

方程两边都有它

论一类被记入学习者账户、却不进入净变化的课堂轮次，及决定它锁死与否的不是障碍高度与延迟长度，而是首次选择的可逆性

教育学 × 化学 × 管理学

作者 王德生 + Claude

栏目 学科通融 · 站外碰撞

篇幅 约2.3万字 · 15页

发表 2026年8月3日

摘 要

课堂顺利完成、学生当场答对，而支持一撤、时间一延、情境一换，成果迅速消失。教育学把病灶放在协助的量与时机，管理学放在公开指标对被测系统的反向塑造，两家的处方互不相容，却共有一个没说出口的前提：可验收的产物属于交出它的那个人。推翻它的材料来自被引来作类比的第三家自己——化学写净离子方程时，把在方程两边同时出现、未进入焦点转化的组分划掉，称旁观离子；它在体系中、浓度可测、维持电中性、缺了它反应甚至不成立，而这些没有一条使它进入净变化。**归属由是否进入焦点转化决定，不由是否在场决定。**由此命名旁观轮次，并做一次必须做的术语修正：本文处理的是同一产物两条路径，属选择性而非速率，承重的那一步不是限速步而是**分岔步**——于是要交还给学生的不是最难的一步，是分岔的那一步；难而不分岔的代做无害，不难却分岔的代做致命。最强的一条结论来自动力学控制与热力学控制：占优的是生成最快的还是最稳定的，取决于**可逆性**而非障碍高度。课堂的首次作答高度不可逆——提交即计分、答案给出即不可未看见，因此延长测验延迟与提高任务难度根本不进入路径选择，这正是延迟测验、撤除支架与增加活动三类改革局部有效、随后被吸收的形式解释。文中另给两个当场可算的读数，并论证一门同时持有「日志活跃是学习代理」与「日志活跃可与学习负相关」两条既有结论的学科，在不承认本文变量时无法自治。

这一篇由哪三个学科的理论体系撞成

第一家主张决定学习的是协助的量与时机——给多则浅层学习，给少则挫败，最优点在中间；第二家主张病不在教学技巧而在公开指标，被测量的结果会反过来改造被测系统，故必须改指标、改责任周期；第三家主张产物不能识别路径，而多条路径中哪一条占优，第一取决于可逆性而非障碍高度。三家不能同为真，且各自的日课正是另一家诊断出的病：调协助量的方子会被评价系统在下一轮压回去（管理判教育），改指标与加延迟测验改的仍是时间与障碍、而这两个量在不可逆条件下不进入选择（化学判管理），而所谓「给第二次机会」若只是减少协助量的一个换算，可逆性就不是另一根轴（教育判化学）。三家均为站外的公开文献，链接直达原始出处，可自行核对。

教育学 · 协助两难与元认知惰性

[Koedinger & Aleven, Exploring the Assistance Dilemma in Experiments with Cognitive Tutors \(<i>Educational Psychology Review</i> 19:3, 2007\); 另见 Fan, Tang, Le, Shen, Tan, Zhao, Shen, Li & Gašević, Beware of Metacognitive Laziness \(<i>BJET</i> 56:2, 2025\); Baker, Corbett, Koedinger & Wagner, When Students “Game the System” \(CHI 2004\)](#)

主张核心变量是协助的量与时机，最优点在给多与给少之间；而这一支自己确立的求助滥用与钻系统空子结果表明，日志上最活跃的那批学生学得最少——同样的协助量落在不同步骤上后果不可比，故它不是一条曲线。

<https://doi.org/10.1007/s10648-007-9049-0>

化学 · 净离子方程、选择性决定步与动力学／热力学控制

[IUPAC <i>Compendium of Chemical Terminology</i> \(Gold Book\) : mechanism, intermediate, catalyst, rate-controlling step, kinetic control, thermodynamic control; 另见 Buskirk, Can Reaction Mechanisms Be Proven? \(<i>J. Chem. Educ.</i> 86:5, 2009\); Carey & Sundberg, <i>Advanced Organic Chemistry</i> Part A \(5th ed., Springer 2007\)](#)

写净离子方程时，在方程两边同时出现而未进入焦点转化的组分被划掉，称旁观离子——它在场、可测、甚至必要，却不在净变化里。这一条正是推翻共有前提的材料；而占优产物由可逆性而非障碍高度决定，这一条给出本文最独有的干预。

<https://goldbook.iupac.org/terms/view/R05139>

管理学 · 指标反应性与制度脱耦

[Espeland & Sauder, Rankings and Reactivity \(<i>American Journal of Sociology</i> 113:1, 2007\); 另见 Meyer & Rowan, Institutionalized Organizations \(<i>AJS</i> 83:2, 1977\); Bevan & Hood, What' s Measured Is What Matters \(<i>Public Administration</i> 84:3, 2006\)](#)

主张公开指标不是中性的镜子，它通过反应性重塑被测组织，故病灶在评价与责任结构；而这一支自己承认，脱耦可以在无人作弊的情况下长期存续——这恰恰说明指标扭曲不能单独说明认知工作在何处被替换。

<https://doi.org/10.1086/517897>

目 录

摘 要.....	5
Abstract	5
一、三条互不相容的诊断	6
二、三家共有的那个前提，与推翻它的材料.....	7

三、核心概念：旁观轮次	9
四、四种亚型：按被外置的那一步分类	10
五、分岔步：一次必须做的术语修正	10
六、形式模型：为什么在不可逆条件下延迟与障碍不进入选择	11
七、三个读数，其中两个当场可算	13
八、三轮循环与四种信息不对称	14
九、八条通道：旁观轮次如何被稳定生产	15
十、假设与判决变量	16
十一、为什么这不是效果余留、协助两难与元认知惰性的另一个名字	17
十二、被强制的不一致	18
十三、外部划界：另外八家	19
十四、站内划界：与本栏两篇先行研究的分工	20
十五、研究设计与操作化	21
十六、干预：交还分岔步，并把首次选择做成可撤回	23
十七、适用边界与三门学科的反向约束	24
十八、证伪条件、竞争解释与一条写死日期的赌注	25
十九、反噬预言与伦理限制	26
二十、自反：本文自身是一次不可逆的阅读	27
二十一、结论	27
注 释	28

摘要

课堂拥有教师、教材、活动、作业、考试与数字平台，教学流程顺利完成，学生当场给出正确答案，而知识在支持撤去、时间延迟或情境变化之后迅速失效。教育学把病灶放在协助的量与时机，管理学把病灶放在公开指标对被测系统的反向塑造。两家的处方互不相容，却共有一个没有说出口的前提：一个可验收的产物，属于交出它的那个人。推翻这个前提的材料来自被引来作类比的第三家自己。化学在写净离子方程时，把在反应式两边同时出现、未进入焦点转化的组分划掉，称为旁观离子；这类组分在体系中、浓度可测、维持电中性、缺了它反应甚至不成立——而这些没有一条构成它进入了净变化。归属由是否进入焦点转化决定，不由是否在场决定。

据此本文命名**旁观轮次**：课堂系统生成了可验收产物，学习者在场、响应并被计分，而决定目标能力形成的那一步由教师、同伴、模板或人工智能承担，制度仍把产物记入学习者账户。本文同时做一次必须做的术语修正：本文处理的是同一产物、两条路径的问题，这在化学里属选择性而非速率，因此承重的那一步不是限速步，而是**分岔步**——两条路径在此分开的那一步。修正后判据随之改变：要交还给学习者的不是最难的一步，是分岔的那一步；难而不分岔的步骤代做无害，不难却分岔的步骤代做致命。

本文最强的一条结论来自化学的动力学控制与热力学控制。占优的究竟是生成最快的产物还是最稳定的产物，取决于**可逆性与时间**，不取决于障碍高度。课堂的首次作答是高度不可逆的：提交即计分，答案一经给出便不可能未被看见，讲解一旦落地学生无法退回讲解之前。在不可逆条件下，延长测验延迟与提高任务障碍并不进入路径选择——这正是“延迟测验”“撤除支架”“增加活动”三类改革局部有效、随后被日常运行吸收的形式解释。由此得到一条与上述三者正交、可单独检验的干预：把首次选择做成可撤回一次，且撤回本身入账。

本文给出二维辨别格、分岔步归属率、**催化比**（同一项帮助在连续轮次中的用量比值，据催化剂在反应后被再生这一性质构造，无须窥测学习者内部即可当场计算）、可逆度、九组可证伪假设与多层研究设计；论证效果余留、协助两难与元认知惰性三家为何各自把病灶放错位置，以及为何一门同时持有“日志活跃是学习代理”与“日志活跃可与学习负相关”两条既有结论的学科，在不承认本文变量时无法自治。

关键词：旁观轮次；分岔步；净离子方程；可逆性；动力学控制；指标反应性；延迟迁移

Abstract

Classrooms reliably produce acceptable products — correct answers, completed worksheets, polished presentations, passing scores — without guaranteeing durable, retrievable, transferable change in learners. Educational research locates the fault in the amount and timing of assistance; management science locates it in the reactivity of public measures. The two prescriptions are incompatible, yet both rest on an unstated premise: that an accepted product belongs to whoever hands it in. The material that overturns this premise comes from

the third discipline itself. When chemists write a net ionic equation, components that appear on both sides without entering the focal transformation are cancelled out and called spectator species; such a component is present, measurable, and sometimes necessary for the reaction to proceed at all — none of which makes it part of the net change. Attribution is settled by entry into the focal transformation, not by presence.

This paper accordingly names the **spectator round**: an instructional episode in which the system generates an accepted product, the learner is present, responsive and credited, while the step that determines formation of the target capability is carried by a teacher, peer, template or generative model. It also makes a required terminological correction: the phenomenon concerns two pathways to one product, which in chemistry is a matter of selectivity rather than rate; the load-bearing step is therefore not the rate-controlling step but the **branching step**. The criterion changes with the correction — what must be returned to the learner is not the hardest step but the branching one.

The paper's strongest result derives from kinetic versus thermodynamic control. Whether the fastest-formed or the most stable product predominates depends on reversibility and time, not on barrier height. Classroom first responses are highly irreversible: submission is scoring, an answer once shown cannot be unseen, an explanation once delivered cannot be undone. Under irreversibility, lengthening test delay and raising task difficulty do not enter pathway selection at all — which formally explains why delayed testing, scaffold fading and activity-based reform prove locally effective and are then reabsorbed. A distinct, orthogonal, separately testable intervention follows: make the first commitment revocable once, and record the revocation. The paper supplies a two-by-two diagnostic matrix, a branching-step attribution index, a **catalysis ratio**, a reversibility parameter, nine falsifiable hypotheses, multilevel designs, and a demonstration that a field holding both "log-visible engagement proxies learning" and "log-visible engagement can be inversely related to learning" cannot remain consistent without the variable proposed here.

Keywords: spectator round; branching step; net ionic equation; reversibility; kinetic control; measure reactivity; delayed transfer

一、三条互不相容的诊断

现代课堂并不缺少学习的外观。教师讲解、学生记笔记、同伴讨论、平台答题、作业批改与考试反馈构成一条高度可见的生产线。课堂结束时，知识点被覆盖，任务被提交，正确率被统计，学生甚至能够复述教师刚刚给出的推理。然而一旦撤去范例与提示，延迟数日，或者把题目换成表面不同而关系同构的情境，很多课堂成果迅速消失。问题不在于课堂是否发生了活动，而在于活动改变了谁、改变了什么、这种改变能否离开原有支持继续运行。

对同一现象，三门学科给出三条不能同时为真的诊断。

教育学把病灶放在协助的量与时机。Koedinger 与 Aleven（2007）把这件事命名为协助两难：给出多于所需的信息会导致浅层学习与自学动机丧失，扣着不给则造成挫败与时间浪费；教学设计的任务是在这条曲线上找到最优点。同一支研究由此长出一整套配套结果——Aleven 等（2006）区分求助回避与求助滥用，Baker 等（2004）刻画学生“钻系统空子”的行为，Fan 等（2024）在随机实验中发现使用生成式模型的一组作文提分最多而知识增益与迁移并无差异，并把它称为元认知惰性。这一支的处方是明确的：调协助的量、改协助的形态、训练学习者的自我调节。

管理学把病灶放在公开指标对被测系统的反向塑造。Kerr（1975）以“奖励甲却指望乙”揭示奖励结构与组织期望的错配；Meyer 与 Rowan（1977）说明正式结构可与核心活动松散耦合而仍因合法性存续；March（1991）指出组织偏向回报更快更确定的利用活动；Espeland 与 Sauder（2007）展示公开指标如何通过反应性重塑组织行为；Bevan 与 Hood（2006）表明目标可以改善报告结果同时诱发替代、选择与博弈。这一支的处方同样明确：病不在教师的教学技巧，在评价与责任结构，故必须改指标、改责任周期、改资源约束。

化学被引来解释路径，而它自己的主张与前两家都不一样。从反应物到产物的过程由步骤、中间体、过渡态与各步动力学共同刻画；复合反应可沿多条路径到达相同产物；催化剂提供不同路径并改变速率；Buskirk（2009）提醒产物与少量宏观读数通常只能支持或排除机理而不能给出唯一机理。更要紧的是这一支关于占优产物的判断：当反应不可逆或时间不足，占优的是生成最快的那个产物；当反应可逆且给足时间，占优的是最稳定的那个。**决定哪一个占优的第一变量是可逆性，不是障碍高度。**

三条诊断不能同为真，而且各自的日课正是另一家诊断出的病。

若管理学对，教育学在开一张制度上无法执行的方子：把协助的量调下去，评价系统在下一轮就会把它压回来，因为压回来对每一个当事人都是局部理性的。若化学对，管理学在换一个同样不起作用的旋钮：改指标、加延迟测验、拉长责任周期，改的仍然是时间与障碍，而在不可逆条件下这两个量根本不进入路径选择。若教育学对，化学没有提供任何新变量：所谓“给第二次机会”不过是“减少协助量”的一个换算，可逆性是协助曲线上的一个坐标，不是另一根轴。

三家必须有两家是错的，或者三家的病灶都放错了位置。本文主张是后者。

二、三家共有的那个前提，与推翻它的材料

教育学争的是产物在多少协助之下被交出，管理学争的是产物被怎样计量与奖励，化学被引来解释产物沿哪条路径生成。三者表面互不相干，却从同一处开始算账：**一个可验收的产物，属于交出它的那个人。**这一条从未被任何一家写成命题，因为它不像一个可以被质疑的主张，更像记账的起点。

推翻它的材料在化学自己手里，而且是一年级教科书里的常识。

溶液中的许多反应，若照总反应式书写，会把并未参与焦点转化的离子一并写进去。化学的处理是把这些在方程两边同时出现的组分划掉，只留下确实发生变化的部分，得到净离子方程；被划掉的那些叫作旁观离子。要紧的是这类组分的处境：它在体系中，浓度可测；它维持电中性，抽掉它溶液根本配不出来；某些情形下缺了它反应压根不能进行。它满足了在场、可测、被计入投料、甚至必要——**而这些没有一条使它进入净变化**。它出现在总反应式里，只是因为总反应式记的是加了什么、剩下什么，不是记谁变了。

把这一条带回课堂，共有前提立刻塌掉。学生在场，学生响应，学生被计分，缺了这个学生这一轮课堂甚至不成立——三家据以完成归属的全部依据，恰好逐条对应旁观离子所满足的那些条件。归属不由在场决定，由是否进入焦点转化决定。而课堂的账本记的正是总反应式：加了什么（学时、教材、活动、工具），产出了什么（答案、作业、分数），中间谁变了不记。

于是出现一个三家都难以单独命名的对象：任务产物已经生成，学生也在场、响应并被计分，但决定产物的那一步由别的主体承担，学生没有沿着目标坐标发生转化。它不是低成绩、不参与或错误答案；它经常表现为顺利、正确与高效。

三门学科在本文中各自占据一个不可互相替代的位置。教育学占据路径解释：执行了什么认知操作，知识便沿什么路径组织为可用能力。化学占据归属规则与占优条件：产物不识别路径，在场不构成归属，而占优路径由可逆性与时间共同决定。管理学占据环境的反向选择：哪种产物被及时奖励，哪条路径就更容易获得时间、技术与合法性。删掉任何一家，剩下两家都无法说明整个自我维持的循环。

必须写明本文的证据等级，以免把跨域的形式相似当成已被证实的因果定律。既有研究能够支持的是局部关系：主动生成、提取练习、延迟测试、指标反应性与工具护栏的效应。化学能够提供的是关于归属与占优的逻辑约束，而不是关于神经层面学习的直接实验证据。"活化障碍"不是一个待测的脑内能量，"中间体"也不是某种特定的神经表征；它们的作用是迫使研究者提出对应的经验问题——哪一步决定路径能否独立运行，哪一项短暂操作没有进入最终答案，哪一种支持改变了障碍却更换了承担者。只有当这些问题被操作化为日志、草稿、口头解释与迁移任务，跨学科结构才具有研究价值。旁观轮次的完整动力模型是本文提出的理论假设，尚需跨学科、多课堂的实证检验。

学科来源	单因主张	能解释什么	自身暴露的缺口
教育学	协助的量与时机决定学习结果	为何生成、提取与互动通常优于被动接收	同等协助量落在不同步骤上，后果不可比，故它不是一条曲线
化学	产物不识别路径；占优路由可逆性与时间决定	为何相同产物不等于相同机理，为何在场不构成归属	体系层的约束不说明课堂里的哪一步在分岔
管理学	可见结果反过来重塑评价与资源环境	为何课堂偏好快、稳、可计量的产物	指标扭曲不能单独说明认知工作在何处被替换

三、核心概念：旁观轮次

本文把学习限定为：在原有支持部分或全部撤去之后，学习者仍能提取、重组并迁移相关知识，且这种能力来自学习者状态的相对持久改变。该定义不排斥合作学习、工具使用与分布式认知，但要求研究者在声称某个学生学会了时，能够区分系统能力与个体能力。

课堂产物是课堂中可被验收的结果，包括正确答案、完成的实验、规范笔记、流畅发言、完整报告、作业提交、考试得分以及由模型辅助生成的文本。产物具有教育价值，但产物本身不携带其生成路径的标签。一个高质量产物可能来自学习者的独立重构，也可能来自复制、提示链、同伴接管、教师逐步领做或模型直接生成。

关键认知中间步骤指从任务输入到产物之间，对目标能力形成具有结构作用的操作。不同学科有不同清单，通常包括：问题表征、相关知识提取、变量与关系识别、策略生成、关键中间推断、错误检测、证据评价、理由表达以及向新情境的映射。它们像化学中不出现在总反应式里的中间体；若不记录这些步骤，课堂只能看到产物，无法判断学习者是否经过了目标路径。

据此，本文将**旁观轮次**定义为同时满足以下条件的课堂微事件：第一，课堂系统生成了可接受产物；第二，学习者在场并参与了表面活动；第三，至少一个决定目标能力形成的分岔步由教师、同伴、教材模板、软件或模型提供；第四，系统把产物或完成状态记入学习者账户；第五，在支持撤去、时间延迟或情境变化后，学习者不能独立再生该步骤。第五条是确认条件，前四条构成过程识别；后文所有假设检验的是构成要素之间的关系，不是用旁观轮次去解释它自己的定义项。

"旁观"不是给学生贴人格标签，也不等于学生没有注意。学生可能投入、认真甚至情绪高涨，却仍然旁观了最承重的那一步——这正是旁观离子的处境：它不是缺席者，它是在场而未进入净变化的那个组分。分析对象是某个操作位置和某一轮任务，不是某类人。一个学生在同一节课中可能先旁观问题表征，后独立完成错误检测；一位教师也可能在一个环节替代学生，在另一个环节成功催化学生。

分岔步承担者	即时产物不合格	即时产物合格
	学习者承担	学习者转化
学习者承担	生成性未完成：错误暴露了当前中间状态，可成为生产性失败	正确产物与独立路径同时出现
外部系统承担	替代失败：代做仍未形成合格产物，问题通常容易被发现	旁观轮次：形式成功、错误信号消失、制度最容易误判

右下角是本文的靶格。它在短期指标上表现为改善，失效不产生即时信号，而且越被标准化越容易自我加固。课堂研究若只比较答对与答错，会把最需要诊断的一格与学习者转化合并。左下角是关键对照：代做而未成功的情形反而安全，因为它留下了信号。这一格的存在同时排除了"帮助有害论"——问题不是帮助的存在，是帮助落在哪一步、以及帮助之后账怎么记。

四、四种亚型：按被外置的那一步分类

旁观轮次并非单一现象。按照被外置的承重操作，可区分表征旁观、策略旁观、诊断旁观与论证旁观。四种亚型可以同时出现，但区分它们有助于说明同样的正确答案为何对应不同的后续缺陷，也能让支持撤除测试对准具体位置。

表征旁观发生在学习者没有亲自确定"这是什么问题"时。教师替学生翻译题意，教材预先标出变量，案例课直接给出问题清单，模型把复杂情境重写为标准题型。学生随后可能顺利计算或讨论，却没有形成从陌生表面中抽取关系结构的能力。典型后果是：原题会做，换一个叙事外壳便不知道从何开始。

策略旁观发生在问题已经被表征，但起点、路径与取舍由外部给定。逐步范例、算法菜单、模板与模型规划都可能承担这一位置。学生执行每一步时并非完全被动，却只在一条已经开通的轨道上运行。后果不是知识点遗忘，而是面对多路径或信息不完整的任务时无法选择，也不能说明为何这一条路优于另一条。

诊断旁观发生在错误检测与修订理由由外部完成。自动批改直接标红，教师立刻指出错因，同伴把错误步骤替换为正确步骤。学生看见修改并能复述解释，却没有经历"不一致从哪里现出、应检查哪一条约束"的搜索。典型后果是同一错误换一种形式再次出现，学习者仍依赖外部报警。

论证旁观发生在结论已经得到，但证据链、反例处理与适用边界由教师、范文或模型补齐。学习者提交的文本可能逻辑严密，课堂也可能展开热烈陈述，但只要理由结构可直接从模板搬运，产物质量就不能证明论证能力。后果是面对新证据时只能重复立场，不能重组理由或承认边界。

四种亚型的严重性取决于课程目标。若课程只要求熟练执行一项标准程序，教师提供问题表征未必构成失败；若目标是建模，表征恰是分岔所在。因而旁观轮次不是按照"帮助多少"判断，而是按照"外部系统是否承担了目标能力的分岔操作"判断。任何编码都必须先公开目标能力与任务分析，否则归属率会沦为任意评分。

混合型课堂还可能出现交替归属：教师完成表征，学生生成策略；系统完成错误检测，学生重写论证。此时不能把整节课判为旁观或非旁观，而应按操作位置记录。轮次化分析的价值正在于，它允许同一学习者既是一个步骤的承担者，也是另一个步骤的旁观者，从而避免把复杂课堂压成教学法标签。

五、分岔步：一次必须做的术语修正

把化学引入学习研究的既有做法，几乎都借用"限速步"这个词：找出最难的那一步，那一步决定整体进程，因此要把它留给学习者。这个借用是错的，而且错得可以被精确指出。

限速步管的是**快慢**。在一条既定路径上，最慢的那一步决定整体速率；改变它，产物出现得更快或更慢，但产物是谁、经由何处，不变。本文处理的却不是快慢问题：同一个正确答案，可以由学习者独立重构而来，也可以由教师逐步领做而来。两条路径、一个产物。这在化学里有专名，属于**选择性**：当多条路径通向不同结局时，决定最终得到哪一个的，是各条路径分开之处的那一步，即选择性决定步；它与限速步常常不是同一步，教科书对二者的区分是明确的。

因此本文把承重的那一步称为**分岔步**：两条路径——学习者承担的路径与外部承担的路径——在此分开的那一步。修正之后，判据整个换了一个：

- 旧判据：把最难的一步留给学生。
- 新判据：把**分岔**的一步留给学生。

两者常常不重合，而且不重合的方向可以预测。**难而不分岔的步骤，代做无害**。长串数值计算、单位换算、格式规范、文献格式整理，这些步骤耗时、易错、令人生畏，却不改变学习者是否形成目标能力：无论谁做，后续的关系结构都一样。**不难却分岔的步骤，代做致命**。把一段陌生叙事读成"这是一个关于两个量同时变化的问题"往往只需十几秒，学生做与教师做在课堂上看不出差别，而它恰恰是建模能力唯一的入口。

这一条直接推翻教育界的一项通行做法：以任务难度为依据分配帮助，把简单的交给学生练手、把困难的由教师带着走。按本文的修正，这套做法系统性地保护了错误的步骤——它把不分岔的难步交给了学生（无益），把分岔的易步交给了教师（有害），并且因为学生确实在费力、教师确实在讲难点，双方都获得了正在认真教学的证据。

分岔步的识别不依赖窥测学习者内部，而依赖任务分析：对目标能力，写出从输入到产物的操作序列，逐一追问"若这一步由外部完成、其余不变，撤除支持后的表现会不会不同"。回答为"会"的，就是分岔步。这个追问可以被预注册，也可以被不同编码者独立作答并计算一致性，因而不是一个只能靠直觉的判断。

需要一并说明的是，本文借用化学的三条约束都是**形式约束**，不是关于人脑的经验主张：产物不识别路径；在场不构成归属；多路径情形下决定结局的是分岔处而非最慢处。三条各自都可以被课堂数据裁决，也各自都可能被裁决为不适用——若某类任务的操作序列上根本不存在可识别的分岔位置，本文对该类任务无话可说。

六、形式模型：为什么在不可逆条件下延迟与障碍不进入选择

化学关于占优产物的判断是这样的：同一批反应物若有两条通路，一条障碍低、生成快，产物较不稳定；另一条障碍高、生成慢，产物更稳定。在低温、短时或不可逆的条件下，占优的是**生成快**的那个，称动力学控制；在给足时间且各步可逆的条件下，体系可以退回来重走，占优的是**更稳定**的那个，称热力学控制。要害在于：从动力学控制切换到热力学控制的开关，是**可逆性**，不是**障碍高度**。若第一步不可逆，无论等多久、把另一条路的障碍降得多低，先形成的那个产物都不会自己转过去。

把这一条形式约束搬到课堂，需要先确认对应关系是否成立，而不是宣布它成立。

对应关系：外部承担的路径成本低、当轮回报快，对应动力学产物；学习者承担的路径成本高、回报延迟，对应热力学产物。**课堂的首次选择是否不可逆，是一个可以直接查验的经验问题**——而答案是压倒性的：提交即计分；答案一经给出便不可能未被看见；教师的讲解一旦落地，学生无法退回讲解之前的状态

重走一遍；模型给出的完整解法进入工作记忆之后，"我自己会不会"这个问题已经无法再问。课堂几乎处处满足不可逆条件。

据此可以把选择写成一个足够简单、但确实能推出东西的形式。设某一轮任务的分岔步归属为 $a \in \{\text{学习者}, \text{外部}\}$ 。设学习者路径的当轮成本为 c_L （时间加失败风险），外部路径当轮成本为 c_E ，由外部路径的存在理由可知恒有 $c_E < c_L$ 。设当轮可见回报为 P （正确率、完成、秩序、满意度），延迟 h 之后的迁移回报为 T ，学习者与教师对延迟回报的折现因子为 κ 。再设**可逆度** $r \in [0,1]$ ，表示首次选择在本轮之内可被撤销并重走的程度。

- 当 $r = 0$ ：一次选择即锁定，本轮的比较只在 $P - c$ 之间进行。由 $c_E < c_L$ 且两条路径的 P 相同（这正是靶格的定义：两条路径给出同一个可验收产物），外部路径**无条件占优**。此时 T 与 κ 不出现在比较式里——**延迟长度与迁移回报的大小根本不进入选择**。
- 当 $r = 1$ ：学习者可以先走便宜路径，再撤回重走。此时被比较的不再是两个互斥选项，而是"走便宜路径并放弃重走"与"走便宜路径后重走"两个序列， κT 第一次进入比较式。
- 因此存在一个门槛 r^* ，使得**只有当 $r > r^*$ 时**，任何以延迟迁移为目标的干预才可能改变均衡。当 $r \leq r^*$ ，加长测验延迟、提高任务障碍、加大迁移在评价中的权重，都只是在改变一个不进入选择的量，其结果是成本上升而均衡不动。

这条推导给出三件原文所无的东西。

第一，它给出一个**次序**：先改可逆性，再改延迟与障碍；颠倒则前者无效。这不是经验建议，是从上式直接读出的。

第二，它给出**对既有失败的解释**。教育改革的典型命运是局部有效、随后被日常运行吸收：增加活动、增加形成性评价、延后测验、撤除支架——每一项都在小样本上有效，放大之后被吸收。通行解释是执行不力、教师培训不足、评价未跟上。按上式，这些解释都不必要：上述四项改的全是 h 、障碍与 T 的权重，而在 $r \approx 0$ 的课堂里这三个量不进入选择，所以它们**在原理上就不该持久有效**，观察到的局部效应更可能来自实验期间不可逆性被临时降低（研究者在场、允许重做、不计分）这一未被记录的副作用。这个解释本身可被检验：若在保持不可逆的前提下严格复制那些干预，效应应当消失。

第三，它给出一条**可证伪的门槛预测**：可逆度与延迟类干预之间应存在交互项，而不只是各自的主效应。若在 r 高低两组中，延迟测验的效应量相同，本条为伪。

必须指出上式的局限。它是一个选择模型，不是学习模型：它说明为什么某条路径被选中，不说明被选中之后学习者内部发生了什么。它假定两条路径的当轮可见回报相同，这一假定只在靶格内成立，在代做失败的那一格并不成立。它把 r 处理为连续量，而课堂里的可逆性往往是若干离散装置的组合。这些都限制了它的适用范围，也各自指出了它可能被推翻的地方。

七、三个读数，其中两个当场可算

理论若只能靠昂贵的延迟测验才能被使用，它在多数课堂里就是不可用的。本文给出两个读数，其中第二个不需要窥测学习者内部，也不需要等待。

读数一：分岔步归属率。对本任务所需的关键操作 k ，记权重 w_k （由课程专家、任务分析与试测共同确定，不由教师单方决定）； o_k 在学习者未获得内容性答案或等价中间结论之前自行发起并完成该操作时记为 1，由外部提供时记为 0，部分共同生成记为 0 到 1 之间。归属率为加权比值。这一读数的弱点必须写明：问题表征等操作部分发生在头脑内部，日志只能给出代理证据，因而必须在部分样本上以出声思考或刺激回忆访谈校验编码是否误读了沉默中的独立生成。

读数二：催化比。这一个来自催化剂的一条性质——催化剂降低障碍并在循环结束时被完整再生，它不进入产物、不被消耗；化学计量试剂则被消耗进产物，因此下一轮必须重新投料。据此得到一条只看外部行为的判据：

> 一项帮助若在下一轮仍完整可用而无须加量，它在起催化作用；若它被消耗进了这一次的产物、以致下一轮必须重新供给同等或更多，它在起替代作用。

操作化为同一项帮助在连续轮次中的用量比值 H_{t+1} / H_t ：提示条数、提示展开层级、教师介入时长、模型调用次数、范例给出数量。催化的曲线应平或下降；替代的曲线单调上升。这个数当场可算、无须延迟测验、无须窥测内部、且几乎全部数据已经存在于现有平台日志中。它的代价是只能给出趋势而不能给出单轮判定，因此它的位置是筛选高风险轮次，再对被筛出的轮次做归属编码与延迟迁移确认。

读数三：可逆度。首次作答提交后可否自行推翻并重走；教师讲解是否覆盖学生原稿；模型答案默认展开还是默认折叠；本轮成绩取首次还是取末次。这四项各记 0 或 1，构成一个粗糙但可核对的四分量指标。它的作用不是精细测量，是让“这节课的可逆度是几”成为一个可以被问出口的问题——目前它在多数课堂里连问都问不出来。

三个读数都不是新的万能指标。它们的价值在于把不可见过程转成可抽样复核的证据，而不是取代迁移测试。若一套编码与延迟迁移没有关系，就应放弃或重做；若它一进入考核便促使教师制造形式上的学生自主，则必须降权。测量服务于识别，不拥有对课堂价值的最终裁决权。

操作类别	典型问题	外部系统常见代做方式	可观察证据
问题表征	这道题究竟在问什么，哪些变量相关	教师直接翻译题意，模型重述并标变量	学习者能否面对新表面自行构造同一关系图
知识提取	哪些原理可用 从哪里开始、为何选这条路	给公式清单、关键词或现成资料定位	无提示时是否能在延迟任务中提取

操作类别	典型问题	外部系统常见代做方式	可观察证据
策略生成		范例照搬、步骤菜单、模型规划	学习者能否提出两条路径并说明选择
中间推断	关键关系如何推出	逐步领做或填空式提示	能否独立写出未提供的中间步骤
错误检测	哪里不一致、为何错	系统直接标红并解释	能否发现新的同构错误
理由表达	答案凭什么成立	提供模板或整段论证	能否用不同语言重建证据链

八、三轮循环与四种信息不对称

旁观轮次不是单次不良教学技巧，而是一套会反过来选择自身条件的循环。为了回答课堂为何长期高投入而低迁移，必须追踪谁先动、另外两项何时跟上，以及这一轮结果如何改变下一轮。

第一轮由管理环境启动。学校、课程与教师面对进度、正确率、课堂秩序、考试成绩与满意度等可见要求。学习者独立生成通常速度慢、方差大、错误多，与可预测的课堂流程冲突。为减少不确定性，环境开始增加范例、提示、标准答案、讲解密度、模板、即时纠错与自动化工具。这一变化未必出于功利或控制欲，它常是教师在时间与责任约束下的合理响应。

第二轮由学习者当前状态与新环境共同选择路径。基础薄弱、先前知识零散的学习者遇到高提示环境时，最容易进入低成本路径：识别教师意图、匹配范例、复制步骤、等待系统纠错、向模型索取答案。该路径快速形成任务产物，却绕开了分岔步。学习者并非懒惰；在一个奖励即时正确、随时提供替代路径且首次选择不可撤回的环境中，选择低成本路径是上一节那条不等式的直接后果。

第三轮由路径与环境共同生成可见表现。因为分岔步已由外部承担，答案更快、更整齐、方差更小，课堂感觉更流畅。学生认为教师讲得清楚，教师认为本节完成顺利，管理者看到进度与成绩，平台记录高完成率。这些表现反过来选择第一轮：原有环境被验证，下一轮获得更多资源与合法性；相反，要求学生先生成、允许错误暴露的课堂可能在短期满意度与覆盖进度上受罚。

循环的关键是时间错位。即时表现的奖励在当场发生，迁移失败要到支持撤去、情境改变或数周后才出现；即使出现，也可能被归因于学生遗忘、动机不足或下位课程教学不力。于是环境不断接收"该路径有效"的快信号，很少接收"学习者没有改变"的慢信号。

轮次	先动变量	可观察变化	随后发生	下一轮的反向选择
第1轮：环境收紧	进度、正确率或秩序 压力上升	提示密度、标准化与代做 资源先增加	学习者可选路径改变	短期稳定被解释为 管理有效

轮次	先动变量	可观察变化	随后发生	下一轮的反向选择
第 2 轮：路 径换道	低基础状态遇到高支 持环境	自发提取与策略生成减少	外部承担率提高、催 化比上行	学习者对替代路径 形成依赖
第 3 轮：表 现确认	外部路径与即时评价 结合	正确率上升、方差下降、 满意度改善	延迟迁移滞后下降	高表现再次强化原 环境

该模型提出的是顺序性预测而非相关性预测：管理压力的上升应早于提示密度上升；提示密度上升应早于学习者自发中间步骤减少；自发步骤减少应早于无支持迁移下降。只要观察顺序系统性相反，本文动力模型就需要修正。

循环之所以稳定，还因为四种信息不对称。**时间不对称**：正确答案与任务完成在本轮可观察，迁移与保持要到下一轮甚至下学期；组织通常把快信号归给当前教学，把慢信号归给学生差异或下一位教师。**可见性不对称**：产物可以被自动记录，中间步骤分散在草稿、停顿、对话与内部提取中；平台最容易保留点击与答案，最难保留问题表征与策略选择。**风险不对称**：允许学生暴露错误让教师当场承担进度、秩序与评价风险，直接给出路径则把风险推迟到未来迁移；当前教师承担前一种风险，却未必承担后一种后果。**归属不对称**：成功产物通常归入学生与教师的共同成绩，迁移失败则容易归入学生个体缺陷；系统获得成功的信用，把未发生的转化外部化给学习者。

九、八条通道：旁观轮次如何被稳定生产

其一，产物归属错记。课堂记录把答案与提交者绑定，不记录答案中哪些分岔步由他人生成。学习管理系统知道谁点击提交，不知道谁完成问题表征；考试知道谁写下结论，不知道结论是否由前一刻的模板直接搬运。

其二，中间步骤从账本中消失。中间体不出现在总反应式中，课堂的认知中间步骤也常不进入成绩单。教师被要求记录完成率和分数，很少记录策略由谁提出、错误由谁识别。越是精细的标准答案，越可能把整个过程压缩成一个终点。

其三，低方差优先于状态改变。管理系统偏好可预测。学习者自主生成带来错误、等待、分歧与不可控提问；外部领做压缩方差。只要评价主要依据课堂秩序与进度，教师就被推动去生产一致产物。

其四，流畅感替代学习证据。讲解清晰、范例相似与提示及时会提高主观流畅感。学习者容易把"我能跟上"判断为"我能独立生成"。Deslauriers 等（2019）的随机研究表明，学生在主动课堂中实际学得更多却主观感觉学得更少——费力感与学习收益可以方向相反。

其五，评价时间窗过短。迁移需要延迟、支持撤除与情境变化来检验，而课堂评价常在教学刚结束、提示仍在工作记忆中时进行。测试越接近范例，越容易测到残留激活与识别。

其六，支架没有退出机制。支架本意是暂时支持，实践中却容易永久存在。每当学生困难，系统就增加一步提示；学生习惯提示后，撤除被视为不公平或低效。催化比在这条通道上单调上升，而这个数几乎从不被人看。

其七，活动化掩盖分岔步外置。课堂改革可能把讲授改造成讨论、抢答、项目或游戏，但若教师预先规定问题框架、关系与评价标准，学生仍只是在既定轨道上填充。行为更活跃，分岔步仍由外部掌握。

其八，生成式模型把替代路径工业化。模型能即时完成解释、规划、纠错、改写与论证，使旁观轮次从偶发变为可规模复制。Bastani 等（2025）的高中数学现场实验显示，开放式工具能显著提高有工具时的练习表现，但工具撤去后，使用无护栏版本的学生可能比从未使用者表现更差；加入以提示而非直接答案为中心的护栏后负面效应明显缓解。要紧的是护栏组的另一半结果：负效应被抵消，**长期增益并未出现**。这与第六节一致——护栏改的是协助形态与障碍高度，没有改可逆性。

这八条也解释了为什么常见改革会被重新吸收。增加课堂活动可能失败，因为活动改变了可见形式，却没有改变分岔步归属。增加脚手架可能失败，因为撤除对象不明确：实践中常按学生是否答对决定是否撤除，学生在高提示下答对，教师便误以为能力已经形成。增加形成性评价也可能失败，因为它仍在原范例、原提示与短时间窗中进行。把课堂数字化可能失败，因为平台最擅长记录终点：若系统保存每次点击却不保存首次表征、提示展开顺序与修订理由，数据规模只会让产物归属错记更精确。教师培训同样可能被吸收：培训要求增加提问、开展合作、使用模型，学校随后把这些做法变成频次指标，教师为达标安排更多表面活动，课堂却仍由教师完成归纳、诊断与结论。

十、假设与判决变量

本文的承重命题可以写成三重否定式：课堂中可迁移学习稀缺，不只是学生参与不足，不只是任务障碍过高，也不只是评价指标发生偏移；它是一种系统产出与个体转化的分离——课堂允许分岔步被外部承担，同时把系统产物记为学习者能力，而这一分离在首次选择不可逆时是路径选择的必然结果。

编号	假设	关键判决变量
H1 归属效应	控制先前知识、时间投入、任务难度与即时正确率后，分岔步归属率越低，延迟迁移越差	归属率→延迟迁移； 控制即时正确率
H2 超越活动分类	在同被归为主动或互动的课堂中，分岔步由外部承担的组，迁移低于由学习者承担的组	同活动类型内的步骤 承担者差异
H3 难度与分岔的分离	等量协助落在分岔步与落在非分岔步，迁移差异显著；而协助总量本身在控制落点后不再显著	协助落点 × 协助总量
H4 可逆性门槛	可逆度与延迟类干预存在交互：低可逆组中延迟测验与提高障碍的效应趋于零，高可逆组中显著	可逆度 × 延迟/障碍干预
H5 可逆性主效应	在时间投入、任务难度与协助总量三者匹配的条件下，仅把首次作答改为可撤回一次并记录撤回，延迟迁移提高	可逆度单独操纵
H6 催化比预测	同一项帮助的轮次用量比单调上升，预测该操作位置的归属率下降与后续迁移下降	催化比→归属率→迁移 测验延迟 × 表面变式

编号	假设	关键判决变量
H7 时间窗调节	即时测验与教学越接近，旁观轮次越难被发现；延迟越长、表面变化越大，即时表现与迁移的分离越明显	
H8 管理压力的中介链	进度、评价或秩序压力对延迟迁移的负向影响，主要由提示密度增加与自发中间步骤减少所中介	压力→提示密度→自发步骤→迁移
H9 轮次顺序	压力上升→外部代做增加→自发步骤下降→即时表现改善→延迟迁移下降的顺序应在时间序列中重复出现	时间先后与轮次反向选择

判决逻辑分三组。H1、H2、H3 检验归属是否具有超出课堂形式与协助总量的解释力；最关键的对照不是讲授对主动学习，也不是有模型对无模型，而是在即时产物、时间与任务难度相近时，只改变分岔步由谁完成。H4 与 H5 是本文最独有的两条，也是与所有既有框架都正交的两条：它们操纵的量在协助两难、效果余留与元认知惰性三家的变量表上都不存在。H6、H8、H9 检验管理循环，要求观察时间顺序而非一次横断相关。

九条假设不要求同一研究一次全部验证。理论可分层接受检验：第一层是归属对迁移的增量效度；第二层是可逆性门槛；第三层是管理循环与轮次顺序。第一层若失败，核心概念应被撤回；第二层若失败，本文相对既有框架的独立贡献消失，应退回为协助两难的一个子类；第三层失败则概念仍可保留为微观诊断，但不能再解释这一现象为何自我强化。

十一、为什么这不是效果余留、协助两难与元认知惰性的另一个名字

新概念若不能与最近邻给出判决性分离，就只是换名。本文的最近邻不是主动学习框架，也不是学习与表现的区分，而是下面三家——它们各自已经覆盖了本文承重命题的一大块，因此必须逐家给出分离线与可裁决的观察。

其一，效果余留。Salomon、Perkins 与 Globerson (1991) 区分"与技术共同产生的效果"和"技术留下的效果"：前者是人机搭档时被重新定义与提升的表现，后者是搭档结束、离开机器之后仍然存在的认知余留。他们进一步主张两者都取决于个体在这场搭档中的**用心投入**。这一区分与本文的双账本几乎同构，出现得比本文早三十余年，任何声称提出新概念的文本都必须先过这一关。

分离线在**病灶的位置**。那一支把病灶放在个体的心理状态：用心则有余留，不用心则无。本文把病灶放在**任务的分岔结构**：一个高度专注、全神贯注跟着教师推导的学生，若分岔步被教师走掉，仍然没有余留；而一个心不在焉、把题目读错、自己撞墙、自己改对的学生，反而有。判决性观察：以注视、时长与自评三路测量投入度并加以控制，只操纵分岔步归属；若余留无差异，本文败，用心投入就是那个变量；若差异仍在，则用心投入不是那个变量，它只是分岔步归属在个体身上的一个相关物。

其二，协助两难。Koedinger 与 Aleven (2007) 把协助的量与时机设为核心变量：给多导致浅层学习，给少导致挫败与时间浪费，最优点在中间。配套的求助滥用与钻系统空子研究已经用过程日志把"学生是否真的在做这一步"操作化了二十年。这一家离本文最近，也最容易把本文吸收为它的一个说法。

分离线在**变量的维数**。协助两难是一条曲线：一根轴（协助的量），两端各有代价，最优点是一个标量。本文主张这是**误置**——同样的协助量，落在分岔步与落在非分岔步，后果不可比，因此它们不是同一条曲线上的两点，而是两条不同的曲线。判决性观察：等量协助分两组投放，一组全部落在分岔步，一组全部落在非分岔步，总量、时机、呈现形式完全匹配；若两组迁移无差异，协助两难的一维刻画足够，本文多余；若有差异且方向与任务分析的预测一致，则“量”这根轴是一个混杂了落点的代理变量。第三节所说的“难而不分岔”与“不难却分岔”两类步骤，为这个实验提供了现成的材料。

其三，元认知惰性。Fan 等（2024）在随机实验中发现，使用生成式模型的一组作文提分最多，而知识增益与迁移与其他组无显著差异，并把这一模式命名为元认知惰性；后续研究已把它做成测量个体倾向的量表。这一家的实证结果几乎就是本文 H1 与 H7 的现成证据，因此它同样构成吸收本文的候选。

分离线在**病灶是人还是任务**。“惰”是一个关于人的词，它把病灶放在学习者的动机与自我调节倾向，于是处方是训练自我调节，工具是测量个体倾向的量表。本文把病灶放在任务的分岔结构与制度的记账方式，因此预测同一个学习者在同一节课内可以在一步上“惰”、在另一步上不“惰”——这不是倾向能解释的。判决性观察：在单一课时内，跨不同分岔位置测同一学习者的归属率；若高度一致（可由单一倾向解释），本文的轮次化分析多余；若随分岔位置系统变化且与预先做出的任务分析吻合，则倾向解释不足。倾向量表在这种情形下测到的不是人格，是那个人所处的任务序列。

收编式结论。三家把病灶各放在一个不同的位置：投入（心理）、协助的量（教学设计）、倾向（人格）。三处都不是那个位置。那个位置是**分岔步的归属，以及它在账上的去向**。接受这一条，三家各自成为特例：效果余留是分岔步归属为学习者时的余留，用心投入是归属的一个心理相关物；协助两难的那根轴是落点在总量上的投影，最优点之所以随情境漂移，是因为不同情境下分岔步在操作序列中的位置不同；元认知惰性是分岔步长期归外之后在个体身上累积出的读数。

三家都不能推出的，是第六节那条：在不可逆条件下，延迟与障碍不进入路径选择。它们推不出来，不是因为疏忽，而是因为三家的变量表上都没有可逆性这一项——协助两难有量与时机没有可撤回性，效果余留有投入没有本轮结构，元认知惰性有倾向没有制度装置。这也是本文可以被独立检验的地方：H5 操纵的量在三家的框架里不存在，因此它的结果无法被三家事后吸收。

十二、被强制的不一致

上一节论证的是本文不可被三家还原。本节论证一件更强的事：有一门正在运行的学科，在**不承认本文变量的前提下无法自治**——不是本文批评它，是它自己的两条既有结论不能同时为真。

学习分析与教育数据挖掘这一行同时持有两条命题。

甲：日志可见的参与度是学习的可用代理。这是仪表盘、预警模型、投入度指标、参与度排名与风险学生识别的全部合法性来源。工程实现上，甲不只要求相关，还要求**单调**：阈值、告警线、排序与推荐都预设“更多的可见活动对应更多的学习”，至少在方向上如此。

乙：日志可见的活跃可以与学习负相关，而且恰恰出现在活跃度最高的那批学生身上。这一条不是外人的指控，是这一行自己确立的结果：求助滥用刻画的是有能力独立完成却大量调用帮助的学生，钻系统空子刻画的是通过快速连点提示直取答案的学生。两类行为在日志上都表现为高频交互、高完成率、短响应时间——按甲的读法，这是模范学生。

甲与乙不是一对需要调和的张力，它们在同一套数据上互相拆台：乙成立意味着代理关系不单调，而甲的全部工程实现预设单调。

该行现有的处置是把乙做成一个待检测的异常类：先识别钻空子行为，再把这些样本剔除或降权，剩下的部分仍按甲处理。这条出路是循环的。要判定某次点击"不是真的在学"，需要一个判据来区分同样的一次点击何时算作学习、何时不算——而这个判据只能是"这一步本该由谁走"。落在非分岔步上的一次提示调用是合理的省时；落在分岔步上的同一次调用是替代。两者在日志上不可分，除非先有任务分析给出分岔位置。**换言之：这一行要让甲成立，必须先承认本文的变量；而一旦承认，甲就不再能以现在的形式成立——参与度不再是学习的代理，它只是一个必须先按操作位置拆开、拆开之后两半符号相反的复合量。**

这不是"应当补充一个调节变量"的建议。它的形式是：任何一个既接受乙、又把参与度当作单调代理来配置阈值与告警的系统，其内部两条命题不能同真；而放弃乙意味着放弃这一行二十年里最被引用的一批实证结果。

同一动作可以对第二个行当再走一遍。生成式模型教育产品的采购与验收普遍以即时正确率、响应速度与使用时长为准。而这一行业在市场材料中反复引用的两项研究——工具撤除后无护栏组低于从未使用组、以及模型组作文提分最大而知识增益与迁移无差异——恰好证明这三个验收数与撤除后的能力可以反号。于是同样的结构再现：验收标准与自己援引的证据不能同真。差别在于，学习分析那一处的不一致出现在两条学术结论之间，这一处的不一致出现在验收标准与举证材料之间，后者更容易被当作营销问题而非理论问题处理，因而也更持久。

需要说明本节主张的边界。本文不主张学习分析这一行的从业者没有意识到这个问题；相反，钻空子检测这一整支研究正是这种意识的产物。本文主张的是，在不引入操作位置这一变量的条件下，这种意识**在形式上无法被兑现**——检测本身需要那个变量。这一主张可以被推翻：若有人给出一个不依赖"这一步该由谁走"的判据，能够稳定区分同一次点击是省时还是替代，本节即告失败。

十三、外部划界：另外八家

除上述三家之外，还有八家覆盖本文的一部分。以下比较使用同一案例作对照：课堂产物正确，分岔步由外部承担，撤除支持后迁移失败。

最近邻	它说到哪一步	本文的分离线	判决性观察
主动学习与参与层级框架	按可见参与与生成产物区分层级	本文追问分岔步的承担者；行为活跃仍可能旁观分岔步	若同为互动课堂，外部承担组与学习者承担组迁移无差异，则本文相对该框架无增量

最近邻	它说到哪一步	本文的分离线	判决性观察
学习与表现的区别	练习中的表现不等于持久学习	本文给出分离发生的微观路径、可逆性开关与管理闭环	若归属率不能在控制即时表现后预测迁移，机制主张失败
认知负荷理论	通过管理三类负荷优化学习	负荷降低可能是催化也可能是替代，关键在分岔步是否被保留	若直接给步骤与先生成后提示在相同负荷下迁移相同，分离线受损
合意困难	增加恰当难度可提高长期保持	本文要的不是难度而是分岔与可撤回；难而不分岔的步骤加难无用	若在分岔步不变的前提下单纯提高难度即可提高迁移，本文的落点主张受损
生产性失败	先尝试后讲授可准备知识结构	本文不要求先失败，关注任意序列中分岔步归属与首次选择的可逆性	若没有失败但保留分岔步仍无法提高迁移，理论适用范围缩小
脚手架与渐隐	支持学习并随能力提升逐步撤除	本文提供渐隐对象与渐隐读数：催化比而非笼统减少帮助	若渐隐非分岔步骤即可获得同等迁移，分岔步概念需修正
认知卸载与分布式认知	认知可分布于人、工具与环境，卸载可提高效率	本文承认系统能力，只反对用系统能力证明个体学习	若研究声称的是团队或人机系统能力而非个体迁移，本文不构成反例
目标置换与制度脱耦	指标成为目标后行为围绕指标重组；正式结构可与核心实践松散耦合	本文无需作弊或虚假指标，产物确实正确，只是由错误承担者完成；并把脱耦推进到任务微过程	若仅在高利害指标下出现、撤去指标即消失，本文可被目标置换吸收

八家之外还有一个更大的共同前提值得点出：以上十一家——包括本文最近的三家——在讨论帮助时，全部把帮助处理为一个**在本轮内已经发生且不可撤销**的事件。协助两难问的是给多少、何时给，合意困难问的是设多难，脚手架问的是何时撤，认知卸载问的是卸给谁。没有一家问：**这次帮助能不能被取消，让学习者退回到接受帮助之前的状态重走一遍**。这个问题之所以从未被问，恰恰是因为在纸笔课堂里它接近于不可能——而在数字环境里它已经完全可实现，且几乎没有人实现它。

十四、站内划界：与本栏两篇先行研究的分工

本栏此前已有两篇触到本文的两根轴，必须分别给出分工线，否则本文读来像自我重排。这两篇比上一节那八家离得更近，其中一篇甚至已经用过“可撤回”这个词。

与《收得回的那一遍》（本栏之三十六）。那一篇从舞蹈排练中的“标记”出发，指出一次练习能修改什么，不由它有多难、多像正式场合决定，而由动作发起之后还能不能被收回决定；它命名撤回余量，并推出执行遮蔽——在不可撤回的区间里练结构，结构缺陷不会以结构缺陷的样子出现，而以执行者能力不足的样子出现。

分离线在**可撤回性调节的是什么**。那一篇的场景里**没有第二个承担者**：舞者做标记时，动作仍然是舞者自己做的，可撤回改变的是这一遍能读出什么，是**可读性**。本文的场景里始终有第二个承担者站在旁边，可撤回改变的是这一步归谁，是**可选择性**。判决性观察分两侧：取一个不存在外部代做路径的任务（即兴、体感、个人技艺），那一篇的预测仍然成立而本文的预测为空，因为没有分岔就没有路径可选；取一个有代做路径但不区分结构与执行的任务（一道只有唯一解法的应用题），本文的预测成立而那一篇的预测为空。两条预测各自有对方为空的地盘，因此不是同一条。

同时必须记下一处**继承**，不能算作本文的发现。那一篇的执行遮蔽推论直接适用于本文第八节的归属不对称：在不可撤回的课堂里，分岔步被代做所产生的缺陷同样不会以"这一步没有被走过"的样子出现，它会以"这个学生基础差"的样子出现，并被随后的加练固化。本文的归属不对称因此是执行遮蔽在归属这一维度上的一次复现，应当记在那一篇账上。

与《那一段谁走的》（本栏之五十一）。那一篇命名哑段——当期不产出任何可被验收之物、而日后一切验收能力由它生成的那段时间，读数是哑段占比与代走率，并明写哑段不可代做，一经代做即不再是它。

分离线在**定义所依据的量**。哑段按**产出的有无**定义，它是那段不产出的时间；本文的分岔步按**路径的分开**定义，它可以完全落在一段有产出的时间之内——把一段陌生叙事读成一个关系结构只需十几秒，而它立刻产生一个可见的表征。因此哑段在本文的坐标里是一个特例：产出为零的分岔步；而本文的分岔步在那一篇的坐标里没有位置，因为它有产出。

更硬的一条来自那一篇自己的定义：哑段一经代做即不再是它。这使哑段在定义上不可能出现在本文的靶格中——靶格要求代做之后产物依然合格。**本文的靶格恰恰是那一篇的定义所排除的那一格**。两篇因而在同一张图上占互补的两块。判决性观察：一个把哑段时间完整排进课表、而在有产出的时段里分岔步仍被全部代做的班级，那一篇预测验收能力保住，本文预测保不住；这个对照两篇都不能自己判，必须实测。

处方的先后关系本文不主张。那一篇的四步以"给哑段一个时间字段"起手，本文以"把首次选择做成可撤回一次"起手；两者不冲突且可能互补，但何者在前尚无依据，本文不假定自己在前。

十五、研究设计与操作化

理论若只能解释已知现象，就仍可能是新名字。旁观轮次必须通过过程数据与支持撤除后的表现被独立识别。建议采用"多地点集群随机试验+微观过程编码+延迟迁移"的组合设计。每个条件至少应包含六个以上同质课堂单元，更稳妥的设计是四所学校、每校八至十二个班级，在班级层随机分配；样本量最终由预注册的功效分析决定。

主实验设五类条件，其中第五类是本文独有的、也是最要紧的一类。

- A 组：常规讲解与练习。
- B 组：先生成后讲解，学生必须先提交问题表征、策略与一项中间推断。
- C 组：高强度逐步提示，确保即时完成。
- D 组：带护栏的模型辅助，只能在学生尝试后提供分层提示。

- **E 组：可撤回条件。**协助总量、时机与形态与 C 组完全一致，唯一差别是首次作答在本轮内可自行推翻并重走，成绩取末次，撤回本身被记录且不计惩罚。

C 与 E 的对照是本文的核心检验，因为它把协助的量、时机与形态全部固定，只留下可逆性一个变量。若两组迁移无差异，第六节的推导为伪，本文相对协助两难的独立贡献消失。B 与 E 的对照回答另一个问题：可撤回是否只是"先生成"的一种实现方式；若两组等效，则可逆性可被归约为生成顺序，本文应据此收窄。

数据分五层。第一层是基线知识与动机；第二层是课堂视频、屏幕日志、聊天记录与草稿，用于编码分岔步承担者并计算催化比；第三层是即时产物；第四层是 48 小时与四周后的无支持迁移，题目表面改变而关系结构保持；第五层是管理环境，包括进度要求、评价权重、教师绩效指标、班级规模与可用支持。部分样本加入出声思考或刺激回忆访谈，用于检验日志编码是否误读了内部生成。

多层模型以学生嵌套班级、班级嵌套学校。核心检验不是简单比较组均值，而是在控制即时正确率后检验归属率是否预测延迟迁移，并用中介模型检验管理压力是否通过提示密度与归属发挥作用。时间序列可用交叉滞后或序列分析检验谁先动。

低成本版本可在现有平台完成，且不需要新增设备。随机抽取每周三个任务，要求系统记录提示展开顺序、答案复制比例、首次独立尝试与错误修订；每两周安排三个不计分、无提示的同构迁移题。**其中可逆性那一项的实施成本几乎为零：**把作业系统的"提交即锁定"改为"本轮内可推翻重做、成绩取末次、撤回入账"，是一次配置修改。多数学校已经拥有提交时间、点击、修改与答题日志，新增成本主要是任务分析与抽样编码。

对一次具体课堂轮次，可以直接问一句不含判断词的话：撤去答案、步骤提示、范例、同伴接管与模型输出后，学习者在 48 小时后面面对表面不同而关系同构的任务，是否独立生成问题表征、分岔推断、错误检测与理由说明中的至少三项？这句话可以转化为任务、日志与编码规则。

场景	即时产物	分岔步可能由谁承担	48 小时撤除支持后的判定
化学概念课	学生能写出平衡移动方向	教师给出粒子层解释与判据	能否在新反应体系中独立画出粒子变化并解释
数学解题	学生得到正确方程与数值	模型完成建模与步骤规划	能否在情境改写后自行建模并指出关键约束
语言写作	作文结构完整、语言流畅	模板或模型提供论点与段落骨架	能否针对新材料自行生成论证结构
管理案例课	小组给出标准战略方案	教师在追问中完成因果归纳	能否对新行业案例提出可反驳的因果链
实验课	实验按时完成、数据规范	讲义规定全部步骤与异常处理	能否面对一个参数变化自行设计控制与诊断

五个场景的答案不应被预设为相同。某节讲授课可能在数学建模上让学生独立生成，而某个项目课可能把问题定义与评价标准全部外置。至少两个场景必须依赖真实日志或迁移测试才能判断，否则判据只是把理论重复一遍。

十六、干预：交还分岔步，并把首次选择做成可撤回

本文不主张教师退出，也不把独立摸索绝对化。化学的约束提醒我们，催化可以降低无关障碍；教育研究也表明，对初学者完全不给指导可能增加无效负荷。改革的核心是区分“帮助学习者跨越某一步”与“替学习者完成分岔的那一步”。前者改变条件，使学习者能够反应；后者更换承担者，使课堂产物在学习者未转化时出现。

第一项原则是把可逆性排在最前。这是本文与既有处方最大的分歧，也是唯一由第六节的推导直接得出、而非由经验建议得出的一条。具体装置有四件，都不需要新技术，只需要改配置：首次作答提交后在本轮内可自行推翻并重走，成绩取末次；撤回本身被记录，不计惩罚，但**只允许一次**；教师的讲解不覆盖学生原稿，两份并列保存；模型答案默认折叠，展开即记为一次不可逆事件，并触发一道等价重入题。第四件是四件里唯一有成本的，其余三件是配置项。

必须写明的是，可逆性不是越高越好。化学在这里给出一条反向约束：工业上大量反应刻意运行在动力学控制之下，因为要的正是那个生成更快的产物；热力学产物不总是想要的那个。翻回课堂，无限可撤回会取消一切承诺，而承诺是判断得以被检验的前提——一个可以无限次改到对为止的作答，不再包含任何关于作答者当前状态的信息。因此处方收窄为**“可撤回一次，且撤回入账”**，不是“随时可改”。这一收窄不是折中，它由本文自己的目的推出：要保留的是学习者当前状态的证据，而证据需要一次不被覆盖的承诺。

第二项原则是生成先于揭示。在给出完整解释、公式或模型答案前，要求学习者提交最小生成物：一句问题表征、一个关系图、一条策略假设或一个可检验中间结论。生成物可以错误，但必须留下学习者当前状态的证据。教师随后针对生成物提供解释，讲授从替代答案变为与实际中间状态对接。

第三项原则是提示阶梯而非答案阶梯。提示从元认知问题开始（你缺少哪一个条件、哪两个量必须同时变化），再到方向提示、局部关系与最终步骤，每次提示记录展开层级。这一原则的验收不看形式，看催化比：若同一类提示的轮次用量单调上升，阶梯已经退化为答案的分期付款。

第四项原则是让错误成为可观察中间体。课堂不要只记录最终改正，而要保存首次表征、错误路径、修订理由与触发修订的信息。错误记录不是惩罚材料，而是判断谁完成了错误检测的证据。若系统直接标红并给出解释，学生只做确认，错误虽然消失，诊断能力却可能没有形成。

第五项原则是支持撤除测试。每个学习单元至少配置一个不计分或低风险的延迟迁移任务，撤去范例、提示与模型。撤除不是为了抓住学生，而是给管理系统提供慢信号。

第六项原则是双账本。课堂同时记录系统产物与学习者转化。系统产物可以包括团队报告、模型辅助文本与协作成果；学习者转化则记录归属率、催化比、可逆度与延迟迁移。两本账都重要，但不能互相冒名。团队完成一个复杂项目可以是优秀的系统能力，却不能自动证明每个成员都获得了同等个体能力。这一区

分也重新解释了"会做"与"会借助工具做": 后者不是低级能力, 而是另一种能力; 错误发生在资格、成绩或课程叙述把两者混为一谈。允许使用模型的课程完全可以评价提示设计、结果核验与人机协同, 但不能仅凭模型辅助成品认证独立论证能力。

第七项原则是管理指标降权而非再造单一指标。若把归属率变成教师排名或学生惩罚依据, 系统会很快制造表面上的自主生成: 故意延迟提示、编造草稿、让学生重复已知步骤。新读数只能用于诊断位置, 不用于给人贴标签; 不得为了提高数字而在安全关键课堂制造无谓失败; 若读数造成不利安排, 必须公开编码规则并提供复核通道。

管理层可做四项改造: 把"覆盖了多少内容"与"学生在无支持条件下能再生多少分岔步"分开报告, 禁止用前者替代后者; 教师评价抽样查看延迟迁移与过程证据, 不要求每节课都生成大量文书; 为生成、等待、错误暴露与修订留出课时缓冲——若进度表没有缓冲, 任何学习者承担机制最终都会被代做吞噬; 采购与课程平台验收增加"答案前尝试""提示层级""首次作答可撤回""支持撤除"与"迁移追踪"五项, 不以即时正确率单独判断教学价值。

教师专业发展也需要重新定位。备课时先回答四个问题: 本单元要形成的能力是什么; 从输入到产物必须经过哪些操作; **哪一步是分岔的** (不是哪一步最难); 撤除支持后用什么变式验证。课堂观察随之改变: 观察者不再统计教师讲了多久、学生说了多少次, 而是抽取若干轮次, 记录问题表征、策略、错误检测与理由生成的发起者。一次高质量观察可以只分析三个问题, 却比一张覆盖几十项行为频次的量表更接近机理。学校可建设学科化的"分岔位置库", 为典型任务标注可能的分岔步、常见替代路径与撤除测试; 该库不是标准脚本, 必须允许教师根据学生状态修订, 并用后续迁移数据修正。

生成式模型教学系统的最低设计规范可写成七条: 在提供内容性帮助前保存学习者的首次问题表征与尝试; 提示分层, 优先采用指向缺口的问题而非交付等价中间结论; 标记哪些操作由系统完成, 使教师能查看外部操作率; 周期性进入撤除模式, 只要求学习者面对表面变化的任务重建路径; 区分"长期可用工具能力"与"课程要求的独立能力"并分别呈现; **答案默认折叠, 展开计为一次不可逆事件并触发等价重入题**; 任何自动化归属判断只能用于任务诊断, 涉及分流、成绩或资源的不利决定必须有人审查、公开规则并允许学习者提交反证。采购方至少应要求供应商报告无支持延迟迁移、提示层级分布、首次独立尝试比例、答案前生成率、催化比曲线, 以及不同基础学生的操作归属差异。

十七、适用边界与三门学科的反向约束

第一, 本文不是反讲授。清晰讲授可以为学习者提供概念工具, 尤其在先前知识不足时。关键是讲授之后, 学习者是否承担了分岔步。若教师提供信息而学生完成重组, 不构成旁观轮次; 若教师连重组也完成, 才进入高风险区。

第二, 本文不是"困难越多越好"。化学的约束迫使本文放弃这一更强主张, 而且放弃得比通常更彻底: 不仅无关负荷、模糊指令与羞辱性失败只消耗认知资源, 而且在**非分岔步上增加难度是纯粹的浪费**——它提高了成本却不改变结局。合意困难那一支若被读成"普遍加难", 与本文相冲突; 若被读成"在决定结局的位置上保留必要的困难", 则与本文相容且本文给出了那个位置的判据。

第三，本文不否定分布式认知。飞行、医疗、工程与复杂组织依赖人机团队，系统能力本身具有独立价值。如果教育目标是学会使用稳定工具完成系统任务，持续卸载可以是合理设计。只有在机构声称个体已经获得可离开工具运行的能力时，才需要个体转化证据。

第四，安全关键场景需要限制试错，而且这里的收窄是硬的。不能为了提高归属权，让初学者在真实病人、危险实验或关键设备上独立探索；可撤回装置在这类场景中同样不适用——真实操作没有撤回键，而在有撤回键的场合训练"可以撤回"本身可能形成危险习惯。处置是：在模拟环境、低风险子任务与预演中保留分岔步与可撤回装置，在真实操作中使用标准化检查并明确禁止撤回。管理学由此削弱了"所有课堂都应最大化自主生成"的价值排序。

第五，学习障碍与无障碍支持需要谨慎处理。读屏、公式表、语言辅助与记忆补偿可能是合理的环境适配，不应被视为代做。判断依据仍是课程声称的目标能力：支持是否替代了该目标的分岔操作，还是仅移除与目标无关的障碍。此处第五节的区分格外有用——绝大多数无障碍支持落在非分岔步上。

第六，可逆性本身有代价。允许撤回会延长任务时间、增加系统复杂度，并可能降低单轮的承诺强度。第十六节已把处方收窄为"一次撤回且入账"，但这个收窄是本文自己推出的、尚未被数据检验的；若实测显示即使一次撤回也显著削弱首次作答的信息量，处方需要进一步收窄，可能收窄到"撤回只在特定单元开放"。

第七，旁观轮次是轮次属性，不是人格属性。教师与学生都处于制度约束之中。把这个概念用于责备依赖型学生或灌输型教师，会把系统问题重新个人化，并刺激新的文书表演。

十八、证伪条件、竞争解释与一条写死日期的赌注

最强的竞争解释。最朴素的一条是：延迟迁移差只因为学生基础弱、动机低、学习时间少或任务太难，外部支持只是对困难的响应而不是导致低迁移的机制。第二条来自认知负荷：逐步提示降低负荷，任何迁移差异都可由负荷或练习量解释。第三条是测试不对齐：所谓迁移下降只是后测测了不同的知识。第四条针对本文最独有的主张：可撤回条件之所以有效，只是因为它变相增加了练习量与在任务上的停留时间。本文必须在控制这些因素后仍观察到效应；就第四条而言，E 组设计中必须把总时间与练习次数与 C 组匹配，或把时间作为协变量。

机制证伪条件。

其一，控制先前成绩、动机、总学习时间、教师效应与任务难度后，若归属权与延迟迁移不存在剂量一反应关系，则核心机制为伪。

其二，在即时正确率相同的任务中，若"外部完成分岔步"与"学习者完成分岔步"两组的无支持迁移长期无差异，则归属机制为伪。

其三，若等量协助落在分岔步与落在非分岔步的两组迁移无差异，则第五节的术语修正没有经验意义，本文应退回协助两难的一维刻画。

其四，若在协助总量、时机、形态、总时间与练习次数全部匹配的条件下，仅操纵首次作答的可撤回性，两组延迟迁移无差异，则第六节的推导为伪，本文相对既有框架的独立贡献消失。

其五，若可逆度与延迟测验之间不存在交互，即低可逆组中延迟测验同样有效，则门槛预测为伪。

其六，若某项帮助的催化比长期单调上升而该位置的归属率与迁移并未下降，则催化比不是可用读数。

其七，若管理压力并不早于提示密度上升，提示密度也不早于自发步骤下降，而顺序稳定相反，则三轮模型为伪。

其八，若旁观轮次只在生成式模型场景出现，在传统讲授、同伴领做、模板写作与实验步骤中无法复现，则跨场景理论范围必须缩小。

其九，若有人给出一个不依赖"这一步该由谁走"的判据，能够稳定区分同一次点击是省时还是替代，则第十二节的强制不一致主张失败。

其中第三条与第四条可用既有平台低成本检验：同一门课两个平行班，协助脚本完全一致，只改作业系统的提交锁定设置；期末同一份闭卷迁移测验。数据全在现有教务系统内，一学期可完成。若控制混淆变量后无效，理论应接受失败，而不是把结果解释为"编码还不够深"。

一条写死日期的赌注。到 2030 年 12 月 31 日之前，若至少三项在不同学科课堂开展的多班级研究报告：在时间投入、任务难度与协助总量三者匹配的条件下，仅改变首次作答的可撤回性即产生无支持延迟迁移的差异，且方向为可撤回组更高，则本文核心判断获得初步支持。以下明确**不算命中**：仅延长测验延迟；仅撤除支架；仅改变模型提示形态或加装护栏；仅报告参与度、满意度或即时正确率；机构自行声明采用了过程性评价或形成性评价；把"允许重做"与其他多项改革打包实施而无法分离贡献。若截至该日期没有出现这种可重复结果，或结果系统性为零，本文应降格为一个启发式比喻。

十九、反噬预言与伦理限制

本文一旦进入管理，最省事的实施方式恰恰可能摧毁它。学校把归属率变成排名，教师便安排学生在镜头前展示自主步骤；平台把提示藏得更深但仍交付等价答案，使催化比在账面下降而实际替代不变；学生生成形式草稿以通过审计；而可撤回装置最容易被做成一个从不被使用的按钮——挂在界面上、写进采购清单、计入合规得分，而默认路径仍是提交即锁定。这是本文预见到的最可能的失败方式，因为它比前三种更省事：另外三种还需要有人演，这一种什么都不用做。

对此本文没有彻底出口，只能规定四条底线。读数指向任务位置而非个人。不得以提高读数为由在安全关键系统制造轮换或风险。任何基于该读数的不利安排都必须公开编码规则、原始证据与复核通道。可撤回装置的验收必须看**实际撤回率**而非功能是否存在——一个撤回率长期为零的系统，与没有这个功能是同一件事，且更有害，因为它已经结清了合规账。

二十、自反：本文自身是一次不可逆的阅读

本文批评课堂把完整产物交给学习者，却也以一篇结构完整的论文把概念、模型、假设与干预一并交给读者。更严重的是，按本文自己的推导，阅读这件事恰好满足不可逆条件：读过便不能未读，一个名字一旦被给出，“我自己能不能命名这个现象”这个问题已经无法再问。于是本文预言，读者从本文取走的最可能是那个生成最快的产物——记住“旁观轮次”这四个字——而不是那个更稳定的产物。

本文能为自己配的唯一可逆装置，是第十一节、第十三节与第十四节那些判决性观察，以及第十八节那九条证伪条件：它们允许读者先自己判一个案例，再回来看本文怎么判。因此本文对自身的撤除测试是：不翻阅正文，能否对一个表面上很活跃的课堂判断其是否发生旁观轮次，指出分岔步在哪里，说明它与脚手架、认知卸载、协助两难与元认知惰性的差别，并给出一个会推翻这个判断的观察。不能做到，发表和阅读都不构成理论学习。

还有一条更不利于本文的自反。第十二节论证学习分析在不承认本文变量时无法自治——而本文自己提出的三个读数，同样面临被做成仪表盘的命运，且第十九节已经预言了它们会怎样被做坏。本文没有理由认为自己的读数比参与度指标更耐受这种命运；能说的只有一句：本文至少把自己会怎样失败写在了正文里，这使得失败可以被指认，而这正是参与度指标当年没有做的事。

二十一、结论

课堂上可迁移学习稀缺，并不意味着课堂什么都没有发生。恰恰相反，课堂往往高效地产生了答案、笔记、作业、项目与分数。问题在于，这些产物属于课堂系统，而制度经常把它们直接记为学习者的状态变化。归属不由在场决定——这是化学在写净离子方程时每天都在做的事，也是课堂账本从未做过的事。

三门学科在这里各出一份不可替代的力。教育学揭示认知路径：执行了什么操作，知识便沿什么路径组织为可用能力。管理学解释环境的反向选择：快、稳、可计量的系统产物为何持续获得资源与合法性。化学则给出两条本文最承重的形式约束：其一，在场不构成归属，产物不识别路径，因此需要一份净方程而不是总反应式；其二，多路径情形下决定结局的是分岔处而非最慢处，而两条路径中哪一条占优，第一取决于可逆性而非障碍高度。

这一理论把课堂评价的核心问题从“学生答对了吗”改为“决定能力形成的那一步由谁完成”；把教学改革从“活动更多”或“难度更高”改为“分岔步交还学习者”；把管理从即时产出单账本改为系统产物与个体转化双账本；把改革的次序从“延后测验、撤除支架”改为“先让首次选择可以被撤回一次，再谈延迟与难度”。最后一条是本文唯一一条与既有处方正交、可单独检验、且实施成本近乎为零的主张，也是本文最可能被推翻的地方。

可迁移的学习并不要求教师沉默、工具退出或学生孤立奋斗。它要求课堂系统的帮助不再以学习者不进入净变化为代价，也要求账本上那个被划掉的位置，终于有人去看一眼。

注释

一、本文所用"净离子方程""旁观离子""选择性决定步""限速步""动力学控制""热力学控制""催化剂再生"等术语均取标准化学教科书与 IUPAC 术语汇编的通行含义，用于提供形式约束，不作为关于神经过程的经验主张。第五节所做的术语修正针对的是既有跨学科文献对"限速步"的借用，不针对化学本身的使用法。

二、第六节的形式推导是一个路径选择模型，不是学习模型。它假定两条路径的当轮可见回报相同，该假定只在第三节所定义的靶格内成立。

三、第十二节的强制不一致主张限于"参与度作为单调代理"这一具体配置，不针对学习分析这一学科的全部工作；该行内部关于钴空子检测的研究正是本文所指问题的自觉产物。

四、本文为理论论文。涉及实证效果、阈值与研究方案的内容需在正式投稿前由作者逐项核验，并按目标期刊规范补充伦理审批与数据计划。

参考文献

- Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education*, 16, 101–128.
- Atkins, P., de Paula, J., & Keeler, J. (2018). *Atkins' Physical Chemistry* (11th ed.). Oxford University Press.
- Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the Cognitive Tutor classroom: When students "game the system." *Proceedings of ACM CHI 2004*, 383–390. doi:10.1145/985692.985741.
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. doi:10.1073/pnas.2422633122.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538. doi:10.1111/j.1467-9299.2006.00600.x.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. doi:10.1146/annurev-psych-113011-143823.
- Brown, T. L., LeMay, H. E., Bursten, B. E., Murphy, C., Woodward, P., & Stoltzfus, M. (2018). *Chemistry: The Central Science* (14th ed.). Pearson.
- Buskirk, A. V. (2009). Can reaction mechanisms be proven? *Journal of Chemical Education*, 86(5), 551. doi:10.1021/ed086p551.

- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67–90.
- Carey, F. A., & Sundberg, R. J. (2007). *Advanced Organic Chemistry, Part A: Structure and Mechanisms* (5th ed.). Springer.
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. doi:10.1080/00461520.2014.965823.
- Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, 116(39), 19251–19257. doi:10.1073/pnas.1821936116.
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113(1), 1–40. doi:10.1086/517897.
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530. doi:10.1111/bjet.13544.
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a Generative Activity*. Cambridge University Press.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. doi:10.1073/pnas.1319030111.
- Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
- IUPAC. (current online edition). *Compendium of Chemical Terminology (Gold Book)*: activation energy; catalyst; intermediate; kinetic control; mechanism; rate-controlling step; reaction coordinate; thermodynamic control. International Union of Pure and Applied Chemistry.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. doi:10.1080/07370000802212669.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772–775. doi:10.1126/science.1199327.
- Kerr, S. (1975). On the folly of rewarding A, while hoping for B. *Academy of Management Journal*, 18(4), 769–783. doi:10.5465/255378.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist*, 41(2), 75–86. doi:10.1207/s15326985ep4102_1.
- Koedinger, K. R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review*, 19(3), 239–264. doi:10.1007/s10648-007-9049-0.

- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. doi:10.1287/orsc.2.1.71.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363. doi:10.1086/226550.
- Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford University Press.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. doi:10.1016/j.tics.2016.07.002.
- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9. doi:10.3102/0013189X020003002.
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176–199. doi:10.1177/1745691615569000.
- Stains, M., Harshman, J., Barker, M. K., et al. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470. doi:10.1126/science.aap8892.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. doi:10.1207/s15516709cog1202_4.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22, 271–296. doi:10.1007/s10648-010-9127-6.
- 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. doi:10.1111/j.1469-7610.1976.tb00381.x.

本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 作者：王德生 + Claude · SDE Universes · 学科通融