

下一步该做什么，有几个答案

论课堂运行的次序应当按倒推的解的个数分段，及为什么「目标明确」与「下一步可算」是两件不相干的事

物理学 × 心理学 × 工程学

作者 王德生 + Claude

栏目 学科通融 · 站外碰撞

篇幅 约1.9万字 · 25页

发表 2026年8月3日

摘 要

课堂如何运行才能让学习真正发生？三门学科给出三条互不相容的次序。物理学的变分传统说，路径不是一步步试出来的，是由两端条件整体决定的，因此定住两端、路径自己出来；心理学的刻意练习传统说反了：路径只能一步一步走，而**学习者本人无法确定自己的下一步**，这正是教练存在的理由；工程学的稳健设计传统说两家都在追一个不该追的东西——**最优点恰恰是最脆弱的点**，要做的是把工作区放在平坦处，代价就是主动放弃最优。刻意练习要最大化的正是稳健设计要牺牲的。三家共同站着而从未说出的那块地是：**从当前状态到目标，「下一步该做什么」是一个有确定答案的问题，分歧只在该怎么求这个答案**。推翻它的材料来自物理学自己：同一个动力系统，**写成初值问题在很弱的条件下解存在且唯一，写成两点边值问题则可能无解、也可能有无穷多解**；物理学与数值分析据此选择方法，而这条在那里回答的是「用打靶法还是直接积分」。**倒推与推进不是同一个动作的两个方向，它们的适定性根本不同**。由此命名**可倒推区间**（唯一／多解／无解三档），判据一句：**把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的「下一步」是同一件事、五件不同的事、还是都说不出来？**一次教研活动即可测。读数与「目标是否明确」正交——目标写得最清楚的课，倒推可以完全无解。文中另推出一条不对称：**错配的代价在两个方向上不相等，因此在档位不确定时应当默认推进而非倒推**。

这一篇由哪三个学科的理论体系撞成

第一家主张实际路径由两端条件与一个整体量共同决定，逐步试探不是求解的方式；第二家主张可迁移的能力只能通过针对当前弱点的、有即时反馈的重复获得，而学习者本人看不见自己的下一步；第三家主张性能最好的那个点也是对参数变化最敏感的点，因此不该追最优，该选对扰动不敏感的区域。三条互不相容：变分说的整体决定要求两端都给定，而刻意练习的工作前提是终点只是一个方向；刻意练习最大化的是特定条件下的表现峰值，而这正是稳健设计要主动削掉的东西。三家均为站外的公开文献，链接直达原始出处，可自行核对。

物理学 · 变分原理，与边值问题和初值问题的适定性差异

[Gray & Taylor, When Action Is Not Least \(*American Journal of Physics* 75:5, 2007\)](#); 适定性一侧见 [Kierzenka & Shampine, A BVP Solver Based on Residual Control and the MATLAB PSE \(*ACM Transactions on Mathematical Software* 27:3, 2001\)](#)

变分表述把路径写成由两端条件决定的整体量，且这一支自己指出实际路径常常只是驻点、不是极小；**而更要紧的一条是它旁边那个不对称——初值问题在很弱的条件下解存在唯一，两点边值问题可以无解、可以有无穷多解**。这条在物理与数值分析里回答的是方法选择问题，从没有人拿它当「该不该先定终点」的证据。

<https://doi.org/10.1119/1.2710480>

心理学 · 刻意练习

[Ericsson, Krampe & Tesch-Römer, The Role of Deliberate Practice in the Acquisition of Expert Performance \(<i>Psychological Review</i> 100:3, 1993\)](#)

主张顶尖水平由结构化训练时数而非总年限决定，而结构化的定义里最关键的一条是**任务由教练根据当前表现指定**；这使「专家能说出下一步」成为这一支能够做研究的前提。它自己承认效应在缺少成熟教练体系与子技能分解的领域显著减弱，而它把这个减弱读作有待改进的现状，不读作一个变量。

<https://doi.org/10.1037/0033-295X.100.3.363>

工程学 · 稳健参数设计

[Nair \(ed.\), Taguchi's Parameter Design: A Panel Discussion \(<i>Technometrics</i> 34:2, 1992\)](#)

主张寻找的不是性能均值最高的点而是性能对不可控扰动不敏感的区域，二者通常相距很远；该讨论本身是这一支内部一场公开的争论。它自己的边界也清楚：稳健设计要求性能可测、扰动可枚举——当目标本身说不清是什么时，「对扰动不敏感」没有定义。

<https://doi.org/10.2307/1269231>

目 录

摘 要	5
一、三条次序，第一步就分开	5
二、物理学这一家：路径由两端整体决定	6
三、心理学这一家：只能一步一步，而且他自己看不见下一步	6
四、工程学这一家：最优点就是最脆的点	7
五、三家共同站着的那块地	7
六、推翻它的材料，来自物理学自己	7

七、命名：可倒推区间	8
八、判据：一句可以拿去问人而不问判断的话	9
九、由此得到的完整次序	9
十、形式部分：错配的代价在两个方向上不对称	10
十一、辨别格	11
十二、走查一节课：三个档位在同一节课里同时出现	12
十三、十处兑现	13
十四、强制不一致：教育目标分类学守不住的两条	14
十五、三处反向约束	15
十六、三家为何各自到不了这一条	15
十七、与既有说法划界	16
十八、与本栏同题篇目的划界	17
十九、与同日另两篇的划界	19
二十、五种最强的竞争解释，与排除它们的设计	19
二十一、分层引用结构	20
二十二、证伪与赌注	20
二十三、反噬与治理	22
二十四、若本判断被证伪，我们会学到什么	22
二十五、自反检验	23
二十六、结论	23
注 释	23
参 考 文 献	24

摘要

课堂如何运行才能让学习真正发生？三门学科给出三条完整而互不相容的次序。物理学的变分传统说，一个系统走的路径不是一步步试出来的，是由两端条件整体决定的，因此正确的做法是把两端定住，路径自己出来。心理学的刻意练习传统说反了：路径只能一步一步走，每一步都要有针对当前弱点的任务和即时反馈，整体最优是走出来的不是算出来的。工程学的稳健设计传统说两家都在追一个不该追的东西：最优点恰恰是最脆弱的点，要做的是把条件场做得足够宽，让系统在扰动下仍然可用，代价就是主动放弃最优。

三条不能同真——刻意练习要最大化的正是稳健设计要牺牲的；而变分说的"整体决定"需要两端都给定，刻意练习却只给定一端。三家共同站着而从未说出的那块地是：**从当前状态到目标状态，"下一步该做什么"是一个有确定答案的问题；分歧只在于该怎么求这个答案。**

推翻它的材料来自物理学自己。同一个动力系统，写成初值问题在很弱的条件下就保证解存在且唯一；写成两点边值问题则可能无解，也可能有无穷多解。物理学与数值分析对此毫不含糊，并据此选择方法——而这条在物理里回答的是"该用打靶法还是直接积分"，是一个计算问题，从来没有人把它当作"该不该先定终点"的证据。**倒推与推进不是同一个动作的两个方向，它们的适定性根本不同。**

由此本文命名**可倒推区间**：学习进程里的一段区间，在其中"要到达那个目标，下一步该做什么"有且只有一个答案。出了这段区间，倒推要么无解、要么多解，而这三种情形要求三种不同的课堂运行方式。判据一句：**把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的"下一步"是同一件事、五件不同的事、还是都说不出来？** 读数**倒推唯一性**可直接测量，数据在多数机构里当天就能采到，且与"目标是否明确"正交——目标写得最清楚的课，倒推可以完全无解。

关键词：课堂运行；可倒推区间；倒推唯一性；边值与初值；刻意练习；稳健设计

一、三条次序，第一步就分开

"课堂如何运行"要的是一条次序：从哪儿起，经过什么，实现什么，在哪一步能插手。三门学科各有一条，而且都不是空谈。

物理学的变分传统给的次序是：定住两端，路径自己出来。这条次序在物理里有极强的地位——力学、光学、场论都可以从一个作用量泛函的驻点条件重新导出。它的教学版本是：把起点（学生现在会什么）与终点（学完要会什么）都写死，中间的安排是可以被求出来的。

心理学的刻意练习传统给的次序相反：路径只能一步一步走。每一步都要有一个刚好超出当前能力的任务、一个即时反馈、一次针对性重复。整体最优不是算出来的，是在无数次局部改进里长出来的。这条次

序背后有一个硬主张：**人无法从终点倒推出自己的下一步**，因为他不知道自己缺什么——这正是教练存在的理由。

工程学的稳健设计传统给第三条，并且把前两家一起否掉：追最优本身是错的。参数空间里性能最好的那个点，往往也是性能对参数变化最敏感的那个点；把系统放在那里，任何扰动都会让它掉下来。正确的做法是选一个性能稍差而对扰动不敏感的区域，把工作点放在平坦处。它的教学版本是：不要把学生训练到只能在标准条件下发挥最好，要把他训练到在条件变化时仍然能用。

三条互不相容。刻意练习要最大化的是特定条件下的表现峰值，这正是稳健设计要主动牺牲的东西；变分说的"整体决定"要求两端都给定，而刻意练习的整个工作前提是终点只是一个方向、不是一个可以代入计算的边界条件；稳健设计则说两端定得越死，做出来的东西越脆。

二、物理学这一家：路径由两端整体决定

把第一家推到最强：一个系统的实际路径不是逐步试探的结果，是由两端条件与一个整体量共同决定的。

这不是一个哲学立场，它有极硬的形式。给定起点与终点，实际路径是作用量泛函的驻点；由此可以导出运动方程，而且导出的方程与从局部受力分析得到的完全一致。这个等价性本身很了不起——它说明"整体决定"与"逐步推进"在这一类系统里是同一件事的两种写法。

它对课堂的主张因此很直接：把学生现在的状态与要到达的状态都刻画清楚，中间该怎么走是可以被求解的，不需要靠试。教学设计里所有"倒推"的做法，其数学骨架就是这一条。

这一家自己知道的一件事，推翻材料一节会用到，这里先按下不表：**这个等价性是有条件的，而条件恰恰落在"两端都给定"上。**

三、心理学这一家：只能一步一步，而且他自己看不见下一步

把第二家推到最强：可迁移的能力只能通过针对当前弱点的、有即时反馈的重复获得，而**学习者本人无法确定自己的下一步。**

这一条的证据非常集中。在有明确评价标准、有长期训练传统的领域里，顶尖水平与训练年限的关系远不如与"结构化训练时数"的关系强；而结构化训练的定义里，最关键的一条正是任务由教练根据当前表现指定，而不是由学习者自选。

这一家对物理学那条次序的诊断是：两端根本没有都给定。学生的当前状态不是一个可以代入的初值——他自己说不清，教师也只能通过表现推断；而终点也不是一个坐标，是一个方向。在两端都是模糊的情况下谈"整体求解"，求出来的是纸面上的路线。

它对稳健设计的诊断更狠：在能力还没有长出来之前谈鲁棒，等于在还没有峰值的时候就开始削峰。

这一家自己承认的边界很清楚，而且写在它自己的文献里：刻意练习的效应在有稳定评价标准、有成熟教练体系、有明确子技能分解的领域最强，在缺少这些条件的领域显著减弱。它把这个减弱解释为"这些领域的训练还不够结构化"——也就是说，它把边界当作了有待改进的现状，而不是当作一个变量。

四、工程学这一家：最优点就是最脆的点

把第三家推到最强：设计的目标不是让性能达到最优，是让性能对不可控因素不敏感。

这一条的形式同样很硬。把可控参数与不可控扰动分开，在参数空间里寻找的不是性能均值最高的点，而是性能方差最小的区域；这两个位置通常不重合，而且往往相距很远。工程实践里大量的经验是：把工作点放在性能曲线陡峭的地方，实验室里表现最好，产线上一塌糊涂。

它对课堂的主张是：不要把学生练到只在标准题型、标准条件、标准提示下发挥最好。要找的是那个"换了条件仍然能用"的平坦区域，哪怕在标准条件下表现差一截。

它对另外两家的诊断合成一句：**两家都在求最优，一家用解析的方式求，一家用迭代的方式求，而最优本身是错的目标。**

这一家自己的边界也清楚：稳健设计要求性能可测、扰动可枚举。当目标本身说不清是什么的时候，"对扰动不敏感"就没有定义——因为你不知道对什么不敏感。

五、三家共同站着的那块地

三家争的是"下一步该做什么"这个答案该怎么求：解析地求、迭代地走、还是干脆不求最优只求平坦。

争论再激烈，三家都默认了同一件事：**这个问题有答案。** 变分说答案可以算出来；刻意练习说答案只能由教练给出但确实存在；稳健设计说答案是一个区域而不是一个点，但区域也是答案。三家的处方全部建立在这个假定上——如果"下一步该做什么"根本没有答案，或者有无穷多个同样好的答案，那么"该怎么求"就不成其为问题。

把这个前提念给三家听，三家都会说这还用说吗。学生现在在这儿，目标在那儿，当然存在一件此刻该做的事。这正是共有前提的签名。

六、推翻它的材料，来自物理学自己

物理学与数值分析对下面这件事毫不含糊。

同一个二阶动力系统，可以写成两种问题。

初值问题：给定初始位置与初始速度，往前积分。在很弱的正则性条件下，解存在且唯一。这是一个适定问题——初值的小扰动只带来解的小变化。

两点边值问题：给定起点位置与终点位置，求中间的路径。这个问题**可以没有解**（在给定时间内根本到不了那个位置），**也可以有无穷多个解**（好几条路径都恰好在那一刻到达），而且即使有唯一解，它对边界条件的敏感度也可能高到没有实用价值。数值上要用打靶法反复试探，且不保证收敛。

这个不对称是数学事实，不是工程经验。而且它就在第二家那一节那个"等价性"的旁边：变分原理确实能导出与局部方程相同的运动方程，但**变分表述本身是一个边值表述**——它假定两端都给定。一旦你只知道一端，那个漂亮的整体求解就退化成一个必须从这一端往前推的问题。

物理学在自己的领域里用这条差别回答的是**该选哪种数值方法**：知道初值就直接积分，知道两端就打靶，打靶不收敛就换方法或者承认这个边值问题不适定。这是一个计算策略问题，被写进每一本数值方法教材。**没有人把它当作"该不该先定终点"的证据——而它恰恰是。**

它推翻的正是共有前提。"下一步该做什么"这个问题不是必然有答案的：它是一个从终点倒推的问题，而倒推问题的适定性不是给定的，是随具体的起点、终点与可用时间变化的。**同一门课，在某些区间上倒推有唯一解，在另一些区间上无解，在第三些区间上有无穷多解——而三家的处方各自只在其中一种情形下成立。**

七、命名：可倒推区间

由此本文命名一类区间。

可倒推区间：一段学习进程，在其中"要到达那个目标，此刻该做什么"这个问题有且只有一个答案。

区间的三种状态：

唯一。问五个懂行的人，五个人说同一件事。这类区间通常出现在有稳定子技能分解、有明确失败模式、有成熟诊断传统的内容上。

多解。问五个懂行的人，得到五件不同的事，而且每一件都说得通、都能到。这不是"没人知道"，是**确实有多条路都能到**。

无解。问五个懂行的人，五个人都说不出一件具体的下一步，只能说出活动的名字（"多读多写""多做练习""培养兴趣"）。这类区间的签名很好认：给出的"下一步"如果换成任何别的目标也一样说得通，那它就不是下一步。

三个条件把这个构念钉住：

它是**知识分布的性质，不是学生的性质**。同一个学生、同一个目标，换一批教师，答案的个数可以从多解变唯一——因为那批人手里有更细的诊断分解。这一条把它与"这个学生学不学得会"彻底分开。

它是**分段的，不是全局的**。一门课在不同阶段落在不同状态里，而且状态会随进度变化；把整门课判成一种状态是错的。

它与目标是否明确正交。这是最容易被误读的一点，值得单说：一门目标写得极其清楚、行为动词与评价标准一一对应的课，倒推可以完全无解——因为"能够批判性地评价一个论证"写得再清楚，也没有人说得出此刻这个学生的下一步是哪一件具体的事。反过来，一门目标只写了一句"把这首曲子弹下来"的课，倒推可以是唯一的——教师一听就知道下一步该练哪四小节。

八、判据：一句可以拿去问人而不问判断的话

> 把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的"下一步"是同一件事、五件不同的事、还是都说不出来？

这句话不含情态词，不问该怎么教，只问答案的个数。它可以在一节课的时间里实测，成本接近零：写一份学生当前作业的匿名样本，附上目标，发给五个人，收回来数一数。

五个场景跑一遍，答案互不相同：

代数变形与解方程。 唯一。五个人会说同一件事，通常连要练哪一类变形都一致。这解释了为什么刻意练习与掌握学习在这类内容上效果最强——它们本来就是为唯一区设计的方法。

写一篇有论证的文章。 多解。五个人会给出五件不同的事：先改结构、先补证据、先删冗余、先明确读者、先重写开头。每一件都能到，而且都有人做成过。这不是标准不清，是路径确实多。

乐器或运动上的一个具体卡点。 唯一，而且一致程度惊人。这一条与第二个场景形成对照：同一个学生在"整体音乐性"这个目标上是多解的，在"这个换把总是慢半拍"上是唯一的。同一门课的不同区间落在不同状态里，这条对照就是证据。

"培养批判性思维"。 无解。五个人给出的是活动名称——多讨论、读一手材料、做辩论。把目标换成"培养创造力"，同样这五句话照样说得通。这是查出来的：教师给出的东西通过不了"换个目标还成不成立"这一关。

临床诊断推理的训练。 多解，且专家之间的路径差异远大于结论差异——已有的专家—新手研究反复记录过这一点，通常被解释为"专家有多种表征"。同一份材料在本文这里是另一个读数：这个区间的倒推有多解。这也是查出来的，不是从命题推出来的。

后两个答案不能从命题推出来，它们是拿现成的文献与现成的教研记录数出来的。

九、由此得到的完整次序

这是一道要"怎么办"的题，因此本文必须交出一条完整的次序，而不只是一个名字。

第一步：测。 在一个单元开始之前，用判据一节那句话测一次倒推唯一性。成本是五个人各花十分钟。得到三档之一。

第二步：按档换运行方式。 三档要求三种不同的课堂运行，而且它们互不通用。

在**唯一区**，用倒推。逆向设计、掌握学习、刻意练习、即时反馈——这些方法在这里全部成立，因为倒推问题在这里适定。这一档的插手点在诊断的精度上：唯一性越高，教练的价值越大。

在**多解区**，不要倒推，要推进加选择。从当前状态出发往前走一小段，看走出了什么，再在若干条路里选。这一档的插手点不在"指定下一步"，在**把多条路同时可见**——让学生看到别人选了不同的路且都走到了。多解区里最常见的失误是教师把自己那条路当作唯一的路讲，这在形式上完全合规，而它把一个多解问题假装成了唯一问题。

在**无解区**，两种方法都空转，必须先动目标。把说不出下一步的目标改写成能说出下一步的形式，改写之后重测。这一步是本文最有用的一处，因为它把一个通常被归为"教师能力不足"的现象重新定位为一个**可测的、目标层面的问题**。若改写之后仍然无解，则这个目标在当前的知识条件下不可教，应当明说，而不是继续用活动填满时间。

第三步：在边界处插手。 区间会移动。一个多解区在学生积累了足够的判断之后可能收敛为唯一区；一个唯一区在学生越过基础之后会展开为多解区。**边界的移动本身是学习的读数**，而且它比表现更早出现——这是证伪一节那条正向读数的来源。

十、形式部分：错配的代价在两个方向上不对称

前面几节给了三档与三种运行方式。本节回答一个前面没有回答的问题：**如果测不准档位，该默认用哪一种？** 这个问题的答案不是"再测一次"——它可以从错配的代价结构里直接推出来，而推出来的结果与通行做法相反。

设一段学习工作，要从当前状态到达目标。记倒推的解的个数为 n 。

在**唯一区用推进法**（ $n = 1$ 而误用推进）。推进法从当前状态往前走一小段，看走出了什么，再选。由于此处只有一条路通到目标，走出来的东西迟早会落回那条路上。损失是**搜索的冗余**：期望多花的步数正比于分支因子与走错后的返回成本，而且它是**有界的**——最坏情形是把该走的那条路晚一点找到。记这个损失为 L_1 ，它随时间线性累积。

在**多解区用倒推法**（ $n > 1$ 而误用倒推）。教师从目标倒推，得到的是 n 条路之一；他会把这一条当作**那一条**讲出来。学生照做，也能到。损失不在效率上——效率甚至更高，因为省掉了选择。损失在**可达集的收缩**：学生学到的是一条路，而不是"这里有多条路以及怎么选"。当条件变化、那一条路不通时，他没有任何东西可以退回来。记这个损失为 L_2 。

关键的不对称在这里： L_1 是效率损失， L_2 是**能力损失**，而且 L_2 在原场表现上完全不可见——学生在这门课的所有考核里都能做对，因为考核也走那一条路。

在**无解区用任何一种**（ $n = 0$ ）。倒推无解，推进没有方向。两种方法的产出都不是学习，而是活动。损失是**全部投入**，记为 L_3 ，且 L_3 与投入时长成正比、与方法无关。

现在把它写成一个决策问题。设档位判定有误差：真实档位为 n ，测得档位为 $\hat{n}$ 。选择方法 $M \in \{\text{倒推}, \text{推进}\}$ 。期望损失

$$E[L] = \sum_n P(n | \hat{n}) \cdot L(M, n)$$

在 $\hat{n}$ 判为唯一而实际可能是多解的情形下（这是最常见的误判方向，因为教师本人熟悉那一条路，很容易把多解读成唯一）：

$$E[L | \text{选倒推}] = P(1) \cdot 0 + P(>1) \cdot L_2 \quad E[L | \text{选推进}] = P(1) \cdot L_1 + P(>1) \cdot 0$$

选推进优于选倒推，当且仅当 $P(1) \cdot L_1 < P(>1) \cdot L_2$ 。

由于 L_2 是能力损失而 L_1 是效率损失，在以迁移为目标的内容上 $L_2 \gg L_1$ ；因此即使 $P(1)$ 相当高——即使你相当有把握这是唯一区——不等式仍然成立。

这推出三条正文前面没有说过的事。

其一，在档位不确定时应当默认推进法，而不是默认倒推法。 这与通行做法相反：教学的默认设置几乎总是倒推（先讲那条路），而把推进当作特殊安排（探究课、项目课）。默认设错的代价不对称——把唯一区当多解区处理，只是慢一点；把多解区当唯一区处理，学生失去的是他本来可以拥有的可选择性，而且这件事在这门课里永远不会被发现。

其二，最危险的不是无解区，是被误判为唯一区的多解区。 无解区的损失 L_3 虽然是全部投入，但它会被发现——学生做不出来，课堂上会有信号。而误判的多解区在所有可见指标上都表现良好。因此本文的辨别格右上那一格（目标明确、倒推非唯一）里，真正的靶子不是无解那一半，是多解那一半：它更常见、更隐蔽、更容易被当作教学质量高的证据。

其三，教师的经验会系统性地把 $P(1)$ 高估。 一位在某个内容上教了十年的教师，他自己那条路已经高度自动化；他给出的“下一步”确实稳定而唯一——**对他自己而言**。判据那一节的问话之所以要问五位教师而不是问一位，正是因为单个专家的内部一致性会被误读为问题的适定性。由此得到一条可检验的预测：**同一份材料，问五位来自不同培养背景的教师，与问五位同一教研室的教师，得到的档位分布应当不同；后者会系统性地偏向唯一。** 这一条可以在一次跨校教研活动里直接测出来，而且它同时是本文读数的一个信度检验。

模型的边界要交代清楚。它假定 L_1 、 L_2 、 L_3 在一段工作内为常数，实际上它们随内容变化很大；它没有给出 L_2/L_1 的具体比值，因此不能算出临界的 $P(1)$ ；它假定误判只有一个方向，而实际上把唯一区误判为多解区也会发生（虽然代价小得多）。上面三条推论只依赖 $L_2 \gg L_1$ 这一个不等式与各量的符号，不依赖具体数值——这是本文认为它们可检验的理由，也是本文对这个模型能力的全部主张。

十一、辨别格

两根轴：目标是否明确，倒推是否有唯一解。

倒推唯一

目标明确 可训练区：方法与问题匹配，刻意练习与倒推设计在这里全部成立。

目标不明确 隐性传承：说不清要什么，但老师一看就知道下一步该练什么。学徒制的常态。

倒推非唯一（多解或无解）

齐整的空转：目标写得漂亮、对齐完美、活动饱满，而没有人说得清下一步。**本文的靶格。**

泥地：既说不清要什么，也说不出下一步。

靶格是右上，三条签名齐：短期指标全部改善（目标清晰度、对齐度、大纲完整度全是满分）；失效不产生任何信号（没有任何检查会去问“这个目标下一步是什么”）；自我加固（越规范化，目标写得越抽象越漂亮，越说不出下一步）。

左下的“隐性传承”格是关键对照，而且它推翻了一个常识：**目标不明确而倒推唯一的课，本文预测其效果优于目标明确而倒推无解的课。**这与“先把目标写清楚”这条通行做法方向相反，且可以直接检验——学徒制、师带徒、乐器教学都落在左下。

十二、走查一节课：三个档位在同一节课里同时出现

三档听起来像是三种课，实际上它们常常同时出现在同一节课的不同时刻。下面走查一节九十分钟的大学生写作课，逐段判档。

第零到二十分钟：讲评上周作业的一个共同问题——引用的证据与所支持的论点之间没有说明关系。教师给出三个例子，逐句改。这一段落在**唯一区**：把某个学生的原句拿给五位写作教师看，问“下一步该做什么”，五个人会说同一件事——在证据与论点之间补一句说明关系的话。倒推适定，教师的示范与练习都恰当，这一段本文没有任何异议。

第二十到四十五分钟：学生就本周主题各自列出提纲。有一位学生的提纲结构完整、每一部分都成立，而整篇没有一条贯穿的判断。把这份提纲交给五位教师，问下一步——答案会是五个：先想清楚你到底要说什么、先删掉一半、先找一个反例、先把最后一段挪到前面、先去读一篇写同一题目的好文章。**这一段落在多解区。**五个答案都能到，而且每一位教师都能举出自己那条路成功的学生。

本文的主张在这里落地：这一段的正确运行不是选一条讲出来，而是让五条同时可见。可执行的做法很朴素——把这五种说法都摆出来，让学生自己选一条走，并要求他记录自己选了哪一条。这不是“尊重多样性”，是因为在多解区里指定一条的代价是学生失去可选择性，而这个代价在这门课的任何考核里都不会显出来。

第四十五到七十分钟：学生按自己选的路修改。仍在多解区，教师巡视时的动作应当是问“你选了哪一条、为什么”，而不是“我建议你先……”。

第七十到九十分钟：教师提出本课程的总目标之一——“形成你自己的声音”。把这句话交给五位教师，问此刻这个学生的下一步——五位都说不出一件具体的、此刻可以开始做的事。给出的会是“多写多读”“找到你真正想说的”“不要模仿”。把目标换成“培养批判性思维”，这三句照样说得通。**这一段落在无解区。**

本文对这一段的处理不是回避，是明说：这个目标在当前的知识条件下说不出下一步，因此这二十分钟里发生的不是学习而是活动。可做的只有一件——把它改写成能说出下一步的形式。一个可行的改写是"在你的三篇作业里找出重复出现的句法习惯，并写清其中哪两条你打算保留"。改写之后重测：五位教师会说同一件事（把三篇并排读、标出重复的结构），档位从无解变成唯一。

这次改写有代价，必须诚实写出来：改写后的目标不等于原来那个目标。"找出重复的句法习惯"不是"形成自己的声音"，它至多是通向后者的一段。本文在反噬那一节已指出，改写与掉换之间没有可靠的分界，而这一节的走查正是那个风险的具体形态。区别只能靠一条：**改写后的目标与原目标要在一个独立的延迟测验上可比**；做不到这一条，就应当承认这是掉换而不是改写。

这节课的三档同时存在，而它的课程大纲把三段写成同层级三个教学环节。这是本文最想说清的一件事：档位不随课程走，随内容与知识分布走；按周次、按单元、按课时组织的教学安排，在结构上无法把三档分开。

十三、十处兑现

数值求解。 边值问题不收敛时改用延拓法或换成初值表述，这是标准操作，有对应的教科书章节；教学里没有对应的动作。

手术训练。 具体入路的下一步唯一性高，整体判断力的下一步唯一性低；训练体系对前者有细致的分解，对后者只有"多上台"。

飞行训练。 程序性科目倒推唯一（检查单本身就是倒推的产物），情境判断科目多解，而模拟机的科目分配恰好反映了这一点——这是查得到的。

编程。 "让这段跑通"唯一，"把这个系统设计好"多解；教学资源的分布与倒推唯一性高度一致，而没有人用这个词。

语言学习。 发音的某个具体音位唯一，"写得地道"无解；后者的教学长期依赖大量输入而非指定下一步，正是无解区的正确应对。

物理竞赛训练。 题型内唯一，跨题型多解；训练体系用大量刷题解决前者，而对后者只有"多想"。

产品设计。 需求已定时倒推唯一，需求未定时无解；行业的应对是先做原型把无解区变成多解区——这正是完整次序第三步的工业版本。

康复训练。 具体动作模式唯一，功能性恢复多解；康复计划的模板结构反映了这一区分。

写作指导。 句子层唯一，篇章层多解，"找到自己的声音"无解；三档在同一门课里同时存在，而课程按周次而不是按档组织。

研究生指导。 技术步骤唯一，选题多解，"成为一个研究者"无解；导师的实际做法是在第三档上用陪伴而不是指导，这在任何培养方案里都没有位置。

十处里有三处（飞行训练的科目分配、产品设计的原型策略、康复计划的模板）已经在实践中按倒推唯一性分了段，只是没有名字。

十四、强制不一致：教育目标分类学守不住的两条

教育目标分类这一支同时持有两条主张，都写在它自己的框架文件里。

甲：教育目标可以被放进一个分层的结构里，每一层都有对应的、可操作化的行为描述。 这是这一支的全部工作产品：认知过程的层级、知识类型的维度、以及为每一格提供的动词表。它的实用价值正在于可操作化——写目标的人可以照着动词表把一条模糊的期望写成一条可评估的陈述。

乙：高层级的目标（分析、评价、创造）比低层级的更值得追求，教学应当向上移动。 这条同样是这一支的核心主张，而且是它在教育实践里最有影响力的一条——几乎所有课程改革文件都引用它来论证"不能只考记忆"。

两条在一处正面相撞，而且相撞的位置恰好是这一支最看重的那一端。**层级越高，倒推唯一性越低。**"回忆某个定义"的下一步，五位教师会给同一个答案；"评价一个论证的可靠性"的下一步，五位教师给出的是活动名称——多读、多讨论、多写。**动词表提供的是目标的写法，不是下一步的算法；把"能够评价"写成一条合规的目标陈述，与知道此刻这个学生该做哪一件具体的事，是两件不同的事，而这一支的整个装置只解决前者。**

于是：甲声称可操作化，乙要求向高层移动，而**可操作化的程度随层级上升而下降**。这一支现行的处置是继续细化行为描述——把"评价"拆成"识别前提、判断相关性、检验充分性"等等。这个处置在形式上保住了两条，代价是：**拆到足够细以致下一步唯一的时候，被拆出来的那些子行为已经落回低层级了**。这不是执行不力，是拆解这个动作本身的性质——每一次拆细都在把一个多解问题换成若干个唯一问题，而"评价一个论证"这件事的难处恰恰在于**选择用哪一组子行为**，而那个选择正是被拆掉的那一步。

二选一。要么承认高层级目标在本文意义上是多解或无解的，因而不能用同一套操作化装置处理——那么甲不能以现在的形式成立，动词表必须按倒推唯一性分区，高层格子里给出的不该是动词而该是**选择的判据**；要么放弃乙，承认可教的只有那些能被操作化到唯一的東西——而这与这一支存在的全部理由冲突。

它不能同时守住两条，除非它承认本文的变量。 一旦承认，"高层级目标"这个说法就必须被拆成两样：可拆到唯一的（照现有装置处理）与拆到唯一就变质的（需要另一套方法），而区分二者需要的正是倒推唯一性这个读数。

这一支已经为此付出了可观察的代价：**大量课程大纲里的高层级目标，其配套的评价方式与低层级目标的评价方式完全相同。** 目标写着"能够评价"，考的是"能够识别下列哪一项是前提"。这一点在任何一次课程审核里都不会被标出来，因为审核的判据是目标与评价是否对齐——而它们确实对齐了，对齐在低层级上。

十五、三处反向约束

心理学削掉"唯一区就该一直用倒推"。 刻意练习在唯一区效应最强，但同一批研究也表明高度专门化的训练会使迁移变窄，并在条件变化时表现下降。因此唯一区的倒推有时间上限，且上限之后必须主动进入多解区。如果没有这一处，本文原本会写下"在唯一区应当尽可能延长倒推式训练"——这句话被削掉了。

物理学削掉"多解是缺陷"。 在控制论与力学里，从一点到另一点存在多条可行路径不是问题，是**可控性强**的表现；解的不唯一意味着可选择的余地大。因此本文必须放弃"多解需要被收敛成唯一"这条价值排序，改为"多解需要被看见并被选择"。这一处动的是价值排序，而且它改变了完整次序第二步的处方形式——从收窄改为展开。

工程学削掉"无解区可以靠稳健设计兜住"。 稳健设计要求性能可测、扰动可枚举；在目标说不清是什么的区间里，"对扰动不敏感"没有定义。因此本文的适用范围收窄：无解区没有现成的方法，只有改写目标这一条路，而改写目标不总是能成功。这一处动的是适用范围，也是本文最不愿意承认的一处。

十六、三家为何各自到不了这一条

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

物理学到不了，因为它的两端总有一端是白送的。 在力学问题里，初始条件通常是给定的——你知道这颗行星现在在哪儿、速度多少。因此"该从哪一端开始"在物理里几乎从不成为一个真问题：有初值就用初值，需要边值时（弹道、最优控制）它清楚地知道自己在解一个更难的问题，并且有一整套数值方法来对付。**适定性对物理学来说是一个计算属性，不是一个决策属性**，因为决策早就被问题本身规定了。而课堂的两端**都不是给定的**：学生当前的状态需要推断，目标只是一个方向。物理学没有遇到过两端都要自己造的情形，因此它的框架里没有"该不该先定终点"这个位置。

心理学到不了，因为它的整个方法论建立在"专家能说出下一步"上。 刻意练习的操作定义里就写着任务由教练根据当前表现指定。这不是一条可选设置，是这一支能够做研究的前提——没有教练指定的任务，就没有"结构化训练时数"这个自变量。因此当某个领域的专家说不出下一步时，这一支的读法只有一种：**这个领域的训练还不够结构化。** 它把"下一步说不出来"当作有待改进的现状，而不是当作一个可以测量、可以分档、可以据此换方法的变量。这不是疏忽——把它当变量，就等于承认自己的方法在一大类内容上不适用，而这一支恰恰是靠"它适用范围比想象的宽"立起来的。

工程学到不了，因为稳健设计的第一步就把这个问题绕过去了。 稳健设计不追踪路径，它选工作区。你问它"从这儿到那儿下一步该做什么"，它的回答是"别走那条线，把参数放在这片平坦区里"。这个回答很有效，但它成立的前提是**目标可测、扰动可枚举**——也就是说，它默认自己站在本文的唯一区或多解区，从来不处理无解区。工程学因此有一套极强的方法来对付"路难走"，而没有任何方法来对付"说不出下一步是什么"，因为在它的世界里，说不出下一步意味着规格没写清楚，而规格没写清楚是上游的事故，不是工程问题。

三处缺口合起来说明这个判断为什么必须由三家撞出来：物理学有适定性这个概念而无两端皆需自造的处境，心理学有逐步推进的完整方法而必须假定下一步存在，工程学有对付难走的整套手段而把说不出下一步排除在自己的问题域外。少任何一家，剩下两家都推不到这一条。

十七、与既有说法划界

最近发展区。 它说到的是：存在一个区域，学生独立做不到而在帮助下能做到，教学应当落在那里。分离线：它刻画的是**难度的位置**，本文刻画的是**倒推答案的个数**；两者正交——一个难度恰当的任务完全可以是无解区的（“写一篇比现在好一点的文章”难度合适而下一步说不出），一个难度过低的任务也可以是唯一区的。判定形式：若把任务难度控制住之后倒推唯一性不再预测结果，则最近发展区覆盖本文；若在难度匹配的任务之间倒推唯一性仍有区分力，则本文对。

刻意练习。 它说到的是：针对当前弱点的、有即时反馈的结构化重复。分离线：它默认下一步存在且由教练给出，因此“下一步有几个答案”在它的框架里不是变量。判定形式：若在多解区强行实施刻意练习（教练指定唯一路径）与允许多路径的对照组结果无差异，则本文的分段多余；若前者在原场表现更好而迁移更差，则本文对。**这是本文最直接的检验。**

逆向设计与掌握学习。 它们说到的是：从目标倒推评估与活动，逐级过关。分离线：它们把倒推的可解性当作给定，整个方法建立在“目标定清楚就能倒推”上。判定形式：若一门目标清晰度极高而倒推唯一性为零的课，其效果不低于目标清晰度中等而倒推唯一的课，则本文的正交主张为伪。本文预测前者更差，而这在辨别格一节里是左下与右上的对照。

形成性评价。 它说到的是：用证据调整教学。分离线：调整的前提是知道调向哪儿；在无解区里证据再充分也无处调。判定形式：形成性评价的效应量应当随倒推唯一性上升而上升；若两者无关，本文的分段对它不增量。

稳健设计（工程学，本文学科之外）。 它说到的是：选对抗动不敏感的工作区。分离线：它是关于**参数选择**的，本文是关于**求解结构**的；一个稳健的工作点仍然要有人说得出怎么走到那里。判定形式：若把“训练是否针对稳健区”控制住之后倒推唯一性不再预测结果，则本文被覆盖。

可控性与可达性（控制理论，方法学占位者）。 把本文的命题剥掉一切教育名词，剩下的形式是“从当前状态能不能到达目标状态、有几条路”——这正是可控性与可达性的形状，它在控制理论里被形式化了六十多年。这是本文最危险的一位占位者，必须正面回答。分离线：可控性是**系统的性质**——给定动力学与输入集合，可达集是客观确定的；本文的倒推唯一性是**知识分布的性质**——同一个学生、同一个目标，换一批更懂行的教师，答案的个数会变。**一个是关于世界的，一个是关于谁在看。** 判定形式：若倒推唯一性只随学生状态与目标变化、而不随评判者群体变化，则可控性覆盖本文，本文只是它的一个通俗说法；若同一份学生样本与同一个目标交给两个不同的专家群体得到不同的唯一性档位，则本文指出的是一个可控性里没有的变量。**这一条可以在一周内实测，而且本文赌后者。**

问题空间与手段一目的分析（认知科学，方法学占位者）。 它说到的是：解题是在状态空间里搜索，用差距缩小指导选择。分离线：它默认目标状态可以被代入以计算差距；无解区的定义正是差距无法计算。判定形式：若人在无解区仍能报告出稳定的"差距感"并据此选择动作，则手段一目的分析覆盖本文的无解档；若报告出的只是活动偏好而非差距，则本文对。

十八、与本栏同题篇目的划界

本栏近期有十余篇同样处理课堂学习的文章，本文必须与它们划清。总的分界先说清楚：其中大多数问的是**为什么很少发生**，要的是一个驱动；本文问的是**该怎么运行**，要的是一条次序。以下只列会被混为一谈的几篇。

与《还在动的时候就读了数》：同一组学科的两两不同取用

本栏同日另有一篇由同样这三门学科撞成的文章，处理同一个问题。距离最近，必须写在最前面并写细。

那篇处理的是**判定的时点**：课堂每隔一两周作一次"会了没有"的判定，而判定的时点由课表决定、不由被判定的那件事的重组时长决定；它主张过早的判定把仍在进行的变化冻在当时的构型上，因而命名了冻结构型，并给出双向翻转率这个读数。

先说两家共用学科的取用点：同一门物理学被两篇拿走的是两件不同的东西。 那篇取的是**玻璃态的物理老化**——被判为冻住的系统，其性质随冻住之后所等的时间持续演化，且演化的起点正是被冻住那一刻的构型；本文取的是**常微分方程边值问题与初值问题的适定性差异**，以及变分表述本身是一个边值表述这件事。前者属于凝聚态与非平衡统计物理，后者属于经典力学与数值分析，两个分支互不引用。工程学同理：那篇取传感回路与可观测性，本文取稳健参数设计。心理学那一家则确有交叠——那篇取提取诱发遗忘，本文取刻意练习，同属记忆与技能获得这一大类；本文承认这一处的距离比另两家近。

再说承重命题的分离。 那篇问的是**什么时候读**——判定的时点相对于弛豫时长是早是晚；本文问的是**读出来的东西有几个答案**——从当前状态到目标，下一步该做什么唯一、多解还是无解。一个是时间轴上的问题，一个是解空间上的问题。二者可以完全分离：一个重组时长极短、判定时点也恰当的单元，其倒推可以处在无解档；一个倒推唯一、方法匹配的单元，其判定也可以早得把学生冻在半路。

双向判定形式。 目标写着"能够评价一个论证"、五位教师都说不出下一步的那个单元，若把判定推迟四周并重测，那篇预测双向翻转率里"不会到会"那一支显著非零、判定过早是病因；本文预测延迟迁移无改善，因为推迟一个没有方向的过程不会给它方向。反过来，一个换把动作的训练单元，倒推高度唯一而教师在第三天就作了判定，本文判此处正该用倒推、无异议；那篇判判定过早。**两篇在这两格上给出不同的诊断与不同的处方**，且两个处方不可互相替代——推迟判定不会让说不出的下一步变得说得出口，改写目标也不会让重组时长变短。

两篇合用时给出一条各自单独说不出的预测：在无解档的单元里，双向翻转率中"不会到会"的那一支应当最高。理由是那里没有人在指定下一步，因而变化只能自发进行，而自发进行的变化恰恰是被过早判定冻

住的那一类。**这条预测两篇单独都推不出来**：那篇没有档位这个变量，本文没有时点这个变量。若实测发现两个读数的相关系数高于零点六，两篇必有一篇是另一篇的代理，届时应保留成本更低的那一个——也就是那篇的双向翻转率，它不需要分母。

与《一条路没有交接点》。这是本栏另一篇同样以“该怎么运行”立题、同样把三家写成三条不同终点的文章，而且它与本文共用物理学与工程学两家。距离最近，必须写在最前面并写细。

先说两家共用学科的取用点：**同一门学科被两篇拿走的是两件不同的东西**。那篇的物理学取的是**盲分析的时点**——判据必须在读出结果之前锁定，历史依据是同一个物理量的公布值曾一段一段朝前一次靠拢；本文的物理学取的是**初值问题与边值问题的适定性差异**——一个在弱条件下解存在唯一，一个可无解可多解，依据是数值方法的标准分类。两处在物理学内部属于不同的分支（前者是实验方法学，后者是常微分方程理论），互不引用，也不能互相推出。工程学同理：那篇取**写进标准的三项独立性**，本文取**稳健设计的最优点脆弱性**。

再说承重命题的分离：那篇处理的是**谁有资格宣布这一段成了**，本文处理的是**要到那儿下一步该做什么有几个答案**。一个是权力问题，一个是适定性问题。**双向判定形式**：一门由完全独立的第三方按事先锁定的判据判定达标的技能课，其倒推唯一性可以极高（那篇读数满分、本文判此处正该用倒推，两篇都判健康，但理由不同且可分开）；而一门达标判定全由授课方自按自判、同时倒推处于无解档的课，那篇诊断为结算权集中，本文诊断为目标未被改写——**两条诊断给出的处方不同且不可互相替代**：交出判定权不会让说不出的下一步变得说得得出，改写目标也不会让自我结算产生信息。**若两个读数的相关系数高于零点六，两篇必有一篇是另一篇的代理**，届时应保留可从既有档案直接算出的那一个，也就是那篇。本文预测两者近乎无关。

与《方程两边都有它》（分岔步·可逆性）。那篇的承重是分岔步的归属。分离线：分岔步的定义要求同一产物有两条路径——**这正是本文的多解区**。因此两篇在这一格上是嵌套关系而非并列，本文诚实写明：**那篇处理的现象只在本文的多解区里存在**，在唯一区没有分岔步可言。本文的增量是把它放进一个三档结构里，并指出唯一区与无解区各需要不同的处理；那篇的增量是在多解区内部给出了“哪一步不能代做”的判据。两篇合用比单用任一篇都强。

与《先动的总是同一边》（节拍比）。分离线：节拍讲结算的频率，本文讲下一步的解的个数。判定：把节拍调到最慢（那篇的处方）而目标仍处于无解档，本文预测无改善——慢下来不会让说不出的东西变得说得得出。

与《总有人替它起步》（结清态·未被归类窗口）。那篇处理产物在成熟之前被归类。分离线：归类是**对已有产物的处理**，倒推唯一性是对**下一步的判定**；一个把每次产物都长时间悬置不归类的班，其倒推档位不因此改变。判定：一门无解档的课延长未被归类窗口，那篇预测续发率上升，本文预测迁移无改善——因为悬置的是没有方向的活动。**这是两篇的判决性对照，可同批测量**。

与《差一条边》（接手图·闭合度）。分离线：那篇的对象是人与人之间的接手关系，本文的对象是下一步的解的个数。**两篇合起来给出一条本文单独说不出的话**：多解区之所以需要“把多条路同时可见”，前提是这个班的接手图能把不同的路传开；图不闭合时，多解区的正确处方也执行不了。

与《正好有人能解决》（清除位·透过率）。分离线：那篇的对象是被消耗的差异，本文的对象是下一步的适定性。判定：一门倒推唯一性极高、教师即时处置能力极强的技能课，那篇预测透过率趋零并判为问题，本文判此处正该用倒推、透过率低不是问题——**两篇在这一格给相反评价**，而这个分歧是可以被检验的：若这类课的学习者在条件变化后表现良好，本文对；若他们在原场满分而换一个条件就垮，那篇对。

十九、与同日另两篇的划界

与《手段取得终点资格的那一段》（教育×化学×艺术）。那一篇处理的是**终点易主这个事件**，读数是易主次数与触发源；本篇处理的是**倒推这个动作在什么条件下有解**，读数是下一步答案的个数。两者数学上独立：倒推唯一的区间里可以从不换位（那正是那一篇的靶格），倒推无解的区间里可以频繁换位而毫无结果（那一篇的漂移格）。**一条共同的推论**：那一篇说终点与手段的角色由信息分布决定，本篇说倒推的可解性由知识分布决定——两条都指向同一个更一般的东西，而三篇都没有把它写实，这是三篇共同交出的洞。

与《这一节课结束时，这个班手里多了什么》（细胞×工程×艺术）。那一篇处理的是学习过程自己造出来的条件，读数是存量。分离线：自建条件是**产物**，倒推唯一性是**约束**。判决性对照：本篇预测自建条件在多解区最容易积累（因为多条路都要留痕才选得出来），在唯一区最少；若实测发现自建条件存量与倒推唯一性正相关，则本篇这条预测为伪，而那一篇不受影响。

二十、五种最强的竞争解释，与排除它们的设计

竞争解释一：其实不就是内容难度吗。 难的内容说不出下一步。排除设计：用独立的题目难度评定测量难度。本文独有的变量与它正交，而且反例在两个方向上都有——一个难度极低的目标可以是无解的（"喜欢上数学"），一个难度极高的技巧动作的下一步专家高度一致。**本文获胜的观察是：在难度匹配的单元之间，倒推唯一性仍有区分力。**

竞争解释二：其实不就是教师经验吗。 经验足的教师说得下一步。排除设计：控制年限与资格。这一条比第一条更难，因为经验确实会提高唯一性判定的准确率。**本文独有的观察在多解区：经验最丰富的五位教师在多解区仍然给出五个不同的答案，而且他们会互相承认对方的答案可行——这与无解区的情形（互相认为对方给的不是下一步）可以分开，而经验解释预测不出这个差别。**

竞争解释三：其实不就是目标写得清不清楚吗。 排除设计：用现成的目标清晰度量表独立评分。本文的正交主张在这里可以直接检验——本文预测存在大量目标清晰度满分而倒推无解的单元（"能够评价一个论证"），也存在目标清晰度中等而倒推唯一的单元（学徒制、乐器教学）。**若清晰度与唯一性的相关系数高于零点七，本文的第二根轴不成立。**

竞争解释四：其实不就是这个学科成熟不成熟吗。 成熟学科有细致的子技能分解，因此下一步说得出。排除设计：在同一学科内部比较不同单元。**本文获胜的观察是：同一门课的不同区间落在不同档位——写**

作课的句子层唯一、篇章层多解、"自己的声音"无解，三者在同一学科同一门课同一位教师。若档位只随学科变化而不随单元变化，本文的分段主张失效，退化为一条关于学科成熟度的分类。

竞争解释五：其实不就是评判者不认真吗。 说不出下一步是因为没细想。排除设计：给评判者充足时间并要求写出理由，再看档位是否变化。本文的预测是无解区的答案在给足时间后仍然是活动名称，且这些答案能通过"换一个目标是否照样说得通"这一关的比例极低。若给足时间后无解区大量转为唯一区，则本文测的是评判者的投入而不是知识分布。

判决可以压缩成一句：若在同一课程、同一学科、控制难度、教师经验与目标清晰度之后，倒推唯一性与运行方式之间仍存在交互效应（唯一区倒推优于推进、多解区方向相反），且同一份材料在不同专家群体间的档位不同，则本机制优于上述五条；若效应被难度、经验或清晰度完全吸收，或档位跨群体一致，则竞争解释获胜。

二十一、分层引用结构

本文的主张分三层，各自的证伪条件不同，可以分层引用。

第一层：可倒推区间作为一个存在物，且它是知识分布的性质而非系统性质。 最强也最容易倒。它的关键证伪条件在证伪条件其二：若同一份材料交给两个不同的专家群体得到相同的档位，这一层就被可控性覆盖，本文只是一个通俗说法。

第二层：按档分段的运行方式作为一个分析工具。 即使第一层的"知识分布"主张被推翻，三档与三种运行方式的匹配仍然可用——那时它退化为一条关于内容性质的分类，仍能指导教学安排。这一层的证伪条件是交互效应：若在唯一区用倒推与在多解区用倒推的效果差异不显著，这一层也倒。

第三层：一条不依赖前两层的经验主张。 目标清晰度极高、而五位教师说不出下一步的那些课程单元，其学生的延迟迁移不优于目标清晰度中等、而下一步可说的单元。这一条不需要"可倒推区间"这个概念，只需要一次教研活动的问卷与一次延迟测验，而且它是三层里最便宜的一条。

最容易检验的是第三层，最容易被推翻的是第一层，而本文最看重的是第一层。 若第三层被证实而第一层被推翻，说明现象为真而本文的刻画错了；若第三层被推翻，三层一起倒。

二十二、证伪与赌注

最强的竞争解释是：其实不就是内容难度和教师经验吗——难的内容说不出下一步，经验足的教师说得出。

本文必须能排除它。难度可以用独立的题目难度评定测量，教师经验可以用年限与专业资格测量。本文独有的变量是**答案的个数**，它与两者都正交：一个难度很低的目标可以是无解的（"喜欢上数学"），一个难度极高的目标可以是唯一的（一个高难度技巧动作的下一步，专家高度一致）；而经验最丰富的五位教师

在多解区仍然会给出五个不同的答案——多解不是无知，是路径确实多，这一点可以通过让专家互相评价对方的答案来区分：多解区里五个人会认为彼此的答案都可行，无解区里五个人会认为彼此的答案都不是下一步。

核心证伪条件：在同一门课程、同一批教师、控制内容难度与教师经验之后，若倒推唯一性与"按档匹配的运行方式"之间不存在交互效应——即在唯一区用倒推与在多解区用倒推的效果差异不显著——则本文的分段主张为伪；届时应把课堂运行的差别归于难度与经验，而非倒推的适定性。

其余七条：

其二，若同一份学生样本与同一目标交给两个不同专家群体得到相同的唯一性档位（跨群体一致），则倒推唯一性是系统性质而非知识分布性质，本文被可控性覆盖，那条分离线失败。

其三，若在多解区中专家互评"彼此的答案都可行"的比例与无解区无差异，则多解与无解不可区分，三档应压缩为两档。

其四，正向读数（顺序）：区间边界的移动应当早于表现变化出现——一个学生的某个区间从多解收敛为唯一，应当先于他在该内容上的成绩提升。若表现变化总是先于边界移动，则边界移动只是结果不是先兆，完整次序第三步失效。

其五，若无解区的目标经改写后唯一性没有变化，则"改写目标"这条处方无效，完整次序第二步的无解档没有可执行内容。

其六，最便宜的一条：**任何一所学校的教研组，用一次教研活动就能测一门课六个单元的倒推唯一性——五份匿名学生样本、五位教师、各十分钟。**所需的一切在多数机构的既有安排里已经存在，不需要新增数据采集。

其七，若不同专家群体规模（三人、五人、九人）给出的档位不稳定，则该读数不可靠，应先解决测量学问题再谈机制。

其八，若在形式部分列出的三个已按唯一性分段的行业（飞行训练、产品原型、康复计划）里，这种分段与结果无关，则本文的兑现是同形而非真兑现。

写死日期的赌注：在二〇三〇年十二月三十一日前，至少两项在同一制度环境内、单元数不少于三十六、预注册的研究，报告倒推唯一性与运行方式之间存在交互效应：在唯一区，倒推式运行优于推进式；在多解区，方向相反。

以下不算命中：只报教师对课程难度的评分；只比较讲授与探究；把"目标清晰度"当作倒推唯一性；用两所学校或两个国家的对比充当交互检验；只报学生满意度或当堂正确率；机构自行说明已按学情分层。

二十三、反噬与治理

倒推唯一性一旦进入考核，最省事的实现方式是把所有目标都改写成唯一区的形式——把"能够评价一个论证"改写成"能够指出论证中的三种常见谬误"。这个改写能把读数做满，而且它在形式上完全符合本文完整次序第三步的建议。问题在于：**改写目标与掉换目标之间没有可靠的分界**。前者是把同一个东西写成可操作的形式，后者是换一个更容易的东西来做，而两者在文档上看起来一模一样。

第二重反噬：无解区一旦被命名，它会成为一个免责的类目。教师可以把任何教不动的东西归入无解区并宣布"这在当前知识条件下不可教"，而本文恰恰给了这句话一个体面的说法。

三条治理底线：倒推唯一性指向内容与知识分布的关系，不指向教师或学生的能力，不得单独用于排名或教师评价；不得为提高读数而在安全关键的训练中把整体判断类目标一律改写成程序性目标；任何据此作出的不利安排——包括判定某目标不可教——必须公开评判者名单、原始答案与复核通道。

我看不到一个彻底的出口。只要这个读数被采用，改写目标就永远比改进教学省事，而改写与掉换的分界只能靠人来判。可用的唯一限制是要求改写前后的目标在一个独立的延迟测验上可比——而这条限制本身也要花钱。

二十四、若本判断被证伪，我们会学到什么

若倒推唯一性与运行方式之间不存在交互效应，我们学到的不是"方法都一样"，而是一件更具体的事：**倒推与推进的差别在教学上不承重，承重的是投入与反馈的质量**。那意味着刻意练习那一支关于"结构化程度"的主张可以推广到它自己划的边界之外，而本文所说的多解区与无解区，只是结构化尚未做到的地方。这会把研究的下一步引向"如何把高层级目标结构化"，而不是"如何为它换一种运行方式"。

若档位跨专家群体一致，我们学到的是：可倒推性是系统的性质而不是知识分布的性质，本文只是可控性的一个通俗说法。那不是一无所获——它会把教学问题正式接到控制理论上，而可达集、可控性格拉姆矩阵这些工具随之可用，这比本文能给的多。

若无解区的目标经改写后唯一性不变，我们学到的是：无解不是目标表述的问题，是这个目标在当前的知识条件下真的没有可教的路径。那意味着一批被广泛写进培养方案的目标应当被明确标注为"不可直接教授"，而这是一个有实际后果的结论——它会改变课程时间的分配方式。

三种失败方式各自指向不同的下一步。

本文交出三个自己解释不了的洞。第一，多解与无解的分界依赖专家互评，而互评本身需要专家在同一个语汇里作答；当两个群体连"下一步"的意思都不一样时，这个分界无法测。第二，本文说无解区必须改写目标，却没有给出改写的方法——那需要另一条完整的次序。第三，本文完全没有处理**档位随学生个体变化**的情形：同一个目标对一个学生是唯一区、对另一个学生是多解区，这在一个班里同时发生，而本文的三档运行方式是按内容分段的，无法按人分段。

二十五、自反检验

按本文自己的判据，这篇文章处在多解区。它交出的是一个读数与一条按档分段的次序，而"读了它之后下一步该做什么"这个问题，五位读者会给出五个不同的答案：先去测一门课、先去反驳强制不一致一节那条与可控性的分离、先去改写一个无解目标、先去做证伪条件其六那个便宜检验、先什么都不做等有人做完。

这五个答案都可行，这正是多解区的签名。按本文自己完整次序第二步的处方，本文此刻不应当指定唯一的下一步，而应当把这几条路同时摆出来——上面那句话就是在做这件事。

本文交出两个自己解释不了的洞。第一，多解区与无解区的分界依赖专家互评，而互评本身需要专家在同一个语汇里作答；当两个专家群体连"下一步"的意思都不一样时，这个分界无法测。第二，本文说无解区必须改写目标，却没有给出改写的方法——那需要另一条完整的次序，本文没有它。

二十六、结论

课堂如何运行才能让学习真正发生，三门学科给了三条不能并存的次序：定住两端整体求解、一步一步局部改进、放弃最优只求平坦。三家共同假定"下一步该做什么"这个问题有答案，分歧只在怎么求。

物理学自己拆掉了这个假定。初值问题在很弱的条件下解存在唯一，两点边值问题可以无解、可以有无穷多解——**倒推与推进不是同一个动作的两个方向，它们的适定性根本不同**。这条差别在物理里回答的是数值方法怎么选，从没有人把它当作"该不该先定终点"的证据。

由此得到一条三家单独都推不出来的判断：**课堂运行的次序必须按倒推的解的个数分段，而不能按内容、按课时或按目标层级分段。** 唯一区用倒推，多解区用推进加选择，无解区必须先改写目标——三种方式互不通用，用错了不会报错，只会在形式上一切合规而什么都没发生。

这句话的可操作形式是一句问话，一次教研活动就能测：把这个学生此刻的状态和这门课的目标交给五位有经验的教师，他们说出的下一步是同一件事、五件不同的事、还是都说不出来？

如果答案是第三种，那么这门课的目标可能写得非常漂亮，与评价对得非常齐，而它的每一节课都在一个没有解的问题上运行。

注释

〔一〕本文所说的"下一步"限于**一件具体的、此刻可以开始做的事**。"多读多写""加强练习""培养兴趣"不构成下一步——判据是：把目标换成同一学科的另一目标，这句话是否照样说得通；照样说得通的，是活动名称不是下一步。

〔二〕判据一节的判据要求五位**有经验的**教师，且形式部分其三指出应尽量来自不同培养背景。人数、经验年限与背景多样性对档位判定的影响是一个测量学问题，本文没有解决它，只在证伪条件其七里把它列为可能推翻本读数的一条。

〔三〕"多解"与"无解"的分界依赖专家互评：多解区里五位教师会认为彼此的答案都可行，无解区里五位会认为彼此的答案都不是下一步。这个分界在两个专家群体使用不同语汇时无法测量，本文在自反一节把它列为自己解释不了的洞之一。

〔四〕本文的形式部分中的 L_1 、 L_2 、 L_3 没有给出测量方法。三条推论只依赖 $L_2 \gg L_1$ 这一个不等式，不依赖赋值；本文不主张它们可被量化。

〔五〕本文的三处反向约束各自削弱了本文的主张，且都由被引学科自己提供。其中物理学那一处（多解不是缺陷而是可达路径多）改变了本文的价值排序，是三处里最重的一处。

参考文献

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman.
- Ascher, U. M., Mattheij, R. M. M., & Russell, R. D. (1995). Numerical Solution of Boundary Value Problems for Ordinary Differential Equations. SIAM.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406.
- Feynman, R. P., Leighton, R. B., & Sands, M. (1964). The Feynman Lectures on Physics, Vol. II, Ch. 19: The Principle of Least Action. Addison-Wesley.
- Gray, C. G., & Taylor, E. F. (2007). When action is not least. *American Journal of Physics*, 75(5), 434–458.
- Kalman, R. E. (1963). Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 1(2), 152–192.
- Kierzenka, J., & Shampine, L. F. (2001). A BVP solver based on residual control and the MATLAB PSE. *ACM Transactions on Mathematical Software*, 27(3), 299–316.
- Macready, G. B., & Dayton, C. M. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Educational Statistics*, 2(2), 99–120.
- Nair, V. N. (Ed.). (1992). Taguchi's parameter design: A panel discussion. *Technometrics*, 34(2), 127–161.
- Taguchi, G. (1986). Introduction to Quality Engineering: Designing Quality into Products and Processes. Asian Productivity Organization.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Wiggins, G., & McTighe, J. (2005). *Understanding by Design* (Expanded 2nd ed.). ASCD.

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