

总有人替它起步

论学习产物要发起下一轮需要一段未被归类的成熟时间，及为什么外部启动越可靠，这条通路越无法被观察到

教育学 × 化学反应动力学 × 管理学

作者 王德生 + Claude

栏目 学科通融 · 站外碰撞

篇幅 约2.0万字 · 28页

发表 2026年8月3日

摘 要

课堂当轮生成了正确答案、概念修正与高参与度，而这些产物很少成为下一轮的原因。对「下一轮该由谁发起」，三家给出三条不能同时为真的答案：教育学答由教师依诊断发起、且反馈越及时越好；管理学答由制度发起，因为下放必被公开指标做成表演；反应动力学答由产物发起，因为不回到条件位的产物只有局部转化、没有传播。三家共有个没说出口的前提——一个产物在它被生成的那一轮就已经具备或不具备接手下一轮的资格。推翻它的材料来自反应动力学自己：**自催化的诱导期**。产物积累到阈值之前速率几乎观察不到，**资格不是产物一生成就有的属性，是一个需要时间达到的状态**；而在诱导期之内测速率，得到的结论是「这个反应不发生」，可它正在发生。带回课堂，结论是刺的：**课堂的一切结算都发生在诱导期之内**。由此给出两个时间量（产物被判定之前的未归类窗口、产物长到能被接手所需的成熟时间）与两个速率，推出五条结论，其中三条方向违反通行判断：**任何缩短未归类窗口的改进都单调降低续发率**——而即时反馈、自动批改、实时正确率、即时解答恰是教育技术二十年的主要方向；**外部启动越可靠，这条通路越不显形**，不是它被损害，是底物在它显形之前已被另一条通路消耗完；以及一条可用现有元分析数据直接检验的：**教学法标签不含这个变量，故同一种教学法内部的差异可能大于不同教学法之间的差异**。文中另论证形成性评价同时持有「证据必须改变教学」与「反馈应当及时」两条处方，在此条件下不能同时成立，且它现有的出路是循环的；并从化学借来一条少见的**正面判据**——加料实验。

这一篇由哪三个学科的理论体系撞成

第一家主张下一轮应由教师依据诊断发起，因为最小指导对新手无效，且反馈应当及时、具体、指向下一步；第二家主张应由制度发起，因为协调、覆盖与责任不可下放，而任何把发起权交给现场的做法都会被公开度量的反应性做成表演；第三家主张应由产物发起，因为自催化物提供的是一条活化能更低的替代成产通路，不让产物回到入口，体系每一轮都得由外部重新点火。三条不能同时为真，且各自的日课正是另一家诊断出的病：教师依诊断作的调整会被评价系统在下一轮压回原计划，而「及时反馈」恰恰是把产物尽快归类的机制（管理判教育）；永远由制度发起的体系没有自催化，它每一轮所需的外部投入不随轮次下降（化学判管理）；把发起权交给一个尚未成形的产物，对新手就是最小指导（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 · 形成性评价、及时反馈与对最小指导的批评

[Black & Wiliam, Assessment and Classroom Learning \(<i>Assessment in Education</i> 5:1, 1998\)](#); 另见 [Hattie & Timperley, The Power of Feedback \(<i>Review of Educational Research</i> 77:1, 2007\)](#); [Kirschner, Sweller & Clark \(<i>Educational Psychologist</i> 41:2, 2006\)](#)

主张证据必须改变后续教学，而调整权驻于教师；同一支又主张反馈应当及时、具体、指向下一步。两条核心处方在同一套操作里互相削减——及时的反馈正是把产物尽快放进类目的那个动作。

<https://doi.org/10.1080/0969595980050102>

化学 · 自催化的动力学签名与诱导期

[Hanopolskyi, Smaliak, Novichkov & Semenov, Autocatalysis: Kinetics, Mechanisms and Design \(*ChemSystemsChem* 3:1, 2021\); 另见 Bissette & Fletcher, Mechanisms of Autocatalysis \(*Angew. Chem. Int. Ed.* 52:49, 2013\); Bentea, Watzky & Finke \(*J. Phys. Chem. C* 121:9, 2017\)](#)

这一支写明自催化由**动力学签名**定义而非由机理定义，且滞后段本身不构成证据——中间体的积累同样会造成滞后段，区分二者必须做加料实验。前一条否掉了一整类基于形态的课堂判断，后一条既给出本文唯一的正面判据，也给出与本栏一篇先行研究的判决性分离。

<https://doi.org/10.1002/syst.202000026>

管理学 · 组织例程与公开度量的反应性

[Espeland & Sauder, Rankings and Reactivity \(*American Journal of Sociology* 113:1, 2007\); 另见 Levitt & March, Organizational Learning \(*Annual Review of Sociology* 14, 1988\); Feldman & Pentland \(*Administrative Science Quarterly* 48:1, 2003\)](#)

主张例程是能力得以跨年度保存的方式，而公开度量会反过来重塑被度量者，故把发起权交给现场不可能稳定。这一支自曝的一处是：同一条反应性论证，同样适用于它自己以及本文所提出的每一个新读数。

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

目 录

摘要	5
Abstract	5
一、三条不能同时为真的答案	7

二、共有前提，与推翻它的材料	8
三、核心概念：结清态	9
四、两个时间量	10
五、两个速率：为什么外部启动越可靠，这条通路越看不见	10
六、五条推论	11
七、二维辨别格与四个读数	13
八、加料实验：一条正面判据	14
九、九种结算装置	15
十、被强制的不一致：形成性评价	16
十一、外部划界：规范全部满足而现象仍不发生	17
十二、站内划界：与本栏两篇先行研究的分工	18
十三、假设与判决变量	18
十四、研究设计与操作化	19
十五、干预：延迟的是归类，不是反馈	20
十六、五类场景与其中两处收窄	21
十七、适用边界与三门学科的反向约束	22
十八、竞争解释、证伪条件与一条写死日期的赌注	23
十九、反噬预言与伦理限制	24
二十、自反	25
二十一、结论	26
注 释	26

摘要

课堂研究已经积累了关于主动学习、提取练习、生成活动、合作讨论与形成性评价的大量证据，却仍然解释不了一个更尖锐的事实：许多课堂在当轮形成了正确答案、概念修正与高参与度，而这些产物很少成为下一轮的原因。对“下一轮该由谁发起”这个问题，三门学科给出三条不能同时为真的答案：教育学答由教师依据诊断发起，并主张反馈越及时越好；管理学答由制度发起，因为协调、覆盖与责任不可下放，且任何下放都会被公开指标做成表演；反应动力学答由产物发起，因为不回到条件位的产物只有局部转化，没有传播。

三家争的是发起位置，而它们共有一个没有说出口的前提：一个学习产物在它被生成的那一轮就已经具备或不具备接手下一轮的资格。推翻这个前提的材料来自反应动力学自己——**自催化的诱导期**。自催化反应在产物积累到阈值之前速率几乎观察不到，只有产物累积起来体系才转入加速；也就是说，产物能不能回到条件位并驱动后续，不是它一生成就有的属性，是一个需要时间达到的状态。在诱导期之内测速率，会得出“这个反应不发生”的结论，而它正在发生。带回课堂，结论是刺的：**课堂的一切结算都发生在诱导期之内**。

据此本文提出两个时间量与两个速率。未被归类窗口 w 是产物生成到它被判为对错、分数、标签或已完成之间的时间；接手成熟时间 m 是产物展开到能被下一轮接手所需的时间；续发的必要条件是 $w > m$ 。外部启动速率 k_1 与产物启动速率 k_2 则决定这条通路能否被观察到。由此推出四条正文之外的结论，其中两条方向违反常识：**任何缩短 w 的改进在 m 不变时单调降低续发率**——而即时反馈、自动批改、实时正确率与即时解答恰好是教育技术二十年的主要进步方向；**外部启动越强越可靠，自续发通路越不显形**，不是因为它被损害，而是因为底物在它显形之前已被非自催化通路消耗完。

本文进一步论证：形成性评价这一行同时持有“证据必须改变后续教学”与“反馈应当及时”两条核心处方，而在 $w > m$ 之下二者在同一套操作里不能同时最大化；它现有的出路（反馈不给分）是循环的，因为要判定一次反馈是归类还是保持开放，只能用“下一轮是否因它而改变”这个判据。本文给出二维辨别格、四个读数、一条来自化学的**加料实验**正面判据、与三家外部近邻及本栏两篇先行研究的判决性分离、九条可证伪命题、多班级研究设计，并把干预定位在一处此前没有被指认的位置：**要延迟的是归类，不是反馈，也不是测验**。

关键词：结清态；未被归类窗口；诱导期；自催化；形成性评价；发起权

Abstract

Research on active learning, retrieval practice, generative activity, and formative assessment has established many ways to improve within-round performance, yet cannot explain why learner products so rarely become causes of what happens next. On the question of who initiates the following round, three disciplines give three answers that cannot all be true.

Education answers: the teacher, on diagnostic evidence—and feedback should be timely. Management answers: the institution, because coordination, coverage and accountability cannot be devolved, and any devolution is reshaped by public measures into performance. Reaction kinetics answers: the product, because a product that does not return to the condition position yields local conversion without propagation.

All three share an unstated premise: that a learner product either does or does not possess, at the moment of its generation, the standing to initiate the next round. The material that overturns this premise comes from kinetics itself—the **induction period** of autocatalysis. Below a threshold concentration of product, an autocatalytic reaction shows almost no measurable rate; only once product accumulates does the system enter acceleration. Standing to drive what follows is therefore not a property a product has on generation; it is a state reached over time. Measure the rate inside the induction period and you conclude the reaction is not happening while it is. Returned to the classroom the consequence is sharp: **every settlement occurs inside the induction period.**

The paper introduces two times and two rates. The unclassified window w runs from product generation to the moment it is judged correct, scored, labelled or marked complete; the maturation time m is what the product needs to become take-up-able; continuation requires $w > m$. External initiation rate k_1 and product initiation rate k_2 jointly determine whether the pathway is observable at all. Four consequences follow, two counterintuitive: any improvement that shortens w monotonically lowers continuation at fixed m —and instant feedback, auto-grading, live accuracy displays and immediate answers are precisely the direction of two decades of educational technology; and the more reliable external initiation is, the less the self-continuing pathway shows, not because it is damaged but because substrate is consumed by the non-autocatalytic route before it can appear. The paper further argues that formative assessment holds both "evidence must change subsequent teaching" and "feedback should be timely," which under $w > m$ cannot be jointly maximised; its existing escape route is circular. It supplies a diagnostic matrix, four measures, a positive **seeding test** borrowed from kinetics, decisive separations from three external neighbours and two prior papers in this column, nine falsifiable propositions, a multi-class design, and an intervention located where none has previously been placed: **what must be delayed is classification, not feedback and not testing.**

Keywords: settled state; unclassified window; induction period; autocatalysis; formative assessment; initiation

一、三条不能同时为真的答案

一间课堂可以非常热闹。学生分组，白板写满答案，教师不断追问，平台实时显示正确率，下课时还能生成一份漂亮的学习报告。以参与次数、当堂测验、任务完成率和满意度为尺度，这节课甚至会被判为高质量课堂。然而第二天仍按预先写好的幻灯片开始：学生上一轮提出的异议没有改变新任务，那个尚未解决的错误没有成为新材料，一段意外的解释也没有进入评价标准。课堂产生了许多局部产物，却没有让这些产物拥有后续的因果力。

主动学习研究已有充分证据说明主动参与通常提升成绩并降低失败率（Freeman 等，2014）；提取练习在长期保持上有显著优势（Karpicke 与 Blunt，2011）。但效果存在不等于概念已经穷尽。Deslauriers 等（2019）发现学生在主动课堂中实际学得更多却感觉学得更少，说明学习的实际发生、主观感受与可见形式三者并不重合。课堂研究已经知道怎样让当轮表现变好，却未必知道怎样让当轮产生的差异取得下一轮的发起权。

问题因此可以收得很窄：**要让课堂上的学习延续下去，下一轮该由谁发起？**对这一问题，三门学科给出三条互相取消的答案。

教育学答：由教师依据诊断发起。理由有两层。第一层是能力层：最小指导在复杂新手任务中会造成过度搜索，学生的自发方向常常是低效的（Kirschner、Sweller 与 Clark，2006）。第二层是制度层：形成性评价的整套设计就是"教师读取学生证据，教师据此调整教学"（Black 与 Wiliam，1998），而这一支同时主张反馈应当及时、具体、指向下一步（Hattie 与 Timperley，2007）。发起权因此合理地留在教师手里，学生产物的位置是被读取的信号。

管理学答：由制度发起。组织学习以例程为基础、依赖历史并面向目标（Levitt 与 March，1988）；课程标准、周进度、作业制度与评分规则把教学组织成可重复运行的例程，而这正是能力得以跨年度保存的方式。更要紧的是这一支的第二条判断：公开度量并非被动描述，它会改变被度量者的注意、资源与行为（Espeland 与 Sauder，2007）。因此任何把发起权交给现场的做法都不可能稳定——一旦"学生发起"被记录、被计数、被检查，它就会被做成表演。发起权只有留在制度手里才不会被自己的反应性吞掉。

反应动力学答：由产物发起。一个产物若不回到条件位，体系只有局部转化；能让体系自加速的唯一机制是自催化，即产物参与促进自身的进一步生成（Bissette 与 Fletcher，2013）。自催化的要点被这一支写得很清楚：**自催化物提供的是一条活化能更低的替代成产通路**（Hanopolskyi 等，2021）。也就是说，产物不只是终点，它可以是下一段路的入口；不让它回到入口，体系每一轮都必须由外部重新点火。

三条答案不能同时为真，而且各自的日课正是另一家诊断出的病。

若管理学对，教育学的处方无法执行：教师依诊断作出的调整会被评价系统在下一轮压回原计划，而"反馈应当及时"这一条恰恰是把产物尽快归类的机制——它让教师的诊断变成一次结算，而不是一次续发。若反应动力学对，管理学在原理上封死了唯一的自加速通路：一个永远由制度发起的体系没有自催化，它每一轮所需的外部投入不随轮次下降，这与"稳定例程保存能力"的主张相冲突。若教育学对，反应动力学

的处方对新手就是最小指导：把发起权交给一个尚未成形的产物，等于让低效搜索决定路径，而这一点已被论证行不通。

三家必须有两家错，或者三家的病灶都放错了位置。本文主张是后者。

二、共有前提，与推翻它的材料

三家争的是发起位置：在教师、在制度、还是在产物。它们却共同站在一块未被审视的地基上——一个学习产物在它被生成的那一轮，就已经具备或不具备接手下一轮的资格。教育学问它够不够好（够好才值得跟进），管理学问它合不合规（合规才进得了流程），反应动力学问它是不是产物（是产物就能回到条件位）。三家问的东西不同，却都把“资格”当成产物的一个当属性：看一眼就能判定，而且判定的时刻可以是它刚出现的那一刻。

推翻这条前提的材料在反应动力学自己手里，而且是这一支最先教给初学者的那一条。

自催化反应的浓度—时间曲线是一条 S 形：起初速率几乎观察不到，因为自催化物浓度接近零；随着产物累积，速率逐步加快；最后随反应物耗尽而放缓。起初那一段叫**诱导期**。这一支把 S 形曲线称作自催化的第一个实验征候，并把它拆成三段：缓慢的引发或诱导段、速率升至极大值的加速段、以及产物多于起始物之后的饱和段。

要害在诱导期那一段的处境。在诱导期之内测速率，得到的结论是“这个反应不发生”，而它正在发生。产物具不具备回到条件位并驱动后续的能力，不是它一生成就有的属性；它是一个需要时间达到的状态。刚生成的那一点产物在数量上不足以改变任何东西，这不说明它没有资格，只说明还没到时候。

这一条一带回课堂，共有前提立刻塌掉。课堂在产物生成的**当轮末端**结算它：答案被判对错，作业被评分，发言被表扬或纠正，讨论在铃声处收束，进度表在下一节重新启动。而按诱导期，那正是它还不具备接手能力的时刻。**课堂的一切结算都发生在诱导期之内**。三家据以判定资格的全部依据——不够好、不合规、只是个局部产物——恰好逐条对应“在诱导期内观测所必然得到的读数”。

同一支还给了本文第二条更硬的材料。这一支明写：**自催化这一现象由它的动力学签名定义，而不是由机理定义**。换言之，判断一个体系是不是自催化的，不能靠观察它的组成与结构，只能靠观察它随时间的行为。带回课堂，这一条否掉了课堂研究最常用的一整类判断方式：**一间课堂有没有续发，不能靠看它像不像学生中心、活动多不多、话语开不开放来回答，只能靠看轮次之间的时间行为来回答**。所有基于课堂形态的判断在原理上不足以裁决这个问题。

三门学科在本文中各占一个不可替代的位置。教育学占据产物的形成：什么条件与路径使学习者当轮生成何种差异。反应动力学占据资格的时间性与显形条件：资格要时间达到，而这条通路能否被观察到取决于两条通路的相对速率。管理学占据结算：哪一种产物在什么时刻被转译为记录，以及记录完成之后由谁重新发起。删掉任何一家，剩下两家都无法说明为什么高投入的课堂长期低续发。

必须写明证据等级。既有研究能支持的是局部关系：主动生成、提取练习、延迟测验、指标反应性与工具护栏的效应。反应动力学能提供的是关于时间结构与可观测性的形式约束，不是关于神经过程的直接实验证据；"诱导期"不是一个待测的脑内时长，"速率常数"也不是某种神经参数。它们的作用是迫使研究者提出对应的经验问题：一个产物从生成到能被接手要多久，它在多久之后被归类，外部启动占了多大比重。只有当这些问题被操作化为课堂日志、录像与教案版本，跨域结构才有研究价值。

学科来源	单因主张	能解释什么	自身暴露的缺口
教育学	教师依诊断发起下一轮，反馈越及时越好	为何最小指导对新手无效，为何证据必须进入教学决策	"及时"与"证据改变教学"在同一套操作里互相削减
反应动力学	产物回到条件位才有传播；资格由时间达到	为何局部转化不等于传播，为何形态判断不足以裁决	时间结构不说明课堂里是哪一道工序在做结算
管理学	制度发起才稳定，下放必被反应性做成表演	为何课堂偏好快、稳、可计量的产物	反应性论证同样适用于制度自己给出的每一个新指标

三、核心概念：结清态

本文把**续发事件**定义为一个可反事实判定的课堂微事件：学习者在第 t 轮生成了差异 $X_{_t}$ （解释、问题、错误、策略或模型），若在其他条件相近时删除 $X_{_t}$ ，第 $t+1$ 轮的活动或条件将发生不同安排，则记该轮发生一次发起权交接。反事实这一条不是修辞上的严格，它有具体的排除功能：教师可以表扬、纠正、记录一个回答，而下一页幻灯片、下一道题与评分标准完全不变——此时产物被处理了，却没有因果改写力。相反，一个错误可能让教师取消预定讲解，要求全班比较两套模型，并把比较结果写进下一次作业的评价依据；此时即使错误本身质量不高，交接也已发生。

据此，本文命名的对象是**结清态**：一个学习产物在达到能够接手下一轮所需的成熟度之前，已被归类为对错、分数、标签或已完成，因而失去取得后续因果力的可能。

三条性质使它区别于邻近的坏情况。第一，它不是"没学会"——当轮可能有高分、好解释与实在的理解。第二，它不是"没反馈"——反馈可能非常及时、具体、指向下一步；恰恰是这种反馈最有效地完成归类。第三，它不是"没参与"——课堂可能高度互动，而互动的每一轮都在当轮末端被清算。它指的是：本轮已经生成了足以改变后续的差异，而它在长到能改变后续之前就被结清了。

"归类"这个词需要精确。它不等于"评价"，也不等于"回应"。归类是指把产物放进一个**已有的类目**并据此结束对它的追问：对或错、几分、哪一档、掌握或未掌握、已完成。一次说"你这个想法有意思，它和刚才那位同学的说法冲突在哪里"的回应不是归类，因为它没有把产物放进已有类目，产物仍处于可展开状态。一次说"很好，答对了"的回应是归类，即使它是正面的、及时的、鼓励性的。**归类的标志不是褒贬，是追问的终止。**

需要一并声明一处收窄。本文把续发权的持有者限定为**可被指认的学习者产物**——个体的、小组的，或一段可追溯到具体发言与文本的共同产物。若把它扩展到"人机协作产物"这类无法指认删除对象的东西，反事实判据立刻失效：删除一个说不清边界的东西，无法判断下一轮会不会不同。人机协作场景仍在本文范围内，但编码时必须把学习者提交的那一部分单独标出，否则该轮不计入。

四、两个时间量

诱导期把一个时间量强加给这项研究：产物从生成到具备接手能力，需要一段不为零的时间。同时，课堂把第二个时间量强加进来：产物从生成到被归类，也有一段不为零的时间。两者的关系决定一切。

接手成熟时间 m ：一个学习产物展开到能够被下一轮接手所需的时间。它不是"学生想清楚所需的时间"，而是产物本身长到一个可被别人接住的形态所需的时间——一个模糊的直觉要长成一个能被比较的主张，一个错误要长成一个能被检验的假设，一个提问要长成一个能被安排为任务的问题。 m 随任务类型、学科与学习者先备水平变化，因此必须实测而不能规定。

未被归类窗口 w ：从产物生成到它被判为对错、分数、标签或已完成之间的时间。它是一个纯外部量，全部证据已经在课堂录像、平台时间戳与教案版本里；测它不需要窥测任何人的内部。

续发的必要条件是 $w > m$ 。这不是一条经验概括，它从两个量的定义直接得出：若产物在长到能被接手之前已被归类，追问已经终止，它不可能取得下一轮的发起权。必要而不充分—— $w > m$ 只是让它有机会，接不接得上还取决于下一轮有没有位置接。

这个必要条件立刻带来一件事："这节课有没有留出时间"这个问题第一次有了一个不含判断词的问法。不是问教师有没有耐心，不是问课堂开不开放，而是问：这一轮的产物出现之后，过了多久被判为对错？这个数当场可算，且与教学法标签正交——一节讲授课可能有很长的 w （教师听完不表态，留到下节课处理），一节热闹的讨论课可能 w 极短（每个发言后教师立即评价）。

m 更难测，但不是不能测。可行的估法是在人为延长 w 的条件下观察续发事件出现在产物生成后多久，取其分布；也可以用"学生自行把一个初始表述改写到能被他人接手的形态"所需的时间作代理。要紧的是承认它是一个待估参数，而不是把它设成一个理论常数。本文对 m 只作一条弱主张：**在需要概念重组的任务上， m 明显大于一次课堂问答的往返时长**。这一条本身可被推翻。

五、两个速率：为什么外部启动越可靠，这条通路越看不见

$w > m$ 解释了产物为什么活不到轮次交界处。它不解释另一件更难堪的事：为什么课堂研究几乎从未观察到自续发的学习链——如果这条通路存在，为什么它在数据里几乎不出现。

答案同样在这一支手里，而且有现成的形式。**两步模型**把 S 形曲线拆成两条并行的通路：一条是慢而连续的外部引发（记为速率常数 k_1 ），另一条是快而自催化的增长（记为 k_2 ）。两条通路消耗同一批底物、生成同一种产物，诱导期的长短有关于 k_1 、 k_2 与初始底物量的闭式表达。

把这两条通路移入课堂，对应关系是清楚的，而且不是比喻：

- k_1 是**外部启动**——教师、教案、进度表、题库、平台算法发起下一轮的速率。它高，意味着每一轮都能被可靠、准时、按计划地启动。
- k_2 是**产物启动**——学习者产物发起下一轮的速率。
- 两条通路争夺同一批底物：一节课的时间、注意与任务位。

由此得到一条本文最不直观、也最容易被误读的结论：

> 当 k_1 足够大时，体系的产出曲线与“没有自催化”的体系无法区分。

理由是形式上的：底物在 k_2 那条通路显形之前，已经被 k_1 那条通路消耗完了。曲线不再是 S 形，而趋近于一条平稳的一阶衰减；诱导期消失，加速段也消失。要紧的是：这时 k_2 并没有变小。自续发的能力没有被损害，它只是不显形。

这条结论解释了三件此前各自被单独归因的事。

其一，课堂研究里几乎看不到自续发链，不构成它不存在的证据。主流课堂的 k_1 极大——每一轮都由铃声、教案、题库准时启动，从不缺席。在这样的体系里，即使 k_2 不为零，观测曲线也不会显出任何自催化特征。以往把“没观察到”读作“没有”，是一个在这个模型下必然发生的误读。

其二，“好教学”与“续发可见”之间存在一处结构性冲突。教师越可靠、备课越充分、进度越准时， k_1 越大，这条通路越不显形。这不是说好教学有害——本文明确不主张这一点；它说的是“好教学”与“看得见续发”是两个不同的目标，且在同一节课里互相遮蔽。任何以“这位教师课上得好不好”为准去挑研究样本的做法，都会系统性地挑到看不见续发的那些班。

其三，要看见它，得先停一段外部启动，而不是减少帮助。化学里判断体系自己走不走，做法是撤掉引发剂看它还动不动，不是稀释所有试剂。翻回课堂：不是减少教师讲解、不是撤除支架、不是降低难度——那些改的是帮助的量；要改的是在某些轮次上不由外部发起。教师可以照常讲解、照常提供材料、照常回答问题；他不做的只有一件事：不决定下一轮从哪里开始。这两件事在日常语言里常被混为一谈，在这个模型里是两个不同的参数。

六、五条推论

前两节给出的是两组量。它们推出的结论有五条，其中三条方向违反教育界的通行判断，因此也最容易被检验；第五条可以用现有的公开数据直接检验，不需要任何新实验。

推论一：任何缩短 w 的改进，在 m 不变时单调降低续发率。

这条从必要条件直接得出。要紧的是它的射程：即时反馈、自动批改、实时正确率显示、答题器的当场统计、生成式模型的即时解答——教育技术二十年的主要进步方向，全部是缩短 w 的技术。它们各自都在自己的目标上有效（提高当轮正确率、减少挫败、缩短教师批改时间），而它们对续发率的作用同号且累

加。这解释了一个长期被读作“技术尚未用好”的现象：数字化程度越高的课堂，产物越快被归类，续发率越低——不是工具用错了，是工具做的正是那件事。

推论二：诱导期使短周期评估系统性地判定续发型设计无效。

若评估窗口短于 m ，续发型设计在评估期内的表现不优，甚至更差——因为它花掉了时间而尚未进入加速段。这解释了为什么“学生中心”类改革的短期试点常常报告零效应或负效应，而这在本模型下不构成对该设计的否定；它是诱导期的必然读数。反过来说，这条也是对本文自己的约束：**本文不能用“效果还没显出来”去解释任何失败**，因为那正是这条推论提供的万能退路。防止它成为退路的办法是事先写死评估窗口的下界——本文在第十八节的赌注里写了这个下界。

推论三： k_1 与 k_2 的相对大小决定可观测性，而非决定能力。

见上节。它给出的可检验形式是：在 k_2 相同的两组课堂中， k_1 高的那组观测到的续发事件更少，而在停止外部启动的轮次上，两组的续发速率应当相近。若停止外部启动之后两组差异反而扩大，说明差的是 k_2 而不是可观测性，本条为伪。

推论四（干预）：要提高续发率，不需要减少反馈，只需要把归类推迟到 m 之后。

这是四条里最要紧的一条，因为它把干预放在了一个此前没有被指认的位置。既有处方都落在别处：延迟测验改的是测量时点，撤除支架改的是帮助强度，增加活动改的是课堂形式，先生成后讲解改的是顺序。它们没有一条动“归类的时刻”。而按第三节对归类的定义，反馈与归类是可分的：指出冲突而不判对错、提问而不给类目、并置两个说法而不宣布哪个成立——这些都是反馈，且都不归类。

因此本文的处方可以写成一句不含判断词的话：**在产物生成之后的一段时间内，只做不终止追问的回应**。这段时间要多长，由 m 的实测决定，不由理念决定。

推论五：教学法标签不含 w ，因此按教学法分组的比较把一个决定性变量留在了组内。

这一条是五条里唯一一条可以用**已有的公开数据**检验的，因此值得单列。

主动学习、提取练习、生产性失败、项目学习、翻转课堂——这些标签描述的是活动的**形式**：谁在说话、用什么材料、按什么顺序。没有一个标签规定产物在生成之后多久被判定。因此同一个标签之下， w 可以从数秒变到数天：一节讲授课的教师可以听完学生的提问后说“这个我们下节课处理”（ w 以天计），一节讨论课的教师可以在每个发言后立即评价（ w 以秒计）。**标签不含这个变量，而按本文的必要条件，这个变量决定产物能否活到交界处。**

由此得到一条锐利且可立即检验的预测：**同一种教学法内部的续发差异，可能大于不同教学法之间的差异**。换言之，教育学几十年来最稳定的一个困扰——各类方法在元分析上平均有效，而效应量的异质性极大且难以归因——在本文的框架下不是噪声，是一个被系统性留在组内的变量所造成的。

这条预测可以分三步检验，成本递增：

第一步，最便宜，且不需要新数据。取一项已有的主动学习或生产性失败元分析，逐篇回看其纳入研究的干预描述与课堂材料，编码可获得的 w 代理（如“是否在活动后立即公布标准答案”“形成性测验的反馈是否即时呈现”“合作讨论后是否由教师当场收束”），检验它能否解释一部分残差异质性。若完全不能，本条为伪。

第二步，中等成本。在一批同标签课堂的现有录像中直接编码 w ，检验其组内方差是否大于跨标签的组间方差。

第三步，最贵。在同一教学法之下随机分配两种归类时刻，看长期结果是否分开——这就是第十四节的 A 组与 C 组。

要写明这条的限度。它不预测教学法标签**没有**作用——那是一个更强也更不可信的主张；它只预测标签所解释的方差与 w 所解释的方差是两块，且后者此前从未被单独取出过。若回看编码发现 w 在同标签研究内部几乎没有变异（即所有采用某方法的课堂都以相近的速度归类），那么本条也失败，而这本身会是一个有价值的发现：它意味着归类速度不由方法决定，而由更上位的东西决定。

五条推论共享一个弱点，必须写明：前四条全部依赖 m 是一个可估的量（第五条不依赖，它只用 w ）。若 m 在同一任务、同一批学习者上分布极散以致无法给出可用的下界，则 $w > m$ 虽然在定义上成立，在实践上不可操作，本文退化为一个描述性框架。这是本文最大的一处未闭合。

七、二维辨别格与四个读数

要把结清态从一个说法变成辨别工具，必须把它与产物质量分开。横轴取当轮产物质量（准确性、解释深度、结构化程度、迁移潜力），纵轴取该产物是否活到了轮次交界处并取得发起权。两轴相互独立：高质量产物可能被立即归类，低质量产物也可能作为反例发起高价值探究。

当轮产物质量		续发权低	续发权高
低	低	传输失灵：没有形成足够差异，课堂靠外部重新启动	探索胚胎：不成熟的错误、问题或模型改变下一轮，具有发展潜力
	高	结清态：答案、解释与参与均良好，产物在长成之前被判定完毕；本文的靶格	链式学习：产物取得后续发起权，并在下一轮改写条件

右下角是靶格，它有三种隐蔽性。短期指标表现为改善，不触发任何警报；失效发生在两轮之间，而两轮之间没有专门日志，因此不产生信号；它会自我加固——越依赖高正确率、准时覆盖与可审计活动，组织越需要快速归类差异， w 越短。左上角是关键对照格：**代做失败或质量低劣的产物反而安全**，因为它留下了信号。这一格的存在排除了“帮助有害论”与“及时反馈有害论”——问题不是反馈的存在，是归类的时刻。

四个读数。前两个当场可算，后两个需要设计。

读数一：未被归类窗口 w 。产物生成到被判为对错、分数、标签或已完成的时长。证据全在课堂录像时间戳、平台提交与批改记录、教案版本里。这个数不需要任何人配合，也不需要窥测内部，且**与教学法标签正交**——这一条是它的价值所在，也是它可被推翻的地方：若 w 与教学法标签或参与次数几乎同义，它就没有理论价值。

读数二：外部启动占比 $k_1/(k_1+k_2)$ 的可观测代理。在一节课的全部轮次中，由铃声、教案、题库或平台准时发起的轮次所占比例。这个数在多数课堂里接近一，而正是这一点使续发不显形。

读数三：续发率 CR。可编码产物中发生反事实交接的比例。编码规则要求同时满足三条：来源可追踪、下一轮发生结构变化、删除该产物则下一轮不同。防伪配套是硬的：教师须在课前提交并锁定路径版本；实际下一轮若发生偏离，必须在日志中标注触发产物；随机抽取事件由第三位编码者复核并报告一致性；**已发生但未记录来源的变化不能自动算作续发。**

读数四：接手成熟时间 m 的估计分布。见第四节。它是四个里最难的一个，也是本文与三家外部近邻的分离全部押注的地方。

四个读数都不是教学质量总分。一个高风险实验课可能有意保持极短的 w ，以保证安全；一个基础记忆单元可能以高准确率与低续发为合理选择。测量服务于识别，不拥有对课堂价值的最终裁决权。

八、加料实验：一条正面判据

到目前为止本文给出的都是必要条件与证伪条件。这一支还给了一件更少见的东西：一条**正面判据**——不是"若某事发生则本文错"，而是"若某事发生则这个体系确实在自续发"。

判据来自这一支自己对 S 形曲线的警告。这一支明写：S 形曲线是自催化的第一个实验征候，但**滞后段本身不构成证据**——其他过程（例如某种中间体的积累）同样会造成滞后段。要证明一个反应是自催化的，必须证明它的速率随产物的加入而升高；做法是在起始时加入少量产物，测初速率，看滞后段是否明显缩短。这叫**加料实验**。

翻回课堂，这条判据是可以照做的：

> 在下一轮开始时，**注入上一轮的一个产物**（一个学生的具体表述、错误或问题），其余条件不变；若该轮到达下一次续发的时间明显缩短，则这个课堂确实在产物驱动之下运行；若时间不变，则它此前表现出的任何"慢启动后来变好"的曲线，来自别的东西。

对照组的设计同样从这条判据抄来：注入**同等长度、同等难度、但不来自上一轮**的一段材料（教师准备的等价例子）。若两组的缩短程度相同，说明起作用的是"多给了一段材料"，不是"这段材料来自上一轮"。这个对照把本文与"多给材料就更好"分开了，而它本身极便宜——两个平行班，一节课就能跑。

这条判据还有一个不常见的性质：**它同时是本文与本栏一篇先行研究的判决性分离**，因为那一篇讲的正是这条警告里所说的"另一种可能"（中间体的积累）。这一点在第十二节详说。

九、九种结算装置

课堂缩短 w 的方式不是一种，而是一套互相加强的装置。列出它们的用处不是道德指控——每一条都在自己的目标上是合理的；用处在于指认它们在同一个量上同号叠加。

其一，答案装置。在学生模型尚未展开时给出标准答案，当轮正确率立即提高，而追问就此终止。判据不是教师给不给答案，是答案给出之后学生先前的产物还会不会改变下一轮。

其二，铃声装置。固定课时与内容覆盖要求把尚在展开的产物切断，下一节从课程计划重新启动。这常被理解为资源限制，它同时是发起权的制度分配。

其三，评分装置。产物被转为分数后，组织获得了可记录结果，继续追踪其展开的动力随之下降。结构新颖但不合量表的回答在这一步损失最大。

其四，救援装置。教师出于善意快速补足步骤、修正表达或替学生完成结构化。它降低挫败并保持节奏，同时把本应由产物发起的下一步交还给教师。

其五，参与计数装置。当举手、发言时长与在线活跃被当作指标，课堂会生产大量可见行为；这些行为若不改变后续，只是参与的产量，且会占用那些确实值得续发的少数差异所需的时间。

其六，粒度装置。课程被切成过细的知识点，每点配固定例题与测验。学生即使形成跨点联系也难以改变下一单元，因为组织的最小工作单元不是问题，而是预先定义的内容块。

其七，合法化装置。最终解释权与风险责任集中在教师，学生产物只有被教师重新表述后才获得地位。这可以控制错误，同时使边缘学生的产物在合法化之前消失。

其八，平台装置。数字平台按预设题库与算法派发下一题。即使系统自适应，它也可能只根据对错与速度调整难度，而不读取产物中的新解释、反例与问题。个性化回答"哪条现有路适合这个人"，续发回答"这个人的产物是否长出了一条原来不存在的下一步"。

其九，即时解答装置。生成式模型直接给出下一步、重写答案或完成推理，任务连续推进，而推进者始终是外部。Bastani 等（2025）的现场实验显示，无护栏工具在练习期提高表现，撤走后使用者可能低于从未使用者；加护栏可缓解。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**。这与本文一致——护栏改的是帮助形态，没有改归类的时刻。

九条共享同一个方向：把 w 压短。而它们在组织里的合理性也共享同一个来源：短的 w 使每一轮可结算、可记录、可审计、可交接。**这不是九个坏习惯，是一套自治的工程**。改其中任何一条而不改归类的时刻，其余八条会把它补回来。

十、被强制的不一致：形成性评价

上一节论证的是结算装置成套存在。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洽——不是本文批评它，是它自己的两条核心处方在同一套操作里互相拆台。

形成性评价这一行同时持有两条命题。

甲：证据必须改变后续教学。这是它的定义。若收集的证据不用于调整接下来的教学，那只是测验，不是形成性评价。这一条不是它的推论，是它据以与总结性评价划界的那条线。

乙：反馈应当及时。这是它最被广泛传播的处方之一，也被大量元分析强化：反馈要具体、指向下一步、并尽快给出，延迟的反馈效力较低。

在 $w > m$ 之下，甲与乙不能同时最大化。乙缩短 w ：及时的反馈把产物尽早放进一个类目（对、错、这一档、还差什么），而按第三节的定义，放进类目就是终止追问。甲要求续发率大于零：产物必须活过 m 才可能改变后续。**越彻底地执行乙，甲越难兑现。**

这一行并非没有察觉。它现有的出路是把反馈与评分分开——只给评语不给分数，或先评语后分数。但这条出路是循环的：要判定一次评语是“仍未归类”还是“已经归类”，需要一个判据，来分辨这次回应是把产物放进了已有类目，还是让它保持开放。而那个判据只能是**下一轮是否因它而改变**，也就是本文的变量。一句“你这里的推理很清楚，下次注意单位”没有分数，却已经把产物归入了“基本正确、有小疏漏”这一档，追问随之终止；一句“你和他的结论相反，能不能各自说说在哪一步分开的”没有评价，却把产物留在了开放状态。两句在形式上都是“不给分的反馈”，在本文的变量上相反。

因此：**形成性评价要让甲成立，必须先承认本文的变量；一旦承认，乙就不能以“越及时越好”的形式成立**——它必须被改写为“越及时越好，但归类不得早于 m ”，而这个限定条件在这一行现有的任何一份处方里都不存在。

这一节的主张有明确边界，也留了退路。本文不主张这一行的从业者不懂这件事；“反馈不给分”这一整支实践正是这种意识的产物。本文主张的是，在不引入“归类时刻”这个变量的条件下，这种意识**在形式上无法被兑现**——判定需要那个变量。这一主张可被推翻：若有人给出一个不依赖“下一轮是否因它而改变”的判据，能够稳定区分一次反馈是归类还是保持开放，本节即告失败。

同一动作可以对第二个行当再走一遍，且更短。生成式模型教育产品的验收普遍以即时正确率、响应速度与使用时长为准，而这行业在材料中反复引用的证据恰好表明这三个数与撤除后的能力可以反号。差别在于，形成性评价那一处的不一致出现在两条学术处方之间，这一处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。

十一、外部划界：规范全部满足而现象仍不发生

"学生产物应当决定下一步"这句话并不新。至少有三支传统把它写在了自己的纲领里，本文必须逐支给出判决性分离，否则只是换名。

其一，响应式教学。这一支主张教学应当从学习者已有的想法出发，教师的专业动作是听懂学生正在想什么并据此决定接下来做什么。它的核心判据几乎就是本文的主张。

其二，协商课程与生成性课程。这一支更彻底：课程不是预先写好再执行的，而是从儿童与学生的探究中长出来的；教师提供材料与结构，方向由探究决定。

其三，知识建构。这一支要求学生对共同体知识的长期进步承担责任，把想法当作可以持续改进的对象；配套的技术环境把学生贴出的表述做成后续探究可以接上的单元（Scardamalia 与 Bereiter, 1991）；共享认识论能动性研究进一步讨论群体如何共同形成知识对象与探究方向（Damşa 等, 2010）。

三支都比本文更早、更完整地主张了"学生产物应当决定下一步"。分离线不在主张，在**分析层次**：三支讲的是**规范与设计意图**，本文讲的是**时间条件**。三支都告诉课堂应当怎样对待学生的想法，没有一支给出"想法要活多久才可能被接手"这个量，因此也没有一支能回答：如果一间课堂完全接受这些规范，它的续发率会不会自动大于零。

本文的回答是不会，而且给出了判决性观察：

> 一间完全采纳响应式教学的课堂——教师全程跟着学生的想法走、从不预设路径、每一次发言都被认真对待——若每次学生发言之后教师立即给出评价性回应（哪怕是赞许的、开放语气的），则 w 极短，本文预测其续发率接近零。

三支会把这样的课堂判为高质量样本；本文预测它落在靶格里。这个反例若实测存在，本文对三支有增量；若不存在——即凡是采纳这三支规范的课堂， w 自然就长——那么"归类时刻"不是一个独立变量，本文被三支吸收。这条对照可以用现有课堂录像回溯编码，不需要新干预。

另有一条形式同构须点名：**累积优势**。它与本文共享"产物成为后续条件"这一形式，但对象不同——它讲个体之间差距的扩大，本文讲轮次之间因果力的传递；且它没有诱导期，也没有归类时刻。两者可以同时成立。

还有一支必须点名的是**生产性失败**。它主张先让学生生成不完善的方案，再进行整合教学；后来的工作把"生产／非生产"与"成功／失败"做成两轴，其中"非生产性的成功"一格与本文的靶格形状相近

（Kapur, 2016）。分离线在第二轴的性质：那一格的判定靠延迟结果回溯，是一个**结果变量**；本文的靶格由 w 与 m 判定，是一个**过程变量**，当场可编码，不必等到后测。因此本文能在结果出来之前指认靶格，而那一支不能；反过来，那一支能判定的"非生产"包含了本文不覆盖的情形（产物活过了 m 、也被接手了，但方向错误）。两者的交集不小，各自的余部也不小。

十二、站内划界：与本栏两篇先行研究的分工

本栏此前已有两篇处理同一问题域，且都用了同一组学科。不作分离，本文读来就是自我重排。

与《方程两边都有它》（本栏之五十四）。那一篇问的是**本轮内这一步归谁**：一个可验收产物由学习者独立完成还是由外部代做，以及首次选择的可逆性如何决定哪条通路锁死。分离线是清楚的：那一篇的变量全部落在**本轮之内**，本文的变量落在**本轮末端到下一轮之间**。两者正交，且本文预测二者**相乘而非相加**——一次由学习者独立完成、且可撤回的首次作答，若在长成之前被归类，它所保住的信息同样进不了下一轮。判决性对照：一个严格执行“分岔步交还学生、首次作答可撤回一次”、而每次作答后立即评分的班，那一篇预测其本轮质量提高，本文预测其续发率仍接近零；两个预测可以同时被检验。

与《每一轮都从新鲜表面开始》（本栏之五十五）。这一篇离得最近，必须最诚实地处理。那一篇命名未结转中间体：每轮真实生成、真实占位，而结算规则只认起点态与终点态，故不设字段；读数是结转率——下一轮的开头是否接上了上一轮留下的某个具体未完成处。

分离线由化学自己给出，这是本文能拿到的最干净的一条。第八节引过那条警告：S 形曲线的滞后段可以来自自催化，**也可以来自某种中间体的积累**；两者在曲线上不可分，区分它们必须做加料实验。**那一篇讲的正是中间体积累那一支，本文讲的是自催化那一支**。两篇因此不是同一个说法的两个名字，而是同一条曲线的两种可能来源——而这一支早在一百年前就知道要把它们分开，并给出了分开的实验。

三条具体分工随之确定。**对象不同**：那一篇的对象是未完成物，本文的对象是任何产物（包括已完成、已答对的）。**读数位置不同**：结转率读在下一轮的开头， w 读在本轮的末端。**因果链位置不同**：本文管产物能不能活到轮次交界处，那一篇管到了交界处有没有人接。判决性对照：一个把开场三分之一严格用于接续上节未完成处、而在本节末端仍即时归类所有新产物的班——那一篇的结转率高，本文的 w 短；本文预测这样的班第二个月的链条仍会断，因为可供结转的东西正在本节末端被消灭。这个预测那一篇作不出，也不与它冲突。

并须记下一处**继承**。那一篇已经写下一句关键的观察：形成性评价必须把中间态归入已知类目才能给出反馈，而那正是一次温和的清零。本文第十节所做的，是把那句话形式化为 w 与 m 的关系，并据此推到内部矛盾。**这一步的起点记在那一篇账上**，本文只承认自己做了形式化与推进。

十三、假设与判决变量

本文的承重命题可写成三重否定：课堂中可迁移学习稀少，不只是学生参与不足，不只是任务难度失配，也不只是评价指标发生偏移；它是**产物在长到能够接手下一轮之前就被归类**，加上外部启动足够可靠以致这条通路即使存在也不显形。

编号	假设	关键判决变量
H1	窗口效 控制先备水平、教师、内容难度、教学时间与当轮成绩后，未被归类	$w \rightarrow$ 续发率 $\rightarrow$ 延迟
应	窗口 w 正向预测续发率与延迟迁移	迁移

编号	假设	关键判决变量
H2 门槛而非线性	w 与续发率的关系应在 m 附近出现拐点, 而非全程线性; w 远小于 w/m 比与续发率的 m 时增加 w 收效甚微	函数形状
H3 正交性	w 与教学法标签、参与次数、当轮正确率的相关应显著低于其与续发率的相关; 否则它只是既有指标的改写	w 与既有指标的相 关矩阵
H4 加料判据	注入上一轮产物的轮次, 到达下一次续发的时间短于注入等价外部材料的轮次	注入来源 (本轮内/ 外) $\times$ 到达时间
H5 可观测性	在停止外部启动的轮次上, 高 k_1 组与低 k_1 组的续发速率应当接近; 若差异反而扩大, 则差的是能力不是可观测性	$k_1 \times$ 停启动条件
H6 归类与反馈可分	同等反馈量下, 不终止追问的回应比终止追问的回应产生更高续发率; 差异不必表现为更高的当轮成绩	回应类型 $\times$ 反馈量 匹配
H7 装置叠加	九种结算装置对 w 的作用同号; 移除其中一种而不改归类时刻, w 的净变化应接近零	单装置移除 $\times w$
H8 诱导期	续发型设计的效应在短于 m 的评估窗口内不出现甚至为负, 在长于若干个 m 的窗口内出现	评估窗口长度 $\times$ 效 应方向
H9 形成性评价的取舍	在同一批教师中, 反馈及时性与续发率呈负相关; 且这一负相关在把"不给分的评语"细分为归类型与开放型之后, 只出现在归类型一侧	及时性 $\times$ 回应类型 $\times$ 续发率

判决逻辑分三组。H1、H2、H3 检验 w 是否是一个独立且有门槛的变量; H3 是最凶的一条——它要求本文的读数与既有指标**不相关**, 这是自设的刀口。H4、H6、H9 检验"归类时刻"这一处的可操作性, 其中 H4 是唯一一条正面判据, H9 直接对准第十节的强制不一致。H5、H7、H8 检验可观测性与叠加结构。

三层可分别接受检验。第一层若失败—— w 不预测续发率——核心概念应被撤回。第二层若失败——归类与反馈不可分——本文的处方失效, 但辨别格仍可保留为诊断工具。第三层若失败, 两速率模型退回为一个说明性的类比, $w > m$ 那一半仍可独立成立。

十四、研究设计与操作化

理论若只能解释已知现象, 仍可能是一个新名字。本文的三个核心量必须能被独立测出, 且要能在既有课堂里测。

主设计是同一制度环境内的多班级随机试验: 同一学校、同一课程标准、相近师资, 至少十八个平行班, 每条件不少于六个班, 持续十二周。四类条件的差别被刻意收窄到一个参数上。

- **A 组**: 常规讲解与练习, 即时反馈照常。
- **B 组**: 延迟测验组。测量时点后移, 其余不变。用于把本文与"延迟测验"分开。

- **C 组：延迟归类组。**反馈量、反馈时机与教师投入与 A 组完全匹配，唯一差别是产物生成后的一段时间内，教师只作不终止追问的回应（指出冲突、并置分歧、要求说明可检验后果），过了该时段再作评价性判定。时段长度按预试估出的 m 设定。
- **D 组：停启动组。**每周有固定比例的轮次不由教案、题库或平台发起，其起点必须取自上一轮的某个可指认产物；其余一切照常，教师照常讲解、照常提供材料、照常回答问题。

A 与 C 的对照是本文的核心检验，因为它把反馈的量与时机全部固定，只留归类时刻一个变量。若两组续发率与延迟迁移均无差异，第六节的推论四为伪，本文的处方失效。A 与 B 的对照回答另一个问题：本文是否只是延迟测验的换名；若 B 组与 C 组等效，则“延迟”这一件事本身足够，归类时刻不是独立变量。C 与 D 的对照分离两个参数： w （归类时刻）与 k_1 （外部启动占比）。

加料实验作为独立的低成本子研究，可以脱离主设计单独跑。两个平行班，同一教师，同一节课：一班的下一轮起点注入上一轮的一个具体学生产物，另一班注入同等长度、同等难度的教师准备材料；测到达下一次续发的时间。这个子研究一节课就能做，不需要新设备，也不需要改变课程，且它检验的是 H4 这条唯一的正面判据。

数据分五层。基线知识与动机；课堂录像、屏幕日志、聊天记录与草稿，用于标记产物生成时刻、归类时刻与轮次交接；即时产物质量；四十八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；管理环境，包括进度要求、评价权重、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，用于检验编码是否把一次开放型回应误读为归类。

w 的编码有一条须写死的规则：**归类时刻取“第一次终止追问的回应”，不取“第一次给出分数”**。教师说“对了”就是归类，即使分数在一周后才出现；教师说“你和他冲突在哪一步”不是归类，即使他随后在自己的本子上记了分。这条规则是全部编码里最容易出错的一处，必须由两名以上盲评编码者独立作答并报告一致性，一致性低于阈值时该批数据作废而不是取平均。

m 由 C 组的预试估出：人为把归类推到很后面，观察续发事件出现在产物生成后多久，取分布的下四分位作为该任务类型的 m 下界。这个估法有一处循环风险——用一个理论所预言的现象去估该理论参数——因此必须预注册估计规则，并在主分析中报告 m 取不同值时结论是否稳健。

最低成本的先导研究不必重组课程。多数学校已有课堂录像、平台日志、教案版本与作业记录；在六个以上同质班级中回溯编码 w 与续发率即可。若同一教师的不同班级在活动水平相近时存在 w 的差异，且该差异先于延迟迁移的差异出现，就得到第一层支持。若不存在，应优先怀疑本文的机制，而不是把失败归因于样本不够。

十五、干预：延迟的是归类，不是反馈

第六节的推论四把干预放在一个此前没有被指认的位置。这一节把它写成可执行的形式，并逐条说明它与既有处方的差别。

第一条，把归类推到 m 之后，而不是把反馈推到 m 之后。这是全部处方的核心，也是最容易被执行错的一条。教师不需要沉默，不需要减少回应，不需要装作没看见错误。他需要做的是在一段时间内只使用不终止追问的回应形式：指出冲突而不判对错（“你和他的结论相反”）、要求说明可检验后果（“如果这样，那接下来应该看到什么”）、并置两个说法而不宣布哪个成立、追问一个步骤的理由。这些全部是反馈，且都不把产物放进已有类目。

第二条，给出一条可操作的判断句。教师在回应之前可以问自己一句不含判断词的话：**说完这句之后，这个学生还需不需要接着想？**需要，则未归类；不需要，则已归类。这一句可以当场用，不需要培训体系。

第三条，在某些轮次上不由外部发起。这与第一条是两个参数。做法是每周固定若干轮次，其起点必须取自上一轮的某个可指认产物，且教师须公开说明这一轮为什么由它触发——学生需要知道不是“教师临时换了题”，而是“上一轮的某样东西改变了后续”。这一条改的是 k_1 ，不是帮助的量。

第四条，建立因果型记录。普通学习档案记录学生做了什么、得了多少分；因果型记录增加一栏：这个产物后来改变了哪一项任务、材料、顺序或评价依据。只有进入后续因果链的产物才记为续发，不是所有发言都记功。这一栏同时是 w 与续发率的原始数据来源。

第五条，把错误从负项改为可选择的启动物。错误不因错误自动获得续发权，也不因错误自动被消除。教师需判断它是否显出稳定模型、普遍混淆或有价值反例。被选中的进入下一轮比较与验证；无价值或高风险的及时终止，并说明终止理由。**透明的终止不同于无声归类：**前者承认产物曾取得地位，只因明确约束而结束；后者把一切重新归入原计划，仿佛没有发生过。

第六条，为生成式模型设“先产物、后回应、再续发”的三步。系统先要求学生提交假设、错误诊断或下一步预测；随后只指出冲突而不给答案；再由学生据冲突修改并决定下一题；系统必须引用该产物说明路径为何改变。撤去学生产物后系统若仍给出同样任务，就不计为续发。要紧的是第二步：**默认关掉即时判定**，把“对/错”的显示延后到该轮结束，这在多数平台上是一个配置项。

第七条，管理层只改三件事，不加考核。为归类的延迟留出课时缓冲；把“覆盖了多少内容”与“有多少轮次由上一轮产物发起”分开报告，禁止用前者替代后者；采购与平台验收增加“即时判定可否关闭”与“产物来源可否追溯”两项。不设每节课的续发次数配额——理由见第十九节。

七条里只有第六条有实施成本，其余六条是规则与配置的改变。这一点本身是可检验的：若在完全不增加教学时间与资源的条件下，仅按第一条与第三条运行的班级与对照班无差异，本文的处方失效。

十六、五类场景与其中两处收窄

数学。学生通过模仿得到正确结果，教师进入下一类题，当轮成绩很高， w 以秒计。延迟归类的做法是：一道题给出答案之前，先收齐三种不同的做法或三种不同的卡点，并置展示而不宣布哪个对，隔一段时间再处理。收窄之处也在这里：基本符号约定、运算安全规则与已证明的逻辑禁步不能因“学生发起”而无限协商，这些位置的 w 应当短。成熟设计不是延长每一个错误，是识别哪些错误显出生成结构。

科学实验。关键对象是异常结果。常规实验要求学生得到预期曲线，异常数据被解释为操作失误并从报告中删除——这是 w 接近零的极端形式。延迟归类的做法是异常结果先不判定为失误，要求学生说明"若它是真的，接下来应该看到什么"。**第一处真收窄在这里：**危险化学品、病原体、强电与高压设备的操作不能为提高续发率而延迟判定，安全上的对错必须立即归类。可行做法是把归类的延迟限制在数据解释侧，而不是操作侧——异常现象可以决定下一轮的分析、模拟或低风险验证，不能改变安全程序。

语文与历史。开放答案多， w 却未必长：教师允许多种解释，随后逐一表扬并继续按篇目推进，表扬本身就是归类。判据是某种解释是否改变了后续的文本选择、证据要求或评价维度。这一类还有一个特有风险：语言表达强、熟悉学术话语的学生更容易让自己的解释活过 m 。匿名文本、书面问题池与延迟回应可以分离表达优势与认知价值。

案例教学。学生提出多套方案，教师最后公布专家方案——这是一次教科书式的末端归类。延迟归类的做法是先让两套相反方案各自说明"在什么条件下你会承认自己错"，再公布。**第二处真收窄在这里：**在真实临床与工程操作中，最终决策权不能因教学实验被模糊；续发应主要发生在模拟、复盘与方案比较阶段。

生成式模型辅助。系统按学生错题自动推送下一题，看起来已经"由产物决定后续"。但若算法只读取正确率、速度与题型标签，它响应的是预设特征，不是产物中的新东西；而且它的即时判定把 w 压到接近零，是九种装置里最有效的一种。加料实验在这一类里最容易做：把上一轮某个学生的具体表述注入下一轮的提示，与注入等价的系统生成材料作对照。

五个场景的答案不应被预设为相同。至少两个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。

十七、适用边界与三门学科的反向约束

其一，本文不是反对及时反馈。这是最容易被误读的一处。及时反馈在纠正程序性错误、防止错误固化、维持学习者信心上都有独立价值，本文不主张削减它。本文主张的是及时反馈与产物续发是两个目标，在同一轮里互相削减，因此需要按任务目标分配：以熟练为目标的任务， w 短是合理的；以概念重组、建模或问题生成为目标的任务， w 必须长于 m 。把本文读成"反馈越慢越好"是把一个条件命题读成了一个方向命题。

其二，反应动力学削掉了"续发越多越好"。传播可以放大副反应，自催化可以失控，抑制与终止是选择性与安全性的必要机制。因此续发权的价值不在最大化，而在让经过检验的差异传播、让高风险与支配性的差异被及时终止。这条反向约束落到了假设层：本文不预测续发率与结果单调相关，而预测条件性关系。

其三，管理学削掉了"制度应当退出发起"。组织协调、安全责任与公共标准决定了某些任务必须集中发起。更要紧的是这一支的反应性论证同样适用于本文自己的读数： w 一旦进入考核，就会被优化成一个数而不是一件事。这条约束直接决定了第十五节第七条为什么不设配额。

其四，教育学要求为初学者保留充分指导。延迟归类不等于最小指导。教师可以先提供范例、支架、词汇与基本程序，再让由学生产物生成的差异进入后续。分界不是"教师控制或学生自由"，是**外部支持是否最终允许产物活过 m** 。

其五， m 的可估性是本文的命门。若 m 在同一任务、同一批学习者上分布极散以致给不出可用下界，则 $w > m$ 虽然定义上成立，实践上不可操作。本文对此没有解决方案，只有一条退路：即使 m 不可精确估计， w 本身仍是一个可测且与既有指标正交的量，辨别格与 $H3$ 仍可独立检验。

其六，结清态是轮次属性，不是人格属性，也不是教师属性。教师与学生都处在制度约束之中。把这个概念用于责备"急于给答案的教师"，会把系统问题重新个人化，并刺激新的表演。

十八、竞争解释、证伪条件与一条写死日期的赌注

最强的竞争解释是时间投入。延迟归类意味着一节课处理的内容变少，优势可能全部来自"在同一内容上花了更多时间"。因此 C 组必须与 A 组匹配总教学时间、教师接触时长与练习剂量，并在共同核心内容上设置等量测验，同时报告覆盖损失。若优势只能通过大幅减少课程目标获得，本文仍有局部价值，不能作为一般原则。

第二条是心理安全。学生只有在不担心羞辱时才会让产物长出来；高续发课堂可能只是心理安全更高。分离线是：在心理安全相近的班级中， w 是否仍预测续发率。若心理安全能完全解释，本文应退回为它的一个过程指标。

第三条是教师适应性。优秀教师本就会跟着学生调整， w 可能只是教师能力的代理。应使用同一教师的交叉设计、课前锁定路径与教师固定效应，比较同一教师在不同规则下是否出现不同的 w 与续发率。

第四条是新奇效应。学生发现自己的问题会改变课程，可能因新鲜而更投入。研究需延长到足以让新奇消退，并测量动机变化。这一条与推论二（诱导期）互为镜像，因此必须同时报告效应的时间曲线：新奇效应应先高后低，诱导期效应应先低后高。**两者方向相反，这使它们可以被分开**——这是本文能给出的最干净的一次分离。

第五条是选择偏差。能生成可续发产物的学生可能本就更强。事件级分析应控制产物质量、先备水平与表达可见度，并专门比较"高质量但被立即归类"与"中等质量但活过了 m "两类的后续命运。

机制证伪条件（合格的证伪句必须同时排除上述五条）：

其一，控制总教学时间、教师效应、心理安全、先备水平、产物质量与内容覆盖后，若 w 与续发率、延迟迁移之间不存在剂量—反应关系，核心机制为伪。

其二，若 A 组与 C 组（反馈量与时机全部匹配，只改归类时刻）在续发率与延迟迁移上均无差异，推论四为伪，本文的处方失效。

其三，若 B 组（延迟测验）与 C 组等效，则归类时刻不是独立变量，本文可被"延迟"这一件事吸收。

其四，若加料实验中，注入上一轮产物与注入等价外部材料对到达下一次续发的时间影响相同，H4 这条正面判据失败，本文只剩证伪条件。

其五，若 w 与教学法标签、参与次数或当轮正确率高度相关（H3 失败），本文的读数只是既有指标的改写。

其六，若在停止外部启动的轮次上，高 k_1 组与低 k_1 组的续发速率差异不缩小反而扩大，两速率模型的可观测性解释为伪。

其七，若有人给出一个不依赖“下一轮是否因它而改变”的判据，能稳定区分一次反馈是归类还是保持开放，第十节的强制不一致主张失败。

其八，若结清态只在生成式模型场景出现，在传统讲授、同伴领做与实验步骤中无法复现，理论范围必须缩小。

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：在总教学时间、反馈量、反馈时机与任务难度全部匹配的条件下，**仅改变归类时刻**即产生无支持延迟迁移的差异，且方向为延迟归类组更高，则本文核心判断获得初步支持。评估窗口须不短于该任务类型 m 估计值的六倍——这一条是为了堵住推论二被用作万能退路。

以下明确**不算命中**：仅延长测验延迟；仅减少反馈总量；仅撤除支架或降低难度；仅改变模型提示形态或加装护栏；仅报告参与度、满意度、当堂成绩或课堂开放度评分；机构自行声明采用了形成性评价、探究式或学生中心；把延迟归类与其他多项改革打包实施而无法分离贡献。若符合上述条件的多项研究均未发现增量效应，本文应被放弃或降格为描述性工具。

十九、反噬预言与伦理限制

本文一旦进入学校制度，最省事的实施方式很可能摧毁它。

第一种是配额。学校要求每节课至少出现三次“由学生产物触发的路径改变”。教师随即学会在教案里形式化地引用学生发言，平台自动生成“由学生问题触发”的标签，续发率表面上升，而发起权仍由指标预先决定。学生产物被用来证明制度达标，不是用来改变制度。

第二种是延迟表演。教师学会在给出判定之前停顿一段可见的时间，形式上满足“延迟归类”，而那段时间里什么也没发生——产物没有被要求展开，只是被搁置。这一种比第一种更难查，因为 w 这个数在它上面完全达标。**唯一的对策是同时记录那段时间里产物有没有变形**：一个被搁置的产物在归类时与生成时一字不差，一个被延迟归类的产物应当在此期间长出可检验后果、边界或与他人主张的关系。

第三种是错误浪漫化。为了反对答案装置，课堂可能延长本应及时纠正的错误，把挫败与混乱当作深度。反应动力学的选择性与终止机制正是为此写进第十七节的。

第四种是责任转嫁。学校借"学生发起"把课程质量、资源不足与支持责任推给学习者。续发权只能是在充分资源与专业保护下获得的位置，不能成为撤走教学的理由。

四种都不能被程序彻底消除。本文能给的只有底线四条：读数指向任务位置而非个人；不得为提高读数在安全关键环节延迟判定；**验收 w 必须同时验收产物在该窗口内是否变形**，否则等于奖励延迟表演；任何基于这些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。

公平问题须单独说。谁的产物更容易活过 m，受语言能力、身份地位、教师期待与同伴网络影响。一个高续发课堂若总由少数强势学生发起，只是把教师中心换成了地位中心。因此研究与实践应同时报告续发来源的分布、被忽视产物的类型与不同群体的获得概率，并为低可见度的表达提供书面、匿名与延迟提交的通道。数据伦理也不能被"研究学习"豁免：课堂录像、平台日志与产物追踪涉及未成年人的行为数据，应坚持最小收集、明确用途、可撤回同意与去标识化；对产物的因果追踪不应演变为持续监控，更不能把尚未成熟的错误永久化为个人标签。

二十、自反

按本文自己的判据，这篇论文正处在靶格里，而且处得比多数文本更彻底。

它给出的是一套已经归类完毕的东西：概念有定义，量有符号，关系有方向，命题有编号，反例有分离线。读者接住的是一个成品。**而按第三节的定义，一个完整的、自洽的、可复述的文本，本身就是一次归类**——它把一批可能长成别的样子想法固定在一个类目里，并因此终止了对它们的追问。本文的 w 在发表的那一刻就归零了。

这不是一句谦辞，它有一个具体后果：本文最可能被接住的是"结清态"这三个字，而不是 w 与 m 的关系；最可能被引用的是那句"课堂的一切结算都发生在诱导期之内"，而不是第八节那条加料实验。名字比关系更容易归类，因此名字传得更快——这正是本文所描述的机制在自己身上的一次运行。

本文能为自己配的唯一装置，是把可以让自己失败的东西写得比结论更具体：第十一节那条判决性观察、第十二节与两篇先行研究的双向对照、第十八节那八条证伪与写死下界的赌注。它们的作用是让读者可以先自己判一个案例，再回来看本文怎么判——即在接住这个成品之前，先让自己的判断长一会儿。

还有一条更不利的自反。本文第十节论证形成性评价在不承认这个变量时无法自治，而本文自己提出的读数同样会被做成仪表盘，第十九节已经预言了它会怎样被做坏。本文没有理由认为 w 比及时反馈更耐受制度化。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，这使失败可以被指认——而那正是"反馈应当及时"这条处方当年没有做的事。

二十一、结论

课堂上可迁移的学习稀少，不是因为课堂没有产物。恰恰相反，课堂高效地生产了答案、笔记、作业、项目与分数。问题在于这些产物在长到能够接手下一轮之前，就已经被判为对错、分数、标签或已完成；而课堂的外部启动足够可靠，以致这条通路即使存在也不显形。

三门学科在这里各出一份不可替代的力。教育学解释产物如何形成，也提供了那条与自己相冲突的处方：证据必须改变教学，反馈应当及时。管理学解释结算为什么成套存在，并把这条约束同时施加给本文自己的读数。反应动力学给出两条最承重的形式约束：**资格不是产物一生成就有的属性，是一个需要时间达到的状态**；以及**这个现象由动力学签名定义而不由机理定义**，因此任何基于课堂形态的判断——像不像学生中心、活动多不多、话语开不开放——在原理上都不足以裁决它。

这一理论把课堂评价的核心问题从"学生答对了吗"改为"这个产物在被判定之前活了多久"；把教学改革从"活动更多""难度更高""测验更晚"改为"归类更晚"；把管理从即时产出单账本改为产物账与交接账双账本；把技术验收从即时正确率改为"即时判定可否关闭"。其中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。

课堂不需要把发起权永久交给偶然性。它需要的只是：在每一轮结束之前，让那个刚长出来的东西，多活一会儿。

注释

一、本文所用"自催化""诱导期""滞后段""速率常数""加料实验"等术语取标准化学动力学的通行含义，用于提供关于时间结构与可观测性的形式约束，不作为关于神经过程的经验主张。第五节的两速率模型借用的是两步动力学模型的结构，不主张课堂过程满足该模型的全部假设。

二、第十节的强制不一致主张限于"反馈应当及时"与"证据必须改变后续教学"两条处方在同一套操作里的取舍，不针对形成性评价这一支的全部工作；该支内部关于"反馈不给分"的实践正是本文所指问题的自觉产物。

三、第十二节所述两篇先行研究均为本栏已刊文章，其判断与读数的表述以原文为准；本文只对分工线与判决性对照负责。

四、本文为理论论文，不报告作者新采集的实验数据。涉及实证效果、阈值与研究方案的内容需在正式投稿前逐项核验，并按目标期刊规范补充伦理审批与数据计划。

参考文献

- 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.
- Bentea, L., Watzky, M. A., & Finke, R. G. (2017). Sigmoidal nucleation and growth curves across nature fit by the Finke–Watzky model of slow continuous nucleation and autocatalytic growth. *The Journal of Physical Chemistry C*, 121(9), 5302–5312. doi:10.1021/acs.jpcc.6b12021.
- Bissette, A. J., & Fletcher, S. P. (2013). Mechanisms of autocatalysis. *Angewandte Chemie International Edition*, 52(49), 12800–12826. doi:10.1002/anie.201303822.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. doi:10.1080/0969595980050102.
- 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.
- Damşa, C. I., Kirschner, P. A., Andriessen, J. E. B., Erkens, G., & Sins, P. H. M. (2010). Shared epistemic agency: An empirical study of an emergent construct. *Journal of the Learning Sciences*, 19(2), 143–186. doi:10.1080/10508401003708381.
- 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.
- Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1), 94–118. doi:10.2307/3556620.
- 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.
- Hanopolskyi, A. I., Smaliak, V. A., Novichkov, A. I., & Semenov, S. N. (2021). Autocatalysis: Kinetics, mechanisms and design. *ChemSystemsChem*, 3(1), e2000026. doi:10.1002/syst.202000026.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487.
- Jones, E., & Nimmo, J. (1994). Emergent Curriculum. *National Association for the Education of Young Children*.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. doi:10.1080/07370000802212669.

- Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, 51(2), 289–299. doi:10.1080/00461520.2016.1155457.
- 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.
- 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.
- Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319–340. doi:10.1146/annurev.so.14.080188.001535.
- Liu, Y., et al. (2023). A nanoscale MOF-based heterogeneous catalytic system for the polymerization of N-carboxyanhydrides. *Nature Communications*, 14, 5598. doi:10.1038/s41467-023-41252-3.
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. doi:10.1287/orsc.2.1.71.
- Merton, R. K. (1940). Bureaucratic structure and personality. *Social Forces*, 18(4), 560–568. doi:10.2307/2570634.
- Robertson, A. D., Scherr, R. E., & Hammer, D. (Eds.). (2016). *Responsive Teaching in Science and Mathematics*. Routledge.
- Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building. *Journal of the Learning Sciences*, 1(1), 37–68. doi:10.1207/s15327809jls0101_3.
- 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.
- Watzky, M. A., & Finke, R. G. (1997). Transition metal nanocluster formation kinetic and mechanistic studies. *Journal of the American Chemical Society*, 119(43), 10382–10400. doi:10.1021/ja9705102.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. doi:10.1207/s15430421tip4102_2.

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