

课堂学习发生学

THE GENESIS OF CLASSROOM LEARNING

账本每完备一分，那一段就被压掉一分。



数值为概念示意，非实测结果

课堂学习发生学

课堂内生对象、运行条件与失效机制的可证伪理论纲领

王德生 著

德麦国际出版社

Demai International Press

出版信息

著 者 王德生

ORCID 0009-0009-8196-0030

出版发行 德麦国际出版社 (Demai International Press)

出版者注册地 新加坡 (德麦国际 Demai International Pte. Ltd.)

出版地 美国 United States

开 本 170×240 mm (16 开)

版 次 二〇二六年八月第一版

封面数值 封面上的 $0.94/1.00/0.87$ 为概念示意, 不代表本书实测结果

I S B N 978-1-970820-90-4

定 价 US\$ 22.00

版权所有, 未经书面许可, 不得以任何方式复制、转载或用于商业用途。

卷首

本书的母题

课堂里决定后果的那些东西，有一大部分是课堂自己造出来的；而课堂的账本只为“不是它造的”那些东西设了字段。于是它每天都在生产它看不见的东西，并把看不见的部分算作零。

而这并非疏漏：**账本每完备一分，那一段就被压掉一分**（在下文结构模型及其余参数不变的前提下）。

全书的形状

一个问题被问了三遍。

What——何谓真正的课堂学习：三组完全不同的学科走了三条互不重叠的路，到了同一个地方。课堂留下的是一件**可再入生成体**：它无人持有，没有结算时点，靠被交出去而继续存在。

How——课堂如何运行：终点会易主（换位区间），“下一步”的答案可以是一个、多个或没有（可倒推区间），条件有一部分是过程自造的（自建条件）。三章各带一条结构模型。

Why——学习为什么少：同一次改变的入口（空窗）、伴生物（悬置连接）、出口（冻结结构型）各有一处熄火，三处的效应**相乘不相加**。

结算——课堂的账本：六个读数、九句判据、三条推导能否收成一式、相乘还是相加，以及这本账被采用之后会怎样反噬。

本书敢让什么判自己输

六个读数若在同一门真实课上采齐后三个以上两两高度相关，本书是一次过度切分；自建条件存量若与教学资源完备度实测为正相关，全书最硬的一段作废；三项增量若与进入次序无关，相乘假设失败。约六十条证伪条件与九条写死日期的赌注，汇总在附录三。

目 录

卷 首	3
数量说明	5
前 言	6
导 论 课堂的账本，与它没有设的那几栏	10
第一编 What：发生的对象	18
第一章 可以被重新发起的那一件事	19
第二章 交到下一个人手里才算数	41
第三章 支架撤掉以后还能重新进去	61
第一编小结 三个名字，一个对象	81
第二编 How：发生的条件是自造的	85
第四章 手段取得终点资格的那一段	87
第五章 下一步该做什么，有几个答案	106
第六章 这一节课结束时，这个班手里多了什么	125
第二编小结 三个旋钮，同一只手	143
第三编 Why：发生的三处熄火	147
第七章 开着的时候没人来	148
第八章 先挡住其余的	166
第九章 还在动的时候就读了数	185
第三编小结 三处熄火，与它们为什么不产生信号	203
第四编 结算：课堂的账本	207
第十章 同一组学科被问了两次	208
第十一章 九对象指标怎样压缩为六个课程主读数	211
第十二章 九句判据合成一份问卷	217
第十三章 三条结构模型能否收成一个式子	221
第十四章 相乘还是相加	225
第十五章 账本的反噬	229
结 语 课堂用一本发现的账本，记一件发生的事	233
附录一 六读数采集手册	235
附录二 九章的学科入口、取用点与最危险的近邻	239
附录三 证伪条件与写死日期的赌注	241
附录四 十五张判断卡片	243
附录五 跨域转移：九章逐条分级	244
附录六 先导研究预注册草案	246
附录七 参考文献	248
后 记	254

数量说明

全书出现五种数量口径。它们在理论上兼容，第十一章第四节有正式区分；此处先给一张压缩链，供第一次阅读时对照。

九个理论对象 → 十三个原始观测量 → 九个对象级指标 → 六个课程主读数 → 三类外部效标

层	数量	是什么	在哪里
理论对象	9	概念层不可还原的九个对象	第一至九章各一，构念矩阵见第十一章第三节
原始观测量	13	各读数展开后的可记录量（如换位率与触发源分布记为两个）	第十一章第一节
对象级指标	9	每个对象各取一个代表性指标	第十一章第一节
课程主读数	6	按采集成本压缩后、一学期一校可采齐的那六个	第十一章第二节、附录一
外部效标	3	撤架后独立表现／四周后自发学习／条件变化后的表现	附录一、附录六

为什么九个对象只压出六个主读数：第一编三个对象的读数都需要一次专门设计的撤架任务，成本最高；第一轮采集里它们择一采作外部效标，不进主读数。这不是数量在变，是同一批东西按成本分了层。

前言

一、它是从九次撞击里长出来的

这本书没有从一个提纲开始。它从九次互不相干的撞击里长出来，而且是倒着长出来的。

最先动手的是最末一个问题：**学习为什么这么少**。一堂课上得很顺，讲清楚了，练到位了，测出来也好看，几周之后换一个表面就什么也不剩——这件事人人都遇见过，人人都能给出解释，而给出的解释一律指向学生、教师或教法。我不满意的不是这些解释不对，是它们都太容易说了。一个每天都在发生、谁都能解释、而谁也没有把它减少一分的现象，多半不是被解释错了，是被安放错了位置。

于是我换了一种做法：不去教育学内部再找一条更好的解释，而是把三门互不相容的学科同时请到同一个问题面前，看它们在哪里撞上。第一次撞出来的是**空窗**——上一轮教学打开的感受期无人到达就关上了；第二次撞出来的是**悬置连接**——为讲清一处而挡住其余，被挡住的那部分在暗处结成了无人拆除的关联；第三次撞出来的是**冻结构型**——还在重组的东西被过早读了数，并按这个读数配置了此后半年。三次撞击各自独立，事后拼起来才发现它们分别落在同一次改变的入口、伴生物与出口。

到这里问题变了。它不再是“为什么少”，而是“**那么它到底该怎么运行**”。于是同样的做法又用了三遍，撞出**换位区间**、**可倒推区间**、**自建条件**——课堂里的终点、次序与条件，各有一部分是这个过程自己临时造出来的。三章都在自己的末尾交出同一个洞：被当作给定的那些东西，有一部分是过程自己的产物；而三章谁也没把这句话写实。

要写实它，必须先回答一个更靠前的问题：**学习究竟是什么样的存在物**，才会有“它自己造出自己的条件”这回事。于是又有了三次撞击，三组学科各走一条路，最后到了同一个地方：课堂真正留下的不是一份记在某人名下、可以在某个时段结清的所得，而是一件在支架撤除之后仍可被重新进入、并且要靠交到下一个人手里才继续存在的东西。

所以本书的编次与它的生成次序正好相反。读者拿到的次序是 **What—How—Why**：先说学习是什么（第一编），再说课堂如何运行（第二编），最后说它为什么熄火（第三编）。这是给读者的次序。给我自己的次序是反的，末编最先写，首编最后写。我把这一点写在前言的第一段，是因为它解释了这本书的一个特征：它的地基是后浇的，因此地基那一编里带着后来才发现的问题的形状。

第四编是新写的，纸上此前没有。前三编九章各自成立，但九章合起来还有一件事没人做：把九个对象、六个读数、三条结构模型放到同一张表上，看它们是不是同一件事的六种说法。这一编是全书增值所在，也是最可能被推翻的一编。

二、这本书不是什么

它不是一本教学法手册。全书没有一处告诉读者应当怎样上课。它给的是读数与判据——先让人能看见，至于看见之后怎么办，那要看具体的班、具体的科目、具体的学期还剩几周。凡是能被直接改写成“五个步骤”的书，多半在写作时就已经把最难的那一段扔掉了。

它不是教育学的综述。九章里被引到的教育学文献远少于化学、细胞学、物理学与工程学的文献总和。这不是轻慢，是做法使然：本书的每一个对象都是在三门学科撞出裂缝的地方命名的，而教育学在其中只占三分之一，甚至有几章根本没有教育学在场。

它不是在批评教师。恰恰相反，全书的承重判断是：那些看不见的东西之所以看不见，不是因为谁不用心，而是因为账本里没有为它们设字段。一个人再用心，也不会去记一栏不存在的账。把结构性的失明写成态度问题，是我最想避开的那种写法。

它不主张取消课程表、评分表与评价。第七章明说“课堂不需要放弃课程表——没有表就没有课”，第九章明说“课堂不需要放弃判定——不判就无法安排”。本书主张的只有一件：**在按表推进、按判定安排之前，先花很短的时间看一眼那几栏空账。**

“发生学”三个字不指一种立场。它在本书里只是一个工序上的区分：一样东西是先有的、外来的、等着被读出来的（那是发现的对象），还是在这个过程里被造出来的（那是发生的对象）。全书要说的事，就是课堂的账本只为前一类设了字段。

三、三条纪律，和它们的代价

第一条：每一章必须落一张判断卡片。不是描述，是判断——一句可以被反对的话。写不出这句话的材料不进书。代价是本书显得薄：许多我很喜欢的现场观察，因为凑不成一句可被反对的话，被扔掉了。

第二条：每一个新名字都要配一个读数和一句不含判断词的问法。“问法不含判断词”是一条很硬的纪律：不能问“这节课有没有真正的探究”，只能问“这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的”。前一句问的是评价者的判断，后一句问的是记录。凡是只能靠专家判断来填的字段，本书一律不设。代价是本书的对象都显得琐碎——它们必须琐碎到可以被清点。

第三条：本书自己接受同一把刀。每一章末尾都有自反一节，第四编末尾有全书的可推翻条件，附录三把九章约六十条证伪条件与写死日期的赌注汇成一张表。代价是这本书给了反对者一份现成的弹药清单。我认为这是它最值钱的部分。

还有一条不算纪律的纪律：**全书不使用任何一个只在我自己的体系内部才有意义的词。**凡出现在书里的名词，要么是目标学科的通用语，要么是本书自造并当场定义、当场给读数的。理由很简单：一个必须先接受一整套前提才能读懂的判断，不构成对任何人的挑战。

四、怎么读

一小时的读法：导论 → 三篇编末小结（第一编末、第二编末、第三编末）→ 结语。这条线索完整，不缺任何一环。

半天的读法：加上附录四的十五张判断卡片，再加第十一章那张六读数总表。读完这两处，读者应当已经可以拿着表走进一间真实的教室。

通读：九章的内部结构是一样的——三家各说各的（各三节）→ 三家为何不能同时为真 → 共有前提 → 推翻它的那件材料 → 命名 → 读数与问法 → 辨别格 → 走查 → 划界 → 证伪与赌注 → 自反。熟悉这个结构之后，可以按需跳读：只关心能拿去用的，读“读数”

“判据”“走查”三节；只关心它站不站得住的，读“共有前提”“推翻材料”“划界”“证伪”四节。

每一章的末尾另有一段**接续**，一句话交代这一章的读数在附录一的哪一节可以采、它在第四编的哪一章被接手、以及它的证伪条件落在附录三的哪一处。跳读的读者可以只顺着这九段走一遍。

两份新附录的用法：附录五把九章的跨域转移逐条排成五步表，第三列写的是每一次转移**明确不能**带过来的东西——要挑本书的毛病，从那一列最省力；附录六是一份可以直接照抄的先导研究预注册草案，它是设计不是结果。

手上有具体议题的读者，建议从附录一的六读数采集手册倒着读回去：先看要采什么，再回到对应的章去看为什么是这几个数。

五、关于引用与体例

先把本书的证据等级说死，免得被读成它不是的东西：**本书提出的是一套可被测量、可被推翻的理论研究纲领，不是已经完成经验定律化的教学处方。**副标题里的“纲领”二字是照这一条定的：九个对象目前全部停在竞争假设层，书名不该比证据站得更靠前。九个对象目前全部停在“竞争假设”这一级（导论第四节给出三级措辞），没有一个达到目标领域内的直接反证。凡在封面、腰封或任何介绍文字里把它说成“已被课堂实证确认”的，都不是本书的主张。

全书引用只到概念层，一般不标页码。理由与我的另几本书相同：**一个假的精确，比一个诚实的模糊危害大得多。**正文一律只写到“某某一支的主张”，不写成看似逐页查过的样子。而附录七那份合并文献表，**我只抽查过其中四条，其余没有逐条回到原始出处复核**——哪几条抽了、抽的结果如何、未复核的部分意味着什么，附录七开头已如实写明。请在引用其中任何一条之前自行核对。

本书**不报告任何新采集的数据**。九章里的全部经验命题、读数与研究方案，都需要通过预注册研究与独立复核来检验。凡本书写“预测”的地方，都是尚未被测的；凡写“已有研究表明”的地方，指的是被引一支自己的报告，而不是对本书对象的检验。这一条我在每一章里都重复了一遍，宁可啰嗦。

书中的三条结构模型（第四章、第五章、第六章各一条）是**结构性的**，不是拟合出来的。它们的作用是把一句自然语言的主张逼到可以看出反直觉后果的地步，而不是给出可代入的数值。第十三章专门讨论它们能不能收成一个式子，结论是：目前不能，且我把不能的地方写清楚了。

六、我最没把握的四处（按信心从低到高）

其一，六个读数从未在同一门真实的课上采齐过。这是全书最大的空缺。如果采齐之后发现其中若干个两两相关系数很高，那么它们里有一部分是另一部分的代理，本书就得自己拆掉几章。我把这条写在最前面，是因为它最可能出否定结果，而它出否定结果的那一天，本书剩下的价值才看得清楚。

其二，第一编三章被我合并成了一个对象。三组完全不同的学科各自走到了几乎同一句话上——我把这件事读作稳健性证据（异源汇聚），但它也可以被读作我把三样不同的东西按到了一起。分界的责任在我，而我并不确定自己分得干净。第一编末尾专门处理这个问题，读者可以在那里判断我是不是在自圆其说。

其三，“效应相乘不相加”目前只对第三编三章成立。那三章各自独立地声明过这一条。把它扩到六项是一次外推，第十四章诚实地写明了外推的条件与它可能失败的方式。若它失败，本书最锋利的一条政策推论——单修任何一项无效——就要收回。

其四，第四编的账本方案自带一个我自己也警惕的风险。它一旦被拿去当考核，就会立刻验证本书第十五章的反噬预言。第十五章因此不是补白，是这一编的必要配件；读第四编而不读第十五章，比不读第四编更糟。

七、一个请求

本书有一道工序我自己完成不了：在同一门真实的课上，把六个读数采齐一次，并把两两相关系数原样报出来。

这件事需要一所学校、一个学期、一位愿意让课被这样看的教师，以及一个不预设结果的编码者。它花不了多少钱，附录一把它压到了一份可以照着做的手册。我请求任何有条件的人去做它，并且——这一点更要紧——在结果不利于本书时照样把它报出来。一本给了六十条证伪条件的书，如果只等来六十次确认，那说明检验的人读错了它的意思。

八、来历与致谢

书里的九章，最初都是独立成文的长篇论文。它们不是为这本书写的，是各自为了回答一个当时看起来很窄的问题而写的。把它们放进同一本书的那个念头，来得比它们晚很多，而且是被它们逼出来的：三章交出同一个洞，三章分别落在同一次改变的三个位置上，三章走了三条完全不同的路却到了同一句话——到第三次巧合的时候，再当作巧合就是懒惰了。

我另有一册以显性框架写成的教育纲领性著作，与本书分属两条线：那一册在纲领层，讨论概念本身如何被命名；本书在证据层，只处理从一节课到一个学期这一个尺度，并且把全部主张压到读数与证伪条件上。两条线可以互相印证，但本书不依赖那一册的任何一条前提；读者不必先读那一册，事实上先读它可能反而不利——本书希望被当作一份可以独立检验的东西来对待。

要谢的人在书里出现过很多次而没有名字：那些让我把课看完的教师，那些在课后老实告诉我“其实我不知道下一步该做什么”的教师。第五章那句判据——把学生此刻的状态和目标交给五位有经验的教师，看“下一步”是同一件事、五件不同的事、还是都说不出来——就是从这样一次交谈里长出来的。当时那位教师以为自己在承认一种无能。他不是。他说出的是这门手艺里最诚实、也最难被记进任何一本账的那一格。

王德生

二〇二六年八月 于新加坡

导 论 课堂的账本，与它没有设的那几栏

一、三间教室

第一间是数学教室。学生刚做完二十道同类型的题，正确率接近百分之百。教师把题里的一个常量改成变量，删掉“先求导再比较端点”那句提示，教室安静下来。学生并没有忘记公式。他们不知道现在应该把什么当作问题，也不知道凭什么判断一条解法已经够好。

第二间是工程教室。学生按流程做完了太阳能校园方案：识别需求、建模、比成本、做演示，作品达到评分表的每一项要求。教师换掉校园、预算与使用者，撤走原来的评分量表。学生仍会操作软件，却把大部分时间花在打听“老师到底要哪个答案”。设计流程留在纸面上，设计判断没有跟着学生离开原任务。

第三间是艺术教室。一次讲评让学生很受触动，他能复述教师关于空间、材料与观看者关系的全部说法。两周后，面对一件不属于原媒介、也没有教师在场的作品，他仍在等某个权威先说出“该怎么看”。体验发生过，感动发生过，判断没有获得重新组织自己的能力。

三间教室有一个共同点，而这个共同点不在学生身上，也不在教师身上：**在事情正在发生的那一刻，三间教室的账本都是满分的。**正确率、参与度、作业质量、流程完成度——每一栏都填得漂亮，每一栏都是真的。出问题的不是记账的诚实，是账本上没有的那几栏。

本书要做的，就是把那几栏找出来，给每一栏一个名字、一个可以清点的读数，和一句不含判断词的问题。

二、同一个问题问了三遍

“课堂学习”这四个字底下压着三个不同的问题，它们各自要的答案形状不同，而混着问是这个领域最常见的一种浪费。

What：何谓真正的课堂学习。它要的不是一份方法清单，是一个能把两节外表相同的课分开的判断——两节课可以有相同的正确率、相同的参与度、相同的作业质量，甚至相同的延迟测验成绩，其中一节形成了学习，另一节只完成了任务。能说清这个差别的，不该再是“更深”“更主动”这类程度词，只能是一个存在论式的回答：那节课里生成了一件此前不存在、课后仍可被重新进入的东西，而这节没有。这是第一编。

How：课堂如何运行才能让学习真正发生。它要的是一条次序——从哪儿开始、经过什么、在哪一步可以插手。这个问题的答案形状是流程，不是定义。三门学科可以各给出一条完整而互不相容的次序，它们冲突的地方比一致的地方有用得多。这是第二编。

Why：学习为什么这么少。它要的是一个驱动——是什么力量使得那件事没有发生。这个问题最容易被回答，也最容易被答坏，因为任何一句关于学生、教师或教法的话听起来都像答案。这是第三编。

三个问题必须分开问，理由不是分类的整齐，而是**它们的错误答案会互相伪装**。用一条次序去回答“是什么”，得到的是把一种教法说成学习的定义；用一个驱动去回答“该

怎么运行”，得到的是把一种病因说成处方。本书的九章里，每一章的开头都要先声明自己在回答哪一个问题，就是为了挡住这种伪装。

而三遍问完之后，出现了一件我没有预料到的事：三个问题的答案落在同一个位置上。“是什么”的答案是“一件课堂自己生成、无人持有、要靠交出去才继续存在的东西”；“怎么运行”的答案是“终点、次序与条件里，有一部分是这个过程自己临时造出来的”；“为什么少”的答案是“每一次改变的入口、伴生物与出口，各有一处产物无人记账”。三句话说的是同一件事的三个侧面。这本书之所以能成为一本书而不是九篇文章的合订本，全部理由就在这里。

三、九次撞击是怎么做出来的

九章的写法是一样的，读者若知道这个工序，读起来会省很多力气，也更容易找到该反对的地方。

第一步，请三门互不相容的学科到场。不是三门相邻的学科——那样只会得到一个平均值。要选三门对同一个问题都能给出完整答案、且答案互相否定的学科。第八章请的是教育学、化学与艺术学；第五章请的是物理学、心理学与工程学；第六章请的是细胞学、工程学与艺术学。选择的标准只有一条：这门学科在这个位置上有没有一条它自己愿意为之辩护的最强主张。

第二步，把三家的主张写到互相不能同真的地步。这一步最费力，也最容易偷懒。写成“三家侧重不同”是没有用的；必须写到教育学每天在做的那件事，正是化学工艺诊断为病的那件事。如果三家不能被逼到互相排斥，说明取的不是它们的承重主张，而是它们的场面话。

第三步，找它们共同站着而从未说出的那块地。三家互相打，说明它们站在同一块地上——争的是这块地上谁对，而不是这块地在不在。这块地就是共有前提。第八章的共有前提是“在一次改变里，被挡住的那部分不参与，参与名单是给定的”；第五章的是“下一步该做什么，是一个总能被求解的问题”；第六章的是“条件是给的”。

第四步，用其中一家自己的材料去推翻这块地。这一步是全书方法上最要紧的一条纪律：**推翻材料不能从外面借，必须是某一家自己承认的、写在它自己教科书里的东西。**第八章用的是化学的酰基迁移——用来挡住一个位点的保护基会自己跑到别处去，于是“挡住”这个动作改写了它本要固定的那张名单；第五章用的是物理学自己的适定性差异——初值问题在很弱的条件下解就存在且唯一，而两点边值问题可以无解、也可以有无穷多解；第六章用的是细胞学的生态位重塑——细胞住进去的那个环境，有相当一部分是住户自己分泌、招募、组织出来的。这三件事在原学科里都只回答一个很技术的问题，没有一件是为了教育而说的。这正是它们有力量的原因。

第五步，命名。被推翻的前提留下一个空位，那个空位上站着的东西需要一个名字。本书对命名有一条很严的自查：**删掉本书的测量办法之后，被说的那样东西是否仍然以同样的方式存在？**若答案是“仍然存在，只是难以测量”，那么本书发明的只是一个新指标，不是一个新对象。九个名字都是按这条自查留下来的。

第六步，给读数 and 一句不含判断词的问法。 名字若不配读数就只是修辞。而读数必须能由不懂本书理论的人来采——这是“不含判断词”的意思：不问“这节课有没有真正的探究”，只问“这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的”。

第七步，辨别格、走查、划界、证伪、赌注、自反。 辨别格用来说明这个对象与既有指标正交（尤其要写出**靶格**：一门每一项既有指标都满分、而本书读数为零的课）；走查是把它按在一节真实的课上走一遍；划界是与最接近的几种既有说法逐条给双向判定；证伪是写出什么样的观察会让这一章作废；赌注是把一个日期写死；自反是问这一章自己不能通过它提出的判据。

这七步走完，一章才成立。九章都走完了。而这套工序有一个副产品，值得单独说：**它使得同一组学科可以被问两次而不重复。** 本书里有三对学科组合各出现了两次——第四章与第八章都是教育学、化学、艺术学，第五章与第九章都是物理学、心理学、工程学，第六章与第七章都是细胞学、工程学、艺术学。两次取用的是同一门学科里互不引用的两个分支（化学的合成方法学与研发策略、物理的玻璃态老化与边值问题适定性、细胞的序贯诱导与生态位重塑），得到的对象完全不同。第十章专门处理这件事：**同一组学科被问两次，问出两样不同的东西，这件事本身是一条关于问题而不是关于学科的信息。**

四、关于“推翻”这个词：跨域转移的五步与三级证据

上一节说，第四步是“用其中一家自己的材料去推翻这块地”。“推翻”这个词在本书里必须先被限定，否则整套方法会被读成一次证据等级上的跳跃。

跳跃是这样发生的：学科甲、乙、丙共有前提 P；乙学科内部的某个材料否定了 P；于是课堂里的 P 也被推翻。这个推理在形式上不成立。**细胞能重塑生态位，不等于一个班一定能自造教学条件；两点边值问题可以无解或多解，不等于课堂里的“下一步”必然具有相同的适定性结构；酰基迁移会改写保护名单，不等于课堂的悬置部分必然结出对应的关联；玻璃态的物理老化，不等于被过早判定的学生一定形成同构的冻结。**

这些材料真正做到的是另一件事，而这件事仍然值钱：**它们证明了 P 不是逻辑必然的。** 在被推翻之前，三家共有的那个前提看起来像是事情的性质；材料一出，它就降格成了一个**可以不成立的假设**——而一个可以不成立的假设，就为目标领域打开了一个竞争模型的位置。

因此本书统一采用**跨域转移五步**，九章的第四步都应按这五步来读：

1. **源领域事实**——原学科确实证明了什么，写到它自己承认的边界为止；
2. **结构映射**——源领域的哪几个关系，对应课堂里的哪几个关系；
3. **不可映射部分**——哪些机制明确不能迁移，须当场写出；
4. **课堂竞争假设**——由映射生成的、与原前提竞争的新预测；
5. **目标领域检验**——什么样的课堂观察才真正支持或推翻这条假设。

按这五步，本书的九次转移**目前全部停在第四步**：第一至第四步已经走完，第五步只给出了设计，没有给出数据。附录五把九章逐条排进这张表，并标出每一章的不可映射部分。

由此，本书使用三级措辞，读者可以据此校准每一处主张的强度：

级别	措辞	含义	本书哪些地方
一级	暴露	源领域材料表明共有前提并非逻辑必然	九章第四步全部属于此级
二级	竞争	由映射生成了一个能与原前提竞争的课堂模型	九个命名对象及其读数
三级	推翻	目标领域内的直接反证	本书尚无一处达到此级

凡本书正文出现“推翻”一词而未加限定的，一律按一级读；真正的三级只能由预注册研究给出。这条限定不削弱本书的主张，它只是把主张放回它实际站着的那一格。一个被降格为“可以不成立”的前提，仍然足以让一整套课堂记录制度失去它自以为拥有的必然性——那正是本书要做的事。

五、全书理论地图

九个对象、六个读数与三条结构模型之间的关系，读者本来要到第十一章之后才看得清。这里先给一张地图。

层	问的是什么	本书的对象	主读数	落在哪一编
对象层	课堂留下的是什么样的存在物	可再入生成体（接力=传递机制，事件再生=显露形态）	再生水平、接力率、持位可换度	第一编
运行层	终点、路径、条件从哪来	换位区间·可倒推区间·自建条件	换位率、倒推唯一性档位、自建存量	第二编
失效层	一次改变在哪三处熄火	空窗（入口）·悬置连接（伴生物）·冻结构型（出口）	对窗率、脱悬率、双向翻转率	第三编
记录层	这些东西怎样才能入账	六读数总表·九句判据·三条结构模型·反噬治理	——	第四编

读法：上面三层是被记录的东西，最下面一层是记录本身；而全书的判断是，前三层里被造出来的那部分，在第四层没有位置。

六、母题

九章合起来说的是同一句话。把它压到最短：

课堂里决定后果的那些东西，有一大部分是课堂自己造出来的；而课堂的账本只为“不是它造的”那些东西设了字段。于是它每天都在生产它看不见的东西，并把看不见的部分算作零。

这句话有三个部件，每一个都可以单独被反对。

部件一：决定后果的东西是过程自己的产物。 第二编三章各给一个位置的证据——终点（第四章：原本作为手段的那一样在某一段里取得了终点资格）、次序（第五章：下一步该做什么有几个答案，这件事本身随知识分布而变）、条件（第六章：一个班在一学期里造出来、随后被当条件用的那些可交接的物）。第三编三章各给一个反面证据——被挡住的那部分自己结出的连接（第八章）、上一轮打开而无人到达的感受期（第七章）、还在动的时候被读的那个数（第九章）。

部件二：账本只为外来物设字段。 这不是比喻。课堂的记录系统——教案、进度表、题库、评分表、成绩册、教学检查表——的每一栏，填的都是可以在事前被写下来的东西：内容点、课时、目标、达标率。可以事前写下来的，一定是先有的、外来的、等着被读出来的。一件由这个过程自己造出来的东西，在原理上无法在事前被写进表格，因为写表格的时候它还不存在。

部件三：因此看不见的部分被算作零。 不是被算作“未知”，是被算作零。这两者的差别是全书最要紧的一处：未知会引出调查，零不会引出任何东西。窗口关上时不产生任何事件，因此在原理上不会以“失败”的形式被发现（第七章）；悬置连接不在原题型里产生任何差错，因此不会被批改抓到（第八章）；一门每一项资源都齐备、执行严格、检查全过的课，学期结束时这个班手里一件自己的东西也没有——而这在任何一份评估里都是满分（第六章）。

七、这不是疏漏：完备度反比

到这里为止，母题还可以被读作一句抱怨：账本没记全，那就把它记全。本书要说的比这狠一层：

账本每完备一分，那一段就被压掉一分——在下文给出的结构模型及其余参数保持不变的条件下。

这一条不是修辞，第六章把它写成了一个式子。设一门课在单位时间里出现“需要一个工具而现有工具不够用”的场合数为 N ，外来条件（教材、题库、评分表、规程、平台功能）对这些场合的覆盖度为 c ，缺口被真的做成可交接工具的转化率为 α ，工具的失效比例为 δ 。于是自建条件的稳态存量是

$$A^* = \alpha(1-c)N / \delta$$

这个式子推出三件正文没说过的话。其一，在 α 、 N 、 δ 不随覆盖度变化这一条件下，存量与外来条件覆盖度线性反相关；放开这一条件后，高覆盖度区间的压制仍然必然，中低区间则可能上升、下降或呈倒 U 形，而且这是设计的必然结果，不是执行不力：每提高一分覆盖度，就按同一比例压掉一分自建。其二， $c \rightarrow 1$ 时 $A^* \rightarrow 0$ ——一门“完备”的课程在结构上不可能有自建条件。所谓完备，定义上就是每一个需要工具的场合都已经有了工具了；因此追求课程建设的完备本身，就是把这个量归零的过程，而这个过程的每一步在任何评估体系里都记为进步。其三， α 与 c 是两个独立的旋钮，本书主张的是抬高 α （允许自造、给一个存放的地方、让它被下一次真的用上），不是压低 c ——这三件事都不减少任何外来资源。

我把这三条推论看作全书最硬的一段，理由是它们把辨别格里的那个靶格从“一种不好的状态”改写成了“一个被正确的努力必然到达的位置”。它是尽职的产物，不是失职的产物。凡是把结构性后果读成态度问题的分析，都会在这里失效。

同样的形状在另外两章里以不同的面貌出现。第四章推出：最优的换位时长与目标端的不确定性无关，而与交还的成功率、转化效率相乘——**换位若不交还，就不是换位，是目标被吞掉**。第五章推出：错配的代价在两个方向上不对称——在唯一区误用推进只是有界的效率损失，在多解区误用倒推却是能力损失，而且这个损失在原场看不见；由此得到一条与教师直觉相反的默认规则：**在长期迁移与可选择性权重较高的内容上，档位不确定时应默认推进**，而且因为经验会系统性地高估唯一性，判据必须问五位教师而不是一位。

三条推导各自独立，形式上互不相同。它们能不能收成一个式子，第十三章正面处理，答案是日前不能——我把三个特例与不能收拢的地方都写清楚了。

八、与最接近的六家分界

一个新名字如果不能与最接近的几种既有说法分开，它就只是那些说法的一次换词。以下六条是全书层面的分界，每一条都给判定形式——不是“我们侧重不同”，而是一个具体的观察，在那里两家给出方向相反或内容不同的判断。

其一，与发生认识论。皮亚杰一支讲的是认识结构在个体身上的阶段性发生，承载物是个体，时标是年。本书讲的是一节课到一个学期这个过程自己产出自己的条件，承载物是一个班这一段时间里的关系与物，时标是周。判定：本书的读数在同一批人、同一学期内应当有剂量-反应关系，而且换一个班就完全不同；阶段则跨个体稳定、跨班级重复。若本书的自建条件存量在同一学校不同班级之间高度一致，那说明它测的是学生的发展水平而不是这个过程的产物，本书这一条就得让位。

其二，与教学的发生原则（按知识的历史发生次序来教）。那一支是一条教学处方：既然某个概念在历史上是这样长出来的，就按这个次序教。本书不是处方，是读数；而且方向不同——那一支讲**内容的发生史**，本书讲**课堂运行自身的发生**。判定极硬：**一门严格按历史发生次序讲授、教师讲得极好的课，本书的六个读数可以全部为零**。若有人证明按发生次序教必然抬高本书的读数，那么本书就是那条处方的一个副产品，应当合并。

其三，与隐性课程。那一支讲学校在正式课程之外传递了价值、规范与社会关系，内容通常是权威、性别、阶层。本书讲的是这个过程造出来的、可以指着看、可以清点、可以交给下一届直接用的物与关系。判定：隐性课程的内容换一所学校大体重复，本书的自建条件换一个班就完全不同；前者可以在一份社会学分析里被事前预期，后者在事前写不出来。

其四，与情境学习和分布式认知。那一支讲认知分布在工具、符号与共同体上，因而不能只在个体头脑里结算——这一条本书完全继承。分离线在**工具从哪来**：那一支讲认知分布在**既有的工具上**，本书讲这些工具有一部分是这个过程**自己造的**。判定：给两个班相同的外部工具，允许自造的那个班在工具撤除后的独立表现应当更好；若两班在撤除后无差异，本书这一条就退化为分布式认知的一个特例。

其五，与指标腐蚀（一个指标一旦成为目标就不再是好的指标）。那一支讲的是度量这个动作会腐蚀所度量之物，病灶在度量的品质与动机。本书讲的是更前面的一层：什么样的东西才可能成为一个字段。病灶在字段结构由结算规则派生，与度量品质无关。判定最干净：取一本从不用于任何考核、无人有动机作假的账本——比如一位教师纯为自己使用的备课记录——本书预测那几栏仍然为空。若在无考核压力处那几栏自动被填上，则本书这一层不成立，回到指标腐蚀足矣。

其六，与形成性评价与学习进阶。那一支要把评价做得更细、更及时、更贴合学习路径。本书对它没有反对，只有一条它自己推不出的后果：细化本身会按比例压掉自建条件。判定：给两个班同样的内容与课时，一个班配全套细目表、评分表与题库，另一个班只给目标与截止时间，学期末清点两个班各自手里可交接的物。本书预测后者更多而当堂表现无差异；若前者也一样多，第六章那个式子就错了。

这六条里，最要紧的是第五条。它是本书与整个“评价改革”传统之间的分界线：那个传统追问账要怎么记才准，本书追问什么东西有资格成为一栏账。

九、全书的形状

第一编（What）：发生的对象。 三章，三组完全不同的学科各走一条路，最后到了同一个地方。第一编末尾把三章合并为一个对象、三条论证路径、三个读数，并处理由此带来的最大风险：异源汇聚究竟是稳健性证据，还是我把三样东西按到了一起。

第二编（How）：发生的条件是自造的。 三章，分别处理终点（换位区间）、次序（可倒推区间）、条件（自建条件）。三章各带一条结构模型，是全书结构性最强的一编。

第三编（Why）：发生的三处熄火。 三章，按入口（空窗）、伴生物（悬置连接）、出口（冻结构型）排列，不按写作次序。三章各自独立地声明过同一件事：三处的效应相乘不相加——入口对不上，另两处做到最好也没有输入；入口对上而伴生物无人拆除，出口读得再准也只是准确地读到一个带着杂物的结果。

第四编（结算）：课堂的账本。 六章，纸上此前没有。第十章处理“同一组学科被问了两次”；第十一章把六个读数放到同一张表上，统一量纲并给出正交性与代理风险；第十二章把九句判据合成一份不含判断词的问卷与编码规程；第十三章处理三条结构模型能否收成一个式子；第十四章处理“相乘还是相加”，并给出全书唯一一条与教改惯行方向正面相撞的政策推论；第十五章处理这本账一旦被拿去考核会发生什么。

结语回到母题：课堂用一本发现的账本，记一件发生的事。

七个附录：六读数采集手册、九章近邻占位者总表（反复出现的那几位才是本书真正的对手）、约六十条证伪条件与写死日期的赌注汇总、十五张判断卡片一览、跨域转移九章逐条分级、先导研究预注册草案、参考文献。

十、本书敢让什么判自己输

按第三条写作纪律，本书在这里先交出自己的可推翻条件。逐章的证伪条件在附录三，此处只写全书层面的三条。

其一，**读数之间的相关**。若六个读数在同一门真实课上采齐之后，其中若干个两两相关系数持续高于零点六，那么它们里有一部分是另一部分的代理，本书就不是九个对象而是三个或两个，应当自己拆掉相应的章。这一条最可能出否定结果。

其二，**完备度反比的方向**。若在同一所学校、同一批教师、内容量相当的课程之间实测发现自建条件存量与教学资源完备度**正相关**，那么第六章那个式子的方向错了，第五节整段作废，全书最硬的一段随之作废。

其三，**相乘还是相加**。若把六项都做到位与只做一项的效果差别，在实测中呈相加而非相乘，那么第十四章那条政策推论——单修任何一项无效——必须收回，而本书从一个结构性判断降格为一份清单。

还有一条不能由本书自己完成的检验，写在这里作为提醒：若有人用“**你没有理解本书的框架**”来回应一条关于读数的技术意见，**请把这件事指出来**。一本主张“看不见的东西被算作零”的书，如果对自己的空账采取同样的处理方式，那它就在当场演示自己批评的那件事。

十一、几个约定

“**课堂**”在本书里指从一节课到一个学期这一段，承载物是一个班、一门课、一位教师与他们之间的关系与物。再大的尺度（一所学校、一个学制、一次改革）本书不作主张，凡越出这一尺度的地方都会当场声明。

“**读数**”指可以由不懂本书理论的人按手册采集的量。“**判据**”指一句不含判断词的问法。“**靶格**”指一门在既有指标上表现良好而本书读数为零的课——每一章都必须写出自己的靶格，写不出的章不成立。

“**发生**”与“**发现**”在本书里是一对工序上的区分，不带任何价值判断：一样东西如果是先有的、外来的、等着被读出来的，处理它的是发现；如果是在这个过程里被造出来的，处理它的是发生。一节课当然两者都有，而且必须两者都有。本书的全部主张只是：**账本只为前一类设了字段**。

最后一个约定关乎语气。本书不报告任何新采集的数据，全部经验命题都需要预注册研究与独立复核来检验。因此凡书中写“预测”处，都是尚未被测的；凡写“应当”处，指的是本书的判断而非已有的证据。这一点我在每一章里都重复了一遍，宁可啰嗦——因为本书批评的那件事，恰恰是把一个尚未成形的东西当作已经结清的东西来处理。

第一编 What：发生的对象

本编要回答的是：一节课结束之后，究竟有没有一样东西被留下来？如果有，它是什么样的存在物？

三章分别由三组完全不同的学科撞出来，走了三条互不重叠的路，最后到了几乎同一句话上。本编末尾处理这件事究竟意味着什么。

课堂研究里有一个位置长期空着。那个位置不在“怎样教更有效”这一侧，也不在“学生学到了多少”这一侧，而在更靠前的地方：**两节外表完全相同的课，凭什么说其中一节留下了东西、另一节没有？**

这个问题不能用程度词回答。说这一节“更深入”“更主动”“参与度更高”，等于承认两节课在同一根轴上，只是刻度不同；而现场经验一再表明它们不在同一根轴上——一节参与度极高、当堂正确率极高、延迟测验也不差的课，仍然可能什么都没留下，而这件事在任何一份记录里都不会以“失败”的面目出现。要把这两节课分开，需要的不是一把更细的尺子，是一个**对象**：一样东西，它在其中一节课里被生成了，在另一节里没有。

本编三章各自去找这样一个对象，用的是三组互不相容的学科。

第一章从教育学、工程学与艺术学出发，这是三组里离课堂最近的一组——三门学科都直接处理“人怎样学会做一件事”。它撞出的对象是**学习事件再生**：课堂真正留下的不是一份内容，而是这个班重新发起一次学习的能力，它可以拆成问题、修订链与判断场三样东西的再生。

第二章换成认知科学、经济学与艺术学，距离拉开一步：经济学根本不关心课堂，它关心的是稀缺资源在制度里怎样配置。它撞出的对象是**接力再生事件**：学习不在某个人身上结清，它必须被交到下一个人手里才继续存在；三家共同默认而从未说出的前提是“学习有主人，也有结算时点”。

第三章换成物理学、心理学与艺术学，距离最远：物理学连“学习”这个词都不用，它谈的是一个系统的响应地形怎样被改变。它撞出的对象是**可再入生成体**：课堂生成的那样东西可以在支架撤除、情境改变、教师不在场之后被重新进入，而它并不持续显露——不显露不等于不存在，这是它与“记住了”最要紧的分别。

三章的次序是**由近及远**。这不是编排上的好看，而是本编的论证形状本身：**离得越远的那一组学科，越没有理由撞出同样东西；而它们撞出了**。第一章的三家都是教育的邻居，它们说到一起去并不奇怪；到第三章，三家里已经没有任何一家是研究人怎样学习的了，而结论仍然落在同一个位置。这三次到达是各自独立完成的，写作时也没有互相参照。

本编因此要冒一个风险，并且必须在末尾正面处理它：**三个名字指的是不是同样东西？**如果是，一本书里不能有三个名字并列；如果不是，三章之间的分界必须写到可以被检验的地步。这是全书作者本人信心最低的两处之一，读者可以在本编末尾判断我处理得是否干净。

第一章 可以被重新发起的那一件事

判断卡片：课堂留下的不是内容，是这个班重新发起一次学习的能力；而它的水平由问题、修订链、判断场三项中**最低的一项**决定，不由平均决定。

三个理论入口：教育学 × 工程学 × 艺术学

提要。课堂学习长期被知识保持、任务成绩、活动参与、流程完成或深刻体验所替代测量，但这些指标都可能在学生没有获得独立学习能力时显得十分漂亮。本章比较教育学关于策略、迁移与“先解题后教学”的研究，工程学关于真实设计、目标——标准——可靠过程——知识资源的研究，以及艺术学关于身体化体验、审美转化和工作室批评的研究。三条理论线索分别看见了学习的结果、路径与条件，却共同预设学习事件已经由课程、教师或平台构成。本章据此提出：课堂学习不是知识保持，也不是流程熟练，也不是体验强烈，而是学生在原有支架撤除和情境改变后，能够重新构成问题、修订链与判断场的学习事件。文章提出“支架撤除——情境变异——事件重构——第三方复核”的观察协议，以及一个不含价值判断词的检验问题：撤掉原例题、教师提示、评分量表与固定材料后，学生能否在陌生情境中提出一个可检验的问题，依据新证据完成两次修改，并让另一组学生复现这条判断链？本章进一步区分课堂表现与学习事件再生，讨论数学、工程、艺术、人文及生成式人工智能课堂中的应用、边界、伦理风险和可证伪预测。

一、引言：课堂越成功，学习越可能被遮蔽

一间数学教室里，学生刚完成二十道同类型题。正确率接近百分之百，板书整齐，课堂节奏流畅。教师把题目中的一个常量改成变量，并删去“先求导再比较端点”的提示，教室突然安静。学生并非忘记了公式；他们不知道现在应该把什么当作问题，也不知道凭什么判断某条解法已经够好。

一间工程教室里，学生按流程完成太阳能校园方案：识别需求、建立模型、比较成本、制作演示。作品达到评分表要求，团队协作也很顺利。教师随后换掉校园、预算和使用者，并撤去原评分量表。学生仍能操作软件，却把大部分时间花在寻找“老师到底要哪个答案”。设计流程留在纸面上，设计判断没有随学生离开原任务。

一间艺术教室里，一次作品讲评引发强烈情绪。学生说自己“第一次真正看见了作品”，也能复述教师关于空间、材料与观看者关系的语言。两周后，面对一件不属于原媒介、也没有教师在场的作品，学生仍等待某个权威先说出“该怎么看”。体验发生过，感动发生过，判断却没有获得重新组织自身的能力。

这三个课堂都可以被现有评价体系判为成功。第一个课堂有成绩，第二个课堂有过程与作品，第三个课堂有参与和转化感。问题不在于这些结果是假的，而在于它们无法回答同一件更严格的事：学习究竟属于谁？它属于学生，还是暂时属于题型、流程、教师、评分表、同伴气氛与材料配置共同搭成的课堂装置？

传统课堂研究常把“学习”定位为个体内部状态的改变：知识增加、记忆保持、策略改善、动机增强、观点转化。社会文化研究把视野扩大到参与、身份与共同体，但在多数

评估中，任务、问题、标准和证据形式依然由研究者或教师预先给出。学生被观察的是如何进入一个已经存在的学习事件，而不是能否使下一次学习事件重新存在。

本章提出一个更苛刻的命题：课堂的使命，不只是让学生在当前事件中表现更好，而是使学生获得重建下一次学习事件的能力。这里的“重建”不是孤立地从零开始，也不是拒绝教师、知识传统、工具和共同体。相反，它指学生能够在新条件下重新组织这些资源：把模糊的不适转成可检验的问题，让行动对证据产生回应，并构成一个他人可以进入、质疑和复核的判断场。

因此，真正的课堂学习不能在课堂最顺利的时候被直接看见。它要在原有支架被有计划地撤去、情境发生结构性变化、原评价语言不再自动适用时，才显露出来。课堂表现回答“在这个场景里完成了什么”；学习事件再生回答“当这个场景不再提供答案时，学生还能使什么重新发生”。两者有关，但不是同一件事。

本章的论证分为五步。第一，分别考察教育学、工程学与艺术学近年的承重理论。第二，说明三种理论各自把握了课堂学习的一面，却在同一具体案例上给出不可互换的判断。第三，揭示三者共同站立的前提：学习事件被当作预先存在的容器。第四，提出“学习事件再生”及其三个构件。第五，给出可操作的观察协议、二维辨别格、课堂设计方案、与既有概念的经验分界、伦理边界和可证伪研究设计。

二、问题的重新定位：不是“学生学到了什么”，而是“学习能否再次发生”

“学生学到了什么”通常要求列出知识、技能、态度或价值观。“学生学得如何”通常要求分数、速度、准确率、作品质量、课堂行为或自我报告。“学生为什么学会”则会追问反馈、认知负荷、先备知识、动机、互动和环境。这些问题都重要，但它们共享一个语法：学习已经被分割成某些可交付的内容，学生是这些内容的持有者，课堂是内容进入学生的通道。

这个语法难以解释一种日益普遍的现象：学生可以获得几乎全部可交付结果，却没有获得下一次学习的发起权。生成式人工智能使这种现象格外清楚。一个学生可以交出结构完整、引用充分、语言成熟的论文，可以展示多个版本的“修改过程”，甚至能依据量表逐项解释作品优点；然而，问题可能由模型确定，修改方向由模型给出，评价标准也由模型生成。所有可见产物都存在，学习事件的发起与组织却发生在学生之外。

这并不只是人工智能带来的问题。传统课堂同样能够提供完整的问题、步骤、范例、标准、材料与反馈。学生在这些条件中表现优异，并不等于能够在条件改变时重新组织它们。换言之，课堂可能把“会在一座已经搭好的桥上行走”误判为“会判断何时需要桥、桥该通向哪里、桥坏了该怎样改、谁有资格说它安全”。

本章因此把学习的最小分析单位从“个体获得的内容”移动到“可重新生成的事件”。一个学习事件至少包含三件事。第一，某种差异被构成为问题：不是所有困惑都成为问题，也不是所有教师提出的问句都被学生真正接管。第二，行动形成修订链：尝试必须因证据而改变，而不是仅仅多做几遍。第三，判断获得一个可进入的场：什么算证据、什么算改进、谁可以质疑、材料如何留下、结论如何被推翻，都有可追踪的安排。

若这三件事全部由教师提供，学生可能在事件中学习，也可能只是借用事件完成任务。若学生能够在教师退出、例题撤去、材料改变、同伴更换或工具关闭后，再次组织出这三件事，学习便不再只是一次课堂结果，而成为能够继续发生的能力。本章把这种能力称为“学习事件再生”。

这个命名刻意避免把学习理解为孤立的心智财产。“再生”不是复制原课堂，而是在约束改变后生成结构同源、内容不同的新事件；“事件”不是瞬间体验，而是问题、行动、证据、标准、他人和材料共同构成的一段可追踪过程。学生不必完全独立，却必须能够改变资源之间的关系。真正的学习者不是“不需要任何人”的人，而是能够知道何时需要谁、需要什么、为何需要，以及这些资源进入之后改变了哪一条判断的人。

三、教育学的最新回答：学习不是偏好匹配，而是策略、迁移与延迟显露

当代教育学已经越来越警惕把学习等同于课堂当下的顺畅表现。Hattie 与 O'Leary (2025) 对十七项元分析的再考察指出，所谓“按照学习风格匹配教学”的因果证据极弱，而许多看似支持学习风格的研究实际上测到的是策略差异。他们主张把注意力从固定偏好移向能够随任务复杂度和学习目标调整的有效策略。这个转向意味着，学习不能由“学生喜欢怎样学”来定义，而要看学生能否选择、组合和改变方法。

“表现”与“学习”的分离提供了第二条线索。练习中即时正确、课堂上反应迅速、教师提示下顺利完成，都可能只是暂时表现；保持、迁移和独立应用往往在延迟测验或新情境中才显露。Soderstrom 与 Bjork (2015) 系统梳理了这一差异：某些让练习显得困难的条件，反而可能产生更持久的学习；某些让练习显得流畅的条件，则可能制造“已经学会”的错觉。课堂若只在内部结算，就容易把支架的作用记在学生名下。

“先解题后教学”与生产性失败把这条分离推进了一步。Sinha 与 Kapur (2021) 的元分析显示，在符合关键设计条件时，让学生先生成多种表征和解法、经历不完整或失败的尝试，再进行有针对性的教学，通常能改善概念理解与迁移。这里最重要的并非赞美失败，而是承认：学生在标准答案出现之前生成的结构，会改变后续解释可以落在哪里。学习不是教师把正确结构放进去，而是学生的初始生成与规范知识发生重新组织。

近年的课堂互动研究又提示，学习不能只从某个静态指标读取。Li 等人 (2024) 发现，翻转课堂中的结构化师生互动与学习结果和脑间同步有关。Xu、Zhang 与 Zhang (2026) 在真实课堂中进行了三个学期的纵向超扫描，收集一百零七名学生、三百九十三节语文和数学课的四千一百七十五份脑电记录。动态模型中，师生耦合的增加能够预测之后的成绩改善；静态模型中，同一时点的较高耦合却可能与较低数学成绩相伴，研究者将其解释为教师对困难学生投入更多注意的选择效应。这个结果提醒我们：单次截面可以把支持强度误读为学习水平，变化轨迹比一个“高值”更有意义。

元认知研究同样把重点放在校准，而不是自信本身。学生可以感觉自己理解得很好，却不能准确预测自己在新任务中的表现；也可以因任务艰难而低估已经形成的结构。高质量学习需要学生比较预期、行动结果与证据，并据此调整策略。于是，教育学给出的强答案可以概括为：学习是能够在时间延迟、任务变化和支架减少后仍然有效的策略性改变。

然而，这个答案仍有一个边界。教育学的迁移测验、延迟测验、未来学习准备测验和元认知任务，通常由研究者提供新问题、材料和评价标准。即使学生被要求选择策略，也

是在一个已经被定义好的任务里选择。生产性失败中的“失败”也由教师预先设计，规范概念与汇总教学在后面等待。教育学非常善于区分“在给定事件中表现”与“在给定新事件中学习”，却较少直接测量学生能否构成这个新事件本身。

教育学自己已经掌握了推开这扇门的材料：即时成功不是学习的可靠证据；学生在规范教学之前的自生成会改变学习；静态高值可能只是教师补偿性投入；元认知的核心是对策略与证据关系的校准。把这些材料放在一起，便得到一个教育学尚未充分展开的推论：若任务、标准和问题始终由外部预先给出，所谓“独立迁移”仍可能只是对另一个预制事件的适应。

四、工程学的最新回答：学习是目标、标准、过程与资源的连续重组

工程教育提供了一种与结果测量不同的语言。工程实践面对的不是答案唯一、条件封闭的习题，而是目标有争议、约束会变化、利益相关者彼此冲突、知识不完整且失败具有成本的开放问题。一个工程方案是否成立，不只看最终产品，还要看目标如何形成、标准如何取舍、过程为何可靠、知识从哪里来，以及方案能否在反馈中继续改进。

Gao 等人（2025）研究工程本科生担任中学生设计导师时的认识变化，提出由目标、理想标准、可靠过程和知识资源构成的 AIRK 框架。研究中的学生报告了六类转变：从学习理论知识转向把知识用于社会需要，从完成作业转向以使用者为中心，从寻找最优解转向接受持续改进，从完成代码转向创造性思考，从技术中心转向沟通与协作，并形成更相互连接的 STEM 知识。这个框架的重要之处在于，它不把“知道”看成静态拥有，而把知识放进目标、判断标准、过程和资源相互制约的设计活动中。

Tang、Jiang 与 Xie（2025）在“太阳能化你的学校”课程中观察七年级学生如何平衡标准与约束。三十八名学生在五天内为校园采用太阳能制定方案并使用技术工具模拟。研究识别出三种不同的设计取向：面向实践、面向性能和面向无关目标。把注意力优先放在真实标准与约束上的团队，设计表现更好。这一发现表明，参加了相同活动、使用了相同工具、走过了相似步骤，并不意味着在执行同一种设计。决定学习质量的不是流程是否出现，而是流程围绕什么目标运转、哪些约束进入判断、哪些证据被忽略。

工程学因此提供了一个强烈的路径观点：学习不是一次正确，而是从目标设定、需求澄清、模型建立、方案生成、原型制作、测试、失败诊断到再设计的可修订序列。任何单一环节都不能独立代表工程理解。尤其是“持续改进”取代“最优答案”之后，学习的价值不再由终点一次结算，而由系统在新约束进入时是否能够改道来判断。

但工程学也暴露出流程主义的危险。设计周期可以被制作成海报，标准可以被写进模板，用户需求可以由教师预填，迭代可以退化为“换颜色——再提交”。学生可能忠实执行全部步骤，却从未决定问题为何值得解决，也未曾承担标准冲突带来的判断。日志越完整，越容易制造“过程已经发生”的印象；若每次修改都只是为了让评分表多一项打勾，所谓迭代只是顺从被分成若干阶段。

工程教育自己的材料已经指出这一点。AIRK 框架把“目标”和“理想标准”与“可靠过程”并列，意味着过程不能独立证明认识变化；“知识资源”被增加为新维度，意味着设计者必须判断还缺什么知识、从谁那里取得、何时足够；从作业导向转向用户导向，意味着问题并非天然存在于题目文字中，而是在与实际使用者和社会需要的接触中重新形成。

Tang 等人的“无关目标”类型更直接说明：同样有设计活动、同样有迭代、同样有作品，路径仍可能围绕错误的中心运转。

于是，工程学把学习从结果推进到路径，却也迫使我们继续追问：路径是谁的？若目标、标准、用户、约束和失败解释全部由教师预装，学生拥有的是流程操作权，不是学习事件的组织权。真正的工程学习不只是执行设计，而是能够在新情境中重新确定何为问题、何为可靠、何为代价、何为足够的证据，并让这些判断经受他人和现实的反驳。

五、艺术学的最新回答：学习是身体、感知与共同判断场的改变

艺术学习对“知识是否被记住”和“步骤是否被完成”保持天然怀疑。一个学生可以背出透视法、和声规则或戏剧结构，也可以完整执行创作流程，却仍未真正看见材料、声音、身体与观看者之间正在发生的关系。艺术教育把学习放到感知被扰动、身体重新协调、情绪获得形式、意义在作品与共同体之间生成的过程里。

Pizzolante 等人（2024）对审美经验转化力量的系统综述提供了重要证据。研究团队从三万九千四百四十项记录中最终纳入二十三篇明确考察艺术与审美长期心理影响的同行评议研究。综述确认艺术与审美经验具有引发转化的潜力，并归纳出若干关键成分：促成条件与预期、认知不协调、顿悟或洞见，以及经验之后的持续影响。这个结果一方面说明艺术经验可能改变世界观、自我理解和行动；另一方面也显示，关于长期转化的严格证据远少于关于即时感受的研究。强烈体验并不自动等于持久学习。

身体化学习研究进一步说明，理解不只发生在语言命题中。动作、姿态、节奏、触觉、空间方向和与环境的协调，能够参与概念形成与身份改变。表演艺术训练中的“知道”经常是一种身体可执行、却难以完全口述的判断。一个舞者调整半拍呼吸，一个画者改变与画布的距离，一个演员让沉默延长两秒，这些修改不是对规则的简单应用，而是身体在材料反馈中重新组织注意。

工作室批评则把艺术学习放进共同判断。Cleary Rodrigues（2025）研究美术工作室中的“crit”，讨论作品意义如何在同伴反馈、共识、判断与异议之间形成。作品并非先有固定含义，再被别人准确读出；意义在作品做了什么、如何与作者和观看者沟通、哪些分歧不能被消除的过程中出现。由此，学习不是私人感受加深，而是学生能够让让自己的选择进入一个可质疑、可回应、又不要求消灭差异的公共场。

艺术学给出的强答案是：学习发生在感知、身体、情绪、材料和共同体的关系被重新编排时。它反对把学习压缩为可复述内容，也反对把判断还原为统一量表。一个人可能不会立即说清学到了什么，却在之后的观看、行动和创作中表现出新的注意结构；反过来，一个人可以说出全部术语，却仍以旧方式观看。

然而，艺术学同样有一处难题。审美转化高度依赖作品、空间、教师语言、同伴气氛和个人生命史。一次工作室批评可以很深，却未必能在另一个媒介、另一个共同体或没有原教师时再度发生。身体化技能可以在熟悉工作室中顺畅执行，离开特定节奏、镜面、伙伴和音乐便失去组织。艺术教育善于指出环境与关系是学习的一部分，却容易把“在这个场中发生过转化”当作“学生已经获得再造此场的能力”。

艺术研究自己也提供了推翻这种等同的材料。系统综述中，关于长期后效的严谨研究十分有限；审美转化需要促成条件，说明体验不是个体内部自动发生；工作室批评中的意义依赖共识与异议的组织，说明判断不属于单一权威。由此可以推出：真正的艺术学习，不是把某次深刻体验保存下来，而是学生能够在新材料、新观众和新限制中重新建立感知差异、修订作品和公共判断之间的关系。

六、三种强答案为何不能互相替代

教育学、工程学与艺术学并非彼此补充得十分和谐。它们在同一个课堂案例上会给出不同结论，而这种分歧不能靠“各说一面”轻易化解。

第一种分歧发生在成绩与路径之间。教育学若依据延迟测验和迁移任务判断学生已经掌握某个概念，工程学仍会追问：新任务的目标、标准和约束由谁确定？学生可能正确地把旧知识迁移到一个研究者设计的新问题，却从未识别真实使用者，也没有决定什么后果值得优化。教育学判为高迁移的表现，工程学可能判为高质量解题而非设计认识。

第二种分歧发生在路径与感知之间。工程课堂可以有完整的迭代记录、用户访谈和原型测试，艺术学仍会追问：学生的注意方式是否改变？材料的反馈是否真的进入判断？如果修改只为了迎合模板，流程再丰富也可能没有发生观看和意义的转化。工程学判为可靠过程的行为，艺术学可能判为形式服从。

第三种分歧发生在体验与可检验性之间。艺术课堂可能产生顿悟、身份改变和强烈的共同体验，教育学会追问这种改变能否保持和迁移，工程学会追问标准、证据与修订是否可追踪。一场让人落泪的演出可以改变学生一晚，却不能据此证明学生在之后能作出更可靠的艺术判断；一次自我报告中的“被改变”也可能无法与期待效应或群体气氛区分。

反过来，教育学和工程学的“可检验”也会遭到艺术学约束。并非所有学习都能被即时语言化，身体知识和模糊感知若被迫提前编码，可能被标准化所损伤。工程学的安全标准与教育学的量化测验也不能取代审美判断中的多义性。因此，本章不能把三者平均成“既要成绩、又要流程、还要体验”。那只会得到一张更长的检查表。

三种答案真正揭示的是三个不可互相吞并的条件：学习必须留下可延迟显露的改变；这种改变必须通过对证据有响应的修订路径形成；修订必须发生在一个由材料、身体、他人和标准共同构成的判断场中。任何一个条件缺席，另外两个都可能制造学习的外观。

七、三者共同站立的隐含前提：学习事件已经存在

尽管三种理论彼此质疑，它们却通常站在同一块未被讨论的地面上：学习事件已经由课程、教师、研究者、工作室或平台构成。学生进入这个事件之后，表现、路径和体验发生变化；研究的任务是测量变化发生了多少、如何发生、是否保持。

在教育实验中，新任务由研究者提供，正确标准由学科共同体预定，迁移距离由实验设计分类。学生可以选择策略，却很少决定什么应当成为下一道问题。在工程课程中，所谓真实问题也往往由项目主题、客户脚本和评分要求划定；学生可以权衡约束，却未必拥有拒绝一个伪问题或重建目标的权利。在艺术工作室中，开放创作比习题自由得多，但展

览制度、教师话语、同伴结构与媒介传统已经构成了什么值得被看见、什么形式可以获得反馈的场。

这个前提之所以稳定，是因为课堂必须组织。没有问题、材料、时间、标准和关系，教学无法开始。于是，“教师提供学习事件”被当成教学的自然条件，而不是一种可能随着学习成熟而逐渐转交的权力。课程研究争论哪种事件更有效，却较少问学生何时能够成为事件的共同构成者。

生成式人工智能让这个前提从幕后走到台前。模型不仅提供答案，还能提供题目、提纲、论证步骤、反例、修改意见、评分量表和模拟读者。过去由课堂装置提供的整套事件，如今可以被一个界面即时生成。若评价仍只看结果、日志或体验，学生几乎可以在没有承担问题形成、证据取舍和判断责任的情况下获得全部可见指标。

因此，三领域共同争论的并不是“哪一个指标最能代表学习”，而是学习能否由某一项外部可见结果单独读出。它们共同假定：只要学习事件已经存在，便可以分别从学生在其中的成绩、路径或体验推断学习。本章要取消的正是“事件已经存在”这一前提。

取消它，并不意味着每个学生都要发明学科、拒绝课程或在无知识状态下自主探索。它意味着把课堂学习的成熟标志从“在教师构成的事件中成功”移到“能够与他人、材料和知识传统共同构成下一次事件”。问题、修订与判断不再只是教学手段，而成为学习本身必须逐步获得的组织能力。

八、用三领域自己的材料推翻这个前提

推翻“学习事件已经存在”并不需要从外部搬来一种浪漫的自由教育立场。三领域自己的研究已经包含了相反证据，只是这些证据通常被用于回答各自内部的问题。

教育学的生产性失败显示，规范讲解之前的学生生成不是噪声，而会改变后续学习结构。若学生自己的表征、比较和失败路径能够决定教学落点，那么学习事件就不是教师讲解开始时才存在。学生的生成已经参与了问题空间的构成。表现与学习的分离又表明，当堂顺利并不能证明事件已被学生接管；延迟和变境之所以必要，正因为原事件中的支架可能代替了学生的组织。

工程学的 AIRK 框架更直接。目标、理想标准、可靠过程和知识资源不是解题前已经固定的四个输入，而是在真实设计中互相改变。使用者的出现可能重写目标，新约束可能改变“理想”的含义，失败可能暴露知识缺口，新知识又会让原问题被重新表述。若目标和标准能够在过程中生成，工程学习就不能被理解为学生进入一个既定任务后执行步骤；任务本身是认识活动的产物。

艺术学关于工作室批评的研究则说明，意义与判断并不预先储存在作品或教师那里。作品的“做了什么”在作者、同伴、观看者和材料之间的反馈中形成；共识与异议不是对既有意义的附加评论，而是意义得以显露的条件。审美转化需要促成条件、认知不协调和事后影响，也说明所谓“体验内容”不是一个早已摆好的对象，而是在关系中发生。

三条材料合在一起，得到一个三家单独都不会自然说出的结论：课堂学习的对象，不只是知识、技能、策略或体验，而是学习事件本身的构成权。学生开始学习某个内容时，

也在学习什么能够成为问题、什么变化算作证据、什么判断可以暂时成立、谁能够进入这条推理、何时需要回到材料重做。

这条结论改变了“教师撤退”的含义。撤去支架不是为了证明学生可以孤立完成，也不是把失败责任推给学生。它是一种诊断：用受控的缺席观察原课堂中哪些结构已经被学生重新组织，哪些仍然只存在于教师、模板或平台。支架撤除之后若学习事件完全消失，不能简单归因于学生能力不足；它也说明课堂把自己的组织功能留在了课堂里。

由此，新理论对象出现了：一种不等同于表现、迁移、流程、体验或自主性的能力——学生能够在条件改变后，使问题、证据响应的修订和可共同检验的判断再次组成一个事件。下一节给出它的严格定义。

九、学习事件再生：一个新的课堂学习定义

本章把“学习事件再生”定义为：在熟悉支架部分撤除、问题表面改变、材料与参与者重新配置之后，学习者能够重新形成一个由可检验问题、证据响应的修订链和可进入的判断场共同构成的活动整体。

这个定义包含三重否定。课堂学习不是知识保持，因为保持的知识可以在教师给定问题之外失去发起作用；不是流程熟练，因为完整流程可以围绕无关目标机械运转；不是体验强烈，因为深刻体验可能依赖一次性场景，不能重新组织后续行动。它们都可能是学习的成分，却没有一个能够单独承担学习的定义。

这个定义也包含三个正向构件。

第一是问题再生。学习者不是只回答一个问句，而是能从材料、情境或行动失败中识别值得处理的差异，把它限定为可检验、可反驳、与学科知识发生联系的问题。问题再生不是随意提问；它要求说明为什么这个问题比邻近问题更承重，什么证据会让问题改写，以及问题与现实约束如何接触。

第二是修订链再生。学习者能组织一段行动，使后续步骤受到前一步证据的约束。修订不是润色，也不是为了“显得做过迭代”而增加版本。至少有一处核心表征、策略、模型、构图、解释或标准因证据而改变，并且学习者能够指出改变前后的差异、触发改变的材料以及被放弃方案的理由。

第三是判断场再生。学习者能让结论进入一个他人可检验的场：证据来源可查，标准的来历可说明，反对者能找到入口，材料能保留足够痕迹，必要时可以修改标准本身。判断场并不要求所有人达成一致，也不要求一切知识都语言化。它要求分歧能够定位到作品、数据、动作、证言、模型或标准，而不是只能诉诸身份、权威或模糊印象。

三个构件不是可以相加补偿的清单。一个学生若提出了好问题，却没有根据证据修改；或反复修改，却没有任何可共享的判断标准；或有完整评价表，却没有真正由情境长出的问题，都没有形成完整事件。学习事件再生的逻辑更接近瓶颈：最弱的一环决定事件能否继续。

可把一个学习事件简记为：

$$\mathcal{L} = \langle Q, V, J \rangle$$

其中，Q 表示问题的形成，V 表示对证据有响应的修订序列，J 表示判断场。给定一次情境变换 T 和一组撤除支架 A，学习者生成新的事件：

$$\mathcal{L}_T = G(T, A; Q, V, J) = \langle Q_T, V_T, J_T \rangle$$

若 Q、V、J 中任何一项仍完全由原课堂复制，事件便只是迁移了表面形式，而未完成再生。为便于研究，可以分别对三项作 0——3 级编码，再用最低项而非平均项表示再生水平：

$$R_T = \min\{q_T, v_T, j_T\} / 3$$

采用最低项有一个理论理由。问题为零时，修订只能围绕外部问题运转；修订为零时，问题只停留在口号；判断场为零时，行动无法与任意主张区分。平均分会让两个强项掩盖一个缺失项，重新制造“课堂很成功”的错觉。

这里的“撤除”必须准确理解。撤除的是诊断所需的熟悉支架，不是撤除残障支持、语言辅助、安全标准、必要知识资源或基本照护。学习者可以查资料、向专家提问、使用工具和寻求同伴，只要能够说明这些资源为什么被引入、改变了哪一步，以及资源之间如何被重新组织。依赖他人不是失败；无法判断依赖的对象、时机与作用，才说明事件仍由外部持有。

十、问题再生：从回答问题到构成可检验的差异

课堂中的“问题”常以问号出现，却并不一定是真问题。教师知道答案、评分表知道步骤、学生知道题型，所谓问题只是启动既定程序的按钮。问题再生要求学生面对一个不完全成形的情境，把其中某种差异转成可处理对象。

问题再生至少经过四个动作。第一，发现不一致：数据与模型不合、作品效果与意图不合、使用者反馈与设计目标不合、史料之间互相冲突。第二，选择承重差异：不是把所有异常都列出来，而是说明哪一处差异若被解释，会改变整体判断。第三，限定可检验形式：确定需要什么材料、比较什么条件、观察什么变化。第四，保留改写接口：新证据出现时，问题本身可以缩小、扩大或改变方向。

数学课堂中的问题再生，不等于让学生自由发明一道题。比如在学习函数极值后，教师给出校园食堂排队数据，但不提供“求最优窗口数”的问句。学生可能提出降低平均等待时间、减少最长等待、平衡员工成本或照顾行动不便者等不同问题。这些问题都使用数学，却对应不同价值标准。真正的再生不在于选到教师心中的答案，而在于学生能说明所选目标、约束和证据之间的关系，并在数据反驳时修改模型。

工程课堂中的问题再生尤其明显。所谓“为学校装太阳能”并不是一个完成的工程问题。要服务哪类使用者？目标是节约成本、减少碳排、增强停电韧性，还是教育展示？屋顶产权、维护能力、阴影、预算与法规如何进入？若这些由评分表全部给出，学生只是求解；若学生必须与场地、人员和标准接触，问题才开始被构成。

艺术课堂的问题再生往往表现为注意的重新分配。学生不是先问“老师喜欢哪种风格”，而是从材料阻力、观看者反应、身体节奏或空间关系中发现作品目前无法承担的张力。例如，一件装置原本试图表达脆弱，却因材料过于光洁而呈现商业橱窗感。问题不是

“怎样让它更好看”，而是“材料的完成度为何取消了作品想保留的不稳定”。这个问题一旦形成，后续选择便不再是任意审美偏好。

问题再生有三种假象。其一是复杂化：把简单题目加上更多背景，不等于形成新问题。其二是个性化：换成自己喜欢的话题，不等于建立可检验差异。其三是批判姿态：说“现有问题设定有偏见”仍不够，必须指出偏见如何进入材料、指标或后果，并提出可以比较的替代设定。

因此，问题再生的最低证据不是问题清单，而是一份“问题形成记录”：最初注意到什么，排除了哪些邻近问法，选择当前问题的依据是什么，什么新证据会迫使问题改写。这份记录可以是文字、图、动作示范、音频或模型，不必统一媒介。

十一、修订链再生：从版本数量到证据驱动的改变

课堂评价常把“有多个版本”当作迭代证据。然而，版本可以只是扩写、润色、换色、调整格式或按教师意见修补。真正的修订链要求前后版本之间存在可解释的因果：某条证据使某个假设、策略或标准失去支撑，学习者因此改变了行动方向。

修订链至少包含五个可追踪位置：初始方案、遇到的阻力、阻力被解释为何种证据、由此改变的核心部分、改变后出现的新后果。缺少任何一项，修订便容易被事后叙事包装。比如学生说“我们根据反馈改进了作品”，却无法指出反馈中的哪一句与作品的哪一处发生关系，便不能证明反馈进入了判断。

工程学的故障树、版本控制、测试日志与设计评审为修订链提供了成熟工具。教育学的错误分析、对比案例、自我解释和检索练习说明，错误只有被解释并改变后续表征时才具有学习价值。艺术工作室中的草图、排练录像、材料样本和同伴批评记录，则让非语言变化获得痕迹。三类工具可以互相约束：工程日志防止艺术变化被神秘化，艺术材料防止工程修订被表格化，教育学的概念测量防止修改只停在作品表面。

本章提出一个“二次改变”最低线：在变异情境中，学习者至少完成两次由不同证据触发的核心修改。第一次修改常只是把旧方法适配新表面；第二次修改更能显示学习者是否理解了问题与标准的关系。比如数学建模中，第一次可能调整参数，第二次才发现目标函数错误；工程设计中，第一次可能换材料，第二次才重写用户需求；艺术创作中，第一次可能移动构图，第二次才改变观看者进入作品的方式。

修订链还必须保留被放弃方案。课堂往往只展示最终答案，失败版本被清理掉。这样做会使判断看起来来自天赋或权威，也让后来者无法知道哪条路为何不通。保留废案不是鼓励堆积材料，而是让修订的方向可检验：若旧方案在新证据下其实仍然成立，学生就必须承认修改可能只是迎合。

修订链的正向读数不是“版本越多越好”，而是证据进入之后，核心决策是否发生了可定位变化。过度迭代可能意味着目标不清、标准漂移或对反馈缺乏筛选。成熟学习者有时会坚持原方案，但必须能说明新证据为何不足以改变它。修订能力包括改变，也包括有理由地不改变。

十二、判断场再生：从“老师说对”到可进入、可异议、可复核

问题和修订若没有判断场，学习仍可能困在私人体验或个人自信中。判断场是由标准、证据、参与者、材料痕迹和复核程序共同构成的空间。它回答五个问题：什么算证据？什么算改进？谁可以提出异议？异议落在哪一件材料上？结论在什么条件下被修改？

课堂中的标准经常以评分量表出现。量表有助于透明，却也可能把判断冻结为外部清单。判断场再生要求学生不仅使用标准，还能说明标准的来源、适用边界和相互冲突。工程中的成本、安全、性能与公平可能无法同时最大化；历史研究中的来源可靠性、代表性与沉默也会冲突；艺术中的统一、多样、形式完成度与开放性并不服从单一尺度。

判断场不等于多数表决。艺术批评提醒我们，异议可能是作品继续生长的条件；科学与工程同样需要少数反例打断共识。可共同检验的意思不是所有人得出同样结论，而是不同结论能够指出各自依赖的材料、标准和推理。当分歧只能归结为“个人感觉不同”，判断场尚未形成；当分歧能够定位为“你把使用者安全置于成本之前，而我认为当前证据不足以证明风险”，讨论才有继续修订的入口。

第三方复核是判断场的关键。另一组学生不需要复制相同答案，但应能根据留下的材料重建：问题如何形成，哪条证据触发哪次修改，结论依赖哪些标准。若第三方只能看到最终作品，无法重建判断链，课堂可能培养了作品生产，却没有培养可传递的学习。

对身体化和艺术性学习，第三方复核不能被误解为强迫全部语言化。舞蹈可以用前后动作片段、节奏标记、身体图谱和同伴模仿来呈现判断；音乐可以用录音版本和演奏选择对照；视觉艺术可以用材料试验与观看路径记录。复核要求证据可进入，不要求证据只以文字存在。

判断场的成熟还表现在标准可以被修订。学生若发现评分量表奖励的“完整”使作品失去风险，或发现工程标准忽略维护者劳动，就需要提出改变标准的依据。但在安全关键领域，学生不能随意废除法规。判断场再生不是标准任意化，而是区分不可谈判的安全底线、暂时采用的专业规范和可以被重新权衡的项目目标。

十三、二维辨别格：课堂表现与学习事件再生不是同一条轴

为了避免把新概念再次变成抽象赞美，本章把课堂表现与学习事件再生分成两条独立轴。横轴是学生在原课堂中的表现，纵轴是支架撤除与情境变化后的事件再生。二者组合成四种状态。

	原课堂表现低	原课堂表现高
变境后再生高	迟效发生 原课堂表现较低，但能在新情境中形成问题、修订和判断。	再生性掌握 既能高质量完成原任务，也能在撤架后重新构成学习事件。
变境后再生低	尚未形成 基础知识、策略、参与条件与事件组织均需继续建设。	借来的成功（关键盲区） 成绩、作品或流程优秀，但问题、路径与标准仍由原课堂持有。

表 1-1 原课堂表现 × 学习事件再生的二维辨别格

右上格是“再生性掌握”：学生既能在原课堂中高质量完成，也能在变境后重新形成问题、修订与判断。左下格是“尚未形成”：原表现与再生都低，需要知识、策略、支持与参与条件的共同建设。左上格是“迟效发生”：原课堂表现不高，但学生在新情境中能生成有效事件，常见于生产性失败、慢熟理解、非语言学习或原评价未捕捉的能力。右下格是本章最要识别的“借来的成功”：原成绩、作品、流程和参与都很好，支架一撤，问题、修订和判断立即失去组织。

“借来的成功”最危险，因为它符合多数形式审查。短期指标显示改善，失效时不产生信号，制度还会自我加固：分数高促使教师继续使用同样题型，流程完整促使课程增加模板，作品漂亮促使学生继续把判断外包，人工智能输出优秀促使平台提供更多自动构成。课堂越正规，事件所有权越难被看见。

二维格也纠正了对低表现学生的快速判断。一个学生可能因语言表达、时间限制、焦虑、文化风格或对原任务缺乏认同而表现不佳，却在开放变境中显示很强的问题形成和修订能力。若只看原成绩，这类学习被漏掉；若只看再生，又会忽略基础知识和执行能力。两条轴必须并置。

十四、观察协议：支架撤除——情境变异——事件重构——第三方复核

学习事件再生不能靠问卷中的“我能独立学习”来测量，也不能靠一次开放任务主观判断。本章提出四阶段观察协议。

第一阶段是支架地图。教师先列出原课堂中实际提供了哪些支持：问题表述、范例、步骤提示、概念表、评分量表、固定材料、同伴角色、教师追问、工具默认设置、人工智能提示词、时间结构和反馈方式。很多支架因长期存在而不可见，地图的作用是分清学生表现中哪些结构来自环境。

第二阶段是情境变异。新任务保留原学习的深层关系，却改变表面对象、目标冲突、材料或参与者。变异必须足够大，使套用原答案失效；又不能大到变成完全不同学科。数学中可保留函数关系而改变目标与约束；工程中可保留能量设计而更换用户和维护条件；艺术中可保留空间张力而更换媒介与观众。

第三阶段是受控撤架。每次撤去两到四项熟悉支架，例如不再给出问题、不提供原量表、更换材料、关闭答案型人工智能、让教师只回答事实查询。撤架应逐步进行，并保留安全、可及性和必要知识支持。目的不是制造无助，而是定位哪些组织功能已经转移给学习者。

第四阶段是事件重构与第三方复核。学习者提交的不是单一答案，而是一组过程证据：问题形成记录、修订链、标准说明、关键材料和一份给复核者的最小入口。另一组学生据此重建判断链，标出无法复现的位置。原组可以回应并修改记录。

整个协议由一句不含价值判断词的问话统领：

撤掉原例题、教师提示、评分量表与固定材料后，学生能否在一个陌生情境中提出一个可检验的问题，依据新证据完成两次修改，并让另一组学生复现这条判断链？

这句话故意不使用“有意义”“充分”“合理”或“真正”等词。它要求观察者去看文件、版本、证据和复核过程，而不是请教师凭印象判断学生是否成熟。

为了提高可比性，可以采用四个辅助读数，但它们只用于研究和形成性反馈，不用于给学生贴固定标签。

第一，问题来源比例：最终承重问题中，有多少核心限定来自学生对材料和后果的分析，而非教师、模板或模型直接给出。第二，证据响应率：所有核心修改中，有多少能够指向明确的新证据。第三，标准可追踪度：关键标准能否说明来源、冲突与边界。第四，第三方重构率：复核组能够从留下的材料中恢复多少关键决策关系。

本章更建议使用剖面而非总分。若研究需要一个最低限度的汇总读数，可把问题、修订和判断各编码为0—3级，取三者最低值作为再生等级。等级定义如下。

等级	状态	观察定义
R0	事件未启动	熟悉问句与提示撤除后不再行动，或只等待外部重新指定任务。
R1	表面复刻	更换对象后复制原问题、步骤与量表；核心关系仍由原课堂提供。
R2	给定任务迁移	能在新问题中应用知识并修改方案，但问题或主要标准仍由外部给出。
R3	部分再生	能形成问题并完成证据驱动修订，但判断场、标准或第三方入口仍不完整。
R4	完整再生	能重新形成可检验问题、两轮以上核心修订与可进入的判断场，并在延迟变境中再次出现。

表 1-2 学习事件再生的形成性等级

延迟复测不可缺少。第一次变境可能因新奇、同伴优势或教师暗示而成功；两到四周后再给另一种表面情境，才能观察结构是否稳定。若学生能更快识别需要什么资源、更早提出可检验问题，并减少对原模板的寻找，才说明再生开始成为可持续能力。

十五、五类课堂中的运行示例

（一）数学：从“会做极值题”到重新构成优化问题

原课堂教授函数极值，学生能完成标准题。变境任务给出校园食堂一个月的排队数据、窗口成本、学生迟到记录和一条无障碍通道要求，不提供目标函数。教师撤去例题和评分量表，只保留数据说明与必要统计工具。

低再生表现是直接询问“要最大还是最小”“用哪个公式”，或自行假设目标却不说明价值后果。中等表现是提出平均等待时间最小化，并完成计算，但忽略极端等待和人员

成本。高再生表现是比较多种问题设定，解释为何当前选择同时考虑平均、尾部风险和可及性；在发现数据受课程时间分布影响后修改模型；让另一组依据其标准与代码复现结论。此时，数学知识不是从题目到答案的通道，而是学生构成公共问题的一部分。

（二）工程：从“走完设计流程”到重新形成目标与标准

原课堂完成太阳能校园设计，教师提供用户、预算、性能指标和评分表。变境任务换成一座老年照护中心，屋顶面积更小，维护人员有限，停电风险更高。学生可以使用原软件与技术资料，但不再得到预制需求清单。

低再生团队会复制原校园方案，只调整面积；中等团队能访谈使用者并修改参数，却仍把“发电量最大”当作唯一理想；高再生团队会发现供电韧性、维护可达性、成本与教育展示之间存在冲突，形成标准权重的理由，在模拟失败后改写目标，并把被放弃方案与风险证据保留下来。第三方不必同意其权重，但能够重建其判断。

（三）艺术：从“被作品打动”到重新生成观看与批评

原课堂围绕一件大型装置进行教师主导的讲评，学生获得丰富术语和强烈体验。变境时，学生面对一段城市噪声、一件旧衣和一处狭窄走廊，不给主题，也不提供原讲评词汇。任务不是“创作一件好作品”，而是让材料之间出现一个可被他人检验的关系。

低再生表现是模仿原装置形式或等待教师解释；中等表现是产生个人表达，却无法说明修改依据；高再生表现是从声音被走廊放大的不适形成问题，通过录音、身体移动和材料悬挂进行两轮修改，并组织同伴从不同路线进入作品。分歧被定位到具体听觉和空间证据，而不是归结为“审美不同”。判断场允许异议继续存在。

（四）历史与语言：从“复述观点”到重新形成证据关系

原课堂学习一场历史事件，教材给出主线叙事和标准材料。变境任务提供相互冲突的日记、报纸、政府档案、后来的回忆和一份缺页记录。学生不被要求“证明哪个版本正确”，而要决定目前最值得追问的差异。

低再生表现是按权威等级排列材料；中等表现是指出偏见，却没有可检验问题；高再生表现是把“谁没有留下记录”本身构成问题，比较材料产生条件，因一条时间证据修改初始叙事，并明确哪些结论只能暂时成立。另一组可以根据来源链重建其判断，也能指出替代解释。

（五）生成式人工智能课堂：从“高质量输出”到确认事件所有权

原课堂允许学生使用生成式人工智能写研究报告。模型可以帮助检索、列提纲、提出反例、修改语言与模拟审稿。若只看最终文本与对话日志，学生很容易获得高表现。变境诊断需要设计“角色切换”而不是简单禁用。

第一轮让模型承担答案生成者，记录学生接受、拒绝和修改的理由。第二轮关闭答案与提纲功能，只允许模型作为批评者：学生必须自行提出问题、选择材料和形成初始标准，模型只能指出缺口。第三轮完全关闭模型，用新材料重构一次较小事件。比较三轮可以区分操作熟练与事件再生。

高再生学生并不必然少用人工智能。他们可能使用得更多，却保持三种作者责任：问题为何这样限定，哪条证据触发了哪次修订，什么标准决定接受或拒绝模型建议。低再生学生即使亲手输入每个字，也可能只是执行模型给出的事件结构。因此，人工智能时代的关键分界不是“人写还是机器写”，而是问题、修订和判断的关系由谁组织、能否在工具角色改变后重建。

十六、课堂设计：从“讲完一课”转向“交出下一次学习的发起权”

学习事件再生不是反对讲授，也不是把所有课堂改成无边界探究。初学者需要清晰示范、概念解释、例题、练习、反馈和知识积累。问题在于，许多课程只完成了知识获得周期，没有安排事件再生周期。学生学会在教师构成的场中行动，却从未经历组织权逐步转移。

本章建议采用“双周期课堂”。第一个周期是获得周期，目标是建立必要知识、基本动作和典型判断。教师在这一周期中可以明确讲授、分解难点、提供范例、纠正误解。第二个周期是再生周期，目标不是再教一次，而是改变情境、撤去部分熟悉支架，让学生重构问题、修订和判断。两个周期不是二选一；没有获得周期，再生容易退化为无知识的意见表达；没有再生周期，获得容易停在依赖性的高表现。

再生周期可以由五个教学动作组成。

第一，保留“未被指定的问题空间”。教师不把所有问题写成清单，而是在材料中留下一个需要学生选择的承重差异。这个空间不能只是“任选题目”，而要有真实约束，使不同选择带来不同后果。

第二，公开支架。教师告诉学生当前提供了哪些帮助：范例解决了什么，量表替代了哪类判断，提示减少了哪种搜索。支架一旦被命名，学生才可能逐步接管；未被命名的支架会伪装成学生自己的能力。

第三，制造结构保持、表面变化的变境。把深层关系保留，把对象、材料、目标冲突或参与者换掉。变境不以刁难为目的，而是让学生发现哪些方法只绑定于原表面，哪些关系能够被重新组织。

第四，安排可恢复的撤架。每次只撤去少量支架，并允许学生在卡住时申请资源。关键不是“不得求助”，而是求助必须带着诊断：当前缺的是事实、方法、反馈、标准还是合作对象？申请资源本身可以成为学习证据。

第五，组织第三方复核与回场。另一组先尝试重建判断链，原组再回到初始任务，说明新事件如何改变了对旧任务的理解。回场使学生看见，学习不是把旧知识搬到新地方，而是新关系反过来改写旧知识的意义。

一节九十分钟的课堂可以这样安排：前二十分钟建立或复习必要知识；接下来十五分钟共同分析一个范例，同时标记教师提供的支架；再用二十五分钟处理变异情境，教师只回答事实与安全问题；随后十五分钟交换材料进行第三方重构；最后十五分钟由学生说明哪项支架已经能够接管、哪项仍需教学。这样的课堂并不减少教师作用，而是把教师作用从持续替代组织转为设计转交时机。

教师专业性也因此被重新定义。优秀教师不仅会解释和反馈，还能识别课堂里哪些结构始终由自己持有，能设计不伤害学生的撤架，能区分“学生不会”与“任务从未把发起权交出”，并能在学生形成错误问题时提供边界而不立即夺回问题。教师不是退出学习事件，而是培养学生成为下一次事件的共同作者。

十七、评价重构：不再用一个分数结算学习

传统评价把学生放在同一个任务中比较表现，便于公平与管理。学习事件再生评价则必须面对不同学生可能提出不同问题、使用不同证据、形成不同标准。若直接把这种差异压成一个总分，新理论会迅速退化为另一张量表。

更合适的方式是建立四类证据档案。第一类是问题来源档案，记录问题如何从材料中形成、经历了哪些改写。第二类是修订档案，保留关键版本、被放弃方案和触发修改的证据。第三类是判断档案，说明标准来源、冲突和边界。第四类是复核档案，记录第三方能够重建什么、在哪里发生分歧、原组如何回应。

评价报告应同时呈现原课堂表现和再生剖面。一个学生可以是“知识准确度高、问题再生低、修订链中、判断场高”，而不是被归为“会学习”或“不会学习”。这种剖面能指导教学：问题再生低可能需要更多观察与问题化训练；修订低可能需要错误分析和证据对照；判断场低可能需要同伴复核与标准讨论。

再生证据必须延迟采集。若变境任务紧接原教学，学生仍可能凭工作记忆复刻。两到四周后、更换教师或同伴、改变材料媒介，可以提高诊断效度。对长期课程，可每四至六周安排一次小型再生任务，而不是每节课都撤架。过度诊断会让学生把“表现独立”当作新的迎合目标，也会占用知识学习时间。

群体学习需要避免把集体成果全部平均到个人。可以记录谁提出了承重差异，谁引入了关键证据，谁使标准被修改，谁完成了整合；但这些记录不应用来抢夺署名。目标是看学习事件如何分布，而不是把每个动作货币化。一个成员可能没有提出问题，却通过质疑使问题改变；另一个成员可能提供材料，使他人的修订成为可能。分布式贡献是学习的一部分。

高风险评价必须保持克制。学习事件再生目前首先是研究构念和形成性工具，不宜直接用于升学、淘汰或教师绩效。任何量化读数都指向“这个位置的组织功能目前由谁持有”，不指向一个人的固定本质。已经依据该读数作出不利安排时，机构应公开编码规则、原始证据和复核通道。

十八、生成式人工智能改变了什么：从答案外包到事件外包

过去，学生可能把计算、检索或语言润色外包给工具；生成式人工智能第一次可以同时外包问题形成、过程组织和评价标准。它能告诉学生“应该研究什么”，给出逐步路线，在每个节点建议修改，并以模拟审稿人身份判定作品是否成熟。于是，课堂最容易观察的全部痕迹都可以由工具制造。

近年的教育技术综述显示，生成式人工智能在 K—12 和高等教育中被广泛讨论为个性化支持、反馈、内容生成与教师助理，同时研究也持续报告准确性、依赖、学术诚信、偏

差与证据不足等问题。仅仅区分“允许使用”和“禁止使用”无法处理事件外包，因为禁止可能损害可及性和真实职业准备，允许又可能使高质量表现与学习完全脱钩。

学习事件再生提供了角色化使用原则。人工智能至少可以有五种角色：答案者、脚手架、批评者、对手和模拟器。答案者直接构成问题与解法，最容易制造借来的成功；脚手架提供部分结构，适合知识获得期；批评者只能指出缺口，不替学生确定修订；对手提出竞争解释，迫使学生澄清证据；模拟器让学生测试工程、社会或艺术后果。课程应逐步把模型从答案者移向批评者、对手和模拟器。

评价人工智能辅助学习时，应进行三项“角色切换测试”。第一，问题切换：模型不再提供选题和提纲，学生能否从新材料中形成承重问题。第二，修订切换：模型只允许提出反例，学生能否决定哪些反例改变核心方案。第三，标准切换：模型不给评分量表，学生能否提出标准、说明来源并接受同伴质疑。三项测试不要求学生离线完成全部工作，而是观察事件结构是否随工具角色变化而崩解。

人工智能日志只能提供有限证据。大量提示词可能表示主动探索，也可能只是反复索取答案；很少提示词可能表示独立思考，也可能是一次性接受模型输出。关键证据仍是问题、证据、修改和标准之间的关系。学生需要标记哪些建议被接受、拒绝或改写，以及理由落在哪项材料上。

人工智能也可能增强再生。它可以迅速生成相反案例、模拟不同使用者、暴露模型假设、把作品置于多种观众面前。若学生掌握事件组织权，工具扩大可变条件；若工具掌握事件组织权，学生只是在更强的课堂装置中移动。区别不在工具强弱，而在角色配置和可撤除性。

十九、与最近的既有理论逐一分界

“学习事件再生”与多个成熟概念相邻。仅仅说它“更综合”没有意义；必须在同一案例上给出相反预测。下表列出十个最近邻及经验分离线。

既有概念	它已经说明什么	判决性分界
迁移学习	把既有知识或策略应用到研究者给出的新任务。	若给定新任务的迁移成绩高，但撤去问句与量表后无法形成问题，则本章区分成立；若标准迁移完全预测事件重构，则本章被吸收。
未来学习准备（PFL）	通过先前活动提高从后续资源或教学中学习的能力。	PFL 允许后续教学与资源成为关键；本章要求在新教学到来之前，学习者先构成可检验事件。
适应性专长	在效率与创新之间切换，面对非例行问题作出灵活回应。	适应性专长通常以成熟实践为背景；本章预测初学者也可能低效率却高再生，且必须重构问题与判断场。
自我调节学习	围绕目标进行计划、监控、策略选择与反思。	当目标与标准被撤去时，高自我调节者若只寻找外部要求而不重构问题，两者即分离。

既有概念	它已经说明什么	判决性分界
生产性失败	先让学生生成不完整解法，再由规范教学整合。	若没有预定汇总教学与标准答案，学生仍能从失败形成新一轮问题与判断，才达到本章所说的再生。
知识建构／认识主体性	学习共同体持续改进公共观念，成员承担认识责任。	本章增加支架撤除与共同体更换：原群体消失后，问题——修订——判断能否在新参与者中重构。
情境学习／实践共同体	通过参与既有实践、身份与共同体关系学习。	本章追问旧共同体不在时，学习者能否进入或构成新的实践，而不只沿既有参与轨迹前进。
转化性学习／审美转化	观点、身份、感知或世界观发生深刻改变。	强烈转化可以没有可复核修订；本章也允许无顿悟却逐步形成可重构判断。两者对同一案例可给相反分类。
建造主义	通过制作可分享的有意义作品促进思考与知识形成。	高质量作品可能由模板、同伴或人工智能组织；关闭这些结构后事件是否崩解，是本章的离线线。
ICAP 参与框架	交互、建构、主动、被动等参与模式预测学习差异。	高度交互但全部由教师脚本控制的活动可低再生；低可见互动的个体工作也可能高再生。参与形态不能替代事件所有权。

表 1-3 “学习事件再生”与十个最近邻概念的经验分界

最接近的概念是迁移与未来学习准备。它们已经把学习从原课堂移向新情境，但大多数测验仍向学生给出新任务、材料和成功标准。学习事件再生把测量再向前推进一步：新情境中什么构成问题、什么证据足以改道、什么标准使结论暂时成立，也进入观察。若标准迁移测验能够完全预测学生在无题目、无量表情境中的事件重构，本章的新构念便被迁移吸收；若出现大量“迁移高、事件再生低”或相反案例，两者才有经验区别。

第二个近邻是自我调节学习。自我调节强调目标设定、计划、监控、策略与反思，但很多研究中的任务目标和评价仍由课程给出。学生可以高效调节完成一个伪问题。本章预测：在目标被撤去时，一部分自我调节水平高的学生会迅速寻找外部要求，而不重新构成问题；另一部分课堂表现一般的学生可能较慢，却能重新形成目标和标准。若这种分离不存在，自我调节足以覆盖本章。

第三个近邻是生产性失败。生产性失败让学生在教学生成解法并从失败中学习，但规范概念和汇总教学通常已经预定。本章并不否定这种设计，而是把它视为学习事件再生的准备条件。判决性差异在于：没有后续规范讲解和预设目标时，学生能否利用失败继续形成新一轮问题与判断。若生产性失败的优势只在教师随后精确汇总时出现，它不等于事件再生。

艺术学中的转化性学习与审美转化也十分接近。它们强调观点、身份和感知方式的深刻改变。本章增加的并不是“更深”，而是另一种可判定要求：转化之后能否留下第三方可进入的修订与判断关系。一个人可能真诚地报告世界观改变，却无法在新材料上作出可追踪判断；也可能没有强烈顿悟，却逐步形成稳定的再生能力。二者对“强体验而低复核”的案例给出不同分类。

二十、三领域反过来给新理论加上的约束

一个跨领域理论若只从三门学科取材料而不接受反向限制，最终会把原学科变成例证仓库。教育学、工程学和艺术学至少迫使本章放弃三条更强、却站不住的主张。

第一，教育学迫使本章放弃“支架越少越好”。初学者的工作记忆、先备知识和策略库有限，过早撤架会把认知资源消耗在无效搜索上，也可能放大背景差异。直接教学、范例、检索练习、反馈和结构化互动仍是必要条件。学习事件再生不是每节课的唯一目标，而是知识获得达到一定密度后，周期性检验组织权是否开始转移。没有这条限制，本章原本可能把独立生成置于所有教学价值之上；现在必须把“厚实教学”放在“受控撤架”之前。

第二，工程学迫使本章放弃“问题和标准都应由学生重建”。在医疗、航空、核能、建筑安全等领域，法规、容差、验证程序和伦理底线不能被课堂开放性取代。学生可以理解标准为何存在、如何操作、何时冲突，却不能为证明再生能力而重新发明安全规则。本章的适用范围因此缩小为：在不可谈判底线之内，对目标、方案、证据和可调整标准进行重构。没有这条限制，再生会变成危险的任意化。

第三，艺术学迫使本章放弃“可复核就必须完全可言说”。身体知识、节奏判断、材料感觉和含混经验可能先于语言，也可能永远不能被完整翻译。第三方复核可以通过动作模仿、版本对照、材料痕迹、声像记录和共同观看完成。本章要求的是判断入口，不是统一话语。没有这条限制，新理论会把非语言学习压成行政文书，并伤害它试图保护的创造性。

三条约束共同改变了价值排序。真正的课堂学习不是最大化独立、最大化开放或最大化可测量，而是在知识继承、安全边界与多样表达之中，逐步转交事件组织权。再生能力越强，也越需要知道何时不该自行重构、何时必须服从已经用生命代价建立的标准、何时应当让不可言说的证据保持其媒介形式。

二十一、适用边界与伦理风险

学习事件再生首先适用于概念理解、设计、探究、创作、论证和复杂实践，不宜覆盖所有课堂目标。基础自动化技能有时需要高频稳定练习；紧急程序有时要求无条件按标准执行；语言初学者可能需要大量模仿。把每一种学习都改成开放事件，会浪费认知资源并制造不必要风险。

对儿童、神经多样性学习者、残障学生和第二语言学习者，撤架不得等同于撤除可及性支持。辅助技术、延长时间、视觉提示、沟通伙伴和情绪安全安排不是需要被“证明可以脱离”的依赖。诊断应撤去的是熟悉答案结构，而不是使学生能够参与的基本条件。

文化差异也需谨慎。公开质疑、主动提问和个人表达在不同文化中的形式不同。问题再生可以通过书面、集体、匿名、图像或行动方式发生，不能把快速口头发言当作发起权。共同体中由长者或专家提供问题也不必然取消学生再生，只要学习者能理解、修改并承担问题与证据的关系。

任何再生读数都不得用于给学生固定分类。它描述某一时点、某一领域、某一支架配置下事件组织功能的位置。教师改变、同伴变化、知识增长和情绪安全都可能使结果改变。机构不得为了让数字好看，在安全关键系统里故意制造轮换、关闭必要工具或安排高代价失败。

新理论还会产生典型的反噬：一旦学校把“问题再生率”“修改次数”列入绩效，最省事的做法将是要求学生提交更多伪问题和伪版本。学习者会学会表演犹豫、制造废案、编写标准冲突，判断场再次被模板持有。这个风险没有简单出口。能做的只是限制量化用途，保留原始材料，使用抽样复核，并让学生有权质疑编码规则。

二十二、可证伪预测与研究设计

“学习事件再生”必须承担被既有概念吸收或被数据推翻的风险。最强的竞争解释是：本章不过把迁移、自我调节、创造力、元认知和认识主体性重新打包。要排除这个解释，需要证明新读数在控制这些变量后，仍预测一种只有事件组织权才能解释的结果。

建议开展一个多点、同质、纵向研究。可在同一教育系统内选择至少六个班级，覆盖数学、工程设计和艺术课程，每个领域至少两个班。所有班级先接受等量的知识教学，再随机进入“常规掌握”或“掌握加再生周期”条件。研究在基线、课程结束、四周后和一个学期后采集原课堂成绩、标准迁移、自我调节、创造力、先备知识、动机、问题形成记录、修订链、判断场与自发项目启动情况。

关键结果不是哪组当堂分数更高，而是四周后面对没有预制问句和量表的新材料时，学生能否构成可复核事件；一个学期后，是否在没有教师要求的情况下发起并完成新的学习项目。编码者应跨学科训练并盲于实验条件，第三方复核由其他班学生完成。统计模型控制教师、班级、先备知识、语言表达与原成绩。

下列条件可以使本理论受挫或被推翻。

编号	竞争解释／风险	触发条件	理论后果
F1	概念冗余	若标准迁移、自我调节、适应性专长与创造力可以完全预测事件再生，且新增读数无增量效度。	撤回独立构念，把协议降为既有变量的组合任务。
F2	机制无效	控制课时、教师质量、动机和先备知识后，再生周期不提高四周后的无预制问题任务表现。	否定“组织权转移”是教学作用机制。

编号	竞争解释／风险	触发条件	理论后果
F3	测量不可复核	跨训练编码者不能稳定识别问题来源、证据触发的核心修订和判断入口。	放弃当前编码体系，承认构念尚不可操作化。
F4	时序预测失败	独立修改先大量出现，问题与标准只在事后访谈中补写，且这种顺序更能预测后续结果。	把“问题先形成”从核心机制中删除。
F5	人工智能外包假说失败	控制能力与使用时间后，完整生成、脚手架和批评者三种角色在关闭工具后的下降没有差异。	不能再主张事件外包独立于一般工具使用。
F6	跨媒介偏差	非语言和身体化学习在未被强迫语言化时无法被第三方复核，而语言流利度主导全部评分。	缩小适用范围，停止宣称覆盖艺术与身体学习。
F7	经验分层不存在	大样本中只出现“表现与再生同步高低”，二维格的两类对角分离极少且不稳定。	二维模型失去必要性，课堂表现可继续作为主要代理。

表 1-4 核心命题的机制证伪条件

还有一条正向时序预测。再生取向课堂中，学生对问题和标准的自发重构应先于独立修订的增长，而独立修订的增长又应先于延迟自发项目的出现。若观察到的顺序相反——学生先大量修改，之后才在访谈中补写问题与标准——则所谓再生可能只是事后叙事。时间顺序比相关更难被共同方法偏差满足。

机制证伪必须控制最朴素解释。比如，若再生组后来更常发起项目，可能只是教师更热情、学生更有动机或课堂时间更多。因此研究需要在同一学校、相近教师培训和等量课时下比较，并测量动机与教师支持。控制这些因素后，若问题——修订——判断最低项不能预测之后的自发项目，本章的机制解释就失败，结果应归因于一般参与度或教师质量。

人工智能提供另一个低成本实验。随机让学生使用三种模型角色：完整生成、只给脚手架、只做批评。控制总使用时间和先备能力后，本章预测完整生成组原课堂产出最高，但关闭模型后的事件再生下降最大；批评者组当堂产出可能较低，撤除后下降较小。若三组的撤除下降没有差异，或完全由自我调节分数解释，事件外包假说不成立。

二十三、自反：本章自身能否通过它提出的判据

本章提出了一个问题、形成了一条修订路径，也给出判断标准和证伪条件，但这并不证明“学习事件再生”已经成立。读者得到的仍是一套相当完整的概念装置：题目、框架、

术语、表格和案例都被预先提供。读者即使能够复述全文，也可能只是在本章构成的事件中高表现。

用本章自己的判据检查，当前文本至少有三个薄弱处。第一，它主要以文字论证为证据，可能低估身体化、关系性和沉默形式的再生。第二，它提出0—3级编码，虽然反复限制高风险用途，仍可能诱发量表化和表演性版本。第三，它尚无独立实证数据，所有课堂案例都是理论构造或研究启发，不能被当作效果证明。

本章真正能够交出的不是“已经被验证的答案”，而是一组可以让它失败的观察。如果未来研究发现，标准迁移、自我调节和适应性专长完全预测所谓事件再生，新概念应被撤回；如果第三方无法可靠识别问题、修订与判断的关系，测量协议应被放弃；如果再生教学只增加学生负担而不改善后续自发学习，教育价值主张不成立。

本章也可能成为自己描述的“借来的成功”。它结构完整、引文丰富、分界清楚，却可能依靠写作者熟悉的理论模板，而未生成可持续研究共同体。只有当不同学科的研究者拿走本章术语、在新场景中重新构成问题、修改协议并发现本章没有预见的失败，它才开始接近自己所定义的学习。

本章作出一条有日期的赌注：到2029年8月3日之前，至少会出现一项预注册、多点、跨两种以上学科且包含不少于六个班级的研究，发现原课堂成绩与学习事件再生出现稳定分离，并且再生读数在控制原成绩、标准迁移和自我调节之后，仍能预测至少四周后的自发学习项目。只有教师自评、学生自我报告、单一班级、教师仍提供问题与量表、或没有延迟结果的研究，不算命中。若届时没有此类结果，或者有严格研究显示增量预测为零，本章的核心命题应被显著收缩。

接续。本章的再生水平 R 取三项中的**最低项**，这条取法在第十二章被用作整份问卷的合成规则； R 的采集属附录一里成本最高的一档，那里给了它在一学期方案中的位置。本章与第二、三章的关系——三个名字是不是同一样东西——由第一编小结判决，判决办法是三个读数的相关结构。本章的证伪条件与写死日期的赌注已汇入附录三。

第二章 交到下一个人手里才算数

判断卡片：学习不在任何一个人身上结清——它必须被交到下一个人手里、被那个人改写并继续用下去，才算存在。

三个理论入口：认知科学 × 经济学 × 艺术学

提要。课堂通常以即时正确率、参与度、作业完成度和作品质量来确认学习，但认知科学表明，快速的任务熟练与缓慢形成的可迁移能力并非同一过程；经济学显示，努力、注意力与坚持会随激励复杂度、同伴规范和自主权配置而改变；艺术学则指出，思考并非先在头脑中完成再被表达，而是在材料、限制、批评和反复制作中形成。三条证据合在一起，推翻了一个共同而隐蔽的前提：学习总能在某个学生、某个时点被单独结算。本章提出“接力再生事件”：一种新辨别在撤除原例与提示后，能够由不同参与者在结构同型的新情境中重新生成，被另一人继续推进，并据此改变下一轮提问、分工或评价。本章给出二维辨别格、无情态词判据、课堂发生机制、跨学科案例、与九类近邻理论的判决性分界、可操作的研究设计及证伪条件。其结论是：课堂学习的最小单位不是个人内部增加了一项知识，而是共同体中出现了一个此前不存在、此后能够继续发生的认知接力回路。

一、引言：一堂“成功”的课，为什么可能什么也没有留下

一间课堂可以在所有可见指标上获得成功。教师的问题设计精致，学生发言踊跃，小组讨论持续，展示作品完整，课堂末测的正确率明显高于课前。录像中的每一分钟都能被编码为“有效互动”，听课者也会写下“目标达成、主体参与、生成丰富”。然而，两周以后，把原来的例子、板书、话术、组员和评分表全部撤掉，换成一个表面不同但结构相同的问题，学生往往只能回忆答案，不能重建产生答案的辨别；有人说出原句，别人却接不下去；教师若不重新提供脚手架，课堂不会从旧经验中长出下一步。此时，那堂课究竟发生了学习，还是只发生了短时的成功表演？

这个问题之所以难，不是因为教育学缺少“学习”的定义，而是因为常用定义分别抓住了不同现象，却把它们当成了同一件事。考试成绩抓住的是当下表现；认知测验试图抓住稳定能力；行为经济学关注注意力、努力与激励之间的配置；合作学习关注互动质量；艺术教育关注制作、批评与形式生成；课堂观察则把参与、表达和任务完成当作代理变量。每一种读数都可能有价值，但它们之间并不存在天然的同一体性。一个学生可以在没有形成可迁移能力时获得高分，可以因公开排名而投入更多努力，也可以制作出一件新颖作品却不能在另一种材料中重建同一关系。反过来，一种深刻变化也可能暂时不表现为正确答案、积极发言或漂亮作品。

因此，“何谓真正的课堂学习”不能再靠给既有指标加权来回答。问题需要从“课堂里出现了多少正确行为”转向“课堂里是否诞生了一种此前不存在、后来还能继续发生的东西”。本章提出：真正的课堂学习，不是知识进入某个学生的头脑，不是学生在给定奖励结构中最优配置注意力，也不是一次具有新意的表达或作品；它是一个**接力再生事件**——一种新的辨别、做法或形式，在撤除原例和提示以后，能够由不同参与者在结构同型的新情境中重新生成，被另一位参与者接着推进，并反过来改变下一轮课堂的提问、分工或评价。

这个定义包含三个不能相互替代的时间环节。第一，必须发生“再生”：不是记住原答案，而是在新表面上重新找到组织问题的差异。第二，必须发生“接力”：新的辨别不能只停留在最初提出者那里，而要被另一个人接住、修改或延伸。第三，必须发生“回写”：接力后的结果要改变下一轮课堂环境，例如改变问题的问法、材料的配置、角色的分工、证据的标准或评价的规则。缺少再生，课堂容易把熟练误当能力；缺少接力，课堂只有私人掌握，没有形成共同学习；缺少回写，互动只是一次性热闹，课堂结构仍按原样重复。

本章的论证不以三门学科的“共同点”为起点，而以它们的冲突为起点。认知科学把学习的判据放在表现背后的能力变化；经济学把学生行为置于稀缺注意力、认知耐力、同伴规范和激励机制中解释；艺术学则主张思考在材料、约束与制作过程中发生。三家对同一课堂现象给出互不相容的归属：它究竟是一个人的内部改变、一种资源配置结果，还是一条由行动和材料共同构成的生成路径？本章将证明，三种答案各自成立的材料，合在一起却迫使我们放弃它们共同默认的“学习有一个稳定主人和结算时点”的前提。

本章有四项贡献。第一，把课堂学习的分析单位从“个人状态”改为“跨主体、跨时点的发生事件”。第二，以再生与接力构成一张二维辨别格，区分服从性表现、私人掌握、脚本流通和真正课堂学习。第三，把“课堂学习是否发生”压缩成一句可以查日志、作业版本和评价规则的观察问题，而不是价值判断。第四，提出一套可被反驳的经验设计：如果接力再生事件不能在控制个人迁移能力、即时成绩、投入时间、动机和同伴网络之后解释后续变化，那么本章的核心命题应当被放弃。

二、三种前沿答案：能力、配置与制作

2.1 认知科学的答案：学习是表现背后的能力变化

认知科学首先要求把“做得更好”与“学得更多”分开。训练任务的分数上升可能来自题型熟悉、策略记忆、速度适应或反馈利用，也可能来自更缓慢形成、能够迁移到新任务的能力。Ericson 与 Klingberg (2023) 分析了 1300 名儿童八周工作记忆训练的数据，提出双过程模型：早期任务特异性提高约占总增益的 44%，平均在训练第三天附近进入平台；另一条较慢的能力变化轨迹则与迁移表现相关。这项结果的意义不在于给“真学习”指定某一种测验，而在于证明同一条上升曲线内部可能混合着性质不同的变化。课堂未测的漂亮增益，不能自行说明学生获得了可在别处使用的能力。

认知科学的第二个推进，是把教师——学生互动从背景变量推进到学习过程的可测结构。Li 等 (2024) 比较了结构化互动的翻转课堂、面授讲授与自发互动、视频讲授与自发互动三种条件。结构化互动条件不仅学习增益更高，而且教师与学生左侧背外侧前额叶的脑间同步更强，脑间同步变化与学习改善相关。Xu、Zhang 与 Zhang (2026) 进一步在真实课堂中取得 4175 次纵向脑电记录，覆盖 107 名中学生、393 节语文和数学课、三个学期。静态模型中，某些数学课上的师生脑间耦合与当期成绩甚至呈负关联；动态模型中，耦合的增加却能预测同一阶段的成绩改善。研究者将静态负关联谨慎地解释为可能存在教师选择效应：教师会把更多互动资源给更需要支持的学生。由此可见，关系的“变化”比关系的静态高低更接近学习轨迹。

认知科学由此形成一个强主张：学习不能用即时表现直接结算，它应当表现为较慢、较稳、能够跨任务或跨时段延续的能力变化；在课堂里，这种变化又与师生互动的动态耦合有关。这个主张能够排除大量假学习，却仍保留一个传统归属：最后发生变化的是“这个学生的能力”。脑间同步是过程指标，教师互动是原因或条件，学习本身最终仍以个人成绩、个人迁移和个人能力为结算点。

2.2 经济学的答案：学习是稀缺认知资源在制度中的积累与配置

经济学不从“知识是什么”开始，而从学生为何把注意力和努力放在某处开始。课堂里的专注、坚持、出勤、练习与冒险都具有机会成本；学生不仅在学习内容，也在对成绩、同伴评价、未来收益、任务难度与自主权做响应。因此，同样的高投入行为可能来自能力成长，也可能来自奖励设计、声誉压力、损失规避或规则不透明。

Brown、Kaur、Kingdon 与 Schofield（2025）把“认知耐力”纳入人力资本：持续进行费力思考的能力本身可以通过学校经验增长。在 1600 名印度小学生的随机实验中，不论训练内容是数学问题还是非学术游戏，只要包含持续的费力认知活动，学生在后续任务中的表现衰减都明显减轻，学校成绩也有所提高。研究还交叉随机设置奖励：奖励能抬高测试开始时的表现，却不能减少随时间出现的下降。这个结果切开了两个课堂中常被混淆的量：愿不愿意一开始用力，与能不能持续用力，并非同一个东西。更重要的是，认知耐力甚至可以由不含学科知识的练习增长，这使“学到了课程内容”与“形成了能够支撑学习的资本”再次分离。

自主权同样不是课堂气氛的装饰。Cullen 与 Oppenheimer（2024）的高等教育现场研究表明，把一部分出勤或课程承诺交由学生选择，能够提高到课率与学科掌握；学生并不总把自主理解为减少约束，很多人会把选择用作自我承诺。与此同时，Abeler、Huffman 与 Raymond（2025）的现场与实验研究显示，激励方案只要稍微复杂，部分属性就会变得不透明；有限理性的参与者会忽略某些条款，进而大幅改变努力投入，甚至相对于完全理性基准过度投入。课堂评分表、徽章、排名、过程分、同伴互评和 AI 反馈组合得越复杂，并不意味着学生获得的信息越充分；复杂度本身会制造一层“看不见的价格”。

经济学由此给出第二个强主张：课堂学习不能脱离资源配置来谈。注意力会疲劳，认知耐力会积累，激励会改变投入，自主权会改变承诺方式，同伴规范会改变公开努力的成本。一个课堂行为是否出现，不仅取决于学生“会不会”，还取决于在此时此地做这件事是否值得、是否安全、是否可见、是否被奖励。这个解释能够揭开“参与度”的经济来源，却也保留自己的归属：学习被结算为个体人力资本、努力供给或未来收益的变化；共同活动主要是形成这些变化的制度环境。

2.3 艺术学的答案：思考不是被表达出来，而是在制作中形成

艺术学对前两种答案都构成挑战。认知科学容易把制作当作内部能力的外部证据，经济学容易把材料、技法与批评当作投入选择；艺术研究则把制作本身看作思考发生的方式。Maras 与 Shand（2023）在艺术教师围绕当代中国艺术作品进行协作推理的行动研究中指出，批判性与创造性思维并不是附加在艺术制作之外的一般能力，而以“制作”的方式显现：判断作品意义的过程包含反复的推理、理论建构、元认知检验，并受实际与理论约束

共同塑造。观点不是先完整存在，再借语言或图像输出；它在对限制的利用、修正和再组织中才成为观点。

Barrett (2026) 的“雕塑工具包”研究进一步模糊了玩、做与学之间的边界。六种普通材料被儿童与成人共同使用，材料不只是承载意图的媒介，而参与知识交换：儿童影响成人，成人影响儿童，材料的阻力、可塑性与组合可能又反过来改变双方的动作和判断。学习不再是人从对象中提取信息，而是人与材料在共同制作中长出新的行动可能。Tufail、Park、Wang 与 Kim (2026) 对传统媒介与数字媒介设计活动的准实验也显示，媒介并非中性通道。数字媒介带来的界面管理和注意力碎片化增加了外在认知负荷；传统材料的物理限制促使学生更多采用问题驱动策略，并在该小样本中产生更高的概念新颖度和更多可持续性考虑。媒介改变的并不是表达效率，而是问题如何被看见。

艺术学由此形成第三个强主张：学习不是一个先成形的内部状态被搬运到作品上，而是在身体、材料、限制、他人批评和反复修改中被组织出来。作品不是学习之后的证据，而是学习发生的场所之一。这个立场能够解释许多“说不清却做得出”“一换材料就看见新问题”的课堂现象，但也有自己的结算习惯：它容易把一次具有生成性的制作视为学习已经发生，即使这种生成只依赖当时的材料、伙伴和情境，日后未必能够再生，也未必能够被另一个人接续。

三、三种答案为什么不能同时充当终审标准

三条理论不是互补的三个角度，而是在争夺同一件事的最终归属。设想一节设计课：学生使用生成式 AI 和纸质材料为社区老人设计导视系统。课堂中，学生不断提出方案，作品迭代迅速，小组争论激烈，最终成果新颖。认知科学会追问：撤掉这个任务与工具以后，学生是否获得了能够迁移的设计辨别？经济学会追问：高投入来自任务价值、自主承诺、同伴声誉还是评分机制？艺术学会追问：新的理解是否在纸张、界面、身体动作、失败草图与批评中生成？三种问题都能得到肯定，却仍可能指向三个不同结论。

第一组冲突发生在认知科学与经济学之间。若一个学生因为奖励提高而投入更多时间、答对更多题，经济学可以准确解释其努力供给变化；认知科学却不能据此承认能力增长。反过来，学生的认知耐力或迁移能力可能增长，但在一个公开努力会被同伴讥讽的班级中，他仍会减少可见投入。于是“没有表现”既可能是没有能力，也可能是能力存在但价格太高；“表现很好”既可能是能力增长，也可能只是激励暂时把努力推高。两家无法共享同一个行为指标。

第二组冲突发生在认知科学与艺术学之间。认知科学倾向于把跨任务稳定性作为较强证据，艺术学却强调独特材料和具体情境的不可替代性。一个学生在泥塑中通过材料塌陷突然理解了“支撑不是隐藏结构，而是可见形式的一部分”，这种理解也许无法被一份脱离材料的标准化迁移题捕捉。若坚持只有抽象迁移才算学习，制作中的生成会被低估；若坚持每一次材料性生成都算学习，则局部偶然、技法熟悉和情境依赖又可能被高估。

第三组冲突发生在经济学与艺术学之间。经济学的课堂设计通常要控制成本、澄清激励、减少无效摩擦；艺术制作却常借助不能提前消除的阻力、含混和失败来产生新形式。对经济学而言，规则不透明可能导致资源错配；对艺术而言，过度透明的任务可能提前封

死探索。若一味优化效率，课堂会把尚未成形的问题压缩成可评分任务；若一味增加开放性，又会把学生推入无穷选择、认知过载和不平等的时间成本。

因此，三种答案不能靠“各占三分之一”来综合。若把能力、投入与制作简单相加，得到的只是一个更复杂的评分表：迁移得分加参与时长加作品新颖度。它依然没有回答，当三项读数彼此冲突时，哪一种变化才构成课堂学习。更根本的问题是，三家虽然争执，却共同站在同一块地上：它们都假定学习最终可以被某个位置持有，并在某个时点完成结算。

四、三家共同默认而没有说出的前提：学习有主人，也有结算时点

认知科学把学习归给学生的潜在能力。教师——学生互动、任务结构和反馈是促成因素，最终读数仍进入学生的记忆、控制、迁移或成绩。经济学把学习归为个体人力资本或资源配置结果，班级规范和激励制度是价格环境。艺术学把思考分布到身体、材料和他人之间，但作品、行动或创作者仍通常提供一个可以辨认的发生位置。三家看似分散，实际共同假定了两件事。

第一，学习有一个稳定的“主人”。即使承认环境、工具和他人重要，研究设计最终仍倾向于问：哪个学生学会了什么？这个问题在很多场景中当然必要，却把跨主体的接续降为外部条件。甲提出一个尚不完整的辨别，乙把它改造成可执行做法，丙在另一情境中重建它，教师再据此改变下轮任务；常规账本会把这些动作分散记录为甲的创造、乙的合作、丙的迁移和教师的教学改进。没有一栏记录“这条理解在谁都没有完整拥有它时，如何跨人活了下来”。

第二，学习有一个可结算的时点。课堂结束、单元测验、作品提交或学期成绩都被当作自然边界。可是任务特异性熟练可能在第三天前快速平台，较慢能力变化随后才显露；一段师生关系的动态变化也可能比某次静态测量更有解释力。艺术制作中的一个形式可能在作品完成后才被另一位学习者重新理解；激励规则今天提高了参与，却可能在下次规则撤除后失效。若学习的时间单位被切在课末，很多尚未完成的变化会被提前宣告完成，很多只在当前条件下成立的表现则会被误认为已经沉淀。

这个共有前提的验收方式很简单：把“学习最终属于某个学生，并可在本课或本单元结束时结算”念给三种研究传统，大多数研究都不会把它当作需要论证的争议命题。它正因为太自然，才构成盲点。真正的突破不能从外面搬来第四种价值立场，而要用三家自己的材料推翻它。

认知科学自己已经提供了最关键的裂缝。2026年的真实课堂纵向研究显示，静态的师生脑间耦合与数学成绩可能呈负关联，而耦合的动态增加却与同阶段进步相关。这个结果并不证明“两个大脑共同拥有一项知识”，也不允许把脑间耦合神秘化为学习本身；但它至少表明，最有信息量的读数可能不是任一方的静态属性，而是关系在时间中的改变。若预测力存在于关系变化中，便不能把关系只当作学生内部学习的外部原因。艺术学关于材料——人之间知识交换的证据、经济学关于同伴规范和制度价格改变努力的证据，又从另外两侧说明：课堂里产生的东西经常没有一个单独拥有者，也没有一个自然终点。

因此，需要把问题从“谁学会了”暂时改写为“什么东西在课堂共同体中获得了继续发生的能力”。这个改写不是否认个人学习，而是为个人账本无法登记的那类事件建立一个存在位置。

五、新判断：课堂学习是一种“接力再生事件”

本章把真正的课堂学习界定为：一种新辨别、做法或形式，在原例、原提示和原分组撤除之后，能够由参与者在表面不同而结构同型的新情境中重新生成；随后被另一位参与者接住、修改或推进；并使下一轮课堂的提问、分工、材料或评价发生可追踪改变。这样的事件称为“接力再生事件”。

这个定义不是把迁移、合作和教学改进并列起来，而是规定它们必须形成一个闭合的时间结构。四个条件缺一不可。

第一，新辨别的产生。学习者必须把原先混在一起的两类情况分开，或把原先割裂的因素组织成一种新关系。它可以是一条数学不变量、一个历史证据标准、一种构图关系、一个实验误差模型、一种程序调试策略，也可以是对自身情绪和行动的重新识别。只有答案增加而辨别方式不变，属于信息获得，不构成这里所说的事件。

第二，延迟再生。原题、原话术、原范例和即时反馈撤除后，参与者要在表面不同的任务里重新生成这条辨别。再生不要求逐字复述，恰恰相反，它要求形式改变而组织关系保留。若只能在教师提示或原材料在场时做出，课堂中出现的是支架依赖表现。

第三，跨主体接力。新辨别必须离开最初提出者，被另一位参与者接住。接住不等于转述，而是能据此增加一个反例、修正一项步骤、迁移到另一媒介、提出更强问题或指出适用边界。若只有原提出者会用，可能已经发生私人学习，但尚未形成课堂学习的公共存在。

第四，环境回写。接力后的结果要改变下一轮课堂。改变可以很小：教师把一个学生提出的反例写进下一次任务，班级把证据来源的权重改进评分表，小组重新分配解释者与质疑者，实验协议增加一项控制，课程讨论从“答案是什么”转为“在何种条件下答案会翻转”。若新辨别不改变任何后续结构，它仍可能有价值，却没有获得继续发生的制度位置。

这四项不是线性加分，而是合取条件。再生为零，接力只是传播脚本；接力为零，再生只是私人掌握；回写为零，成果没有进入课堂的下一轮因果链。正因为任何一项缺失都会改变事件的性质，接力再生事件不是“学习程度更高”的说法，而是一类不同的课堂存在。

它也改变了课堂学习的最小单位。传统最小单位通常是“学生在时点二比时点一多会了一项东西”。接力再生事件的最小单位则是一个跨越至少两个主体、两个时点和两个情境的回路。这个回路没有唯一主人：最初提出者未必是最后重建者，教师未必是发起者，作品或工具也可能承担关键中介。它也没有在一次课末自然结束：只有当新辨别经受撤除、重建、接力和回写，事件才显露为已经发生。

六、二维辨别格：四种看起来都像学习的课堂

为了把定义变成可观察的区分，本章以两条轴构成二维辨别格。纵轴问：原有提示撤除、时间延迟以后，新的辨别能否在结构同型的新情境中再生？横轴问：这条辨别能否被另一位参与者接住并推进，进而改变下一轮课堂？两条轴都不直接使用分数、参与度或作品美观度，因此能够专门处理“所有形式指标都通过，却没有学习发生”的情况。

	不能跨主体接力并改写下一轮	能跨主体接力并改写下一轮
不能延迟再生	服从性表现： 在原题、原提示和原评价下完成任务；条件一撤，做法消失。	脚本流通： 小组能热烈复述、分工展示、相互补句，但换情境后只剩共同话术。
能够延迟再生	私人掌握： 个体能在新情境中独立重建，但别人无法接续，课堂规则不受影响。	接力再生事件： 辨别跨情境重建、跨主体推进，并进入下一轮提问、分工或评价。

左下格最容易识别。学生照例题完成练习，按教师句式回答，评分标准一变便失去方向。它可能是必要的练习阶段，却不能单独证明学习。右上格则是本章定义的课堂学习：它既有个人能够重新生长的能力，又有共同体能够接续和回写的公共生命。

真正危险的是右下格“脚本流通”。它在短期指标上往往最漂亮：学生发言多、合作顺、术语统一、作品完整；它的失败不产生明显信号，因为每个人都能借助同伴、模板和评价语言维持表现；它还会自我加固，教师越依据流畅展示确认成功，下一次越倾向于提供更完整的流程和更统一的话术。生成式 AI 进一步放大了这一格：一个小组可以迅速产出多角度论证、反例和精美幻灯片，却无人能在脱离提示词和模型输出后重建其中的关键辨别。

左上格“私人掌握”则提醒我们，真正的个人学习与真正的课堂学习并不完全同义。一个学生课后独立钻研，能够在新问题中迁移所学，这当然是学习；但如果课堂没有让这项能力被他人接住，也没有因此改变共同活动，那么它主要是发生在课堂场景中的个人学习，而不是课堂作为共同体产生的学习。本章不是贬低这种掌握，而是拒绝用少数个体的掌握替全班结算。

二维格还有一个重要性质：即时成绩可以在四格中都很高。服从性表现可能因题型一致而高分；脚本流通可能因分工和模板而高分；私人掌握可能高分；接力再生也可能高分。因而成绩不是横纵轴的替代物。参与度、作品质量、脑间同步或认知投入同样不能单独替代两条轴。它们是解释事件为何发生的材料，不是事件是否发生的终审读数。

七、一句可以查记录的问题：课堂学习的零情态判据

价值判断式定义最容易在课堂评价中失效。只要问题写成“学生是否真正理解”“互动是否有意义”“迁移是否充分”，不同观察者就会把自己的教育理想填进答案。本章把判据压缩为一句不含“应当、真正、充分、合理、恰当、有意义”等情态词的问题：

十四天后，撤掉原例、原话术和原分组，换一个表面不同而结构相同的问题：班里是否有人能独立重建那条辨别，另一人能接着把它推进，并据此改写下一轮的提问、分工或评价？

这句话可以直接对应数据。独立重建可从无提示作业、口头解释、动作演示或作品版本中判断；另一人接力可从修改记录、引用链、接续发言和跨组迁移中判断；下一轮改写可从教师备课版本、任务说明、评分量规、实验协议或小组角色表中判断。它不要求观察者判断学生“内心是否理解”，而要求检查一个跨时点回路是否留下痕迹。

判据在不同课堂中会产生不同答案，而不是把命题换一种说法重复出来。

场景一：数学证明。原课讨论“抽屉原理”的一个经典例子。十四天后改为网络节点分配问题，不出现“抽屉”一词。甲学生独立重建“对象数超过容器容量便必有重合”的不变量；乙学生指出原结论在允许动态扩容时失效，并提出新的边界条件；教师把“先识别不变量，再列适用条件”加入下一轮证明评分表。这里四项条件齐备。若全班只会说“物品多于盒子”，则是脚本流通。

场景二：科学实验。原课用弹簧测量研究误差。两周后换成温度传感器，设备和数据形式全部不同。学生能否重新区分随机波动、系统偏移和仪器漂移？另一组能否据此增加校准步骤，并使后续实验协议改变？答案不能从原课成绩直接推出，必须查实验日志和协议版本。

场景三：历史证据。原课围绕一份战争日记讨论史料可信度。延迟任务换成企业档案与口述史冲突。学生若只是复述“要多方互证”，仍未再生；若能重新构造“作者位置——形成目的——保存路径——沉默部分”的证据链，另一人据此翻转原叙事，班级再修改证据权重规则，才构成事件。

场景四：艺术工作室。原课在纸浆材料中发现“支撑可以成为表面”。两周后换成声音与空间装置。学生能否在没有纸浆的情况下重建“支撑——显露”关系？另一人能否把它推进为“观众移动本身成为支撑”，并改变下一轮批评会的问题？这里的再生可以是作品和动作，不必首先被语言化。

场景五：经济与管理案例。原课分析一份奖金合同中的不透明条款。延迟任务换成课程积分与游戏徽章。学生能否重新识别“看得见的奖励遮蔽了被忽略的成本”？同伴能否据此修改一套班级激励方案，教师再把“列出未被参与者计入的属性”加入下一次案例分析要求？这需要查方案修订史，而不是听学生声称“我懂了”。

五个场景的可见产物不同：证明中的不变量、实验中的误差模型、历史中的证据链、艺术中的形式关系、经济中的不透明属性。判据没有规定学习必须长成同一种内容，只规定它必须经历相同的时间结构。至少科学实验和艺术工作室的答案无法从命题本身推导，需要真实版本记录；这避免了判据沦为自我证明。

八、接力再生如何发生：六个阶段与一个顺序预测

接力再生事件不是教师宣布一个目标后自动出现，它通常经历六个可区分阶段。六阶段不是固定教案，而是一条用于诊断的时间链。课堂可以从不同内容进入，但若长期跳过其中某一段，事件会退化成二维格中的其他类型。

8.1 第一阶段：原有区分失效

学习的起点不是信息缺口，而是旧区分在一个具体阻力面前失效。数学中的例题方法遇到反例，历史叙事遇到相互冲突的档案，设计方案遇到材料塌陷，程序遇到不能由语法

错误解释的异常，伦理判断遇到两个都合理却不能兼得的价值。教师的任务不是立即填补缺口，而是让失效变得可见，同时控制挫败成本。没有可见失效，学生最理性的选择往往是继续复制已经被奖励的做法。

8.2 第二阶段：把正在形成的东西外置

学生需要把尚未完成的判断放到公共空间中：草图、半成品、猜想、动作、代码、证据表、概念图、口头模型都可以。外置的价值不是展示成果，而是让他人和材料能够对它施加反作用。艺术学在这里提供关键约束：媒介不是中性记录器。纸笔、实体材料、共享白板、编程环境和生成式 AI 会分别突出、遮蔽或改变问题。教师不能只问“用什么工具效率更高”，还要问“这个工具让哪一种错误更容易被看见，哪一种错误被藏起来”。

8.3 第三阶段：让异质反作用进入

一条辨别只有在受到不来自原作者的阻力时才开始成形。阻力可以来自同伴反例、材料限制、数据异常、角色互换、时间延迟或评价标准冲突。这里要避免把“合作”理解为意见叠加。若每个人只补充一条相容观点，课堂会变成信息拼盘；有效反作用必须迫使原判断放弃一部分、改变适用范围或重新排序价值。艺术批评的价值不在赞美与建议，而在作品被另一套观看方式重新组织。

8.4 第四阶段：把发起权交给另一个人

课堂最容易在这里中断。一个学生提出好见解，教师予以总结，然后继续按原教案推进；见解被教师收编，却没有被同伴接力。接力要求改变发起权：下一位参与者不能只评价甲的说法，而要以甲的辨别为起点，完成一个甲没有完成的动作。可以是产生反例、把关系移入另一媒介、给出边界、改变算法、重写问题或设计一次检验。为了防止“强者持续发起、弱者持续附和”，接力对象可在延迟后重新分组，而不是当堂由最熟悉原思路的人继续。

8.5 第五阶段：撤除原条件后的再进入

课堂中的许多成功依赖提示、情境和社会线索。再生必须经历撤除：改变题面、材料、学科语言、组员、评价方式或时间位置，但保留需要识别的深层关系。撤除不是为了增加难度，而是为了分开“记得上一轮做了什么”与“能够重新生成为何那样做”。认知科学关于任务特异性提高与较慢能力变化的分离，为这一阶段提供了直接理由。延迟时间可依据课程性质调整，但必须跨过即时记忆与原情境线索最强的窗口。

8.6 第六阶段：进入下一轮课堂的规则

最后一步不是庆祝迁移成功，而是确认新辨别是否获得制度位置。它有没有改变下一次的问题库？有没有进入新的评分规则？有没有使小组角色重新分配？有没有让教师删去原计划中的一个环节，新增一个由学生发现的问题？如果答案全是否定，新辨别仍依赖参与者的私人记忆，下一届、下一组或下一次备课很容易归零。回写使学习从偶发事件变成共同体可继续利用的环境变化。

六阶段给出一个比相关更严格的顺序预测：旧区分的失效应先于外置，外置应先于跨主体修正，跨主体修正应先于课堂规则改变，规则改变又应先于下一轮更高概率的延迟再生。若观察到教师先修改评分规则，学生随后只是适应新规则；或学生先在测验中迁移，之后才被解释为接力效果，那么本章所说的机制没有得到支持。顺序之所以重要，是因为参与、互动、成绩和作品质量之间很容易出现相关，而“谁先动、谁随后改变”更难由巧合满足。

九、十二个学科场景：同一判据如何产生不同的具体预测

跨学科理论只有在不同领域中产生可被观察的差异，才不只是比喻。以下十二个场景不要求课堂采用同一种活动，而要求每个领域给出一件原理论未必会预测的具体事情。

9.1 数学：从记住证明到重建不变量

若课堂只训练标准证明格式，学生可能在同型题中高分。接力再生理论预测：真正变化首先显露在学生能否把证明压缩成一个不变量或约束，并在题面变化后重新生成；随后，另一位学生应能用反例修改其适用边界。最有价值的课堂记录不是谁先写出完整证明，而是谁的辨别在另一人的证明中改变了形状。

9.2 物理：从套公式到重写实验条件

学生会用公式不等于理解模型。延迟任务若改变测量尺度、摩擦条件或边界，真正学习者应先检查模型成立条件，而不是先寻找对应公式。接力发生时，同伴会把一个人的异常数据转化为新的控制变量，实验协议随之改变。若协议从未因学生数据而修改，所谓探究很可能只是在预设轨道上重演。

9.3 化学：从识别反应到追踪被搬走的条件

在化学平衡与反应动力学教学中，学生容易把“结果改变”归为某个单一因素。接力再生要求他们在新反应体系中重新识别浓度、温度、催化与时间尺度的共同作用，并让另一个小组据此改变取样时点。预测是：真正学习首先改变实验序列，而不只是解释文本。

9.4 历史：从观点多元到证据结构改变

讨论课可以产生很多立场，却不一定改变证据使用。本章预测，真正课堂学习发生时，学生会把另一位同学提出的“档案形成目的”用于重排自己的证据，而不是只在结论段增加一句“也有不同观点”。下一轮评分标准应出现证据来源、沉默和保存路径的具体权重变化。

9.5 语言与写作：从模仿文体到迁移叙述关系

学生可以熟练模仿范文句式。延迟再生的检验应撤掉原主题和关键词，观察其能否重建视角、时间距离或信息遮蔽等叙述关系。接力不是互相润色，而是乙根据甲的叙述结构改写一个完全不同题材，并指出该结构在哪类文本中会失效。

9.6 视觉艺术：从个人风格到可接续的形式问题

艺术课堂常尊重个人表达，因此不愿要求作品“可传递”。接力再生不要求复制风格，而要求形式问题可被别人接住。例如，甲在雕塑中发现“空隙承担重量感”，乙可在摄影或舞蹈中重新组织“空隙——重量”关系。若只能复制甲的形状，不是接力；若乙能改变媒介并推进问题，课堂才获得公共形式。

9.7 音乐：从准确演奏到共同改变听觉标准

合奏中的同步与准确可能来自指挥控制。真正学习发生时，演奏者应能在换曲、换声部或无指挥条件下重建张力处理；另一位成员把这种处理推进为新的排练问题，乐团据此改变对“齐”的评价。预测是：排练记录中的评价词会在事件后发生可追踪改变，而不只是在原曲中表现更好。

9.8 编程：从修复一个错误到改变调试协议

学生借助 AI 可以快速修复代码，却未必形成调试辨别。延迟任务应更换语言或项目，让学生重新区分输入错误、状态错误、并发时序和规格错误。接力发生时，另一位同学以该区分新增测试用例，班级代码审查清单随之改变。若错误只被修复而协议不变，知识仍停留在一次性答案中。

9.9 医学模拟：从完成流程到接住异常信号

在安全关键教学中不能为了观察接力而故意制造危险，但可以用模拟。真正学习不只是按清单完成，而是当异常信号与标准流程冲突时，学员能重建优先级；下一位接手者能从记录中继续判断，而不是从头重做；模拟后的交接模板因此改变。这里的核心读数是“接手成本”，不是表演性流畅。

9.10 职业技能：从手把手复制到跨材料重建

木工、烹饪、护理和设备维护包含大量默会知识。接力再生允许用动作和产物而非语言证明再生。学徒在换材料、换设备或换对象后仍能重建力道、顺序和风险判断；另一位学徒能够观察并改进；工作站的摆放或检查顺序随之调整。若必须由原师傅重新示范，技能尚未获得公共生命。

9.11 公民讨论：从表达立场到修改共同判据

课堂辩论往往把发言数量当学习。本章预测，真正变化不在学生是否改变立场，而在是否形成一条可以被对手使用的判据。甲提出“政策代价落在谁身上”，乙即使保持相反结论，也能用这条问题重写自己的论证；班级随后修改讨论规则，要求每方说明代价承受者。立场未趋同，学习仍可能发生。

9.12 AI 辅助课堂：从答案增产到提示撤除后的再生

AI 可以显著提高产量、语言完整度和方案数量，也可以承担跨主体传播，制造“脚本流通”的强假象。本章预测，使用 AI 的课堂只有在模型输出、提示词和对话记录撤除后，

学生仍能重建关键辨别；另一位学生不是复制提示，而能改变问题；教师据此调整下一轮人机分工，才构成学习。若一切改进都随工具撤除而消失，发生的是外部能力租用。

十二个场景显示，接力再生不是把所有学科变成讨论课。数学可以用证明，医学可以用模拟，艺术可以用作品，职业教育可以用动作。统一的是事件结构，不是表现形式。

十、课堂设计：从“把课讲好”转向“让理解获得下一轮”

接力再生事件不能由某一种教学法保证。讲授、练习、讨论、项目、实验、工作室和AI辅助都可能成为发生条件，也都可能退化为脚本流通。设计重点不在活动名称，而在课堂是否为再生、接力和回写分别留下位置。以下八条原则构成一套最低设计框架。

10.1 先设计延迟任务，再设计当堂活动

常规备课从“这节课做什么”开始，延迟检验则在课程结束后临时补充。更稳妥的顺序是先写下：两周后，表面改变而结构保留的任务是什么？原例和提示要撤掉哪些？再回头决定当堂需要哪些讲授、练习和制作。这样能够迫使教师把教学目标从“学生完成本题”压缩成“学生获得哪条可以重建的辨别”。

10.2 把半成品纳入正式产出

如果课堂只收最终答案，学生最理性的策略是隐藏试错、借用成熟话术、避免暴露不确定性。接力需要可供他人进入的半成品：尚未验证的猜想、失败草图、冲突证据、代码分支、动作片段和未完成模型。评分可以把“留下可被接手的中间状态”单列为证据，但不能奖励关于“我留下了中间状态”的自我声称；必须存在另一位参与者实际接手的记录。

10.3 接力对象与原小组分开

同组成员共享语境、暗号和短期记忆，容易把熟悉误当传递。延迟后重新组队，才能测试新辨别是否脱离原社会线索。可以采用“甲组留下问题包，乙组在不见甲组成员的情况下接手”的方式；接手包应包含必要材料，却故意不提供结论。若乙组必须通过询问原作者才能继续，说明知识尚未形成可接续结构。

10.4 让接力者拥有修改权

很多课堂允许学生“评价”同伴，却不允许他们改变任务、量规或作品方向。没有修改权，接力只会停在评论。教师可预留一个有限而真实的改写接口，例如：每轮由学生提交一条反例进入下轮题库；实验组可以增加一项控制；写作组可以删除一项评分指标；艺术批评会可以更换下一轮的提问。修改权必须有边界，以免课堂目标被任意漂移，但它不能只是象征性建议。

10.5 把首次错误与最终结算脱钩

若首次公开错误直接降低高权重成绩，学生会合理地避免提出尚未成熟的判断。经济学的约束在这里不可忽视：教师不能一面要求冒险，一面用不可逆评分惩罚冒险。可将初次尝试记为低风险记录，把高权重评价放在修订、再生和接力之后；也可以允许学生选择何时把草稿转为正式提交。关键不是取消评价，而是让探索阶段保留可逆性。

10.6 控制摩擦，而不是消灭摩擦

艺术制作需要材料阻力，认知变化需要旧区分失效，但过量摩擦会耗尽工作记忆与认知耐力。Tufail 等（2026）的研究提醒我们，数字媒介可能因界面管理和注意力碎片化增加外在负荷。课堂设计应区分“产生问题的摩擦”与“消耗资源却不产生辨别的摩擦”。前者如数据异常、材料限制和观点冲突；后者如多平台切换、重复登录、模糊指令和无关格式要求。开放不等于把所有复杂度交给学生承担。

10.7 让课堂记录能够追踪“谁接了谁的下一步”

传统记录追踪谁发言、谁提交、谁得分，却很少记录一条想法如何换手。接力记录至少包含四项：前一版本的关键辨别、接力者改动了什么、改动依据来自何处、哪项课堂结构随后改变。记录对象是位置和版本，不是给学生贴“创新者”“追随者”标签。一个人在不同事件中可以发起、接力、质疑或终止，没有固定身份。

10.8 把教师从最终解释者改为回路维护者

教师仍需要讲解、示范和纠错，尤其在新手阶段。但教师的额外职责是防止所有新见解都被自己收回。若学生一提出关键辨别，教师立刻用更完整术语总结，课堂会把发起权重新集中到教师。回路维护包括：延迟总结，让同伴先接；把学生问题带入下轮；识别脚本流通；在认知负荷过高时补充直接教学；在错误可能造成安全风险时中止开放探索。教师不是退出，而是负责让接力既能发生又不失控。

这些原则也解释了为什么“学生中心”不是充分条件。课堂可以把大量时间交给学生，却仍由预设模板控制一切；也可以由教师进行高质量讲授，在关键时点让学生的辨别真正改变后续任务。决定性区别不在谁说得多，而在课堂因谁的发现改变了下一轮。

十一、与九类最近说法的判决性分界

“接力再生事件”包含迁移、生成、合作、制作和共同体变化，因此极易被现有概念吸收。下面逐一给出分界。每一条都以同一案例下的相反判定为标准，而不使用“更全面”“更深入”一类无法检验的说法。

11.1 与学习迁移的分界

迁移研究关心先前学习是否改善新情境表现，尤其是表面不同任务中的结构迁移（Bransford & Schwartz, 1999）。本章接受延迟迁移是必要条件，却拒绝把它当充分条件。若一个学生在新任务中独立成功，没有任何他人接续，课堂提问和规则也未改变，迁移理论判定学习成立，本章判定为私人掌握而非课堂学习。反过来，若没有个体再生，只有共同讨论热闹，本章与迁移理论都不会放行。

11.2 与生成性学习的分界

生成性学习强调学习者通过自我解释、绘图、总结、教学他人等活动，把新信息与既有知识主动组织起来（Fiorella & Mayer, 2015）。若学生能做出高质量自我解释，却在提示撤除后不能重建，或其解释未被另一人推进，生成性活动已经发生，接力再生事件没有发生。若另一人只是复述该解释，也仍停在脚本流通。

11.3 与生产性失败的分界

生产性失败研究表明，先让学生在复杂问题中生成不完善方案，再进行针对性教学，可能促进概念理解和迁移（Kapur, 2008）。本章吸收“失败可以准备未来学习”的洞见，但要求进一步追踪跨主体接力与环境回写。若失败后接受讲解的学生在延迟测验中进步，却无人把同伴的失败结构用于修改下一轮问题，生产性失败判定干预有效，本章仍只承认个人学习。若学生失败很多却没有形成可再生辨别，两种理论都不把失败本身当学习。

11.4 与 ICAP 框架的分界

ICAP 框架按被动、主动、建构、互动四类参与方式预测学习水平，互动活动在满足相互生成条件时通常具有更大潜力（Chi & Wylie, 2014）。本章不以活动类型排序，而以延迟后的事件结果判定。若两名学生当堂互相推理、共同生成新解释，但两周后只能依赖原话术，ICAP 可以把过程编码为高质量互动，本章将其归入脚本流通。若一场短讲授促成后来跨情境再生和跨主体接力，ICAP 会关注其中后续的建构活动，本章则承认讲授可以成为事件的一环。

11.5 与知识建构理论的分界

知识建构理论强调共同体以改进“公共知识对象”为目标，而非只完成个人任务（Scardamalia & Bereiter, 2006）。这是与本章最接近的邻居。分离线在于“公共对象持续改进”与“不同主体在撤除原语境后重新生成”并不等价。若一个在线社区持续编辑同一份概念图，公共知识明显进步，但新成员离开原讨论记录后不能重建其中关键辨别，知识建构理论仍可判共同体知识进步，本章判定尚未形成接力再生。反过来，若新成员能重建并改写后续规则，二者会同时判定成立。

11.6 与情境学习和实践共同体的分界

情境学习强调知识与参与实践不可分离，学习表现为从边缘参与走向更充分参与（Lave & Wenger, 1991）。本章同样反对把知识看作脱离活动的物，但保留跨情境再生这一硬条件。若学徒在本工作站中逐步胜任，却在材料、设备或团队改变后必须等待原师傅重新组织，实践共同体理论可以判其参与身份已经变化，本章不判接力再生完成。这一分界不贬低地方性技能，而是指出课堂若声称培养可迁移判断，必须另给证据。

11.7 与转化性学习的分界

转化性学习关注个体参照框架、意义视角或身份理解发生改变（Mezirow, 1991）。若学生经历深刻反思，重新理解自我与世界，但这一变化未被他人接续，也没有改变课堂后续结构，转化性学习判定可能成立，本章只承认个人转化。若一项公共辨别改变了班级讨论规则，却没有任何人发生身份层面的转化，本章可以判定课堂学习成立，转化性学习未必成立。

11.8 与具身、生成及材料参与理论的分界

具身认知、生成认知和材料参与理论把认知视为身体、行动、工具与环境共同构成的过程（Malafouris, 2013; Hetherington, 2024）。本章的艺术学来源与其高度相近，但

分离线在时间和接力。若一次材料互动产生独特洞见，却不能在另一媒介中再生、不能被另一人继续，也未改变后续课堂，材料参与理论仍可说明认知在制作中发生，本章不把它结算为课堂学习事件。这条分界同时防止本章把每次“新奇体验”都升级为学习。

11.9 与人力资本、理性忽视和激励理论的分界

人力资本理论关心教育如何增加未来生产能力；理性忽视与有限理性研究关心人在信息处理成本下忽略哪些属性；激励理论关心规则如何改变努力。若一项持续练习提高了认知耐力或出勤，学生在长任务中表现更稳，却没有产生某条可重建、可接力并回写课堂的辨别，经济学可以判定人力资本增长，本章不判定特定课堂学习事件成立。反过来，一次接力再生可能没有立即提高考试收益或劳动生产率，本章仍承认学习发生。

九条分界中，知识建构理论最接近本章，因为两者都拒绝把学习完全私有化。但决定性对照仍然存在：公共知识对象可以依靠平台记录、核心成员和持续语境维持；接力再生要求在原线索撤除、参与者更换后重新长出，并进入下一轮规则。若经验研究发现公共知识改进已经必然包含这种再生与回写，那么本章的新概念将被知识建构理论吸收，不应继续单独命名。

十二、三门学科反过来限制新命题

跨领域论证若只从各学科提取支持，而不允许它们削弱新命题，最终会变成包装。认知科学、经济学与艺术学分别迫使“接力再生事件”放弃三条原本更强、却站不住的主张。

12.1 认知科学拿掉“当堂生成即可结算”

若没有认知科学的约束，本章很容易把一次精彩的跨主体生成直接称为学习。双过程训练研究表明，快速任务特异性提高可以与较慢能力变化混在同一上升曲线中。因此，本章必须放弃“课堂中出现新辨别并被同伴推进，就足以确认学习”的强说法，增加延迟再生条件。这个削弱使大量工作坊式高峰体验被暂时搁置：它们可能是学习的起点，但只有延迟后仍能重建，才进入本章的终审。

同时，脑间同步研究只能作为关系变化的证据，不能被当作课堂学习的生物标志。脑间耦合可能受教师选择、共同刺激、动作同步和任务结构影响。本章因此放弃“关系指标越高，学习越真实”的说法。真正承重的是跨时点事件结构，不是某项神经读数。

12.2 经济学拿掉“摩擦越多、接力越频繁越好”

若只从艺术制作和互动理论出发，本章可能赞美一切阻力、延迟和开放接力。经济学提醒：学生的时间、注意力和认知耐力稀缺，复杂规则会产生不透明，公开参与存在同伴成本。本章因此必须把适用范围从“所有教学内容都应组织成接力再生”缩到“值得形成可迁移辨别、且错误成本允许延迟与修订的内容”。基础事实、自动化技能、紧急程序和高风险规范可能更适合直接教学与充分练习。

经济学还迫使本章放弃“接力次数越多越好”的潜在考核逻辑。接力有成本，且频繁换手会损失默会信息。最优课堂不是最大化换手，而是在知识尚未获得公共生命时提供必要接力，在需要深度连续工作时保护稳定持有。接力记录只能用于诊断位置，不能成为个人绩效排名。

12.3 艺术学拿掉“再生必须以语言和抽象规则出现”

若从认知迁移和经济测量出发，本章容易要求学生说出规则、填写表格、完成标准化同型题。艺术学指出，一些辨别存在于触觉、节奏、姿态、空间关系和材料应答中，过早语言化会把它压扁。本章因此放弃“只有能清楚陈述才算再生”的强说法。再生证据可以是动作、作品、演奏、操作序列或对材料的选择，只要另一位参与者能够接住其中关系并推进。

这项让步也改变了评价的价值排序：可陈述性不再高于可重演性。一个学生说得极好却做不出来，可能停在脚本流通；另一个学生难以陈述，却能在新材料中重建关系并使同伴接续，仍可构成事件。评价需要多模态证据，而不能把语言优势误当学习优势。

三项反向约束使新命题更窄，却更可守。没有认知科学，它会把一次生成误当持久变化；没有经济学，它会忽视成本、不平等和激励扭曲；没有艺术学，它会把默会与材料性学习排除在外。跨学科的价值正在于每一门都毁掉了本章原本可能写下的一句更强的话。

十三、如何研究它：事件记录、无量纲读数与伦理边界

接力再生事件首先是一类存在，随后才需要测量。若一开始就把它压成分数，学校会迅速学会生产“好看的接力”：安排学生轮流引用、在教案中预留学生提问、要求每组修改一次量规。形式齐全，事件却可能仍未发生。因此，研究应先保存事件链，再从链中提取辅助读数。

13.1 最小事件记录

一条合格记录至少包含六个字段：

1. **初始失效点**：哪一种旧做法在什么材料或反例前失效；
2. **新辨别版本 A**：最初参与者把什么区分开或重新联系；
3. **撤除条件**：延迟多久，撤掉哪些原例、话术、工具或社会线索；
4. **再生版本 B**：谁在新情境中重新生成，形式发生了什么变化；
5. **接力版本 C**：另一位参与者增加了哪项边界、反例、媒介或步骤；
6. **回写证据**：下一轮任务、分工、材料、协议或评价具体改了哪一处。

只有“学生说受启发”“教师认为讨论有深度”不构成事件记录。版本 A、B、C 可以是文本，也可以是图像、代码、动作视频、实验表或作品。研究者应尽量使用课堂已有记录，避免为了测量引入过重监控。

13.2 两个辅助轴与一个合取指标

可以分别编码“再生强度”和“接力回写强度”。再生强度考察表面差异、提示撤除程度、延迟长度和重建独立性；接力回写强度考察是否跨主体、是否产生实质修改、是否进入下一轮正式结构。两轴保持分开，防止一个总分掩盖脚本流通与私人掌握。

在需要比较班级或教学方案时，可定义“接力再生率”：在预先登记的有效机会中，同时满足延迟再生、跨主体推进与环境回写的事件数，占全部有效机会的比例。这个比率无量纲，数学课可按延迟任务计数，艺术课可按作品批评轮次计数，实验课可按协议修订机会计数。它与考试平均分正交：一个平均分最高的班级可能接力再生率很低；一个仍在

形成新辨别、短期分数普通的班级可能更高。若两者在所有样本中高度重合，新增指标便没有独立价值。

但该比率不能直接用于给教师或学生排名。它指向课堂位置，不指向人的固定品质；不得为了提高数值而在安全关键课堂中安排不必要换手，不得为了制造“学生改写课程”的证据而放弃教师责任；任何已经依据该读数作出不利安排的机构，都应公开编码规则、原始事件链和复核通道。最危险的反噬是：一旦接力再生率进入考核，最省事的达标方法会变成文书中的接力，而不是理解的接力。对此没有纯技术出口，只能限制用途并保留人工复核。

13.3 样本与比较纪律

研究不应拿两个国家、两所制度差异巨大的学校进行宏大比较，然后把差异归因于接力再生。更可靠的起点是在同一学校、同一年级、相近课程进度中选择至少六个班级，最好达到三十个以上班级，记录不同教学条件下的真实变异。关键混淆变量包括先前能力、教师经验、班额、课时、任务难度、评分压力、AI 工具使用、学生流动和社会经济背景。

班级是重要层级，但事件链内部还包含学生、任务和轮次。分析可使用多层模型：学生层控制先前能力与动机，班级层估计教学条件，事件层检验再生——接力——回写的顺序。为了区分关系变化与教师选择，应同时建静态模型和动态模型；若只看某一时点，教师给予低成绩学生更多支持可能制造反向关联。

13.4 低成本数据

多数学校已经拥有可用于初步检验的数据：学习平台的版本历史、作业修订、在线讨论引用、题库修改时间、评分量规版本、实验协议、代码提交记录、教师备课文档与小组重组记录。第一轮研究不需要脑成像或全程录像。神经耦合研究提供的是理论启发，不是学校的常规测量要求。若一种学习理论只有在昂贵设备下才能被看见，它很难成为课堂判断工具。

十四、最强竞争解释与可证伪研究方案

本章面对的最强竞争解释不是“课堂里根本没有互动”，而是更朴素的组合说：所谓接力再生，不过是**延迟远迁移**、**合作学习**和**教师响应性教学**三者同时出现，没有必要假定一种新的事件。这个解释有相当力量。若它能完全解释数据，“接力再生事件”只是旧变量的重新命名，应当撤回。

14.1 四组现场实验

可以在同一学校、同一学科、相近教师与任务难度下，进行班级层面的集群随机实验，至少三十个班级，分为四组：

常规高质量教学组：讲授、练习、反馈与当堂合作完整；

延迟迁移组：增加十四天后的结构同型任务，但不安排跨主体接力；

接力组：安排跨组接手和版本推进，但不把学生结果写入下一轮课程；

完整回路组：同时包含延迟再生、跨主体推进和正式环境回写。

四组都控制总课时、教师培训、作业量和即时测验。主要结果不是一次总分，而是下一单元中新的结构迁移、未见任务中的问题识别、跨组接手成本、学生提出问题进入正式课程的比例，以及课堂规则变化后是否提高后续再生概率。

若完整回路组的优势完全由个体延迟迁移得分与互动次数解释，加入“回写”变量没有增量预测力，本章的核心机制受挫。若接力组与完整回路组没有差异，环境回写不是必要条件。若延迟迁移组已取得全部效果，跨主体接力只是一种教学形式而非事件构成。

14.2 六条证伪条件

第一，增量失效。控制先前能力、即时成绩、投入时间、认知耐力、动机、自主权感知与同伴网络后，接力再生事件对后续未见任务表现没有增量预测力，则核心判断为伪。

第二，顺序失效。若大样本课堂日志显示，环境规则改变经常早于学生跨主体修正，且这种自上而下改变与后续再生同样相关，那么本章提出的“失效——外置——接力——回写——再生”顺序不是承重机制。

第三，换组失效。延迟后重新分组时，所谓接力效果全部消失，而原小组内仍维持高表现，说明观察到的是共享语境或社会记忆，不是可跨主体再生的结构。

第四，激励替代。仅通过提高奖励、公开排名或复杂过程分，即可复制完整回路组在未见任务、规则改写和跨组接手上的全部结果，则事件无需诉诸新辨别的生成，经济激励足以解释。

第五，材料路径失效。在艺术、设计或实验课程中，控制投入时间和先前技能后，不同材料与媒介对辨别形成、接力方向和后续回写没有任何差异，则艺术学关于制作路径构成思考的作用没有兑现，本章的跨域机制应收缩。

第六，低成本记录不一致。若课堂视频编码判定出现大量接力再生，而作业版本、题库、量规和备课记录完全没有对应改变，且后续访谈显示学生也无法重建，则人工观察可能把热闹互动误判为事件。

其中第一条是最便宜的实做检验。学校已有的成绩、作业版本和教学平台日志可以支持初步分析，无需新增高风险操作。第二条是正向顺序读数：真正的事件应表现为学生辨别的跨主体变化早于正式规则改写，规则改写又早于后续再生概率提高。若只观察到相关而没有顺序，不能支持机制。

14.3 被证伪以后能学到什么

若实验表明延迟迁移、合作学习与响应性教学的现有变量已经完全覆盖结果，我们会学到：课堂学习没有必要被提升为一种新的跨主体存在，最合适的理论仍是个人学习加社会条件。干预方向也会改变——学校应分别优化迁移训练、互动结构和教师反馈，而不必维护一条完整接力回路。一个可被这种结果取消的概念，才不是用定义保护自己的口号。

14.4 写死日期的赌注

本章提出如下可检查赌注：到2031年12月31日以前，至少出现一项覆盖三十个以上同质班级、跨越一个学期、经过同行评议的研究，证明在控制即时成绩、个体延迟迁移、

互动次数、投入时间和先前能力以后，跨主体接力与课堂规则回写的联合事件仍能显著预测下一单元的未见任务表现。

以下情况不算命中：只做横断面问卷；只报告师生主观感受；只比较两个高度异质国家或学校；只证明合作学习有效；只发现脑间同步与成绩相关；只要求教师自行声称“课程因学生而改变”；未公开事件编码规则与原始版本链。若截至该日期没有满足条件的证据，本章关于“新事件”的主张应降格为一种课堂设计启发，而非独立理论对象。

十五、不适用边界：不是所有学习都要成为接力

第一，本章讨论的是“课堂学习”，不是所有学习。一个人在独处中阅读、思考、练习，完全可能发生深度学习；它不需要跨主体接力。本章只反对用少数个体的私人掌握替课堂共同体结算。

第二，新手学习需要大量直接教学。对基本事实、符号、核心程序和安全规范，过早要求开放生成会增加无谓负荷。Kirschner、Sweller 与 Clark（2006）关于新手指导的争论提醒我们，最小引导并非普遍优越。接力再生应建立在足够先备知识之上，而不是把教师讲解污名化。

第三，安全关键场景不能为了观察换手率而制造真实风险。医学、航空、化工与应急训练中的接力应在模拟环境或低风险环节进行。规则回写必须经过专业审核，学生发现不能直接覆盖安全标准。

第四，语言表达不能成为唯一入口。神经多样性、语言背景、残障与文化规范会影响公开发言。接力可以通过图像、动作、代码、作品、工具布置和匿名文本发生。若测量只记录口头互动，会系统性高估语言优势者，低估默会与非语言学习。

第五，敏感经验并不适合公共接力。创伤、身份、家庭与健康议题可能需要个人边界。课堂不应以“共同学习”为名迫使学生贡献私人材料，也不能把拒绝公开解释视为没有学习。

第六，环境回写不是学生拥有无限否决权。教师、课程标准和专业知识仍有责任。回写可以是问题、材料或评价细节的改变，不等于每次课堂都由即时偏好决定。真正的开放包含拒绝某项修改并说明理由的制度能力。

这些边界防止新概念吞并所有教育价值。接力再生是判断“课堂作为共同体是否学会了”的工具，不是衡量每一个人、每一项内容和每一种教学法的总尺度。

十六、自反：本章还没有证明它所主张的东西

按照本章自己的判据，这篇文章目前至多完成了“外置”：它把一个判断写成可供阅读的文本，列出结构、案例和证伪条件。它尚未证明接力再生事件已经发生。即使读者认同、引用或复述“接力再生”，也可能只形成脚本流通。只有在原文不在场时，教师或研究者能够在新的课堂问题中重建这一辨别；另一位参与者把它修改成本章没有给出的形式；某项课程设计、评价规则或研究方案因此发生可追踪改变，这篇文章才成为自己所描述事件的一部分。

这个自我定位有两个具体后果。第一，本章不能以出版、流传或被称为“新范式”来证明自己。第二，后续应用者若只制作“接力再生量表”、要求教师打分，却没有保存版本链和环境改写，恰好会成为本章批评的对象：概念被广泛传播，所指事件却没有再生。

本章还存在一个解释不了的洞：某些课堂中，一条辨别可能多年后才在学生的生活里重新出现，且没有留下可追踪的接力记录。它可能是真实学习，却超出学校研究的观察窗口。另一个洞是，集体改变有时来自匿名氛围、无明确发起者的慢变化，事件链难以分段。若强行寻找“谁接了谁”，会重新制造本章试图取消的所有者偏见。未来研究需要允许无中心、低可见度的发生形式。

十七、结论：课堂真正留下的，不是一批答案，而是一条能够继续发生的路

认知科学让我们看到，即时表现与较慢能力变化可以分离，关系的动态改变比静态高低更有信息；经济学让我们看到，努力、注意力、认知耐力与参与受制度价格、复杂激励和自主权配置塑造；艺术学让我们看到，思考并非完成后才被表达，而在材料、限制、批评和制作中形成。三门学科各自不能回答另两门的问题，却共同推翻了一个深藏在课堂评价中的前提：学习总有一个稳定主人，也总有一个可以在课末结算的时点。

接力再生事件由此出现。它不是把个人迁移、合作互动和课程改进相加，而是规定一种新的存在条件：新辨别必须在撤除原条件后再生，必须离开最初提出者被他人推进，必须改变下一轮共同活动。这个事件一旦发生，课堂拥有的就不再只是一批知道答案的人，而是一条能够在不同人、不同材料和不同问题之间继续生长的路。

因此，判断一堂课是否真正发生学习，最锋利的问题不是“学生听懂了吗”“课堂活跃吗”“目标达成了吗”，而是：**十四天后，原来的例子、话术和组员都不在场时，那条辨别能否重新长出；它能否被另一个人接住；课堂会不会因此问出一个昨天还不存在的问题。**

若答案是否定的，那堂课可能完成了教学、练习、激励、合作或创作，却尚未形成本章所说的课堂学习。若答案是肯定的，知识便不再只是被储存的结果，而成为一个共同体继续产生未来的能力。

接续。本章的接力率在第十一章被列入九个读数的总表，归在“需要一次专门设计的任务”那一档。它测的是第一编小结所说的**传递**，与第三章的构成、第一章的显露分工；若三者两两相关高于零点八，本章应与另两章合并——判决与合并后的处理办法写在第一编小结第三节。

第三章 支架撤掉以后还能重新进去

判断卡片：课堂生成的那样东西不持续显露；不显露不等于不存在，而现有测验只查旧答案能否被直接提取，因此系统性地把休眠读成遗忘。

三个理论入口：物理学 × 心理学 × 艺术学

提要。课堂上常见三种被误认作学习的现象：学生当堂答对、延迟后仍能回忆、群体参与高度活跃。它们分别提供表现、记忆与经验的证据，却都不足以单独回答“课堂里究竟生成了什么”。本章将物理学中的“学习留下响应地形与结构印记”、心理学中的“预测误差——提取——迁移路径”以及艺术学中的“具身、媒介化与社会延展的认知”置于同一问题中考察。三类理论彼此冲突，却共享一个隐含前提：学习是一项已被某个学生、某条认知路径或某个课堂场域占有的属性。本章以这些理论自身的材料推翻该前提，提出“课堂可再入生成体”：当原例题、教师提示、工具配置和角色分工被撤去或置换，学习者仍能重建关键关系，并据此改变新情境中的问题、表征、工具、行动或协作结构。真正的课堂学习不是知识量增加，不是一次成绩改善，也不是一场强烈的共同经验，而是这一新型关系存在物的形成。本章给出二维辨别格、无情态词判据、五类课堂读数、十组近邻分界、可证伪研究设计及伦理边界。

一、问题不在“学生学到了多少”，而在“课堂生成了什么”

一间课堂可以非常繁忙：教师讲解清楚，学生记满笔记，提问接连不断，小组讨论热烈，测验分数也立刻上升。所有可见信号都在宣布“学习已经发生”。然而，数周后只要换掉例题表面、撤走教师提示、改变工具或调换同伴，原先的理解便可能迅速消失。课堂并未留下能继续工作的东西，只留下了完成那一节课的路径依赖。

这不是简单的“遗忘”。遗忘仍然假定曾经有一份完整知识存入个体，后来强度下降。更棘手的情况是：当堂的成功从一开始就可能不是学习，而是教师、教材、同伴、界面和任务格式共同搭起的一次性通道。学生沿通道走到了正确答案，却没有获得重新开路的能力。课堂结束后，通道拆除，成功随之撤回。由此，“当堂表现”与“学习”不是同一件事，“延迟记忆”与“课堂学习”也不是同一件事。

本章追问的不是课堂怎样更有效，也不是哪一种教学法优于另一种，而是一个更基础的问题：何谓真正的课堂学习？这个问法要求的答案不是一份方法清单，而是一个可以把两种外表相似的课堂分开的存在论判断。两节课可能拥有相同的正确率、相同的参与度、相同的作业质量，甚至相同的延迟测验成绩，其中一节形成了学习，另一节只完成了任务。能够解释这种差别的，不应再是“更多”“更深”“更积极”之类程度词，而应是课堂里是否生成了一种此前不存在、课后仍能重新进入并继续产生差异的关系结构。

为寻找这一对象，本章选择三个彼此不相容的理论入口。物理学把学习读成系统响应地形的改变；心理学把学习读成预测、错误、提取和迁移的时间路径；艺术学把学习读成身体、媒介、行动与共同体共同构成的经验场。三者各自都能说明课堂的一部分，也都能用自己的证据否定另外两者的充分性。真正的新答案不在三者的平均值里，而在它们无法彼此消化的裂缝中。

二、三种几乎正确的答案

关于课堂学习，最常见的答案可以归为三类。

第一类把学习等同于“留下了东西”。学生能复述概念、执行步骤、识别题型，说明记忆或技能系统发生了稳定改变。这一答案抓住了学习必须具有时间延续性的事实，却很容易把条件反射、格式熟练和题型识别也纳入“理解”。一个学生可以在固定提示下连续做对二十道题，却在题目换一种语言后不知从何下手。稳定并不自动意味着可迁移。

第二类把学习等同于“走过了一条有效的认知路径”。学生先作预测，再遭遇差异，通过解释、提取、反馈和重新编码形成可迁移表征。这一答案解释了为何主动提取常优于重复阅读，也解释了为何错误并非越大越好。然而，它仍倾向于把最终成果放在个体内部：学习被描述为一个人的表征更新，即使这个更新事实上依赖特定教师的追问、同伴的反驳、某种图示和一套角色分工。

第三类把学习等同于“经验方式发生了改变”。艺术教育、具身认知和延展认知提醒人们，理解并非只是一段命题进入大脑，而是看、听、触摸、移动、制作、判断和共同注意的重组。它把课堂从信息传输场恢复为感知与行动的世界。但强烈、开放和富有创造性的经验也可能只发生一次。一次动人的戏剧排演、一场高度同步的讨论、一次让全班惊叹的实验，都可能在第二天无法重建。

三种答案分别守住了稳定、路径和场域。问题是：稳定可以没有迁移，路径可以依赖借来的支架，场域可以没有持续。真正的课堂学习必须同时通过这三种否定，却又不能只是把三项指标相加。三项相加仍然默认“学习”本来就是一个现成对象，只需用更完整的量表把它测出来。本章将论证，课堂学习不是被更准确地测量出来的旧对象，而是在稳定印记、可重构路径与具身共同体相互制约时才出现的新对象。

三、物理学的前沿启示：系统会成为它所学之物

近年的“物理学习”研究把学习从神经网络和算法模型扩展到可调电阻网络、机械网络、磁性材料和其他多体系统。Stern 等人指出，可学习物理网络通过改变内部相互作用来降低任务代价；训练不是只让输出暂时接近目标，而是把任务的代价地形与系统自身的物理地形耦合起来。训练后的网络遭遇随机扰动时，其响应能够反向显露它被调谐去完成的功能（Stern et al., 2025）。更早的工作还表明，学习会在物理系统的 Hessian 结构上留下可辨认的架构印记（Stern, Liu, & Balasubramanian, 2024）。

这一理论给课堂研究带来第一条严格约束：一次正确输出不是学习证据。只有系统面对扰动时的响应组织发生了稳定改变，才能说训练在系统里留下了东西。对应到课堂，教师给出熟悉例题时的正确答案，好比网络在训练输入上的低损失；真正有辨别力的不是这个输出，而是当题目表面、信息顺序或工具约束改变时，学生响应空间的形状是否改变。

任务自适应物理储备池研究进一步显示，同一个物理系统在不同热力学相位中会呈现不同的记忆、非线性和计算特性；通过重构系统所处相位，可以让它适配不同任务（Lee et al., 2024）。这意味着，能力不是一个脱离环境的常数。同样一个学生在纸笔、实验装置、身体动作、图像或同伴协作中，可能拥有不同的可达反应集合。课堂学习若只测一种表面通道，就可能把“系统在该相位里可用”误写成“学生已经拥有”。

物理学的锋利之处在于，它不询问系统是否觉得自己理解，也不询问过程是否投入，而是问：扰动之后，系统还剩下什么？哪些响应模式变得容易，哪些模式变得困难，哪些原本不可达的输出现在能够被稳定产生？课堂研究由此得到一个非心理化判断：撤掉原有输入格式后，学生的反应地形是否被重新塑造。

但是，物理学也会制造一个危险的过度推论。若稳定结构改变就是学习，那么高度自动化的错误习惯、对题型的僵化识别、对教师语气的条件反应，也都可被称为学习。系统确实“成为了它所学的东西”，但它所学的可能只是如何在原训练分布中成功。物理结构能证明留下了印记，却不能单独证明印记具有理解、迁移和自我重构能力。

四、心理学的前沿启示：学习是一条预测——提取——迁移路径

心理学把物理学的静态印记放回时间中。学习不是任意差异留下痕迹，而是学习者基于既有模型作出预测，遭遇结果，计算差异，重新组织表征，并在之后的提取中改变可用性。预测误差理论说明，没有预期就没有可更新的差异；提取研究说明，学习结果不能只看再次输入时的熟悉感，而要看能否在缺少原材料的条件下主动生成。

最新研究也使“制造惊讶”这一流行口号受到限制。Pupillo 等人的实验表明，预测误差对记忆的影响取决于预测结果：当预测正确时，较强误差可能增强后续记忆；当预测错误时，较强误差反而可能损害记忆（Pupillo et al., 2023）。这意味着，课堂中的认知冲突不是越大越好。若新信息与学生已有模型之间没有可工作的桥，冲突只会把经验标记为异常，而不会形成可重建的关系。

2025 年的时间因果学习模型进一步指出，学习者不只在估计“结果与预期差多少”，还在利用事件之间的时间间隔结构推断共同原因（Hamou, Gershman, & Reddy, 2025）。因此，同样的反馈若出现在不同时间位置，会构成不同的学习路径。教师过早纠正，可能使学生从未形成自己的因果结构；反馈过晚，又可能让错误路径先行固化。学习不是把正确答案放在过程末尾，而是安排差异何时出现、由谁发现、如何被重新提取。

提取学习研究则区分了“当下容易”与“长期可用”。重复阅读往往产生流畅感，学生误以为已经掌握；主动提取在当下更费力，却能改变后续保持和迁移。Karpicke 在 2025 年的综合回顾中再次强调，提取不只服务于复制原答案，也可支持推断和在不同测试格式中的应用。真实小学课堂的研究也发现，配合反馈的提取练习相较于重复阅读具有更好的学习结果（Franzoi et al., 2025）。

神经学习研究还显示，奖励预测误差与动作预测误差并非同一信号。Greenstreet 等人发现，动作预测误差信号能够强化重复的声音——动作关联，却不能独自支持奖励导向学习；它需要与奖励预测误差回路协同（Greenstreet et al., 2025）。这对课堂尤其重要：学生“越来越顺手”可能只是动作重复被巩固，不等于知道何时不该继续该动作。

心理学由此给出第二条约束：稳定痕迹必须经过撤除提示后的提取，并在表面改变、关系同构的新任务中重新组织。没有这一条，课堂可能训练出高度稳定却只适用于旧格式的反应。

但心理学也有自己的盲区。许多实验把任务、材料、反馈和社会关系当作外部条件，学习则被归入个体内部的记忆表征。真实课堂中的问题恰恰是：一个人能够提取的东西，

可能由同伴、图示、教室节奏、工具位置和教师话语共同构成。若这些构成因素被视作“噪声”或“支架”，心理学就会把共同完成的认知错误地结算到单个头脑名下。

五、艺术学的前沿启示：理解发生于身体、媒介与共同体

艺术学对“知识在头脑中储存”的图景提出更根本的挑战。艺术经验不只是接收关于作品的命题，而是身体的姿态、感官的选择、媒介的阻力、动作的节奏与共同体的判断规则共同塑造可见之物。Coëgnarts 在 2025 年提出的具身审美认知论强调，身体与感官经验会塑造艺术的智性价值，艺术尤其能够改变知觉学习：学习者不只是知道了新的说法，而是开始看见此前看不见的动态关系（Coëgnarts, 2025）。

这种变化对课堂具有决定意义。学生学会透视，不是记住“近大远小”四个字，而是纸面、手眼动作、空间想象和误差修正形成新的看法；学生学会音乐句法，不只是能说出终止式，而是身体期待、时间张力和听觉分组发生变化；学生理解一首诗，不只是提取主题，而是语音、停顿、意象、历史语境和共同朗读重新组织了意义的入口。

2026 年前后出现的“社会延展心智”美学论述进一步把认知从身体扩展到制度与共同体。Drummond 认为，认知制度是由行动者互动生成的资源池，个体把一部分判断、解决问题和行动能力外置其中；在某些情况下，撤去这一制度，也就撤去了相应认知能力（Drummond, 2026）。艺术世界的批评传统、分类、档案、展览和共同注意，并非只是帮助个体思考的背景，而可能构成思考本身。

Walsh 对 Noë 与西田几多郎的比较也把艺术理解为艺术家、作品与环境之间展开的互动过程。艺术之所以具有认识作用，不只因为它表达一个结论，而因为它扰动并显露生活原本的组织方式（Walsh, 2026）。这为课堂提供第三条约束：真正学习必须改变学习者与对象相遇的方式，包括身体如何进入、材料如何抵抗、他人如何接续、什么被共同体判为值得注意。

然而，艺术学的优势也可能成为它的弱点。一次经验可以极其开放、强烈、共同、具身，仍然未必形成可再进入的学习。某次戏剧排练中，学生在导演、灯光、音乐和同伴推动下产生了突破性的表达；换一组人、换一个空间，他却无法重建那条关系。艺术事件已经发生，学习对象却未必留下。若把经验强度、感动或创造性本身当作学习，便会把不可重现的一次性涌现误写成课堂的持久成果。

六、三类理论为什么不能直接相加

表面看，解决办法似乎很简单：真正学习等于“稳定结构改变+有效认知路径+具身社会经验”。但三项相加仍然没有处理三者之间的冲突。

物理学可以接受一个没有理解的稳定反应，只要系统被调谐并能在扰动下显露功能；心理学必须拒绝，因为条件化反应可能无法迁移。心理学可以接受一个个体在延迟迁移测验上成功的表征更新；艺术学仍会追问，这一更新是否改变了感知、行动和与他人共处的方式。艺术学可以接受一次重新组织身体与世界关系的事件；物理学则会追问，它是否留下可辨别的稳定响应差异。

更重要的是，三者对“学习在哪儿”给出不同答案。物理学把功能放在网络整体的相互作用中，却常以训练后结构作为最终读数；心理学把学习放在个体跨时间的加工路径中；艺术学把认知放在身体——媒介——制度的关系场里。若试图判定谁才是正确位置，讨论会退化为学科争执。因为真实课堂中，三种位置都可能不可缺少。

真正的突破来自另一问：三类理论在争夺学习的所在地时，共同默认了什么？它们都把学习当作一项可被某处占有的属性。要么由网络结构占有，要么由个体记忆占有，要么由当下经验场占有。于是，研究者总在问“谁学会了”“存在哪里”“用什么指标读出”。

可是物理网络的功能并不由任何单个元件持有；它存在于耦合关系与整体地形中。社会延展心智同样表明，某些认知能力在制度被撤去时会一同消失。课堂中也可能存在一种既不完全属于个体、又不等于当下气氛的东西：它在教师、学生、符号、工具、身体和任务约束之间形成，课后可以休眠，遇到新问题又被重新进入。它不是“大家共同知道的内容”，而是一种可以重新生成关系的组织。

七、共同前提的倒塌：学习不必先有一个持有人

传统课堂评价像一本所有权账簿：知识必须记到某个学生名下，能力必须由个人测验结算，合作只被视作帮助个人表现的外部变量。即使研究承认情境、工具和群体重要，也常在最后一步把结果重新归还给个体：环境提供支架，真正的学习仍在头脑里。

物理学习自身的材料足以推翻这一前提。训练后的网络功能由内部相互作用共同构成，随机扰动揭示的是整体被调谐到何种功能，而不是某个元件储存了哪条规则。把一个关键元件单独取出，它并不携带完整功能；功能在关系中。换言之，学习可以拥有可读的稳定性，却没有单一持有人。

艺术学的社会延展心智把这一步推进到人类实践：某些判断能力通过批评传统、符号系统、工具和共同体规则才成立。撤去制度，不只是降低表现，而是取消了相应认知过程的构成条件。于是，“没有一个人独自拥有”不再等于“没有学习”。

由此，课堂学习的核心单位不能再预设为学生内部的知识单元。它可能是一种分布式但可重入的关系存在。它不以某个人是否能孤立复述为唯一判据，也不以当时的集体同步为判据。它的存在方式类似一条已经被共同开出的路：道路不属于任何一个脚步，停用时也可以暂时沉寂；只要路的生成规则仍可被重新唤起，新的行者便能沿它进入并继续开拓。

这里必须避免浪漫化。不是所有合作成果都属于这一类。若少了某位教师便完全不能运行，若新成员只能服从旧分工，若工具一换就失效，那么课堂只有一次性依赖，没有形成可再入对象。新对象必须能在部分构件替换后重组自身，否则它只是临时联盟。

八、新判断：真正的课堂学习是“可再入生成体”的形成

本章提出“课堂可再入生成体”这一概念。它是指：课堂中的人、符号、工具、身体动作、任务约束和判断规则经过教学活动形成一种新的关系组织；当原例题、教师提示、界面、材料或角色分工被撤除和置换时，参与者仍能重新生成其中的关键关系，并用该关系改变新情境中的问题、表征、工具、行动或协作方式。

这一判断可写成三重否定：真正的课堂学习不是学生内部知识量的增加，不是一次任务表现的改善，也不是一场投入、同步或感人的共同经验，而是可再入生成体的出现。

“可再入”不是重复。重复是沿原路径再次到达旧答案；再入是在入口改变后重新生成使路径成立的关系。学生不是记住教师怎样解题，而是在陌生题面中重新发现哪些量必须保持、哪些差异具有因果意义、哪一种表征能使问题可解。

“生成”也不是泛指创造。它要求新关系产生后果：问题被重新提出，工具被重新选择，证据顺序被改写，分工发生改变，或原本不可见的对象变得可见。只说出一个新观点而不改变后续行动，不足以构成生成。

“体”强调这不是个体头脑中的一个属性，也不是课堂气氛。它有边界、有组成、有进入条件、有失活方式，可以被不同成员在不同时间重新激活。某位学生可能先成为入口，另一位学生在之后的任务中接续；某个图示可能退出，身体动作或语言隐喻接替其功能。持有人可以换，关系仍能重组，这正是它与个人能力和一次性集体事件的区别。

可再入生成体一旦形成，课堂就获得一种超出当堂总表现的新能力：它可以在未知问题到来时重新组织自己。这不是“学会更多答案”，而是生成了使新答案可能的关系。

九、它为什么是一类新存在物，而不是一个新指标

一个概念是否真正提出新对象，可以做一个删除测试：删掉本章的测量办法之后，被说的东西是否以同样方式存在？若答案是“仍然存在，只是难以测量”，本章最多发明了一个新指标。可再入生成体的主张更强：传统账本没有为它设置存在位置。

在通常的课堂账本中，只有学生、知识点、成绩、参与行为和教学活动。一次关系由不同成员共同生成、暂时没有人独自占有、在课后休眠、之后被新成员从另一入口重新激活，这一过程会被拆散登记：第一次记为讨论，第二次记为迁移，第三次记为合作，第四次记为个人表现。没有任何一栏承认“同一个关系对象跨构件再次出现”。

因此，可再入生成体不是把已有分数换一种算法，而是把原先被当作余数的东西提升为研究对象。它的边界不由“谁知道”决定，而由“哪些构件被替换后，关键关系仍能重组”决定；它的寿命不由记忆强度决定，而由休眠后能否再入决定；它的变化不由答案数量决定，而由每次再入是否回写自身决定。

如果未来研究证明，所有所谓再入都能被个人先验能力、记忆强度和普通迁移完整解释，那么这一新对象应被取消。反过来，若存在这样的课堂：没有任何个人在第一次结束时拥有完整规则，但通过共享制品、动作和判断传统，后来不同成员能重建关系并产生新行动，那么传统个体账本无法表达它，可再入生成体便获得独立的存在理由。

这里的“新”不是宣称前人从未谈过关系、迁移或共同体，而是指出这些研究通常把它们结算为某项既有能力的原因或条件。本章把可跨构件再生的关系本身作为结果变量。原因与结果的位置发生了改变。

十、可再入生成体的四个构成条件

第一，扰动留痕。教学结束后，系统面对扰动的响应集合必须发生稳定改变。扰动包括换题面、换材料、换顺序、减少提示、改变时间压力或引入冲突证据。若正确率只在原格式上提高，说明通道被熟练化，响应地形未必被改写。

第二，撤架再入。原例题、教师话语、标准步骤和固定同伴撤去后，学习者能够重建关键关系。关键关系不是答案，而是让答案成立的约束：守恒、比例、证据责任、叙事视角、节奏张力、因果方向或评价规则。再入不是从记忆中把整套过程复制出来，而是在缺口又重新找回组织原则。

第三，异境生差。重建的关系必须在新情境中产生一项此前没有的差异。最低限度包括五类之一：提出一个改变任务结构的问题；创造一种更合适的表征；更换工具或测量方式；改变行动顺序；重组协作角色。若学生只是把旧规则套到新例子上，可能构成迁移，却还未充分显示生成体的开放性。

第四，持位可换。课堂层面的生成体不能绑定某一位教师、尖子生或固定软件。至少一项核心功能能够在成员或媒介替换后被接续。这里不是要求任何人都能替代任何人，而是要求组织不把发起权永久锁在原持有人手中。一个团队每次遇到困难都等同一位学生开口，说明团队借用了个人能力，没有形成课堂生成体。

四个条件之间存在反向约束。扰动留痕防止把短暂灵感当作学习；撤架再入防止把稳定习惯当作理解；异境生差防止把复现当作生成；持位可换防止把个人天赋或教师控制误写成课堂成果。缺少任何一项，都可能没有“学习的证据”，却没有本章所定义的课堂学习对象。

构成条件	可观察证据	主要排除的假阳性
扰动留痕	题面、顺序、工具或时间压力改变后，响应集合仍显示结构差异	当堂正确与格式熟练
撤架再入	原例题、提示和固定分工撤除后，关键关系被重新建立	借来的支架与教师代为组织
异境生差	新关系改变问题、表征、工具、行动或协作之一	机械套用与口头新颖
持位可换	成员或媒介替换后，核心功能能被接续	尖子生救场与固定平台依赖

表1 可再入生成体的四项构成条件

十一、二维辨别格：四种外表相近的课堂结果

为了把新对象与常见假阳性分开，可以使用两条轴。横轴问：教学是否改变了延迟后的响应地形，即面对扰动时哪些关系变得可达。纵轴问：在原提示链撤除后，课堂成员能否让关键关系在新情境中重新生成并产生后果。

右下格最容易被误判为优质学习：学生表现稳定、测验分数高、步骤熟练，但换掉题面或撤走支架后无法重新开路。它通过几乎所有常规形式审查，因为短期结果改善，失效在当下没有信号，而且越标准化越容易加固。本章的必要性主要体现在识别这一格。

左上格则提醒人们，创造性事件不等于学习形成。某次讨论借助恰到好处的教师追问和群体氛围产生了全新解释，但延迟后没有任何成员能重建，这是一场真实而有价值的生成事件，却尚未凝结为可再入生成体。

只有右上格同时具备结构留痕与再入生成。它既不会把流畅当学习，也不会把一次性突破当持久成果。

再入生成 / 响应地形	未形成稳定响应改变	形成稳定响应改变
不能在撤架后重新生成	暴露或短暂激活 当堂出现，课后无可重建关系	提示依赖的稳定能力 高分、熟练、可重复，但入口一换即失效
能在撤架后重新生成并产生后果	一次性共同涌现 现场产生新关系，但未形成延迟留痕	可再入生成体 关系在新入口中重建，并改写后续组织

表 2 课堂学习的二维辨别格

十二、一个不含情态词的判据

本章给出如下操作性问句：

撤掉原来的例题、教师提示和同伴分工，换一个表面陌生但关系同构的任务，学生能否自行重建关键关系，并据此改变下一步的问题、表征、工具、行动或协作方式？

这句话不使用“真正、充分、深入、有意义、合理”等判断词。它可以被带到课程文件、课堂录像、学习平台日志、草稿版本和小组分工记录中查验。

判据中的每个部分都有排除功能。“撤掉原来的例题”排除题型识别；“撤掉教师提示”排除教师代为维持因果链；“撤掉同伴分工”排除固定角色借力；“表面陌生但关系同构”把随机难题与迁移任务分开；“自行重建关键关系”排除答案猜测；“改变下一步”排除只会口头解释却不产生行动后果。

判据并不要求学生完全孤立。可再入生成体本来就是分布式的，学生可以使用新工具、与新同伴协作、查询资料。被撤去的是原来那条已知提示链，而不是所有环境。真正要检验的是：系统能否在构件变化后重新组织，而不是个人能否在真空中完成任务。

十三、五种具体读数：同一句判据不会只给“是”或“否”

情境一：数学课上，学生能在原有函数图像题中熟练求导，但把问题改写为“某种药物浓度何时变化最快”，教师不再给坐标轴和公式提示后，学生仍只寻找熟悉题型。读数是“稳定痕迹存在，关系不能再入”。

情境二：历史课进行一次精彩辩论，学生在教师连续追问下重新理解史料立场，现场出现新问题；两周后换一组史料，所有人重新等待教师追问。读数是“场域生成出现，未形成延迟留痕”。

情境三：物理实验小组中，一名学生在传感器失效后改用视频帧估计速度并重建守恒分析，其他成员只能照做。读数是“个体可再入，课堂持位尚未转换”。

情境四：语文课形成一套从叙述可靠性、语调和证据空白切入文本的阅读方式；换一篇陌生小说、调换组长后，学生自行改变问题顺序，并出现不同于原课的解释。读数是“课堂生成体形成并完成一次持位转换”。

情境五：AI辅助写作课中，学生借助固定提示词生成高质量论文；当模型、界面和提示模板改变后，学生能够重新定义问题、设计证据检查并重组人机分工。读数是“媒介替换后完成再入”。若只能寻找原模板，则是“工具依赖的稳定表现”。

五种读数分别为：无再入的稳定、无留痕的涌现、个人拥有的再入、课堂层面的再入、跨媒介的再入。判据因此不是命题的改写，而能把同样“表现不错”的课堂分成不同机制。

十四、课堂学习的时间结构：发生、凝结、休眠与再入

可再入生成体不是一节课结束时突然出现的成品。它具有四阶段时间结构。

发生阶段，原有关系被扰动。学生的预测、习惯或感知方式遇到无法被旧路径吸收的差异。差异必须处在可工作范围：太小只造成同化，太大则成为噪声或挫败。教师的任务不是宣布冲突，而是让学生拥有对差异的初步解释权。

凝结阶段，差异被组织成跨媒介关系。口头解释、图像、动作、公式、材料反馈和同伴反驳之间出现可以互相翻译的结构。凝结不是统一答案，而是形成“哪些变化不改变问题、哪些变化会改变问题”的共同约束。

休眠阶段，课堂停止提供原支架。生成体不持续显露，它可以暂时不被使用。休眠与死亡不同：休眠对象在新线索出现时仍能被重新组织；死亡对象只能靠重教恢复。传统测验常把休眠误作遗忘，因为它只查旧答案是否可直接提取。

再入阶段，新情境提供不同入口。若学习者能重新发现关键关系，并让其改变当前行动，生成体完成一次存在证明。每次再入都会改写它自身：新情境可能暴露旧关系的边界，促使工具、概念或协作结构更新。因此真正学习不是从课堂向生活搬运一个不变包裹，而是一个关系对象在不同情境中反复生成自己。

这一时间结构解释了为何有些课当时看似失败，之后却产生深远学习；也解释了为何有些课当时完美，之后毫无作用。当堂表现主要观察发生和局部凝结，真正的证据来自休眠后的再入。

十五、课堂学习的四种死亡方式

可再入生成体一旦形成，也会死亡。知识的死亡不只表现为个人忘记，还表现为关系失去重新生成自己的条件。课堂中至少有四种常见死亡机制。

第一，提示链寄生。教师为了确保课堂顺利，逐步把提问、线索、表征选择和判断标准全部外包给自己。学生的表现不断改善，但每一次改善都加强对原提示链的依赖。久而

久之，课堂形成的不是再入能力，而是一套对教师微小信号高度敏感的寄生系统。教师一撤离，关系即刻断裂。

第二，结算替代。学校用一次测验把仍在发生的学习提前结算为分数。学生和教师随即把目标从“关系能否继续生成”改为“哪些表现会被记账”。能在新情境中产生差异的探索，因为短期无法计分而被舍弃；能快速提高成绩的格式训练被保留。结算并非只测量学习，它会回过头改造学习对象。

第三，角色锁定。某位学生长期负责提出问题，另一位负责计算，第三位负责美化汇报。团队效率很高，分工却把关键关系封存在固定位置。角色一旦变化，课堂整体能力下降。此时课堂拥有一个可运行组织，却没有持位可换的生成体。

第四，媒介单栖。一个概念只在某种符号、软件或实验装置中成立。学生在该媒介中表现优异，换成语言、图像、身体动作或不同工具便失去关系。媒介单栖尤其容易发生在AI时代：提示模板和界面承担了大量看不见的组织，学生把平台稳定性误作自身学习。

四种死亡方式共同说明，课堂学习不会因为曾经形成就永久存在。再入生成体必须通过不同时点、不同媒介和不同持位的再次发生维持自身。教育质量因此不能只问“教过没有”，还要问“这个关系最近一次由谁、借什么入口、在何种差异中重新活过”。

十六、四个尺度：个人、小组、课堂与制度

“课堂学习”不能把个人、小组、班级和制度混成同一层。可再入生成体可以出现在不同尺度，但每个尺度有不同的存在证据。

个人尺度上，关键是一个学习者在入口变化后重建关系并改变行动。个人可以借助工具和他人，但发起与修正不能完全由外部替代。

小组尺度上，关键是功能分布和持位转换。小组不要求每个人掌握全部内容，却要求问题重述、证据否决、工具选择和协调能够在成员之间续接。小组成果再好，只要核心功能永久锁在一人身上，就只是个人能力的放大器。

课堂尺度上，关键是跨小组传播而不退化为口号。一种关系若能通过公共制品、讨论规范、案例库和共享动作进入不同小组，并在各组中生成不同但可辩护的结果，课堂生成体才成立。若只有统一答案被传播，关系可能已经在传播中死亡。

制度尺度上，关键是跨届、跨教师和跨课程的再入。课程档案、实验传统、工作室制度和评价规则可以成为认知制度，使后来者获得前人没有逐项写下的能力。但制度也最容易把活关系固化为流程。制度尺度的证据不是文件存在，而是新成员能否借文件、物件和实践重新生成问题，并允许新情境回写制度。

四个尺度可以不同步。个人形成而课堂未形成，课堂形成而制度未接住，制度保存而个人无法再入，都很常见。评价时必须先说明判断的是哪个尺度，不能用个人测验替代课堂判断，也不能用学校制度文本证明学生已经学习。

十七、六类课堂中的新解释

以下案例不是教学处方，而是说明同一判断如何在不同学科产生可区分预测。

14.1 数学：从“会解题”到“重建不变量”

数学课堂最常见的假阳性是表征格式被固化。学生记住“看到某词就用某公式”，在训练分布内获得高正确率。可再入生成体要求更高：当符号、图像和现实叙述互换时，学生能重建哪个量保持、哪个关系被约束、哪一步决定可解性。

例如，比例问题不应只在“单价——数量——总价”中被检验。可把同构关系放入地图尺度、溶液浓度、音乐节拍或人口模型，并撤去“比例”标签。若学生改变表征后仍能找到可比单位，说明关系可再入；若只会套公式，说明训练留下稳定通道，却没有生成体。

这一判断还改变错误的地位。错误不是等待教师纠正的偏差，而是显示学生响应地形的探针。两个学生给出同一错误答案，可能来自完全不同的关系组织。教师若直接补步骤，只修输出；若让错误在不同表征中继续运行，才可能暴露需要重构的约束。

14.2 物理实验：从“得到标准结果”到“重组测量系统”

实验课常把“接近理论值”当作学习。可再入生成体则问：装置失灵、噪声增加或理论假设不再满足时，学生能否重新组织测量。

一个小组按说明书测重力加速度，数据漂亮，可能只是顺利执行。另一小组的光电门失效，学生辨认出真正需要的是时间——位置关系，转而用手机视频、标尺和帧率重建测量。若这种改变来自学生对约束的重建，而非教师给出替代方案，实验课才显露生成体。

物理学在此反过来限制艺术化教学：体验强度、开放材料和创意装置不足以替代可检验的响应改变。学生必须能在扰动后留下更稳定的测量选择，而不只是享受探索。

14.3 语文与历史：从“提出观点”到“重排证据责任”

人文学科容易把新颖观点当作生成。本章要求观点改变证据的组织。学生面对陌生文本时，是否会主动区分叙述者、作者、史料生产者和读者位置？是否会因为证据空白而改变问题，而不是用熟悉主题覆盖文本？

一堂课若形成可再入生成体，学生在新材料中会重新安排“谁需要证明什么”。他们可能改变阅读顺序、要求补充来源、把一句看似事实的陈述改写为待核查主张。这里的学习不是记住某个解释，而是证据责任的关系被重建。

艺术学在此提供关键约束：语调、节奏、图像和身体朗读不是装饰，它们可能构成判断入口。但心理学同时要求，入口在换文本后仍能被重新生成。

14.4 艺术课堂：从“完成作品”到“让媒介继续提问”

艺术课堂的假阳性不是只有模仿，也包括一次性灵感。作品漂亮、情绪真挚、展览成功，都不能单独证明课堂学习。

可再入生成体的证据在于：换媒介、换空间或换合作对象后，学生是否还能重建作品中的核心张力，并让新材料反过来改变问题。例如，学生在绘画中探索“遮蔽与显露”，之后转向声音装置时，不是复制视觉构图，而是重新发现静默、距离和听觉方向如何承担遮蔽关系。关系在新媒介中生成了不同形式，说明它不是风格模仿。

与此同时，物理约束要求作品过程留下可辨认的选择变化；心理约束要求学生能在延迟后重新进入，而非只在教师策展下完成。

14.5 项目式学习：从“共同完成”到“发起权换手”

项目式课堂经常产出高质量成果，却把核心认知锁在教师或少数学生手中。最终展示无法显示谁能够在下一轮发起关键改变。

可再入生成体要求“持位可换”。在连续几个项目中，应观察问题重述、工具选择、证据否决和角色重组的发起处是否发生合理换手。若每次都是同一个人救场，团队只是借用了个人能力。若新成员能依据共同形成的关系重新组织任务，课堂层面的学习才出现。

换手不是平均主义，也不是为了指标强行轮换。它指功能能够被接续，而非职位必须轮流。安全关键任务中，经验者仍可保留最终责任，但理由、约束和检测路径必须能被其他成员重新进入。

14.6 AI 课堂：从“答案增产”到“问题与验证结构再生”

生成式 AI 使课堂假阳性更隐蔽。学生可以迅速得到高质量文本、代码、图像和解释，外在成果远高于过去，却可能没有任何关系进入学生或课堂系统。

真正学习的检验不是禁止 AI，而是改变 AI 条件：更换模型、取消历史对话、提供冲突输出、隐藏常用提示模板。学生能否重新定义任务、构造证据、识别模型不确定性并调整人机分工？若能，工具成为生成体的可替换构件；若不能，课堂只是在租用外部能力。

艺术学提醒人们，AI 界面和输出形式会改变感知与创作；心理学要求学生主动提取与验证；物理学要求换工具后响应地形仍保留关键结构。三项约束在 AI 课堂中第一次变得同时不可回避。

十八、与最近既有理论的分界

“可再入生成体”与多种成熟理论相邻。若不能给出同一案例下的相反判断，它就只是改名。下面逐一给出可裁决分离线。

15.1 与提取学习的分界

提取学习关注从记忆中主动生成已学内容及其对保持、推断和迁移的促进。本章承认提取是撤架再入的重要机制，但不把成功提取等同于课堂生成体。

若学生能在延迟测验中准确回忆定义，却在新任务中不改变问题、表征或行动，提取学习可判定学习发生，本章判定尚未形成生成体。反之，若学生不能逐字回忆，却能借新材料重建关键关系并改变方案，本章判定形成了再入，狭义回忆指标可能判低。

15.2 与生成学习理论的分界

Fiorella 与 Mayer 的生成学习强调选择、组织并把新信息与先验知识整合，以支持新问题解决。它与本章最接近。

分离线在于对象归属与构件替换。生成学习通常把生成加工定位于学习者认知过程；本章要求课堂关系在成员、工具或媒介部分替换后仍能重组。若某一学生通过自我解释获得迁移，而班级的关键功能仍锁在教师 and 该学生身上，生成学习成立，课堂可再入生成体不成立。若新成员通过共享制品和判断规则接续发起权，本章所指对象出现。

15.3 与迁移学习的分界

迁移研究问已学知识能否用于新情境。本章把迁移视为必要读数，却不等同于生成体。

若学生把同一规则正确应用到表面不同任务，但问题结构、工具和角色均未改变，这是成功迁移，未必有生成。若新情境迫使学习者修改原关系的适用边界、创造新表征并重组行动，生成体完成了自我更新。观察到机械套用，迁移理论可判成功而本章判不足；观察到关系在再入中被改写时，本章比普通迁移多预测一个结构变化。

15.4 与适应性专长的分界

适应性专长区分效率与创新，强调在新问题中灵活调整知识。它主要描述个人或专业共同体的能力形态。

若一位学生能独立创新解决陌生问题，但其他成员、制品和判断规则没有获得接续功能，适应性专长成立，课堂生成体可能不存在。若核心学生离开后，班级仍能通过共同关系重组任务，本章成立而个人专长不足以解释。

15.5 与 ICAP 框架的分界

ICAP 框架按被动、主动、建构、互动活动预测学习收益，互动通常具有更高潜力。本章不以活动外形判定对象形成。

若课堂对话高度互动、学生彼此推理，却在提示链撤除后不能重建关系，ICAP 可预测较好学习，本章判为场域生成但未凝结。若一段安静的个人建模最终形成可被全班接续的再入结构，活动表面不一定处于最高互动等级，本章仍判课堂学习形成。

15.6 与生产性失败的分界

生产性失败强调在正式教学前让学习者尝试复杂问题，初始失败可为后续理解准备表征。

本章不把失败次序当作必要条件。若学生没有经历明显失败，却通过媒介转换和同伴证据形成可再入关系，本章判定学习发生；生产性失败不是必需。反之，学生经历丰富失败并在教师讲解后答对，但之后不能撤架再入，生产性失败流程完成，生成体未形成。

15.7 与转化学习的分界

转化学习强调意义视角通过反思和话语发生改变，常伴随身份或世界观重构。本章允许没有自觉叙述的学习。

若学生无法陈述“我的观点改变了”，但身体、工具和证据关系在新情境中稳定重组，本章判定形成学习；转化学习可能因缺少明确视角反思而不判定。若学生表达深刻身份转变，却没有任何延迟后的再入行动，转化经验成立，课堂生成体未必形成。

15.8 与具身认知、延展心智的分界

具身与延展认知主张认知由身体、环境、工具或社会制度共同构成。它们为本章提供了场域基础，却不自动给出学习发生的时间判据。

若一组人依靠特定装置和共同体完成复杂判断，延展认知成立；但装置或成员一换，能力消失，本章判定只有当下认知，没有可再入学习。若构件替换后关系仍能重新生成，本章所指对象出现。观察到“一撤环境即消失”时，延展认知解释当下能力，本章拒绝把它结算为课堂学习。

15.9 与物理学习的分界

物理学习以训练后的结构印记和扰动响应为核心。本章借用了它的严格性，却加入了重构和生成要求。

若系统在训练分布附近表现稳定，对扰动也呈现可辨认功能，但不能在结构同构的新问题中重建约束，物理学习成立，课堂生成体不成立。若关系能够通过新表征和新成员再生，即使没有单一固定物理结构承载，本章成立，纯结构印记不足以覆盖。

15.10 与“阈限概念”的分界

阈限概念理论强调某些概念一旦掌握会带来转化、整合，且常被描述为难以逆转。本章不要求学习必然不可逆，也不把概念作为唯一单位。

若学生跨过一个概念门槛却只在个人内部成立，阈限概念成立，课堂生成体未必成立。若一套非命题性的身体——工具——协作关系能被重新进入，即使参与者后来仍可能退化，本章成立而阈限概念不一定适用。

十九、与社会建构主义的分界

社会建构主义强调知识在互动、语言和文化实践中形成，与本章都拒绝把课堂看成单向传输。但是“共同建构过”仍不能证明形成了可再入对象。

若学生在教师精心组织的对话中共同达成理解，而换掉话语脚手架和成员后关系无法重建，社会建构过程真实发生，本章仍判定课堂生成体未形成。反之，若最初的互动并不热烈，却留下可被不同成员和媒介重新进入的关系，本章判定形成学习。分离线因此不是“是否社会互动”，而是互动产物在构件替换后能否继续生成自己。

二十、对课堂设计的重写：教师不是“传递者”，也不只是“促进者”

新判断不会导出单一教学法，却会改变设计目标。

第一，设计可读扰动。教学任务不能只让学生在原格式上成功，应预先安排表面改变、信息缺失、工具故障、角色置换或冲突证据，使响应地形可被读出。扰动不是增加难度，而是改变入口。

第二，保留预测所有权。教师需要让学生在反馈前形成可检验的预测。没有预测，错误只是被告知；有了预测，差异才能指向哪一段关系需要重构。反馈的时点应让学生拥有解释差异的机会，同时避免错误路径长期无约束固化。

第三，安排跨媒介凝结。同一关系至少在两种以上媒介中出现：语言与图像、公式与动作、数据与故事、声音与空间。目的不是丰富呈现，而是检查关系是否能脱离某一种表面载体。

第四，撤除提示链。课堂末尾的“总结”往往再次替学生完成组织。应设置延迟的无提示再入任务，改变问题表面和角色分工。撤架不是把学生推向孤立，而是让原有关系在新构件中重组。

第五，记录发起权。课堂观察不能只记谁回答正确，还要记谁首次重述问题、谁否决工具、谁引入反例、谁改变分工。连续任务中的发起权是否换手，是区分个人借力与课堂生成体的重要证据。

第六，允许关系改写原课程。若新情境揭示原概念边界，课程必须允许学生修改定义、增加例外或改变工具。只允许把课程结论带出去、不允许外部情境回写课程，所谓迁移仍是单向搬运。

二十一、教师工作的真正对象：管理关系的出生与离场

在这一框架中，教师既不是知识搬运者，也不只是提供自由的促进者。教师工作的对象是关系何时出生、由谁暂时持有、怎样脱离原入口、何时接受扰动、如何在不被教师重新接管的情况下完成再入。

首先，教师需要诊断“缺的是什么关系”，而不是只诊断“缺哪个知识点”。同一错误可能来自单位未对齐、证据责任不清、时间顺序混乱或感知不到材料阻力。补讲知识点只会在错误表面覆盖一层正确答案。

其次，教师要保护尚未被语言完全封闭的空白。学生刚形成模糊关系时，过早命名会把探索压缩成术语匹配；完全不命名又可能无法凝结。教师需要在可感、可做、可讨论和可命名之间安排节奏，让关系先以多种形式显露，再进入概念。

再次，教师要安排自己的离场。许多课堂把“教师总结”当作学习完成的标志，实际上总结常把学生正在形成的关系重新收回教师。更高水平的结束方式是留下未完成的异境任务，让下一次再入由学生发起。

最后，教师还要允许课堂反过来改变课程。学生在新情境中发现概念边界、工具缺陷或伦理后果时，课程不能只把它视为偏题。若课堂生成的差异永远不能回写课程，课堂仍是封闭执行系统。

二十二、评价：先确认对象，再计算“再入生成率”

本章的核心对象不是一个分数，但研究与实践需要操作化。可在不把对象还原为指标的前提下，计算“再入生成率”：在若干表面陌生、关系同构且原提示链被撤除的任务中，出现“关系重建+后续组织改变”的次数，占有效再入机会的比例。

一次有效机会必须满足四项记录条件：任务确实改变了表面入口；原提示和固定分工被撤除；关键关系可由独立编码者辨认；后续至少有一项问题、表征、工具、行动或协作改变。只记录最终正确率不够，因为生成可能产生有价值的新路径却暂时得到错误答案；只记录新颖性也不够，因为新颖可能与任务约束无关。

数据可来自多数学校已经存在的材料：学习平台的版本历史、草稿与批注、课堂白板照片、实验日志、小组角色记录、代码提交、教师反馈时点和延迟作业。研究者不必首先安装脑成像设备。2026年的真实课堂纵向超扫描研究显示，师生脑间耦合的动态增加可预

测之后的学业改善，但静态耦合与当下成绩的关系更复杂，可能受到教师选择性关注影响（Xu, Zhang, & Zhang, 2026）。这一结果恰好提醒：课堂关系指标必须看变化与时间顺序，不能把某一时点的“同步”当作学习本身。

再入生成率应与即时分数、延迟回忆、参与度和教师评分并列，而非取代它们。其独特价值在于识别“成绩最好看、再入最差”的反例。如果一个指标与现有成绩高度相关，它大概率只是成绩的代理；只有在高分但提示依赖严重的课堂中给出低读数，它才提供新信息。

伦理上必须写死三条。第一，读数指向课堂位置和关系，不用于给学生贴人格标签。第二，不得为了提高指标，在安全关键任务中强制角色轮换或制造故障。第三，任何依据该读数作出的不利安排，必须公开编码规则、原始证据与复核通道。

二十三、证据的三层结构：现象、机制与存在证明

课堂研究常把三种不同强度的证据混用。

第一层是现象证据：学生成绩提高、回忆保持、参与增加、作品更好。这些证据说明教学产生了效果，却不能说明效果通过何种机制发生，也不能确认生成体存在。

第二层是机制证据：预测误差、主动提取、表征转换、师生耦合或角色重组与后续结果之间存在特定时间关系。机制证据比现象证据更强，但仍可能只是生成体的某一局部过程。

第三层是存在证明：原构件被替换后，关键关系以新的形式重建，并引发后续组织变化。它要求同一对象在至少两个不同时间点和两个不同入口中显露；若只有一次观测，无法区分偶然聪明与可再入存在。

三层应当分开引用。即使“可再入生成体”这一核心概念被证伪，二维辨别格仍可作为区分稳定与生成的分析工具；即使二维格不够完整，“即时成绩不能替代撤架后的迁移”这一经验命题仍可独立成立。分层结构使整篇理论不会因最强命题倒塌而全部失效，也迫使研究者明确每一条数据到底支持到哪一步。

二十四、再入证据的编码协议

为了避免研究者看到任何新颖表现都宣布“生成”，编码应在观察前写明。

第一步，确定关系同构。研究团队需要先说明原任务与新任务共享的约束是什么，并由至少两名学科专家独立确认。只说“都是同一知识点”不够；必须指出哪些表面元素可以变化、哪一项关系若改变就不再同构。

第二步，记录提示链。把教师问题、板书位置、示范步骤、同伴角色、软件默认值和评分语言列成清单。再入任务至少撤除其中两类，并说明学生仍可使用哪些新资源。否则所谓“独立完成”可能只是隐藏提示继续工作。

第三步，编码关系重建的首个可见时点。证据可以是一句问题重述、一张新图、一次测量改变、一个反例或一项分工调整。编码者不能用最终正确答案倒推学生早已理解，而要沿时间序列判断何时首次出现。

第四步，编码后果。关系重建之后，至少一种后续组织必须改变，而且改变与关系之间有可说明的因果连接。学生随意换颜色、换软件或换组员不算生成。

第五步，检查持位。记录发起处与接续处。若每次重建都由同一人完成，结果只能支持个人层面；若另一成员在之后任务中利用公共制品或判断规则发起，才支持小组或课堂层面。

第六步，保留否定案例。研究报告应展示至少两段“表现很好但未形成再入”的材料，以及一段“结果暂时错误但关系已开始重建”的材料。没有否定案例的编码体系容易把理论写进数据。

通过这六步，评价从结果截图转向可审计的发生序列。

二十五、三个领域反过来给新判断加约束

一个跨领域判断若只从各学科取材料而不接受反向削弱，仍然只是借词。

物理学削弱了“只要能再入就算学习”的强主张。临场聪明、群体即兴和偶然联想都可能在新情境中产生差异，但若没有延迟后的稳定响应改变，它们不能构成生成体。本章因此加入扰动留痕与休眠后再入，而不把创造性事件本身当成果。

心理学削弱了“冲突越大、撤架越早越好”的设计冲动。预测误差的效果取决于原预测是否正确、误差强度和反馈时点；过强差异可能损害记忆。本章因此把可工作差异、时间间隔和分层撤架写入边界，不主张以持续挫败制造学习。

艺术学削弱了“所有学习最终都应由个人无工具测验确认”的价值排序。身体、媒介和认知制度在某些能力中是构件件，不是可随意去除的支架。本章所撤除的是原提示链，而非把学习者置于无环境状态。允许使用新工具和新同伴，是因为真正的再入发生在构件替换后的关系重组中。

这三处约束使本章放弃了更强但错误的表述：学习不是越独立越好，不是越意外越好，也不是越稳定越好。它要求稳定、差异和关系三者互相限制。

二十六、不适用的边界与可能的反噬

第一，基础自动化训练并非都需要生成体。字母识别、乘法事实、紧急操作和某些程序性技能首先需要可靠自动化。本章不是否定训练，而是反对把训练成功直接扩张为课堂理解。自动化可以是生成体的构件，却不是充分条件。

第二，并非所有课堂都追求课堂层面的持位转换。个别化矫正、私教和某些专业师徒训练可能以个人能力为目标。此时仍可检验个人再入，但不应强行要求形成班级对象。

第三，安全关键系统中的角色不能为了“换手率”随意轮换。飞行、医疗、实验室危险操作需要资格边界。持位可换指知识和判断路径可接续，不等于责任职位必须随机替换。

第四，文化与艺术中的不可复制经验具有自身价值。一次性事件未形成课堂生成体，并不等于它无价值。本章只回答“是否形成课堂学习对象”，不垄断教育价值判断。

第五，生成体可能传递错误。阴谋论社群、考试投机群体和偏见共同体同样可能形成高度可再入的关系结构。本章的存在判断不等于善的判断。课程还必须引入证据质量、伦理责任和公共可辩护性。

最大的反噬来自指标化。一旦“再入生成率”进入考核，最省事的做法可能是在文书中制造表面换手、预先排练所谓陌生任务、把提示隐藏得更隐蔽，或把学生的新颖发言编码为生成。指标越正式，造假动机越强。本章看不到一个彻底消除这一反噬的出口，只能要求多源证据、公开编码、外部复核，并禁止把单一比例用于高风险决策。

二十七、AI 时代为何使这个问题变得紧迫

在生成式 AI 普及之前，学生作品质量与学生能力虽不等同，二者仍有较强联系。AI 把这种联系迅速拉开：高质量文字、代码、图像和方案可以在极少内部重构的情况下出现。于是，传统课堂最依赖的成果证据开始系统失真。

AI 并不只是增加一个工具，它改变课堂响应地形。提示词、上下文窗口、模型偏好、自动补全和界面默认值共同承担了解决问题、表征选择与语言组织。学生可能获得前所未有的输出，却不知道哪部分关系由模型持有，哪部分由自己或课堂持有。

“可再生生成体”因此提供一个比“是否使用 AI”更有辨别力的问题。更换模型、删除历史对话、给出彼此冲突的回答、限制熟悉模板之后，课堂能否重新生成任务标准、证据链和人机分工？能，则 AI 成为可替换构件；不能，则学习对象仍寄居在平台中。

这也改变教师评价 AI 作业的方式。检测文本是否由 AI 生成越来越不可靠，也不是核心。核心是设计再入任务：要求学生解释为何改变问题、如何验证证据、在模型失效时怎样重组工具。评价从追查“谁写了这段文字”转向辨认“关系在构件替换后是否仍能工作”。

更深的风险是，AI 会把教师也变成提示链的一部分。教师依赖模型生成教案、问题和反馈，课堂表面更加丰富，却可能丧失自身重新组织课程的能力。本章的判据同样适用于教师：当常用模型不可用时，教师能否重建教学关系并让课程继续生成差异？

二十八、对学校评价制度的后果

若学校接受这一判断，评价单位需要发生三项变化。

第一，从“课后立刻结算”转向“保留再入窗口”。课程结束后应安排延迟、变式和构件替换任务，允许学习对象在休眠后显露。没有再入窗口，学校只能评价当堂运行质量，不能评价课堂学习是否形成。

第二，从“成果归个人”转向“区分个人、团队和制度三本账”。个人考试仍然必要，但共同生成的关系不能全部拆成个人分数。学校需要记录哪些能力由公共制品、工作传统和跨届协作承载，同时防止团队成果掩盖个人空缺。

第三，从“标准化消除差异”转向“用受控差异检验结构”。标准化适合比较同一输入下的输出，却不适合检验入口变化后的重组。评价体系应保留一部分结构同构、表面不固定的任务，并公开同构原则而非具体题型。

这三项变化不会自动提高教育质量。它们也可能产生新的文书负担和形式主义。因此，再入证据只能作为抽样诊断，不应把每节课都变成复杂项目，更不应取消基础知识测验。制度改革的目标不是增加一种更时髦的评分，而是承认现有账本看不见某类课堂成果。

二十九、可证伪的研究设计

最强竞争解释不是“传统教学更好”，而是一个更朴素的说法：所谓可再入生成体，只是先验能力、动机、教师质量或已有知识的另一种名称。要使本章成为理论而非修辞，必须在控制这些变量后仍出现独有预测。

核心证伪条件可写为：控制入学先验成绩、一般认知能力、动机、教师效应、任务熟悉度与课堂总投入后，若“原提示链撤除后的关键关系重建”不能额外预测延迟迁移，或其出现不早于问题、表征、工具、行动或协作的改变，则本章机制为伪。届时，课堂学习应重新归因于个体能力或练习量，而非生成体。

研究一可在同一学校选择不少于十二个同年级班级，围绕同一概念设计三次表面不同的任务。第一次正常教学，第二次在两周后撤掉例题格式和固定分工，第三次在六周后更换工具或媒介。编码关系重建出现的时点，并检验它是否先于后续组织改变。若只是结果相关而无时间先后，不足以支持机制。

研究二利用现有数字痕迹。对同一课程的项目版本历史进行序列分析：首次问题重述、表征更换、工具否决、角色重组和性能改善分别何时发生。本章预测，真正形成生成体的项目中，关系重建先于至少一项组织改变，且发起处在后续任务中有更高概率换手。

研究三比较即时高分但再入读数不同的课堂。在控制分数、教师、教材和练习量后，若低再入组与高再入组在表面陌生任务上的表现没有差异，本章失去独立解释力。

研究四检验物理约束：若一次性创造事件在没有任何延迟响应变化时仍稳定预测长期迁移，则“扰动留痕”不是必要条件，本章需缩减定义。

研究五检验艺术约束：若将身体、媒介与共同体全部移除的个人纸笔测验能够完整预测所有新工具和新协作条件下的再入，分布式构成主张被削弱。

研究六是最低成本版本。教师只需保留两类任务：原格式延迟测验与关系同构、表面陌生、撤提示任务，并记录学生改变下一步的证据。多数机构现有作业和平台日志已足以开展初步检验。

任何以两个国家、两个地区的异质比较替代同质多班级研究，都不能构成机制证伪。文化、课程、语言和制度差异太多， $N=2$ 只能提供启发。

三十、六条可直接观察的顺序预测

为了让理论进入课堂研究，本章给出六条顺序预测。

其一，在形成生成体的课堂中，学生对关键关系的首次重述应早于最终成绩改善，而不是在教师公布答案后才出现。

其二，表征、工具或分工的改变应发生在关系重建之后；若改变先发生且与关系无关，它只是探索噪声。

其三，延迟任务中，重新生成关系的学生不一定最先给出正确答案，但更可能在后续修正中到达可辩护结果。

其四，课堂层面的生成体形成后，关键发起处在连续任务中会出现有限换手；完全不换手提示角色锁定，随机频繁换手则可能说明组织尚未凝结。

其五，换媒介时，表面策略的相似度可能下降，但关键约束的可辨认度应保持。若只保留相似形式而约束消失，属于风格复制。

其六，教师提示减少造成的短期成绩下降，可能先于延迟迁移上升。若撤架后即时和延迟结果都下降，且没有关系重建证据，不能把失败浪漫化为“深度学习正在发生”。

六条预测共同要求观察轮次，而不是只看相关。课堂研究应记录这一轮谁先动、其他构件多久跟上、关系改变后怎样回写下一轮入口。

三十一、本章自身是不是它所描述的假阳性

本章全部内容都以可陈述文字呈现，因而天然落在“会说”一侧。读者可以记住“可再入生成体”这个名称，复述三项条件，甚至用它写出一篇评论，却仍未形成本章所主张的学习。

用本章自己的判据检验本章：撤掉本章中的课堂案例和术语，给读者一个陌生组织——例如医院交班、软件故障响应或社区艺术项目——读者能否重新发现“稳定痕迹、撤架再入、异境生差、持位可换”之间的关系，并据此改变分析问题、证据与责任分配？若不能，本章只是制造了一个流畅的新名词。

更严格地说，文章本身无法提供身体、材料和共同体的完整构成条件。它最多制造一个可能入口。真正的存在证明必须发生在读者之后的实践中。因此，本章不能被阅读、被引用或被出版作为理论成功的证据。

这也意味着“可再入生成体”可能被自己的传播方式杀死。一旦被课程标准压缩成四条清单，教师会训练学生在检查表上表演再入，关系重新变成格式依赖。本章没有一个保证性的出口，只能把反例、边界和证伪写进概念，使它持续暴露于构件替换和新情境。

三十二、写死日期的赌注

本章作出如下可失败的经验赌注：到2030年12月31日，至少两项预注册、彼此独立、每项包含不少于十二个真实班级的研究将发现，在控制即时测验成绩、延迟回忆、参与度、先验能力、动机和教师效应后，“提示链撤除后的关系重建”仍能显著预测表面陌生任务中的迁移；并且在时间序列上，关系重建早于至少一种问题、表征、工具、行动或协作方式的改变。

以下情况不算命中：只报告相关而没有时间顺序；只由学校自行宣称采用了新评价；只在实验室个人任务中验证；只比较两个高度异质国家或地区；只发现高分学生再入更好而未控制先验能力；只把新颖回答当作生成。

若到期没有达到上述条件，或高质量研究持续显示关系重建不能提供独立预测，本章的承重判断应被撤回：课堂学习仍可由个人记忆、迁移能力或情境参与的组合充分解释，无需引入新的关系对象。

三十三、结论：课堂留下的不是答案，而是重新开路的存在

真正的课堂学习不是把知识从教师移动到学生，不是把一次正确表现延长到若干周，也不是让课堂变得更热烈、更开放或更有美感。物理学迫使我们寻找扰动之后仍存在的结

构差异；心理学迫使我们检验预测、提取、时间与迁移路径；艺术学迫使我们承认身体、媒介、他人与制度可能构成认知本身。

三类理论合在一起，不是得到三个指标，而是取消一个旧前提：学习不必先被某个个体或某个位置占有。课堂能够生成一种分布式关系对象。它在当堂活动中发生，在跨媒介组织中凝结，在无使用时休眠，在新情境中被重新进入。它的持有人可以变化，构件可以替换，入口可以改变，但关键关系能够重新生成，并继续改写问题、工具、行动和共同体。

因此，何谓真正的课堂学习？答案是：课堂中形成了一个可再入生成体。最简洁的检验不是问学生“懂了吗”，也不是只看分数，而是撤掉原来的例题、提示和分工，换一个表面陌生但关系同构的任务，看学习者能否重建关键关系，并让下一步发生改变。

一堂课的最高成果，不是所有人带走同一个答案，而是课堂离开之后，世界换了入口，他们仍能重新开出路。

这一路并不保证永远正确，却保证错误能够被新的证据重新组织；它不保证每个人拥有相同能力，却让关键功能不再被永久锁死；它不保证课堂经验可以原样复制，却使关系能在差异中再次出生。课堂学习由此从“内容搬运”转向“存在发生”：教育不再只计算带走了多少，而开始辨认什么在离开之后仍能回来，并且每次回来都使世界与学习者同时发生一点改变。

接续。本章的四项构成条件是全书统一对象名的来处，理由见第一编小结第四节；其中**持位可换**在第六章以物的形态重新出现——一件可交接的自建物，就是“发起权不锁在原持有人手里”的物质证据。持位可换度的采集见附录一，本章那句判据被收进第十二章问卷的第一块。

第一编小结 三个名字，一个对象

一、三章各自到达的那句话

第一章：课堂真正留下的，是这个班**重新发起一次学习**的能力；它可以拆成问题、修订链与判断场三项，记作 $\mathcal{Q} = \langle Q, V, J \rangle$ ，而再生水平取三项中的**最低项**而非平均项， $R = \min\{q, v, j\} / 3$ 。

第二章：学习**不在任何一个人身上结清**；它必须被交到下一个人手里、被那个人改写并继续用下去，才算存在。三家共同默认而没有说出的前提是“学习有主人，也有结算时点”，而这两样东西都是账本的假定，不是事情的性质。

第三章：课堂生成的那样东西是一个**可再入生成体**——它有四项构成条件（扰动留痕、撤架再入、异境生差、持位可换），它**不持续显露**，休眠不等于死亡，而传统测验只查旧答案能否被直接提取，因此系统性地把休眠读成遗忘。

二、把它们压到五十字：一个必须正面说的麻烦

如果把三句话各自压到五十字以内，会得到三句几乎可以互相替换的话：**课堂留下的是一件无人持有、支架撤除后仍可被重新进入、并且要靠继续被使用才继续存在的东西。**

这个麻烦必须由我自己先说出来，而不是等读者发现。一本书里不能有三个名字并列指着同一样东西——那不是丰富，那是没写完。

三、两种读法，与判决它们的办法

三章的重合可以有两种读法，而且它们的后果完全相反。

读法甲：这是异源汇聚提供的稳健性证据。三组学科完全不同，取材互不重叠——第一章取的是教育学的迁移与延迟显露、工程学的目标与标准重组、艺术学的共同判断场；第二章取的是认知科学的能力变化、经济学的稀缺资源配置、艺术学的制作中的思考；第三章取的是物理学的响应地形、心理学的预测—提取—迁移路径、艺术学的身体与媒介。九个取用点里，只有“艺术学”这个名字重复出现三次，而三次取的是三样不同的东西（共同判断场／制作中的思考／身体与媒介场）。三条路取材互不重叠而到达同一处，这在方法上叫**异源汇聚**，是一个判断能得到的最好的一种支持。

读法乙：这是同一件事被说了三遍。三章的作者是同一个人，同一段时间里写的，用的是同一套工序。哪怕材料不同，写作者的手是同一只手，很可能在三个不同的入口处各放了一面镜子，照出来的当然是同一张脸。

两种读法都讲得通，而**它们可以被同一个观察判决：**

在同一批课上同时采第一章、第二章、第三章的读数（再生水平 R、接力率、持位可换度），看三者的相关结构。

若三者两两相关系数都高于零点八，读法乙成立——三章说的是同一件事，本书应当只保留一章，另两章降为附录。

若三者相关适中（大约零点三到零点六）且各自与外部效标（延迟迁移、撤架后独立表现）的关联模式不同，读法甲成立——它们是同一对象的三个面，各自捕捉到一部分。

若三者相关接近零，那么它们不是同一样东西，本书的合并是错的，应当拆成三个对象各自成立。

我把话说在前面：**第三种结果对本书最不利，因为它意味着本编的合并是一次修辞而不是一次判断。**而这个观察不贵，用附录一的手册在一门课上跑一次即可。

四、本书的决定：一个对象，三条路径，三个读数

在上述观察做出来之前，本书必须先作一个决定。决定是：**合并**，并且承担合并的风险。

统一对象名取“可再入生成体”。理由有三条，都不是文辞上的。

其一，它是三个名字里最不像教育学的一个。“学习事件再生”与“接力”都可以被现有的教育话语顺畅吸收——前者会被读成“探究式学习的又一种说法”，后者会被读成“合作学习”。而“可再入”这个词在教育里没有位置，它逼着人先问一句“什么叫再入”，再往下读。一个名字最重要的功能是**挡住误读**，不是让人一见如故。

其二，它对“不显露”有**明确处理**。第三章把休眠与死亡分开：休眠对象在新线索出现时仍能被重新组织，死亡对象只能靠重教恢复。这是三章里唯一一处正面处理“看不见的时候它在不在”的地方，而全书的母题恰恰立在“看不见的部分被算作零”上。名字必须承得起母题。

其三，它是三章里唯一给出了完整构成条件并逐条配了排除项的（扰动留痕排除格式熟练、撤架再入排除借来的支架、异境生差排除机械套用、持位可换排除尖子生救场）。一个对象名要能带着它的证伪条件一起走。

另两个名字不废，降格为这个对象的两个侧面，各自承担它单独承担得起的那部分：

接力，是它如何延续的机制。可再入生成体不会自己维持自己；它要么被交出去并被改写，要么在原持有人手里凝固成一件私人技能。第二章的六个阶段与顺序预测，是这个对象的**传递学**。

事件再生，是它显露的形态。这个对象平时不显露，只有在一次新的学习被重新发起时才露一次面； $\mathcal{L} = \langle Q, V, J \rangle$ 与最低项取法，是这个对象的**观测学**。

于是三章不再是三个并列的主张，而是同一样东西的**构成（第三章）—传递（第二章）—显露（第一章）**。本编的次序是由近及远，而它的逻辑次序恰好是倒过来的；读者若愿意重读一遍，从第三章往回读会更顺。

三个读数全部保留，因为它们测的是这个对象的三个不同侧面：再生水平 $R = \min\{q, v, j\}/3$ 测显露；接力率测传递；持位可换度测构成。第十一章把它们与第二、三编的读数放在同一张表上，并处理量纲与代理风险。

五、合并后的定义

课堂学习，是一个班在一段教学里生成的一件可再入生成体：它满足扰动留痕、撤架再入、异境生差、持位可换四项条件；它不持续显露，只在一次新的学习被重新发起时露面；它无人持有，因而不能在任何一个时点对任何一个人结清；它靠被交出去、被改写、被继续使用而继续存在，一旦不再被交出去，它不是被遗忘，是停止存在。

这个定义里有一句话是全书其余部分的地基：它无人持有，因而不能在任何一个时点对任何一个人结清。账本恰恰是按“持有人+结算时点”造出来的。定义与账本在这里正面顶上，而这就是第二编与第三编要走的全部路。

六、一次删除测试

本书对每一个命名都做同一个自查：删掉本书的测量办法之后，被说的那样东西是否仍然以同样的方式存在？若答案是“仍然存在，只是难以测量”，那么发明的只是一个新指标。

可再入生成体过得去这一关，而且过得比较硬：传统账本没有为它设置存在位置。成绩册的每一栏都要求一个人名与一个日期；而这个对象既没有人名（持位可换）也没有日期（不持续显露）。删掉本书的三个读数，它仍然在那儿——只是没有任何一栏可以填它。这正是“看不见的部分被算作零”的原始现场。

七、交给第二编与第三编的东西

交给第二编（How）：本编说这样东西是被“生成”的，但没有说它由什么生成。第二编三章各给一个位置的回答——**终点、次序与条件里**，有一部分是这个过程自己临时造出来的。特别是第六章的自建条件，与本编的“持位可换”是同一件事在物的一侧的形态：一件班级自己造出来、可以交给新人直接用的工具，正是“发起权不锁在原持有人手里”的物质证据。

交给第三编（Why）：本编说这样东西可以不显露，而第三编三章各给一处它被熄灭的位置——入口（窗开着而无人到达）、伴生物（悬置期结成而无人拆除的连接）、出口（还在动的时候就被读了数）。三处都不需要任何人犯错，都在正常运行中发生。

留给第四编的一笔账：本编的三个读数与第二、三编的六个读数之间是什么关系，本编说不出来。第十一章处理。

八、回到母题

本编给母题提供的是它的主语。

母题说“课堂里决定后果的那些东西，有一大部分是课堂自己造出来的”——本编回答的是：被造出来的那样东西，是一件可再入生成体。母题说“账本只为不是它造的那些东西设了字段”——本编回答的是：这个对象在原理上填不进任何一栏，因为它既没有持有人，也没有结算时点，而账本的每一栏都由这两样东西定义。

因此账本不是漏记了它，是没有为它这一类东西留下位置。这两件事的差别，是全书与“把评价做得更细”那一整条路线的分界线。

九、本编的可推翻条件

第一条。若第一章、第二章、第三章的三个读数在同一批课上两两相关系数持续高于零点八，本编的合并是多余的——三章应压为一章。

第二条。若三个读数两两相关接近零，本编的合并是错的——它们不是一个对象的三个面，本书应当拆成三个对象，而全书的架构随之改变。

第三条。若在一门课上，可再入生成体的四项条件全部成立，而学生在撤架后的独立表现与四项条件全不成立的班级无差异，那么这个对象即使存在也不承重，本书没有理由为它写一编。

第四条，也是最需要外人来做的一条。本编把“三组不同学科取材互不重叠而到达同一处”读作异源汇聚证据。若有人能证明这三条路径实际上共用了同一个隐蔽前提——例如三章的艺术学取用点其实是同一支理论的三种表述——那么异源汇聚就是一次伪独立，本编最强的那条支持随之作废。这一条我自己检查过三遍，仍然可能看不见它；能看见它的人只能是从外面来的人。

第二编 How：发生的条件是自造的

本编要回答的是：一节课该怎么运行，学习才可能发生？

这个问题要的是一条次序——从哪儿开始、经过什么、在哪一步可以插手——不是一个定义。三章分别处理次序的三个部件：终点、路径与条件。三章各带一条结构模型。

第一编交出了一件东西：课堂里被生成的那样东西是一件可再入生成体，它无人持有，也没有结算时点。但第一编没有说它由什么生成。本编三章各从一个位置回答，而三章的回答合起来是同一句话的三种形态：**课堂运行里被当作给定的那些东西，有一部分是这个过程自己临时造出来的。**

第四章处理终点。教育学最强的一条主张是先把要到达的表现写清楚，再倒推出评估与活动；化学工艺开发的最强主张正好相反——不要先锁定产物，因为产物形态是条件的函数，先定产物等于先把条件锁死；艺术学的工作室传统给第三条——终点既不是表现也不是条件，是学习者自己那套能继续组织下去的做法。三家不能同真。推翻共有前提的材料来自化学自己：**逆合成分析与条件筛选在化学里并存，而选用判据是“信息在哪一端多就从哪一端开始”**——终点与手段的角色不由事情的性质决定，由信息的分布决定，因而可以互换。由此命名**换位区间**。

第五章处理路径。物理学说路径由两端整体决定，心理学说人只能一步一步走且看不见自己的下一步，工程学说最优点就是最脆的点。推翻材料来自物理学自己：**初值问题在很弱的条件下解就存在且唯一，而两点边值问题可以无解、也可以有无穷多解——倒推与推进的适定性根本不同。**这一条在物理里只回答“该用打靶法还是直接积分”。带回课堂便得到**可倒推区间**与它的三档：唯一、多解、无解。本章的判据是全书最便宜的一个：**把学生此刻的状态与目标交给五位有经验的教师，看“下一步该做什么”是同一件事、五件不同的事、还是都说不出来。**一次教研活动可测。

第六章处理条件。艺术学说只有做出来的那个东西算数，工程学说质量是造出来的不是检出来的，细胞学说命运不写在细胞里、写在它周围。推翻材料来自细胞学自己：**生态位有相当一部分是住户自己分泌、招募、组织出来的——条件是被住在里面的东西造出来的。**由此命名**自建条件**，并给出全书最容易清点的一句判据：**这一节课结束时，这个班手里多了什么，是下一节课可以直接拿来用的？**

三章各推一条形式结论，且都推出了正文本身没有说过的东西：第四章推出最优换位时长与目标端的不确定性无关、而交还失败率与转化效率相乘；第五章推出错配代价在两个方向上不对称，因而在长期迁移与可选择性权重较高的内容上，**档位不确定时应当默认推进**；第六章推出稳态在 α 、 N 、 δ 不随覆盖度变化这一条件下，存量与外来条件覆盖度线性反相关；放开这一条件后，高覆盖度区间的压制仍然必然，中低区间则可能上升、下降或呈倒U形，因而一门“完备”的课程在结构上不可能有自建条件。这三条是全书结构性最强的一段，也是第十三章要处理的对象。

三章还共同交出了一个洞：三条结论显然是同一个更一般的东西在三个位置上的显形，而三章谁也没把这个更一般的东西写实。**本书正是从那个洞开始的**——第一编去回答“那个东西是什么”，第四编去回答“那么账该怎么记”。

最后交代本编与第三编的关系。第四、五、六章与第八、九、七章两两共用同一组学科：教育学·化学·艺术学被问了两次，物理学·心理学·工程学被问了两次，细胞学·工程学·艺术学被问了两次。两次取用的是同一门学科里互不引用的两个分支，问出来的对象完全不同。这件事本身携带信息，第十章处理。

第四章 手段取得终点资格的那一段

判断卡片：终点与手段的角色不由事情的性质决定，由信息的分布决定，因而可以互换；而换位若不交还，就不是换位，是目标被吞掉。

三个理论入口：教育学 × 化学 × 艺术学

提要。课堂如何运行才能让学习真正发生？三门学科各给一条完整的次序，且三条互不相容。教育学的最强主张是先把要到达的表现写清楚，再倒推出评估与活动，路径的终点是可观察的表现，条件与次序都是手段。化学工艺开发的最强主张相反：不要先锁定产物，先把条件空间摸出来，因为产物形态本身是条件的函数，先定产物等于先把条件锁死。艺术学的工作室传统给第三条：终点既不是表现也不是条件，是学习者自己那套能继续组织下去的做法，作品只是它的副产品。三条不能同真——教育学每天在做的那件事，正是化学工艺诊断为病的那件事。三家共同站着而从未说出的那块地是：**一堂课的运行方式，可以由“要到达什么”单独决定；终点一旦确定，中间的次序就只是手段问题。**推翻它的材料来自化学自己。化学内部有两套完整而互相取代的做法：逆合成分析从目标分子倒推路线，条件筛选从条件空间正推产物；化学没有裁定谁对，而是给了一条判据——**信息在哪一端多，就从哪一端开始。**也就是说，终点与手段的角色不由事情的性质决定，由信息的分布决定，因而是可以互换的。由此本章命名**换位区间**：一节课里的一段区间，在其中原本作为手段的那一样取得了终点资格，而原本的终点退为手段；它有真实的开合时刻与可查的触发源，而三家的结算规则都只承认终点态，因此都不设这个字段。判据一句：**这一节课里，有没有哪一刻，原本用来达成目标的那样东西本身成了要达成的东西——是哪一刻，由什么触发的？**配套读数换位率与换位触发源在三个领域量纲不同而比率相同，且与既有指标正交：一门目标写得最清楚、活动与目标对齐得最好、评价与目标一一对应的课，换位率恰恰是零。

一、同一个问题，三条不能并存的次序

“课堂如何运行才能让学习真正发生”是一个要次序的问题，不是一个要定义的问题。它要的答案形状是：从哪儿开始，经过什么，实现什么，以及在哪一步可以插手。

三门学科都能给出一条完整的次序，而且都给得出证据。麻烦在于三条次序的第一步就分开了，而且分开之后再合不回来。

教育学说：先写清楚学生学完之后应当能做什么，再倒推出用什么证据判断他能做到，最后才安排活动。这条次序在教学设计里已经制度化，写进了教案模板、课程大纲与教师培训。

化学工艺开发说：不要先锁定你要什么。同一批起始物在不同条件下会给出不同的产物，产物形态是条件的函数；你先把目标写死，等于在还不知道条件空间长什么样的时候就把大部分可能性排除了。正确的次序是先把条件摸出来，看它能给什么，再决定要哪一个。

艺术学的工作室传统说：两家都把终点放错了地方。作品不是终点，条件也不是终点；终点是学习者手里那套能自己继续下去的做法。一件作品完成得再好，如果做的人下一次仍然要人从头带一遍，这一次就没有留下东西。

三条次序不能并存，而且它们的冲突不是“侧重点不同”。教育学每天在做的那件事——先定终点再倒推——正是化学工艺诊断为病的那件事；而艺术学诊断出的病更狠：先定终点这个动作本身，会把“组织下去的做法”外包给定终点的那个人，因为终点定好了，路怎么走就是照做。

二、教育学这一家：终点是可观察的表现

把第一家推到最强，它的主张是：课堂运行的每一个环节都应当从可观察的目标表现倒推出来。

这不是一个空洞的口号，它有非常具体的工序。先确定学生学完之后应当能在什么条件下做出什么；再确定什么样的证据可以判断他做到了；最后才设计活动与讲授。这个次序的力量在于它切断了一种常见的失败——教师安排了大量活动，学生也确实很忙，而这些活动与要达成的东西之间没有可追溯的关系。倒推使每一个活动都必须交代自己是为哪一条目标服务的。

掌握学习的传统给它加了第二层：目标不但要写清楚，还要分解成可以逐级过关的序列，每一级都设检查点，不过关不往下走。

这一家的证据很硬。目标明确、评估与目标对齐、活动与评估对齐的课程，学生的成绩确实更好，而且这个效应在低先验知识的学习者身上最强。

这一家自己承认的那处麻烦，出现在目标不容易写成可观察表现的时候。“批判性思维”“创造力”“专业判断”这类目标，一旦被写成可观察的行为清单，写出来的东西往往已经不是原来要的那个东西了；而不写成可观察的行为清单，整条倒推次序就无处落脚。这一家的处理方式是继续细化行为描述，或者承认这类目标只能间接评估——两条路都是在原来的次序里找补，没有动那条次序本身。

三、化学这一家：终点是把条件建立起来

把第二家推到最强，它的主张是：要让一个转化发生，真正要做的工作是把条件场调到对的地方；产物是条件的函数，不是可以先行指定的独立变量。

工艺开发的日常正是这样运行的。同一组底物，换溶剂、换温度、换配体、换碱，得到的可以不是同一个产物的不同产率，而是不同的产物。因此工业上大量的开发工作不是“我要这个分子，去找一条路线”，而是“把条件空间铺开跑一遍，看它给什么”。高通量实验把这条次序推到了极端：一次实验里并行跑上千种条件组合，用极小的物料量，先获得条件空间的地图，再决定往哪儿走。

这一家的次序因此是：从条件出发，经过实际得到的产物分布，实现一个可靠的工艺。目标不是起点，是产出物。

它对教育学那条次序的诊断很直接：先定终点倒推，前提是你已经知道从这里到那里存在一条可走的路，并且知道路上每一步大致长什么样。这个前提在成熟工艺里成立，在

探索性工作里不成立；而在不成立的时候硬用倒推，得到的是一条纸面上完美、实际上到处断裂的路线。

四、艺术学这一家：终点是学习者自己的组织方式

把第三家推到最强，它的主张是：工作室里真正在生产的东西不是作品，是做作品的人身上那套能自己继续下去的做法。

工作室教学的结构支持这一点。学生在同一个空间里长时间工作，教师不断走动、短促介入；定期的集体讲评不评分，而是把在场所有人的作品摆在一起看；一件作品完成之后不是结束，是下一件的材料。对这套教学做的系统观察指出，工作室里被反复训练的东西是一组可迁移的做法——观察、想象、坚持、反省、表达、在做的过程中不断评估自己刚才那一步。

这一家对另外两家的诊断是同一条：只要终点被别人定好，“自己组织下去”这件事就没有位置。学生按逆向设计的次序走完一门课，他能做到目标里写的所有事；问他下一次自己该往哪儿走，他没有答案，因为这个问题从来不属于他。而化学那条次序看起来把探索还给了学习者，实际上把探索交给了条件空间——学生仍然是在读一张别人铺好的地图。

五、三家共同站着的那块地

三家争的是**终点该放在哪儿**：表现、条件，还是学习者的做法。

争得再凶，三家都默认了同一件事：**终点是一个可以被指定的位置，指定好了，其余两样就是手段**。教育学指定表现，条件与做法为手段；化学指定条件，产物与路线为手段；艺术学指定做法，作品与环境为手段。三家的争论完全建立在这个假定上——如果终点不是一个可以被指定的位置，那么“终点该放在哪儿”这个问题本身就没有意义。

把这个前提念给三家听，三家都会说这还用说吗。一堂课当然要有个目标，目标当然决定其余安排。这正是共有前提的签名。

六、推翻它的材料，来自化学自己

化学内部并存着两套完整而互相取代的做法，而且它们对同一个问题给出相反的次序。

一套是逆合成分析：从目标分子出发，反向拆解成越来越简单的前体，直到落在市售起始物上。这套做法的整个逻辑就是“先定终点再倒推”，它是二十世纪合成化学最重要的方法论成果，也是教育学那条次序在化学里的精确同构。

另一套是条件驱动的发现：不先定目标，把条件空间跑开，看它给出什么，再从给出的东西里挑。

关键不在于化学同时有这两套——那只是并存。关键在于**化学自己给出了选用哪一套的判据，而这条判据既不是“哪种更科学”，也不是“事情的性质决定”**。判据是：**信息在哪一端多，就从哪一端开始**。目标分子已知、断键策略成熟、每一步的可靠性有文献支撑——这时目标端的信息远多于条件端，倒推最省。反过来，起始物容易得到而产物空间没人走过、每一步的选择性无法预判——这时条件端的信息多，正推最省。

这就把共有前提拆了。终点不是一个由事情本身决定的位置，是一个由信息分布决定的角色；而信息分布会变，因此角色可以互换。一个昨天还是手段的东西，今天可以成为终点；而这个互换不是失误，是正确操作。

化学在自己的领域里用这条判据回答的是“该用哪种研发策略”，是一个方法选择问题。没有人把它当作“终点与手段的角色是可互换的”这一命题的证据——而它恰恰是。

七、命名：换位区间

由此本章命名一类区间。

换位区间：一节课（或一个学习单元）里的一段区间，在其中原本作为手段的那一样取得了终点资格，而原本的终点退为手段；该区间有可查的开合时刻与可指认的触发源。

四个条件，缺一不可。

角色真的换了，不是同时都重要。“既要掌握知识也要培养能力”不是换位，那是两个并列目标。换位的签名是：在这段区间里，原来的终点变成了检验新终点的材料。学生不再是“为了做出这个作品而练手法”，而是“用这个作品检验自己的手法”——作品还在，但它已经不是要到达的东西了。

有开合时刻。换位是一段区间不是一个态，它有开始与结束。结束的方式有两种：交还（原终点重新取得终点资格，而它已经被这一段改写过），或者不交还。

有触发源。换位不是排出来的，是被什么东西触发的。可指认的触发源包括：一个异常结果、一次讲评里出现的对照、一个学生自己发现“我这个做法在别的题上也用得上”。教师宣布“现在我们来关注方法本身”不构成触发——那只是把新终点也指定了一遍。

在三家的账上都不存在。教育学的教案版本只记目标与活动；化学工艺的报告只记条件与产物；工作室没有记录制度。这一段区间在三本账上都不是一样东西，不是记得少，是不设字段。

需要立刻说清一件事：换位不是“目标不清晰”。目标不清晰是从头到尾没有终点；换位是终点明确地易了一次主，且易主的时刻可以指出来。这两件事在读数上完全可分——前者的换位率也是零。

八、判据：一句可以拿去问记录的话

这一节课里，有没有哪一刻，原本用来达成目标的那样东西本身成了要达成的东西？是哪一刻，由什么触发的？

这句话不含任何情态词，不问课上得好不好，只问有没有、是哪一刻、由什么触发。它可以拿去问课堂录像、教案版本、讲评记录与学生的作业稿，不必问任何人的判断。

五个场景跑一遍，答案互不相同：

讲授加习题课。没有。终点从第一分钟到最后一分钟都是“会做这类题”，方法始终是达成它的手段。换位率零。

实验课。有，但通常被压掉。当数据出现异常时，“操作本身对不对”从手段升为要弄清的东西——这一刻就是换位的开口。多数实验课在这里把它关掉：异常被记为操作误差，重做一遍，终点交还得太快，区间长度接近零。

设计与写作工作坊。有，且频繁。学生做到一半发现自己反复用同一种解法，于是这一稿从“要做好的东西”变成“用来看清我那套做法”的材料。这一次换位往往不交还——原来那一稿被放弃了，而它的作用已经完成。

工业工艺开发培训。有，且被写进规程。条件筛选跑完之后修改目标产物，这在开发流程里是正常步骤，有对应的评审节点与文档模板。这一条是查出来的，不是从命题推出来的——化学的工程规程里存在教育的教案模板里没有的那个字段。

代码课与代码评审。有，且留在日志里。当“让测试通过”从终点变成“检验我这个写法站不站得住”的手段时，换位发生了；它在版本历史里表现为一次在测试已经通过之后的重写。这一条同样是查出来的：这种重写在提交记录里可以直接筛出来，而在任何教学评价表里都没有位置。

后两个答案不能从命题推出来。它们说明这个读数在两个与课堂无关的行业里已经有制度化的形式，只是没人拿它来看课堂。

九、两个可以清点的读数

换位率。一个概念单元里，终点易主的次数除以该单元的教学机会数。计数要求两件：角色真的互换（写得出“原来的终点在这段里被用来检验什么”），以及触发源可指认（写得出是什么触发的，教师宣布不算）。

换位触发源分布。每一次换位由什么触发：异常结果、对照出现、学生自己的迁移发现、外部约束变化。这个分布比次数更有用——全部由教师宣布触发的班，换位率再高也只是把新终点也指定了一遍。

正交性实例值得单写一句：一门目标写得最清楚、活动与目标对齐得最好、评价与目标一一对应、进度严格按大纲推进的课，换位率恰恰是零。因为这四件事合起来的意思就是终点从头到尾没有易过主。这说明换位率不是“课程质量”的另一种说法，它与现行的对齐类指标方向相反。

十、形式部分：换位在什么条件下值得做

前面几节是机制描述。本节把它写成能推出正文没说过的东西的形式，否则“信息在哪一端多就从哪一端开始”只是把化学那句话搬了过来。

设一段学习工作需要确定一条从当前状态到目标的路径。把不确定性分在两端：

目标端的不确定性 U_T ：这个目标究竟要求什么、达到它的表现长什么样，还有多少没定下来。

手段端的不确定性 U_M ：可用的做法有哪些、哪一种在这里管用、每一步的可靠性如何，还有多少没定下来。

从目标端倒推，等于假定 U_T 已经足够小，把全部工作放在消解 U_M 上；从手段端正推，等于假定 U_M 的地图值得先画出来，画出来之后 U_T 会被它改写。

一次换位——把原本作为手段的那一样临时提为终点——所做的事，正是把工作从消解 U_M 转为消解 U_T 。它有收益也有代价。

收益：换位期间获得的关于手段的知识，会反过来减少目标端的不确定性。记这个转化效率为 β ：每消解一单位手段端的不确定，带回 β 单位目标端的不确定性下降。 β 的大小取决于这两端有多相关——在手段决定产物形态的领域（化学工艺、设计、写作） β 大，在手段与产物几乎无关的领域（背诵、查表） β 趋近于零。

代价：换位期间原终点被搁置，进度停滞；且交还有失败风险。记搁置一单位时间的进度损失为 c ，交还成功率为 r 。

于是一次时长为 t 的换位，净收益为

$$\Delta V(t) = \beta \cdot \Delta U_M(t) \cdot r - c \cdot t$$

其中 $\Delta U_M(t)$ 是这段时间里**累计被消除**的手段端不确定性，即 $\Delta U_M(t) = U_M(0) - U_M(t)$ ； c 是单位时间的机会成本，全书只有这一个 c 。它是**凹的**——最开始下降最快（第一次认真看自己的做法，收获总是最大），随后边际递减。

由一阶条件 $d\Delta V/dt = 0$ 得到最优换位时长 t^* 满足

$$\beta \cdot r \cdot d\Delta U_M/dt |_{(t=t^*)} = c$$

这个式子推出三件正文前面没有说过的事。

其一，在换位期间目标端不推进这一建模假设下，最优换位时长与目标端的不确定性无关。 U_T 不出现在一阶条件里。这一条的边界必须与结论同时说出：**它是上述假设的后果，不是一条无条件的发现**；若换位本身会改写目标端，一阶条件变为 $\beta \cdot r \cdot d\Delta U_M/dt + d\Delta U_T/dt = c$ ，该结论即不再成立（下一节展开）。这与直觉相反：人们通常认为目标越不清楚，越该停下来琢磨方法。式子说不是——**该不该换位、换多久，由手段端还有多少可学、以及学到的东西能不能带回目标端决定，与目标本身清不清楚无关**。由此得到一条可检验的预测：在目标写得同样清楚的两门课之间，换位率的差异应当由 β 与 c 解释，而不由目标清晰度解释；而在目标清晰度不同的两门课之间，若 β 与 c 相近，换位率应当无显著差异。

其二，交还失败率直接压低最优换位时长，且压得比直觉更狠。 r 与 β 在式子里是相乘的，因此交还成功率下降一半，效果等同于转化效率下降一半。这解释了辨别格右下那个漂移格为什么危险：**在一个从不交还的课堂里，最优换位时长趋近于零——不是因为换位没价值，是因为它的价值取不回来**。处方随之明确：要提高换位率，最有效的一步不是鼓励换位，是把交还做实（明确规定这一段结束后回到哪个终点、以及原终点因这一段被改了什么）。

其三， β 的大小就是本章适用范围的形式表达。在手段与产物几乎无关的内容上， β 趋近于零，一阶条件在任何正的 c 下都无解，最优换位时长为零——**在这类内容上，从不换位是正确的**。这正是反向约束一节里教育学替本章划下的那条界限，现在它不再是一句让步，而是同一个式子在 $\beta \rightarrow 0$ 时的退化。本章没有两套说法，只有一套，参数不同。

最后交代这个模型的边界。它假定 β 、 c 、 r 在一段换位内为常数，实际上它们随内容与学生变化；它没有给出 $U_M(t)$ 的函数形式，因此不能算出具体的 t^* ；它假定换位的收益只经由 U_T 的下降兑现，而忽略了换位本身可能改变学习者的动机。上面三条推论只依赖式子的结构与各量的符号，不依赖函数形式——这是本章认为它们可检验的理由，也是本章对这个模型能力的全部主张。

十一、这个模型的假设与边界

上一节的式子只有在若干假设下成立。把它们逐条写出来，是为了让这个模型可以被针对性地反对。

其一，量的定义。 $\Delta U_M(t)$ 是**累计被消除**的手段端不确定性，不是剩余量 $U_M(t)$ ；两者关系为 $\Delta U_M(t) = U_M(0) - U_M(t)$ 。一阶条件里出现的是它对时间的导数 $d\Delta U_M/dt$ ，**不是 U_M 的导数**——这两者差一个负号，写混会使结论反向。

其二，凹性从哪来。最优解存在唯一，依赖 $\Delta U_M(t)$ 严格凹。这不是数学方便，是一条可被推翻的经验假设：**换位的头几分钟消除的不确定性最多，此后边际递减**。它成立的理由是手段端的不确定性通常集中在少数几处关键判断上；它失败的情形也可以指认——若某处方法的疑难要走到很后面才暴露（例如长周期实验的失败模式）， ΔU_M 会呈 S 形，一阶条件将有多个解，此时“最优时长”须改为“最优区段集合”。

其三，参数是否随时间变化。模型假定 β 、 r 、 c 在一段换位内为常数。这是**局部近似**：一段换位通常只有十几分钟到一节课。跨单元比较时三者都会变，因此本章的式子只用于同一段内的取舍，不用于跨课程比较。

其四，为什么目标端的不确定性不进入一阶条件。这不是发现，是**建模假设的直接后果**：模型假定换位期间目标端不推进（进度停在原处），因此目标端不确定性 U_T 只影响换位的机会成本 c 的大小，不影响 $d\Delta U_M/dt$ 与 c 的比较点。**若换位本身会改变目标端**——例如学生在检验方法时顺带重写了目标——则须把 U_T 写成 t 的函数并加入方程，一阶条件变为 $\beta \cdot r \cdot d\Delta U_M/dt + d\Delta U_T/dt = c$ ，此时“与目标端无关”这条结论不再成立。**这是本章那条反直觉后果的确切适用边界，不是它的反例。**

其五，未建模的三样东西。换位可能改变学生的动机、可能扩大后续的搜索空间、也可能使原目标被重新定义。这三样在当前模型里都不在。它们的共同效果都是抬高换位的真实收益，因此当前式子给出的 t^* 应视为**下界**。

十二、辨别格

两根轴：终点是否明确，以及终点是否易过主。

	终点未易主	终点易过主
终点明确	对齐良好：全部形式审查通过，目标——活动——评价一一对应，而学生下一次仍需人从头带。 本章的靶格。	有交还的换位：终点明确，且在过程中至少一次易主并被改写后交还。
终点不明确	空转：既没有终点也没有换位，活动很多而什么都不结算。	漂移：终点反复易主而从不交还，看起来很活跃，实际是没有任何一样东西被做完。

靶格是左上，三条签名齐：短期指标全部改善（对齐度、完成率、当堂正确率都最高）；失效不产生任何信号（没有人会因为“这门课的终点从没易过主”而报警）；自我加固（越正规化、模板越完整，终点越不可能易主）。

右下的漂移格是关键对照。它排除了“换位越多越好”——漂移格的换位率最高而学习最差，因为从不交还。**换位的价值不在次数，在交还时原终点有没有被改写。**

十三、走查一节课：同一节课的两种走法

抽象的读数需要一次具体的走查，否则编码规则永远停在纸上。下面是一节九十分钟的大学基础化学实验课，同一个内容、同一位教师、两个平行班，逐段对照。

甲班。前十五分钟教师讲解本次实验的原理与操作要点，强调三处易错。随后学生分组操作，教师巡视。第五十分钟左右，第三组的溶液没有出现预期的颜色变化。教师走过去，看了看，说温度低了，让他们把水浴调高五度，随后颜色出现。第七十五分钟各组记录数据、填写报告模板，其中两组的数据超出容许范围，教师让他们照另一组的重做一次。第八十五分钟教师总结，指出全班共同的两个操作问题，提醒下次注意。下课。

按本章的读数：**换位次数零**。从第一分钟到最后一分钟，终点始终是“做出预期的颜色变化并交出合格的数据”。第三组那一刻是一个真实的开口——“我们的操作哪里不同”本可以在那一刻成为要弄清的东西——而它被一句正确的、及时的、专业的诊断关掉了。两组超范围的数据同样：它们本可以成为“这个测量在什么条件下不稳定”的材料，而“照另一组重做”这个处置把它们从体系里移走了。教师的总结是对的，但总结是把已经消失的东西复述一遍，不是把它留下来。

甲班这节课的形式审查全部通过：进度符合、安全无事故、数据完整、报告齐全、课堂问题当堂解决。

乙班。前十五分钟相同。第四十八分钟，第二组同样没有出现颜色变化。教师这次问了一句：“在你们那边，什么和第一组不一样？”这一问不给答案，也不指方向，它把终点从“做出颜色”临时移到了“我们这一组的操作与别人差在哪儿”。第二组花了大约十二分钟排查，发现是加料顺序反了。这十二分钟里他们没有产出任何合格的数据。

关键在于接下来发生的事。教师没有让他们重做一遍了事，而是当场把加料顺序这一项加进了本次实验的记录表——一个原来没有的字段，要求所有组补填。第七十分钟，另有两组发现自己的顺序也和讲义写的不完全一样，虽然他们做出了预期的颜色。第八十五分钟，教师说下一次这个实验的讲义会把加料顺序单独列出来并说明原因，而这个改动的来源是第二组。

按本章的读数：**换位一次，触发源是一个异常结果（不是教师宣布），且完成了交还**。交还的证据有两条：原终点回来了（各组仍然做出并记录了颜色变化），且原终点被改写了（“做出颜色”现在包括“知道顺序为什么重要”）。此外这一次换位留下了一个可查的痕迹：记录表的字段变了，讲义的下一版会变，而这两处改动都能追溯到第二组的那一次排查。

乙班这节课的形式审查有两处不好看：第二组的数据不完整，进度慢了约十分钟。

两班的差别值得逐条摆出来。甲班的教师做对了每一件事，而且做得比乙班更快更准；乙班的教师那一问同样专业，但它把一次可以在三十秒内关闭的事延长成了十二分钟。若只看这一节课，甲班优于乙班：数据齐、进度准、无人卡住。本章的主张是，这个优势的

来源恰恰是那一次换位没有发生，而它的代价要到变式题、到撤除讲义、到下一个实验时才显出来——那时已经没有人会把它与这一节课联系起来。

需要说明的是，乙班那一问不是一种教学技巧。同样一句“什么和别人不一样”，如果教师在问完之后三十秒就把答案说出来，换位没有发生；如果学生找出了原因而记录表一字未改，换位发生了但没有交还，那一次落在漂移格里。**决定它是哪一格的不是那句话，是那句话之后有没有东西被留下来。**

十四、十处兑现

不是类比，是在那个领域里据此能预测出一件具体的事。

工艺开发。条件筛选跑完后改目标产物的项目，比全程锁定目标的项目更少出现后期返工。可从项目档案里核。

软件工程。在测试已通过之后发生的重写次数，与该模块后续的缺陷密度负相关。提交记录可查。

外科训练。从“完成这台手术”到“检验我这套入路”的换位，通常由术中意外触发；训练体系里有没有为这一刻设复盘节点，可查。

乐器教学。一首曲子从“要弹好的曲子”变成“用来听我的手型”的材料，这一次换位在教学日志里通常没有记录，而它是从会弹到会练的分界。

写作。一稿从成品变成“看清我的句法习惯”的材料，这一次换位是修改与重写的分界。

田野调查。研究问题在进入现场之后被改写，这在方法论文献里是正常操作且有专门名称，而在课程设计里没有对应位置。

产品开发。需求在原型跑完之后被改写，敏捷方法的整个价值主张建立在允许这次换位上。

运动训练。比赛成绩从终点变成检验训练方案的材料，这一次换位由成绩与训练日志的偏离触发。

语言学习。一次对话从“要说对”变成“看我哪些结构从来不用”的材料。

实验教学。异常数据从要排除的干扰变成要弄清的对象，这一刻在实验报告模板里没有字段。

十处里有四处（工艺开发、软件、田野、产品）在该行业已有制度化的换位节点与文档模板，而教育里没有。这不是教育落后，是教育的账本从来不为这一段设字段——它只记终点态。

十五、强制不一致：教学设计这门学问守不住的两条

教学设计同时持有两条主张，都写在它自己的教科书与流程规范里。

甲：目标必须在设计开始之前确定，并在实施期间保持稳定。这条不是偏好，是这门学问的操作前提。评估要与目标对齐，活动要与评估对齐，资源与课时要按目标分配；目

标一变，全套对齐关系作废。因此所有主流的设计流程都把目标确定放在最前面，并把“目标漂移”列为需要变更控制的风险。

乙：设计必须经过试用并依据试用数据修改。 这条同样是这门学问自己的核心环节，通常叫形成性试用或原型迭代：先做小规模试行，收集学习者的实际反应，据此修订设计。这一支的全部价值主张是**设计不可能一次做对**。

两条在一处正面相撞：**试用数据经常显示目标本身有问题**——目标要求的那种表现在这个学生群体的现实条件下不构成一个自然的成就单位，或者达到它所需的前置条件根本不在这门课的范围内。这时，乙要求依据数据修改设计，而在这个具体情形里，需要修改的正是目标；甲禁止修改目标。

这门学问现行的处置是把这类发现移出设计流程，归入“需求分析阶段的失误”，退回上一阶段重来。这个处置在形式上保住了两条，代价是：**试用能改的永远只是手段，而试用最常发现的问题在目标上**。于是形成性试用被系统地限制在它信息量最低的那一部分，而这一点在任何流程文档里都不会显示为问题——因为流程规定它就该这样。

二选一。要么承认目标可以在实施期间依据学习者的材料被改写，那么甲不能以现在的形式成立，全套对齐关系需要一个新的概念（本章提供的正是这一个：终点资格可以临时转移并在被改写后交还）；要么坚持甲，那么乙必须被削弱为“试用只用于修改手段”，而这与它自己关于“设计不可能一次做对”的主张冲突——因为一次做不对的地方主要就在目标上。

它不能同时守住两条，除非它承认本章的变量。 一旦承认，“目标漂移”这个风险项就必须被拆成两样：无来源的漂移（应当控制）与由学习者材料触发并被记录的改写（应当鼓励），而区分二者需要的正是换位的触发源与交还记录。

这门学问已经为这件事付出了可观察的代价：**大量课程在实施多年后目标一字未改，而其教材、活动与评价方式全部换过一遍**。目标的稳定不是因为它被验证过，是因为它被保护起来了。

十六、三处反向约束

教育学削掉“换位越多越好”。 低先验知识的学习者需要一个稳定的终点；在他还没有任何可用的做法之前要求他把做法当作终点，只会增加负荷而不产生任何东西。因此换位率有一个随先验知识上升而上升的下限，而不是越高越好。**如果没有这一处，本章原本会写下“课堂应当尽可能频繁地让终点易主”——这句话被这一处削掉了。**

化学削掉“筛选优于倒推”。 当目标已知、路线可靠、条件空间的信息量低时，倒推更省，硬做筛选是浪费。因此本章的适用范围收窄到**条件端信息量大于目标端**的那类内容——成熟技能的训练不在此列。这一处动的是适用范围。

艺术学削掉“换位可以被安排”。 工作室里的换位由作品自己触发，教师能做的只是不把它关掉，不能把它排进课表。因此本章的干预不能以“每单元安排一次换位”的形式实施——那样得到的是宣布，不是换位。这一处动的是干预形式，也是本章最容易被做坏的地方。

十七、三家为何各自到不了这一条

三家的材料都在，判断却没有出现。原因不是谁看漏了，而是每一家的问题结构恰好使这一步成为不必要的。

教育学到不了，因为它的委托关系里终点是被交办的。一门课程的目标通常不由讲这门课的人决定：它来自培养方案、来自课程标准、来自专业认证的毕业要求。教学设计这门学问的整个任务，是在**目标已经给定**的前提下把中间的部分做好。在这个任务里，“终点可以在过程中易主”不是一个被否定的选项，是一个不属于本职范围的问题——教师无权改终点，因此终点会不会易主根本不进入设计的变量表。这一家因此把“目标不清晰”当作唯一的病，而把“目标在过程中变了”当作管理事故。它看不见换位，不是因为证据不足，是因为在它的位置上这件事不该发生。

化学到不了，因为它的体系里只有一个所有者。逆合成与条件筛选的切换在化学里是一个人（或一个团队）的策略选择：我知道哪一端的信息多，我就从哪一端开始。这个切换不涉及任何权力转移，因此化学从来不需要问“谁有资格宣布终点变了”。而课堂里有两个所有者：教师与学生。同样一次终点易主，如果是教师宣布的，学生手里什么也没多；如果是学生的材料逼出来的，那才是本章说的换位。**这个差别在化学的体系里不存在，因此它的判据里没有“由谁触发”这一项——而这一项恰恰是本章的第三个条件。**化学替本章找到了那条判据，却不可能替本章找到那个条件。

艺术学到不了，因为它对记录本身怀有敌意。工作室传统最深的一条信念是：可以被写下来的部分不是那门手艺最要紧的部分。这条信念保护了它——它因此没有把教学压缩成一套可核查的动作清单。但它也使这一家在**原则上拒绝为任何过程建立字段**。换位在工作室里每天都在发生，而工作室永远不会知道自己一学期发生了几次，因为知道这件事需要记录，而记录在这一家看来是把活的东西钉死。本章的读数因此对它是一个冒犯，而不是一个补充。

三处缺口合起来说明这个判断为什么必须由三家撞出来：教育学有终点这个位置而无权动它，化学有互换这个动作而无两个所有者，艺术学有换位的日常而拒绝为它建账。少任何一家，剩下两家都推不到这一条。

十八、与既有说法划界

逆向设计（Wiggins & McTighe，教育学）。它说到的是：从目标倒推评估再倒推活动，能消除活动与目标脱节。分离线：它把终点当作设计前就确定并在整个实施期保持不变的东西，因而没有“终点易主”这个事件类。判定形式：若一门严格按逆向设计实施、终点全程未变的课，其延迟迁移不低于终点易过主并交还的同类课，则逆向设计足够而本章多余；若前者在原场表现更好而在新情境下更差，则本章对。

形成性评价。它说到的是：用学习证据调整教学。分离线：调整的是通往同一个终点的路径，终点本身不进入被调整的范围。判定形式：若把“终点是否被改写”作为协变量加入后，形成性评价对延迟迁移的效应不变，则本章不增量；若效应主要集中在终点被改写过的那些单元，则本章对。

刻意练习。它说到的是：针对当前弱点的、由即时反馈驱动的重复。分离线：它要求终点稳定且可测量，终点易主在它的框架里是训练计划被破坏。判定形式：在终点稳定的技能领域，本章预测换位率与结果无关（刻意练习完全解释）；在终点不稳定的领域，本章预测换位率仍有增量。这一条是本章最可能输的地方，也是最便宜的检验。

项目式学习与探究式学习。它们说到的是：把真实任务作为组织单位，让学生在完成任务中学习。分离线：任务是被给定的终点，学生在其中的自主性落在路径上，不落在终点上。判定形式：若项目式课堂的换位率显著高于讲授式课堂，则二者难以分开；若两类课堂的换位率相近而项目式只是路径更长，则本章指出了它们共同缺的东西。这一条本章预测后者。

目标置换（Merton 1940，社会学·方法学占位者）。把本章的命题剥掉一切教育名词，剩下的形式是“原为手段者取得目的资格”——这正是目标置换的形状，它已经有八十多年历史。这是本章最危险的一位占位者，必须正面回答。分离线：目标置换换的是手段永久性地吞并了目的，是一种失效；本章讲的是手段临时取得终点资格，随后交还，而原终点因这一段被改写。二者的差别落在交还上，而交还是可观察的。判定形式：若观察到换位之后原终点从未被交还、且系统的产出下降，则那是目标置换而本章判该次换位为失效；若观察到交还且原终点被改写、产出上升，则是本章说的换位。本章的辨别格右下角“漂移”格，正是目标置换在课堂里的位置——本章没有绕过它，是把它收进来作为四格之一。

手段——目的分析（Newell & Simon，认知科学·方法学占位者）。它说到的是：解决问题就是不断缩小当前状态与目标状态的差距。分离线：在它的框架里目标状态是搜索的外生参数，整个方法建立在目标固定上；一旦目标可变，“差距”就无法定义。判定形式：若人在解题过程中目标从不改变，手段——目的分析完备；若可观察到目标在解题中被改写且改写之后解题效率提高，则本章指出了它的边界。已有的问题解决研究里有大量后者的记录，只是被归入“问题重构”而不是“终点易主”。

探索与利用（March 1991，组织理论）。它说到的是：资源在探索新可能与利用既有确定性之间分配。分离线：它是一个资源分配问题，两端都是手段，不涉及终点资格的转移。判定形式：若把探索比例控制住之后换位率不再预测结果，则本章被它覆盖；若在探索比例相同的两个班之间换位率仍有区分力，则本章对。

十九、与作者同题诸文的划界

作者另有十余篇同样处理课堂学习的文章，其中三篇已收入本书第三编，其余仍散在书外；本章必须与它们逐一划清，否则只是又一个名字。总的分界线先说清楚：那些文章问的是“为什么很少发生”，本章问的是“该怎么运行”；前者要的是一个驱动，后者要的是一条次序。问法不同，因此读数不可互相替代。以下只列会被读者混为一谈的几篇，每条给双向判定。

与第八章《先挡住其余的》：同一组学科的两不同取向

本书第三编里有一章由同样这三门学科撞成，处理同一个问题。距离最近，必须写在最前面并写细，否则读者有理由认为两章是一件事的两种说法。

那一章处理的是**悬置**：要讲清一处，几乎总要先说一句“暂时不考虑”；它主张悬置期内被挡住的那部分会与新学内容自行结成关联，脱悬后无人拆除，也不进任何账本，因而命名了悬置连接。

先说两家共用学科的取用点：**同一门化学被两章拿走的是两件不同的东西**。那一章取的是**保护基策略与酰基迁移**——用来挡住一个位点的基团会自己跑到别处去，于是“挡住”这个动作改写了它本要固定的那张名单；本章取的是**逆合成分析与条件筛选的并存**，以及化学自己给出的选用判据。前者属于合成方法学里关于官能团操作的一支，后者属于研发策略的一支，互不引用，也不能互相推出。教育学与艺术学同理：那一章的教育学取认知负荷的分配，本章取目标先行与掌握学习；那一章的艺术学取作品的未完成状态，本章取工作室讲评对做法的训练。

再说**承重命题的分离**。那一章的对象是**被挡住的那部分在暗处结成的东西**——它是一个副产物，发生在注意力之外，当事人不知道它形成了；本章的对象是**终点资格的一次显式转移**——它发生在注意力正中央，且要求可指认的触发源与可查的交还。一个是暗的，一个是明的；一个是“没被看的地方长出了东西”，一个是“被看的那样东西换了身份”。

双向判定形式。一堂课每讲一个概念都先说“暂时不考虑”、悬置密度很高，而终点从头到尾没有易过主——那一章有读数，本章为零；反过来，一堂完全不做任何悬置的探究课里，学生把自己的做法提为检验对象并在被改写后交还——本章有读数，那一章为零。两格都存在真实的课堂形态，因此两个读数可同批测量而互不蕴含。

两章合用时给出一条各自单独说不出的预测：若一次换位恰好发生在某个悬置尚未解除的时段内，被悬置项与新取得终点资格的那一样之间会结成更多的连接，且这些连接更难被拆除——因为此刻的注意力正落在方法上而不落在被悬置的内容上。**两者的效应因此应当相乘而不相加**，这一条可以在同一批课堂录像上同时编码两个读数来检验。若实测发现两个读数的相关系数高于零点六，两章必有一篇是另一篇的代理，届时应保留可从教案版本与题库改动直接追溯的那一个。

与《**方程两边都有它**》（**旁观轮次·可逆性**）。那篇按谁走了那一步判归属，承重的是分岔步被代做。分离线：本章不要求任何一步易主。判定形式：教师完整地把分岔步留给学生、学生自己走完并卡住、教师随即讲清——那篇判无害（没有代做），本章判换位窗口被关掉；反过来，同伴代走了分岔步而学生随后把那一段当作检验自己做法的材料，那篇判有害，本章判换位成立。**两篇在这两格上给相反判断，可同批测量**。

与《**每一轮都从新鲜表面开始**》（**未结转中间体·结转率**）。那篇要有**未完成的东西**留下来才谈得上结转。分离线：换位处理的是**角色不是完成度**——一段完全完成的工作照样可以被提为检验的材料。判定形式：一个结转机制完备的班（每轮未完成项全部登记接续），若终点从未易主，那篇读数满分而本章读数为零；一个从不留未完成项的班，若每一节都有一次由学生材料触发的终点易主，本章读数高而那篇为零。**两读数正交**。

与《**先动的总是同一边**》（**节拍比·发起处轮转率**）。那篇的病灶在验收节拍对准了最快的那条驱动。分离线：节拍讲的是**结算的频率**，换位讲的是**终点的身份**。判定形式：把验收节拍调慢到对准最慢那条（那篇的处方）而不动终点归属，本章预测换位率不变——只是同一个零被拖长了。

与《记为负项的那一类改变》（溯改·判据修改）。那篇处理判据被修改这一类事件被记为负项。分离线：溯改改的是**判据**，换位改的是哪一样东西当**终点**；一次换位可以完全不动判据（用原来的标准来检验方法本身），一次溯改也可以在终点从未易主的情形下发生（终点不变而评分细则被反例修订）。判定形式：一个判据频繁被学生反例修订而终点始终是“考出这个成绩”的班，那篇读数高而本章为零。

与《一条路没有交接点》（交接点·独立结算比）。这是作者书外诸文中第一篇与本章同为“该怎么运行”的，因此距离最近，必须写细。分离线：那篇处理的是**谁有资格宣布这一段成了**，本章处理的是**这一段以什么为终点**。二者可以完全分离——一门由完全独立的第三方按事先锁定的判据判定达标的课，其终点可以从头到尾没有易过主（独立结算比满分、换位率为零）；一门达标判定全由授课方自按自判的课，其终点可以在一学期里易主多次（独立结算比为零、换位率很高）。判定形式：若两个读数的相关系数高于零点六，两篇必有一篇是另一篇的代理，届时应保留可从既有档案直接算出的那一个——也就是那篇。本章预测两者近乎无关，因为一个测的是**权力的位置**，一个测的是**角色的身份**。

与《差一条边》（接手图·闭合度）。那篇把接手关系写成一张图，处理的是这张图能否闭合。分离线：那篇的对象是**人与人之间的接手关系**，本章的对象是**终点与手段之间的角色关系**。判定形式：一个接手图完全闭合的班——每个人的产物都被某个人接住——其终点仍可全程未易主（每一次接手都是把同一个终点再推进一段）；反过来，一个图不闭合的班可以在教师与个别学生之间发生真实的换位。**两处合起来给出一条本章单独说不出的话**：换位要真正起作用，需要一张能把换位后的材料接住的图；图不闭合时，换位发生了也会消失在原地。

与《正好有人能解决》（清除位·透过率）。那篇处理的是差异在到达任何位置之前被局部胜任度消耗掉。分离线：那篇的对象是**被消耗的差异**，本章的对象是**被提为终点的手段**；一次换位不需要任何差异出现（学生可以在一切顺利时把自己的做法提为检验对象），一次清除也可以在终点易主之后发生。判定形式：一门透过率为零而换位率很高的课是可能的——所有差异都被当场解决，而终点仍在教师安排下易过主；本章预测这类课的效果介于两者之间，这是两篇合用时的一条新预测。

二十、与本编另两章的划界

本章与本编另两章同答一个问题，采用同一种做法，因此互为最容易被混同的两位。

与第五章《要到那儿，下一步该做什么有几个答案》（物理×心理×工程）。那一章处理的是**倒推这个动作本身在什么条件下有解**，读数是“下一步”的答案个数；本章处理的是**终点易主这个事件**，读数是易主次数。两者数学上独立：倒推唯一性高的区间里完全可以从不换位（这正是本章的靶格），倒推无解的区间里也可以频繁换位而毫无结果（那是本章的漂移格）。判决性对照：一门倒推唯一性最高的技能课，那一章预测应当用倒推、本章预测换位率会趋零且这在成熟技能上不是问题——**两章在这一格上给出方向一致但内容不同的判断，可同批测量**。

与第六章《这一节课结束时，这个班手里多了什么》（细胞×工程×艺术）。那一章处理的是**学习过程自己造出来、随后成为自己条件的那类东西**，读数是存量；本章处理的是**角色互换**。分离线：自建条件可以在终点从未易主的课上积累（例如一个班在纯讲授课

上自发形成的一套记号)，而换位可以在完全没有自建条件的课上发生。两个读数正交。判决性对照：若自建条件存量与换位率相关系数高于零点六，两章必有一章是另一篇的代理，届时应保留可清点的那一个。

二十一、五种最强的竞争解释，与排除它们的设计

前面给出的证伪条件只有排除了下面这几条，才算机制层面的检验，而不只是确认现象存在。每一条都配一个排除设计，且排除设计尽量落在既有数据上。

竞争解释一：其实不就是教师更会随机应变吗。这是最朴素也最难反驳的一条。愿意在异常出现时停下来的教师，通常在别的方面也更好。排除设计：用教案偏离次数独立测量灵活性。本章独有的变量与它正交——一位教师可以频繁偏离教案而终点从未易主（换了活动没换终点），也可以严格按教案走而终点在某一时刻易主（教案里就写着“这一段用你们的方案来检验上周定的标准”）。本章获胜的观察是：在教案偏离次数匹配的两组之间，换位率仍预测延迟迁移。

竞争解释二：其实不就是课上得慢、内容少吗。换位要花时间，慢下来的课自然内容少而思考多。排除设计：控制单位时间的内容覆盖量与总课时。本章获胜的观察是：在内容覆盖量相同的两个班之间，换位率仍有区分力；而若把课上得慢本身就足以解释效应，换位率在控制进度之后应当失去预测力。

竞争解释三：其实不就是学生水平高吗。水平高的学生才提得出能触发换位的材料。排除设计：在基线知识、语言能力与一般认知水平匹配后再看。本章独有的预测是：同一批学生在同一学期的不同单元之间，换位率会显著变化——若换位率主要随学生走而不随单元走，本章把一个结构问题当成了能力问题。

竞争解释四：其实不就是认知参与更深吗。换位期间学生在反思自己的方法，那正是深加工。排除设计：对认知参与做细粒度编码后看增量。本章获胜的案例必须包括两种反向分离：高参与而终点固定（预测迁移较低）；参与量相近而终点易过主并交还（预测迁移较高）。若加入参与变量后换位率完全失效，本章只是把深加工重新命名了一次。

竞争解释五：其实不就是评分者从好结果倒推好过程吗。编码者若知道延迟成绩，容易把前序事件读成换位。排除设计：时间盲法——过程编码在延迟测验揭盲之前完成；教案版本、题库改动与量规修订记录优先于访谈回忆；报告编码分歧率。若盲法之后换位变量的信度很低或效应消失，本章不能用“课堂过程复杂”回避失败。

判决可以压缩成一句：若在同一课程、同一教师或教师固定效应、内容覆盖量、教案偏离次数与先验知识匹配的条件下，由具体差异触发的终点易主先于延迟迁移的改善出现，且交还与改写的比例可解释效应的大部分，则本机制优于上述五条；若效应被灵活性、进度、水平或参与深度完全吸收，或时间顺序相反，则竞争解释获胜。

二十二、分层引用结构

本章的主张分三层，各自的证伪条件不同，可以分层引用。写清这一点是诚实，也是保护——第一层被推翻时，后两层仍然站得住。

第一层：换位区间作为一个存在物。 这是最强也最容易倒的一层。它主张一节课里存在一段可以被识别的区间，其中原为手段的那一样取得了终点资格并在被改写后交还。它的证伪条件是编码信度——若不同编码者对“角色是否真的互换”的判定无法达成一致，这一层就不可操作。

第二层：终点——手段两轴的辨别格作为一个分析工具。 即使第一层的存在物主张被判为过强，这张格子仍然可用：它把“目标是否明确”与“目标是否易过主”分开，而这两件事在现有的课堂评价里被压在一根轴上。这一层的证伪条件是正交性——若两轴高度相关，格子退化为一维。

第三层：一条不依赖前两层的经验主张。 目标写得最清楚、活动与目标对齐得最好、评价与目标一一对应、进度严格按大纲推进的课程，其学生在延迟的变式任务上不优于对齐度中等的课程。这一条完全不需要“换位”这个概念，可以由任何人用现有的对齐度量表与延迟测验直接检验，而且它是三层里最便宜的一条。

最容易检验的是第三层，最容易被推翻的是第一层，而本章最看重的是第一层。 这个顺序本身值得写出来：如果第三层被证实而第一层被推翻，那说明现象是真的而本章对它的刻画是错的；如果第三层被推翻，本章的三层一起倒。

二十三、证伪与赌注

最强的竞争解释是：**其实不就是课上得更灵活、教师更会随机应变吗。**

本章必须能排除它。灵活性可以用教师对预设教案的偏离次数测量；本章独有的变量是**终点资格的转移**，它与偏离次数正交——一位教师可以频繁偏离教案而终点从未易主（换了活动没换终点），也可以严格按教案走而终点在某一刻易主（教案里就写着“这一段用你们的方案来检验我们上周定的标准”）。

核心证伪条件：在同一门课程、同一批教师、控制教案偏离次数、时间投入与学生先验知识之后，若换位率与两周以上的变式迁移之间不存在剂量——反应关系，则本章为伪；届时应把课堂运行的差别归于教学灵活性，而非终点资格的转移。

其余七条：

其二，若换位率与“教师宣布关注方法”的次数高度相关，说明编码没有抓住触发源，只是在测教师的话语，本章的读数失效。

其三，若在终点稳定的技能领域换位率仍强预测结果，本章关于适用范围的收窄是错的，应扩大范围并重写反向约束一节。

其四，正向读数（顺序）：在完整的换位事件中，触发源应当早于教师的任何相关表述出现。若教师表述总是先于触发源，则换位是被安排的，本章的干预形式判断为伪。

其五，若换位之后原终点被改写的比例与结果无关，只有换位次数有关，则“交还与改写”这一条件多余，辨别格应简化为一维。

其六，最便宜的一条：**软件课的提交记录里，测试通过之后发生的重写次数与该模块后续缺陷密度的关系。** 这份数据在多数教学平台与代码托管里已经存在，不需要新的采集。

其七，若不同编码者对“角色是否真的互换”的判定一致性低于可接受水平，本章的核心读数不可操作，应退回为一个分析工具，不再主张剂量——反应。

其八，若在四个已有制度化换位节点的行业（工艺开发、软件、田野、产品）里，那些节点的存在与产出质量无关，则本章的兑现是同形而非真兑现。

写死日期的赌注：在二〇三〇年十二月三十一日前，至少两项在同一制度环境内、平行班不少于十二个、预注册的研究，报告换位率在控制教案偏离次数、时间投入与先验知识之后仍正向预测两周以上的变式迁移。

以下不算命中：只报课堂活跃度或学生满意度；只比较项目式与讲授式；把教师宣布关注方法的次数当作换位；用两个学校或两个国家的对比充当剂量——反应；机构自行说明已实施；与其他改革打包而无法分离。

二十四、反噬与治理

换位率一旦进入考核，最省事的实现方式是：在每个单元里安排一次“方法反思环节”，让所有学生在同一时刻写下自己用了什么方法。这个做法能把数字做满，而它恰恰不是换位——它是把新终点也指定了一遍，触发源是课表。更糟的是它比真换位省事得多，因为真换位要求教师在异常出现时不去关掉它，而这需要放弃当堂的进度。

第二重反噬：为了做高换位率而人为制造异常。教师可以设计一定会出错的任务，让“方法”顺理成章地浮上来。这样得到的换位有触发源，且触发源不是宣布，因而能通过本章的编码规则——本章的读数在这里是可以被骗过的。

三条治理底线：换位率指向课堂的运行结构，不指向教师或学生的稳定品质，不得单独用于排名；不得为提高读数而在安全关键的训练场合人为制造异常；任何依据该读数作出的不利安排，必须公开编码规则、原始证据与复核通道。

我看不到一个彻底的出口。只要这个读数被采用，制造换位就永远比允许换位省事。唯一可用的限制是把高分案例与延迟迁移绑定，并要求每一次换位都能指出一个**不是由本课设计出来的触发源**——而这一条本身也可以被绕过。

二十五、若本判断被证伪，我们会学到什么

一条判断的价值不只在它成立时给了什么，也在它不成立时排除了什么。本章把这一条单列，因为它决定了这项检验值不值得做。

若换位率对延迟迁移没有增量预测力，我们学到的不是“课堂运行方式无关紧要”，而是一件更具体的事：**终点的身份不是课堂运行的承重变量，承重的是终点的稳定性**。那意味着教学设计那条把目标锁死在最前面的次序是对的，而它遇到的所有麻烦（试用发现目标有问题、高层级目标不可操作化）应当在需求分析阶段解决，不应当在实施中解决。这是一个明确的结论，且它会改变现在这种“一边锁死目标一边靠实施中的调整补救”的两头不靠。

若换位率有预测力而交还与否无关，我们学到的是：起作用的是**注意力从终点移开这个动作本身**，而不是它带回了什么。那会把本章的处方从“记录交还”简化为“每单元安排一次方法反思”——也就是本章在反噬那一节判为无效的那个做法反而是对的。这个结果对本章不利，但它比本章的原主张便宜得多，值得优先检验。

若靶格与漂移格的延迟迁移无差异，我们学到的是：终点从不易主与终点乱易主，在结果上是同一件事。那意味着这两根轴合成一根就够了，本章的辨别格应当退回一维，而“目标稳定性”这个现有概念足以覆盖全部现象。

三种失败方式各自指向不同的下一步，这是本章认为这项检验值得做的理由。

本章交出三个自己解释不了的洞。第一，换位的最优频率与内容的什么性质有关，形式部分只说了它随转化效率上升，没有说清哪一类内容的转化效率高。第二，交还与不交还的分界在具体课堂里往往不清楚——一稿被放弃了，原终点算交还了还是没交还？编码规则在这一点上会产生真实分歧。第三，本章完全没有处理**换位对学习者的负担**：把自己的做法当作对象来看是一件消耗很大的事，而本章的整个论证都假定这个消耗可以忽略。

二十六、自反检验

按本章自己的判据，这篇文章的终点从头到尾没有易过主。它开头就写清了要论证什么，中间的每一节都是达成它的手段，结尾交出了同一样东西。它落在自己的靶格里：形式审查全部通过，对齐良好，而读者读完之后手里多了一个名字，课堂的运行方式一处没变。

这个定位有具体后果。本章不能用内部一致性当作证据，不能因为能解释很多场景就宣称成立。它能做的只有把编码规则、竞争解释与失败条件交出来，并且承认：如果这套东西被使用，最先被改的应当是本章的反向约束一节——那三处约束是三家给的，不是本章自己想让步的，因此也最可能在实测中被推翻。

本章交出两个自己解释不了的洞。第一，换位的最优频率与内容的什么性质有关，本章只说了它随先验知识上升，没有说清是哪一类内容更容易产生自发换位。第二，交还与不交还的分界在具体课堂里往往不清楚——一稿被放弃了，原终点算交还了还是没交还？本章的编码规则在这一点上会产生真实的分歧，而这个分歧本章没有解决。

二十七、结论

课堂如何运行才能让学习真正发生，三门学科给了三条不能并存的次序：从表现倒推、从条件正推、以学习者的做法为终点。三家共同假定终点是一个可以被指定的位置。

化学自己拆掉了这个假定。它同时使用逆合成与条件筛选，而选用哪一套的判据是信息在哪一端多——终点与手段的角色由信息分布决定，因而是可以互换的。

由此得到一条三家单独都推不出来的判断：**课堂运行的对错不在选哪个终点，而在有没有一段区间让原为手段的那一样取得终点资格，并在被改写之后把资格交还。**学习不发生在任何一条路径的内部，因为每条路径内部另两样都被降格为手段，而被降格的那一样恰恰是此刻正在成形的东西。

这句话的可操作形式是一句问话：这一节课里，有没有哪一刻，原本用来达成目标的那样东西本身成了要达成的东西——是哪一刻，由什么触发的？

如果一门课的答案永远是“没有”，那么它的目标可能写得非常清楚，活动与评价可能对得非常齐，形式审查可能一项不缺，而它从头到尾只在做一件事：把已经定好的那个位置到达一遍。

接续。本章的一阶条件在第十三章与另两条结构模型并排，那里给出三次统一尝试与它们各自失败在哪。换位率的采集办法、以及“触发源必须同时报、教师宣布不算”这条防造假要求，见附录一与第十二章第六节。本章与第八章共用同三门学科而取用点不同，这件事本身携带的信息由第十章处理。

第五章 下一步该做什么，有几个答案

判断卡片：“下一步该做什么”不是一个总有答案的问题——它可以有一个、多个或没有；而在长期迁移与可选择性权重较高的内容上，档位不确定时应当默认推进，因为错配的代价在两个方向上不对称。

三个理论入口：物理学 × 心理学 × 工程学

提要。课堂如何运行才能让学习真正发生？三门学科给出三条完整而互不相容的次序。物理学的变分传统说，一个系统走的路径不是一步步试出来的，是由两端条件整体决定的，因此正确的做法是把两端定住，路径自己出来。心理学的刻意练习传统说反了：路径只能一步一步走，每一步都要有针对当前弱点的任务和即时反馈，整体最优是走出来的不是算出来的。工程学的稳健设计传统说两家都在追一个不该追的东西：最优点恰恰是最脆弱的点，要做的是把条件场做得足够宽，让系统在扰动下仍然可用，代价就是主动放弃最优。三条不能同真——刻意练习要最大化的正是稳健设计要牺牲的；而变分说的“整体决定”需要两端都给定，刻意练习却只给定一端。三家共同站着而从未说出的那块地是：**从当前状态到目标状态，“下一步该做什么”是一个有确定答案的问题；分歧只在于该怎么求这个答案。**推翻它的材料来自物理学自己。同一个动力系统，写成初值问题在很弱的条件下就保证解存在且唯一；写成两点边值问题则可能无解，也可能有无穷多解。物理学与数值分析对此毫不含糊，并据此选择方法——而这条在物理里回答的是“该用打靶法还是直接积分”，是一个计算问题，从来没有人把它当作“该不该先定终点”的证据。**倒推与推进不是同一个动作的两个方向，它们的适定性根本不同。**由此本章命名**可倒推区间**：学习进程里的一段区间，在其中“要到达那个目标，下一步该做什么”有且只有一个答案。出了这段区间，倒推要么无解、要么多解，而这三种情形要求三种不同的课堂运行方式。判据一句：**把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的“下一步”是同一件事、五件不同的事、还是都说不出来？**读数**倒推唯一性**可直接测量，数据在多数机构里当天就能采到，且与“目标是否明确”正交——目标写得最清楚的课，倒推可以完全无解。

一、三条次序，第一步就分开

“课堂如何运行”要的是一条次序：从哪儿起，经过什么，实现什么，在哪一步能插手。三门学科各有一条，而且都不是空谈。

物理学的变分传统给的次序是：定住两端，路径自己出来。这条次序在物理里有极强的地位——力学、光学、场论都可以从一个作用量泛函的驻点条件重新导出。它的教学版本是：把起点（学生现在会什么）与终点（学完要会什么）都写死，中间的安排是可以被求出来的。

心理学的刻意练习传统给的次序相反：路径只能一步一步走。每一步都要有一个刚好超出当前能力的任务、一个即时反馈、一次针对性重复。整体最优不是算出来的，是在无数局部改进里长出来的。这条次序背后有一个硬主张：**人无法从终点倒推出自己的下一步**，因为他不知道自己缺什么——这正是教练存在的理由。

工程学的稳健设计传统给第三条，并且把前两家一起否掉：追最优本身是错的。参数空间里性能最好的那个点，往往也是性能对参数变化最敏感的那个点；把系统放在那里，任何扰动都会让它掉下来。正确的做法是选一个性能稍差而对扰动不敏感的区域，把工作点放在平坦处。它的教学版本是：不要把学生训练到只能在标准条件下发挥最好，要把他训练到在条件变化时仍然能用。

三条互不相容。刻意练习要最大化的是特定条件下的表现峰值，这正是稳健设计要主动牺牲的东西；变分说的“整体决定”要求两端都给定，而刻意练习的整个工作前提是终点只是一个方向、不是一个可以代入计算的边界条件；稳健设计则说两端定得越死，做出来的东西越脆。

二、物理学这一家：路径由两端整体决定

把第一家推到最强：一个系统的实际路径不是逐步试探的结果，是由两端条件与一个整体量共同决定的。

这不是一个哲学立场，它有极硬的形式。给定起点与终点，实际路径是作用量泛函的驻点；由此可以导出运动方程，而且导出的方程与从局部受力分析得到的完全一致。这个等价性本身很了不起——它说明“整体决定”与“逐步推进”在这一类系统里是同一件事的两种写法。

它对课堂的主张因此很直接：把学生现在的状态与要到达的状态都刻画清楚，中间该怎么走是可以被求解的，不需要靠试。教学设计里所有“倒推”的做法，其数学骨架就是这一条。

这一家自己知道的一件事，推翻材料一节会用到，这里先按下不表：这个等价性是有条件的，而条件恰恰落在“两端都给定”上。

三、心理学这一家：只能一步一步，而且他自己看不见下一步

把第二家推到最强：可迁移的能力只能通过针对当前弱点的、有即时反馈的重复获得，而学习者本人无法确定自己的下一步。

这一条的证据非常集中。在有明确评价标准、有长期训练传统的领域里，顶尖水平与训练年限的关系远不如与“结构化训练时数”的关系强；而结构化训练的定义里，最关键的一条正是任务由教练根据当前表现指定，而不是由学习者自选。

这一家对物理学那条次序的诊断是：两端根本没有都给定。学生的当前状态不是一个可以代入的初值——他自己说不清，教师也只能通过表现推断；而终点也不是一个坐标，是一个方向。在两端都是模糊的情况下谈“整体求解”，求出来的是纸面上的路线。

它对稳健设计的诊断更狠：在能力还没有长出来之前谈鲁棒，等于在还没有峰值的时候就开始削峰。

这一家自己承认的边界很清楚，而且写在它自己的文献里：刻意练习的效应在有稳定评价标准、有成熟教练体系、有明确子技能分解的领域最强，在缺少这些条件的领域显著减弱。它把这个减弱解释为“这些领域的训练还不够结构化”——也就是说，它把边界当作了有待改进的现状，而不是当作一个变量。

四、工程学这一家：最优点就是最脆的点

把第三家推到最强：设计的目标不是让性能达到最优，是让性能对不可控因素不敏感。

这一条的形式同样很硬。把可控参数与不可控扰动分开，在参数空间里寻找的不是性能均值最高的点，而是性能方差最小的区域；这两个位置通常不重合，而且往往相距很远。工程实践里大量的经验是：把工作点放在性能曲线陡峭的地方，实验室里表现最好，产线上一塌糊涂。

它对课堂的主张是：不要把学生练到只在标准题型、标准条件、标准提示下发挥最好。要找的是那个“换了条件仍然能用”的平坦区域，哪怕在标准条件下表现差一截。

它对另外两家的诊断合成一句：**两家都在求最优，一家用解析的方式求，一家用迭代的方式求，而最优本身是错的目标。**

这一家自己的边界也清楚：稳健设计要求性能可测、扰动可枚举。当目标本身说不清是什么的时候，“对扰动不敏感”就没有定义——因为你不知道对什么不敏感。

五、三家共同站着的那块地

三家争的是“下一步该做什么”**这个答案该怎么求**：解析地求、迭代地走、还是干脆不求最优只求平坦。

争论再激烈，三家都默认了同一件事：**这个问题有答案**。变分说答案可以算出来；刻意练习说答案只能由教练给出但确实存在；稳健设计说答案是一个区域而不是一个点，但区域也是答案。三家的处方全部建立在这个假定上——如果“下一步该做什么”根本没有答案，或者有无穷多个同样好的答案，那么“该怎么求”就不成其为问题。

把这个前提念给三家听，三家都会说这还用说吗。学生现在在这儿，目标在那儿，当然存在一件此刻该做的事。这正是共有前提的签名。

六、推翻它的材料，来自物理学自己

物理学与数值分析对下面这件事毫不含糊。

同一个二阶动力系统，可以写成两种问题。

初值问题：给定初始位置与初始速度，往前积分。在很弱的正则性条件下，解存在且唯一。这是一个适定问题——初值的小扰动只带来解的小变化。

两点边值问题：给定起点位置与终点位置，求中间的路径。这个问题**可以没有解**（在给定时间内根本到不了那个位置），**也可以有无穷多个解**（好几条路径都恰好在那一刻到达），而且即使有唯一解，它对边界条件的敏感度也可能高到没有实用价值。数值上要用打靶法反复试探，且不保证收敛。

这个不对称是数学事实，不是工程经验。而且它就在第二家那一节那个“等价性”的旁边：变分原理确实能导出与局部方程相同的运动方程，但**变分表述本身是一个边值表述**——它假定两端都给定。一旦你只知道一端，那个漂亮的整体求解就退化成一个必须从这一端往前推的问题。

物理学在自己的领域里用这条差别回答的是**该选哪种数值方法**：知道初值就直接积分，知道两端就打靶，打靶不收敛就换方法或者承认这个边值问题不适定。这是一个计算策略问题，被写进每一本数值方法教材。没有人把它当作“该不该先定终点”的证据——而它恰恰是。

它推翻的正是共有前提。“下一步该做什么”这个问题不是必然有答案的：它是一个从终点倒推的问题，而倒推问题的适定性不是给定的，是随具体的起点、终点与可用时间变化的。同一门课，在某些区间上倒推有唯一解，在另一些区间上无解，在第三些区间上有无穷多解——而三家的处方各自只在其中一种情形下成立。

七、命名：可倒推区间

由此本章命名一类区间。

可倒推区间：一段学习进程，在其中“要到达那个目标，此刻该做什么”这个问题有且只有一个答案。

区间的三种状态：

唯一。问五个懂行的人，五个人说同一件事。这类区间通常出现在有稳定子技能分解、有明确失败模式、有成熟诊断传统的内容上。

多解。问五个懂行的人，得到五件不同的事，而且每一件都说得通、都能到。这不是“没人知道”，是**确实有多条路都能到**。

无解。问五个懂行的人，五个人都说不出一件具体的下一步，只能说出活动的名字（“多读多写”“多做练习”“培养兴趣”）。这类区间的签名很好认：给出的“下一步”如果换成任何别的目标也一样说得通，那它就不是下一步。

三个条件把这个构念钉住：

它是**知识分布的性质**，不是学生的性质。同一个学生、同一个目标，换一批教师，答案的个数可以从多解变唯一——因为那批人手里有更细的诊断分解。这一条把它与“这个学生学不学得会”彻底分开。

它是**分段的**，不是全局的。一门课在不同阶段落在不同状态里，而且状态会随进度变化；把整门课判成一种状态是错的。

它与**目标是否明确正交**。这是最容易被误读的一点，值得单说：一门目标写得极其清楚、行为动词与评价标准一一对应的课，倒推可以完全无解——因为“能够批判性地评价一个论证”写得再清楚，也没有人说得出此刻这个学生的下一步是哪一件具体的事。反过来，一门目标只写了一句“把这首曲子弹下来”的课，倒推可以是唯一的——教师一听就知道下一步该练哪四小节。

八、判据：一句可以拿去问人而不问判断的话

把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的“下一步”是同一件事、五件不同的事、还是都说不出来？

这句话不含情态词，不问该怎么教，只问答案的个数。它可以在一节课的时间里实测，成本接近零：写一份学生当前作业的匿名样本，附上目标，发给五个人，收回来数一数。

五个场景跑一遍，答案互不相同：

代数变形与解方程。唯一。五个人会说同一件事，通常连要练哪一类变形都一致。这解释了为什么刻意练习与掌握学习在这类内容上效果最强——它们本来就是为唯一区设计的方法。

写一篇有论证的文章。多解。五个人会给出五件不同的事：先改结构、先补证据、先删冗余、先明确读者、先重写开头。每一件都能到，而且都有人做成过。这不是标准不清，是路径确实多。

乐器或运动上的一个具体卡点。唯一，而且一致程度惊人。这一条与第二个场景形成对照：同一个学生在“整体音乐性”这个目标上是多解的，在“这个换把总是慢半拍”上是唯一的。同一门课的不同区间落在不同状态里，**这条对照就是证据。**

“培养批判性思维”。无解。五个人给出的是活动名称——多讨论、读一手材料、做辩论。把目标换成“培养创造力”，同样这五句话照样说得通。这是查出来的：教师给出的东西通过不了“换个目标还成不成立”这一关。

临床诊断推理的训练。多解，且专家之间的路径差异远大于结论差异——已有的专家——新手研究反复记录过这一点，通常被解释为“专家有多种表征”。**同一份材料在本章这里是另一个读数：这个区间的倒推有多解。**这也是查出来的，不是从命题推出来的。

后两个答案不能从命题推出来，它们是拿现成的文献与现成的教研记录数出来的。

九、由此得到的完整次序

这是一道要“怎么办”的题，因此本章必须交出一条完整的次序，而不只是一个名字。

第一步：测。在一个单元开始之前，用判据一节那句话测一次倒推唯一性。成本是五个人各花十分钟。得到三档之一。

第二步：按档换运行方式。三档要求三种不同的课堂运行，而且它们互不通用。

在**唯一区**，用倒推。逆向设计、掌握学习、刻意练习、即时反馈——这些方法在这里全部成立，因为倒推问题在这里适定。这一档的插手点在诊断的精度上：唯一性越高，教练的价值越大。

在**多解区**，不要倒推，要推进加选择。从当前状态出发往前走一小段，看走出了什么，再在若干条路里选。这一档的插手点不在“指定下一步”，在**把多条路同时可见**——让学生看到别人选了不同的路且都走到了。多解区里最常见的失误是教师把自己那条路当作唯一的路讲，这在形式上完全合规，而它把一个多解问题假装成了唯一问题。

在**无解区**，两种方法都空转，必须先动目标。把说不出下一步的目标改写成能说出下一步的形式，改写之后重测。**这一步是本章最有用的一处，因为它把一个通常被归为“教师能力不足”的现象重新定位为一个可测的、目标层面的问题。**若改写之后仍然无解，则这个目标在当前的知识条件下不可教，应当明说，而不是继续用活动填满时间。

第三步：在边界处插手。区间会移动。一个多解区在学生积累了足够的判断之后可能收敛为唯一区；一个唯一区在学生越过基础之后会展开为多解区。**边界的移动本身是学习的读数，而且它比表现更早出现——这是证伪一节那条正向读数的来源。**

十、形式部分：错配的代价在两个方向上不对称

前面几节给了三档与三种运行方式。本节回答一个前面没有回答的问题：**如果测不准档位，该默认用哪一种？**这个问题的答案不是“再测一次”——它可以从错配的代价结构里直接推出来，而推出来的结果与通行做法相反。

设一段学习工作，要从当前状态到达目标。记倒推的解的个数为 n 。

在唯一区用推进法（ $n = 1$ 而误用推进）。推进法从当前状态往前走一小段，看走出了什么，再选。由于此处只有一条路通到目标，走出来的东西迟早会落回那条路上。损失是**搜索的冗余**：期望多花的步数正比于分支因子与走错后的返回成本，而且它是**有界的**——最坏情形是把该走的那条路晚一点找到。记这个损失为 L_1 ，它随时间线性累积。

在多解区用倒推法（ $n > 1$ 而误用倒推）。教师从目标倒推，得到的是 n 条路之一；他会把这一条当作**那一条**讲出来。学生照做，也能到。损失不在效率上——效率甚至更高，因为省掉了选择。损失在**可达集的收缩**：学生学到的是一条路，而不是“这里有多条路以及怎么选”。当条件变化、那一条路不通时，他没有任何东西可以退回来。记这个损失为 L_2 。

关键的不对称在这里： L_1 是效率损失， L_2 是**能力损失**，而且 L_2 在原场表现上完全不可见——学生在这门课的所有考核里都能做对，因为考核也走那一条路。

在无解区用任何一种（ $n = 0$ ）。倒推无解，推进没有方向。两种方法的产出都不是学习，而是活动。损失是**全部投入**，记为 L_3 ，且 L_3 与投入时长成正比、与方法无关。

现在把它写成一个决策问题。设档位判定有误差：真实档位为 n ，测得档位为 $\hat{n}$ 。选择方法 $M \in \{\text{倒推}, \text{推进}\}$ 。期望损失

$$E[L] = \sum_n P(n | \hat{n}) \cdot L(M, n)$$

在 $\hat{n}$ 判为唯一而实际可能是多解的情形下（这是最常见的误判方向，因为教师本人熟悉那一条路，很容易把多解读成唯一）：

$$E[L | \text{选倒推}] = P(1) \cdot 0 + P(>1) \cdot L_2$$

$$E[L | \text{选推进}] = P(1) \cdot L_1 + P(>1) \cdot 0$$

选推进优于选倒推，当且仅当 $P(1) \cdot L_1 < P(>1) \cdot L_2$ 。

由于 L_2 是能力损失而 L_1 是效率损失，在以迁移为目标的内容上 $L_2 \gg L_1$ ；因此即使 $P(1)$ 相当高——即使你相当有把握这是唯一区——不等式仍然成立。

这推出三条正文前面没有说过的事。

其一，在在长期迁移与可选择性权重较高的内容上，档位不确定时应当默认推进法，而不是默认倒推法。这与通行做法相反：教学的默认设置几乎总是倒推（先讲那条路），而把推进当作特殊安排（探究课、项目课）。默认设错的代价不对称——把唯一区当多解区处理，只是慢一点；把多解区当唯一区处理，学生失去的是他本来可以拥有的可选择性，而且这件事在这门课里永远不会被发现。

其二，最危险的不是无解区，是被误判为唯一区的多解区。无解区的损失 L_3 虽然是全部投入，但它会被发现——学生做不出来，课堂上会有信号。而误判的多解区在所有可见指标上都表现良好。因此本章的辨别格右上那一格（目标明确、倒推非唯一）里，真正的

靶子不是无解那一半，是多解那一半：它更常见、更隐蔽、更容易被当作教学质量高的证据。

其三，教师的经验会系统性地把 $P(1)$ 高估。一位在某个内容上教了十年的教师，他自己那条路已经高度自动化；他给出的“下一步”确实稳定而唯一——**对他自己而言**。判据那一节的问话之所以要问五位教师而不是问一位，正是因为单个专家的内部一致性会被误读为问题的适定性。由此得到一条可检验的预测：**同一份材料，问五位来自不同培养背景的教师，与问五位同一教研室的教师，得到的档位分布应当不同；后者会系统性地偏向唯一**。这一条可以在一次跨校教研活动里直接测出来，而且它同时是本章读数的一个信度检验。

模型的边界要交代清楚。它假定 L_1 、 L_2 、 L_3 在一段工作内为常数，实际上它们随内容变化很大；它没有给出 L_2/L_1 的具体比值，因此不能算出临界的 $P(1)$ ；它假定误判只有一个方向，而实际上把唯一区误判为多解区也会发生（虽然代价小得多）。上面三条推论只依赖 $L_2 \gg L_1$ 这一个不等式与各量的符号，不依赖具体数值——这是本章认为它们可检验的理由，也是本章对这个模型能力的全部主张。

十一、两种损失怎样才可比：默认规则的适用条件

L_1 与 L_2 不在同一量纲上——一个是时间与资源，一个是未来可选择性。两者只有先放到同一把尺子上，“ $L_2 \gg L_1$ ”这句比较才有意义，而上一节那条默认规则也才有确定的适用范围。

其一，把两种损失放到同一把尺子上。效率损失是时间与资源，能力损失是未来可选择性；两者只有先映射到同一个效用函数才可比较：

$$L = L_{time} + \lambda \cdot L_{future} + \mu \cdot L_{resource},$$

其中 λ 是这门课对未来迁移与可选择性的价值权重。于是 $L_1 \approx L_{time} + \mu \cdot L_{resource}$ （在唯一区误用推进，多走了几步）， $L_2 \approx \lambda \cdot L_{future}$ （在多解区误用倒推，一条路被指定，其余路未被留痕）。

其二，“ $L_2 \gg L_1$ ”因此不是一条普遍命题，而是一条关于 λ 的命题。它成立当且仅当 λ 足够大——即这门课的价值主要在于学生此后能在多少种情形里重新组织自己的做法。

其三，默认规则必须相应收窄。正文那句“在长期迁移与可选择性权重较高的内容上，档位不确定时应默认推进”，正确的写法是：

在以长期迁移与可选择性为主要目标、且未来能力损失的权重显著高于当下效率损失的内容上，档位不确定时应默认推进。

反过来， λ 小的场合默认规则不成立，且这些场合并不罕见：急救与安全关键操作的程序训练（此刻做对压倒此后可迁移）、考前短期冲刺（时间窗口内不存在“此后”）、一次性合规流程（本来就只有一条路）。在这些内容上默认倒推是对的。

其四，这一收窄不削弱本章的判断。本章真正的发现不是“该用哪一种”，而是两个方向的错配代价不对称，且其中一个方向的代价在原场看不见。不对称性来自 L_{future} 在当期不产生任何可观察差异，与 λ 的大小无关； λ 只决定这条不对称性在具体课程上值不值得据以改变默认。

十二、辨别格

两根轴：目标是否明确，倒推是否有唯一解。

	倒推唯一	倒推非唯一（多解或无解）
目标明确	可训练区：方法与问题匹配，刻意练习与倒推设计在这里全部成立。	齐整的空转：目标写得漂亮、对齐完美、活动饱满，而没有人说得出下一步。 本章的靶格。
目标不明确	隐性传承：说不清要什么，但老师一看就知道下一步该练什么。学徒制的常态。	泥地：既说不清要什么，也说不出下一步。

靶格是右上，三条签名齐：短期指标全部改善（目标清晰度、对齐度、大纲完整度全是满分）；失效不产生任何信号（没有任何检查会去问“这个目标下一步是什么”）；自我加固（越规范化，目标写得越抽象越漂亮，越说不出下一步）。

左下的“隐性传承”格是关键对照，而且它推翻了一个常识：**目标不明确而倒推唯一的课，本章预测其效果优于目标明确而倒推无解的课。**这与“先把目标写清楚”这条通行做法方向相反，且可以直接检验——学徒制、师带徒、乐器教学都落在左下。

十三、走查一节课：三个档位在同一节课里同时出现

三档听起来像是三种课，实际上它们常常同时出现在同一节课的不同时刻。下面走查一节九十分钟的大学写作课，逐段判档。

第零到二十分钟：讲评上周作业的一个共同问题——引用的证据与所支持的论点之间没有说明关系。教师给出三个例子，逐句改。这一段落在**唯一区**：把某个学生的原句拿给五位写作教师看，问“下一步该做什么”，五个人会说同一件事——在证据与论点之间补一句说明关系的话。倒推适定，教师的示范与练习都恰当，这一段本章没有任何异议。

第二十到四十五分钟：学生就本周主题各自列出提纲。有一位学生的提纲结构完整、每一部分都成立，而整篇没有一条贯穿的判断。把这份提纲交给五位教师，问下一步——答案会是五个：先想清楚你到底要说什么、先删掉一半、先找一个反例、先把最后一段挪到前面、先去读一篇写同一题目的好文章。**这一段落在多解区。**五个答案都能到，而且每一位教师都能举出自己那条路成功的学生。

本章的主张在这里落地：这一段正确运行不是选一条讲出来，而是让五条同时可见。可执行的做法很朴素——把这五种说法都摆出来，让学生自己选一条走，并要求他记录自己选了哪一条。这不是“尊重多样性”，是因为在多解区里指定一条的代价是学生失去可选择性，而这个代价在这门课的任何考核里都不会显出来。

第四十五到七十分钟：学生按自己选的路修改。仍在多解区，教师巡视时的动作应当是问“你选了哪一条、为什么”，而不是“我建议你先……”。

第七十到九十分钟：教师提出本课程的总目标之一——“形成你自己的声音”。把这句话交给五位教师，问此刻这个学生的下一步——五位都说不出一件具体的、此刻可以开

始做的事。给出的会是“多写多读”“找到你真正想说的”“不要模仿”。把目标换成“培养批判性思维”，这三句照样说得通。**这一段落在无解区。**

本章对这一段的处理不是回避，是明说：这个目标在当前的知识条件下说不出下一步，因此这二十分钟里发生的不是学习而是活动。可做的只有一件——把它改写成能说出下一步的形式。一个可行的改写是“在你的三篇作业里找出重复出现的句法习惯，并写清其中哪两条你打算保留”。改写之后重测：五位教师会说同一件事（把三篇并排读、标出重复的结构），档位从无解变成唯一。

这次改写有代价，必须诚实写出来：改写后的目标不等于原来那个目标。“找出重复的句法习惯”不是“形成自己的声音”，它至多是通向后者的一段。本章在反噬那一节已指出，改写与掉换之间没有可靠的分界，而这一节的走查正是那个风险的具体形态。区别只能靠一条：**改写后的目标与原目标要在一个独立的延迟测验上可比；**做不到这一条，就应当承认这是掉换而不是改写。

这节课的三档同时存在，而它的课程大纲把三段写成同一层级的三个教学环节。这是本章最想说清的一件事：档位不随课程走，随内容与知识分布走；按周次、按单元、按课时组织的教学安排，在结构上无法把三档分开。

十四、十处兑现

数值求解。边值问题不收敛时改用延拓法或换成初值表述，这是标准操作，有对应的教科书章节；教学里没有对应的动作。

手术训练。具体入路的下一步唯一性高，整体判断力的下一步唯一性低；训练体系对前者有细致的分解，对后者只有“多上台”。

飞行训练。程序性科目倒推唯一（检查单本身就是倒推的产物），情境判断科目多解，而模拟机的科目分配恰好反映了这一点——这是查得到的。

编程。“让这段跑通”唯一，“把这个系统设计好”多解；教学资源的分布与倒推唯一性高度一致，而没有人用这个词。

语言学习。发音的某个具体音位唯一，“写得地道”无解；后者的教学长期依赖大量输入而非指定下一步，正是无解区的正确应对。

物理竞赛训练。题型内唯一，跨题型多解；训练体系用大量刷题解决前者，而对后者只有“多想”。

产品设计。需求已定时倒推唯一，需求未定时无解；行业的应对是先做原型把无解区变成多解区——这正是完整次序第三步的工业版本。

康复训练。具体动作模式唯一，功能性恢复多解；康复计划的模板结构反映了这一区分。

写作指导。句子层唯一，篇章层多解，“找到自己的声音”无解；三档在同一门课里同时存在，而课程按周次而不是按档组织。

研究生指导。技术步骤唯一，选题多解，“成为一个研究者”无解；导师的实际做法是在第三档上用陪伴而不是指导，这在任何培养方案里都没有位置。

十处里有三处（飞行训练的科目分配、产品设计的原型策略、康复计划的模板）已经在实践中按倒推唯一性分了段，只是没有名字。

十五、强制不一致：教育目标分类学守不住的两条

教育目标分类这一支同时持有两条主张，都写在它自己的框架文件里。

甲：教育目标可以被放进一个分层的结构里，每一层都有对应的、可操作化的行为描述。这是这一支的全部工作产品：认知过程的层级、知识类型的维度、以及为每一格提供的动词表。它的实用价值正在于可操作化——写目标的人可以照着动词表把一条模糊的期望写成一条可评估的陈述。

乙：高层级的目标（分析、评价、创造）比低层级的更值得追求，教学应当向上移动。这条同样是这一支的核心主张，而且是它在教育实践里最有影响力的一条——几乎所有课程改革文件都引用它来论证“不能只考记忆”。

两条在一处正面相撞，而且相撞的位置恰好是这一支最看重的那一端。**层级越高，倒推唯一性越低。**“回忆某个定义”的下一步，五位教师会给同一个答案；“评价一个论证的可靠性”的下一步，五位教师给出的是活动名称——多读、多讨论、多写。**动词表提供的是目标的写法，不是下一步的算法；**把“能够评价”写成一条合规的目标陈述，与知道此刻这个学生该做哪一件具体的事，是两件不同的事，而这一支的整个装置只解决前者。

于是：甲声称可操作化，乙要求向高层移动，而**可操作化的程度随层级上升而下降。**这一支现行的处置是继续细化行为描述——把“评价”拆成“识别前提、判断相关性、检验充分性”等等。这个处置在形式上保住了两条，代价是：**拆到足够细以致下一步唯一的时候，被拆出来的那些子行为已经落回低层级了。**这不是执行不力，是拆解这个动作本身的性质——每一次拆细都在把一个多解问题换成若干个唯一问题，而“评价一个论证”这件事的难处恰恰在于**选择用哪一组子行为**，而那个选择正是被拆掉的那一步。

二选一。要么承认高层级目标在本章意义上是多解或无解的，因而不能用同一套操作化装置处理——那么甲不能以现在的形式成立，动词表必须按倒推唯一性分区，高层格子里给出的不该是动词而该是**选择的判据**；要么放弃乙，承认可教的只有那些能被操作化到唯一的東西——而这与这一支存在的全部理由冲突。

它不能同时守住两条，除非它承认本章的变量。一旦承认，“高层级目标”这个说法就必须被拆成两样：可拆到唯一的（照现有装置处理）与拆到唯一就变质的（需要另一套方法），而区分二者需要的正是倒推唯一性这个读数。

这一支已经为此付出了可观察的代价：**大量课程大纲里的高层级目标，其配套的评价方式与低层级目标的评价方式完全相同。**目标写着“能够评价”，考的是“能够识别下列哪一项是前提”。这一点在任何一次课程审核里都不会被标出来，因为审核的判据是目标与评价是否对齐——而它们确实对齐了，对齐在低层级上。

十六、三处反向约束

心理学割掉“唯一区就该一直用倒推”。刻意练习在唯一区效应最强，但同一批研究也表明高度专门化的训练会使迁移变窄，并在条件变化时表现下降。因此唯一区的倒推有

时间上限，且上限之后必须主动进入多解区。如果没有这一处，本章原本会写下“在唯一区应当尽可能延长倒推式训练”——这句话被削掉了。

物理学削掉“多解是缺陷”。在控制论与力学里，从一点到另一点存在多条可行路径不是问题，是可控性强的表现；解的不唯一意味着可选择的余地大。因此本章必须放弃“多解需要被收敛成唯一”这条价值排序，改为“多解需要被看见并被选择”。**这一处动的是价值排序，而且它改变了完整次序第二步的处方形式——从收窄改为展开。**

工程学削掉“无解区可以靠稳健设计兜住”。稳健设计要求性能可测、扰动可枚举；在目标说不清是什么的区间里，“对扰动不敏感”没有定义。因此本章的适用范围收窄：无解区没有现成的方法，只有改写目标这一条路，而改写目标不总是能成功。**这一处动的是适用范围，也是本章最不愿意承认的一处。**

十七、三家为何各自到不了这一条

三家的材料都在，判断却没有出现。不是谁看漏了一眼，是每一家的问题形状恰好让这一步成为多余的。

物理学到不了，因为它的两端总有一端是白送的。在力学问题里，初始条件通常是给定的——你知道这颗行星现在在哪儿、速度多少。因此“该从哪一端开始”在物理里几乎从不成为一个真问题：有初值就用初值，需要边值时（弹道、最优控制）它清楚地知道自己在解一个更难的问题，并且有一整套数值方法来对付。**适定性对物理学来说是一个计算属性，不是一个决策属性**，因为决策早就被问题本身规定了。而课堂的两端都不是给定的：学生当前的状态需要推断，目标只是一个方向。物理学没有遇到过两端都要自己造的情形，因此它的框架里没有“该不该先定终点”这个位置。

心理学到不了，因为它的整个方法论建立在“专家能说出下一步”上。刻意练习的操作定义里就写着任务由教练根据当前表现指定。这不是一条可选设置，是这一支能够做研究的前提——没有教练指定的任务，就没有“结构化训练时数”这个自变量。因此当某个领域的专家说不出下一步时，这一支的读法只有一种：**这个领域的训练还不够结构化。**它把“下一步说不出来”当作有待改进的现状，而不是当作一个可以测量、可以分档、可以据此换方法的变量。这不是疏忽——把它当变量，就等于承认自己的方法在一大类内容上不适用，而这一支恰恰是靠“它适用范围比想象的宽”立起来的。

工程学到不了，因为稳健设计的第一步就把这个问题绕过去了。稳健设计不追踪路径，它选工作区。你问它“从这儿到那儿下一步该做什么”，它的回答是“别走那条线，把参数放在这片平坦区里”。这个回答很有效，但它成立的前提是**目标可测、扰动可枚举**——也就是说，它默认自己站在本章的唯一区或多解区，从来不处理无解区。工程学因此有一套极强的方法来对付“路难走”，而没有任何方法来对付“说不出下一步是什么”，因为在它的世界里，说不出下一步意味着规格没写清楚，而规格没写清楚是上游的事故，不是工程问题。

三处缺口合起来说明这个判断为什么必须由三家撞出来：物理学有适定性这个概念而无两端皆需自造的处境，心理学有逐步推进的完整方法而必须假定下一步存在，工程学有对付难走的整套手段而把说不出下一步排除在自己的问题域外。少任何一家，剩下两家都推不到这一条。

十八、与既有说法划界

最近发展区。它说到的是：存在一个区域，学生独立做不到而在帮助下能做到，教学应当落在那里。分离线：它刻画的是**难度的位置**，本章刻画的是**倒推答案的个数**；两者正交——一个难度恰当的任务完全可以是无解区的（“写一篇比现在好一点的文章”难度合适而下一步说不出），一个难度过低的任务也可以是唯一区的。判定形式：若把任务难度控制住之后倒推唯一性不再预测结果，则最近发展区覆盖本章；若在难度匹配的任务之间倒推唯一性仍有区分力，则本章对。

刻意练习。它说到的是：针对当前弱点的、有即时反馈的结构化重复。分离线：它默认下一步存在且由教练给出，因此“下一步有几个答案”在它的框架里不是变量。判定形式：若在多解区强行实施刻意练习（教练指定唯一路径）与允许多路径的对照组结果无差异，则本章的分段多余；若前者在原场表现更好而迁移更差，则本章对。**这是本章最直接的检验。**

逆向设计与掌握学习。它们说到的是：从目标倒推评估与活动，逐级过关。分离线：它们把倒推的可解性当作给定，整个方法建立在“目标定清楚就能倒推”上。判定形式：若一门目标清晰度极高而倒推唯一性为零的课，其效果不低于目标清晰度中等而倒推唯一的课，则本章的正交主张为伪。本章预测前者更差，而这在辨别格一节里是左下与右上的对照。

形成性评价。它说到的是：用证据调整教学。分离线：调整的前提是知道调向哪儿；在无解区里证据再充分也无处调。判定形式：形成性评价的效应量应当随倒推唯一性上升而上升；若两者无关，本章的分段对它不增量。

稳健设计（工程学，本章学科之外）。它说到的是：选对扰动不敏感的工作区。分离线：它是关于**参数选择的**，本章是关于**求解结构的**；一个稳健的工作点仍然要有人说得出怎么走到那里。判定形式：若把“训练是否针对稳健区”控制住之后倒推唯一性不再预测结果，则本章被覆盖。

可控性与可达性（控制理论，方法学占位者）。把本章的命题剥掉一切教育名词，剩下的形式是“从当前状态能不能到达目标状态、有几条路”——这正是可控性与可达性的形状，它在控制理论里被形式化了六十多年。这是本章最危险的一位占位者，必须正面回答。分离线：可控性是**系统的性质**——给定动力学与输入集合，可达集是客观确定的；本章的倒推唯一性是**知识分布的性质**——同一个学生、同一个目标，换一批更懂行的教师，答案的个数会变。一个是关于世界的，一个是关于谁在看。判定形式：若倒推唯一性只随学生状态与目标变化、而不随评判者群体变化，则可控性覆盖本章，本章只是它的一个通俗说法；若同一份学生样本与同一个目标交给两个不同的专家群体得到不同的唯一性档位，则本章指出的是一个可控性里没有的变量。**这一条可以在一周内实测，而且本章赌后者。**

问题空间与手段——目的分析（认知科学，方法学占位者）。它说到的是：解题是在状态空间里搜索，用差距缩小指导选择。分离线：它默认目标状态可以被代入以计算差距；无解区的定义正是差距无法计算。判定形式：若人在无解区仍能报告出稳定的“差距感”并据此选择动作，则手段——目的分析覆盖本章的无解档；若报告出的只是活动偏好而非差距，则本章对。

十九、与作者同题诸文的划界

作者另有十余篇同样处理课堂学习的文章，其中三篇已收入本书第三编；本章必须与它们划清。总的分界在问法上：那些文章追一个驱动，问**为什么很少发生**；本章追一条次序，问**该怎么运行**。以下只列会被混为一谈的几篇。

与第九章《还在动的时候就读了数》：同一组学科的两两不同取用

本书第三编里有一章由同样这三门学科撞成，处理同一个问题。距离最近，必须写在最前面并写细。

那一章处理的是**判定的时点**：课堂每隔一两周作一次“会了没有”的判定，而判定的时点由课表决定、不由被判定的那件事的重组时长决定；它主张过早的判定把仍在进行的变化冻在当时的构型上，因而命名了冻结构型，并给出双向翻转率这个读数。

先说两家共用学科的取用点：同一门物理学被两章拿走的是两件不同的东西。那一章取的是**玻璃态的物理老化**——被判为冻住的系统，其性质随冻住之后所等的时间持续演化，且演化的起点正是被冻住那一刻的构型；本章取的是**常微分方程边值问题与初值问题的适定性差异**，以及变分表述本身是一个边值表述这件事。前者属于凝聚态与非平衡统计物理，后者属于经典力学与数值分析，两个分支互不引用。工程学同理：那一章取传感回路与可观测性，本章取稳健参数设计。心理学那一家则确有交叠——那一章取提取诱发遗忘，本章取刻意练习，同属记忆与技能获得这一大类；本章承认这一处的距离比另两家近。

再说承重命题的分离。那一章问的是**什么时候读**——判定的时点相对于弛豫时长是早是晚；本章问的是**读出来的东西有几个答案**——从当前状态到目标，下一步该做什么是唯一、多解还是无解。一个是时间轴上的问题，一个是解空间上的问题。二者可以完全分离：一个重组时长极短、判定时点也恰当的单元，其倒推可以处在无解档；一个倒推唯一、方法匹配的单元，其判定也可以早得把学生冻在半路。

双向判定形式。目标写着“能够评价一个论证”、五位教师都说不出下一步的那个单元，若把判定推迟四周并重测，那一章预测双向翻转率里“不会到会”那一支显著非零、判定过早是病因；本章预测延迟迁移无改善，因为推迟一个没有方向的过程不会给它方向。反过来，一个换把动作的训练单元，倒推高度唯一而教师在第三天就作了判定，本章判此处正该用倒推、无异议；那一章判判定过早。**两章在这两格上给出不同的诊断与不同的处方**，且两个处方不可互相替代——推迟判定不会让说不出的下一步变得说得出口，改写目标也不会让重组时长变短。

两章合用时给出一条各自单独说不出的预测：在无解档的单元里，双向翻转率中“不会到会”的那一支应当最高。理由是那里没有人在指定下一步，因而变化只能自发进行，而自发进行的变化恰恰是被过早判定冻住的那一类。**这条预测两章单独都推不出来**：那一章没有档位这个变量，本章没有时点这个变量。若实测发现两个读数的相关系数高于零点六，两章必有一篇是另一篇的代理，届时应保留成本更低的那一个——也就是那一章的双向翻转率，它不需要分母。

与《一条路没有交接点》。书外另有一篇同样以“该怎么运行”立题、同样把三家写成三条不同终点的文章，而且它与本章共用物理学与工程学两家。距离最近，必须写在最前面并写细。

先说两家共用学科的取用点：**同一门学科被两篇拿走的是两件不同的东西**。那篇的物理学取的是**盲分析的时点**——判据必须在读出结果之前锁定，历史依据是同一个物理量的公布值曾一段一段朝前一次靠拢；本章的物理学取的是**初值问题与边值问题的适定性差异**——一个在弱条件下解存在唯一，一个可无解可多解，依据是数值方法的标准分类。两处在物理学内部属于不同的分支（前者是实验方法学，后者是常微分方程理论），互不引用，也不能互相推出。工程学同理：那篇取**写进标准的三项独立性**，本章取**稳健设计的最优点脆弱性**。

再说承重命题的分离：那篇处理的是**谁有资格宣布这一段成了**，本章处理的是**要到那儿下一步该做什么有几个答案**。一个是权力问题，一个是适定性问题。**双向判定形式**：一门由完全独立的第三方按事先锁定的判据判定达标的技能课，其倒推唯一性可以极高（那篇读数满分、本章判此处正该用倒推，两篇都判健康，但理由不同且可分开）；而一门达标判定全由授课方自按自判、同时倒推处于无解档的课，那篇诊断为结算权集中，本章诊断为目标未被改写——**两条诊断给出的处方不同且不可互相替代**：交出判定权不会让说不出的下一步变得说得出口，改写目标也不会让自我结算产生信息。若两个读数的相关系数高于零点六，**两篇必有一篇是另一篇的代理**，届时应保留可从既有档案直接算出的那一个，也就是那篇。本章预测两者近乎无关。

与《方程两边都有它》（分岔步·可逆性）。那篇的承重是分岔步的归属。分离线：分岔步的定义要求同一产物有两条路径——**这正是本章的多解区**。因此两篇在这一格上是嵌套关系而非并列，本章诚实写明：**那篇处理的现象只在本章的多解区里存在**，在唯一区没有分岔步可言。本章的增量是把它放进一个三档结构里，并指出唯一区与无解区各需要不同的处理；那篇的增量是在多解区内部给出了“哪一步不能代做”的判据。两篇合用比单用任一篇都强。

与《先动的总是同一边》（节拍比）。分离线：节拍讲结算的频率，本章讲下一步的解的个数。判定：把节拍调到最慢（那篇的处方）而目标仍处于无解档，本章预测无改善——慢下来不会让说不出的东西变得说得出口。

与《总有人替它起步》（结清态·未被归类窗口）。那篇处理产物在成熟之前被归类。分离线：归类是对已有产物的处理，倒推唯一性是对下一步的判定；一个把每次产物都长时间悬置不归类的班，其倒推档位不因此改变。判定：一门无解档的课延长未被归类窗口，那篇预测续发率上升，本章预测迁移无改善——因为悬置的是没有方向的活动。**这是两篇的判决性对照，可同批测量**。

与《差一条边》（接手图·闭合度）。分离线：那篇的对象是人与人之间的接手关系，本章的对象是下一步的解的个数。**两处合起来给出一条本章单独说不出的话**：多解区之所以需要“把多条路同时可见”，前提是这个班的接手图能把不同的路传开；图不闭合时，多解区的正确处方也执行不了。

与《正好有人能解决》（清除位·透过率）。分离线：那篇的对象是被消耗的差异，本章的对象是下一步的适定性。判定：一门倒推唯一性极高、教师即时处置能力极强的技能课，那篇预测透过率趋零并判为问题，本章判此处正该用倒推、透过率低不是问题——**两篇在这一格给相反评价**，而这个分歧是可以被检验的：若这类课的学习者在条件变化后表现良好，本章对；若他们在原场满分而换一个条件就垮，那篇对。

二十、与本编另两章的划界

与第四章《手段取得终点资格的那一段》（教育×化学×艺术）。那一章处理的是终点易主这个事件，读数是易主次数与触发源；本章处理的是倒推这个动作在什么条件下有解，读数是下一步答案的个数。两者数学上独立：倒推唯一的区间里可以不换位（那正是那一章的靶格），倒推无解的区间里可以频繁换位而毫无结果（那一章的漂移格）。一条共同的推论：那一章说终点与手段的角色由信息分布决定，本章说倒推的可解性由知识分布决定——两条都指向同一个更一般的东西，而三篇都没有把它写实，这是三篇共同交出的洞。

与第六章《这一节课结束时，这个班手里多了什么》（细胞×工程×艺术）。那一章处理的是学习过程自己造出来的条件，读数是存量。分离线：自建条件是产物，倒推唯一性是约束。判决性对照：本章预测自建条件在多解区最容易积累（因为多条路都要留痕才选得出来），在唯一区最少；若实测发现自建条件存量与倒推唯一性正相关，则本章这条预测为伪，而那一章不受影响。

二十一、五种最强的竞争解释，与排除它们的设计

竞争解释一：其实不就是内容难度吗。难的内容说不出下一步。排除设计：用独立的题目难度评定测量难度。本章独有的变量与它正交，而且反例在两个方向上都有——一个难度极低的目标可以是无解的（“喜欢上数学”），一个难度极高的技巧动作的下一步专家高度一致。本章获胜的观察是：在难度匹配的单元之间，倒推唯一性仍有区分力。

竞争解释二：其实不就是教师经验吗。经验足的教师说得下一步。排除设计：控制年限与资格。这一条比第一条更难，因为经验确实会提高唯一性判定的准确率。本章独有的观察在多解区：经验最丰富的五位教师在多解区仍然给出五个不同的答案，而且他们会互相承认对方的答案可行——这与无解区的情形（互相认为对方给的不是下一步）可以分开，而经验解释预测不出这个差别。

竞争解释三：其实不就是目标写得清不清楚吗。排除设计：用现成的目标清晰度量表独立评分。本章的正交主张在这里可以直接检验——本章预测存在大量目标清晰度满分而倒推无解的单元（“能够评价一个论证”），也存在目标清晰度中等而倒推唯一的单元（学徒制、乐器教学）。若清晰度与唯一性的相关系数高于零点七，本章的第二根轴不成立。

竞争解释四：其实不就是这个学科成熟不成熟吗。成熟学科有细致的子技能分解，因此下一步说得出。排除设计：在同一学科内部比较不同单元。本章获胜的观察是：同一门课的不同区间落在不同档位——写作课的句子层唯一、篇章层多解、“自己的声音”无解，三者在同一学科同一门课同一位教师。若档位只随学科变化而不随单元变化，本章的分段主张失效，退化为一条关于学科成熟度的分类。

竞争解释五：其实不就是评判者不认真吗。说不出下一步是因为没细想。排除设计：给评判者充足时间并要求写出理由，再看档位是否变化。本章的预测是无解区的答案在给足时间后仍然是活动名称，且这些答案能通过“换一个目标是否照样说得通”这一关的比

例极低。若给足时间后无解区大量转为唯一区，则本章测的是评判者的投入而不是知识分布。

判决可以压缩成一句：若在同一课程、同一学科、控制难度、教师经验与目标清晰度之后，倒推唯一性与运行方式之间仍存在交互效应（唯一区倒推优于推进、多解区方向相反），且同一份材料在不同专家群体间的档位不同，则本机制优于上述五条；若效应被难度、经验或清晰度完全吸收，或档位跨群体一致，则竞争解释获胜。

二十二、分层引用结构

本章的主张分三层，各自的证伪条件不同，可以分层引用。

第一层：可倒推区间作为一个存在物，且它是知识分布的性质而非系统性质。 最强也最容易倒。它的关键证伪条件在证伪条件其二：若同一份材料交给两个不同的专家群体得到相同的档位，这一层就被可控性覆盖，本章只是一个通俗说法。

第二层：按档分段的运行方式作为一个分析工具。 即使第一层的“知识分布”主张被推翻，三档与三种运行方式的匹配仍然可用——那时它退化为一条关于内容性质的分类，仍能指导教学安排。这一层的证伪条件是交互效应：若在唯一区用倒推与在多解区用倒推的效果差异不显著，这一层也倒。

第三层：一条不依赖前两层的经验主张。 目标清晰度极高、而五位教师说不出下一步的那些课程单元，其学生的延迟迁移不优于目标清晰度中等、而下一步可说的单元。这一条不需要“可倒推区间”这个概念，只需要一次教研活动的问卷与一次延迟测验，而且它是三层里最便宜的一条。

三层里第三层最便宜、第一层最险，而本章押的正是第一层。第三层成立而第一层倒，说明现象是真的而本章给它的刻画错了；第三层一倒，上面两层无处可站。

二十三、证伪与赌注

最强的竞争解释是：其实不就是内容难度和教师经验吗——难的内容说不出下一步，经验足的教师说得出。

本章必须能排除它。难度可以用独立的题目难度评定测量，教师经验可以用年限与专业资格测量。本章独有的变量是**答案的个数**，它与两者都正交：一个难度很低的目标可以是无解的（“喜欢上数学”），一个难度极高的目标可以是唯一的（一个高难度技巧动作的下一步，专家高度一致）；而经验最丰富的五位教师在多解区仍然会给出五个不同的答案——多解不是无知，是路径确实多，这一点可以通过让专家互相评价对方的答案来区分：多解区里五个人会认为彼此的答案都可行，无解区里五个人会认为彼此的答案都不是下一步。

核心证伪条件：在同一门课程、同一批教师、控制内容难度与教师经验之后，若倒推唯一性与“按档匹配的运行方式”之间不存在交互效应——即在唯一区用倒推与在多解区用倒推的效果差异不显著——则本章的分段主张为伪；届时应把课堂运行的差别归于难度与经验，而非倒推的适定性。

其余七条：

其二，若同一份学生样本与同一目标交给两个不同专家群体得到相同的唯一性档位（跨群体一致），则倒推唯一性是系统性质而非知识分布性质，本章被可控性覆盖，那条分离线失败。

其三，若在多解区中专家互评“彼此的答案都可行”的比例与无解区无差异，则多解与无解不可区分，三档应压缩为两档。

其四，正向读数（顺序）：区间边界的移动应当早于表现变化出现——一个学生的某个区间从多解收敛为唯一，应当先于他在该内容上的成绩提升。若表现变化总是先于边界移动，则边界移动只是结果不是先兆，完整次序第三步失效。

其五，若无解区的目标经改写后唯一性没有变化，则“改写目标”这条处方无效，完整次序第二步的无解档没有可执行内容。

其六，最便宜的一条：**任何一所学校的教研组，用一次教研活动就能测一门课六个单元的倒推唯一性——五份匿名学生样本、五位教师、各十分钟。**所需的一切在多数机构的既有安排里已经存在，不需要新增数据采集。

其七，若不同专家群体规模（三人、五人、九人）给出的档位不稳定，则该读数不可靠，应先解决测量学问题再谈机制。

其八，若在形式部分列出的三个已按唯一性分段的行业（飞行训练、产品原型、康复计划）里，这种分段与结果无关，则本章的兑现是同形而非真兑现。

写死日期的赌注：在二〇三〇年十二月三十一日前，至少两项在同一制度环境内、单元数不少于三十六、预注册的研究，报告倒推唯一性与运行方式之间存在交互效应：在唯一区，倒推式运行优于推进式；在多解区，方向相反。

以下不算命中：只报教师对课程难度的评分；只比较讲授与探究；把“目标清晰度”当作倒推唯一性；用两所学校或两个国家的对比充当交互检验；只报学生满意度或当堂正确率；机构自行说明已按学情分层。

二十四、反噬与治理

倒推唯一性一旦进入考核，最省事的实现方式是把所有目标都改写成唯一区的形式——把“能够评价一个论证”改写成“能够指出论证中的三种常见谬误”。这个改写能把读数做满，而且它在形式上完全符合本章完整次序第三步的建议。问题在于：**改写目标与掉换目标之间没有可靠的分界。**前者是把同一个东西写成可操作的形式，后者是换一个更容易的东西来做，而两者在文档上看起来一模一样。

第二重反噬：无解区一旦被命名，它会成为一个免责的类目。教师可以把任何教不动的东西归入无解区并宣布“这在当前知识条件下不可教”，而本章恰恰给了这句话一个体面的说法。

三条治理底线：倒推唯一性指向内容与知识分布的关系，不指向教师或学生的能力，不得单独用于排名或教师评价；不得为提高读数而在安全关键的训练中把整体判断类目标一律改写成程序性目标；任何据此作出的不利安排——包括判定某目标不可教——必须公开评判者名单、原始答案与复核通道。

我看不到一个彻底的出口。只要这个读数被采用，改写目标就永远比改进教学省事，而改写与掉换的分界只能靠人来判。可用的唯一限制是要求改写前后的目标在一个独立的延迟测验上可比——而这条限制本身也要花钱。

二十五、若本判断被证伪，我们会学到什么

若倒推唯一性与运行方式之间不存在交互效应，我们学到的不是“方法都一样”，而是一件更具体的事：**倒推与推进的差别在教学上不承重，承重的是投入与反馈的质量。**那意味着刻意练习那一支关于“结构化程度”的主张可以推广到它自己划的边界之外，而本章所说的多解区与无解区，只是结构化尚未做到的地方。这会把研究的下一步引向“如何把高层级目标结构化”，而不是“如何为它换一种运行方式”。

若档位跨专家群体一致，我们学到的是：可倒推性是系统的性质而不是知识分布的性质，本章只是可控性的一个通俗说法。那不是一无所获——它会把教学问题正式接到控制理论上，而可达集、可控性格拉姆矩阵这些工具随之可用，这比本章能给的多。

若无解区的目标经改写后唯一性不变，我们学到的是：无解不是目标表述的问题，是这个目标在当前的知识条件下真的没有可教的路径。那意味着一批被广泛写进培养方案的目标应当被明确标注为“不可直接教授”，而这是一个有实际后果的结论——它会改变课程时间的分配方式。

三种失败方式各自指向不同的下一步。

本章交出三个自己解释不了的洞。第一，多解与无解的分界依赖专家互评，而互评本身需要专家在同一个语汇里作答；当两个群体连“下一步”的意思都不一样时，这个分界无法测。第二，本章说无解区必须改写目标，却没有给出改写的方法——那需要另一条完整的次序。第三，本章完全没有处理**档位随学生个体变化**的情形：同一个目标对一个学生是唯一区、对另一个学生是多解区，这在一个班里同时发生，而本章的三档运行方式是按内容分段的，无法按人分段。

二十六、自反检验

按本章自己的判据，这篇文章处在多解区。它交出的是一个读数与一条按档分段的次序，而“读了它之后下一步该做什么”这个问题，五位读者会给出五个不同的答案：先去测一门课、先去反驳强制不一致一节那条与可控性的分离、先去改写一个无解目标、先去做证伪条件其六那个便宜检验、先什么都不做等人做完。

这五个答案都可行，这正是多解区的签名。按本章自己完整次序第二步的处方，本章此刻不应当指定唯一的下一步，而应当把这几条路同时摆出来——上面那句话就是在做这件事。

本章交出两个自己解释不了的洞。第一，多解区与无解区的分界依赖专家互评，而互评本身需要专家在同一个语汇里作答；当两个专家群体连“下一步”的意思都不一样时，这个分界无法测。第二，本章说无解区必须改写目标，却没有给出改写的方法——那需要另一条完整的次序，本章没有它。

二十七、结论

课堂如何运行才能让学习真正发生，三门学科给了三条不能并存的次序：定住两端整体求解、一步一步局部改进、放弃最优只求平坦。三家共同假定“下一步该做什么”这个问题有答案，分歧只在怎么求。

物理学自己拆掉了这个假定。初值问题在很弱的条件下解存在唯一，两点边值问题可以无解、可以有无穷多解——**倒推与推进不是同一个动作的两个方向，它们的适定性根本不同**。这条差别在物理里回答的是数值方法怎么选，从没有人把它当作“该不该先定终点”的证据。

由此得到一条三家单独都推不出来的判断：**课堂运行的次序必须按倒推的解的个数分段，而不能按内容、按课时或按目标层级分段**。唯一区用倒推，多解区用推进加选择，无解区必须先改写目标——三种方式互不通用，用错了不会报错，只会在形式上一切合规而什么都没发生。

这句话的可操作形式是一句问话，一次教研活动就能测：把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的下一步是同一件事、五件不同的事、还是都说不出来？

如果答案是第三种，那么这门课的目标可能写得非常漂亮，与评价对得非常齐，而它的每一节课都在一个没有解的问题上运行。

接续。本章的判据是全书最便宜的一个，附录一把它压成一次教研活动的操作步骤，含“五人互评”那一步——多解与无解在答案分布上看起来一样，只有互评能把它们分开。“教师经验会系统性高估唯一性”这一条，在第十二章被用作“教师不应担任编码者”的理由。错配代价不对称能否与另两条推导收成一个式子，见第十三章。

第六章 这一节课结束时，这个班手里多了什么

判断卡片：条件有一部分是这个过程自己造的；而在高覆盖度区间，外来条件每完备一分，这一部分就被压掉一分—— $c \rightarrow 1$ 时，一门完备的课程在结构上不可能有它。

三个理论入口：细胞学 $\times$ 工程学 $\times$ 艺术学

提要。课堂如何运行才能让学习真正发生？三门学科给出三条完整而互不相容的次序。艺术学的作品自足传统说，只有做出来的那个东西本身算数——过程怎么样、条件怎么样都不进入判断，批评面对的是作品。工程学的过程能力传统说反了，而且说得极重：质量不是检出来的，是造出来的；只盯成品是最贵的一种错误，因为等你检出来的时候，产生它的那套过程已经又造了一批。细胞学的生态位传统说两家都把决定性的东西放错了地方：细胞的命运不写在细胞里，写在它周围——把干细胞放进对的微环境，它自己就分化了；换一个微环境，同样的基因组给出完全不同的东西。三条不能同真——艺术学每天在做的那件事，正是工程学诊断为最贵错误的那件事。三家共同站着而从未说出的那块地是：**条件是给定的输入**。艺术学把条件排除在判断之外，工程学把条件当作要控制的自变量，细胞学把条件当作决定性的自变量——三家对条件的地位判断完全不同，而三家都默认它是**先在的、外来的、被提供的**。推翻它的材料来自细胞学自己。干细胞并不只是被生态位决定：它们分泌细胞外基质、招募支持细胞、重塑自己周围的微环境；而类器官更进一步——给一小撮细胞很少的初始条件，它们自己长出结构，并且长出维持这个结构所需的那些条件。**生态位不是住户搬进去的容器，是住户造出来的东西**。这条在细胞学里回答的是“体外培养怎样才能成功”，是一个方法问题，从没有人把它当作“条件与产物的角色可以互换”的证据。由此本章命名**自建条件**：一类由学习过程自己产生、随后成为这个过程的条件而被继续使用的东西。它不是知识，不是能力，不是成品——它是**工具**：一个记号、一条这个班都认的判准、一份可查的错误清单、一件自己做的量具。三家的结算规则都只认成品、过程或环境，因此都不为它设字段。判据一句：**这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的？**读数**自建条件存量**可以逐件清点，且与既有指标正交——一门教材、讲义、题库、评分表、仪器规程全部齐备且严格执行的课，自建条件存量恰恰是零。

一、三条次序，与它们对“条件”的三种态度

“课堂如何运行”要的是一条次序。三门学科各有一条，而且三条对同一样东西——条件——给出三种互不相容的安排。

艺术学的作品自足传统给的次序是：做，然后看做出来的东西。判断只面对成品，因为过程无法被外人重走，条件是每个人自己的事。这条次序的力量在于它拒绝了一切用意图、努力与环境为结果开脱的说法。

工程学的过程能力传统给的次序相反，而且它对第一条的诊断非常重：**质量不能靠检验加进产品里，必须在制造过程中被造进去**。只盯成品意味着你永远在事后发现问题，而在你发现的时候，产生它的那套过程已经又造了一批同样的东西。因此正确的次序是：先把过程稳住，让它有能力持续产出合格品，成品检验只是确认。

细胞学的生态位传统给第三条，并把前两家一起否掉：决定性的东西既不在成品里也不在过程里，在**周围**。同一个基因组，放在不同的微环境里，长出完全不同的细胞。要让某个东西长出来，你要做的不是雕刻它、也不是控制工序，是把它周围的条件建对，然后让它自己长。

三条互不相容。工程学要控制的过程，在艺术学看来正是不可外化的那部分；艺术学只看成品，在工程学看来是最贵的一种错误；而细胞学说两家都在错的层面上使劲——一个雕成品，一个管工序，而决定性的是那个谁也没在管的周围。

二、艺术学这一家：只有做出来的那个东西算数

把第一家推到最强：判断只能面对作品本身。

这条主张有它的道理，而且是硬道理。作者的意图无法核实；过程无法被第三方重走；工作条件千差万别。凡是把这些引进判断的做法，最后都会变成“他很努力”或者“他条件不好”。工作室里的讲评之所以有效，正是因为它把作品摆在那里，所有人看同一个东西。

它对课堂的主张是：让学生做出东西来，然后把东西摆出来一起看。运行的次序是做——看——再做。

这一家自己知道的一处麻烦，在讲评里天天出现：**两个学生做出了同样水平的东西，一个下次还能做出来，一个不能，而作品本身看不出这个差别**。这一家的处理办法是延长时间、多看几件；也就是说，它承认单件作品不够，但它的补救仍然是看更多的作品。

三、工程学这一家：质量是造出来的，不是检出来的

把第二家推到最强：产出的质量由过程的能力决定，成品检验只能筛掉不合格品，不能产生合格品。

这条主张在制造业里已经制度化了几十年。它的核心动作是把注意力从成品移到过程：测量过程本身的稳定性与能力，在过程发生偏移时就介入，而不是等成品出来。它的一句名言是所有相关培训的第一课——**检验不产生质量，质量在制造过程中就已经决定了**。

它对课堂的主张是：不要靠考试来产生质量。考试只能告诉你已经发生了什么。要运行的是过程：让每一步都可观察、可测、可在偏移时介入。

它对艺术学那条次序的诊断是最直接的一条：只看作品等于只做终检，而终检是所有质量手段里最贵、最慢、信息量最低的一种。

这一家自己承认的边界也清楚：过程控制要求过程可重复、输出可测量。当每一次的输出本来就应当不同的时候——这恰恰是艺术与大多数高阶学习的情形——“过程能力”这个量没有定义。

四、细胞学这一家：命运不写在里面，写在周围

把第三家推到最强：一个细胞会变成什么，主要不由它自己决定，由它所处的微环境决定。

这条主张在干细胞生物学里有极强的证据。同样的干细胞，换一套支持细胞、换一套基质硬度、换一组可溶因子，得到的是完全不同的谱系。生态位这个概念之所以立得住，正是因为把细胞搬离原位它就变了，搬回去它又回来了。

它对课堂的主张因此很彻底：不要雕学生，也不要管工序，去建环境。给对的同伴、对的材料、对的时间尺度、对的失败成本，学习会自己发生。

它对另外两家的诊断合成一句：**你们一个在雕成品，一个在管工序，而决定成品是什么的那样东西，你们都当成了背景。**

五、三家共同站着的那块地

三家争的是**条件在判断里该占什么位置**：排除在外、作为要控制的自变量、还是作为决定性的自变量。

争得再凶，三家共享一个假定：**条件是给定的输入**。艺术学把它当作不可控因而不该进入判断的背景；工程学把它当作要稳住的输入；细胞学把它当作要建对的输入。三家对它的重要性判断从零到全部，而三家都默认它**先在于过程、外来于过程、由别人提供**。

把这个前提念给三家听，三家都会说这还用说吗。条件当然是先有的，不然过程在哪儿发生。这正是共有前提的签名。

六、推翻它的材料，来自细胞学自己

细胞学自己知道生态位不是给定的。

细胞造它自己的生态位。干细胞与它周围的细胞分泌细胞外基质，改变基质的成分与硬度；它们招募或诱导支持细胞；它们改变局部的因子浓度、氧张力与力学环境。这不是边缘现象，它是生态位维持自身的常规机制——把细胞拿走，那个生态位会退化。

类器官把这一条推到了极端。从很少的细胞出发，只给一个基础的支持环境，它们自己长出分层的结构、自己分化出多种细胞类型、并且**自己产生维持这个结构所需要的那些局部条件**。整个领域的方法论共识是：你不能把类器官造出来，你只能把初始条件给对，然后让它自己长——而它长出来之后所处的那个环境，与你一开始给的那个已经不是同一个了。

这就把共有前提拆了。条件不是先在的输入，它有一部分是**过程自己的产物**，而且这一部分随后成为过程继续下去的条件。**产物与条件的角色在同一个系统里互换了，而且互换是这个系统能长成它自己的必要条件**。

细胞学在自己的领域里用这条回答的是“体外培养怎样才能成功”——为什么某些细胞离开原位就死，为什么类器官必须自组织而不能被拼装。这是一个方法问题。**没有人把它当作“条件与产物的角色可互换”的证据——而它恰恰是。**

七、命名：自建条件

由此本章命名一类东西。

自建条件：由一个学习过程自己产生、随后被这个过程继续拿来当作条件使用的东西。四个条件，缺一不可。

它是由这个过程自己造的。教师带来的、教材提供的、平台内置的，都不是。判据是可追溯的来源：这件东西第一次出现在哪一次课、由谁、为解决什么而做出来的。

它随后被当作条件用。造出来放在那里没人再碰的东西不算。它必须在后续的工作里被反复调用，成为大家不再讨论就直接使用的东西。

它不是知识，是可以被指着看的东西。一个记号、一张这个班自己整理的错误清单、一条大家都认的判准、一件自己做的量具或模板、一个共享的命名约定。“我们对这个概念理解更深了”不算——那是知识，不是条件。判据很实：它能不能被交给一个新来的人，让他直接用？

它在三家的账上都不存在。作品自足只记成品；过程能力只记过程参数；生态位研究只记环境变量——而环境变量在它们的实验设计里全部是自变量，也就是研究者设定的。这一类东西在三本账上都不是一样东西，不是记得少，是不设字段。

需要说清一件容易混的事：自建条件不是“学生的作品”。一份作业是成品，一份被这个班反复拿来对照的评分范例是自建条件；两者可以是同一份纸，区别在于它有没有被转成后续工作的条件。同一样东西的身份由它之后的用法决定，不由它被做出来时的样子决定。

八、判据：一句可以拿去清点的话

这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的？

不含情态词，不问学生学到了什么，只问多了什么可用的东西。它可以拿去问课堂的实物、共享文件夹、板书照片与作业模板，不必问任何人的感受。

五个场景跑一遍，答案互不相同：

讲授、教材、统一测验的课。零。所有可用的东西——讲义、题库、评分标准、进度表——都是外来的，且课程结束时一件不多。这是最常见的情形，而且它在任何检查中都完全合格。

按讲义走的实验课。接近零。仪器、规程、记录表格、误差处理方式全部外来；学生做完实验，手里多了一份数据和一份报告，两样都是成品，下一次不会被拿来当条件用。

一个班自己形成的代码约定。有，且可清点。命名规则、目录结构、提交信息的写法，一份“我们踩过的坑”清单——这些东西第一次出现有明确的来源，之后被所有人直接使用，新来的人拿到就能用。这一条是查出来的：真实工程团队的约定文档有版本历史，可以看到它是怎么从一次具体的争论里长出来的，而教学中没有对应的记录。

数学课上一个班自己发明的记号。有。教学研究里记录过学生群体自发产生并稳定使用的表示法，而这些记号在课程大纲里没有位置——它们不是教学内容，也不是学习成果，它们是这个班的工具。这一条也是查出来的。

组织工程与类器官实训。有，而且规程里明写。培养规程会写明细胞会改造基质、几天后的基质与初始基质不同、换液策略要据此调整——生物学的操作规程为“体系自己造出来的条件”设了字段，而教学规程没有。这一条是本章最直接的一处兑现。

后三个答案不能从命题推出来。它们说明这个读数在两个与课堂无关的地方已经有制度化的形式。

九、读数：存量、来源比与可交接性

自建条件存量。到今天为止，这门课在用的东西里，有多少件是这个班自己造的。逐件清点，每件写明第一次出现的时间与触发它的那件具体的事。

外来条件占比。在用条件的总件数里外来的比例。这个比例几乎在所有正规课程里都接近百分之百，而这一点从来没有被当作一个数看过。

可交接性。每一件自建条件能否被交给一个新来的人直接使用。这一条把真的自建条件与“我们班的默契”分开——后者无法交接，因此不是条件，是习惯。

正交性实例值得单写：一门教材、讲义、题库、评分表、仪器规程全部齐备且严格执行的课，自建条件存量恰恰是零。因为“全部齐备”的意思就是这个班不需要造任何东西，而“严格执行”的意思是造了也不能用。这说明存量不是“课程完备度”的另一种说法，它与完备度方向相反。

十、形式部分：存量的稳态，与它为什么被完备度按比例压掉

前面几节是机制描述。本节把它写成能推出正文没说过的东西的形式，否则“外来条件充分的课自建为零”只是一句观察。

设一门课在时刻 t 的可交接自建条件存量为 $A(t)$ 。它有产生也有失效。

产生。自建条件在什么时候被造出来？在一件事需要一个工具、而现有的工具不够用的时候。记这门课在单位时间里出现的“需要工具的场合”数为 N （它由内容的复杂度与工作量决定），记外来条件对这些场合的覆盖度为 $c \in [0,1]$ （教材、题库、评分表、规程、平台功能加起来覆盖了多大比例）。未被覆盖的缺口是 $(1-c)N$ 。这些缺口里有一部分会被真的做成可交接的工具，记转化率为 $\alpha \in [0,1]$ ，它由是否允许自造、是否有地方存放、是否有人会去用决定。

于是产生率为 $\alpha(1-c)N$ 。

失效。自建条件会失效：造它的那批人毕业、它所针对的那个问题不再出现、它被一次课程调整废掉。记单位时间的失效比例为 δ 。

于是

$$dA/dt = \alpha(1-c)N - \delta A$$

稳态存量

$$A^* = \alpha(1-c)N / \delta$$

这个式子推出四件正文前面没有说过的事。

其一，稳态在 α 、 N 、 δ 不随覆盖度变化这一条件下，存量与外来条件覆盖度线性反相关；放开这一条件后，高覆盖度区间的压制仍然必然，中低区间则可能上升、下降或呈倒U形，而且这是设计的必然结果，不是执行不力。每提高一分覆盖度，就按同一比例压掉一分自建。这不是说完备的教材有什么问题——它恰恰说明教材做得越好，这个量越低，而两者不可兼得的部分是结构性的。由此得到一条可检验的预测：在同一所学校、同一批教师、内容量相当的课程之间，自建条件存量应当与教学资源完备度呈负相关，且这个负

相关在控制了教师投入与学生水平之后仍然存在。若实测发现两者正相关（资源好的课自建也多），则本模型为伪，届时应把自建归于“整体投入水平”这一共同原因。

其二， $c \rightarrow 1$ 时 $A^* \rightarrow 0$ ：一门“完备”的课程在结构上不可能有自建条件。这不是程度问题，是极限。所谓完备，定义上就是每一个需要工具的场合都已经有了工具了。因此追求课程建设的完备本身就是把这个量归零的过程，而这个过程的每一步在任何评估体系里都记为进步。这一条把辨别格一节的左上格从“一种不好的状态”改写成“一个被正确的努力必然到达的位置”——它是尽职的产物，不是失职的产物。

其三，提高 α 与降低 c 的效果不同，而且只有前者可用。两者在式子里都能抬高 A^* ，但降低 c 意味着故意不提供教材、不给评分表——这是把学生扔进辨别格一节的“荒地”格，且反向约束一节的三处约束都反对它。只有 α 可以动：允许自造、给一个存放的地方、让它被下一次真的用上。这三件事都不减少任何外来资源。式子因此把本章的处方与“少给资源”彻底分开，而这个分开不是靠声明，是靠偏导—— $\partial A^* / \partial \alpha$ 与 $\partial A^* / \partial c$ 是两个独立的旋钮。

其四，一个反直觉的推论：短课程的稳态存量结构性地低于长课程，而且不是因为时间不够。 A^* 与 N 成正比，而 N 是单位时间的场合数，因此 A^* 本身与课程长度无关；但从零起步到接近 A^* 需要的时间为 $1/\delta$ 的量级。若一门课的长度短于 $1/\delta$ ，它永远达不到自己的稳态存量。这解释了一件通常被归因于“短课程内容浅”的事：模块化、短周期、频繁换班的课程结构，即使内容与总课时不变，自建条件存量也系统性偏低——因为存量需要时间长出来，而这个时间由失效率决定，不由内容决定。这一条可以直接检验：同样总课时，分成一个长学期与四个短模块，本模型预测后者的存量显著更低，而两者的当堂表现无差异。

模型的边界要交代清楚。它假定 α 、 c 、 N 、 δ 在一门课内为常数，实际上 c 会随课程建设逐年上升， δ 会因一次课程调整而突然变大；它把所有自建条件当作同质的，而可交接性判据已经说明它们不同质；它没有给出 α 与 δ 的测量方法。上面四条推论只依赖式子的结构与各量的符号，不依赖具体数值——这是本章认为它们可检验的理由，也是本章对这个模型能力的全部主张。

十一、负相关在哪一段成立：把结论所需的条件写出来

上一节的式子把 $(1-c)$ 直接写进了定义，因此“覆盖度越高、自建存量越低”这一条首先是假设的展开，还不是独立发现。本节把模型改写成不预置结论的形式，给出负相关真正成立的条件，并交出一条由此得到的新预测。

其一，把三个参数都写成 c 的函数。外来条件不只是占掉缺口，它还可能改变另外三个量：

$$A^* = \alpha(c, E) \cdot [1 - c] \cdot N(c, E) / \delta(c, E),$$

其中 E 是这门课的其余环境变量（时间、评价制度、班级规模等）。三条可能的正向通道各有其现实机制：教材与平台可能提高转化率 α （有了现成部件，学生更容易把一个缺口做成可交接的工具）；资源可能增加学生看得见的缺口数 N （讲得越多，露出的不足越多）；资源也可能降低失效率 δ （工具有了存放与维护的地方）。

其二，负相关成立的条件。对 c 求导， A^* 下降当且仅当

$$-1/(1-c) + \partial \ln \alpha / \partial c + \partial \ln N / \partial c - \partial \ln \delta / \partial c < 0。$$

左端第一项恒为负且在 $c \rightarrow 1$ 时发散，后三项是外来资源的正向通道。因此可以准确地说：在 c 已经较高的区间，第一项必然压倒其余三项，完备度反比一定成立；而在 c 较低的区间，若资源对 α 或 N 的提升足够强， A^* 可以随 c 上升。

其三，这修正了正文的强度而保住了它的锋刃。“ $c \rightarrow 1$ 时 $A^* \rightarrow 0$ ”这条极限结论不受影响——只要 α 、 N 、 δ 有界， $(1-c)$ 因子就会把存量拖到零。受影响的是中段：本章不再主张自建存量在全区间与资源覆盖度单调负相关，只主张它在高覆盖度区间必然下降，且这个下降是设计的必然结果。

其四，由此得到一条本章原来没有的预测。若上式左端在低 c 区为正、高 c 区为负，则 A^* 关于 c 是单峰的：存在一个使自建存量最大的覆盖度 c^+ ，它既不是零也不是一。这一条可以被实测直接检验——在同一学科、不同资源丰富度的若干门课上采自建存量，看它与覆盖度是单调下降还是倒 U 形。若实测为单调下降，说明正向通道很弱；若为倒 U 形，本章应改写为“存在最优覆盖度”，而“完备的课程没有自建条件”仍然成立。

其五， A^* 与课程长度无关这条推论不受影响，因为它只依赖稳态定义本身：达到稳态所需时间为 $1/\delta$ 的量级，而 δ 的大小与课程长度无关。

十二、辨别格

两根轴：外来条件是否充分，自建条件是否在积累。

	自建存量为零	自建存量在积累
外来条件充分	齐备而空手：一切齐全、执行严格、检查全过，而课程结束时这个班手里一件自己的东西也没有。本章的靶格。	有底盘的生长：外来条件提供起点，班级在其上造出自己的工具并继续使用。
外来条件不足	荒地：什么都没有，也没造出什么。	从零起家：条件匮乏而班级被迫自造，存量高而基础薄。

靶格是左上，三条签名齐：短期指标全部改善（教学资源完备度、规范执行率、进度符合率都最高）；失效不产生任何信号（没有任何检查会问“这个班自己造了什么”）；自我加固（越正规化，外来条件越完备，自造的必要与空间越小）。

右下的“从零起家”格是关键对照，而且它给出一条与常识相反的预测：在外来条件不足而自建存量高的班，与外来条件充分而自建存量为零的班之间，本章预测前者在新情境下的表现更好，尽管后者在原场表现更好。这一条可以直接检验，而且它解释了一个长期被归因于“生源”或“教师热情”的现象。

十三、走查一门课：一件工具是怎么长出来的，又是怎么没长出来的

抽象的存量需要一次具体的走查。下面是同一所学校、同一门程序设计基础课、同一位教师、两个平行班，一学期十六周。

甲班。教材、讲义、题库、评分量规、实验指导书、代码规范文档全部由教研室统一提供，学期初一次性发放。每周的作业从题库中抽取，提交到统一平台，自动判题给出通过与否，教师另按评分量规打分。学生遇到问题在答疑群里问，教师或助教回答。期末考试从同一题库出题。

按本章的读数：**自建条件存量为零，外来条件占比百分之百。**这一点值得逐条确认：代码规范是教研室的，评分标准是教研室的，判题机是平台的，答疑记录留在群里但没有人整理成任何可以被下一个人直接拿起来用的东西。学期结束，班级解散，这个班在这十六周里造出来的可交接物是零件。

甲班的形式审查全部通过，而且是全校范本：资源完备度满分，规范执行率满分，通过率九成七。

乙班。前六周与甲班几乎相同。第七周出现一件小事：三个学生分别在同一类错误上卡了很久——他们把一个循环里的变量作用域搞错了，而自动判题给出的提示只有“答案错误”。其中一位学生把自己排查的过程写成了一段简短的记录，贴在班级共享文档里，标题是“如果输出全是最后一个值，先看这里”。

这件事本身很小。使它成为本章所说的自建条件的，是接下来三件事：第八周另有一位学生直接引用了这段记录，并在下面补了一条自己的情形；第十周这份文档已经有十一条，学生开始在提问前先查它；第十三周一位中途转入的学生拿到这份文档，两天之内自己解决了两个原本要问助教的问题。

按本章的三条读数：**存量至少一件，且可交接性成立**（转入的学生直接用上了，无需任何人讲解）。**外来条件占比下降**——不是因为外来资源减少了，是因为在用条件的总数里多了一件自己的。

第十一周又长出第二件：这个班在一次讨论后自己定了一条不成文但被普遍遵守的约定——提交作业前先把变量名读一遍，能不能让别人一眼看懂。这条约定与教研室的代码规范文档不冲突，但它比那份文档短得多、被使用得多得多。**它可交接**：说给新来的人听，一句话就懂。

学期末，这个班手里有两件东西是十六周前没有的，且都能交给下一届。

两班的差别值得逐条摆出来。乙班的教师没有多做什么，他只做了一件事：把那份共享文档留了下来，并在几次课上提到它。他没有把它做成作业，没有给分，没有要求每人贡献一条——本章反噬那一节说的三种做坏方式，他一种都没做。而这一点恰恰是最难的：**这件事的价值在于它不被考核，而不被考核的东西在一个按周清点的体系里几乎不可能被保护。**

需要说明一个容易混的地方。乙班那份文档如果是教师第二周就拿出来的（内容完全一样、同样好用、同样可交接），本章的读数是零——因为第一条判定条件是“由这个过程自己造”。这一条看起来苛刻，甚至有点不近人情：一份好用的工具，谁造的有什么关系？本章的回答是：造它的那个过程本身就是学习，而拿到一份现成的好工具只是使用。**这个区别在使用体验上完全看不出来，只有在下一次遇到一个还没有工具的问题时才显出来——这也正是本章与分布式认知那条分离线要测的东西。**

十四、十处兑现

类器官培养。 规程为体系自造的基质设字段并据此调整操作；教学规程没有对应字段。

软件团队。 团队自建的约定、脚手架与检查清单有版本历史，可清点；团队交接时这批东西的可交接性直接决定交接成本。

手术团队。 一个团队自己形成的器械摆放与口令约定，换人时是最先丢失的东西，且从不出现在任何培训材料里。

实验室。 一个课题组自己攒的“这台仪器的脾气”清单，是新人上手速度的主要决定因素，而它在任何仪器手册里都没有。

乐团。 一个乐团自己形成的呼吸与起拍约定，是它区别于同水平乐团的东西，指挥换人时最先受损。

厨房。 一个后厨自己定的备料标准与摆位，是出餐速度的主要来源，且不写进任何菜谱。

工地班组。 一个班组自己造的样板与量具，是质量稳定性的实际来源，验收标准里没有它们的位置。

编辑部。 一个编辑部自己积累的体例与常见错误清单，是稿件质量的实际底盘，且它比任何外部体例手册都更被实际使用。

开源项目。 贡献指南与问题模板是项目自己造出来并成为条件的典型，且它们的演化在仓库历史里逐次可查。

学徒制作坊。 师傅与徒弟共同攒下的一套夹具与样件，是这门手艺实际的传承载体，而它在任何技能标准里都不是一个条目。

十处里有四处（类器官规程、软件约定文档、开源项目的贡献指南、编辑部体例）已经有制度化的形式与版本记录，可以直接调取；教育是这十处里唯一一个既大量产生自建条件、又完全不记录它们的地方。

十五、强制不一致：学习成果的归属守不住的两条

现行的学习成果评价体系同时持有两条主张，都写在它自己的规范文件里。

甲：学习成果归属于个体学生，并按个体评价。 这条是整个评价制度的地基。学分、绩点、证书、资格全部发给个人；合作作业里要求区分个人贡献；学术诚信规范的核心正是“这份东西是你自己做的”。没有这一条，评价制度无法运作。

乙：学习是在共同体中发生的，教学应当建设学习共同体、鼓励协作与共享。 这条同样写在几乎所有的课程规范与教学指南里，而且是近三十年教学改革最被广泛接受的一条。

两条在自建条件这个位置上正面相撞，而且相撞的方式与通常说的“合作与竞争的矛盾”不同。**自建条件在物理上是集体持有的**——一套记号、一条判准、一份错误清单，它的价值恰恰在于所有人都用同一套。它无法被分割给个人：把一套命名约定的四分之一归给某个学生是没有意义的。因此在甲的账上，**它不是任何人的成果。**

于是出现一个可以直接观察的后果：一个班花时间去造出一件所有人都在用的工具，这件事在任何一份成绩单上都不留痕迹。它不计入任何人的分数，不进入任何一份能力档案，在课程结束时随着班级解散而失效。乙鼓励学生这样做，甲保证这样做没有回报。

这一体系现行的处置是把集体产物折算成个人贡献——小组作业里的贡献度评分、协作过程的个人反思记录。这个处置在形式上保住了两条，代价是：**折算这个动作本身就把集体持有的东西重新变成了个人持有的东西**，而自建条件的全部价值在于它不被任何人持有。可以折算的那部分（谁写了哪一节）恰恰不是那件工具起作用的部分（大家都认同一套判准）。

二选一。要么承认存在一类学习成果其归属单位是集体而不是个人——那么甲不能以现在的形式成立，评价体系需要一个班级层面的、不可折算给个人的记录项；要么放弃乙，承认共同体只是提高个体学习效率的手段，共同体本身不产生需要被记录的东西——而这与它自己关于“知识在实践中被共同造出来”的主张冲突。

它不能同时守住两条，除非它承认本章的变量。一旦承认，“学习成果”就必须被拆成两样：个体持有的（照现有装置处理）与集体持有且可交接的（需要班级层面的清点），而清点后者需要的正是本章的存量读数。

这一体系已经为此付出了可观察的代价：**几乎所有课程在结束时都不做任何交接。**一个班攒下的东西，下一个班拿不到；下一届从零开始，而这在任何评估里都不会被记为损失，因为损失的那样东西从来不在任何账上。相比之下，同样是一群人在一段时间里共同工作，工程团队、开源项目与编辑部都有交接制度，而且交接的正是本章说的这一类东西。

十六、三处反向约束

细胞学削掉“自建越多越好”。自己造出来的条件同样可以把系统锁死：肿瘤细胞造出维持自身的微环境，体外自组织可以走进一个稳定而错误的构型并且再也出不来。因此自建条件的价值非单调，而且它有一个特有的风险——**自己造的东西最难被自己质疑**。本章因此必须补一条：自建条件需要有可拆的通道，且拆的动作必须由外部触发。**如果没有这一处，本章原本会写下“应当尽可能提高自建条件存量”——这句话被削掉了。**

工程学削掉“过程与自建可以取代验收”。在安全关键的场合，成品验收不可省略，而且自建的工具必须先被验证才能进入使用。因此本章的适用范围收窄到**允许工具在使用中被修正**的场合；在标准必须先于产物的领域，自建条件只能进入辅助层，不能进入判据层。**这一处动的是适用范围。**

艺术学削掉“自建条件都应当共享与清点”。工作室里有一类自建条件是个人的、不可交接的——某个人特有的手法、他自己攒的判断。强行把它们抽出来共享，得到的是一份没人用的文档，而原来那样东西会因为被外化而失效。因此本章的读数只覆盖**可交接的那一类**，另一类明确不在读数范围内，且本章不主张把它转化过来。**这一处动的是读数本身的边界，也是本章最诚实的一处让步。**

十七、三家为何各自到不了这一条

三家的材料都在，判断却没有出现。原因不是谁看漏了，是每一家的问题结构恰好使这一步成为不必要的。

艺术学到不了，因为它的判断单位是一件作品。讲评面对的是摆在那里的那一件东西；它可以延长时间、可以看更多件，但它的每一次判断仍然落在某一件上。而自建条件的特征恰恰是**它不是任何一件作品**——它是一套记号、一条判准、一份清单，是介于作品之间的东西。工作室里这类东西极多（这是本章兑现表里它排在前列的原因），而这一家没有为它设位置：它既不是作品，也不是过程，也不是环境，它是一个在“作品——作者”这个二元里无处安放的第三样。这一家看不见它，不是疏忽，是它的判断单位在原则上不覆盖这个尺度。

工程学到不了，因为它那里工具是外购的。过程能力这一支管的是把给定的过程稳住：量具要经过校准、工装要经过验证、规程要经过评审。**工具在进入过程之前必须已经合格**——这是它的基本纪律，也是它的成就。因此“过程自己造出一件工具并立刻拿来用”在这一家看来不是一个可以研究的现象，是一次违规。它有一整套办法处理工具的引入（验证、确认、变更控制），而这套办法的整个方向是**让工具的来源与过程分离**。它因此不可能问出“这个过程自己造了什么”这个问题——问出来就等于在讨论如何绕过自己的纪律。

细胞学到不了，因为它的实验设计把条件放在自变量那一栏。生态位研究的标准做法是：设定微环境，观察细胞命运。基质硬度、可溶因子、支持细胞类型——这些是研究者拧的旋钮。**当细胞反过来改造这些旋钮时，那在实验设计里叫混杂，叫批次漂移，叫需要控制的干扰。**这一家非常清楚地知道细胞在改造 niche（这正是本章的推翻材料），而它把这件事读作“体外培养难做”的原因，读作需要用换液策略、灌注系统或更好的基质去补偿的东西。它把一个现象当成了一个麻烦。要把它当成一个正的量，需要问“细胞造出了什么、那些东西后来起了什么作用”——而这个问题在一个把条件当自变量的设计里问不出来。

三处缺口合起来说明这个判断为什么必须由三家撞出来：艺术学有大量自建条件而其判断单位不覆盖它，工程学有整套工具纪律而其方向正是切断自建，细胞学有自建的完整证据而把它记在干扰栏里。少任何一家，剩下两家都推不到这一条。

十八、与既有说法划界

建构主义。它说到的是：知识由学习者主动建构，不是被传递的。分离线：它讲的是**内部表征**如何形成，本章讲的是**外部可交接物**如何被造出来并成为条件；一个学生可以充分地建构内部理解而这个班的自建条件存量为零。判定形式：若把学习者的主动建构水平控制住之后自建条件存量不再预测结果，则建构主义覆盖本章；若在建构水平相当的两个班之间存量仍有区分力，则本章对。

知识建构与知识论坛。它说到的是：共同体持续改进公共的知识对象，学生的想法成为后续探究的对象。这是本章最近的一位邻居，必须正面回答。分离线：它的公共对象是**观念**——想法、解释、理论；本章的自建条件是**工具**——可以被指着看、被交给新来的人

直接用的东西。判定形式：若一个班的公共观念持续演化而可交接的工具一件没有增加，知识建构判为成功，本章判存量为零；若一个班的观念讨论平平而攒下了一套在用的判准与清单，两家判断相反。本章预测后者在新情境下的表现更好，而这是可以直接检验的。

实践共同体 (Lave & Wenger)。它说到的是：学习是从边缘参与走向核心的过程，发生在共同体的实践中。分离线：它刻画的是**人的位置移动**，本章刻画的是**物的积累**；一个共同体可以有清晰的参与轨迹而不积累任何可交接物（师徒口传即是）。判定形式：若参与轨迹相同的两个共同体在自建条件存量上无差异，本章不增量。

分布式认知 (Hutchins)。它说到的是：认知过程分布在人、工具与表征之间，不局限在个体头脑里。这是本章最危险的占位者。分离线：它讲的是认知**如何分布在既有的工具上**——它的经典分析对象是一艘船、一个驾驶舱里已经存在的仪表与记录方式；本章讲的是工具是被这个过程自己造出来的，而造这件事本身就是学习。判定形式：**给两个班完全相同的外部工具，分布式认知预测两班无差异；本章预测其中一个班若被允许自造工具，在撤除全部外部工具之后的独立表现更好。**这一条是本章与它之间可以做实验的分界。

生态位构建 (Odling-Smee, Laland & Feldman, 演化生物学·方法学占位者)。把本章的命题剥掉一切教育名词，剩下的形式是“住户改造自己的居所，而改造的结果成为后来者的条件”——这正是生态位构建的形状，它在演化生物学里已经被形式化了二十多年，且明确主张这构成一条与自然选择并列的演化通道。这位占位者极其贴近，必须正面判定。分离线：生态位构建的承重之处是**跨代**——改造的结果作为遗产传给后代，形成生态遗传；本章的承重之处是**同一批人在同一段时间之内**，自建条件对造它的那批人自己起作用。判定形式：若自建条件只在跨届传承时才与结果相关（这一届造的东西对下一届有用，对自己没用），则生态位构建覆盖本章；若在同一学期之内自建条件存量与该班自己的学习结果存在剂量——反应关系，则本章指出了它没有覆盖的那一段。本章赌后者，而这是可测的。

过程能力与统计过程控制 (工程学, 本章学科之外)。它说到的是：控制过程而非检验成品。分离线：它把测量系统当作给定并要求它先被校准；本章说的自建条件里有一大类**就是测量系统本身**——一个班自己定的判准就是它自己的量具。判定形式：若在测量系统完全外来的课堂里自建条件存量仍能预测结果，则两者可分；若存量的效应全部集中在“判准类”自建条件上，则本章实际说的是测量系统的内生性，应当改名并收窄。**这是本章可能需要退让的一处。**

脚手架与撤除。它说到的是：提供支持并逐步撤除。分离线：脚手架是外来的、按计划撤的；自建条件是内生的、不打算撤的。判定形式：若把脚手架撤除的规范程度控制住之后存量不再预测结果，本章被覆盖。

十九、与作者同题诸文的划界

作者另有十余篇同样处理课堂学习的文章，其中三篇已收入本书第三编。总的分界先说清楚：那些文章大多问的是**为什么很少发生**，要的是一个驱动；本章问的是**该怎么运行**，要的是一条次序。以下只列会被混为一谈的几篇，每条给双向判定。

与第七章《开着的时候没人来》：同一组学科的两次不同取用

本书第三编里有一章由同样这三门学科撞成，处理同一个问题。距离最近，必须写在最前面并写细。

那一章处理的是**时序的不对齐**：课堂按内容的逻辑关系排序，而学习者的感受期窗口由上一轮教学打开、有时限、窗外归零；它主张窗口未开而内容到达会产生信号，窗口已开而内容未到达不产生任何信号，因而命名了空窗。

先说两家共用学科的取用点：**同一门细胞学被两章拿走的是两件不同的东西**。那一章取的是**序贯诱导与感受期**——靶组织的感受期窗口本身由更早一轮的信号打开，准备度不是靶方的属性而是上一轮的产物；本章取的是**生态位重塑与自组织**——细胞分泌基质、招募支持细胞、重塑自己周围的条件，少量细胞可以在没有既成生态位的情况下自己长出结构与维持它所需的条件。前者属于发育生物学的诱导一支，后者属于干细胞与类器官一支，两条线在教科书里就是分开的章节。工程学与艺术学同理：那一章的工程学取公差链的分配，本章取过程能力与统计过程控制；那一章的艺术学取材料在制作序列里的抵抗，本章取判断面对作品本身。

再说**承重命题的分离**。那一章的对象是**时机**——窗口何时开、内容何时到，两套时序能否对上；本章的对象是**沉淀**——这个过程造出了什么可以交给新来的人直接用的物。一个是时间对齐问题，一个是资产积累问题。二者可以完全分离：一门时序完全对齐、每个窗口都有相应内容准时到达的课，其自建条件存量可以是零（内容全部是外来的，只是来得正好）；反过来，一门时序错乱、大量窗口空过的课可以攒下一堆工具——**恰恰因为等不到内容**，学生自己做了替代品。

双向判定形式。一个班在教师完全按教材顺序推进、从不错位的条件下，学期末手里一件自己的东西也没有：那一章读数满分，本章为零。一个班因为教材与实际问题脱节而自己整理了一份错误清单并反复使用：本章读数高，而那一章会把这段时间判为大量空窗。**两章在这一格上对同一个班给出方向相反的评价**，这个分歧可以被检验——若这类班的学生在跨情境任务上表现更好，本章对；若他们只是在补一个本可以不出现的洞、长期表现并不更好，那一章对。

两章合用时给出一条各自单独说不出的预测：自建条件本身会打开窗口。一件班级自己造的工具，其被第二次使用的那一刻正是一个感受期的开启——因为使用者此刻正带着一个具体的、由自己造出来的问题去看它。若这一条成立，那么提高自建存量会同时降低空窗率，而这两个读数应当负相关。**这条预测两章单独都推不出来**：那一章没有“窗口由学习者自己打开”这一项（它的窗口全部由上一轮教学打开），本章没有窗口这个变量。若实测发现二者不相关或正相关，本章这条附加预测为伪，而两章的主命题各自不受影响。

与《差一条边》（**接手图·闭合度·催化覆盖率**）。这是书外诸文中与本章最容易被混淆的一篇，必须写在最前面并写细。两篇都在讲“系统自己维持自己”，都从化学或生物的自组织借力，因此距离很近。

分离线在**对象**上：那篇的对象是**关系**——谁的产物被谁接住，构成一张图；本章的对象是**物**——过程造出来、可以指着看、可以交给新人直接用的工具。二者可以完全分离，而且分离的两个方向都有现成的实例。

双向判定形式：一个接手图完全闭合的班——每个人的产物都被某个人接住、催化覆盖率满分——其自建条件存量可以是零：所有的接手都发生在口头，每一次都接住了，而

学期结束时这个班手里一件可交接的东西也没有（研讨型小班的常态）。反过来，一个图不闭合的班——大多数人的东西没人接——可以攒下一份被所有人使用的错误清单：那份清单是被少数人造出来、被所有人取用的，取用不构成接手。**若两个读数的相关系数高于零点六，两篇必有一篇是另一篇的代理**，届时应保留可逐件清点的那一个，也就是本章。

两处合起来给出一条各自单独说不出的话：闭合讲的是这一轮的关系能否自持，存量讲的是这些关系有没有沉淀成不依赖具体人的东西。一张完全闭合的图，只要它的闭合全靠特定几个人，那几个人一走就散；而一件可交接的工具，即使图不闭合也能被下一个人拿起来。**沉淀是闭合的保险，闭合是沉淀的来源**——两个读数应当合用，且本章预测它们的乘积比任一个单独都更能预测跨届的稳定性。

与《每一轮都从新鲜表面开始》（未结转中间体·结转率）。那篇处理的是**未完成的東西**能否被下一轮接上。分离线：结转的对象是一件没做完的工作，自建条件是一件做完了并且可以反复用的工具。判定：一个结转率满分的班——每轮的未完成项全部登记接续——其自建存量可以为零（接续用的是教师的登记表，不是这个班自己造的东西）；反过来，一个从不留未完成项的班可以攒下大量工具。**两读数正交**。

与《正好有人能解决》（清除位·透过率）。那篇处理差异在到达任何位置之前被局部胜任度消耗。分离线：那篇的对象是**被消耗的差异**，本章的对象是**被造出来的工具**。判定：一门透过率极低（一切差异当场解决）的课，其自建存量可以很高——教师造工具、学生用工具，工具是真的，只是不是学生造的；本章的第四条判定条件（由这个过程自己造）在这里正是把两者分开的那道闸，而本章在反噬那一节已承认这道闸最容易被绕过。

与《一条路没有交接点》（交接点·独立结算比）。分离线：那篇的对象是**结算权的位置**，本章的对象是**过程自造的物**。判定：一门由独立第三方判定达标的课，其自建存量可以为零；一门自按自判的课可以攒下很多工具。**两处合用时有一条新预测**：自建条件如果同时被用作判据（一个班自己定的标准），那么它同时抬高本章的读数与降低那篇的读数——**这是两个读数会真正冲突的唯一一处**，本章在与过程能力划界时已经指出这一类“判准类自建条件”是本章可能需要退让的地方，那篇的存在使这个退让变得更有必要。

与《总有人替它起步》（结清态·未被归类窗口）。分离线：那篇讲产物在成熟前被归类，本章讲产物成熟之后有没有沉淀成工具。判定：一门未被归类窗口很长的课，其自建存量仍可为零——悬置了很久，然后消失。**本章预测两者是先后关系而非替代关系**：不被过早归类是产生自建条件的必要条件，但不充分。

二十、与本编另两章的划界

与第四章《手段取得终点资格的那一段》（教育×化学×艺术）。那一章处理的是**终点易主这个事件**；本章处理的是**过程自造的条件**。分离线：换位可以在完全不积累任何工具的课上发生（一次讲评里终点易主而什么都没留下），自建条件可以在终点从未易主的课上积累（一个纯讲授的班自发形成了一套记号）。两个读数正交。**共同证伪条件**：若两个读数的相关系数高于零点六，两章必有一章是另一篇的代理，届时应保留可逐件清点的那一个——也就是本章。

与第五章《要到那儿，下一步该做什么有几个答案》（物理×心理×工程）。那一章处理的是倒推的适定性，是一个**约束**；本章处理的是过程的**产物**。那一章预测自建条件在

多解区最容易积累，因为多条路都要留下痕迹才选得出来；本章对此不作主张，但接受这条预测作为一次外部检验——若实测发现自建条件存量与倒推唯一性正相关（唯一区反而积累更多），那一章那条预测为伪，而本章不受影响。

三章合起来还有一句共同的话，写在这里：三章分别撞出了终点资格可换、倒推可解性可变、条件可自建，三条都是同一个更一般的东西在三个位置上的显形——课堂运行里被当作给定的那些东西，其实有一部分是这个过程自己的产物。本编这三章都没有把这一条写实，这是三章共同交出的洞——本书第四编正是从这个洞开始的。

二十一、五种最强的竞争解释，与排除它们的设计

竞争解释一：其实不就是班级氛围好吗。积极的班自然会攒下点东西。排除设计：用现成的氛围与参与度量表独立测量。本章独有的变量与它正交，而可交接性正是把两者分开的那道闸——一个讨论热烈、参与度极高的班可以一件可交接的东西都没攒下（全部是口头的），一个沉默的班可以攒下一份被所有人使用的清单。本章获胜的观察是：在氛围与参与度匹配的两个班之间，存量仍预测新情境下的表现。

竞争解释二：其实不就是这个班里有几个能干的学生吗。排除设计：控制班内最高水平与水平分布。本章独有的预测是：存量与班内最高水平的相关，应当显著低于存量与“这些东西是否被其他人实际取用”的相关——一件由最强学生造出来但无人使用的东西，按本章的判定条件不计入存量（它不构成条件）。若存量完全由少数人的水平解释，本章测的是个别学生的产出而不是这个过程的自建。

竞争解释三：其实不就是教师放手多吗。排除设计：独立测量教师的指令密度与规定动作占比。这一条与本章的关系比前两条微妙——放手确实提高转化率。本章获胜的观察在于第二个旋钮：在放手程度匹配的两个班之间，是否给了一个存放与取用的位置（共享文档、固定的板书区、可修订的清单）仍能区分存量。若不给存放位置的班也能积累同样的存量，本章关于转化率的主张失去内容。

竞争解释四：其实不就是课程本来资源就少、学生被迫自造吗。这条正是本章自己的模型预测的（存量与覆盖度反相关），因此不能当作反对；要排除的是它的强版本——存量的效应完全来自资源匮乏本身，与自建无关。排除设计：比较三类班——资源充分而允许自造、资源充分而不允许、资源匮乏而不允许。本章预测第一类最好，第三类最差，第二类居中；若第三类不劣于第二类，说明起作用的是资源多少而不是自建。

竞争解释五：其实不就是评分者从好结果倒推吗。清点者若知道该班表现，容易把普通的班级材料读成自建条件。排除设计：清点在结果揭盲前完成；每一件必须给出首次出现的时间、触发它的具体事件与至少三名不同学生的取用记录；报告清点者分歧率。若盲法后存量的信度很低或效应消失，本章不能用“课堂产物难以界定”回避失败。

判决可以压缩成一句：若在同一课程、同一教师、控制氛围、参与度、班内最高水平与教师指令密度之后，可交接的自建条件存量仍正向预测新情境下的表现，且第二次使用的时点早于表现提升的时点，则本机制优于上述五条；若效应被氛围、个别学生水平或资源匮乏完全吸收，或时间顺序相反，则竞争解释获胜。

二十二、分层引用结构

本章的主张分三层，各自的证伪条件不同，可以分层引用。

第一层：自建条件作为一类存在物，且“由这个过程自己造”是它的必要条件。最强也最容易倒。它的关键证伪条件在证伪条件其五：给两个班完全相同的外部工具，若允许自造的那个班在撤除全部外部工具后的独立表现不优于另一个班，则“谁造的”这一条件不成立，本章退化为一个关于工具可得性的主张。

第二层：外来充分度与自建存量两轴的辨别格作为一个分析工具。即使第一层的“必要条件”主张被推翻，这张格子仍然可用：它把“资源是否完备”与“这个班自己有没有东西”分开，而这两件事在现有的课程评估里被压在一根轴上（都记为资源指标）。这一层的证伪条件是形式部分的负相关预测——若两轴正相关，格子退化为一维。

第三层：一条不依赖前两层的经验主张。同样总课时的一门课，分成一个长学期与四个短模块，后者的学生在跨情境任务上表现更差，而两者的当堂表现无差异。这一条完全不需要“自建条件”这个概念，只需要现有的排课记录与一次延迟测验，是三层里最便宜的一条，而且它直接检验形式部分其四那条推论。

由下往上看，第三层几乎不需要前提，第一层则把全部重量都押上了——而本章要的是第一层。若第三层被证实而第一层被推翻，说明现象为真而本章的刻画错了；若第三层被推翻，三层一起倒。

二十三、证伪与赌注

最强的竞争解释是：其实不就是这个班氛围好、参与度高吗——积极的班自然会攒下点东西。

本章必须能排除它。班级氛围与参与度可以用现成量表测量；本章独有的变量是**可交接的自建物的件数**，它与参与度正交：一个讨论热烈、参与度极高的班可以一件可交接的东西都没有攒下（全部是口头的），而一个沉默的班可以攒下一份被所有人使用的错误清单。**可交接性正是把两者分开的那道闸——参与度测的是人的活跃，存量测的是物的积累。**

核心证伪条件：在同一门课程、同一批教师、控制班级氛围、参与度、时间投入与先验知识之后，若自建条件存量与两周以上的、在新情境下的表现之间不存在剂量——反应关系，则本章为伪；届时应把差别归于班级氛围，而非过程自造的条件。

其余七条：

其二，若存量的效应全部集中在“判准类”自建条件上，而工具类、清单类、记号类均无效应，则本章实际说的是测量系统的内生性，命题应改名并大幅收窄。

其三，若自建条件只在跨届传承时与结果相关、对造它的那一届无效，则生态位构建覆盖本章，那条分离线失败。

其四，正向读数（顺序）：一件自建条件被第二次使用，应当早于该班在相关内容上的表现提升出现。若表现提升总是先于第二次使用，则自建条件是学习的产物而不是条件，本章的因果方向反了。**这一条是本章最要紧的一条检验，因为它直接测“它有没有真的成为条件”。**

其五，若给两个班完全相同的外部工具，允许自造的那个班在撤除全部外部工具后的独立表现不优于另一个班，则与分布式认知的那条分界失败。

其六，最便宜的一条：**任何使用版本控制的课程，其仓库历史里可以直接筛出“由学生首次提交、随后被至少三人引用”的文件。**这份数据在多数教学平台与代码托管里已经存在，不需要新增采集，且可交接性可由“新加入者是否直接使用”来判定。

其七，若不同清点者对“这件东西算不算自建条件”的判定一致性低于可接受水平，该读数不可操作，本章应退回为一个分析工具。

其八，若在形式部分列出的四个已有制度化记录的行当里，自建物的存量与产出质量无关，则本章的兑现是同形而非真兑现。

写死日期的赌注：在二〇三一年六月三十日前，至少两项在同一制度环境内、平行班不少于十二个、预注册的研究，报告可交接的自建条件存量在控制班级氛围、参与度与先验知识之后，仍正向预测新情境下的表现；且第二次使用的时点早于表现提升的时点。

以下不算命中：只报班级氛围或参与度；只比较合作学习与讲授；把学生作品的数量当作存量；把教师提供而学生使用的模板计入自建；用两所学校的对比充当剂量——反应；机构自行说明已开展学生自主管理。

二十四、反噬与治理

自建条件存量一旦进入考核，最省事的实现方式是让每个班在学期末交一份“班级学习资源汇编”。这能把件数做满，而它恰恰不是自建条件——那是一次为交差而做的外化，造出来之后不会有人用，可交接性形式上满足而实际使用为零。

第二重反噬更隐蔽：教师会替学生造。一位负责的教师完全可以在第二周就拿出一份“我们班的错误清单”，说成是大家一起整理的。它可用、可交接、被反复使用——本章的三条判据全部满足，只有“由这个过程自己造”那一条被绕过了，而这一条恰恰最难核实。本章的读数在这里是可以被骗过的，而且骗过它的那个人多半出于好意。

三条治理底线：存量指向课堂的运行结构，不指向学生或教师的品质，不得单独用于排名；不得在安全关键的训练中为提高存量而允许未经验证的自造工具进入判据层；任何据此作出的不利安排，必须公开清点规则、每件自建条件的来源记录与复核通道。

我看不到一个彻底的出口。只要这个数被采用，替学生造就永远比让学生造省事，而且看起来更负责。唯一可用的限制是要求每一件自建条件都能追溯到一次具体的、不是由教师发起的触发事件——而这条限制本身也可以被补写。

二十五、若本判断被证伪，我们会学到什么

若自建条件存量对新情境下的表现没有增量预测力，我们学到的不是“班级产物无所谓”，而是一件更具体的事：**工具的来源不重要，重要的是工具的质量与可得性。**那意味着最优的做法是集中力量把教研室的资源做到最好并尽早发给学生——也就是本章靶格里那种课程正是对的。这是一个明确的结论，而且它会终止一批以“让学生自己建构”为名的成本很高的教学安排。

若存量有预测力而“由这个过程自己造”这一条不成立（给两班相同外部工具、允许自造的那班撤除后表现不更好），我们学到的是：起作用的是工具与使用者之间的贴合度，而自造只是达到贴合的一种途径。那会把本章的处方从“让他们自己造”改为“让他们自己改”——允许修改现成工具，成本低得多，且更容易在现行体系里落地。

若存量与外来条件覆盖度正相关（资源好的课自建也多），我们学到的是形式部分那个模型错了：产生率不由缺口驱动，而由整体投入驱动。那意味着自建条件是资源充分的伴随现象，不是它的对立面，本章关于“完备的课程在结构上不可能有自建条件”的推论作废，而这一条恰恰是本章最反直觉、也最容易检验的一条。

三种失败方式各自指向不同的下一步。

本章交出三个自己解释不了的洞。第一，可交接与不可交接的分界在具体清点时会产生真实分歧——一套“大家都知道该怎么做”的做法，写不下来但新人两天就学会了，算不算可交接？第二，本章说自建条件需要可拆的通道且拆的动作必须由外部触发，却没有说清什么样的外部触发是有效的。第三，本章完全没有处理跨届传承：一件自建条件传给下一届之后，对下一届而言它就是外来条件，本章的读数在那一刻把它从分子挪到了分母，而这个挪动是否正确，本章没有论证。

二十六、自反检验

按本章自己的判据，这篇文章造出了什么可以被下一次直接拿来用的东西？

一个名字、三条读数、一张辨别格、一句问话。其中只有那句问话——“这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的”——可以被交给一个没读过本章的人直接使用。其余的都需要先读本章才用得上，因此按本章自己的可交接性判据，它们更接近本章的成品而不是自建条件。

这也意味着本章自己落在了它的靶格附近：外来条件充分（三门学科的文献），产出齐整，而真正可交接的东西只有一件。本章没有理由认为自己比它所描述的那些课堂更耐受这个问题。

本章交出两个自己解释不了的洞。第一，可交接与不可交接的分界在具体清点时会产生真实分歧——一套“大家都知道该怎么做”的做法，写不下来但新人两天就学会了，它算不算可交接？本章没有解决这个分歧。第二，本章说自建条件需要可拆的通道且拆的动作必须由外部触发，却没有说清什么样的外部触发是有效的——那需要另一条完整的次序，本章没有它。

二十七、结论

课堂如何运行才能让学习真正发生，三门学科给了三条不能并存的次序：只看做出来的东西、控制制造它的过程、建好它周围的条件。三家对条件的地位判断从“排除在外”到“决定一切”，而三家共同假定条件是给定的输入。

细胞学自己拆掉了这个假定。细胞分泌基质、招募支持细胞、重塑自己的微环境；类器官从很少的初始条件出发，自己长出结构，并且自己长出维持这个结构所需的条件。生

态位不是住户搬进去的容器，是住户造出来的东西——这条在细胞学里回答的是体外培养怎样才能成功，从没有人把它当作“条件与产物的角色可互换”的证据。

由此得到一条三家单独都推不出来的判断：一个学习过程的条件里，有一部分应当是这个过程自己的产物；一门课如果全部条件都是外来的，它在形式上可以完全合格，而它每一轮都从同一处重新开始。

这句话的可操作形式是一句可以逐件清点的问话：这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的？

如果答案永远是“什么也没多”，那么这门课的教材可能最全、规程可能最细、执行可能最严，而它的每一届学生离开时，留下的和进来时一样多。

接续。本章的稳态式是全书最硬的一段：导论第七节把它提为**完备度反比**，第二编小结第四节说明它在另两章里的同形。自建存量的清点办法与两道闸（必须是可指着看的物、必须由这个过程自己造）见附录一。若实测发现存量与资源完备度**正相关**，本章与依赖它的那几处一并作废——这条列在附录三第一节第二款。

第二编小结 三个旋钮，同一只手

一、三章各自推出了什么

第四章（终点）。净收益 $\Delta V(t) = \beta \cdot \Delta U_M(t) \cdot r - c \cdot t$ ，一阶条件 $\beta \cdot r \cdot d\Delta U_M/dt |_{(t=t^*)} = c$ 。三条反直觉后果：在换位期间目标端不推进这一建模假设下，最优换位时长与目标端的不确定性**无关**（ U_T 不出现在一阶条件里）；交还成功率 r 与转化效率 β 相乘，因此**交还做不实**，换位的价值取不回来，**最优时长趋近于零**； $\beta \rightarrow 0$ 即本章适用范围的形式表达——在手段与产物几乎无关的内容上（背诵、查表），**从不换位是正确的**，这不是让步，是同一个式子的退化。

第五章（路径）。期望损失 $E[L] = \sum_n P(n|\hat{n}) \cdot L(M,n)$ ，关键是 $L_2 \gg L_1$ ：在唯一区误用推进，损失是**有界的效率损失**；在多解区误用倒推，损失是**能力损失**，而且在原场完全**不可见**。三条后果：档位不确定时**默认推进**（与通行做法相反）；最危险的不是无解区而是**被误判为唯一区的多解区**（无解区会被发现，误判的多解区在所有可见指标上都好看）；教师经验会系统性高估 $P(1)$ ，因此判据必须问五位而不是一位。

第六章（条件）。存量方程 $dA/dt = \alpha(1-c)N - \delta A^*$ ，稳态 $A^* = \alpha(1-c)N/\delta$ 。三条后果：在 α 、 N 、 δ 不随覆盖度变化这一条件下，存量与外来条件覆盖度**线性反相关**；放开这一条件后，高覆盖度区间的压制仍然必然，中低区间则可能上升、下降或呈倒 U 形（见下节），且这是设计的必然结果而非执行不力； $c \rightarrow 1$ 时 $A^* \rightarrow 0$ ，一门“完备”的课程在**结构上不可能有自建条件**； α 与 c 是两个独立旋钮，本书主张抬高 α （允许自造、给存放处、让它被下一次真的用上），**不主张压低 c** ——那三件事不减少任何外来资源。

二、三个位置，同一个动作

把三章的对象并排放，会看到它们是同一句话在次序的三个部件上的形态：

部件	被当作给定的	实际上	读数
终点	目标先定，中间只是手段	手段可以临时取得终点资格，再交还	换位率、触发源分布
路径	“下一步该做什么”总有答案	倒推的解可以是一个、多个或没有	倒推唯一性档位
条件	条件是给的	有一部分是这个过程自己造的	自建存量、来源比、可交接性

三行的右侧其实是同一件事：**次序里被当作外生的那些量，有一部分是内生的**。而三行的读数各自独立——这一点不是修辞，是本编的一条硬承诺，下一节给出它的可判决形式。

三、三个读数正交，以及这一条怎样被推翻

- 三章各自写过与另两章的双向判定，合起来是三条：
- 一门倒推唯一性最高的技能课，可以从不换位（换位率为零而倒推档位满分）；
- 一个纯讲授、终点从头到尾没有易过主的班，可以自发攒下一套自己的记号（自建存量高而换位率为零）；
- 一门自建存量为零的课，其倒推档位可以处在任何一档（工具的有无与下一步有几个答案无关）。

三条合起来的承诺是：**三个读数两两之间不应有强相关**。推翻它的观察很直接——若任意两者的相关系数持续高于零点六，两章里必有一章是另一章的代理，届时应保留可清点成本更低、可追溯性更好的那一个。第十一章把这条承诺扩到六个读数。

四、三条推导的共同形状

三条推导形式上毫无关系（一个是一阶条件，一个是期望损失比较，一个是一阶线性微分方程的稳态）。但它们推出的后果有同一个形状，而这个形状比三条推导本身更值钱：

在每一处，把事情做得更好的那个方向，同时也是把这一段压掉的方向。

第六章最直白：课程建设每完备一分，自建条件按比例少一分，而完备的每一步在任何评估里都记为进步。第五章的形态是：**教师越熟练，P(1) 被高估得越厉害**——一个人在某内容上教了十年，他给出的“下一步”确实稳定而唯一，对他自己而言；于是经验越丰富，越容易把多解区当唯一区讲，而这件事在这门课里永远不会被发现。第四章的形态是：**目标写得越清楚、活动与目标对齐得越好、评价与目标一一对应，换位率越趋近于零**——这正是第四章的靶格，而它是教学设计这门学问所要求的理想状态。

三处都不是失职的产物，是**尽职的产物**。这就是导论第七节所说的完备度反比：账本每完备一分，那一段就被压掉一分。本编的贡献是证明了这一条不是一句概括——它在三个互不相干的结构模型里各出现了一次，形式不同而符号一致。

五、与第一编的接线

第一编给出的对象有一项构成条件叫**持位可换**：核心功能不能锁死在某一位教师、某一个尖子生或某一套固定平台上。本编第六章的自建条件，正是这一条在**物**的一侧的形态——一件班级自己造出来、可以指着看、可以交给新人直接拿去用的工具，就是“发起权不锁在原持有人手里”的物质证据。因此第六章的第三条判定条件（是**可指着看**、**可交给新人直接用的物**，不是知识）不是一个技术性限制，它是第一编那条构成条件的可清点化。

第四章与第一编的接线在另一处：**交还**。可再入生成体靠被交出去而继续存在；而换位若不交还，终点就不是被临时借走，是被吞掉。第四章把“交还”写成了 r ，并且证明它与 β 相乘——这是全书第一次把第一编的一条存在条件写进一个式子。

六、与第三编的接线

三章都写过与第三编对应章的划界，合起来给出三条**两编合用才推得出的预测**，任何一编单独都说不出：

1. **换位若发生在某个悬置尚未解除的时段内**（第八章），被悬置项与新取得终点资格的那一样之间会结成更多、更难拆除的连接——因为此刻注意力正落在方法上而不落在被悬置项上；两者的效应相乘。
2. **无解档的单元里**（第五章），第九章的双向翻转率中“不会到会”那一支应当最高——那里没有人在指定下一步，变化只能自发进行，而自发进行的变化恰恰是最容易被过早判定冻住的一类。
3. **自建条件本身会打开窗口**（第六章与第七章）：一件班级自己造的工具，被第二次使用的那一刻正是一个感受期的开启，因为使用者此刻带着一个由自己造出来的具体问题去看它。故自建存量与空窗率应当负相关。

这三条是本编与第三编之间真正的接缝。它们都可以被同一批数据检验，而且**每一条都可能失败**——第三条最脆：若实测发现自建存量与空窗率无关，说明工具的使用并不打开窗口，两编之间就只剩下并列关系。

七、回到母题

母题的第一个部件是“课堂里决定后果的那些东西，有一大部分是课堂自己造出来的”。本编是这一部件的**正面证据编**：终点会易主，路径的解的个数随知识分布而变，条件有一部分由过程自造。三处都不是例外情况，都是常态。

母题的第二个部件是“账本只为不是它造的那些东西设了字段”。本编的三个靶格各是这一部件的一个现场：目标写得最清楚、对齐得最好的那门课，换位率恰恰是零；教师最有经验、下一步给得最干脆的那门课，最可能把多解区讲成唯一区；资源最齐备、执行最严格、检查全过的那门课，学期末这个班手里一件自己的东西也没有。三处的共同点是：**账本上每一栏都满分**。

八、本编的可推翻条件

第一条。三个读数中任意两个相关系数持续高于零点六——则三章中必有一章是另一章的代理。

第二条。若实测发现自建条件存量与教学资源完备度**正相关**，第六章的式子方向错误，本编第四节整段作废，导论第七节随之作废。这是全书最要紧的一次单点检验。

第三条。若在多解档单元里默认使用推进法的班级，其条件变化后的表现与默认使用倒推法的班级无差异，则 $L_2 \gg L_1$ 不成立，第五章的默认规则应当收回。

第四条。若一门课把交还做实（明确规定这一段结束后回到哪个终点、原终点被改了什么）之后，换位的迁移效应没有提高，则 r 与 β 相乘这一条不成立，第四章的处方失去根据。

第五条，针对本编整体。若三条结构模型的后果各自成立，而把三者放在一起时它们互相抵消（例如抬高 α 必然降低换位率），那么本编“三个独立旋钮”的说法是错的，第十四章的相乘假设也随之动摇。这一条只有在六个读数被同时采齐之后才能判定。

第三编 Why：发生的三处熄火

本编要回答的是：学习为什么这么少？

这个问题要的是一个驱动，不是一条次序，也不是一个定义。三章分别落在同一次改变的入口、伴生物与出口上。三章按位置排列，不按写作次序。

关于“学习为什么少”，通行的解释有三类：学生的原有观念没有被充分动摇、教学路径安排不当、评价与进度压制了深度。三类解释都对，都能举出例子，也都被反复实施过。本编不加入这三类，而是提出第四个位置——一个不在学生、不在教师、也不在教法上的位置。

本编三章的共同形状是：它们描述的失效都不需要任何人犯错。没有人偷懒，没有人讲错，没有人评错；每一步都按规范做完，而那件事仍然没有发生，并且**不产生任何信号**。这是它们与前三类解释最要紧的分别：前三类解释的失效都会以某种方式露面（学生答错、进度落后、课堂沉闷），本编的三处熄火不露面。

第七章：入口。上一轮教学在学生身上打开了一个感受期，而本轮内容按学科的逻辑次序排，两套时序互不迁就。窗开着的那几天没有人来，窗关上，什么也没有发生——而**“什么也没有发生”不会以失败的形式被记录**。推翻共有前提的材料来自细胞学自己：**序贯诱导**——靶组织能不能被诱导，取决于更早一轮的信号是否已经把它变成可诱导的；准备度不是靶方的属性，是上一轮的产物，而且有时限，窗外不是打折是归零。读数是**对窗率**。本章最危险的近邻是关键期，分离线在第十三节写足四条。

第八章：伴生物。要让一处发生改变，课堂必须先把其余部分挡住——限定题型、简化情境、给出范例、说一句“这个以后再讲”。被挡住的那部分并没有停在原地：它与刚学的内容之间自行接上一些线，脱悬之后没有人去拆，也不进任何账本。推翻材料来自化学自己：**酰基迁移**——用来挡住一个位点的保护基会自己跑到另一个位点上去，于是“挡住”这个动作改写了它本要固定的那张名单。读数是**悬置密度与脱悬率**。这些线不是错误，它们在原题型里不产生任何差错，因此不会被批改抓到；它们只在换一个表面不同的情境时决定表现，而那时通常已经被归因为“基础不牢”。

第九章：出口。一次重组还在进行的时候被读了数，读出的数被当作此刻的状态，并据此配置此后的条件——分层、选课、期望、给什么任务。推翻材料来自物理学自己：**物理老化**——被判为冻住的玻璃态系统，其性质随冻住之后所等的时间持续演化，而演化的起点正是被冻住那一刻的构型。读数是**窗比与双向翻转率**，后者是全批最便宜的一个：一份现成的卷子加四周间隔即可。

三章各自独立地写下了同一句话：三处的**效应相乘不相加**。入口对不上，另两处做到最好也没有输入；入口对上而伴生物无人拆除，出口读得再准，也只是准确地读到一个带着杂物的结果。这一句是本编交给第十四章的东西，也是全书唯一一条与教改惯行方向正面相撞的推论。

第七章 开着的时候没人来

判断卡片：准备度不是学生的属性，是上一轮的产物，而且有时限；窗关上时什么也不发生，因此这件事在原理上不会以失败的形式被发现。

三个理论入口：细胞学 × 工程学 × 艺术学

提要。课堂按内容的逻辑关系排序：先讲这个，因为后面那个要用它。这条排法从未被认真质疑，因为它看起来只是常识。细胞学主张一次诱导成不成，不取决于信号强度，取决于靶细胞当时在不在感受态；工程学主张一个装配体能不能工作，不取决于每个零件多精确，取决于公差链怎么分配；艺术学主张作品成为什么样子，不取决于作者的意图，取决于材料在制作序列里的抵抗。三家争的是改变由什么驱动，而共同假定了一条没有说出口的话：“准备好了没有”是可以先于动作被读出的属性。推翻它的材料来自细胞学自己——序贯诱导：靶组织的感受期窗口本身是由更早一轮的信号打开的，“准备度”不是靶方的属性，是上一轮的产物。带回课堂便得到本章命名的对象：空窗——由上一轮教学打开、而在其开启期间没有任何相应内容到达的感受期。承重判断是：学习之所以少，不是内容顺序错了、不是学生基础差、也不是教法不对，而是教学序列按内容的逻辑排、感受窗口按上一轮的产物开合，两套时序互不对齐，且课堂只记录前者。本章给出对窗率这一读数、一张二维辨别格、一句不含判断词的问法、与九种相近说法的分界（含关键期这一最危险的占位者）、三处反过来加上的限制、七条证伪条件与一条写死日期的赌注。

一、问题从哪里来

一位教师讲完反比例函数，学生第一次真切地看见“两个量在变，而它们的乘积不动”。这一刻在学生身上打开了一样东西：他现在对“什么东西在变化中保持不动”这类问题特别敏感，随手就能举出例子，也会主动去问别的地方有没有这种情况。

按教材，下一节是反比例函数的图像。图像这一节讲得很好，学生学得也不差。而那个刚打开的口子在两三天之内合上了——不是被压制，是没有东西到达。等到几个月后讲能量守恒、讲化学计量、讲比例推理，那个口子已经不见了，教师面对的是一群“以前学过反比例函数、但看不出这两件事是一回事”的学生。

这不是一个偶然的排课失误。教材按内容的逻辑关系排序——先讲函数的定义，再讲它的图像，再讲它的应用——这条排法在学科内部完全自治，而且是唯一可以事先写下来的排法。问题在于，被打开的那个口子不按内容逻辑开合，它按上一轮实际产生了什么开合。两套时序谁也不迁就谁，而课堂只记录前一套。

窗口关上时不产生任何事件。那几天里课上得很顺，作业正确率正常，因为讲的是图像，而图像不需要那个口子。等到几个月后发现学生迁移不动，归因几乎必然是“基础不牢”“不会灵活运用”——两个归因都指向学生，不指向那次错开。

本章追问的正是这一段。它不报告新采集的数据，是一篇理论性文章；其经验主张需要预注册研究与独立复核检验。

二、细胞学的单因主张：已有状态与信号条件不容，驱动分化路径改变

发育生物学处理过一个形式上完全相同的问题。一块组织能不能被某个信号诱导成某种结构，不取决于信号有多强，也不取决于这块组织的谱系有多“好”，而取决于它在接到信号的那一刻**处不处在感受态**。这个概念在发育学里有近一个世纪的历史，是移植实验的核心读数：同一块供体组织移到同一位置，早两小时成，晚两小时不成（Waddington, 1940; Gurdon, 1987）。

感受态是有时限的。窗口之外，同样浓度、同样时长的信号完全无效——不是打折，是归零。这一条使发育学不得不同时记录两套东西：信号的时空分布，与靶组织的感受期。

这一支的单因主张是：**决定一次诱导是否成功的，是靶方当时在不在感受态，不是信号强度**。时序读数是：先有窗口的开启，随后信号若在窗内到达则分化路径改变，改变后的组织又成为下一轮信号的来源或靶方。

单因锁定成立：这一支不认为“信号强度”与“感受态”是并列因素。在窗外，强度这个变量的斜率为零。

它自己知道、却不往这个方向用的事实是**序贯诱导**：靶组织的感受期窗口本身是由更早一轮的信号打开的。在中胚层诱导等经典体系里，一次诱导的产物正是下一次诱导的感受方，而它的窗口由上一次的强度与时长决定（Gurdon & Bourillot, 2001）。这一支用它来解释发育时序为何如此精确地被编排；**它从来不被当作一条关于“准备度是不是可先于动作读出的属性”的证据**。

三、工程学的单因主张：实际尺寸与装配序列不容，驱动装配条件改变

机械装配提供第二个位置。一台机器装不上，通常不是因为某个零件不合格——每个零件都在图纸公差之内，而它们叠起来就是装不上。决定装配成不成的，是公差**如何沿装配序列累积与分配**，即公差链（Whitney, 2004）。

这一支的单因主张是：**装配失败不是零件不合格，是公差链没有被分配**。时序读数是：先有各零件的实际尺寸分布，随后它们与既定装配序列不容，于是装配条件本身改变——加夹具、改基准、留修配量。

单因锁定成立：这一支不认为提高每个零件的精度是替代方案。提高精度成本按指数上升而累积公差按平方根下降，因此工业上的正解从来不是“都做准一点”。

它自己知道、却不往这个方向用的事实是**选择性装配**：把零件按实测尺寸分组，大孔配大轴、小孔配小轴，用事后配对吸收累积公差。这是量产轴承与发动机的标准工艺。这一支用它来降低成本；**它没有被用来问：一个不能事后配对的系统，靠什么吸收累积误差**。发育不能选配——窗口过了就是过了，没有“等下一批合适的信号来配”这回事。这一条恰恰是本章与工程学之间最锋利的地方。

四、艺术学的单因主张：制作路径与材料条件不容，驱动作品呈现改变

工艺与制作研究提供第三个位置，而且它每天都在处理时间窗。一块陶土只有在**皮革硬**这个含水阶段才能修坯、接坯、刻花：太湿则塌，太干则裂，窗口以小时计。一层油画只有在**可覆时间**内才能安全地上一层：太早则搅动下层，太晚则附着不牢，窗口以天计。

这些窗口在工坊里有名字、有判据、有师徒相传的手感（Rhodes, 1957; Mayer, 1991）。

这一支的单因主张更强：**决定一件作品成为什么样子的，是材料与工具在制作序列中的抵抗，不是作者的意图。**意图在材料面前不断被改写；作品是这一路改写的结果，不是意图的实现（Ingold, 2013）。

时序读数是：先有制作路径与材料当前状态的不容（土太干了、颜料还没干透），随后作品实际呈现出来的样子改变，改变后的样子又决定下一道工序何时可以开始。

单因锁定成立：这一支不认为意图与材料是共同作用的两个因素。它主张意图只有通过材料的抵抗才成为形式，因此抵抗是决定项。

它自己知道、却不往这个方向用的事实是**熟练者与材料的往复**：熟练的工匠不是更能预测窗口，而是随时改变自己的路线以就窗口——不是“等土到皮革硬再来”，而是“土现在到皮革硬了，那就先做这一步”。这一支用它来描述技艺的性质；**它没有被用来问：一个必须按预定工序表推进、无法随窗口改路线的制作体系，会损失什么。**而这正是课堂的处境。

五、三家为何不能同时为真

三家把决定性的驱动放在三个不同位置：细胞学放在已有状态与信号条件的不容，工程学放在实际尺寸与装配序列的不容，艺术学放在制作路径与材料条件的不容。并列起来会得到一句无风险的话：教学要看学生的准备度、要安排好顺序、要顺应实际情况。这句话正确，而任何失败都可以事后归入其中一条。

三家实际上互相取消。

工程学的日课正是细胞学判为不可能的那件事。工程的现实对策是选择性装配：先做出来，实测，再分组配对——把对齐推迟到事后。而发育里没有选配：窗口关上之后，没有任何办法把“当时该来的那个信号”补配给它。若细胞学那一条是对的，工程学的整套误差吸收方案在这一类系统里不可移植；若工程学那一条是对的，细胞学关于窗口不可逆的坚持就是没有找到配对机制而已。

艺术学的日课正是细胞学判为坏工艺的那件事。熟练工匠的做法是**随时改路线以就窗口**——土到皮革硬就先修坯，颜料没干就先去画别的。而发育的时序是被严格编排的，改路线在那里等于畸形。若艺术学那一条是对的，“按预定工序推进”本身就是病；若细胞学那一条是对的，随窗口改路线只能在没有严格依赖关系的系统里成立。

细胞学的处方在另两家眼里不可执行。细胞学的办法是让信号的时序与窗口的时序精确对齐——发育用了几亿年演化出这套编排。课堂做不到：课程表必须事先写出来，而窗口在上一轮结束之前不知道会开在哪儿。

三家必须有两家错，或者三家的病灶都放错了位置。本章主张后者。

六、共有前提，与推翻它的那件材料

三家争的是一次改变由什么驱动。而它们共同假定了一条从未说出口的话：

“准备好了没有”，是一个可以先于动作被读出的属性。

细胞学在移植之前测靶组织在不在感受态；工程学在装配之前测零件合不合公差；艺术学在下一道工序之前摸土到没到皮革硬。三家的整套实践都建立在“先读准备度、再动手”之上。把这句话念给三家听，三家都会说这还用说吗——它不是任何一家的立场，是三家共同站着的地面。

课堂对它的信奉更彻底：前测、分层、先修要求、诊断性评价、学情分析，整套装置都在做同一件事——**在教之前读出学生准备好了没有。**

推翻它的材料在细胞学自己手里，而且是发育学教科书上的常识：**序贯诱导**。靶组织的感受期不是它自带的属性，是**更早一轮信号的产物**；窗口什么时候开、开多久，由上一轮的强度与时长决定。这意味着三件事，逐件都与共有前提相反：

其一，**准备度不是靶方的属性，是上一轮的产物**。测它等于在测上一轮做了什么，而不是在测这个靶方是什么样的东西。

其二，**它是有时限的，且时限也由上一轮决定**。因此“现在准备好了没有”这个问题的答案在几小时内会变，而变化的原因不在靶方身上。

其三，**它会自己关上，且关上不产生任何事件**。窗口关闭时组织没有任何异常，它只是回到了一般状态；只有在事后发现该出现的结构没出现时，才能反推窗口曾经开过又关了。

带回课堂，共有前提塌掉。前测测到的不是这个学生是什么样的学生，是**上一轮教学在他身上打开了什么**；而这个东西会在几天内合上，且合上时全班表现正常。课堂用来读准备度的整套装置，读的都是一个正在变化而它假定不变的量。

七、空窗

据此本章命名一个对象。

空窗：由上一轮教学在学习者身上打开、而在其开启期间没有任何相应内容到达的感受期。

四个必要条件缺一不可。第一，上一轮教学**确实打开了**某一处感受期——它的可观测签名是学习者在此期间自发地举出相关例子、提出相关疑问、或在无关任务里迁移使用该结构。第二，在窗口开启期间，**教学内容按内容逻辑推进到了别处**。第三，窗口关上——自发涌现停止，学习者不再主动使用该结构。第四，**不产生任何信号**：这段时间的课上得正常，作业正确率正常，因为讲的内容不需要那个口子。

三条性质使它区别于邻近的坏情况。

它不由学生的一般能力决定。**基础好的学生同样有空窗**——他的窗口开得更宽、涌现更明显，因而关上时损失更大。这一条使本章与“基础差”划清界线。

它不是打折，是归零。窗外教同样的内容，效果与从未打开时相同。这一条使本章与“最佳时机”这类程度说法划清界线。

它不与内容难度相关。一个很容易的内容可以打开一个很要紧的窗口（反比例函数打开的是不变量意识），一个很难的内容可以什么也不打开。

需要排除两类伪阳性。其一，**窗口未开而内容到达**：这会产生信号——学生听不懂，教师会察觉并处置。它是无效教学，不是空窗。其二，**内容到达而窗口仍在**：这是对上了，不是空窗。

也要排除一类伪阴性：**内容标题相同不等于对上了窗口**。下一节讲“反比例函数的应用”，标题上完全相关，而打开的窗口是“不变量”，讲的是“应用题的解法步骤”——两者不是一回事。**判定的对象是上一轮实际产生的那样东西，不是上一轮的内容标题。**

八、空窗的四种亚型

按被打开的是哪一种窗口，可以分出四类。四类的时长、签名与能接住的内容都不同，因此登记的难度也不同。

不变量窗。学习者第一次看见“变化之中有东西不动”——反比例函数的乘积、化学反应的原子数、守恒律、几何变换里的不变量。它的签名最容易辨认：学习者开始主动举例，而且举的例子跨出本学科。时长以天计。它能接住的内容极广，因为不变量意识几乎是所有定量推理的入口。

反例窗。学习者第一次自己造出一个反例，把一条他刚接受的规则推翻。签名是他开始追问“那如果……呢”，且追问的对象是规则本身而不是题目。时长以天计，但**极易被一次“你想得太多了”**关闭。它能接住的是任何需要边界意识的内容。

可验证窗。学习者第一次不靠权威判断对错——自己代回去验算、自己查一遍、自己做个小实验。签名最难辨认，因为它表现为**追问减少而不是增加**：他不再问“这样对不对”，而是直接去查。时长最短，以小时到一两天计，因为它极度依赖当下那次成功的验证经验。它能接住的是一切关于判据的内容。

结构同形窗。学习者看见两件表面不同的事其实是一回事——行程问题与工程问题、化学平衡与经济均衡、递归与数学归纳法。签名是他开始说“这跟那个是一样的”，且说得哪里一样。时长最长，可达数周，因为同形的识别本身需要慢慢展开。它能接住的是任何跨单元、跨学科的内容，因此它也是被课程表错过得最彻底的一种——跨单元的内容按定义不在下一节课里。

四类的登记难度依次上升：不变量窗与结构同形窗可以从自发举例中直接辨认，反例窗要看追问的对象，可验证窗最难，因为它的签名是一个消失（不再问对不对）而不是一个出现。一个只按“**学生提问多不多**”来判断课堂活跃度的观察表，会把可验证窗记为零。

四类也可以并存。一次成功的推导可以同时打开不变量窗与可验证窗，而两者的时长相差一个数量级——若下一轮只接住其中长的那个，短的那个照样关上。因此登记必须**按类分记**，不能只记“窗口开了”。

九、二维辨别格

横轴取**窗口是否开启**，纵轴取**相应内容是否在窗内到达**。两轴相互独立。

	内容未到达	内容到达
窗口未开	无事发生：按内容逻辑推进，与窗口无关；多数课时属此格，本身不构成损害	无效教学：讲了没用；会产生信号——学生听不懂，教师会察觉并处置
窗口已开	空窗：窗开着，而这段时间在讲别的；不产生任何信号。本章的靶格	对窗：内容落在能被接住的时刻，此处常出现与投入不成比例的快速进展

左下角是靶格，三条性质使它难以被发现。

其一，短期指标表现为改善。空窗期间讲的是不需要窗口的内容，这类内容通常更程序化、更容易讲清、正确率更高。窗口错过得越彻底，那几天的课看起来越顺。

其二，失效不产生任何信号。窗口关闭是一个非事件：没有错误、没有困惑、没有投诉。它的后果要到几个月后才以“迁移不动”的形式出现，而那时它已经无法被追溯到任何一节具体的课。

其三，它自我加固。课堂越是严格按内容逻辑排序、进度越可控、教学越规范，内容到达的时刻就越是由课程表决定而与窗口无关；而保证进度的压力恰恰来自“顺”，即靶格的第一条性质。

右上角的无效教学是关键对照格：同样是错开，窗口未开而内容到达会产生信号，窗口已开而内容未到达不产生信号。两格的后果不对称，而课堂的全部纠错能力只对前一格有效。这一条本身就解释了为什么课堂改革总在处理“太难了”“听不懂”，而从不处理另一格。

十、读数与一句不含判断词的问法

对窗率：一段教学中，落在由上一轮打开的窗内的内容占比。分子的判定需要两步——先在上一轮结束时记录窗口的可观测签名（自发涌现的类型），再看下一轮的起点是不是那一处。

判据写成一句不含判断词的问话：

教案上，后一节课的起点是由前一节的内容序号决定的，还是由前一节实际产生的某样东西决定的？

这句话可以拿去问教案，不需要问任何人的判断，而且答案在绝大多数教案里是前者——这不是发现了什么丑闻，是课程表这件东西的定义。更细的一问是：上一节课结束时，学生身上被打开了哪一处口子？这一节的头二十分钟讲的是不是那一处？

五个场景跑一遍，答案互不相同：

场景	上一轮打开的窗口	下一轮按序讲什么	对窗与否
数学·反比例函数	两个量同时变而乘积不动	该函数的图像	不对；窗口以天计，图像与不变量无关
化学·一次失败的滴定	终点判断的主观性	下一个实验	不对；且失败被记为操作失误，窗口连记录都没有
编程·第一次调通递归	敢于信任自己写的函数	语法与数据结构	不对；而这个窗口极短，一两天内合上
陶艺·修坯	材料在特定含水阶段的可塑性	按周排的下一节课	结构上不可能对窗——皮革硬窗口以小时计而课程按周排；工坊的现有处置是学员自行回来或师傅代修，即代做

场景	上一轮打开的窗口	下一轮按序讲什么	对窗与否
语言·纠正性反馈之后	对某个形式的注意	教材下一课	部分对窗；二语习得里对“即时接受”有成规模的课堂观察数据可查

第四格与第五格是查出来的，不是从命题推出的。陶艺的皮革硬阶段有明确时长，而课程表按周排——**这一格的错配不是执行问题，是排课单位与材料时间尺度之间差了两个数量级**，工坊的应对办法（学员自己回来、或师傅代做）正好是本章预测的两条出路。二语习得对纠正性反馈之后的即时接受有系统的课堂话语编码（Lyster & Ranta, 1997），那批数据里的接受率与后续保持之间的关系，可以直接被重新读作对窗率的一个特例。

十一、十处兑现

其一，感受态对应窗口，因此研究必须在上一轮结束时记录窗口的签名，而不只是记录讲了什么。其二，感受期有时限对应窗口会关上，因此本章预测同一内容在窗内与窗外的效果差异是断崖式的而非渐变的，可用不同延迟的平行组直接测。其三，序贯诱导对应窗口由上一轮产生，因此本章预测**改变上一轮的内容会移动窗口的位置**——这是本章最要紧的一条可操作预测。其四，诱导的强度在窗外斜率为零，对应“窗外加课时无效”，可用等量课时的窗内组与窗外组对比。其五，公差链对应窗口与内容的时序累积误差，因此错配是可以逐轮累加的，不是每轮独立。其六，选择性装对应“事后配对”，因此本章的第一处方不是重排课程表（不可行），是**建立登记与自取机制**：把打开的窗口登记下来，让学习者在窗内自行取用，而不是全班同步。其七，公差不能靠提高单件精度解决，对应“提高每节课的质量不能解决错配”，因此本章预测教师培训对对窗率无效。其八，皮革硬与可覆时间对应窗口有上下界，因此本章预测过早到达与过晚到达同样无效，效应曲线是单峰的。其九，熟练工匠随窗口改路线，对应教师在窗口出现时临时改动教案，因此对窗率与教案的偏离率应当正相关——**一个从不偏离教案的教师，对窗率在结构上接近于零**。其十，材料的抵抗决定形式，对应“学习者实际产生了什么”决定下一轮该讲什么，因此课程表在原理上不能被完整地预先写出。

第三、第六与第九项是本章最依赖的三项。删掉第三项，本章的对象退回“最佳时机”；删掉第六项，处方退回“重排课程”（不可行）；删掉第九项，读数无法从现成材料中取得。

十二、三家为何各自到不了

三家没有到达这个位置，不是因为疏忽。每一家各自的问题结构，恰好使这一步成为不必要的。

细胞学到不了，因为在它的体系里信号的时序也是内生的。发育中的信号与感受态由同一套调控网络产生，因此“两套时序不对齐”这个问题在正常发育里不出现——不对齐的胚胎不能存活，于是这一支从来观察不到一个长期运行的错配系统。它掌握的关于窗口的一切知识，都是在两套时序本来就对齐的体系里取得的。**它的问题结构使错配成为不可观察的，而不是它没想到。**

工程学到不了，因为它可以事后配对。选择性装配这条出路的存在，使“窗口不可逆”在工程学里不成为问题。零件做出来先放着，实测之后再分组配对——时间不是约束，库存就是缓冲。因此这一支从不需要问“如果配不了怎么办”，也就没有理由发展出关于窗口不可逆的任何概念。它给了本章一条极好的处方，却看不见那条处方所要解决的问题。

艺术学到不了，因为它的工作对象持续存在。土还在，画还在，师傅可以等，也可以让学员回来。工坊里从不出现“窗口开着而对象不在场”的情形——对象一直在场，那才是窗口能被利用的前提。因此这一支发展出了极其精细的窗口判据（手感、含水度、指触干），却完全没有发展出“错过了怎么办”的概念，因为它那里错过的代价是这一件报废了，重新做一个就是。

三家的盲区各不相同，交集恰好是课堂的处境：一个窗口内生、不可事后配对、且工作对象在窗口期间不在场的系统。细胞学有内生与不可逆而无错配；工程学有错配与不可逆而可配对；艺术学有内生与在场而对象可弃。三条各占两样，课堂三样都占——这正是课堂里的空窗既普遍又无人指认的原因。

十三、与最接近的几种说法的分界

最近发展区说到的是：存在一个凭借帮助可及、独立不可及的能力带，教学应落在这个带里。分离线在**维度**：那一支说的是**难度带**（能力的上下界），本章说的是**时间窗**（同一难度在不同时刻效果天差地别）。判定形式：把难度完全固定，只改时点——若效果只随难度变化而与时点无关，则本判断错而最近发展区足够；若同一难度在窗内与窗外效果断崖式不同，则反之。

先修知识与前提性知识说到的是：后续学习依赖于前置内容的掌握。分离线在**量的性质**：先修是**存量**（有没有），窗口是**时相**（开没开）。判定形式：取一批先修全部达标而窗口已关的学生——先修观点预测他们能学会，本章预测效果与从未打开时相同。这一对照可以用现成的前测数据加一个延迟设计直接做。

关键期与敏感期（Lenneberg, 1967; Knudsen, 2004）是本章最危险的占位者，因为它说的正是“过了这个时间窗就学不会”，形式几乎相同。分离线必须写足，它是本章全部独立性的所在：**敏感期是发育给定的、以年计的、全人类大致相同的、与教学内容无关的**；本章的窗口是上一轮教学产生的、以天或小时计的、因人因课不同的、且可以被改变上一轮内容而移动。判定形式：若窗口的长度与位置只与年龄相关而与上一轮内容无关，则本章所说的就是敏感期的一个下位情形，应当被它吸收；若在同一批学生身上，改变上一轮内容能够移动窗口的的位置，则两者是不同的东西。**第三项兑现正是为这一条准备的。**

间隔效应与复习时点说到的是：同一内容的重复应当分散安排。分离线在**对象**：那一支安排的是**同一内容**的重复时点，本章说的是**不同内容之间**的时序对齐。判定形式：固定同一内容的所有重复时点，只改变紧邻的另一内容——间隔效应预测无差异，本章预测有差异。

掌握学习（Bloom, 1968）说到的是：不达标就重学，直到达标再前进。分离线在**等待期间发生了什么**：掌握学习严格按内容逻辑排序，并让未达标者停在原地补习。判定形式：本章预测**严格执行掌握学习的班级空窗率更高**——因为等待达标的那段时间里，先前打开

的窗口在关闭，而补习的内容是已讲过的那一处，不是窗口所在的那一处。这是一条相反评价：掌握学习判该班级健康，本章判它损失最大。

先行组织者与预备性教学说到的是：在新内容之前提供一个更高层次的框架。分离线在**方向**：那一支是**向后铺垫**（为下一轮做准备），本章是**向前接应**（接住上一轮已经产生的）。判定形式：把先行组织者的内容固定，只改变它与上一轮实际涌现的相关性——若无差异，本章对它无增量。

物候错配（Visser & Both, 2005）是形式层最贴的上游占位者，来自生态学。把本章的命题剥掉学科名词，剩下的是：一个系统对某类输入的响应能力有时限，而输入的到达时序由另一套逻辑决定，两套时序不对齐。这在生态学里叫物候错配——鸟的育雏高峰与毛虫爆发期因不同的环境线索而各自提前，错开天数与繁殖成功率之间有可拟合的关系。这一支给了本章一件现成的方法论：以错配天数为自变量、以产出为因变量做剂量——反应。分离线在**窗口的来源**：生态学里两套时序各由外部环境线索驱动，彼此独立；本章里窗口由上一轮教学自己产生，因此错配在原理上是可以被同一方消除的——这是生态系统做不到而课堂做得到的。判定形式：若改变上一轮内容不能移动窗口，则两者同构且本章只是一个应用；若能移动，则本章的窗口是内生的，而生态学的模型不适用。

带时间窗的调度问题（Solomon & Desrosiers, 1988）是运筹学里的形式占位者，已有四十年历史。分离线极硬：那一支把时间窗当作给定的约束去优化路线；本章说窗口是上一轮路线的产物，因此它不是一个可以事先写进约束集的量。判定形式：若课堂的窗口可以由学生特征与内容特征事先预测（即可写成给定约束），则本章可被这一支吸收；若窗口只能在上一轮结束后才可观测，则它是内生的，那一支的整套解法在此不适用。

反转模板不适用于本章，必须写明。本章的形式不是“越成功越失败”，是“两套时序不对齐”，因此成功陷阱、能力刚性、内卷化这几个模板与本章无关——一门从未成功过的课程，空窗率同样高。这一条排除是必须写的，因为形式层的命题最容易被误挂到这几个名字上。

划完之后再取一次交集。逐条问这几位各自结构性地看不见什么：最近发展区把能力的边界当作可在任一时刻读出的，因此看不见有效期；先修知识把准备度当作存量，因此看不见它会关上；关键期把窗口的来源当作发育给定的，因此看不见教学产生的窗口；掌握学习把“停下来补齐”当作无代价的，因此看不见等待期间的关闭；先行组织者把铺垫的方向当作只能向后，因此看不见接应。这几行背后是同一件被当作给定的东西：

教学内容之间的先后，由内容自身的逻辑关系决定。

所有派别——学科逻辑派、心理逻辑派、螺旋派、项目派——都在争“按什么逻辑排”，没有一派把“排序的依据来自上一轮实际发生了什么”当作可能。因为一旦承认它，课程表这样东西在原理上就不能被完整地预先写出来：它至多只能写出一个候选池与一套接应规则。这是本章动摇的那块地，也是它最不可能被接受的一处。

十四、一个单元走完：窗开着的那三天

取小学六年级“比与比例”这个单元，跟一个班走三天。

周二，窗口打开。教师讲完按比例分配，一个学生举手说：“那如果两边都加一样多，比就变了吧？”——他自己造了一个反例，而且是对的。教师说这个想得很好，我们后面会讲，先把今天这几道题做完。**这一刻打开的是反例窗**，签名清楚：他追问的对象是规则本身，不是题目。按本章的分类，这类窗口以天计，且极易被一句“后面会讲”关闭。

周三，讲的是比例的基本性质。内容按教材逻辑推进，讲得很好，练习正确率百分之九十二。那个学生在这节课里表现正常，作业全对。**这一天里没有任何信号显示有什么东西正在关上**——他不困惑、不出错、不追问，因为今天讲的内容不需要那个窗口。

周四，窗口关上。他不再提关于规则边界的问题了。不是被压制——教师那句“想得很好”是真诚的鼓励；是没有东西到达。他造反例的那个动作在两天里没有一次被使用，于是它退回背景。此后这个班在整个学期里再没有人主动造过反例。

三个月后，期末。一道题给了一个比例关系并暗含一个不成立的条件，班上绝大多数学生直接套用比例式。批改时这被记作“审题不仔细”。而那个周二举手的学生，此时与其余人没有差别。

如果周三的头二十分钟给了它。代价是二十分钟，收益按本章预测有三项：那个反例被全班检验一次（“两边都加一样多”确实会改变比，这本身是比与差的关键分野）；造反例这个动作被公开认可一次，因而其余人也可能开始造；而周三剩下的时间讲比例的基本性质，二十分钟的损失可以从练习里挤出来。**这二十分钟之所以从未被拿出来，不是因为教师不重视，是因为教案上没有一栏叫“昨天开了什么口子”**——教案上写着的是“比例的基本性质”。

这个推演出三条可以被推翻的具体主张。其一，窗口打开的那一刻有清楚的签名，且教师当场就察觉到了（他说了“想得很好”）——**问题不在识别，在没有位置安放**。其二，窗口关上的那两天里一切指标正常，因此任何以当轮指标为准的监测都读不到它。其三，接住它的成本是二十分钟，而失效的代价在三个月后以“审题不仔细”的形式出现、并被归因给学生。三条都可以在任何一个小学数学教研组用现成的课堂录像查证。

十五、与本编另两章的分工

与第八章《先挡住其余的》。那一章处理的是为使一处改变而作的悬置在悬置期内结成的连接，读数是悬置密度与脱悬率。分离线在**轮次位置**：那一章管一次改变的**伴生物**，本章管它的**入口**。判定形式：一个悬置极少、每一次悬置都被拆解的班级，若它的对窗率为零，那一章判它健康，本章预测它仍然低效——因为拆解发生在窗口已关之后，拆的是一堆无人接的东西。

与第九章《还在动的时候就读了数》。那一章处理的是判定时点相对于重组时长过早，读数是窗比与双向翻转率。分离线在**被错开的两样东西**：那一章错开的是**观测与被观测者**，本章错开的是**内容与感受期**。判定形式：一个把所有判定都推迟到重组完成之后的班级，若它的对窗率为零，那一章判它的读数有效，本章预测它读到的是一批从未被接应过的重组。**更要紧的一条相反预测在那一章里已经写明：对窗做得越好，被启动的重组越慢、越要紧，因而越容易被下一次判定冻住**。因此本章的处方若单独实施，会加重那一章所描述的伤害。

三章合起来才说得出的那句：同一次改变有入口、伴生物与出口三处可以失效；三处的读数数学上独立、可同批测量，而效应相乘不相加。本章单独实施会把更多慢速重组送进过早的判定；第九章单独实施会把判定推迟而不改变入口；第八章单独实施会拆解一批从未被接应的连接。三处的次序是有的：先对窗，再拆解，最后才谈判定时点。这个次序是三篇合起来的产物，任何一篇单独都推不出。

十六、五类场景展开，与其中两处收窄

数学的不变量窗。讲完反比例函数打开的是“乘积不动”，而教材的下一节是图像。图像这一节讲得再好也接不住那个窗口——它需要的是“哪些别的地方也有乘积不动”，不是“这条曲线长什么样”。这一格的错配可以从任何一本教材的目录上直接看出，不需要进课堂。

化学的失败滴定。一次没做准的滴定打开的是“终点判断是主观的”这个窗口，它极有价值——分析化学的整套误差意识都从这里进。而实验课的下一节是下一个实验，且那次失败在实验报告里被记为操作失误。这一格的特殊之处在于：窗口不但没被接住，连打开这件事都没有被记录，因为记录格式里只有成功与失误两栏。

编程的第一次递归调通。打开的是“敢于信任自己写的函数”，这是可验证窗的一个变体，时长极短——一两天内合上。而课程按语法与数据结构推进。这一格是四类窗口里最短的一类，因此也是“下一节课再说”这条通行做法失效得最彻底的一格。

陶艺的皮革硬阶段。第一处收窄在这里：这一格的错配不是执行问题，是排课单位与材料时间尺度之间差了两个数量级——窗口以小时计，课程按周排。教师再用心也对不上。工坊的现有处置只有两条：学员自己回来，或师傅代为修坯。前者要求场地与时间自由，后者是代做。这说明本章的处方在某些场合只能通过改变制度安排实现，不能通过改变教师行为实现，而这大幅削弱了本章作为教学建议的可用性。

语言的纠正性反馈之后。一次纠正打开的是“对某个形式的注意”，时长以秒到分钟计，而二语习得对这个尺度有成规模的课堂话语编码。**第二处收窄在这里：**在秒级窗口上，登记与自取机制完全不可行——没有人来得及登记，也没有池子可取。这一格唯一可行的是教师当场改路线，而那要求教师有偏离教案的权限。因此本章的处方按窗口时长分成两套：天与周尺度的窗口用登记与自取，秒与分尺度的窗口只能靠现场权限，两套的制度前提完全不同。

这两处收窄合起来把本章的处方切成三段：秒级靠现场权限，天级靠登记自取，而小时级（工坊那一格）在现行排课单位下无解。本章对第三段给不出办法。

十七、三个领域反过来加上的限制

工程学削掉了本章最自然的处方，并改变了处方的形态。若只看细胞学，本章会写下一句更强的话：课程表应当按窗口重排。这不可行——窗口在上一轮结束前不可知，而课程表必须事先写出来。工程学给了另一条路：公差链的现实解不是让每个零件都精确，是选择性装配——先做，实测，再事后配对。翻回课堂，处方因此从“重排课程表”变成登记与自取：把每一轮打开的窗口登记下来，形成一个可被取用的池子，让学习者在窗内自

行取用相应内容，而不是要求全班同步对齐。这一处削掉的是本章最想主张的那件事，换上的是一件成本低得多、也现实得多的事。

艺术学削掉了本章的价值排序。若只看工程学，本章会写下另一句更强的话：对窗率越高越好。而工坊的经验相反：陶艺的皮革硬窗口如果每一次都被利用，坯体会因反复操作而失败；油画的可覆窗口全部用满会导致涂层过厚开裂。**有些窗口的正确处置是让它关上——某些理解需要一段无人打扰的沉降期。**因此本章不主张最大化对窗率，只主张**对窗率应当被知道**；具体该对哪几个窗，须由内容目标决定，本章给不出应然数值。这一处直接削掉了一个数值目标。

细胞学削掉了本章的适用范围。并非所有分化都需要外部诱导——有一类是自主特化，靠细胞自身的定域因子完成，与外部信号窗口无关。翻回课堂，适用范围因此收窄到**需要外部输入才能形成的那类学习**；纯熟练化（计算、指法、拼写、操作程序）靠重复完成，与窗口无关，把它算进对窗率是记错了账。这一处收窄了本章的射程，而且收得很实：课堂里熟练化占的比重不小。

十八、不适用的边界

三处反向约束已经削掉了本章的价值排序与部分处方。这一节把边界写全，共五条。

其一，纯熟练化内容不适用。计算、拼写、指法、操作程序，这类内容靠重复完成，不依赖外部输入在特定时刻到达。细胞学那一处反向约束（自主特化不依赖诱导）已经削出这一条，而它在课堂总量里占的比重不小。

其二，安全关键的固定次序不适用。实验室安全、临床操作、机械操作与应急程序的教学次序由风险决定，不由窗口决定。这些场合的处置是在模拟环境中安排取用，而不是打乱真实次序。

其三，单次接触的内容不适用。一次讲座、一场报告、一节公开课，没有下一轮，也就没有对窗的可能。本章要求至少两轮连续。

其四，大班且无法个别取用的场合，处方不适用而诊断仍适用。登记与自取要求学习者能在窗内自行取用相应内容，这需要材料、时间与自由度。在四十人以上的同步课堂里，取用成本可能高到使这条处方无法执行——此时对窗率仍可以被测量、被知道，只是改不动。**诊断与处方的适用范围不同，这一条必须写明**，否则读者会因为处方不可行而连诊断一起丢掉。

其五，目标本身就是“按规定次序执行”的内容不适用。规程、法条、标准作业流程，这类内容的价值恰在于次序固定；按窗口重排等于取消它要教的东西。

五条合起来把射程收到：**有连续轮次、内容依赖外部输入才能形成、窗口时长在天与周之间、且学习者有取用自由度的那一类教学。**这个范围不大，本章认为这是应该的——上一节那三段切分已经说明，本章对其中一整段（小时级窗口）给不出办法。

十九、经验研究设计

本章的核心量是对窗率，而它的检验有一个独特的困难：**必须与间隔效应分开**。因此设计的重心不在样本量，在自变量的选法。

核心设计只有一句话：固定本轮内容与时间间隔，只改上一轮内容。这是唯一能把本章与间隔效应分开的设计。若窗口只是时间间隔的函数，改变上一轮内容不应改变本轮效果；若窗口是上一轮的产物，则改变上一轮内容应当改变本轮效果，且改变的方向可由上一轮所打开的窗口预先预测——**预测必须在实施之前写下并封存**，否则这条检验退化为事后解释。

四类条件。A组按内容逻辑常规推进。B组与C组的本轮内容、课时、教师与间隔全部相同，唯一差别是上一轮内容不同：B组的上一轮打开不变量窗，C组的上一轮打开结构同形窗。**D组**实施登记与自取：每轮末尾登记窗口，下一轮头二十分钟留给登记到的那一处，其余照常。

三处对照各回答一个问题。**B与C**是核心检验——两组的本轮完全相同而效果若有差异，差异只可能来自上一轮所开的窗口。**A与D**检验处方的效应。**B、C与A**合起来检验预先预测的准确性：若能预测方向，说明窗口是可辨认的；若只能事后解释，说明它不是。

窗口签名的编码规程是这套设计里最要紧的一环，也最容易做坏。签名在上一轮结束时编码，编码材料只有那一轮的课堂录像与学生的书面产物；**编码者不得看后续任何材料**，否则会顺着结果去认窗口。四类签名各有判据：自发举例且例子跨出本学科（不变量窗）；追问对象是规则而非题目（反例窗）；询问对错的频次下降而自行核验的动作出现（可验证窗）；主动指出两件事一样并说得哪里一样（结构同形窗）。两名编码者独立作答，一致性低于阈值时该轮作废，不取平均。

对窗率的编码要求本轮的起点可指认到上一轮的一个具体产物——一句具体的学生表述、一次具体的追问、一个具体的例子。指认不到具体产物的，一律不计，无论教案上写了什么。

最低成本的先导研究不需要新干预，也不需要进课堂。取教案的版本记录：一位教师的教案在授课后是否被修改、修改处是否指向上一节的某个具体事件。把这个偏离记录与该班的期末迁移题成绩对上，一个学期之内可以完成。第九项兑现预测两者正相关，而这条检验用的全是现成文件。若在这一层就看不到关联，应当优先怀疑本章的机制。

统计上以轮次为单位嵌套于班级、班级嵌套于教师。核心检验不是组均值比较，而是在控制间隔、难度与教师效应之后，上一轮内容对本轮效果的影响是否显著，且方向是否与预先封存的预测一致。

二十、证伪条件与轮次预测

最强的竞争解释是一句朴素的话：**其实不就是趁热打铁、及时复习吗。**这句话把窗口归约为时间间隔，把对窗归约为间隔效应的一个说法。

排除它需要一个只有本机制才预测得到的变量。本章取的是**上一轮内容对窗口位置的移动**：若窗口只是时间间隔的函数，则改变上一轮的内容不应改变本轮的效果；若窗口是上一轮的产物，则**固定本轮内容与时间间隔、只改变上一轮内容**，本轮效果应当改变，且改变的方向可由上一轮所打开的窗口预测。这是一个干净的独有变量，间隔效应在原理上预测不出它。

据此写出七条证伪条件。

其一，固定本轮内容与间隔、只改上一轮内容，若本轮效果无差异，则核心机制为伪——窗口只是间隔的函数。

其二，若同一内容在窗内与窗外的效果差异是渐变的而非断崖式的，则“归零而非打折”这一条不成立，本章退回“最佳时机”这一类程度说法。

其三，若窗口的长度与位置只与学生年龄相关而与上一轮内容无关，则本章所说的是敏感期的一个下位情形，应被它吸收。

其四，若在窗外增加课时可以达到窗内的效果，则第四项兑现失败，窗口不是一个开关而是一个系数。

其五，若教师培训能显著提高对窗率，则第七项兑现失败，错配是执行问题而非结构问题。

其六，若严格执行掌握学习的班级空窗率不高于常规班级，则本章与掌握学习的相反评价被推翻。

其七，若课堂窗口可以由学生特征与内容特征事先预测，则它不是内生的，本章可被带时间窗的调度问题吸收。

第五条最便宜，且所需材料在多数学校已经存在：教案的版本记录与实际授课的偏离记录都是现成文件，把它们与后续的迁移测验对上即可，一学期之内可以完成，不需要任何新干预。

轮次预测。本章的动力主张不是相关，是一个有次序的循环。

第 t 轮教学结束 → 第 $t+1$ 轮出现某类自发涌现（学生主动举例、主动提问、在无关任务里迁移使用该结构；这是窗口开启的可观测签名）→ 第 $t+2$ 轮若相应内容到达则该处出现与投入不成比例的快速进展、若未到达则该类涌现停止 → 第 $t+3$ 轮的发起处或者转向学生的涌现（对窗），或者继续按内容序号推进（空窗），而后者使下一轮的窗口更难打开。

四步中任何一步顺序系统性相反，模型即需修正：若自发涌现在教学之前就已存在，则窗口不是上一轮产生的；若涌现的停止与内容是否到达无关，则第三步不成立。

二十一、备课可以怎么改：四个动作

本章首先是一项解释，不是教学法。但一旦把空窗看见，备课的动作会变。

第一，上一节末尾留三分钟登记窗口。成本最低的一件：下课前问一句——现在你们最想问的是什么、有没有想到别的地方也有这回事。把答案原样记下来，不评价、不纠正、不展开。这三分钟不需要额外课时，它换的是一份关于窗口的原始记录，而这份记录目前在任何学校里都不存在。

第二，下一节的头二十分钟留给登记到的那一处。不是整节，是头二十分钟。这一条的关键在于它不要求重排课程表——课程表照走，只是每节课的前二十分钟改由上一节的产物决定。二十分钟是本章能想到的最小可行单位：短于此接不住，长于此进度撑不住。

第三，建立候选池而不是重排课程表。这是工程学那一处反向约束的直接后果，也是本章与“按窗口排课”分道的地方。窗口在上一轮结束前不可知，因此课程表在原理上写不出来；能写出来的是一个候选池——每一个知识点配一到两个可以接住不同窗口的入口材料，登记到哪一类窗口就取哪一份。池子可以事先做好，取用是现场的。

第四，给教师偏离教案的权限，并把偏离记录下来。第九项兑现预测：**一个从不偏离教案的教师，对窗率在结构上接近于零。**因此这条处方的第一步不是培训，是允许——允许教师因为学生的某个具体涌现而改动当天的安排，并且把改动记进教案版本而不是当作教学事故。记录同时是读数的来源。

四条里前两条是备课习惯，第三条是教研组的工作，第四条是管理权限。这一点本身可检验：若在完全不改变教案权限的条件下，仅按前两条运行的班级与对照班无差异，本章的处方在缺少第四条时失效。

二十二、反噬预言与伦理限制

对窗率一旦进入考核，最省事的达标办法会摧毁它。

第一种是**把下一节的引入写成“承接上节”**。教案上出现“复习上节内容”“在上节基础上”，即可计入对窗——而承接内容标题不是对窗。判定必须写死在**产物**上：对窗要求本轮的起点可指认到上一轮实际产生的某样具体东西（一个学生的具体表述、一次具体的涌现、一个具体的疑问），不是上一轮的内容标题。

第二种是**制造涌现**。教师可以在上一轮末尾安排一个必然引发某类提问的环节，使窗口的开启变得可预测、可排进教案——于是窗口重新变成给定约束，本章的整个主张被抽空。这一种最难查，因为被制造出来的涌现与自发涌现在记录上完全一样。

第三种是**把登记池做成一份从不被取用的清单**。窗口登记表挂在系统里、进合规检查、计入教研考核，而实际取用率为零。

对此本章看不到彻底的出口。能规定的只有两条：**对窗的计入以可指认的上一轮产物为准；登记池的验收看实际取用率，功能存在而取用率为零，与没有这个池子是同一格。**第二种反噬本章没有对策，只能把它写在这里，让使用者知道自己在做什么。

因为对窗率是一个可以拿去考核人的比率，伦理限制必须写死三条。

第一，这个读数指的是**内容时序与窗口时序的关系**，不是教师的用心程度，也不是学生的悟性。一个严格按大纲执行、对窗率为零的教师，可能正是最尽责的那一个。

第二，**不得为了提高对窗率而在安全关键教学中打乱必须固定的次序**。实验室安全、临床操作、机械操作与应急程序的教学次序由风险决定，不由窗口决定；这些场合的处置是在模拟环境中安排取用。

第三，若学校已据这个读数作出不利安排，**必须公开编码规则、原始证据与复核通道**，并允许被评价者提交上一轮产物的记录作为反证。

二十三、自反

按本章自己的判据，这篇文章的对窗率为零。

它完全按内容的逻辑关系排：先讲三家，再讲共有前提，再命名对象，再给辨别格、读数、分界、限制、证伪。它没有对读者上一轮打开的任何窗口作过对齐，也不可能作——它不知道读者上一轮读了什么，更不知道那次阅读在他身上打开了什么。**本章主张课程表在原理上不能被完整地预先写出来，而它自己正是一份预先写好的清单。**

这一点有一个具体后果，而不只是一句谦辞：**本章最可能被接住的时刻，是读者刚刚经历过一次错开之后**——刚发现学生迁移不动、刚为一次“讲过了怎么又不会”发愁的那几天。而本章出现在读者面前的时刻由出版与检索决定，与那几天无关。因此本章对自己的预测是：**它的绝大多数读者会在窗外读到它，读完觉得有道理，然后什么也不改变。**

本章能为自己做的只有一件：把读数与那条最便宜的检验写得比结论更具体，使它在读者的窗口真的打开时，还能被取用一次。这正是第六项兑现所说的登记与自取——本章把自己放进池子，而不是要求读者此刻就对齐。

本章可以被分层引用。第一层是“空窗”作为一类对象，最强也最易倒。第二层是那张二维辨别格与其中“两格后果不对称”的观察，即使命名不成立仍可作为诊断工具。第三层是一条不依赖前两层的经验主张：**固定本轮内容与时间间隔、只改变上一轮内容，本轮的学习效果会改变。**第三层最稳，也最容易检验。

二十四、为什么常见的改革处理不了这一件

近三十年里，重排课程顺序、螺旋式设计、项目式学习、个性化路径、加强学情分析，每一样都被推行过。本章给出的解释是形式上的：**这五类改革改的全是内容一侧的安排，而窗口一侧一个都没动。**

重排课程顺序改的是内容之间的先后，而窗口不在内容里。无论按学科逻辑排还是按心理逻辑排，排的都是内容；窗口由上一轮实际产生了什么打开，它在排课时不可知。排得再好，两套时序仍然各走各的。

螺旋式设计改的是重访的次数，不改重访的时刻。同一主题在不同学段以更高复杂度重访，重访的时点由学段决定，与窗口无关。因此螺旋设计增加的是碰上窗口的概率而不是对齐；按四类窗口的时长（天到周），跨学段重访碰上的概率接近于零。

项目式学习改的是任务的长度，不改起点的来源。一个持续数周的项目里，起点由项目设计决定，中途每一步由里程碑决定。项目更长，窗口在项目内部打开又关上的次数更多，空窗率可能反而更高——这与该支的自我评价相反，也是一条可查的预测。

个性化路径改的是路线的分叉，不改分叉的依据。算法按作答结果分流，而作答结果记录的是内容点的对错，不是窗口的开合。个性化推送的是“他还不会的那些”，而窗口恰恰开在“他刚刚会了的那一处”旁边——两者在原理上指向相反的方向。

加强学情分析改的是准备度的读数，而准备度不是靶方的属性。这是本章与整个学情分析传统最直接的分歧：前测、诊断、画像，测的都是学生此刻是什么样的学生；而按序贯诱导，此刻的感受态是上一轮的产物，且会在几天内改变。测得越精细，测的仍是一个正在变化而它假定不变的量。

五类改革共享一个特征：**它们全都试图把内容排得更好，而没有一样把“起点由谁决定”当作可以改动的东西。**这不是理念问题——课程表必须事先写出来，是学校这个组织得以运转的条件。本章因此不主张取消课程表，只主张在它之外再留一份候选池。

二十五、分层引用，与三个本章解释不了的洞

本章可以被分层引用。**第一层**是“空窗”作为一类对象，最强也最易倒——只要证明窗口与上一轮内容无关，这一层就没有了。**第二层**是那张二维辨别格，尤其是其中“两格后果不对称”这一观察：窗口未开而内容到达会产生信号，窗口已开而内容未到达不产生信号——即使命名不成立，这个不对称仍然解释了为什么课堂的纠错能力只对其中一格有效。**第三层**是一条不依赖前两层的经验主张：固定本轮内容与时间间隔、只改上一轮内容，本轮的学习效果会改变。第三层最稳，也最容易检验。

有三个洞本章解释不了。

第一个洞：本章说不出窗口的时长由什么决定。四类窗口的时长差了两个数量级（秒到周），本章只能描述这个差别，说不出它由什么决定——是内容的性质、学习者的先备结构，还是打开时那次经验的强度。若时长主要由学习者决定，“登记与自取”这条处方要按人定制，成本会高到不可行。

第二个洞：本章分不清“窗口关上”与“窗口从未打开”。两者的事后表现完全一样：都是没有涌现、都不产生信号。本章用上一轮结束时的签名来判定窗口曾经打开，而签名是概率性的证据——一个内向的学习者可能打开了窗口而不举例、不追问。这使对窗率在沉默的学习者身上系统性低估，而那恰恰是最需要被看见的一批人。

第三个洞：本章无法排除窗口是被观察行为本身制造的。要登记窗口，须在上一节末尾问一句“你们最想问什么”；而这一问本身可能就打开了一个窗口，或把一个尚未成形的涌现固定成一句话。若如此，登记与被登记之物不可分，对窗率测的是登记动作的产物而不是教学的产物。本章没有办法把两者分开——不问就没有记录，一问就已介入。

二十六、三条本章预测而尚无人测的现象

一个判断若只解释已知，仍可能是旧说法的新名字。本章因此列出三条它预测、而据本章所知尚无人测过的现象。每一条不成立都要动本章的一部分。

其一，项目式学习的空窗率应当高于讲授式课堂，而不是更低。通行的预期相反——项目更开放、更以学生为中心，理应更能接住学生的想法。而按本章，项目的长度使窗口在项目内部打开又关上的次数更多，而里程碑与评价量表把每一步的起点都事先定死了，因此每一个打开的窗口都要等到里程碑允许时才可能被处理。这条预测与项目式学习这一支的自我评价直接相反，且检验成本不高：在同一所学校里，对项目班与常规班各做四周的窗口签名编码与对窗率编码即可。若项目班的对窗率显著更高，本条为伪。

其二，一个教师的对窗率应当与他偏离教案的频次正相关，而与他的教龄、学历、教学信念问卷得分无关。这条预测把本章与“好教师更用心”这个最朴素的竞争解释直接分开。按本章，对窗需要在现场改动当天的安排，因此它的必要条件是偏离；而偏离的频次是一个可以从教案版本记录中直接读出的量。若对窗率主要由教龄或信念得分解释，本章所指的就是教师素质的一个别名。这条检验只需要教案的版本记录，多数使用电子备课系统的学校都有。

其三，沉默的学习者其对窗率的低估幅度，应当与班级规模正相关。窗口的签名依赖学习者外显的举例与追问，因此内向者的窗口更难被登记；而班级越大，教师能注意到的

外显信号越集中在少数人身上。这条预测的检验需要一个独立的窗口证据来源——例如书面的课后一分钟提问条，它使沉默者的窗口也留下痕迹。把书面渠道测到的对窗率与课堂口头渠道测到的对窗率相比，两者的差值应当随班额上升。若差值与班额无关，说明沉默者的窗口并未被系统性漏掉，本章关于登记偏倚的担心不成立——这对本章是个好消息，但它同时取消了本章关于“为低可见度表达提供书面通道”这条处方的理由。

三条都不需要新的测量工具，只需要把已经存在的材料按不同的方式记一次。

二十七、结论与一条写死日期的赌注

课堂上学习很少真正到位，不是因为内容顺序排错了，不是因为学生基础差，也不是因为教法不够先进。更靠前的原因是：**教学序列按内容的逻辑关系排，而感受窗口按上一轮实际产生了什么开合；两套时序互不迁就，而课堂只记录前一套。**窗口关上时不产生任何事件，因此这件事在原理上不会以“失败”的形式被发现。

三个领域各出一份不可替代的力。细胞学给出两条最承重的形式约束——**准备度不是靶方的属性，是上一轮的产物；以及它有时限，窗外不是打折是归零。**工程学给出了处方的形态：不是让每一件都精确，是事后配对，因此课堂的第一处方是登记与自取，不是重排课程表。艺术学则削掉了数值目标：有些窗口的正确处置是让它关上，因此本章只主张对窗率应当被知道，不主张它越高越好。

这一理论把备课的问题从“这一节该讲什么”改为“上一节在他们身上打开了什么”；把课程设计从“写一张完整的表”改为“写一个候选池加一套接应规则”；把归因从“学生不会灵活运用”改为“那几天窗开着而我们在讲别的”。

本章给出一条写死日期的赌注。**到 2031 年 12 月 31 日之前**，若至少三项在同一制度环境内、不少于八个平行班开展的研究报告：**在固定本轮内容与时间间隔、只改变上一轮内容的条件下**，本轮的学习效果出现显著差异，且差异的方向可由上一轮所打开的窗口预先预测，则本章的核心判断获得初步支持。

以下明确**不算命中**：仅改变时间间隔；仅改变内容难度；仅比较不同教材或不同教师；仅采用问卷自报“学生兴趣更高”；仅比较两个地区或两个国家；把窗口对齐与增加课时、增加练习量打包实施而无法分离贡献；把教案中的“复习上节”直接计为对窗。若符合上述条件的多项研究均未发现效果差异，本章应被放弃或降格为一个描述性提法。

课堂不需要放弃课程表——没有表就没有课。它需要的只是：在按表推进之前，先花二十分钟看一眼昨天在他们身上开了一个什么口子，然后决定今天要不要先走那一处。

接续。对窗率是六个读数里最难采的一个，附录一为它单列了两步走的记录法，并写明第一步若写成空话（“兴趣提高了”一类），这个读数会虚高到接近满分。本章与第八、九章三处如何**相乘不相加**，见第三编小结与第十四章；本章那条写死到二〇三一年的赌注与六种不算命中的情形，已汇入附录三。

第八章 先挡住其余的

判断卡片：为使一处改变而挡住的那部分不会停在原地——它在悬置期内结成连接，脱悬后无人拆除，也不进入任何账本，却在换情境时决定表现。

三个理论入口：教育学 × 化学 × 艺术学

提要。课堂上学习很少真正到位，通行的解释有三类：学生的原有观念没有被充分动摇、教学路径安排不当、评价与进度压制了深度。本章提出第三类之外的一个位置。要让某一处发生改变，课堂必须先把其余部分挡住——限定题型、简化情境、给出范例、说一句“这个以后再讲”。教育学把这种悬置读作认知负荷的合理分配，化学把它读作合成里的保护基策略，艺术学把它读作作品的未完成状态。三家争的是改变由什么驱动，而共同假定了一条没有说出口的话：**在一次改变里，被挡住的那部分不参与，参与名单是给定的。**推翻这条的材料来自化学自己——酰基迁移：用来挡住一个位点的那个基团会自己跑到另一个位点上去，挡住这个动作改写了它本要保护的那张名单。带回课堂便得到本章命名的对象：**悬置连接**——在悬置期内，被悬置的那部分与新学内容之间自行结成、脱悬后不被拆除、也不进入任何账本的关联。承重判断是：学习之所以少，不是那一处该变的没变，而是每一次悬置都会附带产出一批未被指认的连接，它们在换情境时决定表现，而课堂的账本按内容点记，看不见内容点之间的东西。本章给出悬置密度与脱悬率两个读数、一张二维辨别格、一句不含判断词的问法、与九种相近说法的分界、三处反过来加上的限制、七条证伪条件与一条写死日期的赌注。

一、问题从哪里来

一堂课要讲清楚一件事，几乎总要从一句“暂时先不考虑”开始。物理课说暂时不考虑摩擦，数学课说负数开方以后再说，化学课说这个反应的机理这里不展开，编程课说错误处理先跳过，医学课说这个鉴别诊断到临床再讲。这些句子不是偷懒，它们是教学得以进行的条件——不挡住其余部分，那一处就淹没在其余部分里，什么也讲不成。

问题出在挡住之后。课堂的账本按内容点记：讲了哪几点、测了哪几点、掌握了哪几点。被明确挡住的那部分不在这一轮的账上，等它自己那一轮到了，才被重新记上。两轮之间，这部分被默认为静止的——放在那儿，等着以后来取。

而学生并不按这个账本运行。被挡住的那部分并没有停在原地：它与刚学的内容之间会自行接上一些线。物理课里“暂时不考虑摩擦”的学生，在往后的相当长时间里把“斜面”与“匀加速”直接接在一起，接得又快又稳；等摩擦被正式引入，这条线并不会因为新知识的到来而自动断开，它只是多了一个需要判断何时启用的例外。数学课里被推迟的负数开方，与“方程无解”接成了一条线；复数被引入时，“无解”这个词已经在学生那里承担了太多别的含义。

这些线不是错误。它们不会在原题型里产生任何差错，因此不会被批改抓到；它们只在换一个表面不同的情境时决定表现，而那时通常已经被归因为“基础不牢”。本章要做的，是把它们从“基础不牢”里分出来，给它一个位置、一个读数与一组可以推翻它的观察。

本章不报告新采集的数据，是一篇理论性文章。它的经验主张需要预注册研究与独立复核检验。

二、教育学的单因主张：不相容驱动路径改变

教育学关于深度学习最强的一条单因主张来自概念转变一支：决定学习是否发生的，不是信息是否到达，而是学习者原有的解释与新经验之间是否发生了不相容，以及这种不相容是否达到了迫使原有解释重新组织的程度。Posner 等人给出的四条条件——对原有解释不满、新解释可理解、可信、有生产力——把这条主张写成了一个可以逐条检查的清单（Posner et al., 1982）。

这一支的时序读数是清楚的：**先出现可辨认的不相容，随后出现解释路径的改变，最后才出现可迁移的结果**。它据此把大量教学失败诊断为不相容不足：教师讲得太顺，学生的原有解释从头到尾没有被逼到墙角，因此路径没有理由改变。

单因锁定成立——这一支不认为条件与路径是并列因素。它明确主张，条件安排的全部意义在于制造那次不相容；不相容不发生，条件再好也只是把知识堆在旧结构上。

它自己知道、却不往这个方向用的一件事，在同一支的修正稿里：Strike 与 Posner 后来承认早期模型过于理性主义，学习者的动机、情境与发展状态同样起驱动作用，而且原有概念的整个生态——**包括那些与目标概念看似无关的部分——都参与了对新经验的接纳与排斥**（Strike & Posner, 1992）。这句话被这一支用来说明模型需要变得更丰富；它没有被用来问一个更麻烦的问题：**在教学明确把生态的一部分挡住的时候，被挡住的那部分是否仍在参与**。

三、化学的单因主张：路径与条件的不容驱动产物改变

合成化学处理的是同一个问题的一个更硬的版本。一个带多个官能团的分子，遇到一种对所有位点都无差别进攻的试剂，得到的是一锅混合物。要让其中一处发生反应，办法不是提高驱动力，而是**先把其余位点暂时挡住**：装上保护基，做完这一步，再脱掉。保护与脱保护的完整策略是有机合成的日课，也是它的教科书主干（Wuts & Greene, 2007）。

这一支的单因主张是：**决定一个多官能团分子能被改造成什么的，是哪些位点在这一步被挡住了，而不是试剂有多强、分子当下是什么样子**。选择性来自阻断，不来自驱动。它的时序读数是：先安排保护序列，随后特定产物出现，产物的官能团分布又决定下一步能装什么、脱什么。

这一支自己知道、却不往这个方向用的事实有两件，第二件是本章全部论证的支点。

第一件是代价：每加一对保护与脱保护，总收率下降、步骤增加。有一整支合成化学在追求**不使用保护基的路线**，并且拿出了完整的天然产物全合成作为证明（Baran, Maimone & Richter, 2007）。这一支用它来说明工艺效率，没有用来说明“挡住”这个动作本身有代价。

第二件是**酰基迁移**：装在糖分子某个羟基上的酰基保护基，会在碱性或质子性条件下自己迁移到相邻的羟基上去。挡住二号位的那个东西，一段时间之后出现在三号位。合成

化学把它当作需要避免的副反应，写在操作注意事项里；它从来不被当作一条关于“参与名单”的证据。而它恰恰说的是：**被挡住的那部分不是静止的，挡住这个动作本身会改写它要保护的那张名单。**

四、艺术学的单因主张：已做出来的东西与未走完的路径不容，驱动条件改变

艺术学提供第三个位置，而且它处理的正是“挡住之后”那一段。一件雕塑停在半路上，一幅画的一角留着底色，一件器物的口沿没有修整——它是废品、是草稿、还是作品？这个问题在未完成状态的研究里有成熟的答案：**它算什么，不由作者的意图决定，也不由技术完成度决定，而由它被放进什么样的展示与流通条件决定。**同一块石头在工作室里叫未完成，进了美术馆就叫作品（Schulz, 1975）。

这一支的单因主张是：**已经做出来的那部分与尚未走完的制作路径之间的不容，驱动的是“什么算完成”这个条件本身的改变。**时序读数是：先有半成品与未竟工序之间的张力，随后完成判定的标准改变，改变后的标准又反过来决定往后怎么做——一旦“未完成也可以是作品”被接受，此后的创作会主动留下未竟处。

单因锁定成立：这一支不认为完成判定是意图、技术与制度共同作用的结果。它主张判定权在条件一侧，且是决定性的。

它自己知道、却不往这个方向用的事实是：许多被称为未完成的作品，作者当时并不认为它未完成；未完成这个身份是后世的条件追加给它的。这一支用它来讨论作品身份的历史性；没有用来问：**被判定为“以后再处理”的那部分，在等待期间是否已经与已完成的部分长在一起，以致后来根本无法单独处理。**

五、三家为何不能同时为真

三家把决定性的驱动放在三个不同的位置：教育学放在原有解释与新经验的不容，化学放在序列安排与无差别条件的不容，艺术学放在已做出的部分与未竟路径的不容。若温和地并列，会得到一句没有风险的话：学习需要合适的冲突、合理的顺序与恰当的完成标准。这句话正确，但任何失败都可以事后归入其中一条。

三家实际上互相取消。

化学的日课正是教育学诊断的病根。合成的整套本事是**先挡住其余部分**；而按概念转变，不相容之所以发生，正因为新经验与原有解释生态里的其余部分相遇。把其余部分挡住，等于在原理上取消不相容的发生条件。若化学那一条是对的，概念转变的处方——制造认知冲突——在结构上不可能被执行，因为可执行的教学必须先做一次挡住。

艺术学的日课正是化学判为工艺失败的那件事。让未完成状态保持可见、交由后来的条件判定，在合成里叫作**脱保护没做完就出料**：留在分子上的保护基会改变它此后一切性质。若艺术学那一条是对的，“以后再讲”就不是欠账而是一种正当的作品形态，化学关于代价的整套计算失去对象。

教育学的日课又正是艺术学判为病的那件事。教学要求在单元结束时给出完成判定（会了/没会），而艺术学主张完成判定应当由后续条件给出、不应在制作现场作出。若教育学那一条是对的，未完成研究关于“身份由条件追加”的说法在教学里不适用。

三家必须有两家错，或者三家的病灶都放错了位置。本章主张后者。

六、共有前提，与推翻它的那件材料

三家争的是一次改变由什么驱动。而它们共同假定了一条从未说出口的话：

在一次改变里，谁参与、谁不参与，是给定的；被挡住的那部分在这段时间里不动。

概念转变假定与新经验相遇的是那个相关概念，其余生态是背景；保护基策略假定被装上保护基的位点在这一步就是不参与的；未完成研究假定作品的哪些部分已经定了、哪些还开着，是可以指认的。把这句话摆到三家面前，三家都会觉得不值一提——它不是谁的主张，是三家脚下同一块地面。

推翻它的材料在化学自己手里，而且是操作手册上的常识：**酰基迁移**。装在一个羟基上的保护基，在温和条件下自行迁到相邻羟基。这意味着三件事，逐件都与共有前提相反：

其一，**被挡住的位点不是静止的**。它在被保护期间与相邻位置发生了实际的化学关系。

其二，**改写名单的正是挡住这个动作自己**。若没有装上那个保护基，就没有可迁移的东西；参与名单不是被外力打乱的，是被保护策略自身产生的。

其三，**结果在形式检查上是合格的**。迁移之后分子仍然带着一个保护基、仍然只有一个位点暴露、仍然可以进行下一步反应——只是暴露的不是原来打算暴露的那个。**错误不表现为失败，表现为一个位置不同的正确产物**。

这三条一起带回课堂，共有前提塌掉。挡住某部分内容，不等于那部分内容在这段时间里不参与；恰恰是挡住这个动作，使它与正在学的内容之间生出了本来不会有的关系；而这种关系不产生错误，只产生一个“位置不同的正确答案”。

七、悬置连接

据此本章命名一个对象。

悬置连接：在教学为使某处改变而暂时悬置其余部分的期间，被悬置的那部分与正在学的内容之间自行结成的关联；它在悬置解除后不被拆除，也不进入任何账本。

四个必要条件缺一不可。第一，存在一次**明确的悬置**——教学明白地把某部分放在一边（暂不考虑、以后再讲、这里先按这个来）。第二，悬置期内学习者在被悬置项与新内容之间形成了**可指认的关联**，它体现为一种稳定的解释、一条捷径规则或一个不再被质疑的默认。第三，悬置解除时，教学**只补内容不拆连接**——把被悬置的那部分单独讲一遍，不要求学习者回到当初那道题上把两者重新接一次。第四，**不产生信号**：在原题型内这条连接不导致任何错误。

三条性质使它区别于邻近的坏情况。

它不是“没学会”——当轮可以完全正确、熟练、迅速。它不是“迷思概念”——迷思概念是学习者从课外带进来的，悬置连接是课堂制造的，且它的位置可以从教案里预测。

它不是“惰性知识”——惰性知识说的是知识调不出来，悬置连接说的是知识被调出来时带着不该带的东西。

需要排除两类伪阳性。其一，**明写的近似**：教学说明这是一个在特定条件下成立的近似，并给出近似失效的判据，此时被悬置的不是内容而是复杂度，学习者手里有一个可用的边界。其二，**已完成的拆解**：后续教学要求学习者用被悬置的那一部分重做当初那道题，并说明当初的答案在何种意义上仍然成立，此时连接被拆过一次。

也要排除一类伪阴性：**复习不是拆解**。把被悬置的内容单独讲一遍、测一遍，学习者可以完全掌握它，而当初那条连接原样不动——因为从来没有人要求他把两者放在一起重新判断一次。

八、悬置连接的四种亚型

悬置连接并非一种东西。按被悬置的是什么，可以分出四类；四类的形成机理相同，而拆解成本与失效形态差别很大。

条件悬置。被挡住的是一个使命题为真的条件——暂不考虑摩擦、忽略空气阻力、假设理想气体、设电阻不随温度变化。它结成的连接是“这个结论无条件成立”。要紧的是这条连接的形态：学习者并没有记住一个错误的结论，他记住的是一个**正确的结论加上一个不存在的适用范围**。失效形态因此很特别——他在新情境里不是算错，是**不去查条件**。条件悬置是四类里最容易拆的：回到当初那道题，把条件加回去重算一遍，两个答案并排放着，适用范围就出来了。

范围悬置。被挡住的是定义域——负数开方以后再说、这里只讨论正整数、这个方程无解。它结成的连接不是“暂时不讨论那些情形”，而是“那些情形不存在”。两者的差别在学习者那里几乎不可见，而后果完全不同：等新范围被正式引入，旧结论不会被自动重判，因为在他那里旧结论从来没有过边界。拆解要求回到当初那个“无解”的判断上重新表态——这一步在多数教材里从不出现，复数被引入时通常另起一章。

机理悬置。被挡住的是为什么——记住这个公式就行、机理这里不展开、按这个步骤做。它结成的连接是公式与题型之间的直接映射，且这条映射极其牢固，因为它在原题型里从不出错、反馈永远是正的。失效形态是断崖式的：题型稍变即完全失效，而且失效时学习者无处可查——他手里没有任何可以据以重新推导的东西。机理悬置的拆解成本最高，因为它要求补一段完整的推导，而那通常意味着重上一节课。

判断悬置。被挡住的是判断标准本身的由来——按老师说的这个标准来、先照这个模板写、这样写就能得分。它结成的连接是标准与权威之间的绑定：学习者掌握的不是判断，是“谁说了算”。失效形态最隐蔽——换一个评价者即失效，而这在同一门课里几乎不会发生，因此它可以一直不被发现，直到升学、转学或进入职场。判断悬置是四类里最难拆的，因为拆解要求学习者重新审视一个他从未被允许审视的标准，而这件事在评价权集中的场合对他并不安全。

四类可以叠加。一道物理应用题可以同时悬置条件（不计摩擦）、悬置机理（公式直接给）与悬置判断（答案对不对以标准答案为准）。因此不能把一节课整体判为高悬置或

低悬置，必须**按位置记**：同一节课在一处是条件悬置、在另一处是判断悬置，两处的拆解成本相差一个数量级。

这一分类的实用价值在于它改变了拆解的优先次序。若课时有限，先拆的不该是最难的那一处，而应是**拆解成本最低而失效范围最广**的那一类——按上面的形态，那是条件悬置。判断悬置虽然最严重，但它的拆解需要评价结构本身的改变，不是一节课的事。

九、二维辨别格

要把这个对象从一个说法变成辨别工具，必须把它与教学质量分开。横轴取**悬置密度**：一个单元里明确悬置的次数。纵轴取**脱悬率**：这些悬置里，后来被明确拆解（要求学习者用被悬置项重做当初那件事）的比例。两轴相互独立——讲得越清楚的课堂悬置密度越高，而脱悬率与讲得清不清楚无关。

	脱悬率低	脱悬率高
悬置密度低	直陈型：内容以完整复杂度呈现，学习者自行取舍；当轮吃力，遗留问题少而散	慢工型：悬置少且都被拆过；进度慢，换情境表现稳定
悬置密度高	叠置型：讲解清楚、进度顺利、正确率高，而悬置期内结成的连接从未被拆。本章的靶格	高负荷型：悬置多且都要拆，课时不够用；实际罕见，多数会退回叠置型

- 右下角是靶格，它有三条使它难以被发现的性质。
- 其一，**短期指标表现为改善**。悬置正是使当轮清楚的那个手段。悬置密度越高，讲解越顺、正确率越高、学生的主观理解感越强。任何以当轮读数评价的体系都会奖励它。
- 其二，**失效不产生信号**。悬置连接在原题型内不导致错误。它只在换一个表面不同、结构相同的情境时决定表现，而那时的归因几乎必然是“基础不牢”“不会灵活运用”——两个归因都指向学习者，不指向那次悬置。
- 其三，**它自我加固**。保证进度与正确率的唯一办法是多悬置；悬置越多，可用于拆解的课时越少；脱悬率越低，换情境表现越差；表现越差，越需要用更多练习保住当轮正确率，于是悬置更多。
- 左上角是关键对照格：**内容以完整复杂度呈现的课堂当轮更吃力、正确率更低**，而本章预测它的换情境表现优于靶格。这一格的存在使本章与“讲清楚有害”划清界线——问题不是讲清楚，是讲清楚所必需的那次悬置没有被拆。

十、读数与一句不含判断词的问法

悬置密度从教案与教材直接可数：明确的悬置语句（暂不考虑、先不管、以后再讲、这里只看、按这个来就行）出现的次数。这是一个纯外部量，不需要任何人配合。

脱悬率要求更严：一次悬置算作被拆解，必须满足三条——被悬置项后来被正式引入；学习者被要求回到当初那件事上重做；重做时须说明当初的答案在什么意义上仍然成立、在什么意义上不再成立。三条缺一不可。只讲不接不算，只测不接也不算。

判据写成一句不含判断词的问话：

这学期明确说过“这个以后再讲”的地方，有几处后来真的讲了，并且要求学生把它和当时学的那一处重新接了一次？

这句话可以拿去问教案、教材目录与作业记录，不需要问任何人的判断。五个场景跑一遍，答案互不相同：

场景	典型悬置	后来是否拆解
中学物理·摩擦	本题暂不考虑摩擦	多数教材中摩擦力单独成节，不回溯当初那道题现在怎么算；查目录与习题即可确认
中学数学·负数开方	这个方程无解	复数引入时通常另起一章，很少回到当初那个方程重判
化学·反应机理	机理这里不展开	大学阶段引入，跨学段，几乎不可能回接
编程教学·错误处理	先跳过异常	视课程而定；工程实训类课程常在后期强制回改早期代码，这一格的脱悬率可以很高
陶艺与油画技法	这一层先这样，干了再说	工坊制传统里下一次上手就是回到同一件坯或同一张画，结构上强制回接；查工序单可确认

其中第一格与第五格是查出来的，不是从命题推出的：教材的章节编排与工坊的工序单都是现成文件，任何人都可以核对。第五格尤其要紧——它说明高脱悬率不是理念问题，是工作对象是否持续存在的问题。课堂的作业交上去就不再回来，而坯与画一直在那里。

十一、十处兑现

跨领域的关系只有能在各自领域里推出一件具体的事，才算兑现。

其一，教育学的原有概念对应被悬置项的具体内容，因此研究必须记录教学明确悬置了什么，而不只是记录讲了什么。其二，不相容对应悬置解除时的重新判断，因此拆解的证据是学习者在重做时说出了当初答案的适用边界。其三，四条件里的“生产力”对应拆解后新解释能否处理更多情形，可用变式题直接测。其四，保护基对应课堂的悬置装置，因此它可以被数——教案里的悬置语句就是它的清单。其五，脱保护对应拆解工序，因此脱悬率与合成的脱保护收率同构，都可以按“进去多少、出来多少”计算。其六，保护基迁移对应悬置连接本身，因此本章预测悬置连接的位置可以从被悬置项的相邻内容预测，而不是随机分布。其七，无保护基合成对应“不需要悬置的入口”，因此本章的第一处方不是补拆解，是改入口。其八，未完成状态对应长期悬置项，因此某些悬置不应被拆解，判据是课程是否声称该内容最终要能独立运行。其九，完成判定由条件追加，对应课堂的单元结束判定，因此本章预测把判定推迟到跨单元会改变学习者对“哪些还开着”的认知。其十，作者当时不认为未完成，对应学习者不认为自己欠着账，因此悬置连接不能靠自评问卷测量。

第六、第七与第十项是本章最依赖的三项。删掉第六项，本章的对象退回“没学扎实”；删掉第七项，处方退回“多补课”；删掉第十项，读数退回自评量表。

十二、一个单元走完：从悬置到连接到拆解

抽象的判据要放到一段具体的教学里才看得出成本。取初中物理“力与运动”这个单元。

第一节课作出了三处悬置。教师讲牛顿第一定律，用了小车在水平面上滑行的例子，说了三句话：本题不计摩擦；空气阻力忽略；小车视为质点。三句话都必要——不说这三句，第一定律在中学阶段讲不成。三句分属两类：前两句是条件悬置（可拆），第三句是不可拆的理想化（按边界那一节，判据须改写为“知不知道这是一个模型”）。

第二到第四节课形成了连接。学生做了三十道同型题，全部不计摩擦，全部得出“不受力则匀速”。这期间形成的连接是：**运动状态不变，等同于合外力为零，等同于没有力。**第三个等号是多出来的，教材从未写过它，教师也从未说过它——它是被三十道题的一致经验接上去的。这条连接在这三节课里从不出错，因为每一道题里都没有摩擦力。

第五节课引入摩擦，而通常不拆解。摩擦力单独成节，讲定义、讲方向、讲计算。学生学得很好，测验正确率高。而第一节课那三十道题没有一道被拿回来重算，“合外力为零”与“没有力”之间那个多出来的等号从未被拆开。此时学生手里有两套都正确的东西：一套关于第一定律，一套关于摩擦力，中间没有连线。

失效在两个月后。期末考一道题：木块在水平面上匀速滑动，问受力情况。按第一定律，匀速则合外力为零；按摩擦力那一节，滑动必有摩擦力。相当比例的学生答“不受力”或者答“受重力和支持力”——不是他们不会算摩擦力，是那个多出来的等号还在。批改时这被记作“综合运用能力不足”。

拆解要花多少课时。按本章的判据，拆解要求回到第一节课那道具体的题上，加上摩擦重算一遍，并说出当初那个答案在什么意义上仍然成立（在摩擦可忽略的尺度上成立）。这件事需要的不是一节课，是**十分钟**——因为那道题学生早已熟悉，要做的只是把一个条件加回去。十分钟对一个单元来说是可以挤出来的；它之所以从未被挤出来，不是因为课时不够，是因为**账本上没有这一栏**：教案上写着“摩擦力”这个知识点，没有写着“第一节课那三处悬置”。

这个推演给出三条可以被推翻的具体主张。其一，悬置连接的位置可以从教案预测——本例中它必然出现在“合外力为零”与“没有力”之间，因为这两者在三十道题里从未被分开过。其二，拆解的成本远低于直觉——十分钟，而不是一节课。其三，失效的归因几乎必然指向学习者——综合运用能力不足，而不是第一节课那三句话。三条都可以在任何一个初中物理教研组用现成材料查证。

十三、三家为何各自到不了

三家没有到达这个位置，不是因为疏忽。每一家各自的问题结构，恰好使这一步成为不必要的。

教育学到不了，因为它的因变量止于学习者内部。概念转变的整套方法——临床访谈、概念图、前后测、迷思概念清单——都围绕学习者一侧建立，它问的是“学生原有的是什么、后来变成了什么”。悬置是教学一侧的动作，而这一支从不追问“教学挡住了什么”；教学在它的模型里以“提供不相容经验”这一个变量出现，挡住的动作不占位置。要它看见悬置连接，等于要它把教学的技术性安排纳入概念生态的因果链，而那会使这一支的分析单位从学习者扩到师生两方——它的整套工具在那一刻失效。

化学到不了，因为它的体系里只有一个所有者。分子是实验者的，保护基是实验者装的，迁移了实验者自己承担后果，脱保护做没做完在同一份收率里结清。因此化学不需要“归属”这个变量——挡住与被挡住之间没有利益分歧，也没有账期错位。课堂里有两个所有者：悬置由教师作出，连接长在学生身上，代价由学生在数月之后承担，而三者不在同一本账上。化学的形式约束可以整个搬过来，唯独这一件搬不过来，因为它在化学里根本不构成问题。

艺术学到不了，因为它处理的是作品而不是制作者。未完成状态的研究问的是那件东西算不算作品、由谁的条件判定、在流通中如何被安置，整个论域都在作品一侧。它从不问制作者在等待期间发生了什么变化——一块停在半路的石头对米开朗基罗做了什么，不是艺术史的问题。要它看见悬置连接，等于要它把作品的未完成状态读作制作者身上的一个状态，而那属于另一门学问。

三家的盲区各不相同，但它们的交集恰好是本章的位置：一个由甲方作出、长在乙方身上、由乙方在数月之后承担、而三家的记账单位都不覆盖的东西。教育学的账本记学生，化学的账本记体系，艺术学的账本记作品——没有一本记“甲的一个动作在乙身上留下的东西”。

十四、为什么常见的改革处理不了这一件

课堂改革不缺方案。近三十年里，增加活动、增加反馈、增加练习、分层教学、个性化推送，每一样都被大规模推行过，每一样都在试点里有效，每一样在推广之后回到原处。本章给出的解释是形式上的：这五类改革改的全是悬置之外的变量，悬置密度与脱悬率一个都没动。

增加活动改的是形式，不改归属。讨论、项目、探究、翻转，改变的是谁在说话、用什么材料、按什么顺序。而一个讨论要能进行，问题必须先被限定；限定就是悬置。活动化的课堂通常悬置密度更高——把一个复杂问题改造成可讨论的形式，需要挡住的东西比直接讲授更多。这解释了一个反复被观察到的现象：活动越丰富的课堂，学生的当轮参与与满意度越高，而换情境表现未必更好。

增加反馈改的是速度，不改对象。更快地告诉学生对错，更细地标出错在哪一步，更多地给出改进建议。这一切都在被悬置之外的那部分内容上进行——反馈永远只针对已经进入本轮的东西，因为被挡住的那部分本来就不在作业里。反馈越及时越精细，学生越熟练地在既定范围内作答，越没有理由去碰边界。

增加练习改的是熟练度，不改边界。同型练习提高的是在既定条件下的执行速度与准确率。而悬置连接恰恰在同型练习里从不出错，因此增加练习不仅不能拆解它，还会把它

固化——每一次正确都是对那条连接的一次确认。这一条给出一个可查的预测：**在同型练习量与换情境表现之间，应当存在一个拐点，超过它之后相关转负。**

分层教学改的是难度，不改挡住了什么。把学生按水平分组、给不同难度的任务，而难度的降低几乎总是通过增加悬置实现的——低层组的任务里挡住的东西更多，因此低层组的悬置密度系统性更高、脱悬率系统性更低。这解释了分层教学一个长期的困扰：分层在当轮成绩上有效，而组间差距随时间扩大。按本章，这不是低层学生学得慢，是低层组被挡住的东西更多而拆得更少。

个性化推送改的是路径，不改账本。平台根据错误类型推送下一题，路径因人而异。而推送的依据是答题结果，答题结果记录的是内容点的对错，不是内容点之间的连接。系统越精确地按错误类型分流，越是把学习者固定在他当前那套连接所能应付的题型上——因为凡是他能做对的，都不会被推送。

五类改革共享一个特征：**它们的效果都在被悬置之外的那部分内容上产生，而代价都在被悬置的那部分上积累。**试点期间，研究者在场、课时宽裕、教师有余力回过头补，脱悬率被临时抬高；推广之后课时回到常规，脱悬率回落，此前所有的局部改进同时失效。这不是执行不力，是变量选错了。

十五、与最接近的几种说法的分界

认知负荷理论（Sweller, van Merriënboer & Paas, 2019）说到的是：工作记忆容量有限，教学应当管理内在负荷、削减外在负荷、留出建构负荷。悬置正是削减内在负荷的标准手段，因此这一支为悬置提供了全部正当性。分离线在**账期**：它计算的是当轮的负荷，它的账在下课时结清；本章计算的是悬置期内结成的东西，账在换情境时才到期。判定形式：若在负荷完全匹配的两组之间，只改变脱悬率而换情境表现无差异，则本判断错而负荷理论足够；若有差异，则负荷是当轮的读数，不是全部的读数。

支架与渐隐说到的是：外部支持应随能力提高逐步撤除。分离线在**撤除的对象**：渐隐撤除的是帮助，本章拆解的是被悬置的内容。判定形式：观察一个帮助已全部撤除、而悬置从未被拆解的课堂——渐隐理论判它已经完成，本章预测它在换情境时与从未渐隐的课堂无差异。这样的课堂在实践中极常见，因此这一条可以立刻检验。

惰性知识（Whitehead, 1929; Renkl, Mandl & Gruber, 1996）说到的是：知识存在却调用不出来。分离线在**方向**：惰性知识讲该来的没来，本章讲来了的带着不该带的。判定形式：若换情境失败的学生在提示后能正确调用被悬置项，则是惰性知识；若他能调用却仍按当初那条连接作答，则是悬置连接。这两种表现可以在同一次访谈里分开。

迷思概念研究说到的是：学习者带着先于教学的错误解释进课堂。分离线在**来源**：迷思概念是带进来的，悬置连接是课堂制造的，因此它的位置可以从教案预测。判定形式：若某类错误在使用不同悬置策略的两套教材之间分布相同，则它是迷思概念；若分布随教材的悬置位置移动，则它是悬置连接。教材对比是现成可做的。

螺旋课程（Bruner, 1960）是最贴身的一位，因为它正是“以后再讲”的制度化：同一主题在不同学段以更高复杂度重访。分离线在**重访的形态**：螺旋课程假定重访即重接，本章预测重访通常是新起——在更高的复杂度上重新讲一遍，而不要求学习者回到低复杂

度那一次的具体判断上重判。判定形式：若螺旋重访的班级与直线推进的班级在换情境表现上无差异，则本判断对而螺旋课程的重访机制未兑现；若有差异且差异随重访次数而非随拆解次数变化，则反之。

理想化与模型的必然简化（Cartwright, 1983）来自科学哲学，说到的是：任何模型都必须撇开大量因素，物理定律在严格意义上是假的。分离线在**对象**：它讲模型与世界的关系，本章讲被撇开的部分在学习者身上形成了什么。判定形式：这一支预测任何理想化都同等地“不真”，本章预测理想化之间有系统差别——被撇开项与被保留项在内容上相邻的那些理想化，产生的悬置连接更强。

信息隐藏与架构侵蚀（Parnas, 1972; Perry & Wolf, 1992）是形式层最贴的一位，来自软件工程。把本章的命题剥掉学科名词，剩下的是：**为使一部分可控而暂时排除其余，排除本身在被排除项之间生成了未被声明的关系**。这在软件工程里叫隐性耦合，其长期后果叫架构侵蚀——模块边界划得越清楚、被隐藏的东西越多，未声明的依赖就越容易在边界背后长出来，而它们不违反任何接口约定。分离线在**谁承担**：软件里那些依赖长在代码上，可以被静态分析找出；课堂里它们长在学习者身上，任何静态检查都读不到。判定形式：若课堂的悬置连接能被作业与试卷的常规批改识别，则本章多余，它只是软件比喻；若必须用换情境任务与出声重做才能读到，则两者是同一形式在不同承载物上的两个例子，本章的增量在承载物而不在形式。

成功陷阱与能力刚性（Levinthal & March, 1993; Leonard-Barton, 1992）是模板层的占位者，必须点名。本章的形式一旦剥净，接近“越讲得清楚，换情境越差”这种反转模板，而这个模板早有名字。分离线在**前提**：成功陷阱要求存在一项被推到极端的成功能力，能力刚性要求存在一项曾经带来优势的核心能力；而**一门从来没有成功过的课，悬置密度同样高、脱悬率同样低**——因为悬置是讲课得以进行的条件，不是成功的产物。判定形式：取一批长期成绩低迷的课程，若它们的悬置密度显著低于成绩优异的课程，则成功陷阱足以解释；若两者无差异，则本章所指的**不是成功的代价，是讲授的代价**。

负迁移说到的是：先前学习妨碍后续学习。分离线在**先前学习的身份**：负迁移里的先前学习是一次完整的学习，本章里的先前学习是**一次被声明为暂时的悬置**。判定形式：负迁移预测干扰强度随先前学习的牢固程度上升，本章预测它随悬置的明确程度上升——一句说得越清楚的“这个以后再讲”，产生的连接越强，因为它同时给了学习者一个明确的许可去把那部分放下。这两条预测在方向上可以分开检验。

划完之后再取一次交集。逐条问这几位各自结构性地看不见什么：认知负荷理论把账期当作给定，因此看不见跨轮的连接；螺旋课程把重访当作重接，因此看不见新起；迷思概念研究把错误的来源当作先于教学的，因此看不见教学制造的错误；负迁移把先前学习当作完整学习，因此看不见被声明为暂时的那一类。这几行背后是同一件被当作给定的东西：

教学的最小单位是内容点，而内容点之间的关系是内容点的函数。

所有派别都在争内容点该怎么排、怎么教、怎么测，没有一派把“内容点之间在学习者身上自行长出的关系”当作独立的记账对象——因为一旦承认它，**教学序列这个概念本身就失去对结果的决定权**：同一份序列在不同的悬置安排下产出不同的关系网，而序列本身看不出来。这是本章动摇的那块地。

十六、与本编另两章的分工

本章与本编另两章处理同一个问题，读数与靶格都不同，必须逐一分清。

与第九章《还在动的时候就读了数》。那一章处理的是判定的时点相对于重组速度过早，因而把仍在进行的变化读成零并据此配置后续条件；它的读数是观测窗口与重组时长之比。分离线在位置：那一章管一次改变的出口（改变还没完成就被读数），本章管一次改变的伴生物（改变完成了，而它附带了一批未被指认的连接）。判定形式：取一个把所有判定都推迟到重组完成之后的班级——那一章预测它的读数有效，本章预测若悬置从未被拆解，推迟判定只会更晚地读到同一批悬置连接的后果。两个预测可以同时被检验，且不冲突。

与第七章《开着的时候没人来》。那一章处理的是上一轮教学打开的感受窗口与本轮内容不对齐，因而窗口在无人到达的情况下关上；它的读数是对窗率。分离线在轮次关系：那一章管入口（内容有没有落在能被接住的时刻），本章管入口之后那一次改变附带了什么。判定形式：一个对窗率很高、每一轮都精确接上上一轮涌现的班级，若它同时悬置密度高而脱悬率低，那一章判它健康，本章判它落在靶格。这个对照两章都不能自己判，必须实测。

三章放在一起才说得出一句：同一次改变有入口、伴生物与出口三处可以失效，三处的读数数学上独立、可同批测量，而效应相乘不相加——入口对不上，另两处做到最好也没有输入；入口对上而伴生物无人拆除，出口读得再准也只是准确地读到一个带着杂物的结果。

十七、五类课堂场景，与其中两处收窄

中学物理的摩擦是条件悬置的标准形态。悬置语句在教材里明确可数（本题不计摩擦、忽略空气阻力、设斜面光滑），后续摩擦力单独成节，几乎不回溯当初那些题现在该怎么算。第一处收窄在这里：物理教学中有一类理想化是不可拆的——质点、刚体、点电荷，拆掉之后该层次的物理学不存在。对这一类，本章的判据必须改写：不是“是否回到当初那道题重算”，而是“学习者是否知道这是一个模型、以及它在什么尺度上失效”。这一处收窄不小，因为中学物理里不可拆的理想化占了相当比重。

中学数学的负数开方是范围悬置。它的特殊之处在于跨度极长：初中说无解，高中引入复数，中间隔着两三年。跨度越长，回接的制度成本越高，而回接的价值也越大——复数引入时若能回到当初那个具体的方程重新表态，学习者得到的是一次范围的扩张；若另起一章，他得到的是两套互不相关的说法。这一格可以直接从教材的编排上判断，不需要进课堂。

化学的反应机理是机理悬置，且跨学段。中学说记住这个规律，大学讲机理，中间隔着入学考试。这一格在结构上不可能回接，因为回接需要同一批学习者、同一位教师、同一门课程连续性，而学段之间三者全断。本章对这一格给不出处方，只能指出：跨学段的悬置在现有制度下等于永久悬置。

写作教学的模板是判断悬置，也是四类里最难处理的一格。学生按模板写、按评分标准改、按教师的判断确定好坏；被悬置的是“凭什么这样算好”。第二处收窄在这里：在

评价权高度集中的场合，要求学生重新评判教师给出的标准，可能使他承担不该由他承担的风险——尤其在这个标准同时决定他的成绩时。处置只能是把重判放到低风险场合：同伴互评、匿名样本、往届作品、与另一套标准的对照。这一处收窄改变了本章处方的形态：判断悬置的拆解不能在原评价关系内完成。

作坊制技艺是唯一一格结构上强制回接的。下一次上手就是回到同一件坯、同一张画；材料还在那里，上一次做的每一个决定都以物的形式摆着，不需要任何人提醒。这一格给出本章最有用的一条设计线索，而且它不是理念问题：**脱悬率高低，主要取决于工作对象是否持续存在**。课堂的作业交上去就不再回来，因此每一次悬置都自动变成永久悬置；而坯与画一直在那里，因此每一次未完成都自动进入下一轮的视野。这条线索比“教师应当更重视回顾”有用得多，因为它指向一个可以被改动的制度安排，而不是一种态度。

十八、三个领域反过来加上的限制

化学削掉了本章最自然的一条价值排序。若只看保护基策略，本章会写下一句更强的话：脱悬率应当尽可能高，所有悬置最终都要被拆。而不使用保护基的合成路线（Baran et al., 2007）证明，最好的路线不是保护做得更彻底的路线，是**根本不需要保护的路线**。翻回课堂，处方因此改变形态：第一位的不是补拆解——那会无限增加课时——而是**优先选择不需要悬置的入口**。同一个概念常有多个入口，有的入口必须先挡住三样东西，有的一样都不必挡；后者的当轮讲解更费力，长期账更干净。这一处削掉的是“越拆越好”，换上的是“少挡优先于多拆”。

艺术学削掉了本章的适用范围。若只看化学，本章会写下另一句更强的话：任何未拆解的悬置都是欠账。而未完成状态的研究证明，保持未完成本身可以是构成性的，强行完成会毁掉它。翻回课堂，适用范围因此收窄到**课程声称该内容最终要能独立运行的那些悬置**；对于“知道有这回事就够了”的内容，悬置不需要拆解，把它算进欠账是记错了账。这一处削掉的是本章对适用范围的默认扩张。

教育学削掉了本章的道德色彩。概念转变的修正稿承认，学习者的概念生态里存在大量非目标的连接，它们同样支持后续学习。因此本章只主张悬置连接**不进入任何账本**是问题，不主张它们都有害。有些悬置连接后来成了理解的支点；本章无法预先分辨，也不声称能分辨——它主张的只是：一个没有被任何人指认过的东西，不可能被有意地利用或拆除。

十九、不适用的边界

三处反向约束已经削掉了本章的价值排序与部分适用范围。这一节把边界写全，共五条。写全的用处不是谨慎，是防止读者把一个窄命题用到它不成立的地方。

其一，纯熟练化内容不适用。口算、拼写、指法、键盘、基本操作程序，这类内容不靠概念之间的连接工作，靠的是重复与自动化。在这里挡住其余部分不产生本章所说的连接，因为根本没有“其余部分”可以与它接上。把这类内容算进悬置密度是记错了账。

其二，安全关键内容的保护性悬置不适用。实验室里某些反应的细节、临床上某些操作的风险边界、机械操作中某些禁忌，它们的悬置本身是保护性的。这类悬置的拆解必须

在模拟环境中完成，且拆解的目的不是让学习者重判，是让他知道边界在哪。本章的判据在这里不适用，因为“回到当初那件事重做”在真实场合中不被允许。

其三，目标只要求“知道有这回事”的内容不适用。艺术学那一处反向约束已经削出这一条：并非所有悬置都是欠账。判据是课程是否声称该内容最终要能独立运行；不声称的，悬置就是它的终态，不构成损失。

其四，学习者已有充分先备结构时不适用，且方向可能相反。对一个已经掌握相关结构的学习者，挡住其余部分反而降低干扰、提高效率——这与新手的情形方向相反。因此悬置密度这个读数在专家与新手之间不可比，混在一起算会互相抵消。

其五，单次接触的内容不适用。本章要求悬置与拆解之间有一段可追踪的时间跨度；一次讲座、一场报告、一节公开课，没有下一轮，也就没有拆解的可能，本章对它无话可说。

这五条合起来把本章的射程收得相当窄：它只适用于**有连续轮次、目标要求独立运行、学习者尚未形成相关结构、且内容依赖概念连接的那一类教学**。这个范围小于课堂总量的一半。本章认为这是应该的——一个适用于所有教学的命题，通常什么也没说。

二十、经验研究设计

本章的三个量都必须能在既有课堂里测出来，否则它只是一个说法。

主设计是同一制度环境内的多班级试验：同一学校、同一年级、同一学科、同一课程标准，不少于八个平行班，持续十二周。跨校、跨地区、跨国的比较一律不用——两边在教材、课时、生源与文化上处处不同，所谓控制某个变量在方法上站不住。

四类条件的差别被刻意收窄。**A组**按常规教学，悬置照常作出，不作拆解安排。**B组**在A的基础上刻意提高悬置密度（更多范例、更严格的题型限定、更多“以后再讲”），用作强化对照。**C组**与A的悬置密度、反馈量、课时与教师投入全部匹配，唯一差别是每一次明确悬置都在四周内被拆解一次。**D组**改换入口：对同一批概念，选择不需要悬置或悬置更少的入口教，当轮讲解更费力，课时与A匹配。

三处对照各回答一个问题。**A与C**检验拆解本身的效应，是本章处方的核心检验。**A与D**检验换入口的效应。**C与D**是本章最要紧的一处对照——化学那一处反向约束预测D优于C，即少挡优先于多拆；若实测C优于D，那一处反向约束不成立，本章的处方次序要倒过来。

数据分五层。第一层是基线知识与先前成绩。第二层是教案与教材的悬置语句编码，这一层完全不进课堂，用现成文件即可完成。第三层是课堂录像与学生草稿，用于识别连接的可观测签名。第四层是当轮正确率与常规单元测。第五层是四周与十二周后的换情境迁移任务——题目在表面上改变而关系结构保持。

悬置语句的编码规程必须写死，否则密度这个数会随编码者浮动。计入的语句须同时满足三条：明确指向某项内容；明确表示本轮不处理；且该内容属于本课程后续或相邻课程的范围。“这个太难了”不计入（未指向具体内容），“这个我们不学”不计入（不在课程范围内），“这个以后再讲”计入。两名编码者独立完成首轮编码，一致性低于阈值时该批数据作废而不取平均。

连接的编码用出声重做：给学习者当初那道题，要求他用被悬置项重新做一遍，并说明当初的答案在什么意义上仍然成立。三级评分——能重做且说得出边界、能重做而说不出边界、不能重做。要紧的是编码者不得看该学生的当轮成绩，否则会顺着成绩判。

最低成本的先导研究不必重组课程。只用两样现成材料：教材与教案里的悬置语句（可直接检索），以及同一册教材后续章节与习题里“是否回接”的证据。把这两样与既有的期末迁移题对上，一个学期之内可以完成，不需要任何新干预，也不需要进课堂。若在这一层就看不到关联，应当优先怀疑本章的机制，而不是把失败归因于样本不足。

统计上以学生嵌套班级、班级嵌套教师。核心检验不是比较组均值，而是在控制当轮正确率、练习量、课时与先前成绩之后，检验悬置密度与脱悬率的交互项是否预测换情境表现。为减少“好教师什么都做得好”的选择偏差，优先采用同一教师内比较：同一位教师在不同单元的脱悬率自然波动是否预测对应知识点的迁移。

二十一、证伪条件与轮次预测

最强的竞争解释不是“传统课堂也可以很好”，是一句更朴素的话：**其实不就是没学扎实吗**。换情境表现差的学生，多半原本就掌握得浅；悬置与拆解只是好教学的又一个名字。

排除它需要一个只有本机制才预测得到的变量。本章取的是**悬置密度与脱悬率的交互项**：若只是学得扎不扎实，当轮正确率与练习量就应该解释掉全部差异；本机制预测的是，在当轮正确率、练习量、课时、教师与先前成绩全部匹配之后，**换情境表现仍随交互项变化**——高悬置低拆解的班级最差，而低悬置低拆解的班级不差。后一格是竞争解释预测不出来的：按“没学扎实”，低悬置的班级当轮更吃力、正确率更低，应当更差。

据此写出七条证伪条件。

其一，控制当轮正确率、练习量、课时、教师效应与先前成绩后，若悬置密度与脱悬率的交互项不预测换情境表现，则核心机制为伪。

其二，若低悬置密度且低脱悬率的班级在换情境表现上不优于高悬置密度且低脱悬率的班级，则靶格设定错误，第二根轴应当换掉。

其三，若脱悬（要求回到当初那件事重做）与单纯复习被悬置项在换情境表现上等效，则本章关于“复习不是拆解”的区分没有经验意义，处方失效。

其四，若悬置连接的位置与被悬置项在教材中的相邻内容无关而呈随机分布，则第六项兑现失败，本对象退回“没学扎实”的一个别名。

其五，若同一批学生在使用不同悬置位置的两套教材之下产生分布相同的错误，则本章所指的是迷思概念，不是课堂制造的连接。

其六，若换情境失败的学生在提示后即可正确作答，则本章所指的是惰性知识。

其七，若长期成绩低迷的课程其悬置密度显著低于成绩优异的课程，则成功陷阱这一模板足以解释，本章无独立地位。

第五条最便宜，且所需材料在多数教研室已经存在：两套主流教材的悬置语句可以直接检索，同一批错误的分布可以从既有作业与试卷中回溯统计，一个学期之内可以完成，不需要任何新干预。

轮次预测。本章的动力主张不是相关，是一个有次序的循环，必须写成轮次。

第 t 轮教学作出一次明确悬置 → 第 $t+1$ 轮在学习者的解释里出现被悬置项与新内容之间未被要求的连接（可在出声解释与草稿中观察） → 第 $t+2$ 轮课堂条件因这些连接而改变（教师发现“讲过的又不会了”，处置是增加同型练习或进一步简化） → 第 $t+3$ 轮的发起处从内容序列转到补救序列，而补救序列几乎必然带来更多悬置。

四步中任何一步的顺序系统性相反，本模型即需修正：若补救压力先于悬置出现，则悬置是压力的结果而非原因；若被悬置项与新内容之间的连接在悬置之前就已存在，则它是带进来的而不是制造的。

二十二、课堂可以怎么改：五个动作

本章首先是一项解释，不是教学法。但一旦把悬置连接看见，备课的动作会变，且变的地方比想象中便宜。

第一，把悬置说出来并记下来。成本最低的一件：在作出悬置的那一刻多说一句——我们现在把什么挡住了。这句话不需要额外课时，而它把一个原本不可见的动作变成了一条可以回来找的记录。可验收的证据是教案上出现一份悬置清单，而不是教师说自己很注意。

第二，优先选择不需要悬置的入口。这是化学那一处反向约束的直接后果，也是本章与“多补课”分道的地方。同一个概念常有多个入口：有的入口必须先挡住三样东西，有的一样都不必挡。后者当轮讲解更费力、正确率更低、课堂更不顺，长期账却更干净。备课的问句因此从“怎么讲得更清楚”改为“有没有一个不必挡这么多的讲法”。

第三，把拆解写成动作而不是名称。回顾、复习、串讲、关联，这些名称都不算拆解。一次拆解要被计入，必须留下学习者用被悬置项重做当初那件事的产物，且该产物中出现了对当初答案适用范围的判断。这条规则会被绕过，只是绕过它需要真的让学生做一遍。

第四，让工作对象持续存在。这是工坊那一格的教训，也是本章认为最有杠杆的一条。同一份作业在一学期里被回改三次，胜过三份互不相干的新作业——不是因为回改更认真，是因为回改使每一次悬置都自动进入下一轮视野，不需要任何人记得。这一条不要求教师改变态度，只要求改变作业的流转方式。

第五，把教研的对象从内容序列扩到悬置清单。集体备课时先列这一单元要挡住什么，再列打算什么时候拆、拆到什么程度、哪些不打算拆并说明理由。这份清单同时是第三处反向约束的落地：不打算拆的那些必须写明理由，否则它们会悄悄从“不必拆”滑成“忘了拆”。

五条里只有第四条涉及制度安排，其余四条是备课习惯的改变。这一点本身可检验：若在完全不改变作业流转方式的条件下，仅按第一、二、三条运行的班级与对照班无差异，本章的处方失效。

二十三、反噬预言与伦理限制

脱悬率一旦进入考核，最省事的达标办法会摧毁它。教案里补一句“回顾上节内容”就可以计入拆解，而回顾不是拆解——学习者没有被要求回到当初那道题上重判。平台可

以自动标注“知识点关联复习”，把拆解压成一次推题。教师也可以安排一次形式化的“回头看”，让全班齐声说出当初的答案在什么意义上仍然成立，而没有任何人真的重做一遍。

对此本章看不到彻底的出口。能做的只有把判定写死在**动作**上而不是名称上：一次拆解要被计入，必须留下学习者用被悬置项重做当初那件事的**产物**，且该产物中出现了当初答案适用边界的判断。除此之外的一切形式——回顾、复习、关联、串讲——一律不计。这条规则会被绕过，只是绕过它需要真的让学生做一遍。

因为悬置密度与脱悬率都是可以拿去考核人的比率，伦理限制必须写死三条。

第一，这两个读数指的是**位置**，不是人。悬置是讲授得以进行的条件，密度高不构成对教师的负面评价；把它做成教师排名是对本章的误用。

第二，**不得为了把脱悬率做好看而在安全关键教学中恢复被有意悬置的内容**。实验室安全、临床操作、机械操作与应急程序中，某些内容的悬置是保护性的，拆解须在模拟环境中完成。

第三，若学校已据这两个读数作出不利安排，**必须公开编码规则、原始证据与复核通道**，并允许被评价者提交反证。

二十四、自反

按本章自己的判据，这篇文章做了大量悬置而脱悬率为零。

为了让“悬置连接”这一处清楚，本章把学习动机、个体差异、学科差异、班级规模与教师专业发展全部挂了起来，每一处都用了“本章不处理”这类语句。这些悬置与正文之间已经结成了连接——读者从本章取走的那个对象，将带着“与动机无关”“与学科无关”这类本章从未主张、也没有证据支持的默认。本章不打算在这里补讲它们，因为补讲同样只是复习。

还有一处更难堪。本章主张一个名字会把它所指的东西挡在名字后面；而本章自己产出的正是一个名字。读者最可能取走的是“**悬置连接**”这四个字，而不是**悬置密度与脱悬率的交互项**——名字比关系更容易被记住、被引用、被写进综述。这不是本章的失误，是本章所描述的机制在自己身上的一次运行；本章对此没有办法，只能把可以让自己失败的那七条写得比结论更具体。

二十五、分层引用，与三个本章解释不了的洞

本章可以被分层引用，三层的强度依次递减而稳定性依次上升。

第一层是“悬置连接”作为一类对象。它最强，也最易倒——只要证明悬置期内并不形成额外连接，或形成的连接与被悬置项的位置无关，这一层就没有了。

第二层是那张二维辨别格。即使命名不成立，把悬置密度与脱悬率分成两根轴仍然是可用的诊断工具：它至少能把“讲得清楚”与“讲清楚之后处理了什么”分开来问。

第三层是一条不依赖前两层的经验主张：要求学习者回到当初那件事上用被悬置项重做一遍，其效果不同于单独复习被悬置的内容。这一层最稳，也最便宜——两个平行班、一个单元、一次延迟迁移测验即可检验，且它对本章其余部分成立与否不作要求。

写清可以分层引用，既是诚实，也是保护：第一层被推翻时后两层仍站得住。

有三个洞本章解释不了。它们不是应用细节，而是可能迫使核心定义改变的东西。

第一个洞：本章说不出悬置连接何时是有益的。概念转变的修正稿承认，概念生态里大量非目标连接同样支持后续学习。本章因此只主张它们不入账是问题，不主张它们都有害——但这意味着本章给不出“该拆哪些”的判据，只能给出“哪些拆得起”的成本排序。一个能分辨有益与有害的判据，会把本章的处方从“尽量拆”改成“选着拆”，而本章没有这个判据。

第二个洞：本章说不出悬置连接会不会自行消解。若它在足够长的时间之后自行松脱，本章对长期教育的整个主张就要重写；而现有的迁移研究窗口普遍短于这个时间尺度，因此这一条在可预见的将来无法被检验。本章倾向于认为它不会自行松脱——因为它在原题型里持续得到正反馈——但这只是倾向，不是论证。

第三个洞：本章的两个读数在专家与新手之间不可比，而本章说不出分界在哪。边界那一节已写明方向可能相反，而“充分先备结构”是一个连续量，本章给不出临界点。这使得跨年级、跨学段的比较在方法上不可行，而那恰恰是教育研究最常做的比较。

二十六、三条本章预测而尚无人测的现象

一个判断若只解释已知，仍可能是旧说法的新名字。本章因此列出三条它预测、而据本章所知尚无人测过的现象。每一条不成立都要动本章的一部分。

其一，同型练习量与换情境表现之间应当存在拐点，超过它之后相关转负。通常的预期是单调正相关——练得越多越熟。而按本章，同型练习每一次正确都是对悬置连接的一次确认，因此在拆解之前，练习越多连接越牢。这条预测与“熟能生巧”直接相反，而它只需要在既有的练习量数据与迁移成绩之间做一次非线性拟合即可查，不需要任何新采集。若关系严格单调，本条为伪。

其二，低层分组的悬置密度应当系统性高于高层分组，而这不是任何人有意为之。降低难度的标准手段就是增加悬置——多给范例、限定题型、简化情境。因此分层教学在结构上把悬置密度与学生水平做成负相关，而拆解的课时又优先给了高层组（因为低层组的课时被同型练习占满）。这条预测可以直接从分层班的教案上查，两套教案的悬置语句一数就知道。若两层的悬置密度相当，本条为伪，分层与本章无关。

其三，工作对象持续存在的课程，其脱悬率应当系统性高于作业一交即弃的课程，而与教师的教学理念无关。工坊那一格的教训若成立，那么在同一所学校里，作品类课程（美术、技术、写作工作坊）的脱悬率应当高于试卷类课程，且这个差异在控制教师之后仍然存在。这条预测把本章最有杠杆的一条处方推到了可查的位置：若差异不存在，或差异可由教师个体解释，那么“让工作对象持续存在”就不是一条制度杠杆，本章的第四个动作失去依据。

三条都不需要新的测量工具，只需要把已经存在的材料按不同的方式数一次。

二十七、结论与一条写死日期的赌注

课堂上学习很少真正到位，不是因为课堂没有制造改变。恰恰相反，课堂高效地制造了大量当轮完全正确的改变。问题在于，每一次为使某处改变而作的悬置，都在悬置期内附带产出了一批未被指认的连接；这些连接不产生错误，只在换情境时决定表现，而课堂的账本按内容点记，看不见内容点之间的东西。

三个领域各出一份不可替代的力。教育学解释了改变为何需要不相容；化学给出了两条最承重的形式约束——**要让一处改变必先挡住其余**，以及**挡住这个动作会改写它本要固定的那张名单**；艺术学则限定了适用范围，说明某些悬置本身是构成性的，不该被算作欠账。

这一理论把课堂评价的问题从“学生答对了吗”改为“这次悬置后来被拆过吗”；把教学改革从“讲得更清楚”或“活动更多”改为“少挡优先于多拆”；把教研的对象从内容序列扩展到序列之外那张没人画的关系网。

本章给出一条写死日期的赌注。**到 2031 年 12 月 31 日之前**，若至少三项在同一制度环境内、不少于八个平行班、控制内容总量、课时、教师、先前成绩与当轮正确率之后开展的研究报告：悬置密度与脱悬率的交互项独立预测换情境表现，且低悬置低拆解组不劣于高悬置低拆解组，则本章的核心判断获得初步支持。

以下明确**不算命中**：仅报告“基础不牢”或“迁移能力不足”；仅比较不同教材而不编码悬置语句；仅采用教师或学生自评问卷；仅比较两个地区或两个国家；把拆解与增加课时、增加练习量打包实施而无法分离贡献；把教案中的“回顾”直接计为拆解。若符合上述条件的多项研究均未发现交互效应，本章应被放弃或降格为一个描述性提法。

课堂不需要停止挡住其余部分——不挡就讲不成。它需要的只是：把每一次挡住记下来，并且知道自己欠了几笔。

接续。悬置密度与脱悬率的采集见附录一，其中脱悬率的分子判定很严——后来单独讲不算，必须要求学生回到当初那件事上、用被悬置项重做一遍。本章处理的是一次改变的**伴生物**，它与入口（第七章）、出口（第九章）的次序由第三编小结给出：先对窗，再拆解，最后才谈判定時点。

第九章 还在动的时候就读了数

判断卡片：一次判定若作在重组还没走完的时候，读到的零会被当作结论，并据此配置此后的条件——于是它自我实现。

三个理论入口：物理学 × 心理学 × 工程学

提要。课堂每隔一到两周就要作一次“会了没有”的判定，并据此配置此后的一切条件——补救、分层、跳过、放弃。判定的时点由课表决定，不由被判定的那件事本身决定。物理学主张决定一个系统“有没有在变”的是弛豫时长与观测窗口之比，心理学主张一次提取会主动抑制同类竞争项，工程学主张进不了传感回路的状态就进不了控制。三家争的是变化由什么驱动，而共同假定了一条没有说出口的话：**一个系统里正在变的部分与不在变的部分，可以在观测的那一刻被分开。**推翻它的材料来自物理学自己——物理老化：一个被判定为“不再变化”的玻璃态系统，其性质随被冻住之后所等的时间持续演化，且演化的起点正是被冻住那一刻的构型。带回课堂便得到本章命名的对象：**冻结构型**——一次判定在被判定者身上留下的、并非关于内容而是关于“当时重组进行到哪一步”的状态；此后的慢速变化都从这里出发，而它不进入任何读数。承重判断是：学习之所以少，不是它没有在进行，而是课堂的观测窗口短于重组时长，因此系统性地把慢于窗口的变化读作零，并据此把它冻在当时的构型上。本章给出窗比与双向翻转率两个读数、一张二维辨别格、一句不含判断词的问法、与九种相近说法的分界、一处强制性的自相矛盾、三处反过来加上的限制、七条证伪条件与一条写死日期的赌注。

一、问题从哪里来

一个学生在单元测里做错了三道同类题，被判为“没掌握”，进了补救组。补救组的任务更简单、步子更小、示范更多。四周之后，如果有人拿同一份卷子让他再做一次——而这四周里没有人再教过这部分——他有相当的可能做对。没有人会去做这件事，因为判定已经作出，条件已经按判定配置，那三道题早已从他的课程里消失了。

反过来的情形同样常见：当堂全对、被判为“已掌握”的学生，四周后同一份卷子做不出来，而这四周里同样没有人再教过。

两种情形通常被分别归因：前者叫“开窍了”，后者叫“忘了”。两个词都指向学生，都不指向那次判定。而它们放在一起时露出一件更麻烦的事：**判定翻转是双向的。**若只是遗忘，翻转应当只有一个方向；能从“不会”翻到“会”，说明在判定与复测之间，没有教学的情况下，有东西还在进行。

本章追问的正是这一段：课堂的观测窗口有多长，被观测的那件事需要多长，两者的比是多少，以及当这个比小于一时，一次判定到底做了什么。

本章不报告新采集的数据，是一篇理论性文章；其经验主张需要预注册研究与独立复核检验。

二、物理学的单因主张：已有的样子与观测条件不容，驱动我们对路径的判定改变

凝聚态物理处理过一个形式上完全相同的问题。一块玻璃是固体还是液体？答案不由它的组成决定，也不由驱动力大小决定，而由**弛豫时长与观测时长之比**决定。这个比在流变学里有名字，叫德博拉数（Reiner, 1964）：比值大，物质表现为固体；比值小，表现为流体。同一块沥青，在几秒的观测里是固体，在几十年的观测里会滴下来。

玻璃转变本身也不是一次相变，而是**弛豫时长超过了实验时长**，系统被困在能量地形的一个盆地，无法在观测期内探遍它本该探遍的构型（Debenedetti & Stillinger, 2001）。它并没有停止，只是慢到了读不出来。

这一支的单因主张是：**决定一个系统“有没有在变”的，是弛豫时长与观测窗口之比，不是驱动力的大小**。时序读数是：先有观测窗口的设定，随后同一系统被判为不同的物态，物态判定又决定用哪一套方程去描述它此后的路径。

单因锁定成立：这一支不认为“变化速率”与“观测窗口”是并列因素。它主张在给定观测窗口下，物态是被这个比唯一决定的。

它自己知道、却不往这个方向用的事实是**物理老化**：一块被判定为“冻住了”的玻璃，其密度、模量、蠕变顺应性随**冻住之后等待的时间**持续变化，而且变化的形态由被冻住那一刻的构型决定（Struik, 1978）。这一支用它来预测材料的长期性能与保质期；它从来不被当作一条关于“正在变与不在变可否当下区分”的证据。

三、心理学的单因主张：已有表征与提取路径不容，驱动可及条件改变

记忆研究提供第二个位置，且它的主张与直觉相反。练习一部分内容，不只是让这部分变强，还会**主动削弱同类中未被练到的那些**。这个效应有稳定的实验范式与三十年的复制记录（Anderson, Bjork & Bjork, 1994）：从一个类别中反复提取若干成员，会使同类中未被提取的成员在后测里显著低于从未接触过该类别的基线。

这一支的单因主张是：**决定什么留下的，是提取时的竞争关系；遗忘不是衰退，是提取这个动作自己造成的抑制**。时序读数是：先有一次提取，随后同类竞争项的可及性下降，可及性的改变又决定下一次提取会调出什么。

单因锁定成立：这一支明确主张抑制是提取的直接后果，不是注意力、时间或干扰的副产品，并用独立线索测试排除了这些替代解释。

它自己知道、却不往这个方向用的事实是这一效应的**边界条件**：当学习者把类别内的成员整合成一个相互关联的结构时，提取诱发遗忘会消失甚至反转——提取一个成员反而促进同类（Anderson & McCulloch, 1999）。这一支用它来划定效应的适用范围；它没有被用来问：当一次判定要求学习者提取某几项时，这次判定本身就在改变其余项的可及性——那么被判定为“不会”的那些，究竟是本来不会，还是被这次判定压下去的。

四、工程学的单因主张：控制路径与传感条件不容，驱动系统实际表现改变

控制工程提供第三个位置。一个系统能否被控制，不只取决于有没有执行机构，还取决于它的内部状态能否从可测的输出被重建。不可观测的状态无论多重重要，都进不了反馈

回路 (Kalman, 1963)。工程实践据此把系统分解成可控可观、可控不可观、不可控可观、不可控不可观四块，并对最后一块只做一件事：证明它是稳定的，然后不管。

这一支的单因主张是：决定一个系统实际表现成什么样的，是哪些状态进得了回路，不是这些状态本身有多要紧。时序读数是：先定传感与采样方案，随后可观测子空间被确定，回路只在这个子空间里工作，系统的实际行为随之被塑造成回路看得见的那个样子。

单因锁定成立：这一支不认为“重要性”能替代“可观测性”。一个对性能至关重要而不可观测的状态，在设计上与不存在等价。

它自己知道、却不往这个方向用的事实是采样定理的另一半：若采样频率低于信号最高频率的两倍，得到的不是“没有信号”，而是**混叠**——一个虚假的、比真实变化慢得多的低频信号 (Shannon, 1949)。工程实践的对策是在采样之前加抗混叠滤波器，把快于半采样率的成分先滤掉，宁可承认“这部分我不测”，也不让它伪装成慢变化。这一支用它来保证信号重建的正确性；它没有被用来问：一个采样过疏的评价体系读到的进步曲线，是不是一条混叠产物。

五、三家为何不能同时为真

三家把决定性的驱动放在三个不同位置：物理学放在已有构型与观测条件的不容，心理学放在已有表征与提取路径的不容，工程学放在控制路径与传感条件的不容。并列起来会得到一句无风险的话：评价要及时、要准确、要覆盖重要内容。这句话正确，且任何失败都可以事后归入其中一条。

三家实际上互相取消。

工程学的日课正是物理学诊断的病根。工程的标准做法是：不可观测的那一块，证明稳定即可不管。而按物理学，“看起来不动”恰恰不是稳定的证据——它可能只是弛豫慢于窗口，而被冻住的系统正在老化。若物理学那一条是对的，工程学对不可观测子空间的整个处置方式在原理上不成立，因为它用来判定“稳定”的观测手段本身就是那个窗口。

心理学的日课正是它自己证明会造成损害的那件事。这一支最有力的教学建议是用提取来巩固学习；而它同一套实验证明，提取会主动抑制同类未被提取项。若把判定视为一次大规模提取，则每一次判定都在削弱它没有考到的那些内容。工程学会说这是采样的必要代价，物理学会说这说明观测本身改变了系统——三家在“能不能一边测一边不动它”这件事上不可调和。

物理学的处方在另两家眼里不可执行。物理学的办法是把观测窗口拉长到超过弛豫时长。课堂无法这样做：课时有限，判定必须在单元结束时作出，而按工程学，推迟判定等于降低采样率，只会加重混叠。

三家必须有两家错，或者三家的病灶都放错了位置。本章主张后者。

六、共有前提，与推翻它的那件材料

三家争的是**变化由什么驱动**。而它们共同假定了一条从未说出口的话：

一个系统里正在变的部分与不在变的部分，可以在观测的那一刻被分开。

物理学问变化沿什么路径进行，心理学问什么被记住什么被抑制，工程学问哪些状态进得了回路——三种问法都要求先有一个“此刻在变／此刻不在变”的分类。这句话谁听了都会说这还用说——正因为如此，它从来没有被任何一家写下来过，而三家都踩在它上面。

推翻它的材料在物理学自己手里，而且是材料工程的常识：**物理老化**。一块被判为已经冻结的玻璃态材料，在恒温恒压、无任何外加驱动的条件下，其性质随**等待时间**持续演化；蠕变曲线的形状不随等待时间改变，只沿时间轴平移，平移量是等待时间的幂函数（Struik, 1978）。这意味着三件事，逐件都与共有前提相反：

其一，**被判为“不在变”的系统正在变**。它变得慢，慢到在作出判定的那个窗口里读数为零。

其二，**它此后的变化以被判定那一刻的构型为起点**。冻住的时刻不同，往后的老化轨迹就不同——**判定这个动作参与了后续演化的初始条件**。

其三，**读数为零是关于窗口的，不是关于系统的**。换一个更长的窗口，同一个系统会给出非零读数；这不是测得更准了，是测的本来就是另一个量。

带回课堂，共有前提塌掉。一次判定不是对一个已完成状态的读取，它是在**一段仍在进行的重组中间**，取了一个点，并把这个点当作那个人的状态记下去。而按老化，被记下的那一刻不是中性的：往后的一切慢速演化都从它出发，而课堂据它配置的条件又限制了这条演化能走多远。

七、冻结构型

据此本章命名一个对象。

冻结构型：一次判定在被判定者身上留下的状态，它不是关于内容掌握与否，而是关于当时那次重组进行到了哪一步；此后不受教学的慢速演化以它为起点，而它不进入任何读数。

四个必要条件缺一不可。第一，被判定的那件事的**重组时长长于观测窗口**——学习者当时确实还在进行中。第二，判定被**作出并记录**。第三，**后续条件依判定配置**——分层、补救、跳过、免修、放弃，任何一项都算。第四，**判定与后续条件之间没有再判定**，因此那个构型在此期间无人过问。

三条性质使它区别于邻近的坏情况。

它不是遗忘。遗忘是单向的；冻结构型预测**双向翻转**，且“不会翻成会”的那一支不能由额外练习解释，因为期间没有教学。

它不是测量误差。经典测量理论假定存在一个在测量期间稳定的真值，误差围绕它波动；冻结构型说的是真值本身在以慢于窗口的速度移动，因此两次测量之间的差不是噪声，是位移。

它不是能力差异。同一个学生不同内容上的窗比不同——有的内容重组以分钟计，有的以月计——因此同一个人可以在一处被冻住、在另一处不被冻住。

需要排除两类伪阳性。其一，**判定不配置后续条件**：一次不计分、不分层、不改变任何安排的诊断性测验，即使窗比很小也不产生冻结构型，它只是一次无效读数。其二，**重组已经完成**：若该内容的重组时长确实短于窗口，判定就是有效的，与本章无关。

也要排除一类伪阴性：**推迟判定不等于没有冻结**。把单元测从第二周挪到第六周，若该内容的重组以月计，窗比仍小于一，只是冻得晚一点。延迟本身不是解药。

八、冻结结构的四种亚型

按被冻住的是哪一类重组，可以分出四种。四种的形成机理相同，而复测时的翻转方向不同——这使得亚型分类本身给出一条可查的细化预测。

表征重组冻结。学习者正在把一个问题从表面特征重述为关系结构：从“这是一道行程题”重述为“这是两个量按同一倍数变化”。这个过程以周计。判定落在中途，他被冻在“部分重述”上——遇到新题时用一半旧描述、一半新描述，做对的那些靠旧描述里恰好够用的部分。复测时这一类**最容易出现“不会到会”**，因为重述会在无教学的情况下继续进行。

策略选择冻结。学习者已经掌握两套做法，正在从“只会一套”过渡到“按情形选一套”。判定落在中途，他被冻在“有两套而不知何时用哪套”。这一格的表现极不稳定：同一类题时与时错，取决于哪一套先被想起。教师最常见的处置是要求他“统一用老师教的那一套”——而这恰好把选择这件事从他手里拿走，使冻结变成永久。

判据重组冻结。学习者正在把“老师说对就是对”换成“我能自己验证”。这个过程以月计，中间必然经过一段**权威已被削弱而验证能力尚未建立**的时期。判定落在这一段，他被冻在半信半疑上。这一类复测时**最容易出现“会到不会”**——因为原来靠记忆与信任支撑的正确，在权威被削弱之后失去了支撑，而新的支撑还没长出来。这一格最容易被误读为“退步”，实际上它是四类里最接近成功的一类。

整合冻结。学习者正在把若干孤立结论连成一片：知道它们有关，说不出关系。这个过程时长跨度最大，从数周到数月。判定落在中途，他被冻在“能分别复述、不能一并使用”上。这一类在单项测验里全部通过，只在需要跨项的综合题上失效，因此它被单项判定系统性地漏掉。

四种可以并存，且同一个学生不同内容上处于不同亚型。因此**不能把一个人整体判为“被冻住了”**，必须按内容与重组类型分别记。这一条同时限制了本章的读数：双向翻转率是一个混合量，若不按亚型分层，第一类的“不会到会”与第三类的“会到不会”会互相抵消，总翻转率可能接近零而两个方向都不为零。**分方向报告因此不是精细化要求，是这个读数能不能用的前提。**

九、二维辨别格

横轴取窗比：从该内容首次出现到判定作出之间的时长，除以该内容重组所需的时长。
纵轴取判定是否配置后续条件。两轴相互独立。

	判定不配置后续条件	判定配置后续条件
窗比大（窗口长于重组）	冗余型：读数有效而无人使用；浪费，但不产生损害	对时型：读数有效且被使用，判定与被判定之物同步
窗比小（窗口短于重组）	观望型：读数无效但无后果；关键对照格	冻结型：读数在原理上无效，而此后的一切条件按它配置。本章的靶格

右下角是靶格，三条性质使它难以被发现。

其一，短期指标表现为改善。窗口越短、判定越频繁，即时正确率越高、进度感越强、“及时反馈”这一条越好看。任何以当轮读数评价的体系都会奖励缩短窗口。

其二，失效不产生信号。被冻住的那批人的继续演化不产生任何记录——他们在补救组里的表现由补救任务决定，而补救任务是按冻结结构型配置的，因此永远与判定一致。**判定的自我实现**在这里不是比喻，是记录结构的必然后果。

其三，它自我加固。判定失效表现为“这批学生基础差、需要更多跟踪”，标准处置是提高判定频次；频次一提高，窗比进一步变小。

左下角的观望型是关键对照格：**同样短的窗口，只要判定不配置后续条件，本章预测长期表现不受损害。**这一格的存在使本章与“频繁测验有害”划清界线——问题不是测，是测完之后按它安排了什么。

十、读数与一句不含判断词的问候

窗比需要两个量。分子是从该内容首次出现到判定作出的时长，从课表与教案直接可读。分母是该内容的重组时长，必须实测：在不作判定、不配置条件的条件下，跟踪同一批学习者的表现随时间的变化，取达到平台期所需时长的下四分位作为下界。分母难测是本章最大的一处未闭合，本章不掩饰这一点。

双向翻转率不需要分母，因而是本章更可用的一个读数：把同一份判定在两个时点各做一次，中间不教学，统计判定结果发生翻转的比例，并**分开报告两个方向**。单向翻转（会到不会）是遗忘的签名；**双向翻转，尤其是“不会到会”那一支显著非零，是本章的独有签名。**这个数只需要一份现成的卷子和四周的间隔，任何学校都做得起。

判据写成一句不含判断词的问候：

把这门课上一次作出“会了没有”判定的那份卷子，隔四周原样再做一次、期间不教，有多少人的结果会翻过来？两个方向各占多少？

五个场景跑一遍，答案互不相同：

场景	判定时点	该内容重组时长的量级	窗比
中学数学单元测	教学结束后 1-2 周	分数与比例的概念重组以月计（文献常见口径）	远小于一
编程训练营周考	每周	递归与指针的重组以周计，个体差异极大	接近一，方差极大
医学住院医的当场评估	操作当场	单项技能以次计	大于一
语言能力分级	按课程周期评	官方给出的相邻级别间指导学时以数百小时计（欧洲语言共同参考框架及其配套指南）	远小于一
工业技能认证	一次通过、长期有效	单项操作以天计	远大于一

第四格与第三格是查出来的，不是从命题推出的。语言能力框架的配套材料公开给出了相邻级别之间的指导学时，而多数语言课程按学期评级——**这一格的窗比可以直接算出来，而且算出来的数字与课程设计之间的落差是公开可核的**。住院医评估工具的时点设计同样是公开文件：当场评估与半年一次的里程碑评估在同一套体系里并存，两者的窗比相差两个数量级，而它们被同等地用于配置后续轮转。

十一、十处兑现

其一，德博拉数对应窗比，因此本章的核心读数与流变学的核心读数量纲相同，都是无量纲比，可以跨领域比较。其二，玻璃转变不是相变而是窗口效应，对应“没掌握”不是一种状态而是一个读数，因此研究必须报告观测窗口而不只报告结果。其三，物理老化对应判定后的慢速演化，因此本章预测复测差异随两次测量的间隔呈幂律而非线性，这一条可直接拟合。其四，老化曲线只沿时间轴平移而形状不变，对应被冻住的学习者其后续学习曲线形状与未被冻住者相同而起点不同，可用同一模型拟合两组并只让起点自由。其五，提取诱发遗忘对应判定这一动作本身的抑制作用，因此本章预测**判定所覆盖内容与其同类未覆盖内容之间**存在系统性此消彼长，可在同一份卷子的覆盖与未覆盖题目间直接测。其六，整合作为边界条件，对应本章的处置方向：让被判定内容之间形成结构，可以取消判定的抑制作用。其七，可观测性对应“进不了评价的状态进不了教学决策”，因此本章预测增加评价维度并不改善结果——只要新增维度仍在同一窗口内采样。其八，采样定理对应判定频次与重组频率的关系，因此本章预测提高判定频次不能解决问题，只会把混叠做得更精细。其九，抗混叠滤波对应“判定时明写哪些成分不予判定”，这是本章与“多测”和“晚测”都不同的第三条处方。其十，不可控不可观子空间的处置方式（证明稳定即不管），对应课堂对“这批学生就这样了”的处置，因此本章预测该处置的正当性完全依赖于观测窗口，换一个窗口即失效。

第三、第五与第九项是本章最依赖的三项。删掉第三项，本章的对象退回“遗忘”；删掉第五项，判定的主动作用消失；删掉第九项，处方退回“延迟测验”。

十二、一个班走完：从判定到配置到锁死

取初中数学“分数除法”这个单元，跟一个班走一遍。

第三周作出判定。单元测，班上二十六人不及格。这二十六人里，按本章的亚型分，至少有三类混在一起：一类正在把“除以一个数等于乘它的倒数”从口诀重述为“一个量里有几个另一个量”（表征重组，以周计）；一类已经会两种做法而不知何时用哪种（策略选择，以周计）；一类只是没做够题（重组已完成，纯熟练度问题，以天计）。**卷面上三类完全一样，都是错。**

第四周判定配置条件。二十六人进补救组，补救内容是同型题的分步练习。这个安排对第三类完全正确——他们需要的就是多做题，两周后追上。对前两类是灾难：分步练习把“一个量里有几个另一个量”这个正在成形的重述**替换成了一套操作步骤**，因为分步练习的每一步都是操作，没有一步是关系。第一类学生的重述在这里停住，因为它不再有用武之地。

第六到第十二周，冻结构型在新条件下继续。补救组的任务由判定配置，而判定按操作正确率读数，因此补救组的表现稳定上升——他们确实越来越会按步骤做。判定的自我实现在这里不是比喻：读数上升，正是因为读数所测的正是补救任务所训练的那一样。第一类学生此后遇到需要用关系而不是步骤的题，仍然做不出，而这被记作“基础太差、补了也就这样”。

如果第五周做过一次复测。同一份卷子、四周间隔、期间不教——按本章预测，第一类学生里会有相当一部分翻成“会”，因为重述在这四周里仍在无教学地进行。这一支翻转在现行安排里不可能被观察到，因为第四周就已经配置了补救，四周之后他们做的已经不是原来那份卷子。只要配置一开始，那个反事实就永远拿不到了。

成本对比。复测的成本是一份现成卷子加一节课的时间；配置的成本是八周的补救课时。而现行安排把便宜的那件省掉，把贵的那件做足。这不是谁的疏忽——**教务系统里有“补救组”这个字段，没有“复测结果”这个字段**，因此前者是一个可以走的流程，后者是一件需要有人特意去做的事。

这个推演给出三条可以被推翻的具体主张。其一，同一次不及格里混着重组时长相差一个数量级的几类学生，而判定按同一套规则处置他们。其二，补救内容的性质会决定哪一类被救、哪一类被冻——分步练习救第三类、冻前两类，这一条可以在补救材料上直接看出来。其三，复测的成本远低于补救的成本，而前者从不做、后者一直做，原因是记录结构而非教学理念。三条都可以在任何一个初中数学教研组用现成材料查证。

十三、三家为何各自到不了

三家没有到达这个位置，不是因为疏忽。每一家各自的问题结构，恰好使这一步成为不必要的。

物理学到不了，因为它的体系里没有“据读数配置后续条件”这一步。物理学家读出一块玻璃处于玻璃态，不会因此改变这块玻璃往后能接触什么环境；读数是关于系统的一个描述，它不参与系统此后的处境。物理学确实有观测扰动系统的问题，但那是测量动作本身的能量交换，与本章所说的完全不同：本章说的是**观测的结论**改变了系统此后的输入。这一步在物理学里没有对应物，因为物质不会因为被判定为某种物态而被安排进不同的房间。

心理学到不了，因为它的方法学在结构上排除了本章的对象。提取诱发遗忘的标准范式在二十分钟到一小时内完成，这是必须的——时窗一长，日常生活的干扰就无法控制，效应会被淹没。而本章关心的重组以周或月计。一个以月计的过程在一个以分钟计的范式里根本不可能被观察到，这不是这一支不够努力，是它为了获得因果控制所必须付出的代价。它掌握的关于抑制的一切知识都是在窗内取得的，而窗外的事它在设计上就看不见。

工程学到不了，因为它的系统没有自己的往后。不可观测子空间证明稳定即可不管，这在机器上是对的：一个不可观测且稳定的状态不会因为被这样处置而变成另一个样子。而在人身上，“不可观测且被判为稳定”这件事本身会通过分层、补救、免修、放弃改变这个人往后能接触到什么，于是他确实变成了另一个样子。工程学没有“被判定者”这个角色——它的对象不会因为被如此判定而改变自己此后的轨迹，因此它从来不需要问一次判定做了什么。

三家的盲区各不相同，交集恰好是本章的位置：一个读数在原理上无效、而据它配置的条件使被判定者只能沿这个读数继续。物理学有读数的无效性而没有配置这一步；工程学有配置这一步而对象没有往后；心理学有对象的往后而看不见这个时间尺度。三条各占一半，合起来才是那一格。

十四、为什么常见的改革处理不了这一件

近三十年里，提高反馈频次、引入形成性评价、推行掌握学习、发展自适应测评、增加过程性评价，每一样都被推行过。本章给出的解释是形式上的：**这五类改革改的全是判定的形式与频次，而窗比与“判定是否配置后续条件”一个都没动。**

提高反馈频次改的是采样率，而采样率提不上去。课时、疲劳与教学时间都不允许把判定频次提高到重组频率的两倍以上，因此提高频次只是把混叠做得更精细——得到一条更平滑的伪曲线。这一条给出一个可查的预测：**判定频次与迁移表现之间应当存在拐点，超过它之后相关转负。**

形成性评价改的是判定的用途，不改判定的时点。它要求证据用于调整教学，而调整所依据的证据仍然取自同一个过短的窗口。用一个无效读数去调整教学，调整得越及时，冻结得越彻底。

掌握学习改的是通过的门槛，且把窗比推得更小。不达标就重学、重测，直到通过——这意味着同一个人同一内容上被反复短窗判定，被冻在更多个点上。本章预测严格执行掌握学习的班级双向翻转率反而更高，这与该支的自我评价相反。

自适应测评改的是题目的选择，不改时间尺度。算法根据作答实时调整难度，采样在毫秒级进行——而重组在周与月的尺度上。采样再密，密的也是同一个窗口内的密，跨不过那个尺度。

过程性评价改的是记录的丰富度，不改配置的依据。记录更多的过程数据本身不产生任何损害；损害来自这些数据被用来配置后续条件的那一步，而这一步在过程性评价的设计里从未被单独讨论——它被默认为过程数据的正当用途。

五类改革共享一个特征：**它们都把“判定得更好”当作目标，而没有一样把“判定与后果之间那条线”当作可以拆开的东西。**试点期间，研究者在场、样本量小、后续配置往往被暂缓，损害被临时抑制；推广之后配置自动恢复，此前的改进同时失效。

十五、与最接近的几种说法的分界

遗忘曲线与保持研究说到的是：不复习则表现随时间下降。分离线在**方向**：遗忘是单向的，本章预测双向。判定形式：若两次测量之间的翻转全部是“会到不会”，则本判断错而遗忘足够；若“不会到会”这一支在无教学条件下显著非零，则反之。这一条最便宜，且能一次分开两者。

间隔效应与延迟测验 (Soderstrom & Bjork, 2015) 说到的是：练习分散、测验延迟有利于长期保持。分离线在**被安排的东西**：那一支安排的是**练习**的时点，本章说的是**判定**的时点。判定形式：把练习完全等距固定，只改判定时点与判定后的配置，间隔效应预测无差异，本章预测有差异。

经典测量理论中的重测信度说到的是：同一测验两次施测的相关系数反映测量的稳定性。分离线在**真值**：那一支假定真值在两次施测之间不变，把差异全部归给误差；本章说真值在移动。判定形式：若低重测信度的项目在延长间隔后信度继续下降而**均值不移动**，则是误差；若均值系统性移动且移动方向与初次判定相关，则是位移。

掌握学习（Bloom, 1968）说到的是：不达标就重学，直到达标。分离线在**达标判定的时点**：掌握学习把判定放在教学之后立即作出，反复判定直到通过，因此它的窗比在结构上最小。判定形式：若严格执行掌握学习的班级其双向翻转率低于常规班级，则本判断错；本章预测它反而更高——因为反复的短窗判定把每个人冻在更多个点上。**这是一条相反评价，也是本章最锋利的一条分离线。**

学习率的稳定性（Koedinger, Carvalho, Liu & McLaughlin, 2023）说到的是：在大规模数据中，学习者的学习率惊人地一致，差别主要在起点。分离线在**可观测段**：那一支测的是窗内可被记录的那一段练习——表现关系。判定形式：若在该数据集中把练习序列按“是否伴随一次配置后续条件的判定”分层，两层的起点差异不变，则本章无增量；若被判定过的那一层起点系统性更低而斜率相同，则本章所指的正是“起点差异”的一个来源。**这一条可以在已公开的数据上重新分析，不需要新采集。**

最近发展区说到的是：存在一个凭借帮助可及、独立不可及的能力带。分离线在**维度**：那一支说的是难度的边界，本章说的是读数的有效期。判定形式：把难度完全固定，只改判定时点——最近发展区预测无差异，本章预测有差异。

德博拉数与流变学是形式层的上游占位者，来自物理学之外的工程学科。把本章的命题剥掉学科名词，剩下的是：**以低于对象变化速率的频率观测，得到的不是缺席，而是一个伪像。**这个结构在流变学里叫德博拉数（Reiner, 1964），在信号处理里叫混叠

（Shannon, 1949），在统计学里叫时间聚合偏误，在生态学里叫**粒度与幅度**（Wiens, 1989）——同一片林地，用不同的取样粒度会得到不同的物种——面积关系，而它们不是谁更准，是不同的量。这几位必须点名，且本章的分离线只有一条：**它们讲观测尺度决定所观测对象的类别，本章讲观测之后据此配置的条件参与了对象此后的演化。**判定形式：若判定不配置后续条件时长期表现不受损（左下格），则本章相对这些占位者有增量，因为它们只管读数不管后果；若不作配置的短窗判定同样有害，则本章可被尺度问题吸收。

成功陷阱这一类**反转模板不适用于本章**，必须写明。本章的形式不是“越成功越失败”，是“观测尺度决定所读到的类别”。反转模板要求存在一项被推到极端的能力，而窗比过小是课表的结构性后果，与课程成不成功无关：一门长期失败的课程窗比同样小。这一条排除是必须写的，因为形式层最容易被误认到这个模板上。

划完之后再取一次交集。逐条问这几位各自结构性地看不见什么：遗忘研究把变化的方向当作给定，因此看不见反向翻转；重测信度把真值的稳定当作给定，因此看不见位移；掌握学习把“达标”当作可在教学结束时判定的事实，因此看不见判定本身的时点问题；最近发展区把能力的边界当作在任一时刻可读的，因此看不见有效期。这几行背后是同一件被当作给定的东西：

一个人在某一时刻“会不会某件事”，是一个有确定值的事实。

所有派别都在争“怎么测得更准”，没有一派把“这个值在被测量的那段时间里是否稳定”当作变量——因为一旦承认它在动、且动得慢于测量，“测得准”这件事本身就失去了对象。这是本章动摇的那块地。

十六、一处被逼出来的自相矛盾

上一节论证本章不能被既有说法还原。本节论证一件更强的事：教育测量这一行，在不承认本章变量的前提下无法自治。

这一行同时持有两条命题。**甲**：信度估计要求真分数在测量期间稳定——这不是它的推论，是重测信度这个量得以被定义的前提。**乙**：教育测验测的是学习，而学习是变化。

两条在同一套操作里互相拆台。若被测者在测量窗口内仍在变化，重测信度低就有两个不可分的来源：测量误差与真实变化。这一行现有的处置是**把重测间隔取短**，以“控制真实变化”——而按本章，缩短间隔正是把窗比推小，使判定落在最不该判定的时点上。于是：**为了让甲成立所做的事，恰好使乙所测的那个东西被冻住。**

这一行并非没有察觉。它的应对是引入“稳定性系数”与“等值性系数”的区分，并在手册中提醒间隔的选择须权衡记忆效应与真实变化。但这条出路是循环的：要判断两次测量之间的差异有多少来自真实变化，需要一个独立于测验的变化速率估计；而这一行获得变化速率估计的唯一途径，正是重复施测。**要让甲成立必须假定变化很慢，而变化是否很慢正是待检验的。**

因此这一行面临二选一：承认真值在测量期间移动，则信度的经典定义需要重写为一个带时间尺度的量；或坚持真值稳定，则必须解释为何“不会到会”的无教学翻转会发生——而这一支自己的重测数据里就有这些案例，它们通常被作为异常值处理。

本节的主张有边界，也留了退路：**若在实测中“不会到会”的无教学翻转率低到可忽略（比如低于百分之三），则甲乙的冲突近似不存在，本节失败。**这一条可以用现成的重测数据直接查。

十七、与本编另两章的分工

与第八章《先挡住其余的》。那一章管的是悬置期里自行结成、脱悬后无人拆除的那批连接，读数是悬置密度与脱悬率。分离线在**位置**：那一章管一次改变的**伴生物**，本章管它的**出口**。判定形式：一个把所有判定都推迟到重组完成之后的班级，本章预测其读数有效，那一章预测若悬置从未拆解，推迟判定只是更晚地读到同一批连接的后果。两个预测同时成立且不冲突。

与第七章《开着的时候没人来》。那一章处理的是上一轮打开的感受窗口与本轮内容不对齐，读数是对窗率。分离线在**轮次关系**：那一章管**入口**，本章管**出口**。判定形式：一个对窗率极高、每一轮都精确接上上一轮涌现的班级，若它同时每周作一次配置后续条件的判定，那一章判它健康，本章判它落在靶格——因为对上窗之后启动的那次重组，恰恰是最慢的那一类，也最容易被下一次判定冻住。**这一条是三篇里最反直觉的一处：对窗做得越好，被冻住的东西越要紧。**

三章撮起来才说得出的那句：同一次改变有入口、伴生物与出口三处可以失效；三处的读数数学上独立、可同批测量，而效应相乘不相加。本章单独看会得出“延迟判定”这个处方，而第八章会指出延迟判定并不拆除已结成的连接，第七章会指出延迟判定不改变窗口的开合——三处都对上，才有一次完整的改变。

十八、五类场景中的窗比，与其中两处收窄

中学数学的单元测是窗比远小于一的标准形态。教学到测验通常隔一到两周，而分数、比例、函数这几类概念的重组在文献里以月为量级。这一格的双向翻转率应当最高，而它几乎从不被测——因为单元测的卷子讲评之后就作废了，没有人会隔四周原样再发一次。

编程训练营的周考是方差最大的一格。递归、指针、异步这几个概念的重组时长个体差异极大，从三天到三个月都有。因此同一次周考里，一部分人的窗比大于一、一部分人远小于一，而判定对两部分人同等对待。本章预测这一格的双向翻转率最高且分布双峰。这一格也最容易做实验，因为训练营的课程周期短、班级多、数据全在平台上。

医学住院医的当场技能评估是窗比大于一的一格。缝合、穿刺、插管这类单项操作的重组以次计，当场判定是有效的。第一处收窄在这里：本章对这一格不适用，且必须明说——把本章的处方（明写不予判定）用到这一格，会把本来可以当场判定的东西也悬起来，反而降低训练效率。窗比大的场合，及时判定是对的。

语言能力分级是落差可以直接算出来的一格。欧洲语言共同参考框架的配套材料公开给出了相邻级别之间的指导学时，量级在数百小时；而多数语言课程按学期评级，一个学期的课时通常远低于这个数。这一格的窗比可以从公开文件里直接算，而且算出来的落差不是任何一所学校的执行问题——它是课程周期与能力跃迁尺度之间的结构性差距。

工业技能认证是窗比极大的一格：单项操作以天计，认证一次通过后长期有效。这一格反过来提示，窗比过大同样有代价——认证之后不再复测，被认证者此后的老化无人知晓，而这正是物理学那一处反向约束所说的上界。

第二处收窄由这五格合起来逼出：本章的处方不是“判定越晚越好”，也不是“判定越少越好”，而是窗比应当被知道。知道之后怎么办因格而异：窗比远小于一的格子里，处置是把配置与判定解耦；窗比大于一的格子里，处置是照常判定；窗比极大的格子里，处置是恢复复测。本章给不出一个统一的应然数值，这是它的限度，也是它相对于“多测”与“晚测”这两条口号的全部差别。

十九、三个领域反过来加上的限制

物理学削掉了本章最自然的一条处方。若只看德博拉数，本章会写下一句更强的话：把观测窗口拉长到超过重组时长即可。而物理老化证明，等得越久，系统本身已经因等待而改变——测到的已不是当初那个系统。翻回课堂：判定间隔存在上界，推迟到某个程度之后，测的是“这段时间里他自己长成的样子”，而不是那次教学的效果。这一处削掉的是“越晚越好”。

工程学削掉了本章的另一条自然处方，并改变了处方的形态。采样定理要求采样率高于最高频率的两倍；课堂无法提高判定频次——课时、疲劳与教学时间都不允许，而且提

高频次会加重抑制。工程的现实对策不是提高采样率，是**抗混叠滤波**：在采样之前把快于半采样率的成分滤掉，宁可明写“这部分我不测”，也不让它伪装成慢变化。翻回课堂，处方因此从“多测”或“晚测”变成第三种：**判定时明写哪些成分不予判定**，并在记录里为它们留一个“本次不判”的位置，而不是默认为不会。这一处改的是处方的形态，不是它的强度。

心理学削掉了本章的适用范围。提取诱发遗忘说明，不判定也不是中性的——一次提取就会抑制同类竞争项，而任何形式的练习都是提取。因此本章不能主张“少判定”就更好；它只能主张：**判定之后据此配置条件这一步，是可以被单独改动的那一处**。适用范围因此收窄到“判定会配置后续条件”的场合，对不配置条件的诊断性活动，本章无话可说。

二十、不适用的边界

三处反向约束已经削掉了本章的价值排序与部分处方。这一节把边界写全，共五条。

其一，判定不配置后续条件的场合不适用。这是辨别格的左下角：一次不计分、不分层、不改变任何安排的诊断性活动，即使窗比很小也不产生冻结构型。本章对它无话可说，而这一格在实践中并不罕见——它正是“低风险形成性评价”的定义。

其二，重组时长短于观测窗口的内容不适用。熟练化、程序性技能、事实性记忆，这几类的重组以次或以天计，两周的判定间隔对它们绰绰有余。把这类内容算进窗比是记错了账，而它们在课堂总量里占的比重不小。

其三，安全关键的即时判定不适用。实验室操作、临床技能、机械操作、应急程序中的合格判定必须当场作出，此处不存在推迟的余地。这些场合的处置是在模拟环境中安排复测，而不是延后真实判定；把本章的处方用在这里是危险的。

其四，一次性的资格判定不适用。当一次判定之后不再有可配置的后续条件（结业、颁证、终局性考试），冻结构型的第三个必要条件不满足。它可能仍然不公平，但那属于测量公平的问题，不属于本章。

其五，群体层面的统计判定不适用。本章的对象是个体身上的构型，而群体均值在窗比小的条件下反而可能稳定——因为不同个体的重组处在不同阶段，均值把它们抹平了。用班级平均分来检验本章的机制会系统性地得到零结果，而那不是本章被证伪，是分析单位选错了。

五条合起来把射程收到：**有连续轮次、判定会配置后续条件、内容的重组时长长于判定间隔、且分析单位是个体的那一类教学**。这个范围小于课堂判定总量的一半。

二十一、经验研究设计

本章的三个量都必须能在既有课堂里测出来。

核心设计是双时点复测。同一份判定卷，在原时点与四周后各做一次，中间不教学。“不教学”必须有可核证据，且须三源一致：教案版本、平台日志、作业记录。三源缺一，该样本作废——这一条不能松，因为本章最要紧的那一支（不会到会）恰恰最容易被中间偷偷讲过污染。

四类条件把差别收窄到一个参数上。**A 组**常规判定并据以配置后续条件。**B 组**常规判定但不配置——判定照做、结果照记，而分层、补救、跳过一律不依据它。**C 组**把判定推迟到估计的重组完成之后，并据以配置。**D 组**在原时点判定，但对识别为仍在重组的成分明写“本次不判定”，其余照常配置。

三处对照各回答一个问题。**A 与 B 检验**“配置”这一步是不是损害的来源，这是本章靶格设定的直接检验。**A 与 C 检验**单纯延迟够不够。**C 与 D 是本章最要紧的一处**——工程学那一处反向约束预测 D 优于 C，即抗混叠式的处置优于降低采样率；若实测 C 优于 D，那一处反向约束不成立。

重组时长的估计规程必须预注册，否则它会被事后调成想要的数。做法是在 B 组（判定不配置）中跟踪同一批学习者的表现随时间变化，取达到平台期所需时长的下四分位作为下界，并在主分析中报告该值取不同数时结论是否稳健。**这个估计有一处循环风险**：用一个理论所预言的现象去估该理论的参数。本章不掩饰它，只能靠预注册与敏感性分析降低，不能消除。

双向翻转率的编码要求分方向报告，且“不会到会”这一支须逐例核证期间无教学。编码者不得看该学生的原判定结果之外的任何材料。按亚型分层是强制的——不分层则两个方向互相抵消。

数据分四层：基线知识与先前成绩；判定时点与配置动作的记录（教案与教务系统里现成）；两次判定结果；十二周后的无支持迁移。

最低成本的先导研究不需要任何新采集。在已公开的大规模学习平台数据中，按“该练习序列是否伴随一次配置后续条件的判定”分层，比较两层学习曲线的起点与斜率。若被判定过的那一层起点系统性更低而斜率相同，本章所指的正是“起点差异”的一个来源。这一条完全在既有数据上完成，一个月之内可以跑完。

统计上以判定事件为单位，嵌套于学生、嵌套于班级。核心检验是“不会到会”支的比例与窗比之间的剂量——反应，而不是组均值的比较。

二十二、证伪条件与轮次预测

最强的竞争解释是一句朴素的话：**其实不就是学生忘了、或者本来就没学扎实吗。**

排除它需要一个只有本机制才预测得到的变量。本章取的是**双向翻转中“不会到会”那一支**：在两次测量之间无任何教学的条件下，遗忘解释在原理上预测不出这一支，学得扎不扎实同样预测不出。因此本章的核心证伪条件是：

控制总练习量、内容难度、教师效应与先前成绩后，若“不会到会”的无教学翻转率不显著大于零，或它与窗比之间不存在剂量——反应关系，则本判断为伪。

据此写出七条。

其一，如上；这是核心条。

其二，若窗比小而判定不配置后续条件的班级（左下格）在长期表现上与配置条件的班级无差异，则靶格设定错误，第二根轴应当换掉。

其三，若延迟判定与“判定时写明哪些成分不予判定”两种处置在长期表现上等效，则工程学那一处的处方形态之别没有经验意义。

其四，若复测差异随间隔呈线性而非幂律，则物理老化那一条的形式对应不成立，本章只剩比喻。

其五，若判定所覆盖内容与其同类未覆盖内容之间不存在系统性此消彼长，则提取抑制那一处的兑现失败。

其六，若严格执行掌握学习的班级其双向翻转率低于常规班级，则本章与掌握学习的相反评价被推翻。

其七，若在既有大规模学习数据中，按“是否伴随一次配置后续条件的判定”分层后两层起点无差异，则本章对学习率研究无增量。

第一条与第七条最便宜。第一条只需要一份现成卷子、四周间隔与不教学的约定；第七条完全不需要新采集，只需要在已公开的数据上重新分层。

轮次预测。本章的动力主张不是相关，是一个有次序的循环。

第 t 轮作出判定 → 第 t+1 轮据判定配置条件（分层、补救、跳过）→ 第 t+2 轮被冻住者的慢速演化在新条件下沿冻结构型继续，因而其表现与判定一致 → 第 t+3 轮的发起处从“内容”转到“这批学生的能力”，处置是提高判定频次，窗比进一步变小。

四步中任何一步顺序系统性相反，模型即需修正：若能力归因先于判定出现，则判定是归因的结果而非原因；若配置条件的改变与判定结果无关，则第二步不成立。

二十三、判定可以怎么改：三个动作

本章首先是一项解释，不是评价方案。但一旦把冻结结构型看见，判定的做法会变，且变的地方比想象中便宜。

第一，给“本次不判定”一个字段，并按实际填写率验收。成本最低的一件：在成绩记录里增加一个既非“会”也非“不会”的选项，其含义是“该成分仍在重组，本次不作判定”。这不需要任何新技术，只需要教务系统加一列。要紧的是验收方式——一个填写率长期为零的字段与没有这个字段是同一格，而且更有害，因为它已经把合规账结清了。

第二，复测一次，并分方向报告。一份现成的卷子、四周的间隔、期间不教学的约定，就是本章全部主张的最便宜的一次检验。分方向报告是关键：只报总翻转率，两个方向会互相抵消。这一条不需要任何人相信本章，做一次就知道。

第三，把配置与判定解耦。判定照做——它有别的用处，教师需要它来安排讲评。而分层、补救、免修、放弃这几项配置动作推迟到复测之后。这一条改的不是判定，是**判定与后果之间那条线**，而那条线在多数学校里是自动连上的，从未被人当作一个可以拆开的设计选择。

三条里只有第三条涉及制度安排，前两条是记录方式的改变。这一点本身可检验：若在完全不改变配置流程的条件下，仅按前两条运行的班级与对照班无差异，本章的处方失效。

二十四、反噬预言与伦理限制

窗比与双向翻转率一旦进入考核，最省事的达标办法会摧毁它们。

第一种是**推迟判定而不改任何别的**。窗比在账面上立刻变大，而判定仍然一次性、仍然配置后续条件、仍然把仍在进行的重组冻住——只是冻得晚一点。

第二种是**把复测做成走过场**：同一份卷子隔四周再发一次，而中间悄悄讲过，或者复测不计分因而学生不认真作答。前者使翻转率虚低，后者使它虚高，两者都无法从数据上分辨。

第三种最难查：**把“本次不判定”这一栏做成一个从不被填写的字段**。它挂在评价表上、写进制度文件、计入合规检查，而默认路径仍然是每一项都给出会与不会。

对此本章看不到彻底的出口。能规定的只有一条判定纪律：**复测的有效性以“期间无教学”的可核证据为准**——教案版本、平台日志与作业记录三者一致才算数；以及一条验收纪律：**“本次不判定”这一栏的验收看实际填写率，功能存在而填写率为零，与没有这一栏是同一格**。

因为窗比与翻转率都是可以拿去考核人的比率，伦理限制必须写死三条。

第一，这两个读数指的是**判定的时点**，不是学生的能力，也不是教师的水平。用它给学生贴“慢”或给教师排名，是对本章的误用。

第二，**不得为了把窗比做好看而在安全关键教学中推迟必须立即作出的判定**。实验室操作、临床技能与应急程序中的合格判定必须当场作出，此处的处置是在模拟环境中安排复测，而不是延后真实判定。

第三，若学校已据这两个读数作出不利安排（分层、免修、劝退），**必须公开编码规则、原始证据与复核通道**，并允许被评价者要求一次符合本章条件的复测。

二十五、自反

按本章自己的判据，这篇文章的窗比小于一。

它对一个需要数年才能重组的问题，在一次写作里作出了判定，并把这个判定以完成态交出。它的读者也将在一次阅读里对它作出判定，而这次判定同样会配置后续条件——引用、不引用、纳入综述的哪一格。本章因此正落在自己的靶格里，且落得比多数文本更深：它连一次复测的可能都没有留下，因为一段已经印出去的文字不会在四周后原样再来一遍。

本章的双向翻转率无法自测。能说的只有一句：**本章把可以让自己失败的那七条写得比结论更具体，其中两条不需要任何新数据**。这是本章能为自己配的唯一一件东西——它不能延长自己的窗口，只能把复测的工具交出去。

二十六、分层引用，与三个本章解释不了的洞

本章可以被分层引用。**第一层**是“冻结构型”作为一类对象，最强也最易倒——只要证明无教学期间不发生方向为正的变化，这一层就没有了。**第二层**是那张二维辨别格，即使命名不成立，把窗比与“是否配置后续条件”分成两根轴仍是可用的诊断工具。**第三层**是一条不依赖前两层的经验主张：同一份判定在无教学的两个时点之间会双向翻转，且“不会到会”这一支非零。第三层最稳，也最便宜。

有三个洞本章解释不了。

第一个洞：本章测不准重组时长，而窗比的分母正是它。估计规程有循环风险，敏感性分析只能降低不能消除。这使得窗比在实践中更像一个定性判断而非一个可比的数，而本章的整套论证却把它当作一个比值来用。双向翻转率不依赖分母，因此本章实际能交付的读数只有一个。

第二个洞：本章说不出被冻住之后的演化能不能被逆转。物理老化在材料学里可以被回火消除——重新加热到玻璃转变温度以上再缓慢冷却。人身上有没有对应物，本章不知道，也不敢类推。若存在，本章关于“配置条件使人只能沿冻结结构型继续”的整个主张需要加上一个时限。

第三个洞：本章无法分辨“判定造成的冻结”与“判定只是恰好发生在那一刻”。两者在数据上都表现为翻转与窗比相关。要分开它们，需要一个判定时点被随机分配的设计，而在真实学校里随机分配判定时点几乎不可能获得伦理批准——因为它意味着一部分学生的分层依据被人为推迟或提前。这是本章最实质的一处方法学困难，本章没有解决办法。

二十七、三条本章预测而尚无人测的现象

一个判断若只解释已知，仍可能是旧说法的新名字。本章因此列出三条它预测、而据本章所知尚无人测过的现象。它们不是推论的装饰，每一条不成立都要动本章的一部分。

其一，不及格者的复测翻转率应当高于及格者。通常的预期相反——及格者被认为掌握得更稳，因此更不容易翻转。而按本章，及格意味着重组已经完成或接近完成，翻转空间小；不及格者中混着大量重组进行到一半的人，翻转空间大。因此本章预测**双向翻转率与原判定结果之间存在负相关**，且这一相关在按亚型分层后主要来自表征重组与整合两类。这一条只需一份现成卷子做两次即可查，而它与直觉相反，因此是一个便宜的风险点。

其二，补救的内容性质应当决定翻转的方向。以操作步骤为主的补救材料会把仍在重述的学习者冻住，因而其组内“不会到会”的自发翻转率应当低于不作任何补救的对照组——**补救比不补救更糟，在这一格里**。这一条的实验设计很直接：不及格者随机分为分步补救、关系型补救与不补救三组，四周后复测。若不补救组的自发翻转率不高于分步补救组，本条为伪。

其三，“本次不判定”字段的填写率与教师的学科知识相关，而与教学理念无关。要认出一个学生正在重组而不是没学会，需要认出他那半套重述是什么——这依赖学科知识，不依赖对学生的态度。因此本章预测该字段的填写率与教师的学科背景相关，而与教学信念问卷得分无关。这一条若不成立，说明识别“仍在重组”不需要学科知识，那么本章关于该字段可行性的整套设想需要重写——它将变成一件任何人都能做、也因此任何人都都会做成走过场的事。

三条都不需要新的测量工具，只需要有人把已经在做的事按不同的方式记一次。

二十八、读数的可比性：三处本章自己交代不清的地方

一个读数若不能跨课堂、跨学科比较，它就只是某个人的记录习惯。本章的两个读数在可比性上各有难处，必须写在正文里而不是脚注里。

窗比的分母不可跨学科比较。分数概念的重组时长与递归概念的重组时长不是同一件事的两个数值，它们的“重组”指的可能是不同类型的过程。因此窗比只能在同一学科、同一内容类型内部比较；跨学科的窗比排名没有意义，而这恰恰是最容易被做出来的那种排名。本章对此的处置是：**窗比只用于同一门课的纵向比较，不用于横向排序，并把这一条写进伦理限制。**

双向翻转率的分母依赖判定的难度设置。一份很难的卷子会把大量已完成重组的人也判为不会，从而抬高“不会到会”这一支；一份很容易的卷子则相反。因此翻转率必须与判定的通过率一同报告，单独报告的翻转率不可比。这一条使本章最便宜的那件检验多了一个约束：复测必须用**同一份卷子**，而不是同等难度的另一份——后者引入的差异无法与真实翻转分开。

亚型分层依赖学科知识，因而不可自动化。要判断一个学生正处在表征重组还是策略选择，需要认出他那半套重述是什么，这依赖学科知识。任何试图用作答数据自动分层的做法，都只能按对错模式聚类，而对错模式与重组类型之间没有稳定映射——第一类与第三类的错误在卷面上可以完全相同。这意味着本章的读数在可预见的将来**不能被平台自动产出**，只能由懂这门学科的人手工编码。这一条大幅提高了本章的实施成本，也大幅降低了它被做成考核指标的风险；两个后果一好一坏，本章如实写下，不作取舍。

三处合起来说明：本章交付的不是一个可以填进仪表盘的数，是一套需要人来做编码。这是它的限度，也可能是它唯一的保护。

二十九、一处必须说清的伦理前提：谁有权要求复测

本章交出的最便宜那件检验——同一份判定隔四周原样再做一次——在制度上并不中性，因此必须把它的伦理前提说清，否则它会以最坏的方式被实施。

复测不能是一次新的考试。若复测计入成绩、影响排名、决定分组，它就变成了第二次判定，而按本章，第二次短窗判定只是把人冻在另一个点上。复测的唯一正当用途是**核验第一次判定的有效性**，因此它必须不计分、不排名、不进档案。这一条若守不住，本章的处方比不实施更糟。

要求复测的权利应当在被判定者一侧。一个学生若认为那次判定作得太早，他应当可以要求一次符合条件的复测；而这个权利在现行制度里不存在——申诉的对象是判分的对错，不是判定的时点。本章因此主张一条具体的制度改动：**在任何据判定作出的不利安排（分层、免修、劝退）生效之前，被安排者有一次要求复测的机会**，且复测须满足期间无教学的核证条件。这条改动的成本是一份现成卷子与四周的等待。

复测的结果不得用于追究教师。若“判定被复测推翻”成为教师的失分项，教师的理性反应是把第一次判定作得更保守、更晚、更少——而那会把本章的第三个动作（配置与判定解耦）挤掉，因为解耦要求判定照常作出。读数指向的是**判定的时点与配置之间那条线**，不是任何人的水平。

三条合起来说明一件事：本章所主张的那一件，其可行性完全依赖于复测能否保持低风险。一旦它被卷入问责，它会立刻变成它所批评的那个东西。本章看不到防止这一点的技术办法，只能把它写在这里。

三十、结论与一条写死日期的赌注

课堂上学习很少真正到位，不是因为它没有在进行。它在进行，只是进行得慢于课表。课堂每隔一到两周作一次判定，而许多内容的重组以月计；于是判定系统性地把仍在进行的变化读作零，并据此配置此后的一切条件，使被判定者只能从被冻住的那个构型继续。

三个领域各出一份不可替代的力。物理学给出两条最承重的形式约束——“**有没有在变**”是关于观测窗口的读数，不是关于系统的属性；以及**被判为不变的系统仍在演化，且以被判定那一刻的构型为起点**。心理学说明判定本身不是中性的观测，它会主动抑制没有被它覆盖的那些。工程学则给出了处方的形态：不是提高采样率，是在采样之前明写哪些成分不予判定。

这一理论把课堂评价的问题从“测得准不准”改为“测的时候它停下来了没有”；把改革方向从“更频繁的反馈”改为“为不予判定留一个位置”；把归因从“这批学生基础差”改为“这批学生被冻在哪一步”。

本章给出一条写死日期的赌注。**到 2031 年 6 月 30 日之前**，若至少三项在同一制度环境内、不少于八个平行单元、控制总练习量、内容难度、教师效应与先前成绩之后开展的研究报告：同一份判定在无教学的两个时点之间发生双向翻转，其中“不会到会”这一支显著非零，且该支的比例与窗比之间存在剂量——反应关系，则本章的核心判断获得初步支持。

以下明确**不算命中**：仅报告重测信度系数；仅报告遗忘曲线或保持率；仅延长测验间隔而未核证期间无教学；仅比较两个地区或两所学校；把延迟判定与补充教学、补救辅导打包实施而无法分离；把“本次不判定”栏的存在计为已实施而不查填写率。若符合上述条件的多项研究均未发现“不会到会”这一支，本章应被放弃或降格为一个描述性提法。

课堂不需要放弃判定——不判就无法安排。它需要的只是：在按判定安排一个人的后半年之前，先问一句那次判定是在他停下来之后作的，还是在他还动着的时候作的。

接续。双向翻转率是全书最便宜的读数之一，一份四周前用过的卷子即可，采法见附录一。“不会到会”那一支为什么必须单报、以及它在无解档单元里应当最高这条跨编预测，见第二编小结第六节。本章的伦理要求——任何据判定作出的不利安排必须公开编码规则、原始证据与复核通道——被第十二章第五节收为全书三条伦理之一。

第三编小结 三处熄火，与它们为什么不产生信号

一、三处

位置	熄火形态	推翻材料（来自该学科自己）	读数	靶格
入口	窗开着而无人到达	细胞学·序贯诱导（准备度是上一轮的产物，且有时限）	对窗率	课程表执行完美、进度分毫不差的班

位置	熄火形态	推翻材料（来自该学科自己）	读数	靶格
伴生物	悬置期内结成、脱悬后无人拆除的连接	化学·酰基迁移（挡住这个动作改写了它要固定的名单）	悬置密度、脱悬率	讲解清晰、每处都先“暂时不考虑”的课
出口	还在动的时候被读了数，并据此配置后续	物理学·物理老化（被判冻住者随等待时间持续演化）	窗比、双向翻转率	评价及时、反馈频密、分层精准的课

三行的靶格值得单独看一眼：三处的靶格都是好课。课程表执行到位、讲解层层铺垫、评价及时精准——这三样正是教学管理最想要的东西，而它们恰好分别是三处熄火最容易发生的条件。这不是反讽，是第二编第四节那个形状在失效侧的重现。

二、三处有次序

三章合起来给出一条任何一章单独都推不出的次序：**先对窗，再拆解，最后才谈判定时点。**

理由是位置本身给的。入口没对上，后面两处处理的是一批根本没有输入的东西——拆解一批从未被接应的连接，或者精确地读一个从未开始的重组。悬置未被拆除时把判定推迟，只是更晚地读到同一批连接的后果。因此三章各自的处方**不能按任意次序实施**，这与“改革措施可以并行推进”的常识相反。

三章还各自诚实地写下了单独实施的后果：第七章单独实施，会把更多慢速重组送进过早的判定（对窗做好了，出口没改，等于把更多真变化推到读数下面）；第九章单独实施，会把判定推迟而不改变入口；第八章单独实施，会拆解一批从未被接应的连接。**三处的处方各自都有副作用，而副作用正好由另两处承接。**

三、相乘不相加

三章各自独立声明的那句话，在这里必须写清楚它到底主张什么，因为它是全书最容易被误用的一条。

它主张的是：三处的效果**不是可加的贡献量**，而是**串联的通过率**。记入口的对窗率为 p_1 、悬置被拆解的比例为 p_2 、判定落在重组完成之后的比例为 p_3 ，则一次改变完整走完的比例是这三处**条件通过率的连乘**—— $P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1, E_2)$ ；只有在近似独立时，才可以用三个边际读数 $p_1 \cdot p_2 \cdot p_3$ 代入。它不是 $p_1 + p_2 + p_3$ 的某种加权。

它由此推出一条与教改惯行方向正面相撞的话：**单修任何一项，收益上界由另外两项的条件通过率之积封死**。若 $p_1 = 0.2$ 而 p_2 、 p_3 已经很高，把 p_2 从 0.8 提到 0.95 的全部收益不超过 $0.2 \times 0.15 \times p_3$ ——而这笔投入在任何一份改革方案里都会显示为“评价体系显著优化”。

它**不主张**的东西同样要写明：它**不主张**三项同等重要（乘法对称，但三项的可改动幅度与成本完全不同）；它**不主张**必须三项同时改（次序在上一节已给）；它更**不主张**“不能只改一项”是一条道德指令——在 p_1 已经很高的场合，单修 p_2 完全值得。

这一条的适用范围目前只到本编三章。把它扩到第二编三章、合成六项的乘积，是一次外推，第十四章处理，并且写明了外推可能失败的方式。

四、三章共同的形状：不产生信号

本编三章的失效有一个共同的性质，而这个性质比三个对象本身更值钱：

它们在发生的时候不产生任何可被现有记录系统接收的事件。

窗口关上时什么也没发生，因此在原理上不会以“失败”的形式出现（第七章）。悬置连接在原题型里不产生任何差错，因此不会被批改抓到；它只在换情境时决定表现，而那时已被归因为基础不牢（第八章）。被过早读出的那个数在所有可见指标上都正常，因为后续条件正是按它配置的，于是它自我实现（第九章）。

三处的共同后果是：它们不是被漏记的，是不可能被记的——现有记录的每一栏都需要一个事件来触发，而这三处的定义特征恰恰是不触发事件。这正是导论所说的“看不见的部分被算作零，而不是被算作未知”。未知会引出调查，零不会。

五、与第一编、第二编的接线

与第一编。第一编说可再入生成体不持续显露，休眠不等于死亡。本编第九章正是这一条的失效形态：判定在休眠期读数，把休眠读成了没有，并按“没有”配置了此后半年。两者说的是同一件事的两侧——第一编说“不显露不等于不存在”，第九章说“于是它被读成了零”。

与第二编。第二编三章说终点、路径与条件里有一部分是过程自造的；本编三章说这些自造物在入口、伴生物、出口三处各有一次被熄灭的机会。两编合起来才有完整的图景：东西是被造出来的，而造出来的东西没有账可记，于是它在三个位置上被顺理成章地当作不存在。

三条跨编预测已在第二编小结第六节列出（换位落在悬置期内则效应相乘；无解档单元里“不会到会”那一支最高；自建存量与空窗率负相关）。它们都需要两编的读数同时在场才能检验。

六、回到母题

母题说：课堂每天都在生产它看不见的东西，并把看不见的部分算作零。

第二编证明了“它在生产”，本编证明了“它看不见”，而本编多给了一层：看不见不是因为记录不够细，是因为这三类事情在发生时不产生事件。把记录做得更细，只会更精确地记录那些本来就会产生事件的东西——这正是第九章所说的“判定越及时，冻得越早”。

因此本编的落点不是“要更多观察”，而是：要观察那些不产生事件的位置，必须主动去看，而且要在没有信号的时候去看。第七章的处方是先花二十分钟看一眼昨天在他们身上开了什么口子；第八章的处方是在悬置解除时明确点名拆解；第九章的处方是在按判定安排一个人的后半年之前，先问一句那次判定是在他停下来之后作的，还是在他还动着的时候作的。三条处方都很便宜，而且都不需要任何新的评价工具。

七、本编的可推翻条件

第一条。若三个读数（对窗率、脱悬率、窗比）在同一批课上两两相关系数持续高于零点六，本编三章中必有一章是另一章的代理。

第二条。若在固定本轮内容与时间间隔、只改变上一轮内容的条件下，本轮学习效果没有出现可由上一轮所开窗口预测的差异，第七章的核心判断失去支持（第七章已把这条赌注写死到二〇三一年十二月三十一日，并明确列出了六种不算命中的情形）。

第三条。若在悬置密度匹配的两组之间，明确拆解与不拆解的班级在换情境表现上无差异，第八章的对象即使存在也不承重。

第四条。若双向翻转率中“不会到会”那一支在各类课堂上都接近零——即被判定为不会的学生在四周后几乎从不自行转为会——那么第九章所说的“还在动”就不是一个普遍现象，本章应降格为一个描述性提法。

第五条，针对本编整体。若把三处都做到位的班级与只做一处的班级相比，效果差异呈相加而非相乘，那么“效应相乘不相加”这条本编最锋利的判断作废，第十四章随之作废，而本书从一个结构性判断降格为一份清单。

第四编 结算：课堂的账本

前三编各自成立，而九章合起来还有一件事没人做：把九个对象、九个读数、三条结构模型放到同一张表上，看它们是不是同一件事的几种说法。

本编六章是新写的。它是全书增值所在，也是最可能被推翻的一编——因为它是唯一一处必须对九章同时下判断的地方。

第十章 同一组学科被问了两次

判断卡片：决定一次跨学科碰撞撞出什么的，不是学科的组合，是问题的形状。学科是材料，问题是模具。

一、一个不是巧合的重复

本书九章里，有六章两两共用同一组学科：

学科组	How 题（第二编）	Why 题（第三编）
教育学 × 化学 × 艺术学	第四章·换位区间	第八章·悬置连接
物理学 × 心理学 × 工程学	第五章·可倒推区间	第九章·冻结结构型
细胞学 × 工程学 × 艺术学	第六章·自建条件	第七章·空窗

这个重复最初是尴尬的。同一组学科被拿来回答两个不同的问题，如果撞出来的东西相似，那么第二编与第三编里必有一编是多余的；如果撞出来的东西完全不同，那就需要解释为什么——因为按通常的想法，一次跨学科工作的产出主要由所选学科决定。

结果是后者：六个对象两两之间没有任何一对可以互相还原，而且六章各自写了与对应章的双向判定（见第四、五、六章的“与本编另两章的划界”与第七、八、九章的“与本编另两章的分工”）。本章处理的就是这件事本身：**六个对象的差异是从哪里来的？**

二、同一门学科，两个互不引用的分支

先看取用点。

化学。第八章取的是**保护基策略与酰基迁移**——为固定一个位点而装上的基团会自己迁到别处去。这属于合成方法学里关于官能团操作的一支。第四章取的是**逆合成分析与高通量条件筛选的并存**，以及化学自己给出的选用判据（信息在哪一端多就从哪一端开始）。这属于研发策略的一支。两支在化学内部互不引用，也不能互相推出。

物理学。第九章取的是**玻璃态的物理老化**——被判为冻住的系统，其性质随冻住之后所等的时间持续演化。第五章取的是**初值问题与两点边值问题的适定性差异**。前者属于凝聚态与非平衡态物理，后者属于微分方程与数值方法。两支之间同样没有引用关系。

细胞学。第七章取的是**序贯诱导与感受期**——靶组织的可诱导窗口由更早一轮的信号打开。第六章取的是**生态位重塑与自组织**——住户自己分泌基质、招募支持细胞，造出维持自身所需的条件。前者属于发育生物学，后者属于细胞—微环境相互作用与类器官培养。

工程学与艺术学同理：工程学一次给的是**装配序列与配对**（第七章），一次给的是过程能力与“质量是造出来的不是检出来的”（第六章），一次给的是**稳健设计与最优点最脆**（第五章）；艺术学一次给的是作品的**未完成状态**（第八章），一次给的是**工作室讲评对做法的训练**（第四章），一次给的是**身体、媒介与共同判断场**（第一至三编各有取用）。

第一条结论由此得到：一门学科不是一个答案源，是一批互不引用的分支的集合。跨学科工作的实际单位不是“学科”，是“分支”。这条结论有一个可操作的后果：在检查

一次跨学科工作是否重复时，扫描“同一组学科”是不够的，必须扫描到分支一级；而反过来，两次工作即使标着同样三个学科名，只要取的是不同分支，它们就不是重复。

三、决定产物的是问题的形状

但分支不同只解释了“为什么可以不重复”，没有解释“为什么产物的类型不同”。而类型的差异比内容的差异更整齐：

How 题撞出来的三个对象都是“区间”或“存量”——换位区间有开合时刻与触发源，可倒推区间有三个档位，自建条件有存量、来源比与可交接性。三者都可以在事情正在进行时被指认，都有明确的开始与结束，都可以被安排。

Why 题撞出来的三个对象都是“没有信号的伴生物”——空窗是一个什么也没有发生的时段，悬置连接是一批在暗处结成而无人指认的关联，冻结构型是一个被当作结论的中间读数。三者都在发生时不产生事件，都只在事后、在换情境时才显出后果，都不能被直接安排。

这个差异与学科无关，与问题的形状有关。“该怎么运行”这个问法要一条次序，而次序必然由可指认、可安排的部件组成——问法本身就排除了那些不能被安排的东西。

“为什么很少发生”这个问法要一个驱动，而一个每天都在起作用却从未被处理的驱动，几乎必然是不产生信号的那一类——如果它产生信号，它早就被处理掉了。

第二条结论：学科是材料，问题是模具。同样三家学科，换一个问法，得到的不是一个对象的另一种说法，而是另一类对象。

四、这条结论怎样被推翻

上面那句话很容易说得漂亮，因此必须给出推翻它的办法。

判决设计。取本书的三对学科组合中的任意一对，向它提出本书没有问过的第四个问题——例如“一节课该怎么结束”，或者“什么才算这个班学会了”。由不知道本书结论的人按第三节所述的七步工序做一遍。

若得到的新对象与已有六个中的某一个**高度重叠**（可用同一句判据、同一个读数），则本章为伪：产物主要由学科决定，问法只是换了说法。

若得到的新对象与已有六个都不同，**且它的类型可由问法的形状预测**（要次序的问法给出可安排的部件，要驱动的要问法给出不产生信号的伴生物），则本章得到支持。

若得到的新对象与已有六个都不同，**但类型无法由问法预测**，则本章的前半句（学科是材料）成立而后半句（问题是模具）不成立——那意味着产物由取用点决定，而取用点的选择是任意的，这将使整套工序退化成为一种个人技艺，无法交给别人。

第三种结果对本书的方法最不利，也最可能出现。我把它写在这里，是因为它正是本书方法部分唯一一处没有被检验过的地方。

五、一处必须承认的窄

三对学科组合被重复使用，还有一个不那么体面的解释：**作者的取材面就这么宽**。九章的学科来源集中在化学、物理学、细胞学、工程学、艺术学、心理学、教育学、认知科学、经济学九门上，其中艺术学出现了六次，工程学出现了五次。

我不打算把这件事说成方法上的选择。它有一部分确实是取材面的限制。但有一件事可以据此检验：**若产物主要由学科决定，那么艺术学出现六次就应当使这六章带上某种共同的性质**；而实际上，取自艺术学的六个点（未完成状态、工作室讲评、共同判断场、制作中的思考、身体与媒介、材料参与）分别落进了六个互不相同的对象里，没有一条共同性质。这不是自辩，这是一条可以由读者自己核对的观察——第六节给出核对的位置。

六、六个对象两两不可还原：一张对照

对照	分离线	判决性观察
换位区间 ↔ 悬置连接	显式的角色转移 ↔ 暗处的副产物	悬置密度高而终点从未易主的课：后者有读数，前者为零
可倒推区间 ↔ 冻结构型	下一步有几个答案 ↔ 何时读数	无解档单元里“不会到会”那一支应最高（两章合用才推得出）
自建条件 ↔ 空窗	沉淀的物 ↔ 时机	自建条件被第二次使用时会打开窗口，故两读数应负相关
换位区间 ↔ 可倒推区间	终点的身份 ↔ 倒推的适定性	倒推唯一性最高的技能课：换位率趋零而档位满分
自建条件 ↔ 换位区间	过程的产物 ↔ 角色的互换	纯讲授班可攒下自己的记号（存量高、换位率零）
空窗 ↔ 悬置连接 ↔ 冻结构型	入口 ↔ 伴生物 ↔ 出口	三处效应相乘不相加（第三编小结第三节）

表里每一行都给了一个具体的课堂形态，在那里两个对象给出不同的读数。**六行合起来是本书“九个对象不是一个对象”这一主张的全部证据**。若其中任意一行在实测中不成立——两个读数总是同涨同落——那一行的两章必有一章是另一章的代理。

七、本章交给后面几章的东西

第十一章要把这六个读数（加上第一编的三个）放到同一张表上，处理量纲与采集；第十四章要处理它们的效应是相乘还是相加。而本章给这两章一条前提：**六个对象在概念上不可还原，已经在上表里逐行给出了判决性观察；但概念上不可还原不等于读数上不相关**。两者是两件事，后一件只能靠实测。

第十一章 九对象指标怎样压缩为六个课程主读数

判断卡片：九个对象各有一个指标，而其中只有三个便宜到可以每学期采一次；把它们采齐一次，比再写三章更有价值。

一、九个对象指标

先把九个对象各自的指标完整列一次。为便于比较，一律写成“取值范围·采集方式·最小时间跨度”。

章	读数	取值	怎么采	最小跨度
一	再生水平 $R = \min\{q,v,j\}/3$	0-1	对问题形成、修订链、判断场各作0-3级编码，取最低项	一个单元+一次撤架任务
二	接力率	0-1	记录每件半成品是否被另一人接手并改写	一个单元
三	持位可换度	0-1	撤换核心成员或媒介后，核心功能是否被接续	一个单元+一次撤换
四	换位率、触发源分布	次/课时；分布	记录终点易主的时刻与触发者	一节课
五	倒推唯一性档位	唯一/多解/无解	把学生此刻状态+目标交给五位教师问“下一步”	一次教研活动
六	自建存量、外来占比、可交接性	件；0-1；0-1	清点这个班手里可交接的物	一学期
七	对窗率	0-1	上一轮所开窗口中，被本轮内容接上的比例	两轮内容
八	悬置密度、脱悬率	次/课时；0-1	记录“暂时不考虑”类语句，及其后是否被明确拆解	一个单元
九	窗比、双向翻转率	比值；0-1	判定时点／重组时长；同一份卷子四周后重测的双向变化	四周

九个读数，十三个量。这个数目对一本要求“可以被不懂理论的人采集”的书来说太多了。本章要做三件事：把它压到可以在一学期内采齐的规模；说明哪些量必然相关；给出相关之后怎么办。

二、压成六个：按成本而不是按理论

压缩的依据不是理论重要性，是**采集成本**。理由很简单：一个采不起的读数，无论多重要，在检验里都等于不存在。

按成本分三层：

第一层，一次教研活动或一份现成卷子就能采（最便宜）。

倒推唯一性档位（第五章）：把一份真实的学生作业与该单元目标交给五位有经验的教师，各自写下“下一步该做什么”，看是同一件事、五件不同的事、还是都说不出来。一次教研活动。

双向翻转率（第九章）：一份四周前用过的卷子，原样再测一次，统计“会到不会”与“不会到会”两个方向的比例。四周间隔，无需新工具。

第二层，一个单元的课堂记录就能采。

悬置密度与脱悬率（第八章）：从课堂录音或教案里数“暂时不考虑”“这个以后再讲”一类语句，再看单元结束时是否有一处明确的拆解动作。

换位率（第四章）：数终点易主的时刻。判据是那一句问话——“有没有哪一刻，原本用来达成目标的那样东西本身成了要达成的东西，是哪一刻，由什么触发的”。

第三层，需要一学期或一次专门设计的任务。

自建存量（第六章）：学期末清点这个班手里有多少件可交接的物，其中多少是外来的。

对窗率（第七章）：需要连续两轮的记录，且要能指认上一轮打开了什么。这是六个里最难采的一个，因为“打开了什么”没有现成的记录格式。

（附录一按**采集时间**先后为这六项编号，与本节按成本分层的次序不同，对照时以名称为准。）

第一编的三个读数（再生水平、接力率、持位可换度）都需要一次专门设计的撤架任务，成本属于第三层偏上。本书的建议是：**第一轮采齐只采上面六个；第一编三个读数在同一批课上择一采集（推荐再生水平 R，因为它的编码规程最完整），用作外部效标而不作为对象读数。**

三、九对象构念总矩阵

只靠两两相关系数决定概念的生死是不够的。相关的高低同时受**测量信度、样本范围、课堂类型、分母定义、时间窗口、编码者一致性、天花板与地板效应**的影响：低相关可能只是噪声大，高相关也可能来自一个共同原因。因此在动用任何阈值之前，必须先有一张把九个对象在概念层面区分开的表——**若两个对象在下表的任意一行上取值不同，它们在概念上就不可互相还原，无论读数相关多高。**

对象	本体层级	分析单位	时间尺度	发生位置	必要条件	主要排除对象	主读数	外部效标
可再入生成体	关系存在物	班级／事件	周一学期	全过程	四项构成条件同时成立	普通迁移、记忆保持	持位可换度	撤架后独立表现

对象	本体层级	分析单位	时间尺度	发生位置	必要条件	主要排除对象	主读数	外部效标
接力（传递机制）	过程机制	半成品／人对	单元	交接处	被他人接手并改写	分工合作、小组展示	接力率	延迟变境表现
事件再生（显露形态）	显露形态	一次新学习	单元＋撤架	显露时刻	问题·修订·判断三项俱在	自我调节、适应性专长	再生水平（取最低项）	四周后自发学习
换位区间	过程区间	课堂片段	分钟—单元	终点	角色互换并交还	目标置换、目标漂移	换位率、触发源	变式迁移
可倒推区间	问题结构属性	单元×教师群	单元	路径	下一步的解个数可判	可控性、任务难度	倒推唯一性档位	条件变化后的表现
自建条件	沉淀的物	班级	学期	条件	可指着看、可交接、过程自造	生态位构建、分布式认知	自建存量、可交接性	新情境独立完成
空窗	时序失配	相邻两轮	日一周	入口	窗口开启且无人到达	关键期、遗忘	对窗率	本轮重组量
悬置连接	暗处伴生物	单元	单元	伴生物	悬置期结成且未被拆解	信息隐藏、干扰、误概念	悬置密度、脱悬率	换情境表现
冻结构型	被冻的中间读数	判定事件	四周一学期	出口	判定作于重组完成之前	测量时机与信度、标签效应	窗比、双向翻转率	四周后双向翻转

这张表怎么用。先看“分析单位”与“发生位置”两列：两个对象若在这两列上相同，才有必要担心它们是不是同一个东西（本书里只有第一编三行属于这种情形，其判决已写在第一编小结）。其余各行两两至少有两列不同，因此**概念上的不可还原已经由这张表承担，读数相关只用来判断“测得开与否”，不用来判断“是不是同一个东西”。**

到了实证阶段，正确的工具不是相关系数。应当做的是多特质—多方法分析（同一对象用两种方法测、不同对象用同一种方法测，看方法方差有多大）、探索性因素分析，样本足够时再做验证性因素分析或结构方程模型。第六节那四条保留规则只是**在这些分析做出来之前的临时办法**，不是概念判决的正式程序。本书在此把话说死：**不能只用两两相关系数决定概念的生死。**

四、五个数量口径，永久固定

书中不同位置出现过“九个对象”“九个读数”“十三个量”“六个读数”“第一编三个读数”等说法。它们在理论上兼容，但读者容易以为数量在变。此处把口径分层写死，全书此后一律按此引用：

层	数量	是什么
理论对象	9 个	上表九行，概念层的九个不可还原对象
原始观测量	13 个	各对象读数展开后的可记录量（如换位率与触发源分布记为两个）
对象级指标	9 个	每个对象各取一个代表性指标
课程级主读数	6 个	按采集成本压缩后、一学期一校可采齐的那六个
外部效标	3 个	撤架后独立表现、四周后自发学习、条件变化后的表现

九个理论对象与六个主读数之间的落差不是矛盾：**第一编三个对象的读数成本最高，第一轮采集里只择一采作外部效标**（本章第二节已说明），故进入主读数的是第二、三编的六个。

五、量纲：为什么不做加权总分

六个读数里，四个是比率（脱悬率、对窗率、双向翻转率、换位率的归一化形式），一个是计数（自建存量），一个是有序分类（倒推档位）。它们的量纲不同，时间尺度从四周到一学期不等。

一个几乎必然会被提出的要求是：**做一个总分**。本书明确反对，理由有三条。

其一，加权总分会重新制造它要消除的那个错觉。第一章已经论证过为什么用最低项而不用平均项：平均会让两个强项掩盖一个缺失项。六个读数落在不同的位置上（终点、路径、条件、入口、伴生物、出口），一个总分只会告诉人“整体不错”，而这正是被批评的那本账的做法。

其二，加权需要权重，而权重需要一个本书还没有的东西——各读数与外部效标关联强度的实测值。在此之前，任何权重都是编的。

其三，也是最要紧的一条：六个读数的效应可能是相乘而不是相加（第十四章）。若是相乘，加权总分在数学上就是错的形式——它会把一个 $p_1=0.1$ 而其余五项接近 1 的班评为“良好”，而按乘积算它接近于零。

本书给出的替代形式是一张六格的档位卡：每个读数各自报三档（低／中／高，档界见附录一），六格并排，不合成。读的时候只看一件事：**最低的那一格在哪里。**

六、哪些读数必然相关：一次预注册

本书在数据采集之前，先把自己对相关结构的预测写死。这样做的理由与第五章要求“判据必须在读出结果之前锁定”是同一条：**预测写在后面，就没有检验可言。**

以下预测按本书的机制推出，请在采齐之后原样比对：

1. **自建存量 ↔ 对窗率：预测正相关，约 0.3–0.5。**机制在第二编小结第六节：一件班级自己造的工具，被第二次使用时会打开一个感受期。若实测为零相关，这条跨编预测失败；若高于 0.6，第六章与第七章必有一章是代理。
2. **换位率 ↔ 悬置密度：预测弱正相关，约 0.1–0.3。**两者共享“教师愿意在课中停下来”这个共同原因，但第四章已用教案偏离次数作了正交性论证。若高于 0.6，说明两者测的都是教师的灵活性，两章的独立性都不成立。
3. **双向翻转率 ↔ 脱悬率：预测接近零。**两者机制上无联系。若发现强相关，说明存在一个本书没有指认的共同原因，而那个共同原因比本书的九个对象都更重要。
4. **倒推档位 ↔ 自建存量：预测多解区最高、唯一区最低。**这是第五章与第六章各自写下的对赌：第五章预测正相关（多条路都要留痕才选得出来），第六章不作主张但接受这条检验。**若实测发现唯一区反而积累更多，第五章那条预测为伪，而第六章不受影响。**
5. **对窗率 ↔ 双向翻转率：预测负相关。**对窗率高意味着重组更常被接受，因而更可能在被判定前完成，“不会到会”那一支应当更少地被冻住。这条是第七章与第九章合用才推得出的，两章单独都没写。

总的承诺：六个读数中，两两相关系数持续高于 0.6 的对数不应超过一对。若超过，本书的九个对象是虚高的，应按第六节的规则合并。

七、一学期一校的采集方案

以下是把六个读数在同一门课上采齐一次的最小方案。它假定：一所学校、一门课、两个平行班、一位教师、一个学期（十六至十八周）、一位不懂本书理论的编码者。

第 1 周：建立记录。列出本学期的单元表；为“这个班手里有什么”建一个清单（初始通常为空或只有外来物）。约两小时。

第 2–3 周：采倒推档位。取本单元的一份真实学生作业与该单元目标，交给五位教师，各自独立写下“下一步该做什么”（不许讨论）。记录五个答案的分布。约一次教研活动。

第 3–6 周：采悬置密度与换位率。对四周内的课堂作录音或由观察者计数：每课时“暂时不考虑”类语句的次数；终点易主的时刻与触发者。单元结束时记录是否有明确的拆解动作（脱悬率的分子）。约每周一小时。

第 6 周：第一次卷子。用一份覆盖本段内容的卷子作基线。

第 7–10 周：采对窗率。这是最费力的一项。每单元结束时，由教师用五分钟写下“这一单元在他们身上打开了什么”（一到三条，写具体的问题或困惑，不写“兴趣提高了”）；下一单元结束时，由编码者判断这些条目中有多少被本单元的内容实际接上了。

第 10 周：第二次卷子（与第 6 周同一份），计双向翻转率。

第 11–15 周：继续第 3–6 周的记录；同时择一次课作撤架任务，采第一编的再生水平 R 作为外部效标。

第 16–18 周：清点自建存量。判据是那一句问话——“这一节课/这门课结束时，这个班手里多了什么，是下一节课/下一届可以直接拿来用的”。逐件登记：是什么、谁造的、被用过几次、能不能交给一个新来的人直接用。

总成本估算：编码者约 60–80 人时，教师额外投入约 10 人时，五位教师各 20 分钟。这是全书最要紧的一笔投入，而它比一次公开课的准备成本还低。

八、如果相关太高，怎么办

本书预先写下保留规则，以免届时临时找理由：

规则一：保留可清点成本更低的那一个。在两个高度相关的读数之间，保留第一层／第二层的，弃第三层的。

规则二：保留可追溯性更好的那一个。能留下原始事件记录（谁在哪一刻做了什么）的，优先于只能留下比率的。

规则三：被弃的那一章不删，降为被保留那一章的一个亚型。理由是概念上的不可还原（第十章第六节）与读数上的相关是两件事；读数相关只证明它们测不开，不证明它们是同一个东西。降格保留，等待更好的测量。

规则四：若六个里有四个以上互相高度相关，那么本书的问题不在某一章，在整个架构。届时应当承认：九个对象其实是两三个，而本书的分章是一次过度切分。这一条我写在这里，是因为它是真有可能发生的结果，而一本书如果不预先写下“我在什么情况下承认自己切得太细”，那它在数据面前就只会不断解释。

九、本章没有做的事

本表从未在任何一门真实的课上跑过。六个读数的档界（附录一给出）是按机制推的，不是按分布定的；第四节的相关预测是按机制推的，不是按预实验定的；第五节的成本估算是按每项操作的用时推的，不是实测的。

这一节写在这里不是免责，是定位：**本书到此为止提供的是一份可以被执行的检验方案，不是一份已经通过检验的结论。**九章的全部论证只能做到这一步；再往前一步，必须有人真的走进一间教室。

第十二章 九句判据合成一份问卷

判断卡片：九句判据可以合成一份问卷，九个读数却不能合成一个总分——因为问法是记录的入口，读数是结算的输出；能合的是入口，不能合的是输出。

一、九句原样

九章各给了一句不含判断词的问法。它们的共同规格是：**问的是记录，不是判断**；一个不懂本书理论的人拿着它去查教案、录像、平台日志、草稿版本与分工记录，应当能得到确定的答案。

第一章。给同一个班一次撤架任务：问题、修订链、判断场三项中，最低的那一项处在什么水平——学生能不能从材料或失败里自己识别出值得处理的差异？后续步骤有没有一处因证据而改动，并说得出口改动前后的差别？结论有没有进入一个他人可以找到入口、可以反对的场？

第二章。十四天后，撤掉原例、原话术和原分组，换一个表面不同而结构相同的问题：**班里是否有人能独立重建那条辨别，另一人能接着把它推进，并据此改写下一轮的提问、分工或评价？**

第三章。撤掉原来的例题、教师提示和同伴分工，换一个表面陌生但关系同构的任务：**学生能否自行重建关键关系，并据此改变下一步的问题、表征、工具、行动或协作方式？**

第四章。这一节课里，有没有哪一刻，原本用来达成目标的那样东西本身成了要达成的东西？是哪一刻，由什么触发的？

第五章。把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的“下一步”是同一件事、五件不同的事、还是都说不出来？

第六章。这一节课结束时，这个班手里多了什么是下一节课可以直接拿来用的？

第七章。上一节课结束时，学生身上被打开了哪一处口子？这一节的头二十分钟讲的是不是那一处？（配套的一句：教案上，后一节课的起点是由前一节的内容序号决定的，还是由前一节实际产生的某样东西决定的？）

第八章。这学期明确说过“这个以后再讲”的地方，有几处后来真的讲了，并且要求学生把它和当时学的那一处重新接了一次？

第九章。把这门课上一次作出“会了没有”判定的那份卷子，隔四周原样再做一次、期间不教，有多少人的结果会翻过来？两个方向各占多少？

二、为什么问法可以合，读数不可以合

第十一章拒绝了加权总分。本章却要把九句问法合成一份问卷，两者不矛盾，因为它们处在流程的两端。

问法是记录的入口。它规定去看什么、看哪一段、看谁的记录。九句问法指向九个不同的记录位置（教案的起点依据、课堂语句、终点易主的时刻、教研活动的五份答案、学期末的物件清单、四周前后的两份卷子、撤架任务的表现），彼此不冲突，合起来正好构成一次完整的观察。**合成一份问卷，节省的是走进教室的次数，不是判断的层次。**

读数是结算的输出。把六个输出合成一个数，等于替读的人做了一次他没有授权的加总，而加总的权重本书没有；更要紧的是，若六项的效应是相乘的（第十四章），加总在数学上就是错的形式。

因此本章合的是问卷，不是分数。**问卷末尾没有总分栏。**这一条是本书为数不多的硬性版式要求：一张最后会有人去填总分的表，迟早会变成一张只有总分被读的表。

三、三块，不加权，看最低那一格

问卷分三块，对应三编，每块三问。

第一块：这一段留下了什么（What）。撤架任务一次，问三章的三句。三项各作 0-3 级编码，**取最低项**，不取平均。取最低项的理由第一章已给：问题为零时修订只能围绕外部问题运转；修订为零时问题只停留在口号；判断场为零时行动无法与任意主张区分。一个平均分会让两项强的盖住一项缺的，而那正是本书从头到尾在拆的那种错觉。

第二块：这一段怎么运行的（How）。三句各出一个读数：换位率与触发源、倒推唯一性档位、自建存量与可交接性。**三项分别报，不合并**——它们是三个独立的旋钮（第二编小结第三节），合并会把“三个旋钮”这条主张在表格层面自我否定。

第三块：这一段哪里熄了火（Why）。三句各出一个读数：对窗率、脱悬率、双向翻转率。**三项按串联读：**先看对窗率，若它很低，另两项的数值只说明“在没有输入的情况下另两处做得如何”，不构成改进的依据（第三编小结第二节的次序）。

三块并排，共六格加一个最低项。**读法只有一条：看最低的那一格在哪里。**问卷的全部产出是一句话——“这个班这一段，最先该动的是哪一处”。

四、编码规程

谁来编码。编码者不应是授课教师本人，理由不是信任问题，是第五章那条：一位在某内容上教了十年的教师，他自己那条路已经高度自动化，他给出的“下一步”确实稳定而唯一——**对他自己而言**。同一种自动化会让他看不见自己课上的悬置与换位时刻。教师应当是被问的人，不是编码的人。

盲化到什么程度。编码者不需要读本书；他需要的是附录一那份手册与九句问法。手册的写法刻意做到“不解释为什么”，就是为了让编码不依赖对理论的认同。**若一个读数只有认同本书的人才采得出来，那个读数不合格。**

一致性怎么查。两名编码者独立编码同一批记录的不少于两成，报一致性系数。本书对各项的最低要求不同：计数类（悬置密度、换位率、自建件数）要求高，比率类（对窗率、脱悬率）要求中等，等级类（0-3 级编码）要求最松但必须报出分歧的分布。**低于要求的项不得用于任何比较，只能作为该教师自用的诊断。**

分歧怎么裁。不裁。分歧本身是数据：对窗率上的编码分歧，往往正说明“上一节打开了什么”这件事在记录里本来就没有位置——而这恰是本书要说的事。规程要求把分歧原样保留并报出，不许由第三人拍板消掉。

什么时候采。第十一章第五节给了一学期的日历。这里补一条：**九句里有四句必须在事情正在发生时采**（换位时刻、悬置语句、上一节打开了什么、终点易主的触发源），事后回忆不能替代。这四句决定了这份问卷不能被做成一份期末填报的表格。

五、三条伦理，写死

九章各自写过伦理限制，本章把它们合成三条，作为使用这份问卷的前提。**不接受这三条的场合，本书拒绝提供这份问卷。**

其一，读数指向课堂位置与关系，不指向个人。它不得用于给学生贴标签、分层、排名，也不得用于给教师排名。第九章的整章论证正是关于“一次过早的判定如何配置一个人的后半生”；用它来作判定，是把这本书批评的那件事换个名字再做一遍。

其二，不得为提高读数而制造条件。具体禁止三样：为提高换位率而安排表演性的“学生提议”；为提高自建存量而把教师做好的工具算作班级自造；在安全关键的任务中为提高持位可换度而强制轮换角色或人为制造故障。第三条是第三章明写的。

其三，任何据此作出的不利安排，必须公开编码规则、原始证据与复核通道。这一条是全书唯一一条制度性要求。理由在第九章：一次判定之所以危险，不在它可能错，而在它错了也不会被发现——因为后续条件正按它配置。**复核通道是把“可以被发现”这件事重新装回去的唯一办法。**

六、它会被怎样填坏

一份问卷一旦被使用，就会被填坏。预先写出它会怎样坏，比事后惋惜有用。

坏法一：把“是否发生”填成“是否安排”。教案上写了“这一段用你们的方案来检验上周定的标准”，于是换位率记一次——而课堂上那一段并没有真的易主。**对策：换位率必须同时报触发源**，教师宣布不算，须有一个可指认的触发事件（一个学生的做法、一次异常、一件半成品）。

坏法二：把外来物记成自建。评分表、题库、教师做的模板被记进“这个班手里多了什么”。**对策：自建存量必须同时报外来占比与可交接性**——是不是可指着看的物、能不能交给一个新来的人直接拿去用。

坏法三：把比率的分母做小。悬置密度低不是因为悬置少，是因为记录的人少记了几句“暂时不考虑”。**对策：所有比率类读数必须同时报分子与分母的原始计数**，且原始事件记录须保留到复核期结束。这一条使造假的成本高于改进的成本——这是本书对反噬唯一有效的技术手段（第十五章展开）。

坏法四，最难防的一种：把问卷本身当作教学目标。教师按九句问法去设计课，于是九个读数一起上升，而学生什么也没多得。**对策：本书给不出技术对策，只能给一条制度对策——见第十五章第六节的“失效期”。**

七、本章的可推翻条件

第一条。若两名不读本书、只拿手册的编码者，在同一批记录上得到的一致性系数持续低于可用水平，那么这份问卷不成立——它测的是编码者的理论倾向，不是课堂。这一条最要紧，因为“不含判断词”是本书对这九句话的全部承诺。

第二条。若把九句问法按第三节的分块采集之后，“最低那一格”与外部效标（撤架后独立表现、四周后的自发学习）**没有关联**，那么问卷即使可靠也不承重。

第三条。若在同一批课上，教师自评的结果与编码者的结果高度一致，那么这份问卷的成本可以被大幅削减（教师自评即可），而本章第四节“教师不应编码”的论证是多余的。这是一条对本书**有利的失败**——它会让检验变便宜，因此我特别希望有人去做它。

第十三章 三条结构模型能否收成一个式子

判断卡片：目前不能。三条推导可以被同一句话概括，但概括不是统一；而它们的退化条件确实是同一句话——某一端的不确定性归零时，那一段就不存在。

一、三式并排

第四章： $\beta \cdot r \cdot d\Delta U_M/dt | (t = t^*) = c$ （一阶条件，决定最优换位时长）

第五章： $E[L] = \sum_n P(n | \hat{n}) \cdot L(M, n)$ ，且 $L_2 \gg L_1$ （期望损失比较，决定默认方法）

第六章： $A^* = \alpha(1-c)N / \delta$ （一阶线性微分方程的稳态，决定自建存量）

三式的数学类型互不相同：一个是带凹函数的最优化一阶条件，一个是离散决策问题的期望损失比较，一个是连续时间存量方程的不动点。它们各自推出了正文没有说过的东西，这一点第二编小结已经清点。本章要问的是另一件事：**它们是三个不相干的模型，还是同一个模型的三个投影？**

这个问题不是形式上的洁癖。如果它们是同一个模型的三个投影，那么本书的九个对象背后就只有一个机制，而第十一章第六节的第四条规则会被触发——本书是一次过度切分；如果它们确实不相干，那么“课堂运行里被当作给定的那些东西有一部分是过程自己的产物”这句话就只是一个**家族相似**，不是一条定律。

本章的结论落在中间，而且我把三次失败的尝试原样写出来。

二、第一次尝试：都写成投入—产出最优化

最自然的想法是把三式都写成“某个量的最优水平由边际收益等于边际成本决定”。

第四章本来就是这个形式。第五章可以改写：把“选倒推还是选推进”写成在 $P(1)$ 上的阈值决策，阈值由 L_2/L_1 决定——形式上也是一次边际比较。

失败在第六章。 $A^* = \alpha(1-c)N/\delta$ 不是任何东西的最优水平，它是产生与失效两个流平衡时的存量。没有人在选择 A^* ；它是由 α 、 c 、 N 、 δ 决定出来的。硬要把它写成最优化，就得虚构一个“课程建设者在自建条件与外来资源之间权衡”的目标函数——而第六章的全部论证恰恰是**没有人在做这个权衡**：覆盖度 c 的提高是一件独立地被追求的好事，自建条件的下降是它的副产品，谁也没有把这两样放在一个目标函数里比较过。

虚构一个不存在的目标函数，是把结构性后果重新读成选择问题。这正是本书在导论里说要避开的那种写法。第一次尝试因此不是技术上失败，是它一旦成功就会说错话。

三、第二次尝试：都写成通过率

第二个想法来自第三编：把每一处写成一个通过率，三处相乘。

第五章可以：三档可以写成一个状态到“下一步被正确处理”的转移概率。第四章勉强可以：一次换位可以写成“进入—取得终点资格—交还”的三步通过，其中交还成功率 r 天然就是一个通过率。

失败在时间结构。通过率模型处理的是事件——一次尝试，通过或不通过。第四章的换位不是事件，是区间：它有时长 t^* ，而最优时长正是它要求解的东西；把它压成一个通过率， t^* 就消失了，而 t^* 恰恰是这一章唯一的形式产出。第六章更麻烦： A^* 是存量，不是事件也不是区间，它有历史依赖（今天的存量取决于过去所有时刻的产生与失效）。

事件、区间、存量是三种不同的数学对象，把它们塞进同一个通过率框架，只有先把两个压成一个，而被压掉的正好是它们各自的形式产出。

四、第三次尝试：三个特例的共同结构

第三个想法不再追求形式统一，改追求结构对应：三式是否都在描述“一个内生量被当作外生量”这件事？

这一次站得住。

	被当作外生的量	它实际由什么内生地决定	该量归零时的后果
第四章	终点的身份	由信息在两端的分布决定，可临时互换	$\beta \rightarrow 0$ ：不该换位，换位区间不存在
第五章	“下一步”的存在性	由知识在这个班的分布决定，可以是 0、1 或多个	$P(1) \rightarrow 1$ ：只有一条路，档位问题不存在
第六章	条件的来源	有一部分由这个过程自己产生	$c \rightarrow 1$ ：全部条件外来，自建条件不存在

三行的最后一列是本章唯一真正的收获，下一节单说。但要先把这次尝试的性质说清楚：这是一次结构类比，不是形式统一。三行左侧的“被当作外生的量”分别是一个身份、一个存在性、一个来源——它们不是同一个数学对象的三个取值，而是三种不同类型的东西。把它们排在一张表里，得到的是一个可用的诊断框架，不是一个可推演的模型。

一张能对齐的表不等于一个能推演的式子。两者的区别在于：式子可以推出没被放进去的东西（三式各推出了三条），表只能整理已经放进去的东西。本章不打算把后者说成前者。

五、三者的退化极限是同一句话

上表最后一列值得单独看。三章各自在自己的形式部分里，都推出了一个“这个对象不存在”的极限：

$\beta \rightarrow 0$ （手段与产物几乎无关的内容，如背诵、查表）：一阶条件在任何正的 c 下都无解，最优换位时长为零。在这类内容上，从不换位是正确的。

$P(1) \rightarrow 1$ （真正只有一条路的内容）：倒推唯一，档位判定失去意义，直接倒推即可。

$c \rightarrow 1$ （每一个需要工具的场合都已经有工具）：稳态存量为零，一门完备的课程在结构上不可能有自建条件。

三个极限条件写法完全不同，说的却是同一件事：

当某一端的不确定性归零时，那一段就不存在。

第四章归零的是手段端还有多少可学；第五章归零的是“下一步”的多样性；第六章归零的是缺口。三处都指向同一个位置：本书的九个对象全部长在不确定性还没有被消掉的那一段上；不确定性消掉的地方，它们不是变弱，是不存在。

这一句是本书能给出的最接近“统一”的东西，而它的地位要说清楚：它是一条关于适用范围的陈述，不是一条可以推出新预测的定律。它告诉人这些对象在哪里找得到、在哪里不必找，但它推不出任何一个具体的数值关系。

它还有一个副作用值得写下来：它使本书的适用范围与“教得越熟越好”这条常识正面相撞。一门被打磨得极其成熟的课——手段确定、路径唯一、资源完备——按本书的框架，恰恰是九个对象全部消失的地方。这不是说这样的课不好；这是说本书对它无话可说，而一本书能说清自己对什么无话可说，比它多说三章更值钱。

六、要收成一个式子，还缺什么

如果有人要继续做这件事，本章把缺口指出来。

缺的是一个共同的状态变量。三式各自的状态变量是“手段端的不确定性 U_M ”、“倒推解的个数 n ”、“可交接自建条件存量 $A*$ ”。要统一，需要一个更基本的量，使这三个都是它的函数。

一个候选是：这个班在这一刻手里有什么——把它写成一个集合（工具、记号、已被指认的问题、已结成的连接），三章各是这个集合的一条更新规则：第四章说这个集合里的元素可以临时改变角色（手段→终点→手段）；第六章说这个集合会因缺口而增长、因失效而衰减；第五章说这个集合的构成决定了“下一步”的解的个数。第三编三章则是三条删除规则：窗口关上删掉一批未生成的元素，悬置未拆留下一批未登记的元素，过早判定把一个正在变化的元素冻成常量。

这个方向看起来很整齐，而本书给不出它。给不出的具体原因有三个，写清楚以免后来的人重走：其一，这个集合的元素不同质（一件工具与一条连接不能相加），因此它不是一个可以取模的空间；其二，第五章的 n 依赖的不只是这个班有什么，还依赖教师队伍的知识分布（同一学生同一目标，换一批教师档位会变），而那不是这个班的状态；其三，第三编三章的删除规则里有两条不产生事件，因此在任何以事件为基础的动力学里都写不出来——一个不产生事件的转移，在马尔可夫框架里是隐形的，而这正是本书全部主张的核心难处。

第三个原因最要紧：本书的对象之所以难以形式化，与它们之所以难以入账，是同一个原因。这不是巧合，也不是借口——它是本书母题在数学一侧的重现。

七、本章的可推翻条件

第一条，也是我最希望被触发的一条。若有人真的把三式收成一个式子，并且从中推出至少一条三章各自都推不出的新预测，本章作废。这是本书最受欢迎的一种作废——它意味着九个对象背后确有一个机制，而本书只是把它拆开来看了九遍。

第二条。若有人指出上表的三行左侧其实是同一类对象（例如都可写成“某个二元关系的方向”），那么第四节的“结构类比不是形式统一”这句判断过于保守，本章第二、三节的两次失败应当重做。

第三条。若在实测中发现，在 β 很小、 $P(1)$ 很高、 c 很高的课上（即三个对象都应当不存在的地方），本书的读数仍然显著非零且与效标相关，那么第五节那句“当某一端的不确定性归零时，那一段就不存在”是错的——而它一错，本书对自己适用范围的全部陈述都要重写。

第十四章 相乘还是相加

判断卡片：串联的是**条件通过率的连乘**，不是可加的贡献量。若这一连乘成立，单修任何一项的收益上界由另外两项的条件通过率之积封死——而这与“抓重点、单点突破”这条最通行的改革策略正面相撞。

一、第三编那句话的确切内容

第七、八、九章各自独立地写下了同一句：三处的**效应相乘不相加**。第三编小结把它写成了形式：记入口的对窗率为 p_1 、悬置被拆解的比例为 p_2 、判定落在重组完成之后的比例为 p_3 ，一次改变完整走完的比例是这三处条件通过率的连乘（近似独立时可用 $p_1 \cdot p_2 \cdot p_3$ 代入）。

这个乘法的根据是**位置**，不是统计。三处不是三个并列的影响因素，是同一次改变依次要经过的三道门：没有输入，后面两道门开得再大也没有东西通过；输入进来而伴生物无人拆除，出口读得再准也只是准确地读到一个带着杂物的结果。**串联的通过率天然相乘**，这不需要数据来证明，需要数据来证明的是“这三处确实是串联的”。

二、乘法的确切写法：条件通过率的连乘

上一节说三处“相乘”。这个词必须先被写准：若 p_1 、 p_2 、 p_3 各自是边际比例，直接相乘需要独立性；而位置上的串联只支持**条件概率分解**。因此本书此后一律称之为**条件通过率的连乘**，不称“三个读数相乘”。

其一，正确的写法。记 E_1 为“上一轮打开的窗口被本轮接上”， E_2 为“悬置期结成的连接被拆解”， E_3 为“判定落在重组完成之后”。一次改变完整走完的概率是

$$P(E_1 \cap E_2 \cap E_3) = P(E_1) \cdot P(E_2 | E_1) \cdot P(E_3 | E_1, E_2).$$

这一步不需要任何独立性假设，它只是概率的链式法则。**串联的位置结构支持的正是这个分解，而不是三个边际比例的乘积。**

其二，什么时候可以用边际读数相乘。只有当 $P(E_2|E_1) \approx P(E_2)$ 、 $P(E_3|E_1, E_2) \approx P(E_3)$ 时，实测的脱悬率与窗比才可以直接代入。这是一条**可以被检验的额外假设**，不是位置结构的推论。它有明确的失效方向：一位对窗做得好的教师，很可能同样习惯在单元末尾点名拆解——那么 $P(E_2|E_1) > P(E_2)$ ，直接相乘会**低估**；反过来，若对窗抬高了进入重组的人数，从而使更多人在判定时仍在动，则 $P(E_3|E_1, E_2) < P(E_3)$ ，直接相乘会**高估**。

其三，口径。准确的说法是：

位置结构支持条件概率分解；把分解写成三个边际读数的乘积，则另需一条独立性假设，而这条假设本身要被数据检验。

其四，政策推论因此要加一个限定，但方向不变。“单修一项，收益上界由另外两项的条件通过率之积封死”这条推论，在条件概率写法下变为：**收益上界由另外两项的条件通过率之积封死**。若三者正相关，上界比边际乘积算出的更宽松；若负相关，则更紧。无论哪种，只要 $P(E_1)$ 很低，后两项无论怎样都乘不出可观的通过率——这正是本章的承重结论，它只依赖链式法则，不依赖独立性。

其五，采集上的直接后果。这意味着第十一章那份采集方案必须按个案配对记录，而不是各读一个班级比率：要能对同一批被打开的窗口，追踪它们中有多少被接上、其中又有多少的伴生连接被拆解、其中又有多少在重组完成后才被判定。只有配对数据才能算条件通过率；三个分开采的比率算不出串联。这是本次修订对附录一提出的唯一新增要求。

三、串联与并联：一个判别办法

怎么知道两处是串联还是并联？本章给一个不依赖数据的判别问法，和一个依赖数据的检验。

问法：把其中一处做到满分，另一处做到零分，那件事还能不能发生？若不能，是串联；若还能（只是差一些），是并联。按这个问法：对窗率为零时，无论悬置拆得多干净、判定推得多晚，那一次改变没有发生过——它甚至没有开始。所以第三编三处是串联。

检验：交互项。在观测数据上回归外部效标，看两两交互项的系数。串联预测正交互显著（两项都高时的效果大于两项各自贡献之和）；并联预测交互项不显著。这个检验很弱——观测数据里的交互项容易被共同原因污染——但它是目前能做的。第五节给可行的设计。

四、第二编三章是什么关系：本章说不清楚

第三编三章是串联，这一点比较硬。第二编三章之间是什么关系，本章诚实地说：不清楚。

三种可能都讲得通：

可能一：也是串联。终点、路径、条件是同一次运行的三个部件，缺一不可——终点没有易主的机会、下一步没有答案、手里没有工具，三者任一为零，这一段就跑不起来。

可能二：补偿型。三者可以互相替代一部分——一个自建条件丰富的班，即使换位率低，也可能靠手里的工具走得下去；一个倒推唯一性很高的技能课，本来就不需要换位（第四章明写这是它的靶格而在成熟技能上不是问题）。

可能三：条件依赖。在多解区里三者串联（多条路都要留痕、都要选择、都要工具），在唯一区里三者近乎独立（只有一条路，工具与终点身份都不影响通过）。

我倾向可能三，而这只是倾向。理由是第五章的档位本身就是一个调节变量：它决定了另两项有没有发挥余地。但这个推理是从机制上讲的，没有任何数据支持它，而第十一章第四节已经把相关预测预注册了——那五条预测里没有一条能判决这件事。

本章因此不把六项写成一个六因子的乘积。能写的只有：

一次完整的改变 $\approx (p_1 \cdot p_2 \cdot p_3) \cdot f(\text{换位、档位、自建})$

括号里是串联， f 的形式未知。把 f 写成乘积是外推，写成加权和是另一种外推，本书两种都不做。

五、政策推论，与它的四条限定

若括号里的乘法成立（这一部分证据最强），得到一条推论：

单修任何一项，收益的上界由另外两项的条件通过率之积封死。

举例：若 $p_1 = 0.2$ 而 p_2 、 p_3 都已经很高，把 p_2 从 0.8 提到 0.95 的全部收益不超过 $0.2 \times 0.15 \times p_3 \approx 0.03$ 量级——而这笔投入在任何一份改革总结里都会显示为“评价环节显著优化”，因为 p_2 本身确实提高了 19%。读数上的进步是真的，通过率上的进步几乎没有。

这条推论与最通行的改革策略——集中资源在一个环节上求突破——正面相撞。因此必须同时写下它的四条限定，否则它会被用成一句口号。

限定一：它不主张三项同等重要。乘法对称，但三项的可改动幅度与成本完全不同。对窗率的改动最便宜（第七章：先花二十分钟看一眼昨天开了什么口子），双向翻转率的采集最便宜，而脱悬率的改动需要改变整个单元的收尾习惯。**对称的是数学，不对称的是工程。**

限定二：它不主张必须三项同时改。次序在第三编小结第二节：先对窗，再拆解，最后才谈判定时点。**同时改与按序改是两回事**，这条推论支持后者，不支持前者。

限定三：它不是“不能只改一项”的道德指令。在 p_1 已经很高的场合，单修 p_2 完全值得——上界由另两项的乘积封死，而那个乘积此时接近 1。**推论的正确用法是先测最低的那一项，不是要求全面推进。**

限定四，最要紧：“全面推进”是同一个错误的另一端。若把这条推论读成“必须六项一起改”，那么它带来的是一份更长的检查表、更多的填报、更完备的账本——而本书第五节已经证明，账本每完备一分，那一段就被压掉一分。**一条关于乘法的推论被用成一张更长的表，是本书能想到的最坏的采用方式。**

六、怎么判决：一个可行的设计

理想设计是 2^3 或 2^6 的析因实验，在学校里做不了：一所学校凑不出八个可比的平行班，更不用说六十四个。本章给三个退而求其次的设计，按可行性排序。

设计一（最可行）：阶梯楔形。一所学校的若干个班，按随机次序依次进入三种处理（先对窗、再拆解、最后改判定时点）。每个班都最终接受全部三项，因此没有对照组的伦理问题；串联假设预测效果的增量依进入次序而不同——先进入对窗的班，后续两项的增量应当明显大于先进入判定时点的班。这一条正是“次序有关”的可检验形式，而相加假设预测增量与次序无关。

设计二：观测数据上的交互项。在第十一章的采集方案上多采若干个班，回归外部效标，检验 $p_1 \times p_2$ 、 $p_1 \times p_3$ 、 $p_2 \times p_3$ 三个交互项。**这个检验的证据强度低**，因为共同原因（一位好教师三项都高）会制造伪交互。必须同时测量并控制教师投入与班级基线。

设计三：极端组比较。只比较两类班：三项都在高档的与恰有一项在低档的。串联假设预测恰有一项在低档的班，其效标接近三项都低的班，而不是接近三项都高的班——即“一项塌陷即全塌”。相加假设预测它落在中间。这个比较不需要干预，只需要在足够多的班里采齐读数之后挑出这两类。**样本要求是三个设计里最高的。**

七、若相加成立，本书失去什么

诚实地写清楚失败的后果，比预设成功更有用。

若实测显示三处的效应是相加的，本书失去三样东西：

其一，失去第三编的次序。先对窗再拆解最后判定时点这条次序，其全部根据就是串联；相加意味着从哪一处入手都一样，三章重新变成三个并列的建议。

其二，失去唯一一条政策推论。本书对教改的全部发言权就在“单修一项收益被封死”这一句上。它一倒，本书退回到一份诊断清单——仍然有用，但不再对任何现行策略构成挑战。

其三，也是最重的：失去“位置”这个说法本身。入口、伴生物、出口这三个词的分量全在串联上；如果它们相加，那它们就不是同一次改变的三个位置，只是三个各自独立的影响因素——而“位置”正是本书区别于“影响因素清单”的地方。

因此这一章是全书赌注最集中的一章。它可以被一个阶梯楔形设计在两个学期内判决，而判决的方向本书事先写死：**串联预测增量依次序而不同，相加预测增量与次序无关。**

八、本章的可推翻条件

第一条。阶梯楔形设计中，三项的增量与进入次序无关 → 串联假设失败，第三编小结第二、三节与本章全部作废。

第二条。恰有一项在低档的班，其效标显著优于三项都低的班且接近三项都高的班 → “一项塌陷即全塌”不成立，乘法退化为一种带阈值的加法。

第三条。若三个交互项在控制教师投入与基线后全部不显著，且样本量足以排除小效应 → 相加假设得到支持。

第四条，针对第二编。若实测发现换位率、倒推档位、自建存量之间存在稳定的**补偿关系**（一项低而另一项高的班，效标不低于两项都中等的班），则第三节的“可能二”成立，第二编三章不是三个独立旋钮，而是同一个东西的三种代偿形态——那将是对第二编小结第三节最直接的一次否定。

第十五章 账本的反噬

判断卡片：本书批评账本只为“可以事前写下的东西”设字段，而本书正在提供一批可以事前写下的字段。这不是一个可以靠谨慎绕开的矛盾，只能靠一条失效期来管住。

一、九章各自的反噬预言

每一章都写过自己被采用之后会出什么问题。合起来是一份相当整齐的清单：

第四章：一旦换位率进入考核，会出现“安排好的换位”——教案上写着某一段由学生的方案来检验上周定的标准，课堂上照本宣科地走一遍，而终点从未真的易主。

第五章：一旦倒推档位进入考核，五位教师会在应答前先对一次口径，档位分布随即变得整齐——而这个读数的全部信息恰恰在分歧里。

第六章：一旦自建存量进入考核，教师做的模板会被登记为班级自造；更糟的一种是刻意撤走本该提供的资源，把学生扔进“荒地”格，以制造缺口。

第七章：一旦对窗率进入考核，“上一节打开了什么”会被写成一句永远能对上的空话（如“提高了学习兴趣”），对窗率随即接近满分。

第八章：一旦脱悬率进入考核，会出现形式化的拆解动作——单元末尾加一句“现在我们把之前说过以后再讲的那个补上”，而没有要求学生把它和当时学的那一处重新接一次。

第九章：一旦双向翻转率进入考核，最简单的对策是不再做四周后的复测——一个不采集的读数永远不会难看。

第一至三章：一旦再生水平、接力率、持位可换度进入考核，会出现“接力表演”：半成品被安排给指定的人接手，版本链被补写，判断场被做成一次课堂展示。第二章已经明写：后续应用者若只制作量表、要求教师打分，而不保存版本链与环境改写，恰好会成为本章批评的对象。

九条预言有一个共同的形状：**造假比改进便宜**。这不是道德问题，是成本问题——只要造假更便宜，它就会发生，无论使用者的动机多好。

二、三种反噬形态

九条可以归成三类，三类的严重程度递增。

其一，表演化。读数上升而所指的事情没有发生。这一类最显眼，也最容易防——第十二章第六节的四条对策（触发源、外来占比、原始计数、可交接性）针对的都是它。表演化的特征是**留不下痕迹**：一次真的换位有一个可指认的触发事件，一件真的自建条件可以交给新来的人直接用，而表演出来的没有。要求成对采集（比率必须同时报分子分母、事件必须同时报触发源），造假的成本立刻高于改进的成本。

其二，挤压。为了抬高某一个读数而牺牲更要紧的东西。最危险的是对窗率：一位教师若被要求把对窗率做到0.8，最直接的办法是打乱课程表去追上一轮打开的每一个口子——而第七章从未主张对窗率越高越好，它明写了“有些窗口的正确处置是让它关上”，只主张**对窗率应当被知道**。同类的还有为提高持位可换度而在安全关键任务中强制轮换角

色，这一条第三章已明确禁止。挤压比表演化难防，因为它不是造假：读数确实上去了，代价落在读数之外。

其三，代理化，最深的一层。读数本身成为字段，而字段一旦存在就会被填满——这是本书自己的论证。于是六个读数变成六栏新账，被填、被比较、被排名、被写进检查表；而它们所指的那些不产生事件的东西，重新退回到字段之外。本书最终成为它所批评的那本账的第七到十二栏。

前两类可以用技术手段压住。第三类不能，因为它不是使用不当，它是使用本身。

三、自反悖论

把第三类推到底，得到一个必须正面说出来的矛盾。

本书的母题是：账本只为可以事前写下的、先在的、外来的东西设字段；而课堂里决定后果的那些，恰恰是这个过程正在产出的，因此在原理上写不进任何一栏。

而本书交出了什么？六个读数、九句问法、一份问卷、一份采集手册、一套编码规程、若干档界。这些全部是可以事前写下的东西。如果它们不可以事前写下，它们就不可能被交给一个不懂本书理论的编码者去采——而“可以被不懂理论的人采集”是本书对自己读数的核心要求（第十二章第四节）。

于是本书处在一个真实的两难里：

要让那些东西被看见，就必须为它们造字段；

而一旦造成字段，被填进去的就只是它们当中可以事前写下的那一部分。

这个两难不能靠“我们的字段设计得更好”来化解。第六章的式子已经给出了它的形式： $c \rightarrow 1$ 时 $A^* \rightarrow 0$ 。本书提供的每一个新字段，都在把它自己要测的那个量往下压一点。这不是修辞上的对仗，是同一个机制作用在本书自己身上。

我不打算把它写成一句谦辞。它是一条可检验的预测：若本书的六个读数被广泛采用为常规记录，那么在采用之后的第二个到第三个年头，这些读数与外部效标之间的关联应当显著下降。关联下降不是测量退化，是被测之物退到字段之外。这条预测本书没有办法自己检验，因为它需要本书先被采用。

四、为什么“提高度量品质”不是解药

到这里可以把导论第八节那条最要紧的分界写完。

指标腐蚀那一支（一个指标一旦成为目标就不再是好的指标）给出的处方，最终都落在提高度量的品质与降低作弊的收益上：多用几个指标、加入不易操纵的客观数据、把评价与奖惩脱钩、随机抽查。这些做法都对，本书第十二章第六节四条对策就属于这一类。

但它们只处理前两种反噬。对第三种，它们无效——因为第三种不是度量被腐蚀，是能成为字段的东西被结算规则挑出来了。

结算规则要求两样东西：一个持有人，一个时点。成绩册的每一栏都是“某人在某时的某项”。而本书第一编那句地基说的是：可再入生成体无人持有，也没有结算时点。只要结算规则不变，任何新字段被填进去的时候都会被迫先补上一个持有人和一个时点——于是“这个班手里多了什么”被填成“张三做了一个模板”，“上一轮打开了什么”被填

成“本单元学生兴趣提升”。字段是被结算规则的形状挤出来的，不是被度量者的品质决定的。

因此本书的处方在结算规则一侧，不在度量品质一侧：不结算个人在某一时刻的所得，改为清点这个班在这一段留下了什么、其中多少可以交给下一届直接用。这个改动很小——它不要求取消任何现有考核，只要求在现有考核之外多留一栏不归任何人名下的账。而它在制度上很难，因为一栏不归任何人名下的账，没有人有动机去填。

这就是本书能走到的最远处：它能说清那栏账为什么必须存在，说不清谁 would 去填它。

五、治理三条

在结算规则改变之前，能做的只有三条，全部是防守性的。

其一，读数归采集者持有，不归机构。六个读数的第一用途是一位教师自己看自己这一段哪里最低。制度上只核查是否采集与是否保留原始记录，不核查数值高低，也不跨班比较。这一条一旦被突破，后面两条都失效。

其二，全部读数成对采集。比率必须同时报分子与分母的原始计数；事件必须同时报触发源；物件必须同时报外来占比与可交接性。第十二章第六节已给出四条具体形式。其作用不是防止造假，是使造假的成本高于改进的成本——这是能给出的最强技术保证，也仅止于此。

其三，任何不利安排必须开放复核。编码规则、原始证据、复核通道三样缺一不可。这一条来自第九章：一次判定的危险不在它可能错，而在错了也不会被发现。

六、失效期

三条治理都挡不住代理化。对它，本书只能给一条听起来不像办法的办法：

任一读数一旦进入任何形式的考核，两年后自动作废，须重新设计。

理由是本书自己的论证：字段会被填满，而被填满的字段与它原本要测的东西之间的距离随时间单调增大。与其假装能设计出一个不被腐蚀的字段，不如承认字段有寿命，并把寿命写进制度。

这条设计有三个明显的代价，写在这里以免它被当成聪明话：其一，读数每两年一换，纵向比较就没有了，而纵向比较正是评估最想要的东西；其二，重新设计需要人，而愿意做这件事的人比愿意填表的人少得多；其三，它无法阻止一个机构在两年内把该做的坏事做完。

这三条代价我抵不掉。本章因此不是一份解决方案，是一份定价单——它把采用本书要付的代价列清楚，让人在采用之前就看见。

七、如果本书成功

最后一层，必须由作者自己说。

如果本书成功，它就会被采用；而被采用就会产生它所描述的病。这不是谦虚，是从本书自己的机制里推出来的结论：一个提供了六个可采集读数的框架，其最可能的命运是那六个读数被抽出来做成量表，而九章的论证——包括本章——被跳过。第二章已经预先

写下了这件事：若后续应用者只制作量表、要求教师打分，却不保存版本链与环境改写，恰好会成为本章批评的对象。

对此，本书能做的只有三件，都已经做了：把反噬写进正文而不是附录；把失效期写进制度建议；把“它会被怎样填坏”写在给出问卷的同一章里。

还有第四件，只有读者能做：当你看见这六个读数出现在一张检查表上，而那张表没有要求同时报分母、触发源与可交接性时，请指出那不是本书的东西。

八、全书的可推翻条件

各章的证伪条件在附录三汇总。此处只列判定本书整体成败的五条。

第一条（最可能被触发）。六个读数在同一门真实课上采齐后，其中三个以上两两相关系数持续高于 0.6 → 九个对象是虚高的，本书应按第十一章第六节的规则合并，并公开承认这是一次过度切分。

第二条。自建条件存量与教学资源完备度实测为正相关 → 第六章的式子方向错误，导论第七节、第二编小结第四节与第十三章第五节全部作废。这是全书最集中的一次单点检验。

第三条。阶梯楔形设计中三项增量与进入次序无关 → 相乘假设失败，第三编小结与第十四章作废，本书从一个结构性判断降格为一份诊断清单。

第四条。两名只拿手册的编码者一致性持续过低 → 九句判据“不含判断词”的承诺不成立，第十二章作废，而全书“可以被不懂理论的人采集”这条自我要求随之落空。

第五条，只有外人能做。若本书的异源汇聚（第一编）被证明共用了同一个隐蔽前提，或本书的框架被用来回应一条关于读数的技术意见——“你没有理解本书的框架”——而无人指出，那么本书正在演示它自己批评的那件事。**第一重锁的定义特征就是被锁住的人察觉不到它。**

结 语 课堂用一本发现的账本，记一件发生的事

一节课结束了。教室里留下了什么？

按现有的账本，这个问题有一个整齐的答案：留下了若干个内容点的掌握情况，记在若干个学生名下，附一个日期。这本账记得很认真，每一栏都是真的，而且每一栏都可以被核查。它唯一的问题是：**它的每一栏都要求一个持有人和一个时点，而这一节课里最要紧的那些东西，两样都没有。**

那些东西是这样的。

它们是被造出来的，不是被带进来的。这个班在这一学期里造出了一套自己的记号、一份自己的错误清单、一个“什么算讲清楚了”的默契。终点在某一段里易过主——原本用来达成目标的那样东西，本身成了要达成的东西，然后又被交还，而原来的终点因此被改写。“下一步该做什么”这个问题，在某些单元里有一个答案，在另一些单元里有五个，在还有一些单元里一个也没有；而这不是学生的水平问题，是知识在这群人当中的分布问题。

它们没有持有人。一件班级自己造的工具，属于谁？造它的那个学生已经不用它了，用得最多的是另外三个人，而它真正的价值在于可以交给一个新来的人直接用。一次学习如果只留在一个人手里、没有被交出去、没有被别人改写过，它就还没有完成——它会在那个人身上凝固成一件私人技能，而不是这个班的一件东西。

它们没有结算时点。它们平时不显露。你在任何一个时点去查，都可能查到零；而在支架被撤掉、情境被换掉、教师不在场的那一刻，它们要么重新组织出来，要么没有。休眠不是死亡，可传统测验只查旧答案能否被直接提取，因此系统性地把休眠读成遗忘，把还在动的读成不动。

于是三件事同时成立：这些东西决定后果；它们是这个过程自己的产物；账本上没有它们的位置。第三件不是前两件的疏漏，是前两件的直接后果——**账本的字段由结算规则派生，而结算规则要求一个持有人和一个时点。**

这就是本书全部的内容。它可以压成一句话：

课堂用一本发现的账本，记一件发生的事。

发现的账本为先在的、外来的、等着被读出来的东西设字段：内容点、课时、目标、达标率——它们在写表格的时候就已经存在，因此写得进去。而课堂里正在发生的那件事，在写表格的时候还不存在。它不是被漏记的，是不可能被记的。

更难受的一层在于：这不是一个可以靠把账记全来解决的问题。账本每完备一分，那一段就被压掉一分——这句话的确切分量由第六章那个结构模型给出，条件是其余参数不随覆盖度同时变化；本书不主张它是一条无条件的经验定律。课程建设每提高一分覆盖度，自建条件按同一比例少一分；教师在一门课上越熟练，越容易把多解讲成唯一，而这件事在这门课里永远不会被发现；评价做得越及时、反馈越频密，判定就越可能落在重组还没走完的时候。三处都不是失职的产物，是尽职的产物。**这是本书最不愿意但必须写下的一句：那些把课上好的努力，与那些把这一段压掉的力量，是同一股力量。**

那么该怎么办？

本书没有一份处方，只有一个位置。它给了九个字、六个读数、九句可以拿去查记录的问法，和大约六十条可以让它自己作废的条件。它主张的动作很小：**在按表推进之前，先花二十分钟看一眼昨天在他们身上开了一个什么口子；在按判定安排一个人的后半年之前，先问一句那次判定是在他停下来之后作的，还是在他还动着的时候作的；在这一门课结束时，先清点一下这个班手里多了什么，是下一届可以直接拿来用的。**

三个动作都不需要新的评价工具，都不减少任何现有资源，也都不能被写进任何一张现成的表——因为它们要看的東西，正是那张表没有设字段的那几栏。

最后一句留给“智慧”这个词，本书正文里一次也没有用过它。

如果说一间教室里有什么可以被叫做智慧的东西，那它不在讲得多清楚、排得多合理、评得多精准里。**它在一件更朴素的事情上：这间教室有没有为自己的产物设一栏账。**为自己造的东西记账，任何一本合格的账簿都做不到；而为自己正在造出来、还没有名字、还没有持有人、也还没到结算时候的那些东西留一个位置——这件事没有任何制度会要求你做，做了也没有人会看见，而一节课里真正留下的东西，全在那一栏里。

附录一 六读数采集手册

这份手册可以脱离本书使用。采集者不需要读过本书，也不需要同意本书的任何判断；他需要的只是这几页。手册刻意不解释“为什么”——若一个读数只有认同本书的人才采得出来，那个读数不合格。

全部六项在一门课、一个学期内可以采齐。总成本约：编码者 60–80 人时，授课教师额外约 10 人时，另需五位教师各二十分钟。

通则

一，**一律成对记录**。任何比率必须同时写下分子与分母的原始计数；任何事件必须同时写下触发者；任何物件必须同时写下它是谁造的、能不能交给一个新来的人直接用。只报比率不报原始计数的记录，视为无效记录。

二，采集者不应是授课教师本人。教师是被问的人。

三，**四项必须在事情正在发生时采**（换位时刻、悬置语句、上一节打开了什么、终点易主的触发源），事后回忆不能替代。

四，**不做总分**。六项分别报，读的时候只看最低的那一格在哪里。

五，**原始记录保留到复核期结束**。

读数一 倒推唯一性档位（最便宜，一次教研活动）

采什么。对一个具体单元，把一份真实的学生作业（不是虚构案例）与该单元的目标，交给五位有经验的教师，各自独立写下一句“这个学生下一步该做什么”。**不许事先讨论，不许互相看**。

怎么判档。收齐五份答案后：

唯一档：五份指向同一件事（措辞可不同，动作是同一个）。

多解档：出现两种以上不同的下一步，且五人互评时认为彼此的答案都可行。

无解档：出现两种以上不同答案，且五人互评时认为彼此的答案都不是下一步；或有人直接写“说不出来”。

互评这一步不能省。多解与无解在答案分布上看起来一样，只有互评能把它们分开。

常见错采。只问一位教师（结果必然偏向唯一）；问同一教研室的五位（结果系统性偏向唯一）；用虚构的学生样本（专家会按理想情形作答）。

成本。一次教研活动，每位教师约二十分钟。可一次测一门课的六个单元。

读数二 双向翻转率（最便宜，四周间隔）

采什么。取这门课上一次作出“会了没有”判定所用的那份卷子，隔四周原样再做一次，期间不就该内容教学。

记什么。逐人比对两次结果，分两个方向计数：

“会 → 不会”的人数与占比；

“不会 → 会”的人数与占比（这一支是关键）。

两支分别报，不合并成“一致率”。

档界（参考）。“不会 → 会”这一支：低于 5% 为低档，5%–15% 为中档，高于 15% 为高档。高档意味着上一次判定作在重组还在进行的时候。

常见错采。换一份“平行卷”（题目变了，翻转就无法归因）；期间补课或讲评（等于施加了干预）；只报总一致率（两个方向被合并，全部信息消失）。

成本。一份现成卷子，一次课时，四周等待。

读数三 悬置密度与脱悬率（一个单元）

采什么。课堂上明确说出“暂时不考虑”“这个以后再讲”“先不管”“到后面再说”一类语句的次数与位置。

怎么记。每一条记四项：说的是哪一句、挡住的是什么内容、发生在第几分钟、当时正在讲什么。

悬置密度 = 每课时的悬置语句数。

脱悬率 = 分子 ÷ 分母：

分母：本学期明确悬置过的处数；

分子：其中后来真的讲了，并且要求学生把它和当初学的那一处重新接了一次的处数。

注意分子的判定很严。后来单独讲了那一部分，但没有要求学生回到当初那件事上重做一遍——不计入分子。

常见错采。把“这个不考”记成悬置（那是范围声明，不是悬置）；把教材上的“选学内容”记成悬置（那是课程设计，不是课堂动作）；分子放宽为“后来讲了”。

成本。课堂录音或一名旁听者计数，每周约一小时。

读数四 换位率与触发源（一个单元）

采什么。课堂上“原本用来达成目标的那样东西，本身成了要达成的东西”的时刻。

怎么判定一次换位成立。四条同时满足：

1. **角色真换**——那一段里，班上实际在检验、改进、争论的对象是方法本身，不是原来的目标；
2. **有开合时刻**——说得它从第几分钟开始、到第几分钟结束；
3. **有触发源**——由一个可指认的事件触发（某个学生的做法、一次异常结果、一件半成品）；**教师宣布不算**；
4. **有交还**——那一段结束时回到了原来的终点，并且说得原终点因这一段被改了什么。**没有交还的不计为换位。**

记什么。每次记：开始与结束时刻、触发者是谁、被提为终点的是哪一样、交还时原终点被改了什么。

换位率 = 每课时的成立次数。**触发源分布** = 触发者中学生／异常／半成品／教师各占多少。

常见错采。教案上写着的“方法讨论环节”直接计一次（那是安排，不是发生）；把教师的一句“我们来看看这个方法本身”计为触发源。

成本。与读数三共用同一份课堂记录。

读数五 对窗率（两轮内容，最费力）

采什么。上一轮教学在学生身上打开了什么，以及本轮内容有没有落在那里。

怎么记。分两步，**必须分两次做，不能一次补记**：

1. 上一单元结束当天，由授课教师用五分钟写下一到三条“这一单元在他们身上打开了什么”。要求写**具体的问题或困惑**（例：“好几个人追问为什么负数不能开方”），**不得写状态描述**（“兴趣提高了”“参与度不错”一律作废重写）。
2. 本单元结束时，由编码者判断：上述条目中有几条被本单元的内容实际接上了。“接上”的判定标准是——本单元有一段内容以那条问题或困惑为起点，而不是恰好涉及同一知识点。

对窗率 = 被接上的条数 ÷ 上一轮登记的条数。

档界（参考）。低于 0.2 为低档，0.2-0.5 为中档，高于 0.5 为高档。**注意：本书不主张越高越好**——有些窗口的正确处置是让它关上。这个读数的用途是被知道，不是被最大化。

常见错采。第一步写成空话（这是最常见的失效，且会让对窗率虚高到接近满分）；两步由同一人在同一天补记；把教案里的“复习上节”直接计为对窗。

成本。教师每单元五分钟，编码者每单元约半小时。

读数六 自建存量、外来占比与可交接性（一学期）

采什么。这门课结束时，这个班手里有多少件自己造出来的、可以直接拿来用的物。

怎么记。逐件登记五项：

1. 是什么（一份错误清单、一套记号、一个模板、一张判据表、一段脚本……）；
2. 谁造的（这个班／某几位学生／教师／外来）；
3. 被用过几次，以及第二次被使用是什么时候；
4. 能不能交给一个新来的人直接用（把它交给一个没上过这门课的人，他能不能在不问原作者的情况下用起来）；
5. 它顶替了什么（原本要靠谁或靠哪份材料才能完成的事）。

自建存量 = 第二项为“这个班／学生”且第四项为“能”的件数。

外来占比 = 教师或教材提供的件数 ÷ 总件数。

可交接性 = 第四项为“能”的比例。

判定的两道闸。其一，**必须是可指着看的物，不是知识**——“学生掌握了一套分析方法”不计，“学生整理并在此后反复使用的那张两页纸的分析清单”计。其二，**必须由这个过程自己造**——教师做好的模板不计，哪怕全班都在用。

常见错采。把教师提供的模板登记为班级自造（这是最常见的一种，且会使这个读数完全失效）；只清点数量而不判可交接性；在学期中途清点（存量需要时间长出来）。

成本。学期末一次，约三小时；若课程使用版本控制或共享文档，其历史记录里可直接筛出“由学生首次提交、随后被三人以上引用”的条目。

一学期的采集日历

周次	动作	人时
1	建立记录：单元表、物件清单（初始通常为空）	2
2-3	读数一：五位教师，一次教研活动	$5 \times 0.3 + 1$
3-6	读数三、四：课堂记录	每周 1
6	基线卷子（读数二的第一次）	1
7-10	读数五：每单元教师五分钟+编码者半小时	每单元 0.6
10	复测卷子（读数二的第二次）	1
11-15	继续读数三、四；择一次课作撤架任务（外部效标）	每周 1
16-18	读数六：清点物件	3

采齐之后请做的第一件事：报出六项两两之间的相关系数，包括对本书不利的那些。本书在第十一章第四节已把自己的预测预先写死，请原样比对。

附录二 九章的学科入口、取用点与最危险的近邻

章	三个学科入口	各自的取用点	推翻共有前提的材料	命名	最危险的近邻与分离线
一	教育学 × 工程学 × 艺术学	策略迁移与延迟显露／目标标准过程资源的连续重组／共同判断场	三家自己的材料：事件本身可以不再发生	学习事件再生	知识建构、未来学习准备——分离线在 最低项取法 与撤架后的第三方复核
二	认知科学 × 经济学 × 艺术学	表现背后的能力变化／稀缺认知资源的积累与配置／制作中的思考	三家共同默认“学习有主人、有结算时点”	接力再生事件	学习迁移、生成性学习、ICAP——分离线在 发起权是否交出并被改写
三	物理学 × 心理学 × 艺术学	响应地形／预测—提取—迁移路径／身体与媒介场	学习不必先有一个持有人	可再生生成体	分布式认知、情境学习——分离线在 不显露时它在不在 （休眠≠死亡）
四	教育学 × 化学 × 艺术学	目标先行与掌握学习／逆合成分析与条件筛选并存／工作室讲评对做法的训练	化学自己：信息在哪一端多就从哪一端开始	换位区间	目标置换 （剥到形式层几乎同形）——分离线只靠 交还 ：置换是永久吞并，换位是临时取得资格后交还且原终点被改写
五	物理学 × 心理学 × 工程学	变分原理与两端整体决定／刻意练习／稳健参数设计	物理自己：初值问题解存在唯一，两点边值问题可无解可有无穷多解	可倒推区间	可控性 ——分离线：可控性是系统性质，倒推唯一性是 知识分布性质 （换一批教师档位会变）

章	三个学科入口	各自的取用点	推翻共有前提的材料	命名	最危险的近邻与分离线
六	细胞学 × 工程学 × 艺术学	生态位／过程能力（质量是造出来的不是检出来的）／判断面对作品本身	细胞学自己：生态位有相当一部分是住户造出来的	自建条件	生态位构建 （那家承重在跨代，本章在同一批人同一学期）、 分布式认知 （那家讲既有工具，本章讲工具是过程自造）
七	细胞学 × 工程学 × 艺术学	序贯诱导与感受期／装配序列与事后配对／材料条件对呈现的约束	细胞学自己：感受期是上一轮信号的产物	空窗	关键期 （分离线已写足四条：由上一轮内容决定而非年龄、可被课堂制造、同一学生可多次开合、窗内容与内容匹配而非泛化能力）
八	教育学 × 化学 × 艺术学	概念转变的不相容／保护基策略与酰基迁移／作品的未完成状态	化学自己：酰基迁移改写了它本要固定的名单	悬置连接	信息隐藏与隐性耦合 （本章自承增量在承载物不在形式）、 干扰理论
九	物理学 × 心理学 × 工程学	玻璃态的物理老化／提取诱发遗忘／控制与传感的时序	物理自己：被判冻住者随等待时间持续演化	冻结构型	测量时机与信度、标签效应 ——分离线在读数为零不等于系统为零

读法。最后一列里反复出现的那几位，才是本书真正的对手：目标置换（第四章）、关键期（第七章）、分布式认知（第三、六章两次）。这三处是本书最可能被吸收回既有说法的地方；若要攻击本书，从这三处入手最省力。

附录三 证伪条件与写死日期的赌注

一、五条判定本书整体成败的条件

1. 六个读数中三个以上两两相关系数持续高于 0.6 → 九个对象是虚高的，本书是一次过度切分。（第十一章第六节给出合并规则）
2. 自建条件存量与教学资源完备度实测为正相关 → 第六章的式子方向错误，导论第七节、第二编小结第四节、第十三章第五节全部作废。
3. 阶梯楔形设计中三项增量与进入次序无关 → 相乘假设失败，第三编小结与第十四章作废。
4. 两名只拿附录一的编码者一致性持续过低 → 九句判据“不含判断词”的承诺不成立，第十二章作废。
5. 异源汇聚被证明共用同一隐蔽前提 → 第一编最强的支持作废。（只有外人能做）

二、逐章的核心证伪条件

第一章。若标准迁移、自我调节与适应性专长完全预测所谓事件再生，新概念应撤回；若第三方无法可靠识别问题、修订与判断三者的关系，测量协议应放弃。

第二章。若撤除支架后的“接力”与个人延迟迁移测验高度重合，接力不构成独立对象。

第三章。若四项构成条件全部成立的班，其撤架后独立表现与四项全不成立的班无差异，该对象即使存在也不承重。

第四章。在同一门课程、同一批教师、控制教案偏离次数、时间投入与学生先验知识之后，若换位率与两周以上的变式迁移之间不存在剂量—反应关系，本章为伪；届时应把差别归于教学灵活性。另有七条，其中最便宜的一条是：软件课提交记录里“测试通过之后发生的重写次数”与该模块后续缺陷密度的关系——这份数据在多数代码托管里已经存在。

第五章。在控制内容难度与教师经验之后，若“倒推唯一性 × 运行方式”的交互效应不显著，分段主张为伪。另：若同一份学生样本交给两个不同专家群体得到相同档位，则倒推唯一性是系统性质而非知识分布性质，本章被可控性覆盖。

第六章。在控制班级氛围、参与度、时间投入与先验知识之后，若自建条件存量与两周以上的新情境表现之间不存在剂量—反应关系，本章为伪。**最要紧的一条顺序检验：**一件自建条件被第二次使用，应当早于该班相关表现的提升；若表现提升总是先于第二次使用，则因果方向反了。

第七章。在固定本轮内容与时间间隔、只改变上一轮内容的条件下，若本轮效果没有出现可由上一轮所开窗口预先预测的差异，核心判断失去支持。

第八章。若悬置期内并不形成额外连接，或形成的连接与被悬置项的位置无关，第一层即不成立；若在悬置密度匹配的两组之间，明确拆解与不拆解的班在换情境表现上无差异，本章对象不承重。

第九章。若“不会 → 会”这一支在各类课堂上都接近零，“还在动”就不是普遍现象，本章应降格为描述性提法。

三、写死日期的赌注

章	期限	赌什么	什么不算命中
一	2029-08-03	至少一项预注册、多点、跨两种以上学科、不少于六个班的研究，发现原课堂成绩与事件再生稳定分离，且再生读数在控制原成绩、标准迁移与自我调节后仍能预测四周后的自发学习	教师自评、学生自报、单一班级、教师仍提供问题与量表、无延迟结果
二	2031-12-31	接力读数在控制个人能力后仍预测延迟表现	只有当堂产出、只有自报
三	2030-12-31	撤架—换境设计下再入表现与原课堂表现分离	只换题面不撤支架
四	2030-12-31	至少两项在同一制度环境内、平行班不少于十二个的预注册研究，报告换位率在控制教案偏离次数、时间投入与先验知识后仍正向预测两周以上的变式迁移	未控制教案偏离次数；把换位与增加课时打包
五	—	本章未押日期，改押一次教研活动即可完成的检验（五位教师、五份匿名样本、各十分钟）	—
六	—	本章未押日期，改押版本控制课程的仓库历史：由学生首次提交、随后被三人以上引用的文件	—
七	2031-12-31	至少三项在同一制度环境内、不少于八个平行班的研究，报告只改变上一轮内容时本轮效果出现可预测的差异	仅改时间间隔；仅改难度；仅比较不同教材或教师；仅问卷自报兴趣；把窗口对齐与增加课时打包
八	2031-12-31	明确拆解组在换情境表现上优于不拆解组，且差异不由复习总量解释	把拆解与增加复习打包
九	2031-06-30	双向翻转率中“不会到会”一支在多解档单元里显著更高	期间补课；换平行卷；只报总一致率

附录四 十五张判断卡片

每张一张，是判断不是描述——一句可以被反对的话。九章来自独立成文的原稿，第十至十五章为本书新写。

第一章。课堂留下的不是内容，是这个班重新发起一次学习的能力；而它的水平由问题、修订链、判断场三项中**最低的一项**决定，不由平均决定。

第二章。学习不在任何一个人身上结清——它必须被交到下一个人手里、被那个人改写并继续用下去，才算存在。

第三章。课堂生成的那样东西不持续显露；**不显露不等于不存在**，而现有测验只查旧答案能否被直接提取，因此系统性地把休眠读成遗忘。

第四章。终点与手段的角色不由事情的性质决定，由信息的分布决定，因而可以互换；而换位若不交还，**就不是换位，是目标被吞掉**。

第五章。“下一步该做什么”不是一个总有答案的问题——它可以有一个、多个或没有；而在长期迁移与可选择性权重较高的内容上，**档位不确定时应当默认推进**，因为错配的代价在两个方向上不对称。

第六章。条件有一部分是这个过程自己造的；而在高覆盖度区间，外来条件每完备一分，这一部分就被压掉一分—— **$c \rightarrow 1$ 时，一门完备的课程在结构上不可能有它**。

第七章。准备度不是学生的属性，是上一轮的产物，而且有时限；**窗关上时什么也不发生，因此这件事在原理上不会以失败的形式被发现**。

第八章。为使一处改变而挡住的那部分不会停在原地——它在悬置期内结成连接，脱悬后无人拆除，也不进任何账本，却在换情境时决定表现。

第九章。一次判定若作在重组还没走完的时候，读到的零会被当作结论，并据此配置此后的条件——**于是它自我实现**。

第十章。决定一次跨学科碰撞撞出什么的，不是学科的组合，是问题的形状；学科是材料，问题是模具。

第十一章。九个读数里，只有三个便宜到可以每学期采一次；把它们采齐一次，比再写三章更有价值。

第十二章。能合的是入口，不能合的是输出——九句问法可以合成一份问卷，九个读数不能合成一个总分。

第十三章。三条推导目前收不成一个式子；但它们的退化条件是同一句话——**某一端的不确定性归零时，那一段就不存在**。

第十四章。串联的是**条件通过率的连乘**，不是三个边际读数相乘，更不是可加的贡献量；单修任何一项，收益的上界由另外两项的条件通过率之积封死。

第十五章。本书正在提供一批可以事前写下的字段，而“可以事前写下”正是本书判定为病根的那个性质；这个矛盾无法靠谨慎绕开，只能靠一条失效期管住。

附录五 跨域转移：九章逐条分级

导论第四节给出了跨域转移的五步与三级措辞。此表把九章的第四步逐条排开，第三列不可映射部分是本表存在的主要理由——它写的是每一次转移明确不能带过来的东西。

读法：每一行的第四列都只是一个竞争假设，不是已被证实的课堂事实；第五列写的是它需要什么才能升到三级。截至本版，九行全部停在二级。

表上：源领域事实、结构映射与不可映射部分

章	源领域事实（一级）	结构映射	不可映射部分
一	工程教育中目标、标准、过程、资源被连续重组	重组的对象 ↔ 一次学习的组织要素	工程项目有外部甲方，课堂没有
二	稀缺认知资源在制度中被配置与积累	资源配置 ↔ 半成品在人与人之间的接手	经济学的行为人有明确偏好，学习者没有
三	物理系统的响应地形被扰动改变并留痕	响应地形 ↔ 一个班的反应倾向	物理系统无意向、无社会关系
四	化学中逆合成分析与条件筛选并存，选用依信息分布	起手端的选择 ↔ 终点与手段的角色	化学的产物可被独立检定，教学目标不能
五	初值问题与两点边值问题的适定性不同	适定性 ↔ “下一步”的解的个数	微分方程的解空间有度量，教学路径没有
六	细胞参与塑造自己所处的生态位	住户造条件 ↔ 班级造工具	细胞无意图、无交接行为；生态位重塑跨代
七	细胞的感受期由更早一轮信号打开且有时限	感受期 ↔ 上一轮打开的口子	发育的窗口由基因程序限定，课堂的不是
八	保护基会迁移，改写它本要固定的名单	保护基 ↔ 为讲清一处而作的悬置	化学的迁移有确定反应机理，课堂的没有
九	玻璃态被判冻结后仍随等待时间演化	物理老化 ↔ 判定后仍在进行的重组	材料的演化可连续测量，学生的不能

表下：由映射生成的竞争假设，以及升到三级需要什么

章	课堂竞争假设（二级）	升到三级需要什么
一	课堂留下的是重新发起一次学习的能力	撤架任务上的第三方复核，且再生水平在控制原成绩后仍预测四周后自发学习
二	学习必须被交出去并被改写才继续存在	接率率在控制个人能力后仍预测延迟表现
三	课堂生成一件可再入、不持续显露的生成体	撤架一换境设计下再入表现与原课堂表现分离
四	终点可临时易主并交还，构成换位区间	十二个以上平行班上，换位率在控制教案偏离后预测变式迁移

章	课堂竞争假设（二级）	升到三级需要什么
五	倒推唯一性分三档，错配代价双向不对称	“倒推唯一性 × 运行方式”的交互效应显著
六	条件有一部分由这个过程自造，且与覆盖度存在结构性张力	自建存量与新情境表现的剂量—反应关系，及第二次使用先于表现提升
七	存在无人到达的空窗，且不产生失败信号	固定本轮、只改上一轮，效果差异可由所开窗口预测
八	悬置期结成的连接脱悬后无人拆除	悬置密度匹配下，拆解组在换情境表现上占优
九	判定若作于重组完成之前，读数被冻成结论并自我实现	多解档单元里“不会到会”一支显著更高

这张表最该被反对的地方：第三列写的“不可映射部分”是作者自己指认的，而一次转移真正致命的不可映射之处，往往正是转移者看不见的那一处。若有读者能为任意一行补上第四条不可映射机制，并说明它使该行的竞争假设失效，本书应当就那一行退回一级。

附录六 先导研究预注册草案

本书没有任何新采集的数据。这份草案是把附录一那份采集手册压成一个**最小可执行**的先导研究，供任何愿意做第一轮的人直接照抄。它是设计，不是结果；本书不预告它会得出什么。

一、目的与不追求的东西

目的：在同一门真实课程上，第一次把六个课程级主读数采齐，报告它们的分布、编码者一致性与相关矩阵，并与一个外部效标关联。

明确不追求：不追求正结果；不追求效应量；不追求推广到其他学校。**本研究的成功标准是“采得出来、报得完整”，不是“支持了本书”。**

二、设计

单校、单学科、两个平行班、一个学期（16–18 周）、一位授课教师、两名不读本书正文的编码者。观察性设计，**不设干预**——先导阶段只做测量，不做处方。

三、须预先登记的内容（在采集开始前公开）

1. 六个主读数的操作定义与档界（照抄附录一，不得中途修改）；
2. 编码规程与两名编码者的分工，以及一致性检验方式（计数类用 ICC，等级类用加权 κ ）；
3. 外部效标：本草案取四周后撤架任务上的独立表现；
4. 第十一章第五节那五条相关预测，原样登记；
5. 停止规则：若某读数的编码一致性低于可用水平，该读数只报描述，不进相关矩阵。

四、时间表

照附录一那张一学期采集日历执行。第 1 周建记录，第 2–3 周采倒推档位，第 3–6 周与第 11–15 周采悬置与换位，第 6 周与第 10 周两次卷子采双向翻转率，第 7–10 周采对窗率，第 16–18 周清点自建存量。

本次修订新增一条：第十四章的附注要求**按个案配对记录**——同一批被打开的窗口，须能追踪其中哪些被接上、其伴生连接哪些被拆解、哪些在重组完成后才被判定。三个分开采的比率算不出条件通过率。

五、报告清单（无论结果如何，逐项报出）

每个读数的**原始分子与分母**，不只报比率；

每个事件类读数的**触发源分布**；

编码者一致性系数与**分歧的分布**（分歧本身是数据，不得由第三人裁掉）；

六读数的**相关矩阵全表**，含与第十一章五条预注册预测的逐条比对；

与外部效标的关联，含**不显著的那些**；

采集实际耗时与偏离方案之处。

六、样本量与它的诚实交代

两个班的先导研究不足以做构念区分，也不足以估计相关系数的精度：以两班约六十名学生计，相关系数的置信区间宽到几乎不能排除任何东西。因此本先导研究不能判决第十一章第六节那四条保留规则，只能判断“这套读数是否采得出来”。

真正的构念区分需要：第二阶段六个以上班级、两个学科；第三阶段三十个以上班级，方可做多特质—多方法分析与因素分析。本书把这三阶段写清楚，是为了让第一阶段的结论不被越级使用。

七、伦理

照第十二章第五节那三条：读数指向课堂位置与关系而非个人；不得为提高读数制造条件；任何据此作出的不利安排必须公开编码规则、原始证据与复核通道。先导阶段的读数一律不得进入任何考核。

附录七 参考文献

本表由九章各自的文献表合并去重而成，按著者姓氏字母序排列，去重后共 90 条。

关于本表的性质与可靠性，须如实交代四层。

第一层，它是什么。这是一份来源清单，不是引证索引。正文引用一律只到概念层——引的是某一支的主张，不是某一页的原话——因此表中多数条目并不以著者姓名出现在正文里；经机器对账，正文点名出现的著者约三十余位，其余为概念层来源。用本表去反查“某句话出自第几页”是查不到的，这是体例决定的，不是遗漏。

第二层，抽查过的。近年文献（二〇二四至二〇二六年，共 30 条）是最易出错的一批，从中抽 6 条逐项复核，著者、刊名、卷期、页码与数字对象标识符全部吻合：

1. Xu, Zhang 与 Zhang 关于师生脑间耦合的纵向研究 (*npj Science of Learning*, 二〇二六; 另有二〇二四年起的预印本) ;
2. Brown, Kaur, Kingdon 与 Schofield 关于认知耐力作为人力资本 (*The Quarterly Journal of Economics* 140 卷 2 期 943–1002 页) ;
3. Stern, Liu 与 Balasubramanian 关于学习在物理系统上的结构印记 (*Physical Review E* 109 卷 024311) ;
4. Cleary Rodrigues 关于美术工作室讲评 (*Arts and Humanities in Higher Education* 24 卷 3 期 278–291 页) ;
5. Hetherington 关于物质—对话空间 (*Research in Science & Technological Education*, 二〇二四, DOI 10.1080/02635143.2024.2328766) ;
6. Tufail, Park, Wang 与 Kim 关于认知负荷与约束驱动创新 (*International Journal of Design Creativity and Innovation* 14 卷 2 期 122–149 页) 。

第三层，没有抽到的。其余 84 条未逐条回到原始出处核对。本表的著录信息来自各章写作时所据的文献表，合并时只做了去重与体例统一（连接号一律 en dash、数字对象标识符一律写成链接式、刊名与卷号排斜体），没有重新核库。二〇二五至二〇二六年的条目中，可能有仍处于在线优先或预印本状态而本表按正式出版著录的情况。

第四层，这意味着什么，以及谁该做下一步。由于引用只到概念层，一条著录若有讹误，损失的是可追溯性，不是正文判断的成立与否。但这不构成免责：逐条核验是正式出版前必须完成的一道工序，本书作者的诚实披露不能代替它。在此之前，请在引用本表中的任何一条之前回到原始出处自行核对；若发现讹误，那是我的疏漏，请指出来。

1. Abeler, J., Huffman, D., & Raymond, C. (2025). Incentive complexity, bounded rationality, and effort provision. *American Economic Review*, 115(12), 4404–4437. <https://doi.org/10.1257/aer.20230751>
2. Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
3. Barrett, E. (2026). The sculpture kit: Material pedagogy and entangled making. *Studies in Art Education*, 67(1), 59–79. <https://doi.org/10.1080/00393541.2025.2591489>
4. Bates, G. (2025). “What’s in a name?” Retrieval practice in the primary classroom: A case study. *Education* 3–13. <https://doi.org/10.1080/03004279.2025.2504492>
5. Becker, G. S. (1964). *Human Capital*. University of Chicago Press.

6. Benneyan, J. C., Lloyd, R. C., & Plsek, P. E. (2003). Statistical process control as a tool for research and healthcare improvement. *Quality and Safety in Health Care*, 12(6), 458–464. <https://doi.org/10.1136/qhc.12.6.458>
7. Bereiter, C., & Scardamalia, M. (1993). *Surpassing Ourselves: An Inquiry into the Nature and Implications of Expertise*. Open Court.
8. Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From Learning Processes to Cognitive Processes*.
9. Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61–100.
10. Brown, C., Kaur, S., Kingdon, G., & Schofield, H. (2025). Cognitive endurance as human capital. *The Quarterly Journal of Economics*, 140(2), 943–1002. <https://doi.org/10.1093/qje/qjae043>
11. Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
12. Bursztyn, L., & Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics*, 130(3), 1329–1367.
13. Cappello, N., & Ortiz, D. C. (2025). Creative embodied learning about different identities. *Thinking Skills and Creativity*, 57, 101833.
14. Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243.
15. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
16. Cleary Rodrigues, M. C. (2025). The crit: Making meaning with peers in fine art studio pedagogy. *Arts and Humanities in Higher Education*, 24(3), 278–291. <https://doi.org/10.1177/14740222241311857>
17. Coëgnarts, M. (2025). Art and embodied knowledge: The bodily basis of aesthetic cognitivism. *The Journal of Aesthetics and Art Criticism*, 83(3), 197–210. <https://doi.org/10.1093/jaac/kpaf025>
18. Corey, E. J. (1991). The logic of chemical synthesis. *Angewandte Chemie International Edition*, 30(5), 455–465. <https://doi.org/10.1002/anie.199104553>
19. Cullen, S., & Oppenheimer, D. M. (2024). Choosing to learn: The importance of student autonomy in higher education. *Science Advances*, 10(29), eado6759. <https://doi.org/10.1126/sciadv.ado6759>
20. Cunningham, C. M., & Kelly, G. J. (2017). Epistemic practices of engineering for education. *Science Education*, 101(3), 486–505.
21. Dannels, D., & Martin, K. (2008). Critiquing critiques: A genre analysis of feedback across novice to expert design studios. *Journal of Business and Technical Communication*, 22(2), 135–159. <https://doi.org/10.1177/1050651907311923>
22. Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
23. Dewey, J. (1934). *Art as Experience*. Minton, Balch & Company.
24. Drummond, H. (2026). Aesthetic value, artistic value, and the socially extended mind. *Synthese*, 207, Article 11. <https://doi.org/10.1007/s11229-025-05411-x>
25. Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156.

26. Ericson, J., & Klingberg, T. (2023). A dual-process model for cognitive training. *npj Science of Learning*, 8, 12. <https://doi.org/10.1038/s41539-023-00161-2>
27. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406. <https://doi.org/10.1037/0033-295X.100.3.363>
28. Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, 28, 717–741.
29. Fiorella, L., & Mayer, R. E. (2015). *Learning as a Generative Activity: Eight Learning Strategies That Promote Understanding*. Cambridge University Press.
30. Franzoi, L., et al. (2025). Retrieval practice enhances learning in real primary school settings. *Frontiers in Psychology*, 16, 1632206. <https://doi.org/10.3389/fpsyg.2025.1632206>
31. Gao, L., Jong, M. S.-Y., Chai, C.-S., & Ng, O.-L. (2025). Undergraduate engineering students' epistemic cognition and changes in the course of being engineering design mentors. *International Journal of STEM Education*, 12, 42. <https://doi.org/10.1186/s40594-025-00564-0>
32. Gray, C. G., & Taylor, E. F. (2007). When action is not least. *American Journal of Physics*, 75(5), 434–458. <https://doi.org/10.1119/1.2710480>
33. Greene, J. A., Sandoval, W. A., & Bråten, I. (Eds.). (2016). *Handbook of Epistemic Cognition*. Routledge.
34. Greenstreet, F., Vergara, H. M., Johansson, Y., et al. (2025). Dopaminergic action prediction errors serve as a value-free teaching signal. *Nature*, 643, 1333–1342. <https://doi.org/10.1038/s41586-025-09008-9>
35. Hadwin, A. F., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (2nd ed.). Routledge.
36. Hamou, N., Gershman, S. J., & Reddy, G. (2025). Reconciling time and prediction error theories of associative learning. *Nature Communications*, 16, 10265. <https://doi.org/10.1038/s41467-025-65137-9>
37. Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child Development and Education in Japan*.
38. Hattie, J., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, 1, 16013. <https://doi.org/10.1038/npjscilearn.2016.13>
39. Hattie, J., & O'Leary, T. (2025). Learning styles, preferences, or strategies? An explanation for the resurgence of styles across many meta-analyses. *Educational Psychology Review*, 37, 31. <https://doi.org/10.1007/s10648-025-10002-w>
40. Hetherington, L. (2024). Material-dialogic space as a framework for understanding material and embodied interaction in science education. *Research in Science & Technological Education*. <https://doi.org/10.1080/02635143.2024.2328766>
41. Hetland, L., Winner, E., Veenema, S., & Sheridan, K. M. (2013). *Studio Thinking 2: The Real Benefits of Visual Arts Education*. Teachers College Press.
42. Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
43. Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, 21(1), 45–83. <https://doi.org/10.1080/10508406.2011.591717>
44. Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424.
45. Karpicke, J. D. (2025). Retrieval-based learning. In J. Wixted (Ed.), *Learning and Memory: A Comprehensive Reference* (3rd ed., Vol. 4, pp. 411–441). Elsevier.

46. Kierzenka, J., & Shampine, L. F. (2001). A BVP solver based on residual control and the MATLAB PSE. *ACM Transactions on Mathematical Software*, 27(3), 299–316.
47. Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist*, 41(2), 75–86.
48. Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Prentice Hall.
49. Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60(2), 265–299. <https://doi.org/10.3102/00346543060002265>
50. Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
51. Lee, O., Wei, T., Stenning, K. D., et al. (2024). Task-adaptive physical reservoir computing. *Nature Materials*, 23, 79–87. <https://doi.org/10.1038/s41563-023-01698-8>
52. Li, Q., Wang, D., Xiao, W., Tang, Y., Sun, Q., Sun, B., et al. (2024). Structured interaction between teacher and student in the flipped classroom enhances learning and interbrain synchrony. *npj Science of Learning*, 9, 73. <https://doi.org/10.1038/s41539-024-00286-y>
53. Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist*, 47(3), 232–247.
54. Malafouris, L. (2013). *How Things Shape the Mind: A Theory of Material Engagement*. MIT Press.
55. Maras, K., & Shand, B. (2023). Critical and creative thinking as a form of making in art education. *Studies in Art Education*, 64(1), 23–39. <https://doi.org/10.1080/00393541.2022.2154523>
56. Marzano, D. (2025). Generative Artificial Intelligence (GAI) in teaching and learning processes at the K–12 level: A systematic review. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-025-09853-7>
57. Meyer, J. H. F., & Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines. ETL Project Occasional Report 4.
58. Mezirow, J. (1991). *Transformative Dimensions of Adult Learning*. Jossey-Bass.
59. Møller-Skau, M. (2024). Exploring student teachers' aesthetic learning processes. *Teaching Education*, 35(3), 277–292. <https://doi.org/10.1080/10476210.2023.2278802>
60. Nair, V. N. (Ed.). (1992). Taguchi's parameter design: A panel discussion. *Technometrics*, 34(2), 127–161. <https://doi.org/10.2307/1269231>
61. Nannicelli, T. (2024). Aesthetic education: A perceptual-cognitive model. *European Journal of Analytic Philosophy*, 20, 283–305.
62. Noë, A. (2015). *Strange Tools: Art and Human Nature*. Hill and Wang.
63. Papert, S., & Harel, I. (1991). *Constructionism*. Ablex Publishing.
64. Perera, D., et al. (2018). A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science*, 359(6374), 429–434.
65. Pizzolante, M., Pelowski, M., Demmer, T. R., Bartolotta, S., Sarcinella, E. D., Gaggioli, A., & Chirico, A. (2024). Aesthetic experiences and their transformative power: A systematic review. *Frontiers in Psychology*, 15, 1328449. <https://doi.org/10.3389/fpsyg.2024.1328449>
66. Pupillo, F., Ortiz-Tudela, J., Bruckner, R., et al. (2023). The effect of prediction error on episodic memory encoding is modulated by the outcome of the predictions. *npj Science of Learning*, 8, 18. <https://doi.org/10.1038/s41539-023-00166-x>

67. Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159–179.
<https://doi.org/10.1080/02602930801956059>
68. Sato, T., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459(7244), 262–265.
69. Scadden, D. T. (2006). The stem-cell niche as an entity of action. *Nature*, 441(7097), 1075–1079.
<https://doi.org/10.1038/nature04957>
70. Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed.). Cambridge University Press.
71. Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*. Cambridge University Press.
72. Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.
73. Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective*.
74. Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of the Learning Sciences*, 20(3), 472–506.
75. Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, 22(2), 129–184.
76. Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690.
77. Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, 91(5), 761–798.
<https://doi.org/10.3102/00346543211019105>
78. Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176–199.
<https://doi.org/10.1177/1745691615569000>
79. Stern, M., Guzman, M., Martins, F., Liu, A. J., & Balasubramanian, V. (2025). Physical networks become what they learn. *Physical Review Letters*, 134, 147402.
<https://doi.org/10.1103/PhysRevLett.134.147402>
80. Stern, M., Liu, A. J., & Balasubramanian, V. (2024). Physical effects of learning. *Physical Review E*, 109, 024311. <https://doi.org/10.1103/PhysRevE.109.024311>
81. Struik, L. C. E. (1978). *Physical Aging in Amorphous Polymers and Other Materials*. Elsevier.
82. Tang, H., Jiang, S., & Xie, C. (2025). Solarizing Your School: Engineering design in students' authentic epistemic practices of adopting renewable energy. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-025-10258-5>
83. Trybulec, M., & Howiecka-Tańska, I. (2025). Are interactive exhibits at a science center cognitive artifacts? *Foundations of Science*, 30(3), 651–675. <https://doi.org/10.1007/s10699-024-09959-8>
84. Tufail, M., Park, H., Wang, H., & Kim, K. M. (2026). An empirical study of cognitive load and constraint-driven innovation in the early phase of product design within a digitally mediated medium. *International Journal of Design Creativity and Innovation*, 14(2), 122–149.
<https://doi.org/10.1080/21650349.2026.2617533>

85. Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
86. Walsh, D. (2026). Strange and ordinary tools: The reciprocity of art and life for Nishida Kitarō and Alva Noë. *The Journal of Aesthetics and Art Criticism*, 84(1), 30–43.
<https://doi.org/10.1093/jaac/kpaf048>
87. Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.
88. Xu, X., Zhang, D., & Zhang, Y. (2026). Student-teacher inter-brain coupling forecast cross-semester academic fluctuation in real-world classrooms. *npj Science of Learning*, 11, 30.
<https://doi.org/10.1038/s41539-026-00421-x>
89. Yesilyurt, E., Deniz, H., & Kaya, E. (2024). Exploring epistemic aspects of engineering for K–12 science and engineering education. *Journal of Engineering Education*, 113(2), 439–467.
<https://doi.org/10.1002/jee.20593>
90. Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70.

后 记

一、它最后成了什么

这本书的九章原是九篇各自独立的长文，写它们的时候我并没有打算写一本书。把它们放在一起的念头是被它们自己逼出来的：第三编那三章各自写完之后，同时交出了同一句话——同一次改变有三处可以失效；第二编那三章又各自在末尾交出另一句同样的话——课堂运行里被当作给定的那些，有一部分是这个过程自己的产物。到第三次巧合的时候，再当作巧合就是懒惰。

于是有了第一编去回答“那样东西究竟是什么”，和第四编去回答“那么账该怎么记”。

现在回头看，真正立起来的并不是那九个对象。九个名字里，将来大概会有几个被证明只是另几个的代理——第十一章已经把合并的规则先写好了。**立起来的是那个更朴素的问题：什么样的东西才可能成为一栏账。**这个问题一旦被问出来，教育之外的许多地方都会跟着松动：医院的病历、法院的卷宗、公司的绩效表，每一本账都只为可以事前写下的东西设了字段，而每一处真正决定后果的，都有一部分是那个过程自己正在产出的。本书没有力气去处理那些地方，只在课堂上把这件事做实了一次。

二、装成一本书时做了什么，没做什么

做了的：九章的刊物体例被拆掉——摘要、关键词、英文提要、各自的参考文献被抽出或合并；“本文”改成“本章”；原本指向其他文章的划界，凡对象已收入本书的，一律改成章号互引；第一编三章合并为一个对象名，另两个名字降为它的传递机制与显露形态。此外新写了前言、导论、三篇编末小结、第四编六章、结语与七个附录，约五万字。

改过的：四处结构模型的表述被改写为**带条件的版本**——换位模型里量的定义与一阶条件、默认推进规则的适用范围、完备度反比成立的区间、三处失效的条件通过率连乘。旧的强表述已经删掉，不再以“正文先写强句、后面再附注纠正”的双层形式并存。四处改的是**表述的强度与符号的统一，不是结论本身。**

没做的：九章的论证一字未削。它们原来的判据、读数、辨别格、划界、证伪条件与赌注都原样保留，包括那些我后来觉得写得过满的地方。理由是：**一本由独立成文的东西装起来的书，如果编者把每一处棱角都磨平，读者就再也无法判断哪些结论是各自独立得到的。**异源汇聚之所以能算作证据，前提是那三条路径没有被事后修圆。

也因此，有一条建议我没有照做：把全书压缩百分之八。我先把“重复”量了一遍——真正机械重复的样板（“本章不报告新采集的数据”、“不含判断词”的释义、“这不是价值判断”一类声明）**全书合计不到一千字**；而九章里那些问题的小节（与最近邻的划界、五种最强的竞争解释、证伪与赌注）**标题相同而内容各不相同**，每一章处理的是它自己的对手。要凑够百分之八，就必须动这些地方，而那不是压缩重复，是删论证。**这本书宁可显得啰嗦，也不肯让读者以为某一章的划界是从另一章抄来的。**

三、欠着的三笔账

第一笔，六个读数一次也没有真的采过。这件事我在前言里说过，这里再说一遍，因为它是全书最大的空缺：所有档界、所有相关预测、所有成本估算，都是从机制推出来的，不是从数据来的。附录一把它压到了一份可以照着做的手册，剩下的只能等人去做。

第二笔，第一编的合并可能是错的。三组学科各自走到几乎同一句话上——我把它读作稳健性证据，但它也可以被读作我在三个入口各放了一面镜子。第一编末尾给了判决办法，而在判决之前，那一编是全书最脆的一编。

第三笔，第十三章的失败没有被解决。三条结构模型目前收不成一个式子，我把三次尝试失败在哪儿都写了出来，也写了要收成一个式子还缺什么。那一章诚实，但诚实不等于完成。

四、这本书最可能的坏结局

不是被反驳，是被抽走六个读数、做成一份检查表。

那份表会很好用：六栏、每栏三档、一个总分。它会被印出来，会被填，会被拿去比较，然后在两三年里失去与它原本要测的那件事之间的全部联系——第十五章把这个过程完整地推演过一遍，包括它为什么不能靠“把度量做得更好”来避免。

我能做的只有三件，都做了：把反噬写进正文而不是附录；把失效期写进制度建议；把“它会被怎样填坏”写在给出问卷的同一章里。第四件只有读者能做——**当你看见那六个读数出现在一张检查表上，而那张表没有要求同时报分母、触发源与可交接性时，请指出那不是本书的东西。**

五、最后一句

写这本书的过程里，我改变过一个看法。

一开始我以为，那些看不见的东西之所以看不见，是因为还没有人去看。写到第六章那个式子的时候我才明白不是——**是账本每完备一分，那一段就被压掉一分**；那些把课上好的努力，与那些把这一段压掉的力量，是同一股力量。这一条让整本书变得比我预想的难受，也让它比我预想的值得写。

它不是一个可以靠更用心来解决的问题。能做的只有：**在按表推进之前，先看一眼表上没有的那几栏。**

王德生

二〇二六年八月 于新加坡

作者简介

王德生

德麦国际（新加坡）创办人。

中国科学院计算数学博士及博士后，早年治 Voronoi-Delaunay 剖分、重心 Voronoi 图与有限元方法；后任职于斯旺西大学与南洋理工大学。

近年以发生学为纲，跨近二十个领域写作，已出版专著五十余部。本书属其改姓写作的一支：框架隐在背后，正文只说目标学科自己的话，全部主张压到可采集的读数与可推翻的条件上。

ORCID 0009-0009-8196-0030

窗关上时什么也不发生， 因此这件事在原理上不会以失败的形式被发现。

一堂课上得很顺：讲清楚了，练到位了，测出来也好看。几周之后换一个表面，什么也不剩。这件事人人遇见过，人人能解释，而所有解释都指向学生、教师或教法。

本书换了一种做法：把九组互不相容的学科请到同一个问题面前，看它们在哪里撞上。撞出来的是九个此前没有名字的对象——空窗、悬置连接、冻结构型、换位区间、可倒推区间、自建条件……它们的共同点是：都由课堂自己造出来，都决定后果，而课堂的账本上没有它们的字段。

更难受的一层是：这不是疏漏。可以事前写下的东西才可能成为一栏账，而这些东西在写表格的时候还不存在。于是账本每完备一分，那一段就被压掉一分——那些把课上好的努力，与那些把这一段压掉的力量，是同一股力量。

全书三编九章，另有一编新写的账本：六个可以由不懂本书理论的人采集的读数、九句不含判断词的问法、三条形式推导，以及约六十条可以让它自己作废的条件。

AUTHOR

王德生

德麦国际（新加坡）创办人。中国科学院计算数学博士及博士后，早年治 Voronoi-Delaunay 剖分、重心 Voronoi 图与有限元；后任职斯旺西大学与南洋理工大学。近年以发生学为纲，跨近二十个领域写作，已出版专著五十余部。

课堂学习发生学

王德生 著

ISBN 978-1-970820-90-4

US\$ 22.00



9781970820904