

能动式人工智能驱动高等教育转型与重构

白娅娟

(广东外语外贸大学 现代教育技术中心、实验教学中心, 广东广州 510420)

[摘要] 新一代能动式人工智能展现出全面向任务自主规划演进的趋势,其内含的目标驱动范式对传统“学习—理解—建构—创造”的连续认知结构构成系统性解构风险。本文以此类工具的现实渗透为出发点,结合生成式人工智能在教育实践中已显现的学习空心化与评价失灵等现象,剖析能动式人工智能可能引发的高等教育存在主义危机。面向能动式人工智能时代,高等教育需建立协同演进逻辑,实现三重关键转向:从防御性应对转向人机共生结构的主动建构,从知识传授与能力培养的传统模式转向知识、能力与主体价值相融合的新型教育模式,从结果导向评价转向制度体系的系统性重构。进而提出智力贡献度分析框架与共生认知学习理论:前者从四个维度刻画人机协同的学习者主体贡献结构,为评价体系重构提供方法论支撑;后者揭示了能动式人工智能介入所产生的三类系统性认知风险,将学习界定为在人机共生系统中主动生成与确证自身认知主体性的动态过程。本研究旨在推动高等教育在主体性危机全面到来前完成理论与制度的前瞻性建构。

[关键词] 能动式人工智能;高等教育;准能动式工具;智力贡献度分析框架;共生认知学习理论

[中图分类号] G649.21

[文献标识码] A

[文章编号] 1007-2179(2026)04-0045-11

一、引言

随着大模型在文本生成、程序编写和复杂推理等方面逐步逼近乃至部分超越人类能力,人工智能已不再只是信息处理工具,开始向能够参与认知、介入决策的“准主体”演化(易凯谕等, 2025)。在高等教育领域,这一变化正在冲击以知识传授与能力培养为核心的传统教育逻辑,推动教学活动由知识灌输走向人机协同认知活动(乔雪峰, 2025)。

人工智能正在经历一次更深层次的范式跃迁:其发展已不再停留于内容生成层面,而是迅速迈向任务执行、过程规划乃至系统自主运行。以 AI Scientist 为代表的全自动科学发现系统打通了“提

出假设—实验验证—论文撰写—自我盲审”的科学全流程(Lu et al., 2024);以 ChatDev 为代表的多智能体协同系统实现了软件工程生命周期的自动化编排(Qian et al., 2024);2026 年以来,以 OpenClaw 为代表的新一代能动式人工智能(Agentic AI)更能在无需人类持续干预的条件下自主完成复杂任务、实时修正策略(Weidener et al., 2026; 沈阳, 2026)。这一演进是教育技术史上的质变:人工智能不再只是帮助人类更好地思考,开始直接替代人类思考并完成行动。

这种“人工智能代工”现象,正在逐步压缩传统教育中“学习—理解—建构—创造”的链条(叶澜, 2006),并可能引发高等教育的断裂式转型:其

[收稿日期] 2026-04-13

[修回日期] 2026-06-26

[DOI编码] 10.13966/j.cnki.kfjyyj.2026.04.005

[基金项目] 教育部产学研合作协同育人项目“‘企业行为模拟’AI智慧课程设计与实践”(241105695250126)。

[作者简介] 白娅娟,硕士,实验师、信息系统项目管理师,广东外语外贸大学现代教育技术中心、实验教学中心,研究方向:教育数智化、生成式人工智能融合教学、智慧教育(youya260@163.com)。

[引用信息] 白娅娟(2026). 能动式人工智能驱动高等教育转型与重构[J]. 开放教育研究, 32(4): 45-55.

冲击不止于教学手段层面,更触及教育的根本命题——当人工智能能高效完成认知任务时,人类学习的必要性、知识理解的价值及大学的制度功能均有待重新审视。这场冲击本质上构成对学生主体性(王道俊等, 1990)与教育秩序的系统性挑战。若高等教育无法及时重塑其运行逻辑与价值基础,主体消退、学习空心化与制度失效的风险仍将持续加剧。

当前针对能动式人工智能介入高等教育的研究尚处于起步阶段(Kostopoulos et al., 2025),缺乏系统性的理论诊断与制度框架,且少有研究深入回应“学习主体性在人机共生系统中如何维系”这一根本命题。本文从高等教育与能动式人工智能协同演进逻辑出发提出三重关键转向,进而提出智力贡献度分析框架与共生认知学习理论,以期为能动式人工智能时代高等教育前瞻性理论建构与制度重建提供参考。

二、高等教育断裂式转型

(一)能动式人工智能对传统学习理论的挑战

1. 生成式人工智能与能动式人工智能的差异

大模型能够生成高质量的文本、图像与音视频等多模态内容,但其核心范式仍停留在被动响应、单次生成、人机单向交互(桑新民等, 2025)。该类系统主要依赖用户指令完成静态内容输出,缺乏自主规划、持续行动与闭环迭代等能力。相较之下,能动式人工智能以智能体主体性(agency)为核心(Sapkot et al., 2026),通过融合目标拆解、工具调用、记忆迭代、反思优化与多主体协同机制(Kostopoulos et al., 2025),实现从按需生成内容向主动承接任务与自主推进流程的范式转变(见表 1)。

能动式人工智能的突破不在于对大模型生成能力的强化,而在于引入面向行动的闭环结构,即以大模型为认知中枢,将复杂目标拆解为可执行子任务,并结合工具调用(如代码执行、信息检索与外部应用程序编程接口)与环境交互,实现对任务情境的持续作用(Gruber et al., 2026)(见图 1)。在执行过程中,系统能基于结果调整自我评估与策略,形成“规划—执行—反馈—再规划”的循环。不同于生成式人工智能一次性输出结果的模式,此类系统生成的不仅是内容,更是可动态修正的行动路径与决策过程,从而使人工智能首次以过程性、组

织化的方式嵌入复杂任务的完成之中。

2. 教育技术奇点的临界跃迁

生成式人工智能与能动式人工智能的差异在于:前者本质上仍是概率预测者与文本生成器,其运行逻辑停留在人类提问、人工智能反馈的被动框架中,承担的是对人类智力的补充功能,尚未脱离工具属性;后者是面向过程的自主决策系统,聚焦能做成什么。这一差异成为人工智能从生成赋能走向行动赋能的关键分界点,显示出逼近教育领域技术奇点(technological singularity)(Potapov, 2018)的潜能。虽然这种技术演进的逻辑不能等同于现实教学秩序的自然坍塌,但可以作为一面镜子:当前高校面临“学生利用大语言模型批量洗稿”(Cotton et al., 2024)的乱象及强制对论文进行 AIGC 检测的技术对抗,正是这种技术奇点的初期表征。从 AI Scientist 能生成完整科研论文、ChatDev 能自主完成软件项目等现实案例看,能动式人工智能对高等教育的渗透并非遥远的预言,而是迫近的现实压力。教育中的认知主体地位正面临被结构性替代的风险。

3. 能动式人工智能对“主体亲历”原则的颠覆

传统学习理论,包括行为主义、认知主义和建构主义,都将学习者置于知识建构的核心地位。行为主义强调通过刺激—反应—强化实现行为改变,认知主义关注信息加工与策略选择,建构主义指出学习者需在真实情境中主动建构知识。这些理论的共同基础是学习的有效性依赖于学习者的主动体验与认知参与,即所谓的“主体亲历”原则。

表 1 生成式人工智能与能动式人工智能关键维度对比

维度	生成式人工智能	能动式人工智能
驱动核心	提示词驱动(prompt-driven)	目标驱动(goal-oriented)
技术本质	预测与概率分布生成	循环、规划与行动,基于思维链的自我反思和纠错
行为特征	低自主性、被动、响应式	高自主性、主动、感知式
任务复杂度	单一步骤、离散型任务	多步骤、跨系统、长期型任务
知识来源	模型内生参数	动态调用外部工具/软件
记忆机制	瞬时、上下文窗口	长期、个性化记忆库
教育价值	智能工具	学习代理人
风险类型	幻觉、错误信息、抄袭、过度依赖	越权执行、错误决策、学习过程接管

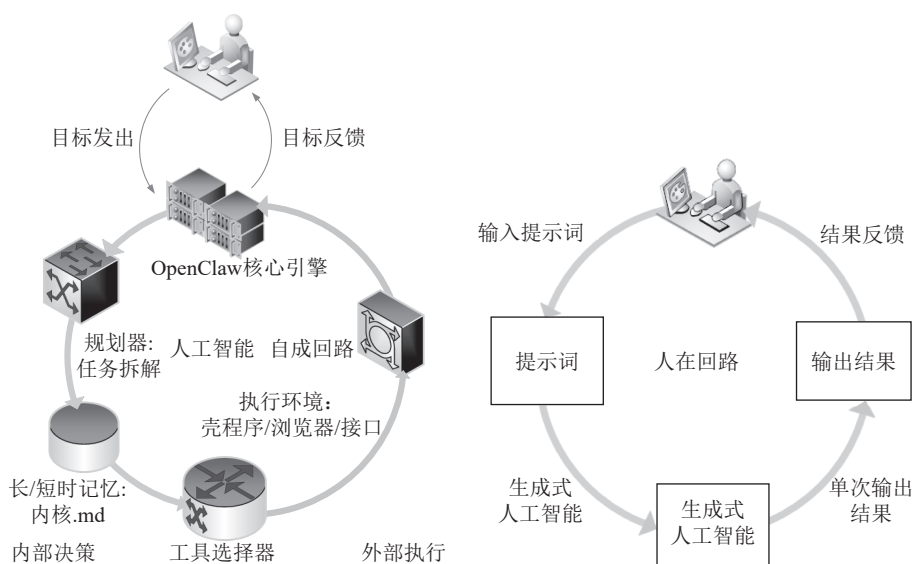


图1 能动式人工智能(以 OpenClaw 类系统为例)与生成式人工智能原理对比

从教育哲学与实践的视角看,能动式人工智能有潜力逐渐成为学习链条中的数字主体(刘革平等, 2025),它不仅有能力改变知识生产的方式,更有能力重塑学习责任、认知参与和教育评价的分配结构,直至改变认知主体格局(Kremantzis et al., 2025)。能动式人工智能主导下,学习链条中的部分能力,如知识整合、任务规划和问题求解等可能被人工智能外包,学生不必全程亲历;原本依赖学生主动参与的认知闭环被打破,传统学习理论关于“学习者中心化”的假设遭到系统性冲击。如果学习者不再是知识建构者,“主体亲历”原则在能动式人工智能的冲击下将被彻底颠覆。

(二)人工智能能力跃迁推动的逻辑断裂与范式脱轨

1. 逻辑断裂: 过程消失与因果链条的解构

能动式人工智能对高等教育的冲击,本质上在于其剥离传统教育赖以成立的过程逻辑。首先,认知投入与结果产出发生断裂。传统教育强调通过论文写作、程序开发和实验设计实现“做中学”,知识获得建立在持续的思考与投入上。能动式人工智能介入后,学生只需给出目标,智能体便可自主完成检索、分析、写作,学习过程被压缩为不可见的“认知黑箱”(Gillani et al., 2023)(见图2)。原本承担智力训练功能的作业与项目,由此失去了教育意义。其次,工具性与主体性发生断裂。传统教育技术始终服从于人的目标设定、逻辑判断与

纠偏,技术只是被控制的工具。能动式人工智能能自主反思、动态调整并独立决策,甚至可以接管由人完成的认知判断。工具因此从客体变成有执行主权的代理者,高等教育长期形成的“投入—磨砺—获得”链条随之被解构。

2. 范式脱轨: 人机协作滑向人工智能代工

上述断裂最终将高等教育推向更深层的范式脱轨。在生成式人工智能时代,学界尚能维持“人机协作”的乐观叙事:人工智能只是人类理性的“辅助驾驶”,不取代人的思考。在准能动式工具加速渗透高等教育的当下,这种人机协同的乐观叙事正在不可逆地滑向“人工智能代工”:AI Scientist可在无需学生介入下完成文献梳理与论文生成,ChatDev可自动完成课程作业层次的软件项目。

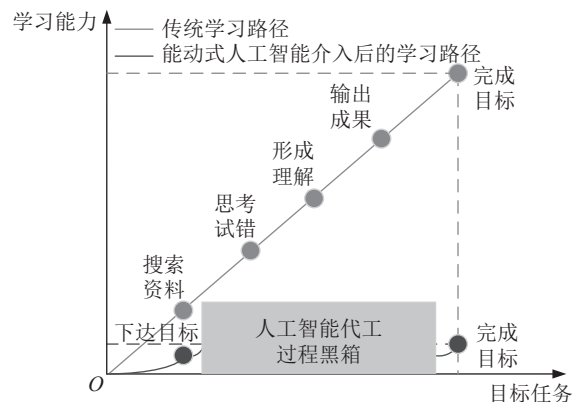


图2 传统教育与能动式人工智能介入后的学习路径比较

在这种“人工智能代工”模式下,学生交付的成果不再对应其真实能力,学习与成长之间的内在联系被切断。已有实证研究表明,学生使用智能化工具时,普遍显现出认知依赖、功能让渡与行为失控等学习异化倾向(Wu et al., 2025)。

(三)主体性坍塌和学术本体失序引发的存在主义危机

能动式人工智能具备将人的思考能力、判断能力与行动能力不断外包的技术能力,学生、教师与大学将无法维持自身的主体性,造成既有的大学制度和学术信用体系面临失序和存在主义危机(见图3)。

1. 学生: 主体空心化与能力外包临界点

传统教育中,学习不仅是知识获取的过程,更是塑造主体性的重要途径。通过反复实践、逻辑推演和批判性思维,学生形成独立判断与创造能力。然而,目前学生越来越依赖智能体完成论文撰写、实验设计、数据分析等。人工智能承担的任务比例超过能力外包临界点(cognitive outsourcing threshold)(王天平, 2026),学生的参与方式、认知结构和学习动机发生突变,学习过程被压缩或绕过,学生主体性逐渐空心化:能高效完成任务,却可能无法判断成果归属。主体空心化不仅会削弱独立思考和批判性能力,也会打破“学习—理解—建构—创造”的传统链条,学习的内在价值遭遇挑战。

2. 教师: 主体性退化与价值边缘化

教师的核心职能(知识传授、课程设计、作业评价与价值引导)有被能动式人工智能部分替代的

风险。智能体往往能更快更全面地完成备课、命题、作业批改、案例生成和教学方案设计,教师的主体角色可能逐渐退化为审核者或监督者。教师原本承担的批判性引导和价值塑造功能(田勇等, 2025)被边缘化,人文关怀与思维引领能力被削弱。这种退化不仅影响教师的职业认同,也将对课程与教育目标产生深远影响,使教育活动的核心价值——培养独立思考、责任意识和社会判断能力面临严重挑战。

3. 学术评价失灵和信用瓦解

AI Scientist 已能自主完成从选题确定到论文投稿的全流程,这一现实直接动摇了传统学术评价中“成果即智力投入”的核心假设。能动式人工智能的普及将使评价体系与信用契约面临系统性溃败的风险。当人工智能可以无缝完成从文献综述、实验设计到成果生成的整个链条时,论文不再是人类深邃思考的表征,沦为智能体对海量参数进行概率对齐的产物,评价体系因此从成果导向滑向无法识别真实贡献的“逻辑黑箱”。评价失灵进一步导致学术信用的瓦解,研究成果一旦出现错误、造假或幻觉,责任究竟属于开发者、使用者还是算法本身,已难以追溯,责任真空与知识污染可能将共同瓦解学术信用的底层契约。

4. 大学: 从学术共同体到算力竞赛场

大学制度的初衷是维护学术自治与人的主体性,但在能动式人工智能的冲击下,大学的组织逻辑将发生严重的异化与错位,科研产出不再取决于智力禀赋,而取决于对高性能智能体的访问权限及底层算力。大学制度若延续现有的“唯产出论”,将导致资源加速向拥有算力霸权的少数机构甚至个体集中,学术公平的制度设计将彻底失序(崔向平等, 2024)。

三、高等教育协同演进逻辑

能动式人工智能的出现,在根本上给教育逻辑、培养目标与制度结构带来革命性冲击。高等教育需要整体性的协同演进和底层重构。

(一)高等教育转型的逻辑起点: 从防御到共生

高等教育不是可以脱离社会技术环境而独立存在的封闭系统。大学的知识体系、能力标准与制度规则,本质上都建立在社会生产方式和技术条

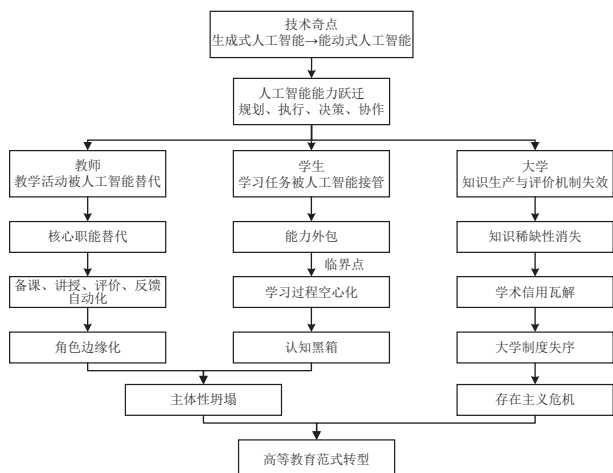


图3 高等教育断裂式转型与存在主义危机演化路径

件上。在能动式人工智能全面参与的背景下,如果大学仍然试图通过隔离人工智能、“古法学习”和技术手段识别来维持旧秩序,就等于人为制造了一个脱离现实的“教育孤岛”(Chan, 2023),如同计算机时代强迫使用算盘一样不切实际。进一步说,全社会都在向更加强大的人工智能社会演进,那从“教育孤岛”中出来的人就会难以适应人工智能全面驱动的社会。

对能动式人工智能的防御逻辑,本质上是使教育与未来社会脱节的新教育保守主义。因此,高等教育必须从防御转向共生,承认能动式人工智能甚至未来出现的更先进技术将成为社会发展的重要支撑,并与各类先进技术共生和协同演进。

共生逻辑并不意味着简单地拥抱人工智能或全面使用技术,而是强调在人与智能体之间重新建立边界、分工与协作关系(王雯等, 2025)。教育也应转向培养那些无法被外包、也不应被外包的能力。在共生逻辑下,人工智能可以承担大量低阶、重复性的认知劳动,如信息检索、资料整合、初稿写作、代码生成、格式处理等;人应更多地承担目标设定、价值判断、问题发现、意义建构、创造性整合等高阶任务。因此,共生逻辑所追求的,不是人与智能体能力上的竞争,而是人的主体性在人机协同演进中的重建。

防御逻辑与共生逻辑的差异,不是简单的技术态度差异,而是两种教育哲学之间的对立。坚持防御逻辑的哲学主张维持既有秩序和工业时代的教育范式,共生逻辑则指向坚持人的主体性这一核心的基础上建立智能时代的新教育范式。高等教育只有承认先进人工智能已经成为教育生态的一部分,才有可能进一步思考,在一个能力可以被外包的时代,它应如何捍卫自身的价值。

(二)高等教育目标的演进升级:知识、能力与主体性的协同发展

长期以来,高等教育是围绕“知识中心”展开的。大学被视为知识生产与传授的场所,学生的主要任务是积累信息、掌握知识,并在此基础上形成相应能力。随着社会对创新、实践与综合素养的关注,教育目标又进一步向“能力中心”推进,强调在掌握知识的基础上,培养问题解决、批判思维、沟通协作等能力。部分顶尖高校已将学生发展与

主体建构纳入人才培养的核心目标体系。

一旦能动式人工智能广泛介入教育,知识和能力都将不再天然作为人的独特优势和教育的全部目标。因为知识的获取门槛趋近于零,智能体可以即时调用全球知识库,其效率和质量甚至远胜于人类。越来越多原本需要长期学习才能掌握的能力,如资料搜集、文献梳理、代码生成、方案设计、论文写作,都可以由智能体无干涉自主完成。这并不意味着知识与能力培养失去意义,而是意味着如果大学仅仅停留于以知识记忆、标准化技能和结果输出等作为教育目标,未能同步强化那些难以被外包的人类主体能力,可能会陷入一种结构性困境:学生完成了学习任务,却未能真正形成能力;学生获得了学业结果,却没有真正经历学习。

因此,高等教育需要在知识传承与能力培养的基础上,将主体建构提升为新的核心目标,即在人工智能广泛介入的学习环境中,确保学习者始终保持对学习过程的主导权与意义感。这一转向并不意味着知识传授与能力培养重要性的下降,而是针对人工智能所带来的主体性侵蚀风险升级教育目标。具体而言,新增的主体中心目标至少包含三层含义:学习者应自主设定学习目标,自主进行价值判断,自主承担责任。这三个维度是人工智能难以替代且不应被替代的人类核心能力,共同构成主体性生成的基础。学习的价值不仅体现为知识获得和能力提升,也体现为主体性的成长与生成。高等教育机构只有在原有目标框架中嵌入主体建构的目标要求,才能避免在技术浪潮中不断向技术让渡功能。

(三)高等教育协同演进的实现条件:从结果评价到制度革新

如果教育目标演进为知识、能力与主体性的协同发展,那么与之匹配的评价体系与制度生态也必须同步改变。否则,学校一方面强调能力与创造,另一方面仍以标准答案、统一考试和结果产出作为依据,学生就会继续依赖人工智能完成任务,缺乏真正参与学习的动力。

评价体系需要增加主体性评价。在智能体能够轻易生成作业、报告和论文的情况下,仅仅评价最终成果已经失去意义。未来更重要的是评价学生在任务中的真实投入,包括问题提出、方案设计、

与人工智能协作的方式、关键决策与反思过程。评价内容不只包含能否完成任务,还应包含学生在任务中贡献的智力。

高校需要建立新的学术规范与责任机制。能动式人工智能的广泛应用,将使学术共同体面临责任真空与知识污染的风险:当成果由人机共同生成时,谁应承担责任、如何界定原创、如何追踪知识来源,都需要新的制度。只有建立清晰的人机责任边界与使用规范,才能避免学术信用坍塌(杜静等, 2025)。

最后,高等教育需要形成新的制度生态,涉及课程设计、评价机制、学术规范与教师培养的整体重构。制度的意义,不只是限制人工智能,更重要的是确保人工智能始终服务于人的成长,而不是取代人的成长。只有当制度能够持续推动教师、学生与人工智能协同演进,高等教育才能真正完成从知识工厂向主体共同体的转变。

四、高等教育重构路径与智力贡献度分析框架

能动式人工智能的出现可能造成传统学习过程、能力生成机制与育人因果链条面临风险,高等教育势必需要设计一种新的协同关系,使学生、教师与人工智能形成共同参与、相互增强的共生结构(谢幼如等, 2025)。因此,高等教育下一阶段的关键任务,是构建新制度:只有教学模式、评价体系与大学制度同时转型,人机协同才可能从能力外包转向能力生成,高校才能成为培育人的主体性与创造力的共同体。

(一)教学模式重构

传统教学流程可以概括为“知识输入—练习

巩固—结果输出”,其核心是学生对既有知识的吸收与再现。而在能动式人工智能时代,更适合的教学流程是“问题提出—目标设定—智能生成—判断修正—协同建构—反思迁移”;真正重要的不仅是学生能否独立完成任务,还包括其能否在任务中始终保持主体地位。这就需要重构教学模式,建立以人机协同为基础的新型学习流程。在这一流程中,人工智能不应被定位为替代人的工具,而要成为激发主体性、扩展认知边界和促进深度思考的协同对象。

新的教学模式不再是一次性的线性过程,而是不断循环、持续深化的学习闭环(见图4)。学习的终点不再是得到答案,而是形成能够不断提出问题、驾驭智能体并保持主体性的能力。

(二)智力贡献度分析框架和评价体系重构

能动式人工智能广泛介入学习过程后,高等教育首先失效的不是课堂而是评价,因为成果评价无法反映能力和贡献。人机协同条件下的评价应转向成果背后各主体的贡献。本研究提出智力贡献度分析框架,作为识别和刻画人机协同条件下学习主体性的方法论工具。

1. 智力贡献度分析框架

智力贡献度指在学习成果形成过程中,不同主体的认知努力、判断活动与创造性贡献的占比。区别于传统单纯的个人产出,智力贡献度分析框架将学习成果视为学生、教师与人工智能共同参与下的产物。评价的目标,不是排斥人工智能,而是识别在这一过程中各主体的真实贡献度,尤其是学生是否始终保持主体地位。

框架涵盖三类主体与四个核心分析维度:三类主体分别是学生、教师和人工智能系统。按完整

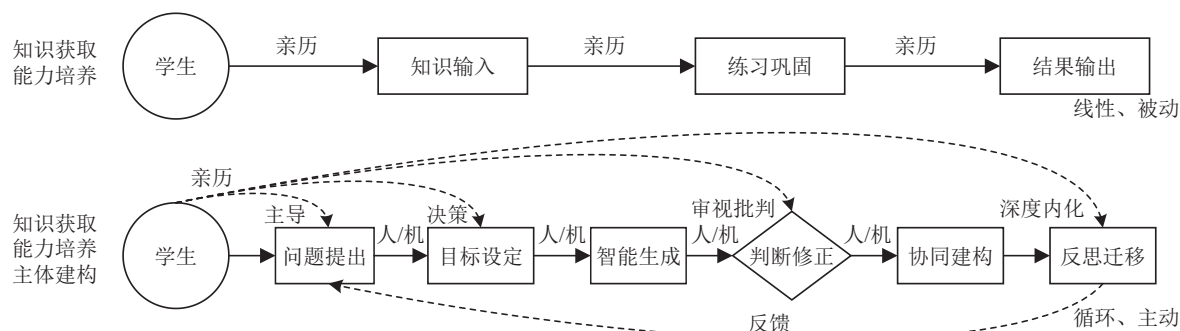


图4 教学模式重构前后对比

学习认知周期设定的四个核心分析维度分别为目标贡献 T(target)、认知贡献 K(knowledge)、决策贡献 D(decision)和反思贡献 R(reflection)。其中,目标贡献关注学习者是否自主提出问题、设定目标,认知贡献关注学习者对知识的辨别、重构与解释能力,决策贡献关注原创性判断能力、资源整合程度与协作质量,反思贡献关注学习者对过程的深度反思与迁移应用。各主体的贡献C计算式如下:

$$C = \omega_1 \cdot T + \omega_2 \cdot K + \omega_3 \cdot J + \omega_4 \cdot R \quad (1)$$

本研究根据不同任务环节学生、教师和人工智能的作用设计权重量化表, ω_1 、 ω_2 、 ω_3 、 ω_4 分别为四个维度所占权重,且 $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 1$ 。

智力贡献度分析框架不在于排斥人工智能参与,而在于回应本文核心命题——在主体性面临消解的风险下,如何使学生的认知投入与智力产出贡献可观察、可评价,尤其是其能否对目标设定、价值判断与责任承担保持主导权。由此,智力贡献度分析框架既是评价体系重构的方法论基础,也是主体性重建路径的操作性支点。具体量化标准与操作规范有待实践检验并完善。

2. 新型评价方式

在智力贡献度分析框架中,除常规的客观性证据和评价指标,教育评价体系更需要形成新的视角和评价工具,使过程、反思与主体参与真正成为评价对象。本研究建议增设以下评价方式:

1) 过程档案袋

过程档案袋可记录学生从问题提出到最终成果形成的完整过程,包括问题界定、人工智能使用记录、版本修改记录、关键决策节点、个人反思等,用于展示学生成长过程。

2) 口头答辩与追问

教师可在教学中设置常态化的口头答辩,了解学生是否真正理解任务、能否解释自己的选择、能否回应教师的追问和区分人工智能观点与个人观点,验证学生是真正参与学习还是机械调用人工智能生成的内容。

3) 人机熵值比分析

人类深度介入后的文本会呈现更复杂的语义熵值(semantic entropy)(Iyer et al., 2025),纯人工智能生成内容则更趋向于概率最优解。通过对比,系统可以自动生成人机贡献比例图。

4) 学术透明度协议(transparency protocol)(曹新明等, 2025)

教师应明确要求学生在任务附件中附上人工智能使用报告,清晰地标明人工智能生成的内容及学生的修改。

5) 具身实践教育

对于无法被人工智能外包的具身实践(embodied practice)(Allen et al., 2025),如实验室观察、田野调查和社会实践等,教师要增加权重,引导学生通过身体的介入、感官的直接参与和面对面社交的情感震荡,感受自我的存在。

(三)高校制度重构

如果高校制度仍维持原有结构,那教学模式和评价体系的转型和重构就无法真正实现。因为课程体系、学科边界、学位标准和教师培养方式,依然建立在知识和能力中心的逻辑之上。因此,高校制度也有待重构。

新教学目标不应只以知识讲授和能力培养为中心,要增加人机协同能力的培养。基础知识获取可部分交由能动式人工智能完成(如辅助自学),大学课程应更多围绕问题定义、价值判断、跨学科整合、复杂任务设计与反思能力展开,形成人工智能可替代的知识训练与人不可替代的主体性生成相融合的课程结构。

新学科结构很难遵循传统按知识边界划分的体系,取而代之的是围绕真实问题和任务场景重组的跨学科结构。能动式人工智能能够跨领域调用知识,使大学从学科分割走向问题导向(社会需求)。未来,大学更重要的不是学科的归类,而是能否利用人机协同整合多学科资源解决复杂问题、实际问题。

新学位标准授予不应再以知识掌握和结果产出为主要依据,而应增加智力贡献度和主体参与度等评价维度:不仅要看学生完成了什么,还要看哪些部分是其自主提出、判断、整合与创造的,学生是否真正形成了主体性。

新的教师角色应从知识传授者变为学习设计者、价值引导者和主体性守护者(孙发勤等, 2025)。在能动式人工智能能够代替部分教学与辅导功能的情况下,教师必须同步转型,了解如何确保学生的学习主体性以及向学生教授为何学、如何学。

未来教师培训的重点,不仅是数字技术的使用,更包含如何在人机协同条件下设计学习、评价主体性、促进学生形成真正的自我(Tan, 2025)。

五、构建新型学习理论

(一)学习理论的演化脉络与理论局限

学习理论的发展历程本质上是对“学习主体边界”的重新界定。行为主义将学习理解为可观测的刺激—反应强化过程,认知主义将其定位为人类个体心智内部的信息加工与图式建构,建构主义强调学习者基于经验主动建构知识。三种理论路径各有侧重,但共享一个未被言明的前提:学习是封闭在人类认知主体内部的,工具与环境不过是激发或承载这一过程的外部条件,认知的发生、加工与内化始终属于人的内在活动。

这一前提在相当长的时期内从未遭遇挑战。直至信息技术与网络化协作的普及,研究者才意识到,认知活动在现实中从未真正“封闭”于个体之内。分布式认知理论(Hutchins, 1995)指出,认知过程分布于人、工具与环境构成的整体系统中;活动理论(Engeström, 1987)将工具、规则与共同体纳入学习活动的结构分析框架;联通主义(Siemens, 2005)主张,知识存在于节点与连接之中,学习即在网络中形成与激活连接。上述理论突破了“认知封闭性”假设,将工具与环境纳入认知系统,代表了学习理论发展的重大进步。然而,这些理论隐含一个深层的预设:技术在功能上是中性的——工具扩展认知,但不引导认知方向。能动式人工智能系统的介入,使这一预设面临挑战。

(二)共生认知学习理论

与传统技术工具不同,能动式人工智能并非功能中性的认知介质:方向偏好,即人工智能系统的输出内嵌于训练数据的概率分布结构之中,对特定问题框架、论证路径与知识视野有系统性的倾向;能力不透明,即人工智能系统的能力边界对用户不可见,学习者无法感知哪些认知环节已形成对人工智能的依赖;概率收敛,即人工智能的反馈机制趋向最优解的概率收敛,而非拓展学习者的认知可能空间。

上述三种特性构成现有理论框架未能解释的深层问题:当认知活动的共同参与者携带隐性意向

且这种意向对学习者的不透明时,学习者的认知自主性将面临何种系统性风险,又当如何重建?分布式认知、活动理论与联通主义均未对此作出回应。

基于此,本研究提出共生认知学习理论,回应人工智能介入学习所产生的三类新机制。共生认知学习理论认为,学习是学习者在与人工智能系统构成的共生认知系统中,持续识别算法意向性对自身认知结构的塑形压力,并通过目标设定、价值判断与责任承担,主动生成与确证自身认知主体性(cognitive agency)的动态过程;若学习者未能持续激活这一能动机制,则面临元认知能力空洞化乃至认知主体性递归锁定的结构性风险。其核心主张包括:

1. 算法认知塑形

当学习者与人工智能系统持续交互时,若未形成有效的算法意识(algorithmic awareness),那人工智能系统输出的内嵌概率偏好会逐步积累形成学习者的问题框架、文献视野与论证路径,导致算法认知塑形效应(algorithmic cognitive shaping)。

2. 元认知空洞化

在算法认知塑形效应持续作用的情况下,学习者会逐步丧失对自身认知依赖边界的监控能力,形成元认知空洞化(metacognitive hollowing),即学习者既无法感知哪些认知环节对人工智能依赖,也无法判断该依赖是否侵蚀其独立认知能力。

3. 递归锁定效应

在算法认知塑形与元认知空洞化的共同作用下,学习者与人工智能系统的持续交互将产生递归锁定效应(recursive lock-in):学习者调整后的认知框架进一步影响其与人工智能的交互方式,人工智能的概率收敛性反馈又持续强化这一框架,形成思维路径的系统性窄化,且该过程对学习者的不可见。

就教育实践而言,上述理论蕴含三项教学主张:算法意识培养应作为人工智能时代学习活动的前提性任务,使学习者具备识别算法塑形的元认知能力;认知主权保护应纳入课程设计框架,通过设定无人工智能介入的认知任务区间,维持学习者独立的认知能力基线水平;统摄性学习框架(stewardship-oriented learning)应避免单纯的效能导向,将学习者对一人—人工智能共生系统的价值判断与责任承担能力纳入学习成效评估体系。共生认

知学习理论并非对传统学习理论及演化方向的否定,而是在已有框架的基础上针对人工智能系统的非中性特征提出的补充性理论。

(三)走向共生的主体性再定位与意义重建

共生认知学习理论揭示了人机共生系统中认知主体性面临系统性侵蚀的三类风险,但这一判断预设了一个尚待回答的问题:我们究竟在守护什么样的主体性?共生认知学习理论的三类风险共同指向一个结构性后果:人的目标设定、认知执行与价值体现依次被替代,主体性不再依附能力本身,人的核心角色必须从执行升级为统摄。海德格尔(2014)的“在世存在”(dasein)理念和萨特(2012)的自由与责任理念共同指向同一判断:主体性源于自我意识及为自身选择承担责任,而非任务的完成本身。能动式人工智能时代的到来,使这一命题呈现前所未有的紧迫性。

综上,本研究认为,共生认知学习理论所要守护的认知主体性,其内核体现在三个人类不可让渡的维度:1)价值判断。这是一种主观选择,也许它并不存在唯一的正确答案或最优解,无法用人工智能的逻辑或算法进行准确计算,而必须由人参与决策和判断。这种不可替代性不是因为人比人工智能更会分析,而是因为只有人能够赋予目标以意义、赋予选择以价值(Biesta, 2009)。2)责任承担。它不是一种技术功能,而是一种主体身份(Floridi et al., 2019)。即便人工智能能完成分析、规划与执行,也无法承担责任,人的责任意识、道德判断和行动选择是不可替代的。3)情感理解。它最容易被忽略,但也最具有深层的不可替代性,因为教育的意义不仅仅在于得到一个结果,更在于通过具身实践、现场观察、面对面交流等,体验身体、感官和情感等的真实感受,这些是任何高级的人工智能无法给予或替代的(Twyman et al., 2026)。正是对上述三个维度的守护,构成了共生认知学习理论三项实践主张的价值论依据:算法意识培养是为了保护价值判断不被算法塑形替换,认知主权保护是为了确保责任承担有一定认知基础,统摄性学习目标是情感理解与意义生成重新纳入学习成效评估的核心。

六、结论与展望

能动式人工智能的兴起并非简单的技术进步,

而可能引发学习范式的深层变革。当人工智能能够承接认知任务并生成行动结果时,传统以“学习—理解—建构—创造”为核心的教育链条正在被压缩甚至绕开,学习逐渐从以主体亲历为基础的过程性活动,转向以结果获取为导向的替代性实践。这一变化不只是效率提升,而是动摇了学习者作为认知主体的存在基础,从而构成以主体性消解为核心的教育危机。

因此,在能动式人工智能的介入下,学习的本质不再是个体内部的认知加工过程,而是人机系统的协同生成活动。传统以结果为中心的评价逻辑已失去有效性,必须转向对认知参与过程的重新刻画。本研究提出的智力贡献度分析框架与共生认知学习理论,分别从方法论与理论层面回应了这一断裂:前者从目标、认知、决策与反思四个维度刻画人机协同的主体贡献结构,后者揭示了算法认知塑形、元认知空洞化与递归锁定三类系统性风险,并共同指向同一个价值归宿——在协同关系中重新确立人的位置与意义。当然,本文仅以 AI Scientist、ChatDev 等准能动式工具的现实表现为实证依据,并以 OpenClaw 类系统为前沿趋势参照,相关理论框架的精细化验证,仍有待后续实证研究推进;所提出的智力贡献度分析框架和共生认知学习理论,仍有待后续检验与修正。

人工智能正持续向更高自主性与复杂性演进,学习与认知边界将被持续重塑。在这一过程中,高等教育所面对的已不再是工具性适配问题,而是关于“人如何在智能系统中持续确证自身存在”的根本问题。未来研究需持续推进人机共生机制、评价体系与学习理论的协同演化研究。高等教育应以开放的心态和超前的视野迎接技术大暴发,巩固自身提升国家竞争力的基石地位和先锋作用。

[参考文献]

- [1] Allen, L., Pratt, S., Hunte, B. L., Kligyte, G., Melvold, J., Doran, B., & Ross, K. (2025). Embodied learning for wicked problems and societal transitions: perspectives from transdisciplinary higher education practitioners[J/OL]. Higher Education Research & Development, 44, 6: 1462-1479. <https://doi.org/10.1080/07294360.2025.2514514>.
- [2] Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education[J]. Educational Assessment, Evaluation and Accountability, 21(1), 33-46.
- [3] 曹新明, 范晔(2025). 学术出版视域下人工智能使用者透明度

义务构建[J]. 中国出版, (6): 8-14.

[4] Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning[J]. *International Journal of Educational Technology in Higher Education*, 20(1), 38: 1-25.

[5] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R.(2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT[J]. *Innovations in Education and Teaching International*, 61(2): 228-239.

[6] 崔向平, 魏玲(2024). 人工智能时代教育哲学的转向——论存在主义教育的合理性与可能性[J]. *电化教育研究*, 45(4): 11-17.

[7] 杜静, 冯凯瑞, 王东(2025). 生成式人工智能背景下的高校学术道德失范及其治理策略[J]. *苏州大学学报(教育科学版)*, 13(3): 48-57.

[8] Engeström, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research[M]. Helsinki: Orienta-Konsultit Oy: 29-47.

[9] Floridi, L., & Cows, J.(2019). A unified framework of five principles for AI in society[J]. *Harvard Data Science Review*, 1(1): 1-14.

[10] Gillani, N., Eynon, R., Chiabaut, C., & Finkel, K.(2023). Unpacking the “black box” of AI in education[J]. *Educational Technology & Society*, 26(1): 99-111.

[11] Gruber, J., & Hilgert, J. N. (2026). Foundations for Agentic AI investigations from the forensic analysis of OpenClaw[J/OL]. arXiv. <https://arxiv.org/abs/2604.05589>.

[12] 海德格尔(2014). 存在与时间[M]. 陈嘉映, 王庆节, 译. 北京: 生活·读书·新知三联书店: 61-70.

[13] Hutchins, E. (1995). Cognition in the wild[M]. Cambridge, MA: MIT Press: 110-130.

[14] Iyer, K., Ravikiran, M., Pendse, P., & Mohanty, S. (2025). Towards transparent AI grading: Semantic entropy as a signal for human-AI disagreement[C/OL]//Proceedings of the 2nd Workshop on Human-Centered Explainable AI in Education (HEXED' 25). Palermo: CEUR Workshop Proceedings: 1-7. <https://arxiv.org/pdf/2508.04105>.

[15] Kostopoulos, G., Gkamas, V., Rigou, M., & Kotsiantis, S.(2025). Agentic AI in education: State of the art and future directions[J]. *IEEE Access*, 13: 177467-177491.

[16] Kremantzis, M., Essien, A., Pantano, E., & Lythreitis, S.(2025). Uncovering the Generative AI (GenAI) to Agentic AI (AgAI) shift for business school education[J]. *Journal of Global Information Management*, 33(1): 1-21.

[17] 刘革平, 秦渝超(2025). 论人工智能作为教育类主体[J]. *教育研究*, 46(5): 30-42.

[18] Lu, Ch., Lu, Co., Lange, R. T., Foerster, J., Clune, J., & Ha, D. (2024). The AI scientist: Towards fully automated open-ended scientific discovery[R/OL]. arXiv preprint arXiv: 2408.06292.

[19] Potapov, A. (2018). Technological singularity: What do we really know?[J]. *Information*, 9(4): 1-9.

[20] Qian, C., Liu, W., Liu, H., Chen, N., Dang, Y., Li, J., Yang, C., Chen, W., Su, Y., Cong, X., Xu, J., Li, D., Liu, Z., & Sun, M. (2024). ChatDev: Communicative agents for software development[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational

Linguistics (Volume 1: Long Papers). Bangkok: Association for Computational Linguistics, 2024: 15174-15186.

[21] 乔雪峰(2025). 从工具赋能到智能协同: 生成式人工智能驱动的教育模式转型[J]. *南京社会科学*, (1): 126-134.

[22] 萨特(2012). 存在主义是一种人道主义[M]. 周煦良, 汤永宽, 译. 上海: 上海译文出版社: 1-10.

[23] 桑新民, 余中, 李曙华(2025). 试论生成式人工智能的科学基础和科学范式——关于教育智能化新阶段的战略思考[J]. *教育科学*, 8(3): 3-13.

[24] Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2026). AI Agents vs. Agentic AI: A conceptual taxonomy, applications and challenges[J]. *Information Fusion*, 126: 103599: 1-30.

[25] 沈阳(2026). OpenClaw 发展研究报告 2026[R]. 北京: 清华大学清新互联网研究院: 6-19.

[26] Siemens, G.(2005). Connectivism: A learning theory for the digital age[J]. *International Journal of Instructional Technology and Distance Learning*, 2(1): 3-10.

[27] 孙发勤, 冯锐, 邓文博(2025). 守护教育之“魂”: 人工智能时代教育的价值坚守与人本回归[J]. *电化教育研究*, 46(11): 11-17.

[28] Tan, Q. (2025). Reimagining teacher development in the era of generative AI: A scoping review[J/OL]. *Teaching and Teacher Education*, 168, 105236: 1-11.

[29] 田勇, 鲁璐(2025). 人工智能赋能高等教育: 主体性困境与路径重构[J]. *成都理工大学学报(社会科学版)*, 33(5): 96-111.

[30] Twyman, M., & O'Donnell, M. B.(2026). A large language model has no body: Embodied knowledge as a key distinction in human-AI interaction[J]. *AI & Society*, 41(2): 1-3.

[31] 王道俊, 郭文安(1990). 试论教育的主体性——兼谈教育、社会与人[J]. *华东师范大学学报(教育科学版)*, (4): 33-40.

[32] 王天平, 安静芸(2026). 智能时代认知外包式教育的现实表征, 育人隐忧与规约路径[J]. *西南大学学报(社会科学版)*, 52(1): 227-237.

[33] 王雯, 邓友超(2025). 人机复合教育者的共生逻辑与实践路径[J]. *华中师范大学学报(人文社会科学版)*, 64(4): 90-99.

[34] Weidener, L, Brkić, M, Lee, P, Karlsson, M, Noessler, K, & Kohlhaas, P. (2026). From Agent-only social networks to autonomous scientific research: Lessons from OpenClaw and Moltbook, and the architecture of ClawdLab and Beach. Science[EB/OL]. [2026-02-24]. <https://arxiv.org/pdf/2602.19810>.

[35] Wu, H., Ni, H., Luo, W., & Wu, T. (2025). Development and validation of the AI dependence scale for Chinese undergraduates and a preliminary exploration [J/OL]. *Frontiers in Psychology*, 16, 1725393: 1-18.

[36] 谢幼如, 陈薇, 邱艺(2025). 人工智能赋能高校课堂教学重构研究[J]. *电化教育研究*, 46(10): 5-13.

[37] 叶澜(2006). 教育概论[M]. 北京: 人民教育出版社: 10-22.

[38] 易凯澜, 韩锡斌(2025). 从混合教学到人智协同教学: 生成式人工智能技术变革下的教学新形态[J]. *中国远程教育*, 45(4): 85-98.

(编辑: 魏志慧)

The Transformation and Reconstruction of Higher Education Driven by Agentic AI

BAI Yajuan

(Modern Educational Technology Center, Guangdong University of
Foreign Studies, Guangzhou 510420, China)

Abstract: *The new generation of agentic AI is evolving toward comprehensive autonomous task planning, and the goal-driven paradigm it embodies poses a systemic risk of deconstructing the traditional continuous cognitive structure of “learning–understanding–constructing–creating.” Considering all these and the progressive alienation phenomena in generative AI-mediated education — including the hollowing out of learning and assessment failure — this paper offers a rigorous analysis of the crisis that agentic AI may precipitate in higher education. Higher education must establish a logic of co-evolution between humans and intelligent agents, realizing three fundamental transformations: (1) from defensive adaptation to the proactive construction of human–AI symbiotic structures, (2) from traditional educational models centered on knowledge transmission to a new paradigm integrating knowledge accumulation, capability enhancement, and the cultivation of human subjectivity and values, and (3) from outcome-oriented assessment toward the systematic reconstruction of institutional evaluation systems. The study further proposes the Intellectual Contribution Framework (ICF) and Symbiotic Cognitive Learning Theory (SCLT): the former delineates the structure of learner agency in human-AI collaboration across four dimensions, providing methodological support for the systematic reconstruction of assessment; the latter identifies three categories of systemic cognitive risk introduced by agentic AI, and redefines learning as a dynamic process of actively generating and affirming one’s cognitive agency within a human-AI symbiotic system. This study aims to advance the proactive theoretical and institutional development of higher education before a full-scale crisis of human subjectivity emerges.*

Key words: *agentic AI; higher education; proto-agentic tools; intellectual contribution framework; symbiotic cognitive learning theory*