# 认知发生环的短路

生成式AI能在问题尚未形成时给出框架、在探索尚未展开时生成方案、在判断标准尚未建立时提供成熟表达，即时产出可能领先于能力形成。本文把学习者围绕真实问题应经历的完整认知发生环被AI跨过、事后不能重建理由或迁移的状态，界定为"认知发生环的短路"。

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在问题尚未充分形成时给出框架，在探索尚未展开时生成方案，并在学习者尚未建立判断标准时提供成熟表达。即时产出由此可能领先于能力形成。本文将学习者围绕真实问题经历的问题占有、方案生成、证据检验、判断承诺、反馈修正与迁移再用，概括为“认知发生环”；将学习者借助 AI 获得可接受成果，却未完成其中若干关键认知转换，且事后不能重建理由、识别错误或迁移到新任务的状态，界定为“认知发生环的短路”。短路不等于一切认知卸载，不以是否独立完成或是否经历痛苦为判据，也不预设封闭的个体心智。它由四种机制共同构成：问题过早闭合、选项空间被预构架、错误反馈被位移、作品质量与个体能力脱钩。文章据此提出“关键转换点”判据，区分脚手架、扩展认知与过程替代，并设计基线生成、延迟调用、反例对抗、理由重建和无辅助迁移五项教学装置。教育需要保护的不是低效率本身，而是学习者形成问题、改变判断、发现错误并在新情境中重新组织行动的能力。

## 一、当作品先于能力成熟

生成式 AI 给大学教育带来的真正难题，并不是学生能否写出合格提示词，也不仅是如何防止代写。更深的变化是，作品质量与学习者能力之间原本并不完美、但仍可辨认的关系正在松动。学生可以在尚不能独立界定问题、组织证据或诊断错误时，得到一份结构完整的论文、一段通过测试的代码或一套看似专业的方案。成果是真实的，参与也可能是真诚的，但成果究竟证明了谁具有什么能力，开始变得不清楚。

把这种现象简单归为“懒惰”会错过问题。许多学生会认真阅读生成内容，主动修改错误，甚至花费比过去更多时间反复与模型对话。他们并非什么都没做，而是认知劳动的分布发生了变化：系统承担了把模糊问题变成清晰选项、把零散材料变成解释框架、把初步判断变成完整表达的关键转换；学生更多承担调用、选择、核对和润色。执行活动增加，并不保证形成活动同样发生。

因此，教育不能只问“AI使用了多少”，而要问“哪些认知转换由谁完成”。如果系统负责了决定任务方向的转换，而学生只能在已生成的空间中确认，成品可能迅速改善，独立迁移却没有同步增长。本文用“认知发生环的短路”描述这种产出与形成相分离的风险。它不是对所有人机协作的否定，而是一项关于学习机制是否仍然闭合的诊断。

## 二、认知发生环：从线性步骤转向可更新的判断循环

### （一）六个环节与五次关键转换

杜威把反思性思维理解为由疑难情境启动、经由假设与推理并接受检验的连续探究。Papert 则借助儿童编程说明，调试不仅修复程序，也使学习者检查自己的想法。Lave 与 Wenger 进一步表明，学习不是从外部接收知识，而是在实践共同体中改变参与方式和身份。三条传统共同提示：教育的对象不只是正确答案，而是学习者与问题、证据、规则和共同体之间关系的改变。

认知发生环可以表述为六个相互递归的环节：问题占有、方案生成、证据检验、判断承诺、反馈修正、迁移再用。所谓“环”，不是要求每项任务机械走完固定程序，而是强调一次判断必须能够被后果检验，并把检验结果带回下一次行动。结论不是终点，而是下一轮问题形成的条件。

### （二）“发生”不等于全部亲手完成

认知从来不是封闭在头脑中的独立生产。书籍、公式、图表、教师、同伴和软件长期参与人的思考；Clark 与 Chalmers 的延展心灵论甚至要求重新审视认知系统的边界。因而，AI进入认知过程本身不能构成短路。一个学生借助模型寻找资料、模拟反方、检查代码或比较表达，可能比完全独立工作形成更强的判断。

关键区别在于，外部支持是否替学习者完成了本应导致能力变化的认知转换。学生不必亲自生成每个句子，却必须能够重新界定问题；不必独自检索所有资料，却必须知道证据为何相关；不必手工完成每一步计算，却必须能发现模型和前提何时不适用。发生的单位不是动作归属，而是判断结构是否在互动中得到更新。

### （三）默会判断与可问责理由

Polanyi 所说“我们所知道的多于我们能够言说的”，提醒我们不能把主体性等同于即时口头解释。专家判断往往包含难以完全表述的模式识别，学生在答辩中卡顿也不必然说明没有理解。然而，默会性不能成为不可检验的避难所。若一个判断确实形成，它通常会在行为上留下痕迹：能够识别相似问题中的关键差异，面对反例时调整，知道何时缺乏把握，并在新情境中重新组织行动。

因此，本文不以“能否完整说出思维过程”作为唯一标准，而以理由、校准和迁移三类证据互相约束。理由显示学习者能否定位关键选择；校准显示其信心是否匹配实际能力；迁移显示形成的结构能否离开原对话继续工作。三者同时薄弱，才有充分理由怀疑发生环被绕过。

## 三、何谓短路：不是快，而是关键转换没有回到学习者

短路可以被严格定义为：学习者借助外部系统获得了任务所需的可接受输出，但至少一个决定后续行动的关键转换主要由系统完成；学习者既不能重建该转换的依据，也不能在条件变化时完成同类转换。定义包含产出、过程和迁移三个层次，缺一不可。

速度快不是短路。熟练者可能迅速完成复杂判断，因为过去的练习已经压缩成默会结构；新手也可能通过高质量示例迅速掌握规律。使用现成框架同样不必然短路，学术训练本来就要求进入既有范式。风险发生在“借用”没有转化为“掌握”：框架决定了学生看见什么，但学生不知道它排除了什么；模型给出结论，学生却不能说明何种证据会使结论失效。

### （一）与认知卸载的区别

认知卸载是借助外部行动或工具降低内部加工负荷（Risko & Gilbert, 2016）。卸载通常具有适应性：把日程交给日历，把常规计算交给软件，可以释放资源处理更高层问题。短路不是卸载本身，而是卸载对象与学习目标错配。例如，课程目标是学习质性编码，学生却把第一次概念生成全部交给模型；被卸载的正是课程希望形成的能力。

### （二）与脚手架的区别

Wood、Bruner 与 Ross 提出的脚手架，指较有能力者控制学习者暂时无法处理的任务成分，使其能够集中处理当前可完成的部分。脚手架的教育价值不只在最后是否撤除，还在支持是否随能力变化而调整，以及学习者是否逐渐承担更多控制。AI若不断输出完整方案，支持强度就可能不再响应学习者的发展，结果是任务完成了，控制权却没有移交。

### （三）与生成效应和合意困难的关系

生成效应表明，在一定条件下，自己生成信息比仅仅阅读更有利于记忆；合意困难研究则强调，某些会降低即时表现、增加加工或提取要求的条件，反而有助于长期保持和迁移。但“困难”并非越多越好。只有当困难与目标能力相关、学习者能够克服并获得有效反馈时，它才具有形成性。短路理论保护的不是挫折，而是与目标能力相对应的有效加工。

## 四、四种短路机制

四种机制可以形成自我强化。问题被过早闭合后，探索只能在预设空间中进行；错误诊断外包后，学生难以知道方案为何失败；由于作品仍然优良，学生与教师又可能高估已形成的能力，下一次更愿意提前调用模型。循环的危险不在某次使用，而在产出成功不断掩盖迁移停滞。

## 五、一个复合教学情境：优秀论文何以不能证明形成已经发生

设想一名社会学本科生准备研究外卖骑手的时间压力。她阅读了若干报道，却不知道如何把兴趣转化为问题，于是请 AI 设计研究框架。模型给出“平台算法的时间规训”、深度访谈和若干预期主题。完成访谈后，她又请模型提炼材料，得到“算法内卷化、身体耗竭、身份模糊”等主题。她逐条回到材料核对，纠正了几处错误，并据此写出质量良好的论文。

这一过程不能仅凭“使用AI”被判为短路。学生核对材料、修正错误和组织写作都是真实的认知活动。问题在于，课程核心目标若是形成研究问题和从材料中生成概念，那么最关键的两次转换——由兴趣到问题、由材料到主题——已经由系统率先完成。学生进行的是框架内验证，却尚未证明自己能够构造框架。

判别需要反事实任务。让学生脱离原对话，用同一批材料提出两个结构不同的解释；加入一段与既有主题冲突的访谈，看她能否改变编码；一周后提供新的小型材料集，要求不调用 AI 完成初步概念化；同时记录她对自己表现的信心。如果她能生成替代解释、吸收反例

并迁移方法，AI可能发挥了脚手架作用；如果她只能复述模型主题，且信心显著高于独立表现，短路假设才获得支持。

这个情境也揭示了评价单位的错误。传统评分把论文当作学生能力的代理，但在人机协作中，作品已经成为复合系统的产出。作品质量适合评价团队结果，不足以单独评价学习者形成了什么。教育评价必须增加对关键转换和迁移能力的直接测量。

## 六、三项强反驳及其理论后果

### （一）观察范例也能学习，为什么必须亲自探索？

人可以通过示范、模仿和比较范例形成能力，学习并不要求每次都从零发现。若把亲历失败说成判断力的唯一来源，就会浪漫化低效率，并忽视新手需要结构化支持。真正需要区分的是“观察成品”与“学习可迁移关系”：高质量范例只有在学习者比较差异、解释选择并尝试变式时，才更可能转化为自己的判断结构。

因此，AI生成范例完全可以具有形成性。教师可以要求模型生成两个都不完美的方案，让学生诊断差异；也可以先展示解决过程，再隐藏部分步骤让学生补全。短路不由信息是否来自 AI 决定，而由学习者是否完成了目标转换决定。

### （二）未来始终有AI，为何还要测无辅助能力？

无辅助测试不是为了恢复一个孤立主体，而是用于识别复合系统中人的最低控制能力。外部系统可能不可用、给出冲突答案、受到偏见影响，或在高风险情境中需要被否决。学习者不必在没有 AI 时复制全部产出，却应能界定问题、识别异常、选择复核路径并在必要时接管关键判断。

因此，“撤除AI后能否独立完成整项任务”不是普遍标准。更合理的是功能性降级测试：当支持减少、模型出错或任务变形时，学生能否维持最关键的判断功能。教育所保护的是可接管性，而不是永远不用工具。

### （三）流畅使用AI是否正在生成一种新主体？

分布式认知和后人类主义提醒我们，主体可能不是拥有全部知识的个人，而是能够组织多种人类与技术资源的关系节点。这个反驳是成立的。未来学习者的重要能力很可能包括任务分解、模型选择、结果整合和系统监督，而不是独自生产每个局部。

但新型协同主体仍需要规范性中心：谁设定目标，谁决定证据标准，谁发现系统性盲点，谁对后果负责。若这些功能也无法定位，“主体分布化”就可能成为责任消失的修辞。认知发生环因此不必封闭在人脑内部，却必须在整个人机系统中存在可追溯的反馈，并确保学习者拥有与其责任相称的否决和接管能力。

## 七、教学设计：把AI从答案通道改造成发生装置

1. 基线生成。在第一次调用 AI 前，学生留下最小基线：当前问题、两个候选解释、最不确定之处。它不追求成熟，只为后续判断变化提供参照。

2. 延迟调用。在关键能力的初学阶段，先安排短时自主尝试，再开放 AI。延迟不是惩罚，而是让学习者形成足以与模型输出比较的内部候选。

3. 反例对抗。把 AI 用作异议者：生成反例、指出证据缺口、改变前提，而不是直接完成框架。学生必须说明哪些挑战改变了判断。

**4. 理由重建。提交成果时，不要求冗长过程日志，只选择两个高杠杆决策点，说明备选方案、采用理由和可能推翻条件。**

**5. 迁移与降级。用短小的新情境测试同类转换；逐步减少支持或植入模型错误，观察学生能否识别、纠偏和接管。**

这些装置应按风险和课程目标选择。语言润色、格式转换和常规检索不需要完整流程；问题界定、模型建立、证据解释和高后果决策则值得保留更多形成性活动。反思也不应退化为合规写作：要求学生保存全部对话，可能只增加负担并奖励事后编造。少量关键节点加行为测试，比海量过程证明更有效。

## 八、评价框架：从作品评分转向双层证据

AI介入后，评价必须同时面对两个对象：复合系统的作品表现，以及学习者自身的能力变化。二者不能互相替代。作品可以在准确性、价值和创新性上极其优秀，而学生尚未形成相应能力；学生也可能在探索中形成了重要判断，却暂时交出不够成熟的作品。

双层证据还要求改变分数解释。AI辅助下的高分不应直接推断为独立能力提高，无辅助表现下降也不必证明协作失败。真正重要的是能力曲线：随着课程推进，学生能否在更少支持下完成关键转换，能否更快发现模型错误，能否把协作中获得的结构迁移到陌生任务。

## 九、研究设计与证伪条件

认知发生环短路必须作为条件性假设接受检验，而不能成为凡使用 AI 皆可套用的批判标签。研究可以采用交叉设计，让同一学习者分别在独立生成、固定范例、AI候选支持和AI全程支持条件下完成结构相近的任务；测量即时作品质量、问题多样性、理由重建、错误发现、信心校准和一周后的无辅助迁移。

最强的总体反证是：在控制先备知识、任务难度和练习时间后，深度 AI 协作不仅提高即时作品质量，而且稳定提高学习者在陌生任务中的问题生成、错误诊断、判断校准和支持减少后的接管能力。如果这种结果可重复出现，就说明 AI 并未绕过认知发生环，而是帮助形成了结构不同但同样有效的新环路。

## 十、结论：教育需要保护的是可更新性

生成式 AI 使大学第一次大规模面对一种新情形：高质量结果可以在相应能力尚未成熟时出现。这个变化打破了作品、学习与作者能力之间的旧代理关系。若教育仍只看成果是否漂亮、工具是否熟练、使用是否合规，就可能把复合系统的成功误判为学习者已经发生的变化。

认知发生环的短路不是关于技术侵入纯粹心灵的故事，而是关于关键转换的控制和反馈是否回到学习者。问题过早闭合、选项空间预构架、错误反馈位移和作品—能力脱钩，使学生可能在每次任务中都取得成功，却没有积累下一次重新开始的能力。

大学需要保护的不是困惑、失败或低效率本身，而是认知的可更新性：学习者能够形成自己的问题，生成未被预先提供的替代方案，让判断接受反例，依据后果修正，并在新情境中重新组织行动。AI可以成为这个循环的一部分，甚至成为强有力的挑战者；但教育必须确保，系统带来的每一次能力扩展，最终都能转化为学习者更强的判断、校准、接管与负责。

最近邻加固：与"认知卸载"的判决性划界。 "认知发生环的短路"最贴身的近邻同样是认知卸载（Risko & Gilbert, 2016）。表面上，学习者借 AI 获得成果却未完成关键认知转换，像是把认知过程卸载给了 AI。但两者切在不同的层面。认知卸载是一个资源转移概念——它衡量多少认知负荷被转移到外部，其判据是"独立完成还是借助工具"，且卸载往往是有益的（省下资源做更高阶的事）。而认知发生环的短路不以是否独立完成、是否经历痛苦为判据，它界定的是一种特定的失败：学习者围绕真实问题本应经历问题占有、方案生成、证据检验、判断承诺、反馈修正、迁移再用这一完整发生环，却在其中若干关键认知转换上被 AI 的即时成熟产出直接跨过，以致事后不能重建理由、识别错误或迁移到新任务。判决性的分离线在于负荷转移与转换缺失：认知卸载预测，只要把负荷交给工具、省出资源，学习就更高效;而本文预测相反——一个把大量负荷卸载给 AI 却仍完整经历了那几个关键认知转换的学习者不构成短路,而一个只卸载了很少、却恰好跳过了"判断承诺"这一转换的学习者已经短路,因为短路的判据不是卸载了多少,而是那几个不可跳过的转换有没有真正发生。认知卸载无法区分"卸载了负荷但转换仍发生"与"卸载不多却跳过了关键转换"——这个转换判据,是本文多出的承重件。

## 参考文献


[1] Bjork, E. L., & Bjork, R. A. Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning[C]//Gernsbacher, M. A., Pew, R. W., Hough, L. M., & Pomerantz, J. R. (eds.). Psychology and the Real World. New York: Worth Publishers, 2011: 56–64.

[2] Clark, A., & Chalmers, D. The extended mind[J]. Analysis, 1998, 58(1): 7–19. DOI: 10.1093/analys/58.1.7.

[3] Dewey, J. How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process[M]. Boston: D. C. Heath, 1933.

[4] Lave, J., & Wenger, E. Situated Learning: Legitimate Peripheral Participation[M]. Cambridge: Cambridge University Press, 1991.

[5] Papert, S. Mindstorms: Children, Computers, and Powerful Ideas[M]. New York: Basic Books, 1980.

[6] Polanyi, M. The Tacit Dimension[M]. Garden City, NY: Anchor Books, 1967.

[7] Risko, E. F., & Gilbert, S. J. Cognitive offloading[J]. Trends in Cognitive Sciences, 2016, 20(9): 676–688. DOI: 10.1016/j.tics.2016.07.002.

[8] Slamecka, N. J., & Graf, P. The generation effect: Delineation of a phenomenon[J]. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4(6): 592–604. DOI: 10.1037/0278-7393.4.6.592.

[9] Vygotsky, L. S. Mind in Society: The Development of Higher Psychological Processes[M]. Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (eds.). Cambridge, MA: Harvard University Press, 1978.

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

把学习者围绕真实问题所经历的六个环节（问题占有、方案生成、证据检验、判断承诺、反馈修正、迁移再用）概括为认知发生环，再把「借助 AI 获得可接受成果、却未完成其中若干关键转换，且事后不能重建理由、识别错误或迁移」界定为短路——这个界定的好处是它不看使用多少，只看关键转换有没有回到学习者。

两处澄清尤其重要：其一，「发生」不等于全部亲手完成，这一句挡住了把本文读成禁令的误解；其二，与认知卸载、脚手架、生成效应／合意困难分别划界，说明短路不是「用了工具」也不是「难度不够」，而是转换的归属出了问题。作者还坦承默会判断与可问责理由之间的张力，这一节让概念保持了应有的粗糙度。

教学设计　把评价点放在关键转换上：让学生说明「你为什么承诺这个判断」「你在哪一步改了主意」，而不是评成品质量。

学习者自检　做完一件事后问三问：我能重建理由吗？能识别其中的错误吗？换个任务还能用吗？三问皆否，即为短路。

工程与设计团队　同一判据适用于借助工具完成的交付：能不能复述判断依据，是区分「做出来了」与「学会了」的唯一实用标准。

机构政策　与其规定 AI 使用比例，不如规定关键转换必须由本人完成并留痕——前者不可执行，后者可查。