

生成式 AI 重塑教育的风险与应对

——来自国际视野的理论反思

韩冰凝

爱丁堡大学

摘要：生成式人工智能与教育融合是大势所趋，但其在教育领域中引发的一些特有风险随之而来：比如其生成内容中潜在的文化偏见、与依赖 AI 工具易带来的学术不端风险和思维惰性等。而学生认知与伦理判断尚在发展，需在教师引导下明确使用的方法来规范行为，在这一过程中需要深刻改变教师教学模式与学生学习生态。一方面，教师可借助生成式人工智能进行课程设计、素材开发与任务生成来提升教学效率与授课内容多样性；同时，该工具也可用于设计评价任务，辅助教师在评价过程中更有效地识别学生是否全部借助生成式 AI 进行回答，从而测试学生的理解深度。另一方面，学生可以通过与生成式人工智能的对话式互动获取即时反馈，从而促进学生的自主学习、元认知调节与批判性思维发展。基础教育亟需在技术赋能与教育伦理之间寻求平衡，教师与学生都需要主动调整自身，以更好地融入这场正在全面推进的教育变革。

关键词：生成式人工智能；人工智能教育；教育教学

引言

人工智能（Artificial Intelligence, AI）指的是构建能够执行智能生物相关任务的系统，例如学习、判断与决策等^[1]。当前，人工智能已广泛应用于自然科学及高新技术领域的研究，并且取得了众多显著成果^[2]。特别是生成式人工智能（Generative Artificial Intelligence, GAI）作为人工智能领域的前沿技术，依托其卓越的自然语言处理和机器学习能力，展现出深刻变革教育生态的巨大潜力^[3]。像 ChatGPT、Kimi、Deepseek 等生成式 AI 工具，通过在大规模人类对话数据集上进行训练，使其能够执行复杂任务并生成类人响应，涵盖语言生成与理解等功能^[4-5]。

生成式 AI 在教育领域的应用已引起了全球范围内的关注和探索，主要集中于对创造力的普遍信任危机、对传统教育环境的改变以及生成式 AI 所产生的教学有效性的讨论^[3,6]，2025 年 4 月，习近平主席在全国政治局第二十次集体学习会中指出，人工智能是造福人类的国际公共产品，呼吁推动人工智能教育全覆盖，包括基础教育与通识教育^[7]。因此。要在 AI 时代顺应潮流蓬勃发展，教育必须有效利用生成式 AI 的变革能力。这就需要明确人工智能在教育中的潜在风险并通过明确各方的使用责任来予以规避；重新定义教学目标、多样化教学材料、整合智能教学方法、采用多方面的方法来评估学生的进步以及促进协作式知识生成。

一、生成式 AI 在教育中引发的潜在风险

随着人工智能兴起而来的隐形歧视、隐私侵犯以

及替代人类等社会问题，已在许多研究中广泛探讨^[8-10]。然而，本文则更侧重于生成式人工智能在教育领域中所引发的争议与关注，特别是聚焦于学习者自主性、批判性思维、学术诚信等与教育密切相关的独特问题。

虽然学术不端行为并不是伴随着 AI 的出现而产生，但近年来 AI 技术的普及显著降低了剽窃的操作门槛。近年来生成式 AI 强大的文本生成能力和数据综合能力不仅冲击了学术诚信体系，更让教育者在评价学生的实际学习成效时遇到更多阻碍^[11]。值得关注的是，未成年人的认知发展尚不完善，而许多生成式 AI 工具在输出内容时，通常不会自动注明数据出处。这种特性带有潜在的诱导抄袭风险，容易致使学生在使用过程中，不自觉地引用未经标注的信息，久而久之形成无意识抄袭的习惯。一旦这种行为模式固化，将对教育的长远发展产生负面影响。

生成式人工智能的过度使用，会给教师与学生带来诸多消极后果，需引起各方重视。对于教师而言，应当防范对生成式 AI 形成过度依赖。研究显示，不恰当的技术倚重，不仅会削弱师生互动在教育中的重要价值，还会加剧因数字资源获取不均而产生的教育公平问题^[8]。与此同时，学生若过度依赖此类工具，批判性思维与创造性思维能力可能逐渐衰退。这是由于借助生成式 AI 获取答案的过程过于便捷，致使学生自主探索、独立思考以及产出原创性结论的动力被削弱。

在数据层面，当人工智能训练数据过度向特定群体倾斜时，极易滋生不公平与歧视现象^[12]。虽然诸如

ChatGPT、Kimi 等生成式 AI 工具, 被普遍认为拥有规模更大、来源更丰富的语言训练数据^[13], 但它们生成的内容仍可能延续社会中固有的刻板印象与不公平偏见^[10]。以 ChatGPT 为典型, 其在训练过程中大量采用以英语和西方文化为主的语料, 因缺乏中国本土数据支撑, 存在明显的西方中心主义倾向^[14]。实际应用中, 该模型输出的案例与论述多以西方范式为导向, 容易导致本土知识体系与文化语境被忽视。

二、生成式 AI 在中学教育中的使用责任

教育的责任本质根植于教育目的的本体论追问。教育学家格特·比斯塔 (Gert Biesta) 在论述中提出“主体化教育”概念, 强调教育应当是一种以培养学生发挥自身“主体性” (subject-ness) 为核心目标的过程, 这一过程被视为教育的根本目的^[15]。在该理论框架下, 批判性思维作为主体性教育的核心要素, 与创造性思维共同搭建起学术探索的基石^[16]。此种教育哲学可追溯至康德启蒙思想的核心理念: “勇于运用自己的理解”, 其本质在于通过教育唤醒主体理性能力的自觉性与实践勇气^[17], 这与当代教育对培养独立思考与批判精神的理念高度契合。

基于上述教育目标, 简单地排斥或禁止 ChatGPT 等生成式人工智能工具显然并非有效的路径^[18]。这一观点也与政府的智能教育政策导向一致。如教育部基础教育教学指导委员会发布的《中小学人工智能通识教育指南 (2025 年版)》, 其中明确要求将 AI 技术深度融入各学段课程体系, 着重培养学生的创新思维与实践能力, 构建全新的中小学人工智能教育生态^[19]; 联合国教科文组织在 2023 年提出的“人本主义人工智能教育”原则, 强调 AI 技术在教育应用中需坚守包容性和公平性的伦理准则^[20]。这些政策均表明当下推动生成式人工智能与教育实践深度融合已成为教育发展的必然趋势。关键在于, 在推进 AI 教育应用过程中, 如何有效保障学生主体性的充分发挥和批判性思维的持续培育。因此, 在将生成式 AI 引入课堂教学时, 教育相关主体需要构建严谨、清晰的使用规范和责任体系, 引导学生明确工具使用的边界与伦理准则, 确保合理应用的同时规避潜在风险。这不仅关乎学生个体的健康成长, 更是实现教育目标、培养具备批判性思维、责任意识与协作能力人才的重要保障。

(一) 教师的使用责任: 有效引导

早在 1762 年, 卢梭就在《爱弥儿》中强调, 教育者的核心使命是为新一代创造作为独立个体平等生存和发展的机会^[21]。在人工智能时代迅速发展的背景下, 这一理念仍然适用: 确保学生能够有效地使用生

成式人工智能工具, 不仅是延续这一教育使命的重要体现, 也是在数字化转型中为他们赢得公平发展空间的关键环节。同时, 教师自身所具备的专业知识、教学经验以及对学生个体特质的深刻理解, 仍是制定合理教学决策的核心, 而这些素养是人工智能无法取代的^[14]。换言之, 教育者与智能工具之间的协调与互补, 才是应对人工智能时代挑战并实现教育目标的可持续发展路径。因此, 在生成式人工智能广泛应用的背景下, 教师应合理利用生成式人工智能开展教学工作, 并建立有效的引导机制, 帮助学生形成对生成式人工智能的规范性认知, 培养学生在规范与指导下正确使用这类工具, 潜移默化地培养学生对人工智能技术使用的批判性审视能力。教师应在课程大纲或评估标准中明确说明生成式 AI 工具可以使用, 甚至鼓励使用这一工具。同时, 教师应为学生提供清晰、可操作的使用规范与指导原则, 明确工具使用的目的与边界, 帮助学生在负责任且符合学术规范的前提下充分发挥其积极作用。工具本身的伦理属性取决于使用主体的行为选择^[16]。当使用者接受系统训练并遵循责任框架时, 人工智能完全可以作为合法学术辅助工具, 而非剽窃的替代品。

此外, 全体教育工作者在日常教学实践中, 必须对生成式人工智能工具内含的西方中心主义倾向保持高度警觉。通过开设地方特色文化课程、补充本土知识案例等方式, 构建更加公平、多元且包容的教育环境。

(二) 学生的使用责任: 主体性发挥

在信息科技高速迭代的当下, 掌握批判性思维、团队协作能力以及数字信息素养等 21 世纪核心技能, 成为学生适应时代发展的必然要求。因此, 在运用生成式 AI 时, 学生需强化自身思辨意识, 对大数据语言模型生成内容保持审慎态度, 改变被动接收信息的习惯, 主动开展批判性分析。同时, 学生应积极履行学习主体责任, 持续提升自主学习和独立判断水平。具体而言, 在借助生成式人工智能获取知识的过程中, 必须严格恪守学术准则, 注重标明信息出处, 杜绝抄袭行为和过度依赖倾向, 让生成式 AI 切实成为推动深度学习、促进个人全面发展的有益工具, 而不是替代独立思考与探索精神的便捷手段。

三、ChatGPT 在教育中的使用策略

如前文所论, 单纯禁止学生使用生成式人工智能作为学习辅助工具难以从根本上解决对其的依赖问题^[23]。伴随生成式人工智能在教育领域的广泛渗透, 其已然成为教育环节中无法分割的重要部分。因此, 针对中学阶段的生成式人工智能使用策略, 重点应放在发挥教师的引导作用上, 推动人机协作, 优化教学

结构, 确保该技术对传统教学形成有益补充, 而非取而代之。在实际应用过程中, 需通过科学规范的实践, 合理把控使用尺度, 使生成式人工智能成为培养学生批判性思维、深化学习效果的有效工具, 避免其异化为学生逃避思考、疏于训练的途径。唯有学生能够批判性地甄别和合理运用 AI 生成内容, 教学成果的科学性与完整性才能得以保障, 学习目标也才能更好地实现。

活动式学习 (Activity-Based Learning) 为生成式人工智能融入基础教育课堂提供了切实可行的教学模式。这一教学模式以学生为中心, 通过任务驱动、合作探究和实践体验, 充分调动学生学习的主动性, 有效提升学习质量与效率^[24]。在教学实践中, 学生参与阅读、写作、讨论及问题解决等主动学习 (Active Learning) 活动, 不仅有助于深化对知识的分析、整合与评价^[25], 还能在过程中不断强化自主学习能力和元认知调节能力。值得关注的是, 活动式学习的特性与生成式人工智能的对话交互、信息支持功能高度适配。学生与生成式人工智能工具互动时, 可实时获取反馈, 并基于此进行及时的自我反思和学习策略调整, 这不仅能激发学生的内在学习动机, 还能优化学习路径。基于以上分析, 下文将围绕中学教育场景, 具体探讨如何将生成式 AI 与活动式学习有机融合, 实现两者的互补增效, 从而推动学生批判性思维发展, 提升学习成效, 助力教育目标的全面达成。

(一) 教师的教学设计策略

尽管生成式人工智能在教育领域的应用边界与伦理争议尚未完全明晰, 但其依托交互对话机制构建的新型教学范式, 切实为教学实践优化提供了技术支撑。从教学策略优化维度来看, 生成式人工智能的应用主要集中于备课与评价反馈两大核心环节^[26], 其中评价反馈策略将在后续章节展开进一步探讨。

在教学设计阶段, 生成式人工智能可作为教师的智慧协作助手。该技术通过整合与分析教学数据, 从多个维度为教师提供反馈信息, 进而协助教师优化教

学设计方案, 提高课堂教学的实际效能。同时, 将生成式人工智能应用于教学场景, 能够有效辅助教师开发丰富多样的课程资源。凭借人工智能卓越的内容生成功能, 教师得以动态更新教学素材, 将最新知识融入课堂, 从而增强课程的吸引力, 推动教学质量的提升。

除此之外, 教师在使用课堂实践时可以使用前文提到的活动式学习为框架来拓展课堂应用场景, 以充分发挥两者的互补优势。模式如下图:

(二) 教师的评价与反馈策略

在教学评价与反馈领域的相关研究表明, 尽管 ChatGPT 等工具能为学生提供具备一定质量的反馈内容, 但调查显示学生群体普遍更倾向于接受教师给予的反馈^[14]。这一现象的成因在于教师反馈不仅具有更强的针对性与真实性, 还能传递出人工智能无法替代的人际互动体验^[27]。此外, 生成式人工智能的广泛应用引发了愈发突出的抄袭问题, 当学生借助 AI 完成学习任务时, 教师往往难以精准判断其真实的学业水平, 这无疑为后续追踪学生学习过程中的问题设置了障碍^[11]。基于此, 在教学评价与反馈环节中, 教师需重点关注如何有效甄别学生对知识的实际理解深度, 避免被人工智能生成的答案所误导。将教学评价与反馈工作完全依托生成式人工智能并非合理选择。更为可行的实施路径是: 在评价与反馈过程中充分发挥教师的主体作用, 将 ChatGPT 作为辅助工具进行合理运用, 在提升反馈效率与个性化水平的同时, 保留教育活动中的人文关怀与专业判断。

生成式人工智能在为教师生成课后练习、测验等评估任务方面具有一定潜力^[28], 能够在一定程度上提升教学效率。但需要注意的是, 人工智能生成的评估任务往往难以全面覆盖具体课程的所有学习目标^[29]。因此, 教师仍需基于对学生实际学情和课程需求的深入把握, 对 AI 生成的任务进行调整与优化, 将其定位为辅助工具, 以确保评估内容与课程目标的高度契合。

在教学反馈与评价环节, 教师应重点关注两个核心维度: 一是评估任务本身的设计, 需保证其针对性与

表 1 与生成式 AI 结合的活动式学习模式

教学阶段	核心活动与内容	实施主体	主要目标 / 功能
课前	学生利用 ChatGPT 生成的预习材料完成知识输入	ChatGPT 生成材料, 教师进行材料选择	实现基础知识的初步掌握, 为课堂高阶活动奠定基础
课中	利用节省的课堂时间开展小组讨论等高阶互动活动	教师组织引导	促进知识内化、批判性思维、协作能力等核心素养发展
课后	教师运用 ChatGPT 生成多样化协作情境; 教师提供即时反馈与个性化指导	ChatGPT 生成情境; 教师反馈	及时巩固学习成果, 满足个体化需求

有效性；二是对学生学习表现的评估^[26]。此外，教师在设计评估任务时，应更加强调布鲁姆认知目标分类中高阶思维层级的能力培养，如应用、分析与创造等方面^[30]，以全面把握学生对知识的深度掌握，引导其在学习过程中展开批判性与创造性思考。以全面了解学生对知识的深度掌握情况，引导学生在学习过程中进行批判性与创造性思考^[31]。为实现这一目标，教师可参考哈拉维（Halaweh）提出的“反向角色交换策略”进行任务设计^[16]，以此有效促进学生高阶思维能力的发展。例如，教师可先利用生成式 AI 围绕特定主题生成文本材料，再要求学生 AI 生成的内容进行批判、评估与补充，具体包括检索并整合更多证据资料、检查其中的逻辑漏洞等，并在此基础上构建自身的知识框架。通过这一教学过程，学生的批判性思维与创造性能力能够得到有效锻炼，学习过程的真实性也能得到保障，从而为教育评估提供了一条可行且富有启发性的实践路径。

（三）学生在学习中的使用策略

生成式 AI 具有强大的交互对话与及时反馈功能，并能够根据学习者的不同给予个性化的反馈，这意味着生成式 AI 工具可以作为虚拟导师为学生的活动式学习提供支持^[26]。活动式学习着重培养学生自主规划任务进程的能力，而生成式人工智能能够依据个体学习节奏和兴趣偏好，灵活调整学习难度与方向，帮助学生自主构建个性化学习策略。这种主动探索的学习模式，显著区别于传统被动接受知识的方式，学生通过这种自我调节的学习行为，能够有效强化元认知能力^[32]。这是被动学习模式难以企及的效果。值得注意的是，该学习过程也是一个培养批判性思维的过程。因此，将生成式 AI 作为活动式学习的辅助手段，在一定程度上可缓解因过度依赖 AI 引发的学习主动性减退和批判性思维弱化等问题。学生也可通过与 AI 对话，锻炼沟通技能，为后续与其他学习者开展交流合作积累经验。

四、结论

综上所述，活动式学习与生成式 AI 在教育场景中融合的核心价值主要体现在三个维度：其一为强化学习者的协作能力；其二是助力学生提升学习主体性，增强内驱力^[24]；其三是培养元认知能力，促使学生通过自我调节策略持续优化学习过程，进而提升学习效果。

需要明确的是，应用生成式人工智能的教育目标并非替代活动式学习的关键实践环节，而是通过促进师生间的沟通与信息交互，为学生提供针对知识内容的即时反馈与多元视角。生成式人工智能的交互式对话功能不仅能依据学生的回答自动生成个性化回应，还能协助教师提升教学与评估效率，优化教学体验。

在生成式人工智能逐渐普及的当下，基础教育领域的参与主体应主动适应这一趋势，而非盲目排斥。教育者需合理引导学生，使其在规范与指导下有效使用生成式 AI 工具，充分释放其在促进自主学习与深度思考方面的潜能；除教育工作者外，学校与社会各界也应共同参与，推动人工智能与教育的深度融合。人工智能 + 教育的发展亟须多方协同：政府和社会应构建相应的政策支持体系，促进教育公平与技术伦理并进；学校则需完善技术基础设施与伦理规程，提升生成式人工智能在教育实践中的适切性与有效性。唯有将人工智能技术与教育目标深层对接，方能持续提升人才培养质量与教育系统效能。

参考文献：

- [1] Korteling J E H, van de Boer-Visschedijk G C, Blankendaal R A M, et al. Human-versus artificial intelligence[J]. *Frontiers in artificial intelligence*, 2021, 4: 622364.
- [2] 李国杰. 智能化科研 (AI4R): 第五科研范式 [J]. *中国科学院院刊*, 2024, 39(1): 1-9.
- [3] 胡姗姗, 陈永军, 王莉, 肖璐. 生成式人工智能背景下教师角色的重塑与困境突围 [J]. *教育观察*, 2025, 14(3): 49-53, 74.
- [4] Sušnjak T. ChatGPT: The end of online exam integrity?[EB/OL]. arXiv:2212.09292, 2022[2025-10-17]. <https://arxiv.org/abs/2212.09292>.
- [5] Xue V W, Lei P, Cho W C. The potential impact of ChatGPT in clinical and translational medicine[J]. *Clinical and Translational Medicine*, 2023, 13(3).
- [6] Corbin T, Dawson P, Nicola-Richmond K, et al. 'Where's the line? It's an absurd line': towards a framework for acceptable uses of AI in assessment[J]. *Assessment & Evaluation in Higher Education*, 2025: 1-13.
- [7] 新华社. 习近平在中共中央政治局第二十次集体学习时发表重要讲话 [EB/OL]. (2025-04-26) [2025-06-26]. https://www.gov.cn/yaowen/liebiao/202504/content_7021072.htm.
- [8] Farrokhnia M, Kazem S, Noroozi O, et al. A SWOT analysis of ChatGPT: Implications for educational practice and research[J]. *Innovations in Education and Teaching International*, 2023, 61(3): 460-474.
- [9] Nyaaba M, Wright A, Choi G L. Generative AI and digital neocolonialism in global education: Towards an equitable framework[EB/OL]. arXiv:2406.02966,

- 2024-06-16[2025-10-17]. <https://doi.org/10.48550/arXiv.2406.02966>.
- [10] Zhuo T Y, Huang Y, Chen C, Xing Z. Exploring AI ethics of ChatGPT: A diagnostic analysis[EB/OL]. arXiv:2301.12867, 2023[2025-10-17]. <https://arxiv.org/abs/2301.12867>.
- [11] Cotton D R E, Cotton P A, Shipway J R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT[J]. *Innovations in Education and Teaching International*, 2023, 61(2): 228 - 239.
- [12] Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Kasneci E, Gasser U, Groh G, Günnemann S, Hüllermeier E, Krusche S, Kutyniok G, Michaeli T, Nerdel C, Pfeffer J, Poquet O, Sailer M, Schmidt A, Kasneci G. ChatGPT for good? on opportunities and challenges of large language models for education[J]. *Learning and Individual Differences*, 2023, 103:102274.
- [13] 舒文韬, 李睿潇, 孙天祥, 等. 大型语言模型: 原理、实现与发展[J]. *计算机研究与发展*, 2024, 61(2): 351-361.
- [14] Cooper G. Examining science education in ChatGPT: An exploratory study of generative artificial intelligence[J]. *Journal of Science Education and Technology*, 2023, 32(3): 444 - 452.
- [15] 比斯塔 G. 《世界中心教育》[M]. 丁道勇译. 北京: 教育科学出版社, 2024.
- [16] Halaweh M. ChatGPT in education: Strategies for responsible implementation[J]. *Contemporary Educational Technology*, 2023, 15(2): ep421.
- [17] Kant I. Beantwortung der Frage: Was ist Aufklärung?[J]. *Berlinische Monatsschrift*, 1784, 12: 481 - 494.
- [18] 兰国帅, 肖琪, 宋帆, 等. 培养人工智能时代负责任和有创造力的公民——联合国教科文组织《学生人工智能能力框架》报告的要点与思考[J]. *开放教育研究*, 2024, 30(5):17-26.
- [19] 教育部基础教育教学指导委员会. 中小学人工智能通识教育指南(2025年版)正式发布[EB/OL]. 北京日报, (2025-05-12) [2025-06-26]. <https://xinwen.bjd.com.cn/content/s68216f0fe4b0ec1c3d96f305.html>.
- [20] Miao F, Holmes W. Guidance for generative AI in education and research[M]. Paris: UNESCO, 2023.
- [21] Rousseau J J. Émile, or On Education[M]. Translated by Barbara Foxley. New York: Basic Books, 1979.
- [22] 刘蕾, 张新亚. 人工智能依赖对创造力的影响与未来教育发展路径的省思[J]. *广西师范大学学报(哲学社会科学版)*, 2024, 60(1): 83-91.
- [23] García-Peñalvo F J. The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic?[J]. *Education in the Knowledge Society*, 2023, 24: e31279.
- [24] Bakhru S A, Mehta R P. Assignment and project activity based learning systems as an alternative to continuous internal assessment[J]. *Procedia Computer Science*, 2020, 172: 397 - 405.
- [25] Patino, A., Ramírez-Montoya, M. S., et al. Buenestado-Fernández, M. Active learning and education 4.0 for complex thinking training: analysis of two case studies in open education[J]. *Smart Learning Environments*, 2023,10(1), 8.
- [26] Lo C K. What is the impact of ChatGPT on education? A rapid review of the literature[J]. *Education Sciences*, 2023, 13(4): 410.
- [27] Jansen T, Höft L, Bahr L, et al. Comparing generative AI and expert feedback to students' writing: Insights from student teachers[J]. *Psychologie in Erziehung Und Unterricht*, 2024, 71(2).
- [28] Choi J H, Hickman K E, Monahan A, et al. ChatGPT Goes to Law School[EB/OL]. SSRN, 2023[2023-03-10]. <https://ssrn.com/abstract=4335905>.
- [29] Al Shloul T, Mazhar T, Abbas Q, et al. Role of activity-based learning and ChatGPT on students'performance in education[J]. *Computers and Education: Artificial Intelligence*, 2024, 6: 100219.
- [30] Bloom, B.S. Taxonomy of Educational Objectives: The Classification of Educational Goals[M]. New York: David McKay Company, 1956.
- [31] 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(7):26-35.
- [32] 杨宁. 从元认知到自我调节: 学习策略研究的新进展[J]. *南京师大学报(社会科学版)*, 2006(4):101-105.