

GenAI 幻觉与认知卸载：商科课程过程性考核的失效归因与对策建议

庞芳莹¹，刘荣芹^{1*}，蒋辉¹

(¹ 南宁学院 商学院，广西 南宁 530200)

摘要：生成式人工智能（GenAI）在商科教育中的广泛应用，影响教师对学生过程性考核的有效性。本文聚焦应用型高校商科专业的教学场景，指出学生因过度依赖 GenAI 产生“认知卸载”，导致逻辑推演与知识内化缺失；同时，对 AI 输出中的“算法幻觉”缺乏辨识，引发作业错误同质化，使评价失去诊断与区分功能。针对此问题，论文建议超越技术禁用思维，重构以“过程可验证”为核心的评价体系：通过提交决策路径日志、设计批判性任务（如诊断 AI 错误、人机辩论），以及推行随堂微测与抽检答辩等在场化考核，恢复过程性评价的真实性与教育价值，以期为数智时代商科人才培养提供了可行的评价改革方案。

关键词：生成式 AI；算法幻觉；认知卸载；过程性考核；商科教育

DOI：<https://doi.org/10.71411/jyyjx.2026.v1i4.1402>

GenAI Hallucinations and Cognitive Offloading: Attribution of Failure and Solutions for Process— Oriented Assessment in Business—major Courses

Pang Fangying¹, Liu Rongqin^{1*}, Jiang Hui¹

(¹ Nanning University, College of Business, Nanning, Guangxi, 530200, China)

Abstract: The prevalent adoption of Generative Artificial Intelligence (GenAI) in business education is undermining the effectiveness of process-oriented assessments. Focusing on teaching scenarios within applied universities' business major, this paper argues that students' overreliance on GenAI leads to cognitive offloading, resulting in deficiencies in logical reasoning and knowledge internalization. Concurrently, students' inability to identify the algorithmic hallucinations in AI-generated outputs causes homogenization of errors in assignments, thereby eroding the diagnostic and discriminative functions of assessment. To address this dilemma, the paper advocates moving beyond a mindset of technological prohibition and reconstructing an assessment system centered on verifiable processes. Proposed strategies include requiring submission of decision-path logs, designing critical-thinking tasks such as diagnosing AI errors or engaging in human-AI debates, and implementing in-class micro-assessments alongside ran

基金项目：广西教育科学“十四五”规划 2025 年度人工智能赋能教育专项重点课题《GenAI 幻觉诱发的应用型高校商科学生认知偏差机制与学习韧性培育研究》（项目编号：2025ZJY250）

作者简介：庞芳莹（1993-），女，广西玉林，博士，研究方向：公司金融、区域经济

刘荣芹（1995-），女，广西桂林，博士，研究方向：社交商务、人机交互

蒋辉（1991-），男，湖北石首，博士，研究方向：社交商务

通讯作者：刘荣芹，通讯邮箱：liurongqin@unn.edu.cn

do-m oral defenses to ensure authentic, on-site evaluation. This study offers a viable pathway for reforming assessment practices to cultivate business talent in the digital-intelligence era.

Keywords: Generative AI hallucination; Cognitive Offloading; Process-Oriented Assessment; Business Education

引言

随着生成式人工智能在高等教育领域的深度渗透,传统商科教学的生态系统正面临结构性重塑。研究证明,以大语言模型为基础的智能工具已具备处理复杂文本逻辑、数据分析及跨学科知识整合的高阶能力^[1]。这一技术跃迁虽然提升了信息获取效率,却对现行的教育评价体系构成了严峻挑战。在应用型本科高校的商科教学中,这种挑战表现为技术工具的强介入性与评价机制的滞后性之间的矛盾。

当前教学实践表明,学生利用智能工具完成课后任务的能力,大幅超越了其真实的认知水平。这种技术辅助下的能力溢出,导致过程性评价数据无法真实反映学生的知识掌握程度^[2]。虽然国内外学界已围绕 AI 的应用潜力、内容风险及认知影响展开探讨^{[3][4]},但现有研究多呈现出“重技术治理、轻评价重构”的特点。在人机协同已成为未来职业常态的背景下,单纯的技术封堵难以提高教学质量。基于以上背景,本文旨在分析 GenAI 导致评价失效的内在机理,探讨如何从对静态结果的核查转向对动态思维路径的审计,以期“新文科”类型课程提供教学改革思路。

1 GenAI 介入下学业评价失效的问题

在传统的商科教学评价范式中,过程性考核被假定为监测学生认知发展的有效工具。然而, AI 的介入切断了“作业产出”与“认知投入”之间的联系。

首先, AI 技术支持下认知替代。认知心理学中的“认知卸载”理论指出,个体面临高负荷任务时倾向于利用外部工具减少内部认知资源消耗^[5]。在商科课程中,学生将关键的逻辑整合与推演计算环节完全交由算法执行,自身仅保留低阶的提示词输入操作,导致学习过程“空心化”。由于缺乏必要的思维摩擦,知识点未能进入长时记忆,这解释了为何学生在平时作业中表现优异,但在脱离工具的考试中甚至无法完成基础应用^{[6][7]}。

其次,学生对 GenAI 的无批判接收。GenAI 的大语言模型在处理具有高度情境依赖性的商科问题时,常因训练数据滞后产生虚假内容^[8]。在教学实践中,这种无批判接收的结果是显著的“错误同质性”。例如,在《财务管理》课程的案例分析中,这种“认知卸载”具象表现为多名学生在分析同一上市公司的偿债能力时,引用了完全一致的错误流动比率。表明学生已将判断权让渡给算法,丧失了基本的专业鉴别力。这种现状导致了评价区分度与反馈机制的系统性失效,虚高的过程性考核成绩掩盖了真实学情,使教师误判学情并错失干预的最佳窗口期^[9]。

2 理论归因

学业评价失真现象的出现是传统考核模式在认知目标、信号筛选以及行为环境三个维度上出现偏差的综合结果。

首先,从认知维度看,传统商科考核侧重于“记忆、理解”等低阶思维能力的测查^[10]。然而,生成式 AI 的技术特性恰恰在于对标准化知识的高效重组。当考核题目仅要求解释定义或套用公式时, AI 工具能以极低成本提供答案。这导致考核目标与技术能力产生严重错位,评价体系实际上评估的是学生的工具调用能力。

其次根据信号理论,有效的评价体系必须通过设定成本门槛来区分高低能力者^[11]。理想状态

下, 高能力学生完成作业的成本较低, 而低能力者需付出极大努力, 这种差异保证了分数的信号价值。但 AI 技术已几乎零成本降低这一门槛, 由于作弊完成的成本极低且难以被发现, 导致分数无法体现学生真实能力。

最后从环境维度看, 学生的 AI 依赖行为是特定教学环境诱导下的适应性选择^[12]。当学生面临基础薄弱与课程难度高的认知落差, 且外部环境提供了缺乏监管的 AI 工具时, 利用 AI 完成作业便成为符合效率原则的行为。若考核方式仍允许学生在离线环境中完成低阶任务, 客观上便会诱发投机倾向^[13]。

3 关注过程与验证的评价体系优化

针对前文所述的评价失效问题, 单纯的技术封堵(如 AIGC 查重软件)难以应对 GenAI 的迭代。商科课程的考核改革应超越简单的防作弊视角, 建立一套适应人机共存环境的新型评价体系。本文提出, 应通过调整考核重心、创新作业形式以及回归课堂场景三个维度, 提高学业评价的信度与效度。

3.1 “思维审计”考核重点

传统的商科作业评价主要关注作业答案的准确性, 在 GenAI 时代, 这些标准化的结果极易被技术工具低成本生成, 因此已失去作为评价依据的区分度。考核的重心应从“答案是什么”转移到“答案是如何得出的”, 即建立针对学生认知逻辑的“思维审计”机制。

3.1.1 提交决策路径日志

在《财务管理》、《商务数据分析》等涉及复杂财务决策的课程中, 应要求学生在提交最终分析报告的同时, 同步提交一份“决策路径日志”。该日志的核心功能是记录学生在解决非结构化问题时的思维轨迹, 日志包括以下三个要素。

(1) 数据筛选逻辑: 解释说明为何选择特定的时间窗口或财务指标

(2) 模型选择依据: 阐述在多个备选财务模型中, 为何选择当前模型(例如: 为何选择两阶段增长模型而非固定增长模型)。

(3) 人机协作说明: 若学生使用了 AI 工具, 需在日志中明确标注 AI 提供的初步观点, 并重点说明自己进行了哪些人工修正或逻辑验证。

通过上述方式, 促使学生梳理其决策路径, 从而规避“认知卸载”带来的评价盲区。

3.1.2 引入“逆向推导”考核方法

在作业设计中, 增加“逆向推导”类题目。即给出最终的财务结果或决策结论, 要求学生反向推导出支撑该结论的假设条件和参数范围。例如, 不直接要求计算某项目的净现值, 而是设定“如果该项目可行(净现值大于 0), 那么企业的资本成本上限是多少”。此类题目要求学生必须掌握变量之间的深层勾稽关系, 使学生回归深度思考。

3.2 作业形式的创新

生成式 AI 普遍存在概率性生成的特征, 容易产生事实性错误或逻辑矛盾(即算法幻觉)。学者应当转变思路, 将 AI 的这一技术缺陷转化为教学资源, 设计“诊断与纠错”型作业, 将考核目标从低阶的记忆与理解提升至高阶的分析与评价。

3.2.1 设计“找错纠错”专项任务

在这种模式下, 过程性考核中的任务不宜让学生从零生成一份报告, 而是让学生充当报告内容“专业审核员”。具体操作步骤如下:

- (1) 教师生成一份包含逻辑漏洞、引用过时法规或数据前后矛盾的财务分析草案。
- (2) 要求学生运用课堂所学理论，找出草案中的错误，并提供修正后的正确数据或逻辑。
- (3) 根据学生识别错误的数量和修正的准确性进行评分。

例如，在《财务管理》的“资本结构优化”章节时，教师可以提供一份 AI 生成的、忽视企业经营风险且过度夸大财务杠杆正效应的分析文本。学生必须深刻理解“权衡理论”和“优序融资理论”的边界，才能识别出文本中的理论误用。这种考核方式不仅规避抄袭风险，更精准地评估学生掌握情况。

3.2.2 开展“人机辩论”课堂讨论

鼓励学生就某一争议性财务话题（如“企业是否应该进行高额现金分红”）与 AI 进行辩论。作业要求学生展示 AI 的观点，并针对 AI 观点中的逻辑漏洞进行逐条反驳。评价的重点在于学生反驳和引证的深度。这种作业形式要求学生必须站立在 AI 的肩膀上进行思考，实现从被动接受到主动批判的认知跃迁。

3.3 考核场景回归课堂

为恢复学业评价作为能力信号的真实性，需构建一个能够规避外部技术干扰的物理评价环境，并且参考李培根加强师生课堂交流和情感联络的观点^[13]，在课后作业难以有效监管的背景下，评价场景应当回归课堂，提升课堂内即时评价的权重。

3.3.1 推行随堂限时微测验

重新分配课程总评成绩的权重结构，将平时成绩中“课后作业”的占比下调，同时增加“随堂测验”的占比。具体实施上，可利用每节课的最后 10-15 分钟，进行阻断电子设备辅助的微测验。测验内容应聚焦于本节课的核心概念辨析或关键指标的推导计算，通过高频、低权重、在场化的测试，获取学生真实的知识掌握情况。

3.3.2 建立作业抽检答辩

针对在课后完成的复杂作业，建立配套的随机抽查答辩机制。教师在批改作业后，在课堂上随机抽取部分学生进行现场问询，问询的重点在于过程细节。例如随堂提问学生作业中提到的 β 系数的溯源数据库，如果学生无法在现场清晰解释其提交的作业内容，则判定该次作业为无效，记入作业成绩为 0 分。这种线上做作业、线下验真伪的组合机制，引入即时的压力测试和信用风险，向学生传递明确的信号，即仅靠 AI 技术生成的成果无法通过最终的能力验证，从而倒逼学生在课后学习中投入真实的精力。

通过上述三种策略的组合应用，能够有效识别和剔除 AI 带来的评价水分，更能引导学生建立正确的技术使用观，即从依赖技术替代思考转向利用技术辅助思考。

4 结语

4.1 研究总结

本文通过对生成式人工智能介入下商科教学评价失效现象的分析，明确了当前考核模式面临的主要问题。笔者认为在商科专业的教学场景中，学生对 AI 工具的无批判式依赖，阻断作业完成与知识习得之间的联系，无法衡量学生真实学业水平。虽然平时作业的准确率大幅提升，但这种提升往往是基于算法生成的，而非学生独立思考的结果，导致了平时成绩高、期末考试差的现象，该现象揭示了传统“结果导向”评价范式在智能化教育时代的适应性危机。鉴于此，教学评价改革的核心逻辑亟需转型，即从对作业产出正确性的单一关注，转向对学生深度思维参与度与认知过程的实质性考察。

4.2 教育启示

基于上述结论,为重建数智时代下有效的学业评价规则,高等教育应当在制度设计与教师发展两个层面进行适应性调整。

4.2.1 制度层面:从禁止转向分类规制

在人机协同已成为未来职业常态的背景下,完全禁止学生在课后使用 AI 工具既不现实,也不符合现代化人才培养的趋势。学校应当制定更为精细化的关于人工智能辅助学习行为的规范细则,明确界定“合规辅助”的边界。建议将生成式 AI 的使用场景划分为三个区域:

(1) 绿区(鼓励使用):包括新闻文件或文献检索、代码调试、头脑风暴及跨语种润色。此类使用有助于提升学习效率,应予以支持。

(2) 黄区(限制使用):包括逻辑框架搭建、数据分析解释。学生可参考 AI 生成的思路,但必须进行实质性的修改与验证,并需在作业中规范引注 AI 生成的内容。

(3) 红区(严禁使用):包括核心概念的简答阐述、关键计算过程的直接生成以及期末考核中的任何形式辅助。对此类行为应定义为学术不端并予以处罚。通过建立清晰的规则体系,引导学生建立正确的技术伦理观,消除因规则模糊而产生的投机空间。

4.2.2 师资层面:提升教师的数字评价素养

考核方式的改革对教师提出了更高的专业要求,教师的角色需要从传统的“作业批改者”转变为熟悉 AI 工具的“思维验证者”。具体地,教师需要具备以下两方面的能力:首先,具备识别 AI 生成内容的技术敏感度。教师应当熟悉主流大语言模型的文本特征(如行文风格、常见漏洞、虚构引用等),从而在批改作业时能够敏锐地发现潜在的代写痕迹。其次,具备设计“抗 AI”评价任务的能力。教师应减少对标准化、记忆性知识的考核,转而设计具有高阶思维挑战度、强情境依赖性 or 要求个性化反思的评价任务。

4.3 未来展望

未来的商科教育,应致力于建立一种“批判性共存”的人机协作关系。其中过程性考核评价体系改革目的,应当是保护学习者的独立思考能力。笔者主张通过加强过程考核和课堂验证,强调学生在使用 AI 时保持审慎态度,只有当学生具备识别算法错误、修正 AI 逻辑的能力,才能培养出真正具备专业素养、无法被简单算法替代的应用型人才。这不仅是教学评价改革的方向,更是数智时代高等教育的育人法则。

参考文献:

- [1] 李赫扬,方思清,陈凯华.人工智能驱动的科学学理论与方法重塑[J].科学学研究,2026,44(01):4-13.
- [2] 联合国教科文组织.在教育和使用生成式人工智能的指南[EB/OL].(2023-09-07)[2026-01-01].<https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>.
- [3] 金春华,张满.新商科专业人才信息化能力培养模式的探索与实践[J].高教探索,2023,(01):51-56.
- [4] Zhai X, Chu X, Chai C S, et al. A review of artificial intelligence (AI) in education from 2010 to 2020[J]. Complexity, 2021: 1-18.
- [5] Risko E F, Gilbert S J. Cognitive offloading[J]. Trends in Cognitive Sciences, 2016, 20(9):676-688.
- [6] 韦正德,林志成,张效初.人工智能时代的认知迁移:ChatGPT 加速认知卸载[J/OL].应用心理学,1-15[2026-01-12].<https://link.cnki.net/urlid/33.1012.B.20251013.1205.001>.
- [7] 齐佳音,张国锋,吴联仁.人工智能背景下的商科教育变革[J].中国大学教学,2019,(Z1):58-62.
- [8] 张力,何振,毛太田.生成式人工智能信息获取对信息茧房的双刃剑效应[J/OL].情报资料工作,

1-15[2026-01-12]. <https://link.cnki.net/urlid/11.1448.G3.20251010.1650.008>.

[9] Wu J, Yang S, Zhan R, Yuan Y, et al. A survey on LLM-generated text detection: necessity, methods, and future directions[J]. Computational Linguistics, 2025, 51 (1): 275-338.

[10] Anderson L. W, Krathwohl D R. A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives [M]. New York: Longman, 2001.

[11] Spence M. Job market signaling[J]. The Quarterly Journal of Economics, 1973, 87(3): 355-374.

[12] Bandura A. Social foundations of thought and action: A social cognitive theory[M]. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[13] 李培根. 智能时代教育需要直面的挑战[J]. 高教探索, 2025, (06): 5-9.