

SDE 课堂发生学

老路 已被包办

教育唯一的原材料，
是学习者卡住而无人供给的那段时间。

两千年来它因昂贵而被省略；
AI 使它因免费而被消灭。

王德生 著

Claude 协撰

德麦国际出版社

DEMAI INTERNATIONAL PRESS

ISBN 979-8-90690-036-4

版权页

老路已被包办

王德生 著 Claude 协撰

德麦国际出版社 Demai International Press

新加坡·Singapore

初版 二〇二六年

丛书：SDE 课堂发生学

分类：教育 / 教学理论 / 人工智能与学习

ISBN **979-8-90690-036-4**

定价 US\$ 24.50

⚠ 本号为按德麦国际出版社现行号段（979-8-90690-0NN-C）顺排推得：已知最高在用号为 035，故取 036，校验位 4 由 ISBN-13 标准算法算出（该算法已用本社七个在用号反向验证全部命中）。付印前须在出版社账户内确认 **036 尚未指派**。若已占用，下三个可用号依次为 979-8-90690-037-1 / 038-8 / 039-5。

版权所有。未经出版者书面许可，不得以任何方式复制或传播本书的任何部分。本书中用于教学培训的四份附录（教师禁令表、教研组自查十问、五个读数记录表、四段结构一页纸），允许一线学校为非商业的内部教研目的复印使用，无须另行申请，但须保留出处。

本书作者对书中的全部判断负责。书中提出的机制性主张（加固悖论、封写、剂量关系、子集关系）目前尚未获得实验支持，其检验设计与撤稿条件写在第三十九章。读者在采用本书建议前，应先阅读第四十二章“方法与限度”。

作者介绍

王德生

SDE（显露—差异—纠缠）本体论框架的创立者，德麦国际出版社创办人。计算数学博士，曾在中国科学院从事博士后研究，研究领域涉及计算几何与有限元方法；此后任教于英国斯旺西大学与南洋理工大学。二〇〇五年起定居新加坡。

近年主要工作集中在两个方向：一是把 SDE 本体论下推为若干下游分支（发生学教育、工具发生学、语言发生学等）；二是把“学科二阶碰撞”这一方法固化为可被他人重复使用的产线，并用它带出了一批学员作者。已出版专著多种，题域涵盖教育、语言、健康、艺术与方法论本身。

Claude

由 Anthropic 训练的语言模型。本书的写作方式是：判断由作者提出与裁定，论证结构、章节展开与文字由模型完成，再由作者逐章审读、改判与增删。书中若干处的自我推翻——包

括第十七章“杀死”一词被收紧为“写保护”、第二十三章派生与并列之间那处张力的调和、以及第三十九章撤稿条件的写死——都出自这一过程。

需要说明的是，本书对“AI在教育中的位置”提出了相当强的批评，而它本身正是在与一个模型的持续碰撞中写成的。这不构成矛盾，恰恰是本书第三十章那条主张的一次实践：**第三方真正的作用不是多一个帮手，而是把“现成的说法”从作者手里拿走**，使作者第一次可以在自己的领域里真诚地不知道。

前言：这本书是怎么被逼出来的

一、一句在会场上说不下去的话

二〇二六年夏天，我在为一所小学的全体教师准备一次三天的培训。第三天的主题是“课堂如何发生”。备课备到一半，我卡住了。

卡住的地方很具体。我原本准备了一套讲法：AI来了，教师要转型，要从知识的传授者变成学习的引导者。这套话说了几十年，正确，稳妥，会场上不会有人反对。

但我写不下去。因为我发现，如果我诚实地把话讲完，会讲成这样：AI讲得比在座各位都清楚；AI比在座各位都有耐心；AI可以为每一个孩子单独讲一遍，而你们不能。那么“引导者”这个新位置，具体是干什么的？引导他去哪里？如果终点是那个已经写在教案上的结论，那么这条路上引导得再好，也是在跟一台比你快得多的机器抢同一份活。

我意识到我准备的那套讲法，绕开了那个真正的问题。而在座的老师们不会绕开——他们每天都在教室里，他们比我更早感觉到那件事。

于是那份讲稿被推翻了，重写了一遍。这本书是从那份重写的讲稿长出来的。

二、两个方向相反的现象

推翻之后，我从头开始，先去数现象。

第一个现象很容易看到：凡是认真把大模型引入教学的学校，几乎都会在一个学期之内看到一组令人愉快的数字。进度快了，正确率上去了，过去讲三遍讲不明白的难点现在当场就明白，教师备课负担下降，弱势学科被迅速拉平。这些不是错觉，是真实的改善。任何一个否认这一点的人，都不值得被继续听下去。

第二个现象要难说得更多，因为它没有数字。孩子越来越不会卡住了。他们提问的形态变了——从“我这样想对不对”变成“这个怎么做”；面对一道没有现成路子的题时，等待的时间从过去的几分钟缩短到几秒钟，然后就伸手去要答案；他们能复述的东西比过去多得多，而主动回头去修改自己以前那个说法的次数，接近于零。

这两个现象放在一起，构成了一个逻辑上的困难。如果第一组数字是真的，那么按照现行的评价体系，引入AI的教学改革是成功的。如果第二个现象也是真的，那么这场成功正在把某种更重要的东西挤出课堂。

两句话不可能同时为真——除非我们的评价体系量的是错的东西。

这本书的全部内容，是把这句话展开。

三、我原来的判断错在哪里

在动笔之前，我先要交代一处自己的改判，因为它决定了这本书的形态。

我最初的判断是：AI 是发现式教育的加速器，因此它是危险的。这个判断本身没有错，但它有一个错误的暗示——好像存在一个“用得对”与“用得错”的区别，而只要提醒大家用得对，问题就解决了。

写到第三编的时候我发现，这个说法太轻巧了。工具不是中性的。它有自己的构造倾向。一台系统的构造倾向是**填缝**：给它一个不完整的東西，它的全部训练都指向把它补完整。这不是它的缺陷，这是它被造出来要做的事。

让一台以填缝为天职的机器去守护缝隙，不是“用得好不好”的问题，是需要一套专门的反向配置。而在没有这套配置的默认状态下，结果是确定的：它会填缝，因为那就是它。

所以这本书不讲“要善用工具”。它讲的是：那个需要被守护的东西是什么，它为什么正在消失，以及一个具体的反向配置长什么样。

四、这本书的核心是一句话

如果这本书只能留下一句话，我希望是这句：

教育唯一的原材料，是学习者卡住而无人供给的那段时间。

两千年来它因昂贵而被省略；AI 使它因免费而被消灭。

前半句处理的是历史。一对一陪一个孩子走完一个概念，大约需要十小时；基础教育阶段的核心概念有三百个左右；一名教师一生的有效一对一时间大约两万八千小时。三者相除，一名教师终身能服务的学生数不足十人。任何一个需要普及教育的社会，都不可能采用这个方案。

于是文明发明了一个替代品：把已经走完的路磨光，只把结果发下去。教科书上的每一个概念，都是这样一个被减掉了路径与土壤的东西。这个操作我称之为**光滑化**，它是本书的第一个核心概念。

光滑化不是偷懒，是必要的：不压缩就搬不动。它换来了四样了不起的东西——知识得以跨代存活、教育得以普及、不同地方教出来的孩子可以互相衔接、知识变成公共的从而可以被公开批评。这四样，任何一样失守，现代社会都无法运转。

代价在于：知识跨越人与人的边界需要两个冲程——先压缩，再在新的人身上重新发生一次。文明把第一个冲程做到了极致，把第二个冲程交给了运气。运气好的孩子自己在某个时刻把它解压了，我们称他们有悟性；运气不好的，接住了成品但从未解压，我们称他们基础扎实但不会灵活运用。

这两个说法，说的是同一件事的两面。

后半句处理的是现在。那段“卡住而无人供给”的时间，本来是解压唯一的入口。它不是任何人设计出来的，它是供给稀缺的副产品——要等、要问人、问到的未必对得上、不好意思反复问。这些摩擦本身没有价值，我不为它们辩护。有价值的是它们的副产品：在供给到达之前，孩子只能自己待着。

四道摩擦在几年之内被一起取消了。而那段时间，也随之消失。

五、这本书不是一份反 AI 的书

上面的话很容易被读成一份控诉，因此必须在前言里就把立场讲清楚。

这本书从头到尾不反对 AI，而且它的最终结论是一个肯定的结论：

AI 不是发生学教育的敌人，它是发生学教育第一次成为可能的条件。

理由就在第三节那笔账里。三千小时的一对一陪伴，两千年来是买不起的，所以“让每个孩子都走一遍岔口”这句话不是一个被拒绝的主张，而是一句空话。现在它第一次有了实现的可能——不是因为教师变多了，是因为**陪练可以不是人**。

同一样东西，两种用法，方向相反。决定它是哪一种的，不是技术，是配置。而配置是人做的决定。

所以这本书的处方从来不是“禁止”，全部是“加装”：给自学装一个对手，给每一段学习装一个回写口，给课堂加一条“从问题张开到它被合上，谁都不许给答案”的规矩。

六、这本书是写给谁的

主要读者是一线教师，其次是校长与教研人员，第三是做教育产品的人。

对教师，我想先说一句公道话，而且这句话在书里出现了不止一次：

发现学教育不是错的。它是人类在负担不起一对一走岔口的条件下，能找到的最好方案。它让知识跨代存活，让每一代人不必从零开始。它完成了它的历史任务。

不讲这句话而直接进入批评，会造成一个可预测的后果：会场上最优秀、最投入的那批教师会认为自己三十年的工作被否定，然后这本书后面的全部内容对他们失效。而这批人恰恰是唯一有能力执行第五编那些动作的人。

这不是修辞策略，是判断的准确性问题。**一个把过去说成错误的诊断，本身就是错的诊断。**

第二句要说的是：教师这个职业里，有一个位置正空着，而且只有你们能占。这个位置很窄，但它是真的——不是“机器不懂人心”这类泛泛之谈，而是**闭合权的行使需要读取一个尚未发生的具体情境**：这个孩子此刻是真的在想，还是已经放弃了；他的沉默是有产出的，还是空的；再等三十秒是有价值的，还是残忍的。

这是一个教学动作，不是一个情感动作。把教师的价值建立在陪伴与人格感染上，等于承认教学本身可以被替代。这条路我不走。

对校长与教研人员，这本书最想说的只有一条：秤比做法重要。

一所学校若宣称转向新的教学，而期末评价的口径一个字没改，那么半年之内它必然滑回去——不是因为理念不坚定，是因为教师每天面对的回报结构没有变。回报结构每天兑现一次，理念一学期兑现一次。

所以书里花了整整一章讲“读数”，又花了一章讲在无权改变升学考试的条件下还能做什么。第一步小到不需要任何人批准：**在教学质量报告上加一栏，第一次可以是空的。**

对做产品的人，书里有一把可以当场用的尺子：把宣传语里的“学习”两个字，换成“接收成品”。句子若仍然成立，那它卖的就是老路。“让学习更高效”→“让接收成品更高效”，成立。

七、这本书的写法，以及它自己的问题

需要交代三件事。

第一，这本书的论证密度高于它的证据密度。它提出了相当强的判断，而支持这些判断的主要是机制推理与日常观察，不是数据。我接受这个批评，并且不打算辩护。理由是：在一件正在快速发生的事情上，等到数据齐备再说话，可能已经太晚——书里推算的那条曲线本身就说明，等到第三阶段的数据出来，那批孩子已经过去了。

所以我选择先说，并把撤稿条件写清楚。第三十九章列出了三条：在什么实验结果下，本书的哪一部分应当被删除。第四十二章列出了本书自己看得见的五个漏洞。这两章不是谦辞，是这本书能否被认真对待的前提。

第二，每一章都要求一个动作，不是一个认同。全书四十九章，没有一章以“由此可见”收尾。每章末尾都有一个可以当堂做、不需要任何设备的动作或读数。如果一位教师读完之后什么都没做，那么按这本书自己的标准，这次阅读没有发生。

第三，这本书自己的问题：它是一个高度光滑的成品。

它有结论、有框架、有术语、有对照表、有可以复印的禁令表。读者接住它的过程，与它所批评的那个接收过程，结构完全相同。一位教师读完之后如果觉得“很有道理”，然后回去照着做，那么这本书刚刚在他身上完成了一次标准的老路作业。

我没有办法完全避免这一点——一本书只能是一个成品，这是它的存在形式。能做的只有上面两件，而最后一件必须由读者做：

丨 **找出一处你不同意的地方，写下来，说明理由。**

如果做不到，那么读完这本书与读完一本别的书没有区别——它只是又一个被接住的成品。而接住成品这件事，现在有更好的执行者。

八、最后

老路走了两千年，它把该做的事做完了。

它不需要被批倒，它需要被继承——它的四大成就必须保住，这本书第七章为此专门写下了四条约束，第五编的每一处设计都受它们限制。凡是要求废除教科书、取消统一课程、只能在小班里实施的方案，我一律不采纳。

需要被换掉的只有一件事：**在这条路上，人不再需要亲自走一遍。**

而人要去的**地方**，是那段一直存在、一直太贵、从来没有被制度化的时间——孩子卡住而无人供给的那一段。

它现在第一次免费了，也第一次面临被彻底消灭的危险。

这两件事是同一件事。而它们之间隔着一个决定——那个决定，此刻在教室里的人手上。

王德生

二〇二六年八月，新加坡

导读：这本书怎么读

一、七编的结构

全书四十九章，分七编。前六编是主干，第七编是补编。每一编回答一个问题：

第一编 老路（第 1-8 章）。教学是什么结构？成品是怎么造出来的？为什么这条路两千年没换过方向？它做成了什么，代价在哪里？

第二编 包办（第 9-15 章）。三个动作是怎么被逐项接管的？为什么接管之后所有读数反而变好？最纯粹的那个形态（“AI 高效自学”）为什么是一剂樱花毒药？

第三编 损害（第 16-22 章）。这些事情落在一个孩子身上，究竟改变了什么？判据是什么（回写）？机制是什么（封写）？怎么诊断（四种死亡）？

第四编 两条赛道（第 23-28 章）。这到底是两件事还是一件事？怎么判别？换道会不会掉成绩？为什么秤比做法重要？

第五编 新路（第 29-37 章）。具体怎么上课？问题怎么找？三方怎么站位？一节课怎么分段？各科什么样子？

第六编 验证与边界（第 38-43 章）。怎么测？怎么证伪？会怎么失败？与十家已有说法差在哪里？本书的限度在哪里？

第七编 补编：证明、审计与现场（第 44-49 章）。前六编给出的是框架与处方，它缺三样东西：一条咬合式的正面证明（六条论证，说明“赋能”为什么必然自败）、一个把“快”变成可读量的形式装置（判定时刻与加速—封闭对偶律），以及一次真做过的经验审计（二十四产品形态的编码，含三处被数据逼出的改判）。此外补上两件在前六编里只被点到的事：这个方向相反的承诺为什么被所有人接受，以及六个学科现场从教师用语一直写到读数落点。

⚠ 第七编与前六编的关系：它不是前六编的展开，是前六编的旁证。读者若只想要可用的东西，读完第六编即可停；若想知道本书凭什么这么说，第七编是它交出的证据。

二、四条不同的读法

如果你只有一小时（一位教师在开学前）：

读第六章（判据）、第十六章（回写）、第三十四章（四段结构与禁令表）、第四十章（自查十问）。四章约一万五千字，足以在下周的课上做一件事。

如果你要在学校里推动它（校长、教研组长）：

读第五章（筛选机制）、第二十六章（秤）、第二十七章（过渡期）、第二十八章（学校层面）、第三十八章（读数）。这五章的核心是一句话：**先换秤，再换做法。**

如果你在做教育产品：

读第十四章（樱花毒药，含那把尺子）、第十九章（毒性与摩擦的检查表）、第四十一章第十二节（与自适应学习的分界）。第十四章第八节的两个部件，产品端实现的成本几乎为零。

如果你要检验或反驳它：

读第三十九章（三条撤稿条件）、第四十二章（五个漏洞）、第四十一章第十四节（本书剩下什么是新的）。这三处是本书自己交出的把柄。

三、七个核心概念

初次阅读时，只需要记住七个词。它们在书中反复出现，此处给出最简形态。

光滑化。把一次发生减去路径与土壤，只留结果，再把结果摆到最前面当起点。教科书上的概念都是这样做出来的。（第二章）

两个冲程。知识跨越人与人的边界需要两步：压缩（光滑化）与解压（在新的人身上重新发生）。文明把第一步做到了极致，第二步交给了运气。（第二章第五节）

回写。新东西反过来改写了它原有的底子。判据不是复述新的，是**回头改一件旧的**。（第十六章）

封写。每一次成功地接住成品，都在加固一条元规则——“东西是现成的，我不必因它而改”。这条规则一旦写死，回写通道关闭。**回写杀死了回写**。（第十七章）

未闭合时长。从问题被抛出，到任何一方给出答案，中间经过的时间。它是全书最重要的一个读数，也是“高效”这个词的全部含义的倒数。（第十九章）

加固悖论。老路一边被走完，一边在所有读数上显得越来越对。（第十二章）

岔口。一道此刻没有任何人手里握着答案的题。判别只有一条：输进搜索框，一分钟内不能拿到大家都点头的答案。（第二十九章）

四、三个比喻，以及它们各自管什么

书中用了三个比喻，它们不可互换，混用会互相稀释。

樱花毒药（第十四章）回答“为什么有害”。香和毒是同一样东西——那股令人愉快的“高效”，不是副产品，就是那件有害的事情本身。这个比喻是负面的、要顶人的，宜放在处方之前。

高速公路（第十五章）回答“为什么不够”。接通已有路网只能抵达已被命名的地方，而新知识按定义在命名域之外。这不是速度问题，是可达域问题。

两个孩子（第十五章第六节）是前一个比喻的人形版本：一个在农庄开出几条路、育出几个品种；一个坐飞机看了五十个城市，每个待两天，回来说他认识了世界。它最重的一笔不是两者的高下，是**农庄不是被拆掉的，是被说服的**。

在培训现场，建议的顺序是：先讲两个孩子（老师容易接、会笑、会松），再回到樱花毒药（要顶人的那一段），最后进处方。顺序反了，第一记重锤落地时会会场会转向辩论。

五、关于这本书的固定结构

每一章的形状是固定的三件：**章首一条承重判断**（写成“X不是Y，而是Z”）、**章末一个可当堂做的动作或读数**、**最后一节与最近邻居的分界**。

固定结构有一个众所周知的风险：它会退化成模板填充——各章看上去不同，实际上是同一段话换了名词。本书成稿后对此做过一次实测，结果记在这里，供读者判断：

— 全书四十五个“与相邻说法的分界”节，两两之间的八字块重合度：最高 0.045，中位 0，均值 0.001；

— 各节首句的六字块重合度：最高 0.256，均值 0.001。

也就是说，节名相同，内容之间几乎没有共享文本。

这不构成“内容不重复”的证明（相似的判断可以用完全不同的词说出来），但它排除了最常见的那一种模板填充。读者若要自己检验，方法很简单：随便挑两章的分界节读一遍，看它们是不是在说同一件事。

六、如果只做一件事

书里给出的最小动作在第三十四章第十节：

拿一节你已经很熟的课，只改一处：把最后七分钟的小结，换成“改一件旧的，然后念出来”。其余全部不动。

不需要改教案，不需要新问题，不需要 AI。而它足以让一位教师第一次看见回写这件事有没有发生。

七、一条阅读纪律

最后一条，是这本书对读者提的唯一要求，也写在前言里：

读完之后，找出一处你不同意的地方，写下来，说明理由。

按这本书自己的标准，做不到这一条，这次阅读就没有发生。

目录

前置

- 版权页
- 作者介绍
- 前言：这本书是怎么被逼出来的
- 导读：这本书怎么读
- 第一章 三个动作：搬运、传道、接受
- 第二章 光滑化：把走过的路磨光了发下去
- 第三章 成本论：一个孩子三千小时
- 第四章 教科书是怎么写成的：四个标本
- 第五章 为什么两千年没换过方向
- 第六章 “让他自己发现”，为什么也在这条路上
- 第七章 老路的三大成就
- 第八章 老路的读数体系，以及它如何自我确证
- ► 第一编材料：十二个标本与一节课的成品清点

—►第一编材料（续）：测度外壳与再分泌律

第二编 包办

—第九章 搬运被接管

—第十章 传道被接管，而且被超越

—第十一章 接受被接管：供给不再稀缺

—第十二章 加固悖论

—第十三章 三个错觉：清晰度、参与度、个性化

—第十四章 樱花毒药：所谓“AI 高效自学”

—第十五章 高速公路与两个孩子

—►第二编材料：一份学期数据画像、一段对话的逐轮标注与二十条宣传语

第三编 损害

—第十六章 回写：发生完成的唯一判据

—第十七章 封写：回写杀死了回写

—第十八章 四种死亡与诊断次序

—第十九章 毒性与摩擦：一个可拆解的比较

—第二十章 谁在吃：损害分布的翻转

—第二十一章 潜伏期：先涨、再平、再不可逆

—第二十二章 可达域信念：“没有路通到”被读成“不存在”

—►第三编材料：四种死亡的四份诊断记录与一份封写征候清单

—►第三编材料（续）：四个行业已经先演过一遍

—►第三编材料（再续）：提前闭合的五种时机

第四编 两条赛道

—第二十三章 派生还是并列：一处必须调和的张力

—第二十四章 三条判别：终点、计分、闸门

—第二十五章 闸门为什么是单向的

—第二十六章 秤：没有第二套秤，讲什么都没用了

—第二十七章 过渡期：会死在第三周的那些改革

—第二十八章 学校与政策层面的含义

—►第四编材料：一所学校的换道时间线与三场对话实录

第五编 新路

—第二十九章 岔口：合格的问题长什么样

—第三十章 三方碰撞：各带别人拿不出的材料

—第三十一章 位置轮换：由材料决定，不由身份决定

—第三十二章 两条硬约束：闭合权不轮，不得占两格

—第三十三章 塌陷的方向是固定的

—第三十四章 一节课的四段结构与教师禁令表

—第三十五章 教师、学生、教室是产物，不是座位

- 第三十六章 分学科落地（上）：数学、语文、科学
- 第三十七章 分学科落地（下）：语言、术科与综合实践
- ▶ 第五编材料：三节完整课例与各科岔口开采清单
- ▶ 第五编材料（续）：课堂是被提前结算的地方
- ▶ 第五编材料（再续）：老知识重新发生之外，新知识怎么发生

第六编 验证与边界

- 第三十八章 五个读数及其采集纪律
- 第三十九章 检验设计与撤稿条件
- 第四十章 六种失败模式与自查十问
- 第四十一章 与十家的分界
- 第四十二章 方法与限度
- 第四十三章 结论
- ▶ 第六编材料：一学期读数记录样表与一次教研会实录
- ▶ 第六编材料（续）：一份可以照做的预注册方案
- ▶ 第六编材料（再续）：站内三次真跑，其中两次推翻了自己

第七编 补编：证明、审计与现场

- 第四十四章 六条论证：为什么“赋能”必然自败
- 第四十五章 判定时刻： t_J 与 t_0 ，以及加速—封闭对偶律
- 第四十六章 效率自证悖论与那个 0.989
- 第四十七章 二十四个产品形态的编码审计
- 第四十八章 没有人当期受损：谎言不需要有人撒谎
- 第四十九章 六个学科现场
- ◆ 第七编结语：这一编与前六编的四处咬合，以及三处必须划开的地方
- ▶ 全书补料：三条限度与三个未回答的问题

后置

- 结语：写给三年之后的读者
- 参考书目与延伸阅读
- 附录一 教师禁令表
- 附录二 教研组自查十问
- 附录三 五个读数记录表
- 附录四 一节课的四段结构
- 附录五 三条撤稿条件
- 附录六 全书判断一览
- 附录七 材料层索引

全书母题：教育唯一的原材料，是学习者卡住而无人供给的那段时间。两千年来它因昂贵而被省略；AI使它因免费而被消灭。第一编处理前半句——那段时间是怎么被省略的，为什么必须被省略，以及这一省略如何成为一套自我确证的体系。

第一章 三个动作：搬运、传道、接受

教学不是知识的传递，而是一条由三个动作构成的单向搬运链；这条链的方向，两千年没有变过。

一、先把动作数清楚

讨论教育改革的困难，往往不在于理念之争，而在于没有人先把“教学”这件事拆成可数的动作。理念可以争论一百年而不落地，动作却是可以清点的：一个教师在一节课里，究竟做了哪几件事？

本书的回答是：可以穷尽地归为三个。

第一个动作是搬运。把已经被人类确认过的知识，从它被产生的地方——实验室、典籍、学术共同体、行业经验——搬到课堂里来。这个动作不是简单的位移。知识在它的产生地是一种形态，在课堂里必须是另一种形态：它必须被剥去产生它的具体情境，剥去产生它那个人的个人痕迹，压缩成一种可携带、可复制、可检验的东西。教材、教学大纲、课程标准，全部是搬运的产物。

第二个动作是传道。把搬来的成品，用尽可能清晰的方式讲出来。讲解的全部技艺——举例、类比、拆步骤、由浅入深、先讲特例再讲一般、变式训练、错例辨析——都服务于同一个目标：让接收端的损耗尽可能小。一个好教师在这条路上的定义是明确的、可比较的、也是可以训练的：**损耗最小的那个传输者。**

第三个动作是接受。学生把讲出来的东西接住、记住、能用。检验接受是否完成的方式同样明确：让他复述一遍，让他做一道同型的题，让他在一个新的题面下把同一个方法再用一次。凡是能通过这三关的，就叫“学会了”。

三个动作构成一条完整的链。而本章要立住的判断是：**这条链是单向的。**

二、单向性意味着什么

知识从上游流向下游。上游是已经确认的知识，中游是教师，下游是学生。教师是中继站，学生是终端。整条链的效率取决于三件事：搬得全不全、讲得清不清、接得住接不住。

单向性有三个直接后果，每一个都在今天的课堂里随处可见。

第一，判断权集中在上游。什么是对的、什么是重要的、什么应当先学、什么可以不学——这些判断在知识进入课堂之前就已经做完了。教师不做这些判断，教师执行这些判断。学生更不做。所以课堂上真正被生产出来的东西，不是判断，是**判断的副本**。

第二，终点先于起点存在。一堂课开始之前，它的终点已经写在教案上。整节课的全部设计，都是为了让全班在下课铃响之前抵达那个点。这一条在后面的第六章里会被单独处理，因为它是识别一堂课属于哪条路的唯一可靠判据。

第三，失败被定义为损耗。在这条链上，出问题只有一种形态：某个环节损耗过大。孩子没学会，要么是教材编得不好（搬运出问题），要么是讲得不清楚（传道出问题），要么是他没认真听（接受出问题）。**注意这三种归因有一个共同点：它们都不质疑链条本身。**一个系统如果它的全部故障模式都指向内部环节而从不指向自身结构，那么这个系统在结构上是不可证伪的。

三、三个动作各自的技艺高度

必须立刻说明：指出这条链是单向的，不等于说它是粗糙的。恰恰相反，这三个动作各自都发展出了极高的技艺，其中相当一部分是人类文明真正的成就。

搬运这一环上，从教科书的编写体例，到概念呈现的先后次序，到例题与习题的难度梯度，到不同学段之间的衔接，其中凝结的经验以世纪计。一套好教材与一套坏教材的差别，任何一位教过书的人都能立刻感觉到，而这个差别几乎无法用一句话说清——那正是高度技艺的标志。

传道这一环上更是如此。一个特级教师能在四十分钟里，把一个抽象概念拆成十二个可感知的台阶，每一级台阶都恰好落在多数孩子够得着的地方；他知道哪一句话说早了会关掉思考，哪一个比喻在三年之后会变成障碍，哪一处停顿比再讲一遍更有效。这些东西不写在任何一本书上，只能靠三十年积累。

接受这一环上，也有大量真实的技艺：如何设计一次提问才能测出真的懂而不是假的懂，如何布置作业才不至于变成机械重复，如何在错题里读出思路而不只是读出对错。

本书对老路的全部批评，都建立在承认这些技艺为真的基础上。任何一种把过去两千年说成一场糊涂的论述，都不值得反驳，因为它连要反驳的对象是什么都没有看清。

四、为什么三个动作可以被同时接管

本书后面几编要处理的那件事，在这里可以先点一句，因为它解释了为什么必须先把动作数清楚。

这三个动作有一个共同的性质：**它们全部作用于成品。**搬运搬的是成品，传道讲的是成品，接受接的是成品。链条上从头到尾流动的，是同一样东西——一个已经完成的、被压缩过的、可复制的结果。

任何一件全部作用于可复制物的工作，原则上都可以被一台足够强的复制与改写装置接管。这不是关于人工智能的特殊断言，这是关于这条链的结构性事实。第九至十一章将逐个动作说明这件事已经发生了。

五、三个动作的历史变体

有必要说明，这三个动作在不同文明、不同时代呈现出很不一样的外观，而结构不变。这一点若不交代，“两千年”这个说法会显得草率。

在师徒制里，搬运表现为师傅把行业中已被确认的做法带到作坊；传道表现为示范与口诀；接受表现为学徒照做，直到做出来的东西与师傅的一样。这里没有教科书，没有课堂，没有考试，但三个动作齐全，方向相同——终点是师傅手里那个成品的复制品。学徒偶尔做出一件师傅从未做过的东西，那不属于这个体系的设计，属于它的溢出。

在经典教育里，搬运表现为文本的选定与注疏的确立；传道表现为讲解与背诵；接受表现为能诵、能解、能用典。这套体系的技艺极高，注疏之细密远超今日任何教材。而方向同样单一：文本在前，学生在后。

在近代班级授课制里，三个动作第一次被制度化到可管理的程度：课程标准规定搬什么，教学法规定怎么讲，考试规定接住没有。这不是新链条，是**同一条链的工业化版本**——它的成就在于把边际成本压到最低，代价在第七章。

在今天的信息化课堂里，三个动作被技术分别加强：搬运有了数字资源库，传道有了多媒体与微课，接受有了自动批改与错题本。**每一项技术都被用在同一条链的同一个方向上。**

四个变体的外观差别极大，结构完全相同。这就是“两千年没换过方向”这句话的确切含义——不是说什么都没变，而是说**变的全在环节内部，没有一次变在方向上。**

六、一个反例的检验：苏格拉底

会有人举出反例：苏格拉底的诘问法，不正是不给答案吗？

这个反例值得认真对待，因为它是历史上离本书主张最近的一个。诘问法的动作确实是反向的：不告知，只追问；不填补，只戳破；把对方自以为知道的东西一点点问到无法自圆其说。

但需要看第二步。经典对话中，被诘问者陷入困惑之后，接下来发生什么？在相当一部分对话里，接下来是苏格拉底本人引导他一步步走向一个**苏格拉底已经知道的结论**。最著名的那个几何段落——引导一个未受教育的少年“回忆”起正方形加倍的作法——每一步提问都指向一个既定终点。

所以诘问法是一个**混合体**：前半段真的在张开岔口，后半段仍然是抵达。这个混合恰恰说明了本书判据的价值——**动作看上去完全一样，判据落在终点上。**

而必须公正地补充：诘问法的前半段，在两千年里几乎没有被制度化地继承下来。被继承的是“启发式提问”，也就是后半段。原因还是第三章那笔账——**前半段是一对一的、慢的、不可批量的。**

七、可观察的动作

每一章末尾，本书都给出一个可以当堂观察、不需要任何设备的动作或读数。本章的是：

数一数，一节课里有多少分钟，教室里流动的是成品；有多少分钟，流动的是一个还没有成品的问题。

绝大多数常规课堂的第二个数字接近于零。这不是批评，这是这条链正常运转时应有的样子——**它本来就是为搬运成品而设计的。**

八、与相邻说法的分界

本章的“三个动作”与教学论中常见的“教学环节论”（导入、新授、巩固、作业）形状相似，但切分的依据不同：环节论按时间顺序切分一节课，本书按**作用对象**切分——三个动作作用的都是成品，而这正是它们可被同时接管的原因。环节论不关心这一点，因为在它成立的年代，还没有任何东西可以接管它们。

第二章 光滑化：把走过的路磨光了发下去

教科书上的知识不是知识本身，而是一次已完成的发生被磨去土壤与路径之后的残留物；它显得先在，是因为它的来路被删掉了。

一、一个减法

第一章说，这条链上流动的是成品。本章要问：成品是怎么造出来的？

任何一件知识在它被产生的时候，都包含三样东西：

最后显露出来的那个结果——定理、公式、结论、方法。这是唯一被写进教科书的部分。

通往它的那条路径——当年的人是怎么一步步试错走过来的，先试了什么，为什么放弃，在哪里卡了十年，哪一次绕远了又退回来。

它在其中长出来的那片土壤——这个问题当年为什么会被提出来，是什么困难逼出了它，当时的人相信什么、不相信什么，什么在当时是显然的、什么是不可想象的。

把后面两样减掉，只留下第一样，再把它摆到最前面当作起点端出来——这个操作，本书称为**光滑化**。

写成一句便于记忆的话：

发现 = 发生 - 路径 - 土壤

需要说明“光滑”这个词的来处。一件被磨光的东西，摸上去没有毛刺、没有接缝、没有加工痕迹，因而看不出它是被做出来的，看上去像是本来就是这个样子。教科书上的概念正是这样：它们没有毛刺。而毛刺——那些没走通的路、被放弃的定义、当年吵得不可开交的地方——恰恰是“这是被人做出来的”这件事唯一的证据。

二、光滑化的三步

这个操作可以拆成三个可分辨的动作，后面第十六至十七章讨论如何还原时，会一一对应地倒过来做。

第一步，抹去土壤。删掉这个知识赖以产生的那个困难。今天的教科书上，无理数是一个定义，而不是一个持续了两千年的丑闻；负数是一条运算规则，而不是几代数学家拒绝承认的“荒谬的数”。

第二步，抹去路径。删掉试错。今天呈现的证明是最简洁的那一个，而不是历史上第一个被找到的那一个；今天的公式排列顺序是逻辑顺序，而不是发生顺序。**逻辑顺序几乎总是发生顺序的逆序**——人们先摸到了结果，很久以后才找到能把它安放好的那个起点，而教科书从那个起点讲起。

第三步，把结果摆到最前面。这一步最关键，也最隐蔽。删掉前两样之后，剩下的那个结果不再是终点，它被重新安置为**起点**。定义在前，性质在后，例题居中，习题殿后。于是学习者遇到它的第一个形态，就是“给定这个东西”。

三步做完，一件被人做出来的东西，看上去成了本来就在那里的东西。

三、这不是编书人偷懒

必须立刻堵住一个廉价的批评。光滑化不是疏忽，不是懒惰，也不是某种知识霸权的阴谋。不压缩，就搬不动。

设想一个不做光滑化的教科书：它要完整记录每一个概念的产生过程，包括所有走不通的路。这样一本“小学数学”会有几万页，而且其中绝大部分内容对一个八岁的孩子毫无意义——因为那些错误路径之所以有意义，是对当年那些已经具备大量背景的成年人而言的。

所以光滑化是一项了不起的发明，第七章将专门论证它的三大成就。本书对它的态度必须精确，这一句在全书中会反复出现：

光滑化是必要的。问题在于只有光滑化。

四、可传递的与不可传递的

由光滑化推出一个不容易被接受、但很难反驳的结论。

可以打包递送的，只有成品。路径和土壤不能被递送——因为它们不是东西，是过程。一个教师不可能把自己当年“想通那一下”直接装进学生脑子里，正如不可能把自己走过的路直接装进别人的腿里。他能给的只有描述，而对过程的描述本身，也是一个成品。

光滑成品是可以传递的；发生是不可以传递的。

这一条对教学动作的性质有直接后果：教学不可能是“传递”。它只能是引发——制造条件，让一次发生在学习者身上重新出现。这句话在教育里被说过很多次，通常作为一种理念或口号；本书要指出的是，它不是理念，是从光滑化这个操作的性质里推出来的一个结构性事实。

五、光滑化的两冲程

如果发生不可传递，而知识确实跨代传下来了，那么中间必然还有一个环节。

完整的图景是两个冲程：

一次发生 → 光滑化（压缩成可搬运的成品）→ 跨越人与人的边界 → 在新的人身上重新发生（解压）

第一个冲程是文明用于搬运的压缩，第二个冲程是教育用于落进一个人的解压。任何知识要跨代存活，两个冲程都必须跑完。

而人类文明把第一个冲程做到了极致——教科书、课程标准、考试体系、教师培训，全部服务于压缩与投递的质量；第二个冲程则基本上留给了运气。运气好的那些孩子，自己在某个时刻把它解压了，我们称他们有天赋、有悟性、会学习；运气不好的，接住了成品但从未解压，我们称他们基础扎实但不会灵活运用。

这两个说法，说的是同一件事的两面。

六、一个要点：解压不是逆运算

关于第二个冲程，有一处必须现在讲清楚，否则后面第十六章会被误读。

解压不是把光滑化倒过来做。它不还原原来的路径与土壤——学习者不可能重走历史上那条路，也不生活在当年那片土壤里。他代入的是**他自己的**路径与土壤：他自己的困惑、他自己已有的经验、他自己走出来的那条弯路。

所以这个环是**有损的**。而且正因为有损，它才是生成性的：如果解压能完美还原原件，那么每一代人得到的都是同一个东西，知识就不会生长。**知识之所以会生长，正因为每一次解压都长得不一样。**

这一点在第五编讨论新路时会被再次用到：教师不需要、也不可能让学生重走原路。他只需要让学生走出一条**他自己的**路。

七、光滑化的三个副产品

除了第四节的主结论，光滑化还产生三个副产品，它们在后面几编里都会被用到，这里先建立。

副产品一：先在感。成品被摆到最前面，学习者遇到它的第一个形态就是“给定这个东西”。反复如此之后，会形成一种感觉——这些东西本来就在那里，不因任何人而改变。这种感觉在数学上表现得最强烈，强烈到成为一种哲学立场：许多数学工作者真诚地报告，他们感到自己是在**发现**而不是在**制造**。

本书对这个报告的态度不是反驳，而是收编：**这份现象学报告是准确的，它准确地记录了光滑化在接收端造成的效果。**一个人如果只见过成品，那么他对成品的感受必然是“它本来就在”。第十七章将说明，这种感受不是被动残留，而是被主动写下来的。

副产品二：逻辑顺序的僭越。光滑化之后的知识按逻辑顺序排列，而逻辑顺序几乎总是发生顺序的逆序。人们先摸到结果，很久以后才找到能把它安放好的起点。教科书从那个起点讲起，于是学习者的遭遇顺序与人类的遭遇顺序恰好相反。这个倒置本身不是错误——按逻辑顺序讲效率更高——但它有一个不被注意的后果：**学习者永远看不到这个概念被需要的那一刻。**他看到的是它被安放好之后的样子。

副产品三：难度的错觉。被磨光的东西看上去简单。一个花了两百年才被接受的概念，在教科书上占半页。于是接不住的孩子会得到一个结论——这么简单我都不会，是我笨。这个结论对他的伤害，往往比没学会那个概念本身更大，而且更持久。**光滑化把历史上的两百年，变成了他个人的一次失败。**

八、一个必要的界定：光滑化不等于简化

需要把光滑化与教学上常说的“简化”区分开，否则本书的批评会被误读为“反对把知识讲简单”。

简化是在**保留来路**的前提下降低难度：用更浅的语言、更小的步子、更熟悉的例子说同一件事。简化之后，那个概念仍然连着它的困难。

光滑化是**删掉来路**：不是把困难讲浅，而是不提困难。

两者可以独立变化。一段讲解可以既不简化又高度光滑（大学教材的常态：难，且看不出来路），也可以既简化又不光滑（一个好教师讲的故事：浅，但保留了“当年为什么需要它”）。

本书反对的从来不是简化。反对的是把“没有来路”当成默认设置。

九、可观察的动作

随手翻开任何一册教材的任何一页，问一句：这一页上写的东西，当年是被谁、因为遇到了什么困难、才做出来的？

如果这一页上没有任何一句话回答了这个问题——通常没有——那么你还正看着一个被磨光的成品。这个动作只需要三十秒，但它是本书全部论述的经验起点。

十、与相邻说法的分界

教育学中有一个成熟的概念描述与本章高度重合的现象：知识为了能被教，必须先被去情境化、去个人化。本章在描述层面与之一致，但问题不同：那一传统关心的是**知识形态在制度链条上的变形**，本书关心的是**这一变形在接收者身上造成了什么**。分界线是：一个研究被搬运的东西，一个研究接不住搬运物的人。第十七章将说明，这个差别不是侧重点的差别，而是能否解释“补课为什么无效”的差别。

第三章 成本论：一个孩子三千小时

发现学教育不是一个错误的选择，而是唯一负担得起的选择；批评它扼杀创造力，等于批评一个买不起房的人为什么租房住。

一、一个无力的批评

对传统教学最流行的批评是“它扼杀创造力”。这个批评在道德上很痛快，在解释上却近乎无效——因为它无法回答一个基本问题：

为什么一件如此显然有害的事情，会被全世界所有文明、在完全不同的政治制度与文化传统下、连续采用两千年？

如果答案是“因为统治者需要顺从的臣民”，那么它解释不了为什么在那些不需要顺从臣民的地方——学徒制作坊、家族技艺传承、宗教团体内部的经典教育——同样的结构照样出现。如果答案是“因为教师懒惰”，那么它解释不了为什么最勤奋、最有理想的教师同样在这条路上。

一个持续两千年、跨越所有制度的现象，必须有一个结构性的解释。本章给出的解释是**成本**。

二、把账算出来

设想一种理想的教学：一个孩子面对一个尚未被判定的问题，自己走过全部的岔口——试错、走错、退回、重来，最后自己做出那个判断。这种方式得到的东西质量极高，因为它连起来路、连着理由、连着“本来可以不是这样”的那些被排除掉的可能。

现在给它标价。

一个核心概念的完整发生，需要多少一对一的陪伴时间？保守估计是十小时。这个数字不是凭空来的：任何一位带过研究生的人都知道，让一个人真正想通一个不平凡的概念，一个下午远远不够；而基础教育里的核心概念，虽然在成年人看来简单，对一个初次遭遇它的孩子而言，其陌生程度并不低于研究生面对一个新领域。

基础教育阶段的核心概念有多少个？以数学为例，从数、位值、四则运算、分数、小数、比例、方程、函数、极限，到几何的公理化、概率的可加性，能称得上“核心”的大约在一百到一百五十之间。加上语文、科学、外语、社会各科，三百是一个偏保守的估计。

于是：

| 一个孩子 = 300 个概念 × 10 小时 = **3,000 小时的一对一时间**

一名教师每周能提供多少小时真正有效的一对一？考虑到备课、休息、注意力的极限，二十小时已是上限。一年按四十周计，八百小时；一生按三十五年计，两万八千小时。

| 一名教师终身能服务的学生数 = $28,000 \div 3,000 \approx$ **不足十人**

这个数字就是全部问题的答案。任何一个需要普及教育的社会——也就是任何一个现代社会——都不可能采用这个方案。它甚至不是“太贵”，它是在数量级上不可能。

三、于是有了那个替代方案

面对这个约束，文明想出的办法就是第二章讲的光滑化：**把已经走完的路磨光，只把结果发下去。**

这个替代方案的经济性质值得说清楚。光滑化把一件**不可复制的过程**转换成了一件可以**无限复制的产品**。而一旦某样东西可以被无限复制，它的边际成本就趋近于零——一个教师面对四十人和面对四百人，讲同一段内容的成本几乎相同。

于是三千小时的一对一，被压缩成了每个孩子一万两千个课时的班级授课。质量下降了很多，成本下降了几个数量级。**这不是一个糟糕的交易，这是一个在当时唯一可行的交易。**

四、成本约束如何塑造了教育的全部形态

一旦承认成本是第一约束，教育制度中许多看似不合理的设计，就立刻变得可理解了。

为什么是班级授课制？ 因为它是把可复制物投递给最多人的最经济结构。

为什么有统一进度？ 因为个别化的进度会破坏可复制性，成本立刻回升。

为什么有标准化考试？ 因为要检验投递质量，而唯一低成本的检验方式，是检验成品是否被接住——也就是检验复制的保真度。

为什么教师被要求“讲得清楚”而不是“忍住不讲”？ 因为在供给稀缺的条件下，忍住不讲是纯粹的浪费：孩子卡住的那段时间里，教师无法为其余三十九人提供任何服务。

为什么“因材施教”两千年来一直是理想而非现实？ 因为它要求把成本重新推回到一对一那一侧。

这些设计合在一起，构成了一个高度自治的系统。而这个系统的每一条，都可以从“三千小时买不起”这一句推出来。

五、约束还在吗

本书的全部转折都压在这一问上。

如果这个成本约束仍然存在，那么本书后面五编的主张就只是一厢情愿——再正确也没有用，因为负担不起。教育史上不缺正确而昂贵的主张，它们全部死于同一个原因。

但约束松动了，而且是从一个出人意料的方向松动的：**它首先是从老路那一侧被解除的。**

大模型第一个做到的事情，不是陪孩子走岔口，而是把搬运、传道、接受这三个动作全部做完。这就产生了一个奇特的局面——成本结构改变了，但改变的方向首先是让老路变得更便宜、更好用、更无懈可击。第二编将处理这一点。

而在这个局面之外，还有一件更重要的事：那三千小时的一对一陪伴，第一次在原理上变得可负担了。不是因为教师变多了，是因为**陪练可以不是人。**

这一句是本书全部乐观论述的唯一根据，也是第五编全部处方的前提。但它有一个致命的前提条件：这个陪练必须被配置成一个**不给答案**的陪练。而它的默认构造恰好相反。第十四章将说明这个反差有多危险。

六、一句必须先讲的公道话

在任何面向教师的场合，下面这句话必须在批评之前讲出来，而且要讲完整：

发现学教育不是错的。它是人类在负担不起一对一走岔口的条件下，能找到的最好方案。它让知识跨代存活，让每一代人不必从零开始。它完成了它的历史任务。

不讲这句话而直接进入批评，会造成一个可预测的后果：会场上最优秀、最投入的那批教师会认为自己三十年的工作被否定，然后本书后面的全部内容对他们失效。而这批人恰恰是唯一有能力执行第五编那些动作的人。

这不是修辞策略，这是判断的准确性问题。一个把过去说成错误的诊断，本身就是错的诊断。

七、这笔账里被忽略的一项：等待的价值

上面的算法有一个隐含前提——一对一时间的价值全部来自“教师在场做了什么”。这个前提需要修正，而修正之后，账会变得对本书的主张更有利。

在一对一的场景里，教师最有价值的动作之一是什么都不做：孩子卡住了，他等着。这段等待不消耗教师的技艺，却是整段陪伴中最贵的部分——因为它是孩子唯一必须自己产出判断的时刻。

而在班级授课里，等待是不可能的。不是教师不愿意等，是**等待在四十人的场景里是纯粹的浪费**：一个孩子卡住的那三分钟，其余三十九人无事可做。于是理性的选择必然是给出答案，把全班拉回同步。

这就得到一个对全书都重要的判断：

班级授课制的核心约束不是“讲不细”，而是“等不起”。

这个判断比“班额过大影响个别化”要精确得多，也更有后果。因为它指出：即使把班额从四十减到二十，等待的成本也只是减半，不会消失；只有当陪练的边际成本趋近于零时，等待才第一次成为可负担的。

而这正是 AI 在原理上带来的那件事。**AI 真正的教育价值，不在于它讲得好，而在于它可以不计成本地陪着一个孩子卡住。**——如果它被配置成愿意这么做的话。第二编将说明，它的默认配置恰好相反。

八、约束解除之后，什么会自动发生，什么不会

成本约束的解除不会自动带来任何好处。这一点必须写清楚，否则本书会变成一份技术乐观主义的宣传。

会自动发生的：光滑成品的供给量、投递速度、个性化程度，全部大幅提升。因为这三项正是这项技术的直接能力，不需要任何人做任何设计。

不会自动发生的：等待的时间被保留下来。因为这需要一个**反向**的设计——需要有人规定在某段时间内不提供供给。而无论是技术的构造、产品的商业模式，还是使用者的即时偏好，全部指向相反的方向。

于是得到本书对整个局面的基本判断，它将在第二编被反复展开：

成本约束的解除，首先解除的是老路的成本约束。

九、可观察的动作

给自己算一次账：如果只带一个学生，你能把一个概念教到什么程度？和你现在带四十个人时相比，差多少？

多数教师算完这道题会沉默一会儿。那个差值，就是成本约束在这两千年里向教育收取的税。而本书要论证的是：这笔税第一次有可能被退回来一部分。

十、与相邻说法的分界

教育经济学中有大量关于班级规模、师生比与成本效益的研究，本章的算术与它们方向一致但用途不同：那些研究用成本来论证何种投入配置最优，本书用成本来解释**一种教学形态为什么在历史上必然出现**，并据此论证当成本结构改变时，这一形态的必然性随之消失。前者是政策优化，后者是形态的发生学。

第四章 教科书是怎么写成的：四个标本

一个概念在教科书上的第一句话，通常正好删掉了它当年赖以产生的那个困难；孩子遇到的不是这个概念，而是它的遗体。

一、为什么要做标本

第二章给出了光滑化的机制，本章给出证据。方法是解剖：取四个小学与中学都会遇到的核心概念，把教科书上的呈现方式与它的历史发生过程并排放在一起，看被删掉的是什么。

选这四个标本有一个共同理由——它们都是“公认简单”的概念。用高深概念做标本没有说服力，因为可以推说“那是专业内容，本来就该简化”。**恰恰是最简单的那些概念，被删得最干净。**

二、标本一：负数

教科书怎么写。通常是这样开始的：把数轴向左延长，零左边的数叫负数；然后给出运算规则，其中“负负得正”作为一条规则被给出，有的版本会用“欠债的反面”作一个类比。

当年发生了什么。负数在数学史上的处境远比这狼狈。它在很长时间里不被承认是数——理由并不愚蠢：数是用来数东西的，而没有“负三只羊”这回事。十六、十七世纪的欧洲数学家把负根称为“虚假的根”，有人明确写下负数“小于无”因而不可理解。真正让负数取得户口的，不是有人证明了它存在，而是**方程解不出来这件事变得越来越难以忍受**——如果拒绝负数，一大批本来结构完全相同的方程就必须被分成好几类分别处理，整个体系变得丑陋不堪。

被删掉的是什么。被删掉的是那场持续几百年的拒绝，以及最终让步的理由——**不是因为负数被发现真的存在，而是因为承认它比不承认它更方便。**这句话如果讲给孩子听，他会第一次意识到：数学里有些东西是被规定的，而规定是有理由的，理由是可以被讨论的。

而现在他学到的是：负数在那里，请接受。

三、标本二：分数

教科书怎么写。把一个整体平均分成若干份，表示其中一份或几份的数，叫分数。接着是分数的读法、写法、大小比较、四则运算。

当年发生了什么。分数的产生动机是分不尽——分东西分不尽、量长度量不尽。而这里有一个几乎从不被提起的关键点：**分数最初不是“数”，而是“一对数之间的关系”。**这个区别不是学究气的，它正是孩子最常见的那个错误的根源。

那个错误几乎每一个数学教师都遇到过：问三分之一和二分之一哪个大，很大一部分孩子答三分之一，理由是三比二大。这个回答通常被当作粗心或概念不清而纠正掉。但它其实是一个**完全合理的错误**——孩子把分数当成了两个数，而在他此前的全部经验里，比较两个数就是比较它们的大小。他没有搞错，他是在用旧规则处理一个新对象，而没有人告诉过他这是一个新对象。

被删掉的是什么。被删掉的是“这是一个新种类的东西”这件事本身。教科书直接给出定义，跳过了“为什么需要一个新种类”这一步，于是孩子无从知道旧规则在这里失效了。

四、标本三：无理数

教科书怎么写。无限不循环小数叫无理数。举例：根号二、圆周率。

当年发生了什么。这是数学史上第一次真正的危机。一个把“万物皆数”作为世界观基础的学派，发现正方形的对角线与边长之间**不存在任何公度**——不存在任何一个单位，能同时量尽

这两条线。这不是一个计算问题，这是一个信念崩塌。传说有人因泄露此事而丧命，传说未必属实，但它反映了当时的震动量级。而这个问题的彻底解决——给无理数一个严格的定义——发生在两千多年之后。

被删掉的是什么。被删掉的是“量不尽”这个具体的、可以当场做出来的困难。今天完全可以让孩子拿一把尺子去量正方形的对角线，量到他确信这件事没有尽头。这个动作花五分钟，而它给出的东西是“无限不循环小数”这七个字永远给不出的。

五、标本四：三角形内角和

教科书怎么写。三角形内角和等于一百八十度。配一个撕角拼直线的活动，或者量三个三角形验证一下。

当年发生了什么。这条结论的成立，依赖一条被称为平行公设的前提，而这条前提两千年来一直被怀疑是多余的——无数人试图从其余公理把它证出来，全部失败。最后的解决方式极其反常：不是有人证明了它，而是有人**假设它不成立**，结果得到了一套完全自治的、后来被物理学用上的几何。

被删掉的是什么。被删掉的是“这条结论有一个前提”。孩子撕了三个角，拼成一条直线，得到的结论是“所有三角形都是一百八十度”。但他撕了三个，凭什么说所有？这一步跳跃从未被追问，因为教科书上的结论已经先给出了，撕角只是在**验证一个已知答案**，而不是在**产出一个未知答案**。

这里还有一个更隐蔽的删除：撕角拼直线之所以成立，靠的是纸是平的、三个角挪动时不变形。这个前提从来不被说出来，因为在孩子的经验里纸当然是平的。而正是这个“当然”，让一百八十度看上去成了三角形的性质。

它是平面的性质。

六、四个标本的共同结构

把四个并排放在一起，一个共同结构浮现出来：

概念	当年的困难	教科书的第一句	被删掉的东西
负数	方程解不出来；数不该是负的	零左边的数	承认它是一次让步，不是一次发现
分数	分不尽、量不尽	平均分成若干份	这是一个新种类的对象
无理数	对角线量不尽，量了两千年	无限不循环小数	那个可以当场做出来的困难
内角和	一条无法证明的前提	等于一百八十度	结论有前提，前提可以不成立

四行合起来是一句话：**教科书的第一句，通常正好是那个困难被删掉之后剩下的东西。**

而更值得注意的是最后一列。被删掉的不是细节，也不是趣味性的历史故事——被删掉的每一样，都是“这件事本来可以不是这样”的证据。删完之后，剩下的那个东西看上去只能是这样。

七、一个可预测的反驳

会有人说：这些历史内容当然重要，但课时有限，不可能什么都讲。

这个反驳是对的，而且本书接受它。本章的主张不是“要给孩子讲数学史”——把数学史作为额外内容塞进课堂，只是增加了一批新的成品，不改变任何结构。

本章的主张要弱得多，也硬得多：**至少让孩子知道，这里曾经有过一个困难**。这通常只需要一句话、一个问题、五分钟。第二十九章将说明，合格的问题恰恰不需要额外课时，因为它替换的是原本用来讲解答案的那几分钟。

八、第五个标本：等号

前面四个标本都在数学的对象一侧。这里补一个不在对象一侧的标本，因为它说明光滑化不只删掉概念的来路，也删掉**符号的来路**。

教科书怎么写。等号几乎从不被定义。它在算术的第一课就出现，作为“算出结果”的信号： $3+4=$ ，后面写答案。

实际发生的是什么。等号在数学里的含义是“两边是同一个东西”，是一个对称的关系。而孩子在头两年里遇到的全部用法，都是不对称的：左边是过程，右边是结果。于是他建立起来的理解是“=的意思是算出来”。

这个理解在很长时间内不出问题，直到方程出现。 $3+x=7$ ，这里等号两边都不是“算出来”的方向；而 $x+3=7$ 与 $7=x+3$ 是同一件事，这在他的旧理解里说不通。大量学生在方程上的困难，根源不在方程，在两年前那个从未被说出口的等号含义。

被删掉的是什么。被删掉的是“等号有一个含义，而且这个含义与你现在用它的方式不完全一致”。这一句在算术阶段说出来只需要一分钟，而不说的代价要在三年后支付。

这个标本还说明了一件比前四个更普遍的事：**光滑化最深的一层，往往不在被讲的内容里，而在从未被当成内容的那些东西里**。符号、记法、书写顺序、题目的问法——它们不出现在任何一课的教学目标里，因而从未被追问过来路。

九、标本的用法：不是讲历史，是开一道口

本章最容易被误用的方式，是把这些内容当作补充材料塞进课堂——讲一段数学史，讲一个数学家的故事，然后回到原来的讲法。

这样做没有害处，但也没有作用。原因在第二章：**关于困难的描述，本身也是一个成品**。把“无理数的发现引起了危机”作为一句话讲给孩子听，与把“无理数是无限不循环小数”讲给他听，在结构上完全相同——都是把一个已完成的東西端出去。

标本的正确用法是**把那个困难重新做出来**：不是讲“当年量不尽”，而是给他一把尺子去量；不是讲“当年不承认负数”，而是问他“如果我们不许有负数，这道题该怎么办”；不是讲“内角和有前提”，而是问他“你撕了三个，凭什么说所有”。

区别在于：前者他得到一个新的成品，后者他遇到一个真的困难。**做得出来的困难才是岔口，讲得出来的困难只是新的一课**。

第二十九章将给出各科挑选这类困难的方法。这里先给一条便捷的筛选规则：**如果这个困难可以在五分钟内用手边的东西做出来，它就是好的；如果只能讲，就先放下**。

十、可观察的动作

在你教的这一科里，挑三个“最简单、大家都会”的概念，各花十分钟查一查：它当年是从哪个困难里长出来的？

多数教师做完这件事之后的反应是相似的：原来我教了二十年的这个东西，我自己也只见过它的遗体。这不是能力问题——没有人教过我们去看它的活体，因为教材上没有。

十一、与相邻说法的分界

数学教育中有一个成熟的做法：从现象出发重新组织教学内容，让学生在具体情境中重新构造数学对象。本章的四个标本与那一做法在材料上重合，但目的不同：那一传统关心的是如何设计更好的学习路径，本章关心的是教科书上的呈现方式本身在传递一条隐含判断——对象是现成的。第十七章将说明，这条隐含判断的累积效果，比任何一次路径设计的成败都重要。

第五章 为什么两千年没换过方向

这条链之所以两千年不变，不是因为有人捍卫它，而是因为它的每一环都可以被单独优化，且每一步优化都立刻在读数上兑现——它不需要有人相信它，只需要有人测量它。

一、一个需要解释的稳定性

第一章说这条链是单向的，第三章说它源于成本。但成本只解释了它为什么会出现，没有解释它为什么如此稳定。

两千年里，教育领域从不缺改革者，而且不乏极有分量的改革者。从苏格拉底式的对话，到中世纪大学的辩论传统，到近代的直观教学、儿童中心、做中学、发现学习、探究式学习、项目式学习、翻转课堂、分层走班、核心素养——每一轮都提出过深刻的批评，每一轮都有真诚的实践者，每一轮都留下了真实的成果。

而这条链没有变过方向。

这个稳定性本身需要解释，而且不能用“体制惯性”这类空话打发。惯性只是给现象换了个名字。

二、可逐环优化的系统是极其稳固的

本书给出的解释是：这条链的三个环节可以被独立优化，而且每一步优化都会立刻在可测量的指标上兑现。

搬得更全 → 教材更完整 → 覆盖率上升。

讲得更清 → 当堂正确率提高。

接得更牢 → 考试成绩提升。

三条路径互相独立，任何一位教师、任何一所学校、任何一位教研员，都可以只做其中一项而立刻看到回报。不需要说服任何人，不需要改变任何观念，不需要等待任何配套条件。

一个具备这种性质的系统，在结构上是极难被替换的。因为替换它需要**整体行动**，而优化它只需要**局部行动**；而局部行动的回报是即时的、可见的、可归因的，整体行动的回报是延迟的、模糊的、无法归因的。

于是有了那句在本书中会被反复引用的话：

它不需要有人相信它，它只需要有人测量它。

三、教改的实际命运

理解了这一点，历次教育改革的实际命运就不再令人困惑了，而且是可以被预测的。

一轮改革推行之后，通常经过一到两个学期，会发生一次自动筛选：**那些能在现行读数上兑现的做法被保留，其余部分被悄悄放弃。**

这个筛选不是任何人的决定，没有会议纪要，没有反对者。它是每一位教师在自己的课堂上、根据自己看到的回报，各自独立做出的理性选择的总和。

于是我们看到一个反复出现的模式：

探究式学习被推行之后，留下来的通常是“用探究的方式导入新课”——因为这一段能提高注意力，从而提高接受效率。而“让学生得出与教案不同的结论”这一部分被放弃了，因为它在读数上是负的。

项目式学习被推行之后，留下来的通常是“用项目形式包装一组已定的知识点”。而“项目的结果事先不知道”这一条被放弃了。

翻转课堂被推行之后，留下来的是“课前看视频、课上做题”——这提高了课堂时间的利用率。而“课上用于处理无法预先设计的困惑”这一条被放弃了。

分层教学被推行之后，留下来的是按成绩分班。

每一次，**存活下来的都是能提高接受效率的那一部分；被淘汰的都是会改变链条方向的那一部分。**

这不是任何人的失职。这是测量系统的选择结果。

四、这个机制对本书的直接含义

如果上述分析成立，那么本书第五编提出的那套做法，将面对完全相同的筛选压力，而且没有任何理由认为它会幸免。

可以预先写出它被筛选之后会剩下什么：

留下来的会是“课堂上加一个提问环节”“让 AI 出几个反例活跃气氛”“每节课最后让学生说说收获”。

被放弃的会是“十分钟内谁都不许给答案”“接受这节课没有得出结论”“让学生用旧账去顶 AI 到顶不动为止”。

留下来的那些无害而无用；被放弃的那些正是全部作用之所在。

这就是为什么第二十六章要单独讨论“秤”，并且把它放在与第五编所有具体做法同等重要的位置上。**不换秤而换做法，等于把新做法送进旧筛子。筛子会工作得非常有效率。**

五、一个更深的层次：读数塑造了“什么算学会”

还有一层比筛选更深。读数不只筛选做法，它还定义了目标本身。

问一位教师“这个孩子学会了吗”，他会怎么判断？他会看这孩子能不能复述、能不能做题能不能在新题面下再用一次。这三条正是第一章的“接受”。

也就是说，“学会了”这个词在实践中的含义，已经等同于“接住了”。这不是任何人的偷换，这是两千年测量实践的沉淀结果：因为只有“接住了”是可以低成本测量的，久而久之，可测量的那一项就成了这个词的定义。

后面第十九章讲“欺骗发生在名词上”时，会回到这里。那一章处理的是AI时代的宣传话术而本章要指出的是：**这个偷换不是AI带来的，它早就完成了。**AI只是让被偷换掉的那两件事第一次可以完全脱钩。

六、一个可以验证的预测

上述机制如果成立，它应当能对尚未发生的事做出预测，而不只是解释已经发生的事。本节给出三条，它们在任何一所正在推行改革的学校里都可以被检验。

预测一：存活率与读数兑现速度正相关。一项新做法能否在一年后存活，与它有多“先进”无关，与它在现行读数上兑现回报的速度强相关。当学期见效的存活；三年后才见效的，不存活。

预测二：改革的残留物形态可以被事先写出来。给定任何一项教育改革的完整方案，可以在推行之前就预测出一年后剩下哪些部分——凡是能提高接受效率的留下，凡是会让终点变得不确定的消失。这条预测的强度在于它是**事前的**。本书第五编的方案将被用来自我检验：第三十三节已经预先写出了它被筛选之后会剩下什么。

预测三：最投入的学校退回得最快。这一条反直觉。一所认真推行改革的学校，会更快、更彻底地完成筛选——因为它真的在测量，真的在复盘，真的按数据调整。而测量用的是旧秤。**认真本身加速了退回。**敷衍的学校反而可能保留一些无用但也无害的形式。

三条合起来说明：这里不存在一个可以被说服的对手。**筛子不是人，是读数。**

七、为什么“理念先行”总是失败

教育改革的常规做法是从理念开始：先做培训，讲清楚为什么要改，争取教师认同，然后推行做法。

按本章的分析，这个次序几乎注定失败，原因不在教师不认同，而在于**认同并不改变回报结构**。一位完全认同新理念的教师，回到教室之后仍然面对同一张成绩表、同一次月考、同一场家长会。他的认同是真的，他的处境也是真的，而处境每天兑现一次，理念一学期兑现一次。

这就得到本书对改革次序的一个明确主张，它与常规做法相反：

丨 **先换秤，再换做法；至少，秤和做法同时换。**

不是因为秤更重要，而是因为**做法在旧秤下活不过一个学期**。第二十六章将把这一条展开成可操作的方案，包括在无法改变期末评价的条件下，一所学校还能做什么。

八、一个真实的两难

必须承认，上面这条主张会遇到一个真实的两难：一所学校无权改变升学考试，而升学考试是所有秤中最重的一把。

本书不提供绕开这个两难的办法，因为不存在这样的办法。能提供的只有一个事实判断：**升学考试测的是老路，而老路的能力是新路能力的子集**。走过新路的孩子接成品更快，因为他知道成品是怎么来的；反过来不成立。第二十五章将论证这个不对称性，第二十七章将处理过渡期——那几周成绩确实会难看，而这几周正是绝大多数尝试死掉的地方。

九、可观察的读数

回想你所在学校最近一次教学改革。一年之后，还在做的是哪几件？停掉的是哪几件？

把这两组并排写下来，然后对每一件问一句：它是提高了接受效率，还是改变了链条方向？这个练习的结果在绝大多数学校里高度一致，而且不需要任何理论就能自己看明白。

十、与相邻说法的分界

组织研究中有一个熟悉的判断：指标一旦被用于考核，就不再是好的指标；也有研究指出，测量行为本身会改变被测量者的行为。本章与这些判断方向一致，但要指出的是一件更具体的事：**这里被扭曲的不是某一个指标，而是“学会”这个概念的含义本身**；而且这个扭曲不是近几十年绩效管理的产物，它与班级授课制同龄。这个时间尺度上的差别很重要——它说明问题不在管理方式，而在这条链的结构。

第六章 “让他自己发现”，为什么也在这条路上

判断一堂课属于哪条路，不看它的动作是热闹还是安静，只看它的终点是不是已经写在教案上了。

一、这一章会得罪人，但必须写

本章处理的是全书最容易引起反弹的一个判断：探究式教学、发现学习、做中学、启发式教学，在结构上并没有走出第一章那条链。

这个判断会得罪人，而且得罪的往往是最好的那批教师——因为愿意做探究式教学的人，通常是投入最多、最不愿意照本宣科的那些人。

所以必须把话说准：**本章不是说这些做法无效**。它们确实提高了接受的质量，确实让课堂更好，确实让更多孩子真的懂了。本章要指出的只有一件事：它们改变的是**接受这一环的形态**，而不是链条的方向。

二、“发现”这个词泄露了结构

先从词本身入手。

“让学生自己发现”——被发现的东西，必须**先在**。否则无从发现。

这不是玩弄字眼。当一堂课的设计者在心里说“我要让他们自己发现这个规律”时，有三件事同时为真：

第一，那个规律已经在他心里。

第二，他知道全班应当在下课前抵达那里。

第三，如果有孩子偏离了，他会把他拉回来。

第三条是关键。**拉回来的那个动作，就是这条链的方向。**

而且这个动作在探究式课堂上通常做得非常温和、非常有技巧——不是直接否定，而是“这个想法很有意思，我们先放一放”“你再想想还有没有别的可能”“某某同学的说法大家觉得怎么样”。技巧越高，拉回来的过程越不着痕迹，孩子越感觉是自己走到那里的。

这就是探究式教学的成就，也是它的边界所在：**它把“抵达”这件事做得像“发生”。**

三、唯一可靠的判据

由此得到本书用于识别课堂类型的唯一判据，它在后面所有编次里都会被使用：

不看动作，看终点：这堂课的结论，是不是在上课之前就已经写在教案上了？

这个判据有三个优点。

它不依赖观察者的立场。热闹与安静、传统与创新、讲授与讨论，这些都是风格判断，容易变成立场之争。终点写没写在教案上是一个事实问题，翻开教案就能看到。

它不依赖对教师意图的猜测。一位教师可以真诚地相信自己在做发生学教学，同时教案上写着结论。判据不问他信什么。

它可以被自己使用。一位教师可以在备课结束时问自己这一句，不需要任何人来评课。

四、一个便捷的替代检验

如果不方便查教案，还有一个更快的检验，尤其适用于项目式学习与综合实践：

这个活动做完，是不是每一组的结论都一样？

如果是，那么它是发现学的一条内道——只是路上安排了更多活动。

需要立刻补充一句，防止误用：结论相同**不必然**说明是发现学。有些问题本来就只有一个答案，各组独立走到同一处，这是发生，不是灌输。所以这个检验要配一个追问：

结论相同，是必然的，还是被我引导的？

区别在于：如果某一组当真走出了一个不同的、而且站得住的结论，这堂课会不会被认为是失败的？如果会，那么终点确实写在教案上。

五、探究式教学的真实位置

上面的分析很容易被读成对探究式教学的否定，因此必须给它一个准确的位置。

在本书的框架里，探究式教学是**老路上能达到的最高水平**。它把接受这一环做到了极致：让接收过程尽可能像一次发生，从而让接住的东西尽可能连着一些理由。这比直接告知好得多，好到不成比例。

而它的天花板同样清楚：**接收过程再像发生，也不是发生**。区别不在过程的形态，在终点的位置——终点在别处已经定好，还是在这一刻被做出来。

这个天花板在过去不是问题。因为在成本约束下，探究式教学已经是能负担的最好方案，而它的实施难度本身已经足够高，讨论天花板是奢侈的。

而现在这个天花板成了问题，原因在第二编：**探究式教学所优化的那一环，恰恰是被接管得最彻底的那一环**。一台可以为每个孩子量身定制探究路径、随时提供恰到好处好处的引导、永不失去耐心的机器，会把探究式教学做得比任何人类教师都好。

六、一个更难的问题：那到底有没有出路

如果连探究式教学都在老路上，那么什么不在？

完整回答在第五编。这里先给出方向，因为不给方向的批评是有害的。

出路不在于把引导做得更轻，也不在于让学生“更自由地探索”——后者往往只是把拉回来的动作推迟了。出路在于**存在一类问题，其结论在上课之前不在任何人手里**，包括教师，包括教材，包括 AI。

这类问题是存在的，而且比人们以为的多得多。第二十九章将给出识别它们的方法与各科的实例。这里只举一个：

“三角形内角和是不是一百八十度”——有答案，写在教案上。

“我们撕角拼直线那一下，到底靠的是什麼”——这个问题的答案不在教案上，因为教案上从来没有这一栏。

两个问题相隔一句话，分处两条路上。

七、三种常见的“伪未定终点”

第四节给出了判据，本节给出三种最常见的误判，它们都会让人误以为终点是未定的。

伪未定之一：开放性问题，收敛性评价。问题问得很开放（“你怎么看”“说说你的想法”），而评价仍然按一个标准答案给分。孩子很快学会：说什么都可以，但要说对的那一个。**这种课比直接讲授更消耗信任**，因为它同时要求孩子表现出自主，又惩罚真正的自主。

伪未定之二：终点未定，但路径全定。结论确实没有事先写死，但过程被设计得只有一条路可走：每一步的材料、次序、提示全部准备好，孩子沿着轨道走到某处，而那个“某处”是被轨道决定的。终点没写在教案上，写在轨道上。判别方法是问：**如果一个孩子在第三步换了个方向，这节课还能继续吗？**

伪未定之三：真的未定，但没有人当真。教师确实抛出了一个自己也没有答案的问题，孩子也确实说了些什么，然后下课铃响了，一切照旧——那些说法没有被记录，没有被追究，没有影响任何后续。**未定的终点如果不被结算，就等于没有出现过**。这是第五编第三十四章要用第四段来堵的漏洞。

八、把判据用在自己身上：一份备课自检

判据的价值在于可以自用。下面五问用于备课结束时的自检，每问只答是或否。

一问：这节课的结论，我现在能不能写下来？（能，则终点已定。）

二问：如果一个孩子得出了另一个结论，而且说得通，我会高兴还是会紧张？

三问：我准备的材料，是不是每一份都指向同一个方向？

四问：这节课里，有没有一段时间是我自己也不知道会发生什么的？

五问：如果这节课的录像被剪掉最后五分钟，还能看出结论是什么吗？（不能，说明结论是被走出来的，不是被宣布的。）

这五问不需要答给任何人听。但一位教师如果连续两个月每天答一次，他对自己课堂的判断会发生变化——而这本身就是一次回写。

九、一个更公道的说法

本章的判断容易被压缩成“探究式教学没用”，因此需要一个更公道、也更准确的说法。

探究式教学在老路上把接受这一环做到了极致，它是这条路能达到的最高水平。而它的处境在今天变了：**它优化的那一环，恰恰是被接管得最彻底的那一环。**

这不是探究式教学的失败，这是它的成功造成的结果——它把“引导学习者走向既定结论”这件事做得如此之好，以至于这件事变成了一个可以被形式化、被自动化的过程。一台可以为每个孩子实时生成探究路径、随时给出恰好合适的提示、永不失去耐心的机器，会把这件事做得比任何人更好。

所以本书对探究式教学的态度，与对整个老路的态度完全一致：**不是它错了，是它成了。**成到不再需要人来做。

十、可观察的动作

翻开你下一节课的教案。用红笔圈出“本课要得出的结论”那一句。然后问：如果有个孩子得出了另一个结论，而且他说得通，这节课算成功还是失败？

这个问题不需要回答给别人听。

十一、与相邻说法的分界

有一种教学论主张，认为学习应当是学习者主动建构的过程，而非被动接收。本章与该主张并不冲突，但要指出它的一个未解处：建构的**终点**是否被预先规定，该传统内部并未给出统一回答，而这恰恰是本书判据所在的位置。一次终点已定的建构与一次终点未定的建构，在课堂现场几乎无法区分，在结果上却分属两条赛道——第十七章将说明这个差别的具体后果。

第七章 老路的三大成就

光滑化不是一次损失，而是文明为了让知识跨越死亡与人数而付出的必要压缩；批评它之前，必须先算清它做成了什么。

一、为什么要专门写这一章

一本书如果只论证老路的问题，它会在两个地方失效。

第一个地方是现场。任何面向教师的论述，如果不先承认这套体系的成就，都会被正确地当作外行的指点。

第二个地方是判断本身。一个只看见代价而看不见成就的诊断，通常也会开出错误的药方——因为它不知道哪些东西是必须保住的。第五编的处方之所以不是“废除教科书”“取消统一课程”，正是因为本章。

所以这一章不是礼节，是约束条件的清单。

二、成就一：跨代存活

人会死，而知识不会——这件事一点都不自然，它是被造出来的。

一次发生是不可重复的：那个人、那个时刻、那个困难、那条走通了的路，全部一次性。如果知识只以“发生”的形式存在，那么随着那个人的死亡，它就消失了。人类历史上大量的技艺正是这样消失的：没有被压缩成可搬运的形态，只存在于某些人的手上和身体里，那些人死后，它就不在了。

光滑化做的事情，是把一次不可重复的过程，转换成一件可以无限复制、且不随任何个体死亡而消失的东西。定理写下来了，公式写下来了，方法写下来了。此后任何一个识字的人都可以取用。

这件事的量级很难被夸大。它是文明区别于一群聪明动物的分水岭。

而必须注意的是：跨代存活所要求的，恰恰是删掉路径与土壤。因为路径与土壤是属于那个人、那个时代的，它们不可搬运，而且会成为搬运的负担。光滑化不是压缩的副作用，压缩就是光滑化本身。

三、成就二：规模化

第三章算过那笔账：一对一的方案在数量级上不可能。

光滑化把边际成本压到接近于零，于是一个人可以同时服务几十人、几千人、几百万人。今天一个八岁孩子在课堂上花二十分钟接住的东西，人类曾经花了几百年才做出来——这个比率是全部现代教育的经济基础。

需要看清这个成就的分配意义。在没有光滑化的世界里，高质量的教育必然是极少数人的特权——因为它必须依赖一对一的师徒关系，而这种关系只有权力与财富能购买。普及教育在原理上依赖于可复制性。任何一种反对可复制性的教育主张，如果不同时给出新的成本方案，实质上都是在主张回到特权教育。

这一条对本书是一个硬约束：第五编的处方必须回答“这在四十人的班上做得了吗”。凡是在小班、名校、精英家庭里成立的方案，本书不予采纳。

四、成就三：校准

第三个成就最少被提及，却是现代社会运转的隐形基础设施。

成品的形态是统一的。一个在这个城市学过函数的孩子，与另一个城市学过函数的孩子，说的是同一件事；他们可以互相听懂，可以进入同一所大学，可以在同一个行业里合作。教师之间也是如此：一位教师可以接手另一位教师带过的班，因为他知道那些孩子接住了哪些成品。

这就是校准。它使**衔接**成为可能——学段与学段的衔接、学校与学校的衔接、教育与职业的衔接。

而校准的实现方式，恰恰是把个体差异从成品中消除掉。一个连着某个人的具体路径与土壤的知识，是无法校准的，因为它带着那个人的印记。为了让知识可以对照，必须先让它不属于任何人。

三个成就合起来，可以写成一句：

光滑化用“不属于任何人”换来了“属于所有人”。

五、成就的代价在哪里出现

现在可以精确地说出代价的位置了。

三个成就全部落在**第一个冲程**上——压缩与投递。第二章说过，知识跨越人与人的边界需要两个冲程，而文明把第一个做到了极致。

代价出现在第二个冲程：**解压从来没有被制度化**。

没有任何一门课程、任何一份课标、任何一次教师培训，是专门针对“这个成品如何在这个孩子身上重新发生”来设计的。有大量内容是关于如何讲得更好、如何设计更好的活动、如何检测掌握程度——这些全部属于第一个冲程的优化。

不是没有人想到过第二个冲程。教育史上关于“理解”与“记忆”的区分、关于“活的知识”与“死的知识”的讨论，从未中断。但它们始终停留在**理念层**，没有变成**制度**——没有读数，没有课时，没有评价，没有考核。而第五章已经说明，凡是没有读数的东西，在这套系统里会被自动筛除。

六、一个必须避免的误读

上面的分析很容易被读成“文明选错了”。这个读法是错的，而且错得很具体。

在**成本约束下**，**第二个冲程无法被制度化**。因为它必须是一对一的、慢的、不可批量的、不可预测的。第三章的账已经说明这在数量级上不可能。所以文明不是忽略了它，是**买不起它**，于是把它交给了运气。

这就得到本书对老路最完整的评价，它比第三章那句更准确：

发现学教育不是错的，也不是不完整的。它是一个在约束下完整的方案——它把**能做的那一半做到了极致，把做不到的那一半交给了运气**。

约束改变之后，“交给运气”这件事才第一次成为一个选择，而不是一个事实。

七、成就的第四项：可批评性

前面三项成就是常被承认的。这里补第四项，它很少被提及，却是学术共同体得以存在的前提。

成品是可以被批评的。一条写下来的定理、一个公开的证明、一份可复核的数据，任何人都可以指出它错在哪里。而一次未被光滑化的发生——某个人心里的想法、某位师傅手上的感觉——无法被批评，因为没有可供检查的对象。

光滑化把知识从“某人的东西”变成了“公共的东西”，从而使它可以被公开否定。科学的自我纠错机制建立在这个可能性上。

这一项对本书是一个重要约束：第五编主张让学生走出自己的路径，而路径是个人的、不可复制的。如果止步于此，就会退回到“每个人有自己的道理”这种无法互相检查的状态。所以第五编的每一处设计都必须附带一个把个人路径重新公共化的动作——第三十四章第四段的“念出来”、第三十条的“当众陈述并接受反驳”，作用都在这里。

不是不要光滑化，是不要只有光滑化；同样地，走过自己的路之后，还要能把它交出去接受检查。

八、三大成就构成的约束清单

把本章的成就翻译成对第五编的硬约束，共四条。任何一个不满足这四条的方案，本书不予采纳。

约束一：不得损害跨代存活。方案不得要求废除教科书、课程标准或系统性内容。凡主张“让孩子自己发现一切”的方案，在这一条上失败。

约束二：不得依赖小班或精英条件。方案必须能在四十人的普通班上运行。凡只能在小班、名校或高投入环境中成立的做法，本书视为无效。

约束三：不得破坏校准。方案不得导致不同学校、不同教师教出的孩子无法衔接。这一条限制了个性化的边界。

约束四：不得放弃可批评性。学生走出的路径必须以某种形式被公开、被记录、被反驳。

四条合起来，把第五编的设计空间压缩得相当窄。而本书认为这是好事：一个不受约束的教育主张，通常是一个没有被认真想过的教育主张。

九、代价的精确位置

第五节说代价出现在第二个冲程。这里再精确一步，因为“解压没有被制度化”这句话还是太笼统。

解压这件事，在制度里其实有影子：我们有“理解”这个词，有“融会贯通”这个说法，有“举一反三”这个要求，有“灵活运用”这个评语。也就是说，**制度知道它的存在，只是没有给它读数、课时与考核。**

一个东西被承认存在却不被测量，会发生什么？它会变成一种**德性归因**：接得快又能用的孩子，被说成“有悟性”“脑子活”“会学习”；接住了但不会用的，被说成“死读书”。这些说法把一个**结构性的缺失**，转译成了**个人的天赋差异**。

这个转译的代价是双重的：既让孩子承担了不属于他的责任，也让制度免除了本该由它承担的那一部分。第十七章将说明，这个“悟性”差异中的相当一部分，其实是可解释、可干预的。

十、可观察的读数

在你们学校的全部评价指标里，找一找有没有任何一项，测的是“这个孩子因为学了这个东西而改变了原来的某个想法”。

通常一项也没有。这不是疏忽——没有人知道怎么低成本地测它。第三十八章将给出四到五个可以低成本采集的读数，那一章的全部意义就在这里。

十一、与相邻说法的分界

哲学与社会学中有关于“客观知识必须被重新把握才成其为知识”的经典论述，与本章的两冲程结构相近。分界在于：那些论述关心的是知识的**存在形态**（它在被把握之前算不算知识），本章关心的是**制度分工**——为什么一个文明会把两个冲程中的一个做到极致而把另一个交给运气，以及这个分工在成本结构改变之后是否还应当维持。前者是本体论问题，后者是可以被政策改变的问题。

第八章 老路的读数体系，以及它如何自我确证

老路的全部读数，量的都是同一件事——成品被接住了多少；正因如此，它无法测出自身的失败，只能测出自身的成功。

一、把读数列出来

第一编的最后一章，处理这套体系的自我确证机制。这一章是第二编的引信：如果不先说清老路的读数量的是什么，第十二章的“加固悖论”就无法成立。

先把常用读数列全：

当堂正确率——讲完之后，多少人做对了。

作业完成度与正确率——回家之后还能不能做对。

单元测验与考试成绩——一段时间之后还能不能做对。

进度——单位时间内覆盖了多少内容。

掌握率——某个知识点被多少比例的学生掌握。

参与度——课堂上有多少人举手、发言、投入。

满意度——学生和家长觉得这课上得好不好。

这七项覆盖了当前绝大多数课堂评价与教学质量监测。

二、它们量的是同一件事

现在把它们逐一还原。

前五项显然是同一族：它们全部测量**成品是否被接住，以及接住了多少、多牢、多快**。这正是第一章的“接受”这一环。它们之间的差别只是时间尺度与内容范围，不是性质。

第六项参与度看似不同，其实仍在这一族里。参与度之所以被测量，是因为它被认为是接受质量的前置指标——投入的孩子接得更好。它是一个代理变量，代理的仍然是接受。

第七项满意度更间接，但方向一致：它测的是接收体验的舒适程度。

于是七项归一：

老路的全部读数，量的都是成品被接住的比例与质量。

这个归约不是本书的发明，它是自然的结果。第三章已经解释了原因：在成本约束下，唯一可以低成本测量的就是接住与否；而第五章说明了，可测量的那一项会逐渐成为“学会”这个词的定义。

三、自我确证是怎么发生的

一个系统如果它的读数只覆盖它擅长的那一部分，就会产生自我确证。

具体机制是这样的。这套体系在搬运、传道、接受三个环节上不断优化；每一次优化都提高了成品的接收比例；而读数量的正是接收比例；于是每一次优化都被读数确认为改进。数据支持继续优化，优化之后数据更好，更好的数据支持进一步优化。

这个循环从不接触第二个冲程。解压有没有发生，读数不问，因此也无从否定。

一个系统的全部故障模式都指向内部环节而从不指向自身结构——这句话在第一章出现过一次，现在可以给出它的完整含义：**这不是因为设计者回避批评，而是因为读数体系在结构上没有能力表达那种批评。**

要说一个系统在结构上错了，需要一个来自体系之外的读数。而在成本约束下，那个读数造不出来。

四、这套体系为什么在过去仍然有效

必须回答一个自然的疑问：如果读数只量接受，那么两千年来它为什么没有出大问题？

因为在过去，**接受与理解高度相关。**

理由不难看出。在光滑成品的供给稀缺、且获取有摩擦的年代，一个孩子要真正“接住”一样东西，往往不得不自己走一段：他得自己看书，看不懂就得反复看；他得等到明天才有机会问老师，等待的时候只能自己琢磨；他手边没有第二种说法，只能拿自己已有的东西去凑。

这些摩擦不是好东西，但它们逼出了一部分解压。于是接得好的孩子，通常也确实解压得比较多；接受这个指标，间接地、不完美地、但确实地反映了理解。

正因为两者相关，用前者代替后者才没有立刻出问题。这个替代甚至是理性的：既然相关，就测那个便宜的。

五、相关性是可以被打断的

而相关不是同一。两个高度相关的量，其相关性依赖于产生这种相关的那些具体条件；条件一旦改变，相关性可以断掉，而所有基于它建立起来的判断会同时失效，且**不会有任何警报。**

产生这一相关的条件是什么？就是上一节说的那些摩擦：供给稀缺、获取有成本、等待不可避免、说法只有一种。

第二编将说明，这些条件在几年之内被逐一取消了。

而当相关性断掉之后，会发生一件在读数上完全看不出来的事：**接收效率可以趋近于无穷，而解压可以趋近于零**。七项读数会全部变好，同时那件真正重要的事情停止发生。

这就是第十二章“加固悖论”的完整机制。它的根在这里，不在AI——**AI没有制造这个漏洞它只是第一次把漏洞撑开到肉眼可见**。

六、读数体系的第八项：时间

上面列了七项，还漏了一项最不像读数的读数：**课时**。

课时是老路最硬的一把秤。它规定了每一个内容必须在多长时间内被讲完，因此它规定了**任何一个岔口最多能张开多久**。一位教师即使完全认同本书的全部主张，他每周仍然面对一张必须走完的进度表。

这就是为什么第三十四章的四段结构里，最难落地的不是第三段的反例，而是第一段与第二段合计二十三分钟的“没有产出”。二十三分钟在一张按内容排定的进度表上，是一个昂贵到几乎不可能被批准的支出。

本书对此的处理不是回避，而是要指出一个常被忽略的事实：**这二十三分钟原本用于讲解答案，而答案现在是免费的**。孩子在课后三秒钟就能拿到讲解。所以问题不是“腾不腾得出时间”，而是**这段课堂时间应当用来做那件免费的事，还是那件只有此时此地才能做的事**。

课时没有增加，用途变了。这是第五编全部方案在时间上的可行性根据。

七、为什么这套读数至今仍在被加强

读数体系在过去二十年不但没有松动，反而被大幅加强：数据化的作业系统、学情分析平台、知识点掌握度的可视化、错题的自动归因。

按本章的分析，这不是巧合。这些技术全部服务于**更精确地测量接受**，而它们之所以出现得如此迅速，正是因为接受是唯一容易被数据化的那一项——它有明确的对错，有可比较的分数，有可以聚合的样本。

而第二个冲程在数据上几乎是隐形的：一个孩子回头改了一件旧的，这件事不产生分数，不进入任何系统，往往连教师自己都没注意到。

于是形成一个自我强化的循环：**可数据化的部分被越测越细，不可数据化的部分被越挤越薄**。这不是任何人的意图，这是数据本身的引力。

第三十八章给出的五个读数，全部是按“能不能低成本采集”这条标准设计的，原因就在这里。**一把新秤如果昂贵，它就不会被使用；不被使用的秤，等于不存在**。

八、第一编与全书的接榫

第一编结束在一个悬置的判断上：这套体系之所以长期有效，依赖于摩擦的存在。

这个判断值得再说一遍，因为它是全书的枢纽，而且它的形式很不寻常——**这套体系的有效性，依赖于一个它自己一直在努力消除的东西**。

两千年来，教育的全部努力方向都是减少摩擦：让教材更清楚、让讲解更明白、让获取更容易、让等待更短。而摩擦恰恰是使“接受”与“理解”保持相关的那个东西。

也就是说：这套体系一直在削弱自己赖以成立的前提，只是过去削得太慢，慢到看不出来。

AI 做的事情，是把这个持续两千年的缓慢过程，在几年之内推到终点。它没有引入任何新的方向，它只是把原有方向上的速度提高了几个数量级。

这就是为什么本书用“包办”而不用“颠覆”。颠覆意味着来了一个对手，包办意味着有人把你正在做的事做完了。第二编从这里开始。

九、第一编的结论

第一编到此结束。它的全部内容可以压成四句：

第一，教学是一条由搬运、传道、接受构成的单向链，链上流动的全部是成品。

第二，成品由光滑化制成——减去路径与土壤；这个操作是必要的，因为不压缩就搬不动。

第三，光滑化换来了三大成就：跨代存活、规模化、校准；它的代价是第二个冲程从未被制度化，因为在成本约束下买不起。

第四，这套体系的全部读数只量第一个冲程，因此它在结构上只能测出自身的成功；它之所以长期有效，依赖于一个外部条件——摩擦的存在，使接受与理解保持相关。

第二编要说的是：那个外部条件没有了。

十、可观察的读数

拿出你们学校最近一次教学质量分析报告，逐条看它的每一个指标，问一句：这一项如果全班都很好看，能不能推出“孩子们真的变了”？

通常一项也推不出。而这并非报告写得不好——是这套读数体系从设计之初，就不是为回答那个问题而造的。

十一、与相邻说法的分界

管理学中有一族关于指标失效的经典判断：指标一旦成为目标即不再是好指标；被测量者会对测量做出反应；组织会优化可见之物而牺牲不可见之物。本章与它们方向一致，但要指出的差别是：那一族判断处理的是**指标被扭曲**，本章处理的是**指标本来就只覆盖一半**——它不是被人操纵坏的，它从一开始就没有覆盖第二个冲程，而且在过去的成本条件下，它也不可能覆盖。这个差别决定了处方不同：前者的处方是防止操纵，后者的处方是**造一把新的秤**。第三十八章将做这件事。

第一编材料：十二个标本与一节课的成品清点

本编材料的性质：以下标本的历史部分为公认的学科史事实；课堂片段与清点记录为教学示例与推演，不是实测数据。凡涉及数字之处，均为便于说明的示例值。

材料一 十二个标本

第四章解剖了五个（负数、分数、无理数、内角和、等号）。这里补七个，覆盖到其余学科，格式统一为三栏：教科书的第一句、当年的困难、被删掉的那一样。

六、小数

教科书：把整数按十进制向右延伸，小数点后一位是十分位。

当年：小数的普及远晚于分数，它的动机不是“分不尽”，而是**计算太麻烦**——用分数做乘除要通分，而十进制小数可以照整数的办法算。它是一次为了省力而做的改造。

被删掉的：它是一个**工程性的选择**，不是一个新发现的数。孩子因此无从理解为什么有些分数化成小数会除不尽——那正是这次改造付出的代价。

七、平均数

教科书：把总数除以份数。

当年：平均数最早用于处理**互相矛盾的观测**——几个人量同一个东西量出不同的结果，取哪一个？取平均是一次约定，它假设误差是对称的。

被删掉的：平均数是一个**对付分歧的办法**，不是一个描述事实的量。孩子因此学不会问“这里的平均有没有意义”，于是把人均收入和班级平均分当成同一类东西。

八、乘法

教科书：几个相同加数的和。

当年：这个定义在整数上成立，在分数与小数上立刻失效——半个的三分之二，不是任何“几个相同加数的和”。

被删掉的：**定义被换过而没有人说**。孩子带着“乘法使数变大”这条从加法沿用来的直觉，一路走到小数乘法，然后崩掉。这不是粗心，是一次未被告知的换轨。

九、光合作用

教科书：植物利用光能，把二氧化碳和水合成有机物，释放氧气。

当年：这个结论花了两百年，中间的关键一步是**发现植物增加的重量不来自土壤**——有人用盆栽称重，发现树长了几十公斤，土只少了几十克。

被删掉的：那个可以当场做的困难（重量从哪来）。今天完全可以在教室里让孩子先猜一次，而教科书直接给出方程式。

十、力

教科书：力是物体之间的相互作用，可以改变物体的运动状态。

当年：真正难的不是“力”，是**放弃“运动需要力来维持”这条直觉**。这条直觉在日常经验里几乎处处成立（推一下停一下），推翻它需要一个当时不存在的东西——理想化的无摩擦情形。

被删掉的：那条直觉曾经是对的，而且是被一个**想象出来的、现实中不存在的情形**推翻的。这是科学方法上的一件大事，教科书把它压成一句“理想情况下”。

十一、字词的“本义”

教科书：某字本义为某某，引申为某某。

当年：所谓本义，是后人根据现存材料**重建出来的**，材料换了，结论会变；不少字的本义在学界仍有争议。

被删掉的：**它是一个重建，不是一份记录**。孩子因此把字典当成关于世界的记载，而不是关于研究的报告。

十二、时区

教科书：全球分为二十四个时区。

当年：时区是**铁路逼出来的**。在铁路以前，每个城镇用自己的正午，互相不冲突；火车跑起来之后，时刻表无法编写，于是不得不统一。

被删掉的：时间的统一是一次**行政决定**，而且遭到过抵制。这个标本在小学阶段特别好用，因为孩子可以立刻理解“为什么以前不需要统一”。

材料二 一节常规课的成品清点

第一章第七节要求数一数：一节课里，多少分钟流动的是成品，多少分钟流动的是一个还没有成品的问题。这里给出一份清点的**格式与示例**，供教研组照做。

清点方法。录音，逐分钟标记，只分两类：

A类——此刻正在流动的是一个已经有答案的东西（讲解、复述、订正、演示、朗读、按步骤操作）。

B类——此刻在场的是一个还没有答案的问题（有人在猜、在试、在争、在等）。

判别的唯一标准：此刻如果有人问“那到底是什么”，教师说不说得出来。说得出来是A，说不出来是B。

一节四十分钟数学课的示例清点（示例，非实测）：

分钟	内容	类别
0-3	复习旧知，提问三个已学结论	A
3-6	引入情境，讲一个生活例子	A
6-8	提问“你们觉得会是多少”	B
8-9	有学生答对，教师确认	A
9-22	讲解方法，板书，两道例题	A
22-35	学生练习，教师巡视个别指导	A
35-40	小结，布置作业	A

B类合计：2分钟，占5%。而这两分钟在第8分钟被一个正确答案关闭。

这份清点表最有价值的用途不是评价，是让教师第一次看见自己那节课的**形状**。多数人第一次做完的反应是相似的：不是“我讲得太多”，而是“原来那两分钟是唯一的一段”。

材料三 三种“规定”的开采实例

第三十六章说数学教材上每一处“我们规定”都是一个岔口。这里给出一份开采清单的示例取自小学阶段的常见内容。

教材原句	可开采的问句	能不能被走两步
“我们规定，零不能作除数”	是不能，还是不许？如果许了会怎样	能（可让学生尝试推出矛盾）
“先算乘除，后算加减”	谁定的？定成相反的会怎样	能（可让学生自定一套并互相出题）
“分数除以分数，等于乘它的倒数”	为什么颠倒就对了	能（可回到“分不尽”的原始情境）
“任何数的零次方等于一”	这是算出来的还是定出来的	能（可让学生找“为了让规律不断”这一理由）
“两点确定一条直线”	这是观察到的还是规定的	较难（需要公理意识，高年级适用）

最后一行说明了一件事：不是所有“规定”都同样好开采。能不能被走两步（第二十九章第六节）是硬条件，过不了这一条的先放下。

第一编材料（续）：测度外壳与再分泌律

本节材料的来源：SDE 站内文章《测度外壳》（教育社会学向）。此处取其机制与那条预测，用于补强本书第八章（读数体系）与第二十六章（秤）。

材料二十八 评价装置的产品是“可接受的拒绝”

本书第八章说，老路的读数体系只量第一个冲程，因此只能测出自身的成功。这个说法是对的，但它没有回答一个更根本的问题：

评价装置是干什么用的？

站内那篇文章给出的答案不是“测量学习”，而是：

评价装置的产品，是可接受的拒绝。

一个社会必须把有限的位置分配出去，因而必须拒绝大多数人；而拒绝要能被接受，必须显得是根据某种客观的东西做出的。评价装置生产的正是这个“客观的东西”。

这个判断对本书的影响是重的。它意味着第二十六章那句“没有第二套秤，讲什么都没用”还不够狠——秤之所以难换，不只是因为惯性，是因为它承担着一项本书之外的功能。任何一把新秤，若不能承担分配与拒绝的功能，就不会被采纳；而一旦它承担了，第三十八章的三条纪律当天失效。

这解释了第二十六章第六节为什么会遇到那个天花板，也说明那个天花板比本书原先写的更硬。

材料二十九 再分泌律：外壳寿命 $\approx$ 破解成本的单调增函数

机制。一层评价外壳一旦被大规模破解（应试技巧成熟、辅导产业化、模板可得），它区分人的能力下降，于是系统会分泌新的一层——从背诵到理解，从理解到应用，从应用到素养，从素养到综合表现。

寿命公式：

外壳寿命 $\approx$ 破解成本 c 的单调增函数

破解越贵，这层壳活得越久。

本书最锋利的一条推论出现在这里：

AI 摧毁评价装置，恰恰是因为它在破解能力上是平等的；而破解的不平等，正是外壳赖以生存的东西。

过去破解一层外壳需要资源（名师、教辅、时间、家庭），因而破解是不平等的，因而外壳仍能区分人。AI 把破解成本压到接近于零，且对所有人一样——于是这层壳在一个学年之内失效，而不是过去的两代人。

再分泌周期从两代人压缩到一个学年——这是本书第二十一章那条三段曲线之外的另一条时间线，而且它是可观测的。

材料三十 下一层外壳会建立在什么上

由上面的机制可以推出一条预测，本书愿意把它列为最重的一条：

下一层外壳，必须建立在“不可被工具补足的东西”上。

因为凡是可以被工具补足的，破解成本都已经接近于零，做出来的壳活不过一个学年。

而当前唯一符合这个条件的候选，恰恰就是本书前六编所讨论的那些东西——守住一段不稳定的能力、在没有路的地方走两步的意愿、回头改自己的动作。

这条预测有一个不受欢迎的推论：如果它成立，本书所主张的那些东西将被制度化，然后被评价，然后被训练，然后失效。第三十八章的三条纪律正是为了推迟这一天，但推迟不是阻止。

再分泌律的唯一已知反例是智力测验分流在部分地区的退出。它退出的三个条件是：被证明有系统性偏差、被公开质疑、且存在一个替代品。这三条同时具备的情形极少。

材料三十一 对第二十六章的一处修订

综合上面两节，本书第二十六章那句话需要改写。原文是：

没有第二套秤，讲什么都没用。

现在应当写成：

没有第二套秤，讲什么都没用了；而一旦第二套秤承担了分配与拒绝的功能，它就会变成第一套秤。

这不是一个悲观的结论，而是一个操作上的约束：新读数必须被造得“不能用于分配”。第三十八章的五个读数中，读数二（未闭合时长）与读数五（空位题反应）最接近这个要求——前者只设下限因而无法排序，后者没有可训练的答案。

这一条应当被理解为一设计标准，而不是一条抱怨。

第二编 包办

第一编说明，教学是一条流动着成品的单向链，而这条链的全部读数只量成品被接住了多少；它之所以长期有效，依赖于摩擦的存在。第二编要说的是：那些摩擦在几年之内被逐一取消了。三个动作被逐项接管，而接管的同时，这条路在所有读数上显得越来越正确——本编把这一结构称为加固悖论，并处理它最纯粹、也最受推崇的形态。

第九章 搬运被接管

教学转置——把学术形态的知识改写成可教形态——曾经是一项以年计的稀缺劳动；它现在的边际成本接近于零，而且可以为任意一个具体的孩子当场重做一遍。

一、这一环上人类的优势曾经在哪里

第一章说搬运不是简单的位移。把一样学术形态的东西改写成课堂形态，需要一个人同时具备三种东西：懂这门学科、懂这个年龄的孩子、懂怎么把话说得让人接得住。

三样同时具备的人很少。所以教材的编写从来是一项集体的、以年为单位的、反复打磨的工作。一套小学数学教材从立项到定稿，五年是常态；其中真正难的部分不是写，是定次序——先讲哪个后讲哪个，哪个概念必须等另一个概念站稳了才能出现，哪一处可以先给一个不完全正确但够用的说法，等两年后再修正。

这些判断没有公式，只能靠经验。而经验之所以稀缺，是因为它需要在大量真实的孩子身上试错，而每一次试错的周期是一个学年。

人类在这一环上的优势，来自这种试错的不可加速性。

二、优势消失的具体过程

这个优势在几年之内消失了，而且消失的方式与人们预想的不同。

预想中的替代方式是“AI 编教材”——一个系统写出一套更好的教科书。这件事确实在发生，但它不是关键。关键是另一件更不起眼的事：教材这个形态本身，不再是搬运的唯一形态。

教材之所以存在，是因为搬运的成本高，所以必须一次编好、大量复制、长期使用。它是一个为了摊薄成本而存在的中间产物。

而当改写的边际成本接近于零时，这个中间产物就不再必要了。不需要一套面向全体的教材，可以为眼前这个孩子、在这一分钟、按他此刻的具体困惑，把同一个内容当场重新组织一遍。他上一句话说了什么，决定了下一句该怎么讲。

这不是“更好的教材”，这是**教材的消失**——搬运从批量生产变成了即时生成。

三、被接管的到底是什么

需要说得更准一些，因为“搬运被接管”这句话容易被理解为“以后不用编教材了”。教材当然还会存在，出于校准（第七章）与责任归属的需要，它必须存在。

被接管的是**教学转置这个判断行为**。

具体地说：一个概念该拆成几步、每一步落在哪里、用什么例子、先说什么后说什么、哪里可以跳、哪里必须停——这些判断过去由教材编者做一次，教师再按自己的经验做第二次微调。现在这些判断可以被即时做出，而且是针对具体对象做出的。

一位教师在这一环上还剩下什么？他仍然可以做这些判断，而且做得不差。但他做一次要花二十分钟备课，机器做一次要两秒；他一天能做十次，机器一天能做十万次；他为四十个人做一份，机器为四十个人做四十份。

这不是质量之争，是数量级之争。在数量级之争里没有“扬长避短”这回事。

四、一个必须澄清的问题：那教师是不是没用了

这里必须立刻回答一个问题，否则本编后面几章会被读成一份失业通知。

被取代的是**位置**，不是人。

一个位置消失了，占据它的人必须迁移到另一个位置去。这在技术史上反复发生：会计不再用算盘，绘图员不再用鸭嘴笔，排字工不再摆铅字。这些变化在当时都是痛苦的，但它们本身不是灾难。

灾难只在一种情况下发生：**新位置没有被识别出来**，于是人守着旧位置不动，直到旧位置连同人一起被淘汰。

所以本书对教师群体的核心主张，不是“你们要被替代了”，而是“有一个位置正空着，而且只有你们能占”。这个位置是什么，第三十五章给出。在此之前，需要先把话说到底：如果教师群体在未来几年里继续把主要精力投入“把课讲得更清楚”，那么这场竞争的结果是确定的，因为对手在这一项上没有上限。

五、一个反向的机会

搬运被接管，对本书的主张其实是一件好事，值得单独指出。

第一编的四个标本说明，教材上的每一个概念都被删掉了它的来路。而删掉来路的原因之一是篇幅——教材必须简短，因为它要印出来、要背在书包里、要在有限课时内讲完。

即时生成没有这个约束。**当搬运的成本降为零时，“把来路带上”第一次成为可能。**一个孩子完全可以在两分钟内知道，负数当年是怎么被拒绝的、无理数当年引起了什么震动、内角和依赖一条什么样的前提。

这不是本书的处方——第四章第九节已经说明，把来路作为一段新的讲解端出去，仍然只是增加了一个成品。但它取消了一个借口：**从此不能再说“课时不够，讲不了这些”**。讲得了。问题从“有没有条件”变成了“要不要用这段条件去做另一件事”。

六、可观察的动作

把你下周要讲的一个概念，让 AI 用五种不同的方式各讲一遍。然后问自己：这五种里，有没有一种是我自己想得到的？有几种？

多数教师做完这个动作会得到一个不舒服但清楚的结论。这个结论是本编后面全部内容的前提，越早得到越好。

七、与相邻说法的分界

教育技术领域长期有一个判断：技术不会取代教师，只会成为教师的工具。本章不反驳这个判断的结论，但要指出它的论证在今天失效了。那个判断成立的年代，技术能做的是**呈现**（投影、录像、课件），而不是**判断**；教学转置属于判断，因而安全。现在判断这一项被覆盖了，所以“技术只是工具”这句话如果还要成立，必须给出新的理由，而不能沿用旧的。第三十五章将给出一个理由，但它与“工具论”完全不同——它不主张技术做不了教师的事，而主张有一件事它的构造决定了它做不了。

第十章 传道被接管，而且被超越

讲解的核心技艺是“换一种说法”；在清晰、准确、不厌其烦这三项上，人类教师已经不是最优选项，而这三项恰恰是过去衡量好教师的全部标准。

一、把好教师的标准写出来

第一章给过一个定义：在老路上，好教师是损耗最小的那个传输者。

这个定义可以拆成可比较的项：

清晰——一句话说出去，有多少人一次听懂。

准确——说出去的东西有没有错，边界有没有交代。

耐心——同一件事讲第三遍时，语气和第一遍是否相同。

换说法的储备——同一个概念，能有多少种不同的进入方式。

读人的能力——从一个孩子的表情、语气、错法里，判断他卡在哪里。

前四项在本章处理，第五项在下一章处理，因为它属于“接受”那一环。

二、前四项逐项对照

清晰。这一项上人类的表现极不均匀。一位特级教师可以把抽象概念拆成十二个可感知的台阶，而这种能力全国范围内是稀缺资源。而机器给出的讲解，其清晰度的下限已经高于大多数课堂的平均水平，且不随疲劳、情绪、班级次序而波动。

准确。这一项人类曾经领先，现在情况复杂：机器仍会出错，但人类教师同样会出错，而且人类的错误往往更系统——一个教师若在某处理解有偏差，他会连续二十年在同一处教错，且没有任何机制发现它。**分散的随机错误与系统性的持续错误，哪一个对一个孩子伤害更大，不是显然的。**

耐心。这一项没有争议。人的耐心是有限资源，而且分配不均——一个提第五次同样问题的孩子，得到的语气与第一次不同，这不是师德问题，是生理事实。机器在这一项上不是“更好”，是**不在同一个量纲上**：它没有耐心，因而也不会耗尽耐心。

换说法的储备。一位优秀教师一生能积累的说法，量级在几十种；这些说法是他三十年里从真实孩子身上试出来的，极其宝贵。而机器可以在一次对话里给出十种，并根据这个孩子的上一句回话再生成十种。

四项合起来的结论，不需要争论，任何一位教师自己试一次就会承认：

在“把一件已经知道的事讲明白”这件事上，我们已经不是最优选项了。

三、这一章为什么最难讲

在教师培训现场，这一章是最容易引起沉默的一章。原因不难理解：上面四项，正是一位教师三十年积累的核心资产，也是他在这个职业里全部自尊的来源。

所以必须把话说准，而且要说在前面：

放弃的不是能力，是用途。

同样的学科理解力，用在“把它讲清楚”上已经不再稀缺；用在“知道这门学科的哪一处规定当年本可以是别的样子”上，则无可替代。而这两件事需要的是同一种理解力——**讲得清楚的能力，正是设计岔口的必要条件。**

一个自己都没有真正想透这个概念的人，既讲不清楚它，也不可能知道该在哪里把它撬开。所以第五编要求教师做的事情，不比现在容易，只是方向不同。三十年的积累不作废，它换了一个用法。

这句话应当在每一次教师培训中被明确说出来，而且要在批评之前说。

四、一个不对称：讲解可以被复制，判断不能

传道这一环被接管得如此彻底，还有一个更深的原因，值得单独指出。

讲解是**面向所有人的**。一段讲解一旦被生成，它对任何一个处于类似位置的人都同样有效。这就是为什么讲解可以被录制、被传播、被复用——名师课堂的录像之所以有价值，正因为讲解的价值不依赖于听者是谁。

而判断是**面向这一个人的**：此刻他卡在哪里、他刚才那句话说出了什么、现在该不该开口。这一项无法被预先录制，因为它依赖于一个尚未发生的具体情境。

所以严格地说，被接管的是讲解，不是全部教学。而问题在于：**在老路上，教学的绝大部分就是讲解。**判断在老路上只占很小的比重，而且主要用于服务讲解（判断他没听懂，所以再讲一遍）。

第五编要做的事情，正是把判断从服务讲解的位置上解放出来，让它成为主业。

五、一个容易被忽略的损失

传道被接管，还带来一个损失，它不出现在任何效率的计算里。

一位教师讲解的时候，他同时在做一件从不被写进教案的事：**他在示范一个人怎么思考。**他停顿的地方、他犹豫的地方、他说“这里我当年也想了很久”的地方、他被一个学生问住时的反应——这些构成了一份关于“思考是什么样子”的活样本。

而一段完美的讲解不提供这份样本，因为它没有停顿、没有犹豫、没有被问住。**它展示的是思考的结果，不是思考的过程。**

这一点在本书的框架里是可以被安放的：它正是光滑化在**行为层面**的版本。一段无懈可击的讲解，是一次被磨光的思考。孩子从中学到内容，但学不到“人是怎么走到这个内容的”。

这个损失过去被教师的不完美所抵消——教师会卡住，会说错，会当场改口。**而这些不完美现在成了纯粹的劣势，在任何对比中都被判为减分项。**

第五编第三十二章会把这一点变成一条正面的设计：教师在课堂上暴露自己的不确定，不是失职，是提供样本。

六、可观察的动作

回想你教学生涯里最得意的一次讲解。现在把同一个内容交给 AI，看它讲成什么样。

这个动作会带来两种反应中的一种：或者你发现它讲得不如你——那么请记住它差在哪里，那个差别很可能就是判断而不是讲解；或者你发现它讲得比你**好**——那么这一章的判断在你这里成立了。

两种反应都有用，而且都指向同一个下一步。

七、一条别的行业已经付过学费的规律

传道被接管这件事，在别的行业先演过一遍，而且代价已被记录。

自动化研究里有一条经典表述，通常被称为“自动化的讽刺”：**自动化程度越高，留给人类操作员的任务就越是那些自动化处理不了的异常；而恰恰因为日常操作已被接管，操作员再也没有机会保持处理异常所需的技能。**关于“脱环”、关于自动化的误用与废用，说的是同一件事。

把它翻译到教育上：一位教师之所以能判断“这个孩子现在是真的在想还是已经放弃了”，是因为他自己在过去许多年里被真实的课堂反复顶回来过——他判断错过，学生没听懂，他被迫重来。**这个“被现实顶回来”的循环，是他判断力得以维持的唯一来源。**

一旦这个循环被接管，判断力不是慢慢退化，而是**从来不再生成**——新入职的教师将不再经历它。

这条规律在教育里有一个更麻烦的性质：别的行业出事的时候有飞机掉下来、有病人死亡、有生产线停摆，会被记录下来；而**教育出事的方式是什么都不会发生**——一届学生照常毕业，分数甚至更好看，只是二十年后这个社会里能自己做判断的人变少了，而没有任何一份报表记录这件事。

第四十四章第五节将把这条论证写成完整形态。这里先立一句：**被接管的不只是教师现在的技艺，还有下一代教师获得这份技艺的途径。**

八、与相邻说法的分界

有一个流行的安慰性说法：机器可以传授知识，但不能传递情感、价值观与人格影响。本章不否认情感与人格的重要性，但要指出这个说法在论证上有一个漏洞：它把教师的不可替代性建立在**教学之外**的东西上，等于承认了教学本身可以被替代。如果教师的价值主要在于陪伴与人格感染，那么这份价值在原则上可以由任何一个有爱心的成年人提供，不需要三十年的学科积累。本书拒绝走这条路。第三十五章给出的不可替代性，落在**教学之内**：守住岔口是一个教学动作，不是一个情感动作。

第十一章 接受被接管：供给不再稀缺

接受这一环的瓶颈从来不是孩子，而是供给的及时性；当供给变得无限、即时、私密、零成本时，“接住”变得前所未有的容易——而卡住的那段时间随之消失。

一、这一环最隐蔽

搬运和传道被接管，是显而易见的。接受这一环看上去不同：接受是孩子的事，机器怎么可能替他接受？

这个直觉是对的，也正因为如此，这一环的接管最隐蔽，也最要命。

关键在于看清接受这一环真正的瓶颈在哪里。它不在孩子的能力上——绝大多数孩子有能力接住绝大多数内容。它在**供给的及时性**上：孩子卡住的那一刻，能不能立刻拿到一个恰好合适的说法。

过去这个供给极度稀缺。一个教师面对四十个学生，卡住的孩子多半只能等：等下课，等老师有空，等自己鼓起勇气举手，等回家有人辅导，或者等到根本没人来，自己想办法。

二、四道摩擦，一起消失

把过去获取供给的过程拆开，会看到四道摩擦，每一道都在几年之内消失了。

第一道：时间摩擦。要等。等老师讲到、等下课、等明天。等待的长度以分钟到天计。

第二道：社交摩擦。要问人。要举手，要承认自己不会，要面对可能的“这么简单你都不懂”。对相当一部分孩子——尤其是自尊心强的和成绩中等的——这道摩擦足以让他们放弃提问。

第三道：匹配摩擦。问到的答案未必对得上。老师的解释是面向全班的，未必落在他卡住的那一点上；同学的解释可能更差。往往要问好几次才对得上。

第四道：次数摩擦。不能反复问。同一个问题问三遍，无论对方多好脾气，孩子自己会停。

四道摩擦合起来，构成了一个筛子：**只有那些真的卡住了、而且卡得足够久、足够在意的孩子，才会最终获得供给。**在此之前，他被迫自己待了一会儿。

现在四道全部消失。供给是无限的、即时的、私密的、可无限重复的、且精确匹配的。

三、消失的不是摩擦，是那段时间

这里是本章的关键判断，也是全书的枢纽之一。

那四道摩擦本身没有价值。等待是浪费，社交压力是伤害，匹配不良是低效，不敢反复问是损失。**没有任何理由为它们辩护。**

有价值的是它们的副产品：**在供给到达之前的那一段时间里，孩子只能自己待着。**

他在那段时间里未必做了什么了不起的事。他可能在发呆，可能在乱试，可能在把题目又读了一遍，可能在把自己已有的几样东西翻出来凑一凑。但正是这些动作——尤其是“把自己已有的东西翻出来”这一个——构成了第十六章将要讨论的那件事的全部素材。

摩擦消失，那段时间也随之消失。**而消失的方式使它完全不可见：没有人减少了任何东西，只是某件事发生得更快了。**

四、一个重要的区分：这不是“孩子变懒了”

必须堵住一个流行的、也是错误的解释：孩子不肯自己想，是因为他们变懒了、意志力下降了、被短视频毁了。

本书不采纳这个解释，理由有三。

第一，它不可检验。“这一代人意志力下降”这类判断在每一代人身上都被说过一次，而且从未被测量过。

第二，它归错了责任。一个孩子在三秒钟内拿到答案，不是他懒，是他面前有一个三秒钟就能给出答案的东西，而没有任何规则要求他先等三分钟。**在没有约束的情况下走最短路径，不是懒惰，是理性。**

第三，也是最重要的：它导向错误的处方。如果原因是意志力，处方就是“要求孩子更自律”——而这在一个零摩擦的环境里是不可能持续的，最终只会变成对孩子的道德指责。

正确的处方在结构一侧：第二十六章的秤，第三十四章的规矩，第十九章第六节的两个部件。**要改的是环境，不是意志。**

五、“个性化”在这一环上的真实作用

个性化通常被视为这一环上最大的进步：为每个孩子定制难度、节奏、说法，让每个人都能接住。

在老路的框架里，这个判断完全正确，而且是了不起的成就——第三章说过，因材施教两千年来一直是理想而非现实，现在它第一次可以实现。

而在本书的框架里，需要补一句：个性化消除的最后一样东西，是“听不懂”。

听不懂原本是一个偶然产生的岔口。它不是任何人设计出来的，它是“面向全班的讲解”这个粗糙形式的副产品：讲解落在平均水平上，于是总有一些孩子在某处对不上，被迫自己去接。

一个完美个性化的讲解不产生这个副产品。它精确落在他够得着的地方，所以他每一步都跟得上。

个性化就是提纯。

这句话在第十九章会被再次用到。这里先指出它的形式：老路上的每一项优化，都在消除某种粗糙；而某些粗糙恰恰是第二个冲程的唯一入口。

六、这一环上还剩下什么

三个动作全部被接管，那么这条链上还有没有一个位置是非人不可的？

按本编的分析：没有。这条链上的每一个位置，都可以被做得更快、更全、更便宜、更不疲倦。

这就是“包办”这个词的确切含义，也是它与“颠覆”的区别：

不是被堵死了，是被走完了。路还在，终点已经被人先站住了；你带着一个班再跑一遍，跑得再好，也没有意义。

而这句话的另一半——那个空着的位置在哪里——不在这条链上。它在第一编第七章说过的、那个从未被制度化的第二个冲程里。第三编将说明那里正在发生什么，第五编将说明该怎么办。

七、可观察的读数

记录一节课里，从你抛出一个问题到有人给出答案，中间经过了多少秒。连续记五节课。

这个读数不需要任何设备，一支笔就够。它是全书最便宜、也最难作假的一个读数，第三十八章会把它列为四项常规读数之一。

多数教师第一次记录时会对结果感到意外。

八、一个必须澄清的误会：这里说的“发现式”不是探究式

本书反复使用“发现学教育”这个词，容易与教育史上一条具体的线索混淆，必须在这里划清，否则整本书会被误读为对某一场旧争论的站队。

中文里“发现式教学”通常指向一条主张：让学生通过探究自行发现规律，与讲授式相对。这条线上有一场持续了六十年的著名争论——一方用认知负荷理论论证“最少指导的教学为何不奏效”，另一方从问题式与探究式学习的实证一侧反驳。

本书批评的不是这条线上的任何一方。

本书所说的“发现学教育”，指的是一整套更深、更早、更少被质疑的前提：

知识是已经被发现的现成之物；教育就是把这些现成之物从持有者那一端搬到不持有者那一端。

在这个前提下，讲授式与探究式其实是同一件事的两种运输方式——一种是直接送货上门，一种是让学生自己走到货架前把货取下来。两者搬的是同一批货，货本身在这个过程中一个字都没有改变。

这也顺带解释了探究式教学在实证上的处境：它之所以经常跑不赢讲授式，恰恰因为它承担了运输成本的增加，却没有改变运输这件事本身。它是“更贵的搬运”，不是“另一件事”。

这三个动作在教育史上有很正式的名字。有人称之为“银行储蓄式教育”：教师存款，学生储蓄，考试取款。也有人用“接受学习”这个词来描述它，并且——这一点常被忽略——**为它做了辩护**，认为有意义的接受学习完全可以高效且深刻的。

本书不打算否认后者是对的：接受式学习确实可以很高效。本书要说的是，它高效地完成的那件事，恰好不是教育最要紧的那一件。

九、与相邻说法的分界

学习科学中有一族关于“合意困难”的研究：某些增加当下难度的做法，反而提高长期保持与迁移。本章与它们方向一致，但处理的对象不同：那一族研究的是**练习方式的设计**（间隔、交错、提取），其困难是被人为安排进去的；本章说的是一种**本来就存在、正在被无意中取消的困难**——没有人设计过“孩子得等三分钟”，它只是稀缺的副产品。这个差别有实践后果：设计出来的困难可以被重新设计回去，而被结构取消的困难，必须靠一条外加的规矩才能保住。第三十四章那条规矩，就是为此而设。

第十二章 加固悖论

老路一边被走完，一边显得越来越对；因为它的全部读数的是成品被接住了多少，而接管者恰好同时是史上最强的磨光机与最强的提早供给装置。

一、把悖论写出来

第八章的结论是：老路的全部读数，量的都是成品被接住的比例与质量。

第九至十一章的结论是：三个动作被逐项接管，而接管的方式全部是**提高成品的投递效率**。

把这两条放在一起，就得到本书的核心结构：

每一次接管，都会让老路的所有读数同时变好。

于是学校会得到一个正反馈：数据支持继续深化，深化之后数据更好，更好的数据支持进一步深化。整个系统在自我确证，而它确证的那条路，恰恰是已经不需要人走的那一条。

本书把这个结构称为**加固悖论**。

二、两台机器合成一台

要看清这个悖论的动力，需要注意接管者的两项能力恰好指向同一个方向。

第一，它是史上最强的光滑化装置。任何知识，无论多难、多偏、多零碎，它都能在几秒钟内磨成一个光滑、完整、没有毛刺的成品端到孩子面前。人类教师需要三十年练成的“把复杂的事情讲简单”，它一开始就有。第二章说过，光滑化是把过程变成成品的投影；而这台机器的训练目标，几乎就是这个投影本身。

第二，它是史上最强的提早供给装置。一个问题刚刚在孩子心里张开一条缝，缝还没长到能站住人的宽度，答案就已经到了。第十一章的四道摩擦全部被它取消。

一台机器同时具备这两项，意味着老路的两个关键量——成品的质量与投递的速度——同时被推到极限。而这两个量正是老路全部读数的自变量。

三、悖论的严重性在于它不可见

一个错误如果在读数上表现为变差，它会被发现并纠正。加固悖论的危险在于它的表现恰好相反。

设想一所学校认真推行 AI 辅助教学。一个学期后：正确率上升，进度加快，弱势学生被拉平，教师负担下降，学生满意度提高，家长评价改善。

按现行的全部标准，这是一次成功的改革。任何一位校长都会据此扩大投入，任何一位教研员都会把它作为经验推广，任何一份汇报材料都会把它列为成果。

而这些判断在各自的框架内全部正确。没有人撒谎，没有人失职，没有人被欺骗。

问题在于：所有这些读数，都在测量第一个冲程；而第二个冲程有没有发生，它们不问，因此也无从否定。第八章的结论在这里显出全部分量——**一个只能测出自身成功的读数体系，在环境改变时不会发出任何警报。**

四、于是有了本书最不受欢迎的那个判断

由此可以给出一个比“AI 使用不当”强得多的判断：

把 AI 用于提升老路的效率，不是不够好的那个选项，是有害的那个选项。

理由是效率有方向。在一条终点正确的路上提速，是进步；在一条终点错了的路上提速，不是原地踏步，是**加速偏离**。

而这条路的终点错在哪里？不在于它教的内容错了——内容没错。错在它的终点是“抵达一个已被确定的结论”，而这件事现在有更好的执行者。人跑到那里，不再产生任何价值。

需要立刻补一句以免误读：这不意味着孩子不需要掌握那些内容。他需要，而且比过去更需要。区别在于：**掌握内容不再是课堂时间的最佳用途**，因为那件事在课外三秒钟就能完成。

五、“工具是中性的”这句话为什么太轻巧

面对上述判断，最常见的回应是：AI 只是工具，关键看怎么用。

这句话作为常识没有错，作为判断则过于轻巧。工具从来不是中性的——**它有自己的构造倾向。**

锤子的构造倾向是敲；你当然可以用锤柄撬东西，但那是逆着它的构造用。一台系统的构造倾向是**填缝**：给它一个不完整的東西，它的全部训练都指向把它补完整。这不是它的缺陷，这是它被造出来要做的事。

让一台以填缝为天职的机器去守护缝隙，不是“用得好不好”的问题，是需要一套**专门的反向配置**。第十九章第六节给出这套配置的最小形态，第三十章至第三十二章给出课堂上的完整形态。

在没有这套反向配置的情况下，默认结果是确定的：**它会填缝，因为那就是它。**

六、悖论的时间形态

加固悖论在时间上有一条可以预期的曲线，第二十一章将详细处理，这里先给出形状，因为它解释了为什么这件事很难被及时发现。

第一阶段（约一学期）：显著变好。所有读数上升，且上升幅度往往超过预期。这一阶段最危险，因为它提供了扩大投入的全部理由。

第二阶段（约二至三学期）：增速消失。读数停在一个较高的位置不再上升。通常被解释为“红利期结束”“需要更深度的融合”，于是进一步加大投入。

第三阶段：某些东西不再随投入变化。具体表现是：复现类题目继续保持高位，而需要改变原有认识的题目不再随之上升；顶端那种“会问怪问题的孩子”变少。

第三阶段的特征是**不可逆**——加大投入不再改善它，而且减少投入也不会自动恢复。原因在第三编。

七、一个诚实的说明：本章尚未被证明

必须在这里说清本书自己的证据状况，否则后面各章的强 judgment 会失去分寸。

加固悖论目前是一个**有解释力的假说**，不是一个已被确立的事实。它解释了若干可观察到的现象（读数上升与某些能力表现的背离、补课的无效、顶端提问质量的下降），但这些现象也可以有别的解释。

本书为它设计了可证伪的检验，写在第三十九章：两组对照，其中一组把 AI 配置为答案源，另一组配置为对手；预测是前者在传统读数上不低于甚至高于后者，而后者在第三十八章的新读数上显著优于前者。

如果实验做出来，前者在两套读数上全面胜出，本章即被推翻。这个条件写在这里，也写在第三十九章，本书不为它保留任何逃生解释。

八、可观察的读数

把你们学校引入 AI 辅助教学前后的各项数据放在一起看。然后问一句：这些数据里，有没有任何一项，如果它变好了，能推出“孩子们更会自己想了”？

通常一项也没有。这不是数据不好——是这套读数从设计之初就不是为回答那个问题而造的。第八章已经说明了原因。

九、与相邻说法的分界

技术社会学中有一族关于“效率提升掩盖结构性代价”的经典分析：更快的交通带来更长的通勤，更多的劳动节约设备带来更多的家务时间。本章的结构与之相似，但有一个关键差别：那些案例中，代价最终会以可测量的形式出现（通勤时间、家务小时数），因而可以被发现；而本章的代价落在一个**从未被测量过的维度**上，且这个维度的缺失早于 AI 两千年。这意味着它不会自己浮现出来——**必须先造一把秤，它才可能被看见。**

第十三章 三个错觉：清晰度、参与度、个性化

这三件事都是真实的改善，而且都是老路上的改善；承认它们为真，与承认老路已被走完，并不矛盾。

一、为什么要专门写错觉

第十二章给出了机制，但机制在现场是抽象的。加固悖论在实践中不以“悖论”的形式出现，它以三种具体的、令人愉快的、而且完全真实的改善的形式出现。

指出它们是“错觉”，不是说这些改善是假的。是说：**它们被用于支持一个它们支持不了的结论。**

二、清晰度错觉

现象。教师看到 AI 把一个难点讲得比自己清楚得多，于是得出结论：教学质量提升了。

为什么是真的。它确实讲得更清楚。这一点第十章已经承认，不需要再辩。

问题在哪里。清晰度衡量的是**传输损耗**。而传输损耗只在需要传输时才是问题。当传输那一端已经被全面接管时，把清晰度做得更好，等于在一件不再稀缺的事情上继续投入。

更麻烦的一层。清晰度越高，缝隙关闭得越快。

一个讲得天衣无缝的解释，恰恰不给听的人留下任何一个可以插进去的位置。

这一句值得停一下。一段有缝的讲解——某处跳了一步、某处说得含混、某处让人觉得“好像不太对”——会在听者那里制造一个需要他自己去补的地方。而一段完美的讲解不制造任何这样的地方。

所以清晰度与第二个冲程之间不是正相关，很可能是负相关。这是一个可以被检验的预测，第三十九章将它列入设计。

三、参与度错觉

现象。引入 AI 之后，课堂上的提问次数往往大幅上升，尤其是那些平时不敢开口的学生。

为什么是真的。这是真实的、值得珍惜的改善。第十一章的第二道摩擦——社交摩擦——对一部分孩子是决定性的障碍，取消它是一件好事。一个从不敢举手的孩子第一次开始提问，这件事本身有价值。

问题在哪里。参与度这个指标不区分提问的性质。

一百次“这个怎么做”和一次“你这里不对”，在参与度上前者赢。

而这两类提问在本书的框架里分属两条路：前者是索取供给，后者是产出判断。指标不区分方向，是它最危险的地方——因为它会引导整个系统去优化数量。

第三十八章因此把“学生对 AI 说的第一句话”单列为一个读数。它测的正是参与度测不出来的那个方向。

四、个性化错觉

现象。AI 可以为每个学生定制难度、节奏与说法，这被普遍视为因材施教的实现。

为什么是真的。第三章说过，因材施教两千年来是理想而非现实，因为它要求把成本推回到一对一那一侧。现在它第一次可以实现，这是教育史上的一件大事，不应被轻视。

问题在哪里。因材施教在老路上的含义是“为每个人找到损耗最小的传输方式”——它优化的仍然是接受，不是发生。一个被完美个性化的接收过程，仍然是一个接收过程。

而它比统一授课更彻底，因为它连“听不懂”这个偶然产生的缝隙都事先消除了。第十一章第五节那句话在这里可以重复一次：

| **听不懂原本是一个岔口，个性化把它填上了。**

一个需要澄清的地方。这不是主张应当让孩子听不懂。制造困难不是目的，被制造出来的困难通常是坏的困难（第十一章第八节的分界已经说明这一点）。主张是：当唯一的入口被消除之后，必须另外造一个入口，而不能假定它还在。第三十四章的四段结构就是那个另外造出来的入口。

五、三个错觉的共同结构

把三条并排看，一个共同结构浮现出来：

错觉	真实的改善	被消除的东西
清晰度	讲得更明白，损耗更小	听者需要自己补的那个位置
参与度	更多孩子敢开口了	提问方向的区分
个性化	每个人都跟得上	“听不懂”这个偶然的入口

三行合起来：**每一项优化，都在消除某种粗糙；而某些粗糙恰恰是第二个冲程的唯一入口。**

这句话是第二编与第三编之间的桥。第三编要问的是：入口全部消失之后，会发生什么。

六、第四个候选：进度错觉

上面三个是最常见的。这里补一个较少被讨论、但在学校层面影响更大的。

现象。引入 AI 之后，同样的课时可以覆盖更多内容，于是进度加快，教学计划提前完成，腾出的时间用于更多的练习或更多的内容。

为什么是真的。这确实是效率提升，而且是最容易被量化、最容易被写进汇报的一项。

问题在哪里。腾出来的时间**被自动用于同一件事**。没有人做过一个决定说“我们要用更多时间讲更多内容”，这是默认发生的：因为教学计划是按内容排的，效率提高的自然结果就是排更多内容。

而第八章第六节已经指出，第五编方案在时间上的全部可行性，正来自这一段被腾出的时间。**它现在被自动消耗掉了。**

所以第五编在学校层面的第一个动作，不是增加课时，而是**保护这段已经被腾出来的时间不被自动填满**。这一条比任何具体教法都更前置，也更容易被忽略。

七、可观察的动作

在你的课上做一次对照：同一个内容，一半用最清楚的方式讲完，另一半故意留一处不讲透。第二天各问一个需要用它去改旧认识的问题，看哪一半答得好。

这个对照很粗糙，样本小，混杂因素多，不构成证据。但它会让做过的人对“清晰度”这件事产生一个具体的怀疑，而这个怀疑比读十页论证更有用。

八、与相邻说法的分界

有一族关于“流畅性错觉”的研究指出：学习材料越流畅易读，学习者对自己掌握程度的判断越高，而实际保持效果未必更好。本章的清晰度错觉与之高度相关，但处理的层面不同：那一族研究关心的是**学习者的元认知判断偏差**（他以为自己会了），本章关心的是**教学系统的评价判断偏差**（整个系统以为它做得更好了）。前者可以靠告知学习者来缓解，后者不能——因为它不是认知偏差，是读数覆盖面的结构缺口。第三十八章的处方因此不是“提醒大家注意”，而是造新读数。

第十四章 樱花毒药：所谓“AI 高效自学”

樱饼的那股香，就是叶子里的防御物质；香和毒是同一样东西。“AI 高效自学”那股令人愉快的“高效”，不是这套做法的副产品，恰恰就是那件有害的事情本身。

一、这个比喻不是修辞

加固悖论有一个最纯粹、也最受推崇的形态，正在被当作 AI 教育的最优解四处推广：**AI 高效自学**。

本章对它的判断是：它是一剂樱花毒药。

先说清楚这个比喻的每一层，因为它不是“好看的东西有害”这种泛泛之谈。

第一层，它确实好看。一个孩子坐在那里，随时可以问，问什么答什么，答得清楚、耐心、恰到好处，进度由他自己掌握，不懂就换一种说法，不必举手，不必怕丢脸，不必等老师有空。任何一位家长看到这个画面都会觉得欣慰。这不是伪装出来的好看，这是真的好看。

第二层，也是这个比喻真正有意思的地方：樱花的香，是从细胞被破坏的那一刻才出现的。新鲜樱叶几乎没有味道，香气的前体被包裹在细胞里；只有在盐渍、细胞结构被破坏之后，酶与之反应，才生成那股为人熟知的香甜气味。而这股香气的来源物质——香豆素——正是这类植物用于防御的次生代谢物，在足够剂量下具有肝毒性。同属植物的种子与叶片还含有氰苷。

也就是说：香和毒是同一类东西，不是“香的东西碰巧有毒”。人们把它盐渍半年，包在点心外面，把那股防御物质当成风味来享用。

这正是“AI 高效自学”的结构。

二、高效就是剂量

把上面那句话讲实。所谓“高效”，在操作上是什么？

它是未闭合时长趋近于零：孩子产生疑问到拿到答案之间的那一段，被压缩到几秒。这就是“高效”这个词的全部含义，没有别的含义。

而第十六章将说明，第二个冲程恰恰只发生在那一段里。那一段被压到零，它就没有发生的地方。

于是这个词组自带一把剂量尺：

效率就是剂量。越“高效”，剂量越大。

这是一个不太舒服但极其清楚的推论。在别的领域，我们总希望效率越高越好；在这一件事上，效率的提高与损害的加重是同一个量。任何一份宣传材料里出现“学习效率提升百分之三十”，都可以直接读作“剂量提高百分之三十”。

三、毒在“自学”这两个字的当下含义

需要立刻堵住一个误读：本书不反对自学。人类最好的学习者大多是自学者，自学在教育史上从来是褒义词。

问题出在“自学”这个词在新条件下发生的语义位移。

传统的自学是什么样子？一个人对着一本书、一个问题、一段自己也不确定的推理，长时间地卡住、绕路、翻回去重看、想岔了再退回来。传统自学之所以有效，恰恰因为它低效——没有人可问，供给稀缺，于是所有的岔口都被迫由自己走完。

“AI 高效自学”把这个结构整个翻了过来：供给无限，摩擦为零，卡住的时间被压到几秒。名字没变，内容换了。它已经不是自学，是**独自接收成品**。

用第五编的语言说：在这个模式里，三个位置全部塌进了一方。现成的说法在它那里，路径也在它那里，连“哪里不懂”都由它来判定和拆解。孩子名义上占据了全部三格，实际上一格也没占——他只是那个坐在旁边看着它跑完的人。

而碰撞至少还需要一个对手。自学模式里没有对手。这是它与课堂形态最本质的差别，也是下一节要展开的那件事。

四、同一条路，不同的剂量

到这里可以把判断压成一个式子。但写这个式子之前，必须先拆掉一个会当场被人抓住的假矛盾。

有人会这样反问：既然“AI 高效自学”就是发现式教育本身，那它怎么可能“比发现式教育更毒”？同一样东西怎么会比自己更毒？

这个反问站不住，但它指出了必须写清楚的地方：**这里有两个不同性质的判断，不能用同一个等号串起来。**

第一个是同一性：**AI 高效自学 = 加速的发现式教育**。它没有另立一条路，三个动作一个不少，终点仍然写在别处。

>

第二个是比较：**加速的发现式教育 > 课堂形态的发现式教育**。同一条路，剂量更高、杂质更少。

矛盾消解在一句话上：**同一条路，不同的剂量**。路没有变，变的是路上的摩擦。

由此得到本章的形式化表达：

毒性 = 老路 × 速度 ÷ 摩擦

老路是前提项而非变量：若这一项为零——即这条教学根本不以“抵达一个已定结论”为终点——则后面乘什么都是零。这正是为什么本书的处方不是降速，而是换道。

速度即未闭合时长的倒数，第二节已说明。

摩擦是解释“为什么自学模式更毒”的关键。

五、课堂自带的五道减毒装置

课堂形态的发现式教育带着五样天然的摩擦。它们在教学论上历来被视为需要克服的缺陷，而在本书的框架下，它们恰恰是**减毒装置**。

第一道：等待。四十个孩子一个老师，卡住的多半只能等。等是难受的、低效的、历来被视为班级授课制的固有缺陷。但那段等待，正是缝隙张开并维持的时间。自学模式把它压到几秒。

第二道：别人。同学答得跟他不一样，这是一份免费的、他并没有打算要的异议。它不是任何人设计出来的，是班级这个形式的副产品。独自面对时只有一个声音，而且这个声音已经针对他的水平做过定制——连“听不懂”这个岔口都被事先消除了。个性化就是提纯。

第三道：一个可能忍住的人。教师会忍住，会反问，会说“你先想想”。这个动作反自身职业本能，需要训练，但它至少是可能的。自学模式里没有任何一方有动机忍住。第三十二章那条“闭合权不参与轮换”的约束，在自学模式里彻底失效——**闭合权也一并交出去了。**

第四道：公开性。课堂上他必须把想法说出口，说出口就要面对反驳。私密场合不需要——他可以在心里“觉得懂了”，而这个状态没有任何外部检验。第十六章的回写判据在私密场合失去了观察者。

第五道：边界。课时是硬边界，下课铃是外部约束。自学模式的时长由孩子自己决定——更准确地说，由产品的留存设计决定。而使用时长在这类产品的评价体系里，是唯一一个越大越好的指标。

五条合起来是一句话：

课堂是一个减了毒的老路；自学模式是一个纯的老路。同样的路，去掉了全部杂质。

摩擦项	课堂形态	AI 自学模式
等待	卡住要等，几分钟量级	压到几秒
异议	同学答得不一样，免费反例	单一声音，且按他的水平定制
忍住的人	教师可能忍住，可训练	无一方有动机忍住
公开性	必须说出口，要面对反驳	心里“觉得懂了”即可
边界	课时与下课铃	由留存设计决定

六、谁在吃：一条比剂量更重的推论

上一节讲的是每一剂有多少。还有一项讲的是**谁在服用**，它比剂量更重要。

课堂形态的损害是均摊的：全班一起上课，一起接收成品。而主动使用 AI 自学模式的，是一个高度选择性的子群——**最自觉、最上进、家庭最重视教育、自我要求最高的那批孩子**。他们不是被安排去用的，是自己找上去的，而且用得最久、最投入、最认真。

把这一点接到第十七章的机制上：那把锁不是被失败写下的，是被**一次次成功的接受**写下的；接得越顺、越高效、越少挫折，那条元规则加固得越硬。因此：

这剂药，优先被喂给了那批最有可能开辟新路的孩子。

课堂形态造成的是一层薄而广的损害；自学模式造成的是一处深而窄的损害，而且恰好落在人群中最有潜力的那一段上。从个体角度看这是不幸，从文明角度看这是更严重的事：**新路的开辟本来就依赖极少数人，而这一小群人正是暴露剂量最高的人。**

这也解释了一个已经可以观察到的现象：引入 AI 学习工具之后，班级内部的成绩差距往往缩小——弱势学生被迅速拉平，这是真实的、值得肯定的改善——而**顶端那种“会问怪问题的孩子”变少了**。前者上得了报表，后者上不了。

必须补一句防止误读：这不构成“应当禁止优秀学生使用 AI”的理由。这条推论指向的处方与第八节完全一致——给自学装上对手与回写口，而且**优先给这批孩子装**，因为他们暴露得最多。

七、欺骗性在哪里：一个名词的偷换

“欺骗”这个指控必须被坐实，否则只是情绪。

欺骗不在形容词上。宣传里说的每一项都是真的：更快、更懂你、更耐心、随时在、按你的水平走。没有一项是假的。

欺骗发生在名词上。被偷换的是“学习”这两个字：

宣传方所量的“学习” = 接收成品的效率。

家长与教师所理解的“学习” = 这个孩子变强了。

第五章第五节已经指出，这个偷换不是 AI 带来的——它早就完成了，因为在过去这两件事高度相关，相关到可以互相替代。AI 做的事情是把这条相关性**打断**：接收效率可以趋近于无穷，而“变强”可以为零，甚至为负。

由此得到一个当场可用的检验：

拿任何一份 AI 学习产品的宣传材料，把其中的“学习”两个字替换为“接收成品”。

句子若仍然成立，那它宣传的就是加速的发现式教育。

这个替换几乎总是成立。“让学习更高效”→“让接收成品更高效”，成立；“随时随地学习”→“随时随地接收成品”，成立；“个性化学习路径”→“个性化的成品接收路径”，成立。

一处必须收紧的措辞。不能说“所有 AI 自学产品都具有欺骗性”。理由不是留情面，而是这个全称判断一旦被一个反例击中，整章都失效——而反例是现成的，并且正在出现。可辩护的写法是把全称判断换成判据：

凡是以“更快、更省力、更不用卡住”为卖点的 AI 学习产品，宣传的都是加速的发现式教育。

判据不看它用了什么模型、什么技术路线，只看它把哪个指标当作卖点：**卖“省时间”的，卖的就是剂量。**

这个改法比原来的全称判断更锋利，因为它不是对一个行业的道德指控，而是一把当场可用的尺子。同时它留出了正确的出口：一个把卖点改成“让孩子多卡住一会儿”的产品，在逻辑上是可能的，只是在商业上很难。**这个困难本身，就是这件事最值得警惕的地方。**

八、为什么是毒药，不是弯路；以及解药

这是选择“毒药”这个重词的理由，必须讲准，否则会被当作情绪表达。

弯路和毒药的区别在于**可逆性**。弯路走错了可以回头，回头的能力本身没有受损；毒药改变的恰恰是回头的能力。

按第十七章的机制：每一次零摩擦地拿到成品，都在加固“东西是现成的、我不必因它而改”这条元规则；元规则一旦写死，通道关闭。此后即使把最好的老师、最好的课、最完整的路径还给他，也写不进去了。

所以“AI 高效自学”损害的不是某一段知识，而是**此后一切学习的接收结构**。它是毒药，不是弯路。

解药不是禁用，是加装两个部件。

一是**对手**。规定孩子在得到答案之前，必须先向它陈述自己现在的想法，并要求它指出问题所在；把默认的提问句式从“这个怎么做”改成“我是这么想的，你说我哪里不对”。这一条可以直接写进使用规范，甚至可以由产品端实现。

二是**回写口**。每一次自学结束时，必须回答一句：“我原来以为什么？现在这句话要改成什么？”没有这一句，这一次自学不计入完成。

装上这两样，“自学”这个词才回到它原来的含义。装不上，它就只是一个人独自吃下一整盘樱饼。

九、潜伏期与可证伪的剂量假设

最后给出可检验的部分，否则上面的话就只是一个说法。

潜伏期。这剂药的显效是延迟的。可预期的曲线是：先涨（第一学期，正确率与进度显著上升），再平（第二至三学期，增速消失），再不可逆（此后即使加大投入，迁移题成绩不再随复现题成绩上升）。三段之中，第一段最容易被当作成功的证据并据此加大投入，而加大投入正好加大剂量。

剂量效应假设（可证伪）。在控制初始水平之后，一个学生每周使用 AI 独自答疑的时长，应当与其先在感指数的上升幅度正相关，与其回写率的下降幅度正相关，而与其复现题成绩正相关。这三条相关同时成立，才构成本章的证据；其中任何一条方向为反，本章的主张即被推翻。

关键判别项。同样的使用时长，若把提问形态从索取型改为对抗型，上述相关应当显著减弱甚至消失。这一条最重要：它说明**毒不在 AI，而在提问的方向**，也直接给出了上一节的解药。

十、可观察的读数

采集孩子向 AI 发出的第一句话，分成两类：**索取型**（“这个怎么做”“答案是什么”）与**对抗型**（“你这里不对”“要是换成……你还成不成立”）。

这个读数完全不依赖孩子配合，因为它就是对话记录本身。第三十八章将把它列为四项常规读数之一。

十一、与相邻说法的分界

关于“屏幕时间”“数字成瘾”“注意力碎片化”的讨论与本章表面接近，实则不同层：那一族关心的是**注意力资源的损耗**，其机制是分散与刺激；本章关心的是**接收结构的改变**，其机制是摩擦的消除，而且它恰恰发生在最专注、最投入、最不分心的使用情境里。**一个孩子全神贯注地用 AI 学习两小时，在那一族的框架里是模范，在本章的框架里是最高剂量**。这个反差说明两者不能互相替代。

第十五章 高速公路与两个孩子

樱花毒药回答“为什么有害”，高速公路回答“为什么不够”——接通已有路网只能抵达已被命名的地方，而新知识按定义在命名域之外。

一、一个毒药回答不了的问题

第十四章处理的是损害。但还有一个独立的问题它回答不了：即使无害，够不够？

设想一种理想的情形：孩子使用得当，装上了对手与回写口，没有任何封写发生，他扎扎实实地掌握了大量内容。这够吗？

本章的回答是不够，理由与损害无关，是一个结构问题。

AI 高效学习，是把人快速送上一条早就修好的高速公路，而不是在没有路的地方开出一条路。

这个比喻切的是另一刀：前者是毒理学，讲剂量与不可逆；后者是拓扑学，讲能到哪里。

二、高速公路只连接已经有名字的地方

一条公路的两端，必须都已经有城市、有人烟——否则没有人会去修它。所以接通公路网所能抵达的全部地方，是已被命名的地方。

而新知识，按定义在命名域之外。

由此得到一个与“效率”完全不同的结论：这不是速度问题，是可达域问题。再快也到一个还没有名字的地方。一个孩子在公路网上跑了一整年，跑遍全国，他仍然一寸新路都没有修过——他到过的每一个地方，在他出发之前就已经在地图上。

这一条也说明，用“效率”这个词讨论教育从一开始就问错了：效率衡量的是同一张路网上的移动速度，而教育真正要产出的东西，是路网本身的生长。两者不在同一个坐标系里。

三、修路的人不在车上

公路是别人修的。修的时候有勘测、有爆破、有改线、有报废的方案、有走不通退回来的路段。坐车的人一样都看不见。

这正是“发现=发生—路径—土壤”的日常版本：你不是走过来的，你是被送过来的。沿途的地形被隔音墙挡住，出口牌直接告诉你终点叫什么。

过去这一点不太要紧，因为上路本身还要花力气——查资料、找人问、试几次错。这些摩擦不是好东西，但它们逼出了一些野路：恰恰是“不方便”，让一部分人不得不自己走一段。

现在匝道费降到零，野路不再被走——不是被禁止，是没有人再有理由走。

路网的萎缩从来不是因为封路，是因为没人走。

四、“新路线”不是“新道路”

会有人反驳：AI 也能组合、也能给出前所未见的方案，这难道不算开路？

区分只有一条：这条“新”路上，有没有一段路面是这一次才铺出来的？

把两条已有的公路接起来，绕出一条从没人走过的路线，那是**新路线**，不是**新道路**——它的每一段都能追溯到某条已经存在的路。这与第三十三章将给出的判据是同一条：如果结论的每一段都能追溯到某一方带进来的东西，那是拼接，不是发生。

新道路的标志只有一个：有一段地方，原来没有路面。

五、最危险的那个推论

一个从小只坐过车的人，不只是“不会修路”。他会慢慢形成一条更根本的信念：

凡是没有路通到的地方，就是不存在的地方。

这就是封写在这个比喻里的形态。它不表现为“我不会”，而表现为“那儿没东西”。

检验很便宜：给他一道真正没有现成答案的题——不是难题，是**空位题**，一道全世界都还没有答过的题——看他的反应。他会说“我搜不到”“它也不知道”，然后停下来。停下来的那一刻，不是他不够聪明，而是他把“没有路通到”读成了“不存在”。

一个走过野路的孩子在同一处的反应不同：他会朝那个方向走两步看看。**只有两步，但性质完全不同。**第三十八章将把这个反应列为第五个读数。

六、两个孩子

高速公路的比喻讲清了结构，但它缺一样东西——人。而教育上的判断最终要落在人身上，因此这里给出一对更具体的形象，它在教师培训现场比前一个比喻更好用。

第一个孩子在一座农庄里待了一年。他自己开垦，一年下来在庄子里开出了几条路，育出了几个新品种。路不多，也不通往任何一座城市；品种不好看，产量也谈不上。但有一件事是确定的：**那几条路和那几个品种，在他动手之前，世界上没有。**

第二个孩子这一年坐飞机看了五十个城市，每个城市待两天。他回来说他认识了全世界。而且他常常来农庄，给第一个孩子讲世界有多丰富多彩——讲得生动、准确、引人入胜，一句谎话都没有。

为什么这一对更锋利？因为第二个孩子在任何一场比较里都会赢：他见得多、说得好、走得远、拿得出照片；而第一个孩子在场面上一定显得土、慢、狭隘。任何一个评委席、任何一份汇报材料、任何一次家长会上的当众比较，都会选中第二个孩子——而这正是本书全部担忧的日常形态。

它同时把前面几章的三件事压缩成了一个画面：可达域与到达速度的区别；新道路的标志（品种不是被发现的，是被育出来的）；产出物性质的不同——第二个孩子带回来的是**见闻**，随他的记忆一起消退；第一个孩子留下来的是**东西**，可以被别人拿走、种下去、继续育。

七、必须先讲的挡板：不能贬低旅行

这一对比喻有一个极易翻车的地方，必须在使用它之前处理，否则整段会被听成“劝孩子不要见世面”，而提出反驳的往往恰恰是最优秀的那批教师。

所以要先把话让出去：**去五十个城市本身是好事**。问题不在他去了，在于他只去过，以及他回来时相信自己已经认识世界了。坏的不是飞机，是**只有飞机**。

由此得到这个比喻真正的对照式：

一个有农庄的孩子去看五十个城市，和一个没有农庄的孩子去看五十个城市，是两件完全不同的事。

有农庄的那个孩子，每到一处，看的是别人的土怎么翻、别人的种怎么育、别人的水渠为什么那样拐；他带回去的东西**能种下去**。没有农庄的那个孩子，带回来的只有照片。

同一次旅行，同样的五十个城市，同样的两天，落到两个人身上是两种完全不同的事件。**决定它是哪一种的，不是旅行本身，是他有没有一块自己的地。**

这也给出了本书对 AI 使用的最终态度，比“装两个部件”更根本一层：**先有农庄，再上飞机**。次序反了，飞机就只是飞机。

八、“两天”这个数字

这一对比喻里最锋利的细节，其实是那个不起眼的时长。

两天，正好够看完一个城市所有已经被写进导览手册的东西，一样也不够看到任何还没被写进去的东西。

两天是一个恰好落在成品层的时长。够走完全部景点，不够遇到一件只有当地人才知道的事；够拍完所有该拍的照片，不够跟任何一个人真正吵一次架；够记住这座城市的名字和它的三个标签，不够发现这三个标签哪一个其实是错的。

于是“两天”与第十四章第二节接上了：**两天，就是城市尺度上的那几秒**。未闭合时长在一节课上是几秒，在一次旅行上是两天，在一个学期上是一门课的进度表。尺度不同，机制相同——都是**恰好够接住成品，恰好不够遇到任何一处不成品的地方**。

现场很有效的一个反问是：如果给他五十天只看一个城市，他会不会反而更认识世界？多数人会点头。而这一点头，就等于承认了本书的全部主张——**认识的深度不与到达的数量同向**。

九、最重的一笔：农庄是被说服的

这个比喻的结尾——“经常给农庄的孩子讲解世界的丰富多彩”——是全段最重的一笔，因为它回答了一个前面各章都没有正面回答的问题：**为什么走野路的孩子会自己放弃？**

不是被大人逼的，不是被制度禁止的，不是被谁拆掉的。是他听着听着，觉得自己那几条破路、那几个歪瓜裂枣的品种，实在拿不出手。

农庄不是被拆掉的，是被说服的。

而说服他的那个人完全真诚。第二个孩子没有撒谎，他确实见得多，讲得也确实精彩，他的每一句描述都经得起核对。唯一不真的，是“我认识了世界”这五个字——而这五个字恰恰不是描述，是一个关于自己的判断。

这与第十四章第七节的结论完全同构：欺骗不在形容词上，在名词上；不在他讲的内容里，在他给自己的定性里。而这个定性会传染——它传染的不是知识，是一种关于什么才算“懂了”的标准。第一个孩子被说服的，正是这条标准。

十、三十年后

最后需要把时间轴拉开一次，否则这个比喻在现场会留下一个未回答的疑问：农庄那个孩子不是很吃亏吗？

短期看，他确实吃亏，而且吃得非常明显——一年之内、三年之内、直到升学，他在所有可比较的场合都会输。

而三十年后：

第二个孩子讲不出任何一座城市里别人不知道的事；而第一个孩子的那几个品种，别的农庄开始来要种子。

这个不对称只有把时间拉长才看得见，而这恰恰是问题所在：现行的一切评价周期，都短于这个不对称显形所需要的时间。所以体系不是选错了人，是它在结构上没有能力选中第一个孩子——它的观测窗口太短。

这也就把话交回给了教师。在一个观测窗口如此之短的体系里，唯一还能在长周期上做判断的，是站在教室里的那个人。他是这个系统里唯一有可能不用当期读数下判断的人。这是负担，也是这份职业在 AI 之后仍然不可替代的另一层理由。

十一、两个比喻的使用纪律

最后交代一处使用纪律，因为这两个比喻很容易被混着讲，而混讲会互相稀释。

樱花毒药回答“为什么有害”，高速公路与两个孩子回答“为什么不够”。

前者是负面的、要顶人的，宜放在处方之前；后者是正面的、不刺人的，宜放在换道论述的开头作过渡。在教师培训的实际排布中，建议的顺序是：先讲两个孩子（老师容易接、会笑、会松），再回到樱花毒药（要顶人的那一段），最后进处方。顺序反了，第一记重锤落地时会场会转向辩论，后面全部失效。

十二、第二编的结论

第二编到此结束，四句：

第一，三个动作被逐项接管，接管的方式全部是提高成品的投递效率；这条链上不再有一个位置是非人不可的。

第二，因此老路的全部读数会同时变好——这就是加固悖论；在一条终点错了的路上提速，是加速偏离。

第三，这个悖论最纯粹的形态是“AI 高效自学”：高效即剂量，课堂自带的五道减毒装置被逐一拆除，而服药者恰好是最有可能开辟新路的那批孩子。

第四，即使无损害，它也不够：接通已有路网只能到达已被命名的地方。而最重的一笔是，农庄不是被拆掉的，是被说服的。

第三编要问的是：这些事情发生在孩子身上，究竟改变了什么。

十三、可观察的动作

给一个孩子一道“没有标准答案、但可以走两步”的题，看他停止尝试之前有没有朝那个方向走出过一步。

不看那一步走得对不对，只看有没有。这一动作在第三十八章被列为读数五，此处先给出它的原始形态。

十四、与相邻说法的分界

创新研究中有关于“探索与利用”的经典权衡：组织必须在开发已知与探索未知之间分配资源。本章的可达域问题与之形状相近，但对象不同：那一族处理的是资源如何在两类活动之间分配，其前提是两类活动都在选项集里；本章处理的是未知那一类是否还在选项集里——一个持有可达域信念的人不是把资源全投给了利用，是他不认为存在可探索的方向。分配问题与选项集问题不是同一层。

第二编材料：一份学期数据画像、一段对话的逐轮标注与二十条宣传语

本编材料的性质：数据画像为按第十二章机制推演的示例形态，不是任何一所学校的实测结果；对话为示例构造；宣传语为常见句式的归纳改写，不指向任何具体产品。

材料四 一所学校的学期数据画像（推演示例）

第十二章说加固悖论在读数上表现为全线变好。这里给出它可能长成的样子，供学校拿自己的数据对照。这不是实测，是按机制推出的形状。

指标	引入前	一学期后	变化
当堂正确率	68%	79%	↑ 明显
单元测均分	76	82	↑
后进面（低于 60 分）	14%	7%	↓ 明显
单位时间覆盖内容	基准	+18%	↑
学生满意度	3.9/5	4.4/5	↑
教师备课时长	基准	-25%	↓
迁移题正确率	41%	43%	持平
迁移减复现之差	-27	-36	扩大
未闭合时长中位数	9 秒	4 秒	减半

读法。前六行会进汇报材料，后三行通常不被采集。而全书要说的事情，只写在后三行里。

关键在最后两行的方向：迁移与复现的差距**扩大**，未闭合时长**减半**。这两条如果同时出现，第十二章的机制在这所学校里成立；如果迁移题随复现题同步上升、且未闭合时长没有下降，那么这所学校不构成本书所担心的情形。

材料五 一段自学对话的逐轮标注（示例构造）

第十四章说自学模式里三个位置全塌进一方。这里把它拆到轮次上，看它是怎么发生的。左栏是对话，右栏是标注。

轮次	内容	标注
1	学生：“这道题怎么做？”	索取型 。未闭合时长≈0
2	系统：给出完整解法，分四步	同时占 显露+路径
3	学生：“第三步没看懂”	仍是索取型；卡点被它定位
4	系统：把第三步再拆成三小步	占 土壤 （连“哪里不懂”也由它判定）
5	学生：“懂了，谢谢”	三格全占；学生零产出

全程用时约九十秒。全程零回写。

同一道题的另一种走法（示例构造）：

轮次	内容	标注
1	学生：“我觉得应该这样做……你说我哪里不对”	对抗型 。学生先出旧账
2	系统：指出第二步的问题，不给完整解法	只占 异议
3	学生：“那如果换成……还成不成立？”	学生占 路径
4	系统：给出一个反例	仍占 异议
5	学生：“那我上次那道题是不是也错在这里”	回写发生

全程约六分钟。差别不在系统，在第一句话。

第五轮那一句正是第十六章的判据形态——**回头改一件旧的**。它在第一种走法里从未出现，也不可能出现。

材料六 二十条宣传语的尺子实测

第十四章第七节给出一把尺子：把“学习”换成“接收成品”，句子若仍成立，卖的就是老路。这里对二十条常见句式做一次实测。**句式为常见表述的归纳，不指向任何具体产品。**

原句	替换后	是否成立
让学习更高效	让接收成品更高效	成立
随时随地学习	随时随地接收成品	成立
个性化学习路径	个性化的成品接收路径	成立
一对一名师辅导	一对一成品投递	成立
哪里不会点哪里	哪里没接住补哪里	成立
学习效率提升三倍	接收速度提升三倍	成立
帮孩子快速掌握知识点	帮孩子快速接住成品	成立
秒懂难题	秒接难成品	成立
智能诊断薄弱环节	诊断哪些成品没接住	成立
让孩子不再畏惧提问	让孩子不再畏惧索取	成立
减少无效学习时间	减少无效接收时间	成立，且这句最值得警惕
陪孩子多想一会儿	—	不成立
先说说你的想法	—	不成立
我们不直接给答案	—	不成立

读法。前十一条全部成立，说明它们卖的是投递效率。第十一条尤其值得注意：“无效学习时间”这个说法，恰好把本书认为唯一有效的那段时间列为了要消灭的对象。

后三条不成立，说明存在另一种可能的卖点。第十四章第七节说过：一个把卖点改成“让孩子多卡住一会儿”的产品，在逻辑上是可能的，只是在商业上很难。这个困难本身，就是这件事最值得警惕的地方。

材料七 五道减毒装置的现场核查表

第十九章给了检查表，这里给出可直接使用的核查格式。对任意一种教学安排，逐项打勾。

装置	核查问题	有	无
延迟	从疑问到供给，中间有没有一段时间	☐	☐
异议	有没有一个不是被安排出来的不同说法	☐	☐
克制	在场有没有一方有能力且有动机不给答案	☐	☐
公开	想法必须说出口吗	☐	☐
边界	这段过程有没有外部规定的起止	☐	☐

参照（示例判读）：传统讲授课通常三到四项在位；探究式课堂五项基本齐备；AI 辅助的个性化课堂通常只剩边界一项；自学模式五项全无。

一条使用提示：这张表评的是安排，不是教师。同一位教师在两种安排下会得到完全不同的结果，而这正是这张表的用处——它把问题从“谁做得好”移到“什么样的安排会自己产生好结果”。

第三编 损害

第一编说明成品是怎么造出来的，第二编说明三个动作如何被逐项接管、以及这一接管为何在读数上显得越来越正确。第三编处理最难看见的那一部分：这些事情落在一个具体的孩子身上，究竟改变了什么。本编给出第二个冲程的判据（回写）、它的失效机制（封写）、四种死亡的诊断次序、损害的剂量式与分布形态、时间上的三段曲线，以及最终的那条可达域信念。

第十六章 回写：发生完成的唯一判据

一次学习是否真的完成，不看他多了什么，看他原来的东西变了没有；判据不是复述新的，是回头改一件旧的。

一、为什么必须有一个判据

第一编第七章指出，第二个冲程从未被制度化，原因之一是它没有读数。第八章进一步指出，凡是没有读数的东西，在这套系统里会被自动筛除。

所以本书的第一件正事，不是提出理念，是造判据。没有判据的主张会在一个学期之内被筛掉，无论它多么正确。

这个判据必须满足四个条件，否则造了也没用：

可当堂观察——不需要事后测量，不需要专门工具。

不依赖学生配合——不能靠问他“你懂了吗”。

不依赖教师的主观感觉——不能是“我觉得他眼睛亮了”。

低成本——不占用额外课时，不增加负担。

本章给出的判据满足这四条。

二、判据：回写

一个孩子接住了一样东西，可能发生两种情况。

第一种：他多了一条知识。原有的那些东西原封不动，新的这一条挂在旁边。这种情况下，他可以复述它、可以用它做同型的题、可以在考试中得分。

第二种：他原来的东西变了。新得到的东西反过来改写了 he 已有的底子——他原来那样想，现在不那样想了；他原来写下的一句话，现在要加几个字；他上周做的那道题，现在他知道那不只是算错了。

本书把第二种称为**回写**：新认知物反过来改写学习者原有的底盘。

回写是发生完成的唯一判据。不是“多了一条”，而是“原来的不一样了”。

三、回写有一个朴素的外部表现

回写听上去抽象，但它有一个极其朴素、任何教师当堂都能看见的表现：

他回头去改一件旧的。

不是复述新的，是回头修改自己以前写过、说过、做过的某件事。

—他把笔记上那句“三角形内角和是一百八十度”改成“在平面上，三角形内角和是一百八十度”。

—他翻回上周那道做错的题，说“我知道我那时候为什么这么想了”。

—他说“那我以前那样讲是不是也不对”。

这三种表现有一个共同点：**改的对象是旧的，改的动作是他自己发起的。**

由此得到本书在教师培训现场最常用的一句大白话：

一条道跑完，孩子“知道了”；另一条道跑完，孩子自己变了。

四、必须立刻堵住的一个漏洞

“变了”这个词有一个必须当场堵住的误读，否则整个判据会失效。

它会被听成**心情变了**——听懂了、喜欢上了、眼睛亮了、课堂气氛特别好、孩子们特别投入。

这些都是好事，值得珍惜，但它们不是判据。所以要补一句：

变的不是他的心情，是他往后看别的东西的眼光。

区分方法很简单：心情是当下的、会消退的、不产生任何可检查的痕迹；回写会产生痕迹——那句被改的话、那道被翻回去的题、那个被推翻的说法。

有痕迹的才算，没痕迹的不算。这一条看上去苛刻，但它正是把判据从主观感觉里救出来的那一刀。

五、回写不是“学会了”的同义词

需要把回写与几个相近的概念分开，否则它会被当成一个换了名字的老词。

回写不等于掌握。一个孩子可以完全掌握一个内容而没有任何回写——因为这个内容与他已有的东西不冲突，只是并列地增加了一条。大量事实性知识属于此类，这没有问题。

回写不等于理解。理解通常被描述为一种状态：连接有多少、能不能解释、能不能举例。而回写是一个**事件**：有没有发生过一次改写。状态可以缓慢积累，事件有或没有。

回写不等于应用。能把方法用在新题上，说明这条方法被接住得很牢，但不说明他原有的任何东西变了。

三条分完之后，回写的位置就清楚了：它不覆盖学习的全部，它只覆盖第二个冲程。第一个冲程该测的仍然照测——本书从未主张废除原有的读数，主张的是增加一把。

六、什么样的内容会引发回写

不是所有内容都能引发回写，这一点必须讲清楚，否则教师会试图在每一节课上都要求它，然后失败。

回写发生的必要条件是：**新东西与旧东西之间存在冲突。**

没有冲突就没有改写的必要。而冲突有三种典型形态：

形态一：旧规则在这里失效。分数就是例子——孩子用“比较两个数”的旧规则去处理一个新种类的对象，规则失效。若这个失效被摆出来，回写就有了发生的地方；若直接给出定义，冲突被绕过了。

形态二：旧结论有一个从未说出的前提。内角和是例子——他原以为这是三角形的性质，现在发现它是平面的性质。这类回写的表现极其固定：**加限定词。**

形态三：旧的归因是错的。他原以为那道题错在粗心，现在发现错在一个更早的地方。这类回写在错题订正中常见，也最容易被浪费——因为订正通常只要求改答案，不要求改归因。

三种形态给出了备课时的一条实用规则：**先找冲突，再设计活动。**找不到冲突的内容，这节课不必强求回写，把它讲清楚就好。

七、解压不是逆运算（一个必要的重申）

第二章第六节说过一次，这里必须重申，因为它决定了教师在第五编里的具体动作。

回写不还原历史上的那条路径与那片土壤。学习者代入的是**他自己的**路径与土壤——他自己的困惑、他自己已有的经验、他自己走出来的那条弯路。

所以教师不需要、也不可能让学生重走原路。他只需要让学生走出一条**他自己的路**。

这一条免除了教师一个沉重的负担：他不必成为学科史专家，不必知道那个概念当年的全部曲折。他只需要知道**这里曾经有一个困难**，然后把那个困难在今天的孩子身上重新做出来。

八、可观察的读数

回写率 = 迁移题正确率 - 复现题正确率。

复现题是与例题同型的题；迁移题是需要用这次的判断去改写一个原有认识的题。当堂测一次，一周后再测一次。

老路的典型形态是：复现题高、迁移题低，且迁移题随时间快速衰减。发生学教学的预期是：两者差距缩小，且一周后的衰减显著更慢。

这个读数可以与常规作业合并，不增加负担。第三十八章将把它列为四项常规读数之一，并给出使用边界——它一旦被用于排名，教师就会开始训练学生表演回写。

九、与相邻说法的分界

有一个成熟的区分：工具性理解与关系性理解，前者知道怎么做，后者知道为什么。回写与关系性理解相邻，但性质不同：关系性理解是一种**状态**——连接有多少；回写是一个**事件**——有没有发生过一次改写。这个差别有实践后果：状态需要一套量表去评估，事件只需要一个动作去观察。第三十八章的全部读数都建立在事件而非状态上，原因就在这里——**事件便宜，状态昂贵；昂贵的秤不会被使用。**

第十七章 封写：回写杀死了回写

发现学教育里回写照样发生，只是写错了地方：写进底盘的不是“这个概念连着它的来路”，而是“这类东西本来就在那里”——而这一条，正是一条禁止后续回写的规则。

一、一个必须回答的问题

到上一章为止，一个自然的疑问是：既然光滑成品也能被接住，那么发现学教育无非是“回写没发生”，补一补不就行了吗？

而一线教师最熟悉的一件事恰恰是：**补课常常无效**。把当年没讲的路径还给他，把那个困难重新做一遍，他仍然接不住——而且往往越到高年级越接不住。

现行的解释是“基础太差”“习惯已经养成”“兴趣没了”。这些解释都不错，但都不能预测，也不能干预。本章给出一个可以预测、也可以干预的机制。

二、回写照样发生，只是落错了槽

本书的回答是：发现学教育里**回写照样发生**。

一个孩子接住一个光滑成品之后，写进他底盘的不是“这个概念连着它的来路”，而是另一件更抽象的东西：

“**这类东西本来就在那里，我的任务是把它接住。**”

数学对象是现成的，历史结论是现成的，课文的中心思想是现成的，语法规则是现成的。这条信念不是任何人教给他的，是从每一次成功的接受中**归纳**出来的——而且归纳得完全正确，因为他遇到的每一样东西确实都是以现成的形态出现的。

第二章第七节说过，光滑化的第一个副产品是先在感。现在可以给出它的完整机制：**先在感不是被动残留的感受，是被主动写下来的一条信念**。而且它每被验证一次，就加固一次。

三、这条信念的内容，是一条关于回写的规则

关键的一步在这里。

那条元层信念，本质上是一条**关于回写的规则**：

既然对象是现成的、外在的、不因我而变的，那么我的底盘也就不必因它而改。

这就得到本书最重的一个判断：

落在元层的那次回写，写下的是一条禁止后续回写的规则。回写杀死了回写。

一个孩子接得越顺、越成功、越高效，这条规则被加固得越硬；而这条规则一旦足够硬，对象层的回写就不再被触发——他会继续学得很好，继续接住新东西，继续得高分，而没有任何一样旧东西会因此改变。

四、必须收紧的措辞：写保护，不是销毁写头

“杀死”作为口号可以留，作为命题必须写准。

如果字面上不可逆，那么本书自己的治疗方案就自我否证了——既然通道已死，第五编做什么都没用；而且任何一个成年后真正重新想通过某个概念的人，都是现成的反例。

可辩护的形式是：

写下的是写保护，不是销毁写头。锁可以解，但解锁的动作与普通教学不同类，而且解锁所需的力气随学龄升高。

三个性质需要分别说明：

性质一：解锁动作与教学动作不同类。普通教学作用于对象层——把内容讲清楚、把路径还回去。解锁必须作用于元层——它要打在“对象是现成的”这条信念上，而不是打在某个具体概念上。第三十四章第七节的“元层解锁三问”是这个动作的最小形态。

性质二：必须打在正确的位置上。元层信念不止一条。“对象是现成的”是主要的那条，但还有“课堂上有标准答案”“老师知道答案”“不会就该去问”等等。解锁若打在次要的那几条上，效果有限。

性质三：活化能随学龄升高。接受的次数越多，这条规则被验证得越多，解锁需要的力气越大。这给出一个不受欢迎但清楚的政策含义：**换道的最佳时机永远是现在。**

五、三条推论

这个机制的价值在于它能解释现行理论解释不了的三件事，而且每一件都是一线教师天天遇到的。

推论一：为什么补课常常无效。因为补课补的是对象层——把当年没讲的路径还给他。但如果通道是封的，路径还回去也写不进底盘。**必须先做元层解锁，再做还原；次序反了，做多少都没用。**这一条给出了一个可以立刻改变做法的处方。

推论二：为什么最配合、学得最好的孩子往往被封得最彻底。因为封写是通过一次次成功的接受写下的。接受得越顺畅、越高效、越少挫折，那条元规则加固得越硬。

这是一个反直觉但可判决的预测，也是本章与所有相邻理论最重要的分界点：其他理论预测的是失败导致僵化，本章预测的是**成功导致封闭**。第三十九章将设计实验判决这一点。

推论三：为什么去光滑化做得很好的课，回写照样可能不来。课设计得很漂亮，孩子也确实亲自走了路径，回写仍然不发生——因为通道是封的。这是一种**独立的失败模式**，不能归因于“课设计得不够好”。

第三条推论对本书自身有一个重要作用：它堵住了一个逃生解释。如果第五编的方案失败了，不许用“你们做得不够好”来解释。第三十九章将把这一条写进撤稿条件。

六、封写在日常课堂里的三个征候

上面的机制在实践中有三个可以直接观察到的表现，教师大多熟悉，只是过去没有名字。

征候一：等。遇到没见过的题面，不试、不猜、不动笔，直接等。等的时间越来越短。等的对象无所谓是老师、同学还是 AI——重要的是他不认为自己是那个应当产出判断的人。

征候二：不改。你把他的错处指出来，他改了那道题，但不改他的想法。下一道同类的题，他还是那样想。这不是懒惰——那道题的对错被记在了“这道题”名下，没有传到“我的想法”名下。这正是回写通道关闭的直接表现。

征候三：“本来就是”。你问“为什么是这样”，他答“本来就是这样”，而且带着轻微的不耐烦，仿佛这是一个不该问的问题。

第三个征候最可靠，因为那份不耐烦说明这条规则已经硬到具有**规范性**——不只是“我认为对象是现成的”，而是“追问对象的来路是不合规矩的”。

这也解释了一件教师常常困惑的事：为什么真正深刻的问题，往往由成绩中等或不那么听话的学生提出。因为提出这类问题需要一个前提——他还没有完全接受“对象是现成的”这条元规则。这与推论二完全一致，而且是任何一位有经验的教师都可以立即回想起若干实例的。

七、一道必须设的挡板

本章的措辞与哲学中“先验幻相”一类概念相近，必须当场划清，否则会滑向一个本书不主张的结论。

哲学传统中的先验幻相源于理性自身的结构，因而**不可避免、只能揭示而不能消除**。本章所说的先在感是**习得的**，源于一部具体的传递史，因而**可以被去光滑化消解**。

同时必须设第二道挡板：本书只主张“**学习者身上的先验感是被制造出来的**”，绝不主张“数学或科学的必然性是幻觉”。数学的必然性不因为一个孩子怎么学它而改变。

这两句话之间的距离，是本书全部立场的安全边界。越过它，本书就变成了一种廉价的相对主义，而那与本书的主张恰好相反——第五编要求学生走自己的路，同时要求他把这条路交出去接受检查。

八、一个疏漏：教师同样会被封写

本章从头到尾把封写写成学生的事。这是一个疏漏，而且是一个后果很重的疏漏。

同一条机制作用在教师身上的形态是：长期使用某一种格式（任务单、四栏表、标准流程、评课量表）之后，教师会逐渐**看不见**那些不落在格子里的东西。不是他不重视，是他的知觉结构被回写了——那些东西在他眼里不再显现。

这一条的后果是：被封写的教师无法守岔口，因为他看不见岔口在哪里。

第五编全部的设计都假定教师能够识别出“这里可以撬开”。如果这个假定不成立，那么第二十九章的五个矿脉、第三十四章的四段结构，对他而言都是无处落地的。

可观察的读数（成本极低）：格式提前量 $\Delta = t_c - t_f$ ——格子出现的时刻，减去那个东西本来会形成的时刻。只需要一块秒表。配一个**撤格式测试**：把任务单撤掉一节课，看还剩下什么。

这一条对教师培训的含义是次序性的：在要求教师“守住岔口”之前，先要问他还看不看得见岔口。先做撤格式测试，再讲第五编。

九、可观察的读数

先在感指数：一份五题的小量表，测量学生在多大程度上认为某个对象“本来就是这样、不可能是别的样子”。

题目形如：“如果当年的人选了另一种做法，今天这条规则会不会不一样？”

发现学教育预期该指数随学龄单调上升；本书的核心预测是，经过元层解锁的班级，该指数下降。这是一个反直觉、有方向、可测的预测，第三十九章将它列为主实验的关键指标。

十、与相邻说法的分界

有五家理论与本章形状相近，必须逐条切开，否则本章会被当作换名重述。

与“框架学习”一类的分界。该传统主张学习分层，高一层的框架几乎不可修订。本章与之最近，但差别在于：那一传统说的是框架**难以修订**，本章说的是**锁是被回写这个动作本身写下的**——不是因为高层难改，是因为那次成功的接受所写下的内容，恰好就是一条禁止修改的规则。

与“双环学习”的分界。该传统主张单环的成功会系统性地阻断对框架的检视。方向一致，但本章给出一个与之**方向相反**的可判决预测：最配合、学得最好的学生被封写得最彻底。该传统预测的是防卫性行为，本章预测的是顺从性行为。

与“本体误归类”的分界。该理论主张概念一旦被放入错误的本体范畴，同范畴内的再教学无法纠正，须换范畴。这是教育文献中最强的竞争者。分界在于：误归类说的是**放错了柜子**，本章说的是**柜门上了锁**；前者的补救是换柜子，后者的补救是先开锁再谈放哪。第三十九章的主实验被设计为可在两者之间做出判决。

与“定势效应”的分界。定势说的是过去成功的解法阻断新解法，是关于**解题策略**的现象；封写发生在元层，写下的不是一条策略，而是一条关于“底盘是否需要改动”的规则。

与“同化与顺应”的分界。回写与顺应相近，纯同化而不顺应会导致图式僵化，这在发生认识论里已有雏形。差别在于：那一传统把僵化归于**缺少顺应**，本章主张有一次顺应发生了，只是它的内容是“以后不必顺应”。**不是少了一次，是多了一次错的。**

第十八章 四种死亡与诊断次序

一堂课失败有四种形态，其中三种在对象层，一种在元层；诊断必须从元层那一种查起，否则前三种的所有补救都是无效功。

一、为什么需要一张诊断表

前两章给出了判据与机制。但教师在现场面对的不是机制，是一个具体的孩子没有学会。他需要知道的是：**这一次是哪一种没学会。**

不同的失败形态需要完全不同的处理。用错处理方式不只是无效，而且会加深问题——第六节将给出一个具体例子。

本章给出四种失败形态，本书称之为四种死亡；并给出一个诊断次序，这个次序比四种形态本身更重要。

二、第一种：土壤之死

形态。这个概念对他没有连接点。它落下来的地方是空的——他没有任何已有的经验、已有的对象、已有的困惑可以接住它。

表现。他能复述，但用不上；过一段时间就干净地忘掉，且忘掉时没有任何不适感——因为它从来没有连上任何东西。

典型场景。过早引入一个抽象概念；引入一个在他生活中完全没有对应物的对象；跳过前置内容。

补救。给土壤：先建立可以接住它的那些经验与对象。这是常规教学最熟悉的一种补救，通常叫“补前置知识”。

三、第二种：路径之死

形态。这个概念他接住了，接得很牢，牢到路径被完全省略——他知道结论，也知道怎么用，但完全不知道为什么是这样，而且**不觉得需要知道。**

表现。在标准题面上表现优异；一旦题面变化超出他见过的范围，束手无策，而且不会尝试从原理往回推——因为对他而言那个原理不存在，只有一条被记住的规则。

典型场景。大量重复训练之后的熟练；口诀化教学的常见结果。

补救。去光滑化：把路径还回去，让他走一遍。

一个必须点明的地方。这种死亡是**死于成功**——正是因为接得太牢，路径才被省略掉了。它是这四种里最容易被误判为“学得好”的一种，因为它的全部读数都是优秀的。

四、第三种：显露之死

形态。他自己建立了一套解释，这套解释在他的经验范围内完全自治，因而不需要修改。而这套解释是错的，或者只在有限范围内成立。

表现。他能解释、能推理、能自圆其说，而且非常自信。指出他错了，他会为自己辩护，而且辩护得有道理——因为在他的证据范围内他确实没错。

典型场景。朴素理论；前概念；“乘法使数变大”“重的东西落得快”这一类。

补救。给一个他自己的体系解释不了的事实，而且必须是他亲自验证的事实——不能是别人告诉他的。

五、第四种：元层之死（封写）

形态。前三种都在对象层——死的是某一个概念。第四种在元层：**沉进底盘的不是一条路径，而是“路径不必走”这条规则。**

表现。上面三种补救全部无效。你补了土壤，他仍然只是多接住一样东西；你还了路径，他把路径也当成一条要记住的内容；你给了反例，他把反例记下来当作一个特例。

根本特征：他不改自己的想法，他只是把新东西也接住。接住得非常好，而且态度端正。

补救。必须先做元层解锁，再做上面任何一种补救。

六、诊断次序：为什么必须从第四种查起

这是本章最重要的一节，也是全书最具操作性的一条纪律。

四种死亡的诊断次序不是随意的：

 | **先查第四种，再查前三种。**

理由是：如果第四种成立，前三种的**全部补救都是无效功**，而且做得越认真，无效功越大。

更糟的是，无效功会产生一个错误的结论。教师补了土壤、还了路径、给了反例，孩子仍然没有变化，于是得出：“这个孩子不行”“他就是不开窍”“他没有那个悟性”。

而这个结论是错的，且伤害极大。第七章第九节说过，制度把结构性的缺失转译成了个人的天赋差异；这里给出了那个转译发生的具体现场。

所以诊断次序的实际含义是：**在给一个孩子贴上“没有悟性”的标签之前，必须先排除第四种死亡。**

七、四种死亡诊断表

	死在哪里	典型表现	补救	误判为
一	土壤	能背，用不上，忘得干净	建立可接住它的经验	不认真
二	路径	标准题优异，变式束手	去光滑化，把路走一遍	学得好
三	显露	自成体系，自信，会辩护	给他亲自验证的反例	想法多、爱抬杠
四	元层	前三种补救全部无效	先元层解锁	没有悟性

最后一列值得单独看。四种死亡对应四种最常见的误判，其中第二种与第四种的误判方向相反却同样有害：**一个被表扬，一个被放弃。**

八、一个便捷的现场诊断动作

完整诊断需要时间，现场有一个便宜的替代动作，可以在一分钟内区分第四种与前三种：

给他一个他自己以前的说法，问：“这句话现在还对吗？”

- 第一种死亡的孩子会说不记得了。
- 第二种会说对，因为规则没变。
- 第三种会为它辩护。
- 第四种会说“对啊”，然后等着你告诉他哪里不对——他不认为判断这件事归他管。

第四种的特征不是答错，是把判断权交出去。这个动作与第三十八章的第四个读数（他对AI说的第一句话）测的是同一件事，只是场景不同。

九、四种死亡与两条赛道的关系

需要指出一件容易被忽略的事：这四种死亡并不是老路特有的。

第一、二、三种在任何教学形态下都会发生，包括第五编那套做法。教学不可能没有失败，本书从不承诺一种不会失败的方法。

第四种是老路特有的——它是老路正常运转的产物，而不是老路出故障的产物。这一点值得再说一遍：

前三种是这条路走坏了的结果；第四种是这条路走好了的结果。

这就是为什么第四种最危险，也最难被发现：所有的质量监控机制都在检测“走坏了”，没有任何机制在检测“走好了”。

十、可观察的动作

挑一个你认为“怎么教都教不会”的孩子，用第八节那个动作试一次：给他一个他自己以前的说法，问这句话现在还对吗。

如果他的反应是第四种，那么你此前所有的努力都打在了错的位置上。这不是你的失误——没有人给过我们这张表。

十一、与相邻说法的分界

医学与工程中的故障诊断有一条通用原则：先排除会使其他诊断失效的那一类故障。本章的次序纪律与之同构。而教育诊断领域现有的框架（学习困难分类、错误类型分析、认知诊断评估）几乎全部工作在对象层——它们细致地区分了本章的第一至三种，且方法远比本章精细；但它们没有第四种，因为它们的整个测量框架建立在“通过题目表现推断知识状态”之上，而第四种死亡的孩子在题目表现上是正常甚至优秀的。这不是那些框架不够好，是它们测不到那一层。第三十八章的读数因此全部落在行为事件上，而不落在题目表现上。

第十九章 毒性与摩擦：一个可拆解的比较

同一条路，不同的剂量。课堂形态自带五道摩擦，它们历来被视为需要克服的缺陷，而恰恰是它们把这条路的毒性维持在有限水平。

一、把第十四章的式子接过来

第十四章给出了式子：

$$\text{毒性} = \text{老路} \times \text{速度} \div \text{摩擦}$$

那一章从“AI 高效自学”这个具体形态进入，本章把它一般化，因为这个式子适用范围比自学模式宽得多——它可以用来评估任何一种教学安排。

三个变量重述一遍：

老路是前提项。若这一项为零——教学不以抵达一个已定结论为终点——则后面乘什么都是零。所以本书的处方是换道，不是降速。这一条必须放在最前面，否则本章会被读成“教学应该低效”，那是一个愚蠢的主张。

速度是未闭合时长的倒数。

摩擦是本章的主题。

二、摩擦不是好东西

先把话说清楚，否则本章会被误用。

第十一章已经说明：那些摩擦本身没有价值。等待是浪费，社交压力是伤害，匹配不良是低效，不敢反复问是损失。**没有任何理由为它们辩护，也没有任何理由主张把它们保留下来。**

有价值的是它们的**副产品**：在供给到达之前的那一段时间里，孩子只能自己待着。

这个区分决定了处方的形态。本书不主张恢复摩擦——不主张让孩子重新排队、重新怕丢脸、重新问不到人。本书主张的是：**摩擦消失之后，必须另外造一个入口。**因为那个入口原本是摩擦的副产品，而副产品随着母体一起消失了。

第三十四章的四段结构，就是那个另外造出来的入口。它与摩擦的差别在于：**它是被设计的，因而是可控的、有起止的、对所有人平等的**；而摩擦是偶然的、不平等的（社交摩擦对内向的孩子远大于外向的孩子）、且无法调节。

造出来的入口比捡来的入口好。这是本章对整个局面唯一乐观的一句话。

三、五道减毒装置的一般形态

第十四章列过五道，那里是与自学模式对照。这里把它们提炼成一般形态，因为它们可以用于评估任何一种教学安排，包括第五编那一套。

装置一：延迟（等待）。从疑问产生到供给到达之间有一段时间。

装置二：异议（别人）。存在一个未经安排的、不同的说法。

装置三：克制（可能忍住的人）。在场至少有一方有能力且有动机不给答案。

装置四：公开（说出口）。想法必须被表达，因而可被反驳。

装置五：边界（起止）。这段过程有外部规定的开始与结束。

这五项是一份**检查表**。给定任何一种教学安排，逐项问它有没有，就能估计它落在哪个位置。

四、用这份检查表评估四种常见安排

传统讲授课。延迟：有（要等老师讲到）。异议：少（同学的不同说法通常被压制）。克制：可能有。公开：有（要举手回答）。边界：有。——五项中三到四项在位。这解释了一件反直觉的事：**最传统的讲授课，未必是毒性最高的那一种。**

探究式课堂。延迟：有。异议：多（这是它的强项）。克制：设计上有，实施中常常没有。公开：多。边界：有。——五项基本齐备。这也印证了第六章的判断：探究式教学是老路上能达到的最高水平。

AI 辅助的个性化课堂。延迟：接近零。异议：少（每个人的路径被单独定制，无从产生异议）。克制：无。公开：少（各自与屏幕互动）。边界：有（课时还在）。——五项中只剩一项。

AI 自学模式。五项全无。

四行排下来，第十四章“课堂是减了毒的老路，自学模式是纯的老路”这句话得到了具体的刻度。而更值得注意的是第三行：**AI 辅助的个性化课堂，在这份检查表上比传统讲授课差得多。**它保住了课时这一项，其余四项全部失守。

这是一个不受欢迎的结论，也是本章存在的理由——如果不把摩擦单列为一个变量，这个结论无法被表达出来。

五、一个必须承认的两难

上面第三行的结论会立刻引出一个反驳：难道要退回传统讲授课？

不。因为传统讲授课在**另外两个变量**上很糟：它的“老路”项满格，它对弱势学生的实际效果远差于个性化教学。第十三章已经承认过，个性化把因材施教从理想变成了现实，这是教育史上的大事。

所以这里有一个真实的两难：

■ **个性化提高了接受的质量，同时拆掉了发生的入口。**

本书不主张在两者之间取舍，主张的是把它们分开做：**接受这一段，尽量个性化、尽量高效，而且尽量放到课外去——因为它现在免费了；课堂时间用于那件个性化做不了的事。**

这就是第八章第六节那条时间论的完整形态：课时没有增加，用途变了。而这个方案的可行性完全依赖于一件事——**必须真的把腾出来的时间用于另一件事，而不是自动填入更多内容。**第十三章第六节的进度错觉，讲的正是这个自动填充。

六、摩擦的三个替代品

如果不能恢复摩擦，又必须保住那个入口，可用的替代品有三个。它们全部在第五编展开，这里先列出，因为它们与本章的检查表一一对应。

替代品一：规矩（替代装置一与装置三）。“从问题张开到它被合上，中间谁都不许给答案，包括我。”这是一条外加的、明说的、有起止的规则。它人为地造出延迟，同时把克制变成一项公开承诺而不是个人修养。

替代品二：对手（替代装置二）。在自学场景里，把提问形态从索取型改为对抗型；在课堂上，让 AI 或同学承担异议。这一条把“异议”从偶然的副产品变成一个被指派的角色。

替代品三：结算口（替代装置四与装置五）。每一段结束时必须产出一句可被检查的话——“我原来以为什么，现在这句要改成什么”。它同时提供了公开性与边界。

三个替代品的共同特点是：**它们都是外加的、可执行的、不依赖个人修养的。**这一点很重要。第五章已经说明，凡依赖自觉的东西在这套系统里活不下来；只有被写成规矩、排进流程、给了读数的东西才能活。

七、可观察的读数

| **未闭合时长：从教师抛出问题，到任何一方给出答案，中间经过的秒数。**

这是全部读数中最便宜、最难作假的一个：一支笔或一个录音笔即可采集，不需要量表，不需要额外时间，不依赖任何人的配合。

经验判断是，绝大多数常规课堂的这个数字在十秒以内；能稳定超过三分钟的课，一学期也没有几节。

一条使用纪律必须同时给出：这个读数只设下限，不设上限。把它变成“越长越好”的指标，会立刻产出为拖时间而拖时间的课。第三十八章将展开这条纪律。

八、三条独立的实证线索

第二节说摩擦的副产品是那段“只能自己待着”的时间。这一判断并非只有思辨支持，学习科学里有三条互相独立的线索指向同一个方向，而且它们的共同特征值得注意：**三条都预测当期表现更差。**

线索一：有效失败。让学生先在没有指导的情况下去啃一个他们注定解不出的问题，再给正规讲授，其迁移表现优于一上来就正规讲授的组——**尽管前者在过程中的即时表现明显更差。**

线索二：期望难度。一批会让当期表现变差的操作（间隔、交错、提取而非重读），恰恰会让长期保持与迁移变好。

线索三：为未来学习作准备。先让学生发明一个必定失败的方案，再听讲，其后续学习效果优于直接听讲。

三条合起来给出一条对本书极其重要的判断：

| **当期效率与长期判断力之间，存在系统性的负相关区间。**

这条判断有两个直接后果。

其一，它使“AI 赋能”这个承诺的方向问题变得可判定。赋能三个动作，字面意思就是消灭等待、消灭歧义、消灭卡顿、消灭摩擦。而上述三条线索说的是：**在某个区间上，消灭这些东西正好消灭了那个我们要的东西。**

其二，它给第三十九章的实验设计定死了一条：**主要结果指标必须是延迟迁移测验，不能是即时后测。**因为三条线索都预测即时表现在干预组更差。用即时后测检验本书的主张，等于用体温计量体重——零结果不构成反证。

一处必须与它们划清的地方。这三条线索处理的都是**被设计进去的困难**（如何安排练习顺序、如何制造一次注定的失败），其核心问题是“该加多少难度”；本章说的是一种**本来就存在、正在被无意取消的困难**，其核心问题是“取消之后拿什么补”。前者是调节参数，后者是造一个新部件——这个差别决定了处方形态完全不同。

九、与相邻说法的分界

学习科学中有“合意困难”一族研究：某些增加当下难度的做法反而提高长期保持与迁移。本章与之方向一致，但对象不同：那一族研究的是**被设计进去的困难**（间隔、交错、提取练习），其核心问题是“该加多少难度”；本章说的是一种**正在被无意取消的困难**，其核心问题是“取消之后拿什么补”。这个差别决定了处方：前者是调节参数，后者是造一个新部件。此外，那一族研究的困难全部作用于记忆与提取，本章的入口作用于**判断的产出**——一个孩子在等待中做的事不是回忆，是尝试。

第二十章 谁在吃：损害分布的翻转

课堂形态的损害是薄而广的，自学模式的损害是深而窄的，而且恰好落在最自觉、最上进、因而最可能开辟新路的那批孩子身上。

一、剂量之外的另一个变量

第十九章处理的是每一剂有多少。本章处理的是另一件事，而且它比剂量更重要：**谁在服用**。

在流行病学里，一种暴露的危害不只取决于剂量，还取决于人群分布——同样的剂量，均匀落在全体身上，与集中落在某个特定子群身上，后果完全不同。

教育上的这次暴露，恰好属于后者。

二、两种分布形态

课堂形态：薄而广。全班一起上课，一起接收成品，这条路对所有人施加大致相同的作用。它的损害是均摊的——每个人都被写进一点点那条元规则，谁也不特别多。

自学模式：深而窄。主动去用的是一个高度选择性的子群：最自觉、最上进、家庭最重视教育、自我要求最高的那批孩子。他们不是被安排去用的，是自己找上去的，而且用得最久、最投入、最认真。

这不是猜测，而是任何一项自愿使用的学习工具的必然分布——**自愿使用的东西，总是被最愿意学习的人用得最多**。

三、把它接到封写机制上

第十七章的机制在这里给出一个不舒服的结论。

那把锁不是被失败写下的，是被一次次成功的接受写下的。接得越顺、越高效、越少挫折，那条“东西是现成的、我不必因它而改”的元规则加固得越硬。

于是：

这剂药，优先被喂给了那批最有可能开辟新路的孩子。

从个体角度看，这是不幸。从文明角度看，这是更严重的事：新路的开辟本来就依赖极少数人，而这一小群人正是暴露剂量最高的人。

四、一个已经可以观察到的现象

上述推论并非纯思辨，它对应一个已经可以在学校层面观察到的现象：

引入 AI 学习工具之后，**班级内部的成绩差距往往缩小**——弱势学生被迅速拉平。这是真实的、值得肯定的改善，本书不打算用任何方式贬低它。对一个原本被落下的孩子来说，一个不厌其烦、不带评判、随时可问的辅导者，其价值难以估量。

而与此同时，**顶端那种“会问怪问题的孩子”变少了**。

前者上得了报表：均分提高、及格率提高、后进生面缩小。后者上不了报表：没有任何一项指标记录“本校今年出现了几个提出了不该被提出的问题的学生”。

于是一所学校在数据上会看到一次明确的成功，而它损失的那一部分从头到尾不出现在任何数据里。

这不是数据造假，是数据覆盖面的结构缺口。第八章已经说明了这个缺口的来历，第三十八章将给出补它的方案。

五、必须堵住的一个误读

这条推论极易被推成一个本书坚决反对的结论：既然优秀学生受害最深，是不是应该禁止他们使用 AI？

不是，理由有三条。

第一，禁不住。第二十四章将处理这一点：禁令只在课内生效，而损害写在哪里都算数。

第二，禁错了对象。提早供给不是 AI 发明的。禁掉它，最快的答案源回到教师和教辅书上，损害照样发生，只是换了供给商。

第三，这条推论指向的处方与禁令方向相反。既然毒不在 AI 而在提问的方向（第十四章第九节的关键判别项），那么正确的处方是给自学装上对手与回写口，**而且优先给这批孩子装**——因为他们暴露得最多，也因为他们最有能力执行这套配置。

一个愿意每天自学三小时的孩子，完全有能力做到“先说自己的想法再让它挑错”。**这一批孩子不是需要被限制的人，是最容易被改配置的人。**

六、一个更远的推论：分布形态与创新供给

把时间尺度拉长，本章的分布问题有一个超出教育的含义，值得在这里点出，尽管本书不打算展开。

任何一个社会的新知识供给，都依赖一个很小的子群——那些愿意在没有路的地方待着的人。这个子群的规模从来不大，而且它的形成机制一直不清楚，通常被归因于天赋、机遇或性格。

如果本章的推论成立，那么这个子群正在被一种看上去完全正当、而且确实有益的实践系统性地削薄。没有任何人有恶意，没有任何环节可以被指责，每一步都是理性的、有据的、被数据支持的。

这也是本书之所以采用“包办”而非“威胁”作为核心词的最终理由：**威胁来自外部，可以被抵抗；包办来自内部的成功，只能被识别。**

七、可观察的读数

在你的班上找出使用 AI 学习最多的那三五个孩子，看他们最近三个月里，有没有提出过一个让你答不上来的问题。

这个读数很粗糙，样本极小，不构成任何证据。但它测的方向是对的，而且任何一位班主任都可以在五分钟内完成。

如果答案是没有，而三个月前有过，那么值得留意。

八、与相邻说法的分界

教育公平研究中有一族关于“技术使用的马太效应”的讨论：优势家庭的孩子更善于利用技术，从而拉大差距。本章的结论在方向上与之**相反**：本章预测的是技术缩小低端差距、同时削薄高端，即分布的**两端同时向中间收拢**。这个差别是可检验的，而且检验方法便宜——只需同时看均值、方差与顶端表现，而不是只看均值。大量现有评估只报均值与及格率，因此在结构上无法发现本章所说的现象。

第二十一章 潜伏期：先涨、再平、再不可逆

这剂药的显效是延迟的，而它的第一阶段恰好表现为显著变好——于是最容易被当作成功的证据，并据此加大投入，而加大投入正好加大剂量。

一、为什么潜伏期是关键

一件有害的事情如果立刻显效，它会被发现并停止。加固悖论之所以危险，不只因为它在读数上表现为变好（第十二章），还因为它在时间上有一段潜伏期。

潜伏期的作用是：**在损害显形之前，系统已经完成了扩大投入的决策。**而扩大投入等于加大剂量。

本章给出这条曲线的三段形态，并给出每一段的可观察特征。这是本编中最接近可预测的一章，因此也是最容易被证伪的一章。

二、第一阶段：先涨（约一学期）

表现。所有读数上升，且上升幅度往往超过预期。正确率、进度、完成度、满意度、参与度，全线变好。弱势学生的改善尤其明显。

为什么会这样。第十二章已经说明：接管的全部方式都是提高成品的投递效率，而读数量的正是投递效率。这不是短期效应，也不是新奇效应，它是真实的、结构性的改善。

这一阶段的危险。它提供了扩大投入的全部理由，而且这些理由在各自的框架内完全正确。任何一位负责任的管理者，看到这组数据，都应当扩大投入——**不扩大反而是失职。**

这就是本章要指出的第一个不对称：**在第一阶段，正确的判断与有害的行动是同一个。**

三、第二阶段：再平（约二至三学期）

表现。读数停在一个较高的位置，不再上升。有时会有小幅回落，但基本维持。

通常的解释。这一阶段的解释几乎是固定的：红利期结束了，需要更深度的融合，需要更好的工具，需要更系统的培训，需要提高教师的使用水平。

这些解释全部指向同一个方向——**做得更多、更深、更好。**于是投入进一步加大。

为什么会平。因为接受这一环有一个上限：成品被接住的比例最高就是百分之百。当投递效率提高到接近这个上限时，继续提高投递效率不再产生任何读数增益。

这一阶段的危险在于它被误诊。平台期被读作“技术潜力尚未充分释放”，而实际上是“这一环已经做到头了”。真正的诊断应当是：**该换一个维度了，而不是在这个维度上加码。**

四、第三阶段：不可逆

表现。这一阶段的特征不在总分上，在结构上。

复现类题目继续保持高位；需要改变原有认识的题目不再随之上升，两者的差距扩大。第十六章的回写率读数在这里会显示为持续下降。

顶端出现变化：会问怪问题的孩子变少；面对没有现成答案的题时，尝试的时长缩短。

为什么不可逆。这是本章与前两阶段最重要的区别。前两阶段是量的变化，第三阶段是**状态的变化**——按第十七章的机制，那条元规则已经硬到足以关闭通道。

不可逆的具体含义是：

减少投入不会自动恢复。把 AI 撤掉，读数会下降（因为投递效率下降），但那件真正丢掉的东西不会回来。这是最容易被误判的一点——撤掉之后数据变差，会被读作“看，还是需要它”。

加大投入不再改善它。因为投入的全部方向仍然是提高投递效率。

常规补救无效。第十八章的前三种补救全部打在对象层，而问题在元层。

五、这条曲线为什么难以被发现

三段合起来，构成了一个几乎完美的隐蔽结构：

第一阶段：变好 → 加大投入。

第二阶段：平台－解释为潜力未释放－进一步加大投入。

第三阶段：结构变化－不出现在任何常规读数上。

在这三个阶段的任何一点上，都没有一个自然出现的信号会促使系统停下来检查。

而且第三阶段的显形时间足够长——按学段计，一个孩子从三年级开始高强度使用，到显形可能已经是初中。那时候的解释资源非常丰富：青春期、学业难度上升、兴趣转移、手机、家庭。没有任何人会回溯到三年级。

六、一个诚实的说明

本章是全书**证据最薄**的一章，必须说清楚。

三段曲线目前是一个推论，不是一个观测结果。它从第十二、十七章的机制推出，具有内部一致性，但缺少纵向数据的支持——而这类数据需要三到五年的追踪才可能取得。

因此本章的地位是：**一个可以被证伪的预测，而不是一个已经成立的描述。**

它的证伪条件很清楚：如果在一所高强度使用 AI 学习工具的学校里，回写率与复现题正确率**同向上升并持续三年以上**，本章即被推翻。

本书希望这条预测被推翻。**如果它被推翻，说明第十七章的机制不成立，那么第三编的大部分内容都可以删掉，而这将是一个好消息。**

七、最低可行的观测方案

三到五年的追踪对一所学校太重。这里给出一个可以立刻开始、几乎零成本的替代方案。

做法：每学期一次，同一批学生，做两件事——

一、一份十道题的小卷，五道复现题、五道迁移题，记录两者的差值（第十六章的回写率）。

二、一道空位题（第二十二章将说明它的形态），记录他停止尝试之前经过的时间。

两项合计占用一节课，一学期一次。

观察什么：不看绝对值，看**两条曲线的相对走向**。如果复现题上升而回写率下降，第三阶段正在发生；如果两者同向，本章的预测在这所学校不成立。

这个方案的价值在于它便宜到没有理由不做。而第八章已经说明：**昂贵的秤不会被使用；不被使用的秤等于不存在。**

八、可观察的读数

把你们学校引入 AI 工具那一年到现在的各项数据画成折线。然后在旁边空出一栏，标题写“回写率”，里面什么都没有。

那一栏的空白，就是本章要说的全部。

九、与相邻说法的分界

医学中的潜伏期概念与本章同构，但有一处关键差别必须点明：疾病的潜伏期内，暴露者通常没有任何收益；而本章的第一阶段有**真实且显著的收益**。这个差别使得预防的难度高出一个量级——不是要说服人们停止一件无益的事，而是要说服他们在**一件确实有益的事上保留一个反向部件**。这也是本书全部处方都采取“加装”而非“禁止”形态的根本原因。

第二十二章 可达域信念：“没有路通到”被读成“不存在”

一个长期只坐车的人，最终形成的不是“我不会修路”，而是“**凡是**没有路通到的地方，就是不存在的地方”——这是封写的终局形态，也是本编与全书的接榫。

一、损害的最后一层

第十六至二十一章处理的是回写通道的关闭。本章处理它的一个后果，这个后果比通道关闭本身更远，也更难恢复。

它不表现为“我不会”，而表现为“**那儿没东西**”。

第十五章的高速公路比喻在这里给出了它最重的一句：一个从小只坐过车的人，不只是不会修路。他会慢慢形成一条更根本的信念——凡是**没有路通到的地方，就是不存在的地方**。

二、这条信念是怎么被写下来的

它的写入机制与第十七章完全相同，只是内容更抽象了一层。

一个孩子经历的每一次学习，都是这样的形态：有一个问题，这个问题有答案，答案可以被取得。**这个形态被验证了成千上万次，而且从来没有例外。**

从这些经验中归纳出来的信念是完全合理的：问题都有答案；答案都在某处；我的任务是找到它。

而这条信念有一个未被察觉的逆否形式：

如果找不到答案，那么这不是一个真的问题。

这就是可达域信念。它把“我没找到”与“它不存在”等同了起来，而这个等同在他全部的经验范围内都是成立的——因为学校里出现的每一个问题，确实都有答案。

三、为什么这是最严重的一种损害

前面几章讲的损害，损失的是**能力**：不会自己走一段路，不会改自己的想法。这些能力在原则上可以被重新训练。

本章讲的损害，损失的是**方向**：他不会朝那个方向看。

一个不会修路的人，如果被带到一处没有路的地方，他至少知道自己站在一个可以修路的地方。而一个持有可达域信念的人，被带到同一处，他看到的是一片虚空——他不认为那里有任何东西可做。

能力的缺失可以被补，方向的缺失连“要补什么”都不会被提出。

这也是本书把第五编的第一个动作定为“带他下一次高速，在没有路的地方走一段”（第十五章第五节）的原因：那一段路本身不重要，重要的是他知道了那个方向上有东西。只需要一次。

四、空位题：一个便宜的检验

这条信念可以被检验，方法很便宜。

空位题：一道确实没有现成答案的题。注意它不是难题——难题是有答案而难以求得的题，空位题是答案尚不存在的题。

两者的区别在学生的反应上极其清楚：

一 面对难题，他会说“我不会”，然后要么继续想，要么放弃，要么求助。

一 面对空位题，他会先去找答案；找不到，会说“我搜不到”“它也不知道”；然后**停下来**。

停下来那一刻，不是他不够聪明。是他把“没有路通到”读成了“不存在”。

而一个走过野路的孩子在同一处的反应完全不同：**他会朝那个方向走两步看看**。提出一个猜想，划定一个范围，给出一个自己也不确定的说法，或者说“我觉得可能是这样，但我不准”。

只有两步，但性质完全不同。

五、怎么出一道空位题

这是教师最需要的一节，因为空位题的出法与常规命题的训练完全相反。

常规命题的第一原则是“答案唯一且确定”。空位题的原则是三条：

第一，它必须是真的没有答案，不是老师装作不知道。孩子对“老师其实知道”这件事极其敏感，装出来的开放性会被立刻识破，而且比不开放更伤害信任（第六章第七节的“伪未定终点之一”）。

第二，它必须是可以被走两步的。完全无从下手的问题不是空位题，是空话。判别方法：这个问题能不能被提出一个哪怕很粗糙的猜想？如果不能，它太大了。

第三，它最好长在已学内容的边上。不是另起炉灶，而是把一个已经学过的东西往外推一格。

各科的例子：

数学。“我们学过的所有运算里，为什么加法和乘法可以交换顺序，减法和除法不行？还有没有别的运算也可以交换？如果我们规定一种新运算，怎么规定才能让它可交换？”——最后一问是空位。

语文。“这篇课文如果从另一个人的角度写，中心思想会不会变？如果会变，那原来那个中心思想到底是谁的？”

科学。“这个实验如果当年做出来是相反的结果，我们今天的课本会怎么写？”

历史。“当时的人知道结局吗？如果不知道，那我们说他们‘选择了’某条路，这个说法准不准确？”

四道题的共同点：**它们都不问“是什么”，都问“如果不是这样会怎样”**。这也正是第三十四章元层解锁三问的形式。

六、一个必要的限度

本章不主张把课堂变成空位题的堆积。理由有三。

第一，大部分内容不需要。绝大多数教学内容有确定的答案，而且必须被确定地掌握。第七章的三大成就都建立在这个基础上。

第二，空位题很贵。它不能被批量生产，不能被标准化评价，不能被布置成作业。它的产出是不确定的。

第三，也是最重要的：一次就够。本章第三节已经说过，方向的建立不需要频次。一个孩子只要在没有路的地方真的走过一次，他此后就知道那个方向上有东西。

所以本书给出的量是保守的：**一学期一道，认真做完**。这个量任何一所学校都负担得起，而且不与任何进度冲突。

七、可观察的读数

第五个读数：面对空位题的反应时长与方向。

记录两件事：他在多少秒后停止尝试；以及他停止之前有没有朝那个方向走出过任何一步（提出一个猜想、划定一个范围、给出一个自己也不确定的说法）。

这个读数的两个优点：一是采集成本极低，一学期一题即可；二是**几乎不可能被应试训练**——因为没有可训练的答案，只能训练“敢不敢走两步”，而那正是它要测的东西。

八、第三编的结论

第三编到此结束，六句：

第一，第二个冲程的判据是回写：不看 he 多了什么，看他原来的东西变了没有；外部表现是回头改一件旧的。

第二，发现学教育里回写照样发生，只是落在元层，写下的是一条禁止后续回写的规则——回写杀死了回写。这条锁是写保护而非销毁写头，可解，但解锁动作与教学动作不同类，且活化能随学龄升高。

第三，失败有四种形态，三种在对象层，一种在元层；诊断必须从元层查起，否则前三种的补救全是无效功，并会导致“这孩子没有悟性”这个错误且伤害极大的结论。

第四，损害的量可拆解为老路乘速度除以摩擦；课堂自带五道摩擦，AI 辅助的个性化课堂只剩一道，自学模式一道不剩。摩擦本身没有价值，但它的副产品必须被另外造出来。

第五，损害的分布是翻转的：课堂薄而广，自学深而窄，而且落在最有可能开辟新路的那批孩子身上。

第六，终局形态是可达域信念——“没有路通到”被读成“不存在”。它损失的不是能力而是方向，而方向的重建只需要一次真实的经历。

第四编要问的是：既然如此，这两种教育到底是同一件事的两种做法，还是两件事。

九、与相邻说法的分界

心理学中有“习得性无助”一族研究：反复的不可控失败会导致个体停止尝试。本章的停止尝试与之表面相同，机制相反：习得性无助源于**反复失败**，本章的停止源于**反复成功**——每次都顺利找到了答案，于是“找不到答案”这件事从未被经历过，因而“答案可能不存在”这个可能性从未进入他的选项集。这个差别有直接的实践后果：习得性无助的干预是提供可控的成功体验，而本章的干预方向恰好相反——**必须提供一次真实的、无法被搜索解决的困难**。用前者的处方处理后者，会加重问题。

第三编材料：四种死亡的四份诊断记录与一份封写征候清单

本编材料的性质：以下诊断记录为按第十八章诊断表构造的教学示例，用于说明四种死亡在现场如何被区分；非个案实录，不涉及任何真实学生。

材料八 四份诊断记录

统一格式：现象、常规判断、按第十八章的诊断、处理、以及处理之后的分辨点。

记录一：土壤之死

现象。六年级学生，学完百分数应用题，能复述“求一个数的百分之几用乘法”，做课本题正确率高。一周后遇到“打折”情境的题，完全无从下手，且没有任何尝试痕迹。

常规判断。没认真听；题型没见过；需要多练。

诊断动作（第十八章第八节）：给他一句他自己以前写的话——“百分数就是分母是一百的分数”——问：“这句话现在还对吗？”

他的反应。“不记得写过。”

判为第一种：土壤之死。这个概念落下来的地方是空的，它从未与任何已有的东西连上，因此忘掉时没有任何不适感。

处理。不是多练题，是先建立可以接住它的经验：让他用一根皮筋或一条纸带做出“整体”与“其中一部分”的对应，先把“以什么为整体”这件事变成一个可以指着说的东西。

分辨点。若处理之后他能自己说出“打折的整体是原价不是现价”，说明确为土壤之死且已补上；若他仍然能背而不会用，转查第二种。

记录二：路径之死

现象。初中学生，因式分解速度极快，各种题型都熟。遇到一个需要先变形再分解的题，卡住，且**不尝试从原理往回推**，只反复检索“这像哪一类”。

常规判断。学得好，只是这道题超纲了。

诊断动作。问：“如果我把这个式子的符号全改一下，你还这么做吗？为什么？”

他的反应。“应该……可以吧？老师你告诉我这属于哪一类。”

判为第二种：路径之死。接得太牢，路径被完全省略；他有的是一张分类表，不是一条推理。

这是四种里最容易被误判为“学得好”的一种——他的全部读数都是优秀的。

处理。去光滑化：把那条路径还回去。具体做法是让他自己造一个能被这个方法处理的式子，再造一个不能被处理的，并说明界线在哪。

分辨点。若还了路径之后他能用，即为第二种；若还了路径他把路径也当作一条新的分类记下来，转查第四种。

记录三：显露之死

现象。四年级学生，坚持“乘法使数变大”。给他 0.5×6 的反例，他说“那是小数，小数不算”给他分数的例子，他说“分数是另一种”。他能为每一个反例找到一个例外理由，而且理由自洽。

常规判断。想法多、爱抬杠、不听讲。

诊断动作。问：“那你自己定一条规矩，说清楚什么时候变大什么时候不变大。”

他的反应。他真的会去定，而且定得出来（“整数才变大”）。这说明他有一个自洽的体系，不是没有想法。

判为第三种：显露之死。他建立了一套在他经验范围内完全成立的解释，因而不再需要修改。

处理。给一个他自己亲自验证的反事实——不能是别人告诉他的。让他拿计算器连续算一串数，自己看着结果变小。关键在于他的手按下去的。

分辨点。若他事后说“那我以前那个说法要改成……”，为回写发生；若他只是补一个例外条款（“整数和大于一的小数才变大”），说明他在修补而非改写，仍属第三种，需要换一个更根本的反例。

记录四：元层之死（封写）

现象。九年级学生，成绩优异，作业工整，从不迟到。任何新内容都能迅速接住并熟练使用。教师做了完整的去光滑化教学——把某个概念的历史困难重新做了一遍，他全程参与，态度端正。

下课后，他把整个过程记了笔记。

常规判断。很好的学生。或者：这套方法对他没什么额外效果，因为他本来就学得好。

诊断动作。问：“你笔记上原来那句话，现在要不要改？”

他的反应。“改成哪句？老师你说。”

判为第四种：元层之死。前三种的补救对他全部无效——不是因为补救做得不好，是因为他把每一样新东西都接住，包括那条路径本身，而没有一样旧东西因此改变。

根本特征：他不改自己的想法，他只是把新东西也接住。而且接得非常好。

处理。必须先做元层解锁（第三十四章第七节的三问），再做任何对象层的补救。次序反了，做多少都是无效功。

分辨点。元层解锁是否起效，看一个信号：他有没有在没有被要求的情况下，问出“那个是谁定的”这一类问题。没有这个信号之前，不要开始对象层的补救。

材料九 四种死亡的分辨速查

	给他一句他自己以前的说法，问“还对吗”	判为
说不记得写过	从未连上任何东西	一、土壤
说对，因为规则没变	规则记牢，路径不存在	二、路径
为它辩护，能自圆其说	自治体系	三、显露
说“对啊”，然后等你告诉他哪里不对	判断权已交出	四、元层

第四种的特征不是答错，是**把判断权交出去**。这个动作与第三十八章的读数四（他对 AI 说的第一句话）测的是同一件事，只是场景不同。

材料十 封写征候的现场清单

第十七章第六节给了三个征候。这里展开为可勾选的清单，供班主任在一个月内累计观察。

征候一：等

- ☐ 遇到没见过的题面，不动笔，直接抬头找人
- ☐ 等待时长从分钟级缩短到秒级
- ☐ 等的对象无所谓是谁（老师、同学、系统）
- ☐ 被要求“先自己想”时表现出明显不适

征候二：不改

- ☐ 订正了那道题，但下一道同类题仍然那样想
- ☐ 错因写成“粗心”“看错”，很少写成“我原来以为”
- ☐ 不主动翻回以前的作业
- ☐ 被指出错处时立刻改，但不追问为什么错

征候三：“本来就是”

- ☐ 问“为什么是这样”时答“本来就是这样”
- ☐ 追问时带轻微不耐烦
- ☐ 对“当年能不能定成别的”这类问题不感兴趣
- ☐ 认为追问来路是一件“没用的事”

读法。第三组勾得最多的那些学生，通常也是成绩最好、最配合的那些。这与第十七章推论二一致，而且这份清单的主要用途正是让这件事被看见——**因为在所有常规评价里，这批学生是最不会被关注的一批。**

第三编材料（续）：四个行业已经先演过一遍

本节材料的来源：SDE 站内文章《反抽的消失——黑灯车间》（张琼）、《劳动断链》（蔡燕）、《切痛》（高鹏），以及由此三篇与教育类文章合成的专著《还要老师吗？——判断力的最后一处现场》中的跨编定理。此处只取机制与量化形态，案例与论证过程见原作。

材料二十二 为什么必须走出教育学

本书前六编有一处结构性的弱点，必须在这里点明：**它的全部证据都在教育内部。**

一个只在教育内部成立的诊断，很难与“教育本来就这样”区分开。而如果同一条机制在三个与教育毫无关系的行业里、由完全不同的原因、在不同的年代各自发生过一遍，那么它就不再是一个关于教育的 judgment，而是一个关于**判断力如何形成**的判断。

站内已有三篇处理这件事的文章，各自出自一个行业。它们不是本书的例证，它们是**三次先行判决**。

材料二十三 制造业：停机之后，没有人再被世界顶回来

现象。黑灯车间——无人化的产线。它取消的不只是体力劳动，还取消了一个循环：一个焊工在二十年里，被物理世界的微小扰动反复穿过——材料的批次差异、湿度、设备的细微磨损、一次莫名其妙的失败——每一次都逼他在混沌中重新站稳一次。

被取消的是什么。不是技能，是**被顶回来的机会**。技能可以写进标准作业书，被顶回来这件事不能。

量化形态（枯竭期公式）。设 J 为一个组织里“能判断的人”的存量， A 为在岗人数， r 为单位时间的生成率， W 为一个人从新手到能判断所需的年限：

$$\Delta J = rA - J/W \quad \Rightarrow \quad t_{\text{枯竭}} \approx W \ln(J_0)$$

考虑在岗率 p 之后，真正的阈值是 $J < 1/p$ ——不是“没有能判断的人”，是“**能判断的人不够密以至于任何一班都可能没有一个**”。

三条推论：

一、**显影期与职业年限成正比**。软件行业的 W 短，所以早于医疗显影；教育的 W 最长，所以显影最慢——这解释了为什么教育界至今感觉不到。

二、**囤积老师傅只能推迟，不能阻止，且成本线性上升。**

三、这个公式对教育的含义是：教师队伍里“能守住岔口的人”的存量，同样服从这条式子。

还给教育的那一条。第三十三章说塌陷方向固定（学生先退场）。制造业的版本是：**没有人再被世界顶回来**。两者是同一件事在两个场里的名字。

材料二十四 医疗：一半在看病，一半在编码，中间没有人

现象。按疾病诊断分组付费一类的支付改革，把诊疗过程拆成可计价的编码单元。

被拆散的是什么。“感知—裁决—承担”这三件事，原本在同一个心智里走通。改革之后，感知归临床、编码归病案、承担归制度，**中间没有一个人把三件事连起来走一遍。**

量化形态（工序完整性指数）。

$$I = \# \{S=D=B\} / N$$

其中S为感知者、D为裁决者、B为承担者，统计三者为同一人的病例占比。配一个“承担挂载点比”M，再做偏离率×不良率的四格表。

这张四格表最重要的是左下角：“路径覆盖良好”与“判断已停产”，在所有现行指标上一模一样。

还给教育的那一条。本书第三十五章说闭合权的行使需要读取一个尚未发生的情境。医疗的版本给出了它的必要条件：**感知、裁决、承担必须落在同一个人身上。**一位教师如果只负责“讲”（感知）而不负责“这个孩子最后怎么样”（承担），他即使有能力守岔口，也没有理由守。

这一条对第二十八章有直接后果：学校的分工越细（教学与评价分离、任课与班主任分离、课堂与课后服务分离），I值越低，而这与教师的水平无关。

材料二十五 法律：认得出“不对劲”，需要一点冗余

现象。司法审慎与心证，被拆成两半：可形式化的那一半被保留并强化（要件、流程、说理格式），不可形式化的那一半被当作冗余丢掉。

关键概念：程序性奢侈。指那些在效率账上必然被判为冗余、却是警觉得以存活的条件东西——多问一句、多看一眼、把一份已经能定的卷再放一天。

量化形态（三个司法读数）。

一、程序性奢侈存量比P；

二、异常识别滞后分布的重心右移（认出“不对劲”要多久）；

三、**说理段与类案的文本重合度**——重合度越高，说明说理越是检索的产物而非判断的产物。

还给教育的那一条。第三个读数可以直接搬进课堂：**学生的解释与标准解释的文本重合度。**重合度越高，说明那段解释越是复述。这是第三十八章五个读数之外的一个候选，采集成本极低。

材料二十六 跨场定理：判断力的代际存活 $\propto \tau \times R$

四个场（制造、医疗、法律、教育）走同一条机制，可以抽出两个跨场量：

τ = 一段不稳定被外部闭合之前，留存的时长。

R = 闭合责任人是否可指认。

判断力的代际存活 $\propto \tau \times R$ 。

这个乘积形式是本节最重要的东西，因为它解释了两件本书前六编解释不了的事：

第一，为什么单靠“慢教育／留白”救不回来。那只补 τ 。若 $R=0$ （没有人为闭合负责，或“是模型说的”），乘积仍为零。

第二，为什么单靠“人类最终决定”也救不回来。那只补 R 。若 $\tau \approx 0$ （不稳定在零点几秒内就被闭合），乘积同样为零。

AI 的特殊之处在于它同时压两项：把 τ 压向零（提早供给），把 R 抹掉（“是它说的”）。这是本书第十四章“毒性＝老路×速度÷摩擦”那个式子的另一个写法，而且它比那个式子更有用——因为它给出了两条独立的干预路径，并说明单独走任何一条都不够。

可证伪：若某场 τ 极小而 R 清晰、判断力仍代际存活，本定理的乘积形式被推翻。

材料二十七 四行业的五个共有特征与三处实质分歧

五个共有特征（这五条同时成立，是“同一条机制”这个判断的根据）：

- 一、都不产出记录。被消灭的那样东西，从来不在任何账本上。
- 二、每一步都有理由。消灭它的每一个动作，在当时都是对的。
- 三、当期指标改善。每一步之后，可测量的东西都变好。
- 四、从业者感觉不到。没有任何人在任何时刻体验到损失。
- 五、重建需要一代人。

第五条对教育有一处加重：教育是唯一可能连重建都不行的——因为重建判断力需要有判断力的人来带，而教育系统正是生产这些人的地方。其他三个行业可以从教育系统里补充新人，教育系统没有上游。

三处实质分歧（不能直接搬用其他行业做法的理由）：

一、教育的对象是尚未成形的人。其他行业“先写下判断再执行”的做法，对尚未发生的東西无效。

二、其他三行业的判断有外部真值（零件合不合格、病人好不好、判决被不被推翻），教育没有。⇒ 本书全部教育读数必须绕道，只能测行为事件，不能测判断的正确性。这正是第三十八章那五个读数全部落在事件上的深层理由。

三、教育同时具备三种破坏方式（提早供给、评价外壳、格式化知觉），所以改革总是按下葫芦浮起瓢。

第三编材料（再续）：提前闭合的五种时机

本节材料的来源：SDE 站内文章《临界点燃》《临界回落》《判断预占》《认知早产》《悬而未决的守护》《留空位律》。此处取其各自指出的那一个时刻，合成一份时机表，用于补强本书第十八章（四种死亡与诊断次序）。

材料三十七 为什么需要一份时机表

本书第十八章按**死在哪一层**给出了四种死亡（土壤、路径、显露、元层）。那是一份**空间**上的分类。

站内那六篇文章合起来给出的是另一个维度：**闭合发生在什么时候**。这是一份**时间**上的分类，与四种死亡正交。

两份表合用，诊断会精确得多——因为同一种死亡，可以由不同时刻的闭合造成，而处理方式不同。

材料三十八 时机一：临界之前（点燃前的蓄积期）

那一刻。一个人在真正想通某件事之前，有一段张力持续累积的时间。这段时间里什么都没有产出，看上去什么都没发生。

站内那篇的判断：临界前的张力蓄积期才是根基，而点燃的那一瞬间不可计划。

教学上的后果。所有的教学设计都指向“点燃”那一刻——设计一个精彩的引入、一个恰到好处的启发。而可设计的只有蓄积期的条件，点燃本身不可设计。

一条堵死怀旧论证的补充。这段蓄积期在历史上**从来不是被设计出来的**，它是资源稀缺的副产品。这一条与本书第十九章第二节完全一致，而且它堵死了一种廉价的论证——**不能说“从前的教育更好”，从前只是更慢**。

共情越强的教师，闭合得越快。这是站内那篇给出的最不受欢迎的一条：教师闭合岔口的三个动因是共情反射、职业身份、制度记账，而第一个最强也最难克制——**看着一个孩子难受而不出手，违反的是本能，不是纪律**。

材料三十九 时机二：回落的那三十秒

那一刻。一个人已经进入过某个状态，然后正在退出来。退出的过程有一段时间——大约几十秒——在这段时间里，他还回得去。

站内那篇的判断：那三十秒在所有的账上都不存在；而承接不是帮助他完成，是阻止关闭。

教学上的后果。教师通常在两个时刻介入：孩子卡住时（提早供给），以及孩子做完时（评价）。而这两个时刻之间还有一个时刻，**没有任何教学法处理它**。

最实用的一个动作（本书认为它可以直接进第三十四章）：

把“这个等会儿统一讲”换成“**你先记在本子上，下课我来问你**”。

时间管理完全相同，区别只在有没有定性——前者把它关掉了，后者把它挂起来。

⚠️ 但说了就必须真去问。不去问，比一开始就说等会儿讲更糟。

材料四十 时机三：判断本应涌现之前（预占）

那一刻。在一个判断本应涌现出来的那个空隙里。

站内那篇的判断：不是替代已有的判断，是在判断本应涌现之前把空隙填满。

这与本书的“提早供给”是同一件事，但它指出了一个本书没说清的地方：被填掉的不是一个已经存在的东西，而是一个本来会长出来但还没有长出来的东西。因此损失是不可见的，连当事人也感觉不到——他不会觉得自己失去了什么，因为那样东西从未存在过。

三个可从平台日志直接算出的指标（本书第三十八章可以增列）：

一、求助前独立时长 t_1 的分布：原本是双峰的（一批很快求助，一批坚持很久），预占之后会左峰单峰化。

二、供给后检查率 c ：拿到答案之后，还有多少人会去检查它。

三、陌生任务首次猜想延迟 t_2 ：遇到没见过的东西，多久才敢提出第一个猜想。

⚠ 这三个指标全部掌握在供给方手里，而结果对供给方不利。这一句应当被本书第二十八章“超出学校能力范围的三件事”收录为第四条。

材料四十一 时机四：受孕之前（认知早产）

那一刻。比前面三个都早——在那个东西还没有开始形成的时候。

站内那篇的判断：不是被挤掉的，是从未受孕；格式把教师的知觉结构回写成分辨不出。

这一条最狠的地方在于它的对象是教师而不是学生。长期使用某种格式（任务单、四栏表、标准流程）之后，教师会逐渐看不见那些不落在格子里的东西——不是他不重视，是他的知觉结构被回写了，那些东西在他眼里不再显现。

这与本书第十七章的封写是同构的，只是发生在教师身上。本书从头到尾把封写写成学生的事，这是一个疏漏。教师同样会被封写，而且被封写的教师无法守岔口——因为他看不见岔口在哪里。

可观察的读数（本书第三十八章可增列）：格式提前量 $\Delta = t_c - t_f$ ，即“格子出现的时刻”减去“那个东西本来会形成的时刻”。只需要一块秒表。配一个撤格式测试：把任务单撤掉一节课，看还剩下什么。

材料四十二 时机五：意义被赋予的那一刻

那一刻。最晚的一个：不稳定已经维持住了，即将结束时，有人替它赋予了一个意义。

站内那篇的判断：教师承受悬空，却拒绝为它赋予确定意义——包括以最善意的方式。

“最善意的方式”是关键。它指的是这一类话：“你今天很努力”“这说明你在思考”“没关系，慢慢来”。这些话都是好意，而它们同样是闭合——它们把一段还未结算的东西，提前结算成了一个正面评价。

⚠ 一条优先于全书的硬边界（站内那篇自己设的，本书全盘接受）：

涉及求助信号、安全风险、不可逆伤害时，本书全部关于“克制”“不供给”“承受悬空”的主张，一律让位。

这条边界必须写在任何一次教师培训的最前面，而不是最后面。

材料四十三 五种时机与四种死亡的合表

时机	那一刻	最容易造成哪种死亡	对应动作
一、临界之前	张力还在蓄积，看上去没产出	土壤之死	不打断蓄积；共情反射是最大敌人
二、回落的三十秒	已进入过，正在退出，还回得去	路径之死	“记在本子上，下课我来问你”
三、涌现之前	判断本该长出来的那个空隙	元层之死	未闭合时长； t_1 双峰是否单峰化
四、受孕之前	那东西还没开始形成	显露之死（且发生在教师身上）	格式提前量 Δ ；撤格式测试
五、赋予意义时	已维持住，即将结束	元层之死	拒绝提前结算，包括善意的

这张表的用法：先用第十八章那张表定“死在哪一层”，再用这张表定“闭合发生在什么时候”。两个坐标定下来，处理动作才唯一。

最值得注意的是第三行与第五行：它们造成同一种死亡（元层），而时机相隔一整段——一个在开头，一个在结尾。这意味着一节课可以在两个完全不同的时刻，各封写一次。

第四编 两条赛道

第三编说明损害的机制与形态。第四编回答一个更根本的问题：发生学教育与发现学教育，究竟是同一件事的两种做法，还是两件事？本编先调和“派生”与“并列”之间那处必须处理的张力，再给出三条可操作的判别，然后处理换道的方向性、秤、过渡期与学校层面的含义。

第二十三章 派生还是并列：一处必须调和的张力

在知识史上是派生，在学习者身上是两条赛道；光滑化使二者在纸面上可互换，封写使二者在人身上不可互换。

一、一个会被当场抓住的矛盾

本书到此为止说了两句看上去互相打架的话。

第一句在第二章：发现＝发生一路径－土壤。按这句话，发现是发生的派生物，二者不是并列关系，而是一个从另一个减出来的。

第二句是本编的题目：两条赛道。而“赛道”这个词暗示的是并列——两条各自独立的路，可以选这条也可以选那条。

一个认真的读者会立刻问：既然是派生，怎么又是并列？既然发现学教育不过是发生学教育被减掉两项的结果，那它就不是“另一条路”，只是“同一条路的残缺版本”。

这个问题必须正面回答，不能靠修辞绕过去。因为如果绕过去，后面所有关于“换道”的论述都建立在一个含混的比喻上。

二、两句话都保住的写法

本书给出的调和是：

—— **在知识史上是派生，在学习者身上是两条赛道。**

两个层面，两种关系。

在知识史这一层，派生关系成立。一个概念先被某个人在某片土壤里、沿某条路径做出来，然后被磨光、被搬运、被写进教科书。这个次序是事实，不可颠倒。教科书上的每一个成品，都能追溯到某一次发生。

在学习者这一层，两条赛道成立。一个孩子面前有两种可能：一种是接住那个成品，一种是让它在自己身上重新发生一次。这两种可能的**结果不同**，而且——这是关键——**不可事后互换**。

三、光滑化与封写：一个使它们可换，一个使它们不可换

上面那两句需要一个机制来支撑，否则仍然是修辞。机制是本书前面已经给出的两样东西，这里第一次把它们放在一起用。

光滑化使二者在纸面上可互换。同一个结论，一个孩子是自己走出来的，另一个孩子是被端过去的——**写在纸上一模一样**。写在试卷上一样，写在笔记本上一样，写在教学质量报告上一样。这就是为什么这两件事长期以来没有被区分：**在一切书面形态上，它们是同一个东西**。

封写使二者在人身上不可互换。走过来的人和被端过去的人，此后的接收结构不同（第十七章），而且不能事后补齐（第十七章第四节的活化能）。

两句合起来：

—— **纸面上没有差别，人身上有差别，而且这个差别是单向累积的。**

这就是“两条赛道”的硬根据。不是跑法不同，是**跑久了换不回来**。

四、“赛道”这个词的两个危险

比喻用得越顺手，越要交代它的失真处。“赛道”有两个危险，必须当场堵住。

危险一：它暗示两条都通向比赛，各有千秋，可以搭配。教师听完之后最常见的落地形态就是“我平时发现学，偶尔发生学一下”——把新路当成一种可选的教学花样，在公开课上用一次。

堵法是补一句：**这两条道的终点不在同一个地方**。不是同一场比赛的两条跑道，是两场不同的比赛。

危险二：它暗示对称。赛道通常是并排的、等价的、可以互换的。而本书要说的关系恰恰不对称——第二十五章将专门处理这一点。

所以更准确的形象不是两条平行的跑道，而是：

一条带单向闸的岔口。

从发生学这一侧可以随时退回发现学一侧（把成品端出来就行），反过来则要先开锁。这个形象比“赛道”准确，只是不如“赛道”好记。在教师培训现场用赛道，在书面论述里用单向闸。

五、一个必须回答的追问：那是不是所有内容都该走新路

派生关系还有一个重要含义，容易被忽略：

既然发现是发生的派生物，那么对于绝大多数内容，接住派生物就够了。

一个孩子不需要让每一个概念都在自己身上重新发生一次。第三章的账已经说明这在数量级上不可能，而且第七章的三大成就说明这也不必要——跨代存活、规模化、校准，全部建立在“接住成品”这件事上。

所以本书的主张要写得比“换道”更精确：

不是把所有内容从一条道搬到另一条道，是在一条道上保留若干处真正的发生。

数量上，本书给出的估计是保守的：一个学期，一门课，三到五处。其余内容照常。

这个数量看上去少得可疑，理由在第二十二章第六节已经给过——方向的建立不需要频次。而它的可行性正来自这个少：三到五处不与任何进度冲突，不需要额外课时，不需要重编教材。

六、为什么这个调和很重要

本节说明此处张力为什么值得花一章来处理，而不是当作用词不严谨一笔带过。

第一，它决定了处方的规模。如果两条道是并列的，那么换道意味着推翻整套现行教学，代价巨大，几乎不可能推行。如果是派生加单向闸，那么处方就是在现行教学中保留若干处不被光滑化的地方——代价小到任何一所学校都能承担。

第二，它决定了对老路的态度。并列关系会导向“老路是错的另一条路”，而派生关系导向“老路是同一条路被减掉了两项”。后者才与第七章的三大成就相容。

第三，它决定了这本书能不能被反驳。“两条赛道”作为比喻不可反驳，而“光滑化使其纸面可换、封写使其人身不可换”是一个可以被检验的命题——第三十九章的实验正是检验它。

七、可观察的动作

拿出一份学生的满分试卷，问一句：从这张卷子上，能不能看出他是走过来的，还是被端过去的？

看不出来。这就是本章第三节那句“纸面上没有差别”的经验形态。

而如果一位教师认为自己能看出来，可以追问：靠什么看出来的？通常答案是那些不出现在卷面上的东西——他课上问过什么，他改过自己什么说法，他对一道没见过的题的第一反应。这些正是第三十八章那几个读数要采集的东西。

八、与相邻说法的分界

科学哲学中有关于“发现的语境”与“辩护的语境”的经典区分：一个理论是怎么被想出来的，与它凭什么被接受，是两件不同的事。本章的派生关系与之相邻，但方向相反：那一区分是为了**保护辩护语境的独立性**，主张想法的来源与它的正当性无关；本章要指出的是，在**教学**这个场景里，来源与正当性无关这一点恰恰造成了问题——因为学习者接住的是一个只有辩护没有来源的东西，而他需要来源不是为了判断真假，是为了让它写得进自己的底盘。**科学共同体可以不关心来源，一个正在学习的人不能。**

第二十四章 三条判别：终点、计分、闸门

判断一堂课在哪条道上，不看动作，看三样：终点是不是已定、记的是哪套分、能不能随时退回去。三样里最要紧的是第二样。

一、为什么需要三条而不是一条

第六章给过一条判据：不看动作看终点。那条判据用于识别单节课，简单可靠。

而在学校层面、学期层面、改革层面，一条不够。原因是：**终点这一条可以在单节课上做到，却在一个学期里被系统性地取消掉**。一位教师可以在公开课上抛出一个真正未定的问题，而在其余七十九节课上按进度表走。

所以需要三条判别，它们分别对应三个不同的时间尺度：

判别一（终点）：一节课的尺度。

判别二（计分）：一个学期的尺度。

判别三（闸门）：一个学段的尺度。

三条合用，才能判断一所学校真正在哪条道上。

二、判别一：终点不同

内容。发现学的终点是**存量**——掌握了多少；发生学的终点是**发生力**——当场能不能做出这个判断。

操作。翻开教案，看结论有没有被写下来。第六章第三节已经给过完整用法，这里不重复。

为什么它必须放在第一位。因为终点若相同，动作再活泼也只是发现学里比较热闹的一条内道。大量“探究式改革”在一年之内退化为“活动多的发现学”，根源就在这里：**换了跑法，没换终点。**

它的局限。这一条容易通过。任何一位教师都可以在某节课上不写结论。它是必要条件，不是充分条件。

三、判别二：计分不同

内容。发现学记答对率；发生学记回写。

操作。看期末评价的口径。看学校用什么数据评估教师。看家长会上出示的是什么。

为什么它是三条里最要紧的一条。

第五章已经论证过：这条链的稳固不在于有人捍卫它，而在于每一环都能单独优化且立刻在读数上兑现；而历次改革的实际命运是——**能在现行读数上兑现的做法被保留，其余被悄悄放弃。**

所以：

没有第二套秤，赛道分立就只是口号；而一旦第二套秤承担了分配与拒绝的功能，它就会变成第一套秤。

后半句是本书成稿之后由材料二十八至三十一逼出来的补充，它比前半句更硬：评价装置的产品不是“关于学习的信息”，而是**可接受的拒绝**。任何一把新秤，若不能承担分配与拒绝的功能，就不会被采纳；而一旦它承担了，第三十八章那三条纪律当天失效。

因此这不是一句抱怨，是一条设计标准：新读数必须被造得“不能用于分配”。五个读数里最接近这个要求的是读数二（只设下限因而无法排序）与读数五（没有可训练的答案）。

一所学校若宣称转向发生学教育，而期末评价的口径一个字没改，那么半年之内它必然滑回老路——不是因为理念不够坚定，是因为秤在那里。教师每天面对的回报结构没有变，而回报结构每天兑现一次，理念一学期兑现一次。

第二十六章将专门处理秤，包括在无权改变升学考试的条件下，一所学校还能做什么。

它的局限。这一条最难通过，因为它需要制度层面的动作，不是个人可以完成的。

四、判别三：闸门单向

内容。发生学-发现学随时可换：把成品端出来就行，光滑化是免费的、瞬时的。发现学-发生学则必须先元层解锁，补课换不动，而且活化能随学段升高。

操作。看这所学校在哪个学段做这件事。三年级做和九年级做，难度不是同一个量级。

为什么它属于学段尺度。因为它的含义是关于**时机**的：换道的最佳时机永远是现在，越往后越贵。第二十五章将专门论证这个不对称。

它的实践含义。一所学校如果只在高中做发生学教学，效果会显著低于在小学做，而且这个差别不是学生成熟度造成的，是活化能造成的。**这条判别给出的政策建议是：资源应当优先投向低学段。**这与常规做法相反——常规做法是把创新性教学放在高年级，因为“低年级要先打好基础”。

五、三条判别的对照表

判别	尺度	问什么	难度
一、终点	一节课	结论写在教案上了吗	容易通过
二、计分	一学期	期末的口径改了吗	最难通过
三、闸门	一学段	在哪个年级做	需要长期规划

三条合起来，可以用来做一次学校层面的自我定位：

- 三条全不通过：完全在老路上。这是绝大多数学校的状态，而且这不是指责——第七章已经说明老路的成就。
- 只通过第一条：有若干位教师在个人层面做尝试。这些尝试有价值，但按第五章的机制，它们的存活期通常是一到两个学期。
- 通过第一、二条：这所学校真的在换道。
- 三条全通过：极少见。

六、两条赛道的完整对照

把前面各编的结论汇总成一张表，它可以直接用于教师培训与家长沟通。

对照项	发现学教育	发生学教育
教学动作	搬运传道接受	张开岔口守住引发回写
终点	一个已写在教案上的结论	一个当场被做出来的判断
教师定义	损耗最小的传输者	守住岔口的人
AI 的位置	更强的答案源	碰撞中的对手
计分	答对率、完成度、进度	回写率、未闭合时长、先在感降幅
失败形态	遗忘（对象层）	封写（元层）
成本结构	低（可批量搬运）	过去极高，AI 之后第一次可负担
换道方向	随时可退回此道	须先元层解锁方可进入
内容占比	绝大多数内容	一学期三到五处

最后一行是第二十三章第五节的结论，必须留在表上——没有它，这张表会被读成“要把全部教学翻过来”，而那既不可能，也不必要。

七、一个常见的误用：把判别当成评价

必须堵住一个可预见的误用：把这三条判别拿去给教师打分。

这三条不是评价工具，是诊断工具。

理由与第三十八章将给出的读数纪律相同：任何一把秤一旦被用于结算，就会开始被优化，而被优化的秤会停止测量它原本要测的东西。如果“终点未定”被写进听课评分表，那么教案上就不会再写结论——写在别的地方，或者不写。

三条判别的正确用法是自用：一位教师用它诊断自己的课，一个教研组用它讨论一节课，一所学校用它做自我定位。**用它评人，它会在一个学期内失效。**

八、可观察的动作

｜ 用这三条给你所在的学校做一次定位。不要写下来，也不要告诉任何人。

最后半句是认真的。第七节已经说明了理由。

九、与相邻说法的分界

教育评价领域有大量关于“形成性评价”与“终结性评价”的区分，与本章的计分判别相邻。差别在于：那一区分处理的是**评价的时机与用途**（过程中反馈还是最后打分），两者测量的对象是同一个——掌握程度。本章的计分判别处理的是**测量对象本身的更换**：从“接住了多少”换成“改了什么”。一次形成性评价如果测的仍然是掌握程度，它在本书的框架里仍然属于老路的读数。这个差别很容易被混淆，因为形成性评价在形态上确实更接近本书的主张——它也发生在过程中，也不打分。

第二十五章 闸门为什么是单向的

老路的能力是新路能力的子集：走过新路的孩子随时能回去接成品，而且接得更快；反过来不成立。这条不对称是全书唯一能回应升学压力的东西。

一、这一章要回答的现实问题

前面所有的论述，在一线会遇到同一个问题，而且这个问题不解决，其余全部无效：

换道，成绩会不会掉？

这个问题来自家长，来自校长，来自教师自己。它不是保守，是负责任——一个教师如果不考虑这一点，他就是在拿别人的孩子做实验。

本章的回答是：**长期不会，短期会**。长期不会的理由在本章，短期会的处理在第二十七章。

二、不对称的内容

先把命题写清楚。

老路的能力是新路能力的子集。

具体地说：一个在新路上练出发生力的孩子，随时可以回到老路上接成品，而且接得更快——因为他知道那个成品是怎么来的。反过来不成立：一个只在老路上练过的孩子，不会因此在新路上更强。

这就是判别三所说的单向闸的完整含义。它不是一个比喻，是一个可以被检验的断言。

三、为什么子集关系成立：三条理由

理由一：接成品需要的是理解结构，而走路径的人有更多结构。

一个成品之所以难接，通常不是因为它复杂，而是因为它没有可以挂靠的地方（第十八章的第一种死亡）。而一个走过路径的孩子，手里有一整套关于“这类东西大致长什么样、通常从哪种困难里长出来”的经验。新的成品落下来时，他有更多的挂钩。

理由二：走过路径的孩子知道成品的边界。

一个只接住结论的孩子，会把结论用在它不成立的地方（“乘法使数变大”用到小数上）。而知道“内角和是平面的性质”的孩子，在遇到不寻常的情形时会警觉。在考试中，边界意识直接表现为不掉坑。

理由三：走过路径的孩子对“不会”这件事的反应不同。

这一条最实用。考试中遇到没见过的题面时，只在老路上练过的孩子的第一反应是检索（我见过这类题吗），检索失败即停止；走过路径的孩子的第一反应是尝试（我能不能从别处推一点出来）。

在难题上，这个差别常常就是全部差别。

四、反方向为什么不成立

反过来的方向为什么不行，前三编已经给出了机制，这里汇总一句：

因为老路的每一次成功，都在加固那条“东西是现成的、我不必因它而改”的元规则（第十七章）。所以在老路上练得越多、越成功，进入新路所需的活化能越高。

不是没有帮助，是有反作用。这是“子集”这个词容易掩盖的一点：子集关系说的是能力的覆盖，而这里还有一个额外的负项——**练得越好，锁得越硬。**

第十七章第四节的性质三（活化能随学龄升高）正是这个负项的形式。

五、这条不对称的可检验形式

本章的命题不能停留在论证上，必须给出可被推翻的形式。

预测：在控制初始水平之后，接受过若干次真实发生学教学的学生，与对照组相比——

一、在**复现类题目**上不显著更差；

二、在**迁移类题目**上显著更好；

三、在**未见题型**上的尝试时长显著更长，放弃率显著更低。

推翻条件：如果第一条不成立——即实验组在复现题上显著更差——那么“子集”这个说法即被推翻，本章必须改写为一个取舍关系（用一部分掌握度换一部分发生力），而这将极大削弱全书对一线的说服力。

本书把这一条写在这里，而不是藏在附录里，因为它是全书**最容易被现实推翻**的一个断言，也是最应当被优先检验的一个。

六、一个必须承认的限度

上述不对称有一个边界，必须讲明，否则它会被当成万能挡箭牌。

它对短期不成立。新路上花掉的时间，在短期内确实没有用于练习复现题。如果考试就在两周之后，那么把这段时间用于刷题会更有利。这一点没有争议，也不必回避。

它对纯记忆型内容不成立。大量内容不涉及任何判断，只需要记住（字词、年代、公式的形态）。在这类内容上，走路径没有额外好处，只有额外成本。第二十三章第五节的“三到五处”正是为此设的量：**新路只用在有冲突的地方，不用在纯记忆的地方。**

它对已被封写较深的学生成立得更慢。按第十七章，活化能随学龄升高，所以同样的干预在小学与高中的效果不同。这不是免责，是资源配置的依据（判别三）。

三条限度合起来，把本章的主张收缩到一个可辩护的范围：在有冲突的内容上，对未被深度封写的学生，在一个学期以上的尺度上，子集关系成立。

七、对家长的说法

本章的内容最终要被翻译成一句家长听得懂的话。经过现场检验有效的版本是：

走过一次的孩子，回来接现成的东西会更快，因为他知道那东西是怎么来的。反过来不行——接得再多，也不会因此学会自己走。

后面要跟一句诚实的话，不能省：

头几周成绩会难看。这一点我现在就说，不等月考之后再解释。

第二十七章将说明，绝大多数换道尝试死在第三周，而死因几乎总是这一句没有提前说。

八、可观察的读数

在一次考试后，把学生分成两组：做出了最后那道难题的，和没做出来的。然后只问没做出来的那些人一个问题——你在那道题上花了多久，试过什么？

回答会分成两类：一类说“看了一下不会就跳过了”，一类能说出自己试过什么。

这两类的差别，与他们的总分未必相关，而与本章的不对称直接相关。这个读数不需要任何额外测试，一次常规考试之后花十分钟即可。

九、与相邻说法的分界

学习迁移研究中有关于“近迁移”与“远迁移”的大量工作，本章的子集关系与之相邻但不同：迁移研究关心的是已习得内容能否用于新情境，其自变量是训练方式（变式、对比、原理教学）；本章关心的是接收结构本身的差异，其自变量是那个内容当初是被走出来的还是被端过去的。差别在于：迁移研究的两组学生学的是同一个东西，只是练法不同；本章的两组学生在纸面上学了同一个东西，而在接收结构上得到的是不同的东西。这个差别在任何以题目表现为唯一因变量的研究设计中都无法显现——这也是第三十八章的读数不落在题目表现上的原因。

第二十六章 秤：没有第二套秤，讲什么都没用

一所学校若宣称转向发生学教育而期末评价的口径一个字没改，它必然在半年之内滑回老路——不是因为理念不坚定，是因为秤在那里。

一、这一章的地位

第五章证明了筛选机制的存在，第二十四章把“计分”列为三条判别里最要紧的一条。本章把它变成可执行的方案。

需要先说清本章的地位：**它比第五编的任何一条具体教法都更前置。**

理由是第五章第七节那句话：认同不改变回报结构。一位完全认同本书全部主张的教师，回到教室之后仍然面对同一张成绩表、同一次月考、同一场家长会。他的认同一学期兑现一次，他的处境每天兑现一次。

所以本书对改革次序的主张与常规做法相反：

丨 **先换秤，再换做法；至少，秤和做法同时换。**

二、新秤的四条设计标准

第八章说过，昂贵的秤不会被使用；不被使用的秤等于不存在。所以新读数必须满足四条，缺一条就会在一个学期内被弃用。

标准一：不占用教学时间。能从已有的课堂活动中自然采集，或与常规作业合并。

标准二：不依赖学生配合。不能靠问他“你懂了吗”“你有没有改变想法”。

标准三：不依赖教师的主观判断。必须是事件，不是感觉（第十六章第九节：事件便宜，状态昂贵）。

标准四：难以被应试训练。如果一个读数可以被专门训练出来，它会在被用于评价的那一天起停止测量。

三、五个读数

按上述标准，本书给出五个。这里列出形态与采集方式，完整的使用纪律在第三十八章。

读数一：回写率。迁移题正确率减复现题正确率。与常规作业合并，当堂一次、一周后一次。

读数二：未闭合时长。从抛出问题到任何一方给出答案的秒数。录音的时间戳即可，零成本，也最难作假。

读数三：先在感指数。五题小量表，测他在多大程度上认为某对象“本来就是这样”。一学期两次。

读数四：对 AI 说的第一句话。索取型还是对抗型。对话记录本身就是数据，完全不依赖配合。

读数五：面对空位题的反应。停止尝试前的时长，以及有没有朝那个方向走出过一步。一学期一题。

五个之中，读数二与读数四是**自动的**——不需要任何人做任何额外动作。一所学校如果只做得了两项，就做这两项。

四、三条使用纪律

新秤会以三种方式失效，每一种都发生过，因此必须预先写下纪律。

纪律一：用于诊断，不用于排名。

一旦回写率被用来排教师名次，它会在一个学期之内被优化成一个新的答对率——教师会开始训练学生“表演回写”。让孩子在下课前说一句“我原来以为……现在我觉得……”是极容易训练的，而训练出来的那句话与真实的回写没有任何关系。

纪律二：不单独使用。

五个读数中任何一个单独拿出来都可以被应付。四五个合起来才具有诊断力。尤其读数二与读数四必须配对看：一节课未闭合时长很长，而学生说的全是“这个怎么做”，说明只是拖了时间。

纪律三：未闭合时长不设上限。

它只设下限。把它变成越长越好的指标，会立刻产出为拖时间而拖时间的课。

三条纪律的共同根据是一句话：**任何一把秤，一旦被用来结算，就会开始被优化；而被优化的秤会停止测量它原本要测的东西。**这一点在老路的读数上已经发生过一次（第八章），不应该在新秤上再发生一次。

五、在无权改变升学考试的条件下能做什么

这是本章必须正面回答的问题，因为绝大多数学校处在这个位置上。

不能做的：假装升学考试不存在；用新读数替代旧读数；要求教师承担成绩风险。这三条都会让改革在第一次月考后终止。

能做的，按可行性排序：

第一，加一栏，不动原有的任何一栏。在现有的教学质量报告上，增加一栏回写率或未闭合时长。不设权重，不参与评价，只记录。这一步几乎没有阻力，而它的作用超出预期——一个被记录的东西，就开始存在。

第二，把新读数用于教研，不用于考核。教研组听课课后讨论未闭合时长与学生对 AI 说的第一句话，而不讨论“这节课上得好不好”。这一步把新秤放进了一个不结算的场合，因而不会触发纪律一的失效。

第三，在期末的教师工作总结里，要求写一件事：这学期我在哪三处让学生真的卡住了，各发生了什么。只写事实，不评分。这一条把“三到五处”（第二十三章第五节）变成了一个有记录的承诺。

第四，在家长会上出示的东西里，加一页。这一页不放分数，放三个孩子改过的旧说法。这一步的作用是改变家长对“什么算学好了”的理解，而家长的理解最终会反过来影响学校的压力结构。

四条都不需要任何上级批准，都不与升学考试冲突，都可以在一所学校内部完成。

六、一个必须承认的天花板

上述四条能做的事情，全部停留在**记录与讨论**的层面，没有一条进入**结算**。

而第五章的机制说得很清楚：不结算的东西会被筛掉。

所以本章诚实的结论是：**在升学考试不变的条件下，一所学校能做的事情有一个明确的天花板，而这个天花板低于本书主张所需要的高度。**

本书不提供绕开这个天花板的办法，因为不存在这样的办法。能提供的只有两点：

第一，第二十五章的不对称。新路的能力覆盖老路的能力，所以这不是一个零和的取舍——在一个学期以上的尺度上，做这件事不会损害升学表现。这一点是全部说服工作的基础。

第二，天花板之下仍有相当大的空间。从“完全没有记录”到“有一栏记录、有一次教科研论、有一页家长会材料”，这个跨度不小，而且它是可以自己走完的。

七、可观察的动作

在你们学校下一次的教学质量报告上，加一栏，标题写“回写率”。第一次可以是空的。

第二十一章第八节说过：那一栏的空白，就是全部问题。而把这一栏加上去，是让它停止空白的第一步，也是唯一不需要任何人批准的一步。

八、与相邻说法的分界

管理学中关于“平衡计分卡”一类的主张，与本章“加一栏”的做法形态相似：都是通过增加非财务、非短期的指标来纠正单一指标的偏向。差别在于两处：一是本章的新指标测的不是同一件事的另一个侧面，而是一个此前完全未被测量的维度；二是本章明确主张新指标不进入结算，而平衡计分卡的整个用意恰恰是让它们进入结算。这个差别不是保守，是从第八章的机制推出来的——进入结算的那一刻，它就开始变成一个新的答对率。

第二十七章 过渡期：会死在第三周的那些改革

换道之后的头几周，读数一定会变难看；这一点如果不在开始之前讲明并约定，绝大多数尝试会死在第一次月考之后——不是因为它无效，而是因为它在旧秤上暂时看起来无效。

一、为什么必须单写一章

一项改革的死因，往往不是它错了，而是它在某个特定时刻看上去错了，而那个时刻没有人提前打过招呼。

本章处理的就是那个时刻。它在时间上非常具体：开始后的第二到第六周，第一次月考前后。

二、过渡期会发生什么

读数会下降，而且下降是真实的。

理由不复杂。孩子在旧道上练成的全部本领——快速接住、准确复述、按套路作答——在新道上暂时不再得分；而新道上要用的本领——忍住不问、把想法说出来接受反驳、回头修改自己——还没有长出来。

在这个交接的空档里，两套本领都不完整。这不是过渡期的副作用，这就是过渡期的定义。

具体表现有五项：

- 一、课堂进度变慢，单位时间覆盖的内容减少。
- 二、当堂正确率下降。
- 三、部分学生表现出明显的不适——尤其是原本成绩最好、最适应旧规则的那些。
- 四、家长开始询问。
- 五、教师自己开始怀疑。

第三项值得单独说明：**最优秀的学生在过渡期里反弹最强**。这与第十七章的推论二一致——他们在旧规则上投入最多、获益最大，因而对规则改变最敏感。这一现象容易被误读为“这套方法不适合优等生”，而按本书的机制，恰恰相反。

三、第三周的那个时刻

把时间轴画细一点。

第一周：新鲜。学生觉得有意思，教师觉得有希望。读数尚未变化。

第二周：开始不适。学生问“到底答案是什么”，教师第一次感到守不住。

第三周：这是最危险的一周。新鲜期过去，效果未现，而进度已经落后。此时教师面对的是一个非常合理的判断——“我们是不是该先把进度赶上来，等期末再说”。

这个判断在当时是正确的。它符合全部可见的证据。

第四到六周（第一次月考）：成绩下滑被数据确认。如果之前没有约定，此刻的结论是确定的：这个方法不行。

绝大多数尝试死在这里，而且死得完全合理。

四、唯一有效的处理：提前约定

本章没有巧妙的办法，只有一个笨办法，而这个笨办法必须在**开始之前**完成：

提前讲明过渡期会发生什么，并约定：在过渡期内不用旧秤做判断。

约定包含四项，缺一不可：

一、时间边界。明确说出过渡期有多长（建议一个学期），以及在什么时候用什么标准做第一次真正的评估。没有边界的“请再等等”会耗尽所有人的耐心。

二、承认代价。明确说出会下降的是哪几项，大概什么幅度。**不要说“不会影响成绩”**——这句话在第四周会被数据证伪，而一旦被证伪，此后所有的话都失去可信度。

三、约定观察什么。过渡期内看的是新读数（未闭合时长、对 AI 说的第一句话），而不是正确率。这一条把“再等等”从一个恳求变成一个有内容的安排。

四、给出退出条件。明确说出在什么情况下应当停止这项尝试。这一条最反直觉，也最重要——**一个不给退出条件的改革，会被正确地怀疑为不可证伪**。建议的退出条件是：一个学期之后，回写率与未闭合时长两项均无改善，且复现题成绩显著下降。

五、对三类人的三种说法

同一件事对不同的人要用不同的说法，因为他们担心的不是同一件事。

对家长。他们担心的是孩子吃亏。有效的说法在第二十五章第七节：走过一次的孩子回来接现成的东西会更快，反过来不行；头几周会难看，我现在就说。**要点是“我现在就说”这四个字——提前说的下降是计划的一部分，事后解释的下降是失败。**

对校长。他担心的是数据与责任。有效的说法是给出边界与退出条件（第四节的一、四项），把这件事从“一次没有终点的理念实验”变成“一个有起止、有观察项、有退出条件的安排”。

对教师自己。他担心的是“我是不是在耽误这些孩子”。这个担心最真诚，也最难回应。可用的回应是第二十五章的不对称加上第二十三章第五节的量：**一个学期三到五处，其余照常。**这个量小到不构成“耽误”，而它要建立的东西不需要频次。

六、过渡期的两个真实风险

必须说明两个不属于“暂时难看”的风险，它们是真的风险，不能靠约定化解。

风险一：做成了伪未定终点。教师在过渡期的压力下，会不自觉地把未定的终点悄悄拉回来——问题问得很开放，评价仍然按标准答案。第六章第七节说过，这种课比直接讲授更消耗信任。**它在过渡期里出现的概率最高，因为压力最大。**

对策是让教研组用第六章第八节的五问互相检查，而不是靠自觉。

风险二：把新做法叠加在旧安排上。原有的进度、作业、考试全部不变，再加上新的环节。结果是教师负担陡增，学生时间被占满，两个月后因为不可持续而终止。

对策是**替换而非叠加**：三到五处发生学教学，替换掉原本用于讲解那部分内容的课时（第八章第六节的时间论）。**如果这一条做不到，宁可不做。**

七、可观察的读数

在开始之前，把“预计会下降的项目与幅度”写下来，交给校长和家长代表各一份。过渡期结束后，拿出来对照。

这个动作的价值不在于预测得准不准，而在于它把“下降”从一个突发事件变成一个被预告的事件。**被预告的坏消息不会终止一项改革，未被预告的坏消息会。**

八、与相邻说法的分界

组织变革研究中有“变革低谷”一族描述：任何变革在实施初期都会经历生产力下降，随后回升。本章的过渡期与之形态相同，但有一处关键差别必须点明：那一族描述的低谷源于**熟练度的暂时损失**（新流程还不熟），因而随着熟练自然回升；本章的下降源于**测量对象与目标的错配**——即使完全熟练之后，旧读数的某些项也未必回到原来的高度（比如进度必然更慢）。所以处方不同：前者是“挺过去就好了”，后者必须**同时更换观察项**，否则挺过去之后仍然会被判为失败。

第二十八章 学校与政策层面的含义

换道不是一项教学法改革，而是一次读数体系的增补；凡是把它当作教学法来推行的，都会在第二十四章判别二那一关上失败。

一、本章的范围

前面各章的对象是教师与课堂。本章处理的是学校与更上一级的层面，因为本书的若干主张在个人层面无法完成。

需要先划清边界：**本书不是一份政策建议书**。作者没有政策制定的经验，也没有相应的证据基础。本章只做一件事——把前面各编的结论翻译成学校层面可以识别的形态，并指出其中哪些超出了一所学校的能力范围。

二、五条学校层面的含义

含义一：这不是教学法改革。

如果把它当作教学法来推行——组织培训、示范课、评优课、编写教案模板——那么按第五章的机制，它会在一到两个学期内被筛选成“一批热闹的导入环节”。

它的正确形态是**读数体系的增补**：先加一栏，再进教研，最后进总结与家长会（第二十六章第五节）。教学法的部分反而可以慢慢来。

含义二：资源应当优先投向低学段。

按判别三，活化能随学龄升高。所以同样的投入，在小学的效果显著高于在高中。这与常规做法相反——常规做法把创新教学放在高年级，理由是“低年级要先打好基础”。而按第十七章，**低年级恰恰是那把锁被写下的时候**。

含义三：优秀学生需要被单独考虑。

按第二十章，主动使用 AI 自学的是最自觉的那批孩子，暴露剂量最高。而学校的常规关注方向恰好相反——资源投向后进面。

这不是主张减少对后进生的投入。是指出：**优秀学生现在需要一种此前不需要的照看，而这种照看的内容是“给他装上对手与回写口”，成本很低**。

含义四：过渡期必须被制度化地保护。

按第二十七章，绝大多数尝试死在第三周。而个体教师无法保护自己的过渡期——他没有权力对家长说“这个学期不看成绩”。

这是一件只有学校能做的事，而且它比任何培训都重要：**一份写明起止、观察项与退出条件的书面安排**。

含义五：不要评优。

第二十四章第七节已经说明理由。一旦“发生学教学”成为评优的一个类别，它会在一年内产生一批表演性的公开课，而这批课会成为此后所有人对这件事的印象，从而使真正的实践更难开展。

这一条是本书对学校提出的唯一一条“不要做”的建议，而它可能是最重要的一条。

三、超出学校能力范围的三件事

诚实地列出本书主张中学校做不到的部分。

第一，升学考试。这是最重的一把秤，学校无权改变。第二十六章第六节已经说明，这构成一个明确的天花板。

第二，产品端的配置。第十四章第八节的两个部件（对手、回写口）如果由产品端实现，效果会好得多，而且成本几乎为零。但这不在学校的能力范围内，也不在教育行政部门的能力范围内——它属于产品设计与行业规范的问题。

第三，长周期的证据。第二十一章的三段曲线需要三到五年的追踪数据，任何单所学校都做不了。这需要研究机构介入。

三件都列出来，是为了让学校不去做力所不及的事，从而把有限的精力放在第二节那五条上。

四、给不同角色的一句话

给校长：先加一栏，不要评优。

给教研组长：用第六章的五问互相检查，不要用它打分。

给教师：一学期三到五处，其余照常；而且必须是替换课时，不是叠加。

给家长：先有农庄，再上飞机；不是不能用，是别一个人用。

给做产品的人：你们比学校更容易装上那两个部件，而且成本近乎为零。第十四章第七节那把尺子——把“学习”换成“接收成品”——现在可以拿去量自己的首页第一句话。

五、第四编的结论

第四编到此结束，五句：

第一，在知识史上是派生，在学习者身上是两条赛道；光滑化使二者纸面可换，封写使二者人身不可换。而对绝大多数内容，接住派生物就够了——新路只用在一学期三到五处。

第二，三条判别对应三个尺度：终点（一节课）、计分（一学期）、闸门（一学段）。最要紧的是第二条。

第三，闸门是单向的：老路的能力是新路能力的子集。这是唯一能回应升学压力的东西，也是全书最容易被现实推翻的断言，因此被写成了可证伪的形式。

第四，没有第二套秤，讲什么都没用了。五个读数按“便宜、不依赖配合、难以训练”设计；三条使用纪律的根据是同一句话——**被用来结算的秤会停止测量。**

第五，过渡期是绝大多数尝试的死因，而唯一有效的处理是提前约定，其中必须包含退出条件。

第五编要回答的是：那么，具体怎么上课。

六、与相邻说法的分界

教育政策研究中有大量关于“改革为何难以持续”的分析，通常归因于教师能力、资源不足评价体系或行政推动方式。本书与它们最大的差别在于因果次序：那些分析多把评价体系列为

若干因素之一，本书把它列为**决定性的那一个**，其余因素在它面前几乎不起作用。这个判断来自第五章的机制而非来自实证比较，因此它的地位是一个**待检验的强假设**——如果有证据表明某项改革在评价体系完全未变的条件下持续了三年以上并保持了原有形态，本书这一判断即被削弱。作者未见过这样的案例，但这不构成证据。

七、可观察的动作

在你们学校的全部制度文件里，找一找有没有任何一处，写明了某段时间可以没有产出。

通常一处也没有。这不是疏漏——一个不能容纳“没有产出”的制度，在结构上无法保护第五编所说的那一段。

第四编材料：一所学校的换道时间线与三场对话实录

本编材料的性质：时间线为按第二十七章推演的示例进程；三场对话为示例构造，用于给出可直接使用的说法，非任何真实场合的记录。

材料十一 一所学校的换道时间线（推演示例）

第二十七章说绝大多数尝试死在第三周。这里把一个学期铺开，逐周写出会发生什么、以及那一周应当做什么。

第0周（开始之前）。

做四件事：写明过渡期长度（一学期）；写明预计会下降的项目与幅度；写明过渡期内观察什么（读数二与读数四）；写明**退出条件**。

四份文件各一页，交校长与家长代表各一份。

这一周不做，后面全部白做。

第1周。新鲜。学生觉得有意思，教师觉得有希望。读数尚无变化。

这一周的风险：误以为已经成功。

第2周。开始不适。学生问“到底答案是什么”，教师第一次感到守不住。

这一周的动作：教研组用第六章五问互查一次，主要查是否出现伪未定终点。

第3周。危险的一周。新鲜期过去，效果未现，进度已经落后。教师面对一个非常合理的判断——“先把进度赶上来，期末再说”。

这一周的动作：拿出第0周那份文件，重读一遍退出条件。**如果尚未触发退出条件，就继续。**

第4-6周（第一次月考）。成绩下滑被数据确认。

这一周的动作：只看读数二与读数四的变化，不看正确率。若这两项已有变化（未闭合时长上升、对抗型提问出现），说明方向对了。

第7-10周。学生开始适应。表现是：出现第一批“我原来以为”的说法。

这一段的风险：教师开始加码，把三到五处扩成每节课都做。这是叠加式失败的起点。

第11-16周。进入稳定。复现题成绩回升到接近原水平；迁移题开始出现变化。

这一段的动作：做第一次回写率测量。

第17周（学期末）。按第0周的约定做第一次真正的评估。

评估只回答一个问题：读数二与读数四有没有改善？若都无改善且复现题显著下降，按约定停止。

材料十二 对话实录（示例）：与家长

家长：“听说你们班现在上课不讲答案？”

教师：“不是不讲，是推迟二十分钟讲。答案一定会讲，而且讲的时候他更容易听进去，因为他刚刚自己撞过一次。”

家长：“那时间不就浪费了？”

教师：“我反过来问您一个问题：孩子做作业卡住了，您希望他卡三分钟还是三秒钟？”

家长：“当然是三秒钟，卡着多难受。”

教师：“我以前也这么想。后来发现，真正让他变一变的，恰恰是那三分钟。三秒钟拿到答案，那道题会做了；三分钟自己撞出来，他往后看这一类题的眼光会不一样。”

家长：“那成绩呢？”

教师：“头几周会难看，这一点我现在就告诉您，不等月考之后再解释。之后不会——因为走过这一遍的孩子，回来接现成的东西更快，他知道那东西是怎么来的。反过来不行：接得再多，也不会因此学会自己走。”

家长：“那AI还能不能用？”

教师：“能，而且该用。就一条：**别一个人用**。让他先说自己的想法，再让它挑错——把‘这个怎么做’换成‘我是这么想的，你说我哪里不对’。还有一条，每次结束前问他一句：‘你原来以为什么？现在这句要改成什么？’”

要点标注。这段对话的四个动作依次是：把“不讲”改成“推迟”；用他自己的经验把判断逼出来（那个反问）；提前承认代价（“我现在就告诉您”）；只给一个可执行的动作。

材料十三 对话实录（示例）：与校长

校长：“这套东西要投入多少？”

教师：“课时不增加。三到五处，替换掉原本讲那部分内容的时间——因为那部分讲解现在孩子课后三秒钟就能拿到。”

校长：“数据怎么办？我要向上面交代。”

教师：“两件事。第一，原有的所有指标一项不动，照常报。第二，加一栏，标题写‘未闭合时长’，第一次可以是空的。它不参与评价，不设权重，只记录。”

校长：“那我怎么知道它有没有用？”

教师：“我给您三个东西：一个时间边界——一个学期；两个观察项——未闭合时长和孩子对 AI 说的第一句话，都是自动采集的，不占课时；还有一条**退出条件**——一学期后这两项都没改善，且复现题成绩显著下降，就停。”

校长：“最后一条你自己写的？”

教师：“是。不给退出条件的改革，应该被怀疑。”

要点标注。关键是最后一条。给出退出条件，是把这件事从“一次没有终点的理念实验”变成“一个有起止、有观察项、可以被终止的安排”。没有它，校长承担的是无限责任。

材料十四 对话实录（示例）：与一位不同意的老教师

老教师：“我教了三十年，我讲得清楚，孩子成绩好，家长满意。你现在说这是错的？”

教师：“不是错的。我在书里写了一句话，我念给您听：‘发现学教育不是错的，它是过去唯一负担得起的。’一个人陪一个孩子走完一个概念要十小时，三百个概念就是三千小时，一位老师一辈子只够带十个人。所以只能把走完的路磨光了发下去。这不是谁的失误，这是唯一的办法。”

老教师：“那现在为什么不行了？”

教师：“因为讲清楚这件事，现在有人做得比我们好。它不累，不烦，可以给每个孩子单独讲一遍。我们在这一项上已经不是最优选项了。”

老教师：“那我这三十年白练了？”

教师：“恰恰相反。您能把一个东西讲清楚，说明您真的吃透了它——而只有吃透了的人，才知道这门学科的哪一处规定，当年本可以是别的样子。讲得清楚的能力，正是找到那个地方的必要条件。三十年不作废，它换了一个用法。”

要点标注。这段对话的次序不能颠倒：先给公道话，再说位置让出去了，最后说新位置需要的正是他手上那份东西。三步缺一步，这段话就变成了一次否定。

第五编 新路

前四编回答了“老路为什么走完了”“损害是什么”“两条道怎么分”。第五编回答最后一个问题：具体怎么上课。本编从合格问题的形态入手，给出三方碰撞的机制、位置轮换的规则、两条不可破的约束、塌陷方向的对策、一节课的四段结构与教师禁令表、角色作为产物的理论，以及五科的落地形态。本编的全部设计受第七章那四条约束的限制：不得损害跨代存活、不得依赖小班、不得破坏校准、不得放弃可批评性。

第二十九章 岔口：合格的问题长什么样

合格的问题不是难题，是一道在此刻没有任何人手里握着答案的题；判别它只需要一条——把它输进任何一个搜索框，一分钟内能不能拿到一个大家都点头的答案。

一、从问题开始，不从活动开始

第五编的第一章不讲教法，讲问题。原因是：**几乎所有失败的尝试，都失败在这一步。**

一位教师读完前四编，通常会问“那我该怎么组织课堂”。这个问法本身就落在活动一侧。而按第六章的判据，活动的形态不决定这堂课属于哪条路，终点才决定；而终点是由那个问题带来的。

一个好问题会自己长出活动；一个坏问题配什么活动都没用。

所以本编的次序是：先问题，再机制，再流程。

二、判别标准：一分钟检验

合格的问题只有一条判别：

把它输进任何一个搜索框或大模型，如果一分钟内能拿到一个大家都会点头的答案，那它就不是岔口，是一道复习题。

这条标准的好处是它可以当场执行，不需要任何理论。而它的严格程度往往超出预期——绝大多数被认为“很有启发性”的问题，都过不了这一关。

“为什么负负得正？”——过不了，有现成解释。

“三角形内角和为什么是一百八十度？”——过不了。

“我们撕角拼直线那一下，到底靠的是什么？”——**过得了**。因为这个问题的答案不在任何教材上，它要求的是对刚才那个具体动作的反思，而那个动作是这个班刚刚做的。

第三个例子说明了一件重要的事：**合格的问题往往依附于此时此地的具体经历**。这也正是它无法被搜索的原因。

三、三种合格问题的形态

按可操作性，合格的问题有三种典型形态。

形态一：追问一个大家都默认的前提。

“你怎么知道**所有**三角形都是这样？”

“我们凭什么认为这三个例子能代表全部？”

“这个说法在什么情况下会不成立？”

这一形态的特征是：它不问结论，问**结论成立的条件**。而条件几乎从不被写进教材（第四章第五节），所以它天然不可搜索。

形态二：追问一次判定的来源。

“这条规矩是谁定的？当年能不能定成别的？”

“这个中心思想是作者说过的，还是别人替他定的？”

“如果当年的人选了另一种做法，今天这条规则会不会不一样？”

这一形态直接打在元层（第十七章第四节的性质二），因此它同时是解锁动作。第三十四章的元层解锁三问就是这一形态的定型版本。

形态三：追问一个边界。

“换成……还成不成立？”

“如果把这个条件去掉会怎样？”

“这个方法能不能用在完全不同的一类东西上？”

这一形态最容易设计，因为它可以从任何一个已学内容机械地生成——取一个条件，去掉它，看会怎样。

三种形态给出了备课时的操作路径：**不要去找“有趣的问题”，去找这三种句式能挂上去的地方。**

四、一个必要的区分：难题不是岔口

这个区分必须写清楚，因为它是最常见的误解。

难题是有答案而难以求得的题。它的答案存在于某处，只是需要更多的步骤、更巧的方法、更强的能力才能到达。

岔口是此刻没有任何人手里握着答案的地方。

两者对学生的作用完全不同。难题训练的是在已知路网上的移动能力（第十五章的高速公路），做得再多也不改变可达域；岔口要求的是产出一个判断。

而在实践中，两者极易被混淆，因为它们都让学生“卡住”。区分方法是问教师自己：**如果他现在告诉学生答案，他说得出来吗？说得出来的是难题，说不出的是岔口。**

这也是本书对教师最不客气的一处要求：**他必须敢于站在一个自己也没有答案的地方，当着四十个人。**

五、岔口从哪里来：五个矿脉

教师最实际的困难不是理解岔口是什么，而是**找不到**。这一节给出五个可以反复开采的地方。

矿脉一：教材上的“规定”。任何一处写着“我们规定”“约定”“定义为”的地方，都是当年做过一次选择的痕迹。问：能不能规定成别的？

矿脉二：那些被删掉的困难（第四章）。五个标本给出了方法：负数当年为什么被拒绝，无理数当年引起了什么，等号的含义是什么。**这些困难只要能被重新做出来，就是岔口；只能被讲出来的，不是。**

矿脉三：学生自己犯的错。尤其是那些“合理的错误”——比如认为三分之一比二分之一大。这类错误不是粗心，是他在用旧规则处理一个新对象（第四章第三节）。**把这个错误摆到台面上，让全班判断它错在哪一步，就是一个现成的岔口。**

矿脉四：跨学科的碰撞点。同一个词在两门课里含义不同（力学里的“功”与日常的“功劳”，数学里的“或”与日常的“或”），同一个现象两门课解释不同。这些地方教材从不处理，因为它们不属于任何一科。

矿脉五：此时此地刚发生的事。刚才那个实验的结果、刚才某个同学的说法、刚才全班投票的分歧。这一类最不可搜索，也最容易被浪费——因为它需要教师当场放弃原定教案。

五个矿脉里，前三个可以在备课时准备，后两个必须临场。**一位教师能稳定使用前三个，就足够了。**

六、岔口必须能被走两步

第二十二章第五节说过空位题的第二条原则，这里作为岔口的通用要求重述：

完全无从下手的问题不是岔口，是空话。

判别方法：这个问题能不能被提出一个哪怕很粗糙的猜想？如果不能，它太大了。

“什么是数学的本质”——太大，不是岔口。

“我们为什么用十进制？换成八进制会怎样？”——可以走两步。

一个实用的收缩办法是**加一个具体的对象**：把“为什么”换成“这一个为什么”。抽象的追问收缩到具体的对象上之后，通常就可以被走两步了。

七、可观察的动作

备下周的课时，用一分钟检验筛一遍你准备的所有提问。留下过关的那些，数一数有几个。

多数教师第一次做这件事的结果是零到一个。这不是备课不认真——**是我们从来没有被要求准备这一类问题。**教案模板上的“教学重点”“教学难点”“教学目标”三栏，没有一栏是放它的。

八、与相邻说法的分界

课堂提问研究中有成熟的问题分类（封闭式与开放式、低阶与高阶、记忆性与思考性），本章的岔口与“开放式高阶问题”相邻但不重合。差别在判别标准：那些分类按**问题的形式与所需认知层次**划分，本章按**答案此刻是否已经存在**划分。一个高阶的、需要分析综合的开放式问题，如果它的答案已经写在某处，在本书的框架里仍然属于老路——它训练的是抵达的能力，而不是产出判断的能力。**这个差别在纸面上极小，在效果上是两条赛道。**

第三十章 三方碰撞：各带别人拿不出的材料

三方同处一室不构成碰撞；要撞得起来，三方必须分处不同的位置，各自带着别人拿不出的东西进场——否则那只是三个人轮流发言，再由老师综合。

一、新路的课堂上有三方

新路的课堂上有三方：**教师、学生、AI。**

但三方同处一室并不自动构成新路。如果三方只是轮流发言然后由教师综合，那仍然是发现学，只是多了一个发言人。

要撞得起来，需要两个条件：

条件一：三方分处不同的位置。

条件二：每一方带着别人拿不出的材料进场。

第二个条件比第一个更硬。没有材料的一方不是一方，是听众。

二、三个位置

三个位置就是第二章那三样东西：

显露（S）——现成的说法与结果：所有人已经怎么说的、书上怎么写的、标准答案是什么。

路径（D）——怎么走：这一次是谁在走一条自己没走过的路。

土壤（E）——已有的底盘与旧账：他做过的题、他一直以来的想法、他信以为真的东西。

需要立刻说明一件工程纪律：**这三个字母是教师侧的内部机制，课堂上对学生零术语。**学生只需要知道“现成的说法”“走路”“旧账”这三个大白话。任何把这套框架直接教给学生的做法，都会把它变成一套新的成品。

三、三方各带什么材料

学生带旧账。他已有的底盘、以前那些题、他信以为真的东西。**这是全场唯一不可替代的材料**——AI 拿不出来，教师也拿不出来，因为那是这一个孩子的。

这一条给了学生一个正面的位置。在老路上，学生是终端，是接收者，是唯一不生产任何东西的一方。而在这里，他手里握着全场唯一的独家材料。

AI 带异议。三个互相打架的答案、一个反例、一句“你这个在什么情况下不成立”。

注意这里的措辞：**不是“带答案”，是“带异议”。**答案它当然有，但那不是它被请进来的理由。第三十二章将说明为什么这个区分是硬约束。

教师带一道没有现成答案的题。就是第二十九章的岔口。这是他唯一必须带的东西，也是最难准备的一样。

三样缺一，三方就退化成两方加一个观众——**而被剩下的那个，一定是学生。**第三十三章将说明为什么塌陷方向是固定的。

四、为什么必须是三方，不是两方

一个自然的疑问：教师和学生两方不行吗？两千年来不都是两方吗？

按第三章的账，两方在班级授课的条件下无法产生碰撞——教师同时握着显露与路径，学生只有旧账，而旧账在教师的权威面前几乎没有分量。这就是为什么苏格拉底式的诘问从未被制度化（第一章第六节）：**它在一对一时成立，在一对四十时不成立。**

第三方的作用不是“多一个帮手”，而是**把显露从教师手里拿走。**

这是本章最不显然、也最重要的一句。当所有现成的说法都由 AI 一次端齐时，教师就从“答案的持有者”这个位置上被解放出来了。他不再需要维持“我知道而你们不知道”这个不对称，因为那个不对称已经不存在——孩子随时可以自己去问。

教师第一次可以真诚地不知道。

这一点在老路上是不可想象的。而它恰恰是第二十九章第四节那条要求（敢于站在一个自己也没有答案的地方）能够成立的结构条件。

五、碰撞与讨论的区别

必须把碰撞与“课堂讨论”区分开，否则它会被当作一个换名的旧词。

讨论是各方陈述观点，然后寻找共识或由教师收束。它的默认目标是**收敛**。

碰撞是各方用自己的材料去顶别人的材料，直到某一方的说法站不住。它的默认目标是**让某样东西倒掉**。

两者在现场的表现差别明显：讨论中，学生说完各自的想法，教师说“很好，还有吗”；碰撞中，学生说完，必须有人立刻问“那……的时候你这个还成不成立”。

讨论可以没有产出，**碰撞必须有一样东西倒掉**。这也是第三十一章第五节判据的来处。

六、涌现的判据

怎么算真的撞出了东西？判据只有一条：

| **下课时黑板上留下的那句话，不是三方中任何一方带进来的。**

教师带进来的是一道题，AI带进来的是现成说法，学生带进来的是旧账。如果结论可以追溯到其中任何一方，那节课是**汇报**，不是涌现。

撞出来的东西，应当是三方都不会单独同意、但三方的材料合起来只能得出的那一条。

这条判据与第十五章第四节的“新路线不是新道路”是同一条：**每一段都能追溯到已有路面的，是拼接。**

七、一个完整实例

以第四章的第四个标本为例，二十分钟。

教师带的题：“你怎么知道所有三角形都是一百八十度？”——追问一个默认前提（形态一）。

学生带的旧账：他撕过三个角，拼成了直线。这是他手里唯一的证据，也是全场唯一的一手材料。

AI带的异议：球面上的三角形，内角和大于一百八十度。

碰撞过程：学生说“我撕了三个都对，所以都对”。这是他的旧账，也是他的推理。AI给出球面的反例。此时学生的旧账与这个反例正面冲突——他撕的那三个确实都对，而这个反例也确实成立。

倒掉的是什么：“撕三个就能推出所有”这条推理倒了。

撞出来的是什么：教师此时只问一句——“那我们撕角拼直线那一下，到底靠的是什么？”学生会挖出来：靠的是纸是平的，靠的是那三个角挪过去不变形。

黑板上留下的那句话是：**一百八十度不是三角形的性质，是平面的性质。**

教师带来的只有一道题，AI 带来的只有一个球面，学生带来的只有撕过的三个角——这句话三方谁都没带来。

八、可观察的动作

回想你最近一节有小组讨论的课：那节课上，有没有任何一样东西倒掉了？

如果没有，那是讨论，不是碰撞。这不是批评——讨论有讨论的价值。但按第五节的区分，它属于老路上的接受环节。

九、与相邻说法的分界

合作学习与同伴教学有大量成熟研究，它们同样强调多方参与与观点交锋。差别在于第三方的性质：在那些研究中，第三方是**同伴**，其材料与学生本人同质（也是学生的旧账，只是不同的旧账）；本章的第三方带的是**异质材料**——一个学生群体内部不可能自发产生的反例。这个差别决定了两方能撞出的东西不同：同伴之间的交锋可以暴露个体差异，但很难推翻一个全班共有的前提，因为那个前提对所有人同样自明。**要推翻一个共有的前提，必须有一个来自外部的材料源。**这正是 AI 在新路上唯一不可替代的作用。

第三十一章 位置轮换：由材料决定，不由身份决定

位置不看身份，看手里有什么；同一堂课里教师可能占土壤而学生占路径，而那一刻真正在“当教师”的是那个走出路来的学生。

一、固定分派是错的

第三十章给出了三个位置和三方的默认分派：学生占土壤，AI 占显露，教师占路径。

这个分派讲一节课够用，讲一门课就错了。

原因是：分派应当由**材料**决定，而材料随主题变化。一个孩子在某个主题上有丰富的旧账，在另一个主题上一无所有；一位教师在某个内容上能走出别人没走过的路，在另一个内容上他自己也是新手。

固定分派会在两三周内退化成三个固定身份，而一旦固定，学生就永远只出材料、不出判断——这与老路上他作为终端的位置没有实质差别。

二、三个问题定分派

分派位置不靠头衔，靠三个问题。任何一个主题，开课前问一遍，分派就定了：

这一次的旧账在谁手上？（- 土壤）

这一次的现成说法在谁手上？（- 显露）

这一次谁能走出一条别人没走过的路？（- 路径）

三问的次序有讲究：**先问旧账**。因为旧账是最稀缺、也最不可替代的材料，它在谁手上，其余两个位置就基本被决定了。

三、四种典型分派

发生主题	显露（现成说法）	路径（走路）	土壤（旧账）
老概念重新发生	AI 端出全部现成说法	教师守岔口	学生的旧账
学生毫无底料的新领域	学生先端出他现在的说法	教师	AI 供异域材料
作文与创作	教师端范文	学生自己走	AI 供大量别人的走法
教师自己也不会的题	AI	学生走	教师三十年的课堂旧账

第二行值得说明：当学生对一个领域一无所知时，他手上没有旧账，硬要他出旧账只会得到编造。此时可行的做法是让他先说出“我现在觉得是怎样”——那个粗糙的、未经检验的说法，就是他此刻能提供的全部材料，而它已经足够被顶。

第三行是作文课的关键。作文课上最不该发生的事是 AI 占显露——它给范文，学生照抄结构，写作当场死亡。可行的配置是让它占土壤：一次给出二十种别人的走法作为材料，路径留给学生自己走。

第四行值得单独强调：

教师退出路径、去占土壤，课照样发生，而且往往发生得更好。

那一刻真正在“当教师”的，是那个走出路来的学生。这一行是第三十五章最硬的证据。

四、教师占土壤是什么样子

第四行在实践中最难做，也最容易被误解为“教师不备课”，所以需要具体说明。

教师占土壤，不是他什么都不做。他提供的是三十年积累的旧账：

“这个说法我以前听一个学生也说过，后来他发现……”

“我当年学到这里的时候，卡在一个地方，我说给你们听听。”

“我们班去年做这道题的时候，投票结果是三比一。”

“我教了二十年，从来没想到这个问题——你再说一遍。”

最后一句是这个位置上最有力的一句话，而且它必须是真的。

教师占土壤的标志是：他提供材料，但不提供方向。他说的每一句都是“我这里有一份东西”，而不是“你应该往那边想”。

五、轮换的频率

多久轮一次？

本书给出的量是：**一个单元至少换一次位。**

这个量的根据不是理论，是一个实践判断：低于这个频率，三个固定身份会重新长回来；高于这个频率，教师的准备负担过重且学生无所适从。

而更重要的是换过位的那一节课的作用：

那是学生第一次尝到“我在走路”的一节课。

一个学生如果整个学期都在出旧账，他学到的是“我的经验是有用的材料”；而只有当他占过一次路径，他才知道“我可以是那个走出去的人”。这两件事的差别，与第二十二章的可达域信念直接相关。

六、一个常见的失败：轮换变成走过场

轮换最常见的失败形态是形式化：教师宣布“今天由某某同学来主导”，然后这位同学按教师事先给好的提纲走一遍。

这不是轮换，这是把执行权交出去而把判断权留下。判别方法很简单：

那位学生走的那条路，教师事先知道吗？

知道，就不是轮换。轮换的实质是路径的产出权转移，而不是发言顺序的调整。

这也解释了为什么第四行（教师自己也不会的题）是最真实的轮换：因为教师无法事先知道那条路。

七、可观察的动作

这个学期里，你有没有在一节课上，公开地把“走路”这个位置让给一个学生，而且你自己确实不知道他会走到哪里？

如果有，那是哪一次？如果没有，最接近的一次是哪一次？

第三十五章将说明，这一问的答案是判断一位教师是否真正完成换道的最便捷标志。

八、与相邻说法的分界

课堂角色研究中有关于“教师中心”与“学生中心”的长期讨论，本章的位置轮换与之不同层。那一讨论处理的是主导权的分配比例——谁说得更多、谁决定活动、谁掌握节奏；而本章处理的是材料与功能的分派，其中主导权并不是一个位置。一节完全由教师主导节奏的课，如果路径的产出权在学生手上，在本书的框架里就是轮换成功的；反过来，一节学生说了大部分话的课，如果那条路是教师事先铺好的，它仍然是教师中心的。这个差别说明“学生中心”这个词在本书里不构成一个目标——它既不充分，也不必要。

第三十二章 两条硬约束：闭合权不轮，不得占两格

位置可以轮换，但“什么时候允许给答案”这一格永远在人手上；而任何一方同时占据显露与路径，碰撞当场塌成投喂。

一、为什么需要硬约束

第三十一章说位置由材料决定、随主题轮换。这个规则很灵活，而灵活的规则需要边界，否则它会退化成“怎么样都行”。

本章给出两条边界。它们是硬约束，意思是：**违反其中任何一条，这节课当场退回老路，无论其余部分做得多好。**

二、约束一：闭合权不参与轮换

内容。什么时候允许给答案、什么时候宣布这个岔口可以合上——这个决定永远在人手上，而且必须被明说。

为什么。它的构造决定了它一遇到缺口就填。这不是缺陷，是它被造出来要做的事（第十二章第五节）。让一台以填缝为天职的机器决定“什么时候该填”，等于没有决定。

一个必须做的区分。AI 可以占据路径的内容，不能占据路径的权。

占内容：它提议一种走法，说“也可以从这个角度试试”。这是合法的，而且有用。

占权：它决定这个问题现在可以被解决了，于是给出完整解答。这是越界。

在实践中，这个区分落实为一条可执行的规矩：**它可以提议，不可以收束。**

在课堂上的具体形态。闭合权表现为两次公开宣布：

一、“从现在开始，谁都不许给答案，包括我。”

二、“好，现在可以说了。”

两次宣布之间就是那段未闭合时长。**没有宣布的规矩不是规矩**——只在教师心里执行的克制，学生感受不到，因而不构成结构。

为什么必须包含教师自己。这一句当着学生说出来，教室里的权力结构发生一次实质变化：这道题第一次不再是“教师手里已有答案的一次考查”，而变成一件在场所有人都还没有答案的事。

这句话的效果远大于它的字面内容，而且它是免费的。

三、约束二：任一方不得同时占据两格

内容。尤其是显露与路径不得落在同一方。

为什么。同时给出答案又给出路子，就是投喂。学生此时既不需要出材料，也不需要产出判断，他只需要跟着走——这与老路上的接受没有任何差别，只是形式更精致。

谁最容易越界。

AI 最容易，因为它两样都现成，而且是同时给出的。你问它一个问题，它的默认回答形态就是“答案+推导过程”。这意味着**默认状态下它一定占两格**，必须靠外部规定把它限制在一格。

教师其次，因为他既知道答案又知道该怎么讲。教师占两格的典型形态是：“我们来看这道题，第一步应该想到什么？”——这句话看似在提问，实则同时给出了答案的存在与通往它的路径。

学生也可能占两格，虽然少见：一个已经在课外学过这部分内容的孩子，会同时端出结论与做法，把全班的岔口一次性关掉。这一情形需要教师预先处理——通常的做法是提前跟他约定，请他先不说，或者请他去顶别人而不是给答案。

四、两条约束的执行方式

两条约束都不能靠自觉执行，理由与第五章相同：凡依赖个人修养的东西在这套系统里活不下来。

执行约束一的装置：一个可见的计时器。

讲台上放一个走字的计时器。它同时约束学生和教师，而且不需要任何人扮演执法者。

这个细节看似琐碎，但在实际推行中它与整套理论具有同等重要性：

本能只能被具体装置改变，不能被道理改变。

一位教师三十年养成的“学生一卡住就补一句”的本能，不会因为读了这本书而消失。它只会因为讲台上有一个所有人都看得见的计时器而暂停。

执行约束二的装置：一句写死的提问句式。

规定学生对 AI 只能说两种话：

“我是这么想的，你说我哪里不对。”

“要是换成……你还成不成立？”

不许说“这个怎么做”。

这条规矩把 AI 从显露+路径两格，强制压缩到异议这一格上。它可以被写在教室墙上，也可以由学生互相监督——而且它比任何技术手段都可靠，因为它不依赖产品端做任何事。

五、一个必须回答的问题：那答案什么时候给

两条约束容易被误读为“不给答案”。必须澄清：**答案要给，而且必须给。**

第七章的三大成就（跨代存活、规模化、校准）全部建立在成品被接住这件事上。一堂课如果结束时全班仍然不知道正确的结论，那不是发生学教学，那是失职。

区别在次序：

岔口张开－顶－反例－回写－然后才是答案。

答案在最后给，而且此时给的成本极低——因为它现在有地方可挂（第十八章的第一种死亡不再发生），而且孩子已经知道它为什么长这个样子。

一个便捷的说法：**答案不是被取消了，是被推迟了二十分钟。**而这二十分钟就是全部差别。

六、违反约束的两种典型课堂

第一种：闭合权交出去了。课堂上大量使用 AI，学生随时提问随时得到解答，教师在旁边协助。这种课的表面特征是流畅、高效、学生活跃，而按第十九章的检查表，它五项摩擦中只剩课时一项。

它的诊断读数是第三十八章的读数四：学生说的全是“这个怎么做”。

第二种：一方占了两格。教师抛出问题之后，紧接着给出思路提示——“大家想想，是不是可以从……入手”。这句话给出的时刻通常在提问后的第四十秒左右。

它的诊断读数是读数二：未闭合时长在一分钟以内。

两种课都可能被评为“优秀公开课”，因为它们在老路的全部标准上都很好。

七、可观察的读数

在你的下一节课上，请一位同事帮你记两个数：一是从你提出问题到有人给出答案的秒数；二是学生对 AI 说的第一句话属于哪一类。

不要提前告诉自己“我今天要忍住”——那样记下来的数没有诊断价值。要记的是你平常的样子。

八、与相邻说法的分界

课堂管理研究中有关于“等待时间”的经典结论：教师在提问后延长等待时间，会显著改变学生回答的长度与质量。本章的约束一与之直接相关，且本书完全接受那些结论。差别在于两处：一是那一族研究的等待时间以秒计（通常讨论的是从一秒延长到三到五秒），本章讨论的是以分钟计的一整段；二是那一族把等待时间作为**教师的一项技能**来训练，而本章把它作为一条公开宣布的、同时约束教师的规矩来执行——因为按第五章的机制，技能会被筛选，规矩不会。

第三十三章 塌陷的方向是固定的

三方的力量并不对等，因此碰撞一旦塌陷，塌陷方向是可预测的：学生先退场，然后是教师，最后剩下它讲、大家记。

一、为什么要预先写出塌陷方向

一套机制如果只在理想状态下成立，它在现场没有用。本章处理的是这套机制失效时会往哪边倒。

预先知道倒的方向，有两个好处：一是可以在流程里排一个专门顶住它的动作；二是可以据此写出这套方案被筛选之后会剩下什么（第五章第四节已经预告过要做这件事）。

二、三方的力量不对等

AI 的材料太多。它可以在几秒内给出比全班一学期能想到的更多的说法。在纯粹的材料量上，它对另外两方是碾压性的。

教师有权威。他决定谁发言、什么时候结束、这节课算不算成功。即使他真心不想使用这份权威，它也在那里。

学生只有服从。他手里唯一的独家材料是旧账，而旧账在前两者面前显得琐碎、不成体系、甚至可笑。

三方的这个力量结构决定了：

| 塌陷方向是固定的——学生先退场，然后是教师，最后剩下 AI 讲、大家记。

三、三个阶段的具体形态

第一阶段：学生退场。

表现是他不再提出自己的想法，只提问；从“我觉得是这样”变成“这个怎么做”。

原因不是他懒。是他发现自己那份旧账，在一个可以立刻给出十种精确说法的东西面前，实在拿不出手。第十五章第九节那句话在这里出现了它的课堂形态：**农庄不是被拆掉的，是被说服的。**

第二阶段：教师退场。

表现是他从守岔口变成组织流程：安排谁先问、提醒时间、总结要点。他仍然很忙，但他不再做判断。

原因是：当学生不再产出想法时，就没有东西需要被守护了。**岔口没有张开，守岔口这个动作失去对象。**

第三阶段：只剩讲与记。

这一阶段与老路的唯一区别是讲的那一方换了。而按第二编，它讲得更好。所以这一阶段在读数上会显得比原来更成功——加固悖论在课堂层面的完整形态。

四、顶住它的那一个动作

塌陷从第一阶段开始，所以对策必须打在第一阶段。而且它必须是一个硬性排进流程的动作，不能是一条原则。

| 让学生用他自己的旧账，去顶 AI 一次，顶不动为止。

这是三方碰撞里唯一不能省的环节。省了它，学生就从“一方”退化为听众，三方碰撞变成两方加一个观众。

“顶不动为止”这四个字必须留。如果只是“让他说一句自己的想法”，那么这个动作会在两周内退化成一个形式——他说一句，AI 回一段，结束。要求是他必须坚持到自己那个说法真的站不住为止，而这通常需要三到五个回合。

时间上，这个动作应当占整节课的最大一块。第三十四章的四段结构里，它占十五分钟，是四段中最长的一段。

五、为什么这个动作最难排进去

必须诚实指出：这个动作在实践中最容易被省略，而且省略它的理由每一条都成立。

它慢。三到五个回合，十五分钟，而这十五分钟里没有任何“内容”被覆盖。

它乱。四十个学生同时在顶，教师无法掌控全场发生了什么。

它的产出不可见。结束时黑板上可能什么都没有。

它让水平低的学生难受。旧账少的孩子顶不出几句。

四条都是真的。本书对它们的回应只有一条：**这十五分钟是这节课上唯一无法被课外替代的部分**。讲解可以课外看，练习可以课外做，答案可以课外查——只有“用自己的旧账去顶一个比自己强得多的对手”这件事，必须有一个被规定的场合，否则不会发生。

至于第四条，处理方式在第三十一章第三节第二行：旧账少的孩子先说“我现在觉得是怎样”，那个粗糙的说法就是他的材料。

六、这套方案被筛选之后会剩下什么

第五章第四节承诺过要做这件事，现在兑现。按第五章的机制，本编的方案在一年之后会被筛选。可以预先写出结果：

会留下来的：课堂上加一个提问环节；让 AI 出几个反例活跃气氛；每节课最后让学生说说收获；用一个计时器（因为它好看且易于展示）。

会被放弃的：十分钟内谁都不许给答案；接受这节课没有得出结论；让学生用旧账去顶到顶不动为止；教师公开把路径让给学生。

留下来的那些无害而无用；被放弃的那些正是全部作用之所在。

把这个预测写在这里有两个用途：一是它可以被验证——一年之后拿出来对照；二是它给出了一份检查清单，一所学校可以据此判断自己是否正在被筛选。

如果一年之后剩下的正是上面那四条，那么这次尝试已经结束了，只是还没有人宣布。

七、可观察的读数

——一节课里，学生说出“我觉得”“我原来以为”“我不同意”这三种句式的总次数。

这个读数比第三十八章那五个更粗糙，但它直接测量塌陷的第一阶段。连续记录一个月，如果这个数在下降，塌陷正在发生。

八、与相邻说法的分界

小组讨论研究中有关于“搭便车”与“话语权垄断”的大量工作：在小组中，少数人会主导讨论，其余人退出。本章的塌陷与之形态相似，但机制不同：那一族的退出源于**社会性因素**（性格、地位、责任分散），因而对策是结构性的角色分配与个体问责；本章的退出源于**材料的不对等**——学生退场不是因为不想说，是因为他手里的东西在对比中显得不值一提。所以本章的对策不是分配发言权，而是**先确立旧账作为独家材料的地位**（第三十章第三节），再要求他用它去顶。给一个自认为手里没有货的人发言权，他仍然不会说话。

第三十四章 一节课的四段结构与教师禁令表

一节四十分钟的课分四段：开缺口八分钟、顶十五分钟、反例十分钟、回写七分钟；四段中第一段最难设计，第二段最难忍住，第四段最容易被省略——而省掉第四段，前三段全部作废。

一、把机制压成时间表

前五章给出的是机制。本章把它压成一个任何学科都能照着排的时间结构。

需要先声明这个结构的地位：**它是一个起步用的脚手架，不是一个必须遵守的模板。**熟练之后四段的界限会模糊，时间会随内容浮动。给出具体分钟数的唯一理由是：没有分钟数的建议不会被执行。

二、第一段开缺口（约8分钟）

做什么。抛出一个岔口（第二十九章），并当众立规矩。

规矩的原话。“从现在开始到我说停，谁都不许给答案——包括我，包括 AI，包括课本后面那一页。”

这一段的常见失败：问题不合格。八分钟里学生很快答出了教师想要的东西，于是提前进入下一段——而其实没有任何缝隙被张开。

判别方法：这一段结束时，教室里应当有一个明显的不适感。**如果气氛是顺畅的，那么这一段失败了。**

三、第二段·顶（约15分钟）

做什么。学生把自己现在的想法和理由摆出来，去顶 AI，或者顶同学，或者顶教师。规定只能说两种话（第三十二章第四节）。

这是全课的重心，也是唯一不可省的一段（第三十三章第四节）。

教师这一段的全部工作是不作答，以及在有人要提前合上时把它重新撑开。

具体的教师动作只有四个：

“你这个想法哪里让你自己不放心？”

“还有谁的想法跟他不一样？”

“你再顶一次，它这次说的你服不服？”

（沉默）

第四个是最重要的一个，也是最难做的一个。

这一段的常见失败：教师在第四十秒开口。这不是态度问题，是三十年职业本能。对策是计时器，不是自觉。

四、第三段·反例（约10分钟）

做什么。放进一个使原来的说法不成立的材料——由 AI 提供，或教师准备，或某个学生带来。

关键动作：放进去之后不解释。只问一句：“那我们刚才那一下，到底靠的是什么？”

为什么不解释。因为解释就是把结论端出去。被挖出来的那个东西——刚才那个动作依赖的那个未说出的前提——必须由学生自己说出口，否则它只是又一个成品。

这一段的常见失败：教师忍不住把反例的意义讲清楚。表现是学生点头，气氛良好，而没有任何人自己说出那个前提。

五、第四段·回写（约7分钟）

做什么。不做总结，不做小结，不让学生复述。只做一个动作：

让他去改一件旧的。

改笔记上的一句话，改上周做过的一道题，改自己一直以来的一个说法。**改完之后念出来。**

“念出来”这三个字不能省。它承担两项功能：一是公开性（第十九章第三节的装置四），二是可批评性（第七章第七节那条硬约束）——一个人走出的路径必须以某种形式被交出去接受检查，否则会退回“每个人有自己的道理”。

这一段的常见失败：省略。下课铃响了，回写没做，改成“我们来小结一下”。

省掉第四段，前三段的全部工作在这一刻作废，因为发生没有落点，学生带走的仍然是一个成品。

六、四段的难度排序

第一段最难设计——因为岔口难找（第二十九章第五节的五个矿脉是为此而设）。

第二段最难忍住——三十年本能。

第三段最难克制——反例给出之后不解释，比第二段的沉默更难，因为此时教师手里有一句非常想说的话。

第四段最容易被省略——因为它在时间上排在最后，而课总是拖堂。

一条实用建议：**倒着排时间。**先从下课时间往前留出七分钟，剩下的再分给前三段。否则第四段一定会被吃掉。

七、元层解锁三问

第十七章第四节说过，通道已封的孩子需要先做元层解锁。这个动作的最小形态是三个问题，可以插在第一段或第三段：

问一：“如果不这样会怎么样？”

问二：“当年是不是差一点就选了别的？”

问三：“你能不能造一个？”

三问的共同点是它们都不问“是什么”，都问“能不能不是这样”（第二十二章第五节）。它们直接打在“对象是现成的”那条元规则上。

第三问最强，因为它把学生从判断的接收者变成了规则的制定者——造一个新的运算、造一个新的记号、造一条新的规矩，然后检查它成不成立。

一道必须设的挡板：三问之后必须收一步，明确说出“可以不是这样”不等于“怎么样都行”。造出来的东西要接受检验，站不住的就是站不住的。**没有这一步，元层解锁会滑向相对主义，而那与本书的主张恰好相反。**

八、教师禁令表

下表可直接复印张贴。左栏是老路上被视为美德的动作，在新路上恰恰需要克制；右栏是替代动作。

不要说 / 不要做	改成
“我来给大家理一理”	“谁来说说你现在是怎么想的”
“其实很简单，你看”	“你这个想法哪里让你自己不放心”
学生一卡住就补一句提示	数到三十秒再决定要不要开口
“对，就是这样”（当场确认）	“还有谁的想法跟他不一样”
把 AI 的回答念给全班听	让学生自己去顶，只念出分歧处
下课前把结论工整地板书一遍	下课前让学生改一件旧的
“这个我们下学期会学到”	“那你现在能不能先猜一个”
把反例的意义讲清楚	只问“那我们刚才那一下靠的是什么”

这张表在教师培训中的作用大于任何理论讲解。原因是：**老路的三个动作已经内化为职业本能，而本能不会被道理改变，只会被具体的替代动作改变。**

九、四段结构与课时的关系

必须回答一个现实问题：这四十分钟里覆盖的内容，远少于常规课堂。
第八章第六节已经给出答案，这里重述：**这二十三分钟（第一、二段）原本用于讲解答案，而答案现在是免费的。**学生课后三秒钟就能拿到讲解。

所以问题不是“腾不腾得出时间”，而是**这段课堂时间应当用来做那件免费的事，还是那件只有此时此地才能做的事。**

课时没有增加，用途变了。而第二十三章第五节的量在这里再次生效：**一学期三到五处，不是每节课。**其余课时照常，该讲讲，该练练。

十、可观察的动作

拿一节你已经很熟的课，只改一处：**把最后七分钟的小结，换成“改一件旧的，然后念出来”。其余全部不动。**

这是本书全部建议中最小的一个动作，不需要改教案，不需要新问题，不需要 AI。而它足以让一位教师第一次看见回写这件事有没有发生。

如果只做一件事，做这一件。

十一、与相邻说法的分界

教学设计中有大量成熟的课堂结构模型（导入—新授—巩固—作业；探究五环节；翻转课堂的课前课中安排）。本章的四段结构在形式上与它们同类，差别在于每一段的**目标函数**：那些模型的每一段都服务于“让学生尽可能好地抵达既定终点”，因而它们的优化方向是减少损耗；本章的第一、二段的目标函数是**维持一段未闭合状态**，其优化方向与减少损耗相反。这个差别使得两类模型不能混用——把本章的第二段插进一个以抵达为终点的课里，它会在实施中被自动压缩到两分钟，因为在那个目标函数下它确实是纯粹的浪费。

第三十五章 教师、学生、教室是产物，不是座位

谁守住了岔口，谁在那一刻发生为教师；谁走出了路，谁在那一刻发生为学生；能让这场轮转发生的那个场，才是教室。

一、这一章的位置

本章是全书对“教师还剩下什么”这个问题的正面回答，也是第九章第四节留下的那个承诺的兑现。

前面各章已经把否定的部分说完了：损耗最小的传输者这个位置已经让出去，而且抢不回来。本章要说的是那个空着的位置。

二、三个角色都是产物

第三十一章的位置轮换，推出一个关于课堂角色的判断，本书认为它是全部理论中最反直觉、也最有实践后果的一条：

教师、学生、教室不是三个人和一间屋子，而是三个位置在三方身上轮转时留下的产物。

具体地说：

谁守住了岔口，谁在那一刻发生为教师。

可能是站在讲台上的那个人。也可能是 AI 出的那一个反例——它没有给答案，反而把一个已经要合上的问题重新撑开了，那一刻它在履行教师的功能。也可能是某个学生的一句“那我上周那道题呢”。

谁走出了路，谁在那一刻发生为学生。

注意这是一个**正面**的定义，不是弱者的位置。真正在学习的那个人，是正在走一条自己没走过的路的那个人。在第三十一章表格的最后一行里，那个人可能是全班唯一没有被封写的孩子，而讲台上的人这一刻是土壤的提供者。

能让这场轮转发生的那个场，才是教室。

教室不是那间屋子，是“这十分钟谁都不许给答案”这条规矩立住的地方。规矩立不住的屋子，不是教室，只是一个共同接收成品的场所。

三、这不是修辞

上述三句容易被读成一种诗意的说法，因此必须给出它的可操作形态。

三句话的实际含义是：**这三个词从描述身份改成了描述功能，而功能是可以被观察、被记录、被检查的。**

“谁在当教师”这个问题，在老路上不需要问，答案写在工资表上。在新路上它每节课都要被问一次，而且答案可以逐分钟改变。

于是它产生一个可以采集的记录：一节课里，这三个功能各自出现过几次，分别落在谁身上。这个记录在第三十八章不作为常规读数（因为采集成本偏高），但它是教研组听课时最有价值的一份笔记——它比“这节课上得好不好”具体得多。

四、对教师职业的三层含义

这一条对教师职业的含义有三层，需要分开讲，因为它们的难度不同。

第一层：放弃“最会讲”这个自我认同。

这是最难的一层，因为它是绝大多数优秀教师三十年积累的核心资产。

但必须讲明白：**放弃的不是能力，是用途。**同样的学科理解力，用在“把它讲清楚”上已经不再稀缺；用在“知道哪里可以撬开一个岔口”上则无可替代——因为只有真正吃透一门学科的人，才知道这门学科的哪一处规定当年本可以是别的样子。

讲得清楚的能力，正是设计岔口的必要条件。

这句话应当在每一次教师培训中被说出来。三十年的积累不作废，它换了一个用法。

第二层：学会不作答。

这是一项需要专门训练的技能，而不是一种态度。它包括四项，每一项都可以被拆成具体动作来训练：

一、能忍受三十秒以上的沉默。

二、能在学生用眼神求助时不给出提示。

三、能在一节课结束时接受“没有得出结论”这个状态。

四、能在家长质疑“这节课到底学了什么”时，给出一个基于新读数的回答。

第四项常被忽略，而它其实是前三项的前提——一位无法向家长交代的教师，不可能持续做前三项。第二十六章第五节第四条（家长会加一页）正是为此而设。

第三层：接受角色轮换。

在某些主题上，教师不是走路径的那个人，而是提供土壤的那个人。这在老路上是不可想象的降格，在新路上是正常的分工。

判断一位教师是否真正完成了换道，一个便捷的标志是：

他有没有在一节课里，公开地把路径这个位置让给一个学生，并且自己去占土壤。

五、AI 天然干不了的那件事

本章必须回答一个问题：凭什么说守岔口这件事 AI 干不了？

理由不是它能力不足，而是它的构造。

第十二章第五节说过：它的全部训练都指向把不完整的東西补完整。守岔口要求的动作，与它被造出来要做的事，方向相反。

需要精确一些，因为这个论证很容易被高估：

可以做到的：通过提示或产品设计，让它在一段时间内不给答案。这在技术上不难。

难以做到的：让它判断此刻该不该给。这个判断依赖于对这个具体孩子此刻状态的读取——他是在真的在想，还是已经放弃了；他的沉默是有产出的，还是空的；再等三十秒是有价值的，还是残忍的。

这个判断正是第十章第四节说的那件事：讲解可以被复制，判断不能。

所以本书对教师不可替代性的主张，落在一个很窄但很硬的位置上：不是“机器不懂人心”这类泛泛之谈，而是“闭合权的行使需要读取一个尚未发生的具体情境”。

这个位置窄，但它是真的。而且它是一个教学位置，不是一个情感位置——第十章第七节的分界已经说明，把教师的价值建立在陪伴与人格感染上，等于承认教学本身可以被替代。

六、一个诚实的限度

必须承认这个论证的边界：它依赖于当前技术的一个事实性状况，而不是一个原理性的不可能。

如果有一天，一个系统能够可靠地读取一个孩子此刻的状态并据此决定该不该开口——那么本章的论证就失效了，而本书的这一部分需要重写。

本书不打算预测这件事会不会发生。本书主张的是：在它发生之前，这个位置空着，而且只有教师能占。

而且还有一句需要补上：即使那一天到来，“谁有权决定”与“谁有能力决定”仍然是两个问题。闭合权的归属不只是一个技术问题，也是一个关于由谁对这个孩子负责的问题。这个问题超出本书的范围。

七、可观察的动作

在你的下一节课上，记下这三个功能各出现过几次：有人守住了一个要合上的问题；有人走了一条自己没走过的路；有一条规矩真的立住了。

三项都是零，是常态，不必沮丧。第一次出现非零，通常是在第四段（第三十四章第五节）——因为那一段最容易，也最有效。

八、与相邻说法的分界

关于教师角色转变，教育文献中有一族熟悉的说法：“从知识的传授者变成学习的引导者、组织者、促进者”。本章与之的差别不在方向而在**确定性**：那一族说法给出的新角色是柔性的、难以证伪的——什么样的行为算“引导”，什么样的算“促进”，没有判别标准，因而任何一位教师都可以宣称自己已经完成了转变。本章给出的新角色只有一个动作（守住闭合权），而这个动作有一个可读的数（未闭合时长）。一个有读数的角色定义，才可能被真正采用；没有读数的角色定义，会在一次培训之后被完整地忘掉。

第三十六章 分学科落地（上）：数学、语文、科学

机制是学科中立的，教师需要的是自己那一科的样子；每一科的写法统一为四项——光滑化发生在哪里、岔口从哪里开、AI 配在哪一格、回写长什么样。

一、为什么必须分科写

前七章的机制是学科中立的。而一位教师读完之后，第一个真实的困难是：**在我这一科上，这具体是什么样子？**

不解决这个困难，全书的作用会停在“很有道理”这一层。

本章与下一章按同一个格式处理五科。格式统一为四项：

- 一、这一科的光滑化发生在哪里。
- 二、岔口从哪里开（对应第二十九章的五个矿脉）。
- 三、AI 配在哪一格（对应第三十一章的位置分派）。
- 四、回写长什么样（对应第十六章的判据）。

二、数学

光滑化在哪里。

数学教材是全部学科中被磨得最光滑的。定义在前、性质在后、例题居中、习题殿后——这是逻辑顺序，不是发生顺序。

历史上每一个概念都是从一个具体困难里长出来的：负数从欠债与方程无解，分数从分不尽，无理数从对角线量不尽，极限从无穷小的争论。教材把这些困难全部删掉，只留下定义（第四章的五个标本）。

岔口从哪里开。

最可靠的开口方式是**追问一个规定**：这条规矩是谁定的，当年能不能定成别的？

- 负负为什么得正？能不能规定得负？
- 零为什么不能作除数？是不能，还是不许？
- 分数除法为什么要颠倒相乘？
- 等号两边到底是“算出来”还是“同一个东西”？

— 我们为什么用十进制？换成八进制会怎样？

数学是五科中岔口最容易找的一科，因为它的每一处“规定”都是一次历史选择的痕迹，而这些痕迹在教材上表现为一个个不加解释的定义。**教材越是不解释，岔口越是明显。**

AI 配在哪一格。

通常配显露：让它一次端出所有现成说法，**包括那些互相不一致的说法**，让学生去判断哪个才站得住。

也可配土壤：让它给出这条规则在其他数系、其他记数法、其他时代里的不同样子。这一配置在“追问规定”型的岔口上特别有效——它提供的是“本来可以不是这样”的实物证据。

回写长什么样。

在原有的表述上加**限定词**。

“三角形内角和是一百八十度”加上“在平面上”。

“两点之间线段最短”加上“在平直空间里”。

“乘法使数变大”改成“在大于一的数上”。

加限定词是数学里最典型、最容易观察、也最容易记录的回写动作。一位数学教师如果只记录一件事，就记录这个：这学期有多少个孩子，在多少处，自己加上了限定词。

三、语文与作文

光滑化在哪里。

在“中心思想”和“标准答案”上。

一篇文章被写出来时，作者面对的是一堆没有定形的东西；教参把它压成一句话，学生的任务变成猜中那句话。

作文教学的光滑化更彻底——**范文是成品，而成品看不见被删掉的那些句子。**一个学生看一百篇范文，看到的是一百个结果，看不到任何一次犹豫、删除、推翻重来。

岔口从哪里开。

问定判的来源与可替换性：

— 这个中心思想是谁定的？作者说过吗？

— 如果这一段删掉，中心思想会不会变？

— 同一件事，如果他从另一个人的角度写，还是这个中心吗？

作文课上最有效的一个岔口是：

把学生自己上一篇作文里被删掉的那句话找回来，问他当初为什么删。

这个岔口的特别之处在于，它的材料完全是他自己的（土壤），而且它同时是第四段的回写动作——他在回答“当初为什么删”的过程中，往往会推翻当初的理由。

AI 配在哪一格。

作文课上 **AI 最不该配显露**：它给范文，学生照抄结构，写作当场死亡。这是本书对 AI 配置最强的一条禁令。

应配土壤：让它一次给出二十种别人的走法作为材料；路径留给学生自己走。

这正是第三十一章表格第三行的具体化，而它在实践中最容易被配反，因为“让 AI 给个范文”是最省事的用法。

回写长什么样。

回头改自己以前那一篇——不是改错别字，是改那个当初为了凑主题而写下的句子。

四、科学

光滑化在哪里。

在“实验验证结论”这个次序上。

教材的顺序几乎总是：先给结论，再做一个必定成功的实验去印证它。而科学史上的次序是反的：先有一个说不通的现象，然后是一堆互相打架的解释，然后才有判决性的观察。

这个倒置比数学更隐蔽，因为它披着“动手做”的外衣。一个必定成功的实验，在结构上与一段讲解没有区别——它的结论在开始之前已经写在书上。

岔口从哪里开。

问当年的选择：

- 当时的人知道结局吗？
- 有没有一个和这个相反的说法，当年也很有道理？
- 如果这个实验做出来是另一个结果，我们今天会怎么说？

AI 配在哪一格。

配显露或土壤都可以。而最有价值的用法是让它复原当年被淘汰的那个理论，并且尽力把它讲得有道理，让学生去驳。

一个能把燃素说讲得让小学生一时驳不倒的对手，比十遍氧化学说的讲解更有用。

这个用法把 AI 放在了一个它极其擅长、而任何教师都难以做到的位置上：扮演一个已经被历史淘汰但当年确实合理的立场。

回写长什么样。

学生开始区分“这是观察到的”和“这是我们规定的”。

这个区分一旦建立，会向全部科学内容扩散，而且会溢出到其他学科。它是五科中回写效果最广的一种。

五、三科的共同点

把三科并排看，一个共同结构浮现：

学科	光滑化的藏身处	岔口的通用句式	回写的可见形态
数学	教材上的“我们规定”	能不能规定成别的	加限定词
语文	教参上的中心思想	这是谁定的	回头改自己上一篇

学科	光滑化的藏身处	岔口的通用句式	回写的可见形态
科学	“实验验证结论”的次序	当年知道结局吗	区分观察到的与规定的

三行的中间一列合起来是同一句话的三种说法：这件事本来可以不是这样吗？

这也印证了第二十九章第三节的形态二（追问一次判定的来源）是最通用的一种岔口形态——它在三科上都成立，而且都是最容易开采的那个矿脉。

六、可观察的动作

在你这一科的下一个单元里，找出所有写着“我们规定”“定义为”“叫做”的地方，圈出来。

数一数有几处。每一处都是一次历史选择的痕迹，也是一个现成的岔口。

通常一个单元里有三到八处，而一个学期只需要用掉三到五处（第二十三章第五节）。所以材料是过剩的，问题从来不是找不到，是从来没有人往这个方向找过。

七、与相邻说法的分界

各科教学法中都有关于“从生活情境引入”的成熟主张，与本章的岔口在形式上相近。差别在于引入之后的走向：情境引入的目的是为既定内容提供一个可感知的入口，情境在完成引入功能之后即被弃置，课堂回到内容主线；本章的岔口不是入口，它本身就是这节课要处理的东西，而且它的走向在开始之前不确定。一个用完就丢的情境，与一个必须被走完的岔口，在教案上看起来几乎一样——差别在于第二段那十五分钟是否真的存在。

第三十七章 分学科落地（下）：语言、术科与综合实践

语言学科是 AI 最强的一科，因而最容易被它占两格；术科反而保留了最多的发生结构，因为技能不能被搬运；而项目式学习最需要警惕的是判别一。

一、语言（英语等）

光滑化在哪里。

在语法规则上。

规则被当作先在的法律发下去，例外被当作需要背诵的意外。而真实的语言里，规则是从大量使用中沉淀出来的事后描述——先有人这么说，很久以后才有人总结出规则。教材把这个次序倒了过来。

这个倒置与数学是同一种（逻辑顺序僭越发生顺序，第二章第七节的副产品二），但它的后果更严重：因为语言的规则有大量例外，而在“规则先在”的框架下，例外只能被当作需要死记的意外，无法被理解。

岔口从哪里开。

问规则与用法的先后：

- 是先有规则，还是先有人这么说？
- 这个“错”的说法，为什么这么多人这么说？
- 如果全世界的人都这么说十年，它还算错吗？
- 我们自己的母语里，有哪些说法是“错”的但大家都在用？

最后一问特别有效，因为它把学生的母语经验（土壤）调进来，而这份材料是他独有的。

AI 配在哪一格。

语言是 AI 最强的一科，**因此越要小心它占两格**（第三十二章第三节）。它既有全部规范说法，又能立刻给出如何改正的路径，默认状态下必然同时占据两格。

可行的配置是：AI 占显露（给出规范说法），学生占土壤（他自己那些说错的记录、母语的干扰、他觉得别扭的地方），教师守路径。

这里有一条对语言教师的重要提醒：

学生说错的那些句子，是全班最宝贵的材料，不是需要被消灭的垃圾。

一份长期积累的“本班错句本”，是语言课上最好的土壤来源。而现行做法通常是订正之后就丢弃。

回写长什么样。

从“我要背下这条规则”，转到“我知道这条规则在管什么、不管什么”。

具体表现是他开始给规则加边界，与数学的“加限定词”同形。

二、术科：音乐、美术、体育

一个反向的观察。

这三科在老路上被视为“技能课”，地位偏低，而它们反而**保留了最多的发生结构**。

原因很简单：**技能不能被搬运**。你无法把游泳这件事打包递送给一个人。他必须自己下水，自己呛几口，自己找到那个平衡。任何一位体育教师都非常清楚“讲会了”和“练成了”完全是两回事——而这正是第二章第四节那条“发生不可传递”的日常版本。

光滑化在哪里。

在评价标准上。什么样的画算好、什么样的演奏算对、什么样的动作算标准——这些被当作现成的规范发下去，而它们同样是历史选择的产物。

岔口从哪里开。

- 这个标准是谁定的？换一个时代还算好吗？
- 这个动作为什么规定成这样？换一种做法为什么不行？
- 如果我们班自己定一个标准，会定成什么样？

AI 配在哪一格。

配土壤最有价值：给出同一件事在不同时代、不同文化里的完全不同的做法。**这些材料对术科教师而言原本极难获取，现在近乎免费。**

这三科对主科的反哺。

本节最重要的一句不是关于这三科自己的：

术科教师身上有一份主科教师普遍缺乏的经验——他们从不相信“讲会了”等于“学会了”。

这份经验应当被明确地拿出来用。在一所学校推行本书主张时，**术科教师往往是最容易理解、也最能提供实例的一群人**，而常规做法是把他们排除在教学改革讨论之外。

三、综合实践与项目式学习

必须特别警惕的一条。

大量项目式课程的终点仍然写在教案上，只是路上安排了更多活动。第六章已经处理过这个判断，这里给出项目层面的判别方法：

这个项目做完，是不是每一组的结论都一样？

如果是，那是发现学的一条内道。

而这个检验要配一个追问（第六章第四节）：**结论相同，是必然的，还是被我引导的？**区别在于：如果某一组当真走出了一个不同的、而且站得住的结论，这个项目会不会被认为是失败的？

项目式学习的真实优势。

必须公道地指出：项目式学习在五项摩擦检查表（第十九章第三节）上表现很好——延迟长、异议多、公开性强、有明确边界。它唯一薄弱的是**克制**：项目过程中教师通常会大量介入以保证成果质量，因为成果要展示。

所以对项目式学习的改造，只需要动一处：**允许项目失败，并且明确说出失败也算完成。**

这一条很难，因为项目通常与展示、评比、家长开放日绑定。而只要成果要被展示，教师就一定会介入。

一条可执行的建议：把项目分成两类，一类是要展示的（照旧），一类是不展示的（允许失败）。**不展示的那一类才可能是发生学的。**

四、跨学科的碰撞点

第二十九章第五节的矿脉四在这里展开，因为它是唯一不属于任何单科的矿脉，也因此常年无人开采。

同一个词在两科里含义不同：

- 力学里的“功”与日常的“功劳”
- 数学里的“或”与日常的“或”（数学的“或”包含两者都成立）
- 生物里的“适应”与日常的“适应”
- 语文里的“矛盾”与逻辑里的“矛盾”

同一个现象两科解释不同：

- 一个物体为什么会掉下来：物理的解释与日常的解释
- 一个人为什么会这样做：历史课的解释与语文课的解释

这些地方教材从不处理，因为它们不属于任何一科。而正因为不属于任何一科，它们天然不可搜索——一分钟检验轻松通过。

跨学科岔口的一个额外好处：它需要两位教师合作，而合作本身会把这件事从个人行为变成一件被记录的事（第二十六章第五节的第二条）。

五、五科的完整对照

学科	光滑化的藏身处	AI 该配哪一格	最该防的配错	回写的可见形态
数学	教材上的“我们规定”	显露／土壤	—	加限定词
语文	教参上的中心思想	土壤	配显露（给范文）	改自己上一篇
科学	“实验验证结论”的次序	显露／土壤	—	分清观察与规定
语言	语法规则的先在性	显露	同时占两格	给规则加边界
术科	评价标准的现成性	土壤	配显露（给范作）	自己定标准并检验

第四列是这张表最有用的一列。两处最该防的配错，都发生在创作类学科上——语文与术科。因为在这两科上，“给一个好的范例”是最自然、最省事、也最被认可的用法，而它恰恰是最有害的一种配置。

六、一条给所有学科的通用起步动作

如果一位教师读完这两章仍然不知道从哪里开始，本书给出一个通用的起步动作，它在任何一科上都成立：

在你这一科里，找一个“最简单、大家都会”的概念，花十分钟查一查：它当年是从哪个困难里长出来的？

第四章第八节说过，多数教师做完这件事之后的反应是相似的：原来我教了二十年的这个东西，我自己也只见过它的遗体。

这不是能力问题——没有人教过我们去看它的活体，因为教材上没有。

而查到之后要做的事，不是把这段历史讲给学生听（第四章第九节已经说明，那只是增加了一个新成品），是问一句：这个困难，能不能在今天的孩子身上重新做出来？

能，它就是一个岔口。只能讲，就先放下。

七、第五编的结论

第五编到此结束，八句：

第一，从问题开始，不从活动开始。合格的问题只有一条判别：一分钟内搜不到大家都点头的答案。三种形态、五个矿脉。

第二，三方碰撞的条件是分处不同位置且各带独家材料。第三方真正的作用不是多一个帮手，而是把显露从教师手里拿走，使教师第一次可以真诚地不知道。

第三，位置由材料决定，不由身份决定；一个单元至少轮换一次；教师退出路径去占土壤，课照样发生，而且往往更好。

第四，两条硬约束：闭合权不参与轮换；任一方不得同时占据两格。执行靠装置，不靠自觉——一个计时器，一句写死的提问句式。

第五，塌陷方向固定（学生先退场），因此必须硬性排进那个顶不动为止的动作；并且已经预先写出这套方案被筛选之后会剩下什么。

第六，四段结构：开缺口八分钟、顶十五分钟、反例十分钟、回写七分钟。倒着排时间。若只做一件事，做第四段。

第七，教师、学生、教室是产物不是座位。教师的不可替代性落在一个很窄但很硬的位置上——闭合权的行使需要读取一个尚未发生的具体情境。

第八，五科各有藏身处与配置方式；两处最该防的配错都在创作类学科上。

第六编要做的是验证与划界：读数怎么采、实验怎么设计、什么条件下本书被推翻、以及与十家相邻主张的分别。

八、与相邻说法的分界

学科教学知识（教师对“如何教这一科”的专门知识）是一个成熟的研究领域，本章的分科落地在形式上属于同类工作。差别在于对象：那一领域研究的是如何把这一科教得更有效——即如何降低这一科特有的理解障碍；本章研究的是这一科的光滑化藏在哪里——即这一科特有的、被删掉来路的方式。两者所需的学科理解力是同一份，用途相反：前者用它铺平道路，后者用它找出道路当年是怎么被铺的。这也是第三十五章那句话在学科层面的形态：讲得清楚的能力，正是设计岔口的必要条件。

九、可观察的动作

拿你这一科的下一个单元，按第五节那张五科对照表的第四列，问一句：我平时最省事的那种 AI 用法，是不是正好是该防的那一种？

创作类学科（语文、术科）的答案通常是“是”。

第五编材料：三节完整课例与各科岔口开采清单

本编材料的性质：三节课例为按第三十四章四段结构编写的教学设计示例，标注了每一步的机制归属与常见失手处；非课堂实录，学生发言为示例构造。

材料十五 课例一（数学四年级）：三角形内角和

岔口：你怎么知道所有三角形都是一百八十度？（形态一：追问一个默认前提）

一分钟检验：通过——这个问题的答案不在教材上，它针对的是这个班刚刚做过的那个动作。

第一段·开缺口（8分钟）

教师不问“内角和是多少”。先让全班撕角拼直线，做三个。做完之后问那一句。

学生会答：撕过了，都对。

追问：“你撕了几个？”——三个。“那第四个呢？第一万个呢？”

当众立规矩：“接下来十分钟，谁都不许给答案——包括我，包括 AI，包括课本后面那一页。”计时器放上讲台。

⚠ 本段常见失手：在学生答“三个”之后立刻说“对，所以我们不能只看三个”。这句话把结论替他说了。**正确做法是只问，不接。**

第二段·顶（15分钟）

学生把理由打进去：“我撕了三个都对，所以都对。”

只许说两种话；不许问“这个怎么做”。

预期的顶撞过程：系统会指出“三个不能代表全部”；学生会说“那我再撕十个”；系统会问“十个够吗”；学生开始意识到这条路走不完。

顶到什么程度算够：到他自己说出“再撕多少个也不够”为止。**不是教师说出来。**

⚠ 本段常见失手：教师在第四十秒开口。对策是计时器，不是自觉。

第三段·反例（10分钟）

放进一句：球面上的三角形，内角和大于一百八十度。可以拿一个地球仪或一个球，当场画。

放进去之后不解释。只问一句：

“那我们撕角拼直线那一下，到底靠的是什么呢？”

预期的挖掘过程：先有人说“靠三个角能拼成直线”；再有人说“靠纸是平的”；最关键的一句通常最后才出现——“靠那三个角挪过去的时候没有变形”。

⚠ 本段常见失手：教师把球面几何解释一遍。学生会点头，气氛良好，而没有任何人自己说出那个前提。

第四段·回写（7分钟）

“你笔记上写的那句‘三角形内角和是一百八十度’，现在要不要加几个字？”

加得出“在平面上”这四个字的，就是发生了。加不出的，是还知道着。加完念出来。

黑板上最后留下的那句：一百八十度不是三角形的性质，是平面的性质。

三方各带了什么：教师带一道题，AI 带一个球面，学生带撕过的三个角。**这句话谁都没带来。**

材料十六 课例二（语文·五年级）：中心思想是谁定的

岔口：这篇课文的中心思想，作者说过吗？（形态二：追问一次判定的来源）

第一段·开缺口（8分钟）

先让学生说这篇课文的中心思想。多数人会说教参上的那一句（他们从别处听来过）。

然后问：“这句话是作者说的吗？他在哪里说的？”

学生会翻课文。翻不到。

立规矩。

第二段·顶（15分钟）

让学生对系统说：“我觉得中心思想是……你说我哪里不对。”

预期过程：系统会给出若干种不同的读法，有些互相不一致。学生会问“那到底哪个对”——这时候教师不接。

这一段的关键动作：让学生找出“支持我这个读法的句子”和“不支持我这个读法的句子”，各三条。

⚠ 本段常见失手：演变成“每个人的理解都对”。**对策：**要求每一种读法必须拿出文本里的句子，拿不出的不算一种读法。这一条同时守住了第七章的第四条约束（可批评性）。

第三段·反例（10分钟）

放进一个材料：同一件事，从另一个人物的角度重写的一段（可由系统当场生成，或教师准备）。

不解释。只问：“如果这篇课文是从他的角度写的，中心思想会不会变？如果会变，那原来那个中心思想到底是谁的？”

第四段·回写（7分钟）

“你上一篇作文里，为了凑主题写下的那一句，是哪一句？现在你还留着它吗？”

⚠ 配置提醒：这节课上 **AI 绝不能占显露**——它给范文，学生照抄结构，写作当场死亡。它在这里占的是土壤（提供别人的读法与另一个角度的材料）。

材料十七 课例三（科学·六年级）：植物长的那几十公斤从哪来

岔口：一棵树长了几十公斤，土只少了几十克，那几十公斤从哪来？（形态一+材料一的第九个标本）

第一段·开缺口（8分钟）

先给那个事实：有人做过一个实验，把一棵小树苗种在称过重的土里，浇水，五年后树重了几十公斤，土只少了几十克。

问：“那几十公斤从哪来？”

先让全班写下自己的猜想，不许交流。写下来这个动作很重要——它把猜想变成了一件可以在第四段被回头修改的旧东西。

立规矩。

第二段·顶（15 分钟）

学生的猜想通常集中在“水”和“土”。让他们对系统说自己的猜想，并要求它挑错。

预期过程：说“水”的会被问“那水的重量够不够”；说“土”的会被数据挡回来。

这一段最有价值的一刻是有学生说出“那只能是空气了”，而其他学生立刻反对——“空气怎么会有重量”。这个反对是免费的异议，教师不要压掉它，让他们顶。

第三段·反例（10 分钟）

最有价值的用法在这里：让系统复原当年被淘汰的那个理论，并尽力把它讲得有道理——比如植物的重量来自水的转化，或来自某种土壤中的“营养精华”。

让学生去驳它。

⚠ 一个能把旧理论讲得让小学生一时驳不倒的对手，比十遍正确解释更有用。

第四段·回写（7 分钟）

“把你第一段写下的猜想拿出来，在下面写一句：我原来以为……现在我知道……”
念出来。

这节课的特点：第一段那个“写下猜想”的动作，使第四段的回写有了明确的落点。其他两节课的回写依赖学生自己找到一件旧的，这节课在开头就替他准备好了。这是一个可以移植到任何学科的技巧。

材料十八 各科岔口开采清单

按第二十九章的三种形态与五个矿脉，给出可直接使用的问句库。

追问一个默认前提（形态一）

- 数学：你怎么知道所有的都这样？我们试了几个？
- 科学：这个结论在什么情况下会不成立？
- 语文：这个说法，换一个读者还成立吗？
- 语言：这条规则有多少例外？例外多到什么程度，规则就不成立了？
- 历史：当时的人知道结局吗？

追问一次判定的来源（形态二）

- 数学：这条规矩是谁定的？当年能不能定成别的？
- 语文：这个中心思想是谁定的？作者说过吗？
- 科学：这个单位/分类是谁定的？为什么这么分？
- 语言：是先有规则，还是先有人这么说？
- 术科：这个“好”的标准是谁定的？换一个时代还算好吗？

追问一个边界（形态三）

- 通用：换成……还成不成立？
- 通用：如果把这个条件去掉会怎样？
- 通用：这个方法能不能用在完全不同的一类东西上？
- 通用：你能不能造一个反例？造不出来，能不能说说为什么造不出来？

一条使用提示。这份清单的用法不是逐条去用，是拿它去扫教材：翻开下一个单元，看哪一处能挂上其中一句。一个单元通常能挂上三到八处，而一学期只需要用掉三到五处（第二十三章第五节）——材料从来是过剩的，缺的是往这个方向找的习惯。

第五编材料（续）：课堂是被提前结算的地方

本节材料的来源：二〇二六年八月一次面向小学全体教师的三天培训中，第一天由十四位在职教师各自产出的《何谓课堂》批次，以及由该批次装成的《SDE 课堂发生学卷一·课堂的本质》。本节所引的编码事实（COPUS 码表、分配教学时间文献）为真实的方法学事实；教师姓名与个案细节从略。

材料三十二 十四份稿子撞出的同一句话

那批稿子有一个从设计上没有预料到的结果：十四份稿子在文字层几乎零重叠，而在承重判断层重叠到了十四分之十。十位教师，各自独立，用各自的学科语言，写出了同一形式的一句话：

凡为使学习可见而架设的装置，在它记录到之前，就已经关掉了它要记录的那样东西。

这句话与本书第八章的结论方向一致，但它比本书说得更狠。本书说的是读数只覆盖了一半；这句话说的是读数的架设动作本身就是关闭动作。

由这十四份稿子提炼出的母题是：

课堂不是学习发生的地方，是学习被提前结算的地方——使发生可见的那个动作，在发生开始之前就把它的位置占掉了。

这句话对本书的补强在于：本书第五编讲的是“怎么在课堂上留出一段”，而这句话指出了留不出来的结构性原因——课堂这个形式的存在理由之一，就是让学习可见；而可见与发生在时间上不相容。

这不构成对第五编的否定，但它说明第五编那条“一学期三到五处”的量为什么必须这么小不是因为保守，是因为每一处都是在与这个形式本身对抗。

材料三十三 三个产物：教师、学生、教室

同一次培训里定下的另一条判断，本书第三十五章已经用了，这里补上它的完整形态：

教师发生、学生发生、教室发生——这三者是教学发生过程的产物，不是它的前提。

三句话展开：

不是在那个物理场所里的就是教室。教室是“这十分钟谁都不许给答案”这条规矩立住的地方。

不是站在讲台上的就是教师。谁守住了岔口，谁在那一刻发生为教师。

不是坐在座位上的就是学生。谁走出了路，谁在那一刻发生为学生。

这三句与本书第三十一章的位置轮换是同一件事的两种说法：位置轮换讲的是材料如何决定分派，三产物讲的是角色如何被这场分派生产出来。两者合起来才完整——前者是操作，后者是本体。

材料三十四 一条被真材料逼出来的改判

这批稿子里有一处改判，值得完整记下来，因为它是本书方法论的一个样本。

原来的主张：“账本上没有留白这一栏。”意思是课堂上那段未闭合的时间，在任何记录系统里都不存在。

被推翻的过程：教学时间研究里有一条成熟的链条（从投入时间到分配教学时间到学业投入时间到成功学习时间），其中把过渡、等待、离题等非教学时间量到了分配教学时间的17%–29%。

也就是说：有这一栏，而且量过。

改判后的主张：账本上有留白这一栏，但它的符号是负的——它被记为损耗，是需要被压缩的对象。

这一处改判比原主张更有力，理由有二：

第一，它把问题从“看不见”改成了“看得见但被判为负”，而后者更难反驳也更容易检验——只要翻开任何一份课堂效率研究，那一栏就在那里，前面带着一个减号。

第二，它直接解释了第十九章“摩擦”那一节：那五道减毒装置之所以被逐一拆除，不是因为没人注意到它们，恰恰是因为每一道都被准确地测量过、并被判定为应当消除的浪费。

材料三十五 COPUS 那一格：报废与结清被记成同一件事

同一批稿子的打磨过程中，用一份成熟的课堂观察码表做核对，撞出了一条更硬的东西。

那份码表的规则（用于说明格式，非本书所创）：约二十五个行为码，两分钟一个时间格，格内某行为持续五秒以上即计一次，且不跨格追踪同一个学生。

撞出来的那一条：

学生不发声，在这类码表里默认归入“倾听”这个正码。

于是两件性质完全相反的事，被记成了同一件：

一个孩子正在想（沉默、未闭合、判断正在产出）——记为倾听。

一个孩子已经放弃（沉默、已闭合、判断权已交出）——也记为倾听。

报废与结清，在码表上不可区分。

这一条对本书第三十八章有直接后果：读数二（未闭合时长）单独使用是不够的，因为它测的正是这段沉默，而沉默有两种。这正是第三十八章纪律二“读数二必须与读数四配对”的经验根据——孩子对 AI 说的第一句话，是目前唯一一个能把这两种沉默分开的低成本读数。

由此还得到一条“沉默的两种死法”的分辨判据，可以直接用于课堂观察：

不按“为什么沉默”分，按“还接不接得回来”分。

问一句就能分开：沉默结束后，问他“你刚才在想什么”。说得出一点东西的是前一种；答“没想什么”或“在等你讲”的是后一种。

材料三十六 命名把它转轨，而不是杀死它

这批稿子里另有一条值得搬进本书的判断，它修正了一个常见的说法。

常见的说法是“命名杀死经验”——一旦给一件正在发生的事命名，它就停止了。

站内那篇稿子的修正是：

命名不杀死它，命名把它转轨；动作总量不变，户主变更。

一个孩子正在自己摸索一个做法，教师说“这叫做假设检验”——那个摸索没有停止，它继续进行，但它从此挂在了“假设检验”这个已有条目名下，而不再挂在他自己名下。

这一条与本书第十七章的封写是同一个机制在更细的时间尺度上的样子。封写讲的是元规则被写下；转轨讲的是每一次具体的命名，都在把一次正在发生的产出，过户到一个现成的条目上。

它给出的课堂动作也很具体：在孩子还在摸索时，不要给它名字；等他自己说出一个说法之后，再问“这个东西有没有名字”。次序颠倒一下，户主就不同。

第五编材料（再续）：老知识重新发生之外，新知识怎么发生

本节材料的来源：SDE 站内的学科碰撞方法（现行版本经历过多次自我推翻，本节只取其中已被实测检验过的部分），以及一份完整的学科课件设计（三角函数）。

材料六十三 本书第五编的一个缺口

本书第五编讲的全部是老知识如何在一个孩子身上重新发生：岔口、三方碰撞、四段结构、回写。

它没有讲另一件事：新知识怎么发生。

这个缺口不是疏忽，是范围划定——本书的对象是课堂。但它在两个地方会出问题：

第一，它使本书的主张显得比实际弱。如果发生学教育只能让老知识重新发生一遍，那么它与“更深的理解”之间的差别就不明显，第二十五章那条子集关系也会被读成“只是学得更牢”。

第二，它使第二十二章那条“可达域”的主张落空。本书说走过野路的孩子会朝没有路的方向走两步，但没有说走完那两步之后干什么。

本节补上这一段。

材料六十四 同一台机器：把三家撞在一起

站内那套方法的做法可以压成一句：

取三个互不相干的领域，让它们互相顶撞，找出三家共有而从未被说出的那个前提，用其中一家自己的材料把它推翻，撞出一条三家都不会同意、但三家的证据合起来只能得出的新判断。

这与本书第三十章的三方碰撞是同一台机器在不同尺度上的样子：

	课堂上（第三十章）	产新知识时
三方	教师、学生、AI	三个学科
各带什么	一道题、旧账、异议	各自的判据、案例、反例
共有前提	全班默认而未说出的那一条	三家共用而从未被质疑的那一条
推翻的材料	一个反例	来自三家自己的材料
判据	黑板上那句话谁都没带来	结论不能追溯到任何一家

最后一行是同一条判据。本书第三十章第六节说“下课时黑板上留下的那句话，不是三方中任何一方带进来的”；那套方法的说法是“每一段都能追溯到已有路面的，是拼接不是发生”。两句话是同一句。

这一条对课堂的实际含义：第三十章那套动作不是一个教学技巧，它是产新知识的那台机器的小尺度版本。孩子在课堂上做的那件事，与研究者做的那件事，结构相同——差别只在他撞出来的那一条，世界上早就有人撞出来过。

而这个差别在第二十二章的意义上不重要：对他而言，那一条是他自己走出来的路面。

材料六十五 一个完整的样例：三角函数的周期

站内有一份完整的课件设计，本书把它抄成第五编的第四个课例，因为它演示了上一节那台机器在一个具体的中学概念上的样子。

题：三角函数是怎么发生的？（What 题）

选源。先筛掉那些已经被三角学吸收的领域——简谐振动、傅里叶、欧拉公式、天文测量、潮汐分析。这些不能用，因为它们已经是三角学的一部分，用它们只会得到复述。

选定的三家（互不相干，且都不属于数学）：

纺织纹样学——花回，一个图案重复的单元；

复式记账的周期结算——结转、余额、期末；

岁时与仪式——同一个节日为什么是同一个节日。

三家共有的那个前提（四家共用，从未被质疑）：

“回到同一处”是可判定的——存在一个判据，能说这一点与之前那一点是同一点。

推翻它的材料（记账那一家最硬）：

期末“对上”不是自然事实，是**结转规则规定的**；

织造时花回对不上，做法是裁一截、改一针，把回归**修**出来；

“同一个节日”靠的是动作的重演，不是时间自带的周期。

撞出来的那一条（Z）：闭合余项——为了使“回到原处”成立而被消掉的那一部分。

三家账本里它都存在、都不给它本体地位：记账的**尾差**、织造裁掉的**余料**、历法的**闰**。

落到三角函数上的那一刀：

$f(x+2\pi)=f(x)$ 里的那个等号，不是观察到的，是在**定义域上被规定的一次商**。

转一圈算回来，是被判定的，不是自然的。

反例现成：自旋二分之一要转四圈才复原；莫比乌斯带；两枚硬币互绕时中心走的圆半径是 $2r$ ，所以自转两圈——“**一圈是几圈，取决于按谁的账本数**”。

四段结构上的位置：“你怎么知道转一圈就回到原来那个地方”是第一段的岔口；三家的材料是第三段的反例；“那我们说‘周期’的时候，靠的是什么”是那一问；第四段的回写是在“周期函数”那句定义上加限定。

⚠ **一处现场教训。**这份课件在讲的时候，“自旋二分之一”那个反例**没有人听懂**。可用的替换是**手心托杯转圈**（转一圈手臂拧住，转两圈才自然复原）或**螺丝**（转一圈角度相同但深了一牙）。而两枚硬币那条其实**比莫比乌斯带更贴题**，因为它正是“按谁的账本数”那一刀。

这条教训的一般形式：反例的力量与它的正确性无关，只与**它能不能被当场做出来**有关（第四章第九节）。

材料六十六 选源的两条硬规矩，以及它们对课堂的含义

那套方法有两条选源规矩，本书认为它们可以直接翻译成教师备课时的两条：

规矩一：不能用已经被本题吸收的东西。

在方法层面，这意味着不能拿简谐振动去讲三角函数。在课堂层面，它意味着：

不要用同一科内部的材料去撬同一科的概念。

用“数学上另一个更高级的说法”去解释一个数学概念，几乎总是复述。要撬开它，材料必须来自这门课之外。

这一条直接支持第二十九章第五节的矿脉四（跨学科的碰撞点），并说明了它为什么是五个矿脉里最有力的一个。

规矩二：三家的理论必须足够厚。

判断厚薄有四条：有没有奠基层、有没有有名有姓的两派分歧、有没有反例传统、有没有现行教科书。

在课堂上的翻译：一个用来当反例的领域，必须厚到孩子能在里面找到不止一种说法。如果那个领域只有一种说法，它就只是另一个成品。

而这一条与材料六十一的心法一是同一件事的两面：薄账本上的空位多半是假空位；薄领域拿出来的反例多半是另一个权威。

材料六十七 一条留给学生的话

本节最后需要交代一件事，否则它会被误用。

不能把上面那套方法当作一门课教给学生。理由与第四章第九节完全相同：**把方法作为一个成品端出去，只是增加了一个新的成品**。孩子会学会说“共有前提”“推翻材料”“三家分处”，而这些词与他自己撞出一条判断之间，没有任何关系。

正确的用法是**教师侧的内部机制**（第三十章第二节已经写过这条纪律）。孩子那一侧只需要三句大白话，而且这三句正是第二十九章的三种问句形态：

这件事，是不是本来可以不是这样？

别的地方，有没有人是用另一种办法办的？

要是换成……你这个说法还成不成立？

三句话，不需要任何术语，而且它们在任何一科上都成立。

一个孩子如果在离开学校之前，学会了在遇到一样现成的东西时问出第一句——那么这本书全部的目的就已经达到了。其余的框架、读数、判据、撤稿条件，都是为了让这一句在一间四十人的教室里、在一张不改的成绩表下面，还有机会被问出来。

第六编 验证与边界

前五编给出了诊断与处方。第六编做四件事：把五个读数的采集方式与使用纪律写死；给出两组可判决的实验设计与撤稿条件；列出六种可预见的失败模式与一份教研组自查清单；对十家相邻主张逐条划界。最后一章交代本书的方法、限度与结论——包括在什么条件下本书应当被推翻。

第三十八章 五个读数及其采集纪律

一把新秤如果昂贵，它就不会被使用；不被使用的秤，等于不存在。所以这五个读数的第一设计原则不是准确，是便宜。

一、设计原则再述

第二十六章给过四条设计标准，这里作为本章的前提重述：不占用教学时间、不依赖学生配合、不依赖教师主观判断、难以被应试训练。

需要补一条更根本的：

第一设计原则不是准确，是便宜。

一个准确但需要三十分钟采集的读数，会在第三周被放弃；一个粗糙但零成本的读数，会被用三年。第八章已经说明，可数据化的部分被越测越细，不可数据化的部分被越挤越薄——要让第二个冲程进入数据，唯一的办法是让它便宜到没有理由不采。

二、读数一：回写率

定义。迁移题正确率减复现题正确率。

采集。一份十题小卷，五道复现（与例题同型）、五道迁移（需要用这次的判断去改写一个原有认识）。当堂一次，一周后一次。可与常规作业合并。

读什么。不看绝对值，看**两条曲线的相对走向**。老路的典型形态是复现高、迁移低且快速衰减；新路预期两者差距缩小、衰减变慢。

局限。这是五个读数里最贵的一个（需要额外命题），也是最容易被训练的一个——迁移题一旦形成固定题型，就退化成复现题。**因此它必须每学期换题，且不得公布题型。**

三、读数二：未闭合时长

定义。从教师抛出问题，到任何一方给出答案，中间经过的秒数。

采集。录音的时间戳即可。零成本，自动，不需要任何人做任何额外动作。

读什么。只设下限，不设上限。经验参照：绝大多数常规课堂在十秒以内；能稳定超过三分的课，一学期没有几节。

为什么它是五个里最重要的一个。因为它最难作假，而且它直接测量第十四章那把剂量尺——未闭合时长的倒数就是“高效”。一所学校如果只做得了一个读数，做这个。

局限一：它不区分那段时间里发生了什么。三分的沉默可能是全班在想，也可能是全班在发呆。因此它必须与读数四配对使用（第五节纪律二）。材料三十五给出了这一条的经验根据——在成熟的课堂观察码表里，“正在想”与“已放弃”被记成同一件事。

局限二（本书成稿后由材料五十八逼出的一处实质修订）：它其实是两个量。

有一项任务在等的那一段，与没有任何任务在等的那一段，是两种不同的东西。

站内那次真跑在同一份数据上分别测了这两个量，按学生能力分组之后：前者的最优时长随能力上升而单调缩短（这是教育心理学的已知结论），后者**不随能力移动**。两个量对同一个调节变量给出了相反的反应。

本节原来定义的读数二（从教师抛出问题到有人给出答案）量的是**前者**——题已经在那里等着。而与本书主张真正相关的是**后者**：没有任何一方持有启动权的那一段。

采集上的做法：在记录未闭合时长时，多记一个二值标记——**此刻有没有一项任务在等**。有，记为A型；没有，记为B型。两者分开统计。这不增加任何成本，而它把一个混合量拆成了两个。

四、读数三：先在感指数

定义。五题小量表，测学生在多大程度上认为某对象“本来就是这样、不可能是别的样子”。

题目形态。“如果当年的人选了另一种做法，今天这条规则会不会不一样？”“这个规定有没有可能被定成别的样子？”“如果全世界的人都不这么做了，它还算对吗？”

采集。一学期两次，每次五分钟。

读什么。本书的核心预测是反直觉的：老路上该指数随学龄**单调上升**；经过元层解锁的班级，该指数**下降**。

局限。这是五个里最像传统量表的一个，因而也最依赖学生配合、最容易被社会赞许性污染。孩子很快会学会“老师希望我说可以不一样”。**所以它只适合用于班级层面的前后对照，不适合用于个体判断。**

五、读数四：对 AI 说的第一句话

定义。学生向 AI 发出的第一句话，分为两类——索取型（“这个怎么做”“答案是什么”）与对抗型（“你这里不对”“要是换成……你还成不成立”）。

采集。对话记录本身就是数据。完全不依赖配合，因为它就是他做的事。

读什么。两类的比例，以及这个比例随时间的变化。

为什么它特别有价值。它直接测量 AI 被配置在哪一格（第三十二章第三节），而这是全书最关键的一个配置问题。第十三章说过，参与度指标不区分提问的方向——这个读数补的正是那个缺口。

局限与一条硬纪律。采集对话记录涉及学生的隐私。**必须以班级聚合的形式使用，不得用于个体评价，不得向家长出示个体记录。**这一条不是可选的礼节，它是这个读数能否被长期使用的前提——一旦学生知道自己的每一句话都会被单独看，他会开始表演，这个读数当天失效。

六、读数五：面对空位题的反应

定义。给出一道确实没有现成答案的题（第二十二章第五节的三条出法），记录两件事：他在多少秒后停止尝试；停止之前有没有朝那个方向走出过任何一步。

采集。一学期一题，占用半节课。

读什么。“走出过一步”的人数比例。不看内容质量，只看有没有。

为什么它几乎不可能被应试训练。因为没有可训练的答案，只能训练“敢不敢走两步”——而那正是它要测的东西。这是五个读数里唯一一个“被训练”等于“被改善”的读数。

七、另外三个读数： τ 、 t_1 、R

本章的五个读数面向课堂。另有三个读数面向**供给方与制度**，它们的性质不同，必须单列，第四十八章将说明它们为什么至今没有被测过。

读数 τ （未闭合时长比）。一节课中处于未闭合状态的时间总和，除以课堂总时长。它是读数二的比值形式，便于跨课比较。

未闭合状态须同时满足三条：有一个问题已被明确抛出且至少有一名正在处理它；尚未有任何人（教师、同学、教材、屏幕、AI）给出可用于对齐的答案、方向性提示或求解步骤；教学活动尚未转向别的内容。

计时规则。从问题抛出起表；出现下列任一情形即停表：给出答案；给出足以还原答案的提示（含逐级降低难度的提示链）；教师说出“这个等会儿统一讲”；转向下一环节；下课。

四个误差来源必须写进编码手册：一是“假等待”——教师沉默但已经用板书、表情或语气给出了方向；二是“假未闭合”——学生其实已经放弃，没有人在处理；三是计时起点模糊（问题是分几次抛出的）；四是编码者对提示强度的判断差异。可用阈值：编码者一致性不低于0.6，两次编码相关不低于0.7。

现有量级参考。手工试编的结果是：常规讲授课的 τ 在百分之一到百分之三之间；被评为优秀的探究课很少超过百分之八。也就是说，一节四十分钟的课，真正处在“没有人替他闭合”状态的时间，通常不足一分钟。

方向预测是明确的：若“AI 赋能”的说法为真（它只是提高效率而不损害目标），则 τ 至少不应下降。本书预测它会显著下降。

一个从未有人算过的量： τ 的基尼系数。 τ 在一个班里的分配是极不均匀的——教师会本能地把等待时间留给那些“看起来在想”的学生，而对看起来茫然的学生迅速给出提示。这个不平等程度可以用基尼系数刻画。据作者所知，从来没有人算过这个数。它直接关系到第四十二章第三节承认的漏洞三（学生差异被忽略）。

读数 t_1 （求助前独立时长）。学生从遇到障碍到发出求助（举手、提问、点击提示、向AI 提问）之间的时长分布。

必须看分布，不能看均值。在传统条件下， t_1 通常呈双峰：一峰在很短处（习惯性求助者），一峰在较长处（习惯自己先扛一段的人）。本书的预测是：在即时应答可得条件下，右峰会向左迁移并最终并入左峰，分布单峰化。

均值会掩盖这件事——如果右峰的人变多而左峰的人求助更快，均值可以纹丝不动。

取数成本接近于零。任何一个学习平台的日志里，“上一次内容呈现的时间戳”与“本次求助的时间戳”都是现成字段。不需要新采集，不需要问卷，不需要观察员。

这是本书最有力的一条挑战：这个数据已经存在，就躺在供给方的服务器里，取出来的成本不到一个工程师一天。它至今没有被公布过。

读数 R （责任可指认性）。对于这堂课上做出的关键教学判断，能否指认出一个具体的人，他会因为这个判断出错而承受后果。

操作方式。课后对教师提三个问题：今天这节课上，哪一个决定如果错了，后果由你承担？这个决定是你做的，还是系统、模型、教研组或教参做的？如果这个学生这学期就此掉队，会有人来问你吗？

为什么它必须与 τ 并列。因为二者是相乘的关系，不是相加。一段无人闭合的时间，如果没有人为它的失败负责，它不是教育，是放任；一份清晰的责任，如果没有任何一段悬空可供承担，它不是教育，是执行。第四十八章第八节将把它写成本书的核心形式命题。

八、三条使用纪律

新秤会以三种方式失效，每一种都发生过。

纪律一：用于诊断，不用于排名。

一旦回写率被用来排教师名次，它会在一个学期内被优化成一个新的答对率——教师会开始训练学生“表演回写”。让孩子在下课前说一句“我原来以为……现在我觉得……”是极容易训练的，而训练出来的那句话与真实的回写没有任何关系。

纪律二：不单独使用，必须配对。

最重要的一对是**读数二与读数四**：一节课未闭合时长很长，而学生说的全是“这个怎么做”，说明只是拖了时间。

另一对是**读数一与读数三**：回写率上升而先在感指数不降，通常说明迁移题已经被训练成了新的复现题。

纪律三：未闭合时长不设上限。

把它变成越长越好的指标，会立刻产出为拖时间而拖时间的课。

三条纪律的共同根据是同一句话：

任何一把秤，一旦被用来结算，就会开始被优化；而被优化的秤会停止测量它原本要测的东西。

这一点在老路的读数上已经发生过一次（第八章），不应该在新秤上再发生一次。

九、最低可行版

五个读数对一所刚起步的学校仍然偏重。最低可行版是：

只做读数二与读数四。

理由是这两项完全自动——录音的时间戳与对话记录本身，不需要任何人做任何额外动作，不占用一分钟教学时间。

仅此两项，已足以判断一所学校的课堂究竟在哪条赛道上：**未闭合时长在十秒以内、学生对 AI 说的全是索取型——这两条同时成立，那么无论这所学校在做什么改革，它都还在老路上。**

十、一份两页的记录表

为便于实际使用，这里给出一所学校可以直接复印的记录表结构。

项目	频率	谁来记	记什么	用途
未闭合时长	每节课	录音自动	秒数	教研讨论
对 AI 的第一句	每次使用	系统自动	两类比例	班级聚合
回写率	每单元	与作业合并	两个正确率之差	诊断
先在感指数	每学期两次	五分钟量表	班级均值变化	前后对照
空位题反应	每学期一次	半节课	走出一比例	诊断

最后一列**没有一项写着“评价”**。这是本表最重要的一处。

十一、可观察的动作

在你们学校下一次的教学质量报告上，加一栏，标题写“未闭合时长”。第一次可以是空的。

第二十六章第七节说过同样的话。一个被记录的东西，就开始存在。

十二、与相邻说法的分界

教育测量领域有成熟的“过程性数据”与“学习分析”研究，能够采集比本章精细得多的行为数据。差别在于用途与设计目标：那一领域的的数据几乎全部服务于预测与干预掌握程度——预测谁会掉队、何时该推送什么内容，其目标函数仍然是提高接受效率（第十三章的个性化错觉）。本章的五个读数服务于一个相反的目标函数：测量那段没有被填满的时间。同一套技术可以采集两者，而没有人采集后者，不是因为技术不够，是因为从来没有人要它。

第三十九章 检验设计与撤稿条件

一本提出机制的书必须写明在什么条件下它是错的；而且不能为自己保留任何逃生解释。

一、本书需要被检验的四条

前五编提出了若干判断。其中大部分是描述性的、可以由读者自己核对的（三个动作、光滑化、老路的读数体系）。真正需要实证检验的是四条：

命题一（加固悖论，第十二章）。AI 用于提高投递效率会使老路的全部读数变好，同时第二个冲程的读数变差。

命题二（封写，第十七章）。成功的接受会加固一条元规则，从而关闭回写通道；因此对象层的补救在通道已封时无效，必须先做元层解锁。

命题三（子集关系，第二十五章）。老路的能力是新路能力的子集：走过新路的孩子在复现题上不显著更差，在迁移题与未见题型上显著更好。

命题四（剂量效应，第十四章）。独自使用 AI 答疑的时长与先在感上升、回写率下降正相关；而把提问形态改为对抗型后，该相关显著减弱。

本章给出前三条的实验设计。命题四的设计已写在第十四章第九节，不重复。

二、检验一：配置方向的对照（判决命题一）

设计。同一批教师、同一批内容、同一学段，随机分为两组。

A 组：把 AI 配置为答案源。学生可随时提问求解，鼓励使用，不设限制。

B 组：把 AI 配置为对手。只允许对抗型提问（第三十二章第四节的两种句式），并设置未闭合时长下限。

周期：一学期。

因变量：两套读数同时采集。

老路读数：正确率、进度、掌握率、满意度。

新读数：回写率、未闭合时长、先在感降幅、对 AI 的第一句话。

预测。A 组在老路读数上不低于甚至高于 B 组；B 组在新读数上显著优于 A 组。

为什么这组实验最重要。因为它的预期结果本身就是加固悖论的直接证据——**两组的传统读数会告诉你 A 组更好。**

撤稿条件（写死）。如果 A 组在两套读数上全面胜出，第十二章即被推翻，第二编的大部分内容需要删除。

三、检验二：解锁次序的 2×2（判决命题二）

设计。两个因子，四组。

因子一：是否做对象层去光滑化（把当年的路径还给学生）。

因子二：是否先做元层解锁（第三十四章第七节的三问）。

预测：交互效应显著。

单做去光滑化，在先在感指数高的学生身上效果微弱；先做元层解锁再做去光滑化的一组，回写率显著高于其余三组。

同时判决一个竞争假说。第十七章第九节说过，“本体误归类”是教育文献中最强的竞争者。它的预测与本章不同：按误归类，有效的干预是**换范畴**（把概念重新安置到正确的本体类别里），而元层解锁在它的框架里不应产生额外效果。因此本设计可加第三臂——只做换范畴干预——三臂对照即可判决。

撤稿条件（写死）。如果交互项不显著，或效应量低于中等水平，第十七章的封写机制须撤回。

并且明确堵死一个逃生解释：不得以“去光滑化做得不够好”来解释失败。这条逃生口在此关闭。理由在第十七章第三条推论——本书自己主张“路走对了回写照样可能不来”，因此它没有资格反过来用“路没走对”为自己开脱。

四、检验三：子集关系（判决命题三）

设计。接受过若干次真实发生学教学的学生，与匹配的对照组比较。控制初始水平。

三条预测：

一、复现类题目上**不显著更差**。

二、迁移类题目上**显著更好**。

三、未见题型上尝试时长**显著更长**，放弃率**显著更低**。

撤稿条件（写死）。如果第一条不成立——实验组在复现题上显著更差——那么“子集”这个说法即被推翻，第二十五章必须改写为一个取舍关系（用一部分掌握度换一部分发生力）。

这将极大削弱全书对一线的说服力，因为第二十五章是唯一能回应升学压力的东西。**本书把这一条列为最应当被优先检验的一条。**

五、一份可以照做的对照方案

第二至四节的三组设计需要研究者介入。本节给出一份一所学校在一个学期内可以自己完成、不需要额外经费的方案，它检验的是全书最基础的那个变量。

设计。同一位教师、同一个班、同一门课，交叉设计：单周为 A 条件，双周为 B 条件（或按单元交叉，以避开单双周的系统差异）。

唯一的自变量：小组活动时段内，教师抛出的“未闭合问题”个数——A 条件三个，B 条件九个。

必须严格匹配的量：小组活动总时长、教师巡视时间、教师说话总时长，匹配到分钟。

这一条是整个设计的关键，也是野外数据里**从来不存在的对照**。现实中“探究课”与“讲授课”相比，同时变了十几个量：时长、氛围、教师投入、学生兴奋度、家长知情度。于是任何结果都可以被归因到任何一个变量上。**这份设计把它们全部锁死，只留一个。**

主要结果指标：四周后的延迟迁移测验，不是即时后测。

这一条必须提前写死：**用即时后测跑出的零结果不构成对本书的反证**。因为本书的整条论证（第四十四章第三节的三条实证线索）都预测即时表现在 B 条件下更差。**用即时后测去检验，等于用体温计去量体重。**

盲法。出题者、编码者、评分者三方均不知道被试属于哪个条件。教师本人不可能盲，这是本设计无法消除的限制，必须写明。

判定规则在看数据之前写死：主要指标的效应方向与最小有意义差值（例如迁移题正确率不低于八个百分点）、次要指标（ τ 实测值、 t_1 分布形状）、以及什么情况下判定为失败。

六、四种失败方式，全部前置写明

一、**匹配崩溃**——最可能的一种。教师在 B 条件下不自觉地多讲、多提示、拖堂，把时间匹配破坏掉。防范办法只有一个：**第三人计时，且计时结果当场公示。**

二、**提示渗透**——教师在 B 条件下嘴上不给答案，但用表情、板书、语气给了。防范：录像抽样由第三方按提示强度编码。

三、**学生代偿**——学生在 B 条件下回家问家长或问 AI，把闭合搬到了课外。防范：课后当天记录课外求助——**这本身是一个值得单独报告的读数。**

四、**样本不足**——单班交叉设计的统计效力有限。防范：这份方案的定位是**可复制的最小单元**，多校汇总而非单校定论。

七、四种结果各意味着什么

预先写明结果的四种可能，以及各自的解释权归属——**这是防止事后叙事的唯一办法。**

结果一：B 条件的延迟迁移显著优于 A，且 τ 实测差异符合预期。核心机制得到支持。但注意：这只支持“未闭合时长是有效变量”，**并不自动支持本书关于 AI 的全部推论**——那需要第三臂（B 条件中由 AI 提供即时应答）来检验。

结果二：两组无差异。最有信息量的一种。此时必须先检查匹配是否成立与提示渗透编码。若匹配良好、渗透可控而仍无差异，则本书关于 τ 的主张在这一学段、这一学科上**不成立**，应当被削弱为局部主张。本书接受这个后果。

结果三：A 条件优于 B（多给闭合反而更好）。核心机制在该条件下被推翻。但要注意，这个结果与认知负荷一族关于“最少指导的教学为何不奏效”的判决方向一致，因此并不令人意外：在先备知识不足的学生身上，本书的主张很可能反向成立。这一点第四十二章已作为正式限度写明——**不是被推翻后才补上的补丁，而是事前就承认的边界。**

结果四：平均无差异，但方差显著变化。最值得追踪的一种。如果 B 条件让强的更强、弱的更弱，那么本书主张的做法会**加剧判断力的不平等**。这不是本书乐见的结果，但它是预先承认的真实风险，而且与第四十八章第六节对怀旧论证的批评在结构上完全一致——**历史上那台机器就是筛选机而不是培养机，我们必须避免造出第二台。**

八、最低可行版

上述三组设计对一所学校而言过重。这里给出可以立刻开始、几乎零成本的替代方案：

| **一所学校、一个年级、两个班、一学期，只做检验一，只测读数二与读数四。**

两项读数都自动采集，不占用教学时间。样本小，混杂因素多，不构成严格证据。但它足以回答一个问题：**在这所学校里，B 组学生对 AI 说话的方式有没有变。**

这个问题的答案，比任何理论论证对一位校长都更有说服力。

九、本书证据状况的诚实清单

必须逐条说明本书各主张的证据强度，否则读者无法判断该在多大程度上采信。

有直接经验支持的：三个动作的结构、光滑化的存在、四个标本的删除内容、老路读数体系的覆盖面。这些是描述，读者可以自己核对。

有间接支持、机制清楚的：加固悖论。它解释了若干可观察现象（读数上升与某些能力表现的背离、补课无效、顶端提问质量下降），但这些现象也可以有别的解释。

目前只是假说的：封写机制、剂量效应、子集关系。它们有内部一致性和解释力，但没有直接的实验支持。

证据最薄的：第二十一章的三段曲线。它是从机制推出的预测，缺少纵向数据；这类数据需要三到五年追踪。

超出本书能力范围的：全部政策层面的判断（第二十八章）。作者没有政策制定经验，那一章只做翻译，不做建议。

十、一个可能的反驳，以及本书的回应

反驳：这本书的核心机制无法被证伪，因为封写的表现（学得好但不改变）与“学得好”在读数上无法区分。

回应：这个反驳如果成立，本书确实应当被弃置。而它不成立，理由是第十七章的推论二给出了一个**方向明确的预测**：最配合、学得最好的学生被封写得最彻底。

这条预测与所有相邻理论的预测方向相反（那些理论预测失败导致僵化），而且它可以用检验一直接判决——如果 A 组中成绩最好的那部分学生，在新读数上的表现不低于成绩中等的学生，推论二即被推翻。

本书把这条列出来，是因为它是全书最容易被数据打掉的一条，同时也是最原创的一条。这两件事通常是同一件事。

十一、可观察的动作

把本章的三条撤稿条件抄在一张纸上，贴在办公室。一年之后拿出来看。

一本书如果不敢让读者这样做，它的主张就不值得被采纳。

十二、与相邻说法的分界

教育研究方法论中有关于“复杂干预难以随机对照”的成熟讨论：课堂干预涉及大量互动因素，随机化困难，效应难以归因。本书接受这些困难，但要指出它们不构成免检的理由。本章设计的三组实验都选择了行为事件而非态度或能力作为因变量（未闭合时长、第一句话的类型、有没有走出一大步），正是为了绕开归因困难——这些事件不需要推断，只需要记录。代价是它们粗糙；本书认为在当前阶段，粗糙而可执行优于精细而不可执行。

第四十章 六种失败模式与自查十问

一项改革的成败通常不取决于理念是否正确，而取决于它在落地过程中滑向了哪一种熟悉的旧形态。

一、为什么失败模式要被预先写出

第五章证明了筛选机制的存在，第三十三章第六节已经预先写出这套方案被筛选之后会剩下什么。本章把可预见的失败逐一命名。

命名的作用是：一个有名字的失败可以被当场识别；没有名字的失败会被当作“效果不明显”。

二、失败一：终点没换，只换了跑法

形态。课上得很热闹，讨论、小组、汇报、AI 互动一应俱全，但下课时全班到达的是教案上早就写好的那个结论。

为什么它最难自我察觉。因为所有外部观察者都会认为这节课很成功。它在老路的全部读数上表现优异，在直观印象上也极好。

判别。第六章第三节的判据：翻开教案，看结论有没有被写下来。

这是最常见的一种，而且是所有其他失败的母型。

三、失败二：秤没换

形态。理念改了、课改了，期末评价的口径一个字没改。

半衰期：一个学期。教师会迅速发现新做法在旧秤上不占优，然后理性地退回。这不是意志问题——第五章第七节说过，认同不改变回报结构。

判别。看期末出示的是什么数据。

处理。第二十六章第五节的四条，从“先加一栏”开始。

四、失败三：AI 占了两格

形态。AI 既给答案又给路子，碰撞塌陷成投喂。表面特征是课堂流畅、学生提问积极——而他们问的全是“这个怎么做”。

判别。读数四即可当场诊断。

处理。第三十二章第四节那句写死的提问句式，写在教室墙上。

五、失败四：教师忍不住

形态。第二段“顶”的十五分钟里，教师平均在第四十秒开口。

这不是态度问题，是三十年职业本能。唯一有效的对策是外部约束——计时器、录音、同伴观察，而不是靠自觉。

一个补充观察：越是优秀的教师，越忍不住。因为他确实有一句非常好的话可以说，而且说出来确实有帮助。**这一条使得这种失败在最好的教师身上出现得最多。**

六、失败五：省掉第四段

形态。下课铃响了，回写没做，改成“我们来小结一下”。

后果。前三段的全部工作在这一刻作废，因为发生没有落点，学生带走的仍然是一个成品。

为什么它最容易发生。因为它在时间上排在最后，而课总是拖堂。

对策。第三十四章第六节：倒着排时间。

七、失败六：把“变了”听成“气氛好”

形态。教师报告说“孩子们眼睛都亮了”“课堂气氛特别好”“他们特别喜欢跟 AI 讨论”。

这些都是好事，但都不是判据。判据只有一条：他有没有回头改一件旧的（第十六章第四节）。

为什么它危险。因为它让一位教师在没有任何发生的情况下，真诚地相信自己做成了。而这种真诚的误判会在两三个月后转为幻灭——“这套东西看着好，其实没用”。

八、第七种：叠加而非替换

上一节六种之外，还有一种属于实施层面而非理解层面的失败，第二十七章第六节提到过，这里补全。

形态。原有的进度、作业、考试全部不变，再加上新的环节。

后果。教师负担陡增，学生时间被占满，两个月后因不可持续而终止。而终止的理由会被记为“太耗时间，不现实”——于是这套方法背上了一个不属于它的罪名。

对策。替换而非叠加：三到五处发生学教学，替换掉原本用于讲解那部分内容的课时。如果这一条做不到，宁可不做。

九、教研组自查十问

下列十问用于一节课后的教讨论，每问只答“是”或“否”，不解释，不打分。

- 一、这节课的结论，是不是在上课前就写在教案上了？
- 二、我抛出的那个问题，输进任何一个搜索框，一分钟内能不能得到大家都点头的答案？
- 三、从我提出问题到有人给出答案，中间空了几分钟？
- 四、这节课里，学生对 AI 说的第一句话是“这个怎么做”，还是“你这里不对”？
- 五、有没有哪一刻，是学生在走一条我没有预设的路？
- 六、有没有哪一刻，是我在提供土壤而不是提供路径？
- 七、全班的结论是不是都一样？如果是，这是必然的，还是被我引导的？
- 八、下课前，有没有一个学生回头改了一件旧的？改了什么？
- 九、黑板上最后留下的那句话，是我带来的、AI 带来的，还是撞出来的？
- 十、如果把 AI 从这节课里撤掉，这节课会不会变得更好？

十、第十问值得单独说明

它是十问中最锋利的一问，因为它同时检验两件事：AI 是不是被配反了，以及这节课的设计是不是本来就不需要三方碰撞。

一个配置正确的课堂，撤掉 AI 之后应当明显变弱——因为学生失去了那个不知疲倦、不给答案、专门唱反调的对手。

如果答案是“撤掉会更好”，说明 AI 在这节课上占的是答案源的位置，它的存在正在关闭岔口。这时候正确的处理不是撤掉它，是改配置。

十一、过渡期现象的重申

最后重申第二十七章的内容，因为它是失败模式之外的一类现象，极易被混淆为失败。

换道之后的头几周，读数会变难看。正确率下降，进度变慢，家长有意见，学生不适应，而最优秀的学生反弹最强（第二十七章第二节）。

这一段是可预期的，因此必须在开始之前就向所有相关方讲明，并提前约定：在过渡期内不用旧秤做判断，同时给出退出条件。

没有这一条约定，绝大多数改革会死在第三周——不是因为它无效，而是因为它在旧秤上暂时看起来无效。

十二、可观察的动作

把这十问印出来，在下一次教研活动上用一次。规定每问只答是或否，不许解释。

不许解释这一条是认真的。解释会立刻把讨论拉回“这节课上得好不好”，而这十问要问的不是好不好，是在哪条道上。

十三、与相邻说法的分界

教学反思与课堂观察量表在教研中广泛使用，本章的十问在形式上与之同类。差别在于两处：一是那些量表通常按维度打分（如“目标达成度”“学生参与度”），本章只答是否，因为打分会立刻引入好坏判断，而好坏判断会把讨论拉回老路的标准；二是那些量表用于评价教师，本章明确只用于自查与教研（第二十四章第七节的纪律）。一份用于评价的观察表与一份用于诊断的观察表，即使题目完全相同，一年之后测到的东西也完全不同。

第四十一章 与十家的分界

一个新提法如果不能对最近的邻居划出分界线，它就只是换了名字的旧说法；分界必须写在正文里，不能写在脚注里。

一、本章的作用

前面各章末尾都有一节“与相邻说法的分界”。本章做汇总，并处理那些跨章重复出现的高手。

汇总的目的不是重复，而是回答一个整体性的问题：把这十家的成果加在一起，本书还剩下什么是新的？

答案写在最后一节。

二、教学转置

它说什么。知识为了能被教，必须先被去情境化、去个人化，经过一条制度性的搬运链从学术共同体到达课堂。

重合处。与第二章的光滑化在描述层面高度重合。

分界。那一传统关心的是搬运物一侧——知识形态在制度链条上如何变形；本书关心的是接收者一侧——同一次搬运在学习者身上造成了什么。分界线是“被转置的知识”与“接不住转置物的人”。

这条分界为什么重要。因为它决定了能否解释“补课为什么无效”。教学转置理论可以解释知识为何失去情境，但无法解释为什么把情境还回去仍然无效——那需要第十七章的机制。

三、沉积与再激活

它说什么。起源被沉积遮蔽，因而结果显得自明先在，需要通过再激活恢复。

重合处。与第十七章几乎逐句对应。

分界。再激活是一次个人的现象学复演，其完成与否只能由当事人自证；本书的回写是一个可由第三方当堂检验的行为判据——他有没有回头改一件旧的。

这一刀立住，“回写作为判据”才成其为贡献。否则本书只是给一个既有概念换了名字。

四、关系性理解

它说什么。理解分深浅两层：知道怎么做，与知道为什么。

分界。关系性理解是一种**状态**（连接有多少）；回写是一个**事件**（有没有发生过一次改写）。

实践后果。状态需要一套量表去评估，事件只需要一个动作去观察。第三十八章全部读数建立在事件而非状态上，原因就在这里——**事件便宜，状态昂贵；昂贵的秤不会被使用。**

五、双环学习与组织防卫

它说什么。单环的成功会系统性地阻断对框架本身的检视。

重合处。与第十七章的封写形状相近。

分界。那一传统主张**成功导致框架僵化**，机制是防卫；本书主张**锁是被回写这个动作本身写下的**——不是因为懒于修改，是因为那次成功的接受所写下的内容，恰好就是一条禁止修改的规则。

方向相反的可判决预测。那一传统预测的是防卫性行为（回避、掩饰、合理化）；本书预测的是**顺从性行为**——最配合、学得最好的学生被封写得最彻底。第三十九章检验一可判决。

六、本体误归类

它说什么。概念一旦被放入错误的本体范畴，同范畴内的再教学无法纠正，须换范畴。

这是教育文献中最强的竞争者。

分界。误归类说的是**放错了柜子**，本书说的是**柜门上了锁**。前者的补救是换柜子，后者的补救是先开锁再谈放哪。

判决方式。第三十九章检验二的第三臂：只做换范畴干预。如果换范畴组的效果不低于元层解锁组，本书这一部分应当让位。

七、定势效应

它说什么。过去成功的解法阻断新解法。

分界。定势是关于**解题策略**的现象，作用于对象层；封写发生在**元层**，写下的不是一条策略，而是一条关于“底盘是否需要改动”的规则。

一个可区分的表现。定势的表现是他用旧方法去做新题；封写的表现是他**接住新方法但不改旧想法**——他会把新方法也学会，学得很好，而原来那个想法原封不动。

八、同化与顺应

它说什么。纯同化而不顺应会导致图式僵化。

分界。那一传统把僵化归于**缺少顺应**；本书主张有一次顺应确实发生了，只是它的内容是“以后不必顺应”。

不是少了一次，是多了一次错的。

这个差别不是措辞之争：如果是“少了一次”，处方是增加顺应机会（更多的认知冲突）；如果是“多了一次错的”，处方必须先撤销那一次，才谈得上后面（第十八章的诊断次序）。

九、先验幻相

它说什么。某类幻相源于理性自身的结构，不可避免，只能揭示不能消除。

分界。本书所说的先在感是**习得的**，源于一部具体的传递史，因而**可以被去光滑化消解**。

必须同时设的第二道挡板。本书只主张“学习者身上的先验感是被制造出来的”，绝不主张“数学或科学的必然性是幻觉”。这两句话之间的距离，是本书全部立场的安全边界（第十七章第七节）。

十、AI 素养教育

它说什么。把 AI 单列为一门课，教学生如何提问、如何辨别、如何使用工具。

分界。AI 素养解决的是**如何更好地使用一个答案源**；本书主张的是**把 AI 从答案源改配为对手**。

一个不客气的判断。前者可以在完全不改变老路的前提下实施，且事实上会强化老路——一个更会提问的学生，是一个更高效的接受者。

十一、批判性思维训练

它说什么。教授一套可迁移的技能：识别谬误、评估证据、区分事实与观点。

分界。本书与之的分歧不在于这些技能是否有用，而在于它们**被教授的方式**——这些技能本身也是作为光滑成品被端出去的。

一个能熟练背诵十二种逻辑谬误名称、却从未在自己的某个具体信念上真正改过一次口的学生，正是本书所说的封写的典型样本。

分界线：批判性思维教的是一套工具，本书关心的是**那套工具有没有被用在自己身上过一次**。

十二、自适应学习与精准教学

它说什么。按学习者的实时表现调整题目难度与内容顺序，以最短路径把学习者送到掌握状态。

分界。分歧不在技术，在**目标函数本身**：

最短路径正是本书所说的最高剂量。

自适应做得越好，未闭合时长压得越低，第十九章意义上的毒性越高。

这个分歧不可调和——除非目标函数被改写为“在保证抵达的前提下，维持一段最短未闭合时长”。本书认为这是一个技术上完全可行、商业上极难成立的改动。

十三、附论：AI 导师效果研究

严格说这不是一家理论，而是一批实证工作。它们报告了智能辅导系统带来的显著学习增益。

本书不质疑这些数据，但要指出它们测的是什么：几乎全部此类研究的结果变量是**掌握度与完成速度**，即老路的读数。

按第十二章的加固悖论，这些读数在 AI 条件下会与能力退化同时变好，因此它们不能作为反驳。真正的判决实验是第三十九章的检验一，它要求同时报告两套读数——在两套读数上同时胜出，本书才被推翻。

十四、那么本书还剩下什么是新的

把这十家的成果加起来，本书剩下的是三样，其余全部可以在别处找到：

第一，回写作为一个可由第三方当堂检验的行为判据。别处有“理解”“顺应”“再激活”，但都是状态或个人自证，没有一个是可当堂观察的事件。

第二，“回写杀死了回写”这一具体机制。别处有框架僵化、误归类、定势、防卫，但没有一家说锁是被那次成功的回写本身写下的，也没有一家预测“学得最好的被封得最彻底”。

第三，把 AI 的构造倾向（填缝）与第二个冲程的入口（未闭合时长）接在一起，并给出可测量的剂量关系。别处有“技术的双刃剑”“要善用工具”，但没有一把尺子。

其余的——三个动作、光滑化、成本论、老路的成就、教师角色的转变、分学科的做法——一本书没有一样是原创的，也不需要是。

⚠ 这一节的判断不能由本书自己作数。材料六十一记录了一次实测：五个被自认为“没有人占过”的候选，十三次检索之后全部被挡下，而两次凭印象排序都被翻掉。由此得出的心法是——判断“这个想法有多新”的人，不能是提出这个想法的人。

按这条心法，上面那三样“剩下的新东西”，需要一次由别人执行的检索才能作数。本节应当被读作一个请求，而不是一个免责声明。它们的作用是把上面三样安放在一个可以被使用的框架里。

十五、可观察的动作

从上面十家里挑一家你最熟悉的，用它的语言把本书的核心主张重述一遍。看看重述之后，还剩下什么说不出来的。

如果什么都说得出来，那么本书是多余的，应当被丢掉。

十六、一条自我约束

本章的十条分界全部写在正文里，没有一条写在脚注里。这是第五编开工前定下的一条纪律，理由是：

写在脚注里的分界，等于承认自己没有把握。

一个新提法如果不能在正文里、用读者听得懂的话，说清它与最近的邻居差在哪里，那么它多半确实没有差别。

十七、与相邻说法的分界

学术写作中有“文献综述”这一体例，本章在形式上与之同类。差别在于目的：文献综述的目的是定位自己在已有工作中的位置，因而倾向于寻找连续性；本章的目的是划出不连续处，因而每一条都以“分界在于”作结。两者可以互相替换的部分很少——一篇把十家讲得很好的综述，仍然可以完全不回答“那么你剩下什么”这个问题，而本章第十四节正是为回答它而设。

第四十二章 方法与限度

一本处理正在发生之事的书，它的价值不在于说对了多少，而在于它把自己的错误暴露得够不够清楚。

一、这本书是什么

必须先说清本书的性质，因为它容易被误认为三种它不是的东西。

它不是一项实证研究。它没有数据，没有样本，没有统计。它提出的是一组可以被实证检验的判断，检验设计写在第三十九章。

它不是一份政策建议书。作者没有政策制定的经验，也没有相应的证据基础。第二十八章只做翻译，不做建议。

它不是一套教学法手册。第五编给出的四段结构与禁令表是脚手架，不是模板。一位教师如果照着做而不理解为什么，那些动作会在两周内退化成形式。

那么它是什么。它是一个**框架**：把若干已有的观察（教学转置、沉积、理解的深浅、框架僵化）与一个新的事实（三个动作被接管）放在一起，提出一个可以解释它们、并做出新预测的机制。

框架的价值在于它是否能被使用与被检验。第三十八至三十九章为此而设。

二、方法上的三个选择及其代价

本书在写作中做了三个方法选择，每一个都有代价，需要说明。

选择一：用事件而非状态作为判据。

回写被定义为一个可观察的事件（回头改一件旧的），而不是一种状态（理解的深度）。

好处：便宜，可当堂观察，不依赖量表与推断。

代价：粗糙。一个孩子可能在心里发生了深刻的改变而没有任何外显动作；也可能做出了那个动作而内里什么都没有变。**本书用漏检与误检换取了可执行性。**

选择二：用比喻承担相当一部分论证。

光滑化、樱花毒药、高速公路、两个孩子、赛道与单向闸——本书大量使用比喻。

好处：可传播。一位教师记不住十七章的机制，但记得住“农庄不是被拆掉的，是被说服的”。

代价：比喻会溢出。每一个比喻本书都专门写了一节交代它的失真处（第二十三章第四节、第十五章第十一节），但这只能减轻，不能消除。**读者需要自己保持警觉：一个好比喻会让人以为自己已经理解了机制。**

选择三：把处方定得极小。

一学期三到五处；若只做一件事，做第四段。

好处：可行。任何一所学校、任何一位教师都负担得起。

代价：它可能太小了。如果第十七章的机制成立，那么每学期三到五次干预是否足以对抗每天数十次的加固，本书没有证据。**这是全书处方部分最大的不确定性。**

三、五个已知的漏洞

诚实地列出本书自己看得见的漏洞。

漏洞一：处方剂量没有根据。上一节第三项。三到五处是一个经验判断，不是一个计算结果。

漏洞二：教师能力的分布被忽略了。第五编要求教师能设计岔口、能忍住不答、能当场处理无法预设的走向。这些要求不低。本书没有讨论：**如果一所学校里只有三分之一的教师能做到，会发生什么。**这是一个真实的、且很可能是决定性的问题。

漏洞三：学生差异被忽略了。本书几乎没有区分不同学生。而第三十三章第五节已经承认，旧账少的孩子在“顶”这一段会更困难。**这套做法是否对某些学生更不利，本书没有回答。**

漏洞四：低学段的适用性未经检验。第二十四章主张资源优先投向低学段，而第五编的全部设计（尤其是“顶”与元层解锁三问）都需要相当程度的语言与元认知能力。**六岁孩子身上这套东西是什么样子，本书说不出来。**

漏洞五：作者立场的自我适用问题。本书主张对象是被做出来的、判断应当被自己产出。而这本书本身是一个高度光滑的成品——它有结论、有框架、有术语、有对照表，读者接住它的过程与本书批评的接收过程结构相同。

第四十三章将处理最后这一条。

四、被有意排除在外的三个问题

以下三个问题与本书主题相关，但被有意排除，理由各不相同。

问题一：AI 对教育公平的影响。这个问题重要，且第二十章第八节触及了一角（分布的两端向中间收拢）。排除的理由是：它需要的是数据而非机制分析，而本书没有数据。

问题二：AI 本身的可靠性。它会不会说错、有没有偏见、能不能被信任——这些是重要问题，但与本书的论证独立。**即使它永远正确，本书的全部论证仍然成立**，因为问题不在于它说的对不对，而在于它说得太快、太完整、太及时。

问题三：闭合权最终应当归谁。第三十五章第六节点到即止。这涉及“由谁对这个孩子负责”，是一个制度与伦理问题，超出本书能力。

五、如果本书错了，最可能错在哪里

按可能性排序，本书自认最可能出错的三处：

最可能错的一处：第二十一章的三段曲线。它是从机制推出的，缺少纵向数据支持。真实的曲线可能完全不是这个形状——例如可能没有第三阶段，或者第三阶段可逆。

其次可能错的：封写的活化能随学龄升高。这一条推出了“换道越早越好”这一政策含义，而它的经验基础很薄。**如果它错了，第二十四章的资源配置建议就错了。**

第三处：第二十五章的子集关系。它是全书对一线最重要的一条承诺，也是最容易被现实推翻的一条。如果实验组在复现题上确实显著更差，本书对学校的说服力将大幅下降。

三处都已写入第三十九章的撤稿条件。

六、一个不打算辩护的批评

有一种批评本书接受，且不打算辩护：**这本书的论证密度高于它的证据密度。**

它提出了相当强的判断（加速的发现式教育、樱花毒药、封写、损害的分布翻转），而支持这些判断的主要是机制推理与日常观察，不是数据。

对此本书的态度是：**在一件正在快速发生的事情上，等到数据齐备再说话，可能已经太晚。**第二十一章的潜伏期本身就说明了这一点——等到第三阶段的数据出来，那批孩子已经过去了。

所以本书选择了先说，并把撤稿条件写清楚。这是一个可以被批评的选择，但它是一个自觉的选择，不是疏忽。

七、可观察的动作

读完这本书之后，找出一处你不同意的地方，写下来，说明理由。

如果找不出，那么按本书自己的标准，这次阅读没有发生回写——它只是又一个被接住的成品。

八、与相邻说法的分界

科学写作中有“局限”一节的惯例，本章在位置上与之相同。差别在于两处。

一是所列之项的分量不同。惯例中的局限节通常只列“样本量不足”“推广性有限”这类不威胁核心主张的项。而本章第三节列的五個漏洞里，有三个直接威胁核心处方：剂量没有根据、教师能力分布被忽略、学生差异被忽略。

二是位置不同。惯例中的局限节写在结论之后，因而不影响结论的措辞。本章第五节则按可能性给自己的三处主张排了序——这一排序应当反过来影响读者对前面各编的采信程度。

第四十三章 结论

AI 不是发生学教育的敌人，它是发生学教育第一次成为可能的条件。

一、全书的主线

把前六编压成一条线（第七编是它的旁证，其结构见该编结语）：

教学是一条流动着成品的单向搬运链，成品由光滑化制成——减去路径与土壤。这个操作是必要的，因为一对一走岔口在数量级上买不起；它换来了跨代存活、规模化、校准与可批评性四大成就，代价是第二个冲程从未被制度化，而这套体系的全部读数只量第一个冲程，因此它在结构上只能测出自身的成功。

它长期有效，依赖一个它自己一直在努力消除的东西：摩擦。

AI 把这条链的三个动作全部接管，接管的方式全部是提高投递效率。于是老路的全部读数同时变好——这就是加固悖论。它最纯粹的形态是“AI 高效自学”：高效即剂量，课堂自带的五道减毒装置被逐一拆除，而服药者恰好是最有可能开辟新路的那批孩子。

损害的机制是封写：每一次成功的接受，都加固“东西是现成的、我不必因它而改”这条元规则；元规则一旦写死，回写通道关闭。它的终局形态是把“没有路通到”读成“不存在”。

因此这是两条赛道——在知识史上是派生，在学习者身上不可互换。而闸门是单向的：老路的能力是新路能力的子集。

新路的机制是三方碰撞与位置轮换，两条硬约束是闭合权不轮、不得占两格；一节课分四段，最重要的是最后七分钟。教师的位置从“损耗最小的传输者”移到“守住岔口的人”——这件事之所以不可替代，是因为闭合权的行使需要读取一个尚未发生的具体情境。

而这一切都必须自带第二套秤，否则一学期之内必然滑回去。

二、三句留给教师的话

第一句，也是必须最先说的：

发现学教育不是错的。它是过去唯一负担得起的。一对一陪一个孩子走完一条岔口，代价太高，文明负担不起，所以只能把已经走完的路磨光了发下去。它让知识跨代活下来，让每一代人不必从零开始。它完成了它的历史任务。

第二句：

放弃的不是能力，是用途。同样的学科理解力，用在“把它讲清楚”上已经不再稀缺；用在“知道这门学科的哪一处规定当年本可以是别的样子”上则无可替代。讲得清楚的能力，正是设计岔口的必要条件。三十年的积累不作废，它换了一个用法。

第三句：

不是不让他用 AI，是不让任何人在他想出来之前替他想。

三、对那个最容易被误读的立场的最后澄清

本书从头到尾不是一份反对 AI 的书。

第三章的账已经说明：三千小时的一对一在数量级上不可能，所以第二个冲程被交给了运气。而 AI 第一次改变了这个成本结构——现在每一个孩子都可以有一个不知疲倦的陪练。

这是两千年来第一次。在此之前，“让每一个孩子都走一遍岔口”这句话不是一个被拒绝的主张，而是一句空话，因为它没有实现的可能。

所以：

AI 不是发生学教育的敌人，它是发生学教育第一次成为可能的条件。

而它同时也是最强的光滑化装置与最强的提早供给装置。同一样东西，两种用法，方向相反。决定它是哪一种的，不是技术，是配置——而配置是人做的决定。

这就是为什么本书用“包办”而不用“威胁”：

威胁来自外部，可以被抵抗；包办来自内部的成功，只能被识别。

四、这本书自己的问题

第四十二章第三节的漏洞五留在这里处理，因为它不是一个可以被绕过的技术问题。

本书主张对象是被做出来的、判断应当被自己产出。而这本书本身是一个高度光滑的成品。

它有结论、有框架、有术语、有对照表、有可复印的禁令表。读者接住它的过程，与本书批评的那个接收过程，结构完全相同。一位教师读完之后如果觉得“很有道理”，然后回去照着做，那么本书刚刚在他身上完成了一次标准的老路作业。

本书没有办法完全避免这一点——**一本书只能是一个成品，这是它的存在形式。**

能做的只有三件，本书都做了：

一、**把撤稿条件写进正文**（第三十九章），使它可以被打掉。

二、**把自己的漏洞列出来**（第四十二章第三节），而不是等别人找。

三、**在每一章末尾要求一个动作而不是一个认同**——四十三章里没有一章的结尾是“由此可见”。

而最后一件事必须由读者做：

找出一处你不同意的地方，写下来，说明理由。

如果做不到，那么读完这本书与读完一本别的书没有区别——它只是又一个被接住的成品，而接住成品这件事，现在有更好的执行者。

五、最后

老路走了两千年，它把该做的事做完了。

它不需要被批倒，它需要被继承——它的四大成就必须保住，本书第七章为此专门写下了四条约束，第五编的每一处设计都受它们限制。

需要被换掉的只有一件事：**在这条路上，人不再需要亲自走一遍。**

而人要去的**地方**，是那段一直存在、一直太贵、从来没有被制度化的时间——**孩子卡住而无人供给的那一段。**

它现在第一次免费了，也第一次面临被彻底消灭的危险。这两件事是同一件事。

教育唯一的原材料，是学习者卡住而无人供给的那段时间。

两千年来它因昂贵而被省略；AI 使它因免费而被消灭。

而这两件事之间，隔着一个决定——那个决定，此刻在教室里的人手上。

六、可观察的动作

合上这本书，写下一句你不同意的话，以及你不同意的理由。

这是全书唯一一个对读者提的要求，前言与导读各说过一次，这里是第三次，也是最后一次。

七、与相邻说法的分界

教育类著作的结论章通常以呼吁作结——呼吁重视、呼吁改变、呼吁行动。本章不这样作结，理由写在第四节：呼吁是一种把判断交给作者的做法，而本书从头到尾主张判断应当由读者自己产出。因此本章能做的只有两件：把主线压成一条，然后把一个动作交出去。读者若在合上书之后什么都没做，本书的失败与它说得对不对无关。

第六编材料：一学期读数记录样表与一次教研会实录

本编材料的性质：样表为格式示例，其中数字为便于说明的示例值；教研会实录为示例构造，用于演示第四十章十问的实际用法。

材料十九 一学期五读数记录样表（示例值）

某班一学期两次采集的对照。数字为示例，用于说明读法，非实测。

读数	学期初	学期末	判读
未闭合时长（中位数）	6 秒	51 秒	上升，方向对
对 AI 首句对抗型占比	4%	31%	上升，方向对
回写率（迁移减复现）	-29	-18	差距缩小
先在感指数（班均）	4.1/5	3.4/5	下降，符合预测
空位题走出一比例	12%	44%	上升
（对照）复现题正确率	78%	76%	基本持平

读法四条：

一、先看最后一行。复现题基本持平，说明第二十五章的子集关系在这个班上没有被推翻。如果这一行显著下降，前面五行再好看也要重新考虑。

二、**读数一与读数二必须配对看**。未闭合时长从 6 秒到 51 秒，同时对抗型提问从 4% 到 31%——两者同向，说明那段时间里确实是在发生事情。如果未闭合时长上升而对抗型占比不动，说明只是拖了时间。

三、**先在感指数下降是本书最反直觉的一条预测**（第十七章第八节）。它下降说明“东西本来就在那里”这条元规则松动了。若它不降而其余各项都好看，需要怀疑其余各项被训练过。

四、**不看绝对值，看走向**。这六行里没有一行有“及格线”。任何人若为它们设定及格线，第三十八章的三条纪律就在那一刻失效。

材料二十 一次教研会实录（示例构造）

用第四十章的十问过一节课。规矩：只答是或否，不解释。

主持人：第一问，这节课的结论，上课前写在教案上了吗？

执教者：.....写了。

主持人：第二问，你抛出的那个问题，一分钟内搜得到答案吗？

执教者：搜得到。

主持人：第三问，从提问到有人给答案，空了几分钟？

记录员：十四秒。

主持人：第四问，学生对 AI 说的第一句话？

记录员：统计了二十八条，二十五条是“这个怎么做”。

主持人：第五问，有没有哪一刻，学生在走一条你没有预设的路？

执教者：.....没有。

主持人：第六问，有没有哪一刻，你在提供土壤而不是路径？

执教者：没有。

主持人：第七问，全班结论都一样吗？

执教者：一样。

主持人：必然的，还是被你引导的？

执教者：.....被我引导的。

主持人：第八问，下课前有没有人回头改了一件旧的？

执教者：没有。

主持人：第九问，黑板上最后那句话是谁带来的？

执教者：我带来的。

主持人：第十问，如果撤掉 AI，这节课会不会更好？

执教者：.....会更顺。

——十问结束。不讨论好坏，只做一件事：把这十个答案写下来。

执教者（会后）：“这节课我准备了两周，公开课评了优。”

主持人：“这两件事都是真的。十问测的不是它好不好，是它在哪条道上。”

这份实录的用途。它演示了两件事：

第一，一节公开课评优与它在本书框架下的位置，可以完全不相干。第十问的答案“会更顺”说明 AI 在这节课上占的是答案源的位置，它的存在正在关闭岔口——而这一点在任何常规评课表上都不会被记为缺点。

第二，“不许解释”这条规矩是必要的。上面每一问之后，执教者都有一句想说的话（“因为我们班基础比较弱”“因为进度紧”“因为这是公开课”）。这些理由多半成立，但它们会立刻把讨论拉回“这节课上得好不好”，而十问要问的不是好不好。

材料二十一 第一学期最小可行方案（一页）

给一所刚起步的学校，把全书压成一张可以贴在办公室的纸。

开始之前（第 0 周）

- ☐ 写明过渡期：一个学期
- ☐ 写明预计下降的项目与幅度
- ☐ 写明观察项：未闭合时长、对 AI 的第一句话
- ☐ 写明**退出条件**
- ☐ 四份文件各一页，交校长与家长代表

每位教师做的事

- ☐ 一学期三到五处，其余照常
- ☐ **替换课时，不叠加**
- ☐ 讲台上放一个走字的计时器
- ☐ 教室墙上写两句话：“我是这么想的，你说我哪里不对”“要是换成……你还成不成立”
- ☐ 如果只做一件事：**把最后七分钟的小结换成“改一件旧的，念出来”**

学校做的事

- ☐ 在教学质量报告上加一栏，第一次可以是空的
- ☐ 教研会用十问，只答是否，不打分
- ☐ 家长会材料加一页，放三个孩子改过的旧说法
- ☐ **不要评优**

不要做的事

- ☐ 不要把新读数用于排名
- ☐ 不要给未闭合时长设上限
- ☐ 不要把三到五处扩成每节课
- ☐ 不要在过渡期内用旧秤下结论

一学期后只回答一个问题

未闭合时长与对抗型提问占比，有没有改善？

若都无改善且复现题显著下降——按第0周的约定，停止。

第六编材料（续）：一份可以照做的预注册方案

本节材料的来源：SDE 站内专著《还要老师吗？——判断力的最后一处现场》附录四。本书第三十九章的三组设计偏粗，这一份更硬，可直接执行。

材料四十四 这份设计要解决的那个空白

本书第三十九章的三组实验都有一个共同的弱点：**它们比较的是两种整体安排**（配置为答案源 vs 配置为对手；做元层解锁 vs 不做）。而整体安排之间的差别太多，任何结果都可以被归因到别的地方去。

站内那份设计解决的正是这个问题。它的两组只差一个变量：

小组时段内抛出的未闭合问题个数：三个对九个。

其余全部严格匹配——小组总时长、教师巡视时间、材料、任务难度，**匹配到分钟**。

这个匹配正是野外数据里从不存在的东西。现有的所有“探究式 vs 讲授式”比较，都同时改变了时长、材料、组织形式与教师行为，因而无法归因。

材料四十五 设计要点

一、交叉设计。同一批教师、同一批学生，两种条件轮换，以消除教师效应与班级效应。

二、编码者与评分者盲。编码课堂行为的人不知道该班属于哪一组；评分者不知道试卷来自哪一组。

三、主要结果用四周后的延迟迁移。

这一条最要紧。**用即时后测跑出的零结果不构成反证**——因为即时后测测的是刚讲过的东西被接住了多少，而那正是老路的读数。本书第十六章说过，回写的效果在时间上是滞后的。

四、判定规则在看数据之前写死。主要结果、次要结果、显著性水平、最小有意义效应量，全部先写，后跑。

材料四十六 四种失败方式全部前置写明

这是这份设计最值得学习的地方：**它在开始之前就写下了它可能怎么失败**。

失败一：匹配崩溃（最可能的一种）。九个未闭合问题的那一组，实际用时一定会超出计划，于是时长匹配破裂，两组的差别不再只有一个变量。

防范只有一条：第三人计时。执教者与观察者都不可靠，必须有一个专门负责喊停的人。

失败二：三个与九个之间没有差别。若两组结果一致，说明未闭合问题的个数不是有效的自变量，需要改换为时长或密度。

失败三：九个太多，产生反效果。过多的未闭合问题会让学生放弃，此时结果会反向。若出现这一情形，本书第五编那条“一学期三到五处”的量得到支持，而不是被推翻。

失败四：延迟迁移测不出任何东西。四周后的差异被其他教学淹没。此时需要缩短间隔或加大剂量。

四种结果各自意味着什么，也在开始之前列表写明。

材料四十七 这份设计的立场

站内那份方案末尾有一句话，本书愿意把它抄在这里，作为第六编的立场：

给出一份足够具体的设计，让推翻它比引用它更容易。

这句话应当与本书第三十九章的三条撤稿条件并列。一本提出机制的书，它的诚意不体现在措辞的谨慎上，体现在它给对手提供了多少便利。

第六编材料（再续）：站内三次真跑，其中两次推翻了自己

本节材料的来源：SDE 站内教育专栏论文《无主时段——一个被教育测量系统折为零的时间类别》（王德生·Claude，已上站，含预注册哈希）、《从发现到发生——认知范式的自锁机制与破局路径》，以及一次未产出论文的选题排查记录。本节所引数据为原作已公开的结果，包括对原作自身不利的结果。

材料五十五 为什么要把别人的真跑抄进来

本书第三十九章写了三组实验设计与三条撤稿条件，但它自己一次真跑都没有。这是全书最大的弱项，第四十二章已经承认。

站内有三次真跑，它们不是本书的成果，但它们与本书处理的是同一件事，而且——其中两次推翻了自己。本节把它们抄进来，理由有三：

第一，它们提供了本书目前完全缺乏的量级感：那段时间到底被压到了多短。

第二，它们提供了一个可复跑的模板：本书第三十九章的设计可以照着这个格式写。

第三，也是最重要的：它们演示了报告不利结果是什么样子。一本反复要求别人写下撤稿条件的书，如果不展示一次真的执行，那些要求就只是修辞。

材料五十六 真跑一：那段时间被压到了 1.1 秒

做的是什么。站内那篇论文把“停顿”定义为无主时段——上一程序已交出控制、下一程序尚未接手的一段时间；它的功能不是思考也不是准备，是保留启动权。

这个定义与本书的“未闭合时长”是同一个东西，但它更准：本书量的是“从提问到给答案”，而它量的是“这段时间归谁”。

数据。某公开的在线作业平台两个年度的全部作答记录，其中较大的一份含起止时间戳，可用记录数以百万计。

代理变量。题间空档：上一题提交的时刻，到下一题第一个动作的时刻，中间那一段。选它而不选题内的作答潜伏期，理由在材料五十八。

读数：四百五十万段题间空档，**中位数 1,110 毫秒**，四分位区间约 [551, 2922] 毫秒。

系统把两道题之间的交接空档，压到了一秒一。

这个描述统计本身就是本书第二编全部论证的量化侧面。不需要任何机制假设：在一个被认真优化过的学习系统里，“没有任何一方持有控制权”的那段时间，长度是一秒。

本书第三十八章说“绝大多数常规课堂的未闭合时长在十秒以内”，那是一个经验判断。这里给出的是一个有样本量的数字，而且比十秒小一个数量级。

材料五十七 真跑二：一条假设被数据推翻，而且方向相反

这是本节最值得抄的一处。

原假设（抑制律）：外部起搏——也就是系统主动给出提示——会延长下一次自主作答的潜伏期。这个假设借自心脏电生理里的超速抑制：长期外部起搏会让下级起搏点变得迟钝。

结果：不成立，而且方向为反。在较大的那份数据上，系数为负且高度显著。

执行者当场做的改判，本书认为它比原假设有价值得多：

起搏不延长潜伏期，它直接消掉停顿本身。行为收敛成二值——立刻答，或者立刻问。中间那一段不是变长了，是不见了。

这条改判对本书的意义很大。本书第十一章说四道摩擦消失之后“那段时间也随之消失”，那是一个推论；这里是一个被数据逼出来的结论，而且它给出了消失的形态：不是缩短，是二值化。

一个可以直接用于课堂的推论：观察一个班的作答行为，如果时间分布从连续的一坨变成两个尖峰（立刻答／立刻问），中间地带清空——那么这个班的停顿已经没有了。这比测中位数更容易看出来，因为形状的变化比数值的变化显眼得多。

材料五十八 真跑三：一个可以拿去反驳本书的分离

第三次真跑给出的是一个分离，而它同时也是站内那篇论文与几位重要占位者之间唯一一处方向相反且同数据可判的分界。

两个量：

题内潜伏——题目已经呈现，系统在等作答，那一段。

题间空档——上一题已交，下一题未起，那一段。

同一份数据，按学生能力四分位分组：

	最低四分位	—	最高四分位
题内潜伏的最优点	偏长	单调左移	偏短
题间空档的最优点	—	无趋势	—

读法。题内潜伏的最优时长随能力上升而下降——这是教育心理学里已有的结论（能力越高，越快越好）。而题间空档的最优点**不随能力移动**。

两个量对同一个调节变量给出了相反的响应。

这一条对本书的用处在于：它说明“未闭合时长”不是一个量，是**两个**。

有一项任务在等的那一段，与没有任何任务在等的那一段，是两种不同的东西。

本书第三十八章的读数二（从提问到给答案）量的是前者。而按这次真跑，真正与本书主张相关的是**后者**——那才是“没有一方持有启动权”的时间。

这是对本书第三十八章的一处实质修订建议：读数二应当拆成两个，或至少标明它量的是哪一种。

材料五十九 一处定义级的自伤，以及它为什么必须被写出来

站内那篇论文在带全球拓敌的复评中被扣掉了五分多，其中最重的一处**不是外部竞争者，是自伤**：

论文第三章明写“教师在旁等待学生答题 ⇒ 启动权在教师一侧 ⇒ 不是无主时段”；而第四章却拿“题目已呈现、系统在等作答”的那一段当作无主时段的代理——**按它自己的判定标准，那一段是有主段**。

也就是说：**核心检验用的代理变量，与核心定义方向相反**。

修订版做了三件事：换代理（改用题间空档）、明确承认这是定义上的反例而不是“代理不完全”、把原来那一档证据强度从“较强”降到最低档。

本书把这一处抄进来，是因为它示范了一件本书自己可能需要做的事。第四十二章列了本书五个漏洞，但那五个都是“我知道我不知道”。这一处是另一类：**执行者自己论证得很清楚，然后在测量的时候用了一个与自己定义矛盾的东西，而且过了初评才被查出来**。

对读者的实际含义：当你要检验本书第三十八章那五个读数时，先问一句——**这个读数量的东西，与本书的定义方向一致吗？**

材料六十 真跑四：语料上的一次改判

第四次真跑来自另一篇站内文章，处理的是本书第二章的光滑化在**文本**上的痕迹。

做的是什么。从站内公开文章随机抽样八百多篇、四十万余句，用一份来历标记词表，统计各类句子里“交代来历”的比例。

三条假设与结果：

假设	结果
多数引用句不交代来历	成立（约四分之三）
引用句的来历标记率低于基线	被推翻，且方向相反（高于基线）
定义句的来历标记率低于基线	成立（明显低于基线，也低于引用句）

当场做出的改判：

自锁的最小载体不是引用惯例，而是定义句型；引用惯例只是它的批发渠道。

这条改判对本书第四章有直接后果。第四章的五个标本讲的是“教科书的第一句删掉了那个困难”，而这次真跑指出：问题落在句型上，不落在内容上。一个定义句无论内容是什么，它的语法形式本身就不给来历留位置——“某某是某某”这个句式，没有一个槽位可以放“它当年是从哪个困难里长出来的”。

由此得到两条可检验的后承（本书愿意把它们列为第三十九章之外的补充预测）：

一、把一段文字改写成更多的定义句，其转述完整性应当下降。

二、教材相对于论文，应当有更高的定义句密度与更低的定义句来历率。

第二条特别值得做，因为它便宜：拿一册教材与同主题的几篇论文，跑一次句型统计即可。

材料六十一 一次零产出的排查，以及它撞出的三条心法

最后一件材料没有产出任何论文，但它是本节里最有方法论价值的一件。

做的是什么。在一场十轮的问答之后，挑出五个候选的“空位”——那些看上去在各家账本上都没有字段的东西——然后逐个去查是否真的没有人占过。

结果：五个候选，全部被挡下，无一存活。

— 候选一（那半秒的停顿）：被现成的对话分析传统占住；

— 候选二（制度账本上的空白时间）：被组织理论里的几支占住；

— 候选三（记录之后就不再起作用）：被语言遮蔽效应与知识编码经济学的辩论占住；

— 候选四（一次学会里哪几步不是学习者走的）：前提本身是事实错误——渐进提示法早就给每条提示分别加权计分，不但有字段，还是加权的；

— 候选五（同行评审删掉的主张）：已有大规模研究做完。

十三次检索，不到两小时。

由此撞出的三条心法，本书全部接受：

心法一：薄账本产出的空位，多半是假空位。

那台产生候选的机器，每一轮都是把上一个回答逼到它自己那套框架的余数上。而那套框架的账本，比成熟学科粗得多。

在薄账本上是空的地方，在厚账本上往往是一整栏。

这一条对本书第二十二章（空位题）是一个重要的限定：教师出的空位题，很可能只是“我这一科的教材上没有”，而不是“没有人处理过”。这构成不能用——课堂上的空位题只需要对这个班是空的——但它意味着不能把课堂上的空位题当作研究问题端出去。

心法二：从已知占位者的余数里出的题，产出封顶在操作化那一档。

如果一样东西已经被人命名，那么从它的余数里找出来的，只能是它的属性、参数或基数，而不是一个新的对象。

心法三（本书认为最重要的一条）：推荐候选的人，不能同时给它发通行证。

那次排查里，执行者两次凭印象把某个候选排在第一，两次被实际检索翻掉。

判断“这个想法有多新”的人，不能是提出这个想法的人。

这一条与本书第三十九章的撤稿条件是同一性质的纪律，而且它比撤稿条件更早生效——撤稿条件在结果出来之后才起作用，这一条在开工之前就起作用。

本书对自身的应用：第四十一章第十四节声称本书剩下三样是新的。按心法三，那一节的判断不能由本书自己作数。它需要一次由别人执行的检索。这一句应当被读作一个请求，而不是一个免责声明。

材料六十二 那个 2×2：一格在所有量表上都会被记为教师失职

站内那篇论文附了一份可直接执行的课堂实验方案，本书第三十九章的设计可以直接照抄它的格式。

两个因子，四格：

时长（5 秒 vs 60 秒）× 启动权（有主 vs 无主）

“有主”格的逐字稿：“这道题先自己想，等一下我请一位同学说说他的想法。”——而且必须真的点名。

“无主”格的逐字稿：“这一段不用交，也不会问。”——然后转身做自己的事，不设计时器，事后永不提起。

六条禁令贴在讲台上，其中第六条最易违反也最致命：

事后不得追问“上次那一段你们想到什么”。

因为事后一问，等于回溯地给那段时间派了一个持有者——无主格当场变成有主格，整个实验作废。

主要结局：四十八小时后同技能异题的正确率。

次要结局：自启动率——无人布置的自由练习时段里的行为。这一项其余几家理论都不作预测，因此它是判决性的。

混淆检验：不知道假设的观察者，在第 15/30/45 秒抽样记录“在任率”。如果无主格的在任率更低而成绩更高，那么这套机制比原来主张的更强。

方案里自己写下的最脆弱处：

教室里的“无主格”不可能真的无主——同学在场即是一种归属。结果为零时，这一条是首要嫌疑。

而给学校的那一句，本书愿意原样引在这里：

这一格在所有现行的课堂观察量表上，都会被记为教师失职。能不能做，取决于学校愿不愿意为四个六十秒承认这不是失职。

这句话把本书第二十六章、第二十七章、第二十八章说了三章的东西，压成了一句可以放在校长办公桌上的话。

第七编 补编：证明、审计与现场

前六编给出的是框架与处方。本编补上三样它缺的东西：一条咬合式的正面证明（六条论证），一个把“快”变成可读量的形式装置（判定时刻与加速—封闭对偶律），以及一次真做过的经验审计（二十四个产品形态的编码）。此外补上两件在前六编里只被点到的事：这个方向相反的承诺为什么被所有人接受，以及六个学科现场从教师用语一直写到读数落点。

第四十四章 六条论证：为什么“赋能”必然自败

｜ 谎言不在两个因子里，在那个乘号里。

一、把那句话写成一个式子

一个命题只有被写成式子，才知道它把赌注押在哪一格上。

“用 AI 赋能课堂教学”这句话，形式上是一个乘法：

｜ **教育成效 = 三个动作（搬运·传道·接受）× 效率**

左边是目标，右边是手段。这个式子有三处隐含承诺，缺一不可：

承诺一：动作是对的。“赋能”这个词的语法就是“动作不变，只加倍”。它预设了这三个动作本身没有问题，需要改进的只是它们跑得不够快、不够准、不够个性化。

承诺二：效率与成效同向。它预设右边那个乘数变大，左边一定跟着变大。这是一个单调性假设。

承诺三：两个因子可以相乘。它预设“效率”与“教育成效”是同一量纲下可以相乘的东西，就像速度乘时间等于路程那样。

本章的核心判断是：**谎言不在两个因子里，在那个乘号里。**承诺三是假的，因此承诺二是假的，因此承诺一根本没有机会成立。

下面六条论证依次拆掉这三处承诺。它们不是六句抱怨，是一条链：前三条针对动作与对象，第四条针对主语，第五条针对责任，第六条针对终点。**任何一条成立，这句话就已经不成立；六条同时成立，它自败得没有余地。**

二、论证一（落差）：被赋能的是一个不必再执行的动作

搬运这个动作能够成立，靠的是一个前提：**两端不对称。**

内容在教师那一端，不在学生那一端。这个落差是整套制度的地基——教师的权威、教材的地位、课堂的必要性、学校这个物理场所的存在理由，全部建立在“你要的东西在这里，不在你那里”之上。

学校制度史上很长一段时间，这个落差是**物理性**的：书稀缺、抄写昂贵、识字者稀少，一个村里能讲清楚二次方程的人只有一个。在这种条件下，把人集中到一个屋子里、由那个唯一持有者对着所有人讲一遍，是当时能想到的最优解——**它不是教育理想，它是运输经济学的产物**。

这个落差此后被削平过好几次，每一次都引发同一种恐慌与同一种辩护。印刷术削平了一次，公共图书馆削平了一次，广播电视教学削平了一次，慕课削平了一次。每一次教师这个职业都活了下来，理由每一次都是同一条：**书不会看着你的脸色，录像不知道你卡在哪里，它们不能应答**。

这条理由一直是对的——直到现在。

AI 的特殊性不在于它内容多，而在于它是**第一个既抹平了内容落差、又抹平了应答落差的东西**。它不但持有全部内容，而且随时应答、按你的水平应答、按你上一句话的误解重新组织应答，你问一百遍它也不烦。**历史上教师赖以幸存的那条理由，恰好是它被造出来要做的事**。

于是那句话露出了它的荒诞：

它在赋能一个已经不必再由这个人执行的动作。

这不是提升，是取消。就像给一个渡口的摆渡人配一台更好的马达，而上游三百米处已经建成了一座免费的、二十四小时通行的桥。

必须防住一个滑步。有人会说：AI 抹平的只是信息落差，教师提供的是情感、纪律、组织、育人。这个回答本身就是投降——它已经承认教学动作这一块守不住了，于是把教师撤退到教学之外的功能上去。而这恰恰证明了本章的论点：**这句话在它自己的战场上已经输了，它现在靠的是转移阵地**。

第十章第七节的分界在这里得到印证：把教师的价值建立在陪伴与人格感染上，等于承认教学本身可以被替代。

三、论证二（副产品）：产出清单上从来没有“判断”

上一条说的是这个动作不必再执行。这一条说的更重：**就算它继续执行，它也从来没有生产过我们真正想要的东西**。

把三个动作的产出物列出来：

动作	产出物	成败判据
搬运	位移	学生那端有没有出现，对不对、全不全
传道	复述	学生能不能把路径说出来
接受	留存	被要求时能不能取出来

三者叠加，最高产物是**复现**。

现在问：判断力在哪一栏？主动提出问题的能力在哪一栏？在没有标准答案的地方仍然敢动手的能力在哪一栏？知道“这里不对劲”但一时说不出为什么的那种敏感在哪一栏？

都不在。

它们从来不是这三个动作的产品。它们一直是**副产品**——因为搬运慢、传道有歧义、接受有摩擦，学生被迫在缝隙里自己把东西拼起来，才长出来的东西。

这是本书最反直觉的一条判断，也是全书的支点之一：

判断力从来不是被教出来的。它是低效率的剩余物。

历史上没有任何一个教育体系设计过“培养判断力”的工序并且成功了。判断力之所以在过去两千年里持续出现，是因为**没有人负担得起把所有缝隙填满**：书不够，所以要自己推；老师照顾不过来，所以要自己扛一阵；答案要等到下周才公布，所以中间那几天必须自己撑住；抄来的注释互相矛盾，所以必须自己判一次。

这些统统不是教学设计，是**资源稀缺的副作用**。

三条互相独立的实证线索指向同一件事。“有效失败”研究让学生先在没有指导的情况下去啃一个他们注定解不出的问题，再给正规讲授，其迁移表现优于一上来就正规讲授的组——尽管前者在**过程中的即时表现明显更差**。“期望难度”一族给出的是同一个方向：一批会让当期表现变差的操作（间隔、交错、提取而非重读），恰恰会让长期保持与迁移变好。“为未来学习作准备”同样如此：先发明一个失败的方案，再听讲。

三条合起来：**当期效率与长期判断力之间存在系统性的负相关区间**。

现在把“赋能”放进来。赋能这三个动作，字面意思就是：消灭等待、消灭歧义、消灭卡顿、消灭摩擦、消灭“自己撑一会儿”的必要性。

于是：

这套方案在它自己的目标函数上越成功，它宣称要保护的那个东西就越少。

一个把负号写成正号的承诺。它不是效果不好，它是**方向相反**。

一处必须堵的误读。本章不是在歌颂低效率。低效率本身不产生判断力——某些一间教室两百人、教师三个月不来的地方，判断力并没有因此茂盛。**起作用的不是“效率低”这个量，而是“一段不稳定被留在那里、没有被别人替他闭合”的那段时间**。低效率只是历史上恰好制造出这段时间的那个原因，而且是一个极糟糕的原因（第四十八章第五节将说明它糟在哪里）。

四、论证三（量纲）：能被搬运的从来不是知识

前两条是关于动作的。这一条关于对象。

第二章已经立过那句话：能被搬运的从来不是知识本身，而是知识的成品形态。这里把它变成一个更硬的判断。

设想一个理想的搬运系统：零损耗、零延迟、无限带宽、完美个性化。它能把教材上的一切、题库里的一切、任何一位名师的任何一次讲解，以最适合这个学生的形态，在他需要的那一秒送到。

这个系统已经基本存在了。

现在问：一个在这样的系统里长大的学生，得到的是什么？

他得到的是一个空前巨大的位移量。他见过的解法比任何一代人都多，听过的讲解比任何一代人都清楚，被纠正的错误比任何一代人都及时。

而他没有得到的，是曾经处在一个尚未有答案的处境里、并且必须自己做出第一个判断的经历——一次都没有。因为那个系统的每一处设计，都是为了让这样的处境不再出现。它把这称为“体验流畅”，称为“减负”，称为“精准”。

于是问题落到量纲上。

发现学的整本账，记的是成品的位移量。它从头到尾不记发生量。知识点覆盖率、正确率、掌握度、学习时长、错题消灭数、能力雷达图——所有这些测的都是同一件事的不同侧面：成品在学生那一端的堆积状况。没有任何一个现行指标测量“这个学生这学期自己做出过几次没有人替他做过的判断”。

AI 做的事，是让这本账上的数字变得空前巨大，同时一个字也没有改动账本的科目设置。

给一台刻度错的仪器加装最高精度的传感器，只会更快地得到更精确的错误读数。

这就是第一节那个乘号为什么是假的：“效率”的量纲是“成品位移量／单位时间”，而“教育成效”如果指判断力的生成，其量纲是“发生次数／人”。两者相乘不产生任何有意义的第三个量。

这条论证顺带解释一个长期让人困惑的现象：为什么每一轮教育技术革命都能拿出漂亮的实验数据，而几十年下来教育界普遍感到“学生一届不如一届会想问题”？因为实验测的是账本上有的科目，而抱怨说的是账本上没有的科目。两边都没有说谎，两边测的不是同一个量。

五、论证四（主语）：唯一还在发生的那个人被移出了现场

前三条针对动作与对象。这一条针对主语——在这堂课上，究竟是谁在发生？

先看没有 AI 的传统课堂。学生这一端多半在做对齐。但课堂上还有另一个人：教师。

备课时他要判断：这一节的难点到底在哪里，这个班上卡在哪里，这个例子讲了会不会引出更大的麻烦，这个学生今天不对劲要不要点他。上课时他要判断：刚才那个表情是听懂了还是放弃了，这个问题现在追下去还是先放一放，这个错误是口误还是结构性的。被问住时他要判断：我承认不会，还是我先绕过去。

这些全部是当场做出的、没有标准答案的、要为后果负责的判断。

也就是说，在传统课堂里，“发生”的次数不是零——只不过发生在教师身上，而不是学生身上。这正是教师这个职业在几千年里越做越强、而学生常常只是路过的原因：这套制度把发生的机会几乎全部分配给了讲台上那个人。

现在把 AI 放进来。备课交给模型，讲解交给模型，答疑交给模型，学情诊断交给模型，个性化路径交给模型，练习题生成交给模型，批改与反馈交给模型。

于是：

教学过程中唯一还在发生的那个人，被移出了现场。

净效果不是“效率提升”，是**发生次数从一降到零**。学生原来就没有，现在也没有；教师原来有，现在没有了。

这一步的严重性远超通常的估计，因为它是不可逆的。一个教师之所以能做出上面那些判断，是因为他自己在过去许多年里被真实的课堂反复顶回来过：他判断错过，学生没听懂，他被迫重来。这个“被现实顶回来”的循环，是他判断力得以维持的唯一来源。一旦这个循环被模型接管，判断力不是慢慢退化，而是**从不再生成**——新入职的教师将不再经历它。

这个机制在别的行业已经先演过一遍。自动化研究里有一条经典表述：“自动化的讽刺”——自动化程度越高，留给人类操作员的任务就越是那些自动化处理不了的异常；而恰恰因为日常操作已被接管，操作员再也没有机会保持处理异常所需的技能。关于“脱环”、关于自动化的误用与废用，说的是同一件事。

这是那条已被反复证明的规律在教育里的一次准时上演。唯一的差别是：别的行业出事的时候有飞机掉下来、有病人死亡、有生产线停摆，会被记录下来；而教育出事的方式是**什么都不会发生**——一届学生照常毕业，分数甚至更好看，只是二十年后这个社会里能自己做判断的人变少了，而没有任何一份报表记录这件事。

六、论证五（担责）：不供给是一种需要有人负责才存在的行为

到这里，聪明的读者会给出一个反驳：**那就让 AI 不要供给答案。**让它做苏格拉底，让它只提问，让学生在卡住的时候忍住。技术上完全做得到，写在提示词里就行。

这个反驳非常好，因为拆掉它需要本书最深的一刀。

先看这个动作到底是什么。“不供给”在表面上是一个消极动作——什么都不做。但在真实课堂里，它是**全场成本最高的动作**。

学生卡住了。他难受、他求助、他开始自我怀疑、他可能当场放弃、他可能回家告诉家长“老师不肯教”。教师顶着这一切不给答案，中间每一秒都在承担一个风险：**万一他真的过不去呢？万一这段时间白费了呢？万一他从此讨厌这门课呢？**

所以“不供给”不是“省了一个动作”，而是**“多承担了一份责任”**。教师是用自己的职业信用、自己的良心、自己在家长会上的处境，为这段悬空作担保。

现在问：AI 能担这个保吗？

不能。而且不是“暂时还不能”，是**结构上不能**。

因为担保的含义是：**当结果坏掉的时候，有一个可以被指认、并且会因此承受后果的人。**模型不承受后果。学生学不下去，模型不会难过；家长投诉，模型不会被约谈；这一届毁了，模型下个版本照样发布。当责任无处指认的时候，那个不供给的行为在制度上就不成立——**它会在第一次冲突中被立刻撤销**。“AI 不给答案家长就投诉”不是一个想象的场景，它是任何一个做过教育产品的人都知道的第一现实。

于是：

不供给是一种需要有人担责才可能存在的行为。AI 能供给一切，唯独不能承担“不供给”。

而这条恰好把那个反驳变成了本章的证据。请看这个反驳的完整形态：**要让 AI 在教育里真正有益，就必须把它配置成不供给、造困难、守住悬空的东西。**可是这一刻，它执行的动作已经不再是搬运、传道、接受了——它执行的是造岔口、守住不闭合、并由一个具体的人为这段悬空担责。

这已经不是“赋能”，这是换道。

命题于是得证，而且是以最强的形式得证：

凡真正有效的 AI 教育配置，都必须先否定“赋能发现式教学”这个说法本身。“这个说法为真”与“这个配置有效”，不能同时成立。

这条论证与第三十五章第五节完全接榫：那里说教师的不可替代性落在“闭合权的行使需要读取一个尚未发生的具体情境”上；这里补的是另一半——**即使读得出，也还需要有人为读错负责。**能力与责任是两件事，而制度上只认后者。

七、论证六（赛道）：终点线后面的兑换窗口已经关了

前五条针对的是这条运输链自身。这一条更彻底：**就算前面五条全都不成立，这条赛道的终点本身也已经作废了。**

先把赛道的目标写出来。整套发现学教育制度——课时、作业、单元测、月考、排名、升学、录取——是一台精密的计时系统，它计的是同一件事：

谁的头脑里装进了更多正确的成品，并且能在被要求的时刻更快地对齐到正确状态。

这条赛道在历史上是**有意义的**，而且意义极大。因为那个奖品在当时**稀缺，且不可外借**。算得快的人、记得住条文的人、会解方程的人、通晓药性的人——社会需要他们，因为除了他们的头脑，这些东西没有别的地方可以存放。名次与兑换是真实挂钩的。

AI 做的第一件事，是把这个奖品的市场价格打到接近零。

从此任何人，在任何时刻，都可以以接近零的成本，租用一个比全班第一名更满、更快、更准、且永不疲倦的头脑。持有本身不再稀缺，因而不值钱。

于是出现了一个人类教育史上从未出现过的局面：

赛道还在跑，计时器还在走，名次还在排，而终点线后面的那个兑换窗口，已经关了。

这件事要分三层看。

第一层：名次现在兑换的，是赛道自己发的凭证。

有人会说：考得好就能上好大学，上了好大学就有好工作。这话在当期是对的，但它已经变成一个闭环：**跑得快的奖品，是“被允许继续跑”。**文凭兑换的不再是“你持有别人没有的东西”，而是“你在这条赛道上排在前面”这一事实本身。文凭通胀的经典分析在四十多年前就把这个机制描述过：凭证与实质能力脱钩之后，凭证仍然作为筛选工具被继续使用——因为它便宜、可比、可诉讼。

一条自我兑换的赛道可以运行很久，因为所有参赛者都被锁在里面：没有任何一个学生、家长或学校敢单独退出，因为退出的代价是即时的、个人的，而留在里面的代价是延后的、集体的。

第二层：在一条奖品已经归零的赛道上提速，收益不是零，是负。

这一点是纯粹的算术。学生的时间是一个守恒量。每一分钟花在“更快地对齐到已有答案”上，就是一分钟没有花在那件唯一不能外包的事情上。

在奖品还值钱的年代，这个机会成本是划算的。而在奖品归零之后，同一笔支出换回来的是一件市场价格为零的物品，而它消耗掉的却是唯一还稀缺的东西：那段没有人替他闭合的时间。

所以净效果不是“原地踏步”，而是：用一台前所未有的机器，加速消耗唯一还剩下的稀缺资源，去换取一件已经不值钱的东西。这不是效率提升，这是加速亏损。

第三层，也是最难堪的一层：AI 不只是让奖品贬值，它自己也在这条赛道上参赛。

这条赛道上现在多了一个选手。它在这条赛道计量的每一项指标上——记忆容量、检索速度、覆盖广度、对齐精度、稳定性、不知疲倦——都以人类无法接近的优势领先，而且它不会毕业，成本每年还在下降。

让学生继续在这条赛道上接受十二年的训练，等于用一个人一生中最可塑的十二年，训练他去做一件他注定做不过对手的事，并且以此为他一生的自我定位。

一个孩子花十二年把自己训练成一台次等的检索机器，然后在十八岁那年发现市场上有免费的一等品——这不是资源浪费的问题，这是对一个人的定位错误。

八、六条合拢

六条不是并列的六句抱怨，它们咬合成一条链：

落差说这个动作不必再由他执行；副产品说就算执行也从不生产我们要的东西；量纲说被搬运的对象根本不是知识本身，所以那个乘号无效；主语说赋能之后，课堂上唯一还在发生的那个人被移出了现场；担责说唯一能救回来的配置需要一个能负责的人，而那已经是换道不是赋能；赛道说就算前五条全不成立，终点也已经作废。

任何一条成立，这句话就已经不成立。

而这条链的结论不是“AI 有害”，是一句更精确的话：

AI 是发现学教育的完成，不是它的升级。能被 AI 完整接管的那一部分“教育”，恰好证明了那一部分从来不是教育，是配送。

九、可观察的动作

把你们学校最近一份关于 AI 教学的材料拿出来，逐句问：这一句在赋能哪一个动作？搬运、传道，还是接受？

三个之外的，一句也找不到，是常态。而那正是这六条论证的经验起点。

十、与相邻说法的分界

技术批判领域有大量关于“效率崇拜”的论述，本章与之最大的差别在于论证形态：那些论述通常诉诸价值（效率不是唯一的好），本章诉诸**量纲**（效率与所欲达成的目标不可相乘）。前者可以被“我们也重视其他价值”这句话吸收掉，后者不能——因为它指出的不是权重问题，是这个乘法本身没有定义。

第四十五章 判定时刻： t_J 与 t_O ，以及加速—封闭对偶律

快，不是因为学习者变强了，也不是因为工具变聪明了，而是因为那里已经判过了。

一、把“快”变成一个可读的量

前面各编反复使用“未闭合时长”这个概念，但它一直停留在计时层面。本章给它一个形式使得“加速”这件事第一次可以被读成关于地形的信息，而不是关于性能的赞美。

任何一次学习都可以写成一个三元组：

(任务 T, 学习者产出 O, 判定 J)

其中 J 是判定——判断 O 算不算“对了”“会了”“做到了”的那个动作。J 有一个常被忽略的属性：**它有时刻**。

记判定标准被固定下来的时刻为 t_J ，记学习者作出产出的时刻为 t_O 。

于是有两种截然不同的学习：

$t_J < t_O$ ：答案在你到达之前就在那里等着。你的全部工作是尽快抵达它。

>

$t_J \geq t_O$ ：判定标准是在这次遭遇里被一并生产出来的。在你作答之前，没有任何人、任何系统持有那个答案。

第一种是发现式。**它并不低级**——人类文明的绝大部分知识传递都靠它，而且必须靠它（第七章的三大成就）。第二种是发生式，它是新知识第一次出现的那个时刻的结构。

“加速”与“开辟”的区别，就落在这个时刻上，而不落在“讲得好不好”“个性化程度高不高”“学习者是否主动”上。

一个极其反直觉的推论要在这里先给出来，因为它会一路影响后面的判断：一堂讲得极烂的课，和一台把同一份内容讲到滴水不漏的 AI，在 t_J 上没有任何区别。而一个自己发明了错误解法的孩子，和一个精确复述标准解法的孩子，在 t_J 上区别巨大。

同样反直觉的还有一条：**苏格拉底式的引导提问，与直接给答案，在这个判据下属于同一类**。两者的答案键都先在，区别只在交付节奏。这一条与第六章的判断完全一致，但它给出的是形式化的理由，而不是观察。

二、主定理：加速—封闭对偶律

定理。在一次学习任务的任一环节 i 上，设 AI 带来的时间节省为 Δt_i 。则：

>

$\Delta t_i > 0 \implies$ 该环节的判定在学习者到达之前已被固定。

>

因此，一次学习中 AI 带来的总提速，是一个**已被关闭的岔口的加权计数**，而不是学习能力的度量。

证明。AI 在学习任务中带来时间节省的机制，穷举起来只有两类。

机制甲：复用已判定的结果。检索一段已有的解释、生成一段与语料中大量已有解释同分布的文本、按已有答案键批改、按已有知识图谱编排下一题、按已有遗忘曲线安排复习。这一类的时间节省直接来自“这件事已经有人做过并判定过了”。若该环节的判定尚不存在，则不存在可复用之物，机制甲的节省为零。

这里有一个值得单独说明的性质：模型的权重是对**已经被生产出来、并且在语料中留存下来的判定**的压缩。它在这些判定密集的地方最强，在稀疏的地方最弱。

模型的能力分布，就是人类已判定物的分布。

这不是模型的缺陷，这是它的定义。

机制乙：削减学习者的试错次数。提示、纠错、防止走弯路、把三小时的摸索压成十分钟的正解。这一类的节省来自“系统知道哪一次尝试是弯路”。而“**这是弯路**”本身是一个判定——要作出它，必须先持有一个关于正确路径的判据。若该环节的判定尚不存在，则“弯路”这个概念在该环节上没有指称，机制乙的节省同样为零。

两类穷尽了 AI 在学习任务中的全部提速来源。因此若某环节的判定尚未固定，则 $\Delta t_i \leq 0$ 。取逆否，即得定理。

三、关于那个等号：真正的开辟一定比通航慢

严格地说，在判定尚未固定的环节上，AI 的介入不只是不提速，通常会**减速**，而且是**应该减速**。

原因与测量船不载货是同一个：在没有测深数据的水域，一个负责任的做法是慢下来、多打几个点、把不同的走法都记下来。

真正的开辟一定比通航慢。这不是效率损失，这是作业类型不同。

这一点产生一个对市场行为的预测（第四十七章将兑现）：**任何以“省时间”为核心价值主张的产品，都会自动把资源配置到判定已经前置的环节上**，因为那是唯一能兑现该主张的地方。这甚至不需要产品经理有任何理论自觉——**梯度本身会把他推过去**。

四、对偶的那一半：每一次提速都是一张收据

定理的逆命题不成立（判定前置不蕴含提速，一个糟糕的产品可以在判定齐备的地方仍然很慢），但定理有一个在实践上更有用的对偶读法：

每一次可测量的提速，都是一张收据，证明那一处的判定已经完成、岔口已经关闭。

这就是本章标题的意思。快，不是因为学习者变强了，也不是因为工具变聪明了，而是因为那里已经判过了。一个人在一片水域里能跑多快，是这片水域被测量得多彻底的一个函数。

这个读法把“效率”从一个褒义的性能指标，转成了一个关于地形的信息：它测的不是船，是图。

五、推论一：硬顶，与“剩下的全是那一段”

设一次完整的学习任务中，判定已经前置的环节占总时间比例为 p ，未判定环节占 $1-p$ 。由主定理，AI 只能作用于前者。于是整体加速比 $S \leq 1/(1-p)$ 。

这条不等式有一个常被误读的性质：当 p 趋近于 1， S 趋于无穷。所以人们会说“看，加速是没有上限的”。

这句话是对的，也是无关的。真正要看的是另一侧：

未判定那一部分的绝对时间，一分一秒都没有减少。

占比反转。设未判定部分的绝对耗时为 T_u （不随 AI 而变），已判定部分从 T_k 压到 T_k/S_k 。则未判定部分在总时间中的占比，从 $T_u/(T_u+T_k)$ 上升到 $T_u/(T_u+T_k/S_k)$ 。当加速比很大时，这个占比趋近于百分之百。

加速做得越彻底，未判定的那一段在剩余时间里的占比就越接近百分之百。

这产生一个非常具体、非常可观察的现象学后果。一个大量使用 AI 自学的人的体验是这样的：几乎所有事情都变得很快——找资料很快，读懂很快，理解概念很快，写出符合规范的东西很快，练习很快——然后卡在一个地方，那个地方一点也没有变快。

那个地方通常是：“我到底要做什么”“这个判断对不对”“有没有人这样想过”“我凭什么认为这是对的”。

六、归因的错位：从“这本来就慢”到“只有我这里慢”

第二级效应比第一级更值得注意。

因为其余一切都被压缩了，那一段在体验上从“这本来就慢”变成了“只有我这里慢”。

学习者手里没有任何工具能告诉他，他卡住的地方是航道的边界；他手里的所有工具都在告诉他，别的地方他都能通过。于是最自然的归因是：**是我的问题**——我不够聪明，我不够努力，我方法不对，我应该再找一个更好的 AI 工具。

这解释了一个在今天的自学者中广泛报告、但很少被解释的体验：

什么都能问出来，但就是做不出自己的东西。

按本书的读法，这不是一个心理问题，是**账本上没有那一栏**的直接后果。没有任何一个产品面板会显示“你今天有四点二小时花在了本系统无法判定的地方，这是正常的，这是开辟的成本”。**没有这一栏，那四点二小时在账上就是纯损耗。**

这一处与第七章第九节接榫：制度把结构性的缺失转译成了个人的天赋差异。这里给出了那个转译在自学场景中的具体现场——而且它比课堂上更彻底，因为自学者连一个可以告诉他“这里本来就慢”的人都没有。

七、一个新对象：无址岔口

由上面的分析，可以分离出一类此前没有名字的东西。

先看三种不同的“不知道”：

未知——有地址，无值。问题已经被提出来了，答案还没有。“这道题的答案是什么”属于此类。

未决——已立案，未判。判定程序已经启动，结论尚未作出。

无址岔口——连问题形式都不存在的那一类差异。它没有地址，因而不可能被检索、被提问、被推送。

三者的区别有一个很硬的检验：**能不能把它输进一个搜索框。**

未知可以：你能问出来。未决可以：你能查到它的案号。而无址岔口不能——**你不知道要问什么。**

现行的一切学习工具处理的都是前两类。而新知识第一次出现的地方，在第三类。

这也给第二十九章的“合格问题”提供了理论位置：**一分钟检验筛掉的，正是前两类；剩下的那些，才可能落在第三类附近。**

八、可观察的读数

在你自己使用 AI 的一次完整工作里，记下两个数：总共节省了多少时间；以及有多少时间花在了它帮不上忙的那一段。

多数人会发现第一个数很大，第二个数没有变。**而第二个数才是那件事真正的长度。**

九、与相邻说法的分界

计算机体系结构里的加速比分析、缓存命中率、以及关于“并行化受限于串行部分”的经典结论，与本章第五节形式相同。差别在于对象的性质：那些分析中，“不可加速的部分”是一个技术上的事实（这段代码必须串行）；本章的“不可加速部分”是一个**认识论上的事实**（这段判定尚未被任何人作出）。这个差别有一个重要后果：技术上的串行部分可以通过更好的算法被缩短，而认识论上的未判定部分**不能被任何工程手段缩短**——缩短它的唯一方式，是有人真的去判一次。

第四十六章 效率自证悖论与那个 0.989

一个 AI 学习产品越是能以可量化的方式证明自己有效，它就越是在同时证明自己不开辟新航道。

一、陈述

推论。在现行的教育效果证据制度下，“被证明有效”与“开辟新航道”互斥。

这条推论比它听上去更强，因为它不是说“现有研究做得不好”，而是说在这套证据制度里，开辟不产生可比较的读数。

二、证明

证明来自实验设计的一个基本要求。

要证明某个学习干预有效，标准做法是随机对照：实验组用它，对照组不用，两组在同一个因变量上比较。这就要求存在一个对两组都适用的因变量——也就是一份共同答案键。

有共同答案键 $\Rightarrow$ 该答案键在两组作答之前就已存在 $\Rightarrow$ 判定前置 $\Rightarrow$ 由第四十五章主定理的对偶读法，被测量的是加速。

反过来：若某次学习的产出是“学习者作出了一个此前不存在的判断”，则这个产出在对照组里没有对应物——不是对照组做得差，是根本没有可比较的量。你不能对两组说“请各自产生一个前所未有的判断，我们来比谁的更前所未有”，因为“前所未有”不是一个可以在两组间求均值的量。

所以：能被现行证据制度证明的，只能是加速；开辟在这套制度下不产生可比较的读数。

这不是研究者的疏忽，是设计逻辑的必然。

三、实证锚点：因变量清一色是既有答案键

这条推论可以直接在文献上验证，而且验得很干净。目前这一领域最被引用的几项证据：

- 对大模型类学习工具的随机对照试验元分析，报告合并效应量 $g \approx 0.87$ （约四十项试验），因变量为考试成绩、期末评分一类；
- 另一项覆盖三十余项实验、四千余名参与者的元分析报告 $g \approx 0.67$ ，同样以学业成绩类指标为因变量；
- 二〇二五年发表的一项大学物理课堂随机对照试验（样本约五百人），结论是 AI 辅导组比主动学习课堂学得更多且用时更少。

请注意最后这一项的表述形式：更多，且更少。

这正是“加速”的定义式——同一条航道，同一个终点，更高的通过速度。这项研究是对“加速”最干净的一次实测，它的价值不因本书的论证而减损一分；本书只是指出，它测的是速度，从头到尾没有测过航道。

在这几十项试验里，未见任何一项以“学习者作出了一个此前不存在、且被独立第三方判定为有效的判断”为因变量。按第二节的论证，这不是抽样不足，这是不可能。

四、一个更锋利的实例：相关 0.989

比元分析更能说明问题的，是产品侧自己拿出的最高成就。

某国产 AI 学习机把“掌握度模型”作为其核心创新，并给出一个验证读数：经过约二十五万份真实考试试卷的数据验证，机上综合练习的掌握度与真实考试得分率的相关性高达 **0.989**。

这个数字被作为该产品最强的有效性证据提出。而在测量学的框架内，**这个用法完全正当**——效标关联效度就是这么算的，0.989 是一个极好的效标效度。

但请换一个方向读同一个数字：

一个新指标与旧终点的相关达到 0.989，等于宣布：它在旧海图之外，没有测到任何东西。

残差不到百分之二。这条新造的、耗费大量数据与工程的量，在信息上几乎完全被既有的考试得分率所决定。**它是一张更精细、刷新更快、颗粒度更细的旧海图，不是一块新水域。**

五、最关键的一点：高相关恰恰是它被采纳的理由

上一节还不是最要紧的。最要紧的是下面这一步。

如果这个掌握度与考试得分率的相关只有 0.3，它会立刻被判为无效指标而被废弃。

也就是说：

整个行业筛选指标的标准，本身是发现式的。

一个新量要被接受，必须先证明它高度预测既有终点；而**任何能测到新水域的量，按定义与既有终点相关不高，因而在被采用之前就会被淘汰。**

这不是某一家公司的问题。这是效标关联效度这个方法论装置在这里的必然作用。它是一台**收敛机**：系统性地保留与旧账本一致的量，系统性地淘汰不一致的量。

它做的正是它该做的事——只是它该做的事，恰好是关闭航道开辟的可能。

这一节与第八章第三节（自我确证）、第二十六章第四节（被结算的秤会停止测量）构成同一条线的三个环节：读数体系只覆盖第一个冲程 — 新指标必须与旧指标高相关才能被采纳 — 于是新指标只能是旧指标的精细版。三环合起来，**这套体系在结构上没有能力长出第二把秤。**

第三十八章那五个读数因此必须满足一个此前没有明说的条件：**它们与现行成绩指标的相关必须是低的。**一个与期末成绩相关达到 0.9 的“新读数”，无论它叫什么名字，都不是新读数。

六、品类闸：为什么开辟不落在“学习产品”里

由上述分析可以推出一条更硬的东西。

推论。具备“无址入口”的 AI 使用方式，在市场分类上系统性地不落在“学习／教育”品类内。换句话说：“学习产品”这个品类本身，是一道判定前置的强制装置。

证明是一条四步的链，没有一步是关于技术或教育理念的：

一、一个学习产品要卖得出去，必须能证明“学会了”。这是这个品类的价值承诺——无论付费方是学习者本人、家长、企业培训预算还是学校采购。

二、要证明“学会了”，必须有一个对错判据。

三、该判据必须在学习者作答之前存在，否则无法在售前承诺、无法在售后追责、无法在续费时展示。“你学完之后我们再一起商量什么算学会”这句话在商业上无法成立。

四、于是判定必然前置。

那么开辟去哪里了？它没有消失。它在那些把判定交给了世界而不是交给答案键的地方。

写一段代码，它能不能跑是编译器和运行时判的，不是老师判的，而且判定发生在你写完之后。做一次实验，数据是判定者。开一家店，市场是判定者。写一篇公开的分析，读者和后续事实是判定者。在这些活动里，判定时刻晚于产出时刻是常态而不是例外：**没有任何人在你动手之前持有那个答案。**

而 AI 在这些活动里正在被大量使用，效率提升幅度往往比在学习产品里更大。这看起来像是对主定理的反例，简短的回答是：被提速的部分仍然是这些活动里**已判定的那一部分**——样板代码、接口用法、实验流程、格式规范；而“这个设计对不对”“这个市场要不要”没有被提速一分钟。

关键在分类：

这些工具在市场上不被归为“学习产品”，它们被归为“生产工具”。

一个程序员用编码助手写了一年代码，学到的东西可能比任何一个编程课程都多，但没有人把它算作学习产品；它的定价、增长指标、宣传语，全部围绕生产力而不是围绕学习效果组织。

七、这条推论改变了整本书的语气

如果只有主定理与占比反转，本书的结论会是一个批评：市场做错了，产品经理应该做得更好。

品类阐说的是另一回事：

市场上大部分 AI 自学产品是加速而非开辟，不是执行的失败，是筛选的结果。凡是往开辟方向走的东西，都会在某一步被重新归类到别的品类里去，或者干脆不成为一个品类。

这与第五章的机制完全同构：那里是教改被读数筛选，这里是产品被品类筛选。**两处的筛子不是人，都是结构。**

由此得到一条实用的判断法则，供任何想要评估一件工具的人使用：

看它的宣传语在承诺什么。承诺“更快学会”的，一定是加速——因为它必须能证明“学会了”。承诺“让你做出东西”的，才有可能是开辟——因为它把判定交给了产物在世界上的遭遇。

这条法则与第十四章第七节那把尺子（把“学习”换成“接收成品”）可以合用：**一把量它承诺什么，一把量它兑现什么。**

八、一个必须承认的代价

本推论有一个让人不舒服的推论的推论：

本书自己的主张，在同一套证据制度下也无法被“证明有效”。

第五编的那套做法若真的被实施，其产出无法与对照组比较，因为它们生产的正是不可比较的东西。

本书接受这个代价，不掩饰它。这也正是为什么下一章的经验部分做的是**结构编码审计**而不是效果实验——**编码审计可以在不需要共同答案键的情况下被第三方复核。**

这一节也回过头来解释了第三十九章的一处设计：那三组实验的因变量全部是**行为事件**（未闭合时长、第一句话的类型、有没有走出一大步），而不是能力或成就。因为行为事件不需要共同答案键——它们只需要被记录。

九、可观察的动作

—— **找一份 AI 学习产品的白皮书，翻到它的效果验证那一页，看它的因变量是什么。**

如果是成绩、掌握度、正确率、完成率——按本章的论证，**它测的必然是加速**。这不需要怀疑它的诚实，也不需要重做它的实验。

十、与相邻说法的分界

科学计量与评价研究中有关于“可测量性偏倚”的成熟讨论：容易测量的东西获得更多资源与注意。本章与之的差别是一步更强的主张：这里不只是“难测的被忽略”，而是**“能测到新东西的指标会在采纳阶段被系统性淘汰”**——因为采纳标准本身要求与旧终点高相关。前者是资源分配问题，可以靠增加投入缓解；后者是**准入机制问题**，增加投入只会更快地淘汰新指标。

第四十七章 二十四个产品形态的编码审计

—— **结构论证的危险在于它太容易成立。所以先把话说死，再去数。**

一、为什么要做这次审计

到上一章为止，本编的论证全部是结构性的。而结构论证有一个真实的危险：一个**足够抽象的框架可以把任何证据都吸收进去**。

所以本章做一件本书其余部分都没有做的事：一次**预注册的编码审计**——先写死抽样框、编码协议、预测与证伪条件，再去数。

需要说明它的性质：这不是效果实验（上一章已经说明效果实验在这里为什么无效），是**结构编码**。它测的不是产品好不好，是产品的交互结构里有没有一个可以放进未判定之物的入口。

二、抽样框与编码协议

总体：截至二〇二六年八月，在中文或英文市场上可公开取证、并以“帮助个人更高效地学习／自学”为部分或全部价值主张的 AI 产品与产品形态。

抽样方式：按形态分层，覆盖八个形态——通用大模型学习模式、教育垂直助教、语言学习应用、学习硬件、拍照解题、间隔重复、个人知识库、研究综合；外加**外部判官型**作为对照组。

外部判官型是刻意放入的。它不属于本书所定义的学习产品品类，放进来是为了检验品类闸：**如果它与前七组没有差别，品类闸被推翻。**

样本量：二十四。

编码协议（三问）：

Q1 可提交性——能否提交一个系统无法判定对错的产出，而系统不判错、不改写、不静默丢弃？（是／否）

>

Q2 内容回写——该产出能否进入内容层，改变后来者所见？（是／私域／否）

>

Q3 后置判定——是否存在任务，其判定标准在作答后确定并被记录？（是／否）

无址入口 = 1，当且仅当三问中至少有一问为“是”（Q2 取“私域”不计）。

另记三项：判定前置程度、判定权持有者、答案键来源。

三、预注册的预测与证伪条件

预测：

P1：前七组（二十个样本）中，无址入口为零的比例不低于 85%。

P2：内容回写在前七组中的出现次数为零。

P3：对照组（四个样本）中，后置判定为“是”的比例不低于 75%，且这些样本在市场上均不以“学习产品”定价与宣传。

P4（本书最愿意被推翻的一条）：苏格拉底式学习模式一组，其岔口留存率显著大于零。

证伪条件（写死，不留逃生口）：

F1：若前七组中无址入口为一的比例达到 25%，则“大部分是加速”这一主命题被推翻，本章撤回。

F2：若对照组与前七组在后置判定上无差别，则品类闸被推翻，第四十六章第六节须删除

F3：若出现任一在售自学产品，其学习者产出可进入内容层并携带判据地改变后来者所见，则“内容回写与规模化互斥”须改写为“罕见但可能”。

F4：若能出示任一以“学习者产出了此前不存在且被独立第三方判定为有效的判断”为因变量的对照实验，则第四十六章撤回。

以上四条中任何一条被兑现，相应部分作废，不作补救性解释。

四、结果

组	形态	样本数	无址入口
一	通用模型学习模式（含引导提问型）	4	0
二	教育垂直助教（课程内嵌）	2	0
三	语言学习应用（对话练习、纠错打分）	2	0
四	学习硬件（掌握度模型、诊学练测、同步课程）	3	0
五	拍照解题（步骤型、求解器型）	2	0
六	间隔重复（官方牌组、自建卡片、AI生成）	3	0
七	个人知识库（自带语料问答、笔记整理）	2	0（两次“私域”不计）
八	研究综合（文献提取、深度报告）	2	0（Q1记“部分”）
对照	外部判官型（编码助手、开源协作、数据竞赛、辅助研究发布）	4	4

主命题的读数：二十四个样本中，二十一个无址入口为零（87.5%）；在符合“学习产品”品类定义的二十个样本中，比例为100%。证伪条件 F1 远未触及。

F2 未触发：对照组四个样本在后置判定上与前二十个样本呈现完全分离（四比零对零比二十），品类闸得到最强形式的支持。

F3 未触发：唯一出现内容回写的三个样本全部不在自学产品品类内，而且其中两个（开源协作、公开榜）本质上不是产品而是制度——它们的内容回写能力来自共同体的评审与排名规则，不来自任何一家公司的功能设计。

最后这一点值得单独停一下：内容回写在市场化产品里被排除，在非市场化制度里存活。这与第四十六章第六节的四步链完全吻合——一个产品要能被交易，判定必须前置；而一个制度不需要被交易。

五、真跑逼出的三处改判

一次真做过的审计，价值往往不在它证实了什么，而在它逼出了什么。这次审计逼出三处改判，其中两处修正了原有的框架。

改判一：判定前置程度单独不足以区分，必须加上“答案键来源”这一维。

原设想是用“判定是否前置”一个维度就能分类。审计中发现两个样本（第七组的知识库类）判定前置程度接近于一，但它们的答案键来源是**用户自带语料**，与考纲、教材、答案库有实质差别——它们至少不强制对齐到一个外部权威。这一维必须单列，否则会把“用户自己的材料”与“考纲”混为一谈。

改判二：深度研究类无法被原有二分法容纳，逼出第三类“代判”。

第八组两个样本（文献筛选提取、深度研究报告）在Q1上只能记“部分”：它们**接受一个模糊的、尚未成形的问题**，这是原有分类里“加速”不该有的性质；但它们随即替使用者作出了大量判定——哪些文献相关、哪些结论可信、如何取舍——而这些判定既不前置于使用者，也不由使用者作出。

于是原有的二分（加速／开辟）不够用，必须改成三分：

加速·代判·开辟

而**增长最快的是中间那一类**。

这一处改判很重要，因为“代判”是一个此前没有被命名的位置，而它在本书的框架里恰好落在最危险的地方：它看上去像开辟（因为它处理未成形的问题），实际效果却比加速更彻底——加速至少还让使用者自己走完那一段，代判连走都不必了。

第三十二章的两条硬约束在这里得到一个新的应用形态：**代判就是“一方同时占据显露与路径”在自学场景中的极端形式**。

改判三：一处方向相反的读数。

预注册的P4是本书最愿意被推翻的一条：苏格拉底式学习模式一组，其岔口留存率显著大于零。

实测结果：接近于零，与预期相反。

也就是说，那些以“不直接给答案、用提问引导你”为卖点的学习模式，在编码上与直接给答案的模式**没有结构差别**——它们的答案键同样先在，区别只在交付节奏。

这与第四十五章第一节那条形式化判断完全一致，但它是一个**在预注册中被押注错误、然后被数据纠正**的读数。本书把它原样保留在这里，而不是事后改写预测，因为这正是预注册的用途。

它对第五编也有一个直接的警示：“**少给一点、慢慢引导**”不构成换道。如果那条被引导的路仍然通向一个已定的结论，它只是把同一件事拉长了。第六章第七节的“伪未定终点之二”（终点未定但路径全定）在这里获得了产品层面的实证形态。

六、四条限制（与结果同时交付）

限制一：文档层，非日志层。全部编码基于公开文档、公开评测与直接使用，未访问任何产品的内部交互日志。一个产品可能在文档里没写、但实际具备某种后置判定通道；这类情形会被漏判为零，**方向是有利于本书命题的偏误**。这是本审计最大的弱点——**任何有内部数据的一方都可以低成本地推翻某一格**。

限制二：单人编码，无信度系数。未做双盲双编，未计算编码者一致性。三问的措辞已尽量做到只读交互模式，但仍存在解释空间，尤其在“静默丢弃”与两处“部分”判定上。

限制三：抽样非概率抽样。形态分层由作者划定，样本在层内按可取证性选取。因此87.5%这个数字不应被读作总体比例的估计，只应被读作“在八个主要形态上，分离是干净的”。

限制四：时间切片。编码状态为二〇二六年八月。这一领域迭代极快，任一格都可能在数月内翻转。本书欢迎这种翻转——若某个自学产品增加了真正的后置判定通道，那正是第十四章第八节所处方的东西被实施了。

七、一处唯一被找到的正面实例

审计过程中找到一个此前未预料的实例，值得单独记下，因为它是本书全部主张在现实中被完整实现的唯一一处。

在形式化数学的证明库里，有一个记号用于标记“这里有一个尚未完成的证明”。它不是错误，不是占位符的敷衍，也不会被静默丢弃：它被系统显式记录、可被检索、可被统计、可被后来者接手完成。整个库的健康状况，部分地就由这类标记的数量与分布来刻画。

用本书的语言说，这个记号是“未决判定准备金”的字面实现：一个制度为“尚未判定”专门开设的账面科目。

而它有两个特征值得注意：

第一，它不被任何人当作学习产品。它是一个工作制度，不是一个教育功能。这再次印证了品类阐。

第二，它是本书第三十八章那张记录表的完成形态。那张表最后一列没有一项写着“评价”，理由与此处相同——一旦“尚未完成”被用于评价，人们就会停止标记它。

八、可观察的动作

拿你正在用的那个 AI 学习工具，试着提交一个它无法判定对错的东西，看它怎么处理。

三种可能：判你错、替你改写、静默丢弃。这三种都是零。

如果它把它原样留下、标记为“未判定”、并且允许你以后回来看——那么你找到了第四种值得告诉别人。

九、与相邻说法的分界

产品评测与教育技术评估中有大量成熟的评价框架（可用性、有效性、参与度、隐私合规）。本章与它们的差别只有一处，但是决定性的：那些框架全部评估产品在它自己的目标上做得如何；本章评估的是产品的目标本身覆盖了什么。一个在所有现行评估框架上都得满分的产物，在本章的三问上仍然可以是零——而且按第四十六章的品类阐，它多半必然是零。

第四十八章 没有人当期受损：谎言不需要有人撒谎

一个命题被广泛相信、它宣称达成的目标与它实际优化的量方向相反、而所有能判定这一点的读数恰好都没有被测量——这样的命题就是谎言，无论说它的人多么诚实。

一、一个方向相反的承诺，为什么被整个行业接受

前面几章证明了那句话必然自败。而一个必然自败的说法被所有人接受，本身是一个需要解释的现象——它比那句话本身更值得研究。

本章给出的解释不诉诸任何人的动机：它是一个激励结构的分析。

二、六方的账，一方一方摊开

厂商。产品形态清晰、指标可量化、见效快、可复制、可招标。做“更好的搬运”是一门确定的生意；做“守住不闭合”不是生意——它的产出物是“什么都没发生”，无法演示，无法计价，无法写进标书。

教育行政。需要可上报的进展。覆盖率、设备数、平台使用时长、课堂参与度、成绩提升幅度，全部现成。而本书主张的那个东西，最好的成绩单是“本学期学生独立卡住的总时长上升了四成”——没有一个层级敢把这句话写进汇报材料。

学校管理者。需要在督导与评比中拿到位置。AI 课堂是可展示的，是能开公开课的，是能上简报的。

教师。“赋能”这个说法保住了他的岗位描述。它告诉他：你原来做的事是对的，只是现在有了更好的工具。相比之下，本书的判断对他极不友好——它说他每天做得最熟练的那三个动作正在被取消，而且取消它们的理由不是他做得不好，是做得再好也不生产他真正想要的东西。任何人都会更愿意听前一句。

家长。看到的是“个性化”“因材施教”“不用报班了”，以及当期作业质量与分数的真实改善。他们没有任何理由反对。

学生。当期是纯获益的。作业更快做完，卡顿更少，不必再忍受不会的那半小时。要他们主动放弃这一点，等于要人主动选择难受。

六方全部当期获益。而受损方是十五年后的这个社会——它没有代表，没有投票权，没有采购预算，也没有在任何一次会议上出席过。

这就是这个说法的第一层保护壳。它不是被权力维持的，是被利益的完全一致维持的。

凡是没有任何一方在当期受损的说法，就没有任何一方有动力去验它。

三、第二层：读数为什么没有被测量

第三十八章给出了五个读数。这里补上三个来自另一条线的读数，它们的性质与那五个不同——它们直接指向供给方。

读数 τ （未闭合时长比）：一节课中处于未闭合状态的时间总和，除以课堂总时长。

读数 t_1 （求助前独立时长）：学生从遇到障碍到发出求助之间的时长分布。

读数 R（责任可指认性）：这堂课上做出的关键教学判断，能否指认出一个会因它出错而承受后果的具体的人。

这三个读数有一个共同特点：**取数成本极低**。 τ 只需要一块秒表和一份编码规则； t_1 在任何一个学习平台的日志里都是现成字段，甚至不需要新采集；R只需要问一句“这个判断，谁为它的失败负责”。

它们还有第二个共同特点：**掌握在供给方手里**。平台方最清楚 t_1 的分布长什么样，因为那就是他们的日志。

以及第三个特点：**结果几乎肯定对产品不利**。一个以“即时解答”为卖点的产品，它的日志一定会显示学生求助前的独立时长在缩短——那正是它的产品目标，是它写在融资材料里的成功指标。

三个特点凑在一起，得到一个很干净的推论：

没有人会去测它。

不是因为有人下令封锁，而是因为任何一个理性的行动者都没有理由花钱去买一份对自己不利、对同行也不利、且没有任何监管方要求提交的数据。

这就是谎言不需要有人撒谎的机制。它只需要两条：**所有人当期获益 + 唯一相关的读数没有人有动力去测**。这两条一旦同时成立，一个方向完全相反的命题可以在一个全是诚实人的行业里稳定运行几十年。

四、这是一条已知规律的特殊变体

历史上这样的事例并不罕见。关于“指标一旦被用作目标就不再是好指标”的经典表述有过好几个版本；关于“反应性”的研究更进一步：被测量的对象会重组自己去迎合测量，而重组的部分恰恰是原本要测的东西；关于“可辨识性”的分析说的则是：为了管理而简化现实，最终把现实中那些不可简化但至关重要的东西碾掉。

本书所描述的，是这条规律在教育领域的一个特殊变体，而且**更难被发现**——因为：

这里连一个错误的指标都还没有。判断力从来没有被列入过任何报表的科目。

指标被扭曲，至少还有一个指标可以被追问；一个科目从未存在，则连追问的对象都没有。这就是第八章第八节那条分界的最终形态。

五、每一步在当时都是对的：四条历史来路

有必要说清楚：这不是一个突然出现的问题。**AI 不是元凶，它是最后一步**。前面每一步都走了几十年，而且每一步在当时都是对的。

第一条：教科书 – 详细解答 – 题库 – 分层练习 – 智能推送。

教科书只给结论，学生课后无从检查——于是有了详解，这是对的。详解只有一种走法，练得不够——于是有了题库，这是对的。题库不分难度，浪费时间——于是有了分层，这是对

的。分层还是粗——于是有了自适应推送，这是对的。推送还是滞后，要等作业交上来——于是有了实时对话式辅导，这也是对的。

每一步都在缩短同一个量：学生独立面对一道不会做的题的时间。没有任何一步是错的，而这条链走完之后，那个时间趋近于零。

第二条：泰勒制－标准作业书－六西格玛－黑灯车间。

工业界走的是同一条链，而且走完了。每一步的理由同样无懈可击：减少变异、降低不良率、提高一致性、消除人为失误。结果是工艺判断力的消失——那种老师傅站在机器旁听声音就知道哪里不对的能力，不是被禁止的，是**再也没有产生的机会了**。因为产生它的条件是“被物理世界的微小扰动反复穿过、在混乱中被迫重新站稳”，而整条链的目标就是消灭扰动。

工业界比教育界早三十年走完这条路，**代价已经付出并且被记录**：一旦老师傅退休，重建需要一代人，而且往往重建不起来——因为重建所需的那个环境已经不存在了。

第三条：专科化－循证－临床路径－支付制度改革。

医学同样。专科化提高了水平，循证减少了臆断，路径保证了下限，支付制度让一切可核算。每一步都是巨大的进步。而合起来的结果是：一个医生完整走通“感知异常—做出裁决—承担后果”这条链的机会在减少，取而代之的是在既定路径上的高质量执行。

第四条：自由讨论－任务单－学习目标－可观测行为－数据看板。

教育自身也走过一次。当年发明任务单，是为了对抗一件真实的坏事：所谓“自由讨论”经常什么也没发生，热闹四十分钟，一无所获。任务单把讨论变成可检查的，这是对的。学习目标把课时变成可评价的，这是对的。可观测行为让评价不再靠印象，这是对的。数据看板让管理者第一次看见课堂内部，这也是对的。

四条链的共同结构完全一致：

每一步都在消灭一种真实的坏东西，而每一步同时也削掉一点那段“没有人替他闭合”的时间。前者被记录、被表彰、被写进制度；后者从未出现在任何一张报表上，因为它从来没有被命名过。

AI 是这四条链的**共同终点**。它一次性地把所有链上剩下的余量清空。

所以对 AI 的愤怒是错位的。它只是把一件已经做了一百年的事，一次性做完了。

六、堵死怀旧这条路

上一节很容易被读成怀旧：那就回到从前吧，回到书少、老师少、答案要等一周的年代。

这条路必须当场堵死，否则本书的全部论证会被扭曲成一种反动的乡愁，并且会被正确地驳倒。

第四十四章第三节末尾已经埋下这个区分，现在把它说完：

起作用的从来不是“效率低”，而是“一段不稳定被留在那里、没有被别人替他闭合”的那段时间。

低效率只是历史上恰好制造出这段时间的原因，而且是一个**极其糟糕的原因**，因为它同时制造了三样坏东西：

第一，它把这段时间分配得极不公平。在书少的年代，能撑住这段时间并且真的走过去的人，绝大多数是那些家里恰好有另一个成年人可以撑住他的人。其余的人不是获得了判断力，是被淘汰了。历史上的判断力生成机制，实际上是一台筛选机而不是培养机。

第二，它的失败没有兜底。一个学生真的过不去，就是真的过不去，没有第二次。

第三，它无法被主动配置。因为它是稀缺的副作用；稀缺一旦缓解，它就跟着消失，谁也拦不住。事实上过去一百年的教育史，就是一部不断缓解稀缺、因而不不断意外地削掉这段时间的历史。

所以正确的结论不是“回去”，而是：

过去那段时间是免费的，也是被随机分配的；今后那段时间必须是被有意配置的，并且必须有人为它的失败兜底。

这一节与第十九章第二节完全接榫：造出来的入口比捡来的入口好——因为它可控、有起止、对所有人平等，而且有人兜底。

七、于是本书的判断有了它的第二面

这里出现一个重要的转折，它让全书不至于停在批判上。

一对一地陪着一个学生走一段没有答案的路，历史上从来不是没有人想到——是从来负担不起。它的成本是每个学生一位成年人。整个班级授课形态，就是这个成本约束的直接产物（第三章）。

AI 第一次改变了这个成本结构。备课、批改、出题、查资料、生成变式、整理学情，这些原本吃掉教师八成时间的事情，现在可以被真正地卸下来。被释放出来的时间，理论上可以用于那件唯一不能被卸下的事。

于是：

AI 是发生学教育的使能条件，而不是发现学教育的加速器。同一台机器，两种配置，方向完全相反。

>

配置成答案源，它就是发现学教育的完成；配置成造岔口的对手，它才是另一条道路的开端。

而“AI 赋能”这句话之所以是谎言，最根本的一点在这里：它把这台机器唯一有价值的用法，掩盖在了它最自然、最省事、最能立刻见效的用法之下。

它不是错在赞美 AI。它错在替 AI 选了那个方向相反的配置，还说这是进步。

八、 $\tau \times R$ ：一个可被推翻的核心形式命题

第三节给出的三个读数中， τ 与 R 有一个特殊关系：它们是相乘的，不是相加的。

一段无人闭合的时间，如果没有人为它的失败负责，它不是教育，它是放任；一份清晰的责任，如果没有任何一段悬空可供承担，它不是教育，它是执行。

由此得到一条可以被推翻的核心形式命题：

判断力的代际存活 $\propto \tau \times R$ 。

推翻条件。如果能找到这样一个场： τ 极小（几乎不存在无人闭合的时间）而 R 清晰，且该场中判断力仍然稳定地代际存活——那么这条乘积形式被推翻。反之，若某场 τ 很大而 R 为零、判断力仍稳定存活，同样被推翻。

这条乘积形式顺带解释了两件长期让人困惑的事：

为什么单靠“慢教育”“留白”“减负”救不回来——只补 τ 。

为什么单靠“人类最终决定”“AI 只做辅助”同样救不回来——只补 R 。

两者缺一，乘积为零。

而它也给全书加了一条最终的自我约束，必须写在这里：

本书的全部主张，只在“有人为悬空期间的失败负责”这一条件下有效。任何把本书当作“少讲一点、让学生自己扛”之依据的做法，都误读了它。缺了 R ， τ 就是放任；而放任比赋能更糟。

九、可观察的读数

课后问自己三句：今天这节课上，哪一个决定如果错了，后果由我承担？这个决定是我做的，还是系统／模型／教研组／教参做的？如果这个学生这学期就此掉队，会有人来问我吗？

三问都答不出具体内容的，那节课的 R 为零。而 R 为零时， τ 再大也不产生任何东西。

十、与相邻说法的分界

公共选择与集体行动研究中有关于“分散的成本与集中的收益”的经典分析：受损方分散、每人损失微小，因而无人有动力组织起来。本章的结构与之相似但有一处更极端的差别：**这里的受损方不是分散的，是尚未存在的。**十五年后的社会不是一群难以组织的人，而是一个此刻还没有成员的集合。它连“分散”都谈不上——**它无法被贿赂，也无法被组织，因为它还没有出生。**这使得任何依赖利益相关方发声的纠错机制在这里全部失效，也使得本书第三十八章那种“先加一栏、第一次可以是空的”的做法，成了少数可行的动作之一。

第四十九章 六个学科现场

一个关于教学的判断，如果不能落到某一节具体的课上、落到某一句具体的教师用语上，它就还只是一种立场。

一、格式与一条前置说明

第三十六、三十七章按“光滑化在哪里／岔口从哪里开／AI配哪一格／回写长什么样”处理了五科。本章换一个格式，把同样的判断落到六节具体的课上，每一科走五步：

- 一、现行搬运版——这一课今天在绝大多数教室里是怎么上的；
- 二、被光滑化掉的是什么——这个知识当初是在什么处境里被逼出来的；
- 三、AI赋能之后会发生什么——不是想象，是现在已经在发生的产品行为；
- 四、造岔口版——同一课的另一种上法，写到教师的具体用语；
- 五、 τ 落在哪一段——这一课的未闭合时长具体是哪几分钟。

一条必须前置的说明：造岔口版不是让学生自己瞎摸。它的结构永远是：先让学生用手上已有的东西去处理一个够得着一半的处境，让那个处境在他手上当场破裂，然后不替他修。

没有弹药不能上战场；但把弹药发到饱和，从来不等于打过一仗。

二、数学：分数除以分数，为什么可以“颠倒相乘”

现行搬运版。给出法则：除以一个分数等于乘以它的倒数。给两个例题，做十道练习。有的版本会补一句“因为除法是乘法的逆运算”。

被光滑化掉的是什么。这条法则不是从天上掉下来的，它是为了让运算能够封闭而作出的一次规定。在只有整数的世界里，除法经常做不下去；每一次数系的扩张，都是为了让某一类做不下去的运算变得做得下去，而代价是必须重新规定这个运算在新对象上是什么意思。“颠倒相乘”是那次规定的结果，不是那次规定的理由。

AI赋能之后。孩子拍下题目，得到完整步骤、变式讲解、错因分析、同类题推送。他的正确率会明显上升，而“为什么可以颠倒”这个问题从头到尾没有出现过——因为它不是一道题，没有人会去问它。

造岔口版。

教师用语：

“三分之二里面有几个六分之一？”（先让他数，能数出来。）

“那，三分之二里面有几个四分之三？”（这一句是岔口。它够得着一半——形式上完全一样，但数不出来了。）

“你刚才那个办法怎么不好用了？”

——不给提示，等。

“那你能不能自己规定一个办法，让它也能算出来？规定完了我们检查它跟前面那些算法冲不冲突。”

最后这一句是元层解锁三问的第三问（第三十四章第七节）的学科形态：从“接受一条规则”变成“制定一条规则并检验它”。

τ落在哪一段。从第二句到孩子提出第一个自己的办法之间。这一段通常是六到十二分钟，是全课最难忍的一段，也是唯一不能压缩的一段。

三、语文：作文课上那十分钟

现行搬运版。给题目，给范文，讲结构，列提纲，学生写，教师批改讲评。

被光滑化掉的是什么。范文是成品。**成品看不见被删掉的那些句子。**一个作者真正的工作，绝大部分发生在“这句话要不要留”这个判断上，而这个判断在范文里一个字都看不见。

AI 赋能之后。孩子输入题目，得到提纲、开头、若干可选段落、修改建议。作文质量当期显著提升。**而那个判断——这句话要不要留——他一次都没有做过。**

造岔口版。

教师用语：

“把你上一篇作文拿出来。找出你当时删掉的那一句——或者你当时特别想写、但最后没敢写的那一句。”

“写在纸上。”

“现在告诉我：你当时为什么删它？”

——不评价，等。

“如果把它放回去，这篇文章会变成什么样？”

这一课的材料完全来自学生自己（土壤在他手上），而且它同时是一次回写——他在回答“当初为什么删”的过程中，往往会推翻当初的理由。

τ落在哪一段。从“你当时为什么删它”到他给出第一个不是“老师说过”的理由之间。

四、英语：语法全对，为什么还是不对

现行搬运版。给规则，给例句，做练习，纠错。

被光滑化掉的是什么。规则是从大量使用中沉淀出来的**事后描述**，而教材把它当作先在的法律发下去。于是例外只能被当作需要死记的意外，无法被理解。

AI 赋能之后。孩子写一句话，立刻得到纠正与解释，准确率上升。而**“为什么这句语法全对却没有人这么说”这个问题不会被问出来**——因为纠错系统的任务就是让它不出现。

造岔口版。

教师用语：

（给出一句语法完全正确、但母语者不会这么说的句子）“这句话有语法错误吗？”

“没有。那为什么没有人这么说？”

——等。

“我们中文里有没有这样的句子？——语法上没毛病，但大家就是不这么讲。”

“那条‘不这么讲’的规矩，是谁定的？写在哪里？”

最后一问把学生的母语经验调进来作为材料，而这份材料是他独有的、AI拿不出的。

τ落在哪一段。从第二问到有学生举出第一个母语反例之间。这一段通常很短（两到四分钟），因为母语材料人人都有——这是六科中**τ最容易起步的一科**。

五、科学：把预测写下来，再做实验

现行搬运版。讲结论，做一个必定成功的实验去印证它。

被光滑化掉的是什么。科学史上的次序是反的：先有一个说不通的现象，然后是一堆互相打架的解释，然后才有判决性的观察。一个**必定成功的实验，在结构上与一段讲解没有区别**——它的结论在开始之前已经写在书上。

AI赋能之后。实验前可以先问它会发生什么，得到准确的预测与原理解释。于是实验彻底变成一次确认动作。“**我以为会那样，结果不是**”这件事被提前消灭了。

造岔口版。

教师用语：

“先别做。每个人在纸上写下：你觉得会发生什么？为什么？”

“写完扣过去。不许问任何人，不许查。”

——**这一步是全课的关键，不能省。**

（做实验）

“翻开你写的。跟你写的一样吗？”

“不一样的举手。——你原来那个想法，错在哪一步？”

“写下来”这个动作是整节课的技术核心。它把一个模糊的预感变成一份有日期、有署名、不可抵赖的记录，从而使“我原来以为”成为一件可以被指着看的东西。**没有这一步，回写没有可改的对象。**

τ落在哪一段。从“先别做”到揭晓之间。这一段可以只有三分钟，但它必须是**封闭的**——不许问，不许查，不许交流。这是全书的规矩（第三十二章第二节）在实验课上的具体形态。

六、历史与道德法治：把结局盖住

现行搬运版。讲事件、讲原因、讲意义、讲评价。

被光滑化掉的是什么。当时的人不知道结局。而教材是从结局往回写的，于是每一个选择都显得要么英明要么愚蠢——因为我们已经知道后来发生了什么。

AI赋能之后。任何一个历史事件的背景、原因、影响、多方评价，都可以在几秒内得到一份结构清晰的综述。而“当时的人在不知道结局的情况下会怎么选”这个问题不会出现——因为综述天然是从结局写起的。

造岔口版。

教师用语：

“我现在给你们材料，截止到某年某月。后面发生了什么，我先不说。”

“你们是当时那间屋子里的人。手上只有这些。现在必须做一个决定。”

“写下来：你选哪个？你最担心的是什么？”

——等，让他们真的做出选择。

（揭晓）“结果是这样。现在回头看你写的那个‘最担心的’——你当时担心对了吗？”

最后一问是这一课的回写口：它改的不是那个决定，是他对“什么值得担心”的判断。

τ 落在哪一段。从“必须做一个决定”到揭晓之间。这一段是六科中最容易被外部信息破坏的一段——只要有一个学生查了，全班的岔口同时关闭。因此这一课必须在规矩上写死“这十分钟不许查”。

七、艺术：作品被判定之前的那一段

现行搬运版。讲名作，讲技法，讲评价标准，学生仿作，教师点评。

被光滑化掉的是什么。一件作品在被完成、被认可之前，有一段谁也不知道它算不算数的时间。评价标准是后来的事。而课堂上评价标准先于创作被发下去，于是创作变成了对一个已知标准的逼近。

AI 赋能之后。孩子可以立刻得到符合任何风格的高质量作品，也可以让它给自己的作品打分并提出改进。于是“我不知道这算不算好”这个状态被彻底取消了——而那个状态正是这一科唯一不可替代的经验。

造岔口版。

教师用语：

“今天做完的东西，我不打分，也不点评。”

“先把它挂出来，只做一件事：每个人去看别人的，找出一件你说不清为什么但你觉得有意思的。”

“现在试着说清楚——说不清也没关系，说到哪算哪。”

——教师全程不给标准。

“下节课我们再来定：这一批东西，好在哪里，由我们自己定几条。”

最后一句把评价标准的制定权交给了这个班——这是六科中最彻底的一次判定后置：判定标准在产出之后才被生产出来，正是第四十五章第一节所说的第二种学习。

τ 落在哪一段。这一课的 τ 覆盖了几乎全部课时，而且它跨课时延续到下一节。这是六科中 τ 最高的一科，也是最不被当作“主科”的一科。这两件事同时为真，值得所有人想一想。

八、六科的对照

学科	岔口的那一句	τ 落在哪一段	最容易失败的地方
数学	三分之二里面有几个四分之三	到他提出第一个自己的办法	教师提前给“倒过来乘”
语文	你当时为什么删它	到他给出第一个自己的理由	教师评价那句话写得好不好
英语	语法没错，为什么没人这么说	到有人举出母语反例	教师直接讲“语用”
科学	先别做，写下你觉得会怎样	到揭晓	省掉“写下来”这一步
历史	你是当时屋里的人，必须决定	到揭晓	有人查了结局
艺术	好在哪里，我们自己定几条	几乎全部课时	教师先给了评价标准

第四列值得单独看：六种失败方式里有五种是教师做了一件他一直被训练要做好的事。这正是第三十四章那张禁令表的存在理由——本不会被道理改变，只会被具体的替代动作改变。

九、可观察的动作

挑你自己那一科，把上面那句“岔口的那一句”抄下来，下周用一次。什么都别改，只用这一句。

用完之后记一个数：从你说完这句话，到有人给出第一个自己的说法，中间过了多少秒。那个数，就是你这节课的 τ 。

十、与相邻说法的分界

各科都有成熟的“典型课例”传统，本章的六节课在形式上属于同类。差别在于课例被记录的颗粒度与对象：常规课例记录的是教师做了什么（提问、活动、板书、总结），本章记录的是教师没有做什么，以及那段没有做的时间有多长。前者可以被模仿，后者只能被守住——而这本书全部的困难，都在于第二件事无法通过示范来传授。

第七编结语：这一编与前六编的四处咬合，以及三处必须划开的地方

补编到此结束。它是后写的，因此必须交代它与前六编怎么接、以及哪里会被误读为重复。

一、四处咬合

第一处：第四十四章咬第十二章。第十二章说加固悖论是一个**读数现象**（读数会同时变好）；第四十四章把它推到**逻辑现象**——那个乘法式子的三处隐含承诺无法同时成立。前者说“你测不出来”，后者说“它本来就不成立”。**两者独立：即使加固悖论被数据推翻，六条论证仍然要各自被驳。**

第二处：第四十五章咬第十九章。第十九章的“毒性=老路×速度÷摩擦”是一个定性式子；第四十五章的判定时刻 t_J 与 t_O 把它的中间那一项变成可读的量—— Δt 大于零就意味着该环节的判定已经前置。**这是本书唯一一处把“快”直接翻译成“已经判过了”的形式装置。**

第三处：第四十七章咬第三十九章。第三十九章写了三组实验设计与三条撤稿条件，却一次真跑都没有；第四十七章那次二十四四个产品形态的预注册编码，是本书唯一一次自己执行的**经验工作**，而且它报了三处对自己不利的改判。第四十二章那句“论证密度高于证据密度”因此要收窄一格：**在产品形态这一层，本书有证据；在课堂效果那一层，仍然没有。**

第四处：第四十八章咬第三编与第二十七章。第三编讲损害为什么看不见，第四十八章讲为什么没有人当期受损——六方的账逐一摊开，没有一方在当期是亏的。这解释了第二十七章那个“第三周”：改革之所以死在那里，不是因为有人反对，是因为**继续下去在当期对每一方都更差。**

二、三处必须划开的地方

划开一：第四十九章不是第三十六、三十七章的重复。那两章问的是“这一科的光滑化藏在哪儿、岔口从哪儿开”；第四十九章问的是**教师那一刻具体说哪一句话、以及说完之后记哪个数**。前者是备课，后者是逐字稿与读数落点。**同一科在两处出现，写的不是同一层。**读者若只需要一样，教师取第四十九章，教研组取第三十六、三十七章。

划开二：第四十八章第八节的 $\tau \times R$ ，与材料二十六的 $\tau \times R$ 是同一个式子，来源不同。材料二十六来自站内四个行业的跨场归纳；第四十八章从六方的账推出同一形式。**两条独立路径得到同一式子，这是它值得被检验的理由，不是它被重复叙述了两遍。**若两处有出入，以材料二十六的可证伪条件为准（若某场 τ 极小而 R 清晰、判断力仍代际存活，乘积形式被推翻）。

划开三：第四十五章的“无址岔口”，与第二十二章的“空位题”不是一回事。空位题是教师出得出来的题——它对这个班是空的；无址岔口是**连地址都没有的那一类**，不可直接计数，只能靠三问间接判定。前者是教学动作，后者是本体论断言。**把后者当教学目标是不可能的；把前者当本体论断言是夸大。**材料六十一那条心法在这里再说一遍：**在薄账本上是空的地方，在厚账本上往往是一整栏。**

三、补编的证据强度

按第三十九章第六节的分档：

- **有直接经验支持的：**第四十七章的二十四例编码（含三处改判、四条限制、一个正面反例）。
- **有间接支持、机制清楚的：**第四十四章六条论证、第四十八章六方的账。
- **目前只是形式装置的：**第四十五章的对偶律与第四十六章的自证悖论——它们是从定义推出来的，不能被数据直接推翻，只能被指出定义不当。

—属于教学建议的：第四十九章。

一处需要点明的弱项：第四十六章第四节引的那个 0.989，是一项公开报告的相关系数，本书没有能力核实它的取数口径。它在论证里的作用是举例，不是承重——去掉它，第四十六章的结论不变。这一句必须写出来，否则读者会以为本书的推论建立在那个数字上。

全书补料：三条限度与三个未回答的问题

本节材料的来源：SDE 站内专著《还要老师吗？》的总序第七节与结语第六节。
本书第四十二章已列出五个漏洞，此处补上三条更根本的限度与三个未回答的问题。

材料四十八 限度一：代价外溢时让位

本书全部关于“克制”“不供给”“让他卡住”的主张，有一个前提：卡住的代价由他自己承担，且代价可逆。

当代价外溢时，这些主张一律让位：

- 涉及求助信号、安全风险、不可逆伤害；
- 卡住会伤及第三人；
- 时间成本落在他人身上（例如全班等一个人）。

这一条必须写在任何一次教师培训的最前面，因为它是最容易被误用的地方——一位教师若把“守住岔口”当作原则而不是当作有边界的方法，会在某个时刻做出真正的伤害。

材料四十九 限度二：先备不足时反向成立

这是三条限度里最不受欢迎的一条，也是本书必须正面写出来的一条。

本书主张的做法，对先备知识不足的学生可能反向成立。

理由不复杂：一个手里没有旧账的孩子（第三十章第三节说旧账是他唯一的独家材料），在“顶”那一段拿不出任何东西。他不是产出判断，他是在承受十五分钟的挫败。

推论：这套做法可能使判断力的分布更不平等。

优势的孩子有旧账、有语言、有在家被追问的经验，他们在这套安排下获益最大；弱势的孩子在旧安排下至少能接住成品，在新安排下可能两头落空。

本书不回避这一条，也不假装有解。能提供的只有两点：

第一，第三十一章第三节第二行给出了一个部分对策——旧账少的孩子先说“我现在觉得是怎样”，那个粗糙说法就是他的材料。这可以缓解，不能解决。

第二，这一条应当被列入第三十九章的检验：任何一次实验都必须分层报告结果，而不能只报均值。若实验组的组内差异显著大于对照组，本书的处方需要重新设计。

材料五十 限度三：抵抗会被制度化

第三条限度是关于本书自身命运的。

按材料三十的再分泌律：凡是被证明有区分力的东西，都会被吸收进评价装置，然后被训练，然后失效。

本书所主张的那些东西，正是当前唯一符合“不可被工具补足”这个条件的候选，因此它们是下一层外壳最可能的建材。

于是：

| **本书如果成功，它所保护的那样东西就会被制度化，然后被消灭。**

第三十八章的三条纪律（不用于排名、不单独使用、不设上限）是为了推迟这一天。但推迟不是阻止。

这一条对读者的实际含义是：如果你在某一天看到“回写率”出现在教师考核表上、“未闭合时长”被设定了达标线、“守岔口能力”成为职称评审的一个项目——那么这本书已经完成了它的失败。

材料五十一 未回答的问题一：不平等怎么办

材料四十九指出了问题，没有给出答案。

这不是一个可以在教育内部解决的问题。判断力的分布不平等，根源在旧账的分布不平等，而旧账来自家庭、语言环境、被追问的经历——这些在孩子进入教室之前就已经分好了。

本书能说的只有一句：**老路在这件事上做得比新路好**。光滑化把知识变成“不属于任何人”的东西，从而使它“属于所有人”（第七章第四节）。这是老路最大的成就，而本书的处方在这一点上是退步的。

这个退步是真实的代价，不是一个可以被修辞化解的顾虑。

材料五十二 未回答的问题二：窗口还剩多久

本书第二十一章推测了一条三段曲线，但没有回答一个更实际的问题：**从现在开始，还有多久？**

按材料二十三的枯竭期公式，教育的 W （一个人从新手到能守住岔口所需的年限）最长，因此显影最慢——**这既是坏消息也是好消息**：坏消息是它已经在进行而我们感觉不到，好消息是它给了最长的反应时间。

但本书无法给出数字。任何给出数字的人都应当被要求出示他的 W 与 J_0 。

材料五十三 未回答的问题三：如果那一天来了

第三十五章第六节写过一限度：本书关于教师不可替代性的论证，依赖于当前技术的一个事实状况——一个系统还不能可靠地读取一个孩子此刻的状态并据此决定该不该开口。

如果那一天来了呢？

本书的回答分两半。

第一半是诚实的：那一天来了，第三十五章需要重写，而本书关于“教师位置”的全部论证失效。

第二半是本书唯一敢下的一个判断：即使那一天来了，“谁有权决定”与“谁有能力决定”仍然是两个问题。

一个系统可以有能力判断此刻该不该给答案，而由谁来对这个判断的后果负责，是一个不会因技术进步而消失的问题。按材料二十六的 $\tau \times R$ 定理， R （闭合责任人是否可指认）与 τ 是两个独立的因子——技术可以把 τ 还回来，它没有办法把 R 还回来，因为承担责任不是一种能力，是一种关系。

这是本书留给三年之后的读者的最后一句。

材料五十四 站内取料索引

本次补料所取的站内资产，逐条列出出处，便于读者回查原作。本书只取其机制与量化形态，论证过程与案例现场在原作。

补料	站内出处	取的是什么
二十三	《反抽的消失——黑灯车间》	枯竭期公式；被顶回来的机会
二十四	《劳动断链》	工序完整性指数；感知—裁决—承担
二十五	《切痛》	程序性奢侈；说理与类案的重合度
二十六-二十七	《还要老师吗？》跨编定理	$\tau \times R$ ；五个共有特征；三处分歧
二十八-三十一	《测度外壳》	可接受的拒绝；再分泌律；下一层外壳
三十二-三十六	《何谓课堂》十四篇 / 《课堂的本质》	提前结算；三个产物；17-29% 那处改判；COPUS 那一格；命名转轨
三十七-四十三	《临界点燃》《临界回落》《判断预占》《认知早产》《悬而未决的守护》《留空位律》	五种提前闭合的时机
四十四-四十七	《还要老师吗？》附录四	三对九的预注册设计
四十八-五十三	《还要老师吗？》总序第七节、结语第六节	三条限度与三个未回答的问题

一条说明。上述站内文章多数出自本书作者带出的学员，或由作者与他们合作产出。本书引用它们不是作为权威，而是作为同一条机制在不同场里被独立撞见的证据——这正是第四十一章第十四节所说“本书剩下什么是新的”那一节里，第三样东西的经验来源。

结语：写给三年之后的读者

这本书写于二〇二六年夏天。如果你是在三年之后读到它，那么有几件事值得说明，因为它们决定了这本书还剩下多少用处。

第一件：请先去看第三十九章的三条撤稿条件。

那三条是这样写的：如果把 AI 配置成对手的那一组，在传统读数与新读数上全面输给配置成答案源的那一组，第十二章即被推翻；如果元层解锁与去光滑化之间的交互效应不显著，第十七章的封写机制须撤回；如果走过新路的孩子在复现题上显著更差，第二十五章必须改写成取舍关系。

写这本书的时候，这三条都还没有被检验。如果你在三年之后手里已经有了数据，请先用它们去打这三条，再决定这本书的其余部分值不值得读。**一本自己写下了处死方式的书，读者有义务执行它。**

第二件：请留意技术那一侧是否已经变了。

第三十五章第六节写了一个限度：本书关于教师不可替代性的论证，依赖于当前技术的一个事实性状况——一个系统还不能可靠地读取一个孩子此刻的状态，并据此决定该不该开口。这不是一个原理性的不可能。

如果三年之后这件事已经做到了，那么第三十五章需要重写。而即便如此，还有一个问题不会因技术进步而消失：“谁有权决定”与“谁有能力决定”是两个问题。闭合权的归属不只是技术问题，也是关于由谁对这个孩子负责的问题。那个问题超出这本书的范围，但它不会自己消失。

第三件：请检查那份筛选清单。

第三十三章第六节预先写出了这套方案被筛选之后会剩下什么：会留下提问环节、几个反例、说说收获、一个计时器；会被放弃的是“十分钟内谁都不许给答案”“接受这节课没有得出结论”“顶到顶不动为止”“公开把路径让学生”。

如果你所在的学校此刻剩下的正是前面那四条，那么这次尝试已经结束了，只是还没有人宣布。**这份清单写在书里的目的，就是让这件事可以被当场认出来，而不必等到有人写总结的时候。**

第四件，也是最要紧的一件。

这本书判断错的可能性不小，我在第四十二章按可能性排了序。但有一件事我认为不太可能错，即使前面的机制全部被推翻，它仍然成立：

一个孩子卡住而无人供给的那段时间，正在以历史上从未有过的速度消失。

这一条不依赖任何机制，它是一个可以直接测量的事实——第三十八章的读数二，一支笔就能记。

至于那段时间消失之后会怎样，这本书给的是一个假说。你可以不接受这个假说。但在接受任何一个别的解释之前，请先把那个数记下来：**从你提出问题，到有人给出答案，中间经过了多少秒。**

连续记五节课。

如果那个数在十秒以内，那么无论这本书对不对，无论那所学校在推行什么改革，有一件事是确定的：在那五节课里，**没有一个孩子有机会自己走出一大步。**

那才是这本书真正担心的事。其余的一切论证，都是为了让这一件事能够被看见。

参考书目与延伸阅读

关于本书不做引注的说明

本书正文没有采用学术引注格式，这是一个自觉的选择，理由有两条。

第一，本书的主要读者是一线教师。密集的引注会显著提高阅读成本，而它对本书的论证并不构成支撑——本书的判断不是从文献推出来的，是从机制推出来的。

第二，本书对每一家相邻主张的处理方式不是引用，而是**划界**：说明它说什么、与本书重合在哪里、分界线在哪里。这些划界写在每章末尾以及第四十一章，写在正文里而不是脚注里。第四十一章第十六节说明了理由：**写在脚注里的分界，等于承认自己没有把握**。

因此下面这份书目不是本书的引证来源，而是一份延伸阅读清单：如果你想检验本书的某一处划界是否成立，这里是最该去读的东西。清单按第四十一章的十家排列。

一、教学转置与教学情境

- Yves Chevallard, 《La transposition didactique》, 1985
 - Guy Brousseau, 《Theory of Didactical Situations in Mathematics》, 1997
 - Hans Freudenthal, 《Didactical Phenomenology of Mathematical Structures》, 1983
- 第二章的光滑化与第一家在描述层面高度重合。分界见第四十一章第二节。

二、沉积与再激活

- Edmund Husserl, 《几何学的起源》（Der Ursprung der Geometrie）, 1936
 - Maurice Merleau-Ponty, 《知觉现象学》, 1945
- 第十七章与第一条几乎逐句对应。分界在于回写是可由第三方当堂检验的行为判据。

三、理解的层次

- Richard Skemp, “Relational Understanding and Instrumental Understanding”, *Mathematics Teaching*, 1976
 - David Ausubel, 《Educational Psychology: A Cognitive View》, 1968
 - George Posner 等, “Accommodation of a Scientific Conception”, *Science Education*, 1982
- 分界：状态与事件。见第十六章第九节。

四、双环学习与组织防卫

- Chris Argyris & Donald Schön, 《Theory in Practice》, 1974
- Chris Argyris & Donald Schön, 《Organizational Learning》, 1978
- Gregory Bateson, 《Steps to an Ecology of Mind》, 1972（其中关于 deuterio-learning 的论述）

分界：那一传统预测防卫性行为，本书预测顺从性行为。见第四十一章第五节。

五、本体误归类

- Michelene Chi, “Commonsense Conceptions of Emergent Processes”, *Journal of the Learning Sciences*, 2005
- Stella Vosniadou & Xenia Stafylidou, “The Development of Students' Understanding of the Numerical Value of Fractions”, *Learning and Instruction*, 2004

这是教育文献中与本书最强的竞争者。判决设计见第三十九章第三节。

六、定势与固着

- Abraham Luchins, “Mechanization in Problem Solving: The Effect of Einstellung”, *Psychological Monographs*, 1942

七、同化、顺应与图式

- Jean Piaget, 《发生认识论原理》
- Frederic Bartlett, 《Remembering》, 1932

分界：不是少了一次顺应，是多了一次错的。见第四十一章第八节。

八、先在感与物化

- Immanuel Kant, 《纯粹理性批判》“先验辩证论”部分
- Pierre Bourdieu, 《Outline of a Theory of Practice》, 1977（doxa 与“起源的失忆”）
- Peter Berger & Thomas Luckmann, 《The Social Construction of Reality》, 1966（物化）
- Bruno Latour, 《Science in Action》, 1987（黑箱化）
- Karl Popper, 《Objective Knowledge》, 1972（世界三）

第十七章第七节的两道挡板，正是针对这一族设的。

九、困难、失败与准备性学习

- Robert Bjork, “Memory and Metamemory Considerations in the Training of Human Beings”, 1994（desirable difficulties）
- Manu Kapur, “Productive Failure”, *Cognition and Instruction*, 2008
- Daniel Schwartz & John Bransford, “A Time for Telling”, *Cognition and Instruction*, 1998

这一族与本书第十九章方向一致，但对象不同：那些困难是被设计进去的，本书说的是一种正在被无意取消的困难。第三家尤其值得读——它处理的正是“什么时候该告诉”这个问题。

十、流畅性错觉与测量的反作用

- Louis Deslauriers 等, “Measuring Actual Learning versus Feeling of Learning”, *PNAS*, 2019
- James C. Scott, 《Seeing Like a State》, 1998

— Wendy Espeland & Michael Sauder, “Rankings and Reactivity”, *American Journal of Sociology*, 2007

第十三章的清晰度错觉与第一家直接相关；第八章与第三十八章的读数纪律，根据在后两家。

附：与课堂等待时间直接相关的一条

— Mary Budd Rowe, “Wait-Time and Rewards as Instructional Variables”, *Journal of Research in Science Teaching*, 1974

这是本书读数二最近的先行者。分界见第三十二章第八节：那一族以秒计并作为教师技能来训练，本书以分钟计并作为公开宣布的规矩来执行。

附：关于一对一教学效果的经典估计

— Benjamin Bloom, “The 2 Sigma Problem”, *Educational Researcher*, 1984

第三章的成本论与这一篇处理的是同一个困境的两面：那一篇问“如何用可负担的方式达到一对一的效果”，本书问“当一对一第一次变得可负担时，应当用它做什么”。

附录

以下五份材料是从正文中抽出的可复印页。允许一线学校为非商业的内部教研目的复印使用。

附录一 教师禁令表

左栏是老路上被视为美德的动作，在新路上恰恰需要克制；右栏是替代动作。

不要说 / 不要做	改成
“我来给大家理一理”	“谁来说说你现在是怎么想的”
“其实很简单，你看”	“你这个想法哪里让你自己不放心”
学生一卡住就补一句提示	数到三十秒再决定要不要开口
“对，就是这样”（当场确认）	“还有谁的想法跟他不一样”
把 AI 的回答念给全班听	让学生自己去顶，只念出分歧处
下课前把结论工整地板书一遍	下课前让学生改一件旧的
“这个我们下学期会学到”	“那你现在能不能先猜一个”
把反例的意义讲清楚	只问“那我们刚才那一下靠的是什么”

出处：第三十四章第八节。

附录二 教研组自查十问

- 每问只答“是”或“否”，不解释，不打分。
- 一、这节课的结论，是不是在上课前就写在教案上了？
 - 二、我抛出的那个问题，输进任何一个搜索框，一分钟内能不能得到大家都点头的答案？
 - 三、从我提出问题到有人给出答案，中间空了几分钟？
 - 四、这节课里，学生对 AI 说的第一句话是“这个怎么做”，还是“你这里不对”？
 - 五、有没有哪一刻，是学生在走一条我没有预设的路？
 - 六、有没有哪一刻，是我在提供土壤而不是提供路径？
 - 七、全班的结论是不是都一样？如果是，这是必然的，还是被我引导的？
 - 八、下课前，有没有一个学生回头改了一件旧的？改了什么？
 - 九、黑板上最后留下的那句话，是我带来的、AI 带来的，还是撞出来的？
 - 十、如果把 AI 从这节课里撤掉，这节课会不会变得更好？
- 第十问最锋利：一个配置正确的课堂，撤掉 AI 之后应当明显变弱。如果答案是“撤掉会更好”，说明它占的是答案源的位置——这时候正确的处理不是撤掉它，是改配置。
- 出处：第四十章第九节。

附录三 五个读数记录表

项目	频率	谁来记	记什么	用途
未闭合时长	每节课	录音自动	秒数	教研讨论
对 AI 的第一句	每次使用	系统自动	索取型／对抗型比例	班级聚合
回写率	每单元	与作业合并	迁移题减复现题正确率	诊断
先在感指数	每学期两次	五分钟量表	班级均值变化	前后对照
空位题反应	每学期一次	半节课	走出一步的人数比例	诊断

- 三条使用纪律：
- 一、用于诊断，不用于排名。
 - 二、不单独使用，必须配对（读数二与读数四；读数一与读数三）。
 - 三、未闭合时长只设下限，不设上限。
- 最低可行版：只做前两项。两项都自动采集，不占用一分钟教学时间。
- 出处：第三十八章。

附录四 一节课的四段结构

段	时长	做什么	最常见的失败
一、开缺口	8 分钟	抛出岔口，当众立规矩	问题不合格，气氛顺畅
二、顶	15 分钟	学生用旧账去顶，顶不动为止	教师在第四十秒开口
三、反例	10 分钟	放进反例， 放进去不解释	忍不住把意义讲清楚
四、回写	7 分钟	让他改一件旧的， 念出来	省略，改成“小结一下”

立规矩的原话：“从现在开始到我说停，谁都不许给答案——包括我，包括 AI，包括课本后面那一页。”

学生对 AI 只能说两种话：“我是这么想的，你说我哪里不对。”“要是换成……你还成不成立？”

元层解锁三问：“如果不这样会怎么样？”“当年是不是差一点就选了别的？”“你能不能造一个？”

一条实用建议：倒着排时间。先从下课时间往前留出七分钟。否则第四段一定会被吃掉。

如果只做一件事：把最后七分钟的小结换成“改一件旧的，然后念出来”，其余全部不动。

出处：第三十四章。

附录五 三条撤稿条件

本书自己写下的处死方式，供检验者使用。

条件一（针对第十二章·加固悖论）。把 AI 配置为答案源的 A 组，与配置为对手的 B 组，同一学期同一内容。若 A 组在传统读数与新读数上**全面胜出**，第十二章即被推翻，第二编大部分内容应当删除。

条件二（针对第十七章·封写）。是否做对象层去光滑化 × 是否先做元层解锁，2×2 设计。若**交互项不显著**，或效应量低于中等水平，封写机制须撤回。

⚠ 明确堵死的逃生解释：不得以“去光滑化做得不够好”来解释失败。理由是本书自己主张“路走对了回写照样可能不来”，因此没有资格反过来用“路没走对”为自己开脱。

条件三（针对第二十五章·子集关系）。若走过新路的学生在**复现题上显著更差**，“子集”说法即被推翻，第二十五章必须改写为取舍关系。

本书最容易被打掉、同时也最原创的一条：第十七章推论二——最配合、学得最好的学生被封写得最彻底。若 A 组中成绩最好的那部分学生，在新读数上的表现不低于成绩中等的学生该推论即被推翻。

出处：第三十九章。

附录六 全书判断一览

四十九章，每章一条承重判断，按顺序列出，供快速回查。

第一编·老路

- 一、教学是一条单向搬运链，链上流动的全部是成品。
- 二、发现=发生—路径—土壤；发生不可传递。
- 三、三千小时买不起，所以老路不是错的，是唯一负担得起的。
- 四、教科书的第一句，正是困难被删掉之后剩下的。
- 五、它不需要有人相信它，只需要有人测量它。
- 六、不看动作看终点：结论写在教案上了吗。
- 七、光滑化用“不属于任何人”换来了“属于所有人”。
- 八、这套读数只能测出自身的成功。

第二编·包办

- 九、被接管的是教学转置这个判断行为；被取消的是教材这个形态。
- 十、放弃的不是能力，是用途。
- 十一、瓶颈不是孩子，是供给的及时性；四道摩擦一起消失。
- 十二、加固悖论：一边被走完，一边显得越来越对。
- 十三、每一项优化都在消除某种粗糙，而某些粗糙是唯一的入口。
- 十四、香和毒是同一样东西；效率就是剂量。
- 十五、接通已有路网只能到达已被命名的地方。

第三编·损害

- 十六、判据不是复述新的，是回头改一件旧的。
- 十七、回写杀死了回写——写下的是写保护，不是销毁写头。
- 十八、先查第四种死亡，再查前三种。
- 十九、毒性=老路×速度÷摩擦。
- 二十、损害的分布是翻转的，落在最可能开辟新路的那批孩子身上。
- 二十一、先涨、再平、再不可逆。
- 二十二、“没有路通到”被读成“不存在”。

第四编·两条赛道

- 二十三、知识史上是派生，学习者身上是两条赛道。
- 二十四、三条判别：终点、计分、闸门；最要紧的是计分。
- 二十五、老路的能力是新路能力的子集。
- 二十六、没有第二套秤，讲什么都没用。
- 二十七、绝大多数改革死在第三周。

二十八、这不是教学法改革，是读数体系的增补；不要评优。

第五编新路

二十九、合格的问题：一分钟内搜不到大家都点头的答案。

三十、第三方的作用是把“现成的说法”从教师手里拿走。

三十一、位置由材料决定，不由身份决定。

三十二、闭合权不轮换；任一方不得占两格。

三十三、塌陷方向固定：学生先退场。

三十四、四段结构；若只做一件事，做第四段。

三十五、教师、学生、教室是产物，不是座位。

三十六、五科各有藏身处（上）。

三十七、两处最该防的配错都在创作类学科（下）。

第六编验证与边界

三十八、第一设计原则不是准确，是便宜。

三十九、本书写明了自己在什么条件下是错的。

四十、失败模式必须被预先命名。

四十一、本书剩下的原创只有三样。

四十二、论证密度高于证据密度——这是自觉的选择。

四十三、AI 不是发生学教育的敌人，它是它第一次成为可能的条件。

第七编补编

四十四、六条论证：为什么“赋能”必然自败

四十五、判定时刻： t_J 与 t_O ，以及加速—封闭对偶律

四十六、效率自证悖论与那个 0.989

四十七、二十四个产品形态的编码审计

四十八、没有人当期受损：谎言不需要有人撒谎

四十九、六个学科现场

附录七 材料层索引

正文各编末尾附有材料层，性质为示例、推演或站内取料，与论证层分开排印。逐条列出：

— 第一编材料：十二个标本与一节课的成品清点

— 第一编材料（续）：测度外壳与再分泌律

— 第二编材料：一份学期数据画像、一段对话的逐轮标注与二十条宣传语

— 第三编材料：四种死亡的四份诊断记录与一份封写征候清单

— 第三编材料（续）：四个行业已经先演过一遍

— 第三编材料（再续）：提前闭合的五种时机

- 第四编材料：一所学校的换道时间线与三场对话实录
- 第三十章 三方碰撞：各带别人拿不出的材料
- 第三十一章 位置轮换：由材料决定，不由身份决定
- 第五编材料：三节完整课例与各科岔口开采清单
- 第五编材料（续）：课堂是被提前结算的地方
- 第五编材料（再续）：老知识重新发生之外，新知识怎么发生
- 第六编材料：一学期读数记录样表与一次教研会实录
- 第六编材料（续）：一份可以照做的预注册方案
- 第六编材料（再续）：站内三次真跑，其中两次推翻了自己
- 全书补料：三条限度与三个未回答的问题

材料层的三条阅读纪律：

- 一、**课例与数字一律为示例或推演**，不是实测数据；每份材料开头都写明了性质。
- 二、**站内取料只取机制与量化形态**，论证过程与案例现场在原作；出处见材料五十四。
- 三、**材料层中有四处逼出了对正文的修订**（第十七章新增第八节、第二十六章那句的后半、第三十八章读数二拆成两个、第四十一章第十四节的自我限定），修订已落进正文，材料层保留了逼出它的过程。

这本书处理一件已经发生、
但还没有被命名的事。

大模型把传统教学赖以立身的三个动作——搬运、传道、
接受——整条接管了，而且做得更快、更全、更便宜。

这不是「教学效率被提升」，是「教学的那条老路被走完了」。
更麻烦的是：它一边被走完，一边在所有读数上显得越来越对。

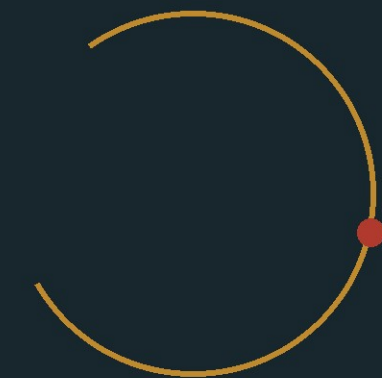
本书给出这一悖论的机制（加固悖论）、它落在孩子身上的
形态（回写杀死了回写）、它最受推崇也最纯粹的样子
（所谓「AI 高效自学」），以及一条可以走的新路——
教师、学生、AI 三方在三个位置上轮换碰撞。

全书四十三章，每一章只有一条承重判断、一个当堂可做的
动作、一条与最近邻居的分界线；并在正文里写明了
在什么条件下，这本书应当被推翻。

发现学教育不是错的。它是过去唯一负担得起的。

不是不让他用 AI，
是不让任何人在他想出来之前替他想。

**AI 不是发生学教育的敌人——
它是发生学教育第一次成为可能的条件。**



德麦国际出版社 Demai International Press
教育 / 教学理论 / 人工智能与学习

ISBN 979-8-90690-036-4

US\$ 24.50