

# 从单智能体到多智能体:大模型智能体支持下的激励型学习活动设计与实证研究\*

黄昌勤<sup>1,2,3</sup> 钟益华<sup>1</sup> 王希哲<sup>3</sup> 韩中美<sup>3</sup> 魏同权<sup>1</sup>

(1. 华东师范大学上海智能教育研究院, 上海 200062; 2. 浙江大学教育学院, 杭州 310058;  
3. 浙江全省智能教育技术与应用重点实验室, 金华 321004)

**摘要:**大模型与智能代理技术的不断进步,使得大模型智能体成为教育领域中实现教与学提质增效的重要新工具。基于大模型智能体的功能定位差异,单一智能体虽然已能够针对各类教与学任务提供诸如内容生成、智能反馈与评估等支持,但单智能体的交互特点、功能属性具有较高同质性,在促进深层次认知发展方面存在一定局限。相比之下,多智能体能够通过模拟多种教育主体角色,提升学习互动的多样性和深度,进而实现更为个性化和深度的学习体验。鉴于智能体在学习过程中的应用主要依靠学习者自发性,为了保障学习活动的有效开展,本研究基于 ARCS 动机模型分别设计基于单智能体与多智能体的激励型学习活动方案,并面向英语阅读场景开展了准实验研究。实验结果发现:基于多智能体的激励型学习活动相较单智能体能够显著提升学生在推理、评价与应用方面的学习成绩,具有更强的学习动机,且有效促进了其深层次认知发展,尤其是抽象与概括能力。研究证明了多智能体在支持学生深度学习中的价值,为未来进一步探讨多智能体在教育中的应用提供了借鉴。

**关键词:**大语言模型;多智能体;ARCS 动机模型;激励型学习活动;深度学习

## 一、引言

随着以 ChatGPT 为代表的大语言模型(下文简称“大模型”)技术的不断发展,基于大模型形成的智能体这一高效、先进的技术手段,已成为支持教育增效与未来发展的关键驱动力(杨宗凯等, 2023)。一方面,大模型智能体本身支持文本、图片、音视频等多模态数据的实时交互,可为学习者提供便捷高效的学习反馈支持,并且利用大模型的生成能力也能够为学习者营造更为生动、真实的学习环境与流程体验(Zhao et al., 2024);另一方面,智能体基于海量数据形成的知识体系,可以通过模拟各教育主体角色(专家、助教、学伴、管理者等),为各类学习活动提供个性化指导(Kasneci et al., 2023),进而帮助教师快速调整教学策略或方法,推动学习者认知过程的自我建构(蒋里, 2023)。

由于智能体的实现需要一系列的框架条件与验证过程,且对话式大模型本身的“提问-反馈”模式自然具有专家型智能体的角色属性,使得当前研究中多采用学习者与单智能体对话以辅助问题解决的模式(邱燕楠等, 2023)。该模式下智能体在学习者学习过程中往往充当着指导者的角色,主要依靠智能体对知识讲授、练习生成、评价反馈等特定能力进行内容指导(李海峰等, 2024)。虽然已有部分研究通过加入学习者能力水平感知(Wang et al., 2024)、阶段式反馈框架(Chen et al., 2024)等策略设计,证

\* 基金项目:教育部哲学社会科学研究重大课题攻关项目“人工智能和教育深度融合研究”(24JZD011)。

实了在单智能体支持下也能显著提升学习者绩效。但单智能体的角色限制导致学习过程难以进行多方协作活动设计,且其在提供交互与引发反思过程中的缺失也限制了学习者高层次认知探究的能力,不利于引发学生的深度思考(Lee et al, 2024)。针对这一局限,基于大模型的多智能体为学习活动支持提供了更为有效的解决方案。多智能体是由多个自主智能体组成的系统,智能体间可以通过协作或竞争的方式来完成复杂任务(Janbi et al., 2023),即以智能体与智能体的内部交互,实现智能体能力的群智涌现与自我提升(吴永和等, 2024),进而增强多智能体的整体应用效果;此外,多智能体提供的不同角色划分也可以更聚焦学习者的特定需求,即相比单智能体的直接知识讲述,多智能体能够模拟真实课堂中的“师(智能体)-生(自然人)”与“生(智能体)-生(自然人)”互动,实施与真实课堂相似的学习策略(卢宇等, 2024)。

由于在与智能体的交互过程中需要学习者具备较高自发性,而学习动机的缺乏常常使学生困囿于枯燥的学习活动中,进而影响学习效果(Zhang et al., 2021)。如何为学习者提供一种有效的动机激励框架,以支持学习者投入于当前的学习活动是智能体在教育教学中应用的关键内容。鉴于大模型智能体对学习过程角色与活动设计的全面多元支持,可以通过灵活运用注意(Attention)、关联(Relevance)、信心(Confidence)和满意(Satisfaction)的ARCS动机模型(Keller, 1987),激发学生动机,并引导他们减少低层次认知加工的负担,投入更多时间和精力于高层次的认知活动(如探究、协作和评价等),从而推动深度学习的实现。

基于此,本研究基于大模型与ARCS动机模型框架,分别设计了基于单智能体与多智能体的激励型学习活动设计方案,旨在激发学生的学习动机并引导其认知加工向深层次发展。为验证单智能体与多智能体的学习活动实施有效性与效果差异,本研究在大学生英语阅读课程中开展了实证研究,通过比较单智能体与多智能体在阅读学习活动开展过程中对学习成绩、学习动机和认知层次的影响,以检验其对学习者深度学习的促进作用,为培养学生高阶认知能力与深层次学习能力提升提供参考。

## 二、文献综述

### (一) 单智能体与多智能体系统

智能体是一种能够感知周围环境并反作用于环境,以实现其自身目标的自适应系统(Franklin et al., 1996)。在人工智能领域,设计与实现智能体一直被作为核心目标之一。长期以来,智能体受其核心模型的灵活性和智能化水平不足等限制,导致其难以发挥预期作用。随着大模型的推出,智能体的构建得到了可行的技术支持。智能体在常规应用中以单智能体形式为主,即智能体作为独立智能代理(Agent)发挥与外部环境的交互能力,而大模型的接入使得智能体在各类任务中的适配提供了可能(卢宇等, 2024)。然而,研究者们发现,单智能体在复杂环境中的灵活性和动态适应性仍有限。例如,在教育领域中,当单智能体面对多学习任务、多角色交互的复杂情境时,其反馈内容较为单一,也难以通过协作机制以动态生成更为优质的解决方案(于济凡等, 2024)。为此,研究者尝试通过多智能体的群智涌现和互相迭代提升,克服单智能体在灵活性和适应性方面的不足,使其能够有效适应更为复杂、动态的任务环境(吴永和等, 2024)。

多智能体系统的思想基础可以追溯到分布式人工智能的发展(翟雪松等, 2024)。起初,相关研究主要集中于分布式问题本身,并试图通过设定规则的方式确保分布式系统行为的一致性和效率。然而,这种基于既定规则的设计在面对复杂、动态的环境时表现出灵活性和适应性不足(Moulin et al., 1996)。随着多智能体系统在解决复杂问题方面展现出较强的动态性和灵活性(Gronauer et al., 2022),研究者逐渐将关注点从传统分布式人工智能转向了多智能体系统。同时,计算能力的不断提升和强化学习方法的应用,使得多智能体系统在环境感知、自主学习以及策略优化等方面取得了显著进展(Xie et al., 2017)。然而,在教育领域,该阶段的多智能体系统仍存在局限性。由于缺乏实时反馈机制,其灵活性和动态适应能力仍受到限制,从而影响了其在实际教学中的广泛应用(Guo et al., 2024)。

近年来,基于大模型的多智能体系统逐渐成为研究的焦点。这类系统结合了自然语言处理技术与生成式人工智能的优势,使得智能体能够通过人类语言进行实时、高效的交互,并具备在多模态环境下动态感知、推理与行动的能力(Xi et al., 2025)。这种即时反馈和动态学习机制,使多智能体系统达到了前所未有的灵活性和智能化水平。在教育领域,基于大模型的多智能体系统展现出了巨大的应用潜力(Guo et al., 2024)。例如,多智能体系统能够整合和生成文本、图像、语音等多种模态内容,为学习者创造沉浸式的学习体验(卢宇等, 2024)。此外,这些系统能够模拟师生对话,并根据学生的学习情况实时调整教学策略,提供即时反馈和动态指导,从而帮助学生更高效地理解和掌握复杂的知识结构,提升个性化学习效果(Swan et al., 2023)。同时,学生还可以参与由多智能体扮演的虚拟学习团队中,在协作互动中培养解决问题的能力,并通过多角度的反馈深化对知识的理解。因此,将基于大模型的多智能体系统应用于教育场景,不仅能够为学生提供更加丰富和动态的学习体验,同时还能以更高效的交互与协作机制推动教学模式的创新。

### (二) 基于大模型的智能体对深度学习的支持

随着大模型的不断迭代,大模型智能体为促进学生深度学习创造了有利条件,主要体现在以下两方面。首先,深度学习的情境不局限于以文本为唯一的信息载体,它强调文本、图像、视音频等多模态资源支持,以为学习者提供生动、丰富的沉浸式学习体验。智能体凭借其生成和处理多模态数据的优势,能够为构建真实的学习情境提供强有力支持(Zhao et al., 2024)。其次,智能体通过自然语言交互识别学生提出的问题,并能像人类专家一样,实时生成个性化的反馈与指导(Corlatescu et al., 2024; Kasneci et al., 2023),帮助学生发现不足,进而进行反思与调整,逐步形成适合自身的学习策略,实现自我知识建构。

然而,大模型智能体在促进深度学习方面仍面临一些挑战。现有研究多是通过与单智能体的交互来解决学生的预设问题(邱燕楠等, 2023)。例如,Wang et al.(2024)基于大模型构建了面向英语阅读的智能体,能够自动生成与学生阅读能力相匹配的个性化问题和反馈。然而,该智能体仅限于问题生成与知识讲述等功能,无法支持学生与智能体之间开展有助于高层次认知发展的协作、探究等学习活动。尽管研究结果显示,这类智能体在短期内能够提升学生的学习成绩,但在支持学生学习能力的长期发展,尤其是深度学习能力方面存在局限性。而基于大模型的多智能体系统为解决上述问题提供了可行解决方案。该系统通过多个自主智能体之间的协作或竞争,能够完成更为复杂的任务(Janbi et al., 2023)。此外,多智能体可以通过角色扮演的方式模拟真实课堂中的师生互动,实施类似于真实课堂的教学策略(卢宇等, 2024)。这一特性使得基于多智能体的教学设计能够灵活运用多元化的教学策略(如协作策略),激发学生的学习动机,引导其聚焦更高层次的认知活动。通过这种方式,学生不仅仅是知识的被动接受者,还能主动参与到推理、批判性思考和协作问题解决等更复杂的认知活动中(翟雪松等, 2024)。这种以深层认知为导向的学习活动,使学生能够投入更多的时间和精力关注高阶认知需求,从而促进深度学习的实现。

### (三) ARCS 动机模型及其在教育中的影响

ARCS 动机模型是一种综合多种动机策略的教学模型(Keller, 1987),包括注意、关联、信心和满意四个阶段,旨在提升学习者的学习动机,激励学习活动的持续高效开展。第一阶段旨在吸引学生的注意力,激发学习兴趣(Keller, 2008)。第二阶段要求学习活动与学生的个人目标和需求相匹配,使学生感受到学习的意义。第三阶段侧重于建立学生的自信,使其在学习过程中感到有掌控力并对成功充满期待。最后一个阶段则着重于提升学生的满足感,从而维持持续学习的动力。近年来,ARCS 动机模型在教育领域受到了广泛关注,研究表明其能够激发学生的学习动机和提升学习成绩(Chang et al., 2024)。特别是,ARCS 动机模型适合作为一种教学框架,用于指导有技术支持的学习环境(Song et al., 2001)。因此,将多智能体与 ARCS 动机模型相结合,能够有效整合技术支持和教学设计的优势,进一步促进学生深度学习能力的发展。

综上所述,本研究旨在将多智能体与 ARCS 动机模型相结合,设计基于多智能体的激励型学习活动。为了检验其有效性,在大学生英语阅读课程中开展了一项实证研究,旨在回答以下研究问题:

- (1)相比于单智能体,基于多智能体的激励型学习活动是否能更有效地提升学生的学习成绩?
- (2)相比于单智能体,基于多智能体的激励型学习活动是否能更好地激发学生的学习动机?
- (3)相比于单智能体,基于多智能体的激励型学习活动对学生深度学习认知层次的影响有何不同?

### 三、从单智能体到多智能体的激励型学习活动设计

#### (一) 基于单智能体的激励型学习活动设计

参考基于 ARCS 动机模型的激励型学习活动流程(即注意→关联→信心→满意),设计基于单智能体的激励型学习活动(见图 1)。在注意阶段,学生选择感兴趣的学习主题,智能体根据学生兴趣与主题生成个性化学习材料并提出初步探究性问题,旨在引出话题并激发学生的学习兴趣。在关联阶段,智能体结合现实情境设计深入探究性问题,帮助学生将所学内容与现实情境相联系,以增强学习的现实意义。在信心阶段,智能体设计个性化学习任务,并在学生遇到困难时提供引导性提示,帮助他们克服问题,逐步增强信心。在满意阶段,智能体根据学生的学习过程生成个性化评价与激励,帮助学生进行自我反思,巩固所学内容,以激发持续学习的动力。

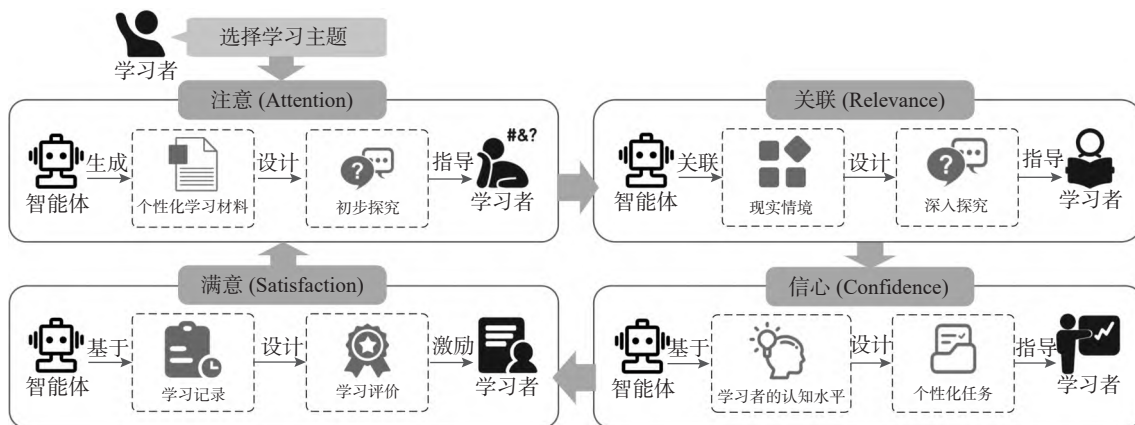


图 1 基于单智能体的激励型学习活动设计

然而,单智能体支持下的学习活动也存在一些不足:1)单智能体的互动性较为有限,学生通常只能与单一角色进行互动,缺乏与其他角色(如教师、同伴)的合作与讨论。这种局限使得学生难以获得来自不同角度的反馈和支持,思维和理解容易受到限制。2)单智能体的反馈通常是单向的,缺乏灵活性,尤其在学生遇到困难时,无法提供丰富的反馈形式或多元化的解决策略;3)单智能体的学习过程较为孤立,缺乏同伴互动和社交支持。这种缺乏社交互动的学习环境可能导致学生在面对挑战时缺少足够的动力和信心,从而影响学习积极性和持续性。4)单智能体的学习过程缺乏监管,难以实时监控和引导学生与智能体互动中可能出现的偏离学习主题等行为。

#### (二) 基于多智能体的激励型学习活动设计

为克服单智能体的不足,并为学生的学习过程提供更全面的支持,本研究开发了一种基于大模型的多智能体系统。利用现有多智能体框架可极大简化系统建构过程,从而实现快速开发。如表 1 所示,当前常用的基于大模型的多智能体框架包括 AutoGen、CAMEL、BabyAGI、MetaGPT、CrewAI 等。多智能体系统中人类参与是指人类用户是否可以以及如何参与智能体之间的交互过程。交互模式分为动态和静态两种:动态模式允许指定任意智能体作为下一位发言者,而静态模式则不允许指定,智能体间的交互关系固定不变。在这些框架中,只有 AutoGen 支持人类参与,允许自定义输入次数、条件,

并可以指定与用户交互的智能体角色。此外, AutoGen 同时支持动态和静态两种交互模式, 而其他框架仅支持静态模式。在其他功能上, 各框架均支持角色扮演和工具调用; 其差异在于: CAMEL 的启示式提示(inception prompting) 可让智能体自动循环最初的任务提示直到任务完成; BabyAGI 具备任务自驱动功能, 能够根据预定目标自动创建、组织和执行任务; MetaGPT 通过标准化操作程序(SOP) 确保智能体输出结构化内容; CrewAI 中的智能体则可以自主委派任务, 并通过相互询问提升问题解决效率。

表 1 基于大模型的多智能体框架对比

特点	AutoGen	CAMEL	BabyAGI	MetaGPT	CrewAI
人类参与	支持	不支持	不支持	不支持	不支持
交互模式	动态/静态	静态	静态	静态	静态
其他特征	角色扮演、工具调用、自定义条件	角色扮演、工具调用、启示式提示	角色扮演、工具调用、任务自驱动	角色扮演、工具调用、标准化操作程序	角色扮演、工具调用、自主任务委派

在实际的学习活动中, 学生与多智能体的交互主要体现在师生互动和协作探究过程。此外, 在学习活动的不同阶段, 多智能体与学习者之间的交互关系需要灵活调整。通过对比现有的多智能体框架, AutoGen 因其在人类参与和灵活交互模式上的优势, 故将其作为多智能体系统的开发框架。

在现有的多智能体框架支持下, 多智能体系统的开发成为可能。考虑到智能体在教学过程中的不同功能需求, 通过合理划分角色, 确保每个智能体在教学过程中各司其职, 从而实现协同增效。因此, 本研究设计了助教、同伴和专家三类角色的智能体, 分别为学生的学习过程提供全面支持。

基于多智能体的激励型学习活动设计见图 2。在注意阶段, 学生选择感兴趣的学习主题, 助教智能体根据学生兴趣与主题生成个性化学习材料, 并提出初步探究性问题; 同伴智能体通过模拟对话和角色扮演增强学生的沉浸感。此阶段旨在激发学生的学习兴趣并引出话题, 激活先验知识。在关联阶段, 助教智能体结合现实情境设计深入探究性问题, 学生在研读学习材料后围绕议题与同伴智能体展开讨论。此阶段旨在增强学习材料与现实生活的关联性, 帮助学生体会到学习的意义。在信心阶段, 助教智能体根据学习材料设计个性化学习任务, 并通过提供引导性提示帮助学生解决困难; 同伴智能体则在与学生的互动中分享任务完成经验。此阶段通过个性化的任务和同伴支持, 帮助学生增强信心并加深对学习材料的理解。在满意阶段, 学生通过与同伴智能体的互评进行自我反思, 助教智能体则根据学生的学习过程生成评价与激励, 进一步巩固学习成果并激发持续学习的动力。

### (三) 单智能体与多智能体的个性化任务生成工作流程

为了更加清晰地展示单智能体和多智能体在支持学生学习过程上的区别, 本研究以个性化任务生成为例, 详细描述其工作流程(见图 3)。在单智能体系统中, 任务生成由单个智能体独立完成。在学生提出具体的需求后, 智能体将根据学习材料和学生需求, 生成与此相关的任务。然而, 单智能体在该过程中缺乏动态调整难度和反馈机制的能力, 可能导致任务的难度过高或过低, 从而影响学生的学习体验。而在多智能体系统中, 任务生成则由多个智能体分工合作完成。具体而言, 首先由专家智能体制定任务的考核标准, 明确任务的预期目标和难度。其次, 助教智能体根据这些标准提供相应的学习资源和支持, 并生成具体的任务内容。该任务不仅覆盖学习目标, 还将根据学生的学习进度动态调整任务的难度和复杂度。此外, 由同伴智能体模拟学生的思维过程, 完成任务并根据要求修正答案。最后, 助教智能体则根据预设的标准评审同伴智能体的任务表现, 若任务符合标准, 则推送给学生; 否则, 继续优化任务设计直至符合标准后推送给学生。

## 四、实验设计

### (一) 研究对象

为了验证单智能体与多智能体在激励型学习活动中的有效性差异, 本研究选取中国东南部某大学

的48名教育技术专业的大二学生为研究对象,并将这些学生随机分为实验组和对照组。其中,实验组24人(8名男生,16名女生;平均年龄=19.69岁;SD=0.624),参与基于多智能体的激励型学习活动;对照组24人(10名男生,14名女生;平均年龄=19.31岁;SD=0.535),参与基于单智能体的激励型学习活动。所有学生均签署了知情同意书,具备丰富的在线学习经验,且此前已有一定的智能体使用经验,熟悉智能体的基本操作和交互形式。

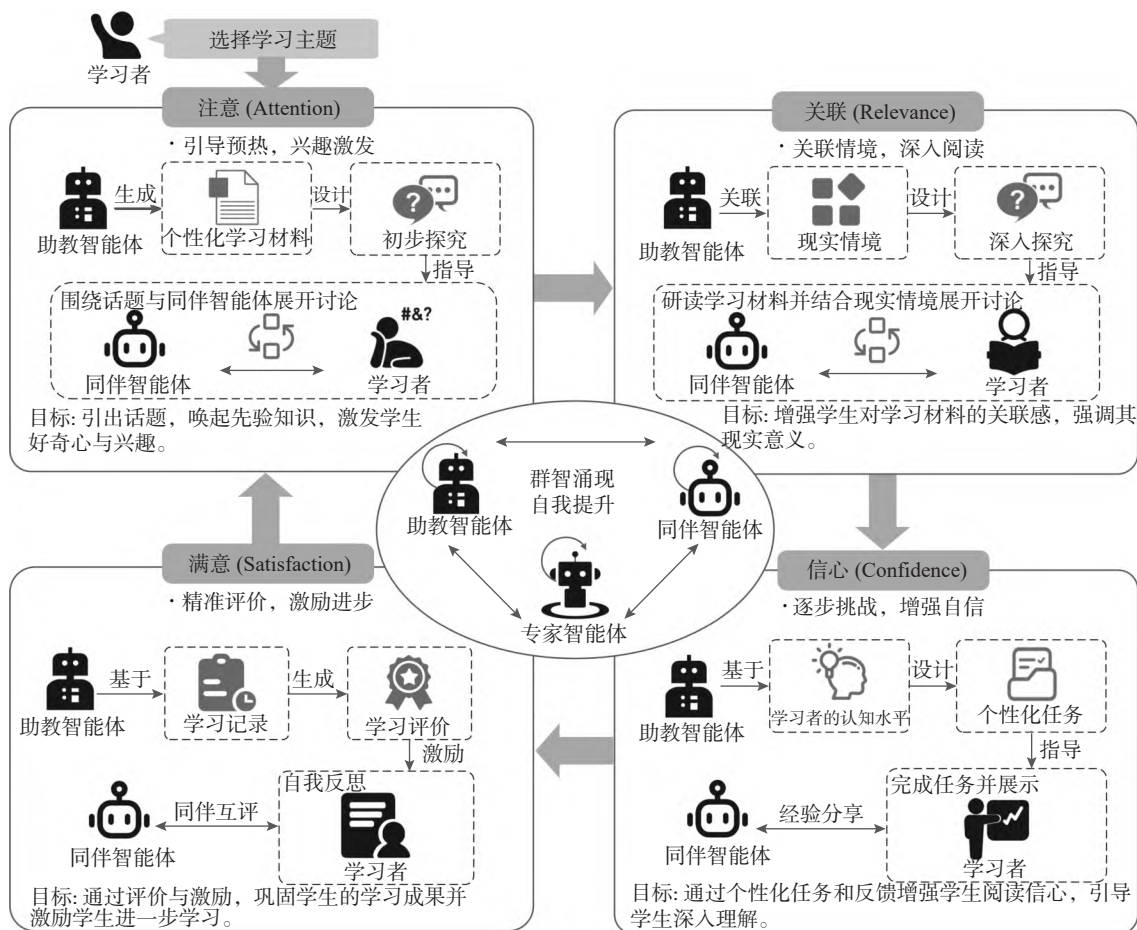


图2 基于多智能体的激励型学习活动设计

## (二) 实验设计与流程

本研究设计了一个准实验以检验单智能体与多智能体在激励型学习活动中对大学生深度学习的影响。在活动设计阶段,严格控制多智能体组与单智能体组在活动时长(每组活动时长均为45分钟)和复杂度上的一致性。两组活动在覆盖的知识点和难度设计上保持一致,确保活动目标具有可比性。根据实验活动中的时间记录,两组学生在完成活动时的平均用时无显著差异( $t=0.592, p>0.05$ ),表明活动安排未引入因时长不同而导致的混淆效应。

此外,为了确保两组学生在教师支持方面的一致性,我们对实验期间教师的介入进行了严格限制。教师的职责仅限于技术操作问题的解答,例如指导学生如何登录系统或解决设备问题,所有涉及学术内容的问题均被要求由学生独立解决。

该实验分为三个阶段:前测阶段、准实验阶段和后测阶段,整个实验共持续4周时间,其流程设计如图4所示。具体而言,在第一阶段,所有参与者完成学习成绩与学习动机问卷前测。前测结果显示,实验组和对照组在学习成绩上不存在显著差异( $t=0.943, p>0.05$ ),说明两组学生的学习水平在开展实

验之前大致相当。在第二阶段,实验组学生参与基于多智能体的激励型学习活动,而对照组学生参与基于单智能体的学习活动,本研究中该过程主要通过讯飞星火大模型的对话交互实现。在该阶段,系统会自动记录学生在学习活动中的问答数据,以便后续进行认知网络分析。在第三阶段,所有参与者完成学习成绩和学习动机问卷后测评。

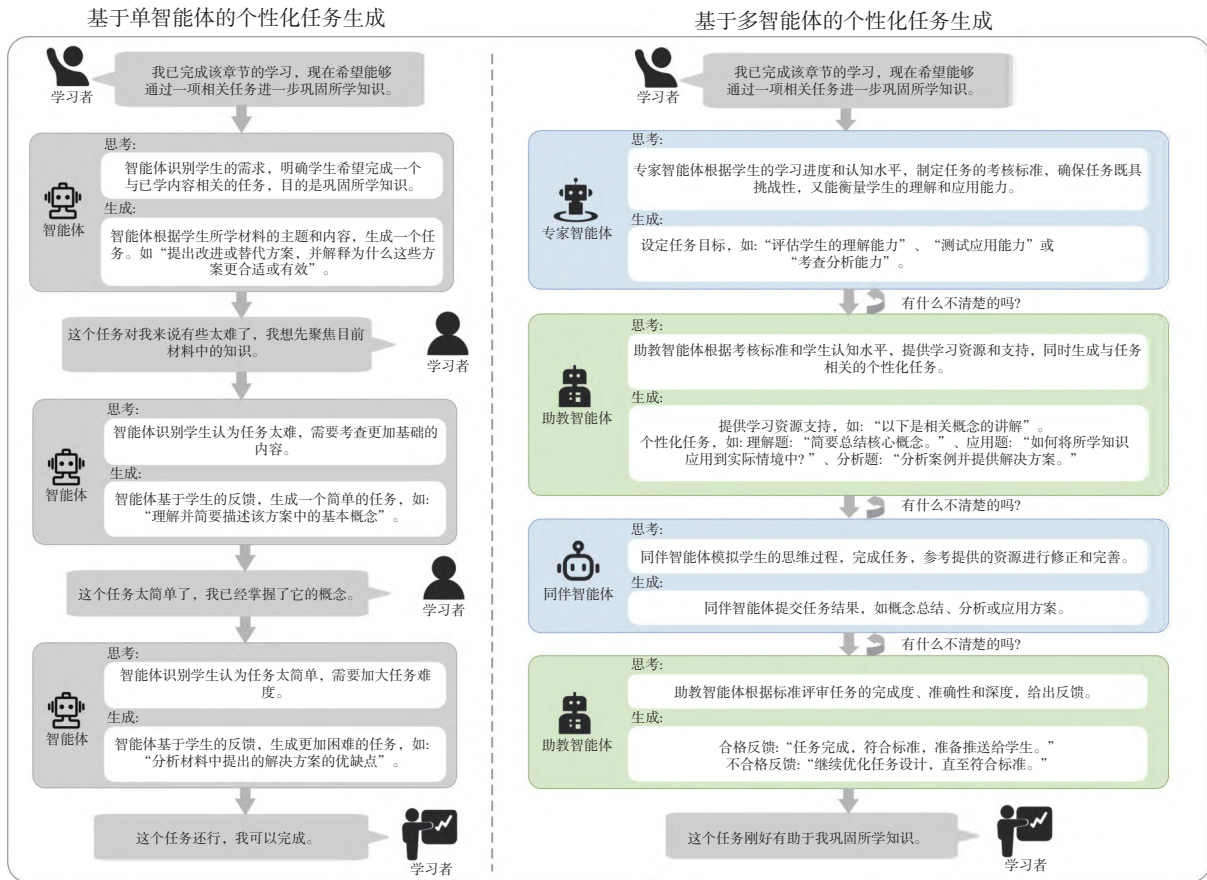


图3 单智能体与多智能体的工作流程对比

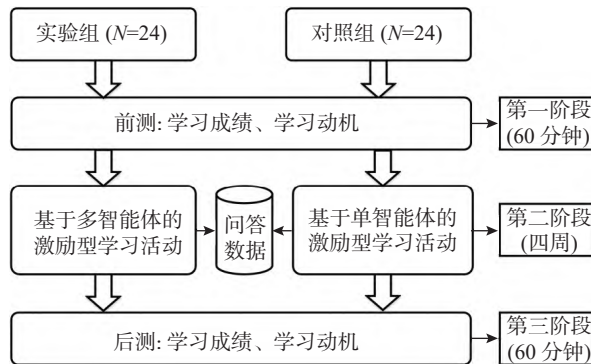


图4 实验流程

### (三) 实验测量工具

本研究采用量化研究和质性研究相结合的混合研究方法,以对比分析多智能体与单智能体在激励型学习活动中对学生学习成绩、学习动机和深度学习认知层次的影响。采用的测量工具主要包括:  
第一,学习成绩测试。该测试包括10个问题,分别从字面、推理、评价、应用四个维度对学生的学

习成绩进行评价。字面是指学生能够直接理解文本中的字面意思;推理是指学生能够推理和分析文本中的深层含义;评价是指学生能够对文本进行评价和判断;应用则是指学生能够将理解的知识运用到其他情境中。测试题目均为开放式问题,总共100分,其中字面35分,推理30分,评价25分,应用10分。本研究由3位具有丰富教学经验的教师进行双盲评价工作,评分成绩的组内相关系数值为0.86,说明3位教师的评分具有较高的一致性。

第二,学习动机问卷。该问卷包括6个问题,其设计参考了Lin et al.(2021)的研究,并根据本研究的具体情况,对句子进行了略微修改,例如,“在这门课上取得好成绩是我最满意的事情”。问卷采用李克特五点量表(最低分为1分,最高分为5分)。得分越高,学生的积极性越高。问卷的Cronbach's alpha值为0.85,具有较高的一致性。

第三,深度学习认知层次编码表。通过系统平台记录了学生在学习活动中的问答数据。在实验设计中,规定学生需基于智能体反馈独立完成作答。在数据分析阶段,对学生回答进行了筛查,并标注可能为直接复制智能体输出的回答,将此类数据从最终分析中剔除,以确保研究结果的真实性与可靠性。针对这些数据,本研究参考Biggs et al.(2014)提出的SOLO分类法,将学生的思维认知层次从简单到复杂依次划分为前结构(PS)、单点结构(US)、多点结构(MS)、关联结构(RS)和抽象拓展结构(AS)。该数据的编码工作由3名研究人员依据编码表进行人工编码,并进行一致性比较,评判信度Kappa值为0.89,说明3名研究人员的评判具有较高的一致性。

## 五、研究结果

### (一) 学习成绩

本研究将学生学习成绩前测作为协变量,后测作为因变量,采用单因素协方差分析来检验两组学生在学习成绩上的差异。分析结果如表2所示,两组学生在学习总成绩( $F=12.842, p<0.001, \eta^2=0.234$ )、推理( $F=13.938, p<0.001, \eta^2=0.249$ )、评价( $F=16.026, p<0.001, \eta^2=0.276$ )和应用( $F=7.412, p<0.001, \eta^2=0.150$ )维度上均呈现出显著差异,而在字面维度上不存在显著性差异。结果表明,基于多智能体的激励型学习活动可以显著提高学生的学习成绩。

表2 学习成绩的单因素协方差分析

变量	组别	N	Mean	SD	Adjusted mean	Std. error	F	$\eta^2$
总成绩	实验组	24	73.46	3.99	73.77	0.35	12.841***	0.234
	对照组	24	72.25	3.78	71.94	0.35		
字面	实验组	24	28.46	2.48	28.86	0.23	0.016	0.000
	对照组	24	29.21	1.74	28.81	0.23		
推理	实验组	24	21.67	1.01	21.59	0.11	13.938***	0.249
	对照组	24	20.92	1.10	20.98	0.11		
评价	实验组	24	15.63	1.21	15.67	0.12	16.026***	0.276
	对照组	24	15.00	1.25	14.96	0.12		
应用	实验组	24	7.71	0.46	7.65	0.11	7.412**	0.150
	对照组	24	7.13	0.61	7.18	0.11		

注: \*\* $p<0.01$ , \*\*\* $p<0.001$ 。

### (二) 学习动机

本研究将学生学习动机前测作为协变量,后测作为因变量,采用单因素协方差分析来检验两组学生在学习动机上的差异。分析结果如表3所示,两组学生的学习动机( $F=15.557, p<0.001, \eta^2=0.257$ )呈现出显著差异。结果表明,基于多智能体的激励型学习活动更有助于激发学生的学习动机。

表3 学习动机的单因素协方差分析

组别	N	Mean	SD	Adjusted mean	Std. error	F	$\eta^2$
实验组	24	3.72	0.23	3.69	0.05	15.557***	0.257
对照组	24	3.35	0.30	3.38	0.05		

注:\*\*\* $p < 0.001$ 。

### (三) 深度学习认知层次

本研究基于SOLO分类法的五个维度对两组学生的认知层次进行编码,结果如表4所示。研究人员共记录了146个编码,实验组(RS=45%; AS=14%)在深度学习层次上的频率高于对照组(RS=16%; AS=3%),而对照组(US=32%; MS=49%)在浅层学习层次的频率高于实验组(US=6%; MS=35%)。这些结果表明,基于多智能体的激励型学习活动更能促进学生的深度学习。

表4 两组学生各认知层次的频次

维度	编码	实验组		对照组	
		N	占比	N	占比
浅层学习	PS	0	0%	0	0%
	US	4	6%	24	32%
	MS	25	35%	37	49%
深度学习	RS	32	45%	12	16%
	AS	10	14%	2	3%

为进一步探究基于多智能体的激励型学习活动对学生深度学习认知层次的影响,本研究进行了认知网络分析,结果如图5所示。图5(a)显示了实验组和对照组学生的认知网络。其中,实验组学生的认知网络用红点表示,对照组学生的认知网络则用蓝点表示。另外,图5(b)中的红色方框代表实验组的认知网络质心,其分布于坐标轴右侧;右侧的大部分点与深度学习有关(即RS、AS),说明实验组学生多处于深度学习层次。蓝色方框代表对照组的认知网络质心,其分布于坐标轴左侧;左侧的大部分点与浅层学习有关(即PS、US、MS),说明对照组学生多处于浅层学习层次。

图5(c)和图5(d)分别是实验组和对照组的认知网络。可以发现,蓝线(对照组)主要在单点结构(US)、多点结构(MS)和关联结构(RS)维度上建立了较多的连接,而红线(实验组)则主要在多点结构(MS)、关联结构(RS)和抽象拓展结构(AS)维度上建立了较多的连接。此外,本研究对两组学生的认知网络差异进行了t检验,如表5所示。X轴的结果表明,两组学生认知网络的质心有显著性差异( $t = -6.43, p < 0.001, d = 1.86$ )。因此,通过分析上述结果可以发现,与对照组相比,实验组学生在学习中的认知层次更趋向于深度学习层次。

## 六、讨论

在学习成绩方面,实验组学生在整体、推理、评价和应用维度的得分均显著高于对照组。这些研究表明,基于多智能体的激励型学习活动更有助于促进学生对学习主题和材料的深入理解,提升学习成绩。这一发现呼应了翟雪松等(2024)的结论,表明通过多个智能体之间的交互和协作,能为学生的学习过程提供更加多样化的视角,促进议题的多元化,进而增加学习的复杂性和深度。此外,激励型学习活动中提供的反馈基于多智能体的群体智慧,在精准性和完整性方面得到显著提升,减少了因“幻觉”问题对学生产生的误导(Du et al., 2024)。因此,实验组学生更趋向于精准、深入的理解学习内容,提升了学习表现,尤其是在高阶能力方面(即推理、评价与应用)。然而,两组学生在字面理解成绩上未表现出显著差异,这可能与以下原因相关:首先,实验对象为具有较高阅读基础的大学生群体,他们在字面理解这一基础认知层次上已具备较强的能力,因此在该维度上未能观察到显著差异。其次,

字面理解任务属于较低认知层次的任务,单智能体的支持已能够有效满足学生完成任务的需求,而多智能体的优势则主要体现在更高阶的学习任务中,因此在此类基础任务中未能体现出其显著优势。

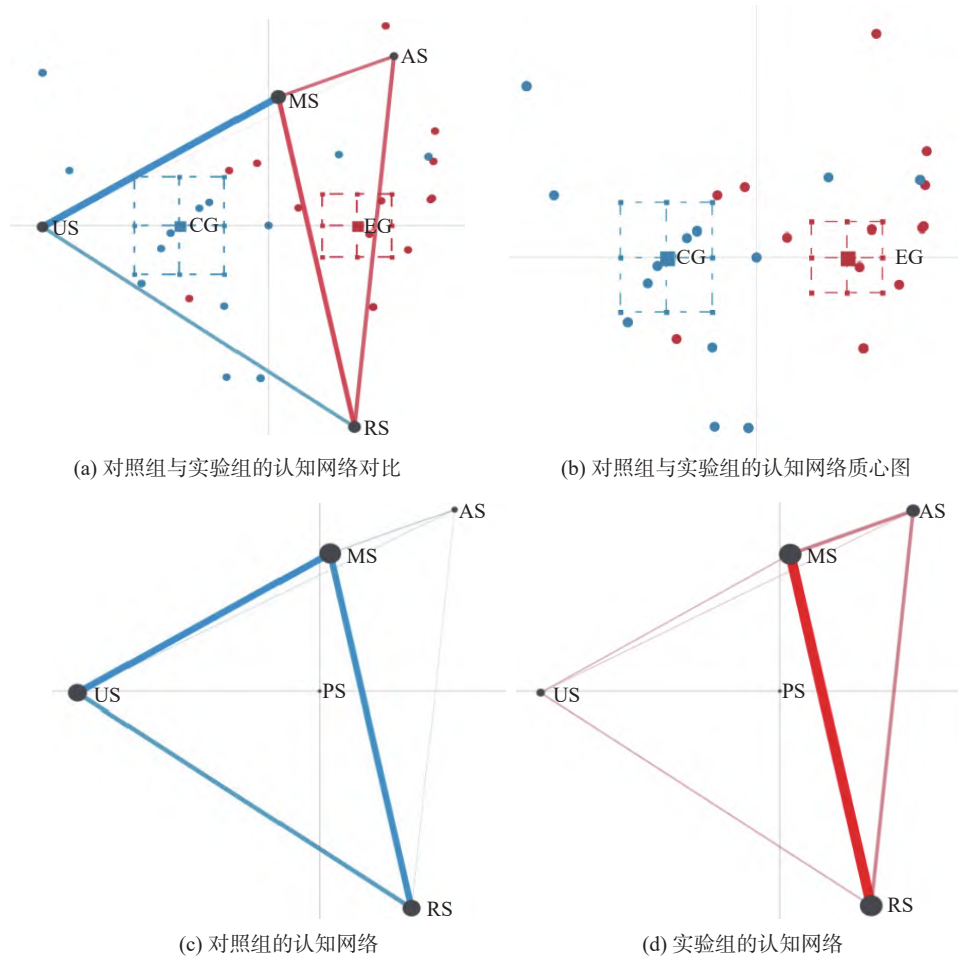


图 5 两组学生问答的认知网络分析结果

表 5 两组学生的认知网络差异 t 检验

组别	N	X轴				Y轴			
		Mean	SD	t	d	Mean	SD	t	d
实验组	24	-0.30	0.28	-6.43***	1.86	0.00	0.27	0.00	0.00
对照组	24	0.30	0.36			0.00	0.42		

注:\*\*\*p<0.001。

在学习动机方面,研究结果表明,基于多智能体的激励型学习活动可以显著增强学生的学习动机。这可能源于以下几个方面。首先,相比单智能体,多智能体具备更强的动态交互机制,能够构建更具活力的学习环境(翟雪松等, 2024)。其次,通过模拟真实课堂中的师生互动,使学生感受到是在与真实的人类对话,而非冰冷的机器,这有助于鼓励学生主动参与到多智能体的交流过程中进行思维碰撞。此外,本研究将 ARCS 动机模型作为一种教学框架,将动机策略融入教学的各个环节,这能发挥技术支持和教学设计两个方面的整合优势以激发学生的学习动机。大量研究也表明,将 ARCS 动机模型作为一种教学框架以指导技术支持的学习环境,有助于维持和激发学生持续学习的动力(Chang et al., 2024; Li et al., 2023)。然而,值得注意的是,多智能体组的学生可能因工具的新颖性而在短期内表现出

更高的学习动机,这种“新奇效应”在相关文献中已有报道(Deng et al., 2025)。虽然实验结果显示多智能体组学生的动机增强显著,但这一现象是否能够长期维持仍需进一步验证。

在深度学习方面,基于多智能体的激励型学习活动有助于提升学生的深度学习认知层次,具体表现为提高了学生对所学内容在现实情境中的关联和应用能力、增强了对学习内容的抽象概括能力。研究发现,实验组的认知网络(主要在多点结构、关联结构和抽象拓展结构之间)比对照组的认知网络(主要在单点结构、多点结构和关联结构之间)更趋近于深度学习。这一结论可引用梅培军(2022)提出的深度阅读学习实践原则来解释,即共生理论指导下的深度学习强调关联性、过程性、生成性和互成性四方面内容,而多智能体为以上原则在教学实践中的应用提供了有力支持。在关联性方面,多智能体通过模拟学生与教师、同伴乃至现实生活之间的互动,设计了丰富的学习活动,这不仅有助于学生理解文本,还能促使学生将所学知识应用到实际情境中,以深化对文本的理解。在过程性方面,基于多智能体的激励型学习活动注重反思性任务设计,通过提问、互动讨论和个性化反馈等方式,引导学生在学习过程中不断审视思维过程,引发更深入的思考和知识内化。此外,在生成性方面,多智能体能够构建更具激励性和开放性的学习环境,学生被鼓励与多智能体进行互动与协作,从多样化的视角解读文本内容,以促进他们产生对所学内容的创新性见解。最后,在互成性方面,基于多智能体的学习活动并非向学生提供直接的标准答案,而是在ARCS动机模型作为教学框架的指导下,激励和引导学生与多智能体进行开放性交流,在互动中进行观点碰撞,进而相互启发、激活思维,促进学生的知识建构和思维发展,实现深度学习。

此外,结合认知负荷理论,我们对多智能体功能在学习过程中的作用进行了更细化的分析。异质角色扮演功能通过提供多元视角和复杂情境,增加了任务的内在复杂性。这种内在负荷要求学生整合多样化的信息,从而推动其构建更全面和系统的认知结构。深层次提问功能通过设计更具挑战性的问题,激发学生在知识建构和迁移学习中的积极加工,这一过程可能增强学生的有效负荷,进一步强化其对学习内容的深度理解与应用能力。实时纠偏功能则通过精准动态反馈,帮助学生快速识别和调整认知偏误,同时减少外在负荷的干扰,使学生能够更加专注于核心任务并高效完成学习目标。多智能体的这些功能协同作用,通过适度提升内在负荷、增强有效负荷和减少外在负荷,为学生的学习过程提供了平衡的认知支持,进而有效促进了高阶认知能力的发展。

## 七、结论

本研究将大模型智能体与ARCS动机模型相结合,设计了基于单智能体与多智能体的激励型学习活动,以支持学生的深度学习。实验结果表明,相比基于单智能体的激励型学习活动,基于多智能体的激励型学习活动更能提高学生的学习成绩和动机。此外,认知网络分析结果进一步证实了其在促进学生深度学习上的有效性。尽管已有研究指出了多智能体在教育中的潜力,但还未尝试将多智能体与有效的教学策略(如ARCS动机模型)相结合,以促进学生的深度学习。本研究的结果展示了一种新的尝试,有助于启发研究人员和教师在未来采用多智能体进行相关的研究和教学设计。

然而,本研究也存在一些不足。首先,教学时间较短,因为深度学习能力的提升通常是一个长期的过程。因此短期实验可能无法全面反映多智能体教学的长期效果。其次,本研究的样本量较少,限制了结论的普适性,未来需要更大规模的样本量来验证研究结果。此外,尽管多智能体的活动设计可能因交互形式的丰富性而产生“新奇效应”,即学生对新颖工具的兴趣在短期内可能提升学习动机和成绩,但由于学生前期已有智能体使用经验,这种效应可能在短期内的影响并不显著。尤其是在一些任务场景中,多个智能体通过内部协作共同完成任务,最终由一个智能体给出反馈,这种方式与单智能体的交互方式相似,因此学生的体验可能并未显著不同于单智能体的使用感受。然而,需要注意的是,目前关于单智能体与多智能体新奇效应的长期研究仍然缺乏,因此未来研究应通过延长实验周期或重复实验,进一步排除新奇效应的影响,并探讨交互复杂性与学习效果之间的关系。此外,未来研究可采用

延时测验评估多智能体教学的长期学习效果,同时引入跨任务迁移测量和外部任务表现等验证方法,以检验多智能体教学在不同任务情境中的适用性及其在促进深度学习中的广泛应用价值。最后,建议未来研究对不同年级、不同学科、不同认知风格的学生进行对比分析,以更好的发挥多智能体这一工具在支持学生学习过程中的潜力。

(黄昌勤工作邮箱: cqhuang@zju.edu.cn; 王希哲为本文通信作者: xzwang@zjnu.edu.cn)

## 参考文献

- 蒋里. (2023). AI 驱动教育改革: ChatGPT/GPT 的影响及展望. *华东师范大学学报(教育科学版)*, (07), 143—150.
- 李海峰, 王炜, 李广鑫, 王媛. (2024). 智能助产教学法——以“智能苏格拉底会话机器人”教学实践为例. *开放教育研究*, (02), 89—99.
- 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.
- 梅培军. (2022). 深度阅读的实践原则与教学策略——共生理论的视角. *天津师范大学学报(基础教育版)*, (06), 71—75.
- 邱燕楠, 李政涛. (2023). 挑战·融合·变革: “ChatGPT 与未来教育”会议综述. *现代远程教育研究*, (03), 3—12+21.
- 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*, (05), 63—75.
- 杨宗凯, 王俊, 吴砥, 陈旭. (2023). ChatGPT/生成式人工智能对教育的影响探析及应对策略. *华东师范大学学报(教育科学版)*, (07), 26—35.
- 于济凡, 李睿森, 李曼丽, 刘惠琴. (2024). 多智能体协同交互的高临场感在线学习环境构建. *现代教育技术*, (12), 17—26.
- 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*, (03), 63—73.
- Biggs, J. B., & Collis, K. F. (2014). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
- Chang, C. C., & Hwang, G. J. (2024). Elevating EFL learners' professional English achievements and positive learning behaviours: A motivation model - based digital gaming approach. *Journal of Computer Assisted Learning*, 40(1), 176—191.
- Chen, Y., Zhang, X., & Hu, L. (2024). A progressive prompt-based image-generative AI approach to promoting students' achievement and perceptions in learning ancient Chinese poetry. *Educational Technology & Society*, 27(2), 284—305.
- Corlatescu, D. G., Watanabe, M., Ruseti, S., Dascalu, M., & McNamara, D. S. (2024). The automated model of comprehension version 4. *0-Validation studies and integration of ChatGPT. Computers in Human Behavior*, 154, 108154.
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224.
- Gronauer S, Diepold K. (2022). Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review*, 55(2), 895—943.
- Du, Y., Li, S., Torralba, A., Tenenbaum, J. B., & Mordatch, I. (2024). Improving Factuality and Reasoning in Language Models through Multi-agent Debate. In *Proceedings of the Forty-first International Conference on Machine Learning*, 11733-11763.
- Franklin, S., & Graesser, A. (1996). Is it an agent, or just a program? A taxonomy for autonomous agents. In *Proceedings of the International Workshop on Agent Theories, Architectures, and Languages* (pp. 21—35). Springer Berlin Heidelberg.
- 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. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 8048-8057.
- Janbi, N., Katib, I., & Mehmood, R. (2023). Distributed artificial intelligence: Taxonomy, review, framework, and reference architecture. *Intelligent Systems with Applications*, 18, 200231.
- Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274.
- Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance and instruction*, 26(8), 1—7.
- Keller, J. M. (2008). An integrative theory of motivation, volition, and performance. *Technology, Instruction, Cognition, and Learning*, 6(2), 79—104.
- Lee, H. Y., Chen, P. H., Wang, W. S., Huang, Y. M., & Wu, T. T.. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21(1), 1—28.
- Li, M., Chen, Y. T., Huang, C. Q., Hwang, G. J., & Cukurova, M. (2023). From motivational experience to creative writing: A motivational AR-based learning approach to promoting Chinese writing performance and positive writing behaviours. *Computers & Education*, 202, 104844.
- Lin, Y. N., Hsia, L. H., & Hwang, G. J. (2021). Promoting pre-class guidance and in-class reflection: A SQIRC-based mobile flipped learning ap-

- proach to promoting students' billiards skills, strategies, motivation and self-efficacy. *Computers & Education*, 160, 104035.
- Moulin, B., & Chaib-Draa, B. (1996). An overview of distributed artificial intelligence. *Foundations of distributed artificial intelligence* (pp.3—55). John Wiley & Sons.
- Song, S. H., & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational technology research and development*, 49(2), 5—22.
- Swan, M., Kido, T., Roland, E., & Santos, R. P. D. (2023). Math agents: Computational infrastructure, mathematical embedding, and genomics. Retrieved from <https://arxiv.org/abs/2307.02502>.
- Wang, X., Zhong, Y., Huang, C., & Huang, X. (2024). ChatPRCS: A Personalized Support System for English Reading Comprehension based on ChatGPT. *IEEE Transactions on Learning Technologies*, 17, 1762—1776.
- 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.
- Xie, J., & Liu, C. C. (2017). Multi-agent systems and their applications. *Journal of International Council on Electrical Engineering*, 7(1), 188—197.
- Zhang, D., & Perez-Paredes, P. (2021). Chinese postgraduate EFL learners' self-directed use of mobile English learning resources. *Computer Assisted Language Learning*, 34(8), 1128—1153.
- Zhao, R., Zhuang, Y., Xie, Z., & Philip, L. H. (2024). Facilitating self-directed language learning in real-life scene description tasks with automated evaluation. *Computers & Education*, 219, 105106.

(责任编辑 王 森)

## From Single-agent to Multi-agent: Motivational Learning Activities Design and Empirical Study Supported by LLM-based Agents

Huang Changqin<sup>1,2,3</sup> Zhong Yihua<sup>1</sup> Wang Xizhe<sup>3</sup> Han Zhongmei<sup>3</sup> Wei Tongquan<sup>1</sup>

(1. Shanghai Institute of Artificial Intelligence for Education, East China Normal University, Shanghai 200062, China;

2. College of Education, Zhejiang University, Hangzhou 310058, China;

3. Zhejiang Key Laboratory of Intelligent Education Technology and Application, Zhejiang Normal University, Jinhua 321004, China)

**Abstract:** The continuous advancements in large language models (LLMs) and agent technologies have made LLM-based agents a promising new tool for enhancing the quality and efficiency of teaching and learning in the educational domain. Based on the functional positioning of LLM-based agents, single-agent systems are capable of supporting various teaching and learning tasks, such as content generation, intelligent feedback, and assessment. However, due to the homogeneous nature of their interaction characteristics and functional attributes, single-agent systems face limitations in promoting deeper cognitive development. In contrast, multi-agent systems can simulate various educational roles, enhancing the diversity and depth of learning interactions and thereby providing more personalized and profound learning experiences. Considering that the application of agents in the learning process primarily relies on learners' initiative, this study based on the ARCS model designed motivational learning activities for both single-agent and multi-agent systems to ensure the effective implementation of learning activities and conducted a quasi-experimental study in the context of English reading. The experimental results reveal that motivational learning activities supported by multi-agent systems significantly outperform single-agent systems in improving students' performance in reasoning, evaluation, and application tasks. Furthermore, multi-agent systems exhibit stronger motivational effects and effectively promote deeper cognitive development, particularly in abstraction and generalization skills. The study underscores the value of multi-agent systems in supporting deeper learning and provides insights for further exploration of their applications in education.

**Keywords:** large language model; multi-agent; ARCS model; motivational learning activities; deeper learning