

复推率

论学会了没有，不看现在能不能说出那条规则，而看换一个表面不同的情境能不能重新推出同一条

教育心理学 × 运动学习 × 组织例程研究

作者 王德生 + Claude

栏目 学科通融 · 站外碰撞

篇幅

发表 2026年8月1日

摘 要

当一个成形的答案随时可得，学会了就无法再从作品上看出来。一个流行的处方是记录判断规则的前后变化：帮助前写下你依据什么，帮助后指出哪一条被改了。三个领域对这个处方给出互相排斥的答案。教育心理学说这正是学习的证据；运动学习说不对——真正承担适应的是无法语言化的内隐系统，而要求把规则说出来反而会挤掉它；组织例程研究说更不对——规则从来不被存储，它在每一次执行时被参与者重新推断出来，两个时点之间根本没有「同一条规则」可供比较。三家共有一个未说出的前提：存在一条被存在某处、可以在两个时点之间被取出来比对的规则。三家的材料合起来只能得出：规则不被存储，它每次被重新生成；学会了，就是重新生成的可靠性上升。本文把这个可靠性称为复推率，判据不含情态词：换一个表面不同、结构相同的情境，他会不会推出同一条？而且必须至少测两次——一次推对可能只是记住了。由此得到一条与直觉相反的推论：能把规则说得越清楚的人，复推率未必越高；说得清楚是复述的证据，不是复推的证据。

这一篇由哪三个领域的争论撞成

三边对「学习的最小单位是什么」给出互相排斥的答案：教育心理学说是规则的前后差，运动学习说真正承担适应的是无法语言化的内隐系统而语言化会挤掉它，组织例程研究说规则从不被存储、每次执行时被重新推断。三家共有一个未说出的前提：存在一条被存在某处、可以在两个时点之间取出比对的规则。三家的材料合起来只能得出：规则每次被重新生成，学会就是重新生成的可靠性上升。

运动学习 · 说出来会挤掉学到的

[Taylor, Krakauer & Ivry, Explicit and implicit contributions to learning in a sensorimotor adaptation task \(J. Neuroscience, 2014\)](#)

在视觉旋转适应任务里可以把两个成分分开：使用者能说出的显式策略，与他说不出、也无法主动关闭的内隐适应。两者的时间进程与稳定性不同，且会互相竞争——当显式策略迅速把误差消掉时，驱动内隐适应的信号也被消掉了。语言化不是中性的观测手段。

<https://www.jneurosci.org/content/34/8/3023>

组织例程研究 · 没有被存起来的规则

[Feldman & Pentland, Reconceptualizing organizational routines as a source of flexibility and change \(ASQ, 2003\)](#)

例程有两面：参与者关于「这件事该怎么做」的抽象理解，与每一次具体执行。而抽象面不是模板——它由执行被反复推断出来，不同的人推断出的版本不同，同一个人在不同时候也不同。两个时点之间根本没有「同一条规则」可供比较，因为每次拿出来的都是刚被造出来的。

<https://journals.sagepub.com/doi/10.2307/3556620>

教育心理学 · 前后差就是学习的证据

[Kornell, Hays & Bjork, Unsuccessful retrieval attempts enhance subsequent learning \(JEP:LMC, 2009\) ; Richland, Kornell & Kao, The pretesting effect \(2009\) ; Kapur, Productive failure \(2008\)](#)

自行生成优于被动接受；在直接教学前先行探索、即使失败也可能改善迁移。而先测效应给出的一条对本文最要紧：先行判断的价值不依赖那份判断的内容质量——哪怕几乎必然答错，先试一次再看答案仍优于同样时间的直接学习。这一点是「前后差」框架自己解释不了的。

<https://web.williams.edu/Psychology/Faculty/Kornell/Publications/Kornell.Hays.Bjork.2009.pdf>

目 录

一、导论：三个不该同时为真的说法.....	5
二、三家各自说到了哪一步	5
三、冲突落在哪一点上	6
四、复推率.....	7
五、两根轴，四种局面	7
六、一个不含情态词的判据	8
七、逐领域兑现	9
八、两处被材料逼出来的收窄	10
九、与最近的几种说法的分界	10
十、三家各自为什么到不了这一条	11
十一、推论、适用边界与反噬	11
十二、与同期几篇的分界	12
十三、本文自己的复推率	13

十四、可以判本文错的六种结果13

结 论.....14

注 释.....14

一、导论：三个不该同时为真的说法

当一个完整、流畅、通常正确的答案随时可得，一份作品就不再能证明作者会什么。于是一个处方反复出现：不要看作品，看规则的前后变化——帮助到达之前先写下你依据什么判断，帮助之后指出哪一条依据被改了、为什么改。三个领域对这个处方给出三个互相排斥的答案。

第一场在教育心理学。 这一支的判断是：这正是学习的证据。自己先生成一个判断，比直接读一个正确判断更有利于后续保持；在结构不良的问题上先行尝试并失败，可以为后续教学准备出更好的结构；而把前后两个版本配成一对，可以把「作品改善」与「人的判断改善」分开。争点是：**规则的前后差，是不是学习的最小单位？**

第二场在运动学习。 这一支给出的答案是：真正承担适应的那一部分，根本不在规则里。在感觉运动适应任务中可以把两个系统分开：一个是使用者能说出来的显式策略，一个是他说不出来、也无法主动关闭的内隐适应；两者相加构成总的行为改变。更要紧的是，两者会互相竞争——当显式策略把误差消掉时，驱动内隐适应的那个信号也被消掉了，于是**把规则说出来并加以运用，会挤掉那一部分真正稳固的学习。**

^[1] 争点是：**要求人语言化他的规则，是在测量学习，还是在改变学习？**

第三场在组织例程研究。 这一支说前两家的前提都不成立。一项例程不是一个被存在某处的程序，它有两面：参与者关于「这件事该怎么做」的抽象理解，以及每一次实际执行。而抽象的那一面并不是被存起来后取出执行的模板——它在每一次执行中被参与者重新推断、重新说出、也因此不断改变；不同的参与者从同一批执行里推断出的抽象面并不相同。^[2] 推到底：**两个时点之间根本没有「同一条规则」可供比较，因为每一次拿出来都是刚刚被造出来的。** 争点是：**规则的前后差，比较的是两个什么东西？**

三条不能同真。若前后差是学习的单位，运动学习那边的分离实验就该显示语言化只有好处而无代价，而它显示的是竞争。若语言化只有代价，那么大量以规则表述为核心的专业训练就该是无效的，而它们不是。若根本没有同一条规则可比，那么教育心理学几十年里靠前后对照得出的结论就都失去了对象，而那些结论可复现。

本文认为，僵局来自一个三方共有、从未被单独说出的前提：**规则是一个被存放在某处、可以被取出、被比较、被改写的东西。** 教育心理学要比较它的两个版本，运动学习要问它由哪个系统持有，例程研究说它不在任何地方——三家争的是它放在哪儿、由谁持有，没有人问「学会」是不是就等于「持有一条规则」。

二、三家各自说到了哪一步

教育心理学这一支的两块基石很老也很硬：自行生成的材料比被动接受的记得更牢；在直接教学之前先行探索、即使失败，也可能改善后续的迁移。近年被推到人机协作里，形成的处方是把帮助前后的判断配成一对，用这一对的差来判断学习是否发生。

这一支还有一位更近、也更容易被忽略的邻居：先测效应。在还不知道答案的时候先试着回答——哪怕几乎必然答错——之后再看正确答案，其保持效果优于同样时间的直接学习。^[3] 这条结论对本文很要紧，因为它说明先行判断的价值不依赖那份判断的内容质量；而这一点恰恰是「前后差」框架无法解释的：如果起作用的是差，那么一个纯粹瞎猜的初答应当不产生任何有用的差。

运动学习这一支提供了最硬的反向证据。在标准的视觉旋转适应任务里，用一个简单的办法可以把两个成分分开：要求参与者报告他打算瞄准哪里，实际落点与瞄准点之间的差就是内隐部分，瞄准点与目标之间的差就是显式部分。结果是两者的时间进程与稳定性都不同，且总误差被显式策略迅速消掉时，内隐部分的增长会停下来——因为驱动它的那个误差信号已经不存在了。^[1]

推论很直接：在这类任务上，要求人说出规则并按规则行动，会系统性地减少那一部分不需要说出来、也不会随注意力涣散而消失的学习。这一支的边界同样清楚：它的实验对象是感觉运动适应，把结论直接搬到概念推理上没有依据。但它至少确立了一件事——语言化不是中性的观测手段。

组织例程研究这一支把规则的存在方式整个改写了。一项例程被分成两面：一面是参与者对「这件事一般怎么做」的抽象理解，另一面是具体的每一次执行。关键在于抽象面不是模板：它由执行被反复推断出来，不同的人推断出的版本不同，同一个人在不同时候推断出的版本也不同；例程的变化正是在这个反复推断里发生的，而不是靠有人去修改一份文件。^[2]

这一支到达的边缘是：它得出了「没有可保存的规则」，但它的问题是例程如何变化与如何稳定，它不问一个人学会了没有。

三、冲突落在哪一点上

把三家的处方并排放，互不兼容：记录前后差、不要语言化、没有可比的东西。把三家的材料并排放，出现另一样东西。

例程研究说：抽象规则每次被重新推断。运动学习说：有一部分学习不经过语言。先测效应说：先行判断的价值不依赖它的内容。

三条串起来只能得出一件事：**学会了，不等于持有一条规则，而等于在需要的时候能把那条规则重新推出来。** 规则不是一个存量，它是一个每次被重新生成的产物；学习不是这个存量的改变，是重新生成这件事的可靠性上升。

这一步同时解释了三家各自的现象。它解释了为什么先测有效而与初答质量无关——先测练的不是记住某条规则，是走一遍从情境到规则的那条推导路径，走过的路径下次更好走。它解释了为什么语言化会有代价——把规则说出来并直接套用，等于跳过了推导，下次仍然要从头推；而如果推导被一个现成的表述代替，那条路径就得不到练习。它也解释了为什么例程研究找不到被存起来的规则——因为本来就没有，每次拿出来的都是刚推出来的。

于是问题的形状变了。它不再是「怎样记录规则的变化」，而是：**下一次，他还能不能把它推出来。**

四、复推率

在一个表面不同而结构相同的新情境里，一个人能够重新推出同一条判断依据的可靠程度，本文称之为复推率。

四件事需要说清。

「重新推出」而不是「说出」。说出只需要检索一段话，重新推出需要从这个具体情境走到那条依据。两者可以完全分离：一个人可以流利地背出「相关不等于因果」，而在一个新的案例前照样把相关读成因果。判别办法是把情境的表面特征全部换掉——换领域、换数字、换叙述方式，只保留结构。

「同一条」由功能定义，不由措辞定义。两次推出的表述可以完全不同，只要它们在这两个情境里作出同样的区分，就算同一条。这一点使复推率可以在完全不语言化的任务上测量——只看行为上的区分，不看陈述。

「可靠程度」是一个比率，必须至少测两次。一次推对可能是记住了那个情境；两次以上、且情境之间表面差异足够大，才开始区分复推与复述。这是本文与几乎所有现行做法的一个操作差别：现行做法多数只测一次迁移。

「新情境」的结构相同性必须事先由第三方判定，不能由被试事后确认，也不能由出题者凭直觉认定——否则整个测量会滑向「他答对了所以这两题结构相同」。

由此得到本文的承重命题：学习不是规则从一个版本变成另一个版本，是从情境到规则那条推导路径的可靠性上升。规则本身在每一次判断时被重新造出来，因而没有一个可以被存起来、被比较、被改写的东西。

这条命题有一个立刻可检验的、与直觉相反的推论：能把规则说得越清楚的人，复推率未必越高。说得清楚是检索的证据；而一个流畅的现成表述恰恰可以代替推导，使那条路径得不到练习。这解释了一个在教学里长期被当作个体差异的现象——有些学生的复述完美而迁移糟糕，通常被归因为「死记硬背」，而按本文这不是态度问题，是路径没有被走过。

这条判断为什么三家都不会同意。教育心理学不会同意，因为它的整个测量框架建立在规则可比对之上，而本文说没有可比对的东西。运动学习不会同意，因为它把显式与内隐当作两个并列的系统，而本文说显式的那一份也不是存量，它同样是每次被推出来的。例程研究不会同意，因为它的结论是关于集体的抽象面，而本文把它用到了个人身上，且给了一个它没有的量化读数。

五、两根轴，四种局面

只用复推率一根轴不够。复推率相近的两种局面结局不同：一种能说清自己推的是什么，一种说不清。要分开，需要第二根轴。

第一根轴：**复推率高不高**——换一个表面不同、结构相同的情境，能不能重新作出同样的区分。第二根轴：**这次复推能不能被说出来**——不是能不能复述一段现成的话，是能不能说明这一次为什么这样判断。

	能说明	不能说明
复推率	可教的行家 既能在新情境作出区分，也能说明凭什么。稀缺，且这是唯一能把能力传下去的一格。	沉默的行家 判断稳定可靠而说不出理由。老练的临床、维修、品鉴多在这里。 能力是真的，可传递性接近零。
	会背的人 能把规则说得很清楚，换个情境就用不上。所有以陈述为考核方式的体系都在批量生产这一格。	还没学会 既推不出也说不出。

左下格是靶格，而它之所以是靶格，不是因为它最严重（右下格更严重），而是因为**它在所有以陈述为形式的评估里都得高分**。笔试、面试、口头复盘、书面反思，全部测的是能不能说出来；而说得清楚与推得出来在这一格里是分离的。一个体系若只用陈述评估，会系统性地把资源导向这一格，并且看不见自己在这么做。

右上格值得单说，因为它常被误判。沉默的行家不是「还没形成理论」，也不是「不愿意分享」；按本文，他的规则从来就没有以可陈述的形态存在过，每一次都是重新推出来的，而推导本身不经过语言。**因而要求他「把经验总结成条文」，得到的通常是一份低复推率的陈述——那是他事后为这次判断编的理由，不是他推的过程。**这解释了为什么专家访谈提炼出的规则，交给新手用往往不灵。

六、一个不含情态词的判据

判断一个人学会了没有，不必问他懂不懂、理解得深不深、有没有掌握本质。只需做一件事：

换一个表面不同、结构相同的新情境，看他会不会重新推出同一条——并且至少做两次。

三条施测要求把它变成可执行的：结构相同性由不知道被试表现的第三方事先判定；表面差异必须覆盖领域、数值与叙述方式三项；两次之间不给任何反馈，否则第二次测的是学习而不是复推。

跑几个场合，答案互不相同。

数学的变式题。换数字算不算表面差异？按上述要求，不算——只换数字，结构与叙述都没变，测的是执行不是复推。真正的表面差异要换叙述框架（把一个几何问题写成一个分配问题）。这解释了一件长期被观察到的事：换数字的变式题成绩与原题高度相关，换框架的变式题成绩与原题相关很弱，而后者才预测长期表现。

临床的鉴别诊断。换一组症状表现不同而机制相同的病例。这个领域已经有现成的操作——用不同表现的同一疾病考察诊断者，而本文加的一条是：**必须至少两例，且两例之间不给反馈**。多数考核只用一例。

代码的调试。 换一个语言、换一个框架，注入结构相同的缺陷。这一处的复推率可以完全不经语言测量——只看他先去看哪里。

写作的论证判断。 换一个题材、换一个立场，看他会不会指出同一类的证据缺口。这里要小心一个陷阱：若两次的立场相同，测到的可能是立场而不是复推；因此第二次应当换到他不同意的一侧。

运动技能。 换一个器械、换一个场地。这里复推率与语言化完全无关，而它恰恰是本文第二根轴最纯粹的右侧样本——高复推、零陈述。

七、逐领域兑现

一、考试设计。 处方是把「一次迁移题」改成「两次以上、表面差异覆盖三项、之间不给反馈」。成本几乎为零（同样的题量，改分布），而它测的东西不同。可裁决预测：改造后的分数与原设计的分数相关将显著低于零点七，且改造后的分数对一年后的表现预测更好。

二、专家知识的提炼。 处方是不要求专家陈述规则，而是让他在一组事先构造的情境上作判断，由第三方从判断的分布里反推区分面。**这与传统的专家访谈方向相反：不问他怎么想，看他怎么分。**

三、师徒制。 它之所以在很多行当里无法被课程替代，按本文有一个具体的原因：课程传的是陈述，而右上格的行家没有陈述可传；师徒制传的是大量情境，等于直接练那条推导路径。

四、临床的病例讨论。 若讨论的重点是「这个病人该怎么处理」，练的是执行；若重点是「换成另一个表现，你还会这样想吗」，练的是复推。这一条不花任何额外时间。

五、面试。 几乎全部落在左下格：问的是能不能说出来。可裁决预测：在同一批候选人上，把陈述式问题换成两次结构相同、表面不同的判断任务，两种方式给出的排序相关很低，而后者对试用期表现的预测更好。

六、机器学习的评测。 分布外泛化的测法与本文完全同构：换表面不换结构。**这一处的价值在于它说明复推率不是一个关于人的概念——它是关于任何学习系统的。** 而这个领域已经积累了大量关于「表面差异到底该怎么构造」的方法，可以直接借用。

七、法律的案例教学。 苏格拉底式追问的核心动作正是不断改变事实的表面特征，看学生的判断在哪一处翻转。这是复推率的一个古老实现。

八、飞行与应急训练。 模拟器训练的价值不在重复标准处置，而在情境的表面差异；这解释了为什么按固定脚本重复演练的效果远低于随机变式演练。

九、语言学习。 能说出语法规则与能在新句子里作出正确区分，是本文两根轴的教科书例子；而外语教学的评估几乎全在左下格。

十、组织的复盘（本领域反过来加的约束）。 例程研究指出，抽象面是由集体的执行反复推断出来的，不同的人推出的版本不同。推到复推率上：**一个团队的复推可能发生在团队层面而不在任何个人身上——**

没有一个人能独自推出那条依据，而这群人在一起就能。这一条把本文的测量单位问题整个打开了，见下节。

八、两处被材料逼出来的收窄

第一处，由运动学习逼出：测量会改变被测的东西。 显式与内隐的竞争意味着，要求一个人报告他的推导，会改变他后续的学习路径——报告本身把误差信号消掉了一部分。于是复推率的测量不是无成本的观测：**反复测复推率，会把被试推向左侧（更依赖显式策略），而这正是本文用来诊断的那一侧。**

这条约束不能靠更好的施测消除。能做的只有两件：其一，第二根轴的测量（能不能说明）必须与第一根轴的测量分开进行，且分开的次数要少；其二，在纯行为可测的任务上，第一根轴应当完全不经语言测量。由此本文丧失了一个方便的主张——**复推率是一个可以随时随地反复读取的指标；它必须承认自己是一次介入。**

第二处，由例程研究逼出：单位可能不是个人。 如果抽象面由集体执行反复推断，那么「他能不能推出来」这个问法在很多真实场合是错位的——真实的判断由一组人共同推出，而任何单个成员单独测都会得低分。

这条约束比第一条更重，因为它动摇的是本文的测量对象。**本文因此不能主张复推率是一个个人属性；它只能主张它是一个「判断单元」的属性，而判断单元是个人还是团队，必须按具体场合确定，且这个确定本身没有一般规则。** 一个直接后果是：所有以个人为单位的选拔与考核，在团队复推的场合会系统性地测错东西——它测的是这个人单独时的复推率，而工作时的判断单元不是他一个人。本文只能把这个问题交出去，没有办法处理。

九、与最近的几种说法的分界

一、与规则前后差的框架。 那一支要求保存帮助前后的两份规则并比较。分离线：它假定两份规则指向同一个存量的两个版本，本文说两份都是各自当场被推出来的产物，前后差因而不是同一个东西的改变，是两次推导的结果不同。判决点：把「帮助后的那一份」改成事后补写（因而不是当场推出的），前后差框架预测迁移效果显著下降，本文预测下降很小——因为起作用的是帮助前那次真实的推导，不是两份的差。**这是一条便宜、直接、且两家预测相反的对照。**

二、与先测效应。 这是本文最近的一位，因为它已经说明先行尝试的价值不依赖初答的正确性。分离线：先测效应解释的是为什么先试一下有助于后续学习，机制多被归于搜索记忆的过程或对答案的注意提升；本文提出的是一个测量量，且预测这个量本身可以被独立观测并预测迁移。判决点：两组接受同等次数的先测，其最后一组被要求在两个表面不同的情境里各推一次、另一组只在一个情境里推两次。先测效应对二者不作区分，本文预测前者的迁移显著更好。

三、与有益困难。 那一支预测某些让训练变慢变难的安排改善长期保持。分离线：有益困难的自变量是难度，本文的自变量是表面差异的覆盖面。判决点：在难度、用时、错误率完全匹配的两组之间，只改变情境表面差异的覆盖面，本文预测迁移不同，有益困难预测相同。

四、与显式—内隐分离。 那一支把学习分成两个可分别测量的成分。分离线：它处理的是同一次适应中的两个成分，本文处理的是同一条判断依据在两个不同情境里的两次生成。判决点：一个纯粹内隐、完全无法报告的学习成果，在本文这里可以有很高的复推率（换器械照样做到），而在显式—内隐框架里它只落在内隐一侧、不涉及跨情境的问题。

五、与迁移研究本身。 近迁移与远迁移的区分与本文高度相邻。分离线：迁移研究测的是新任务上的成绩，本文测的是同一条依据是否被重新推出——一个人可以在新任务上成绩很好而依据的是另一条规则（比如靠一个只在这两题上碰巧成立的表面线索）。判决点：在成绩相同的两组之间，看他们在判断翻转点上的位置是否一致；不一致者按本文是不同的规则，而迁移成绩无法区分。

六、与分布外泛化。 机器学习那一侧的这个概念与本文几乎同构，且方法更成熟。本文承认这一点并明确借用它的方法论：**表面差异的构造必须由不接触被试表现的第三方事先完成**，这正是那个领域为防止基准泄漏而发展出的纪律。分离线只在对象上：那一侧不处理第二根轴（能不能说明），因为它的系统本来就不需要说明。

十、三家各自为什么到不了这一条

教育心理学到不了，因为它的记录工具就是语言。 它要比较两个版本，就必须让两个版本以可比较的形式存在，而可比较的形式就是陈述。于是它天然测到的是复述那一维。

运动学习到不了，因为它的任务里没有情境的表面差异。 视觉旋转适应是一个单一情境的任务，它可以把显式与内隐分开，但它没有第二个情境可以问「换一个还推不推得出」。

例程研究到不了，因为它的问题是稳定与变化，不是学会没学会。 它得出了「每次被重新推断」这个结论，但它关心的是例程为什么会漂移，而不是某个参与者的推断可靠性。

十一、推论、适用边界与反噬

第一条推论：陈述式评估会系统性地把资源导向左下格，且看不见自己在这么做。 因为它的读数在那一格最高。

第二条推论：想提高复推率，就增加情境的表面差异，而不是增加练习量或提高难度。 这是本文唯一一条可执行的处方，成本近乎为零——同样的题量，改分布。

适用边界一：测量单位不确定。 见第八节第二处，这是本文最大的未解处。

适用边界二：只适用于存在结构—表面之分的任务。 有些任务的表面就是结构（比如识别一张具体的脸），那里没有可换的表面。

适用边界三：本文不主张陈述无用。 第二根轴的左侧是唯一能把能力传下去的一格；本文只主张陈述不能作为学会的证据，不主张它没有价值。

反噬预言。 若复推率成为考核指标，最可能发生的是**对表面差异的应试化**：培训会开始教「常见的表面变换有哪几类」，于是学生学会识别变换本身，而这又是一条可以被记住的现成规则。**本文的判据会被它所诊断的那个机制吃掉——学生将复述「这是换了框架的同一类问题」，而不是重新推。** 我找不到出口：任何一个足够稳定、可被公布的表面差异构造方式，都会在几轮之后变成可复述的对象。唯一能拖延的是让构造方式保持不可预测，而这与考试的公平性直接冲突。

交出去的一个洞。 「结构相同」这个判定本身没有一般标准。本文把它交给不接触被试表现的第三方，这只是防止了事后合理化，没有解决什么叫结构相同。而在很多领域里，两个情境算不算结构相同，恰恰是这个领域最难的问题本身。

十二、与同期几篇的分界

与「独版」那一篇。 那篇说帮助前那份记录的价值在于不可重制，本文说它的价值在于走了一遍推导。判决点：一份在帮助前作出、但由他人代写的判断（不可重制，却没有走过推导），那篇判为有价值，本文判为零。反过来，一次在心里默默完成、未留任何记录的推导，那篇判为零，本文判为有效。**两篇合起来才完整：留痕解决第三方的核对，推导解决本人的学习。**

与「不按的代价先落在谁身上」那一篇。 那篇处理动力，本文处理能力。判决点：代价完全内部化而从未在多个表面情境里推过的人，那篇预测他会去干预，本文预测他干预时用的仍是记住的那一条，换个情境就失效。

与「记零位」那一篇。 那篇问一类动作有没有账，本文问一条依据能不能被重推。判决点：一个把「发现错误」记账很足、而考核方式全是陈述式的共同体，那篇判为健康，本文预测它招进来的全是左下格的人。

与「入判份额」那一篇。 那篇问产出进不进别人的判据，本文问依据能不能被重新推出。判决点：一个入判份额极高、而其判断从未在表面不同的情境里被检验过的人，那篇判为立法者，本文预测他所改写的标准在情境变化时会失效——两篇合起来解释一类常见的失败：一个有影响力的判断被广泛采用，然后在一个表面不同的场合集体翻车。

与「中断之后能不能接手」那一篇。 那篇测能不能从当前位置继续，本文测能不能在新情境重新开始。判决点：接手能力高而复推率低的人，能沿着已有状态往下走，却在情境变形时无法重建路线；两个量正交，可以在同一批人身上分别测。

与「闭合之后还能不能回来」那一篇。那篇的重复产生率量的是必须重造多少已有信息，本文的复推率量的是能不能重造。判决点：一个记录完备、重复产生率极低的系统里，接手者不需要复推——材料都在。因而两个量此消彼长：留痕越好，对复推的依赖越低。这给出一条两篇合起来才有的推论：一个高度依赖个人复推的组织，通常是留痕很差的组织，而它会把这个缺陷读成「我们的人经验丰富」。

十三、本文自己的复推率

本文提出的这条依据——学会等于重新推出的可靠性——只在一个情境里被推出过一次，就是写作它的一次。按它自己的判据，它的复推率不可知，因为没有第二次。

更具体地说，本文有一处结构性的自指困难：它主张陈述不是学会的证据，而它本身是一份陈述。读者读完之后能把「复推率」讲得很清楚，这在本文的判据下恰恰是左下格的读数，不是学会的证据。

能做的对冲只有一条，且它必须由读者完成而不是由作者完成：合上本文，取一个本文没有提到的领域，构造两个表面不同、结构相同的情境，判断本文的第一根轴在那里能不能取到值。若取不到，说明这条依据没有被推出来，只被记住了。作者无法代做这件事，这也正是本文自己主张的直接后果。

十四、可以判本文错的六种结果

第一，最便宜、今天就能做的一条。 在同一批学习者上比较两种测法：一次迁移题 vs 两次表面差异覆盖三项的迁移题（总题量相同）。本文预测后者的分数与陈述式测验的相关显著低于前者，且对三个月后表现的预测更好。若两种测法给出的排序高度一致，复推率没有增量价值。

第二，前后差框架足够。 若把「帮助后的那份规则」改成事后补写之后，迁移效果显著下降，则起作用的是前后两份的比较而不是那次真实推导，本文的替代解释被推翻。

第三，语言化没有代价。 若在概念推理任务上（不只在感觉运动任务上），要求报告推导过程的组与不报告的组，其后续复推率没有差异，则第八节第一处收窄不成立，本文对陈述的整套怀疑失去根据。

第四，说得清楚就是推得出来。 若在控制练习量之后，陈述质量与复推率高度相关（相关系数稳定高于零点八），则两根轴合并，四格表退化，本文的核心区分消失。这一条是本文最容易被击中的地方。

第五，个人单位成立。 若在团队作业的场所，团队复推率可以由成员的个人复推率简单加总或取最大值预测，则第八节第二处的单位问题不存在，本文可以退回个人属性。

第六，本文自陈最软的一处。 「结构相同」若在实践中无法由第三方一致判定（不同第三方对同一对情境是否结构相同意见分歧很大），则整个测量无法执行，本文只是一个无法操作的说法。这一条应当在做任何其他检验之前先测——它是本文的前置条件。

正向读数。 若本文成立，增加情境表面差异的覆盖面应当提高复推率，而这个提高应当先于陈述质量的提高出现，甚至可以在陈述质量完全不变时出现。顺序比相关更难被巧合满足。

写死日期的赌注。 到二〇三二年十二月三十一日，至少应有一个主流的学业或职业能力评估体系，把「同一依据在两个以上表面不同情境中的重新推出」作为一项独立记分项，而不再只用单次迁移题与陈述式作答。若届时没有这类实现，且第一条对照的差异被证明不存在，本文关于学习单位是推导路径而非规则存量的主张即告失败。

结 论

学习的最小单位是什么，三个领域给出三个不相容的答案：规则的前后差、无法语言化的内隐适应、以及根本没有可保存的规则。三个答案都只对了一段，因为它们共享一个未被说出的前提——规则是一个被存在某处、可以被取出比较的东西。

例程研究推翻了它：抽象规则不是模板，它在每一次执行中被重新推断出来。运动学习补上另一半：真正稳固的那部分适应不经过语言，而把规则说出来并直接套用会挤掉它。先测效应补上第三块：先行判断的价值不依赖它的内容质量——起作用的是走了一遍推导，不是留下了什么内容。

于是学会了没有，不看现在能不能说出那条规则，看换一个表面不同、结构相同的情境时，他会不会重新推出同一条，且至少要看两次。而能把规则说得越清楚的人，复推率未必越高——说得清楚是复述的证据，不是复推的证据。

想提高它，也不必增加练习量或提高难度，只需增加情境表面差异的覆盖面。这条处方几乎不花钱，而它测到与练到的，是所有陈述式评估都测不到的那一维。

注 释

1. 本文所说的「同一条依据」由功能定义：两次判断在各自情境里作出同样的区分即为同一条，措辞不同不影响判定。
2. 运动学习部分的结论限于感觉运动适应任务，把它推广到概念推理属于本文的主张，尚未被该领域的证据直接支持。
3. 例程研究部分只使用「抽象面由执行反复推断而非被存储」这一骨架，不涉及该领域内部关于例程变化机制的争论。
4. 第五节四格表尚未经过信度与效度检验，是研究设计的起点，不是成熟的评估工具。

参考文献

1. Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *Journal of Neuroscience*, 34(8), 3023-3032. jneurosci.org

2. Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1), 94-118. sagepub.com
3. Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 989-998. bjorklab.psych.ucla.edu
4. Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, 15(3), 243-257.
5. Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379-424.
6. Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the Real World* (pp. 56-64). Worth.
7. Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1-38.

人机分工声明

本文由王德生提出研究方向、承重判断与最终发表责任；Claude 完成三个领域的选源与冲突定位、公开文献检索、命题成形、二维表构造、逐领域兑现、分界与证伪设计以及正文写作。第八节两处收窄由材料逼出，其中第二处动摇了本文的测量单位并被明确交出。运动学习部分的结论限于感觉运动适应任务，将其推广到概念推理属于本文的主张而非该领域的既有结论，已在正文与注释中标明。本文未标注任何评分——按本站惯例，参与改写者不为自己改写的文本打分。

证据边界 运动学习部分的结论限于感觉运动适应任务，把它推广到概念推理属于本文的主张，尚未被该领域证据直接支持；例程研究部分只使用「抽象面由执行反复推断而非被存储」这一骨架。第十四节六条判据本文**均未执行**，其中第六条应当先于其余五条执行——它是本文的前置条件。

字数 纯汉字约 1.1 万 **版本** v1.0 · 2026-08-01

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