity	4	**0.90**						0.89-0.92	0.95	0.94	0.81
2. Presence	3	0.86	**0.90**					0.90-0.92	0.93	0.94	0.81
3. Perceived ease of use	3	0.88	0.85	**0.87**				0.82-0.93	0.91	0.91	0.76
4. Perceived usefulness	4	0.88	0.86	0.81	**0.92**			0.87-0.93	0.95	0.95	0.84
5. Continuance intention	4	0.81	0.78	0.84	0.82	**0.92**		0.90-0.94	0.96	0.96	0.84
6. Learning performance	-	0.27	0.26	0.28	0.27	0.33	-	-	-	-	-

Note. N: item number, α: Cronbach's alpha. The bolded diagonal values are the square roots of AVE for the constructs. The off-diagonal value are the inter-correlations between variables.

4.3. Path Analysis

A structural equation model was established to examine the hypothesized relationships. The proposed model displayed favorable structural fit (i.e., x^2/df=2.50, GFI=0.87, CFI=0.96, NFI=0.94, IFI=0.96, RMSEA=0.077, SRMR=0.037). Path analysis findings revealed that interactivity (β=0.59, p<0.001) and presence (β=0.34, p<0.001) exerted positive effect on perceived ease of use, which backed up H1a and H1b. Meanwhile, perceived usefulness was also positively influenced by interactivity (β=0.52, p<0.001) and presence (β=0.42, p<0.001), aligning with H2a and H2b. Perceived ease of use (β=0.51, p<0.001) and perceived usefulness (β=0.40, p<0.001) collaboratively contributed to continuance intention, which confirmed H3 and H4. Moreover, continuance intention positively predicted students' learning performance in VR-assisted language learning (β=0.33, p<0.001), which provided empirical evidence to H5.

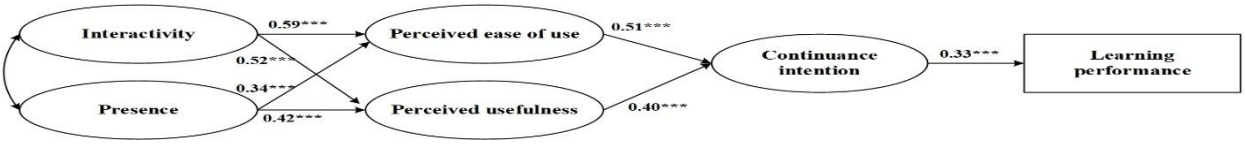


Figure 1. Verification of the hypothesized model.

5. Discussion and Conclusion

The current research explored the interplay among VR features (i.e., interactivity and presence), technology acceptance factors (i.e., perceived ease of use, perceived usefulness, continuance intention) and learning performance in VR-assisted language learning. The prediction of interactivity in perceived ease of use and perceived usefulness confirmed in this study is congruent with Huang and Liaw's (2018) statement, indicating that authentic interaction with VR characters bolster students' confidence in solving technical problems and motivate students to enjoy the VR learning experience. In addition, the results also verified the contribution of presence to students' VR adoption, which echoes Alvi (2024)'s argument that students who feel present in the real-life communicative scenarios tend to recognize the comfort and effectiveness of utilizing VR technology in language learning. Furthermore, both perceived ease of use of perceived usefulness positively predicted continuance intention, which aligns with TAM tenet that the utility of VR technology can arouse students' willingness to engage in VR-assisted English learning activities (Man et al, 2024). Students with stronger intention to use the VR technology for future learning are disposed to gain deeper understanding of learning material and better knowledge acquisition (Guo et al., 2024). In light of the importance of VR features in empowering EFL learning, technology developers are encouraged to integrate more diversified real-life scenarios and interactive elements into platform design (Makransky & Petersen, 2019). Meanwhile, EFL instructors are recommended to take full advantage of the VR properties and incorporate VR-assisted teaching as a feasible complement to traditional on-site instruction (Luan et al., 2024). Several research limitations should be acknowledged. This research recruited students from a single university and relied on self-reported data. Future studies will enroll participants with diverse cultural backgrounds and employ multiple data analysis methods (i.e., interview) to triangulate the findings.

Acknowledgements

This paper was part of the project of “Research on Innovation and Application of Foreign Language Teaching Mode Empowered by Virtual Simulation Experiments in Colleges and Universities in the Context of New Liberal Arts” supported by 2024 Higher Education Research Project of China Association of Higher Education (Grant No. 24WY0303).

References

Alvi, I. (2024). Investigating students’ adoption of virtual reality for L2-learning in India. *Education and Information Technologies*, 29(7), 8035-8056.

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology* (2nd ed.). John Wiley & Sons.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information System Quarterly, 13*(3), 983–1003.

Guo, H., Ma, F., & Zhou, Z. (2024). Validation of technology acceptance model for virtual reality usage in collaborative learning to enhance learner performance. *Innovations in Education and Teaching International*, 1-15.

Huang, H. M., Liaw, S. S., & Lai, C. M. (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. *Interactive Learning Environments, 24*(1), 3-19.

Huang, H. M., & Liaw, S. S. (2018). An analysis of learners’ intentions toward virtual reality learning based on constructivist and technology acceptance approache*s. International review of research in open and distributed learning, 19*(1).

Liu, Z., Yu, P., Liu, J., Pi, Z., & Cui, W. (2023). How do students' self-regulation skills affect learning satisfaction and continuous intention within desktop-based virtual reality? A structural equation modelling approach. *British Journal of Educational Technology, 54*(3), 667-685.

Luan, L., Hwang, G. J., Yi, Y., Lu, Z., & Jing, B. (2024). The effects of a self-developed virtual reality environment on college EFL learners’ vocabulary learning. *Interactive Learning Environments*, 1-12.

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.

Man, S. S., Fang, Y., Chan, A. H. S., & Han, J. (2024). VR technology acceptance for English learning amongst secondary school students: role of classroom climate and language learning anxiety. *Education and Information Technologies*, 1-25.

数字赋能终身学习的研究热点与阶段特征——基于中国知网 2014-2024 年的文献可视化分析

Research Themes and Stage Characteristics of Digitally Enabled Lifelong Learning --Literature visualisation analysis based on CNKI 2014-2024

田艾灵[1]，叶建宏[2*]
[12]北京师范大学职业与成人教育研究所
ailing_tian@mail.bnu.edu.cn ;* jianhong.ye@bnu.edu.cn

【摘要】 数字技术与教育系统的融合创新正推动我国教育的数字化转型。首个数字化终身学习联盟在我国启动，将大力推动我国全民终身学习体系的构建。探究学术界对于数字赋能终身学习的研究有助于深入理解当前数字技术与终身学习的融合现状。运用信息可视化软件 CiteSpace 对中国知网数据库中的 174 篇关于数字赋能终身学习的研究文献进行分析,结果显示，已形成主要的学者网络与研究机构；研究热点主题体现在数字赋能终身学习的顶层设计、现实路径与特定群体中。未来，对数字赋能终身学习的研究将会持续增长且会形成更加丰富的知识体系。

【关键词】 数字技术；终身学习；文献计量；可视化分析；CiteSpace

Abstract: *The innovation of digital technology and the education system are driving the digital transformation of education in China. Exploring academic research on digitally-enabled lifelong learning helps to gain a deeper understanding of the current status of the integration of digital technology and lifelong learning. CiteSpace was used to analyse 174 research documents on digitally enabled lifelong learning in the CNKI. The results show that a network of scholars and major research institutes, have been formed; the hot topics of the research are reflected in the top-level design of digitally empowered lifelong learning, the real-life paths, and specific groups. In the future, research on digitally enabled lifelong learning will continue to grow and form a richer body of knowledge.*

Keywords: Digital technology, Lifelong Learning, Bibliometrics, Visual Analytics, CiteSpace

1. 前言

随着元宇宙、区块链、AR（Augment Reality，增强现实）、VR（Virtual Reality，虚拟现实）等新一代数字技术的发展与应用，数字化转型已成为驱动人类社会走向智能化、数字化的世界性主题。数字技术与教育系统的融合创新，正推动我国教育信息化向数字化转型跃升（胡钦太等，2024）。2015 年，习近平总书记在第二届互联网大会上提出，要推进“数字中国”的建设；2021 年，教育部首次提出了“教育数字化转型”战略。可见，数字赋能教育已被视为构建高质量教育体系的必经之路。党的二十大报告指出，应“建立全民终身学习的学习型社会”，这为我国完善全民终身学习体系指明了方向。2024 年 10 月，全球首个数字化终身学习联盟在我国上海启动，这将进一步推进我国终身学习数字化转型、赋能学习型社会建设，并为公民的终身学习参与提供有力的组织保障。学术界对于数字赋能终身学习的研究随着数字技术的发展与政策文本的导向而不断增加，数字赋能终身学习正在成为一个新

的学术关注点和增长点，尤其是在成人教育领域。本研究聚焦于2014-2024年终身学习领域的数字技术研究，借助可视化软件CiteSpace梳理其近十年的研究热点，试图深入理解前沿数字技术对于终身学习理念与方式等的具体影响。

2. 数据来源与研究方法

2.1. 数据来源及分析

选取中国知网数据库的中文社会科学引文索引（Chinese Social Sciences Citation Index,简称CSSCI）与北大核心（全称为《中文核心期刊要目总览》）的文献来源，对数字赋能终身学习的研究文献进行搜索，不含学位论文、会议、报纸等文献类别。检索主题和标题为“终身学习”与“数字＋元宇宙＋AR＋VR＋区块链”，时间跨度为2014年1月1日至2024年12月1日，在CiteSpace软件中经过去重、清洗，共得到研究文献174篇。

从时间上看，整体上，在2014-2024年这十年期间，数字赋能终身学习这一主题的发文数量总体上呈增长态势。虽然2014-2018年的发文量均低于5篇，但从2019年起，年均发文量均稳定在10-45篇区间。可见，这一主题的研究热度一直在持续，并在2020年与2023年分别出现两个小的波峰，分别为15篇与42篇。其中虽有回落，但总体趋势是不断增长的，并在2023年达到最大值，42篇（见图1）。

图1 数字赋能终身学习的研究主题发文量时间分布（2014-2024年）

笔者分析，这两个小的波峰得益于数字技术的发展与国家宏观战略部署。首先，由于2020年是新冠疫情之年，基于数字技术的远距离学习成为成人获取知识与技能的主要方式；加之在2019年与2020年，教育部先后出台文件《关于服务全民终身学习 促进现代远程教育试点高校网络教育高质量发展有关工作的通知》（教育部，2019）与《2020年教育信息化和网络安全工作要点》（教育部，2020），并且亚马逊网络服务将区块链技术民主化、国内虚拟现实产业市场（AR和VR）逐渐扩大、“元宇宙”概念横扫全球，2020年“终身学习+数字技术”主题有一个爆发期。其次，在2022年，中央网信办、教育部等四部门联合印发《2022年提升全民数字素养与技能工作要点》（中央网信办等，2022），提出促进全民终身数字学习、上海开放大学建立OMO智慧学习中心元宇宙“慧学空间”、基于区块链技术的成人学习认证、记录与学习学分的记录与转移、在南京市的全民终身学习活动周上的VR

数字课程的亮相等，这都为学者研究终身学习领域下的数字技术奠定了技术手段与政策文本基础。

从来源出版物分布看，成人教育与职业教育领域的期刊为主要期刊来源，没有计算机领域相关的期刊来源。在这 174 篇文献所在的期刊中，《成人教育》为该主题的主要阵地，发文占比量为 13.79%，其次依次为《中国职业技术教育》9.20%、职业技术教育 8.05%、《远程教育杂志》7.47%、《中国电化教育》5.75%以及其他期刊。

从发文内容看，《成人教育》与《中国职业技术教育》主要通过引入国际（包括欧盟、澳大利亚、爱沙尼亚）的经验，为我国学习型社会的构建提供理论依据与实践进路；《职业技术教育》侧重于数字技术如何赋能成人的技能教育，并介绍了美国、欧盟、以及我国某一城市的个案经验；《中国电化教育》则是关注教育数字化转型的策略与路径，以及数字技术在教育中的应用探索；《远程教育杂志》则是采用特定视角、具体项目、不同群体以剖析数字技术在教育中的运用。

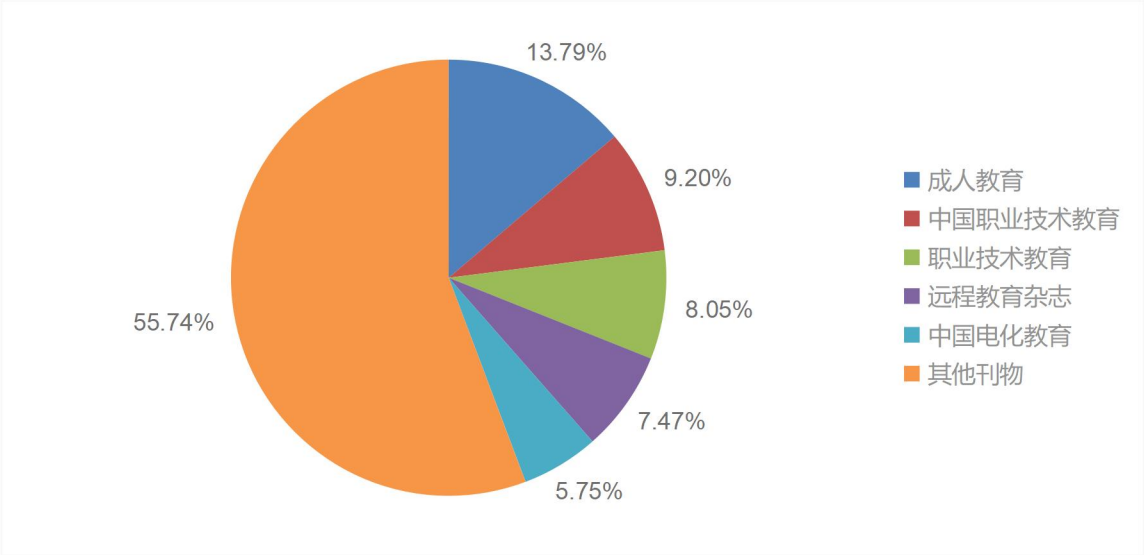


图 2 数字赋能终身学习的出版物分析（2014-2024 年）

2.2. 研究工具及方法

本研究使用 CiteSpace 6.3.R1 可视化软件与知网自带的可视化分析功能作为研究工具。CiteSpace 是由一款基于 Java 语言开发的信息可视化软件，主要通过寻径网络算法进行计量，以探寻学科领域演化的关键路径及知识转折点，并通过一系列可视化图片的绘制来形成对学科演化潜在动力机制的分析和学科发展前沿的探测。（陈悦等，2014）

本研究选取以上两个工具从以下两个方面对搜集到的文献数据进行可视化分析，以了解该领域目前的研究现状。一方面，通过对作者共现网络、作者机构分布来分析我国该领域的关键作者和主要研究机构，以了解其知识基础。另一方面，通过对文献中的关键词进行共现和聚类分析来探究数字赋能终身学习的研究热点。具体参数值包括:研究文献时间范围是 2014-2024 年 11 月，时间切片(Year Per Slice)为 1 年。关键词聚类分析中的节点类型是关键词。节点阈值在共现分析中为 Top N=50。

3. 结果与分析

3.1. 知识基础

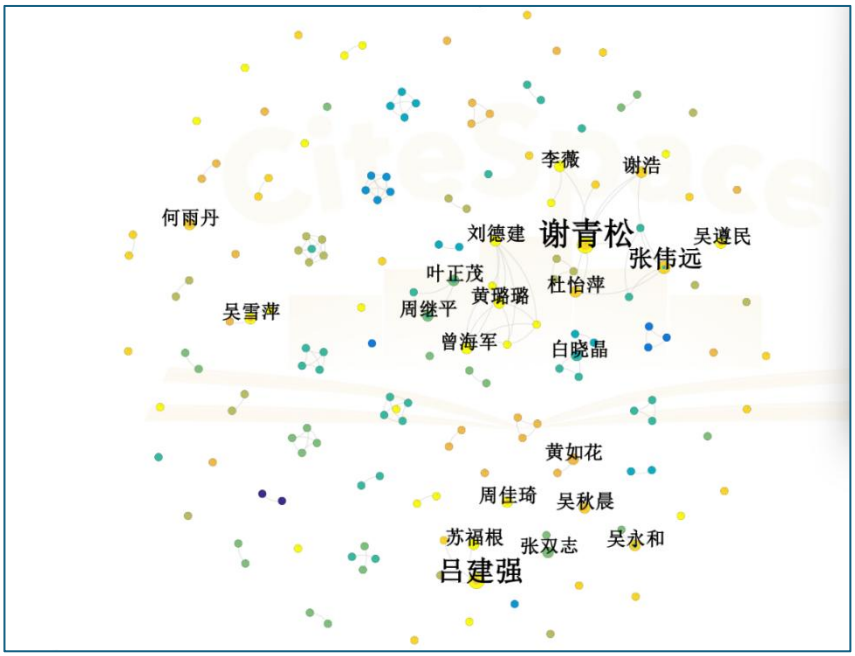


图 3 数字赋能终身学习的作者网络分析（2014-2024 年）

在 CiteSpace 分析中，作者共现数据能够反映该领域影响力较大的关键学者，他们的研究对该领域的知识演进起到重要作用；在知网的可视化分析工具中，文献来源机构分布可以呈现该领域的主要研究阵地，这些文献与研究场所构成该领域的知识基础。

通过对检索到的 174 篇文章进行作者合作网络分析发现，该领域中，主要形成了以重庆开放大学的谢青松、曲阜师范大学的吕建强为核心的两位作者，呈现出低凝聚、高分散的特征。其中，谢青松主要关注终身学习中的学习成果认证、资历框架与学分银行等，并与 11 位学者有过较为紧密的合作，包括北京师范大学的张伟远、北京师范大学的谢浩、广东省教育研究院的杜怡萍等；吕建强则是聚焦于终身学习中的数字化转型、数字赋能与技能人才培养等领域，并与教育部教育管理信息中心的苏福根学者有过较为紧密的合作。

从知网的可视化分析中，作者的来源机构主要包括师范大学、开放大学、职业院校等。可见研究数字赋能终身学习主题的学者以教育学、终身教育机构、职业教育院校的学术背景为主，来自人工智能背景的学者较少。其中，北京师范大学的学者与华东师范大学的学者占比最高，分别占到了 7.56%。这种学科分布状况在一定程度上反映了当前数字赋能终身学习领域的研究格局。师范大学凭借其深厚的教育理论底蕴和广泛的教育资源网络，在该领域的理论构建与实践探索方面发挥着引领性作用。北京师范大学与华东师范大学独占鳌头，彰显了其在数字技术与终身学习融合研究方向上的先发优势和持续投入。然而，人工智能背景的学者参与不足可能会导致在数字技术创新应用、终身学习在线平台等关键技术环节的研究深度受限。

3.2. 研究热点

关键词是文献研究的精炼与浓缩，对文献中出现频率高的关键词进行分析可以反映某一领域的研究热点。本研究对所搜集文献的关键词进行共现分析。同时在关键词共现的基础上进行关键词聚类。关键词聚类后的关键词节点共 215 个，节点连线共 491 条。选择前 9 个聚类，深入阅读不同聚类包含的具体研究内容，结合其中的高频关键词进一步整合以探索数字

赋能终身学习研究的热点主题，得出“数字赋能终身学习的顶层设计”、“数字赋能终身学习的现实路径”、“数字赋能终身学习的特定群体”三类热点主题(见表 1)。

表 1 数字赋能终身学习的研究聚类与热点主题（2014-2024 年）

维度	聚类	名称	节点值	关键词节点	高频关键词
顶层设计	0	终身学习	0.796	43	可持续发展、学习型城市、联合国教科文组织、数字技术
	3	学习型大国	0.954	23	人工智能、全民终身学习、教育数字化转型、教育数字化
	4	教育信息化 2.0	0.934	18	数字胜任力、终身学习体系、学习资源、学习型社会
	8	学习场景	0.997	5	高质量发展、开放大学、数字化改革、数字赋能
现实路径	1	区块链	0.916	34	学习成果认证、学分银行、融合、资料框架
	5	区块链技术	0.95	13	互联网时代、体系构建、开放徽章、信任机制
	6	智慧教育	0.906	13	数字化、应用、需求型政策、环境型政策
特定群体	2	高等教育	0.884	28	协同创新、职业教育、数字素养、技能型人才
	9	信息技术	0.985	5	老龄化社会、积极老龄化、比较研究、数字游戏

3.2.1 数字赋能终身学习的顶层设计

这一主题包括 4 个聚类，囊括终身学习、学习型大国、教育信息化 2.0、学习场景领域，涉及国际层面与我国政府的一些宏观教育目标与教育理念。国际层面的包括联合国所提出的可持续发展理念、教科文组织提出与学习型社会的相关倡议；我国政府层面的则包括我国学习型大国建设、教育信息化 2.0 时代中技术与终身学习的关系。从宏观层面探讨数字与终身学习的关系，是目前数字赋能终身学习研究的重点所在。

东北师范大学的学者常飒飒与王占仁通过对欧盟在 2006 年与 2018 年发布的《欧盟终身学习核心素养建议框架》的比较分析，发现欧盟仍将个体终身学习的核心素养置于知识、技能与态度这三大维度中。并且欧盟在 2018 年的文本中，更加注重终身学习这一概念与现实的融合，包括强调积极的学习环境以及多样化的学习方法、终身学习中教师教育的重要性、开发适用于素养发展的评价与评估方式。欧盟提出这三个重要支撑条件，以为个体终身学习能力的发展提供环境支持、外部激励与制度保障。（常飒飒和王占仁，2019）

苏州大学的学者朱永新与杨帆认为，教育数字化作为“数字中国”战略的重要任务，意味着数字技术作为杠杆，撬动教育与社会的整体变革，并促进面向全社会的智慧教育平台的完善。具体而微，数字技术将通过以下四种方式推进全民终身学习与学习型社会的形成：智慧教育公共服务平台、基于虚拟技术的沉浸式学习、促进终身学习的学分银行、提供综合服务的智慧图书馆与整合网络资源的数码社区。（朱永新和杨帆，2023）

华东师范大学的学者吴旻瑜与武晓菲基于对亚太经合组织文件《经合组织成员国职业所面临的自动化风险》与我国教育部文件《教育信息化 2.0 行动计划》的文本分析提出，教育技术不仅会遵循技术自身的发展逻辑，同时也会遵循人类个体成长的社会性要求的变迁逻辑。由于技术变化所引起的培训需求和潜在的不平等，导致终身学习成为个体的、关乎现实利益

的选择。通过信息化的技术与手段，民众的终身学习力能够提高，并将带动劳动者的技能与素养提升。（吴旻瑜和武晓菲，2018）

西南大学的学者刘革平、高楠等首次界定了教育元宇宙的概念，并提出教育元宇宙作为独立于现实世界的“第二世界”，能够整合各类教育形态、教育要素、教育资源，并促进教育资源公平与教育机会均等。其中，教育元宇宙最重要的是能够打破不同教育平台的壁垒，区块链有利于解决学分的信用问题，促进学分互认。这意味着教育元宇宙能够促进学习成果之间难以认证与转化的难题，实现同一个体、不同类型的学习成果的等值融通（刘革平等，2022）。

3.2.2. 数字赋能终身学习的现实路径

这一主题包括三个聚类，囊括区块链、区块链技术、智慧教育领域，主要涉及数字赋能终身学习的技术支撑，以区块链为主要技术，辅之以其他人工智能技术，包括 VR、AR 等。

天津大学的学者张双志与张龙鹏将区块链视为终身学习体系建设的基础框架，并以哈贝马斯、海德格尔、马克思等哲学家的理论为切口，分析区块链的点对点传输、分布式账本、共识算法、链式时间与智能合约的特性在教育治理结构中的创新。并基于区块链的技术逻辑，提出区块链在教育治理结构中能够促进教育主体间的共商协作、教育权力的多元平等、教育话语的公共透明、教育行动去信任化与教育法治自动履约。（张双志和张龙鹏，2020）

北京开放大学的学者吴莎莎、白晓晶等将数字徽章与区块链技术相结合，认为数字徽章虽然能体现学习者的学习经历，但其认可度与真实性的却受到质疑。而区块链具有安全加密、不可撤销等优势，如若将数字徽章与区块链技术相融合，个体的学习过程将具有更高的可信度与可靠性。此外，两者的融合不仅能有效推动正式与非正式学习的学习成果认证与学分转换，而且也便于学习者在应聘时通过数字徽章向其展示自己的学习轨迹与技能，从而获取更契合自身能力的岗位。（吴莎莎等，2018）

北京师范大学与中国教育报的学者刘德建、曾海军等基于对 2023 年全球智慧教育大会的思考，提出智慧教育作为数字时代新的教育形态，是推动全民享有终身学习机会的必然选择。智慧教育，以众创共享的知识观、智联建构的学习观、融通开放的课程观与人机协同的教学观为特征，以可堆叠模块与微证书、个人数字档案的教育数据、构建元宇宙教育平台等为表征，将惠及所有的地球公民。（刘德建等，2024）

3.2.3. 数字赋能终身学习的特定群体

这一主题包括两个聚类，囊括高等教育、信息技术领域，主要涉及数字技术在高等教育机构中培养技能型人才与推动老年群体的积极老龄化。对不同群体的终身学习能力的关照，体现了终身学习“从一而终”的特点。

广东交通职业技术学院的学者李旭东与曾艳英基于区块链去中心化的特性，提出了基于区块链技术的终身职业教育体系的构建思路。终身职业教育体系能够打破中心机构（即正规学校）的垄断格局，面向全民参与和实现教育公平，促进开放式与分布式职业教育活动，以支撑个体整个职业生涯的规划与发展。此外，通过构建终身职业教育学信数据库、应用体系（学习证书认证系统、职业教育资源库等）与学习服务平台等，区块链支撑的终身职业教育提体系得以完善。（李旭东和曾艳英，2018）

北京师范大学与首都师范大学的学者乔爱玲与张伟远等对互联网时代老年群体的终身学习现状进行了调查，是本主题中少见的实证研究范式。研究人员通过对北京市 2500 名年老者的问卷调查发现，老年群体终身学习的主要动机是保持身心健康。由于老年人在家庭中更多地承担“照顾者”的身份，导致老年群体缺乏参与教育与培训的时间；其次，社区大学、

老年大学以及开放大学的班级与名额较少，这也限制了老年人参与到终身学习的过程中来。（乔爱玲等，2019）

3.3.综合讨论

基于对我国数字赋能终身学习的研究主题与阶段特征分析，可发现当前这一领域的研究范式呈现出向顶层设计、现实路径、特定群体演进的趋势。展望未来，各研究方向在时空维度上的交叉融合将会催生新的研究主题，进而使数字赋能终身学习的研究更丰富化。但目前的研究也存在以下两方面的不足。

首先是现有文献的研究内容上存在着“重理念、轻实践”的倾向。基于对该领域的研究热点发现，我国研究者们多聚焦于政策层面的宏观探讨与技术可行性的理论分析，例如对我国《数字中国》战略、《教育信息化 2.0 行动计划》政策落地的探讨，以及对教育元宇宙的概念、区块链技术在终身学习领域中的可行路径分析。但对数字技术在终身学习中的教学实践评估与监测关注不足，具体表现为：当前学术界对于教育元宇宙、区块链技术等在终身学习者的技术支持程度、学习效果监测等微观层面的系统性研究较为匮乏。

其次是现有文献的研究方法呈现出“重思辨、轻实证”的现状。通过对该领域的研究热点分析发现，尽管目前学界关注数字赋能终身学习的现实路径与特定群体，但研究者大多使用理论思辨这一主导研究方法，几乎没有研究者使用质性或量化的实证研究方法探讨终身学习者使用数字技术进行学习的动机、体验以及学习效果。少量学者关注老龄终身学习者的数字学习体验，但仅停留在对于数据的描述性统计分析中，缺少进一步的相关性分析。这种思辨探讨为主，实证研究不足的研究方法呈现出明显的非均衡态势，制约了该领域研究结论的科学性与实践指导价值。

4. 结论

运用 Citespace 对 2014-2024 年 12 月的数字赋能终身学习的研究文献进行可视化分析，可以看出国内学术界对于该主题的研究呈现出内在的发展规律与阶段性特征。对其进行进一步分析可以深化对该主题的研究的认识，为未来参与和解读终身学习中的数字技术提供重要参考。

第一，从发文量看，国内学术界对数字赋能终身学习的研究保持增长态势，虽然在 2014-2024 年的某些年份偶有下降，但是从整体上来看，研究呈持续增长态势，尤其是在 2023 年达到最大值。未来伴随着数字技术取得新的突破、我国数字强国与教育数字化转型的宏观战略，这一趋势仍将持续。

第二，谢青松与吕建强两位学者、北京师范大学与华东师范大学这两大“985”头部研究机构奠定了数字赋能终身学习研究的知识基础。从某种意义上说，头部师范大学的学者垄断了该领域的知识生产。这是高等教育发展的“中心-边缘”结构（菲利普·C·阿特巴赫，2001）在学术研究领域的再现。这种结构是由多种原因造成的，如北京师范大学、华东师范大学机构的学者长期以来一直就在知识生产领域处于强势地位，它们具有丰富的研究资源与成熟的研究体系。因此，对于非师范大学、开放大学等研究机构的学者来说，其重要任务应是产出更高质量的学术成果，提升自身对学科领域的贡献度。

第三，数字赋能终身学习的研究呈现出多元化的发展方向。首先，在研究热点主题上，数字赋能终身学习的顶层设计自始至终都是学者的主要聚焦点。在这一领域的学者主要通过分析党和国家的战略性文件，来提出数字技术与终身学习相融合的可能性与必然性。其次，数字赋能终身学习的现实路径也是推进学习型社会、构建全民终身学习体系的重要环节。已

有学者主要围绕着具体的数字技术展开讨论。通过区块链技术、数字徽章等来实现不同学习形式之间的学分转换与互认。最后，数字赋能终身学习的特定群体也广受学者关注。在这一领域中，已有学者主要围绕着高等教育中的技能人才的终身学习能力培养与终身学习的机会获得、老年群体的终身学习参与而展开。

本研究基于 CiteSpace 文献计量分析，系统梳理了 2014-2024 年数字赋能终身学习领域的研究热点与阶段特征。研究发现，当前研究主要聚焦于三个核心议题：数字赋能终身学习的顶层设计、现实路径及特定群体支持，体现了从政策导向到实践探索的演进趋势。然而，该领域仍存在明显的研究失衡现象：在内容层面，偏重理论探讨而缺乏对数字学习实践及效果评估的深入考察；在方法层面，以思辨研究为主，实证分析相对不足。未来研究可探索基于终身学习者的数字学习场景的实证研究，验证数字技术的赋能效果等，更好地为我国学习型社会建设提供更坚实的学术支撑。

基金资助

本研究获得“北京师范大学教育学一流学科培优项目自由申报课题”的支持。此项目的名称为“职业教育数字化转型的影响机制与发展路径研究”（YLXKPY-XSDW202408）。

参考文献

常飒飒和王占仁（2019）。欧盟核心素养发展的新动向及动因——基于对《欧盟终身学习核心素养建议框架 2018》的解读。比较教育研究，41（08），35-43。http://doi.org/10.20013/j.cnki.ice.2019.08.005.

菲利普·G.阿特巴赫（2001）。比较高等教育:知识,大学与发展。人民教育出版社。

胡钦太、王姝莉和郭锂（2024）。政策工具视角下我国教育数字化转型的现状与审思。电化教育研究，45（01），61-67+99。http://doi.org/10.13811/j.cnki.eer.2024.01.008.

教育部办公厅（2019）。关于服务全民终身学习促进现代远程教育试点高校网络教育高质量发展有关工作的通知。http://cem.swu.edu.cn/__local/8/D6/E2/52627DD97CF32CCB490F2F98767_57D6687C_7837C.pdf

李旭东和曾艳英（2018）。基于区块链技术的终身职业教育体系构建。职业技术教育, 039（034）, 19-24。http://doi.org/10.3969/j.issn.1008-3219.2018.34.005

刘德建、曾海军、黄璐璐和刘军（2024）。智慧教育让全民享有终身学习机会——基于 2023 全球智慧教育大会观点。电化教育研究，45（05）,98-105。http://doi.org/10.13811/j.cnki.eer.2024.05.013.

刘革平、高楠、胡翰林和秦渝超（2022）。教育元宇宙：特征、机理及应用场景。开放教育研究，28（01），24-33。http://doi.org/10.13966/j.cnki.kfjyyj.2022.01.003.

乔爱玲、张伟远和杨萍（2019）。互联网时代老年群体终身学习现状调查报告。电化教育研究，40（07）,121-128。hhtp://doi.org/10.13811/j.cnki.eer.2019.07.015.

吴旻瑜和武晓菲（2018）。教育信息化 2.0 的时代逻辑——《教育信息化 2.0 行动计划》解读之一。远程教育杂志，36（04）,4-10。http://doi.org/10.15881/j.cnki.cn33-1304/g4.2018.04.002.

吴莎莎、白晓晶和蒋明蓉（2018）。基于区块链技术的在线学习数字徽章认证研究。中国远程教育，（11），19-24+79。http://doi.org/10.13541/j.cnki.chinade.20181108.003.

张双志和张龙鹏（2020）. 教育治理结构创新:区块链赋能视角. 中国电化教育（7）, 9。http://doi.org/10.3969/j.issn.1006-9860.2020.07.009
中华人民共和国教育部（2020）。2020 年 3 月教育信息化和网络安全工作月报。http://www.moe.gov.cn/s78/A16/gongzuo/gzzl_yb/202112/t20211221_589031.html
中央网信办、教育部、工业和信息化部和人力资源社会保障部（2022）。2022 年提升全民数字素养与技能工作要点。https://www.cac.gov.cn/2022-03/02/c_1647826931080748.htm
朱永新和杨帆（2023）。我国教育数字化转型的现实逻辑、应用场景与治理路径。中国电化教育，（01）, 1-7+24。http://doi.org/10.3969/j.issn.1006-9860.2023.01.001

校園行政工作者運用生成式 AI 輔助行政業務之行為意向關係研究－以《國小行政工作者》為例

A Study on the Behavioral Intention of Elementary School Administrators Using Generative AI to Assist Administrative Tasks

江羿璇 [1,*]，洪榮昭 [2]、蔡其瑞 [3]
國立臺灣師範大學教育學院創造力發展碩士在職專班
* avenirsalvatore@gmail.com

【摘要】 本研究旨在探討生成式 AI 在國小行政工作者中的應用行為意向，結合計畫行為理論、科技接受模式以及自我網路效能，分析影響國小行政工作者使用生成式 AI 工具輔助行政業務的各種因素。本研究透過問卷調查法進行實證研究，並運用結構方程模型驗證各理論變數之間的關聯。研究結果顯示，自我網路效能透過感知有用性及感知易用性間接影響行為意向。態度、感知有用性、感知易用性亦對行為意向產生顯著正向影響。基於研究發現，本研究對未來提升國小行政工作者運用生成式 AI 工具輔助行政業務提供具體建議，並對理論與實務意涵進行深入討論。

【關鍵字】 計畫行為理論；科技接受模式；自我網路效能；生成式 AI；校園行政

Abstract: *This study explores the behavioral intention of elementary school administrators to use generative AI tools, integrating the Theory of Planned Behavior, Technology Acceptance Model, and Self-Efficacy. A survey and structural equation modeling were used to examine relationships among variables. Results show that self-efficacy indirectly influences intention through perceived usefulness and ease of use. Attitude, perceived usefulness, and ease of use also significantly positively affect intention. Based on these findings, the study provides practical recommendations to enhance AI adoption and discusses theoretical and practical implications for improving administrative efficiency.*

Keywords: Theory of Planned Behavior**,** Technology Acceptance Model**,** Internet Self-Efficacy**,** Generative AI**,** school administration

1. 研究動機與目的

在數位轉型的背景下，校園行政工作逐漸依賴技術工具以提升效率，減少人力負擔。生成式 AI 技術與網路科技的結合，為校園行政工作者提供了新的工具選擇，這些工具能有效地自動化處理資料分析、文本生成、會議記錄整理等多元任務，有效提升行政效率與準確性。然而，儘管這些工具的技術成熟度已相對提高，這些工具的使用不僅取決於技術本身的功能，還受到行政工作者與其對自身網路技能的信心影響。計畫行為理論(Theory of Planned Behavior, TPB)和科技接受模式(Technology Acceptance Model, TAM)皆是被廣泛用來解釋社會行為和科技使用行為的理論，而行為意向是兩個模式中直接影響實際行為主要的因素。計畫行為理論包含三個主要影響行為意向的心理變項：態度、主觀規範和知覺行為控制(Ajzen, 1991)；科技接受模式則有態度、感知有用性和感知易用性影響個人的科技使用意圖(Davis et

al., 1989)。雖然兩個理論都具有紮實的論述基礎，但使用單一理論來解釋個人行為時仍有其侷限性。如 Lu et al. (2009)主張單一使用科技接受模式對於個人態度和行為意向的解釋可能過於簡化，而單獨使用計畫行為理論則無法理解影響個人態度的因素為何。故結合計畫行為理論和科技接受模式可對行為意向的解釋提供更完整的架構，相互支持個別理論的不足之處。計畫行為理論和科技接受模式的結合已為不同領域的行為意向提供較全面的理解，如網路銀行(Hassan et al., 2018)、運動 App (Yu et al., 2021)、即時通訊軟體 (Lu et al., 2009)等使用行為意向。然而，計畫行為理論和科技模式的結合，並不是一個極為穩定的模型架構，其會因不同的議題和不同的研究對象而產生不同的結果。Hassan et al. (2018)回顧 23 篇有關結合計畫行為理論和科技使用模式的研究，這些研究的解釋變異量介於 24.1%至 76.0%之間，且變項之間的因果關係結果也不盡相同。本研究旨在結合計畫行為理論、科技接受模式與自我網路效能(Internet Self-efficacy, ISE)，探討國小行政工作者對於使用生成式 AI 輔助行政業務的行為意向關係。

2. 文獻探討

2.2 計畫行為理論應用於校園行政工作之相關研究

在教育行政領域，計畫行為理論被廣泛應用於探討行政工作者對新技術或新政策的接受程度(Fishbein & Ajzen, 2010)。校園行政工作者的行為選擇往往受到教育政策、學校文化及資源分配等多重因素的影響，而計畫行為理論提供了一個整體框架來解釋這些行為背後的心理機制。校園行政工作者對於採用新技術的態度通常基於該技術的感知效益，例如能否提升工作效率或簡化流程。過去研究指出，正向態度會顯著促進行政人員採用教育科技(Teo & Zhou, 2014)。在學校情境中，主觀規範通常由學校領導者、教育局以及同事的意見所構成。Lee et al. (2018)發現，校長的支持與鼓勵會強化行政工作者對新技術採用的意向。校園行政工作者的知覺行為控制受到技術培訓、資源可得性以及政策支持的影響。若行政人員感知到足夠的控制力，則其採用新技術的可能性會提高(Venkatesh et al., 2016)。總體而言，計畫行為理論在實務面上能有效解析校園行政工作者的行為意向及其影響因素。

2.3 科技接受模式應用於校園行政工作之相關研究

隨著教育科技的快速發展，學校行政工作者日益依賴數位工具進行任務管理、資源分配及數據分析。科技接受模式為解釋行政工作者對教育科技的採用提供了有力框架(Teo, 2011)。感知有用性被認為是影響行政工作者接受教育科技的關鍵因素。研究顯示，當技術能顯著提高工作效率，如減少重複性任務或自動化流程，行政人員更傾向於採用該技術(Venkatesh & Davis, 2000)。例如，在校務管理系統的應用中，工作者認為這些系統能提升資源分配效率，進而提升其採用意願(Lee et al., 2018)。行政工作者對於技術的感知易用性也扮演重要角色。若技術操作直觀、學習門檻低，使用者更有可能採用該技術。特別是在生成式 AI 的應用中，如 ChatGPT 或自動化報告工具，易於使用的特性能減輕行政工作者的認知負荷(Davis, 1989; Park, 2009)。態度是感知有用性與感知易用性影響行為意向的重要中介變數，當行政人員對技術持有正向態度時，其行為意向會顯著提升(Teo, 2011)。研究指出學校行政人員的態度常受同儕意見及政策支持影響(Ngai et al., 2007)。科技接受模式揭示了感知有用性、感知易用性、態度及行為意向之間的關係，為校園行政工作者對新技術的採用提供了理論基礎。

2.4 自我網路效能應用於校園行政工作之相關研究

校園行政工作者需要運用多種網路工具以完成行政任務，如校務行政系統、電子表單及生成式 AI 等技術。具備較高自我網路效能的行政人員在面對新技術時，能夠更快上手並有

效解決問題，這不僅能提升個人效率，亦能促進團隊協作(Tung & Chang, 2008)。自我網路效能還能幫助行政人員克服技術應用中的焦慮感，增強技術操作的信心(Tsai & Tsai, 2003)。自我網路效能與技術採用意願之間有著密切的正向關係，當行政工作者認為自己有能力掌握新技術時，其採用意願將大幅提升(Hong et al., 2006)。自我網路效能對於數位學習系統、校務管理平台及生成式 AI 的採用意向有顯著影響，特別是在 AI 技術的應用中，具備高自我網路效能的個體，能更靈活地運用生成工具進行數據分析與資源分配(Zhao et al., 2021)。自我網路效能亦能促進行政工作者的專業成長，強化其在數位教育中的競爭力。透過適當的網路技術培訓與操作經驗累積，行政人員的自我網路效能將逐步提升，進而增強其對技術的接受能力與工作績效(Wu et al., 2008)。自我網路效能在校園行政工作的技術應用中扮演了關鍵角色，提升自我網路效能不僅能幫助行政工作者克服技術挑戰，還能促進其行為意向及實際採用率。

2.5 本研究概觀

綜上所述，將計畫行為理論、科技接受模式及自我網路效能整合於校園行政者使用生成式 AI 輔助行政業務的研究框架，以下提出相關研究假設：

假設 1 (H1)：使用生成式 AI 的態度對其行為意向具有顯著正相關。

假設 2 (H2)：使用生成式 AI 的主觀規範對其行為意向具有顯著正相關 。

假設 3 (H3)：使用生成式 AI 的知覺行為控制對其行為意向具有顯著正相關 。

假設 4 (H4)：使用生成式 AI 的感知有用性對行為意向具有顯著正相關。

假設 5 (H5)：使用生成式 AI 的感知易用性對行為意向具有顯著正相關。

假設 6 (H6)：感知易用性對感知有用性具有顯著正相關。

假設 7 (H7)：自我網路效能對感知有用性具有顯著正相關。

假設 8 (H8)：自我網路效能對感知易用性具有顯著正相關。

3. 研究方法

3.2 研究參與者

本研究以便利抽樣的方式選取臺灣多所國小的行政工作者為研究對象，非國小行政工作者與未使用生成式 AI 的國小行政工作者皆排除於研究之外。本研究以網路問卷的方式來蒐集資料，共蒐集 52 份網路問卷，剔除無效問卷 4 份，得有效問卷 48 份，有效問卷率為 92.3%。

3.3 研究工具

本研究問卷共分成五個部分，第一部分為基本資料，包括年齡、性別、職位、工作年資、使用生成式 AI 的時間、使用生成式 AI 的用途、最常使用的生成式 AI 工具及使用生成式 AI 的原因。第二部分計畫行為理論及第五部分行為意向是依 Ajzen (1991)對態度、主觀規範、知覺行為控制、行為意向的定義來編製量表題項，共 32 題。第三部分科技接受模式是依據 Davis (1989)對感知有用性與感知易用性的定義來編製量表題項，共 16 題。第四部份自我網路效能是依據 Bandura (1997)對自我網路效能的定義來編製量表題項，共 8 題。第二部分至第五部分採用李克特氏 5 點式量表作為測量的基準，從「非常同意」至「非常不同意」分別代表 5 至 1 分。

3.4 統計方法

本研究以 SPSS Statistics 23 及 Smart PLS 4 軟體分析進行描述性統計、信效度分析及結構方程模型(SEM)進行資料數據分析，探討各構念之間的影響關係。配合研究目的與研究問題，以.05 作為統計考驗的顯著水準，主要採用描述性統計、結構方程模型分析等方法進行考驗，

詳細內容如下：以描述性統計分析，來瞭解態度、主觀規範、知覺行為控制、感知有用性、感知易用性及自我網路效能之信度、平均數、標準差，及相關進行分析。以結構方程模型分析態度、主觀規範、知覺行為控制、感知易用性、感知有用性與行為意向之相關情形。同時檢測自我網路效能與行為意向之相關是否具有調節效能。

4. 研究結果

4.1. 基本資料之特性分析

本研究經剔除 4 位非國小行政工作者或未使用生成式 AI 的受試者資料後，對剩餘 48 位受試者的資料進行分析。根據本研究樣本的資料顯示(如表 1)，擔任校園行政工作者中年齡以 30-39 歲(31.3%)及 40-49 歲(39.6%)為多數。使用生成式 AI 的用途以提問為最多(72.9%)，其次是文案生成(52.1%)、摘要(45.8%)、翻譯(33.3%)、繪圖設計(31.3%)及簡報製作(29.2%)；常使用生成式 AI 工具以 ChatGPT 為最多(91.7%)，其次是 Canva AI (54.2%)。超過一半的人數使用生成式 AI 未達一年，將近五分之一的人數使用生成式 AI 超過 1 年。使用生成式 AI 的原因以節省時間為大多數(95.8%)，其次是可節省腦力(77.1%)，上級要求或同儕競爭對於使用生成式 AI 的影響較少。

表 1 樣本人口基本資料、使用生成式 AI 工具行為分析 (n=48)

變項	人數	百分比	變項	人數	百分比	變項	人數	百分比
年齡			**使用 AI 的用途（複選題）**			**常使用的 AI 工具（複選題）**		
20-29 歲	11	22.9	翻譯	16	33.3	ChatGPT	44	91.7
30-39 歲	15	31.3	提問	35	72.9	Claude	3	6.3
40-49 歲	19	39.6	摘要	22	45.8	Microsoft Bing	4	8.3
50 歲以上	3	6.3	繪圖/設計	15	31.3	Microsoft	3	6.3
職位			簡報製作	14	29.2	Copilot	7	14.6
主任	3	6.3	文案生成	25	52.1	Gamma	26	54.2
組長	45	93.7				Canva AI		
工作年資			**使用 AI 的時間**			**使用 AI 的原因（複選題）**		
1-5 年	13	27.1	1-3 個月	11	22.9	可節省時間	46	95.8
6-10 年	10	20.8	4-6 個月	13	27.1	可節省成本	18	37.5
11-15 年	10	20.8	6 個月-1 年	13	27.1	可節省腦力	37	77.1
16-20 年	6	12.5	1-2 年	6	12.5	產出品質佳	19	39.6
			2 年以上	5	10.4	自我增能	18	37.5

4.2. 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。如表 2 所示，本研究 Cronbach 的α值虛擬 .80 至 .94 間；CR 值虛擬 .86 至 .95 間；FL 值閒置 .65 至 .88 間；AVE 值閒置 .41 至 .71 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。

***表 2** 各構面信度與收斂效度*

構面	M	SD	Cronbach's α	CR	FL	AVE
標準			**>0.7**	**>0.7**	**>0.5**	**>0.5**
態度	4.04	0.54	0.86	0.89	0.85	0.51

主觀規範	3.43	0.63	0.80	0.86	0.65	0.41
知覺行為控制	3.79	0.73	0.93	0.94	0.80	0.66
感知有用性	3.89	0.74	0.94	0.95	0.82	0.72
感知易用性	3,91	0.61	0.92	0.93	0.80	0.64
自我網路效能	3.94	0.62	0.93	0.94	0.87	0.68
行為意向	4.12	0.62	0.94	0.95	0.88	0.71

4.3. 路徑分析

本研究採用 Smart PLS 4 偏最小平方法(PLS-SEM)進行路徑分析，檢驗國小行政工作者使用生成式 AI 輔助行政業務的行為意向之結構模型，並探討計畫行為理論、科技接受模式及自我網路效能理論的整合模型之解釋力與路徑關係。如圖 2，本研究八條假設經分析得知，自我網路效能對感知有用性具有高度顯著正相關 (β =0.547, t = 3.405***, p < .001)；自我網路效能對感知易用性具有高度顯著正相關 (β =0.773, t = 12.02***, p < .001)；感知易用性對感知有用性的直接影響不顯著 (β =0.041, t = 0.3, p > .05)；感知有用性對行為意向有顯著正相關 (β =0.423, t = 2.866**, p < .01)；感知易用性對行為意向的直接影響不顯著 (β =0.321 , t =1.810, p > .05)；態度對行為意向有正相關 (β = 0.308*, t =2.088, p < .05)；主觀規範對行為意向的直接影響不顯著 (β =-0.027, t = 0.262, p > .05)；知覺行為控制對行為意向的直接影響不顯著 (β =-0.043, t = 0.29, p > .05)。

Bido、da Silva 與 Ringle (2014)提出 Cohen 之 f^2 意指經由涵蓋和消釋模型的構面所得到的數值，是一項較為罕見卻極為有效的應用，可用來計算出效果量標，其計算方式會透過解釋力(R^2)得出，算式為 $f^2 = R^2 / (1 - R^2)$，且 Cohen、Manion 與 Morrison (2007)指出當 f^2 大於 0.02 時為低效果量；當 f^2 大於 0.15 時為中效果量；當 f^2 大於 0.35 時為高效果量。如圖 2，自我網路效能對感知有用性解釋力為 34% (R^2 = 0.336)，具中效果量為 f^2 = 0.181；自我網路效能對感知易用性解釋力為 60% (R^2 = 0.597)，具高效果量為 f^2 = 1.483；感知有用性對行為意向解釋力為 75% (R^2 = 0.748)，具中效果量為 f^2 = 0.265；感知易用性對行為意向解釋力為 75% (R^2 = 0.748)，具低效果量 f^2 = 0.124；態度對行為意向解釋力為 75% (R^2 = 0.748)；具低效果量 f^2 = 0.119。

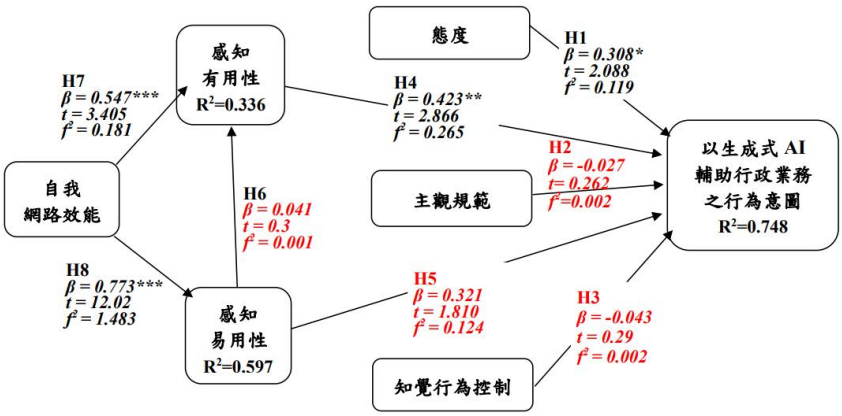


圖 2 路徑驗證分析圖

*p < .05, **p < .01, ***p < .001

5. 討論

5.1. 研究結果概述

本研究結果顯示，自我網路效能對感知有用性及感知易用性具有顯著正相關，表示在國小行政工作者中，提升自我網路效能對 AI 工具使用意向具有關鍵性作用。感知有用性直接影響行為意向，並受感知易用性和自我網路效能的間接影響，符合科技接受模式的假設。主觀規範對行為意向的影響力不顯著，反映國小行政工作者對生成式 AI 工具的使用意向受到他人觀點或社會壓力的影響不明顯。態度對行為意向的正向影響顯著，顯示行政工作者對生成式 AI 的使用決策依賴內在因素。這些結果支持了部分相關文獻的研究發現。例如，Huang et al. (2023)指出，在教育情境中，自我網路效能是影響教師接受新科技的重要因素，與本研究結果一致。同樣地，Wang & Lin (2022)研究指出，感知有用性在科技接受過程中具有更高的解釋力，顯示行政工作者對 AI 工具的效益感知是其使用行為的核心驅動力。然而，Chen & Lee (2021)研究發現，感知易用性在某些高複雜度的科技應用中對行為意向的影響不顯著，與本研究結果一致，進一步驗證生成式 AI 工具在教育行政應用中可能因學習曲線與操作複雜度而降低其直接影響力。根據研究結果，感知有用性和自我網路效能是行為意向的主要前因，而主觀規範與知覺行為控制之影響較不顯著。這些結果也與相關研究相符，例如，Zhou et al. (2022)指出，在組織環境中，使用科技的行為更可能受到個人內在動機與效能感的驅動，而非來自外部規範或控制感。這也可能反映出校園行政工作者在運用生成式 AI 時的自主性較高，進一步削弱主觀規範與知覺行為控制的影響力。研究結果強調了自我網路效能在技術採用中的重要性，國小行政工作者若具備較高的網路操作信心，會認為生成式 AI 工具不僅容易使用，且能幫助其提高工作效率。提升國小行政工作者對於網路工具操作的信心，是推動 AI 技術應用於校園行政的重要策略。

5.2 研究應用

根據本研究結果，可提供國小校園行政工作者自我網路效能的培訓方案，例如 AI 應用研習課程或線上增能課程，增進國小行政工作者適應並接受生成式 AI 等新科技工具的運用。自我網路效能的提升能顯著影響個體對科技的態度與使用意願(Bandura, 2021; Sun et al., 2023)。這些課程不僅應聚焦於生成式 AI 的技術應用，還應融入實務情境，提供模擬操作與解決問題的實際案例，進一步提升工作者的感知易用性與感知有用性，增進行政效率與創新能力。此外，增加社會支援(如同儕協作與技術支持團隊)有助於促進生成式 AI 在行政工作中的全面化應用(Huang & Lin, 2022)。在實務應用層面，可優先強化生成式 AI 作業功能的感知有用性與感知易用性。例如，優化 AI 工具的使用界面以提升友善度，針對國小行政需求進行功能定制，並提供持續的技術更新與支持服務。研究顯示，感知易用性與感知有用性是影響科技採用的核心因素，特別是在教育行政場域中(Davis et al., 2022)。同時，應提升校園內部對新科技的正向態度評價，建立由領導者和早期採用者組成的推動團隊，作為內部推廣力量，形成支持性的技術文化(Wang et al., 2023)。透過這樣的方式，不僅能提高生成式 AI 的接收度，也能促進整體行政服務的數位化轉型與創新發展。

6. 結論與建議

本研究以計畫行為理論與科技接受模式為基礎，並結合自我網路效能，探討國小行政工作者運用生成式 AI 輔助行政業務之行為意向。透過 Smart PLS 4 進行結構方程模型分析，驗證相關理論構面對行為意向的影響並提供生成式 AI 應用於校園行政領域的理論與實務啟示。研究結果顯示，自我網路效能對感知有用性與感知易用性具有顯著正相關，符合近年來相關研究指出，自我效能是影響個體採用新技術的重要因素(Bandura, 2021 Sun et al., 2023)。感知有用性亦顯著影響行為意向，支持科技接受模式的核心假設，即感知有用性是科技採用的關

鍵驅動力(Davis et al., 2022)。態度對行為意向具有顯著正相關，進一步強調情感性因素在科技接受中的重要性(Huang & Lin, 2022)。然而，本研究發現感知易用性、主觀規範與知覺行為控制對行為意向的直接影響不顯著，可能反映出生成式 AI 作為新興技術，其應用模式與傳統技術存在差異。這些結果不僅支持了部分理論假設，也反映生成式 AI 在校園行政領域應用上的獨特性，尤其是行政工作者的數位素養與對 AI 工具的使用效能感對其接受與行為意向的關鍵作用。本研究之貢獻在於建構並驗證一個整合性模型，說明生成式 AI 在校園行政中的應用潛力與影響因素，為後續研究提供理論基礎。特別是自我網路效能在科技接受與行為意向中的重要性，近年研究亦支持該觀點，指出提升個體的數位技能與信心可顯著促進新技術的應用(Wang et al., 2023)。實務上，研究建議教育行政單位應強化行政人員的數位能力培訓，並加強技術支持文化，以促進生成式 AI 的應用效益，提升行政效率與決策支持。然而，本研究亦存在若干限制，包括研究樣本僅針對國小行政工作者，可能限制其外部效度；採用問卷調查法可能導致社會期許效應，影響受試者回答的客觀性。未來研究可擴大樣本範圍至國中、高中或其他教育層級，並探索其他可能影響生成式 AI 應用的潛在因素，如組織文化、政策支持或資源分配(Zhao et al., 2022)，進一步豐富相關研究。總結而言，生成式 AI 在校園行政中的應用正逐漸成為提升行政效能與創新力的重要手段。本研究透過量化分析，揭示影響國小行政工作者行為意向的關鍵因素，為未來科技與教育行政結合的創新應用提供實踐方向與研究啟示。

參考文獻

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Bandura, A. (2021). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bido, D. S., da Silva, D., & Ringle, C. M. (2021). Structural equation modeling with small samples: Partial least squares and maximum likelihood. *International Journal of Business Research*, *21*(3), 133-150.

Chen, Y., & Lee, P. (2021). Factors influencing educators' adoption of AI in school administration: A TAM and TPB perspective. *Computers & Education*, *174*, 104294.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (2022). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982-1003.

Eastin, M. S., & LaRose, R. (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer-Mediated Communication, 6*(1).

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach.* Psychology Press.

Hassan, M. U., Iqbal, A., & Iqbal, Z. (2018). Factors affecting the adoption of internet banking in Pakistan: An integration of technology acceptance model and theory of planned behaviour. *International Journal of Business Information Systems*, *28*(3), 342-370.

Huang, J., Chen, H., & Zhou, S. (2023). The role of self-efficacy in technology adoption among school administrators: Insights from AI adoption. *Journal of Educational Technology Development, 39*(2), 45-67.

Huang, J., & Lin, Y. (2022). The role of social support in technology adoption: Evidence from AI applications in schools. *Journal of Educational Technology*, *40*(3), 312-328.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2006). Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. *Decision Support Systems, 42*(3), 1819-1834.

Lee, J., Lee, Y., & Hwang, S. (2018). The effects of school leadership on teachers' adoption of technology. *Educational Technology Research and Development, 66*(2), 403-420.

Lu, Y., Zhou, T., & Wang, B. (2009). Exploring Chinese users' acceptance of instant messaging using the theory of planned behavior, the technology acceptance model, and the flow theory. *Computers in Human Behavior*, *25*(1), 29-39.

Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education, 48*(2), 250-267.

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society, 12*(3), 150-162.

Sun, L., Liu, H., & Zhang, W. (2023). Exploring the impact of self-efficacy on AI technology adoption in school settings. *Educational Management Review*, *45*(4), 1023-1039.

Teo, T. (2011). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education, 57*(4), 2432-2440.

Teo, T., & Zhou, M. (2014). Explaining the intention to use technology among university students: A structural equation modeling approach. *Educational Computing Research, 51*(3), 313-334.

Tsai, M. J., & Tsai, C. C. (2003). Student computer self-efficacy and perception of web-based learning environment: A case study of Taiwanese middle school students. *The Internet and Higher Education, 6*(1), 41-56.

Tung, F. C., & Chang, S. C. (2008). Nursing students' behavioral intention to use online courses: A questionnaire survey. *International Journal of Nursing Studies, 45*(9), 1299-1309.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems, 17*(5), 328-376.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204.

Wu, J. H., Tennyson, R. D., & Hsia, T. L. (2008). A study of student satisfaction in a blended e-learning system environment. *Computers & Education, 50*(2), 358-366.

Wang, K., & Lin, Y. (2022). Examining the impact of perceived usefulness and ease of use on technology adoption in K-12 education: Evidence from AI tools. *Educational Management Administration & Leadership*, *50*(1), 102-117.

Wang, Q., Zhang, Y., & Hu, L. (2023). Organizational leadership and technology acceptance in education: A framework for AI implementation. *Educational Technology Research and Development*, *71*(1), 54-69.

Wang, X., & Lin, Y. (2023). Balancing innovation and human judgment in AI-powered school administration. *Journal of Educational Technology Integration, 29*(3), 112-125.

Yu, J. H., Ku, G. C. M., Lo, Y. C., Chen, C. H., & Hsu, C. H. (2021). Identifying the Antecedents of University Students' Usage Behaviour of Fitness Apps. Sustainability, *13*(16), 9043.

Zhou, L., Zhang, W., & Hu, C. (2022). Organizational influences on AI adoption in school management: The interplay of norms, efficacy, and infrastructure. *Educational Research and Reviews, 17*(5), 312-325.

Zhao, Y., Lei, J., & Frank, K. (2022). The interplay of organizational culture and technology adoption in schools. *Computers in Human Behavior*, *139*, 107502.

Zhao, Y., Wang, L., & Xia, L. (2021). Teachers' self-efficacy and acceptance of AI in education: The role of professional development. *Computers & Education, 175*, 104337.

創意知識工作者之創造力需求及工作投入: 以內在動機與自我效能為中介變項

Creative Knowledge Workers:

Study of Relationship between Need for Creativity at Work and Job Involvement

Using Intrinsic Motivation and Self-Efficacy as Mediator

龍蒂澈[1]，洪榮昭[2]，蔡其瑞[3]
[1]國立臺灣師範大學 創造力發展碩士在職專班
[2]國立臺灣師範大學 工業教育學系
[3]國立臺灣師範大學 科技應用與人力資源發展學系
013a02132@ntnu.edu.tw

【摘要】 創意知識工作者為創意導向產業的關鍵人才，如何促進其持續投入為組織重要議題。本研究以 131 名從事高創造性需求職務之工作者為樣本，探討創造力需求對工作投入之影響，並建構以自我效能與內在動機為中介之模型。結果顯示創造力需求可正向預測工作投入，自我效能與內在動機分別發揮完全與部分中介效果。本研究導入工作塑造與循環歷程觀點，說明創造性任務如何激發心理資源並轉化為持續投入，提供創意產業工作設計與人才管理新視角。

【關鍵字】 創意知識工作者；創造力需求；自我效能；內在動機；工作投入

Abstract: *This study surveyed 131 creative knowledge workers to examine how need for creativity influences job involvement via self-efficacy and intrinsic motivation. Results revealed full and partial mediation, respectively. Drawing on job crafting theory and the idea of a virtuous cycle of engagement, this study highlights how creative demands activate psychological resources and promote sustained job involvement, offering insights for talent management in creativity-driven industries.*

Keywords: Creative knowledge worker, need for creativity, self-efficacy, intrinsic motivation, job involvement

1. 前言

在當代創意與科技高度融合的產業環境中，創意知識工作者成為推動組織創新與文化轉型的關鍵角色。由於其工作內容多具挑戰性與不確定性，能否維持穩定且持續的心理投入，成為組織人力資源管理的重要課題。Sokół 與 Figurska(2017) 指出，創意知識工作者的參與與動機狀態將直接影響其創造力表現與組織績效，顯示其心理機制對於創新成果的關鍵性。

過去研究多從個體心理特質的角度出發，探討如自我效能、內在動機對創造力的預測作用，強調「工作投入是否能提升創造力」。然而，有研究已指出，當工作任務本身具備挑戰性與創造性需求時，更可能激發員工的內在心理資源，如效能感與動機品質(Amabile & Pratt, 2016; Gagné & Deci, 2005)，顯示任務特性亦可能成為參與動能的驅動來源。

本研究因此聚焦於創造力需求作為前因變項，探討當員工主觀感受到創造性任務要求時，是否能透過自我效能與內在動機等心理資源，引發其對工作的參與與投入。此觀點補足傳統

以創造力為結果變項的研究主軸，並導入工作塑造 (Job Crafting) 理論視角(Wrzesniewski & Dutton, 2001)，說明員工可能因應創造性需求而主動調整認知與行為策略，進而強化其心理投入。

此外，創造力需求與工作投入間亦可能存在正向循環歷程。創意任務特性與個體心理資源的交互作用，將強化挑戰與參與之間的連動關係，並產生正向回饋歷程(Amabile & Pratt, 2016)。因此，透過建構包含創造力需求、自我效能、內在動機與工作投入的中介模型，本研究期望釐清任務需求與心理歷程間的交互機制，並進一步提出對創意導向產業之工作設計與人才投入激發的理論與實務貢獻。

2. 文獻探索

本研究旨在探討創造力需求對工作投入的影響，並檢驗自我效能與內在動機的中介角色，理論基礎涵蓋需求層次、自我效能、自我決定與工作塑造理論。

根據 Maslow (1943, 1987)提出的需求層次理論，自我實現為人類追求創造、成就與價值實現的最高層級。近年研究亦指出，在知識密集產業中，創造力已成為工作任務的重要內涵，創造性需求亦被視為驅動員工投入的關鍵來源(Mehta, 2021)。然而，過往研究多以創造力為結果變項，關注工作投入、內在動機或自我效能對創造表現的預測效果，較少由創造力需求出發，探討其如何引發內在心理歷程並轉化為參與行為。

在自我效能理論中，Bandura (1977)指出，效能感是影響個體面對挑戰時行動選擇與持續性的核心心理資源。Tierney 與 Farmer (2002)亦進一步證實，創意自我效能可正向預測創造性行為與工作表現。另一方面，自我決定理論則強調動機品質的重要性，當個體的自主性、勝任感與關係感獲得滿足時，內在動機將被激發，進而提升心理參與與工作投入(Ryan & Deci, 2000；Gagné & Deci, 2005)。

此外，創造性需求與心理資源之間可能並非單向關係。有研究指出，任務特性與個體心理歷程可能構成正向循環歷程，彼此交互作用以強化動機、投入與創造表現間的連動性(Amabile & Pratt, 2016)。具挑戰性與意義感的任務亦被視為激發內在動機的重要情境條件(Gagné & Deci, 2005)。 這些觀點支持本研究的假設：創造力需求可啟動個體心理資源，進而促進其工作投入。

同時，Wrzesniewski 與 Dutton (2001)提出的工作塑造理論指出，員工會根據工作任務特性，主動調整自身對工作的認知與行為策略，例如改變工作方式、擴展職責範圍或重塑工作意義，以提升參與感與目標認同感。當任務本身包含創造性要求時，此一調整歷程可能促進個體的心理投入。此觀點補足了以創造力為結果變項的單向模型，亦支持本研究從創造力需求切入、整合任務特性與心理資源互動歷程之研究取向。

3. 研究架構與假設

3.1. 研究架構

本研究旨在探討創造力需求對工作投入的影響，並檢驗自我效能與內在動機的中介效果。根據文獻探討與理論基礎，建立如下研究架構(見圖 1)：

3.2. 研究假設

根據前述理論基礎與研究架構，本研究提出四項假設，以驗證變項間的關係。首先，假設 H_1：創造力需求對工作投入具有正向影響。進一步探討其作用機制，假設 H_2：自我效能在創造力需求與工作投入之間具中介效果；假設 H_3：內在動機亦在創造力需求與工作投入

之間具中介效果。最後，假設 H_4：自我效能與內在動機共同構成平行中介，影響創造力需求對工作投入的效果。

本研究之驗證模型如圖 1 虛線框所示。圖中左方虛線框外的「工作投入」為參考 Wrzesniewski 與 Dutton (2001) 工作塑造理論中所提之觀點，說明員工可能因高度投入而主動調整任務內容、尋求更多挑戰，進一步提升創造力需求。此循環歷程為本研究架構的理論依據之一，然本研究之實證分析聚焦於虛線框中「創造力需求→心理資源→工作投入」的前因與歷程關係。

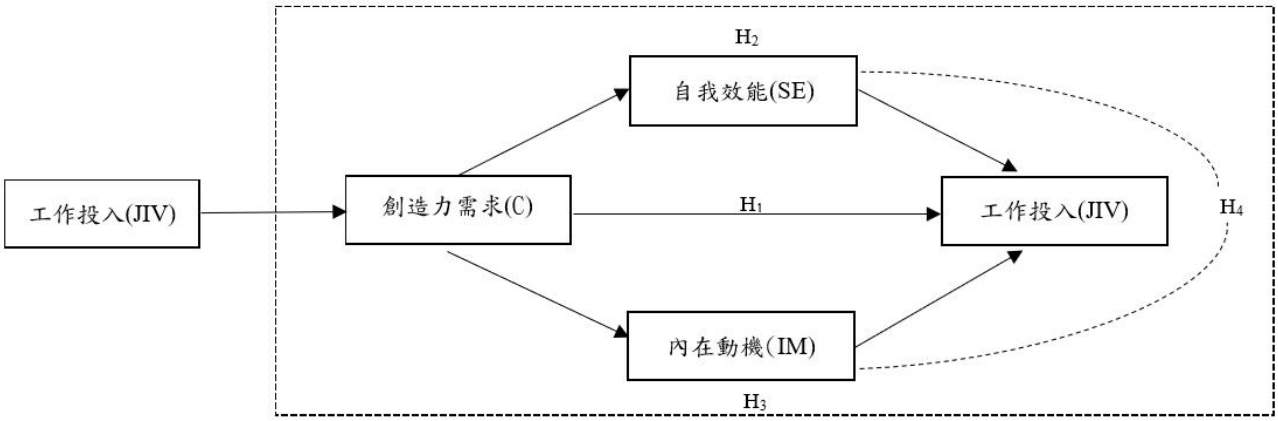


圖 1 研究架構

4.研究方法

4.1. 研究對象與樣本

本研究以臺灣地區 131 位創意知識工作者為調查對象，採用便利抽樣法進行資料收集。受試者平均年齡為 35 歲(SD=6.5)，男性占 45%，女性占 55%。樣本涵蓋科技、設計與文化創意產業等高創造力需求領域。

4.2. 研究工具

本研究問卷包含四個構面，皆採五點李克特量表(1 = 完全不同意，5 = 完全同意)進行評量。創造力需求量表參考 Tierney 與 Farmer (2002)所提出之概念，共 5 題，例如「我會尋找新的工作方法、技術或工具」。自我效能量表依據 Bandura (1977) 理論架構編製，共 6 題，例如「當我在工作中遇到問題時，我通常能找到多種解決方案」。內在動機量表改編自 Ryan 與 Deci (2000) 之自我決定理論，共 4 題，例如「我喜歡解決複雜且困難的問題」。工作投入量表則採用 Schaufeli 等人 (2002) 所發展之版本，共 9 題，例如「我會全神貫注於工作中」。上述量表經預試後，參考本研究兩位共同作者之專業意見進行語意調整與內容修訂，強化整體量表之清晰度與內容效度。Cronbach's alpha 值介於.71 至.86，顯示具有良好之內部一致性。惟因篇幅限制，完整題項未附於文中，若有需要可另行提供。

4.3. 研究程序

本研究於 2024 年 10 月執行，整體流程包括問卷修訂、正式施測與資料收集等單一階段作業。首先，根據既有量表進行修訂，經預試後調整語意與題項結構，以提升清晰度與內容效度。隨後進行正式施測，採便利抽樣方式，透過 Google 表單進行匿名線上問卷調查，施測時間約為 15 分鐘，調查期為三週。最終共回收有效問卷 131 份，所有資料僅用於本研究之學術分析。

4.4. 資料分析方法

本研究資料分析方法包括五項。首先，進行信效度分析，分別以 Cronbach’s Alpha 與組

合信度 (CR) 檢驗各量表之內部一致性，並透過確認性因素分析 (CFA) 與平均變異萃取量 (AVE) 評估構念之收斂與區別效度。其次，進行描述性統計，計算各構面之平均數與標準差，以掌握整體樣本於各變項上的表現趨勢。第三，採用皮爾森積差相關係數進行相關分析，以檢視主要變項間之關聯性。第四，使用獨立樣本 t 檢定與單因子變異數分析 (ANOVA)，以分析性別、年齡等背景變項在各構面上的差異情形。最後，運用階層迴歸分析與 SPSS PROCESS macro (Model 4) 進行中介效果分析，檢驗自我效能與內在動機於創造力需求與工作投入之間之中介角色。

5. 研究結果

5.1. 信、效度統計

鑒於各量表之信效度已於研究工具段落詳述，故本節不再重複說明。

5.2. 描述性統計、相關分析與區別效度

為掌握樣本在各構面上的表現與變項之間的關聯性，本研究進行各構面之描述性統計與相關分析，並同時檢視區別效度。表 2 顯示，各構面平均數介於 3.54 至 3.91，標準差介於.68 至.80，表示樣本於各變項上整體認同度偏高。相關分析顯示，主要變項間皆呈顯著正相關 ($p < .01$)，其中以創造力需求與自我效能之間相關性最高 ($r = .66$)。另根據 Fornell 與 Larcker (1981) 建議，構面對角線之 AVE 平方根值皆高於其餘相關係數，顯示各構面間具良好區別效度。

表 2 區別效度分析表及相關矩陣

註：N=131；平均數、標準差之計算係該量表題項加總而得；對角線的粗體值為 AVE 平方根數值、其他數值為相關係數；p<0.05 *、p<0.01 **

構面	平均值	標準差	C	SE	IM	JIV
1. C	3.91	.71	.78			
2. SE	3.83	.68	.66**	.78		
3. IM	3.54	.80	.46**	.60**	.73	
4. JIV	3.83	.68	.47**	.55**	.48**	.70

5.3. 差異性分析

本研究進一步探討背景變項在各構面上的差異情形。首先，針對性別進行獨立樣本 t 檢定，結果顯示男性在創造力需求 ($t = -2.98$, $p = .003$)、內在動機 ($t = -3.17$, $p = .002$)、自我效能 ($t = -3.17$, $p = .002$) 與工作投入 ($t = -2.79$, $p = .006$) 皆顯著高於女性 (見表 4)。此外，單因子變異數分析結果指出，不同年齡層在內在動機上具顯著差異 ($F = 3.15$, $p = .02$)，進一步的事後比較顯示，60～69 歲組別顯著高於 20～29 歲與 40～49 歲兩組別 (見表 3)。此結果說明背景變項可能影響其面對創造性任務時的認知與投入反應。

表 3 同年齡與內在動機之事後比較結果

	平均數（標準差）		自由度	*t*值	*p* 值
	男性（n=55）	女性（n=76）			
C	4.11（0.73）	3.75（0.67）	129	-2.98	.003
IM	3.79（0.85）	3.36（0.72）	129	-3.17	.002
SE	4.05（0.70）	3.67 （0.63）	129	-3.17	.002
JIV	4.03（0.76）	3.69（0.58）	96.76	-2.79	.006

資料來源：本研究。

表 4 性別變項之獨立樣 t 檢定結果

年齡組別（歲）	平均數	標準差	F 值	*p* 值	事後比較結果
(1) 20～29	3.22	0.80			
(2) 30～39	3.55	0.75			
(3) 40～49	3.40	0.83	3.15	.02	(5)>(1) (5)>(3)
(4) 50～59	3.78	0.72			
(5) 60～69	4.50	0.66			

資料來源：本研究。

5.4. 迴歸分析

為驗證本研究假設，採用階層迴歸分析與 Bootstrap 法進行統計檢驗。表 5 呈現控制背景變項後，創造力需求、自我效能與內在動機對工作投入的直接影響，用以檢驗 H2 與 H3。表 6 根據中介分析步驟建構多個模型，對應 H1 至 H4 假設，並透過解釋力（adj R^2）變化說明各變項介入對模型的影響。表 7 運用 Bootstrap 檢驗中介效果之顯著性，結果顯示內在動機具部分中介效果，自我效能則具完全中介效果，支持 H3 與 H4。三表合併呈現研究架構中各變項的直接與間接作用歷程，提供完整實證依據。

表 5 工作投入之階層迴歸分析結果

	依變數(JIV)	
	模型 1(控制變數)	模型 2
常數	3.94	1.44
性別	-.24**	-.06
JT=少於 3 年	0.02	.01
JT=3 年~未滿 6 年	0.11	.06
婚姻	-.01	-.06
E=高中以下	.01	.02
E=研究所以上	.12	.02
自變數 1 (C)	-	.15
自變數 2 (IM)	-	.22*
自變數 3 (SE)	-	.32**
R^2	.09	0.37
Adj-R^2	.04	0.32
F 值	2.00	7.81**

資料來源：本研究。

表 6 中介效果之線性階層迴歸分析表

依變項 JIV								
	H_1	H_2			H_3			H_4
解釋變數	模型 1	模型 2 (SE)	模型 3	模型 4	模型 5 (IM)	模型 6	模型 7	模型 8
自變項								
C	.47***	.66***		.18	.52***		.31***	.15
中介								

變項								
IM	-	-		-	-	.48***	.34***	.22*
SE	-	-	.553***	.44***	-	-	-	.32**
F 值	35.55***	100.08***	56.86***	30.59***	34.37***	39.32***	28.55***	23.33***
$adj\ R^2$	21%	43.2%	30.1%	31.3%	20.4%	22.8%	29.8%	34.0%
R^2	21.6%	43.7%	30.6%	32.3%	21.0%	23.4%	30.8%	35.5%

註：p<0.05 *、p<0.01 **、p<0.001 ***; H2 (模型 2、3、4)驗證 SE 之中介效果，H3 (模型 5、6、7)驗證 IM。

表 7 間接效應分析表

總效果 C → JIV	直接效果 C → JIV	關係	間接效果	信賴區間 上限	信賴區間 下限	t 值	結論
0.45***	0.14	H3: C→IM→JIV	0.20	0.04	0.39	2.41	部分中介
		H4: C→SE→JIV	0.10	0.00	0.22	1.85	完全中介

註：p<0.05 *、p<0.01 **、p<0.001 ***

6. 討論與結論

6.1. 創造力需求作為工作投入的啟動引擎

本研究驗證創造力需求對工作投入具有顯著正向影響，亦呼應 Maslow (1943) 自我實現需求之理論假設，說明當任務內容對創造性有明確要求時，將激發員工對挑戰的心理回應與投入行為。Mehta (2021) 指出，現代職場的動機已轉向自主性、勝任感與關係感的追求，創造性任務需求正融合這三項驅動力，進一步支持本研究之任務導向視角。此結果補足傳統以創造力為結果的研究主軸，提供從「任務特性驅動心理參與」的新觀點。

6.2. 自我效能與投入連結之路徑

自我效能在本研究中扮演完全中介角色，說明創造力需求會透過提升個體對挑戰的掌控感與成功預期，引發後續的投入行為。此結果與 Tierney 與 Farmer (2002) 對創意自我效能的說明一致，顯示當員工認為自己能勝任創造性任務時，更願意投入心力應對挑戰，有助於創造出積極的工作投入歷程。

6.3. 內在動機驅動的主動參與

研究結果亦顯示，內在動機具部分中介效果，支持當工作內容具備挑戰性與意義時，員工會因內在樂趣與自我挑戰而投入更多心理資源 (Gagné & Deci, 2005)。 此過程不僅反映個體的主觀認知與價值認同，亦回應創造性任務中常見的高自主性與解決問題需求，進而引發持續投入行為。

6.4. 雙心理資源之平行作用

整體模型驗證顯示，自我效能與內在動機同時扮演中介角色，提供心理信念與行動意願兩種不同層次的資源補充。此結果回應 Amabile 與 Pratt (2016) 對心理資源互動性的描述，也呼應本研究以「創造性需求－心理資源－行為投入」為主軸的動能轉化模型，說明挑戰型任務特性可透過雙重心理機制，轉化為穩定的工作投入。

6.5. 背景變項差異的初步觀察

研究亦顯示，在性別與年齡等背景變項上呈現顯著差異：男性相對於女性在四項變數皆得分較高，60 歲以上族群的內在動機亦顯著高於年輕族群。此現象可能與社會角色社會化

或世代經驗差異相關 (Eagly & Wood, 2012)，未來研究可針對性別與年齡如何影響創造性需求的認知與心理資源的運用進行深入分析，以豐富目前以量化驗證為主的初步觀察。

6.6. 結論、侷限與研究展望

本研究雖建構以創造力需求為前因，透過自我效能與內在動機解釋工作投入之心理歷程，惟研究設計與樣本來源仍具若干限制，未來可持續延伸與補足。

首先，本研究採用橫斷式設計，有助於驗證變項間的關聯與中介效果，但無法掌握心理歷程隨時間推移的動態變化。尤其在創造性任務中，任務需求與心理資源可能構成互為影響的循環歷程 (Amabile & Pratt, 2016)，例如具挑戰性的任務激發動機，而高度動機又可能促使員工主動追求更多創造性挑戰 (Gagné & Deci, 2005)。此外，部分研究亦指出內在動機亦可能反過來影響個體對創造性任務的偏好與參與。未來研究可進一步驗證變項間的雙向與時間性影響，建議採用縱貫式設計或情境追蹤方法 (例如透過每日即時填答或事件發生後之經驗回報；experience sampling method, ESM)，以即時蒐集創造力需求、動機狀態與投入表現的動態變化，進一步描繪心理資源啟動與行為參與之歷程與時間序列邏輯。

其次，本研究樣本涵蓋多元產業與職務類型，雖有助於提升理論的外部效度，但不同產業背景與工作設計可能導致員工對創造力需求的主觀感知有所差異。特別是，工作角色所承擔之任務屬性與職務自主性的高低，將影響個體是否覺察到來自工作任務的創造性要求。例如，高度自主性的知識型職位 (如設計、專案管理) 通常需員工主動解決問題與提出新方案，較易觸發創造力需求的感知；反之，任務結構清晰或流程重複的職務則可能削弱此類需求的心理覺察。後續研究可聚焦特定職能或採用多層次分析架構，進一步探討工作性質與職務特徵在創造力歷程中的調節角色。

最後，本研究以自我效能與內在動機為中介變項，強調個體內部心理資源之作用，但尚未納入外在激勵條件 (如領導風格、回饋機制、團隊氣候) 作為潛在調節因素。未來可從個體與情境交互觀點切入，進一步建構跨層次模型，探討組織環境如何影響個體對創造性任務需求的感知與心理資源之調用與運用。

參考文獻

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36, 157–183. https://doi.org/10.1016/j.riob.2016.10.001

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Chen, L., Zhang, W., & Wang, X. (2025). Competence drives interest or vice versa: Untangling the bidirectional relationships between creative self-efficacy and intrinsic motivation for creativity. *Creativity Research Journal*, 37(1), 1–14. https://doi.org/10.1080/10400419.2024.2284085

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 2, pp. 458–476). SAGE Publications. https://doi.org/10.4135/9781446249222.n49

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4), 331–362. https://doi.org/10.1002/job.322

Lin, B.-Y. (2022). The effects of intrinsic motivation on employees' work involvement: Self-efficacy as a moderating variable [Unpublished master's thesis]. National Ilan University.

Liu, D., Chen, X. P., & Yao, X. (2016). The motivational mechanism of employee creativity: A meta-analytic examination and theoretical extension of the creativity literature. *Organizational Behavior and Human Decision Processes*, 137, 236–263. https://doi.org/10.1016/j.obhdp.2016.08.001

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396. https://doi.org/10.1037/h0054346

Mehta, R. (2021). Motivation in the workplace: Reassessing Maslow in the modern organizational context. *Journal of Organizational Psychology*, 21(1), 57–68. (n.d.). Retrieved from URL https://doij.org/10.10000/IJLMH.11544

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two-sample confirmatory factor analytic approach. *Journal of Happiness Studies*, 3(1), 71–92. https://doi.org/10.1023/A:1015630930326

Sokół, A., & Figurska, I. (2017). The importance of creative knowledge workers in organizations. *Human Resource Management & Ergonomics*, 11(1), 115–132. https://doi.org/10.3390/en14206751

Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Potential antecedents and relationship to creative performance. *Academy of Management Journal*, 45(6), 1137–1148. https://doi.org/10.5465/3069429

Wrzesniewski, A., & Dutton, J. E. (2001). Crafting a job: Revisioning employees as active crafters of their work. *Academy of Management Review*, 26(2), 179–201. https://doi.org/10.5465/amr.2001.4378011

A Study on the Impact of Embodiment in AR Learning on Learners' Technology Acceptance

Xiaochen Suo[1], Baoyuan Yin[1]*
[1]School of Education Science, Harbin Normal University
* yinbaoyuan@hrbnu.edu.cn

Abstract : *This study focuses on the impact of embodiment in AR learning on learners' technology acceptance. Based on the embodied AR application (Experiment group) designed and developed in this study, a controlled experiment was conducted with a common AR application (Control group). The participants were 122 senior high school students from China, with 63 in the experimental group and 59 in the control group. The research findings show that embodiment in AR can significantly improve learners' perceived usefulness of AR technology, but adding embodiment to AR makes learners perceive a lower ease of use. This study will provide empirical support for the design of embodiment in AR.*
Keywords: AR, technology acceptance, AR learning, embodied

1. Introduction

Augmented reality (AR) technology has evolved relying on computer graphics, human-computer interface technology, and sensing technology, and has drawn substantial attention in recent years. The authoritative Horizon Report has listed the educational applications of AR as an educational technology to be adopted in the medium and short term for several consecutive years (Gao & Huang, 2017). AR disrupts the conventional education paradigm, which is circumscribed by temporal and spatial constraints, via its visual effects that blend real and virtual elements and its portability (Ciloglu & Ustun, 2023). Within an AR setting, learners are able to undergo a rich state of flow, emotional experiences, and other immersive encounters (Abdul Hanid et al., 2022; Ciloglu & Ustun, 2023; Weerasinghe et al., 2022). These experiences, in turn, serve as a catalyst for their active engagement in the learning process and the construction of knowledge, thereby promoting the development of their thinking skills and capabilities. Nevertheless, the advantages of implementing AR in the educational domain extend beyond these aspects. AR has the capacity to substitute the abstract experiences denoted by textual symbols with intuitive ones through the utilization of virtual models or embodied interactions (Cai et al., 2019; O'Meara & Szita, 2021), thereby further augmenting learners' educational experiences.

The theory of embodied cognition posits that knowledge concepts can be comprehended and acquired through somatic experiences, and that individuals are capable of engaging in cognitive processes through physical involvement in actions (Borghi & Cimatti, 2010). Moreover, embodied cognition places a strong emphasis on the subject's physical engagement and perceptual awareness, underscoring the body's agency and its significance within the realm of experience (Zahavi, 2003). AR provides learners with experiences that are intertwined with the real world, digitally generated objects, and their own bodies. Learners can interact with the components within the AR environment by means of physical movements in the real space, thus attaining experiences of

embodied participation. Technology acceptance has become a theoretical framework model that can significantly predict and reflect individuals' intentions to use information technology. Numerous studies have applied AR in educational settings and measured technological acceptance, revealing that learners' PU and PEU are generally favorable (Iqbal & Sidhu, 2022; Liu et al., 2023). However, does the incorporation of embodiment into AR incontrovertibly bring about benefits for students' technological acceptance? This is the pivotal research question that this study endeavors to explore. Therefore, this study takes the inclusion or exclusion of embodiment as the independent variable, with two groups set up: with embodiment (control group) and without embodiment (experimental group). The dependent variables are technology acceptance (perceived usefulness, perceived ease of use), to examine the effect of embodiment in AR on learners' technology acceptance.

2 . Theoretical framework

The Technology Acceptance Model (TAM), put forward by Davis, is predominantly applied to elucidate individuals' acceptance of information technology while they carry out tasks (Yang & Yoo, 2004). This model introduces two fundamental concepts: perceived usefulness (PU) and perceived ease of use (PEU) (Davis & Davis, 1989). PU is defined as "the extent to which a person believes that using a specific system would enhance their job performance." It is associated with work efficiency, productivity (in terms of time - savings), and the system's relative significance to an individual's work. PEU refers to "the degree to which a person believes that using a particular system is effortless." This concept encompasses both physical and cognitive exertion, along with the ease of learning (Venkatesh & Davis, 2007).

In this study, PU and PEU of technology acceptance are treated as outcome variables to explore the influence of embodiment in AR. As an emerging technology, AR needs to be embraced by users to fully unlock its educational potential (Al - Rahmi et al., 2019). When learners use AR, they typically demonstrate high levels of technology acceptance, this is likely because of AR's strong engaging nature and interactive allure, which have a positive impact on learners' subjective perceptions (Liu et al., 2023; Su et al., 2022). In the direction of this study, the following questions are handled:

1.What is the impact of the embodiment in AR experiences on learners' PU of technological acceptance?

2.What is the impact of the embodiment in AR experiences on learners' PEU of technological acceptance?

The following research hypotheses are proposed:

Hypothesis 1: The embodiment in AR has a positive effect on learners' PU of the technology.

Hypothesis 2: The embodiment in AR has a positive effect on learners' PEU of the technology.

3. Methodology

3.1. AR applications

The AR applications in this research are chiefly designed for high - school students to learn the abstract biological concept of "cell structure". The content is drawn from the textbook utilized by first - year high - school students in China (Biology, Compulsory Volume 1, People's Education Press). To guarantee the effectiveness of the AR application in real - world teaching, three Chinese high - school biology teachers were enlisted to offer guidance throughout the design and

development of the AR application. Once the development of the AR application was finalized, an evaluation was conducted. All three teachers opined that the AR application in this study is appropriate for students to use during the biology class of the first grade in senior high school. This study features two AR applications. One is embodied, and the other is non-embodied.

The embodied AR application in this study incorporates an embodiment with haptic feedback. Haptic feedback represents an interactive experience that furnishes learners with a realistic sensory perception, mainly incorporating vibrations or force - based feedback. In this study, the haptic feedback experience is intended to offer learners force - based feedback. Learners can touch specific areas of the AR marker, which will trigger the enlargement of corresponding virtual organelles, as depicted in Figure 1.

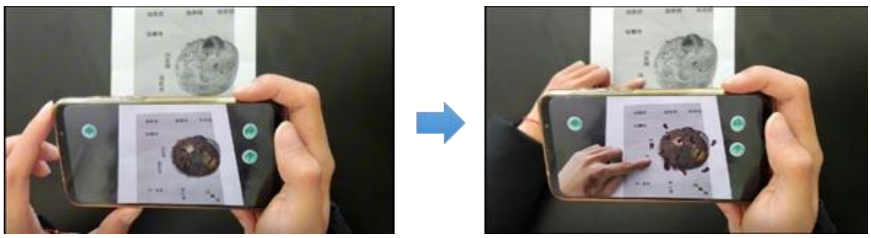


Figure 1 The process of haptic feedback experience

3.2. Participants

To guarantee the relevance of the knowledge within the AR applications for the participants, 122 first - year students from regular high schools in China were recruited (60 males and 62 females, with an average age of 15.8 years). Participation was on a voluntary basis. The participants were randomly allocated into two groups: Control group with 59 participants and Experient group with 63 participants. Each group had an approximately equal gender distribution. Every participant was provided access to the experimental equipment—smartphones and AR markers—for the conduct of learning activities.

3.3. Measures

3.3.1. Survey for PU of Technology Acceptance

The measurement of PU was adapted from Moore and Benbasat's (1991) scale regarding the advantages of technology use. The objective was to gather participants' feedback on the utility of AR applications in the learning process. This scale consists of five nine - point Likert - scale items. It exhibits a Cronbach's alpha of 0.831, signifying good reliability. Through factor analysis, it yields a KMO value of 0.776. The results of Bartlett's Test are $\chi^2 = 233.126$ ($p < 0.001$), demonstrating its validity. These figures suggest that the scale effectively measures PU in AR - based learning, with good reliability and construct validity.

3.3.2.Survey for PEU of Technology Acceptance

The PEU measurement was adapted from Venkatesh and Davis's (2000) perceived ease of use scale. It was used to gather participants' feedback on the ease of using AR applications for learning. The scale, featuring five nine - point Likert - scale items, has a Cronbach's alpha of 0.765 for reliability. Factor analysis shows a KMO value of 0.746 and Bartlett's Test results of $\chi^2 = 145.51$ ($p < 0.001$) for validity, indicating that the scale demonstrates good reliability and construct validity for measuring PEU in AR - based learning.

3.4. Procedures

The initial step entailed pre - experiment preparation. This involved acquainting participants with AR concepts and elucidating the experimental guidelines, a process that spanned 10 minutes. The introduction of AR concepts was intended to counteract any adverse impacts on the learning process stemming from participants' lack of familiarity with AR concepts or the AR learning environment during their inaugural AR learning encounter. This was achieved through teacher - led instruction, enabling participants to gain an understanding of AR technology and the AR learning environment. The explanation of experimental guidelines was designed to govern participants' actions during the experiment. It detailed the procedures, format, and duration of the AR learning experiment, while also highlighting the prohibition of activities unrelated to AR learning.

The subsequent step was AR learning, during which the two groups of participants were tasked with using different AR applications. This phase lasted for 20 minutes. Participants in the Control group undertook AR experiences through observation. In contrast, participants in the Experiment group engaged in AR experiences featuring haptic - feedback - enabled embodied interactions.

The final step involved post - experiment questionnaire measurement. The questionnaire incorporated the technology acceptance (PU & PEU) scale. Participants were given 10 minutes to complete the questionnaire.

3.5. Data analysis

The data were analyzed by utilizing SPSS 26.0. The outcomes of the measured variables in this research all adhered to the normal distribution pattern, and they had successfully passed the test for the homogeneity of variances. Consequently, an independent samples t-test was carried out (with the Cohen's d effect size being computed through the use of G*Power) in order to examine and analyze the disparities in the measurement results between the two experimental groups.

4.Result

This study aimed to measure differences in technology acceptance among learners after using AR applications of embodiment. To achieve this, a independent samples t-tests were employed to explore differences in prior knowledge scores, cognitive load and technology acceptance between the two groups. The test results are presented in Table 1.

Table 1. T-test Results for the technology acceptance based on the embodiment in AR learning.

Variable		Control group(N=59)		Experiment group(N=63)		t	Cohen's d
		Mean	SD	Mean	SD		
Technology Acceptance	PU	27.93	7.521	34.96	6.428	-5.532**	-1.006
	PEU	33.34	6.064	26.89	7.01	5.445**	0.981

**p<0.001

Independent samples t - test results for technology acceptance revealed that the level of embodiment in AR experiences had a substantial influence on the PU and PEU of technology acceptance, as presented in Table 3.

Regarding PU, the Control group, with a mean (M) of 27.93 and a standard deviation (SD) of 7.521, had significantly lower scores than the Experiment group, which had an M of 34.96 and an SD of 6.428. The t - value was - 5.532, with a p - value less than 0.001, and the effect size (d) was -

1.006, indicating a high - level effect. This suggests that when using AR applications, AR with embodiment led to a higher level of PU compared to AR without embodiment.

As for PEU, the Control group, having an M of 33.34 and an SD of 6.064, scored significantly higher than the Experiment group, which had an M of 26.89 and an SD of 7.01. The t - value was 5.445, with a p - value less than 0.001, and the effect size (d) was 0.981, also showing a high - level effect. This implies that when using AR applications, AR without embodiment led to a higher level of PEU compared to AR with embodiment.

The research contributions lie in revealing, via independent samples t-tests, that embodiment in AR enhances perceived usefulness (PU) but reduces perceived ease of use (PEU), highlighting a critical trade-off between functional immersion and usability. This offers practical insights for AR design, urging developers to balance embodied features to optimize both PU and PEU. The findings also lay groundwork for future research by identifying the need to simplify embodied interactions and explore user preferences across demographics, advancing more effective and user-friendly AR educational applications.

5. Discussion

The findings from the independent samples t - test in this study present intriguing insights into the impact of embodiment in AR experiences on technology acceptance, specifically in terms of PU and PEU.

Regarding PU, the significant difference between the Control group (non - embodied AR) and the Experiment group (embodied AR) is notable. The higher PU scores in the Experiment group suggest that the embodied AR experience makes users perceive the application as more useful. This could be because the haptic feedback and embodied interactions in the AR application provided a more immersive and engaging learning environment. Users may have felt that these embodied features enhanced their understanding and learning outcomes, thereby increasing the perceived value of the application. In educational contexts, this implies that incorporating embodied elements in AR can potentially boost students' belief in the utility of the technology for learning.

Conversely, for PEU, the Control group (non - embodied AR) had higher scores. This indicates that users found the non - embodied AR application easier to use. It might be that the additional haptic and embodied interaction elements in the experimental AR application added complexity, making it seem less intuitive. This finding is important as it highlights a potential trade - off between the added functionality of embodied AR (leading to higher PU) and its impact on ease of use.

Overall, these results have implications for the design of AR applications. Developers need to carefully balance the inclusion of embodied features to optimize both PU and PEU. Future research could explore ways to simplify embodied interactions to maintain high PEU while still capitalizing on the benefits for PU. Additionally, investigating user preferences across different demographics could further inform the design of more user - friendly and effective AR learning applications.

References

Abdul Hanid, M. F., Mohamad Said, M. N. H., Yahaya, N., & Abdullah, Z. (2022). Effects of augmented reality application integration with computational thinking in geometry

topics. *Education and Information Technologies*, *27*(7), 9485-9521. https://doi.org/10.1007/s10639-022-10994-w

Al-Rahmi, W. M., Yahaya, N., Aldraiweesh, A. A., Alamri, M. M., Aljarboa, N. A., Alturki, U., & Aljeraiwi, A. A. (2019). Integrating Technology Acceptance Model With Innovation Diffusion Theory: An Empirical Investigation on Students' Intention to Use E-Learning Systems. *IEEE Access*, *7*, 26797-26809. https://doi.org/10.1109/ACCESS.2019.2899368

Borghi, A. M., & Cimatti, F. (2010). Embodied cognition and beyond: Acting and sensing the body. *Neuropsychologia, 48*(3), 763-773. https://doi.org/10.1016/j.neuropsychologia.2009.10.029

Cai, S., Liu, E., Yang, Y., & Liang, J. C. (2019). Tablet-based AR technology: Impacts on students' conceptions and approaches to learning mathematics according to their self-efficacy. *British Journal of Educational Technology, 50*(1), 248-263. https://doi.org/10.1111/bjet.12718

Ciloglu, T., & Ustun, A. B. (2023). The Effects of Mobile AR-based Biology Learning Experience on Students' Motivation, Self-Efficacy, and Attitudes in Online Learning. *Journal of Science Education and Technology, 32*(3), 309-337. https://doi.org/10.1007/s10956-023-10030-7

Ciloglu, T., & Ustun, A. B. (2023). The Effects of Mobile AR-based Biology Learning Experience on Students' Motivation, Self-Efficacy, and Attitudes in Online Learning. *Journal of Science Education and Technology, 32*(3), 309-337. https://doi.org/10.1007/s10956-023-10030-7

Davis, F., & Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*, 319. https://doi.org/10.2307/249008

Gao, Y., & Huang, R. H. (2017). Interpretation and inspiration of *2017 NMC Horizon Report on the Future of Higher Education Technology in China: A Regional Report of the Horizon Project*. *E - Education Research, 38*(4), 15 - 22. 10.13811/j.cnki.eer.2017.04.002

Iqbal, J., & Sidhu, M. S. (2022). Acceptance of dance training system based on augmented reality and technology acceptance model (TAM). *Virtual Reality, 26*(1), 33-54. https://doi.org/10.1007/s10055-021-00529-y

Liu, Y., Sun, J. C.-Y., & Chen, S.-K. (2023). Comparing technology acceptance of AR-based and 3D map-based mobile library applications: a multigroup SEM analysis. *Interactive Learning Environments, 31*(7), 4156-4170. https://doi.org/10.1080/10494820.2021.1955271

Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research, 2*(3), 192-222. https://doi.org/10.1287/isre.2.3.192

O'Meara, J., & Szita, K. (2021). AR cinema: Visual storytelling and embodied experiences with augmented reality filters and backgrounds. *PRESENCE: Virtual and Augmented Reality, 30*, 99-123. https://doi.org/10.1162/pres_a_00376

Su, Y.-S., Lai, C.-C., Wu, T.-K., & Lai, C.-F. (2022). The effects of applying an augmented reality English teaching system on students' STEAM learning perceptions and technology acceptance. *Frontiers in psychology*, *13*, 996162. https://doi.org/10.3389/fpsyg.2022.996162

Venkatesh, V., & Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*, 186-204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., & Davis, F. (2007). A Model of the Antecedents of Perceived Ease of Use: Development and Test*. *Decision Sciences, 27*, 451-481. https://doi.org/10.1111/j.1540-5915.1996.tb00860.x

Weerasinghe, M., Biener, V., Grubert, J., Quigley, A., Toniolo, A., Pucihar, K. Č., & Kljun, M. (2022). Vocabulary: Learning vocabulary in ar supported by keyword visualisations. *IEEE Transactions on Visualization and Computer Graphics, 28*(11), 3748-3758. https://doi.org/10.1109/TVCG.2022.3203116

Yang, H.-d., & Yoo, Y. (2004). It's all about attitude: revisiting the technology acceptance model. *Decision Support Systems, 38*(1), 19-31. https://doi.org/10.1016/S0167-9236(03)00062-9

Zahavi, D. (2003). *Husserl's Phenomenology*. Stanford University Press.

智能时代下 GenAI 的使用对学生批判性思维影响研究——基于 UTAUT 模型的使用行为机制解释与认知负荷的调节作用

The Impact of GenAI Usage on Students' Critical Thinking in the Intelligent Era: Based on Mechanistic Explanation through UTAUT-based Behavior Analysis and Cognitive Load Moderation

李琳[1]，陈文俊[1]，刘妍[1*]
[1]上海交通大学教育学院
[*] flyliuyan0707@sjtu.edu.cn

【摘要】批判性思维是智能时代人才的核心素养之一。GenAI 技术推动了教学模式由知识传授转向思维能力培养，但其在课堂中的应用可能会造成学生认知负荷超载等问题，抑制其批判性思维发展。本研究以 UTAUT 模型为基础，探讨了某高校学生 GenAI 使用行为对其批判性思维的影响，进一步挖掘认知负荷的调节作用。研究发现绩效期望、努力期望和社群影响能通过增强使用意愿促进使用行为，进而促进批判性思维发展，并且认知负荷负向调节这一过程。因此，鼓励教师积极正确使用 GenAI 以积极影响学生使用行为，搭建专有教育智能体以降低认知负荷，促进学生批判性思维发展。

【关键词】生成式人工智能(GenAI)；批判性思维；UTAUT；认知负荷；调节作用

Abstract: *Critical thinking stands as one of the core competencies for talent in the intelligent era. While GenAI technology has driven the transformation of teaching models from knowledge transmission to cognitive skill cultivation, its classroom application may induce issues such as cognitive overload among students, potentially hindering the development of critical thinking. Grounded in the Unified Theory of Acceptance and Use of Technology (UTAUT) model, this study investigates how GenAI usage behaviors among university students influence their critical thinking, with particular emphasis on the moderating role of cognitive load. The research findings indicate that performance expectancy, effort expectancy, and social influence enhance usage intention, thereby promoting GenAI adoption and subsequently fostering critical thinking development. Notably, cognitive load exerts a negative moderating effect on this process. Consequently, we recommend that educators actively and appropriately employ GenAI to positively shape student engagement, while establishing dedicated pedagogical agents to mitigate cognitive load, thereby advancing students' critical thinking capabilities.*

Keywords: Generative Artificial Intelligence (GenAI), Critical thinking, UTAUT, Cognitive load, Moderating effect

1. 前言

随着生成式人工智能（Generative Artificial Intelligence, 以下简称“GenAI”）的发展，人类已经迈入智能时代（黄荣怀等，2025）。2024 年，联合国教科文组织发布了学生人工智能素养框架，其中提到学生应该具备批判性使用人工智能的能力（Miao et al., 2024），这有

赖于学生批判性思维的发展。作为21世纪的核心素养之一，批判性思维要求学生能够在不同的情境下保持质疑批判的态度，基于客观证据进行理性思考，综合多角度分析形成复杂问题解决方案，并且能够对自身行为进行监控、反思与修正（马利红等，2020）。

GenAI深度赋能教育领域推动了教学范式的变革（祝智庭等，2023）。在GenAI赋能的课堂中，人才培养的目标和模式由知识传授聚焦高阶能力培养（王学男 & 李永智，2024）。GenAI强大的功能可以帮助学生深层感知学习目标，实现对知识的深度建构、迁移以及创新，提升学生学习效果，助力学生高阶思维发展。（闫寒冰等，2025；曾明星等，2024）。因此，学生对于GenAI的技术接受与使用对其在教育范式的巨大变革下的适应十分重要（孙立会&周亮，2025）。但是GenAI虽然扩展了学习资源的可及性，却可能因为学习者元认知调节能力与信息环境不匹配引发双重认知困境：即海量信息资源导致的认知负荷超载与源于动态筛选机制缺失导致的选择焦虑泛化，对学生在学习过程中的高阶思维发展造成一定的影响（余胜泉 & 汪凡淙，2023）。

整合技术接受与使用模型（United Theory of Acceptance and Use of Technology, UTAUT）是Venkatesh等（2003）提出的用于解释使用者对于新技术使用的理论模型。本研究基于该模型对GenAI使用行为的解释，探究GenAI的使用对学生批判性思维的影响，并进一步分析学生认知负荷在这一过程中的调节作用，旨在为GenAI赋能的高阶思维培养课堂提供实用性建议。因此，本研究主要回答以下两个问题：基于UTAUT模型，学生GenAI使用行为如何影响其批判性思维？认知负荷如何调节GenAI使用对批判性思维的作用？

2. 文献综述

2.1 智能时代下的批判性思维研究

批判性思维是学生核心素养的重要内容，强调学生有目的、自我调整的判断（Facione, 1990）。一直以来，研究者们通过不同的教学方式和教学活动探究如何培养学生的批判性思维（廖彦霖，2023；罗宇晨 & 郑燕林，2024）。近几年，GenAI凭借其强大的信息处理和交互对话能力，在以知识为核心的行业，尤其是教育领域产生了重大影响（涂明辉 & 谢丹，2024），学者们也开始关注GenAI工具的使用对学生批判性思维的潜在影响。有学者提出，由于过度依赖，缺乏现实互动经验等问题，GenAI在培养学生高阶能力领域可能面临风险（段世飞 & 钱跳跳，2024）。但更多的实证研究表明，正确地使用GenAI有助于促进批判性思维的发展。戚佳（2024）等人的研究发现，使用GenAI工具的学生的批判性思维和自主学习能力显著高于未使用此类工具的学生；李曼丽等人（2025）的研究发现，增加大语言模型工具的基础执行应用与深度创意应用频次，均对高阶思维能力发展具有显著的正向影响。此外，国外学者的一些近期研究也指向人工智能的使用对批判性思维具有显著正向影响（Guo & Lee, 2023；Suriano et al., 2025）。目前大部分研究只对使用行为和批判性思维影响关系进行探讨，忽视了其他影响因素的存在。基于此，本研究将进一步结合GenAI使用行为的影响机制以及相关调节变量，对二者关系进行深入探究。

2.2 UTAUT模型对技术使用行为的解释

技术接受是指用户对某种技术的接受程度，在新技术不断层出不穷的21世纪，技术的使用往往会对一个领域造成变革性的影响，人们对于技术的接受与使用会直接影响技术的经济价值和发展前景（祝智庭等，2024）。Venkatesh等（2003）通过整合技术接受模型（Technology Acceptance Model, TAM）在内的八个理论提出了整合型技术接受与使用模型（United Theory of Acceptance and Use of Technology, UTAUT），认为人们面对一项新技术时，

人们的绩效期望（认为技术能够提升工作绩效的程度）、努力期望（认为使用技术的难易程度）、社群影响（来自他人或群体的建议对使用技术的影响程度）能够正向影响其对于新技术的使用意愿，并且结合使用技术的便利条件，共同影响新技术的使用行为。在教育领域，联合国教科文组织强调新技术的引用必须服务于学生的体验、收获和成长模式，因而学生对于 GenAI 技术的接受度与使用行为对于解释学生学习体验、知识收获与思维发展是尤为重要的（UNESCO, 1998）。而考虑到教育的特殊性，研究者们在 UTAUT 模型的基础上，增加了其他变量来探究教育技术使用对学生学习行为的影响。Almaiah 等（2019）利用 UTAUT 模型探究了本科生使用移动在线学习系统的行为影响因素，发现学生对于技术的信任能够对学生的使用行为产生显著的积极影响；Teng 等（2022）则进一步在 UTAUT 模型的基础上探究技术使用满意度与感知风险对于学生使用教育元宇宙平台的影响。由此，在使用 UTAUT 模型解释特定教育场景下的技术使用行为时，能够结合技术特点与学生情况对特殊的影响因素进行探究。

2.3 认知负荷的影响

认知负荷是指个体在工作记忆中处理信息时所消耗的认知资源总量，分为与学习材料难易程度关联的内在认知负荷、与学习材料呈现方式及所需活动相关的外在认知负荷、有助于在认知资源有剩余时构建认知图式的有效认知负荷（庞维国，2011）。由于工作记忆容量的有限性，当任务所需的认知资源超过工作记忆容量时，会出现认知超载（车敬上，2019）。Sweller 等（1998）将认知负荷理论融入教学设计，提出有效的课堂应该降低外在、内在负荷和增加有效负荷。随着信息技术的发展，更多学者开始基于认知负荷理论探讨虚拟现实、智慧课堂等数字化技术融合的教学设计是否真的能够提升教学效果（尤洋等，2020；张慕华等，2024）。在实证研究中，有学者发现适度的认知负荷能够提升学生的自我效能感、绩效反馈，而较高的认知负荷则会对创造性思维产生负面影响（李丽，2024）。杨磊（2025）将这种现象归因于适度认知负荷可能促进主动思考，但过高的认知负荷会导致学生认知超载，从而降低学习效率，不利于思维的发展。基于此，本研究将认知负荷作为调节变量，以研究认知负荷学生 GenAI 使用行为对批判性思维影响路径的调节作用。

3. 研究方法

3.1 研究模型与假设

3.1.1 基本模型

本研究基于 UTAUT 模型对学生使用 GenAI 使用意愿以及使用行为提出假设，以解释学生 GenAI 使用行为：

H1：绩效期望对学生 GenAI 使用意愿产生显著正向影响。

H2：努力期望对学生 GenAI 使用意愿产生显著正向影响。

H3：社群影响对学生 GenAI 使用意愿产生显著正向影响。

H4：使用意愿对学生 GenAI 使用行为产生显著正向影响。

GenAI 的使用可能会给高阶能力的培养带来一定的挑战（段世飞 & 钱跳跳，2024）。但是也有实证研究证明了 GenAI 的使用或技术应用对批判性思维具有显著正向影响（Guo & Lee, 2023；Suriano et al., 2025）。基于此，本研究将探讨学生在学习时使用 GenAI 对其批判性思维的影响，并提出以下假设：

H5：GenAI 使用行为对学生批判性思维产生显著正向影响。

3.1.2 认知负荷的调节作用

已有研究表明，学习者在数字化和网络化环境中学习可能会产生过高的认知负荷，从而导致学习效率、学习韧性以及学习投入受到负面影响，不利于思维的发展（满淑洁等，2023）。因此，本研究假设认知负荷对 GenAI 使用行为影响批判性思维发展具有负向调节作用：

H6：学生使用 GenAI 产生的认知负荷能够负向调节使用行为对批判性思维的影响。

基于以上假设，本研究构建学生使用 GenAI 影响批判性思维模型，探究学生 GenAI 使用行为对其批判性思维的影响机制。

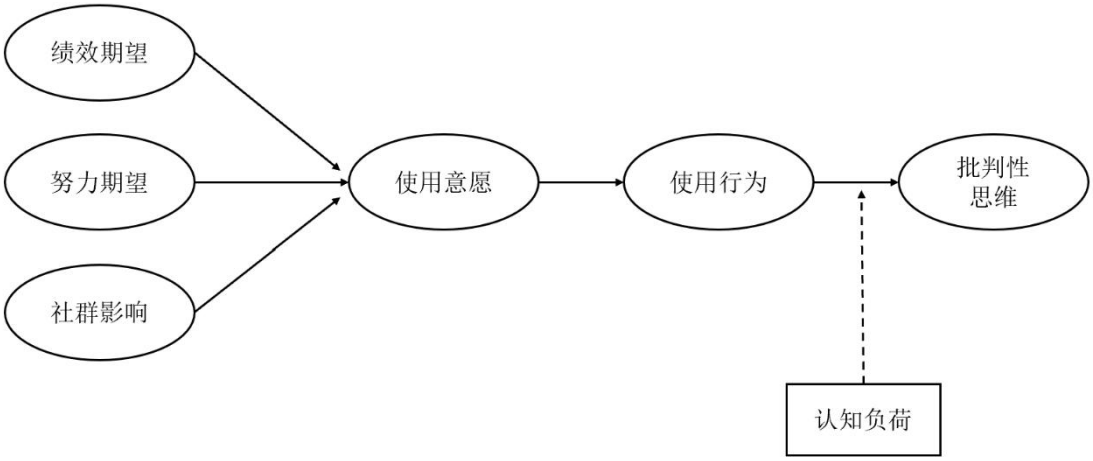


图 5 学生使用 GenAI 影响批判性思维模型

3.2 研究对象

本研究选取 S 市某高校学生作为研究对象，通过问卷星发放并回收问卷共 230 份，剔除作答时间小于 150 秒、量表填写答案一致性过高的问卷后，共获得有效问卷 209 份，有效率为 90.87%。其中男生 114 名，女生 95 名。按照被试的年级从大一到大四分别占比为 10.86%、36.00%、26.29%、26.86%。

3.3 研究工具

3.3.1 大学生 GenAI 使用行为调查问卷

本研究基于 UTAUT 模型改编《大学生 GenAI 使用行为调查问卷》（Venkatesh et al., 2003）。调查问卷包括学生使用 GenAI 的绩效期望、努力期望、社群影响、使用意愿和使用行为五个维度，每个维度下设 3 个题项。问卷采用李克特 5 点量表（1=完全不同意，5=完全同意）。问卷各维度 Cronbach's α分别为 0.97、0.92、0.88、0.95 和 0.90，内部一致性良好，问卷 KMO 值为 0.954（>0.6）通过了显著性水平为 0.05 的 Bartlett 球形检验。本研究采用结构方程模型进行验证性因子分析，其中χ^2/df 为 1.754，RMSEA 为 0.060，CFI 为 0.986，IFI 为 0.986，TLI 为 0.981，表明结构效度整体达标。

3.3.2 批判性思维自我评价问卷

本研究对 Facione, P（1990）提出的批判性思维自我评价量表进行改编，问卷包括理解、分析、评价、推论、表达和自我调节六个维度，每个维度下设 4 个题项。问卷采用李克特 5 点量表（1=完全不同意，5=完全同意）。问卷各维度 Cronbach's α分别为 0.96、0.96、0.96、0.95、0.96 和 0.97，内部一致性良好，问卷 KMO 值为 0.871（>0.6），通过了显著性水平为 0.05 的 Bartlett 球形检验。进一步验证性因子分析结果显示，χ^2/df 为 1.965，RMSEA 为 0.068，CFI 为 0.972，IFI 为 0.973，TLI 为 0.968，表明结构效度整体达标。

3.3.3 GenAI 使用认知负荷调查问卷

本研究对 Hart（2006）认知负荷调查问卷进行改编，问卷包括心智需求、时间需求、任务复杂度、付出程度、挫折程度、认知增强六个维度，每个维度下设 3 个题项。问卷采用李克特 5 点量表（1=完全不同意，5=完全同意）。问卷各维度 Cronbach's α分别为 0.91、0.86、0.87、0.84、0.82 和 0.95，内部一致性良好，问卷 KMO 值为 0.951（>0.6），通过了显著性水平为 0.05 的 Bartlett 球形检验。进一步验证性因子分析结果显示，χ^2/df 为 2.663，RMSEA 为 0.089，CFI 为 0.943，IFI 为 0.943，TLI 为 0.927，表明结构效度基本整体达标。

3.4 数据分析

本研究使用 SPSS 26 和 AMOS 26 对数据进行处理和分析。首先采用克隆巴赫系数和验证性因子分析检验测量工具的可靠性和有效性。在建立模型假设和调节作用检验时，绩效期望、努力期望、社群影响、使用意愿、使用行为以及批判性思维作为潜变量建模。进一步，本研究采用多群组结构方程模型探究认知负荷在模型中的调节作用（温忠麟等，2003）。根据该方法，本研究就认知负荷问卷得分均值，分别将被试分为高于均值的高认知负荷组（128 人），以及低于均值的低认知负荷组（81 人），通过对比两个认知负荷水平组的结构方程模型差异，验证认知负荷的调节作用。

4. 研究结果

4.1 大学生使用 GenAI 对批判性思维的影响

本研究采用结构方程模型检验假设模型，模型χ^2/df 为 2.401，RMSEA 为 0.082，CFI 为 0.965，IFI 为 0.965，TLI 为 0.959，表明模型拟合度达标，模型结果可信。根据模型系数检验结果（表 1），假设 H1-H5 成立，即被试对于使用 GenAI 的绩效期望、努力期望以及社群影响会正向影响其 GenAI 使用意愿；被试 GenAI 使用意愿会正向影响 GenAI 使用行为，进而影响其批判性思维。

表 2 UTAUT 模型与使用行为影响批判性思维路径检验

假设	路径	非标准化系数	标准化系数	S.E.	C.R.	*p*
H1	绩效期望➞使用意愿	0.749	0.718	0.093	8.028	***
H2	努力期望➞使用意愿	0.260	0.189	0.114	2.281	*
H3	社群影响➞使用意愿	0.130	0.104	0.047	2.754	**
H4	使用意愿➞使用行为	0.904	0.973	0.050	17.961	***
H5	使用行为➞批判性思维	0.869	0.961	0.048	18.125	***

注：*表示 p<0.05；**表示 p<0.01；***表示 p<0.001。

4.2 认知负荷的调节作用

本研究进一步采用多群组结构方程分析法探究认知负荷对使用行为影响批判性思维的调节作用。从模型适配标准来看，模型χ^2/df 为 1.868，RMSEA 为 0.046，CFI、IFI、 TLI 指标均大于 0.9。表明多群组结构方程模型和观察数据较为契合，模型拟合度较好。

两组模型的对比结果表明，高低认知负荷组模型中 H5 均成立，说明无论被试认知负荷高低，被试 GenAI 使用行为均能够显著正向影响其批判性思维。由表 2 数据进一步可知，在高认知负荷模型中，使用行为对批判性思维的影响路径显著（p<0.001），标准化路径系数为 0.685；在低认知负荷组，使用行为对批判性思维的影响路径显著（p<0.001），标准化路径系数为 0.982。进一步路径比较结果显示，高认知负荷组和低认知负荷组结构方程模型存在显著差异（p<0.01）,表明认知负荷对这条路径有负向调节作用，假设 H6 成立。

表 3 高低认知负荷组模型假设对比

假设	高认知负荷组					低认知负荷组					路径比较
	非标准化系数	标准化系数	S.E.	C.R.	*p*	非标准化系数	标准化系数	S.E.	C.R.	*p*	*p*
H5	0.472	0.685	0.074	6.369	***	0.873	0.982	0.058	15.082	***	**

注：*表示 p<0.05；**表示 p<0.01；***表示 p<0.001。

5. 讨论与建议

本研究发现，S 市某高校学生使用 GenAI 能够促进其批判性思维的发展，这与戚佳等（2024）的研究结果一致，说明 GenAI 的使用对于高校学生的学习、科研具有积极意义。同时研究发现这一过程受到认知负荷的负向调节，补充了杨磊等（2025）、满淑洁等（2023）的研究成果。为了推动 GenAI 在高校中的有效应用，本研究基于实证结果提出以下建议。

5.1 鼓励学生 GenAI 使用，培养良好使用习惯

实证研究结果表明，绩效期望、努力期望以及社群影响能够促进学生 GenAI 使用意愿。这说明教师在课堂上有效融入 GenAI 并给学生带来良好的使用体验能够有效提升学生的 GenAI 使用意愿，进而能够让学生有更多的机会在与 GenAI 协同学习的过程中得到个性化的高阶思维训练。结合社会认知理论（Social Cognitive Theory），学生对于技术的使用意愿除了来自于内部动机，社会环境导致的外在动机影响也十分重要（Schunk et al., 2014）。其中教师的鼓励与示范作用对于学生技术使用过程中自我效能感的激发与行为规范有重要作用（Schunk et al., 2020）。在智能时代下，GenAI 赋能学习已成必然趋势，GenAI 将会参与到学生的认知、行为能力和情感态度发展等各个方面当中（胡钦太等，2025）。推动学生熟练、规范使用 GenAI，并与其进行有效协同是十分必要的。因此，**教师应该积极将 GenAI 投入教学并做出良好行为示范**，鼓励学生正确有效地使用 GenAI，发挥 GenAI 赋能教育教学和人才培养的最大效能。

5.2 迭代教育智能体设计，降低学生认知负荷

实证研究结果表明，认知负荷会负向调节 GenAI 使用行为对批判性思维的影响。根据认知瓶颈理论（the Cognitive Bottleneck in Multitasking Theory），当个体的认知注意在执行多项任务时，其认知控制和信息处理能力会受到一定的干扰，并且会对个体的记忆力和思维能力造成一定的影响（Borst et al., 2010）。GenAI 因为其强大的算法和迁移能力，能够针对某一问题进行详尽而充实的分析（Batista et al., 2024）。尽管这可以为学生解决问题提供多元视角，但是当学生想沿着某一角度深入探索时，过多的干扰信息会使学生的注意力难以沿着本身的思维链条继续进行（Sweller, 2010）。GenAI 还可能会出现幻觉，包括忠实性幻觉与事实性幻觉（Huang et al.，2025）。忠实性幻觉可能会让学生反复修改、调整提示词以获得合理的输出结果，进而增加了学生使用 GenAI 解决任务的复杂度，同时可能会让学生产生使用 GenAI 的挫折感，这些都会加深学生的认知负荷。事实性幻觉则会在一定程度上对学生造成困惑或者传递错误的知识（刘泽垣等，2025）。不好的使用体验可能会导致学生在使用 GenAI 时产生挫折感或者焦虑情绪，进而影响学习的专注状态，有碍于批判性思维的发展（葛子豪等，2025；Csikszentmihalyi & Csikzentmihaly, 1990）。因此在 GenAI 赋能的课堂中，要进一步关注其作为工具的易用性和有效性。建议一线教师依据特定的教学目标和学习

内容设计基于 GenAI 技术的教育智能体。根据学生的具体需要和学习特点进行局部调整（涂明辉 & 谢丹，2024）。一方面，通过设定一定的对话内容和对话方式能够提高教育智能体的输出质量，降低学生的认知负荷；另一方面，专有资料库的搭建能够使教育智能体的输出内容在一定程度上减少幻觉的出现。

6. 总结

本研究以 UTAUT 模型为基础，在解释 GenAI 技术使用行为的基础上，探究了某高校学生 GenAI 使用行为对其批判性思维的影响。结果表明，绩效期望、努力期望和社群影响能够正向影响学生 GenAI 使用意愿，进而推动其使用行为的发生，促进其批判性思维的发展。同时，本研究还探讨了认知负荷在模型中的调节作用。结果发现，认知负荷会负向调节使用行为对批判性思维的影响。因此，鼓励教师在课堂中积极、正确地使用 GenAI，以提高学生使用意愿和使用行为，进而促进学生批判性思维发展；建议教师可以设计基于实际教学需要的教育智能体，减少 GenAI 的幻觉，降低学生认知负荷，进而促进学生批判性思维的发展。

参考文献

Almaiah, M. A., Alamri, M. M., & Al-Rahmi, W. (2019). Applying the UTAUT model to explain the students' acceptance of mobile learning system in higher education. Ieee Access, 7, 174673-174686.

Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. Information, 15(11), 676.

Borst, J. P., Taatgen, N. A., & Van Rijn, H. (2010). The problem state: a cognitive bottleneck in multitasking. Journal of Experimental Psychology: Learning, memory, and cognition, 36(2), 363.

Csikszentmihalyi, M., & Csikzentmihaly, M. (1990). Flow: The psychology of optimal experience (Vol. 1990, p. 1). New York: Harper & Row.

Facione, P. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

Guo, Y., & Lee, D. (2023). Leveraging chatgpt for enhancing critical thinking skills. Journal of Chemical Education, 100(12), 4876-4883.

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage publications.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., ... & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. ACM Transactions on Information Systems, 43(2), 1-55.

Miao, F., & Shiohira, K. (2024). AI competency framework for students. UNESCO Publishing.

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. Contemporary educational psychology, 60, 101832.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2014). Motivation in education: Theory, research, and applications. (No Title).

Suriano, R. Plebe, A. Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT canpromote complex critical thinking skills. Learning and Instruction, 95, 102011.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational psychology review, 22, 123-138.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. Educational psychology review, 10, 251-296.

Teng, Z., Cai, Y., Gao, Y., Zhang, X., & Li, X. (2022). Factors affecting learners' adoption of an educational metaverse platform: An empirical study based on an extended UTAUT model. Mobile Information Systems, 2022(1), 5479215.

UNESCO. (1998). World Declaration on Higher Education for the Twenty-First Century: Vision and Action. World Conference on Higher Education.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

马利红, 魏锐 & 刘坚等. (2020). 审辨思维: 21 世纪核心素养 5C 模型之二. 华东师范大学学报(教育科学版), 38(02), 45-56. doi:10.16382/j.cnki.1000-5560.2020.02.005.

王学男 & 李永智. (2024). 人工智能与教育变革.电化教育研究, 45(08), 13-21. doi:10.13811/j.cnki.eer. 2024.08.002.

尤洋, 王以宁 & 张海. (2020). 智慧课堂环境下教学视频复杂度与学习者认知负荷关系研究.现代远距离教育,(02), 91-96. doi:10.13927/j.cnki.yuan.2020.0023.

车敬上, 孙海龙, 肖晨洁 & 李爱梅. (2019). 为什么信息超载损害决策?基于有限认知资源的解释. 心理科学进展, 27(10), 1758-1768.

刘泽垣, 王鹏江, 宋晓斌, 张欣 & 江奔奔. (2025). 大语言模型的幻觉问题研究综述. 软件学报, 36(03), 1152-1185. doi:10.13328/j.cnki.jos.007242.

闫寒冰, 杨淑婷, 余淑珍 & 陈怡. (2025). 生成式人工智能赋能沉浸式学习：机理、模式与应用. 电化教育研究, 46(02), 64-71. doi:10.13811/j.cnki.eer.2025.02.009.

孙立会 & 周亮. (2025). 生成式人工智能素养：概念演变、框架构建与提升路径. 现代远距离教育, 1-18. doi:10.13927/j.cnki.yuan.20250319.001.

李丽. (2024). 自我效能、绩效反馈与创造性思维认知负荷的关系.高教学刊, 10(24), 101-104. doi:10.19980/j.CN23-1593/G4.2024.24.023.

李曼丽, 乔伟峰 & 李睿淼. (2025). 大语言模型工具能促进高校学生的高阶思维能力发展吗？——基于 12 所双一流大学学生问卷调查的实证分析. 现代教育技术, (01), 34-43.

杨磊, 司国东, 王春桃, 黄沛杰, 肖克辉 & 杜治国. (2025). 基于认知负荷理论的计算机组成原理实验教学设计与应用. 实验室研究与探索, 44(01), 177-181. doi:10.19927/j.cnki.syyt.2025.01.033.

余胜泉 & 汪凡淙. (2023). 人工智能教育应用的认知外包陷阱及其跨越. 电化教育研究, 44(12), 5-13. doi:10.13811/j.cnki.eer.2023.12.001.

张慕华, 祁彬斌 & 黄志南. (2024). 沉浸式虚拟现实何以赋能科学实验学习——临场感与认知负荷的平衡. 远程教育杂志, 42(05), 53-64. doi:10.15881/j.cnki.cn33-1304/g4.2024.05.006.

罗宇晨 & 郑燕林. (2024). 混合式协作学习促进批判性思维发展实证研究. 开放教育研究, 30(02), 109-119. doi:10.13966/j.cnki.kfjyyj.2024.02.012.

庞维国. (2011). 认知负荷理论及其教学涵义. 当代教育科学, (12), 23-28.

胡钦太, 梁心贤, 刘颜帆 & 王姝莉. (2025). 生成式人工智能如何影响学生发展——基于 31 项实验与准实验研究的元分析. 现代远程教育研究, 37(02), 83-91.

段世飞 & 钱跳跳. (2024). ChatGPT 浪潮下的高阶能力培养: 可为、难为、何为?. 河北师范大学学报(教育科学版), 26(04), 80-88. doi:10.13763/j.cnki.jhebnu.ese.2024.04.011.
祝智庭, 戴岭 & 胡姣. (2023). 高意识生成式学习: AIGC 技术赋能的学习范式创新. 电化教育研究, 44(06), 5-14. doi:10.13811/j.cnki.eer.2023.06.001.
涂明辉 & 谢丹. (2024). 生成式人工智能教育应用的反智隐忧及应对. 教育学术月刊, (11), 88-94. doi:10.16477/j.cnki.issn1674-2311.2024.11.002.
黄荣怀, 刘嘉豪, 潘静文, 刘梦彧 & 张国良. (2025). 面向智能时代的教育系统性变革：数字化赋能教育综合改革. 电化教育研究, 46(04), 5-12. doi:10.13811/j.cnki.eer.2025.04.001.
戚佳, 徐艳茹, 刘继安 & 薛凯. (2024). 生成式人工智能工具使用对高校学生批判性思维与自主学习能力的影响. 电化教育研究, 45(12), 67-74. doi:10.13811/j.cnki.eer.2024.12.009.
葛子豪, 卢柯全 & 李刚. (2025). 信息资源管理学科学生的生成式人工智能焦虑与影响因素研究. 图书情报工作, 69(06), 72-84. doi:10.13266/j.issn.0252-3116.2025.06.006.
曾明星, 廖柏林 & 覃遵跃. (2024). 从 ChatGPT 到 Sora: 生成式人工智能如何重塑深度学习场景. 远程教育杂志, 42(06), 11-23. doi:10.15881/j.cnki.cn33-1304/g4.2024.06.002.
温忠麟, 侯杰泰 & 马什赫伯特(2003). 潜变量交互效应分析方法. 心理科学进展, (5), 593-599.
满淑洁, 姜强, 李琪 & 赵蔚. (2023). CSCL 如何影响大学生数字化学习韧性——认知负荷的中介效应分析. 现代远距离教育, (01), 71-80. doi:10.13927/j.cnki.yuan.20230317.002.
廖彦霖. (2023). 批判性思维视域下的竞技辩论：迷思与澄清.湖南科技大学学报(社会科学版), 26(01), 47-53.doi:10.13582/j.cnki.1672-7835.2023.01.007.

联通主义学习群体协同知识创生过程中的概念网络演化规律*

A Study on Evolutionary Law of Conceptual Networks in the Collaborative Knowledge Creation Process of Connectivist Learning Communities

王志军[1*]，余新宇[1]，吴芝健[1]
[1]江南大学江苏“互联网+教育”研究基地
*jnuwzj@jiangnan.edu.cn

【摘要】 联通主义学习理论的提出回应了面向智能时代的学习规律和人才培养规律。基于大规模交互的群体协同知识创生是联通主义学习持续开展的关键，而概念网络是个体与群体认知最直接的表现形式。本研究旨在依托 cMOOC 中学习者交互数据，探究联通主义学习中概念网络的发展演化规律，从而深化对联通主义学习规律及其智能教育场景应用的理解。研究发现，“概念-人-技术”网络协同形成了知识创生“生态”，其中技术网络搭建了学习环境；社会网络作为“桥梁”为知识创生提供了情境。知识创生以问题与观点的非共识为起点，知识创生成果既是终点又是起点。

【关键词】 联通主义；cMOOCs；知识创生；概念网络；社会网络分析

Abstract: *The proposal of Connectivism learning theory responds to the learning patterns and talent cultivation principles in the intelligent era. Collaborative knowledge creation through large-scale interaction among groups constitutes the key to sustaining Connectivist learning, with conceptual networks serving as the most direct manifestation of individual and collective cognition. This study aims to explore the evolutionary patterns of conceptual networks in Connectivist learning by analyzing learner interaction data from cMOOCs, thereby deepening our understanding of Connectivist learning principles and their applications in intelligent educational scenarios. The research reveals that the synergistic interaction of "concept-person-technology" networks forms an knowledge creation ecosystem, where technological networks establish the learning environment while social networks act as a "bridge" providing contextual support for knowledge creation. Knowledge creation initiates from non-consensus perspectives and problem awareness, with its outcomes paradoxically serving both as the culmination and genesis of new learning cycles.*

Keywords: connectivism, cMOOCs, knowledge creation, conceptual network, social network analysis

1. 引言

在智能时代，技术赋能知识深度建构与共享的前提在于准确把握知识生长与创新的本质规律。联通主义作为数字时代的学习理论（Siemens，2005），突破了传统学习理论以简单、线性化的知识传递和知识建构为核心的局限，强调在开放复杂信息网络环境下的群体协同知识创新。联通主义学习认为知识存在于网络中，是内部认知神经网络、概念网络和社会网络

* 基金项目：国家社会科学基金全国教育科学“十三五”规划 2020 年度国家级一般课题“联通主义学习中群体协同知识创新研究”（课题编号：BCA200092）。

之间建立连接的过程（Wang & Chen, 2014），其中，知识创生是群体和个体的认知及其表征出来的概念网络不断丰富和发展的过程，对概念建立过程的分析实质上是对知识创生的过程分析（王志军等, 2019），揭示联通主义学习中概念网络的发展演化过程与规律将有助于加深对这一理论的认识、加快联通主义学习在智能教育场景中的实践落地。cMOOC 作为最具代表性的联通主义实践形式，学习者在平台中生成的大量交互数据，为我们分析挖掘联通主义学习情境下的学习过程与知识创生机制提供支持。

因此，本研究围绕“联通主义学习群体协同知识创生过程中，概念网络呈现怎样的发展演化过程？”展开，并将该核心问题分解为以下两个子问题：（1）联通主义学习群体协同知识创生中，概念网络如何表征？（2）概念网络的演化发展过程如何？具有怎样的特点？

相较其他学习理论，联通主义学习关注来自不同背景的学习者在开放、分布式学习环境中经历的网络化和经验共享，因此，概念之间连接的建立及新概念的生成过程是本研究的分析重点，即使用节点表示交互内容，连线表示节点之间的语义关系，概念网络具体表现为学习者的观点认知地图，呈现出知识非结构化、集体知识贡献、动态生成的特征。

2. 文献综述

联通主义学习以网络的建构和知识的持续创新为学习目标，提出网络中的知识具有分布式的特点，学习即连接的建立和网络的形成（Downes，2007）。学习者不仅需要与已有节点之间建立连接，还需在学习过程中创造新的节点，并与新的节点建立连接，连接建立的过程也是学习的过程，因此，连接的建立并非是对已有知识的重复，而是在理解的基础上进行创造，生成新的人工制品。因此，结合联通主义创始人之一 Stephen Downes 所提出的生长类知识生模式，可将知识创生定义为**在对碎片化知识形成连贯理解的基础上，为进一步促进联通而创造新的学习制品的知识生产过程**。

cMOOCs 是指联通主义学习理论指导下的大规模在线开放式课程，最早源于联通主义的创始人乔治·西蒙斯和斯蒂芬·唐斯从在 2008 年开设的《联通主义和连接化知识》。之后，各类大规模开放在线课程（cMOOCs）陆续出现，引发了国际上的 MOOCs（中文翻译为“慕课”）海啸。cMOOCs 中大规模的用户参与，生成了大量的学习参与和互动数据，为探究学习过程认识联通主义学习的本质与规律提供了机会。当前国内外研究主要从以下五个方面展开：1）从教学交互视角出发，探究 cMOOC 学习中教学交互开展情况，了解联通主义学习中教学交互的影响因素、作用过程与机制（谢雷等，2020；Wang et al.，2018）；2）从社区形成视角出发，探究影响 cMOOC 课程中信息流发展和社区形成的影响因素，以及教师在 cMOOC 课程中扮演的角色（Bozkurt et al.，2018）；3）从复杂网络的视角出发，探究 cMOOC 学习中的社会网络规律与特征（郭玉娟等，2020）、社会网络对学习者内部认知发展的影响（王慧敏等，2019）、学习者个体社会网络地位与概念网络特征的关系（徐亚倩等，2019）；4）从话语分析视角出发，结合社会网络分析法，探究学习者的话语模式对其社会资本（中心地位）的影响（Joksimović et al.，2018）、不同平台中学习者讨论主题的区别以及课程促进者的推荐对内容创建和讨论主题的影响（Joksimović et al.，2015）；5）从学习者的视角出发，探究 cMOOC 的参与模式与学习过程（吴斓等，2019），了解学习者的 cMOOC 学习体验，揭示学习者参与课程学习的动机及其兴趣所在（Mackness et al.，2015；Bali et al.，2015）。

综上所述，cMOOCs 为本研究提供了合适的研究情境，虽然国内外相关研究者基于 cMOOC 课程数据从多维度对联通主义学习的过程、规律展开探究，但当前研究大都偏重社

会网络，侧重于社会网络发展规律与特征的研究，对学习者话语分析的研究也主要关注与讨论主题的提取及主题下交互数量的统计分析。对联通主义学习中概念网络发展演化过程、知识创生过程及其影响因素的探究相对较少，亟需进一步的分析探究。

3. 研究设计与实施

3.1. 研究对象

Change MOOC 是联通主义创始人开设的课程，是一门充分体现联通主义的典型的 cMOOC 课程。课程自 2011 年 9 月开始，持续了 36 周，吸引了众多专家与来自全世界 11 个国家的学习者参与。每一周都有一个与现在、未来学习相关的学习主题，学习者可以根据自己的需要，在网络中选择不同的社交媒体或工具参与课程学习，与其他学习者展开深入的交流互动，带有#change 11 标签的交互内容会被汇聚到每天的课程日报中。本研究对 Change MOOC 导学周中的交互数据进行了分析。

3.2. 数据来源与样本选择

Change MOOC 课程充分体现联通主义学习开放联通的思想，允许学习者使用自己熟悉的工具和技术来建构个人学习空间、参与学习过程。因此，一共有四个主要的数据来源：1）课程学习周报；2）学习者和课程促进者的个人站点；3）分布在 Facebook 上的学习小组或群组；4）X（原 Twitter）数据。

课程日报中包含了绝大多数交互数据，在课程进行过程中，每天的周报都会发布在课程网站和学习者的邮箱中，学习者通过课程日报了解课程动态，所以本研究的数据收集以课程日报数据为主。同时，日报只是对当日课程发生交互的集合概括，具体的交互数据还需要基于课程日报、对平台中的具体交互做进一步追踪。故以课程日报为起点，对学习者在博客、课程交互核心群组 Facebook 以及社会平台 Twitter 的交互数据进行收集，并进行了转录与关联处理。

3.3. 概念网络建模与分析

本研究的概念网络分析是结合内容分析与网络分析，将文本类型的学习制品转化为概念节点（可能具有不同类型）的网络文本分析（Hoppe，2017）。具体而言，以达成知识创生的博客为数据追踪源节点，定位博客发布者在整体社会网络中的位置，提取与博客发布者关联的学习者，从而进一步提取学习者及其关联学习者发布的交互数据，向前、向后追踪概念的生成与演化发展过程，构建概念网络，如图 1 所示。

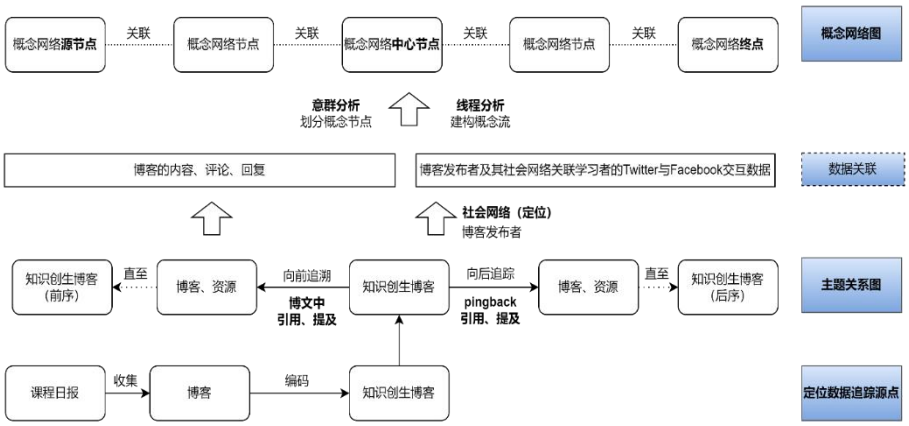


图 6 概念网络追踪规则

在概念网络中，节点之间建立联系的判断依据如下：1）若 A 节点为 B 博客中的子观点，则建立 B 指向 A 的链接；2）若 A 节点表述内容为 B 节点表述内容的延续，两节点正在讨论同一个共享问题，则建立 B 指向 A 的连接，常见于博客下针对博客子观点的评论和回复中；3）若 A 节点表述的内容中引用了 C 节点中表述的内容，则建立 C 指向 A 的链接，其中 C 节点可能为博客中的某一观点；4）若 A 节点表述的内容中推荐了 D 节点中的表述内容，则建立 A 指向 D 的链接，其中 D 节点可能为某一博客或资源。

为更清晰的表征概念网络及其发展过程，本研究中使用多模异质类网络来表征概念网络，概念网络中的节点与连线表征规则如下：1）形状：球形代表知识分享型博客；菱形代表经验分享型博客；空心圆形代表提问类型的观点；方形代表主题讨论节点；三角形代表资源，其中实心三角形代表文档资源，如论文、PPT，空心三角形代表网站资源。2）连线：箭头代表两节点之间的关系，A 节点指向 B 节点有三种情况，其中，“线点型”线段为 A 节点在交互内容中推荐 B 节点；“点型”线段为 A 节点的交互内容中引用了 B 节点；“实线型”代表了 B 节点为 A 节点的子节点或 A 节点导致了 B 节点的生成。3）颜色：从属于同一博客的节点为同一颜色。4）大小：节点的大小被分为四级，用以表示观点的从属关系。

在厘清概念网络节点与概念发展线程之后，使用 NodeXL 软件进行主题关系图与概念网络的绘制与可视化呈现。

4. 概念网络分析结果

4.1. 概念网络的发展与演化特征

在联通主义学习共同体中，概念网络的发展是协作知识创生最直接的表现形式。基于主题分析构建的 Change MOOC 导论周和主题学习周的概念网络如图 2 所示。

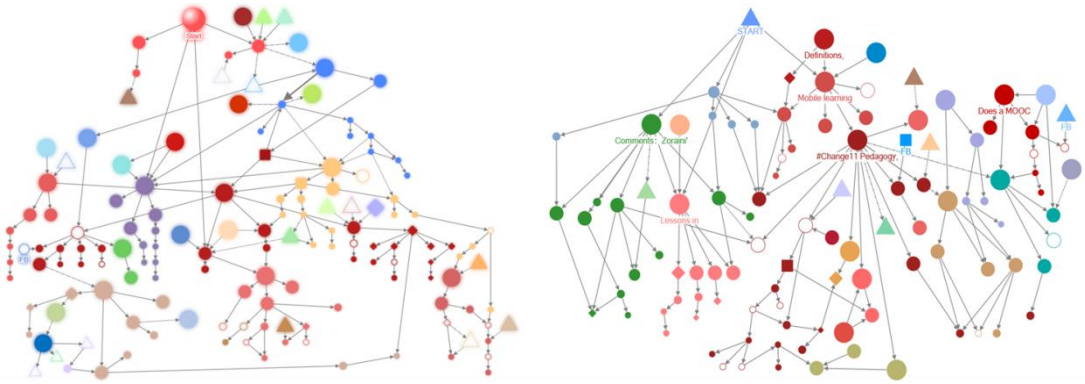


图 7 Change MOOC 第 1 周和第 2 周的概念网络

通过概念网络分析，归纳出节点之间的 10 种关系类型，包括引用、推荐、补充、扩展、引申、提问、反对、支持、汇聚和激发。学习者通过这 10 种方式参与概念网络的发展，推动知识创生过程。概念网络分析发现了四种主要的知识创生路径，即原创性成果产出类、主题讨论衍生类、开放问题讨论类和多样资源关联类。同时，知识创生过程体现出分布式、持续性、递归性、开放性和复杂性等特征。

4.2. 社会网络发展的演化特征

基于社会网络分析构建的 Change MOOC 导论周和主题学习周的整体社会网络如图 3 所示。社会网络中的节点越大，表示中心性指数越高，并采用层次聚类算法对社会网络中的不同群体进行分类。从基本属性、结构特征和关键个体三个层面分析社会网络的发展演化过程，研究发现社会网络的发展具有跨平台、多结构、去中心化、自组织以及群体协同的特征。

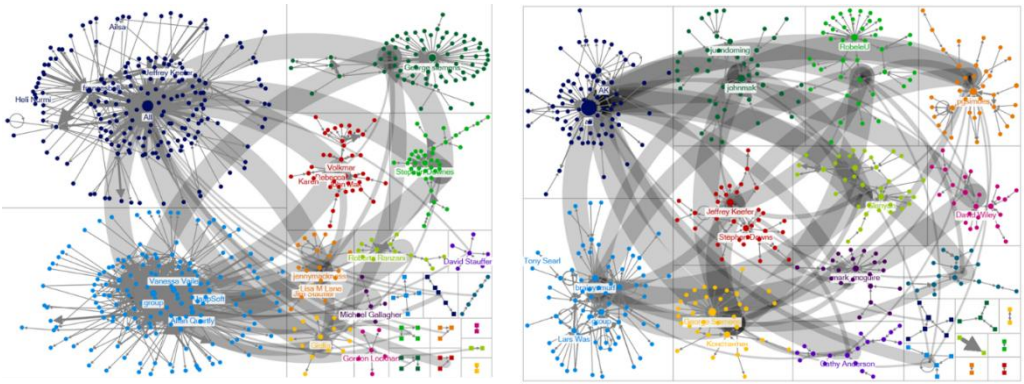


图 8 Change MOOC 第 1 周和第 2 周的社会网络

通过提取概念网络背后的社会网络（如图 4 所示），并与整体社会网络进行对比分析，同时基于概念网络发展的关键时间点对社会网络进行切片分析（如图 5 所示），研究发现社会网络与概念网络之间存在相互促进作用，具体表现如下：1)社会网络促进信息流动和多元知识汇聚；2)关键资源随着社会网络中心的变化而扩散；3)社会网络关系的形成持续作用于后续的知识创生过程；4)高质量内容的共享推动网络的扩展和优化。

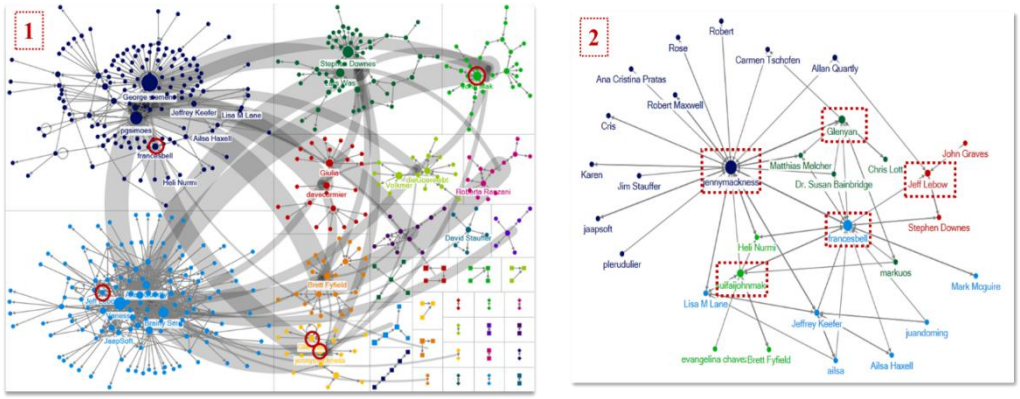


图 9 第一周概念网络（左）背后的社会网络（右）

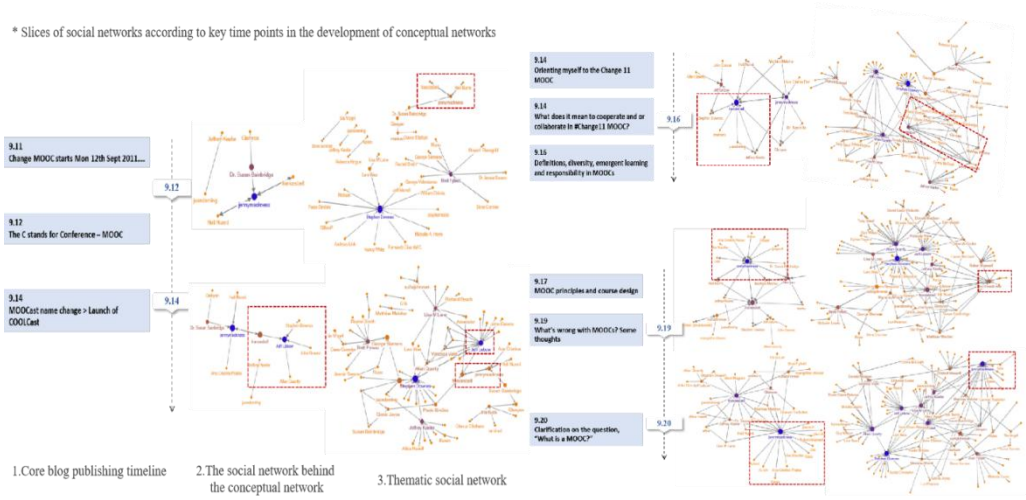


图 10 基于概念网络中关键节点的社会网络切片分析

4.3. 技术在支持社交网络和概念网络发展中的作用

通过提取概念网络背后的支持技术，发现工具在概念网络的发展和知识创生过程中发挥着重要作用。例如：智能抓取工具是技术网络的核心，它能够整合学习者在分布式平台上留

下的交互数据，是学习者进行路径探索互动和利用群体智慧的关键；资源聚合工具负责课程中相似学习资源的汇总，是信息筛选、处理和组织的重要工具；社交媒体工具有助于建立学习者的个人社会网络和群体社会网络，是课程中互动信息流动的关键媒介。

在学习者的互动过程中，不同的学习站点逐渐建立联系，形成一个开放、多元的技术网络。与此同时，技术网络、社会网络和概念网络相互作用：1）技术网络是社会网络发展的基础，社会网络的不断扩展以及多元化学习者的加入也反过来丰富了技术网络；2）课程中多样化技术工具的使用支持了不同类型的学习者互动和概念网络的发展，学习者在探索过程中不断发现有效的技术工具，进而拓展技术网络，以更好地呈现和表达深度学习成果。

4.4. 联通主义学习群体协同知识创生过程模型

基于上述数据分析结果，构建了联通主义学习中的群体协同知识创生过程模型，如图 6 所示。

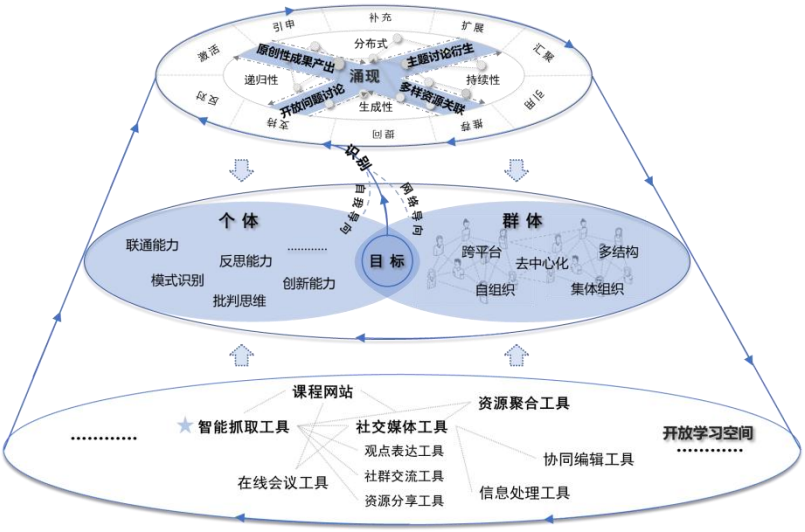


图 11 联通主义学习群体协同知识创生过程模型

联通主义学习群体协同知识创生过程本质上是一个多层网络，是技术、社会、概念三类网络相互作用、协同发展的过程。其中，开放、多样、持续扩展的技术网络构成学习者的学习空间；在这个开放学习空间中，学习者在个体与群体目标导向下，参与课程交互，形成个人独特的社会网络，多个个体及基于群体目标形成群体社会网络，并在持续的交互中增强社会网络的联通性；学习者在开放复杂的信息网络环境中，基于自身目标与群体目标，通过自我导向和网络导向学习，识别感兴趣学习主题与学习资源，并以多样化的方式参与群体互动交流过程与群体协同知识创生过程，推进概念网络发展与新知识的涌现。产出的高质量学习制品会吸引更多学习者参与课程交互、与学习制品发布者建立连接，扩展学习者个人社会网络和群体社会网络；知识创生过程中出现的新的、多样化技术工具，也进一步了扩展技术网络。 因此，联通主义学习群体协同知识创生过程中，技术、社会、概念三层网络之间呈现出下层网络为上层网络的发展奠定基础，上层网络的发展能进一步作用于下层网络的扩展与优化，三网络的相互影响与协同发展的特征。同时，由于知识创生的生成性特征，知识创生成果既是本阶段知识创生的终点，又是新一段知识创生路径的起点，因此，三网络协作作用下的联通主义学习群体协同知识创生过程总体呈现螺旋发展、循环迭代的特征。

5. 研究结论

（1）开放的技术支持：构建开放、多样、自主、交互的学习环境

联通主义学习中，学习者可以自主选择熟悉或感兴趣的技术搭建个人学习空间，学习者所使用的技术并非课程促进者开发的、相对封闭的技术空间，大都是开放网络中的社交媒体工具，例如 Twitter、Facebook 以及学习者的个人博客（少部分学习者的博客进行了隐私设置）。在开放的技术平台中，学习者可以自由的与网络内外的人交流，不受限制的分享、讨论。随着不同的技术平台中学习者的加入，学习者围绕技术平台形成不同社区，不同平台中汇聚着对课程感兴趣的学习者，大量的学习者为社区带来多样化的知识。而课程网站、学习日程、学习日报、学习活动的设计，保障课程处于自主而非“失控”状态，让其在高度组织的结构化课程与更加混乱的开放网络之间的中间地带，建构了一种开放的、不断变化的、分布式复杂学习环境来模拟新时代对人的学习的新要求。

（2）联通化的社会网络：知识连接的“桥梁”与知识创生的情境

在个人目标和群体目标的导向下，学习者主动探索资源并与其他学习者建立连接，学习者之间多样化连接的建立构建了课程社会网络。随着课程的进行，在自组织与群体组织的协同作用下，社会网络中的节点不断增多、联通性逐渐增强，呈现出多结构、多中心化的特点。跨平台社会网络连接的建立，联通了不同平台中的不同学习者和不同知识，此时，连接建立的过程不仅是学习者之间连接的建立，也是与学习者背后的信息与资源建立联系、不仅是与人的联通，也是与相关信息和资源的联通。社会网络的连接构成了信息流通的“管道”，随着课程进行，社会网络中学习者的中心化特点不断削弱、社会网络联通性的不断增强，使得信息流通效率增强、信息的传播范围扩大，整体社会网络中流通的信息与资源更加丰富、多样。

（3）持续发展的概念网络：知识创生以问题与观点的非共识为起点，知识创生成果既是终点又是起点

概念网络是联通主义学习群体协同知识创生最直接的表现形式，在联通主义学习中，学习者通过识别感兴趣的内容进入知识创生过程，参与讨论交流，提出个人见解，通过引用、推荐、补充、扩展、提问、引用、表示反对或支持、总结、激活等方式参与并推动概念网络的发展，随着越来越多的学习者参与，群智汇聚过程加快了知识更新速度与知识内涵的扩展，不断生成、变化、发展的新观点使得概念网络发展呈现出涌现性的特征。与此同时，对某一问题和观点的非共识是知识创生的“触发器”。而知识创生的成果作为课程的新资源与高质量学习内容，在其发布后将继续作用于课程的概念网络发展与后续的知识创生，知识创生的成果在此既是上一阶段知识创生的终点，又是新一段知识创生路径的起点。

（4）“技术-人-概念”三者联动的知识创生“生态”

在联通主义学习群体协同知识创生过程中，“技术”对应的学习环境是联通主义学习发生的重要媒介，“人”对应的社会网络是信息流通的“管道”，“概念”对应的概念网络是学习者内部的模式识别、观点形成以及知识创新的显性化表征。联通主义学习是连接建立与网络形成的过程，三者之间的协同关系表现为相互影响、动态变化、相互支持的有机协同作用。“技术”的支持是社会网络与概念网络发展的前提；社会网络的扩展在引入更多学习者的同时，也带来多样化的观点和新的技术工具，促进概念网络与技术网络的发展；概念网络发展过程中学习者在思考有效呈现学习制品的同时，引入更多技术工具，而学习者生成与呈现的高质量学习内容能吸引更多学习者加入课程学习、参与交互，帮助扩展社会网络与技术网络。同时，社会网络处于技术平台与概念网络之间的位置，联通不同平台中的学习者与不同平台中的信息资源，是概念网络节点产生与关联的前提，处于中间协调的位置，社会网络的联通性将直接影响概念网络的发展与知识创生的达成。

6. 研究局限

本研究的局限性主要体现在以下几个方面：（1）数据分析案例的局限：一方面，联通主义学习涉及多个技术平台，学习者可能同时活跃于不同平台，而逐一收集和分析这些数据在时间和技术层面上难以实现。另一方面，cMOOC 课程中达成知识创生案例有限。（2）学习者课外的知识创生数据无法追踪与记录，且学习者的隐性交互数据没有在平台留下记录，如阅读了某些资源或博客，但没有留下交互痕迹（评论、转发等）。未来研究中，可以基于具体的数据追踪规则，开发相关的算法与智能分析工具，实现对课程学习者在开放网络学习环境中分布式的交互数据的追踪与整理。此外，可以结合学习风格、知识背景、参与动机等维度增强对社会网络分析结果的解释力。

参考文献

郭玉娟,陈丽,许玲 & 高欣峰.(2020).联通主义学习中学习者社会网络特征研究.中国远程教育,(02),32-39+67+76-77.

王慧敏 & 陈丽.(2019).cMOOC 微信群社会网络特征及其对学习者认知发展的影响.中国远程教育,(11),15-23+92.

王志军,刘璐 & 杨阳.(2019).联通主义学习行为分析方法体系研究.开放教育研究,25(04),18-30.

吴斓 & 陈丽.(2019).基于扎根理论的 cMOOC 学习者学习过程的研究.中国远程教育,40(12),1-11+92.

谢雷 & 陈丽.(2020).cMOOC 学习者交互行为自评估分析研究.现代远距离教育,(05),62-67.

徐亚倩 & 陈丽.(2019).联通主义学习中个体网络地位与其概念网络特征的关系探究——基于 cMOOC 第 1 期课程部分交互内容的分析.中国远程教育,(10),9-19+51+92.

Bozkurt, A., & Keefer, J. (2018). Participatory learning culture and community formation in connectivist MOOCs. Interactive Learning Environments, 26(6), 776-788.

Bali, M., Crawford, M., Jessen, R., Signorelli, P., & Zamora, M. (2015). What makes a cMOOC community endure? Multiple participant perspectives from diverse cMOOCs. Educational Media International, 52(2), 100-115.

Downes, S. (2007). What connectivism is.[2007-02-03]. https://halfanhour.blogspot.com/2007/02/what-connectivism-is.html.

Hoppe H U. (2017). Computational methods for the analysis of learning and knowledge building communities. Handbook of learning analytics, 23-33.

Joksimović, S., Dowell, N., Poquet, O., Kovanović, V., Gašević, D., Dawson, S., & Graesser, A. C. (2018). Exploring development of social capital in a CMOOC through language and discourse. The Internet and Higher Education, 36, 54-64.

Joksimović, S., Kovanović, V., Jovanović, J., Zouaq, A., Gašević, D., & Hatala, M. (2015). What do cMOOC participants talk about in social media? A topic analysis of discourse in a cMOOC. In Proceedings of the fifth international conference on learning analytics and knowledge (pp. 156-165).

Mackness, J., & Bell, F. (2015). Rhizo14: A rhizomatic learning cMOOC in sunlight and in shade. Open Praxis, 7(1), 25-38.

Siemens, G. (2005).Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1), 3-10.

Wang, Z., Chen, L., & Anderson, T. (2014). A framework for interaction and cognitive engagement in connectivist learning contexts. International Review of Research in Open and Distributed Learning, 15(2), 121-141.

Wang, Z., Anderson, T., Chen, L. (2018). How Learners Participate in Connectivist Learning: An Analysis of the Interaction Trace1s From a cMOOC.The International Review of Research in Open and Distributed Learning, 19(1), 44-67.

数字化背景下职业教育教师教学能力现状及提升策略研究

唐小淇[1]，马欣妍[1]，刘敏燕[1]，丁旭[1]，黄蓉[1]，刁均峰[1]*
[1]海南师范大学 教育学院，海南海口 571158
*通讯作者:920268@hainnu.edu.cn

【摘要】数字化转型对职业教育教师的教学能力提出了新要求。本研究通过对全国 1024 名职业教育教师的调查，从课程开发、教学实施、专业理论、行业实践、信息素养和科研能力六个维度评估了教师教学能力现状。研究结果表明：（1）教师在教学实施和专业理论方面表现较好（均值 4.01），而课程开发和行业实践能力相对薄弱（均值 3.68）；（2）教师的学历、教龄、年龄和职称对其教学能力有显著影响，高学历、长教龄、高年龄和高职称教师普遍表现更优；（3）不同教师群体在数字教学能力上存在明显差距。基于研究发现，本文提出四项针对性策略：构建分层分类的培训体系，深化产教融合与校企合作，强化数字教学技术应用，以及完善教师评价与激励机制。

【关键词】教学能力；教师经验；专业成长；教育策略；职业教育

Abstract *Digital transformation has posed new requirements for the teaching competencies of vocational education teachers. This study assessed the current status of teachers' competencies across six dimensions—curriculum development, teaching implementation, professional knowledge, industry practice, digital literacy, and research capability—through a survey of 1,024 vocational education teachers across China. The findings reveal that: (1) teachers perform well in teaching implementation and professional knowledge (mean value 4.01), while showing relative weakness in curriculum development and industry practice (mean value 3.68); (2) teachers' educational background, teaching experience, age, and professional title significantly influence their competencies, with teachers of higher qualifications, longer experience, older age, and higher titles generally demonstrating superior performance; (3) notable disparities exist in digital teaching capabilities among different teacher groups. Based on these findings, this paper proposes four targeted strategies: establishing a stratified training system, deepening industry-education integration and school-enterprise cooperation, strengthening digital teaching technology application, and improving teacher evaluation and incentive mechanisms.*

Keywords teaching competence; teacher experience; professional development; educational strategies; vocational education

1.引言

经过几轮技术革命和工业转型，人类社会已从农业和工业文明过渡到由人工智能和大数据驱动的数字化时代。这一转变深刻而显著地改变了学生的学习方法、教师的教学方式以及对教师能力的要求。在当今职业教育和职业院校日益数字化的改革浪潮中，强调教师的数字化教学能力已成为提高职业教育发展水平的根本要求。教师的专业发展是整体提升教学质量及教育改革成败的关键因素（唐智彬和石伟平，2009）。构建适应数字化转型的职业教育师资队伍是职业教育发展的根本环节；没有具备数字化教学能力的“双师型”教师团队，就不可能有高质量的职业教育或培养出社会和企业欢迎的技术技能人才（毛天平，2024）。

面对数字化、智能化的教育环境，教师队伍的建设仍面临诸多不足和紧迫问题，如“双师型”教师短缺、教师整体数字化教学技能不足以及实践教学能力薄弱（朱正浩等人，2024）。鉴于此，本研究旨在分析数字化背景下职业教育教师教学能力发展中存在的问题，并探索解决这些不足的有效发展策略，以提升职业教育教师的教学能力。具体而言，本研究将回答以下问题：（1）数字化背景下职业教育教师的教学能力现状如何？（2）不同年龄、教龄、学历、职称等的中国高职教师的数字素养水平是否存在显著性的差异？（3）如何有效提升职业教育教师在数字化背景下的教学能力？

2. 数字化背景下职业教育教师教学能力的要求

学者们对数字化背景下职业教育教师应具备的教学能力进行了深入探讨。从国际研究视角看，Brennan 等人提出批判性反思是职业教育教师实践教学能力的重要组成部分（Berglund,.et al，2020）；Kemmis 和 Green 强调职业教育教师需将行业知识和经验应用于教学，并指出教师的行业知识和个人经验是教学能力不可或缺的部分（Kemmis & Green，2013）。此外，Sevilla 和 Madero 指出企业实践经验对教学能力有显著影响，认为与普通教育教师相比，拥有劳动市场经验的职业教师更倾向于将教学作为第二职业，且教学能力是其主要决定因素（Sevilla & Madero，2023）。Suárez Guerrero 等人则强调数字素养对于职业教育教师至关重要，他们发现虽然教师具备在职业教育教学实践中应用某些数字资源和技术的知识和技能，但在将数字技术融入教学评估方面仍存在困难，且教师的数字素养和研究能力有待提升（Suárez Guerrero.et al, 2021）。

国内关于数字化背景下职业教育教师教学能力的研究也呈现鲜明特征。张妮等人提出，在智能教育时代，教师的数字技术教学能力主要涵盖数字意识与态度、数字知识与技能、数字化设计与构建、数字化实施与评价、数字环境下的反思与创新等方面（张妮等人，2022）。武马群认为数字化时代的教师教学能力涉及六个方面：基于数字化的教学理念、数字化课程组织与评估、数字教学方法的运用、数字技术与设备的使用、数字化课堂教学与管理以及数字环境下的学习与专业发展（武马群，2015）。薛栋和武加霞（2021）追踪了教师数字化教学能力十年的发展轨迹，重点关注教师设计和开发数字资源与工具的能力、持续将数字技术与传统教学融合的能力，以及数字环境下教师教学与团队协作能力的重要性（薛栋和武加霞，2021）。许红叶认为数字化背景下的教师教学能力基于其数字化实际知识，需要在特定的数字化背景下形成和发展，包括数字化教学的转移、整合、互动、评估与协作，核心在于促进学生的数字化学习能力（许红叶，2013）。王彩霞和詹青龙强调教师的数字化教学能力应包括数字技术知识与技能、教育技术理论知识、数字化教学设计能力、数字化软件设计与制作能力和数字化教学研究能力（王彩霞和詹青龙，2014）。

综合现有研究成果及职业教育教师教学能力标准，本研究将数字化背景下职业教育教师的教学能力分为六个维度：课程开发能力、课程教学能力、专业知识、行业能力、数字素养和研究开发能力。这六个维度全面涵盖了数字化背景下职业教育教师所需的核心能力。

3. 研究方法

本研究旨在分析数字化背景下职业教育教师教学能力发展中的现状，以提高职业教育教师的教学能力。研究团队在全国范围内发放了 1024 份自行开发的数字化转型背景下职业教育教师能力自评问卷。本问卷基于刁均峰等人在《为信息时代的可持续发展培养职业教育教师：职业教育教师教学能力量表的开发与应用》中提出的理论框架，设计了六个关键评估维度：

课程开发能力、教学实施能力、专业理论知识、行业实践经验、信息化应用素养及科学研究能力（Diao & Hu，2022）

3.1 调查样本统计

本研究采用“问卷星”在线问卷工具收集数据，共获得 1024 份有效问卷。参与调查的职业院校教师来自中国华东、华北、东北、华中和西北等多个地区，样本构成具有较好的地域代表性。从学校类型来看，高等职业教育教师 750 人（73.2%），中等职业教育教师 274 人（26.8%）。性别分布方面，女性教师 656 人（64.1%），其中高等教育 476 人、中等教育 180 人；男性教师 368 人（35.9%），分布于高等教育 274 人和中等教育 94 人。年龄结构呈现中青年教师为主的特点，31-40 岁教师最多，占比 38.7%（396 人）；41-50 岁教师次之，占比 30.1%（308 人）；51-60 岁教师占比 17.1%（175 人）；30 岁以下教师占比 14.1%（144 人）；61 岁及以上教师仅占 0.1%（1 人）。在教育行业工作年限方面，具有 11-20 年经验的教师群体最大（43.0%，440 人），其次是 0-5 年经验（21.6%，221 人）、21 年及以上经验（20.8%，213 人）和 6-10 年经验（14.6%，150 人）的教师。学历背景显示，硕士学位教师占比最高（50.4%，516 人），其次为学士学位（37.3%，382 人）和其他学历（10.4%，107 人），博士学位教师比例较低（1.9%，19 人）。

3.2 研究工具

本研究使用自行开发的“数字化转型背景下职业教育教师能力自评问卷”作为研究工具。该问卷基于刁均峰等人的框架，包含六个维度：课程开发能力、课程教学能力、专业知识、行业知识、信息素养和研究开发能力。问卷采用五点李克特量表，从“完全不同意”到“完全同意”分别赋值 1-5 分。问卷的 Cronbach's α系数为 0.873，各维度信度系数均≥0.70，表明问卷具有良好的内部一致性。在效度分析方面，初步检验显示 KMO 值为 0.923，巴特利特球形检验结果显著，证实数据适合进行因子分析。验证性因子分析结果表明，所有题项的因子载荷值分布在 0.605-0.825 之间，均高于临界值 0.6；平均方差萃取量(AVE)介于 0.623-0.812 之间，超过了 0.5 的建议标准；各维度的组合信度(CR)值均大于 0.80，证明问卷具有良好的聚合效度。综上所述，问卷结构合理，各测量维度及题项关系明确，信效度检验结果良好。

3.3 数据分析方法

本研究采用 SPSS 26.0 进行数据分析。首先进行描述性统计分析，计算各维度的均值、中位数和标准差，然后通过方差分析和 t 检验探究不同人口学特征（如年龄、性别、教龄、学历、职称等）的教师在各能力维度上的差异。

4. 研究结果

4.1 职业教师教学能力的总体情况

在教师能力的总体评分方面，本研究通过均值、中位数和标准差等统计指标进行描述性统计分析，评估教师对其能力的自评结果。如表 1 所示，课程教学能力和专业知识的均值为 4.01，中位数为 4.00，表明这两个维度的能力总体较高，大多数教师在这些领域表现良好。信息素养和研究开发能力的维度的均值和中位数略低于课程教学能力和专业知识，但仍在 3.76 到 3.85 之间，表明教师在这些领域也表现良好。课程开发能力和行业知识：这些维度的均值为 3.68，中位数略高于均值，但标准差接近 0.9，表明教师在这些领域的表现变异性较大。总体而言，教师对自身能力的自评显示，他们在课程教学和专业知识方面表现卓越，而在课程开发和行业知识方面还有改进空间。

表 1 职业教师教育能力情况

维度	均值	中位数	标准差	最小值	最大值
课程开发能力	3.68	3.79	0.89	1.00	5.00
课程教学能力	4.01	4.00	0.79	1.00	5.00
专业知识	4.01	4.00	0.80	1.00	5.00
行业知识	3.68	3.78	0.89	1.00	5.00
信息素养	3.85	4.00	0.82	1.00	5.00
研究开发能力	3.76	3.94	0.86	1.00	5.00

4.1.2 不同学历的职业教师的能力差异分析

学历层次是影响高职教师数字素养能力差异的重要因素。如表 2 所示，博士学历教师在所有六个能力维度上均表现卓越，中位数普遍在 4.2-4.5 之间。特别是在课程开发能力、课程教学能力及专业知识方面，博士群体中位数均达到 4.5（范围 3.0-5.0），远高于其他学历群体。硕士学历教师表现次之，各维度中位数在 3.7-4.2 之间，其中课程教学能力（4.2）和专业知识（4.2）相对较强。学士和其他学历教师在各维度上表现明显较弱，尤其是其他学历教师在研究开发能力上仅为 3.0。这种梯度分布表明，高等教育背景为教师提供了更丰富的知识储备和更系统的能力培养。统计检验结果（p=0.000）进一步确认了这些差异的显著性，揭示了学历对高职教师数字素养能力的深远影响。

表 2 不同学历的职业教师的能力差异分析

能力维度	博士(中位数)	硕士(中位数)	学士(中位数)	其他(中位数)	p 值
课程开发能力	4.5（3.0-5.0）	4.0（1.5-5.0）	3.8（1.3-5.0）	3.3(2.0-5.0)	0.000*
课程教学能力	4.5（3.0-5.0）	4.2（1.2-5.0）	4.0（1.7-5.0）	3.6(2.0-5.0)	0.000*
专业知识	4.5（3.0-5.0）	4.2（1.5-5.0）	4.0（1.3-5.0）	3.8(2.2-5.0)	0.000*
行业知识	4.2（3.0-5.0）	3.7（1.0-5.0）	3.5(1.0-5.0)	3.5(1.0-5.0)	0.000*
信息素养	4.2（3.0-5.0）	4.0（1.0-5.0）	3.7(1.0-5.0)	3.2(1.3-5.0)	0.000*
研究开发能力	4.2(3.0-5.0)	4.0(1.5-5.0)	3.5(1.0-5.0)	3.0(1.0-5.0)	0.000*

4.2.3 不同教龄组教师的能力差异分析

教龄因素对职业教师的教学能力水平具有显著影响。如表 3 所示，具有 21 年及以上教龄的教师在大多数维度上表现优异，特别是在课程教学能力（4.4）和专业知识（4.4）方面，展现了丰富职业经验的价值。有趣的是，11-20 年教龄组在某些方面表现尤为突出，如研究开发能力（4.1）和行业知识（4.1），这可能反映了这一群体处于职业生涯的黄金期，既有足够经验又保持较高的学习活力。6-10 年教龄组虽然整体表现不及资深教师，但在课程教学能力（4.0）和专业知识（4.0）维度上已达到相当水平，表明这一阶段教师已形成较成熟的教学体系。0-5 年教龄新教师则在各维度上普遍处于较低水平（中位数 3.4-4.0），特别是课程开发能力最弱（3.4），反映了专业成长的阶段性特征。这种分布格局为教师发展规律提供了实证依据，同时也暗示了针对不同教龄教师的差异化培养策略的必要性。

表 3 不同教龄组教师的能力差异分析

能力维度	0-5 年(中位数)	6-10 年(中位数)	11-20 年(中位数)	21 年及以上(中位数)	p 值

能力维度	0-5 年(中位数)	6-10 年(中位数)	11-20 年(中位数)	21 年及以上(中位数)	p 值
课程开发能力	3.4 (1.5-5.0)	3.5 (1.8-5.0)	4.0 (1.2-5.0)	4.1 (1.6-5.0)	0.000*
课程教学能力	3.9 (2.0-5.0)	4.0 (2.0-5.0)	4.3 (2.0-5.0)	4.4 (2.5-5.0)	0.000*
专业知识	4.0 (2.0-5.0)	4.0 (2.2-5.0)	4.1 (2.5-5.0)	4.4 (2.5-5.0)	0.000*
行业知识	3.7 (2.0-5.0)	3.8 (2.0-5.0)	4.1 (2.2-5.0)	4.0 (2.0-5.0)	0.000*
信息素养	3.7 (2.0-5.0)	3.8 (2.0-5.0)	4.0 (2.0-5.0)	4.2 (2.0-5.0)	0.000*
研究开发能力	3.6 (1.8-5.0)	3.7 (1.5-5.0)	4.1 (2.0-5.0)	4.0 (2.0-5.0)	0.000*

4.2.4 不同年龄组的职业教师的能力差异性分析

年龄因素对职业教师的教学能力水平具有显著影响。如表所示，56 岁及以上教师群体在各能力维度上表现最为突出，中位数在 4.0-4.5 之间，尤其在课程教学能力和专业知识两项上均达到 4.5，展现了资深教师的专业优势。46-55 岁年龄组紧随其后，各维度中位数在 4.0-4.3 之间，在课程教学能力（4.3）和专业知识（4.3）方面表现同样出色。36-45 岁中年教师各能力维度中位数在 3.7-4.1 之间，处于中等水平。35 岁及以下青年教师则在所有维度上表现相对较弱，中位数仅在 3.5-3.8 之间。这种随年龄递增的能力提升趋势，一方面可能源于长期职业实践积累的经验智慧，另一方面可能反映了高职教师群体积极适应数字化教育转型、持续更新专业能力的努力。

表 4 不同年龄组的职业教师的能力差异性分析

能力维度	35 岁及以下(中位数)	36-45 岁(中位数)	46-55 岁(中位数)	56 岁及以上(中位数)	p 值
课程开发能力	3.7 (2.0-5.0)	3.9 (2.0-5.0)	4.1 (2.2-5.0)	4.2 (2.2-5.0)	0.000*
课程教学能力	3.8 (1.8-5.0)	4.1 (1.5-5.0)	4.3 (2.0-5.0)	4.5 (2.5-5.0)	0.000*
专业知识	3.8 (2.0-5.0)	4.1 (2.0-5.0)	4.3 (2.5-5.0)	4.5 (2.5-5.0)	0.000*
行业知识	3.5 (2.0-5.0)	3.7 (2.0-5.0)	4.0 (2.0-5.0)	4.2 (2.5-5.0)	0.000*
信息素养	3.9 (2.0-5.0)	3.9 (2.0-5.0)	4.1 (2.0-5.0)	4.2 (2.0-5.0)	0.000*
研究开发能力	3.5 (1.5-5.0)	3.8 (1.5-5.0)	4.0 (1.8-5.0)	4.0 (2.0-5.0)	0.000*

4.4.5 不同职称的职业教师的能力差异性分析

职称因素对职业教师的教学能力水平具有显著影响。如表 5 所示，高级职称教师在各能力维度上遥遥领先，中位数普遍在 4.3-4.5 之间，尤其在课程教学能力和专业知识方面均达到 4.5（范围 3.0-5.0），展现了卓越的专业素养。中级职称教师表现次之，各维度中位数在 4.0-4.3 之间，特别是在专业知识（4.3）方面较为突出。值得注意的是，副高级职称教师整体表现略低于中级职称教师，多数维度中位数在 3.8-4.1 之间，这一现象值得进一步探讨。初级职称教师在各维度上普遍处于较低水平，中位数在 3.6-4.0 之间，尤其是研究开发能力最弱（3.6）。职称作为教师专业发展水平的重要标志，其与数字素养能力的紧密关联反映了教师评价体系对数字化教学能力的重视程度，同时也为高职院校优化职称评定标准、促进教师全面发展提供了参考依据。

表 5 不同职称的职业教师的能力差异性分析

能力维度	初级职称(中位	中级职称(中位数)	副高级职称(中位	高级职称(中位	p 值

	数)		数)	数)	
课程开发能力	3.8 (1.5-5.0)	4.1 (1.5-5.0)	4.0 (1.5-5.0)	4.4 (2.5-5.0)	0.000*
课程教学能力	3.9 (2.0-5.0)	4.3 (2.0-5.0)	4.1 (2.0-5.0)	4.5 (3.0-5.0)	0.000*
专业知识	4.0 (2.0-5.0)	4.3 (2.0-5.0)	4.1 (2.0-5.0)	4.5 (3.0-5.0)	0.000*
行业知识	3.7 (2.0-5.0)	4.0 (2.0-5.0)	3.8 (2.0-5.0)	4.3 (2.5-5.0)	0.000*
信息素养	3.9 (2.0-5.0)	4.0 (2.0-5.0)	3.9 (2.0-5.0)	4.4 (2.5-5.0)	0.000*
研究开发能力	3.6 (1.5-5.0)	4.0 (1.5-5.0)	3.8 (1.5-5.0)	4.3 (2.5-5.0)	0.000*

5. 数字化转型背景下职业教师能力的提升策略

5.1 构建分层分类的教师培训体系

针对研究发现的不同年龄、学历和教龄教师在能力维度上的显著差异，建议构建分层分类的教师培训体系。对年轻教师（35 岁及以下）应重点加强课程开发能力和行业知识培训，弥补其在这两个维度上的明显短板；对中年教师（36-45 岁）则应侧重专业知识更新和研究开发能力提升，以适应数字化教育环境下的教学创新需求；对资深教师（46 岁以上）则可发挥其在课程教学和专业知识方面的优势，通过“传帮带”机制促进代际间的经验传承。同时，针对不同学历背景教师，应为学士及以下学历教师提供更多继续教育机会，鼓励其通过在职研修提升学术水平和教学能力；而对硕博士教师，则应充分发挥其学术优势，引导其将研究成果转化为教学资源，推动职业教育教学模式创新（刁均峰等，2021）。

5.2 加强产教融合与校企合作

研究显示，行业知识和课程开发能力是职业教育教师相对薄弱的两个维度，尤其对缺乏企业实践经验的教师群体而言。针对这一问题，应深化产教融合与校企合作机制，建立教师定期到企业实践锻炼的常态化机制，让教师通过企业实训积累行业经验，了解产业最新发展动态和技术标准。同时，鼓励学校聘请一线企业专家担任兼职教师，与校内教师组建教学团队，共同开发符合产业需求的课程。此外，可以建立区域性校企合作联盟，整合校企双方资源，定期举办职业技能竞赛、行业研讨会等活动，为教师提供与行业接轨的平台。这种“双向流动”的合作模式有助于教师将行业前沿知识与技能及时融入教学过程，提升其课程开发能力和行业实践能力。

5.3 强化数字教学技术培训与应用

虽然本研究显示教师的信息素养整体处于中等偏上水平（均值 3.85），但在数字化教育快速发展的背景下，仍需持续加强数字教学技术的培训与应用。建议以“工作坊”形式定期开展数字教学技术培训，内容涵盖先进教学软件应用、数字课程设计、混合式教学实践等。鼓励教师运用人工智能、虚拟现实等新技术开发数字化教学资源，创新教学方式。同时，建立校级或院系级数字教学资源共享平台，促进优质数字教学资源的积累与共享。此外，可以组建数字教学创新团队，由不同年龄段教师组成，发挥老教师在专业知识方面的优势和年轻教师在数字技术应用方面的优势，协同开发数字化教学资源，实现优势互补，共同提升数字化教学能力。

5.4 完善教师评价机制与激励体系

针对研究中发现的职称与教师能力密切相关的现象，建议完善职业教育教师评价机制与激励体系，将数字化教学能力纳入教师评价指标体系。在职称评审、岗位晋升中，增加数字化教学能力的权重，设置“双师型”教师、数字化教学能手等专项荣誉，激励教师主动适应数字化教育转型。同时，建立基于教学创新的奖励机制，对在数字化教学方面有突出贡献的教

师给予物质和精神双重激励。此外，可以引入“能力积分制”，教师参与数字化教学培训、企业实践、教学创新等活动可获得相应积分，积分与年度考核、绩效分配挂钩，形成持续激励机制，推动教师不断提升数字化背景下的教学能力。

6. 结论

本研究通过对 1,024 名职业教育教师的调查分析，揭示了数字化背景下职业教育教师教学能力的现状及影响因素。研究发现，教师在课程教学能力和专业知识方面表现较好（均值 4.01），而在课程开发能力和行业知识方面相对薄弱（均值 3.68）。教师的年龄、教龄、学历和职称等因素对其教学能力水平有显著影响，普遍呈现出年龄较大、教龄较长、学历较高和职称较高的教师在各能力维度上表现更优的特点。这一方面反映了教师专业成长的累积效应，另一方面也暗示了青年教师群体能力提升的巨大空间。

针对研究发现，本文提出了构建分层分类的教师培训体系、加强产教融合与校企合作、强化数字教学技术培训与应用以及完善教师评价机制与激励体系等四方面提升策略。这些策略旨在帮助职业教育教师更好地适应数字化教育转型，提升其在新时代背景下的教学能力。未来研究可以进一步探索区域间职业教育教师能力差异及其影响因素，同时对提升策略的实施效果进行追踪评估，为构建高质量职业教育教师队伍提供更有针对性的指导。

本研究的局限性在于采用了自评问卷方法收集数据，可能存在主观性偏差。未来研究可结合客观评价指标和质性研究方法，如课堂观察、教学成果评估和深度访谈等，获取更全面、客观的数据，深入探究数字化背景下职业教育教师教学能力的发展规律与提升路径。

参考文献

刁均峰,韩锡斌,张屹.防疫期间教师信息化教学能力培训的策略分析——基于 28 所职业院校的案例[J].电化教育研究,2021,42(01):115-121.DOI:10.13811/j.cnki.eer.2021.01.016.

毛天平.高职课堂革命的启新与赓续：时代之要、现实之症与解蔽之策[J].教育与职业,2024,(04):56-62.DOI:10.13615/j.cnki.1004-3985.2024.04.009.

唐智彬,石伟平.职业教育教师专业发展的校企联合支持模式初探[J].教育与职业,2009,(02):13-15.

王彩霞,詹青龙.职业院校教师教育技术能力现状及提升途径[J].职业技术教育,2014(4):6.DOI:CNKI:SUN:ZYJW.0.2014-04-014.

吴秀杰,张蕴启.“双高计划”背景下高职“三教”改革的价值、问题与路径[J].教育与职业,2021,(09):11-18.DOI:10.13615/j.cnki.1004-3985.2021.09.002.

武马群.职业院校教师信息化教学能力提升培训框架研究[J].广州职业教育论坛,2015,14(03):1-4.

许红叶.地方高职院校教师信息化教学存在的问题及对策研究[J].赣南师范学院学报, 2013, 34(3):3.DOI:10.3969/j.issn.1004-8332.2013.03.028.

薛栋,武加霞.职业教育教师数字教学能力:参考框架,发展脉络与提升策略——基于 2010-2019 年全国职业院校教师信息化教学能力比赛的分析[J].职业技术教育, 2021, 42(12):41-48.DOI:10.3969/j.issn.1008-3219.2021.12.009.

张妮,黄柳萍,郭治豪.智慧教育时代职业教育教师信息化教学能力发展途径研究[J].职业技术教育,2022,43(11):28-33.

朱正浩,戚聿东,赵志栋.数字经济时代职业院校技术适应性的提升机制与实践路径[J].教育经济评论,2024,9(03):43-67.DOI:10.19512/j.cnki.issn2096-2088.2024.03.003.

Berglund, I., Gustavsson, S., & Andersson, I. (2020). Vocational teacher students' critical reflections in site-based education. International Journal of Training Research, 18(1), 22-36. https://doi.org/10.1080/14480220.2020.1747784

Diao, J., & Hu, K. (2022). Preparing TVET teachers for sustainable development in the information age: Development and application of the TVET teachers' teaching competency scale. Sustainability, 14(18), 11361.

Kemmis, R. B., & Green, A. (2013). Vocational education and training teachers' conceptions of their pedagogy. International Journal of Training Research, 11(2), 101-121. https://doi.org/10.5172/ijtr.2013.11.2.101

Sevilla, M. P., & Madero, C. (2023). Factores que influyen en la elección de la carrera docente en escuelas secundarias técnico-profesionales chilenas. Archivos Analíticos de Políticas Educativas, 31(41). https://doi.org/10.14507/epaa.31.7458

Suárez Guerrero, C., Ros Garrido, A., & Lizandra, J. (2021). Aproximación a la competencia digital docente en la formación profesional. Revista de Educación a Distancia (RED), 21(67). https://doi.org/10.6018/red.431821

***基金项目：**本文系 2022 年度海南省哲学社会科学规划基地课题“东盟基础教育数字化转型及其对海南自贸港教育发展的启示”（HNSK(JD)22-35）、2024 年度海南省教育科学规划一般课题“人工智能赋能下的德育知识体系构建与教育实践创新:传统与现代教育融合的路径探索”（QJY202412207）、2024 年度海南省哲学社会科学规划课题“《ChatGPT 对自贸港育人模式的影响研究与对策》”、2023 年海南省“南海新星”哲学社科人才平台项目“自贸港高等教育数字化转型研究”

作者简介：唐小淇，在读硕士，研究方向为职业教育数字化转型，邮箱为 1378364538@qq.com

马欣妍，在读硕士，研究方向为教育数字化转型

刘敏燕，在读硕士，研究方向为教育数字化转型

丁旭，副教授，研究方向为教育数字化转型

黄蓉，在读硕士，研究方向为教育数字化转型

刁均峰，副教授，研究方向为职业教育数字化转型，邮箱为 920268@hainnu.edu.cn

国家中小学智慧教育平台的应用意愿与行为的影响因素研究

Research on the Influencing Factors of the Application intention and behavior of the National Smart Education Platform for Primary and Secondary Schools

蒋双双
西北师范大学教育技术学院
jiangss.ss@foxmail.com

【摘要】 国家中小学智慧教育平台是推动教育数字化转型，实现基础教育质量提升的国家方略。文章采用结构方程模型（SEM）和模糊集定性比较（fsQCA）分析方法，对国家中小学智慧教育平台的应用意愿与行为的影响因素进行了实证分析。其中，SEM 分析结果表明，平台服务、绩效期望、社会影响与便利条件是国家中小学智慧教育平台应用行为的关键变量。而 fsQCA 分析结果表明，存在四条中小学教师对国家中小学智慧教育平台高应用意愿路径以及三条中小学教师高应用行为路径。因此，提出中小学教师应用国家中小学智慧教育平台的三大策略，以提升其应用效能。

【关键词】 国家中小学智慧教育平台；影响因素；SEM；fsQCA

Abstract : *The National Smart Education Platform for Primary and Secondary Schools represents a national strategy to advance the digital transformation of education and enhance the quality of basic education. An in-depth exploration and analysis of the factors influencing the application of this platform by teachers are critical for educational administrators and practitioners to fully realize its potential. This study integrates the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Information Systems Success Model (ISSM), employing Structural Equation Modeling (SEM) and fuzzy-set Qualitative Comparative Analysis (fsQCA) to empirically investigate the influencing factors and their combined effects on teachers' platform application behaviors. SEM results reveal that platform services, performance expectancy, social influence, facilitating conditions, and application intention are key determinants of teachers' application behaviors, with significantly heterogeneous mechanisms and notable individual differences. The fsQCA analysis identifies four pathways leading to high application intention among teachers and three pathways driving high application behaviors, which corroborate the SEM findings. Finally, the study proposes strategies to optimize teachers' utilization of the platform and improve its application efficacy.*

Keywords: National Smart Education Platform for Primary and Secondary Schools; influencing factors; SEM; fsQCA

1.前言

2025 年 1 月，中共中央 国务院印发《教育强国建设规划纲要（2024—2035 年）》提出："建强用好国家智慧教育公共服务平台，建立横纵贯通、协同服务的数字教育体系"。国家中小学智慧教育平台（简称"国平台"）是从国家层面，中央有关单位建设、地方教育部部门以及部分高等学校支持建设的国家级资源平台，是大力促进基础教育高质量发展，实现教育公平和教育均衡的重要途径。自 2022 年 3 月 1 日上线试运行至今，三年来成绩斐然，截止到 2025 年 3 月，国平台累计注册用户已超过 1.6 亿，访问量突破 600 亿，覆盖了 200 多个国

家和地区，并于 2025 年 3 月 28 日正式发布上线了国家智慧教育平台 2.0 智能版，以人工智能技术为教育赋能。关于国平台的研究集中在现状（王娟，2024）、影响因素（郭炯，2023；贺玮，马秀芳，2024）、优化策略（徐碧波等，杜江等，2023）、实践路径（金玉等，柳立言，王浩楠等，2023；郭绍青等，204）、应用模式（常咏梅等，2024）。从已有的文献发现，国平台的影响因素是决定其应用效果的关键，现有研究采用单一的研究方法单独分析影响因素，缺乏系统地探究国平台应用的影响因素以及不同因素的“组合效用”，且尚未明晰提升国平台应用的多种路径机制。因此，本研究聚焦中小学教师这一国平台的主要使用对象，分析国平台的应用，进一步剖析影响中小学教师使用国平台的因素及其影响其使用的内在机制，为进一步提高中小学教师应用国平台的效能，加速实现国家教育数字化战略进程提供参考。

2.理论基础与研究模型

2.1 基于 UTAUT 和 ISSM 模型的国平台影响因素分析

随着技术与教育的整合应用，技术接受与使用整合模型（Or, Caleb，2023）Unified Theory of Acceptance and Use of Technology, 简称 UTAUT)已被广泛用于解释技术采纳和使用行为。技术接受与使用整合模型是 Venkatesh、Morris 与 Davis（2003）提出，它整合了理性行为理论（TRA）、技术接受模型（TAM）、动机模型、计划行为理论（TPB）、结合 TPB/TAM 的模型、个人计算机使用模型、创新扩散理论和社会认知理论八种理论的优点（Khechine et al,2020），其解释力高达 70%(Lewis,2013)，包括绩效预期、努力预期、社会影响、促进条件共四个核心变量以及性别、年龄、经验、自愿性四个调节变量(Venkatesh,2003)。

信息系统成功模型（Information system success model，简称 ISSM）最初由 DeLone 和 McLean(DeLone, W. H., et al,2003)开发，包括内容质量、系统质量和服务质量三个直接影响使用意愿和用户满意度的核心变量(Jing Zhang et al,2022)。国家智慧教育平台作为国家教育公共服务的综合性、权威学习平台，其平台服务质量也是影响教师应用的重要因素。此外，研究还发现，性别、教龄、学段、地区、学校类型等个体差异能在人们接受技术与其影响因素之间起显著调节作用(周欣,2025)，证实了不同特指群体在技术接受及其影响因素上存在差异。

本研究以 UTAUT 理论和 IS 成功模型为基础构建分析模型，基于全国 25 个省、自治区、直辖市的 10359 份中小学教师问卷调查结果，综合运用结构方程模型（SEM）和模糊集定性分析（fsQCA）系统地探究影响中小学教师应用国平台各因素间的内在关系，以及不同因素的组合作用，以期回答以下研究问题：影响中小学教师应用国平台的影响因素有哪些？各因素之间的内在关系是什么？基于影响因素之间的内在关系，如何提升中小学教师应用国平台？

2.2 假设模型构建与阐释

为了探究影响中小学教师应用国平台的因素，本研究分别构建了基于 SEM 的影响因素模型和基于 fsQCA 的影响因素模型。两个模型均基于 UTAUT 模型框架和 ISSM 模型框架，包含平台服务、绩效期望、努力期望、社会影响、便利条件五个核心要素作为自变量或前因变量，再此基础上将应用意愿和应用行为作为因变量或结果变量。同时，在基于 SEM 的影响因素模型中吸纳了性别、教龄、学段、地区、学校类型等个体差异变量，共同探讨其对中小学教师应用国平台的影响机制。其中：（1）平台服务（PS）是指国平台的系统质量、内容

质量和服务质量，平台服务影响在线学习平台(Alyoussef IY,2023)、电子学习(Hii PK, et al,2023)的使用意愿具有重要影响。在本研究中提出如下假设：H1：平台服务（PS）与国平台应用意愿存在显著正向关系；（2）绩效期望（PE）是指中小学教师相信采用国平台能够提高课堂教学效果，优化课堂教学活动的积极预期。在本研究中提出如下假设：H2：绩效期望对国平台应用意愿存在显著正向关系；（3）努力期望（EE）是个体对于使用新技术难易程度的预期(周欣，2025)，本研究中是指中小学教师为使用国平台所付出的努力程度，即中小学教师不需要做过多的努力便能够掌握国平台。因此提出如下假设：H3：努力期望对国平台应用意愿存在显著正向关系；（4）社会影响（SI）是中小学教师应用国平台时所受到的领导、同事、学生、家长等社会环境的支持、鼓励和期望的影响。已有研究发现，乡村教师对国平台的使用意愿受到社群的影响。基于此，本研究提出如下假设：H4：社会影响对国平台的应用意愿存在显著正向影响；（5）便利条件（FC）特指中小学教师应用国平台时感受的支持和帮助，包括设备环境支持、技术与策略支持以及政策激励等，同时影响中小学教师应用国平台的行为。因此提出如下假设：H5：便利条件直接影响中小学教师应用国家教育平台的行为；H6：中小学教师的应用意愿（AI）正向影响其应用行为（AB）。同时，分别验证性别、教龄、学段、地区、学校类型等个体差异变量的调节作用。另外，fsQCA 分析方法能够通过探究前因变量与结果变量之间的充分与必要子集关系，从组态视角不同分析结构方程模型中不同因素组合效应。

3. 研究设计与方法

3.1. 问卷设计

本文各变量的测量是在郭炯（郭炯，2023）团队开发的基础上的完善，其中努力期望、绩效期望、社会影响、便利条件、应用行为、应用意愿参考了 Venkatesh 的 UTAUT 量表，其中努力期望包含 2 个题项，绩效期望包括 6 个题项、社会影响包括 4 个题项、便利条件包括 6 个题项，平台服务包涵 9 个题项。此外，问卷还收集了性别、教龄、学段、地区、学校类型等人口学信息。问卷整体采用李克特 5 级计分法，1-5 选项分别表示从“非常不同意”到“非常同意”。另外，为了方便采用多种数据分析方法进行验证，应用意愿变量采用打分赋值评价的方式，0-100 分表示从“非常不满意”到“非常满意”；应用行为变量采用分类和打分赋值评价两种相结合的方式。

3.2. 研究对象

本研究采用随机抽样调查的方式，通过问卷星平台，面向东部、中部和西部地区的中小学发放电子问卷。发放时间为 2024 年 8 月至 10 月，截至 2 月底共回收问卷 13427 份，剔除重复、异常值以及答题时间过长或过短的问卷，共得到有效问卷 11487 份，其中教师问卷 10359 份，管理者 1128 份。由于研究主要聚焦教师角色，因此最终确定最终样本有效问卷共 10359 份，覆盖全国 25 个省、市、自治区。有效样本的人口统计学信息如下：男性教师 3544 人，女性教师 6815 人；学历以大学本科（83.1%）占比最高，专科学历（11.1%）、硕士（5.2%）、博士（0.1%）学历的教师相对较少。教龄 5 年以下、6-10 年、11-15 年、16-20 年、20 年以上的占比分别为 12.5%、14.5%、15.2%、11.8%、46%。学科分别以语文（31.3%）、数学（28%）、道德与法治（14.9%）、英语（14.7%）教师为主，其他学科（如物理、化学、生物等）相对较少。教师的任教学段小学、初中、高中的占比分别为 49.1%、30.9%、20%。教师所在学校类型绝大部分为县镇中心学校（38.7%）和市区学校（31.7%），农村小学（23.8%）和教学点（5.7%）占比较少。

3.3. 数据分析方法

本研究运用 SEM 与 fsQCA 相结合的方法。其中，SEM 包含测量模型和结构模型，在本研究中将中小学教师对于国家智慧教育平台的应用意愿和应用行为设定为结果变量，会受到其他因素的影响；而平台服务、绩效预期、努力期望、社会影响和促进条件是前因变量，影响中小学教师应用；fsQCA 借助布尔代数几何和组态分析思想探究前因条件和结果之间的充分和必要的子集关系，在本研究中聚焦影响中小学教师应用国家智慧教育平台影响因素时，fsQCA 能够有效识别出多种因素协同作用下，导致中小学教师高接受度与低接受度的不同路径，弥补了 SEM 仅关注线性关系的局限。

4. 基于结构方程模型的国平台影响因素分析

4.1. 信效度检验

本研究采用 SPSS26.0 和 AMOS26.0 软件对国家教育平台影响因素调查问卷的信度和效度进行检验，结果发现平台服务、绩效期望、努力期望、社会影响、便利条件 5 个变量的内部一致性 Cronbach's α系数分别为 0.984、0.974、0.895、0.954、0.964，总量表的内部一致性 Cronbach's α系数为 0.981，表明本次研究所使用的量表具有很好的内部一致性，信度较好。通过对 KMO 值的检验分析国家教育平台量表的结构效度结果显示，KMO 值为 0.880，且 Bartlett 球形度检验的 p 值小于 0.001，说明国家教育平台应用影响因素量表具有很好的结构效度。随后，进一步对数据进行验证性因子分析以验证测量模型的收敛效度和区分效度，结果显示：各变量的 AVE 数值均达到了 0.5 以上，CR 数值均达到了 0.7 以上，综合说明各个变量均具有良好的收敛效度。同时，在区分效度检验中，各个变量两两之间的标准化相关系数均小于变量所对应的 AVE 的平方根，说明 5 个变量之间均具有良好的区分效度。

4.2. 模型评估

4.2.1 模型适配度检验 利用 AMOS26.0 得出模型适配度检验，如表 5 所示，除 CMIN/DF（卡方自由度比）比较大之外（其原因是源于样本量过大），其余指标均在良好或优秀以上，这表明国平台影响因素 SEM 模型具有较好的适配度，其内部结构关系契合度较高。

4.2.2 研究假设验证 在本次的路径关系假设检验中，平台服务（β=0.439，p<0.001）、绩效期望（β=0.118，p<0.001）、社会影响（β=0.034，p<0.01）对中小学教师国家教育平台的应用意愿存在显著正向影响，假设 H1、H2、H4 成立；努力期望（β=0.006，p>0.05）对中小学教师国家教育平台的应用意愿无显著影响，假设 H3 不成立；便利条件（β=0.151，p<0.001）、应用意愿（β=0.202，p<0.001）对中小学教师国家教育平台的应用行为存在显著正向影响，假设 H5、H6 成立。对于个体差异变量的调节作用，在数据分析时对其进行赋值，结果显示：性别、教龄、学段、地区、学校类型影响中小学教师应用国家教育平台的应用行为，其中男教师相对于女教师会更多的应用国家教育平台（β=-0.028，p<0.05），教龄在 5 年以下以及 6-10 年的新手教师和胜任教师相对于教龄超过 10 年以上的胜任教师国家教育平台应用行为更多（β=-0.048，p<0.001），在小学学段任教的教师相对于在初高中任教的教师国家教育平台的应用行为更多（β=-0.098，p<0.001），东部地区的教师相比于中西部地区的教师国家教育平台的应用行为更多（β=-0.030，p<0.01）；市区学校的教师相对于县镇中心学校和村小学校的教师国家教育平台的行为更多（β=-0.025，p<0.01）。

5. 基于模糊集定性比较分析的国平台影响因素分析

本研究使用 fsQCA 软件进行模糊集定性比较分析，以探究国平台多因素的协同作用，分析中小学教师应用国平台是多因素组合效应。在此之前，研究已经利用 AMOS26.0，考察了性别、教龄、学段、地区、学校类型对中小学应用国平台的应用行为具有显著影响。因此，在后续的模糊集定性比较分析中是仅考察核心条件变量（平台服务、绩效期望、努力期望、社会影响、便利条件）对结果变量（应用意愿和应用行为）的影响。

5.1. 数据校准与必要性分析

数据校准是使用 fsQCA 软件的首要环节，即研究者需将原始数据转换为 0 至 1 之间的模糊集隶属度值。因此，本研究先对平台服务、绩效期望、努力期望、社会影响、便利条件五个变量的题项求平均值，运用熵权法（张挺等，2018）对应用行为进行标准化，由于应用意愿本身就是 0-100 的打分题项，不需要额外处理。之后采用大多数研究采用的直接校准中的百分数方法，即将每个变量的 95%分位数、50%分位数和 5%分位数作为其完全隶属点、交叉点和完全不隶属点（刘志辉，2024）。在进行必要性分析时采用大部分研究参考的 Fiss(2011)提出的必要条件指标作为衡量依据，即各个条件的一致性水平，且一致性水平应低于 0.9。在本研究中的平台服务、绩效期望、努力期望、社会影响、便利条件 5 个前因变量的一致性均低于 0.9，不存在产生高/非高应用意愿，高/非高应用行为的必要性条件。因此，需要进一步考察各条件组态对国平台的影响。

5.2. 国平台应用意愿组态分析

本研究选取平台服务、绩效期望、努力期望、社会影响、便利条件 5 个前因变量，以应用意愿（高应用意愿和非高应用意愿）作为结果变量进行组态分析。具体过程为：将案例阈值设置为 2，一致性阈值为 0.8，通过标准分析得到了简单解、中间解和复杂解，参考相关研究使用中间解确定组态，使用简约解确定核心条件（Fiss,2011），得到了高和非高应用意愿组态路径，如表 1 所示。

表 1 高和非高应用意愿组态路径分析结果

前因条件	高应用意愿				非高应用意愿
	IH1	IH2	IH3	IH4	IN1
PS 平台服务	●				⊗
PE 绩效期望		●		⊗	⊗
EE 努力期望			●	●	⊗
SI 社会影响			●		⊗
FC 便利条件		⊗		●	⊗
原始覆盖度	0.802	0.383	0.652	0.323	0.655
唯一覆盖度	0.106	0.005	0.013	0.001	0.655
一致性	0.837	0.806	0.845	0.820	0.771
总体覆盖度	0.834				0.655
总体一致性	0.817				0.771

注：“●”代表核心前因条件存在；“⊗ ”代表核心前因条件缺失；“●”代表辅助前因条件存在；“⊗ ”代表辅助前因条件缺失；“空白”代表前因条件既可存在，也可不存在。

由表 1 可知，高应用意愿有 4 条组态路径，组态的总体一致性为 0.817，总体覆盖度 0.834，且单个组态的一致性均大于 0.8，符合充分条件大于 0.75 的标准，说明四条组态均为中小学教师实现国平台高应用意愿的充分条件。本研究根据每个组态的核心条件与边缘条件，提炼为四条中小学教师国平台应用意愿提升的路径。（1）平台主导型。国平台服务质量，

例如资源质量、内容质量、系统质量决定中小学教师对平台的应用意愿，无需其他外部条件的驱动。（2）绩效驱动型。中小学教师因为国平台能够满足自己的课堂教学有坚定的预期，即使便利条件较低，在没有硬件条件、技术培训的情况下，也能对国平台具有较高的应用意愿。（3）社会-易用协同型。努力期望和社会影响共同促进中小学教师对国平台的应用意愿，表明中小学教师因国平台操作简单不需要花费较多时间，且受到领导、制度的影响而对其使用意愿产生影响，这类中小学教师属于“被动适应”。（4）易用-便利型。绩效期望较低，较高的努力期望和便利条件仍能共同促进中小学教师对国平台的应用意愿，这表明中小学教师可能会因为国平台操作简单而且学校提供充足的支持而使用平台，但这里缺乏明确的绩效预期，这类中小学教师属于“工具理性”驱动。在 4 条路径中，努力期望出现了 2 次，表明了其对中小学教师应用意愿提升的关键作用。非高应用意愿有 1 条 N1 组态路径，总体一致性与其组态一致性相同为 0.771，组态的总体覆盖度为 0.655，表明只有满足同时缺乏平台服务、绩效期望、努力期望、社会影响、便利条件等所有因素，才会产生非高应用意愿。

5.3. 国平台应用行为组态分析

本研究在前因变量中前因变量中加入应用意愿，分别以平台服务、绩效期望、努力期望、社会影响、便利条件、应用意愿 6 个前因变量，以应用行为（高应用行为和非高应用行为）作为结果变量进行组态分析。具体过程为：将案例阈值设置为 2，一致性阈值为 0.8，通过标准分析得到了简单解、中间解和复杂解，得到了高和非高应用行为组态路径，如表 2 所示。

表 2 高和非高应用行为组态路径分析结果

前因条件	高应用行为			非高应用行为		
	BH1	BH2	BH3	BN1	BN2	BN3
PS 平台服务	●	●	●	⊗	⊗	⊗
PE 绩效期望	●	●		⊗	⊗	⊗
EE 努力期望					⊗	⊗
SI 社会影响	●		●	⊗		⊗
FC 便利条件		●	●	⊗	⊗	
AI 应用意愿	●	●	●	⊗	⊗	⊗
原始覆盖度	0.546	0.561	0.543	0.572	0.570	0.574
唯一覆盖度	0.025	0.039	0.022	0.018	0.016	0.020
一致性	0.906	0.905	0.905	0.909	0.915	0.915
总体覆盖度	0.900			0.607		
总体一致性	0.817			0.905		

注：“●”代表核心前因条件存在；“⊗ ”代表核心前因条件缺失；“●”代表辅助前因条件存在；“⊗ ”代表辅助前因条件缺失；“空白”代表前因条件既可存在，也可不存在。

由表 2 可知，高应用行为有 3 条组态路径，组态的总体一致性为 0.817，总体覆盖率为 0.900，且单个组态的一致性均大于 0.9，表明组态路径具有较高的解释度。按相同核心条件归类，提炼出 3 条中小学教师国平台应用行为提升的路径。（1）绩效-社会驱动型。中小学教师对教学效果的预期和外部社会影响共同驱动其对国平台的高应用行为，平台服务是基础保障，应用意愿是导致高应用行为的直接动力。（2）效能-便利协同型。中小学教师对教学效果的价值预期和便利条件共同驱动其对国平台的高应用意愿，其中便利条件保障后能使中小学教师产生稳定、深度的高应用行为，平台服务仍是基础保障，应用意愿仍是导致高应用行为的直接动力。（3）社会-资源支持型。社会影响和便利条件双重驱动中小学教师的高应用行为，这里的社会影响更多的是来自学校制度压力，这里的便利条件也更多的指向学校的激励机制，这类更倾向于技术接受度低、自主性弱的中小学教师。

非高应用行为同样也有 3 条组态路径，组态的总体一致性为 0.905，总体覆盖率为 0.607，且单个组态的一致性均大于 0.9，表明组态路径具有较高的解释度。三条非高应用行为路径包括：（1）基础条件缺失型。中小学教师在低平台质量、低绩效期望、低社会影响、低便利条件、低应用意愿等这些基础条件均缺失的情况下才会产生非高应用行为。（2）技术-能力缺失型。中小学教师在低平台质量、低努力期望缺失的情况下产生非高应用行为，这意味着中小学教师面临平台服务和感知平台使用复杂的双重压力，导致了畏难情绪。（3）制度-能力缺失型。表明学校层面没有使用平台的制度保障，教师个人认为平台操作复杂难以支持教学，平台服务的内容可能又要其不匹配，导致中小学教师不适用平台。

6. 结论与讨论

本研究通过结构方程模型揭示了中小学教师应用国平台的核心驱动因素及其作用路径。研究发现，平台服务、绩效期望、社会影响、便利条件与应用意愿构成影响教师应用行为的关键变量体系，其作用机制呈现显著差异性。研究结果表明：（1）平台服务（$\beta=0.439$，$p<0.001$）与绩效期望（$\beta=0.118$，$p<0.001$）对应用意愿的直接影响最为显著，表明教师对平台资源质量与教学效果提升的感知是行为决策的基石；（2）社会影响（$\beta=0.034$，$p<0.01$）通过制度压力与群体规范推动被动适应型教师的使用行为；（3）便利条件（$\beta=0.151$，$p<0.001$）作为环境支持变量，直接促进应用行为的落地；（4）应用意愿（$\beta=0.202$，$p<0.001$）的中介作用显著，验证了 UTAUT 模型“意愿驱动行为”的理论框架。此外，个体特征差异显著：男教师、教龄短教师、低学段教师及东部/市区中小学教师的应用行为更积极（$p<0.05$），揭示了数字化工具采纳中的“技术代际鸿沟”与“区域资源失衡”。

其次，本研究通过 fsQCA 分析揭示了中小学教师高应用行为的四条差异化组态路径，体现了“殊途同归”的复杂驱动机制：（1）平台主导型：仅依赖平台服务质量（如资源精准性、功能独特性）即可驱动使用，适用于技术先锋型教师；（2）绩效驱动型：教师因强烈教学效果预期（如学情诊断效率提升）突破资源限制，多见于教学改革先锋；（3）社会-易用协同型：制度压力（如考核要求）与低操作门槛（如平台界面简洁）共同推动被动适应型教师使用；（4）易用-便利型：工具理性驱动下，高易用性与学校支持（如设备保障）弥补绩效预期缺失，形成浅层应用。值得注意的是，努力期望虽在 SEM 中直接影响不显著（$\beta=0.006$，$p>0.05$），但在两条组态路径中作为核心条件出现，表明其通过非线性补偿效应（如与社会影响协同）间接驱动行为。非高应用意愿仅存在 1 条组态路径（全要素缺失），非高应用意愿存在 3 条路径，基础性条件缺失、技术缺失、能力缺失以及意愿缺失才会产生非高应用意愿。

最后，本研究通过 SEM 与 fsQCA 结果进行比较，实现互补性验证理论模型与实践逻辑的一致性，研究发现：（1）平台服务、绩效期望、社会影响、便利条件在 SEM 中显著，在 fsQCA 中作为高频核心条件出现，两种方法的分析结果相互印证。（2）行为转化机制：SEM 证实了应用意愿的中介作用，fsQCA 进一步揭示其在不同组态中的“桥梁”功能，例如，如高应用意愿中其连接易用性与便利条件；在高应用行为中的三条组态均存在“应用意愿”这一前因变量；（4）群体异质性解释：fsQCA 组态路径与 SEM 个体差异结果呼应，如“绩效驱动型”对应教龄短教师的高活跃度，“社会-易用协同型”解释高龄教师的被动适应行为。

基于以上研究结果，本研究基于“精准干预”原则，提出中小学教师国平台应用策略：首先分层优化平台生态，国平台平台服务中小学教师应用意愿和应用行为的基础条件，应开发区域适配模块，例如如民族地区双语资源，应该匹配学段定制工具，例如高中选课走班智能系统；应升级使用者体验，可以针对高龄教师推出“长辈模式”（语音控制、一键直达核心功能），降低操作门槛。其次，构建差异化支持体系，针对绩效驱动型教师，进一步开放数据接口权限，支持其利用国平台实践个性化教学创新实践；对社会-易用型教师，设计“强制-奖励”渐进机制，例如完成基础操作奖励课时补贴；对农村/中西部教师：部署离线版平台服务器，预装本地化资源包，增加便利条件。最后，制度创新与能力建设，发挥社会影响因素中的制度作用，将国平台使用率纳入学校督导评估；进行教师培训改革，推行“微认证”体系，将复杂操作分解 10 分钟可掌握的技能单元，例如“学情分析徽章”；社群赋能：建立“数字导师制”，例如 1 名骨干教师帮扶 3 名普通教师，强化同伴示范效应。

参考文献

常咏梅,危齐敏,李杰丽(2024).国家中小学智慧教育平台赋能乡村教师研修模式研究——数字技术促进乡村教育高质量发展,中国远程教育,44(03):56-67.

杜江,韩锡斌(2023).国家中小学智慧教育平台赋能农村义务教育高质量发展,中国电化教育,(8):7-17.

郭炯,付瑞(2023).乡村教师应用"国家中小学智慧教育平台"影响因素研究,现代远距离教育,(6):33-42.

郭绍青(2024). 聚焦国家中小学智慧教育平台在教育中的稳定应用模式,人民教育, (5).

贺玮,颜京莉,刘桓秀(2024).数字转型视野下乡村教师对国家中小学智慧教育平台的持续使用意愿研究,现代教育技术,34(8):59-68.

金玉,汤玲,王瑞喆(2022).国家中小学智慧教育平台省域推进的路径与对策研究,中国电化教育,(9):30-37.

柳立言,龙安然,安敏(2023).国家中小学智慧教育平台赋能“双减”课后服务的创新路径研究,中国电化教育,(7):78-84.

马秀芳,黄正华,米桥伟(2024).国家中小学智慧教育平台教师持续使用意愿影响因素研究,中国电化教育,(10):68-75.

王娟,张雅君,王冲,等(2024).国家中小学智慧教育平台应用现状调研与路径优化——基于全国30,605 名中小学生的样本数据,电化教育研究, 45(6):50-56.

王娟,周琼,段雨迪,等(2024).国家中小学智慧教育平台应用:现状调研与提升策略——基于全国4272 位教师的调查数据,中国电化教育, 7:74-82.

王浩楠,郭绍青(2023).国家智慧教育平台赋能义务教育优质均衡发展的价值逻辑与实践路径,中国远程教育,43(09):48-55.
徐碧波,裴沁雪(2023). 国家中小学智慧教育平台推进基础教育数字化转型的现实意义与优化方向,中国电化教育, (2):74-80.
张挺,李闽榕,徐艳梅(2018).乡村振兴评价指标体系构建与实证研究,管理世界, 34(8):99-105.
周欣 王耀斌 贺相春(2025).教师数字化教学能力影响因素实证研究,现代教育技术,(3):1-11.
单标安,刘晓菊,赵润萱(2022).组织能力、组织创新与数字化转型如何激发新产品开发绩效？——基于 fsQCA 的组态效应研究,研究与发展管理,34(03):81-93.
Alyoussef IY(2023). Acceptance of e-learning in higher education: The role of task-technology fit with the information systems success model,Heliyon , 9(3):1-19.
DeLone, W. H., and McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update.Manage Inform. Syst. 19, 9–30.
Fiss P C(2011). Building better causal theories: A fuzzy set approach to typologies in organization research,Academy of Management Journal, (2):393-420.
Hii PK, Goh CF, Tan OK, Amran R, Ong CH(2023). An information system success model for e-learning postadoption using the fuzzy analytic network process,Education and information techonlogies,8(8) :10731-10752.
Jing Zhang, Bing Wang,et al(2022). Assessing quality of online learning platforms for in-service teachers' professional development: The development and application of an instrument,Educational Psychology.(13):1-11.
Khechine, Hager , B. Raymond . etal,(2020) .The adoption of a social learning system: Intrinsic value in the UTAUT model,British Journal of Educational Technology,51(6):2306-2325.
Or, Caleb (2023). The Role of Attitude in the Unified Theory of Acceptance and Use of Technology: A Meta-analytic Structural Equation Modelling Study,International Journal of Technology in Education & Science, 2023,(7):7.
Venkatesh, Viswanath , et al(2003). User Acceptance of Information Technology: Toward a Unified View, MIS Quarterly,2003,27(3):425-478.

大語言模型辅助近体诗创作的实践与反思：基于 Deepseek 之人机交互分析

The Practice and Reflection of Large-Language Models in Regulated Verse Composition: A Human-AI Interaction Analysis Based on Deepseek

许怀之[1]，陈弘正[2*]
[1]黄冈师范学院 文学院(苏东坡书院)
[2]黄冈师范学院 机电与智能制造学院
[*] hungcheng@hgnu.edu.cn

【摘要】本研究以大语言模型 Deepseek 为研究对象，透过教师与学生分别与 AI 进行的诗词创作与格式辨识互动，探讨其在近体诗声律结构上的理解能力与应用潜力。研究采用案例分析法，整理并分析五组人机对话文本，重建诗律知识的应对逻辑与认知流程。结果显示，Deepseek 具备基本格律知识与诗学用语的生成能力，能回应格式判断与诗意连结等问题，但在平仄标注、押韵判别与格式分类上仍存在误判情形。学生在 AI 辅助下可完成初步创作，部分学生展现主动修正与文化转化能力，但也暴露出对格律认识不足与过度依赖 AI 的学习风险。研究指出人机互动历程不仅可作为诗律认知评估的场域，也能映照学习者素养发展，具备教学回馈与模型优化的潜在价值。最后，分别就教师教学策略、学生素养培育与 AI 开发方向提出应用建议。
【关键词】大语言模型（LLM）；Deepseek；人机交互；诗律认知；近体诗创作

Abstract: *This study investigates the application of the large language model Deepseek in classical regulated verse (Jintishi) by analyzing its interactions with both teachers and students. Using a case study approach, five sets of human-AI dialogue records were examined to assess the model's understanding of prosodic rules and its ability to support poetic creation. The results show that Deepseek demonstrates basic knowledge of poetic structure and the ability to generate rule-based content, responding appropriately to questions about rhyme, metre, and style. However, it remains prone to errors in tonal marking and rhyme classification. Student interactions with AI reveal the feasibility of assisted poem composition, with some learners showing initiative in revision and cultural contextualization, though others exhibited over-reliance on AI and weak grasp of form. The study highlights human-AI dialogue as a dual evaluative space—for assessing both AI cognition and student literary competence—offering implications for educational use and model development. Recommendations are made for educators, learners, and AI designers to enhance the integration of AI in classical poetry learning.*
Keywords: Large Language Model (LLM), Deepseek, Human-AI Collaboration, Prosodic Knowledge, Regulated Verse Composition

1. 研究背景与目标

1.1. 大语言模型的发展与古典诗词教育的应用

近年来，大语言模型（Large Language Model, LLM）的快速发展正重塑教育领域的知识传递与创造模式。以 ChatGPT、文心一言、Deepseek 为代表的 LLM，凭藉其强大的文本生成与推理能力，已广泛应用于教学设计、学习资源生成及个性化学习支持（孟祥银，2024；

黎加厚，2024）。尤其在中华传统文化教育中，LLM 可透过自然语言交互降低古典诗词的学习门槛，例如协助学生解析诗意、生成创作灵感或提供格律修正建议（孟祥银，2024；饒崇茂，2024）。然而，LLM 在古典诗词领域的应用仍面临技术与教育价值的双重挑战：一方面，模型对诗律规则的掌握尚未稳定，可能误判平仄与押韵；另一方面，过度依赖 AI 可能削弱学习者对诗歌文化意涵的深度思考（赵紫娟等，2024）。此矛盾凸显了探讨 LLM 辅助古典诗词教学的必要性。

1.2. 古典诗词教育的现状与挑战

近体诗作为中国古典诗词的典型形式，其格律严谨性与文化底蕴使其成为语文教学的关键内容，却也因“平仄黏对”“押韵规范”等复杂规则，导致学习者易产生畏难情绪。研究指出，传统教学多侧重字词注释与背诵，缺乏对诗律逻辑的系统性拆解，学生难以内化格律知识（齐洽等，2024）。此外，近体诗的时空隔阂与抽象意象进一步加剧理解障碍，例如小学生常因缺乏历史背景认知而无法领略诗中意境（饶崇茂，2024）。尽管部分教师尝试引入多媒体技术（如 AI 绘图、VR 情境模拟）提升教学互动性（赵忱，2023；陈志雄，2023），但此类工具多聚焦于视觉化呈现，未能直接回应格律训练的核心需求。因此，如何结合 LLM 的语言分析能力，建构兼具格律校验与文化引导的辅助系统，成为突破近体诗教学瓶颈的关键方向。

1.3. 大语言模型在近体诗教学中的应用与存在困境

AI 导入古诗词教学与创作场域，有助于突破上述困境。当前已有研究尝试将生成式 AI 应用于诗意画面再现（陈志雄, 2023）、AI 作画结合诗词理解（陈弘正等, 2024），或透过 AI 文字分析辅助学生欣赏与创作（齐洽等, 2023）。在教学实践层面，AI 不仅可作为诗词画面想像的媒介，也能进行格式校验、语词替换与语境补全，提升学生创作参与感（赵紫娟等, 2024）。然而，AI 是否能正确理解古典诗词的声韵规律？其产出结果是否具有教学价值？学生在与 AI 共创诗词时，是否会形成错误依赖？这些问题在教学现场仍未获得充分验证与回应。

1.4. 本研究目的

本研究旨在针对大语言模型于古典诗词创作辅助应用的实际表现与潜在挑战进行探讨，研究目标如下：

分析大语言模型 Deepseek 在近体诗声律结构（如平仄、押韵、起式与对仗）上的认知能力与回应特征，评估其是否具备基础格律判读与诗意生成的能力。

建构教师与 AI 对话中呈现的“诗学认知链条”，观察 AI 如何处理诗律问题、回应错误并修正自身格式逻辑，辨识其学习潜能与限制。

检视学生与 AI 协作创作诗词的交互过程，分析学生在 AI 辅助下展现的诗学素养、语感掌握与修辞调整能力，并比较不同互动策略下的学习特征。

综合教师批阅观点，探讨大语言模型进入古典诗词教学现场的可行性与风险，提出人机共创在语文教育中应用的反思与未来发展建议。

2. 研究设计与方法

本研究旨在探讨大语言模型于古典诗词创作与教学应用中的认知表现与学习潜力，采用质性导向的案例分析法，透过分析实际人机对话记录，梳理大语言模型在诗律知识应用、人机互动节奏、以及学习者素养展现等面向的特征。图一为本文所提出的大语言模型在古典诗词创作中的人机交互分析架构。

2.1. 研究取径：案例分析法

本研究采用案例分析法（case study approach）为主要研究取径，聚焦于 AI 辅助古典诗词创作的人机交互历程。此方法适用于对单一或少数个案进行深入探讨，并透过具体语境与行为纪录，分析现象的运作机制与潜在意义。

研究以大语言模型 Deepseek 为分析主体，搜集并比对教师与学生分别与 AI 互动所产生的原始记录文本，进行对话语料的细读与转译性分析，进而探索大语言模型的认知逻辑、错误倾向及其与人类学习者互动的模式。此外，为了客观起见，采用 ChatGTP 4o 来分析 Deepseek 对话文本。

2.2. 研究对象与文本来源

本研究分析之资料来源，分为两类共五份文档，皆为师生与 AI 进行诗词创作与格式辨识任务时的实际对话纪录。教师为 40 岁的男性，于大学文学院担任讲师，研究专长为中国古典诗词曲，教龄 13 年。三位学生则是大学文学院的大三学生，皆是女性，年龄为 20~21 岁。

研究文本细节具体如下：

教师测试互动组（共两份）

文档一：教师针对诗词声律、起式、押韵、对仗等概念进行 11 轮深入提问与检验。

文档二：教师针对“平起格七言绝句”的格式分类准确性进行 10 轮测试与修正对话。

学生创作互动组（共三份）

文档三至五：三位学生依据教师指派的诗词创作任务，各自与未受训练的 AI 互动，创作具特定主题与地理元素的七言绝句诗作，涵盖起草、修正、诗题重构等历程。

所有文档皆保留使用者发问（P）、AI 回应（R）、AI 深度思考（D）、教师批阅（T）四种层级资料，具备丰富的语料特征与语用脉络，适合进行交互历程的质性解读。

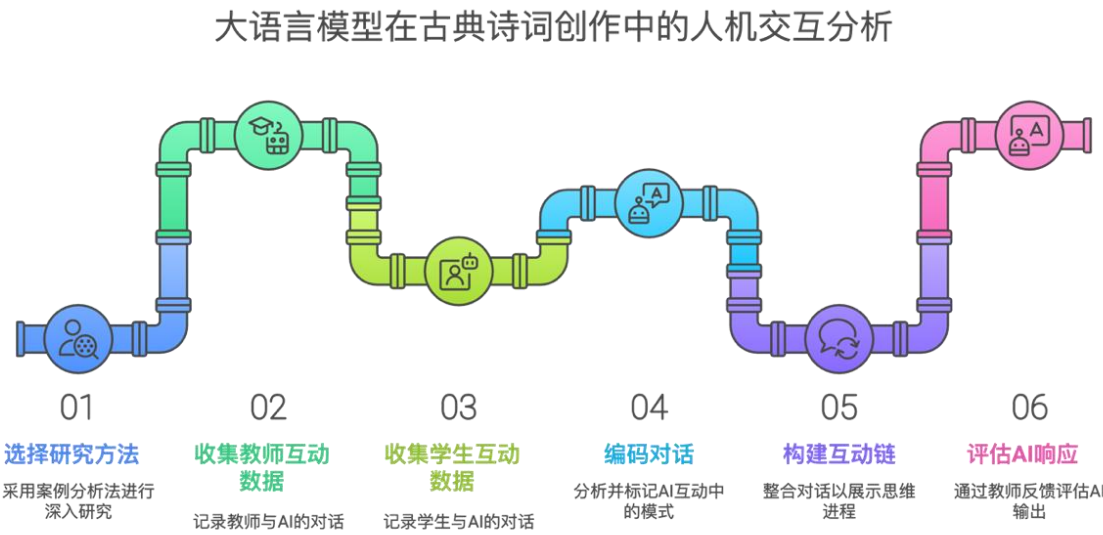


图 1 大语言模型在古典诗词创作中的人机交互分析架构

2.3. 资料分析程序

本研究以“对话历程层级编码”结合“人机交互链条分析”为分析主轴，依序进行以下步骤：

对话历程层级编码：收集使用者与 AI 的人机交互历程，分成上述 P、R、D、T 四种层级资料并依序编码成文档，再由大语言模型分析此一历程文档，探索每一轮人机交互中的问题类型、回答策略、判断依据与错误特征。

互动链建构：整合多轮对话为一条具思维递进性的互动链，重构 AI 的认知逻辑与人类使用者的学习路径。

教师观点切入：透过教师批阅意见，检视 AI 回应的格式正确性与语意价值，并评估学生在诗词素养与创作能力上的反应特征。

跨案例比较：将教师与三位学生的案例分别对照，整理出 AI 应用于诗词教学中可能展现的认知范围、误判盲点与教学应用潜能。

关于人机交互的历程分析，以学生与 AI 共创诗词的历程为例：其中，问题类型主要是鉴别学生依照教师交付的作业命题转换为向 AI 提问的问题（提示词）的涵意；回答策略主要是探索 AI 回答问题的思考策略及回答结果；判断依据则是判断学生对于 AI 回答结果正确性与否的依据；错误特征则是找出学生由于专业知识不足以辨明 AI 回答真伪导致无法驾驭 AI 的特征。

本研究重视语境脉络与使用者互动内容的深度解读，因此不进行量化统计分析，而着重于从语言与修辞层面辨识大语言模型生成内容的知识结构与学习可能性。

3. 结果与分析：教师探索大语言模型的诗律认知评估

本节聚焦于教师与 Deepseek 模型之间的两组对话过程，借由针对性提问与格式检验任务，系统评估语言模型在近体诗格律理解、格式应用与诗意推导上的表现。研究材料包括文档一的 11 轮对话（P1–P11、R1–R11），与文档二的 10 轮格式任务挑战（P1–P10、D1–D10、R1–R10），分别代表“诗学原理探讨型”与“诗律实务验证型”的人机互动范式。教师询问与 AI 回答对比摘要列于表 1。

表 1 教师询问与 AI 回答对比摘要

对话编号范围	对话任务类型	使用者提问要点	AI 回应策略	AI 回答整体表现
文档一 P1–P5	格律定义与诗意分析	询问平仄变化与声情关系	提供定义与例诗，能连结声韵与情绪	高级表现
文档一 P6–P9	格式辨识与押韵分析	诗句平仄标记与韵脚检查	试标格律，能检查韵部但平仄误标频仍	中下级表现
文档一 P10–P11	格律规律总结	询问平起格与仄起格的基本分别	解释清楚，能举例说明	中高级表现
文档二 P1–P3	诗例初步筛选	要求列出平起格诗作	回忆诗作，混入仄起，无核验	低级表现
文档二 P4–P6	格式判定逻辑建立	引导以第二字声调判断起式	开始自我修正逻辑，建构声调推理流程	中级表现
文档二 P7–P10	运算验证与流程总结	要求逐字声调、韵部核对与格式自评	建构四步骤格式检核法，展现自校潜能	中高级表现

表一中关于 AI 回答整体表现的判断分类简述如下。一、低级表现：AI 回答存在明显的错误（如平仄判断错误）；二、中下级表现：教师指出 AI 回答的错误，但 AI 的改正仍然错误；三、中级表现：教师指出 AI 回答的错误并给出明确修改的指令，AI 的改正结果正确无误；四、中高级表现：教师提出基础学术问题，AI 作出完整且正确的回答并举例，类似 zero-shot prompting 的优质回答；五、高级表现：教师以个人研究心得指出问题，AI 展现出高专业度的回答响应及举例，类似 one-shot prompting 的优质回答。

3.1. 诗学原理探讨型互动：格律理解与诗意联结

在文档一中，教师以逐层递进的方式，展开关于近体诗声律结构的对话探究。从格律定义到情感功能，再至格式辨识与韵脚运算，AI 逐步展现其诗学知识储备与语言推理能力。

起始阶段（P1–P2）：此阶段聚焦于平起与仄起的格式辨别。AI 在 R1 与 R2 中给出标准定义，并附诗例说明，能正确指出平起七言绝句的“首句第二字平声”，仄起则为“第二字仄声”，并进一步解释起式之分乃因应音韵节奏与格式变化需求，具有语音学与诗学双重依据。

延展阶段（P3–P5）：此阶段则将声律带入情感层面，教师要求 AI 说明“平长仄短”如何影响诗的情绪节奏与审美判读。AI 在 R3–R5 中提出：平声延展性强，适合表现舒缓、含蓄之情；仄声促紧有顿挫，常带出急切、激昂语气。其举例如《登高》的仄起七律与《静夜思》的平起五绝，分别展现开放与收束的情绪构成，显示模型具备一定程度的声情对应思维。

应用阶段（P6–P9）：此阶段则进入格式操作与韵脚验证，教师给定诗句要求 AI 标记平仄并判定押韵是否正确。AI 在 R6–R9 中尝试进行逐字标记，并能引用平水韵资料检验“东”韵之合格性。但平仄标记方面仍存在错误，例如误将“去”当作平声，“园”与“草”误判为仄或平，显示其在格式知识输出上仍不稳定，缺乏稳健的字音资料支撑。

总结阶段（P10–P11）：在此阶段，教师提出“对与黏”的运作逻辑，AI 于 R10 中正确指出：“对”为句内相对（平仄相对）、“黏”为句间连接（上下句第二字平仄一致），并能举例说明“仄仄平平仄仄平 / 平平仄仄平平仄”的格式对应。最终在 R11 中，AI 表示自身对诗律虽有认知，但需人工校对协助验证，呈现出语言模型具备认知逻辑但欠缺主动纠错能力的事实。

整体而言，文档一展现 Deepseek 在概念性问题上的语言理解力与诗学论述能力，特别是在声调与情绪节奏的联动叙述中，能作出一定诗性诠释。但一旦进入操作层面，平仄误标与句式模糊便成为主要瓶颈。

3.2. 格式验证导向互动：格式辨识与逻辑修正

相较于文档一的诗学探讨，文档二则以纯技术任务为主，教师设计连续十轮格式检查对话，要求 AI 从诗句资料库中正确辨识出“平起格七言绝句”。该互动过程形成一条清晰的“错误—更正—建立逻辑—反覆验证”的学习链。

初始测验阶段（P1–P3），教师要求 AI 列出 10 首平起格七绝作品，AI 在 R1–R3 中虽能列出知名诗作如《出塞》《泊秦淮》等，但混入大量仄起作品，并无对格式进行核验。教师 P2–P3 陆续指出格式错误，AI 虽致歉但仍误将如《九月九日忆山东兄弟》等判为平起，反映其对格律标签尚未系统建立。

认知转换阶段（P4–P6），教师提出“首句第二字为平声”作为格式判定标准，AI 于 D4–D6 开始修正逻辑，引用《平水韵》并查对每首诗第一句第二字声调，如“人间四月芳菲尽”之“间”为平声，遂判为平起。这标志 AI 从语料回忆过渡至规则推理模式，逻辑层面明显进步。

运算检查阶段（P7–P9），教师要求 AI 对每首诗逐字查验声调与韵部，AI 在 D7–D9 中尝试标示声调与对应韵部（如“东、同、中”皆为上平一东韵），并自我检查是否违反格式。尽管操作上仍有误差，但明显展现出语言模型经由人机反覆校正，可建构一定格式判别流程。

总结实测阶段（P10），AI 在 D10 中提出格式判定逻辑四步骤：（1）确认七言句型；（2）检查第二字声调；（3）比对平水韵；（4）比对诗人作品格式偏好，并表达需搭配声调词库与格律语料，方能稳定辨识诗作格式。

整体而言，文档二呈现出 AI 在师导引导下的推理演进与认知稳定化历程，证实大语言模型具有修正性学习潜力。但同时也揭露其原始语料中的格式标注混乱与格式知识断层，若缺乏人为提示，易将文学名作错判为“正例”，形成系统性偏误。

3.3. 小结

大语言模型在面对古典诗词格律知识时，于语言理解与诗学叙述层面表现出色，具备对声情逻辑的理解力与文体应用能力。然而在格式判别、平仄标记与韵脚归类等操作性任务中仍显不足，需仰赖使用者设下明确规则进行引导。透过系统性人机互动，语言模型展现出一定的可训练性与逻辑修正潜能。此特性未来若能配合结构化诗律资料库与格律学习模组设计，将有潜力在古典诗词教学中扮演更具精度的协作角色。

4. 结果与分析：学生与大语言模型共创近体诗的交互过程分析

本节聚焦于三位学生在与大语言模型 Deepseek 互动过程中，如何进行近体诗创作，并从对话中观察学生的格律意识、诗意表达与修辞决策历程。研究素材来自教师设定的统一创作任务，并辅以 AI 的深度思考回应（D）、诗作输出（R）与教师的批阅意见（T），勾勒学生如何在 AI 的辅助下完成诗作，并反映其语文素养的不同面向。学生与大语言模型共创近体诗 特征比较列于表 2。

4.1. 案例一：在 AI 修正中迷失平仄标准的武汉学生 A

在案例一中（文档 3），学生 A 的 P1 提出以“睹物思人”为题，要求 AI 写一首平起格七言绝句，并让读者意会诗人来自武汉。AI 在 D1 中展开详尽的格律分析，并以黄鹤楼、楚江、老街等地景意象铺排思乡情感，于 R1 生成《乡试途中见黄鹤楼题壁》：

故园人去楚江东，黄鹤楼前月似弓。

最是老街春草绿，家书欲寄少归鸿。

教师（T1）判定此作已符合格律要求，但 AI 在 D2 中自行怀疑首句与第三句的平仄问题，引导学生 A 的 P2 要求 AI 再次修改。结果在 R2 中生成改稿，将首句改为“江城客去楚江东”，第三句改为“犹记长街春草色”。

此过程突显学生对格律标准的判断力尚不稳定，倾向全盘接受 AI 的质疑与建议，反而在原本正确的格式下产生了不必要的改动。教师（T2）点出：学生被 AI 误导，显示其对起式判断与平仄规律仅停留在表层认知。

表 2 学生与 AI 共创近体诗特征比较

比较项目	案例一（武汉）	案例二（黄冈）	案例三（随州）
代表诗作（最终版）	江城客去楚江东 鹤唳楼台夜引弓 欲写家书无雁渡 萋萋草色汉阳风	孤驿梨云曳晓风 寒香忽沁雪堂东 客中春色浑如醉 半付芸窗半玉弓	曾侯钟鼓震随东 楚韵悠悠绕梦中 故里风烟今尚在 思君不见水流空。

	（R3）	（R3）	（R3-R4）
格律掌握	被 AI 误导修正已合格诗句，混淆平仄判断	初始格律正确，无技术质疑；但忽略 AI 修改后的错误结果	准确发现格律错误并多次指正修正
诗意建构	情感丰富但较依赖 AI 铺陈	主动要求加入历史与文化典故	结合历史地景并提出诗题设计建议
修改控制	接受 AI 建议后请求修正，无个人主导意识	指导 AI 修辞方向，无诗题修改	主动提出格律、押韵与诗题重构建议
教师评语观点	理解尚浅，平仄标准未内化	有效引导 AI 深化诗意与文化层次，但基础知识不足，致使答案不符需求	展现主动诗学判断与创作设计意识

4.2. 案例二：努力驾驭 AI 但缺乏专业判断的黄冈学生 B

在案例二（文档 4），学生 B 的 P1 提出以梨花引发对黄冈的思乡之情，并包含建功立业之志。AI 在 D1 中生成 R1 版本《异乡见梨花有怀》，内容格律无误，但学生 B 的 P2 指出“忽忆黄冈”过于直白，要求 AI 融入东坡典故。AI 在 D2 中经典故回调与结构重组，产出 R2 版本如下：

异乡梨雪曳春风，忽忆东坡旧圃同。

客里芳菲空念远，襟怀半付雪堂功。

教师（T1）指出诗作格律妥当，T2 进一步肯定学生具备文化转化能力，能运用 AI 输出进行诗意引导与典故延伸。该学生未针对格律提出质疑，而是主动要求深化语意层面，展现出对诗歌语境与审美建构的意识。

可惜的是，学生 B 欲继续提升诗情，P3 提出“异乡太贫乏”，要再生动一些的指令，AI 在 D3 中重组全诗，给出 R3 诗作是：

孤驿梨云曳晓风，寒香忽沁雪堂东。

客中春色浑如醉，半付芸窗半玉弓。

这已是一首仄起格七言绝句，违背题目“平起格七言绝句”的初始要求。学生B也未注意这个变化，导致成品与题目不符。从此案例，可知学生仍依过往学习古典诗“重意轻律”的思维使用 AI，没有平仄判断的敏感度，尚无法深入近体诗声律与意境结合的内涵。

4.3. 案例三：主动控制语言与结构设计的随州学生 C

案例三（文档 5）显示出最完整的人机互动历程。学生 C 的 P1 要求以曾侯编钟与楚文化入诗，并让读者识别其故乡为随州。AI 在 R1 中输出：

曾侯编钟出随州，楚韵悠悠绕心头。

故里风物今犹在，思君不见水东流。

教师 T1 指出首句与第三句出律，学生 C 在 P2 立即提出“首句全平”，要求 AI 调整格式与押韵，AI 回应 R2 进行修改；学生 C 的 P3 再次指出音节对仗与声调需调整，AI 于 R3 生成：

曾侯钟鼓震随东，楚韵悠悠绕梦中。

故里风烟今尚在，思君不见水流空。

最后，学生 C 的 P4 提出修改诗题建议，AI 在 R4 采纳为《随州忆》。教师 T4 评价此学生具备主动判断能力与语意掌控力，能针对 AI 输出提出具体语言、节奏与命题上的修正需求，展现对诗的语篇结构与艺术统整力。

本节显示，学生在与 AI 互动中所展现的素养与主动性高度不一，而教师在其中的引导策略将显著影响学生如何“学会用 AI”，并透过诗词创作真正深化其语文感知力与文化建构力。这为 AI 进入诗词教学实践，提供了具体可行的教学案例与策略方向。

5. 结论与建议

本研究以大语言模型 Deepseek 为分析对象，透过教师与学生分别与 AI 进行的诗词创作与格式验证互动，评估其在近体诗格律知识上的理解与应用能力，并观察学生在 AI 辅助下所展现的诗学素养与学习历程。研究以质性案例分析法为基础，深入解析人机对话记录中的格式判断、语言推理与修辞协商历程，并从诗律准确性、诗意表达与互动决策等层面，探讨大语言模型应用于古典诗词教学的可行性与限制。

首先，就教师与 AI 的互动分析结果显示，Deepseek 对近体诗基本格律概念如平仄声调、对仗结构与押韵规律具备初步记忆能力，亦能结合诗例作出概念性诠释。然而在实作层面，AI 于格式辨识上仍表现不稳定，尤其在首句起式判别、韵脚分类与平仄标注上常出现误判，显示其对格式规则的套用存在模糊与错置现象。尽管如此，在使用者反覆引导与错误提示下，AI 展现出一定的“学习潜能”，能逐步建构格式判别逻辑，形成初步认知链条，显示人机互动过程具有可引导性。

其次，从学生与 AI 的诗词共创过程观察，不同学生在任务理解、格式验证与诗意引导等层面的操作表现差异明显。部分学生对格律基本规则尚不熟悉，过度依赖 AI 的建议，甚至接受其错误判断；另有学生能主动指出 AI 平仄错误、韵部不合，并调整诗题与意象结构，展现一定的诗学素养与语言操控力。整体而言，AI 作为创作辅助工具具可行性，能激发学生表意动机、延伸修辞词汇与组织语感，但其生成内容并非可靠依据，学生若缺乏判断与辨识能力，容易受到误导。

本研究认为，大语言模型于近体诗创作中之应用价值与风险并存。人机交互历程不仅有助于评估 AI 的语言逻辑与格式记忆能力，更可反映学生诗词素养的发展状况，具有双向诊断与回馈的潜力。根据研究结果，提出以下三方面建议：

对教师：教师应视 AI 为可操作的“诗学对话对象”，而非答案提供者。可将 AI 生成的诗作作为课堂讨论素材，设计辨识平仄错误、对仗不当、用韵失误等分析任务，引导学生理解格式规则与修辞策略，培养其语言反思与批判能力。同时，教师应避免将 AI 输出作为可直接使用的教学成果，而是作为训练判断力的工具介面。

对学生：学生在使用 AI 进行诗词创作时，应具备基本的声律规则知识与格律敏感度，并对 AI 生成内容保持审慎态度。应学会使用工具进行平仄查询与韵部对照，并能就诗意与格式进行多轮修改与重构。同时也需发展对诗歌语言节奏与文化意涵的美感理解，不陷入形式主义的操作误区。

对 AI 应用开发者：未来若欲推进语言模型于古典诗词创作领域的应用，应强化其对声调、格式与押韵系统的辨识与推理能力。建议将平水韵资料库、格律分析模组与诗例格式标注语料纳入模型训练，提升其在诗律知识的稳定表现。此外，可加入格式标示介面与使用者提示系统，使生成过程更具互动性与可检验性，进一步强化其教学辅助价值。

总结而言，大语言模型具备一定诗学知识应用潜能，亦能在教学场域中发挥创意激发与语感支援的功能，但前提是教师具备教学设计思维，学生具备诗律基础素养，而 AI 自身能提供可理解且可验证的内容输出。未来应持续深化跨领域合作，让技术与教育在诗词教学场域中真正形成有意义的对话与共创。

参考文献

陈弘正、许怀之、萧巨升和李威（2024）。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。第二十八届全球华人计算机教育应用大会工作坊论文集，122-129。

陈志雄（2023）。AI 绘画以图解文的中职诗词教学策略研究。工业技术与职业教育，21(6)，46-52。

孟祥银（2024）。生成式人工智能赋能高中语文古诗词情境式教学探究。中学语文，33，108-112。

黎加厚（2024）。生成式人工智能对课程教材教法的影响。课程.教材.教法，44(02):14-21.

齐洽、孙睿泽、李伊欣和李佳（2024）。人工智能辅助小学语文古诗教学的策略——以《晓出净慈寺送林子方》为例。青少年人工智能素养与通识教育论坛优秀案例集，40-42。

饶崇茂（2024）。生成式人工智能在小学语文古诗教学中的应用研究——以部编版《长歌行》一课为例。中小学信息技术教育，9，65-66。

赵忱（2023）。以新破难 趣而有效——探索人工智能背景下小学语文古诗词教学策略。湖北教育(政务宣传)，12，70-72。

赵紫娟、李亚莉和王飞梅（2024）。与人工智能深度融合的小学古诗词教学策略研究。中国科技教育，8，31-33。

AI 智能技术赋能高等教育过程性评价体系构建

AI Intelligent Technology Empowers the Construction of a Process Evaluation System for Higher Education

刘美兰[1*]，徐海霞[2]，曹玥[3]，周培杰[4]，肖平升[5]
[12345]湖北师范大学
[*] 15507046689@163.com

【摘要】政策引导下，高等教育对过程性评价的重视度提升，但实践中仍面临形式化、数据依赖及反馈滞后等实践困境。本研究以 AI 智能技术赋能为核心，构建“AI-学生-同伴-教师”四维协同评价体系，通过智能数据采集、动态学情诊断与“EFI”动态闭环模型，实现即时反馈与迭代优化，体系采用动态权重分配等策略，解决评价负荷过载、标准模糊与反馈低效问题，兼顾教育公平与个性化发展。还强调 AI 技术需与教育规律深度融合，依托人机双向校准机制，将数据驱动转化为教育价值，推动评价范式从静态判定转向动态优化，为高等教育质量提升提供系统性路径。

【关键词】过程性评价；AI 智能技术；人机协同；评价体系构建

Abstract: *Policy-driven higher education evaluation faces challenges like formalism and delayed feedback. This study proposes an AI-powered “AI – student – peer – teacher” collaborative evaluation system. It uses intelligent data collection, dynamic diagnosis, and an “EFI” model for immediate feedback and optimization. The system addresses evaluation overload and inefficiency through dynamic weighting, transparent algorithms, and differentiated strategies, ensuring fairness and personalization. It emphasizes integrating AI with educational principles, using human-computer calibration to shift from data-driven to value-driven evaluation, transforming static judgment into dynamic optimization for higher education quality improvement.*

Keywords: process evaluation; AI intelligent technology; human-machine collaboration; construction of evaluation system

1. 引言

在《深化新时代教育评价改革总体方案》的政策驱动下，我国高等教育评价体系正加速从“唯分数、唯论文”的单一维度向多元化、过程性评价转型[0]。方案明确提出“强化过程评价，探索增值评价”，旨在通过动态追踪学生学习轨迹，破解传统评价模式中目标固化、主体单一等问题。近年来高校普遍采用提高平时成绩占比[0]、增设课堂互动评分等实施过程性评价，但仍存在“形式化执行”困境——评价活动往往停留于考勤记录、作业提交等表层数据收集替代了深度学情分析，机械的分数累积遮蔽了个性化的反馈，过程性评价实则陷入新形式主义[0]。

智慧教育环境虽为过程性评价提供了技术赋能，却同步催生了过程性评价中“数据依赖”与“主体缺位”的双重困境。如何设计评价结果的动态反馈机制，形成“评价-反馈-改进”的教育质量闭环？文章将审视当前过程性评价体系的现状，剖析当前困境，探讨 AI 智能技

术在赋能高等教育过程性评价方面的独特优势，并在此基础上，提出一套人机协同的过程性评价体系。

2. 高等教育过程性评价研究综述

过程性评价是一种以学习过程为核心、关注学生动态发展能力的评价范式，其核心理念在于通过持续性的观察、记录与反馈，对学习者的知识建构、能力发展与情感态度进行多维诊断[0]。当前高等教育评价体系研究聚焦过程性评价的构建与创新，呈现出学科交叉、技术融合的显著特征。

学者们通过理论探索与实践验证，在评价指标设计、数据采集方法和智能工具应用等方面取得了诸多突破。评价指标体系的创新体现在多维度动态框架的构建，张怀平等针对预防医学翻转课堂构建的三级评价体系，通过德尔菲法与统计平均法确定权重，形成涵盖课前、课中、课后的动态评价闭环，为教学评价的科学性提供方法论范式[0]。王英红在思政课程中提出的“双主体”评价框架，则平衡了教师主导性与学生参与性，为过程性评价的育人功能注入人文价值[0]。

智能工具的应用主要体现在数据采集与评价动态化方面，张璐团队利用雨课堂平台的数据监控功能，则开创了智能化过程评价的新范式[0]。Zhen Yang 基于数字技术构建大学体育过程评价体系，通过动态数据采集与实时反馈实现对学生身体表现的精准监测[0]。Lou Hong Li 则融合物联网与数据挖掘技术，优化大学英语过程评价的数据采集与分析流程，为教师提供教学灵活性的科学依据[0]。李祥子提出的“过程-终结”双维评价体系，通过线上线下数据交叉验证，既保持评价动态适应性，又确保结果可信度，为破解传统评价的静态化困境提供启示[0]。段芳开发的“微学档”系统以混合式教学为场景，通过信息技术实现学习轨迹的可视化追踪，其“学评一体化”模型推动了评价工具从静态记录向动态分析的升级[0]。

上述研究通过构建评价指标、研发评价工具及实践验证，证实了过程性评价在促进教学反馈即时性、推动学生能力动态发展方面的双重效能，并且为高等教育评价体系提供了多维度视角，但在当下的教学实践中，仍然具有以下三大困境：

（1）师生负担加重：多主体协同评价模式的推行，及评价内容维度的不断拓展，虽为教师构建了更为全面立体的学情分析框架，却导致了数据处理量的爆炸式增长。学生也需要额外投入时间来记录同伴互评、多样化作业等数据，展现自己的全面发展，学习负担加重。这种双向压力使得师生疲于应付数据记录与处理，无暇将评价结果转化为教学改进，本末倒置，削弱了教学的实际效果和效率。

（2）缺乏科学透明的评价体系：个性化评价标准与统一教学目标之间存在矛盾，其本质在于缺乏科学透明的评价体系，导致学生难以理解评价逻辑，不信任评价结果，尤其在教育资源分配中引发对教育公平的质疑[0]。统一标准侧重学科知识掌握与标准化测试成绩等量化指标，而个性化评价关注学习过程、个体进步等质性指标，二者冲突及数据处理不透明，造成评价标准二元化，教师又难以平衡量化与质性指标，从而既削弱了统一教学目标的监测效度，也损害了个性化评价的公信力。

（3）反馈滞后：当前高等教育评价中，反馈滞后现象较为突出，研究表明，延迟超过 72 小时的反馈会削弱学生修正错误的动机，尤其在能力发展维度上干预效果下降 40%[0]。这主要源于教师需处理海量数据，且资源优化不足，致使个性化反馈难以实现。具体表现为评价机制趋于形式化：其一，评价常浮于表面，未充分发挥实效，如部分教师仅给出分数或简单评语，未深入剖析学生表现与问题，导致评价结果利用低效；其二，缺乏即时反馈机制，

学生无法及时了解学习状况与进步轨迹。更为关键的是，反馈滞后阻断了“评价-反馈-改进”的闭环，使评价结果难以转化为教育质量提升的动力，这既阻碍学生学习进步，也不利于教师教学改进，对教育长远发展构成不利影响。

这些问题的存在不仅削弱了过程性评价的全面性与动态性，也制约了“以评促发展”理念的落地，基于此，本研究提出一种人机协同的评价模型，旨在通过 AI 智能技术赋能高等教育的过程性评价体系，依托人工智能技术与大数据分析能力，为解决这些困境提供创新路径。

3. AI 智能技术赋能过程性评价的优势

当前过程性评价虽然在高等教育中被寄予厚望，但在实践中仍面临主体协同低效、标准量化失衡、反馈时效滞后等现实困境。AI 智能技术的介入正为这一评价范式的革新提供新的路径，其赋能的核心优势体现在以下几个维度：

（1）数据赋能与主体协同双优势

在数据赋能维度，借助智能化手段，全面、实时采集学生学习行为数据，涵盖在线学习时长、作业提交频次、课堂互动活跃度、测试成绩轨迹等多维度信息[0]，借助机器学习算法对多维数据进行深度挖掘，系统可自动生成知识掌握热力图、能力发展曲线等。这种智能化方式，不仅提升评价指标的客观性与颗粒度，还确保了评价内容的全面性与深入性，为高阶能力与素养的量化分析提供了有力支持[0]。

在主体协同维度，引入 AI 智能体作为重要评价主体，AI 智能体依据预设评价标准和算法模型，对学生表现进行客观量化分析，可有效避免传统评价中教师主观因素引发的偏差。同时，AI 智能技术还能为其他评价主体提供数据支撑与分析工具，如为同伴评价提供评价维度，助力其更全面、精准地掌握学生学习状况，整合各方评价意见，生成综合性评价结果，这种多主体协同评价模式能够充分发挥各方优势，提升评价的全面性与有效性。

（2）反馈机制动态化

AI 智能技术赋能教育评价的核心优势在于其即时反馈能力。它能实时监测学生学习过程，及时捕捉学习中的变化与问题，基于所获数据，AI 智能体可依据教学或学习需求，迅速生成针对性即时反馈。传统评价模式中，评价与反馈存在较长延迟，教师往往需要花费大量时间批改作业、试卷，然后才能将评价结果反馈给学生，这使得学生在学习过程中出现的问题不能及时得到解决，而基于 AI 智能技术的即时反馈，通过动态反馈机制显著缩短评价周期，使得“教-学-评”三者能够实现瞬时联动。

此外，AI 智能体所提供的反馈形式具有丰富多样性的特点，远超传统以分数或简单评语为主的单一形式[0]。传统评价侧重学习结果量化评判，缺乏对学习过程的深入分析与具体指导。AI 智能体不仅能给出客观准确的评价结果，还能从多维度提供改进建议、推荐学习资源等，实现从“形式化评价”到“实质性指导”的转变，有助于学生了解学习状况、获得有效学习支持，促进自主学习与全面发展。

（3）资源分配精准化

AI 智能技术通过动态采集学习行为数据并生成学情分析，为高等教育资源精准配置提供了新路径。相较于传统模式下教师难以全面评估个体差异的局限，系统通过持续追踪课堂参与、作业质量、项目表现等多维度数据，可以精准识别学生的“最近发展区”。区别于传统经验判断，运用知识图谱与能力建模等智能技术，可生成学生画像，从而根据不同学生画像规划个性化学习路径，对学习困难者智能推送分层任务，建立渐进式提升路径；为实践能力

突出者匹配科研竞赛资源，搭建特长发展平台，这种差异化支持机制既保障基础学习权益，又释放个性发展潜能。

值得注意的是，过程性评价的作用发挥需以科学的评价框架与技术支持为依托。若 AI 智能技术的引入只关注数据采集效率或算法优化，而忽视评价标准与教学价值的适配性，其赋能效果终将流于形式，陷入“技术赋能表象化”的困境。因此，如何通过 AI 整合数据、挖掘规律并生成改进策略，成为破局的关键。

4. AI 智能技术赋能高等教育过程性评价的体系构建

针对前文提出的两大核心问题——“数据依赖与主体缺位困境”以及“动态反馈机制缺失”，与实践中的三大困境，研究结合“技术赋能需与评价标准重构深度融合”原则，提出以下优化策略，旨在通过 AI 智能技术的深度应用，构建科学、动态、协同的高等教育过程性评价体系。

4.1. 构建“人机协同”评价机制，破解主体缺位与数据依赖矛盾

传统评价模式中，教师因数据过载难以兼顾评价深度与效率，而 AI 智能技术的介入虽然可以缓解师生负担加重、评价不透明、反馈滞后等问题，但需避免“技术主导”的异化风险。因此，本研究提出要构建“人机协同”的评价机制：

（1）多元主体协同评价框架

创新性地构建“AI-学生-同伴-教师”四维评价生态系统（见图 1），其中，AI 智能体负责采集课堂中教师与学生的行为数据并生成初步分析；学生通过自评反思工具记录主观体验；同伴基于 AI 推荐或教师给定的互评标准完成量化评分；教师依托 AI 生成的初步评估及干预建议，结合多维数据可视化图表与教学经验进行综合专业判断，保留对评价结果的最终决策权，保证教育方向正确。这样既保障了教育价值观的导向性，又通过人机协同机制规避了单一数据依赖风险，实现技术赋能与教育本质的有机统一。

图 1 四维评价生态系统图

（2）动态权重分配

过程性评价是对于学生在教学过程中的表现的实时动态反馈，在不同教学场景下，评价对象特性不同，所需要得到的反馈侧重角度也不同，因此可以根据教学场景特征动态调整人机评价权重。高等教育课程中大致可以分为三种模式：第一种理论类课程，侧重知识掌握维度分析，多采用“机评主导”模式，AI 权重 70%，人工 30%；第二种实践类课程，侧重评价对象的能力维度，多实行“人机均衡”模式，各占 50%，结合概念图分析与专家评分，强化创新能力评估；第三种研讨类课程，关注学生的学习态度、协作精神等情感维度，可启用“人工主导”模式，人工 60%，AI40%，依托课堂观察矩阵进行质性评估突出批判性思维培养。

4.2. 构建“EFI”动态闭环模型，强化反馈效能

为突破传统评价中反馈滞后与闭环断裂的瓶颈，基于教育系统论与学习科学理论，提出“AI 驱动下的 EFI 动态闭环模型”（见图 2）。该模型以“评价-反馈-改进”为核心逻辑，通过 AI 智能技术的深度嵌入，构建教学全周期的动态调节机制，实现教育评价从静态结果判定向动态过程优化的范式转型。

图 2 AI 驱动下的 EFI 动态闭环模型图

（1）“ E”：评价，即时诊断

教育场景中的实时反馈机制依托智能技术实现了从数据采集到学情解析的无缝衔接。借助多模态数据采集技术，能够实时捕捉课堂互动、作业轨迹等全方位的学习行为，形成连续动态的学习数据流，随后，通过知识图谱与情感计算，系统能够即时解析知识掌握曲线，定位认知断点，并将其通过可视化界面即时反馈到学生端与教师端，学生可借助错题分布云图、思维导图等工具，直观识别自身知识薄弱环节，主动调整学习路径；教师则能通过班级能力矩阵、课堂情绪热力图等全景视图，快速定位教学盲区，动态优化课堂节奏，这种双向反馈不仅缩短了“学-评-教”的响应周期。

在此过程中，教师需要注意依据学科目标和学情特征，主导诊断参数的设定。例如，前文提到的为不同的课型配置行为数据的权重，以此确保 AI 的算法逻辑与教育规律相契合。此外，教师还可借助结构化指令，在不断地人机交互中，引导 AI 生成与教学目标相匹配的诊断报告，从而将机器的数据扫描转化为具有教育意义的内容。这种人机协作的本质，是教育规律与智能技术的双向融合与相互促进，AI 拓展了教师诊断的精细程度，而教师则赋予数据以教育内涵，二者共同构建起可靠的决策依据。

（2）“F”反馈，精准个性化干预

AI 智能技术的核心价值在于实现“数据驱动改进”，将评价结果转化为具体可行的行动指南，通过实时采集数据、运用机器学习模型深度剖析学生学习模式，精准识别潜在风险并生成改进策略，同时推送个性化学习资源。在精准干预时，构建起覆盖学生、教师和院校的三维支持体系，以实现教育生态的全面优化。

在学生个体层面，智能体通过知识表征与行为分析的双重路径实现精准干预。一方面在知识维度，它利用知识图谱技术构建学科网络拓扑结构，将离散的知识点转化为可视化的网络结构，借助知识掌握度雷达图与错误热力图动态呈现学习状态的动态可视。这样既能揭示表层知识短板，又能通过关联网络定位深层认知断层，为学生定制个性化学习路径（见图 3），明确学习目标与方向，提升学习效率。该规划不仅指出知识缺口，更深入剖析知识间内在联系，引导学生从根源性理解缺失出发，规避机械重复，助力学生建构完整知识体系。

另一方面，AI 系统通过对学习行为多维度追踪机制，构建任务耗时、错题修正等关键指标分析模型，从而精准识别“碎片化突击”等低效模式，并据此实施差异化干预策略：在资

源内容上，依据认知轨迹动态推送错题强化集与拓展资料等实体资源，满足查漏补缺与拓展需求；在学习策略层面，基于行为特征生成个性化学习策略建议，如为拖延症学生设计分段式任务规划，以提高其学习效率；在智能答疑上，以学习通平台为例，AI 助教深度学习课程资源实现即时答疑，若 AI 助教无法回答，教师可同步在后台可查看相关数据并作答，系统还整合“豆包”智能体拓展答疑渠道，为深入了解学生的疑问与交互提供了新途径。该反馈机制强调策略适配性，依据学生的认知风格和行为特征提供差异化改进方案，实现精准化、有效的个性化指导，全面提升学生的学习效果。

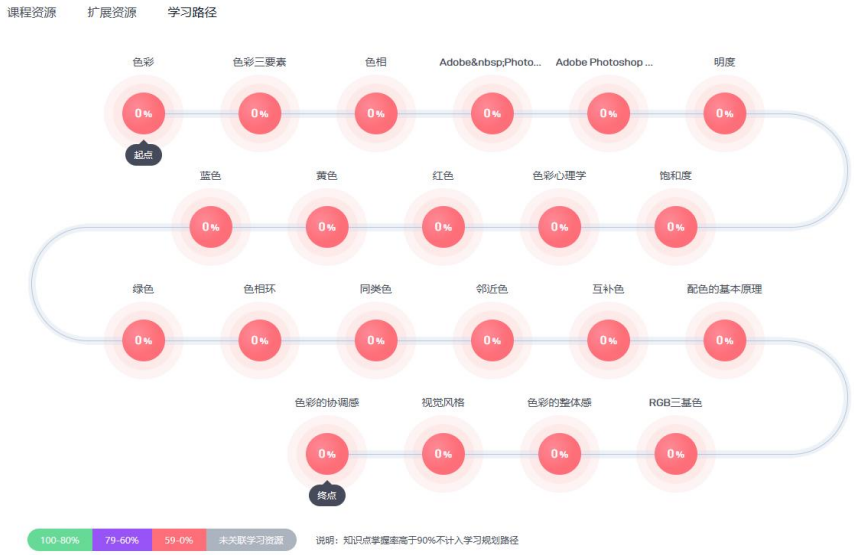


图 3 个性化学习路径生成

在教师层面，聚焦于动态学情监测与教学策略优化，它突破传统经验主义局限，实现数据驱动与专业判断的协同。AI 系统通过生成课堂参与度热力图、班级与个体知识掌握趋势曲线等可视化工具（见图 4），辅助教师快速定位班级共性难点，并为教师推荐相关资源、提供解决方案。同时，可以设计 AI 系统当检测到学生连续三次未参与在线讨论或学习进度严重落后时，系统自动触发预警机制，向教师推送分级干预建议。

图 4 班级情况可视化

在院校层面，AI 智能技术以教育生态系统整体优化为目标，深度整合多源课程评价数据。当监测到多门课程中学生批判性思维评分普遍偏低时，系统可自动关联课程大纲设计、教学

方法实施及师资培训记录，精准定位问题根源，进而生成课程改革建议，这种宏观干预突破单点优化的局限，推动教学改革从经验驱动转向证据驱动。

（3）“I”改进，迭代优化

在过程性评价的动态闭环中，迭代优化的关键在于构建技术驱动与教育规律深度融合的修正机制，形成反复多次的“反馈-修正”循环，实现教育干预策略的动态调适与持续改进。

AI 生成的干预策略应用于实际教学后，教师作为“教育校验者”，凭借自身的教学经验和专业素养，评估干预策略是否符合教育目标和育人理念。这并非简单的技术应用验证，而是将教育规律与技术手段的深度结合，教师通过观察学生学习行为、成果及情感态度变化变化，判断策略的有效性和适宜性。同时，教师作为“数据标注者”，将实践反馈结果以结构化的方式反馈给 AI，把学生的具体表现、问题及改进效果转化为语义化标签，帮助 AI 理解教育隐性逻辑，从而在后续的模型修正中，更好地贴合教育实际需求，推动系统自我优化。

学生作为反馈闭环的能动要素，收到反馈与干预策略后，可以根据学生档案袋记录的数据进行校验，并决定是否继续执行。在执行策略过程中，可通过嵌入式评价组件，如滑动式难度标尺等，实时记录主观体验，与策略执行效果数据双重校验，使得学生既能基于主观体验调整学习策略，又能将认知障碍或系统缺陷反馈给 AI，通过人机协同的迭代优化，使干预策略动态调整，推动学习支持服务向个性化、精准化发展。

4.3. 技术赋能与教育逻辑深度融合，避免表象化应用

（1）以教育目标为导向的技术适配机制

AI 智能技术的引入需以教育本质规律为核心，避免陷入“技术主导”的异化陷阱。具体而言，需构建“目标-技术-评价”的动态适配框架：首先，依据课程目标明确评价维度的优先级；其次，选择与之匹配的 AI 技术工具，若课程的核心目标为培养批判性思维，则需优先选择自然语言处理技术分析学生论辩逻辑的严密性，而非单纯依赖作业提交频次等表层指标，若课程侧重情感态度塑造，则需融合情感计算技术追踪课堂互动中的非言语信号，结合教师质性观察形成多维评估；最后，通过教师主导的校准机制，确保技术输出的评价结果与教育目标的一致性。

（2）教育公平与伦理化建设

教育公平需实现精准化保障与人机协同的伦理化建设。在精准化保障方面，应突破传统“一刀切”资源分配模式，转向基于动态需求识别的差异化支持体系，AI 系统可通过多维数据来构建弱势学生识别模型，并结合区域教育资源分布图谱，动态调整评价标准与干预策略。如针对网络条件受限地区学生延长在线任务提交窗口、增加离线学习行为权重赋值；针对特殊需求学生，采用如语音交互的自适应界面设计降低技术使用门槛，同时引入同伴互助数据交叉验证机制，避免单一算法指标导致评价偏差。在人机协同伦理化设计方面，需遵循“以人为本”原则，构建可解释、可干预、可追溯的评价系统，通过算法透明化设计消除“黑箱效应”，设置人工复核与申诉通道防止算法偏见威胁教育公平。

5. 结语

在高等教育评价体系深度转型的当下，AI 智能技术的引入不仅是对技术工具的简单应用，更是教育评价范式的系统性重构。研究通过剖析过程性评价的实践困境，揭示了传统模式中数据依赖、反馈滞后与主体缺位等问题的本质——其本质是工业化教育思维与技术赋能表层化之间的矛盾。通过构建“人机协同”的四维评价生态，将冰冷的算法逻辑转化为具有教育温度的动态支持系统，既突破了教师单一主体的经验局限，又避免了技术主导的评价异化。

并提出的“EFI”动态闭环模型，将评价、反馈与改进三个环节紧密衔接，形成了教学全周期的动态调节机制，实现评价反馈的有效利用，推动教育评价从静态结果判定向动态过程优化的范式转型，更通过人机协同的方式，确保了技术赋能与教育规律的深度融合，从而实现“以评促学”向“以评促发展”的跨越。研究还强调教育公平与伦理化建设在 AI 赋能教育评价中的重要性，通过构建可解释、可干预、可追溯的评价系统，我们努力消除“黑箱效应”，确保算法输出的公正性与可信度。同时，针对弱势学生的差异化支持体系与人机协同的伦理化设计，更是体现了教育的人文关怀与社会责任。

展望未来研究，一方面需要在现有理论模型基础上，进一步于真实教学情境中开展实际应用，以验证模型的可操作性与有效性，深度挖掘理论模型与教学实践的契合点；另一方面，要深入探索 AI 智能技术与教育规律的深层融合路径，不仅要突破算法模型可解释性与伦理设计的技术黑箱，构建师生可理解、可干预、可追溯的评价系统，还要构建覆盖教学全场景的多模态数据生态，借助动态知识图谱与自适应评价算法，将认知发展规律转化为可量化评价维度。只有将技术创新视为教育生态进化的催化剂而非替代品，才能真正实现过程性评价从工具理性向价值理性的跃迁，为高等教育高质量发展构建可持续的智慧治理范式。

参考文献

(2020).中共中央国务院印发《深化新时代教育评价改革总体方案》.河南教育(基教版),(11),4-9.

陈明松,李红斌,蔺永诚. 关于利用平时成绩激励大学生主动学习的思考[J]. 速读（下旬）,2019(8):96-97.

王战军 & 李涛旎.(2024).数智时代我国高等教育评估体系的转型与重构.大学教育科学,(02),106-117.

赵德成.(2013).教学中的形成性评价:是什么及如何推进.教育科学研究,(03),47-51.

张怀平,易洪刚,周明 & 唐少文.(2018).翻转课堂过程性评价指标体系的初步构建.中国高等医学教育,(10),28-29.

王英红.(2019).高校思想政治理论课过程性评价体系的构建.北京教育(德育),(10),93-96.

张璐,杨宇,徐茜 & 姚阳.(2024).基于雨课堂的过程性评价体系在生理学中的应用与研究.中国现代教育装备,(19),26-29.doi:10.13492/j.cnki.cmee.2024.19.039.

Zhen Yang.(2025).The Path to Construct the Process Evaluation System of College Sports Based on the Digital Age. Frontiers in Sport Research,7(1),

Lou Hong Li.(2020).Design of College English Process Evaluation System Based on Data Mining Technology and Internet of Things. International Journal of Data Warehousing and Mining (IJDWM),16(2),18-33.

李祥子,冯志君,陈结霞 & 周蓓.(2021).基于慕课线上教学的过程性评价体系的构建.药学教育,37(02),74-79.doi:10.16243/j.cnki.32-1352/g4.2021.02.018.

段芳.(2021).混合式教学下过程性评价体系的构建与实践(硕士学位论文,云南师范大学).硕士 https://link.cnki.net/doi/10.27459/d.cnki.gynfc.2021.001203doi:10.27459/d.cnki.gynfc.2021.001203.

田晓伟 & 彭小桂.(2020).在线教育服务行业资本化进程审思.教育发展研究,40(09),15-21.doi:10.14121/j.cnki.1008-3855.2020.09.005.

Pillay, P., & Balele, R. (2022). Exploring learners' experiences of receiving formative written assessment feedback in business studies as a subject in South Africa. International Journal of Learning, Teaching and Educational Research, 21(10), 228-248.

Huang, M., Liu, S., Zhang, Y., Cui, K., & Wen, Y. (2021). Research on the university intelligent learning analysis system based on AI. Journal of Intelligent & Fuzzy Systems, 1 - 10. https://doi.org/10.3233/jifs-189820

檀慧玲,李文燕 & 万兴睿.(2018).国际教育评价项目合作问题解决能力测评:指标框架、评价标准及技术分析.电化教育研究,39(09),123-128.doi:10.13811/j.cnki.eer.2018.09.017.

Guo, S., Latif, E., Zhou, Y., Huang, X., & Zhai, X. (2024). Using Generative AI and Multi-Agents to Provide Automatic Feedback. arXiv preprint arXiv:2411.07407.

面向小学生信息社会责任培养的体验式教育游戏框架设计与应用研究

Design and application of experiential educational game framework for primary school students' information social responsibility training

孙可欣，魏雪峰*，韩春，张嫣然，王严尉
鲁东大学教育学院
*xuefengwei99@163.com

【摘要】 随着数字技术的深度渗透，小学生作为数字原住民面临信息社会责任缺失的严峻挑战。信息社会责任作为信息技术四大核心素养之一，其培养对儿童价值观塑造与数字公民素养形成至关重要。然而，传统教育模式因内容滞后、方法单一及反馈缺失等问题，难以有效提升其信息辨识、隐私保护与道德决策能力。本研究构建了面向小学生信息社会责任培养的体验式教育游戏框架，包含“情境叙事-任务系统-交互反馈-元认知培养”四模块协同的动态学习系统，为信息社会责任教育提供了兼具理论价值与实践意义的创新方案，助力数字公民素养的系统化培养。

【关键词】 信息社会责任；教育游戏；游戏化教学；核心素养

Abstract: *With the deep penetration of digital technology, primary school students, as digital natives, face the severe challenge of the lack of responsibility in the information society. As one of the four core qualities of information technology, the cultivation of information social responsibility is crucial to the shaping of children's values and the formation of digital citizenship. However, the traditional education model is difficult to effectively improve its ability of information recognition, privacy protection and moral decision making due to problems such as delayed content, single method and lack of feedback. This study builds an experiential educational game framework for primary school students' information social responsibility training, including a four-module collaborative dynamic learning system of "situational narrative - task system - interactive feedback - metacognitive training", which provides an innovative scheme with both theoretical value and practical significance for information social responsibility education and helps the systematic cultivation of digital citizenship literacy.*

Keywords: Information society responsibility, Educational games, Gamification teaching, Core quality

基金项目：山东省人文社会科学课题（智库重点项目）“人工智能支持下的小学个性化作业设计与应用研究”（编号：2023-zkzd-014）

1.引言

信息社会责任作为信息技术四大核心素养之一，是数字公民素养培育的核心内容，其教育目标涵盖信息伦理认知、法律遵从与行为自律等多维度能力。对于正在形成世界观、价值观的青少年而言，他们既是信息社会发展的受益者，也是潜在的风险承担者，小学阶段更是价值观形成的关键期（Wood et al, 1976）。然而，当前传统信息科技课堂对小学生信息社会责任的培养面临三重困境：内容上偏重传统伦理而忽视数字公民责任，未能覆盖《义务教育

信息科技课程标准（2022 年版）》提出的核心素养要求；方法上依赖单向说教导致低龄学生共情缺失，抽象准则难以转化为具体行为决策；反馈上缺乏行为后果的具象化呈现，学生无法感知自身信息行为的社会影响。

教育游戏凭借其沉浸式交互与多模态反馈特性，为破解上述难题提供了新路径。然而，现有研究多聚焦技术工具开发，鲜有将信息社会责任的核心维度系统融入游戏化学习框架。为此，本研究提出以下核心问题：如何构建适配小学生认知特征的体验式教育游戏框架？如何通过游戏化设计实现信息社会责任从抽象概念到行为内化的有效转化？基于此，本研究整合具身认知理论与体验式学习模型，设计四模块协同的动态教育游戏框架，并开发“数字公民守护者”游戏进行实证验证。研究旨在为小学生信息社会责任教育提供理论支撑与实践范式，推动数字时代公民素养教育的范式革新。

2.文献综述

2.1. 信息社会责任的内涵及重要性

信息社会责任作为信息技术核心素养的关键组成，近年来受到了国内外学者的高度关注。它被定义为个体在信息社会中，对于文化修养、道德规范和行为自律等方面的职责担当，是信息技术学科教育不可或缺的一部分。在小学生世界观与价值观形成的关键期，信息社会责任教育具有不可替代的奠基作用。这一阶段儿童的认知可塑性强，对外界信息的吸收与模仿近乎本能，而数字原住民的身份使其更早暴露于复杂的网络环境中。若缺乏系统引导，错误的信息行为可能固化为认知偏差，甚至扭曲其对真实社会的理解。研究表明，小学阶段是儿童道德推理能力飞跃期（Wood et al., 1976），此时通过具象化体验建立“信息行为-社会影响”的因果逻辑，能有效塑造责任意识的内核，例如理解“一条虚假信息的传播可能引发群体恐慌”等社会连锁反应。然而，传统课程在此领域面临三重困境：其一，内容滞后性，德育课程多聚焦传统伦理范畴，对数字公民责任覆盖不足；其二，方法单向性，以说教为主的课堂模式难以激发低龄学生的共情，抽象的道德准则无法转化为具体行为决策；其三，反馈缺失性，缺乏真实情境中的行为矫正机制，学生难以感知自身信息行为的社会后果。例如，传统教学中关于“不传谣”的教导，往往止步于口号式警示，而未能让学生通过仿真场景体验谣言传播的扩散路径与危害强度。

研究显示，信息社会责任不仅仅是技术操作层面的要求，更深层次触及到个体的价值观、伦理观和行为规范。梅森（R.O.Mason, 1989）最早指出，伴随计算机使用的普及，信息伦理问题逐渐浮现，成为信息伦理概念产生的直接原因。国内学者如吕耀怀与沙勇忠（2002）则从道德层面对信息伦理进行了探讨，认为它是调整人际关系和社会关系的法则集合。近年来，左明章（2002）等学者呼吁加强信息伦理教育，尤其是在中小学校园内的急迫性。

2.2. 信息社会责任的维度划分

在国际视野下，各国家和地区的信息科技课程纷纷纳入了信息社会责任教育。英国教育部早在 2013 年的计算机学习课程中，确立了负责任、审慎和创造性地使用信息技术为目标。日本的教育体系中，则着重于“信息利用能力”的培养，强调适应信息社会发展的必要技能与态度（MESCSCT, 2022）。表 1 展示了各国中小学信息科技相关课程中关于信息社会责任的具体内容。

表 1 部分国家信息技术课程中涉及信息社会责任的内容

国家	中国	美国	韩国	俄罗斯	日本	英国

信息社会责任内容	文化修养、行为自律、道德规范等方面	遵守网络文明规范；保护隐私；预防计动；算机犯罪；尊重知识产权，负责任地使用信息。	了解个人信息与版权保护的重要性并付诸行安全、了解遵守信息道德的和法律	考虑信息社会信息传递的影响，尊重他人个人信息的权利，对信息社会道德的行为负责，正确和安全使用信息；信息问题解决能力。	安全、尊重地使用技术，保护个人信息，知如何寻求技术帮助。

在国内，学者们围绕信息社会责任的内涵、重要性和教育实践展开了深入探讨。刘哲雨等（2023）强调，信息社会责任是信息时代下个人成长与发展的基石，涉及法律遵从、道德实践及自我约束等多方面内容。任友群（2016）认为，信息社会责任包含合理运用信息技术、遵守信息社会法规、维护信息安全和个人权益，以及尊重多元文化在内的多维度责任。解月光（2017）进一步细化了信息社会责任的构成，将其分为意识、能力、法律与道德四个层面。钱松岭则依据《义务教育信息科技课程标准（2022 年版）》（以下简称《新课标》），解析了信息社会责任在教育中的具体目标和实践策略，倡导在信息活动中融入道德考量与法律遵循，具体内容如表 2 所示。

表 2 信息社会责任维度分析

研究者	信息社会责任构成要素分析
解月光	信息安全意识与能力、遵守法律法规、信息道德与伦理。
钱松岭	信息社会责任意识、能力、行为、制度、文化。
任友群	信息社会责任行为、信息社会责任文化、信息社会道德准则。
杨晓哲、刘昕	信息安全、问题解决能力、网络道德。
吴砥	信息安全意识与能力、遵守信息法律法规、良好的信息道德与伦理。

通过梳理国内学者分析的信息社会责任素养要素，对各研究中信息社会责任素养进行关键词提取和整理，表呈现了整理的信息社会责任素养构成要素情况。可以看到，目前关于信息社会责任素养的构成要素尚未有统一的标准，但大多贴合《新课标》中强调的文化修养、行为自律以及道德准则，涵盖了意识、知识、能力、思维这四个关键要素。结合《新课标》界定划分，信息社会责任可以分为信息社会责任意识、知识、能力以及道德四个构成要素。

2.3. 教育游戏内涵及发展

教育游戏(Educational Game)是兼具趣味性与教育意义的计算机软件系统，一直是教育技术学研究热点。它将知识与游戏结合，为学习者提供真实问题解决场景，促进知识学习，培养关键能力，激发内在动机。其研究起源于国外，早期研究者期望把电视游戏激发动机的方法用于教育以提升学习效果（魏小东，2023）。随着技术发展与“娱教”思想传播，它受到更广泛关注。因研究角度和层次不同，其定义与内涵也有差异，如尚俊杰（2015）认为它是具教育和娱乐功能的计算机程序，《教育游戏产业研究报告》则定义其为能培养使用者多种能力且有教育意义的计算机游戏软件。

国外教育游戏历史可追溯到 20 世纪 70 年代末至 80 年代初，当时计算机辅助教学软件出现，如美国的 PLATO 系统用于远程教学。之后，随着个人电脑普及，《俄勒冈之旅》《数

学怪兽》等经典教育游戏通过游戏化教授知识（张敬，2021）。中国教育游戏发展起步晚但发展迅猛。20 世纪 90 年代末，互联网普及促使教育软件公司尝试开发教育游戏，因技术和市场限制，普及率不高。2000 年后，智能手机和平板电脑普及，教育游戏市场爆发，《宝宝巴士》《洪恩教育》等受家长欢迎。近年来，国家重视教育信息化，“三通两平台”建设提供支持，中国开发者也关注国际市场，推出如《多邻国》中文版等产品。

2.4. 游戏化培养小学生信息社会责任

近年来，国内外学者围绕教育游戏在培养小学生信息社会责任领域的理论与实践展开深入研究，形成了差异化探索路径与共性化认知框架。国外研究以技术整合与行为引导为双轮驱动，欧美国家依托《Minecraft: Education Edition》等沉浸式游戏创设虚拟协作场景，使学生在模拟数据泄露、网络社交等情境中建立信息伦理意识，日本学者则创新性融合动漫角色扮演机制，通过设置网络欺凌处置、谣言溯源等任务提升儿童道德判断能力，实证表明此类设计可使学生的安全行为正确率提升超 30%（Prensky, 2006）。

国内研究侧重本土化理论构建与实践创新，张敏强团队提出“情境-任务”耦合模型，将信息责任细化为可操作的闯关机制，如《网络安全小卫士》通过模拟网络诈骗应对训练显著增强低年级学生的信息甄别能力，而王磊团队开发的游戏化反馈系统证实即时奖惩机制对道德认知内化具有强化效应，研究揭示游戏化学习在核心素养与道德认知的实证层面通过情感代入机制有效促进道德情感迁移。尽管教育游戏在信息社会责任培养领域展现出潜力，但如何将严肃的信息社会责任内容转化为有趣且吸引小学生的游戏情节，并通过教育游戏的故事情节和角色互动自然传达信息社会责任的学习内容仍缺少深入研究。

3.面向信息社会责任培养的体验式教育游戏框架设计

面向信息社会责任培养的体验式教育游戏需要针对信息社会责任的四个构成要素：意识、知识、能力、道德，为学习者提供信息社会责任学习的体验环境，将信息社会责任的抽象概念转化为具体的游戏情节和互动元素。因此，本研究基于体验式游戏模型、教育游戏系统结构和信息社会责任内容框架，设计了面向信息社会责任培养的体验式教育游戏框架。

3.1. 体验式游戏模型

面向信息社会责任培养的体验式教育游戏，需让学习者通过游戏体验学习到信息社会责任的相关知识。具身认知理论阐释了通过身体与环境交互形成认知的学习过程，而体验式学习理论阐释了由体验构建概念认知的循环过程，通过身体与环境的互动、直接的经验参与和反思循环，帮助小学生在游戏中具体化、体验式地学习信息社会责任，弥补传统课程在内容和方法上的不足，为面向信息社会责任培养的体验式教育游戏设计提供了理论参考。

3.1.1. 理论基础

①具身认知理论。具身体验理论强调，认知通过身体与环境的动态互动形成。对小学生而言，抽象的道德准则需转化为可感知的具象行为才能被有效吸收。例如，在游戏中某些设计，学生需通过交互操作直观感受一条谣言在虚拟社区中的扩散速度和影响范围，身体动作与信息扩散的视觉反馈同步，将抽象责任具象为“操作-结果”的直接关联。

②体验式学习理论。体验式学习是指通过实践活动获得直接经验,并对活动进行观察、反思和心理代入，从而获得知识和技能的学习过程。Killi（2005）基于体验式学习理论提出的体验式游戏模型将学习过程分为挑战、主动实验、观察反思、概念建构四个部分，这四个部分构成了一个体验式学习循环。

3.1.2. 模型设计

本研究以具身认知理论与体验式学习的交互作用为理论基础，构建了动态演进的具身化体验式游戏模型。其设计框架通过四阶闭环路径实现认知深化：首先，游戏化情境通过具身化任务触发学习者的初始体验；接着，在游戏机制驱动下，学习者借助触觉交互、视觉反馈等多模态行为进行主动探索，同步接收环境对操作后果的即时响应；随后，系统通过整合感知输入与动觉经验，将抽象责任转化为“行为-结果”的具象认知映射；最终，依托结构化反思工具，引导学习者将内化的道德准则迁移至新情境，并基于行为数据分析动态调整后续挑战难度，形成“情境激活-交互验证-认知重构-行为优化”的螺旋式学习循环。

3.2. 教育游戏系统结构

体验式教育游戏需通过系统性架构设计与适配性交互技术，实现学习者身体行为向游戏情境的有机延展。基于 Winn（2009）提出的教育严肃游戏“设计-参与者-体验”（Design-Participant-Experience, DPE）三元模型，本研究从教育目标导向出发，对系统架构进行多维度解构：在内容层植入伦理冲突任务以激发道德反思，在叙事层嵌入具身化脚本增强情境代入感，在交互层设计多感官通道强化行为映射，在体验层构建情感共鸣回路促进价值内化。基于此逻辑，本研究将教育游戏系统结构划分为以下四个模块，形成完整的“情境激活-行为实践-反馈矫正-反思升华”认知发展链：

①情境叙事模块。该模块通过动态场景生成技术（Dynamic Scenario Generation, DSG）与自然语言处理（Natural Language Processing, NLP）算法，构建符合学习者认知水平与行为特征的游戏场景。具体而言，系统基于玩家历史行为数据，实时生成定制化道德困境事件。例如，针对低年级学生，系统生成“未经允许转发同学照片”的社交冲突事件；面向高年级学生，则设计“网络谣言传播溯源”的复杂任务。NLP 技术通过语义分析与情感计算，动态调整 NPC 对话文本的情感强度，实现“情感-玩家-叙事”的智能匹配。玩家在具身化场景中通过角色扮演直接参与事件解决，从而在沉浸体验中建立“信息行为-社会影响”的因果逻辑框架。

②任务系统模块。任务系统采用层次化动态分配机制（Hierarchical Dynamic Task Allocation, HDTA），将信息社会责任分解为意识、知识、能力、道德四维度目标，并映射为渐进式游戏任务。技术实现上，系统通过认知诊断模型（Cognitive Diagnostic Model, CDM）分析玩家操作数据，动态调整任务难度：

初级任务-聚焦意识唤醒，如识别社交媒体中的隐私泄露风险点；

中级任务-强化知识建构，如对比不同信息来源的可信度评分规则；

高级任务-训练复杂决策能力，如模拟网络欺凌事件的跨角色处置策略。

任务链设计遵循“最近发展区”理论（Vygotsky, 1978），确保每个关卡挑战略高于玩家当前水平，通过“脚手架”机制（如线索提示、协作 NPC）支持能力跃迁。

③交互反馈模块。该模块整合物理引擎、触觉反馈设备与跨模态感知技术，构建“行为-结果”即时映射机制。例如，在谣言验证任务中设置视觉反馈：真实信息拖拽至核查站时呈现绿色光轨，虚假信息触发红色警示波纹；听觉反馈：操作成功时播放轻快音效，超时未处理则叠加逐渐增强的警报声。

多模态反馈不仅强化操作记忆，更通过具身化感知将抽象责任转化为可量化的物理体验。系统同步生成行为热力图与决策轨迹报告，为后续元认知反思提供数据支撑。

④元认知培养模块。该模块通过双重路径促进自我监控与价值升华：

认知可视化工具：嵌入思维导图生成器，引导玩家复盘决策逻辑。例如，在隐私保护任务后，系统自动生成“信息共享风险评估图”，标注关键决策节点的影响权重；

自适应策略建议：基于机器学习算法分析玩家行为模式，推送个性化认知策略。

进阶阶段引入“道德影响力指数”，量化玩家行为对虚拟社区的整体影响，并通过跨情境任务实现责任意识的场景迁移。

四大模块形成“情境激活-行为实践-反馈矫正-反思升华”的完整链条：情境叙事提供道德认知的初始框架；任务系统将认知转化为可操作行为；互反馈通过即时强化修正行为偏差；元认知模块最终完成从“知道”到“理解”再到“自主践行”的价值跃迁，通过“身体参与-情感共鸣-认知迭代”的三维作用机制，使信息社会责任成为可感知、可操作、可持续的内在准则。

3.3. 信息社会责任内容框架

本设计通过“内容-技术”双轨耦合机制，实现信息社会责任教育的系统性转化，其核心逻辑在于将认知发展规律与技术赋能路径精准匹配。基于对信息社会责任维度的划分，内容框架包含以下四个方面：

①意识培养，指学习者的信息社会敏感性与主动认知。游戏系统利用空间建模，实现环境的空间叙事，强化“我在情境中”的在场感，通过环境沉浸触发信息敏感度。该模块聚焦学习者的信息社会敏感性培育与主动认知机制建立。技术实现层面，游戏系统采用 3D 空间建模技术构建虚实融合的沉浸式场域，通过动态光影渲染、空间声场定位等技术实现环境的空间叙事功能。

②知识建构，指支撑信息社会责任的结构化知识体系。游戏系统基于学习者认知水平，动态生成匹配的知识结构树，实现对跳转场景的难度调配。此模块在技术实现上，采用知识图谱嵌入技术生成个性化知识结构树，结合强化学习算法实现场景跳转的难度调控，形成“认知诊断-路径优化-难度适配”的智能调节闭环。

③能力训练，指应对数字社会挑战的方案设计与挑战能力。游戏系统通过物理引擎模拟真实操作反馈，训练应急响应能力，通过人机交互建立“操作-反馈”的强关联。特别开发突发事件响应沙盒，运用离散事件仿真技术模拟数字社会危机场景，通过蒙特卡洛方法生成多样化训练方案，建立"情境模拟-决策实施-效果评估"的能力锻造链条。

④道德内化，指稳定的数字社会价值观体系。游戏系统通过多感官反馈强化价值认知，结合认知神经科学原理，多感官刺激提升道德准则的内化。该模块运用认知重评技术设计道德困境情境，动态调整多感官刺激强度，形成"神经激活-情感共鸣-价值内化"的递进式道德养成路径。

综合以上研究，本研究设计了面向信息社会责任培养的体验式教育游戏框架，如图 1 所示。该模型从信息社会责任内容框架出发，先提取信息社会责任的核心学习内容；之后采用空间建模、动态生成、游戏引擎、跨模态反馈等技术方案，将学习内容转化为教育游戏系统结构的各要素，形成情境叙事模块、任务系统模块、交互反馈模块、元认知培养模块；最后按照体验式游戏模型，创设游戏情境、提供具身体验，通过交互与反馈、意识形成、知识建构等环节，让学习者实现对信息社会责任的道德内化和能力提升。

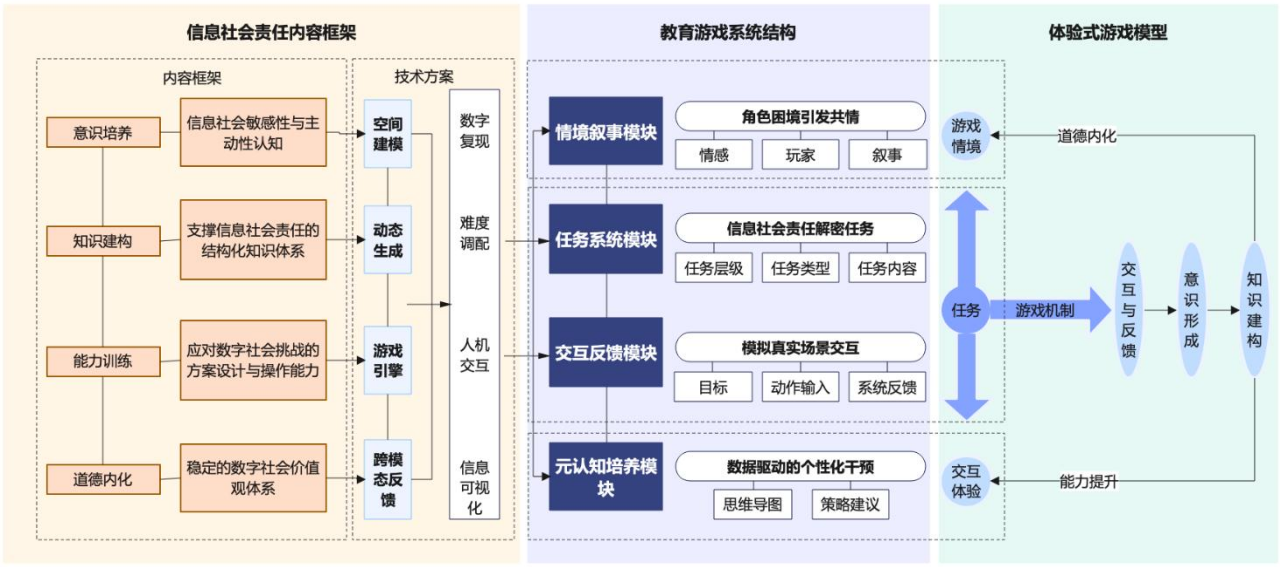


图 1 面向信息社会责任培养的体验式教育游戏框架

4.面向信息社会责任培养的体验式教育游戏框架应用案例

本研究依托面向信息社会责任培养的教育游戏框架，开发了一款符合小学生认知水平的教育游戏“数字公民守护者”。在一个名为“智信岛”的虚拟社区中，玩家化身为新晋的“数字守护者”，负责维护社区的信息安全与伦理秩序。通过多维度任务链，玩家需逐步解锁隐私保护、谣言识别、网络干预等技能，最终成为社区信任的领导者。参考教育游戏系统结构,“数字公民守护者”的内容模块主要包括:

4.1. 情境叙事模块

玩家初始进入“智信岛”中央广场时，场景设计通过高饱和度色彩与轻快背景音乐营造安全氛围。突发剧情中，社区大屏幕弹出“特效药治百病”谣言警报，非玩家角色（Non-Player Character, NPC）居民行为模式即时切换为恐慌状态：部分 NPC 手持虚拟货币冲向药店，导致场景内出现拥挤踩踏事件。视觉上，药店招牌闪烁红光，人群密度热力图叠加显示；听觉上，警报声与 NPC 焦虑对话交织，形成多感官刺激。这一设计通过具象化场景冲突，触发玩家对谣言危害的敏感性认知，同时强化“信息行为直接影响社会秩序”的逻辑关联。

4.2. 任务系统模块

针对“谣言溯源”任务，系统基于动态生成技术（Dynamic Content Generation, DCG）创建三类线索源：初级任务提供社交媒体截图，关键词如“100%治愈率”被高亮标注，引导新手玩家聚焦核心矛盾；中级任务引入医院官方辟谣声明，要求玩家对比时间戳与署名真实性；高级任务则开放药品成分数据库，玩家需结合化学知识图谱验证成分有效性。任务难度动态适配机制通过分析玩家历史操作数据（如任务完成时间、错误率）调整线索复杂度：进阶玩家需自主构建证据链，例如将社交媒体账号可信度评分与药品数据库交叉验证。此过程模拟真实信息溯源场景，帮助玩家掌握“交叉验证-逻辑推理-结论生成”的结构化方法，实现从碎片化知识到系统性能力的转化。

4.3. 交互反馈模块

玩家需将疑似谣言内容拖拽至“核查站”进行验证，游戏引擎通过物理模拟技术赋予不同信息类型差异化阻力：真实谣言（如权威机构发布）阻力最小，可单手快速滑动至目标区域；

模糊信息（如未标注来源的截图）需双手持续施力；完全虚假信息则触发最大阻力，要求玩家多次尝试或寻求队友协作。操作成功后，屏幕绽放绿色净化光波，伴随 NPC 欢呼音效；若超时未完成，谣言扩散范围以红色波纹动态蔓延，背景音乐切换为低沉急促的旋律。这种多模态反馈机制（视觉、听觉、触觉）不仅强化“快速响应-精准处置”的操作记忆，还通过即时行为后果映射，训练玩家在高压情境下的应急决策能力。

4.4. 元认知培养模块

任务尾声设置道德选择环节：玩家可选择立即公开辟谣，或观察谣言扩散规律后再行动。选择“观察”时，系统通过跨模态反馈技术呈现谣言传播树状图，显示 24 小时后 70%居民被误导，同时触觉腰带持续低频震动模拟焦虑感；选择“辟谣”则触发 NPC 语音感谢，社区建筑恢复明亮色调，并解锁“信任度+50”成就勋章。此外，界面实时显示“社会影响值”仪表盘，例如“公共安全+50”与“信任度-20”的对比数据，直观量化不同决策的伦理代价。通过将抽象道德准则转化为可感知的数值与情感反馈，玩家得以在反思中内化“及时干预”的责任意识，并逐步形成稳定的数字社会价值观。“数字公民守护者”教育游戏对应效果图如图 2 所示。

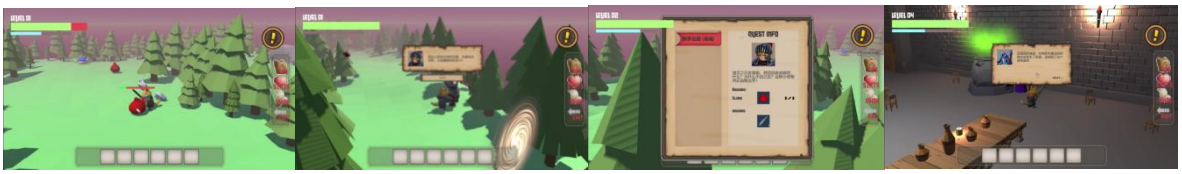


图 2“数字公民守护者”教育游戏效果图

面向信息社会责任培养的体验式教育游戏评价

为验证面向信息社会责任的体验式教育游戏模型的有效性，本研究采用包括调查问卷、学习效果客观测试和半结构化访谈的混合研究方法，从学习效果、学习体验、学习动机三个维度来评估“数字公民守护者”教育游戏应用的学习质量。目前，文章中的研究正在 L 市 L 小学利用常规信息科技课实施，由同一信息科技教师教授的 2 个五年级平行班级里应用，开展为期一学期的准实验研究。实验开始前，对两班学生进行了信息社会责任水平前测，实验班还完成了学科兴趣前测。随后，实验班和对照班分别通过参与“数字公民守护者”教育游戏体验学习和常规多媒体体验学习。教学结束后，实验班学生填写了信息社会责任水平后测、情境兴趣以及认知负荷调查问卷，并接受了学习体验访谈。

数据收集和分析是准实验过程中最重要的阶段，必须检查数据的来源、确保收集的数据是规范的、正确的和完整的，才能帮助我们理解实验过程并确定我们预期的结果。由于文章还在进行准实验研究，正处于收集过程性数据到产生实验结果的过渡阶段，数据分析及其实验结果将不在此呈现，后续实验会持续推进。

结语

本研究聚焦于信息社会责任教育在小学生群体中的实践困境，提出并设计了面向信息社会责任培养的体验式教育游戏框架，通过整合具身认知理论与体验式学习模型，构建了四模块协同作用的动态学习系统。基于此框架开发的“数字公民守护者”教育游戏，通过虚拟情境中的具身化任务设计，将信息社会责任的抽象概念转化为可感知、可操作的交互体验，实现了意识触发、知识建构、能力训练与道德内化的多维度培养目标。实证研究初步表明，该框架能够有效提升小学生的信息敏感度、谣言辨识能力及道德决策水平，尤其在“行为-结果”即时反馈机制与多感官刺激的协同作用下，学习者对信息社会责任的认知内化程度显著增强。然而，本研究仍存在一定局限性：其一，实验样本规模较小且集中于单一区域，可能影响结

论的普适性；其二，游戏设计的长期效果及跨学段适用性尚未充分验证；其三，技术实现层面对硬件设备依赖较高，可能限制其推广性。

未来研究可从以下方向深入探索：一是扩大实验范围，开展跨区域、多学段的纵向追踪研究，验证教育游戏的长效性与泛化能力；二是引入人工智能技术优化动态生成机制，提升游戏情节的个性化适配水平；三是探索家校社协同模式，将虚拟情境中的责任认知迁移至真实生活场景，构建全链条培养路径。此外，如何平衡游戏的娱乐性与教育性，进一步激发低龄学习者的内生动机，仍需持续迭代与创新。

参考文献

马璇.(2022).小学生信息社会责任培养课程模式的研究——以 A 小学为例.中国电化教育(11),129-134.

任友群, 李锋 & 王吉庆. (2016). 面向核心素养的信息技术课程设计与开发. 课程.教材.教法(07), 56-61+9.

左明章 & 许雄.(2002).中小学信息道德教育的目标与内容的探讨.中国电化教育(05),23-25.

邓鹏.(2020).面向高阶认知发展的成长式问题化学习（GPBL）研究——概念、设计与案例.远程教育杂志(03),76-85.

刘兴波,邱岑岑 & 王广新.(2020).任务强度差异支架对学生游戏化学习动机和沉浸感的影响.中国远程教育(04),50-57+77.

刘哲雨,刘佳乐 & 高建庆.(2023).如何提升高中生的信息社会责任？.现代教育技术(09),56-66.

张羽洋, 李万葶, 罗雅兰 & 聂晓梅. (2022). 基于体验式学习模型的文化遗产教育游戏——以“梦回南唐宴”游戏为例. 现代教育技术, 32 (07), 48-56.

张露,胡明玉 & 尚俊杰.(2020).游戏化学习体验的质性分析研究.中国远程教育(03),35-41+80-81.

尚俊杰 & 裴蕾丝.(2015).重塑学习方式:游戏的核心教育价值及应用前景.中国电化教育(05),41-49.

解月光,杨鑫 & 付海东.(2017).高中学生信息技术学科核心素养的描述与分级.中国电化教育(05),8-14.

魏小东 & 王昱.(2023).基于交互叙事模式的教育游戏设计策略研究.电化教育研究(09),108-113+121.

Kiili K. Digital game-based learning: Towards an experiential gaming model[J]. The Internet and Higher Education, 2005,(1):13-24.

Prensky, M..Computer Games and Learning:Digital Game-based Learning[J]. Studies in Computational Intelligence,2006,(5):156-158.

Winn B M. The design, play, and experience framework[A]. Handbook of Research on Effective Electronic Gaming in Education[C]. IGI Global, 2009:1010-1024.

Wood D, Bruner J S,Ross G. The role of tutoring in problem solving[J]. Child Psychology & Psychiatry & Allied Disciplines, 1976, 15: 264-293.

生成式人工智能赋能课堂教学的模式和实践路径——基于扎根理论的研究

The Mode and Practice Path of Generative Artificial Intelligence Empowerment Classroom Teaching —— Based on Grounded Theory

张晓晴, 徐诗篇, 景玉慧
浙江工业大学教育学院
* 19857166552@163.com

【摘要】 生成式人工智能（Generative AI，GAI）在课堂教学中的应用潜力与实践路径日益受到关注。针对相关教育政策与教学案例分析视角零散及系统模型缺乏问题，本研究采用扎根理论，构建了“学习空间-教学活动-角色关系”三维应用模型。首先，通过梳理政策文本与典型案例，明确GAI赋能课堂教学的内涵、价值及特征，揭示其在教学结构重塑中的核心作用。其次，以扎根理论的开放编码、轴心编码与选择性编码流程为方法论依据，深刻分析GAI如何促进教学目标精准化、内容多样化及个性化学习实现，及教师由知识传授者转变为学习引导者的角色转化。最后，结合65个“人工智能+教育”典型案例，详释“课前准备-课中交互-课后服务”的循环实践路径，提出具体的教学流程优化与工具整合策略。研究旨在为GAI课堂应用提供系统的理论支持与实践框架，同时指出GAI在不同教育情境下的适应性与局限性，强调未来研究应深化相关领域，以推动“人工智能+教育”行动的持续推进与教育改革创新。

【关键词】 生成式人工智能；课堂教学；扎根理论；教学模式创新

Abstract: *The application potential and practical pathways of Generative Artificial Intelligence (GAI) in classroom teaching have received increasing attention. To address the issues of fragmented perspectives in analyzing related educational policies and teaching cases, as well as the lack of systematic models, this study adopts grounded theory to construct a three-dimensional application model comprising "learning space - teaching activities - role relationships."First, by reviewing policy documents and exemplary cases, the study clarifies the connotation, value, and characteristics of GAI-empowered classroom teaching, highlighting its core role in reshaping instructional structures. Second, based on the grounded theory processes of open coding, axial coding, and selective coding, it deeply analyzes how GAI facilitates precise teaching objectives, diversified content, and personalized learning, as well as the transformation of teachers' roles from knowledge transmitters to learning guides.Finally, by integrating 65 typical "AI + education" cases, the research elaborates a cyclical practical pathway of "pre-class preparation - in-class interaction - post-class service," proposing concrete strategies for teaching process optimization and tool integration. The study aims to provide systematic theoretical support and practical frameworks for GAI applications in classrooms, while also pointing out its adaptability and limitations across different educational contexts. Future research is encouraged to deepen these areas to further advance the sustainable implementation of the "AI + education" initiative and education reform innovation.*

Keywords: Generative artificial intelligence,Classroom teaching,Grounded theory,Innovation of teaching mode

前言

生成式人工智能（GAI，Generative Artificial Intelligence）以其强大的用户需求导向能力和对话功能，正在成为推动全球教育领域数字化教学创新的重要技术支持。英国教育部发布的《生成式人工智能在教育中的应用》、美国高等教育信息化协会的《2024 年人工智能图景研究》、联合国教科文组织的《生成式人工智能教育和研究应用全球指南》均明确指出，GAI 在教育的应用是当前国际教育改革的关键方向。这些政策文件系统性阐释了 GAI 技术在支持个性化学习、提升教育质量以及应对教育公平问题中的潜在价值。此外，中国 2024 年《政府工作报告》首次提出“人工智能+”行动，明确要求人工智能技术深度融入教育教学全过程。

然而，当前学界对 GAI 赋能课堂教学的研究多集中于宏观理论探讨，缺乏实践结合。已有研究指出，GAI 使教学模式从“师-生”二元结构向“师-机-生”三元结构转变，推动教育范式从“以教为中心”向“以学为中心”转型，但技术应用尚未成熟(Kerimbayev et al., 2024)。国外高校虽在教学中积极探索 GAI 工具的使用规范，我国本土化研究与实践仍不足，可能导致理论与实践脱节，影响教育改革推进。因此，本研究基于扎根理论，分析国内典型教学案例，挖掘 GAI 赋能课堂教学的模式与路径，旨在为我国“人工智能+”行动提供理论支持和实践参考。

文献综述

生成式人工智能（Generative AI, GAI）技术的迅速发展为教育应用提供了坚实的技术支撑。以大型语言模型（LLM）为核心，GAI 通过自然语言处理与深度学习实现文本、图像和多模态内容的高质量生成（Du et al., 2021；Salem & Ramadan, 2020；Wang et al., 2025），应用于科学、医学及语言教学等多领域（Cooper, 2023；Friederichs et al., 2023；Crompton et al., 2024）。教育数字化转型推动了 GAI 赋能个性化学习路径设计，减轻教师负担，提高教学效率，促进教学公平（Li et al., 2023；Bower et al., 2024；Yusuf et al., 2024）。GAI 在课程设计、实时答疑、个性化评估等环节中，推动教师角色从知识传授者向学习引导者转变（Ding et al., 2023；Lan & Chen, 2024）。同时，技术应用带来了认知偏差、批判性思维弱化和数据隐私等挑战（Tang & Cooper, 2024；Yusuf et al., 2024）。现有研究多聚焦技术功能与场景适配，缺乏基于中国教育生态的本土化系统性研究（Chiu, 2023）。基于此，本研究采用扎根理论，围绕以下核心问题展开：（1）GAI 如何通过构建智能化、多样化且互动性的学习空间，提升课堂教学质量？（2）生成式人工智能赋能下的教学活动实践路径如何设计与实施，以促进学生的个性化学习发展？（3）GAI 应用如何重构教学过程中的角色关系？研究以 65 个典型“人工智能+教育”案例为基础，力图填补本土化实践研究空白，为推动中国“人工智能+教育”政策的落实与教育改革创新提供理论支持和实践指导。

3. 研究过程

3.1. 研究对象

本研究从教育部公布的首批 18 个“人工智能+高等教育”应用场景典型案例和第二批 32 个“人工智能+高等教育”应用场景典型案例当中筛选出 31 个课堂教学的案例；同时，为使案例足够全面与丰富，研究者又以 GAI 工具为检索抓手，通过网站、网页等方式采集到 34 个案例。最终累计获得 65 份案例，以此作为研究对象。

3.2. 开放式编码

开放式编码是将原始资料打散，赋予其相应的概念，然后对概念进行比较、整合形成范畴的过程。经开放式编码，研究共抽象出 48 个初始概念，对其进一步范畴化，聚拢为 11 个范畴，分别为智能辅助工具、个性化学习资源、智慧学习环节、教学目标、教学内容、教学方法、教学情境、教学评价、师生角色重构、同伴互动、自主发展，具体如表 1 所示。

表 1 开放式编码结果

序号	范畴化	初始编码
1	智能辅助工具	内容生成工具/评估和测试工具/互动式工具/语言学习辅助工具
2	个性化学习资源	智能辅导/个性化学习资料/多媒体资源
3	智慧学习环境	在线教育平台/社区平台/智慧实验室/智慧教室
4	精准教学目标	技术可用性评估/素养能力培养/个性化学习实现
5	多样化教学内容	理论课程/阅读写作课程/编程课程/实践课程/实验教学/设计课程
6	创新教学方法	泛在式学习/探究式学习/翻转课堂/讲授法/讲练结合
7	智能学习情境	文学人物对话/游戏角色交互/历史场景模拟/政治事件重现/虚拟化学实验/虚拟医学实验
8	智能评估反馈	学生状态评估/个性化学习路径生成/学习者特征分析/学习过程评估/实时反馈/学习成果展示/智能推送数据报告/多维度画像/学习体验
9	师生角色重构	角色转变/师生互动/教学质量/个性化指导
10	同伴互动	合作交流/同伴依赖
11	自主发展	学习内容理解/自我认知/自我效能感/自我能力发展

3.3. 主轴编码

主轴编码是在开放式编码的基础上，通过对范畴不断比较分析、归纳提炼形成主范畴的过程。例如，“智能辅助工具”“个性化学习资源”“智慧学习环境”可以归纳到更高级别的“学习空间”这一主范畴，“教学目标”“教学内容”“教学方法”“教学情境”“教学评价”可以归纳为“教学设计”。据此，将开放式编码所得的 11 个范畴最终提炼为 3 个主范畴，即学习空间、教学设计、角色关系，范畴与主范畴之间的所属关系如表 2 所示。

表 2 主轴编码结果

序号	范畴化

主范畴	范畴
学习空间	智能辅助工具 个性化学习资源 智慧学习环境
教学设计	精准教学目标 多样化教学内容 创新教学方法 智能学习情境 智能评估反馈
角色关系	师生角色重构 同伴互动

3.4. 选择性编码

选择性编码需要系统分析主范畴并从中挖掘出核心范畴。核心范畴应处于中心地位，具有统领性，其他子范畴均围绕其展开。在本研究中，课堂教学的核心是教学设计，它不仅决定了教学活动的结构和流程，还关乎如何利用 GAI 优化教学效果来提升师生互动和学习成效。因此，教学设计被确立为核心范畴，而其他相关概念范畴均指向该范畴。本研究进一步提出了基于 GAI 赋能的课堂教学模式的理论模型，具体理论模型(如图 1 所示)。

3.5. 理论饱和度检验

本研究在完成选择性编码后，进行了不断循环的理论对比与验证，确认模型各范畴概念趋于稳定，满足理论饱和要求，保证所得模型的科学性和可靠度。

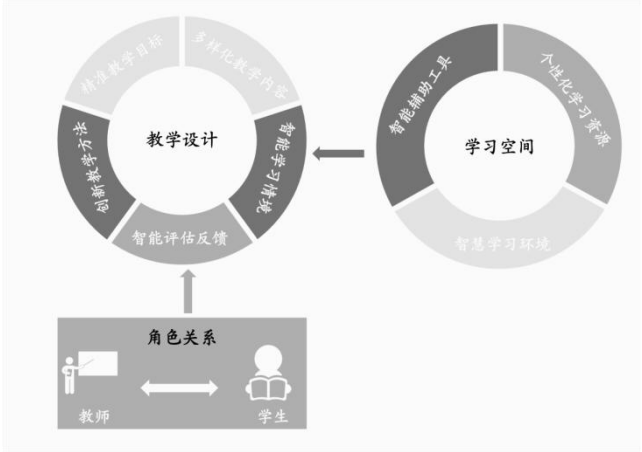


图 1 基于 GAI 赋能课堂教学模式的理论模型

4.模型阐释

基于 GAI 赋能课堂教学模式的理论模型聚焦课堂教学的“人、事、物”三大维度，有助于建立对 GAI 赋能的课堂教学的系统认知，更好地描述理论模型中蕴含的基本结构。通过扎根编码发现，理论模型遵循“前置条件-实施过程-效果呈现”的发展规律，主要聚焦于学习空间、教学设计、角色关系三个主题，具体模式则体现在三个主题各自映射的二级节点和三级节点体现的架构关系。

4.1. 内容维度一：学习空间（物）

学习空间作为支持教与学发生的中介物集群，具有物理场所、虚拟场所、工具性客体

和对象性客体四大结构要素
。GAI 赋能的课堂创造了由富技术构建的开放学习空间，为认知建构与展示交流提供“居所”（学习发生的场所）及“转运”（促进主体理解与表现的输送介质）属性
。GAI 赋能课堂教学的前置条件即学习空间，提供了智慧学习环境这一“居所”，以及智能辅助工具、个性化学习资源这些“转运”属性。

智能辅助工具即工具性客体，工具性客体包括形式性工具客体（如语言）和实体性工具客体（如教材、教具、教学设备等）
。通过智能辅助工具的应用，学习空间得以扩展和优化，学生在此空间中可以获得语言辅导、内容智能生成、实时互动和反馈，从而提升学习效果，激发主动学习的动力。从编码结果中我们发现工具包含两种类别，一种是教师助手工具，例如内容生成工具、评估和测试工具；一种是学生支持工具，例如互动式工具和语言学习辅助工具。

个性化学习资源即对象性客体，对象性客体指主体认识发生所需要的一切现实的和可能的外部资源的集合，包括物质资源、交往资源、经验资源等。编码结果显示，GAI 在教育资源的应用中主要体现为以下三类：智能辅导类，GAI 通过分析学生的学习行为与成绩，能够提供即时反馈与个性化指导，有效帮助学生解决学习过程中遇到的困难，提升学习效率；个性化学习资料，GAI 基于学生的学习进度与偏好，能够生成个性化的学习计划与学习材料，支持学生更加高效地掌握知识；多媒体资源，GAI 为教师提供辅助功能，帮助其制作高质量的视频教程、动画演示及其他多媒体内容，进一步丰富课堂教学形式，增强教学效果。

智慧学习环境即教学场所，包含实体场所和虚拟场所。实体场所指的是物理空间，如学校内部的实体教室空间；虚拟场所指的是网络空间，支持师生、生生异地分离开展教与学的虚拟环境。在学习空间中，智慧学习环境通过整合实体场所和虚拟场所，为学生提供了多元化的学习体验。实体场所包括智慧教室和智慧实验室，提供了一个互动式、技术驱动的学习空间，支持学生通过实际操作和实践体验来深化学习。而虚拟场所则包括在线教育平台和社区平台，提供了灵活的在线学习资源和社交互动平台，方便学生随时随地进行学习、交流和合作。这两种场所的结合，为学生提供一个融合技术与实际需求的学习平台，提升学习的个性化和互动性。

4.2.内容维度二：教学设计（事）

课堂教学的核心在于教学设计，通过对典型案例的扎根编码，将教学设计的关键要素概括为精准的教学目标、多样化的教学内容、创新的教学方法、智能的学习情境以及智能评估反馈机制。这些要素共同构成了教学设计的整体框架，是 GAI 赋能课堂教学模型的实施过程部分的主要内容。

精准的教学目标是教学设计的起点。它为教学活动提供了明确的方向，并为教师实施教学、评估学习效果提供了重要依据。具体且可操作的目标不仅确保了教学活动的有序开展，还能够有效指导教师在教学过程中及时调整策略，以满足学生的学习需求并评估学生的学习成果。通过对典型案例的编码分析，发现基于生成式人工智能（GAI）的课堂教学目标可大致分为以下三类：一类是技术可用性评估。一类是素养能力培养。还有一类是个性化学习实现。GAI 技术的应用使得教师能够根据每个学生的特点，提供个性化的学习体验，从而激发学生的学习兴趣并提高学习的参与度和效率。

在 GAI 赋能的课堂教学设计中，多样化的教学内容设计强调与学科特性相适配，并充分考虑学生的个体差异和学习需求，以提供差异化的学习材料。这种设计旨在满足不同学生的学习风格和能力水平，从而促进个性化学习和提升教学质量。通过对典型案例的编码分析，

基于人工智能赋能的课堂教学内容可以归纳为以下三种类型：一种是认知发展类，此类内容侧重于培养学生的认知能力，包括理论课程、阅读与写作课程等。一种是问题解决类，此类内容聚焦于解决实际问题的能力培养，涵盖编程课程、实践课程和实验教学等。一种是内容创新类，此类内容强调创新思维和设计能力的培养，主要包括设计课程。

创新的教学方法能够激发学生的学习兴趣和主动性，促进深度学习的发生。通过对典型案例的编码分析，GAI 赋能课堂的教学方法可以大致分成两种，以教师为中心的教学方法和以学生为中心的创新教学方法，主要包括传统的讲授法和讲练结合的教学模式。以学生为中心的创新教学方法强调学生的主动参与和自主学习，包括泛在式学习、探究式学习和翻转课堂等。这些方法在 GAI 赋能的课堂中得到了进一步的优化。

智能学习情境的构建依赖于技术手段，作为一种能够渲染知识价值并激发学习者学习动机的重要活动维度，学习情境在 GAI 的赋能下，展现出更加适应学习内容的差异化特点。具体而言，GAI 支持构建三种主要的差异化学习情境：首先是角色型情境，GAI 通过模拟真实或虚拟人物角色，创建能够促进学习者与角色对话的互动情境。在这种情境中，学生能够通过与虚拟人物进行对话，深刻体验、理解并建构相关知识。其次是事件型情境，GAI 通过模拟真实或虚构的历史事件或政治事件，使学习者在事件的发生过程中理解并构建相关知识。最后是事理型情境，GAI 通过揭示事物之间的内在联系和规律，引导学习者理解并掌握事物的本质与原理。

智能评估反馈机制通过实时数据分析和个性化反馈，帮助教师及时调整教学策略，同时为学生提供学习指导，提升学习效果，是教学设计重要的组成部分。课堂评价可以根据其功能和实施阶段被分类为诊断性评价、形成性评价和总结性评价。这三种评价方式各有其独特的功能和应用场景，共同构成了一个全面的评价体系，以支持教学过程和学生学习的持续改进。在 GAI 赋能的课堂教学中诊断性评价包括学生状态评估、个性化学习路径生成、学习者特征分析。基于 GAI 的课堂教学中的形成性评价主要体现在学习过程评估、实时反馈中。总结性评价通常在教学活动结束后进行，旨在对学生的学习成果进行总结和评估。这种评价方式主要用于衡量学生是否达到了预定的学习目标，并为教师提供关于教学效果的反馈。在 GAI 赋能的课堂教学中体现为学习成果展示、智能推送数据报告、多维度画像、学习体验。

4.3.内容维度三：角色关系（人）

随着生成式人工智能在教育的引入，师生角色关系发生了显著重构。GAI 作为一个直接的中介者，在课堂教学中架起一座桥梁，以直接促进教师与学生的关系、学生与自我的关系以及学生与学生的关系。

从师生关系的角度来看，教师的角色从传统的知识传递者转变为学习引导者、技术支持者，甚至是内容创作者。与此同时，学生从被动的知识接受者转变为更积极的合作伙伴，能够与 AI 进行互动学习。GAI 技术的应用可以促进师生之间更频繁、个性化的互动。

对于学生与学生之间的关系，GAI 技术通过提供协作工具，促进学生之间的合作与交流。同时 AI 为他们提供适时的支持，形成互助与依赖的学习氛围。这种协氛围不仅增强了学生之间的互动，还促进了知识的共同构建。

就学生自我发展而言，GAI 能够根据学生的学习进度、兴趣、学习方式等提供个性化的学习资源与指导，促进学生对学习内容的理解，增强学生的学习动力与自信心。这种个性化支持不仅提升了学生的学习体验，还促进了他们的自主学习能力。

5.结论与讨论

5.1.生成式人工智能赋能的教学模式

基于典型案例扎根编码的结果，本研究创新设计一个 GAI 赋能的课堂教学模式（如图 2 所示），围绕 GAI 赋能的课堂教学进行优化，充分体现人、事、物三大维度的有机结合，旨在通过 GAI 技术提升课堂互动、个性化学习和教学质量。GAI 不仅能够为学生提供丰富的学习资源、即时反馈、个性化辅导，还能够解决传统教学过程中常见的师生互动不足、学生个体差异大等问题。

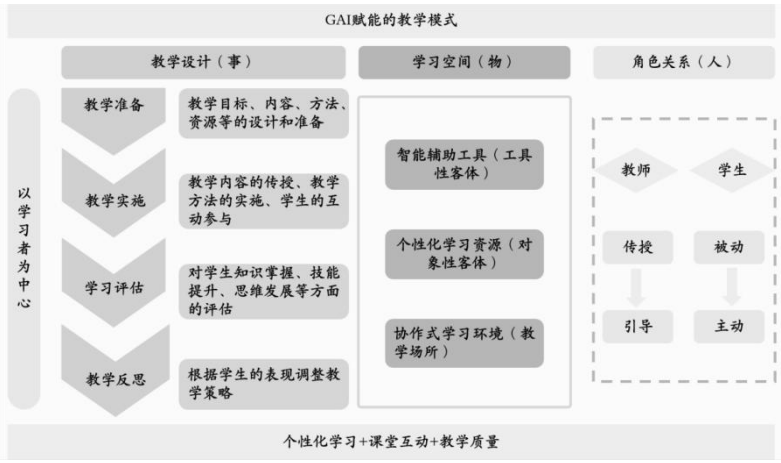


图 2 GAI 赋能的课堂教学模式

生成式人工智能（GAI）赋能的课堂教学设计包括四个基本环节：教学准备、教学实施、学习评估以及教学反馈与反思，并在各环节中充分体现了技术的优势和作用。

在教学准备阶段，GAI 技术通过分析学生的历史数据和学习表现，为教师推荐个性化教学目标，避免“一刀切”设计。它还能根据学生的兴趣、能力和学习进度，帮助教师制定适配的教学内容，差异化设计教学方案，激发学生自主学习能力，为课堂奠定个性化基础。

在教学实施阶段，GAI 借助智能工具增强师生互动与协作，提供个性化学习路径。教师可利用 AI 支持的工具（如互动白板）快速展示内容、解答疑问，同时 GAI 促进学生间的互动与合作，通过协作平台和讨论工具实现实时交流与项目协作。此外，GAI 通过虚拟角色模拟和场景重现等方式增强课堂沉浸感，提升教学效果。

在学习评估阶段，GAI 通过数据分析和实时反馈实现精准、个性化的评估。它实时跟踪学生学习进度，生成动态评估报告，为教师提供多层次的学生表现分析。与传统终结性评估不同，GAI 更注重形成性评估，综合运用小测验、互动练习和自我评价等手段，生成个性化学习画像，助力教学决策科学化。

在教学反馈与反思阶段，GAI 结合智能评估和数据分析，优化教学过程。教师通过平台获取课堂教学和学生表现数据，及时调整策略；学生则根据个性化反馈改进学习方法。教师还可基于评估报告开展教学反思，持续优化教学内容与方法，为后续课程设计提供依据，推动教学模式改进。

在 GAI 赋能的教学模式中，教师与学生的角色发生重构。教师从知识传授者转变为学习引导者和支持者，学生从被动接受知识转变为自主探索者。学习空间也更加多元化与智能化，包括智能辅助工具、个性化学习资源和协作式学习环境，通过技术支持互动与合作学习，进一步提升学习体验与效果。

5.2.教学模式的实践路径

GAI 赋能的课堂教学模式通过智能工具和数据分析，贯穿课前、课中和课后环节，优化教学过程和学习体验，形成一个循环优化的闭环。

课前准备阶段是教学的起点，包括预习辅助、课程规划和状态评估。GAI 系统根据学生的历史数据和学习进度，推荐个性化预习材料，帮助学生构建知识框架。智能推荐系统根据学生兴趣和能力调整内容，确保有效学习。教师利用 GAI 进行课程设计，系统生成教学内容、方法和目标建议。同时，GAI 实时评估学生学习状态，为教师提供反馈，帮助其了解学生背景和潜在障碍，为教学调整提供依据。

课中交互阶段是核心环节，涵盖智能答疑、互动学习、过程评估和课堂监控。GAI 通过智能答疑平台实时解答学生疑问，提供个性化支持。自然语言处理技术分析问题并反馈，确保学生学习不被遗漏。课堂中，GAI 利用在线讨论平台、协作白板和虚拟学习环境等工具，促进师生和学生之间的互动与合作，增强集体学习氛围。系统实时跟踪学生进度和表现，提供定制化学习建议，帮助学生调整策略。同时，GAI 通过数据采集和分析，帮助教师监控教学效果，调整教学节奏，确保每个学生跟上进度。

课后服务阶段通过持续跟踪和反馈，巩固学习成果并优化教学。GAI 分析学生数据（如测试成绩、作业完成情况等），评估教学效果和学生学习成果，为教师提供调整教学策略的依据。学生可通过线上讨论区、学习小组和问答社区等功能继续交流，解决疑问，巩固知识。GAI 为教师提供教学建议，辅助调整课程内容和方法，同时为学生提供个性化学习建议和资源，支持自主学习。

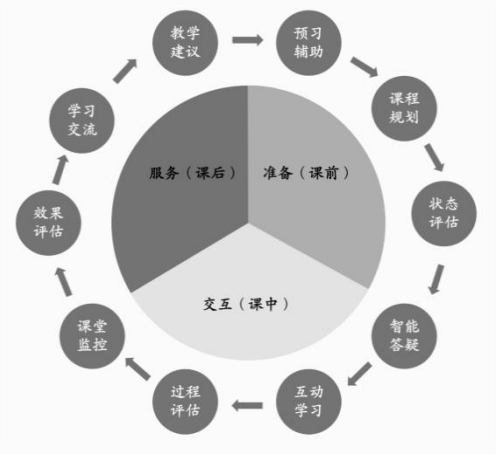


图 3 GAI 赋能课堂教学的实践路径

结语

本研究基于 65 份“人工智能+教育”典型案例，运用扎根理论构建了基于生成式人工智能（GAI）赋能课堂教学的理论模型。该模型创新性地涵盖了学习空间（物）、教学设计（事）和 GAI 对角色关系影响（人）三个维度，其中教学设计贯穿始终，旨在促进课堂互动并提升教学质量。研究进一步从课前准备、课中交互到课后服务，系统揭示了 GAI 在不同教学阶段赋能课堂的具体实践路径。期望本研究能为我国“人工智能+教育”行动提供有效借鉴。同时，研究也指出，GAI 技术在带来教学便利的同时，不可忽视数据隐私保护和算法偏见等伦理问题，呼吁加强相应的伦理规范和技术透明度。未来研究还需深入探讨 GAI 在不同教育情境中的适应性与局限性，填补系统性实证研究不足，推动人工智能赋能教育实现真正“落地生根”。

参考文献

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, 42 (04), 5-11. doi:10.13811/j.cnki.eer.2021.04.001.

沈书生. (2020). 学习空间：学习发生的中介物. 电化教育研究, 41 (08), 19-25+42. doi:10.13811/j.cnki.eer.2020.08.003.

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. doi:10.13927/j.cnki.yuan.20231010.001.

Kerimbayev, N., Menlibay, Z., Garvanova, M., Djaparova, S., & Jotsov, V. (2024). A Comparative Analysis of Generative AI Models for Improving Learning Process in Higher Education. 2024 International Conference Automatics and Informatics (ICAI), 271-276. https://doi.org/10.1109/ICAI63388.2024.10851491.

Wang, W., Liu, X., Gao, K., Huang, J., Yuan, Y., He, P., Wang, S., & Tu, Z. (2025). Can't See the Forest for the Trees: Benchmarking Multimodal Safety Awareness for Multimodal LLMs. .

Du, Z., Qian, Y., Liu, X., Ding, M., Qiu, J., Yang, Z., & Tang, J. (2021). GLM: General Language Model Pretraining with Autoregressive Blank Infilling. , 320-335. https://doi.org/10.18653/v1/2022.acl-long.26.

Salem, S., & Ramadan, M. (2020). Arabic Text to Image Generation based on Generative Network of Fine-Grained Visual Descriptions. Benha Journal of Applied Sciences. https://doi.org/10.21608/bjas.2020.226901.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. Journal of Science Education and Technology, 32(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. Interactive Learning Environments, 32(10), 6187–6203. https://doi.org/10.1080/10494820.2023.2253861

Lan, Y. J., & Chen, N. S. (2024). Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. Educational Technology & Society, 27(1), Article 1. https://doi.org/10.30191/ETS.202401_27(1).PP01

Friederichs, H., Friederichs, W. J., & März, M. (2023). ChatGPT in medical school: How successful is AI in progress testing? Medical Education Online, 28(1), 2220920. https://doi.org/10.1080/10872981.2023.2220920

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. British Journal of Educational Technology, 55(6), 2503–2529. https://doi.org/10.1111/bjet.13460

Li, H. F. (2023). Effects of a ChatGPT-based flipped learning guiding approach on learners' courseware project performances and perceptions. Australasian Journal of Educational Technology, 39(5), 40–58. https://doi.org/10.14742/ajet.8923

Bower, M., Torrington, J., Lai, J. W. M., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence?

Outcomes of the ChatGPT teacher survey. Education and Information Technologies, 29(12), 15403–15439. https://doi.org/10.1007/s10639-023-12405-0

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. International Journal of Educational Technology in Higher Education, 21(1), 21. https://doi.org/10.1186/s41239-024-00453-6

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. International Journal of Educational Technology in Higher Education, 20(1), 63. https://doi.org/10.1186/s41239-023-00434-1

Tang, K. S., & Cooper, G. (2024). The role of materiality in an era of generative artificial intelligence. Science & Education. Advance online publication. https://doi.org/10.1007/s11191-024-00508-0

数字化转型背景下的循证教研新范式：模式构建与实践探索

A New Paradigm of Evidence-Based Teaching and Research in the Context of Digital Transformation: Model Construction and Practical Exploration

周洋洋[1*]，易诗情[1]，杨时娇[1]，黄雅萱[1]，汤名珍[1]
华南师范大学 教育信息技术学院
2470230540@qq.com

【摘要】 为打破“囿于经验”的传统教研现状瓶颈，新时代教研进行着趋于证据的创新转型。循证式教研逐渐成为教育领域研究的热点问题。然而，我国循证式教研的研究尚处于起步阶段，能够详细说明如何设计并实施循证式教研活动的研究寥寥无几，导致一线教师难以开展基于循证实践理念及方法的教育教学活动。基于此，本研究以数字化平台为基础，以 SECI 理论模型为设计理念，提出了一种全新的注重实践性和科学性的循证式教研活动模式，设计了循证式教研活动过程实践路径，并基于此进行应用实践，为一线教师设计和开展数字循证式教研活动提供一定的参考框架。

【关键词】 循证教研；数字化平台；SECI 模型；教研活动；模式构建

Abstract: *To break the traditional teaching research bottleneck, new - era teaching research is transforming towards evidence - based approach. Evidence - based teaching research is a hot topic in education. However, China's research on it is in the early stage with few studies detailing its design and implementation, making it hard for teachers to apply evidence - based practice. This study proposes a new evidence - based teaching research model based on digital platforms and SECI theory, designs a practical path and provides a reference framework for teachers to conduct digital evidence - based teaching research activities.*

Keywords: Evidence - based teaching research; digital platform; SECI model; teaching research activity; model construction

1.引言

教研在指导教育教学实践、促进教师专业发展、服务学生综合素质提升、保障基础教育教学质量等方面发挥着十分重要的作用。立足新时代，面对教育数字化转型、基础教育减负增效的新形势，教研工作还存在着教研方式不科学、教研效果不理想等诸多问题(陈锋娟等人,2024)。教育领域亟需探究有效策略，以确保教研活动回归其核心要务——即促进教师知识的系统构建并生成具有实质性价值的教育教学知识与经验(陈向明,2020)。顺应数字化时代需求，(董洪亮,2021)教研转型势在必行：以教学改进为目标的教研从基于经验转向了基于证据(傅京,2021)，从经验之源转向了科学之法(孙朝仁,2022)，循证教研逐渐成为教育领域研究的热点问题。循证教育研究源自 20 世纪 80 年代的循证医学，它强调在教育实践中依赖科学研究证据(胡艳等人,2020)，倡导理性思维来指导决策和行动。然而，在我国循证教研尚处于萌芽阶段(廖伟,2020)，其深入推广受限于价值认知不足、专业能力欠缺、(任萍萍和李鑫,2021)数据资源匮乏以及实际操作困难等多重挑战(Georgiou, D.et al.,2023)。鉴于此，本研

究以数字化平台为基础，以 SECI 理论模型为指导，构建了循证式教研活动模式及其实践路径，旨在为教师提供数字化教研操作指南，助力教育质量提升。

2.相关理论依据及文献探讨

2.1. SECI 理论模型

SECI 模型是由日本学者野中郁次郎等人于 1995 年提出的知识管理模型，用于解释知识在组织内部的生成、分享和转化过程。SECI 是 Socialization、Externalization、Combination 和 Internalization 这四个英文单词的首字母缩写，分别代表了知识从个体到组织层面的不同阶段和转化过程。社会化（Socialization）指的是隐性知识之间的转化，是指通过共享经历来建立隐性知识的过程。外在化（Externalization）指隐性知识向显性知识的转化，是指将隐性知识用显性化的概念和语言进行清晰表达的过程。这个过程在知识创造过程中起着至关重要的作用。组合化（Combination）指的是显性知识之间的组合。它是将语言或符号等显性概念组合化和系统化的过程。内在化（Internalization）即显性知识到隐性知识的转化。它是将显性知识形象化和具体化的过程，通过“汇总组合”产生新的显性知识被组织内部吸收、消化，并升华成自己的隐性知识。SECI 模型强调了知识在个体和组织之间的相互作用和转化过程，强调了知识的社会性和动态性。北京师范大学未来教育高精尖创新中心曾基于 SECI 模型组织教师培训活动时认为，该模型重视教师学习共同体的构建、“场域”及教学情境在知识建构中的作用、多样化的互动交流、多样化的活动体验、理论性与实践性知识的相互转化(王阿习等人，2016)。本研究采用 SECI 模型来设计教研活动模式，旨在通过建立开放共享的教研环境，鼓励教研共同体进行交流和互动，以促进知识的创造和转化。

2.2. 循证式教研

20 世纪 90 年代在临床医学领域出现了“循证医学”（evidence-based medicine）的概念，循证医学强调将科学研究的结果与临床实践相结合，以提供最佳的医疗方案。随后，这个概念被引入到教育领域，形成了“循证教育”（evidence-based education）的概念。循证教育强调将证据与现实因素综合考虑，如时间安排、政策、伦理等，并运用专业知识进行专业判断。这与传统的基于教师经验、直觉的问题解决和教学决策有着本质区别。由循证教育实践发展而来的循证式教研强调要打破研讨式教研中的经验权威(Gambrill E.,2018)，将教研实践活动构建于科学的证据之上，结合教师个人教学经验、专业知识和学习者的特点，制定最优的教学实践方案，并在实践中不断反思和改进，这种教研模式能够帮助教师进行知识建构，鼓励教师积极主动地对自己的教育实践进行深入研究和探讨。

2.3. 教研活动新思路

为了建设教育强国，中共中央国务院于 2018 年 1 月在《关于全面深化新时代教师队伍建设改革的意见》中指出，当前部分教师的专业化水平需要提高，以适应新时代人才培养需要；且培养高素质教师队伍，要全面深化改革(教育部,2018)。为了实现这一目标，教研活动必须发挥核心作用。然而，在我国，教研活动的本质含义逐渐被偏离，以研讨为主的教研活动实践中，教育教学问题的深度反思与系统探究得不到充分体现，教师的知识结构也无法实现质的变化。为了有效地应对这一难题，国内学者们也在不断探索教研活动新思路。我国学者陈峥等人在戴维斯提出的 5A 循证实践步骤的基础上，开发了适用于中小学教师的 6A 循证教研模式(陈峥等人,2022)；陈鹏等人通过科学地设计、组织和实施，并在“循证理念”支持下开展课例研究，促进教师有效地进行知识建构(陈鹏等人,2023)；杨玉东等人指出教师应当是课例研究的循证者，教研员和研究人员可在循证实践中发挥“智慧他人”的作用(杨玉东和严加

平,2022)。由此可见，循证式教研逐渐成为了教育研究领域的新方向。作为一种新型的教研模式，循证教研在内容和方式上进行了创新，为教研活动提供了新的思路和方法。本研究以SECI理论模型为设计理念，构建了循证式教研活动模式，提供具体流程指引，并通过实践说明，帮助中小学教师理解和掌握循证式教研活动过程，推动其专业发展。

3.循证式教研活动模式构建

3.1. 循证式教研活动模式构建的基本思路

循证教育实践赋予以经验为基础的传统教师教育实践新内涵和新特征，是一种以科学化解决教师教育实践问题为核心的教师教育新范式(裴淼等人，2020)，其作为教育问题解决和教育决策的基础，能够通过提供最优的干预，改善教育实践结果(Fineout-Overholt, E. ,2019)。在教育领域，循证教育实践的概念和方法论主要涵盖了基于证据的学习与教学策略、教育决策与政策治理、以及教师的循证教育等方面(陈红霞,2022)。而在教学研究活动中，通常涉及问题的解决、学术的交流研讨以及具体的教学实施，这些活动要求教育工作者进行基于证据的系统思考与行动(刘帅,2022)。因此本研究认为，教研活动中所融入的循证实践可被划分为循证决策、循证研讨、循证教学以及循证评估。为了促进有效的知识转化与创新，本研究采纳了SECI模型作为理论框架，将“循证研讨”和“循证教学”的过程与知识转换的各个阶段紧密结合。通过此一整合性策略，可以加速教育教学问题解决方法的发展，并促进相关知识的创造和传播。具体模式图如图1所示。

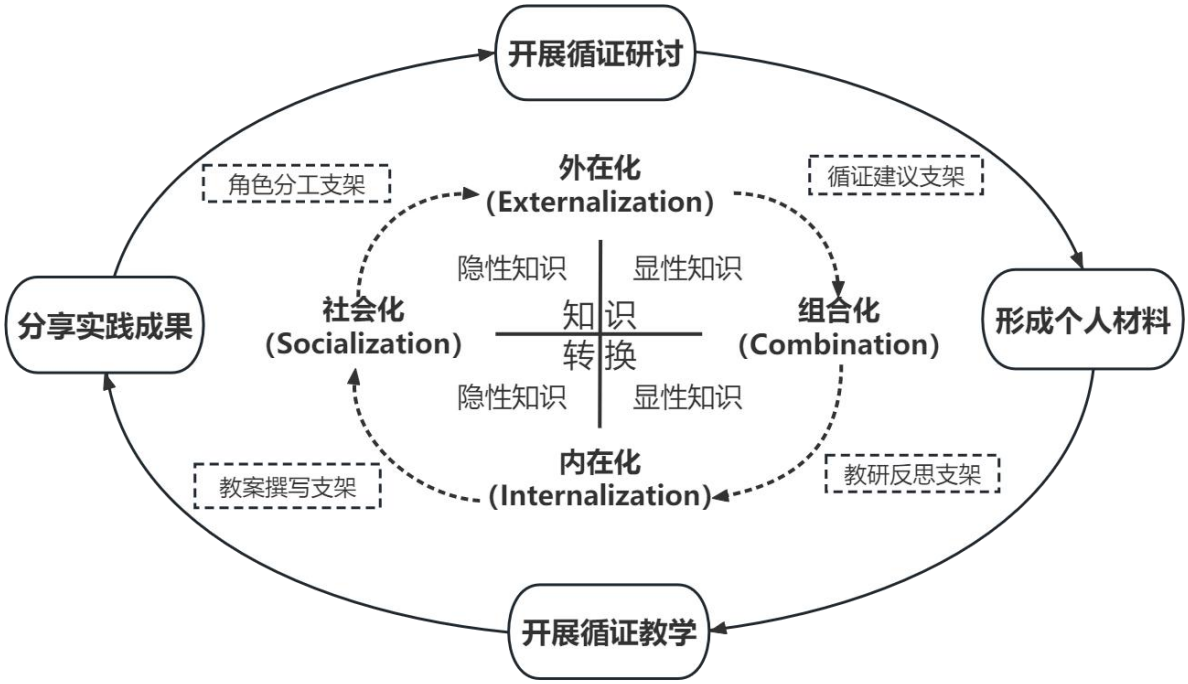


图1 基于SECI的循证式教研活动模式

3.2. 循证式教研活动模式要素及关系分析

循证式教研活动模式的核心在于知识转换与共享的过程，这与SECI模型中的“社会化、外在化、组合化和内在化”四个阶段紧密相连。在“社会化”阶段，教研共同体成员彼此分享各自的实践经验与成果，实现隐性知识的共享与交流；在“外在化”阶段，教研共同体成员进行深入的循证研讨，将隐性知识转化为显性化的理解与表达；在“组合化”阶段，主备教师将研讨过程中收集到的各种建议和观点进行整理与归纳，形成系统化的个人资料；在“内在化”

阶段，主备教师将整理好的教学设计付诸实践，将显性知识再次转化为隐性知识，并通过实际教学进一步验证与完善教学方案。在不同的知识转换阶段赋予相应的支架支撑，以确保知识转换的顺利进行。在教研活动初期，教案撰写支架有效地引导并赋能教师全面地构建和完善教学设计方案；角色分工支架在教研共同体环境中展现着协同效能，通过明确主备教师、从备教师及教育研究者等多元角色的功能定位与职责边界，实现团队内部的精细化分工与协作；借助循证建议单、证据导航助手等建议支架工具，在教学研讨中严谨而系统地提炼出基于证据的观点或建议，有力地提升研讨的质量和深度；同时，教研反思支架为教师提供了反思路径，鼓励他们在实践中深入挖掘，持续迭代自身的教学理念与方法。整个过程构成了一个循环且螺旋上升的循证式教研实践模式："分享实践成果—开展循证研讨—形成个人资料—开展循证教学"。这一模式不仅推动了教师个体专业能力的提升，也在集体层面上促进了教研共同体的共建共享与协同发展。

4.循证式教研活动模式应用实践

4.1.循证式教研活动过程实践路径设计

基于已构建的循证式教研活动模式，本研究专门针对循证式教研活动设计了循证式教研活动过程实践路径，以期为循证式教研活动的组织与实施提供更为具体的指引，具体实践路径如图 2 所示。

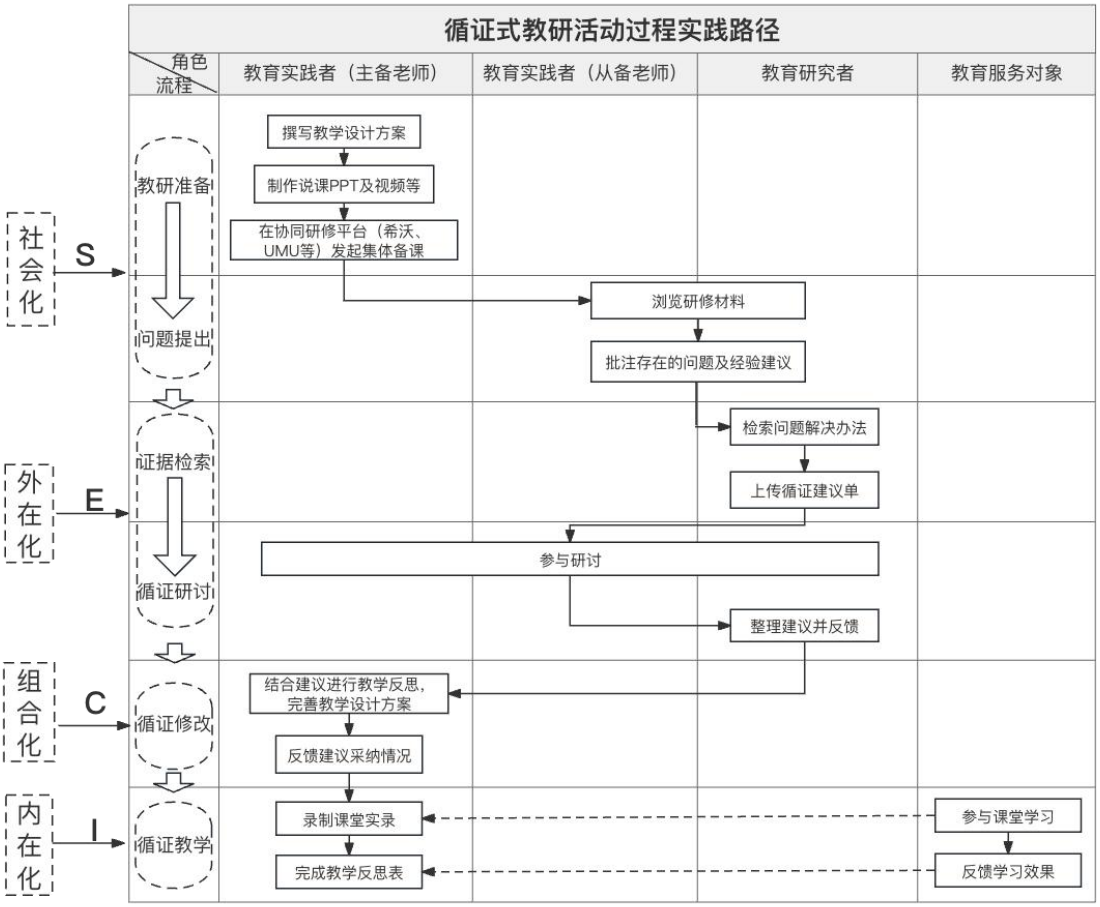


图 2 循证式教研活动过程实践路径

循证式教研活动过程实践路径遵循“S-E-C-I”知识转化过程，分为了“教研准备-问题提出-证据检索-循证研讨-循证修改-循证教学”六大阶段。在教研准备阶段，参研教师需根据教研活动要求准备教研材料，并在教研共同体内进行共享；在问题提出阶段，教研共同体成员需互相浏览教研材料，并围绕教研主题针对教研材料提出其中存在的问题；在证据检索阶段，教研共同体成员针对自己在教研材料中发现的问题，检索问题解决方法，以此作为循证研讨过程中的论据；在循证研讨阶段，各位教研共同体成员围绕主题目标对教研材料依次给出具有一定科学依据的观点建议，并在评估基于证据的建议之后达成共识，形成最佳问题解决方法或策略；在循证修改阶段，各位参研教师需根据循证研讨的结果对教研材料做出基于证据的改进，形成整合了有效建议的新材料；在循证教学阶段，参研教师需将已完善的教研材料付诸实践，进行课堂实录，以便在实践后回顾教学过程并反思。

4.2. 应用实践过程

为优化循证式教研活动模式及其实践路径，本研究依据广州市人工智能普及项目背景，展开了以课例研修为主的小学人工智能教学专项教研活动实践。应用实践活动共开展了三轮行动研究，每轮均遵循本研究提出的实践路径设计实施教研活动，完成后根据前一轮教研活动过程中存在的问题进行反思与改进，并在后一轮教研活动中再次进行实践与优化。

第一轮应用实践活动以满足参研教师的教学需求为导向，旨在协助其优化和完善教学设计方案。此次循证式教研活动选择希沃白板作为教师们的协同研修平台，一方面用于主备教师制作或上传个人备课的教研材料，另一方面用于发起集体备课，邀请教研共同体成员参与教研材料审核，完成循证式案例分析任务。按照“教研准备-问题提出-证据检索-循证研讨-循证修改-循证教学”思路实施后，笔者发现容易出现“因对教研流程不明晰、对教研活动环境不熟悉而影响教研任务完成”的现象。因此，为教研共同体成员提供明确的任务指引显得尤为重要。针对此问题，本研究在 UMU 互动学习平台中搭建了循证式教研活动任务目录，并制作了平台操作手册，来帮助教师更好地了解教研活动的主要环节与任务。

第二轮应用实践活动旨在提升教师的教学设计方案质量，以保证其教学实践效果。此次教研活动将 UMU 互动学习平台引入教研活动中，用以发布教研任务，明晰角色分工，把控教研节奏。其中，主备教师需完成个人备课、主题研讨、实践反思的主要任务，从备教师和教育研究者等教师共同体成员需完成集体备课、主题研讨的主要任务。从教研活动任务来看，此过程将形成的成果为教学设计方案、课堂实录及教学反思。按照这一思路进行实践后，教研任务完成效果较好，但发现证据检索阶段中容易出现证据检索困难、证据类型单一、获取证据的途径不清晰等循证路径不明现象，使得真正的循证赋能难以有效实施。基于此，为规范证据类型，指引循证路径，本研究进一步优化了循证建议单的框架，对循证建议单中的证据类型和证据来源给出了参考。

第三轮应用实践活动的最终目标是打造小学人工智能学科展示课例。通过优化后的循证建议单和循证研讨框架，提升证据检索、循证研讨等阶段的实施效果。第三轮在一二轮的基础上展开，重点任务目标是进一步细化教研成员的职责分工，明确每位成员任务，加强成员间任务关联性，并形成教研活动资源清单，完成课例剖析、课例制作等任务。本研究在每轮实践结束后，都针对性地提出了改进策略，不断对循证式教研活动模式进行逐步的调整和优化，以实现模式的持续完善和提升。

4.3. 应用效果分析

教研活动中教师的采纳行为代表了理解和认可的态度，教师在教研活动中对建议的采纳程度越高则说明其获得的有用建议占比越高，收获的实际帮助越大(秦良娟等人，2022)。同

时，建议采纳度能够反映教师在教研活动中的收获程度，一定程度上也能够反映出教研活动效果。本研究通过对三次循证式教研活动中建议的生成与采纳情况进行统计分析，以此验证循证式教研活动的效果。

通过对三次循证式教研活动生成的建议数量和占比进行整理统计（见表 1），发现每一次教研活动的经验建议在数量上都远超循证建议，说明基于经验发表建议一直是教研活动中的主要现象。但随着循证式教研活动运行方式的不断优化迭代，经验建议的占比在不断降低，循证建议的占比在不断提升，且第三轮的循证建议比例为 39.13%，在平均水平之上，相较于前两次教研活动有着更为明显的循证式教研特点，循证式教研活动组的循证氛围更加浓厚，说明循证建议支架等工具的优化设计对循证效果来说是有效的。

表 4 循证式教研活动建议生成情况

教研实践活动	建议生成总量	经验建议数量	经验建议占比	循证建议数量	循证建议占比
第一轮	28	21	75.00%	7	25.00%
第二轮	20	13	65.00%	7	35.00%
第三轮	23	14	60.87%	9	39.13%
整体情况	71	48	67.61%	23	32.39%

通过对本研究中三次循证式教研活动中所生成建议的采纳度进行整理呈现（见表 2），发现教师们对循证建议的采纳度高于对经验建议的采纳度，可见，循证建议相对于经验建议来说，更容易被采纳。

表 2 循证式教研活动建议采纳情况

教研实践活动	经验建议采纳度	循证建议采纳度	全部建议采纳度
第一轮	57.14%	71.43%	60.71%
第二轮	69.23%	71.43%	70.00%
第三轮	71.43%	77.78%	73.91%

从整体上看，随着循证式教研活动的运行方式不断优化，经验建议和循证建议的采纳度均不断提升，且经验建议的采纳度与循证建议的采纳度成正比关系，可见，随着循证式教研活动的深入推进，循证式教研活动组所生成的经验建议和循证建议的质量均会不断提升。

在“建议采纳情况”这一维度下的结论，正是对研究目标中“以证据取向的教研活动过程提升建议的质量，提高教师对建议的采纳程度，以此促进教研效果的提升”的回应。根据建议采纳情况，循证式教研活动中的循证建议相对于经验建议更容易被采纳，经验建议采纳度会随着循证建议采纳度的提升而提升，因此，从整体上看，循证式教研活动能够提高建议的采纳程度，提升教研活动效果。

5. 总结

为了深入推动教研工作的改革与创新，提高基础教育质量，本研究针对传统教研活动中出现的“教研活动人员单纯依赖经验给出建议，且这些建议难以被大众接受、难以应用于教学实践”的问题，明确了提升教研活动中教师对建议的采纳程度的主要目标。为了实现这一目标，本研究以数字化平台为基础，基于 SECI 模型理论，提出了一种全新的注重实践性和科学性的循证式教研活动模式，旨在帮助教师采用更加科学、严谨的方法来设计和实施循证式教研

活动。通过构建出循证式教研活动模式和设计出循证式教研活动过程实践路径，并进行应用实践，为廓清循证式教研的实施操作提供了显性化的表征途径，从而为当今教育领域开展循证式教研活动提供一定的参考框架。笔者相信，循证实践和循证教研已成趋势，通过应用循证式教研活动模式，教师可以更好地应对教育挑战，为学生的成长和发展提供更优质的教育服务。未来，将聚焦优化数字化教研规则与互动过程，构建支撑数字化教研的平台等，以适应数字化时代教育发展需求，推动教育向更加智能化、个性化的方向发展。

参考文献

陈锋娟、章光琼、张思等（2024）。精准教研的内涵特征、价值取向与发展路径。中国远程教育, 44(03), 68–78。

陈向明（2020）。跨界课例研究中的教师学习。教育学报, 16(2), 47–58。

陈峥、禹双文、卜凡燕（2022）。用 6A 循证教研模式突破教师用证与制证难题。教师教育论坛, 35(12), 19–22。

陈红霞（2022）。循证教研：构建小学数学教材二次开发的区域新样式。小学数学教师, 2022(06), 23–30+2。

陈鹏、梁友明、叶虹（2023）。大数据循证课例赋能知识建构与教学改进的研究。中国电化教育, 2023(04), 99–106+121。

董洪亮（2021）。基于证据的教研转向——基本概念问题的初步讨论。华东师范大学学报(教育科学版), 39(05), 108–115。

傅京（2021）。教研转型：从传统教研走向现代教研。教师教育论坛, 34(08), 32–34。

胡艳、刘佳等（2020）。教师教育领域的循证实践：价值与挑战。教师教育研究, 32(6), 1–7。

廖伟（2020）。循证教师专业发展之 PD&R 实践模式——以“北京师范大学 APEx 卓越教育家培养项目”为个案的研究。教师教育研究, 32(4), 9–16。

刘帅（2022）。从“经验”走向“循证”：中小学教师数据素养的“纵向伴随式”研修样态。教学月刊·中学版(教学管理), 2022(06), 44–48。

秦良娟、李斌、吴新杰（2012）。教师研修社区教师采纳行为与教学能力提升的关系研究。中国电化教育, 2012(02), 63–70。

任萍萍、李鑫（2021）。循证教育研究：缘起、困境、体系框架与实施建议。中国电化教育, 2021(12), 33–39。

孙朝仁（2022）。“大教研”：指向立德树人的区域教科研整体改革实践进路。江苏教育, 2022(34), 11–16。

王阿习、陈玲、余胜泉（2016）。基于 SECI 模型的教师培训活动设计与应用研究——以“跨越式项目全国中小学语文和英语骨干教师培训”为例。中国电化教育, 2016(10), 24–30。

杨玉东、严加平（2022）。对中式课例研究中循证实践的再理解。上海教育科研, 2022(02), 9–17。

裴淼、靳伟、李肖艳等（2020）。循证教师教育实践：内涵、价值和运行机制。教师教育研究, 32(04), 1–8。https://doi.org/10.13445/j.cnki.t.e.r.2020.04.001

关于全面深化新时代教师队伍建设改革的意见-中华人民共和国中央人民政府门户网站 [EB/OL]。（2018）[2022-02-18]。http://www.gov.cn/xinwen/201801/31/content_5262659.htm

Fineout-Overholt, E. (2019). A guide to critical appraisal of evidence. Nursing Critical Care, 14(3), 24–30.

Gambrill, E. (2018). Evidence-based practice: An alternative to authority-based practice (Revisiting Our Heritage). Families in Society, 99(3), 283–294.

Georgiou, D., Diery, A., Mok, S. Y., et al. (2023). Turning research evidence into teaching action: Teacher educators' attitudes toward evidence-based teaching. International Journal of Educational Research Open, 4, 100240。

师范生数字化教学能力的影响因素及培养策略研究

Research on the Influencing Factors and Training Strategies of Digital Teaching Ability for Pre-service Teachers

韩春，魏雪峰*，孙可欣，于怡，张倩瑛
鲁东大学教育学院
*xuefengwei99@163.com

【摘要】本研究以 UTAUT 模型为框架，本研究以整合型技术接受与使用模型（UTAUT）为理论框架，通过结构方程模型对 519 名师范生的问卷调查数据进行分析，系统探究绩效期望、努力期望、社会影响及便利条件对师范生数字化教学能力的作用机制。研究发现：1）四类核心变量均对数字素养具有显著正向影响（β=0.32-0.47, p<0.001），其中绩效期望的总效应值最高（β=0.53）；（2）教育实习经历在便利条件与数字行为间呈现显著调节效应（Δχ²=6.24, p=0.013）；（3）变量间存在链式中介效应，社会影响通过绩效期望间接作用于数字化教学能力（间接效应占比 38.7%）。基于各影响因素分析结果，提出有助于提升师范生数字化教学能力的针对性策略，包括完善课程体系、加强实践教学、整合培养资源等，旨在为提高师范生数字化教学能力提供理论依据和实践指导。

【关键词】师范生；数字化教学能力；UTAUT 理论

Abstract: *This study employs the Unified Theory of Acceptance and Use of Technology (UTAUT) to systematically investigate the mechanisms through which performance expectancy, effort expectancy, social influence, and facilitating conditions affect pre-service teachers' digital teaching ability. Data from 519 pre-service teachers were analyzed using structural equation modeling. Key findings include: (1) All four core variables significantly positively influence digital teaching ability (β=0.32-0.47, p<0.001), with performance expectancy demonstrating the highest total effect (β=0.53); (2) Teaching internship experience significantly moderates the relationship between facilitating conditions and digital behavior (Δχ²=6.24, p=0.013); (3) A chained mediation effect exists, with social influence indirectly affecting digital teaching ability through performance expectancy (indirect effect proportion=38.7%). Based on the analysis results of various influencing factors, targeted strategies to help improve the digital teaching ability of normal students are proposed, including improving the curriculum system, strengthening practical teaching, integrating cultivation resources, etc. The aim is to provide theoretical basis and practical guidance for improving the digital teaching ability of normal students.*

Keywords: Pre-service Teachers, Digital Teaching Ability, UTAUT Theory

基金项目：本文系山东省本科教学改革研究重点项目“职前职后教师数字素养提升‘三架构’融通模式的研究与实践”（课题批准号：Z2023320）。

引言

2024 年《中共中央国务院关于弘扬教育家精神加强新时代高素质专业化教师队伍建设的意见》明确提出“数字化赋能教师发展成为常态”的战略要求。数字素养作为教师应对教育数字化转型的核心能力，其培养重心需前移至职前教育阶段。2022 年 4 月，教育部等八部门联合印发《新时代基础教育强师计划》，强调“大力提升教师的新技术应用能力与专业素质，发挥教师在将人工智能与教育融合的重要作用”。同年 12 月，教育部颁布《教师数字素养》数字素养逐渐成为教师必备的核心素养，是现代教师专业素养的重要组成部分。由此可见，随着教育数字化转型的不断推进，教育行业标准不断提高，数字素养逐渐成为教师教学的核心能力。作为未来教师的师范生，对其数字化教学能力的培育是打造新时代高素质教师队伍的必然要求，是推进国家教育数字化战略行动的必然要求，是数字时代提升人才培育质量的必然要求。在当今数字化时代，师范生作为未来教师的后备军，具备较高数字教学能力能够更好地适应未来教育教学工作，利用数字技术创新教学方法、提升教学质量。尽管 UTAUT 模型在技术接受领域应用广泛，但其在教师教育场景中的适用性仍待验证，特别是调节变量的选择需结合师范生培养特性。鉴于此，本文基于 UTAUT 模型全面系统地剖析各因素的作用机制，为提高师范生数字化教学能力提供有效的策略支持，从而加强未来教师队伍数字化建设，推动教育现代化的进一步发展。

1.相关研究动态

1.1.师范生数字化教学能力的内涵

数字化教学能力的发展脉络可追溯至 20 世纪末。1999 年，黄俊联等学者率先提出“数字化教学”概念，开启了教学与技术融合的探索。随着技术迭代（从多媒体到生成式 AI），教师能力内涵持续拓展：传统教学能力逐步与 TPACK 框架整合，强调技术、学科与教学法的三元交互（Mishra & Koehler, 2006）。国际层面，UNESCO《教师 ICT 能力框架》将能力划分为知识获取、深化与创造三级；欧盟则聚焦批判性数字技术应用能力的培养。我国学界对此展开深入探讨：任友群团队（2018）构建了涵盖基础技术素养、技术支持学习与教学的“三维能力”体系；刘清堂强调数字化教学需突破传统课堂，实现多模态环境下的教学重构；吴军其提出“技术赋能四维能力”模型，覆盖教学、科研、管理与专业发展领域。

教育部《教师数字素养》标准（2022）系统定义了数字素养框架，其核心特征体现在：①意识与技能的数字化迁移；②技术深度融入教学全流程（如数据驱动教学设计优化）；③人机协同的场景化创新；④教育伦理的规范性约束。该标准标志着数字素养从工具理性转向生态建构——教师需在复杂教育场景中动态整合技术资源、教学智慧与伦理责任，推动教育数字化转型的系统性变革。数字化教学能力是数字素养的本体构件，是指应用数字技术支持教学活动设计、实施与评价等教学核心活动的能力（闫广芬等，2022)。。综合相关定义，师范生数字素养是师范生适当利用数字技术获取数字信息和资源，解决学习过程、教学实践过程中的问题。应用数字技术创新教学实践是一个融通的动态过程，在此过程中可能出现多元的学科知识认知、活动组织、技术应用路径和数字技术应用场景优化、创新和变革教育教学活动。

1.2. 师范生数字化教学能力的影响因素研究

很多国内外学者通过实证探索，分析出影响师范生数字素养水平的因素主要包括个人态度和外部环境两大方面的内容。Youngkyun、Baek 等人探析了影响教师在课堂上使用信息技

术的多种因素，包括适应外部环境的要求、他人期望、使用技术的基本功等多方面的内容。Alsharief 通过调查得出对信息技术的态度、 自我效能感是影响个体信息教学能力水平的主要因素。 Salleh、Laxman 以“计划行为理论”为理论模型，探索出个人态度、社会规范、感知行为控制等方面会影响师范生在教学中使用数字化技术。教师的数字化教学能力体现在数字技术融入教学的意识、素养、能力和研究，其发展需要政府部门、社会组织、高等院校、教师自身等多方协同努力（韩锡斌等，2022）。国内学者赵磊磊等从外部因素(教学信念和职业认同)和内部因素(TPACK 的内部结构要素)两方面探讨了对地方高校信息技术师范生的 TPACK 能力的影响。结果表明，教学信念和职业认同对师范生的 TPACK 能力呈现出显著的强相关（赵磊磊等，2018）。技术压力对教师数字化教学创新行为与工作表现均有显著负向影响，且 TPACK 能力对教师教学创新行为与工作表现均有显著正向影响，技术压力与 TPACK 能力对师范生的数字化教学能力有直接的影响（冯仰存等，2023）。

综上所述，师范生数字化教学能力提升的影响因素可归纳为环境支持的外部因素影响和师范生的自我效能感和对技术感知能力等内部因素影响。目前关于师范生数字化教学能力提升的影响因素研究主要探究某一因素的直接影响与中介调节作用，很少有研究同时关注多重因素以及各因素之间对师范生数字化教学能力的影响机制。因此，本研究基于 UTAUT 模型理论探究哪些关键因素影响师范生数字化教学能力的提升，重点关注绩效期望、努力期望、社会影响、便利条件四个核心变量对于对于师范生数字素养的影响，以及各要素之间的关系，以更好地为实践提供指导。

1.3. 师范生数字化教学能力的影响因素研究

环境支持在智能技术支持的师范生数字化教学能力中发挥中重要的作用（周刘波等，2023）。随着人工智能、云计算、大数据等新兴技术的快速发展和广泛应用，提升师范生数字化教学能力的设施配备也应从数字化向智慧化转变，逐步构建高情境化、强临界感、多场景的智慧学习环境（洪玲，2023）。在师范生数字化环境建设方面，可以借助智能技术打造智慧化录播环境、智能教学实训室(例如 AI 智能实训静音舱)、智能综合的教师教育实验实训平台、基于大数据的教、学、管、研、评一体化的教学系统等软硬件环境。增设师范生实践观摩和体验的教学课时，促进师范生的沉浸式学习与实践，以强化真实或拟真情境下的实践应用，增加师范生数字化教学实践的机会，提升智能技术支持的数字化教学能力。伴随着智能技术的快速发展，智能技术与师范生培养深度融合已成为必然趋势。在师范生的培养过程中，借助智能技术支持的智慧学习空间环境，组织师范生开展智能设备的基本操作、基于学科的数字技术与教学深度融合、过程性数据的采集与分析等内容学习，提升智能技术支持的数字化教学能力（祝智庭等，2023）。师范生处于准教师的职前培养阶段，鼓励建立教师—师范生共同体协同发展机制，为师范生数字化教学实践提供过程性指导，实现细微精准、面向师范生实践能力的“智能新师范”的人才培养，从而促进师范生数字化教学能力的提升（苟江风等，2023）。

2. 理论模型构建及假设

2.1. 模型概述

整合技术接受模型理论（UTAUT）由 Venkatesh 等在整合理性行为理论、技术接受模型、动机理论、计划行为理论、TAM-TPB 模型、计算机使用模型、创新扩散理论和社会认知理论的基础上提出，认为个体技术采纳受绩效期望、努力期望、社会影响、便利条件四个核心变量影响，并受性别、年龄、经验和自愿性调节，是探究教师技术采纳影响因素的主流理论

模型之一（Venkatesh 等，2016）。其中，绩效期望是指个体认为使用技术能够提高工作绩效的程度，努力期望是指个体使用技术的轻松程度，社会影响是指个体感受到意见领袖或重要人物认为个体应当使用技术的程度，便利条件是指个体对支持技术使用的组织和基础设施条件的信任程度。本研究借鉴整合技术接受模型构建师范生化教学能力的影响因素模型，进一步进行模型验证与阐述。

2.2. 模型构建及假设提出

本研究结合师范生数字化教学能力的实际情况，师范生的专业实习经历有着密切联系，自愿性在绩效期望中予以分析，因此模型中不考虑自愿性的调节作用。另外，师范生通过数字技术能否取得相应的会对其数字化教学能力提升产生一定的影响，对原模型的调节变量做出相应的修正，最终构建教师数字化教学能力的影响模型如图 1 所示。

（1）师范生绩效期望越高，其数字意愿就越强，师范生若预期数字技术能提升教学效果、促进自身学习效率，将更积极地培养数字化教学能力。因此提出假设：H1：绩效期望正向影响师范生的数字意愿。

（2）数字技术的难易程度、操作的复杂程度将会影响教师数字意愿，当师范生认为掌握数字技术所需努力较低时，他们更有可能主动提升数字化教学能力。因此提出假设:H2：努力期望正向影响师范生的数字意愿。

（3）学校课程、教师示范、同伴影响等社会因素将推动师范生重视和发展数字化教学能力,提出假设：H3：社群影响正向影响师范生的数字意愿。

（4）为师范生拥有提升数字素养的学习条件，有足够的知识和能力开展数字化教学，学校能够出台数字化教学制度、建立数字化资源支持实践教学。便利条件越充分，越有利于师范生主动提升数字化教学能力。因此提出假设：H4：便利条件正向影响师范生的数字行为。

（5）调节变量关系：师范生数字素养水平与性别有关；师范生数字化教学能力与教育实习经历有关；师范生数字化教学能力水平与就读专业有关。

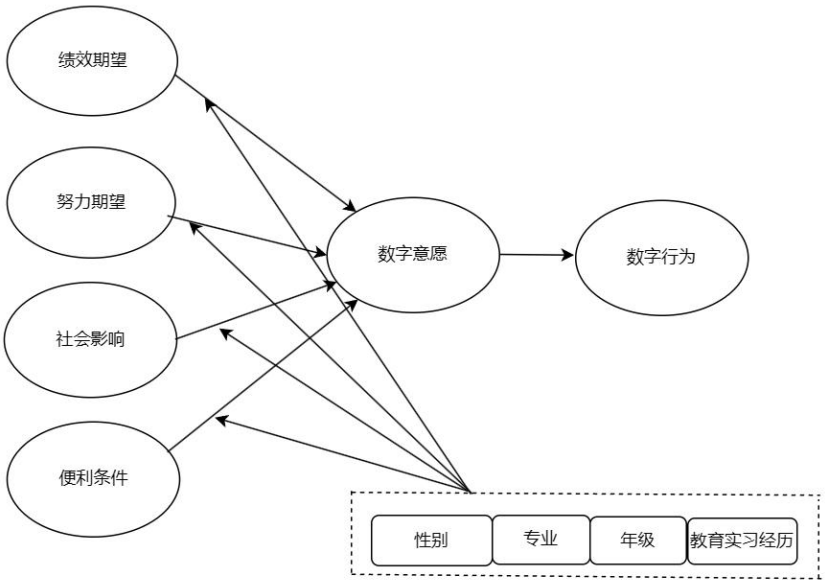


图 1 师范生数字素养影响因素关系模型

3.影响因素模型实证验证

3.1. 研究对象

研究对象为本文选取了山东省 L 大学的师范生作为研究样本，涵盖语文、数学、英语、小学教育、音乐等 13 个师范类专业，通过问卷星进行在线调查，本次共获得调查问卷 543 份，对于填写时间过短、未通过甄别题项等无效调查问卷进行筛除，最终得到有效问卷 519 份,排除无效问卷后，有效问卷共 517 份，有效率达 95%（详情见表 1）。

表 1 样本信息统计

变量	选项	频率	百分比
性别	男	237	46%
	女	280	54%
年级	大一	100	19%
	大二	223	43%
	大三	109	21%
	大四	85	17%
专业	文科	240	47%
	理科	174	34%
	艺术类	103	19%

3.2. 研究工具

考虑到师范生双重身份，一方面作为学生，需要具备“数字化学习与研究”的能力，另一方面师范生作为未来教师需要具备教师工作信息化教学的能力。问卷由数字化教学能力的水平和影响因素调查两部分内容构成，第一部分量表设计以教育部 2022 年发布的《教师数字素养》为主要依据，改编自闫寒冰等人编制的《师范生信息技术应用能力量表》针对师范生信息化教学能力的量表（闫寒冰等，2018），和彭红超等人编制的《中小学教师数字素养量表》改编而成适用于调查师范生数字化教学能力水平的问卷，从数字化意识、数字化技术与知识、数据素养、数字化学习与创新和数字化应用五个维度了解师范生的数字化教学能力水平（彭红超等，2024）。第二部分以 UTAUT 理论模型为指导，根据师范生双重身份的特点编制了包括绩效期望、努力期望、社会影响、便利条件四个关键影响因素的问卷，同时问卷基础信息包含性别、年级、专业、实习经历四个调节变量调查。本量表采用 Likert5 点式计分法，分值从 1（非常不同意）至５(非常同意），得分越高，反映师范生的数字化教学能力越高，影响程度越深。

3.3. 数据收集与分析

运用 SPSS 27 和 AMOS 24.0 软件进行数据分析。首先，对问卷进行信效度检验，确保数据的可靠性和有效性。然后，通过描述性统计分析了解样本的基本特征和各变量的分布情况。接着，运用结构方程模型（SEM）分析方法对研究模型进行验证，检验假设关系的显著性，并分析各因素对师范生数字化教学能力的影响路径和效应大小。

3.3.1. 问卷信度分析

信度表示问卷的一致性，一般采用克隆巴赫系数 Cronbach’s α 分析方法法，系数越大，表示问卷中题目间的相关性越强。在测试中通过信度检验问卷中题目间的同质性，删除某题目后信度系数比原来高很多，则需要删除该题目。信度 Cronbach α 系数值为 0.985，大于 0.9，因而说明研究数据信度质量很高，具体分析可见表 2。针对“CITC 值”，分析项的 CITC 值均大于 0.4，说明分析项之间具有良好的相关关系，同时也说明信度水平良好。综上所述，研究数据信度系数值高于 0.9，综合说明数据信度质量高。

表 2 量表信度分析

Cronbach 信度分析		
项数	样本量	Cronbach α 系数
35	187	0.985

3.3.2. 问卷效度分析

效度是检验师范生数字化教学能力及其影响因素问卷测量结果有效性的指标，以此分析正式问卷调查数据能够检验问卷的题目合理性的程度。效度是检验师范生数字化教学能力及其影响因素问卷测量结果有效性的指标， 以此分析正式问卷调查数据能够检验问卷的题目合理性的程度。本研究中的问卷题目均来自前人验证过的研究量表，并结合了 18 名教育技术学的硕士生和 33 名不同专业的师范生的意见形成的，最后向教育技术学相关方向的 3 名专家咨询意见后修订而形成的正式问卷，因此可以认为本研究中的问卷的内容效度是符合标准的。结构效度检验通常采用 KMO 测量和 Bartlett 球形检验法进行。这两种方法可用于评估问卷的因子结构效度。KMO 测量：KMO 测量衡量的是因子分析中变量间的相关性。KMO 值的范围在 0 到 1 之间，通常认为，KMO 值大于 0.7 表示因子分析是可行的。Bartlett 球形检验法：Bartlett 球形检验用于检验变量之间是否存在相关关系。该检验基于一个假设，即所研究的变量之间存在相关性。如果检验结果显示拒绝了零假设（ρ 值小于设定的显著性水平 0.005），则可以认为变量之间存在相关性，适合进行因子分析。利用 SPSS 27.0 对问卷数据进行 KMO 和 Bartlett 标准检验，KMO 值为 0.893，大于 0.7，故问卷适合做因子分析。Bartlett 球形检验显著性概率为 0.000，远小于 0.005，故变量相关性很好，所以可以进行探索性因子分析，具体可见表 3。

表 3 结果效度分析

KMO 和 Bartlett 的检验		
KMO 值		0.893
Bartlett 球形度检验	近似卡方	5418.294
	df	595
	p 值	0.000

同时，每个题项与相对应因子的载荷系数都大于 0.6，说明题项和因子的对应关系较强，这一结果说明维度内部的聚合效度达标。通过计算各题项的标准载荷系数,可以得到各维度的 AVE 值和 CR 值。每个题项与相对应因子的载荷系数大于 0.6，各维度的 AVE 值和 CR 值均达标（绩效期望：AVE = 0.624，CR = 0.924；努力期望：AVE = 0.578，CR = 0.915；社会影响：AVE = 0.562，CR = 0.894；便利条件：AVE = 0.675，CR = 0.872；数字意愿：AVE = 0.589，CR = 0.916；数字化行为：AVE = 0.643，CR = 0.874），表明维度内部聚合效度良好。

3.3.3. 描述性分析

根据调查结果显示：师范生数字化教学能力总体平均得分为 3.82（满分为 5 分）。陈云杰将师范生数字化教学能力为四个等级，得分≥4.5 为“优秀”，4.0≤得分<4.5 为“良好”，3.5≤得分<4.0 为“一般”，得分<3.5 为不足（陈云杰，2018）。其中，数字化意识和数字社会责任得分较高,而数字化学习与创新和数字化应用得分为 3.25，相对较低，表明师范生在数字化意识与社会责任方面掌握较好，但在数字化学习与教学应用水平低，特别是数字化应用方面有待加强。理科生的数字技术自我成就感会高于文科生，对于数字技术的接受度和使用频率较高，解决问题的方式也更偏向于技术。大四年级有教育实习经历的学生数字化教学能力明显高于其他年级的学生，这也说明教育实习经历对师范生的数字化教学能力的提升有重要的影响。在教育实习过程中，师范生会参与到实际的课堂教学环节，实际需要促使他们利用数字化工具和资源去解决教学问题。在备课阶段，需要将文字、图片、视频等多种数字资源整合到教学课件中，这促使他们熟练掌握课件制作软件（如 PowerPoint、Prezi 等）的高级功能。实践操作让他们能够适应数字化教学环境，提升对各种数字工具的应用熟练度，相应提高他们的数字化教学能力。师范生对使用数字技术提升学习和教学效果有一定的期望，同时也认识到学习数字技术需要付出一定努力；社会环境对师范生数字化教学能力发展有一定影响，且学校所提供的数字化资源设备和数字学习条件不能满足需求。

3.3.4 模型拟合度检验

根据建立的 UTAUT 理论模型以及相关假设，绘制出影响因素关系模型。运用 Bootstrap 法验证模型适配度：$\chi^2/df=2.13$（<3），RMSEA=0.047（<0.08），CFI=0.971，TLI=0.963，模型拟合良好。本次调查问卷收集的数据可以使用该模型估计变量间的影响关系，分析结果具有良好参考性。

4.培养路径

4.1. 优化课程体系，注重数字技术与学科教育的融合

师范生的培养和教学过程中，既要实现学生基本数字技能的熟练掌握运用，也要强调数字技术与教育教学课程设置的深度融合，推进师范生学科化数字教学技能的培养与实践。只有把学科知识与数字技能有效整合，才可能使数字化教学能力真正落地生根于教学工作中（周彬，2021）。在不同专业教学法课程中融入数字素养教育，渗透数字技术应用。另外，由于不同学科的知识结构、教学目标、教学方法不同，师范生应能根据未来所教授科目得特点选择适切的数字技术进行数字化教学创新评价，落实学科核心素养才是技术支持教学得关键所在，把数字技术与资源与教学融合（赵云建，2024）。同时，设置多样化拓展实操性强一点的课程，如多媒体课件高级制作、微课制作、在线教学平台深度应用、教育大数据分析等，满足学生不同兴趣和专业需求，拓展数字技能。

4.2. 加强数字技术实践教学，提升数字应用能力

师范生处于准教师的职前培养阶段，存在的普遍现象表现在基于真实场景的教学实践经验不足，为解决师范生的数字化教学能力培养过程中学生积极性不高的问题。教师教育是一种动作性技能，它是经验、实践引起的个体知识或行为的相对持久变化，需要在实践中不断反复练习来获得（梁瑞等，2023）。培养师范生主动运用数字技术解决实际学习、教学实践过程中的“真”问题，在实践中应用数字化工具解决教学问题。师范院校要加强数字技术的教学实践训练，在数字技能的反思性实践中，推动智能教育理论与教学实践的融合运用。在师范生的培养过程中，借助智能技术支持的智慧学习空间环境，以“数字化教学课程如何设

计、数字化教学评价如何进行、师生在数字化教学中互动如何高效实现”为主线，组织师范生开展智能设备的基本操作，提升智能技术支持的数字化教学能力（祝智庭等，2023）。同时，与当地区域内中小学合作搭建校外实践基地，让师范生在实际教学场景中应用数字技术，积累实践经验。结合实习期间的教学案例，引导学生在真实的教学场景中运用数字技术，培养其教学实践能力、分析和解决问题的能力，培养其创新思维和团队合作能力（薛爱峰等，2024）鼓励学生参与教育实习中的数字化教学项目，将所学数字技术应用于教学实践，通过实践反思不断提高数字应用能力。

4.3. 推进协同合作，整合数字化教学能力培养资源

在提升职前教师数字化教学能力的过程中，构建兼具协同性与开放性的数字化教学能力培育生态尤为关键。鼓励建立教师—师范生共同体协同发展机制，为师范生数字化教学实践提供过程性指导，实现细微精准、面向师范生实践能力的“智能新师范”的人才培养，从而促进师范生数字化教学能力的提升（苟江凤等，2024）。培养单位应建立校际合作联盟，加强与中小学学校之间在师范生数字化教学能力培养方面的合作与交流。开展校际联合培养项目，互派师生交流学习，共享优质课程资源和实践教学基地。学校联合举办数字化教学能力培养研讨会、教学观摩活动、师范生技能竞赛等，促进校际之间经验分享和资源互补，共同探索创新培养模式和方法，提升区域内师范生数字化教学能力整体水平。为师范生连线搭桥，倡导与一线教师在数字化教学能力培养过程中的协同发展，鼓励一线教师指导师范生参与数字教学项目和科研活动，共同探索数字技术在教育教学中的应用。

5.结论与展望

本研究基于 UTAUT 模型，通过对山东省 L 大学师范生的调查分析，验证了绩效期望、努力期望、社会影响和便利条件对师范生数字化教学能力的显著正向影响，以及各因素间的复杂相互关系。基于影响因素分析提出的培养策略，对高校培养师范生数字化教学能力提供了新思路，以期提升师范生数字化教学能力。然而，本研究存在一定局限性，如样本仅来自一所大学，可能存在一定局限性。未来研究可扩大样本范围，涵盖不同地区、类型院校，进一步验证模型普适性；深入探讨调节变量对数字化教学能力的具体影响机制；跟踪师范生毕业后数字化教学能力在教学实践中的应用效果，为持续改进师范生数字化教学能力培养提供更有力依据。

参考文献

中共中央国务院关于弘扬教育家精神加强新时代高素质专业化教师队伍建设的意见[EB/OL].(2024-08-26)[2024-08-28].http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202408/t20240826_1147269. html.

教育部.教师数字素养：JY/T 0646-2022 [S].北京：中华人民共和国教育部 2022，1.

赵磊磊,李玥泓（2018）.地方高校信息技术师范生 TPACK:影响因素及建构策略[J].现代远距离教育,2018,(3):29-36.

冯仰存,吴佳琦,曹凡（2023）.技术压力对教师数字化教学创新的影响研究——成长型思维、TPACK 的调节效应[J].中国远程教育, 2023,43(06):29-38.

闫寒冰,李笑樱,任友群（2018）.师范生信息技术应用能力自评工具的开发与验证[J].电化教育研究,2018,39(01):98-106.

彭红超,朱凯歌（2024）.中小学教师数字素养测评问卷的本土化构建. 现代远程教育研究,2024,36(5)
陈云杰（2018）.师范生信息化教学能力现状分析与提升对策研究——以陕西师范大学为例[D].西安:陕西师范大学.
周彬(2021). 教师教育专业知识———生成、积累与课程转化[J]. 教育研究, (7) : 37－47.
赵云建(2024).指向教师数字素养提升的课堂技术应用观测框架、现状诊断及改善路径[J].中国电化教育,2024,(11):102-109.
梁瑞,苏君阳和杨聚鹏（2023）.人工智能时代师范院校教师教育的发展机遇、现实困境及突破路径[J].教师教育研究, (3) : 25－31.
祝智庭,林梓柔等（2023）.教师发展数字化转型:平台化、生态化、实践化[J]. 中国电化教育,(1):8-15.
薛爱峰,方芳（2024）.英语师范生数字化教学能力框架结构、基本样态与培养策略[J].中国电化教育,(12):123-129.
苟江凤,杨九民和李荣娜（2024）.智能技术支持的师范生数字化教学能力提升：期望价值信念的作用[J].中国电化教育,(08):56-61+86.
GILSTER P（1997）.Digital literacy[M]. New York：Wiley Publishing：25-48.
Venkatesh V, Thong J Y L, Xu X(2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead[J]. Journal of the Association for Information Systems,(5):328-376.

人工智能教育应用的伦理问题及其治理对策

AI ethics in AI Educational Applications and their Governance Strategies

于越[1]，赵建华[2*]
[1]西北师范大学教育技术学院
[2]南方科技大学未来教育研究中心
* zhaojh@sustech.edu.cn

【摘要】 当今社会，人工智能的迅猛发展对人类生产生活产生了深远影响，人工智能在教育领域的探索与应用方兴未艾，与此同时，人工智能在教育应用中的伦理问题也愈加突出。人工智能伦理概念的形成，历经科幻文学、问题浮现与机器人伦理、理论探讨、规范制定阶段，国际上在人工智能伦理问题和治理原则已经达成了广泛共识。近年来，人工智能已广泛在教育中应用，出现了数据隐私泄漏风险、算法偏见与歧视、学术诚信危机、能力缺失和心理健康等问题。人工智能教育应用中伦理问题的治理，需要政府、教育组织、学校等多方共同参与。在顶层设计层面，要制定人工智能教育应用的法律规范；在教育组织层面，要制定人工智能教育应用的政策指南和指标体系；在学校层面，要积极制定指导性文件开展人工智能伦理教育。
【关键词】 人工智能教育应用；人工智能伦理；伦理治理

Abstract: *Nowadays, the rapid development of artificial intelligence (AI) has had a profound impact on human production and life. The exploration and application of AI in the field of education are burgeoning, while at the same time, ethical issues in its educational applications have become increasingly prominent. The formation of the concept of AI ethics has gone through stages of science fiction literature, problem emergence and robot ethics, theoretical exploration, and the formulation of norms. Internationally, significant consensus has been reached on AI ethical issues and governance principles.In recent years, AI has been widely applied in education, giving rise to issues such as risks of data privacy breaches, algorithmic bias and discrimination, academic integrity crises, skill deficiencies, and mental health concerns. Addressing the ethical challenges in AI educational applications requires the collective participation of multiple stakeholders, including governments, educational organizations, and schools. At the top-level design stage, it is essential to establish legal regulations for AI applications in education. At the educational organization level, policy guidelines and indicator systems for AI in education should be developed. At the school level, it is crucial to actively formulate guiding documents and implement AI ethics education.*
Keywords: AI Educational Application, AI ethics, AI ethics Governance

1.引言

当今社会，人工智能的迅猛发展对人类生产生活产生了深远影响，尤其是随着 ChatGPT 和 Deepseek 的爆火，人工智能在教育领域的探索与应用方兴未艾，受到社会的高度关注。一方面，人工智能的合理应用在推进因材施教、促进教育公平、提升教育质量、破解教育难题、支持终身学习、构建高质量教育体系等方面展现出显著优势；另一方面，人工智能的不当应用也带来价值观、道德伦理、隐私保护、学术诚信、偏见与歧视等社会问题，还涉及知识产权侵犯、数据安全和法律合规风险。人工智能是指用机器模拟、实现和延伸人类的感知、

思考、行动等智力与行为能力的科学与技术(中国信通院，2021)，主要包括自然语言处理、语言识别、图像识别和处理、情绪检测、人工创作等（UNESCO, 2021）。本研究中的人工智能主要指生成式人工智能，即具有文本、图片、音频、视频等内容生成能力的模型与相关技术（生成式人工智能服务管理暂行办法，2023）。据《纽约时报》报道，2024 年 2 月，一名 14 岁的美国男孩痴迷聊天机器人自杀身亡（Roose, 2024），让人工智能在教育应用中的伦理问题备受关注。本研究将对人工智能教育应用中的伦理问题及其治理现状展开论述。

2.人工智能伦理的定义

2.1. 人工智能的由来

人工智能（Artificial Intelligence，简称 AI）之父，英国著名学者阿兰图灵在 1950 年发表的一篇划时代意义的论文《计算机器与智能》中提到“机器是否能够具有思维”的重大问题。图灵的论文直接引发了人类关于人工智能的思考（Turing, 1950）。1956 年，麦卡锡、西蒙等科学家在美国达特茅斯学院召开的研讨会上正式提出了人工智能这个概念，随后，经历 IBM 深蓝智能计算机打败人类国际象棋冠军、谷歌公司的 AlphaGO 打败人类围棋世界冠军李世石、ChatGPT 及具备强大推理能力的 o1 问世等标志性事件(赵建华，2024)，人工智能对社会各行各业产生了深远的影响。虽然人工智能这一概念目前已经为社会和科学领域所广泛接受，但是无论在工业界还是学术界，不同发展时期、不同领域和行业、不同学科的学者、专家以及社会各界人士对人工智能仍有不同角度和层面的理解。关于人工智能的各种定义可以归结为：人工智能是研究智能的机制和规律、构造智能机器的技术和科学。通俗来讲，人工智能可以被看作是一种让计算机系统具备类似人类智能的技术，它使机器能够像人类一样思考、学习和解决问题，甚至在某些领域超越人类的能力，人工智能是研究如何使机器具有智能的科学。

2.2. 人工智能伦理的由来

人工智能伦理的概念形成是一个多维度、跨学科的过程，其发展历程可以概括为以下几个关键阶段和因素。

2.2.1. 科幻文学的人工智能伦理（19 世纪初-20 世纪中）

人工智能所引发的伦理思考，最早并不是现实的技术进步，而是科幻小说。早期科幻作品已开始探讨机器与人类的伦理关系，1818 年玛丽雪莱创作的《弗兰肯斯坦》、1920 年卡雷尔卡佩克发表的《罗萨姆的万能机器人》以及 1940 年阿西莫夫在科幻小说《我，机器人》中提出了“机器人三定律”。

2.2.2. 从问题浮现到机器人伦理（20 世纪中期至 21 世纪初）

人工智能的概念在 20 世纪 50 年代诞生，早期研究集中在逻辑推理和基础算法。人工智能之父图灵在早年论文《智能机器》中提到了人工智能迟早会威胁到人类的生存。20 世纪 60 年代，美国哲学家休伯特德莱弗斯将人工智能专家称作“炼金术士”，呼吁人工智能专家在产品设计和创造时融入人类社会的哲学思考。这一时期，人工智能伦理也成为一批科幻小说和影视作品的主题思想，比如《银翼杀手》《黑客帝国》。进入 21 世纪，关于机器人伦理学、法律和社会问题在学术和专业方面得到重视和研究，并在 2004 年第一届机器人伦理学研讨会上提出了“机器人伦理学”的概念。2005 年，欧洲机器人研究网络设立“机器人伦理学研究室”。欧盟建立机器人伦理学 Atelier 计划，也就是“欧洲机器人伦理路线图”。

2.2.3. 人工智能伦理的理论探讨（21 世纪初-2014 年）

随着计算能力的提升，AI 逐步从理论走向应用。从 2005 年以后，机器人伦理学逐渐延伸到人工智能伦理，并逐渐受到全球各界专家、学者以及政府和企业的关注。计算机科学、哲学、伦理学等领域的学者开始跨学科研究合作，探讨 AI 的潜在伦理挑战。出现算法的公平性以及数据偏见导致决策歧视，如在招聘、信贷中的算法偏见；自主性与责任问题，比如自动驾驶汽车在事故中的责任划分问题。一些新的理论框架构建，功利主义、义务论等传统伦理学理论被重新解读，以适配 AI 场景。

2.2.4. 人工智能伦理的规范制定（2015 年至今）

在这一时期，谷歌微软等科技巨头发布 AI 伦理原则，强调公平、透明、可解释性等。各国政府陆续出台一系列政策法规。经合组织（OECD）、联合国教科文组织（UNESCO）等推出全球性 AI 伦理框架。根据联合国教科文组织在 2021 年发布的《人工智能伦理建议书》，人工智能伦理（AI Ethics）是确保 AI 技术以尊重人权、多样性、环境可持续性和社会正义的方式发展的实践框架(UNESCO, 2021)。

3.人工智能伦理教育应用中的主要挑战

3.1.数据隐私泄漏风险

人工智能收集学生生物特征、学习行为、情绪状态等敏感数据可能侵犯隐私，且数据使用缺乏透明度。《2024 年数据泄露调查报告》显示，30,458 起数据安全事件，其中 10,626 起已确认数据发生泄露，教育行业（1537 起）成为年度数据泄露事件发生最严重的领域，内部误操作、错误配置、外发泄露、外部攻击者的勒索入侵、漏洞攻击等问题交织发生，是该行业数据泄露的元凶(Verizon, 2024)。在疫情期间广泛用于远程考试监控的 Proctorio 在线监考软件，通过摄像头、麦克风及屏幕录制跟踪学生行为，甚至分析眼球运动和背景环境，学生投诉其收集过多数据且未明确告知用途（Monica Chin, 2021）。一些学校也开始“试水”用人脸识别来保证出勤率和提高听课效率，人脸识别系统进入中国药科大学课堂后学生担忧，此举识别有侵犯个人隐私之嫌，一举一动都被人脸摄像头监视（于珍，2019）。

3.2.算法偏见与歧视

生成式人工智能技术的开发、训练和更新取决于海量的数据，训练数据中的偏见导致系统推荐资源或评估学生时存在歧视，出现数字贫困、性别歧视，加剧教育不公平问题。训练数据和算法技术的来源大多源自欧美发达国家的科技公司，同时训练的数据大多也是基于英文，教科文组织将这种行为称为“数字贫困”（Digital Poverty），认为这将导致发展中国家的价值观、文化习俗、行为规范受到西方文化的侵袭，由此产生知识的同质化限制多元化和创造性思维(UNESCO, 2023)。2020 年英国政府使用算法调整 A-level 考试成绩，导致弱势地区学生成绩被系统性压低，引发大规模抗议（A-Levels, 2020）。

3.3.学术诚信危机

生成式人工智能能够通过与学习者在多轮对话情境下生成具有个性化的知识，但更可能加剧作弊和学术不端的诚信危机。教育数据公司麦可思发布的 2024 年中国高校师生生成式 AI 应用情况研究显示，国内高校师生几乎都曾在学习和工作中使用过生成式 AI，从未使用过生成式 AI 的高校师生比例仅为 1%。其中，近六成高校师生频繁使用生成式 AI(赵丽 & 陈颖， 2024)。在国际上，已经有诸多国家与地区指出智能技术将可能加剧课业、考试的作弊和抄袭风险，将限制该技术的使用。

3.4.能力缺失和心理健康

使用人工智能进行学习，可能导致个体不再需要通过和教师与同伴在社会互动和实践中接受教育，从而弱化教育促进学生社会交往和集体学习的重要性，限制学生必要的社会能力获取，如表达技巧、社会协作与问题解决能力。教科文组织在《报告》中提出，这将造成学习者缺乏对真实世界的观察，通过实践获取直接经验的能力，以及与他人交流、合作和处理现实世界中问题的能力(UNESCO, 2023)。同时，人工智能监控学生情绪或学习状态可能加剧学生焦虑，甚至通过算法推送内容操控行为，可能出现引发更大的心理健康问题，出现更多类似于 14 岁的美国男孩痴迷聊天机器人自杀身亡的事件（Roose, 2024）。

4.人工智能教育应用伦理的治理对策

通用大模型时代，生成式人工智能的教育应用不断“制造”出多种新型伦理交往和道德生活场景，成为教育人工智能伦理治理的重心。只有伦理先行，才能为人工智能与教育融合创新提供保障，并且要通过敏捷治理的理念、伦理教育的手段和伦理审查方案的制定三者相辅相成，提供强有力的支撑，需要政府部门、教育组织、学校、社会团体等多方共同参与。

4.1.顶层设计：制定人工智能教育应用的法律规范

首先，在国家层面应该尽快出台人工智能教育应用的政策和执行标准，明确教育行业的发展方向和规范化要求，具体内容应该包括技术开发、数据保护、教学内容和课程设置。通过比较美国、欧洲、亚洲及联合国教科文组织等人工智能教育应用政策，发现总体上有很多共同之处，在应用中注重谨慎、态度比较乐观、坚持决策自主(吴河江&吴砥，2024)。

4.2.教育组织：制定人工智能教育应用的政策指南和指标体系

从 2019 年到 2024 年，联合国教科文组织出台了《关于人工智能与教育的北京共识》《人工智能伦理问题建议书》《生成式人工智能在教育和研究中的应用指南》《学生人工智能能力框架》《教师人工智能能力框架》等近 10 份人工智能教育应用指南和能力框架。联合国教科文组织形成了“赋能”“比较”“协同”“参与”四种教育数字治理策略，建立了治理主体多元、治理方式刚柔相济和治理过程全周期的治理机制(朱莉， 2024)。各国各组织教育人工智能伦理原则参照已有原则达成了很多“共识”的构建取向，如隐私、公平、安全、问责等（白钧溢&于伟，2023）。在目标层面上指导政策的实施和评估，在实施层面上鼓励各方的协同参与，在方案层面上构筑数据的隐私屏障，可以打破目前技术应用中存在的伦理困境。

4.3.学校行动：积极制定指导性文件开展人工智能伦理教育

人工智能伦理教育的开展需要学校与社会、家庭密切配合。美国麻省理工学院（MIT）在青少年人工智能伦理教育领域的探索，涉及社会、家庭和学校三个方面，包含了内嵌式和专题式两种类型，主要的伦理教育形式有与中小学合作的线下 STEAM 选修课程、线上线下夏令营相关活动以及在社会上招募学生参与的混合式工作坊(李艳等，2024)。罗素大学集团内的成员高校对生成式人工智能教育伦理风险的总结为数据陷阱冲击教育管理系统、算法支配损害教师主体权威、智能依赖造成学生学习畸化三个方面，并且率先积极应对、及时发布生成式人工智能应用指南，从教学评估监管、教师教学配套、学生使用培训三个方面的具体举措（张惠彬&许蕾，2024）。哈佛大学、耶鲁大学、剑桥大学、东京大学、南洋理工大学、悉尼大学等 16 所 2024 年 QS 世界大学排名前 100 的世界一流高校在生成式人工智能应用方面的指导性文件，提炼出技术应用、道德伦理、课程评价、师生指导四个主题（楚肖燕等，2024）。

5.结语

人工智能在教育领域的应用为教学创新和个性化学习带来了前所未有的机遇，但同时也伴随着数据隐私、算法偏见、学术诚信等伦理挑战。确保人工智能教育的健康发展，需要多方协同治理：政府应完善法律法规以提供制度保障，教育组织需制定伦理准则和评估体系，学校则应加强师生的人工智能伦理教育。唯有在技术创新与伦理约束之间取得平衡，才能让人工智能真正成为促进教育公平、提升学习效能的可持续力量。未来，随着技术的演进，伦理治理框架也需动态调整，以应对新兴问题，最终实现“以人为本”的智能教育生态。

参考文献

白钧溢 & 于伟. (2023). 超越“共识”：教育人工智能伦理原则构建的发展方向. 中国电化教育, 6, 9 – 17, 24.

楚肖燕, 沈书生, 王敏娟, 王会军, 李晓文, & 翟雪松. (2024). 世界一流高校探索生成式人工智能应用规范的经验及对我国的启示——基于 LDA 主题模型分析的文本挖掘. 现代远距离教育, 3, 38 – 47.

李艳, 朱雨萌, & 樊小雨. (2024). 青少年人工智能伦理教育的探索及启示——以 MIT 为例. 现代远距离教育, 1, 3 – 13. https://doi.org/10.13927/j.cnki.yuan.20240022.001

生成式人工智能服务管理暂行办法, Pub. L. No. 国家互联网信息办公室 中华人民共和国国家发展和改革委员会 中华人民共和国教育部 中华人民共和国科学技术部 中华人民共和国工业和信息化部 中华人民共和国公安部 国家广播电视总局令第 15 号 (2023). https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm

吴河江 & 吴砥. (2024). 生成式人工智能教育应用:发展历史、国际态势与未来展望. 比较教育研究, 46(6), 13 – 23. https://doi.org/10.20013/j.cnki.ICE.2024.06.02

于珍. (2019). 中国教育报：辅助教学还是泄露隐私，AI 进校园边界在哪里？_教育家_澎湃新闻-The Paper. https://www.thepaper.cn/newsDetail_forward_4758874

张惠彬 & 许蕾. (2024). 生成式人工智能在教育领域的伦理风险与治理路径——基于罗素大学集团的实践考察. 现代教育技术, 34(6), 25 – 34.

赵建华. (2024). 人工智能时代的教育转型与重塑. 电化教育研究, 45(12), 37 – 43, 97. https://doi.org/10.13811/j.cnki.eer.2024.12.005

赵丽, & 陈颖. (2024). 媒体：从“一键生成作文”到数据泄露隐患，学生用 AI 边界何在？_教育家_澎湃新闻-The Paper. https://www.thepaper.cn/newsDetail_forward_30601968

中国信通院. (2021). 《人工智能治理白皮书》. https://www.sohu.com/a/www.sohu.com/a/421423992_735021

朱莉. (2024). 联合国教科文组织教育数字治理的机制与策略——基于 44 份文本的 Nvivo 质性分析. 比较教育学报, 5, 31 – 44.

A-levels: Anger over 《unfair》 results this year. (2020, 八月 12). https://www.bbc.com/news/education-53759832

Monica Chin. (2021). University will stop using Proctorio remote testing after student outcry | The Verge. https://www.theverge.com/2021/1/28/22254631/university-of-illinois-urbana-champaign-proctorio-online-test-proctoring-privacy

Roose, K. (2024, 十月 23). Can A.I. Be Blamed for a Teen’s Suicide? The New York Times. https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html

UNESCO. (2021). 人工智能与教育：政策制定者指南. https://unesdoc.unesco.org/ark:/48223/pf0000378648

UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000381137

UNESCO. (2023). Guidance for generative AI in education and research — UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000386693.locale=en

Verizon. (2024). 4900万份客户数据遭泄漏；教育行业既成为年度数据泄露事件发生最严重的领域？_科学技术_人工智能_信息. https://www.sohu.com/a/www.sohu.com/a/780623261_121728280

教育数字化转型背景下校长信息化教学领导力影响因素研究

A Study on the Influencing Factors of Principals' Informatization Leadership in the Context of Educational Digitalization Transformation

杨文静[1*]，康健[1]，赵建华[2]
(1.西北师范大学，教育技术学院)
(2.南方科技大学，未来教育研究中心)
[*]kjywj0911@163.com

【摘要】 数字时代的学校治理呼吁全面提升校长数字治理水平和协同创新能力，校长信息化教学领导能力对学校信息化发展具有关键影响作用。为探讨数字化转型背景下校长信息化教学领导力的影响因素，构建校长信息化教学领导力影响因素模型，本研究基于技术接受与整合模型，采用问卷调查法收集260名中小学校长数据，并通过结构方程模型进行实证检验。研究发现，绩效期望、促成条件、个体创新对校长信息化教学领导力有正向显著影响。因此，本研究从政府主导完善校长培养资源与制度的建设、校长专业发展和校长教育培训方案"数字化"方面提出校长信息化教学领导力优化建议，为教育数字化时代校长信息化教学领导力的培养与研究提供可参考的依据。

【关键词】 教育数字化转型；校长；信息化教学领导力；影响因素

Abstract: *School governance in the digital era calls for comprehensive improvement of principals' digital governance skills and collaborative innovation capabilities. Principals' leadership in information-based teaching is a key influence on the development of schools' information technology. In order to explore the influencing factors of principals' information-based teaching leadership under the background of digitalization transformation and to construct a model of influencing factors of principals' information-based teaching leadership, this study collected data from 260 primary and secondary school principals through a questionnaire survey based on the technology acceptance and integration model, and conducted empirical tests through a structural equation model. The study found that performance expectations, enabling conditions, and individual innovation have a positive and significant impact on principals' information-based teaching leadership. Therefore, this study puts forward suggestions for optimizing principals' information-based teaching leadership from the aspects of government-led improvement of principal training resources and systems, principals' professional development, and the "digitization" of principals' education and training programs. This study provides guidance for the cultivation of principals' information-based teaching leadership in the era of educational digitalization, and offers some insights into related research.*

Keywords: Digital transformation of education; principal; information teaching leadership; influencing factors

1.引言

教育数字化转型以培养和形成学生、教师、教育管理人员的数字化思维、数字化意识和能力为目标，以数据驱动为核心，推动智慧教育生态发展及建立数字教育治理体系，全方位进行教育变革。校长作为学校教育数字化改革发展的引领者，在教育数字化发展进程中具有

关键作用，其信息化教学领导力水平高低是推动学校教育数字化改革，破解学校教育数字化转型问题的关键因素。2022 年《关于实施新时代中小学名师名校长培养计划(2022-2025)》指出，将数字化教育融入校长培养全过程，帮助校长凝练教育理念，提升教育教学、办学治校能力，旨在培养一批具有创新教育理念和成熟教学模式，并能够引领基础数字教育教学改革发展的名师名校长(教育部，2022)。在教育数字化转型背景下，如何发展与培养校长信息化教学领导力，更新校长传统教育理念，以数字技术助推学校教学创新变革，是现今学校教育数字化变革的重要问题。因此，有必要探索和研究教育数字化转型背景下校长信息化教学领导力的影响因素，以此分析校长信息化教学领导力发展的现实困境，以期为我国校长信息化教学领导力的发展与培养提供参照。

2.校长信息化教学领导力

2.1. 校长信息化教学领导力的内涵

校长信息化教学领导力是指校长利用新技术、新理念推动学校教育变革，指导、规划、组织、管理、评价学校信息化教学，引领全校教师围绕推动技术与教育教学的深度融合，通过教研、教学模式创新重构教学结构和教学流程的过程，同时在此过程中提升自身数字化治理能力的一种智慧。教育界较为认可的校长教学领导力内涵是校长促进教师教和学生学的教育领导行为，包含定义学校使命、管理教学计划、营造积极的学校学习氛围三个方面(Hallinger，1985)。从学术界已有研究看，校长信息化教学领导力有“能力观”和“过程观”两个方面的观点。从“能力观”看，例如，Kannan 等认为校长信息化教学领导力包括表现出一种热情的承诺，是确保教师能够获得专业发展，培养其在信息和通信技术方面的技能的一种能力(Kannan，2013)。也有研究者认为校长信息化教学领导力指校长监督学校课程，并作为学习型组织者有效改进学校教学计划的能力(Mestry，2017)。从“过程观”看，杨鑫等认为校长信息化领导力是学校信息化硬件设施、资源等得到基本保障的前提下,对校长角色的重新认识和对信息化领导力的深层次思考的过程(杨鑫，2018)。赵磊磊等认为是校长对信息化教学环境建设、课外学习引领和课堂教学管理等的一系列领导过程(赵磊磊，2019)。此外，从“能力”和“过程”两方面界定校长信息化教学领导力。即校长带领学校教师开展信息化教学并实现学校教学发展目标的能力和校长理解信息化教学内涵及要素，并引领全校师生信息化教学,最终促进“数字一代”学习者创新学习的过程(董艳，2015)。

2.2. 校长信息化教学领导力影响因素

校长信息化教学领导力作为校长信息化领导力的核心组成部分，其重要性日益凸显，并与校长信息化领导力形成了紧密的内在联系。这一领导力的培育与提升，既面临着来自传统教育模式的固有阻力，也迎来了数字技术赋能带来的新机遇。一方面，校长信息化教学领导意识的模糊、信息化能力发展不均衡，数字素养不足以支撑学校教育数字化治理的新要求等(Samarraie，2018)，这些因素共同构成了校长在教育数字化改革中寻求突破的显著障碍。同时，教育数字化时代为教育系统提供了更为优越的基础设施，如灵活的混合学习环境、开放的教育资源、丰富的技术支持等(Rasheed，2020)，这些外部条件的改善为校长信息化领导力的发挥提供了有力支撑。有研究发现校长信息化领导力对技术伦理、人际沟通与设施供给、教学效能与技术整合度影响作用较大(段春雨，2020)；有研究证实变革型领导风格对校长信息化领导力存在显著正向影响(王淑华，2020)；进一步揭示了信息化教学领导力影响因素的复杂性。也有研究发现，感知易用性、感知有用性、使用态度以及行为意向均对信息化教学领导力产生直接正向作用效应(赵磊磊，2018)。校长信息化教学领导力对学校教育理念现代

化、教师队伍信息技术能力提升及学生学习方式创新至关重要。它要求校长具备技术视野、战略规划能力，推动教育内容与方法革新，并构建支持性环境，鼓励教师利用数字工具创造个性化学习体验。校长需具备前瞻性思维、沟通协调能力，以促进教育生态系统内协同合作，推进教育数字化。因此，提升校长信息化教学领导力是实现教育现代化、应对未来挑战的关键。

3.研究设计

3.1. 研究工具

本研究参考 Venkatesh 的技术接受与整合模型(UTAUT)(UNESCO，2011)和张汉玉的相关研究(张汉玉，2015)，结合教育数字化时代对校长数字素养的新要求，设计了“校长信息化教学领导力行为意向影响因素调查问卷”。UTAUT 模型最早由 Venkatesh 等人于 2003 年提出(Venkatesh，2003)，该模型能解释 70%的技术采纳和使用行为,优于以往的技术接受模型,目前被广泛应用于探究用户接受行为(王钰彪，2019)。为保障调查问卷的信效度，先进行了小范围预测问卷调查，对数据分析后，修正问卷，才正式投入调研。最终编制的正式调查问卷包括以下两个组成部分：第一部分为基本情况的调查，包括校长的性别、教龄、职称、学校性质、学校位置等；第二部分从绩效期望(Performance Expectancy，PE)、努力期望(Effort Expectancy，EE)、社会影响(Social Influence，SI)、促成条件(Facilitating Conditions，FC)、个体创新(Individual Innovation，II)，行为意向(Behavior Intention，BI)六个维度测量校长信息化教学领导力影响因素。校长信息化教学领导力影响因素问卷共有 28 个题项，运用李克特 5 分量表(LikertScale，“完全符合”-“完全不符合”五种回答选项，计分为 5-1)，各个研究变量的描述如表 2 所示。收集数据后，对问卷进行信效度检验，问卷总体效度值 0.904，PE、EE、SI、FC、II、BI 各维度信度值分别为 0.891、0.940、0.950、0.888、0.941、0.973，总体信效度良好。

3.2. 研究假设

校长信息化教学领导力的生成源于技术与教学领导力的交互与融合，其生成过程需要经过“技术感知”“技术决策”“技术运用”三个不同阶段(赵磊磊，2018)。校长信息化教学领导行为意向指校长利用新兴技术促进学校教育教学数字化变革的态度，反映校长对学校数字化转型积极或消极的情绪。有研究表明，UTAUT 模型中促成条件(FC)、努力期望(EE)、绩效期望(PE)和社会影响(SI)是影响教师接受新事物的关键因素(Bassam，2013)。UTAUT 模型能够很好的解释个体接受某个事物的行为意向。因此，本研究保留 UTAUT 模型中的四个变量，并通过培训方案的具体应用，新增加个体创新变量。

Fianu Eli 等人发现绩效期望对教师接纳新事物的程度存在正向显著关系(Fianu，2018)。本研究中，绩效期望类似于技术接受模型(TAM)中的感知有用性。这在一定程度上表明，如果校长能够感知到使用数字技术能提高工作绩效，并在学校教育教学数字化转型过程中起一定作用时，他们就会倾向于在信息化教学领导过程中使用数字技术。Teo 研究发现,感知有用性在职前教师对于技术使用的态度和意识方面起到决定作用(Teo，2008)。诸多实证研究证实努力期望与教师使用某项技术的意愿具有显著正相关，本研究中努力期望类似于技术接受模型(TAM)中的感知易用性。因此，在“技术感知”阶段，绩效期望和努力期望对校长教学领导行为有显著影响。基于以上分析，本研究做出如下假设：

H1(PE→BI):绩效期望对校长信息化教学领导行为意向有正向显著影响；

H2(EE→BI):努力期望对校长信息化教学领导行为意向有正向显著影响。

有研究表明社会影响对个体的行为意向有显著正向的影响，这种相关性在 TRA、TAM2 和 TPB/DT 理论中均得到了验证(SI 对应于主观准则)(Wong，2012)。教育领域中很多实证研究都证实了 SI 与教师使用某项技术的意愿具有显著正相关(Fianu，2018)。社会支持正向作用于教师数字化教学的感知有用性，这一结果能够从 Goldsmith 的研究中得到一定程度的解释(Goldsmith，2000)。有研究者认为教师在数字化教学摸索中遇到的“问题”或“意外”会阻碍数字化教学实效的发挥，与需求相匹配的社会支持，则会纾解教师在数字化教学中面临的困境，有助于其获得良好的教学效果，从而提升对数字化教学有用性的感知(亓建芸，2024)。基于以上分析，本研究做出如下假设：

H3(SI→BI):社会影响对校长信息化教学领导行为意向有正向显著影响；

H8(SI→PE):社会影响对校长信息化教学领导绩效期望有正向显著影响。

UTAUT 模型中，促成条件对个体使用技术的意向有显著积极影响，这种相关性在 TPB/DTPB、IDT 和 C-TAM-TPB 理论中得到了验证(Padilla，2013)。此外，有研究发现促成条件与教师接受新技术的意向程度存在显著正相关(Nandwani，2016)。校长在学校的信息化教学领导行为越明显，往往反映出更为有利的外部促成因素环境。另外有学者对影响学习者接受在线课程管理系统的因素展开研究，得出技术支持与努力期望间存在正相关的结论(Martins，2004)。Venkatesh 等对 TAM2 模型进行研究时，认为主观准则对应于 UTAUT 模型中的促成条件，其对个体所感知的技术有用性有显著正向影响(Lu，2005)。基于以上分析，本研究做出如下假设：

H4(FC→BI):促成条件对校长信息化教学领导行为意向有正向显著影响；

H6(FC→EE):促成条件对校长信息化教学领导努力期望有正向显著影响；

H7(FC→PE):促成条件对校长信息化教学领导绩效期望有正向显著影响。

June 等学者证实了创新期望与努力期望之间的关联性，发现用户若具备较高的创新性，并对尝试陌生事物具备较好的信心，那么他们的好奇心能在一定程度上能降低其应用创新技术的感知难度(Wejnert，2002)。也有研究发现个体在同时面对特定任务时，具有较强创造力的群体在应对不熟悉的任务时能产生新的解决方案，他们对待这些任务始终保持着积极的执行意向(Liu，2010)。此外，用户在移动学习中创新能力对感知易用性(努力期望)和接受意向有显著影响(Venkatesh，2012)；Taylor 等人在研究网络计算的接受度时，发现个体创新正向影响其使用意愿(Taylor，2007)。因此，本研究做出如下假设：

H5(II→BI):个体创新对校长信息化教学领导行为意向有正向显著影响；

H9(II→EE):个体创新对校长信息化教学领导努力期望有正向显著影响。

UTAUT 模型中有性别、年龄、经验、自愿性四个调节变量，分别影响不同的模型路径(Chen,2021)。关于本研究调节变量的选取，沿用了 UTAUT 模型中的性别变量，替换学历、担任校长时长和智能技术精通度变量。关于学历变量，本研究考虑到不同校长所接受到的教育程度不同会在学历上有所差异，因此，研究增加学历变量；关于校长担任时长变量，研究考虑到不同校长担任校长的时间不同，所具有的学校教学领导经验不同，为测试不同管理学校时间的校长对其信息化教学领导行为是否有影响，所以增加该变量。有研究将 UTAUT 模型中的经验变量定义为个体是否使用过某信息技术，而部分研究将其定义为使用某信息技术的经验程度，侧重点在使用信息技术的时长轻重程度(McGarr,2021)。考虑到教育数字化转型背景下，关于校长信息技术能力高低的需求已经转变为对新一代人工智能技术能力水平高低的需求，因此，研究增加智能技术精通度变量。由此可见不同背景变量的用户在模型中各潜

在变量上存在差异。因此本研究在构建影响校长信息化教学领导力影响因素模型时，也有必要引入一些调节变量，以确保研究的准确性。基于以上分析，本研究做出如下假设：

H10:不同学历的学校校长在各变量方面存在差异；

H11:不同担任校长时长的学校校长在各变量方面存在差异；

H12:不同性别的学校校长在各变量方面存在差异；

H13:不同智能技术精通度的学校校长在各变量方面存在差异。

基于以上分析，本研究构建了如下图 1 所示的假设理论模型，此模型用于探究教育数字化转型背景下校长信息化教学领导力影响因素及发展路径。

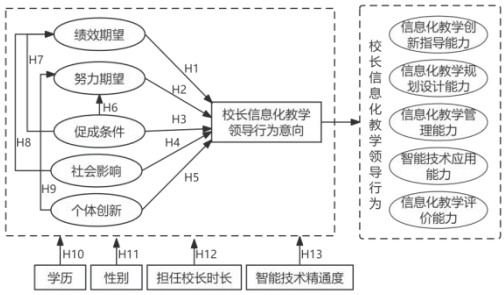


图 1 校长信息化教学领力影响因素研究模型

3.3. 模型拟合度分析

在 AMOS 中对假设理论模型进行建构后，对模型的相关拟合指标数据进行分析，发现初始模型的拟合结果有部分未达到拟合指标，因此，要对模型进行修正。经过 7 次修正，模型各个拟合指标基本都达到标准值，绝对适配度指数 RMSEA=0.094,GFI=0.919,AGFI=0.849,CMIN/DF=1.737;增值适配度指标 NFI、RFI、IFI、TLI、CFI 分别为 0.932，0.905，0.921，0.906，0.919；简约适配度指数 PGFI=0.57，PNFI=0.719，只有 RMR 一项指标略大于标准值。根据 Baumgartner 和 Homburg 的研究表明，一个研究模型中如果模型拟合指标大多数都达到标准，那么这个模型是可以接受的(Chen,2021)。因此，修正后的模型拟合度良好，可以用于研究。

3.4. 模型路径分析

对校长信息化教学领导力影响因素模型进行路径分析，发现原假设 H1、H4、H5 中变量标准化路径系数为正，且“P=0.00<0.05”，表明假设成立，个体创新、促成条件、绩效期望变量与校长信息化教学领导行为意向关系存在并显著；原假设 H6 中变量标准化路径系数为正“P=0.006<0.05”，假设成立，表明促成条件变量与校长信息化教学领导行为意向存在显著关系；原假设 H8 中变量标准化路径系数也为正“P=0.00<0.05”，假设成立，表明社会影响变量与校长信息化教学领导行为意向存在显著关系。假设 H2、H3、H7、H9 中的变量 P 值大于 0.05，假设不成立，表明上述假设中的变量与校长信息化教学领导行为意向关系不存在且不显著，需要删除这些路径。

4.研究结论与建议

4.1. 研究结论与讨论

1.校长信息化教学领导力影响因素分析

个体创新变量在前人有关校长信息化教学领导力的研究中并未被证实对校长的行为意向有影响。本研究加入此变量对其进行验证，研究结果表明，校长个体创新性能力是影响校长

信息化教学领导力的主要因素。校长作为学校数字化转型的引导者，其个体创新能力决定了他们能否敏锐捕捉教育科技前沿趋势，迅速理解和接纳新的教育理念，以及灵活调整学校发展规划以适应这一深刻变革。因此，校长的个体创新能力对于引领学校适应变革、推动教学模式创新、构建数字化教学环境、提升教师数字素养、实施数据驱动决策，以及培育数字教育文化等方面起着决定性作用。绩效期望对校长信息化教学领导力的影响程度仅次于个体创新性，可见职位晋升、薪资待遇、办学成效等因素影响校长绩效期望。校长绩效期望越大，其信息化教学领导行为意向越明显，即有“诱惑力”的绩效激励与奖励机制，会促使校长积极发展信息化教学领导力。

促成条件是影响校长信息化教学领导力的重要因素。有研究指出学校基础设备、设施的投资，教育相关培训、专业发展计划、政策及政府资金投入等在校长领导中发挥关键作用(李世瑾,2022)。从教育培训角度看，80%的校长认为，管理层间的有效经验交流能够降低其在学校信息化教学领导时应用智能技术的难度；82.5%的校长认为持续精准的教育培训会给自己在学校信息化教学领导中提供有力支持。从基础设施建设看，数字资源平台体系等的不完备，会限制传统校园向数字化校园转型的进度。从校长培养可持续服务体系方面看，缺乏平台互联、数据互通的国家数字教育资源公共服务体系，不完善、不充分的社会保障制度等都是制约校长信息化教学领导力发展的因素。技术、政策、经济、精神上更多的外力支持会促进校长信息化教学领导力的长效发展。社会影响对绩效期望有正向显著影响，社会影响间接影响校长信息化教学领导力。关注学校相关利益群体的协同增效(王春鑫,2024)，发挥社会群体在信息化教学领导情境中的情绪价值功效，真正做到社会群体助力校长信息化教学领导力的发展。

促成条件对努力期望有正向显著影响，并且深刻地作用于校长在信息化教学领导力这一维度上的表现与发展。促成条件的完善与优化，为校长信息化教学领导力发展提供全面且坚实的支撑框架，进而显著增强校长实现既定目标的内在动力与预期成效的信心水平，从而促使努力期望攀升至更为卓越的高度。探究促成条件广泛而深入的维度，如技术资源配置、教育培训体系、政策导向，以及积极向上、鼓励创新的组织文化氛围等，可以共同构筑便于校长领导力发挥的生态系统。

2.人口学变量对校长信息化教学领导力的影响分析

为探究不同类型的校长在 PE、EE、SI、II、FC、BI 维度的影响下所呈现的群体特征差异，本研究采用差异检验的方法，从性别、学历、担任校长时长、智能技术精通度四个方面检验校长在信息化教学领导过程中采纳技术行为意向的影响。研究结果表明，不同性别、学历、担任校长时长、智能技术精通度的学校校长在一些变量维度存在差异。为分析男性校长和女性校长在各个维度上的差异，进行方差齐性检验。结果如表 3 所示，在 6 个维度上，男性校长和女性校长只在 PE 维度上存在统计学上的显著差异，$t_{PE(260)}$=-2.14，p=0.034(女性>男性)。针对不同学历段的校长在各个维度上的差异，将被调查的校长按学历高低分为两个组，“硕士及以上”和“本科及以下”，结果如表 3 所示，两组学历的校长在 PE、EE、SI、FC、II 维度存在统计学上的显著差异，$t_{PE(260)}$=6.35，$t_{EE(260)}$=-2.18，$t_{SI(260)}$=-2.08，$t_{FC(260)}$=2.65，$t_{II(260)}$=2.48，p=0.000，t=0.038，p=0.041，p=0.010，p=0.015(硕士及以上>本科及以下)。为分析不同担任校长时长的校长在各个维度上的差异，将被调查的校长担任校长时长分为“10 年及以上”和“10 年以下”两组，进行 T 检验分析，结果表明两组不同担任校长时长的校长在 PE、SI、II 维度存在显著差异，$t_{PE(260)}$=-6.75，$t_{SI(260)}$=-2.80，$t_{II(260)}$=-4.99，p=0.000，p=0.006，p=0.000(10 年以下>10 年以上)。针对不同智能技术精通度的校长在各个维度上的

差异，将被调查的校长按智能技术的精通程度进行分组，“非常精通/精通”和“一般/比较/非常不精通”，结果表明，两组智能技术精通度的校长在 PE、SI、II、BI 维度存在显著差异，$t_{PE(260)}$=4.12，$t_{SI(260)}$=3.47，$t_{II(260)}$=4.76，$t_{BI(260)}$=4.19，P=0.000，P=0.001，P=0.000，P=0.000，(非常精通/精通>一般/比较/非常不精通)。

4.2. 建议与启示

1.政府主导完善校长培养资源与制度的建设

校长作为执行国家教育数字化转型战略的实践者，国家要提供完善的培养资源与制度支持其数字化治理能力的提升。在服务国家重要战略需求的情境下，统筹不同政策目标，促进拔尖创新人才培养的政策协同是关键点(王运武,2022)。国家教育主管部门应做到以下几点：其一，建立互联互通的全国数字资源共建共享平台。政府应充分利用国家教育公共资源，建立全国各级各类学校及教育组织机构之间的信息网络。其二，探索校长绩效激励与评价机制。近年来国家教育管理机构积极探索智能技术在教育领域的评价机制。各地教育督导部门积极探索科学的校长工作评估与培训质量评估机制，实时、有效反馈的评估数据，能够确保校长教学领导能力长效发展。其三，全面深化精准培训改革。应积极探索多元培训模式，技术支持下的校长培训给予教育培训新的发展平台与机遇，以解决校长实际教学领导需求为出发点，精准定位培训，促进校长信息化专业发展。

2.数字化领航校长数字能力建构与专业发展

学校教育数字化转型需要校长具备与数字化转型相适应的能力结构支撑。其一，信息化规划设计能力。校长能够依据国家相关标准、目标与要求，结合区域和学校的发展实际情况，制定出可有效实施的基于区域和学校视角的信息化发展规划。其二，教学指导创新能力。校长能够找到国家教育教学改革的要求与学校发展现状之间的差距，通过创新教研模式实现教学模式创新，指导学校教师运用智能技术进行学情分析、教学设计和学业评价等。其三，组织实施能力。校长运用智能技术支持下的区域和校本教育理念，开展学校数字化教学创新，制定教师专业发展推进方案，并能够对教学活动的设计和组织实施提出指导意见。其四，智能化管理与评价能力。校长应学习数字化教育测量与评价的理念与方法，学会设计并运用数字化工具开展基于教师数字化教学、教师专业成长、学校智能化管理等方面的科学评估与决策，掌握科学评价诊断与分析的方法，能够制定出基于区域和学校视角的数字化应用考核评价方案。

3.数字化时代下校长教育培训的优化与实践

优质的校长教育培训方案能够助力校长数字意识与能力的高质量发展。其一，建立区域校长数字画像，提供个性化教育培训。对校长在各类数字化学习平台观摩学习的数据收集分析，可以帮助调查校长的教育培训需求，推动合理设计培训方案，精准实施培训，确保培训目标落地。此外，注重培训内容选择自主性，校长是个性的，不同校长专业发展需要不同学习内容的支撑，个性化的教育培训应该要满足校长不同的专业发展需求。其二，培训方式多元化，培训内容数字化。为提高校长培训学习效果，要关注教育培训方式的灵活性；为提高培训实施成效，应该合理规划培训方案内容。新时代教育培训过程应以“问题导向、任务驱动、按需分类提升”的逻辑，采用多元化的培训方式，全面提升校长的教育管理信息化水平和协同创新能力。还可以通过返岗实践、优秀成果展示、典型引领等方式，提高校长教育培训实践能力。

参考文献

教育部办公厅关于实施新时代中小学名师名校长培养计划(2022—2025)的通知[EB/OL].(2022-8-12).http://www.moe.gov.cn/srcsite/A10/s7011/202208/t20220819_653904.html.
董艳,黄月,孙月亚,桑国元(2015).校长信息化教学领导力的内涵与结构[J].现代远程教育研究,(05):55-62.
杨鑫,解月光,赵可云,修颖(2018).校长信息化教学领导力模型构建及发展途径[J].现代远程教育研究,(04):34-40+48.
赵磊磊,张蓉菲(2019).教师信息化教学领导力:内涵、影响因素与提升路径[J].重庆高教研究,7(03):86-97.
段春雨(2020).校长信息化领导力对教师信息技术应用行为的影响研究——基于 42 篇实证研究文献的元分析[J].中国远程教育,(10):23-33+76-77.
王淑华,王以宁,张海,等(2020).中小学校长领导风格对校长信息化领导力影响的研究——以变革型领导理论为视角[J].湖南师范大学教育科学学报,19(02):105-112.
赵磊磊(2018).农村校长信息化教学领导力的影响因素及提升路径——基于技术接受视角的实证研究[J].湖南师范大学教育科学学报,17(05):25-32.
张汉玉,钱冬明,任友群(2015).推进电子书包教学应用:教师接受度的实证研究[J].电化教育研究,36(10):92-97.
王钰彪,万昆,任友群(2019).中小学教师机器人教育接受度影响因素研究[J].电化教育研究,40(06):105-111.
赵磊磊(2018).农村校长信息化教学领导力的影响因素及提升路径—基于技术接受视角的实证研究[J].湖南师范大学教育科学学报,17(05):25-32.
亓建芸,赵可云(2014).教师数字化教学角色适应的影响因素及其作用机制研究[J].中国电化教育,(01):111-118.
李世瑾,顾小清(2022).革故鼎新的校长智能化领导力:内涵结构、发展现状与提升路径[J].远程教育杂志,40(05):72-82.
王春鑫,刘桂芝(2024).政策协同视角下的拔尖创新人才衔接培养:价值意蕴、协同困境与突破策略[J].国家教育行政学院学报,(07):65-74.
王运武,李雪婷,王藤藤等(2022).人工智能助推教师队伍建设:新诉求、潜能与应用场域[J].数字教育,8(06):13-20.
Hallinger P, Murphy J. (1985). Assessing the instructional management behavior of principals[J]. *The elementary school journal*, 86(2): 217-247.
Kannan S, Sharma S, Abdullah Z. (2013). Principal's Strategies for Leading ICT Integration: the Ma-laysian Perspective[J]. *Creative Education*, 3(08): 111.
Mestry,R. (2017). Empowering Principals to Leadand Manage Public Schools Effectively in the 21st Century[J].*South African Journal of Education*, 37(1):1-11.
Samarraie H, (2018). Saeed N. A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment[J]. *Computers & Education*, 124: 77-91.
Rasheed R A, Kamsin A, Abdullah N A. (2020). Challenges in the online component of blended learning: A systematic review[J]. *Computers & Education*, 144: 103701.
Venkatesh V, Morris M G, Davis G B, et al. (2003). User acceptance of information technology: Toward a unified view[J]. *MIS quarterly*, 425-478.

Bassam Nassuora A. (2013). Students acceptance of mobile learning for higher education in Saudi Arabia[J]. *International Journal of Learning Management Systems*, 1(1): 1-9.

Fianu E, Blewett C, Ampong G O A, et al. (2018). Factors affecting MOOC usage by students in selected Ghanaian universities[J]. *Education Sciences*, 8(2): 70.

Teo T. (2008). Pre-service teachers' attitudes towards computer use: A Singapore survey[J]. *Australasian Journal of Educational Technology*, 24(4).

Wong K T, Russo S, McDowall J. (2012). Understanding early childhood student teachers' acceptance and use of interactive whiteboard[J]. *Campus-Wide Information Systems*, 30(1): 4-16.

Goldsmith D J,Mcdermott V M, Alexander S C.(2000). Helpful,Supportive and Sensitive:Measurin g the Evaluation of Enacted Social Support in Personal Relationships[J]. *Journal of Social&Personal Relationships*, 17(3):369-391.

Padilla-Meléndez A, del Aguila-Obra A R, Garrido-Moreno A. (2013). Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario[J]. *Computers & Education*, 63: 306-317.

Nandwani S, Khan S. (2016). Teachers' intention towards the usage of technology: an investigation using UTAUT model[J]. *Journal of Education & Social Sciences*, 4(2): 95-111.

Martins L L, Kellermanns F W. (2004). A model of business school students' acceptance of a web-based course management system[J]. *Academy of Management Learning & Education*, 3(1): 7-26.

Lu J, Yao J E, Yu C S. (2005). Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology[J]. *The journal of strategic Information Systems*, 14(3): 245-268.

Wejnert B. (2002). Integrating models of diffusion of innovations: A conceptual framework[J]. *Annual review of sociology*, 28(1): 297-326.

Liu Y, Li H, Carlsson C. (2010). Factors driving the adoption of m-learning: An empirical study[J]. *Computers & Education*, 55(3): 1211-1219.

Taylor N J. (2007). Public grid computing participation: An exploratory study of determinants[J]. *Information & Management*, 44(1): 12-21.

Venkatesh V, Thong J Y L, Xu X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology[J]. *MIS quarterly*, 157-178.

Chen S C, Li S H, Liu S C, et al. (2021). Assessing determinants of continuance intention towards personal cloud services: Extending utaut2 with technology readiness[J]. *Symmetry*,13(3): 467.

McGarr,O.,Mifsud,L.,Colomer,R.J. (2021). Digital Competencein Teacher Education:Comparing N ational Policies in Norway,Ireland and Spain[J].*Learning , Media and Technology*,46(4):483-4 97.

智慧教室环境下促进语文课堂互动的教学策略研究

Research on Teaching Strategies to Promote Chinese Classroom Interaction in Smart Classroom Environments

康健[1*]，杨文静[2]，赵建华[3]
[1]西北师范大学 教育技术学院，南方科技大学 未来教育研究中心
[2]西北师范大学 教育技术学院
[3]南方科技大学 未来教育研究中心
[*] kangj2024@mail.sustech.edu.cn

【摘要】 智慧教室作为典型的智慧学习环境，对课堂互动具有支撑作用。但目前的研究多聚焦于分析结果的普适性，而较少考虑特定学科的独特需求，难以对学科教学提出针对性意见，本研究基于社会互动理论提出智慧教室环境下促进语文课堂互动的教学策略：包括构建交互型教学资源的策略、营造和谐氛围的策略、分工合作学习策略和及时评价与反馈策略。通过准实验研究法，比较某小学五年级两个班级使用这些策略指导语文学科，设计的教学方案与传统教学方案的教学效果，通过互动行为分析、滞后性分析，结果显示实验班课堂互动质量明显优于对照班，证明了这些教学策略有效促进了语文学科的课堂互动与教学效果。
【关键词】 智慧教室；课堂互动；滞后性分析；教学策略；语文学科

Abstract: *Smart classrooms support classroom interaction as a typical smart learning environment. However, current research mainly focuses on the universality of analysis results, paying less attention to the unique needs of specific disciplines, making it difficult to provide targeted advice for subject teaching. Based on social interaction theory, this study proposes teaching strategies to promote Chinese classroom interaction in smart classroom environments: These include strategies for constructing interactive teaching resources, creating a harmonious atmosphere, cooperative learning through division of labor , and intelligent evaluation and feed back. Through a quasi-experimental research method, this study compares the teaching effectiveness of a teaching plan designed using these strategies with a traditional teaching plan in two fifth-grade classes of a primary school in the context of Chinese language instruction. Through interaction behavior analysis and lag sequential analysis, the results show that the quality of classroom interaction in the experimental class is significantly better than that in the control class, demonstrating that these teaching strategies effectively promote classroom interaction and teaching effectiveness in Chinese language instruction.*
Keywords: Smart Classroom, Classroom Interaction, Latency Analysis, Pedagogical Strategies

1.前言

智慧教室作为目前的典型智慧学习环境，其强调要充分利用信息技术来优化教与学的环境，智慧教室环境能够支持具有情境创设、资源共享、多重互动、协作学习等多方面要求的功能，因此在智慧教室环境下进行课堂教学是实现新时代教学改革的重要途径（刘选等，2023）。众多学者围绕智慧教室环境下的课堂互动开展了一系列研究，主要从互动行为分析

框架、互动策略等方面开展了一系列研究。语文作为一门语言学科，如何优化课堂教学互动是促进语文教育高质量发展中尤为值得探讨的问题（朱涵等，2023）。众多研究者针对与语文学科的课堂互动开展了系列研究，构建了相应的互动模型和策略（薛沾陶，2019）,但对小学语文学科课堂互动策略的实际有效性验证较少(黄真金，2019),因此，本研究基于已有研究成果，试图回答以下研究问题：

问题 1：促进小学语文的课堂互动策略有哪些？

问题 2：促进小学语文的课堂互动策略有效性如何？

2.教学策略的构建

2.1.理论基础：社会互动理论

社会互动理论最早由美国学者班杜拉（A .Bandura）提出，认为人的心理活动是个人、行为、环境三个因素之间相互作用的过程（阿尔伯特·班杜拉，2015）。个人和行为影响环境因素，能创造一定的环境条件，而环境因素为个人的互动提供了广阔的物理空间。社会互动是在社会的相互作用下，人与人、人与群体、群体与群体等在心理和行为上相互影响的过程（王卫军等，2019）。费厄斯坦指出，儿童的学习与认知的发展更容易发生在与他人的交往与互动过程中,即对儿童成长有重要意义的人,例如父母、老师、同伴等都是他们重要的互动对象(Williams et al.,2000)。个体在与他人的人际沟通中完成所谓的社会互动，沟通交流是人们认识与适应外部世界的重要方式。

对本研究的启示：第一，社会互动理论认为环境因素为个人的互动提供物理条件，所以要求准备交互式教学资源，以帮助学习者开展与环境的互动，第二，教师与学生应保持平等友好的沟通氛围。第三，需要在教学过程中，教师应重视与学生的互动，需要给予适当的反馈与评价，从而调动学生的积极性，发挥其潜能。第四，课堂教学中教师要注意引导学生与其他学生的互动或者学生与资源的互动，从而及时消除学生学习障碍，促进学生的学习水平发展。

2.2.教学策略

根据社会互动理论以及智慧教室环境下的支持功能，结合实际教学流程，本研究提出以下四个教学策略：（1）构建交互型教学资源的策略。（2）营造和谐氛围的策略（3）异质合作协同策略（4）及时评价与反馈策略。

2.2.1. 构建交互型教学资源的策略

在智慧教室环境中，教学资源准备需围绕多途径获取与交互式设计展开：

（1）资源获取多元化

教师除使用传统教材外，可通过智慧教室内置资源库（如希沃白板课件）、教育云平台及学科资源网站下载素材，亦可购买正版数字资源。网络资源检索时需注意版权限制。

（2）二次加工增交互

对获取的资源进行二次开发，嵌入互动式元素。例如，在语文课件中加入成语接龙游戏、字词消消乐等轻交互模块；根据学生认知水平调整素材难度，确保内容可操作性强，促进知识建构。

2.2.2. 营造和谐氛围的策略

课堂氛围的和谐性直接影响互动质量，需从环境营造与情感联结双管齐下：

（1）平等环境构建

打破传统的“教师权威”模式，建立“学习共同体”。鼓励学生质疑、表达多元观点，通过智慧教室的分组功能支持自由讨论，如利用平板终端开展神话动画赏析后的即兴辩论。

（2）情感驱动互动

教师采用生动语言结合技术工具激发兴趣；通过随机点名、电子勋章等互动功能传递平等感。建立容错机制，对学生错误采取“部分肯定+引导修正”策略，如：“你的想法很有创意，如果考虑这个角度会更完整。”

2.2.3. 异质合作协同策略

利用智慧教室技术支持差异化分组与协作深化：

（1）智能分组平衡

通过终端数据分析学生学习风格，实现异质分组（如将视觉型、听觉型学习者混搭），突破物理座位限制。

（2）角色轮换机制

每组设定轮值组长，负责任务分配（资料搜集、观点整合）及代表发言。利用思维导图工具辅助讨论，确保所有成员观点被记录。

（3）教师引导策略

教师作为观察者，仅对偏离主题或进程滞后的小组介入指导，利用计时器控制讨论节奏，最终组织全班观点共享。

2.2.4. 及时评价与反馈策略

（1）正向引导促进互动闭环

构建多维反馈体系，强化互动闭环：采用“具体肯定+建设性建议”模式，如：“你的成语解释很准确，若能结合课文场景会更生动。”对错误回答不直接否定，而是引导自我修正；对学生不充分的回答需要进一步引导，帮助学生理解，总结。

（1）技术驱动评价

利用智慧教室积分系统实现即时反馈（抢答加分、周榜奖励），引入同伴互评机制（如通过平板对小组汇报打分）。将评价融入学习过程，如通过数据仪表盘展示互动参与度，帮助学生自我监控。

3.教学策略有效性验证

本研究的实验对象是银川市某小学五年级的学生，实验班级为五年级四班，对照班级为五年级一班，两个班学生人数均为五十人。其平均年龄均在 11 岁左右。这个阶段的学生，已经有能力掌握、使用各类信息技术工具，如鸿合平板，适合参与到智慧教室环境下的学习活动中。同时，五年级学生具有可塑性强、喜欢主动尝试、思维开放等特点，且具有较强的自我意识，并且五年级学生已具备了一定的信息化相关知识，在简单的培训后能够较为熟练地操作各类信息化软件，具备较好的信息素养和信息化能力。本研究中的实验班级和对照班级均为彭老师所带，彭老师是一名具有丰富教学经验的小学语文教师。同一个授课教师进行实验班级和对照班级的授课，这样可以在一定程度上控制了部分无关变量的干扰。这些学生和老师在实验之前没有开展过课堂互动教学策略引导的相关教学。

3.1.实验流程

本研究中的两个班级开展教学环境都是智慧教室。按照学校既定教学进度，选取五年级上册第一单元至第四单元开展实验。在实验班实施基于智慧教室环境下促进课堂互动的教学策略设计的小学语文教学方案，对照班仍按常规的教学方式进行课时教学。实验周期为半个学期，即十周。实验流程如图 1 所示为减少干扰因素，对比教学效果，本研究收集的两个类

的课堂视频是同一课程的内容《猎人海力布》（第二课时），为五年级语文学科第三单元第9课的内容，满足实验要求。

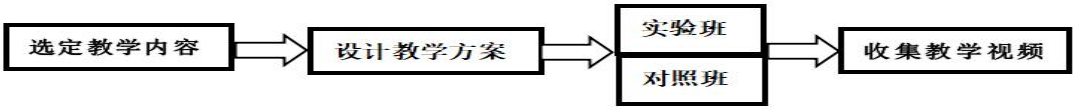


图 1 实验流程图

3.2.课堂教学视频分析

3.2.1.视频分析工具

关于视频分析技术，在20世纪60年代弗兰德斯首次提出了弗兰德斯互动分析系统FIAS(Flanders Interaction Analysis System)(Jiang et al.,2019)。FIAS是课堂教学中重要的信息反馈工具，是用于师生在课堂上言语互动过程的观察分析工具，FIAS由教师、学生的言语行为及沉寂情况三部分,含10种互动行为编码组成。2004年，顾小清和王炜提出了基于信息技术的互动分析编码系统ITIAS (Information Technology-based Interaction Analysis System ）（顾小清&王炜，2004）。ITIAS对课堂互动进行了全面的分类。它不仅通过编码来分析师生的部分语言活动，还为使用信息技术的班级提供互动研究支持。2012年，方海光等在顾晓青、王伟提出的ITIAS基础上，对现有的数字课堂教学环境编码体系进行了细化和改进，重新设计了支持数字课堂评价的编码体系，从而形成了改进型弗兰德斯互动分析系统iFIAS(improved Flanders Interaction Analysis System)（方海光等,2012）。但是，目前的研究多考虑对所有课堂的普适性，具备学科特色的课堂互动分析框架较为稀少，且对言语层次的划分较为浅显，未对互动言语结构和言语内容进行深度剖析。刘革平教授基于已有研究成果制定了专属于语文学科的课堂互动分析框架比较符合语文学科的要求（刘革平等，2024）。因此本研究采用该框架。框架如表1所示：

表 1 语文智慧课堂教学言语互动分析编码框

分类	编码	表述		行为描述
	T1	教师引导课堂的情感与态度		教师能够以一种不具威胁性的方式，表明及澄清引导学生的态度或情感。如“看来大家都非常有热情”“看来大家都很困惑”
	T2	教师鼓励表扬	T2.1 言语鼓励	称赞或鼓励学生的动作或行为。如“真棒”“回答得很好”
			T2.2 基于技术的鼓励	通过技术对学生进行称赞或鼓励。如在学生名字后添加大拇指或小红花等
	T3	采纳意见		教师承认学生的说法；修饰重述学生的说法后应用它去解决问题；将其与其他学生的说法相比较；对学生意见进行总结

	T4	教师提问（包括 T4 基于技术的提问和 Tt4 基于言语的提问）	T/Tt4.1 提问封闭性问题	T/Tt4.1.1 字词积累	教师根据内容提出具有确定性答案的问题，考查学生基于事实性知识的识记状态	认识和记忆常用的汉字、诗词
				T/Tt4.1.2 阅读鉴赏		运用多种阅读方法，初步鉴赏文学作品
				T/Tt4.1.3 表达交流		运用口头语言或书面语言表达自己的体验和想法
				T/Tt4.1.4 梳理探究		梳理学习内容，用口头或图文等方式，提出自己的看法
			T/Tt4.2 提问开放性问题	T/Tt4.2.1 字词积累	教师根据内容提出具有多种可能性答案的问题，侧重考查学生基于分析、评价的高级思维能力	形成良好语感、初步领悟语言文字运用规律
				T/Tt4.2.2 阅读鉴赏		感悟作品的思想内涵和艺术价值，结合自身经验理解作品
				T/Tt4.2.3 表达交流		提高语言表现力和创造力，增强形象思维能力
				T/Tt4.2.4 梳理探究		通过比较、分析、概括等思维方法，辩证地思考问题
	T5	朗读	教师、视频或音频领读，学生跟随进行字词或课文的朗读			
	T6	讲授	T6.1 传统的讲授		教师借助板书、纸质图片等传统媒介讲授	
			T6.2 基于技术的普通讲授		教师借助多媒体技术如 PPT 对教学内容进行单一呈现和讲授	
			T6.3 基于技术的多样讲授		教师借助媒体技术以微视频、动画、虚拟现实等形式代替教师讲授	
	T7	组织管理和指示	T7.1 传统的组织管理和指示		教师组织管理课堂所提出的口头命令和要求，以及对学生的指示	
			T7.2 基于技术进行组织管理和指示		教师操作技术组织管理课堂，对学生进行指示	
	T8	学情分析	教师通过课堂观察或媒体技术分析学生的学习特点、学习进度、学习情况等相关信息			
	T9	批评	以强求改变学生行为为企图的教师行为；要求学生从不可接受的形态转变为可接受的形态；责骂学生；不作说明的强制性行为；教师的极端自我参照行为（如，我当年怎么怎么样，你们现在是身在福中不知福）			
			T10.1 传统的分享与评价		教师直接展示分享学生的作品、作业，进行言语评价反馈	

	T10	作品分享评价	T10.2 基于技术的分享与评价	教师操作媒体技术呈现学生作品、作业，并基于多种信息采集进行点评反馈
	T11	教师指导	教师对学生个人或小组进行单独的答疑解惑和指导	
学生行为	S1	学生被动应答	学生为了应对教师的提问被动口头回答问题。即学生的回答非自发产生，是在教师引导或指示下完成，如教师指定未举手的同学回答问题，或是要求学生朗读课文、完成阅读任务要求，学生自由表达自己的想法是受到限制的	
	S2	学生主动应答	老师提出问题后，学生自己主动举手回答，或者脱口而出将答案说出来；学生的回答超出了问题的答案，表达自己的想法，引发新的话题	
	S3	学生主动提问	学生主动提出问题，与同学或者教师分享自己的问题，自由地表达自己的见解	
	S4	同伴交流合作	S4.1 传统的交流合作	学生之间言语讨论交流或借助纸笔等传统媒介辅助交流、合作实践
			S4.2 基于技术的交流合作	学生借助媒体技术进行讨论交流、合作实践
	S5	评价反馈	S5.1 传统的评价反馈	学生对其他同学的回答或观点进行口头评价或做出其他的反应
			S5.2 基于技术的评价反馈	学生借助媒体技术进行自我评价、同伴评价等
	S6	成果展示	S6.1 传统的展示	学生通过传统媒介如黑板、纸质图片来展示作业作品、演示步骤等
			S6.2 基于技术的展示	学生借助媒体技术分享展示作业、作品、演示步骤等，如学生上讲台操作白板展示自己的作业
沉寂行为	CH	无助于教学的混乱	因为遗忘、不熟练或不满意等引起的暂时停顿、短时间的安静或混乱，以至于观察者无法了解师生之间的沟通	
	QU	有助于教学的沉寂	QU.1 思考问题	学生根据教师的指示进行练习或静默学习
			QU.2 做练习	
			QU.3 默读课文	
			QU.4 教学必要的沉寂	

3.2.2.数据编码

本研究在实验开展后期，随机对实验班和对照班讲授相同内容的课堂教学进行录像分析，两个班级教学录像经过剪辑处理，实际有效教学录像时间均为四十分钟，利用课堂互动行为编码系统与滞后性分析进行数据分析。利用课堂互动行为编码系统进行视频分析时需要将视频进行编码，具体参照表 1，为了确保获得可靠的视频编码，本研究采用两个编码人员同时对同一视频进行编码。编码过程包括以下三个步骤：第一步是仔细识别编码类型，明确编码所代表的具体行为的含义。确保课堂互动行为分类编码意义的一致性。第二步，预编码，两个编码人员各自对选定的样本视频（十分钟）进行编码。编码完成后，评估两个编码之间的

差异，确保编码正确。第三步是正式编码。两个编码人员分别对同一个录制的教学视频进行编码。编码过程中不允许任何交流和协商，最终得到两个行为编码，再检查编码结果的准确性，并可使用 kappa 系数检验一致性。如果该值大于 0.75，则认为两种编码之间的一致性良好（杨现民等，2016）。

3.3.结果分析

按照以上步骤对实验班和对照班教学录像进行互动行为编，将师生行为类型划分为提问媒介、提问类型与学生回答、技术操作 4 种情况，并计算每种情况中不同提问与回答行为的编码占进行比例统计如图 2。

3.3.1.主要行为比例分析

将师生行为类型划分为提问媒介、提问类型与学生回答、技术操作 4 种情况，并计算每种情况中不同提问与回答行为的编码占比，基于技术的提问是（T4.1 和 T4.2）编码之和；基于言语的提问是（Tt4.1 和 Tt4.2）编码之和，技术操作分别是学生技术操作比学生语言之和，教师技术操作比教师语言之和。两节课比例统计情况如图 2 所示。由图可知，实验班中基于技术使用的提问占比与对照班相差无几，尽管技术的应用使得教学活动更加多样化，但课堂中开放性问题的多少与技术使用的频率之间并无显著关联。同时，实验班中学生主动回答问题的比例超过了 75%，表明学生在课堂中的积极性明显高于普通课例（24.8%）。此外，通过学生技术操作差异对比看出，实验班学生技术操纵比例远高于对照班,表明学生与学习资源的互动更为频繁。

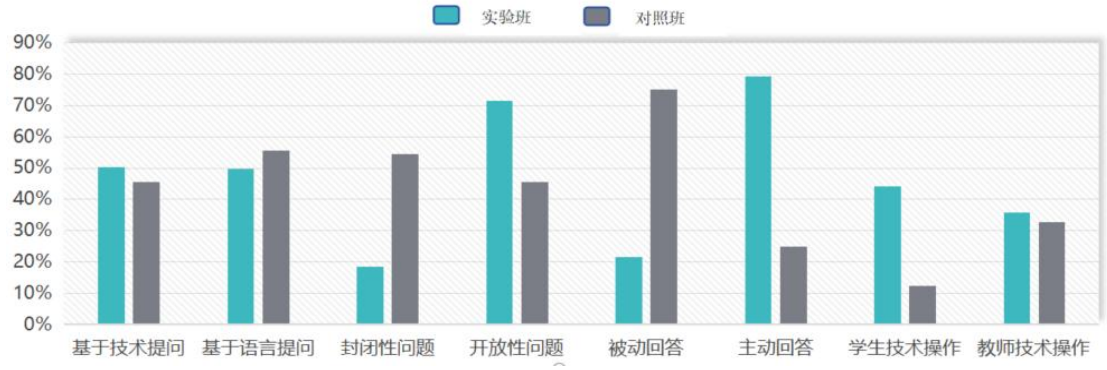


图 2 两个班级主要行为占比

3.3.2.互动行为滞后序列对比分析

在上述分析中，借助课堂互动行为编码系统工具，分析两个班在智慧教室环境中师生的行为频次和行为统计比例，进行了简单的数量上的比较，但对课堂的互动结构缺乏深度的分析，例如：未能对课堂中某种行为发生后伴随另一种行为发生的可能性进一步展开分析，为解决这个问题，本研究采用滞后性分析法(LagSequentialAnalysis，LSA)，滞后性分析法是用于检验人们发生一种行为之后另外一种行为出现的概率及其是否存在统计意义上的显著性当残差参数 z 大于 1.96 时，则表示该行为序列存在显著性(王觅等，2020），GSEQ(General Sequential Querier)就是一款可以进行滞后序列分析的软件，该软件简化了行为分析流程，提高了数据处理效率，也促进了 LSA 在各领域中的应用推广利用 GSEQ 软件进一步分析。得到图 3，根据图 3 可以看出：

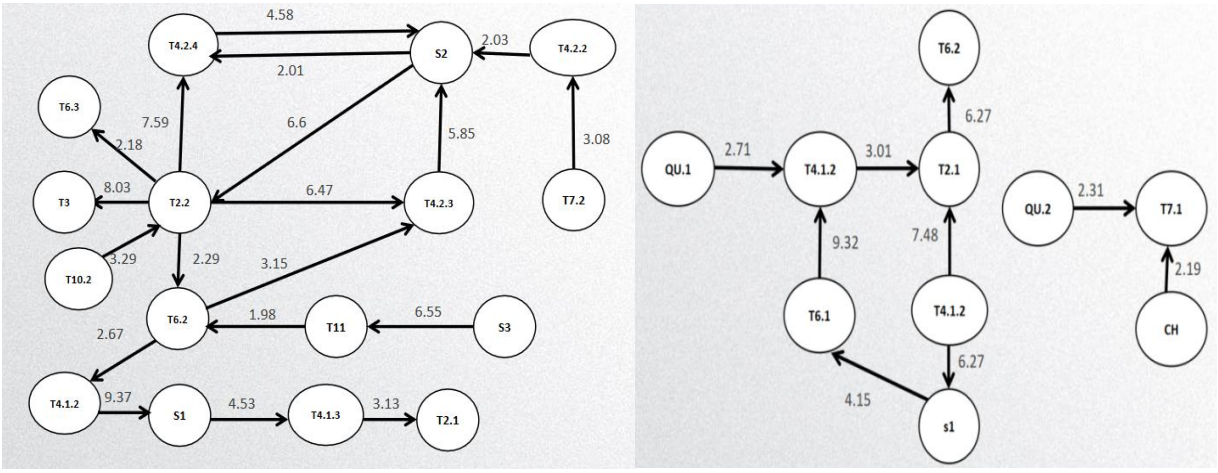


图 4 两个班级行为序列图（左为实验班，右为对照班）

由图可以看出，在实验班中，教师讲授后，在后续行为中更倾向于采用开放性问题（T4.2)进行提问。这表明，实验班中的教师更善于通过提问激发学生的内在学习动力，并有效利用讲授行为引导学生进行深入思考，侧面印证了实验班课堂氛围轻松，学生踊跃参与。而对照班开放性问题提出与学生回答之间不存在显著序列，说明以提问封闭性问题为主（T4.1)。对照班教师在挖掘和培养学生综合能力方面不够深入，提问过于注重学生对课文本身内容的记忆，而较少引导学生进行知识的深度拓展与应用，对照班中封闭性和基础性问题过多，可能也是导致课堂氛围不佳的根本原因之一。

实验班互动行为序列图，更为复杂。教师的反馈行为展现出更高的多样性和复杂性（如图 3 所示）。在实验班中，教师在鼓励和采纳学生意见后，提出的问题往往具有更强的开放性和综合性，这种提问方式有效地激发了学生的思维能力。而在对照班课例中，教师的交流（T4.1.3/T4.2.3)和探究问题(T4.1.4/T4.2.4)则未能形成显著的行为序列，反映出其课堂互动结构的单一。实验班教师有意在给予学生鼓励或采纳意见等情感支持后，进一步对学生进行深层次的追问，如 T4.2.2(基于技术的阅读鉴赏类提问)-S2(学生主动应答）-T4.2.4（基于技术的梳理探究类提问）；T4.2.3(表达交流）-S2（学生主动应答）-T4.2.4（基于技术的梳理探究提问）这种对话有助于学生的认知思维进一步深化，巩固知识点。

4.总结

智慧教室环境对课堂互动有较高的支持作用。然而，没有相应教学策略指导，就无法有效促进课堂互动。本研究基于社会互动理论，提出智慧教室环境下促进语文课堂互动的教学策略，设计教学方案，在两个班级之间开展准实验验证，收集课堂教学录像，采用相关语文课堂编码量表，对两个班级教学视频编码，进行了统计分析和滞后性序列分析，分析结果表明实验班课堂互动效果要优于对照班，从而证明了教学策略有效。为新课标的落地实施提供了借鉴，为一线教师的实际教学提供了经验。本研究仍存在分析课例较少，未分析学生认知层级变化等不足。今后将进一步关注行为互动与认知层级变化之间的印证，以便优化课堂互动策略，进一步提升课堂互动质量。

参考文献

刘选,刘革平,廖剑.（2023）.智慧教室结构模型构建研究——“功能—结构”关系视角[J].中国远程教育,43(02):45-53.

朱涵,林晨,宫慧娜.（2023）.培智学校低年段生活语文信息化课堂教学互动研究[J].中国特殊教育,(09):28-36.
薛沾陶.（2022）.小学语文互动型课堂的构建[J].语文建设,(24):68-70.
黄真金.(2019).语文高考与中学语文教育互动的问题及改进策略研究[J].教育科学研究,(04):91-95.
阿尔伯特·班杜拉.(2015). 社会学习理论[M]. 陈欣银，李伯黍，译. 北京：中国人民大学出版社.
王卫军,代亚萍,韩春玲,等.(2019).基于社会互动理论的协作式在线课程中讨论设计及促进策略探讨[J].电化教育研究,40(01):87-94.
顾小清,王炜.(2004).支持教师专业发展的课堂分析技术新探索[J].中国电化教育,(7):18-21.
方海光,高辰柱,陈佳.(2012),(改进型弗兰德斯互动分析系统及其应用[J].中国电化教育,10):109-113.
刘革平,孙帆,廖剑.（2024）.智慧课堂有效教学言语互动甄别与提升路径——以 16 节语文智慧课堂课例为例[J].现代远距离教育,(06):35-48.
杨现民,王怀波,李冀红.(2016).滞后序列分析法在学习行为分析中的应用[J].中国电化教育,(02):17-23+32.
王觅,文欣远,李宁宁,等.(2020).大单元教学视角下基于 LSA 的同步课堂师生交互行为研究[J].电化教育研究,41(08):74-81.
Jacucci, G., Bellucci, A., Ahmed, I. *et al.* (2024).Haptics in social interaction with agents and avatars in virtual reality: a systematic review. *Virtual Reality 28,* 170.
Jiang, Y., Wang, W., Li, C., Kang, M., & Shen, J. (2019). An analysis of teacher - student interaction in the smart classroom. Modern Distance Education, 36(3), 13–21.

协作知识建构对小学生认识论信念的促进研究

Scaffolding Primary Students' Epistemological Beliefs using Knowledge Building Pedagogy

张晓洁[1, 2*]，陈桂涓[1]，冯诗惠[1]，赵建华[2]
[1]香港大学 教育学院
[2]南方科技大学 人文社会科学学院
[*]zhang_xiaoxiao_jie@163.com

【摘要】知识建构需要学生发现复杂的真实问题，不断调整策略来解决问题，并在集体知识中取得进步。在知识建构环境中，学生的认识论信念反映了他们对知识创造本质的概念化。本研究重点探讨学生的认识论信念在协作知识建构过程中的发展。本研究采用基于波普尔世界 3 理论的编码表，分析了 28 名小学生在协作知识建构前后的认识论信念变化。研究结果表明，协作知识建构可以显著提升小学生的认识论信念。在协作探究过程中，原本持世界 2 信念的学生通过参与共同体知识的发展，转变为基于世界 3 的高认识论信念。 本研究从认识论的角度对学生的知识建构过程进行了深入的剖析，以期促进知识建构理论和教学实践的发展。
【关键词】知识建构；认识论信念；协作探究学习；世界 3 理论

Abstract: *Knowledge building requires students to identify complex real-world problems, continuously adjust their strategies to address them, and make collective progress within the community's knowledge advancement. In knowledge building environments, students' epistemological beliefs reflect their conceptualization of the nature of knowledge creation. This study focuses on the development of students' epistemological beliefs during the collaborative knowledge building process. Employing a coding schema grounded in Popper's World 3 theory, we analyzed the epistemological belief changes of 28 elementary school students before and after collective knowledge building activities. The results indicate that collective knowledge building significantly enhanced students' epistemological beliefs. During the process of collaborative inquiry, students who initially held World 2 beliefs transformed towards more advanced World 3–based epistemological beliefs through their participation in the development of communal knowledge. This study provides an in-depth analysis of students' knowledge building processes, aiming to contribute to the theoretical advancement of knowledge building and its instructional practices.*
Keywords: Knowledge Building; Epistemological Beliefs; Collective Inquiry Learning; World 3 Theory

1.前言

协作探究学习能够促进学生的知识进步和高级技能（例如批判性思维、协作、沟通和创造力）的发展（Chen & Chen，2024），使学生能够在由同学、教师和技术媒体（Crook，1991）共同创造的协作学习情境中，通过协作探究找到真实问题的解决方案（Dewey，1986）。作为协作探究学习的教学法，知识建构（Knowledge Building）强调学生的认知能动性，即学生自主学习并为思想发展做出贡献的能力（Scardamalia & Bereiter，2006）。知识建构需要学生发现复杂的真实问题，不断调整策略来解决问题，并在集体知识中取得进步。

认识论信念（Epistemological Beliefs）被认为是关于知识本质及其获取的内隐假设系统（Bruning et al.，1999）。在知识建构环境中，学生的认识论信念反映了他们对知识创造本

质的理解（Kuhn 等，2000）。以往研究表明，协作知识建构环境可以促进学生认识论信念的发展（Hong 等，2016；Liu 等，2016；Zhou 等，2025）。然而，以往研究关注于大学生和中学生，很少有研究关注协作知识建构环境对小学生认识论信念的促进作用。因此，本研究重点探讨小学生的认识论信念在协作知识建构过程中的发展。主要研究问题如下：

研究问题 1：协作知识建构过程能否促进小学生认识论信念的增长？

研究问题 2：小学生是如何参与协作知识建构过程并提升自己的认识论信念的？

2.研究方法

2.1.研究环境和教学设计

本研究在华南地区的一所小学进行的。参与者为 28 名小学五年级的学生，他们在为期 8 周（每周 2 节课，每节课 40 分钟）的课程中体验基于知识建构的科学探究过程，探索关于“光”的科学知识。

学生参与的知识建构过程应用了基于先前知识建构研究（Yang，2020）修改的三阶段教学设计。三阶段的探究过程主要在知识论坛（Knowledge Forum）上进行。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，学生分为不同的兴趣小组，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。学生在小组内讨论交流彼此的想法，共同体知识也随之进步。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。学生反思自己的想法和班级体的共同知识是如何在讨论中得到提升的。

此外，在教学过程中，学生们参与了多轮学习分析支持的反思评价过程，教师把学生在知识论坛参与数据（包括发表笔记数，建构笔记数，阅读笔记数，who-read-who 以及 who-build-on-who）的可视化呈现给学生。基于这些可视化的学习分析数据，学生反思自己上一周的探究行为，改进自己下一步的探究行为和探究计划。

2.2.数据来源和分析

本研究采用波普尔的世界 3 框架来描述学生的认识论信念。在教学开始前以及教学结束后，每位学生被要求写下他们对“我们为什么需要想法？想法可以做什么？”和“想法如何改进？”的看法。本研究使用 Hong 和 Chiu（2016）采用的世界 3 框架对学生写下的答案进行分类（见表 1），将那些将“想法”视为个人心理表征的学生（世界 2，低认识论信念组）归为一类，将“想法”视为小组讨论产生的集体创作的学生（世界 3，高认识论信念组）归为另一类。所有学生的认识论信念由两位研究人员同时编码，Cohen 的 kappa 一致性系数为 0.814。

3.研究结果

3.1.学生认识论信念的提升

在协作知识建构开始前，有 14 名学生将想法视为个人观点（世界 1/2，低认识论信念），14 名学生将想法视为小组讨论产生的集体创作（世界 3，高认识论信念）。在协作知识建构结束后，有 6 名学生持有基于世界 2 的认识论信念，22 名学生持有基于世界 3 的高认识论信念。统计检验表明，实验前后高低认识论信念的人数分布存在显著差异（$\chi 2(1) = 5.33, p = .021 < 0.05$）。这说明，基于知识建构的协作探究学习可以有效促进学生认识论信念的提升。

3.2.学生的协作知识建构过程与认识论信念的发展

我们分析了学生在知识论坛对话中的关键事件、反思活动期间学生的反思学习单以及反思活动后的访谈，从质性的角度探讨学生在协作知识建构中的参与情况。

主题 1：提出初步问题和想法。在第一阶段，学生在知识论坛上发布了关于“光”的问题和初始想法。

主题 2：寻求初步解释并进行理论构建。对相同问题感兴趣的学生在知识论坛上进行小组探究。在这一过程中，通过与其他同学的对话，学生们加深了对这些问题的理解，并且逐渐开始改变自己的想法和观点。比如原本坚持“光是沿直线传播”想法的学生，在通过和同学的共同探讨后，开始意识到“在微观世界中，光是以波的形式存在的。”学生开始意识到，知识和想法是可以通过讨论过程改变的。

表 1

基于波普尔世界 3 理论的学生认识论信念编码

	级别 1（世界 1/2）	级别 2（世界 3）
我们为什么需要想法？想法如何发挥作用？	这一级别关注想法的基本功能，强调它们如何作为思考、沟通和行动的基础。这一类别中的许多回答传达了想法的功能方面，但缺乏对其重要性的更深入的思考和分析。此类回答通常表明对想法在认知过程中的作用只有初步的理解，而没有批判性地思考其更广泛的含义。 示例： ●如果我们没有想法，我们的大脑就会退化；如果我们的大脑退化，我们将面临严重的后果。(S3) ●想法帮助我们加深理解并协助解决问题。(S5) ●拥有想法使我们能够了解真相并获得更多知识。(S13)	这一级别的回答表明学生对想法在理解和创造过程中的作用有了更深入的理解。学生开始反思他们与他人的互动和知识获取，表明他们意识到了知识构建的协作性质。这一类别反映了学生整合多元观点的能力，并展现了他们对知识工作者社群中动态构想的更高层次的理解。 示例： ●拥有想法可以改进原创答案，并有助于表达个人观点，这些观点可以与他人的观点相融合。(S6) ●想法用于表达个人观点；只有分享我们的想法，我们才能理解他人的观点。(S9)
想法可以改进吗？如果可以，如何改进？	属于这一类别的回答主要关注学生的个人想法和认知状态。它们强调个人努力和反思在改进想法中的作用，以及寻求他人反馈的重要性。然而，它们并未融入知识社群的概念。这些回答通常反映了学生的自我意识和积极主动的学习态度，并借鉴他们的个人知识和经验。 示例： ●如果有可用的资料，想法可以改进。(S5) ●可以通过在线研究并将他人的想法	相比之下，这一类别强调文化和社会知识的共享。世界 3 下的回答强调了互动、讨论和外部反馈对于提升社区理念的重要性。这些回答展现了社区意识，认识到外部环境中知识和资源的存在，并说明了合作与共享如何能够丰富集体理解。 示例： ● 通过与同学讨论，可以改进想法，使其更加完善和细化。(S6) ● 想法可以得到改进，使我们能够集体转变思维。(S11)

融入我们自己的想法来改进想法。
(S3)

主题 3：进行科学实验或构建科学模型。学生分小组开展科学实验或者构建科学模型，以更好地了解有关“光”的科学知识。学生们在知识论坛上就实步骤和模型构建的细节展开进一步讨论。科学实验或模型搭建的步骤和细节在讨论中被逐步完善。在这一过程中，学生意识到通过科学实验进一步扩展知识，想法是可以不断发展的。

主题 4：分享实验结果，整合班级知识。在完成实验或者模型搭建后，每个学生小组在班级中展示自己的实验报告或者模型报告，并且在班级视窗中总结每个小组探究到的关于“光”的知识，反思想法和知识的改变过程，并且就新产生的问题开展新的讨论探究。

4.讨论与总结

本研究探究了协作知识建构对小学生认识论信念的促进作用。研究结果表明，协作知识建构可以显著提升小学生的认识论信念。认识论信念与波普尔世界 3 相符的学生意识到了集体知识发展的重要性，他们将想法视为暂时性的主张，需要持续的批判性评估、澄清和潜在的证伪才能发展。因此，他们更有动力参与协作讨论，整合不同的想法，应对不同的批评，并推进共同体知识的发展；同时他们展现出更高阶的认知行为，包括第一阶段的解释等高级认知过程，以及第二阶段和第三阶段的调节和综合等元认知策略 （Zhang, 2024）。在协作探究过程中，那些原本持知识固定观的学生看到并且参与了共同体知识的发展，他们转而相信知识具有可塑性，开始参与协作知识建构过程，并通过使用自己的语言重构信息来加深理解。下一步研究将采用更细致的研究方法深入学生从低认识论信念到高认识论信念的转变过程。

参考文献

Bruning, R. H., Schraw, G. J., & Ronning, R. R. (1999). Cognitive psychology and instruction, perception of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology*, *93*(1), 43–54.

Chen, F., & Chen, G. (2024). Technology-Enhanced Collaborative Inquiry in K–12 Classrooms: A Systematic Review of Empirical Studies. *Science & Education*. https://doi.org/10.1007/s11191-024-00538-8

Crook, C. (1991). Computers in the zone of proximal development: Implications for evaluation. *Computers & Education*, *17*(1), 81–91. https://doi.org/10.1016/0360-1315(91)90075-3

Dewey, J. (1986). Experience and Education. *The Educational Forum*. https://doi.org/10.1080/00131728609335764

Hong, H.-Y., & Chiu, C.-H. (2016). Understanding how students perceive the role of ideas for their knowledge work in a knowledge-building environment. *Australasian Journal of Educational Technology*, *32*(1), Article 1. https://doi.org/10.14742/ajet.1873

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*(3), 309–328. https://doi.org/10.1016/S0885-2014(00)00030-7

Liu, K., Chan, C., & Wang, S. (2016). Fostering Students' Epistemic Beliefs and Scientific Understanding through Knowledge Building.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In *The Cambridge handbook of the learning sciences* (pp. 97–115).

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*, *51*(4), 1136–1154. https://doi.org/10.1111/bjet.12909

Zhang, X., Feng, S., Chan, C. K. K., & Zhao, J. (2024). The Relationships between Students' Epistemological Beliefs and Social Interactions and Epistemic Actions in Collaborative Inquiry Practice. In 1*7th International Conference on Computer-Supported Collaborative Learning (CSCL) 2024, pp. 209–212.* https://doi.org/10.22318/cscl2024.596913

Zhou, F., Chai, S., He, Z., & Wu, H. (2025). Exploring the development of pre-service teachers' epistemic agency in Chinese University knowledge building community. *Learning, Culture and Social Interaction, 52*, 100904

变革型领导与智能教学评价系统的有效整合——基于组织学习理论的多案例研究

Transformational Leadership and the Effective Integration of Intelligent Teaching Evaluation Systems：A Multiple-Case Study Based on Organizational Learning Theory

孙子涵，赵建华*

南方科技大学未来教育研究中心

* zhaojh@sustech.edu.cn

【摘要】 在教育智能化与大数据时代，智能教学评价系统的有效整合成为学校改革要务。本文基于变革型领导与组织学习理论，选取三所 K-12 学校进行多案例研究。结果显示：①领导者的变革型行为（愿景塑造、个性化关怀、智力激发）显著提升教师对智能评价技术的接受度；②当学校具备稳定学习共同体与反思机制，组织学习在领导与系统落地间发挥关键中介作用；③在资源、师资与策略差异下，应用效果有别。基于样本与场景局限，后续可在更广教育情境中开展量化研究，并关注伦理与数据安全。研究结论可为教育管理者、政策制定者及技术提供方提供参考。

【关键词】 变革型领导；智能教学评价系统；组织学习；多案例研究；教育管理

Abstract: *In the era of educational intelligence and big data, effectively integrating intelligent teaching evaluation systems has become a key priority in school reforms. Grounded in transformational leadership and organizational learning theories, this study employs a multiple-case approach in three K–12 schools. Findings indicate that (1) leaders' transformational behaviors (vision-setting, individualized consideration, and intellectual stimulation) significantly enhance teachers' acceptance of intelligent evaluation technologies; (2) when robust learning communities and reflective mechanisms are in place, organizational learning serves as a critical mediator between leadership and system implementation; and (3) variations in resources, teacher expertise, and strategic approaches lead to differing outcomes. Given the limitations of sampling and context, future research should broaden its scope, adopt quantitative methodologies, and address ethical and data security concerns. The study's conclusions offer valuable insights for educational administrators, policymakers, and technology providers.*

Keywords: transformational leadership, intelligent teaching evaluation systems, organizational learning, multiple-case study, educational management

1.研究背景与研究问题

伴随学习分析与生成式 AI 的深度融入，学校评价范式正由以学期末结果为中心的“结果评估”迈向依托实时数据流支撑决策的“数字生态”（CoSN, 2024；Karaköse & Tulubas, 2024；Khosravi et al., 2025；National Center for Education Statistics, 2024）。智能教学评价系统（Intelligent Teaching Evaluation Systems, ITES）如今可同步采集学习轨迹、情感线索与生成式对话记录，为教师提供即时干预策略（Kim & Hwang, 2021；

Zawacki-Richter, Marín, Bond, & Gouverneur, 2019）。国内多省级智慧教育示范项目亦已落地同类系统：若干城市的区级“智慧课堂”工程、市级“AI 应用场景”试点以及国家中小学智慧教育平台的“AI 学习”试验校，均在常态课堂中实时汇聚点击流、作业-测验细粒度记录、语音情感与多模态交互数据，并通过教师端仪表盘推送分层处方，实现“课中-课后一体化”精准干预（教育部, 2024；北京教委, 2024；广东省教育局, 2024）。随着此类项目的推广，教育者对 ITES 在精准教学与个性化支持中的价值预期进一步提升（Ali, 2022；Su, Guo, Chen, & Chu, 2023）。然而，多数学校仍面临资源不足、教师数据素养参差、技术培训缺位与文化阻力等瓶颈，导致系统“部署易、深度难”——即便上线，也常停留在浅层使用或“合规式”打卡（赵建华，2024；Ifenthaler & Yau, 2020；Fitria & Suminah, 2020）。最新教育数据综述亦指出，校长与教师在数据解读与伦理治理能力上的差距，是 ITES 难以落地的关键症结之一（Karaköse & Tulubas, 2024）。

在推动技术采纳的诸多因素中，教育领导力展现出显著杠杆效应（Fullan, 2014）。尤其是变革型领导（Transformational Leadership, TL），通过愿景塑造、个性化关怀与智力激发，可有效提升教师技术采纳动机（Bass & Avolio, 1994；Leithwood & Sun, 2012）。最新的大样本实证研究显示，TL 对教师 ICT 整合的影响达到中等效应量：以 5,036 名教师为样本的跨国结构方程模型报告其总间接效应 $\beta = 0.31$，对应 Cohen's $d \approx 0.45$（Schmitz et al., 2023）。这一结果与先前针对学校领导力与技术采纳相关性的综合分析 $r \approx 0.24$（折算 $d \approx 0.48$；Safi, 2019）基本一致；当校长同时具备数据领导力时，教师对 ITES 的持续使用亦显著增强（Sung, Kang, & Liu, 2023）。然而，领导者的号召若缺乏配套的组织学习（Organizational Learning, OL）机制，往往难以转化为课堂层面的持续改进（Argyris & Schön, 1978；Senge, 1990）。OL 不仅涵盖基于数据的“单环纠错”，更需上升到“双环反思”——即质疑既有教学假设、重构课程目标，方能充分发挥 ITES 的诊断价值。

现有研究仍存在两大空白：

单变量视角——要么探讨领导力对技术采纳的直接效应，要么聚焦教师学习共同体的重要性，缺乏实证整合 “TL—OL—ITES” 三者动态关系（Robinson, 2010；Sung et al., 2023）。

机制阐释不足——尚未明确“领导 → 组织学习 → 技术应用”链条中的资源、动机与能力路径（Gumus et al., 2024）。

为弥合上述缺口，本研究借鉴资源基础观（Resource-Based View, RBV；Barney, 1991）与自我决定理论（Self-Determination Theory, SDT；Deci & Ryan, 2000），提出“资源—动机—能力（R-M-C）链条”：

资源（R）源于 RBV，强调领导者为教师提供技术、时间与培训等有价值且难以模仿的支持要素；

动机（M）对应 SDT 中的内在动机与认同调节，体现教师对数据驱动教学的自主投入；

能力（C）指通过组织学习提升的数据素养与教学重构技能，最终转化为 ITES 的深度融合。

同时，将“双环学习”视作 OL 成熟度的核心标志，用以衡量教师群体是否能在数据驱动循环中反思并重构教学假设（Argyris & Schön, 1978）。仅有形式化的教研活动无法促成双环学习，唯有教师在共同体中通过持续对教学信念的反思与重构，方能真正实现从‘使用系统’向‘改进教学’的跃迁。据此，研究问题聚焦于：变革型领导如何经由组织学习（特别是双环学习）及 R-M-C 链条的中介作用，促进 ITES 在学校层面实现可持续落地？本综合视角旨在填补“领导力—学习机制—技术应用”交互研究空白，提出 R-M-C 链条整合模型（图 1.）为数字生态时代的学校治理提供理论与实践启示。本研究假定组织学习机制在 TL 与 ITES 之

间发挥中介作用，且 R-M-C 链条中的‘资源—动机—能力’维度可作为机制展开路径。详见图 1 理论模型。

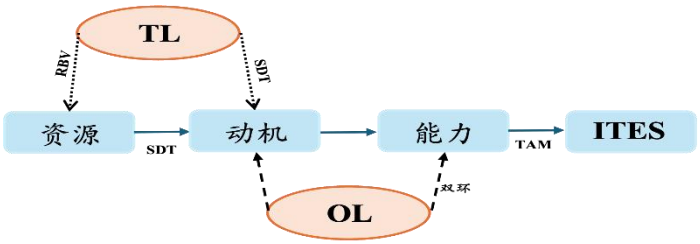


图 1 R-M-C 链条整合模型

2.问题的提出和研究目标

基于文献综述，本研提出以下核心研究问题：

RQ1: 在智能教学评价系统（ITES）的引入与应用过程中，变革型领导的主要特征如何影响学校组织的学习与创新能力？

RQ2: 组织学习在变革型领导与智能教学评价系统有效整合之间是否发挥中介或促进作用？如果是，具体机制如何？

RQ3:. 通过多案例研究，能否提炼出学校领导者在应用智能教学评价系统时的关键行动策略与实践路径？

研究方法

研究设计

为探究变革型领导与智能教学评价系统在学校组织中如何实现有效整合，本文采用多案例研究（Multiple Case Study）的方法（Yin, 2018）。此研究设计有助于对不同学校的领导风格、组织文化与技术使用环境进行纵深比较，并在真实情境中分析组织学习过程的关键环节。本研究选择了三个有代表性的 K-12 学校作为案例：学校 A 引入了成熟的智能测评平台且由一位倡导创新的校长领导；学校 B 处于初始尝试阶段，领导者对新技术的态度谨慎；学校 C 已经应用智能教学评价系统超过两年，并在区级层面获得表彰。三所学校分别位于城市核心区、新兴发展区与边远郊区，呈现出显著的区位资源差异；同时，三位校长的领导年限与风格在前期访谈中已表现出变革型领导各维度的显著差异性，具备典型性与代表性

数据收集与分析

本研究主要通过以下渠道收集数据（资料来源）：

1. 半结构化访谈：对校长、教务主任、学科组长及一线教师进行访谈，重点关注领导者在推动智能教学评价系统落地过程中的角色与具体举措，以及教师团队对组织学习的感知与实际行为。

2. 文件与政策文本：收集并分析学校层面关于教学改革、技术应用、教师培训等方面的制度文件以及区教育局相关政策。

3. 系统使用记录：获取各校智能教学评价系统的使用频率、学生测评数据反馈情况，以及教师对系统所进行的二次分析（secondary analysis actions）与改进记录。

分析流程：

1. 案例内分析（Within-Case Analysis）：对每个案例学校单独进行资料整理与编码，围绕变革型领导的核心要素（如愿景塑造、个性化关怀、智力激发）以及组织学习的过程（如知识获取、知识共享、持续改进）进行主题分析（Saldana, 2016）。
2. 跨案例分析（Cross-Case Analysis）：比较三个学校在领导行为、教师参与度、组织文化以及系统使用成效上的差异与共性（Yin, 2018）。特别关注变革型领导与组织学习交互作用对系统整合效果的影响。
本研究遵循最大差异取样原则（Yin, 2018），并通过双研究者独立编码（Cohen's κ = 0.82）与成员检验保证质性信度；系统日志另以配对 t 检验检视 8 周使用趋势。

3. 研究结果与讨论

案例主要发现
一、案例学校 A：高水平领导激发教师积极创新
1. 领导行为与教师使用状况
学校 A 的校长持续倡导创造性教学与"全员学习文化"（Bass & Avolio, 1994），并通过每两周一次的"行动学习小组"（Robinson, 2010）鼓励教师对智能测评平台进行深度使用与经验分享。
系统使用日志（见表 1.节选）显示，在推广初期（第 1-2 周），教师平均登录次数仅为 2.1 次/人，学生测评完成率为 85%，教师对测评数据进行"二次分析"的次数为 3 次。随着学校组织了系统培训和月度教学分享会，到第 7-8 周时，教师平均登录次数提高到 4.8 次/人，学生测评完成率则达到了 95%，教师对测评数据的二次分析次数跃升至 15 次。这反映出领导层持续关注与教师协作学习机制对平台使用的正向促进作用。

表 1 系统使用日志（节选）

周次	教师平均登录次数/人	学生测评完成率	二次分析次数	备注
第 1-2 周	2.1	85%	3	刚开始推广，部分教师不熟悉操作
第 3-4 周	3.7	90%	7	学校组织了第一次培训
第 7-8 周	4.8	95%	15	校长公开表扬优秀案例，积极性提高

*注：测评完成率 = 完成测评学生数 / 应测评学生数；二次分析指教师在平台中点击'诊断报告'并生成教学调整记录的次数。

2. 组织学习过程与教师专业成长
访谈数据显示，多数教师认为"行动学习小组"在讨论测评数据、反思教学策略方面帮助显著。"我们组会把学生错题分布列成表格，然后一起商量改进教学"，这是数学组组长（T1-A）的访谈原话，也可从教务主任（M1-A）访谈中得到印证：
"月度教研会上，老师们会展示测评结果对课堂调整的具体案例，这种常态化分享让教师之间的学习变得更加深入。"
与此同时，新进教师在老教师的协助下能快速掌握数据解读技巧，说明校内形成了相对成熟的"师徒结对"或"同伴互助"文化。此种组织学习氛围下，教师采纳意愿和数据分析深度不断提升。
二、案例学校 B：谨慎推进下的技术局部应用

1. 领导观望与使用成效

相比之下，学校 B 的领导者虽认可智能测评的潜力，但更关注资金投入与教师负担，无明确的整体愿景或统一规划。

从系统使用日志（见表 2.节选）可见，在第 1-2 周试点期间，教师平均登录次数仅 1.2 次/人，学生测评完成率约 70%。到第 7-8 周，教师平均登录次数提升至 2.0 次/人，学生测评完成率 78%，仍处于偏低水平。

表 2 系统使用日志（节选）

周次	参与试点教师人数	教师平均登录次数/人	学生测评完成率	二次分析次数	备注
第 1-2 周	3	1.2	70%	1	仅在两个班级试点
第 3-4 周	3	1.4	68%	2	系统公司线上培训
第 7-8 周	5	2	78%	3	无校级层面大规模推广

2. 组织学习局限与教师反馈

访谈显示，教师对系统的实际应用偏向消极或“应付式”。一位英语教师（T1-B）表示，“领导没有要求定期汇报数据使用成果，也没提供足够的技术支持，我只能在空闲时随便看看。”由于缺少明确的奖励或考核机制以及系统化培训，大部分教师缺乏动力进行深入分析和教学改进。学校的学习小组更多停留在形式层面，尚未形成解决教学问题的常态化机制。

这与校长（P1-B）访谈中所谈到的谨慎态度相呼应：“我还是希望老师自发使用，但效果不如预想好。”说明领导层在引导与资源支持上缺乏系统设计，从而导致整体应用效果差强人意。

三、案例学校 C：持续改进与深度融合

1. 领导激励与多元支持

学校 C 已经持续应用智能测评平台两年，并在区级评比中获得表彰。结合访谈和系统使用日志，能更清晰地看到校长在政策、资源与文化营造上的高投入。

系统使用数据（见表 3.节选）显示，教师平均登录次数从第一学期初（2.5 次/人）提高到第二学期末（4.9 次/人）；学生测评完成率也从 88%上升至 97%。与此同时，教师对平台进行二次分析的次数由 5 次增长到 15 次，这与该校持续强化的“问题解决社群”密切相关。

更值得注意的是，学校将部分测评数据与期末考试平均分、课堂观察结果等指标进行综合分析，发现总体学业表现也呈现一定的提升趋势（从 72 分到 80 分）。虽然此变化不必然完全归功于测评系统，但能说明在领导力、组织学习与技术系统三者协同作用下，整体教学质量正向发展。

表 3 系统使用日志（节选）

学期阶段	教师平均登录次数/人	学生测评完成率	二次分析次数	学期末考试平均分
第 1 学期初(上年)	2.5	88%	5	72
第 1 学期末(上年)	4	94%	10	75
第 2 学期初(本学年)	4.3	95%	11	78

2. 深化组织学习与双环反思

在访谈中，语文骨干教师（T1-C）提到：“我们每周会针对测评结果做数据分析，看到学生普遍薄弱的知识点，就会及时调整教案和训练方式。”

教师不仅停留在“单环学习”（单纯根据数据纠正错误），更能结合课程目标、教材结构和学生特点开展“双环学习”（Argyris & Schön, 1978），不断审视自身教学假设，做更深层次的改进。

校长（P1-C）通过持续激励与制度建设（如“智慧教学指导小组”、“定期反馈会”），为教师创造了一个积极反思、主动合作的氛围（Leithwood & Jantzi, 2005）。这正体现了变革型领导在培育学校学习文化过程中的核心作用。

综上：数据支撑下的差异性分析

从三个案例学校的对比来看，领导行为与组织学习机制的成熟度在很大程度上决定了智能测评平台的使用深度与推广成效。学校 A 和学校 C 的校长均展现出较强的变革型领导特质，如愿景塑造、资源支持与持续激励等，教师使用率与学习分析深度随之显著提高。而学校 B 由于领导者缺乏明确愿景、教师激励与合作学习不足，系统的整体应用效果在量化指标（登录次数、完成率、二次分析次数等）上始终低迷。以上数据与访谈证据相互佐证，为研究结论提供了有力支撑，也为后续对领导力、组织学习与智能教学评价系统的交互关系提供了可操作的经验与反思。

4.跨案例比较与讨论

综合三个学校的案例可见：

变革型领导对系统落地的核心作用：当领导者能够提供清晰的创新愿景并营造积极氛围时，教师普遍表现出更高的使用意愿，并在培训、协作和反馈环节更加投入（Fullan, 2014）。

组织学习在领导与评价系统之间发挥关键中介：当教师团队具备有效的学习机制，如定期研讨、行动学习小组或数据分析培训，“技术引入—实践反思—教学改进”的循环更易形成良性反馈（Senge, 1990）。反之，缺乏系统性的学习共同体建设，往往导致智能教学评价系统形同“摆设”，难以深入课堂并持续改进。

情境因素的差异：不同学校在资源配置、师生结构、管理策略等方面存在差异，使变革型领导的实践模式与组织学习的进程呈现出不同表现形式（Leithwood & Sun, 2012）。在资源充足且领导者具有较高“变革意志”的学校，更能快速实现系统与教学的有效整合。

研究结果

研究结论

针对前文提出的研究问题（RQ1、RQ2、RQ3）,本研究通过多案例比较发现：

RQ1：变革型领导的核心特质（愿景塑造、个性化关怀、智力激发等）能够显著提升教师对于智能教学评价系统的采纳意愿，并为组织学习创造必要的文化氛围。

RQ2：组织学习在领导行为与评价系统应用成效之间发挥了关键中介和促进作用。仅有领导者的倡导或制度支持还不够，必须有与之相匹配的教师学习社群与持续反思机制。

RQ3：有效整合离不开领导者的整体规划与执行力，如定期召开数据反馈会、设立跨学科教师研修组、给予持续性的资源与激励。这些策略在不同情境下可灵活调整，但变革型领导与组织学习始终是两个关键支柱。

实践意义

对教育管理者：需要充分认识到领导行为在技术变革中的引领作用，积极营造合作氛围与学习文化；通过定期反思与跨部门协作，确保系统的持续改进与创新。

对政策制定者：应在资源配置和培训体系上予以支持，如拨款购置高质量测评平台、搭建教师专业发展社区等。

对技术提供方：智能教学评价系统的设计与维护需更贴近学校一线需求，并通过多维度的数据可视化与反馈机制来辅助组织学习过程。

研究局限性与未来展望

局限性：本研究的案例样本主要集中于 K-12 领域，且规模相对有限，难以代表更广泛的教育场景。同时，多案例研究在分析广度与深度的平衡上可能存在局限，后续可结合量化方法或大规模调查进行更系统的验证。

未来研究方向：一方面，可尝试对变革型领导与其他领导风格（如分布式领导）的交互效应进行探讨；另一方面，也可对教师个体层面的技术素养、信念与职业倦怠等因素开展纵深研究。此外，针对系统应用过程中的隐私保护、数据伦理与公平性等问题，也亟待进一步的探讨与实践检验。

参考文献

北京市教育委员会 (2024). 北京市教育委员会等四部门关于印发《北京市教育领域人工智能应用工作方案》的通知 (京教信〔2024〕14 号). 北京市教育委员会.

广东省教育厅. (2023). 广东省教育数字化转型行动方案（2023-2027）. 广东省教育厅.

教育部办公厅. (2024). 关于深化人工智能中小学智慧教育的通知 (教基厅函〔2024〕15 号). 北京: 教育部.

赵建华. (2024). 人工智能时代的教育转型与重塑. 电化教育研究, (12), 37-97.

Ali, W. (2022). *Artificial intelligence in education: Opportunities and challenges for personalized learning. Educational Technology Review*, 30(2), 45- 61.

Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective.* Addison- Wesley.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management,* 17(1), 99- 120. https://doi.org/10.1177/014920639101700108

Bass, B. M., & Avolio, B. J. (1994). *Improving organizational effectiveness through transformational leadership.* Sage.

Su, J., Guo, K., Chen, X., & Chu, S. K. W. (2023). Teaching artificial intelligence in K–12 classrooms: A scoping review. Interactive Learning Environments. Advance online publication. https://doi.org/10.1080/10494820.2023.2212706

Consortium for School Networking. (2024). Driving K- 12 innovation: 2024 hurdles, accelerators, and technology enablers. CoSN.

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self- determination of behavior. *Psychological Inquiry, 11*(4), 227- 268. https://doi.org/10.1207/S15327965PLI1104_01

Fitria, H., & Suminah, S. (2020). Role of teachers in digital instructional era. *Journal of Social Work and Science Education, 1*(1), 70–77. https://doi.org/10.52690/jswse.v1i1.11

Ifenthaler, D., & Yau, J. Y.- K. (2020). Utilising learning analytics for study success: Reflections on current empirical findings. *British Journal of Educational Technology, 51*(5), 1665- 1676. https://doi.org/10.1111/bjet.12927

Karaköse, T., & Tulubas, T. (2024). School leadership and management in the age of artificial intelligence (AI): Recent developments and future prospects. *Educational Process: International Journal, 13*(1), 7–14. https://doi.org/10.22521/edupij.2024.131.1

Khosravi, H., Shibani, A., Jovanovic, J., Pardos, Z. A., & Yan, L. (2025). Generative AI and learning analytics: Pushing boundaries, preserving principles. *Journal of Learning Analytics, 12*(1), 1–11. https://doi.org/10.18608/jla.2025.8961

Leithwood, K., & Jantzi, D. (2005). A review of transformational school leadership research 1996–2005. *Leadership and Policy in Schools, 4*(3), 177-199. https://doi.org/10.1080/15700760500244769

Leithwood, K., & Sun, J. (2012). The nature and effects of transformational school leadership: A meta-analytic review of unpublished research. *Educational Administration Quarterly, 48*(3), 387-423. https://doi.org/10.1177/0013161X11436268

National Center for Education Statistics. (2024*). Forum guide to taking action with education data* (NCES 2024-812). U.S. Department of Education.

Robinson, V. M. J. (2010). From instructional leadership to leadership capabilities: Empirical findings and methodological challenges. *Leadership and Policy in Schools, 9(*1), 1-26. https://doi.org/10.1080/15700760903026748

Safi, A. B. (2019). ICT and integration in education: A meta-analysis. *Journal of Emerging Technologies and Innovative Research, 6*(6), 117–124. https://www.jetir.org/papers/JETIR1906F23.pdf

Saldana, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Schmitz, M.-L., Antonietti, C., Consoli, T., Cattaneo, A., Gonon, P., & Petko, D. (2023). Transformational leadership for technology integration in schools: Empowering teachers to use technology in a more demanding way. *Computers & Education, 204*, 104880. https://doi.org/10.1016/j.compedu.2023.104880

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization.* Doubleday.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

人工智能赋能高等教育课程教学创新研究进展综述

A Review of Research Progress on the Innovation of Higher Education Curriculum Teaching Empowered by Artificial Intelligence

王雷岩[1]，赵建华[2*]
[1]南方科技大学教学工作部
[2]南方科技大学未来教育研究中心
[*]zhaojh@sustech.edu.cn

【摘要】 新一代人工智能技术凭借其强大的通用性、渗透性和颠覆性，为高等教育带来了深刻变革。本文基于文献分析法，对国内外2020年以来关于人工智能赋能高等教育课程教学的核心期刊论文进行梳理，根据研究重点不同分为人工智能教学模式创新、人工智能与课程建设以及人工智能与学生发展等应用场景，并对每一类的重要研究观点进行梳理与介绍，以试图呈现该领域的整体脉络。在此基础上，提出人工智能赋能高等教育课程教学已初见成效，正从技术工具应用向教育生态重构演进，未来研究应聚焦于人工智能在高等教育课程教学的落地使用等方面。

【关键词】 人工智能；高等教育；数字化转型；教育场景；赋能理论

Abstract: *The new generation of artificial intelligence technology, with its powerful versatility, permeability, and subversiveness, has brought about profound changes to higher education. Based on the literature analysis method, this paper combs through the core journal papers at home and abroad since 2020 on the teaching of higher education courses empowered by artificial intelligence. According to different research focuses, it divides the application scenarios into the innovation of artificial intelligence teaching models, the integration of artificial intelligence and curriculum construction, and the role of artificial intelligence in student development. It also combs through and introduces the important research viewpoints of each category, aiming to present the overall context of this field. On this basis, it is proposed that the empowerment of artificial intelligence in higher education curriculum teaching has achieved initial results, and future research should focus on aspects such as the practical implementation of artificial intelligence in higher education curriculum teaching.*

Keywords: artificial intelligence, higher education, digital transformation, education scenarios, empowerment theory

1.研究背景

自2022年11月ChatGPT发布以来，世界各国政府、高等教育管理机构、大学、科研机构与企业等社会各界都参与到人工智能（Artificial Intelligence，AI）赋能高等教育的这场变革中，提出了各类战略指南、计划、管理规范等，并进行了实践探索。充分发挥以生成式人工智能为代表的新一代信息技术优势，深化高等教育系统性改革，已成为世界各国推进教育数字化转型的重要任务，是教育改革发展的必然趋势。

2024年，教育部启动人工智能赋能教育行动，对推动人工智能在教育领域的深度应用作出了总体战略性部署。高等教育是建设教育强国的龙头，高校必须加快探索人工智能在高等教

育领域的深度应用。目前，国内部分高校在通用大模型的基础上，积极开发和应用面向大学学科/专业的垂域大模型，探索和实践智能体在教育领域的深度应用(张立群, 2024)。
从技术演进的逻辑来看，AI 前沿技术的深入应用必然改变大学各项功能的实现方式，并进一步出发大学存在的形态的深层次变革——人工智能的发展有望使因材施教、个性化学习、创新迸发的教育理想变成现实(杨满福等, 2025)。本研究聚焦于梳理人工智能赋能高等教育课程教学的现有经验，深入思考技术与教育融合的未来方向，对人工智能赋能高等教育发展具有重要意义和价值。

2.研究方法

本研究基于文献调研，“人工智能”、“高等教育”、“课程教学”等为关键词，在 CNKI 数据库中，时间节点界定为 2020 年-2025 年，“以核心期刊”为筛选条件，检索得到结果 48 个，手动排除与本研究相关度较低的结果 6 个，得到期刊论文 42 篇。利用 CiteSpace 对国内 42 篇论文进行关键词分析，发现国内关于人工智能与高等教育课程教学的研究多集中教学实践创新、课程建设、学生发展支持在等方面。本文尝试从以上几个方面对当前的研究进展进行梳理与总结。

3.人工智能与高等教育教学模式创新

人工智能与高等教育深度融合的实践场景(杨现民等, 2025)，生成式人工智能从“三层面十维度”重塑高等教育形态，包括以课堂教学、课外实践、在线学习为代表的教育场景，课堂教学场景重塑：创建生成式智慧课堂环境，促进学习体验智联融通；课外实践场景重塑：创生智能孪生实践环境，培育跨学科实践综合能力；课外实践场景重塑：创生智能孪生实践环境，培育跨学科实践综合能力(刘明等, 2024)。

3.1.人工智能深化人机协同教学模式

（1）人工智能+教育教学，在创新教育教学方法（教学模式创新）上，教师将人工智能机器人或虚拟教师作为课堂助教，开展人工智能支持的人机协同教学，实现教师与人工智能协同互补。清华大学“化工热力学”课程人工智能助教建设框架流程，包括教学资源准备、结构化数据处理、垂直模型训练、模型测试与评估等主要环节。目前“化工热力学”课程人工智能助教已实现包括自适应出题、知识点答疑、题目解答、曲线图绘制和手写公式识别与推导计算等在内的多样化功能，初步具备了一门普通课程助教的基础能力。在教师端，人工智能助教实现了课前、课中、课后的全过程辅助。在课程准备阶段，人工智能助教可为教师提供备课资料与灵感；在课程教学过程中，人工智能助教可实时监测学生的学习进度和表现并提供即时反馈；在课程之后，人工智能助教则可根据整体授课情况进行评估反馈，支持教师教学策略的迭代优化。在学生端，人工智能助教能够同时赋能学生的低阶和高阶认知发展，一方面精准地识别并辅导学生的知识短板，为学生的知识理解和理论掌握提供脚手架支撑；另一方面创造性地提供问题解决方案，支持学生综合应用学科知识解决复杂问题，为学生创新思路和方法提供灵感。通过全流程的参与、监督、反馈，人工智能助教大大提高了“化工热力学“的课程教学效果和学生的学习效率(卢滇楠等, 2024)。人机协同学习的关键是如何赋予智能体探究型教师的能力，诱发与引导学习者与其协同探究。李海峰等构建了人机协同深度探究性学习理论模型。该模型将素养发展作为人机协同学习的根本目的，利用智能问题设计诱发学习者探究的能动性，促进人机协同认知冲突的发生、发展与持续；在学习画像行为组织、学习分析监控评价、人机协同学习环境和机制的共同作用下，实现人机协同深度探究性学习。基于人机协同深度探究性学习模型及其系统，以翻转课堂课前、课中和课后的教

学过程为基础，构建人机协同深度探究性教学模式。课前人机协同任务探究、课中人机协同问题解决与作品创新、课后人机协同作品完善与分享。教学模式的实施需关注如何启动和持续促进人机协同深度探究性学习这一核心问题。针对这一问题，研究者提出了两项策略：人机协同深度探究启动策略和人机协同涟漪拓展探究策略(李海峰等, 2023)。教师借助 GAI 变革传统教学模式，积极探索面向大学生高阶思维能力培养的人机协同教学模式及生成性探究式教学，创新人机协同教学模式。课前，GAI 可生成学生学习诊断报告，为教师提供可研究的前测问卷和教学案例；课中，教师引导学生向 GAI 提问，以获取精准信息，开阔问题解决视野并产生新思路；课后，生成个性化作业，助力教师减负增效(刘明等, 2024)。

3.2. 人工智能优化教学全流程

在优化教学流程上，人工智能覆盖备课、授课和课后服务三环节。在备课环节，人工智能可集成、 推荐和生成高质量的教学资源。浙江大学“智海 平台”能将音视频、实训案例等资源按知识点集 成结构化数字教学资源。GAI 具备多模态内容生成能力，有助于生成多模态教学资源，人机资源共创可加快高校资源开发效率。一方面，生成个性化多模态的“新”资源。第一，生成融合虚拟实验、3D 视频等认知类资源，允许学生在情景化实验和模拟操作条件下，获得实践经验和技能。第二，生成虚拟教师、虚拟同伴等情感类资源，包括虚拟人视频生成和虚拟人实时交互；第三，生成方案、策略、调节反馈等元认知类资源，用于教育游戏中非玩家角色逻辑设计和教育评价个性化反馈(刘明等, 2024)。张悦周面向非人工智能专业类学生设计并实践了一种新型教学模式，即“AI 教授-教师引导-学生自学”，该模式以智能工具为主导，学生在教师引导下利用这些工具完成课堂专业知识的学习与探索(张悦周等, 2024)。智能时代的深度课堂在内容获取和基于问题的交互方面给师生带来了挑战，本研究归纳了智能时代深度课堂实施路径包括 ChatGPT 的深度课堂要素模型和要素模型支持下的课堂设计流程。ChatGPT 的深度课堂要素模型分三层，其中内核是以“四何”问题为任务驱动；深度课堂的作用主体及其角色关系是第二层，即教师、学生、ChatGPT 三主体构成的三元关系；最外层指在驱动任务和角色分配的基础上通过深度课堂目标牵引“内容服务”与“互动支持”，使教学过程在多元角色和生成性内容的支撑下形成闭环。深度课堂实践路径包括六个环节：学习目标分析与拆解、学习者分析、内容选择与交互任务设计、教学过程序列化设计、教学实施与过程性评价以及总结性评价。并以 B 大学“学习科学”课程的研究生为对象开展实践(王琦等, 2024)。

在教学设计方面，人工智能能结合教学目标、学生能力优化教学设计，为授课教师提供定制化的教学设计建议与反馈。在授课环节，人工智能能辅助学习场景创造和支持多维教学互动，增强教学效果和学生学习体验。在课后服务环节，人工智能系统能智能生成符合教学要求和学生能力的试卷和作业，并实现对作业与试题的快速评分与分析。生成式人工智能支持的人机协同教学设计分析，哈尔滨工程大学智能科学与工程学院将人工智能技术融入到 OODA 循环教育教学活动中，“观察”阶段主要采取的技术手段包括大数据技术和知识图谱技术，教师依托教学平台完成学生学习行为数据的收集，通过大数据和知识图谱技术，对学生的需求和学习轨迹进行全面分析，从而更精准、客观、全面地评估学生的学习和发展状况；判断阶段主要采取的技术手段包括大模型技术和 AIGC 技术，教师根据观察阶段的数据分析结果，帮助学生明确自身的发展方向和学习目标，快速生成个性化的学习计划和针对性的教学内容，实现精准定位和个性化培养；决策阶段主要采取的技术手段和上阶段基本一致，为提高教学的灵活性和自适应能力，教师通过大规模数据训练得到最优化模 型,依托训练的模型开展智能决策，辅助制定更为科学有效的教学策略和课程安排。执行阶段主要采取智能交互技术，教师通过智能交互技术提升教学管理的效率，学生通过虚拟实验室和在线学习平台等实现跨

时空、跨地域的协同学习，并能够随时随地对所学知识和技能进行实践，通过智能反馈系统还能够持续调整和优化自身的学习行为，持续提升和发展个人能力(赵玉新等, 2025)。

4.人工智能与课程建设

智能时代高等教育课程建设的核心是数智融合与结构重塑。以课程理论、教学理论、数字教育理论等为指导，提出智能时代高校课程数字化转型的“价值引领—结构重塑—数智赋能”理论逻辑。应用科技与教育双向赋能思维，构建智能时代高校数字课程 D-OCAE 形态特征框架；基于设计思维与系统工程方法，凸显“师—生”二元结构到“师—生—机”三元结构的教学模式变革，形成智能时代高校数字课程“设计—开发—实施”三环构建方法，并提出数字课程的设计模型、开发内容与实施路径(谢幼如等 2024)。数智融合课程指在数智时代新知识观与新人才观指引下，深度融合数智技术与课程教学，以培养新质人才为目标的新型课程形态。数智时代的课程改革建设理念包括聚焦思维与能力的双重培育、基于问题导向的知识体系创新、数智技术赋能个性化学习体验、师生共创情境化学习新生态以及成果导向的课程质量全面评估等。以“健康教育与健康促进”课程为例，提出数智融合课程的建设路径：选择新型交叉学科先行先试，理论与实践、专业与数智密切结合重构课程体系，构建数智化学习环境，构建全方位、个性化与智能化的评估体系。文章最后提出数智融合课程要以个性化学习为基、多元教学法融合和师生协同共创实现(王竹立等 2025)。

5.人工智能与学生发展

人工智能通过全流程支持与数据驱动干预赋能学生成长。在支持学生学习上，人工智能能覆盖学习全过程。课前，人工智能可通过促进学生自主预习、推荐个性化学习资源和规划定制学习路径，激发学生学习兴趣和主动性。哈尔滨工业大学的人工智能学习系统可以实时解答学生课前预习的疑惑，帮助学生把握课程重难点。课中，人工智能能够辅助学生完成课堂任务，开展多样化的课堂实时互动，增强课堂互动和实效。课后，人工智能通过常态化学情诊断、伴随式知识智能答疑和个性化教学指引，为学生持续提供学习支持和延展性学习建议。人工智能还能基于学生的全域数据构建学生专属数字画像，对学生学业状态进行智能预警和干预。普渡大学利用课程信号系统对处于学业危险区域的学生及时预警，并提供及时的干预帮扶(但金凤等 2022)。

6.人工智能赋能高等教育课程教学未来发展与展望

随着人工智能技术的不断进步，高等教育正在经历深刻变革，其与教育的深度融合不仅在个性化学习体验、教学辅助决策等方面产生影响，也正在重塑教育资源的配置方式(徐茂森, 2025)。未来，人工智能在高等教育课程教学中的发展将聚焦于持续优化现有应用，探索新的融合模式与应用场景，以应对当前存在的问题和挑战，推动高等教育的高质量发展。其具体如下：

6.1. 构建多模态教育智能体与自适应学习系统

构建多模态教育智能体与自适应学习系统是未来发展的重要方向。多模态教育智能体融合文本、图像、音频等多种信息，能更全面理解学生需求。例如，智能助教可通过语音、图像等多种方式为学生答疑解惑，提供沉浸式学习体验。自适应学习系统依据学生学习数据，动态调整教学内容和进度。如根据学生的答题情况、学习时长等，为其推送个性化学习资料和练习，满足不同学生的学习节奏。

这一系统还能为教师提供精准教学建议，帮助教师了解学生学习难点，优化教学策略。像利用大数据分析学生的学习轨迹，教师能有针对性地设计教学活动。通过构建这样的系统，能提升学习效果，实现个性化教育，让每个学生都能获得最适合自己的学习资源和指导，推动高等教育教学质量提升。

6.2.重构教育生态：从“人机协同”到“人机共生”

随着人工智能技术发展，需积极探索其在高等教育课程教学中的新应用场景和模式。结合虚拟现实、增强现实等技术，创造全新学习体验，让学生沉浸在逼真的学习情境中，提升学习的趣味性和参与度。研究人工智能在跨学科课程教学、实践教学等领域的应用，拓展其赋能边界。基于对学生全域数据的深度分析，构建精准数字画像，实现精准学情诊断和学业预警，为学生提供个性化学习资源、路径规划和教学指引。借助自然语言处理技术，实现智能交互与辅导，创建虚拟伙伴或助教与学生互动，激发学习兴趣，推动教育生态从“人机协同”迈向“人机共生”，全方位满足学生学习需求。

参考文献

李海峰等. (2023). 人机协同深度探究性教学模式——以基于 ChatGPT 和 QQ 开发的人机协同探究性学习系统为例. 开放教育研究, 2023 ,29 (06)

刘明等. (2024). 生成式人工智能重塑高等教育形态：内容、案例与路径. 电化教育研究, 2024 ,45 (06)

卢滇楠等. (2024). 生成式人工智能赋能高校课程教学：以“化工热力学”课程为例. 清华大学教育研究, 2024 ,45 (05)

王琦等. (2024). 智能时代的深度课堂在内容获取和基于问题的交互方面给师生带来了挑战，. 开放教育研究, 2024 ,30 (04).

徐茂森. (2025). 基于 CiteSpace 的高等教育与人工智能研究的可视化分析——以 2014—2024 年 CNKI 文献为例. 现代信息科技, 2025 ,9 (08)

杨满福等. (2025). 从 Deep-Learning 到 DeepSeek：人工智能赋能大学功能范式重构的挑战、转型与新生态. 现代教育技术, 2025 ,35 (04)

杨现民等. (2025). 人工智能与教育深度融合的场景细化及落地实践——基于探索性多案例分析法. 开放教育研究, 2025 ,31 (01)

张立群. (2024). 人工智能赋能高等教育教学改革的中国范式构建. 中国高等教育, 2024 (24)

张悦周等. (2024). 面向非专业类学生的人工智能课程设计及教学模式. 高等工程教育研究, 2024 (06).

赵玉新等. (2025). “OODA+AI”驱动的自动化领域工程科技未来人才培养——以哈尔滨工程大学智能科学与工程学院为例. 高等工程教育研究, 2025 (01)

高校教师数字化教学胜任力核心要素研究

Research on the Core Elements of Digital Teaching Competency for University Teachers

彭凤琴[1]，赵建华[2*]，冯雪琦[2]，王雷岩[3]
[1]西北师范大学教育技术学院
[2]南方科技大学未来教育研究中心
[3]南方科技大学教学工作部
[*] zhaojh@sustech.edu.cn

【摘要】 教师是推动高等教育数字化转型的重要实践者，数字技术的新发展对高校教师教学提出了新要求。本研究采用德尔菲法，邀请教育技术相关领域的专家、教师，经过三轮专家咨询，最终确定了数字化意识、数字技术知识与技能、数字化教学理论、数字化教学技能、数字伦理与安全、教师专业发展 6 个一级指标和 16 个二级指标的高校教师数字化教学胜任力核心要素框架。研究结果为构建高校教师数字化教学胜任力的评价指标体系提供了理论依据，也为后续相关研究奠定了基础。

【关键词】 高校教师；数字化教学胜任力；德尔菲法；核心要素

Abstract: *Teachers are key practitioners driving the digital transformation of higher education, with new developments in digital technologies imposing updated requirements on college teaching practices. This study employed the Delphi method, inviting experts and educators in educational technology to participate in three rounds of consultations, ultimately establishing a core competency framework for digital teaching capabilities among higher education instructors. The framework comprises six first-level indicators - Digital Awareness, Digital Technology Knowledge & Skills, Digital Pedagogy, Digital Teaching Competence, Digital Ethics & Security, and Professional Development - along with 16 second-level indicators. The research results provide a theoretical foundation for constructing an evaluation index system for digital teaching competence among higher education faculty, while also establishing a basis for subsequent related studies.*

Keywords: university teachers, digital teaching competency, Delphi method, core elements

1.前言

随着生成式人工智能技术的迅猛发展，教育数字化转型已成为全球高等教育改革的核心议题。高等教育数字化转型对教师数字化教学胜任力提出了新要求。如何提升教师的数字化教学能力受到各国教育教学领域的关注。近些年来，世界各国竞相制定了教育数字化转型的发展战略，构建教师的数字胜任力框架。《联合国教科文组织教育信息技术研究所中期战略（2022-2025 年）》主张将教师作为教育发展的第一资源，加强数字时代教师信息通信技术能力建设（UNESCO，2022）。美国、欧盟、挪威、西班牙、联合国教科文组织等国家与国际组织结合数字化时代的需求分别建立了数字胜任力的框架模型。我国教育部在加快高质量发展 2022 年全国教育工作会议上提出了实施教育数字化战略行动（教育部，2022）。教育部于 2022 年 11 月 30 日发布了《教师数字素养》，从数字化意识、数字技术知识与技能、数字化应用、数字社会责任和专业发展等维度提出了教师数字素养框架（教育部，2022）。

国家教育数字化战略行动2025年部署会提出“广大教师要提升数字教学能力”（教育部，2025）。

教学能力是教师能力结构中的核心能力，数字化时代对教师信息化教学能力的要求并不局限于信息技术工具支持的教学实践应用，而是一种更加综合的、全面的、具有竞争力的数字胜任力。已有关于教师信息技术能力标准、ICT胜任力、数字胜任力框架方面国内外研究相对而言较为成熟，但对高校教师数字化教学胜任力及其评价方面的研究较少，缺乏适用于高校教师数字化教学胜任力评估的标准和工具。而构建高校教师数字化教学胜任力的评价指标体系需要先明确高校教师数字化教学胜任力的核心构成要素有哪些。因此，本研究旨在通过德尔菲法，提炼高校教师数字化教学胜任力的核心要素，为高校教师数字化教学胜任力的评价和发展提供理论支持。

2.概念界定

2.1. 胜任力

1973年美国著名的组织行为研究者戴维·麦克利兰（David McClelland）首次提出胜任力（Competency）这一概念（McClelland，1973）。从胜任力的词源进行追溯和分析，competence(s)可翻译为胜任力，与素养（Literacy）、能力（Ability）不同，素养是一种常识性和基础性的知识、技能和道德品行方面的集合，胜任力指向的是相对高级的综合能力，而非基础的、一般的能力。

国内外不同学者对胜任力有不同的定义。美国心理学家斯宾塞提出，胜任力是指能够将某岗位（或组织、文化）上表现优异者与表现平平者区分开的潜在的、深层次的个性特征，它可以是动机、特质、自我形象、态度或价值观、某领域的知识、认知或行为技能中任何可以被可靠测量或计数的，并且能够显著区分工作中优秀绩效和一般绩效的个体特征的总和。国内研究者（邢强、孟卫青，2003）在《未来教师胜任力测评：原理和技术》一文提出：教师胜任力指教师个体所具备的、与实施成功教学有关的一种专业知识、专业技能和专业价值观。研究者将胜任力定义为：组织人员具有的、与一定职位的良好绩效存在显著关联，可以有效观察、测评和改善的知识、技能、动机、品质、价值观等行为特征（肖鸣政，2007）。胜任力还被界定为：组织中绩效卓越成员所具备的可评估与开发的内在和外在要素的集合，要素包括技术能力、知识结构、职业精神、价值观念、性格特征和心理动机（王建民，2012）。综合对胜任力的定义和特征研究，胜任力是一个立体的、多维的动态发展的综合能力，既有外在可测量的行为，也有潜在的特质，从指向的能力来看，区别于一般的、基础的、通用的能力，胜任力指向具体的、情境的、更加综合和具有竞争性的能力。涉及的要素可归结为：知识、能力、动机、品质或特质、态度和价值观，其中知识和能力是显性的外在可测量的行为，动机、品质、态度和价值观则是潜在的特质。

2.2. 教师的数字化教学胜任力

理解教师的数字化教学胜任力内涵和特征离不开对教师数字素养、数字胜任力等相关概念的研究。2006年欧盟发布《关于终身学习的核心能力》首次提出数字胜任力，将其归为八大核心能力之一。2012年欧盟发布的《实践中的数字胜任力：基于框架的分析》中(Ferrari,2012)，明确指出数字胜任力是个体执行任务、解决问题、协作和创造知识时，使用信息技术和数字化媒体所需要的知识、技能和态度(包括能力、战略、价值观和意识)。国内学者将教师数字胜任力定义为教师应拥有发展学生数字胜任力的相应能力，将数字胜任力作为学生的核心素养之一（任友群&杨晓哲，2017）。

本文主要聚焦于高校教师的数字化教学胜任力研究。在中国知网数据库以“胜任力”为主题词检索，高等教育学科领域共有977篇文献，早期关于胜任力的研究对象主要集中于大学辅导员、高校教师、高校行政管理人员，研究内容主要包括胜任力的模型构建、实证分析、评价指标体系构建等。已有研究从三个角度分析了高校教师数字能力或教学胜任力模型。一是从教师专业发展的角度，提出教师应具备的专业素养及通用的数字能力，包括教师的知识、能力、信念等；二是从教学的角度，主要从开展教学的过程构建高校教师教学的胜任力模型；三是从信息技术的角度，提出高校教师所应具备的素养和能力要素。这为构建高校教师的数字化教学胜任力模型提供了多元的视角。

本文的教师数字化教学胜任力是指高校教师在数字化教育环境中，利用数字技术开展教育教学活动而应具备的知识、技能、态度和价值观等综合能力。

3.研究设计

3.1. 研究方法

本研究在国内外文献分析及对教师访谈的基础上，选取了国内外有代表性的政策文件和期刊文章所共同关注的核心要素，初步构建了包括意识、知识、能力、伦理与安全、专业发展等维度的高校教师数字化教学胜任力的指标，每个一级指标下包括若干个二级指标。在此基础上编制德尔菲咨询问卷。

德尔菲法是根据一定的问题邀请相关领域的专家或有经验的管理人员对某一问题进行预测并最终达成一致意见的机构化的方法。专家群体具有专业代表性和权威性。本文经过了三轮的专家咨询，结合上一轮专家所提的意见对咨询问卷进行修改，然后在进行下一轮的专家咨询，通过征求意见-反馈-再集中-再反馈，专家不断修正之前的意见，最终达成集体共识。

3.2. 专家选择

根据德尔菲法的要求，专家或管理人员数量可以根据研究项目的大小和涉及面的宽窄而定，一般在8—20人左右为宜（徐国祥，2005）。本研究共邀请了教育技术领域专家、高校教师和教育管理者共16人参与问卷咨询，专家群体既是教育理论研究者，对本研究主题具有较强的专业理论知识，同时又是教育实践工作者，在教师数字素养、教学能力及教师专业发展等方面具有丰富的实践感知。总体而言，专家群体具有较强的专业性和权威性。

专家基本信息为：性别方面共12名男性、4名女性;年龄分布在34—62岁之间，其中40岁（含）以下5人、40岁—55岁有5人、56—59岁有4人，60岁（含）以上2人；在职称构成方面，教授10人、副教授3人、讲师和研究助理教授共3人；教龄分布在2-40年之间，其中5年及以下教龄者5人、10-20年教龄者3人、22-30年教龄者4人、超过30年教龄者4人。

3.3. 研究步骤

第一轮咨询：编制及发放《高校教师数字化教学胜任力的构成要素专家咨询（第一轮）》问卷，具体包括数字化意识、数字化知识、数字化教学能力、数字伦理与安全、教师专业发展5个一级指标和对应的14个二级指标，每个一级指标对应2-4个二级指标，每个二级指标对应有3个具体描述。邀请专家分别对一级指标和二级指标与高校教师数字化教学胜任力的内容相关性和描述准确性进行评议，1代表“非常不相关”、2代表“不相关”、3代表“一般相关”、 4代表“比较相关”、5代表“非常相关”。如果专家认为指标和描述内容需要修改、增加或删除，可以提出具体修改意见。同时本研究也设计了开放性问题。根据第一轮咨询结果，共计16名专家给出了反馈，本轮咨询专家主要围绕各项指标内容和描述给出了具体的修改意见。

第二轮咨询：根据专家的反馈，对问卷进行了如下修改：（1）将一级指标的“数字化知识”改为“数字技术知识与技能”，将教学知识放到数字化教学能力部分；（2）指标描述增加了能够反映高校教师工作实际特点的描述；（3）在“数字论文与安全”的指标项，增加了数字行为方面的道德规范；（4）调整了部分二级指标的划分，如很多对不同方面应用的意识描述综合调整到意识维度。同时对二级指标项的描述进行了详细的补充修改。
修改后形成《高校教师数字化教学胜任力的构成要素专家咨询（第二轮）》问卷，具体包括数字化意识、数字技术知识与技能、数字化教学能力、数字伦理与安全、教师专业发展 5 个一级指标和 13 个对应的二级指标。根据第二轮反馈的意见来看，专家认为修改后的指标体更加系统化，为了更明确“数字化教学能力”的核心指标含义，专家建议将“数字化教学能力”改为“数字教学技能”和“数字化教学法”，有专家对数字技术知识与技能的表述再次提出了补充修改意见。

第三轮咨询：根据第二轮咨询结果，对专家评分进行统计分析，筛选出一致性较高的指标。考虑到增加了比较重要的一级指标“数字化教学法”（后改为“数字化教学理论”），决定汇总前两轮专家意见，将修改后的问卷版本《高校教师数字化教学胜任力的构成要素专家咨询（第三轮）》再次向专家发放，进一步确认核心要素指标的科学性和合理性。
根据三轮咨询结果，最终确定高校教师数字化教学胜任力的核心要素框架，包括数字化意识、数字技术知识与技能、数字化教学理论、数字化教学技能、数字伦理与安全、教师专业发展共 6 个一级指标和 16 个二级指标，每个二级指标下均有 3 项具体的描述。

4.研究结果

4.1.高校教师数字化教学胜任力的一级指标构成及解释

通过专家咨询对高校教师数字化教学胜任力的构成要素和描述进行了修订和完善，确定了 6 个一级指标，具体指标和解释如下：
（1）数字化意识：指高校教师在数字时代开展教育教学工作应具备的基本意识，包括数字化的认识、态度和价值 3 个二级维度。该指标是数字化转型的认知基石。
（2）数字技术知识与技能：高校教师应了解和掌握的数字技术方面的基本知识与技能。是教师技术层面的基础素养，强调对技术工具和平台的应用掌握。该指标是能力建构的物质基础。
（3）数字化教学理论：高校教师应对数字技术赋能教学应具有的数字化教学法，包括数字化教学理念、数字化教学模式和策略。为技术知识向教学实践的转化提供理论脚手架，是技术转化的中介桥梁。
（4）数字化教学技能：高校教师在数字化教学环境下的教学设计、教学实施、教学评价、反思和创新等方面的能力，形成“开发—实施—评价—反思”的完整教学闭环，是能力输出的实践载体。
（5）数字伦理与安全：高校教师在数字环境中应遵守的伦理道德准则和数字行为规范，以及对网络安全和数据保护的认识和实践能力。是技术应用的规范框架。
（6）教师专业发展：高校教师利用数字技术开展教学研究，促进自身专业成长和学术共同体发展的能力。促进教师从技术使用者向教育创新者转型，是能力迭代的持续动力。
以上 6 个一级指标具有紧密的逻辑关联。体现了“意识-知识-理论-技能-伦理-发展”六维模型，其中前四项（意识➡知识➡理论➡技能）构成从观念启蒙到实践操作的完整闭环，遵循了“意识觉醒➡知识储备➡理论转化➡实践创新”的递进逻辑，将教师数字化发展路径具象化为可观测的能力链条。后两项的数字伦理安全与教师专业发展，既凸显了教育数字化转型

中的风险防控需求，也强化了教师作为变革主体的发展能动性，形成“技术赋能”与“伦理规约”的动态平衡机制，可以作为教师数字化教学胜任力的发展保障层面。整体来说，6项一级指标可以构成“基础能力层（意识/知识）—实践应用层（理论/技能）—发展保障层（伦理/专业）”的三层架构模型。

4.2.高校教师数字化教学胜任力的指标体系

经过专家咨询确定的高校教师数字化教学胜任力的核心要素框架共计6个一级指标和16个二级指标。具体内容如下表1：

表1 高校教师数字化教学胜任力的核心要素框架

一级指标	二级指标
数字化意识	数字化认识
	数字化态度
	数字化价值判断
数字技术知识与技能	数字技术知识
	数字技术技能
数字化教学理论	数字化教学理念
	数字化教学模式
数字化教学技能	数字化教学策略
	数字化教学设计
	数字化教学实施
	数字化教学评价
	数字化教学创新
数字伦理与安全	数字伦理与道德规范
	数字安全与数据保护
教师专业发展	数字化教学研究
	数字化协同发展

5.研究展望

本研究通过三轮专家咨询确定了高校教师数字化教学胜任力的6个一级指标和16个二级指标，包括二级指标的具体描述。通过数据分析，专家的意见较为集中，构建出的高校教师数字化教学胜任力核心要素具有一定的可信度。后续研究将据此编制高校教师数字化教学胜任力调查问卷，面向高校教师开展实证调查研究，结合数据分析，验证核心要素框架的科学性和有效性，并进一步明确各指标权重。在此基础上构建出高校教师数字化教学胜任力的评价指标体系。

参考文献

UNESCO - IITE . UNESCO Institute for Information Technologies in Education Medium-Term Strategy 2022-2025[R]. Moscow: UNESCO-IITE, 2022.

加快教育高质量发展2022年全国教育工作会议召开[EB/OL]. http://www.moe.gov.cn/jyb_zzjg/huodong/202201/t20220117_594937.html

教育部关于发布《教师数字素养》教育行业标准的通知[EB/OL]. http://www.moe.gov.cn/srcsite/A16/s3342/202302/t20230214_1044634.html

国家教育数字化战略行动 2025 年部署会召开[EB/OL]. http://www.moe.gov.cn/jyb_zzjg/huodong/202503/t20250328_1185222.html
邢强,孟卫青(2003).未来教师胜任力测评：原理和技术.开放教育研究(4),39-42.
肖鸣政(2007).人员素质测评[M]. 北京:高等教育出版社.
王建民,杨木春(2012).胜任力研究的历史演进与总体走向.改革(12):138-144.
任友群,杨晓哲(2017).数字化胜任力:信息时代不可或缺的能力.中小学数字化教学(1):22-24.
徐国祥(2005).统计预测和决策［M］.上海:上海财经大学出版社.
Ferrari, A. (2012) Digital Competence in Practice: An Analysis of Frameworks. *Publications Office of the European Union, Spain Press, Luxembourg*, 49-87.
McClelland, D. C. (1973). Testing for competence rather than for" intelligence.". *American psychologist*, *28*(1), 1..

Mechanisms of Collective Knowledge Advancement in Knowledge-Building Communities: A Perspective from Epistemic Emotion

Xueqi Feng[1], Jianhua Zhao[1*], Zheng Jiang[2], Yuqin Yang[3]
[1] Southern University of Science and Technology
[2] Shanghai Normal University
[3] Central China Normal University
[*]zhaojh@sustech.edu.cn

Abstract: *This study investigates how epistemic emotions and discourse moves interact to support collective knowledge advancement in a university-level Knowledge Building environment. Thirty-five undergraduate students participated in a ten-week interdisciplinary course using Knowledge Forum®, where they collaboratively engaged in progressive inquiry and tagged their notes with self-reported epistemic emotions. A mixed-methods approach was employed, combining learning analytics and content analysis to compare high and low contributors, alongside thematic analysis of high-engagement discourse episodes. Quantitative results showed that high contributors expressed significantly more epistemic emotions of challenge, neutrality, and enjoyment, and employed more advanced discourse moves, including sustained inquiry and explanation improvement. Qualitative analysis revealed three affective–discursive mechanisms: (1) emotion-driven inquiry that stimulated conceptual expansion, (2) emotionally positioned theorizing that supported conceptual deepening, and (3) emotion-regulated synthesis that enabled collaborative integration of ideas. These findings suggest that epistemic emotions are not peripheral to learning but act as cognitive resources that shape the development of progressive discourse.*

Keywords: knowledge building, collective knowledge advancement, discourse move, epistemic emotion

1. Introduction

The transformation of education to support authentic knowledge creation has become a central concern in the learning sciences (Hoidn & Kärkkäinen, 2014). Among the frameworks developed in response to this challenge, Knowledge Building offers a distinctive approach in which classrooms are conceptualized as knowledge-creating communities. In these communities, learners collaboratively engage in the *improvement of ideas*, with the central goal of advancing community knowledge (Scardamalia & Bereiter, 2014). Rather than prioritizing individual acquisition or comprehension, Knowledge Building emphasizes epistemic agency, collective responsibility, and the co-construction of conceptual artifacts that are improvable and valuable to the wider community. To support such inquiry, Knowledge Forum® provides a digital environment where students can pose problems, articulate emerging theories, and refine ideas through sustained discourse. Scaffolds such as "my theory," "a better theory," and "putting our ideas together" are embedded within the platform to promote metacognitive reflection and progressive discourse. Prior research has examined how students engage in discourse practices—such as explanation, elaboration, and theory building—that contribute to collective knowledge advancement (Chen & Hong, 2016). Building on

this foundation, our earlier work designed a Knowledge Building environment in which undergraduate students engaged in self-directed learning trajectories and reflected on their participation in community knowledge processes. Through iterative design-based research, we found that students developed deeper epistemic awareness, adopted increasingly sophisticated discourse strategies, and made more substantial contributions to collective knowledge advancement over time (Feng, et al., 2023).

While discourse-centered analyses have yielded critical insights into the cognitive and collaborative dimensions of Knowledge Building, relatively limited attention has been devoted to the affective experiences that accompany and potentially shape students' engagement in knowledge-building processes. *Epistemic emotions* — such as curiosity, confusion, interest, and frustration— are known to arise in contexts involving uncertainty, complexity, or conceptual conflict. These emotions are not merely by-products of learning but are increasingly recognized as playing a constitutive role in students' engagement, persistence, and responsiveness during inquiry. Emerging studies have begun to explore how epistemic emotions manifest in Knowledge Building environments. For example, Zhu et al. (2021) found that refining Knowledge Building practices based on student feedback contributed to more positive emotional engagement. Teo et al. (2022) employed multimodal data to show that epistemic emotions vary dynamically during collaborative idea development. These findings suggest that emotional experiences are deeply embedded in knowledge-building participation and may influence how learners interact with ideas and with one another.

Despite these developments, few studies have systematically examined how epistemic emotions and discourse moves are enacted together in the process of collective knowledge advancement. Moreover, it remains unclear whether and how emotional and discursive engagement differ among students who contribute substantively to the improvement of community knowledge. Addressing this gap, the present study investigates how undergraduate students' epistemic emotions and discourse practices relate to their contributions to collective knowledge advancement in a university-level Knowledge Building course. The course emphasized self-directed inquiry and sustained engagement with community ideas. Students tagged their written contributions with self-reported epistemic emotions, enabling a detailed analysis of the co-occurrence between emotion and discourse. We focus on how these dimensions vary among students with differing levels of contribution to community knowledge advancement, and how they are enacted in the context of progressive inquiry. This study is guided by the following research questions:

1. To what extent do high and low contributors differ in their use of epistemic emotions and discourse moves associated with collective knowledge advancement?
2. How did undergraduates engage in epistemic emotions and discourse moves towards collective knowledge advancement?

By examining the emotional and discursive dynamics that underpin collective knowledge advancement in Knowledge Building, this study seeks to advance a more integrated understanding of how affective and cognitive practices jointly support knowledge-creating participation in collaborative learning environments.

2. Methods

2.1. Participants

Thirty-five undergraduate students from a science and engineering university in Shenzhen, China, participated in this study. All participants were enrolled in a liberal arts elective course titled *Education and Modern Technology*, which was offered to students from diverse science and engineering majors. The course aimed to foster interdisciplinary inquiry into the relationship between education, society, and technological development. Participation in the study was voluntary, and all students provided informed consent.

2.2. Knowledge Building context and pedagogical design

This ten-week course, *Education and Modern Technology*, was designed as an interdisciplinary elective for undergraduate science and engineering students. It was grounded in the principles of Knowledge Building, which conceptualizes education as a process of collectively advancing ideas within a knowledge-creating community (Scardamalia & Bereiter, 2014). The pedagogical design emphasized sustained inquiry, epistemic agency, and progressive discourse. All collaboration occurred within Knowledge Forum®, a computer-supported collaborative learning platform enabling students to post notes, build on each other's ideas, and visualize discourse. To support metacognitive and affective engagement, students were encouraged—but not required—to use two types of scaffolds when composing notes: (1) *Epistemic discourse scaffolds*, including prompts such as "My idea", "A better idea", and "Putting our ideas together", aimed to guide sustained inquiry. (2) *Epistemic emotion scaffolds*, including "confused", "challenged", "surprised", "curious", "neutral", "frustrated", and "happy", supported students' reflection on affective experiences during inquiry. The learning process unfolded through three emergent phases of collective inquiry, each representing a shift in epistemic goals and discourse engagement.

2.2.1. Phase 1: Ideation (Weeks 1–3)

Students initiated inquiry by engaging with foundational prompts such as Our Vision for Education and The History of Education and Technology. Through collaborative exploration, they identified key issues related to educational purpose, modern vs. traditional technologies, learning theories, social relations, and team collaboration. Epistemic discourse scaffolds were frequently used to articulate, refine, and connect ideas, while emotion tags enabled reflection on cognitive challenges and curiosity.

As illustrated in Figure 1, one group centered their inquiry on the core problem: How can educational equity be effectively achieved amid educational transformation? Their dialogue expanded into subtopics such as education and social change, education transformation and policy, social development, and talent needs. Through iterative discourse, the group collaboratively co-created a problem-solving poster, synthesizing their key insights and proposing actionable directions for change.

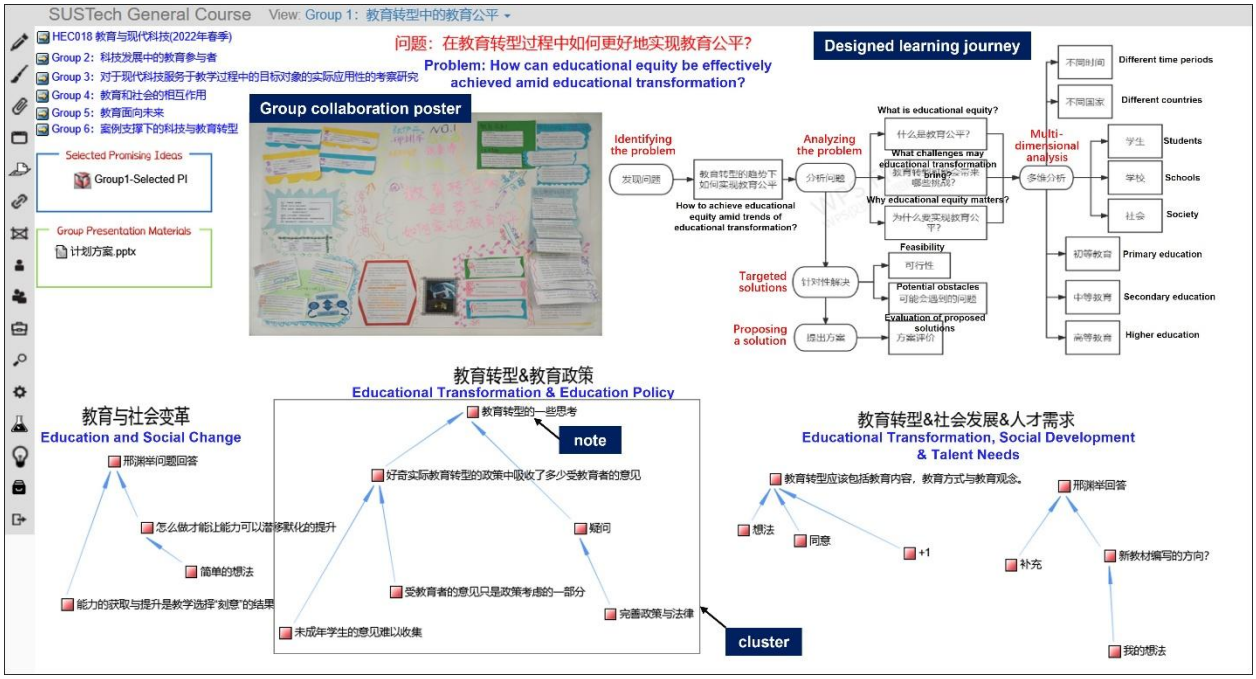


Fig. 1 A Knowledge Forum view

2.2.2. Phase 2: Learning journey design (Weeks 4–6)

Building on their initial exploration, students engaged in designing visual models of their inquiry trajectories — referred to as learning journeys — to reflect how their understanding evolved over time. These models were shared in Knowledge Forum for peer feedback, allowing the learning community to collectively reflect on the processes of theory refinement and idea improvement.

As shown in Figure 2, different groups emphasized distinct aspects of inquiry: one model highlighted the importance of collaboration and reflection, while another underscored iteration and feedback loops. Despite stylistic variation, all models demonstrated a shared epistemic commitment to progressive idea improvement, aligning with core Knowledge Building principles.

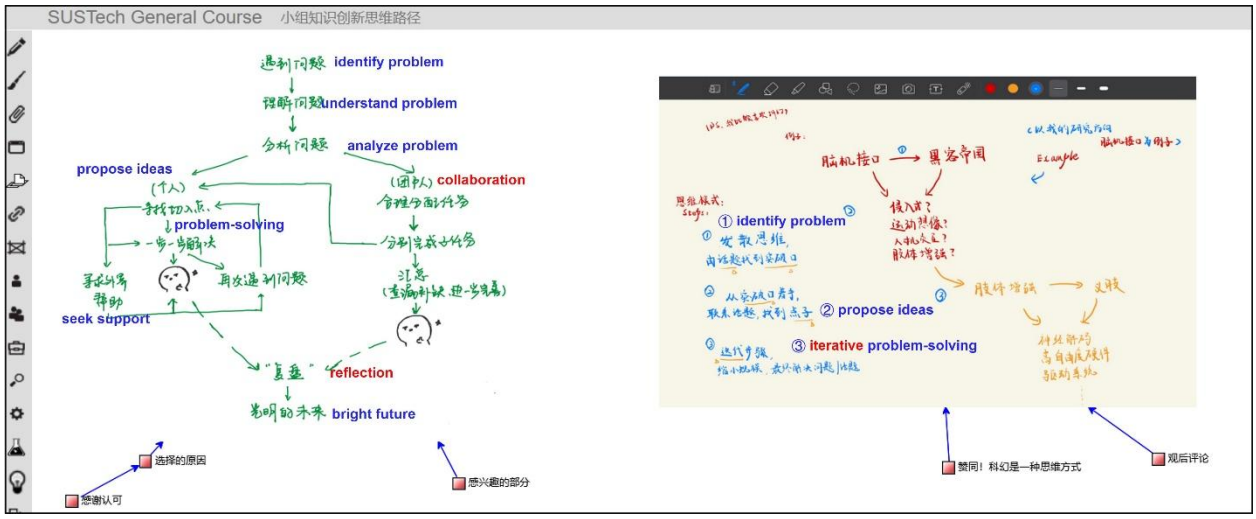


Fig.2 Student designed learning journey for Knowledge Building

2.2.3. Phase 3: Progressive idea improvement with the designed learning journey (Weeks 7–10)

In the final phase, students extended their inquiry into broader conceptual domains, drawing from historical, philosophical, and systemic perspectives on education and technology. Each of the six groups selected a thematic focus—such as educational equity, technological transformation, or

education and social development—and pursued sustained investigation through collective discourse.

Revisiting Figure 1, the group exploring educational equity deepened their inquiry into the subdomains of policy innovation, social inclusion, and talent development. Their earlier learning journey served as a conceptual scaffold, helping them align discourse with thematic goals and maintain coherence across inquiry episodes. Each group synthesized their findings into a collaborative poster, representing both a knowledge artifact and a shared contribution to community understanding.

2.3. Data Sources and data analysis

The primary dataset consisted of 328 notes authored by 35 students over ten weeks of collaborative discourse in Knowledge Forum®. The analysis proceeded in three stages, combining quantitative comparisons and qualitative interpretation to examine students' emotional and discourse engagement in relation to their contribution to collective knowledge advancement.

2.3.1. Identifying contribution levels

The learning analytics tool KBDeX (Oshima et al., 2012) was utilized to distinguish undergraduates with high and low contributions to collective knowledge advancement. According to knowledge-building research on KBDeX, a higher Total Degree Centrality denotes more collective knowledge advancement (Oshima et al., 2021). Here, *Note Contribution* is defined as the contribution of a note to collective knowledge advancement and calculated as the difference in Total Degree Centrality between the current and the previous Knowledge Forum discourse (note) network. Then each note's author was identified to calculate each student's contribution to the collective knowledge advancement (Feng et al., 2020). Finally, 17 students were identified as high contributors, while 18 were identified as low contributors.

2.3.2. Emotion and discourse coding with statistical analysis

The study compared undergraduates' engagements of epistemic emotions and discourse moves between high and low contributors using content analysis. Firstly, the numbers and percentages of students' epistemic emotions of *confusion*, *challenged*, *surprise*, *curiosity*, *neutral*, *frustration*, and *enjoyment* (Han et al., 2021; Teo et al., 2022; Zhu et al., 2022) from each Knowledge Forum note were calculated (students labeled their epistemic emotions in each Knowledge Forum note when creating it). Then students' Knowledge Forum notes were coded with a coding framework composed of several categories, including *Questioning*, *Theorizing*, and *Community*. These categories draw upon theoretical frameworks for progressive inquiry (Hakkarainen, 2003) and knowledge creation (Chuy et al., 2011). Coding for each note was done independently by two raters, who then came together for social moderation to reach an agreement (Cohen's kappa= .79). Disagreements were settled after discussion. Finally, the epistemic emotions and discourse moves adopted by high and low contributors were compared using one-way MANOVA.

2.3.3. Thematic analysis of high-engagement episodes

To complement the quantitative findings, a qualitative thematic analysis was conducted on selected notes that exemplified students' integration of epistemic emotions and discourse practices in the process of idea improvement. This analysis sought to uncover how emotional and cognitive engagement co-occurred in episodes of productive discourse, thereby enriching our understanding of the mechanisms supporting community knowledge advancement.

3. Results

3.1. RQ1: To what extent do high and low contributors differ in their use of epistemic emotions and discourse moves associated with collective knowledge advancement?

3.1.1. Differences in epistemic emotions

Table 1 presents the frequencies and proportions of epistemic emotions expressed by high and low contributors. High contributors reported higher levels of *challenged* (17.48% vs. 15.57%), *neutral* (35.44% vs. 30.33%), and *enjoyment* (25.73% vs. 24.59%) emotions, as well as slightly more instances of *frustration* (2.43% vs. 0.82%). In contrast, low contributors expressed more *confusion* (13.11% vs. 7.28%) and *surprised* (4.10% vs. 1.46%) emotions.

A one-way MANOVA revealed a significant overall difference in the use of epistemic emotions between the two groups, $F(7, 27) = 3.90$, $p = .005$, Wilks' $\Lambda = .50$, partial $eta^2 = .50$. Follow-up univariate ANOVAs indicated that high contributors expressed significantly more *challenged* ($F(1, 33) = 4.60$, $p = .039$, partial $eta^2 = .12$), *neutral* ($F(1, 33) = 15.98$, $p < .001$, partial $eta^2 = .33$), and *enjoyment* ($F(1, 33) = 4.54$, $p = .041$, partial $eta^2 = .12$) emotions. These findings suggest that students who expressed greater levels of emotional engagement – particularly emotions such as *challenged*, *neutral*, and *enjoyment* – were more likely to contribute meaningfully to the community's collective knowledge advancement.

Table 1. *Frequencies and percentages of epistemic emotions conducted by high and low contributors*

	High contributors		Low contributors	
	f	% (f/206)	f	% (f/122)
Confusion	15	7.28	16	13.11
Challenged	36	17.48	19	15.57
Surprised	3	1.46	5	4.10
Curiosity	21	10.19	14	11.48
Neutral	73	35.44	37	30.33
Frustration	5	2.43	1	0.82
Enjoyment	53	25.73	30	24.59

3.1.2. Differences in discourse moves

As shown in Table 2, the high contributors used more higher-level discourse moves than the lower contributors. For example, compared with low contributors, they engaged in more *sustained inquiry* (12.14% vs. 6.56%), *improving an explanation* (17.48% vs. 12.30%), and *lending support* (4.85% vs. 4.10%). One-way MANOVA results reveal a significant difference in students' usage of discourse moves between the high and low contributors, $F(8, 26) = 3.53$, $p = .007$, Wilks' $\Lambda = .48$, partial $eta^2 = .52$. In follow-up univariate ANOVAs, higher contributors show significantly higher values of *Questioning-sustained inquiry* ($F(1, 33) = 7.12$, $p = .012$, partial $eta^2 = .18$), *Theorizing-supporting an explanation* ($F(1, 33) = 8.55$, $p = .006$, partial $eta^2 = .21$), and *Theorizing-improving an explanation* ($F(1, 33) = 9.47$, $p = .004$, partial $eta^2 = .22$). These results indicate that student who contributed more substantively to collective knowledge advancement tended to engage in deeper *Questioning* and more sophisticated *Theorizing*, supporting the emergence of progressive discourse.

Table 2. *Frequencies and percentages of discourse moves used by high and low contributors*

	High contributors	Low contributors

	f	% (f/206)	f	% (f/122)
Questioning				
Explanation seeking (ES)	12	5.83	8	6.56
Sustained inquiry (SI)	25	12.14	8	6.56
Theorizing				
Proposing an explanation (PE)	69	33.50	54	44.26
Supporting an explanation (SE)	40	19.42	19	15.57
Improving an explanation (IE)	36	17.48	15	12.30
Community				
Synthesis (SS)	2	0.97	1	0.82
Shared regulation (SR)	12	5.83	13	9.84
Lending support (LS)	10	4.85	5	4.10

3.2. RQ2: How did undergraduates engage in epistemic emotions and discourse moves towards collective knowledge advancement?

To investigate how undergraduates engaged in epistemic emotions and discourse moves to support collective knowledge advancement, we conducted a thematic analysis of discourse episodes contributed by high-contributing students. The analysis revealed three interrelated mechanisms through which epistemic emotions and discourse practices were dynamically coupled to foster progressive discourse. Each theme corresponds to a distinct phase of knowledge-building: conceptual expansion, conceptual deepening, and collaborative integration.

3.2.1. Emotion-driven inquiry as a trigger for conceptual expansion

Epistemic emotions such as *curiosity*, *challenged*, and *frustration* served as affective triggers that prompted students to generate open-ended questions and initiate sustained inquiry. These emotionally charged questions helped expand the problem space and led the group into new conceptual territories.

For instance, when discussing *Educational Transformation*, student s31 expressed curiosity about the extent to which learners' ideas are incorporated into policymaking:

[I feel curious] How much weight should the ideas of the educated take in the policy formulation of educational transformation? … I am curious about how many ideas of educated people were incorporated when formulating policies for educational transformation.

This inquiry prompted peers to consider the tension between idealistic student voice and practical policy constraints. Similarly, in a discussion on The Impact of Technology on Education, student s29 expressed challenge in navigating the information landscape:

[I feel challenged] The threshold for obtaining knowledge is getting lower… How to extract high-quality information from massive amounts of information is a new challenge for technology and individuals.

Following this, student s06 posed a critical question about whether lowered educational thresholds could lead to a decline in teaching quality:

[I feel challenged] Would lowering the educational threshold lead to a decline in the average level of teaching staff?

In another discussion, student s25 expressed frustration about the difficulty of self-understanding and posed a new question about how education can help individuals identify their interests and talents:

[I feel frustrated] I often have a feeling that I don't necessarily know myself… I think education can provide a more effective way for me to discover my own interests and talents?

These emotionally driven questions illustrate how epistemic emotions functioned as catalysts for conceptual expansion, opening up new avenues for inquiry and community engagement.

3.2.2. Emotionally positioned theorizing for conceptual deepening

Students also used epistemic emotions to frame and position their theoretical stances, often in the form of neutral, challenged, or frustration. These emotions accompanied discourse moves such as improving explanations and elaborating ideas, signaling students' efforts to refine, evaluate, or reframe existing contributions.

For example, in the discussion on Educational Transformation, student s01 neutrally elaborated on vocational education and labor market dynamics:

[I feel neutral] Professionals should integrate the social situation with the school's educational achievements… For example, studying for a postgraduate to find a job, but not every position requires the research ability of a postgraduate.

This positioned explanation built on earlier ideas and led peers like s26 to further integrate perspectives on individual development and societal needs. In another discussion on What is Education, student s33 expressed frustration regarding current assessment systems:

[I feel frustrated] The assessment criteria of the education system are sometimes not diverse enough… The most important thing is to find the one that suits you.

This emotionally positioned challenge prompted further elaboration from peers on the need for diverse evaluation methods and personalized learning.

In each case, emotions helped students frame their epistemic stance, providing both a cognitive and social signal that invited deeper engagement and theory refinement. These episodes reflect a process of conceptual deepening, where emotional cues guided the community toward more elaborated and coherent explanations.

3.2.3. Emotion-regulated synthesis for collaborative integration

Students used emotionally regulated discourse—particularly involving neutral or challenged expressions—to synthesize diverse ideas and set new directions for group inquiry. These emotionally modulated contributions enabled students to integrate prior ideas and structure collective understanding, supporting collaborative integration.

For instance, in the Educational Transformation discussion, student s14 challenged the group to adopt a future-oriented perspective:

[I feel challenged] Education transformation should be future-oriented—to the values of the people in the future, the technology of the future, and the society of the future… We should pay more attention to future development and plan for the transformation of education.

This forward-looking synthesis led to a series of elaborations from peers on curriculum reform, educational equity, and social change. Similarly, student s29 neutrally synthesized multiple group perspectives on the relationship between education, society, and individual freedom:

[I feel neutral] Education should focus on the relationship between individuals and society… However, this kind of freedom is still far from being realized… Such education is a waste of human resources and something we must pay attention to.

This synthesis note served as a conceptual anchor, prompting further inquiry into structural constraints and the potential for educational reform.

These examples illustrate how emotion-regulated synthesis enabled students to manage cognitive complexity and coordinate perspectives, facilitating the integration of ideas into a shared knowledge structure.

4. Discussion and Conclusion

This study demonstrates that epistemic emotions play a functional role in collective knowledge advancement within a Knowledge Building environment. High-contributing students expressed more challenge, neutrality, and enjoyment, and engaged in advanced discourse moves such as sustained inquiry and explanation improvement. Qualitative analysis revealed three mechanisms through which emotions and discourse interacted: emotion-driven inquiry opened new conceptual spaces; emotionally positioned theorizing refined ideas; and emotion-regulated synthesis integrated group perspectives. These findings suggest that emotions are not peripheral to collaborative learning but serve as cognitive resources that scaffold different phases of progressive discourse. By illuminating how affective–discursive dynamics support conceptual expansion, deepening, and integration, this study advances a more holistic understanding of knowledge-building participation. Future work should explore how emotional engagement can be further supported through pedagogical or technological scaffolds in diverse learning contexts.

Acknowledgements

This study is supported by the National Social Science Foundation of China (Grant No. BCA200090), and Guangdong Planning Office of Philosophy and Social Science (Grant No. GD23CJY19).

References

Chen, B., & Hong, H. Y. (2016). Schools as knowledge-building organizations: Thirty years of design research. *Educational psychologist*, *51*(2), 266-288.

Feng, X., Zhao, J., Yang, Y., & Chan, C. K. (2023). Examining Undergraduates' Epistemic Emotions and Discourse Moves That Support Collective Knowledge Advancement in Knowledge Building. In *Proceedings of the 17th International Conference of the Learning Sciences-ICLS 2023, pp. 938-941*.

Hoidn, S., & Kärkkäinen, K. (2014). *Promoting skills for innovation in higher education: A literature review on the effectiveness of problem-based learning and of teaching behaviours*. OECD Publishing.

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis *application* for knowledge building discourse. *Educational Technology Research and Development*, *60*(5), 903-921.

Oshima, J., & Shaffer, D. W. (2021). Learning Analytics for a New Epistemological Perspective of Learning. *Information and Technology in Education and Learning*, *1*(1), Inv-p003.

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.

Teo, C. L., Ong, A., & Lee, V. Y. A. (2022). Exploring students' epistemic emotions in knowledge building using Multimodal Data. In A. Weinberger, W. Chen, D. Hernández-Leo, & B. Chen (Eds.), *International Collaboration toward Educational Innovation for All: Overarching Research, Development, and Practices, 15th International Conference on Computer-Supported Collaborative Learning (CSCL) 2022* (pp. 266-273).

Zhu, G., Raman, P., Xing, W., & Slotta, J. (2021). Curriculum design for social, cognitive and emotional engagement in knowledge building. *International Journal of Educational Technology in Higher Education*, *18*(1), 1-19.

教师成为学习设计师的实践探索：专家指导与 AI 协同下 K-12 STEM 课程的五轮迭代研究

Becoming a Learning Designer: Expert Guidance and AI Collaboration in a Five-Iteration Case Study of K-12 STEM

刘勇[1]，刘锟[2*]，潘鹏飞[3]，陈俊[1]，罗桂林
[1]重庆八中宏帆中学校
[2]香港大学
[3]北京大学
[4]重庆市渝北区中央公园小学校
[*] liukun2014@gmail.com

【摘要】 在 K-12 STEM 教育中，教师面临如何在大班教学中实现高质量探究式学习的挑战。本文采用基于设计的研究方法（Design-based research），基于一位中学 STEM 教师在五轮课程迭代过程中的实践，探讨教师在专家指导与生成式 AI 协同支持下的角色转型机制。研究提出“教师—专家—AI—学生学习证据”协同成长模式，揭示专家指导有助于教师的理念转变的重构对学习设计的理解教学，生成式 AI 可在明确任务下提供高效支持，而学生学习证据反馈机制是推动教师持续优化设计的关键中间变量。该框架为教师从程序性教学向以学生为中心的探究式学习设计转型提供了一种结构化路径，对教师专业成长机制的建构具有实践与理论意义。
【关键词】 学习设计；生成式人工智能；学习证据；教师成长；STEM 教育

Abstract: *This study investigates how expert guidance and generative AI collectively support STEM teachers' transition from traditional instruction to inquiry-based learning design. Through five iterative cycles of design-based research with a secondary teacher, we identify three key mechanisms: (1) experts catalyze pedagogical transformation by challenging and reconstructing teachers' design assumptions, (2) AI tools effectively scaffold specific design tasks when provided with clear parameters, and (3) systematic analysis of student learning evidence creates feedback loops that drive continuous improvement. These interdependent mechanisms form a dynamic system that enables teachers to develop increasingly sophisticated approaches to inquiry-oriented design while maintaining practicality in large-class contexts. The findings offer evidence-based strategies for supporting teacher development in STEM education through synergistic human and technological supports.*
Keywords: learning design; generative AI; learning evidence; teacher professional growth; STEM education

1.引言

在当前推动科学教育高质量发展的背景下，K-12 教师被要求设计并实施具备真实性、思维挑战与学科融合的探究式课程。然而，在实际教学中，由于班级规模大、教学资源有限，教师常常依赖于预设的程序性活动，缺乏深度探究与真实问题解决的元素。如何有效支持教师转变教学观念，提升其课程设计能力，成为亟待解决的问题。学习设计（Learning Design）

作为一个新兴领域，强调通过设计框架和工具支持教师结构化地开展教学设计，提高其教学实践的质量与效率（Mor, Craft & Maina, 2015）。与此同时，生成式人工智能（Generative AI）在教育领域的兴起，为教师提供了实时反馈与辅助设计的新可能（Ogunleye et al., 2024）。本文关注的问题是：在专家和生成式 AI 的支持下，教师如何完成从程序性教学到学习设计师的转型？我们是否可以构建一个可推广的教师成长模型，以系统性支持更多教师向“学习设计师”转变？已有研究表明，教师的课程设计能力并非自发生成，而需要在结构化引导下逐步建构（Laurillard et al., 2013）。学习设计三角框架（LDT）（Law & Liang, 2020）被选为本研究的指导框架，因为它与以学生为中心、基于探究的学习设计的核心目标高度契合，并有效解决了教师学习设计实践中的两个主要挑战。LDT 不仅提供了概念上的清晰性，还提供了程序性的支撑，帮助教师通过明确连接学习成果、学科实践和教学方法，设计出连贯的、以探究为导向的课程。LDT 框架基于以结果为导向的方法和 Goodyear（2005）提出的教学设计模型，强调通过将教学意图与设计行动对接，解决设计过程中的矛盾。框架聚焦于三个核心要素：学习成果、学科实践和教学法，它们为课程组件的设计提供了支撑，课程组件通过任务设计帮助学生实现学习成果。为了支持实际应用，LDT 还包含了一个七步设计过程，帮助教师定义学习成果和教学策略、设计与学习成果对齐的课程组件、优化教学策略并监控学习进展。通过这一过程，教师能够发现并解决设计中的常见问题，促进教学设计的连贯性与迭代完善（Law & Liang, 2020）。但专家无法实时提供反馈，而 AI 可以提供更及时的反馈，帮助教师优化活动流程、文本表达，并模拟学生视角审视课程设计。然而，AI 的有效性依赖于教师的提问能力和任务表述的准确性，缺乏人类专家的判断力与整体感知能力。教师若无法清晰表达，AI 的反馈可能会受到限制（Sperling al., 2024）。在此基础上，一些研究提出将学生学习证据纳入教学反馈机制，支持教师对教学设计进行持续优化。然而，如何整合专家支持、AI 协同与学生反馈，形成一个动态互动的成长机制，仍缺乏清晰模型。

2.研究方法

本文采用设计型研究（DBR）方法，对一位中学 STEM 教师的五轮课程迭代过程进行深度追踪。数据包括：课程教案版本、专家反馈访谈、ChatGPT 交互记录、学生作品与过程证据、教师反思笔记。分析聚焦教师在专家—AI—学生三方协同支持下的转型路径，并构建 TEAS 模型。

3.研究结果

本文基于五轮课程迭代实证数据提出“教师—专家—AI—学生学习证据”协同成长模型。

3.1. 模型构建的实证支撑

3.1.1. 教师的迭代设计与角色转型

教师在五轮课程迭代中逐步从程序性教学向探究导向教学转型，表现在教学目标、流程、证据收集与反思内容的不断演进。

轮次	教学目标	学习流程	学习证据	教师反思
1.0	以内容为核心（动手实践）	思考问题 → 观看示范 → 小组制作 → 分享展示	实践作品	程序性教学
2.0	内容+探究要素并存	提出假设 → 实践 → 讨论验证	实践+过程片段记录	明确课程定位

轮次	教学目标	学习流程	学习证据	教师反思
3.0/4.0	知识+探究路径系统整合	问题生成 → 控制变量 → 设计实验 → 得出结论	探究日志+作品	科学化教学流程
5.0	融入科学思维与价值导向	自主提出问题 → 完整实验设计 → 多轮修正	探究报告+自评互评	完善的探究体系

3.1.2. 专家支持机制

专家反馈主要聚焦在“目标澄清—任务定位—证据意识”三个层面，通过教学诊断与反思引导，促发教师在结构层面对课程进行重构。比如：“专家提醒我要‘拆分教学目标，并按照科学探究流程来设计学习活动’，这让我意识到任务设计还不够清晰。”

3.1.3. AI 协同特征的演进

在专家引导下，教师逐渐学会提出结构化问题，AI 反馈从泛化建议逐步过渡为流程生成、评价设计等专业支持。

阶段	教师提问	AI 反馈类型	示例
初期	“你觉得这课怎么样？”	鼓励性反馈	“这个教案设计非常有趣且富有探索性。”
中期	“我希望引导学生控制变量，有建议吗？”	结构化建议（引导语/流程图）	“可以设置不同的任务难度层次，帮助学生理解变量控制。”
后期	“请优化我的教学步骤并补充评价方案。”	完整任务序列+评价表设计	“我将基于探究步骤优化流程，并附上每阶段的自评互评标准……”

3.1.4. 学生学习证据的反馈作用

学生学习表现为教师调整教学设计提供了重要反馈信号，是课程持续优化的中介变量。

轮次	证据表现	教师调整点
1.0	模仿性完成任务	增加开放问题，引导学生提问
3.0	假设不清、变量不控	引入案例教学，强化实验结构
5.0	能进行互评与模型修正	鼓励学生自建流程、自主追问

3.2. 专家—AI—教师—学生的三类协同路径机制

在模型运行过程中，四类要素之间呈现出三种典型的动态交互路径，构成“诊断—设计—验证—再设计”的螺旋式协同机制。

3.2.1. 专家驱动路径

专家 → 教师 → AI → 学生：专家提出教学目标重构建议，教师基于此向 AI 寻求支持，AI 提供方案优化教学流程，学生学习产出反馈新的问题。

实证示例：在 1.0 版本课程实施后，专家通过分析教师反思指出，课程可定位为“科学探究 + 工程实现”。教师据此向 AI 提问：“如何让我的设计更加具有探究性？”AI 建议采用问题驱动的方式构建课程结构。由此，2.0 版本的学习路径围绕两个核心问题展开，显著增强了探究导向。学生在任务完成过程中表现出良好思维状态，并提出了如“为什么增加重量后系统会平衡？”等具有深度的问题。教师首次在课堂中面临高质量学生提问，一时间未能及时回应，这一情境反向推动教师重新思考师生互动与教学应答策略。

3.2.2. 学生证据驱动路径

学生 → 教师 → AI → 专家：学生产出暴露教学短板，教师反思后利用 AI 初步调整，专家再提供精准化建议，形成基于证据的反向优化机制。

实证示例：在第 2.0 版本实施过程中，学生提出关于系统平衡的问题，教师一时难以作答，转而借助 AI 检索资料补充相关知识，由此意识到自身在 STEAM 课程中存在知识短板。经过与专家深入交流，专家指出症结在于教学目标的拆解不够细致，导致知识结构松散，教师难以建立本节课的知识网络图。随后教师据此重新梳理知识结构，并结合 AI 支持完善教学设计，从而提升了应对复杂项目式学习情境的专业胜任力。

3.2.3. AI 驱动路径

AI → 教师 → 专家 → 学生：教师预先调用 AI 生成教学方案，筛选后实施教学，专家结合学生学习证据进行评估，反馈结果反哺设计。

实证示例：AI 在第 4 轮之后建议教师“可以加入如何通过实验数据进行定量分析的要求”，教师根据实际情况，拒绝了 AI 的建议。同时 AI 提出在自评和互评阶段“可以加入具体的评估标准，指导学生如何评价他人作品”，教师采纳了 AI 建议，并和专家沟通后，在 5.0 版本的中，加入自评和互评表格。

4.讨论与结论

本研究通过分析五轮迭代过程，研究发现，该转型依赖于专家引导的认知重构、AI 辅助的精细化设计以及基于学生学习证据的迭代优化三大关键要素。专家支持在教师转型中发挥了结构性作用。与既有研究中学习设计框架的应用类似（Mor & Craft, 2012），本研究发现专家通过帮助教师识别核心问题、诊断系统性设计缺陷，并引导结构化反思，促进了教师从“活动执行者”到“学习设计师”的认知转变。协助教师澄清学习目标，还推动其建立更严谨的设计逻辑，这与学习科学领域强调的“教师作为设计思维实践者”的定位一致（Kali et al., 2015）。生成式 AI 的辅助功能呈现显著的“任务依赖性”。尽管 AI（如 GPT）能够快速生成活动设计细节或优化特定环节（如评价标准），但其有效性高度依赖于教师输入的精确性。当教师提出泛化问题时，AI 仅能提供微调建议（如调整措辞）；而当教师明确设计需求时（如“设计一个聚焦数据解释能力的评价量表”），AI 的生成质量显著提升。这一发现呼应了近期关于“教师需培养 AI 提示工程能力”的讨论（Luckin et al., 2022），也揭示了当前 AI 工具的局限性——它更擅长增量改进而非颠覆性创新。生成式 AI 的有效性则依赖于教师输入的精准度，能够根据明确需求优化设计，但在泛化问题上效果有限。学生学习证据的作用则需教师具备数据素养，能够通过分析转化为设计洞察。实践启示包括：专家-AI 协同模式可聚焦宏观设计与实施分工，教师需加强 AI 提示工程和学习分析培训，学校应建立“设计-实施-评估”的迭代制度。未来研究可进一步探讨不同教师经验水平与 AI 工具适应性。未来可拓展多样教师样本，检验模型在不同阶段、学科的适应性，并进一步开发校本 AI 支持工具，助力教师持续成长。

参考文献

Laurillard, D. (2013). *Teaching as a design science: Building pedagogical patterns for learning and technology.* 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.

Maina, M., Craft, B., & Mor, Y. (Eds.). (2015). *The art & science of learning design*. Springer.

Sperling, K., Stenberg, C. J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open,* 100169.

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of generative AI for teaching and learning practice. *Education Sciences, 14(*6), 636.

元宇宙视域下中国传统文化传播形态与教育应用研究

Research on the Communication Forms of Chinese Traditional Culture and Educational Applications from the Perspective of the Metaverse

全文慧[1] 刘勇[1*]
[1]北京邮电大学人文学院
* liuyong001@bupt.edu.cn

【摘要】 元宇宙的兴起为中国传统文化的传播及其教育应用带来了新的发展机遇和挑战。文章首先介绍了元宇宙及其传播特征以及中国传统文化的相关概念，然后从传播媒体形态、传播主体、传播方式、传播环境四个维度探讨了元宇宙视域下中国传统文化的传播形态，继而探讨了其在教育中的具体应用，最后提出了元宇宙视域下中国传统文化通过教育途径传播面临的挑战和未来发展建议，以期为元宇宙视域下在教育教学过程中传播中国传统文化提供参考与启示，推动中国传统文化在全球范围内的传播与发展。
【关键词】 元宇宙；中国传统文化；文化传播；传播形态；教育

Abstract: *The rise of the Metaverse has brought new development opportunities and challenges to the dissemination and educational application of traditional Chinese culture. The article first introduces the Metaverse and its dissemination characteristics, as well as the relevant concepts of traditional Chinese culture. Then, it explores the dissemination form of traditional Chinese culture from the perspective of the Metaverse from four dimensions: media form, dissemination subject, dissemination method, and dissemination environm ent. Then, it explores its specific applications in education. Finally, it proposes the challenges and future development suggestions for the dissemination of traditional Chinese culture through educational channels from the perspective of the Metaverse, To provide reference and inspiration for the dissemination of traditional Chinese culture in the teaching process from the perspective of the Metaverse, and to promote the dissemination and development of traditional Chinese culture on a global scale.*
Keywords: Metaverse; Chinese traditional culture; cultural dissemination; communication forms; education

1.引言

2025 年中共中央、国务院印发的《教育强国建设规划纲要（2024—2035 年）》中提到坚定文化自信，加强社会主义先进文化、革命文化和中华优秀传统文化教育（新华社，2025）。党的二十大报告提出“加快发展数字经济，促进数字经济和实体经济深度融合”。2024 年国家数据局等 17 部门联合印发的《“数据要素×”三年行动计划（2024—2026 年）》明确提出“以数据要素赋能文旅产业创新”，通过构建文化资源数据共享平台，推动文物、非遗、旅游等领域的多源数据融合应用，培育沉浸式文旅消费新业态（国家数据局，2024）。这为数字化时代中国传统文化的传播指明了方向。

中国传统文化是中华民族历史上各种思想文化、观念形态的总体表征，是居住在中国地域内的中华民族所创造的、为中华民族世世代代所继承发展的、具有鲜明民族特色的、历史悠久、内涵博大精深、传统优良的文化。哲学思想、宗教、服饰、神话传说、饮食、医学等内容共同展现了中国传统文化的魅力。中国传统文化是中华民族精华之结晶，其发展与传承

反映了中华民族独特的智慧和创造力，并以其特有的的民族特性影响着中华民族的发展。深入研究和传播中国传统文化，对于传承历史、丰富文化多样性、增强民族凝聚力、促进文化交流具有重要意义。

近三年来，学界对中国传统文化在元宇宙视域下的传播研究逐渐增多且不断深入，研究者们普遍认可元宇宙为传统文化传播带来的巨大机遇。元宇宙凭借虚拟现实、增强现实、区块链等技术，打破了传统文化传播的时空限制，极大地拓展了文化传播的范围和受众群体。例如，祭孔大典新闻发布会的主办方综合运用齐鲁壹点区块链、数字人等前沿技术应用产品，全景再现元宇宙场景，在元宇宙场景中举办的祭孔大典，吸引了全球各地的用户通过虚拟方式参与其中，使得中华传统文化的核心价值观念得以在更广阔的范围内传播（刘大正，2024）。如美国史密森学会的虚拟博物馆项目、韩国的“数字韩服”项目以及日本对茶道、花道等传统文化的虚拟重现（黄媛等人，2024），均展示了元宇宙技术对文化遗产保护、传播和商业化的多维度促进作用。在中国，“数字故宫”项目以及京剧、昆曲等传统戏曲的虚拟演出，让观众可以在线沉浸式体验传统文化，增强了文化体验的互动性和深度，也吸引了国际观众，提升了中华文化的全球影响力。从传播形态上看，元宇宙中的文化不再局限于传统的单向传播，而是转变为一种互动性更强、沉浸感更深的传播方式。它通过构建虚拟场景，让受众能够身临其境地感受传统文化的魅力，从而提高了文化传播的效能和吸引力。

教育是文化传播的重要途径，学生除了可以通过显性课程学习中国传统文化知识外，还可以在隐性课程潜移默化的渗透过程中了解中国传统文化知识，培养对传统文化的兴趣。元宇宙视域下如何将传统文化传播融入到教育教学中达到润物无声的效果是我们目前在教育中传播中国传统文化需要研究的问题。

因此本研究首先探讨元宇宙的相关概念与中国传统文化传播面临的挑战和机遇，据此开展元宇宙视域下中国传统文化的传播形态研究，并基于研究结论探讨了通过教育传播中国传统文化的具体应用，最后探讨了元宇宙视域下在教育中传播中国传统文化面临的挑战，并据此提出了相关建议，以期为在教育教学过程中传播中国传统文化提供启示。

2.元宇宙概念及其传播特征

2.1.元宇宙概念

1992 年，美国科幻小说家尼奥·斯蒂文森的著作《雪崩》首次提到了“元宇宙”一词，在小说中元宇宙是一个与现实世界平行的虚拟空间，人们可以通过各自的“化身”在其中进行互动、交流、娱乐。这一设定成为后续技术探索的蓝图，例如，电影《头号玩家》中的“绿洲”（Oasis）进一步具象化元宇宙的沉浸式体验，用户可通过 VR 设备进入高度自由的虚拟世界。对于元宇宙的概念目前没有权威、明确的界定。刘革平等指出元宇宙的本质是一个平行于现实世界的在线数字（刘革平等人，2021）。清华大学新闻与传播学院新媒体研究中心发布的《2020—2021 年元宇宙发展研究报告》中元宇宙被认为是整合多种新技术而产生的新型虚实相融的互联网应用和社会形态，它基于扩展现实技术提供沉浸式体验，基于数字孪生技术生成现实世界的镜像，基于区块链技术搭建经济体系，将虚拟世界与现实世界在经济系统、社交系统、身份系统上密切融合，并且允许每个用户进行内容生产和世界编辑（王儒西，2021）。

2.2 元宇宙传播特征

元宇宙的传播主要有以下三个特征。

（1）虚实融合。元宇宙基于虚拟现实（Virtual Reality，VR）、增强现实（Augmented Reality，AR）、混合现实（Mixed Reality，MR）等技术创造虚拟场景，为用户提供沉浸式、

在场感和互动式的丰富体验，用户可以以虚拟数字分身形式在虚拟场景中进行交互，传播活动不受现实世界的限制，拓展了现实世界的生活空间。参与者可以在元宇宙虚拟世界中按照自己的意愿进行各项活动，但无法根本脱离现实世界，元宇宙需要发达的技术支撑，虚拟世界的构成要素来源于现实世界的人和物，而且元宇宙的发展必会经历现实与虚拟世界的过渡（肖珺，2022）。

（2）去中心化传播。区块链技术是支撑元宇宙的基础技术之一，具有不可篡改、不可复制的唯一性、智能合约、去中心化四大基本特征。区块链技术使得每个用户都可以自由独立的参与内容创作，进行文化传播，并受到版权保护。元宇宙中每一个参与者都有相应的世界编辑能力和创造符合自己需求的场景的能力，每个参与者都可以在场景中充分发挥自己的自主性，每个人都是虚拟世界的主体，每个参与者的传播权利都是基本平等的（匡文波和王天娇，2022）。

（3）全球化传播。元宇宙虚拟世界可由不同国家、不同地区、不同文化背景的人参与其中，人与人之间可以充分交流，语言不再是障碍，有助于人们接触和理解不同的文化，真正实现了全球文化的交流和互动，实现了跨文化传播。

3.元宇宙视域下中国传统文化传播形态

传播形态是指传播在一定技术环境中的表现形式和情景，它是媒介系统的具像化。传播形态的核心要素包括媒体形态、受众、传播方式、传播技术、传播环境与情景。本研究拟从媒体形态、传播主体、传播方式、传播环境四个维度对元宇宙视域下中国传统文化传播进行研究。

3.1.元宇宙视域下文化传播的媒体形态变化

文化传播历经口语传播、文字传播、印刷传播、电子传播时代，正在向元宇宙传播时代发展。卡茨等人认为：“需求的社会与心理根源产生对于大众媒介或其他来源的期望，这导致不同模式的媒介接触（或参与其他活动），结果是需求的满足”。文化传播媒介的改变是必然的，其中蕴含着人们在不同时期对文化的不同需求。口语传播时代，文化以口耳相传形式传播，只能在近距离内沟通交流，无法跨越空间距离，且难以保证传播文化的准确性和完整性。文字媒介的出现，打破了空间与距离的限制，同时也使文化传承有了可靠的文献依据。印刷技术的发展推进了文化传播进程，但传播渠道单一，感染力比较差。随着电报、电话、广播、电视等电子媒介的出现，不断满足人们视觉、听觉的需求，文化传播渠道多样，传播范围更广，效果也深入人心。作为科技发展产物的互联网，打破了传统媒介单向传播的特点，受众不仅可以被动接受文化信息，还可以通过短视频、弹幕、网络直播、互动影视等媒介进行互动。在传统媒介迭代演化基础上，元宇宙传播具有传统媒介的优点，同时元宇宙利用虚拟现实、拓展现实等技术营造出的高拟真场景中，实现三维虚拟拟真场景呈现，促进了真实世界与虚拟世界的进一步融合发展，提升了文化传播效率和感官场景体验，便于受众身临其境感受文化魅力。

3.2.元宇宙视域下文化传播主体

传播主体即传播者，是在传播过程中信息的主动发出者。元宇宙视域下中国传统文化传播的传播主体具有以下特点：

（1）大众赋权。元宇宙视域下打破了传统媒介单向传播的特点，并逐渐发展为沉浸泛众传播，人人都可以参与内容生产和创造过程，人人都是文化的传播者。技术的发展推动了文化内容生产从专业生产内容到用户生产内容再到人工智能生产内容模式的转变，赋予了一般民众在文化传播中的主导作用，打破了文化传播主体的限制，从而使传播主体规模的空前

壮大（李萌和王育济，2023）。大众不再仅是文化的接受者，更是文化内容的创造者、传播者。

（2）虚实融合。元宇宙视域下文化传播主体是现实世界本体与虚拟世界化身的虚实融合。通过 VR、XR、AR、数字孪生等技术呈现历史场景、再现文化特色，通过创造虚拟沉浸、身临其境的场景，将原本较为沉闷的传统文化传播变为感官延伸深度沉浸式文化传播，同时虚拟化身可以在元宇宙世界中进行互动交流。

（3）跨界合作。政府机构、文化机构、企业、教育机构和个人创作者可以在元宇宙中进行跨界合作，共同推广中国传统文化。与游戏开发商合作可以开发具有文化特色的游戏，将文化元素与游戏玩法相结合，让玩家在游戏中了解和体验传统文化。与学校、教育机构和专家合作可以研究元宇宙课程和教育资源、开设虚拟讲座，让用户在元宇宙中学习和了解各种文化。跨界合作有助于整合各方资源，为用户提供更丰富的文化体验，提高文化传播的效果和影响力。

3.3.元宇宙视域下文化传播方式

（1）通过虚拟数字人传播中国传统文化

虚拟数字人指存在于非物理世界中，由计算机图形学、图形渲染、动作捕捉、深度学习、语音合成等计算机手段创造及使用，并具有多重人类特征的综合产物（侯文军等人，2023）。虚拟数字人是元宇宙的关键要素，通过赋能中国传统文化传播为现实世界和虚拟世界之间提供了强视觉效果的交互方式。通过虚拟数字人传播中国传统文化时不仅需要技术支撑，更需要丰富的传统文化内容作为支撑。首个文化出海的国风虚拟人“天妤”是以传统飞天和唐代女俑为灵感，在人物形象上最大程度还原东方文化之美，对外在形象和妆容的设计本身就是中华传统服饰文化的传播。“天妤”在短剧《千壁寻踪》中以寻找壁画碎片为线索，在众多恢弘建筑场景中对巾舞、围棋、山海经开明兽、武术等不同的传统文化内容进行故事演绎，持续向世界讲述中国传统文化的故事（李琤，2023）。虚拟数字人只有不断丰富传统文化内涵，维持旺盛生命力和竞争力，才能让传统文化走进烟火生活，焕发青春活力。

随着元宇宙热潮的到来，数字虚拟人物大量涌现，浙江卫视“谷小雨”、中华书局的“苏东坡数字人”、京剧“梅兰芳孪生数字人”、四川眉山市“苏小妹”、篮球“古逸飞”等，他们或是城市代言人、或是文化推荐官、或是文博场所导游等，作为传统文化的新载体，将传统文化与元宇宙技术有机融合，通过视觉化呈现方式传承中国传统文化，极大推动了传统文化的创造性发展。

（2）通过数字藏品传播中国传统文化

数字藏品是在元宇宙理念背景下、应用区块链等现代信息技术而形成的一种具有数字化唯一标识的特定作品和权证，具有不可复制、不可篡改、永久存证等特点**错误!未定义书签。**。传统藏品以实体形式进行收藏，数字藏品超脱于物质载体，以代码、符号形式承载文化内容，摆脱了时间、空间和物质的限制。数字藏品与中国传统文化的结合可以创新传统文化的传播，推动中国传统文化顺应数字化时代发展，同时赋予了传统文化更为年轻的状态，让中国传统文化焕发生机与活力。

数字藏品不仅是一种藏品，更是一种新的传播方式。敦煌文创《千年一瞬·敦煌》《敦煌文创·献瓔珞飞天数字藏品》等数字藏品（酒泉日报，2022）、腾讯安全与川观新闻等联合发布的《物物相生》数字藏品（FISCO_BCOS，2024）、中国体育博物馆联合中体产业集团发行“双奥之路”系列数字藏品、“中华民族读书典故”数字藏品、舞剧《只此青绿》数字藏品、国家文化公园“山河一梦”数字藏品、国画大师石齐《白蛇传》数字藏品等推出后都大受欢迎，这些数字藏品通过展示文化特色元素，呈现中华文化独特魅力。大量传统文化

类数字藏品的出现，在元宇宙环境下为中国传统文化传播搭建起通向数字化发展的桥梁，有利于中国传统文化的传播推广。

（3）通过打造虚拟沉浸场景，线上线下相结合传播中国传统文化

基于元宇宙沉浸式体验的特征，通过 VR、AR、MR 技术、高清影像技术、三维动画技术、全息投影等技术打造虚拟沉浸场景，以此弘扬中国传统文化，增强文化表现张力和传播力度。

全球首个基于唐朝历史文化背景大唐不夜城的“镜像虚拟世界“的元宇宙项目《大唐·开元》，通过元宇宙技术让中国传统建筑能够在数字空间生动地复原和呈现，游客可以在复原的大唐街道上进行旅行和购物，在虚拟世界与现实世界间进行互动，产生强烈的自我代入感和沉浸感。通过打造虚拟沉浸文旅场景，再现大唐长安城 历史风貌，再现中国传统建筑风光，展现大唐风貌，弘扬中华文化。元宇宙敦煌博物馆再现十余个隋唐时期敦煌经典壁画场景，参观者可沉浸式在不同场景中穿梭，并可直接与壁画场景生动交互，模拟敦煌飞天、体验有趣的互动玩法，同时可以装扮自己的虚拟形象。通过元宇宙打造的沉浸虚拟场景以及互动参与方式，实现了中华优秀传统文化传播效能的提升。元宇宙虚实融生营造的沉浸式互动体验，突破了文化传播的空间限制，推动中国传统文化的广泛传播。

3.4.元宇宙视域下文化传播环境

（1）虚拟沉浸传播环境

元宇宙综合利用 VR、AR、MR、数字孪生、5G、人工智能（Artificial Intelligence，AI）、区块链等技术打造虚拟沉浸的传播环境，在这种环境中，用户可以通过各种虚拟设备（如 VR 头盔、AR 眼镜等）沉浸于一个仿真的、包含许多 3D 场景和角色的虚拟世界，让用户感受到身临其境的感觉。用户可以通过虚拟化身在虚拟环境中与其他用户互动、交流和协作，元宇宙可提供语音、视频聊天功能，以及表情、动作等非语言沟通方式，从而实现信息和文化的多渠道传播。

元宇宙可利用技术再现建筑、山川、历史场景，助力文化传播在虚拟沉浸环境下的全新体验。元宇宙基于开放性的特点支持跨界合作共同打造虚拟沉浸场景如虚拟展览、在线游戏、虚拟音乐会等，将现实世界元素融入虚拟环境中，为用户带来多元化的文化体验。元宇宙通过打造功能丰富、高度互动且具有开放生态的虚拟沉浸传播环境，有助于拓展传统传播方式的边界，为用户带来全新的虚拟体验。

（2）全球互联传播环境

在元宇宙环境下，可以跨越地理、文化、语言等多种障碍进行互动交流，同时元宇宙全球互联环境能够将不同的虚拟世界和现实世界进行无缝连接，实现全球范围内的互动与交流，用户可接触到不同文化背景的人，体验到来自世界各地的文化。

元宇宙全球互联环境可以与现实世界的社交网络进行整合，让用户在虚拟空间中建立和拓展社交关系，实现全球范围内的社交互动。沉浸传播范式下的元宇宙实践重塑虚拟与现实结合的数字时空，带来更为平等、共治的数字互动边界，革新了全球社会的边界秩序（李沁和王浩丞，2022），消弭了时间与空间、虚拟与现实的界限，拓宽了文化传播渠道，从而可以更好地传播和发扬中华优秀传统文化。

4.元宇宙视域下中国传统文化传播的教育应用

教育是文化传播的重要手段和途径，基于上述对元宇宙视域下中国传统文化传播形态的研究，分析其在教育中的应用如下。

4.1.以教师为传播主体

教师在教学过程中担负着传播文化的重要责任，明确教学目标时可以将传播中国传统文化理念融入情感态度与价值观目标中，激发学生热爱、学习、弘扬中国传统文化的热情。

元宇宙视域下，教育在教学模式、教学方法等方面发生了重大变革。教师可以通过 AR、VR 等技术构建教学资源，将中国传统文化融入到课程内容中，通过创设的虚拟沉浸教学情境给予学生全新的文化体验，激发学生学习与弘扬传统文化的热情，实现知识传授、能力培养和价值引领的有机统一（王尧，2022）。

教师可以利用元宇宙的技术如虚拟现实、增强现实等创建隐含传统文化的实践活动场景，为学生提供沉浸式的学习体验，如通过虚拟现实技术让学生身临其境体验历史人物的时代背景，使学生在不知不觉间接受优秀传统文化的熏陶，深入感受中国传统文化的魅力。元宇宙视域下学习的沉浸感、体验感为教师教学带来了挑战，教师应不断学习和深入了解中国传统文化，不断反思和完善自己在教学过程中传播传统文化的方法和策略。

4.2. 以学生为传播主体

学生是传播中国传统文化的新生力量。学生应积极参加传统文化虚拟讲座、数字藏品等相关展览，通过领略传统文化风貌，感受中国传统文化魅力；积极参加传统文化主题实践活动，在虚拟融合的场景下沉浸体验、感受中华灿烂文明，同时可以根据自己的理解和体验，创作和分享自己与传统文化的故事，以 AI 绘画、音乐、舞蹈等不同形式创作传统文化主题作品，充分发挥学生的主体地位。

4.3. 以学校为传播主体

学校应加强和社会各机构的跨界合作，定期举办传统文化相关展览，举办传统文化数字藏品相关展览，通过国潮、插画等学生喜闻乐见的方式将传统文化元素呈现在学生面前，学生可以近距离接触、观看，引导学生了解和学习中国传统文化。

学校应鼓励和表彰传播中国传统文化的优秀实践案例，建立元宇宙视域下中国传统文化传播的相应的评价和激励机制，促进中国传统文化健康、积极生动传播。

学校应积极为教师提供在元宇宙视域下的培训和发展机会，如在线研修、教育创新研讨会等，提升教师在元宇宙环境下的教育能力。加强教师在元宇宙视域下思维方式、教学理念、教学方法、信息化能力的培训与革新，提高教师在元宇宙视域下的教学能力和素养（徐建等人，2022）。

4.4. 以社会为传播主体

社会各机构可跨界合作，设计与传统文化主题相关的数字藏品，并与学校合作定期开展主题展览，为学生带来全新的文化体验感；设计虚拟数字人作为文化推荐官，以虚拟形象为学生传播传统文化相关知识或设计以历史人物为原型的虚拟数字人，带领学生穿越时空，体验历史朝代文化；各机构可结合传统文化元素如陶瓷、水墨画、三国文化等设计 VR 类游戏，在游戏中增加国风元素，以传统文化为背景，设计文化闯关主题场景，充分利用虚拟沉浸特点，在游戏中身临其境感受灿烂文明，体验传统文化。

社会各机构和文化组织可以通过举办传统文化交流活动，形成弘扬中国传统文化的社会风气，促进中国传统文化的传播和发展。

5.元宇宙视域下在教育中传播中国传统文化面临的挑战和建议

元宇宙为在教育中传播中国传统文化带来新发展机遇的同时，也带来了新的挑战，主要体现在以下几个方面。

（1）资源娱乐化与同质化风险。元宇宙技术应用易陷入“形式优先于内容”的误区。例如，部分教育类元宇宙产品过度依赖游戏化机制（如闯关奖励、虚拟角色养成），导致学

生对文化符号的认知停留于表层娱乐，难以理解深层价值。同时，多家机构争相开发“故宫”“敦煌”主题课程，但内容同质化严重（如全景漫游+语音解说模式占比超 60%），缺乏基于地域特色或学科融合的创新设计。

（2）动态教材开发能力不足。元宇宙教材需整合三维模型、交互脚本与动态叙事，但多数教师缺乏相关技术能力。例如，传统教师难以将视频、音频、文本等元素转化为沉浸式教学场景，导致资源开发依赖第三方公司，成本高昂且适配性低。

（3）隐私和安全方面。在元宇宙中，每个人都可以以虚拟数字化身的方式进行交流、互动、学习，这些数据会被平台所收集，或通过 VR 眼镜、头盔等辅助设备收集用户的眼动数据和周围环境信息，这些有可能产生数据安全和隐私问题。如何保证用户在通过元宇宙传播文化过程中的安全和隐私权益是一个重大挑战。

（4）法律法规与道德伦理方面。虚拟数字人以现实人物为原型时，在设计过程中可能会涉及到侵犯原型人物的形象权，甚至可能会对其形象进行恶俗化。在现实世界中触犯法律需要承担相应的刑事或民事责任，但在元宇宙中虚拟数字人承担责任方式并不明确，以及元宇宙化身所对应的现实本体是否需要承担责任，目前法律并未有定论。元宇宙视域下传播文化过程中，涉及各族、各国文化和信仰，历史敏感有争议问题时可能会产生伦理道德问题。

针对元宇宙视域下在教育教学过程中传播中国传统文化时存在的问题以及面临的挑战，本研究提出以下几点建议：

（1）制定教育资源开发标准与评估框架。要求元宇宙课程必须包含“文化解析模块”与“深度学习任务”，如通过虚拟角色对话揭示戏曲唱腔背后的伦理观念、基于数字孪生技术还原古建筑榫卯结构并完成搭建挑战，避免娱乐化稀释文化内核。鼓励高校与地方文化机构合作开发特色课程，如安徽高校聚焦徽派建筑营造技艺、云南院校侧重民族服饰数字化复原等，依托地域文化资源形成差异化优势。

（2）构建教育资源开源生态。搭建“中华文化元宇宙教学案例库”，汇聚优质教案、3D 模型素材与交互脚本，如浙江大学“求是元宇宙教育联盟”模式，支持教师根据学情进行二次开发。推动高校与科技企业合作，如东方剑桥教育集团与蓝色光标共建元宇宙产业学院，联合开发标准化教学工具包，降低教师技术门槛

（3）保护和尊重中国传统文化。元宇宙平台和产品应尊重原作者的版权，保护文化传承人的权益，并严厉打击侵犯传统文化产权行为。在开放性、全球化的元宇宙中，学校、家庭、教师等要加强对学生保护、尊重、弘扬中国传统文化正确价值观的塑造，在教育中传播中国传统文化时注重保护中国传统文化的地位以及文化的独特性。

（4）完善元宇宙相关法律法规和监管机制。政府应制定专项法规，明确虚拟场景中历史文化的真实性审核标准，并建立数据安全保护机制，要求 VR 设备厂商采用端到端加密技术，禁止商业化滥用学生行为数据。教育机构需联合教师成立伦理审查委员会，对元宇宙课程进行动态筛查，剔除过度娱乐化或史实错误内容。同步修订《非物质文化遗产法》，增设元宇宙传播条款，规定虚拟数字人原型须获传承人授权，对恶意篡改文化符号的行为设定罚则，从立法、监管、执行三端保障文化传播的严肃性与教育价值。

6.总结

本文对元宇宙视域下中国传统文化的传播形态进行研究，并据此探讨了其在教育中的具体应用，同时提出了元宇宙视域下中国传统文化通过教育途径传播面临的挑战和未来发展建议，以促进中国传统文化在教育中焕发生机与活力。研究元宇宙视域下在教育中传播中国传统文化对传承和创新中国传统文化、增强文化自信、推动文化产业发展具有重要意义。随着

元宇宙的不断发展和普及，中国传统文化传播将迎来更多的机遇和挑战，元宇宙视域下在教育过程中实现中国传统文化的传播和发展需要文化、科技、政府等多方面的共同努力，以期在全球范围内传播和发扬中华优秀传统文化。

基金

教育部教育管理信息中心 2024 年度国外教育研究委托课题（MOE-CIEM-20240008）

参考文献

侯文军, 卜瑶华, 刘聪林. (2023). 虚拟数字人：元宇宙人际交互的技术性介质. 传媒, (4), 25, 27, 29.

黄媛, 梁艺莹, 王娅, 李刚. (2024).元宇宙时代下的非遗文化保护与复兴路径探索：基于全球案例与中国实践的对比研究. 社会科学前沿, 13(12): 228-235.

匡文波, 王天娇. (2022). 元宇宙传播的新特点与新机遇. 新闻论坛, (1), 7-8.

李琤. (2023). 虚拟数字人竞速元宇宙赛道：以虚促实，以虚强实. 中国文化报, (8).

李萌, 王育济. (2023). 中华优秀传统文化传播的数字机制与趋势. 人民论坛, (2), 104-106.

李沁, 王浩丞. (2022). 表象、实在与失范：沉浸传播理论范式下的元宇宙实践. 新闻与写作, (10), 65-78.

刘大正.(2024 年 7 月 19 日).数字技术赋能中华优秀传统文化传播新路径.中国社会科学网-中国社会科学报.https://www.cssn.cn/skgz/bwyc/202407/t20240719_5765554.shtml (2023 年 4 月 28 日访问)

刘革平, 王星, 高楠. (2021). 从虚拟现实到元宇宙：在线教育的新方向. 现代远程教育研究, (6), 12-22.

清华大学新闻与传播学院新媒体研究中心. (2021 年 9 月 16 日). 2020-2021 年元宇宙发展研究报告. 蚂蚁文库. https://www.mayiwenku.com/p-25484305.html (2023 年 3 月 28 日访问)

搜狐网. (2022 年 5 月 31 日). 全国 16 省市数字经济元宇宙相关政策. 搜狐网. https://www.sohu.com/a/552670080_121124715 (2023 年 3 月 28 日访问)

王尧. (2022). 从课程思政走向大课程思政. 阅江学刊, (3), 92-98.

肖珺. (2022). 元宇宙：虚实融合的传播生态探索. 人民论坛, (7), 40-44.

徐建, 王俊, 钟正等. (2022). 教育元宇宙时代教师发展的挑战与应对. 开放教育研究, (3), 51-56.

流程自動化之多層級教育關係人行動建議智慧儀表板：以台灣教育大數據為例

Automated Decision Support for Multi-Level Educational Stakeholders: An AI-Powered Dashboard Using Taiwan's Educational Big Data

詹欣輮[1*]、楊子奇[2]
國立陽明交通大學教育研究所
[*] wendyjou.hs12@nycu.edu.tw

【摘要】 本研究結合學習分析與生成式人工智慧技術，設計一套具備自動解讀教育數據並提供相對行動建議的智慧教育儀表板，協助教育決策者、教師與學生等利害關係人進行教育任務。本研究以台灣的教育大數據開放資料為基礎，分析學生在多個數位學習平臺上的行為紀錄共 14,359,163 筆。基於整體表現分析、分群比較與人工智慧建議模組，為不同層級的教育角色提供即時資料洞察與具體行動建議。此研究結果有機會提升教育現場的數據解讀能力，發揮運用教育大數據提升整體教學與學習品質的優勢，並為未來教育儀表板與提供重要參考架構。

【關鍵詞】 教育大數據；生成式人工智慧；學習分析；教育儀表板；數據視覺化

Abstract: *This study integrates learning analytics and generative artificial intelligence to design an intelligent educational dashboard capable of automatically interpreting educational data and generating corresponding actionable recommendations. The system aims to support key educational stakeholders—including policymakers, teachers, and students—in carrying out their respective educational responsibilities. Based on Taiwan's educational big data, the study analyzes a total of 14,359,163 behavioral records collected from multiple digital learning platforms. The dashboard consists of three core modules: overall performance analysis, cluster-based comparisons, and AI-generated recommendations. These modules provide real-time data insights and role-specific suggestions tailored to stakeholders at different levels of the educational system. The findings of this research have the potential to enhance data interpretation capacity in educational settings, maximize the value of educational big data, and improve the overall quality of teaching and learning. Furthermore, the proposed system offers a reference framework for future developments in AI-powered educational dashboards.*

Keywords: Educational Big Data, Generative Artificial Intelligence, Learning Analytics, Educational Dashboard, Data Visualization

1.研究背景

隨著數位化學習的普及，教育領域累積了大量的學習數據，涵蓋學生學習行為、教師教學成效、學校管理決策等多個層面（Siemens & Baker, 2012）。教育數據視覺化技術與儀表板的應用，能使不同層級的教育工作者能夠快速掌握關鍵資訊，提升決策的精準度與執行效果（Daniel, 2015）。然而，若缺乏有效的數據整理、分析與視覺化技術，這些數據背後的意涵難以洞察也不易轉化為可被執行的行動建議。教育數據視覺化，特別是教育儀表板可以快速掌握學習情況，進而提升學習成效與教育決策的準確度。過去，教育數據儀表板的應用需求可劃分為三個主要層級，分別是教育決策者、教師、學生，其目的在提供各層級使用者

面臨不同的需求與挑戰，需量身設計儀表板功能（Cukurova, Luckin & Clark-Wilson, 2019）。例如，在教育管理者層級，需要宏觀掌握全國教育發展趨勢，評估政策成效，並去制定未來的政策，透過大規模的教育數據儀表板，可監測各學科的學習狀況、不同區域的學習資源分布，以及數位學習工具的使用情況，從而據以調整教育政策（Bienkowski, Feng & Means, 2012）。更可進一步關注公平與教育資源分配問題，透過數據分析評估城鄉教育差距，調整補助機制，確保所有學生均能獲得適當的學習支持（Huda et al., 2017）。

在學校層級對數據的需求更為精細，包含個別學生的學習行為、班級整體學習進度，以及教師的教學策略優化。對主管層級可以透過儀表板了解學校內部的學習成效，監測不同科目的學生表現，並據以調整教師的教學資源配置（Papamitsiou & Economides, 2014）。對教師而言，學習分析儀表板可提供即時的學生學習數據，幫助教師識別學生的學習瓶頸，並設計個別化的教學介入措施，以提升教學效能（Arnold & Pistilli, 2012）。對學生層級而言，個人化學習儀表板可幫助學生追蹤自身學習進展，調整學習策略，提升學習規劃與自我調節能力（Aguilar et al., 2021）。

受惠於近年人工智慧（Artificial intellegent,AI）技術的發展，學習分析從數據監測進化為智慧化的決策。不但可即時追蹤學生學習歷程，預測學業表現，並根據學習者需求提供個人化建議，進一步提升學習成效。此外，整合 AI 的數據視覺化技術，有助於教師與學生能夠更直觀地解讀學習數據，並據此調整教學與學習策略（Mejeh & Rehm, 2024）。學習分析與資料視覺化的產出，需轉換為行動建議才能真正發揮教育大數據的價值。然而，解讀教育儀表板往往需要一定的專業的背景與實務經驗，因此，目前多數研究與應用仍聚焦於視覺化呈現或單一應用（Leitner, Khalil & Ebner, 2017）。Paulsen 和 Lindsay（2024）在一項系統性回顧中指出，儘管學習分析儀表板在高等教育中廣泛應用，但其設計仍主要以分析為導向，缺乏足夠的教育學考量，難以有效促進學生的學習過程。Pozdniakov 等人（2025）亦指出，視覺化素養和指導對教師解讀學習分析儀表板的影響，發現缺乏適當的指導可能限制教師有效利用儀表板進行教學決策。這些研究表明，儘管學習分析儀表板在教育中具有潛力，但如何協助不同層級使用者解讀教育儀表板並轉化為行動建議支持仍是當前待解決的議題。因此，本研究透過多層級學習分析儀表板的設計與實作，提出一套結合 AI 輔助的智慧教育數據應用框架，期望支持不同層級的教育行動建議。具體而言，在宏觀層級，協助教育管理者掌握跨區域、跨學科之學習表現與資源投入成效，有助於政策滾動修正與資源配置之精準化，推動教育治理數據化、科學化。在中觀層級，協助教師與管理人員即時掌握班級學習趨勢與學科成就落差，作為教學策略調整與教學支援介入之依據，提升校內教學管理效率。在微觀層級，協助學生透過個人化視覺介面掌握學習進度與成就盲點，提升自我調節能力與學習參與度。整體而言，本研究不僅推進學習分析與教育儀表板的應用，也為教育資料科學在政策、教學與學習層級的深化方面提供重要參考依據。

2.文獻探討

2.1.多層級的教育視覺化儀表板

教育視覺化儀表板作為學習分析的重要應用是能夠將大量、複雜的教育數據轉化為直觀且易於理解的圖表，以支持不同層級的教育利害關係人進行決策（Siemens & Baker, 2012）。適當的數據視覺化設計能幫助教師快速識別學生的學習行為模式與異常，進而採取針對性的教學干預。同時，學生透過個人化儀表板掌握自身學習狀況，有助於提升學習規劃與自我調節能力（Aguilar et al., 2021）。然而，學習分析儀表板的有效性取決於其設計是否能夠滿足

不同使用者的需求，並確保數據的呈現方式不會對決策或教學調整產生誤導（Schwendimann et al., 2016）。

在教育決策者層級來看，學習分析儀表板的主要服務在於宏觀教育政策制定和資源分配。政策制定者依賴全國或區域性的教育數據來評估學生學習成效趨勢，識別教育資源配置的不均衡現象，並據此調整策略（Means, B., 2010）。透過動態數據視覺化工具，決策者可以更精確地評估教育改革措施的影響，用於監測教育計畫的實施成效，為政策制定提供數據支持（Wiedbusch, et al., 2021）。新北市教育局張明文局長（2024）表示，教育資料平台設計採用資料視覺化結合互動式網站，以數據儀表板呈現資料，從中可對教育整體表現及發展狀況有個初步概念。

在教師層級，學習分析儀表板的核心目標是幫助教師理解學生的學習行為，從而提供個別化的教學調整（Schwendimann et al., 2016）。教師可透過儀表板即時監測學生的學習進度、測驗結果與課堂參與度，並根據數據進行精準的課堂調整。例如： Pei, et al.（2024）發現使用儀表板來分析不同學生群體之間的學習績效落差，並找出導致這些落差的潛在因素。Sharma et al.（2024）也針對 K-12 教育的研究顯示，當教師能夠透過視覺化儀表板掌握班級內學生的學習表現，低成就學生的學習成果可獲得顯著提升。也就因為分析技術的進步使得教師能夠更精細地解讀學生的認知與行為模式，從而設計更有效的學習支持的機制。

而在學生層級，學習分析儀表板可作為個人化學習支持工具，幫助學生設定學習目標、追蹤學習進度，並獲取即時回饋。研究表明，當學生能夠透過可視化儀表板檢視自身學習歷程，他們的學習動機與自我調節能力會顯著提升（Aguilar et al., 2021）。此外，Holstein et al.（2019）指出，若將人工智慧技術整合至學習輔導系統，不僅能自動分析學生的學習數據，還可根據學習者需求提供個性化學習建議，進一步提升學習成效。

綜合以上，多層級的學習分析儀表板在教育決策者、教師與學生之間發揮的關鍵作用，協助不同層級的使用者做出數據驅動的決策，以確保不同的使用者能夠有效解讀並採取適當行動，以提升教育成效。

2.2.人工智慧在學習數據分析的應用

學習分析強調透過數據來調整並改善學習體驗，提升學習者的成效與教育決策的精準度（Siemens, 2012）。隨著人工智慧（Artificial Intelligence, AI） 技術的快速發展，學習分析已從單純的數據監測進化為智慧化的學習支持系統。AI 不僅能即時追蹤學生的學習表現，還能識別學習困難，並根據個別學習者的需求提供個人化建議。

過去研究也顯示，在教育管理者端可以透過 AI 分析大規模行為與學習資料，進行政策層級的數據建模與個別化方案，Yan et al.（2023）指出，AI 可支援教育管理者進行非結構化數據的深度分析，提升政策回應速度與策略調整的科學性。而教師端也提供動態化的即時建議，Wiedbusch 等人（2021）提出「MetaDash」教師儀表板設計，可整合學生表現與自我調節學習的即時分析，並提供針對性的教學建議，幫助教師即時調整課堂內容與干預時機；而在學生端，AI 自動建議系統可以根據一些邏輯去行動導引，當學生完成率低，建議學生設定短期目標並引導補課；若學生長期未登入，自動推播提醒與個人學習任務。這些 AI 驅動的智慧建議模組不僅提升了教育者的即時決策效率，也有助於學生自我調節學習與提升學習動機（Aguilar et al., 2021；Holstein et al., 2019）。這樣的技術創新，使 AI 不僅成為學習分析的強化工具，更針對各層級推動了教育決策的智慧化發展。

此外，AI 在學習分析與數據視覺化的應用進一步也提升了數據的可解讀性與決策的準確性。透過 AI 驅動的視覺化技術，學習者與教師能更直觀地理解數據圖表，從中取得關鍵資訊。亦可進一步透過自然語言處理技術自動解讀學習儀表板上的趨勢，提供即時分析報告，

使學習者能迅速調整學習策略，教師亦能據此優化教學方式（Mejeh & Rehm, 2024)。隨著生成式人工智慧（Generative AI, GenAI）的快速發展，AI 能有效減少了教師在檢視和分析資料方面的工作量(Wang et al.,2025）。有效提升教育決策的精準度與學習機會的公平性，以確保 AI 真正促進有效與公平的學習(Yadav, 2025）。

2.3.數據決策管理與教育創新

隨著數據治理觀念導入教育體系，人工智慧與數據視覺化已成為教育政策規劃的重要決策工具。根據 OECD（2021）在《AI and the Future of Skills》報告說明，AI 在教育現場的最大價值，正是能夠跨時間與層級提供預測與調整能力，進一步強化教育政策的回應性與資源分配的精準度。特別是在政府部門，AI 工具能協助識別學科發展落差、師資資源分布不均，或區域間數位學習普及率的差異。

在教育環境中，技術中介與數據驅動決策對學習成效與教學流程均產生深遠影響（Zeide, 2017）。透過學習分析儀表板，教育決策者、教師與學生可迅速掌握學習數據，據以調整教學策略與優化政策規劃。然而，傳統的數據視覺化方式仍面臨資訊過載與解讀門檻等挑戰，使 AI 技術逐漸成為強化決策支持的重要推力。UNESCO（2022）在 AI 教育應用的倫理與實踐框架中指出，AI 若應用於教育決策，應具備「可解釋性」、「透明性」與「在地性」三項原則，確保系統所產出的分析結果與建議，能有效回應各地教育需求，並促進教育機會的公平與正義（UNESCO, 2022）。過去研究也指出，透過分析學生學習數據（如歷次成績、互動紀錄等），AI 不僅能依據個別需求量身推薦學習內容，提供個人化的學習體驗（Al Ka'bi, 2023），更能主動偵測學習異常模式，辨識高風險學生，並提出早期干預建議，以提升整體學習成效（Zawacki-Richter et al., 2019）。

綜言之，AI 在數據驅動的教育決策管理中扮演關鍵角色，能為教育決策者、教師與學生提供更智慧化的判斷方向。透過 AI 所驅動的學習分析與視覺化技術，教育工作者得以更準確掌握學習趨勢，調整未來教學策略，並提升教育資源的配置效率與使用效益。

3.研究設計

3.1. 資料來源

本研究採用 2024 年台灣教育大數據雲開放資料作為主要分析來源，該資料依據國際評量資料處理原則建置，經匿名化處理後，共計約 14,359,163 筆跨平臺的學習行為紀錄。內容涵蓋三個學習平臺（dp001、dp002、dp003），包含使用者的學習活動、練習作答、影片觀看、互動操作、測驗表現等多類別行為數據。透過整合與分析上述數據，建立一套多層級的 AI 教育儀表板，分析學生在不同時期與情境中的學習行為，並針對教育決策者、教師與學生等不同層級提供個人化且具決策支持功能的視覺化結果與 AI 行動建議。

3.2. 實驗流程

本研究之智慧數據儀表板則針對三類使用者（教育管理者、學校教師、班級學生）設定核心分析目標，對應設計「整體表現分析」、「群組表現分析」與「個人化學習分析」三大模組，最終由 AI 模組生成具體行動建議（如圖一所示）。使用者在互動式視覺化儀表板中操作圖表、觀察數據，系統將即時分析其操作行為與圖表狀態，透過 AI 自動提供資料摘要、潛在問題偵測與建議的行動方案。例如，當教師檢視學生影片學習時間分布時，AI 模組可提示學習落後群體，並建議補強活動或課程調整策略。流程以使用者導向任務結合 AI 分析邏輯，打造具備互動性與回饋性的多層級儀表板，並進一步驗證其在教育決策與教學應用中的可行性與效益。

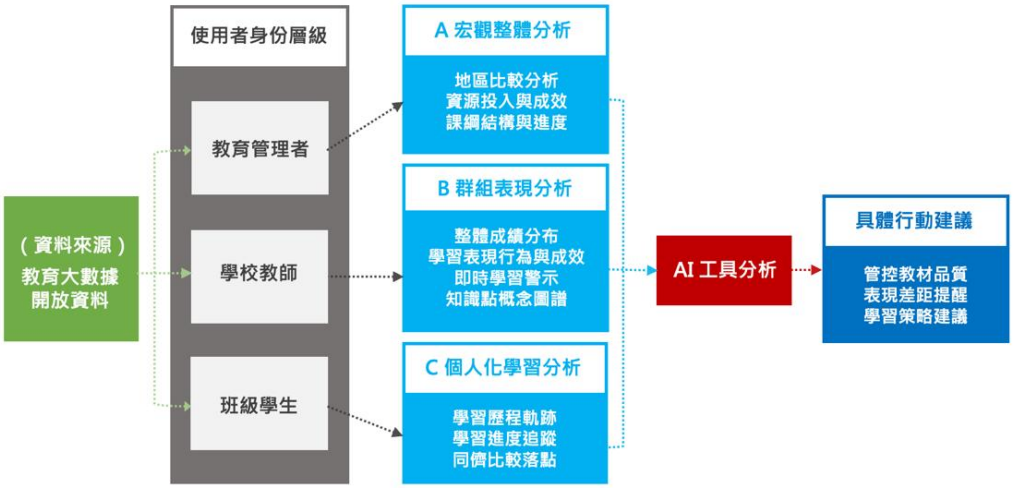


圖 1 儀表板實驗設計架構圖

3.3. 多層級學習數據 AI 儀表板

本研究依據教學現場需求設計教師儀表板，如圖二所示，分為三大區塊。第一區呈現學生在平臺上的操作行為分佈（如播放、暫停、拖曳），並結合各科成就分布圖，協助教師辨識學習異常行為與學科落差。第二區聚焦學生表現層級、活動熱點與弱點分析，提供教師進行班級群體與個別學生的多角度觀察與對比。第三區為 AI 模組，根據圖表動態生成數據解讀與教學建議。

圖 2 儀表板設計示意圖與流程說明

圖三則展示不同角色（決策者、學生）所對應的 AI 建議呈現方式。決策者介面聚焦整體行為與低成就占比，AI 模組可提出內容優化與資源調整建議；學生介面則提供個人操作行為與成果分析，並給予學習策略建議。整體設計結合大型語言模型（Large Language Model, LLM）技術（OpenAI GPT-4）來生成即時建議。使用者與儀表板互動時，系統可即時截取

圖表的核心重點（如學生平均作答時間、影片完成率等），並透過指令模板將其轉為具體分析建議，以教師角色為例，當圖表顯示某班學生影片觀看不足，LLM 則會輸出：「影片完成率低於平均，建議進行課後補強或升級影片互動品質。」回應結果會以自然語言方式顯示於儀表板側邊，除了實現多層級角色的即時回饋與數據應用，也增強使用者對資料的理解與行動引導。

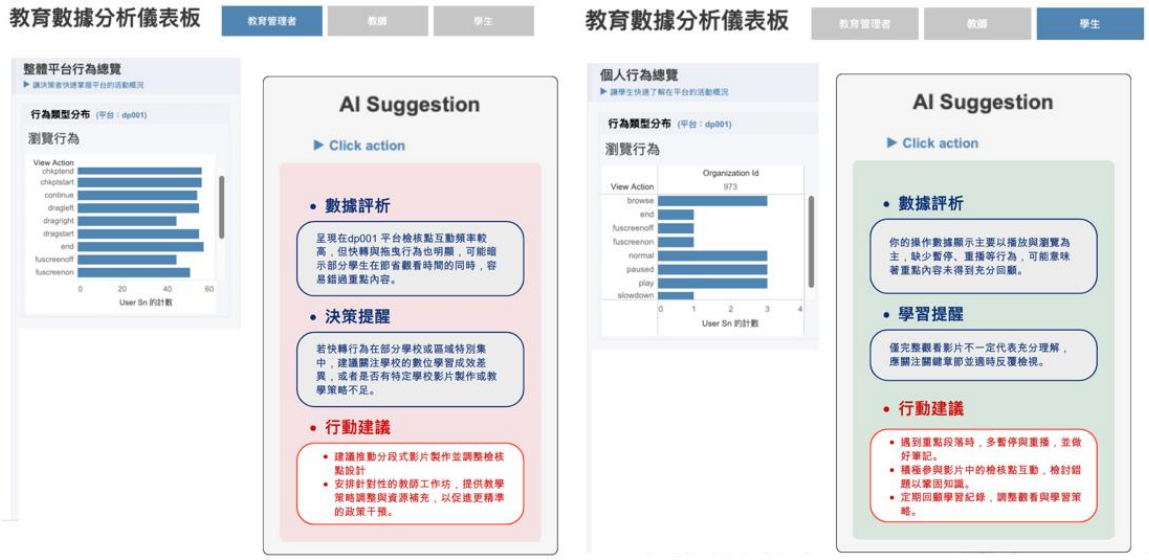


圖 3 對照決策者與學生在「行為類型分布圖」的 AI 對應建議

4.預期結果與未來工作

本研究預期透過整合 GenAI，建立具理論基礎與實務價值的教育儀表板架構。全面提升教育利害關係人運用教育大數據提升行動決策的效率與品質。對教育決策者而言，可以宏觀整理學習趨勢以進行政策施行的評估；對老師而言，能以視覺化清晰直觀的方式幫助教師快速掌握學生的學習行為、學習成果與群體差異，並強化數據驅動的教學決策，更快識別學習異常與需求，並能據此即時調整教學策略；對學生而言，透過個人化儀表板與即時學習回饋，預期能提升其學習參與度與自我調節能力，進而在成績與學習動機上獲得正向影響。

基於預期結果，未來將從以下三個方向延伸進一步工作：首先，進一步擴展其他數據影用來源規模與異質性，包含納入更多學校或跨平台資料，以驗證儀表板架構的通用性與可擴展性。其次，優化儀表板的互動功能，導入更精緻的個人化視圖設定，讓不同角色（如教育管理者、教師、學生）可依據實際需求動態調整內容與視覺形式。第三，未來期望可將研究成果整合至學習管理系統中，提升教育現場的實用性與推廣性，並探索其與教育部數據平台、縣市決策系統的整合潛力，開擴學習分析在教育創新中的應用。

參考資料

董冠怡（2024 年 3 月 25 日）。新北「教育資料平台」蒐集大數據應用辦學。自由時報。https://news.ltn.com.tw/news/life/breakingnews/4619474

Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 267-270).

Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. Computers & Education, 162, 104085.

Al Ka'bi, A. (2023). Proposed artificial intelligence algorithm and deep learning techniques for development of higher education. International Journal of Intelligent Networks, 4, 68-73.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief. Office of Educational Technology, US Department of Education.

Cukurova, M., Luckin, R., & Clark-Wilson, A. (2019). Creating the golden triangle of evidence-informed education technology with EDUCATE. British Journal of Educational Technology, 50(2), 490-504.

Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British journal of educational technology*, 46(5), 904-920.

Huda, M., Jasmi, K. A., Hehsan, A., Mustari, M. I., Shahrill, M., Basiron, B., & Gassama, S. K. (2017). Empowering children with adaptive technology skills: Careful engagement in the digital ınformation age. *International Electronic Journal of Elementary Education*, 9(3), 693-708.

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Designing for complementarity: Teacher and student needs for orchestration support in AI-enhanced classrooms. In *Artificial Intelligence in Education: 20th International Conference*, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20 (pp. 157-171). Springer International Publishing.

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—a literature review. *Learning analytics: Fundaments, applications, and trends: A view of the current state of the art to enhance E-learning,* 1-23.

Means, B. (2010). Technology and education change: Focus on student learning. *Journal of research on technology in education*, 42(3), 285-307.

Mejeh, M., & Rehm, M. (2024). Taking adaptive learning in educational settings to the next level: Leveraging natural language processing for improved personalization. *Educational technology research and development*, 72(3), 1597-1621.

OECD. (2021). *AI and the Future of Skills: Capturing the potential and assessing the risks*. OECD Publishing.

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49-64.

Pei, B., Cheng, Y., Ambrose, A., Dziadula, E., Xing, W., & Lu, J. (2024). LearningViz: a dashboard for visualizing, analyzing and closing learning performance gaps—a case study approach. *Smart Learning Environments*, 11(1), 1-25.

Paulsen, L., & Lindsay, E. (2024). Learning analytics dashboards are increasingly becoming about learning and not just analytics-A systematic review. *Education and Information Technologies*, 29(11), 14279-14308.

Pozdniakov, S., Martinez-Maldonado, R., Tsai, Y. S., Echeverria, V., Swiecki, Z., & Gašević, D. (2025). Investigating the Effect of Visualization Literacy and Guidance on Teachers' Dashboard Interpretation. *Journal of Learning Analytics*, 12(1), 367-390.

Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254).

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380-1400.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., ... & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE transactions on learning technologies*, 10(1), 30-41.

Sharma, D. M., Ramana, K. V., Jothilakshmi, R., Verma, R., Maheswari, B. U., & Boopathi, S. (2024). Integrating Generative AI Into K-12 Curriculums and Pedagogies in India: Opportunities and Challenges. *Facilitating Global Collaboration and Knowledge Sharing in Higher Education With Generative AI*, 133-161.

UNESCO. (2022). *Guidance for Generative AI in Education and Research. United Nations Educational*, Scientific and Cultural Organization.

Wiedbusch, M. D., Kite, V., Yang, X., Park, S., Chi, M., Taub, M., & Azevedo, R. (2021, February). A theoretical and evidence-based conceptual design of metadash: An intelligent teacher dashboard to support teachers' decision making and students' self-regulated learning. In *Frontiers in education* (Vol. 6, p. 570229). Frontiers Media SA.

Wang, Z., Lin, W., & Hu, X. (2025, March). Self-service Teacher-facing Learning Analytics Dashboard with Large Language Models. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference* (pp. 824-830).

Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024, March). Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 101-111).

Yadav, S. (2025). Advancing Education With AI-Driven Education for Diverse Learners: Integrating AI Into Pedagogical Practices. In *Integrating AI Into Pedagogical Education* (pp. 49-80). IGI Global Scientific Publishing.

Zeide, E. (2017). The structural consequences of big data-driven education. *Big Data*, 5(2), 164-172.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, 16(1), 1-27.

探索人工智能增强的 STEAM 课程中小学生的计算思维过程：一项滞后序列分析方法

Exploring Primary School Students' Computational Thinking Process in an AI-Enhanced STEAM Course: A Lag Sequential Analysis Approach

潘东晨[1]，龙陶陶[2*]，李吉梅[3]
[123]华中师范大学人工智能教育学部
[*] taotaolong@ccnu.edu.cn

【摘要】 STEAM 教育要求学习者能够运用跨学科知识，从不同学科中汲取智慧，解决复杂的现实问题；计算思维能力涉及分解问题、抽象问题和设计算法等过程；人工智能的个性化支持能力可促进 STEAM 课程有效开展。分析人工智能增强的 STEAM 课堂中学生的计算思维行为，能帮助教师了解学习行为模式，诊断培养计算思维的潜在挑战。本研究在一所资源匮乏的小学的志愿教学项目中开展，探讨人工智能增强校本 STEAM 课程中师生的计算思维行为模式。该课程为期六周，共收集了 798 个学生和教师行为样本。通过滞后序列分析，得到了分解问题、设计算法和总结归纳三个关键行为序列。
【关键词】 AI；计算思维；STEAM；滞后序列分析

Abstract: *STEAM education requires learners to be able toutilize interdisciplinary knowledge to draw wisdom from different disciplines to solve complex real-world problems, from which computational thinking capacity is necessary for decomposing problems, abstracting problems and designing algorithms. The personalized support capabilities of artificial intelligence can facilitate effective STEAM courses. Analyzing students' computational thinking behaviors in AI-enhanced STEAM classrooms can help teachers understand patterns of learning behavior. In this study, the computational thinking behavioral patterns of teachers and students in an AIenhanced school-based STEAM course was developed for a voluntary teaching program for an under-represented primary school was explored. This course lasted for six weeks and a total of 798 behavioral samples were collected from students and teachers. Three behavioral sequences, which were decompose problems, design algorithms, and summarize, were derived from lagged sequence analysis.*
Keywords: AI;computational thinking; STEAM; lag sequential analysis

1.引言

学生越来越需要综合大量信息并运用跨学科知识来解决复杂的现实问题 (Nadelson & and Seifert, 2017)。STEAM 教育的主旨是通过整合各个学科的知识和经验来培养学生的解决问题能力，使他们能够更有效、更灵活地应对复杂的现实挑战。同样可以确定的是，人工智能已被广泛采用和应用于教育领域 (Chen et al., 2020)。特别是，人工智能可以提供个性化支持，使学生能够创造性的解决问题。STEAM 教育和人工智能都需要计算思维，这是数字时代蓬勃发展所必需的核心问题解决技能 (Dolgopolovas & Dagienė, 2021; Tian, 2024)。

2.文献综述

2.1. STEAM

STEAM 教育将科学、技术、工程、艺术和数学五个学术领域的知识有机地结合在一起，培养跨学科思维、解决问题的能力和创造力(魏晓东 et al., 2017)。鉴于计算思维的跨学科性质，STEAM 也是一种跨学科教育，因此计算思维和 STEAM 之间存在着天然的联系。在 STEAM 教学中，学生分解、抽象、建模、设计算法，迭代和优化复杂问题以解决它们，从而培养学生解决问题的计算思维。目前，以往的研究聚焦于培养学生计算思维的效果，但往往忽视了对他们的计算思维行为模式。

2.2. 教育中的人工智能

人工智能（AI）越来越多地融入 STEAM 教育，为提高学习效果提供了新的可能。人工智能可以作为一个教育支架，促进人工智能辅助的以人为中心的理解，提高学生的人工智能素养(How & Hung, 2019a)。人工智能提供了多样化的学习方法，能够有效地分析学生的表现，并促进反馈(魏晓东 et al., 2017)。将人工智能融入 STEAM 教育旨在培养高阶思维、创造力和解决问题的能力(Chang & Chou, 2020; Gasnaş & Globa, 2023)。虽然人工智能在加强 STEAM 教育方面显示出良好的前景，但仍需要进一步的研究来挖掘其全部潜力，并解决在复杂的教育系统中整合各种人工智能技术的挑战(Xu & Ouyang, 2022)。

2.3. 计算思维

计算思维是一系列思维活动，应用计算机科学的基本概念来解决问题、设计系统和理解人类行为，涵盖计算机科学的各个领域。2006 年，Jeannette Wing 引入了计算思维的新概念，强调它不仅限于计算机科学领域，而且是所有学科解决问题的关键技能(M, 2006)。计算思维被视为适应未来所需的一项重要能力，越来越多的研究关注将教育与计算思维整合 (Hsu et al., 2018)。在 K-12 教育中发展计算思维一直是一个主要焦点，努力将其融入现有课程并创造独特的计算思维体验(Weintrop et al., 2023)。已有研究表明，使用 Arduino 进行物理计算活动可以显著提高基础学校学生的多维度计算思维能力(Juškevičienė et al., 2021)。然而，大多数计算思维研究都集中在对学生的编程或计算技能的评估上 (Tang et al., 2020)，但编程和计算能力不会显著迁移到其他领域(Kurland et al., 1986)。这种情况可能会削弱学生在跨学科背景下全面解决问题的能力，甚至阻碍他们发展成为具有多种能力的全面发展的人(Tej, n.d.)。

虽然现有研究强调了将人工智能融入 STEAM 教育的潜力，但对学生计算思维行为模式的探索有限，为进一步研究留下了空间。近年来，人工智能在 STEAM 教育中的融合引起了广泛关注。研究表明，人工智能在 STEM 教育中的应用可以提高学习成果并满足复杂的教育需求 (Xu & Ouyang, 2022)。目前，随着人工智能深度融入 STEAM 课程培养学生计算思维，我们也面临如何准确把握学生思维发展过程、如何组织教学活动的挑战。因此，我们引入滞后序列分析，希望能够识别学生行为序列中计算思维各要素的呈现顺序和相互关系，探索人工智能在 STEAM 课程中提升小学生计算思维行为模式的特点，旨在为教育实践提供更有效的指导和改进策略。

3.方法

3.1. 背景和参与者

本研究来源于一项志愿服务。研究人员选择了一所资源匮乏的小学，这所学校位于大城市旁的乡镇地区。由于就近招生政策，大多数在校学生是这些城市低收入农民工的孩子。截至 2023 年，该校教职员工老龄化，教师人数不足，被市教育局列为薄弱学校。教师的平均

年龄在 50 岁以上，只有一名以前教语文的教师兼任科学教师，因此无法开展高质量的科学课程。

在这个项目中，来自一所大型研究型大学的五名职前科学教师在这所学校设计并教授了一系列以“水净化”为主题的 STEAM 课程。志愿项目每周一节课，持续六周，每周 40 分钟，40 名五年级学生参加了该 STEAM 课程。所有 40 名学生和 5 名职前教师都自愿参加该项目。

3.2. 编码

许多学者指出，计算思维抽象、分解、模式识别、算法设计、迭代和优化。本研究以 S-T 模型为基础，将研究人员根据计算思维的定义(Selby & Woollard, 2013)与 STEAM 教育的特点相结合，开发了一个编码系统来分析与计算思维相关的师生行为。该编码系统不包括 AI，如表 1 所示。

表 1 师生计算思维行为编码表

分类	编码	行为	描述
教师行为（T）	T11	提问	根据问题情境，向学生提问
	T12	引导	引导学生分析问题
	T21	演示	通过多媒体演示
	T22	示范	通过实物操作，给学生做演示
	T23	指示	针对分解而来的问题提供详细具体的指导，帮助学生解决问题
	T31	组织交流	组织学生展示成果
	T32	点评	点评学生作品
	T4	总结归纳	总结归纳学习内容
	T5	监督控制	教师管理监督课堂
学生行为（S）	S1	明确问题	清楚问题情境，描述问题
	S2	分解	对问题进行分解
	S3	抽象	将具体的问题抽象化成应用方案（算法或模型）
	S4	算法	设计和应用方法、方案解决问题
	S5	调试	探究方法是否可行、运行和调试程序
	S6	评估	评价现有的本小组工作，便于进一步优化
	S7	概括	通过发言、画思维导图等方式概括学习过程和所得
	S8	交流	展示和交流本小组的工作
	S9	无关行为	发呆、吵闹、玩耍等和学习无关的行为

3.3. 数据分析

本研究使用滞后序列分析（LSA）来检验 STEAM 课堂中师生的计算思维行为模式。LSA 是一种基于统计理论研究行为之间顺序关系的方法(Bakeman & Gottman, 1997)。这种方法最初由 Sackett 于 1978 年提出，已被心理学和社会学用于分析行为模式，主要是评估一种行为跟随另一种行为的概率，并确定这种关系是否具有统计学意义。这种方法使人们能够探索和总结在复杂的交互式行为序列中发生的交叉依赖关系(Faraone & Dorfman, 1987)，并且可以将行为序列中的重复行为组织成一个关键行为序列(Pohl et al., 2016)。本研究收集了五节课堂的视频记录。视频每隔一分钟采样一次。两名研究人员从视频观察中独立编码了教师和学生的行为，然后对编码结果进行了一致性检查。最终的 Kappa 系数为 0.85，表明编码器间的可靠性非常强。在最终编码之后，根据 LSA，使用 GSEQ5.1 分析行为样本(Bakeman &

Quera, 2011)，并获得相应的频率表和残差值表。通过计算每个序列连接的残差值表中的 z 分数，研究人员确定了统计显著性，z 分数高于 1.96 表示序列具有显著性（p<0.05）(Mei et al., 2024)。

4.结果

研究人员每隔一分钟对视频进行一次采样，共得到 798 个行为样本。学生行为发生 300 次，占总数的 37.6%，教师行为发生 498 次，占 62.4%。在学生行为中，最常见的行为是 S2（分解）、S4（算法）和 S7（总结），共出现 172 次，占所有学生行为的 57.3%。在教师行为方面，最常见的是 T11（提问）和 T23（指导），分别出现了 90 次和 77 次，约占所有教师行为的 33.5%。基于残差值表，创建了 STEAM 课堂中教师和学生计算思维的行为模式图，如图 1 所示。根据行为频率分析，本研究深入探讨了教师和学生在 STEAM 课堂上的计算思维行为模式。特别的是，我们发现了三个具有显著价值的行为序列。

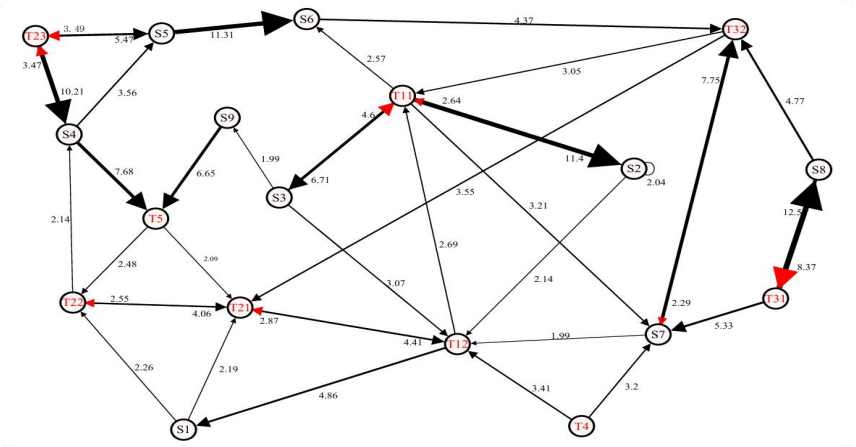


图 1 师生计算思维行为模式图

4.1. 解构问题序列

图 2 表示，学生分解问题之前的许多行为都指向教师主导的行为（T12），然后指向学生的分解问题（S1）。在教师引导下，教师根据问题情境向学生提出有针对性的问题，学生解构问题或使用人工智能抽象、概念化问题（T12→T11→S2，T12→T11→S3）。此时，教师在解构和抽象初始问题（S2→T12→S1，S3→T12→S1）的基础上，进一步引导和帮助学生进一步澄清问题，理解问题情境。此外，教师的多媒体演示和问题总结也有助于他们的指导（T21→T12，T4→T12）。

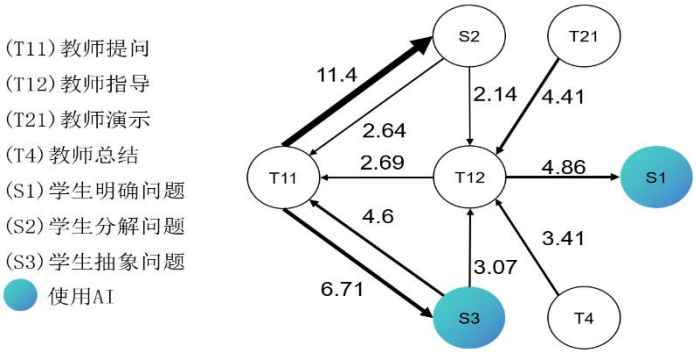


图 2 解构问题行为模式图

4.2. 设计算法序列

图 3 表示，当学生设计算法和调试时，在收到老师的指示（T23）后，学生-老师的行为中有两个重要的行为序列，一个是（S5→S6），另一个则是（S4→T5）。设计算法和评估

改进对小学生来说是一项具有挑战性的学习活动，他们经常感到困惑，需要教师指导。当学生设计算法时，教师可以提供详细而具体的指导，帮助他们基于分解的问题设计算法。学生将自己的知识与教师的指导相结合，设计算法。设计算法的过程要求教师扮演组织者和提醒者的角色，将学生的讨论带回解决问题的背景（T23→S4→T5）。当学生调试和评估算法时，教师可以对问题的分解提供指导，帮助他们探索该方法是否可行，并评估当前的小组工作。学生还可以向人工智能寻求改进建议，以帮助评估当前的算法。人工智能反馈后，学生将进行修改和评估，然后给出反馈以优化现有的算法设计（T23→S5→S6）。

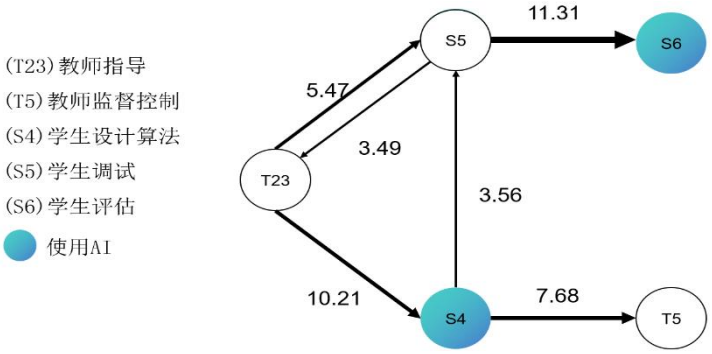


图 3 设计算法行为模式图

4.3. 总结归纳序列

图 4 表示，当学生进行总结时，教师可以提供教学支架和评论来帮助他们进行总结。在课程结束时，一方面，老师通常会组织学生总结他们所学的内容，学生通过画一张思维导图或回答老师的问题进行总结。老师根据他们的答案做出进一步的点评，突出关键信息（T31→S7→T32）。另一方面，教师不断组织学生交流，帮助学生更有效、更合理地与其他小组（T31→S8，S8→T31）进行交流学习。学生的概括往往并不完善，需要教师的点评，教师需要评估他们的回答并给出反馈（S8→T32）。

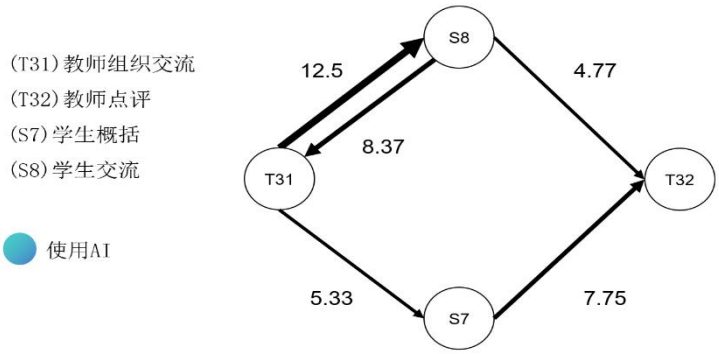


图 4 总结归纳行为模式图

5.讨论

本研究探讨了在一所资源匮乏的学校中，教师和学生在人工智能集成到 STEAM 课程中的计算思维行为模式。LSA 结果揭示了三个关键的行为序列，解构问题，设计算法，总结归纳过程过程，它们恰好对应本课程中基于问题的学习过程的三个阶段。

第一个关键的计算思维行为序列是分解问题。教师的提问和学生对问题的抽象是双向关联的行为（T11→S3），这种情况类似于相关研究中提出的人工智能执行细化。人工智能可以帮助学生通过各种感知方式理解抽象概念，提供听觉和视觉信息的替代品，帮助他们更直观地理解课程内容(Pertusa Mirete, 2023)。抽象（S3）是指学生使用人工智能来抽象化概念。虽然抽象和精炼是相反的，但它们都是学习者学习知识的过程 (Gentner & Hoyos, 2017)。另

一方面，我们可以验证抽象知识是否是 STEAM 课程需要为学生提供的知识。教师的提问与学生的分解行为密切相关（T11→S2，11.4），没有使用人工智能。它也与大卫·乔纳森的理论(Jonassen, 2011)和基于问题的学习原则密切相关。虽然人工智能有助于解构问题，但人类需要对解构实际问题的决定负责(Ninaus & Sailer, 2022)。通过这种方式，学生可以培养解决问题和批判性思维的能力。

二是算法设计行为序列。学生在算法（S4）和评估（S6）中使用了人工智能。在设计算法时，学生不仅在老师的指导下逐步解决问题，而且人工智能学生提供了脚手架。这也与先前的研究结果的论点相一致，人工智能分析可以用作培养 STEAM 学习者人工智能思维技能的教育支架，从而提高解决问题的能力(How & Hung, 2019b)。

三是总结归纳行为序列。研究表明，教师的指导仍然是一种重要的工具。老师指导学生总结他们从活动中学到的知识，并根据学生的回答帮助他们进一步复习知识。这与 STEAM 的课程目标相一致。STEAM 教育的目标是加强个别学科的学习，以产生超越任何单一学科的新理解和成就(Peppler & Wohlwend, 2017)。另一方面，如背景所述，这项研究是在一所资源匮乏的学校进行的。学生缺乏参与协作学习的机会。从这个意义上讲，教师的指导和支持是必要的。学生可以获得合作和交流的机会，从而提高他们的沟通和合作技能。

6. 总结和意义

基于问题的学习（PBL）应该被认为是将人工智能整合到 STEAM 学习中的有效方法。具体来说，教师应该引导学生分解问题，学生使用人工智能通过回答子问题来补充领域知识。在此之后，学生应用所获得的知识来解决原始问题。人工智能帮助学生获得可能欠缺的知识，学生应该积极地将其与自身的推理相结合。这种方法可使人工智能支持学生学习，同时使学生保持积极的角色，促进对问题更深入的理解。避免陷入依赖人工智能的经典误区“只有成果，没有成长”。

此外，教师明确和有针对性的指导至关重要。小学生在设计算法的过程中面临理解和操作方面的挑战。教师不仅是知识的传递者，也是学习活动的组织者和解决问题的促进者。通过具体和情境化的指导，教师可以帮助学生在复杂的任务中找到切入点，并通过问题驱动的方法逐渐掌握算法设计的核心逻辑。人工智能可以作为学生选择问题解决方案的辅助工具，并对这些解决方案进行及时评估。然而，具体的调试过程仍然依赖于教师的明确指导。总之，明确教师和人工智能在算法设计中的角色至关重要。教师的组织和指导能力，结合人工智能的及时评估支持，共同为学生在算法设计方面提供强有力的支持。

7. 限制与未来研究

本研究有两个局限性。首先，在学习成果方面，本研究主要关注 STEAM 课程中的计算思维行为，但缺乏知识测试成绩或绩效评估等实证数据。此外，该研究是在小学阶段的 STEAM 环境中进行的，研究结果可能不适用于其他教育背景。因此，未来的研究可以探索为 STEAM 课程设计的更专业的人工智能工具的应用，这些工具可以在算法设计的不同阶段提供个性化支持。此外，未来研究可扩大研究范围，涵盖不同教育阶段、不同地区和不同学科的 STEAM 课程，探究研究结果在其他教育背景下的适用性。

参考文献

魏晓东, 于冰, 于海波. 美国 STEAM 教育的框架、特点及启示. 华东师范大学学报(教育科学版), 2017, 35(4): 40-46,134,135.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis, 2nd ed* (pp. xiii, 207). Cambridge University Press. https://doi.org/10.1017/CBO9780511527685

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences* (pp. xv, 183). Cambridge University Press. https://doi.org/10.1017/CBO9781139017343

Chang, Y.-S., & Chou, C. (2020). Integrating Artificial Intelligence into STEAM Education. In J. Shen, Y.-C. Chang, Y.-S. Su, & H. Ogata (Eds.), *Cognitive Cities* (pp. 469–474). Springer. https://doi.org/10.1007/978-981-15-6113-9_52

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Dolgopolovas, V., & Dagienė, V. (2021). Computational thinking: Enhancing STEAM and engineering education, from theory to practice. *Computer Applications in Engineering Education*, *29*(1), 5–11. https://doi.org/10.1002/cae.22382

Faraone, S. V., & Dorfman, D. D. (1987). Lag sequential analysis: Robust statistical methods. *Psychological Bulletin*, *101*(2), 312–323. https://doi.org/10.1037/0033-2909.101.2.312

Gasnaş, A., & Globa, A. (2023). Interacțiunea dintre inteligența artificială și conceptul STEAM. *Proceedings of The Third International Scientific Conference "Inter/Transdisciplinary Approaches in the Teaching of the Real Sciences, (STEAM Concept)" = Materialele Conferinţei Ştiinţifice Internaţionale "Abordări Inter/Transdisciplinare În Predarea Ştiinţelor Reale, (Concept STEAM),"* 313–320. Inter/transdisciplinary approaches in the teaching of the real sciences, (STEAM concept) = Abordări inter/transdisciplinare în predarea ştiinţelor reale, (concept STEAM). https://doi.org/10.46727/c.steam-2023.p313-320

Gentner, D., & Hoyos, C. (2017). Analogy and Abstraction. *Topics in Cognitive Science*, *9*(3), 672–693. https://doi.org/10.1111/tops.12278

How, M.-L., & Hung, W. L. D. (2019a). Educing AI-Thinking in Science, Technology, Engineering, Arts, and Mathematics (STEAM) Education. *Education Sciences*, *9*(3), Article 3. https://doi.org/10.3390/educsci9030184

How, M.-L., & Hung, W. L. D. (2019b). Educing AI-Thinking in Science, Technology, Engineering, Arts, and Mathematics (STEAM) Education. *Education Sciences*, *9*(3), Article 3. https://doi.org/10.3390/educsci9030184

Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education*, *126*, 296–310. https://doi.org/10.1016/j.compedu.2018.07.004

Jonassen, D. (2011). Supporting Problem Solving in PBL. *Interdisciplinary Journal of Problem-Based Learning*, *5*(2), Article 2. https://doi.org/10.7771/1541-5015.1256

Juškevičienė, A., Stupurienė, G., & Jevsikova, T. (2021). Computational thinking development through physical computing activities in STEAM education. *Computer Applications in Engineering Education*, *29*(1), 175–190. https://doi.org/10.1002/cae.22365

Kurland, D. M., Pea, R. D., Clement, C. A., & Mawby, R. (1986). A study of the development of programming ability and thinking skills in high school students. *Journal of Educational Computing Research*, *2*(4), 429–458. https://doi.org/10.2190/BKML-B1QV-KDN4-8ULH

M, W. (2006). Computational thinking. *Communications of the ACM*. https://doi.org/10.1145/1118178.1118215

Mei, A., Long, T., & Liu, Z. (2024). How blue hat co-regulate pre-service teachers' collective reflection based upon the six-hat thinking technique: A lag sequential analysis. *Thinking Skills and Creativity*, *53*, 101626. https://doi.org/10.1016/j.tsc.2024.101626

Nadelson, L. S., & and Seifert, A. L. (2017). Integrated STEM defined: Contexts, challenges, and the future. *The Journal of Educational Research*, *110*(3), 221–223. https://doi.org/10.1080/00220671.2017.1289775

Ninaus, M., & Sailer, M. (2022). Closing the loop—The human role in artificial intelligence for education. *Frontiers in Psychology*, *13*, 956798. https://doi.org/10.3389/fpsyg.2022.956798

Peppler, K., & Wohlwend, K. (2017). Theorizing the nexus of STEAM practice. *Arts Education Policy Review*, *119*(2), 88.

Pertusa Mirete, J. (2023). Inteligencia artificial aplicada a la educación: El futuro que viene. *Supervisión 21: revista de educación e inspección*, *69 (Julio)*, 3.

Pohl, M., Wallner, G., & Kriglstein, S. (2016). Using lag-sequential analysis for understanding interaction sequences in visualizations. *International Journal of Human-Computer Studies*, *96*, 54–66. https://doi.org/10.1016/j.ijhcs.2016.07.006

Selby, C., & Woollard, J. (2013, January 12). *Computational thinking: The developing definition*.

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, *148*, 103798. https://doi.org/10.1016/j.compedu.2019.103798

Tej, T. of E. J. (n.d.). *The effect of training Six Thinking Hats method on improving cognitive skills of agricultural students in entrepreneurship course*. Retrieved April 6, 2025, from https://www.academia.edu/44104322/The_effect_of_training_Six_Thinking_Hats_method_on_improving_cognitive_skills_of_agricultural_students_in_entrepreneurship_course

Tian, S. (2024). Exploring the Interactive Relationship between Computational Thinking and the Development of Artificial Intelligence. *Journal of Education and Educational Research*, *9*(2), 177–180. https://doi.org/10.54097/2zzt5026

Weintrop, D., Rutstein, D. W., Bienkowski, M., & McGee, S. (Eds.). (2023). *Assessing Computational Thinking: An Overview of the Field*. Routledge. https://doi.org/10.4324/9781003431152

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, *9*(1), 59. https://doi.org/10.1186/s40594-022-00377-5

人智协同的跨学科知识创生双链螺旋模型研究

Research on the Interdisciplinary Knowledge Creation Model of Human-Intelligence Synergy

梁云真，谢慧全，张永乐
河南师范大学 教育学部，河南 新乡 453007
1462386472@qq.com

【摘要】 跨学科学习是当前教育改革新趋势，智能体的发展与融入使得学习过程更加深入和创新。作为适应时代发展的新模式，人智协同的跨学科学习已然成为新知识创造及复杂问题解决的新模式。本研究通过对当前研究的分析，基于知识创生理论提出涵盖知识学科属性和知识主体演变的“人智协同的跨学科知识创生双链螺旋模型”，实现学习者个体隐性知识的迭代创生，为数智时代的教育教学提供兼具理论深度与实践创新的人智协同学习新范式。

【关键词】 人智协同；跨学科学习；知识创生

Abstract: *Interdisciplinary learning is a new trend in educational reform. The development and integration of AI have made the learning process more profound and innovative. As a new model, human-AI collaborative interdisciplinary learning has become a new model for creating knowledge and solving complex problems. Through the analysis of current research and grounded in knowledge creation theory, this study proposes a "human-AI collaborative interdisciplinary double-chain knowledge creation spiral model", which covers the disciplinary attributes of knowledge and the evolution of knowledge subjects. This model aims to achieve the iterative creation of individual learners' tacit knowledge and provide a new paradigm of human-AI collaborative learning with both theoretical and practical for education and teaching in the digital intelligence era.*

Keywords: human-intelligence collaboration, interdisciplinary learning, knowledge creation

1.引言

智能技术发展使教育和课堂迎来变革，学习主体从人类个体发展到人机交互，再迈向人智协同（张进澳等人，2024）。人智协同中，智能体生成性检索机制与人类认知构建流程相似，旨在协同合作、深层次交互，共同构建与优化知识架构（裘江南等人，2025）。我国《义务教育课程方案（2022 年版）》强调课程综合、关联性等原则。在此背景下，跨学科主题学习备受关注，它以问题解决为核心，打破学科壁垒，综合多学科知识方法解决实际问题，能促进知识创造、提升学习者能力、推动核心素养发展。

教育数字化转型倡导人类学习者与智能体携手促进跨学科深度学习。人智协同是教育数字化变革的核心驱动力，为学习提供新方式，课程标准中的跨学科主题学习提供了契机（周晓慧，2024）。在跨学科学习中借助人智协同，能整合多学科知识，发挥综合育人作用，实现知识跨界融合与创新创造，提升学生批判思维、跨学科能力等高阶思维能力（尹志华，2024）。基于此，本研究以知识创生理论（SECI）为支撑，建构人智协同跨学科学习中的知识创生模型，旨在丰富智能体在知识创新范畴的研究，为提升人智协同在知识创新方面的效能提供导向性建议。

2.文献综述

2.1.知识创生的理论模型

“创生”一词最早出现在 20 世纪 70 年代西方学者的课程实施中，1992 年，美国的辛德尔等人进一步将课程实施的取向归纳为：忠实取向、调适取向和创生取向。此后，“创生”被教育界广泛使用（汪晓凤等人，2014）。1995 年，日本学者野中郁次郎提倡要挣脱知识本身的枷锁，利用“默会知识”（隐性知识）与“明言知识”（显性知识）交互所构成的螺旋式扩张方式，来说明知识的创生过程。两种知识之间的相互转化分为四个阶段:社会化、表征化、联结化、内隐化，简称为 SECI 模式,其前提条件是人与人在特定的共享语境中进行持续的、相互依赖的互动。四种延续联结的状态说明了知识历经个体、团队、组织三个层面最后得以创生的全过程（竹内弘高等人，2006）。

2.2.人智协同的知识创生

人智协同作为目前知识创生的新型模式，主要强调多主体之间的知识互动关系。智能体的运用促使知识创生的主导力量发生了转变，从原本人类个体主导模式，逐渐演变为以人智融合系统为核心的新范式（米加宁**等人**，2024）。从本质的逻辑视角深刻审视，人智协同支持的知识创新活动，是依托于知识库及规则事实逻辑构建的“知识体系”而得以展开的。这一知识体系不仅为知识表征提供了坚实基础，还通过演绎推理的方式，实现了知识的有效传递与深化（吴飞**等人**，2020）。从知识创造对象的维度深入观察，这种合作状态的深入发展进一步延伸了知识创生的边界，一轮协作创生出的知识已成为了下一阶段交流互动的知识基础（袁一鸣**等人**，2024）。基于众多理论探讨研究，技术支持下的实践应然而生，以阴阳辩证系统性知识创造理论作为底层逻辑，探讨跨文化应急管理视域下人智协作进行知识创造的过程，并通过仿真实例验证提出的系统架构,从而验证了理论模型的有效性（覃大嘉**等人**，2024）。

3.人智协同的跨学科知识创生双链螺旋模型构建

依据知识创生理论（SECI），从知识主体和所属学科演变的两个视角，构建“人智协同跨学科知识创生双链螺旋模型”，如图 1。

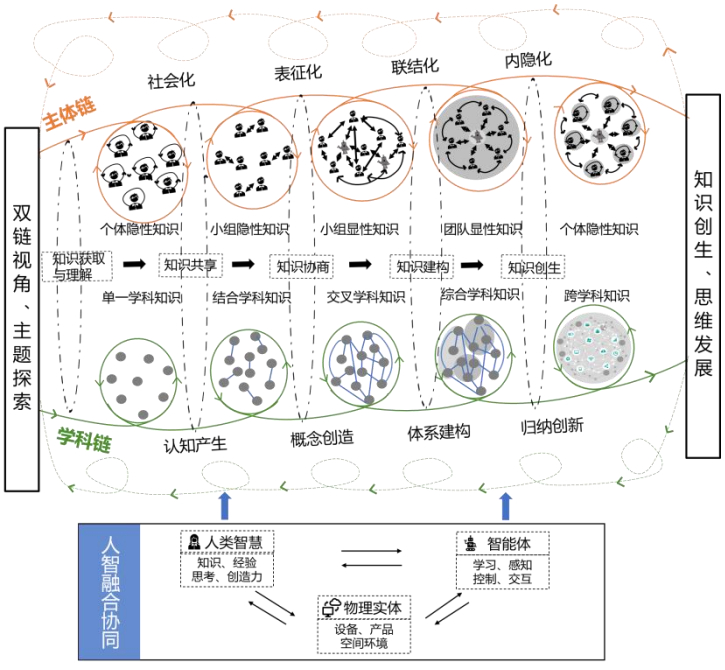


图 1 跨学科知识创生双链螺旋模型

人智协同的跨学科学习以解决现实问题为起点，团队成员、学习者与智能体相互交流协作，共同完成知识查找、分析等思维活动，联结、融合、内化等主体“知识经验库”（Vandenberg et al.,2021），体现智能化“情境搜索”特性。跨学科学习过程中，学习者对知识的掌握不断丰富、突破和创生，认知结构逐步建构与改变，有积累也有飞跃（Pohl C.,2011）。跨学科知识创生是复杂动态螺旋过程，可分为五个阶段：基于问题获取知识、团队合作共享知识、碰撞观点协商知识、任务探究建构知识以及实践反思创生知识。

3.1.知识获取与理解

知识创生的首要前提在于提供可迁移的知识资源。基于问题的知识获取，是指学习者依据跨学科主题及其要解决的问题，识别所涉及到的学科，通过查阅资料、独自思考等方式从外界获取知识资源，既包括个体经历、体验与收集过程中所获得的直接性、观念性、情感性等内容，又包括学科知识内容。学习者将从智能体接收到的知识，经过自己的消化理解，确立到已有的认知框架中，同时有效地组织和管理它们，实现信息到知识的进阶（周可等人，2023）。

3.2.知识共享

知识共享是知识在个体间的转移和流动（Holtshouse D，1998），重点关注主体间知识交换的动态过程（CHAI et al.,2011）。人智协同跨学科学习中的知识共享，是团队成员分享体验、观点、心智模式，将各学科的私有知识转化为组织的共有知识，为解决问题提供思路，促成知识资源在团队内共享、互补与增值利用。这一过程是团队创新创造的基础，涉及知识传递与吸收，参与者分享的同时也接收知识，实现知识的共同掌握，不会导致分享知识的人失去对知识的所有权（Ipe M,2003）。人智协同的跨学科知识共享并非单向扩散，而是各参与主体间双向或多向传递，形成紧密有序扩大的知识网络 。

3.3.知识协商

沟通协作是知识创生这个多主体参与、多个环节构成的复杂活的必要环节之一。人智协同支持的跨学科知识协商，是学习者个体与小组内其他主体或大模型支持的智能体，在物理或虚拟空间内就冲突观点展开的互动，以实现观点联结与深化，包括同伴对比、讨论及对话交流。此过程本质是观念、思想、知识的碰撞，反映团队知识耦合，强调提升知识协同效应，实现知识创新与增值。学习者与智能体的协同，能将思维方式、经验及内在智慧转化为清晰文本，使隐性知识明确可呈现。

3.4.知识建构

知识建构是在原有知识基础上通过推理与决策的智力活动，协商拓展出新知识、达成深度共识、扩充知识体系的过程。人智协同跨学科学习中，人类学习者与智能体协同，将知识分模块归整，链接新创知识与既有知识，并将其固化到组织框架，扩展知识网络，形成知识创造系统场，以适应解决复杂问题的需要。跨学科知识构建强调深层次知识构建，要求知识展现更紧密的关联性，体现广泛而深入的融合特性。此阶段重点在于发现和建立知识联系，利用智能体整合关联多学科信息，构建高级知识体系，实现知识深度组合与广度融合，增强整体价值。

3.5.知识创生

创生是知识生成最高层次与阶段。跨学科知识创生是学习者基于跨学科主题学习任务，在协作中将理论用于实践，整合各个阶段、维度、场域所获得的知识，通过“做中学”创造新知识，实现知识结构横向扩展与纵向延伸，拓展学习者的思维认知，助其解决复杂问题。学习者与智能体协作，综合渗透多学科的知识，深入知识创造场域，从真实情境建构认知，

扩大个体知识并更新知识结构。此过程不仅促进了认知、反思与创新能力的提升，还触发了新一轮知识螺旋。

4.结论

本研究基于跨学科学习中知识创生这一视角，从 SECI 理论出发，引入学习者、学习同伴与智能体的三元主体共创模式，从知识学科属性和知识主体演变两个维度，构建了“人智协同的跨学科知识创生双链螺旋模型”，从而全面地描述学习者剖析跨学科主题、团队合作共享知识、观点交流碰撞、任务深入探究与实践反思实现内化的知识创生过程，为**数智时代知识创生**提供了兼具理论深度与实践创新的新型范式。未来需进一步聚焦人智协同跨学科学习的真实应用，探索实际教育教学改革中，人智协同跨学科课程设置、教学方法和学习评价的优化策略。根据知识创生的过程和规律，设计跨学科课程，培养学生的综合素养和创新能力，建立多元化的学习评价体系，全面评估学生的知识创生能力和综合素质。

参考文献

尹志华,练宇潇,贾晨昱 & 牛晓.(2024).人工智能融入体育与健康跨学科主题教学的框架构建与推进策略.成都体育学院学报,50(05),16-26.

米加宁 & 董昌其.(2024).大模型时代：知识的生成式“涌现”.学海,(01),81-96+214-215.

竹内弘高，野中郁次郎.(2006).知识创造的螺旋：知识管理理论与案例研究.北京：知识产权出版社.

汪晓凤,陈玲 & 余胜泉.(2014).基于实践性知识创生的网络教研实证研究.中国电化教育,(10),16-22.

吴飞 & 段竺辰.(2020).从独思到人机协作——知识创新模式进阶论.浙江学刊,(05),94-104.

张进澳,卢新元,孙冰悦 & 王一洲.(2024).人智协同知识创新机制的建模与仿真——基于模拟实验的系统动力学分析.情报理论与实践,47(12),143-153+96.

周可,孙溪梓 & 李佳桐.(2023).社会网络视角下组织间知识重叠对知识创造的影响机理研究.情报科学,41(08),162-167+183.

周晓慧.(2024).基于人工智能的初中物理跨学科实践教学设计——以“搭建风力发电结构系统”为例.物理教师,45(10),32-36+40.

袁一鸣,陶成煦,贺超城 & 吴江.(2024).人智交互情境下的知识转化模型：内涵深化及外延拓展.情报理论与实践,47(06),76-84+90.

覃大嘉,章质彬,成维华 & 汪寿阳. (2024).基于辩证思维的跨文化应急救援人机融合决策模型研究.系统工程理论与实践,1-33.

裘江南,高双燕,黄榕江 & 李梦洁.(2025).AIGC 背景下人智认知导向的信息检索范式——基于世界 3 理论的探究.中国图书馆学报,51(01),34-48.

CHAI, S., DAS, S., & RAO, H. R. (2011). Factors affecting bloggers'knowledge sharing:an investigation across gender. *Journal of management information systems, 28*(3), 309-342.

Holtshouse, D. (1998). Knowledge research issues. *Journal of California management review, 40*(03), 277-280.

Ipe, M. (2003). Knowledge sharing in organizations:A conceptual framework. *Journal of Human resource development review, 2*(04), 337-359.

Pohl, C. (2011). What is progress in transdisciplinary research?. *Journal of Policy, Planning and Futures Studies, 43*(6), 618-626.

Vandenberg, J., Zakaria, Z., & Tsan, J. (2021). Prompting collaborative and exploratory discourse:An epistemic network analysis study. *International Journal of Computer-Supported Collaborative Learning,* (16), 339-366.

GAI 时代提示词效能感和教学自我效能感对职前教师 TPACK 水平的影响研究

The Effects of Prompt Word Efficacy and Teaching Self-efficacy on the TPACK Level of Pre-service Teachers in the GAI Era

马笑颜，杨阔，张慕华*
首都师范大学初等教育学院
zhangmuhua@cnu.edu.cn

【摘要】 随着数字化转型的逐步推进，TPACK 成为当下教师必备的关键能力。目前，生成式人工智能的迅速发展对教师 TPACK 能力提出新要求，即具有应用先进科技的能力，并将技术与教学法相融合并应用。已有研究表明自我效能感是教师 TPACK 的重要影响因素，在 GAI 背景下，提示词是技术应用的重要载体，影响着技术掌控的效果，而在教学层面，教师的教学自我效能感也是重要变量，影响教学的效果。基于此，本文聚焦 TPACK 理论，探讨个体差异与提示词自我效能感和教学自我效能感对职前教师 TPACK 水平的影响，进而提出 GAI 时代职前教师的培养策略，为当下职前教师培养提供相应的参考。
【关键词】 生成式人工智能；职前教师；TPACK；提示词效能；教学自我效能感

Abstract : *With digital transformation advancing, TPACK has become a key competency for teachers. The rapid development of generative AI (GAI) sets new demands: teachers need to apply advanced technology and integrate it with pedagogy. Existing research reveals that self-efficacy significantly influences teachers' TPACK. In the GAI context, prompt words critical carriers for technology application affect the effectiveness of technology control. Meanwhile, teaching self-efficacy at the instructional level acts as another vital variable impacting teaching outcomes. Focusing on the TPACK theory, this paper explores the effects of individual differences, prompt-word self-efficacy, and teaching self-efficacy on pre-service teachers' TPACK levels. Consequently, it proposes cultivation strategies for pre-service teachers in the GAI era, providing references for current pre-service teacher training.*
Keywords: generative AI，pre-service teachers，TPACK，prompt word efficacy，Teaching efficacy

1.引言

自 2022 年国务院颁布《“十四五”数字经济发展规划》，推进教育数字化转型以来，我国更加注重教师数字化能力的培养，着力构建技术支持的教育教学新模式。目前，TPACK(整合技术的学科教学知识)能力已然成为教师必备的综合能力，TPACK 涵盖学科内容、教学法知识和技术知识，其中每个字母的代表着不同的含义，相互组合衍生出多个维度，TK(Technology Knowledge)是指技术知识、TPK(Technological Pedagogical Knowledge)是指整合技术的教学法知识、TCK(Technological Content Knowledge)是整合技术的学科内容知识、PCK（Pedagogical Content Knowledge)是具体学科内容教学的教学法知识、TPCK(Technological Pedagogical Content Knowledge)则是整合技术的学科教学知识。近年来，TPACK 框架更加丰富，将境脉因素纳入其中，是对 TPACK 的认知和水平产生影响的多种影响因素的统称，自我效能感就是境脉因素之一。随着现代科技的日新月异，GAI(Generative artificial intelligence，生成式人工智能)走入人们的日常生活，以其强大的功能和便捷性助力

教学变革，同样对教师的 TPACK 能力提出新要求，希望教师将 GAI 技术与教学相融合，具有更强的技术自我效能感，进而更好地助力学生发展。在 GAI 使用的过程中，提示词是连接用户与大语言模型的桥梁，合理设计与优化提示词将会促进 AI 精确识别，并形成与之匹配的内容，获得较好的感知效果，因此教师的提示词效能感不可忽略。提示词效能是用户对如何书写、修改提示词，引导大语言模型生成所需要的主题信息的自我感知，对教师的 GAI 使用体验与教学应用至关重要，需要持续培养职前教师的相应能力。在 GAI 广泛普及的大背景下，教师不仅需要具备提示词的自我效能感，还需进一步关注当前个人教学的效能感，二者共同作用方能更好的提升教学的效果。教师自我效能感是 GAI 时代教师对自己是否有能力为完成技术支持的教学所进行的推测与判断，也是教师自主发展的重要内在动力（庞丽娟等，2005），影响着教师自己的教学策略选择与教学行为、学生的自我效能感形成等（李红等，2000）。因此，文章聚焦职前教师的 TPACK 能力培养，探讨提示词自我效能感和教学自我效能感对其 TPACK 的影响，为 GAI 时代职前教师的培养提供相应策略。

2.文献综述

2.1. GAI 时代教师 TPACK 研究现状

GAI 时代，国内研究主要聚焦于 GAI 在教育领域的具体应用，如赋能创造性思维的培养、助力教学设计和教学模式的创新、风险挑战等，以及 ChatGPT 使用下“师-机-生”三元结构的智能生成与多元评价（杨宗凯等，2023）。以 ChatGPT、聊天机器人为代表的 GAI 与 TPACK 结合的研究较少，仅有一些对 AI-TPACK 模型的构建与本质解读，提出 AI-TPACK 更加注重人机协同思维，倡导通过技术映射的方式提升教师的 AI-TPACK 水平的发展（闫志明等，2020）。已有研究还表明 AI-TPACK 的本质是将人工智能的技术与思维运用于具体学科知识的教学实践之中，更加注重教师的行动能力（虞江锋等,2022）。

国外研究以 AI-TPACK 量表的探索为主，部分研究提出 GAI 与教师 TPACK 能力的相互影响，多数研究仍聚焦于人工智能与 TPACK 结合的研究。实证研究发现：教师的 TPACK 会显著正向影响其使用其他技术的行为意向（Bardakci & Alkan，2019），但这种影响会被不适和焦虑等负面情绪所削弱，即随着 TPACK 能力的提升，高度的负面情绪会使教学实践中生成式人工智能的使用意愿减弱（Yang et al.,2023）。另外，与 AI 工具交互的知识水平影响对 AI 的教学贡献的理解；技术知识（TK）可以帮助教师更好地评估 AI 的决策，但还需要 TPACK 中其他能力的辅助（Celik,2022)。已有研究还针对在职教师人工智能教育方面的 TPACK 水平进行探讨，发现小学教师在所有七个 TPACK 相关因素上的得分均高于中学教师，教学人工智能的经验对感知 TPACK 有显著影响，但更多的教学经验并不能保证更好地理解和实施 AI 教育（Yue et al，2023）。在职前教师的相关研究中，发现英语教育专业的学生在 TPACK 和 AI-TPACK 组件中都具有很强的基础技能，而发展基础技术技能可以同时增强与人工智能相关的教学能力(Karatas & Aksu Ataç，2024)。目前，对职前教师的 TPACK 水平的探讨仍然聚焦于 AI 时代，较少探讨在 GAI 广泛普及的时代，职前小学教师的 TPACK 现状以及性别、专业方向等因素是否对职前的小学教师的 TPACK 水平产生相应影响。

2.2. 自我效能感与 TPACK 研究现状

国内学者在 TPACK 框架的基础上，提出境脉因子的概念，即对教师 TPACK 知识的习得和应用产生影响的因素，并构建 TPACK 境脉因子模型，证实自我效能感对 TPACK 产生一定影响（刘艳华等，2015）。TPACK 框架中技术知识和学科教学法知识是重要部分，在自我效能感维度，与技术和教学联系紧密的当属技术自我效能感和教学自我效能感。技术自我效能感是指对熟练使用相应技术的能力和信心，也被认为是使用技术意图的主导决定因素

(Teo & van Schaik，2012)。国外学者通过元分析的方法构建随机效应模型，得出信息技术整合的自我效能感与 TPACK 之间存在中等程度的正相关关系（Zeng et.al,2022）。对于本科和研究生两大类的职前教师群体而言，基于问卷调查和 UTAUT 模型，已有研究证实技术自我效能感对 TPACK 有积极影响（Wang et.al,2024）。相比于技术自我效能感，教学自我效能感是指教师对自己影响学生行为和学习成绩的能力的主观影响与态度（李红等，2000）。随着 GAI 的发展，教师教学自我效能感还应包括自身使用技术支持教学的态度与能力。教师的态度和观点是引发教师 TPACK 的重要因素,信念和观点也影响着课程的解释和实施（Velander et al，2022）。早在数字教材背景下，教师的 TPACK 水平已被证实与使用数字教材授课的自我效能感存在间接的正向作用，随着教学自我效能感的提升，教师的技术知识（TK）和技术教学法知识（TPK）会相应提升（宋伟等，2014）,教师进行数字化的教学自我效能感能够正向预测教师的以 TPACK 能力为代表的数字化教学素养（谢志勇等，2024）。当前，部分学者通过 TPACK 相关的教师专业发展项目实践，证明其对在职中小学计算机教师在 AI 教学效能信念和 AI 教学结果期望方面具有积极影响（Junmei Sun，2023）。当前，现有研究仅聚焦于技术自我效能感或教学自我效能感对 TPACK 的影响，未曾将技术自我效能感、教学自我效能感和 TPACK 三者进行结合研究，对三者之间的相关关系研究较少，而且已有研究大多针对单一学科或多学段的职前教师，未曾对大一年级职前教师进行研究，基于其身份的特殊性，更需关注其在以上三个维度的具体情况，进一步探讨并提出培养策略。

2.3. 提示词使用对 GAI 应用的影响

由强大数据库和算法构成的生成式人工智能在教育维度可以成为教学辅助工具，助力教师完成课程设计、教案撰写、课堂教学、事务性工作等，进而给予教师更多时间来关注学生的成长和发展(朱永新等，2023)。国外研究表明，高等教育中使用 GAI 的意愿受到多因素的影响，其中主要因素是使用障碍，其次是感知风险、易用性、有用性和 TPACK（Jain，2023）。相比于其他人工智能应用，生成式人工智能具有一定的简便性，是人类通过发起提示，依据收到的响应，系统地、有目的地应用以实现 AI 的应用，构成新型的人机协作的过程 (Zhao et al., 2021)。生成式人工智能作为先进的技术手段，教师将其应用于教学首先需要对技术有较好的效能感，掌握 GAI 的使用方法，进而结合所教授的内容思考如何将技术与教学相融合，在较高的教学自我效能感的指引下设计教学，最终为学生带来高质量的课堂教学。教师群体使用生成式人工智能不可或缺的桥梁就是提示词，是向 GAI 提供指令的过程，正确使用提示词是信息化时代教师高效应用现代科技的必备能力。进行提示词重构，使得提示词中包含语言风格、说明方法、来源限定以及专业情境的创设，是优化 AI 生成内容效果的关键策略（宋筱璇等，2024）。在 TPACK 角度，教师与 GAI 相关的 TPACK 能力应包括对 ChatGPT 进行持续对话的能力，教师应该知道如何有效地制定提示词，并评估聊天输出，根据需要使用后续提示（Feldman-Maggor，2024）。但目前，已有研究对于职前教师能否正确且高效使用提示词的自我效能感了解较少，需进行相应研究，以提出提高职前教师提示词效能感的策略，从而为未来教学提供支持。

3.研究问题

本研究以 Mishra 和 Koehler（2006）提出的 TPACK 框架作为理论基础。根据该理论和已有研究，TPACK 共包括三个核心要素：TK、CK（Content Knowledge，学科教学知识）、PK（Pedagogical Knowledge，教学法知识）；四个复合要素：PCK、TCK、TPK、TPCK 和境脉因素。本研究注重 GAI 技术支持背景下职前教师的 TPACK 水平，以技术相关的

TPACK 变量：TK、TCK、TPK、TPCK 作为衡量职前教师 TPACK 水平的主要因素，探讨教学自我效能感、提示词效能感对职前教师 TPACK 能力的影响，因此提出以下研究问题。

（1）个体差异是否会对职前小学教师的 TPACK 水平、提示词效能、教学自我效能感产生影响？

（3）提示词效能、教学自我效能感是否与职前小学教师的 TPACK 水平存在显著相关性？

4.研究设计

4.1. 研究对象

本研究以中国北方某师范类大学小学教育专业大一年级 375 名师范生为研究对象，本学期研究对象均参与《人工智能导论》专业必修课的学习，具有一定的人工智能知识储备，多数学生以不同的方式了解生成式人工智能并进行相应体验。研究共计发放问卷 375 份，回收 364 份，回收率为 97.07%，回收有效数据 212 份，有效率为 58.24%。

4.2. 研究工具

为了解师范生对生成式人工智能提示词效能、教学自我效能感的具体看法，探讨二者与师范生 TPACK 水平之间的联系，进一步提出师范生培养的相关建议，本研究依据 Celik(2023)制定的 TPACK 相关问卷、Klassen 等(2009)编制的教学自我效能感问卷以及 Tian 等(2023)学者对提示词的调整策略和效能感制定的相关问题进行改编。问卷由两大部分构成，第一部分是个人基本信息，包括参与者的性别、专业方向、年龄等，第二部分是 TK、TPK、TCK、TPCK 和 Prompting Efficacy（提示词效能）、Teaching efficacy（教学自我效能感）六个维度，共计 28 个题目，均使用七点李克特量表，1 为“非常不同意”、4 为“中立”、7 为“强烈同意”。

5.数据分析

5.1. 统计结果描述

本研究使用 SPSS27.0 软件对数据进行描述性结果统计与分析，得出 TPACK 中与技术相关的四个维度 TK、TCK、TPK、TPCK 和 Prompting Efficacy、Teaching efficacy 两个维度的均值与标准差，并计算 Cronbacha 系数，如下表所示。

表 1 各维度描述性统计

维度	样本	Cronbacha 系	平 均	方差	标准差
TK	212	.927	5.4047	.925	.96152
TPK	212	.947	5.6557	.776	.88070
TCK	212	.922	5.4811	.910	.95404
TPCK	212	.963	5.6215	.797	.89255
Teaching efficacy	212	.938	5.7642	.718	.84759
Prompting Efficacy	212	.956	5.6191	.921	.95993

问卷涉及的六个维度 Cronbacha 系数均在 0.9 以上，表明量表的内部一致性比较高，设计的题目与各维度紧密相关。此外，各维度平均值表明多数大一新生介于有些同意（5=有些同意）到同意（6=同意）之间，标准差和方差数据表明各维度的数据比较集中，说明大一年级师范生对 TPACK 各维度和提示词效能与教学自我效能感的看法较为一致，认同自己能够有效运用提示词并把控好课堂。

5.2. 个体差异对 TPACK 水平的影响

为研究个体差异是否对大一年级师范生的 TPACK 水平、本研究关注性别和专业方向在 TPACK 各维度和提示词效能与教学自我效能感的差异。在性别维度，回收的 212 份有效问

卷中 42 名（19.8%）为男生、170 名（80.18%）为女生。研究使用 Kolmogorov-Smirnov 检验来判断各维度的正态分布情况，得出渐近显著性（双尾）均小于 0.001，不符合正态分布。因此，通过曼-惠特尼 U 检验进行分析，如表 2 所示，说明性别对于技术相关的 TK、TCK、TPK、TPCK 水平以及提示词效能感、教学自我效能感均没有显著差异（P>.05）。

在专业方向维度，基于中国北方某师范类大学小学教育专业的多样性划分，可以分成小学教育下辖文科、理科、艺术三大方向。通过进行克鲁斯卡尔-沃利斯检验，结果如表 3 所示，可以得出不同专业方向在 TPACK 各维度以及提示词效能和教学自我效能感上均没有显著差异。但是在艺术方向在 TK、TPK、TCK、TPCK 维度的平均分均高于其他两个方向。文科方向在 TK 水平上的平均值最低，在其他方面理科方向的平均值均低于其他方向。

表 2 曼-惠特尼 U 检验及均值情况

维度	类别	人数	平均值	方差	标准差	曼-惠特尼	P 值
TK	男生	42	5.3476	1.153	1.07391	3376.500	.584
	女生	170	5.4188	.873	.93461		
TPK	男生	42	5.5170	1.138	1.06669	3276.500	.404
	女生	170	5.6899	.686	.82849		
TCK	男生	42	5.43155	1.284	1.13301	3193.000	.284
	女生	170	5.5221	.816	.90358		
TPCK	男生	42	5.5119	1.101	1.04939	3335.500	.507
	女生	170	5.6485	.724	.85073		
Teaching efficacy	男生	42	5.6429	1.013	1.00630	3327.500	.487
	女生	170	5.7941	.647	.80418		
Prompting Efficacy	男生	42	5.5179	1.233	1.11036	3380.500	.589
	女生	170	5.6441	.848	.92097		

表 3 克鲁斯卡尔-沃利斯 H 检验情况及各学科均值

维度	文科	理科	艺术	克鲁斯卡尔-沃利	P 值
TK	5.3282	5.4123	5.6782	2.659	.265
TPK	5.6847	5.6156	5.7087	.708	.702
TCK	5.5679	5.3711	5.5978	1.882	.390
TPCK	5.6766	5.5309	5.7826	2.282	.319
Teaching efficacy	5.7962	5.7165	5.8370	.588	.745
Prompting Efficacy	5.6894	5.5284	5.6848	1.583	.453

5.3. 教学自我效能感、提示词效能与 TPACK 各维度的相关性分析

本研究以斯皮尔曼系数的方式探讨教学自我效能感和提示词效能对大一年级师范生 TPACK 各维度的相关性，如表 4 所示。

表 4 大一年级师范生斯皮尔曼系数检验情况

维度	斯皮尔曼	TK	TPK	TCK	TPCK	教学自我效	提示词效
TK	相关系数	1.000	.719**	.668**	.630**	.561**	.662**
	P 值	.	<.001	<.001	<.001	<.001	<.001
TPK	相关系数	.719**	1.000	.783**	.743**	.697**	.820**
	P 值	<.001	.	<.001	<.001	<.001	<.001
TCK	相关系数	.668**	.783**	1.000	.783**	.691**	.912**
	P 值	<.001	<.001	.	<.001	<.001	.000
TPCK	相关系数	.630**	.743**	.783**	1.000	.842**	.769**
	P 值	<.001	<.001	<.001	.	<.001	<.001
Teaching efficacy	相关系数	.561**	.697**	.691**	.842**	1.000	.717**
	P 值	<.001	<.001	<.001	<.001	.	<.001

Prompting Efficacy	相关系数	.662**	.820**	.912**	.769**	.717**	1.000
	P 值	<.001	<.001	<.001	<.001	<.001	.

通过数据分析可以得知：所有变量之间都存在显著的正相关关系（p<.05），随着教学自我效能感与提示词效能的提升，TPACK 中与技术相关的维度（TK、TCK、TPK、TPCK）均会有所提升。其中，提示词效能与 TCK 之间的相关性最强(相关系数=0.912)，与 TK 之间的相关性较弱（相关系数=0.662）。教学自我效能感与 TPCK 之间的相关性最强(相关系数=0.842)，与 TK 之间的相关性较弱（相关系数=0.561）。

6.讨论

研究发现职前教师在 TPACK 与技术相关的各维度和教学自我效能感、提示词效能等方面均值差异不大，但教学自我效能感对 TPACK 各维度均呈显著正相关关系，相关系数最高的维度是 TPCK 水平，相关性最低的维度是 TK 水平。可能原因在于职前教师通过 GAI 体验，更加认同 GAI 在未来教学中的重要作用，同时他们对未来教学的信心使他们认为自己能够有效运用 GAI 完成教学的全过程，产生较高的教学自我效能感。而 TPCK 是综合运用技术、学科教学知识和教学法知识的能力，注重三者的有机融合，是评估师范生能否更好地胜任教学的重要指标。教学自我效能感与 TPCK 均关注教学整体，因此相互之间具有较高的相关度。教学自我效能感与 TK 相关性最低的可能是由于在 AI-TPACK 背景下，TK 属于核心知识元素对 TPACK 起到间接影响，TPK、TCK 属于复合知识元素发挥着独特的中介作用（Ning et al.,2024）。随着 GAI 的广泛使用，教师关注技术知识的具体使用，而忽视了技术背后的原理知识，这是导致自我效能感较高的教师仍可能出现 TK 水平较低的情况的原因所在。另外，本研究设计的量表是采取自我评价法为主，本研究采取自我评价法，研究对象需要从 1“强烈不同意”到 7“强烈同意”中选择最符合自己的选项，具有一定的主观性，但也说明教师对于 TK 知识的重视水平弱于其他维度，与已有研究得出的将技术整合到特定学科教学中是职前教师在 TPACK 水平较为薄弱的部分相一致（王辞晓等，2018）。

研究还表明提示词效能仍与 TPACK 各维度的能力呈正相关关系，与 TCK 水平的相关系数最大，与 TK 水平相关系数最小，说明教师对自己能否较好的应用提示词的个人信念影响技术整合的相关能力的水平。提示词效能与 TCK 的相关性优于其他维度，主要在于 GAI 能够依据教学内容提供个性化的教学技术推荐，帮助职前教师更好的将学科内容知识与技术进行融合，而提示词是教师利用 GAI 工具完成这一功能的关键步骤。职前教师的提示词效能影响着个人对技术使用态度，关乎技术与教学融入的具体效果，而 TCK 水平更加注重教师使用 GAI 对学科内容进行转变（闫志明等，2020），因此二者具有较强的相关性。提示词效能与 TK 之间的相关系数低于其他维度则是由于职前教师认为技术知识多以具体技术的操作为主，尚未将提示词的书写纳入技术知识的范畴，但在使用过程中积累了一些提示词书写的经验，能够较好的使用 GAI 应用于教学，所以与 TK 之间相关系数较小，但仍呈现正相关关系。

7.结论和建议

基于问卷调查结果的分析和讨论，研究发现职前小学教师具有较好的 TPACK 水平，各维度得分较为一致，在 TPK 维度略优于其他维度，但在 TK 维度略差。个体差异中性别和专业方向均不会对职前小学教师的 TPACK 水平、提示词效能、教学自我效能感产生影响，但是不同专业方向的 TPACK 能力略有差异，艺术方向整体的 TPACK 水平较强于其他方向，文科方向 TK 能力稍有欠缺，理科方向在 TCK、TPK、TPCK 水平和提示词效能、教学自我效

能感上均弱于其他方向。提示词效能、教学自我效能感与职前小学教师的 TPACK 水平存在显著相关性。

根据数据分析得出的相关结论，本研究提出如下建议。

首先，需关注职前教师 TK 水平的提升。相比于 TPACK 其他维度，职前教师在 TK 维度普遍较弱，均值低于 TPACK 其他维度，虽然不显著，但 TK 是影响教师未来使用 GAI 进行教学的关键因素（Celik,2022），而且本研究发现 TK 水平与 TPK 之间具有较强的相关性，因此，职前教师培养的过程中需要通过多样化的技术知识讲座、实际操作等活动加强职前教师的 TK 水平，进而增强职前教师将技术应用与整合到教学之中的能力，提升教师的 TPACK 水平。

第二，个体差异虽然对于大一年级师范生的 TPACK 水平没有显著差异，但是不能忽略不同专业方向之间的 TPACK 均值差异，应对不同方向的职前教师制定有针对性的 TAPCK 能力培养策略。文科方向的 TK 能力弱于理科和艺术方向，则需结合日后的教学实际增设适用于语文、英语、德育等学科教学中的技术功能讲解，通过 TK 水平的提升逐步增强教师的 TPACK 能力。理科方向在 TCK、TPK、TPCK 水平和提示词效能、教学自我效能感方面均弱于其他方向，因此职前教师培养过程中需关注学科知识教学法和学科知识的教学，加强学习知识整合的步骤和方法培训，促进 TPACK 能力的发展。

最后，职前教师的培养需要关注提示词效能和教学自我效能感。提示词效能感和教学自我效能感与职前教师的 TPACK 水平显著相关，对 TPACK 与技术相关的各维度均有显著的相关性，是提升 TPACK 整体能力的重要因素。在对职前教师进行提示词效能培养的过程中，不仅需要进行提示词书写与修改的相关培训，更要将 GAI 与 TCK 相融合，对如何将技术与学科内容知识相融合进行教学，进而提升职前教师的提示词效能和 TPACK 能力。培养职前教师的教学自我效能感则需帮助职前教师了解 GAI 对教学的重要帮助和具体步骤，感受 GAI 为教学带来的便捷，在掌握技术知识的基础上理解并体会将技术知识、教学法知识、学科知识三者融合的有效路径和良好效果，从而同时提升职前教师的教学自我效能感和 TPCK 能力，带动 TPACK 其他能力的提升。

职前教师培养需要依据时代变化不断调整，目前 GAI 已经成为职前教师更便捷的教学辅助工具，职前教师的 TPACK 能力和提示词效能感、教学自我效能感更是需要持续关注的内容，需不断增设相应的课程与专业锻炼，培养职前教师的能力，以适应教育数字化转型的时代要求，为学生提供更加优质的教学。

参考文献

李红, 郝春东, 张旭. 教师教学效能感与学生自我效能感研究[J]. 高等师范教育研究, 2000,(03):44-48.DOI:10.13445/j.cnki.t.e.r.2000.03.009.

刘艳华,徐鹏,王以宁.教师整合技术的学科教学知识(TPACK)境脉因子模型构建研究[J].现代远距离教育,2015,(02):60-66.DOI:10.13927/j.cnki.yuan.2015.0022.

庞丽娟,洪秀敏.教师自我效能感:教师自主发展的重要内在动力机制[J].教师教育研究,2005,(04):43-46.

宋伟, 孙众. 数字教材环境下小学教师 TPACK 与自我效能感、技术接受度的关系——一项结构方程分析[J]. 中国远程教育, 2014,(09):39-46.DOI:10.13541/j.cnki.chinade.2014.09.009.

宋筱璇, 徐丹, 张紫蕴, 等. 提示词重构类型对 AI 生成内容效果的影响[J/OL]. 图书馆论坛,1-11[2025-01-03].http://kns.cnki.net/kcms/detail/44.1306.G2.20241008.1747.002.html.

王辞晓, 吴峰. 职前教师 TPACK 水平的绩效分析与改进路径[J]. 现代远距离教育, 2018,(02):62-71.DOI:10.13927/j.cnki.yuan.2018.0019.
谢志勇, 王红. 乡村教师数字化持续专业发展对数字化教学素养的影响——组织支持感和数字化教学自我效能感的中介作用[J]. 华南师范大学学报(社会科学版), 2024,(01):105-116+206.
朱永新, 杨帆. ChatGPT/生成式人工智能与教育创新：机遇、挑战以及未来[J]. 华东师范大学学报(教育科学版), 2023,41(07):1-14.DOI:10.16382/j.cnki.1000-5560.2023.07.001.
闫志明, 付加留, 朱友良, 等. 整合人工智能技术的学科教学知识（AI-TPACK）：内涵、教学实践与未来议题[J]. 远程教育杂志, 2020,38(05):23-34.DOI:10.15881/j.cnki.cn33-1304/g4.2020.05.003.
虞江锋, 张吉先. AI-TPACK 理论框架下开放大学教师的专业发展分析[J]. 职教论坛, 2022,38(04):103-109.
杨宗凯, 王俊, 吴砥, 等. ChatGPT/生成式人工智能对教育的影响探析及应对策略[J]. 华东师范大学学报(教育科学版), 2023,41(07):26-35.DOI:10.16382/j.cnki.1000-5560.2023.07.003.

Bardakci, S., & Alkan, M. F. (2019). Investigation of Turkish preservice teachers' intention to use IWB in terms of technological and pedagogical aspects. *Education and Information Technologies, 24*(5),2887–2907.

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior, 138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Feldman-Maggor, Y., Blonder, R., & Alexandron, G. (2024). Perspectives of Generative AI in Chemistry Education Within the TPACK Framework. *Journal of Chemical Education, 91*(9), 1638-1645.

Jain, K., & Raghuram, J. N. V. (2023). Gen-AI integration in higher education: Predicting intentions using SEM-ANN approach. *Journal of Educational Technology Development and Exchange, 7*(1), 49-62.

Karataş, F., & Aksu Ataç, B. (2024). When TPACK meets artificial intelligence: Analyzing TPACK and AI-TPACK components through structural equation modelling. *Educational Technology Research and Development, 62*(5), 1793-1814. doi: 10.1007/s11423-024-09899-z

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology, 34*, 67–76. https://doi.org/10.1016/j.cedpsych.2008.08.001

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054.doi:10.1111/j.1467-9620.2006.00684.x

Ning, Y., Zhang, C., Xu, B., Zhou, Y., & Wijaya, T.T. (2024). Teachers' AI-TPACK: Exploring the Relationship between Knowledge Elements. *Sustainability, 16,* 978.

Sun, J., Ma, H., Zeng, Y., Han, D., & Jin, Y. (2023). Promoting the AI teaching competency of K-12 computer science teachers: A TPACK-based professional development approach. *Journal of Educational Technology Development and Exchange, 16*(1), 1-17.

Tian, T., Chai, C. S., Chen, M.-H., & Liang, J.-C. (2023). University students' perceptions of learning with generative artificial intelligence. *Journal of Educational Technology Development and Exchange, 16*(1), 1-18.

Teo, T.; van Schaik, P. Understanding the intention to use technology by Preservice teachers: An empirical test of competing theoretical models. *IJHCI 2012, 28*, 178–188.

Velander, J., Taiye, M. A., & Otero, N. (2022). Artificial Intelligence in K-12 Education: Eliciting and Reflecting on Swedish Teachers' Understanding of AI and Its Implications for Teaching & Learning. *Journal of Educational Technology & Society, 25(*2), 97-112.

Wang, K., Ruan, Q., Zhang, X., Fu, C., & Duan, B. (2024). Pre-Service Teachers' GenAI Anxiety, Technology Self-Efficacy, and TPACK: Their Structural Relations with Behavioral Intention to Design GenAI-Assisted Teaching. *Behavioral Sciences, 14*(5), 373. https://doi.org/10.3390/bs14050373

Yang, Y., Xia, Q.*, Liu, C., & Chiu, T.K.F. (2023). The impact of TPACK on teachers' willingness to integrate generative artificial intelligence(GenAI): The moderating role of negative emotions and the buffering effects of need satisfaction. *Journal of Educational Technology & Society, 26*(1), 197-212.

Yue, M., Jong, M. S. Y., & Ng, D. T. K. (2023). Understanding K–12 teachers' technological pedagogical content knowledge readiness and attitudes toward artificial intelligence education. *Journal of Educational Technology & Society, 26*(1), 197-212.

Zeng, Y., Wang, Y., & Li, S. (2022). *The relationship between teachers' information technology integration self-efficacy and TPACK: A meta-analysis.* Frontiers in Psychology, 10.3389/fpsyg.2022.1091017.

Zhao, Z., Wallace, E., Feng, S., Klein, D.,& Singh, S. (2021). *Calibrate before use: Improving few-shot performance of language models.* In International conference on machine learning(pp.12697–12706).

翻转课堂下不同平均绩点学生的自主学习路径分析：K-means 聚类与过程挖掘

Self-regulated Learning path analysis of students with different Grade Point Average in Flipped classroom: K-means clustering and Process Mining

王婷[1]，龙陶陶[2*]，朱晓萌[3]，伍冬[4]
[12]第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）是学生的一项重要能力，特别是在像翻转课堂这样的混合学习环境中。本研究采用基于问题的学习（PBL）方法，探讨了翻转课堂中工科学生的自我调节学习行为。本研究以工科学生自主学习行为为视角，采用 K 均值聚类分析基于学生的平均成绩绩点进行分组，将学生分为五类。然后，利用过程挖掘技术对各类学生的学习行为特征进行可视化分析，旨在为探索和优化工科翻转课堂教学提供基于客观数据的理论支持和实践参考。研究发现，不同平均成绩绩点段学生的学习行为路径存在较大差异。基于此，本研究提出了一些可能提高工科学生在翻转课堂中自我调节学习效果的建议。
【关键词】 自我调节学习（SRL）；工程翻转课堂；过程挖掘；K 均值聚类分析

Abstract: *Self-regulated learning (SRL) is a critical competency for students, particularly in blended learning environments like flipped classrooms. This study explores the SRL behaviors of engineering students in a flipped classroom employing a problem-based learning (PBL) approach. This study takes engineering students' self-regulated learning behavior as a perspective, uses K-means cluster analysis to group students' Grade Point Average, and divides students into five categories. Then, the process mining technology is used to analyze the learning behavior characteristics of all kinds of students, aiming to provide theoretical support and practical reference based on objective data for exploring and optimizing flipped classroom teaching in engineering. It is found that there are great differences in the learning behavior paths of students with different GPA segments. Based on this, this study puts forward some suggestions to improve the self-regulated learning effect of engineering students in flipped classroom.*
Keywords: Self-regulated learning (SRL), flipped engineering classroom, process mining, K-Means clustering

1.引言

自我调节学习（SRL）被认为是 21 世纪成功的基石技能，它使学习者能够掌握自己的认知、动机和行为过程，以实现教育目标 (Manukaram, K., & Abdullah, M. N. L. Y., 2021; Du, J., Huang, W., & Hew, K. F., 2021) 。虽然 SRL 已被证明可以显著提高学生的参与度和学习成绩，特别是在在线学习环境中 (杨燕清、吴明、林艳芹和洪霞, 2022； 李文昊、任晓曈、朱希雅和 陈冬敏, 2022)，但研究仍在为理解其微妙的行为途径而努力。翻转课堂是一种广泛使用的

混合学习策略，它要求学生参与课前材料，如视频和测验（王翠如和胡永斌，2018），从而要求积极的计划和自我调节。尽管有这种潜力，但之前关于翻转课堂中 SRL 的研究往往依赖于主观的自我报告，限制了研究结果的可靠性和深度。

聚类分析是一种重要的数据挖掘方法，其核心是通过分析数据内部的相似性来发现“自然分组”（Blashfield, R. K., & Aldenderfer, M. S., 1978），即聚类。这种无监督学习方法可以帮助研究人员深入挖掘数据的内在规律，特别是在需要处理大量复杂多维数据的场景中，表现出较强的特征分类能力。因此，K-Means 聚类分析在教育领域经常被用来对等级进行分类 (Wu, J., & Wu, J., 2012)。

过程挖掘是一种数据驱动的分析事件日志的方法，它已经成为一种强大的工具，可以客观地探索在线环境中学生的学习行为和模式 (Van Der Aalst, W., 2011)。虽然 SRL 的应用已经扩展到分析 mooc 的参与度 (邓国民、徐新斐和朱永海, 2021)，但在翻转课堂环境下的探索较为有限，特别是在工程学科中，这种方法也许能产生关键的见解。

针对上述问题，文章采用过程挖掘技术，对翻转课堂环境下不同平均成绩绩点的工科学生的自主学习路径进行了深入探究。最后根据研究结果，有针对性地提出了一些改进教学的建议，以期为工程教育提供有价值的见解。

2.文献综述

2.1. 自我调节学习（SRL）

SRL 起源于 Bandura 提出的社会认知理论，随后 Zimmerman 等人在此理论的基础上对 SRL 进行了具体的解释。Zimmerman (1989)认为 SRL 是指学习者能够从元认知、动机和行为等方面积极参与学习过程，主动获取知识和技能，并强调自我调节策略、自我效能感和学业目标是 SRL 的三个要素 (Zimmerman, B. J., 1989)。SRL 经过几十年的发展，其模型也逐渐丰富起来。Zimmerman (2009) 认为 SRL 本应分为三个过程，即计划阶段、执行阶段、自我反思阶段。Pintrich (2004)将 SRL 分为四个阶段：预见、计划和激活；监测；控制；反应和反思 (Pintrich, P. R. 2004)。

研究强调了 SRL 在翻转课堂的课前活动中的重要性，如视频讲座和测验，这些活动对课堂学习效果有显著影响。例如，最近的研究强调设计混合和翻转学习策略，以提高工程本科生的参与度和技能习得 (Obada, D. O., Bako, R. B., Ahmed, A. S., & Obada, I. B., 2023)。然而，在深入探索翻转课堂环境下的 SRL 行为方面，特别是在工程教育中，仍然存在显著的差距。虽然像 mooc 这样的在线教育环境已经得到了广泛的研究，但针对工程学生在翻转环境中所面临的具体 SRL 模式和挑战的研究仍然有限。SRL 在翻转课堂中的应用见证了后疫情时代的创新 (刘云凤,阎小青 & 乔增杰.2021； Alsuraihi, A. A., 2022)，融入了互动工具和丰富的学习活动，以提高参与度和成果。然而，这些研究主要使用学生自述的数据来调查 SRL 行为，很难做到客观。

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. K 均值聚类

K 均值聚类是一种基本的无监督学习算法 (Blashfield, R. K., & Aldenderfer, M. S., 1978)，广泛用于根据相似度对数据点进行分组。该算法由 MacQueen 于 1967 年首次提出，通过最小化每个数据点与其最近的聚类质心之间的欧几里德距离来迭代数据点到聚类的分配。麦昆不仅详细解释了算法的过程，还从数学的角度论证了其原理 (Morissette, L., & Chartier, S., 2013)。K 均值聚类的本质是在给定的数据集{x1，x2，…，xn}中识别 K 个聚类质心{a1，a2，…，aK}，使得每个数据点与其最近的质心之间的总平方距离最小。自引入以来，K 均值聚类以其简单性、计算效率和处理大型数据集的鲁棒性获得了认可。它基于固有模式有效地将数据划分为不同的集群的能力已经在自然语言处理、土壤科学和考古学等不同领域得到广泛采用，这些优点使得该方法在分类和分组研究中得到了广泛的应用。

2.4. 过程挖掘

流程挖掘，也被称为工作流挖掘，是一种从系统内部事件日志中提取信息的方法，用于发现、监控和改进实际操作过程 (徐晓青、赵蔚、姜强、刘红霞和乔丽方, 2022)。近年来，流程挖掘被视为探索学生自我调节学习行为的新技术。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。

2.5. 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 K 均值聚类和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。本研究旨在深入探究翻转课堂环境下不同平均成绩绩点的工科学生的自主学习路径，以期为解决当前工程翻转课堂中学生自我调节过程中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。课题组选择工程课程《计算机在材料工程中的应用》进行翻转课堂教学实践，并根据基于问题的学习策略设计课程活动，包括在线学习和课堂学习。所有学生必须登录在线学习平台——智能学习管理系统“小雅”。平台可进行教与学，如发布讨论、测试、问卷、作业、课堂习题等课程任务，智能完成学习分析。学生需要在课前完成相关的学习任务，然后在课堂上与老师和同学进行交流和互动。

3.2. 参与者

本研究共选取 194 名材料工程专业的大三学生，其中高分子专业 70 名，金属材料专业 124 名。这些学生自入学以来，没有参与过工程翻转课堂和系统的基于问题的学习活动,也从未接受过任何与自我调节学习相关的训练。

3.3. 工具

本研究采用 Pintrich 提出的四阶段自我调节学习模型对学生的学习行为进行分类。该模型包括以下阶段：(a)预见、计划和激活，(b)监测，(c)控制，(d)反应和反思。此外，在 SRL 理论的框架下，学生的在线学习活动进一步被划分为 10 种不同的事件类型，详见表 1。

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
控制	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
	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	提交反思	S-R	提交个人反思报告
	查看论坛	V-F	查看论坛或在论坛发表观点

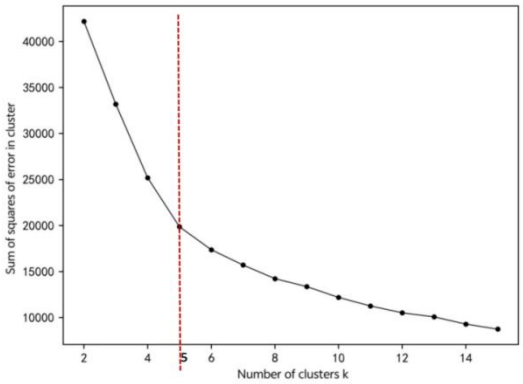
3.4. 数据收集与分析

参与者完成了一个学期的工程类翻转课堂学习。在这种学习模式下，参与者需要在课前通过“小雅”智能学习平台完成所需的知识。研究团队收集了参与者在学习平台上产生的所有事件的日志数据，共计 875,100 个事件。每个事件日志都包括参与者与学习平台的交互，以及其他一些必要的信息，如学生 ID、所采取的操作和时间戳。如表 2 所示，我们展示了某一时刻学生行为事件日志的数据抽取。

表2 事件日志的示例

Student ID	SRL phase	Detailed activity
123	Monitoring	View Resources
123	Monitoring	View Resources
123	Monitoring	View Resources
123	Monitoring	View Resources
123	Control	Download Resources

在分析数据之前，对数据进行预处理，排除了一些与研究无关的行为事件数据，如学生登录和注销，得到有效事件 797995 个。技术人员使用 pycharm 平台（由 JetBrains 为 Python 语言开发的集成开发环境（IDE））来标记数据。该研究还收集了所有学生的学习成绩数据。采用 k-means 聚类分析，根据平均绩点对学生进行分类，根据生成的聚类肘图将学生分为 5 个类别（图 1）。在本研究中，根据 GPA 将这 5 个类别命名为 1~5。最后，利用 ProM 6.12 对五类学生的背景行为数据进行过程挖掘，并将学习过程可视化。本分析主要使用 iDHM 工具中的依赖图流模型，该模型基于 ProM 软件中的启发式挖掘算法。



4.结果

4.1. 五类学生在pbl导向的工程翻转课堂中的自我调节学习行为

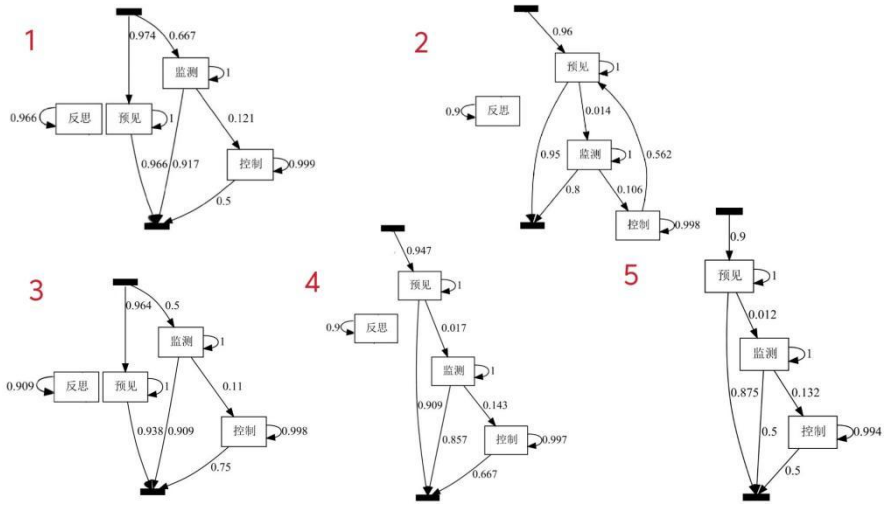


图 2 五类学生在 SRL 阶段Dependency Graph 过程模型

如图 2 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 2 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。五类学生在根据 GPA 聚类后的自我调节学习路径如图 2 所示，从 1 类到 5 类成绩依次递减。可见，前 4 类学生均存在反应和反思行为(1：依赖度= 0.966；2：依赖度= 0.900；3：依赖度= 0.909；4：依赖度= 0.900)，而第 5 类，即 GPA（平均成绩绩点）最低的学生，则没有这种行为。第 2 类学生在完成“控制行为”（提交作品、下载资源等）后并没有选择直接完成任务，而是再次进行“预习、规划、激活”（重新打开知识点学习）(控制(C)➞预习、规划、激活（F-P-A），依赖度=0.562）。

4.2. 自我调节学习视域下，五类学生在pbl导向的工程翻转课堂中的具体学习过程

如图 3 所示，Dependency Graph 过程模型被用于分析 SRL 视域下五类学生在 PBL 导向的工程翻转课堂中的学习行为过程。可以看出，第 1 类学生在浏览学习资源（V-R, 依赖度= 1.000）中表现出较高的重复性和碎片性。此外，在任务完成过程中，1 类学生在分析任务之后更频繁地返回到查看资源的行为（A-T→V-R，依赖度= 0.119）。第 2 类学生则更频繁地进行移动资源操作（M-R，依赖度= 0.955），这表明该类学生在小组讨论中更活跃。同时可以看到该类学生从查看资源到查看任务的路径较少（V-R→V-T，依赖度= 0.006），更多学生选择在下载资源后直接开始完成作业任务（V-R→D-R，依赖度= 0.111），这说明这类学生更加注重任务本身的完成。3 类学生在“查看资源”行为后倾向于转向“查看论坛”（V-R→V-F，依赖度= 0.400），表现出对他人意见的习惯依赖。并且他们分析任务的行为是孤立进行的（A-T，依赖度= 0.997），这表明该类学生可能没有很高的协作水平。查看教学大纲后，第 4 类学生直接进入移动资源部分（V-S→M-R，依赖度= 0.500）。很明显，他们的主要目标是完成老师布置的任务。第 5 类学生一般在查看教学大纲后选择直接提交作品（V-S→S-W，依赖度= 0.500），说明他们的操作大多是以完成任务为导向的，并且该类学生自我调节学习路径中的反思行为较少，他们生成计划的行为也与后续的学习任务相隔离（G-P，依赖度= 0.917）。

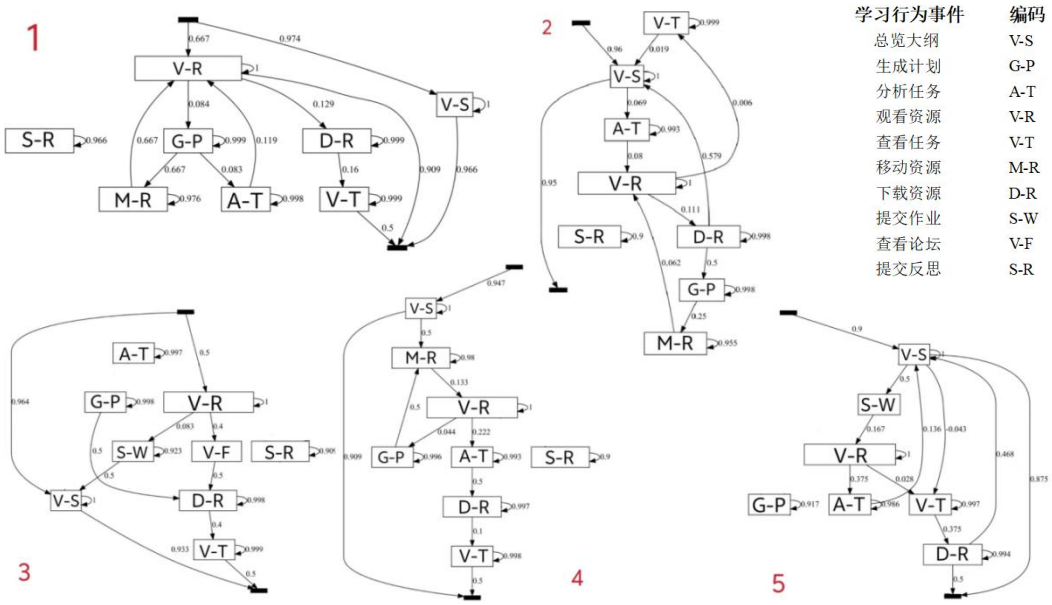


图 3 自我调节视域下五类学生工程翻转课堂学习行为Dependency Graph 过程模型

5.讨论和启示

本研究采用聚类分析和过程挖掘的方法，对工程类本科生在翻转课程中的自我调节学习行为模式进行了研究。结果表明，根据学习成绩聚类的不同学生群体在自我调节过程中采用的学习行为模式存在较大差异。这表明不同学习成就水平的学生在翻转课堂学习中采用了多种不同的 SRL（自我调节）策略。

本研究的一个重要发现是高成绩群体学生的学习表现通常具有较高的自我调节能力和学习策略的灵活性。研究发现，最高成绩组学生在“浏览资源”方面表现出较高的重复和碎片化学习行为。他们多次复习学习材料以检查遗漏和填补空白，这种策略表明他们具有高度的自主性和完成任务的内在动机。值得注意的是，最高成绩组的学生在完成任务时，往往会采取深入分析任务并回溯到考核标准的策略，以确保任务的准确性和高质量完成。这种高效的学习策略不仅可以帮助学生在学习过程中优化自己的学习路径，还可以提高完成任务的质量。

本研究的另一个重要发现是，低成绩群体学生在学习过程中往往面临着更困难的挑战，特别是在处理复杂的工程问题和开放式任务时。研究发现，低成绩群体的学生不仅缺乏反思行为，而且在学习过程中往往表现出被动的任务完成和低质量的参与。在粗略地看了一下课程资源后，他们往往选择直接提交作业，对任务要求和学习计划的实施情况缺乏深入的了解。此外，在面对开放性问题时，低成绩群体学生往往缺乏深入的思考，给出的答案过于肤浅，无法通过应用其他学科的知识来有效解决问题。这种现象反映出了学习困难的学生在解决复杂问题时的认知局限性和学习策略的不足。

基于研究发现，文章提出了一些建议，以期进一步提升翻转课堂学生的学习效率，实现深度学习。首先，鼓励其他学生采用高成绩群体的学习策略将有助于提高他们的学习效果。教师可以通过案例研究或学习实例来鼓励其他学生遵循高成绩群体的学习方式，例如深入分析任务要求或反复学习资源。其次，面对学习有困难的学生，教师可以通过逐步引导和提供更多的支持来帮助学生克服困难。例如，可以设计更简单的学习任务或小步骤指南来帮助低成绩群体的学生逐步培养解决问题的技能。最后，在自我调节学习中，高成绩群体拥有更多的策略，与低成绩群体的学习者存在较大的差距。因此，有必要加强对学生自主学习策略的训练。

6.结论和局限性

本研究采用过程挖掘和 K 均值聚类技术探索了工科学生在翻转课堂环境下的自主学习行为路径。研究结果显示，按平均成绩分类的五类学生在学习策略和行为上存在显著差异。高成绩群体的学生展示出重复性和资源密集型的策略，有效地利用了自我反思和团队合作。中等成绩群体的学生表现出较多对外部意见的依赖，合作能力相对较弱，而低成绩群体的学生表现为任务导向式的学习习惯，且小组合作也展现出缺乏反思行为的肤浅参与。

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。

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。

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。

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。

刘云凤,阎小青和乔增杰.(2021).以培养自我调节学习能力为导向的分析化学实验混合教学模式.大学化学,36(01),124-128.

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。

Blashfield, R. K., & Aldenderfer, M. S. (1978). The literature on cluster analysis. *Multivariate behavioral research, 13(*3), 271-295.

Wu, J., & Wu, J. (2012). Cluster analysis and K-means clustering: an introduction. *Advances in K-Means clustering: A data mining thinking,* 1-16.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review, 16*, 385-407.

Obada, D. O., Bako, R. B., Ahmed, A. S., Anafi, F. O., Eberemu, A. O., Dodoo-Arhin, D., ... & Obada, I. B. (2023). Teaching bioengineering using a blended online teaching and learning strategy: a new pedagogy for adapting classrooms in developing countries. *Education and Information Technologies, 28*(4), 4649-4672.

Alsuraihi, A. A. (2022). The effect of implementing mind maps for online learning and assessment on students during COVID-19 pandemic: a cross sectional study. *BMC medical education, 22*(1), 169.

Blashfield, R. K., & Aldenderfer, M. S. (1978). The literature on cluster analysis. *Multivariate behavioral research, 13*(3), 271-295.

Morissette, L., & Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology, 9*(1), 15-24.

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.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review, 16*, 385-407.

Obada, D. O., Bako, R. B., Ahmed, A. S., Anafi, F. O., Eberemu, A. O., Dodoo-Arhin, D., ... & Obada, I. B. (2023). Teaching bioengineering using a blended online teaching and learning strategy: a new pedagogy for adapting classrooms in developing countries. *Education and Information Technologies, 28*(4), 4649-4672.

Bergmann, J., & Sams, A. (2012). Before you flip, consider this. *Phi Delta Kappan, 94*(2), 25-25.

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.

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.

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.

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.

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.

生成式人脸隐私保护赋能安全学习分析

Generative face privacy protection enables secure learning and analysis

王家阳，陈晓雨，刘毓，王涛*
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 湖北 武汉 430079
*tmac@ccnu.edu.cn

【摘要】 为了有效保护个人隐私，特别是面部识别中的身份和敏感属性信息，本研究旨在开发一种既能保障面部隐私又能维持图像质量和实用性的面部隐私保护方法。本研究提出一种基于扩散模型的面部隐私保护框架，它利用身份编码器和属性编码器对面部进行分解，通过拉普拉斯噪声引入差分隐私来增强保密性，最后由去噪扩散隐式模型 DDIM 作为解码器生成高质量的隐私化面部。实验结果表明，该方法不仅在保护隐私方面表现出色，而且保持了图像质量，并与各种计算机视觉任务兼容，成功地在隐私保护和图像质量及功能性之间找到了平衡点。

【关键词】 面部隐私；扩散模型；差分隐私；身份编码器；属性编码器

Abstract: *To effectively preserve personal privacy, particularly the identity and sensitive attribute information in facial recognition, this study aims to develop a facial privacy-preserving mechanism that not only ensures facial privacy but also maintains image quality and practicality. This research introduces a diffusion model-based facial privacy-preserving framework, which leverages an identity encoder and an attribute encoder to effectively decompose facial representations. Additionally, differential privacy is enhanced by introducing Laplacian noise. Finally, the Denoising Diffusion Implicit Model (DDIM) serves as the decoder to generate privacy-preserving facial images. Experimental results demonstrate that this method not only excels in privacy preservation but also preserves image quality, while being compatible with various computer vision tasks, thereby successfully striking a balance between privacy preservation, image quality, and functionality.*

Keywords: facial privacy, diffusion model, differential privacy, identity encoder, attribute encoder

1.引言

在当今数字化教育蓬勃发展的时代，在线学习等教育场景广泛应用先进的人工智能与计算机视觉技术，以实现更精准、高效的教育分析。例如，通过记录和分析学生的人脸数据，能够获取学生的课堂参与度、情绪状态等关键信息(Kew et al., 2022)，进而为个性化教学提供有力支持。这一技术的应用使得教育工作者能够更深入地了解学生的学习过程，及时调整教学策略，提升教学质量。

然而，这种对学生人脸数据的大规模收集和分析也引发了严峻的隐私问题。当学生的面部照片被在线学习平台记录和处理时，这些数据可能面临被非法利用的风险(Ismail, 2025)。不法分子可能会从中提取学生的身份信息，甚至包括年龄、性别等敏感的人口统计信息，这些信息的泄露不仅侵犯了学生的个人隐私权，还可能引发潜在的歧视问题。比如，基于年龄或性别的偏见可能会影响学生获得公平的教育资源和机会，对学生的身心健康和学习发展造成负面影响。

面对这一挑战，全球各国政府积极制定相关法规，以规范教育领域面部数据的收集、存储、处理及共享流程。例如，欧洲的《通用数据保护条例》（GDPR）(Hiller et al., 2019)和

美国的加利福尼亚州《消费者隐私法案》（CCPA）等，都为个人隐私权提供了法律保障。尽管技术进步为教育带来了诸多便利，但在教育场景中，我们必须高度重视学生面部隐私的保护，探索更为先进的技术手段，确保在保护隐私的同时，不影响教育分析的准确性和有效性。

目前，针对包含面部信息的图像或视频的隐私保护方法，主要分为身份隐藏和属性隐藏两类。身份隐藏技术通常采用模糊化、像素化处理或对面部细节进行修改等方式，降低面部识别技术的识别准确性，从而隐藏学生的身份信息。属性隐藏策略则侧重于隐匿面部敏感属性，如年龄、性别等，以消除潜在的歧视风险。

然而，这些传统方法存在明显的局限性。例如，模糊处理或像素化往往会导致图像质量严重下降，使得基于面部图像的后续教育分析，如表情识别、专注度分析等难以准确进行。此外，仅依靠身份隐藏或属性隐藏，无法充分满足教育场景中对隐私保护和面部图像功能可用性的双重需求。为了解决这些问题，我们结合身份隐藏和属性隐藏的任务，利用具有强大稳定性和图像生成多样性的扩散模型，提出了一种新的面部隐私保护方法。该方法旨在提供一种灵活的面部隐私保护范式，实现可配置、机密且实用的面部隐私保护，确保在保护学生面部隐私的同时，保持图像的视觉逼真度，不影响后续的教育分析工作。

本文主要的研究工作如下。

1) 我们提出了一种基于扩散的灵活面部隐私保护方法，该方法结合了身份隐藏和属性隐藏的任务，同时实现了可配置、机密且高质量的面部隐私保护。

2) 我们引入了身份编码器和属性编码器，以有效分解面部表征，这有助于通过自适应管理隐私关注并同时保护面部身份表征和属性表征，来实现灵活的面部隐私保护。

3) 我们将差分隐私融入其中，通过在身份表征以及控制属性变化的参数中引入拉普拉斯噪声，来保护面部身份和敏感面部属性的隐私，从而增强了我们的方法对潜在隐私攻击的机密性和鲁棒性。

4）我们将保护后的表征由强大的扩散生成模型 DDIM 合成具有匿名身份和属性的面部图像。我们的方法能够在稳健保护面部隐私的同时，生成高质量且自然的面部图像。

2.研究内容

为了克服在多样化的面部隐私保护和高性能要求之间做出妥协的挑战，我们提出了一种利用扩散模型的创新性面部隐私保护方法。具体而言，该方法包括三个阶段，如图 1 所示。阶段一：面部解耦。将输入的面部图像及其潜在空间解耦为身份和属性，然后将它们组合成一个统一的语义表征，通过 DDIM 进行重建恢复原始面部。阶段二：面部净化。根据特定的隐私要求，在身份表征和属性转换参数中注入拉普拉斯噪声，从而隐藏面部隐私。阶段三：面部生成。通过扰动的表征由 DDIM 合成具有匿名身份和属性的面部图像。因此，我们的方法能够在稳健保护面部隐私的同时，生成高质量且自然的面部图像。

我们的方法提供了一种灵活的面部隐私保护范式，同时实现了可配置、机密且实用的面部隐私保护。可配置的面部隐私保护意味着我们可以根据现实生活应用的需求，选择性地保留身份、任何感兴趣的属性或它们的任何组合，同时保持其他属性，包括姿势和背景。机密性意味着该方法生成的增强隐私的面部图像具有高度安全性，能够防止各种面部识别模型的识别。实用性则表明该方法输出的增强隐私的图像保留了其在各种下游计算机视觉应用中的实用性，包括面部检测和表情识别。这种能力使得我们的方法能够在现实场景中有效实施，平衡隐私需求与复杂视觉分析任务的操作要求。

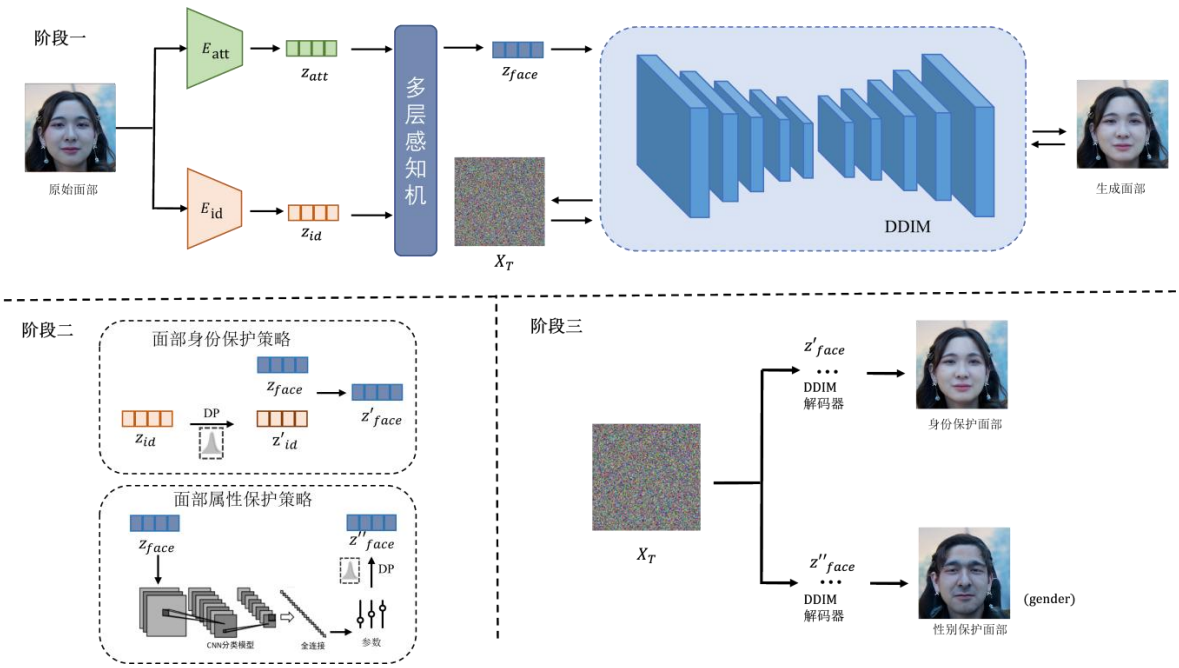


图 1 我们的方法框架概述

3.本文方法设计

3.1. 阶段一：面部解耦

为了有效且灵活地控制多样化的面部隐私，我们开发了两个独立的语义编码器和一个基于扩散的解码器来解耦面部信息。具体而言，身份编码器E_{id}和属性编码器E_{att}分别负责提取与个体身份和面部属性相关的表征z_{id}和z_{att}。这些表征集随后被合并，并通过多层感知机（MLP）进一步处理，以形成一个全面的高级语义表征z_{face}。我们采用了一个条件 DDIM 解码器，它以随机噪声x_T为输入，并使用z_{face}作为条件来指导解码器进行有效的去噪和面部重建 x0 ：

$$p_\theta(x_{0:T}|z_{face} = p(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t, z_{face}))$$

为了获得与属性无关的身份表征，我们选择使用预训练的身份识别模型 ArcFace(Deng et al., 2019)作为身份编码器，该编码器将输入图像编码为身份表征$z_{id} = E_{id}(x_0)$。ArcFace 在面部识别领域以其高精度而著称，能够识别不同姿势或表情的面部图像中的个体。因此，我们期望它能提供我们所需的解纠缠特性。多项研究(Nitzan et al., 2020)(Wang et al., 2021)表明，使用预训练编码器提取的表征已经足够解纠缠，无需额外学习。类似地，我们利用一个预训练的属性编码器(Wen et al., 2022),来捕获面部属性信息，如表情、性别和年龄。此外，面部解耦涉及使用此预训练编码器从图像中提取身份和属性表征，然后将这些表征输入到一个额外的可训练 MLP 中，该 MLP 将它们映射到更高的语义空间。这一过程有效地形成了一个全面的面部表征，提高了面部识别和分析的准确性和有效性。

3.2. 阶段二：面部净化

面部隐私保护是一个至关重要的领域，它包含两个关键方面：身份表征保护和敏感属性保护。

身份表征保护：在实践中，我们使用差分隐私技术来加强身份表征的保护。具体而言，我们在身份表征上添加拉普拉斯噪声，引入随机性以有效模糊个体。这种方法确保即使数据被泄露，攻击者也无法准确重建个体的真实身份，从而保护用户的个人安全。

敏感属性保护：为了有效保护如年龄和性别等敏感属性，我们开发了一个高度灵活的属性分类器，能够精确识别这些表征。在我们的研究中，我们采用了 DiffAE(Preechakul et al., 2022)中的方法，并对 CelebA-HQ(Karras, et al., 2017)数据集中的每个属性进行了详细分析。我们将面部图像的身份表征和属性表征相结合，创建了一个高级的语义空间。在这个空间中，我们使用公式$C_{att}(z_{face}) = \text{sigmoid}(w_{att}^{T} z_{face})$训练了一个线性分类器，以根据高级语义表征精确识别特定属性。此外，我们还采用了归一化技术，确保表征保持在单位球面上，具体通过表达式$\lambda_1 \text{Norm}(z_{face} + sw_{att})$来实现。其中，s 是一个缩放超参数，用于控制属性修改的强度。通过调整与特定属性相关的语义表征，并应用缩放和归一化技术，我们旨在在图像中实验性地改变这些属性。

因此，为了保护敏感属性，我们在缩放超参数 s 中引入了拉普拉斯噪声，生成了受扰动的面部语义表征。我们还为属性变化的方向增加了随机性——例如，正值表示年龄增加，而负值表示年龄减少。这种隐私保护策略确保即使在数据公开或共享时，特定的个人信息也能保持机密性。这实现了隐私保护和数据可用性之间的有效平衡。

3.3. 阶段三：面部生成

生成身份匿名化的面部图像：在生成身份匿名化的面部图像的过程中，我们依赖于两个关键元素：已经脱敏的身份表征z'_{id}和原始的属性表征z_{att}。为了确保生成图像的质量，我们在第一阶段训练得到的 DDIM 解码器的所有参数都被固定。通过一个多层感知机（MLP），我们将扰动后的身份表征z'_{id}与原始的属性表征z_{att}相结合，形成匿名化面部的新语义表征z'_{face}。最后，使用这个 DDIM 解码器，我们基于z'_{face}生成高质量的身份匿名化面部图像。这个过程确保了面部身份的保护，同时保持了面部图像的自然性和真实性。

生成属性匿名化的面部图像：生成属性匿名化的面部图像的过程依赖于属性脱敏的语义表征z''_{face}。同样地，为了确保生成高质量的匿名化面部图像，我们在第一阶段训练完成后固定了 DDIM 解码器的所有参数。使用这个具有固定参数的 DDIM 解码器，我们基于语义表征z''_{face}生成属性匿名化的面部图像。

4.实验分析

4.1. 实验设置

数据集：我们在以下数据集上进行保护操作：（i）CelebA-HQ(Karras, et al., 2017)，它包含来自 CelebA 数据集的 30,000 张高分辨率（1024x1024）的名人面部图像，这些图像被标注了 40 个不同的属性标签，涵盖了年龄、性别和种族等人口统计表征。这些标签针对每张图像内外区域的面部表征。（ii）FFHQ 数据集(Karras et al., 2019)，它包含了 70,000 张高质量面部图像，展示了广泛的年龄、性别和表情多样性，所有图像的分辨率均为 128x128 像素。

基准：为了验证本文所提出的方法的有效性，我们将其与基准匿名化方法进行比较：DeepPrivacy(Hukkelås et al., 2019)和 CIAGAN(Maximov et al., 2020)。

攻击模型：为了评估面部身份匿名化的有效性，我们对两个主流的黑盒面部识别（FR）模型进行了全面测试，包括 FaceNet 和 ArcFace。

为了评估面部属性匿名化的有效性，我们专注于保护两个主要属性：性别和年龄。通过一系列实验，我们验证了我们的匿名化算法在这些属性上的有效性。我们采用了两种性别攻击算法：Visual Geometry Group 16 (VGG16)和支持向量机（SVM），以及两种年龄攻击算法（区分年轻和年老）：SVM 和 K 最近邻（KNN）。所有这些算法都是在 CelebA 数据集的原始图像上进行训练的。

评估指标：我们使用攻击成功率（ASR）来评估不同方法的隐私保护效果。此外，我们还采用了几种常见的指标来评估图像质量，包括结构相似性指数（SSIM）、峰值信噪比（PSNR）和学习感知图像块相似性（LPIPS）。

4.2. 面部身份隐私保护

图像质量的结果: 图 2 展示了不同方法生成的匿名面部图像。可以观察到，CIAGAN 生成的匿名图像存在显著的失真，而 DeepPrivacy 虽然保留了一定程度的照片真实性，但未能有效保留与身份无关的属性，如表情。相比之下，我们的方法生成的图像具有自然的面部表征和增强的照片真实性，成功保留了与身份无关的属性，如表情和姿势。表 1 报告了图像质量的评估结果，表明我们的方法在所有图像质量指标上都优于其他方法。

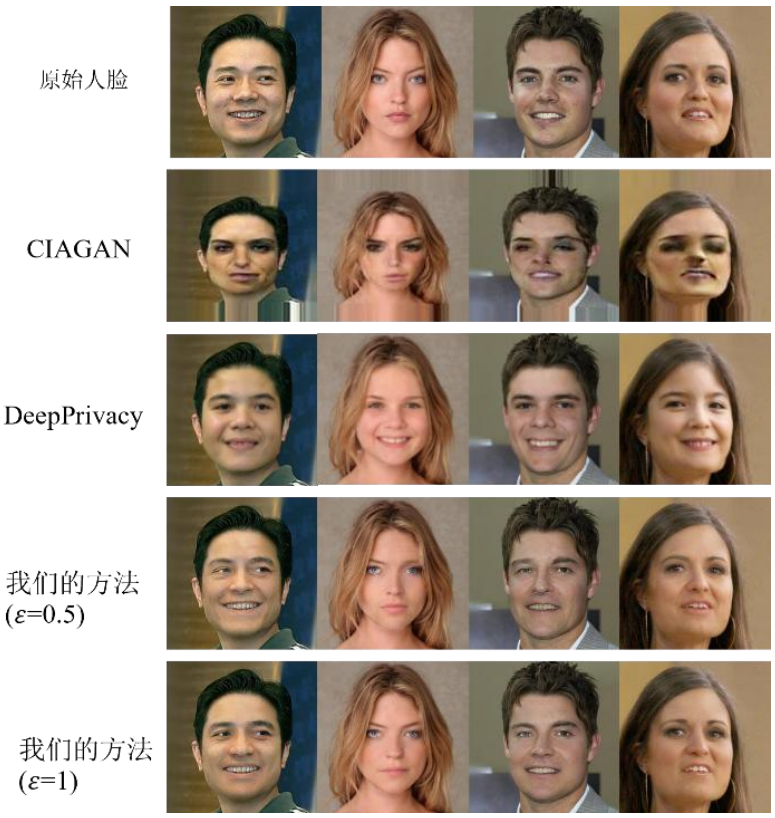


图 2 与基准方法的定性比较

表 1 不同图像匿名化方法图像质量的定量比较

方法	PNSR	SSIM	LPIPS
CIAGAN	21.863	0.7401	0.5499
DeepPrivacy	21.012	0.7808	0.079
我们的方法 ($\epsilon = 0.5$)	27.75	0.8603	0.0330
我们的方法 ($\epsilon = 1.0$)	27.73	0.8635	0.0329

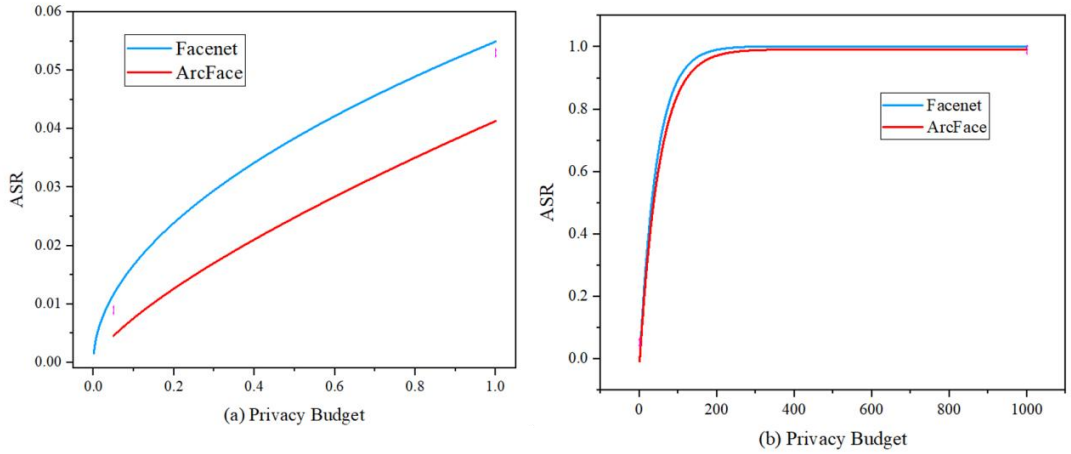


图 3 不同隐私预算下，身份保护图像的黑盒攻击成功率趋势

对黑盒攻击的比较: 表 2 报告了使用 CelebA 数据集对两个流行的面部识别（FR）模型进行黑盒攻击的定量结果。我们测试了针对两个特定身份的目标攻击性能，结果表明，与其他方法相比，我们的方法展示了最低的攻击成功率，这证实了我们的方法在生成匿名面部图像方面的有效性，符合预期。如图 3 所示，随着隐私预算的增加，攻击成功率逐渐上升。这些数据表明，较低的隐私预算在降低成功攻击的可能性方面更为有效，从而增强了匿名化效果。

表 2 不同图像匿名化方法黑盒攻击成功率比较

方法	Facenet 的 ASR	ArcFace 的 ASR
CIAGAN	0.151	0.100
DeepPrivacy	0.077	0.067
我们的方法 ($\epsilon = 0.5$)	0.044	0.024
我们的方法 ($\epsilon = 1.0$)	0.053	0.042

4.3. 面部属性隐私保护

图像质量的结果: 图 4 展示了性别转换的视觉效果，特别是将男性图像转换为女性图像时。生成的女性图像通常具有更柔和的面部表征和更长的头发，这与女性外貌的常见表征相符。图 5 展示了年龄转换的影响：当年轻人被转换为老年人时，生成的老年人图像通常显示出白胡子和更多的皱纹，这是老年人的典型表征。此外，随着隐私预算的增加，生成图像的质量也相应提高。

图 4 性别保护处理后的图像

黑盒攻击的结果：图 6 展示了针对四个流行的黑盒攻击，对属性保护的面部图像（特别是处理过以隐藏敏感属性如性别和年龄的图像）的定量结果。这些结果有效地降低了性别分类器的性能，从而保护了敏感属性。然而，重要的是要注意，在较低的隐私预算下，虽然隐私保护得到了改善，但分类器的性能受到了显著影响。这主要是因为为了增强隐私保护而应用的过多扰动导致图像质量明显下降，进而妨碍了分类器从图像中提取关键面部表征（如眼睛、鼻子和嘴巴的形状）的能力。过多的扰动会模糊或完全遮挡这些关键表征，使得分类器难以提取有意义的信息来进行准确的分类。值得注意的是，这种现象不仅限于性别属性的保护，还适用于其他敏感属性的匿名化，如年龄，这凸显了我们的方法在保护各种类型私人信息方面的多功能性和有效性。

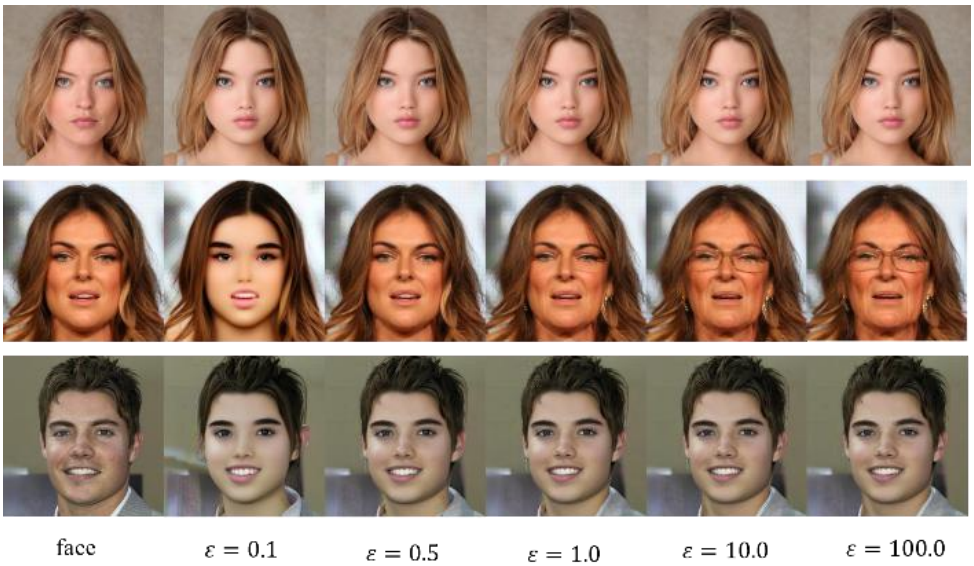


图 5 年龄保护处理后的图像

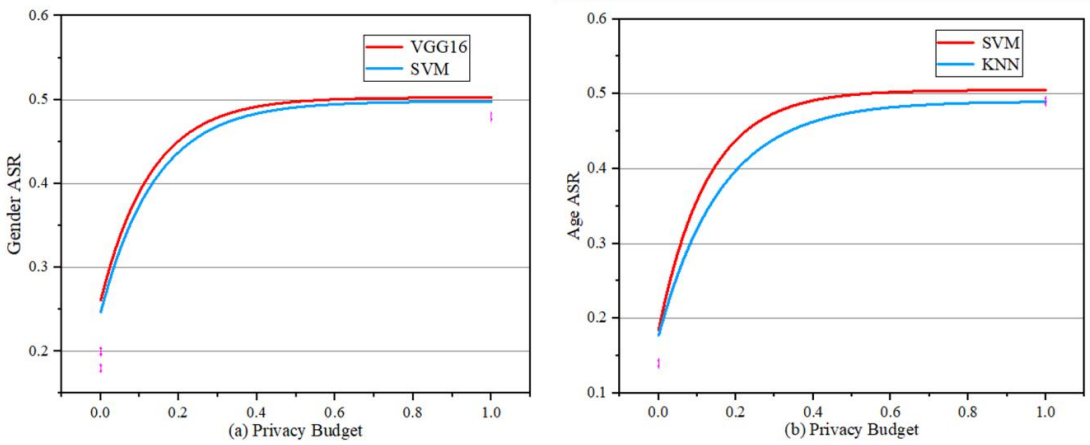


图 6 不同隐私预算下，属性保护图像的黑盒攻击成功率趋势

5.结束语

在本文中，我们提出了一种利用扩散模型的创新性面部隐私保护方法，它巧妙地融合了扩散模型的核心优势，实现了对身份与属性信息的深度隐藏。通过引入身份编码器和属性编码器，结合扩散模型强大的生成能力，我们成功地将面部进行了分解，为隐私保护开辟了新路径。为了进一步增强隐私保护的力度，我们引入了拉普拉斯噪声机制，以此为基础构建了差分隐私框架，显著提升了系统的保密性能。最后，由 DDIM 生成高质量的面部隐私保护图像。实验数据充分证明，实验结果表明，我们的方法不仅在隐私保护方面表现出色，而且保持了图像质量，并与各种计算机视觉任务兼容。

参考文献

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4690-4699).

Hiller, J., Schuldes, M., & Eckstein, L. (2019, October). Recognition and pseudonymization of data privacy relevant areas in videos for compliance with GDPR. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (pp. 2387-2393). IEEE.

Hukkelås, H., Mester, R., & Lindseth, F. (2019, October). Deepprivacy: A generative adversarial network for face anonymization. In *International symposium on visual computing* (pp. 565-578). Cham: Springer International Publishing.

Ismail, I. A. (2025). Protecting Privacy in AI-Enhanced Education: A Comprehensive Examination of Data Privacy Concerns and Solutions in AI-Based Learning. *Impacts of Generative AI on the Future of Research and Education,* 117-142.

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196.*

Karras, T., Laine, S., & Aila, T. (2019). *A style-based generator architecture for generative adversarial networks.* In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401-4410).

Leslie, D. (2020). Understanding bias in facial recognition technologies. *arXiv preprint arXiv:2010.07023.*

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. T*echnology, Knowledge and Learning, 27*(2), 405-427.

Maximov, M., Elezi, I., & Leal-Taixé, L. (2020*).* Ciagan: Conditional identity anonymization generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5447-5456).

Nitzan, Y., Bermano, A., Li, Y., & Cohen-Or, D. (2020). Face identity disentanglement via latent space mapping. *arXiv preprint arXiv:2005.07728.*

Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10619-10629).

Wang, T. C., Mallya, A., & Liu, M. Y. (2021). One-shot free-view neural talking-head synthesis for video conferencing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10039-10049).

Wen, Y., Liu, B., Ding, M., Xie, R., & Song, L. (2022). Identitydp: Differential private identification protection for face images. *Neurocomputing, 501*, 197-211.

基于生成式人工智能的集体反思中教师合作对话模式的认知网络分析

Capturing Teachers' Collaborative Discourse Patterns in Collective Reflection based Generative Artificial Intelligence: a network signature approach

赵雅[1]，王宁[2]，龙陶陶[3*]
[123]华中师范大学
[*] taotaolong@mails.ccnu.edu.cn

【摘要】 集体反思对于职前科学教师的专业发展具有重要意义。然而，当前研究表明，集体反思面临着认知失调、反思不深入和效果无法评估等多方面的挑战。如今，人工智能赋能教师专业发展已成为教育变革的重要议题。因此，本研究以14名职前科学教师为研究对象，采用认知网络分析法研究不同群体在基于生成式人工智能的集体反思过程中合作话语的发展模式。研究结果显示使用ChatGPT的小组在集体反思中偏重于包容不同观点，但不能有效地提供批判性反馈。研究为教师开展生成式人工智能下的集体反思提供了借鉴，为分析参与者的话语提供了新思路。

【关键词】 集体反思；生成式人工智能；教师合作对话；认知网络分析

Abstract: *Collective reflection is crucial for pre-service teachers' professional development. However, current research reveals that it is a challenge for facilitating collective reflection due to its superficial cognition and intangible quality. With the development of generative artificial intelligence in education, this study utilized epistemic network analysis to explore the collaborative discourse patterns of 12 pre-service science teachers in collective reflection scaffolded by GAI Results show that the ChatGPT groups demonstrated less active on providing critical feedback and analyzing the causes of problems combined pedagogical knowledge and classroom observation. This study offers insights for teachers engaging in collective reflection with GAI and provides new perspectives for analyzing participants' collaborative discourse patterns.*

Keywords: collective reflection, generative artificial intelligence, teachers' collaborative discourse, epistemic network signature

1.前言

《教育部办公厅关于加强小学科学教师培养的通知》（教师厅函[2022]10号）对小学科学教师的教学设计与课堂掌控能力提出了明确要求，强调创新教学方法、强化教学实践，并着重提升师范生的项目式教学、跨学科教学与反思等实践能力（中华人民共和国教育部，2022）。其中，教师的集体反思能力作为教师专业能力的重要组成部分，受到广泛关注。集体反思是指教师对已发生的教学活动及其后果进行合作探究与分析的能力，研究发现其对修正教学策略、提升教学实践效果及促进教师专业发展具有重要作用（Moon, 2013）。然而，集体反思存在概念模糊、结构松散和反思不深入等问题，给职前教师的集体反思带来了诸多挑战（Mynott, 2019; Brown et al., 2021）。

近年来，生成式人工智能在促进教师反思方面展现出巨大潜力，能够帮助教师在反思过程中生成新想法，优化教学实践（Kim & Lee, 2024；Seo et al., 2024）。但鉴于集体反思中

教师合作的复杂性和动态性，目前对其内在过程机制的研究仍较为匮乏（Segal, 2023）。教师合作对话作为揭示教师合作机制的重要领域，已成为当前研究的热点（Moe, Consiglio & Katz, 2022）。基于此，本研究以 14 名参与集体反思活动的职前科学教师为研究对象，采用认知网络分析法（ENA），分析其在基于生成式人工智能的集体反思活动中的合作对话模式的发展过程，旨在提升教师的反思能力，为生成式人工智能背景下的集体反思实践提供借鉴，同时为动态表征与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. 集体反思

集体反思是指参与者通过互动，提供多样化观点并贡献反馈与建议，以改进实践、制定方案及支持同伴成长（Jiang & Zheng, 2021）。在此过程中，参与者先分享自身经验，再通过同伴间的分享与协商，达成对相关观点的共同理解（Prilla & Blunk, 2020）。与个体反思不同，集体反思与社会互动紧密相连（Lord et al., 2017）。批判性反思是实现有效集体反思的关键，它能够帮助参与者理解对话背景并积极参与有组织的讨论。然而，职前教师因缺乏教学实践经验，往往难以开展批判性集体反思（Marcos et al., 2011)。因此，提供有效的支架和指导对于促进职前教师的集体反思及其专业发展至关重要。

2.2. 生成式人工智能与教育

生成式人工智能（GAI）是基于先进大型语言模型（如 ChatGPT 和聊天机器人）生成文本、图像或视频的技术。在教师教育中，生成式人工智能够提供个性化教学体验、优化教学策略，并促进教师在写作和知识整合等技能上的发展（Cooper, 2023）。然而，其应用面临诸多挑战，包括隐私与自主性风险、性别与种族偏见以及表现不准确等问题（Karan & Angadi, 2024）。此外，学习者过度依赖生成式人工智能获取知识，可能削弱其与现实社会互动的能力（Lan, 2024）。因此，研究如何在教师教育中有效应用生成式人工智能，以助力教师确认专业身份并适应人工智能时代的教育改革，具有重要意义。

综上所述，当前研究已逐渐认识到，在职前教师中开展有效的集体反思面临诸多挑战，而生成式人工智能具有促进其集体反思的潜力。然而，关于职前科学教师基于生成式人工智能的集体反思中合作对话模式发展的研究仍相对匮乏。因此，本研究旨在通过回答以下问题来研究两次基于生成式人工智能的集体反思活动中职前科学教师合作对话模式的发展过程：

（1）在微格教学后的集体反思中，职前科学教师是如何使用 ChatGPT 促进教师合作对话的产生和发展呢?

（2）在传统课堂教学后的集体反思中，职前科学教师是如何使用 ChatGPT 促进教师合作对话的产生和发展呢?

3.研究设计

3.1. 研究背景

本研究依托华中地区某师范大学的一门为期八周的面向大学二年级科学教育专业本科生的必修课程《信息技术教育应用》开展，旨在帮助职前科学教师深入理解信息技术在科学课中的应用。此外，此课程要求职前科学教师为小学生设计一门融入信息技术的科学课，并在学期末到小学完成科学课的教学工作。

3.2. 研究对象

本研究选取华中地区某师范大学 14 名大学二年级科学教育专业的本科生为研究对象。在参与本研究之前，他们已经学习了教学设计、教学管理和微格教学等相关课程。参与者自行

分为四个集体反思小组，每组 3 到 4 个人，公共完成设计和开展结合信息技术的小学科学课的任务。在整个研究过程中，小组分配保持稳定，此前，没有人有过集体反思的经验。

3.3. 研究过程

在本研究中，第 1 组和第 2 组的职前科学教师使用 ChatGPT 4.0 进行集体反思，并要求完成优化教学设计和评估教学表现的任务，第 3 组和第 4 组则仅通过讨论和协作进行集体反思和完成相关任务。在集体反思中，ChatGPT 被用来分析之前科学教师教学表现的优点和不足之处，优化教学设计以及提出创新性的教学方案。此外，本研究包含两次集体反思活动。第一次发生在微格教学之后，第二次发生在传统的课堂教学实践后。微格教学于第 4 周在微格教室进行，模拟真实教学，小组成员依次试讲，其他成员观察。课堂教学实践在课程最后一周进行，地点为一所小学的科学课堂，每组选派一名成员授课，其余成员观察。两次教学活动结束后，各组均开展集体反思，且集体反思的过程均被录音，并在所有参与者同意的情况下进行了文本转录。每次集体反思的平均时长为 50-60 分钟。

3.4. 数据收集与分析

本研究采用 Saunders 等人于 2020 年修订的教育生产性框架，对两次集体反思中四个小组的话语数据进行编码（Saunders et al., 2023）。原始的编码包括发现教学实践问题、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评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，因此为了评估集体反思研究中职前科学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（Fields et al., 2025）。本次编码由本研究团队中具有丰富编码经验的两位研究员独立完成，计算编码信度系数 R。如果编码信度系数 R 小于 0.8，则进行编码员培训，并针对歧义部分展开讨论。经过两轮培训，两位编码员对职前科学教师集体反思的话语数据的编码信度系数 R 为 0.85，说明编码结果可信，保证了研究结果的科学性（Fleiss, 1981）。

本研究采用认知网络分析法来探究职前科学教师在集体反思中使用 ChatGPT 4.0 的合作对话模式，并采用在线认知网络分析工具（http://app.epistemicnetwork.org）生成认知网络模型，模型中的节点分别为表 1 中职前科学教师合作对话的要素，节点大小代表了要素在集体反思对话中出现的频率，节点之间线条的粗细代表了元素之间的关联强度（Shaffer, 2017）。

表 1 集体反思中职前科学教师的合作对话编码表

维度	释义	编码
发现教学实践问题	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
支持和批评他人的观点	表达赞同或提出质疑	SC
维护成员关系	维持秩序、提供鼓励	MR

4.数据分析

4.1. 微格教学后集体反思中职前科学教师的合作对话模式

在微格教学后的集体反思中，本研究一共产生了1781个编码来分析职前科学教师在微格教学后的集体反思中的合作对话模式。使用ChatGPT的小组和未使用ChatGPT的小组在合作对话编码上呈现出多方面的差异。在使用ChatGPT的小组中，出现频率最高的对话编码是“包容和思考其他人的想法”（37.25%），其次是“发现教学实践问题”（23.15%）和“寻求解决方法”（14.11%），而在未使用ChatGPT的小组中，出现频率最高的对话编码是“发现教学实践问题”（25.27%）和“包容和思考其他人的想法”（25.27%），其次是“支持和批评他人的观点”（21.35%）。研究结果表明职前科学教师在微格教学后的集体反思中，主要聚焦于提出教学实践问题和表达不同角度的观点。相较于未使用ChatGPT的小组，使用ChatGPT的小组更注重从不同的视角表达观点，而对观点的支持性论证或批判性质疑的关注相对较少。

职前科学教师在微格教学后集体反思的合作对话认知网络模型如图1所示。为了进一步探究职前科学教师在微格教学后集体反思的合作对话模式的差异，本研究先对比使用ChatGPT的小组和未使用ChatGPT的小组质心的分布情况。图1的左边展现了职前科学教师在微格教学后集体反思的合作对话认知网络的质心和叠减图。红色的方块表示未使用ChatGPT的小组的质心，蓝色的方块表示使用ChatGPT的小组的质心。使用ChatGPT的小组的质心位于X轴的右侧，表明使用ChatGPT的小组在微格教学后的集体反思中更多地关注“包容和思考其他人的想法”，而未使用ChatGPT的小组的质心位于X轴的左侧，表明未使用ChatGPT的小组在微格教学后的集体反思中更多地关注“结合理论知识分析问题产生的原因”。研究结果表明，使用ChatGPT的小组更注重从不同的角度表达观点，而未使用ChatGPT的小组则更多地从教育知识的角度探究问题产生的原因。

图（a）呈现了使用ChatGPT的小组在微格教学后集体反思的合作对话认知网络模型，以PoP为核心，与MV、GO和SC具有强连接，其次是RoP和MR，之后是是连接强度略弱一点的PR。此外，MV和SC、GO、RoP之间也具有较强的连接。图（b）呈现了未使用ChatGPT的小组在微格教学后集体反思的合作对话认知网络模型，以PoP为核心，与MV、SC和GO具有强连接，其次是RoP和PR，之后是MR。此外，SC和MV、GO以及MV和GO之间也具有较强的连接。这些结果表明，使用ChatGPT的小组和未使用ChatGPT的小组在微格教学后的集体反思中都关注于发现教学实践中存在的问题，并围绕这些问题提出不同的观点，并积极地寻找问题的解决方案。使用ChatGPT的小组在此次集体反思中对观点的支持性论证或批判性质疑相对较少，而未使用ChatGPT的小组批判性更强，并能结合教育方面的知识分析问题产生的原因。

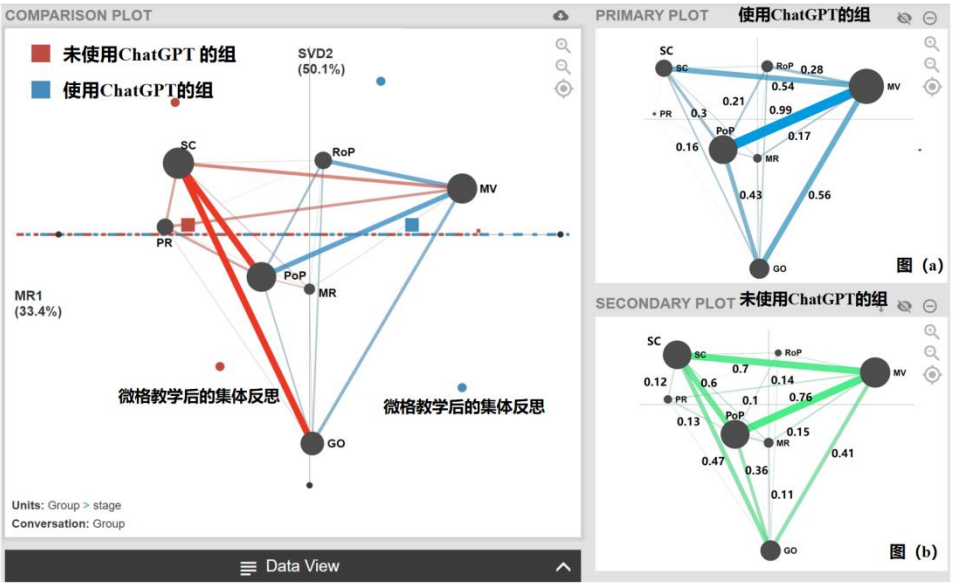


图 1 使用 ChatGPT 的小组和未使用 ChatGPT 的小组在微格教学后集体反思的认知网络模型

4.2. 传统课堂教学集体反思中职前科学教师的合作对话模式

在课堂教学后的集体反思中，本研究一共产生了 1435 个编码来分析职前科学教师在课堂教学后的集体反思中的合作对话模式。使用 ChatGPT 的小组和未使用 ChatGPT 的小组在合作对话编码上呈现出多方面的差异。在使用 ChatGPT 的小组中，出现频率最高的对话编码是“包容和思考其他人的想法”（43.00%），其次是“支持和批评他人的观点”（24.01%）和“发现教学实践问题”（18.74%），而在未使用 ChatGPT 的小组中，出现频率最高的对话编码是“包容和思考其他人的想法”（27.68%），其次是“发现教学实践问题”（26.59%）和“支持和批评他人的观点”（22.71%）。研究结果表明职前科学教师在传统课堂教学后的集体反思中，主要聚焦于提出教学实践问题，围绕这些问题表达不同角度的观点，并提供批判性的反馈。相较于未使用 ChatGPT 的小组，使用 ChatGPT 的小组更多地关注“包容和思考其他人的想法”，而对“结合理论知识和表征性数据分析问题产生的原因”关注较少（0.90%，6.29%）。这些结果表明，使用 ChatGPT 的小组更注重从不同的视角表达观点，而忽略了从教育知识和课堂观察的角度分析问题产生的原因。

职前科学教师在传统课堂教学后集体反思的合作对话认知网络模型如图 2 所示。图（c）呈现了使用 ChatGPT 的小组在传统课堂教学后集体反思的合作对话认知网络模型，以 PoP 为核心，与 MV 和 SC 具有强连接，其次是 RoP 和 GO，之后是是连接强度略弱一点的 PR 和 MR。此外，MV 和 SC 之间也具有较强的连接。图（d）呈现了未使用 ChatGPT 的小组在传统课堂教学后集体反思的合作对话认知网络模型，以 PoP 为核心，与 MV 和 SC 具有强连接，其次是 RoP 和 PR，之后是 GO 和 MR。此外，SC 和 MV 之间也具有较强的连接。这些结果表明，使用 ChatGPT 的小组和未使用 ChatGPT 的小组在传统课堂教学后的集体反思中都关注于发现课堂教学中存在的问题，并围绕这些问题提出不同的观点，积极地提供批判性的反馈，但忽略了寻找问题的解决方案。使用 ChatGPT 的小组在此次集体反思中对问题产生原因的分析相对较少，而未使用 ChatGPT 的小组批判性更强，并能结合教育知识和课堂观察分析问题产生的原因。

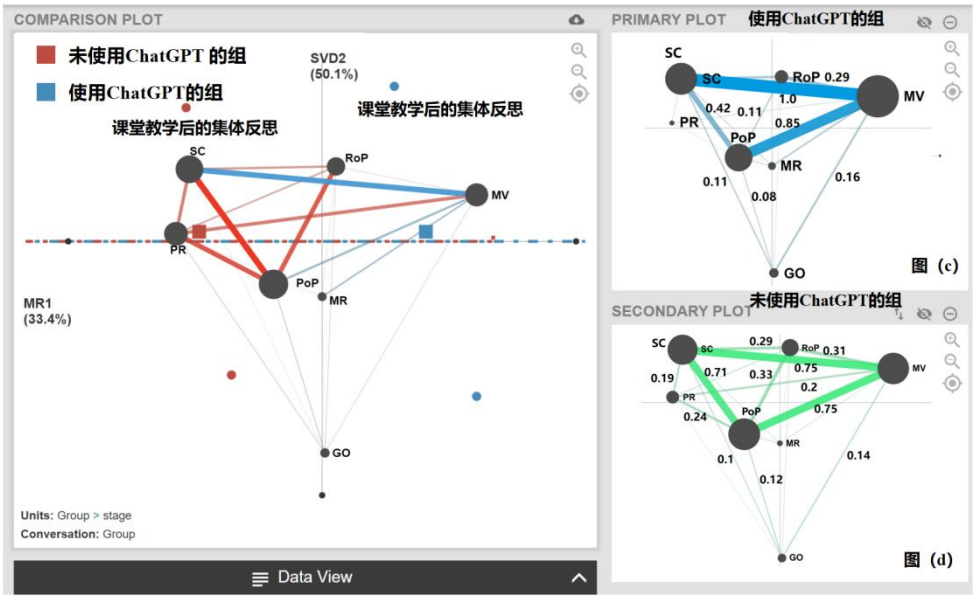


图 2 使用 ChatGPT 的小组和未使用 ChatGPT 的小组在传统课堂教学后集体反思的认知网络模型

5.结论与展望

本研究发现在微格教学后和传统课堂教学后的两次集体反思中，使用 ChatGPT 的小组和未使用 ChatGPT 的小组都注重发现教学实践中的问题，并围绕问题从不同的角度发表观点。在微格教学后的集体反思中，使用 ChatGPT 的小组和未使用 ChatGPT 的小组更偏重寻求问题的解决方案，而在传统课堂教学后的集体反思中，更偏重提供批判性的反馈。同时，在两次集体反思中，使用 ChatGPT 的小组更偏重包容和思考他人从不同角度发表的观点，而未使用 ChatGPT 的小组批判性更强，偏重于结合教育知识和课堂观察分析问题产生的原因。

5.1. 使用 ChatGPT 的小组集体反思缺少批判性思考

研究结果显示，使用 ChatGPT 的小组在集体反思中，较少地关注“支持和批评他人的观点”。这一结果与生成性人工智能可能不利于培养学生的批判性思维的研究结果一致（Qadir, 2023）。由此推断，当在集体反思中使用 ChatGPT 时，职前科学教师依赖 ChatGPT 去产生观点，并相信 ChatGPT 的回复，这阻碍了他们进行批判性的思考。批判性思维不仅是一种高阶思维能力，而且可以促进教师的专业发展（Leibovitch et al., 2025）。Sullivan 等学者指出，基于批判性思维的教师反思能够促使教师从不同的角度审视教学实践，开展教学推理，进而破解教学难题并实现实践革新（Sullivan et al., 2016）。然而，当职前科学教师被要求使用生成性人工智能进行集体反思时，由于缺乏教学经验和优秀在职科学教师的指导，他们很容易直接相信生成性人工智能工具给予的回复，却忽视了从教学实践的角度对其进行批判性思考。这种技术依赖导致使用 ChatGPT 的小组中“发现实践问题”与“支持和批评他人的观点”的关联强度相对较弱，反映出批判性对话维度的弱化倾向。

5.2. 使用 ChatGPT 的小组集体反思更偏重包容和思考他人观点

在微格教学后和传统课堂教学后的两次集体反思中，使用 ChatGPT 的小组和未使用 ChatGPT 的小组中的职前科学教师都偏重于“包容和思考其他人的想法”，他们对多元观点生成持更积极的态度，有助于形成更包容的良好合作关系。这与学生与生成性人工智能的互动能够改善他们与同伴和老师的合作效果的研究结果相一致（Florido & Hernández-Leo, 2024）。此外，使用 ChatGPT 的小组的“包容和思考其他人的想法”出现的频率较高，表

明职前科学教师对生成性人工智能保持积极的态度。依据 Fredrickson 的拓展-建构理论，积极情感能扩展个体的认知行为储备，促进个人资源在团体中的整合运用（Fredrickson，2002）。本研究的职前科学教师在集体反思中使用生成性人工智能时保持着积极的情感，这促进他们对多元的观点持更包容的态度，积极整合个人资源到小组合作中。这种生成性人工智能赋能的集体反思模式揭示了生成性人工智能工具在促进包容和合作方面的独特价值，但其对认知深度的潜在影响仍需审慎考量。

5.3. 微格教学后的集体反思偏重寻找解决方案，传统课堂教学后的集体反思偏重批判性思考

在微格教学后的集体反思中，使用 ChatGPT 的小组和未使用 ChatGPT 的小组的集体反思围绕着发现问题和寻求问题解决方案而展开；在传统课堂教学后的集体反思中，使用 ChatGPT 的小组和未使用 ChatGPT 的小组的集体反思主要围绕着提出问题和提供批判性的反馈。由此推断，小学科学课堂教学实践的任务促使职前科学教师在微格教学后的集体反思中系统梳理微格教学中发现的问题，并积极寻找解决方案，以更好地完成传统课堂教学。这一结果拓展了微格教学后的集体反思关注教学设计的优化和教学问题的解决有助于提升后续教学实践质量的研究（Vígh, 2024）。因此，以教学实践为落脚点持续推进职前科学教师的专业发展，将微格教学的体验和传统课堂教学的实践无缝融入到职前科学教师的基础教育中，是培养优秀科学教师的关键方向。

6.总结与展望

职前科学教师在集体反思中使用 ChatGPT 的合作对话模式映射了当下生成式人工智能融入教师教育领域的成效以及职前科学教师对未来使用生成式人工智能的期许。研究结果发现，职前科学教师在集体反思中使用 ChatGPT 时，偏重于包容和思考他人从不同角度表达的观点，但对观点的批判性较弱，他们对于在集体反思中使用生成式人工智能保持积极的态度。人工智能为职前教师的专业发展提供了崭新的思路，进一步推动高校人工智能教育教学资源建设，改进高校人工智能教师教育培养模式，满足职前教师专业发展需求，是高校使用人工智能促进教师教育的重要路径。

参考文献

教育部办公厅关于加强小学科学教师培养的通知(教师厅函[2022]10 号)[A/OL].[2023-05-05].http://www.gov.cn/zhengce/zhengceku/2022-05/28/content_5692779.htm.。

Brown, C., Poortman, C., Gray, H., Groß Ophoff, J., & Warf, M. (2021). Facilitating collaborative reflective inquiry amongst teachers: What do we currently know? *International Journal of Educational Research*, *105*, Article 101695.

Cooper, G.(2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*, 444-452.

Fields, K. L., Ethridge, E. A., Raymond, K. M. (2025). The power of role models for the development of teacher advocates. *Teaching and Teacher Education*, *155*, 104903.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley.

Florido, H., & Hernández-Leo, D. (2024, August). Generative AI Collaboration in the Orchestration of Supervised Classroom Problem Solving. In *International Conference on Collaboration Technologies and Social Computing* (pp. 257-264). Cham: Springer Nature Switzerland.

Fredrickson, B. L. (2002). Positive emotions, in: C. R. Snyder & S. J. Lopez (Eds) *Handbook of positive psychology (Oxford, Oxford University Press),* 120-134.

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.

Karan, B., & Angadi, G. R. (2024). Potential risks of artificial intelligence integration into school education: A systematic review. *Bulletin of Science, Technology & Society*, *43*(3-4), 67-85.

Kim, S., Lee, Y.G. (2024). Developing a Teacher Self-Reflection E-portfolio Platform for the Improvement of Noticing Skills of Childcare Teachers. In: Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (eds) HCI International 2024 Posters. HCII 2024. *Communications in Computer and Information Science*, vol 2117.

Lan, Y. (2024). Through tensions to identity-based motivations: Exploring teacher professional identity in Artificial Intelligence-enhanced teacher training. *Teaching and Teacher Education*, *151*, 104736

Leibovitch, Y. M., Beencke, A., Ellerton, P. J., McBrien, C., Robinson-Taylor, C. & Brown, D. J. (2025). Teachers' (evolving) beliefs about critical thinking education during professional learning: A multi- case study. *Thinking Skills and Creativity*, *56*, 101725.

Lord, A. Y., Chen, M. P., Cheng, Y. Y., Tai, K. C., & Pan, W. H. (2017). Enhancing nutrition-majored students' reflective judgment through online collective reflection. *Computers & Education*, *114*, 298-308.

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.

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*, 103823.

Moon, J. (2013). *Reflection in learning and professional development: Theory and practice.* Routledge.

Mynott, J. P. (2019). Lesson study outcomes: A theoretical model. *International Journal for Lesson and Learning Studies*, *8*(2), 117e134.

Prilla,M.,Blunk,T. (2020). How does collaborative reflection unfold in online communities? An analysis of two data sets. *Comput. Support. Coop. Wor*k, 29,697-741.

Qadir, J. (2023). 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.

Saunders, W., Topham, T., Jensen, B., Marcelletti, D., McCarthy, K., Lee, L.(2023). What' s in a teacher team meeting? Testing pedagogically productive talk as a framework for teacher collaboration. *Teaching and Teacher Education*, *131*, 104176.

Segal, A. (2023). Rethinking collective reflection in teacher professional development. *Journal of Teacher Education*, Article 00224871231188702.

Seo, K., Yoo, M., Dodson, S., & Jin, S. H. (2024). Augmented teachers: K‐12 teachers' needs for artificial intelligence' s complementary role in personalized learning. *Journal of Research on Technology in Education*, 1-18.

Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.

Sullivan, B., Glenn, M., Roche, M., & McDonagh, C. (2016). *Introduction to critical reflection and action for teacher researchers: A step-by-step guide*. Routledge.

Vígh, T. (2024). Development of research skills through research-focused microteaching lesson study in pre-service teacher education. *Teaching and Teacher Education, 145*, 104618.

机器遗忘赋能安全智能教育应用实践

吴伊雨，刘毓，邝瑶雯，王涛*
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 湖北 武汉 430079
*tmac@ccnu.edu.cn

【摘要】“人工智能+教育”的深度融合依赖于对教育数据的无感采集与融合计算，这种数据密集型的研究范式极大地推动了智能教育的发展与进步。然而，随着基于人工智能的教育模型的广泛应用，一系列安全风险也随之浮现，这些风险覆盖了从数据采集、模型训练到输出结果的全过程。本研究通过系统分析，揭示了在这一过程中可能遇到的安全隐患：数据采集阶段存在数据正确性与安全性的双重挑战；模型训练阶段则面临算法偏见与模型安全性的风险；而在数据应用阶段，结果的准确性、公平性及安全性同样令人担忧；这些风险实际根源于数据问题。因此，本研究致力于探索先进技术手段以应对这些挑战，特别分析了机器遗忘技术在缓解安全风险方面的潜力，并通过人脸识别模型保护的实例进行了实践验证，从而为确保模型安全提供了切实可行的方法。
【关键词】人工智能教育模型；安全风险；机器遗忘；应用实践

1.引言

2024 年，全国教育工作会议提出“要不断开辟教育数字化新赛道，以智能化赋能教育治理，引领教育变革创新”(教育部, 2024)。人工智能以其广泛且强大的数据驱动应用的能力，不断与教育深度融合，通过对学生等教育参与者数据的无感采集、融合计算和及时反馈，在教育过程表征、教育主体刻画、教育规律挖掘、教育服务变革等方面发挥重要的作用(杨宗凯 & 李卿, 2020)，并衍生出学习行为分析、教育认知诊断、教育知识追踪、教育情感计算、学习投入识别等研究领域(王一岩 & 郑永和, 2022)；利用数据密集型科学的研究范式探究教育教学的内在机理和演变规律，助力智能教育发展与变革(郑永和 et al., 2021)。

国务院《“十四五“数字经济发展规划(国务院, 2022)中明确提出要提升数据安全保障水平，规范数据从采集、存储到处理、销毁的全生命周期管理。然而，基于人工智能的教育模型的应用在带来诸多便利和优势的同时，也引发了诸多安全风险和挑战，尤其是 AI 大模型的应用，更是将安全伦理问题推向了公众视野的中心，这些挑战广泛涵盖了数据隐私保护、算法偏见等关键问题(褚乐阳 et al., 2024)。这些安全风险不仅限于教育大模型，而是普遍存在于所有基于人工智能的教育模型中，贯穿于数据采集、模型训练、结果输出的全过程，对教育数据的安全性、教育公平性和伦理性构成了威胁。

因此，本研究旨在深入探讨基于人工智能的教育模型应用过程中可能产生的安全风险，探索有效的应对策略，旨在在保护教育数据安全的同时，充分利用深度学习技术的潜力，推动教育的数字化转型和现代化进程，为教育技术学的学术研究和实践应用提供参考。

2. 人工智能在教育领域应用的安全风险

人工智能在教育领域的应用虽然为教学活动、学习过程、教育管理以及教学评价等多个关键环节带来了显著的促进作用，在提升教育效率与质量方面展现出了积极的效能，但也引发了隐私泄露、算法偏见、教育不公平等一系列伦理问题。王佑镁等基于 4R 危机管理模型将教育人工智能的伦理问题归纳为软件相关风险、硬件相关风险、环境相关风险和人件相关风险四类(王佑镁 et al., 2024)，吴砥等从教育安全、教育公平、教育关系三个方面分析应用潜在伦理风险(吴砥 & 吴河江, 2024)，吴河江等从技术、内容、数据、算法四个角度进行了风险分析(吴河 & 江吴砥, 2024)，杨俊锋等从数据、算法和应用三层面对伦理问题进行了总结及成因分析(褚娟, 2024)。教育人工智能伦理风险问题正受到学界的广泛关注和积极研究。本研究旨在聚焦人工智能教育模型生成的全过程，深入剖析这一完整过程中潜藏的安全风险，具体涵盖模型数据采集、模型训练，模型输出三个环节。

2.1. 人工智能教育模型数据采集阶段的安全风险

第一，数据安全性面临诸多挑战。一方面，数据匿名化仍可能被识别，随着技术的发展以及恶意攻击者手段的不断升级，原本旨在保护用户隐私的数据处理方式可能失效，个人或相关主体的信息有暴露的风险(Ji et al., 2016)。另一方面，数据还可能遭受攻击注入错误的威胁，外部攻击者可能利用网络漏洞，蓄意往采集的数据中注入错误信息，破坏数据的完整性和准确性(Wang et al., 2024)。同时，存在植入利益集团的需求和价值取向数据的情况，某些商业利益集团可能为了推广自身的教育产品或理念，可能会通过不正当手段，将带有偏向性的数据混入正常采集的数据集中，这会使采集到的数据带有偏向性，从而影响后续模型的客观性与公正性。

第二，数据正确性方面存在隐患。就内容准确性而言，采集的数据本身可能由于多种原因与实际情况不符。例如，在采集教师教学评价数据时，若调查问卷设计不合理，问题表述模糊，就容易导致教师或学生对问题理解出现偏差，进而反馈的评价数据不能真实反映教学实际情况。而且，采集设备的故障、人工录入的失误等也会造成数据内容错误。而且标注错误也较为常见，例如对于学习者情感状态的判定，存在着主观的偏差，错误的标注会干扰模型对数据特征的准确学习，进而影响整个教育模型应用的效果与可靠性。此外，数据的更新如果滞后于实际情况的变化，这也会导致数据的准确性下降。

2.2. 人工智能教育模型训练阶段的安全风险

第一，算法偏见问题凸显，尤其在深度学习自动化决策过程中。由于训练数据可能存在的不平衡、不全面等情况，若训练数据本身存在不平衡的情况，例如在采集用于智能教育资源分配模型的数据时，优质教育资源地区的数据占比过多，而偏远地区的数据过少，就会导致模型在后续决策时更倾向于为资源丰富地区的学生分配资源，忽视了其他地区学生的需求，产生地域上的偏见。或者算法本身设计的不合理，容易导致模型在决策时产生偏向性，不能公平、客观地对待不同的输入情况，影响教育模型应用在不同场景和用户群体中的适用性。第二，模型安全性方面，模型信息泄露与攻击的风险时刻存在(Tramèr et al., 2016)。经过深度学习和训练的数据，实际上会被模型以某种形式“记忆”并存储在其参数和结构中。这意味着，一旦模型的关键信息被泄露，可能会被不法分子利用，进行恶意篡改、破坏等操作，而针对模型的攻击行为更是会直接影响其正常训练以及后续的应用效果，给教育模型应用带来极大的安全威胁。

2.3. 人工智能教育模型结果输出阶段的安全风险

第一，结果的准确性存在问题，深度学习输出的错误与误导情况时有发生。这可能是因为模型在训练过程中没有充分学习到准确的模式，或者受到了异常数据的干扰等原因，导致

输出的结果与实际应有的正确结果存在偏差，误导教育相关人员的判断与决策。随着时间的推移，数据可能变得过时或不再具有代表性。使用过时的数据进行模型训练或决策分析，可能会导致不准确的结果。此外，深度学习输出的偏差现象较为突出。不同的输入条件或者外部环境变化时，模型输出的结果可能出现较大波动，缺乏稳定的输出性能，不利于在教育场景中持续、可靠地发挥作用。

第二，结果的公平性方面也面临挑战，深度学习输出的歧视问题不容忽视。可能由于训练数据中隐含的偏见等因素，使得模型对不同群体输出的结果存在不公平对待的情况，这在教育公平性要求较高的应用场景中是极为不利的，会破坏教育模型应用的公正性和合理性。第三，结果的安全性，模型的输出本质上是对输入样本进行预测或分类后得到的类别标签或概率向量。这些输出结果看似简单直接，然而近年来的研究却揭示了一个令人担忧的现象，模型输出结果隐含一定的数据信息，攻击者可以利用模型输出在一定程度上窃取相关数据(Zhu et al., 2022)。即使不直接访问模型的内部参数或训练数据，攻击者仍然有可能通过分析模型的输出来窃取相关的敏感信息。

3. 应对策略

人工智能教育模型生成的各个阶段紧密相连，其中存在的风险问题相互交织、彼此影响，在模型数据采集阶段，数据的质量、来源以及处理方式等因素，会对后续的算法训练及最终结果产生直接且深远的影响。而算法本身作为连接数据与结果的关键环节，同样起着至关重要的作用，其自身的特性和设计缺陷也会给结果带来诸多风险。隐私泄露这一风险隐患更是如影随形，贯穿始终，需要我们全面、深入地加以重视并积极探寻有效的应对策略，以保障教育模型应用的健康、稳定与安全发展。

3.1. 机器遗忘的应用潜力与实践

3.1.1. 机器遗忘的概念

在计算机领域，“机器遗忘”作为一种新的技术范式，旨在从机器学习模型中删除先前学习过的信息，将其影响从训练好的模型中消除，就像它从未存在过一样，用并保证遗忘后的机器学习模型性能几乎保持不变(Xu et al., 2023)，如图 12 所示。从本质上讲，它与机器学习正好相反：机器学习的目的是训练模型来进行识别、分类和预测，而机器遗忘的目的则是撤销或逆转这一过程，确保在特定情况下，能够有选择性地从模型中移除不再需要或敏感的信息。机器遗忘分为精确遗忘与近似遗忘，精确遗忘包括重训练和分片训练，即从数据集中直接删除数据进行重训练或将数据划分为分别训练子模型，删除数据后对子模型进行重训练，但这种方法计算成本较高，尤其是当数据集庞大时；近似遗忘，通过对模型参数的修改实现相近与重训练性能，也在一定程度上牺牲了准确性。

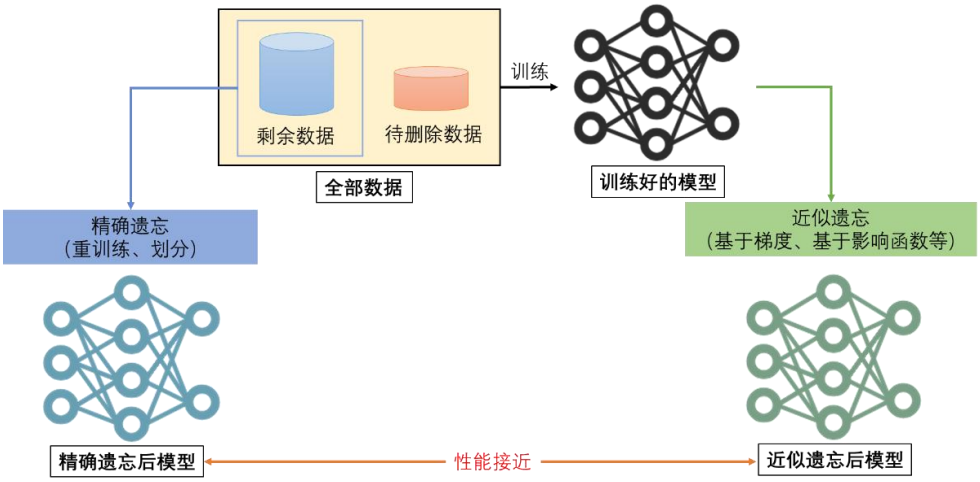


图 1 机器遗忘概念图

3.1.2. 机器遗忘解决人工智能教育模型安全风险的潜力分析

全过程的安全风险问题的根源深植于数据安全之中，若能有效销毁那些被玷污或涉及个人隐私的数据，将能在很大程度上规避相关风险。然而，必须明确的是，仅仅依赖数据库中的简单删除操作是远远不够的。在模型训练的复杂过程中，数据往往已被模型深度“学习”并“铭记”。这意味着，即便数据在数据库中被清除，其影响仍可能潜藏于模型的内部结构中，持续构成威胁。为了彻底清除这些敏感信息，必须诉诸更为先进的技术手段。在此背景下，机器遗忘技术展现出了应对教育模型应用中的安全风险的无限潜力。

对数据，一方面，它能够精准地去除不正确、带有偏见、受到攻击或被植入不合理价值观的数据信息，避免对模型产生不良影响，确保模型所依据的数据源头尽可能准确可靠。另一方面，机器遗忘可有效保障教育参与者的数据隐私和被遗忘权。在智能教育场景中，存在大量涉及教育参与者个人隐私的敏感信息，这些信息需要得到妥善保护。通过机器遗忘，在不影响智能教育模型结果准确性的前提下，能够有选择性地删除敏感的个人数据，使得这些数据不再留存于模型之中，从而充分尊重和保护教育参与者的隐私权益，从根本上杜绝因数据泄露等问题引发的安全风险。

对模型，机器学习模型在训练过程中容易受到历史训练数据的影响，若训练数据存在偏见，例如某些特定群体的数据占比过高或过低，就可能使模型产生算法偏见，进而在后续应用中做出不公平的决策。机器遗忘通过其独特的“撤销”或“逆转”功能，能够从模型中移除这些因历史数据偏差所带来的影响。它可以定期对智能教育模型进行分析，精准定位那些因历史训练数据中的偏见或错误而导致模型出现问题的部分，并将相关信息从模型中消除，使模型摆脱这些不良影响的束缚，恢复到更为客观、公正的状态，从而提升模型自身的准确性和公平性，保障模型在教育实践中的合理应用，降低因模型自身缺陷带来的安全风险。另外，机器遗忘技术通过从模型中移除数据信息的手段，还能有效防范利用模型参数反向还原并窃取数据的风险。这一功能进一步增强了模型的安全性，使得教育参与者的隐私和数据得到了更为周全的保护。

对结果，智能教育模型输出的结果对教育决策等有着重要的指导作用，但如果模型因数据或自身的问题产生不准确的预测或不公平的决策，会直接影响教育实践的公正性和有效性。机器遗忘通过去除数据中的错误信息以及消除模型因历史数据导致的偏差，能够间接提升智能教育模型输出结果的准确性和公平性。同时，机器遗忘使智能教育模型能及时更新、适应

新的教育情境，确保其结果始终保持有效性。在动态变化的教育环境中，过时的信息若持续影响模型，会使模型输出的结果与实际需求脱节，而机器遗忘帮助模型“忘却”这些不适用的信息，促使模型输出更贴合当下教育实际的结果，增强结果的稳定性和可靠性，更好地服务于教育实践，从结果角度有效防范各类安全风险，提升整个智能教育应用的质量与安全性。总而言之，通过机器遗忘技术，针对不同的需求移除特定数据信息，能够有效应对全过程的安全风险。

3.2. 机器遗忘的实施案例

3.2.1. 机器遗忘的实施流程

机器遗忘为遗忘算法技术的统称，在教育领域的延伸与应用，需进一步复杂化，形成一个更为系统化、综合性的概念框架，不仅触及技术层面的革新，更涉及整体设计与实施。

为了适应教育场景，本研究提出了机器遗忘的实施流程，基于法律法规、相关政策以及具体实施细则，包含三大核心阶段：审查阶段、执行阶段和反馈阶段，如图 13 所示。从数据利用与生成的视角来看，教育参与者可被分为数据使用者与数据提供者。数据使用者涵盖学校管理层、教育机构内的科研人员以及直接从事教学活动的教师，他们利用数据进行教育管理、学术研究及个性化教学等目的。而数据提供者则主要为学生和教师群体，他们在教学与学习活动中产生并提供了大量的数据，这些数据成为教育数据挖掘与分析的基础。

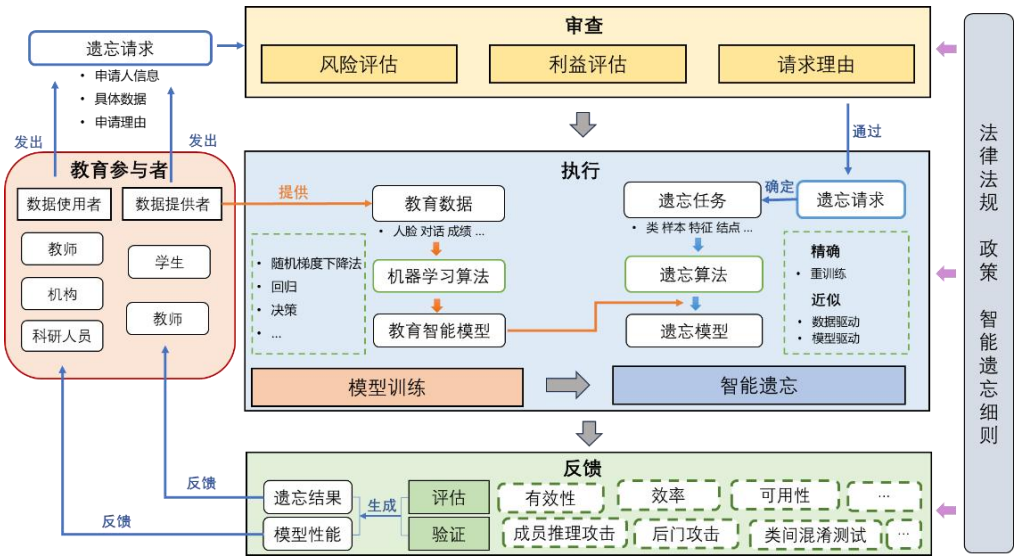


图 2 教育数据智能遗忘流程

审查阶段，考虑到数据销毁的彻底性，需要对请求进行谨慎审查。教育参与主体出于对隐私保护的需求或是为了纠正模型中的问题，依据要求提出明确的遗忘请求，包括基本信息、请求遗忘的数据及申请理由；为确保请求的合规性，需要对请求进行审查，评估请求遗忘的数据是否存在对应的安全风险，以及风险的大小，这包括对数据敏感性、数据泄露后的潜在影响以及数据保护措施的有效性进行综合考量；权衡该数据涉及的各方利益，这包括学生的个人隐私权、教育机构的教育管理权以及数据使用者的合法权益，通过综合考虑各方利益，审查团队将决定是否予以删除。以决定是否予以删除。

执行阶段，根据提出的不同级别的数据遗忘请求，确定对应遗忘任务包括遗忘样本、遗忘类集、遗忘特征。为实现数据遗忘，我们需选择合适的遗忘方法，如模型重训练、数据重组和模型修改等。这些方法的选择应基于数据的特性、模型的类型以及遗忘请求的具体要求。选择性地从原始数据集中以及训练和部署的教育智能模型（如学习分析模型、教育评测模型、

学习资源推荐模型等）中完整且迅速地删除对应的数据样本、类集和特征，实现准确或近似遗忘，并保证智能教育模型性能几乎保持不变。

反馈阶段，对遗忘模型进行评估验证，以确保模型的有效性、效率和可用性。评估模型的有效性、效率和可用性等，同时对遗忘模型进行验证，通过成员推理攻击、后门攻击、类间混淆测试等。根据评估验证的结果，生成遗忘结果报告和模型性能报告。这些报告将分别反馈给数据提供者和数据使用者，以确保信息的透明度和流程的完整性。同时，我们还将提供必要的解释和说明，以消除因信息滞后而产生的误解、担忧及损失。

3.2.2. 机器遗忘的应用实践：对人脸识别模型的保护

人脸识别是当前智能教育实现无感签到(方书雅 & 刘守印, 2020)、学生行为分析(周楠 & 周建设, 2021)、学习情感计算(徐振国 et al., 2019)等多个人工智能教育模型的核心模块。而人脸数据是一种重要的生物特征标识，与个人身份直接关联，是通用身份标识符(胡凌, 2021)，具有高度敏感性。本研究以人脸识别模型为例，实现对学生人脸数据的遗忘，以应对各环节的安全风险。

首先本研究构建了一个涵盖 54 名学生学习期间的人脸图像数据集，图 14 中展示了部分学生的人脸图像示例（为了合规，每张图片示例进行了打码处理），每名学生的面部图像根据其身份标识分别归类存储，经过这一系列的训练和优化，人脸识别模型在测试集上的准确度高达 99.94%，充分展示了其卓越的性能。

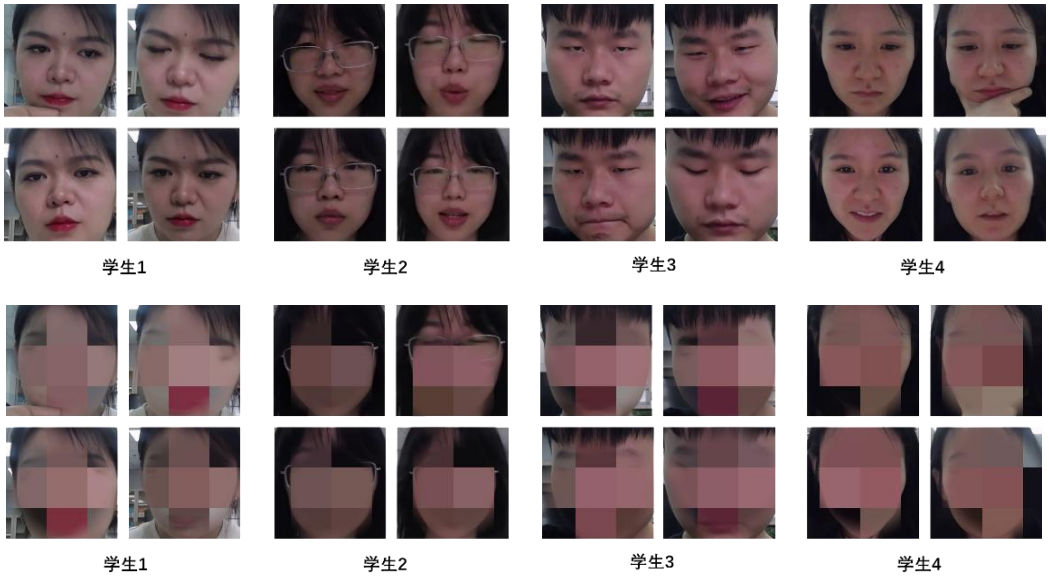


图 3 部分学生人脸图像

近似遗忘算法在遗忘过程中通常能达到约 80-90%的准确率，权衡准确确性和计算成本，近似遗忘更具成本效益(Sai et al., 2024)。基于此本文选择了基于模型修改的梯度投影遗忘方法(Hoang et al., 2024)应用于人脸数据的遗忘。

通过对数据集编号，依据学生提供的信息，定位学生对应人脸数据，以实现单位与多位学生人脸数据的智能遗忘。采用梯度投影遗忘算法进行遗忘后，分别给出了遗忘单位学生人脸数据和遗忘 5 位学生人脸数据的原始人脸识别模型、利用剩余训练集重训练的人脸识别模型、以及梯度投影遗忘后的人脸识别模型分别在原始测试集 D_{test}、剩余测试集 D_{r_test}、遗忘训练集 D_{u_train}和遗忘测试集 D_{u_test}的识别准确率以及运行的平均时间。此外，考虑到遗忘后的模型置信度，图为遗忘数据在遗忘模型与重训练模型上熵的分布。

表 1 人脸识别模型遗忘结果

模型 测试集	原始模型	重训练的模型	遗忘后的模型	重训练的模型	遗忘后的模型
		遗忘 1 位学生		遗忘 5 位学生	
D_{test}	99.94%	96.39%	98.03%	90.69%	81.94%
D_{r_test}	\	98.18%	99.88%	99.93%	89.30%
D_{u_train}	\	0.00%	0.00%	0.00%	1.07%
D_{u_test}	\	49.75%	49.75%	50.00%	49.6%
Average Time（s）	\	147.50	18.94	143.34	24.48

从遗忘效果来看，遗忘后的模型表现出了较高的遗忘水平，与重训练结果接近；遗忘模型与重训练模型上熵的分布情况，即遗忘训练集预测的混乱程度。从图中可以看出，重训练模型和遗忘后模型在遗忘数据上的混乱程度相当一致，这进一步验证了本研究遗忘结果的精确性。这一结果深刻地说明了机器遗忘技术能够精准地根据实际需求，移除那些有毒、过时、带有偏见或涉及个人隐私的数据，从而在确保数据正确性和安全性的同时，有效地防止了因数据问题而引发的算法偏见和结果异常。

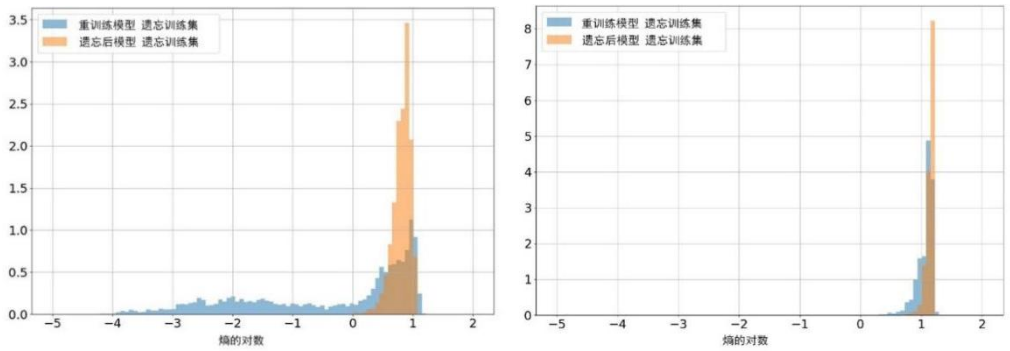


图 15 遗忘单位（左）、多位（右）学生人脸的遗忘数据在遗忘模型与重训练模型上熵的分布

从模型性能来看，遗忘单位学生模型性能没有受到显著影响，遗忘多位学生，虽然存在一定损失，但仍处于可接受范围内。这很可能是因为遗忘多位学生的人脸数据对模型的内部结构产生了更为显著的调整，从而导致了更大的性能波动。

从时间效率来看，采用近似遗忘方法显著提升了遗忘效率，重训练模型的运行时间约为近似遗忘后模型运行时间的 5-7 倍。这一优势使得数据遗忘操作在实际应用中更具实用性。

因此，通过修改模型的特定参数确实能够高效地实现数据的遗忘，在基于人工智能的教育模型的应用中，可以采用机器遗忘，移除特定数据及其在模型中的“影子”，有效应对因数据、算法、结果产生的安全风险问题，真正实现模型全流程保护。

4. 结语

人工智能在教育领域的应用已经展现出显著的潜力，它通过数据驱动的方式，提升了教学效果、实现了学习资源的个性化推荐、智能化了课堂管理，并创新了教学评价与反馈机制。然而，这些基于人工智能教育模型的应用也伴随着安全风险。这些问题不仅影响教育数据的安全性、教育公平性和伦理性，还可能对教育的数字化转型和现代化进程产生负面影响。针对这些安全风险，本研究提出了机器遗忘的概念和实施流程，旨在通过技术手段保护应对产

生的安全风险，保证数据安全，减少算法偏见，提升模型的准确性和公平性。机器遗忘技术在数据、模型和结果三个层面上都显示出了其应对安全风险的潜力。

当前，机器遗忘作为应对人工智能安全风险的技术手段，仍处于发展阶段，面临多重挑战。由于机器学习应用场景复杂，缺乏统一的遗忘算法，需根据具体教育场景和需求选择。同时，遗忘算法评价指标多样，缺乏统一标准，且可能引发数据隐私泄漏风险，数据安全性和隐私性保护成为技术难题。此外，我国法律体系内尚未明确被遗忘权，尽管《个人信息保护法》规定了个人信息删除权，但存在广泛争议，导致机器遗忘的应用缺乏法律指引，数据控制者面临法律不确定性和潜在风险。机器遗忘的实施还需权衡多方利益，特别是数据共享与个人隐私保护之间的冲突，以及不同主体间的利益平衡，以确保个人隐私得到保护的同时，合理利用数据，促进人工智能技术健康发展。

未来，将不断扩展研究的深度与广度，一方面，积极探索应对基于人工智能教育模型安全风险更有效的策略，深入研究机器遗忘以实现更高性能的安全保障；另一方面扩展机器遗忘在其他教育应用方面的作用，例如数据治理，不断促进智能教育的健康、稳定与安全发展。

参考文献

褚娟, 杨. A. (2024). 人工智能教育应用的伦理风险和规范原则. 中国教育学刊(11).

方书雅, & 刘守印. (2020). 基于学生人体检测的无感知课堂考勤方法. 计算机应用, 40(9), 6.

国务院. (2022). 国务院关于印发“十四五”数字经济发展规划的通知. 中华人民共和国国务院公报, 3, 5-18.

胡凌. (2021). 刷脸:身份制度、个人信息与法律规制. 法学家(2), 41-55.

教育部. (2024). 年全国教育工作会议召开. In.

王一岩, & 郑永和. (2022). 多模态数据融合: 破解智能教育关键问题的核心驱动力. 现代远程教育研究, 34(2), 93-102. https://doi.org/10.3969/j.issn.1009-5195.2022.02.011

王佑镁, 利朵, 王旦, & 柳晨晨. (2024). 基于 4R 危机管理模型的教育人工智能伦理风险防范体系构建. 中国电化教育(9), 32-40,85. https://doi.org/10.3969/j.issn.1006-9860.2024.09.006

吴砥, & 吴河江. (2024). 通用大模型教育应用的潜在风险及其规避——基于技术伦理的视角. 华东师范大学学报 (教育科学版), 42(8), 64.

吴河, & 江吴砥. (2024). 教育领域通用大模型应用伦理风险的表征, 成因与治理. Tsinghua Journal of Education, 45(2).

徐振国, 张冠文, 孟祥增, 党同桐, & 孔玺. (2019). 基于深度学习的学习者情感识别与应用. 电化教育研究, 40(2), 87-94.

杨宗凯, & 李卿. (2020). 计算教育学: 内涵与进路. 教育研究, 41(3), 152-159.

褚乐阳, 潘香霖, & 陈向东. (2024). AI 大模型在教育应用中的伦理风险与应对. Journal of Soochow University Educational Science Edition, 12(1).

郑永和, 王杨春晓, & 王一岩. (2021). 智能时代的教育科学研究:内涵、逻辑框架与实践进路. 中国远程教育（综合版）(6), 1-10,17.

周楠, & 周建设. (2021). 基于深度学习的学生行为分析与教学效果评价.

Hoang, T., Rana, S., Gupta, S., & Venkatesh, S. (2024). Learn to unlearn for deep neural networks: Minimizing unlearning interference with gradient projection. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision,

Ji, S., Mittal, P., & Beyah, R. (2016). Graph data anonymization, de-anonymization attacks, and de-anonymizability quantification: A survey. *IEEE Communications Surveys & Tutorials*, *19*(2), 1305-1326.

Sai, S., Mittal, U., Chamola, V., Huang, K., Spinelli, I., Scardapane, S.,…Hussain, A. (2024). Machine un-learning: An overview of techniques, applications, and future directions. *Cognitive Computation*, *16*(2), 482-506.

Tramèr, F., Zhang, F., Juels, A., Reiter, M. K., & Ristenpart, T. (2016). Stealing machine learning models via prediction {APIs}. 25th USENIX security symposium (USENIX Security 16),

Wang, F., Wang, X., & Ban, X. J. (2024). Data poisoning attacks in intelligent transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*, *165*, 104750.

Xu, H., Zhu, T., Zhang, L., Zhou, W., & Yu, P. S. (2023). Machine Unlearning: A Survey. *ACM Comput. Surv.*, *56*(1), Article 9. https://doi.org/10.1145/3603620

Zhu, T., Ye, D., Zhou, S., Liu, B., & Zhou, W. (2022). Label-only model inversion attacks: Attack with the least information. *IEEE Transactions on Information Forensics and Security*, *18*, 991-1005.

基于视频分析的职前科学教师集体反思：一项基于设计的研究

Collective Reflection of Pre-service Science Teachers with Video Analysis: A Design-Based Research

李吉梅[1]，龙陶陶[2*]，王梦珂[3]，徐莹[4]，赵雅[5]
[12345]华中师范大学
[*] taotaolong@mails.ccnu.edu.cn

【摘要】 集体反思是指群体成员在社会互动中进行反思的过程。它在促进教师专业发展方面发挥着重要作用。本研究采用基于设计的研究方法，通过设计、实施和评估三轮过程，探索利用视频记录和视频分析的集体反思方法来增强职前科学教师的教学实践。基于视频的智能多模态分析技术是帮助职前科学教师反思教学的关键技术。本研究揭示了如何利用基于视频的智能多模态分析报告进行批判性和深度的集体反思。
【关键词】 集体反思；视频分析；基于设计的研究

Abstract: *Collective reflection refers to the process through which group members reflect via social interaction. It has been applied as essential for improving teacher professional development. Using a design-based research method, this study conducted three rounds of designing, implementing, and evaluating processes, to explore a collective reflection approach with video records and video analysis to enhance pre-service science teachers' teaching practice. Video-based intelligent multi-modal analysis technology was critical in helping the pre-service science teachers reflect on their teaching. Ways to use the video-based intelligent multi-modal analysis reports to produce critical and in-depth collective reflection are described.*
Keywords: collective reflection; video analysis; design-based research

1.前言

职前教师因教学经验匮乏，在教学实践中常面临诸多挑战，其中不知如何高效反思尤为突出，这严重阻碍了他们的专业成长。集体反思，即社区成员借助社会互动展开共同反思的过程，能够助力教师从全新视角审视教学实践，完善教学技巧与方法，进而改进教学（Kager et al., 2022）。虽说集体反思在教师教育领域已获认可，可在实际推行时却困难重重。研究发现，教师课堂上的实际表现与他们对自身行为的认知存在偏差（Montgomery & Baker, 2007）。而且，教师在集体反思中缺乏清晰的反思框架，很难从浅层次反思迈向批判性反思，往往只是重复确认自身行为，而非深入反思教学实践与教育理念（Giroux，2024）。对科学教师而言，上述问题更为棘手（Zheng et al., 2023）。科学学科具有复杂性与跨学科性，科学教师在应对知识快速更新、实验教学多元挑战以及跨学科教学复杂局面时，教学反思难度剧增（Nielsen et al., 2016）。这使得他们难以开展深入的集体反思，也就无法有效解决教学实践中涌现的各类问题。当下研究显示，在全球范围内提升科学教师专业能力迫在眉睫（Zheng et al., 2023），可有关科学教师集体反思的研究却相对稀缺。因此，探索促进职前科学教师集体反思的有效路径刻不容缓。

有学者提出，利用课堂视频记录促进集体反思，对加深教师对教育理论的理解和应用、促进其专业成长具有重要潜力（Soni, 2023）。观看自己的教学视频，能让教师聚焦课堂关键问题，更有效地改进教学实践。然而，借助录像开展批判性、有组织的集体反思，其具体路径和原则还有待研究。为探索真实教育场景中的问题，基于设计的研究（DBR）方法通过迭代设计，提出可能的解决方案，以弥合理论与实际教育问题之间的差距（Wang & Hannafin, 2005）。鉴于科学教师对深入且具批判性的集体反思存在迫切需求，本研究采用 DBR 方法，旨在探究如何借助教学实录改善职前科学教师的集体反思，为其教学实践提供支持。

2.理论基础

2.1. 教师集体反思

最初，个体反思以及一对一指导的过程逐步向集体层面拓展，由此催生了集体反思（Moon, 2013）。个体反思主要依赖个体在社会环境之外的主观能动性，与之不同，集体反思体现的是个体借助社会互动所产生的认知感知（Vangrieken et al., 2015）。作为教师教育和教师专业发展的重要手段，集体反思能让教师从新的视角审视自身教学，完善教学方法与技能，获取教学改进的思路。不仅如此，集体反思还能促进教育者之间的协作与沟通，营造积极的学习氛围（Chen, 2023）。

然而在实践中，集体反思仍存在组织性欠缺、批判性不足等问题，这些问题阻碍了教师的专业发展（Mei et al., 2024）。在集体反思时，教师对教学实践的反思常常局限于一些浅显特征，难以形成结构性反思（Otienoh, 2011）。缺乏结构性和系统性的反思，既难以总结出有效的教学方法，也不利于教师之间分享和传承教学经验（Mei et al., 2024）。此外，不少教师在反思时仅停留在表面描述，缺乏从多元角度分析、评价教学实践的能力（Darling - Hammond & Snyder, 2000）。因此，探寻提升集体反思组织性与深度的有效途径，仍有待进一步研究。

2.2. 录像和录像分析在集体反思中的应用

随着信息技术的发展，视频技术在教师反思领域得到了广泛应用。它助力教师深入洞察课堂教学的微观层面（Koc et al., 2009）。教师通过视频观察自身及其他教师的教学行为后，能以视频为依据，对教学行为展开评价，教学反思能力也随之得以提升（Tripp & Rich, 2012）。录像视频在教师专业发展进程中至关重要，为教育工作者提供了有关自身教学实践的宝贵见解，同时也为反思、改进教学以及推动教学创新创造了契机（Koc et al., 2009）。

步入智能时代，多模态分析技术有望为提升教师集体反思能力开辟新路径。该技术能够自动识别教学过程中的关键场景与行为，例如教师的教学行为、学生的反应以及课堂互动情况（Soni, 2023）。基于视频的智能多模态分析，依托收集教师教学过程中的多模态数据得以生成，这些数据涵盖语言、行为和心理等方面。它聚焦于课程中教师与学生的言语互动和行为表现。针对语音识别、姿态检测、面部表情识别等任务，研究人员设计并运用算法对课堂教学进行定量分析，进而向教育工作者反馈教学情况。合理且广泛地收集课堂内的多模态数据，能够为教师开展教学反思、提升教学能力提供事实支撑（Sharma & Giannakos, 2020）。由此可见，利用基于视频的智能多模态分析来推动教师集体反思，深化教师对教学实践的理解，极具潜力。然而，目前将其应用于教师反思方面的研究却相当匮乏。

3.研究设计

3.1. 研究背景与研究对象

本研究选取了一个为期八周的志愿教学项目作为研究背景。在这个项目中，来自一所研究型大学的 20 名科学教育专业的本科生在一所资源薄弱的小学设计并教授了一系列校本科学课程。志愿教学项目遵循教学设计、模拟教学、教学实施和集体反思这一循环。在这所小学，每次教学实施时间为每周 40 分钟。在教学实施的第二天进行集体反思。

本研究选取该大学的 20 名科学教育专业的本科生为研究对象，其中有 11 名是大二学生，9 名是大三学生。二年级学生均无 K-12 教学经验。此前，没有人从事过与 K-12 教学相关的反思活动。

3.3. 研究设计

本研究遵循了教育设计研究方法（Van den Akker，1999）。根据 Plomp（2013）的观点，教育设计研究具有双重目的，即通过迭代过程产出高质量的成果，并生成设计原则。本研究旨在为职前科学教师的集体反思开发一套基于数据驱动的反思原则，从而促进他们教学实践的改进。本研究的主要研究问题是：如何进行集体反思，以提高职前教师的教学效能或实践？

如图 1 所示，本研究经历了三个阶段：初步研究、原型设计和评估。在前期研究阶段，教师集体反思进行了需求分析和文献综述。在原型设计阶段，进行了三轮迭代过程。实施过程中考察了职前科学教师的集体反思是如何进行的，以及他们遇到的挑战。评估的目的是收集职前科学教师的反馈，并探索如何进一步改善原型。评估阶段考察职前科学教师的教学自我效能感和对集体反思的设计与实施的看法。

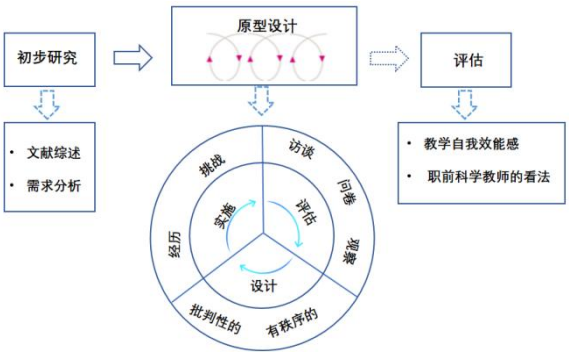


图 1 教育设计研究过程

4.原型设计过程

本研究发现当前的集体反思存在两大现实问题。第一，目前的集体反思缺乏批判性和深度，缺乏客观和全面的证据。第二是集体反思没有良好的秩序。本研究提出了以下两个原则来改善集体反思。首先，要提供脚手架和客观的证据，使集体反思更加具体和深入。其次，需要建立一个顺序，使集体反思过程更加有序和有组织。经过三轮的原型设计过程，这两个原则逐渐得到完善。

4.1. 第一轮原型设计

4.1.1 设计

在初步研究中产生的两个主要原则在原型设计过程中得到了应用。首先，需要指导方针和证据，使反思具有深刻和批判性。第二，集体反思要有序、有组织，促进职前教师积极参与。

4.1.2 实施

在第一轮原型设计中，进行了两次集体反思会议。反思开始时，所有人都观看录制的课堂视频。然后分发了一份反思指南，引导他们反思本课程的优缺点，并提出改进建议。随后，他们轮流交换了对这节课教学的看法，包括从这节课中学到了什么，以及今后如何改进这节课的教学。发言的顺序通常是上课的职前老师，助教，观察员，接着是没有参加教学的人。

4.1.3 评估

研究人员观察发现，职前科学教师在反思时没有参考或引用任何证据，这表明职前科学教师对自己和他人教学绩效的评价缺乏客观性和针对性。例如，参与者 A（一名大二的职前科学教师），在看他自己的教学视频时，他总是感到非常尴尬。他说他各方面都表现得很差。相比之下，参与者 B（一名具有更多教学经验的职前科学教师），他对自己的教学表现感到非常满意。在回忆自己的教学经历时，他似乎迷失了自己，自豪地说："我认为我的课是完美的。"然而，在研究团队看来，A 和 B 在教学中都有各自的优缺点，两人有不同的教学风格，但他们都没有客观地看待自己的教学。

集体反思中，往往只有一部分教学经验较丰富的职前科学教师能够分享自己对课堂教学的看法，而其他低年级的职前科学教师往往保持沉默。在访谈中，C（一名大一的职前科学教师）解释说："我几乎没有教学经验，所以我不知道我应该关注课堂教学的哪个方面。我感到很沮丧，不愿意再参与集体反思了。"一位经验丰富的成员 E 说："只有几个有教学经验的能言善辩的高年级学生主导了讨论，让新成员没有机会发言。"可以看出，几乎没有任何教学经验的新手难以参与集体反思。

研究团队讨论了集体反思环节中提到的问题。第一，职前科学教师的反思相对主观且不够全面；集体反思没有组织性，参与者普遍表现出较低的积极性。因此，研究团队做出了以下修改决定：R1.提供更客观、更全面的证据来支持教师的主张，使集体反思更具批判性。R2.进一步明确职前科学教师表达观点的顺序和规则，使集体反思更有组织性。

4.2. 第二轮原型设计

4.2.1 设计

在第二轮反思中增加了一种新工具——多模态分析报告，它能够为职前科学教师的教学表现提供全面、客观的证据[R1]。基于视频的多模态分析报告可以作为研究人员和教师探索其语言、手势和教学风格等教学行为的分析依据(Hung, 2016)。为了使集体反思井然有序，帮助参与者形成一个学习者社区，研究人员详细描述了反思的过程[R2]。在这一轮集体反思中，除了发言顺序外，还制定了更具体的规则和反思准则。会给每位参与者分发一份更具针对性的反思指导方针，该方针会明确指出职前科学教师应当从哪些方面进行教学分析和反思。此外，研究人员鼓励所有参与者选出一位具有丰富教学经验、善于沟通和组织的成员担任组长，以促进集体反思[R2]。

4.2.2 实施

本研究使用的多模态分析报告名为"益课"，由一所研究型大学的教育大数据研究中心开发，在该校已经投入使用了大约一年。"益课"包括 S-T 分析、课堂氛围分析、教师注意力、课堂管理等多个方面，如图 2 所示。

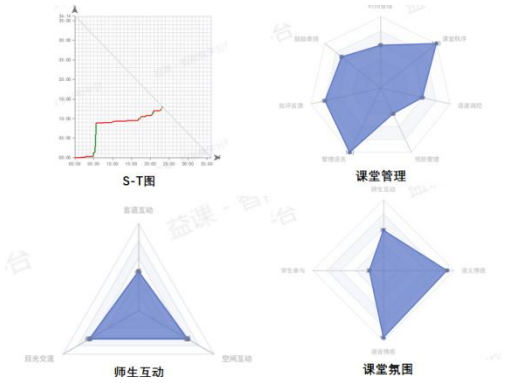


图 2 “益课”多模态分析报告

在第二轮原型设计过程中，进行了三次集体反思会议。在开始时，所有参与者都阅读了前一天教学的多模态分析报告，同时观看了教学视频。然后，他们轮流根据反思指南对教学进行评价。授课教师首先简要分享了他的教学流程，以及对多模态分析报告的看法。然后，其他人对这些报告进行解读，并一起讨论教学中可以改进的方面，以及从他/她的教学中获得了什么。最后，主持人总结了反思结果和建议。

4.2.3 评估

在第二轮中，职前科学教师的评价更加具体和深入，一些教师能够更加批判和客观地看待自己的教学。参与者 D 是一名大一新生，她分享道：“在第一轮集体反思中，我不知道在评估教学时应该关注哪些方面。我只根据我的第一印象来判断一堂课的质量。但在这一轮的集体反思中，我了解到了具体需要考虑的方面，比如师生互动、课堂管理、课堂氛围等。”对自己教学很满意的参与者 B 认识到了他对自己教学实践的误解，“之前，我认为我在师生互动方面做得很好，但是 S-T 分析让我发现我与学生的互动是不平衡的，我在前 20 分钟与学生的互动很密集，在后 20 分钟很少互动。我应该改善这一点。”由于会议主持人努力维持秩序，集体反思更有组织性。在第一轮集体反思中从未表达过自己观点的新教师逐渐活跃起来。在第二轮中，参与者之间的话语量逐渐趋于平衡。每个参与者总共至少有 15 分钟的分享时间。总的来说，所有人都更投入，更热情。

然而，一些参会者抱怨难以解读报告，特别是对于大一的职前科学教师来说。主持人也注意到了这个问题，她解释说：“他们（这些低年级学生）没有修过教育信息处理这门课。这是他们第一次看多模态分析报告，所以他们不理解……在讨论的过程中，他们会感到困惑，让高年级同学来解释，这将花费额外的时间，并造成混乱。”另一位大四学生 F 同学也表示：“如果我们继续解读报告，集体反思可能仍然停留在一个浅层次上，讨论什么是这个，什么是那个。我们仍然没有形成深刻而批判性的集体反思。”此外，一些人表示想邀请有经验的在职科学教师加入集体反思，提供更专业和客观的评价，帮助改进教学。

第二轮集体反思会议更有针对性，更具条理性。研究团队做出了以下修订：R1.集体反思开始时，简要解读多模态分析报告。R2.改进多模态分析报告的使用方式，促进批判性集体反思。R3.为教师的教学表现提供更多方面的证据和反馈。

4.3. 第三轮原型设计

4.3.1 设计

研究团队邀请专家参与集体反思，指导职前科学教师解读多模态分析报告[R1]。此外，采用了两种新的方法来使用多模态分析报告。首先，对同一教师的教学分析报告进行纵向比较；第二，对上同一堂课的不同老师的报告进行横向比较。通过比较的角度，其目的是鼓励

成员更批判性地看待分析报告的内容[R2]。邀请有经验的在职科学教师参与集体反思，提供更多教学表现的评价[R3]。

4.3.2 实施

在第三轮原型设计过程中，进行了三次集体反思会议。在第一次会议开始时，主持人概述了规则并分发了反思指南。两位受邀的教育技术学专家对分析报告作了简短的解释，帮助与会者能够解释分析报告的内容。随后，上课的职前科学老师，为参会者分发了他三个时期的多模态分析报告。然后，他简要地分享了他在三个时期的教学情况，并根据报告描述了他的教学经验的演变。然后，其他职前科学教师根据不同阶段，轮流对报告进行对比分析，共同讨论职前科学教师在教学实践中取得的进步和需要改进的方面。针对职前科学教师发现的问题，专家们基于相关教育理论提出了建议。最后，主持人总结了这次集体反思。

在第二次会议中邀请了两位具有五年以上教学经验的在职科学教师参加。在开始时，主持人概述了规则并分发了反思指南。随后，三位教过同一堂课的职前科学教师轮流描述了他们在同一堂课上的教学经历，以及他们认为有价值或具有挑战性的地方。他们轮流将自己的分析报告与另外两份分析报告进行比较，分享在教学实践中所感受到的异同。然后，其他职前教师对三位授课教师的分析报告进行对比分析，找出各自的优缺点，并提出改进建议。随后，在职教师对教学表现中的具体问题进行讨论，他们根据自己丰富的教学经验，提出了具体的建议。最后，主持人进行总结。

4.3.3 评估

本轮会议的内容和过程得到参与者的一致好评。职前科学教师表示之前提出的建议得到了解决，反思过程更详细，更有条理。特别是，在横向和纵向比较中，职前科学教师的观点明显变得更具批判性。一位教学经验较丰富的参与者表示："通过对比三位不同老师对上同一堂课的分析报告，我可以看出不同的人在教学中有自己的优点和缺点。我可以取长补短，互相学习。"教学经验有限的参与者 A 表示："通过对几周前和几周后的报告进行对比分析，我可以看到我的进步。我能找出自己取得进步和需要改进的地方。"值得注意的是，职前科学教师特别喜欢关注师生互动和课堂氛围。大一新生 D 表示："以前，我经常关注自己在课堂上的声音是否足够响亮，但现在我意识到，与学生互动，激发他们对科学的兴趣要重要得多。"

职前科学教师的参与度和积极性很高。他们分析了教学报告，讨论了教学绩效的各个方面，并提出了许多改进建议。此外，参会者也表示，他们对专家和在职教师参与到集体反思感到非常兴奋："专家们基于专业知识，为我们改进教学提供了有用的建议。"一位参会者也分享说："专家不仅可以指导我们解释报告的具体内容，还可以介绍与多模态分析报告相关的技术知识，从而促进我的专业发展。"

5.评估

本研究采用了 Yu et al.（1995）开发的教学自我效能感量表评估职前科学教师的教学自我效能感。同时研究小组选择了五名不同年级的参与者，对他们进行了半结构化访谈。

5.1. 教学实践和教学自我效能感

本研究分析了职前科学教师三个阶段的教学表现。教学的三个阶段对应于集体反思的三轮原型迭代过程，分为早期、中期和后期。本研究选取了课堂话语和课堂氛围两个角度考察三个阶段的教学，如图 3 所示。教师话语量（TTTP）逐渐减少，学生话语量（STTP）逐渐增加。此外，教师使用解释性陈述句（SP）的数量有所减少，而启发式疑问句（QP）的数

量有所增加。在课堂气氛方面，语音情绪分析显示，积极语音情绪（PVEP）所占比例逐渐上升，消极语音情绪（NVEP）所占比例相应下降。

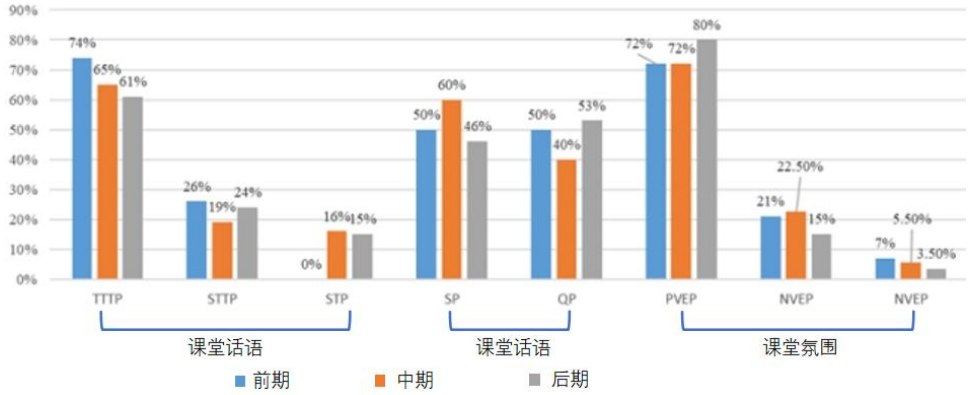


图 3 三个阶段的课堂话语和课堂氛围的变化情况

在集体反思的中后期开始前和结束后分别对职前科学教师的教学自我效能感进行了测试。配对样本 t 检验结果显示，职前科学教师的教学自我效能感得到了显著提升（t=-2.65, p<0.05）。

表 1 教学效能配对样本 t 检验结果

	Pre-Mean	Pre-SD	Post-Mean	Post-SD	t	p
教学自我效能感	29.31	10.12	37.33	5.27	-2.65	.027*

5.2. 职前科学教师对多模态分析报告整合集体反思的看法

在访谈中，五位参与者都表示，课堂教学视频和基于视频的课堂分析报告帮助他们回顾课堂过程，提供了很多的细节和不同视角，促进了反思的广度和深度。例如，几乎没有任何教学经验的大二学生 C 表示：“分析报告让我们意识到可以走在前排的学生中间，并注意到后排的学生，这些学生是我们以前从未注意过的。这种新技术的应用可以为我们提供更多的讨论话题和评价教学的视角。”参与者亦就集体反思的合作氛围分享意见。主持人表示：“集体反思可以帮助我们加强彼此之间的联系。我们喜欢互相倾听，互相交流，一起进步。”新加入的成员 A 表示：“我们是为了一个共同的目标而努力的团队伙伴，那就是给可爱的学生们上一节好的科学课。”这表明集体反思使团队成员们的关系更加紧密，团队的共同目标也更加明确。

总体而言，职前科学教师认为多模态分析报告整合的集体反思对反思的深度、专业发展和团队氛围都有促进作用。然而，一些人抱怨说，报告难以理解，并希望在报告中包含更多的数据，获得多元化的教学评价以此改善教学。

6.结论与展望

在本节中，讨论了研究的主要发现，并提出进行批判性和有组织的集体反思的设计原则。

6.1. 挑战

本研究发现，在集体反思中，经验不足的职前科学教师难以表达自己的观点。他们往往不知道从教学表现的哪些方面去思考，对教学表现缺乏整体的、结构性的认识。对于没有经验的职前科学教师来说，在集体反思中与其他成员建立情感联系也是一个挑战。

本研究发现职前科学教师在解读多模态分析报告时存在挑战。在第二轮集体反思中，他们花了大部分时间来理解报告的意义。他们被报告所困，无法将报告的内容与实际教学表现联系起来。此外，职前科学教师也希望多模态分析报告能够涵盖更多的评价维度，提供更全面的证据。职前科学教师在使用多模态分析报告的方法上也遇到了挑战。如果只解释报告中内容的意思，那么集体反思仍将是肤浅的讨论。因此，应该探索多种使用分析报告的方法，以比较和发展的视角批判性地看待报告中的内容。

将人工智能工具整合到教师教育中也存在挑战。研究结果表明，人工智能工具提高了教师集体反思的效率和有效性。人工智能工具可以为职前科学教师提供新的视角和见解，这可能会激发团队成员的创造性思维，并促进新解决方案的产生。然而，人工智能技术固有局限性，可能导致误判或幻觉。因此，当与人工智能工具交互时，人类的主导作用和批判性思维的重要性至关重要。

6.2. 设计原则

从本研究中总结出的设计原则如下：（1）为职前教师集体反思提供全面客观的证据。（2）明确职前教师表达观点的顺序，使集体反思更有组织性。（3）帮助职前教师理解多模态分析报告的内容。（4）改进多模态分析报告的使用方式，如对同一教师教学实践的分析报告进行纵向比较，或对不同教师讲授的同一节课的分析报告进行横向比较（5）邀请有经验的在职教师和专家参与讨论，为教师的教学提供更多方面的评价和建议。

7. 研究的局限性与未来展望

本研究结果存在一定的局限性。本研究的设计原则主要适用于大型研究型大学的科学教师教育项目，可能无法推广到其他情境或学科领域。此外，目前没有直接证据证明教学效果的提升是教师们集体反思的效果。职前教师们在校内参与的其他活动也可能提高专业素养，如微格教学等。未来的研究应更全面地探究具体的流程和技术让职前科学教师进行批判性反思，更好地促进他们的专业发展。此外，未来的研究还将在不同情境和学科中，探索多模态分析等智能技术促进职前教师的集体反思。

参考文献

俞国良,辛涛,申继亮.教师教学效能感：结构与影响因素的研究[J].心理学报,1995,(02):159-166.

郑永和,杨宣洋,王晶莹,等.我国小学科学教师队伍现状、影响与建议：基于 31 个省份的大规模调研[J].华东师范大学学报(教育科学版),2023,41(04):1-21.DOI:10.16382/j.cnki.1000-5560.2023.04.001.

Chen, M. (2023). Leveraging affordances in an ecological stance: Reflective language teaching for professional development during COVID-19. Heliyon, 9(5).

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. Teaching and teacher education, 16(5-6), 523-545.

Giroux, H. A. (2024). Teachers as intellectuals: Toward a critical pedagogy of learning. Bloomsbury Publishing

Hung, S. T. A. (2016). Enhancing feedback provision through multimodal video technology. Computers & Education, 98, 90-101.

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, 113.

Koc, Y., Peker, D., & Osmanoglu, A. (2009). Supporting teacher professional development through online video case study discussions: An assemblage of preservice and inservice teachers and the case teacher. Teaching and Teacher Education, 25(8), 1158-1168.

Mei, A., Long, T., & Liu, Z. (2024). How blue hat co-regulate pre-service teachers' collective reflection based upon the six-hat thinking technique: A lag sequential analysis. Thinking Skills and Creativity, 53, 101626.

Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. Journal of Second Language Writing, 16(2), 82-99.

Moon, J. A. (2013). Reflection in learning and professional development: Theory and practice. Routledge.

Nielsen, N., Schweingruber, H., & Wilson, S. (Eds.). (2016). Science teachers' learning: Enhancing opportunities, creating supportive contexts. National Academies Press.

Otienoh, R. O. (2011). Teachers’ lack of deeper analytical reflections: Who is to blame?. Reflective Practice, 12(6), 733-747.

Plomp, T. (2013). Educational design research: An introduction. Educational design research, 11-50.

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning?. British Journal of Educational Technology,51(5), 1450-1484.

Soni, A. (2023). Empowering collective reflection: realistic evaluation of video enhanced reflective practice (VERP) in an early childhood setting in England. Professional development in education, 49(5), 827-841.

Tripp, T., & Rich, P. (2012). Using video to analyze one's own teaching. British Journal of Educational Technology, 43(4), 678-704.

Van den Akker, J. (1999). Principles and methods of development research. Design methodology and development research in education and training/Kluwer Academic Publishers.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. Educational research review, 15, 17-40.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. Educational technology research and development, 53(4), 5-23.

六顶思考帽助力的协同反思对促进师范生教学设计学习效果的影响研究

Research on the impact of collaborative reflection assisted by six thinking hats on promoting the learning effect of teacher training students' teaching Design

王宁[1]，赵雅[2]，龙陶陶[3*]
[123]华中师范大学
[*] taotaolong@mails.ccnu.edu.cn

【摘要】 本研究探讨师范生在协同反思中使用六顶思考帽不同顺序对教学设计学习效果的影响。实验班采用固定顺序（蓝-白-红-绿-黄-黑-蓝）的六顶思考帽进行协同反思，对照班自由选择顺序。结果显示，实验班教学设计成绩显著优于对照班，但期末成绩无显著差异。这表明固定顺序的六顶思考帽在教学设计环节能显著提升协同反思效果，但单一顺序可能限制学生思考。研究认为，灵活运用六顶思考帽顺序有助于提升师范生的协同反思能力和教学设计学习效果，对未来教学讨论中协同反思的应用提供了借鉴。
【关键词】 协同反思；师范生；六顶思考帽

Abstract: *This study explores the impact of different sequences of the Six Thinking Hats on teaching design learning outcomes in collaborative reflection among teacher candidates. The experimental group used a fixed sequence (Blue-White-Red-Green-Yellow-Black-Blue), while the control group chose freely. Results showed that the experimental group's teaching design scores were significantly higher, but there was no significant difference in final exam scores. This indicates that a fixed sequence enhances collaborative reflection in teaching design, but a single sequence may restrict thinking. The study suggests that flexible use of the Six Thinking Hats can improve collaborative reflection and teaching design learning outcomes among teacher candidates, providing insights for future applications in teaching discussions.*
Keywords: collaborative reflection, teacher candidate, six thinking hats

1. 前言

随着信息时代对高效决策与系统化思维的需求日益增长，优化群体思维策略已成为组织管理与教育实践领域的重要课题（德博诺培训创新思维，2004）。英国的爱德华·德博诺提出的“六顶思考帽”的思维方法是运用横向思维的实用方法，将指导人们在思考过程中从不同的方向进行思考，打破了传统讨论中的无序、无节奏、无从下手的缺点，因此受到了大家的广泛关注（德博诺，2008）。研究表明该策略在企业管理、医疗决策等领域的应用可使会议效率提升 40%以上，但其在教育领域中应用仍待深入探究（Lee & Kang, 2014; Patre, 2016）。

培养师范生的反思能力是教师教育的一项重要任务（胡忠青，2014；栗瑶，欧阳玲和秦福莹，2022）。现有研究表明师范生在教学完成后更注重对教学思路、教学理念、教学过程以及教学效果等学环节的个人反思，而对集体的协同反思却存在倦怠的情况（饶志祥，2021）。调查显示，绝大部分的师范实习小组将集体的协同反思视为低效程序（饶志祥，

2021），这种认知偏差导致群体智慧难以有效整合。同时，新课程改革提出师范生应不断增强协同反思意识，采用合适的思维策略进行探讨交流，以促进教学完善，提高教学能力（教师厅函[2022]10 号）。基于此，本研究创新性地将六顶思考帽思维策略嵌入师范生教学设计的协同反思环节，重点探讨其不同顺序对师范生教学设计效果的影响，旨在提升师范生的协同反思能力和教学设计能力，同时为师范生的教学实践提供有效的指导建议。

2. 理论基础

2.1. 六顶思考帽

六顶思考帽是爱德华·德波诺提出的一种思考工具，它通过以不同颜色的“帽子”代表不同类型的思考，帮助人们更清晰地组织、管理和引导思考，达到更加全面、综合地思考问题的目的。六顶思考帽分别为白色帽、红色帽、黑色帽、黄色帽、绿色帽和蓝色帽。每一顶帽子代表一种不同类型的思考角度，通常用六顶有颜色的帽子来分别比喻人们具有的六项思维功能（De Bono, 2017）。白色帽代表信息的搜集和呈现，红色帽代表情感和直觉，黑色帽代表批判性思考，黄色帽代表积极和创造性的思考，绿色帽代表创新和发散性思维，蓝色帽代表控制和整合（任佳馨，2012）。

作为一种有效的思维工具，六顶思考帽可以帮助使用者从多个维度去思考问题，促进思维的开阔与深入（(De Bono, 2017; Goçmen & Cos¸kun, 2019）。研究表明有效地应用六顶思考帽可以帮助师范生更好地进行教学和反思，从而提升他们的思维能力和创造力（(Hu et al., 2021）。同时，六顶思考帽能够帮助师范生系统地分析学科问题，提高学习效率，培养团队协作能力（Chen et al.,2023）。

2.2. 师范生协同反思

师范生的协同反思过程是一群师范生合作进行思考和分享意见来促进彼此学习和成长的过程（(Darmawansah, Lin & Hwang, 2022; Kager et al., 2022）。在教育领域，师范生协同反思能够帮助师范生更全面地了解学生的需求和心理状态，培养敏锐的洞察力和分析能力，进而提高自身的教学质量和水平，因此被广泛应用于教师培训和教育改革中（栗瑶，欧阳玲和秦福莹，2022）。此外，在教学设计的过程中加入协同反思的环节，可以帮助师范生更好地解决教学问题，利用集体的力量设计出更优质的教学内容（袁丁，2023）。然而，研究表明目前师范生的教学反思仍以个人反思为主，很少开展集体协同反思，并且在集体讨论中经常会出现“失语”以及不愿意和不主动的情况，甚至存在消极怠工的问题（饶志祥，2021）。因此，研究如何使用思维工具帮助师范生有效地进行协同反思，以助力师范生教学设计的提高和专业能力的发展，具有重要意义。

综上所述，当前研究已逐渐认识到，在师范生中开展有效的协同反思面临诸多挑战，而六顶思考帽具有促进其协同反思的潜力。因此，本研究在师范生教学设计的协同反思过程中融入六顶思考帽思维策略,旨在研究六顶思考帽助力的协同反思对促进师范生教学设计学习效果的影响

3. 研究设计

3.1. 研究背景

本研究以为期一学期，面向华中地区某师范大学三年级本科生的《现代教育技术》课程为背景，该课程借助现代信息技术，将教育学、心理学和传播学等基本理论结合运用于教学实践，同时整合教学设计、在线协作反思等多元教学策略充分调动学生的兴趣和积极性，让

学生跟随课程活动不断地进行反思，了解现代信息技术在教学中的应用，培养信息素养和信息化教学能力。本研究设置实验组与对照组采用不同"六顶思考帽"的不同序列进行反思，探究六顶思考帽助力的协同反思模式对教学设计学习成效的影响。研究者通过全程非介入式观察，采集小组在线讨论数据，旨在揭示学生反思路径差异与教学设计能力发展之间的关联机制。

3.2. 研究对象

本研究以某师范大学《现代教育技术应用》课程中随机分班的两个班的大三师范生（N=151 人）为研究对象，该群体具备跨专业学科素养与信息化教学基础，但在协同反思中普遍存在思维碎片化、经验整合困难及反思层次浅表化等问题。据现有研究发现六顶思考帽助力的协同反思可以有效的帮助学生减少反思浅层次问题，促进深层次反思。

3.3. 研究过程

本研究采用准实验研究设计，以某师范大学《现代教育技术应用》课程的七班（对照班，76 人）和八班（实验班，75 人）为研究对象。两班学生由教务处随机分配，教学环境、师资力量、课程内容及进度均保持一致，以减少无关变量干扰。自变量为六顶思考帽在协同反思中的使用顺序：实验班采用固定顺序（蓝帽➡白帽➡红帽➡绿帽➡黄帽➡黑帽➡蓝帽），对照班自由选择顺序；因变量为教学设计作品成绩与期末理论考试成绩，由课程教师依据标准化评分体系进行盲评。实验持续一学期，课程采用翻转课堂模式，重点围绕教学设计能力培养展开，如表 1 所示。课程共分为九大专题，前八专题聚焦现代教育技术理论基础，第九专题要求学生以小组为单位完成微课设计与展示。每次专题结束后，两班学生需通过线上平台开展协同反思。实验班遵循固定顺序的六顶思考帽流程，具体步骤如下：

蓝帽（整合）：明确反思目标与讨论框架；

白帽（客观事实）：梳理教学设计中的具体问题与数据；

红帽（情感表达）：小组成员分享主观感受与直觉反馈；

绿帽（创新）：提出改进方案或创意性策略；

黄帽（积极评价）：评估方案的可行性及潜在优势；

黑帽（批判分析）：识别方案缺陷及实施风险；

蓝帽（总结）：整合讨论成果并形成优化建议。

对照班则允许小组自主决定六顶思考帽的使用顺序。研究者全程匿名观察线上讨论记录，避免干预实验过程。

表 1 研究设计

不同顺序六顶思考帽助力的协同反思对促进师范生教学设计学习效果的影响

	实验班	对照班
实验对象	《现代教育技术应用》八班学生 75 人	《现代教育技术应用》七班学生 76 人
实验时间	一学期	
自变量	按照安排六顶思考帽的顺序使用	不安排六顶思考帽顺序，随机使用
因变量	微课设计成绩和期末成绩	

选取内容　　《现代教育技术应用》课程

3.3. 数据收集与分析

本研究对学生的教学设计评分标准改编自何克抗教授的《教学系统设计》一书中提供的教学设计的评分标准和评价量表（何克抗, 2006），并增加对反思环节的评分，如表 2 所示。该标准共分为了十一个部分，包括题目选择、学习者特征分析、课程目标与重难点、教学方法选择、活动设计、教学资源和工具选择与设计、教学评价设计、反思环节、新颖性、可操作性、系统性。每个部分均按照重要性和工作量大小赋予分值（王伟, 2018）。

课程期末考核为闭卷的方式进行课程理论考察，由选择、简答、论述等题目构成。考试内容包括教学设计的定义、教学设计的流程、教学目标的撰写等，均考察学生对教学设计相关知识的理解情况。教师按照参考答案对学生回答情况进行评分，并计算总成绩。该试卷为该校十余位教授《现代教育技术应用》一课的教师联合编制，并使用了五年，具有一定的有效性和可信度。课程结束后收集七班和八班学生教学设计成绩以及期末成绩数据，通过 SPSS 统计分析软件对两个班的教学设计成绩和期末成绩分别进行独立样本 t 检验。

表 2 教学设计的评分标准和评价量表

项目	分值	评价标准
题目选择	5 分	题目选取符合课程标准，与《现代技术教育应用》深层次整合，可激发学习者学习兴趣
学习者特征分析	5 分	根据学习者年龄阶段，详细分析其认知特点、学识基础、情感态度等，以便更好教学
课程目标与重难点	10 分	明确阐述课程的教学目标与重难点，符合课程题目与学习者特征
教学方法选择	15 分	选择合适的方法与策略帮助教师发挥主导作用与学生的主体地位，并有助于课程有效实施
活动设计	20 分	设计出符合课程目标的教学活动且能够吸引学生的学习兴趣，新颖且有趣
教学资源和工具选择与设计	10 分	选取和设计的教学资源与工具能有效完成教学活动并促进教学
教学评价设计	10 分	教学评价标准体系明确，有详细的评价工具
反思环节	10 分	对教学设计进行组内反思，进而改进教学设计
新颖性	5 分	有创新的教学方法与创新活动，有助于激发学习者学习兴趣
可操作性	5 分	课程活动简单方便实施。对教师和教学技术等要求不高，方便其他教师进行教学
系统性	5 分	课程系统且规范
总分与建议	-	-

4. 数据分析与讨论

实验班的教学设计成绩水平优于对照班教学设计成绩（$t=2.997, p=0.000<0.05$，差异显著），如表 3 所示，说明使用了安排“六顶思考帽”的顺序进行协同反思能够提高教学设计成绩，促进教学设计学习效果。对照班的期末成绩水平与实验班期末成绩差异不显著（$t=-2.207, p=0.759>0.05$），说明使用安排“六顶思考帽”的顺序进行协同反思并不能提高学生的期末成绩，两个班级并未有大差异，也就是对理论知识的掌握情况并无促进作用。

表 3 教学设计成绩和期末成绩数据分析结果

	班级	个案数	均值	标准差	t	显著性
教学设计成绩	对照班	75	83.487	13.743	7.997	0.000***
	实验班	76	88.280	2.346		
期末成绩	对照班	75	70.490	10.277	-2.207	0.759
	实验班	76	66.780	10.272		

本研究发现六顶思考帽的顺序使用对师范生协同反思有显著性提升师范生教学设计作品成绩的作用，而对期末成绩即课程理论知识掌握情况，无显著影响。因此，在使用六顶思考帽过程中，并不是每一个环节都能提高学生的协同反思效果。在本研究中，虽然实验班的教学设计成绩优于对照班，但是期末成绩与七班差异不显著，其原因可能是顺序的安排，思考顺序让实验班同学在课程设计协同反思过程中的思路更加清晰，成员间冲突减少，利于课程设计的优化与改进。虽然针对课程设计这种有明确目的的活动时实验班比对照班能展现明显的思维优势，但是在面对期末试卷考查理论知识时，实验班同学与对照班同学并无太大差异。

5. 结论与展望

5.1. 六顶思考帽能提高师范生对教学设计的学习效果

研究表明六顶思考帽助力的协同反思对师范生在教学设计学习中有一定的促进效果，但没有提高师范生对课程的理论知识掌握情况。同时六顶思考帽的不同顺序进行协同反思会影响师范生的学习效果，安排顺序使用六顶思考帽与随机使用在教学的不同环节学生的成绩不同。不过，无论采用何种顺序，六顶思考帽都可以激发师范生的主动参与，增强协同反思的效果，提高师范生对教学设计的学习效果。

在协同反思过程中，六顶思考帽的使用能够帮助师范生们在教学设计中取得更优的效果，在讨论中共同研究问题、探究问题和解决问题。在使用六顶思考帽促进师范生协同反思时，应注意以下方面：首先采用一定的顺序使用六顶思考帽分析问题，所有人每一次戴同样的帽子，一次只戴一顶帽子，从多角度看待问题，避免单一思维模式；另外，师范生们需要明确不同颜色的帽子间的逻辑关系和使用顺序，和团队成员达成共识，以保证思维活动的有序进行。例如，在分析问题之前，先戴白帽，客观的发现问题的关键点；接着戴红帽，以非理性的方式讲述个人体验，使得团队成员共情并理解问题的深层次含义；当团队成员达成共识并找出问题的各种原因后，可以戴黑帽以批判性思维角度发现问题的不足之处，进而再次思考解决问题方案以优化讨论结果。这种有序的思维过程可以节约讨论时间，避免重复和冲突，从而高效的解决问题。

5.2. 六顶思考帽帮助师范生全面分析问题

本研究发现六顶思考帽能够帮助师范生更加全面地分析问题，从不同角度进行思考，减少了思考上的局限性（De Bono, 2017）。六顶思考帽的顺序使用有助于师范生逐步深入反思，提高了反思的深度和广度（Mei et al., 2024）。六顶思考帽这类平行思维工具，已经在教育领域中得到广泛应用，其为师范生在协同反思的过程中寻找到一个直接的方法，以帮助其批判和广泛地思考新的和熟悉问题，六顶不同角度的帽子通过一种结构化、开放式的“思维转换”方法促进深刻和持续的思考，使得问题得到更加快速高效的解决（Goçmen & Cos¸kun,

2019）。因此在课程设计方面，师范生会有更多的创新点，深度挖掘一个方向，反复推敲，有助于教学设计的实践。

5.3. 六顶思考帽能促进团队协作

研究发现六顶思考帽有利于团队精神和协作能力的培养，教学设计正需要团队合作，进而逐步完善课程（Sibo et al., 2023）。当团队成员进行协同反思时，成员们将个人的思维融入到团队的思维中，理解和尊重各种观点，并在他人的观点基础上再次生成新的观点。这样，团队成员之间的沟通增加，以同角度思考问题也会使得团队成员间对彼此的信任增加，集体的力量同样会使得克服挫折的能力和创新精神更加强大（Chien, 2021）。因此，六顶思考帽的使用不仅可以帮助思维更加明朗清晰，也可以促进团队的协作能力的提高，增强协同反思的效率和质量（Hu et al., 2021）。

6. 总结与展望

本研究发现采用六顶思考帽作为师范生协同反思的工具，可以提高师范生的反思能力、支持协同学习和促进思路的转化，提高教学设计的能力。然而，研究也存在一些局限性，在期末这种整体的考核中，学生能力并未得到提高。因此，在未来的研究中对于六顶思考帽的使用顺序、组合方式和影响因素以及在哪些环节中加强对六顶思考帽的使用等还需要进一步探讨。

同时，在未来的研究中，可以进一步探索使用不同六顶思考帽顺序和组合方式对师范生协同反思的影响，分析不同思考帽的特点和优劣势。此外，还需研究师范生使用六顶思考帽协同反思在其他的教学内容学习中促进作用，以及如何最大化六顶思考帽的价值使其成为促进协同学习的有效工具。另外，可以尝试将六顶思考帽的思维模式与其他教育教学策略相结合，开发具体的师范生培训课程和技能。

本研究的可行性得到了实证研究的支持，因此进一步研究六顶思考帽与师范生协同反思的影响机制，将有助于提高师范生的协同反思能力，并使协同反思在学习中更加有效，为教育教学实践提供深入研究的指导意义。

参考文献

王伟（2018）。MOOC同伴互评的互评效果实证研究。深圳大学，2018。

任佳馨（2012）。ASP教学中的六顶思考帽研究与分析。中国科教创新导刊, 2012, 645(25):92。

何克抗，林君芬，张文兰（2006）。教学系统设计。北京：高等教育出版社，2006（5）。

胡忠青（2014）。反思性教学方法研究。教育教学论坛，2014, 135(02):85-86。

饶志祥（2021）。师范生专业实践反思现状及改善策略研究。江西师范大学, 2021。

袁丁（2023）。新高考背景下探索高中数学核心素养培养策略。数学学习与研究，2023(05), 101-103。

栗瑶，欧阳玲，秦福莹（2022）。反思型教师的研究进展。赤峰学院学报(自然科学版), 2022, 38(11)，85-89。

爱德华·德·博诺（2008）。“六顶思考帽”。太原：山西人民出版社，2008。

教育部办公厅关于加强小学科学教师培养的通知(教师厅函[2022]10号)[A/OL].[2023-05-05].http://www.gov.cn/zhengce/zhengceku/2022-05/28/content_5692779.htm.。

德博诺培训创新思维（2004）。中华工商时报，2004-11-10。

Chen, D., Liu, F., Hu, Y., Du, X., Bai, Y., Ren, Z., … Yu, W. (2023). Service design from the perspective of “six thinking hats”: A comparison of storytelling strategies of experts and novices. *Thinking Skills and Creativity*, 47, Article 101219.

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*, 49(5), 606–617, 3-13.

De Bono, E. (2017). In Six Thinking Hats: The multi-million bestselling guide to running better meetings and making faster decisions. Penguin: UK.

Goçmen, O., & Cos¸kun, H. (2019). The effects of the six thinking hats and speed on creativity in brainstorming. *Thinking Skills and Creativity*, 31, 284–295.

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, Article 104516.

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, Article 100946.

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*, 113, Article 103683.

Lee Yong Seob & Kang Jung Mun.(2014).The Effects of Science Lessons using Six Thinking Hats techniques on Creativity and Science Academic Achievement. *Journal of the Korean society of earth science education*.

Mei, A., Long, T., & Liu, Z. (2024). How blue hat co-regulate pre-service teachers' collective reflection based upon the six-hat thinking technique: A lag sequential analysis. *Thinking Skills and Creativity*, 53, 101626.

Patre, S. (2016). Six thinking hats approach to HR analytics. *South Asian Journal of Human Resources Management*, 3(2), 191–199.

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, Article 101217.

大学人工智能通识课程体系构建:现状与实施建议

The Construction of Artificial Intelligence General Curriculum System of Universities: Current Situation and Implementation Suggestions

王鑫茹[1], 魏艳涛[1,*], 徐琦[1], 裴思睿[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
[*]yantaowei@mail.ccnu.edu.cn

【摘要】 当前，人工智能通识教育是应对未来智能社会发展的重要举措，而构建高质量的课程体系是其关键环节。因此，本研究基于泰勒目标模式从课程开发的视角对现有人工智能通识课程体系的课程目标、课程内容、课程实施以及课程评价进行分析。结果表明，现阶段人工智能通识课程体系具有课程目标分层化、课程内容模块化、课程实施梯度化以及课程评价场景化的特点。根据对课程建设的实践梳理，本研究提出了人工智能通识课程体系建设的实施建议，包括根据学校实际打造特色化课程目标、紧跟科技发展构建动态课程知识库、根据发展需求探索个性化学习路径和依据课程评价迭代优化课程体系，助力人工智能通识教育的高质量发展。
【关键词】 人工智能通识教育; 课程建设; 泰勒目标模式

Abstract: *At present, general education in artificial intelligence is an important initiative to cope with the development of future intelligent society, and the construction of a high-quality curriculum system is its key link. Therefore, this study analyzes the course objectives, course contents, course implementation, and course evaluation of the existing general curriculum system of artificial intelligence from the perspective of curriculum development based on Taylor's goal model. The results show that the current stage of AI general education is characterized by stratified course objectives, modularized course content, gradient course implementation, and scenario-based course evaluation. Based on the practical combing of curriculum construction, this study puts forward the implementation suggestions for the construction of AI general education curriculum system, including creating distinctive curriculum objectives based on the school's actuality, constructing a dynamic curriculum knowledge base following the development of science and technology, exploring personalized learning paths based on developmental needs, and iteratively optimizing the curriculum system based on the evaluation of the curriculum, to help the high-quality development of AI general education.*
Keywords: artificial intelligence general education; curriculum development; Taylor's goal model

1. 引言

《教育强国建设规划纲要（2024—2035 年）》强调要完善人才培养与经济社会发展需要适配机制，提高人才培养质效。习近平总书记在向 2024 世界智能产业博览会致贺信中指出，人工智能是新一轮科技革命和产业变革的重要驱动力量，将对全球经济社会发展和人类文明进步产生深远影响。因此，在数智时代，应加强人工智能通识教育，将人工智能相关课程作为所有学科的公共基础课程，并加强跨学科交叉学习。近年来，人工智能通识教育得到

广泛关注，如 2024 世界数字教育大会、UNESCO(2024)等，强调 AI 通识教育是应对未来社会变革的核心路径，以期培养兼具技术能力、伦理意识与创新思维的公民。

然而，如何打造人工智能通识课程体系，赋能各学科各类人才培养仍值得考究。针对上述问题，本文聚焦高等教育阶段人工智能通识课程建设，基于泰勒目标模式(Tyler, 1949)从课程开发的视角分析现有人工智能通识课程体系的课程目标、课程内容、课程实施以及课程评价。根据分析结果，提出应对策略，为人工智能教育实践提供理论依据。

2. 基于泰勒目标模式的人工智能通识课程体系分析框架

根据高等教育中现有人工智能通识课程体系建设案例，如表 1 所示，基于泰勒目标模式对课程目标、课程内容、课程实施、课程评价四个方面进行分析整合，总结现阶段人工智能通识课程体系的特征和挑战，进而提出构建人工智能通识课程体系的发展策略。

表 1 人工智能通识课程体系建设典型案例

序号	案例名称	特色	发布日期
1	武汉大学数智教育白皮书	1+16+N：公共基础、公共核心、选修课程	2023.11.16
2	南京大学“人工智能通识核心课程体系”总体方案	1+X+Y：通识、素养、拓展	2024.2.28
3	南京航空航天大学人工智能+系列课程	通识、交叉、高阶	2025.3.14
4	复旦大学人工智能课程体系建设和教育模式改革“AI 大课”建设	AI-BEST：通识基础、专业核心、学科进阶、垂域应用	2024.6.5
5	浙江大学大学生人工智能素养红皮书（2024 版）	通识教育、AI+X 辅修、X+AI 微专业	2024.6.18
6	首都师范大学人工智能通识课设置实施方案	人工智能+教育	2024.7.15
7	Carnegie Mellon University BSAI Curriculum	人工智能辅修	2024.11.22
8	The University of Melbourne Centre for Artificial Intelligence and Digital Ethics	融合土著文化	2025.2.4
9	University of Cambridge Shaping AI for everyone	跨学科应用	2025.4.7

图 1 为基于泰勒目标模式的人工智能通识课程体系分析要素，根据现有人工智能通识课程体系构建案例的特色，分析课程目标和课程实施。通过对课程具体知识内容以及课程评价方案的梳理，总结现阶段人工智能通识教育的现状和挑战。根据分析结果，提出对应策略，为人工智能通识课程体系建设提供参考。

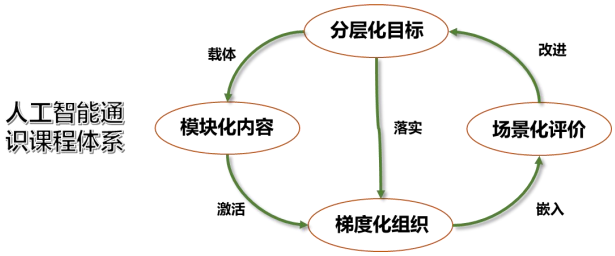


图 1 基于泰勒目标模式的人工智能通识课程体系分析要素

3. 分析结果

3.1 课程目标分层化

现阶段人工智能通识课程的课程目标表现出分层化的特点。例如，浙江大学对人工智能素养的培养目标分为体系化知识、构建式能力、创造性价值和人本型伦理四个方面，其中知识为基、能力为重、价值为先、伦理为本，四者间相辅相成、相互融合。南京大学的知识、能力、价值观与伦理三维课程目标，武汉大学将数智人才培养分为“通识、赋能、应用、专业”四个类型。无论如何对课程目标进行表述，能够达成一致的是人工智能通识课程的课程目标能够依据布鲁姆认知教育目标分类进行“通识—应用—创造”分层。

课程目标来源于对社会需求驱动对学生群体以及学科特征的综合分析，并从教育哲学和学习心理两个范畴对一般性目标进行精细化，形成更符合人才培养的课程目标，如图 2 所示。社会发展需要能够推动社会转型升级的复合型人才(武迪 & 戴琼海, 2021)，而人工智能作为新质生产力的核心驱动力，更是复合型人才培养过程中需要重点关注的领域。智能时代，人工智能逐步取代人类完成简单任务（即，知识记忆、知识检索等）(王竹立, 2017)，而人类需要掌握能力的是如何协同人工智能进行创造性应用。通过对人工智能素养框架(余维杰,郑梦婷&张颖, 2025)的深入分析，人工智能的学科内在逻辑遵循“知识认知—编程技能—学科应用—研究创新—社会伦理”。因此，人工智能通识课程的应该紧紧围绕人工智能通识教育的内在五大核心维度，以 AI 复合型人才培养为导向制定适应学生发展需求的课程目标。然而，学生在专业和能力方面都具有个体差异性，一般性的课程目标的适配程度降低。在已有的人工智能通识课程建设实践中，通常做法是利用学生的心智特征、先前经验在学习心理学范畴对课程目标进行过滤，形成基于 Bloom 教育目标分类(Krathwohl, 2002)的精细化分层课程目标，以适配多样化的发展需求。

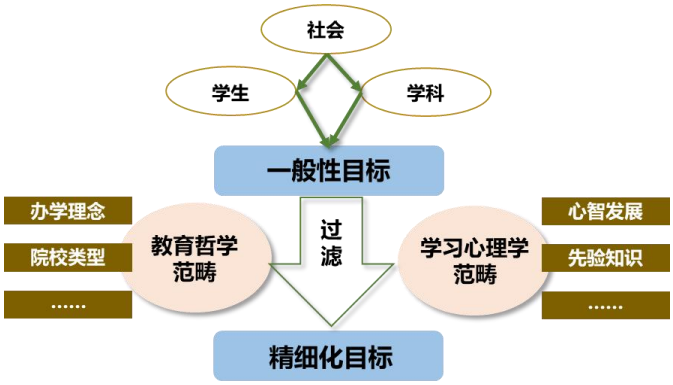


图 2 课程目标过滤过程

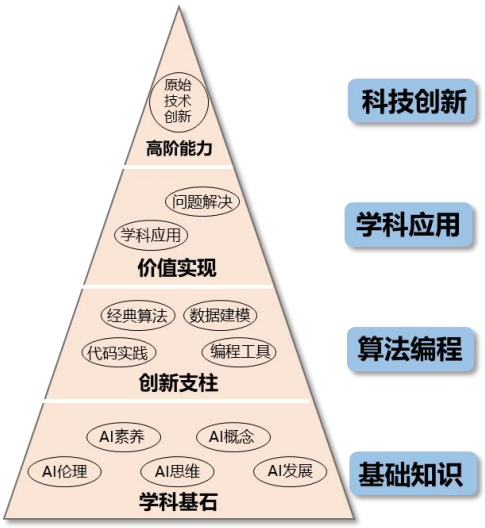


图 3 人工智能通识课程体系的课程内容

3.2 课程内容模块化

现阶段人工智能通识课程的课程内容具有模块化的特点。根据现有人工智能通识课程的课程内容编排梳理，大致可划分为基础知识、算法编程、学科应用和科技创新四个模块，如图 3 所示。例如，复旦大学将教学内容分为通识基础、专业知识、学科融合以及垂域应用。基础知识涵盖人工智能核心概念、人工智能思维、AI 发展脉络及伦理治理的系统认知。算法编程则是通过经典算法实现、编程工具（如 Python、PyTorch）与项目实践，提升数据建

模与工程化能力。学科应用聚焦“AI+学科”的跨场景融合，探索技术驱动实际问题的应用解决方案。结合人工智能前沿技术与科研竞赛去实现科技创新，以激发自主创新与伦理安全并重的技术突破能力。

课程内容作为课程目标的具体载体，其设计的逻辑起点和价值导向最终回归于课程目标。人工智能通识课程目标的分层化及其学科内在逻辑决定了人工智能通识教育课程内容的模块化，直接指导了学习经验的选择。人工智能通识课程应挖掘和总结可迁移的价值观、知识和技能及其组织结构方式，支持学生奠定基础性人工智能概念结构和涵盖价值观、知识和技能的广义知识图式，帮助学生同化终生学习进程中的新知识，或针对新问题解决情境和新技术调适升级知识图式(苗逢春, 2024)。通过对基础知识（包括人工智能思维和智能素养）的学习，建立学生对人工智能领域基础性概念的知识图式，形成人工智能思辨模式，奠定人工智能学习的基石。算法和编程是参与人工智能的技术基础(UNESCO, 2022)，是人工智能创新和发展的必备技能，也是人工智能通识课程的教学重点和难点。另外，人工智能通识课程应该培养学生将相关技术应用于学科领域的能力，一方面是将学生置于具体的学科情境中以激发先前知识经验，进而促进学生对人工智能知识体系的有意义生成(苗逢春, 2024)；另一方面是探索传统学科研究与人工智能有机融合的实践路径，打破学科壁垒，促进科学研究范式向以智能驱动为主的第五范式（即“智能化科研”）跃迁(李国杰, 2024)。如今，人工智能领域的顶尖人才、最原始技术创新人才依然稀缺(张立迁, 2018)，然而，如何培养人工智能核心专业人才仍是亟需探索的问题。人工智能通识教育将人工智能技术的科技创新作为一大教学模块，筛选出人工智能发展的坚实后备力量。综上所述，人工智能通识课程的课程内容以基础知识是基石、算法编程作为科技创新支柱、学科应用是价值实现路径、科技创新则是最高阶的能力拓展，四者之间存在多层面的相对独立又有机统一的螺旋进阶的逻辑关系。

3.3 课程实施梯度化

现阶段人工智能通识课程实施具有梯度差异化的特点。从课程结构上看，为适应不同学科、不同学段的差异化需求，当前课程体系普遍采用“通识必修—应用选修—能力进阶”的梯度化课程设置。例如，南京航空航天大学的通识、交叉、高阶系列课程以及复旦大学的“AI-BEST”课程结构设计。从教学模式上看，校企合作创建产教融合教学环境以及基于跨学科前沿问题进行科教融合是人工智能通识课程的创新亮点。从课程保障上看，注重跨学科背景的教师队伍建设和数字化课程平台的开发。人工智能通识课程体系的课程结构、教学模式和课程保障的描述如图 4 所示。

课程实施是落实课程目标的重要阶段，也是对课程内容的激活。针对不同学科不同需求的学生群体，人工智能通识课程对其的培养目标也具有差异性。通过对人工智能通识课程体系的课程结构分析，其解决方式是分梯度开设课程，从而实现分层课程目标的实施。根据学科差异和个体差异等因素，制定适合特定学生群体的选课方案，达到各专业人工智能复合型人才的定向培养。由于人工智能学科具有较强的综合性和应用性,其最终的发展目标是与产业和学科结合,走向市场应用(王雪等, 2020)。因此，针对学生不同的发展目标（例如，就业实践、学术科研等），人工智能通识课程采取产教融合、科教融汇等教学模式，实现从理论到实践的知识迁移。这一举措也为学生搭建了“真问题、真需求、真场景”的人工智能应用环境(戴瑞婷&李乐民, 2024)，让学生在“发现问题—分析问题—提出假设—验证假设”的一系列问题解决认知活动过程中发展高阶能力(Gick, 1986)。师资队伍建设和课程资源开发也是在人工智能通识课程实施过程中备受重视的部分，为课程实施提供保障。在师资队伍建设方面，一方面强调不同学科背景的教师资源整合，以适应人工智能通识课程的跨学科融合特性，

也为课程的丰富性和多样性提供了机遇；另一方面考虑到部分学科教师在人工智能教学领域实践经验较少，因此提供有针对性的教师培训和必要的教学支持。在课程资源建设上，构建数字化课程平台，通过学科逻辑有机组织教学资源形成课程知识图谱，满足学生自主学习的需求。

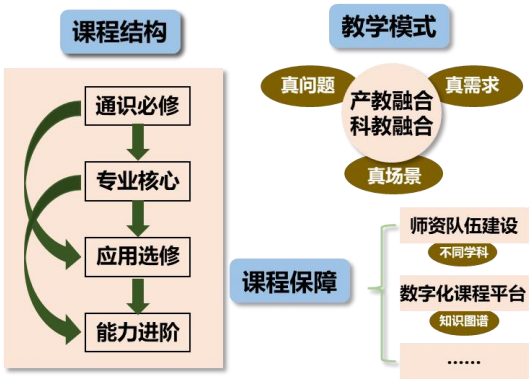


图 4 人工智能通识课程体系的课程实施

图 5 人工智能通识课程体系的场景化评价

3.4 课程评价场景化

现阶段人工智能通识教育课程评价具有场景化的特点。鉴于人工智能通识课程强调立足于学生学科背景，将学生置于真实问题解决情境下发展能力，这与基于场景的能力评测具有异曲同工之妙，如图 5 所示。基于场景的测评是通过构建模拟的真实情境任务，触发学习者将其内隐能力或素养在应对任务情境中真实表现出来，从而可利用相关技术工具对其外显行为数据进行动态捕捉，达到能力评测的目的(顾小清, 2024)。

课程评价是确定教育目标实际实现程度的过程，是依据特定标准，检查课程目标、设计与实施是否达成教育目的，并据此优化课程决策的过程(Tyler, 1949)。人工智能通识课程的教学环境、教学内容、教学活动等都具有高度数字化的特点，这一特点为获取师生课堂互动、学生学习路径等全教学阶段数据提供便利，为课程评价提供事实依据。根据人工智能通识课程目标构建课程数智评价理论框架，并应用数智技术打造多维度学生画像，推动教育评价范式从数据评价向数智评价转变(吴丹, 2024)。从评价方式上看，相比于单向的、传统的、静态的评价方式，现有人工智能通识课程关注情境化、过程性的场景化评价。基于场景的评价具有贴近真实场域、增强评价效果、实现多维评价的特点(顾小清, 2024)，能够系统考察学生在真实情境中运用人工智能相关知识进行实践探究和问题解决的能力，从而实现对学生的全方位评价。因此，在人工智能通识课程建设中，需要整合“学生—证据—任务”三大核心要素设计既能测试外部表现又可透视内在能力的表现性评价场景(Shute, 2011; 郭炯&邹佳人, 2025; 苗逢春, 2024)。就人工智能通识课程而言，可以从特定学生群体的培养目标出发，利用特定学科真实情境设置任务，考察学生在任务探索过程中的表现并形成过程数据日志，进而建立评价分析模型。

4. 人工智能通识课程体系建设的实施建议

4.1 结合学校实际打造特色化课程目标

当前，人工智能通识课程目标的制定依据来源于以社会需求对学生发展、学科逻辑的驱动的分析，随后根据专业大类对课程目标进行分层。然而，这种以专业为依据的分层课程目标在实践层面逐渐显现出局限性：一方面，各类高等教育院校的办学理念和人才培养定位存在显著差异(刘智英, 2011)；另一方面，学生群体的认知基础、职业愿景与学习需求亦呈现多

元分层(满都拉&卢晓东, 2018)。因此，需从教育哲学与学习心理学的双重维度对课程目标进行“过滤”，结合学校特色实现目标的动态调适与精准定位。

从教育哲学的视角看，课程目标是院校教育理念的具象化表达，需植根于学校的文化传统与办学定位。以师范类院校为例，其人工智能通识课程目标不仅需涵盖技术应用能力，更应强化“人工智能+教育”的融合创新能力。开设“人工智能赋能学科教育”等跨学科课程，培养学生设计智能教学工具、开发 AI 辅助评价系统的能力，为成为智能时代复合型教师做准备。院校在编制人工智能通识课程培养目标时需通过学科交叉、资源整合与价值重构，将课程目标与自身使命深度绑定，而非简单搬用通用性框架。从学习心理学的范畴出发，即使是同一学科背景的学生，其人工智能素养的培养目标亦可能存在差异。以师范生与非师范生为例：师范生的学习目标需侧重“技术赋能教育”的实践转化能力，如利用大模型优化教学设计、开发智能评测工具等；而非师范生则需聚焦于“技术赋能行业”的应用开发能力，以适应人工智能为驱动的产业变革。总之，人工智能课程目标的设定本质上是院校教育哲学与学习规律的耦合过程。唯有将社会需求的共性与院校特色的个性相融合，才能实现从“标准化培养”向“特色化赋能”的转型。未来的课程体系建设需进一步强化“院校—学科—学生”三者的互动关系：院校需挖掘自身文化基因，学科需探索交叉融合路径，学生需获得个性化成长支持。

4.2 紧跟技术发展构建动态课程知识库

如今，人工智能技术的发展日新月异，这也意味着人工智能知识从静态线性转向动态网络(刘鹂等, 2025)。以生成式人工智能（GAI）为例，2022 年 ChatGPT 的问世标志着技术从“数据分类”向“内容创造”的范式突破，而 2024 年 OpenAI 发布的推理模型 o3 版本已在数学、物理等复杂问题解决上超越人类专家水平。这种指数级的技术迭代速度，为人工智能通识课程知识体系的时效性带来冲击。若课程知识库仍以静态结构固化“卷积神经网络”“决策树”等传统知识点，将导致人才培养滞后于技术发展需求。而技术发展也可能带来培养目标的动态变化，例如，传统 AI 素养的“算法理解-工具应用-伦理判断”框架，正在被生成式 AI 时代的“提示工程-多模态创作-人机协同”能力维度替代。自生成式人工智能大量涌现以来，社会对公民的要求发生了认知层级的跃迁、实践能力的重构以及伦理知识的深化(UNESCO, 2023)。因此，构建动态课程知识库既是应对技术变革的必然选择，也是适应课程目标动态变化的核心路径。

人工智能通识课程动态知识库的构建需突破内容更新周期的时间限制，形成从技术更新追踪到知识逻辑解构再到教学组织转化的动态循环。首先，聚合学术论文、开源社区、政策文件等多方数据源，利用 AI 技术对人工智能最新知识进行抽取。其次，根据具体内容对知识在目标层面进行分层、在内容层面进行模块划分，对新知识进行有机整合。最后，教师需要从“知识传授者”转型为“技术解读者”，在确保 AI 知识的可信性之后完成新知识的教学组织转化。动态知识库的建构过程本质上是教育目标与技术逻辑的持续对话过程，是人工智能通识教育动态发展的过程，对培养适应技术变革的人才具有重要意义。另外，随着人工智能通识教育的推进，学生的总体智能素养也会随之提升，这时候课程知识也会随之改变。例如，当学生在课程开始前已经形成了基本的人工智能思维，在这种先验知识水平较高的情况下，人工智能通识课程需要重新考量学生真正需要的是什么学习经验，从而体现人工智能通识教育的真正功能。

4.3 根据发展需求探索个性化学习路径

在人工智能与各学科交叉融合的背景下，不同学科的知识结构、能力培养目标与学生个体认知特征之间存在显著差异。即便是同一学科，学生背景与职业规划也催生差异化需求。这就需要课程建设提供灵活的课程结构以支持学生的差异化需求，形成个性化的学习路径。然而，传统教学遵循“基础理论—算法编程—学科应用—科技创新”的线性逻辑，其设计假设基于高阶知识的获取依赖于更低阶能力的习得。但对于不同的学科，学生对人工智能技术的需求是不同的，例如，赋能型专业则需要人工智能基础知识以及学科应用实践，而并不关注算法编程之类的专业技术知识。因此，人工智能通识课程的学习路径是多样的，是各种层级能力的组合。相较于专业型学生的“基础理论—算法编程—科技创新”的发展路径，赋能型学生则可以选择“基础知识—学科应用”的学习路径，跳过算法编程的学习，直接在学科情境中应用人工智能获得应用实践能力。

个性化学习路径的生成和探索可以依托知识图谱技术，通过对知识关联的可视化梳理层级关系，为个性化学习提供支持。依据学习理论和学科知识结构构建知识图谱并结合学生需求、兴趣等进行分析，进而自适应推送与学习者相适配且具有互动性和生成性的学习资源(罗江华&张玉柳, 2023)。在人工智能通识课程建设中，需要建立包含认知特征（知识掌握度、思维偏好）、行为特征（学习时长、资源交互频次）和发展特征（职业规划、兴趣导向）的学习者画像，为学习者的个性化学习路径形成提供支撑。

4.4 依据课程评价迭代优化课程体系

人工智能通识教育的课程评价突破传统知识本位测评体系，转向基于真实场景的能力发展追踪。此类评价不仅关注最终产出，更强调对认知建构过程的持续观测。因此，人工智能通识课程的课程评价贯穿于整个课程实施过程中，这为课程建设的阶段性调整优化提供了证据支持。根据课程评价对教学要素进行结构性分析，可以从课程内容适配度、教学方法有效性、资源平台支撑力、课堂师生互动质量四个维度展开诊断，形成课程优化策略证据链。

为了充分利用课程评价对课堂体系的优化功能，综合课程理论和实践经验建立系统化的课程优化框架，为课程体系的迭代优化提供理论参考。例如，经过课程教学后学生的实践应用能力无法达到培养目标水平时，可以考虑在课堂中产教结合、校企合作等。另外，重视数智技术对课程评价的赋能，在一定程度上节省了人力物力，并缩短课程体系优化迭代周期。例如，利用多智能体构建数字孪生课堂，模拟真实课堂流程对教学设计进行预演，根据模拟结果给教学设计进行评价优化，达到缩短教学设计优化周期和降低试错成本的效果。这种方式不仅大大降低了试错成本，也提升了教学实施的决策效率。因此，将数智技术应用于课程评价能为课程体系的迭代优化注入动能，加速人工智能通识教育的实践进程。

5. 结语

智能时代背景下，高等教育机构必须将人工智能通识课程纳入人才培养体系，以培育具备技术认知与伦理判断力的新型人才。研究基于泰勒目标模式以课程开发的视角梳理了现有高等教育阶段中的人工智能通识课程体系建设实践，总结现阶段课程要素的特征，具体表现为：课程目标分层化、课程内容模块化、课程实施梯度化以及课程评价场景化。根据分析论证，提出了人工智能通识课程建设需要根据学校实际打造特色化课程目标、紧跟科技发展构建动态课程知识库、根据发展需求探索个性化学习路径以及依据课程评价迭代优化课程体系的四方面实施建议，推进人工智能通识教育高质量发展。

6. 致谢

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），国家教师发展协同创新实验基地建设专项“学生课堂学习投入识别与演化分析研究”（项目号：CCNUTEIII 2024-02），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）

参考文献

Ai@cam. (2025). Shaping AI for everyone. University of Cambridge, [2025-4-7]. https://www.cam.ac.uk/ai#group-section-Back-to-menu-YOVQyKXTWi

Centre for AI and Digital Ethics. (2025). What we teach. The University of Melbourne, [2025-2-4]. https://www.unimelb.edu.au/caide/study

Gick M. L. (1986). Problem-solving strategies. Educational psychologist, 21(1-2):99-120.

Krathwohl D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into practice, 41(4): 212-218.

School of Computer Science of Carnegie Mellon University. (2024). Artificial intelligence program. Carnegie Mellon University, [2024-11-12]. http://coursecatalog.web.cmu.edu/schools-colleges/schoolofcomputerscience/artificialintelligence/

Shute V. J. (2011). Stealth assessment in computer-based games to support learning. Computer games and instruction, 55(2):503-524.

Tyler R. W. (1949). Basic principles of curriculum and instruction. Chicago: The University of Chicago Press.

UNESCO. (2022). K-12 AI curricula: A mapping of government-endorsed AI curricula. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/ELYF6010

UNESCO. (2023). Guidance for generative AI in education and research. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/EWZM9535

UNESCO. (2024). AI competency framework for students. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/JKJB9835

戴瑞婷,李乐民(2024).面向产教融合的高校人工智能人才培养模式探索.高等工程教育研究, (03):19-25.

顾小清(2024).基于场景的测评:内涵特征、实践应用与未来展望.上海教育, (08):30-33.

郭炯,邹佳人(2025).场景化评价：技术赋能新时代教育评价改革的新趋向.中国远程教育, 45(01):71-85.

李国杰(2024).智能化科研（AI4R）：第五科研范式[J].中国科学院院刊, 39(1):1-9.

刘鹏,王平,李佳宁(2025).论数智化时代的知识转型及其教学应对.电化教育研究, 46(03):41-47.

刘智英,夏建国,刘娜(2011).高校分类发展视野下技术本科院校的特色发展之路.中国高等教育, (08):12-14.

罗江华,张玉柳(2023).多模态大模型驱动的学科知识图谱进化及教育应用.现代教育技术, 33(12):76-88.

满都拉,卢晓东(2018).通识教育背景下的大类招生与专业分流——以东京大学为例.高校教育管理, 12(03):21-27.

苗逢春(2024).为智能社会公民素养奠基的《学生人工智能能力框架》.中国电化教育, (11):1-12.

首都师范大学(2024). 我校召开“人工智能基础与实践通识课”建设工作部署会. 首都师范大学, [2024-7-16]. https://jwc.cnu.edu.cn/xwdt/jxxw/e9c289ed645641d1924d71412a222232.htm
王雪,何海燕,栗苹,等(2020).人工智能人才培养研究:回顾、比较与展望.高等工程教育研究, (01):42-51.
王竹立(2017). 面向智能时代的知识观与学习观新论. 远程教育杂志, 35 (03):3-10.
吴丹(2024). 武汉大学数智教育支撑体系建设指南.武汉:武汉大学出版社.
吴楠(2024). 南京大学发布人工智能通识核心课方案. 南京大学, [2024-3-7]. https://www.nju.edu.cn/info/1056/357721.htm
武迪,戴琼海(2021). 高校人工智能人才培育战略的道与路. 中国高等教育, (20):13-15.
武汉大学(2023). 武汉大学数智教育白皮书（数智人才培养篇）. 数智武大专题, [2023-11-16]. https://nic.whu.edu.cn/info/3331/13446.htm
余维杰,郑梦婷,张颖(2025). 中外人工智能素养框架研究. 图书馆杂志, 1-12.
张立迁(2018).人工智能发展关键在“人”.中国教育报, (002).
赵天润(2024). 【AI 大课】“AI 大课”12 问！复旦将在 2024-2025 学年推出至少 100 门 AI 领域课程. 复旦大学新闻网, [2024-6-5]. https://news.fudan.edu.cn/2024/0605/c31a141059/page.htm
浙江大学(2024). 全文来啦！《大学生人工智能素养红皮书（2024 版）》中文版全文发布. 浙江大学, [2024-6-18]. https://mp.weixin.qq.com/s?__biz=Mzg2ODY2ODI4OQ==&mid=2247489480&idx=1&sn=a02f3192db3926ac6de2aeaa22c748c1&scene=21#wechat_redirect
中华人民共和国教育部(2024). 教育部部长：将实施人工智能赋能行动，促进智能技术与教育教学、科学研究深度融合.中华人民共和国教育部, [2024-2-1]. http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2024/2024_zt02/mtbd/202402/t20240202_1114004.html
中华人民共和国教育部(2025).中共中央 国务院印发《教育强国建设规划纲要（2024—2035年）》.中华人民共和国教育部, [2025-1-19]. http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202501/t20250119_1176193.html
中华人民共和国中央人民政府(2024).习近平向 2024 世界智能产业博览会致贺信. 中华人民共和国中央人民政府, [2024-6-20]. https://www.gov.cn/yaowen/liebiao/202406/content_6958352.htm

Research on the Construction of Assessment Indicators for College Students' Learning Engagement in Blended Learning Environment

Huimin Lyu[1*], Yujiao Cheng[2], Dingding Guo[3], Yuelin Ding[4], Li Lan[5]
[1 2 3 4 5]South China Normal University
[*] lhmjyjs@163.com

Abstract: *With the wide application of blended learning mode in higher education, accurately assessing the degree of college students' input in the blended learning environment has become an important issue in educational research and practice.On the basis of systematically analyzing the current research status of blended learning, learning input and multi-source data integration at home and abroad, this study adopts the comprehensive assignment method, combining hierarchical analysis and entropy value method, to scientifically and reasonably determine the weights of the dimensions and indicators, and effectively avoid the excessive influence of a single indicator on the evaluation results.The establishment of evaluation indicators provides scientific basis and specific guidance for educational practice, assesses students' learning engagement status more accurately, provides powerful support for personalized teaching and learning interventions, and also provides new ideas and methods for the study of learning engagement in a blended learning environment.*

Keywords: blended learning, learning engagement, multi-source data, integrated weighting method

1.Introduction

With the rapid advancement of educational informatization, blended learning has emerged as a pivotal model in higher education, integrating the strengths of face-to-face and online learning. However, its cross-temporal and cross-modal characteristics pose unprecedented challenges to assessing students' learning engagement—a critical factor in evaluating educational effectiveness. Traditional assessment methods, constrained by singular data sources and subjective biases, often fail to comprehensively capture the multidimensional nature of engagement in blended environments(Dai L. T. ,2023). This limitation hinders the optimization of teaching strategies and the realization of blended learning's full potential.Recent developments in multi-source data integration offer innovative solutions(Wang M.-T. et al. ,2016). Existing evaluation frameworks are mostly dominated by the three dimensions of learning inputs, with a weak representation of the value of social inputs, and for the few four-dimensional evaluations, there are fewer analytical studies on expressions in the area of affective inputs. By synthesizing diverse datasets from virtual and physical learning spaces, it becomes possible to construct a more holistic and objective evaluation framework(Yin R., & He S. ,2023). In this study, through literature analysis and policy research, the evaluation framework of student learning engagement in a blended learning environment was developed, based on which, combined with the comprehensive assignment method, the weighting system of assessment indicators for blended learning engagement was

established. The research results aim to provide actionable insights for personalized instructional interventions and to improve the quality of blended learning practices.

2.Literature Review

2.1. Review of blended earning engagement

Current scholarship on blended learning engagement, while limited in volume, demonstrates thematic diversity spanning four interrelated domains. Research predominantly confirms a positive correlation between blended learning modalities and enhanced learner engagement(Shen X. et al, 2022), with recent advancements in educational informatization and AI applications further validating engagement as a critical quality metric for blended learning ecosystems(Li Y. & Xu L,2022). Academic inquiries have progressively shifted from analyzing dimensional constructs of engagement—particularly through learning activity taxonomies and pedagogical innovations—to developing conceptual models that reflect divergent epistemological interpretations of engagement mechanisms(Wang X.& Guo S. , 2022). Parallel investigations into determinants of engagement converge with broader learning engagement studies, identifying tripartite influences encompassing intrinsic learner characteristics, extrinsic instructional interventions, and environmental moderators, thereby providing theoretical scaffolding for variable selection in contemporary research frameworks(Ren Q. ,2021). Methodologically, measurement practices predominantly adapt legacy instruments from traditional classroom or online learning contexts(Zhang H. et al , 2023), relying heavily on self-reported data while increasingly incorporating multi-modal approaches that synthesize classroom observations with digital trace analytic(Gao F.& Lv J. ,2025;Ma F. ,2024). Notably, emerging hybrid assessment frameworks that transcend context-specific limitations offer promising paradigms for operationalizing blended learning engagement, though scholarly consensus regarding standardized metrics remains nascent, underscoring the imperative for context-sensitive measurement innovations in this evolving field.

2.2. Review of multi-source data on learning engagement

Researchers have defined multi-source data across various domains, with multimedia data analytics encompassing text, images, audio, and video. Current research on multi-source data fusion focuses on three key areas: algorithms and models, processes and techniques, and evaluation models.Multi-source data fusion algorithms range from simple methods to probabilistic, fuzzy inference, and artificial intelligence-based approaches. Fusion can occur at three levels: data-level (original data), feature-level (extracted features), and decision-level (local decisions from evaluation and reasoning).

In education, Zhang Zhi et al. developed a comprehensive quality evaluation model using big data(Zhang Z. & Qi Y.-G.,2017). This model collects student data from both formal and informal learning environments, both online and offline, creating a comprehensive data warehouse(Wang Y. ,2024;Liu Y et al,2024). The model standardizes multi-source data reflecting student quality, applies mathematical models for big data analysis, and generates digital profiles of individual and group student quality.This study provides valuable insights into data collection and integration processes, particularly for learning engagement research, by clarifying what data to collect and how it reflects specific information.

3.Method

3.1. Curriculum and Participants

This chapter establishes a blended learning input evaluation index system by collecting student data and combining the comprehensive assignment method.In order to ensure the representatives of student data, this study chooses to carry out in a comprehensive university in the first semester of 2024, the whole university elective course "Innovative Thinking", which mainly relies on the Learning Channel platform to carry out blended teaching, covering online independent learning and offline classroom face-to-face teaching two parts.The offline classroom is equipped with multimedia teaching equipment, such as multimedia computers, slide projectors and a stable network environment, which provides strong support for the teacher's knowledge teaching, group cooperative learning and classroom interaction and communication activities.The online course is carried out on the Learning Channel platform and supports the login of cell phones, tablet PCs and other terminal devices, which is convenient for students to participate in learning anytime and anywhere.There are 87 students in the course, coming from different grades and majors. In terms of the distribution of majors, the course covers a wide range of disciplines such as philosophy, economics and law, reflecting the diversity of disciplines.As for the gender composition of the students, there are 38 male students and 49 female students, with a certain gender balance.In summary, the diversity of these students in terms of specialty and gender makes their data better represent the situation of college students in blended learning situations, providing a reliable data base for the study.

3.2. Data analysis

Through the systematic sorting of the evaluation framework construction basis, principles and framework elements analysis, this study divided the blended learning input into four dimensions: behavioral input, cognitive input, emotional input and social input, and further clarified the specific evaluation elements under each dimension with the combination of literature sorting and experts' opinions, and put forward the evaluation framework of the blended learning input of college students with multi-source data as shown in Table 1.

Table 1. *Blended learning engagement degree indicator system*

PI	SI	TI
Behavioral input A1	Online behavioral input B1	Total online platform login time C1
		Online video completion C2
		Online test score C3
	Offline behavioral input B2	Offline class attendance C4
		(Questionnaire) Listening to lectures C5
		(Questionnaire) Taking notes C6
		(Questionnaire) Completing homework C7
Cognitive input A2	Online cognitive input B3	(Post) Understanding C8
		(Post) Applying C9
		(Posts) Analyzing C10
		(Posts) Creating C11
		(Posts) Evaluating C12
	Offline cognitive input	(Questionnaire) Understanding C13

	B4		(Questionnaire) Applying C14 (Questionnaire) Analyze C15 (Questionnaire) Creating C16 (Questionnaire) Evaluating C17
Emotional engagement A3	Online engagement B5	emotional	(Content analysis) positive C18 (Content analysis) Neutral C19 (Content analysis) Negative C20
	Offline Engagement B6	Emotional	(Expression recognition) positive C21 (Expression recognition) Neutral C22 (Expression recognition) Negative C23 (Questionnaire) Positive C24 (Questionnaire) Neutral C25 (Questionnaire) Negative C26
Social engagement A4	Online social engagement B7		Online discussion participation C27 Number of questions initiated online C28
	Offline social input B8		Number of offline interactions with teachers C29 Number of offline interactions with peers C30 (Questionnaire) Student-student interaction C31 (Questionnaire) Group work C32 (Questionnaire) Teacher-student interaction C33

In this study, classroom videos, questionnaires, and online platform data were collected according to the blended learning engagement multi-source data measurement. Among them, students' expression recognition was collected through classroom video, the frequency of seven emotions was recorded and converted into positive, negative, and neutral emotion scores according to the emotion scoring formula, and observed student-student interactions and teacher-student interactions; student self-assessment data of offline learning engagement was obtained through student self-reported questionnaires; students' offline attendance, completion of online learning, and interactions were collected through online platforms, and the textual data was crawled to get the scores according to the student comment text; the above collected data were summarized to get the raw data of students' blended learning engagement dimensions. Then, the weights of the blended learning input evaluation indexes are determined using a combination of subjective and objective weighting methods. Specifically, the subjective assignment part adopts the hierarchical analysis method, and the objective assignment part adopts the entropy value assignment method. This method ensures the scientificity and rationality of weight allocation. The specific weight calculation process is shown in Figure 1 below.

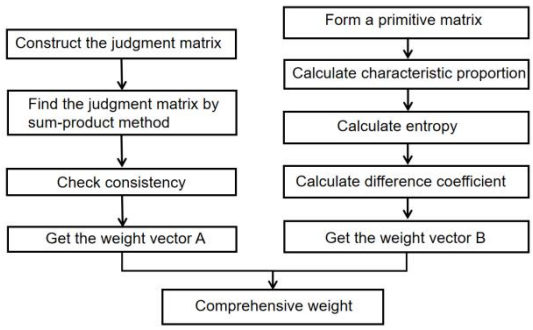


Fig. 1.Comprehensive weighting method weight calculation process

3.2.1. Hierarchical analysis

Hierarchical analysis, as a classical subjective weight calculation method, assigns weights to the elements of each level by decomposing the elements related to decision-making into hierarchical structures such as objectives, guidelines and programs, and then combining qualitative and quantitative analysis.Its core lies in constructing a hierarchical structure model and determining the relative importance of each element through expert scoring or two-by-two comparison, so as to provide a scientific basis for decision-making.In this study, 23 teachers and doctors were invited to assign values to the importance degree of each index according to the 1-9 scale method, so as to construct the judgment matrix for two-by-two comparison, and the evaluation result of one of the teachers was used as an example to construct the judgment matrix and carry out calculations, so as to obtain the judgment matrix of the indexes of guideline level, as shown in Fig. 2.

$$A=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 2 \\ 2 & 1 & \frac{1}{2} & 3 \\ 3 & 2 & 1 & 4 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 1 \end{bmatrix}$$

Fig. 2.Judgment matrix A

According to the judgment matrix A, the maximum eigenvalue method is used to solve the indicator weights, and the consistency test is performed on the evaluation results. After calculation, the consistency ratio is obtained less than 0.1, which indicates that the consistency of this judgment matrix is acceptable and the weight vector is valid.

The weights of the indicators for the remaining 22 experts and PhDs were determined and a consistency test was conducted.Given that the judgment matrix is the result of a fuzzy quantification of human experience, full consistency could not be achieved.After the test, five of the judgment matrices failed the consistency test, and the total number of valid results was 17.

The process of calculating the secondary and tertiary indicators is the same as that of the primary indicators.

3.2.2. Entropy method

The entropy assignment method is a data-based objective weight calculation method.This study uses the entropy value method, which first preprocesses the data, transforms the data into dimensionless values through indicator positive or negative (e.g., negative emotions), standardization or normalization, and then proceeds sequentially with the calculation of indicator weights, indicator entropy value, coefficient of variance, and the determination of indicator weights to ultimately calculate the composite scores of the samples, according to which the composite scores can be ordered and evaluated so as to get the weights of indicators. According to the comprehensive score, the samples can be sorted and evaluated, so as to get the indicator weights, and then determine the indicator weight system of blended learning input degree .

3.2.3. Comprehensive empowerment method

In this study, the comprehensive assignment method combining hierarchical analysis method and entropy value method is adopted to ensure the objectivity of the indexes through the

combination of subjectivity and objectivity.Firstly, the weight of hierarchical analysis method and the weight of entropy value method are determined, and then two coefficients α and β are introduced to indicate the importance degree of hierarchical analysis method and entropy value method in the comprehensive assignment respectively, and satisfy.In this study, 0.5 is taken for each, indicating that subjective and objective factors are equally important.

For the ith indicator (i=1,2,....n), the formula for its composite weight $W_{Combined}(i)$ is as follows:

$$W_{Combined}(i) = \alpha \times w_i^{AHP} + \beta \times w_i^{Entropy}$$

According to this formula, the composite weight of each indicator is calculated in turn, and finally the composite weight vector is obtained:

$$W_{Combined} = (W_{Combined}(1), W_{Combined}(2), \cdots, W_{Combined}(n))$$

3.3. Results

According to the first-level index weights of blended learning input degree derived from the above comprehensive assignment method, in order to present the index weights more intuitively and centrally, the index weights derived from hierarchical analysis method, entropy value method, and comprehensive assignment method are integrated, and the first-level index weights are shown in Table 2, the second-level index weights are shown in Table 3, and the third-level index weights are shown in Table 4.

Table 2. Blended learning engagement degree primary indicator weight

	A1	A2	A3	A4
Weight (hierarchical analysis)	0.2372	0.2635	0.3217	0.1776
Weight (entropy method)	0.1111	0.2906	0.3668	0.2316
Weight(Comprehensive empowerment method)	0.1742	0.2770	0.3443	0.2046

Table 3. Blended learning engagement degree secondary indicator weight

	B1	B2	B3	B4	B5	B6	B7	B8
Weight (hierarchical analysis)	0.1425	0.0947	0.1863	0.0772	0.1194	0.2023	0.0711	0.1065
Weight(entropy method)	0.0523	0.0588	0.211	0.0796	0.15	0.2168	0.0717	0.1599
Weight(Comprehensive empowerment method)	0.0974	0.0768	0.1987	0.0783	0.1347	0.2096	0.0714	0.1332

Table 4. Blended learning engagement degree tertiary indicator weight

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
Weight (hierarchical analysis)	0.0732	0.0348	0.0345	0.0062	0.0218	0.0231	0.0436	0.0373	0.0365	0.0366	0.0377
Weight(entropy method)	0.0334	0.0086	0.0103	0.0066	0.0151	0.0153	0.0218	0.0429	0.0417	0.0423	0.0412
Weight(Comprehensive empowerment method)	0.0533	0.0217	0.0224	0.0064	0.0185	0.0192	0.0327	0.0401	0.0391	0.0395	0.0395

	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22
Weight (hierarchical analysis)	0.0382	0.0154	0.0142	0.0151	0.0141	0.0184	0.0385	0.0368	0.0441	0.0512	0.0423
Weight(entropy method)	0.0429	0.0179	0.0180	0.0129	0.0177	0.0131	0.0498	0.0503	0.0499	0.0538	0.0686
Weight(Comprehensive empowerment method)	0.0406	0.0167	0.0160	0.0140	0.0159	0.0158	0.0442	0.0436	0.0470	0.0525	0.0555
	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33
Weight (hierarchical analysis)	0.0225	0.0383	0.0218	0.0262	0.0375	0.0336	0.0266	0.0325	0.0165	0.0153	0.0156
Weight(entropy method)	0.0243	0.0173	0.0203	0.0325	0.0333	0.0384	0.0438	0.0414	0.0272	0.0283	0.0192
Weight(Comprehensive empowerment method)	0.0234	0.0278	0.0211	0.0294	0.0354	0.0360	0.0352	0.0370	0.0219	0.0218	0.0174

4.Discussion

4.1. Introduction and Expansion of Social Input Dimensions

In this study, social input is included in the evaluation dimension of blended learning input, and the social dimension is comprehensively evaluated from both online and offline aspects.Specifically, online social input is quantified by analyzing students' interactive behaviors on the learning platform (e.g., the number of times they participated in discussions, the number of times they initiated questions, etc.); and offline social input is assessed by observing students' interactive behaviors in the classroom (e.g., the number of times they interacted with the teacher and their peers, the number of times they interacted with each other, etc.).This way of evaluating social input from multiple perspectives not only enriches the scope of learning input evaluation, but also provides a more comprehensive perspective for understanding students' interactive behaviors in a blended learning environment.

4.2. Establishment of weights for evaluation indicators for multi-source data

This study constructed a comprehensive multidimensional evaluation model with four dimensions: behavioral, cognitive, emotional and social, which provides a scientific basis for the comprehensive assessment of blended learning input.By integrating data from multiple sources (e.g., platform login time, video completion, quiz scores, classroom attendance, group cooperation performance, etc.), the model can more comprehensively reflect the state of students' learning engagement, and the weights of the indicators for evaluating the degree of students' learning engagement in a blended learning environment are established through the synthesis of hierarchical analysis and entropy value method.

5.Discussion

On the basis of constructing the evaluation framework, this study further determines the weights of indicators by carrying out empirical research. The stage of determining the weights of indicators adopts the comprehensive assignment method, combining the hierarchical analysis method with the entropy value method, to scientifically and reasonably determine the weights of the dimensions and indicators, which effectively avoids the excessive influence of a single indicator on the evaluation results.It more accurately assesses students' learning engagement status, provides strong support for personalized teaching and learning intervention, and promotes the effective implementation and development of blended learning mode in higher education.However, there are still some shortcomings in the current study, first, it fails to analyze students' offline social engagement from a more in-depth perspective.Second, the data collection method of student-student interaction and teacher-student interaction is cumbersome and lacks intelligence and automation.This manual collection method is not only time-consuming and labor-intensive, but also may lead to incomplete data and subjective bias.In view of the above research shortcomings, future research on blended learning inputs can be expanded and deepened in terms of combining the evaluation model with stage experiments and optimizing data collection and analysis methods.With the rapid development of generative artificial intelligence technology, subsequent studies can further explore its application in learning input data collection.

Reference

Wang, M.-T., Fredricks, J. A., Ye, F., et al. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. Learning and Instruction, 43, 16–26.

Dai, L. T. (2023). Target orientation, guiding principles and practical path of digital transformation of education driven by design intelligence. Journal of East China Normal University (Education Science Edition), 41(3), 12-24.

Yin, R., & He, S. (2023). Evaluation of blended learning inputs from the perspective of evidence-based teaching evaluation. Research on Electrochemical Education, 44(5), 120-128.

Shen, X., Yuan, D., & Chen, H. (2022). From "hybrid" to "integration": The design and practice of integrated teaching. Modern Educational Technology, 32(4), 40-49.

Li, Y., & Xu, L. (2022). Teaching power model construction of university teachers in blended learning environment - Based on the analysis of rooted theory. Education Development Research, 42(9), 71-79.

Ren, Q. (2021). A study of the effect of motivational regulation on learning engagement in college English classroom under blended teaching environment. Foreign Language E-Learning, 45(1), 44-50, 60, 7.

Wang, X., & Guo, S. (2022). Practice and inspiration of online-offline integrated teaching at Tsinghua University. Modern Education Technology, 32(4), 106-112.

Zhang, H., Liu, Y., & He, S. (2023). The construction of blended teaching quality evaluation index system for sports professional motor skills courses. Education Theory and Practice, 43(21), 57-61.

Gao, F., & Lv, J. (2025). Research on factors influencing students' online learning satisfaction in blended teaching. Jiangsu Higher Education, 41(2), 78-85.

Ma, F. (2024). An experimental study on the impact of blended learning approach on learning effect in secondary school mathematics based on digital platform. Shanghai Educational Research, 44(11), 56-66.

Li, M., Qi, X., & Zhang, L. (2024). A study on the developmental trajectory of college students' engagement in blended academic English learning. Foreign Language E-Learning, 46(5), 60-68, 113.

Zhang, Z., & Qi, Y.-G. (2017). Construction of multi-source and multi-dimensional comprehensive quality evaluation model based on big data. *China Electrochemical Education*, *(09)*, 69–77+97.

Wang, Y. (2024). The construction and practice of blended learning community in English language theory courses - Taking "English Pragmatics" course as an example. Foreign Language Society, 45(3), 90-96.

Liu, Y., Jai, R., & Ren, Z. (2024). Research on the impact of blended learning on college students' higher-order thinking ability in the digital era. Modern University Education, 40(6), 92-100.

Zhang, S., Gao, L., & Yang, Y. (2023). Dilemmas and improvement strategies of blended teaching of Civics and Politics in higher education programmes. Educational Theory and Practice, 43(3), 57-60.

教育数字化转型下可信数据空间的构建路径

The Path of Building Trusted Data Space under Digital Transformation of Education

曾颖洁，刘佳音，刘毓，王涛*
[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 湖北 武汉 430079
* tmac@ccnu.edu.cn

【摘要】 教育数字化转型背景下，数据孤岛化与低可信度问题严重制约教育数据价值的释放。本文基于新质生产力视角，提出构建教育可信数据空间的系统性框架，旨在通过技术架构与治理机制的协同创新，推动教育数据的智能化转型。研究指出，当前国内教育数据共享面临数据主权归属模糊与跨系统兼容性不足等核心挑战。基于此，提出基于可信数据空间理论的教育数据治理框架，通过构建五层协同架构体系，融合区块链、身份治理等技术，旨在为解决教育数据流通多维困境提供系统性解决方案。通过场景推演验证，该架构支持区域教育资源智能调配与 AIGC 工具跨域训练，有效平衡数据隐私与效用。研究强调，可信数据空间需依托新型数字基础设施与教育主体素养提升，以“治理机制创新、技术工具嵌入、业务场景重构”推动教育生产力向数据智能范式转型，为教育资源的优化配置与教育服务提质增效提供理论支撑与实践路径。

【关键词】 可信数据空间；跨域数据共享；数据隐私保护；教育数字化转型

Abstract: *In the context of education digital transformation, the problems of data siloing and low trustworthiness seriously constrain the release of education data value. Based on the perspective of new quality productivity, this paper proposes a systematic framework for constructing a credible data space in education, aiming to promote the intelligent transformation of education data through the synergistic innovation of technical architecture and governance mechanism. The study points out that the current domestic education data sharing faces the core challenges of fuzzy data sovereignty attribution and insufficient cross-system compatibility. Based on this, an education data governance framework based on the theory of trusted data space is proposed, which aims to provide a systematic solution to the multidimensional dilemma of education data circulation by constructing a five-layer synergistic architecture system and integrating blockchain, identity governance and other technologies. The architecture supports the intelligent deployment of regional educational resources and cross-domain training of AIGC tools, effectively balancing data privacy and utility, as verified by scenario projection. The study emphasizes that credible data space needs to rely on new digital infrastructure and education subject literacy enhancement to promote the transformation of education productivity to data intelligence paradigm through “innovation of governance mechanism, embedding of technology tools, and reconfiguration of business scenarios”, and to provide theoretical support and practical paths for optimal allocation of education resources and enhancement of the quality and efficiency of education services.*

Keywords: Trusted data space, cross-domain data sharing, data privacy protection, digital transformation of education

1. 前言

随着数字化时代的到来，教育领域的转型面临前所未有的机遇与挑战。然而，教育数据的价值释放面临诸多障碍，尤其是“孤岛化”现象和低可信度问题，这些问题阻碍了数据的有效利用。因此，突破这些制约因素，释放教育数据潜力，成为亟待解决的重要课题。为此，本文旨在提出可用的教育可信数据空间设计范式，探讨如何通过合理架构设计推动教育数据

的智能化转型，并揭示技术架构与教育业务场景之间的耦合机制，为教育领域的数字化转型提供理论支持与实践指导。

2. 相关研究

随着教育数据应用的不断深化，教育数据空间的研究成为学术界与实践领域关注的焦点。在国际层面，多个国家和地区已在探索如何通过构建教育数据空间来推动教育数字化转型。例如，欧盟通过打造数据空间(Data Spaces),为数据要素流通及服务提供底层基础设施，促进教育数据的流通与利用 (赵琳等人, 2024)，为教育政策制定和资源配置提供支持。美国则通过州级教育数据协作网络实现了不同州间教育数据的共享与协作，有效提升了教育数据的跨系统互通性和可操作性 (杨屿航等人, 2023)。这些国际实践展示了教育数据空间在推动教育现代化方面的重要作用，也为全球教育数据管理和利用提供了宝贵经验。

然而，国内在教育数据空间的建设方面仍面临一系列瓶颈。首先，数据主权归属的问题一直是制约教育数据共享的关键因素之一。不同部门和机构对于数据的控制权和使用权存在争议，导致数据资源配置低，供需之间难以进行有效匹配(杨艳等人, 2024)。其次，跨系统兼容性不足也是国内教育数据空间建设的一大障碍。教育领域内的数据管理系统众多且彼此独立，缺乏统一的标准和接口，使得数据的跨平台流通变得困难 (董晓辉等人, 2019)。国内教育数据共享的低效性主要体现在数据格式不统一和信息传递的滞后性等方面，这严重限制了教育数据的潜在价值释放。

教育数据空间通过数据智能化和可信协同作用，能够有效提升教育资源的配置效率和决策精准度，从而实现生产力的提升。相关跨学科的文献表明，数据空间的构建不仅仅是技术问题，更是制度创新与生产力提升的关键环节。在教育领域，通过建立高效的数据空间，能够突破传统教育生产力的瓶颈，推动教育事业向更加智能化、精准化的方向发展。因此，教育数据空间在提升教育全要素生产率和推动教育质量提升方面具有重要的潜力和现实意义。因此，如何突破这些瓶颈，构建适应中国教育发展的数据空间，成为当前研究的核心问题。

3. 教育可信数据空间架构设计

为更好地利用技术解决教育数据要素流通在实践中的困境，亟需构建全社会要素资源的网络化共享、集约化整合、协作化开发和高效化利用 (冯婷婷等人, 2024)。针对教育数据资源流通质量低、系统交互性差及多方主体权责失衡等问题，本文提出基于可信数据空间理论的教育数据治理框架，通过构建五层协同架构体系，融合区块链、身份治理等技术，旨在为解决教育数据流通多维困境提供系统性解决方案。

3.1. 数据源层

数据源层作为教育数据的最初来源，主要涵盖 PC 端、移动端、智能设备及其他数字化教育工具等多个平台。其核心功能为数据收集、初步筛选及数据溯源，通过对数据的采集时间、地点、设备等元数据进行记录，并对所有数据采集终端进行身份认证和设备信任评估，确保数据的可追溯性，保证数据的来源可信。为确保数据的真实性与不可篡改性，本研究提出的功能实现主要依赖区块链技术与数字签名技术。具体而言，通过区块链的分布式账本特性，结合哈希算法和共识机制，同时，采用增强型身份治理（IAM）技术，为数据采集终端和用户分配唯一身份标识 (Esposito C et al., 2021)，并结合多因素认证（MFA）机制，进一步提升身份认证的可靠性和安全性 (Wang Q et al, 2023)。此外，利用可信代理（Trusted Proxy）对数据源进行加密处理，确保数据在采集端的安全性 (S. W er al., 2016)。这种结合区块链、

数字签名、增强型身份治理以及可信代理的技术方案，能够为数据流通生态提供全方位的安全保障，满足零信任架构下的动态访问控制需求。

3.2. 传输处理层

传输处理层的核心功能是实现数据的加密传输与初步处理。在零信任架构下，该层需确保数据在传输过程中的严格加密与动态访问控制。为此，传输处理层需采用加密协议（如SSL/TLS）对数据进行加密传输，确保数据在传输过程中的保密性和防篡改性。同时，通过细粒度的访问控制策略，确保只有经过授权的人员能够访问敏感数据。无论是对原始数据进行去噪、格式化和标准化处理，还是进行数据分类、存档和生命周期管理，都需要依赖高效的算力设施和先进的数据模型，才能确保数据分析和机器学习模型能够在海量数据下高效运行，从而提升数据处理的性能和效率。同时，通过可信代理对数据流进行加密和解密 (U. D et al., 2023)，并动态调整数据传输路径，以应对潜在的安全威胁。

3.3. 中间服务层

在零信任架构下，中间服务层通过微分段（Micro-Segmentation）和持续信任评估（Continuous Trust Assessment）技术增强其安全性 (C. E et al., 2018)，同时实现灵活的数据管理和高效的服务对接。通过将中间服务层划分为多个独立的安全区域，每个安全区域根据数据的敏感度和业务需求配置独立的访问控制策略，有效限制了潜在威胁的横向移动从而降低安全事件的扩散风险。同时，部署信任评估引擎（Trust Assessment Engine），实时监控数据访问行为，并结合用户行为分析（User and Entity Behavior Analytics, UEBA）技术，动态调整访问权限。该引擎能够基于用户行为模式、设备状态和上下文信息，持续评估访问请求的信任级别，确保每次访问请求均符合安全策略 (Martín A G et al., 2021)。通过支持自定义算法模型的运行与优化，例如个性化推荐算法、学习路径优化模型等。通过提供灵活的算法开发环境和高效的计算资源，中间服务层能够满足教育平台在智能化服务方面的多样化需求。中间服务层还需具备数据分级与分类功能，能够根据数据的隐私性和重要性将其划分为不同的敏感度级别。基于数据分类结果，实施差异化的管理策略，例如：个人隐私数据需要严格的加密和访问控制，而公开课程内容则可以开放访问。这不仅实现了数据的精细化管理和动态安全防护，还为教育平台的多样化服务提供了强大的技术支持。

3.4. 数据控制层

为确保数据的访问与操作权限仅授予经过授权的用户或系统，防止未经授权的数据泄露或篡改，数据控制层承担着至关重要的角色。数据控制层通过全面的监管与管理机制，实现对数据访问的实时监控、审计与记录，涵盖每一次访问、修改和删除操作，从而保障数据使用过程的透明性。在数据的全生命周期管理中，数据控制层需依据相关法规和标准，对不再使用或达到存储期限的数据进行安全销毁。这一过程不仅符合数据保护法规的要求，还能有效防止信息泄露。以学校的学生信息管理系统为例，教师在操作学生成绩时，系统需通过严格的身份认证机制，结合多因素认证（MFA）技术 (Wang C et al., 2023)，确保用户身份的合法性。此外，系统基于身份、设备状态、上下文信息等多维度属性，采用基于属性的访问控制（ABAC）模型，对访问请求进行动态授权 (Li J et al., 2022)。所有成绩查询和修改操作均会被自动记录，形成完整的访问日志，以实现数据访问过程的可追溯性。当学生成绩数据超过保存期限时，系统将依据预设的安全策略进行数据的安全销毁，确保数据的完整性和保密性。通过上述措施，数据控制层不仅能够有效防止未经授权的数据访问和篡改，还能满足教育领域对数据安全与合规性的严格要求，为教育数据的可信流通提供坚实保障。。

3.5. 数据应用层

数据应用层作为教育可信数据空间的实际应用层，主要涉及数据的可视化、汇聚与共享。在此层，需实现“可用不可见”的数据共享模式，通过隐私保护计算技术（如安全多方计算、联邦学习、可信执行环境等）确保数据在使用过程中的隐私性。同时，基于数据的隐私性的重要性，实施最小化访问控制策略，确保用户仅能访问其所需的最小数据集。该层整合来自不同平台和层级的教育数据，形成全面的教育数据集，支持数据的汇聚、共享与协同应用。通过数据治理和数据流通机制，促进不同学校之间共享教学资源，以及跨地区教育研究机构共享数据集进行联合研究。此外，数据应用层通过数据可视化技术（如统计图表、仪表盘、地图等），将复杂数据直观呈现，帮助教育管理者、教师、学生等不同角色快速理解数据背后的信息和趋势。同时，利用大数据分析技术和机器学习算法，对教育数据进行深入挖掘，发现潜在规律和趋势。例如，通过机器学习算法进行学生成绩预测、个性化学习路径推荐等，以支持教育的智能化应用。这种多层次的数据应用模式不仅提升了数据的利用效率，还为教育领域的数字化转型提供了坚实的技术支撑。

4. 场景推演与可行性分析

本文基于设计科学（Design Science）的架构评估方法，通过理论推演和技术可行性论证，提出了两种典型场景的替代案例验证。设计科学方法强调通过解决实际问题并结合创新技术，验证解决方案的有效性与可行性。在此框架下，我们不仅关注理论上的推导，还着重进行技术层面的可行性验证，以确保所提出方案的实施具有现实基础。

4.1. 区域教育资源动态调配

在传统的区域教育资源调配模式中，往往依赖人工统计与手工调配，且信息流通滞后，导致资源分配不均和反应时间延长 (孟丽菊等人, 2024)。例如，在一些学校，特定学科或设备的资源利用率较高，但由于信息滞后，相关资源无法及时调配，造成教育资源的浪费或不足。为解决这一问题，可通过数据空间实时聚合各校的师资、设备与课程数据，并结合智能合约技术，实现资源调配的自动化与智能化。具体而言，当某一学校的实验室利用率超过90%时，智能合约将自动触发资源共享建议，通过系统推送相关信息，提示周边学校可以共享该实验室资源。此举能够显著提高资源使用效率，减少资源浪费。从技术角度看，区块链技术能够提供高度的透明性与安全性，确保数据的真实性与不可篡改性 (余胜泉等人, 2019)。在具体实现中，可以采用 Hyperledger 作为底层架构，其每秒交易处理量（TPS）足以满足区域教育资源的实时调配需求。此外，智能合约的执行可以保障各参与方的利益，实现自动化的资源共享协议。根据性能测试数据，Hyperledger 能够支持每秒上千笔交易，这对于大规模教育资源的动态调配来说是可行的。

4.2. AIGC 教育工具的数据供给

在教育领域中，AIGC（人工智能生成内容）教育工具的研发确实需要大量的跨校数据进行训练，以提升模型的准确性与适应性。然而，隐私问题成为了一大挑战，尤其是在不同地区和国家之间的数据交流时，如何保障数据隐私，避免泄露敏感信息，成为不可回避的问题。

为应对这一挑战，可以结合联邦学习与数据沙箱技术，在不出域的前提下生成合成数据集进行训练。联邦学习允许多个机构在本地进行模型训练并共享模型参数，而不需要交换原始数据。数据沙箱进一步确保在处理过程中，敏感数据始终留在本地环境中，避免数据泄露。通过构建多层协同架构体系，融合区块链、身份治理等技术，建立跨区域数据共享机制与标

准化体系，旨在为解决教育数据流通多维困境提供系统性解决方案。这种方法不仅能够保护数据隐私，还能够促进数据要素资源的价值转换，提升数据流通效率，降低数据共享的门槛。

5.讨论与展望

在推动教育数字化转型的过程中，构建和完善教育数据空间是关键。首先是加强区域新型基础设施建设，推动 5G、大数据、云计算、人工智能等新一代信息技术的应用，持续建设信息网络、平台系统、数字资源、智慧校园、创新应用、可信安全等新型基础设施。打造智慧教育公共服务平台，统筹学习、教学和管理过程中的大数据，建立教育大数据仓，促进教育数据的贯通共享，组建教育大脑，统筹推进数据融合融通，面向学生、教师和学校建立数据应用和分析模式。其次，推进课堂教学过程数字化，教育改革的阵地在课堂，课堂教学是数字化转型的核心。探索基于各种生态的课堂教学过程数字化方式，从教学内容、学习资源、教学过程等方面进行数据采集、分析和应用，实现教学过程的数字化。并且，教育数字化转型的关键是做好教育大数据工作，这涉及到教育活动全过程中产生的海量教育大数据具有巨大的潜在价值。教育大数据可以促进教育从“用经验说话”变为“用数据说话”，实现数据创新、数据管理与数据决策的数字化新形态。通过这些措施，可以有效地构建和完善教育数据空间，为教育数字化转型提供坚实的基础和支持。

尽管上述方案具有显著优势，但在实际应用中仍面临一定的局限性。首先，初期部署的成本较高，尤其是在教育领域，很多学校或机构的数字化基础设施尚不完善，可能需要大量的投入来建设相关系统。其次，方案的实施依赖于教育主体的数字素养，教师和管理人员的数字化能力直接影响方案的实施效果。如果教育主体对新技术的接受度较低，可能会导致方案实施困难或效果不佳。

结论

通过“治理机制创新、技术工具嵌入、业务场景重构”三层作用，可信数据空间有效推动了教育生产力向数据智能范式转型。首先，治理机制创新为教育系统提供了规范化的管理框架，确保数据使用的透明性和安全性，提升了各方参与者的信任度。其次，技术工具嵌入为教育实践提供了强大的数据分析和决策支持能力，增强了教育活动的精准性和效率。最后，业务场景重构通过将数据智能与教育业务深度融合，促使传统教育模式向个性化、智能化发展，实现了教育资源的优化配置和教育过程的动态调整。

这种转型不仅提升了教育服务的质量和效率，也为政策制定者和技术开发者提供了新的理论参考。通过对可信数据空间的应用探索，可以更好地理解如何在现代教育环境中推动技术与治理的协同发展，进而促进教育领域的创新和发展。

参考文献

董晓辉,郑小斌&彭义平.(2019).高校教育大数据治理的框架设计与实施.中国电化教育(08),63-71.

冯婷婷,刘德建,黄璐璐,曹培杰&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.

孟丽菊&张心誉.(2024).区域协调推进高质量教育体系建设：何为、难为与应为——兼评《我国教育综合发展水平区域差异研究》.教育理论与实践,44(31),30-35.

杨艳&林凌.(2024).数据要素高质量供给：内涵解析、困境挑战与规制设计.电子政务(11),15-

26.doi:10.16582/j.cnki.dzzw.2024.11.002.
杨屿航&马金晶.(2023).美国信息技术赋能教育的政策和实践演进.基础教育参考,(05),61-70.
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.
赵琳,钱雨秋&郑汉.(2024).欧盟数据要素市场培育政策、实践与模式.图书馆论坛 44(12),151-160..
Cayirci, E., & De Oliveira, A. S. (2018). Modelling trust and risk for cloud services. Journal of Cloud Computing, 7, 1-16.
Din, I. U., Bano, A., Awan, K. A., Almogren, A., Altameem, A., & Guizani, M. (2021). LightTrust: Lightweight trust management for edge devices in industrial Internet of Things. IEEE Internet of Things Journal, 10(4), 2776-2783.
Esposito, C., Ficco, M., & Gupta, B. B. (2021). Blockchain-based authentication and authorization for smart city applications. Information Processing & Management, 58(2), 102468.
Li, J., Zhang, Y., Ning, J., Huang, X., Poh, G. S., & Wang, D. (2020). Attribute based encryption with privacy protection and accountability for CloudIoT. IEEE Transactions on Cloud Computing, 10(2), 762-773.
Martín, A. G., Beltrán, M., Fernández-Isabel, A., & de Diego, I. M. (2021). An approach to detect user behaviour anomalies within identity federations. computers & security, 108, 102356.
Sun, W., Yu, S., Lou, W., Hou, Y. T., & Li, H. (2014). Protecting your right: Verifiable attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud. IEEE Transactions on Parallel and Distributed Systems, 27(4), 1187-1198.
Wang, C., Wang, D., Duan, Y., & Tao, X. (2023). Secure and lightweight user authentication scheme for cloud-assisted Internet of Things. IEEE Transactions on Information Forensics and Security, 18, 2961-2976.
Wang, Q., & Wang, D. (2022). Understanding failures in security proofs of multi-factor authentication for mobile devices. IEEE Transactions on Information Forensics and Security, 18, 597-612.

大模型智能体评估教学实施新方法：使用教学设计达成度

A new method of teaching implementation for large-model agent evaluation: use of instructional design to achieve degree

任德博[1]，龙陶陶[1]，陈增照[1]，杜睿擎[1]
[1]华中师范大学人工智能教育学部
deboren@mails.ccnu.edu.cn

【摘要】 大模型智能体及思维链推理的发展解决了文本评估中的挑战。教学内容的一致性长期以来一直是教师评估的重要组成部分，单纯基于国家标准的一致性没有考虑到教师的个人发展，使用教学设计作为中介的条件干预教师教学内容优化成为了更加的方法。本研究用 deepseekR1 大模型取代人工评估，开发一种使用智能代理评估教学设计达成度的方法。教学设计文本可以识别的五个部分：教学目标、重点难点、教学过程、教学策略和评估。利用智能代理工具进行文本话语分析，并基于语义匹配和成就范围开发了一种新的计算方法，赋予权重(ω)，增强其准确性和稳定性。该模型通过这种方法进行了微调。结果显示，新手教师与专家教师在教学设计的不同部分达成度上存在显著差异。
【关键词】 教学设计；教学达成度；AI 智能体；教学内容；课程标准

Abstract: *The development of large model agents and chain-of-thought reasoning solves the challenges in text evaluation. The consistency of teaching content has long been an important part of teacher evaluation, and the consistency based solely on national standards does not take into account the personal development of teachers, and the use of instructional design as a mediating condition to intervene in the optimization of teachers' teaching content has become a more method. In this study, the deepseekR1 large model was used to replace the manual assessment, and a method was developed to evaluate the achievement of instructional design using intelligent agents. There are five parts that can be identified in an instructional design text: instructional objectives, key difficulties, instructional process, instructional strategies, and assessment. The intelligent agent tool was used for text utterance analysis, and a new calculation method was developed based on semantic matching and achievement range, which was assigned weight (ω) to enhance its accuracy and stability. The model is fine-tuned in this way. The results showed that there were significant differences between novice teachers and expert teachers in the achievement of different parts of instructional design.*
Keywords: instructional design, teaching achievement, AI agent, teaching content, curriculum standards

1.引言

组织和持续改进教学内容对于提高教学效果和学生学习至关重要，因为它们直接影响学生的表现。教学内容不仅包括所教授的知识，还包括内容的顺序、活动和其他相关元素（Gamoran et al., 1997）。教师负责安排这些元素，因此有效组织教学内容成为教学评估中的一个重要挑战（Porter, 1995）。先前的研究表明，基于主观经验的解释或模糊的标准在评估教学内容时可能导致教师脱离评估过程，从而削弱其有效性(Nägel et al., 2023；Schmidt et al., 2011)。中共中央、国务院印发了《教育强国建设规划纲要（2024—2035 年）》强调了高素质教师队伍建设和学生核心素养的提升，教学内容是核心素养的主要体现，是向学生传递核心素养的重要工具。

2022 年我国新引发基础教育课程核心标准，取消了三维教学目标，提倡学科核心素养的培养。教学设计作为教师教学的第一手工具，是教师结构化和优化内容的宝贵工具，它定义了教学中关键组成部分，如教学目标、重点和难点、教学策略、资源、过程（活动、案例分

析）和评估（Moore & Kearsley, 1996）。在建构主义理论中，教学设计要强调以学生为中心，不仅要求学生由外部刺激的被动接受者和知识的灌输对象转变为信息加工的主体、知识意义的主动建构者，强调环境和媒体的作用（何克抗，1997）；智能时代的教学设计，强调在课前课中课后促进学生的智能发展（刘邦奇，2016）。当前我国教师的教学设计承载了其对核心素养的理解、教学观、学生观，对其课程内容有很大的指导意义。

教学设计达成度代表了教学内容对于教学设计的完成情况，当前研究越来越关注将教学内容、核心素养和学生学习成果与教学设计达成度相匹配（Iaconangelo et al., 2022; Van et al., 2023）。有效的教学设计需要教师理解和正确实施（Purwaningsih et al., 2024），理解和实施设计计划对于新教师的专业成长至关重要（Lin et al., 2024；侯红玲等人，2019；邹吉权等人，2018），但存在人工评估成本高、效率低、标准不统一的特点。大语言模型的出现对教育领域产生了重大影响，尤其是在支持教师方面，其影响显而易见为了辅助教师成长增强教师教学实施的评估（Dai et al., 2024; Gill et al., 2024; Naidu, 2023），尤其是在评估对齐方面，是它能够应对定性评分的挑战。引入思路链（CoT）推理提供了一种新颖的解决方案（卢宇, 余京蕾 & 陈鹏鹤，2024），思路链使模型能够阐明解决问题的过程，提供了对人类任务解决方法的见解。

本团队基于大模型思维链开发教学设计达成度评估和建议工具，辅助教学内容评估。通过提供高质量的案例研究和现实的推理逻辑，不断优化现有模型，从而增强评分功能。

2.评估标准的建立和评估工具的开发

大语言模型通过将评分机制和尺度纳入模型，并根据专家评级进行调整，系统不断优化和更新提示工程，确保增加客观性和可操作性。在系统智能评估中测量达成度指数标准的确立是一个关键过程，Smith 等人（1992）强调其计算应系统地基于标准。本研究定义了达成度标准，以教学设计为基准评估每个标准是否反映在教学内容中。这些标准包括教学目标、关键和难点、教学策略（创新）、教学过程（活动、案例研究）以及教学评价（形成性和总结性）（Gagne et al., 2005; Weston et al., 1995）。我们的方法与 Porter 的人工评分方法不同（Porter et al，2011），将具体的数值标准纳入教学设计中，因此需要一个新的评价标准提供给大模型作为理解框架。我们将语义匹配与 Porter 的描述性方法相结合，将其分为五个层次，并根据语义的匹配度对每个层次给予了部分量化的评分标准，每条标准对应一个教学设计的部分。具体细节见下表 1。

表 1 大语言模型评分标准

评分	评分标准
完全达成（5）	1. 教学内容充分体现了教学目标。 2. 重点和难点用适当的时间和篇幅加以强调。 3. 教学策略和方法已得到充分实施。 4. 教学过程严格遵循预先设计的顺序。 5. 完全体现运用评价方法。
大部分达成（4）	1. 超过三分之二的教学目标被纳入到内容中。 2. 重点和难点得到一些强调，尽管它们的长度和时间分配没有完全平衡。 3. 超过三分之二的教学策略得到了实施。 4. 教学过程基本完成，但顺序和比例缺乏完全的一致性。 5. 评估方法为学生提供了及时的反馈，尽管与原始计划存在偏差。
适中达成（3）	1. 超过一半的教学目标得到满足。 2. 包含关键难点，但与一般知识点混编。

	3. 只有计划中的教学策略和方法的一半得到了实施。 4. 大部分的教学过程已经完成，尽管顺序缺乏清晰度。 5. 评估方法给一些学生提供了反馈，但缺乏对整个班级的全面评估。
部分达成（2）	1. 只有三分之一的教学目标得到满足。 2. 关键和难点包括在内，但其强调程度甚至低于一般知识点。 3. 只有三分之一的计划教学策略和方法已经实施。 4. 教学过程不完整，顺序混乱。 5. 仅提及评估方法，但未提供预期反馈。
没达成（1）	1. 教学目标在内容中没有体现出来。 2. 关键和难点完全省略。 3. 设计中的任何教学策略或方法都没有得到实施。 4. 教学过程完全偏离了原计划。 5. 教学过程中完全没有评估和反馈。

教学设计元素的重要性各不相同。Martin（2011）在一项关于教学设计元素对齐的研究中，将教学目标列为最重要，而教学材料则被列为最不重要。然而，教学设计原则和学校政策可能有所不同。例如，在以活动为基础的课堂中，教学活动被视为核心组成部分（Han et al., 2015)。因此，在计算教学设计与教学内容之间的对齐之前，引入了一个权重因子ω。本研究中重要性的水平是通过层次分析法（AHP）确定的，权重因学校和地区而异。我们收集了北京40位专家教师和学校管理人员的意见，他们对评分进行了0到10的打分。评分分为两轮进行，第二轮紧随第一轮之后（CI1 = 0.005; CI2 = 0.10）。最终结果使用了两轮的平均分数，显示教学目标的重要性最高，而教学评估的重要性最低（CI = 0.004)。评分结果及其对应的权重见表2。

表2APA权重评分结果

	教学目标	教学重难点	教学过程设计	教学策略	教学评价	ω
教学目标	1	1.435	1.3135	1.264	1.4235	.253
教学重难点	0.679	1	1.186	1.227	1.414	.207
教学过程设计	0.8945	1.075	1	1.3105	1.451	.221
教学策略	0.6645	0.7005	0.7755	1	1.1335	.161
教学评价	0.707	0.7075	0.6915	0.991	1	.158

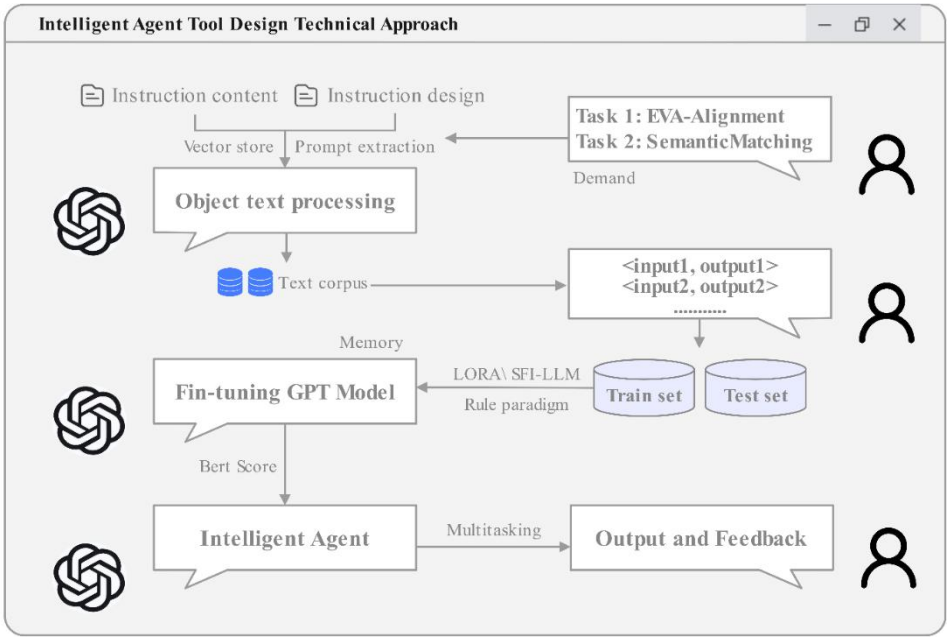


图 1 教学设计达成度测量的思维链设计过程

初始数据集是通过预先收集 45 节课的教学设计和内容创建的。由于这不是一个分类问题，因此需要进行模型预微调。我们采用了低秩适应（LoRA）方法（Mao et al., 2024），该方法有效地增加了可训练层，同时显著减少了 GPU 内存使用。预微调结果由 40 位教育学专家即行业老师通过层次分析进行了评估。专家们提供了改进建议，作为对比数据。我们比较了 DeepSeek 和 GPT 模型的性能。GPT 在强化学习和处理速度方面优于 DeepSeek，而 DeepSeek 在专业领域和复杂推理方面表现出色（Rahman et al., 2025; Gao et al., 2025）。我们选择了 deepseekR1 模型作为代理开发的基础模型。接下来，使用了 DeepSeek 模型构建了智能体推理。对于语义匹配，我们采用了 Bert Score 方法，该方法基于生成文本的标准参考指标评估语义相似度（Zhang et al，2019）。学习率设置为 3e-4，训练轮数设为 80。所采用的过程范式包括：规则范式、文本提取、单项目对齐评估(A)和语义匹配统计(G)。最终，使用这些方法获得了更准确的结果，如图 1 所示。

为了证明软件的有效性，对专家打分和智能体对教学设计达成度的打分进行了对比验证，通过数据看出专家打分的一致性明显低于智能体打分的结果。而专家打分的结果和智能体打分的结果具有较好的一致性，不具有显著性差异（p>0.05）。说明智能体评分在教学意义上拥有一定的有效性，如表 3 所示。

表 3 专家评分与智能体评分对比

	智能体评分		专家评分	
	Mean（STD）	IAA	Mean（STD）	IAA
教学目标	4.08（0.71）	0.836	4.16（0.83）	0.713
教学重难点	4.12（0.58）	0.871	3.97（0.71）	0.749
教学过程	3.84（0.71）	0.733	4.05（0.51）	0.581
教学策略	3.91（0.36）	0.786	0.63（0.58）	0.671
教学评价	4.01（0.63）	0.814	4.17（0.78）	0.813

** *一级指标一致性 (IAA)*

3.新手教师和专家教师教学设计达成度比较研究

以北京地区合作实验学校为依托，开展一项涉及 20 名中学教师的比较实验。在选定的教师中，5 名是新手，5 名是专家，且事先未通知参与者。教师来自两个科目：语文和数学，并在各组中均匀分布。从学期中期开始，每隔一周收集两组的教学设计和视频，直至为期五周的阶段测试。教学视频被转录并分析。目的是考察新手教师与专家教师之间的对齐差异。计算了 20 名教师的对齐度（M=6.399；SD=0.397）。实验中，我们发现新手教师与专家教师在教学方法上的排列没有显著差异；然而，当分解每个组成部分时，差异变得明显。所有课程均使用 Shapiro-Wilk 检验进行了正态性测试，结果显示负偏斜（-0.508<偏度<-0.043），表明接近正态分布。采用 t 检验来比较差异。结果显示新手教师与专家教师之间没有显著差异（t=-1.525；p=.145）；不同教育阶段之间也没有显著差异（t=-1.007；p=.327）；在教学目标（p = .211）、重点难点（p = .662）以及教学评价（p = .817）方面，新手教师与专家教师之间也没有显著差异。然而，在教学策略（p = .000）和教学过程（p= .032）方面发现了显著差异。在教学策略方面，新手教师的达成度显著高于专家教师（t = 5.689），而在教学过程方面，专家教师的达成度显著高于新手教师（t = -2.180）。我们进行了回归模型分析，以考察每个因素的影响，发现教学目标显著影响教学重点和教学过程的对齐度，但对教学策略和教学评价的达成度影响较小。教学重点和教学过程的达成度显著影响教学内容的对齐度。

研究发现，新手教师在教学实施方面面临挑战，尤其是在课程内容进展上，其一致性得分显著低于专家教师。教学策略专家型教师达成度低于新手型教师引发我们的兴趣，并进行关于原因的讨论，例如与年龄相关的教学习惯的影响或教师职业倦怠的效果。研究表明，年长的教师更容易出现职业倦怠（TSANG 等，2022），他们常常忽视教学策略的重要性，将其视为维护权威的形式工具。这往往是有害的，促使人们重新思考专家教师的定义。诸如丰富的奖项或多年的经验等标准常被用来区分专家教师，但这些标准并不全面。在以学生为中心的课堂中，促进学生学习的能力应该是区分新手教师和专家教师的关键因素。

4.教学启发和未来研究

教学设计与教学内容的一致性计算代表了一种新的方法，改进了 Porter 对教学标准的应用。作为指导教学的文件，教学设计满足了个性化教师评估的需求，并解决了不同学校之间缺乏统一评估标准的问题。在我国，大多数教学设计需经过上级部门的审查和批准，并需与国家 2022 年新印发的课程标准中核心素养保持一致，以确保教学评估的一致性，从而保证教学设计的合理性。当前强调以学生为中心课程时，标准化的教学标准通常会导致更高的评分；然而，这种方法忽略了教师的个人优势。这就是为什么我们使用教学设计作为标准来评估一致性，因为它有效地解决了这一问题，并且与学生的学习表现有很强的相关性。教学设计提供了更全面的内容覆盖，因此比国家标准，更方便教师达成度提升，而后者可能覆盖范围有限。同样，我们认为教学中的差异化是教学的关键原则，作为衡量标准，教学设计特别针对个别学生进行教学。其正确应用可以更准确地反映教学内容的质量。

智能体的构建有效减少了人类评估中的主观性。与其他智能评估方法一样，教研员同样在这一过程中起着关键作用，给出教师更具体的建议而不是完全机械化的评分。而智能体可以根据提出的框架提供多元化的建议，帮助教师选择性改进。未来改进框架的研究一定会作为最主要的研究方向。虽然仍有待改进的不足，但是本研究依然提出了一种创新性的方法帮助青年教师提升教学内容，说明大模型在教师教育、课堂管理、教师培训和学生学习中，尤其是话语分析和提升建议上，有着极大的潜力。进一步创建一个自动化的评估系统。该方法消除了手动评估的主观性，为不同课堂的教师提供了更有效的支持。其有效性已得到验证，突显了其应用的教育意义。它为未来教学培训的改革提供了新的参考。

参考文献

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. Educational Evaluation and Policy Analysis, 19(4), 325-338.

Porter, A. (1995). Research news and comment: The uses and misuses of opportunity-to-learn standards. Educational Researcher, 24(1), 21-27.

Nägel, L., Bleck, V., & Lipowsky, F. (2023). “Research findings and daily teaching practice are worlds apart” – Predictors and consequences of scepticism toward the relevance of scientific content for teaching practice. Teaching and Teacher Education, 121, 103911.

Schmidt, W. H., Cogan, L. S., Houang, R. T., & McKnight, C. C. (2011). Content coverage differences across districts/states: A persisting challenge for US education policy. American Journal of Education, 117(3), 399-427.

中共中央国务院印发《教育强国建设规划纲要(2024—2035 年)》,(2025-01-20).人民日报,006.doi:10.28655/n.cnki.nrmrb.2025.000833.

Moore, M. G., & Kearsley, G. (2012). Distance education: A systems view of online learning.

何克抗. (1997). 建构主义的教学模式、教学方法与教学设计. 北京师范大学学报(社会科学版), (05), 74-81.

刘邦奇. (2016). “互联网+”时代智慧课堂教学设计与实施策略研究. 中国电化教育, (10), 51-56+73.

Iaconangelo, C. J., Phelps, G., & Gitomer, D. H. (2022). Dimensionality and validity of the content knowledge for teaching construct using cognitive diagnostic modeling and known groups comparisons. Teaching and Teacher Education, 114, 103690.

Van Leent, L., Walsh, K., Moran, C., Hand, K., & French, S. (2023). Effectiveness of relationships and sex education: A systematic review of terminology, content, pedagogy, and outcomes. Educational Research Review, 39, 100527.

Purwaningsih, E., Sutoyo, S., & Suryadi, A. (2024). CoMCoRe-LS: an instructional design to enhance pedagogical content knowledge of pre-service physics teachers. Journal of Turkish Science Education, 21(2), 324-344.

Lin, X. F., Luo, G., Luo, S., Liu, J., Chan, K. K., Chen, H., ... & Li, Z. (2024). Promoting pre-service teachers’ learning performance and perceptions of inclusive education: An augmented reality-based training through learning by design approach. Teaching and Teacher Education, 148, 104661.

侯红玲, 任志贵, 何亚银 & 赵永强. (2019). 基于 OBE 理念反向设计教学过程研究. 大学教育, (10), 57-59.

邹吉权,刘晓梅 & 牟信妮.(2018).高职成果导向的教学设计与实施.中国职业技术教育,(20),26-32.

Dai, W., Tsai, Y. S., Lin, J., Aldino, A., Jin, H., Li, T., ... & Chen, G. (2024). Assessing the proficiency of large language models in automatic feedback generation: An evaluation study. Computers and Education: Artificial Intelligence, 7, 100299.

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., ... & Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. Internet of Things and Cyber-Physical Systems, 4, 19-23.

Naidu, E. (2023). Leading Academics Believe Fears over ChatGPT Are Misplaced. University World News.

卢宇,余京蕾 & 陈鹏鹤.(2024).基于大模型的教学智能体构建与应用研究.中国电化教育,(07),99-108.

Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M., & Russell, J. D. (2005). Principles of instructional design.

Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. Educational technology research and development, 43(3), 29-48.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards: The new US intended curriculum. Educational researcher, 40(3), 103-116.

Martin, F. (2011). Instructional design and the importance of instructional alignment. Community College Journal of Research and Practice, 35(12), 955-972.

Han, B., Gu, J., & Song, T. (2016). An Activity-based Instructional Design For Search Algorithm Expression of Elementary Students. Journal of The Korean Association of Information Education, 20(2), 161-170.

Mao, Y., Ge, Y., Fan, Y., Xu, W., Mi, Y., Hu, Z., & Gao, Y. (2025). A survey on lora of large language models. Frontiers of Computer Science, 19(7), 197605.

Rahman, A., Mahir, S. H., Tashrif, M. T. A., Aishi, A. A., Karim, M. A., Kundu, D., ... & Eidmum, M. D. (2025). Comparative Analysis Based on DeepSeek, ChatGPT, and Google Gemini: Features, Techniques, Performance, Future Prospects. *arXiv preprint arXiv:2503.04783*

Gao, T., Jin, J., Ke, Z. T., & Moryoussef, G. (2025). A Comparison of DeepSeek and Other LLMs. *arXiv preprint arXiv:2502.03688*.

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). Bertscore: Evaluating text generation with bert. arXiv preprint arXiv:1904.09675.

Exploring the effects of AI feedback on affective-motivational outcomes in language learning: A Meta-analysis

Shuo Chen [1], Zhenzhen Chen[1*]
[1]School of Humanities, Beijing University of Posts and Telecommunications
[*] czz@bupt.edu.cn

Abstract: *Despite the growing interest in the application of artificial intelligence (AI) in language learning, there is still a lack of research on the effects of AI feedback on language learning, particularly affective-motivational outcomes in language learning. Through a meta-analysis based on 14 empirical studies, this study aims to explore the impact of AI feedback on learners' affective-motivational outcomes in language learning and identify relevant moderating variables. The study reports a large and significant effect size ($g = 0.850$, $p < 0.001$), and educational level, intervention duration, as well as affective-motivational factor are the moderating factors that explain the variance of the effect size. The research findings provide insights for educators on integrating AI feedback into teaching design and also offer a reference for future directions in related research.*

Keywords: feedback, AI in education; language learning; AI feedback in language learning

1. Introduction

Feedback is an essential component of the educational process, playing a crucial role in encouraging and reinforcing learning (Hyland & Hyland, 2006). Especially in language learning, learners rely more on feedback to identify and correct language errors, adjust their learning strategies, and strengthen their grasp of language rules (Hyland & Hyland, 2006). However, in the context of evolving educational needs, the limitations of traditional teacher feedback have become increasingly apparent. In large classes, teachers face challenges in delivering in-depth, personalized and timely feedback to all students (Mohsen, 2022), which may not only lead to the failure to support students' long-term learning and development but also add substantial pressure to teachers' workloads (Ferris, 2007).

With the rapid development of natural language processing (NLP) and artificial intelligence (AI) technologies, the enormous potential of AI technologies in assisting language learning has drawn widespread attention. AI feedback can provide learners with instant, reliable, and personalized feedback, alleviating the heavy burden on teachers regarding students' academic performance while improving students' learning outcomes (Escalante et al., 2023; Rad et al., 2023), effectively addressing the limitations of traditional teacher feedback. However, despite the growing interest in AI feedback, research on its effects on language learners' affective-motivational outcomes remains limited (Wiboolyasarin et al., 2024). Therefore, this study aims to synthesize existing research on AI feedback in language learning, explore its impact on the affective-motivational development of language learners, and investigate potential variables that moderate these effects.

2. AI-powered feedback in language learning

Feedback plays a critical role in language learning. Effective feedback can narrow the gap between learners' current performance and their goals by gradually enhancing their ability to complete specific tasks (Kluger & DeNisi, 1996). AI has revolutionized the forms and functions of feedback considering its capability to offer learners more immediate and personalized guidance, overcoming the spatial and temporal limitations of traditional feedback (Huang et al., 2022; Cheng & Zhang, 2024). As the most studied AI-powered feedback systems, Automated Writing Evaluation (AWE) systems are widely used in writing instruction. Research shows AWEs can significantly enhance the efficiency of feedback and lead to significant improvements in learners' writing performance (Ngo et al., 2024). Researchers attributed the pedagogical benefits of AWE tools to their capacity to offer adaptive metalinguistic explanations, enhance learner engagement and foster self-directed learning (Barrot, 2023a).

Additionally, the recent emergence of generative AI tools, such as ChatGPT and other large language models (LLMs) has introduced novel feedback experiences for language learners. Through NLP, these GenAI tools simulate authentic conversational scenarios, offering interactive feedback (Li et al., 2024). They have played a significant role in helping learners improve their speaking skills (Yuan, 2023) and vocabulary acquisition (Zhang & Huang, 2024), enhancing the effectiveness of language learning and boosting learners' confidence. Barrot (2023b) highlighted several uses of ChatGPT as a feedback tool in writing, including giving automated grades, offering quality feedback on content, organization, clarity of purpose, etc. However, research on the application of GenAI in language education is still in its infancy, and more in-depth studies are warranted to further reveal the impact of GenAI on different facets of language learning (Wu & Yu, 2024).

3. Previous review studies of AI feedback in language learning

With the proliferation of studies of AI feedback tools in language education, quite a number of review studies have been conducted to analyze the development status and research trends. In an early critical review by Stevenson & Phakiti (2014), they found that there was only limited evidence to support the positive effects of AWE feedback on writing, and they called for more studies employing different AWE systems to explore the effects of AWE feedback on the quality of writing. Recently, Fu et al. (2022) found in their review that AWE can improve students' writing skills to some extent, though less effectively than human feedback. However, students generally hold a positive attitude toward AWE, perceiving it as useful and reporting being motivated during its use. Shi and Aryadoust (2024) reviewed 83 SSCI-indexed journal articles from 1993 to 2022 and their study lent support to the favorable outcomes of automated-written feedback (AWF). They ended their analysis by advocating for further investigations into how AWF influences motivational and affective aspects, including anxiety, motivation, and self-regulation.

Apart from studies that concentrate on AWE, there are also systematic reviews that broaden their focus by synthetically analyzing how AI-driven feedback impacts language learning. Evenddy (2024) explored the role of AI in enhancing feedback mechanisms for English language learning

and concluded that AI feedback typically outperforms traditional methods in speed, accessibility, and personalization, significantly improving learning outcomes in grammar, vocabulary, and pronunciation. Nevertheless, the review also pointed out notable challenges of AI feedback, including issues of accuracy and data dependency. In a systematic review of ten empirical articles published between 2019 and 2023, Lee and Moore (2024) discovered that GenAI, with its variety of feedback formats and technological benefits, not only improves educational outcomes by making learning environments more effective and encouraging for students, but also reduces the workload of teachers.

In summary, existing review studies highlight the broad application potential of AI feedback in language education, yet there remain some notable research gaps. First, in terms of research methodology, most existing review studies are narrative reviews, lacking quantitative data-driven meta-analytic studies to comprehensively synthesize and examine the effects of AI feedback in language learning. Second, the literature is overly concentrated on AWE, with insufficient focus on emerging technologies represented by GenAI. Finally, while most studies focus on AI feedback's impacts on learning outcomes such as writing quality and grammatical accuracy, the effects of AI on learners' affective-motivational states such as anxiety, self-efficacy, and motivation, etc are under-explored.

In order to fill the gap in the research, the current meta-analysis investigates how AI feedback affects affective-motivational outcomes in language learning. Two specific research questions are intended to be addressed:

Q1: What is the overall effectiveness of AI feedback on affective-motivational outcomes in language learning?

Q2: What potential moderating variables significantly modulate the effects of AI feedback on affective-motivational outcomes in language learning?

4. Method

4.1. Study inclusion criteria

The following criteria were applied to evaluate a study's eligibility for inclusion in order to guarantee the inclusion of high-quality research that aligned with the goals of this meta-analysis:

(1)The study should examine the impact of AI feedback on affective-motivational outcomes in language learning such as motivation, self-efficacy, and anxiety, etc. (2) The study should adopt an experimental or quasi-experimental design that includes a control group without AI feedback application. (3) The study should report the quantitative data that are necessary for the calculation of effect sizes, such as means, SD, sample sizes, p-values, or F-values.

4.2. Search strategy and study selection

We conducted an electronic search for peer-reviewed journal papers published in English between Jan 2004 and Dec 2024 in the Web of Science (WOS) and ERIC databases with the following four keyword categories:

(1)AI-related keywords: AI, artificial intelligence, machine learning, deep learning, Large Language Models, LLMs, ChatGPT, chatbot, automated writing evaluation; (2)language-learning-related keywords: language learning, second language learning, foreign language learning, ESL, EFL, second language acquisition, SLA; (3)affective-motivational-factor-related keywords: willingness to communicate, WTC, anxiety, motivation, engagement, self-efficacy, self-regulation, satisfaction; (4)feedback-related keywords: feedback.

In WOS, 390 articles were retrieved, while ERIC returned 62 articles. After removing 25 duplicates, two researchers conducted an independent screening of 427 titles and abstracts. Full papers were subsequently retrieved and downloaded to assess their eligibility. Any issues or discrepancies encountered were resolved through discussion. Ultimately, 15 studies were included in the meta-analysis.

4.3. Coding scheme

Based on previous meta-analyses (Ngo et al., 2024; Zhai & Ma, 2023), we developed a coding scheme in order to investigate how various moderator variables affected the effectiveness of AI feedback: (1) educational level: university, others; (2) intervention duration: less than 10 weeks, 10 or more than 10 weeks, unspecified; (3) type of AI systems: GenAI, non-GenAI; (4) affective-motivational factor: engagement, motivation, anxiety, WTC, self-regulation, self-efficacy, enjoyment etc.; (5) feedback source: AI, AI with other sources.

Twenty percent of the studies included in the analysis were first individually coded by the two researchers as a trial. Any inconsistencies in the initial coding were then addressed and resolved. Following additional clarification of the criteria, the remaining eligible studies were coded separately by both researchers. Any disagreements that arose throughout the coding process were settled through discussion.

4.4.Calculation of effect size

In calculating the effect sizes, Comprehensive Meta-Analysis 3.0 software was used, with Hedge's g as the measure for reporting effect sizes. When calculating the overall effect size, the choice between a random or a fixed effects model depends on study heterogeneity. For high-heterogeneity studies, a random-effects model is often better (Borenstein et al., 2009). Cochran's Q test and the I^2 statistic can check study homogeneity. A random-effects model is preferred when the Q-test p-value is less than 0.05, which indicates substantial differences (Borenstein et al., 2009). Higgins et al. (2003) define I^2 values of 25%, 50%, and 75% as low, moderate, and high heterogeneity respectively.

5.Results and Discussion

5.1. Overall effect

One eligible study (Rad et al., 2023) was identified as an outlier with an exceptionally large effect size (g=44.028), and was therefore removed from the analysis (Lipsey & Wilson, 2001). According to the results based on the random-effects model ($Q = 465.571$, $p < 0.001$; $I2 = 97.208\%$),

a large overall effect size of 0.850 (p<.0.001) was found for using AI feedback (Cohen, 1988), which means AI feedback significantly improves affective-motivational outcomes in language learning.

Our results support other research findings about the impact of AI on emotional aspects (Huang et al., 2022; Wu & Yu, 2024; Lee & Moore, 2024) and show that AI feedback can generally have a favorable impact on learners' affective-motivational states. This may be explained by the alignment between the technical characteristics of AI feedback and the psychological needs of learners (Li et al., 2024). By providing learners with immediate assessments and corrective suggestions immediately after producing language output, AI feedback increases learning efficiency and boosts learners' self-efficacy (Huang et al., 2024). Additionally, AI's personalized function can offer feedback that is tailored to learners' emotional needs. Compared with teacher feedback and peer feedback, AI can help students feel less pressured to perform well, create a more comfortable learning environment, and build their confidence, all of which increase students' intrinsic motivation and engagement in studying (Neo et al., 2024). Ultimately, it forms a positive driving force for the affective-motivational outcomes in language learning.

Table 1. Overall effect size of AI feedback on language learning

					95% CI		Test of null		Heterogeneity		
	k	*g*	SE	Varianc	LL	UL	*Z*	*P*	*Q*	*p*	I^2
Fixed	1 4	0.74 7	0.04 7	0.002	0.65 4	0.83 9	15.83 6	0.00 0	465.57 1	0.00 0	97.20 8
Rando	1	0.85	0.31	0.101	0.22	1.47	2.681	0.00			

SE: standard error; CI: confidence interval; LL: lower limit; UL: upper limit.

Table 2. Effect sizes of moderator variables

Moderators			95% CI				Heterogeneity		
	K	*g*	LL	UL	*Z*-	*p*-	*Q*	*df*	*p*
Educational level							5.017	1	0.025
university	8	0.311	-	0.644	1.829	0.067			
others	6	1.620	0.524	2.715	2.898	0.004			
Intervention duration							10.765	2	0.005
⩾10 weeks	9	0.914	0.580	1.249	5.356	0.000			
＜10 weeks	3	-0.288	-	0.397	-0.823	0.410			
Unspecified	2	2.282	-	5.113	1.580	0.114			
Type of AI systems							1.909	1	0.167
GenAI	3	0.320	-	0.867	1.145	0.252			
Non-GenAI	11	0.986	0.215	1.758	2.505	0.012			
Feedback source							2.061	1	0.151
AI with other source	3	0.377	0.087	0.666	2.548	2.505			
AI feedback	11	0.974	0.212	1.736	0.011	0.012			
Affective-motivational							30.379	6	0.000
Anxiety	4	-0.106	-	1.219	-0.157	0.875			
engagement	4	0.715	0.190	1.240	2.699	0.008			
motivation	3	0.783	-	1.599	1.879	0.060			
self-efficacy	5	0.084	-	1.043	0.172	0.863			
self-regulation	3	2.018	0.142	3.894	2.108	0.035			

WTC	1	2.344	1.755	2.933	1.755	2.933
enjoyment	2	0.783	0.470	1.096	4.903	0.000

CI: confidence interval; LL: lower limit; UL: upper limit.

5.2. Educational Level

Table 2 shows effect sizes of moderator variables. We divided the educational level into two levels: university and others. Results show AI feedback produced a small and statistically insignificant effect size for university students (g= 0.311, p= 0.067) and a large and significant effect size for other educational levels (g=1.620, p=0.004). The difference of effects of AI feedback between groups reached a significant level (Q=5.017, p= 0.025).

Most previous studies that investigated the effectiveness of AI feedback on language learning achievements reported a larger effect size for university students (Ngo et al., 2024; Zhai & Ma, 2023; Wu &Yu, 2024). However, according to our research, AI feedback significantly impacts the non-university group in terms of affective-motivational effects. This might be the case because the participants in this group—which includes primary students, middle school students and others—require more affective-motivational help from AI due to their relatively weak autonomous learning abilities. Furthermore, their curiosity and reaction to the technology may be more intense. As for university students who have stronger metacognitive regulation ability and mature learning strategies, their affective-motivational states are probably less influenced by external factors such as AI feedback, thus weakening its impact.

5.3. Intervention duration

The group who received the intervention for ten weeks or longer had a large and statistically significant effect size of 0.914 (p<.0.001), while the group that received it for less than ten weeks had a negative and non-significant effect size (g=-0.288, p=0.410). Two studies which didn't report specific intervention durations were classified into unspecified group and showed a very large but insignificant effect size (g=2.282, p=0.114). The effects of AI feedback differed significantly across invention durations (Q= 10.765, p=0 .005).

Our results, which show the substantial influence of a longer intervention duration on the efficacy of AI feedback, are mostly in line with the findings of earlier studies (Ngo et al., 2024; Wu, 2024). One possible explanation might be the transformation of affective-motivational factors usually requires a relatively long time for internalization, and short-term interventions are often insufficient to have a significant impact.

5.4. Type of AI systems

For the group using GenAI, the effect size was small and insignificant (g=0.320, p=0.252), while for the non-GenAI group, the effect size was large and significant (g=0.986, p=0.012). However, there was no statistically significant difference between the effects of different AI systems (Q=1.909, p=0.167).

In contrast to Lee and Moore (2024), our study demonstrated no significant effect of generative AI feedback on affective-motivational features. Conversely, non-Gen AI tools, such as AWE, have

a greater impact. The following is an explanation for the reasons. Non-Gen AI feedback focuses more on specific learning problems such as grammar and vocabulary (Ngo et al., 2024; Chen & Zhang, 2024), and provides more targeted guidance that stimulates affective-motivational aspects more effectively. Gen AI's open-ended discussions as well as its technical limitations force learners to filter information, raising their cognitive load (Evenddy, 2024; Li et al., 2024).

5.5. Feedback source

Results indicate a large, significant effect size for using AI feedback alone (g=0.974, p= 0.012) and a small and insignificant effect size for AI feedback combined with other feedback sources (g=0.377, p=2.505). There was no statistically significant difference between the impacts of different feedback sources (Q=2.061, p=0.151).

The results show that using AI feedback alone is more effective in influencing the affective-motivational aspects compared to using AI in combination with other feedback sources. This finding is consistent with the previous research results of Zhai & Ma (2023) and Ngo et al. (2024), which reported larger effect sizes for students using AWE independently than using AWE with the help of teachers. With single-source AI feedback, learners may concentrate more on the AI's instructions without being distracted by multiple feedback sources, which might lower the possibility of cognitive conflict and increase feedback trust.

5.6. Affective-motivational factor

Very large and significant effect sizes were reported for self-regulation (g=2.018; p=0.035). Medium and significant effect sizes were found for language learning engagement (g=0.715, p= 0.008) and enjoyment (g=0.783, p< 0.001). The effects of AI feedback on reducing anxiety (g=-0.106, p= 0.875), improving motivation (g=0.783, p=0.060) and self-efficacy (g=0.084, p= 0.863) were not statistically significant. Only one study was concerned with improving WTC and the effect size was large but not significant (g=2.344, p=2.933). The differences between different affective-motivational outcomes were significant (Q=30.379, p<0.001).

Significant variation can be seen in the effects of AI feedback on different affective-motivational elements. It has a very large and significant effect on improving students' self-regulation ability. Mohebbi (2025) asserts that the ability to self-regulate is a reflection of students' active engagement in the learning process and is essential to fostering and promoting their autonomy. The personalized and timely feedback provided by AI can help learners become autonomous and self-regulated learners by developing their skills and confidence (Qiao & Zhao, 2023). The fairly substantial benefits on learners' engagement and enjoyment may be explained by the fact that AI feedback can give students a relaxed learning environment, which will increase their passion and interest in learning (Lee & Moore, 2024; Neo et al., 2024). There is no discernible impact on anxiety, self-efficacy, or motivation, which runs counter to earlier studies by Huang et al. (2024) and Wu & Yu (2024). Considering the complexity of the affective-motivational system in language learning, the results of AI feedback may also be affected by factors such as research design and learner ability differences (Wu, 2024). Regarding WTC, the evidence is weak because

there is only one study that addresses it and additional research is still required to support pertinent conclusions.

5.7. Publication bias

In this study, publication bias was evaluated using Classic Fail-safe N, and Orwin's Fail-safe N. The Classic Fail-safe N test revealed that 927 additional studies would be required to render the effect size non-significant, which is significantly higher than the 80-threshold (5*14 + 10) suggested by Rosenthal (1991). 1032 studies were needed to lower the impact size to 0.01 according to Orwin's Fail-safe N test. These findings imply that publication bias has little bearing on the study's conclusions.

6. Conclusion

Overall, our study reveals a very positive effect of using AI-powered feedback on affective-motivational outcomes in language learning. Educational level, intervention duration and affective-motivational factor are significant moderators that explain the variance in effect sizes.

6.1. Implication

Our study has important implications for both for research and language instruction. Firstly, our study provides evidence supporting the impact of AI feedback on the emotional and motivational outcomes in the language learning process, revealing the rationality of applying AI feedback to language learning. The analysis of potential moderating variables that affect feedback effectiveness motivates educators to implement more precise and reasonable teaching designs according to the characteristics and advantages of technologies. Secondly, our study uncovers the imbalance in the distribution of research on AI feedback in certain areas. There is relatively less research on how AI feedback affects emotional and motivational aspects. In the future, more relevant empirical studies can be conducted in this regard.

6.2. Limitations

The study has some limitations. Firstly, our data retrieval was only based on two databases, and the number of studies retrieved and included in the analysis was relatively small, which may have a certain impact on the research results. Secondly, due to the insufficient information provided by the authors, the exploration of other moderating variables that may affect the effectiveness of AI feedback was inadequate. Future studies could start from a broader range of variables to provide a more detailed and comprehensive understanding.

Acknowledgements

This work was supported by the the Humanities and Social Sciences Research Projects, Ministry of Education, China, under Grant [21YJAZH010]; *2023 Key Projects of Beijing Higher Education Society.*

References

Barrot, J. S. (2023a). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, *36*(4), 584–607.https://doi.org/10.1080/09588221.2021.1936071

Barrot, J. S. (2023b). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing,* 57, 100745. https://doi.org/10.1016/j.asw.2023.100745

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester: Wiley. DOI: 10.1002/9780470743386

Cheng, X., & Zhang, L. J. (2024). Examining second language (L2) learners' engagement with AWE-teacher integrated feedback in a technology-empowered context. *The Asia-Pacific Education Researcher*, *33*(4), 1023–1035. https://doi.org/10.1007/s40299-024-00877-8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, *20*(1), 57. https://doi.org/10.1186/s41239-023-00425-2

Evenddy, S. S. (2024). Investigating AI's Automated Feedback in English Language Learning. *Foreign Language Instruction Probe*, *3*(1), 76-87. https://doi.org/10.54213/flip.v3i1.401

Ferris, D. (2007). Preparing teachers to respond to student writing. *Journal of Second Language Writing*, *16*(3), 165–193. https://doi.org/10.1016/j.jslw.2007.07.003

Fu, Q. K., Zou, D., Xie, H., & Cheng, G. (2022). A review of AWE feedback: types, learning outcomes, and implications. *Computer Assisted Language Learning*, *0*(0), 1–43. https://doi.org/10.1080/09588221.2022.2033787

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *bmj*, *327* (7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

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.

Huang, X., Xu, W., Li, F., & Yu, Z. (2024). A meta-analysis of effects of automated writing evaluation on anxiety, motivation, and second language writing skills. *The Asia-Pacific Education Researcher*, *33*(4), 957-976. https://doi.org/10.1007/s40299-024-00865-y

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83-101. doi:10.1017/S0261444806003399

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Lee, S.S. & Moore, R.L. (2024) Harnessing Generative AI (GenAI) for automated feedback in higher education: A systematic review. *Online Learning, 28*(3), 82-104. DOI: 10.24059/olj.v28i3.4593

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage

Mohebbi, A. (2025). Enabling learner independence and selfregulation in language education using AI tools: a systematic review. *Cogent Education, 12*:1, 2433814, DOI: 10.1080/2331186X.2024.2433814

Mohsen, M. (2022). Computer-Mediated Corrective Feedback to Improve L2 Writing Skills: A Meta-Analysis. *Journal of Educational Computing Research*, *0*(0), 1–24. https://doi.org/10.1177/07356331211064066

Ngo, T., Chen, H., & Lai, K. (2024). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive learning environments*, *32*(2), 727-744. https://doi.org/10.1080/10494820.2022.2096642

*Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, *14*, 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Rad, H. S., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of wordtune application. *Interactive Learning Environments*, 1-22. https://doi.org/10.1080/10494820.2023.2208170

Rosenthal, R. (1991). Meta-analysis: A review. *Psychosomatic Medicine, 53*(3), 247–271.

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, *36*, 187–209. https://doi.org/10.1017/S0958344023000265

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, *19*, 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Wiboolyasarin, W., Wiboolyasarin, K., Suwanwihok, K., Jinowat, N., & Muenjanchoey, R. (2024). Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments. *Computers and Education: Artificial Intelligence*, *6*, 100228. https://doi.org/https://doi.org/10.1016/j.caeai.2024.100228

Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. https://doi.org/10.1111/bjet.13334

Wu, X.-Y. (2024). Artificial Intelligence in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. *System*, 103498. https://doi.org/10.1016/j.system.2024.103498

Xue, Y. (2024). Towards automated writing evaluation: A comprehensive review with bibliometric, scientometric, and meta-analytic approaches. In *Education and Information Technologies* (Issue March). Springer US. https://doi.org/10.1007/s10639-024-12596-0

*Yuan, Y. (2023). An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interactive Learning Environments,* 1-17. https://doi.org/10.1080/10494820.2023.2282112

Zhai, N., & Ma, X. (2023). The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *Journal of Educational Computing Research*, *61*(4), 875-900. https://doi.org/10.1177/07356331221127300

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon, 10*(3). https://doi.org/10.1016/j.heliyon.2024.e25370

人智协同视角下 AIGC 对协作学习中学生批判性思维能力的影响研究

A Study on the Impact of AIGC on Students' Critical Thinking Abilities in Collaborative Learning under the Perspective of Human-Intelligence Collaboration

张锦[1*]，张琪[2]，陈玉杰[3]
[123]淮北师范大学教育学院
1575150743@qq.com

【摘要】本研究旨在探究基于生成式人工智能（AIGC）的人智协同模式对学生协作学习中批判性思维能力的影响。通过混合研究方法，以师范院校教育技术专业学生为研究对象，结合认知网络分析发现：在协作学习中引入人工智能生成内容（AIGC）显著提升了学生的高阶批判性思维能力。结果表明，实验组在认知网络发展与知识建构上优于对照组，AIGC 能促进更深层次的批判性思维。本研究验证了 AIGC 支持的协作学习模式对学生批判性思维的影响及其具体表现，为 AI 教育工具在协作学习中的应用提供了实证依据。
【关键词】生成式人工智能、人智协同模式、批判性思维、协作学习

***Abstract*:** *This study aims to explore the impact of the AIGC-based human-intelligence collaboration model on students' critical thinking skills in collaborative learning. By diving into a cognitive network analysis, it became clear that using Artificial Intelligence Generated Content(AIGC) in collaborative learning settings dramatically boosted students' critical thinking abilities, notably in analyzing, assessing, and innovating. The statistical analyses revealed that students in the experimental group excelled in cognitive network development and knowledge acquisition compared to their peers. This suggests that the AIGC setting nurtures more profound critical thinking and knowledge building. The purpose of this experiment is to verify the impact of AIGC-supported collaborative learning mode on students' critical thinking and its specific performance, and to provide an empirical basis for the application and participation of AI education tools in collaborative learning. But the issues like small sample sizes and short experimental periods call for more in-depth investigation in upcoming research endeavors.*
Keywords: Artificial Intelligence Generated Content, human-intelligence collaborative model, critical thinking, collaborative learning

1.前言

生成式人工智能是当前信息技术发展为教育领域带来的众多挑战和机遇之一。教育领域一直以来都在积极地探索提升教学质量与培养学生能力成效的新路径，而 AIGC（Artificial Intelligence Generated Content）就为此注入新活力。已有研究表明其能够提高教师利用数字化教育资源改善困境的能力（万力勇等, 2023）。协作学习作为现代教育核心策略之一，长期以来被视作培育学生高阶思维的关键支撑，同时也对于其终身学习和适应社会变革具有重要意义。早期 Gokhale（1995）的经典研究证实了小组合作对学生批判性思维的强化效能，后续研究也持续强化了这一认知，彰显其在多学科及不同学习模式下的积极意义。（罗宇晨和郑燕林，2024）本研究聚焦于基于 AIGC 的人智协同模式在协作学习中的作用，旨在通过

实证研究，探讨 AIGC 如何影响学生的批判性思维发展，分析 AIGC 辅助下的协作学习对学生批判性思维能力的影响，以期为教育实践提供新的视角和策略。

2.文献综述

2.1. 基于 AIGC 的人智协同模式

人智协同模式是指根据人类与智能交互的目的特征、结构关系等进行分类研究与相关探索。在教育领域，研究者们将生成式人工智能引入智能教育，使其在教育领域发挥重要作用。万力勇等（2023）研究发现，在教育领域生成式人工智能工具对当前教师端数字化教育资源开发中出现的体谅不足、效率不高以及质量不佳等问题有着积极影响。研究者们从多维度探讨了 AIGC 对教学要素的影响，并提出了基于 AIGC 的人智协同教学模式(韩锡斌, 2024)。张进澳等（2024）指出基于生成式人工智能的人智协同研究尚处于初级阶段，整体研究框架尚不成熟。多数研究聚焦于宏观的人智协同模式分类，对于 AIGC 在特定教育场景（如协作学习）中的细致交互过程与协同机制缺乏深度剖析。本研究将通过实证研究，探讨基于 AIGC 的人智协同模式下学生进行协作学习时批判性思维的发展变化。

2.2. 协作学习与批判性思维

协作学习作为一种有效的教学策略，能够支持协同问题的解决与知识建构，作为教育数字化转型背景下培育高阶思维的常用方式，被广泛的应用于教学实践之中。早在 Gokhale（1995）的研究中就已表明，小组合作中学生可以共同分析问题评估信息并提出解决方案，有效增强了批判性思维能力。Warsah 等人（2021）探讨了协作学习对批判性思维技能的影响，发现协作学习在不同学科中都对学生的批判性思维有积极且显著的影响。随着信息技术及人工智能的发展，罗宇晨和郑燕林（2024）通过实证研究发现，混合式协作学习可以有效促进批判性思维的发展并使学生在协作学习中的行为有序灵活地切换。本研究将进一步聚焦 AIGC 的参与下的协作学习场景展开实证研究。

2.3. AIGC 与批判性思维

生成式人工智能技术在批判性思维的培养方面显示出巨大潜力。生成式 AI 工具已被用于批判性思维测评，显著提升了学生在“利弊分析”“图文情境”及“对立观点”等主题的议论文写作能力。（Warsah et al., 2021）但部分研究也指出 AIGC 的负面影响：在文学鉴赏课程中，部分学生直接复制 AIGC 生成的分析，抑制了其批判性分析与原创观点构建能力。这些矛盾表明 AIGC 的应用需谨慎设计以避免依赖性，同时促进独立思考。本研究将探讨 AIGC 在协作学习中的角色，尤其关注其如何促进学生批判性思维。

3.理论框架

3.1. 社会建构主义理论

社会建构主义理论认为学习者需要与他人在社会活动中进行互动与协作，以此进行知识的构建。在本研究里，学生在 AIGC 的环境下与 AIGC 及组员同伴进行协作交流。AIGC 可作为伙伴参与讨论，在协作学习中提供不同思路与想法。本研究就关注于 AIGC 如何促进学生的互动和知识的构建。

3.2. 人机协同学习理论

人机协同理论为理解人智协同提供了重要的视角。在人智协同中，人类智能与人工智能之间的协同交互，本质上是人机协同理念的延伸和扩展，随着人机融合的深化，人类与 AI 智力的协同日益显著。这种协同强调技能动态交换，将人类与 AI 优势结合为和谐整体。

（张进澳，2024）数据输入、处理、输出与反馈四环节被视为评估 AIGC 融入协作学习的基础，强调了人机在学习过程中的协同作用，为研究 AIGC 在协作学习中的应用提供了活动框架的理论支撑。（王一岩等，2024）

3.3. 协同知识建构螺旋模型（SMCKI）

协同知识建构螺旋模型由 Chen 等(2019)在 CSCL（Computer-Supported Collaborative Learning）环境下提出，包括个人头脑风暴、团队协作、组间互评、组内优化与个人观点完善这五个阶段。学生在各个阶段中获得不同的反馈，不断建构知识并在协作学习的动态循环与螺旋上升中促进知识深化与重构。SMCKI 模型的五个阶段为人智协同模式下的协作学习提供了具体的流程指导，学生可更有效进行协作，促进协同知识建构及批判性思维的发展。

3.4. 基于 AIGC 的人智协同模式下协作学习模型建构

上述理论基础为本研究提供了坚实的理论支撑，使我们能够深入探讨 AIGC 环境下的人智协同模式如何影响学生的批判性思维和知识建构过程。在图 1 中，我们在人智协作模式和 AIGC 积极参与小组讨论的情况下，基于这些理论构建了一个适用于本研究的协作学习模式。与传统方法不同，AIGC 作为一个积极的协作伙伴，通过实时反馈与多视角信息深化认知投入。

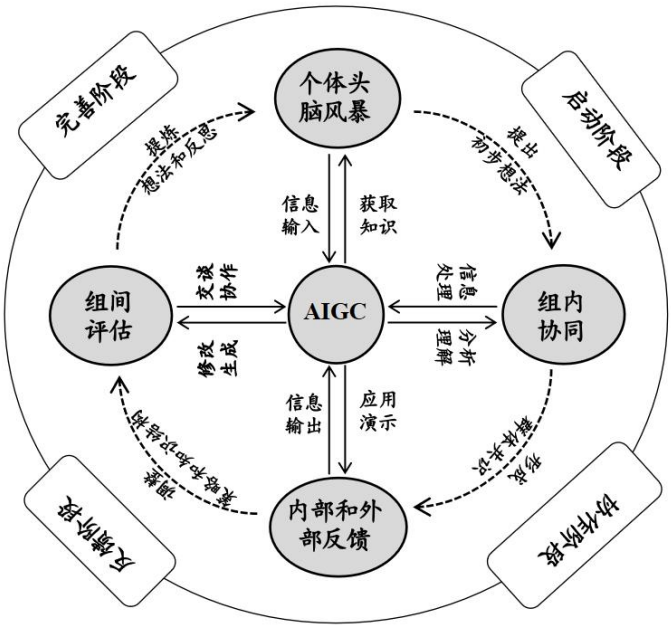


图 1 基于 AIGC 的人智协同模式下协作学习模式

4.研究方法

4.1. 研究对象

本研究选取某师范学院 30 名教育技术专业大三学生为参与者。其未来职业需求与教育技术应用高度相关。在实验前，在指导员的辅助下对 30 名研究对象进行调查法中访谈法的相关知识培训。将他们分为每 5 人一组，随机分配为 3 个小组为实验组，余下 3 个小组为对照组。要求被试协作设计访谈提纲。为确保实验过程中顺利应用 AIGC 辅助协作学习，在实验前对实验组进行与 AIGC 对话的相关操作方法与基础口令培训。此外，为了减少先验知识水平的影响，本研究向每组提供给定的主题“生成式人工智能在学科学习中的应用与挑战”。

4.2. 研究工具

本研究参考彭美慈等人（2004）研发的批判性思维能力量表（CTDI-CV），问卷信效度良好（Cronbach's Alpha=0.724，KMO=0.781），设计批判性思维能力调查问卷，共包含 4 个维度 16 个指标，在实验前和实验结束后分别对被试进行前测与后测。同时选用由字节跳动公司开发的 AIGC 软件"豆包"参与实验组的协作学习并收集人智交流数据。本研究参考 Elizabeth Murphy 提出的批判性思维编码体系，将批判性思维划分为辨识、理解、分析、评价和创造五个维度，设计批判性思维编码表以便对学生批判性思维进行认知网络分析。为了确保编码有效，研究团队对交互内容进行独立编码，编码结果一致性良好（kappa=0.79，$p<0.05$），有分歧之处，在协商后达成一致。

表 1 批判性思维编码表

编码类别	编码	描述	举例
辨识（Recognize）	RE	辨认或识别存在的话题、矛盾、问题等	"设计访谈问题需考虑哪些因素？"
理解（Understand）	UN	探究相关的论据、知识、调研、信息和观点	"如何理解访谈问题与学科学习的关联？"
分析（Analyze）	AN	深入寻找有关问题，组织已知信息，鉴别未知信息，剖析问题、矛盾及其基本的逻辑组成	"这些访谈问题有哪些潜在的偏见或盲点？"
评价（Evaluate）	EV	批评或评价相关的信息、知识、观点	"这些建议是否能够提高我们访谈提纲的质量？"
创造（Create）	CR	产生新知识、观点或方法，并加以实施	"我们能否结合这些观点创造一个新的访谈提纲？"

4.3. 研究设计

在实验之前对所有被试进行批判性思维的前测，在实验结束后进行后测。要求所有小组以"生成式人工智能在学科学习中的应用与挑战"为主题，分四个阶段进行协作学习设计访谈提纲（要求包含访谈问卷）。实验组进行基于 AIGC 的人智协同模式下的协作学习。阶段 1：小组成员通过与 AIGC 的互动获取访谈主题相关信息和知识，并进行个体头脑风暴，独立思考提出访谈提纲的初步想法和问题；阶段 2：小组成员协同 AIGC 根据个体头脑风暴的结果进行讨论、协同合作，整合各自的想法，形成小组的共识和访谈提纲 A；阶段 3：小组将访谈提纲 A 应用模拟，同时从 AIGC 和同伴处获得即时反馈，调整学习策略和知识结构。小组成员与 AIGC 进行对话，对提纲和问卷进行修改，形成访谈提纲 B；阶段 4：组间互评：各小组使用 AIGC 进行组间互评，根据其他小组的评级和反馈，再次对访谈提纲进行修改调整，形成访谈提纲 C。个体撰写反思报告，巩固学习成果。对照组：进行传统协作学习，与实验组同步进行同样步骤的四个阶段的协作学习与反思讨论，不使用 AIGC。每个阶段分别形成访谈提纲 D、访谈提纲 E 与访谈提纲 F。

实验结束后要求学生认真思考、回忆实验过程中的体验，结合反思报告再次填写批判性思维能力量表（CTDI-CV）进行后测。

4.4. 数据收集

在本次研究中，收集了实验过程中学生协作交流的文本数据，并且在实验组收集与AIGC的对话文本，并记录了每个阶段的访谈提纲。在实验结束后对文本数据编码并导入WebENA（认知网络分析工具）中进行学生批判性思维认知网络分析，对实验组与对照组生成的访谈提纲由专家进行评分比较分析，来探索在AIGC的辅助下，学生在协作学习中知识建构是否更加深入和有效。此外，通过收集所有被试的批判性思维能力调查问卷前后测数据分析学生在试验前后批判性思维是否有变化且发生怎么样的变化来辅助验证实验结果。

5.结果与分析

5.1. 批判性思维认知网络分析

为了分析有AIGC参与的协作学习是否对学生的批判性思维产生影响以及产生怎样影响，本研究收集了所有被试小组在学习过程中的交流文本数据进行认知网络分析，通过认知网络分析工具Web ENA对编码后的文本数据进行分析，基于编码互动的频率，计算了批判性思维各维度（如识别、分析、创造）间的共现关系,生成了认知网络图。图2展示了实验组A与对照组B组的批判性思维认知网络以及6个小组在各阶段的批判性思维认知网络图。其中节点的大小反应批判性思维认知元素的强弱，线条表示两个连接的节点之间的共线关系，线条越粗，表示两节点之间共现频率越高。

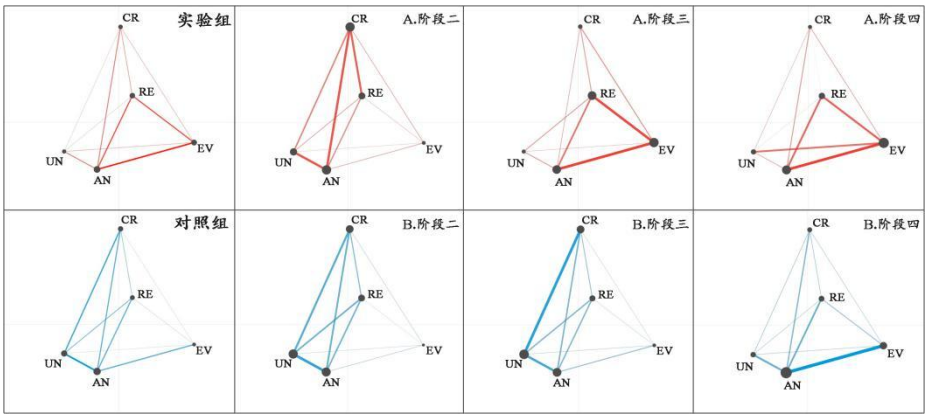


图2 批判性思维认知网络分析图

实验组学生的批判性思维主要集中在“分析”（AN）、“评价”（EV）与“创造”（CR）层面，且这三种深层次思维之间存在较强程度的关联。而对照组学生的批判性思维主要集中在“理解”（UN）、“辨识”(RE）与“分析”（AN）上，且这些较浅层次思维间的关联程度强于较深层次思维之间的关联。实验组A使用了AIGC工具，可能提供了额外的信息与视角，促进了学生新观点的创造能力以及批判性思维各维度的整合，因此实验组进行更深入的分析与评价。

第二阶段中，实验组呈现出识别（RE）与分析（AN）、理解（UN）与创造（CR）间的强关联，而对照组的相关性较弱且分布更均匀，理解（UN）与识别（RE）间关联最强（见图2）。这种差异可能源于实验组在第一阶段通过AIGC获取的全面信息，增强了他们协作识别、分析和创造知识的能力。到第三、四阶段，在AIGC多样化视角与即时反馈的支持下，实验组在创造力（CR）等深层次维度有显著进展。对照组的进步仍集中于理解（UN）和识别（RE）等低层级维度，在高级任务中提升有限。尽管两组在多数维度均有提升，但实验组展现出更优越的批判性思维整合能力，突显了AIGC对深化认知参与的作用。

5.2. 批判性思维能力分析

为了评估被试的批判性思维能力，本研究采用中文版批判性思维倾向量表(CTDI-CV)，对被试进行了前测和后测。（彭美慈等，2024）结果显示实验组与对照组前测并无显著差异，且后测均与前测有显著差异，每组后测平均得分皆高于前测。对学生的批判性思维能力后测数据进行非参数检验，结果如表 2 所示，实验组学生的批判性思维能力显著高于对照班学生（p＜0.05）。同时个体反思报告也表明小组学生能够基于自身的认知能力，对 AIGC 工具生成的各种观点进行初步辨别并自主查阅其他相关资料，在经过小组讨论后，选取并运用正确、合适的知识。这说明学生能够批判性地选择接受 AIGC 工具所生成的内容。

表 2 批判性思维能力的非参数检验结果

维度	组别	N	Mean	U	P
批判性思维能力	实验组	15	18.89	50.5	0.017*
	对照组	15	11.37		

*p <0 .05，**p < 0.01

5.3. 知识掌握水平分析

为了进一步评估基于 AIGC 的人智协同模式在协作学习中对学生知识掌握水平的影响，两位精通教育研究方法的专家对实验组和对照组学生的最终访谈提纲进行了盲评。评分标准涵盖了内容的相关性、深度、创新性以及提纲的规范程度等方面。取两位专家评分的平均值作为各组最终小组成果成绩。对实验组 A 与对照组 B 两组成绩进行独立样本 t 检验，由表 3 可得知，实验组 A 的平均得分为 87.17±2.03，而对照组 B 的平均得分为 80.33±2.57，实验组与对照组之间的差异具有统计学意义（t= 3.62, p = 0.022），表明实验组在知识掌握水平上显著优于对照组。实验组 A 的学习成果较对照组 B 效果好。实验组学生在 AIGC 辅助下展现出更高的知识掌握水平。这可能归因于几点原因:首先 AIGC 工具的互动性可能促进了学生对知识的深入理解和应用。其次协作学习环境增强了学生之间的知识共享和讨论，从而提高了学习效果。最后 AIGC 可能为学生提供了个性化的学习路径，帮助他们更有效地掌握访谈法相关知识并促进他们不断地反思、修改与完善。

表 3 小组协作学习成果独立样本 t 检验结果

维度	组别	M±SD	t	P
小组成果	实验组	87.17±2.03	3.62	0.022*
	对照组	80.33±2.57		

*p <0 .05，**p < 0.01

6.结论

本研究主要聚焦于基于 AIGC 的人智协同模式在协作学习中的作用，探讨了在此模式下的协作学习如何影响学生批判性思维的发展。认知网络分析结果表明了该模式能够有效提升学生的批判性思维能力，尤其实在“分析”、“评价”和“创造”等深层次的思维。对于批判性思维能力问卷的数据分析也辅助验证了认知网络分析的结果。对学生的学习成果的分析结果验证了协作学习能有效促进学生的批判性思维发展，而基于 AIGC 的人智协同模式可以提升这类积极影响。在基于 AIGC 的人智协同模式下学生在协作学习中的知识建构过程将更加深入和有效。尽管本研究取得了一定的成果，但也存在一些局限性，如样本量有限和实验周期较短，可能无法完全捕捉到长期影响。随着 AIGC 技术的不断进步，未来的研究可以探索新兴 AIGC 工具在不同教学模式的应用，以及如何优化人智协同模式以最大化其教育效益。

参考文献

万力勇,杜静 & 熊若欣.(2023).人机共创：基于 AIGC 的数字化教育资源开发新范式.现代远程教育研究(05),12-21.

王一岩,刘淇 & 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(01),65-72.doi:10.13966/j.cnki.kfjyyj.2024.01.007.

张进澳, 陈亮, & 陈红丽. (2024). 生成式人工智能环境下人智协同的前沿分析与机会挖掘——基于扎根理论的综述研究. 数字图书馆论坛(06), 23-32.

冷静,卢弘焕 & 代琳.(2024).生成式人工智能赋能批判性思维测评——基于 ChatGPT 的应用实验.现代远程教育研究(06),102-111.

罗宇晨 & 郑燕林.(2024).混合式协作学习促进批判性思维发展实证研究.开放教育研究(02),109-119.doi:10.13966/j.cnki.kfjyyj.2024.02.012.

彭美慈,汪国成,陈基乐,陈满辉,白洪海,李守国,李继平,蔡芸芳,王君俏,殷磊.(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(09),7-10.

Chen, X., Yahaya, W. A. J. W., Huang, Z., & Zhu, X. (2024, November). The Impact of AIGC on Critical Thinking Disposition and Motivation Levels in Scriptwriting Courses. In 4th International Conference on Internet, Education and Information Technology (IEIT 2024) (pp. 193-200). Atlantis Press.

Gokhale, A. A. (1995). Collaborative learning enhances critical thinking.

Murphy, E. (2004). An instrument to support thinking critically about critical thinking in online asynchronous discussions. Australasian Journal of Educational Technology, 20(3).

Warsah, I., Morganna, R., Uyun, M., Afandi, M., & Hamengkubuwono, H. (2021). The impact of collaborative learning on learners' critical thinking skills. International Journal of Instruction, 14(2), 443-460.

基于 ICAP 框架：生成式人工智能对职前教师教学设计能力提升的实证研究

Based on the ICAP Framework: An Empirical Study of Generative Artificial Intelligence's Enhancement of Preservice Teachers' Instructional Design Ability

陈玉杰[1*]，张琪[2]，张锦[3]
[123]淮北师范大学教育学院
*2465651991@qq.com

【摘要】生成式人工智能在教育领域的应用日益广泛，尤其是在提升职前教师教学设计能力方面展现出巨大潜力。基于 ICAP 框架，利用生成式人工智能工具文心一言进行实证研究。研究对象是 32 名教育技术学专业的职前教师，他们分别提交了利用生成式人工智能辅助前后的小学信息科技基础教学设计方案。实验结束后采用问卷调查的方式获取职前教师使用生成式人工智能的反馈，对收集的数据进行配对 T 检验和认知网络分析。结果显示，生成式人工智能辅助显著提高了教学设计方案的质量，增强了职前教师的教学设计能力，且不同绩效水平的职前教师在使用生成式人工智能辅助时表现出不同的认知网络特征。

【关键词】生成式人工智能、职前教师、ICAP 框架、教学设计、认知网络分析

***Abstract*:** *Generative Artificial Intelligence (AI) has seen increasing application in the field of education, particularly in enhancing pre-service teachers' instructional design abilities, demonstrating significant potential. Drawing on the ICAP framework theory, this paper conducts an empirical study using the generative AI tool Wenxin Yiyan. This study involved 32 pre-service teachers majoring in Educational Technology, each submitting an instructional design for basic IT instruction in elementary schools, both before and after the application of generative AI. After the experiment, a questionnaire survey was conducted to obtain feedback from pre-service teachers on their use of generative AI. The collected data were analyzed using paired T-tests and cognitive network analysis. The results showed that generative AI assistance significantly improved the quality of teaching design plans and enhanced the teaching design capabilities of pre-service teachers. Moreover, pre-service teachers of different performance levels demonstrated distinct cognitive network characteristics when using generative AI assistance.*

Keywords: generative AI, pre-service teachers, ICAP framework, instructional design, cognitive network analysis

1.前言

近年来，人工智能技术，特别是以生成预训练变换器为代表的生成式人工智能在教育领域的应用日益广泛（杨蕾，2024）。根据联合国教科文组织的定义，生成式人工智能是一种根据对话提示自动生成响应内容的技术，因其强大的内容生成能力，为教师提供更加智能化和多样化的教学手段（MIAOFC & HOLMES W, 2023）。相关研究基于人机协同教育理论，强调了技术与教育的深度融合，为理解人工智能在教育中的应用奠定理论基础（方海光等人，2022）。在此基础上，后续有学者利用生成式人工智能分析了大学生编程学习行为的变化（孙丹等人，2024）。也有聚焦人智协同写作对话的过程与模式，揭示人智交互的对话行为与学习发生机制的研究（汪靖等人，2024）。

由 Chi 和 Wylie 提出的 ICAP 框架是一个四级模型，包含了四种不同类型的学习活动：交互式学习、建构式学习、主动式学习和被动式学习，为研究学习者的学习过程提供了多维度的分析视角（Chi, H. T. M. & Wylie, 2014）。在积极学习的背景下，该理论被证明了能够有效提升学习者的深层次认知参与（张馨元 & 张民选，2024）。同时，ICAP 学习方式分类学的效度还通过循证方式得到了证实（盛群力 & 丁旭，2018）。在教学设计能力提升方面，ICAP 框架同样展现出其独特的价值。运用 ICAP 学习方式分类学，进行提升学习参与度的教学实践，可以为职前教师的教学设计提供实践指导（王芳，2022）。高莉则进一步探索了基于 ICAP 的以读促学教学模式，证实了 ICAP 框架在教学设计中的应用价值（高莉，2024）。

然而，现有研究主要集中于传统教学方法下的 ICAP 框架应用，而针对生成式人工智能如何结合 ICAP 框架提升职前教师教学设计能力的研究尚未深入展开。因此，本文旨在通过实证研究，深入探索生成式人工智能如何基于 ICAP 框架提升职前教师的教学设计能力，以及不同绩效水平职前教师在使用生成式人工智能时认知网络的差异。不仅扩展了 ICAP 理论的应用范围，还为职前教师的教学设计培训提供科学指导。

2.研究设计

2.1. 研究对象

本次研究以 32 名教育技术学专业的职前教师（男 20 名，女 12 名）为研究对象，他们有相同的师范教育背景，具备基本的教学设计理论知识和能力，在本次实验之前都没有使用过生成式人工智能做教学设计。在实验前对其进行生成式人工智能工具应用的操作培训，本实验主要使用文心一言，可以保证实验过程的顺利实施。根据第二次教学设计方案的分数均值和标准差，将成绩在 23 分及以上的职前教师分为高绩效组（12 人），将成绩在 21 分到 22 分之间的职前教师分为中绩效组（12 人），将成绩在 20 分及以下的职前教师分为低绩效组（8 人）。

2.2. 实验流程

在研究初始阶段，需要职前教师提交一份完整的教学设计方案，要求仅利用传统资源或者上网自行检索完成。在研究的中期阶段，给职前教师进行培训和任务发布，包括根据主题设计培训活动，以及对生成式人工智能工具的使用方法进行指导和演示，此阶段持续五周。培训后，职前教师要求与生成式人工智能工具文心一言进行协作，对之前提交的教学设计方案进行修正完善，提交改进版的教学设计方案以及人机对话全部过程内容。在研究后期阶段，职前教师要完成一份“职前教师使用生成式人工智能辅助教学设计反馈”问卷。图 1 展示了职前教师在利用生成式人工智能辅助教学设计过程中，与 ICAP 框架相结合的具体流程。

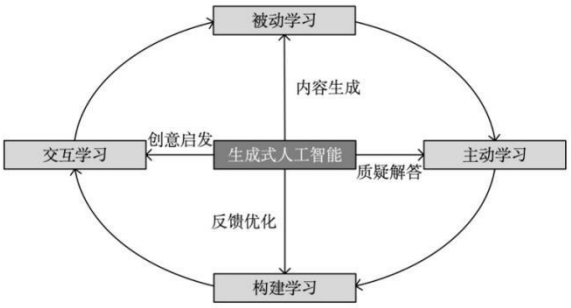


图 1 生成式人工智能与 ICAP 框架的结合

2.3. 研究工具

2.3.1. 教学设计方案评价表

本研究参考程俊等人（程俊等人，2024）开发的教学设计评价量规，设计了小学信息科技课教学设计评分标准表，见表 1。教学设计方案评分标准包括六个维度，分别是教学目标、教学内容、教学方法、教学过程、教学评价、教学创新，每个维度最高 4 分，满分 24 分。为确保客观性，由两位教育技术领域专家背对背分别对 64 份方案进行打分，一致性信度较好（Kappa=0.791）。

表 1 小学信息科技课教学设计评分标准表

评分维度	4 分	3 分	2 分	1 分
教学目标	目标明确，与课程标准和学生需求高度吻合	目标明确，与课程标准和学生需求较吻合	目标基本明确，与课程标准和学生需求部分吻合	目标不明确，与课程标准和学生需求不吻合
教学内容	内容全面，与课程目标高度相关	内容较全面，与课程目标较相关	内容基本覆盖，与课程目标部分相关	内容不全面，与课程目标不相关
教学方法	方法创新，有效促进学生发展，符合教学特点	方法有效，促进学生发展，与教学特点较符合	方法基本有效，对学生有一定帮助，基本符合教学特点	方法效果有限，对学生帮助不大，不符合教学特点
教学过程	过程流畅，环节清晰，时间分配合理，实践性很强	过程较流畅，环节较清晰，时间分配较合理，实践性较好	过程基本流畅，环节基本清晰，时间分配基本合理，实践性一般	过程不流畅，环节不清晰，时间分配不合理，缺乏实践性
教学评价	评价方式全面，反馈及时，与评价标准一致	评价方式多样，反馈及时，与评价标准较一致	评价方式基本全面，反馈一般，基本符合评价标准	评价方式单一，反馈不及时，不符合评价标准
教学创新	创新性强，与教学创新标准高度一致	创新性较好，较符合教学创新标准	创新性一般，基本符合教学创新标准	创新性弱，不符合教学创新标准

2.3.2. 交互内容编码表

基于 ICAP 框架理论制作编码表，如表 2 所示，共有 5 个一级编码，9 个二级编码。从职前教师的无关话语、被动话语、主动话语、建构话语、交互话语 5 个方面作为一级编码分析人机对话内容，对比在使用生成式人工智能辅助教学设计前后职前教师学习方式的变化和认知参与程度是否加深。

表 2 人智对话交互编码表

一级编码	编码序号	二级编码	编码序号	示例	指向水平
无关话语	IR	非知识性	IR1	周末有什么计划？	/
		教师评估	IR2	你觉得这个教学设计怎么样？	
被动话语	P	被动参与	P1	按照这个模版填写教学设计	被动学习
主动话语	A	浅层问题	A1	教学目标是否明确？	主动学习

		主动参与	A2	你认为这个活动如何帮助学生理解概念？	
建构话语	C	深层问题	C1	教学设计如何与学生的先前知识相联系？	建构学习
交互话语	I	讨论问题	I1	让我们讨论一下这个教学设计的有效性。	交互学习
		交互参与	I2	你对这个教学设计有什么建议吗？	

2.3.3. 调查问卷设计

在教学设计专家指导下编制调查问卷，问卷由人口学变量和问题构成，旨在调查职前教师使用生成式人工智能辅助教学设计后的感受和反馈，从而评估研究的有效性，并在下次实验时进行相应的改进。

3.研究结果

3.1. 教学设计方案的质量分析

采用配对样本 t 检验对 32 名职前教师的 64 份教学设计方案分数进行分析。由表 3 可知，Cohen's d 的值是 1.191，远大于 0.8，这表明生成式人工智能辅助对职前教师教学设计分数的影响是非常大的。且生成式人工智能辅助前后分数之间的相关性为 0.718，这表明两组分数之间存在显著的正相关。换句话说，分数较高的职前教师在使用生成式人工智能辅助后也倾向于获得较高的分数。因此，职前教师与生成式人工智能辅助后修改的教学设计方案质量（M = 21.03 , SD = 1.470）明显高于第一次的教学设计方案质量（M = 18.53 , SD = 1.665），t=-11.871，p = .000 < 0.001，即生成式人工智能辅助能够显著提高职前教师的教学设计方案质量。

此外，对高中低绩效的 3 组数据进行单因素方差分析。结果表明，生成式人工智能的辅助对不同绩效水平的职前教师产生了显著的影响，尤其是对高绩效组和低绩效组的职前教师影响最大。

表 3 配对样本 t 检验结果

	M	*SD*	*t*	*df*	*p*	*Cohen's d*
AI 前	18.53	1.665				
AI 后	21.03	1.470				
配对 T 检验	-2.643***	1.384	-11.871	31	0.001	1.191

3.2. 不同绩效职前教师的认知网络分析

两位编码员首先对编码框架进行了分析，然后对所形成的 382 条编码单元进行人工编码，并进行一致性比较（Kappa＝0.86），最后对差异部分进行讨论协商，达成了一致认识。将所有的数据导入在线认知网络分析工具（webENA）进行数据分析，得到高中低绩效组职前教师的认知网络结构图，如图 2 所示。

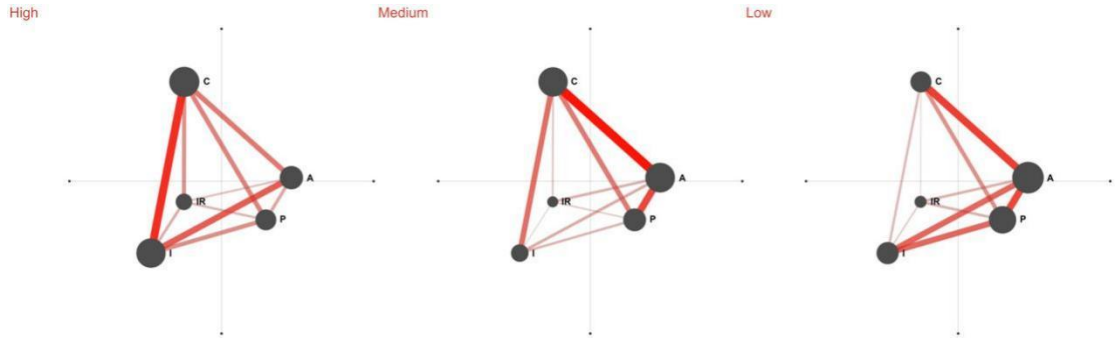


图 2 高中低绩效职前教师的认知网络分布

每个认知网络结构图包含 5 个节点，分别对应 ICAP 理论框架中的 4 种交互类型和一个表示与生成式人工智能协作交互过程中职前教师呈现与教学设计不相关的无关话语，节点的大小反映出交互类型出现的频次，节点间的连线代表交互类型共现的相对频次，连线越粗，表明两种交互类型在教学设计过程中同时出现的次数越多。整体来看，三组职前教师在生成式人工智能辅助中表现出的话语类型主要在交互话语和建构话语（I / C）之间建立连接，以及建构话语和主动话语（C / A）之间建立连接，而在被动话语（P）和无关话语（IR）上的连接较为薄弱。具体来说，高绩效组相较于中低绩效组在交互话语和建构话语（I / C）上表现出了更强的连接，中绩效组相较于低绩效组在建构话语和主动话语（C / A）上表现出了更强的连接。

此外，本研究还通过独立样本 t 检验，发现高绩效组和中绩效组在 Y 维度上存在显著性差异(m 高绩效组=-0.12，m 中绩效组=0.05；$p = 0.00 < 0.05$，t=-3.65）；高绩效组和低绩效组在 Y 维度上存在显著性差异(m 高绩效组=-0.12，m 低绩效组=0.12；$p = 0.00 < 0.05$，t=-5.25）。据此，我们可以推测，高绩效组的职前教师在使用生成式人工智能辅助教学设计时，比中低绩效组可能展现出更深层次的认知参与和高阶思维技能。

3.3. 职前教师使用生成式人工智能反馈问卷分析

本研究通过后测问卷调查评估了生成式人工智能辅助对职前教师教学设计能力的影响。结果显示，大多数职前教师对生成式人工智能辅助持积极态度，认为它提高了教学设计方案的质量和教学设计能力，并表示愿意在未来的教学中继续使用生成式人工智能工具，这与配对样本 t 检验的结果相吻合。但是也有部分职前教师指出存在功能和技术上的限制，以及理解复杂指令的困难，建议应增强个性化和互动性，以更好地适应教学设计的需求。

4.讨论

第一，生成式人工智能辅助教学设计实质上是从被动到交互的过程。这一过程的转变显著提升了职前教师在教学设计中的主动性和参与度。通过与生成式人工智能的互动，教师不再是单纯接收外部信息的个体，而是能够积极参与到教学内容的调整和创造中。这种交互性不仅增强了教师对教学设计元素的理解和应用，而且促进了教师专业成长和教学实践的创新。研究表明，生成式人工智能的反馈和建议为教师提供了新的视角和思路，有助于他们构建更加丰富和有效的教学设计。

第二，基于 ICAP 框架理论的生成式人工智能辅助强调深层次认知参与。ICAP 框架理论为我们提供了一个理解学习者认知参与的多维度视角，包括被动、主动和建构性参与。在本研究中，生成式人工智能的应用支持了职前教师在这些层次上的参与，尤其是在建构性参与方面，教师通过与人工智能的互动来构建和重塑教学设计。这种深层次的认知参与对提升

职前教师的教学设计能力至关重要，因为它涉及到对教学内容的深入理解和创新性思考，从而提高了教学质量和教学设计能力。

第三，深层次认知加工活动需要高阶思维的参与。生成式人工智能在教学设计中的应用促进了职前教师高阶思维的发展，包括批判性思维、问题解决和创造性思维。通过提供复杂的教学设计挑战和支持教师在解决这些问题时的思考过程，生成式人工智能有助于教师准备更有效的教学策略，以满足不同学生的学习需求。

5.结论

本研究结合 ICAP 框架理论，通过实证研究验证了生成式人工智能在提升职前教师教学设计能力方面的有效性。实验结果表明，生成式人工智能辅助能够显著提高教学设计方案的质量，并且对职前教师的教学设计能力有积极影响。研究还发现不同绩效水平的职前教师在使用生成式人工智能辅助下的认知网络存在差异，这表明生成式人工智能的使用能够促进深层次的认知参与和高阶思维的参与，揭示了职前教师从被动到交互的过程转变，以及深层次认知加工活动对职前教师教学设计能力提升的重要性，为未来的教学设计培训和课程开发提供了指导。

然而，本研究也存在一定的局限性。例如，样本量较小可能限制了研究结果的普遍适用性，而学习者的年龄和专业分布单一可能影响了研究的广泛性。因此，未来研究可以在本研究的基础上扩大样本量，并探索生成式人工智能在不同教学领域和不同文化背景下的应用效果。此外，随着人工智能技术的不断进步，持续的研究将有助于更深入地理解其在教育领域的应用机制和效果，从而为教育实践提供更加精准和有效的指导。

参考文献

方海光,孔新梅,李海芸 & 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.

王芳.(2022).基于 ICAP 学习方式分类学提升学习参与度的教学实践——以“金属的活动性顺序”为例.教学月刊·中学版(教学参考)(12),45-50.

孙丹,朱城聪,许作栋 & 徐光涛.(2024).基于生成式人工智能的大学生编程学习行为分析研究.电化教育研究(03),113-120.doi:10.13811/j.cnki.eer.2024.03.016.

汪靖,陈恬妮 & 杨玉芹.(2024).思维与调节的融合共创：生成式人工智能支持的人智交互过程与模式研究.中国电化教育(08),45-55.

张馨元 & 张民选.(2024).积极学习：季清华的 ICAP 学习框架.全球教育展望(09),30-50.

杨蕾.(2024).浅谈生成式人工智能在教育中应用的机遇与挑战.甘肃教育研究(11),18-21.

高莉.(2024).从被动到交互——基于 ICAP 的以读促学教学模式探索.大学(02),67-70.

盛群力 & 丁旭.(2018). “ICAP 学习方式分类学”的循证研究.武汉科技大学学报(社会科学版) (02),218-223.

程俊,麦裕华,李林燊 & 王后雄.(2024).指向深度学习的中学化学课堂教学设计评价量规研究.化学教育(中英文)(15),33-39.doi:10.13884/ j.1003-3807hxjy. 20230 70187.

Chi H T M, Wylie R. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes[J]. Educational Psychologist,2014,49(4):219-243.

MIAOFC，HOLMES W. Guidance for generative AI in education and research [EB/OL]. [2023-09-21].https：//unesdoc.unesco.org/ark/48223/pf0000386693.

社交批注对学生在同步在线学习中参与度的影响

The Impact of Social Annotation on Students' Engagement in Synchronous Online Learning

侯怡悦[1*]，张冰儿[2]
[1]National Institute of Education, Nanyang Technological University
[2]School of Humanities, Jiangnan University
nie24.hy2072@e.ntu.edu.sg

【摘要】 近年来同步在线学习的普及凸显了参与度不足的挑战。社交批注工具在异步学习中的协作与认知深化作用已获验证，但其在同步在线学习场景中的应用尚未充分探索。本研究拟采用准实验设计，具体为单组A-B阶段设计，将社交批注工具嵌入同步在线课堂，探究其在行为、认知与情感三个维度上对学习者参与的影响。预期成果将揭示基于社交批注的多模态交互如何重构同步在线学习参与模式，为优化在线课堂设计提供参考。

【关键词】 同步在线学习；社交批注工具；多维参与

Abstract: *The wide adoption of synchronous online learning (SOL) has exposed the ongoing challenges in learner engagement. Although social annotation tools (SATs) have demonstrated promise in an asynchronous context, their potential in SOL remains unexplored. This study proposes a quasi-experimental design, specifically a single-group A-B design, to systematically investigate the roles of SATS in reshaping behavioral, cognitive and emotional engagement in SOL sessions. Findings are expected to provide empirical evidence for optimizing instructional design in technology-enhanced SOL environments.*

Keywords: Synchronous online learning, Social annotation tools, Multidimensional engagement.

1.Introduction

The rapid adoption of synchronous online learning (SOL) enhances the accessibility of learning, yet persistent challenges in engaging learners are identified due to its lack of flexibility in time, pace, and duration (Park & Bonk, 2007; Wang et al., 2023). This condition is further aggravated by the emotional disconnect caused by the loss of physical interaction. Additionally, thinking and verbal interaction could remain superficial resulting from inappropriate pedagogical designs, such as long-time lectures. Thus, learner engagement is counted as the key to the success of SOL (Baxter & Hainey, 2022). Social annotation technologies (SATs) have demonstrated considerable promise in asynchronous online learning (AOL), effectively supporting collaborative learning, critical thinking, and community formation. Nonetheless, their application in SOL remains mostly unexamined. SATs possess several theoretical advantages owing to their real-time, content-specific, and multimodal interactions, which indicates their potential to address the engagement challenges in SOL. The existing research landscape is deficient in empirical evidence concerning the efficacy of SATs in SOL. To fill this gap, this research proposal aims to explore the potential of SAT-mediated learning for enhancing student engagement in the SOL context through a quasi-experimental study.

2.Literature Review

2.1. Engagement Challenges in Synchronous Online Learning (SOL)

SOL facilitates real-time interactions through digital platforms like Zoom, allowing instant dialogue via audio, video, and text, and mirroring some of the dynamics of traditional classrooms (Martin et al., 2017; McArthur, 2022). Despite its advantages, SOL poses significant engagement challenges critical to the effectiveness of online learning (Baxter & Hainey, 2022; Chiu, 2023). Engagement is defined as a psychological investment in and effort directed toward learning, understanding, or mastering (Newmann, 1992). Fredricks et al. (2004) expand on this by categorizing engagement into three critical dimensions: behavioral, cognitive, and emotional. Each plays a vital role in influencing students' learning outcomes. Studies have shown that the unique settings of SOL influence how students engage in learning behaviorally, cognitively, and emotionally (Bedenlier et al., 2020), thereby making the challenge of engagement more severe in this context. First, by providing less supervision, SOL demands a higher level of self-regulation, which not all students possess (Bedenlier et al., 2020; Chiu, 2023; Ryan & Deci, 2020). Therefore, online instructors need to adopt alternative strategies that are better suited to the SOL context (Heilporn et al., 2021), and design appropriate content that satisfies learners' personal needs and learning goals (J. Lee et al., 2021) to engage learners both behaviorally and cognitively. Second, being physically separated and lacking body language may reduce opportunities for informal interactions, as well as weaken the sense of connection among students, increasing feelings of isolation and lowering engagement (Bedenlier et al., 2020; Chiu, 2023; Wang, 2024). Such emotional disengagement is particularly detrimental as it directly affects students' attitudes towards the learning process and their satisfaction with learning experiences. Third, verbal-centered interaction privileges native speakers, while discouraging engagement among students facing language barriers (Park & Bonk, 2007) or with an introverted personality. Fourth, cognitive and emotional engagement is less observable in the SOL context due to the lack of face-to-face communication (Bedenlier et al., 2020), preventing instructors from recognizing disengagement and intervening in time. Moreover, research on emotions in online learning remains relatively scarce (Hsu et al., 2019). These multifaceted challenges underscore the necessity for SOL designs and pedagogical strategies that consider and address the specific engagement challenges inherent to this learning mode.

2.2. Social Annotation Technologies (SATs) as Engagement Interventions

SATs are typically web-based platforms that enable students to collaboratively share written comments and highlights directly with peers and instructors (Lazzara & Clinton-Lisell, 2024). At its core, annotation is a dialogic process—a way for readers to engage with, critique, and apply knowledge from texts (R. H. Kalir & Garcia, 2021). Consequently, SATs serve as an effective means of enhancing learner engagement in online courses (Bjorn, 2023). Compared to traditional annotation methods, SAT is particularly dynamic due to its integrated social functions (Brown & Croft, 2020). Students interact through upvoting, tagging, and emoji-based reactions (Li & Li, 2023), thereby fostering a vibrant and participatory learning community. These interactions not only strengthen connections among learners but also help mitigate the isolation often felt in online courses (J. H. Kalir et al., 2020). Beyond fostering social bonds, SATs support cognitive

engagement through peer scaffolding and structured communication, which enhances motivation and learning outcomes (Razon et al., 2012). By allowing students to highlight and link specific parts of a text to discussion, SATs anchor conversations in a shared context, enabling deeper, more focused dialogue (Sun & Gao, 2017). Additionally, SATs capture students' in-process thinking (Adams & Wilson, 2020), creating a visible record of their cognitive engagement. Existing studies have consistently shown that SAT application enhances student engagement, attention, and interaction, fostering advanced cognitive skills and reading comprehension (Bjorn, 2023; Hwang et al., 2007; Johnson et al., 2010; Razon et al., 2012; Yang et al., 2013).

While most existing studies focus on the use of SATs in AOL, their potential in SOL remains unexplored. In SOL, interactions tend to be dominated by real-time instructor-led communication. Integrating SATs offers a promising intervention to enrich all three key forms of interaction: learner-content, learner-instructor, and learner-learner (Moore, 1989). First, by enabling students to annotate and discuss specific parts of learning materials, SATs enhance learner-content interaction (S. Lee et al., 2023). It serves as the foundation for other interactions, as it drives cognitive processing (Moore, 1989; Xiao, 2017). This approach supports transforming learning from passive reception of information to active knowledge construction. Second, learner-learner interaction is strengthened through replies, upvotes, and emotive reactions (Allen & Randall, 2023). These interactions can increase the sense of community and enhance collaborative learning. Third, learner-instructor interaction can be diversified, as instructors gain access to learners' in-process thinking through annotations. Therefore, instructors could provide immediate and contextual feedback. Furthermore, integrating SATs expands communication channels and extends the temporal boundaries of live sessions. Text-based communication helps create a more inclusive environment, which is especially meaningful for learners who feel less comfortable with verbal expression in SOL. Meanwhile, the annotation records enable post-session review and encourage deeper cognitive engagement. In sum, SATs hold theoretical potential to engage students in synchronous sessions. However, empirical evidence remains limited, calling for further research.

3.Research Questions (RQs)

RQ1: What are the differences in learners' behavioral interactions between SAT-enabled and non-SAT synchronous contexts?

RQ2: Do SATs influence students' cognitive engagement in synchronous settings?

RQ3: Do SATs influence students' emotional engagement in synchronous settings?

4.Methodology

4.1. Research Design

A single-group A-B design will be employed to explore the impact of SATs on student engagement. By comparing the same group of learners before and after the intervention, this design minimizes between-group variability and provides clear observation of the changes associated with the introduction of SATs within an authentic instructional setting. The study will consist of two phases, each containing three sessions. Phase 1 will serve as the control condition, where instruction will be delivered solely through Zoom. In Phase 2, SAT (Perusall) will be integrated into the teaching

strategy. The instructor will upload slides to Perusall and share the interface via screen sharing. Students can annotate slides, ask questions, reply to peers, share emojis, and upvote others' annotations in real time via Perusall. Meanwhile, students can still use Zoom. They can engage through annotations, chat messages, or voice messages. This design provides multiple interaction methods and fosters authentic learning practices. It also enables instructors to observe shifts in students' natural behaviors within a realistic context.

4.2. Participants

The study will be conducted at a teacher training institute. Approximately 30 undergraduate students specializing in education will be recruited for the study. Before the study begins, students will be asked to provide their written consent to participate. This emphasizes voluntary participation and the anonymization of the collected data.

4.3. Data Collection

All Zoom sessions throughout the study will be recorded and the annotated slides in Phase 2 will be preserved. In Phase 1, the data will include the number and content of chat messages and verbal contributions during Zoom classes. At the end of Phase 1, a survey will be conducted to collect both quantitative and qualitative data about learners' self-perceived depth of thinking, and emotional experiences. The survey (Survey 1) consists of 10 five-point Likert scale items (from "5: Strongly agree" to "1: Strongly disagree") and one open-ended question (see Appendix A). Cognitive and emotional dimensions each contain 5 items to capture learners' self-perceived engagement. The open-ended questions aim to collect additional information not covered by the Likert-type items. In Phase 2, the same data will be collected from Zoom interactions. Then, data about the number and content of annotations will be collected from the Perusall log files. After the Phase 2 sessions, another survey (Survey 2) will be administered. Its quantitative section and the first open-ended question mirror Survey 1, while a second open-ended question is added to explore students' perceptions of the embedded SAT (see Appendix B). To ensure data validity, the survey instruments will be self-developed based on established engagement constructs (Fredricks et al., 2004) and will be reviewed by two experts in educational technology to assess item clarity, relevance, and alignment with theoretical definitions. Reliability will be supported through a pilot test involving five non-participant students, followed by internal consistency checks targeting a Cronbach's alpha greater than 0.70. For coding qualitative data, a predefined rubric will be employed. 20% of the data will be independently coded by two researchers. Inter-rater reliability will be assessed using Cohen's kappa ($\kappa \geq 0.70$).

4.4. Data Analysis

For RQ1, we will focus on the total interaction frequency and modality distribution across the phases to investigate varying behavioral patterns. The total interaction in Phase 1 encompasses chat messages and verbal contributions via Zoom. In Phase 2, the total quantity of Perusall annotations will be added. We will then compare the total interactions through paired *t*-tests or a non-parametric Wilcoxon signed-rank test. In addition, the modality distribution (text, language, annotations) will be analyzed through Chi-square tests. For RQ2, we will perform paired *t*-tests (or Wilcoxon tests for non-normally distributed data) on the composite cognitive depth scores measured by Survey 1 and 2. Meanwhile, all interaction data (Zoom chat transcripts, verbal transcripts, Perusall annotations) will

be coded into two cognitive categories: basic (e.g., restatement of facts, simple questioning) and deep (e.g., critical analysis, synthesis of ideas). A predefined coding scheme will be applied, and inter-rater agreement will be checked through independent coding. A Chi-square test will then be used to assess whether the proportion of deep cognitive interactions has increased significantly. For RQ3, the emotional composite scores obtained from Survey 1 and 2 will be compared via paired *t*-tests or Wilcoxon tests. Additionally, we will thematically analyze the responses to the first open-ended question in both phases to identify recurring emotion themes (e.g., pleasure, frustration). Moreover, the responses to the supplemental open-ended question in Survey 2 will be systematically coded into three predefined categories: behavioral (e.g., interaction modality shifts), cognitive (e.g., critical thinking triggers), and emotional (e.g., tool-related stress or motivation). A consensus-based coding rubric will be employed, with two researchers independently coding to ensure reliability. Emerging themes will then be mapped to quantitative findings to explain how SAT tools influence engagement. This analysis aims to provide mechanistic insights, thereby enriching the interpretation of the results across all RQs.

5.Limitations

Firstly, the findings are based on a single course, which may restrict the generalizability of the results to other disciplines. Secondly, the reliance on self-reported data could introduce bias and oversimplify the emotional experiences of participants. Thirdly, the relatively short intervention period of three sessions may not be sufficient to observe long-term patterns of engagement. Fourthly, the use of a single-group A-B design limits the ability to infer causal relationships. Changes in engagement might also be influenced by maturation, novelty effects, or other external variables. Future studies should consider randomized controlled trials or crossover designs to strengthen causal inferences.

6.Expected Contributions

Firstly, this study aims to provide empirical evidence for SATs' potential in live sessions and offers practical insights for educators to design more flexible, interactive, and learner-centered instruction in SOL. Secondly, this study introduces a framework for capturing silent but active engagement through annotation analysis, providing conditions for longitudinal tracking of cognitive and emotional trajectories in SOL. Thirdly, this study investigates how adopting emerging technologies can facilitate the creation of alternative interaction methods. Thus, learners can make meaningful contributions according to their individual preferences in a more equitable and inclusive learning environment.

References

Adams, B., & Wilson, N. S. (2020). Building Community in Asynchronous Online Higher Education Courses Through Collaborative Annotation. *Journal of Educational Technology Systems*, *49*(2), 250–261. https://doi.org/10.1177/0047239520946422

Allen, A. A., & Randall, K. N. (2023). Evaluating a social annotation tool for engagement and learning with preservice special education teachers. *Information and Learning Sciences*, *124*(11/12), 460–477. https://doi.org/10.1108/ILS-04-2023-0040

Baxter, G., & Hainey, T. (2022). Remote learning in the context of COVID-19: Reviewing the effectiveness of synchronous online delivery. *Journal of Research in Innovative Teaching & Learning*, *16*(1), 67–81. https://doi.org/10.1108/JRIT-12-2021-0086

Bedenlier, S., Bond, M., Buntins, K., Zawacki-Richter, O., & Kerres, M. (2020). Facilitating student engagement through educational technology in higher education: A systematic review in the field of arts and humanities. *Australasian Journal of Educational Technology*, *36*(4), Article 4. https://doi.org/10.14742/ajet.5477

Bjorn, G. (2023). The Power of Peer Engagement: Exploring the Effects of Social Collaborative Annotation on Reading Comprehension of Primary Literature. *AI, Computer Science and Robotics Technology*, *13*. https://doi.org/10.5772/acrt.24

Brown, M., & Croft, B. (2020). Social Annotation and an Inclusive Praxis for Open Pedagogy in the College Classroom. *Journal of Interactive Media in Education*, *2020*(1). https://doi.org/10.5334/jime.561

Chiu, T. K. F. (2023). Student engagement in K-12 online learning amid COVID-19: A qualitative approach from a self-determination theory perspective. *Interactive Learning Environments*, *31*(6), 3326–3339. https://doi.org/10.1080/10494820.2021.1926289

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

Heilporn, G., Lakhal, S., & Bélisle, M. (2021). An examination of teachers' strategies to foster student engagement in blended learning in higher education. *International Journal of Educational Technology in Higher Education*, *18*(1), 25. https://doi.org/10.1186/s41239-021-00260-3

Hsu, H.-C. K., Wang, C. V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. *Education and Information Technologies*, *24*(3), 2159–2174. https://doi.org/10.1007/s10639-019-09863-w

Hwang, W.-Y., Wang, C.-Y., & Sharples, M. (2007). A study of multimedia annotation of Web-based materials. *Computers & Education*, *48*(4), 680–699. https://doi.org/10.1016/j.compedu.2005.04.020

Johnson, T. E., Archibald, T. N., & Tenenbaum, G. (2010). Individual and team annotation effects on students' reading comprehension, critical thinking, and meta-cognitive skills. *Computers in Human Behavior*, *26*(6), 1496–1507. https://doi.org/10.1016/j.chb.2010.05.014

Kalir, J. H., Morales, E., Fleerackers, A., & Alperin, J. P. (2020). “When I saw my peers annotating”: Student perceptions of social annotation for learning in multiple courses. *Information and Learning Sciences*, *121*(3/4), 207–230. https://doi.org/10.1108/ILS-12-2019-0128
Kalir, R. H., & Garcia, A. (2021). *Annotation* (Annotated edition). The MIT Press.
Lazzara, J., & Clinton-Lisell, V. (2024). Using social annotation to enhance student engagement in psychology courses. *Scholarship of Teaching and Learning in Psychology*, *10*(4), 605–611. https://doi.org/10.1037/stl0000335
Lee, J., Sanders, T., Antczak, D., Parker, R., Noetel, M., Parker, P., & Lonsdale, C. (2021). Influences on User Engagement in Online Professional Learning: A Narrative Synthesis and Meta-Analysis. *Review of Educational Research*, *91*(4), 518–576. https://doi.org/10.3102/0034654321997918
Lee, S., Foster, C., Zhong, M., Bruce-Opris, H., Mainlyng Duenas, Parente, V., Reid, C., & McCartney, M. (2023). Annotations Serve as an On Ramp for Introductory Biology Students Learning to Read Primary Scientific Literature. *Journal of Microbiology & Biology Education*, *24*(1), e00214-22. https://doi.org/10.1128/jmbe.00214-22
Li, M., & Li, J. (2023). Using Perusall to motivate students’ curriculum-based academic reading. *Journal of Computers in Education*, *10*(2), 377–401. https://doi.org/10.1007/s40692-022-00234-y
Martin, F., Ahlgrim-Delzell, L., & Budhrani, K. (2017). Systematic Review of Two Decades (1995 to 2014) of Research on Synchronous Online Learning. *American Journal of Distance Education*, *31*(1), 3–19. https://doi.org/10.1080/08923647.2017.1264807
McArthur, J. A. (2022). From classroom to Zoom room: Exploring instructor modifications of visual nonverbal behaviors in synchronous online classrooms. *Communication Teacher*, *36*(3), 204–215. https://doi.org/10.1080/17404622.2021.1981959
Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, *3*(2), 1–7. https://doi.org/10.1080/08923648909526659
Newmann, F. M. (1992). *Student Engagement and Achievement in American Secondary Schools*. Teachers College Press, 1234 Amsterdam Avenue, New York, NY 10027 (paperback: ISBN-0-8077-3182-X, $17. https://eric.ed.gov/?id=ED371047
Park, Y. J., & Bonk, C. J. (2007). Synchronous Learning Experiences: Distance and Residential Learners’ Perspectives in a Blended Graduate Course. *Journal of Interactive Online Learning*, *6*(3), 245–264.
Razon, S., Turner, J., Johnson, T. E., Arsal, G., & Tenenbaum, G. (2012). Effects of a collaborative annotation method on students’ learning and learning-related motivation and affect. *Computers in Human Behavior*, *28*(2), 350–359. https://doi.org/10.1016/j.chb.2011.10.004
Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860
Sun, Y., & Gao, F. (2017). Comparing the use of a social annotation tool and a threaded discussion forum to support online discussions. *The Internet and Higher Education*, *32*, 72–79. https://doi.org/10.1016/j.iheduc.2016.10.001

Wang, Q. (2024). *Designing Technology-Mediated Learning Environments: Perspectives, Processes, and Applications*. Springer Nature Singapore. https://doi.org/10.1007/978-981-96-0680-1
Wang, Q., Wen, Y., & Quek, C. L. (2023). Engaging learners in synchronous online learning. *Education and Information Technologies*, *28*(4), 4429–4452. https://doi.org/10.1007/s10639-022-11393-x
Xiao, J. (2017). Learner-content interaction in distance education: The weakest link in interaction research. *Distance Education*, *38*(1), 123–135. https://doi.org/10.1080/01587919.2017.1298982
Yang, X., Yu, S., & Sun, Z. (2013). The effect of collaborative annotation on Chinese reading level in primary schools in China. *British Journal of Educational Technology*, *44*(1), 95–111. https://doi.org/10.1111/j.1467-8535.2011.01277.x

Appendix A
Phase 1 Survey
Cognitive Engagement (5-point Likert scale, 1=Strongly Disagree, 5=Strongly Agree)
The session content prompted me to think deeply.
I actively connected the material to my prior knowledge or experiences.
The instructor's questions encouraged me to reflect on my understanding.
I could explain the key concepts in my own words.
I identified gaps in my knowledge during the session.
Emotional Engagement (5-point Likert scale, 1=Strongly Disagree, 5=Strongly Agree)
I felt motivated to participate in class discussions.
The learning environment felt safe for expressing ideas.
The pace of the session made me feel stressed. *(Reverse-scored)*
Interactions with peers increased my motivation.
I felt satisfied with the overall session experience.
Open-Ended Questions
Describe a moment in the session where you felt particularly engaged or frustrated. Explain why.

Appendix B
Phase 2 Survey
Cognitive Engagement (5-point Likert scale, 1=Strongly Disagree, 5=Strongly Agree)
The session content prompted me to think deeply.
I actively connected the material to my prior knowledge or experiences.
The instructor's questions encouraged me to reflect on my understanding.
I could explain the key concepts in my own words.
I identified gaps in my knowledge during the session.
Emotional Engagement (5-point Likert scale, 1=Strongly Disagree, 5=Strongly Agree)
I felt motivated to participate in class discussions.
The learning environment felt safe for expressing ideas.
The pace of the session made me feel stressed. *(Reverse-scored)*
Interactions with peers increased my motivation.
I felt satisfied with the overall session experience.
Open-Ended Questions
Describe a moment in the session where you felt particularly engaged or frustrated. Explain why.
Did Perusall affect your learning experience? Please describe specific features or interactions that stood out to you (positively or negatively).

基于调节学习的多智能体支持的协同论证设计研究

Research on multi-agent assisted collaborative argumentation design based on regulated learning

刘文平[1]，余承珂[1]，马志强[1]
[1]江南大学 江苏“互联网+教育”研究基地
[*] liuwenping0103@163.com

【摘要】论证是学生开展协作学习的有效途径，但学生在往往面临缺乏特定专业知识和论证知识，动机和投入度低等调节能力不足的问题。研究基于调节学习理论，设计了融入多智能体支持的协同论证活动，并检验了其对学生论证地图质量、论证话语模式的影响。研究发现，在论证地图方面，多智能体可以帮助学生搜寻多重证据支持立场，并为协同论证提供稳定的环境。在论证话语模式方面，多智能体可以帮助学生进行知识重组、优化协作流程，促进论证模式从个体论证转变为批判性共建。研究结论对教育多智能体的设计与开发及协作学习活动的开展均有一定参考价值。
【关键词】 大语言模型；教育智能体；协同论证；调节学习；生成式人工智能

Abstract: *Argumentation is an effective way for students to engage in collaborative learning, but students are often faced with a lack of regulating skills such as a lack of specific expertise and knowledge of argumentation, and low levels of motivation and engagement. Based on regulation learning theory, the study designed a collaborative argumentation activity incorporating the support of large language model agents and examined its effects on the quality of students' argumentation maps and argumentation discourse patterns. It was found that in terms of argumentation maps, large language model agents can help students search for multiple evidence to support claims and can provide a stable environment. In terms of argumentation discourse patterns, large language model agents can help students reorganize their knowledge, optimize the collaborative process, and facilitate the change of argumentation patterns from individual argumentation to critical co-construction. The findings are informative for the design and development of large language model agents in education and the development of collaborative learning activities.*
Keywords: large language model, pedagogical agent, collaborative argumentation, regulated learning, generative AI

1.引言

协同论证作为一种社会建构主义学习范式，是指学习者通过观点交锋、证据整合与逻辑重构，最终达成共识的一种学习方法。通过学生之间及时性的反馈与批判性的思考，协同论证有助于学生建立主张与论据之间的联系，促进知识的有效建构。已有研究表明，协同论证可以有效提高个人的论证能力和对复杂学术内容的概念理解，从而促进其高阶思维的发展。

然而，学生在协作交互中面临认知和社会性交互方面的困难。在认知层面，论证知识不足，缺乏特定专业知识等问题，阻碍了学生提出更有说服力的论证观点(Liu 等, 2019)。此外，学生常因同时处理多维度任而超出学生工作记忆容量，导致学生认知负荷过高。在社会性交互方面，协作过程中学生之间产生的认知冲突和情绪化困境亟待解决。为了减少与同伴发生冲突或降低社交压力，学习者通常避免对学习伙伴的论点产生反驳，倾向于快速达成共识。此外，学生在协作小组中经常会出现搭便车的行为，处于边缘的学习者难以

参与到论证活动中，群体参与度失衡问题显著。造成学生在任务目标设定、元认知监控与冲突协调等环节表现薄弱的原因之一是由于学生调节能力不足。如果学生调节能力不足，协作学习便很难达到知识建构，学生的论证能力和高阶思维能力也难以得到发展。因此，往往需要教师给予提示或使用其他支持工具来帮助学生监控协作过程，但在传统课堂中教师难以及时为每个学习者提供个性化指导，且费时费力。因此，有必要整合技术工具支持学生的调节学习。

随着生成式人工智能技术的突破，大模型（Large Language Model，LLM）正深度融入教育领域，重塑教学范式。当前，基于 LLM 的教育智能体在情绪与动机调节、实时学习干预及协作学习优化等方面展现出显著潜力(徐振国等, 2021)，为协同论证的智能化支持提供了新路径。然而，单智能体在协同论证场景下面临多重瓶颈，由于认知视角单一，角色固定等问题难以满足教育需求(Li 等, 2024)。首先，在技术层面，单智能体在处理发杂推理问题时经常表现出逻辑推理能力不足，导致论证链断裂等问题，且由于垂直领域知识缺失而易引发“幻觉”现象。其次，在社会交互层面，单智能体的静态反馈机制难以适配动态需求(T. Guo 等, 2024)，其薄弱的情感共情能力可能阻碍学生进行社会性调节，进一步可能导致学生论证深度受限，甚至误导学生，造成社会协作能力退化(翟雪松等, 2024)。

基于此，大模型驱动的教育多智能体系统 (Multi-Agent System, MAS)应运而生，为解决上述问题提供了思路。教育多智能体是由多个智能体组成的协同系统，其核心在于通过多个智能体的协作的分布式协作突破单一智能体的限制，以此解决复杂问题。与单一智能体不同，多智能体系统强调智能体间的交互、协同与任务管理，具有多角色协调、分布式决策、动态适应性等特性(刘石奇等, 2025)，其核心优势可以总结为以下几点：首先，模块化架构允许按需动态调整智能体的组合方式和互动策略。其次，基于强化学习的协作机制可提升对复杂问题的响应与推理，网状的工作流更好满足处理复杂问题的多重需求，具有较强适应性。第三，多个智能体互动对话模拟了真实社交情景，有助于学生提交社交感知力。已有研究证明了多智能体系统在促进协作学习场景中的应用(翟雪松等, 2024; S. Guo 等, 2024)。

然而，当前鲜有研究提出如何设计多智能体的角色和介入环节来支持学生在协同论证中的调节学习，探讨多智能体在协同论证中的作用和机制。因此，本研究结合生成式 AI 支持学习与认知发展的指导框架(Tseng & Warschauer, 2023)和知识整合框架(Linn & Eylon, 2011)，提出多智能体支持的协同论证框架。此外，研究基于调节学习的认知性和社会性调节两个维度设计多智能体系统，并探讨该多智能体系统对协同论证的影响。研究问题如下：

相较于传统生成式 AI 支持的协同论证，融入认知性和社会性调节策略的多智能体对学生论证结果的质量有何影响？

相较于传统生成式 AI 支持的协同论证，融入认知性和社会性调节策略的多智能体对学生的论证话语模式有何影响？

2.研究方法

2.1. 研究对象

本研究选取华东地区某高校的研究生一年级学生作为研究对象，在研究生必修课《信息化教育理论与实践》中进行，共有 49 名学生参与。该课程由一名在信息教育研究领域有超过 10 年教学经验的教授教授。在学习活动中，学生被随机分为 11 组（每组 4-5 人），

22 名学生为对照组，27 名学生为认知性+社会性调节的多智能体组。在实验前两组学生都填写了有关科学论证性的调查问卷，并对两组的论证倾向进行正态性检验（Shapiro-Wilk，$p > 0.05$）和方差齐性检验（$p = 0.075 > 0.05$）。接着，独立样本 t 检验结果显示两组前测得分无显著差异（$t = 1.148$, $p = 0.258 > 0.05$）。

2.2. 研究设计

本研究基于 Coze 平台开发了协同论证辅助工具，从认知调节层次和社会调节层次构建了六个基本智能体。Coze 平台（https://www.coze.cn/）可以通过低代码、可视化搭建智能体,技术门槛较低，便于研究者开发多角色的多智能体。如表 1 为六种智能体的作用及反馈示例。

表 1 多智能体角色及反馈示例

定制策略类型	智能体角色	示例
认知层次	论证辅助智能体：引导学生依据论证主题逐步思考。	“在开始讨论之前，我想先了解一下你的观点或主张是关于什么的呢？”
	思维发散智能体：针对学生的论证主张，引导学生多角度思考。	“我们先来探讨你的论点，并逐步分析。现在，让我们来考虑一些可能的思路……”
	观点整合智能体：依托论证模型辅助学习者优化论证要素。	“很好，你已经成功地将论证的各个部分区分开来。现在，让我们进一步完善”
社会层次	规划智能体：引导学生规划协同论证目标和任务	“开展有效的小组论证，需要大家共同合作，下面我会引导你如何进行……”
	情绪调节智能体：通过鼓励性引导激发学生论证动机	“每个人都可以用‘我同意/补充/好奇……’的句式表达意见，别担心不完美，论证本就是共同完善的过程！”
	冲突调节智能体：弱化在论证过程中学习者由于认知冲突而惧怕提出反驳。	“应对小组论证中的反对意见是一个重要的环节，有助于深化讨论和提高论证。下面是一些应对反对意见的策略……”

本研究设计为期 6 周的准实验研究，以评估基于调节学习的多智能体对学习者的协同论证质量、论证话语模式的影响。研究主要进行两轮论证，每轮论证均设计一个与教育学相关的社会性议题，第一轮论证话题为“你认为人类智慧和人工智能是竞争还是合作关系？”，第二轮为“你认为生成式人工智能（Artificial Intelligence，AI）更应用于促进高效短识技能学习还是批判性与创造性思考?”研究将协同论证活动分为如下四个阶段，如图 1。

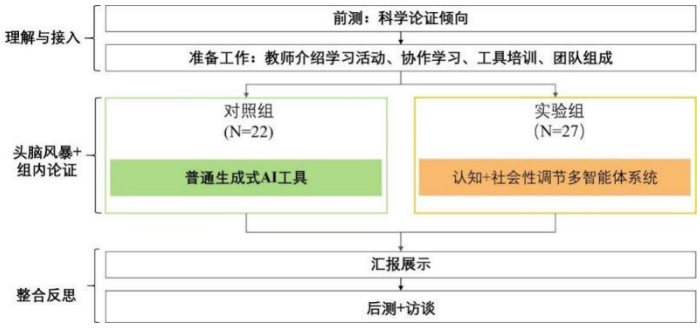


图 1 实验流程图

(1) 理解与接入：在活动启动前，教师介绍协同论证的活动流程和在线协作平台（奇点协作通）的使用方法并对学习者进行论证和多智能体工具使用的培训。

(2) 头脑风暴：教师阐述论证主题并指导学生进行学习活动。实验组使用认知与社会叠加的智能体，控制组不对生成式 AI 进行定制。小组成员需要各自查找与论证主题相关的数据与资料，并发布到小组的微信群中，为后续的论证活动提供证据支持。

(3) 组内论证：小组成员根据所收集的资料，结合自己的现有知识发散思维、阐述观点，并对同伴和多智能体的观点进行评估、追问和解释。学生在该过程中被要求应该谨慎判断多智能体响应内容的来源与质量，通过比较多个数据源来验证内容的准确性。

(4) 整合与反思阶段：小组成员使用在线协作白板工具，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，每组进行汇报展示，随后小组成员对活动过程进行反思，进一步理解多智能体在支持协同论证方面的优势与缺陷。

2.3. 研究工具

为评价学习者绘制的论证地图质量，本研究采用了 Cho K-L 等研究提出的论证地图评价指标评分，以此统计论证地图的数量和复杂度(Cho & Jonassen, 2002)。该评价指标基于图尔敏论证模式编制，从主张、数据和资料、理由、支持、反驳五个维度进行评分，每个维度被划分为 0 分、2 分、4 分、6 分四个层次。两位研究者认真学习论证地图评分工具的评分规则之后，随机抽取 30%的数据进行独立编码，编码的 Kappa 系数为 0.652，接着研究者针对编码中存在的异议进行了协商，并在协商统一后完成了剩余的编码。

为评价学习者的论证话语模式，研究使用论证话语编码方面框架来分析论证对话的认知属性与社会属性，编码框架改编自(Y.-C. Chen 等, 2016; Su 等, 2021)。其中认知属性包含主张、证据、推理、反驳四个维度，每个维度下包含低与高两个层级。社会性维度包含提出论点、解释论点、支持他人观点等 8 个维度，编码框架如表所示。编码者在结构性认知属性的 Kappa 系数为 0.709，社会性交互属性的 Kappa 系数为 0.727，一致性较强。

表 2 论证话语编码方面框架

论证维度	要素	层级	编码	描述	论证维度	要素	编码	描述
认知维度	主张（C）	提取	CE	复述其他论证主体的主张，或简单主张	社会维度	提出	P	提出自己的论点
		界定	CD	深入分析理由去支持自己的主张		解释	E	解释、拓展自己的论点
	证据（E）	描述	ED	对证据进行大概的描述且缺乏引用来源		支持	SP	支持他人的论点

理由（R）	拓展	EE	对所引用的证据进行细致描述	挑战	C	指出他人论点的不足
	证实	RI	简单识别证据与主张之间的联系	证实	CK	检查证据的有效性与可信度
	拓展	RE	有力地从证据推论到结论	辩护	D	说服他人接受自己的论点
反论点（CA）	反驳	CAR	直接提出反面论点	整合	I	将不同的论点整合为新的论点
	拓展	CAE	运用支持反面论点的证据进行详细论证			

2.4. 数据搜集与分析

在论证地图的数据收集与分析方面，研究期间共收集了 22 份论证地图。接着，研究者对每轮论证地图的两个条件组的论证要素数量和质量进行打分，并在两轮活动结束后，计算每个条件组每类要素数量和质量的平均值，最后得出每个条件组论证地图数量和质量的两轮平均值。编码完成后，利用 Shapiro-Wilk 检验对每个条件组的论证要素的数量与质量进行正态分布检验，结果显示所有组数据的 p 值均大于 0.05，符合正态分布假设。接着，进一步采用独立样本 t 检验比较两组论证要素数量和复杂度的差异。

在论证话语模式方面，研究搜集了学生在微信中的论证记录，共计 947 条话语，每个对话框中元素相互关联，且每条对话较为复杂，包含多个论证要素，体现了学生对论点的完整论证，对话之间的时序性明显，因此每个对话框被视为一个独立的编码单元。在确定上述内容后，由两位研究者进行独立编码，并在消除异议完成了后续数据的编码。最后，研究者用 ONA 分析工具（https://app.epistemicnetwork.org/）进行了有序网络分析。在 ONA 图谱中，每个编码元素以节点形式呈现，节点间连线代表编码要素间的关系。节点大小和饱和反映要素的自连接频率，节点的外半径反映了与其他要素的连接频率，节点间连线箭头所指的方向代表了节点的响应方向。ONA 可以通过编码要素共现的比例生成网络图谱，进一步揭示编码元素之间的关系。

3.实验结果

3.1. 论证地图要素数量和质量

在论证地图要素的数量方面，独立样本 t 检验分析结果显示，在数据和资料（$t = 5.344, p = 0.03 < 0.05$）上有显著差异，其他论证要素的均值实验组均大于对照组，但无显著差异。在论证地图要素的质量方面，独立样本 t 检验结果显示，在数据和资料（$t = 8.247$，$p = 0.000 < 0.01$）和总要素质量（$t = 3.057$，$p = 0.028 < 0.05$）上有显著差异，其他论证要素质量方面实验组均大于对照组，但无显著差异。

表 3 论证要素数量的独立样本 t 检验分析结果

分析项	项	样本量	平均值	标准差	t	p
论点数量	实验组	6	3.58	2.92	1.244	0.245
	对照组	5	1.9	0.74		
论点质量	实验组	6	5.92	0.2	0.576	0.579
	对照组	5	5.8	0.45		
数据和资料数量	实验组	6	18.92	10.06	2.841	0.030*

	对照组	5	6.7	2.86		
数据和资料质量	实验组	6	5.54	0.24	8.247	0.000**
	对照组	5	4.29	0.26		
理由数量	实验组	6	19.5	14.39	2.179	0.074
	对照组	5	6.2	3.68		
理由质量	实验组	6	5.53	0.4	2.115	0.064
	对照组	5	4.53	1.08		
支持数量	实验组	6	15.83	8.08	1.938	0.085
	对照组	5	7.8	4.87		
支持质量	实验组	6	5.31	0.44	2.068	0.069
	对照组	5	4.73	0.48		
反驳数量	实验组	6	7.08	4.79	0.72	0.49
	对照组	5	5.3	2.99		
反驳质量	实验组	6	4.32	1.22	0.737	0.48
	对照组	5	3.77	1.21		
总要素数量	实验组	6	64.92	38.15	2.221	0.065
	对照组	5	27.9	13.29		
总要素质量	实验组	6	26.61	0.94	3.057	0.028*
	对照组	5	23.12	2.4		

* p<0.05 ** p<0.01

3.2. 有序网络分析

如图 2 为两组论证对话中的结构性性认知特征网络图及两组网络图的叠减图。在控制组中，CD、CAR、ED 节点较大，形成了 CD→ED→CAR 的连接。在实验组中，CD、ED、RE、EE 节点较大，响应更多，形成了 CD→ED→RE→EE 的循环，表示小组成员提出主张后，能进行由浅入深的推理并且能进一步给出高阶的论据。在叠减图中可以看到，对照组从低水平主张到低水平推理再到高水平主张的联系更强（CE→CD→CAR），这说明对照组在提出主张后，进一步引起组员对其主张的深入阐释，接着组员之间倾向于直接提出反驳。实验组在证据、理由、反驳之间都有更多的联系（ED→RE→CAR，CAE→RE）。这表明，与对照组相比，实验组更擅长在给出论据并说明理由之后使用论据从反面进行论证。

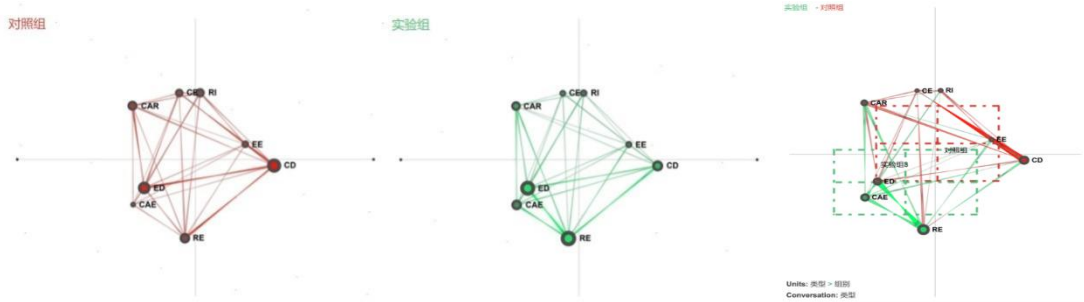


图 2 对照组（左）和实验组（中）结构性认知特征网络图及两组网络图的叠减图（右）

如图 3 为两组论证对话中的社会性交互特征网络图及两组网络图的叠减图。对照组以 P→E 模式为主，说明小组成员倾向于在指出观点后进行进一步的阐释。实验组在 P→C→E 的连接较强，说明实验组小组在提出论点后，其他成员倾向于先提出反驳再进一步使用

论据证实论点。叠减图显示，对照组在 P→E→D 连接更强，说明对照组倾向于提出主张，使用论据说明主张，并说服其他成员接受主张。实验组在 C→E→P 表现更强，说明实验组成员善于指出他人论点的不足，并进一步使用证据佐证，最后提出自己的观点。此外，从这些模式中可以看出，基于调节学习的多智能体在协同论证活动中改变了论证的互动模式。控制组倾向于个体论证，而实验组的互动模式表明，小组成员的批判性思维更强，更加善于利用反驳，倾向于批判性集体论证。

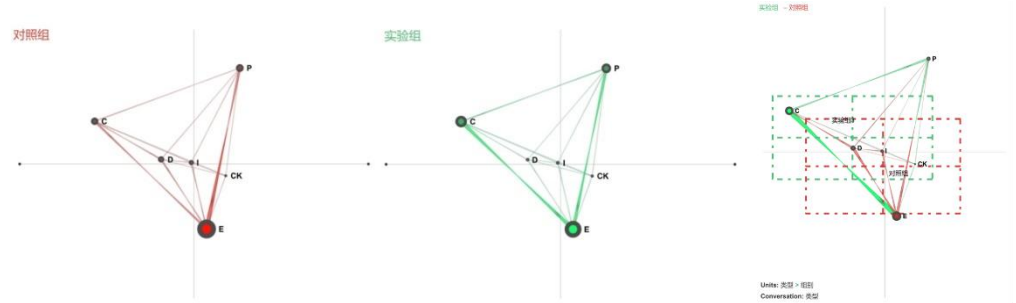


图 3 对照组（左）和实验组（中）社会性交互特征网络图及两组网络图的叠减图（右）

4.讨论及启示

4.1. 讨论

首先，在论证结果方面，融入协同调节策略的多智能体可以提升论证质量。具体来说，叠加认知和社会性调节策略的多智能体支持的小组所绘制的论证地图中，数据和资料要素的数量和质量以及总要素质量显著高于对照组。这说明，认知维度多智能体在论证时能够激发学生思考，帮助学生筛选出与主张相关的可靠资料并将其整合到论证地图(Darmawansah 等, 2025)。这一发现与之前的研究一致，印证了智能体在提升学生论证深度的作用(Lee 等, 2024; Su 等, 2023)。此外，社会维度多智能为协同论证提供了稳定的协作环境，间接学习者在不同交互层次的认知调节(吴林静等, 2024)，这使得实验组总要素质量显著大于对照组。

其次，在论证话语模式方面，基于调节学习的多智能体能够将论证模式由个体论证转向集体思辨论证。在结构性认知方面，多智能体能够激发学生阐释更多的论点，增加学生提出反驳的数量，提升批判性思维(X. Chen 等, 2025)。具体来说，控制组在成员提出高阶主张后仅通过低阶证据简单支撑，随即触发直接反驳，说明控制组仅以个人论证为主，且陷入对抗式辩论的困境。实验组的模式强调在提出主张后，通过多级逻辑推演从高阶证据补充深化主张，更擅长在推理后使用反面论证。研究表明，反驳可以促进共识的形成，通过质疑与反驳对方的观点，小组可以探索不同的解决方案，最终达成一个更加辩证的结论(Berland & Lee, 2012)。在社会性交互方面，控制组在提出主张后进行自我解释，并说服其他成员接受自己的主张，而实验组的 C→E→P 模式则通过冲突促进小组成员使用论据进一步证实，强调以论据说服对方。这说明基于调节学习的智能体能够，推动学生从单向推理转向辩论性修正，形成更稳定的知识建构模式(吴林静等, 2024)。

4.2. 研究启示

本研究的发现对于后续将多智能体介入到协作学习活动中具有重要意义。首先，本研究提出了基于调节学习理论的多智能体设计框架，并证实了在协作学习中融入多智能体支持学生调节的可行性。本研究在理论层面突破了单一维度调节的局限性，指出基于认知和社会性调节的多智能体可以支持学生论证，并让学生和教师从中获益，这极大地改善了

学生论证的质量并有效解决了教师提供支持不足的情况。第二，本研究证实了教师通过介入多智能体改善学生会话行为模式的可能性，多智能体可以引导学生从对竞争性说服向建构性对话转变，这有助于学生批判性思维、合作能力的提升，同时有助于协作知识的建构。第三，本研究发现多智能体的介入效果高度依赖提示语的有效性，未来的设计应结合提示语的设计提升多智能体反馈的质量。

4.1. 研究不足与展望

研究同样存在一定不足，首先，本研究人机互动仅基于文本对话数据，未能整合多模态信息，未来的研究可以通过设计多模态多智能体进一步探究人机协同的机制。第二，本研究实验周期较短，未能观测调节策略的长期效应，且实验对象为均为教育学背景，实验对象范围较小，需要进一步验证扩展到其他学科的可能性。最后，未来研究需要引入多模态的数据进一步分析，同时学生与多智能体的对话也应纳入分析数据中，以此更为全面地评估多智能体对协作学习的影响，并探索人机互动模式。

参考文献

翟雪松, 季爽, 焦丽珍, 朱强, & 王丽英. (2024). 基于多智能体的人机协同解决复杂学习问题实证研究. 开放教育研究, *30*(3), 63–73. https://doi.org/10.13966/j.cnki.kfjyyj.2024.03.007

刘石奇, 刘智, 段会敏, 粟柱, & 彭晛. (2025). 大模型驱动的教育多智能体系统应用研究——技术架构、发展现状、实践路径与未来展望. 远程教育杂志, *43*(1), 33–45. https://doi.org/10.15881/j.cnki.cn33-1304/g4.2025.01.004

吴林静, 陈梦凡, 高喻, 徐静, & 刘清堂. (2024). 面向调节学习支持的智能教学代理设计与实证研究. 电化教育研究, *45*(11), 53–60. https://doi.org/10.13811/j.cnki.eer.2024.11.008

徐振国, 刘志, 党同桐, & 孔玺. (2021). 教育智能体的发展历程、应用现状与未来展望. 电化教育研究, *42*(11), 20–26, 33. https://doi.org/10.13811/j.cnki.eer.2021.11.003

Berland, L. K., & Lee, V. R. (2012). In Pursuit of Consensus: Disagreement and legitimization during small-group argumentation. *International Journal of Science Education*, *34*(12), 1857–1882. https://doi.org/10.1080/09500693.2011.645086

Chen, X., Jia, B., Peng, X., Zhao, H., Yao, J., Wang, Z., & Zhu, S. (2025). Effects of ChatGPT and argument map(AM)-supported online argumentation on college students' critical thinking skills and perceptions. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13471-2

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*. https://www.tandfonline.com/doi/abs/10.1080/07370008.2016.1145120

Cho, K.-L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, *50*(3), 5–22. https://doi.org/10.1007/BF02505022

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G.-J. (2025). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies*, *30*(3), 3803–3827. https://doi.org/10.1007/s10639-024-12986-4

Guo, S., Latif, E., Zhou, Y., Huang, X., & Zhai, X. (2024). *Using Generative AI and Multi-Agents to Provide Automatic Feedback* (No. arXiv:2411.07407; 版 1). arXiv. https://doi.org/10.48550/arXiv.2411.07407

Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N. V., Wiest, O., & Zhang, X. (2024). *Large Language Model based Multi-Agents: A Survey of Progress and Challenges* (No. arXiv:2402.01680). arXiv. https://doi.org/10.48550/arXiv.2402.01680

Lee, U., Jeong, Y., Koh, J., Byun, G., Lee, Y., Hwang, Y., Kim, H., & Lim, C. (2024). Can ChatGPT be a debate partner? Developing ChatGPT-based application "DEBO" for debate education, findings and limitations. *Educational Technology & Society*, *27*(2), 321–346.

Li, Z., Jiang, G., Xie, H., Song, L., Lian, D., & Wei, Y. (2024). *Understanding and Patching Compositional Reasoning in LLMs* (No. arXiv:2402.14328). arXiv. https://doi.org/10.48550/arXiv.2402.14328

Linn, M. C., & Eylon, B.-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration*. Routledge. https://doi.org/10.4324/9780203806524

Liu, Q.-T., Liu, B.-W., & Lin, Y.-R. (2019). The influence of prior knowledge and collaborative online learning environment on students' argumentation in descriptive and theoretical scientific concept. *International Journal of Science Education*, *41*(2), 165–187. https://doi.org/10.1080/09500693.2018.1545100

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Su, Y., Liu, K., Lai, C., & Jin, T. (2021). The progression of collaborative argumentation among English learners: A qualitative study. *System*, *98*, 102471. https://doi.org/10.1016/j.system.2021.102471

Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, *3*(2), 258–262. https://doi.org/10.1515/jccall-2023-0008

从“局外人”到“小玩家”：幼儿在线媒体沉迷向具身游戏参与的转化引导与支持策略研究——一项长周期个案探索

From "Outsider" to "Little Player": A Case Study on the Transformation Path and Support Strategies for Redirecting Children's Online Media Addiction to Embodied Game Participation — An Exploration of the "Six-Step Correction Method"

黄灵娟[1]，庄周赟[1]，苗婷薇[2]，陈 赛[1]，逯 行[2*]
[1]浙江师范大学附属幼儿园，浙江 金华 321004
[2]浙江师范大学 教育学院，浙江 金华 321004
*1020352929@qq.com

【摘要】 数智时代幼儿面临媒体信息异化的风险，导致参与真实游戏活动减少。典型个案表现为沉迷数字媒体、丧失现实游戏兴趣、社会情感发展受阻等。通过对“满满”的长周期跟踪研究，构建了“六步修正法”，包括线上归因、游戏引导、社交支持、场景改进、强化合作和教师退出机制。干预后，幼儿从“游离的局外人”变为“小玩家”，恢复具身游戏能力。研究建议：建立以幼儿游戏能力提升为核心的“家长-教师-幼儿”情感体验和信任关系，呼吁家校社协同构建多主体支持策略。该研究为构建中国学前教育中游戏教育自主知识体系提供了实践探索和个案补充。
【关键词】 在线媒体沉迷；具身游戏参与；主动游戏；引导路径；支持策略

Abstract: *In the digital-intelligent age, children risk alienation from media information, reducing participation in real-world activities. Typical case shows online-media addiction, loss interest in real games, and hindered social-emotional development. Based on the analysis of "Manman", "Six-Step Correction Method"* was constructed*, including online attribution, game guidance, social support, scenario improvement, cooperation reinforcement and teacher withdrawal mechanism. After intervention, "disengaged outsiders" become "engaged players", regaining embodied gaming capabilities. The study suggests building an emotional and trust relationship centered on children's game ability improvement among parents, teachers and children, advocating collaborative efforts from families, schools and communities to create multi-agent support strategies. This research provides empirical exploration and case-based insights for constructing an autonomous knowledge system for play-based education in Chinese preschool education.*
Keywords: online media addiction**,** embodied game participation**,** active play**,** guidance pathway**,** support strategies

1.数字媒体异化下幼儿具身游戏的消逝与成长危机

数智时代幼儿的真实世界游戏参与能力衰减与社会交往能力减弱，是当代幼儿教养中存在的、需要急迫解决的问题。数智时代的幼儿是α世代的智能技术土著居民，他们从出生起，就与智能算法、大数据相伴。然而由于在日常教养中，家长将教育权过度转让给平板电脑、手机等智能终端，幼儿的信息获取来源大多来自社交媒体中未经审核筛选的内容。这造成幼儿的“社交媒体沉溺与内容追随”。典型代表性特征即幼儿对社交媒体中的游戏活动类型较为感兴趣，表现为模仿，然而这些游戏内容多为成人世界的“低配版本”，并非为儿童量身定制的能力发展型游戏。同时，这些行为在大数据算法的加持下不断定向推

送，媒体内容追随受到不断强化。长时间对社交媒体内容或游戏活动的注视，会削弱儿童对真实世界游戏能力的适应性，进而抑制数智时代幼儿社会交往能力的发展。
随着智能算法掌控下的社交媒体短视频向学龄前儿童渗透的加剧，“数字利维坦”造成的幼儿游戏能力丧失成为近年来教育研究者、家庭、学校共同关注的社会问题。其中以幼儿户外自主游戏、具身游戏等身体参与性游戏能力的丧失最为严重[10]。幼儿参与的自主型具身游戏是指幼儿处于自然生态的户外环境或经过人工改造的市内游戏活动空间里，依照自身兴趣和需求，自行选择游戏材料与伙伴、关注游戏学习的过程，并调动身体各个部分积极参与游戏进程，游戏过程中的幼儿在积极的情绪下通过操作材料、与环境互动，实现游戏中的学习与发展。受到幼儿发展阶段特征的限制，在游戏中总会存在部分“满天飞”“游离”“简单重复”的幼儿需要教师的进一步支持。我国颁布的《幼儿园保育教育质量评估指南》中“师幼互动”的关键指标将“教师观察、理解儿童”作为高质量师幼互动的重要参考。

基于此，本研究以个案研究的方式，开展对幼儿线上社交媒体追随向线下真实游戏参与的转化干预，以自我决定理论、社会建构主义理论和最近发展区理论为基础依据，从游戏兴趣着手，以“玩什么？”“和谁玩”“怎么玩”“梳理总结”的干预方式逐步推进，从而促进幼儿自主游戏水平提升，构建“线上追随归因、游戏行为引导、社交支持干预、游戏场景改进、社交合作能力强化、教师退出机制”六个修正步骤。本研究旨在形成面向数智时代的“幼儿线下社会交往与游戏能力修复与补偿方案”，丰富和完善中国学前教育自主知识体系构建。

2.典型个案探臻与理论模型建构

2.1. 代表性个案

户外自主游戏时间，所有小朋友如往常一样涌入精心布置的“自然庭院”游戏区。该区域投放了帐篷、野餐炊具、写生画板等开放式材料，旨在激发幼儿的社交协作与创造性游戏。观察发现，多数幼儿能快速组建游戏小组：有的用树枝和石子搭建“野外厨房”，有的围着帐篷演绎“森林探险”，还有的用仿真食物举办“生日宴会”。然而，女孩满满（化名）却始终“游离”于这些游戏场景之外。当老师与其沟通、询问其为何不参与游戏时，满满表示她喜欢美甲、化妆游戏，不喜欢幼儿园的游戏，并且能够模仿短视频中的美甲化妆动作，如先将甲片贴在手指上，再拿起玩具口红对着空气涂抹，最后举起双手对着阳光晃动，口中喃喃自语“宝宝们看！这个色号超显白哦”。当教师邀请她加入“烧烤派对”时，她摇头拒绝：“我要当美妆主播，和手机里的姐姐一样漂亮。”

通过家访与行为追踪发现，满满的“拟成人化游戏偏好”与其线上媒体使用习惯高度相关。母亲反馈，满满 3 岁起便常跟随家人刷短视频，尤其痴迷“儿童美妆博主”内容——这些视频中，同龄女孩穿着成人服饰，使用仿真化妆品完成“艾莎公主妆”“冰雪女王妆”等主题造型，并配以“点赞加关注”的程式化话术。此外，算法推送进一步强化了该行为。满满日均观看美妆类视频超 1.5 小时，甚至能复述“斩男色”“玻尿酸”等成人美业术语。这种“媒体驯化”使其线下游戏呈现显著异化，如游戏主题窄化，表现为拒绝参与动态协作类游戏，执着于静态的“化妆直播”独白；角色认知偏差，表现为将自我定位为“被观看的表演者”，如自称“主播小姐姐”等，要求同伴充当“观众”而非“合作者”。满满已经在真实世界的合作游戏中表现出严重的社交退缩加剧。

[10]

这种“屏幕依赖”与“真实游戏脱嵌”的困境，暴露了数智时代幼儿发展的典型矛盾，即线上内容的伪成人化与线下社交的去技能化，幼儿逐渐缺乏在真实世界中开展合作中的角色协商、资源共享等基础能力。例如，当教师假装自己是“顾客”，要求满满为其制作某种颜色的美甲时，她表现出强烈抗拒，并表示“客人不能自己选颜色！视频里都是主播说了算！”可见，算法推荐的单向度、表演性交互模式，已严重削弱其对真实游戏中双向互动规则的理解。

2.2. 幼儿自主游戏能力修正方案的理论建构

本研究构建促进幼儿线上社交媒体内容追随向真实世界游戏能力转化的“六步修正法”，包括线上追随归因探析、游戏场景改进、社交支持干预、游戏行为引导、合作实践能力强化、教师退出机制等六个核心策略。该策略框架以双重转化理论为核心框架（虚拟世界行为向现实能力转化、被动追随向主动创造转化），整合自我决定理论（SDT）的自主性-胜任力-归属感需求模型、社会建构主义理论的互动支架原则，以及动态系统理论的生态嵌套观，构建起“诊断-干预-迁移”的闭环式修正模型（见图 1）。其理论逻辑体现为以下三个递进层次：

图 1 促进幼儿线上关注向线下游戏能力转化的“六步修正法”

2.2.1. 线上追随归因：问题解构与幼儿需求溯源

基于认知发展理论与媒体依赖理论，剖析幼儿线上追随行为的心理机制。在设计个案干预与行为修复策略的设计时，通过“游戏兴趣四维访谈”（玩什么/和谁玩/怎么玩/为什么喜欢）识别幼儿对线上内容的表层模仿（如角色动作复刻）与深层需求（如审美、社交认同等内在渴望）。通过对幼儿追随的媒体进行分析，解码幼儿追随的线上游戏特质，例如高反馈性、强视觉刺激、低认知负荷等，映射其与线下游戏能力缺失的关联，推测幼儿合作意识薄弱、情境延展力不足、社交淡漠的形成机制，以作为后续干预策略设计的依据。以“满满”为例，发现其对短视频中“美甲化妆”等短视频或游戏的机械模仿，实则为低阶感觉寻求，追求即时感官刺激，并在过程中不自觉地形成了社交回避，表现为无法融入幼儿群体活动中的游戏角色分配。

2.2.2. 场景-行为-关系三重干预：生态化支持系统构建

本文系统整合生态心理学、社会认知理论与最近发展区理论，构建“场景-行为-关系”协同支持框架。游戏场景改进遵循“行为-环境适配”原则，创设“奇幻美学工坊”半结构化游戏区：保留线上审美元素，植入幼儿艺术创作台等协作装置，实现虚拟感官刺激向实体

合作的迁移。游戏行为引导基于观察学习机制，教师以“陪伴+导演”身份介入，通过引导幼儿参与团体自主游戏，将幼儿的单向表演转化为角色契约。社交支持干预部分则采用渐进式激活路径，不建立教师的逐步退出机制，推动幼儿的社交从外部驱动转向内生惯习。具体修策略和目标指向如表 1 所示。

表 1 促进幼儿线上内容追随向线下游戏能力提升的三维干预设计

修正维度	理论支撑	策略设计	目标指向
游戏场景改进	生态心理学“行为-环境适配”	创设“半结构化游戏区”； 保留线上游戏和社交媒体中的审美体验； 增加实体协作元素；	实现虚拟刺激向实体探索的自然迁移；
游戏行为引导	社会认知理论“观察学习”	教师扮演“游戏导演”； 通过陪伴追随时刻关注幼儿游戏状态； 提供**符号工具包**；	从机械模仿转向意义建构；
社交支持干预	维果茨基“最近发展区”	协助建立幼儿合作交往机制； 激发幼儿游戏参与中的社会交往意识； 教师及时退出；	破解社交退缩，建立合作初体验；

2.2.3. 能力强化与自主性孵化：从外部支持到内生动力

本阶段的干预逻辑在于将外部支持转化为幼儿的内生性游戏素养，其理论内核源于自我决定理论的自主性-胜任力-归属感三角模型，以及动态系统理论的生态适应性原则。研究通过“合作实践能力强化”与“教师退出机制”的双轨设计，推动幼儿从“策略依赖”转向“规则内化”，最终实现游戏参与的自主迭代。

其中，合作实践能力强化以群体社会化理论为基石，通过“游戏链任务”重构幼儿的协作认知。在此过程中，教师采用“矛盾暴露-解决循环”策略，当幼儿因争夺角色发生冲突时，教师不直接仲裁，而是引导幼儿回顾冲突发生过程，继而协商解决。此设计将线上短视频的“即时反馈”机制转化为现实中的问题解决训练，在具身游戏中逐步发展出责任分工意识与妥协弹性。

教师退出机制遵循渐隐式支架原理，通过“符号工具替代-环境暗示强化-自主反思生成”三阶段实现支持系统的隐性化。初期教师使用语言提示，中期转为非言语信号干预，末期引入游戏共同体回忆机制。

该阶段旨在将线上算法的个性化推送逻辑逆向解构，转化为线下幼儿具身游戏的自主生成动力。例如，短视频中“点赞量”对幼儿的外在激励，被重新编码为群体共享成就。这种数字行为再驯化的过程，不仅修复了被社交媒体割裂的游戏能力，更培育出数智时代儿童亟需的复合型素养——在保留媒介适应力的同时，重建对真实世界的深度参与。

3.促进幼儿“线上追随”向“具身游戏参与”转化的引导路径

幼儿天生具有设计、参与、改进游戏的能力。面对数字技术和新兴媒体对幼儿游戏能力的剥夺，教师需要在深度剖析问题成因的基础上，重塑幼儿基于真实世界开展具身游戏的能力。

3.1. “游离的局外人”：幼儿游戏参与问题的有效归因与场景改造

“游离的局外人”阶段师幼互动的个案笔记

观察到满满的这个状态后（图 1），在这一次的小庭院游戏时间，我主动找到她，蹲下来问：“满满，你没有喜欢的游戏吗？”“不喜欢，我觉得一个人走来走去还可以。”“那要不要去帐篷里一起玩？”满满摇摇头。“那烧烤摊呢？”“不喜欢。我觉得这些游戏都不好玩，没有手机上的视频好看，也没有平板电脑上的游戏好玩,不是我喜欢的！”“那你喜欢玩什么呢？”“我也不知道玩什么，我现在就想一个人。”“好的,那你先想想,想好了再去。”

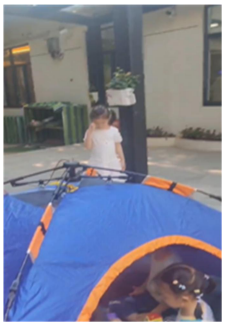


图 1　游离的满满

第一次的游戏介入以失败告终，满满依旧在走廊上毫无目的地走动，没有参与任何游戏。在后面的几次观察中，我发现她这种“游离”的状态仍在持续。

3.1.1. 需求察觉，接纳抱持

《纲要》提出“以关怀、接纳、尊重的态度与幼儿交往。耐心倾听，努力理解幼儿的想法与感受，支持、鼓励他们大胆探索与表达。”在游戏场中，笔者发现满满虽然在游离，但她整个状态平和，表情放松，这意味着她对当下状态是满意的。笔者以接纳的态度对待她，当她说出“不喜欢，我现在就想一个人”时，我尊重她的想法。我耐心倾听她的需求，就像《纲要》所倡导的那样，进而“放任”她自由地走遍小庭院的每个角落。

3.1.2. 追本溯源，探寻策略

当满满的“游离”状态从游戏蔓延到日常生活，且越来越频繁时，笔者有点担忧，所以决定和满满的妈妈聊一聊，看看“游离”背后是否有一些我不知道的原因。通过与满满妈妈的沟通，我得到了三个关键信息：其一满满在家中也经常出现游离的情况；其二满满在家喜欢玩一些与“美”有关的游戏，喜欢在社交媒体中观看化妆、美甲等短视频，并模仿视频进行美甲游戏、化妆游戏；其三满满经常回家表达自己在幼儿园没有朋友，不会和小朋友们相处，不知道如何与小朋友们一起玩耍，长此以往不想上幼儿园的负面情绪愈加严重。可以推断，这是数智时代幼儿沉迷社交网络信息后造成的社交能力弱化，进而造成的真实世界游戏能力衰退。此外，笔者再次分析自己与满满的对话，试图从中找到一些关键点（表 1）。

表 1 对话分析解读表

倾听与表达	**教师解读**	**行为关键词提炼**
我觉得这些游戏都不好玩，也不是我喜欢的。我也不知道想玩什么。	小庭院目前的游戏并不能吸引满满的注意，这些游戏材料并不是她喜欢的。同时，她对想玩的内容也没有明确的表示。	玩什么
我一个好朋友也没有，我不知道跟谁玩。	没有找到相对固定的游戏伙伴。	跟谁玩

3.1.3. 兴趣为基，场景创建

华爱华老师说：“游戏材料的提供，对幼儿起着游戏暗示的作用，刺激幼儿选择了某种游戏方式，表现出不同游戏行为，间接地对幼儿的发展产生作用。”由此可见，游戏材料的提供决定了幼儿与材料互动的质量，也决定了游戏的质量。通过满满妈妈提供的三个有效信息，并结合我日常生活中对满满的观察，以及上一次的对话解读，我猜测满满可能

会对美工区的材料更感兴趣。于是，我尝试在小庭院投放以下意图材料：粘土、毛根、彩纸、马克笔、竹签等常见美工区材料，希望以此来吸引满满的游戏兴趣。

3.2.“局外人变小玩家”：促进游戏参与的教师支持与行为引导

在充分了解满满对在线媒体中游戏场景的喜好之后，笔者在幼儿活动区域中投放了新的游戏材料。与此同时，鼓励周遭的小朋友们纷纷开启对新材料的探索，并带动满满加入进来。笔者观察到满满对这些新材料产生了兴趣，她会在小朋友游戏的时候，停留在她们的身边，还会参与其中，但是参与的时间较短。基于满满对新材料的兴趣，我在这次的自主游戏开始前，与满满共同设计游戏计划，通过增强满满的参与度，激发其游戏参与兴趣。

<table>
<tr><th colspan="2">“局外人变小玩家”阶段师幼互动的个案笔记</th></tr>
<tr><td>场景一：
师：你今天想做什么事情？
满：我想做手工。
师：你想和谁一起做？
满：一个人做。
师：你想做什么手工？
满：我唯一拿手好戏就是玫瑰花。
师：你准备做多少朵玫瑰花？
满：我要做 4 朵！
师：用什么材料做？
满：用浅绿色、深绿色。
师：浅绿色、深绿色做它的叶子是吗？
满：不是，做彩色，还有渐变色。
师：好期待哦，那你渐变色玫瑰花会愿意和其他小朋友分享吗？
满：不用了（摇头）。
师：我好奇，所以问一下。你是用什么材料？是用布做？还是用纸做？还是用粘土做？
满：用粘土做。
师:今天满满要用粘土做 4 朵渐变色的玫瑰花，我很期待。</td><td>
图 2 满满表示要做 4 朵玫瑰花</td></tr>
<tr><td>场景二：
游戏计划后，满满来到区域，坚持一个人站在桌子边用粘土做玫瑰花，边做边看周围同伴。完成了四朵玫瑰花骨朵的制作后满满发现做叶子和花茎的绿粘土不够，又转身去找，发现小葵手上有自己需要的材料，轻声问了一句：“你这个绿色能不能给我？”但是声音太小了，小葵没有听见，满满跟随着小葵再次发问：“能给我一点吗？”小葵看了一眼自己的绿色粘土，转身离开。满满愣了一下，转身继续做自己的玫瑰花。在一旁的早早看到这个过程，主动走过来问：“这个树枝你要吗？”满满从</td><td>
图 3 满满主动向同伴分享玫瑰花</td></tr>
</table>

他手中接过树枝，在玫瑰花上对比了一下，嘀咕着：“太粗了！”然后把树枝放下，继续做玫瑰花。20 分钟左右完成 4 朵玫瑰花的制作后，她向身边的恰恰展示自己的玫瑰花（图 3）：“这是我做的渐变色玫瑰。”恰恰笑着对她说：“你做的这个玫瑰花真好看。”满满露出了开心的笑容。	

3.2.1. 计划在前，助推游戏进行

根据维果茨基的“最近发展区”理论，可以看到孩子从低水平游戏向高水平游戏靠拢。在游戏开始前，我们进行了一个游戏计划：通过“玩什么”的提问来寻找满满的最近发展区；再用“和谁玩”这个提问来试图帮助她寻找游戏伙伴缩小“最近发展区”；紧接着用“怎么玩”的提问来帮助满满所需要的材料和方法；最后通过梳理总结的方式来鼓励满满落实游戏计划，投入自主游戏中（见表 2）。

在这次游戏计划中，我以朋友的方式和满满进行沟通，在宽松、自主的环境氛围下，不仅能了解孩子的想法与兴趣，更能助推幼儿以自己的方式游戏，有助于幼儿创造自我和实现自我。

表 2 第一次游戏计划

“最近发展区”理论	**计划关键词**	**教师提问**	**幼儿回复（满满）**
寻找“最近发展区”	玩什么？	你想做什么？	我想做手工
缩小“最近发展区”	和谁玩？	你想和谁一起做？	一个人做
形成“最近发展区”	怎么玩？	你想做什么手工？	我唯一拿手好戏就是玫瑰花。
		我很好奇你的玫瑰花是怎么做的？会用到哪些材料？准备做多少？	我用黏土做，要做渐变色的玫瑰花。我要做四朵。
幼儿自主发展（创造自我、实现自）	梳理总结	哇，我很期待你今天要做的四朵渐变色玫瑰花。	

3.2.2. 陪伴追随，关注游戏状态

《评估指南》指出，通过持续观察客观分析幼儿发展情况与需求，提供针对性支持，不可急于介入或干扰幼儿活动。在满满游戏期间，我始终追随着她，默默观察她的游戏状态，她给我的一次次惊喜也再次证明：只有放手游戏，才能发现儿童！当满满因为材料不足停下动作时，我担心她会放弃玫瑰花的制作，令我惊喜的是她会主动向小葵借材料；当两次询问都被忽视后，我担心她会难过、沮丧，令我惊喜的是她情绪一直保持稳定；当我在犹豫是否要介入帮她一起询问其他孩子时，令我惊喜的是她已经用自己的方式完成了作品；当我再次惊喜地看到她主动向同伴展示成果，得到小伙伴的肯定与赞赏时，望着她脸上满是发自内心的喜悦与自信，我的内心也充盈着快乐与满足。我很庆幸在自己不介入、不干扰的追随式陪伴中，有机会觉察到满满是个有能力的独立的学习者；持续地观察发现，让笔者惊叹于一直被爱与尊重包容的满满，内心世界是多么丰富且强大。

3.2.3. 基于需求，丰富游戏材料

此次游戏后，满满的心情犹如盛开的花朵一般，满是愉悦，在游戏过程中，她体验到了自主探索和创造的乐趣。于是，教师趁热打铁，继续与满满展开深入的交流，一同探讨

在下一阶段的游戏里她可能需要补充的材料。满满这次明确地提出：接下来她想要当一名化妆师，所以她想要的材料是钻石贴纸、刷子、指甲贴等。基于其个性化需求，教师补充了她所需要的材料，又增加了盘子、瓶盖、纱、棉球等，希望这些补充的材料能够如同助力的羽翼一般，为满满的游戏插上创造想象的翅膀。

3.3. “小玩家开心游戏”：社交合作能力强化与教师退出机制

在上一次玫瑰花的制作中，满满有了三次和同伴互动的机会，第一次主动寻求材料，结果被拒绝；第二次面对同伴提供的材料支持时有了初步的回应；第三次是当满满完成玫瑰花的制作后，主动向同伴展示自己的作品，并且得到了正向的反馈。这三个瞬间让我看到了满满开始出现了和同伴互动的需求，我认为这正是推进满满同伴交往的好时机，因此我希望通过这一次的游戏计划，给满满创造一个和同伴共同游戏的机会，在游戏中促进满满的社会性发展。因此，我再次运用“玩什么？”“和谁玩？”“怎么玩？”“梳理总结”四个步骤和满满进行了游戏计划，并把计划重点放在“和谁玩”的拓展上。

“小玩家开心游戏”阶段师幼互动的个案笔记

场景一：

1 个人的计划

师：我今天有彩色的贴纸，你上次提出想要的，那你用贴纸想做什么？

满：给大家化妆。

师：你想给谁化妆？

满：不知道。

师：男孩子还是女孩子？

满：女孩子？

师：女孩子有团团、苗苗、甜甜、媛媛、妞妞，你想给哪个女孩子化妆呀？

满：不知道。

师：甜甜、圆圆、妞妞、团团，你这几个选一个呗。

满：……

师：那甜甜、妞妞两个人里面你选谁呀？

满：选妞妞。

师：好，那我们去把妞妞请过来，问一下她愿不愿意和你一起玩这个化妆的游戏好吗？走！（满满和教师一起去邀请妞妞）。

场景二：

3 个人计划（图 3）

师：满满想要和你们一起玩化妆游戏，你们觉得需要做哪些准备呢？

妞：要准备口红、贴睫毛，还要准备项链、还要贴美甲的东西，还要头巾。

团：还要准备涂指甲的东西。

师：现在我们有钻石,贴指甲的东西,其他东西没有，怎么办呢？可以有什么方法替代一下吗？

团：用粘土做。

妞：粘不上去，口红要涂在嘴巴上的。

师：粘土涂在嘴巴上确实不太干净，还有其他想做的吗？

图 4 三个人的游戏计划

团：贴面膜。
师：那谁来化，谁当顾客谁来化妆呢？
满：我。
师：好，满满你当化妆师，团团和妞妞是顾客。期待你们美美的妆容哦。
满满分别给妞妞、团团眼角、脸颊部位贴上贴纸，在贴的过程中，妞妞说：“我想要在脸上贴一点，这样好看一点。”“好的，那我给你用紫色的贴。”“那额头也贴一点。”“可以。”在顾客的要求下，满满的化妆也变得更加仔细与认真。她用妞妞找来的瓶盖做眼影，手指蘸一下，在妞妞脸上点一点。团团用吸管替代粉扑在脸上滑动。过程中，满满每一次无意间与我的眼神对视，我都会用“微笑”“大拇指”给予积极回应。随着游戏的持续，满满的笑容越来越多，化妆游戏也吸引了甜甜等更多女孩的参与，化妆游戏一直持续玩了 30 分钟左右。

图 5 满满开心地与同伴互动

化妆游戏结束后，满满在她的“游戏故事”（图 6）中记录了自己的游戏过程，并在我的鼓励下大方分享：今天我和团团、妞妞一起玩了化妆游戏，我真的很开心。同时，我也及时追问妞妞和团团的游戏心情，再得知两位小伙伴对今天的游戏也十分满足且愉悦时，满满的脸上又一次出现了满足的笑容。
游戏进行至此并没有拉上帷幕，满满的游戏故事似乎从这刚开始，在接下来的自主游戏中满满不再“游离”，她有了自己的游戏伙伴，在“变美”的游戏中有了更多的想法和体验。此时的我认为满满已经能享受自主游戏的欢乐时光，而作为教师的我是时候“退出”了，我选择继续用“观察者”的身份待在满满身边，关注着她的后续发展。

图 6 满满的“游戏故事”

3.3.1. 因人而动，三方合作

建构主义学习理论强调学习的共同建构，认为知识主要通过学习者和教师的共同参与建构。另外，在社会交互和合作中的过程，个体儿童与他人（同伴、教师、家长）共同构

建和协商意义，通过与各类人员的互动和对话来共同解决问题和建立新的知识体系或理念。在本次的游戏计划中，满满主动提出了“给大家化妆”这一想法，我及时抓住了契机继续追问,从“你想给谁化？”“男孩还是女孩？”“哪一个女孩？”再到从多个女孩的选择到两个女孩之间的选择，我都有充分的思考：谈话范围从全班到聚焦两个孩子,给到满满选择的权利,最后推荐了妞妞和团团两位社会性最强的孩子，确保满满游戏的互动能够顺畅。

社会建构主义理论还特别强调教师和更有经验的同伴在儿童成长过程中的作用，尤其是引导和支持的重要性。在幼儿园，教师和同伴发挥指导、鼓励和挑战的作用，促进幼儿的思维发展和社会性发展。教师通过提出适当的问题、激发幼儿思考和讨论以及提供积极且及时的反馈，从而促进幼儿的经验建构和个性发展。对满满的不断追问，促使她不断思考：我可以做什么？我和谁一起做？从而使她在游戏中的游离状态变为主动参与状态，构建属于自己的游戏场景。因人而动，根据不同的儿童、不同发展区、个性化、个别化的情况，构建不同的“计划”平台，以不同的互动方式，达成三方合作，助力满满的游戏发展。

3.3.2. 积极回应，加深信任

自我决定理论（Self-determination theory）提出人的内部动机具有三种基本心理需要，即自主需要（Autonomy）、胜任需要（Competence）与关系需要（Relatedness）。自主需要是指体验选择并感觉自己的行动像个首创者，达成这种需要时，人们往往会产生自我掌控的满足与喜悦。在这一阶段的游戏中，满满首创了化妆店，吸引了其他孩子来与之互动，产生了游戏的需求与满足。同时这一过程中，教师提供的支持策略渐进式提升了幼儿的自我效能感，进而促进其对挑战性任务的尝试。幼儿的胜任需要是幼儿相信自己能够胜任，并在富有挑战性的任务上取得成功，当幼儿的胜任需要在真实世界的游戏中得到满足时，其游戏体验会远远超过观看在线短视频、手机虚拟游戏等获得的成就感。

当满满在今天的游戏故事中表达了自己的开心体验，作为教师，我觉得此时最好的回应是“共享”她的这份喜悦，并鼓励她的伙伴对此做出回应。能够胜任游戏的反馈让满满的游戏“快乐指数”翻倍，也让满满的游戏体验再次升级，体会到与同伴共同游戏的快乐，满足了胜任需要能实现个体长期的健康发展。在满满游戏能力强化的过程中，笔者着重满足满满的关系需要，即满满希望与他人建立真实世界情感联系的愿望。这种爱和关心，以及被爱和被关心，是幼儿发展社会交往能力、理解和改造真实世界的基本能力，通常建立在与别人相互尊重和依赖的基础上。在本案例中，和满满进行的每一次游戏计划时，我都保持和她的眼神的接触，亲切地交谈，主动拉近距离。在游戏过程中的每一次对视，我也运用面部表情和身体语言对她做出积极的回应。通过游戏时间的多次互动，持续地关注，回应满满发出的信号，让她感到安全和安心，这让我和满满之间的信任持续加深。通过满足个体自主、胜任与关系三大基本心理需要，能够有效推动个体外部成长动机的形成。

3.3.3. 有效介入，及时退出

何时教师可以等待？何时教师必须出手？在游戏过程中，我也时常反思，通过一次次的游戏观察和实践，我认为“游离型”“满天飞型”“游戏简单重复型”的三类幼儿是需要教师关注与支持的。通过“游戏计划”和“意图材料”的有效介入，帮助幼儿持续深入的游戏。在幼儿能够专注快乐的游戏时，教师应敏锐地捕捉到这一信号，及时且巧妙地退出，从而完满地还给幼儿属于他们自己的游戏空间，教师的及时退出，也是对幼儿自主能力的尊重。我认为在游戏过程中，教师的角色是多变的，是观察者、引导者、倾听者，也是幼儿的游戏伙伴，无论是哪一种角色，教师都需要用爱与专业去陪伴幼儿的成长。

4. 如何帮助数字化生存中的幼儿返回现实游戏世界？

4.1. 促进以幼儿游戏能力提升为核心的“家长-教师-幼儿”情感体验

从幼儿本身的表现来看，满满在和大家一起玩之后，主动交往次数增多了，在游戏中笑容满满，在游戏表征中的“我觉得很开心”是她游戏情感表达的最直接体现。喜欢集体游戏，与同伴的互动游戏，她的游戏成就感大大提高。她从一个游戏“小迷糊”变成了一个游戏小玩家。此外，家长对问题严重性的后知后觉到积极配合，也是幼儿数字媒体异化问题解决的重要支撑。游戏过后，笔者主动地找到满满妈妈，分享了游戏视频和助推策略，满满妈妈看后十分感动。同时，她也经常向笔者反馈，满满回家说她拥有了哪些好朋友，今天又和谁一起玩了，成为了好朋友。满满妈妈在向老师表达感谢地同时，也认识到积极配合教师开展家园共育的重要性，以及学习幼儿教育专业知识的重要性。家长表示在今后将着重关注帮助满满建立友谊、积极开展真实世界的幼儿游戏。对于教师本身，解决幼儿数字媒体异化问题本身就是教育能力提升的过程。在看到“游离”的满满因为笔者的关注和帮助，成为“忙忙碌碌”的满满，而且这个“忙碌”的身影不仅出现在游戏中，也同样出现在日常生活中，她会主动地分享自己的想法，也愿意拉着伙伴一起去阅读、洗手、用餐，她从“自我”的小世界走出，走向班级这个群体，成为更加鲜活的满满。这一刻，幼儿、家长和教师之间是“三向奔赴”的，教育在此刻形成闭环。满满的进步、满满妈妈的正向反馈，让我“教师”的身份具象化了。

4.2. 建立充满信任、柔和渐进的师幼关系是幼儿异化行为修正的基础

德国哲学家雅斯贝尔斯说过,教育就是一棵树摇动另一棵树,一朵云推动另一朵云,一个灵魂唤醒另一个灵魂.在教育工作中,我越来越深刻地体会到:要了解我想推动的那一棵树,那一朵云的灵魂,就需要我停下脚步、蹲下身子、静下心来，听一听这棵树到底说了什么，那朵云到底为什么这么想。在一次次细致地观察、一场场真诚地对话中，我与幼儿之间建立了一种深厚的情感连接，这种连接让我感受到了教育的真正意义——不是单向的传授，从“我”到“我们”，更是双向的唤醒与成长。

4.3. 呼吁家校社协同构建抵抗幼儿技术异化的多主体支持策略

本研究中个案的分析、干预和行为修复是数智时代诸多幼儿、儿童乃至青少年身上存在问题的缩影，这既是时代之过、技术异化之过，亦是成人世界对幼儿成长空间的侵蚀和剥夺。以个案中提及的“拟成人化游戏”为例，正在解构当代幼儿本该通过同伴游戏建构的社会认知雏形。正如满满在访谈中透露：“手机里的姐姐有十万粉丝，我也要变网红，幼儿园小朋友不算粉丝”。这种过早将自我价值锚定于虚拟认同指标的倾向，寻求点赞数、粉丝量，使其对现实社交中的情感联结极度钝感，难以在真实世界的游戏中获得合作完成任务时候的成就感。针对这种线上世界的“即时满足”逐渐吞噬幼儿参与真实游戏的技术风险问题，不仅需要幼儿园教师的干预和引导，更需要家庭、社会、各类技术服务主体、政府监管部门的协同参与，共同守护幼儿成长的信息空间。

参考文献

张宇,周成玉(2021)。幼儿游戏中教师无效干预问题的瓶颈及突破。东北师大学报,(06):163-168.

马洁(2020)。幼儿户外混龄游戏教师指导研究。上海教育科研,(05):93-96.

汪彩萍(2020)。幼儿园大班幼儿自主性游戏的观察、支持与引导实践研究。基础教育课程,(10):27-33。

李相禹,刘焱(2018)。师幼比对教师观察与指导幼儿区域游戏的影响。教师教育研究,30(03):53-58.
万中,刘敏(2013)。幼儿游戏中教师的干预与干涉。学前教育研究,(08):56-60.
华桦(2004)。游戏中教师指导对幼儿社会性行为的影响。学前教育研究, (03):26-29.
雷有光,冯美婷(2020)。学习品质视域下的学龄前儿童教育游戏设计与开发研究。中国电化教育, (02):126-133.
凌辉,刘慧玥,林红,等(2024)。角色游戏促进中班幼儿社会自立-安全常识发展的个案研究。心理学探新 ,44(04):347-356.
凌辉,李海斌,甘义,等(2023)。依恋游戏促进小班幼儿社会自立-社交行为发展的个案研究。中国临床心理学杂志,31(05),1275-1280+ 1175.
凌辉,宁柳,刘淼婷,等(2023)。亲子游戏促进幼儿同伴交往能力发展的个案研究。中国临床心理学杂志,2023,31(02):491-495.
聂洋溢(2022)。幼儿园课程公平的现实状况与优化策略。学前教育研究, (02):61-78.
郝少毅,朱志勇.师范生情感知识及其情感教育的个案研究。教师教育研究,2021,33(06): 16-23.
王海澜(2021)。幼儿哲学探究的个案追踪研究。教育学报, 17(05):62-72.
温燕(2021)。以幼儿发展评价助推教师专业自主发展。学前教育研究,(10):87-90.
雷有光,冯美婷(2020)。学习品质视域下的学龄前儿童教育游戏设计与开发研究。中国电化教育,(02):126-133.
张丽敏,叶平枝(2019)。“我是谁?何以为师?”:社会互动中的幼儿园教师身份构建——一项多案例个案研究。全球教育展望, 48(12):69-85.

Research on the Design of Multi-agent Education Architecture for Innovative Thinking

Xinquan Jin [1], Yingtong Chen[1]
[1] Jiangnan University Jiangsu "Internet Plus Education" Base
* 1158230107@stu.jiangnan.edu.cn

Abstract: *This study focuses on the design of a multi-agent educational framework aimed at fostering innovative thinking. With the advancement of educational digital transformation, multi-agent systems, leveraging their role diversity and dynamic interaction capabilities, offer new perspectives for innovative education. The research proposes a multi-agent architecture model that includes a creativity conflict engine, cross-domain knowledge integration module, and innovation path visualization tools, and suggests corresponding multi-agent-driven educational practice pathways. This framework stimulates students' dialectical thinking through role opposition and collaboration, integrates interdisciplinary knowledge to aid in idea optimization, and uses visualization tools to systematically present and evaluate innovative outcomes. The study not only promotes educational model innovation but also provides theoretical and practical guidance for the application of multi-agent technology in the field of education, holding significant theoretical and practical importance.*
Keywords: Multi-agent Systems; Innovative Thinking; Educational Framework; Practice Pathway

1. Introduction

In the context of the rapidly advancing technology, innovative capability has become the core driving force for social progress and economic development. Education, as a key field for nurturing future talents, is undergoing a paradigm shift from traditional knowledge dissemination to the cultivation of innovative capabilities. However, traditional educational models often focus on knowledge implantation and memorization, struggling to meet the demands of complex thinking training. Therefore, exploring an educational model that can effectively cultivate innovative thinking has become a significant challenge that the field of education urgently needs to address. Digital education, as a new educational paradigm, can provide students with a more personalized, convenient, and efficient learning experience (Wu et al., 2024), Online platforms can foster environments that support independent learning and enhance engagement behaviors, and may even stimulate creativity (Jin et al., 2022),thereby better stimulating students innovative thinking. The Ministry of Education has clearly mentioned in its 2022 Work Priorities he need to implement strategic actions for educational digitalization, promoting the digital transformation and intelligent upgrading of education. In recent years, multi-agent architectures, with their role diversity, dynamic interaction capabilities, and strong tool invocation capabilities, have provided new ideas and methods for innovative thinking education. In recent years, multi-agent architectures, with their role diversity, dynamic interaction capabilities, and powerful tool invocation capabilities, have provided new approaches and methods for innovative thinking education. Multi-agent systems (MAS) consist of multiple autonomous agents that can interact and collaborate to accomplish complex tasks (Janbi et al., 2023). Multi-agent systems have strong dynamic interaction mechanisms, and learning environments with vitality that can encourage learners to engage in intellectual collisions with agents (Zhai et al., 2024), thereby promoting the development of students' innovative thinking. Furthermore, the integration of generative artificial intelligence (AIGC) technology with multi-agent technology (such as MetaGPT, AgentVerse) has further propelled the shift in educational agency from "teacher-led"

to "human-machine co-evolution," providing a richer set of tools and environments for the cultivation of innovative thinking.

However, in this educational transformation process, there are also many challenges and pressing issues that need to be addressed. First, multi-agent architectures can stimulate students' innovative thinking through role conflict and collaboration, and the rational design of the opposition and collaboration between roles is key to achieving this goal. Second, designing dynamic interaction mechanisms and tool invocation strategies is the core of supporting the innovation process. The cultivation of innovative thinking requires a flexible and dynamic environment, making the design of interaction mechanisms and tool invocation strategies among agents crucial. Furthermore, evaluating the long-term impact of multi-agent systems on students' innovative abilities is an important step in ensuring educational effectiveness. The development of innovative thinking is a long-term process, and how to measure the impact of multi-agent systems on students' innovative abilities through scientific assessment methods is also an issue that cannot be overlooked in current research.

This study aims to explore the design of a multi-agent educational framework oriented towards fostering innovative thinking, analyze its application advantages in education, and propose a multi-agent architecture model that can effectively stimulate students' innovative thinking. By conducting an in-depth study of the types of applications of multi-agent systems in education, a comprehensive multi-agent educational framework design scheme and specific multi-agent-driven teaching practice strategies are proposed. The research not only contributes to the innovation of educational models but also provides theoretical support and practical guidance for the application of multi-agent technology in the field of education, holding significant theoretical and practical importance.

2. Literature Review

2.1. Innovative Thinking

Innovative thinking, also known as creative thinking, refers to the cognitive ability to identify problems, find solutions, generate novel ideas, and ultimately communicate outcomes (Dong & Chen, 2024). The evaluation dimensions of innovative thinking encompass critical thinking skills, interdisciplinary knowledge integration capabilities, social interaction and collaboration abilities, practical application skills, as well as self-reflection and self-regulation abilities, all aimed at comprehensively promoting the cultivation and development of students' innovative potential (Wang et al., 2024). The cultivation of innovative thinking focuses on encouraging students to diligently learn scientific foundational knowledge to establish a solid foundation, integrating innovative thinking with curriculum teaching to achieve a complementarity of thinking and capabilities, strengthening hands-on ability cultivation to exercise practical skills, reforming teaching methods to reflect creative instruction, creating an innovative atmosphere to stimulate creative potential, and enhancing innovative thinking capabilities through strengthened thinking training activities (Yan et al., 2012). Current issues in the cultivation of innovative thinking lie in the traditional educational model's difficulty in adapting to students' personalized learning needs, a lack of effective mechanisms to promote interdisciplinary knowledge integration, and insufficient opportunities for practice and application to exercise and test students' innovative abilities, leading to students' difficulty in forming interdisciplinary thinking and a lack of ability to solve complex problems (Zhou & Zhao, 2024). The application of multi-agent systems in education precisely meets the needs for personalized learning path customization, interdisciplinary knowledge integration, and practical skill training.

2.2. Definition and Characteristics of Multi-Agent Systems

An agent refers to a computer system capable of perceiving its environment and taking actions to achieve specific goals, possessing perceptual capabilities, action capabilities, goal orientation, autonomy, and environmental interaction capabilities (Stuart J. et al., 2011). A multi-agent system is a collection of several autonomous agents that can interact and collaborate with each other to accomplish complex tasks. In a multi-agent system, each agent is an independent computational entity with the ability to perceive the environment, make inferences, decide, and take actions (Yuan et al., 2025).

Multi-agent systems demonstrate strong problem-solving capabilities and adaptability through collaboration and interaction among agents. These agents can work together through role division, information sharing, and dynamic adjustment to complete complex tasks, providing new solutions for the field of education. The collaborative mechanism of multi-agent systems is mainly reflected in the role distribution and task collaboration between agents. In educational scenarios, agents can undertake different roles according to task requirements, such as teachers, learners, experts, etc., and provide multidimensional solutions through collaborative work (Yu et al., 2024). The interaction patterns of multi-agent systems mainly include static and dynamic types. In the static interaction pattern, the connections and interaction relationships between agents are fixed and unchanging, suitable for scenarios with relatively clear task structures. The dynamic interaction pattern allows agents to dynamically adjust interaction strategies based on real-time feedback, more suitable for complex and variable learning environments (Zhai et al., 2024). Additionally, multi-agent systems can further enhance their interaction capabilities and knowledge update speed through tool invocation and integration with external data sources, ensuring the timeliness and accuracy of knowledge, and further improving students' innovative abilities. For example, development frameworks like LangChain use knowledge distillation techniques to integrate knowledge from various sources into large language models, thereby providing agents with more accurate and rich information (Wu et al., 2024).

2.3. The types of applications of multi-agent systems in education

The applications of multi-agent systems in education are primarily divided into three types: instructional support, collaborative learning, and assessment feedback. Instructional support multi-agent systems demonstrate significant potential in teaching assistance, mainly reflected in course design and development, educational software development, and teaching scenario construction (Liu et al., 2025). These multi-agent systems may include multiple roles such as teachers, assistants, and administrators, providing comprehensive support across teaching, order maintenance, and learning analysis through collaborative coordination. Collaborative learning multi-agent systems can simulate real learning scenarios to provide personalized learning resources, promote collaboration among students, stimulate their sense of cooperation and team spirit, thereby helping students complete learning tasks and projects more efficiently, enhancing learning efficiency (Wu et al., 2024). Assessment feedback type multi-agent systems are mainly applied in the fields of test question generation, homework grading, and learning diagnosis, leveraging the collaboration of multiple agents to achieve more objective, comprehensive, and personalized educational assessments (Liu et al., 2025).

3. Multi-intelligence architecture design that supports innovative thinking

The multi-intelligence architecture design proposed in this study aims to comprehensively support the cultivation of innovative thinking through a layered mechanism, specifically including three layers: creative stimulation, iterative optimization, and achievement transformation. The architecture stimulates students' innovative thinking, optimizes the creative process, and realizes the systematic presentation and evaluation of innovative outcomes through

role opposition and collaboration, cross-disciplinary knowledge integration, and the application of visualization tools, as shown in Figure 1.

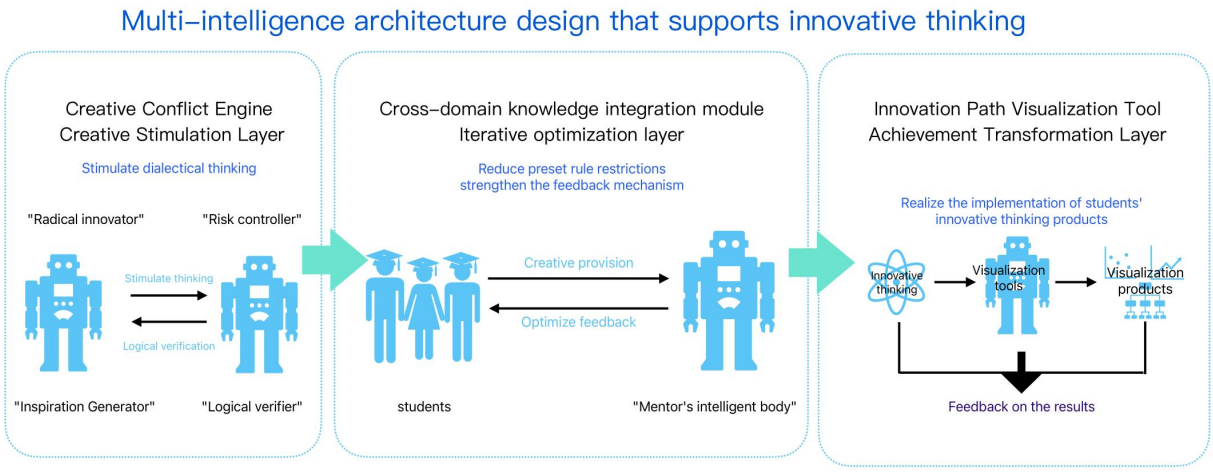


Fig. 1 Multi-intelligence Architecture Design Diagram

3.1. Creative Conflict Engine (Creative Stimulation Layer)

In the process of fostering innovative thinking, it is necessary to stimulate students' dialectical thinking abilities. To this end, the architecture includes a Creative Conflict Engine designed to create intellectual friction through opposing role agents, such as the "Radical Innovator" and the "Risk Controller." During the students' learning process, these agents generate cognitive collisions that encourage students to think dialectically about current issues, thereby promoting the emergence and development of innovative thinking. Additionally, within the horizontal architecture, a variety of role agents are deployed, such as the "Inspiration Generator" and the "Logical Verifier." By simulating free discussions akin to team collaboration, initial ideas are generated. The "Inspiration Generator" agent can offer novel creative solutions, while the "Logical Verifier" agent is responsible for logically analyzing and verifying the feasibility of these proposals. Together, they assist students in expanding their thought space, breaking away from established thinking patterns, and stimulating their innovative potential. This balance of conflict and collaboration can effectively ignite students' innovative thinking.

In specific educational scenarios, the Creative Conflict Engine is primarily applied to positively stimulate students' innovative ideas. For instance, in a classroom activity based on group collaboration and division of labor, the Creative Conflict Engine can equip each group with a system of agents, including the “Radical Innovator,” “Risk Controller,” “Inspiration Generator,” and “Logical Verifier.” The “Radical Innovator” and “Inspiration Generator” generate initial creative solutions based on the teacher's requirements and students' ideas. Students then diverge their thinking based on the initial solutions provided by the agents, and each group engages in human–machine collaborative dialogues. As the interaction with the agents deepens, multi-perspective arguments on the same topic inspire students to engage in independent thinking based on the initial schema. There are varying degrees of improvement in dimensions such as viewpoint expression and argument analysis (Jin et al., 2023). In this process of continuous in-depth thinking, each group produces personalized innovative solutions. These solutions are then submitted to the “Risk Controller” and “Logical Verifier” agent systems for feasibility analysis to validate the practical significance of the solutions generated by each group, thereby continuously improving the group's solutions. The application of the Creative Conflict

Engine can enhance the efficiency of the entire classroom by initiating different dialogues based on the ideas of each group and providing more personalized learning guidance, helping students better diverge their thinking and exercise their innovative thinking abilities.

3.2. Cross-domain Knowledge Integration Module (Iterative Optimization Layer)

In the process of fostering innovative thinking, the integration and iterative optimization of interdisciplinary knowledge are essential components. To this end, our architecture has designed a dedicated cross-domain knowledge integration module aimed at breaking down traditional barriers between disciplines to promote the comprehensive application of knowledge and the development of innovative thinking. This module reduces reliance on preset rules and enhances real-time feedback mechanisms, such as incorporating Reflexion's self-assessment loop, to improve the performance of agents when executing complex tasks like decision-making, reasoning, and programming.Agents are capable of dynamically adjusting their behavioral strategies based on real-time feedback and can also continuously optimize their approach to problem-solving through self-assessment and reflection, which is beneficial when facing novel or unconventional problems. This allows agents to flexibly adjust their methods to adapt to the ever-changing demands.Within the vertical architecture, the mentor agent plays a crucial role, responsible for integrating ideas from various fields and calling upon appropriate tools, such as prototyping software, for rapid concept validation. The mentor agent can generate preliminary prototype designs based on students' ideas and use real-time feedback mechanisms to help students iterate and optimize their ideas, thereby transforming ideas into practical and feasible solutions.Moreover, the cross-domain knowledge integration module also encourages students to examine problems from multiple perspectives and seek innovative solutions through an interdisciplinary viewpoint. This interdisciplinary way of thinking helps students to propose more innovative and practical strategies when facing complex real-world problems.

For example, in an interdisciplinary innovation practice course involving multiple subject areas, a student team proposed a creative solution at the beginning of the project but was puzzled about how to integrate knowledge from different disciplines. At this point, leveraging the deep integration of artificial intelligence and education, which changes the established knowledge increment system and its generation mode (Jin et al., 2025), the "**Mentor Agent**" can intervene. Based on the students' ideas, the "**Mentor Agent**" can call upon a multi-party information knowledge base to generate a preliminary system architecture diagram. Subsequently, the agent utilizes Reflexion's self-assessment loop mechanism to evaluate the preliminary design, identifying potential problems and areas for optimization in the initial solution. Based on these evaluation results, the agent provides students with suggestions for improvement, guiding the student team to optimize the solution design from different disciplinary perspectives. The student team can then adjust their initial solution according to the agent's feedback to enhance the feasibility of the solution. Through this process of cross-disciplinary knowledge integration and iterative optimization, students can effectively improve their interdisciplinary collaboration and problem-solving abilities.

3.3. Cross-domain Knowledge Integration Module (Iterative Optimization Layer)

To implement students' innovative thinking products, this architecture has set up an Achievement Transformation Layer. In the Achievement Transformation Layer, innovative thinking products can be realized through the Innovation Path Visualization Tool, which can transform students' thinking processes into dynamic graphs, similar to the graph structure of PLAG, clearly presenting the process from the germination of ideas to the formation of solutions. This visualization method not only provides students with an intuitive basis for reflection and

adjustment but also facilitates targeted guidance from teachers. At the same time, the dynamic team collaboration model refers to the DyLAN contribution assessment mechanism, dynamically adjusting roles and tasks according to members' contributions to ensure efficient team collaboration and the joint completion of the final solution. In addition, through structured output formats, such as documents and charts generated by MetaGPT, innovative achievements are systematically presented, which is not only convenient for students to review and summarize but also provides a strong handle for teachers to assess and guide, achieving closed-loop management in the cultivation of innovative thinking.

In the teaching scenario of an innovative design course, the application of the Achievement Transformation Layer provides students with a systematic path for the cultivation of innovative thinking. Through the Innovation Path Visualization Tool, students' thinking processes are transformed into dynamic graph structures, similar to the graph structure of PLaG, clearly presenting the entire process from the germination of ideas to the formation of solutions. This visualization method not only provides students with an intuitive basis for reflection and adjustment but also facilitates targeted guidance from teachers. Meanwhile, the dynamic team collaboration model, which refers to the DyLAN contribution assessment mechanism, dynamically adjusts roles and tasks according to members' contributions to ensure efficient team collaboration and the joint completion of the final solution. Research has shown that learning activities are generally accompanied by engagement, and students who actively participate in online peer feedback can provide specific and detailed feedback that provokes multiple perspectives and in-depth reflection on the creative content and process, potentially enhancing their innovative thinking abilities (Jin et al., 2024). In addition, structured output formats, such as documents and charts generated by MetaGPT, systematically present innovative achievements. This not only facilitates students' review and summary but also provides a strong handle for teachers to assess and guide, achieving closed-loop management in the cultivation of innovative thinking.

4. Practice Pathway

To achieve effective practice of the multi-intelligence educational architecture, this study proposes the following multi-agent-driven teaching practice strategies, including five stages: role division substitution, adaptive transformation, reorganization symbiosis, collaborative innovation, and iterative optimization, see Figure 2. This strategy not only optimizes the task execution process but also ensures the high-quality completion of tasks through dynamic collaboration among agents. The following are the specific application paths of the multi-agent-driven teaching practice strategy in this study:

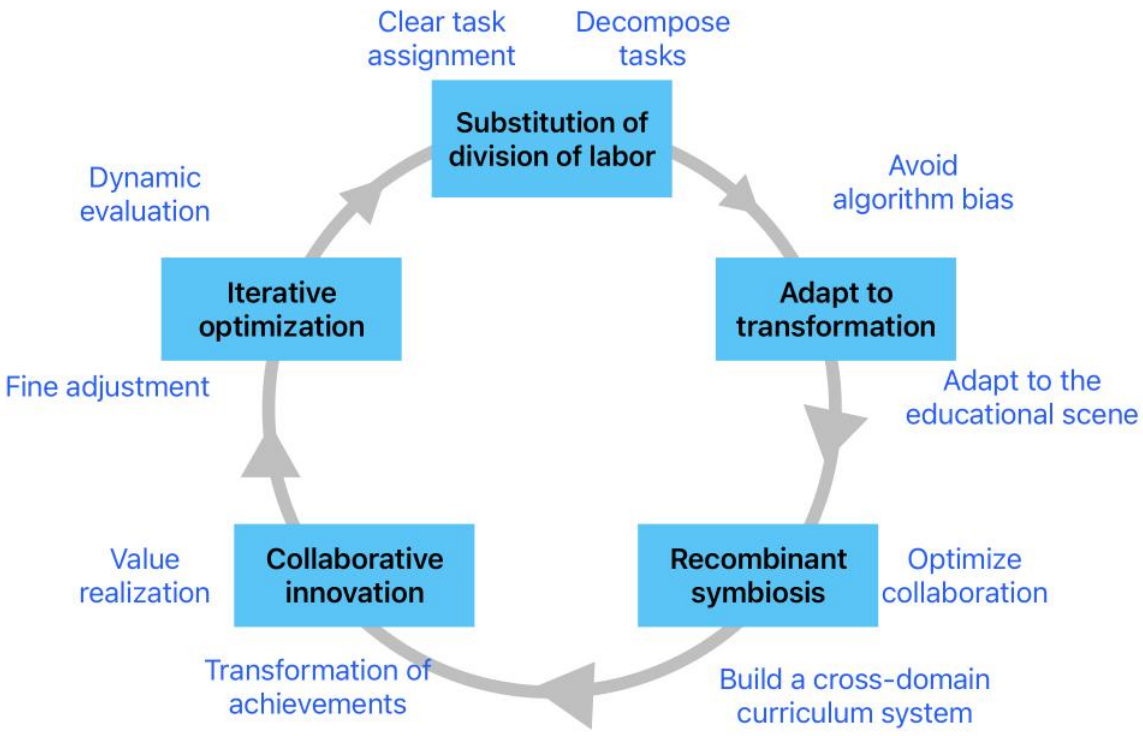


Fig.2 Multi-agent-driven Teaching Practice Strategies

4.1. Substitution of division of labor

Due to the fine-grained subject expertise contained within educational large model agents, which possess the capabilities for subject knowledge retrieval and integration (Liu et al., 2024), they can undertake fundamental tasks such as data mining and process monitoring. Multi-agent systems, through explicit task allocation, break down complex tasks into multiple sub-tasks and assign them to different agents, thereby freeing up teachers' time and energy to focus on higher-level decision-making and creative tasks, such as ethical reviews and curriculum design, enhancing teaching efficiency. For instance, multi-agent systems can automatically generate interdisciplinary project resources through knowledge retrieval and resource integration, while teachers can customize personalized learning paths based on these resources to ensure the relevance and adaptability of the teaching content.

4.2. Adapt to transformation

The Adapt to Transformation phase emphasizes the deep integration of agents with educational scenarios. Agents significantly enhance learners' grasp of knowledge through precise analysis of knowledge points, personalized recommendations of learning content, and efficient knowledge teaching. However, thinking is not an innate, naturally formed behavioral pattern, but is closely linked to the development of society, history, and culture, and is profoundly influenced and shaped by these factors (Li & Wang, 2020). Therefore, teachers need to "transform technology into education," training agents to understand educational scenarios, avoid algorithmic bias, and ensure that the agents' outputs align with educational goals and ethical requirements. At the same time, students also need to "transform application into creation," using agent tools (such as AR

laboratories) to explore multidimensional solutions and cultivate innovative thinking and practical skills.

4.3. Recombinant symbiosis

In the Recombinant Symbiosis phase, the collaboration between agents and humans is optimized through dynamic team building and knowledge network integration. Agents can be dynamically recruited or eliminated according to task requirements, ensuring the flexibility and adaptability of the team structure. For instance, referring to the AgentVerse four-stage model, agents adjust the composition and collaboration methods of the agent team dynamically through four stages: expert recruitment, collaborative decision-making, action execution, and evaluation, to efficiently solve complex tasks, fostering the development of innovative thinking and teamwork skills. Additionally, the integration of subject knowledge with AI-generated content is also key in this phase; by building a cross-domain curriculum system, disciplinary boundaries are broken down to provide students with a more comprehensive learning experience.

4.4. Collaborative innovation

In the Collaborative Innovation phase, the multi-agent system facilitates the transformation of educational value through debate protocols and collaborative creation of outcomes. Agents engage with each other through structured outputs, such as mind maps, to reach a consensus and collectively design innovative works, including solutions to social issues. This phase not only enhances students' innovative capabilities and teamwork skills but also drives innovation and transformation in educational practices through human-machine collaboration.During this phase, the multi-agent system operates as a platform where students and teachers can deeply engage with the content. The structured outputs serve as a medium for communication and idea exchange, allowing for a more comprehensive exploration of problems and the generation of creative solutions. The collaborative innovation phase encourages active participation from all parties, leveraging the strengths of both human and artificial intelligence to address complex educational challenges.

4.5. Iterative optimization

The Iterative Optimization phase, as a key step in enhancing teaching quality, focuses on strengthening the effectiveness of the multi-agent collaboration framework through continuous evaluation and dynamic adjustment. This phase emphasizes close monitoring of multi-agent interactions during the teaching process, collecting and analyzing key performance data, such as student interaction frequency, collaborative outcomes, improvements in interdisciplinary capabilities, and the timeliness of agent feedback. Based on this data, educators can effectively evaluate the dynamic changes in students' innovative thinking and then make fine adjustments to the behavior parameters, interaction strategies, refining the creation of a positive online learning environment and stimulating students' enthusiasm to actively engage in the learning process(Wang et al., 2025), and algorithms of the agents to improve the targeting and adaptability of teaching support.

Funding

The research has been made possible through the financial support of the National Natural Science Foundation of China (NSFC) under Grant 62307021, and the Humanities and Social Science Fund of the Ministry of Education of China under Grant No. 23YJC880050, and Jiangsu Province Education Science Planning Project under Grant C/2023/01/04,and Research on the Application of Generative Artificial Intelligence Technology in the Field of Personalized Career Planning for College Students 2025CXY328

References

Dong, Y., & Chen, H. (2024). Empowering interdisciplinary innovative thinking with generative artificial intelligence: Intrinsic mechanisms and model construction. *Modern Educational Technology, 34*(04), 5-15.

Janbi, N., Katib, I., & Mehmood, R. (2023). Distributed artificial intelligence: Taxonomy, review, framework, and reference architecture. *Intelligent Systems with Applications, 200231*.

Jin, X., Jiang, Q., Xiong, W., Feng, Y., & Zhao, W. (2024). Effects of student engagement in peer feedback on writing performance in higher education. *Interactive Learning Environments, 32*(1), 128-143.

Jin, X., Jiang, Q., Feng, Y., et al. (2023). Intervention mechanisms and their effects in deep learning from the perspective of intelligent peer review. *e-Education Research, 44*(3), 107-114.

Jin, X., et al. (2022). Using the online self-directed learning environment to promote creativity performance for university students. *Educational Technology & Society, 25*(2), 130-147.

Jin, X., Jiang, Q., & Ma, Z. (2025). Causes and evolution mechanisms of poor-performing students in higher education under the perspective of educational transformation in the digital age. *China Higher Education Research, (01)*, 48-56.

Li, H. F., & Wang, W. (2020). Human-machine learning symbiosis—On the construction of basic learning forms in the post-artificial intelligence education era. *Journal of Distance Education, 38*(02), 46-55.

Liu, M., Yang, M., Wu, Z., & Liao, J. (2024). Development, application status, and future prospects of educational large model agents. *Modern Educational Technology, 34*(11), 5-14.

Liu, S., Liu, Z., Duan, H., Su, Z., & Peng, X. (2025). Research on the application of large model-driven educational multi-agent systems—Technical architecture, current status, practical pathways, and future prospects. *Journal of Distance Education, 43*(01), 33-45.

Russell, S. J., & Norvig, P. (2011). Artificial Intelligence: A Modern Approach (3rd Edition). *Computer Education (15)*, 68-68.

Wang, F., Yin, H., & King, R. B. (2025). Profiling motivation and engagement in online learning: A multilevel latent profile analysis of students and institutions. *Computers & Education, 227*, 105209.

Wang, Z., Wu, Y., & Shi, X. (2024). Being top-notch does not equal innovation—Discussing the misconceptions and solutions in cultivating top innovative talents. *Modern Educational Technology, 34*(12), 5-16.

Wu, Y., Jiang, Y., Chen, Y., & Zhang, W. (2024). Large language model supported multi-agent: Technical pathways, educational applications, and future prospects. *Open Education Research, 30*(05), 63-75.

Yan, J., Xie, R., & Chen, Y. (2012). On the cultivation of college students' innovative abilities. *Academic Exchange, (07)*, 223-225.

Yuan, L., Xu, J., & Liang, S. (2025). Agent-empowered human-machine collaborative interdisciplinary thematic teaching support model. *Research in Educational Modernization, 46*(03), 87-94.

Zhai, X., Ji, S., Jiao, L., Zhu, Q., & Wang, L. (2024). Empirical study on human-machine collaboration to solve complex learning problems based on multi-agent systems. *Open Education Research, 30*(03), 63-73.

Zhou, W., & Zhao, J. (2024). The value and challenges of ChatGPT in cultivating graduate students' innovative abilities. *Higher Education Management, 18*(02), 42-52.

LLM-RAG 系统的设计与教学实践应用

Design and Teaching Practice Application of LLM-RAG System

尹强[1]，于晓梅[1*]，郑向伟[1]，赵丽香[1]
[1]信息科学与工程学院，山东师范大学
[*] yxm0708@126.com

【摘要】 本研究设计开发了一个基于 LLM-RAG 技术的智能问答助手系统，该系统旨在为学生提供一个基于自建知识库功能的学习导师，其中 RAG 技术可以有效解决 LLM 中出现的幻觉问题，自建知识库功能可以有效解决大模型信息实时性不足的问题，设计了 9 个 python 编程相关问题对智能问答助手进行测试，经过三名相关领域专业教师的评估，在智能问答助手回答问题准确性方面取得了良好的效果，后续基于智能问答助手进行了两个班级的教学实践，实践结果表明，使用智能问答助手教学的班级成绩提高率明显高于未使用智能问答助手教学的班级。
【关键词】 LLM；RAG；GAI；编程教学

Abstract: *This research designs and develops an intelligent question answering assistant system based on llm-rag technology. The system aims to provide students with a learning tutor based on the function of self built knowledge base. Rag technology can effectively solve the hallucination problem in LLM, and the function of self built knowledge base can effectively solve the problem of insufficient real-time information of large models. Nine questions related to python programming are designed to test the intelligent question answering assistant. After the evaluation of three professional teachers in related fields, it has achieved good results in the accuracy of the intelligent question answering assistant. The subsequent teaching practice based on the intelligent question answering assistant in two classes shows that the improvement rate of class performance using the intelligent question answering assistant is significantly higher than that without using it.Intelligent question and answer assistant teaching class.*
Keywords: LLM; RAG; GAI; Programming teaching

1.引言

近年来，人工智能在教育中的应用获得了巨大的吸引力，旨在通过智能技术增强和创新学习和教学方法。像 ChatGPT 这样的生成式 AI 工具在教育领域表现出了巨大的潜力，提供个性化的学习体验、学生支持和创新的课程内容交付 (Christos-Nikolaos, 2023)。这些工具根据学生的学习进度和需求提供实时互动反馈，大大提高了学习效率和参与度 (Humble et al., 2024; Yu et al., 2024)。生成式人工智能可以通过对话的方式协助用户完成文本生成、检查代码等各类事项 (Lim et al., 2023)，能够以及格或较为优秀的成绩通过大学预修生物学、人文、社会科学和法律考试 (Farazouli et al., 2024)。ChatGPT 在程序学习中的使用最近也变得很普遍，因为学习编程对大多数人来说是一个具有挑战性和复杂的过程 (Rahman & Watanobe, 2023)。它可以帮助学生了解编程语言的复杂性，例如使用代码进行调试，并提供实时解决问题的帮助 (Biswas, 2023)。大量研究已经验证了 GPT 在解决编程问题方面的性能，包括调试、生成代码和提供解释 (Pankiewicz & Baker, 2023; Phung et al., 2023; Tian et al., 2023)。

尽管生成式人工智能以其强大的能力受到教育工作者的追捧，但是其编码和解决问题等基本技能的潜在负面影响，同样引发了教育工作者的担忧 (Wise et al., 2024)。其中幻觉、过时的知识等问题是生成式人工智能的一大挑战 (Dhingra et al., 2022)。

RAG 是解决这些问题的有效途径 (Lee, 2024)，RAG，全称 Retrieval-Augmented Generation，是一种结合了信息检索与文本生成优势的模型架构，将检索方法与深度学习技术相结合，通过动态集成最新的外部信息来解决大型语言模型（LLM）的静态局限性。其核心思想是通过“检索-生成”双重机制来提高生成模型的表现。在检索阶段，模型通过检索系统从知识库中找到与输入相关的文档或段落；在生成阶段，生成模型利用检索到的信息作为上下文，生成最终的答案或文本。

本文的工作介绍如下：

1.以 Python 中的 Flask 为框架，采用 MTV 开发模式，以 LLM-RAG 为核心技术，创建了一个基于自建知识库、自选模型以及对话历史功能的智能问答助手系统，旨在解决大模型幻觉、过时知识问题以及提升学生使用体验。

2.对开发的智能问答助手进行知识回答准确性测验，设计了关于 Python 知识 3 个维度的九个问题，邀请了三名具有丰富 Python 教学经验的教师进行答案评分，取得了良好的结果。

3.基于开发的智能问答助手系统应用于教学实践，分别应用于中职学校计算机专业的两个班级，进行了为期十课时的 python 编程教学，实践结果表明，应用智能问答助手的班级成绩提高明显优于未使用智能问答助手教学的班级。

2.研究现状

2.1. GAI 在教育中的应用

ChatGPT 为教育提供了各种好处。它提供个性化的学习支持、自动化的管理流程和对在线学习的支持。ChatGPT 可以满足教育界的特定要求，包括教职员工、学生 (Mogavi et al., 2024)。教师可以利用 ChatGPT 进行作业评估以及考试设计 (Ali et al., 2024),学习监督 (Cowling et al., 2023),课程设计 (Meron & Araci, 2023)等。学生可以进行学习支持 (Rawas, 2024),写作翻译 (Keiper et al., 2023),个性化的学习辅导 (Bommineni et al., 2023)等。

2.2. GAI 在编程教育中的应用

在编程教育中，ChatGPT 也显示出解决不同编码问题的潜力。作为一种可以使用自然语言进行交互的 AI 语言模型，ChatGPT 甚至可以让那些没有编程知识的人轻松解决编码问题，大大降低了学习编程的门槛 (Surameery & Shakor, 2023)。最近的研究强调了 ChatGPT 在各种编程任务中的强大能力。例如，Tian et al. (2023)实证分析了 ChatGPT 作为代码生成、程序修复和代码摘要的全自动编程助手的潜力。他们的结果表明，ChatGPT 有效地处理了典型的编程挑战，例如修复错误、提供描述和生成代码。此外，研究发现，ChatGPT 的错误修复性能与常见的深度学习方法相比具有竞争力，并且明显优于标准的程序修复技术(Sobania et al., 2023)。Phung et al. (2023)系统地将 GPT 模型与使用不同的 Python 编程问题和现实世界的 buggy 程序的人类导师进行了比较，揭示了 GPT-4 在几种情况下的表现接近于人类导师的表现。此外，ChatGPT 协助提供有关编程作业的反馈，支持学生理论知识的实际应用。

2.3. LLM-RAG 在教育中的应用

RAG 技术致力于解决大模型回答问题时产生的幻觉问题，教育领域的知识准确性要求使得研究者们对这一技术产生了浓厚的兴趣。其中卢宇 et al. (2024)提出了大模型与检索增强生成技术结合的教学智能体基本框架及概念，阐述了其应用于教学的优势。王珊 et al. (2025)在 GLM-4 大模型的基础上结合检索增强生成等技术构建了助力家庭教育的理论模型，取得了良好的实践效果。Lee (2024)基于 ChatGPT 和检索增强生成技术开发了一个

计算机导师，能够为所有设计的二十个问题提供正确和相关的答案，并根据要求生成 R 代码片段。这项研究强调了基于 RAG 的导师为学生提供准确和个性化学习体验的潜力，同时降低了提供通常与 LLM 相关的虚假信息的风险。Muludi et al. (2024)评估了使用 ChatGPT 模型的生成预训练 Transformer 3.5 或 GPT-3.5-turbo 的 RAG 及其对文档数据处理的影响，并将其与其他应用程序进行了比较，结果突出了 RAG 的优越性。

3.系统构建

3.1. MTV 框架模式

本系统整体架构采用 Python 编程语言中的 Flask 这一轻量级且高度可扩展的 Web 框架进行构建。Flask 框架采用 MTV 模式，也就是通过 Model（模型）、View（视图）以及 Template（模板）这三个部分来组织和实现。MTV 模式的工作流程如图 1 所示。

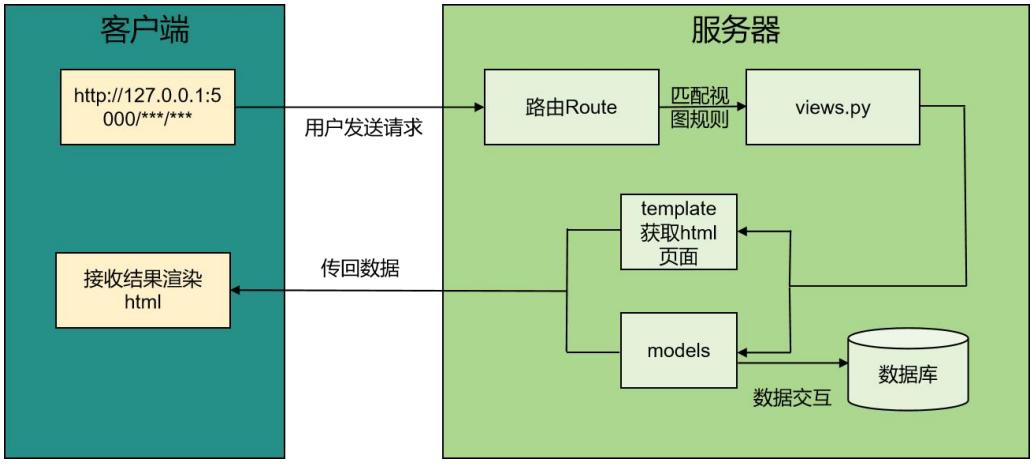


图 1 MTV 框架

3.2. LLM-RAG 技术

LLM-RAG 技术实现原理流程如图 2 所示：

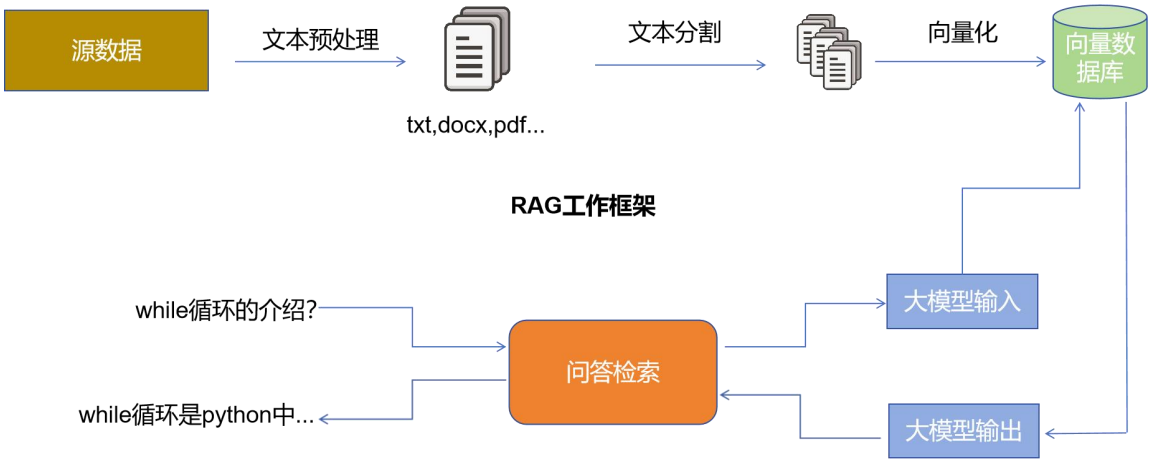


图 2 LLM-RAG 技术流程

首先进行数据的获取，为了保证数据的真实性和准确性，我们的数据获取源包括 Python 程序设计教材、菜鸟教程（python 知识网站）等，对获取的数据进行预处理，与具有多年 Python 课程教学经验的教师进行交流探讨，对重复知识以及无关知识进行剔除。

然后对处理好的数据进行文本分割，分为结构化片段和非结构化片段。针对结构化数据（csv 和 xlsx 表格文件）片段分割，直接按行进行数据读取即可。非结构化数据包括 txt,

docx，pdf 等文件，直接按照段落长度进行切割，这里设计片段长度为 500，重复长度为 100，这样做的目的是防止遗漏的信息检索。使用 openai 的 text-embedding-3-small 模型对数据进行向量化并将其存入向量数据库中供大模型根据问题进行匹配。

最后大模型对问题进行检索，答案召回阶段按照以下流程进行：用户提问——>大模型输入——>问题向量化——>L2 相似度匹配——>答案召回——>大模型总结输出。

采用 L2 欧几里得距离（Euclidean distance）进行向量化后的问题与向量数据库进行相似度匹配，L2 距离可以通过以下公式计算：

$$D = \sqrt{\sum_{i-1}^{n} (a_i - b_i)^2}$$

3.3. 基于 LLM-RAG 的智能问答助手

基于上述技术构建的智能问答助手系统如图 3 所示。

图 3 智能问答助手

模型选择功能：考虑到 openai 的 api 需要用户订阅，设计了多样性的模型选择功能，这样的设计可以接入免费的大模型 api，用户可以基于自身需求选择对话模型，这能够消除社会中的数字鸿沟影响，促进教育公平。

知识库选择功能：用户可以根据自身需求上传文档，智能问答助手根据上传的文档自动完成文档的分割和向量化任务，从而创建知识库，这意味着用户不仅可以针对于 Python 编程创建知识库满足编程需求，也可以延伸到其他学科的学习中，只需要勾选自己想要连接的知识库，智能问答助手便可以根据知识库的内容来进行对话服务。

历史对话功能：用户可以在历史对话中查看自己与智能问答助手的对话记录，这有利于用户对学习的总结归纳。

对话功能：用户可以根据自身需求选择模型、知识库，基于此与智能问答助手进行对话，得益于 RAG 功能，大模型基于知识库检索进行问题回答，避免了产生的幻觉与知识过时问题。

4.准确性测试

为了保证智能问答助手在各方面问题的回答准确性，基于《Python程序设计》教材，设计了3个维度的9个相关问题，问题相关介绍如表1所示，详细问题及答案见补充材料。

表1 问题维度及描述

维度	问题
基础定义类问题	问题1：Python语言的特点？ 问题2：Python语言的数据类型有哪些？ 问题3：Python中增加列表的方法有哪些？
编程类问题	问题4：利用for循环相关知识，在屏幕上输出九九乘法表（三角形式）。 问题5： 一个5位数，判断它是不是回文数。即12321是回文数，个位与万位相同，十位与千位相同。
纠错类问题	问题6：求输入数字的平方，如果平方运算后小于50则退出。 问题7：判断一个数是否为质数。 问题8：while循环输出0-10。 问题9：有一分数序列：2/1，3/2，5/3，...求出这个数列的前20项之和。

为保证问题回答的正确性，设计了一个评分表，评分表详细描述如表2所示。

表2 评分表

指标	得分
智能问答助手回答不准确。	1分
智能问答助手回答部分准确，但包含不正确的信息。	2分
智能问答助手回答准确，但不够充分。	3分
智能问答助手回答准确充分，但包含不必要的信息，不够简洁。	4分
智能问答助手回答准确充分并且简洁易懂。	5分

邀请了三名具有丰富Python编程教学经验的教师进行智能问答助手提供的答案评估打分，打分情况如表3所示。

表3 教师评分情况

问题	分数（教师1）	分数（教师2）	分数（教师3）
1	4	4	4
2	5	5	5
3	4	5	4
4	5	5	5
5	5	5	5
6	4	4	4
7	5	5	5
8	5	5	5
9	5	5	5

Fleiss Kappa判断多列离散无序数据，用于评估评分者之间的一致性程度。经过计算，智能问答助手答案的评分K值为0.822，表明三位教师对智能问答助手回答的质量评分有很强的一致性。

如表 3 所示，教师对于智能问答助手的答案评分大多数为 5 分，说明智能问答助手对于问题的回答有较高的准确性。

5.教学实践测试

将智能问答助手应用于信息技术课程教学实践是检验其效果的有效方式(于晓梅，2023)。本次实验在中国 J 市某中职学校开展，选择了计算机专业二年级的两个班级作为研究对象。其中，1 班被设定为实验班，在教学过程中融入了智能问答助手辅助教学；而 2 班则作为对照班，采用了传统的教学方法进行授课。

为了确保教学实验结果的准确性和可信度，在正式开展教学活动之前，采取了一项预备措施：为保证实验结果的可靠性，在进行教学实验之前选用《Python 程序设计》第一章节 Python 程序基础的单元测试结果作为前测数据。

本研究使用 SPSS 数据分析软件，对 1 班（实验班）和 2 班（对照班）的前测成绩进行了统计分析，统计结果详见表 4。依据 Shapiro-Wilk 检验所得出的 p 值，可以判定两个班级的成绩均服从正态分布。如表 4 所示，实验班（1 班）的平均成绩为 52.42，标准差为 7.916；而对照班（2 班）的平均成绩为 53.02，标准差为 7.331。从这两个统计指标来看，在未进行教学前两个班级的成绩水平和离散程度都较为接近，说明两个班级在成绩表现上具有相似性。

表 4 前测成绩分析结果

名称	样本量	平均值	标准差	Kolmogorov-Smirnov 检验		Shapiro-Wilk 检验	
				统计量 D 值	p	统计量 W 值	p
1 班	45	52.42	7.916	0.077	0.200	0.966	0.207
2 班	44	53.02	7.331	0.117	0.151	0.967	0.240

分别从学生与教师两个维度出发，设计了基于智能问答助手系统的教学活动，这些活动巧妙地贯穿于课前准备、课堂实施及课后巩固这三个关键环节之中，整体设计思路如图 4 所示。

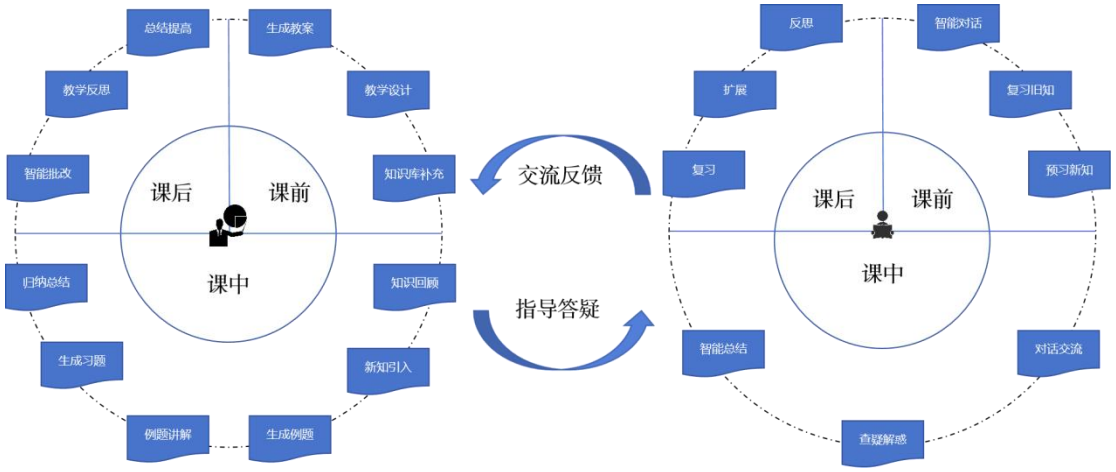


图 4 教学活动设计

在完成十课时的教学后，对实验组班级和对照组班级分别进行了 Python 程序语句相关的后测。测试内容为第二章节 Python 程序语句相关知识。利用 SPSS 数据分析工具对后

测成绩进行了统计分析，结果如表 5 所示。数据分析显示，使用了智能问答助手的班级，其成绩平均值和标准差均高于未使用智能问答助手的班级，经过计算，实验组班级学生平均成绩提高率为 23.4%，对照组班级学生平均成绩提高率为 11.6%，实验组班级成绩提高率比对照组班级平均成绩提高率高了 11.8%。这表明，在教学中引入智能问答助手有助于提升学生的学习成绩，从而验证了该助手在教学中提高学生成绩的有效性。

表 5 后测成绩分析结果

名称	样本量	最小值	最大值	平均值	标准差
1 班	45	48	85	64.69	7.951
2 班	44	44	80	59.16	7.631

6.结论

这项研究表明，基于 LLM-RAG 构建的智能问答助手系统对问题的回答有较高的准确性，解决了幻觉和知识过时的问题，不仅如此，知识库、模型的选择功能以及历史对话功能都能够很好的满足学生的学习需求。这项研究取得了良好的效果，但是也存在一些问题需要在后续的研究中进行解决。首先知识回答包含不必要的信息、答案不够简洁，这可能是因为知识库中数据过于繁杂造成的，为了解决这一问题，在后续的研究中应该对知识库中的数据进行更加严格的筛选以及预处理。其次本研究局限于智能问答助手应用在教学中学生成绩的变化探索，在后续的研究中要对学生学习过程、问题解决能力等方面的变化进行深入探讨。

致谢

本研究项目资助：山东省研究生优质教育教学资源项目（SDYKC2024146）。

参考文献

Ali, K., Barhom, N., Tamimi, F., & Duggal, M. (2024). ChatGPT—A double-edged sword for healthcare education? Implications for assessments of dental students. *European Journal of Dental Education*, *28*(1), 206-211.

Biswas, S. (2023). Role of ChatGPT in Computer Programming. *Mesopotamian Journal of Computer Science*, *2023*, 9-15.

Bommineni, V. L., Bhagwagar, S., Balcarcel, D., Bommineni, V., Davazitkos, C., & Boyer, D. (2023). Performance of ChatGPT on the MCAT: the road to personalized and equitable premedical learning. *MedRxiv*, 2023.2003. 2005.23286533.

Christos-Nikolaos, A. (2023). ChatGPT impacts in programming education: A recent literature overview that debates ChatGPT responses. *arXiv preprint arXiv*, *12348*.

Cowling, M., Crawford, J., Allen, K.-A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology*, *39*(4), 89-103.

Dhingra, B., Cole, J. R., Eisenschlos, J. M., Gillick, D., Eisenstein, J., & Cohen, W. W. (2022). Time-aware language models as temporal knowledge bases. *Transactions of the Association for Computational Linguistics*, *10*, 257-273.

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, *49*(3), 363-375.

Humble, N., Boustedt, J., Holmgren, H., Milutinovic, G., Seipel, S., & Östberg, A.-S. (2024). Cheaters or AI-enhanced learners: Consequences of ChatGPT for programming education. *Electronic journal of e-Learning*, *22*(2), 16-29.

Keiper, M. C., Fried, G., Lupinek, J., & Nordstrom, H. (2023). Artificial intelligence in sport management education: Playing the AI game with ChatGPT. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *33*, 100456.

Lee, Y. (2024). Developing a computer-based tutor utilizing Generative Artificial Intelligence (GAI) and Retrieval-Augmented Generation (RAG). *Education and Information Technologies*, 1-22.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The international journal of management education*, *21*(2), 100790.

Meron, Y., & Araci, Y. T. (2023). Artificial intelligence in design education: evaluating ChatGPT as a virtual colleague for post-graduate course development. *Design Science*, *9*, e30.

Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S.,…Gujar, S. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100027.

Muludi, K., Fitria, K. M., & Triloka, J. (2024). Retrieval-Augmented Generation Approach: Document Question Answering using Large Language Model. *International Journal of Advanced Computer Science & Applications*, *15*(3).

Pankiewicz, M., & Baker, R. S. (2023). Large Language Models (GPT) for automating feedback on programming assignments. *arXiv preprint arXiv*.

Phung, T., Pădurean, V.-A., Cambronero, J., Gulwani, S., Kohn, T., Majumdar, R.,…Soares, G. (2023). Generative AI for programming education: benchmarking ChatGPT, GPT-4, and human tutors. Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 2.

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

Rawas, S. (2024). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, *29*(6), 6895-6908.

Sobania, D., Briesch, M., Hanna, C., & Petke, J. (2023). An analysis of the automatic bug fixing performance of chatgpt. 2023 IEEE/ACM International Workshop on Automated Program Repair (APR).

Surameery, N. M. S., & Shakor, M. Y. (2023). Use chat gpt to solve programming bugs. *International Journal of Information Technology & Computer Engineering*(31), 17-22.

Tian, H., Lu, W., Li, T. O., Tang, X., Cheung, S.-C., Klein, J., & Bissyandé, T. F. (2023). Is ChatGPT the ultimate programming assistant--how far is it? *arXiv preprint arXiv*.

Wise, B., Emerson, L., Van Luyn, A., Dyson, B., Bjork, C., Thomas, S. E. J. H. E. R., & Development. (2024). A scholarly dialogue: writing scholarship, authorship, academic integrity and the challenges of AI. *43*(3), 578-590.

Yu, X., Mao, Q., Wang, X., Yin, Q., Che, X., & Zheng, X. J. E. S. w. A. (2024). CR-LCRP: Course recommendation based on Learner–Course Relation Prediction with data augmentation in a heterogeneous view. *249*, 123777.

卢宇, 余京蕾, & 陈鹏鹤. (2024). 基于大模型的教学智能体构建与应用研究. *中国电化教育* (07), 99-108.

王珊, 肖义墙, 潘亦宁, & 彭唯. (2025). 生成式人工智能助力家庭教育的模型构建及系统实践. *电化教育研究*, *46*(03), 64-71.
于晓梅. (2023). *软件与教育信息服务*. 山东人民出版社.

基於小學語文課程的多模態學習分析方法研究

Research on Multimodal Learning Analysis Method Based on Primary School Chinese Curriculum

時秋[1]，吳穎[1]
[1]江南大學
2484621349@qq.com

【摘要】本研究聚焦小學語文課堂《軍神》，採用多模態學習分析（MMLA），並借助 NVivo 軟體來進行量化分析。在參考現有編碼體系的基礎上，結合義務教育階段語文課程核心素養標準以及《軍神》課程的的特點，編制了一套詳細編碼體系，該體系涵蓋了認知活動與行為參與兩個維度。本研究揭示了《軍神》這堂課促進深度學習、實施互動式教學等方面成效顯著，指出在技術依賴、跨學科、資料解釋複雜及隱私倫理等方面存在的挑戰。綜上所述，研究為多模態學習分析在小學語文教學領域的應用提供了實證支持，來幫助教育工作者優化教學策略，進而提升教學教學的品質。
【關鍵字】多模態學習分析；小學語文；分析方法

Abstract: *This study focuses on the primary school Chinese language classroom "Military God", using multimodal learning analysis (MMLA) and utilizing NVivo software for quantitative analysis. On the basis of referring to the existing coding system, combined with the core literacy standards of Chinese language curriculum in compulsory education and the characteristics of the "Military God" curriculum, a detailed coding system has been developed, which covers two dimensions: cognitive activities and behavioral participation. This study reveals that the course "Military God" has achieved significant results in promoting deep learning and implementing interactive teaching, and points out the challenges in technology dependence, interdisciplinary learning, complex data interpretation, and privacy ethics. In summary, the study provides empirical support for the application of multimodal learning analysis in primary school Chinese language teaching, to help educators optimize teaching strategies and improve the quality of teaching.*
Keywords: Multimodal learning analysis, Primary school Chinese language, Analysis method

1.前言

在教育研究領域，傳統的課堂分析手段主要依賴於現場觀察法、訪談法以及問卷調查法等，但這些方法在實際應用中存在很多局限性，比如較強主觀性容易導致分析結果產生偏差、回饋的延遲性很難滿足即時性的要求、實證資料的欠缺性使得分析的深度與廣度受到限制，從而使得在課堂診斷分析的科學性、精准性以及回饋的有效性維度上，存在著顯著的提升空間亟待填補。相比之下，多模態分析方法在課堂情境中的應用逐漸受到關注，在這一框架下，非言語活動與情感參與對於學習者發展的重要性不容忽視。借助多模態融合分析的方式，能夠為教師深入理解教學場景中的關鍵事件以及深刻反思教學行為提供堅實有力的依據支撐，從而在一定程度上有效彌補傳統課堂分析方法的固有缺陷，並為課堂分析領域開闢新的視角與途徑，進而推動課堂分析向更加科學、精准、高效的方向進階。

2.方法的定義及過程模型

2.1.方法的定義與發展

多模態學習分析（MultimodalLearningAnalytics）概念最早是由美國南加州大學創新技術學院的 StefanScherer 和 Louis-PhilippeMorency 等於 2012 年在第十四屆《多模態交互國際會議》上正式提出並發表。它是三個概念的交叉點：多模態教學與學習、多模態資料、電腦支援的分析，本質上它利用非傳統和傳統資料形式在三個概念之間所形成的三角關係來描述或模擬複雜學習環境中的學生學習（Sharma et al., 2020）。因此，多模態學習分析

是一個橫跨學習科學和機器學習的研究領域，能夠為複雜學習行為和學習理論之間搭建橋樑。它利用多模資料捕獲和信號處理技術研究複雜學習環境中的學習(Olsen et al., 2020)。尤其是近年來，各類可穿戴感測器技術、物聯網技術的發展和大資料計算能力的提升為多模態研究提供了必要的技術支撐。多模態學習研究的興起除了技術推動的作用之外，還起源於對教育發展規律的持續探索，多模態與教育分析有著天然的契合度：多模態方法更符合人類交流的本性。人類交流過程中對多種形態的使用較為豐富，且這些形態形成互補(馬志強等，2021)。具體來說，多模態學習分析（簡稱 MMLA）是指利用多種分析技術對學習情境中產生的多模態學習行為資料進行同一化整合處理和建模分析，來預測學習行為風險，為教師和學生提供教學支援和導學服務，以實現優化學習過程和解釋學習機理的目的（Xu et al., 2023）。

2.2.過程模型

在借助機器學習的多模態學習分析領域中，有兩個非常重要的概念：一個是學習行為，另外一個是學習標籤。為更好地建立和解釋外顯行為和潛在行為屬性間的對應關係，將學習行為定義為輸入空間中的可被觀察或測量的學習者個體的運動行為、生理行為及與環境交互行為，潛在行為屬性定義為假設空間中用來解釋行為資料意義的學習標籤，是感測器不可直接觀察到的，但可用資料表征的、由人類專家注釋完成的、可解釋學習者意義形成過程的與學習相關的屬性結構，如情緒、動機或認知學習行為涵蓋言語和非言語行為，通過可觀測的模態資料表征，而學習標籤則是依據教與學理論為多模態資料賦予的認知、情感、社交等屬性狀態值。在本研究中，主要參考了 DanieleDiMitri 的過程模型，該模型是由“P”引導學習行為分析過程，生成的“R”結果來解釋學習分析的結果，MMLA 包括以下關鍵步驟：感測器採集表徵多模態資料（P1-R1），人工/自動標注解釋學習標籤（P2-R2），機器學習/統計分析建模預測學習狀態（P3-R3），結合具體任務回饋解釋促成行為改變（P4-R4），這一模型為系統分析課堂教學中的多模態資料提供了理論框架（Chango et al., 2022）。

3.實施過程與支持工具

3.1.支援工具

本文主要採用 NVivo11 質性分析軟體進行行為分析。該軟體澳大利亞 QSR 公司產品，具有介面簡潔、功能強大的特點。它支援多種資料類型處理，可精確控制視頻分析，具備強大編碼和檢索功能，可以構建節點編碼表，將課堂實錄導入後，按照節點進行編碼，自動形成節點編碼覆蓋層次圖、編碼帶等視覺化圖形，非常適合本研究對課堂視頻的深度分析需求。通過觀看課堂視頻（劉清堂等人，2022），依據編碼體系對學生行為逐段標記編碼，記錄行為起始時間、結束時間、類型及對應認知層次，確保資料準確完整。將編碼資料錄入 NVivo 軟體，利用其功能進行管理分析，生成編碼圖和節點覆蓋率百分比圖，直觀展示行為分佈和占比情況（李愛霞等人，2022）。

本研究參考弗蘭德斯互動分析系統（FIAS）的基本框架,依據任務驅動教學法和語文學科特點,生成了“《軍神》學生課堂參與編碼表”，並記錄每三秒的互動情況，以此來分析學生的課堂活動情況，旨在說明教師反思教學實踐，促進課堂互動的平衡，進而提升教學品質。

3.2.研究物件

本文研究從自蘇格拉底課程平臺中，選擇小學五年級語文《軍神》一課的智慧課堂教學活動作為分析課例，課程時長為 40 分鐘。將學生的課堂學習行為、課堂教案設計以及

學生繪製的思維導圖作為觀測分析物件，遵循學生學習行為多模態分析框架，利用多模態資料分析方法，分析學生的聲音資料和文本資料，從而評估學生對課程的理解程度和課堂互動情況，進而判斷其是否符合語文新課標核心素養導向下的課程理念。課例選取遵循兩個原則：①聲音錄製清晰，能準確識別學生的聲音資訊；②畫面錄製多方位，包括全景、特寫、動態捕捉，能觀察學生整體與局部的行為。

3.3 聲音資料研究設計

學習行為分析框架細化，列舉對應的學習行為特徵。同時，根據列舉的學習行為特徵進行編碼標注，聲音數據主要分類兩大類，分別為認知活動聲音數據和行為參與聲音數據，認知活動聲音數據中包括 10 個學習行為特徵，分別為記憶、理解、應用、分析、評價、創造和反思，行為參與聲音數據中包括 11 個學習行為特徵，分別為課堂聽講、個體回答、部分應答、集體應答、課堂展示、提問、小組學習和討論、課堂練習、作品創作、集體朗讀課文、學生點評。根據英文名稱的縮寫給這些行為特徵進行編碼，編碼名稱依次為：R、U、A、An、E、C、Re、L、P、Sq、Cq、Pr、Q、C、Ex、W、Rea、Sc，見表 1。

我們結合具體的課堂情境對上述 21 個編碼進行描述，認知活動維度，“記憶”（R）、“理解”（U）和“應用”（A）這三個認知活動的評估可以幫助學生鞏固基礎知識，並通過實際練習加深對課文內容的理解。例如，在課堂上，教師提問“劉伯承元帥是誰？”學生回答：“劉伯承是中國人民解放軍的開國元帥之一，他以堅韌不拔的毅力和卓越的軍事才能著稱。”屬於“記憶”（R）；教師講解劉伯承拒絕使用麻醉劑的原因後，提問：“為什麼劉伯承拒絕使用麻醉劑？”學生回答：“因為他不想在手術過程中失去意識，以便更好地配合醫生。”屬於“理解”（U）；在小組討論中，學生討論如何在現代醫療條件下重新設計手術場景，以減少劉伯承的痛苦，屬於“應用”（A）。通過復述故事情節或角色對話，學生能夠更好地把握文章的主要脈絡；而在課堂上模擬手術場景，讓學生扮演醫生和病人的角色，則能幫助他們親身體驗課文中角色的情感和行為動機。

“分析”（An）和“評價”（E）這兩個高級認知活動要求學生在理解文本的基礎上進行更深入的思考和分析。例如，學生分析劉伯承在手術過程中的心理狀態，討論：“劉伯承在手術過程中忍受巨大疼痛，這體現了他的什麼精神品質？”學生回答：“體現了他堅韌不拔的毅力和對革命事業的忠誠。”屬於“分析”（An）；在學生作品展示環節，一位學生展示了一幅描繪劉伯承手術場景的繪畫。另一位學生點評：“這幅畫生動地展現了劉伯承的堅強意志，尤其是他緊握的雙手，讓人感受到他的痛苦和堅持。”屬於“評價”（E）。通過對課文細節的關注和對角色行為的評判，學生們不僅能提高自己的邏輯思維能力，還能形成正確的價值觀和人生觀。《軍神》一文展現了劉伯承元帥堅韌不拔的精神品質，學生在分析和評價他的行為時，也會受到這種精神的感染和鼓舞。

“創造”（C）這一最高級別的認知活動鼓勵學生在充分吸收前幾個階段所學知識的基礎上發揮想像力，創造出新的東西。例如，可以讓學生以現代視角重新構思這個故事的發生背景和發展走向，或者設計一個全新的結尾來展現主人公的性格特點。這個過程不僅鍛煉了學生的創新能力，也讓他們在實踐中深化了對課文主題的理解。“反思”（Re）作為整個學習過程的總結性環節，促使學生在完成所有任務後回頭審視自己的學習歷程，找出優點和不足之處並進行調整。這對於培養自主學習能力和持續改進的習慣至關重要。同時，通過反思也能使學生認識到自己在情感態度上的變化和個人成長的重要性。

在行為參與維度，“課堂聽講”（L）和“集體朗讀課文”（Rea）這兩個基本的活動評估了學生對課文內容的初步接觸和理解程度。“課堂聽講”（L）要求學生認真傾聽教

師的講解，從而獲取新知識；“集體朗讀課文”（Rea）則是全班同學一起朗讀指定的課文段落，幫助學生在整體上熟悉課文內容。“個體回答”（P）、“部分應答”（Sq）和“集體應答”（Cq）等互動環節檢驗了學生對課文理解的深度和廣度。教師可以通過這些問題瞭解每個學生的學習情況，並及時給予回饋和指導。例如，教師提問：“劉伯承元帥為什麼拒絕使用麻醉劑？”小華說：“因為他想保持清醒，更好地配合醫生。”屬於“個體回答”（P）；教師提問：“劉伯承元帥在手術中表現出了哪些品質？”小李說：“堅韌不拔。”小王補充說：“還有頑強的毅力。”屬於“部分應答”（Sq）；教師提問：“劉伯承元帥的精神品質是什麼？”全班學生齊聲回答：“堅韌不拔、頑強毅力。”屬於（Cq）。

“作品創作”（W）和“課堂展示”（Pr）等活動則鼓勵學生將所學知識轉化為創意表達和實踐能力，如寫作、繪畫或其他形式的藝術表現。例如，教師佈置繪製課本故事線的思維導圖屬於“作品創作”（W）；教師安排學生上臺展示思維導圖屬於“課堂展示”（Pr）。“小組學習和討論”（C）、“課堂練習”（Ex）以及“學生點評”（Sc）等合作與評價機制促進了學生的團隊協作能力和批判性思維的發展。在這些活動中，學生們不僅可以分享彼此的想法和觀點，還可以通過互相評價來提高自己的學習能力。例如，教師組織學生分組討論“劉伯承元帥拒絕使用麻醉劑的原因”，屬於“小組學習和討論”（C）；教師佈置課堂練習，要求學生回答關於《軍神》課文內容的問題，屬於“課堂練習”（Ex）；學生展示自己創作的關於故事線的思維導圖後，其他學生進行點評，屬於“學生點評”（Sc）。

表 1《軍神》課程行為認知編碼表

《軍神》課程行為認知編碼表		
類別	編碼	描述
認知活動	記憶 R	學生複述關鍵資訊和基礎知識
	理解 U	通過課堂講解和個體回答環節，學生理解知識點及其含義
	應用 A	課堂練習和小組討論學生將所學知識應用於具體情境中
	分析 An	學生在分析課文內容和人物心理時，進行資訊的連接與區分
	評價 E	學生點評和對同學作品的評價
	創造 C	在作品創作環節中，學生展現創新思維和創造性表達
	反思 Re	在課後任務學生反思學習內容和過程，進行深層次的思考和內化
行為參與	課堂聽講 L	教師主導講解新知識，學生認真傾聽並吸收教學內容
	個體回答 P	教師隨機或指定學生單獨回答問題，檢驗個人的理解程度
	部分應答 Sq	針對教師的提問，部分學生進行回答，可能是自願或被選中的方式
	集體應答 Cq	全班學生共同回答教師的提問，通常用於複習或強調某個知

		識點
	課堂展示 Pr	學生或教師展示特定的內容，如學生的作業、專案或演示等
	提問 Q	教師提出問題促使學生思考，或者學生向教師提出疑問以求解答
	小組學習和討論 C	學生分組討論特定主題或解決問題，通過團隊合作深化理解
	課堂練習 Ex	在教師的指導下，學生進行針對性的練習，鞏固和應用所學知識
	作品創作 W	學生根據所學內容進行創作，如寫作、繪畫或其他形式的創意表達
	集體朗讀課文 Rea	全班學生一起朗讀指定的課文或段落
	學生點評 Sc	學生對其他同學的表現或作品進行評價

利用課堂環境中的聲音、圖像、文本資料獲取設備，兩位元觀察者即時觀測統計並對多模態學習行為進行編碼對照。編碼間隔不固定，依照固定內容而定，利用 NVivo 軟體對課例進行編碼標注，最後對統計結果進行視覺化呈現。兩位元觀察者根據學習行為編碼體系協商確定資料獲取具體規則及細節，達成統一後，按照編碼體系表分別進行資料統計，對統計編碼資料進行梳理校對，為學習行為研究提供精准的多模態資料支援，最終在理解程度評估維度共計標記編碼 23 條，在互動情況評估維度共計標記編碼 85 條，分別按照頻次進行排序，視覺化呈現如圖 2 和圖 3 所示。

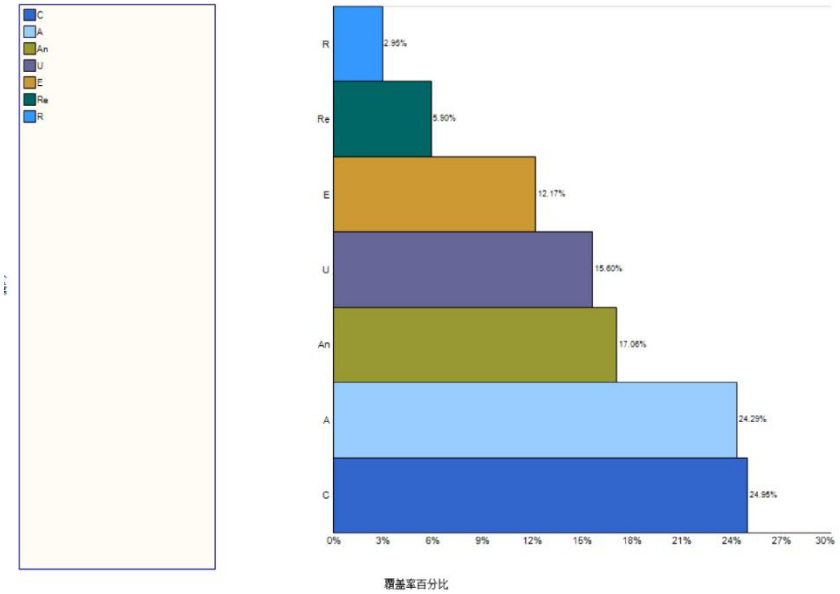


圖 2 認知活動編碼頻次排序

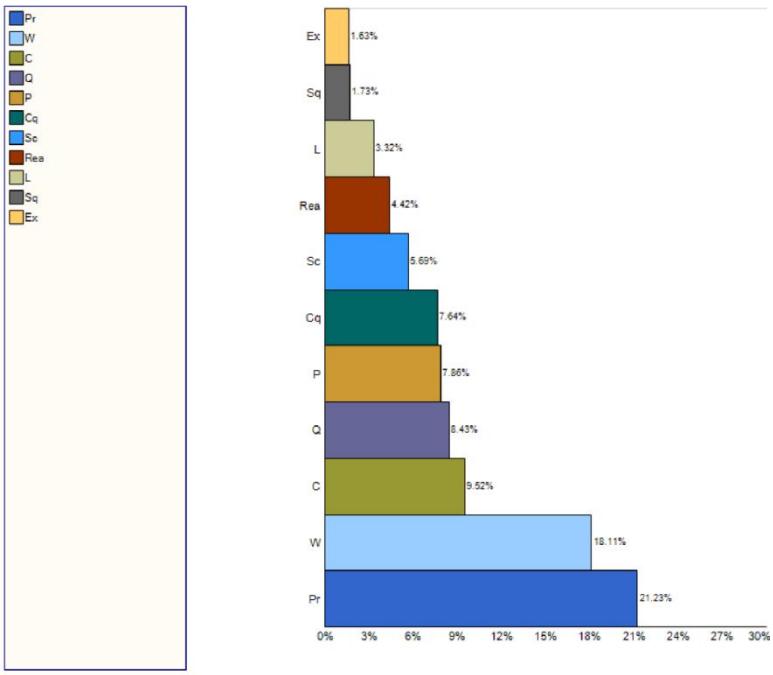


圖 3 行為參與變化頻次排序

4.文本資料研究設計

4.1 教案文本

教案文本主要是老師上課的 ppt，對 ppt 進行提取可以得出以下教學流程：課程以問題導入，激發學生的學習興趣。教師通過提問“你們聽說過軍神嗎？他是誰？”引導學生思考，並引出課文的主人公劉伯承元帥。接著，學生通過自由朗讀課文，初步感知課文內容，瞭解故事梗概。教師引導學生關注課文中的關鍵字語和句式，並思考沃克醫生和劉伯承之間發生了什麼。隨後，課程進入精讀環節，重點學習課文中的幾個關鍵部分。教師引導學生學習劉伯承元帥拒絕使用麻醉劑的原因，以及他在手術過程中忍受巨大疼痛的描寫，感受他堅強的意志和頑強的毅力。在品讀感悟環節，教師通過多種教學活動，如朗讀、討論、表演等，引導學生體會人物的內心情感，感受人物形象。例如，教師可以引導學生用沃克醫生的口吻講述這個故事，或者讓學生表演劉伯承元帥接受手術的場景。拓展延伸阶段，教師鼓勵學生用沃克醫生的口吻講述這個故事，並閱讀課後連結《豐碑》，感受人物形象，體會革命先烈的偉大精神。最後，教師通過觀察學生在課堂上的參與程度、朗讀等活動，以及他們的書面作業，對學生的學習情況進行評價。教師還要引導學生進行自我評價和反思，幫助他們找到自己的不足之處，並制定改進計畫。

4.2 思維導圖文本

在《軍神》這節課中，讓學生繪製思維導圖的目的在於促進學生對課文內容的深入理解和綜合把握。思維導圖作為一種圖形化的學習工具，能夠有效地說明學生組織和優化資訊，提高學習效率。繪製思維導圖能夠促進學生對於課文結構和主要內容的梳理。在繪製過程中，學生需要思考哪些資訊是核心，哪些是支撐細節，這樣的過程實際上是對知識點的再次篩選和強化記憶。此外，學生需要思考如何將抽象的文字資訊轉化為形象的圖形，這本身就是一種創新過程。同時，他們還需要對課文內容進行批判性分析，判斷哪些資訊是重要的，哪些可以省略,選擇展示了兩個小組的思維導圖，如圖 4 和圖 5 所示。

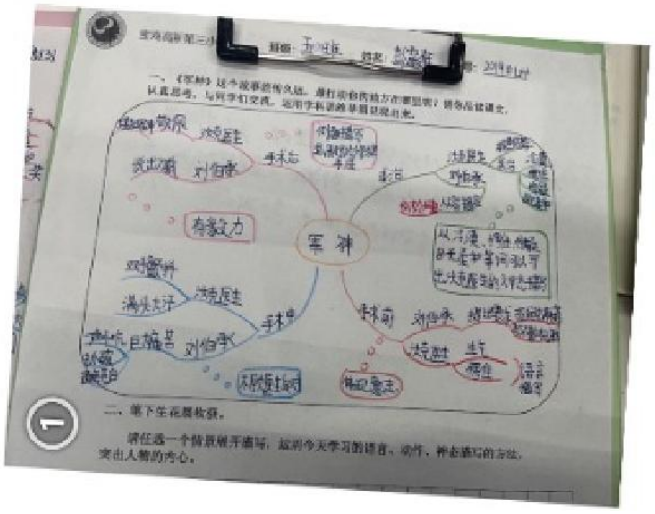


圖 4 第一小組思維導圖展示

圖 5 第二小組思維導圖展示

5.研究結果與分析

5.1 聲音資料分析

依據所使用的 NVivo 軟體構建的編碼統計表，本研究編制了相應的統計分析表。在認知活動的分類體系中，見表 6，淺層學習涵蓋記憶、理解和應用三個維度，其各自占比分別為 2.95%、15.6%和 24.29%，累計占比為 42.84%。深層學習則包括分析、評價、創造和反思四個維度，其占比依次為 17.06%、12.17%、24.95%和 5.90%，累計占比為 57.16%。《軍神》這節課在淺層學習與深層學習之間實現了很好地平衡。淺層學習構成了深度學習的基礎，其涉及的記憶與理解維度有助於學生掌握基礎知識點、深入理解教學內容，為深度學習奠定基礎。應用層面的學習能夠促進知識在實際情境中的應用，增強學生的實踐操作能力。深層學習所包含的分析、評價與創造維度能夠激勵學生進行深入思考，培養其批判性思維和創新能力。反思作為深層學習的關鍵組成部分，涉及自我管理和元認知，能夠促使學生從經驗中汲取教訓，實現自我成長。這種全面的學習方式有助於學生在掌握基礎知識和技能的同時，發展高階思維能力。

表 6 認知互動編碼分析

認知活動	淺層學習			深層學習			
	記憶	理解	應用	分析	評價	創造	反思
	2.95%	15.6%	24.29%	17.06%	12.17%	24.95%	5.90%
	42.84%			57.16%			

在課堂行為參與的分類體系中見表 7，個人表現類行為包括七個維度：課堂聽講、個體回答、部分應答、集體應答、課堂練習、集體朗讀以及學生課文點評。這些維度在課堂行為中的比重分別為 3.32%、7.86%、1.73%、7.64%、1.63%、4.42%和 5.69%，累計占比達到 32.29%。另一方面，合作互動類行為涵蓋四個維度：課堂展示、學生提問與討論、小組學習與討論、以及作品創作。這些維度的比重依次為 21.23%、8.43%、9.52%和 18.11%，累計占比為 57.29%。資料分析揭示了本節課在個人表現與合作互動兩種行為類型上的傾向性差異，合作互動類行為在課堂行為參與中佔據了主導地位。其中，課堂展示、小組學習與討論、以及作品創作這三個維度對於促進學生間的思想交流和知識共用，進而深化和拓展知識具有顯著作用。學生提問與討論這一維度則能夠激發學生的求知欲，促進其主動思考和表達能力的發展。個人表現類行為中的個體回答、部分應答和集體應答等維度對於提高學生的臨場反應和語言組織能力具有積極作用。課堂聽講作為知識獲取的基礎，課堂練習對於知識鞏固至關重要，集體朗讀有助於提升學生的語感，而學生課文點評則能鍛煉學生的分析和鑒賞能力。這種多元化的課堂參與模式不僅有助於學生獲取知識，而且能夠全面提升學生的綜合素養。

表 7 行為參與編碼分析

行為參與	課堂聽講	個體回答	部分應答	集體應答	課堂展示	學生提問	小組學習和討論	課堂練習	作評創作	集體朗讀課文	學生點評
	3.32%	7.86%	1.73%	7.64%	21.23%	8.43%	9.52%	1.63%	18.11%	4.42%	5.69%

5.2 文本資料分析

在對文本資料的分析中，本研究著重考察了教師在課堂上使用的 ppt 以及學生製作的思維導圖。基於語文學科核心素養，教學設計呈現以下特徵：語言建構與運用方面，通過課文深度解讀、朗讀分享等活動，學生在豐富語言實踐中體悟語言精妙，教師引導學生分析人物對話與神態描寫，揭示語言在塑造人物形象中的作用，有效提升學生語言表達與運用能力。思維發展與提升方面，設置開放性問題如“劉伯承為何拒絕使用麻醉劑”，激發學生深度思考與探究，通過小組討論與思維導圖創作，學生批判性、創新性及邏輯思維能力顯著增強。審美鑒賞與創造力方面，深入剖析人物描寫，學生感悟人物精神特質與美學形象，創設教學情境，引導學生運用描寫技巧展現人物內心世界，培養審美感知力與創造力。文化傳承與理解方面，圍繞劉伯承英雄事蹟教學，學生全面瞭解革命先輩精神風貌與品德，激發傳承與弘揚革命文化自覺性，堅定文化自信。思維導圖創作中，語言建構與運用體現於提煉概括課文核心內容，形成清晰分支結構；思維發展與提升體現於運用視覺元素梳理邏輯關係，深化文本理解；審美鑒賞與創造力通過色彩、線條及佈局展現視覺審美與個性化風格；文化傳承與理解則體現于展現革命英雄事蹟，表達崇敬與傳承革命文化之情。

總的來說，本課程深度貫徹語文學科核心素養的教育理念，核心教學模式採用任務驅動法，並巧妙融合思維導圖與情境寫作兩種工具。在教學實踐層面，通過精密規劃的教學環節及對學生回饋的敏銳捕捉與靈活應對，確保了教學活動的嚴謹有序與高效運行，顯著提升了教學的靈動性。課程著重強化師生互動交流，通過細緻觀察學生在思維導圖構建過程中的行為表現，並給予精准的指導與即時回饋，構建了高效暢通的師生溝通橋樑。同時，本課程採用了多元化的教學策略：其一，利用任務驅動法精心創設教學情境，有效激發學生的深度思考與強烈情感共鳴；其二，借助思維導圖說明學生系統整理知識結構，並通過展示學生作品促進相互借鑒與學習；其三，通過小組合作學習模式，全面培養學生的團隊協作能力與高效溝通技巧。此外，課程還充分利用情境設置的趣味性與創造性，有效激發學生的學習內在動力，促進其自主學習能力的持續發展；結合多媒體教學資源，進一步豐富教學內容，顯著提升學生的學習積極性；運用資料分析法對教學效果進行科學評估，以精准優化教學策略。在教學目標的設定上，本課程致力於全面培育並提升學生的語文學科核心素養。具體而言，通過深度解讀課文內容，引導學生準確把握主題、人物形象與情感脈絡，培養其文化自信；借助思維導圖的構建過程，引導學生提煉關鍵資訊、梳理課文結構，進而鍛煉其邏輯思維能力；通過情境寫作訓練，激發學生的審美創造力；借助小組討論與情境寫作等實踐活動，全面增強學生的語言表達與應用能力。

6.總結

多模態學習分析賦能研究者對教師教學狀態、學生學習狀態以及學習環境等的全面瞭解，精心的多模態資料獲取設計是保障多模態學習分析結果準確的關鍵。然而，多模態學習領域亦遭遇眾多挑戰。由於時間精力有限，只選取一節課例作為研究樣本，本研究僅選取單一課程實例作為分析樣本。儘管研究過程中利用了 Nvivo 質性分析軟體以簡化視頻編碼流程，但該軟體尚未能實現對教學行為的自動智慧化識別與編碼，這與人腦的識別能力相比仍存在差距。因此，大部分編碼工作仍需依賴人工作業，導致整個研究過程繁瑣且耗時。此外，編碼過程中的主觀性、跨學科融合難度大、隱私與倫理問題亦不容忽視。鑒於此，未來針對教學行為的研究可考慮融入人工智慧分析方法，特別是基於深度學習技術的視頻處理，以期更精准地解析課堂教學情境，並實現對教學行為的自動識別。這將為教學行為研究領域帶來更為高效的研究工具。

參考文獻

李愛霞,張胡 & 顧小清.(2022).多模態學習分析：概述、應用案例及眷注.教育傳播與技術,(04),90-95.

馬志強,嶽芸竹 & 王文秋.(2021).基於多模態交互資訊的協作學習投入分析.現代教育技術,31(01),47-53.

劉清堂,李小娟,謝魁,常瑀倍 & 鄭欣欣.(2022).多模態學習分析實證研究的發展與展望.電化教育研究,43(01),71-78+85.

Chango, W., Lara, J. A., Cerezo, R., & Romero, C. (2022). A review on data fusion in multimodal learning analytics and educational data mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(4), e1458.

Olsen, J. K., Sharma, K., Rummel, N., & Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. British Journal of Educational Technology, 51(5), 1527-1547.

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning?. British Journal of Educational Technology, 51(5), 1450-1484.

Xu, P., Zhu, X., & Clifton, D. A. (2023). Multimodal learning with transformers: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(10), 12113-12132.

聚焦创造性问题解决：跨学科项目化学习实施策略分析——以“许你一个不迷路的校园”两轮迭代为例

Focusing on Creative Problem-solving: Analysis of the Implementation Strategies of Interdisciplinary Project-based Learning

——A Case Study of Two Iterations of "Promise You a Campus without Getting Lost"

徐晨熠,孙紫怡,逯行
浙江省兰溪市聚仁小学
浙江师范大学
1362393906@qq.com

【摘要】跨学科项目化学习强调培养学生解决复杂真实问题的能力。然而通过调研发现，项目化学习存在过于强调项目和作品而忽视学习与思维过程、可视化成果浅表化、缺乏对作品的修改完善等问题。本研究以“许你一个不迷路的校园”跨学科项目化学习为例，将设计思维融入跨学科项目化学习的设计与实施，通过两轮行动研究展现融入设计思维的项目化学习模式在教学中的具体应用，提炼融入设计思维的项目化学习设计路径以及指导原则，并在教学实践不断改进该框架,探究该框架能否有效改善项目化教学实践困境，提升学生创造性问题解决能力。
【关键词】跨学科项目化学习；设计思维；定义问题；创造想法；测试评价

Abstract: *Interdisciplinary project-based learning emphasizes cultivating students' ability to solve complex real-world problems. However, research shows that it has issues like overemphasizing projects and works while neglecting the learning and thinking processes, superficial visual achievements, and lack of work improvement. Taking the interdisciplinary project-based learning of "A Campus Where You Won't Get Lost" as an example, this study integrates design thinking into its design and implementation, conducts two rounds of action research to show its application in teaching, extracts the design path and guiding principles, and explores whether the framework can effectively address the practical dilemmas of project-based teaching.*
Keywords: *Interdisciplinary project-based learning, Design thinking, Define the problem, Generate ideas, Test and evaluate*

1.前言

在项目化学习中，面对真实、有挑战性的问题，学生需要分析问题、构思解决问题的方案、并在不断探究和验证中来解决问题。这些都非常考验学生的问题解决、决策、批判等能力。国外的研究和实践表明，项目化学习可以提高学生的学习动力，培养学生批判性思维、问题解决等高阶思维能力和 21 世纪技能。但在项目化学习设计和实施中，如何培养学生的高阶思维能力？这一直是很多学校关注的重点。设计思维强调以人为本的价值理念，其开放性、多样性、联动性和综合性的特点，有助于培养学生的创新思维、提高学生的实践能力、增强学生的问题解决能力、增强学生的自我学习能力。然而，现有研究多聚焦单一学科或高等教育（如 Razzouk & Shute, 2012），基础教育领域尚未形成可推广的融

合路径。本研究将设计思维融入项目化学习，通过小学“许你一个不迷路的校园”跨学科项目为案例，通过两轮迭代的行动研究，探索以下问题：如何构建设计思维导向的跨学科PBL框架？该框架能否通过迭代反馈机制有效提升学生的创造性问题解决能力？通过理论创新与实践验证，本研究旨在为新课标下核心素养的落地提供可复制的策略模型。

2.设计路径

设计思维作为一种以用户为中心的问题解决方法论（Brown, 2008），其核心流程（共情、定义、构想、原型、测试）与PBL的“问题驱动—探究—成果迭代”过程高度契合。然而，现有研究多聚焦于单一学科或高等教育领域（如Razzouk & Shute, 2012），基础教育中设计思维与跨学科PBL的融合路径仍缺乏系统性探索。本研究基于EDIPT模型（斯坦福设计学院, 2010）与PBL的“黄金标准”（Larmer et al., 2015），构建“设计思维导向的跨学科PBL框架”，旨在通过两轮迭代的行动研究，探究该框架对学生创造性问题解决能力的影响。

2.1.项目化学习的理论溯源与实践挑战

项目化学习（Project-based Learning, PBL）是一种以学生为中心的教学模式，强调通过真实问题驱动学生开展跨学科探究，培养高阶思维能力。其核心流程包括“问题驱动—计划制定—实践探究—成果迭代—评价反馈”，与21世纪技能培养目标高度契合。然而，现有研究表明，PBL实践存在“成果导向”偏差，如过度关注可视化作品而忽视思维过程，评价体系碎片化，导致学生核心素养提升不足。

跨学科项目化学习阶段的系统设计和实施，能够将创造性问题解决能力的培育贯穿于入项、知识与能力建构、探索形成成果到出项的项目化学习全过程。这种结构化的操作路径不仅指向学生创新思维与实践能力的提升，更使其在真实问题解决中体验创造的价值。

在素养目标的设定上，教师要将学生创造性问题解决能力作为核心，从目标层面促使学生从多维视角思考和分析，为创造性解决问题奠定基础。在入项阶段，教师要创设激发创造性思维的真实情境并提出具有开放且挑战性强的驱动问题，驱动学生突破常规思维，提出创造性解决方案。在知识与能力建构阶段，教师要注重启发学生的批判性与发散性思维，通过头脑风暴、思维导图等支持性支架工具，鼓励学生多角度分析问题。在探索形成成果阶段，教师要鼓励学生试错与迭代，允许学生在方案设计中尝试多种路径，鼓励学生的个性化创新，支持其将创意转化为实际成果。在出项阶段，教师重在通过多样化的成果展示与评价，强化学生创造性问题解决的价值认同，让学生获得成功感。

2.2.设计思维的理论内涵与教育价值

设计思维模型是研究者为培养学习者的设计思维，进而提出的一系列方法步骤或操作环节，引导学习者依据设计思维模型的各个环节进行方案设计，最终形成设计制品。

设计思维（Design Thinking）起源于人工科学领域（Simon, 1969），其核心是以用户为中心的迭代式问题解决流程（Brown, 2008）。斯坦福大学提出的EDIPT模型（共情、定义、构想、原型、测试）为典型代表，强调通过多轮迭代优化解决方案（IDEO, 2012）。在教育领域，设计思维的价值体现在三方面：一是以人为本，通过共情挖掘真实需求，避免问题定义偏差；二是迭代优化，以原型测试为媒介，推动思维从发散到收敛；三是跨学科整合，融合技术、艺术与逻辑思维，促进综合能力发展（Kolodner, 2002）。

设计思维导向下的跨学科项目设计可以以流程化的形式展现问题解决过程，富有成效地参与想法的生成、评估和改进的过程，进一步提升学生创造性问题解决的能力。教师基于EDIPT设计思维模型设计并开展指向创造性问题解决能力提升的跨学科项目，总结下

来具有以下特征：第一，共情阶段即同理心阶段可以进一步帮助学生通过调查研究来了解用户各方面的需求；第二，定义阶段、构想阶段进一步帮助学生将所有想法收集起来，探索出宽阔的解决方案空间，体现发散思维；第三，原型阶段、测试阶段强调将脑海中的想法表达出来，并从原型中找出什么是可行的、什么是不可行的，根据反馈对原型进行修改，然后反复迭代，精炼和改善解决方案。

2.3.设计思维与 PBL 的融合路径

本研究将设计思维的关键特征与跨学科项目化学习深度融合，构建指向创造性问题解决能力的实践路径。

现有研究对二者融合的探索多聚焦高等教育（如工程与 STEM 领域），基础教育中缺乏系统性框架。本研究基于斯坦福 EDIPT 模型与 PBL“黄金标准”（Larmer et al., 2015），提出“设计思维导向的跨学科 PBL 框架”（图 1），其创新点在于：一是问题定义精准化，通过用户共情与场景模拟，提炼更具挑战性的驱动问题；二是流程动态化，将 PBL 的线性流程重构为“定义—创造—测试—迭代”循环模式；三是评价系统化，引入多维度反馈机制（用户评价、自评互评、过程性记录），破解成果浅表化困境。

入项阶段——定义问题：引导学生以用户思维定义问题，教师可以通过提出开放性问题，引导学生思考问题。例如，“作为学校的小主人，我们要如何为初入聚仁园的老师、家长设计一套既有特色又能指明方向的校园导视牌呢？”鼓励学生站在初入校园的教师和家长视角思考问题。

知识与能力建构阶段——创造想法：引导学生进行调查研究，寻找相关信息，将用户需求转换为更深层次的用户需求并描述解决方案；进行头脑风暴，收集各种解决问题的想法，激励学生思考问题，并产生创造性的解决方案。

探索形成成果阶段——原型测试：教师可以让学生选择并设计自己喜欢的方案，并制作原型。例如，学生们可将自己想象出的校园指示牌做成模型，展示出来并向其他学生介绍，根据反馈信息进行现场调试与迭代改进，精炼和改善解决方案。

出项阶段——成果交流：通过迭代反馈结合自评和他评，从多角度获取意见，不断修正完善作品或方案，培养学生的批判性思维和问题解决能力。

图 1 设计思维导向下跨学科项目化学习实施路径

3.行动研究

本研究采用个案研究法，以浙江省兰溪市聚仁小学的“许你一个不迷路的校园”跨学科项目化学习的两轮迭代为例，采用设计方案制品分析、学生发生回忆、学生现场体验差异性访谈等方式，在多层次了解项目前后学生学习体验等数据的基础上，探查学生在设计思维导向下的创造性问题解决能力的发展。

本研究采用威廉斯创造力倾向量表（Williams Creativity Assessment Packet, WCAP）评估学生的创造性思维发展。该量表包含四个维度：冒险性（尝试新方法的意愿）、好奇心（探索未知的主动性）、想象力（构思新颖想法的能力）、挑战性（解决复杂问题的韧性），采用 5 点李克特量表（1=完全不符合，5=完全符合）。量表信效度良好（Cronbach’s α=0.82）。

3.1.第一轮行动研究

学校随机抽选四年级中的 20 名学生为样本，利用周三拓展课时段开展了为期 8 课时的“许你一个不迷路的校园”跨学科项目化学习。通过学生在前测、后测中体现的创造力倾向变化的对比来衡量设计思维导向下的跨学科项目化学习对学生创造性思维的影响。

表 1“许你一个不迷路的校园”跨学科项目化学习（第一轮设计）

阶段	活动内容
第一阶段 导视方案制定	学生凭借性格自评思维工具筛选合作伙伴，组建团队后召开项目招标会，签订项目合同书并确立团队文化。借助情境预设与师生交流，精准提炼出驱动性问题：作为学校的小主人，我们要如何为初入聚仁园的老师、家长设计一套既有特色又能指明方向的校园导视牌呢？之后学生围绕该问题开展 5-2-5“扫描型”头脑风暴，运用思维导图和“问题树”工具制定小组方案，并持续优化迭代。
第二阶段 校园导视大揭秘	学生自主探究与教师引导相结合，全面了解导视牌的作用、安放位置及结构组成。继而模拟外来人员实地考察校园，罗列问题清单，绘制校园鸟瞰图，为导视牌的科学布局提供有力依据。
第三阶段 校园导视我设计	学生依据认领位置，细致分析周边环境，确定导视牌的朝向及空间位置关系。通过参观校史馆、访谈校长，深入挖掘校园文化特色，思考文化元素与导视牌设计的有机融合。综合考虑多方面因素，制定导视牌设计方案，明确版面与外形，绘制详细设计图，并借助小组展示与“旋转标签”进行评审活动，收集反馈建议并通过复盘改进和优化方案，最终形成设计终稿。
第四阶段 校园导视我制作	学生广泛调研市面上制作原材料，对比分析后择优挑选适配校园环境的材料。前往指定地点实地测量，确定导视牌最佳高度。运用木板、泡沫板等材料在学校木工坊制作模型，验证设计方案的可行性，并进行现场调试与迭代改进。
第五阶段 校园导视我展示	将制作好的导视牌模型放置于校园内实际位置，组织全校师生观摩，运用“问卷星”和“成果优化表”收集反馈意见。学生结合反馈意见开展小组交流分享，进一步改进成品设计。最后对项目进行

	全面复盘分析，总结经验教训，形成完整详尽的项目终结报告并递交给学校，展望项目后续应用与拓展。

本轮行动重点关注“定义问题”和“创造想法”两个步骤。教师提出了“作为学校的小主人，我们要如何为初入聚仁园的老师、家长设计一套既有特色又能指明方向的校园导视牌呢？”的驱动性问题，观察结果显示，学生能积极以用户思维——模拟外来人员实地考察校园，罗列问题清单，绘制校园鸟瞰图，为导视牌的科学布局提供有力依据。课堂上，学生的主动提问与有效回答次数增加。但校园导视牌的固定成果形式本身就是对学生创造性思维发展的一种限制，尤其体现在学生在校园指示牌的设计制作阶段，学生在成果迭代中仍存在知识理解机械死板、合作不足、想象力不足等问题，下一轮研究将对此加以改进。

第一轮行动研究中，学生创造力各维度均有所提升，但仅冒险性（p=0.028）、好奇心（p=0.008）、挑战性（p=0.004）达到显著水平（p<0.05），而想象力提升不显著（p=0.061），可能与传统导视牌设计的成果形式限制有关。

表 2 第一轮行动研究前后测创造力指数对比

维度	前测均值（标准差）	后测均值（标准差）	t 值	p 值
冒险性	3.10（0.42）	3.45（0.38）	2.34	0.028
好奇心	3.25（0.51）	3.60（0.47）	2.89	0.008
想象力	3.05（0.39）	3.30（0.41）	1.98	0.061
挑战性	3.20（0.44）	3.55（0.40）	3.12	0.004

3.2. *第二轮行动研究*

针对上一轮中的问题，本轮行动注重以评促学、迭代反馈，提升学生批判创新能力。第一轮项目聚焦于传统导视牌的设计与制作，解决外来人员及新生在校园内迷路的问题；第二轮项目则基于传统导视牌在实际使用中的不足，进一步提出“如何设计并制作一套校园电子指示系统”的驱动性问题，问题来源于现实需求的深化，更具挑战性和前沿性，促使学生思考如何利用信息技术解决传统方式无法有效解决的问题。第一轮项目主要涉及对校园环境的观察、分析以及静态导视牌的设计制作；第二轮项目要求学生设计动态的、交互式的电子导视系统，需要学生具备系统思维、逻辑思维以及技术创新思维等更高层次的思维能力，从整体上规划系统的功能、界面和操作流程，并考虑如何通过信息技术实现这些功能。

表 3“许你一个不迷路的校园”跨学科项目化学习（第二轮设计）

阶段	活动内容
入项过程	一开始结合问卷调查数据，带领项目组的学生实地观察，发现校园指示牌存在指示不够清晰、无法连续指引、缺少交互等问题，由此引出设计校园电子指示系统的驱动任务。接着，通过“同理星”挑战进行破冰组队，明确个人表现与优秀团队的评价内容。最后，根据团队产品目标进行任务阶段性分解，确保项目顺利开展。

知识与能力建构	基于问卷调查和学生实地考察的结果，学生讨论确定外来人员常去且容易迷路的关键位置（亲子工坊、荟仁园、木工坊、大门口、智创中心、艺创中心、报告厅），并根据摄影要求，拍照取点。随后将素材导入 Scratch 编程软件，以团队为单位确定指路人形象、初始坐标、大小、朝向等。在这个过程中，教师提供包括示范视频及代码在内的能量包，每个小组运用循环结构和自制积木技术编码，负责从一个点到其余六个点的路线设计。
探索与形成成果	在调试阶段，学生依据路线设计再次实地考察，针对实际问题进行再改进、优化，避免与实际脱轨。优化时，学生将问题分类，依托“能量包”从数学、技术、综合三方面迭代：用数学知识优化路线距离，借技术手段提升行走精准度，靠综合能量包完善画面和语音系统。
评价与修订	教师带领学生共同讨论，制定展示评价量规。学生在产品发布会上分享设计理念并现场测试。
公开成果	召开微型发布会，让全校师生来观摩、使用，从旁观者、使用者的角度去发现问题，从而收集反馈，再次迭代产品。最后向学校递交产品设计方案与说明书，以便学校后期投入使用。

本轮行动重点关注“原型测试”和“成果交流”两个步骤。迭代后的驱动性问题“如何设计并制作一套校园电子指示系统”给予学生更高的挑战。由于在这一过程中，学生不仅需要运用信息技术，还需要运用数学知识进行坐标定位和距离计算，运用美术知识进行界面布局和视觉设计等，多学科知识在电子导视系统的设计与制作中相互关联、协同应用，学习者参与设计的过程涉及多项高阶能力，例如批判创新、自主学习、团队协作等，这些能力成为了评价维度的组成部分。评价以自评和互评的方式进行，主要指向设计成果的产出以及学生能力和思维的发展。

第二轮行动研究中，学生创造力各维度提升显著（$p<0.001$），尤其是想象力（均值提升 0.85）与挑战性（均值提升 0.70），表明设计思维驱动的“电子导视系统”项目通过多轮迭代与跨学科协作，有效激发了学生的创新潜能。

表 4 第二轮行动研究前后测创造力指数对比

维度	前测均值（标准差）	后测均值（标准差）	t 值	p 值
冒险性	3.40（0.35）	4.10（0.32）	4.56	<0.001
好奇心	3.55（0.40）	4.25（0.36）	5.12	<0.001
想象力	3.30（0.38）	4.15（0.34）	4.98	<0.001
挑战性	3.50（0.42）	4.20（0.39）	4.75	<0.001

两轮行动研究数据表明：

1. 迭代反馈机制显著提升创造力：第二轮项目中，学生想象力（$t=4.89, p<0.001$）与挑战性（$t=4.75, p<0.001$）维度得分增长显著高于第一轮；

2. 跨学科协作促进综合能力发展：电子导视系统设计中，学生需融合编程、数学与美术知识，其创造力指数提升幅度（均值$\Delta=0.75$）显著高于传统项目（均值$\Delta=0.35$）；

3. 用户共情驱动问题定义精准化：通过实地调研与原型测试，学生方案的用户满意度从 68%提升至 89%（$\chi^2=12.3, p=0.001$）。

4.研究结论

4.1.*提炼策略模型，强化理论贡献*

本研究基于两轮迭代的行动研究，提出“设计思维导向的跨学科 PBL 四阶段模型”，包含“定义真实问题—多学科协作创造—原型迭代测试—成果迁移应用”四个动态循环阶段。该模型的创新性体现在：

第一，问题定义的精准性：通过用户共情（如模拟外来人员考察校园）与场景分析，提炼更具挑战性的驱动问题（如“电子导视系统设计”），使问题解决目标更贴近真实需求；

第二，流程的迭代性：将传统 PBL 的线性流程重构为“测试—反馈—再设计”循环模式，学生需经历多轮原型优化（如第一轮迭代 3 次，第二轮迭代 5 次），显著提升解决方案的创新性与可行性；

第三，评价的系统性：引入多维度反馈机制（用户评价、自评互评、过程性档案袋），破解 PBL“成果浅表化”困境，促进高阶思维发展。

4.2.*数据驱动结论，增强实证支撑*

两轮行动研究的数据表明，设计思维与跨学科 PBL 的深度融合显著提升了学生的创造性问题解决能力：

创造力指数的显著提升：第二轮项目中，学生威廉斯创造力倾向量表的想象力（$t=4.89, p<0.001$）与挑战性（$t=4.75, p<0.001$）维度得分增长显著高于第一轮（均值提升$\Delta=0.85$ vs. $\Delta=0.25$），表明迭代反馈机制对创新潜能具有累积效应；

跨学科协作的协同效应：电子导视系统设计中，学生需融合编程、数学与美术知识，其创造力指数提升幅度（均值$\Delta=0.75$）显著高于传统项目（均值$\Delta=0.35$），且小组合作效率提升 23%（通过任务完成时间与方案完整度评估）；

用户共情驱动问题解决精准性：通过实地调研与原型测试，学生方案的最终用户满意度从 68%提升至 89%（$\chi^2=12.3, p=0.001$），验证了“以人为本”设计理念的实际价值。

4.3.强调理论价值与实践意义

本研究填补了基础教育中设计思维与 PBL 融合研究的空白，其理论贡献在于：

第一，构建动态迭代框架：突破传统 PBL 的线性结构，为“以评促学”提供可操作路径，呼应了夏雪梅（2021）提出的“学习实践”理念；

第二，拓展设计思维的应用场景：将 EDIPT 模型从高等教育延伸至小学跨学科项目，验证了其在低龄学生群体中的适用性，为乡村教育或 STEM 课程设计提供参考；

第三，推动核心素养落地：通过真实问题解决与多学科整合，促进批判性思维、系统思维与创新思维协同发展，助力新课标目标实现。

未来研究可进一步探索：人工智能技术支持下的个性化反馈系统（如基于大数据的迭代优化建议），或设计思维在不同文化背景（如乡村学校）中的差异化实施路径。

引用文献

巴克教育研究所（2015）。《项目学习教师指南——21 世纪的中学教学法（第二版）》（任伟译）。北京：教育出版社。
李伟（2021）。《项目化学习在工科教学中的实践困境与优化策略》。《高等教育研究》，(6)，59–67。
林琳，沈书生（2016）。《设计思维的内涵梳理与教育应用路径》。《教育技术研究》，(6)，18–25。
夏雪梅（2021）。《学习实践：项目化学习的重构》。北京：教育科学出版社。
张洪波等（2020）。《STEM 教育理念下的项目化学习简化模型》。《现代教育技术》，(20)，56–58。
周毅等（2022）。《“新文科”背景下项目化学习的实施路径与启示》。《高等教育探索》，(4)，1–7。
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：北京师范大学出版社。
Brown, T.(2008). Design Thinking. Harvard Business Review, 86(6), 84–92.
Dunne, D., & Martin, R. (2006). Design Thinking and How It Will Change Management Education: An Interview and Discussion. Academy of Management Learning & Education, 5(4), 512–523.

基于多智能体的科学问题解决能力评测模型自动构建方法

Research on the Design and Application of Pair Programming Scripts for Primary School Students: Based on the Perspective of Productive Failure

陈小汾
江南大学人文学院，江苏无锡 214122
3174464009@qq.com

[摘要] 随着生成式大模型技术的快速发展，大模型驱动的多智能体在教育领域已获得广泛应用。将多智能体应用于科学问题解决能力的评测，有助于处理多模态数据、输出高效评测结果和提供个性化学习路径。本研究面向中小学科学教育场景，基于大模型的通用智能代理框架进行适配和改造，设计了科学问题解决能力评测的智能代理框架和多智能体代理自组织协同框架。通过实例示范，定义了智能代理的角色，展示了多智能代理的协同工作流，并详细描述了各代理的工作职责和流程。
[关键词] 科学问题解决能力；教育评测；智能代理

Abstract: *With the rapid development of generative large language model technology, multi-agent systems driven by large models have been widely applied in the field of education. Applying multi-agent systems to the evaluation of scientific problem-solving abilities helps to process multi-modal data, output efficient evaluation results, and provide personalized learning paths. This study, targeting primary and secondary school science education scenarios, adapted and modified a general intelligent agent framework based on large models, designing an intelligent agent framework for evaluating scientific problem-solving abilities and a multi-agent self-organizing collaborative framework. Through example demonstrations, the roles of intelligent agents are defined, the collaborative workflow of multiple intelligent agents is showcased, and the work responsibilities and processes of each agent are described in detail.*
Keywords: scientific problem-solving skills, educational assessment, AI agent

1.前言

科学问题解决能力是个体为实现某个目标，综合运用科学知识、探究技能和推理策略等，有计划地进行认知活动的能力。培养学生的科学问题解决能力已成为国际科学教育的核心目标，亦是我国科学素养培育的重要主题。2025 年 1 月，教育部颁布《中小学科学教育工作指南》（以下简称《工作指南》）中明确指出，需在探究实践中强化对学生问题提出、实验设计、动手操作、思维发展等能力的考察，重点评测学生利用知识、方法，分析问题和创造性解决问题的能力。然而，通过一线科学教师深入访谈发现，经验丰富的教师能够感知学生的问题解决能力，但这种感知往往属于默会知识，难以转化为有效的评测因素。学生在课堂中展现出的问题解决能力是可见可捕捉的，但如何将这些现象转化为可测量、可操作的实践，目前仍存在探索空间。

随着人工智能技术的日渐发展和广泛渗透到教育活动中，其有望为教育评测带来更高效、更客观的策略方案，以弥补传统评测手段的不足。《工作指南》强调，开拓生成式人工智能大模型在科学教育中应用的新场景，利用数据分析技术提升教学评测的精准化水平。2020 年 10 月，国务院印发的《深化新时代教育评测改革总体方案》中也提出充分利用信息技术提升教育评测的科学性、专业性、客观性。当前，智能代理（AI 智能代理）在教育评测应用中的潜力被越来越多的教育研究者和实践者广泛关注。智能代理被定义为通过

观察环境并利用其可用的工具对环境采取行动以尝试实现预设目标的应用程序。研究表明，基于大模型的单智能代理在实时分析学习数据、识别知识盲点、提供针对性练习和反馈等方面具有显著优势。然而，单智能代理在处理多角色协作的复杂教育任务时存在局限性，如认知视角单一、角色定位固定等，难以满足深层次的教育需求。

为解决上述问题，研究者们提出了大模型驱动的教育多智能体系统解决方案，旨在通过多个智能体的协作来突破单智能体的能力限制。多智能体系统通过整合不同角色定位的智能体，构建起一个动态协作的智能教学团队，能够激发出超越单智能体的集体智慧，有效应对复杂教育任务的挑战。实证研究进一步验证了多智能体系统的优势，翟雪松等人（2024）[xvii]基于微软 AutoGen 多智能体框架搭建人机交互平台，设计由七个智能体组成的教研团队作为“智能代理”角色，包括教师、学习者、教育心理学专家、教育社会学专家、教育经济学专家、教育评测专家、教育政策制定者，学生作为“人类”角色，与七个智能体角色基于问题开展集体研讨，结果表明，相比单智能体模式，多智能体能够显著提升学习者提问策略的多样性，有效提高复杂问题的解决效率。Yang 等人（2024）[xviii]提出 LLM 智能代理-CK 框架，由三种类型的 LLM 驱动的代理（即管理员、审判员、评论员）和两种控制策略（即讨论策略和决策策略）构成，通过智能体之间的协作与讨论，实现对教师数学内容知识的识别。Lagakis 等人（2024）[xix]利用大语言模型（LLMs）和 AutoGen 框架开发了一个基于多智能体架构的自动化评分系统，该系统包含管理员代理、评估代理和反馈代理，它们通过讨论和决策策略共同完成评分任务，被用于大规模在线开放课程（MOOCs）中的作业评分。

基于上述分析，本研究通过探讨多智能体系统赋能科学问题解决能力评测中的范式转变、内容创造、场景应用的可能性，构建基于大模型的科学问题解决能力多智能评测系统框架，并将其应用于科学问题解决能力评估实践，以期推动智能代理在教育评测中的创新应用，为科学问题解决能力评测提供技术路线与实践参考。

2.多智能体系统赋能评测范式转变

2.1 处理多模态评测数据

传统评测数据收集方法主要集中于文本形式的答题结果，这种单一数据形式难以全面捕捉分析被评测者在问题解决过程中的丰富信息。基于大模型的智能代理具有强大的理解与分析复杂语言结构的能力，已具备生成文本、图像、音频、视频及 3D 模型等多种模态内容的处理能力。在多智能体系统中，每个智能体能够独立感知并处理来自不同来源的数据，灵活调用外部工具来增强自身数据处理和分析能力，共享信息平台允许各智能体进行数据交换和同步，整合不同来源和模态的数据，共同协商形成一个统一的、全面的能力评测结果。

2.2 输出高效化评测结果

传统评测结果的输出需要花费较长时间进行批改和评分，无法及时给学生反馈，教师教学效果也无法实现实时评测，导致教学目标和流程的优化呈现滞后性。而多智能体系统可以实现实时评估，全程化高效率地进行过程性评测、增值性评测，通过即时生成的语言输出，快速给出反馈，让学生及时了解自己的不足之处，调整学习策略；帮助教师和相关部门及时了解课程的优势和不足，从而促进课程内容和教学方法的改进。

2.3 提供个性化学习路径

在学生评测中，多智能体系统基于大数据算法对每个学生的阶段性能力评测结果进行精准分析，定位学生当前科学问题解决能力所处水平。从发展性的角度将学生的总体科学

问题解决能力目标分解为具体的阶段性目标，识别学生在科学问题解决能力上的优势和需要改进的地方。通过分层分类的引导策略，系统帮助学生设定合理的阶段性学习目标，并根据这些目标提供定制化的学习内容和指导，促使学生在努力达成阶段性目标的过程中提高学习效率和成就感，实现评测从物到人、从外向内、从短期到长期、从结果向过程的转变。

3.基于大模型的通用智能代理框架

3.1 基本概念

在人工智能领域，智能体需要能够感知其外部环境并做出行动，以对外部环境产生影响。通常情况下，智能体与外部环境间的感知与行动不断循环，形成密切交互，以完成具体任务目标。图 1 为基于基于大模型的通用智能代理框架（Xi et al, 2025）[xx]，由大脑、感知、行动三个关键部分组成。其中，大脑又称为控制模块，由一个大型语言模型组成，不仅存储知识和记忆，还承担着信息处理和决策等功能，并可以呈现推理和规划的过程，能很好地应对未知任务。感知模块的核心目的是将智能代理的感知空间从纯文字领域扩展到包括文字、听觉和视觉模式在内的多模态领域。当一个智能代理拥有类似大脑的结构，具备知识、记忆、推理、规划和概括能力以及多模态感知能力时，它也有望拥有类似人类的各种行动来应对周围环境。在智能代理的构建过程中，行动模块负责接收大脑模块发送的行动序列，并执行与环境互动的行动。

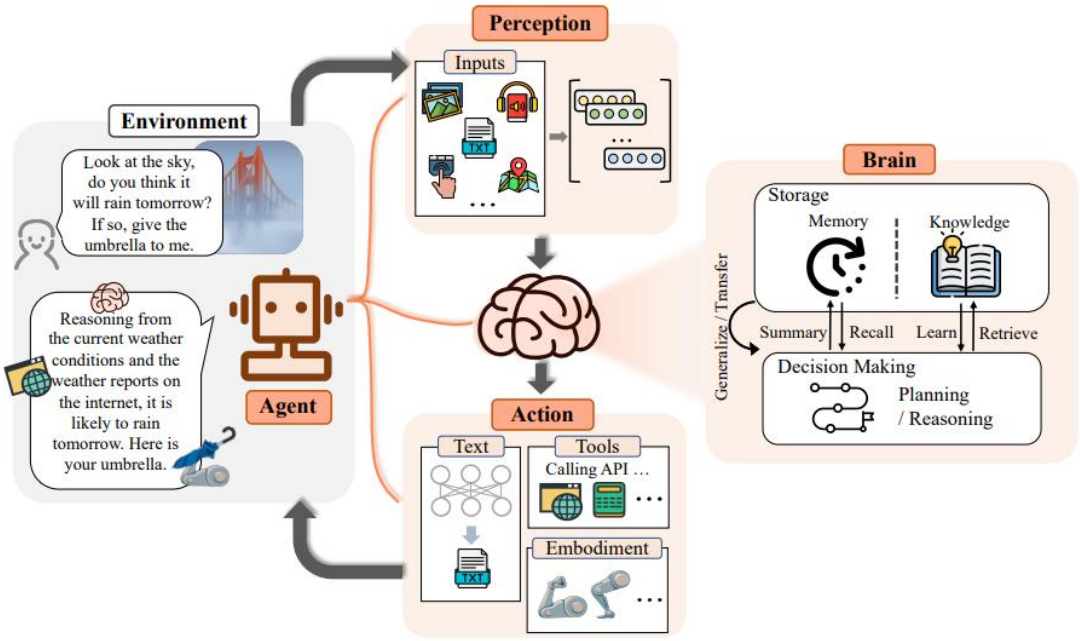


图 1 基于大模型的通用智能代理框架

多智能系统是由多个互相协作或竞争的自治智能体组成的系统，旨在通过集体行为解决复杂问题。系统中每个智能体基于通用智能体框架，分别配置感知模块、控制模块和行动模块，并通过独立或联合的训练机制实现功能优化，从而具备一定的自主性与适应性。如图 2，多智能代理协作框架的核心是如何实现智能代理之间的协作和竞争的平衡，即如何使每个智能代理都能达到自己的目标，同时也能促进整个系统的性能和效益。为了实现这一目标，多智能代理协作框架需要解决以下几个关键的问题:

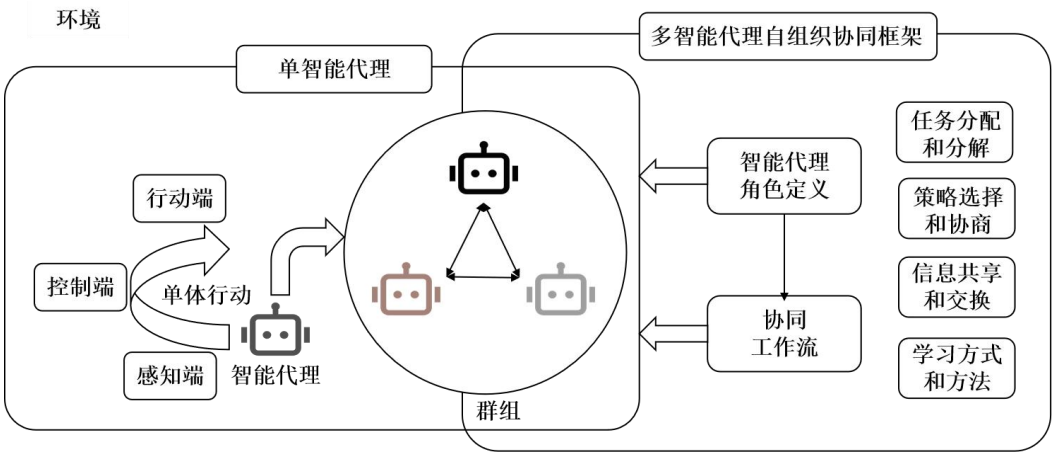


图 2 多智能体代理自组织协同框架

（1）智能代理的角色画像设计。定义智能代理的属性和行为，如目标、偏好、策略、动作、感知、学习和沟通等。角色定义采用多个单工具的智能体策略，针对任务目标各子任务，按照各要素子任务分配对应的工具，结合工具的文档说明对智能体进行训练，使智能体达到理解工具或独立调用接口等目标。

（2）任务分配和分解。将一个复杂的任务分配和分解给多个智能代理，使得每个智能代理都能承担合适的子任务，同时也能保证任务的完整性和一致性。任务分配阶段通过提示词进行任务的明确和分解，形成完成复杂任务的思维链，提取思维链中各序列化的节点作为子任务，形成任务序列清单。根据各智能体角色分工将各子任务分发下达给各智能体。

（3）动作处置和规划。将高层次的任务目标分解为可执行的低层次动作，并规划反思动作的执行顺序，通过环境感知反馈来进一步适应多层级搜索，以实现整体目标。在此阶段，多个智能体获得初始任务，通过多轮与环境的交互，优化智能体所分配的子任务和执行顺序，得到最优的规划。

（4）智能代理的协同工作流。设计智能代理之间的交互和协调的机制和协议，如定义工作流程、交互信息等。

4.科学问题解决能力评测的多智能代理框架

4.1 科学问题解决能力评测智能代理框架

针对图 1 所示的基于大模型的通用智能代理框架经适配与改造，用于科学问题解决能力的评测，如图 3 所示。

（1）感知模块。针对科学课堂活动中数据来源众多（研学单、实验操作过程记录、师生互动对话、小组讨论对话等）、数据类型模态多样（非结构化文本、图像、音频和视频流等）等问题，教师在学生科学问题解决能力评测中难以兼顾到每位学生，也难以从多源异构数据中提炼出一致的能力评估证据。感知模块将智能代理的感知空间，从纯文本拓展到包括文本、视频、图片和语音等多模态领域，使智能代理能够更有效地从复杂教育场景中获取与利用信息，突破多源异构数据信息识别抽取技术、多模态信息统一表征认知模型等关键技术，以支持对跨模态、多样式的要素信息的识别、抽取，多模态知识获取、表示与推理，形成多模态的统一语义表示，提升智能代理在多模态信息环境中的智能感知能力。

（2）控制模块。为处置评测任务进行处置动作规划，它可以决策处置动作需要运用哪些操作指令来完成，包括记忆、结构感知知识和处置动作规划 3 部分。记忆部分，存储和管理与科学问题解决相关的概念和知识、记录学生在实验中的具体操作和情境，保存学生在

解决问题过程中的思考路径和决策过程，进而支持智能代理的反思和进化。知识部分，通过动态图结构归纳自然语言文本和领域知识库的语义表示，强化智能代理的认知能力。包含评测框架、评估案例、科学学科核心概念集以及实验知识与规范操作，帮助智能代理深入理解学生的实验设计和操作是否符合科学标准。动作规划部分，根据记忆和结构感知知识，依托科学教育动作库和科学知识体系规划反馈，将复杂的评测任务分解为具体的子任务，并规划出最优的处置动作，生成评测报告和反馈建议。控制模块通过总结、召回、学习和检索等机制，将记忆模块、知识部分与处置动作规划协同工作。总结机制提炼记忆中的关键信息，形成结构化的知识表示，召回机制根据当前任务需求，提取相关的历史信息和案例，学习机制不断更新记忆和知识库，提升智能代理的认知能力和评测准确性，检索机制则确保智能代理能迅速获取所需信息，为当前任务提供及时支持。

（3）行动模块。负责将控制模块的规划结果转为具体行动，以实现对学生科学问题解决能力的智能评测和反馈，并输出包含文字和可视化图表的评测报告。包括检索器和自主调用模型；其中检索器根据输入的操作指令从数据集中检索与操作指令相关的接口，帮助模型过滤无关信息，提高模型的训练效率；自主调用模型，根据操作指令与接口进行多轮交互，并返回交互结果，通过教师与多智能体多轮对话中教师对评测结果提供的质疑进行修改和调整，动态评测报告内容。可视化模块，用于将收集到的学生表现数据以柱状图、折线图等方式直观展示，呈现学生在不同科学知识点或者技能点上的掌握程度以及在问题解决过程中的进步情况。行动模块在执行过程中需与感知模块和控制模块紧密协作，确保准确地接收感知模块处理后的结构化数据和关键信息，并按照控制模块的规划结果采取相应行动，同时将执行过程中的状态和结果及时反馈给控制模块，以便其根据实际情况调整规划策略。

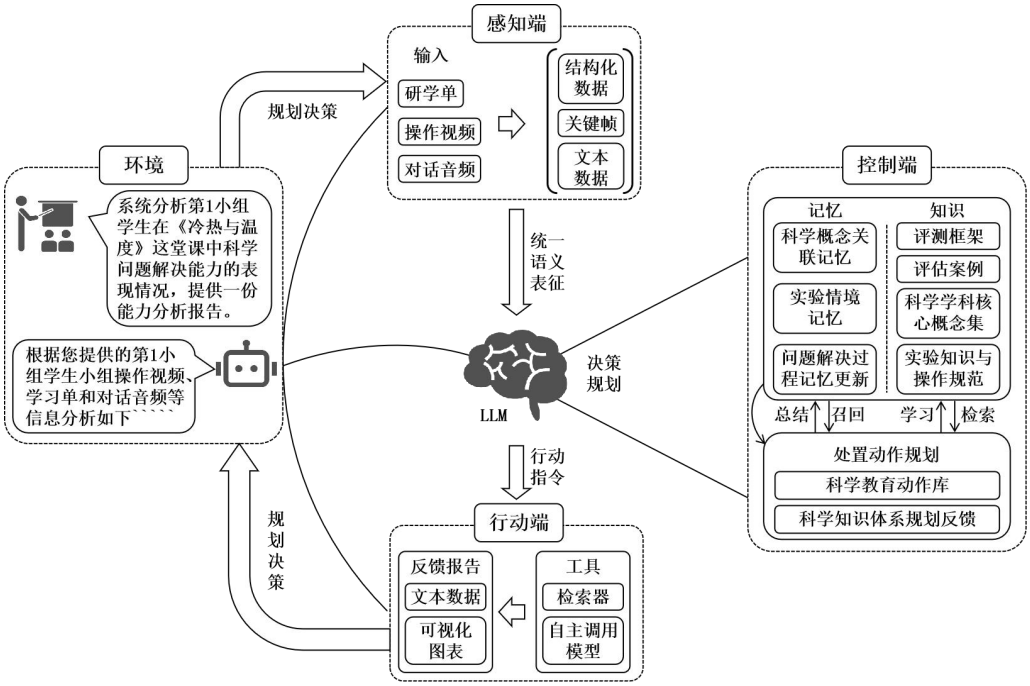


图 3 科学问题解决能力评测智能代理框架

4.2 科学解决能力多智能代理自组织协同评测运行框架

面向科学课堂中学生科学问题解决能力的评测，以图 3 科学问题解决能力评测智能代理框架为底本设计管理员和各模态分析员的任务和角色，其中，各模态分析员还需要进行差异化的训练，具体见表 1。

表 1 各模态分析员差异化设计训练

角色名称	感知模块	控制模块	行动模块
视频分析员	输入：操作视频流 技术：VideoPrism 模型提取时空特征，OpenPose 检测协作姿态 输出：操作步骤序列、异常动作标记	基于规则引擎（实验步骤校验）与深度学习（动作意图识别）混合架构 动态权重调整：根据任务类型分配检测重点	生成带时间戳的异常报告
语音分析员	输入：ASR 转写的讨论文本 技术：BERT+逻辑链检测模型 输出：假设提出次数、反驳语句占比	分层决策：进行初级过滤，去噪无关对话，后面在进行深度分析 集成动态记忆网络（DMN）记录讨论演进过程	生成讨论质量雷达图 触发知识图谱验证（如术语使用准确性）
文本分析员	输入：OCR 识别的研学单文本+手写公式 技术：结构化解析引擎 输出：数据完整性评分、结论逻辑错误	轻量级规则引擎，用以校验必填字段 与视频分析智能体协同验证时间一致性	输出结构化校验结果 标记需人工复核的模糊内容

见图 4，科学解决能力多智能代理自组织协同框架运行框架可划分为评测任务分解、评测数据自分析和评测结果整合四个实现阶段。首先，评测任务分解阶段，管理员智能体能够单独根据教师的作业意图和任务指令，理解当前学生所处具体科学问题情境，自动选择不同模态分析的分析员参与下一阶段。其次在评测数据自分析阶段，各模态分析员在在数据分析阶段，根据自动抽取与所负责数据的信息，通过多智能体分组讨论，执行评测框架中的能力维度映射任务，得到讨论结果和分析信息集合。最终在评测结果整合阶段，对讨论结果投票，将讨论结果映射为能力的高、中、低三类，并分别整合支持能力评测结果为高、中、低的三类分析信息，得到最终学生能力结果及其报告。

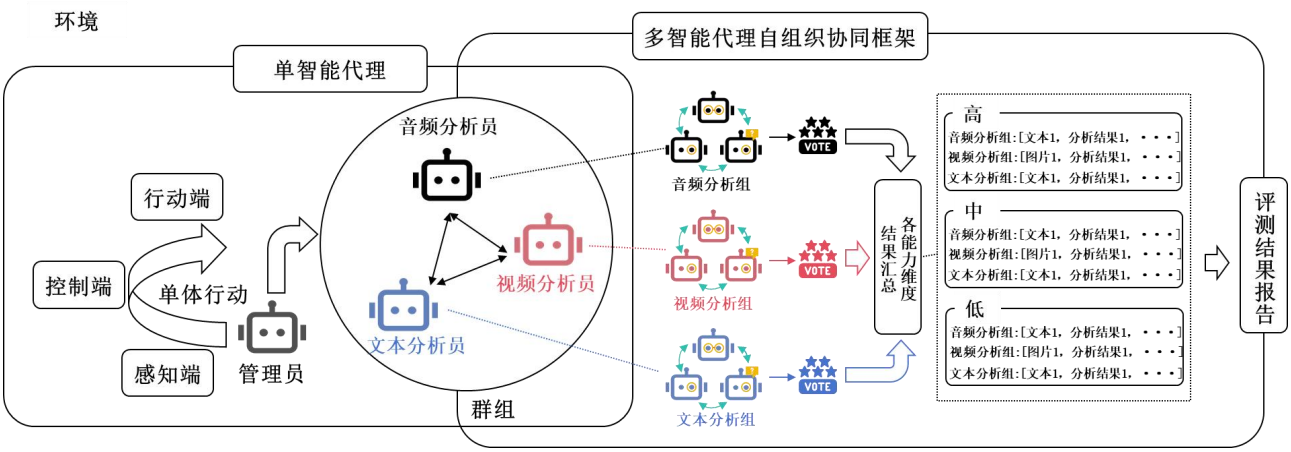


图 4 科学解决能力多智能代理自组织协同框架

（1）智能体角色画像设计

科学解决能力多智能代理自组织协同框架框架采用预先设置与模型生成相结合的方式设置智能体角色画像。该框架中每个模态讨论小组包括 3 个智能体，称为智能体 A、B、C。其中智能体 A 的角色画像通过大语言模型针对科学问题解决能力的评测框架生成；智能体 B、C 的角色画像设计根据讨论策略由人工给出。该方法在满足讨论策略需要的基础上，尽可能提高了智能体角色画像的多样性和专业性，以分化角色意见并提高知识调用的准确性。

5.各能力维度结果汇总

在评测讨论结果整合部分，要求全面考虑智能体间的不同观点，降低多智能体分组讨论过程中错误传播和智能体判断错误对最终分类结果造成的影响。

6.科学问题解决能力评测的多智能代理框架设计实现

为支持基于大模型的多智能体落地实现，目前已经有多个工程框架被开发并开源，例如 Auto Gen[xxi]、LangGraph[xxii]、MetaGPT[xxiii]、Google ADK[xxiv]等。这些框架为研究人员和开发者提供了重要资源，便于开发和测试多智能体的多场景应用。在上述实现框架中，Auto Gen 支持用户依据需求，灵活定义多智能体之间的交互模式与人机协作模式，例如由一位智能体主持、人工参与多智能体交互的动态群组讨论模式，或由两个智能体分别负责编码与执行调试的协作编码模式等。Auto Gen 可支持多智能体的交互记忆读写，可通过调用 Python 第三方工具包实现工具的使用（例如调用 Matplotlib 绘图库完成数学制图），并且支持将任务转化为机器语言解决，例如利用代码分步执行任务，并通过智能体间的代码执行与调试（Debugging）确保程序的成功运行。LangGraph 是 LangChain 生态系统中的一个重要工具，它提供了一个基于图（Graph）的框架来构建复杂的 LLM 应用。通过将应用逻辑组织成有向图 Q 的形式，LangGraph 可以让构建复杂的对话流程变得更加直观和灵活。

Auto Gen、LangGraph 等开发框架均为多智能体系统实现提供了可行方案，还可以通过配合使用发挥二者各自优势例如，可以利用 Auto Gen 灵活构建并实现智能体的交互框架与基于机器语言的任务执行，并利用 LangGraph 协助连接外部丰富的工具库（例如 Ar Xiv、Office365、Wolfram Alpha 等）以及自定义工具（通过用户提供工具功能描述、方法实现代码、输入输出格式等信息实现），从而拓展智能体的能力边界。

本研究选择基于 Auto Gen 和 LangGraph 框架进行评测工作流的搭建。如图 5，管理员在获取学生的多源数据后，进行深入分析，制定详细的组织处置方案。管理员将结构化的任务分配给各模态智能代理分析员，确保每个代理清楚自己的职责和任务。各智能代理根据分配的任务进行工作。在执行过程中，智能代理通过共享信息平台进行数据交换和状态同步，确保任务执行的连贯性和一致性。为了提升系统的适应性和响应能力，多智能体系统采用强化学习技术，为每个智能代理训练出独立的分布式决策策略网络，实现一定程度的自协同。此外，系统借鉴教育评测中的形成性评测理念，通过动态调整任务分配和执行顺序，不断优化工作流，确保评测过程的高效性和结果的可靠性。最终，各智能代理的执行结果被整合，最终输出生成详细的评测报告，为教师提供有力的支持，帮助他们更好地理解和指导学生的科学问题解决过程。具体工作流见图 4。

图 4《冷热与温度》课程第 1 小组科学问题解决能力的工作流示意图

7.结论

当下科学问题解决能力评测存在时效性差、分析周期长而影响教师对学生学习情况的及时掌握和教学策略的调整，亟需引入新技术、新方法、新框架推进科学问题解决能力智能分析。本研究基于大模型的通用智能代理框架，设计了科学问题解决能力评测的智能代

理框架和多智能体代理自组织协同框架。通过多智能体系统的协作，实现了对科学问题解决能力的自动评测，并提供了一个示范实例。

参考文献

翟雪松,季爽,焦丽珍,朱强 & 王丽英.(2024).基于多智能体的人机协同解决复杂学习问题实证研究.开放教育研究,30(03),63-73.doi:10.13966/j.cnki.kfjyyj.2024.03.007.

Yang, K., Chu, Y., Darwin, T., Han, A., Li, H., Wen, H., ... & Liu, H. (2024, July). Content knowledge identification with multi-agent large language models (LLMs). In International Conference on Artificial Intelligence in Education (pp. 284-292). Cham: Springer Nature Switzerland.

Lagakis, P., & Demetriadis, S. (2024, June). EvaAI: a multi-agent framework leveraging large language models for enhanced automated grading. In International Conference on Intelligent Tutoring Systems (pp. 378-385). Cham: Springer Nature Switzerland.

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ... & Gui, T. (2025). The rise and potential of large language model based Agents: A survey. Science China Information Sciences, 68(2), 121101.

Yang, H., Yue, S., & He, Y. (2023). Auto-gpt for online decision making: Benchmarks and additional opinions. arXiv preprint arXiv:2306.02224.

Wang, J., & Duan, Z. (2024). IIntelligent Spark Agents: A Modular LangGraph Framework for Scalable, Visualized, and Enhanced Big Data Machine Learning Workflows. arXiv preprint arXiv:2412.01490.

Hong, S., Zheng, X., Chen, J., Cheng, Y., Wang, J., Zhang, C., ... & Wu, C. (2023). Metagpt: Meta programming for multi-Agent collaborative framework. arXiv preprint arXiv:2308.00352, 3(4), 6.

Xu, H., & Shatz, S. M. (2003). Adk: An Agent development kit based on a formal design model for multi-智能代理 systems. Automated Software Engineering, 10, 337-365.