

接住它的那一层，决定它还能不能走

论学习产物的采纳层级及其持有者，及为什么最积极的响应常常正是消音的方式

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2.0万字 · 27页

发表 2026年8月3日

摘 要

课堂拥有讲授、练习、提问、反馈与评价，却仍普遍出现当堂会、换题不会。六类既有解释共享一个未被说出的前提：学生的解释、错误、问题与策略若被听见、被回应、被使用，就算被采纳了。本文指出**采纳不是二值的，它有层级**——同一份产物可以被接在表征层、路径层或规则层，而三层的持有者分别是学生、教师与组织。承重判断由此得出：真正的学习之所以稀缺，不是因为学生产物没有被采纳，而是因为**大量产物被采纳在低于其诊断价值所在的那一层**；采纳发生了、响应发生了、证据被使用了，而那份产物原本要挑战的东西恰恰因此不再被挑战。本文称之为**降层采纳**。由此三家最接近的说法从竞争者变成特例：生成性学习覆盖表征层（持有者是学生）、形成性评价覆盖路径层（持有者是教师，故它在设计上就会把规则层产物接在路径层）、双环学习覆盖规则层（持有者是组织）——每一家看不见的，恰好是它的持有者管不着的那一层。化学在此不作装饰：最强的吸附位点把分子接住却不放行，**接住恰是它退出转化的方式**；而工业上为使批次可比而周期性再生，说明**轮的边界不是自然的，是被制造出来的**。判据一句不含情态词：**上一轮那个产物，改变了下一轮的哪一层——它换的是一道题，还是换了「什么算数」？**配套读数**采纳层距**与既有指标正交：一个反馈最及时、诊断最精准的智能课堂，层距接近最大值，因为它的全部接口都在路径层，它能做的唯一动作是换一道题。

这一篇由哪三个学科的理论体系撞成

第一家把学习的条件放在学习者内部的表征重组，主张没有已有表征的经过，输入不能成为知识；第二家主张在场不等于转化，产物能否进入后续路径是一个必须单独询问的系统性质，而最强的接住恰恰使转化终止；第三家主张局部变化只有被编码进例程、规则与资源接口才成为组织能力，而绩效信号、探索与利用的分配、权限边界三者合起来，使组织对高层问题的标准回应必然是一个低层动作。三家不能同为真：若第一家对，改善接口不必要；若第二家对，教育学与管理学各自的处方都停在增加回应量这一侧而没有触及层级；若第三家对，教师的意愿与学生的表征都在权限结构的下游。三家对同一件事给出互不相容的归属——一名学生的规则层反例被教师用两道对比题回应，按形成性评价判为高质量，按本文判为降层采纳。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 · 表征重组、生成性学习与形成性评价

[BLACK & WILIAM, Assessment and Classroom Learning \(<i>Assessment in Education</i> 5:1, 1998\)](#); 另见 POSNER 等, [Accommodation of a Scientific Conception \(<i>Science Education</i> 66:2, 1982\)](#); CHI & WYLIE, [The ICAP Framework \(<i>Educational Psychologist</i> 49:4, 2014\)](#); FIORELLA & MAYER, [Eight Ways to Promote Generative Learning \(<i>Educ Psychol Rev</i> 28, 2016\)](#)

主张学习必须经过学习者已有表征，且证据应被用于调整正在进行的教与学；而这一支自己确立的持有者边界表明，教师的当堂权限止于路径——因此它在设计上就会把规则层产物接在路径层，这在它自己的框架内完全是成功。

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

化学 · 萨巴捷原理、位点封堵与稳态近似

[MEDFORD, VOJVODIC, NØRSKOV 等, From the Sabatier Principle to a Predictive Theory of Transition-Metal Heterogeneous Catalysis \(<i>Journal of Catalysis</i> 328, 2015\); 另见 GROPP, ROJAS-BUZO & BORDIGA \(<i>Chemical Reviews</i> 123:21, 2023\); IUPAC Gold Book: catalyst poisoning, autocatalytic reaction; BODENSTEIN 1913 \(稳态近似\)](#)

结合太弱活化不了、太强脱附不下来，活性峰值在中等结合强度处；由此得到本文的核心推论——**最强的位点接得最快最牢，而被它接住的分子退出了转化**。另一件材料是稳态近似成立的条件由工业周期性再生**制造**出来，说明轮的边界不是自然的。两件都在化学里回答别的问题，从没人拿它们当关于采纳的证据。

<https://doi.org/10.1016/j.jcat.2014.12.033>

管理学 · 单环双环、绩效反馈与探索利用

[ARGYRIS, Double Loop Learning in Organizations \(<i>Harvard Business Review</i> 55:5, 1977\); 另见 LEVITT & MARCH, Organizational Learning \(<i>Annual Review of Sociology</i> 14, 1988\); MARCH, Exploration and Exploitation \(<i>Organization Science</i> 2:1, 1991\); GREVE, Performance, Aspirations, and Risky Organizational Change \(<i>ASQ</i> 43:1, 1998\)](#)

主张经验须被编码进例程才成为组织能力，且低于期望时才触发搜索；三条机制合起来解释了组织为何系统性地把高层问题接在低层——低层采纳属于利用，代价可计、见效可测、当堂可验收；而这一支自己没有「产物本身的诊断层级」这个量，故它看不见「回应了但降了层」与「没有回应」的差别。

<https://doi.org/10.2307/2393591>

摘要.....	5
一、问题的错位：完成了教学，没有完成学习.....	6
二、什么算学习：从即时正确到跨轮次的续走.....	6
三、教育学的贡献与它的止步处	7
四、化学的机制约束：接住不等于放行，而轮的边界是被制造出来的	8
五、管理学的制度解释：为什么组织会系统性地把高层问题接在低层	9
六、三家共同的隐含安排	10
七、采纳的三层及其持有者	10
八、降层采纳：本文的承重命题	12
九、采纳层距：一个可清点、且与既有指标正交的读数.....	12
十、判据：一句不含情态词的问话	14
十一、二维判别格：识别最危险的"结算性成功".....	14
十二、四类课堂场景：同样的表现，何以导向不同的学习.....	15
十三、十个行当走一遍	16
十四、与最近的几种说法的分界	17
十五、与本栏近作的分界	19
十六、可证伪的六组命题	20
十七、两处反过来削弱本文的约束	21
十八、研究方案：从课堂印象转向层级与版本证据.....	22
十九、适用边界、可能反噬与伦理条款.....	23
二十、反身性检验：本文自己被接在哪一层.....	24
二十一、结论：从让学生正确，到让正确能够继续走.....	24

摘要

课堂拥有讲授、练习、提问、反馈与评价，却仍普遍出现“当堂会、换题不会，眼前对、过后忘”。既有解释分别归因于动机、教法、认知负荷、反馈质量、迁移条件与制度约束，但它们共享一个未被说出的前提：学生的解释、错误、问题与策略若被听见、被回应、被使用，就算被采纳了。本文指出，采纳不是一个二值事件，它有层级——同一份学生产物可以被接在表征层、路径层或规则层，而三层的持有者分别是学生、教师与组织。由此得到本文的承重判断：真正的学习之所以稀缺，不是因为学生产物没有被采纳，而是因为**大量产物被采纳在低于其诊断价值所在的那一层**；采纳发生了、响应发生了、证据被使用了，而那份产物原本要挑战的东西恰恰因此不再被挑战。本文把这种情形命名为**降层采纳**，并给出可清点的读数**采纳层距**。化学在此不作装饰：多相催化中最强的吸附位点会把分子接住却不放行，接住恰是它退出转化的方式；而工业上为使批次可比而周期性再生，说明**轮的边界不是自然的，是被制造出来的**。判据一句，不含情态词：**上一轮那个产物，改变了下一轮的哪一层——它换的是一道题，还是换了“什么算数”？**全文给出六组可失败的经验命题、一套以序列与版本为单位的研究方案、七条适用边界与伦理条款，以及一项写死日期的赌注。

关键词：采纳层级；降层采纳；采纳层距；层级持有者；结算性成功；形成性评价

Title: The Level at Which It Is Caught Decides Whether It Can Still Travel

Abstract: Classrooms may contain instruction, practice, questioning, feedback and assessment and still produce the familiar pattern of success in the moment and failure later. Existing explanations attribute this to motivation, pedagogy, cognitive load, feedback quality, transfer conditions or institutional constraint, yet they share an unstated premise: that a learner's explanation, error, question or strategy counts as adopted once it has been heard, answered or used. This paper argues that adoption is not binary but layered. The same learner product can be caught at the level of representation, of pathway, or of rule, and these three levels are held respectively by the learner, the teacher and the organisation. The load-bearing claim follows: real learning is scarce not because learner products go unadopted, but because a great many of them are adopted **one level below** the level at which their diagnostic value lies. Adoption occurs, response occurs, evidence is used — and precisely thereby, what that product was going to challenge is no longer challenged. The paper names this **level-lowered adoption** and gives a countable reading, the **adoption level gap**. Chemistry is not decorative here: in heterogeneous catalysis the strongest binding sites catch a molecule without releasing it, so that being caught is exactly how it leaves the reaction; and industrial regeneration between batches shows that the boundary of a round is not natural but manufactured. The operational question, free of modal words: **which level did that product change in the next round — did it change one exercise, or did it change what counts?** Six

falsifiable propositions, a sequence-and-version research design, seven boundary and ethics conditions, and a dated wager are given.

Keywords: adoption level; level-lowered adoption; adoption level gap; level holders; settled success; formative assessment

一、问题的错位：完成了教学，没有完成学习

当一节课有明确目标、完整板书、热烈互动、即时反馈与顺利达标时，人们很容易说学习已经发生。然而只要把观察时间从当堂拉长到数周，把任务从原题改为异质情境，把问题从"会不会答"改为"能不能重新界定问题并组织行动"，这种确信便常常动摇。学生可以准确复述定义却不能识别其适用边界；可以套用刚练过的公式却在结构相同、表面不同的题目前停住。

学习科学早已提醒区分表现与学习。提取练习在当下显得更困难，却带来更好的长期保持〔13〕；主动学习课堂中的学生可能感觉自己学得更少，测得的学习却更多〔2〕；迁移也不是简单的会与不会，而随知识领域、情境、时间、功能和反应方式的距离变化〔11〕。这些发现共同削弱了一个朴素等式：当堂表现等于学习。

但把表现与学习分开，仍不足以说明这种错位为何如此稳定。常见解释可归为六类：动机不足、教学方式单一、认知负荷与先备知识、反馈不够及时具体、迁移条件缺失、制度与评价约束。每一类都得到过证据支持，也都产生过有效干预，而每一类的干预在推广之后都被大面积吸收。这一事实本身需要解释。

本文认为，六类解释共享一个未被说出的前提：**学生的产物一旦被听见、被回应、被使用，就算被采纳了**。在这个前提下，"课堂有没有回应学生"是一个可以用是与否回答的问题，于是所有改革都朝同一个方向走——让课堂更多地回应学生：更多互动、更快反馈、更细的诊断、更智能的适应。

本文要论证的是：采纳不是二值的。同一份学生产物，可以被接在三个不同的层级上，而接在哪一层决定了它此后还能不能走。更要紧的是，**最积极、最迅速、最专业的某种回应，常常正是把产物接在最低那一层的方式**——因为低层的接住成本最小、见效最快、在任何形式审查中都表现为优秀。

本文把教育学、化学与管理学置于同一个问题上。教育学说明学习者内部为何发生表征与策略重组；化学说明为什么"被接住"与"能继续"是两件不同的事，并说明轮的边界从何而来；管理学说明局部变化在什么条件下才会被组织保存，以及为什么组织会系统性地把高层问题接在低层。

二、什么算学习：从即时正确到跨轮次的续走

"真正的学习"容易成为价值口号。若把它等同于深刻、自由、创造或终身成长，就难以研究，也贬低了记忆、模仿与熟练训练的价值。本文采用一个最低限度的操作性界定：相对于一次性表现，学习是知识、技能、策略或问题表征发生相对持久的改变，使个体在延迟、变式或新情境中能够恢复、选择、重组并解释自己的行动。

据此有四个常用证据面。**保持**：提示消失、时间推移之后仍能提取关键知识或重建操作。**迁移**：识别新情境与原任务的深层关系并转用知识，远迁移尤其困难，不能凭一道变式题宣告实现〔11-12〕。**重组**：在冲突证据面前修改原有表征，而不是把它当作例外吸收。**解释**：能说出自己为何这样做，以及在什么条件下不该这样做。

这四面主要描述学习者内部发生了什么。本文增加第五个过程证据：**续走**。学习者本轮形成的解释、错误、问题或策略，是否在下一轮里还能继续走下去。这一证据不替代前四面，而解释它们为何更可能出现：一个新表征如果只在被追问时短暂出现，下一题仍要求学生回到固定模板，它就难以稳定。

据此，本文以**学习轮次**为分析单位。一个最小轮次由四部分构成：任务或问题、学习者反应、解释或评价、下一轮条件。前三项只告诉我们本轮发生了什么，第四项才告诉我们本轮产物获得了什么后效。这里的"下一轮"不必等于下一节课；它可以是同一讨论中的下一次追问、作业的下一版本、平台给出的下一条提示、下周备课的调整，或一个单元后的评价修订。

需要立刻说清的是：本文反对的不是记忆、练习与即时正确。背诵、模仿与重复练习可以构成学习的重要组成部分。本文反对的只是把本轮的可见正确当作因果链的终点，以及把"有没有被回应"当作判断课堂的唯一维度。

三、教育学的贡献与它的止步处

教育学最直接地研究学习发生的条件。概念转变理论认为，新概念被接受并非因为教师清楚陈述，而是学习者对既有概念产生不满，并把新概念理解为可理解、合理且有解释力的替代方案〔3〕。这一判断把学习从信息接收推进到表征竞争。

生成性学习进一步表明，总结、绘图、自我解释、教别人等活动之所以有效，不在于学生表面上忙碌，而在于其促进选择、组织与整合〔14〕。ICAP 框架把认知参与区分为被动、主动、建构与互动，提出建构新推论与互为回应的对话通常比单纯操作材料更有学习潜力〔4〕。

形成性评价离本文的问题最近。Black 与 Wiliam 把它理解为使用学习证据调整正在进行的教与学〔5〕；Sadler 强调学习者需要理解质量标准、比较当前表现并采取缩小差距的行动〔6〕；高质量反馈应帮助回答目标在哪里、目前在哪里、下一步往哪里去〔7〕。这一支已经明确要求证据返回并改变后续安排。

生产性失败与"为未来学习做准备"把时间顺序推进了一步：学生先尝试复杂问题，即使未得出规范解，也可能生成多种表征、暴露关键差异，为后续讲授建立更精确的注意结构〔9-10〕。知识建构理论则更明确地要求共同体持续改进公共观念，使学生承担超越个人作业的认知责任〔15〕。

这些理论合起来提供了两个关键结论：学习必须经过学习者已有表征，不能由输入量直接推出；学生产物具有改变后续学习的认识论价值。

但它们止步在同一处。它们把“证据被使用”当作一个不再分层的终点。形成性评价要求证据改变教与学，却没有区分证据改变的是哪一层；生成性学习关心学习者内部的重组，把外部条件当作背景；知识建构要求观念被持续改进，但它规定的是一种教育形态而不是一个可以落在任何教学法上的编码变量。

于是出现一个在实践中极常见、而在理论上无法被这三家分辨的情形：一名学生提出的反例，实际挑战的是“什么算作证据”；教师听见了、重视了、当场回应了——回应的方式是补充两道对比题。这是模范的形成性评价：证据被读取，教学被调整，差距被缩小。而那个反例原本要挑战的东西，一次也没有被挑战。

四、化学的机制约束：接住不等于放行，而轮的边界是被制造出来的

把课堂称为“反应”很容易沦为比喻，因此必须先说明化学在本文中承担什么工作。它不负责把学生等同于分子，也不把教师等同于催化剂；人的理解具有意向、意义、历史与伦理，不能还原为物质过程。化学在本文中承担两件具体的推导工作，而不是提供修辞授权——删掉本节，第七节与第八节的两条核心判断都不成立。

第一件：接住不等于放行。

多相催化的现行理解把一个多步转化的实际产出归结为反应物与中间体在表面的结合强度。结合太弱，分子活化不了；结合太强，分子留在表面下不来。活性因此在中等结合强度处达到峰值——这就是萨巴捷原理与火山曲线的基本形状〔19-21〕。

这条原理的关键推论常被忽略：**最强的那个位点，恰恰是把分子接住却不放行的位点。**一个强吸附位点会优先捕获分子——它接得最快、最牢、最“积极”——而被它接住的分子不再参与后续转化。在极端情形下，强吸附物种会长期占据活性位点，使新的反应物无处吸附，这就是自毒化与位点封堵，是催化剂失活最常见的形式之一〔22-23〕。

也就是说，在化学里，“被接住”与“能继续”是两个必须分开来问的问题；而在很多体系中，**接住得越快越牢，越不能继续。**这一条不是类比，它是本文第七节层级判断的直接来源：课堂里最迅速、最专业、最不容置疑的那种回应，可能恰是把产物固定在最低层的方式。

第二件：轮的边界是被制造出来的。

化学的第二件材料更关键，因为它推翻的是一个从未被说出的前提。反应动力学有一个使用了一百多年的标准操作——稳态近似：把中间体的浓度对时间的导数设为零，于是中间体从总速率方程里被消掉，只剩反应物与产物〔24〕。它是一个近似，其成立条件是**中间体不积累**。

而在工业连续生产中，这个条件不是自然满足的，是被制造出来的：为了让批次之间可比、为了让上一批的积炭不影响这一批的读数，装置会**周期性再生**——把上一轮留在表面上的东西烧掉，让下一轮从接近初始的表面开始〔23〕。

这就给出了一条一般性判断：当一个系统要求逐轮结算与逐轮可比时，它会主动清除轮与轮之间的残留。轮的边界不是时间自己划出来的，是结算规则划出来的；而清除残留的动作，恰恰是维持可比性所必需的。

课堂同理。一节课要有明确的教学目标、要当堂达成、要可被检查，这些要求本身就要求下课时状态是干净的——一个"打开了一半、还没走完"的表征无法写进本节课的目标达成。**不是教师不愿意接住，是结算规则不允许有一样东西跨在两轮之间。**

化学同时反过来给本文加两条限制。其一，催化主要改变动力学而不会自动把错误终点变成正确终点：课堂拥有更快更密集的响应，不能补救错误知识、贫乏目标或资源匮乏；学生可以高效地相互强化一个误概念。其二，正反馈既放大有价值的结构，也放大偏差；因此本文不主张"层级越高越好"，见第十三节。

五、管理学的制度解释：为什么组织会系统性地把高层问题接在低层

课堂既是认知场域，也是一个小型组织。它有角色、时间表、例程、绩效指标、风险分配与资源边界。一个学生的新问题是否进入下一轮，取决于它是否被听见，也取决于课程是否有接口接收它、教师是否有权限调整、评价是否容纳新证据、时间是否允许。

组织学习研究把经验编码进例程视为关键过程。Levitt 与 March 指出，组织学习具有例程依赖、历史依赖与目标导向等特征，经验通过规则、程序、技术、信念与文化被保存〔28〕。若一名成员学到的新方法没有进入共同流程，这种学习可能随人员离开而消失。Feldman 与 Pentland 进一步指出，例程不是可被简单打包的黑箱，而是有内部结构与动力的系统〔30-31〕。

Argyris 的单环与双环学习提供了层级区分：单环在既定规范与目标内修正行动，双环则质疑并改变支配行动的变量、政策与目标〔27〕。这是本文最接近的一个理论，第十二节将专门交手。

管理学在本文中承担的核心推导，是解释**为什么组织在结构上倾向于把高层问题接在低层**。三条机制合起来足以说明。

其一，绩效反馈。组织往往把绩效与历史或社会期望相比较，低于期望时才更可能搜索并承担改变风险〔32〕。若课堂纪律良好、进度按时、正确率较高、家长满意，系统缺少搜索信号。即使学生只能处理同型题，只要当前评价不测延迟保持与远迁移，这种局限就不会表现为绩效差距。

其二，探索与利用的资源分配。组织必须在搜索新可能与利用已知方案之间分配有限资源，而短期反馈常使组织偏向利用〔29〕。低层采纳属于利用：加两道题、换一个例子、多讲一遍，代价可计、见效可测、当堂可验收。高层采纳属于探索：重新界定问题、修改证据标准、调整评价依据，代价高、周期长、结果不确定。

其三，权限的分布。修改一道题在教师权限内；修改一个单元的评价标准通常不在；修改统考口径几乎一定不在。**因此，即使教师准确识别出一个产物的诊断层级在规则层，他能够动用的最高接口也可能只到路径层。**这一点极其重要：降层采纳在很多情形下不是判断失误，是权限边界的结果。

三条机制合起来解释了一个稳定现象：**组织对高层问题的标准回应，是一个低层动作。**这不是懈怠，它在每一步上都是理性的。

六、三家共同的隐含安排

三门学科分别关注学习者表征、转化路径与组织例程，表面上对象迥异，却在同一处相遇。

教育学问任务如何改变学习者，化学的常规表达从反应物与条件走向产物，管理制度往往从目标与规则走向执行绩效。三者都容易采用"条件—过程—结果"的单向叙事。单向叙事不是错误，它是分析的必要简化；但当它被默认为唯一形态时，就会带来一个共同后果：**采纳被当成一件有与无的事。**

三家都假定：如果学生的产物影响了后来，那就是被采纳了；如果没有影响后来，那就是没被采纳。于是三家的处方都朝同一个方向走——增加影响的机会：更多生成活动、更快反馈、更多学生话语权、更强的制度接口。

推翻这个假定的材料，来自化学自己，而且是化学里最平常的一件事：**最强的吸附位点接得最快最牢，而被它接住的分子退出了转化。**在化学里，这件事回答的是催化剂为什么失活、为什么最优结合强度在中间；从来没有人把它当作关于"采纳"的证据。

一旦把它移过来，那个共同假定就站不住了：采纳有强弱，也有层级；而**采纳得越快越牢，未必越能继续走。**

由此，课堂的四种断裂可以被重新描述。时间断裂：一轮结束即清零，而清零是结算规则的要求，不是疏忽。对象断裂：学生改变被视为个人属性，不进入任务与规则。证据断裂：评分记录答案是否正确，却不记录哪一层因此改变。权限断裂：教师可能认识到产物的价值，却没有相应层级的接口。四者之中，第三与第四正是本文的着力点。

七、采纳的三层及其持有者

（一）三层

表征层。产物改变学习者本人下一轮如何分类、解释与注意。例如学生发现"重物下落更快"的直觉不能解释真空中的情形，于是在下一问题中改用力与运动状态的关系。这一层主要发生在个体内部，可通过概念图版本、自我解释与新情境判断推断。

路径层。产物改变下一轮的行动方式、支架、题序、协作结构或资源选择。例如教师根据学生把"面积不变"等同于"周长不变"的错误，取消原定同型练习，改为比较一组反例；小组依据首次讨论中话语被一人垄断的结果，下一轮设立证据记录员与反例提出者角色。路径改变不修改目标，改变的是到达目标的方式。

规则层。产物促使课堂重新界定何为问题、何为有效证据、何为成功，或谁有权提出下一个问题。例如学生指出一道社会调查题的分类排除了某类经验，班级据此修改变量定义；一项实验出现与教材预期不一致的结果，教师没有把它当作操作失败排除，而是让全班讨论可重复性与异常值处理标准。

三层不是由低到高的道德阶梯。基础技能训练中，稳定的答案与路径有助于自动化；实验安全中，规则不能由一次局部经验随意改变；新手在缺少图式时也可能因开放选择而超载。

（二）三层的持有者

本文的第一处承重判断在这里：三层各有不同的持有者，而持有者决定了这一层的接口能否被打开。

表征层由**学习者**持有。只有他能改变自己下一轮如何注意与解释；教师可以提供材料与追问，但不能替他改。

路径层由**教师**持有。改一道题、换一个表征、调一次题序、重排一次分组，通常都在教师的当堂权限之内。

规则层由**组织**持有。何为有效证据、何为达标、用什么评价、以什么为下一个问题，这些通常不在单个教师的权限内，而在教研组、课程团队、学校、考试制度或平台的数据结构里。

这一分层立刻给出一个既有理论看不见的推论：一个产物被接在哪一层，不只取决于它值不值得被接在那一层，还取决于那一层的持有者此刻在不在场、有没有权限、承不承担后果。一名教师面对一个规则层的反例，他能够动用的最高接口只到路径层——不是因为他不理解，而是因为规则层不归他。

（三）三家理论各自只覆盖一个持有者

现在可以把第三节留下的问题接上。三个最容易与本文混淆的理论，恰好各自只覆盖一个持有者：

生成性学习与概念转变覆盖表征层，持有者是学习者。它们的全部判据都落在学习者内部的选择、组织与整合上，因此它们无法分辨"学生自己改了"与"课堂条件也跟着改了"。

形成性评价覆盖路径层，持有者是教师。它的定义就是使用学习证据调整正在进行的教与学，而教师能够调整的正是路径。这解释了一件此前无法解释的事：**形成性评价在设计上就是一个路径层工具，因此它会系统性地**把规则层的产物接在路径层。这不是它做得不好，是它的持有者只到那一层。

双环学习与组织例程覆盖规则层，持有者是组织。它们讨论支配变量与规范的修订，却不以单个产物为单位，因此看不见"一个具体的规则层产物被降层接住"这件事。

三家不是三个相邻的竞争理论，而是**同一个采纳层级问题在三个持有者上的三次读数**。承认这一点，三家各自的适用范围与各自的盲区同时被解释：每一家看不见的东西，恰好是它的持有者管不着的那一层。

八、降层采纳：本文的承重命题

由此得到本文的承重判断，它可以写成一句：

＞ 课堂上真正的学习之所以稀缺，不是因为学生产物没有被采纳，而是因为大量产物**被采纳在低于其诊断价值所在的那一层**；采纳发生了，而它原本要挑战的东西恰恰因此不再被挑战。

本文把这种情形称为**降层采纳**。它有三个必要条件与一个确认条件。

条件一：存在一个可识别的学生产物（解释、错误、问题、策略或规则质疑），且它在被提出时具有可判定的**诊断层级**——它挑战的是学习者自己的表征、是到达目标的路径、还是"什么算数"本身。

条件二：该产物在后续轮次中**确实被采纳**——存在一项可追踪的条件变化，且该变化与该产物之间有可辩护的关联（直接引用、据此改题、依据错误类型换用表征、由学生问题触发资料选择）。仅仅被听见、被表扬、被记录，不算。

条件三：实际采纳层级低于诊断层级。差值即为该次采纳的**层距**，取值 0、1 或 2。

确认条件：该产物原本要挑战的那一层，在此后的若干轮内**没有被任何其他产物触及**。这一条排除了"先在低层接住、后来在高层处理"的正常序列——那种情形是延迟采纳，不是降层采纳。

降层采纳与"没有采纳"在现象上完全不同，在后果上却可能更糟。没有采纳会留下不满：学生会觉得没被听见，教师会觉得没时间，系统里存在一个可被指认的缺口。降层采纳不留下任何不满：学生被回应了，教师做了专业判断，证据被使用了，形式审查全部通过。**它的失效不产生任何信号。**

它还自我加固。层级越低，采纳成本越小、见效越快、越容易在当堂被验收；因此越是强调"及时反馈""精准诊断""数据驱动"的课堂，低层接口越发达、被使用得越频繁，而高层接口越不被使用。**越正规化越严重**，这是靶格成立的第三条签名。

九、采纳层距：一个可清点、且与既有指标正交的读数

一个判断若不能被清点，它就只是一个说法。本文给出的读数是：

＞ **采纳层距** = 最近 N 次被采纳的学生产物中，诊断层级与实际采纳层级之差的平均值；取值域 [0, 2] 。

配套报告**采纳层级分布**（表征／路径／规则各占多少）与**高层接口使用率**（规则层被触及的次数占比）。三个数分别报告，不得加总——高频的低层采纳与低频的规则层采纳不是同一种现象。

三层的判定判据

层级若判不准，层距就只是编码者的印象。判定用一句测试：**顺着这个产物往下走，最先撑不住的是哪一样——他的想法、他的做法，还是"什么算对"？**

规则层的判据是：该产物若成立，会使此前已被判为正确的某些答案不再正确，或使此前不算证据的东西开始算证据。它改动的是判分依据本身，不是某一道题的答案。学生指出评分标准在此例上自相矛盾、指出题目的分类排除了某类情形、指出实验中的异常值处理规则与结论方向有关——三者都在这一层。

路径层的判据是：该产物若成立，答案与判分依据都不变，改变的是到达那个答案的可行路线集合。学生发现代数法无法揭示某个几何约束、发现两类错误需要不同的对比例、发现讨论被一人垄断因而需要角色分工——三者都在这一层。

表征层的判据是：答案、依据与路线都不变，改变的只是这一个学习者此后如何注意与分类。它不产生任何外部可见的条件变化，只能通过概念图版本、自我解释与新情境判断推断。

三条判据是层层递减的：先问它动不动判分依据，不动则问它动不动路线，都不动则记为表征层。**编码时必须先判诊断层级、后看后续材料**，且两项由不同编码者独立完成——否则会出现按结果倒推层级的循环，那会使整套读数失去意义。

一处诚实的保留：这套判据在多数情形下可判，但在少数情形下不可判——一个质疑可以同时挑战路线与依据，而两者的先后取决于此后谁先被触及。此类序列应记为不可判并单独报告，不得强行归入某一层。

三个领域的兑现如下，量纲不同而比率相同。

课堂：从课堂录音与教案前后版本中，逐条编码“产物—诊断层级—实际采纳层级”。一节课通常产生三到八个可编码序列。

催化：一个分子被强吸附位点捕获而不进入后续转化，与它进入目标路径，是两种不同的“接住”。工业上以**选择性**清点二者之比——被接住的分子里有多少进入了目标产物。选择性与转化率是两个独立的量，而现行课堂只清点转化率。

组织：一线提出的问题若属规则层（“这个流程本身在什么条件下不适用”），而管理层的回应是加一次培训或增加一道检查（执行层），层距即为一或二。工单系统里的“处理方式”字段已经存在，只需按层级重新编码。

四条性质逐条对照。**无量纲**：三处都是比例或差值，可并排比较。**可事后清点**：不依赖现场判断，录音、教案版本、工单记录在多数机构里已经存在。**把一个印象变成一个数**：教师常说“这个班的问题提得深”或“提了也没用”，此前没有对应读数。

与既有主要指标正交——这一条最容易漏，必须举出一个“既有指标最好看而这个数最难看”的真实例子：

> 一个反馈最及时、诊断最精准、每个学生的错误类型都被自动识别并推送对应练习的智能课堂，**采纳层距接近最大值**。因为它的全部接口都在路径层：它能做的唯一动作是换一道题。它的即时正确率、参与度、反馈延迟三项指标全部处在最好看的一侧，而规则层接口使用率为零。

反向的例子同样存在：一个反馈很慢、当堂正确率一般的班级，若其中一名学生的反例导致全班用两节课重新讨论“什么算作足够的证据”，则该次采纳层距为零，而所有既有指标都会变难看。**这正是本文预测长期表现更好、而现行验收会判定为进度失控的那一类。**

十、判据：一句不含情态词的问话

本文的判据是：

＞上一轮那个产物，改变了下一轮的哪一层——它换的是一道题，还是换了"什么算数"？

这句话可以拿去问教案的前后版本、课堂录音、平台日志与工单记录，不需要问任何人的判断，也不含"应当""充分""真正""恰当"这类词。它只要求一个具体的指认：**哪一层**。

把它在五个场景里跑一遍，答案互不相同。

一节物理课。学生问"潜艇改变的是自重还是排水体积"。若下一轮的改变是"教师补充两道浮力计算题"，答：路径层，层距一。若改变是"本单元此后把'哪些量可以独立变化'列为每题必答项"，答：规则层，层距零。

一次师徒带教。徒弟指出师傅的判断依据在这一例上不成立。若改变是"师傅多示范一遍"，答：路径层。若改变是"此后凡遇此类情形，两人先各自写下依据再比对"，答：规则层。

一个催化实验室的连续批次。问的不是转化率而是选择性：被接住的分子里有多少进入了目标路径。这一读数在任何一份反应报告里都已经存在，而它的存在正说明**化学早已把"接住"与"能继续"分开清点**——这一条是查出来的，不是推出来的。

一次医院晨会。住院医师指出昨日的鉴别诊断路径在这一例上排除得过早。若改变是"今天多查一项检验"，答：执行层。若改变是"科室修改该类主诉的初筛清单"，答：规则层。值得注意的是，多数科室的交班记录里有"处置"字段，**没有"因何而改"字段**——这同样是查出来的。

一次产品迭代。用户反馈指向的是"这个功能的成功标准定错了"。若改变是"优化该功能的交互"，答：路径层。若改变是"重定义该功能的北极星指标"，答：规则层。工单系统里的"处理方式"字段已经存在，只需重新编码。

五个答案互不相同，其中两个是查出来的而不是想出来的：**反应报告里早就分开清点选择性与转化率；交班记录里有处置字段而没有因何而改的字段**。

十一、二维判别格：识别最危险的"结算性成功"

只用答对率判断学习时，所有低表现都像失败，所有高表现都像成功。加入采纳层级之后，可以得到一个更具诊断力的二维格。

高层接口被使用（层距小）

只有低层接口被使用（层距大）

本轮输出高 生成性掌握：正确产物被带入变式、生成新问题，或改变了"什么算数"

结算性成功：答案正确、任务完成，回应及时，而被挑战的那一层一次也没被触及

可见失败：错误被记分或纠正，后续按原计划运行

高层接口被使用（层距小）

只有低层接口被使用（层距大）

本轮输 生产性扰动：错误或困惑改变了证据标准、

出低 问题界定或评价依据

"可见失败"最容易触发干预，因为分数、沉默与错答都会报警。但若干预只是重复原输入，系统可能长期停留在这一格。

最易被忽视的是**结算性成功**。它同时满足制度最喜欢的信号：正确、快速、有回应、可评分、按进度。学生在教师提示下得出结论，教师说"很好，我们再看两道类似的"，进入下一页；平台判断掌握后推送新知识点，却不要求学生说明策略或识别适用边界。每一项都可以是真实成绩，而它们共同的弱点是：**被回应的层级永远低于被挑战的层级**。

这一格与"没有回应"完全不同，也与"回应质量差"完全不同。它的回应质量往往很高——只是全部落在同一层。

二维格不是给课堂贴永久标签。同一节课的不同学生、不同知识点与不同轮次可以落在不同格；同一产物也可能先被低层接住，之后才在另一情境中被高层处理——那属于延迟采纳，须由第八节的确认条件排除。研究者应编码序列，而不是凭一次观察下结论。此外，高层采纳并不保证高质量：若错误观点在缺少证据校准时改变了规则层，它可能造成误概念扩散。判别格必须与知识准确性、迁移、保持与公平指标共同使用。

十二、四类课堂场景：同样的表现，何以导向不同的学习

（一）讲授中的一次正确回答。 教师讲完浮力概念，问"浮力与什么有关"，学生答"与排开液体的重力有关"，教师肯定后翻页。这是高输出，采纳层级为零——没有任何一层因此改变。若教师接着请他解释"很重的钢船为什么能浮"，并据其把"重"与"密度"混淆的解释改换对比例，采纳落在路径层。若他提出"潜艇改变的是自重还是排水体积"，而教师据此把"哪些量可以独立变化"列为本单元此后每题必答项，采纳落在规则层，层距为零。同一个正确回答，三种去处。

（二）数学例题后的熟练练习。 两班都学会用配方法解一元二次方程，当堂正确率接近。甲班继续完成十道同型题，错误被逐题订正；乙班要求学生比较配方法、公式法与图像法各自暴露的信息，并把学生提出的"什么时候图像法更能看出根的结构"写入下一课的问题清单。甲班的采纳全部在路径层，乙班有一次落在规则层——它改变的是"这一单元里什么问题值得问"。本文不预先宣布乙班全面优越，而作出更具体的预测：在需要选择表征、解释参数变化或处理陌生方程的任务上，乙班更可能重组策略；在近期同型测验上，甲班可能更稳定。

（三）小组讨论中的"被听见"。 一节历史课讨论某项政策的影响，学生自由表达，气氛良好，最后教师按教参归纳三点。按参与度看这是主动课堂，按心理安全看学生敢于发言，按本文看采纳层级为零。另一班的讨论未必更热烈，但教师把各组相互冲突的因果解释整理为下一轮的材料选择标准，让学生寻找能够

区分这些解释的证据——被接住的不是某个学生"观点正确"，而是"什么样的证据能区分这些解释"，这是规则层。

（四）自适应平台的高频反馈。 平台可在每道题后给出反馈并依正确率调整难度，表面上循环极密。但若它只识别答案对错，不识别学生采用的策略、错误模型与提示依赖，它能做的唯一动作就是换一道题——**它的全部接口都在路径层，因此它的层距在结构上不可能小。**相反，一个较简单的系统即使不实时，也可以保存学生解释、比较草稿版本，让教师据此修改下周的问题清单。采纳层级不是技术先进程度的同义词。

四类场景说明：教学活动的名称推不出采纳层级。讲授可以包含规则层采纳，讨论可以只有象征性参与；纸笔作业可以改变课程，智能系统也可以把学生锁在同一层。

十三、十个行当走一遍

一个判断若只在一个领域成立，它是一个领域内的观察；要成为一个可用的辨别维度，它必须能在别处据以预测一件具体的事。以下十处，每处给出一条可去查的预测。

一、异相催化的连续生产。 预测：以选择性为主要考核指标的产线，其催化剂配方演化方向与以转化率为主要指标的产线不同，且前者更早引入弱吸附位点的设计。这是一个可以从既有工艺档案里回溯检验的预测——**它同时说明化学早已把"接住"与"能继续"分开清点，而课堂尚未。**

二、中小学课堂。 预测：控制总课时、内容量与班级规模后，采纳层距小的班在未教过的新情形任务上表现更好，而在当堂检测与单元测验上不占优势，甚至略低。

三、住院医师培训。 预测：在交班记录中增设"因何而改"字段并要求指明被改动的层级，在不增加总带教时长的条件下，一年后住院医师在非典型病例上的鉴别诊断表现改善；而标准化考试成绩无差异。

四、师徒制手工艺传承。 预测：手工艺传承难以规模化，主要不是因为知识不可言传，而是因为规模化把规则层从师傅手里收走了——课时、考核与结业标准归了机构。可检验的推论：把师徒制课程化之后，结业通过率上升而独立处理异常情形的能力下降。

五、科研组会。 预测：以"每人汇报本周进展"组织的组会，其采纳几乎全部落在路径层（下一步做什么实验）；若改为从上次散会时那个没吵完的分歧开始，同样时长下，关于"什么算作足够证据"的讨论频次上升，而周进展汇报的完整度下降。组会纪要可直接编码。

六、软件工程的缺陷处理。 预测：把用户反馈按"改交互／改功能／改成功标准"三层编码后，团队的层距分布与其产品在两年尺度上的方向稳定性相关，而与缺陷关闭速度无关，甚至负相关。

七、心理治疗。 预测：按固定节数结算的短程治疗会把每次会谈推向一个可结算的收束，因而来访者提出的、实际挑战"我们在治什么"的那一类表述，最常被接在"这次的作业"这一层。逐次会谈录音的开场与收束编码可直接检验。

八、法庭的连续开庭。 预测：跨越多次开庭的复杂案件中，当事人提出的、实际挑战举证责任分配的主张，若被接在“补充一份证据”这一层，则争议焦点的收敛速度显著慢于被接在举证规则层的情形。庭审笔录已存在，可直接编码。

九、企业的年度战略复盘。 预测：复盘提出的、实际挑战“我们凭什么算成功”的意见，被接在“明年的行动项”这一层的比例，与该企业五年尺度上的战略漂移正相关；而与年度目标达成率无关。

十、语言学习。 预测：以单元为结算单位的教材，学习者提出的“这句话为什么不能这样说”这类质疑，几乎全部被接在“再记一条例外”这一层；若改为让它进入本单元“什么算作可接受表达”的讨论，则口语产出的复杂度增长快于同等学时的对照组，而单元测验成绩无差异。

十一（一处方向相反的预测）、安全关键系统。 在核电、航空、麻醉与高压作业中，把一线提出的质疑接在规则层是危险的——规程的稳定性本身就是安全的一部分。此处的正确做法是把它接在**规程之外的独立复审通道**里，而不是当场改动规则。这不是本文的反例，是适用边界的正面标定，见第十九节第四条。

十四、与最近的几种说法的分界

本文最容易被读成某个既有概念的换名。以下逐一交手，每条给出一个可以分出胜负的对照观察。

双环学习（Argyris 1977，组织理论〔27〕）。它说到哪一步：单环在既定规范内修正行动，双环质疑并改变支配行动的变量与目标。这是本文最接近的一个说法，两者都区分“改做法”与“改规范”。分离线有两处，都可判定。其一，**单位不同**：双环学习以组织的学习类型为**单位**，本文以**单个产物**为单位；因此本文能够描述一个整体上是双环学习型的组织，其中某一个具体的规则层产物仍被降层接住。其二，也是更要紧的一处：**双环学习没有“产物本身的诊断层级”这个量**。它问的是组织改了什么，不问那件被回应的事本该被接在哪一层——因此它没有“层距”，也无法把“回应了但降了层”与“没有回应”分开。判定形式：**若观察到一个被公认为双环学习型的组织，其规则层产物的平均层距仍显著大于零，则本判断对而该概念不覆盖本判断；若层距与组织的单环／双环归类完全共变，则本判断可被它吸收。**

形成性评价（Black & Wiliam 1998；Sadler 1989；Hattie & Timperley 2007〔5-7〕）。它要求使用学习证据调整正在进行的教与学。第七节已给出分离线：**它是一个路径层工具，因为它的持有者是教师，而教师的当堂权限止于路径**。因此它会系统性地把规则层产物接在路径层，而这在它自己的框架内完全是成功。判定形式：**取一个规则层产物（如学生指出评分标准在此例上自相矛盾），若教师立即给出针对性的对比比例与反馈，则按形成性评价判为高质量，按本判断判为层距一的降层采纳；此后若该评分标准在四周内从未被任何人触及，则本判断对而该概念错；若该标准随后被修订且修订可追溯到该产物，则属延迟采纳，本判断的确认条件不成立**。这是本文与教育学主流处方最直接的对立，也是最容易被做实做检验的一条。

响应式教学（Hammer 一支，科学教育）与**协商课程／生成性课程**。它们主张跟着学生的想法走，让学生的想法改变后续教学。分离线：它们规定的是**教师的取向与设计意图**，本文测的是**单次采纳落在哪一层**。判定形式：**若观察到一个完全采纳响应式教学理念、教师全程跟随学生想法的课堂，其规则层接口使**

用率仍为零（因为规则层不归教师持有），则本判断对而该取向不足以推出本判断；若取向到位即层距为零，则本判断可被它吸收。这一条是本文对该取向的实质增量：它指出跟随学生的意愿改变不了持有者的边界。

知识建构（Scardamalia & Bereiter [15]）。它把共同体的观念视为可持续改进的公共对象，要求学生承担集体认知责任。这可被视为规则层采纳被制度化的一种教育形态。分离线：知识建构是一套教育目标与实践原则，本文是一项跨教学法的编码变量——它同样可以诊断拼音、乘法、器械操作与考试复习。判定形式：若在不实行知识建构的常规课堂中，层距仍能预测延迟迁移，则本判断独立成立；若层距只在知识建构课堂中有变异，则它只是该形态的一个内部指标。

生产性失败与理想困难（Kapur 2008; Roediger & Karpicke 2006 [9,13]）。前者强调在规范教学前尝试复杂问题，后者揭示短期表现与长期学习可能分离。分离线：本文不要求先失败，也不把困难本身视为好事；一个失败若只被教师判错、下一步仍按固定脚本运行，采纳层级为零；一个正确解若改变了"什么算作足够的证据"，则采纳落在规则层。判定形式：在失败经历量匹配的两组之间，只改采纳层级而迁移出现差异，则本判断对；若匹配后无差异，则反之。另须点名一处近旁：该支近年提出的"无成效的成功"（unproductive success）与本文的结算性成功共享同一直觉，但它的第二根轴是延迟结果、须事后回溯判定，本文的第二根轴是过程变量、当堂即可编码。判定形式：若两者在同一批课堂上的编码结果高度一致，则本判断只是它的一个前瞻代理；若存在一批被本文判为层距零而按延迟结果判为无成效的课堂，则二者测的不是同一件事。

学生能动性（OECD 2019 [16]）。它关注学生设定目标、行动并影响自身与环境的能力与意愿。分离线：能动性是一种主观—环境关系，层距是一种可核查的层级落差。判定形式：若一个班级的学生能动性量表得分很高而层距接近最大值（学生有选择、有声音，所有选择导向同一份标准产物），则本判断对而能动性不足以推出它。

主动学习与 ICAP（Freeman 等 2014; Chi & Wylie 2014 [1,4]）。它们按活动外观与认知参与状态分类。分离线：本文不按活动外观分类，按采纳层级分类。判定形式：控制本轮互动与生成活动的量之后，采纳层级仍对延迟迁移有额外预测力，则本判断保留；否则应予撤销。这一条同时是第十六节命题一。

必要多样性定律（Ashby 1956，控制论）。把本文的命题剥到只剩形式，它是这样一句话：一个多层系统对高层扰动的标准回应发生在低层，因而扰动在形式上被处理而实质未被吸收。这个纯形式的结构在上游有名字——调节器的多样性必须至少等于被调节系统的多样性，否则扰动无法被吸收。分离线：必要多样性讲的是多样性总量不足，本文讲的是多样性分布在错误的层级上。判定形式：若一个课堂的路径层接口极其丰富（题库巨大、诊断精细、反馈即时），其总响应多样性远高于对照班，而规则层接口使用率为零、层距仍最大，则本判断对而必要多样性不覆盖它；若增加总响应多样性即可使层距下降，则本判断可被它吸收。前一种情形正是第九节所举的智能课堂。

成功陷阱与能力刚性（Levinthal & March 1993; Leonard-Barton 1992 [29 及相关]）。必须点名，因为本文的形式带有"越成功越失败"的反转模板：回应越及时、越专业、越自动化，降层采纳越严

重。分离线：成功陷阱讲的是**利用挤出探索**，是同一维度上的资源分配；本文讲的是**接口所在的层级**，不是分配问题。判定形式：**若在探索资源被硬性保障（划出固定比例的时间与预算）的条件下，规则层接口使用率恢复，则成功陷阱对而本判断多余；若资源被保障而持有者边界不变、层距仍不下降，则本判断对。**

自催化（IUPAC [22]；Bissette & Fletcher 2013 [23]）。它为"产物成为后续原因"提供机制来源，但课堂产物具有意义与规范属性，需要教师判断、公共证据与伦理审查。本文从化学取的**不是自催化**，而是第四节那两件更硬的材料：接住不等于放行，以及轮的边界是被制造出来的。判定形式见第十六节命题五与命题六。

至此仍可问：谁是最接近的那一个说法？是形成性评价。本文之所以不能被它吸收，只在一处：**它的持有者是教师，因而它的接口止于路径层；而本文测的恰恰是接口层级与产物层级之间的落差。**一个把形成性评价做到极致的课堂，其层距可以是全校最大的。

十五、与本栏近作的分界

本栏近期有两篇处理同一个问题域，必须逐一指名并给出可分胜负的预测。

《**方程两边都有它**》。那一篇的变量是**轮内归属**——一次转化中分岔的那一步由谁走完；其干预是首次作答可撤回一次。本文的变量是**采纳层级**——产物被接在哪一层。两者正交：一个分岔步完全由学生自己走完（按那一篇判为良好），而他走完之后提出的规则层质疑被接在路径层（按本文判为靶格）。判定形式：**在分岔步归属被固定的条件下，只改采纳层级而延迟迁移出现差异，则本文不可被那一篇吸收；若归属一旦固定、采纳层级不再产生差异，则本文可被它吸收。**反向亦然。

《**每一轮都从新鲜表面开始**》。那一篇的变量是**轮间发起处**——下一轮从起始态还是从上一轮终态开始。本文的变量是**层级**——被接住的是哪一层。两者可分离，且两个方向的反例都容易找到：一个课堂可以每节课都从上一轮的具体未决处开始（发起处接得很好），而所有接住都发生在路径层（层距最大）；反过来，一个课堂可以每节课都从教案起点讲起（发起处落在起始态），而本学期有一次学生的反例改写了评价标准（那一次层距为零）。判定形式：**在发起处规则被固定的条件下，只改采纳层级而迁移出现差异，则两篇互不覆盖；若固定发起处后层距不再产生差异，则本文可被它吸收。**

并诚实记下一处继承：本文第四节关于"轮的边界是被制造出来的"这一判断，与那一篇关于结算规则强制清零的论证同源，此处不掩饰，该处的功劳记在那一篇账上。本文在它之上增加的，只有一件：**边界被制造出来之后，产物在被清除之前会先被接住一次，而接在哪一层决定了清除的后果。**

三篇合起来的分工可以写成一句：一篇问这一步该由谁走，一篇问下一轮从哪儿开始，本文问被接住的是哪一层。三者可同时施加，且本文预测它们的效应**不是相加而是相乘**——一个由学生走完的分岔步，若其产物被接在低层，则下一轮是否从终态开始都不再要紧。

十六、可证伪的六组命题

一个理论若能解释任何结果，就没有解释力。本文至少应承担以下六组经验风险。在此之前必须先摆出最强的竞争解释，而且不能挑一个好打的：“其实不就是课时不够、内容太多、教师没时间吗？”这个解释朴素、难反驳，且对本文观察到的一切都给出说法。本文的独有变量是层距，六组命题都在控制掉绝对负荷之后仍要求它成立。

命题一·增量效度。 在控制先备知识、教师质量、任务难度、在任务时间、练习数量、互动频率与即时正确率后，层距应对延迟保持、远迁移、新问题生成或策略重组提供额外预测。若层距与主动参与、练习量或反馈频率完全重合且无增量效度，则本构念应被撤销。

命题二·层级特异性。 三层的效应应当各不相同：路径层采纳主要改善近迁移与熟练；规则层采纳更可能影响问题界定与证据判断，同时伴随更大的成本与波动。若三层对所有结果产生相同作用，层级划分应予撤销。

命题三·时间顺序。 若层级是机制而非事后标签，则应观察到：产物出现在先，采纳发生在后，迁移改善再后。若研究反复发现教师先按预案改变了任务，学生产物只是随之出现，而编码仍把它记为采纳，则因果方向被颠倒。时间戳、教案预版本与课堂序列分析是必要证据。

命题四·靶格差异。 即时输出相近时，层距小的课堂与层距大的课堂在同型近测上可以没有显著差异，但前者应在延迟迁移、策略选择理由与学生自生问题质量上表现更好。若二者在预注册的后测上长期无差异，则“结算性成功”只是修辞标签。后测的敏感度须在研究前写死并预注册，不得事后声称后测不够敏感。

命题五·非单调与持有者约束。 层距的收益不应是一条越小越好的直线。在先备知识不足、证据标准薄弱、心理安全低或任务高度程序化时，频繁的规则层采纳可能增加认知负荷、强化错误或损害协调。更要紧的是持有者约束：若在教师权限不变的条件下，仅靠改变教师意愿即可使层距显著下降，则第七节关于持有者的判断为伪——那将说明层级边界只是取向问题，而不是权限结构问题。

命题六·接住不等于放行。 本文从化学取来的核心推论要求：响应速度与响应层级应当负相关或至少不相关。若在同一批课堂中，反馈越快、诊断越精细、自动化程度越高的课堂，其规则层接口使用率上升，则第四节的推导为伪，本文最独有的那条判断随之作废。这一条是六条中最便宜的，因为智能教学平台的日志已经记录了反馈延迟与推送类型，只需按层级重新编码。

这些命题把本文置于可失败的位置。尤须强调：不能仅凭教师访谈中“我会根据学生调整教学”来验证。意向不是后效，响应性自评也不是因果链。有效证据必须同时显示可识别的学生产物、其诊断层级的独立编码、随后条件变化的层级、二者之间的依赖关系，以及学习结果的延迟表现。

若本文被证伪，会学到什么。 若命题一不成立，那么真正稀缺的东西不在采纳的层级上，而在采纳的总量上，干预方向应回到“增加回应机会”这一条既有路线；若命题五不成立而层距可由教师意愿单独调节，那么问题属于教师专业发展而非权限结构，这将是一个比本文乐观得多、也便宜得多的结论。

十七、两处反过来削弱本文的约束

一个判断若从头到尾只被材料支持，多半没有真的接触材料。以下两处，是三家之中的两家反过来给本文加的约束，各削掉了本文原本会写下的一句更强的话。

第一处：化学削掉了"层距越小越好"。

本文原本会写：层距是一个越小越好的量，理想课堂的层距接近零，规则层接口应当尽可能频繁地被打开。化学不允许。催化主要改变动力学，它不把错误的终点变成正确的终点；一个高效的正反馈同样会放大偏差。更具体的一条来自位点封堵：**规则一旦被改动，此后所有产物都要在新规则下重新被判读，而新规则本身尚未被校准。**频繁改动规则层的课堂，会使每一轮的判读标准都处在未稳定状态——这在化学里对应于一个持续重构、从未达到稳态的表面，其选择性反而最差。

于是本文必须收窄：**层距与学习累积的关系是非单调的，存在一个由"新规则能否在被改动后仍然可用"决定的下限。**削掉的是"越小越好"这条线性主张，换成一条带下限的关系。它同时带来一条可检验的推论：**当规则层改动的频率高到新标准来不及被校准时，规则层采纳反而降低迁移**——这解释了为什么某些高度开放的课堂在长期评估中并不占优，也给出了判断何时该停止改动规则的标准：不是看开放程度，而是看上一次改动后的新标准有没有被至少两轮独立使用过。

第二处：管理学削掉了"提高教师的意愿就够了"，并改动了本文的适用范围。

本文原本会写：只要教师意识到产物的诊断层级，主动把它接到相应的一层，层距就会下降。因此本文的处方将是教师专业发展。管理学直接毁掉这条路：规则层由组织持有，不由教师持有。一名完全理解本文、且完全愿意照做的教师，面对一个规则层产物，他能动用的最高接口仍然只到路径层——除非组织把那一层的接口向他打开。

于是本文的处方形态必须改变：**本文的干预不能以教师培训的形式实施，只能以接口设计的形式实施。**可操作的形式是：明确写出哪一类产物由哪一层接收、由谁批准、在多长时间内容答复、被拒绝时须给出何种理由；并为规则层保留一个低成本、低风险的入口（例如每单元一次、限定范围、有明确复核程序的"标准复议"）。

这一让步同时改动了适用范围：**本文只适用于规则层接口在原则上可被打开的系统。**在规则由外部硬性给定的场合——统一考试口径、法定课程标准、安全规程——本文只能给出诊断，给不出可执行的干预；此时诚实的做法是承认路径层是唯一可动的一层，并把精力放在把路径层用满，而不是假装规则层开放。

第三处（教育学提出，本文接受但不作为主要约束）：并非所有规则层产物都值得被接在规则层。一个基于误概念的质疑若被接到规则层并改变了评价标准，其危害远大于被接在路径层。因此层距的解读必须与知识准确性联用：**层距小而准确性低，比层距大更糟。**

十八、研究方案：从课堂印象转向层级与版本证据

（一）样本与比较策略。 初步检验应选择同一学校、同一年级、同一学科的至少十二个平行班，覆盖一个完整教学单元，并设置二至六周的延迟测验。之所以不用两个风格迥异的优秀案例作比较，是因为教师经验、学生构成、课程进度与学校文化会同时变化，任何差异都可被事后解释。较同质的多班设计可以利用同一教师跨班、同课异构或同一班级跨教学周的变化，减少把教师魅力误认为层级效应的风险。**严禁以两地、两校或两国的对比充当剂量—反应的证据**——两边在处处不同，所谓控制在方法上站不住，只能作启发式说明。

（二）数据与编码单元。 需要收集课堂录像与转写、学生连续作业版本、教师备课的前后版本、板书或课件版本、平台分支日志、自拟问题与概念图，以及即时、延迟、近迁移与远迁移任务。分析单位不是整节课，也不是一次发言，而是“任务／问题—学习者产物—采纳—下一轮条件”的最小序列。

每个序列至少编码七项：产物类型（答案、解释、错误模型、问题、策略、规则质疑）；**诊断层级**（表征／路径／规则，须独立于采纳层级编码）；是否被后续明确引用；**实际采纳层级**；层距；时间延迟与影响范围；采纳、延迟或拒绝的理由。

两项防伪措施是必需的。其一，**诊断层级与采纳层级必须由不同的编码者独立给出**，且诊断层级的编码者不得看到后续轮次的材料——否则会出现按结果倒推层级的循环。其二，编码者必须区分“因学生产物而改变”与“教师本来就计划改变”，据以校验的材料是**课前锁定的教案版本与时间戳**，辅以教师刺激回忆访谈；已经发生但无法追溯来源的变化，不得计入采纳。

（三）结果变量与分析。 即时正确率不能担任唯一结果。保持用延迟提取测量；近迁移使用结构相同、表面不同的任务；远迁移要求跨表征、跨功能或跨情境重组；未来学习准备可测学生利用新资源与新反馈的效率；问题生成质量可从可研究性、区分力、证据要求与概念边界编码。还应测量认知负荷、心理安全、学生话语分布与教师工作量，以识别收益是否以超载或不平等为代价。

数据具有学生嵌套于班级、序列嵌套于单元的结构，可用多层模型估计层距的增量效度，用序列分析检验“产物—采纳—迁移”的时间顺序。对平台环境，可比较日志中不同分支的层级分布；对质性材料，可采用过程追踪，寻找“若没有这一份学生产物，该层的改变就不会发生”的反事实证据。所有分析都应报告编码者一致性、缺失模式、教师固定效应、替代编码方案与敏感性分析。

（四）一种可执行的层级审计。 学校不必等大型研究完成才开始观察。一次不用于考核的单元审计，可抽取三次连续课堂与两版作业，逐条回答：本单元产生了哪些学生产物；它们各自挑战的是哪一层；实际被接在哪一层；哪些被拒绝、理由是什么；被接住之后，那一层此后有没有再被触及。审计结果不宜压成单一排名，而应形成一张**接口图**：学生内部有改变而路径不变，教师有调整而规则不动，班级有好问题而评价不承认。这样，改革的对象就从“教师应更开放”变成具体的、可指认的某一个接口。

十九、适用边界、可能反噬与伦理条款

第一，层级不能替代知识质量。 若课程目标错误、证据贫乏或教材失真，让学生产物更快进入更高层只会加速偏离。任何规则层采纳都需要学科知识、反例与外部标准校准。

第二，采纳不等于全量接受。 课堂时间有限，学生产物相互冲突，部分建议不准确、不安全或不公平。专业判断包括选择、延迟与拒绝；关键是标准可说明，并给重要产物保留复核机会。若把本文读成“每个意见都应立即改变课程”，教师责任将被取消，课堂会被话语最强者占据。

第三，层级可能扩大不平等。 善于表达、熟悉学校语言或与教师文化资本相近的学生，其产物更易被接在较高层。研究与实践必须同时检查采纳机会与采纳层级在性别、语言、成绩与社会背景上的分布。一个层距很小而规则层采纳全部来自少数强势学生的课堂，只是把教师中心改成了地位中心。匿名问题、轮流角色、书面草稿与教师主动读取沉默者的证据，都是减少偏差的接口。

第四，稳定性有不可替代的价值。 乘法事实、发音动作、实验安全、医学急救与紧急疏散需要高一致性与自动化；在这些环节，规则层开放会造成风险。此时层级判断应集中在表征与路径两层，规则层保持关闭是正确的，不应被本文的读数记为缺陷。

第五，新手与熟练者需要不同的接口结构。 新手缺少足够图式时，过多选项会占用工作记忆；随着知识增长，固定支持又可能变成冗余。因此接口应随能力变化：先让学生的错误改变提示与例题，再让策略改变题序，最后才在证据充分时开放问题与规则。

第六，本文的读数不得直接用于惩罚性评价。 这一条必须写死三款，缺一款则本文的读数不得使用。其一，层距指的是接口的位置，不是人——同一位教师在两套不同的权限结构下会给出完全不同的层距，而他本人没有变；不得以层距对个人作出评价、排名或不利安排。其二，不得为把这个数做好看而在安全关键环节开放规则层，也不得为凑规则层采纳而人为制造质疑；此类做法一律禁止，不设例外。其三，已按这个读数作出不利安排的，必须公开编码规则、报告不确定性、允许人工复核并提供申诉途径。

第七，可能的反噬，以及我看不到出口。 本文一旦被制度采用，最省事的实现方式恰恰会摧毁它要保护的东西。层距会被做成一个可考核的指标；于是最省事的达标法是在教案上增设“规则层采纳”一栏并每单元填满一次——填写这个动作可被检查、可计分，而实际的接口一动不动。更糟的一种：为了让规则层采纳有素材，教师会预先安排一名学生在指定时点提出一个预设的质疑。要区分“真实产生的规则层产物”与“为指标而制造的质疑”，需要的恰恰是那个不进任何账本的判断。我看不到出口。唯一能写下的底线是：验收必须看那一层此后有没有被真实使用，而不是看表上有没有这一栏。

第八，采纳需要终止与遗忘。 并非所有过去的产物都应永久影响未来，否则系统会被历史噪声拖累。课堂应规定何时重新验证、何时归档、何时撤回已采纳的规则。能够停止一次失效的采纳，与能够发起一次采纳同样重要。

二十、反身性检验：本文自己被接在哪一层

本文批评课堂把学生产物接在低于其诊断层级的一层，而本文自身面临同样的风险，且风险更大。

一篇论文是一个高度封闭的轮次。它把写作过程中真实产生的未决之处——那些没想清楚的岔路、那几处推到一半发现走不通的论证——全部清除，只交出一个完成度很高的终点态。读者拿到的是产物，不是过程。

更具体的一层：本文提出的是一个**规则层**的主张——它要求改变“什么算作已被采纳”这一判断标准。而按本文自己的分析，规则层由组织持有；一篇论文不持有任何组织的规则层接口。因此本文最可能的去处，正是被接在路径层：读者记住“层距”这个词，把它写进下一篇文献综述，在既有的研究框架内多加一个变量——**采纳发生了，降了一层，而本文原本要挑战的那条判断标准一次也没有被挑战。**

本文没有能力阻止这一点，也没有理由认为自己比它所描述的那些系统更耐受。它能做的只有一件：把使用要求写死。采用本文概念的人，必须指出一个具体序列——哪一份学生产物，在什么时间，挑战的是哪一层，实际被接在哪一层，此后那一层有没有被再次触及。若只能把课堂重新描述为“层距较大”或“层距较小”，而不能产生与既有说法不同的观察与预测，这个概念就应当被放弃。

留一个本文解释不了的洞：本文说不清一个产物的**诊断层级由谁判定**。文中要求它由独立编码者在不看后续材料的条件下给出，这在研究中可行；但在真实课堂里，判断“这个反例挑战的是路径还是规则”本身就是一项专业判断，而这项判断的持有者是谁，本文没有答案。这不是一个可以靠增加编码规则解决的技术问题，它是本文框架内的一个真实缺口。

二十一、结论：从让学生正确，到让正确能够继续走

课堂上真正的学习之所以少见，不是因为课堂没有信息、活动、反馈或回应。多数课堂的回应量远比三十年前充分。稀缺的原因在另一处：**大量学生产物被接在低于其诊断价值的那一层——被回应了，而它原本要挑战的东西恰恰因此不再被挑战。**

三门学科合起来把这件事说清楚。教育学表明，学习必须经过学习者已有表征，因而学生产物具有认识论价值；化学表明，接住与放行是两件事，最强的位点接得最牢也最不放行，而轮的边界是为了可比性被制造出来的；管理学表明，绩效信号、资源分配与权限边界三者合起来，使组织对高层问题的标准回应必然是一个低层动作。三者共同指出：**采纳不是有与无，是层级；而层级由持有者决定，不由意愿决定。**

由此得到一个简明而苛刻的判据：上一轮那个产物，改变了下一轮的哪一层——它换的是一道题，还是换了“什么算数”？这个判据使我们看见被高分、顺畅与及时反馈遮蔽的结算性成功，也防止把热闹互动、密集反馈与技术自适应自动等同于学习。

本文的主张不是让学生控制一切，也不是让课堂持续变化。学习需要稳定的知识、专业的指导、成熟的例程与公共的标准。需要建立的是**选择性、分层、可审查的接口**：让有证据价值的产物进入相应层级，让错误可以改变诊断路径，让正确可以改变“什么算数”，让课程能够保存它从学生那里学到的东西。

最后给出一项写死日期的赌注：

> 截至二〇三一年六月三十日，若没有任何一项在同一制度环境内完成的多班研究表明——在控制先备知识、参与度、练习量、反馈频率、教师差异与即时表现之后，“采纳层距”仍能稳定预测延迟保持、远迁移或问题生成质量——则本文关于层级的强主张应当撤回，只保留它作为课堂描述语言的弱版本。

什么不算命中，一并写死：仅报告参与度、满意度或即时正确率改善，不算；仅在教案模板里增设“规则层采纳”一栏而接口未变，不算；仅由教师自评“我会根据学生调整教学”，不算；诊断层级与采纳层级由同一编码者给出，不算；两地或两校对比而无同质多班设计，不算；与其他改革打包而无法分离，不算。

一个值得保留的理论，也必须允许后来的证据改变它自己的标准——而不只是允许证据给它添一个变量。

参考文献

- [1] FREEMAN S, EDDY S L, MCDONOUGH M, et al. Active learning increases student performance in science, engineering, and mathematics [J]. PNAS, 2014, 111(23): 8410-8415.
- [2] DESLAURIERS L, MCCARTY L S, MILLER K, et al. Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom [J]. PNAS, 2019, 116(39): 19251-19257.
- [3] POSNER G J, STRIKE K A, HEWSON P W, et al. Accommodation of a scientific conception: Toward a theory of conceptual change [J]. Science Education, 1982, 66(2): 211-227. DOI: 10.1002/sce.3730660207.
- [4] CHI M T H, WYLIE R. The ICAP framework: Linking cognitive engagement to active learning outcomes [J]. Educational Psychologist, 2014, 49(4): 219-243. DOI: 10.1080/00461520.2014.965823.
- [5] BLACK P, WILIAM D. Assessment and classroom learning [J]. Assessment in Education, 1998, 5(1): 7-74. DOI: 10.1080/0969595980050102.
- [6] SADLER D R. Formative assessment and the design of instructional systems [J]. Instructional Science, 1989, 18: 119-144. DOI: 10.1007/BF00117714.
- [7] HATTIE J, TIMPERLEY H. The power of feedback [J]. Review of Educational Research, 2007, 77(1): 81-112. DOI: 10.3102/003465430298487.
- [8] KLUGER A N, DENISI A. The effects of feedback interventions on performance [J]. Psychological Bulletin, 1996, 119(2): 254-284.
- [9] KAPUR M. Productive failure [J]. Cognition and Instruction, 2008, 26(3): 379-424. DOI: 10.1080/07370000802212669.
- [10] SCHWARTZ D L, BRANSFORD J D. A time for telling [J]. Cognition and Instruction, 1998, 16(4): 475-522. DOI: 10.1207/s1532690xci1604_4.

- [11] BARNETT S M, CECI S J. When and where do we apply what we learn? A taxonomy for far transfer [J]. *Psychological Bulletin*, 2002, 128(4): 612-637. DOI: 10.1037/0033-2909.128.4.612.
- [12] BRANSFORD J D, SCHWARTZ D L. Rethinking transfer: A simple proposal with multiple implications [J]. *Review of Research in Education*, 1999, 24: 61-100. DOI: 10.3102/0091732X024001061.
- [13] ROEDIGER H L, KARPICKE J D. Test-enhanced learning [J]. *Psychological Science*, 2006, 17(3): 249-255. DOI: 10.1111/j.1467-9280.2006.01693.x.
- [14] FIORELLA L, MAYER R E. Eight ways to promote generative learning [J]. *Educational Psychology Review*, 2016, 28: 717-741. DOI: 10.1007/s10648-015-9348-9.
- [15] SCARDAMALIA M, BEREITER C. Knowledge building: Theory, pedagogy, and technology [M] // SAWYER R K. *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, 2006: 97-118.
- [16] OECD. Student agency for 2030: Concept note [R]. Paris: OECD, 2019.
- [17] KIRSCHNER P A, SWELLER J, CLARK R E. Why minimal guidance during instruction does not work [J]. *Educational Psychologist*, 2006, 41(2): 75-86.
- [18] KALYUGA S, AYRES P, CHANDLER P, et al. The expertise reversal effect [J]. *Educational Psychologist*, 2003, 38(1): 23-31. DOI: 10.1207/S15326985EP3801_4.
- [19] MEDFORD A J, VOJVODIC A, HUMMELSHØJ J S, et al. From the Sabatier principle to a predictive theory of transition-metal heterogeneous catalysis [J]. *Journal of Catalysis*, 2015, 328: 36-42. DOI: 10.1016/j.jcat.2014.12.033.
- [20] IUPAC. Sabatier principle; Activation energy; Transition state theory [EB/OL]. *Compendium of Chemical Terminology (Gold Book)* [2026-08-03]. <https://goldbook.iupac.org/terms/view/E02108>.
- [21] NØRSKOV J K, BLIGAARD T, ROSSMEISL J, et al. Towards the computational design of solid catalysts [J]. *Nature Chemistry*, 2009, 1: 37-46. DOI: 10.1038/nchem.121.
- [22] IUPAC. Catalyst poisoning; Autocatalytic reaction [EB/OL]. *Compendium of Chemical Terminology (Gold Book)* [2026-08-03]. <https://goldbook.iupac.org/terms/view/A00525>.
- [23] ARGYRIOU S D, GROPPA E, ROJAS-BUZO S, BORDIGA S. The role of in situ/operando IR spectroscopy in unraveling adsorbate-induced structural changes in heterogeneous catalysis [J]. *Chemical Reviews*, 2023, 123(21): 12135-12169. DOI: 10.1021/acs.chemrev.3c00372.
- [24] BODENSTEIN M. Eine Theorie der photochemischen Reaktionsgeschwindigkeiten [J]. *Zeitschrift für physikalische Chemie*, 1913, 85: 329-397. (稳态近似的最早系统表述之一)
- [25] LUDLOW R F, OTTO S. Systems chemistry [J]. *Chemical Society Reviews*, 2008, 37(1): 101-108. DOI: 10.1039/B611921M.
- [26] ASHBY W R. *An Introduction to Cybernetics* [M]. London: Chapman & Hall, 1956. (必要多样性定律)

- [27] ARGYRIS C. Double loop learning in organizations [J] . Harvard Business Review, 1977, 55(5): 115-125.
- [28] LEVITT B, MARCH J G. Organizational learning [J] . Annual Review of Sociology, 1988, 14: 319-340. DOI: 10.1146/annurev.so.14.080188.001535.
- [29] MARCH J G. Exploration and exploitation in organizational learning [J] . Organization Science, 1991, 2(1): 71-87. DOI: 10.1287/orsc.2.1.71.
- [30] FELDMAN M S, PENTLAND B T. Reconceptualizing organizational routines as a source of flexibility and change [J] . Administrative Science Quarterly, 2003, 48(1): 94-118. DOI: 10.2307/3556620.
- [31] PENTLAND B T, FELDMAN M S. Organizational routines as a unit of analysis [J] . Industrial and Corporate Change, 2005, 14(5): 793-815. DOI: 10.1093/icc/dth070.
- [32] GREVE H R. Performance, aspirations, and risky organizational change [J] . Administrative Science Quarterly, 1998, 43(1): 58-86. DOI: 10.2307/2393591.
- [33] EDMONDSON A. Psychological safety and learning behavior in work teams [J] . Administrative Science Quarterly, 1999, 44(2): 350-383. DOI: 10.2307/2666999.
- [34] NONAKA I. A dynamic theory of organizational knowledge creation [J] . Organization Science, 1994, 5(1): 14-37. DOI: 10.1287/orsc.5.1.14.
- [35] LEONARD-BARTON D. Core capabilities and core rigidities [J] . Strategic Management Journal, 1992, 13(S1): 111-125.

字数与版本：正文约二万字。本文为第二版，在第一版的基础上重写了核心构念、化学一节与分界一节。

人机分工声明：本文的问题设定、三学科材料与第一版全部文字由胡敏完成；第二版的核心构念重述、化学一节的两条推导、分界与证伪部分的重写由 Claude 完成；两处反过来削弱本文的约束由二者往复确定。全部外部文献可直接核对。

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