教育作为判断情境的设计

生成式AI使成品质量越来越难单独证明学生经历了学习过程，大学教学需从"是否产出正确答案"转向"能否校准来源、识别风险、为判断承担责任"。本文提出"判断情境设计"：教师有控制地组织来源不一、表面可信的材料，要求学生在不知结论时形成怀疑、核验、修正并说明责任边界。

作者 陈晓艳 · SDE 学员 · 发表于2026年7月25日 · 全球文库校订版

生成式人工智能能够快速生成结构完整、语言流畅的专业文本，使成品质量越来越难以单独证明学生经历了相应的学习过程。由此，大学教学需要从“是否产出正确答案”进一步转向“能否校准信息来源、识别关键风险并为判断承担责任”。本文提出“判断情境设计”：教师有控制地组织来源不一、质量不同且表面可信的材料，要求学生在不知道结论的情况下形成怀疑、核验证据、修正立场并说明责任边界。这一设计并非全新的教育理论，而是把认知学徒制、支架教学、生产性失败、错误学习与认识警觉研究重新组合到生成式 AI 的信息环境中。文章将其操作过程概括为“设标—布疑—初判—核验—反馈—迁移”六个环节，并论证三个必要限制：困难必须与学习者先备知识匹配；材料来源的暂时隐去必须经过伦理审查并在课后说明；教学目标不是训练普遍怀疑，而是形成与证据质量相匹配的信任。本文最后提出过程量规、对照研究方案和证伪条件，使“认识警觉”从修辞性目标转化为可观察、可比较的教学结果。

一、AI 架空的不是知识，而是成品作为学习证据的可靠性

围绕 AI 与大学教学的争论，常在“教学生使用”与“限制学生使用”之间摆动。前者强调就业能力和效率，后者担心学术诚信与认知依赖。两种立场都容易把注意力集中在工具本身，却忽略了评价链条的变化：当学生可以借助 AI 生成一份结构完整、措辞专业的作业时，成品与学习过程之间原本并不牢固的对应关系进一步变弱。

因此，需要修正的不是“教育等于交付”这一单一哲学命题，而是更具体的证据假设：一份合格成品，能够在多大程度上证明提交者理解了问题、检验了证据并愿意承担结论？AI 并未使知识教学失去意义，也没有使讲解、示范和练习过时；它使教师更难仅凭最终文本判断学生完成了哪些认知工作。把问题说成“交付走到尽头”过于绝对，真正被削弱的是成品的诊断效度。

这一区分决定了教学改革的方向。如果问题只是学生会不会使用工具，增加提示词训练即可；如果问题是成品不能充分显示判断过程，那么课堂必须产生新的证据：学生在哪里停下来，怀疑基于什么，如何核验，为什么改变，以及最终结论由谁负责。本文把围绕这些证据组织课堂的做法称为“判断情境设计”。

二、理论定位：它不是凭空出现的新概念

把学习置于有后果的问题情境，并非 AI 时代才出现的思想。Dewey（1938）早已强调经验的教育价值取决于经验的质量及其对后续经验的影响。判断情境设计继承这一点，但把问题进一步限定为：当信息源能够生成高度流畅、质量不一的答案时，教师如何组织经验，使学生学会校准信任而非只完成任务。

（一）认知学徒制与支架教学：让判断过程可见

认知学徒制强调建模、教练、支架、表达、反思和探索，使专家原本隐蔽的认知过程能够被学习者观察和练习（Collins，Brown，& Newman，1989）。支架研究则讨论教师如何根据学习者当前水平提供临时帮助，并随着能力提高逐步撤除（Wood，Bruner，& Ross，1976）。原稿把这些传统概括为“善意支撑”，再与“不可靠对手”作极性对立，既低估了它们包含的探索和撤架，也夸大了新概念的独立性。

判断情境设计的增量不在于否定支架，而在于改变一部分支架的对象。教师不只帮助学生获得正确答案，还要帮助他们判断何时需要怀疑、向何处寻找证据以及何时可以合理信任。对于初学者，完全撤开教师往往只会制造迷失；有效设计需要把支持从“告诉结论”转为“支持核验”。

（二）生产性失败与错误学习：困难只有被反馈接住才有价值

Kapur（2008）的生产性失败研究表明，先让学习者尝试解决复杂问题，再进行有针对性的教学，可能促进更深理解；Metcalfe（2017）对错误学习的综述也显示，错误在获得及时、可信的纠正时可以成为重要学习资源。这些研究直接挑战原稿中“判断只能从内部自行长出”的绝对说法。困难本身不会自动产生能力，错误也不会自动变成学习；产生价值的是先备知识、尝试、反馈与反思之间的组

织。

因此，AI 生成的瑕疵文本可以成为教学材料，但它不是天然的“磨刀石”。如果错误过于隐蔽，学生可能把错误记住；如果错误过于明显，任务只剩机械找错；如果教师从头到尾不提供反馈，学生无法校准自己的判断。判断情境的核心不是制造陷阱，而是设计可被发现、可被验证、可被纠正的认识困难。

（三）认识警觉：目标不是不信，而是校准信任

Sperber 等人（2010）用“认识警觉”描述人们面对交流信息时评估内容与信息源可靠性的机制。它为本文提供了比“警觉器官”更精确的理论语言。认识警觉不是普遍怀疑，更不是把 AI 预设为敌人；它要求根据证据、来源、任务风险和可核验性调整信任程度。成熟的学习者既能发现值得怀疑之处，也能在证据充分时接受可靠建议。

（四）AI 哲学文献能支持什么，不能支持什么

Dreyfus（1972）对早期人工智能的规则主义假设提出具身与情境性批评；Searle（1980）的中文屋论证质疑句法操作是否足以构成理解；Bender 等人（2021）则警示大型语言模型可能以流畅形式复制偏见、遮蔽训练材料与意义来源。这些文献都是真实而重要的思想资源，但它们不能共同推出“当前及未来 AI 必然在所有实践判断上失败”。前两者是持续存在争议的哲学论证，后一篇主要讨论语言模型的社会技术风险，而非证明永久性的能力上限。

本文因而采用更稳健的立足点：在当前教育应用中，AI 输出可能正确，也可能错误；它不能作为需要承担职业责任的法律、医疗或工程主体；输出质量会随模型、提示、数据和任务而变化。无论未来技术如何进步，学习者仍需对自己采用的证据和提交的结论负责。教学设计应围绕这一责任结构，而不是押注某项技术永远做不到什么。

三、核心概念：判断情境设计及其理论增量

判断情境设计，是教师围绕一个具有真实专业意义的判断节点，有控制地组织表面可信但证据质量不一的材料，让学生经历“接受—迟疑—核验—更新—承担”的过程。材料可以由 AI 生成，也可以来自教材中的简化、过时规范、媒体报道或人类专家的错误意见。AI 的特殊性在于它能够低成本、大规模地产生高度流畅且质量不稳定的材料，但它不是这一教学法成立的唯一条件。

这一概念真正增加了三个焦点。第一，它把“来源校准”纳入专业判断：学生不仅判断内容对不对，还判断谁说的、依据什么、风险多高。第二，它把“不确定性管理”纳入评价：学生要区分已经证实、合理推测与尚待核验。第三，它把“责任归属”放在流程末端：使用工具不取消署名者对最终结论的责任。三个焦点共同把课堂从找错竞赛提升为证据与信任的校准训练。

四、六环教学流程：设标—布疑—初判—核验—反馈—迁移

1. 设标：选择真正需要判断的节点。教师先确定专业任务中哪一步不能只靠形式正确。例如，法学中的适用前提，医学中的检查结果可信度，工程中的情境参数与规范适配。目标必须可观察，不能只写“培养主体性”。

2. 布疑：混合正确、错误与不确定材料。材料不能全是陷阱，否则学生会形成“AI 必错”的错误先验。更合理的组合包括可靠结论、可纠正错误、证据不足陈述和价值冲突，让学生练习选择性信任。错误应由教师或学科专家复核，避免引入新的事实风险。

3. 初判：要求学生先记录信任程度和理由。学生在查证前标注“接受、保留、怀疑”及其信心等级，并指出最关键的证据缺口。初判不是为了奖励直觉，而是为后续更新提供比较基线。

4. 核验：使用多种来源完成交叉检查。学生查原始法规、临床指南、规范条文、数据或权威数据库，必要时再次使用 AI 生成反例，但不得把另一次生成当作事实证明。核验过程要记录搜索路径和证据层级。

5. 反馈：公开真相并解释错误机制。教师说明哪些内容准确、哪些错误、哪些无法确定，并讨论错误来自遗漏条件、过度概括、来源失真还是价值冲突。没有这一环，困难容易变成挫败或错误记忆。

6. 迁移：更换表面特征，检验判断是否可复用。学生面对新的材料与来源，在没有“这里有错”提示的情况下再次完成审查。只有在新情境中表现改善，才能说明形成了可迁移的认识警觉，而非记住了上一题的答案。

五、三类课堂示例：从“找错”转向“校准证据”

（一）法学：适用前提而非法条编号

教师准备一份合成的法律意见书，其中多数引文真实，少数段落遗漏构成要件或把特定司法解释错误推广到另一类事实。学生先标出可信程度，再查验现行法规与完整裁判理由。评分重点不是找出几个错误，而是能否说明：哪一条前提缺失、该缺失如何改变结论、应使用何种原始来源核验。

（二）医学：检查结果的可信度而非背诵诊断

在已经完成相应临床技能训练后，教师提供一份急性腹痛记录：症状、检验与体格检查并不完全一致，其中某项体征虽记录为阴性，但操作条件和记录者经验不清。学生不能仅据该体征排除疾病，而要追问检查是怎样完成的、结果可靠性多高、还需要什么证据。该练习必须由具备临床资质的教师审核，并明确属于教学模拟，不能作为真实诊疗建议。

（三）工程：规范计算与现场条件的对应

教师提供一份形式完整的结构计算书，其中部分取值符合一般规范，却未充分反映具体地点、材料批次或节点构造。学生需要把计算式与项目条件逐项对应，区分“公式运算正确”与“模型假设适用”。优秀答案不仅指出不匹配，还提出进一步试验、细部分析或现场核查方案。

三个例子的共同结构并不是“AI 缺法感、缺身体、缺现场”，而是专业结论依赖材料之外的条件：来源、测量过程、适用范围、现场参数和责任后果。人类学习者也会遗漏这些条件。把错误专属于 AI 会削弱理论，真正需要训练的是对任何高流畅度信息源保持证据敏感。

六、必须正面处理的五个边界

（一）隐去来源涉及教学欺骗

暂时不说明材料由 AI 生成，可能提高任务真实性，却也涉及知情、信任和课程伦理。教师应提前在课程说明中告知“部分练习会暂时隐去材料来源”，确保活动不会影响学生现实权益，并在活动结束后立即说明与讨论。高风险专业、心理脆弱群体或涉及真实个人信息时，不宜采用隐匿设计。

（二）困难必须与先备知识匹配

专家的“味道不对”往往来自长期经验压缩后的模式识别，新手并不天然拥有。让知识不足的学生独自面对高度隐蔽的错误，可能把困难变成羞辱或猜谜。初学者需要核验清单、来源层级表和示范；中阶学生可以获得有限提示；高阶学生才适合在无提示材料中自主定位问题。教师撤开应是渐进的，而非原则性的完全退出。

（三）不能把不信任误当批判性

若课程持续提供错误 AI 文本，学生可能学会逢 AI 必反，而不是提高判断。认识警觉的指标应当包含“正确接受可靠信息”的能力。任务材料需要同时包含正确、错误和不确定内容，评分也要惩罚无证据否定。最终目标是校准，而非怀疑最大化。

（四）过程痕迹同样可以被生成

要求提交六百字反思并不足以证明判断真实发生，因为 AI 同样可以生成一份逼真的思考过程。更可靠的证据应组合使用：限时初判、版本记录、随机口头追问、证据地图、同伴质询和新的迁移任务。任何单一痕迹都可能被伪造，多个时间点和多种表现形式的证据更有诊断力。

（五）教学材料需要持续更新

不存在“AI 因固定本质必然犯错”的永久题库。模型能力、检索增强和领域工具不断变化，昨天有效的错误可能明天消失。教师应围绕稳定的专业风险设计任务——遗漏适用条件、混淆证据等级、忽视测量过程、责任主体不清——而不是围绕某个模型的暂时弱点。每次使用前都应重新测试和复核材料。

七、评价与研究：怎样证明学生真正学会了

判断情境设计不能只用“学生停住了”“课堂更投入”等印象证明有效。评价至少应覆盖发现、核验、更新、校准和迁移五个维度。

较强的研究设计可以比较三组学生：常规讲授组、明确告知“请找错”的审查组、来源与质量混合的判断情境组。三组接受相同知识教学，并在即时测验和延迟迁移任务中比较错误识别率、可靠信息接受率、信心校准、核验路径质量与口头辩护表现。还应记录先备知识，以检验该设计是否只对高水平学生有效。

这一主张可以被证伪。如果判断情境组只表现出更强的不信任，却没有提高准确率、信心校准和迁移；如果其优势在去掉教师提示后消失；或者活动显著增加焦虑、扩大先备知识差距，那么该设计就不能被宣称为优于常规教学。证伪条件应落在可测量的学习结果上，而不是社会是否奖励“愿意停下来的人”这类无法直接归因于课堂的宏大判断。

八、结论：课堂要培养的是有根据的信任

AI 时代的大学课堂并没有只剩下一块“机器无法进入”的人类领地。知识讲授、专家示范、技能练习和工具协作仍然重要。真正变化的是，流畅成品不再足以证明理解，专业教育必须把来源判断、证据核验、不确定性管理和责任承担变成可见的学习活动。

判断情境设计的价值，不在于把 AI 塑造成注定犯错的敌人，也不在于期待学生从陷阱中神奇地长出某种器官。它通过精心选择的材料、适度困难、独立初判、交叉核验、明确反馈和迁移练习，让学生逐步学会一件更朴素也更困难的事：既不因语言流畅而轻信，也不因来源可疑而盲目拒绝；在证据充分时接受，在条件缺失时停住，在风险上升时提高核验标准。

教育最终要保留的不是对人的独特性的浪漫想象，而是现实世界中的责任结构。工具可以提供候选答案，却不能替提交者承担法律、医疗、工程和公共决策的后果。大学应当让学生在毕业前反复练习：我凭什么相信？还有什么没有被核验？如果结论错了，我能否说明自己做过怎样的审查？当这些问题成为专业习惯，AI 的便利才可能转化为能力，而不是遮蔽能力。

最近邻加固：这一重组产生了什么单一成分都不具备的东西。　本文诚实地自承"判断情境设计"是认知学徒制、支架教学、生产性失败、错误学习与认识警觉五者的重组，而非全新理论。但一个重组要具备独立价值，必须产生一种任何单一成分都不具备的涌现属性——否则它只是清单。本设计的涌现属性，是"在不知道结论的前提下为自己的判断承担责任"这一要求：认知学徒制示范专家的判断过程，但学徒始终可对照专家这个已知答案；支架教学在已知目标下递减支持；生产性失败允许失败，但失败与否事后有标准答案可判；错误学习研究从错误中的学习，但错误是相对于正确而言的。这五者无一例外都预设了一个可供最终对照的正确答案存在于教师端。而在生成式 AI 的信息环境里，判断情境设计的要害恰恰是取消了这个教师端的正确答案：学生面对来源不一、表面皆可信的材料，在无人（包括教师）预先知道结论的情况下形成怀疑、核验、修正并说明责任边界。判决性的分离线在于是否存在可对照的正确答案：五个来源成分都在"有正确答案可最终对照"的框架内运作，而判断情境设计生产的是"在无正确答案可对照时仍须为判断负责"的能力——这种能力不是五者之和，而是抽掉了它们共同预设的那个答案锚点之后才浮现的新要求。重组之所以不只是清单，正在于此。

参考文献

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 453–494). Lawrence Erlbaum Associates.

Dewey, J. (1938). Experience and education. Macmillan.

Dreyfus, H. L. (1972). What computers can’t do: A critique of artificial reason. Harper & Row.

Kapur, M. (2008). Productive failure. Cognition and Instruction, 26(3), 379–424. https://doi.org/10.1080/07370000802212669

Metcalfe, J. (2017). Learning from errors. Annual Review of Psychology, 68, 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Searle, J. R. (1980). Minds, brains, and programs. Behavioral and Brain Sciences, 3(3), 417–424. https://doi.org/10.1017/S0140525X00005756

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. Mind & Language, 25(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

这一篇的分寸最见作者的老实：它明说「判断情境设计」并非全新的教育理论，而是把认知学徒制、支架、生产性失败、错误学习、认识警觉几条既有线索按新的评价条件重组。AI 架空的不是知识，而是成品作为学习证据的可靠性——这一句是全文的支点，也是它区别于泛泛谈 AI 教学的地方。

设计本身写得可执行：教师有控制地组织来源不一、质量不同且表面可信的材料，要求学生在不知道结论的情况下形成怀疑、核验证据、修正立场并说明责任边界。而它的理论增量也说得准确——上述五种成分都预设「有正确答案可对照」，判断情境设计取消了这个答案锚点，于是「无正确答案时仍须担责」成为一项在抽掉共同预设之后才浮现的新要求。

大学教学　把作业从「写一篇结构完整的论文」改为「在一组真伪混杂的材料里作出判断并说明你为何相信」。前者已被 AI 架空，后者尚未。

评价与学术诚信　评分对象从成品移向判断链：怀疑在哪一步产生、用什么核验、立场如何修正——这些是 AI 难以代劳且可被追问的。

职业培训　同一设计适用于任何需要现场判断的岗位培训：给出表面可信但相互冲突的信息，要求学员承担取舍。

教师备课　难点不在出题，在准备材料——需要真正来源不一、质量不同且都看起来可信的一组材料。