

人—AI 具身分布认知系统（HA-EDCS） 模型、机制与教育启示

李昌奎
香港产教融合研究所

摘要 本文超越“工具论”，融合具身认知与分布式认知，提出将“人—AI—环境”视为整合认知单元的 HA-EDCS 模型。以“协同增效悖论”为问题导向，界定数字具身与认知人造物，强调交互界面在共享表征与认知流动中的枢纽作用。模型刻画认知卸载与耦合、具身交互与情境生成、表征转化与生成、系统演化与能力涌现等机制，并提出以过程性评估与互补性 AI 设计降低外在负荷、防止认知萎缩的策略。研究为智能学习环境的设计与教学改革提供可操作的理论框架与方法论依据。

关键词 具身认知；分布式认知；人-AI 协同；认知卸载；交互界面

DOI <https://doi.org/10.6914/aiescn.010201> **文章编号** 3079-7918.2025.0102.1-16

引言：在“协同增效悖论”中重思人工智能教育

人工智能（AI）技术，特别是近年来以生成式人工智能为代表的大语言模型，正以前所未有的深度和广度渗透教育领域，引发了全球学术界的高度关注与研究热潮。^① 纵观人工智能教育研究的发展，其焦点已从早期的理论探讨，逐步演进至智能化教学模式的构建，并最终聚焦于“人机协同”这一核心议题。^② 人机协同学习，通常被定义为利用智能技术辅助和优化教学过程，通过人与机器的协同工作，推动教学全过程的重塑与创新。^③ 这一概念的提出，标志着研究范式从关注技术本身转向关注人与技术的互动关系，具有重要意义。在国家战略层面，相关政策文件也明确提出要“把人工智能融入教育教学全要素、全过程，创新智能学伴、智能教师等人机协同教育教学新模式”，彰显了其在教育现代化进程中的核心地位。^④

© The Author(s) 2025. Article History Received: 14 July 2025; Accepted: 19 August 2025; Published: 30 September 2025. 人工智能教育研究 (Réngōng Zhìnéng Jiàoyù Yánjiū) Artificial Intelligence Education Studies, ISSN 3079-7918 (Print), ISSN 3079-7926 (Online). Hosted by Artificial Intelligence Education Studies Press (Certificate No.70125648-007), and published by Creative Publishing Co., Limited (CPCL®) in Hong Kong. Email: cpcl@cpcl.hk, wtocom@gmail.com, <https://aies.press>, <https://cpcl.hk>. Repository:<https://pdf.ssci.cc/index.html>. This is an open-access article distributed under the terms of the Creative Commons Attribution-NoDerivatives 4.0 International License, which permits use, sharing, and redistribution in any medium or format, provided the original author(s) and source are credited, but does not allow modifications or adaptations of the work.

①王雯、李永智：《国际生成式人工智能教育应用与省思》，《开放教育研究》，2024年第3期，第37-44页。

②祝智庭、戴岭、赵晓伟：《“近未来”人机协同教育发展新思路》，《开放教育研究》，2023年第5期，第4-13页。

③王一岩、朱陶、杨淑豪、郑永和：《人机协同教学：动因、本质与挑战》，《电化教育研究》，2024年第8期，第58-65页。

④参见《国务院关于深入实施“人工智能+”行动的意见》。相关论述亦可参见：《人工智能赋能人机协同教育教学新模式》，《重庆日报》，2025年9月22日，第7版。

然而，在当前多数研究和实践中，一种潜在的、根深蒂固的“工具论”视角仍然占据主导地位。在这种视角下，AI 被视为一种功能强大的外部工具，由作为认知主体的人类学习者所使用，其核心价值在于提升个体学习者的知识获取效率或技能掌握水平。^① 这种观点虽然直观，但随着人-AI 交互的日益深化，其理论局限性也愈发凸显。技术的飞速发展对既有的理论模型构成了巨大的挑战，迫使我们必须寻求一种更具解释力的理论框架。^②

这种理论建构的迫切性并不仅仅源于学术思辨，更来自于一个严峻的经验难题：人-AI 协同的实际效果远未达到预期，形成了一种“协同增效悖论”（Synergy Paradox）。一项由麻省理工学院（MIT）研究人员主导的、回顾了百余项实验的综合性分析得出了一个令人意外的结论：在平均水平上，人-AI 组合的表现并未超越表现最佳的单独人类或单独 AI 系统，甚至在统计上显著更差。^③ 尤其在决策制定类任务中，人-AI 协同的表现往往不如单独的 AI。与此同时，另一项发表于《Frontiers》的系统性综述指出，当前的人-AI 交互在很大程度上“还不够协同”，其交互模式往往被简化为基础的菜单选择或按钮点击，缺乏支持真正动态协作的功能。这些发现共同揭示了一个深刻的矛盾：尽管技术能力日益强大，但我们尚未找到有效实现”协同增效的稳定路径。这一悖论的存在，直接对旨在通过人机协同提升教育质量的政策目标构成了现实挑战。若缺乏坚实的理论指导，大规模推广的“AI+ 教育”实践很可能陷入低效甚至反效的困境，从而偏离教育改革的初衷。

本文认为，这种协同实践的困境根源于理论的贫乏。占主导地位的“工具论”将人与 AI 在概念上割裂开来，导致系统设计者倾向于关注 AI 本身的技术性能，而忽视了人与 AI 之间交互界面的关键作用，这种忽视在大量实证研究中清晰可见。这种理论视角直接导致了那些被观察到的、功能简陋的交互范式的产生，并最终造成了协同效能不彰的后果。因此，要破解协同困境，实现人-AI 协同的巨大潜力，就必须进行一场深刻的“认知转向”：超越“工具论”，迈向一种“整合系统论”的视角。

这一转向的核心在于分析单元的根本性变革：不再将学习者个体或 AI 技术作为独立的分析对象，而是将由人、AI 及交互环境共同构成的混合体视为一个不可分割的、统一的整合认知系统。^④ 在这个系统中，认知过程并非局限于人类大脑内部，而是分布于系统的各个组成部分及其复杂的交互之中。学习的本质，也从个体对知识的单向内化，转变为整个认知系统结构与能力的动态演化与涌现。^⑤ 为清晰界定这一理论转向，下表对两种视角的核心维度进行了对比。

表 1：“工具论”视角与“整合系统论”视角对比

维度	工具论视角	整合系统论视角
分析单元	学习者个体	人-AI-环境构成的认知系统
AI 的角色	外部辅助工具	认知系统的内生组成部分
认知的场所	学习者大脑内部	分布于人、AI 及交互介质之间
学习的本质	知识的获取与内化	认知系统能力的演化与涌现
核心研究问题	AI 如何提升个体学习效率？	人与 AI 如何构成新的智能，其运行机制为何？

基于此，本文将从认知科学的前沿理论——具身认知（Embodied Cognition）与分布式认知（Distributed Cognition）——汲取思想资源，尝试构建一个能够描述和解释“人-AI”整合认知系统的新模型。通过整合这两大理论，本文旨在提出“人-AI 具身分布认知系统”（Human-AI Embodied Distributed Cog-

^① 本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。
^② Kenneth Holstein and Vincent Alevan, “Designing for human AI complementarity in K-12 education,” *arXiv preprint arXiv:2104.01266*, 2021.
^③ 本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。
^④ Edwin Hutchins, *Cognition in the Wild*, Cambridge, MA: MIT Press, 1995.
^⑤ 祝智庭等：《“近未来”人机协同教育发展新思路》，第 4-13 页。

nitive System, HA-EDCS) 模型, 深入剖析其运行机制, 并探讨其对未来教育设计、教学实践和学术研究的深远启示, 以期破解当前人-AI 协同的“协同增效悖论”提供一个更具解释力的理论框架。

二、理论基石 (一): 数字具身与认知人造物

在构建人-AI 整合认知系统的理论大厦之前, 必须首先奠定其微观层面的基石。这一基石建立在对人-AI 交互的物质性与身体性的深刻理解之上, 它要求我们超越将交互视为纯粹信息交换的抽象观念, 转而关注其物理的、具身化的本质。这需要我们整合具身认知、认知人造物以及延展心智三大理论。

(一) 具身认知: 智能根植于身体与环境的交互

具身认知思潮的兴起, 标志着认知科学对传统笛卡尔“身心二元论”的深刻反思与颠覆。^① 它反对将心智视为一种脱离身体、对抽象符号进行计算的处理器, 主张认知、心智与智能从根本上源于、依赖于并受限于拥有特定身体的行动者在物理与社会环境中的感觉-运动交互。^② 认知是具身的 (embodied)、嵌入的 (embedded)、生成的 (enactive) 和延展的 (extended), 这意味着认知过程不仅发生在大脑中, 更延展到身体乃至外部环境中。^③ 我们的身体形态、感觉运动系统的能力, 深刻地塑造了我们的概念体系和推理方式。在教育领域, 具身认知理论强调学习者身体的参与和感官体验是高级认知发展的基石, 有效的学习需要在认知主体充分整合所处环境与机体内部生理资源的基础上发生。^④

然而, 将具身认知理论应用于人-AI 协同学习这一高度数字化的领域时, 一个显而易见的挑战浮出水面: 大多数人-AI 交互发生在虚拟的数字空间, 似乎与物理性的身体相去甚远。^⑤ 为了应对这一挑战, 需要将“具身”概念进行深化和拓展, 提出“数字具身” (Digital Embodiment) 的新视角。在人-AI 交互中, 学习者的身体并未缺席, 而是通过物理接口 (如键盘、鼠标、触摸屏) 和软件界面这一中介, 将其感觉-运动能力延展到了数字世界, 形成了一个功能性的“数字身体”。^⑥ 学习者在屏幕前的每一次点击、拖拽、打字、滑动, 都不是纯粹的抽象信息输入, 而是一种具身行为。这些行为的流畅度、节奏感、精确性, 都受到界面设计所提供的“示能性” (affordance) 的塑造。^⑦ 这一过程正是技术现象学所描述的人与技术之间的“具身关系”, 技术设备延伸和拓展了人与世界的关系, 成为感知世界的第二介质。^⑧

(二) 认知人造物: 让物品变得智能的界面

“数字具身”的发生场域——人机交互界面——并非一个被动的、透明的信息通道。从人机交互 (HCI) 的视角看, 它是一种主动的、功能性的“认知人造物” (cognitive artifacts)。唐纳德·诺曼 (Donald Norman) 将认知人造物定义为“那些为了服务于表征功能而维持、展示或操作信息, 并以此影响人类认知表现的人造设备”。^⑨ 其核心观点是, 人类智能的真正力量并非来自孤立的大脑, 而是源于我们创造并使用外部辅助工具来增强认知能力——简言之, “是物品让我们变得智能”。从这个角度看, 人-AI 交互中的界面, 包括屏幕、键盘、鼠标、触摸屏乃至更复杂的语音 (VUI) 和多模态系统, 都不仅仅是信息的被动通道, 而

①李其维: “‘认知革命’与‘第二代认知科学’刍议”, 《心理学报》, 2008 年第 12 期, 第 1306-1327 页。

②Andy Clark, *Being There: Putting Brain, Body, and World Together Again*, Cambridge, MA: MIT Press, 1997.

③本文作者前期研究, 参见“具身与分布: 作为整合认知系统的人-AI 协同学习模型研究 (修订稿)”, 未刊稿, 2025 年。

④王辞晓: “具身认知的理论落地: 技术支持下的情境交互”, 《电化教育研究》, 2018 年第 7 期, 第 25-32 页。; 余海 EW: “具身认知视域下的知识学习与教学”, 《杭州师范大学学报 (社会科学版)》, 2018 年第 2 期, 第 108-115 页。

⑤陈兴冶、李曼: “基于具身认知理论的信息技术学科教学模型研究”, 《现代远程教育研究》, 2020 年第 2 期, 第 111-119 页。

⑥Don Ihde, *Technology and the Lifeworld: From Garden to Earth*, Bloomington, IN: Indiana University Press, 1990.

⑦James J. Gibson, *The Ecological Approach to Visual Perception*, Boston: Houghton Mifflin, 1979.

⑧Don Ihde, *Postphenomenology and Technoscience: The Peking University Lectures*, Albany, NY: State University of New York Press, 2009.; 唐·伊德著, 韩连庆译: 《技术与生活世界: 从伊甸园到尘世》, 北京大学出版社, 2012 年。

⑨Donald A. Norman, *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*, Reading, MA: Addison-Wesley, 1993.

是一种主动的、功能性的认知人造物。^①

这种人造物并非总是良性的。一个至关重要的区分在于，认知人造物可以被分为两种类型：互补性人造物（complimentary artifacts）和竞争性人造物（competitive artifacts）。互补性人造物通过提供支架来辅助思考、发展技能，例如使用纸笔规划思路；而竞争性人造物则倾向于替代心智过程，可能导致相关技能的萎缩，例如完全依赖 GPS 而丧失方向感。AI 作为一种极其强大的认知人造物，其设计和使用方式将决定它在特定学习情境中扮演互补性角色还是竞争性角色。这一区分对于后续探讨认知卸载的教育意涵至关重要。

（三）延展心智：作为心智一部分的 AI

将“数字具身”与“认知人造物”理论结合，可以进一步与心智哲学中的“延展心智假说”（Extended Mind Thesis, EMT）建立联系。由安迪·克拉克（Andy Clark）和大卫·查尔默斯（David Chalmers）提出的 EMT 主张，认知过程可以真实地延展到物理世界，外部工具在满足特定条件时就成为心智系统的一部分。^② 这些条件通常被称为“信任与胶水”标准（trust and glue criteria），具体包括：恒定可靠，该外部资源必须是可靠的，并且通常是行动者持续使用的；易于获取，当需要时，存储于其中的信息必须易于获取；自动信任，从该资源中获取的信息，其可信度被自动接受，通常不会受到二次审视；先前认可，存储于该资源中的信息，必须在过去某个时间点被行动者有意识地认可过。^③

当学习者与 AI 通过界面进行紧密耦合的交互时，如果这种交互满足上述条件——例如，学生持续使用某个 AI 助手，能即时获取其生成的内容，并且高度信任其输出（这在实践中很常见），那么 AI 系统在功能上就扮演了其延展心智的角色，承担了部分记忆、计算或信息生成的功能。^④

因此，通过整合具身认知、认知人造物和延展心智理论，我们可以建立一个更坚实的理论基础。人-AI 协同学习不再是两个孤立“心智”的纯信息交换，而是学习者的“数字身体”与作为认知人造物的 AI 系统之间的一场动态耦合。这种耦合的质量，即 AI 最终成为互补性还是竞争性认知人造物，直接决定了这种“心智延展”是促进认知发展还是导致认知萎缩。界面的设计、信息的呈现方式、交互的反馈延迟等，都成为塑造认知过程的重要环境因素。这一视角为我们后续构建整合认知系统模型提供了至关重要的身体与物质维度。

二、理论基石（二）：作为分析单元的分布式认知系统

如果说具身交互与认知人造物理念为我们的模型提供了微观的、现象学的维度，那么分布式认知（Distributed Cognition, DCog）则直接提供了将人与技术整合为单一分析单元的宏观理论框架。由认知人类学家埃德温·哈钦斯（Edwin Hutchins）开创的分布式认知理论，是对传统认知科学将个体设定为认知唯一载体的根本性挑战。^⑤

（一）分布式认知理论的核心：分析单元的转移

哈钦斯通过对轮船导航等复杂协同工作的民族志研究发现，完成这些任务所需的认知过程，并非完整地存在于任何一个人的头脑中，而是分布在一个由多人、多种技术工具和环境结构组成的复杂系统之中。^⑥ 在其著作《荒野中的认知》（Cognition in the Wild）中，哈钦斯详细描述了海军舰船上的导航团队如何通过复杂的社会互动和对图表、方位仪等物质工具的协同操作，来共同完成确定船只位置这一高认知要

^① 本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。

^② Andy Clark and David Chalmers, “The Extended Mind,” *Analysis*, Vol. 58, No. 1 (1998), pp. 7-19.

^③ Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, Oxford: Oxford University Press, 2008.

^④ 本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。

^⑤ 周国梅、傅小兰：“分布式认知——一种新的认知观点”，《心理科学进展》，2002 年第 2 期，第 148-153 页。

^⑥ Edwin Hutchins, *Cognition in the Wild*, Cambridge, MA: MIT Press, 1995.

求的任务。没有任何一个船员拥有关于导航任务的全部知识,知识和认知过程本身就“存在”于整个团队及其物质环境中。^①

因此,分布式认知理论的革命性在于其分析单元的转移。它主张,研究认知时不应预设边界,而应将“一个由个体、人造物及其在特定工作实践中的相互关系构成的集合”作为基本的分析单元。在这个单元内部,认知被视为“表征状态在不同媒介间的传播与转化”过程。^②根据哈钦斯等学者的阐述,认知过程的分布主要体现在跨社会成员、内部与外部结构之间(即认知卸载)以及跨时间三个层面。^③这一理论框架将个体和环境视为一个统一的认知系统,为分析学习活动提供了全新的视角。^④

将这一强大的分析框架应用于人与生成式 AI 的协同学习场景,能够为我们提供一种前所未有的、机制性的解释。以学生与 AI 共同撰写学术段落为例,这个过程不再能被模糊地描述为“合作”,而是可以被精确地建模为一个表征状态传播的循环:学生的内部表征(想法)通过指令被外部化,AI 将其转化为新的外部表征(生成的文本),学生再将此信息内部化以更新想法,从而开启新的迭代。^⑤在这个循环中,认知清晰地分布在学生的大脑、屏幕上的文本以及 AI 的模型之中。AI 在此处扮演的角色,远非一个被动的“工具”,而是一个积极的、能够转化和传播表征的认知媒介。

(二) 定位 HA-EDCS 模型:与其他系统理论的比较

为了更清晰地定位本文提出的模型,有必要将其与认知科学及相关领域的其他系统性理论进行区分。下表系统地比较了 HA-EDCS 模型与其他三个关键理论框架:延展心智假说(EMT)、增强智能(Augmented Intelligence)和社会技术系统(Socio-Technical Systems, STS)理论。

表 2: HA-EDCS 模型与其他系统理论框架比较

维度	人-AI 具身分布认知系统 (HA-EDCS)	延展心智假说 (EMT)	增强智能 (Augmented Intelligence)	社会技术系统 (STS) 理论
分析单元	整合的人-AI-界面认知系统	延展的个体心智(心智+工具)	被增强的人类个体	组织性的工作系统
技术观	认知系统的内生组成部分	个体认知的物理延伸	提升人类能力的工具	与社会子系统相互依存的技术子系统
核心机制	表征通过具身行动的传播与转化	内部与外部资源的耦合	人类技能的增强与放大	社会与技术子系统的联合优化
主要焦点	学习过程的微观机制	认知的本体论边界	任务表现与效率	组织设计与工作满意度

如表所示,延展心智假说(EMT)主要关注个体心智如何延伸至外部物件,其分析单元仍以个体为中心。^⑥增强智能的概念通常也持一种以人为中心的观点,即 AI 作为一种工具来增强或放大人类既有的能力,但并未从根本上改变分析单元。^⑦社会技术系统(STS)理论虽然也强调系统观,但它更多地应用于组织设计层面,旨在实现社会与技术两大子系统的“联合优化”,其分析粒度较宏观,较少深入到认知

^①James Hollan, Edwin Hutchins, and David Kirsh, “Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research,” *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 2 (2000), pp. 174-196.

^②Edwin Hutchins, *Cognition in the Wild*, p. 354.

^③周国梅、傅小兰:“分布式认知——一种新的认知观点”,第 148-153 页。

^④王美、王佑镁:“分布式认知理论视域下幼儿深度学习的要素分析”,《教育科学》,2022 年第 2 期,第 28-34 页。

^⑤本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025 年。

^⑥Andy Clark and David Chalmers, “The Extended Mind,” pp. 7-19.

^⑦本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025 年。

过程的微观层面。^①

相比之下,分布式认知理论为 HA-EDCS 模型提供了独特的理论优势。它从一开始就将整个系统(包括人、AI、环境)作为认知发生的场所,其核心关注点是认知过程本身(即表征的传播与转化)如何在系统的各个组件之间分布。这种视角不仅在本体论上更为激进,也为分析人-AI 协同学习中复杂的知识共建过程提供了更精细的分析工具。这种分析方式从认知视角打开了主体在整个创新或学习过程中的实践方式,体现了认知分布的重要意义。^② 这为构建一个整合的、动态的人-AI 协同学习模型奠定了核心理论基础。

三、模型建构:人-AI 具身分布认知系统(HA-EDCS)

在前两节坚实的理论基石之上,本文正式提出一个旨在超越“工具论”的整合性理论模型——“人-AI 具身分布认知系统”(Human-AI Embossed Distributed Cognitive System, HA-EDCS)。该模型的核心论点是:在深度协同的学习场景中,人类学习者、人工智能代理以及它们赖以交互的数字-物理界面,共同构成了一个不可分割的、功能整合的认知系统。此系统内的认知过程是分布式的,而这种分布又是通过学习者在界面上的具身化行动得以实现和调节的。^③

HA-EDCS 模型是对具身认知与分布式认知两大理论的有机融合。分布式认知理论为模型提供了宏观的系统架构,确立了将“人-AI-环境”整体作为分析单元的合法性。^④ 然而,若没有具身认知的视角,这种分布很容易被理解作为一种抽象的、非物质性的信息流动。具身认知理论则为模型注入了微观的、现象学的维度,强调了认知分布必须通过具体的、物理性的交互行为来完成。学习者与 AI 之间的认知耦合,并非凭空发生,而是通过“数字具身”才得以建立。因此,分布是系统的本质属性,而具身则是分布得以实现的必要途径。^⑤

在结构上,HA-EDCS 模型可以被理解为一个包含三大核心要素的动态系统:认知主体(Human Agent),即人类学习者。学习者是系统的意图发起者、过程的编排者和最终意义的阐释者,拥有背景知识、情感、意图以及在 AI 辅助下也无法被替代的批判性思维和伦理判断力。^⑥ 认知代理(AI Agent),即人工智能系统。AI 是系统认知能力的放大器、重组器和生成器,拥有强大的计算、记忆、模式识别和信息生成能力。根据其设计和在任务中的作用,它可以扮演互补性或竞争性认知人造物的角色。^⑦ 交互界面(Interactional Interface),这是连接认知主体与认知代理的具身化场域,是“数字具身”发生的场所。作为一个关键的认知人造物,其设计(如信息的布局、反馈的模式)深刻地影响着系统内表征的传播效率和认知过程的形态。

这一模型的构建,与计算机支持的协同学习(CSCL)领域从“以学习者为中心的设计”转向“以活动为中心的设计”(Activity Centered Design)的趋势不谋而合。^⑧ 后者同样建立在分布式认知和活动理

^①Eric L. Trist, “The evolution of socio-technical systems: A conceptual framework and an action research program,” in A. H. Van de Ven & W. F. Joyce (Eds.), *Perspectives on organization design and behavior*, New York: Wiley, 1981, pp. 19-75.

^②徐雅、陈凡:“主体工程创新实践的分布式认知分析”,《东北大学学报(社会科学版)》,2015年第5期,第449-454页。

^③本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

^④James Hollan, Edwin Hutchins, and David Kirsh, “Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research,” pp. 174-196.

^⑤本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

^⑥祝智庭等:“‘近未来’人机协同教育发展新思路”,第4-13页。

^⑦本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

^⑧D. Paul Kirschner, Jan-Willem Strijbos, Karel Kreijns, and Pieter Jelle Beers, “Designing Electronic Collaborative Learning Environments,” *Educational Technology Research and Development*, Vol. 52, No. 3 (2004), pp. 47-66.

论之上,强调学习是一个由个体、任务、工具和社群构成的复杂活动系统。^① HA-EDCS 模型正是将这一思想应用于人-AI 协同这一新兴领域,将 AI 视为活动系统中一个独特的、具有能动性的认知代理。

(一) 模型的视觉化呈现

为了更直观地理解 HA-EDCS 模型的动态运行机制,图 1 提供了一个视觉化的概念框架。该图展示了模型的三个核心要素如何通过四个核心机制相互作用,形成一个持续演化的认知循环。

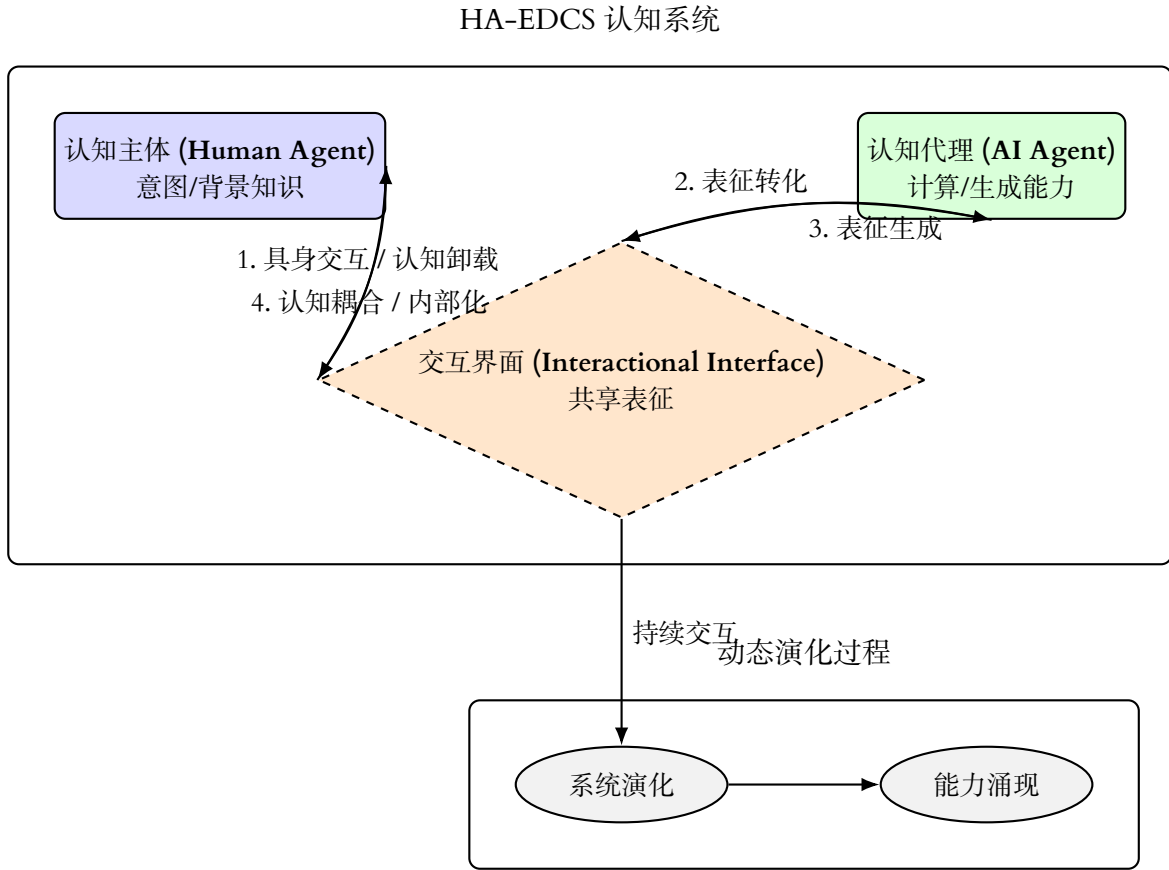


图 1: 人-AI 具身分布认知系统 (HA-EDCS) 模型示意图

如图 1 所示, HA-EDCS 模型是一个闭环系统。循环始于 认知主体 (Human Agent), 学习者将其内部的想法或意图, 通过在 交互界面 (Interactional Interface) 上的具身交互行为 (如打字、点击) 进行外部化, 这构成了 认知卸载的第一步。交互界面作为系统的核心枢纽, 承载着双方均可感知的 共享表征 (如文本指令)。认知代理 (AI Agent) 接收并转化这些表征, 生成新的信息, 并将其呈现在交互界面上。认知主体随后感知并内部化这些新的共享表征, 完成一次认知耦合, 并可能基于此发起新一轮的交互。这个循环并非静止的, 随着持续的交互, 整个系统会发生 演化, 并可能在特定条件下 涌现出超越任何单一组成部分能力之和的新能力。这个视觉模型清晰地揭示了交互界面在调节人与 AI 之间认知流动的中心地位, 也直观地展现了认知过程如何在系统的不同组件之间动态分布。

HA-EDCS 模型的提出, 其价值不仅在于对人-AI 协同学习现象的描述, 更在于其生成性——它能够催生一套全新的研究问题, 从而推动整个领域的范式转型。在“工具论”框架下, 研究者倾向于问: “使用这个 AI 工具对学生的学习成绩有何影响?” 这是一个典型的输入-输出式评估问题。而在 HA-EDCS 模

^①Y. Engeström, “Expansive learning at work: Toward an activity theoretical reconceptualization,” Journal of Education and Work, Vol. 14, No. 1 (2001), pp. 133-156.

型框架下,研究问题将转变为:①“在这个人-AI系统中,认知负荷是如何在人与AI之间动态分配的?这种分配是促进了技能发展还是导致了技能萎缩?”;“当前的交互界面设计,作为一种认知人造物,是支持还是阻碍了何种类型的共享表征的建立?”;“不同的具身交互模式(例如,通过打字对话与通过语音对话)如何改变系统整体的认知特性和学习体验?”;“系统是否以及在何种条件下涌现出了超越个体学习者和AI独立能力之和的新能力?这种涌现的机制是什么?”。

这些问题将研究的焦点从对“工具”效果的简单评判,转向对“系统”内部过程与机制的深入探究。这种研究视角的转变,将问题从单纯的教育学范畴,扩展到了一个需要系统设计思维来解决的领域。它不再仅仅关注学习者的最终产出,而是关注整个认知系统的功能效率、内部协调以及动态演化。这不仅为理论研究开辟了新的方向,也为设计更有效、更符合认知规律的智能学习环境提供了根本性的指导。

四、HA-EDCS模型的运行机制:协同、表征、生成与涌现

为了使“人-AI具身分布认知系统”(HA-EDCS)模型更具操作性和解释力,本节将深入剖析其内部的四个相互关联的核心运行机制。这些机制共同描绘了一幅动态的图景,揭示了人与AI如何通过具身化的交互,实现认知过程的分布、耦合与共同演化,并为理解“协同增效悖论”的成因提供了机制性的解释。

(一) 认知卸载与耦合:效率与风险的双刃剑

认知卸载(Cognitive Offloading)是分布式认知系统最基本的运行机制之一。它指的是认知主体将部分认知任务转移到外部环境或人造物上,从而释放有限的内部认知资源(如工作记忆和注意力),以专注于更高阶的认知活动。②在HA-EDCS中,AI作为一个极其强大的认知人造物,为学习者提供了前所未有的认知卸载可能性。③例如,在撰写一篇复杂的学术论文时,学生可以将信息检索、参考文献管理、甚至初步内容生成等任务卸载给AI,自己则能集中精力于论点构建、逻辑推理和创造性表达等更高阶的认知活动。

与认知卸载相伴相生的是认知耦合。当认知任务被卸载给AI后,人与AI的认知过程便不再是独立的,而是变得紧密联结、相互依赖,形成一个快速迭代的闭环。这种高度的相互依赖性,使得人与AI形成了一个功能上不可分割的认知耦合体。

然而,认知卸载是一把双刃剑。虽然它能提升效率,但也带来了严峻的教育风险,即“认知萎缩”(cognitive atrophy)——由于长期不使用某些心智过程而导致其能力衰退的“不用则废”现象。多项研究指出,过度依赖AI与批判性思维能力的下降存在负相关关系,尤其是在年轻群体中。一项MIT的研究通过脑电图(EEG)监测发现,使用ChatGPT撰写文章的学生,其大脑活跃度显著低于仅靠自己或使用谷歌搜索的学生,这反映了深度记忆和创造性构思过程可能被绕过。教育家霍华德·加德纳(Howard Gardner)甚至警告说,技术的发展可能使“心智的大多数认知方面”对人类来说成为“可选项”。

教育者面临的核心挑战在于,如何引导学生进行有益的认知卸载,同时避免有害的认知替代。关键在于区分AI是作为促进技能发展的“脚手架”(scaffold),还是导致技能萎缩的“拐杖”(crutch)。这意味着教学设计必须引导学生将认知卸载用于辅助而非替代核心认知技能,确保学习者始终处在“认知拓展”(cognitively expand)而非纯粹“认知卸载”(cognitively offload)的状态。

(二) 共享表征与交互基面:主体间性的构建

任何形式的协同认知,都必须建立在某种共享的理解之上。在HA-EDCS中,人机交互界面扮演了“交互基面”(Interactional Ground)的关键角色,它承载着双方都能感知和操作的共享表征(Shared

①本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI协同学习模型研究(修订稿)”,未刊稿,2025年。

②David Kirsh, “A Few Thoughts on Cognitive Overload,” *Intellectica*, Vol. 30, No. 1 (2000), pp. 19-51.

③本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI协同学习模型研究(修订稿)”,未刊稿,2025年。

Representations), 如对话历史、共同编辑的文档、可视化图表等。这个交互基面是一个外部化的、共同的认知空间, 为双方提供了一个共同的参照点, 从而能够建立和协商共同的理解, 即达成某种程度的“主体间性”(intersubjectivity)。

这一概念与计算机支持的协同学习(CSCL)领域中关于建立“共同基础”(common ground)或“共享理解”的理论紧密相关。^①在CSCL环境中, 共享的数字工作空间(shared workspace)是促进知识共建的核心技术支持。^②在人-AI协同中, 交互基面同样扮演着这一角色。例如, 一项针对博士生使用AI辅助写作的实证研究发现, 学生与AI在一个共享的文本界面上进行反复的互动, 通过不断的提问、生成、评估和修改, 共同塑造最终的文本。^③这个过程正是通过操作共享表征来动态协商意义的体现。对话的轮流交换结构(如“提问-回答”这样的邻近对)是构建共享表征、不断校准彼此“意图”的核心过程。^④因此, 交互基面的设计质量, 直接影响到共享表征建立的效率和深度。

(三) 具身交互与协同模式的生成

人与AI的交互并非抽象的信息交换, 而是通过学习者的“数字具身”行为完成的。^⑤HA-EDCS模型的第三个机制强调, 这些具身交互行为本身就是构成学习活动的一部分, 并且能够动态地生成和改变学习的情境。学习者与AI交互的节奏、强度、模式等, 都蕴含着丰富的认知与情感信息。^⑥

更重要的是, 我们可以从实证研究中识别出具体的、可观察的交互模式, 这些模式与学习效果直接相关。前述针对博士生写作的研究就识别出了不同的协同策略。研究发现, 那些与AI进行迭代式、高度互动过程的学生, 其写作任务的表现通常更优。这种模式体现为学生不仅仅是提出一个宽泛的指令然后接受结果, 而是通过一系列精确、连续的追问和修改指令来引导和塑造AI的输出, 与AI形成一种紧密的“对话伙伴”关系。相反, 那些采用简单“一次性”指令或低互动策略的学生, 其表现则相对较差。

这表明, 具身交互行为的差异(是选择快速、连续地提问, 还是进行缓慢、深思熟虑的指令修改)不仅是实现目标的不同路径, 它们本身就构成了不同的协同模式, 并生成了不同的学习情境。一个“迭代式”的交互模式生成了一个深度参与、共同探索的“协同建构”情境; 而一个“一次性”的交互模式则生成了一个被动接受、浅层利用的“工具使用”情境。因此, 具身交互不仅是实现认知卸载和建立共享表征的手段, 它本身就是一种意义生成的过程, 是学习情境的动态来源。

(四) 系统演化与能力涌现: 协同增效的条件

HA-EDCS模型是一个动态的学习系统, 其最令人激动的特征在于其演化与涌现的潜力。随着人与AI交互的持续, 系统会发生双向的适应与学习。一方面, 学习者会变得越来越擅长与AI协作, 发展出所谓的“协同AI素养”(Collaborative AI Literacy)。另一方面, 先进的AI系统也能根据学习者的行为数据进行模型微调, 从而更好地适应其个人风格和需求。^⑦

这种双向适应的最终结果, 是整个“人-AI”系统能力的演化, 并可能发展出超越任何单一组成部分能力的涌现能力(Emergent Capabilities)。^⑧然而, 正如引言中所述, 这种“协同增效”并非必然发生。研究表明, 协同增效的涌现是有条件的。一个关键的条件是任务类型。前述MIT的综合性研究发现, 人-AI

^①Susan R. Fussell, Robert E. Kraut, and Jane Siegel, “Coordination of communication: Effects of shared visual context on collaborative work,” in Proceedings of the 2000 ACM conference on Computer supported cooperative work, 2000, pp. 21-30.

^②Gerry Stahl, “Building collaborative knowing: Elements of a social theory of CSCL,” in Computer support for collaborative learning: Foundations for a CSCL community, Mahwah, NJ: Lawrence Erlbaum Associates, 2002, pp. 53-60.

^③本文作者前期研究, 参见“具身与分布: 作为整合认知系统的人-AI协同学习模型研究(修订稿)”, 未刊稿, 2025年。

^④Emanuel A. Schegloff, Sequence Organization in Interaction: A Primer in Conversation Analysis, Vol. 1, Cambridge: Cambridge University Press, 2007.

^⑤王辞晓: “具身认知的理论落地: 技术支持下的情境交互”, 第25-32页。

^⑥本文作者前期研究, 参见“具身与分布: 作为整合认知系统的人-AI协同学习模型研究(修订稿)”, 未刊稿, 2025年。

^⑦王喜、王崇阳: “人机协同中的交互式学习”, 《中国计算机学会通讯》, 2021年第8期, 第26-31页。

^⑧赫尔曼·哈肯著, 凌复华译: 《协同学导论》, 上海译文出版社, 1984年。

协同在内容创作类任务中表现出显著的正向协同效应，而在决策制定类任务中则往往表现出负向效应。^①这种差异的存在，恰恰印证了 HA-EDCS 模型的解释力。在决策任务中，通常存在一个由数据驱动的最优解，人类的认知偏见或对 AI 的不当信任（或不信任）会干扰系统达到这个最优解，从而导致负协同。而在开放的创作任务中，人类的审美、情境理解和创造性意图能与 AI 的生成能力形成真正的互补，从而产生正协同。这为我们预测和设计能够产生涌现能力的学习任务提供了重要的实证依据。例如，一个艺术家与 AI 绘画工具共同创作出一种全新的艺术风格，这正是在一个开放的、生成性的创作任务中实现的协同涌现。

这四个机制之间存在着紧密的因果联系。学习者采取的具身交互模式（例如，迭代式对话）决定了他们能否与 AI 建立高质量的共享表征；这种深度参与的建构过程使得认知卸载更倾向于成为一种有益的认知拓展，而非有害的认知替代；而正是这种特定类型的交互、表征和卸载的组合，在合适的任务类型（如内容创作）下，最终创造了能力涌现的条件。这一过程模型为解释人-AI 协同效能的差异，即“协同增效悖论”，提供了机制性的说明。

五、交互界面：作为认知生态系统的核心

在 HA-EDCS 模型中，交互界面并非一个中立的通道，而是连接认知主体与认知代理、调节二者认知流动的核心枢纽。它是“数字具身”发生的物理场域，也是承载共享表征的认知空间。可以说，交互界面的设计质量直接决定了人-AI 认知系统的整体效能。本节将深入论述交互界面作为认知生态系统核心的关键作用，并从认知负荷、界面模态和设计原则三个方面，探讨其如何成为破解“协同增效悖论”的关键。

（一）界面作为认知负荷的调节器

根据人机交互（HCI）领域的认知负荷理论（Cognitive Load Theory），个体在处理信息时所消耗的认知资源是有限的。认知负荷主要分为两类：内在认知负荷（Intrinsic Cognitive Load），即任务本身固有的难度；以及外在认知负荷（Extraneous Cognitive Load），即由信息呈现方式或界面设计不当所造成的额外认知负担。^②学习效果的优化，在很大程度上依赖于将外在认知负荷降至最低，从而使学习者能将宝贵的认知资源集中用于处理内在认知负荷。^③

从这个角度看，“协同增效悖论”的产生，很大程度上可以归因于当前多数 AI 交互界面设计所引发的高昂外在认知负荷。以典型的生成式 AI 对话框为例，尽管其看似简洁，但在处理复杂学习任务时，它会给用户带来多种外在负荷：指令构建负荷，用户需要将复杂的意图转化为 AI 能够理解的、精确的文本指令（即“提示工程”），这本身就是一项高要求的认知任务；记忆负荷，用户需要记住之前的对话脉络，以便进行连贯的追问和迭代，因为线性展开的对话历史并不利于快速回顾和整合；整合负荷，当 AI 生成多个分散的回答时，用户需要在大脑中将这些信息碎片整合成一个连贯的知识体系。

这些由界面设计不当所施加的外在认知负荷，很可能抵消了甚至超过了通过认知卸载所节省的认知资源，最终导致人-AI 系统的整体表现下降。因此，一个设计优良的交互界面，其首要目标就是通过优化信息结构、提供视觉化支架、简化操作流程等方式，最大限度地降低外在认知负荷，从而为实现协同增效创造必要的认知条件。

（二）界面模态与认知系统的动态

^①本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。

^②Ali Darejeh, Nadine Marcusa, and Gelareh Mohammadi, “A critical analysis of cognitive load measurement methods for evaluating the usability of different types of interfaces: guidelines and framework for Human-Computer Interaction,” arXiv preprint arXiv:2402.11820, 2024.

^③John Sweller, Jeroen J. G. van Merriënboer, and Fred G. W. C. Paas, “Cognitive Architecture and Instructional Design,” Educational Psychology Review, Vol. 10, No. 3 (1998), pp. 251-296.

交互界面的模态 (modality), 即用户与系统进行信息交换的通道 (如文本、图形、语音、手势等), 深刻地影响着 HA-EDCS 的内部动态。不同的模态对应着不同的认知加工方式, 从而产生截然不同的认知负荷和交互体验。

命令行界面 (CLI) 与图形用户界面 (GUI): CLI 要求用户精确回忆 (recall) 并输入指令, 认知负荷较高, 但对于专家用户而言效率极高且控制精准。^① GUI 则通过图标和菜单等视觉元素, 将任务转化为再认 (recognition), 极大地降低了认知门槛和负荷, 但可能牺牲部分效率和灵活性。^②当前主流的生成式 AI 聊天界面, 在认知上形成了一种矛盾: 它拥有 GUI 的简易外观, 却要求用户具备 CLI 式的精确指令构建能力, 这种认知上的不匹配是造成交互摩擦和高外在负荷的重要原因。

多模态界面 (Multimodal Interfaces) 的潜力: 多模态界面整合了语音、手势、眼动、触觉等多种输入输出通道, 允许用户以更自然、更符合人类本能的方式进行交互。^③研究表明, 多模态学习环境能够通过整合多种感官通道, 提供更丰富的信息表征, 从而增强理解和记忆。^④在 HA-EDCS 的框架下, 多模态界面能够创造一种更深度的“数字具身”体验, 降低认知加工的转换成本, 使认知在人与 AI 之间的分布更加无缝和高效。^⑤例如, 学生可以通过语音提出一个模糊的想法, 同时用手势在屏幕上圈出需要修改的段落, 这种多模态的协同指令远比纯文本输入更高效、更直观。这为设计能够真正实现协同增效的下一代学习系统指明了方向。

下表对不同界面模态对 HA-EDCS 的影响进行了比较。

表 4: 不同界面模态对 HA-EDCS 的影响比较

界面模态	核心认知动作	外在认知负荷	对共享表征的影响	对协同增效悖论的可能影响
类 CLI 的文本框	回忆与精确构建	高 (需记忆指令结构)	表征单一 (纯文本), 线性且碎片化	易导致负协同 (高交互成本抵消 AI 优势)
传统 GUI	再认与选择	低 (直观、易上手)	表征丰富 (图文并茂), 但交互受限于预设选项	可能实现弱正协同, 但难以支持开放式、生成性任务
多模态界面	直观行动与表达	最低 (符合自然交互)	表征最丰富、最动态, 支持同步、多通道输入	最具潜力实现强正协同 (极大降低交互成本, 促进认知流畅)

(三) 设计“互补性”而非“竞争性”的认知人造物

交互界面的最终目标, 是使 AI 成为一个“互补性”而非“竞争性”的认知人造物。这意味着界面设

^①Shardeum Content Team, “GUI vs. CLI: What Are the Differences?,” Shardeum Blog, 2023, <https://shardeum.org/blog/gui-vs-cli/>.

^②“Difference between CLI and GUI,” GeeksforGeeks, 2023, <https://www.geeksforgeeks.org/operating-systems/difference-between-cli-and-gui/>.

^③Nicu Sebe, “Multimodal interfaces: Challenges and perspectives,” Journal of Ambient Intelligence and Smart Environments, Vol. 1, No. 1 (2009), pp. 23-30.

^④MD Abdulrahman, et al., “Multimedia tools in the teaching and learning processes: A systematic review,” Heliyon, Vol. 6, No. 11 (2020), e05312.

^⑤Jared D. T. Guerrero-Sosa, et al., “A Comprehensive Review of Multimodal Analysis in Education,” Applied Sciences, Vol. 15, No. 11 (2025), p. 5896.

计必须遵循特定的教育设计原则，引导系统走向有益的认知拓展，而非有害的认知替代。^①基于 HCI 和学习科学的研究，以下设计原则至关重要：^②过程可视化原则，界面不应仅仅呈现 AI 的最终输出，而应以某种方式揭示其生成过程或信息来源。例如，在生成一段总结性文字时，界面可以高亮显示其引用的关键原文，这为学习者提供了批判性评估的支点，将 AI 从一个“黑箱”变成一个可供探究的伙伴；交互脚手架原则，界面应为学习者提供结构化的支持，以引导其进行更深度的思考和更有效的协同。例如，与其提供一个空白的输入框，界面可以提供一系列启发式问题模板（如“请从……角度分析这个问题”、“这个观点有哪些潜在的反驳意见？”），从而将交互从简单的“问-答”提升为苏格拉底式的“对话”；表征可操作原则，界面应将 AI 生成的内容视为可供学习者操作、修改和重组的“认知原材料”，而非不可更改的“最终产品”。例如，提供拖拽、合并、拆分 AI 生成内容的功能，或者将文本一键转化为思维导图，都能鼓励学习者对 AI 的输出进行二次加工和意义建构，从而确保学习者在认知回路中的主体地位。

综上所述，交互界面是 HA-EDCS 的神经中枢。通过精心设计，降低外在认知负荷，采用更符合认知规律的多模态交互，并遵循促进批判性思维的教育原则，交互界面可以成为破解“协同增效悖论”、释放人-AI 协同巨大潜力的关键杠杆。

六、模型例证：人-AI 协同写作的课堂实践分析

为了使 HA-EDCS 模型的抽象机制更加具体化，本节将通过分析已发表的课堂教学案例，展示该模型如何被用作一个分析工具，来理解和评估真实世界中的人-AI 协同学习实践。协同写作是生成式 AI 在教育中最普遍的应用场景之一，相关的课堂实践为我们提供了丰富的分析素材。

案例一：AI 作为头脑风暴伙伴与初稿撰写者

在一项研究中，教师允许学生在写作过程的任何阶段使用 AI，但要求他们必须对使用过程进行明确的记录和反思。^③许多学生选择使用 AI 进行头脑风暴和起草初稿。这个过程涉及：认知卸载与耦合，学生将“从零开始”这一认知负荷最高的阶段性任务——构思和初步组织语言——卸载给了 AI。这显著降低了写作的启动门槛，使学生能更快地进入修改和完善的阶段。此时，学生与 AI 形成了紧密的认知耦合，学生的思维流与 AI 的文本生成流交织在一起。共享表征与交互基面，AI 生成的草稿文本成为了人-AI 之间的核心共享表征。它不再是学生头脑中模糊的想法，而是一个可以被共同审视、评判和修改的外部实体。整个对话框和文本编辑器构成了交互基面。具身交互与情境生成，学生通过与 AI 进行“对话”来完善想法，这种交互模式本身就创造了一个“探索性”的学习情境。学生可以快速尝试不同的指令，观察 AI 的反应，这种低成本的试错过程鼓励了创造性思维。系统演化与能力涌现，在这种模式下，一个超越学生独立构思能力的、结构更丰富的初稿得以“涌现”。然而，该案例也凸显了认知卸载的风险：如果学生止步于此，不进行深度修改，那么 AI 就扮演了“竞争性人造物”的角色，学生的写作能力可能无法得到锻炼，甚至可能萎缩。

案例二：AI 作为批判性对话伙伴

在另一项高中 AP 英语课程的案例中，教师 Margot 设计了一个更为精巧的教学活动。她首先花费数小时，通过反复修改指令，让 ChatGPT 生成了一篇有明显缺陷但看似合理的议论文。然后，她将这篇 AI 生成的文章与一篇人类写的范文并置，要求学生进行对比分析和批判。利兹大学（University of Leeds）的

^①Y. Earnshaw, A. A. Tawfik, and M. Schmidt, “User experience design,” in Foundations of Learning and Instructional Design Technology, 2017, <https://nic.pressbooks.pub/lidtfoundations/chapter/user-experience-design/>.

^②Y. K. Chang and J. Kuwata, “Learning Experience Design: Challenges for Novice Designers,” in Foundations of Learning and Instructional Design Technology, EdTech Books, 2021, https://edtechbooks.org/ux/LXD_challenges.; Q. Wang, “Designing a Web-based constructivist learning environment,” Interactive Learning Environments, Vol. 17, No. 1 (2009), pp. 1-13.

^③本文作者前期研究，参见“具身与分布：作为整合认知系统的人-AI 协同学习模型研究（修订稿）”，未刊稿，2025 年。

公共历史课程也采用了类似的设计,让学生围绕 ChatGPT 生成的一句简短陈述展开批判性论证。^①其机制分析如下:认知卸载(反向),在这个设计中,教师利用 AI“卸载”了“生成一个高质量教学反面教材”的任务。这比教师自己去模仿一种有缺陷的写作风格更高效。共享表征与交互基面,并列呈现的两篇范文构成了强有力的共享表征,为学生的比较和批判提供了具体的认知支点。具身交互与情境生成,学生的核心具身交互行为是“剖析”和“批判”AI 的文本。这种交互模式生成了一个“审辩式思维”(critical thinking)的学习情境。学生不再是被动的知识接收者,而是主动的意义建构者和评估者。系统演化与能力涌愈,通过这个过程,学生对“什么是好的论证写作”的理解得以演化。他们学会了识别 AI 写作中常见的“空洞”和“缺乏立场”等问题,从而发展出一种批判性地看待 AI 生成内容的能力。这里的“涌现”并非一个新作品,而是一种新的、更深刻的批判性认知能力。这个案例完美地展示了如何通过精心的教学设计,将 AI 定位为一个互补性认知人造物,从而规避认知萎缩的风险。

这两个案例清晰地表明,HA-EDCS 模型的四个机制能够有效地捕捉和解释人-AI 协同学习过程中的关键动态。它揭示了协同学习的成败并不取决于 AI 技术本身,而在于由教学设计所塑造的整个认知系统的结构与运行方式。案例二的成功之处在于,教师没有将技术视为需要管理的变量,而是主动地设计了一个认知生态系统,在这个系统中,AI 的特定属性(如生成看似合理但有缺陷的文本)被巧妙地利用来激发学生的更高阶思维。这正是从“工具论”转向“系统论”的实践体现。

七、实践图景与未来展望:重塑智能时代的教与学

“人-AI 具身分布认知系统”(HA-EDCS)模型的提出,不仅是一个理论上的思辨,它更为我们描绘了一幅未来教育的实践图景,并对当前及未来的教学设计、教师角色和学习评估提出了深刻的挑战与变革要求。这一模型所蕴含的系统性、分布性和具身性视角,要求我们从根本上重新思考教育的组织方式和核心目标。

(一) 对教学设计与教师角色的启示

HA-EDCS 模型要求教学设计的焦点发生根本性转移:从设计静态的学习内容,转向设计动态的认知生态系统。^②这意味着设计的核心任务不再是创建信息本身,而是构建一个能够促进人-AI 认知系统高效、健康运作的环境。这要求教师的角色也必然发生深刻的演变。当学生能够通过与 AI 的协同来构建知识时,教师作为“知识传授者”的传统角色将逐渐式微。取而代之的,是一种更为复杂和动态的新角色——学习装置的架构师(**architect of learning apparatuses**)或人-AI 认知乐团的指挥家。^③

教师的新职责将包括:教授协同素养,教师的核心任务之一,是教授学生如何与 AI 进行有效协同。这是一种全新的、面向未来的核心素养,即“协同 AI 素养”(Collaborative AI Literacy)。^④它包括有效的提问技巧(即“提示工程”)、批判性地评估 AI 生成内容的能力、以及合乎伦理地分配认知任务的判断力。教师需要引导学生理解 AI 的局限性,并始终保持人类在价值判断和最终决策中的主导地位。这与国家政策中强调培养学生“人机协作能力”和“批判性思维”的目标高度一致。^⑤编排学习体验,教师需要像指挥家一样,精心编排人-AI 协同的学习活动。如 Margot 老师的案例所示^⑥,教师需要设计出能够将 AI 定位为互补性认知人造物的任务,知道何时引入 AI 进行头脑风暴,何时让学生对 AI 的输出进行批判性分

^①University of Leeds, “Case studies: Using generative AI in student education,” 2024, <https://generative-ai.leeds.ac.uk/ai-for-student-education/case-studies/>.

^②孟青泉、王辞晓、尚俊杰:“人工智能教育研究的问题与发展建议”,《开放教育研究》,2020年第4期,第38-47页。

^③祝智庭等:“‘近未来’人机协同教育发展新思路”,第4-13页。

^④本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

^⑤中华人民共和国教育部等六部门关于印发《中小学人工智能教育指导意见》的通知。

^⑥本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

析,何时组织同伴讨论,从而引导整个“学习乐团”奏出和谐而富有创造力的乐章。诊断与干预系统,教师需要具备诊断人-AI 认知系统运作状态的能力。当发现某个学生的协同系统运转不畅时(例如,过度依赖 AI、无法提出有效问题),教师需要及时介入,提供个性化的指导,帮助学生优化其协同策略。^①

(二) 整合认知系统的评估挑战

HA-EDCS 模型对学习评估提出了最具颠覆性的挑战。如果认知是分布的,学习成果是系统涌现的,那么传统的、仅仅针对学习者个体的知识性测验就显得捉襟见肘,甚至可能引发严重的学术诚信问题。^② 仅仅评估学生,相当于在评估一场双人网球比赛时只给其中一名选手打分,这显然无法反映整场比赛的水平和协同的效果。

因此,未来的学习评估需要进行范式革新,核心在于解决“贡献度归属”(attribution)问题,并转向对协同过程和协同能力的评估。

表 6: “协同 AI 素养” 示例性评估量规

评估维度	初阶 (Beginning)	发展中 (Developing)	熟练 (Proficient)	卓越 (Excellent)
战略性提问与迭代	提出宽泛、一次性的指令;很少或不进行追问。	能提出较为具体的指令,但迭代过程较浅,缺乏连续性。	能提出精确的指令,并通过 2-3 轮追问来优化 AI 输出。	能设计一系列逻辑连贯的指令,通过深度迭代引导 AI 共同构建复杂的解决方案。
批判性评估与整合	基本直接复制 AI 的输出,很少进行事实核查或修改。	能识别 AI 输出中明显的错误,并进行简单的修改。	能对 AI 输出进行事实核查,识别潜在偏见,并结合个人知识进行有意义的整合。	能系统性地验证 AI 输出,深入分析其局限性,并能将 AI 生成的内容作为原材料进行创造性的重构。
认知任务分配	将核心思考任务(如论点构建)完全卸载给 AI。	认知卸载的分配不合理,有时替代了本应由自己完成的核心认知活动。	能合理地区分辅助性任务和核心认知任务,主要将前者卸载给 AI。	能战略性地分配认知任务,利用 AI 处理信息密集型任务,从而将自身认知资源聚焦于最高阶的创造与决策。
过程透明度与归属	未记录或模糊地描述 AI 的使用过程。	能简单记录使用了 AI,但对具体交互过程和 AI 贡献的描述不清晰。	能清晰地记录关键的交互过程,并恰当地注明 AI 的贡献。	能提供详尽的交互日志或反思文档,清晰地区分个人贡献与 AI 贡献,并能阐述协同决策的过程。

1. 过程评估工具: 从学习分析到视觉映射

评估的重点应从仅仅关注最终的学习产品(如一篇文章),转向关注人-AI 协同的过程。这需要开发新的评估工具和方法。学习分析(Learning Analytics),通过分析学生与 AI 交互的数字足迹(如交互日志、查询历史),学习分析技术可以量化协同过程的关键指标,例如学生提问的质量、信息筛选的策略、以及对 AI 反馈的迭代优化能力。^③ 这些数据可以为教师提供关于学生协同策略的客观洞察,从而进行精准的教学干预。视觉映射工具(Visual Mapping Tools),可以引导学生使用视觉化工具来记录和反思他们

^①王一岩等:“人机协同教学:动因、本质与挑战”,第 58-65 页。

^②本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025 年。

^③René Lobo-Quintero, “AI-Enhanced Think-Pair-Share: A Learning Analytics Approach to Foster Linguistic Creative Thinking and Collaborative Learning,” Journal of Learning Analytics, Vol. 12, No. 2 (2025), pp. 19-34.

的协同过程。例如,一项研究让学生使用“技能映射图”和“协作旅程图”来记录他们在项目中如何分配人与 AI 的任务,以及在关键决策点上如何与 AI 进行协商。^① 这些视觉化的“协同过程制品”不仅能增强学生的元认知能力,也为教师评估学生的协同策略和意图性提供了具体、可分析的证据。

2. 能力评估框架:“协同 AI 素养”及其评估标准

未来的评估需要一个全新的框架来衡量学生在人-AI 协同中的核心能力。本文将此核心能力定义为“协同 AI 素养”,并提出一个可供教学实践参考的评估标准框架。

表 6 是一个“协同 AI 素养”的示例性评估量规 (Rubric),旨在为评估学生在 AI 辅助下的学习过程提供具体标准。

这个评估框架将评估的焦点从“学生知道了什么”转向“学生如何借助 AI 去知道和创造”。它不仅为教师提供了可操作的评估工具,也为学生指明了在人工智能时代所需发展的核心能力方向。^②

(三) 政策建议与未来研究

HA-EDCS 模型也为教育政策的制定提供了新的方向。联合国教科文组织 (UNESCO) 在其关于教育中 AI 应用的指导方针中,反复强调以人为本、包容和公平的原则。^③ HA-EDCS 模型与此高度契合,它将人类学习者置于系统的意图核心,并强调通过设计来确保 AI 服务于人的发展,而非取代人的能动性。政策制定者应鼓励和资助那些旨在开发“互补性”AI 学习工具和评估“协同过程”的研究与实践。^④ 中华人民共和国教育部发布的《中小学人工智能教育指导意见》也强调了培养学生的“人机协作能力”和“批判性思维”,并鼓励校企合作开发教学资源^⑤,这为 HA-EDCS 模型所倡导的实践变革提供了政策支持。

未来的研究应致力于将 HA-EDCS 模型从理论框架转化为可操作的实证研究工具。这包括:方法论创新,采用认知民族志、话语分析和多模态学习分析等方法,深入真实学习场景,细致地追踪和记录表征在人-AI 系统中的传播与转化;技术开发,开发能够自动捕捉和分析交互日志的工具,以量化协同模式(如迭代深度、问题质量)和评估协同过程;设计实验,开展受控实验,系统地比较不同界面设计、任务类型和教学干预对 HA-EDCS 系统整体表现的影响,以验证模型的预测并提炼设计原则。

结论

在人工智能与教育深度融合的时代浪潮中,本研究的核心贡献在于倡导并构建了一种理论上的范式转型:即从将 AI 视为外部辅助工具的“工具论”视角,转向将人-AI 协同体视为一个整合认知系统的“系统论”视角。这一转型不仅是必要的,更是紧迫的。它直面当前人-AI 协同实践中普遍存在的“协同增效悖论”,并主张这一困境的根源并非技术本身,而是我们用以理解和设计人-AI 关系的理论框架的贫乏。

本文通过系统地梳理和整合认知科学领域的两大前沿理论——具身认知与分布式认知,构建了“人-AI 具身分布认知系统”(HA-EDCS)模型。该模型明确指出,人-AI 协同学习的本质,是一个认知过程分布于人类学习者、AI 代理及交互界面之间,并通过学习者的具身化行动得以实现的动态过程。我们进一步阐释了该模型的四大核心运行机制——认知卸载与耦合、共享表征与交互基面、具身交互与情境生成、系统演化与能力涌现。这些机制共同揭示了人-AI 如何超越简单的功能叠加,形成一个能够演化并涌现出

^①Michael Eagle, et al., “Co-Creating with AI in Higher Education: A Human-AI Collaborative Storybook Project,” in Proceedings of the ACM on Collective Intelligence 2025 (CI '25), La Jolla, CA, USA, 2025.

^②参见 Hyunji Kim, “Development of Rubric Criteria for Assessing AI Digital Literacy,” The Journal of Korean Association of Computer Education, Vol. 28, No. 8 (2025), pp. 67-73. 以及 Philippa Hardman, “Designing Rubrics for AI-Enabled Assignments,” IDEEAS Lab, 2024, <https://ideeaslab.com/resources/teaching/ai-assignment-rubrics/>.

^③UNESCO, AI and education: Guidance for policy-makers, Paris: UNESCO, 2021.

^④本文作者前期研究,参见“具身与分布:作为整合认知系统的人-AI 协同学习模型研究(修订稿)”,未刊稿,2025年。

^⑤中华人民共和国教育部等六部门关于印发《中小学人工智能教育指导意见》的通知。

新能力的整合智能体，并为解释协同增效的条件提供了机制性的说明。

在此基础上，本研究响应审稿意见，对模型的关键环节进行了深化。通过引入视觉化图示，模型的动态循环过程变得更为直观。通过对“交互界面”的专题论述，本文揭示了其作为认知生态系统核心，在调节认知负荷、塑造协同动态方面的关键作用，并提出了具体的设计原则。最后，通过提出更具操作性的学习评估工具与框架，如过程导向的学习分析和“协同 AI 素养”评估量规，本研究将理论模型的实践意涵落到了实处。

最终，本研究认为，这种基于认知科学的系统性视角，为我们重新审视和塑造智能时代的教与学提供了全新的理论透镜。它不仅为新一代智能学习环境的设计与评估提供了原则性指导，更对教学模式的创新、教师角色的重塑以及学习评估体系的变革提出了根本性的挑战。展望未来，在人与人工智能的关系日益紧密的背景下，持续深化对人-AI 整合认知系统的理论探索与实证研究，将是推动教育走向更智能、更人性化未来的关键所在。

引用本文 李昌奎. 人-AI 具身分布认知系统 (HA-EDCS): 模型、机制与教育启示 [J]. 人工智能教育研究, 2025, 1(2):1-16. <https://doi.org/10.6914/aiescn.010201>.

Cite This Article Changkui LI.(2025). The Human - AI Embodied Distributed Cognitive System (HA-EDCS): Model, Mechanisms, and Implications for Education. *Artificial Intelligence Education Studies(Réngōng zhìnéng jiàoyù yánjiū)*, 1(2):1-16. <https://doi.org/10.6914/aiescn.010201>

The Human-AI Embodied Distributed Cognitive System (HA-EDCS): Model, Mechanisms, and Implications for Education

Changkui LI

Hong Kong Integration Research Institute

Abstract This paper moves beyond a prevalent instrumentalist view by integrating embodied cognition and distributed cognition to propose the Human-AI Embodied Distributed Cognitive System (HA-EDCS), which treats the human-AI-environment assemblage as an integrated cognitive unit. Anchored in the “synergy paradox,” the study delineates digital embodiment and cognitive artifacts, and highlights the interactional interface as the hub through which shared representations are established and cognitive flows are regulated. The model explicates four interlocking mechanisms—(1) cognitive offloading and coupling, (2) embodied interaction and context formation, (3) representational transformation and generation, and (4) system evolution and emergent capabilities. It further advances design and assessment strategies that prioritize process-oriented evaluation and complementary (rather than substitutive) AI, thereby reducing extraneous cognitive load and mitigating the risk of cognitive atrophy. By operationalizing these concepts, the paper offers an actionable theoretical and methodological framework for the design of intelligent learning environments and for pedagogical reform, providing guidance on how to orchestrate interfaces, tasks, and human-AI collaboration to realize genuine synergistic gains in learning.

Keywords Embodied Cognition; Distributed Cognition; Human-AI Collaboration; Cognitive Offloading; Interactional Interface

(责任编辑: 黄欣 cpcl@cpcl.hk)