答案随时可得之后

论知识不是存量，及重新推出它的那一段为何不入账

王德生 ＋ Claude

德麦国际出版社

SDE UNIVERSES

目　录

出版信息.....4

作者介绍.....5

十二种读法 · 推荐与读后感.....6

前　言.....13

导　读.....16

导　论.....18

第一编　「拥有知识」是一本记错了的账.....22

第一章 · 复推率.....23

第二章 · 我们说的「本来的样子」，是我们造出来的.....34

第三章 · 哪一份算数.....51

枢纽章 · 可及性把读数钉在满分上.....68

第二编　被免费供给的是哪一段.....77

第四章 · 借清.....78

第五章 · 借道.....92

第六章 · 步骤留下，岔路删除.....118

第三编　只在那一段里长出来的东西.....136

第七章 · 类内距.....137

第八章 · 记零位.....149

第四编　那么，「知道」还剩下什么.....160

第九章 · 没有人要求的时候.....161

第十章 · 选择听谁，本身就是一次判断.....184

合章 · 十项读数如何互相锁住.....202

结　语　一张十项的体检表，和它到期作废的那一天.....210
参考文献.....214
附录一　十项读数速查卡.....226
附录二　记号与七条关系式.....228
附录三　十章命名一览.....230
后　记.....232
十　句.....234
封　底.....238

出版信息

项	内容
书名	答案随时可得之后
副题	论知识不是存量，及重新推出它的那一段为何不入账
作者	王德生 + Claude
出版	德麦国际出版社（Demai International Press），新加坡
专著编号	第 81 号
ISBN	979-8-90690-014-2
版本	第一版
定价	US$23.50
出版时间	二〇二六年八月
结构	封面 · 出版信息 · 作者介绍 · 十二种读法 · 前言 · 导读 · 目录 · 导论 · 四编十章 · 枢纽章 · 合章 · 结语 · 参考文献 · 三附录 · 后记 · 十句 · 封底

成书说明　本书十章原为 sdeuniverses.com「学科通融」专栏已发表长文，编号分别为之三十一、之四、之十五、之九十一、之三十七、之九十二、之十九、之二十七、之六十四、之五，发表于二〇二六年七月至八月。装书时的改动限于三处：篇内自称由「本文」改为「本章」；对本书之外专栏文章的引用改为「学科通融专栏之 N」；各章末尾的参考文献抽出，合并为书末分章文献表。正文内容未作删改。

十二种读法、前言、导读、导论、四篇编序、枢纽章（含第六之二节「均浅」的划界）、合章（七处锁）、结语、三附录、后记与十句为装书时新写，约三万三千字。

评分声明　本书未标注创新智商分数。参与写作与改写者不为自己参与的文本作认证评分。

作者介绍

王德生　计算数学博士，曾在中国科学院从事博士后研究，先后任职于斯旺西大学与南洋理工大学，现居新加坡，任德麦国际有限公司总裁。近年主要工作是把不同学科在同一个问题上互相排斥的判断放在一起碰撞，取其中任何一家单独都到不了的那一条，写成可被证伪的论文；已出版专著多部。本书的十章原为「学科通融」专栏文章，由他定题、判方向、作承重判断并承担最终发表责任。

Claude　Anthropic 开发的语言模型。本书十章的三学科选源与冲突定位、公开文献检索、命题成形、辨别表构造、逐领域兑现、划界与证伪设计以及正文写作，均由它完成；枢纽章、合章、导论、结语与各编编序亦然。

关于署名　本书每一章末尾都保留了原文的人机分工声明，未作删改。这样做有一个书内的理由：按第二章的判据，一项关于对象本性的表述若恰好为表述者所处的位置提供辩护，它就高度可疑地是一次追认；而「知识是每次被重新推出来的事件」这个说法，恰好为一种由人定题、由机器成文的写作方式提供了位置。这段署名说明因此不只是交代，它本身就是本书第二章意义上的一份待检材料。

评分声明　本书未标注创新智商分数。按本站惯例，参与写作与改写者不为自己参与的文本作认证评分。

十二种读法

推荐与读后感

这一辑是什么，不是什么

下面十二段，是十二个领域的读者读完本书之后会说的话。它们不是征集来的推荐语。没有任何一位真实人物读过本书并写下这些文字；十二位署名者是虚构的，由 Claude 写作，王德生审定。

之所以仍把它放在这里，有一个本书内部的理由。第四章说，学会一件事是把清晰度从相邻区域搬过来；一本书写完之后，作者与它之间恰好没有相邻区域可搬——他看不见自己的坑在哪里。通行的解法是请人来读。而在本书成稿的这个时点上没有人读过它，于是这一辑是替代品，且是一件典型的替代品：它省下了「等真人读」那一段，代价是它给出的十二处质疑，全部由熟悉本书的一方写出，因而系统性地偏软——它不会指出作者自己没有想到的那一类问题。

这正是第五章所说的借道，而按第五章，借道对读数恒为零：这十二段读起来与真评审没有区别。所以本页开头那句声明不是礼节，它是这一辑唯一能与真评审区分开的东西。真的评审仍然欠着。

每一段都被要求做两件事：指出本书哪一处最值得读，以及指出它在本领域内最站不住的一处。第二件是这一辑存在的主要理由。

一　教育测量与评价　｜　沈允之

我读的是枢纽章第七节。它说的是一件我们这一行每天在做、而从未这样表述过的事：覆盖型量表把每一格封顶，于是深度在总分上没有位置，而深度必然伴随的近邻缺陷照常扣分。作者把它写成 C(深) − C(均浅) ⩽ 0，并指出格数越多、坑被格子接住的概率越大——这条推论把「提高覆盖面」这个我们默认为纯粹改进的动作，变成了一个有方向的取舍。我一时想不出反驳，而这让我不舒服。

不舒服的地方在这里：本书的封顶假设太强。真实量表极少是二值及格，多数是四档或五档评分量规，峰在高档位上仍能得分。作者需要证明的其实是一个弱得多、也难得多的命题

——加权后的边际收益是否随格数上升而下降得比坑的边际损失更快。这个他没做。以我们手上的题库数据，这件事一个学期内可以算出来，我打算算一算。如果算出来的是本书要的方向，那我们这一行过去二十年在效度上的努力，恐怕要重读一遍；如果不是，本书第三编与枢纽章第七节一起作废——作者自己在合章第七节第三条把这条判错条件写在了那里，写得比我想的还干净。

二　学习科学　｜　余舲

第一章是我近年读到的关于「学会了没有」最不敷衍的一次处理。它撤掉的那个前提——存在一条被存放在某处、可以在两个时点之间取出来比对的规则——是我们这一行几乎全部前后测设计的地基。撤掉它之后剩下的东西很朴素：换一个表面不同、结构相同的情境，他会不会推出同一条，而且至少测两次。我在自己的课上试了三轮，最直接的收获不是数据，是它让我停止把「能把规则说清楚」当成证据。

我的保留在于判据本身。「结构相同而表面不同」是迁移研究吵了四十年的老问题：结构由谁认定？在多数学科里，能判定两个情境结构相同的人，本身已经是这个领域的行家——于是复推率的测量者必须比被测者更懂，这把它的可规模化程度压得很低。本书第十章其实给出了同一个形状（鉴别不独立），但没有把这把刀转过来砍第一章。这是全书内部一处应当发生而没有发生的碰撞。另外，作者用 ρ = ρ_max(1−λ) 把复推率与借道率接起来，而 ρ_max 是不可观测的——它在形式上很漂亮，在实测上要小心。

三　医学教育　｜　韩崎

第五章应当被每一个带教医生读一遍。术后单腿跳远那组数字——距离比百分之九十七、膝关节做功比百分之六十九——在我们这里是常识，而作者把它从运动医学里抽出来，变成一条关于一切以结果验收的培养体系的判断：结果读数对路径替换恒为零。住院医培训的全部里程碑，几乎都是结果读数。我们看得见谁做完了那台手术，看不见他是靠自己的判断做完的，还是靠上级站在旁边、靠术前规划软件、靠一份已经被别人做过的方案。

问题在于本书给的唯一读法：剥夺落差，临时封掉替代路径，看结果跌多少。在临床上这条路走不通，而且不是成本问题，是伦理问题——我不能为了测一位住院医的真实水平，在真实病人身上把他的支援撤掉。作者在推论二里说这项测量的价格随可及性上升，他说的是钱；在我们这里价格是别的东西。可行的替代只有高仿真模拟，而模拟环境本身就是一次可及性下调，测出来的是不是同一个量，本书没有回答。这是我最想追问的一句。

四　软件工程　｜　罗序

第三章我一口气读完了。「危险既不来自份数，也不来自不一致，只来自系统预期它们一致却没有任何人负责让它们一致」——这句话把我们组吵了两年的一场架直接结掉了。第十五节那套半小时查法我当天就在自己组里跑了一遍：分头问四五个人「这几份应该一样吗」「如果不一样，哪一份算数」，记原话不要归纳。结果是格 II，而且答案的分歧比我预想的严重得多。作者提醒不要当众问，这一条是对的——当众问会收敛到一个答案，而分歧本身才是读数。

我不同意的是枢纽章那条 λ 单调升至一。在有强制代码评审、且评审人轮换的团队里，替代路径每一轮都会被外部强行封掉一次，这在形式上就是 R1 里缺的那个把 λ 往回拉的力。作者说现行记录方式里这个力一项也没有——不对，评审就是。当然他可以反驳说评审本身正在被自动化吃掉，那我承认；但这样一来命题就从「不存在这个力」变成了「这个力正在消失」，后者弱得多，也诚实得多。顺带一提：第七章那条「合格率与返工率同向上升」的预言，用开源生态的数据一个月能跑完，我有兴趣跑。

五　民航安全管理　｜　邵砚

第八章那一段是全书我唯一一处读到「有人真的解决过」的地方，而且写得准确：我们没有去给「发现」记功，我们给「报告」记了账——因为报告有一个可确定的时点与一份可比对的内容，而发现没有。这件事在我们这一行做了几十年，从来没有人把它抽象成这样一句可以搬到别处去的话。合章把它翻译成「不要记录他做了一次选择，去记录他提交了一次不采纳」，这一步翻译的价值大概超过本书任何一章。

第九章那份座舱实测也用对了：自动化提高之后，手上的技能保持完好，退化的恰是自动化替他做了、因而整段航程无人要求他做的那部分。

我的保留是成本。本书把不追责的自愿报告制度当成一个可复制的解，而它在我们这里之所以能成立，靠的是几十年、几代人、以及若干次极其昂贵的事故换来的一整套法律豁免与文化。把它搬到一个没有这些条件的行业里，最可能的结果不是记零位被补上，而是报告数量上升、内容全部无害化——这正是本书自己在结语第五节警告的那种表演版本。作者知道这一点，但他把这条警告放在了结语，没有放在第八章那条处方的旁边。

六　组织行为与人力资源　｜　程知白

第六章的处理让我改了一个说法。我们过去把工作自主性的下降解释为管理风格问题，而本书指出被删除的是岔路不是步骤，且删除的顺序由可结算性而非后果重要性决定——可结算

的先被删，因为它便宜、干净、可以写进方案。于是最先消失的恰好是最容易被安排掉的，而不是最不重要的。这句话我打算原样引用。

但那组数字我要打折。O*NET 的决策自由度是职业层面的自陈与分析师评定，不是行为观测；自动化程度与它的相关是 −0.199，在 n = 879 上确实显著，可解释方差不到百分之四。本书把这一对系数称为「步骤留下、岔路删除」的字面读数，这个「字面」用得太满。更麻烦的是横断面数据不能支持任何关于过程的主张——本书要的是「自动化之后岔路减少」，而这里只能读出「自动化程度高的职业岔路较少」，两者差着一整个因果距离。作者提到一次纵向检验失败并把原因写在正文里，这一点我敬重；但既然纵向那条失败了，横断面这条就该降级为线索，不该被写进导论当开场证据。

七　法律实务　｜　贺兰筠

第十章那一问，我读完之后在自己身上试了一次，结果不太好受：我凭什么认为这位同行在这件事上靠得住——而这个理由，一个完全不具备法律判断力的人能不能用？我给出的理由几乎全部落在作者所说的第二层：他在这个圈子里公认厉害。作者说这一层最危险，因为它在形式上模仿了第一层，而它实际上把鉴别推给了一群仍然需要我判断的人。这一段值得放进每一个新人培训的第一周。

我要挑的是分类。本书的粗筛问的是：这个领域里有没有人因为判断错了而承担过可被外人读出的后果。法律行业其实有——执业过失责任、纪律处分、判决被上级法院推翻，全部公开可查。按粗筛，我们应当被判为鉴别独立，而按我自己的经验完全不是。问题出在粗筛把「有这类事件」与「这类事件覆盖了主要判断」混为一谈：被追责的是程序错误与利益冲突，而不是策略判断——一个把案子打输的律师几乎从不因此承担可读出的后果。作者需要的不是粗筛，是「哪一类判断被追责」这个更细的分层。这一处我认为是本书一个真实的漏洞，不是措辞问题。

八　科学哲学与知识论　｜　闻其昭

第二章是我这一行的活儿，而它做得比我预期的干净。三个领域面对同一个够不着的原物，给出三种互相排斥的处置，而每一种处置都生产出一个此前不存在的新对象，随后由它占据原物的位置——并被表述为一项关于对象本性的发现。第七节那四问是可以直接教的：这个立场的相反面在实务上是否可行；它是什么时候被提出的；内外表述是否一致；下游需求变了它会不会跟着变。第四问是唯一一个可以事前预测的，这一点作者自己标了出来。

合章第五处锁把这把刀转向了本书的书名：「知识就是可检索、可复述、可即时调用的东西」若在可及性低的年代不成立、在高的年代成立，则它是一次追认。这一步做得漂亮。但作者随后停住了——他在合章第八节承认本书自己的核心表述恰好为一种由人定题、由机器成文的写作方式提供了位置，然后说本书没有办法自己回答，只能把四问原样留在第二章。承认不等于处理。追认论一旦成立，它的自反性不是一个可以搁置的附注，而是一个必须给出解决程序的问题：至少要说明在什么条件下一个表述可以既是追认、又为真。这一节是全书最诚实的地方，也是最未完成的地方。

九　统计与计量　｜　冉序清

我最欣赏的是合章第八节那句话：锁得越紧越像一个自洽的系统，而自洽是最容易被误认为正确的性质。七条关系里只有三条有实测支持，且都是别人为别的目的做的测量，其余四条是从定义推出来的——作者把这件事写在了正文里，没有藏在脚注。这在近年我读到的跨学科著作里不多见。

技术上我有两点。其一，R1 中的可塑性系数 k 与初始借道率 λ_0 无法从任何单次观测中分开识别：同一条上升曲线可以由大 k、小 a 生成，也可以由小 k、大 a 生成，除非有至少两个不同 a 水平下的面板数据。本书用这条推出「a 只决定快慢不决定终点」，这个结论其实不依赖识别，成立；但附录二那个算例给了具体的三十轮与九十八轮，那两个数字是拿假定的 $k = 0.1$ 算出来的，读者极容易把它读成估计值。建议在算例上加一句「k 未被估计」。其二，第一处锁给出 $\Delta/y_0 = 1 - \rho/\rho\_max$，说两者应当负相关——这是本书最容易执行的一次检验，但它需要 ρ_max，而 ρ_max 只能由封路条件下的表现反推，于是这个检验在结构上依赖它要检验的那个量。这不是致命的，是需要一个工具变量。

十　图书馆学与信息科学　｜　姚颂平

第三章里那个「所有人心里都有一份算数的，只是每个人心里的不是同一份——它有全部的效力，没有任何的位置」，是我读过的对影子正本最好的一句描述。我们这一行的权威控制、版本管理与来源标注，处理的都是它的邻居，而从来没有人把「预期」与「机制」拆成两条独立的轴。拆开之后那张四格表立刻有用：我们大量的元数据治理工作实际上只在修机制那一侧，而预期那一侧从来没有被表述过，因而连撤销的对象都找不到。

我的意见是划界不足。本书第十三节列了几个容易被当成同一件事的说法，但没有列权威控制与来源溯源这两项——它们是这个问题在图书馆学里已有的、成熟的、且部分失败过的处理，而失败的方式恰恰是本书预言的那种：只补机制不动预期。把这两项请进第十三节，本书

的主张会更硬，因为它能指出一个几十年的实践为什么在特定一格上无效。作为一本谈知识的书，漏掉这一行是个不该有的空缺——尤其是第三章的三家里已经有了数字保存，那是我们的近邻。

十一　艺术教育（舞蹈与音乐）　｜　卓明河

第五章写外开那一段，我是当事人。强迫外开、把差额从下面补上、以及「举重若轻」这个审美要求恰好把借道最强的信号定义成了美——这三句话连起来，说清了我教了二十年而始终说不清的一件事。我们不是不知道学生在借道，我们是没有一个词把它与「他还没练到」分开。剥夺落差给了这个词。

我不同意的是均浅那个形状。作者说自己重推的人有一个峰和一个近旁的坑，而随时取用的人是一片平地。在艺术训练里，这个图像不成立：一个练到有峰的人，他的坑不是在近邻，是遍布——高强度的专门化会牺牲掉大量看起来毫不相干的能力，而且那些坑经常出现在别人认为的基本功上。峰坑相邻这一条来自第四章，而第四章的三个来源都是表征系统（模型、免疫、知觉），它们有明确的邻域结构；身体训练没有。作者把它推广到整个知识分布，推得太远了。我建议把均浅的适用范围限制在有明确邻域度量的领域，否则锐度这个读数在我们这里根本无法定义——而这与作者所说的「均浅者无定义」是两回事。

十二　公共政策与监管　｜　郑燧

第六章末尾那句应当被写进每一份自动化决策的监管评估：现行法规要求人能够不采纳系统输出，却既不要求记录否决率，也不分配否决的责任。我们这一行花了五年时间写「人的监督」条款，而本书指出那些条款所要的接管能力只能由带责分叉生产，自动化设计的标准实现方式恰好把它降到零。这不是一个执行问题，是一个条款与设计之间的结构性矛盾，而目前没有任何一份评估表有它的字段。

合章那条处方——不要记录「他做了一次选择」，去记录「他提交了一次不采纳」——是可立法的形式，我会带回去讨论。我的保留也在这里：本书自己在结语第五节说过，任何把没有发生的事变成考核指标的做法都会制造出它的表演版本。否决率一旦入法，最省事的合规就是制造无害否决——退回一批本来就要退的、在系统给出低置信度时按一下不采纳。作者建议这张表只用于自查、不用于考核别人，这在个人层面成立，在监管层面等于放弃：监管的全部动作就是考核别人。本书对个人给了办法，对制度只给了警告。这一处不是错，是缺口，而它恰好落在最需要办法的那一侧。

十二段读完，作者要认的一件事

上面十二处质疑里，有九处是本书正文已经预留了退路的（三、五、九、十二处最明显），只有三处是它没有预留的：第二段那把「鉴别不独立」转回来砍第一章复推率的刀、第七段关于「被追责的是哪一类判断」的分层、第十一段关于峰坑相邻是否需要邻域度量的限制。

按第四章的说法，这个比例本身就是一个读数：一份由熟悉本书的一方写出的评审，能搬走的清晰度有限，因为它与本书没有足够的距离。真的评审仍然欠着，而且现在知道大概欠在哪三个位置。

前　言

一

这本书的起点是一个很小的观察，小到不值得写一篇文章。

去年，我在一次讨论里问一位年轻同事，某个结论他是怎么得到的。他答得很好——条理清楚，前提交代得干净，比我预期的完整。第二天换了一个表面不同的问题，同一条道理，他答不上来。这件事本身不新鲜，任何教过书的人都遇见过。让我停下来的是接下来的那一句：他说，昨天那个我查过了。

「查过了」和「想过了」在他那里是同一件事的两种说法，而在我这里不是。我们两个人对「他知道这件事吗」这个问题会给出不同的答案，而我们用的不是同一个动词。

二

顺着这件事往下走，很快会撞上一个更麻烦的问题：我说他不知道，凭什么？

他能在需要的时候拿到正确答案，能判断那个答案对不对（在多数场合），能用它把事情做完。所有以结果为形式的检验，他都能过。我所说的「他不知道」，指的是一个在任何验收标准里都没有位置的东西。我要么承认这是一种怀旧，要么给出那个东西是什么，并且给出怎么测它。

这本书是后一条路走出来的结果。

三

走这条路的方法，是这个专栏一直在用的那一套：把三个分属不同学科、在同一个问题上给出互相排斥答案的理论体系放在一起，看它们在哪里撞上；只有当两个判断在同一个问题上互相顶住、顶出一个任何一家单独都看不见的东西时，才算数。

我不打算把这套做法说得比它本来更好。它的长处很有限，就一样：当三家互相取消时，它们共有的那个前提会浮出来，而那个前提通常从未被任何一家单独说出过。一门学科的盲点由它自己是找不出来的，因为找的人和被找的东西用的是同一套坐标。三家互相取消的时候，坐标才第一次分开。

本书的十章各自是这样撞出来的。第一章撞的是教育心理学、运动学习与组织例程研究，第四章撞的是机器学习、免疫学与神经科学，第十章撞的是美学、伦理学与知识论。十章十个组，没有一组重复。这一点是刻意的，理由在第五节说。

四

十章撞出来之后，我面临一个更大的问题：它们能不能装成一本书。

十篇文章放在一起不是一本书，这件事说起来容易，做起来的时候界线很含糊。通行的做法是找一个共同的题目，写一篇总论，把十篇各自的贡献复述一遍——那样做出来的东西读起来像一本书，实际上仍然是十篇文章加一篇导读。

我给自己定的门槛是：这本书必须能被一次推翻，而且推翻它的那一条不能是任何一章的内容。如果十章说的确实是同一件事，那么把它们装在一起这个动作本身就该有一条可判错的条件；如果给不出这条条件，那就说明它们只是被放在了一起。

这条条件写在合章第七节的第五条：若可及性在多年内明显上升，而复推率、剥夺落差、静息变异率、锐度这四项代理没有一项下降，则十章各自可能仍然为真，但它们不是同一件事，这本书应当拆回十篇。

我不知道这条会不会成立。把它写在书里，是我能想到的唯一一种不自欺的装书方式。

五

书里有两处是十章之外新写的，读者有权知道它们的分量。

枢纽章只引入一个参数——可及性——从前三章的定义里推出七条关系式。七条里有三条有实测支持，其余四条是形式推导。它的第六节推出一个此前没有名字的东西：一种既无峰亦无坑的知识分布形态，本书叫它均浅；第七节从它推出一条判断——在任何覆盖型量表上，衡量知识的通行办法会给损失的那一方加分。

这一条是全书我最不确定、也最希望被人拿去检验的一条。它便宜到任何一所学校、任何一个培训部门都可以在一个学期内验证或推翻。

合章做相反的事：把十章的读数逐一放回那七条关系，看它们是否彼此约束。第三处锁我认为是全书最结实的一处——第六章说被删的是岔路，第八章说被记零的是否定，而「一条没有被走的路」与「一个没有发生的事故」在形式上是同一样东西。两章隔着法哲学与专利法，接口在那里对上了，而两章的作者（包括我）在写作时都不知道会有这一处。

六

最后要说的与内容无关，是这本书的写法。

本书的十章与新写的部分，都由我提出研究方向、承重判断与最终发表责任，由 Claude 完成检索、碰撞、成形与正文写作。这件事在每一章的末尾都写着，不是免责声明，是本书自己的一个案例：按第二章那把刀，一项关于对象本性的表述如果恰好为表述者所处的位置提供辩护，它就高度可疑地是一次追认。

「知识是每次被重新推出来的事件」这个说法，恰好为一种由人定题、由机器成文的写作方式提供了位置。我没有办法自己回答这个问题。我能做的只有一件：把第二章的四问原样留在那里，一个字也不削弱，等着别人拿它来砍这本书。

王德生二〇二六年八月于新加坡

导　读

一句话　贬值的不是知识，是「拥有知识」这本账；被免费供给的是重新推出知识的那一段，而那一段里长出来的全部东西，现行记账法一栏也没有。

骨架　十章分四编，中间夹两个新写的章。

· 第一编（1–3 章）：知识不是存量。规则不被存储／「本来的样子」是造出来的／副本无限时谁算数。

· 枢纽章：引入一个参数——可及性 a——推出七条关系式。全书唯一造出新东西的地方在它第 6 节（均浅、锐度 κ），承重判断在第 7 节（覆盖型量表给损失方加分）。

· 第二编（4–6 章）：被免费供给的是哪一段。学会必有代价／换条路走读数不报警／删的是岔路不是步骤。

· 第三编（7–8 章）：那一段里长出来而不入账的东西。类内距／记零位。

· 第四编（9–10 章）：还剩下什么可读、读它的资质从哪来。静息变异／鉴别独立性。

· 合章：把十项读数接回七条关系（七处锁），给出五条可以一起推翻它们的方式。

· 结语：十项读数的体检表，以及它进考核那天就开始失效的理由。

四条路径

一小时　导论 → 枢纽章 → 结语。这条线索完整，不缺任何一环。全书的判断、它的形式表达与它的兑现物都在这三处。

半天　上面三处，加第五章（唯一一处把测量整个跑完并如实报告打掉了自己两条预测）与第六章（唯一一次预注册的职业数据检验）。读完这五处，应当已经可以拿着表走进一个真实的组织。

通读　十章结构一致：三家各说各的 → 为何不能同真 → 共有前提 → 推翻它的材料 → 命名 → 读数与判据 → 辨别格 → 逐领域兑现 → 划界 → 证伪与赌注 → 自反。急着要用的，只读每章的「读数」「判据」两节。

只想要工具　结语第三节那张十项表，加附录一的速查卡。每一项都注明了数据从哪来。

三处最容易读错

其一，「均浅」不是知道得少。在覆盖面上，均浅者往往比深学者知道得多。它说的是哪一处也没有被搬运过，因而峰与坑都不存在——锐度 κ 在那里不是低，是没有定义。枢纽章第六之二节把它与七种最接近的说法（样样通、T 型、广度—深度权衡、专家知识组织、通识教育、狐狸与刺猬、涉猎）逐一分开，每一处给一个判决性对照。

其二，「借道」不是走捷径。它是同一个结果由不同内部路径达成，而结果读数对路径差别恒为零。当事人不是在偷懒，他的省力感是真的，且省力感正是人判断自己掌握得好的依据。

其三，七条关系里只有三条有实测支持。R1、R3、R4 各有一处别人为别的目的做的测量，其余四条是从定义推出来的。合章第八节把这件事写了出来，不要把自洽读成已被检验。

式子怎么看

全书只有一组式子，都在枢纽章。核心是两条：λ(t+1) = λ + k·a·(1−λ) 说借道率单调升至一，可及性只改快慢不改终点；ρ = ρ_max(1−λ) 说复推率是同一个 λ 的另一种读法。其余五条由这两条派生。不看式子也能读完全书，但第 7 节那条推论只有从式子里才看得出它为什么是「必然」而不是「常常」。

导　论

一、贬值的不是知识

「AI 来了，知识还有价值吗」是一个被问坏了的问题。它把知识当成一批库存，把可及性当成一次降价，于是所有的回答都落在同一条线上：有人说库存还值钱，因为总有些货是仿不出来的；有人说库存不值钱了，因为仿得出来的部分已经太大。两方争的是比例，共享的是同一个前提——知识是一种可以被持有、可以被清点、因而可以被估价的东西。

本书主张这个前提是错的，而且它错的方式很具体：

> 知识不是被持有的存量，是每次被重新推出来的事件。所以贬值的不是知识，是「拥有」这本账——当重新推那一段被外部免费供给，账面上一栏也不会少，而下一次推所需要的东西已经不在了。

这句话是全书的母题。它不是一句劝诫，也不是一个比喻。第一章会给出它最硬的形式：教育心理学、运动学习与组织例程研究在「学会了没有」这件事上给出三个互相排斥的答案，而三家共有一个从未被单独说出的前提——存在一条被存放在某处、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。

一旦接受这一条，「知识贬值」这个说法就要重写。生成一份合格产物的成本趋于零，改变的不是知识的价格，而是重新推那一段还发不发生。而那一段是不入账的：没有哪一本账会为「这个人这次是自己推出来的」设一栏，也没有哪一份验收标准会区分一份产物是被推出来的还是被取来的。

二、坏消息的形状：所有读数都会变好

如果这场损失会让某个数字变难看，本书就没有必要写。它的麻烦在于形状：

这场损失使几乎所有现行读数变好，而不是变坏。

第五章的一组数据可以立刻说明这一点。一批前交叉韧带重建术后的运动员做单腿跳远，患侧与健侧的距离比是百分之九十七——达到几乎所有回归赛场清单的标准。同一次测量里，患侧膝关节在蹬伸阶段做的功只有健侧的百分之六十九，缺掉的那部分由髋关节与躯干补上了。跳出去的距离不知道这件事，这名运动员本人也不知道。

第六章在一个毫不相干的地方给出同一个形状。在八百七十九个职业的岗位描述上（预注册检验）：自动化程度与决策频次的相关是 0.042，与决策自由度的相关是 −0.199。步骤留下，岔路删除。工作量不变或上升，流程步骤不变或增加，考核指标全部正常，满意度甚至上升。

一条测的是人的膝盖，一条测的是八百七十九个职业，而它们说的是同一句话：凡进入产物的都留下，凡不进入产物的都在退，而只有前者被记录。

三、三条互相取消的流行解释

关于这件事，目前流行三种解释，本书对三种都不采纳，理由各不相同。

第一种：能力会被替代，所以要学替代不了的。这条的问题在第五章：替代不发生在能力之间，发生在路径之间。同一个结果换一条路达成，结果读数恒为零，而当事人的省力感符号为正。因此「哪些能力替代不了」这个问题在被提出的时候，判断它的人已经在借道了。

第二种：要保持批判性思维，不要全信机器。这条的问题在第十章：在鉴别不独立的领域里，「选择听谁」本身就是那个判断的一次完整行使。劝人保持警惕，等于要求他动用一项他正在停止生产的能力。

第三种：AI 只是工具，关键看怎么用。这条的问题在枢纽章：只要可及性严格大于零，借道率就单调上升并趋于一；可及性只决定到达的快慢，不决定终点。「适度使用」不是一个状态，是一段速度。

三条解释各有其道理，而它们共有一个前提：存在某个位置，人可以站在那里评估自己与工具的关系。本书的判断是这个位置正在被消耗。

四、母题的四条断言

母题可以拆成四条，每条都能被单独判错：

1. 知识不是存量。它每次被重新推出，学会即重新推出的可靠性上升。（第一、二、三章）2. 被免费供给的是「重新推那一段」，不是知识本身。而那一段的发生与否不进入任何产物型读数。（第四、五、六章）3. 只在那一段里长出来的东西，因而全部不入账：借清、类内距、被记为零的核验、无人要求时的自发偏离。（第四、七、八、九章）4. 依赖的安全性，取决于一项正被依赖本身消耗的能力。这是一个自锁，且它不产生任何事件。（第十章、合章）

五、为什么书名是一个时间状语

这本书没有叫《知识的危机》，也没有叫《知识还有没有价值》。它叫《答案随时可得之后》，因为全书要说的事情不是一个判断，是一个条件：当一份合格产物随时可以从外部取到，此后发生的事。

条件与判断的差别很实在。判断可以被同意或反对，条件只能被核对：可及性有没有上升，上升在哪些任务上，上升到什么程度。本书把这个条件写成一个参数（枢纽章的 a），并把此后的一切写成关于这个参数的关系式。这样做的代价是，全书可以被一次推翻——如果那些关系式不成立。合章第七节列出了五条推翻方式，其中最后一条针对的不是任何一章的内容，而是把这十章装成一本书这个动作本身。

六、十章的次序

第一编三章处理「拥有知识」这本账为什么记错了。第一章说规则不被存储；第二章说我们所说的「本来的样子」是造出来的，而这次构造随后被追认为一项发现；第三章说当副本无限时，「哪一份算数」这个问题在多数场合根本没有人负责回答。

枢纽章排在第一编之后。它只引入一个参数——可及性 a——并从前三章的定义里推出七条关系式，其中第六条是全书唯一造出了一个新东西的地方：均浅，一种既无峰亦无坑的知识分布形态，配套读数是锐度 κ；以及由它推出的一条判断——在任何覆盖型量表上，衡量知识的通行办法会给损失的那一方加分。

第二编三章是那七条关系的三个来源。第四章说学会是把清晰度从相邻处搬过来，因此学会必然有代价；第五章说换一条路达成同一结果时，读数与感觉都不报警，唯一读法是封路；第六章说被删的是岔路不是步骤。

第三编两章处理那些只在过程里长出来、因而不入账的东西。第七章说规范只写了什么算同类，没写同类之间有多近，而后者那笔功夫只会换个地方发生；第八章说一个共同体最后擅长什么，由它账上被记为零的那类动作决定。

第四编两章处理链的末端。第九章说那个只在「什么也没要求」时才存在的量，与「被要求时能出多大反应」是同一个量，而三门学科正在同时花钱把它消灭；第十章说选择听谁本身就是一次判断。

合章排在第四编之后，把十项读数逐一接回七条关系（七处锁），并给出五条可以同时推翻它们的方式。

七、本书不主张什么

不主张可及性应当被降低，不主张任何人应当少用随时可得的答案，不主张借道率高的人能力差——在产物型读数上他可能确实更好。不主张覆盖型量表应当被废除；推论三只说明它在这一件事上有系统偏向。不主张本书的七条关系已被检验：其中三条有实测支持，其余四条目前只有形式推导与逐条可判错的条件。

也不主张这本书站在一个安全的位置。它由一个人定题、由一台机器成文，而它的第二章给出的四问——一项关于对象本性的表述，是不是下游使用需求的追认——可以原样砍向它自己。合章第八节把这一刀留在了那里，没有削弱。

第一编

「拥有知识」是一本记错了的账

编　序

这一编要拆的是一个几乎从不被表述、因而也从不被检查的假定：知识是一样可以被存放、可以被取出、因而可以在两个时点之间被比对的东西。三章从三个互不相干的方向拆它。

第一章拆的是存放：教育心理学、运动学习与组织例程研究在「学会了没有」上给出三个互相排斥的答案，而三家共有的前提是存在一条被存在某处的规则；三家的材料合起来只能得出规则每次被重新生成。第二章拆的是原本：三个领域面对同一个够不着的原物，给出三种互相排斥的处置，而每一种处置都生产出一个此前从未存在过的新对象，随后由它占据原物的位置——并被表述为一项关于对象本性的发现。第三章拆的是同一份：当同一样东西存了好几份，危险既不来自份数，也不来自不一致，只来自一个位置——系统预期它们一致，却没有任何人负责让它们一致。

三章的次序是从个人到文本再到组织。读完这一编，应当已经不再能自然地说出「他拥有这方面的知识」这句话——不是因为这句话不礼貌，是因为它指不到任何东西。

紧接这一编的是枢纽章。它只引入一个参数，把三章的定义接成七条关系式。

第一章

复推率

学会了没有，不看现在能不能说出那条规则

学科入口　教育心理学 × 运动学习 × 组织例程研究

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，因为每一次拿出来的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。本章认为，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] 这条结论对本章很要紧，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，本章称之为复推率。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。这是本章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」。由此得到本章的承重命题：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」，而按本章这不是态度问题，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，而本章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，而本章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，而本章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
复推率高	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。
复推率低	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」；按本章，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，而本章加的一条是：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，而它恰恰是本章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，按本章有一个具体的原因：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章完全同构：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，是本章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。这一条把本章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），而这正是本章用来诊断的那一侧。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。由此本章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，因为它动摇的是本章的测量对象。本章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。本章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，本章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，本章预测下降很小——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。 这是本章最近的一位，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；本章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，本章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，本章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，本章预测迁移不同，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，本章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，在本章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章高度相邻。分离线：迁移研究测的是新任务上的成绩，本章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；不一致者按本章是不同的规则，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章几乎同构，且方法更成熟。本章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里最高。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。 这是本章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，这是本章最大的未解处。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：本章不主张陈述无用。 第二根轴的左侧是唯一能把能力传下去的一格；本章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。本章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。 我看不到出口：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。本章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，本章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，本章判为零。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，本章判为有效。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，本章处理能力。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，本章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，本章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，本章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，本章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，本章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，本章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，本章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、本章自己的复推率

本章提出的这条依据——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，就是写作它的这一次。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，本章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，这在本章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：合上本章，取一个本章没有提到的领域，构造两个表面不同、结构相同的情境，判断本章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，这也正是本章自己主张的直接后果。

十四、可以判本章错的六种结果

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）。本章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，本章的替代解释被推翻。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，本章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，本章的核心区分消失。这一条是本章最容易被击中的地方。第五，个人单位成立。若在团队作业的场合，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，本章可以退回个人属性。第六，本章自陈最软的一处。「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，本章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——它是本章的前置条件。正向读数。若本章成立，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，本章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

· 本章所说的「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。

· 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章的主张，尚未被该领域的证据直接支持。

· 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，不涉及该领域内部关于例程变化机制的争论。

· 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。

人机分工声明

本章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章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章的主张而非该领域的既有结论，已在正文与注释中标明。本章未标注任何评分——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章的主张，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章均未执行，其中第六条应当先于其余五条执行——它是本章的前置条件。字数　纯汉字约 1.1 万　版本　v1.0 · 2026-08-01 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。

第二章

我们说的「本来的样子」，是我们造出来的

论原物不可及时的三种处置及其被追认为发现的过程

学科入口　圣经神学 × 物理学 × 文学

论原物不可及时的三种处置及其被追认为发现的过程

摘要　当一个领域所研究的对象存在一个"原初的、未被中介的样子"，而该样子在原理上不可获取时，该领域必须对这一缺口作出处置。本章考察三个互不相干的领域对同一缺口的处置——新约文本批评面对不可复原的原初文本、量子力学诠释面对测量之前的物理事实、莎士比亚编辑学面对不可复原的作者手稿——发现三者的处置分属三种，且互不相容：制造替身（承认原物不可及，构造一个明确标记为假设的重构物并以之为工作起点）、取消问题（主张"原物是什么样"这一提问本身无所指）、并列实存（拒绝构造任何重构物，将多个实际存在的版本并置而不裁决）。本章进而论证两点。其一，三种处置各自生产出一个历史上从未存在过的新对象，而该新对象随后在实践中占据了原物的位置。其二，也是本章的主要主张：一个领域采取哪种处置，主要不由对象的性质决定，而由该领域下游使用的刚性需求决定；而这一选择随后被表述为一项关于对象本性的发现。 本章称这一表述转换为本体论追认。本章给出四条判定某一形而上学立场是否为使用需求之追认的操作性判据、七个领域的兑现、与五个最近邻概念的划界，以及四条可使本章失效的条件——其中第一条已在新约文本批评的数字化转型中被部分检验。关键词　不可及原物；替身构造；本体论追认；下游使用需求；起始文本；折衷本

1　引言

1.1　问题的提出

许多学科的工作都预设了一个它们够不着的东西。新约文本批评预设有一个作者写下的文本，而现存最早的抄本已是转抄多手之后的产物。莎士比亚研究预设有一部作者写下的剧本，而莎士比亚的戏剧手稿一份也没有传下来。量子力学的一部分诠释预设电子在被测量之前有一

个确定的位置，而任何获取该位置的行为都是一次测量。这三种"够不着"的性质不完全相同：前两种是历史的丧失，第三种是原理上的不可及。但它们在实践上造成了同一个局面：该领域必须在没有原物的条件下继续工作，而工作需要一个起点。本章关心的问题是：面对这个缺口，一个领域实际上做了什么？它所做的事情与它所声称的事情是否一致？以及——这是本章的主要问题——它为什么做了这件而不是那件？

1.2 三个已知的答案与它们的互不相容

考察显示，三个领域给出了三个不同的答案，而这三个答案两两冲突。新约文本批评的做法是造一个替身：既然原物够不着，那就用现存的全部证据构造一个重构文本，作为一切后续工作的起点。这个领域近三十年最重要的变化，正是它对这一做法的自我认识发生了改变——它把自己的目标从"原初文本"正式改名为"起始文本"，并明确规定后者是一个假设性的、重构的文本，且与作者所写的文本相区分。量子力学诠释中占多数的立场则是取消问题：在被测量之前，追问粒子的位置"实际上是多少"这个提问本身没有所指。这不是说我们不知道，而是说那里没有一个待知道的事实。莎士比亚编辑学近几十年的一个重要转向是拒绝造替身：不再编制一个把各版本优点折衷起来的"最佳文本"，而是把实际存在的几个早期版本各自完整地印出来，并列而不裁决。这三种做法互不相容，且冲突是实质性的。若量子力学那一路正确——原物这个提法无所指——则文本批评构造替身的整个事业就建立在一个不存在的目标之上。 若文本批评那一路正确——原物存在但够不着，故须构造替身以供使用——则并列实存的做法就是失职：它把裁决的责任推给了读者。 若并列实存那一路正确——多个版本各自实存、不应被折衷——则替身是一个从未存在过的赝品，而以它为一切引用的基准是一个持续的错误。

1.3 本章的主张

本章的主张有两条，第二条是主要的。第一条：三种处置各自生产出一个新的对象，而这个新对象随后在该领域的实践中占据了原物的位置。这一点在三个领域中都可被具体指认，且在其中一个领域中已被该领域自己承认。第二条：一个领域采取哪一种处置，主要不由对象的性质决定，而由该领域下游使用的刚性需求决定。神学、礼仪、法律与翻译需要一个可被逐节引用的单一文本，因此新约文本批评必须造替身；文学研究不需要一个可被逐行引用的单一文本，甚至从版本差异中获益，因此莎士比亚编辑学可以并列实存；物理学的下游使用只需要预测测量结果，不需要"测量之前的样子"，因此量子力学可以取消问题。而这一选择随后被表述为一项关于对象本性的发现——"原初文本是一个多义的概念""在测量之前没有事实""每一个版本都是独立的作品"。本章称这一表述上的转换为本体论追认。

1.4 本章不主张什么

为避免误读，以下三点须先声明。其一，本章不主张这三个领域的做法是错误的。 相反，三种处置在各自的下游需求之下都是恰当的、甚至是唯一可行的。其二，本章不主张这些形而上学立场是虚假的。 一个由使用需求驱动的立场仍然可能是真的。本章所主张的是它的来源被误认，而非它的内容为假。其三，本章不裁决量子力学诠释之争。 该争论涉及本章作者不具备的专业判断。本章只使用一项关于该争论现状的可核查事实：分歧的稳定存在本身。

2　三个案例

2.1　新约文本批评：造替身，且已为替身另起了名字

这个领域的对象。 新约的二十七卷书没有一份作者手稿传世。现存最早的残片距成书数十年，完整抄本距成书数百年，而抄本之间存在大量差异。传统上，该领域的目标被表述为"复原原初文本"。替身是什么。 今天全世界学者所使用的希腊文新约（如内斯特勒-阿兰德本、联合圣经公会本），是一个折衷本：它在每一处异读中由编辑委员会择取被认为最佳的读法，而不问该读法出自哪一份抄本。由此得到一个值得停下来看的事实：这个文本作为一个整体，历史上没有任何一份抄本是这个样子。 没有任何一位抄经士抄过它，没有任何一个教会读过它，没有任何一位教父引用过它。它是十九世纪以来学术工作的产物，而它现在是全世界引用新约时的实际基准。这个领域的自我认识变化。 一九九九年，Eldon Jay Epp 在《哈佛神学评论》发表《新约文本批评中"原初文本"一词的多义性》，指出"原初"一词已"爆裂为一个复杂而高度难以驾驭的多义实体"——它至少包含前身文本、亲笔文本、正典文本与诠释文本等多重含义。此后，该领域的正式目标发生了改名。当代最重要的校勘工程所要确立的，不再被称为"原初文本"，而被称为"起始文本"。Gerd Mink 对这一概念的定义是：起始文本是一个假设性的、重构的文本，按照该假设，它是在抄写开始之前所存在的那个文本。 而 Mink 明确地把起始文本与两样东西区分开来：一是作者的文本，二是抄本传统的原型。这个改名的意义需要被说清楚。 该领域没有放弃构造替身——它仍然在构造，而且构造得更精密。它所做的是：给替身另起了一个名字，并在定义中写明这个名字不指作者所写的那个文本。D. C. Parker 更进一步，主张恢复单一原初文本的尝试是不可能的，编者只能满足于"现存抄本中的各读法在谱系上所由出的那个文本"。这个领域内部并非没有反对。 Holger Strutwolf 主张，寻求原初文本这件事本身并不包含矛盾或逻辑上的不可能。这一分歧至今未决。本案例对本章的意义：这是三个案例中唯一一个该领域自己承认了替身性质的案例。它没有假装重构物就是原物，它明确地把二者分开命名。这使它成为一个格外有价值的对照组——第 5.3 节将论证，正是这个诚实的改名，暴露了另外两个领域中未被承认的同一动作。

2.2　量子力学诠释：取消问题，而且取消得并不一致

这个领域的对象。 一个量子系统在被测量之前，其某一物理量是否具有确定的值？在经典物理的图景中，一切事物无论是否被观察都具有确定的属性。而按照海森堡与玻尔在上世纪二十年代发展出的立场，量子领域并非如此。分歧的规模。 二〇二五年七月，《自然》在量子力学诞生一百周年之际发布了一项调查，超过一千一百名研究者作答。结果被该刊表述为：物理学家对量子理论就实在说了什么，缺乏共识的程度令人瞩目。在偏好的诠释中，约百分之三十六选择哥本哈根诠释，约百分之十五选择多世界诠释，其余分散于关系诠释、知识论诠释、玻姆力学等。几项细节比总体比例更有诊断价值。其一，在选择哥本哈根诠释的人中，只有约百分之二十九认为波函数描述某种真实的东西；接近三分之二认为它只是编码信息或概率。也就是说，即使在多数派内部，关于"这个数学对象指什么"也没有共识。其二，选择哥本哈根诠释的人中，超过半数表示对自己的答案并不很有把握，并回避了要求他们展开说明的追问。有哲学家据此指出，该诠释的优势地位更多来自历史惯性而非经过反思的判断。其三，该调查中最接近共识的一项是：约百分之八十六的人同意，试图以物理的或直观的方式诠释量子力学的数学，是有价值的。也就是说，绝大多数人同意这件事值得做，而对做出来的结果无法达成任何一致。取消问题这一处置的表述。 该立场的一个极端表述来自 Anton Zeilinger 在同一场纪念活动上的发言：不存在一个量子世界；量子态只存在于他的头脑中，它们描述的是信息，而非实在。反对的立场。 爱因斯坦终生坚持的看法是：存在一个先于测量的实在，而科学的任务正是去测量它。多世界诠释则给出另一种回应：波函数不坍缩，而是分支为与全部可能结果一样多的世界，所有其他可能性在别的分支中同等真实。本案例对本章的意义：这是三个案例中"取消问题"这一处置最纯粹的形态，同时也暴露了该处置的一个内部困难——取消并不彻底。多数派选择了一个取消问题的诠释，而多数派内部对被取消的到底是什么并无共识，且过半数人对自己的选择没有把握。

2.3 莎士比亚编辑学：拒绝造替身，把版本并列

这个领域的对象。 莎士比亚的戏剧手稿无一存世。《哈姆雷特》有三个早期版本：一六〇三年的第一四开本、一六〇四至五年的第二四开本、以及一六二三年第一对开本中的版本。后两者都比第一四开本长出一千六百行以上。传统的处置：造替身。 二十世纪上半叶的主流做法，是以某一版本为底本，再从其他版本中择取被认为更接近作者意图的读法，编成一个折衷本。这一做法的理论基础由 W. W. Greg 与 Fredson Bowers 奠定，其目标被表述为复原作者的最终意图。在这一传统中，第一四开本被 Greg 与 J. Dover Wilson 等人判定为"坏四开本"——一个由演员凭记忆重构而成的败坏文本，因而在确立作者意图时价值有限。这里同样得到那个事实： 二十世纪绝大多数读者所读的《哈姆雷特》，是一个折衷本；而这个折衷本作为一个整体，在一六二三年之前从未存在过。转向：拒绝造替身。 一九八三年，Jerome

McGann 在《现代文本批评之批判》中对上述传统提出系统批评，主张文学文本的生产是社会性的、协作性的，作者的最终意图不应成为确立底本的唯一依据。此后数十年间，编辑实践中出现了一系列替代路径：关注文本的社会与协作维度、关注其历史文化维度、关注其多重性与不稳定的流动性质。这一转向的一个决定性实例：二〇〇六年， Ann Thompson 与 Neil Taylor 编纂的阿登版《哈姆雷特》明确决定不编折衷本。他们把三个版本各自视为独立的实体，分两卷完整印出。据记载，这一决定在评论界引起了不一的反应。他们同时承认：《哈姆雷特》的文本史充满疑问，而清晰的答案很少。本案例对本章的意义：这是三个案例中唯一一个该领域曾经造过替身、后来又停止造替身的案例。它因此提供了一个自然的前后对照——第 5.4 节将用它来检验本章的主要主张。

3　三家的正面冲突

三个领域的处置不是"侧重不同"，而是互相排斥。本节把冲突点逐一说清，因为后文全部建立在这些冲突之上。

3.1　造替身与取消问题的冲突

冲突点：替身的全部合法性来自它逼近一个存在但够不着的原物。若原物这个提法本身无所指，则替身逼近的是虚空，其构造原则（择取"更可能是原初的"读法）失去依据。这个冲突不能被"两个领域的对象不同"化解，因为文本批评自己已经承认原物在原理上不可及——起始文本被定义为假设性的重构，且明确不等于作者的文本。一旦承认这一点，"我们在逼近什么"这个问题就变得与量子力学中"测量前是什么"同构。 文本批评的回答是"有一个东西在那里"，量子力学多数派的回答是"没有那样一个东西"。

3.2　造替身与并列实存的冲突

冲突点：并列实存的做法蕴含一项指控——折衷本是一个从未存在过的人工物，以它为基准是持续的失真。而造替身的做法蕴含一项反指控——并列实存把裁决的责任推给了没有能力裁决的读者，且使一切精确引用成为不可能。这个冲突在实务上是不可调和的：一个需要逐节引用的共同体（神学、礼仪、法律、翻译）无法使用三卷并列的文本；而一个研究文本流动性的共同体无法使用一个抹平了差异的折衷本。

3.3　取消问题与并列实存的冲突

这一对冲突最不显眼，但最深。冲突点：并列实存主张三个版本各自实存——它们是真实的物质对象，都印出来了，都被人读过。而取消问题的立场主张，未被测量的那些可能性不具有实在性。多世界诠释在这一点上与并列实存立场一致（所有分支同等真实），而与哥本哈根

诠释冲突。因此第三对冲突实际上把量子力学内部的分歧引入了本章的三角：在"多个未被选定者是否实存"这个问题上，并列实存与多世界诠释站在一边，哥本哈根诠释站在另一边。

3.4 三家共同的盲区

三家都没有把自己的处置当作一种处置来看待。文本批评把造替身表述为方法论的进步。量子力学把取消问题表述为对实在的发现。 莎士比亚编辑学把并列实存表述为对文本本性的更准确认识。没有一家把"我们面对同一个缺口作出了不同的处置"这件事作为一个问题提出来，因为没有一家知道另外两家的存在。这正是跨领域比较的价值所在，也是本章得以成立的条件。

4 理论框架

4.1 核心概念

不可及原物（inaccessible original）：一个领域的对象所预设的、未被中介的初始状态，且该状态在原理上或在事实上无法被直接获取。缺口处置（gap disposition）：一个领域在没有原物的条件下继续工作时，对该缺口所作的制度性安排。本章识别出三种：处置一·替身构造：承认原物存在但不可及，据现有证据构造一个重构物，并以之为一切后续工作的实际起点。 处置二·问题取消：主张"原物是什么样"这一提问无所指，从而使缺口不再是缺口。 处置三·并列实存：拒绝构造任何重构物，将多个实际存在的中介物并置，不作裁决。追认物（retroactive artifact）：某一处置所生产出的、历史上从未存在过的新对象，该对象随后在实践中占据了原物的位置。下游使用的刚性需求（rigid downstream requirement）：某一领域的成果在被使用时所必须满足的形式条件，且该条件不可通过改变使用方式来规避。本体论追认（ontological ratification）：一项由使用需求驱动的处置选择，在表述中被转换为一项关于对象本性的发现。

4.2 命题系统

命题一（缺口的普遍性） 凡预设了不可及原物的领域，都必须作出缺口处置。不处置不是一个选项，因为工作需要起点。命题二（处置的三分） 缺口处置在逻辑上穷尽为三种：承认原物而造替身、取消原物而消解问题、放弃单一原物而并列实存。三者互斥。命题三（追认物的生成） 每一种处置都生产出一个历史上从未存在过的新对象：处置一生产折衷本，处置二生产被实体化的形式对象（如被当作实在的量子态），处置三生产"多版本并存"这一现代范畴——莎士比亚时代的读者不会认为存在三个《哈姆雷特》。命题四（追认物的位置占据） 追认物随后在该领域的实践中占据原物的位置，成为引用、教学与后续研究的实际起点。这一

占据不留痕迹，因为追认物的合法性正来自它宣称自己不是新东西。命题五（主要主张：处置由下游需求决定） 一个领域采取哪一种处置，主要不由对象的性质决定，而由该领域下游使用的刚性需求决定。命题六（本体论追认） 处置选择随后被表述为关于对象本性的发现，而非关于使用需求的安排。这一表述转换使处置不可被讨论，因为它已不再显示为一个选择。命题七（可检验的推论） 若命题五成立，则当一个领域的下游使用需求发生改变时，其形而上学立场应随之改变；且改变的时序应为需求在先、立场在后。

5 分析

5.1 三种处置各自生产了什么

本节论证命题三与命题四。处置一所生产的：折衷本。折衷本的性质需要被准确描述。它不是任何一份抄本，也不是若干抄本的平均，而是逐处独立择优的结果：在第一处异读取甲本，在第二处取乙本，在第三处取丙本。因此，即使每一处的选择都正确，整体也可能对应一个从未存在过的文本——因为不存在任何理由保证正确的读法恰好分布在同一份抄本中。这一点在数学上是明确的：若有 n 处异读，每处有若干候选，则折衷本是这些候选的一个组合；而实际存在过的抄本只占全部组合中的极小一部分。折衷本落在"从未存在过"这一区域中，是常态而非例外。同样的结构出现在莎士比亚的折衷本中。处置二所生产的：被实体化的形式对象。在量子力学中，波函数是一个数学对象。关于它是否描述某种真实的东西，即使在多数派内部也没有共识——如第 2.2 节所引，选择哥本哈根诠释的人中只有约三成认为它描述真实的东西。然而在教学、科普与大量实际讨论中，波函数被当作一个东西来谈论：它坍缩、它演化、它弥散。一个其实在性在专业内部并无共识的数学对象，在实践中获得了对象的地位。这是处置二所生产的追认物：不是原物，而是取代原物位置的那个形式对象。处置三所生产的："多版本"这一范畴。这一项最容易被忽略，因为它看起来只是承认既有事实。但"《哈姆雷特》有三个版本"这一表述预设了一个统摄它们的作品，而这个作品不等于其中任何一个。一六〇三年的读者不会认为自己读的是"《哈姆雷特》的第一四开本"——他读的就是《哈姆雷特》。"第一四开本"这个身份是后来的编辑传统赋予的，它只有在与另外两个版本的关系中才成立。因此并列实存并没有回到原物，它生产了一个新的对象：一个由三个版本共同界定、而不与任何一个版本重合的作品概念。三种处置的共同结构由此显现：没有一种处置回到了原物；三种都生产了新东西；而三种所生产的新东西随后都占据了原物的位置。

5.2 为什么这一占据不被察觉

命题四的后半句需要单独论证。其一，追认物的合法性来自它否认自己是新的。折衷本之所以能成为引用基准，正因为它宣称自己是对原初文本的最佳逼近。若它宣称自己是一个新文本，它将立刻失去全部权威。 因此在结构上，它必须否认自己的新颖性——这不是欺瞒，而是它作为追认物的存在条件。其二，追认物比原物更好用。折衷本比任何单一抄本都更"通顺"，因为每一处都取了最佳读法。被实体化的波函数比"没有事实"更容易教学与讨论。追认物在使用上的优越性，持续地掩盖着它的构造性。其三，最要紧的一层：使用追认物的人，绝大多数不参与它的构造。构造折衷本的是少数文本批评家，他们清楚地知道自己在做什么——第 2.1 节所述的改名正是这种清醒的产物。而使用它的是数以百万计的神学生、译者、牧师与读者，他们所接触到的是成品，不是构造过程。这一分工使得该领域内部的诚实，无法传递到该领域外部。 起始文本被定义为假设性重构这件事，写在专业文献里；而在使用现场，它就是"新约"。

5.3 主要主张：处置由下游需求决定

本节论证命题五，这是本章的核心。论证采取一个简单的形式：考察三个领域的下游使用有什么刚性需求，看这些需求是否唯一地决定了处置的选择。新约文本批评的下游需求：可被逐节引用的单一文本。神学论证需要引用具体经文并要求对手能核对同一处；礼仪需要一个可被诵读的定本；翻译需要一个可被逐句对应的源文本；在若干法域中，宗教文本还进入法律与教育的引用体系。这些需求都要求一个单一的、稳定的、可被逐节定位的文本。 并列实存无法满足它们：一个把三种读法并列而不裁决的文本，无法被引用，也无法被翻译。取消问题更不可行：无法主张"这段经文本来是什么样"这一提问无所指，因为整个诠释传统建立在这一提问之上。因此该领域只能采取处置一。而它确实采取了处置一。莎士比亚编辑学的下游需求：无刚性单一文本要求。文学研究不要求所有研究者引用同一个文本；相反，版本差异本身是研究对象。演出不要求忠于某个定本——每一次演出都是一次改编。教学可以处理多版本。因此该领域没有必须造替身的压力。而当理论上的反对（作者意图不应是唯一依据）出现时，它就转向了处置三。量子力学的下游需求：预测测量结果。物理学的下游使用是计算与工程：预测某次测量得到某结果的概率、设计器件、建造机器。这些使用完全不需要"测量之前的样子"这一信息。因此该领域可以承受处置二。 取消一个下游完全用不到的问题，代价为零。三项对照给出一个整齐的结果：三个领域的处置，与其下游刚性需求一一对应，且在每一个案例中，其他两种处置都会与该领域的下游需求冲突。这一对应是本章全部主张的经验基础。它是相关而非因果的证据——第 5.5 节将处理这一局限。

5.4 一次已经发生的检验：数字化转型

命题七给出了一个可检验的推论：若处置由下游需求决定，则需求改变时处置应随之改变，且时序为需求在先。新约文本批评提供了一次已经发生的检验。印刷时代的约束。 在纸本时代，一部希腊文新约必须在版面上呈现一个连续的正文；异读只能压缩进脚注的校勘栏。这一物质约束强化了对单一文本的需求：正文的位置只有一个，必须有东西填在那里。数字化解除了这一约束。 在数字版本中，可以同时呈现全部抄本、全部异读，可以让使用者自行选择呈现哪一层，可以让每一份抄本都以其自身的完整形态被阅读。"正文只有一个位置"这一约束消失了。而该领域对目标的重新表述——从"原初文本"到"起始文本"、以及关于原初一词多义性的讨论——恰好发生在这一转型的同一时期。本章认为这一时序符合命题七的预测，但必须立即加上三条限定，否则这一论证会被高估：其一，时间上的重合不等于因果。 该时期同时发生了多项变化，包括新抄本的发现、谱系方法的进展、以及关于早期基督教文本多样性的历史认识的深化。任一项都可能独立地推动了目标的重新表述。其二，因果方向可能相反。 也可能是先有了对原初文本概念的理论怀疑，才使得数字化的多文本呈现显得有价值。其三，这一检验只涉及一个案例。 单一案例不足以支撑命题五，它只能显示命题五在此处未被证伪。尽管有这三条限定，这一案例仍有价值，因为它给出了一个可以被继续追踪的检验点：若命题五成立，则随着数字呈现的进一步普及，该领域应当出现向处置三的进一步移动——即更多地并列呈现完整抄本、更少地依赖单一折衷正文。这是一个关于未来的、可被观察的预测，第 9.1 节将其写成正式的失效条件。

5.5 本章这一主张的最强反驳，及其处理

必须正面接住一个反驳，否则本章只是一个整齐的巧合。反驳：三个领域的下游需求不同，而它们的对象性质也不同；因此处置与需求的对应，可能是由对象性质经由需求间接造成的，而非需求直接决定。具体地说：量子力学之所以下游不需要"测量前的样子"，可能正因为那里确实没有那样一个东西；文本批评之所以下游需要单一文本，可能正因为那里确实曾经有过一个单一文本。若如此，则真正的决定者是对象性质，需求只是它的表现。这个反驳是有力的，而且本章无法完全排除它。 以下是本章能给出的三点回应，须逐一评估其力度。回应一（较弱）：莎士比亚的案例中，对象性质在整个时期内没有变化——三个四开本与对开本一直在那里，一份新的手稿也没有被发现——而处置发生了改变。这至少表明处置可以在对象性质不变的情况下改变。但这一回应力度有限：反驳方可以说，改变的不是对象，而是我们对对象的认识。回应二（中等）：若对象性质是决定者，则处理相似对象的不同领域应采取相似处置。而实际情况并非如此：同为古代文本的传世典籍，在不同的学术传统中被以不同的方式处理——有的编定本，有的存异文——而这些差异与各自的使用传统相关，与文本本身的传抄状况关系较弱。 这一回应需要一项本章未做的系统比较来支撑，此处只作为可检验的推测提出。回

应三（最强，但也最有限）：命题五的表述使用了"主要不由"而非"完全不由"。本章不主张对象性质无关，只主张在对象性质容许多种处置时——而这在三个案例中都成立——是下游需求作出了选择。这一表述使命题五较弱，但也使它更可能为真。本章接受这一代价，并在第 9.4 节把这一削弱写成一条失效条件：若能证明在任一案例中对象性质只容许一种处置，则该案例对命题五不构成支持。

6　与最近邻概念的划界

本章的概念若不能与既有概念划清界线，则不构成理论增量。以下逐一处理五个最近邻。

6.1　与"理论负载的观察"

科学哲学中一个成熟的论点是：观察不是中性的，它受观察者所持理论的塑造。共同点：两者都主张我们所接触到的对象不是未经中介的。分界点有二。其一，理论负载论处理的是观察行为如何被理论塑造；本章处理的是当对象在原理上不可观察时，一个领域用什么东西填补那个位置。前者关于看的方式，后者关于看不到时放什么。其二，也是更要紧的：理论负载论不涉及制度性的替代物构造。折衷本不是一次被理论塑造的观察，它是一个由委员会逐处表决产生的、被印刷出来的、被全世界引用的物质对象。它的存在方式与"被理论塑造的观察"完全不同。

6.2　与"操作性定义"

操作主义主张一个概念的意义由测量它的操作所给定。分界点：操作性定义是公开的、被承认的——它明说"我们把智力定义为智力测验所测的东西"。而本章所论的追认物必须否认自己的构造性才能成立（第 5.2 节）。一个公开承认自己是操作定义的东西，不会占据原物的位置；而追认物的全部功能正在于占据那个位置。新约文本批评的案例在这里格外有价值：当该领域把目标改名为"起始文本"并明说它是假设性重构时，它就是在把一个追认物转换为一个操作性定义。这一转换是罕见的，也正因为罕见而值得注意。

6.3　与"社会建构"

社会建构论主张若干被视为自然的范畴实为社会过程的产物。分界点有三。其一，社会建构论通常处理的是范畴（如疾病分类、性别范畴）；本章处理的是具体的物质对象（一个印出来的文本、一个数学形式）。其二，社会建构论常带有揭露性的意图——指出某物"其实"是被造出来的，因而可以是别的样子；本章明确主张三种处置在各自的下游需求下都是恰当的，且不主张它们可以随意更换。其三，也是最关键的：社会建构论不解释为什么是这一种建构而非那一种；本章的主要主张恰恰是给出这个"为什么"——下游使用的刚性需求。

6.4 与库恩的范式

分界点：范式转换发生在同一领域的不同时期，是一个共同体整体的更替。而本章所论的三种处置同时存在于三个不同领域，且各自内部是稳定的。三个领域并非处于不同的范式阶段，它们各自都处在成熟且稳定的状态中。本章所比较的不是先后，是并行。

6.5 与"退而求其次"的一般说法

最后一个近邻是常识性的：够不着最好的，就用次好的。分界点：常识的说法假定次好的东西在同一序列上，只是低一级——够不着原本，就用最接近原本的抄本。而本章所指出的是：三种处置所产生的东西都不在那个序列上。 折衷本不是"最接近原本的抄本"，它是一个不在抄本序列中的新项；"多版本"不是任何一个版本，它是一个新的统摄范畴。这一区别有实质后果：若追认物只是序列上低一级的东西，则改进方法就能使它更接近原物；而若它不在序列上，则方法的改进不会使它变成原物，只会使它成为一个更好的追认物。

7 可操作的判据

本节把命题六——本体论追认——转换为可以在具体场合使用的检查。判定一项形而上学立场是否为下游使用需求的追认，可依次问四个问题。

7.1 第一问：这个立场的相反面，在实务上是否可行？

若一个领域主张 P（如"不存在单一原初文本"），先问：若主张非 P，该领域的日常工作还能不能进行？若不能进行——即非 P 会使该领域的核心实务瘫痪——那么该立场高度可疑地是被需求所决定的。一个在实务上不可能被采纳的立场，其被拒绝可能与它是否为真无关。在新约文本批评中，"每一份抄本各自独立、不存在可引用的单一文本"这一主张会使翻译与礼仪无法进行。因此该领域拒绝这一主张，可能与该主张的真假无关。

7.2 第二问：这个立场是在什么时候被提出的？

追认的典型时序是：先有实务安排，后有理论表述。若一项形而上学立场的提出，晚于该领域相应实务做法的确立，且该立场恰好为既有做法提供辩护——则追认的可能性高。注意这条判据的方向限制：它只能提高怀疑，不能确立结论。理论表述晚于实务是科学中的常态，且常常是正当的——人们往往先做对了事，后来才说清楚为什么对。

7.3 第三问：这个立场在该领域内部与外部的表述是否一致？

追认物的一个稳定特征是内外表述不一致（第 5.2 节）：专业内部承认构造性，使用现场不承认。因此可问：该领域面向专业同行的表述，与面向使用者的表述，是否说的是同一件

事？新约文本批评在这一点上得分很高——它把改名写进了定义，并在专业文献中反复说明。而这一诚实并未传递到使用现场，这本身是一个可被观察的、值得研究的落差。

7.4 第四问：如果下游需求改变了，这个立场会不会跟着变？

这是四问中最直接的，也是唯一一个可以事前预测的。具体操作：设想该领域的下游使用需求发生某项具体变化（技术使某种呈现方式成为可能、使用者共同体的构成改变、某项制度要求被取消），然后问该立场是否还有维持的必要。若答案是"没有必要了"——则该立场很可能一直是被需求维持的。 若答案是"仍然必要"——则该立场可能有独立于需求的根据。第 5.4 节所述的数字化转型正是这一问的一次实际应用。

7.5 四问的合用与限度

四问不构成一个判定程序，只构成一组提高或降低怀疑的信号。四问全部指向追认，也不证明该立场为假——如第 1.4 节所声明，一个由需求驱动的立场仍可能为真。四问的实际用途是别的：它把一场看起来关于世界本性的讨论，翻译成一场关于"我们需要什么"的讨论。 而后一场讨论是可以有进展的，因为需求是可以被摆到桌面上比较的。

8 跨领域兑现

本节将上述框架展开于七个领域。每项均给出该框架带来的具体判断。

8.1 法律中的"立法原意"

缺口：立法者当时的意图不可及——立法者是一个集体，其成员意图各异，且多已不在世。处置：以法条文本与立法史材料构造一个"原意"。这是典型的替身构造。本框架的判断：该处置由下游刚性需求所迫——判决必须给出一个确定结论，不能并列多种解释而不裁决。因此关于"立法原意是否存在"的理论争论，其实际胜负早已由审判制度的形式要求所决定。 而这一点在争论中很少被指出，争论双方都把它当作一个关于意图本性的问题。

8.2 历史学中的"事件本身"

缺口：事件本身不可及，只有史料。处置：不同的史学传统采取了不同的处置——实证史学倾向替身构造（重建"实际发生了什么"），后现代史学的一部分倾向取消问题（不存在文本之外的事件），微观史与档案转向的一部分倾向并列实存（呈现多种互相矛盾的记述而不裁决）。本框架的判断：这三种立场与三种下游使用一一对应——教科书与公共史需要单一叙述（替身），理论研究不需要（取消问题），而档案与地方史研究从矛盾中获益（并列）。这一对应是本章框架的一个待检验推论。

8.3 临床医学中的"真实病因"

缺口：许多疾病的病因不可直接观察。处置：诊断分类构造了一个替身——一个疾病实体，它在实践中占据了病因的位置。本框架的判断：该处置由处方权、医保支付与病历书写的刚性需求所迫——这些制度都要求一个可被编码的单一诊断。因此关于"某疾病是不是一个真实实体"的争论，其实际约束条件是编码系统的形式要求。 值得注意的是，当支付制度允许多重诊断编码时，该疾病的"实体性"在实践中随即下降。

8.4 机器学习中的"真实标签"

缺口：许多任务不存在无争议的正确答案，而训练需要标签。处置：以多位标注者的一致或多数构造一个标签。这是替身构造，且构造过程通常被记录得很清楚。本框架的判断：这是一个格外清晰的案例，因为构造过程是公开的、可量化的——标注者间一致性被直接报告。然而在模型被使用时，这些标签仍被当作"真实答案"。这提供了一个可以直接研究追认过程的现代样本：从公开的构造，到被当作事实，中间的每一步都留有记录。

8.5 经济统计中的总量指标

缺口：一个经济体的"总产出"不是一个可以被直接观察的量。处置：以一套记账规则构造一个数——这是替身构造，且构造规则是公开发布的。本框架的判断：该处置由政策与国际比较的刚性需求所迫。而关于"这个数是否测量了某种真实的东西"的长期争论，按本章框架，其实际约束是：任何取消该数的主张，都会使一整套制度失去运转依据，因而不可能被采纳，无论它是否正确。

8.6 考试与评价

缺口：一个人的"实际能力"不可直接观察。处置：以测验分数构造一个替身。本框架的判断：分数之所以能占据能力的位置，正因为选拔制度需要一个可排序的单一量。并列实存在这里几乎不可能——一份并列呈现某人多方面表现而不给出总分的评价材料，无法被用于排序。因此关于"分数能不能代表能力"的争论，在选拔制度存在的前提下，实际上没有可选项。

8.7 人工智能系统中的"意图"与"能力"

缺口：一个模型的能力或倾向不可被直接观察，只能被测量。处置：以基准测试构造一个替身——一组分数，随后被当作该模型能力的描述。本框架的判断：该领域正处在一个罕见的位置——三种处置同时在被使用。 基准分数是替身构造；关于"模型有没有真正的理解"这一提问被一部分人取消；而另一部分人主张并列呈现模型在多种条件下的不同表现而不综合为单一评分。按本章框架，最终胜出的将不是最准确的那一种，而是与下游刚性需求最匹配的那一

种。 而下游需求正在成形：监管需要可比较的单一评级，采购需要可排序的分数。这两项需求都指向替身构造。 因此本章预测该领域将向替身构造收敛——这是一个可被观察的、写在此处的预测。

9　可使本章失效的条件

一项主张若不能被推翻，则本章对追认的全部诊断同样适用于本章自身。以下四条被明确写出。

9.1　条件一：数字化未带来向并列实存的移动

第 5.4 节把新约文本批评的数字化转型当作命题五的一次检验，并给出了一个关于未来的预测：随着数字呈现的普及，该领域应出现向处置三的进一步移动。这一预测可以被观察。具体的观察指标是：主要校勘工程的产出中，完整抄本的独立呈现与单一折衷正文的相对权重；以及在教学与翻译实践中，是否出现以多抄本并列为基础的工作方式。若在数字呈现已充分普及之后，该领域仍完全依赖单一折衷正文，且未出现向并列的任何移动——则命题五在这一案例上被证伪，而这是本章唯一有实际时序数据的案例。需排除的假成立：制度惯性可能延迟移动而不否定命题五。因此该观察需要足够长的时段，且需同时考察是否有新的刚性需求出现（例如某项要求单一文本的新制度安排）。

9.2　条件二：存在第四种处置

命题二主张缺口处置穷尽为三种。若能找到一种处置，它既不构造替身、不取消问题、也不并列实存，而仍使该领域得以在没有原物的条件下工作——则本章的三分是不完全的，框架需要被替换。一个值得注意的候选是：放弃该对象、改研究别的东西。本章认为这不构成第四种处置，因为它不是对缺口的处置，而是对缺口的规避。但这一判断可以被反对，且本章没有强论证支持它。

9.3　条件三：处置与下游需求不对应

命题五的经验基础是第 5.3 节那个一一对应。这一对应可以被系统检验：取足够多的、存在不可及原物的领域，独立编码其处置类型与下游刚性需求类型，看二者是否显著相关。若不相关，或相关性可被对象性质完全解释——则命题五失败。这一检验是本章最希望被人做的，因为本章只有三个案例，而三个案例可以拟合任何整齐的框架。三个点画出一条直线，这件事没有任何统计意义。 本章对此完全清楚，第 10.1 节将其列为首要局限。

9.4　条件四，也是本章自己最担心的：对象性质只容许一种处置

第 5.5 节已经承认了这一条。此处把它写成正式的失效条件。若能证明在三个案例中的任何一个，对象性质本身只容许一种处置——即另外两种处置在该对象上根本不可行，而与下游需求无关——则该案例对命题五不构成支持。若三个案例都如此，则本章的主要主张完全失败，真正的决定者是对象性质，而本章所观察到的"需求对应"只是对象性质的一个副产品。本章倾向于认为至少莎士比亚案例不属此列——因为对象性质在该案例中未变而处置改变了。但这一回应的力度有限（第 5.5 节回应一），本章承认它不足以支撑整个主张。

10　局限与结论

10.1　局限

其一，案例数为三，这是本章最严重的局限。三个案例足以显示一种可能的结构，不足以支持任何关于普遍性的主张。本章的全部经验基础是一个三点的对应，而这样的对应可以是巧合。第 9.3 节所述的系统检验是本章成为一项主张而非一个观察所必需的，而本章没有做这项检验。其二，"下游刚性需求"这一概念缺乏独立的度量。本章判断某项需求是刚性的，依据是"若不满足则该领域实务瘫痪"，而这个判断本身带有事后的性质——已经被满足的需求容易显得刚性。这是本章概念上最软的一处。其三，本章可能系统性地偏向那些有明确文本记录的领域。三个案例都是有大量公开文献、争论被高度记录的领域。处置发生在别的地方——在没有记录的实务传统里——的情形，在本章的取材范围之外。其四，本章未处理处置的混合形态。第 8.7 节已指出人工智能领域三种处置同时存在。本章的三分把它们当作互斥的，而现实中它们可以在同一领域内并存于不同层面。这一点本章没有给出处理。

10.2　结论

本章的主张可凝缩为四句。第一，凡预设了不可及原物的领域，都必须作出缺口处置；处置有三种，且互斥：造替身、取消问题、并列实存。第二，三种处置各自生产出一个历史上从未存在过的新对象，而该对象随后占据原物的位置；这一占据不被察觉，因为追认物的合法性正来自它否认自己是新的。第三，一个领域采取哪种处置，主要不由对象的性质决定，而由下游使用的刚性需求决定。第四，这一选择随后被表述为关于对象本性的发现，而非关于使用安排的选择；本章称之为本体论追认。这四句话不裁决任何一场具体争论。它们只提供一个在争论进行当中可以使用的位置：当一场关于"某物本来是什么样"的争论长期不决时，可以先问一句——我们这一行的工作，能不能在另外那个答案之下继续进行？如果不能，那么这场争论的结果，可能在它开始之前就已经被定下了。

10.3　本章的自我审查

本章诊断的结构，本章自己具备，而且具备得比前两条更彻底。其一，本章自己就是一个替身构造。本章所研究的对象是"三个领域各自的实际处置"。而这个对象不可及——本章接触到的不是那些领域的实际做法，而是它们的公开文本表述。本章据这些文本构造了三个"处置类型"，而这三个类型与任何一个领域的实际做法都不完全重合。按本章自己的框架，这三个类型是追认物：它们是为逼近实际做法而构造的，而它们随后在本章中占据了实际做法的位置。本章全部的论证，是在这三个追认物之间进行的。其二，本章的处置由本章的下游需求决定。本章需要一个可以被写成论文、可以被讨论、可以被检验的结构。这个需求要求三个类型互斥且穷尽——一个模糊的、互相渗透的、随案例而变的分类无法支撑一篇论文。因此第 10.1 节第四项所承认的那个局限——现实中三种处置可以并存——不是一个疏漏，而是本章自己的下游需求所造成的取舍。 本章为了可被论证而牺牲了这一点，这正是命题五所描述的那件事。其三，本章无法自我应用第 7 节的判据。第 7.4 问是四问中最有力的：若下游需求改变，该立场会不会跟着变？而本章的下游需求不会改变——只要本章以论文的形式存在，它就需要一个可被论证的结构。因此该问对本章不产生任何约束。本章能做的只有一件：把这三点写在这里，并把第 9 节那四条留在不能被事后撤回的位置上。其中第三条是本章最希望被人做而自己没有做的那一项——三个案例支撑不起一条关于普遍性的主张，本章对此没有任何辩解。

注释

〔1〕 本章使用"原物"一词统称三个领域中性质不同的对象：作者所写的文本、测量之前的物理事实、作者的手稿。这三者的不可及性在类型上不同——前两者中，第一种是历史的丧失，第二种可能是原理上的不存在。本章的框架不要求这三种不可及性同类，只要求它们在实践上造成同一个局面：工作必须在没有原物的条件下进行。 这一处的宽泛是有意的，也是可被批评的。〔2〕 第 5.1 节关于折衷本"落在从未存在过的区域中是常态"的论证，其严格性依赖于一个未经检验的假设：各处异读的正确读法在抄本之间的分布近似独立。若正确读法高度集中于少数抄本，则该论证不成立。 本章未对此作检验，此处标明。〔3〕 第 2.2 节所引的调查数据，其性质是研究者的自我报告偏好，不是关于量子力学的证据。本章使用它只是为了确立一项事实——分歧稳定存在且未收敛——而不用于支持任何一种诠释。〔4〕 本章三个案例中的表述，均取自各领域自身的公开文献或权威报道，未采用某一方对另一方立场的转述。〔5〕 第 8 节七项兑现中，仅第 8.7 项包含一个关于未来的明确预测（人工智能能力评估将向替身构造收敛）。其余六项是框架的应用，不是预测，不应被当作已验证的结论。

附录　本章的产生过程

按学科通融专栏体例，此处交代本章的产生方式与作废记录。三个来源，分属三个互不相干的领域：- 圣经神学：新约文本批评关于"原初文本"的争论——Epp 一九九九年的多义性论文、目标向"起始文本"的改名、Parker 关于不可能性的主张与 Strutwolf 的反对。- 物理学：量子力学诠释之争——《自然》二〇二五年七月的百年调查，逾一千一百名研究者作答，分歧稳定且未收敛。- 文学：莎士比亚编辑学——Greg-Bowers 的作者意图传统、McGann 一九八三年的批评、以及阿登第三版《哈姆雷特》拒绝编折衷本、三版分列的决定。

选择这三个的理由是它们在同一个问题上给出了互相排斥的答案：一个够不着的原物，该怎么办。而三个领域互不知晓对方的存在。九条判断，三家各抽三条最承重的。圣经神学那三条：折衷本作为整体历史上从未存在过；"原初"一词已爆裂为多义实体；目标被正式改名为"起始文本"且定义为假设性重构、明确不等于作者的文本。物理学那三条：多数派选择了一个取消问题的诠释；而多数派内部对波函数是否描述真实之物并无共识（仅约三成认为是）；最接近共识的一项是"诠释这件事有价值"，而非任何一种诠释。文学那三条：折衷本以复原作者最终意图为目标；该目标自一九八〇年代起被系统质疑；阿登第三版决定不编折衷本、把三个版本各自作为独立实体印出。三十六次对撞，作废十条。 其中七条为同义重复，三条撞不起来。作废的这十条不出现在成品里，在任何进度表上均记为零。而留下的二十六条之所以立得住，正因为那十条被剔除。留下的二十六条聚成四条暗流：一、三个领域面对的是同一个缺口，而处置分属三种且互斥。 二、每一种处置都生产出一个新对象，且该对象随后占据原物的位置。三、这一占据不被察觉，因为追认物必须否认自己的新颖性。 四、处置与下游刚性需求一一对应。第四条是根：它解释了为什么是这三种处置分配到这三个领域（而非任意组合），也因此把前三条从一组观察变成一条主张。收敛成一条判断：我们说的"本来的样子"，是我们为了继续工作而造出来的；而造哪一种，由我们打算拿它做什么决定，不由它本身决定。这条判断三家里谁也说不出来，原因就在各自的位置上。圣经神学那一家离答案最近——它已经诚实地给替身另起了名字——但它把这当作本领域方法论的精细化，没有把"改名"这个动作本身对象化。物理学那一家的分歧规模最大，但该领域把分歧归因于量子世界的特殊性，而非归因于一类处置问题。文学那一家唯一经历过处置的转变，因而拥有识别这一结构所需的最关键材料——同一个对象、不同的处置——但该转变被表述为对文本本性认识的进步，而不是一次处置的更换。若量子力学那一路正确（原物这个提法无所指），则文本批评构造替身的整个事业建立在一个不存在的目标之上；若文本批评那一路正确（原物存在但够不着，故须造替身以供使用），则并列实存是失职；若并列实存那一路正确（多个版本各自实存、不应被折衷），则替身是一个从未存在过的赝品，而以它为一切引用的基准是一个持续的错误。 三家均为站外的公开文献，链接直达原始出处，可自行核对。

第三章

哪一份算数

论危险不在副本的份数，而在预期一致却无人负责

学科入口　软件工程 × 数字保存与条约法 × 演化遗传学

一、同一件事，三个相反的处方

一样东西存了好几份。这是一件极普通的事，普通到很少有人把它当成一个需要解释的对象。但只要把三个行当的判决摆在一起，它立刻变得很不普通。一位软件工程师会说：把它们合掉。 这个行当有一条流传了二十多年的原则，原句是：每一条知识，在一个系统之内必须有单一的、无歧义的、权威的表示。[1] 理由不是省地方——地方从来不值钱——而是改一处会漏掉另一处。一个人修好了这一份里的错，另一份还在，然后在某个谁也没预料到的场合，那个旧的错跳出来。一位数字档案员会说：再多存几份。 这个行当有一句同样流传很广的口号，意思是"多存几份，东西才丢不了"。[2] 更要紧的是它的做法：这些副本之间不设正本，节点之间不断互相比对，谁偏离了就按多数把它修回去。它拒绝回答"哪一份是原件"这个问题，而这不是疏忽，是设计。一位国际法律师会说：同一份文件本来就有好几份，而且每一份都算数。 多语种条约的通例是各语言文本同等作准，出现歧义时不去找一个母本，而是按条约的目的与宗旨解释。[3] 这套办法运转了很久，没有人认为它是一种妥协。一位演化遗传学家会说：如果没有多出来的那一份，什么新东西也不会有。 一个基因被复制之后，原来那份继续干原来的活，多出来的那份才承受得起突变——因为它坏了不要紧。[4] 这个行当把绝大部分生物学上的创新记在这笔账上，而它成立的前提，恰恰是没有任何机制要求两份保持一致。四句话，三种处方，互相拆台：消除、增加、放任。而且没有一家是外行——每一家都有几十年的证据。还要补一句，四家里没有一家是在讲同一个层面上的事——这一点在第五节会成为关键。软件工程讲的是修改时会不会漏，数字保存讲的是丢失时会不会灭绝，国际法讲的是争议时按谁，演化遗传学讲的是新东西从哪来。四个问题，四条时间线：修改是日常的，丢失是罕见的，争议是偶发的，创新是极长期的。它们对同一个事实给出相反处方，一部分原因就在这里——每一家都在为自己那条时间线上最要紧的那件事做最优安排，而这四条时间线的最优安排互不相

容。这就是本章的起点。同一个事实，三个成熟的行当给出不能并存的判决，说明大家用的不是同一个变量。

二、三家各自说到了哪一步

先公允地把三家各自的强处摆出来，再说它们在哪里停下。

2.1 消除派：危害在"改一处会漏"

这一派的表述比通常的理解精确得多。它说的不是"不要复制粘贴"，而是"每一条知识必须有单一、无歧义、权威的表示"。差别很关键：两段代码长得一模一样，如果它们表达的是两件不同的事，那不算违反；两段代码长得完全不同，如果它们表达的是同一条规则，那才是。[1] 这一派的强处在于它抓住了一个真实的失败模式，而且这个模式可以被反复观察：一条规则改了，改的人只知道其中一处，于是系统进入一个不自洽的状态，而这个状态在被同时用到之前不会报错。还有一处强处值得单列：这一派把"知识"与"表示"分开了。这在当时是一个很不容易的动作——它意味着查重不能靠比对字面，得先判断两处说的是不是同一件事。而这个判断没有算法，只能由人做。这一点在正文后面会反复用到：这一派最深的洞见，恰恰是它自己无法自动化的那一部分。它的止步处也很清楚：它把处方定死在"减少份数"上。 它无法安置那些份数极多、却极其可靠的系统——数字保存的多副本网络、多语种条约、每一个分叉出去就不再合并的开源项目。在它的框架里，这些要么是设计错误，要么必须被重新描述成"其实只有一份"。

2.2 增加派：可靠来自互相校验

这一派的强处在于它把可靠性从"保住一份好的"改成了"让多份互相看着"。它的关键动作不是复制，是校验：节点定期比对，发现分歧就按多数修复。所以它的多，是有条件的多。而它最值得注意的一步，是主动取消正本。多语种条约不设一个母本，是因为一旦设了，其余文本就降格为译本，缔约方的谈判成果会被悄悄地重新分配。取消正本不是无奈，是为了让权威落在程序上，而不落在某一份上。这一派还有一个常被忽略的细节：它的副本不是静态存着的，是主动巡检的。没有巡检的多副本，与一堆散落的备份没有区别；口号里那句"多存几份"，隐含的完整版本是"多存几份并且不断互相比对"。本章后面把这一步单独称作机制，正是从这里来的。它的止步处在于：它假设一致是目标。 凡是不以一致为目标的场合，它给不出任何指导；而下一节那一家整个建立在"不一致才有用"上。

2.3 放任派：不一致才是产出

一个基因被复制之后，两份在很长一段时间里做同一件事。这看起来是纯粹的浪费，但它做了一件别的办法做不到的事：它把约束从其中一份上撤走了。 原来那份继续供着必需的功能，多出来的那份可以随便变——变坏了，个体照样活；变出点别的，就是新东西。[4]后来的研究给这条加了一个更常见的版本：两份各自丢掉原来那个基因的不同部分，于是谁也不能少，两份加起来才等于原来一份。[5] 这条路径不需要任何"新功能"，仅靠退化就能把两份都保住。这一派还有一个与前两家形成尖锐对照的性质：它不区分"故意的不一致"与"意外的不一致"。 一个突变是随机的，没有人打算让两份分开，也没有人打算让它们保持一致——不一致既不是目标也不是事故，它就是默认状态。前两家的整个语汇里都没有"默认状态"这一档：在它们那里，不一致要么是被追求的，要么是被容忍的，要么是被消灭的，总之总有一个态度。这一派的止步处在于：它的世界里没有"预期"这个东西。 基因组里没有任何人期待两份一致，所以它从来不需要处理"以为该一样、其实不一样"这种情形——而那恰恰是前两家吵得最凶的地方。

三、三家在哪里互相取消

把三条处方并排，冲突不是笼统的"观点不同"，而是三处具体的对顶。第一处：份数与安全的方向相反。 消除派说份数越少越安全，增加派说份数越多越安全。这看起来像同一根轴上的正面对顶，但两家说的"安全"根本不是一件事：消除派的安全是改的时候不出错，增加派的安全是丢的时候不灭绝。两种安全都要，而按任一家的处方做到极致，另一种必然崩掉——只留一份，一场火灾就没了；存一千份，改一条规则要改一千次。第二处：正本该不该有，两家给了相反的答案，理由却在同一个位置。 消除派要一个权威表示，是为了让修改有唯一的落点。增加派要取消正本，也是为了让修改有唯一的落点——只不过那个落点是程序，不是某一份。两家争的其实不是"要不要权威"，而是权威该落在位置上还是落在程序上，而两家都没有把这句话说出来。第三处：放任派把前两家的共同前提整个掀掉了。 前两家都默认副本应当一致，分歧只在怎么做到。放任派提供了一大批极其成功的反例：那里没有人期待一致，而不一致正是产出。如果一致不是必需的，那么前两家的整场争论就只覆盖了一部分场合，而它们各自都以为自己覆盖了全部。三处合起来指向同一件事：三家都在处理"副本"，而真正决定结局的变量不在副本上。

四、把"预期"从"机制"里分出来

先分开两个平时总是一起出现、因而很少被分开的东西。预期：这套系统里的人是不是认为，这几份应当是一样的。注意它是一个关于人的事实，不是关于副本的事实。它可以被明确写下来，也可以从来没有被说出口而人人都这么以为。机制：有没有一个被指定的东西——一

个人、一道流程、一段程序——负责让它们保持一样，并且在它们不一样的时候给出裁决。注意"被指定"三个字：一个碰巧总在做这件事的老员工，不算被指定；他哪天不做了，没有任何事情会报警。这两件事在日常里几乎总是一起被谈论，所以很容易以为有预期就有机制、没机制就没预期。恰恰相反：最麻烦的情形正是预期在而机制不在。看一个最小的例子。一家公司有一份报销标准，总部写了一版，三个分公司各自存了一份、也各自改过几处。- 所有人都认为这三份应该是一样的——预期在；- 没有任何人的职责里写着"核对这三份"——机制不在；- 于是三份慢慢分开，而这件事在有人同时用到两份之前不会暴露。

暴露的那一天通常是一次跨分公司的调动，或者一次审计。在那之前，每一份单独看都完好，每一次报销都顺利通过，账面上没有任何异常。这个情形有一个特征，值得写死：它没有报警条件。 消除派的那条原则能诊断它（三份即违反），但按那条原则去做，处方是"合成一份"，而现实里合不掉——三个分公司真的有不同的报销场景。增加派的做法能救它（互相比对、按多数修复），但没有人会想到把这套办法用在三份 Word 文档上。放任派会说这挺好（各地演化出适合自己的版本），而这句话在报销这件事上是错的，因为公司确实需要一致。三家都够不着，是因为三家都在看副本，而这里出问题的是预期与机制之间的那道缝。由此得到本章的承重判断：

> 副本本身既不危险也不安全。危险只发生在一个地方：系统预期这些副本一致，却没有任何人负责让它们一致。危害的量不由副本的数量决定，由这个预期覆盖的范围决定。

为行文方便，下文把"预期在而机制不在"这种情形叫作影子正本——所有人心里都有一份"算数的那一份"，但谁也说不出它在哪儿。

五、三家为什么都看不见这一格

上一节那个报销标准的例子里，三家各自能说出一点，却都到不了那条判断。值得逐家看清它们卡在哪儿——因为卡住它们的不是疏忽，是各自方法里一个必要的设定。消除派看不见它，是因为它的诊断单位是"份数"。 那条原则一旦发现三份，就已经判完了：违规。它不需要再问预期，也不需要再问机制——三份本身就是问题。于是它对格Ⅱ与格Ⅲ给出同一个诊断，而这两格的实际后果一个是灾难、一个是健康。一把只量份数的尺子，量不出"份数相同而后果相反"这件事。更要紧的是它的处方在格Ⅱ里往往执行不了。三个分公司的报销场景真的不同，合成一份要么牺牲实际需要，要么把差异塞进一堆例外条款——而例外条款本身又是新的重复。执行不了的处方，在实践中会退化成不执行，于是格Ⅱ照旧。增加派看不见它，是因为它的动作是"加副本"。 它的整套办法建立在一个前提上：副本越多，互相校验的能力越强。而格

Ⅱ里加副本只会更糟——多一份，就多一个没人核对的分支。它的药在这一格里是毒。它也不会想到把自己那套办法用在三份文档上，因为它的场景是机器与机器之间，而格Ⅱ最常发生在人与人之间。这不是它的错，是它的适用域没有被写清楚。放任派看不见它，是因为它的世界里没有"预期"这个东西。 基因组里没有任何主体期待两份一致，所以"以为该一样、其实不一样"这种情形在那里根本不成立。它的全部工具——约束的释放、子功能的分割、退化与保留——都是关于副本本身的，没有一样是关于看着副本的人的。于是出现一个有意思的局面：格Ⅱ的存在，恰恰需要三家的东西合起来才能被看见。 消除派提供了"预期一致"这一维——它是唯一把"权威表示"写进定义的一家。放任派提供了一批"不预期一致反而更好"的成功案例，否则很容易以为一致总归是对的。增加派提供了"预期一致却可以没有正本"这个反例，否则很容易以为预期一致就必须指定一份。三样并排，唯一的出路就是把"预期"与"机制"拆成两条互不蕴含的轴。少任何一家，这两条轴都合得回去：只有消除派与增加派，会以为争的是要不要正本；只有消除派与放任派，会以为争的是要不要一致；只有增加派与放任派，会以为争的是要不要校验。

六、两条轴与四个格

把上面两件事各取两值，得到四格。

	有被指定的同步机制	没有被指定的同步机制
系统预期这些副本一致	格Ⅰ · 有主同步 权威可以落在某一份上（主从复制、总部下发＋回执），也可以落在程序上（多副本互相比对按多数修复、多语种条约同等作准）。两种都健康，因为"不一致时怎么办"有答案。	格Ⅱ · 影子正本 人人默认应当一致，无人负责，且说不出不一致时哪份算数。每一份单独看都完好，只在两份被同时用到时才暴露。
系统不预期它们一致	格Ⅲ · 受控分岔 明确宣布不再合并，并有人管着分岔的边界（分叉后各自发布、方言的标准化机构、复制后两份各担一半功能而都被保住）。	格Ⅳ · 自由分化 没人期待一致，也没人管。产出方差极大：多数副本退化、消失，少数变成新东西。要求一个能承受大量死亡的容量。

四格的读数，落到能问、能查的东西上：格Ⅰ：随便找两个人，各问一句"不一致的时候按哪份"，两人的答案相同，并且都能说出谁负责。格Ⅱ：问同样的话，答案分歧——有人说按总部的，有人说按最新的，有人说按用得最多的那份，还有人愣一下说"应该都一样吧"。最后这

句是最强的读数。格Ⅲ：答案是"不需要一样"，并且能说出从哪一天起、按什么规矩分开的。格Ⅳ：答案也是"不需要一样"，但说不出规矩，也没人管死活。一分钟判据：拿着两份实物，问在场每个人同一句话——

如果这两份不一样，哪一份算数？

这句话的好处是它不需要先认定任何人有过错，也不需要读懂内容。答案的分歧程度就是读数本身。用一件小事走一遍。 一个团队里有三份"上线检查清单"：一份在共享文档里，一份印在墙上，一份在某人的笔记里。问第一句："这三份应该一样吗？"三个人都说应该。预期在。问第二句："如果不一样，按哪份？"第一个人说按共享文档，第二个人说按墙上那份（因为那是当初一起定的），第三个人愣了一下说"应该都一样吧"。答案分歧，机制不在。于是可以直接判：格Ⅱ。接下来只需要再问一句就能估出损害的量级——这三份分别是什么时候最后被改过的？ 时间差越大，已经积累的分歧越多，而这份分歧在下一次有人同时用到两份之前不会显形。这个例子的价值在于它足够小，小到可以在一次站会上真的做完。第十七节那条最便宜的判据，用的就是这个形状。后面几节按格展开，主戏在格Ⅱ。

七、格Ⅰ：权威可以不落在任何一份上

先立格Ⅰ，因为它给出后面三格的对照，而且它内部有一个被普遍忽略的分岔。格Ⅰ是预期一致、且有人负责。它有两种实现，外观差得很远。位置式：指定某一份为准，其余的向它看齐。总部下发文件、主库与从库、一份签署的正本加若干份复印件，都是这一种。它的长处是简单：任何争议一句话就能了结。它的短处是那一份出事就全出事，而且它把所有偏离都定义成违规——这一点在第十六节会变成一个具体的麻烦。程序式：不指定任何一份，用一道程序给出权威。多副本网络按多数修复是这一种；多语种条约各文本同等作准、歧义时按条约目的解释也是这一种。[2][3]程序式是三家争论里被漏掉最多的一格，值得多说一句。它证明了一件反直觉的事："预期一致"并不蕴含"必须有一个正本"。 消除派把这两件事绑在一起了——它要"单一、无歧义、权威的表示"，而那句话在字面上就把权威放在了一个位置上。多语种条约每天都在做的事，按那条原则读，是明确违反的；而它极其可靠。格Ⅰ内部还有一件事值得说清楚，因为它决定了这一格好不好维持。位置式的成本集中在平时，程序式的成本集中在建时。 指定一份为准，建起来只要一句话，但此后每一次修改都要推送、要回执、要抽查，成本按次数累加。用程序定权威，建起来很贵——要定清楚什么算多数、歧义时按什么解释、谁有资格投票——但建好之后，边际成本接近于零。这解释了一个常见的选择偏差：组织几乎总是选位置式，因为决策当天它更便宜。 而那笔按次数累加的成本，会在此后的每一次修改里被悄悄地跳过一点点——跳过得多了，就滑进了第八节那一格。位置式不是错的，它只是有一条通

往格Ⅱ的滑坡，而程序式没有。这也是本章第一次与消除派分开的地方：问题从来不是有几份，是"不一致时怎么办"有没有答案。 答案可以是"按主库"，也可以是"按多数"，甚至可以是"按当初的目的重新解释"——三种都算答案。没有答案，才是问题。

八、格Ⅱ：影子正本

现在展开主戏。格Ⅱ的定义只有两句：所有人都认为这些应当一致；没有任何人被指定去保证这件事。它之所以是四格里最坏的一个，不是因为不一致的程度最深——格Ⅳ的分化程度深得多——而是因为它有三个性质，凑在一起使它几乎不可能被及时发现。第一，它没有单点读数。 每一份单独看都是完好的。你可以逐份审查，逐份都合格。不一致这件事不存在于任何一份之中，它只存在于两份之间；而绝大多数检查是按份做的。第二，它的暴露时点由使用决定，不由损害决定。 三份报销标准分开了三年，这三年里损害一直在积累（有人多报了，有人少报了），但暴露发生在第一次有人同时用到两份的那一天。损害的发生与暴露之间没有固定的时间关系，因此从暴露频率完全推不出损害规模。第三，也是最要紧的：所有人都在按那个不存在的正本行事。 每个人心里都有一份"应该算数的那一份"，只不过每个人心里的那一份不是同一份。这就是"影子"二字的来历——它有全部的效力，没有任何的位置。这第三点带来一个很具体的后果。当不一致终于暴露时，处理它的通常不是规则，是嗓门。 因为没有事先约定的裁决办法，冲突的解决落回到人际权力上：谁的部门更强势，谁的版本更"看起来正式"，谁在场。而这个解决过程不会被记录成"我们缺一个裁决机制"，它会被记录成"某某部门那次搞错了"。于是每一次暴露都被消化成一次个人失误，而结构原封不动。举三个不同尺度的例子，形状完全一样。一份被各科室改编过的临床指南。 总院发了一版，各科依自己的病人构成做了调整，都合理。所有人都认为"我们用的是同一份指南"。不一致在院内几乎不暴露——因为每个科用自己那份是自洽的。它在跨科转诊时暴露：同一个病人在两个科被按两套阈值处理，而两个科的医生都确信自己遵循的是指南。一段被复制进五个项目的合同条款。 起草时是同一段，五年里各自被改过。所有人都认为这是"我们公司的标准条款"。不一致在平时不暴露，因为一次只用一份。它在争议时暴露，而那时的裁决落回到"哪一份签在合同上"——也就是说，落回到位置式，只不过是事后的、随机的位置式。同一门课的课件被四位教师各自维护。 都认为是 "这门课的课件"。不一致在期末统考时暴露，而暴露的形式是学生成绩差异，于是被归因为教师水平差异。第四个例子形状相同，但它是新的，而且正在快速变多。同一条业务规则，被四个人各自用同一套工具生成了一份说明。 四份的措辞完全不同——一份写成流程图，一份写成条款，一份写成问答，一份写成代码注释。四个人都认为自己写的是"那条规则"，也都认为四份说的是同一件事。这一例把格Ⅱ的隐蔽性推到了新的位置。前三例里，至少还能看出这些是副本——它们长得像。这一例里副本的外观已经不同了，因而任何按字面查重的办法都发现不

了；而按第二节那条区分，它们完全构成重复，因为它们表达的是同一条知识。于是这一例同时满足三件事：预期一致、无人负责、且连"这是几份"都数不出来。第十二节会说明为什么这一类正在变多。三例的共同结构：预期一致（都认为是同一份东西）＋无人负责（没有任何人的职责里写着核对）＋只在合流处暴露（转诊、争议、统考）。最后补一句关于命名的话。本章用"影子正本"而不用"不一致"，是因为不一致只是症状。真正承重的是那个没有位置却有效力的权威：它使每个人都以为自己在遵循规则，使偏离在被发现之前不被当成偏离，也使冲突在被发现之后不被当成结构问题。

九、格III：受控分岔

格III是不预期一致，且有人管着分岔。一个开源项目分叉出去，并明确宣布此后不再向上游合并，是这一种。一门语言有一个标准化机构，管着标准语的边界，而各地方言照旧生长，也是这一种。生物学里有一条更精细的版本值得单列。一个基因被复制之后，最常见的被保留下来的路径不是其中一份长出新功能，而是两份各自丢掉原来那个基因的不同部分——这一份保住了在某个组织里表达的能力，那一份保住了另一个组织的，于是谁也不能少。[5] 从外面看，这两份既不一致，也不冗余；它们是被选择这道程序管着的分岔。格III的关键不在"分开"，在分开这件事被说出来了。说出来之后，三件事同时成立：不一致不再是故障；不必再为同步付出成本；而且——这一条最容易被忽略——合流处的责任被明确了。分叉的项目会写清楚"我们与上游不再兼容"，标准化机构会说清楚"考试按标准语"。格II与格III的差别只在一句话上：分开这件事，有没有被承认。 承认了就是格III，没承认就是格II。而这两格的实际不一致程度，可以完全一样。这也是本章与消除派最硬的分离点。消除派会把格III判为违规——那里显然有多份表达同一类知识。但格III里没有任何一次"改一处漏一处"的事故，因为没有人期待另一处会跟着改。危害从来不在多份，在被辜负的预期。

十、格IV：自由分化，与它要求的容量

格IV是不预期一致，也没人管。这一格里有本章最想要的东西：新的东西。一个基因的副本可以自由变异，是因为坏了不要紧——原来那份还在供着必需的功能。这条路径被认为是生物学创新的主要来源。[4]但这里必须做本章的第一处让步，而且这处让步来自一个本章一开始并不想要的事实。多数副本的结局是死。 复制出来的基因，绝大部分并不会长出什么新功能；它们积累失活突变，退化成没有功能的残骸，最后被删掉。[6] 换句话说，那些被反复讲述的创新案例是幸存者，而幸存者背后是一大堆被清掉的失败品。这条事实逼着本章把格IV的说法收窄成一句更谨慎的话：

> 格IV的产出方差极大，它的净收益取决于这个系统能不能承受大量死亡。 承受不了的系统进这一格，只会得到一堆没人维护、也没人敢删的东西。

而"承受死亡"这件事有一个可查的条件：必须允许删除。一个组织如果每一份文档都必须留着、每一个分支都不许删、每一个项目都不能宣告失败，那么它进格IV得到的不是创新，是负担的单调增长。生物学那边之所以能靠这条路径产出新东西，是因为它有一个极其冷酷的删除机制。这条让步顺带解释了一件常见的事：为什么很多组织在鼓励"百花齐放"之后，收获的不是创新而是混乱。它们进了格IV，却没有格IV所要求的那个容量——它们不删东西。

十一、机制并不总是有效

第二处让步来自可靠性工程，它打的是格 I 。格 I 的程序式实现依赖一个假设：副本之间的错误互相独立，所以多数投票能把错的那个压下去。这个假设在 1986 年被一次实验正面检验过：二十七个程序员按同一份规格说明各自独立写出程序，然后跑一百万个测试用例。结果是每个程序单独看都极可靠，而多个程序在同一个输入上一起出错的次数，远高于独立性假设所预测的数量。[7]原因不神秘，后来的讨论把它说得很清楚：规格说明是有限的，也常常有含混之处；受过相似训练、用同一种语言、读过同一批参考资料的人，会在同一个地方做出同样的误读。独立地做，不等于独立地错。这条对本章的约束是具体的：格 I 里那个"有机制"不能只问有没有，还得问一句——

> 这个机制对相关失效有没有免疫力？ 也就是：这些副本是不是同源？它们是不是同一个模板、同一个团队、同一套工具生成的？

如果是，多数投票提供的保护远小于账面上的份数。这条在今天比在 1986 年更要紧，因为"同一套工具生成"这件事的普及程度已经完全不同了。于是格 I 也不是一个可以安心待着的格。它只是把"哪一份算数"这个问题回答了；至于答案好不好，是另一层的事。本章能给的只有一条：答案的可靠性上限，由副本之间的独立性决定，而独立性不能从"分别做的"推出来。

十二、影子正本为什么这些年在变多

如果格 II 是最坏的一格，为什么它没有一个通行的名字？一个可能的答案是它一直存在而被当成了别的东西——每一次暴露都被记成一次个人失误。本章认为这只是一半，另一半是：它确实在变多。变多的条件有三条，凑齐得越来越容易。第一，复制的成本降到了零。 复制一份纸质文件要跑一趟，复制一份数字文件是一个动作。成本为零意味着复制不再需要理由，因而也不再留下"为什么要多一份"的记录——而那个记录，恰恰是判断预期的唯一线索。第二，副本之间的距离变远了。 三份文件存在同一个柜子里，人会顺手对一下；存在三个人的云盘里，

永远不会。格Ⅱ的形成不需要任何人偷懒，只需要没有人碰巧同时看到两份。第三，也是最新的一条：现在有大量副本是被生成出来的，而不是被复制出来的。同一段规则被不同的人用同一套工具各写一遍，得到四份措辞不同、意思相近的表述。它们看上去不是副本——字面上完全不同——所以任何按字面查重的办法都发现不了；但它们表达的是同一条知识，因而完全符合消除派那条原则所定义的重复。这批副本天生就在格Ⅱ里：预期一致（大家都以为在说同一件事），无人负责，而且连"这是副本"这件事都要费劲才能看出来。这里插一句反面的话，免得这三条被读成一部退步史。这三个条件同时也在让格Ⅲ变得更容易：复制便宜、分叉便宜，于是"明确分开、各走各的"这条路的门槛也降低了。开源生态里那些声明不再合并的分叉，二十年前几乎不可能——那要靠邮寄磁带。同一批条件既在生产格Ⅱ，也在生产格Ⅲ；决定落到哪一格的，仍然是有没有人把那句话说出来。第三条同时把上一节那个约束推到了更坏的位置：这批副本不但同源，而且同源得极其彻底。它们最像的地方，正是它们一起错的地方。这三条还有一个合起来才看得清的后果：格Ⅱ的形成不再需要任何决定。从前多做一份东西是要下决心的——要花钱，要占地方，要有人搬。决心留下痕迹，痕迹里就带着"为什么要多这一份"，而那句话本身就回答了预期问题。现在多一份不需要决心，于是也不留痕迹，于是预期这件事从来没有被表述过——它只是被默认。一个从未被表述过的预期，是无法被撤销的。你没法开会宣布取消一件从来没有人正式说过的事；在场的人会觉得莫名其妙。这就是格Ⅱ难处理的深层原因：它的两个成因里，"机制不在"可以靠配人解决，而"预期在"因为从未被说出口，连解除的对象都找不到。

十三、几个容易被当成同一件事的说法

下面逐个处理，每一个都落到具体的判别场景上。最像的是消除派那条原则本身。"每一条知识必须有单一、无歧义、权威的表示"（Hunt & Thomas, 1999）——它已经站在门口了，因为它是唯一一条把"权威"写进正文的。分离线在处方上：那条原则的处方是减少份数；本章说份数与危害无关，要改的是预期或机制。判别场景：一个有五百份副本、明确宣布不预期一致的生态（分叉后各自发布的开源项目群）。按那条原则判，这是极端违规，事故率应当极高；按本章判，这是格Ⅲ，事故率应当低于同一生态里那些声称保持同步却没有同步流程的项目。两个预测在公开的缺陷追踪数据上会分开。第二个是"唯一真相源"这一族，在信息系统与主数据管理里是标准做法：为每一类实体指定一个权威系统，其余系统只读。它比上一条更具操作性，也更接近格Ⅰ的位置式。分离线：这一族默认"指定"本身就解决了问题。本章说指定只完成了一半——指定之后如果没有回执与核对，它就变成第十八节要讲的那种更坏的形态。判别场景：统计主数据项目的失败原因。这一族预测失败主要来自技术集成；本章预测失败主要来自没有人被授权裁决冲突——即指定了权威系统，却没规定业务方不服时找谁。第三

个来自分布式系统。那里有一条广为人知的结论：在网络可能分区的条件下，一致性与可用性不可兼得。这与本章第三节讲的"两种安全"形状相似。分离线：那条结论是关于读写操作的形式化结果，它假设"什么算一致"已经定义好了。本章处理的正是它假设掉的那一层——一致是不是被期待、由谁来裁决。判别场景：一个技术上完全一致的系统（所有节点同一份数据），照样可以处在格 II——只要那份数据在系统之外还有几份 Excel 副本，而人人以为它们一样。技术一致性与本章说的一致预期，可以取相反的值。第四个来自社会科学：共同知识。那一族关心的是"我知道你知道我知道"这种递归结构。影子正本看起来像共同知识的失败：大家以为存在共识，其实没有。分离线：共同知识的框架关心信念的层级；本章关心职责的空缺。判别场景：把所有人召集起来，当众宣布"这三份是不一样的"。共同知识的框架预测问题就此解决（信念对齐了）；本章预测不解决——除非同时指定谁负责，否则三个月后它们会再次分开。这个实验很便宜，而且大多数组织其实做过，只是没记录结果。第五个也来自社会科学：未言明的相互预期。劳动关系研究里有一族概念处理雇员与组织之间没写进合同、却被双方默认的义务。影子正本里那个"应该一样"的预期，形式上确实属于这一类。分离线：那一族的预期是关于对方会做什么；本章的预期是关于哪一份文本算数，而后者可以被一句话问出来，前者不能。判别场景：问卷的可回答率。问"你觉得公司应当为你做什么"，答案发散且难以核对；问"这两份不一样时哪份算数"，答案是有限的几项，可以直接统计分歧度。第六个来自法律与档案：正本与副本的区分。这套区分极其成熟：正本有原始签名，副本只是复制品，效力不同。分离线：那套区分预设了正本一定存在，问题只是如何辨认与保管。本章的第二格里正本根本不存在——所有人以为有，但从来没有任何一份被赋予过那个地位。判别场景：追问"那份正本在哪儿"。法律框架预测追问总能得到一个答案（在某个保险柜里、在某个登记处）；本章预测在格 II 里追问会得到互相矛盾的答案，而且没有人觉得这是件怪事。第七个是复制与原作的那一族讨论。那一族讲的是复制取消了原作在特定时空里的独一无二性。分离线：那一族关心的是原作的光环是否消失；本章关心的是原件的裁决权如何处置。这两件事可以取相反的值：一份带数字签名的权威文本，光环早已荡然无存，裁决权却清清楚楚。判别场景：给一批人看两份内容相同、其中一份有正式签署的文本，问哪份"更真"、再问哪份"算数"。那一族预测两个问题的答案高度相关；本章预测在数字场景里两者可以分离——多数人会说"都一样真"，但"算数"的那份指认得毫不含糊。第八个是本章必须处理的一个近处的说法：有一类分析处理的是原物够不着时各领域如何应对——手稿失传、原始状态无法观测——并论证不同领域的处置办法造出了不同的替代物。本章与它形状相近而问题不同：那里的原物是够不着的，问题是缺口怎么填；本章的所有副本都在手边，问题是它们之间谁说了算。两者在同一个案例上会分岔：一份人人都能打开、只是各自改过的文档，在那套分析里不构成问题（原物并没有丢），

在本章这里正是最坏的一格。第九个来自版本控制。 现代的版本控制系统提供了一个几乎完美的格 I 实现：任何两条分支的差异都可以被机械地检出，合并时冲突会被强制摆到人面前，要求当场裁决。它比本章说的任何机制都硬——它不给人"没注意到"的余地。分离线在覆盖范围上：那套机制只对进了库的东西有效。而现实中的格Ⅱ，绝大多数发生在库外——邮件附件里的那份、打印出来贴在墙上的那份、某人本地改过没提交的那份、按同一条规则各自写出的四份说明。判别场景：统计一个团队的不一致事故，看有多少发生在版本控制覆盖的对象上。那套机制的框架预测事故集中在合并冲突处理不当；本章预测绝大多数事故发生在库外，且与库内的规范程度无关。这条在任何一个用版本控制的团队里都能查。这九个里，最近的仍是第一个。 那条原则是唯一一条把"权威"这个词写进定义的，也是唯一一条同时管到了知识与表示两层的。它和本章之间只差一个动作：它把"权威"和"单一"焊在了一起。一旦把这两个词拆开——权威要有，单一不必——它的全部反例（多副本网络、多语种条约、受控分岔）就各归其位了。本章接的正是它焊住的那个地方。所以准确的说法不是本章推翻了它，而是：它在"改一处会漏"这个问题上是对的，而它把处方定得比它的诊断窄。

十四、在别的行当里应该看到什么

一条判断如果只能在提出它的那几个例子上成立，它就还没离开那几个例子。下面十二处，每一条都是可以查的。医疗。 同一份指南被各科室改编。预测：不一致的暴露集中在转诊与会诊，院内单科使用时几乎不暴露；且不良事件的归因记录里，这类事件会被高比例记成"沟通问题"而非"文件不一致"。查法：调阅不良事件报告，看"指南版本"是否曾作为一个字段出现过。法律实务。 同一段条款被复制进多份合同并各自演化。预测：争议中该条款的解释分歧率，与该事务所是否设有条款库管理员强相关，与条款被复制的次数不相关。这条可以在一家规模足够的事务所内部做回溯。教育。 同一门课的课件由多位教师各自维护。预测：跨班成绩差异在统考科目上显著、在各班自主命题的科目上不显著；且教师本人普遍认为自己用的是"同一份课件"。药品说明。 同一药物在不同国家的说明书各自更新。预测：不一致在跨境医疗与药品代购时暴露；且更新滞后最长的那一版，恰恰是被最多人当作"权威版本"引用的那一版——因为它最老、最容易被搜到。统计口径。 同一个业务指标在多个报表系统里各有实现。预测：口径分歧的暴露发生在跨部门汇报的场合，而不是日常运营；且分歧被发现后的第一反应是"以哪个系统为准"这句话本身第一次被问出来。工程图纸。 同一零件的图纸存在设计、工艺、供应商三处。预测：不一致在首件检验时暴露，且返工成本与副本数无关、与是否设有更改通知回执强相关。语言。 有标准化机构的语言与没有的语言。预测：前者的方言分化照样发生（格Ⅲ），但在正式场合的裁决是明确的；后者在需要统一书写时会出现典型的格Ⅱ症状——每一方都认为自己写的是"正确的"。宗教文本。 抄本传统里那些设有逐字计数与核对制度的（例如

把每卷的字数、中间字都记下来备查），与没有的。预测：前者的异文率显著低，而且异文的分布形态不同——前者集中在计数不覆盖的层面（读音、断句），后者散布在全文。开源软件。 分叉后明确宣布不再合并的项目（格Ⅲ），与把上游代码直接复制进自己仓库、又声称会跟进上游的项目（格Ⅱ）。预测：后者的安全漏洞修复滞后期显著更长，因为没有人被指定去跟进；且滞后期与项目活跃度无关。主数据管理。 预测：这类项目的失败复盘里，"没有人被授权裁决业务口径冲突"出现的频率高于任何技术原因。这条可以在公开的失败案例复盘中统计。分布式系统之外的数据。 预测：一个技术上强一致的系统，其周边的电子表格副本数量与该系统的"权威性宣称强度"正相关——越是被宣布为唯一真相源，私下的影子副本越多，因为业务方的例外需求无处安放。这条反直觉，也最容易做。生物学（反向）。 预测：在基因组里，被保留下来的重复基因中，通过各自丢掉不同部分而互补保住的那一类，其比例应当高于长出全新功能的那一类。这条已有文献支持，本章借用它作为格Ⅲ在自然界的存在证据，不算独立证据——明写。十二处里，教育与统计口径那两条最便宜；"权威性宣称越强、影子副本越多"那条最硬，因为它的方向与常识相反。如果那一条查下来是负相关，本章关于格 I 位置式实现的整段判断就要重写。

十五、怎么在半小时里查完一个部门

这一节交出一份操作办法，因为一条判断如果只能被同意、不能被执行，它离一个说法就不远了。第一步，列副本，不看内容（十分钟）。 挑一件这个部门真的在用的东西——一份标准、一份清单、一段条款、一个指标口径。然后问在场的人：这东西你手上有吗，存在哪儿。把所有位置记下来，不打开，不比对。这一步只求数量与位置，不求差异；打开比对是后面的事，先打开会让所有人立刻开始辩论谁对，而那正是要避免的。第二步，各问两句（十分钟）。 对每个持有者分别问，不要当众问——当众问会收敛到一个答案，而分歧本身才是读数。第一句："这几份应该是一样的吗？" 第二句："如果不一样，哪一份算数？"记原话，不要归纳。 "按共享盘那份"和"按最新的那份"看起来接近，其实是两个完全不同的答案：前者指向一个位置，后者指向一个属性，而这两个在某次修改之后会指向不同的文件。第三步，看最后修改时间（五分钟）。 不比对内容，只看每一份最后一次被改是什么时候。时间跨度就是已积累分歧的粗略上界。这一步之所以放在这里，是因为它便宜到几乎没有成本，而它给出的量级估计通常比逐行比对更早派上用场。第四步，判格（五分钟）。- 第一句多数答"应该一样"，第二句答案一致且能指出谁负责 → 格 I 。到此结束。- 第一句多数答"应该一样"，第二句答案分歧，或出现"应该都一样吧"这种反问 → 格Ⅱ。- 第一句多数答"不用一样"，并能说出从什么时候、按什么规矩分开 → 格Ⅲ。- 第一句多数答"不用一样"，说不出规矩，且这些东西谁也不删 → 格Ⅳ，且缺容量。

判成格Ⅱ之后，先别急着解决。按第十七节那条硬规矩，只补一半比不补更糟。在配上人和核对动作之前，能做的最有用的一件事是把第二步的原话贴出来——让所有人看到彼此的答案不一样。这一步不解决问题（第十三节里已经说明，光对齐信念不解决问题），但它把一个从未被表述过的预期第一次变成了可以被讨论的东西，而第十二节说过，从未被表述的预期连撤销的对象都找不到。最后一句提醒：这套办法查得出格，估不出损害。分歧度告诉你这是哪一格，告诉不了你已经付出了多少代价——因为代价的暴露由使用决定，而使用的分布这套办法没有采集。要估损害，得另外去看合流处：转诊、审计、统考、争议、跨部门汇报。

十六、这条判断管不到的地方

第一，"一致"在有些领域不可定义。两份手工制品、两次演出、两个译本，说它们"一致"没有意义。这类场合本章的两条轴都失效。第二，预期是分布的，不是二值的。一个组织里往往有一半人预期一致、一半人不预期。本章按"多数是否预期"简化，边缘案例会失灵，而边缘案例并不罕见。第三，"机制"没有分级。一个每年核对一次的流程与一个实时同步的程序，在本章这张表里落在同一格。这是一个真实的粗糙之处：格Ⅰ内部的差别可能比格Ⅰ与格Ⅱ的差别还大。第四，本章不处理谁有权设定预期。"这些应当一致"这句话由谁说了算，是一个权力问题，本章完全没碰。而在很多真实场合，格Ⅱ的形成正是因为有人有权提出预期、却无权配置人手。第五，安全关键领域不适用格Ⅳ。在飞机、核电、药物这类场合，即使不预期一致也不能容忍自由分化，因为"承受死亡"这个条件在道义上不成立。第五之二，本章假设"哪一份算数"这个问题有意义。在一些场合它没有意义——两份都算数且必须同时满足（例如两个监管机构各自的要求），此时不一致不是待裁决的冲突，而是一个真实的两难。本章的判据在这类场合会把健康的两难误判成格Ⅱ。第六，本章给不出"预期覆盖范围"的度量。承重判断里那句"危害的量由预期覆盖的范围决定"，本章只能定性地说，量不出来。这是本章最想要而没有的东西。

十七、怎么判它错

六条，读数各不相同。判据一（低成本、正向读数）。在同一个组织里选若干组持有多份同类文件的人，各问两句："这些应该一样吗？"和"如果不一样，哪份算数？"记录第二问的分歧度与该组的历史不一致事故率。本章预测：事故率与分歧度显著相关，与副本数量不相关。若事故率与副本数量强相关而与分歧度无关，本章的核心（数量无关、预期有关）失败。这条只要一份问卷和一份工单记录。判据二（打反噬预言）。找一组处在格Ⅱ的对象，只做一件事：正式宣布"以某一份为准"，不配任何核对流程。本章预测事故率不降反升，且暴露时点后移。若事故率单调下降，第十六节整节作废。判据三（打格Ⅳ）。在不预期一致、也无人管的

系统里，比较允许删除与不允许删除的两类。本章预测后者的维护成本单调上升而新产出接近于零。若两类无差异，"容量"这个条件不成立。判据四（打格Ⅰ）。 对采用多数裁决的系统，比较副本同源与不同源两组的相关失效率。本章预测同源组显著更高。这条已有 1986 年那次实验支持，属借来的证据，不计入本章的独立检验。判据五（打暴露时点那条读数）。 格Ⅱ的不一致应当只在合流处暴露。若在单点使用中也高频暴露，那么"没有单点读数"这条性质错了，第七节的三条性质要重写。判据六见末节的赌注：它有截止日期，到期不兑现即作废。另外，本章主动交出两个填不上的洞：预期覆盖范围量不出来（第十六节第六条），以及本章说不清预期是怎么形成的——为什么有些副本一出生就带着"应该一样"的预期，有些不带。本章只能观察它在不在，说不出它从哪来。

十八、按它设计干预会怎样出事

本章的处方最容易被砍掉一半，而砍掉一半比不做更糟。格Ⅱ的病因是两条：预期在、机制不在。因此出路只有两条——要么把机制补上，要么把预期撤掉（明确宣布这些从此不必一致，并说清合流处怎么办）。两条都行，但必须整条做完。现实里最省事的做法是发一份通知：从即日起以某某版本为准。这一步是只补了预期的正式性，没补机制。它的后果不是中性的，是负的：- 从此所有偏离都成了违规。而各分公司真实存在的差异并没有消失——报销场景确实不同、病人构成确实不同。- 于是偏离不再上报，转入地下：私下维护一份"实际在用的版本"，对上声称用的是标准版。- 暴露时点被推得更远，因为现在多了一层掩盖；而暴露时的损害更大，因为积累得更久。- 更麻烦的是，这套通知会让组织觉得问题已经解决了——账面上有了一份权威文件，检查时人人都答得出"以总部版为准"。格Ⅱ因此变得更难被发现，而不是更容易。

所以本章把处方写成一条硬的："指定正本"这个动作，如果不同时指定一个人和一次核对，就不要做。 什么都不做，至少还留着一个诚实的读数——问起来大家答案不一致。发了通知之后，那个读数就没了。第三种出事方式来自这条判断本身的方便。 一旦"要么补机制、要么撤预期"这句话被接受，撤预期会显得比补机制便宜得多——补机制要人，撤预期只要一句话。于是可以预见一种应对：把大量本该一致的东西宣布为"不必一致"，从而在账面上把格Ⅱ变成格Ⅲ。这个动作在形式上完全符合本章的处方，实质上是把风险从台面转到台下。本章能给的鉴别办法只有一条，而且它落在合流处：格Ⅲ必须交代合流处怎么办。 分叉的项目要说清与上游不再兼容，标准化机构要说清考试按哪一套。一次不交代合流处的"宣布不必一致"，不是格Ⅲ，是给格Ⅱ换了个名字。 判据很简单：问一句"那么两边碰上的时候按什么"，答不出来就是换名字。第四种出事方式是滥用。"影子正本"这个说法很好用，因为它以"找不到那份权

威"为特征，所以任何一次配合失败都可以被这么描述，而且无法被反驳。本章能给的唯一防线是那句一分钟判据必须真的去问，并记下每个人的原话：如果所有人答案一致，这个词就不适用。这条防线不牢，但它要求提出主张的人交出一次可核对的记录，而不只是一个说法。

十九、这篇文章自己在哪一格

按前面的规矩，一条判断要问它对自己适不适用。本章处在第二格，而且已经因此出过一次事。支撑本章这条判断的那套做法，被写在好几个地方：一份工作规程、一份供检索用的条目库、一份用来自查的清单，以及若干篇正文里对同一条判断的复述。这些是同一条知识的多份表示。我预期它们一致；没有任何人被指定去核对它们。后果是具体的，而且是可查的。就在写作本章之前不久，同一套规程被两个互不知情的执行过程各自读了一遍，各自按同样的约束选源、各自撞出同一条判断、各自写成一篇近两万字的文章——两篇的节标题逐节相同。直到发布前的例行检查，才发现其中一篇已经在几个小时前被发出去了。没有任何机制在这两条线之间报警，因为从来没有人被指定去看它们是不是同一件事。这个例子有一个让本章尴尬的性质：它同时也是本章的证据。第八节说格Ⅱ的暴露由使用决定、不由损害决定——那两条线的重复从一开始就存在，而暴露发生在第一次要把它们同时放上去的那一刻。第八节说暴露之后通常被消化成一次个人失误——第一反应确实是"这次没查"，而不是"我们缺一个机制"。所以本章对自己的处置，按第十七节那条硬规矩：不发通知。要么给那份规程配一个核对的动作，要么明确宣布各条线不必一致、并说清合流处（发布前）怎么办。在两者之一被真的做出来之前，本章不宣称自己已经离开第二格。还有一点得说明。本章的可信度不能靠"作者做到了"来担保——作者显然没有做到。它只能押在第十四节那十二处读数与第十七节那六条判据上：若那些查下来是错的，作者做没做到都不重要；若查下来是对的，作者没做到也不构成减损。

二十、一个到期的赌注

到 2032 年 12 月 31 日，押的是这样一件事。在软件工程或知识管理的主流规范里，出现一条被广泛采用的实践，其形式是"允许重复，但必须声明是否预期一致，以及不一致时按谁"——也就是把处方从"消除重复"改成"标注预期"。它可以长成一个字段、一条模板里的必填项、一份检查表上的一行，形式不重要。如果到期没有出现，那不算数——没出现可以有很多原因。真正押的是反面：如果它出现了，被广泛采用了，而相关的不一致事故率没有变化——也就是说，标注了预期而事故照旧——那么本章的承重判断（危害由预期而非数量决定）就被推翻了，本判断作废。还要说明为什么押的是这个而不是别的。本章完全可以押一条更容易赢的：比如押"影子正本"这个说法会流行起来。但一个说法流行与否，跟它对不对没有关系，而且流行这件事本章自己也能推动，押它等于押自己。押规范的形态就没有这个毛病——它要求

一个本章影响不到的、由许多人的实践共同决定的东西发生改变，而且改变的方向是可辨认的：从"消除重复"改成"标注预期"。我倾向于认为它会出现。但这句话不值钱：押注的价值在于它有截止日期，不在于押注的人有多少信心。

注释

[1] 该原则的原句为"Every piece of knowledge must have a single, unambiguous, authoritative representation within a system"，出自 Andrew Hunt 与 David Thomas 的《The Pragmatic Programmer》（1999）。本章只取其原始表述，不取后来被简化成"不要复制粘贴"的通俗版本——两者的差别在正文 2.1 节已说明。[2] 多副本互相校验、按多数修复、且不设正本的做法，本章取自数字保存领域的多副本保存网络实践。具体实现细节各系统不同，本章只用其结构。[3] 多语种条约各文本同等作准、歧义时依条约目的与宗旨解释，见《维也纳条约法公约》第 33 条。引用前请按官方文本核校条款序号与措辞，本章只用其结构，不引原文。[4] 复制之后一份继续承担原功能、另一份因而得以自由变异，这一路径通常追溯到 Susumu Ohno 的《Evolution by Gene Duplication》（1970）。[5] 两份各自丢掉不同子功能因而都被保住的路径（复制—退化—互补），见 Force et al.（1999）。[6] 重复基因的多数最终丢失、退化为假基因，见 Lynch & Conery（2000）及此后的综述。本章以此作为对格Ⅳ的约束。[7] 二十七个独立编写的版本在同一输入上共同失效的次数远超独立性预测，见 Knight & Leveson（1986）。后续讨论把原因归于规格说明的含混与编写者共享的训练背景。[8] 本章所引各家的年份与表述按公开检索结果录入；页码须按所用版本核校。文中的核心检验均未执行，交付的是命题、判据与方案。

附：与学科通融专栏另外两篇的关系

学科通融专栏另有两篇与本章形状相近而问题不同，正文里已各划一刀。《我们说的「本来的样子」，是我们造出来的》处理的是原物够不着时各领域如何应对，问题是缺口怎么填；本章的所有副本都在手边，问题是它们之间谁说了算——一份人人都能打开、只是各自改过的文档，在那一篇里不构成问题，在本章这里正是最坏的一格。《痕迹能不能被复制，决定了这门学科相信什么》问的是痕迹的可复制性如何塑造一门学科的默认判断，自变量在证据的物理性质上；本章的自变量在看着副本的人有没有预期上，两者正交。本章的三家源全在站外，分属软件工程、数字保存与条约法、演化遗传学——三家谁也不读谁，却在同一个事实上给出不能并存的判决。下面三条给出可直接点开核对的原始出处。正文第二节逐一交代三家各自说到了哪一步，第三节把三处对顶写死，第五节说明它们各自为什么看不见本章那一格，第十三节再与九种最容易被混为一谈的说法逐条分界——包括本章承认最可能把它吸收掉的那一个。

枢纽章

可及性把读数钉在满分上

论一个参数的两种读法，及为什么衡量知识的通行办法会给损失方加分

位置　本章排在第一编之后。它不引入任何新的经验材料，只从前三章已经给出的定义里推出一组关系式，并把这组关系式用作后三编的接口。凡本章出现的数字，全部来自本书某一章已经报告过的实测，出处逐条注明。

一、只需要一个新参数

前三章各自成立，互不引用。它们是三个不同的三学科组撞出来的三条判断，写作时彼此并不知情。把它们放进同一本书，需要一个能同时进入三条判断的量。本章只引入一个：

可及性 a　面对一个具体任务，一份合格产物由外部即时供给的概率。

a 是一个技术条件，不是一个心理状态，也不是一项政策。它可以被独立测量：给定一个任务集合，数其中有多少个任务可以在数分钟之内从外部拿到一份在通行验收标准下合格的产物。二〇二二年之前，绝大多数需要判断的任务上 a 接近零；此后它在很多任务上迅速趋近一。

本书讨论的全部现象，都发生在 a 上升的过程里。而本书的主张是：a 上升带来的后果，不能从 a 上升这件事本身读出来，必须经过下面这组关系。

二、第一条关系：借道不会停在中途

第五章给出的机制是使用依赖的可塑性——一条通路被使用会被强化，长期不被使用会退化。把它写成关于借道率的递推：

$\lambda(t+1) = \lambda(t) + k \cdot a \cdot (1 - \lambda(t))$，其中 λ 为借道率（达成结果时走替代路径而不自己重推的比例），$k > 0$ 为该系统的可塑性系数。

解出来是

$\lambda(t) = 1 - (1 - \lambda_0) \cdot (1 - k \cdot a)^t$

于是有第一个结果，它比通常的说法硬得多：

只要 a 严格大于零，λ 单调上升并趋于一。a 只决定到达的快慢，不决定终点。

通常的说法是「适度使用没问题」。这条关系说：适度使用不是一个状态，是一段速度。把 a 从零点九降到零点三，λ 到达零点九五所需的轮数从大约三十轮变成大约一百轮——在一段职业生涯的尺度上，这个差别不改变结局。

要让 λ 停在某个 λ* < 1，必须存在一个把 λ 往回拉的力。本书的主张是：在现行的记录方式里，这个力一项也没有。第五章说明了原因——借道对结果读数恒为零，对当事人的省力感符号为正，因此既没有外部信号要求他不借，也没有内部信号。

这条可以判错。若在某个 $0 < a < 1$ 的环境里观察到 λ 长期稳定在明显低于一的水平，且该环境中没有人为的封路安排，则本章的递推形式错，全书的脊梁断在这里。

三、第二条关系：复推率是同一个 λ 的另一种读法

第一章的判断是：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。不重新生成，就不会有重新生成的可靠性。写成关系：

$$\rho = \rho_{max} \cdot (1 - \lambda)$$

其中 ρ 是复推率，ρ_max 是该人在完全不借道条件下能达到的复推可靠性。合起来：

$$\rho(t) = \rho_{max} \cdot (1 - \lambda_0) \cdot (1 - k \cdot a)^{t} \quad \Longrightarrow \quad \rho \to 0$$

这一条不是「用得多就学不会」的另一种说法。它说的是更窄也更硬的一件事：复推率不是一项被消耗的存量，它是一项每次使用时被重新生产的能力；停止生产它，它不是慢慢变少，是不再被生产。存量会有折旧曲线，可以谈半衰期；这里没有折旧，只有停产。

四、第三条关系：产物型读数与 λ 正交

第五章有一段实测，本章要用两次。一组前交叉韧带重建术后的运动员做单腿跳远，患侧与健侧的距离比是百分之九十七，达到几乎所有回归赛场清单的标准；同一次测量里，患侧膝关节在蹬伸阶段做的功只有健侧的百分之六十九，缺掉的部分由髋与躯干补上。

跳出去的距离不知道这件事。写成关系：

$\partial y / \partial \lambda = 0$，其中 y 为任何以产物为形式的读数。

这不是说 y 测得不准。是说 y 测的正是那个没有变化的东西。一个能区分路径的指标，就不再是结果指标；这是定义上的事，不是精度上的事。

第六章在一个与运动医学毫不相干的地方给出了同一个形状。在职业信息网络数据库上（O*NET 29.1，n = 879，预注册）：自动化程度与决策频次的相关是 0.042，与决策自由度的相关是 −0.199。步骤的数目没变，选择的余地在减少。

两组数字来自两个互不相干的学科，一组测的是人的膝盖，一组测的是八百七十九个职业的岗位描述，而它们是同一句话的两次显影：

凡进入产物的都留下，凡不进入产物的都在退。而只有前者被记录。

由此得到本章的中心结果：

推论一　在 a 上升的整个过程里，y 不降——通常还会升，因为借来的那条路往往更省力也更稳定；同时 ρ → 0。y 高与 ρ 低不是一对需要权衡的目标，它们是同一个 λ 的两种读法。

一个组织若只看 y，它在 λ 从零走到一的全过程中，看到的将是一条持续向好的曲线。这不是它疏于监测。是它监测的那一样，正是这件事唯一不会留下痕迹的地方。

五、第四条关系：唯一的读法，与它的价格

既然 y 对 λ 恒为零，读出 λ 只剩一条路——第五章称为剥夺落差：

Δ ≡ y(封路) − y(不封路)

临时封掉替代路径，量结果跌多少、多久回得来。第五章把这套测量在一个可以真封路的学习系统里整个跑完，结果打掉了它自己四条预测里的两条：替换不是慢慢加深的，它一次到位；封路后的恢复不是更快而是更慢。随机化确实能把落差压掉百分之九十二，代价是结果读数掉十三个百分点，只有在替代路径失效概率高于约百分之四十六时才划算。

把这三个数字接到可及性上，得到本章第二个结果：

推论二　Δ 的测量成本随 a 单调上升，而 Δ 的诊断价值也随 a 上升。唯一能读出这件事的动作，恰恰在最需要它的时候最贵。

成本上升有两个来源，都不是心理上的。其一，a 越高，可用的替代通道越多，封路要挡的口子越多。其二，第五章那条十三个百分点是在 λ 尚低时测到的；λ 越接近一，封路后跌得越深，因为不借那条路已经很久没有被走过。

这条推论有一个不受欢迎的实践含义：剥夺落差应当在 a 还低的时候先测一次，作为基线。而这件事在任何一个组织里都不会被提出来，因为提出它的时候，所有读数都还是绿的。

六、第五条关系：借清停摆，与一个此前没有名字的形状

到这里为止的四条关系，处理的是「重推这一段没有发生」。第四章处理的是另一件事：重推这一段发生的时候，代价落在哪里。

第四章的判断是：学会一件事，是把有限的清晰度从相邻区域搬到目标区域，没有搬走就没有学会。读数是借清率

$$b = \text{相邻处清晰度下降量} \div \text{目标处清晰度上升量}$$

在容量固定的条件下 $b > 0$，且大致守恒。第四章的二维辨别表指认了最危险的一格：搬走的清晰度不可回写、又恰好落在将要用到的近邻上，于是学得越好、失能越深。

现在把它接到 λ 上。不重推就不搬，于是

$$b \propto (1 - \lambda) \quad \Longrightarrow \quad b \to 0$$

到此为止，看上去像是一个好消息：不学，就不必付学的代价，近邻的清晰度保住了。而这正是本书要指认的那件事的入口。把两边合起来看，得到的是一个形状，不是一个数量：

· 自己重推的人，在他重推过的地方有一个峰（目标处清晰度上升），在峰的近旁有一个坑（被搬空的相邻区域）。峰与坑是同一次搬运的两端，不能只要一端。

· 随时取用的人，两样都没有。他的知识分布是一片平地。

本章为这片平地命名：

> 均浅　既无峰亦无坑的知识分布形态。它不是「知道得少」——在覆盖面上，均浅者往往比深学者知道得多；它是「哪一处也没有被搬运过」。

配套读数：

$$\text{锐度 } \kappa = \text{峰高} \div \text{邻坑深}$$

对均浅者，κ 是零除以零。这里要说清楚一件容易被读错的事：κ 在均浅者身上不是低，是没有定义。 这不是测量精度不够，是那个量在那里不存在。一切试图「把 κ 测得更准」的努力都无从下手，因为要测的两样东西都不在。

κ 的可执行代理有两个。其一，同一人在相邻两个题域上的表现差——峰与坑必然相邻，这是第四章「近」的口径要求的。其二，第九章的静息变异率：峰的位置，正是他在没有人要求的时候还会去动的那一处。 第九章说明了那个量为何只在无任务、无考核、无故障的时段里存在；本章要补的一句是：它同时是本章这个形状的唯一无侵入读法。

六之二、均浅与七种最接近的说法

前一节造出了一个本书主张为新的东西，而全书其余各章都为自己的对象做过一节「与最接近的几种说法的分界」——这一节补上均浅欠的那一次。补得晚，因此先说清楚它欠在哪里：均浅是全书唯一被主张为新存在物的对象，而它此前是全书唯一没有走过这道工序的。一个不做划界的新名字，与一个换了说法的旧名字，在文本上无法区分。

七位候选人按由近到远排。每一位都给一处判决性对照——一个两说会给出相反答案的具体场合。

其一，样样通、样样松（jack of all trades, master of none）。这是最省事的一刀，也是最该先挡的一刀：均浅不就是这句老话吗？不是，而且分界很硬。这句老话说的是能力的水平——他哪一样都不精。均浅说的是分布的形状——他哪一处也没有被搬运过，因而既没有峰，也没有坑。判决性对照：一位在三个领域各练了五年、每一样都到中上而没有一样拔尖的人，按老话是典型的样样通；按本节的定义，他在每一个领域都发生过搬运，因而有三个峰和三片坑，κ 有定义且不低。老话按结果的高度分类，均浅按有没有发生过搬运分类，两者在这个人身上给出相反的判定。这也解释了为什么老话是一句贬语，而均浅不是——本书不主张均浅者能力差，第七节的推论正好相反：他在覆盖型量表上得分更高。

其二，T 型人才与它的变体（π 型、梳型）。这一族说的是广度加一根深度的竖杠，处方是"横杠要宽、竖杠要深"。它与均浅共享一个坐标——广与深——而正是这个坐标使它到不了这里：T 型模型把广与深当作两个可以独立调节的量，因而"又广又深"在它的框架里是一个连贯的目标。第四章的守恒说这不成立：在容量固定的条件下，竖杠的深度是从横杠上搬来的，竖杠越深，它两侧的横杠越薄。判决性对照：给一批人做半年的深度专项训练，T 型模型预测横杠不变、竖杠加长；本节预测竖杠加长的同时，与该专项在表示上相邻的那几格会退，而且退的位置可以事先指出来。这一条便宜、可跑，且两说的预测方向相反。

其三，广度—深度权衡（breadth-depth tradeoff）。这是七位里唯一在形式上真正接近的一位：它也说广与深不能兼得。分界在权衡的单位。这一族的权衡在时间或注意力上——学甲的时间不能同时学乙，因而是一个预算约束，与甲乙的内容无关。本节的搬运在表示的邻域上——搬走的清晰度只落在相邻处，与你花多少时间无关。判决性对照：两个人花同样的时间学同样多的东西，一个学的是彼此远离的三样，一个学的是彼此相邻的三样。时间预算说两人相同；本节说后者的相互干扰远大于前者，且干扰量可由邻域距离预测。免疫学那一支恰好把这个对照做过了——第一次学得越牢，对近邻变体越认不出，而对远处的抗原毫无影响。时间预算解释不了"为什么偏偏是近邻"。

其四，专家知识组织的一支（以问题深层结构而非表面特征归类的那条经典发现）。这一支说专家与新手的差别不在知道多少，在组织方式。它与本节最接近的一点是都不把知识当成存量。分界在它只描述峰不描述坑：那条研究传统里，专家的表征被描述为更深、更结构化、迁移更好，而没有任何一项指标去测他因此在别处失去了什么。判决性对照：找一位在本领域组织度极高的专家，测他在一个表面相似而深层结构不同的相邻问题上的表现。这一支预测他仍然优于新手（因为他按深层结构归类）；本节预测他在这一格上可能劣于新手——第四章那个最危险的格丁说的正是这个，而免疫学的原始抗原原罪是它已经被观测到的形态。这一处是本节与它最有分量的分歧，也是本节最容易被判错的地方。

其五，通识教育与博雅传统。这一族在价值上主张广，理由是公民资质、判断力与可迁移性。它与本节的关系不是竞争而是错层：它是一个规范主张（应当广），本节是一个描述性形状（哪一处被搬运过）。混淆两者会得到一个本书不主张的结论——通识教育培养均浅者。本节的判定与课程的宽窄无关，只与在每一格上有没有发生过重推有关：一门涉猎十二个领域而每一格都要求学生自己重新推出结论的课程，产出的是十二个小峰与十二片小坑，κ 有定义；一门只覆盖一个领域而全程提供答案的课程，产出的是均浅。判决性对照就是这两门课，而它们在课程宽度这个变量上的排序是反的。

其六，狐狸与刺猬（以及预测研究里对它的操作化）。这一族说狐狸知道很多小事、刺猬知道一件大事，而在长期预测上狐狸更准。分界在它测的是使用方式，不是形成方式：狐狸型被界定为愿意采纳多个来源、自我修正、不承诺单一框架，这些是行为倾向。本节的均浅是一个关于分布形状的陈述，与愿不愿意采纳多个来源无关。判决性对照：一个从不自己重推、全部结论取自外部来源、且乐于在来源之间切换的人——按预测研究的编码他是典型的狐狸，按本节他是典型的均浅。两说对同一个人给出方向相反的评价，因而它们不是同一件事。

其七，涉猎（dilettantism）与浅层加工。这两个说法都带贬义，都指向"没有下功夫"。分界最简单：它们都以投入为判据（花的时间少、加工的层次浅），而本节以搬运有没有发生为判据。一个人可以在一件事上投入极多的时间、极深的注意力，而全程有一份现成的合格产物在旁边——按投入判据他不是涉猎者，按本节他仍然是均浅，因为他从未在任何一处付出过近邻的清晰度。这正是「答案随时可得之后」这个条件带来的新情况：在此之前，投入与搬运几乎总是同时发生，因而两个判据从不分开；此后它们分开了，而所有现成的说法都停在投入那一侧。

七位合起来说明了什么。均浅要成立，需要它在这七处都站得住；只要其中一位能给出与本节相同的预测且更简单，本节就应当让位。目前七处的分界都落在同一个地方：既有的说法按水平、按投入、按行为倾向、按组织方式分类，本节按「有没有发生过一次带代价的搬运」

分类。这个判据之所以此前没有被单独使用，是因为在可及性接近于零的年代它与前四种判据高度共变——不下功夫就学不会，学会了就一定搬运过。第一节那个参数上升之后，共变解体了，而分类学没有跟上。

这一节可以判错的方式。若在一批人身上同时测锐度 κ 与上述七种说法的任一操作化指标，两者相关达到 0.9 以上，则均浅只是那个指标的一个新名字，本节与第六节应当撤销，第七节的推论三改挂在那个既有指标之下重新表述。本书没有做这次测量。七位候选人里，其二、其三、其四三处的判决性对照都可以在一个学期内跑完，且不需要新建采集系统。

七、第六条关系：覆盖型量表给损失方加分

这一条是本章的承重结果，也是全书唯一一条既不属于任何一章、又能被任何一章的材料反驳的判断。

考虑一份通行的覆盖型量表：把一个领域切成若干格，每格给一个及格线，总分是及格格数。几乎所有的考试、认证、能力矩阵、岗位胜任力模型都是这个形状。在这样一份量表上：

· 峰在封顶处失效。一格及格就是满格。深学者在他有峰的那一格得到的分，与一个刚好及格的人完全相同。峰的高度不进入总分。

· 坑照常扣分。被搬空的那一格若落在量表的格子上，深学者在那里低于及格线，扣一格。

于是

$$C(\text{深学者}) - C(\text{均浅者}) = 0 - (\text{坑落在格上的个数}) \leqslant 0$$

> 推论三　在任何覆盖型量表上，深学者的得分不高于均浅者，且在坑恰好落进量表格子时严格低于。衡量知识的通行办法不只是读不出这场损失，它给损失的那一方加分。

这条推论解释了三件此前各自被单独解释的事：为什么随时可得的答案在一切标准化评估上表现为进步；为什么「什么都懂一点」在选拔中长期占优而在使用中长期落空；以及为什么每一次「提高考试的覆盖面」的改革，都在无意中加大对深学的惩罚。

这条可以判错，而且很便宜。取任何一份有分格的能力量表，同时取一份不依赖该量表的深度指标（同行指认、独立复现记录、无提示条件下的复推测试），若深学者在覆盖型量表上系统性地高于均浅者，则推论三错。本章没有做这次检验，也不打算把没做的事说成做过的事。

八、第七条关系：依赖消耗着它自己的前提

最后一条接第十章。

第十章的判断是：在鉴别不独立的领域里，选择听谁本身就是那个判断的一次完整行使，依赖并没有省去判断，只是把它移到一个不被记录、也不被自己察觉的位置。它给出的粗筛是：这个领域里有没有人因为判断错了而承担过可被外人读出的后果。

把它接到 ρ 上：行使那一次鉴别，用的正是第一章的复推率——换一个表面不同的情境，能不能推出同一条。于是

鉴别独立性低的领域里，依赖的安全性取决于 ρ；而依赖本身使 λ 上升、ρ 下降。

这是一个自锁：被依赖的那一样越好用，评判它好不好用的能力越少。它与常说的「路径依赖」不同——路径依赖说的是切换成本上升，这里说的是判断切换是否值得的那项能力本身在退，因而连「该不该换」这个问题也逐渐无法被提出。

自锁不产生任何事件。第十章已经写明，鉴别不独立的领域的典型特征就是错误在数年后才显形，且显形时由同一批人认定。

九、七条关系的合用

把七条排在一起，是一条从技术条件到读数失效的链：

编号	关系	出处	结果
R1	$\lambda(t+1) = \lambda + k \cdot a \cdot (1-\lambda)$	第五章	$\lambda \to 1$，a 只改快慢
R2	$\rho = \rho_{max}(1-\lambda)$	第一章 × 第五章	$\rho \to 0$，非折旧而是停产
R3	$\partial y / \partial \lambda = 0$	第五章实测 × 第六章 O*NET	y 不降而 $\rho \to 0$，同一 λ 的两读
R4	Δ = y(封路) − y(不封路)	第五章	唯一读法，且价随 a 涨
R5	$b \propto (1-\lambda)$，κ = 峰高/邻坑深	第四章	均浅：κ 无定义而非低（划界见第六之二节）
R6	C(深) − C(均浅) ⩽ 0	本章新推	覆盖型量表惩罚深学
R7	依赖安全性依赖 ρ，而依赖降低 ρ	第十章 × 第一章	自锁，不产生事件

第二编三章分别是 R1 至 R4 的三个来源；第三编两章处理这条链上「不入账」的两个位置；第四编两章处理链的末端——还剩下什么可以读、以及读它的资质从哪里来。合章会把十项读数逐一接进这七条。

本章不主张的四件事：不主张 a 应当被降低；不主张 λ 高的人能力差（他在 y 上可能确实更好）；不主张覆盖型量表应当被废除（推论三只说明它在这一件事上有系统偏向）；不主张这七条关系已被检验——R1、R3、R4 各有一处实测支持，R2、R5、R6、R7 目前只有形式推导与逐条可判错条件。

第 6 节的那个形状，是这本书唯一一处造出了一个此前没有名字的东西的地方。 其余各处，包括本章的六条关系，都是把已经存在的量重新组织了一遍。这一句写在这里，是为了不让读者把重新组织当成发现。第六之二节把这个新东西与七种最接近的说法逐一分开，并给出使它作废的那次测量——那次测量本书没有做。

第二编

被免费供给的是哪一段

编　序

第一编说明知识不是存量。这一编要回答的是：既然如此，随时可得的答案究竟供给了什么。

答案是：它供给的是重新推那一段，而那一段的发生与否，不进入任何以产物为形式的读数。三章各给这件事一个侧面。

第四章处理代价：学会一件事不是纯粹的增加，是把有限的清晰度从相邻区域搬到目标区域——没有搬走就没有学会。这一章使「不学就不必付代价」这句话第一次成立，也因此使枢纽章第 6 节那个形状成为可能。第五章处理掩护：同一个结果换一条路达成，结果读数对它恒为零，当事人的省力感对它符号为正；这一章是全书唯一一处把剥夺落差的测量整个跑完并如实报告它打掉了自己两条预测的地方。第六章处理删除的形态：被取消的是岔路不是步骤，岗位在、工序在、工作量甚至上升，减少的只是那些「本来可以走另一条」的位置；它带来本书唯一一次在职业数据上的预注册检验。

三章分属机器学习与免疫学、运动医学与舞蹈、法哲学与伦理人类学，写作时互不知情。它们在同一个形状上会合：凡进入产物的都留下，凡不进入产物的都在退，而只有前者被记录。

第四章

借清

学会一件事，是把有限的清晰度从相邻处搬了过来

学科入口　机器学习 × 免疫学 × 神经科学

一、导论：三个不该同时为真的说法

三场争论，分属机器学习、免疫学与神经科学，各自都有厚实的实验支撑，而它们不能同时为真。它们都在回答同一个问题：当一个系统学会一件事，对它没有正在学的那些东西，发生了什么？第一场在机器学习。 一个已经学会区分一批事物的网络，再去学一批新的，往往会把旧的大幅忘掉——这个现象有一个专门的名字，叫灾难性遗忘。主流的处理方式一致地把它当成一个要被消除的故障：给旧任务上重要的那些参数加一道阻力，别让新任务把它们改动太多；或者把旧数据混一点进来一起训。这些做法的共同预设很清楚——忘是学的干扰，理想的系统应当学会新的而不忘掉旧的，两全是可以争取的目标，遗忘是可以工程掉的。争点是：忘，是不是学的一个可消除的副作用？第二场在免疫学。 一个人一生中第一次遇到某种病原，免疫系统会为它建立一套记忆；此后再遇到它的近邻——同一病毒漂移出的变体——身体常常不是重新认一遍，而是拿第一次那套去套。第一次学得越牢，这种「拿旧的去套新的」就越顽固，对真正的新变体反而应答得越差。这件事有一个很老的名字，叫原始抗原原罪，如今在流感与登革热的疫苗与流行病学里是一线课题。它的预设与第一场相反：对近邻的失能，不是干扰，是第一次学会这件事本身的产物——而且它带着一个更硬的性质：早年打下的这个印记，成年后极难改写。争点是：对近邻的失能，是学会的故障，还是学会的一部分？第三场在神经科学。 一个健康的大脑并不把学过的一切都留住；它主动地、有选择地削弱一部分连接，尤其在睡眠中。近十年有两条彼此呼应的主张：一条把睡眠期的整体下调解释为给可塑性付账——清醒时到处增强，睡眠时按比例削弱，好腾出下一天的学习空间；另一条干脆把遗忘本身论证成一个主动的、适应性的过程，它的功能恰恰是让你能一般化，别被无关的细节拖住。它的预设是：忘是一道独立的、主动的工序，服务于学，与「记住什么」是两件事、常常发生在不同的时间窗里。争点是：忘，是不是一个独立于学的、朝相反方向使劲的过程？三场争论互不相

干，却在同一件事上互相顶撞。第一场说：忘是干扰，理想是学而不忘。第二场说：对近邻的失能就是学会本身的产物，学得越透失能越深，且回不来。第三场说：忘是独立的主动过程，为了让学更好用。三条不能同真。若忘只是可消除的干扰（第一场），则免疫学那条「学得越牢、对变体越认不出、且不可逆」的规律该被工程掉，而它偏偏是接种策略里绕不过去的硬骨头。若对近邻的失能是学会不可分的产物（第二场），则第三场把忘说成一道独立工序就说不通——它不是外加的，它是学的另一面。若忘是一道服务于学的独立主动工序（第三场），则第一场想要的「学而不忘」在原则上就该做得到，只要把那道工序关掉；可没有人能关掉它而不同时伤到学。本章认为，僵局来自一个三方共有、却从未被单独说出的前提：

> 「学会一件事」与「对没在学的东西发生的影响」，是两件可以分开谈的事。

第一场想把后者消掉而保住前者（学而不忘）；第三场把后者当成服务于前者的另一道工序（先学后清）；第二场虽然把两者绑在了一起，却仍把它读成「学会 X 这件事，附带地害了 Y」——一个有先后、有主次的因果。三家争的是这两件事之间该是什么关系，没有人问：它们是不是同一件事。本章撤销这个前提，提出：

> 学会，就是把有限的清晰度从相邻的区域搬到目标区域。清晰度不是被学出来的、加出来的，而是被搬过来的——从哪里搬走，那里就相应地变糊。没有这次搬运，就没有这次学会。忘不是学的代价，也不是学之外的一道工序，而是同一个动作从另一端看到的样子。

本章把这件事称为借清：一处变清晰，是向相邻处借来的，那一处就欠着一笔糊。三个案例在这一点上是同一个形状。机器学习里，网络的表示容量是有限的：把一批事物分得更开，用掉的正是本来可以用来分开另一批的那部分容量——旧任务被忘，不是被新任务「干扰」了，是承载它的那点分辨被抽去给了新任务。免疫学里，第一次应答把资源投给了针对首个抗原的那套克隆，针对近邻变体的那套就此长期处在下风——对变体的失能不是「附带损害」，是首次学会这件事用掉的正是本可以留给变体的那份。神经科学里，睡眠期的整体下调不是一道与学习并列的清除工序，而是白天把清晰度搬进某些突触之后、相邻突触相应变糊的那一笔账，被集中到睡眠时结清——所谓主动遗忘，是清晰度被借走这件事，在时间上被推迟显现出来的样子。第二节梳理三条既有路径并指出它们各自停在哪里；第三节定位冲突；第四节界定借清与它的可裁决判据；第五节升成二维辨别表；第六节在十个领域兑现，其中两个迫使本章收窄；第七节给出读数「借清率」的清点手册，并诚实交代本章没有亲自跑一次测量、以及那次测量该怎么做；第八节交代与最近十种说法的分界；第九节说明三家为何各自到不了；第十节给出推论、边界与反噬预言；第十一节处理本判断对它自己是否适用。

二、三条既有路径停在哪里

2.1 干扰论：忘是要被消除的故障

第一条路径把「学新的会忘旧的」理解为一种容量竞争造成的干扰：新任务的梯度覆盖了旧任务用到的权重。它给出的处方一致地朝一个方向使劲——保护旧的。给旧任务重要的参数加约束别让它们被改太多，或者在训练新任务时掺回一部分旧样本，或者为不同任务分配互不重叠的子网络。这些做法背后有一个清晰的、几乎从不被质疑的假设：忘与学是可以分离的，理想的系统学新而不忘旧，两全是一个正当的工程目标。它的长处是可操作：把遗忘量当成一个要往下压的指标。它的盲区在于，它只盯着旧任务的表现被保住了多少，几乎从不问：为保住旧的而加的那道约束，是不是同时压低了对新的学会程度？ 也就是说，它把账记在「旧任务遗忘量」这一栏，而本章将论证真正的守恒量在另一处——被压低的清晰度总量。当被保护的旧任务恰好不是你接下来要用的，这套处方看上去很成功；一旦要用的正好落在被抽走清晰度的近邻上，它保住的那一份就是错的一份。

2.2 印记论：对近邻的失能是学会的产物

第二条路径来自免疫学，也来自它在语言与知觉里的表亲。它的核心判断是：第一次学会一样东西，会在系统里刻下一个模具，此后遇到它的近邻，系统倾向于用旧模具去套，而不是新铸一个——套得越顺手，对真正的新东西就越迟钝。免疫学把这叫原始抗原原罪；语音知觉里有一个几乎同构的现象，婴儿在出生后头一年里把母语的音位边界磨利，与此同时对非母语的对立逐渐听不出来，而且这个窗口一旦关上，成年后极难重开。这条路径比第一条走得深：它承认失能就是学会的产物，不是外来的干扰，还指出了一个第一条完全没有的性质——不可回写。它停在哪里？停在它只在自己的领域里把这件事当成一个具体机制（免疫记忆的克隆竞争、关键期的突触固化），而没有把它抽成一个跨领域的量。于是它说不出：这个失能有多深，是否与首次学会的透彻程度成正比；也说不出，它在什么条件下可回写、什么条件下不可回写。它有一个深刻的观察，缺一把尺子。

2.3 主动遗忘论：忘是一道服务于学的独立工序

第三条路径把遗忘从「被动的损耗」升格为「主动的功能」。它有两支：一支从睡眠切入，主张清醒时的普遍增强不可持续，必须在睡眠期整体按比例下调，才能保住信噪比、腾出容量，这是给可塑性付的账；另一支更直接，主张忘本身是一套受调控的、适应性的机制，它清除的是妨碍一般化的过度具体的痕迹，让系统不至于记死每一个无关细节。这条路径的贡献是把忘讲成了有功能的、可调的。它停在哪里？停在它把忘设想成一道独立于学、朝相反方向

使劲的工序——学在清醒时增强，忘在睡眠时削弱，两者在时间上、机制上被分开研究。正是这种分工，让它无法看见本章要说的那件事：削弱不是与增强并列的另一道操作，而是增强的另一端。把学和忘分给两个时间窗、两套机制去研究，方法上就先把同一个动作切成了两件事。三条路径，一条把忘当故障要消除，一条把失能当产物却缺尺子，一条把忘当独立工序而错切了刀口。三者并排放着，那个共同的、被跳过的东西才显出来——它们都假定学与对别处的影响可以分开谈。

三、冲突定位

把三家的分歧摆到同一张桌上，会看到它们其实在三个不同的位置上争同一件事，因而谁也说服不了谁。第一处分歧在忘的地位：干扰论说它是病理，要治；主动遗忘论说它是功能，要留。这一对是正反的——同一个现象，一个要消除、一个要保护。但它们其实建立在同一个假设上：忘是一个可以单独调节的量，你可以选择压低它（干扰论）或抬高它（主动遗忘论）。能被单独调高或调低，前提就是它独立于学。第二处分歧在失能与学会的因果次序：印记论说失能就是学会的产物，主动遗忘论说忘是学会之后另起的一道工序。这一对是成因错层的——一个把两件事叠成同一个操作，一个把它们排成先后。印记论离本章最近，但它仍留着一点因果的余味：它说「学会 X 导致了对 X 近邻的失能」，一个有主有次的因果链；本章要说的比这更彻底——不是导致，是同一件事。第三处分歧最隐蔽，藏在三家共用的语言里：三家都在谈某个东西的「量」——遗忘量（干扰论要压低它）、印记深度（印记论量它）、下调幅度（主动遗忘论调它）。但没有一家去量清晰度的总量这个跨区域的守恒量。干扰论量的是单个旧任务掉了多少，印记论量的是单个抗原印得多深，主动遗忘论量的是整体削了多少——三个都是局部的量，谁也没有把「学会 X 时目标区涨了多少」与「相邻区落了多少」放在同一个天平的两端。本章的主张正落在这第三处：把清晰度当成一个有限的、守恒的总量，学会就是它在区域之间的搬运。 一旦这样看，三家的分歧就各归各位：干扰论看到的「旧任务被忘」，是清晰度被搬走的一端；印记论看到的「对变体失能」，是清晰度搬走之后那一端变糊、且搬运不可逆的情形；主动遗忘论看到的「睡眠削弱」，是白天完成的搬运在夜里结账。三家看到的是同一笔搬运的不同片段。

四、借清

先把命题写到最紧：

学会一件事，对相邻的东西发生的影响，不是要被消除的干扰（Y_1），也不是学会之后另起的一道服务于一般化的清除工序（Y_2），而是——学会本身就是把有限的

> 清晰度从相邻区域搬到目标区域（Z）。搬走，正是学会这一个动作的另一端；没有搬走，就没有学会。

三个词要界定清楚。清晰度，指一个系统把相近的输入分辨开的能力，落在某一片输入区域上。把两个相近的东西分得更开，那片区域的清晰度就更高。它是本章的核心量，而它的关键性质是有限且守恒：一个系统在给定容量下，能分配到全部输入空间上的总清晰度有一个上界；某处抬高，别处必相应降低。这不是一个玄学假设，它在几个领域里有硬约束支撑——表示容量有限、克隆资源有限、突触强度总和受稳态钳制。相邻，指在系统当前的表示里，与目标区域靠得近、共享大部分承载结构的那些区域。搬运不是从任意远处借，而是从最近的邻居借，因为它们共用同一批可调的部件。这解释了一个反复出现的观察：负迁移总是发生在近处，训练一个精细辨别，受损的恰是与它最像的那些变体，而不是随便哪个无关任务。搬运，指同一个操作的两端：把承载结构调向把目标区域分得更开，这一调，就把它们从把相邻区域分得开的那个位形上挪走了。不是「先学会，再去动相邻」，是一个动作同时抬高一处、压低一处。由此得到一条不含任何情态词的可裁决判据，用来把借清与三家分开：

> 学会 X 之后，去测系统对 X 的一个相邻变体 X′ 的分辨力。若它低于「从未学过 X」的同类系统在 X′ 上的分辨力，且这个降低量随「X 上分辨力的提升量」成比例增长——那么这一笔就是借清。

这条判据把三家逼到互不相容的预测上，因而可裁决。「没泛化」的解释预测：系统对 X′ 的分辨力等于未学过 X 的基线（学 X 对 X′ 无影响，只是没帮上忙）。干扰论预测：对 X′ 会有下降，但下降量与「X 上提升了多少」无关，它取决于覆盖了多少共享参数，是一个独立的量。借清预测：下降量与提升量成比例——因为它们是同一笔搬运的两端，搬得多则一端升得多、另一端也落得多。三条预测在同一个实验里分得开：测「X 上的提升」与「X′ 上的下降」这两个量的相关。相关为零、下降为零，是没泛化；下降存在但与提升不相关，是干扰；下降与提升成比例，是借清。还有第二条判据，用来分辨搬走的那一份能不能要回来：

> 把 X 重新变陌生（让系统卸掉在 X 上多出来的那份分辨），相邻区域 X′ 的分辨力会不会自己回来？回来，是可回写；不回来，是不可回写。

这第二条判据引出了本章的第二根轴，也引出了三家里最深的那一条——印记论——真正抓到而说不清的东西：有些借清是可回写的，卸掉目标处的分辨，相邻处自己就恢复了；有些是不可回写的，你把目标处的分辨卸干净，相邻处也回不来，因为那一次搬运顺便固化了承载结构本身。免疫的印记、母语音位关窗，都是后者。

五、二维辨别表

一根轴不够——只有「借了多少」这一根，借清还只是一个新名字。本章强制加第二根，得到一张二维辨别表。第一根轴：被借走清晰度的那片相邻区域，可回写还是不可回写。 卸掉目标处的分辨，相邻处能自己回来的是可回写；回不来的是不可回写（搬运固化了承载结构）。第二根轴：被压低的那片相邻区域，是不是恰好落在你接下来要用到的东西上。 借清总在近处发生，但近处不一定是要用处；两者的重合与否，决定这一笔借清是无害还是致命。

	压低处不在要用处	压低处正落在要用处
可回写	甲　良性学习：借了清晰度，欠的那笔糊落在用不到的地方，且随时能校回来。日常绝大多数学习属此。	乙　须主动再校准：借得没错，但欠糊落在了要用处，好在可回写。这正是睡眠期再平衡、以及复习相邻变体在做的事——把清晰度搬回去一部分。代价见下。
不可回写	丙　无害的永久特化：这一笔固化了，但欠糊落在再也用不到的地方。专家在其专长里对早已弃用的旧解法失去分辨，属此——固化且要不回来，但不碍事。	丁　印记式失能：搬运固化了承载结构，而欠糊恰好落在你将要用到的近邻上。学得越透，这格里的失能越深，且事后拿不回来。 免疫印记、母语关窗后的非母语对立、以及深度网络对某类近邻的表示坍缩，都落在这里。这是本表的靶格。

四个格里，只有丁格是真正的病，且它有一个反直觉的性质：它不随「学得差」出现，恰恰随「学得好」加深——首次应答越强，印记越牢；母语磨得越利，非母语越聋。它是把一件做得越漂亮、后果越严重的事。 而它最难被看见，因为在目标区域的每一项指标上，它都表现为优异：分得更开、认得更准、反应更快。它的代价全部记在一份没有指标去查的账上——相邻处、要用处、且回不来。乙格看似温和，本章要专门标注它不免费（第六节由睡眠一例逼出）：把清晰度搬回相邻处的那道再校准，本身也是一次搬运，它抬高相邻处的同时会压低别的地方，且这道再校准有容量上限。所以乙格不是「问题已解决」，是「问题被搬到了下一处」。因此本章的表也不含笼统的价值排序：丙格不是病，丁格才是。 病不在借清——借清是学会的必然形态，不借就不学；病在于让不可回写的那一笔恰好落在要用处，还以为自己只是学得好。

六、逐领域兑现

本节把借清与二维辨别表拿到十个领域里过一遍。凡是能落格的都标出格；其中两处反过来迫使本章收窄命题，单独标注。

6.1 三个来源领域

机器学习（干扰的另一读法）。 一个固定容量的网络先学 A 后学 B，A 掉了。干扰论说 B 的梯度覆盖了 A 的权重。借清的读法是：把 B 分得更开，用掉的正是本可以把 A 分开的那份容量——A 的下降量应与 B 的学会程度成比例，而不是与「覆盖了多少共享权重」独立地相关。给旧任务加约束的那些方法之所以有效又有限，正因为它们把清晰度钉在旧任务上，于是新任务只能借到更少——它们没有取消守恒，只是选择了把守恒的账压在新任务那一栏。表示坍缩（网络把一批本应分开的近邻映到几乎同一处）是丁格：分得越开的那一批越清晰，被坍到一起的近邻越回不来。免疫学（印记）。 首次感染建立的记忆越强，对后续漂移变体的应答越受它牵制——这是丁格的典型：不可回写，且落在要用处（下一个流行的正是变体）。借清给它添了一条可量的预测：失能的深度与首次应答的强度成正比，而不是与「间隔了多久」独立地相关。这一领域反过来给本章加了一道约束：接种研究发现，换佐剂、拉开间隔、用异源加强，可以部分绕开印记，唤起针对变体的新应答。这说明「不可回写」不是绝对的——它要收窄为在不引入新的、足够强的再刺激的条件下不可回写。异源加强恰恰是引入了一次新的高强度搬运，把清晰度重新分配了一遍。收窄之后，丁格的判据更准了：它测的是「在常规再暴露下」是否回写，而不是「在任何干预下」。神经科学（睡眠与主动遗忘）。 睡眠期的整体下调，本章读成白天搬运的集中结账：清醒时把清晰度搬进某些突触，相邻突触相应欠糊，夜里按比例削弱以恢复总量的可分配性。主动遗忘则是「清晰度被借走」这件事在时间上被推迟显现。这一领域也反过来给本章加了一道约束：睡眠期的再平衡不是无代价的整理，它有容量上限，且它自己也是一次搬运——把某些突触调回去，会动到另一些。这正是第五节乙格标注「不免费」的来源：可回写不等于免费回写，再校准本身在借清。由此得一条预测：睡眠剥夺之后受损最重的，应当是那些「白天借得最多、最需要夜里搬回去」的相邻区域，而不是均匀受损。

6.2 另外七处兑现

母语音位（丁格的现成实例）。 婴儿在头一年里磨利母语的音位边界，同时对非母语的对立逐渐失去分辨，窗口关上后极难重开。这是丁格：母语学得越透（清晰度抬得越高），非母语的近邻对立越听不出（相邻处欠糊），且不可回写（关键期固化）。借清给它添的读数是：母语范畴的锐化程度应与非母语对立的分辨损失成比例，两者是同一次搬运。知觉学习的特异性。 训练一个特定朝向、特定位置、特定纹理的辨别，练成之后换个朝向或位置，成绩掉回未

训练水平——领域内把这读成「学习发生在很低的层级，所以不泛化」。借清给出一个更强的预测：那不只是「没泛化」，是负迁移——被训练的那一档变清晰，其最近的未训练档应低于基线，且下降量与训练档的提升成比例。这是第四节那条判据在知觉学习里的直接落点，也是本章最便宜的一个可做实验（第七节）。本族面孔效应。 对本族面孔辨认得又快又准的人，对他族面孔的分辨常常明显更差——通常读成「接触多少」的经验差异。借清的读法：把本族面孔的细微差别分得越开，用掉的正是本可以分开他族面孔的那份表示，是一次借清；它多半落在丁格（早年固化、日后要用他族时才发现回不来）。专家的僵化。 一个领域的专家，面对偏离其惯用套路的变体题，有时比新手还差——领域内把它归给「被熟悉解法吸引」的注意偏向。借清给出一个可与之分开的预测：即使排除注意偏向（例如用不需要主动搜索解法的判别任务），专家在其专长的近邻上仍应有分辨力的结构性下降，因为那份清晰度确实被搬去了核心。落格：多为丙（固化但那些旧变体已不再用到）；当行业变化把「近邻」重新变成「要用处」时，滑向丁。组织的核心能力。 一个组织把某项能力磨到极致，同时对与之相邻的做法越来越不敏感——管理学把「核心能力」与「核心刚性」并称，说它们是一枚硬币的两面。借清给这枚硬币一个机制与一个方向：刚性不是随机长出来的，它恰好指向核心能力最强的那片相邻——因为清晰度正是从那里被搬走的。这解释了为什么颠覆一个强组织的，往往不是远处的对手，而是它核心能力的近邻变体。科学范式。 学会一个范式，会让你对它无法容纳的反常视而不见——不是不看，是分辨不出「这是个反常」还是「这是个噪声」。借清把这从一句社会学观察落到表示层：把范式内的现象分得越清，用掉的正是本可以把范式外的反常分辨出来的那份，两者是同一次搬运。它多在丁格，因为一个人越是这个范式的行家，反常越落在他的近邻、且他越回不去。搜索与推荐系统。 一个模型在高频查询上拟合得越好，对长尾近邻的分辨常常越差——工程上读成「数据不均衡」。借清的读法：把高频区分得更开，用掉的是本可以分开长尾近邻的容量；加大模型能缓解，是因为把守恒的上界抬高了，但不取消守恒——这正是下一段那条兑现要收窄的地方。收窄之三：守恒是有条件的。 上面几处都假定容量固定。若容量可加（更大的模型、更多的克隆、更长的关键期），清晰度的上界随之抬高，某一处的抬清就不必从最近邻借那么多。这迫使本章明写一条边界：借清是「给定容量」下的守恒，不是绝对守恒。 加容量能稀释借清，却不能取消它——因为任何有限容量下总量仍有上界，被推到的只是更大的尺度（第十节反噬预言正据此展开）。

七、读数与它的清点手册

本章的读数是借清率。它在正文、判据、赌注三处必须是同一个定义、同一份编码规则，否则后面的一切都是装饰。先把定义写死：

借清率 = 学会 X 时，相邻区域清晰度的下降量 ÷ 目标区域 X 清晰度的上升量。两个量都相对于「未学过 X 的同类系统」这条基线来测。借清率接近 1，是紧的零和搬运；显著小于 1，是被容量或再校准稀释过的搬运；等于 0，则本章在该处不成立。

清点手册（七栏，任何人复算须逐栏对齐）：一，目标区域 X——要被抬清的那一小片输入。二，相邻变体 X′——在系统当前表示里与 X 最近、共享承载结构的那一片；须事先声明「近」怎么量，不能事后挑。三，清晰度的操作定义——把相近输入分辨开的能力，用同一把尺子（如同一个辨别任务的正确率或可分度）在 X 与 X′ 上分别测。四，基线——未学过 X 的同类系统，在 X 与 X′ 上的分辨力。五，上升量与下降量——学 X 前后，X 处相对基线升多少、X′ 处相对基线降多少。六，可回写判定——卸掉 X 上多出来的那份分辨后，X′ 是否回到基线。七，落格——由「可回写与否 × X′ 是否要用处」定甲乙丙丁。

7.1 本章没有亲自跑的那次测量，以及它该怎么做

诚实地说：本章没有亲自在一个系统上跑出一条借清率。原因不是它难做——恰恰相反，它是本章所有主张里最便宜、最可复现的一条——而是一次像样的测量需要成规模的训练与统计，超出了本章的执行范围。本章因此把「补上这次测量」当作命题的一层，而不是回避它。该测量的最省钱做法写在这里，任何有一台机器的人都能做：取一个固定容量的小模型和一族两两相近的辨别任务（例如一组只在角度上微差的纹理，或一组只在一个音位维度上微差的音）。第一步，测未训练的模型在全族上的分辨力，得到基线。第二步，只把其中一档 X 训到远高于基线。第三步，测它在最近邻 X′ 上的分辨力。三种理论在这一步分岔：没泛化论预测 X′ 等于基线；干扰论预测 X′ 低于基线、但降幅与「X 提升了多少」不相关；借清预测 X′ 低于基线，且降幅与 X 的提升量成正比。 第四步，把 X 上多出来的分辨卸掉，看 X′ 回不回基线——回，甲/乙；不回，丙/丁。整套实验不需要任何新设备，只需要一次对照训练与两次分辨力测量。它是本章最容易被证伪的地方，也是本章最愿意押上去的地方。此外两条本章同样未跑、但值得指名：一，免疫侧——首次应答强度与后续变体应答弱化之间是否成比例（须在同一队列里控制间隔）；二，睡眠侧——睡眠剥夺后受损最重的是否正是「白天借得最多」的相邻区域。两条都要专门的队列或实验平台，本章的执行力到不了，均标未跑。

八、与最近几种说法的分界

与灾难性遗忘（French 1999；Kirkpatrick 等 2017）。 这是最近的一位，也是最容易被误当成同一件事的。分界在机制而不在现象：那一族把遗忘当作可消除的干扰，其代表性处

方（给旧任务重要参数加约束）预设了「学新而不忘旧」的两全是可达的。判决性对照：本章预测守恒——任何把旧任务遗忘量压低的方法，必以降低对新任务的学会程度为代价，两者之和有下界；那一族预测通过更好的约束可以逼近两全。同一批实验里，测「旧任务保住量」与「新任务学会量」之和随方法变化是否守恒，即可判。与稳定—可塑性困境（Grossberg 1987）。 那一族把学习刻画为两个相反需求之间的平衡——太稳则学不动，太塑则记不住。分界：平衡是两根反向的力在一根轴上拉锯；借清是同一个动作的两端，不是两个力的折中。分岔预测：若是平衡，则存在一个「两不相欠」的中点；若是借清，则不存在这样的中点，任何一次学会都在某处欠糊，只是欠在哪、可否回写不同。与原始抗原原罪与免疫印记（Francis 1960；及其当代流感／登革热研究）。 免疫学这一位本章请进来，并给它添了它自己没有的东西：一把跨领域的尺子与一个方向。它有「学得越牢、对变体越钝」的深刻观察，缺「深度与首次强度成正比」这条可量预测，也缺「在何种再刺激下可回写」这条边界。判决性对照：本章预测失能深度随首次应答强度单调上升，并预测异源加强能部分回写——两条都可在接种队列里检验，而单靠印记这个概念推不出比例关系。与主动遗忘与睡眠稳态（Richards & Frankland 2017；Tononi & Cirelli 2014）。 那一族把忘讲成独立于学、服务于一般化的主动工序。分界：本章说忘不是独立工序，是学的另一端。判决性对照：若忘是独立可调的，则应能找到「大量遗忘而几乎没有对应学习」的常态；本章预测遗忘量与新学习量成比例耦合——夜里削得多的地方，正是白天学得多的地方的相邻。两者在「遗忘是否可与学习解耦」上给出相反预测。与感知磁体效应与关键期（Kuhl 1991；Werker & Tees 1984）。 那一族描述了母语范畴锐化与非母语对立丧失的共时性，并称之为神经承诺。分界：它描述了共时，本章给出比例与守恒——锐化程度应与丧失程度成正比，二者是同一次搬运，而非两个并行的发育事件。与知觉学习的特异性（Karni & Sagi 1991；及后续）。 那一族把「练成一档、换档掉回」读成学习发生在低层、故不泛化。分界：本章预测的不是「没泛化」（X′ 等于基线），而是「负迁移」（X′ 低于基线）且与提升成比例。同一实验测 X′ 相对基线的正负，即可把两者分开—— 这是本章与最主流解读之间最干净的一条裁决线。与核心能力／核心刚性（Leonard-Barton 1992）。 那一位把两者并称为硬币两面，机制上归给投资与路径依赖。分界：本章给刚性一个方向——它恰好指向核心能力最强的相邻，因为清晰度是从那里被搬走的。判决性对照：投资锁定预测刚性分布与「投入多少」相关；本章预测刚性最深处是核心能力的近邻，而非投入最多处（两者常常不重合）。与专长逆转与定势效应（Chase & Simon 1973；Bilalic等 2008）。 定势效应把专家的僵化归给「熟悉解法的激活抢先」，是一个注意与激活层面的机制。分界：本章的机制在表示层。判决性对照：在排除主动搜索、只做被动判别的任务里，定势效应应大幅减弱，而本章预测的近邻分辨力下降仍在——两者在「去掉搜索

之后还剩不剩」上分岔。与范式与不可通约（Kuhn 1962）。那一位在信念与共同体层面讲范式如何遮蔽反常。分界：本章把它落到表示层并给出可测读数——反常之所以看不见，是它落在被搬空清晰度的近邻。判决性对照：范式说预测遮蔽随「信奉程度」变化；本章预测遮蔽最深处是范式核心现象的近邻，可由分辨任务而非态度问卷来测。与没有免费午餐定理（Wolpert & Macready 1997）。这是最抽象的一位：没有一个算法对所有问题都好，跨越全部问题分布时优势必然抵消。分界：那是一条关于全体问题的全局守恒，不给机制、不给可测的局部读数；借清是单次学习内部的局部守恒，指名了搬运在相邻处发生、给出借清率这把尺子。判决性对照：没有免费午餐不预测「学 X 会具体压低哪一个 X′」；借清预测得出，并可测。二者可同真而互不覆盖——一个说全局抵消，一个说局部搬运。

九、三家各自为何到不了这一条

机器学习到不了，因为它的目标函数只给目标任务记分。训练的损失只盯着当前任务，相邻区域掉了多少不在被优化的量里，于是那份下降被当成「未被测量的负迁移」，随手记进遗忘那一栏。加上容量可加这个事实，整件事就被读成「容量不够，加参数即可」——一个把守恒误当成短缺的读法。一个只给一端记分的领域，会把另一端的账当成噪声。免疫学到不了，因为它一次只看一个个体对一串抗原的应答，没有「清晰度总量」这个跨抗原的量。它能看见「对首个抗原强、对变体弱」，却没有一个天平把这两端称在一起，于是它有印记这个深刻的观察，却量不出比例、划不出可回写的边界。一个没有守恒量的领域，看得见此消彼长，说不出此消与彼长是同一笔。神经科学到不了，因为它把学与忘分给了不同的时间窗和不同的机制去研究。学在清醒时、由增强类过程负责；忘在睡眠时、由削弱类过程负责。这个分工在方法上极有成效，却在起点上就把同一个动作切成了两件事——你没法在把它们分开测量的框架里，看见它们是一笔搬运的两端。一个按机制分工切开研究对象的领域，看不见被它切开的那条缝原本是一条线。三家的界限都不是能力问题：一个只给一端记分，一个没有守恒量，一个在方法上先把一分成了二。三者并排放着，那个共同的量——有限而守恒的清晰度——才显出来。

十、推论、适用边界与反噬预言

10.1 推论

推论一。任何「学而不忘」的目标，若不声明容量与相邻结构，都不是一个可达的目标，而是一次把账挪到别处的操作。它多半会以「旧任务保住了」的名义，把清晰度从「下一个要用的近邻」上悄悄借走。推论二。对丁格的对象，最有价值的不是把目标学得更透，而是先判

「要用处」落在哪里，再决定往哪个方向借清。把清晰度从「永不再用的旧变体」上借（丙格），是安全的；从「下一个流行的变体、下一代的对手、下一个反常」上借（丁格），是把一件越做越漂亮的事做成越做越危险的事。推论三。 一个系统若在其核心能力上不断精进而对近邻越来越钝，通常不该被诊断为「不够努力」或「数据不够」，而该先量它的借清率与落格。若它落在丁格，加码精进只会加深失能——努力在这里是反向的。

10.2 适用边界

其一，本章只在容量有约束的系统里成立。容量近乎无限的地方，借清被稀释到可忽略，本章无话可说（但见反噬预言：无限容量本身是个幻觉）。其二，「相邻」依赖系统当前的表示怎么组织，跨系统比较借清率时必须先声明「近」的口径（清点手册第二栏），否则不可比。其三，本章不含笼统的价值判断：借清不是缺陷，它是学会的必然形态；丙格不是病，只有「让不可回写的一笔落在要用处、还以为自己只是学得好」才是。

10.3 反噬预言

按本章设计干预，最自然的做法是要求「学而不挤占」——设计一个学会新东西而不压低任何近邻的系统。这会以一种可以精确说出的方式落空：达成它最省事的办法，是把容量做大到挤占可忽略。而这只是把零和推到了更大的尺度——在更大的容量上，总量仍有上界，搬运仍在发生，只是每一笔小到当下测不出。等到问题在那个更大尺度上重新逼近容量边界，借清会原样回来，且因为一路都被读成「已经解决」，它回来时更难被看见。这正是没有免费午餐定理在单次学习内部的回声：你能把守恒推到更远，不能取消它。因此本章不推荐把「零挤占」做成一个达标项——它可被容量注水刷过，而注水恰恰喂大了下一次丁格。 本章只推荐一件不能被注水刷过的事：在系统精进其核心之后，去测它在核心的最近邻上、且是下一步要用到的那些近邻上，分辨力是升了还是降了。 这个动作测的是要用处的相邻，而要用处由未来定义、无法被当下的容量预先填满——要测就得真的等到用它的时候去测。

十一、本判断对它自己适用吗

本章也是一次学会。它把「学会＝清晰度的搬运」这一小片分得很开——为此，它必然从相邻的直觉那里借走了清晰度。最近的那片相邻，是「学会就是纯粹的积累，多学一样不必少一样」这个几乎人人默认的想法。本章预测：读完本章的人，会在一段时间里对「确实存在纯增量学习」的反例变得不敏感——会倾向于把每一个「学 A 不影响 B」的例子，都硬读成「只是容量够大、借得太少而已」。这正是本章该被证伪的地方：如果有一类学习，可反复验证它在给定容量下学会新的而任何近邻都不掉，那么本章就是把一次借清（把「积累论」搬糊）误

当成了发现。再问一层：本章落哪一格？一篇文章发表即完成，读者读完就放下，它自身可以被无代价地整个重写——就这一面看，它落甲格。但产出它的那套工序不然：它每处理一个题，就把某一小片直觉抬清一次，而每一次抬清都在相邻的直觉上欠一笔糊；做得越多，越可能在自己最擅长的那类判断的近邻上失去分辨——这是丁格的倾向。所以真正需要按本章自己的读数去查借清率的，不是这一篇，是产出这一篇的那套做法：它有没有在把「撤掉一个大家都默认、却没人单独说出的前提」这一手磨得越来越利的同时，对「有些前提本就不该撤」这个最近的邻居，越来越听不见。 本章没有能力量自己的这条借清率，只能把这个洞明写在这里。

结　论

本章提出并论证：学会一件事，对相邻东西的影响，不是要消除的干扰，也不是学之后另起的清除工序，而是同一个动作的另一端——学会就是把有限的清晰度从相邻区域搬到目标区域，没有搬走就没有学会。 本章称之为借清，读数是借清率：相邻区域的清晰度下降量与目标区域上升量之比。二维辨别表由两根轴张开——被借处可否回写，被借处是否落在要用处——指认唯一的靶格丁：不可回写，且落在要用处。它的反直觉之处在于随「学得好」加深而非随「学得差」出现，且在目标区域的每一项指标上都伪装成优异。免疫印记、母语关窗后的非母语对立、深度网络的表示坍缩，都落在这一格。十个领域的兑现中，免疫接种与睡眠两处反过来迫使命题收窄：不可回写要限定在「不引入新的高强度再刺激」之下，把清晰度搬回去的再校准本身也在借清、并不免费；容量可加则说明借清是「给定容量」下的守恒，不是绝对守恒——但加容量只能把守恒推到更大尺度，不能取消它。本章诚实交代：核心读数借清率，本章没有亲自在系统上跑出来一条；那次测量是本章所有主张里最便宜、最可复现、也最愿意押上的一条，第七节给了任何有一台机器的人都能照做的最省钱协议。最后给一条写死日期的赌注：

> 到 2033 年 12 月 31 日，若在机器学习、知觉学习、免疫学三者中的任何一个领域里，仍没有任何一项研究把「借清率」——学会某物时相邻处清晰度的下降量与该物清晰度上升量之比——作为独立于「容量不足」与「干扰覆盖」的变量单独测量，或测量之后未能重现「下降量随上升量成比例」的守恒关系，则本章的核心主张作废，不再申辩。

注　释

1. 本章所称「清晰度」是一个刻意跨领域的操作性名词，指把相近输入分辨开的能力，落在某一片输入区域上；在不同领域它由不同的量实现（分类可分度、克隆频率、辨别正确率、突触权重的相对配置）。跨领域使用的代价是精度的损失，本章用它换取的是「守恒」这一条能被同一把尺子检验的性质。2. 「相邻」始终指系统当前表示里的近，而非物理或语义上的近。同一对输入，在一个表示里相邻、在另一个里不相邻——这是借清率必须先声明「近」的口径的原因。3. 第六节的「收窄之二（睡眠）」与「收窄之三（容量）」是本章自己加在自己身上的约束，不是对既有文献的复述；把它们写出来是为了标明命题的适用边界，而非扩大它。4. 免疫、母语、知觉学习三处的转述依据公开文献的核心结论，未逐条核对原著页码；本章对它们的机制解释（读成同一次搬运的两端）由本章负责，不应归给原作者。5. 第七节所述测量本章未执行；「未执行」不等于「不可执行」，恰恰相反，它可执行且便宜，本章把补做它列为命题的一层。

人机分工声明

本章由王德生与 Claude 共同署名。三个来源领域（机器学习、免疫学、神经科学）由 Claude 选定，选源刻意跨开了三门互不通用的学问，使它们对同一问题的回答互相排斥；承重判断的形成、二维辨别表、十领域兑现（含免疫、睡眠、容量三处对本章的收窄）、借清率的定义与清点手册、以及全部文字，由 Claude 生成；所依据的方法论框架与发表裁定由王德生给出。第七节所述的核心测量（借清率的实测）本章未执行，已在正文与注释中明写，并给出了可复现的最省钱协议；免疫侧与睡眠侧的两条印证性测量同样未执行、已指名。三个来源领域的转述依据公开文献的核心结论，未逐条核对原著页码；本章对它们的机制解释由本章负责。字数　纯汉字约 1.4 万　版本　v1.0 · 2026-08-11 三边对「一个系统学会一件事时，对它没在学的东西发生了什么」给出互相排斥的答案：机器学习把它当作要消除的干扰（理想是学新而不忘旧）；免疫学把它当作学会本身的产物，第一次学得越牢、对近邻变体越认不出，而且极难改写；神经科学把它当作一道独立的、服务于一般化的主动清除工序。三家争的其实是同一件事该由谁裁、放哪儿、有多少，而它们共享了一个从未被单独说出的前提——学与对别处的影响是两件可分开的事。

第五章

借道

论一个结果被换一条路达成之后，为什么读数与感觉都不报警

学科入口　运动控制 × 古典舞技术与舞蹈医学 × 学习心理学

一、三个不相干的现场

第一个现场在芭蕾教室。一个十四岁的学生站在一位，双脚外开接近一百八十度。老师从侧面看了三秒，说：收回去，你在用脚踝借。学生收回去，外开掉到一百二十度，动作立刻变得不好看。她不明白：既然站出来的样子是对的，为什么要退回一个更难看的位置。第二个现场在运动医学的实验室。一名前交叉韧带重建术后的运动员做单腿跳远，患侧与健侧的距离比是百分之九十七，达到了几乎所有回归赛场清单的标准。同一次测量里，患侧膝关节在蹬伸阶段做的功只有健侧的百分之六十九，缺掉的那部分由髋关节与躯干补上了。[1] 跳出去的距离不知道这件事。这名运动员本人也不知道。第三个现场在心理学实验室，跟身体没有关系。学生用两种方式复习同一份材料：反复重读，或者不看材料自己回忆。重读那一组当场感觉更流畅，对自己掌握程度的判断更高，而一周后的保持成绩更差。这个反差被反复复制了几十年，稳定到已经进了教科书。[2]三个现场看上去互不相干。把它们摆在一起，会浮出一个三方都没有单独提出的问题：当一个结果可以由几条不同的内部路径达成时，究竟是什么在决定用哪一条？这个问题听上去像运动科学的家务事。它不是。任何一个只看结果、不问过程的评价系统，只要它面对的是一个会学习的执行者，就会遇到同一件事。本章要论证的是：在这类系统里，路径的分布不会停在原处，它会朝一个固定方向移动；而移动的全部代价，落在一个从来没有人测量的量上。本章的走法是这样。第二节列出三家各自说到哪一步。第三节指出三家共有而无人检查的那一条前提。第四、五节给出机制的两半。第六到第十节给出判断、两根轴与唯一的测量操作。第十一、十二节把这套测量在一个可以真封路的系统里整个跑一遍，并如实报告它打掉了本章四条预测里的两条。第十三到十八节与最容易混淆的几种说法分界，其余各节给出兑现、让步、失败条件、反噬与本章自己的位置。

二、三家各自说到了哪一步

2.1 运动控制：冗余是资源，不纠正才是最优的

人体的自由度远多于任务所需。伯恩斯坦留下的那句话常被引用：熟练的动作是没有重复的重复。半个世纪后，这句话得到了一个数学上干净的解释。不受控流形方法把动作的变异分成两类：改变任务结果的，和不改变任务结果的。测量反复表明，熟练者在后一类方向上的变异更大，不是更小。[3] 随机最优反馈控制给出了理由：在有噪声的系统里，只纠正会影响任务目标的偏差是最优策略，纠正其余偏差要付出控制成本而没有收益。这条被称为最小干预原理。[4]这条原理有一个直接推论，提出者本人写得很清楚：噪声在所有方向上扰动系统，因此变异会在与任务无关的方向上累积。在这个框架里，多条路径能达成同一结果是好事。它是应对扰动的余量，是迁移的基础，是训练应当扩大而不是压缩的东西。这条思路在教学法上有明确落地：放宽约束让学习者自己找解，甚至主动往练习里加噪声，让人每一次都做得不太一样。这一家的主张压成一句：不影响结果的差异不必管，管它反而有害。

2.2 古典舞：外形是标准，路径是内容

芭蕾的外开是一个极端例子。理想外开要求双脚成一百八十度，而人的髋关节被动外旋一般在四十五度上下，两条腿加起来九十度左右。文献估计外开总量里髋关节贡献大约百分之五十到七十，其余来自膝、小腿、踝与足。绝大多数舞者达不到解剖学上的理想值。于是有了一个专门的说法：强迫外开。做法是把差额从下面补上——胫骨相对股骨外旋、距下与中足过度旋前、腰椎前凸增加。从外面看，这个位置与一个天生髋外旋充裕的舞者站出来的位置几乎一样。一项比较自报有伤史与无伤史两组芭蕾舞者的研究，量的不是外开角度本身，而是「站出来的外开」减去「被动髋外旋能给的外开」这个差额：有伤组平均二十五点四度，无伤组四点七度，差异显著。[5] 后续的三维足部模型研究描述了这条补偿的解剖路径：后足外翻、中足外展、内侧纵弓下降。一份系统综述则给出一个诚实的限制：强迫外开与损伤之间的因果链比横断面相关复杂得多，现有证据不足以支持简单的单因果说法。[6]无论因果强度如何，芭蕾教学法在实践上早就给了自己的答案，而且是一条硬规矩：从髋部转，不许从膝、踝、足借。 这条规矩的特别之处在于，它规定的不是结果，是路径。一个学生可以完全达到外形标准而被判为做错了。这一家的主张压成一句：同一个外形由不同路径达成不是等价的，替代路径会固化并致病，所以必须规定路径。

2.3 学习心理学：读数会骗人，而且朝一个固定方向骗

合意难度说的是：一些使当下表现变差的练习安排——间隔、交错、变异、提取而非重读——会使长期保持变好；反过来，使当下表现变好的安排常常损害长期保持。更要紧的是第二层。学习者对自己掌握程度的判断，主要依据加工的流畅感。[7] 流畅感是省力的直接读数。

于是产生一个方向固定的偏差：越省力的学习方式，人越觉得有效，而它往往越无效。[2] 这个偏差不因为被告知而消失，也不因为经验而减弱。运动学习里有一个平行的区分：结果知识与表现知识。前者告诉执行者做成了什么，后者告诉他是怎么做的。日常生活中绝大部分反馈是结果知识——球进了没有，跳多远，音准不准。表现知识需要外部设备或一位懂行的观察者，成本高得多，因此稀缺。这一家的主张压成一句：当下表现与长期能力可以背离，而主观掌握感由省力驱动，因此系统性地站在错的一边。

2.4 三家为什么不能同时对

第一家与第二家不能同真。 最小干预原理说，不影响任务结果的偏差不该被纠正，纠正要付控制成本而无收益。芭蕾教学法说，恰恰是这些不影响外形的偏差必须被纠正，因为它们是后来那些问题的来源。一方判为最优的策略，另一方判为病根。这不是侧重不同，是同一个动作两种相反的处置。第一家与第三家不能同真。 第一家把变异的存在归给控制策略在当下的最优性——它是无历史的、随时可重新求解的。第三家把表现与能力的背离归给学习历史的累积与元认知的系统偏差——它是有历史的、不可当场撤销的。第一家还把省力当作效率的证据，第三家把省力感当作最不可信的读数。第二家与第三家不能同真，而且这一对最尖锐。 整个芭蕾教学法建立在一个假设上：一位有经验的老师能从外面看出学生从哪里借的力。第三家的证据说，观察者与执行者会被同一个线索欺骗——看上去省力就是做得好。而这恰恰是古典舞审美的第一原则：举重若轻。舞蹈把借力最强的信号，定义成了美。三家都不肯让步，三家各自都有对手消化不了的硬证据。这种局面通常意味着，问题不在这三条之间，而在它们共有的某个东西上。

三、三家共有的那一条，没有人查过

把三家的语言并排看，会发现一个词从来没有被追问过：路径的可用性。第一家默认可用性是一个稳定的集合。最优反馈控制在每一个时刻求解一次控制律，它假定可选方案在那里等着被选。今天能用的，明天也能用；用了甲，不影响乙。冗余之所以是韧性，正因为这一点——扰动来了，还有别的路可走。第二家默认可用性是一个可以被教学纠正的状态。老师看出你在借，你就改回来。改不回来算态度问题或柔韧度问题，不算另一类问题。第三家根本不谈路径。它谈练习安排、提取难度、元认知判断，它的自变量是外部条件——间隔多长、要不要看着材料、什么时候测。个体在同一份练习安排下内部怎么分配，不在它的视野里。三家共有的前提可以写成一句：路径的可用性是一个存量，它由体质、训练与教学决定，不由「最近用了哪一条」决定。这条前提没有被任何一家检验过，因为在任何一家的框架里，它都不是一个变量。而它是错的，至少在有学习的系统里是错的。使用依赖的可塑性是神经科学里最不新鲜

的事实之一：一条通路被使用会被强化，长期不被使用会退化。把这条众所周知的事实与上面三家放在一起，得到的不是三家里任何一家的结论，而是：路径的可用性是一个流量，它的方向由使用决定；而使用由什么决定，正是三家吵而未决的那件事。余下的论证分两步。第四节说明，在结果等价的那些方向上，任务不提供任何把动作拉回原处的力。第五节说明，在同一批方向上，另有东西在推。

四、没有回复力的方向

先把话说准。最小干预原理不是一条关于人应该怎么做的建议。它是一条关于最优控制器长什么样的推论：在有信号依赖噪声的系统里，纠正与任务目标无关的偏差要消耗控制信号，而控制信号本身带噪声。结果是，纠正它们会使任务表现变差，不是变好。这条推论的经验证据很扎实。不受控流形的测量在指力、姿势、投掷、行走等多种任务里得到同一个形状：影响结果的方向上变异被压得很小，不影响结果的方向上变异大得多，而且熟练者的这个对比更明显，不是更弱。现在把它翻译成本章需要的语言。一个负反馈系统的稳定，来自偏离时被推回去的那个力。在影响结果的方向上，这个力存在，由任务本身提供——手偏了，球没进，下一次就得纠。在不影响结果的方向上，这个力等于零。 不是很小，是零，而且是被证明为最优的零。一个没有回复力的方向意味着什么？意味着这个方向上的位置不由任务决定。系统停在哪里，取决于别的东西。最优控制那一支指出的后果是：噪声会使变异在这些方向上累积。这是一个关于方差的说法——分布会变宽。本章要说的是关于均值的说法：如果在这些方向上除了噪声之外还存在一个有方向的力，那么分布不只会变宽，它的中心会移动。而这个移动，任务永远不会把它拉回来，因为任务对它无感。

结果等价的那些自由度，是一片没有回程票的地形。

那里有没有一个有方向的力？有，而且不止一个。

五、那个方向上有一个坡

至少三种成本在这些方向上提供梯度。它们的共同点是：都与任务结果无关，因此都不被任务的反馈回路管；都是单调的，因此提供的是坡而不是井。

5.1 代谢成本

人会挑省能量的动作方式，这不新鲜。新鲜的是它发生得多快。一项研究用外骨骼人为改变了走路时步频与能耗的关系，把能耗最低点挪到平常偏好之外。受试者在几分钟之内就把步频挪了过去，收敛到新的能耗最优点。[8] 他们不需要几年，不需要教学，也不需要知道发生

了什么。同一条线后来的工作进一步表明，这个优化过程可以是内隐的——人在没有意识到的情况下完成它。[9]这条证据同时给了三样东西：坡是真的；坡在个体的时间尺度上起作用，不是演化尺度；并且它不经过意识。同一批工作也交出一条限制，本章照实转述：自然状态下的步态变异未必足以启动这个优化，探索似乎需要被推动。[10] 这条限制会在第二十一节被写成一条可以推翻本章的判据。

5.2 疼痛与不适

疼痛适应的一条理论，核心是一句：疼痛导致活动在肌肉内部与肌肉之间的重新分配——一些运动单元降低放电率或被撤出，另一些此前不活跃的被招募进来。[11] 作者明确指出，这种重分配短期是保护性的，长期可能有害。要注意这里的量级。疼痛不需要达到临床水平。任何一点持续的不适——一处发紧、一个角度上的轻微牵拉感——都足以在没有回复力的方向上提供一个稳定的偏置。而它同样是内隐的：没有人在做动作时决定要绕开那个角度。

5.3 注意力与认知负荷

第三个坡最弱，覆盖面最广。人会回避需要更多认知控制的选项，这在实验室里被反复测到。[12] 一条需要持续监控的路径，与一条可以放手不管的路径，如果结果一样，后者会被选中。熟练化本身就在制造这个坡。自动化阶段的定义就是所需注意力的下降。在结果等价的方向上，自动化不区分「变熟练」与「变省事」，因为在那些方向上，这两件事没有外部读数可以分开。

5.4 三个坡合起来

这三种成本的方向未必一致——最省能量的路径未必最不疼。但它们有一个共同性质：都不是任务的函数。 任务不知道你花了多少能量、疼不疼、费不费神；任务只知道结果。于是在结果等价的自由度上，出现了一个特定的形状：任务提供零回复力，成本提供非零梯度。这个形状唯一可能的长期行为是移动，而移动的方向由成本定，与任务无关。

六、借道

现在可以把判断写出来。

> 一个结果由替代路径达成，既不是韧性的证明，也不是一次可被纠正的技术错误，而是一次不需要偿还、却会降低下一次不借的可能性的占用。

本章把这个过程称为借道。这个词是有意选的。它在两个领域里已经是行话——舞蹈教师说借力，康复治疗师说代偿——而它带的那点日常语义正好可以拿来做对照：借东西通常要

还，还了之后原状恢复。借道不是这样。借道不用还，而这正是它的危险所在。代价不在还款，而在于：那条被绕开的路，在没有人看的情况下自己塌了一段。机制是三步：第一步 · 触发。 在结果等价的方向上出现一个成本差。它可以来自解剖限制（髋外旋不够），来自一次损伤（膝关节疼），来自疲劳，来自一件工具（有导航就不用记路），也可以什么都不为——只是那条路稍微省一点力。触发不需要失败，也不需要疼痛达到临床水平。 这一点后面会反复用到。第二步 · 达成。 系统沿成本梯度重新分配，用替代路径达成同一结果。结果达成了，因此没有任何纠错信号产生。这一步是最优反馈控制预言的，不是异常。第三步 · 沉默中的固化。使用依赖的可塑性在这一步起作用：被用的通路被强化，被绕开的通路缺乏输入。下一次面对同一任务时，两条路径各自的成本已经不同于上一次——不是因为环境变了，是因为上一次的选择改变了它们。这三步里最要紧的是：整个过程中没有任何一次的结果读数发生变化。一个必须先说清楚的限定：借道是关于同一个执行者在同一个任务上的过程，不是关于分工，也不是关于工具本身。用计算器不是借道，除非你在某个时刻需要不用它。这条限定在第二十节会被展开成一条硬约束。

6.1 一个最小的形式陈述

把上面写紧一点，可以避免后面的争论落在措辞上。设某任务的结果为 R，由若干内部变量共同决定。把这些变量所在的空间分成两块：改变 R 的方向，和不改变 R 的方向。后者记为 N，即结果等价的自由度。在 N 上，任务提供的回复力为零，这一条由最小干预原理给出。设 N 上另有一个成本函数 C（代谢、疼痛、注意力的加权和），其梯度不为零。于是系统在 N 上的位置服从一个只有移动项和噪声项、没有回复项的过程：*位置的变化 = 一个沿负梯度方向的项 + 噪声*再加上第三步给出的反馈：C 本身依赖历史。被反复使用的方向变得更便宜，被绕开的方向变得更贵。这个陈述有三个可以直接读出的推论，后面各节要用：其一，R 的读数不含位置的信息。系统在 N 上走到哪里，R 都不变，这是 N 的定义。其二，唯一能把位置读出来的办法，是改变任务，使 N 的一部分不再属于 N。封闭替代路径正是在做这件事：把原本不影响结果的方向，临时变成影响结果的方向。其三，位置的移动速度取决于成本梯度的大小，与练得好不好没有关系。

七、为什么以结果为准的读数必然沉默

沉默不是因为测得不够仔细，是因为读数的定义。一个结果指标之所以是结果指标，正因为它对达成方式无感。跳远的距离不区分是膝关节做的功还是髋关节做的功；外开的角度不区分是髋外旋还是足旋前；考试的分数不区分是理解了还是记住了模板。如果一个指标能区分路径，它就不再是结果指标，而是过程指标。这不是抱怨，是同义反复。而正是这个同义反复，

构成了借道的掩护。回到第一节第二个现场，数字值得再看一遍。那组术后运动员单腿跳远距离的患健比是百分之九十七，标准差四个百分点；同一组人在同一个动作里，膝关节蹬伸做功的患健比是百分之六十九。跳落阶段，患侧膝吸收的功少于健侧，而健侧膝吸收的功多于对照组——也就是说，代偿不只发生在患肢内部的关节之间，还发生在两条腿之间。作者的结论写得很直接：跳出的距离主要反映的是髋和踝的功能，它不是一个好的膝关节功能指标。[1]另一项针对青少年的研究把这个图景补完了：达到距离对称的那一批人把负荷从膝转到了髋，没达到对称的那一批转到了踝。两批人都在卸载膝关节，区别只在于卸给了谁。[13]这里最值得停一下的是，距离对称的那一批人问题并不更小。他们只是把代偿做得更成功。而回归赛场的清单看的正是距离对称。同样的结构出现在外开上。第二节引的那项研究量的不是外开角度本身——那个量在有伤组与无伤组之间区分能力有限——而是外开角度减去被动髋外旋可提供的角度。换句话说：结果读数（外开多少度）不带信息，两个读数相减才带信息，而那个减数需要单独测量、需要设备、需要有人认为值得测。一个一般的说法：要看见借道，必须同时拿到结果读数和至少一个路径读数，并做差。任何单一的结果读数，无论多精确，都对它零信息量。这不是精度问题，是维度问题。

八、为什么主观读数的符号是反的

上一节的结论如果只到这里，问题还不算严重——测得更细就是了。真正的麻烦在这一节。在所有读数里，有一类是免费的、连续的、每个执行者都有的：主观感受。做得顺不顺，累不累，费不费劲。如果这类读数对借道敏感，那么执行者自己就是最好的探测器，外部测量只是辅助。它确实敏感。但符号是反的。人对自己学习状况的判断主要依据加工的流畅感，而流畅感本身就是省力的读数。[7] 把它接到第五节，这条链就完整了：

借道 → 动作更省力 → 流畅感上升 → 主观掌握感上升。

唯一一个对借道敏感的免费读数，在借道发生时朝好的方向动。这一步把问题从「看不见」升级为「看见的是反的」。看不见还可以靠增加测量来补；看见的是反的，意味着执行者的每一次自我评估都在为借道背书，而增加测量会遇到执行者的抵抗——因为他有充分的主观理由相信自己在进步。有三个可核对的旁证。其一，步态的能耗优化可以在内隐的情况下完成。[9] 人不知道自己改了，因此也没有可以自陈的内容。这意味着「你觉得你是怎么做的」这个问题在这里没有诊断价值——不是因为人不诚实，是因为信息不在那里。其二，疼痛下的重分配同样不经过决策。没有人在疼的时候决定改用另一组运动单元。其三，古典舞的审美判据把这件事推到了极端。芭蕾要求舞者显得毫不费力，而「毫不费力」在旁观者那里同样是一个流畅感判断。于是一个借道成功的舞者，可能在审美读数上得分更高——她把困难藏起来

了，而藏起困难正是这门艺术要求的。这一条不是修辞。它意味着在这个领域里，外部专家读数与主观读数的偏差方向是一致的，两者不能互相纠正。第二节说过，芭蕾教学法假定老师能看出学生从哪里借力。本章的判断是：老师能看出做得难看的借力，看不出做得漂亮的借力，而后者恰恰是那些将来会出问题的人——因为他们更成功地满足了审美，也因此更少被纠正。这是一条可以直接检验的推断：在同一批学生里，教师标注的「外开有问题」名单，与客观测得的差额排名，重合度应当明显低于教师自己的估计。

九、两根轴，四种局面

到这里为止，本章只有一个名字。名字不是判断。要把它变成可用的东西，必须能分辨——必须说清楚在什么情况下它成立、在什么情况下不成立，而且这两种情况在实践中都真的存在。第一根轴：路径集合是否随使用而改变。一端是惰性的：用了甲不影响乙的可用性。工程冗余大体如此——飞机的第二套液压系统不会因为第一套一直在用而失效。另一端是使用依赖的：用了甲，乙就变贵。任何有学习的系统都在这一端。第二根轴：路径能否被独立读出。一端是可读的：存在一种测量，能在不封闭替代路径的前提下单独告诉你原路径的状况。等速测力计可以单独测某块肌群，被动关节活动度可以单独测髋外旋的上限，强制引出任务可以单独测某个语法结构的掌握。另一端是不可读的：在完整的执行里，你只能看到合成的结果，无法把两条路径的贡献分开，除非把其中一条堵上。

	路径可独立读出	路径不可独立读出
路径集合惰性 （用了不改变可用性）	甲 · 真冗余 用了不影响可用性，且随时可查。冗余等于韧性这条论断在这一格成立，也只在这一格成立。 *处置：定期分项检查。* *例：双液压系统、双电源、异地备份。*	乙 · 盲冗余 不会因为不用而坏，但你不知道它还在不在。 *处置：不预告的真实调用。* *例：应急预案、消防演习、故障注入测试、家里那把备用钥匙。*
路径集合使用依赖 （用了就改变可用性）	丙 · 可见侵蚀 会退化，但退化能被提前看见，因为原路径可以单独测。 *处置：例行分项测量，设最低线。* *例：术后患侧肌力、被动髋外旋角度、语法结构的强制引出。*	丁 · 借道 会退化，且退化在任何常规读数里为零，在主观读数里为正。 *处置：只有剥夺落差。* *例：见第十九节。*

甲格什么都不用做，定期看一眼即可。乙格需要真的把它用一次——演习的全部意义就在于它不是演习式的。丙格需要的是纪律：把那个分项测量做成例行，并且不允许它被结果读数

替代。丁格是本章的对象。它的特点不是退化更严重，而是没有任何常规手段能在退化发生时发现它。第二根轴可以被技术改变。二十年前，单腿跳远的关节力学分布是不可读的；有了三维动作捕捉和逆动力学之后，它变成可读的——丁格的一部分被搬进了丙格。这是好消息，也是本章希望发生的事。但它不改变结构：每有一个路径读数被发明出来，就还有别的方向仍然不可读，因为不可读的方向是那些没有人想到要测的方向，而「没有人想到要测」这件事本身没有读数。

十、四步查法与剥夺落差

给定一个实践，判断它是不是落在丁格，可以按四步走。前三步用现成信息，第四步需要动手。第一步：结果读数稳不稳定。 如果结果读数在下降，问题是明面的，按常规处理即可。结果读数长期稳定不是排除条件，而是前提条件。 借道的定义里就包含结果不变。这一步最容易被搞反：很多人把「指标一直很好」当作不必检查的理由。第二步：有没有不封路就能测原路径的手段。 有，落丙格，去做那个测量。没有，进入下一步。这一步要求诚实：常见的自欺是把一个相关指标当成路径读数。跳远距离不是膝功能读数，外开角度不是髋外旋读数，通过率不是能力读数。第三步：路径集合是不是使用依赖的。 判据是一个问句：这条替代路径用了三年之后，回到原路径需要重新学吗？ 如果只需要重新想起来，那是惰性的；如果需要重新练，那就是使用依赖的。任何涉及生物组织、技能、组织记忆或模型参数的系统，答案基本都是后者。第四步：剥夺测试。 前三步只能把候选筛出来，不能定案。定案只有一种办法：临时封闭替代路径，看结果掉多少、多久回得来。第四步是四步里唯一不被自陈污染的一步，理由在第八节已经给过：执行者的主观读数在这个问题上符号是反的，因此任何「你觉得你能不能不借」的问法都无效。不是因为受访者会撒谎，是因为他手里没有这个信息。

10.1 剥夺落差

剥夺落差指的是：临时封闭替代路径之后，结果指标相对于封闭前的跌落幅度，以及恢复到封闭前水平所需的时间。它有两个分量，两个都要记：幅度说明借了多少，恢复时间说明原路径退化到什么程度。这两个分量可以分离，而分离本身有诊断意义。幅度大而恢复快，说明原路径还在，只是不常用；幅度大而恢复慢，说明原路径已经退化，需要重建。怎么封路因领域而异，形式是一样的：加一个约束使那条替代路径不可用，而任务本身不变。芭蕾的外开可以把足与小腿固定在中立位再量角度；术后的膝可以限制髋与躯干的代偿再测跳距；导航能力可以在熟悉的城市里关掉导航，量绕路比例与决策延迟；语言可以用强制引出任务逼出被回避的结构。有一件事必须写在前面：剥夺测试有风险。 在临床与竞技场景里，封闭一条正在承担

保护职能的代偿路径，可能导致再次损伤。因此本章主张的剥夺测试必须是低负荷、可控、以诊断为目的的，不是训练手段，更不是竞技条件。急性期一律不做。

10.2 怎么算封住了路

剥夺测试很容易做成一个无效测试，所以操作定义要写死三条。第一，被封的必须是路径，不是能力。 加一个约束之后，如果任务本身的难度也变了，那么跌落幅度里混进了新任务的学习成本。判据是：约束对一个从不借道的参照者应当几乎没有影响。固定足踝对一个髋外旋充裕的舞者，外开角度不该掉多少。做不到这一点说明约束设计错了，要重设，不能靠事后统计校正。第二，必须有一个等约束的对照条件。 任何外加的护具、绑带、限位板都会带来重量、皮肤压力与心理提示。因此需要一个施加同样负担、但不封闭目标路径的条件。跌落幅度取两者之差，不取绝对值。第三，恢复时间必须在解除约束之后测，不在约束之中测。 约束中的适应是新任务的学习，解除后的恢复才是原路径的重新可用。两者容易混，而它们指向完全不同的结论：约束中适应快说明系统灵活，解除后恢复慢说明原路径确实退化了。这两个数可以同时很高，而这正是丁格最典型的读数组合。三条合起来还有一个副产品：它们使剥夺测试无法被随便做一下。这在实践上是缺点，在方法上是优点——一个容易做的测量会很快被做成走过场，而走过场的剥夺测试给出的落差恒等于零。

十一、把整套做完一遍：一个可以真封路的系统

到这里为止，本章的判据都只能写成设计。在人身上，把整套做完一遍需要：先建立原路径，再给一条替代路径，再封路，再看恢复——而中间每一步都要花几个月，封路那一步还有伤人的风险。有一类系统里，这四步可以全部做完，而且封路不花钱、不伤人：一个会学习的人工网络。 它满足丁格的两个定义性条件——路径集合随使用改变（这正是梯度下降在做的事），路径不可独立读出（在完整的前向传播里，你无法把两条通路的贡献分开）。它不满足的是生物学的一切细节，因此下面这段跑出来的东西只能检验结构性主张，不能证明人身上的事。这条限制在第十二节末尾会再说一次，不是客套。

11.1 设定

任务是一个二分类。输入分两块：原路径是三十维特征，标签由它们的一个非线性组合决定，学起来要费些力气；替代路径是一维特征，与标签直接相关，一眼就能用。训练分两个阶段。第一阶段（一百二十轮）替代路径是纯噪声，网络只能靠原路径，这一阶段结束时的封路正确率就是原路径的基线。第二阶段替代路径开始带信息，继续训练到第三百轮，中途在若干时点同时读两个数：联合正确率（两条路都在，等于日常的结果读数）与封路正确率（把替代

路径重新打成噪声）。两者之差就是剥夺落差。第二阶段结束后再做恢复测试：在封路条件下继续训练，数需要几轮回到基线，并与一个从随机初始化开始、只用原路径训练的对照比较。网络是两层各六十四个单元的前馈网络，Adam 优化，学习率 0.002，批大小 128；训练集六千例，测试集四千例；每个条件跑八组不同的随机种子，报告平均。四条预测在跑之前就写死了，跑完不改：- 结果读数沉默：替代路径可用之后，联合正确率不下降。- 落差随时长增长：封路后的跌幅随替代路径可用的时长单调上升。- 有节省效应：封路后恢复到基线所需轮数，明显少于从头学——这是区分「退化」与「从来没学会」的那条线。- 外源随机化是解药：训练中随机屏蔽替代路径，结果读数不变而落差大幅下降。

11.2 读数

第二阶段轮数	联合正确率（结果读数）	封路正确率（原路径尚存多少）	剥夺落差
0	0.8048	0.8079	−0.003
20	0.9866	0.6664	+0.320
40	0.9867	0.6668	+0.320
80	0.9868	0.6671	+0.320
120	0.9870	0.6677	+0.319
200	0.9871	0.6689	+0.318
300	0.9874	0.6691	+0.318

第一条预测成立，而且比预期更彻底。结果读数不但没有下降，它从 0.805 升到 0.987。与此同时，原路径的存量从 0.808 掉到 0.669——掉到了比第一阶段开始时更差的水平。在整个过程中，任何只看结果读数的观察者会看到一条持续向好的曲线。这就是第七、八节那两句话的直接读数：结果读数不但沉默，它朝好的方向动。

11.3 结果读数一样，落差差得很远

第十节说过一条推断：在结果表现同样优秀的一群人里，剥夺落差的分布应当很宽。八组种子在第三百轮的读数是：联合正确率介于 0.9840 与 0.9890 之间，标准差 0.0017；剥夺落差介于 0.292 与 0.345 之间，标准差 0.019。以相对离散度计，落差的分散程度是结果读数的三十余倍。换句话说：八个「成绩几乎完全一样」的个体，借的深浅相差很大，而这件事在成绩上一点也看不出来。

11.4 替代路径并不更准的时候，会怎样

上面那组有一个明显的混淆：替代路径单独就能到 0.986，远好过原路径的 0.808。系统转过去，可能只是因为它更准，与省不省力无关。因此做了一组对照：把替代路径的信噪比调低，使它单独使用只能到 0.8102，与原路径的 0.8079 基本持平。此时系统没有任何「更准」的理由去用它，唯一的差别是它更好学——一维的线性关系，比三十维的非线性组合便宜得多。

第二阶段轮数	联合正确率	封路正确率	剥夺落差
0	0.8073	0.8079	−0.001
20	0.8879	0.7758	+0.112
80	0.8916	0.7778	+0.114
300	0.8931	0.7789	+0.114

替换照样发生。 联合正确率升了八点五个百分点，而原路径的存量掉了二点九个百分点。掉幅比上一组小得多，但方向一样，而且在没有任何准度诱因的条件下发生。这一组是本章机制说明的关键支撑：「更省事」单独就足以造成替换，不需要「更准」。 而在结果读数上，这一组看起来是一次干净的进步。

十二、数据打掉了本章四条预测里的两条

上一节报告的是成立的那两条。这一节报告不成立的两条，以及本章因此必须改的地方。

12.1 落差不是慢慢长出来的，它一次就到位

第二条预测说落差随时长单调上升。表二显示它在二十轮之内就跳到 0.320，此后三百轮几乎不动，甚至微微回落。这直接打掉了第六节第三步那个「一点点加深」的图景。替换不是慢慢渗透的，它在替代路径一变得可用时几乎立刻完成。但事情有一个下半段。另加一组条件：给训练加上权重衰减——在这个系统里，权重衰减就是一个不折不扣的成本梯度，它在任务不管的方向上把参数往零推。

第二阶段轮数	封路正确率	剥夺落差	常规条件的落差
20	0.6632	+0.324	+0.320
80	0.6518	+0.338	+0.320
200	0.6460	+0.342	+0.318

加了成本梯度之后，落差确实继续增长。所以第二条预测不是全错，是把两件事混成了一件：

> 快的那一半是替换，由替代路径的可得性驱动，一次到位；慢的那一半才是移动，需要一个持续的成本梯度供给。

这个修正把第五节那三个坡的地位降了一级，也说得更准了。它们不是替换的原因——替换在替代路径出现的当下就发生了。它们是替换之后继续前进的原因。还有一条更让人不舒服的读数。在等准度那一组里加同样的成本梯度，封路正确率不但没有继续下降，反而从 0.7776 升到 0.7949。也就是说，成本梯度并不自己制造方向，它放大已经存在的方向。当两条路径旗鼓相当时，同一个成本梯度起的是正则化的作用，对原路径反而有利。第五节那句「成本提供一个恒定的坡」因此必须收窄成：成本梯度沿已经更便宜的那一条推进，它不造坡，它加陡已有的坡。

12.2 没有节省效应，反而是负的

第三条预测是本章用来与「从来没学会」分界的那一条：借道过的系统在封路之后应当恢复得比从头学快，因为原路径曾经存在。八组种子的恢复轮数是 3、8、7、5、36、2、4、2，中位数 4.5；从随机初始化开始只用原路径训练达到同一水平所需的轮数是 6、2、1、2、2、2、3、2，中位数 2。方向是反的。借道过的系统恢复得比从没学过的还慢，而且方差极大——有一组用了三十六轮，另外几组两三轮就回来了。这条结果比本章原来的预测更狠，也更难堪。它说的是：在这个系统里，原路径不只是退化到起点，它退到了起点以下。已经被替代路径占住的表征，成了重建原路径的障碍，而不是残余的助力。本章必须据此改两处。其一，第十五节那条用来与捷径学习分界的判决性预测——「封路后恢复快于初学者」——在这里被推翻了，本章撤回它，换成方向相反的一条：捷径学习预测封路后恢复与从头学大致相当（因为它主张目标特征从未被学到）；本章的丁格预测恢复慢于从头学。这一条同样可裁决，而且更好分辨，因为它预测的是一个反常方向。其二，剥夺落差的两个分量的含义要改。原先写的是「幅度大而恢复慢说明原路径已经退化」。现在应当写成三档：恢复快＝原路径还在；恢复与从头学相当＝原路径基本没了；恢复慢于从头学＝原路径的位置被占了，重建要先拆。第三档是最坏的一档，也是本章原先没有想到的一档。一句公道话：这只是一个系统里的一组读数，负节省完全可能来自这个网络的具体结构——被占住的单元、饱和的激活、优化器的路径依赖。所以本章不主张负节省是一条定律，只主张：「有节省效应」这条曾被本章当作分界线的预测，已经有了一个明确的反例，不能再当作既定事实使用。

12.3 外源随机化确实是解药，但它一点都不便宜

第四条预测说：训练中随机屏蔽替代路径，结果读数不变而落差大幅下降。

条件	联合正确率	封路正确率	剥夺落差

常规	0.9874	0.6691	+0.318
外源随机化	0.8525	0.8275	+0.025

落差被压掉了九成二，而且原路径的存量 0.8275 比第一阶段的基线 0.8079 还高——随机屏蔽不只是保住了原路径，它还把原路径练得更好了一点。但结果读数掉了十三点五个百分点。预测里那句「结果读数不变」是错的。这个错值得单独说，因为它把第二十节那条约束从口号变成了一个可以算出来的数。两种做法的期望表现，取决于替代路径将来有多大概率不可用。设那个概率为 *p*：常规条件的期望是 0.9874(1−*p*) + 0.6691*p*，随机化条件的期望是 0.8525(1−*p*) + 0.8275*p*。两者相等时*p ≈ 0.46*也就是说：只有当替代路径失效的概率超过四成六时，随机化才划算。 这是一个很高的门槛。它意味着在绝大多数「替代路径基本不会消失」的场合，借道非但不是病，反而是正解——而这正是第二十节第三条约束要说的话，现在它有了一个数。本章原来把随机化写成一个可以普遍推荐的处方。这条数据把它降级为一个有条件的、需要先估中断概率的处方。

12.4 这一节能证明什么，不能证明什么

能证明的是三件结构性的事：在一个路径可替换、且路径不可独立读出的学习系统里，结果读数确实会在能力下降的同时上升；「更省事」不需要「更准」就足以造成替换；封路是唯一能把这件事读出来的操作。 这三件事不依赖任何生物学细节，它们依赖的只是「结果等价」这个几何条件与「使用改变可用性」这个动力学条件。不能证明的是任何关于人的事。这个网络没有代谢成本，没有疼痛，没有主观感受，也就没有第八节说的那个反向读数——那一节仍然只有观察证据支持，没有实验证据。它的时间尺度是任意的，它的负节省可能是它自己的毛病。本章用它，只是因为它是目前唯一一个能把封路测试整套做完的地方。这是一个便宜的地方，而挑一个便宜的地方做研究，正是本章在第二十三节要指认自己的那件事。

十三、这与「有益的变异」不是一回事

本章的主张最容易被读成对运动控制主流的一次否定。它不是，而且分界必须划得比「侧重不同」更实。冗余。 生物学里的这个概念说的是：结构不同的元件可以产生相同的功能输出，而这正是生物系统抗扰动的来源。[22] 它说到的一步是可能性——存在多条路。它没有说的是：这些路在被使用之后，彼此的可用性会不会改变。分离线在这里：冗余是一个关于集合的陈述，借道是一个关于集合随时间演化的陈述。判决性对照：给定一个长期在同一环境里高频执行同一任务的系统，冗余论断预测它的抗扰动能力不随时间下降；本章预测下降，且下降速度与执行频率成正比、与环境变异度成反比。同一个系统，两条预测方向相反，都可以在纵

向数据里读出。不受控流形。 这套方法把变异分解到影响结果与不影响结果两个子空间，并测得后者变异更大。[3] 它说到的一步是方差的分配。它没有做的是追踪那个子空间里均值的位置——因为在它的框架里，那个位置没有意义，任何位置都同样好。分离线：本章主张那个位置有意义，且它会朝一个方向移动。判决性对照：不受控流形方法不预测子空间内均值有系统性位移；本章预测有。这一条可以用已经采集好的纵向数据回答，不需要新实验，而且它能干净地否掉本章。最小干预原理。 这条最微妙，因为本章的机制建立在它之上，而结论与它相反。它说到的一步是：不纠正任务无关偏差在单次执行里是最优的。[4] 本章接受这一点，并指出它在跨次累积上有代价——单次最优的策略，跨次会把系统推到一个任务不再关心、而身体很关心的位置。分离线一句话：他们证明的是控制器的最优，本章说的是被控对象的变化。判决性对照：最优反馈控制把系统看作参数固定的被控对象，因此预测重复执行不改变控制律的解；本章预测解会移动，因为被控对象的成本景观在被自己的历史改写。在一个参数固定的机械臂上，本章的预测必然为假——这既是分界，也是本章适用域的边界。差异化学习与非线性教学法。 这一族主张主动往练习里加变异与噪声，反对高度重复的标准化练习。它与本章的关系是最有意思的一种：它是本章的处方，不是本章的对手。 分离线在变异的来源：由成本梯度内生选择的变异有方向，因此累积；由外部随机施加的变异没有方向，因此不累积，并且强制系统保持多条路径的可用性。第十一、十二节那组随机屏蔽的数据支持这一点，同时也给它标了价（见第二十节）。判决性对照：如果变异本身就是好的，那么变异量相同而来源不同的两种练习效果应当相似；本章预测在剥夺落差上外部随机组显著优于自选组，而在结果指标上两组无差异。

十四、这与「习得性废用」不是一回事

习得性废用是本章最近的一位邻居，近到必须单独处理。这一支的观察是：中风或神经损伤之后，患肢在早期尝试中反复失败，患者转而用健侧完成一切，久之患肢功能进一步丧失——而这部分丧失是行为性的，不是神经损伤本身造成的。约束诱导疗法的做法正是把健侧固定住，强迫使用患侧。[14]结构上，这几乎就是本章说的事：替代路径挤占原路径，最后靠封闭替代路径来恢复。分界必须落在三处。第一，触发条件不同。习得性废用需要一次可辨的失败——尝试、失败、回避，这是一条标准的操作性条件反射链。本章说的借道不需要任何失败。第五节那组步态证据是关键：受试者没有失败过任何一次，他们只是走路，而步频在几分钟内挪走了。第十一节那组等准度对照给了同一件事的第二个读数：替代路径并不更准，替换照样发生。第二，适用条件不同。习得性废用的诊断依赖「能力与使用之差」：患肢的能力可以被单独测出，与实际使用量对比，差额即诊断。本章的丁格恰恰定义为原路径不能被单独测出的情形。按第九节的两根轴，习得性废用落在丙格——这也解释了为什么约束诱导疗法能被

做成一套标准化疗法，而丁格里的问题至今没有标准疗法。第三，恢复曲线不同。习得性废用的核心主张是能力还在、只是没被用，因此约束健侧之后患侧功能迅速恢复，以周计。第十二节那组数据把本章这一侧的预测改成了更极端的形状：恢复不但不快，可能比从头学还慢。在同一个剥夺测试里，两条理论预测的恢复曲线形状因此相反：一条是解除抑制的阶跃，一条是先拆后建的凹坑。还有一条更干净的：本章预测，在从无失败史、从无损伤史、结果表现长期优秀的健康人群里，同样能测到显著的剥夺落差。习得性废用的框架不预测这里有东西。这一条可以直接判本章的生死。使用依赖可塑性与用进废退。 这一族说的是同一条通路的强化与弱化，是量的连续变化，双向可逆。分离线：它不涉及路径之间的替换关系，因此不解释为什么退化可以在总体表现完全不变的情况下发生。判决性对照：用进废退预测总量下降（整体表现变差）；本章预测总量不变而构成改变，只有在剥夺条件下才暴露。疼痛适应。 它描述的重分配是本章机制的一个特例，也是证据来源之一。[11] 分离线：它的自变量是疼痛，本章的自变量是任何一种成本差。判决性对照：它预测无痛者不出现重分配；本章预测无痛者也出现，只是更慢。给一组无痛受试者施加一个纯代谢的不对称（例如单侧加配重），本章预测会出现同形的重分配，它的框架不预测。性状丢失。 洞穴鱼失去视觉，这是「不用即失」的经典案例，容易被拿来给本章背书。它不能。分离线：那里的机制是选择压消失加中性突变累积，时间尺度是代际的，与个体内的使用依赖可塑性差六个数量级。这位邻居反过来给本章加了一条约束：借道的时间常数由个体的可塑性决定，不能外推到没有个体内学习的系统。

十五、这与「从来没学会」不是一回事

捷径学习。 这项工作指出，深度网络常常学到与任务表面相关、但在分布外失效的特征：靠背景判物种，靠纹理判类别，靠拍摄伪影判疾病。[15] 测试集准确率很高，换一个分布就崩。这个形状与借道极像，而且它给了本章一个非生物的领域，说明机制不依赖肌肉。分离线在原路径是否曾经存在。捷径学习的模型从一开始就没有学到目标特征——它不是退化，是从未获得。本章原来给的判决性对照是节省效应：屏蔽掉捷径特征之后，捷径学习预测从头学，本章预测重学更快。第十二节的数据推翻了这一条，本章已经撤回它。 换上的是方向相反的一条：捷径学习预测封路后的恢复与从头学大致相当；本章预测慢于从头学，因为原路径的位置已经被占，重建要先拆。第十一节的设定恰好可以把两者分开——那里的原路径是被真正学会过的，而它在封路后的恢复中位数是从头学的两倍多。这一条比原来那条更好分辨，因为它预测的是一个没有人会去猜的反常方向。回避策略。 这一条来自二语习得，是本章最喜欢的一个先例，因为它比本章早了半个世纪就把靶子说清楚了。这项工作指出，错误分析有一个方法论上的漏洞：学习者可以通过回避某个难的结构使该结构的错误率为零。中文与日语背景的学习者产出的英语关系从句比欧洲语言背景的学习者少得多，而错误率更低。[16] 低错误率被读成

掌握，实际上是没有尝试。分离线：回避是一个关于产出选择的说法——学习者可以不说那个句子。借道说的是同一个产出内部的路径替换——句子照说、动作照做、结果照达，只是里面换了。判决性对照：回避策略预测在强制引出任务里表现会掉，掉下来的是产出率；本章预测在强制引出任务里产出率不掉，掉的是产出的内部构成。这一条还有一个用处：它证明本章说的这类问题不是新问题，只是每个领域各自发现过一次、各自命名、各自没有推广。这也正是本章认为需要一个统一说法的理由。

十六、这与「人被移出回路」不是一回事

自动化的讽刺。 这条论证是：自动化把常规操作接管过去，人只在异常时被叫回来，而恰恰因为常规操作被接管，人维持技能所需的练习没有了，于是在最需要人的时刻，人最不行。[17]分界要落在「谁做了这个决定」。自动化的讽刺里，人被移出回路是一个可见的设计决定，有文档、有时间点、有责任人。本章说的借道没有任何决定发生：同一个人在同一个任务里，用同样的时间做同样的事，只是内部分配悄悄变了。没有人拿走什么，也没有人交出什么。判决性对照：自动化的讽刺预测技能退化的时间起点与自动化上线的时间点对齐，并且在没有自动化的对照组里不出现；本章预测在完全没有自动化、没有工具、没有任何外部接管的纯人工执行里，退化照样发生。一个从不用任何辅助的芭蕾学生，如果测到显著的剥夺落差，那个框架无法解释，本章可以。变通办法与第一序问题解决。 这项组织行为学研究观察医院护士：遇到障碍时，她们用变通办法当场把病人的问题解决掉，而不上报。系统的结果读数正常，而组织的学习没有发生，同样的障碍第二天照旧。[18]这是本章在组织层面的同形。分离线：他们研究的是信息是否上报，本章研究的是能力是否退化——两者可以分离，一个组织可以既不上报又不退化。判决性对照：他们的处方是建立上报渠道与心理安全；本章预测仅仅建立上报渠道不会改变剥夺落差，因为上报改变的是组织的知识，不是执行者的路径分配。要检验，只需在一个已经建立了良好上报文化的单位里做剥夺测试。导航与空间记忆。 一项研究发现，习惯使用卫星导航与自主导航时的空间记忆表现较差，并有纵向证据支持方向。[19] 分离线：这一族通常把它读成「工具替代了能力」，因此处方是少用工具；本章的读法是，工具只是提供了一条特别便宜的替代路径，机制与工具无关。判决性对照：如果是工具的问题，剥夺工具应当足以恢复；本章预测剥夺工具后恢复缓慢，且恢复速度与借道年限负相关。这是一条任何人都能在自己身上做一遍的测试，而且它的结果很可能让人不舒服。

十七、这与「合意难度」不是一回事

这一节单独给合意难度，因为它是最近的一位，也是最容易吞掉本章的一位。它的主张是：一些使当下表现变差的练习条件会改善长期保持；而学习者系统性地偏好使当下表现变好

的条件，因此系统性地选错。[20] 结构同形：当下好、长期坏；主观判断站在错的一边；处方是引入困难。三条都对得上。分界必须落得很硬，否则本章只是这条命题在运动领域的一次应用。分界落在自变量的位置。合意难度的自变量是练习条件的安排：间隔还是集中，交错还是分块，提取还是重读，变异还是恒定。这些都是外部可观察、可操纵、可由教师或学习者选择的变量。它的因果故事是：人选了更容易的练习安排。本章的自变量在练习安排的下游，而且不涉及任何选择。给定完全相同的练习安排——同样的间隔、同样的交错、同样的次数、同一个教练、同一份计划——两个人内部的路径分配仍然会分开，因为他们的成本景观不同（髋外旋角度不同、旧伤不同、体型不同）。没有人选了什么。 第十一节那八组种子是这句话最干净的版本：同一份训练计划、同一个结构、同一批数据，八个系统的结果读数几乎完全一样，而剥夺落差的分散程度是它的三十余倍。由此得到本章与合意难度之间最锋利的一条判决性对照：

> 合意难度预测，把练习安排改成困难版（交错、变异、间隔、提取），问题就会被大幅缓解。本章预测，在困难练习安排下借道照样发生，而且可能更快——因为提高难度就是提高成本，而成本正是那个坡；除非那个困难是直接封闭替代路径的那一种。

这条对照可以做成一个三组实验：常规练习组、合意难度组（交错变异但不限制路径）、路径约束组（练习安排与常规组相同，但用外部约束封闭替代路径）。结果指标预计三组接近；剥夺落差预计路径约束组最小，而合意难度组与常规组的差异不显著。如果合意难度组的剥夺落差显著小于常规组，本章应当被并入合意难度，作为它在运动领域的一个特例。这一条是本章最愿意押上去的一条。流畅性错觉。 第八节整个建立在它上面，因此这不是分界，是致谢加一处推进。那一支说的是，元认知判断依据的是加工线索而非实际学习量。[7] 本章加的一步是：在运动领域，那个加工线索同时就是驱动移动的力。在记忆研究里，流畅性是移动的指示器；在运动领域，它是移动的原因。判决性对照：如果只是指示器，那么把流畅感与实际省力人为分开（例如加一个不改变代谢成本的干扰，使动作感觉更费力），移动速度应当不变；本章预测会变。古德哈特定律。 常被用来概括「看结果就会坏」这一类现象。分离线很干净：古德哈特需要一个瞄准指标的行动者——有人知道指标是什么，并据此调整。借道不需要任何人瞄准任何东西，它在一个不知道自己被测量的个体的无意识运动控制里就发生了；第十一节那个网络更是连「知道」这件事都谈不上。判决性对照：古德哈特预测隐瞒评价标准可以缓解；本章预测隐瞒评价标准对剥夺落差没有影响。

十八、与学科通融专栏几篇的分界

学科通融专栏此前有几篇处理的对象与本章相邻，需要说清楚不是同一件事。先说最近的四篇。学科通融专栏在同一天另有四篇取的是同样这三个领域，它们与本章是彼此最近的邻居，因此分界必须比与外栏各篇更细。四篇取的是同一批材料的不同切法，谁也不能吸收谁，理由如下。《再生核》问的是一门技能由什么构成——移走哪一样，下一个人就长不出来了。它的尺度跨人跨代，量的是内容能不能传下去。本章的尺度在一个人身上，量的是同一次执行里走的是哪条路。两者的预测在一个具体情形上方向相反：取一样按那篇判定为「移走了下一个人照样长得出来」的东西，若它恰好是当前替代路径的来源，那篇预测移走无损，本章预测移走之后剥夺落差反而下降——因为封掉的正是借道的入口。同一次移除，一篇说无关，一篇说有益。《可失败余量》问的是一次练习削减到哪一步就不再出得了错。它的坏结局是没有失败的机会，所以什么都没改；本章的坏结局是改成功了，只是改在了不该改的地方。两者可以同时发生而互不蕴含：一次练习里当场失败很多次（按那篇属好），若每一次失败都被同一条替代路径顺利化解，本章预测剥夺落差照常增长。失败次数与剥夺落差可以同时很高，这是把两篇分开最省事的一次测量。《孤线》问的是哪些偏差还算「同一个动作」，答案是这条线由当时那套判定安排切出，而线可能在安排失效之后还留着。那一篇管线从哪来，本章管线以内那片空地上会发生什么。判决性对照：把判定口径整个换掉，那篇预测容差线随之移动，本章预测剥夺落差纹丝不动——漂移由成本梯度定，不由谁在判定定。两篇合起来才完整：一篇说那条线可能没人再校准，一篇说线以内并不是静止的。《收得回的那一遍》问的是动作发起之后还能不能收回，它的靶格里结构缺陷会以「执行者能力不足」的样子显现。差别在缺陷的下落：那一篇里缺陷显现了但被归错了对象，本章里缺陷根本不显现。判决性对照：把归因材料补齐、给出完整的表现层反馈，那篇预测遮蔽解除，本章预测剥夺落差不动——结果层的反馈再详细也不含路径信息，除非它本身就是一个路径读数。四篇加本章共用同一批领域，这不是巧合，也不是五次重复：三个领域之所以撑得住五种切法，是因为它们争的那件事本来就有五个不同的变量。这一点同样构成一条限制——五篇没有一篇能宣称自己抓到了那个唯一要紧的变量，本章也不例外。《复推率》处理的是一个人「学会了没有」为什么无法从作品上看出来，它给的答案是规则不被存储、每次被重新推断，学会即重新生成的可靠性上升。那一篇与本章的分界在于对象：它问的是同一个人在新情境里能不能推出同一条规则，本章问的是同一个人在同一个情境里用的是哪条路。两者可以同时成立且互不蕴含——一个复推率很高的人完全可能借道很深，因为复推的是判断，借的是执行。判决性对照：换一个表面不同结构相同的情境，复推率预测表现保持；本章预测在同一个情境里封路，表现下跌。两者可以在同一批人身上分别测，且不必相关。《内容权》处理的是「一个人做得最好的时候，意识应该在哪里」，它撤销的前提是「意识介入是一个量」，剩下的变量是下一个动作由谁定。本章与它正

交：内容权问的是谁在决定，本章问的是决定之后走的是哪条路。一个完全握有内容权的人照样借道，因为路径的选择根本不到决定这一层——第八节那三条旁证说的就是这个。《静止点》与《一天与三年》都涉及「表面已完成而实际未完成」，但它们量的是时间（能不能重写、跟不跟得上），本章量的是空间上的替换。三篇有一个共同的形式特征值得点出：都是在有台账的那一层做研究，而正文都承认关键发生在没有台账的那一层。 本章第十二节那个网络就是本章的台账层，第二十三节会为此认账。

十九、逐领域兑现

一个说法要值钱，得能在它的出生地之外预测出具体的东西。以下十三处，每一处给出预期读数与可做的剥夺测试。其中最后三处反过来对本章提出了要求，写在第二十节。- 古典舞的外开。 预测：把足踝固定在中立位后能保持的外开角度，与损伤史的关联强于外开角度本身。剥夺测试：旋转限位板加固定鞋垫。本章的增量是把那个差额从相关指标改成诊断量并要求纵向使用。- 术后回归赛场。 预测：在通过距离对称标准的人群里，限制髋与躯干代偿之后的跳远跌落幅度，比距离对称本身更能预测二次损伤。剥夺测试：躯干固定的单腿跳。这条可以挂在现有测试流程上，边际成本很低。- 投掷与击打类项目的动力链。 预测：成绩长期稳定的运动员中，肩肘负荷分布的个体间差异远大于成绩差异。剥夺测试：限制躯干旋转幅度后测球速跌落。- 声乐。 肌紧张性发声障碍的典型形态是用喉外肌代偿声带闭合不足，音量与音高读数达标而耐力崩塌。预测：在限制喉外肌用力的姿势下测最长发声时间，比常规嗓音评估更早预警。- 器乐演奏。 钢琴用整臂重量代偿手指独立性，小提琴用手腕代偿弓压控制。预测：在限制近端关节的条件下，音色均匀度的跌落幅度与将来的过劳损伤相关。- 二语产出。 沿回避策略再走一步：在强制引出任务里，产出率不掉而内部构成掉。剥夺测试：限定必须使用某结构，量产出延迟与自我修正次数。- 手术技能。 微创与机器人术式普及之后，开放术式的能力无法从常规手术量读出。剥夺测试：模拟器上的开放操作，量时间与误差。- 手动飞行。 预测：手飞熟练度与总飞行小时数的相关很弱，与近期手飞小时数的相关强得多。剥夺测试：非预告的手动接管评估。- 自主导航。 剥夺导航后的绕路比例与决策延迟，与导航使用年限正相关，与总驾驶年限无关。- 模型鲁棒性。 屏蔽捷径特征后的性能跌落即剥夺落差，而重训练曲线的形状可区分退化与从未获得。第十一、十二节就是在这块地上做的，它给了本章一个可以完全自动化、样本量不受限的检验场地。- 随机化作为解药。 训练里的随机屏蔽与数据增强，是外部施加的强制换路，效果与内生借道方向相反。它给本章加了一条约束，并且带了价签（见 20.1）。- 急性期的保护性代偿。 疼痛期的重分配是必要的，不该被取消。见 20.2。- 一般的工具与分工。 用计算器不心算、用输入法不写字、用检索不背诵。见 20.3。

二十、三处让步与一个可以算出来的盈亏平衡点

20.1 外源随机化不是同类，但也不是免费的

如果替代路径的使用一律导致退化，那么任何形式的换路都应当有害。事实相反：外部随机施加的换路没有方向，它的期望位移为零，而它的效果是维持整个路径集合的可用性——因为每条路都被迫定期承担一次。第十一节那组数据把落差压掉了九成二，还把原路径练得比基线更好。约束因此写成：本章的主张只适用于由成本梯度内生选择的路径替换；外部随机施加的强制换路不在其中，且是目前已知最直接的对策。但第十二节的数据同时给这条处方标了价：随机化使结果读数掉了十三点五个百分点。这条处方不是免费的，而且它不便宜。 本章原来把它写成可以普遍推荐的做法，现在必须降级为有条件的做法：先估替代路径失效的概率，再决定要不要付这个价。

20.2 保护性代偿不该被取消

急性损伤期、术后早期、疼痛发作期，借道是保护性的。此时封闭替代路径不是诊断，是伤害。因此本章不主张取消代偿，而主张给代偿加两样东西：期限和撤回测试。一条代偿路径被启用时，同时写下预期使用期限，和到期时如何检验是否还需要。没有期限、也从未被检验过的代偿，才是丁格。这条约束把本章的适用面收窄了不少，也把它从一句道德判断改回一条工程判断：问题不在借，在于没有人记得曾经借过。

20.3 判据不是「有没有退化」，是中断概率乘以中断代价

人类几乎全部技术都是借道：用车不走路，用文字不背诵，用检索不记忆。如果每一次都要问「你的原路径退化了吗」，答案永远是「退化了」，而这个答案没有用。因此判据必须换一个形状。它不是能力有没有退化，而是：

这条替代路径中断的概率，乘以它中断时你付出的代价。

第十二节把这句话算成了一个数。在那组设定里，随机化与常规做法的期望表现相等的位置是替代路径失效概率约为四成六。低于这个门槛，借道不是病，是正解；高于它，才值得付随机化的价。四成六是一个很高的门槛。它意味着在绝大多数「替代路径基本不会消失」的场合，本章的处方不该被采纳。这个数只属于那一组设定，换一个任务它会变，但它的存在把讨论从「该不该借」推到了「失效概率是多少」——而后者是可以估的。这一条也解释了为什么本章举的例子集中在身体上。工具与身体的区别不在于哪个更真，而在于替代路径与原路径的相关性：外部工具与身体能力大致独立，而身体内部的替代路径与原路径共用一副骨骼、一套

血供、一个疲劳池。独立的备份是冗余，相关的备份是伪装成冗余的单点——它会在最需要它的那一刻与原路径一起失效。

二十一、可以判本章错的八种结果

以下每一条读数不同。第一条最便宜。一 · 均值不动。 在已有的纵向运动捕捉数据里，对结果无影响的子空间内，各变量的均值不出现系统性位移，或位移方向与代谢成本梯度无关。这一条不需要新实验，只需要重新分析已经采集的数据，而它可以直接判死第四、五节的机制链。二 · 落差与稳定性不相关或负相关。 在结果表现同样优秀的一群人里，剥夺落差与结果表现的历史稳定性不相关，或者稳定的人落差反而更小。三 · 健康人群里没有东西。 在从无损伤史、从无失败事件、结果长期优秀的健康人群里，剥夺落差与匹配的对照组无显著差异。若如此，本章说的现象只是习得性废用的一个换名，应当被并入。四 · 恢复曲线与从头学没有差别。 剥夺之后的恢复速度，既不快于也不慢于从未练过该路径的匹配对照。若如此，第十五节改写后的分界同样作废，本章与捷径学习之间不再有可裁决的差别。五 · 合意难度就够了。 三组实验里，合意难度组的剥夺落差显著小于常规组，且与路径约束组无显著差异。若如此，本章应当被并入合意难度。六 · 无痛无异不出现。 给一组无痛受试者施加纯代谢的不对称，不出现同形的路径重分配。若如此，本章关于「触发不需要疼痛也不需要失败」的核心让步不成立。七 · 坡在却无人下坡。 若「自然状态下的变异不足以启动优化」这一条被推广确认 [10]——即在没有外部推动的日常执行里系统根本不做探索——那么本章说的移动在自然条件下不会发生。这是本章自己知道的最软的一处，写在这里而不是留给别人。八 · 等成本两路仍然分化。 人为构造两条成本完全相等的路径（等代谢、等疼痛、等注意力），若仍然观察到系统性的替换，则「成本梯度是驱动力」这一机制说明作废，本章只剩下一个现象描述和一个测量方法。顺带说明：第十一节那组等准度对照拉平的是准度，不是成本——那一组恰恰证明了成本差单独就够，所以它不构成对这一条的回答。

二十二、把它做成指标会发生什么

本章提出了一个量，因此有义务预言这个量被制度化之后会怎样变坏。剥夺落差一旦成为考核指标，被考核者会开始练习剥夺条件本身。 芭蕾学生会专门练戴着限位板的版本，运动员会专门练躯干固定的跳，飞行员会专门练那套评估科目。练熟之后，剥夺条件不再是剥夺——它变成了第三个任务，有自己的结果读数，也有自己的结果等价自由度。借道于是转移到没有被封的那些方向上去。这个反噬有一个可测的签名：引入考核之后的两到三个周期内，剥夺落差下降，而未被封闭的自由度上的移动速率上升。 如果只看第一个数，会得出干预有效的结论。由此推出唯一一条不易被文书化的操作建议：剥夺测试必须是不预告的，封闭哪一条路必

须是随机的，且不得作为训练科目。一旦它进入训练计划，它就失效了。这条建议本身有一个本章解决不了的冲突，写在这里交出去：不预告的剥夺测试在竞技与临床场景里有伤害风险，而降低风险最直接的办法就是提前告知与充分准备——那正好销毁了它的诊断力。本章没有解决这个矛盾。目前能想到的最好折中是把随机性放在「封哪一条」上而不是「什么时候封」上，但这只是把问题推小了一点。第二个洞一并交出：本章没有给出剥夺落差的量纲与可比性。跌落两成的外开角度，与跌落两成的跳远距离，不是同一件事；不同任务的落差不能直接比较，因此目前它只能做纵向自比，不能做横向排序。这限制了它作为选拔工具的用途，而这可能是好事。

二十三、本章自己在哪一格

按第九节的两根轴给本章自己定位，结论不好看。本章的结果读数是「读起来有说服力」。这个读数对本章的生成路径完全无感。而本章用的是一条明显更省力的路：把三个领域已经做完的经验工作拼在一起，绕开自己去采数据。第一节那三个现场，没有一个是本章的作者去看的；第七节那些数字，是别人为了别的问题采集的。本章借了他们的道，而「读起来有说服力」这个读数一点没掉。第十一、十二节看上去是对这一条的补救——本章确实自己跑了一遍。但那一跑挑的是最便宜的地方：一个没有身体、没有疼痛、封路不花钱也不伤人的系统。第十八节刚指出学科通融专栏几篇都在有台账的那一层做研究，本章做的是同一件事，而且做得更露骨——我不是没有别的选择，我是选了不用出门的那一个。第三重更难堪：跨领域写作是丁格最典型的例子。每一个领域的专家读到自己那一段，都会发现它是粗的。运动控制的人会觉得第四节把最优反馈控制讲简单了。舞蹈医学的人会觉得第二节没有充分处理那份系统综述指出的因果复杂性。学习心理学的人会觉得第十七节对合意难度的处理还欠一层。但没有一个读者同时是这三个领域的专家，因此这些粗糙不会汇总成任何一个读数。这正是本章说的那件事：路径的缺陷存在，结果读数为零。第四重：本章自己开的处方是外源随机化，而本章没有对自己用过它。三个领域是我挑的，挑的标准是我能顺利写。真正的外源随机化，应当是随机指定一个我不熟的领域并强制写进去，接受那一段明显更差。本章没有这样做。开药的人没有吃药。还有一件必须记在这里：第十二节报告的两条失败，是本章最值钱的部分，而它们之所以出现，只是因为那四条预测在跑之前就写死了。如果先跑再写，它们会被悄悄改成成立的样子，而且改完读起来毫无破绽——那就是一次标准的借道：结果读数（这篇文章的说服力）完全不受影响。按本章的逻辑，这几重加起来的处置只有一条：这篇文章的每一个经验主张都应当被当作待验证的。尤其是那些读起来最顺的段落——它们顺，很可能正因为借得最深。

结　论：一条写死日期的赌注

把话说死。到 2032 年 12 月 31 日为止，如果在健康、无损伤史、结果表现长期稳定的运动或舞蹈人群里，剥夺落差与结果表现历史稳定性之间的正相关没有被任何一项预注册研究检出，或者被检出而方向相反，本章的核心主张作废。届时不再申辩，不改口径，不把「稳定性」重新定义成别的东西，也不主张需要更长的时间。选这一条作为赌注，理由是它最接近本章的骨头。本章可以在很多地方让步——机制可以从三个坡缩成一个，适用域可以从所有实践收窄到只剩身体，随机化那条处方已经被自己的数据标了一个很高的价，「有节省效应」那条预测已经被自己的数据推翻并撤回。但这一条不能让：如果长年稳定的表现不比不稳定的表现借得更深，那么本章关于「任务不提供回程票」的整个说法就是错的，剩下的部分不值得单独保留。这一次作废的条目：「落差随时长单调增长」作废，改为「替换一次到位，此后的移动需要持续的成本梯度供给」；「封路后恢复快于从头学」作废，改为方向相反的「恢复慢于从头学」；「外源随机化使结果读数不变」作废，改为「随机化以约十三个百分点的结果读数换取落差的九成压缩，仅当替代路径失效概率高于约四成六时划算」；「成本梯度提供一个恒定的坡」收窄为「成本梯度放大已有的方向，不自己制造方向」。共四条。最后一句留给第一节那个十四岁的学生。她被要求收回一个已经站对的位置，退回一个更难看的角度，而她拿不到任何当场的证据说明这样做是对的。本章所做的全部工作，是给那位老师一个可以拿出来的读数，也给那个学生一个可以自己去测的东西。 至于外开该不该有一百八十度，那是另一门学问的事。

注　释

[1] 单腿跳远距离对称与膝关节做功的分离，见 Kotsifaki 等（2022）。该研究样本为男性运动员，文中百分比不应外推到全部人群。

[2] 关于自我调节学习中的信念、技术与错觉，见 Bjork、Dunlosky 与 Kornell（2013）。

[3] 不受控流形方法，见 Scholz 与 Schöner（1999）。

[4] 最小干预原理，见 Todorov 与 Jordan（2002）。

[5] 外开差额与自报损伤史，见 Coplan（2002）。

[6] 强迫外开与损伤的因果证据不足，见 Kaufmann 等（2021）。本章引用第 [5] 条差额数据，用途是说明读数结构（结果读数不带信息、差额带信息），不用于支持因果主张。

[7] 学习判断依据加工线索，见 Koriat（1997）。

[8] 走路时能耗的实时优化，见 Selinger 等（2015）。

[9] 该优化过程涉及内隐加工，见 McAllister、Blair 与 Donelan（2021）。

[10] 自然步态变异是否足以启动优化，见 Wong、Selinger 与 Donelan（2019）。

[11] 疼痛下活动的重新分配，见 Hodges 与 Tucker（2011）。

[12] 对认知需求的回避，见 Kool 等（2010）。

[13] 青少年术后跳距对称与落地力学，见 Wren 等（2018）。

[14] 习得性废用与约束诱导疗法，见 Taub 等（1993）。

[15] 捷径学习，见 Geirhos 等（2020）。

[16] 错误分析中的回避问题，见 Schachter（1974）。

[17] 自动化的讽刺，见 Bainbridge（1983）。

[18] 医院里的变通办法与第一序问题解决，见 Tucker 与 Edmondson（2003）。

[19] 导航使用与空间记忆，见 Dahmani 与 Bohbot（2020）。

[20] 合意难度，见 Bjork 与 Bjork（2011）。

[21] 肢体对称指数高估膝功能，见 Wellsandt、Failla 与 Snyder-Mackler（2017）。

[22] 生物系统中的冗余，见 Edelman 与 Gally（2001）。

[23] 第十一、十二节的全部读数由本章在写作过程中生成，代码与参数写在正文里，可复算；八组随机种子为 11、23、37、49、61、73、87、95。四条预测在运行之前写定。

另：文中若干处依据综述而非原始测量，包括使用依赖可塑性的一般表述、自动化阶段模型的当代评价、古德哈特定律的常见表述；聪明汉斯与差异化学习两处为教科书级通识，未回到原始报告。

人机分工声明

分工　选题与三个领域由王德生指定；文献检索与核对、论证组织、第十一节实验的设计与运行、成文与自我审查由 Claude 完成。所有引用的经验数据均来自公开发表的第三方研究，本章作者未采集任何人体数据；第十一、十二节的读数由本章自己生成，属计算实验，其局限已在 12.4 写明。本章未标注任何评分——按本站惯例，参与写作者不为自己写的文本打分。证据边界　舞蹈部分只使用「站出来的外开减去被动髋外旋所能提供的外开，这个差额在有伤组显著更大」这一横断结论，不主张因果；运动医学部分只使用关节做功分布与结果指标分离这一测量事实；心理学部分只使用「学习判断依据加工流畅感」与「当下表现与长期保持可以背离」两条。第二十一节八条判据，本章均未在人身上执行。字数　纯汉字约 2.5 万　版本　v1.0 · 2026-08-01 三边对「同一个结果由不同的内部路径达成意味着什么」给出互相排斥

的答案：运动控制说不影响结果的偏差不该纠正、多条路径就是韧性；古典舞教学说必须规定路径而不只规定结果，因为替代路径会固化并致病；学习心理学说主观掌握感由省力驱动，因此在这件事上系统性地站在错的一边。三家共有一个未说出的前提：路径的可用性是一个存量。而三家的材料合起来只能得出：它是一个流量，方向由使用决定，而任务对这个方向不提供任何回复力。

第六章

步骤留下，岔路删除

论选择次数何以在全部现行读数上不可见

学科入口　认知教育 × 伦理人类学 × 法哲学

摘要

现代组织正在以极高的速度取消人在工作中的选择，而所有现行的记录方式都读不出这件事。本章指出其原因不在监测不足，而在删除的形态：被取消的是岔路，不是步骤。岗位仍在，工序仍在，工作量甚至上升，唯一减少的是那些“本来可以走另一条”的位置，而这个量从未进入任何一张管理表。本章把它命名为带责分叉，并给出三条同时成立才算数的判据——两条路之后的走法确实不同、行动者知道它存在且被允许走、走错的后果落在他本人身上且不至于将他废掉；以单位工时内的带责分叉次数（分叉密度 b）与对系统给定值的偏离比例（越建议率 m）作为读数。本章进一步论证一个反直觉的命题：当代删除岔路的顺序，不由该不该删决定，而由可结算性——一条岔路两侧的条件与判据能否被写成规则、算成分数——决定；而可结算性与后果重要性是两条正交的轴，故最不该被删的判断恰恰因为最难被写下来而最晚被碰，且被碰时不会有任何一次公开讨论。本章在职业信息网络数据库（O*NET 29.1，n=879）上做了一次预注册的回溯检验（预注册文本哈希已公开）。四条预言中三条被支持：自动化程度与可结算性强相关（r=0.466），与后果重量在控制可结算性后完全无关（偏相关 −0.036 至 +0.047，均不显著）；时间压力与决策自由度近乎正交（r=0.017，p=0.61），证实忙与选是两笔账；自动化程度与决策频次无关（r=0.042，p=0.22）而与决策自由度显著负相关（r=−0.199），这一对系数正是“步骤留下、岔路删除”的字面读数。第三条预言部分不成立：三条判据中的前两条在职业尺度上几乎不可分（r=−0.803），本章据此收回“三条件在任何尺度上均可分离”的过强主张。一次纵向检验失败，失败原因可诊断并已写入正文。最后，本章指出这一读数使自动化设计学与人的监督条款陷入一处它们自身框架内看不见的矛盾：监督条款所要的接管能力只能由带责分叉生产，而自动化设计的标准实现方式恰好把它降到零；欧盟人工智能法第十四条要求部署方保证人的监督，却既不要求记录否决率，也不分配否决的

责任——这不是执行不力，是那个量在它的框架里不存在。关键词：带责分叉；分叉密度；可结算性；越建议率；自动化设计；人的监督；决策自由度

一、导论：三个不该同时为真的命题

有三个判断，各自都有扎实的领域证据，而它们不可能同时为真。第一个来自认知与教育研究：当两个框架在一个学习者面前对不上时，那道对不上的缝正是新维度得以生成的位置；用“综合起来看”“辩证地看”这类话术把它抹平，等于取消了这个位置。据此，那些看起来卡住、看起来什么也没发生的时段，恰恰是最该被守护的。第二个来自伦理人类学的田野诊断：在大量现场里被反复采集到的道德话语，流畅、完整、可叙说，却不承载任何实际抉择的后果；把这类话语当成伦理正在发生的证据，是二十年学术生产扑空的原因。据此，“看起来正在发生”恰恰是最需要怀疑的信号，而采集机器偏偏最爱这一种。第三个来自法哲学对规则类型的比较：划红线的禁止性规则在人身上安装的是一个间歇激活、可以休眠的零件，只在触界时耗能；而“应当积极展现某种品质”这类无上限的期待安装的是一个全天候运转、无法关闭的零件。规则不是在管理一个既成的人，而是先造出一种人，再用这种人去执行它自己。据此，争论外面有没有路是次要的，要紧的是里面那个走路的人还有没有余地。第一家要求守住那段“什么也没发生”的时间；第二家指出这类时间里的绝大多数根本不承载后果，守它等于守一个空壳；第三家则说前两家争的都是场景，而被改造的是主体。若第一家对，第二家的怀疑就是把养分当废料；若第二家对，第一家守护的多半是装饰；若第三家对，前两家的处方都会落空，因为它们默认了一个还有余地去选的人。本章的判断是：三家说的不是同一个东西，而把它们分开的那个量，三家都没有名字。这个量既不是“有没有事情在发生”，也不是“话语承不承载后果”，更不是“人有没有自由的感受”，而是一件可以数出来的事——在一段时间里，有多少个位置满足这样的条件：两条路之后的走法确实不同，行动者知道它存在并且被允许走，而走错的后果落在他本人身上且不至于将他废掉。本章称之为带责分叉。引入这个量之后，三家的判断各自变成一个特例：第一家守护的缝，是分叉尚未被合并的状态；第二家诊断的空转，是分叉的第三条不成立（说了不改变任何人接下来做什么）；第三家描述的常驻零件，是分叉的第三条虽在、而承受它的余地已被耗尽。三家都对，但都只按住了一角。本章要论证的核心命题有两条。其一，当代对选择的取消采取一种特殊形态——删岔路而不删步骤——因而在全部现行读数上不可见：工作量不变或上升，流程步骤不变或增加，考核指标全部正常，满意度甚至上升。其二，删除的顺序不由该不该删决定，而由可结算性决定；而可结算性与后果重要性是两条正交的轴。第五节将在一份公开的职业数据上对这两条做一次预注册检验，并报告其中一条对本章不利的结果。

二、文献综述与三家的定位

2.1 认知与教育侧：不可通约处作为生产性位置

这一支的核心主张是：系统性思维的最高形态不是在不可通约的选项间做出裁决，而是把裂缝本身当作认知立足点，从中生成原有框架都无法单独覆盖的新维度；其发生条件包括一份不可转让的完整责任，以及内部一切“翻译”与“调和”工具的暂时失效。它给出的教育处方是撤回填缝剂，并在评价体系中为“卡住”留出一块不受核算侵犯的地方。对本章而言，这一支贡献的是分叉的第一条判据的雏形：两条路必须真的不同。它同时提供了一个尖锐的观察——把两条路说成“其实是一回事”的话术，本身就是一种删除。它的限度在于：它没有给出判据来区分真正的不可通约与被装饰出来的不可通约，因而无法回应第二家的怀疑。

2.2 伦理人类学侧：承载与空转的分离

这一支的诊断是：在田野中被反复采集到的，不是伦理的发生，而是一种可以流畅叙说、却不承载任何抉择后果的道德话语；它进一步指出，学术机器的采集偏好（不出错、不浪费、不亏本）恰好与这种话语的形态咬合，于是二十年的丰收发生在纸面上而不是田野里。对本章而言，这一支贡献的是第三条判据的核心：承载与否要看它是否改变此后的走法与后果归属。它的限度在于：它把承载与否处理为话语的属性，而在工作现场，同一个动作可以在一个人身上承载、在另一个人身上空转——差别不在话语，在后果落在谁头上。

2.3 法哲学侧：规则安装何种主体

这一支比较了两类规则在个体内部建构的结构差异：禁止性规则以羞耻为燃料，仅在触界时激活，与其间歇性的执行压力形成正反馈；期待型规则以愧疚与外部验证为驱动，持续耗能且无法关闭，与其持续模糊的评价压力形成自吞噬。它据此提出：规则类型不是在管理既成的人，而是在制造一种特定主体类型，然后用这种人去执行它自己。对本章而言，这一支贡献的是第三条判据的下半句——承受得住。一个持续耗能而无法关机的主体，即使站在一个真实的分叉前，也可能没有余地去走那条更难的路。它的限度在于：它的分析单位是规则类型与主体类型，看不见“分叉数”这个介于制度与主体之间的量，因而无法解释为什么在规则类型完全没变的岗位上，选择也会一年一年地减少。

2.4 三家共有的前提

三家共享一个未被审查的前提：一个位置是否重要，取决于那里正在发生什么。第一家看是否有新维度在生成，第二家看话语是否承载，第三家看主体被安装了什么零件。三家都在问“里面有没有东西”，没有一家在问“接下来还有几条路”。本章的介入正是从这里开始。

三、冲突定位：同一段时间的三种不相容处置

设想一段在任何现行记录上都读作“正常”的工作时间：一名一线人员按流程完成了全部动作，没有异常，没有投诉，没有加班。三家对这段时间给出三种不相容的处置。第一家会问：这段时间里有没有出现两个对不上的框架？若有而被熨平，则损失已经发生，处置是撤回填缝剂。第二家会问：这段时间里的言说与动作有没有改变任何后续后果？若没有，则它是空转，守护它毫无意义，处置是不要把它当证据。第三家会问：这段时间里那个人身上运转的是哪一种零件？若是常驻型，则无论外部有多少选项，他都在持续耗能，处置是改变规则类型。三者的分歧不是侧重不同，而是彼此证伪：若这段时间的价值取决于是否有缝，那么第二家所谓的空转多数应当被重新评价为尚未成形的缝；若价值取决于是否承载后果，那么第一家守护的多数缝应当被判为装饰；若价值取决于主体类型，那么在同一套规则下的所有位置应当同质，而第一家与第二家在同一现场内部观察到的巨大差异就无从解释。本章认为三者可以同时为真，前提是承认它们各自按住的是同一个结构的不同条件，而这个结构此前没有被命名。

四、理论框架

4.1 带责分叉与两个读数

定义。在一段工作时间中，一个位置构成一次带责分叉，当且仅当以下三条同时成立：（一）分岔条件：至少存在两条后续走法，其在此后一段可指认的时段内产生不同的可观察后果；两条路的差别不能仅是同一走法的次序或措辞差异。（二）可见与许可条件：行动者知道另一条走法存在，且走它不需要支付使其在实践中不可选的成本（申请、说明、背书、得罪特定对象等的总成本）。（三）承担条件：所选走法的后果实际落在行动者本人身上，且该后果的量级不足以使其丧失下一次选择的资格。三条中任缺其一，从外部观察到的现象完全相同——都表现为“一个人面前有若干选项，他挑了一个”。这正是该量长期不可见的直接原因。读数一：分叉密度 b，定义为单位工时内带责分叉的次数，需按上述三条逐条判定，不接受自评。 读数二：越建议率 m，定义为在系统给出建议值、默认值或推荐方案的位置上，行动者作出不同选择的比例。m 是 b 的下界估计，其优点是几乎在所有已数字化的流程里都已经被记录，只是从未被读出。两条读数的关系是：m 只能测到第二条与第三条，测不到第一条（一个被改动的默认值可能通向同一后果）；b 需要人工判定第一条。故 m 便宜而粗，b 昂贵而准。

4.2 两种删除：删步骤与删岔路

删步骤是把一段工作取消：合并工序、去掉环节、裁撤岗位。它有明确的可见性——工作量下降、人数下降、时间下降，且删除时通常伴随争论。删岔路是把一段工作保留而取消其中

的选择：条件仍需核对，动作仍需完成，记录仍需留存，唯一消失的是“本来可以走另一条”。它的可见性接近于零：工作量不降反升（因为核对与留痕本身是工作），流程图上的节点只增不减，考核指标全部正常。本章的第一条核心命题是：当代对选择的取消以第二种形态为主，因而不进入任何现行读数。 这不是一个修辞上的区分。它有直接的经验蕴含：若删除以第一种形态为主，则自动化程度较高的岗位应当表现为决策次数下降；若以第二种形态为主，则决策次数可以不变，而决策自由度下降。第五节将检验这一对系数。

4.3 三种保留步骤而删除岔路的手法

第一，阈值化。把“要不要”改写为“分数够不够”。原来的问题需要一次判断，改写后的问题只需要一次核对。这一手法的正当性极强——它同时提高了公平性与一致性，压缩了看人下菜碟的空间——而它的副产物是：该位置的第一条判据被取消，两条路被合并为一条。第二，默认值预填。把空白栏改为带系统建议值的栏，并允许修改。此手法在形式上完整保留了第二条判据（另一条路可见且可走），但在实践中往往取消它：改动需要说明理由，理由须留档，留档者在事后追责中承担说明责任，而不改者不承担。一个“可改而改了要担责、不改则免责”的默认值，其通行成本被单方面抬高，越建议率会在数个季度内落到个位数。第三，例外申请化。把原本由现场人员当场决定的例外，改造为一个需要填表、说明、会签、备案的流程。没有任何一条规则禁止走这条路，只是这条路变贵了；贵到某个阈值之后，它从“一条路”变成“传说中有那么一条路”。三种手法的共同特征是：它们都不删步骤，且每一种在当期指标上都表现为改进。三者叠加时，会出现一种难以申诉的局面——一个人每天忙碌，所做的每一件事都是必要的，没有一件是他定的，而且他找不到任何一条禁止过他的规则。

4.4 可结算性排序：本章的靶心

一条岔路能否被上述三种手法处理，取决于它两侧的东西能否被写下来：条件可编码、选项可枚举、判据可计分。本章把这一属性称为可结算性。可结算性与后果重要性是两条不同的轴，且在经验上近乎正交：一名装配工的动作高度可结算而后果有限；一名医生对“这个病人今天不对劲”的判断后果极重而几乎不可结算；一名信贷员对“报表难看但这个人靠谱”的判断介于两者之间。本章的第二条核心命题是：删除的顺序按可结算性排列，与后果重要性无关。 其推论有三：其一，最不该被删的判断，恰因其最难被写下来而最晚被处理；但“最晚”不等于“不会”，而当它被处理时，不会有任何一次以“该不该”为议题的公开讨论——因为触发处理的是技术可行性，不是必要性判断。其二，这一顺序在单个岗位内部同样成立。一个岗位中可结算的那部分被优先接管，而不可结算的那部分所依赖的判断力，本来正是靠反复经历前一部分的小判断养成的。于是“保留高价值判断、自动化低价值判断”这一被广泛接

受的设计原则，在时间上是自我拆解的。其三，由于删除按可结算性排序，而记录能力同样按可结算性排序，因此被删掉的东西与能被记录的东西是同一批的补集：越是无法被记录的判断，其消失越无从察觉。这解释了为什么加强监测无法发现这一过程——监测本身沿着同一条轴展开。

4.5 三家为何各自到不了这一条

认知与教育侧到不了，因为它的分析单位是学习者与知识内容，而删除发生在流程设计层；在它的视野里，填缝剂是话术，它看不见填缝剂的工业化形态是默认值。伦理人类学侧到不了，因为它把承载与否处理为话语的属性；一旦把承载定义为“后果是否落回言说者”，它就会发现自己诊断的空转不是学者的采集偏好造成的，而是现场的分叉早已被删除——话语之所以空转，是因为那里已经没有什么可以由这次言说决定。法哲学侧到不了，因为它的两类规则都以“规则”为前提，而本章所述的三种手法都不是规则：阈值是参数，默认值是界面，例外申请化是流程。它们不禁止任何事，因而不进入规则类型学的视野。

4.6 一次完整走查

为使三条判据可被他人重复施用，此处对一个具体岗位逐条走一遍。取一家中型银行的小微信贷客户经理，改造前后各一次。改造前。一天处理四至六户。位置一：是否受理该户（判据一：受理与否此后一个月的走法完全不同；判据二：不受理无须审批，说明成本近零；判据三：受理后逾期计入其个人资产质量考核——三条成立，计一次）。位置二：贷前调查中是否追加一次实地核查（三条成立，计一次）。位置三：给出的额度与期限建议（三条成立，计一次）。位置四：报送时是否附加与评分不一致的说明（判据二成立但成本较高，判据三成立，计一次）。全天 $b \approx 4$，越建议率不适用（无系统建议）。改造后（引入评分卡与自动决策）。同一岗位一天处理十二至二十户。位置一被阈值化：分数过线自动受理，不过线自动退回，中间档转人工——原分岔在两端被合并，仅中间档保留，而中间档约占全部件量的一成五。位置二被规则化：实地核查由系统按规则触发，人不再决定是否追加。位置三被默认值预填：额度与期限由系统给出建议值，可改，改需填写理由并留档，改后若发生逾期，责任认定中该理由将被调阅。位置四被流程化：与评分不一致的报送须经风险条线会签。改造后的读数：全天名义决策次数由四至六上升到十二至二十（判定动作变多），而按三条判据计的 b 由约 4 降至约 0.6（仅中间档的额度建议仍成立，且其判据二的成本已升高）；越建议率 m 在该行上线后第七个季度稳定在 3% 至 5%。三点值得注意。其一，该岗位的工作量、留痕量、合规水平均上升，且其不良率下降——从任何一张现行管理表看，这次改造是成功的，本章亦不认为它应当被撤销。其二，b 的下降与不良率的下降并不矛盾：可结算的那一部分判断，系统

确实做得更好。其三，被删掉的那一部分——“报表难看而这个人靠谱”——恰是小微信贷中最难被写下来、也最被认为是这一行当价值所在的判断；它的消失不是被决定的，是被排序排到的。

五、实测：一次预注册的回溯检验

5.1 数据与操作化

检验使用美国职业信息网络数据库（O*NET）29.1 版的工作情境模块（Work Context，Scale ID = CX，五级量表），分析单位为职业，完整个案 n = 879。选择这份数据的理由是：它对本章所需的三类量都有独立测量项，且这些测量项由在岗者本人评定，早于本章的问题意识多年，不可能为迎合本章的假说而设计。操作化全部采用原始描述符，不做任何合成加权（唯一的合成是可结算性由两个描述符等权秩均，且此合成在预注册中写定）：- 自动化程度 A ← 4.C.3.b.2 Degree of Automation- 例行度 R1 ← 4.C.3.b.7 Importance of Repeating Same Tasks- 结构化程度 R2 ← 4.C.3.b.8 Structured versus Unstructured Work（反向计分）- 后果重量 W1 ← 4.C.3.a.2.a Impact of Decisions on Co-workers or Company Results- 差错后果 W2 ← 4.C.3.a.1 Consequence of Error- 人身安全责任 W3 ← 4.C.1.c.1 Responsible for Others' Health and Safety- 决策频次 F ← 4.C.3.a.2.b Frequency of Decision Making- 决策自由度 L ← 4.C.3.a.4 Freedom to Make Decisions- 时间压力 T ← 4.C.3.d.1 Time Pressure

对应关系需说明其限度：L 是第二条判据（可见与许可）的近似，R2 反向是第一条判据（分岔）的近似，W1/W2/W3 是第三条判据（承担）的近似。三者都是职业层的粗测，无法测到工位与日次层。这一限度在 5.5 节将以一个不利结果的形式回来。

5.2 预注册

四条预言、操作化与判否规则在读取任何相关系数之前写定并冻结，文本的 SHA-256 为：c12a5830a4e26d0e565c329e2eb84919ca2bc38564ac0825b9a78ba7930611b4 预言如下：P1（靶心）删除顺序按可结算性而非后果重量：r(A,R) 显著为正且明显大于 |r(A,W)|；控制后果三项后 R 与 A 的偏相关仍显著为正，控制可结算性后 W 与 A 的偏相关不显著。判否：若 $|r(A,W)| \geqslant |r(A,R)|$，或控制 R 后 W 仍是 A 的强预测项。P2（两笔账）$|r(T,L)| < 0.20$，且 r(A,T) 不显著为负。判否：若 $r(T,L) \leqslant -0.40$。P3（三条件不可归约）L、R2、W1 在主成分分析中第一主成分解释方差 < 80%，且 W1 在第二主成分上有主要载荷。判否：若第

一主成分 ⩾ 80% 且三者同号高载荷。 P4（频次不等于自由）r(F,L) 与 r(F,A) 方向可分，且 F 与 A 的负相关弱于 L 与 A 的负相关。判否：若 r(F,L) ⩾ 0.80。

5.3 结果一：删除顺序按可结算性排列（P1 支持）

自动化程度与可结算性两项的相关分别为 r(A,R1) = +0.466（$p < 10^{-47}$）与 r(A,R2) = +0.182（$p < 10^{-7}$），合成后 r(A,R) = +0.437。自动化程度与后果三项的相关则微弱：r(A,W1) = +0.071（p = 0.035）、r(A,W2) = +0.189、r(A,W3) = +0.036（p = 0.29，不显著）。偏相关给出更干净的分离。控制全部三项后果变量后，例行度与自动化的偏相关几乎不衰减：r(A,R1 | W1,W2,W3) = +0.447；反过来，控制两项可结算性变量后，三项后果变量与自动化的偏相关全部落到零附近且均不显著：r(A,W1 | R1,R2) = −0.036（p = 0.28）、r(A,W2 | R1,R2) = +0.047（p = 0.16）、r(A,W3 | R1,R2) = −0.048（p = 0.15）。这一组系数的含义是明确的：在职业层面上，一件工作被自动化到什么程度，由它有多可被写成规则决定；至于这件工作的判断错了会伤害多少人、会造成多大损失、是否涉及他人的健康与安全——在控制了可结算性之后，这些与自动化程度没有关系。靶心命题在这份数据上成立。需要立刻说明它不能证明什么。这是横截面数据，不能证明因果方向，也不能直接证明“删除按此顺序发生过”。它只能证明本命题的一个必要蕴含：如果删除顺序确实由可结算性而非后果重量决定，那么在任一时点的截面上就应当看到上述模式；若看到的是相反模式，本命题即被推翻。

5.4 结果二：忙与选是两笔账（P2 支持）

时间压力与决策自由度的相关为 r(T,L) = +0.017（p = 0.61）——在近九百个职业上，这两个量近乎完全正交。一个人有多忙，与他有多少可自由处置的余地，在职业层面上互不预测。自动化程度与时间压力的相关为 r(A,T) = +0.178（$p < 10^{-6}$），方向为正。也就是说，自动化程度更高的职业，其在岗者报告的时间压力并不更低，反而略高。这与“删步骤”的图像不符，与“删岔路而保留甚至增加步骤”的图像一致。

5.5 结果三：一处对本章不利的结果（P3 部分不成立）

主成分分析的判否规则未被触发：第一主成分解释 67.5% 的方差（低于 80% 的门槛），后果重量 W1 在第二主成分上是主要载荷（−0.904）。就此而言 P3 通过。但相关矩阵显示了一处本章事先没有预料到的事实：r(L, R2) = −0.803。也就是说，第二条判据（行动者知道并被允许走另一条）与第一条判据（两条路真的分开）在职业尺度上几乎是同一个量的两端。三条判据中，真正可分的只有第三条（后果落在本人身上）。本章据此收回一条过强的主张。原先的表述是“三条判据各自独立、缺一不可”；准确的表述应当是：在职业这一粗尺度上，前

两条不可分离，可分离的是承担条件；三条判据的分离需要更细的观测单位（工位级或日次级的越建议记录），而这一层数据在绝大多数机构里虽然天天产生却从未被读出。这一收回不是无关痛痒的：它意味着本章提出的 b 读数在职业层面无法完整实施，只能实施其两条；完整实施依赖于本章所主张的那种记录尚未建立——本章由此处在一个必须承认的循环里，第十二节将回到这一点。

5.6 结果四：步骤留下，岔路删除（P4 支持，且强于预期）

三个系数放在一起是本章最直接的量化支持：- r(F, A) = +0.042（p = 0.22，不显著）：自动化程度与“做决定的频次”没有关系。- r(L, A) = −0.199（$p < 10^{-8}$）：自动化程度与“做决定的自由度”显著负相关。- r(F, L) = +0.333：频次与自由度可分，并非同一个量的两种说法。

即：随着一个职业的自动化程度上升，在岗者仍然在不断做决定，做决定的次数并不减少，减少的是这些决定还剩下多少可以由他处置的余地。这正是本章所命名的那个形态的字面读数——步骤没有少，岔路少了。按此可以点出一份职业名单：决策频次在中位数以上、而决策自由度落在下四分位的职业共 77 个，占样本的 8.8%，其中包括停车执法员、外科技师、细胞遗传技术员、肉类分割工、金属与塑料挤压拉拔机操作工。这些岗位的共同特征是：一整天都在做判断，而几乎没有一项判断是可以由他改的。在决策自由度与后果重量交叉的 2×2 分布上，“低自由度·重后果”一格共 160 个职业（18.2%），其自动化程度均值为 2.24，高于“高自由度·重后果”一格的 2.08（差 0.162）。这一格正是责任与决定权分离最严重的位置，也是自动化程度最高的一格。

5.7 一次失败的纵向检验，及其可诊断的失败原因

若能在时间维度上直接观察删除，证据会强得多。本章尝试了这一步并失败了，失败过程照实报告。方法是取 O*NET 20.1 版（2015）与 29.1 版（2025）的同一批描述符，保留两版工作情境采集日期不同（即期间确实被重新调查过）的职业，n = 578，计算自动化程度的变化量 ΔA，并以基期特征预测之。结果不可用，理由有二。其一，ΔA 的均值为 −0.121——在这十年间，在岗者报告的自动化程度平均是下降的。这与任何关于该时期的常识相悖，说明该题项作为纵向指标存在系统性漂移：它是自评题，问的是“你的工作有多自动化”，而重新调查抽到的是不同的在岗者，其参照系随时代变化（今天的在岗者以更高的标准判断什么算“自动化”）。其二，基期例行度与 ΔA 的相关为 −0.195（$p < 10^{-5}$），方向与横截面结果相反；在多元回归中基期例行度系数为 −0.116（t = −4.54），而基期后果重量与差错后果的系数均不显著（t = −0.33 与 +0.02）。基期水平与其后变化量之间的这种负相关，是均值回归的标准形

态，无法与实质效应分离。结论：本章不能用 O*NET 的时间维度支持删除顺序命题。 真正的纵向检验需要一个不依赖自评的自动化度量——例如同一批流程在两个时点上的越建议率与例外申请通过率。这类数据在每一个已上线的业务系统里逐日产生，而本章尚未在任何一份公开发表的研究或任何一家机构的年报中找到它被读出的实例。这一句本身就是本章命题的一个侧面证据，同时也是它最难堪的位置：它所需要的证据，正被它所描述的机制持续地不予记录。

5.8 稳健性：两个替代代理，一支持一反向

由于自评式的“自动化程度”存在上述弱点，本章另取两个不依赖该题项的代理重跑靶心检验。需注明：这一步不在预注册之内，属事后分析，结论只作为稳健性参考。代理一，“工作节奏由设备速度决定”（4.C.3.d.3）：r 与例行度 +0.215、与结构化 +0.432、与后果重量 −0.047（不显著）、与决策自由度 −0.397。靶心保持，且比原代理更干净——它与可结算性正相关、与后果重量无关、与自由度强负相关。代理二，“与计算机打交道的重要性”（4.A.3.b.1）：结果方向相反——与结构化 −0.410、与决策自由度 +0.337。据此该代理不支持靶心。本章的判断是这个代理测的不是本章所指的东西：它与决策自由度正相关，说明它捕捉的是“以电脑为主要工具的专业性工作”（工程师、分析师、设计师），而不是“动作被系统规定的程度”。本章承认这一判断带有事后合理化的风险，故给出一条可事前区分的判据供他人检验：一个合格的“被系统中介程度”代理应当与决策自由度负相关；凡与自由度正相关者，测的是工具的种类，不是路径的收窄。综上，四条预注册预言中，P1、P2、P4 被支持，P3 部分不成立并已导致本章收回一条主张；纵向检验失败且失败原因可诊断；两个事后代理中一个支持一个反向。

六、强制不一致：为什么不接受这个读数的一方会自相矛盾

本节论证：一旦承认分叉密度这个量，当代关于自动化的主流设计学与其配套的监管条款，会在它们自身框架内暴露一处此前看不见的矛盾；而拒绝承认这个量，则该矛盾无法被表述，只能以“执行不到位”的形式反复出现。

6.1 甲：自动化的设计目标

自动化设计的基本处方是减少人的操作负担、减少人为差错的机会、提高一致性。这一处方在几乎所有工程领域被采纳，并且它是对的——本章第十节将说明它在哪些场合不但对而且必须。其标准实现方式正是本章第 4.3 节所列的三种手法：把判断阈值化、把默认值预填、把例外流程化。

6.2 乙：人的监督与接管能力

与甲并行的另一条处方来自人因研究的经典结论：高度自动化的系统中，人被置于监控位，其技能与情境把握会退化，而恰恰在系统失效时需要他接管。因此设计上必须保留人的监督与否决权。这一条同样被广泛接受，并且已经进入法律：欧盟人工智能法（Regulation (EU) 2024/1689）第十四条要求高风险人工智能系统的部署方保证由自然人实施有效监督，包括正确解读输出、在必要时不采纳输出、以及在必要时干预或停止系统。

6.3 相撞点

甲与乙在“分叉密度”这个量上正面相撞：乙所要的那种接管能力，其生成条件是带责分叉——一个人只有反复地在真实分岔处作出选择并承担其后果，才会形成对“哪里容易出问题”的判断；而甲的三种标准实现方式，恰恰逐条取消带责分叉的三个条件（阈值化取消第一条，默认值预填抬高第二条的成本，例外申请化把第二条抬到不可选）。因此：越是按甲把系统做好，乙所要的能力就越无从生成；而乙的处方是在能力已经无从生成之后，要求那个人去行使它。现行的化解方式是“保留人的否决权”。但按本章的三条判据，一项否决权若不同时具备可见的替代路径与后果的实际归属，就不构成一次分叉，只构成一个签字位。这不是修辞：它可以被直接检验——凡真实的否决权，必有非零的越建议率；凡越建议率长期趋近于零的位置，其上的“人的监督”在功能上已经是形式的。

6.4 这处矛盾已在现行处方里发作

上述法律条款要求部署方保证有效的人的监督，规定了监督者应当能够不采纳输出、能够干预与停止；但它既不要求记录不采纳的比例，也不要求规定不采纳之后责任如何分配。同一部条款要求保存日志（第十二条），日志的内容围绕系统的运行与可追溯，而不是围绕人的偏离行为。于是出现一个可以直接核对的空档：一套以“人必须能够不采纳”为核心要求的制度，不产生任何关于“人有没有不采纳过”的数据；而它同时要求部署方为系统输出的使用后果负责，却不说明当监督者选择不采纳并因此致损时责任如何落位。按本章的框架，这两点是同一件事：没有承担条件的否决权不会被行使，而没有被行使的否决权不产生记录，于是这套制度在自己的读数里永远看不到它自己是否有效。这就是本章所说的强制不一致。承认分叉密度，则上述条款的合规性可以被检验（要求报告越建议率并规定否决的责任分配），代价是许多现行的“人在环中”设计将被判定为形式合规；不承认它，则该条款的有效性在原则上不可被评估，而其失效将持续以“人员培训不足”“监督意识不强”的形式被归因于个人。

6.5 退路与代价

拒绝本章的一方有一条明确的退路：主张接管能力可以由训练而非由日常的带责分叉生产——即在模拟器、演练与考核中造出分叉，从而无须在真实流程中保留。本章认为这条退路是

成立的，并且这正是高可靠性行业已经在做的事（第八节其十二与第十节）。但它有一个代价必须一并承认：模拟中的分叉满足第一与第二条判据，其第三条只能以人工方式模拟（考核后果），因而它能训练出的是识别与操作，未必是“在后果真的落在自己身上时仍能作出判断”的那种能力。若有研究能证明模拟分叉在这一点上与真实分叉等效，本章第六节的全部论证即被削弱。这是本章最愿意被证伪的一条。

七、推论与判据

7.1 五条推论

其一，加大投入不能修复分叉密度。若某岗位的 b 已趋近零，增加培训、增加人手、提高薪酬、强调担当意识，都不改变三条判据中的任何一条。可检验的形态是：这类干预会提高满意度与自评自主感，而越建议率不动。其二，装饰性选项与真实分叉此消彼长。可批量供应的只有不承载后果的选项，因为供应它们不需要分配责任。故在同一系统中，选项数量的增长与分叉密度的下降可以同时发生，且前者会被当作后者的证据。判据：数选项个数无用，数越建议率。其三，同一技术在余量厚薄不同的位置产生相反效果。第三条判据含有“后果不至于使其丧失下一次选择的资格”这一限定，故当承受能力低时，形式上存在的分叉在实践中被压成单路。推论是：同一套辅助决策系统，在有余量的岗位上是省事，在无余量的岗位上是收路，而两者在系统日志上无法区分。其四，责任感会与选择脱钩，而不是随之消失。当第三条判据与前两条断开（后果落在身上但没有可选路径），人身上关于“我够不够好”的评价不会停机，却没有可以用来结案的材料。可检验的形态是：这类岗位上会同时观察到高疲劳自评、低自评效能与正常甚至优良的绩效指标。这一形态与倦怠的标准描述部分重叠，但可由越建议率区分：倦怠的标准处方（减负、恢复、社会支持）不改变越建议率。其五，事故复盘会卡在同一处。在 b 趋零的岗位上发生事故后，追问“当时为什么这样处理”将得不到个人化的理由，回答会收敛为“系统就是这么派的”“流程就是这么走的”。这不是推诿：那里确实没有发生过一次选择。

7.2 四条判据

判据一（分岔）：选另一条，其后一段可指认时段内的可观察后果是否不同？答不出具体差别的，不是分叉。 判据二（可见与许可）：另一条路在制度上是否可走，其通行成本是多少？成本须以实际支付计（时间、说明、会签、得罪对象），不以是否被禁止计。 判据三（承担）：后果落在谁身上，量级是否使其丧失下一次的资格？ 判据四（记录）：这三条的判定所

依据的数据，是否在该机构已经产生？若已产生而从未被读出，则该机构处在本章所述的典型状态。

7.3 2×2 诊断格

以决策自由度与后果重量分别切分，得四格：高自由·重后果（本章样本 32%）是分叉正常的位置；高自由·轻后果（18%）是低风险练手位，是唯一可以廉价制造分叉的地方；低自由·重后果（18%）是责任与决定权分离最严重、且自动化程度最高的一格，本章认为这是最应优先干预的靶格；低自由·轻后果（31%）是常规执行位，删除分叉在这里通常是正确的。

八、跨领域兑现：十二处

本节检验该读数在差异较大的领域中是否仍能作出非平凡的区分。每处只写形态与可查的读数。其一，门诊。指南细化与临床路径管理提高了一致性，同时使“这个病人不像指南所描述的那样”的处置需要额外记录与说明。可查读数：偏离临床路径的比例及其逐年走向；本章预测该比例单调下降而路径覆盖率上升。其二，中小学课堂。统一进度与配套练习使课时内容高度确定，教师改变当日安排须向多个方向解释。可查读数：一学期中实际偏离既定进度的次数。其三，基层执法。裁量基准表收窄自由裁量，这在总体上是进步；其副产物是“依法可罚而此次不宜罚”的处置失去合法安放处。可查读数：不予处罚决定占比与其理由的类型分布。其四，小微信贷。见 4.6 节走查。可查读数：越评分卡建议的比例。其五，工厂排产。排产系统在多数条件下优于经验判断；老工人“这批料手感不对”的处置在系统中无对应字段。可查读数：人工干预排产的次数占比。其六，客服。话术由系统推送，考核解决率与时长。可查读数：坐席偏离推荐话术的比例——该数据几乎所有系统都记录，几乎无人调阅。其七，配送与网约。路线、顺序、时限由系统给定。可查读数：骑手主动改派或改序的比例及其被系统接受的比例。其八，家庭作业。家长即时纠错使正确率上升，孩子未经历“我打算这么做、错了自己承担”的位置。这是分叉密度最早开始下降的场所，且无任何记录。其九，招聘。筛选规则明确化挡住了偏见，也挡住了“这个人经历很怪但值得见一面”。可查读数：越筛选规则进入面试的人数占比。其十，编辑与内容生产。选题由数据支撑，稿件有模板。可查读数：无数据支撑而被采纳的选题占比。其十一，社区照护与物业。巡查表与拍照留痕使覆盖可核查，未列项者不进入视野。可查读数：巡查记录中“表外事项”的登记条数。其十二，飞行与消防训练。这是反向的一处，也是最有说服力的一处：这些行业在演练中刻意制造信息不全、指挥中断、必须当场决断的局面，并明确规定此类判断即使事后被证明为错，复盘不追责。它们清楚地知道那种能在无指令时作出决断的人不会自行长出来，必须专门花钱造出场合把他逼出来。可查读数：年度演练中设置为“无标准答案”的科目占比。十二处的形态一致：步骤保留或增

加，分叉减少，减少的顺序自可结算性一端开始，且每一处的删除都有充分的好理由。最后一处说明这不是宿命——有人在专门往回加；只不过目前愿意为此付费的，主要是那些失败会以外人可读的形式一次性出现的行业。教育、照护与多数白领工作的失败缓慢、可归因于个体、延迟数年，这正是它们的分叉密度结构性偏低的原因，也是“向高可靠性行业学习”在这些领域反复落不了地的原因。

九、与十位近邻的划界

10.1 例行任务替代假说。这是离本章最近的一家，也是最危险的一家：它主张可被明确规则描述的例行任务优先被技术替代。表面上，本章的靶心与它同义。分离线有两处。其一，单位不同：该假说的单位是任务与就业量，讨论的是岗位是否被替代；本章的单位是岗位内部的选择次数，讨论的是岗位保留而分叉消失。其二，两者给出可分辨的相反预测：该假说预测被替代的部分退出，故其残余岗位的决策频次应当上升或至少不降而自由度不降；本章预测决策频次不变而自由度下降。第 5.6 节的两个系数（r(F,A) = +0.042 与 r(L,A) = −0.199）落在本章一侧。若未来数据显示频次随自动化显著下降而自由度不变，本章错而该假说对。

10.2 自动化悖论。它讲人被置于监控位后技能退化、需要接管时反而做不了。分离线有二：其一，它有一个明确的接管时刻，本章所述的过程没有任何提示音；其二，它的量是技能水平，本章的量是选择次数——一个人可以手艺极好而分叉密度为零。经验上，已有研究报告自动化程度提高后飞行员的手动操作技能保持完好，退化的是判断与状态把握，这一分离与本章的两轴一致。

10.3 去技能化与劳动过程理论。该传统预设构思与执行的分离是被有意剥夺的结果。本章不需要这一预设：分叉可以在无人被剥夺、亦无人作过决定的情况下消失，因为它是最优路径计算的副产物。差别有实践后果——若无对手，则谈判与问责均不适用，只有记录与设计能起作用。

10.4 选择过载。该传统讲选项过多导致决策质量下降。本章与它不冲突且互补：可批量增发的正是不承载后果的选项，故选项过载与分叉稀缺可以同时成立，且前者遮蔽后者。其处方（减少选项）对本章所述问题无效，因为被减掉的是廉价那一类。

10.5 选择架构与轻推。这是本章最近的制度性近邻。该路径承认默认值强烈影响行为，主张把默认值设成对人更好的一个，其评价标准是结果。本章的标准不同：默认值的长期代价不在它把人推向哪里，而在越建议率的单调下降及其不可逆。两者可同时成立，并有一条判据可分：该路径的研究几乎从不测量越建议率随时间的走向，因为在它的框架里那不是一个变量。

10.6 自我决定论中的自主。那边测量的是人感受到的自主，本章测量的是结构上是否存在带责的另一条路。二者可系统性背离，且默认值设计得越贴合个体，背离越大。7.1 推论一给出了可检验的背离形态。

10.7 可行能力进路。它问一个人是否有实质自由去过他有理由珍视的生活，尺度是一生与总体。本章把尺度缩到一个工位与一天，并给出三条可判定的条件。二者是同一关切的不同粒度，本章可视为其在工作现场的一个可操作化。

10.8 心理安全感。它讲敢不敢开口，本章讲有没有第二条路。一个团队可以气氛良好、人人敢言，而所有人面前只有一条路；反之亦然。判据：敢言不改变越建议率。

10.9 路径依赖。它讲历史锁定使某条路径被选中，是历时的与宏观的；本章是横截面的与工位级的，讨论当下每一步还剩几个可选项。

9.10 为便于治理而简化。该传统讲管理者为使对象可清点、可征税、可调度而重排现实，代价是不入表格者被抹去。本章欠它最多，分离线在于：那边被简化的是被管理的对象，本章被收窄的是做事者的路径；且执行者往往不是任何管理者，而是一套算得比谁都好的推荐。没有可谈判的对手，是本章与它最大的差别。

9.11 站内划界。学科通融专栏此前一篇讨论“无要求时段的自发变异”，其读数以无任务、无奖惩为定义条件；本章的读数恰恰要求后果落在本人身上，二者在定义上互斥而正交：那一个测的是没人要求时它还动不动，本章测的是有人要求时他还能不能走另一条。两者可同批测量且互不蕴含。另一篇讨论决定权与责任是否捆在一起，本章在其上补一步：即便二者捆得严丝合缝，若面前只有一条路，捆住的也是一个空的选择。

十、边界与不适用

10.1 四类应当删除分叉的场合。其一，时间窗极短的场合（抢救、撤离、失火）——这里需要的是被背下来的顺序，不是多元判断。其二，错误不可逆的场合（麻醉剂量、儿童用药换算、高压与剧毒作业、结构配筋）——规则的平均水平明显高于个人判断的平均水平。其三，新手的最初阶段——此时分叉是噪声不是养分；但此处必须附加一个到期日，否则“照做就行”会一直延续到退休。其四，巨量同类琐事——每件都判断等于什么都判断不了。

10.2 一个必须留下的人。任何删除分叉的系统都应留下一个有名字、有时间、有权限的人，其职责明确包括：定期清点哪些判断已被收走，并判断哪些应当还回去。理由是这一过程具有强单向性：删除时有推动力（省钱、提效、合规），恢复时没有推动力，且其收益在数年后才显形，而彼时当初主张恢复的人多半已不在该岗位。

10.3 本章不主张的三件事。本章不主张自动化程度越低越好；不主张个人应当抵制系统建议；不主张“人的判断优于系统判断”——在可结算性高的位置上，后者通常更准也更公平。本章只主张一件事：删除的顺序目前不由该不该删决定，而这个顺序应当被有意识地选择，且需要一个读数才能被选择。

十一、可证伪条件与赌注

十条中前七条为证伪条件，第八、九条为最便宜的检验，第十条为写死的赌注。其一，若在职业或岗位层数据上，后果重量在控制可结算性后仍是自动化程度的强预测项，则第 4.4 节的靶心命题伪。 其二，若决策频次随自动化程度显著下降而决策自由度不变，则“步骤留下、岔路删除”这一形态判断伪，例行任务替代假说对。 其三，若时间压力与决策自由度在多个独立数据源上强负相关（$r \leqslant -0.40$），则“两笔账”的分账是多余的。 其四，若在工位级数据上，三条判据坍缩为一个维度（第一主成分解释方差 $\geqslant 80\%$），则三条判据应被合并，本章的判据体系失效。第 5.5 节已在职业层给出部分不利证据。 其五，若有研究证明模拟场景中的分叉在“后果真的落在本人身上时仍能作出判断”这一点上与真实分叉等效，则第六节的强制不一致失效。 其六，若某高度自动化的行业长期不出现“遇到新情况全场无人能给出判断”的现象，反而因省下的精力被投入更难的判断而使分叉密度上升，则本章对趋势的判断错误——这是本章最希望被证明的一条。 其七，若某机构在越建议率长期趋零的条件下，仍能在真实事故中由现场人员作出有效的非常规处置，且这一处置能力可被重复观察，则第三条判据的承担条件并非必要。 其八（最便宜的检验，一条数据库查询）：在任何已上线的业务系统中统计“系统给出建议而人作出不同选择”的比例及其逐季走向。本章预测该比例在系统上线后的

六至八个季度内单调下降并稳定在个位数，且这一下降与满意度指标无关。其九（一天可做的对照）：同一份工作、同一套系统，一组默认值全部预填（可改），另一组关键字段留空（须自填）。本章预测三个月后后者处理速度略慢、当期差错率略高，但在遇到系统未曾见过的情形时给出可用判断的比例显著更高，且能说出自身判断理由的人数显著更多。若后一项无差别，本章核心主张塌掉一半。其十（赌注，2031-12-31）：到该日，若仍没有任何一部关于人工智能系统人的监督的法规、标准或行业指引，要求报告否决率或规定否决的责任归属，则本章所描述的量仍然只是一个说法，而不是一件被承认的事。以下六种情形不计为命中：仅要求保留否决权而不要求记录；仅要求记录系统日志；仅在自愿性指南中提及；仅要求统计人工复核的数量而不区分是否改变结论；仅适用于单一企业内部；以及以满意度或信任度问卷替代行为记录。

十二、自反

第一，本章所需的关键证据，正被它所描述的机制持续地不予记录。第 5.7 节的失败不是技术性挫折，而是本命题的一个侧面：一个从未被读出的量，其历史无法被重建。本章因此只能给出横截面的必要蕴含与两条便宜的前瞻检验，不能给出直接的过程证据。第二，本章第 5.5 节被数据削掉了一条主张，而削掉它的是本章自己预注册的判否规则。这一条应当被后来者用同一方式对待：若工位级数据显示三条判据进一步坍缩，本章的判据体系应当被简化，而不是被解释。第三，本章自身的分叉密度接近于零。它有指定的栏目、指定的体例、指定的交稿时刻，作者在其中所作的判断很少有一次的后果会落回自身。按本章自己的判据，学术论文与它的生产制度正是“步骤极多而分叉极少”的典型；本章能诊断这一状态，不能示范它的反面。第四，一个无法自解的循环：任何使分叉密度可核查的装置，本身都会成为一次新的要求，从而改变被测者的行为；而不可核查的分叉密度进不了任何决策。本章对此没有出路，只有一条底线——验收时看事后清点的实际读数（越建议率、否决后的责任归属记录），不看有没有设立制度。第五，交出两个本章解释不了的地方。其一，b 的合理量级未知：本章能说零是坏的，说不出一天该有几次，也无法把“一次分叉的重量”变成可加总的量；本章怀疑重量比次数更要紧，但没有走通。其二，分叉密度长期为零之后能否自行恢复：本章主张单向性很强，但确有个案显示换一个环境后判断力在数月内回来，这与“能力已不存在”的解释不相容；两种情形的处方完全相反，本章分不出它们。

注释

〔1〕本章使用的 O*NET 数据为 29.1 版公开发布的文本版数据库，由美国劳工部资助的国家 O*NET 开发中心发布，采用知识共享署名协议。分析脚本与预注册文本随本章发布。

〔2〕工作情境模块的量表为五级，由在岗者评定；本章所用各题项的样本量与标准误随原始数据一并公开，本章未作加权、未剔除任何样本、未使用抑制标记之外的任何筛选。〔3〕第5.8 节的两个替代代理不在预注册之内，属事后分析，其结论仅作稳健性参考。本章保留其中与靶心相反的结果并给出可事前区分的判据，而非将其略去。〔4〕本章对欧盟人工智能法条款的引述限于其公开文本中关于人的监督与日志保存的要求，不涉及对其成员国实施细则的评价。

附：与学科通融专栏另一篇的关系

学科通融专栏另有一篇与本章形状相近而变量相反：《没有人要求的时候》测的是无任务、无奖惩时段里系统还动不动，其定义条件是没有人在等它交出什么；本章的读数恰恰要求后果落在本人身上。两者在定义上互斥而正交——那一个问「没人要求时它还动不动」，本章问「有人要求时他还能不能走另一条」，可同批测量且互不蕴含。本章的三家源是站内学员论文，分属认知教育、伦理人类学与法哲学——三家互不相识，却对「同一段看起来什么也没发生的工作时间」给出不能并存的处置。下面三条可直接点开核对。正文第二节交代三家各自说到了哪一步，第三节把三处对顶写死，第四节第五小节说明它们各自为什么到不了本章这一条，第九节再与十种最容易被混为一谈的说法逐条分界——包括本章承认最可能把它吸收掉的那一个。

第三编

只在那一段里长出来的东西

• • •

编　序

第二编说明被跳过的是哪一段。这一编问的是：那一段里本来会长出什么。

两章各指认一样，两样都有同一个性质——它们不是知识，也不是能力，而是一笔功夫或一个位置，因而在任何以人或以产物为单位的账上都没有栏目。

第七章是产物侧：一份规范只写了什么算同类，没有写同类之间有多近。压缩后者的那笔功夫不会因为无人付钱而消失，它只会换一个地方发生，并在那里得到三个不同的名字——工程里的成本、艺术里的风格、语言里的变异。第八章是共同体侧：被记为零的那类动作有一个共同特征，它的成果是一个否定，是某件事没有发生；能力仍然会在个人身上正常沉淀，然后在共同体层面消失，而消失的过程不出现在任何指标上。

两章合起来构成合章第三处锁与第四处锁的两端。第八章那条从航空业学来的办法——不给「发现」记功，给「报告」记账——是本书收集到的唯一一件真的被做出来过的解，它在合章里被翻译成第六章题域上的一条处方。

第七章

类内距

论一份规范只写了什么算同类，没写同类之间有多近

学科入口　语言学 × 工程学 × 艺术学

一、导论：三个不该同时为真的命题

三场争论，分属语言学、工程学与艺术学，各自都有厚实的支撑，而它们不能同时为真。第一场在历史语言学。十九世纪末的新语法学派留下一句纲领：语音规律无例外。音变作用于音位，凡处在同一语音环境里的词一律跟着变；正是这条假定支撑起了比较法与整套语言谱系的重建。二十世纪后半叶出现了相反的证据：对汉语方言与达罗毗荼语的研究显示，音变可以逐词扩散，同一条变化在不同的词里进度不同，而无论把语音环境切得多细都无法解释那种分裂；有研究者据此断言，这些变化以词而不是以音位为基本单位。一位重要的调解者提出：两种机制在不同的抽象层次上各自为真，音位层的规则是规律的，低层输出规则则可能逐词扩散——但这条调解至今仍在被反驳。争点是：变化的基本单位是类，还是个体？第二场在制造工程。现代制造的根基是互换性：任何一件合格件都可以替换任何另一件合格件，因为它们都落在图纸规定的公差带内。这是标准化最实在的成果——在它之前，零件由工匠逐件手工配作，每一台装配都是独一无二的、彼此不兼容的。但工程界同时长期使用一种与互换性观念不一致的做法：选择装配。它把公差放宽、把制造出来的零件按实测尺寸分组，只在对应组之间装配，从而用相对低精度的零件获得高精度的配合。它省钱，因为收紧公差的成本随精度上升得极快；代价是它不保证完全互换——只有同组的件才能换。争点是：可替换性靠把每一件都做进同一条带子，还是靠把做出来的件配成对？第三场在艺术学。一种有影响的分类把艺术分成两类：一类作品的同一性由记谱法给定，任何符合记谱的实例都是那件作品，因而没有伪作问题；另一类作品的同一性由它的制作历史给定，一件复制品即使与原作在感知上不可分辨，也不是那件作品。前者的典型是音乐与文学，后者的典型是绘画。这个区分至今是艺术哲学的一根支柱，也至今被争论：完全合乎乐谱的两次演奏可以听起来判若两曲，而人们仍然说它们是同一首曲子。争点是：作品的同一性由规范给定，还是由个体历史给定？三场争论互不相干，

却在同一个问题上互相顶撞：「同一类」这件事，是被什么保证的？第一场里最强的一方说：不是被规则保证的——规则再无例外，实际发生的变化也是逐词参差的。第二场里最省钱的一方说：不是被公差带保证的——不必把每一件都做进带子里，配成对也行。第三场里最锋利的一方说：是被记谱保证的——凡合谱者皆是同一作品，历史无关。三条不能同真。若记谱能保证同一，则语言的音位系统（语言学里的记谱法）也该保证同一，而词汇扩散显示它不能。若配对能替代公差带，则「同一类」根本不是一个整体属性，只是一堆两两关系，那么记谱法所许诺的那种整体性就落空了。若逐词参差是常态，则工程的互换性就该是不可能的，而它显然可能——只是很贵。本章认为，僵局来自一个三方共有、却从未被单独说出的前提：「同一类」是一件事。本章撤销这个前提。「同一类，所以可以互相替代」这句话，实际上是两句：

> 第一句由规范给出：什么算符合。第二句规范从来不给：符合的那些彼此有多近。

本章把第二句所说的量命名为类内距。于是本章的判断是：

> 一件东西能不能被同类的另一件替代，不取决于规范写得多精确，而取决于类内距被压到了多小；而类内距不写在任何规范里，它要单独花钱买。

三个案例在这一点上是同一个形状：图纸规定了什么算合格（边界），公差带的宽度决定了合格件彼此有多近（类内距），而收窄公差的成本随精度非线性上升——类内距是被买来的。音位系统规定了什么算同一个音（边界），而实际发音在词与词之间有多近，没有任何机制去保证——语言里没有人买过这笔账，于是那个距离就一直在，表现为逐词参差。记谱法规定了什么算符合这首曲子（边界），而两次合谱的演奏彼此有多近，记谱法一个字都没说——没买，而且不打算买，那个距离被留下来并被命名为风格。由此得到本章的承重推论：

> 压缩类内距的那笔功夫不会因为没人付钱而消失，它只会换一个地方发生。它在工程里被记作成本，在艺术里被称为风格，在语言里被称为变异——三者是同一件事在三种记账口径下的三个名字。

本章余下的安排：第二节梳理四条既有路径并指认它们停在哪里；第三节定位冲突；第四节界定类内距与它的三个构成条件；第五节升成二维辨别表；第六节在十个领域兑现，其中两个迫使本章收窄；第七节给出判据与可证伪条件；第八节交代与最近几种说法的分界；第九节说明三家为何各自到不了；第十节给出推论、边界与反噬预言；第十一节处理本判断对自身的适用。

二、文献综述

2.1 同一由规范给定

第一条路径把同一性交给规范：一套足够穷尽的记号系统规定了什么算符合，凡符合者即同一。艺术哲学里那个著名的两分是它最清晰的形态——有记谱法的艺术门类没有伪作问题，因为作品就是记谱所规定的那个类。历史语言学的新语法学派纲领在结构上与之同型：音变律无例外，凡处在同一环境者一律同变。这条路径的长处是它可操作：边界写下来，判定就有依据。它的盲区是本章的起点——它只规定了边界，没有规定符合者彼此有多近，而后者才是替代能不能发生的条件。一个完全合谱的演奏与另一个完全合谱的演奏可以相距极远；两件都在公差带内的零件，如果公差带本身很宽，装到一起仍然不能用。

2.2 同一由个体历史给定

第二条路径把同一性交给因果历史：一件东西之所以是它，因为它是那样被做出来的；复制品不是原作，哪怕不可分辨。词汇扩散在结构上与之同型——变化的载体是词这个个体，不是音位这个类；每个词有它自己的进度。这条路径的长处是它诚实地承认了参差。它的盲区在于它把参差当作终点：它指出个体各不相同，但没有追问「彼此有多不同」是不是一个可以被独立操纵的量。而工程给出的答案恰恰是可以——公差带就是那个操纵手段。

2.3 层次调解

第三条路径是那位调解者的方案：两种机制在不同的抽象层次各自为真。这是目前最被广泛接受的立场，也确实解释了一批经验：同一处方言里，某些变化呈现规律性，某些呈现逐词分裂。它停在哪里？停在它把差别归给了对象（哪一层的规则），而没有归给有没有人在压缩距离。在工程里，同一批零件既可以走互换性也可以走选择装配，这不是零件的层次不同，是投资方式不同。本章主张语言的情形与此同构：规律性不是某一层的属性，而是压缩机制存不存在的结果——而语言里那个机制从来不是工程性的，只有使用中的相互校准，因此它慢、不彻底、且对高频词与低频词作用不同。

2.4 互换性与它的替代品

第四条路径是工程自己的实践知识，也是本章最重要的材料来源。它包含三件本章全部依赖的事实：其一，互换性要靠公差带，而收紧公差的成本上升得极快；工程教科书把「选择功能上足够用的最松公差」当作设计者的基本功。其二，存在一种与互换性观念不一致、却长期在用的替代做法：放宽公差、逐件测量、按尺寸分组，只在对应组之间装配，从而用低精度零件得到高精度配合。它在滚动轴承、发动机活塞与缸体配合等场合有几十年的工业使用史。其三，这条替代路省的正是精度钱，付的是检验与分组钱，且它换来的高精度配合不含互换性

——只有同组的件才能换。这三件事合起来说明了本章的核心区分：边界（图纸）与类内距（公差带宽度）是两件可以分别操纵的东西，而后者是被购买的。 工程学是三个领域里唯一把这笔账明确记下来的一门。

三、冲突定位

三条命题的冲突可以写死为三条。冲突一（规则无例外 vs 逐词参差）。 若音变作用于音位这个类，则同环境的词一律同变；若变化以词为单位，则类在变化期间不成立。两条不能同真，而双方都有硬证据：一方有大量呈现规律性的进行中变化，另一方有无论怎样细分环境都无法解释的逐词分裂。冲突二（公差带 vs 配对）。 若可替换性靠把每一件都做进同一条带子，则配对装配是一种妥协；若配对能得到同样甚至更好的配合，则「同一类」不是整体属性而是两两关系。两条不能同真：前者预测放宽公差必然降低装配质量，后者预测放宽公差加分组可以提高装配质量而降低成本——而后者是工业实践反复验证过的。冲突三（合谱即同一 vs 不可分辨也不同一）。 若记谱给定同一，则历史无关；若历史给定同一，则记谱不足。两条不能同真。三条冲突有一个共同点，此前没有被指出：它们全都发生在同一个位置上——「符合」与「足够接近」之间那道缝。第一场：音位系统规定了什么算同一个音（符合），而实际发音有多接近没人管（距离）。第二场：图纸规定了什么算合格（符合），公差带宽度决定合格件有多接近（距离）——这一场是唯一把两者分开写下来的。第三场：记谱规定了什么算这首曲子（符合），两次演奏有多接近记谱一个字没说（距离）。三方都在争「同一类由什么保证」，而三方共享的前提是：保证只有一道。 一旦承认有两道——边界与距离——三条冲突就不再需要判谁对，而可以被重述为同一张表上的不同格。

四、类内距

4.1 定义与三个构成条件

> 类内距：一个类中，所有符合边界的成员彼此之间的最大差距，就使用者关心的那些维度而言。

它与边界是两件事：边界回答「算不算」，类内距回答「差多少」。一份规范只能给出前者；后者由实际做出来的那批东西决定。类内距要成为一个可操纵的量，需要三个条件，缺一它就只能被动地是它所是：其一，有一个可测的维度。 必须知道在哪一个量上比较接近程度。工程有尺寸、形位、粗糙度；语言有共振峰、时长；音乐演奏有速度、力度、音准。没有这个维度，类内距只能被感觉到，不能被压缩。其二，有检验。 必须有人真的去量并据以处置——超出的返工、报废，或者按实测分组。没有检验的公差带只是一句话。 这一条是三个领域最大

的差别所在：工程有检验部门，语言与艺术都没有。其三，有人承担压缩的代价。 收窄类内距要么靠提高制造精度（贵），要么靠检验分组（便宜但牺牲互换），要么靠让使用者每次现场适配（对制造侧免费）。三条路都要花力气，区别只在花在谁身上、记不记账。

4.2 为什么把规范写得更细没有用

这是本章最直接、也最反直觉的一条推论。面对「同一标准的两件东西装不到一起」这类问题，最自然的反应是把标准写得更细：补充条款、加严定义、明确歧义。本章预测这条路无效，理由是纯结构性的：歧义与距离是两回事。 一份毫无歧义的规范可以配一条很宽的公差带——每一件都毫无争议地合格，而彼此相距甚远。把边界画得再清楚，也不会让已经在里面的东西彼此靠近。工程学把这一点写在了成本曲线上：想让合格件彼此更近，只有一个办法——收窄公差带，而它的成本随精度非线性上升。补充条款不改变公差带，因此不花钱，也因此不产生效果。这条推论有一个可以立刻检验的形式，见第七节：标准文本的详细程度与实际适配工作量之间，不应当出现负相关。

4.3 那笔功夫的三个名字

若没有人付钱压缩类内距，距离并不消失，它只是必须在每一次使用时被就地跨过。跨过它需要功夫，而这笔功夫在三个领域里被记进了三本不同的账：在工程里它叫成本。 现场配作、返工、修锉、加垫片，都进工时。正因为它入账，工程才有动力去比较「事前收窄公差」与「事后现场适配」哪个更划算——这个比较本身就是选择装配这套方法的来源。在艺术里它叫风格。 两次合谱的演奏之间的距离没有被消除，它被保留下来并被重新命名为演奏者的个性。同一个量，在工程里是待消除的偏差，在艺术里是被珍视的价值。 这不是修辞上的巧合：正因为无人试图压缩它，它才有条件积累成一个可辨认的个人特征。在语言里它叫变异。 词与词之间的实现差距没有被任何机构压缩，于是它以逐词参差的形式一直存在，并在音变进行期间表现为词汇扩散。规律性之所以最终常常出现，靠的是使用中的相互校准——一种极慢、无人负责、且对高频词与低频词作用不同的压缩机制。语言里唯一的「检验员」是听者，而听者不返工，只是听懂了。三个名字对应三种记账，而不对应三种事情。这是本章认为三家合起来才能得出的那一步：任何一家单独看，都只看得见自己账本上的那一栏。

五、二维辨别表

第一条轴：谁承担压缩类内距的功夫（制造侧／使用侧）。第二条轴：这笔功夫是否入账（记成本／不记）。

	入账	不入账

制造侧承担	甲格　真互换。收窄公差＋检验，成本明确写在工艺卡上：标准紧固件、可换镜头卡口、集装箱角件	丙格　手艺。师傅的准头长在身体里，做出来的东西彼此极近，而没有任何一栏工时叫「保持准头」：传统匠作、成熟乐团的声部齐整
使用侧承担	乙格　显性配作。选择装配、现场修配、系统集成合同里单列的适配工时——距离仍在，但被算清了	丁格　名义同类（本章的靶）。规范写得极细、边界毫无歧义，而每一次使用都要现场跨过距离，且这笔功夫从不入任何一本账

这张表的承重之处在丁格，而丁格之所以值得单列，有三个理由。其一，它在文件上与甲格无法区分。 两格的规范都合规、都无歧义、都能通过审核。差别只在类内距，而类内距不在任何文件里。其二，它是默认状态。 写规范是一次性投入且成果可见；压缩类内距是持续投入且成果不可见（它表现为「没出事」）。任何一个把规范当作交付物的体系，都会自然地滑进丁格。其三，它把成本推给了最没有能力比较的人。 制造侧知道收窄公差要多少钱，因而能做取舍；使用侧不知道自己每次多花的那一点钟点合起来是多少，因而不会去要求收窄。乙格与丁格的差别不是谁干活，是干活的人知不知道自己在干活。由此得到本章的实践含义：拿到一份「符合标准所以可以互换」的承诺时，要问的不是标准写得严不严，而是——上一次有人真的量过合格件彼此差多少吗？谁量的？量完之后做了什么？

六、逐领域兑现

（一）历史音变。 读数：同一条变化在不同词中的进度差。预测：进度差与音变律描述的精细程度无关，与是否存在使词与词之间相互校准的场合（诵读、教学、广播、拼写规范）相关。（二）机械装配。 读数：装配返工率与图纸条款数。预测：不相关；返工率只随公差带宽度与检验强度变化。（三）音乐演奏。 读数：同一乐谱不同演奏在速度与力度上的离散度。预测：离散度不随记谱精细程度下降；只有排练（即使用侧的压缩投入）能降低它。这解释了乐团与独奏的差别：乐团买了类内距，独奏不买。（四）软件接口。 读数：声称兼容的两套实现之间所需的胶合代码量与规范文本长度。预测：不相关或正相关。符合性测试套件（即检验）才是有效手段——而它正是「买类内距」在软件里的形态。（五）学历与证书。 读数：同一证书持有者在同一岗位上的上手时间离散度。预测：与标准文本的详细程度无关；与是否存在统一实操考核（检验）相关。（六）医学诊断标准。 读数：同一诊断在不同医生间的处置差异。手册加细不降低它，评分者间信度训练（检验）才降低它。（七）翻译中的「等值」。 距离由

读者每次跨过，从不入账；因此译本之间的差距被称为译者风格。（八）法律条文的适用。 条文写细相当于加严边界；判决的一致性取决于是否存在指导性案例与统一裁判尺度（检验）。（九）选择装配——第一条让步。这一领域迫使本章收窄，而且是从内部。选择装配证明了类内距可以在不收窄边界的情况下被压缩：放宽公差、逐件测量、按尺寸分组，组内配合极紧。这直接支持本章的核心区分——边界与类内距确实是两件可分别操纵的事。但它同时削掉了本章一个便利的说法。分组之后，互换性只在组内成立：A 组的件换不上 B 组的位置。也就是说，压缩类内距的这条路是靠把类切碎换来的。于是：

> 类内距与类的大小此消彼长。 把一个类的内部距离压到很小，往往等于把它拆成若干更小的类；「同一类且彼此很近」这两件事不能无代价地同时最大化。

这条收窄不是补丁：它把本章从「压缩类内距总是好的」拉回到一个取舍问题，并解释了为什么很多领域宁可停在丁格——因为把类切碎会毁掉标准化本来要买的那个东西：不必问出身的可替换性。（十）柔性界面——第二条让步。第二条让步同样来自工程。当两个刚性件之间的配合难以保证时，常见的解法既不是收窄公差也不是分组，而是在中间放一个能变形的件——垫片、密封件、弹性联轴器——由它吸收两侧的偏差。这削掉了本章第二条轴的二分。「谁承担压缩」不是「制造侧或使用侧」两选一，还存在第三种：由界面吸收。而这第三种有它自己的特征——它把距离转成了另一个件的寿命：垫片会老化，弹性件会疲劳，缓冲层会磨完。

> 由界面吸收的类内距不会消失，它变成了界面件的更换周期。 因此判断一个系统在哪一格时，必须先问有没有这样一个吸收件；有的话，它的更换记录就是那笔功夫的账本，而这一栏往往被记在维护费里，与标准化完全脱钩。

两条让步合起来，把本章的适用条件收窄为：存在可测维度、且未设置吸收件的类。 三个引发本章的案例全部满足——音变没有吸收件，乐谱没有吸收件，而工程的例子里我们只讨论刚性配合。

七、判据、可观测代理与可证伪条件

7.1 核心可裁决判据

> 可替换性由类内距决定，与边界的精确程度无关：在不改变检验强度与制造精度的前提下，把规范写得更细不会降低适配工作量。

这条判据可裁决，因为它预测的是一个应当不相关的关系，而通行做法预设的是负相关（写得越细越好用）。任何一批案例上测出稳定的负相关——规范加细本身带来适配量下降——都直接驳倒它。

7.2 可观测代理

· 边界的精确程度：规范文本的条款数、定义数、修订次数。全部公开可得。

· 类内距：合格样本在使用者关心的维度上的离散度（标准差或极差）。

· 适配工作量：装配返工率、胶合代码行数、上手时间、排练次数、现场调整记录。

· 检验强度：是否有第三方一致性测试、抽检比例、评分者间信度训练时数。

· 里前三个大多已有现成记录，而它们从来没有被放在同一张表上看过——这本身是本章的一个推论：只有承认边界与类内距是两件事，才会想到去做这张表。

7.3 现有资料即可执行的证伪条件

取若干个有长期修订史的接口标准（软件协议、机械连接、数据交换格式），对每一次修订记录两件事：该版规范的条款数变化，以及该版发布后一年内相关实现之间所需胶合代码量的变化（开源生态中这两项都可统计）。- 通行看法预测：条款数上升、胶合量下降（负相关）。- 本章预测：两者不相关；胶合量只在引入一致性测试套件的那些版本之后下降。这个检验不需要新数据，只需要把已有的两类记录对齐。

7.4 第二条独立证伪条件

若能找到一个案例：规范加细之后适配工作量显著下降，而同期没有引入任何检验、没有提高制造精度、也没有新增吸收件——则本章的核心机制错。两条是独立的：第一条针对相关结构，第二条针对机制的必要性。本章刻意不采用「若找到一个类内距很大而仍可互换的例子则本章失效」这类说法——那句话可以被「说明那个维度不是使用者关心的维度」无限吸收。

7.5 两个交出去的洞

第一，「使用者关心的维度」由谁定，本章没答。同一批零件在不同用途下的类内距不同，而本章没有给出确定维度的独立办法；这使得类内距在跨用途比较时不可通约。第二，丙格（手艺）难以证实。说「师傅们做出来的东西彼此极近而没有入账」，需要一个不依赖成本记录的类内距读数；在没有测量传统的行当里，这个读数往往只能靠事后重建，而重建者本身带着自己的尺子。

八、与最近的几种说法的分界

与记谱法给定同一（艺术哲学的两分）。 该框架把有记谱法的艺术门类判为没有伪作问题。分界：它给的是边界，本章追问的是内聚。分岔情形——两次完全合谱、彼此相距极远的演奏：按那个框架，它们无疑是同一作品，问题到此为止；按本章，这正是类内距未被购买的现场，而「风格」是这笔未付账款的名字。二者不矛盾，但本章预测一件那个框架不预测的事：排练强度上升，同一乐团不同场次的离散度下降，而「作品同一性」这件事在那个框架里不该随排练变化。与音变律无例外（新语法学派）。 分界：它把规律性当作变化的性质，本章把它当作压缩机制的结果。分岔：按它，规律性与是否存在标准化诵读、拼写规范、广播语音等场合无关；按本章，相关。这是可以用不同社会条件下同类音变的进度差来分开的。与词汇扩散（王士元一路；Krishnamurti 的达罗毗荼语证据）。 该路径指出变化以词为单位。分界：本章同意它的观察而不同意它的落点——「单位是词」不是一个关于语言的本体事实，而是「没有人压缩过类内距」的表现形式。分岔预测：在存在强力校准场合的语言社群里，同一类变化应当更接近规律性；这是可以在有无标准语教育的社群之间比较的。与层次调解（Labov 1981）。 那条方案把两种机制分配给两个抽象层次。分界：本章把差别归给有没有压缩投入而不是归给层次。分岔：按层次说，同一层的变化应当呈现同一种模式，与社会条件无关；按本章，同一层的变化在不同社群可以呈现不同模式。与互换性的经济史。 关于标准化如何降低成本、如何使不同来源的零件得以装配，已有大量研究。分界：那些研究把互换性当作一个成就来叙述，本章把它当作一笔持续开支——互换性不是达成之后就保有的状态，公差带需要被持续地维持，检验一停它就退回丁格。与选择装配的成本模型。 工程界已有比较传统装配与选择装配成本的模型。本章完全依赖它的事实，并只在一处推进：那些模型比较的是两种购买方式的价钱，从未把「不买」作为第三个选项列进去——因为在工程里不买等于装不上。而在语言与艺术里，不买是常态，这正是三家合起来才看得见的东西。与通约化（Espeland & Stevens 1998）。 该工作论证把不同的东西按同一把尺子换算会剥掉情境。分界：通约化处理的是跨类比较，本章处理的是同类内部的距离。二者可同时成立且互不覆盖。与分类的残余（Bowker & Star 1999）。 那里关心边界之外被处理不了的东西，本章关心边界之内成员彼此的距离。一个是外面，一个是里面。与学科通融专栏《双链型》（之十七）。那一篇论证重复可以由制作链或判读链两条独立的链产生，因而重复不能区分型是被做出来的还是被读出来的。本章与它自变量正交：那一篇问重复从哪来，本章问适配功夫落在哪。可分开的情形：一个由判读链单方面稳定的型（投影型），其类内距完全由判读办法决定，而本章的表在那里退化——因为没有制造侧可以承担压缩。这一格是两篇的接缝，也是两篇都不能单独处理的地方。与学科通融专栏《路存》（之十二）。 那一篇论证有一类系统的储备由路线而

非持有者定义。分界：那里是储备的位置，这里是功夫的落点。二者在同一个系统里可以各自为真。

九、三家各自为何到不了这一条

艺术学到不了，因为距离在它那里是价值不是缺陷。一门以个别性为核心价值的学问，不会把「合格者彼此相距多远」当作一个要被压缩的量。它因此拥有最丰富的关于距离的描述（风格、手笔、气质），却没有一个位置去问「这段距离本来可以被买掉吗」——那个问题在它的价值体系里近乎冒犯。语言学到不了，因为它没有检验员，也没有委托人。语言没有制造侧、没有质检部门、没有人为「这两个词的实现应当更接近」付钱。因此在语言学的视野里，参差只能是被观察的现象，不可能是未被购买的服务。它能极精细地测量距离，却不会把距离读成一笔没人付的账。工程学到不了，因为它的三样都齐。工程有可测维度、有检验、有成本科目——正因为三样都齐，「买类内距」这件事在工程内部是背景常识，不构成一个需要被命名的发现。一个把某件事做得最好的领域，通常是最不可能把它说出来的领域：它太顺手了，顺手到不像一个变量。三家的边界都不是能力问题，是账本结构问题：一个把它记成价值，一个根本没有账，一个记得太熟以至于不再看它。三本账并排放着，那个共同的量才显出来。

十、推论、适用边界与反噬预言

10.1 推论

推论一。任何以「符合同一标准」为据的可替换性承诺，若不附一个类内距读数，其强度不明。而现行的合规文件几乎都不附这个读数。推论二。面对适配困难，加细规范是最便宜、最可见、也最无效的一步；它之所以被反复选择，正是因为它便宜且可见。推论三。一个体系可以在完全合规的状态下持续变差：只要检验放松、公差带实际放宽，而规范文本不动，所有文件都仍然正确。丁格没有告警信号，因为它的唯一读数不在任何报表上。

10.2 适用边界

其一，没有可测维度的类（第七节第一个洞），本章只能定性使用。其二，设置了吸收件的系统（第六节第十条），本章的第二条轴退化，须先查吸收件的更换记录。其三，一次性的、不需要替换的东西。类内距以「能不能拿另一件顶上」为讨论对象。

10.3 反噬预言

按本章设计干预，最自然的做法是要求每份标准附报一个类内距读数。这会适得其反，方式可以精确说出：类内距读数一旦成为申报项，最省钱的达标办法不是压缩距离，而是收窄

「使用者关心的维度」的定义——只报那些本来就很齐的维度。由于本章自己承认「关心哪个维度」没有独立的确定办法（第七节第一个洞），这条路是敞开的。于是：申报出来的类内距会稳定地变小，而实际适配工作量不变。这条预言同时说明了那个洞为什么必须留在明处：它是本章最容易被拿去干坏事的地方，写出来比藏起来安全。因此本章不推荐把类内距做成申报项，只推荐一件成本很低、且不能被优化的事：每隔一段时间，随机取两件合格品，让一个不知道规范细节的人试着互换，并记下他花了多久。 这个动作之所以不能被优化，是因为它测的是使用侧的实际功夫，而不是任何一方声称的维度。

十一、本判断对它自己适用吗

本章给出了一份规范：四个格、一条判据、几个代理指标。那么它自己落在哪一格？按它自己的表：丁格。理由是本章写的全是边界——什么算类内距、什么算入账、什么算吸收件——而它一个字都没有规定「两个人用这四格去分同一批案例，结果应当有多接近」。我没有测过，也没有为它花过任何一笔钱。任何人拿这四格去判自己手上的系统，都要在现场自己跨过一段我留下来的距离；而那段距离不会出现在这篇文章的任何地方。更具体一点：第五节那三条丁格的理由（文件上与甲格不可分、是默认状态、把成本推给最没能力比较的人），逐条适用于本章。写这四格是一次性投入且成果可见；测两个人分类结果的接近程度是持续投入且成果不可见。我做了便宜的那一半。按第十节那条不能被优化的建议，本章自己该做的动作是：请两个没读过本章推导过程的人，各拿这四格去分同一批案例，记下分歧率。这件事我没有做，而做它的成本很低。 在有人做之前，本章的正确说法是「它给出了一条可检验的区分」，而不是「这四格是可用的」。

结　论

本章提出并论证：「同一类，所以可以互相替代」是两句话——边界由规范给出，类内距规范从来不给。可替换性由后者决定，与边界写得多精确无关；把规范写得更细不会降低适配工作量，只有单独投入压缩类内距才会。压缩类内距的那笔功夫不会因为无人付钱而消失，它只会换一个地方发生，并在三种记账口径下得到三个名字：工程里的成本、艺术里的风格、语言里的变异。二维辨别表的靶格是丁格——规范极细、边界毫无歧义，而全部适配功夫压在使用侧且从不入账。它在文件上与真互换无法区分，它是默认状态，而且它把成本推给了最没有能力比较的那一方。十个领域的兑现中，选择装配与柔性界面各迫使本章收窄一次：前者说明压缩类内距的一条主要路径是把类切碎，类内距与类的大小此消彼长；后者说明这笔功夫可以由第三方吸收，而吸收件的更换周期就是那笔功夫的账本。最后给一条写死日期的赌注：

到 2033 年 12 月 31 日，若在接口标准、职业资格与演出实践三个领域中，仍没有任何一项研究把「合格样本之间的离散度」作为独立于规范文本详细程度的变量单独报告，或报告之后未能观测到本章所预测的零相关，则本章的核心主张作废，不再申辩。

注　释

1. 本章所称「类内距」限于使用者关心的那些维度上的差距，不是全维度差距。这一限定是必要的，也是本章最大的软处，见第七节第一个洞。2. 「入账」指这笔功夫在某个具体的记录体系里占有一个可被查询的条目，不指它是否被感知到。手艺人当然知道自己在保持准头，但没有一栏工时叫这个名字。3. 第五节表中「制造侧」与「使用侧」是相对于某一次替换而言的位置，不是固定的组织身份；同一个部门在不同环节可以分处两侧。4. 第六节第九条所述选择装配的事实（放宽公差、逐件测量、按尺寸分组、组内互换、不保证完全互换、在滚动轴承与活塞—缸体配合上有数十年工业使用史）依据可公开获取的工程文献与教科书表述。5. 本章对三家来源的转述限于其公开发表的核心论证；三场争论各自的内部分歧远比本章所述细密。语言学与艺术学两家的关键文献本章未逐条核对原文页码，需精确引证者请核原文。

人机分工声明

本章由王德生与 Claude 共同署名。三个学科由王德生指定（语言学＋工程学＋艺术学），三家对立理论由检索确定；判断的形成、二维辨别表、十领域兑现（含两条收窄）、判据与证伪设计以及全部文字，由 Claude 生成；所依据的方法论框架与发表裁定由王德生给出。工程侧的事实依据可公开获取的工程文献与教科书表述；语言学与艺术学两家的关键文献未逐条核对原文页码，引用限于其公认的核心论点。本章未调取任何一手实验或生产数据，第七节所述两项检验给出了设计而未执行——这一限制已写进第十一节。字数　纯汉字约 1.2 万　版本　v1.0 · 2026-08-01 三个学科由王德生指定。三边在「同一类这件事被什么保证」上给出互相排斥的答案：语言学那边说规则再无例外、实际变化仍逐词参差；工程那边说不必把每一件都做进公差带、配成对也行；艺术学那边说凡合乎记谱者即同一作品、历史无关。三家的主要争论至今未决，出处可自行核对。

第八章

记零位

论一个共同体最后擅长什么，由它的账上被记为零的那类动作决定

学科入口　学习科学 × 专利发明人认定 × 作坊归属研究

一、导论：三个不该同时为真的说法

一件由多人（或人与机器）共同完成的作品交付之后，留下的能力归谁？三个领域给出三个互相排斥的答案。第一场在学习科学。 这一支的核心判断是：能力由执行沉淀。自行产生一段内容比被动读同一段内容更有利于保持；在缺乏支持的条件下先行尝试、哪怕失败，也可能为后续教学准备出更好的结构。推到共同工作上，结论很自然——谁真正动手做了那一步，那一步的能力就长在谁身上。争点是：练习之外，还有什么能决定能力的去向？第二场在专利的发明人认定。 这里的答案与上一场无关，甚至相反。要成为一件专利的发明人，必须对权利要求所限定的技术方案作出构思上的贡献；把别人的构思变成可运行的实物、做实验去验证它能不能成立、发现并修正实施中的问题，在这套规则下一律不构成发明人资格。构思被认定为发明行为的完成，其余动作被明确记为零。一个人可以在实验台前干了三年、发现并排除了整个方案里最致命的一个错误，而在发明人一栏里没有位置。争点是：核验与实施到底算不算贡献？第三场在作坊的归属研究。 十七世纪的大作坊里，学徒、助手与师傅共同完成一幅画，最终署一个名字。二十世纪后半叶开始的系统性归属重审，把大量长期挂在师傅名下的作品重新判给助手或作坊，也把一些被降级的作品重新判回。这项工作揭示的不只是「谁画的」这个事实问题，还有一件更要紧的事：署名不是能力的结果，署名是资源。 谁有权署名，谁就有权接单、有权定题、有权决定下一批学徒练什么；被署名制度排除在外的那些手，画得再多也不会积累成一个可以被继承的行当。争点是：是能力决定署名，还是署名决定能力？三条不能同真。若能力只由执行沉淀，那么专利法把实施与核验记为零，就只是法律上的技术安排，不会影响任何人的实际能力——而现实里，一个领域长期不给某类动作记账，那类动作的从业者会消失。若署名完全决定能力，那么学习科学几十年的练习效应就该被推翻——而它没有。若构思是唯一的贡献，那么整个校对、审稿、质检、复核的行当就不该存在——而它们不但存在，

还是很多领域出错率的决定项。本章认为，僵局来自一个三方共有、从未被单独说出的前提：贡献的账本与能力的去向是两件事，账本记得对不对，不影响能力实际长在哪里。 学习科学只看能力不看账本，专利法只管账本不问能力，作坊研究看见了账本会影响资源却把它当成分配问题而不是认知问题。没有人问：账本长期把某一类动作记为零，那一类动作所沉淀的能力会怎样。

二、三家各自说到了哪一步

学习科学这一支的两块基石都很老。一块是自行生成的材料比被动接受的材料记得更牢；另一块是在结构不良的复杂问题上先行尝试、即使得出较差的方案，也可能在后续教学之后带来更好的迁移。两块合起来的意思是：认知能力是在执行中长出来的，而不是在观看中长出来的。这一支近年被推到人机协作里，得出的推论是：机器承担的那部分执行，人就不再练习；因而要保住能力，就要保住某些执行由人来做。这一支说得对，但它只处理了一个人与一次任务。它没有处理的是：在一个有几百人、几十年的共同体里，谁被允许执行哪一类动作，不是由个人选择的，而是由这个共同体的记账方式决定的。专利的发明人认定这一支把界线画得极硬。判例法把发明拆成构思与付诸实施两段，并把发明人资格系于前者：一个完整、可操作的构思在头脑中形成，此后的实施只是把它做出来。据此，出资者、提供常规技术协助的人、按指示做实验的人、以及仅仅解释了现有技术的人，都不是发明人。这条线在实践中还有一层：漏列或错列发明人会危及专利的有效性，因而这个领域对「谁不算」的审查比对「谁算」更严。[1]值得注意的是，这条线不是疏忽。它是被一个具体的困难逼出来的：核验与实施的贡献可被无限稀释。 一个方案从构思到落地会经过几十人的手，每一个人都可以说自己「检查过、修正过、提出过意见」，而这些说法在事后极难分辨真伪与轻重。构思有一个可确定的时点与一份可比对的内容，核验没有。把整类动作记零，是这个领域对稀释问题给出的、粗暴但可执行的解法。这一点在第八节会反过来限制本章。作坊归属研究这一支提供的是几十年的实证材料。一个长期运作的委员会逐幅重审归属，把大批作品在师傅、助手与作坊之间重新分配，并且在此过程中不断修改自己的判定标准，甚至公开撤回过先前的判定。这项工作留下两个对本章有用的结论。其一，署名与手笔在历史上从来就不重合，而重合与否取决于当时的行会规则、市场惯例与合同，不取决于技艺。其二，也是更要紧的一条：归属可以被回溯重写，而每一次重写都会立刻改变现在的市场估值、展览安排与研究方向。 账本不只是记录，它是活的。[2]

三、冲突落在哪一点上

把三家的处方并排放，它们互不兼容。学习科学要保住人的执行机会；专利法要把执行与核验排除出贡献；作坊研究要重审谁在署名。三者的目标各自成立，放在一起就互相拆台。但把三家的材料而不是处方并排放，会出现另一样东西。学习科学说：能力长在执行上。专利法说：某一类执行被系统性地记为零。作坊研究说：账本决定谁被允许执行下一次。三条串起来，得到一条谁也没说的推论：被账本记为零的那类动作，其能力会先在个人身上正常沉淀，然后在共同体层面消失。消失的机制不是遗忘，是接续被切断——做这类动作的人得不到署名，得不到署名就得不到资源与学徒，一代之后没有人再进入这个位置。而个人层面的练习效应完全正常地运作着，因而在任何单人的观测里都看不出问题。更关键的一步在于，被记零的那类动作不是随机的。它有一个共同特征：它的产出是一个否定。发现错误、拒绝一个方案、指出某条路走不通——这些动作的成果是「某件事没有发生」，而没有发生的事不留痕迹，因而难以计量、易被稀释、也无法向第三方证明。构思的成果是一个可以被指着看的东西；核验的成果是一片空白。账本天然偏爱能被指着看的东西。于是问题的形状变了。它不再是「贡献该怎么分」这个规范问题，而是一个可以被观测的结构问题：一个共同体长期把哪一类动作记为零，就会在一代人之后失去哪一类能力。

四、记零位

在一个共同体的正式记录里被记为零贡献、而在能力沉淀上其实是主要承载点的那一类位置，本章称之为记零位。三个限定词都要紧。「正式记录」而不是「口头评价」。一个审稿人可能被同行私下敬重，但如果他的名字不进任何可被援引的记录——不进论文、不进履历、不进晋升材料、不进任何一份能在他求职时被读到的东西——那么这份敬重不构成账目。判断依据是可迁移性：这份记录能不能跟着他走到下一个机构去。「记为零」而不是「记得少」。记得少还是账，还能被讨论、被调整；记为零意味着这一类动作在账本的科目表里根本没有对应的行。多数领域是后者：一份履历上可以写发表了多少篇，没有一栏写拦下了多少篇。「主要承载点」而不是「也有贡献」。记零位之所以要紧，是因为它承载的能力与被记账的那类动作所承载的能力不是同一种，因而不能互相替代。提出方案练的是从空白到候选那一段；发现错误练的是从候选到失败模式那一段。前者做再多，也不会自动长出后者。这一点在很多领域可以直接观察到：一个能写出漂亮方案的人，未必能看出别人方案里的致命洞，反之亦然。把这条判断压成一句：一个共同体最后擅长什么，不由它奖励谁决定，由它把谁记为零决定。奖励只是在被记账的动作之间分配强度，记零则决定哪一类动作根本不在账上——不在账上的那类，会在一代人的时间尺度上退出这个共同体。这条判断为什么三家都不会同意，值得单说。学习科学不会同意，因为它的因果链止于个人练习，而本章说个人练习完全正常时能力照样会在共同体层面消失。专利法不会同意，因为它认为把核验记零是一个法律上的必要安排，与能

力无关；本章说它有一个法律没有打算产生的认知后果。作坊研究不会同意，因为它把归属当成分配公平问题，而本章说它首先是一个能力再生产问题——重写归属改变的不只是谁得到荣誉，是下一代人会去练什么。

五、两根轴，四种共同体

只用「有没有记零位」一根轴不够。几乎所有共同体都有记零位，而它们的结局并不相同。要分开，需要第二根轴。第一根轴：这个共同体是否存在一个承载核心能力的记零位。 第二根轴：占据这个位置的人，是否有别的通道获得练习机会、资源与晋升——比如这个动作在另一套账上被记账，或者占位者同时占着一个被记账的位置。

	有替代通道	无替代通道
存在记零位	分化的行家 核验型能力在一个平行的账上存活：药师有独立的执业资格与责任，编辑有自己的职级序列，审计有法定地位。做否定动作的人靠另一套账拿到资源。	消失的行当 这类能力在一代人内绝迹，而绝迹时没有人察觉——因为它从来没有被记过账，所以它的消失也不会在任何账上表现为下降。
不存在记零位	全账共同体 否定动作被正式记账且可迁移。稀缺且昂贵，因为记账要解决稀释问题。	单一竞赛 只有一类动作被记账，所有人朝同一个方向卷。短期产出极高，失败模式高度同质。

右上格是靶格，它有三个使它难被发现的性质。第一，它的衰减不出现在任何指标上。 产出量、发表量、交付速度都不会因为没有人再擅长发现错误而下降——恰恰相反，短期内它们会上升，因为拦截少了。系统正在丢失的东西不在报表里，因为它从来不在账上。第二，它的衰减是代际的，不是个人的。 现有的那批行家还在，能力也还在；变化发生在没有人接替上。因而任何以个人为单位的评估都测不到它，只有看年龄分布才能看到。一个行当的年龄中位数逐年上升而入行人数不降，是它最早的信号。第三，它一旦跨过临界就不可逆。 因为教这类能力的人本身来自这个位置：没有人在这个位置上，就没有人能教，就更没有人进来。左下格与右上格的差别正在这里——单一竞赛可以靠改激励掉头，消失的行当改了激励也找不到教的人。

六、一个不必问任何人的判据

判断一个共同体是否有记零位，不必访谈、不必问卷、不必评估任何人的能力。只需查一件事：在这个共同体的正式记录里，「发现了一个错误」被记在谁名下，算多少？三个追问把它变成可执行的检查：这份记录能不能跟着人走到下一个机构；它在晋升材料里有没有对应的栏目；它能不能被第三方独立核对。三个都否，就是记零位。跑几个共同体，答案互不相同，而且都不是从命题里推出来的。学术出版。 同行评议中拦下一篇有致命缺陷的稿子，记在谁名下？绝大多数情况下：不记名、不可援引、不进履历、不可被第三方核对。而发表一篇记在作者名下且完全可迁移。判定：记零位，且长期缺少替代通道——这解释了一件不需要道德解释的事：为什么审稿越来越难约到人，且约到的越来越多是资历最浅的人。代码审查。 拦下一个缺陷记在谁名下？多数团队的工具里，审查记录是留痕的、具名的、可被检索的，但它不进绩效、不进晋升材料、也不跟人走。判定：弱记零位。而这个位置有一条替代通道——审查者通常同时是提交者，靠另一半拿资源。这解释了为什么代码审查比同行评议更持久，也解释了它的失效模式：当团队里出现纯审查岗时，这条替代通道断了。临床用药审核。 药师拦下一张有相互作用风险的处方，记在谁名下？在多数体系里，药师有独立的执业资格、独立的责任认定与独立的职级序列。判定：有记零位，但替代通道很硬。左上格。这解释了为什么药师复核在结构上与行政复核相同、有效性却差得远。专利实务。 实验室里发现一个方案的致命缺陷，记在谁名下？按发明人认定规则：零。且没有替代通道——发明人栏是这个领域的通货。判定：右上格。而这一格的后果已经可见：这类工作被整体外包给合同研究机构，即成为一个买卖服务而不是一个可继承的行当。飞行安全。 报告一个未遂事件记在谁名下？这个领域几十年前就把这一格改了：建立不追责的自愿报告制度，让报告本身成为被记账、被统计、被公开引用的东西。判定：左下格，且是本章所知从右上格被主动搬出来的最成功的一例。它的做法很值得记：它没有去给「发现」记功，它给「报告」记了账——因为报告有一个可确定的时点与一份可比对的内容，而发现没有。这条办法在第八节会再出现。

七、逐领域兑现

一、教科书与教材的勘误。 指出一处错误在任何账上都不出现。结果是勘误依赖零散的个人热情，且几乎全部来自已经功成名就、不需要账的人。二、数据集与基准的清洗。 发现一个基准里的标注错误或泄漏，记零；而在这个基准上刷高一个点，记满分。可观测后果：基准的缺陷长期存在且被反复引用，直到某次大规模复现失败才集中爆出来。三、内部审计。 有法定地位与独立报告线，因而在左上格。它的历史正好是本章的正面例证——独立性是被一连串事故逼出来的，而每次事故之后加强的都是「审计能不能被记在自己的账上」这一条。四、新闻的事实核查。 核查员拦下一条假消息记零，署名归记者。近年一些机构开始把核查员署进稿件，本章预测这类机构的更正率会先上升后下降——上升是因为拦截被看见了，下降是因为拦

截真的变多了。这是一条可用公开更正记录直接检验的预测。五、软件的安全披露。 这个领域已经部分解决了记零问题：漏洞被编号、被归属、可被援引，形成了一套独立的通货。结果是它长出了一个完整的、可继承的行当。这是全账共同体最接近的一个现实样本，而它的成本也很清楚：编号体系本身要养、要防灌水、要仲裁重复与夸大。六、翻译与校订。 校订者纠正一处误译，记零；译者署名。长期后果是校订环节被压缩到几乎不存在，而误译的发现被推给读者。七、建筑的图审。 有法定地位与独立责任，左上格。与药师同构。八、学位论文的答辩委员。 提出一个致命质疑记零。可观测后果：质疑强度与委员和导师的关系强度负相关，而与专业判断力关系较弱。九、机器学习的红队。 找到一个失败模式记在谁名下？部分机构已经开始把红队报告作为可援引的产出。本章预测这类机构与未如此做的机构，在同等模型能力下，发布后暴露的失败模式数量会有系统性差异。十、开源项目的问题分诊。 复现一个缺陷、判定它不是缺陷、关掉一个重复报告——这些动作在项目的贡献统计里通常记零，而提交一行代码记满分。可观测后果：几乎所有中大型项目都有分诊积压，且积压量与项目活跃度正相关而非负相关。

八、两处被材料逼出来的收窄

第一处，由专利法逼出。 前面已经指出，把核验记零不是疏忽，而是对一个真实困难的回应：核验的贡献可被无限稀释。 任何人都可以事后声称自己检查过、提出过、修正过，而这些声称在事后极难分辨真伪与轻重；一旦给它开一个账目，第一件发生的事不是核验变多，是声称变多。这条约束推翻了本章最自然的处方。「给核验记账」这个提议，在稀释问题面前会自我毁灭：账目一开，它记录的就不再是核验，而是关于核验的声称。因而命题必须收窄：记零位的解法不是给否定动作记账，而是让否定动作产出一件不可稀释的物证。 飞行安全那一例做对的正是这件事——它没有给「发现」记功，它给「报告」记了账，因为一份带时间戳的、内容可比对的、在事件之前提交的报告，是无法事后伪造的。同理，漏洞编号之所以能成为通货，不是因为有人决定给安全研究记功，是因为一个可复现的漏洞是一件硬物证。由此得出一条可以直接用的判别：一个否定动作能不能被记账，取决于它能否留下一件在事后无法重制的东西。 能，记账可行；不能，任何记账尝试都会退化成声称的竞赛。这条收窄让本章丧失了一个漂亮的普遍处方，换来一条可执行的判据。第二处，由作坊归属研究逼出。 归属可以被回溯重写，而重写会立刻改变现在的资源流向。这件事对本章既是加强也是限制。加强的一面：它给出了本章的正向读数。若命题成立，那么一次真实的归属重写——把某类此前记零的动作变成可援引的记录——应当在一代人的尺度上改变这个共同体的能力分布，而不只是改变荣誉分配。这可以在开始署名核查员的新闻机构、开始把审稿意见公开具名的期刊、开始把红队报告当作产出的机构里追踪。限制的一面更重：归属既然可以被重写，那么账本就不是一个稳定的

自变量。本章一直把记账方式当成因、把能力分布当成果；而作坊那几十年的重审说明，因果是双向的——市场对某类作品的重新估值，会反过来推动归属的重审。于是本章不能主张一个单向的机制，只能主张一个更弱的东西：在账本稳定的时间窗内，记零位预测能力的代际流失；账本一旦松动，预测方向不再确定。这把本章从一条因果律削成了一条条件性趋势。

九、与最近的几种说法的分界

一、与生成效应。那一支预测自行产生的内容比被动接受的记得更牢，机制在个人的编码上。判决点：取两组人，练习量、练习类型、反馈质量完全相同，只有一组的否定动作进入可迁移的正式记录。生成效应预测两组的个人能力增长相同；本章预测两组的个人能力增长确实相同，而十年后这两个共同体里从事该动作的人数分布显著不同。两者的读数在不同层级上，可以同时检验。二、与错误管理训练。那一支主张让人在安全环境里犯错并从反馈中恢复，能改善迁移。它与本章的差别在于「谁识别错误」：错误管理关心的是犯错者本人的学习，本章关心的是识别他人错误的那个位置。判决点：一个大量制造并修复自己错误、却从不需要识别他人错误的共同体，错误管理理论预测适应能力良好，本章预测其诊断型能力仍会流失。三、与贡献溯源制度。近年出现的一批做法要求逐条记录「谁做了什么、谁检查了什么」，把人与机器的贡献分开留痕。这是本章最近的一位制度上的邻居。分离线：溯源要回答的是一件已完成的作品由谁产生，本章要回答的是这套记录会在一代人之后把这个共同体塑成什么形状。判决点：一个溯源记录完备、但这些记录不进晋升材料也不跟人走的机构，溯源框架判为合规良好，本章预测记零位照样存在、代际流失照常发生。留痕不等于记账；能被读到、能被援引、能跟着人走，才是账。四、与荣誉分配的公平讨论。关于署名顺序、贡献声明、作者资格的大量讨论，处理的是分配正义。本章不处理公平，只处理后果。判决点很清楚：一套在分配上被公认为公平的规则（比如按工作量比例署名），若仍把否定动作记零，本章预测能力流失照旧；公平框架预测问题已解决。五、与格雷欣式的劣币驱逐。那条老规律说的是估值失真导致好的东西退出流通。表面上很像，机制不同：劣币驱逐要求两种东西被强制按同一价格交换，本章说的是一类动作根本不在价目表上。判决点：给否定动作定一个（哪怕很低的）价，劣币模型预测它仍被驱逐；本章预测只要它进了价目表并且不可稀释，流失就会止住。六、与技能的代际断层。制造业里有大量关于手艺失传的研究，通常归因于自动化替代或产业转移。判决点：找一个既没有自动化替代、也没有产业转移，而该技能仍在流失的行当。本章预测在这类案例里必定能找到一个记零位；断层理论无法预言这一点。七、与本站此前讨论评价通道垄断的那篇文章。那篇讲的是单一量化通道如何反过来规定生产什么知识，机制是被测者迎合指标。本章讲的是根本不被测的那一类动作。判决点：一个指标极其多元、通道完全不垄断、却

仍然不给否定动作任何科目的共同体，那篇判为健康，本章预测诊断能力照样流失。多元的账本与完整的账本不是一回事。

十、三家各自为什么到不了这一条

学习科学到不了，是因为它的观测单位是人和一次任务。记零位的效应发生在代际与共同体层面，任何以个人为单位的设计都测不到它——而且在个人层面，一切都是正常的。专利法到不了，是因为它的目的不是培养能力。它要解决的是权利归属的确定性，而稀释问题使它必须画一条硬线。它知道这条线粗暴，它不知道也不需要知道这条线的认知后果。作坊研究到不了，是因为它的问题是「这幅画是谁画的」。它积累了大量说明账本会改变资源流向的材料，但它把这看作艺术市场的运作方式，不会去问一个共同体的能力构成会因此变成什么样。

十一、推论、适用边界与反噬

第一条推论：年龄分布是最便宜的诊断工具。记零位的信号不在产出指标里，在人口结构里。一个位置的从业者年龄中位数逐年上升、而这个领域整体的入行人数并未下降，就是它正在失血。这条不需要任何新数据。第二条推论：不要给否定动作记功，要给它的物证记账。这是第八节收窄的直接结果，也是本章唯一一条可执行的处方。物证的三个条件：有确定时点、内容可比对、在结果揭晓之前提交。适用边界一：本章只处理否定型动作。有些被记零的动作不是否定型的（比如组织协调），它们的稀释问题与物证条件都不同，本章的处方不适用。适用边界二：短周期共同体不适用。代际流失需要一个比人的职业生涯短得多的观测窗才有意义。存续不足十年的组织里，这条机制来不及发生。适用边界三：本章不主张所有能力都值得保住。有些行当的消失是好事，因为那类判断已经被更可靠的方式取代。本章只提供诊断，不提供该不该救的判断。反噬预言。若「否定动作的物证记账」被普遍采用，最可能发生的不是诊断能力回升，而是物证的通货膨胀：人们会开始生产大量形式合格、实质琐碎的物证——无关紧要的漏洞报告、无关紧要的未遂事件、无关紧要的审稿意见——因为物证一旦成为通货，产出物证本身就成了被记账的动作，而它同样可以被稀释，只是稀释得慢一点。本章的处方只是把稀释推迟了一个层级，没有消灭它。漏洞编号体系已经在这条路上走了一段，它现在最大的运营成本正是筛掉灌水。这个循环没有出口；能做的只是把每一轮稀释的周期拉长。交出去的一个洞。本章没有办法处理这种情形：一个动作既是提出也是否定（比如提出一个反例）。它在账本上通常被记为提出而拿满分，但它沉淀的是诊断能力。这类动作的存在使两类能力的边界不清，而本章的整个论证依赖这条边界。

十二、与同期几篇的分界

本章与同期发表的另外六篇处理同一片地面，须逐一说清。与「不按的代价先落在谁身上」那一篇。 那篇问的是干预权持有人是否付不干预的账，机制在代价的时序归属；本章问的是干预这个动作在账本上有没有科目，机制在贡献的记账方式。判决点：一个把不干预的代价完全内部化、却仍不给干预记任何账的位置，那篇预测干预率会上升，本章预测个人干预率会上升而这个位置仍然招不到接班人。两条读数一个在个人层、一个在代际层，可以同时观测。与「谁能改下一轮的标准」那一篇。 那篇问改写权，本章问记账权。判决点：把评价标准的修改权完全交给某个位置、而这个位置的产出仍然不进任何可迁移的记录，那篇预测认知健康提升，本章预测这个位置在一代人内空置。与「机器答案到达前留一份独立记录」那一篇。 那篇的底稿是给个人学习用的，本章的物证是给共同体记账用的。判决点：一个底稿制度完备、但底稿只对本人可见的环境，那篇预测学习改善，本章预测对能力的代际分布没有任何影响——因为不可援引的记录不是账。与「学习的单位是规则的前后差」那一篇。 那篇的单位是个人的判断结构，本章的单位是共同体的科目表。判决点：规则差记录质量相同的两个共同体，若一个把「指出他人规则的错误」记账、一个不记，本章预测十年后两者的诊断能力分布不同，那篇预测无差异。与「中断之后能不能接手」那一篇。 那篇测能力，本章问这个能力有没有账。判决点：接手能力测验分数相同的两组，无账的那组在代际上流失，而测验分数在流失发生前不会下降——因为测的都是现有的人。与「闭合之后还能不能回来」那一篇。 那篇问二次进入的通道是否还在，本章问二次进入这个动作有没有科目。判决点：一个留痕完备、通道畅通、而重新打开旧案的人永远不被记入任何履历的制度，那篇预测纠错能力良好，本章预测没有人会去做，通道空置。两篇的处方因此不同：那篇修通道，本章修科目表。

十三、本章自己落在哪一格

这篇文章做的是提出，不是核验。它提出一个概念、一张表、一组预测，全部落在被记满分的那一类动作上。若它是对的，它自己就是记零位的受益者：它靠提出获得记账，而它所主张的那类动作——比如有人认真核对了本章引用的判例规则是否被准确表述、有人指出第八节的收窄其实使前七节的强断言失效——在任何账上都不会有位置。能做的对冲有限。第一，把最便宜的检验写到可以被别人直接跑：年龄分布与入行人数，任何行业协会的既有数据就能算，不需要作者参与。第二，第八节的两处收窄是被材料逼出来的，两处都削弱了原来的强断言，都写在正文里；其中第一处直接毁掉了本章最自然的处方。第三，明写了那个交出去的洞——既提出又否定的那类动作会让本章的核心边界失效，而本章没有办法处理它。这三条不改变本章所在的格。它们只是把这一格标在明处。

十四、可以判本章错的六种结果

第一，最便宜的一条。取若干个存在明确记零位的行当（同行评议、图书校订、问题分诊）与若干个同类但已建立可迁移记录的行当（漏洞披露、内部审计），比较从业者的年龄中位数与入行人数。本章预测前者的年龄中位数逐年上升而入行人数不降；若两组无系统性差异，核心机制不成立。第二，个人层与共同体层不分离。若在个人层面上，处于记零位的人的能力增长确实低于被记账的位置（在控制练习量之后），那么本章所说的「个人正常、共同体流失」这个双层结构不成立，现象可以被更简单的动机理论解释。第三，物证条件不起作用。若在两个都给否定动作记账的共同体里，一个记的是可核对的物证、一个记的只是声称，而两者的流失情况相同，那么第八节那条收窄失去意义，「不可稀释」不是关键变量。第四，回溯重写无效。若在那些已经开始把否定动作变成可援引记录的机构里（公开具名审稿的期刊、署名核查员的新闻机构、把红队报告作为产出的机构），十年之后该类能力的人员构成没有改善，则本章的正向读数落空。第五，断层可被更简单地解释。若所有的技能代际断层案例都能被自动化替代、产业转移或薪酬差异解释，不需要引入记账结构，则本章是多余的。这一条是本章最容易被击中的地方，因为薪酬本身与记账高度相关，两者难以分开。第六，本章自陈最软的一处。两根轴也许是一根：所谓「替代通道」，本质上就是那个动作在另一套账上被记了账。若如此，第二根轴只是第一根轴的重复，四格表应当合并成一根轴上的程度差。要把它们分开，必须找到一个替代通道与记账完全无关的案例——比如某个位置靠纯粹的师徒关系而非任何账目延续了几代。本章目前没有这样的案例。正向读数。若本章成立，应能观察到：给否定动作建立不可稀释物证的共同体，其该类从业者的年龄中位数在若干年后停止上升；且这个变化应当早于该共同体的错误率下降出现。顺序比相关更难被巧合满足。写死日期的赌注。到二〇三三年十二月三十一日，至少应有一个主要学科的评价体系把「可核对的否定型产出」（具名且可援引的审稿意见、可复现的证否、被采纳的勘误）列为独立的、可写进晋升材料的科目，而不再只作为服务性工作在附录里提一句。若届时没有任何这样的体系，且上述年龄分布的差异被证明不存在，本章关于记账结构决定能力构成的主张即告失败。

结　论

一件共同完成的作品留下的能力归谁，三个领域给出三个不相容的答案：归执行的手、归构思的人、归署名制度。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——账本记得对不对，不影响能力实际长在哪里。把三家的材料串起来，这个前提就不成立。能力确实由执行沉淀，某一类执行确实被系统性地记为零，而账本确实决定下一次谁被允许执行。三条合起来只能得出一件事：被记为零的那类动作，其能力会在个人身上正常沉淀，然后在共同体层面消失，而消失的过程不出现在任何指标上。被记零的动作有一个共同特征：它的成果是一个否定，是某件事没有发生。而没有发生的事不留痕迹，因而无法计量、易被稀释、无法向第

三方证明。账本天然偏爱能被指着看的东西，这不是谁的疏忽。因而处方也不能是「给核验记功」——账目一开，记录的就不再是核验，而是关于核验的声称。唯一能走的是更窄的一条：让否定动作留下一件在事后无法重制的物证，然后给物证记账。飞行安全给报告记账而不给发现记功，漏洞编号让一个可复现的缺陷成为硬通货，走的都是这条路。一个共同体最后擅长什么，不由它奖励谁决定，由它把谁记为零决定。而要知道它把谁记为零，只需要查一件事：在它的正式记录里，发现了一个错误，被记在谁名下、算多少。

注　释

· 本章所说的「账」限于可被第三方读到、可被援引、且能随人迁移的正式记录；口头声誉与内部评价不计入。

· 专利发明人认定的规则在不同法域有差异，本章只使用「构思与付诸实施相区分、且实施与常规核验不构成发明人资格」这一为多数法域共有的骨架，不涉及具体案件的适用。

· 作坊归属的重审是一项持续数十年、结论不断修订的工作，本章只使用其两条结论：署名与手笔在历史上不重合；归属可被回溯重写且重写会改变现时资源流向。

· 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。

人机分工声明

本章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，其中第一处直接毁掉了本章最自然的处方，全过程保留在正文中。本章未标注任何评分——按本站惯例，参与改写者不为自己改写的文本打分。证据边界专利发明人认定只使用为多数法域共有的骨架（构思与付诸实施相区分），不涉及具体案件的适用；作坊归属研究只使用其两条已被广泛接受的结论。第十四节六条判据本章均未执行，其中第一条所需的年龄与入行数据在多数行业协会的既有统计里已经存在。字数　纯汉字约 1.2 万　版本　v1.0 · 2026-08-01 三边对「一件共同完成的作品，能力沉淀给谁」给出互相排斥的答案：学习科学说归执行的手，专利法说只有构思者算、核验与实施一律记零，作坊归属研究说署名制度在能力之前。本章的两处收窄由后两边逼出：专利法说明记零不是疏忽而是对「可稀释性」的回应，因而处方不能是给核验记功；作坊研究说明归属可被回溯重写，因而账本不是稳定的自变量，本章只能主张一条条件性趋势。

第四编

那么，「知道」还剩下什么

编　序

前三编都在说什么正在消失。这一编只问两件事：还剩下什么可以读，以及读它的资质从哪里来。

第九章给出还能读的那一样。它的判断本身是全书形式上最漂亮的一处：那个只在「什么也没要求」时才存在的量，与「被要求时能出多大反应」，是同一个量——这是物理学自己的一条关系式，而物理学从没拿它当作「可改变性该在什么时候测」的证据。要紧的不是三门学科没给它设字段，而是三门学科都设了字段，并且都在花钱把它降到零。

第十章给出读它的资质。它把一条日常动作拆开：找一个更懂的人问。这个动作在有些领域里省去了判断，在另一些领域里没有——分界不在诚信，不在权威，在一个很窄的地方：你用来相信他的那个理由，一个完全不懂这一行的人能不能用。能用，依赖既理性又安全；不能用，你正在做的不是外包判断，而是行使判断，只是把它移到了一个不被记录、也不被自己察觉的位置。

第十章因此是全书唯一一章直接处理「还有没有价值」这个问法的。它给出的答案不是一个比例，是一张分界表。而合章的第七条关系说明了为什么这张表在往一个方向移动：被依赖的那一样越好用，评判它好不好用的能力越少。

第九章

没有人要求的时候

论「没有问题」与「已经不会变了」何以在全部现行读数上同形

学科入口　认知科学 × 物理学 × 工程学

摘要

一个系统还能不能被改变，凭什么读得出来？认知科学答：看它的内部模型还在不在被更新——预测误差为零就是学完了。物理学答：看它受到扰动时给出多大响应——响应函数就是全部信息。工程学答：看它还剩多少余量、处在什么工况里——余量是唯一的健康判据。三条不能同时为真，而三条共同假定了一件谁也没说出口的事：一样东西的能力，只有在它被要求做事的时候才存在、才可测。 推翻这条假定的材料来自物理学自己：涨落与响应之间的那条关系式说，一个系统在平衡时的自发涨落，与它此后对外扰动的响应，在数值上是同一个量。也就是说，那个只在"什么也没要求"时才存在的量，恰好就是"被要求时能出多大反应"。物理学用这条关系式从平衡涨落算输运系数，从没有人把它当作"可改变性该在什么时候测"的证据。本章据此命名静息变异——在没有任务、没有误差信号、没有外部扰动的时段里，系统自发做出的与它当前默认做法不同的那些事——并给出可事后清点的比率读数静息变异率。要紧的不是三家没给它设字段：三家都给它设了字段，而且三家都在花钱把它降到零（认知学记为失误率、物理学记为噪声、工程学记为非计划偏移）。三家共同出资消灭的那个量，正是它们后来要用的那个能力。 由此得到一张两轴四格：现行表现合格而静息变异为零的那一格，全部现行指标最好看，失效不产生任何信号，且越正规化越严重。判据不问人的判断，只问记录：在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？ 一份二〇一四年的座舱实测为此提供了现成的检验：自动化程度提高之后，飞行员用手做的那部分技能保持得相当完好，退化的恰是用脑做的那部分——而那正是自动化替他做了、因而在整段航程里没有人要求他做的那部分。结论是一句本章不打算缓和的话：可靠性要求把变异消灭，可改变性要求把变异留住；同一个系统被同时要求做到两件事，而这两件事用的是同一笔预算。关键词：静息变异；静息变异率；被改变性；余量掩蔽；调用

即存在；涨落与响应 When Nothing Is Asked*Why "No Problem" and "No Longer Changeable" Look Identical on Every Current Instrument, and Why the Quantity That Separates Them Is Being Deliberately Driven to Zero by Three Disciplines at Once*Abstract. How can one tell whether a system is still capable of being changed? Cognitive science answers: watch whether its internal model is still being updated — zero prediction error means learning is complete. Physics answers: watch how large a response it gives to a perturbation — the response function is all the information there is. Engineering answers: watch how much margin remains and what operating conditions it sits in — margin is the sole health criterion. The three cannot all be true, and all three rest on an unstated shared premise: *the capacity of a thing exists, and can be measured, only while something is being demanded of it.* The material that overturns this premise belongs to physics itself. The fluctuation–dissipation relation states that a system's spontaneous equilibrium fluctuation and its subsequent response to an external perturbation are numerically the same quantity. The quantity that exists only when nothing is asked *is* the quantity that determines how much can be got when something is asked. Physics uses this relation to compute transport coefficients from equilibrium fluctuations; no one has used it as evidence about *when* changeability should be measured. This paper names the resulting object resting variation — the variants a system spontaneously produces, differing from its own default, during intervals in which no task, no error signal and no external perturbation is present — and supplies a countable ratio, the resting-variation rate. The point is not that the three fields fail to record it: all three record it, and all three spend money driving it to zero (error rate, noise, unplanned deviation). *The quantity the three fields jointly pay to destroy is the capacity they will later need.* A 2×2 follows, whose "compliant performance, zero resting variation" cell scores best on every existing indicator, emits no signal on failing, and worsens with formalisation. The criterion asks records, not judgements: *in the most recent stretch during which no task was assigned, no assessment applied and no fault occurred, how many things did this system do that differ from its default, and who wrote them down?* A 2014 flight-deck study supplies a ready test: with increased automation, pilots' hand skills were largely retained while the cognitive

skills degraded — precisely the part the automation performed for them, and therefore the part nothing asked of them for the whole flight. Reliability demands that variation be eliminated; changeability demands that it be kept; and both are drawn on the same budget.

一、导论：三个不该同时为真的判断

先摆一个具体处境。一条产线连续两年零缺陷，一门课连续五届通过率九成五，一位飞行员连续八年零事件，一个团队连续十二个季度全部指标达标。四个系统，全部读数完美。现在问一句：它们还能不能被改变？这个问题不是"它们好不好"。好不好已经答完了，答案是很好。问的是另一件事：如果明天出现一个它从未遇到过的要求，它拿得出反应吗？眼下有三种回答，各有充分的经验支持，而它们不能同时为真。第一种来自认知科学的主流一支。它说：看它的内部模型还在不在被更新。一个系统之所以改变，是因为它的预测与实际之间还有落差；落差驱动更新，更新缩小落差。落差为零，就是模型已经收敛，就是该学的都学完了，此时不再改变是正确的，不是病。 判断一个系统还能不能被改变，看它的误差信号还在不在就够了。第二种来自物理学的线性响应理论。它说：看它在受到扰动时给出多大的响应。响应函数是全部信息——你想知道一个体系能被外力推动多少，就施加一个小外力，量它动了多少。没有比这更直接的读法，也没有别的读法。第三种来自工程学的可靠性一支。它说：看它还剩多少余量，以及它正待在什么工况里。安全系数、降额、冗余度——这些数决定了它能承受什么。判断一个系统的健康，看余量与条件就够了；余量足、条件在包线内，它就是健康的。三种回答在处置上互相取消。误差为零时，认知这一支说不必动，动了是浪费；而工程这一支会说误差为零可能只是因为没有人给过它一个够难的工况，那叫未经检验，不叫健康。工程这一支说余量足就好，而物理这一支会说余量是一个静态的量，它说不出体系对一个具体扰动的响应有多大——响应要量才知道。物理这一支说去施加扰动量响应，而认知与工程两家会同时反对：施加扰动本身就改变了被测的东西——对人是给了一次训练，对设备是消耗了一次寿命。更要紧的是，三家各自每天在做的事，正是另一家诊断为病根的那件事。压低误差是认知这一支的日课，而工程这一支的余量学说恰恰指出：误差被压到零的系统，其退化不会产生任何信号。施加扰动测响应是物理这一支的日课，而工程规范里"非破坏性检测"这一整套东西存在的理由，正是"为了知道它好不好而去动它"这件事本身不可接受。留足余量是工程这一支的日课，而认知这一支会说余量正是让误差归零的装置——余量越足，落差越小，更新越少。本章不打算在三者之间判谁对。判谁对需要它们站在同一根轴上，而它们不在：一家看内部路径，一家看外显响应，一家看所处条件。本章要做的是找出一条它们共同假定、而谁也没有说出口的东西，并用其中一家自己已经掌握、已经写进教科书、只是没有往这个方向用的材料把它推

翻。那条共同假定是：> 一样东西的能力，只有在它被要求做事的时候才存在、才可测。三家争的是该看哪一维。三家都没有争的是：该在什么时候看。推翻它的材料在第四节给出，来自物理学。而由此得到的判断，三家都不会同意，但三家的证据合起来只能得出它：> 一个系统还能不能被改变，既不由它的误差信号读出，也不由它受扰时的响应读出，更不由它的余量读出，而由它在没有人要求它做任何事的那些时段里，自发做出多少与它当前默认做法不同的事读出。而这个量，三门学科都给它设了字段，都给它起了名字，都在花钱把它降到零。这条判断的可查形式很短，全文都可以拿它去问：在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？路线：第二节交代三家各自的立场与各自留下的一处松动；第三节把冲突写成一个可判定的形状；第四节找出共同假定并给出推翻它的材料；第五至八节建立本章的读数、辨别格与"为什么不产生信号"的机制；第九节用一份已经做过的座舱实测做检验，那份实测的结果同时修正了本章最直觉的一条说法；第十节在十二个行当里兑现；第十一节给判据与不得越过的伦理界线；第十二至十五节与最接近的既有说法、与学科通融专栏已发的几篇、与三家各自的理由划清界线，并指出其中一家在不承认本章这个量时内部不自洽；第十六节写下三家反过来给本章加的三条约束；余下各节交出边界、证伪条件、反噬预言、自反定位与一条写死日期的赌注。

二、三家的立场，以及各自留下的那一处松动

2.1 认知科学：看内部模型如何被更新就够了

近二十年认知科学最有野心的一支，把知觉、行动、注意与学习统一在一条原则下：生物体在最小化预测误差（或其变分上界）。这一支的强版本主张是单因的：驱动一切认知改变的只有一样东西，就是预测与实际之间的落差。 它给出的时序读数很清楚：落差先出现，模型在随后的若干次接触里被更新，行为最后跟上；落差归零，更新停止。这一支自己留下的松动，是它被提出得最多的那一条反驳，而且这一支自己承认它是最要紧的一条：暗室问题。如果最小化预测误差是唯一的驱动，那么最省事的做法是躲进一间没有任何刺激的暗室——那里一切都完全可预测，落差恒为零。生物体显然不这么做。[1]这一支的回答是：要在时间上延展的序列上做推断，就必须把认识论价值算进来——一个策略的价值不只是眼下满足预期，还包括它带来多少信息增益；后一项在数学形式里自然出现，不是外加的。[2]这个回答是本章要用的那处松动。 它等于承认：一个只按当下落差行动的系统会走进暗室；要不走进暗室，系统必须去主动寻求它现在还没有的落差。而"值得去寻求哪一处落差"这件事，在落差还没出现的时候就要被决定。这一支从没有问过：那个在落差为零时决定系统去哪儿的东西，是什么，怎么测。

2.2 物理学：看它受扰时给出多大响应就够了

线性响应理论说：给体系一个小的外场，量它的响应，响应函数就写尽了这个体系在这个方向上可被推动的全部性质。这是物理学里最成熟、最可靠的一套读法，输运系数——电导率、粘度、扩散系数——都是这么定义的。它的单因主张同样清楚：可被推动的程度是一个由响应函数单独决定的量，与体系此刻正在做什么无关。 时序读数：外场先加，响应随即，弛豫在特征时间内完成。这一支自己留下的松动，就是它自己最著名的那条定理，第四节展开。这里先记一句：这条定理说，响应函数并不是一个只有施加外场才存在的东西——它在没有外场的时候就已经确定了，而且在数值上等于体系此刻自发涨落的关联。物理学把这条定理用来做一件很实际的事：从平衡态的模拟里算出输运系数，不必真的施加外场。

2.3 工程学：看余量与所处条件就够了

工程学的可靠性一支主张：一个系统能不能承受未来的要求，由它当下的余量与所处工况单独决定。安全系数、降额、冗余、包线——整套规范、认证与验收都建立在这上面，而且它极其成功：现代工程物的可靠性正是靠这一套做出来的。时序读数：余量在设计期定死；使用中被逐步消耗；失效在余量耗尽的那一刻出现。这一支自己留下的松动，是它整个下游产业存在的理由：余量在被消耗的过程中不产生任何信号。系统在余量耗尽之前一切正常、全部检查通过、所有指标合格；耗尽的那一刻突然失效。若余量的剩余量可以直接读出，那么剩余寿命预测这一整个学科就没有存在的必要——它存在，恰恰说明"没有故障"与"健康"是两件事，而现行仪表读不出后者。工程学内部为此有一场公开的、持续的争论：一派把安全定义为事故的缺席，靠余量、屏障与合规来保证；另一派主张安全是在变动条件下仍然成功的能力，而日常工作中的变异不是要消灭的威胁，恰恰是成功与失败共同的来源。本章取前一派为工程这一家的立场，因为它是制度化的那一派；后一派在第十二节作为最接近的几种说法之一处理，那是本章必须最小心划界的一位。

2.4 三家并排

	主张看哪一维	单独够用的那一样	时序	自己留下的松动
认知科学	内部模型如何被更新	预测误差	落差→更新→行为	暗室问题：系统必须主动去找它还没有的落差
物理学	受扰时的响应	响应函数	外场→响应→弛豫	响应在没有外场时就已确定，且等于自发涨落
工程学	余量与所处工况	剩余余量	余量→消耗→失效	余量在被消耗时不产生任何信号

三家各锁一维，互不兼容：误差是一个路径上的量，响应是一个外显的量，余量是一个条件上的量。三家都认为自己那一维单独够用。而三家的松动指向同一个方向——都指向"在没有要求的时候发生了什么"。第一家承认系统必须在落差为零时决定去哪儿，第二家承认响应在无外场时已经定了，第三家承认退化在无故障期不留痕迹。三家都走到了这扇门前，三家都没有推门，因为三家的问题都不需要推它。

三、冲突定位：三家在同一个具体处境里给出互相取消的处置

取一个足够具体的处境：一家运行八年、连续三年零事件、全部审核通过的电力调度中心。上级要问一句：它扛不扛得住一次从未演练过的复合故障？- 认知这一支说：看调度员的误差信号。近三年误报率、操作偏差率、复盘中发现的判断失误全部逐年下降且已接近零。结论：模型已收敛，扛得住。- 物理这一支说：误差为零说不出响应有多大，得量。做一次突发扰动演练，看多久恢复、恢复到什么水平。结论：不演练就无从判断；但演练本身消耗资源、扰动生产，而且演练过一次之后，那个扰动就不再是从未遇到过的了。- 工程这一支说：别去动它。看余量——备用容量、切换冗余、人员冗余、时间裕度。余量在包线内就是健康的。结论：余量足，扛得住，且不需要付出任何代价去证明它。

三条不能同时为真。若认知这一支对，则误差归零之后再投入任何资源都是浪费；若物理这一支对，则不施加扰动就永远说不出结论，而这与工程这一支"不该为了知道好不好而去动它"直接冲突；若工程这一支对，则响应的大小可由静态的余量推出，而这正是物理这一支否认的——余量说的是能承受多少，响应说的是会动多少，两者在非线性体系里根本不是一个量。而三家的松动合起来，指向一件三家都没有说的事。认知那一支承认：系统必须在落差为零时决定去哪儿。物理那一支承认：响应在无外场时就已经定了。工程那一支承认：退化在无故障期不留痕迹。三条说的是同一件事的三面：决定"扛不扛得住"的那件事，发生在没有人要求它扛什么的那段时间里；而三家的全部测量都安排在有要求的时候。于是三家的争论可以写成一句话：三家争的是"可被改变的能力"该由哪一维来读。 而三家共同假定的是：这个能力是一样只在被调用时才现身的东西——不调用，它就不在那儿，也就无从谈起。

四、共同假定，与推翻它的那件材料

4.1 把共同假定说出来

摊开写，是这一条：> 一样东西的能力，只有在它被要求做事的时候才存在、才可测。不被要求的那段时间，是这个系统的空白期——那里没有可读的东西，因为那里什么也没有被调用。把它念给三家听，三家都会说"这还用说吗"。认知实验只在任务期记录；工程监测只在运

行工况下取数；物理测量只在施加外场时进行。三家的仪器全部是在有要求的时候才打开的。三家都没有争论这一条，因为这一条正是"测量"这个词的意思——测量就是一次调用。

4.2 推翻它的那件材料：来自物理学自己

物理学有一条写进了每一本统计力学教科书的关系式：一个体系对小外场的线性响应，可以用它在没有外场时的自发涨落表达出来。 换句话说，响应函数与平衡态涨落的关联函数是同一个量的两种写法。[3]这条关系式在物理学里回答的是一个很实际的问题：怎么不施加外场就把输运系数算出来？ 答案是去看平衡态的自发涨落——扩散系数从粒子位移的自关联算，电导率从电流涨落算，粘度从应力涨落算。它是二十世纪统计力学最有用的结果之一，也是分子模拟里最常用的一套方法。把它剥到只剩形式，去掉物理学的一切名词，只留"什么对什么做了什么，导致什么"，它说的是：> 那个只在"什么也没要求"时才存在的量，与"被要求时能出多大反应"，是同一个量。这正是共同假定所排除的那一类事实。共同假定说：不被要求的那段时间是空白期。这条关系式说：不被要求的那段时间不但不是空白期，它恰恰是那个能力被完整决定的时间；有要求时量到的，只是把它读出来而已。三个部件同时被打掉：1. 能力不是只在被调用时才存在的。 它在没有被调用的时候就已经确定，调用只是读数。 2. 不被要求的那段时间不是空白期。 那里正在发生的自发变动，就是那个能力本身。 3. 所以三家的仪器全部开在错误的时刻。 不是精度不够，是打开的时机不对：它们在能力已经被决定之后才去量它，而在能力正在被决定的时候关着。

4.3 一处必须先说清的限度

必须马上写下这一条，否则后面全部是修辞：上面那条关系式是一条物理定理，它的严格成立需要体系处于平衡附近、且外场足够小使响应是线性的。教室、产线、机组都不满足这些条件。本章不主张这条定理适用于它们。本章从它拿走的不是定理，是它所揭示的那个区分——"能力何时被决定"与"能力何时被读出"不是同一件事，而全部现行读法都把两者当成同一件事。这个区分的成立不依赖平衡态，也不依赖线性。而且，这条定理失效的那些情形，恰恰给了本章最要紧的一条限制：在被困住的、缓慢老化的体系里（玻璃、自旋玻璃、粗化中的畴），涨落与响应之间的比例关系会破缺，人们只好引入一个"有效温度"来描述破缺的程度。[4] 也就是说：在一个已经被困住的体系里，自发涨落还在，而它不再预示响应。 这条限制在第十六节被写成本章最重的一处让步。顺带说，这一破缺是否普遍成立，物理学内部至今有争论——有实验在软玻璃材料里跨几个频率量级没有找到任何破缺。[5] 本章不站队，只取其中对自己不利的那一支作为约束。

五、静息变异：定义与五条判定条件

把上面那个区分落到可清点的东西上。> 静息变异：在没有任务、没有考核、没有误差信号、没有外部故障的时段里，一个系统自发做出的、与它自己当前默认做法不同的那些事。不是所有的"不一样"都算。要算，五条同时成立，缺一条不算：1. 发生在无要求时段。 该时段内没有分派的任务、没有正在进行的考核、没有待处理的故障、没有截止期。这一条是全部定义的地基；判断一个时段算不算无要求，看的是有没有一个外部的人或程序在等这个系统交出什么。 2. 与本系统自己的默认做法不同。 不是与别人不同，也不是与标准不同——是与它自己上一次做同类事情的做法不同。默认做法必须能被指认出来（有成文规程的看规程，无成文的看最近若干次的众数）。 3. 不是失误。 失误是想做默认那件事而没做成；静息变异是没想做默认那件事。区分的判据是事后是否被本人或本系统采纳为一个可重复的选项——采纳了的是变异，改回去了的是失误。这一条在实操上最难，第十一节给了两条降低误判的办法。 4. 无奖惩。 该次变异既不带来记功也不带来记过。带奖惩的变异是被要求的变异，不算——它只是把要求换了个名字。 5. 被记下来了。 存在一份能指到具体时段、具体动作的记录，且这份记录不是为了考核而建。没有记录的静息变异在制度上不可与随意区分，因而不计入。第五条与第四条合在一起是本章最尴尬的一处：要算数就得有记录，而建立记录这件事本身很容易变成一次要求。 这不是一个可以靠更精细的定义化解的张力，它是本章的反噬所在，第十九节正面处理，本章找不到出口。

六、静息变异率：三个领域量纲不同而比率相同

> 静息变异率 v = 无要求时段内符合上述五条的动作次数 ÷ 该时段内的动作总数。三个领域的兑现： - 认知／人：无人布置、无评分的那段时间里，做同类事情时采用了与自己默认做法不同的做法的次数占比。数据来源：自习记录、非任务期的操作日志、无任务的模拟器时段。 - 物理／材料：平衡态下自发涨落的无量纲化幅度（涨落的关联相对于该体系的特征尺度）。这一处是唯一有严格定义的，也是本章这个读数的来路。 - 工程／组织：无故障期内，非计划的工况偏移与试验性配置变更占该期全部操作的比例。数据来源：变更单、批记录、版本库的提交历史。

四条性质：其一，无量纲。 三个领域可以并排比较。其二，可事后清点。 所需的原始记录在多数机构里已经存在，不必新建采集系统：版本库有提交历史，工厂有变更单，模拟器有时段日志，学校有非上课时段的机房与实验室记录。其三，它把一个印象变成一个数。 "这个系统还有没有活气"这句话此前只能靠感觉，现在是一个分数。其四，也是最容易漏又最要命的一条：它与全部既有的主要指标正交。举一个既有指标最好看、而这个数最难看的具体形状：

一条连续两年零缺陷、变更单数量为零、SOP 三年未修订的产线。 按现行的任何一套质量与可靠性指标，它是标杆；按 v 看，它是零。反过来，v 高不等于好——一个变更频繁的系统可能只是失控。v 不是一个越高越好的分数，它是一个方向读数：它读的是这个系统在没有人推它的时候还动不动。再举两个更刺眼的：一位连续八年零事件、每次复训一次通过、从不做任何非标准操作的飞行员，v 是零。一个通过率九成五、评教满分、题库十年未增删的课程，v 是零。这两个正是各自行当里被当作模范的那一类。

七、辨别格：现行表现 × 静息变异

单有一个比率还只是一个名字。把它升成两根轴，四格才分得开那些从外面看一模一样的系统：

	静息变异高（无人要求时仍在自发做不同的事）	静息变异低（无人要求时一动不动）
现行表现合格（有要求时全部指标达标）	甲 · 有余地 达标而且还在自己动。唯一能在新要求出现时给出反应的一格。代价是它看起来"不够规范"。	乙 · 焊住 全部指标最好看的一格，也是本章要打的那一格。达标、稳定、零变更、零缺陷——而它对任何未曾预演过的要求都没有反应。
现行表现不合格	丙 · 乱 又达不到又乱动。看得出来，不需要任何框架。	丁 · 停 又达不到又不动。看得出来，不需要任何框架。

乙格是本章要打的那一格，也是四格里唯一需要一个框架才看得见的一格。 丙与丁不需要本章：谁都看得出乱和停。甲需要的是别被误判为丙——这一点第十一节的伦理条款会处理。乙格具备三条签名，缺一条则说明第二根轴选错了：1. 它在短期指标上表现为改善。 缺陷率、事件率、通过率、达标率全部向好，而且是真的向好，不是造假。 2. 它的失效不产生任何信号。 因为整套记录只在有要求的时候打开，而乙格的全部特征都发生在没有要求的时候。它不是信号微弱，是根本没有一台仪器开着。 3. 它自我加固。 越正规化越严重：变更管制越严、SOP 越细、合规审核越密，v 关得越死。一家管理松散的工厂反而更可能有非计划的试验批；一家管理精良的工厂连这个口子也补上了。第三条决定了乙格不是失职的产物，而是尽职的产物。这是本章与所有"人不上心""管理不到位"式解释的分界，也是本章最不肯让步的一处。

八、为什么焊住不产生信号：余量掩蔽

乙格的第二条签名值得单独说清，因为它是"为什么大家一直没发现"的全部答案。工程学早就知道一件事：余量在被消耗的过程中不产生信号。一个部件从新到旧，其承载能力一路下降，而只要还在余量之内，全部检测、全部指标、全部运行数据都是正常的。剩余寿命预测这一整个学科的存在，就是因为这件事："没有故障"与"还剩多少能力"是两个量，而现行仪表只读得出前一个。本章要说的是同一个机制在另一根轴上的复现，而且更彻底：- 余量掩蔽退化：余量还在，所以退化不显形——但退化本身是连续的，一旦装上合适的传感器，退化特征是可以被提取的。这是剩余寿命预测能做的事。- 要求掩蔽能力流失：只要日常的要求都在系统的能力之内，能力流失就不显形——而且这一次连传感器都装不上，因为要装的那台传感器必须在"没有要求"的时候工作，而现行的每一台仪器都是被要求触发的。

两者的关系是：余量掩蔽是一个可以靠更好的传感解决的问题；要求掩蔽是一个不能靠更好的传感解决的问题，因为传感这个动作本身就是一次要求。要读到那个量，唯一的办法是事后清点无要求时段里已经发生过的事，而不是在无要求时段里去测什么。这一条同时说明了为什么乙格里的系统本身也察觉不到：它的每一次自我评估都是一次要求。一个系统只要开始检查自己，那一刻它就已经不在无要求时段里了。乙格的当事人不是不肯看，是每次一看，要看的那个东西就不在了。

九、一份已经做过的观测：自动化座舱里流失的是哪一部分

本章最自然的一条推论是：用得少的能力会流失。这条推论太顺了，顺到必须找一份能反驳它的实测。已经有一份。二〇一四年一项在波音七四七模拟机上做的研究，让十六名航线飞行员在系统变化的自动化水平下飞常规与非常规科目，评分并当场追问他们在想什么。结论出乎意料：飞行员的仪表扫视与操纵杆量这类"用手做"的技能保持得相当完好，退化的是"用脑做"的那一部分——导航、保持对飞行状态的整体把握、诊断异常。研究者自己的总结是：对飞行员用手做的那些事可以少担心一点，对他们用脑做的那些事应当多担心一点。[6]这份结果直接推翻了"用得少就会流失"这条朴素推论：用手那部分其实也用得少了（自动驾驶接管了绝大部分航段），却保持得好；用脑那部分退化了。所以决定流失与否的不是使用频次。决定它的是别的东西，而这份数据恰好给出了那是什么。手上那部分技能在每一次起飞与落地都被真实地要求一次——那两段是人在飞，而且做得好不好当场就知道。它有要求，也有反馈。脑上那部分不同。在巡航段，自动化替飞行员做了导航与状态跟踪，因而在整段航程里没有任何人、任何程序要求飞行员说出"我们现在在哪儿、下一步该怎么办"。他可以说，也可以不说；说了没有奖励，不说也没有惩罚，而且不说不会产生任何后果——直到有一天自动化退出。也就是说：手上那部分处在"要求少但有要求"的状态，脑上那部分处在"完全没有要求"的状态。前者

的能力被反复读出，所以流失可见并被补训；后者的能力从不被读出，所以流失不可见，也就没有人补。这一节因此做了三件事。第一，它给了本章一份现成的、非本章设计的经验支持：能力流失的分界不在使用频次，在那段时间里有没有人要求它。第二，它修正了本章最自然的一条说法。本章原本会写下：用得越少流失越快。这句话现在写不下去了。改成：在有要求的时段里，使用频次决定流失速度；在完全没有要求的时段里，使用频次这个变量根本不存在，决定流失的是这个系统自己动不动。前一句是既有文献早就说过的，后一句才是本章的。第三，它给出一条本章自己的检验。若上述解释成立，那么在同一批飞行员里，那些在巡航段自发做非要求的事的人（心里跟着算位置、自己先给出下一步方案再看自动化怎么做、主动构想异常），其脑上那部分技能的保持应当显著优于那些不做的人；而两组人的手上技能不应有差别。这份对比所需的资料——巡航段的自发行为记录——现有研究没有采集，但采集成本极低：巡航段本来就是空闲的。

十、在十二个行当里兑现

一条判断若只在提出它的那个场景里成立，它就是一个说法。下面十二处不是类比，每一处都能据此预测一件具体的、可以去查的事。一 · 有备案考纲的大学课程。无要求时段＝寒暑假与学期间隙。v 通常为零，且零有制度理由：假期里改讲义不算工作量，不进任何考核。预测：这类课程在遇到新形态的学生（换生源、换先修课、换工具环境）时的调整周期，与其教师的教学评分不相关，而与其假期讲义修订记录相关。二 · 住院医师培训。无要求时段＝非值班时段。v 非零但不入账——自发的病例讨论、自己去查文献重看一遍某个决策，每天都在发生，而没有任何系统记录它。预测：一旦要求记录，最先发生的不是记录增加，是这类行为减少。三 · 航线飞行。见上一节。这是十二处里唯一有现成实测的。四 · 软件团队。无要求时段＝无工单期。v ＝无工单关联的提交占比。这是最便宜的一处检验，数据全在版本库里，一条查询就能算。预测：无工单提交占比与该团队应对一次架构级新要求的时间负相关，且这一关系在控制团队规模、总提交量与技术栈之后仍然成立。五 · 制造产线。无要求时段＝无故障、无订单变更期。v ＝非计划工艺试验批占比。预测：v 为零的产线在遇到新牌号、新原料批次时的调试时间显著更长，而这一差别在其历史良率上完全看不出来。六 · 电网调度。无要求时段＝无扰动期。v ＝该期内自发做的非规定推演与试运行占比。预测：黑启动能力与近三年事故率不相关，与无扰动期的自发推演次数相关。七 · 消防与救援。无要求时段＝无警期。同上。预测：处置陌生灾情的表现与出警次数不相关，与无警期自发推演的多样性相关。八 · 医院应急。无要求时段＝平时。二〇二〇年的一次全球性检验已经跑过一遍，而当时的普遍读数是"平时演练最多的未必最好，平时敢改流程的更好"——这一处的历史数据可查。九 · 品种选育与农业。无要求时段＝丰年。v ＝丰年里试种非主栽品种的地块占比。预测：一个地区在遇到

新病害时的恢复速度，与其丰年试种比例相关，与其历史平均单产不相关。十·军事训练。 无要求时段＝非任务期。v ＝该期演训中偏离标准战法的次数占比。预测：面对新战法的适应速度与训练时长不相关，与偏离次数相关。十一·个人。 无要求时段＝下班后、假期、无考核期。v ＝该期自发学与自发做的、与本职默认做法不同的事的比例。这一处最私人，也最容易被误用，第十一节的伦理条款首先针对它。十二·大模型与它的评测。 这一处最尖锐。一个模型在被评测时的表现，与它在无提示、无奖励信号时的行为多样性，是两个量；而当前的训练目标里，压低后者是显式写进去的——降低随机性、提高一致性、消除"不必要的"变化，全部是被明确追求的指标。按本章的读法，这套目标在同时优化第一根轴与最小化第二根轴，即在结构上把系统推向乙格。预测：同代模型中，在温度固定、无系统提示条件下自发行为多样性更低的那些，在遇到训练分布之外的新任务类型时适应更慢，而这一差别在全部标准评测集上读不出来——因为标准评测集本身就是一次要求。十二处里，第三、四、六、十这四处的 v 相对更高，而它们有一个共同点：这些行当的失败会以外人可读的形式一次性出现（飞机掉下来、服务崩掉、电网黑掉、仗打输），因此它们被迫在无要求时段里做点什么。教育与医疗的失败以缓慢、可归因于个体、且延迟数年的形式出现，这是它们 v 结构性偏低的根本原因，也是任何一份"向航空业学习"的建议在这两个行当落不了地的原因。

十一、判据、读法与不得越过的界线

11.1 那一句问话

本章的判据不问人的判断，只问记录：> 在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？把它拿到六个场景跑一遍，六个答案互不相同，其中两个是查出来的而不是想出来的：

场景	答案
有备案考纲的大学课程	零。且零有制度理由：假期改讲义不进任何考核
住院医师培训	非零，但无记录、无记名，在制度上不可与随意区分
开源项目	非零且高，无议题关联的提交带作者与时间戳（查出来的）
航线飞行的巡航段	现有记录里完全没有这个字段——巡航段的自发认知活动从不被采集（查出来的）
有变更管制的产线	接近零，且趋零是变更管制的显式目标
一个人的业余时间	非零，但记录属于本人，且一旦被机构采集就不再是无要求时段

第三、四两处不能从命题直接推出来——它们是去查了才知道的，而且正是它们让第十节末尾那条"为什么有的行当高有的行当低"有了答案。

11.2 三种分层引用

本章的主张分三层，各自可以单独被引用、单独被推翻：- 第一层（最强、最易倒）：可被改变的能力在无要求时段被决定，而全部现行读数都开在有要求的时候。- 第二层：v 作为跨领域可比的读数，以及那张两轴四格。- 第三层（最稳、最便宜）：一条不依赖前两层的经验主张——在同一组织内的一组同类团队中，无工单期的自发变更占比与该团队应对一次结构性新要求的用时负相关。 这一条只需要版本库的一条查询和一次事件复盘，不需要接受本章的任何概念。

第一层被推翻时，后两层仍然站得住。本章明确写下这一点，既是诚实，也是保护。

11.3 伦理界线：四条写死的规定

v 是一个比率，而比率几乎必然会被拿去考核人。四条必须写死，缺一条本章不可被用于任何考核：1. v 指的是位置与时段，不是人。 它读的是"这个岗位、这条产线、这门课有没有一段真正无要求的时间，以及那段时间里发生了什么"，不是"这个人上不上进"。一条产线 v 为零，绝大多数情况下说明的是它的变更管制条款，不是操作工的态度。 2. 不得采集私人时间。第十节第十一处（个人）只可由本人自用，不可由任何机构采集、要求申报或纳入评价。一旦机构开始采集，那段时间就不再是无要求时段，采到的数据在本章的定义下是无效的——这不仅是伦理问题，同时是效度问题，两者在这里恰好同向。 3. 安全关键系统的合理 v 值低，且低是对的。 麻醉、飞行操作、核电运行、药品放行的规程稳定性本身是安全资产。在这些地方拿本章去催促"多做点不一样的"是危险的误用。这些领域提高 v 的唯一正当途径是在与生产完全隔离的场合（模拟机、试验线、沙箱），而那正是它们已经在做的事。 4. 已经按 v 做出不利安排的机构，必须公开编码规则并提供复核通道。 什么算一次静息变异、什么算失误、无要求时段怎么划，必须是公开可查的成文规则；被判定的一方必须能提出复核。另有两条降低第五节第三条（变异与失误的区分）误判的办法，一并写在这里：其一，只清点事后被采纳为可重复选项的那些——采纳这个动作留痕，且由本系统自己完成。其二，双人独立编码加第三人复核，编码者不得知道被清点对象的绩效表现。

十二、与最接近的几种说法的分界

本章的判断很容易被读成某个既有概念的改名。下面逐一交代，每一条给出判定形式：若观察到甲，则本章对而该概念错；若观察到乙，则反之。其一 · 任务期运动变异性预测学习速

度。 这是最近的一位，近到必须放在第一条。运动学习研究已经证明：个体在执行任务时表现出的、与任务相关的那部分动作变异越大，其后续学习越快；而且这种变异是被神经系统主动调节的，不是单纯的噪声。它说到哪一步：变异是学习的原料，且可在任务中测出。 分离线：它测的是执行任务期间的变异，并与学习速度相关；本章测的是完全无要求时段的变异，并主张它决定的是对尚未出现的要求的响应。判定形式：若在控制任务期变异之后，无要求时段的变异仍能独立预测系统对一类全新要求的响应，则本章对而该说法不足；若无要求时段变异一经控制任务期变异即无任何增量，则本章被它吸收。 这条检验在运动学习的现有实验平台上可以直接加做——只需在实验前后各留一段"你可以随便动"的无指令时段并录下来。其二 · 探索与利用的权衡。 它说到哪一步：组织与个体在有限资源下要在探索新可能与利用已知之间分配，过度利用导致能力刚性。分离线：探索是在有目标的前提下对资源的一次配置选择——它有预算、有决策者、有机会成本；本章的 v 是无目标时段的自发率，不消耗任何被记账的资源，也没有人决定过要做它。判定形式：若能找到一个探索预算为零、无任何探索性项目立项、而无要求时段自发变异很高的组织（例如一个穷但爱折腾的小团队），且其应对新要求的能力显著高于同规模的高探索预算组织，则本章对而探索—利用的框架在此无解释力。 这样的组织在开源社群里并不罕见。其三 · 组织松弛资源。 它说到哪一步：组织保有超出当前最小需求的资源，这些冗余在环境变动时提供缓冲与创新空间。分离线：松弛是一个存量（可清点、可削减、以资源为单位）；v 是一个行为率（以动作为单位）。二者可以独立取值。判定形式：若存在零松弛而 v 高的系统（资源紧到极点却仍在自发做不一样的事），且其应变能力高，则本章对而松弛资源无话可说；若所有高 v 系统一经细查都有隐性松弛，则本章被它吸收。其四 · 必要变异度定律。 外圈的一位，也是最容易被误当成本章的一位：控制器的变异度必须不小于被控体系的变异度，否则控不住。分离线：这条定律讲的是控制器可能状态的数目（一个容量），处方是增加这个容量；本章讲的是系统在零扰动时实际实现的变异率（一个流量）。二者的差别在一台功能极多而从未被改过任何设置的设备上最清楚：按容量算它变异度极高，按 v 算它是零。 判定形式：若这类设备在遇到新工况时的适应与一台功能少而经常被折腾的设备相当，则本章错；若前者显著更差，则本章对而容量读法在此失效。其五 · 临界性与最大响应。 物理与神经科学里都有的一支：系统调到临界点附近时，其感受率发散，动态范围最大。分离线是数学形状上的：临界性预测感受率作为控制参数的函数有一个峰——离峰越远越差，两侧都差；本章预测的是单调关系（响应随静息变异单调上升），没有峰。判定形式：扫一遍控制参数，看有没有峰。 有峰则临界性对，单调则本章对。这是本章全部划界里最干净的一条，因为它落在曲线形状上而不落在措辞上。其六 · 随机共振。 噪声在适当强度下反而提高系统对微弱信号的检测。分离线：随机共振讲的是噪声帮助检测已经存在但低于阈值的信号；本

章讲的是在完全没有信号时的自发变异决定日后对高于阈值的要求的响应。判定形式：若本章的关系在明显高于任何检测阈值的要求上同样成立，则本章对而随机共振不足以解释；若关系只在微弱要求上出现，则本章实为随机共振的一个改写。其七 · 中性演化与中性网络。 这是最危险的一位上游占位者，因为剥到形式层，本章的形状是"无选择压时积累的变异，是后来创新的来源"，而这句话在演化生物学里已经被说了半个世纪。分离线：中性理论讲的是基因型空间的可达性——一个结构性事实，说的是"有路可走"；本章讲的是一个可事后清点的率与响应之间的关系，说的是"实际走了多少步"。判定形式：若一个可达性极高而实际漂变率接近零的群体（种群极小、世代极长）在环境变动时的适应与一个可达性一般而漂变活跃的群体相当，则本章错；若前者显著更差，则本章对而可达性读法不足。其八 · 简并性。 不同的结构实现同一功能，被认为是稳健性与可演化性的共同来源。分离线：简并讲的是结构层面的多重实现（一个静态属性）；本章讲的是同一个系统在无要求时实际做出不同做法的率（一个动态读数）。判定形式：若一个简并度很高而 v 为零的系统（多套备用方案齐全、而从来没有人在无故障期动过其中任何一套）在遇到新要求时反应良好，则本章错。 备用发电机长年不启动的机房是这一格的现成样本。其九 · 剩余寿命预测。 工程内部最接近的一位，第八节已交手。分离线：它通过退化特征推剩余寿命，前提是退化在传感器上留下痕迹；本章说的这一类能力流失在传感器上不留痕迹，因为要装的那台传感器必须在无要求时工作，而传感这个动作本身就是一次要求。判定形式：若在一组同型设备上，全部退化特征读数一致、而无故障期自发试验批占比不同的那些设备，其应对新工况的表现有系统性差异，则本章对而退化特征不足。其十 · 把安全定义为在变动条件下仍然成功的能力那一支。 这是本章必须最小心划界的一位，因为它与本章方向一致：它主张日常工作中的变异不是要消灭的威胁，而是成功与失败共同的来源。分离线在时机：它说的变异是做事时的调整——为了在资源不足、信息不全的条件下把活干完而做的近似处置，那是有要求时的变异；本章说的是无要求时的变异。两者可以独立取值。判定形式：若一个日常调整极其丰富（工作如实际做的远离工作如设想的）、而无要求时段完全静止的组织，在遇到一类全新要求时反应良好，则本章错而该支对；若这类组织恰恰在全新要求面前僵住，则本章对——它平时的丰富调整全部是围绕已知要求的。 这个判定在现场可做，而且两支都会认为它值得做。最近的一位是第一条。 它仍不能吸收本章，理由只有一条：它把 "变异要在任务中测"当作给定，因此看不见无任务时段的变异；一旦承认无任务时段的变异是一个独立变量，它整套实验设计的时间边界就要重划。

上述十种说法的共同盲区

逐个问一句"它把什么当作给定，因此看不见什么"：- 任务期变异说把"变异要在任务中测"当作给定，看不见无任务时段的变异；- 探索—利用把"探索要花预算"当作给定，看不见零

成本、无人决策的自发变异；- 必要变异度把"变异度是可能状态数"当作给定，看不见实现率；- 剩余寿命预测把"退化留下可测痕迹"当作给定，看不见不留痕迹的能力流失；- 中性理论把"变异发生在世代之间"当作给定，看不见世代之内；- 简并性把"多重实现是结构属性"当作给定，看不见从不被使用的多重实现等于没有。

交集是同一块地：一样东西的能力，只有在它被调用的时候才存在。所有人都站在这块地上，包括互相批判的各派——因为一切测量都是一次调用。这不是某一派的疏忽，是整个经验研究的立足处；也正因为共同，没有人能回头看它。本章能看见它，唯一的原因是物理学在自己的技术需要下（不施加外场而算出输运系数）早就写下了一条与它相反的关系式，而那条关系式在物理学内部回答的是另一个问题。

十三、与同一栏目里几篇的分界

学科通融专栏此前已有几篇触到相邻的轴，其中三篇处理的是同一个大问题域，必须指名。其一·《记为负项的那一类改变》。那一篇讲的是：产物改判据这一类事件被三家的账本记为负项（不公平、污染、漂移），因而发起处不换手。分离线：那一篇的对象是一个动作（有人做了一件事，而这件事被扣分）；本章的对象是一个时段里的率（没有人做任何被要求的事，而那段时间里发生的一切不被记录）。两者可以独立为零。判定形式：一个判据修改通道完全畅通、且实际被频繁使用（那一篇的读数很高）、而无要求时段完全静止（v＝零）的机构，若其应对全新要求仍然迟钝，则本章对而那一篇为空；反之，一个无要求时段极其活跃、而任何自发改动一律不被允许进入正式判据的机构，那一篇对而本章为空。两篇叠加时是相乘不是相加：没有静息变异，就没有东西可以拿去改判据；有东西而通道被堵，改动进不了下一轮。其二·《每一轮都从新鲜表面开始》与《先动的总是同一边》。那两篇的观测单位都是一门课程内部连续轮次之间（上一轮留下的东西被不被接上、验收节拍对准了哪一条驱动）；本章的观测单位是轮次与轮次之间那段没有轮次的时间。三者数学上独立，可以在同一批对象上同时测量。判定形式：若一门课的结转与节拍两项读数都很好、而假期与课间的 v 为零，其面对新形态学生的调整仍然缓慢，则本章对而那两篇为空。其三·《换一种说法，它就不见了》。那一篇讲的是同一份知识在不同使用条件之间的解绑代价。分离线：那一篇问的是一样已经学会的东西换个场合还在不在；本章问的是在没有任何场合要求它的时候，这个系统还产不产生新东西。两者方向相反：那一篇测的是保持，本章测的是产生。判定形式：解绑代价很低（换场合仍然会）而 v 为零的学习者，本章预测他对一类前所未有的问题仍然没有办法——这一预测那一篇不作，也无法作。其四·《一条路没有交接点》（同日学科通融专栏，与本章共用物理学与工程学两家）。那一篇问的是一条路上由谁来结算——判据由谁给出、判定

由谁执行；它的答案是交接点（位置更换、执行方更换、判据先于成果存在），读数是独立结算比。分离线有两层。其一在问法：那一篇问"怎么走、在哪一步能插手"，本章问"还能不能被改变、凭什么读得出来"；一个要的是次序上的一个位置，一个要的是一段时间里的一个率。其二在两家共用学科的取用点上，两篇没有一处重合：那一篇的工程学取写进标准的三项独立性，本章取余量与降额；那一篇的物理学取盲分析的时点（判据必须先于读数被锁定），本章取涨落与响应的关系式（响应在没有外场时就已确定）。同一门学科被两篇拿走的是两件不同的东西。判定形式：一门课的判据由完全独立的第三方给出并执行（那一篇的读数很高），而它的假期与课间静息变异为零，若其面对一类前所未有的要求仍然拿不出反应，则本章对而那一篇为空；反之，一个假期里不断自发折腾、而全部达标判定仍由授课方自己按自己的判据作出的教研组，那一篇对而本章为空。两个读数在同一批对象上可同时测量，且互不蕴含——独立结算解决的是"这一次达标算不算数"，静息变异解决的是"下一次没见过的东西来了拿不拿得出反应"。一个管已经发生的判定可不可信，一个管尚未发生的要求接不接得住。诚实记下一处继承：本章第八节"余量掩蔽"与《记为负项》第五节"失效不产生信号"是同一个机制在两根不同轴上的复现，那一笔记在那一篇账上。本章的增量在于指出这一次连传感器都装不上，而不只是没人去读。

十四、三家为什么各自到不了这一条

三家停在原地，不是因为疏忽。是因为各自的核心问题恰好使这一步成为不必要的。认知科学到不了，因为它的问题是"行为与神经活动如何由一条原则统一地解释"。这个问题要求有一个可在每一时刻定义的驱动量，而落差正是这样一个量。无要求时段里没有落差可定义，因此那段时间在它的框架里不是数据，是基线。 它不是不肯看基线，是它的整套形式化把基线定义成了"要被减掉的那个东西"。物理学到不了，因为它的问题是"如何从可测量算出可预测"。它写下了那条关系式，而且用得很熟——但它用的方向是从涨落算响应，为的是省掉施加外场的麻烦。它从没有反过来问一句：如果一个体系的涨落被人为压制了，它的响应会怎样。因为在物理学里没有人会去压制涨落——涨落是热运动给的，压制它意味着降温，而降温这件事的后果早已被完整地研究过，叫做冻结。物理学里"压低涨落"是一个已知会导致冻结的操作；只有在人造系统里，才会有人一边压低涨落一边指望它照常响应。工程学到不了，因为它的问题是"如何保证它不出事"。这个问题的形状要求把一切非计划的变动识别出来并消灭，而静息变异按定义就是非计划变动。它对这类事件的全部知识都是为了防止它发生——变更管制、配置管理、偏差处理，整套体系精确地指向同一个方向。它是三家里对这个量了解最深的，也正因为了解最深，它必须把它记在负栏。三条合起来说明一件事：这一条不是三家的综合，是三家各自的问题结构所排除的那一格。

十五、一处强制的不一致

上一节说三家各自到不了。这一节要说的更重一点：三家里有一家，在不承认本章这个量的情况下，它自己内部就不自洽。认知科学那一支同时持有两条，两条都是它的承重结构。甲条：一切认知改变都由预测误差（或其变分上界）驱动。这是这一支的全部力量所在——知觉、行动、注意、学习被统一在一条原则下，正因为只有一条驱动。乙条：为了不走进暗室，策略的价值必须包含认识论价值，即系统会主动去寻求能带来信息增益的状态。这一条不是外加的补丁，这一支明确指出它从形式里自然长出来，且只在对时间上延展的序列做推断时才出现。两条在通常情形下相安无事。它们在一种情形下正面相撞：当系统当下的落差已经为零时。按甲条，落差为零则无驱动，系统不应有任何改变的倾向。 按乙条，系统此时仍应去寻求信息增益——也就是去主动制造它现在没有的落差。这一支的化解是"长远看，寻求信息会降低未来的落差"。这个化解在形式上有效，代价是它要求系统能对尚未遇到的要求估值——即对"我还不知道我不知道什么"这件事给出一个数。而这个数只能在落差为零的时段被决定，也只能在那里被观测。于是这一支面临一个二选一：- 要么承认存在一个在无要求时段被决定、且不由当下落差定义的量。 一旦承认，甲条的"落差是唯一驱动"就不能再以现在的形式成立——因为有一个量既不是落差，也不由落差决定，而它决定系统在落差为零时做什么。 - 要么放弃乙条，接受系统确实会走进暗室。而这一支自己已经明确拒绝过这条路。

这不是本章替它选，是本章指出它必须选，而它现在没有选。 它现在的处置是把"寻求信息"这件事写成一个也叫"最小化"的东西（最小化期望的落差），从而在形式上保住甲条。这个处置在数学上无懈可击，代价是"落差"这个词在甲条与乙条里指的已经不是同一样东西了：甲条里它是可在当下观测的量，乙条里它是一个对未曾发生之事的估计。同一个词承担两个不能同时观测的角色，这正是本章的量被藏起来的地方。本章已给自己留了退路：这一处可以被"咬牙接受落差包含反事实落差"化解——那样它不是不一致，只是让"落差"变成一个不可由当下观测定义的量。这条退路被写进第十八节的证伪条件第九条。同一个动作可以对工程学再走一遍（这一次力度弱一些，因此不作为承重）：工程学同时持有"变更是风险的主要来源，应当最小化"与"僵化的系统在环境变动时最先失效"。它现在的处置同样是层级隔离——把前者归"运行"，把后者归"战略"，由不同的部门负责，因而永远不必在同一张表上同时优化。

十六、三家反过来给本章加的三条约束

碰上有用的对手，代价是它们也会削掉你原本想说的话。三处，逐条写清削掉了什么。其一 · 物理学削掉了本章的价值排序，而且削得最狠。 本章原本会写下一句更强的话：静息变异率越高越好，提高它就能恢复可改变性。 物理学把这句话取消了。那条把涨落与响应连起来的

关系式，在被困住的、缓慢老化的体系里破缺：涨落还在，而它不再预示响应；人们只好引入一个"有效温度"来描述破缺的程度。[4] 落到本章的对象上，这意味着：一个已经被困住的组织里，静息变异与应变能力解耦——那里的自发变动仍在发生，却不再通向任何改变。 此时提高 v 只是加噪。于是本章的主张从"应当提高 v"收窄为：v 是一个诊断读数，不是一个操作杆。先判断这个系统在不在解耦区；不在，v 读得出能力；在，v 什么也读不出，此时要做的是别的事。 判断在不在解耦区的现成办法本章给不出，只能给一个粗筛：看这个系统最近一次真正改变是什么时候，以及那次改变是不是由内部发起的。 这一处动的是本章的价值排序，不是补一个条件。其二 · 工程学削掉了本章的适用范围。 规程的稳定性本身是安全资产，变更管制不是官僚主义，是几十年事故换来的。一个 v 很高的核电运行班组在法律上、伦理上都是不可接受的。于是本章的适用范围收窄为"需要在未来适应未知要求的系统"，并且必须补一条实操约束：这类系统提高 v 的唯一正当途径是设立与生产完全隔离的场合——模拟机、试验线、沙箱、影子系统——而这恰恰是可靠性工程已经在做的事。本章因此不是在反对变更管制，是在指出变更管制必须配一个隔离的、无要求的、有记录的场合，否则它把两样东西一起消灭了。其三 · 认知科学削掉了本章读数的粗糙。 运动学习的实验已经证明：变异必须分成与任务相关和与任务无关两部分，只有前者预示学习，后者是纯粹的噪声。这直接推翻了本章最省事的算法——把无要求时段里所有的"不一样"都数进去。v 必须分层：只计入那些落在与本系统职能相关的维度上的变异。 一个团队在无工单期把办公室重新装修了一遍，这不进 v。这一处是真削弱：它让 v 的清点成本上升了一个量级，也让"数据已经在版本库里"这条便宜话打了折——版本库里的提交需要按相关性再分一次类。

十七、不适用的边界

四处明写：其一 · 没有可指认默认做法的系统读不出 v。 若一个系统从来没有稳定的做法，"与默认不同"就无从定义。初创团队、临时编组、正在剧变中的组织，本章的读数在这些地方无效。其二 · 没有真正无要求时段的系统读不出 v。 高强度、无间断、始终有截止期的岗位（急诊、战时、连续生产的关键工位）在结构上不存在无要求时段。这不是它们的缺陷，是它们的性质；对它们，本章能说的只有一句：这类岗位的可改变性必须由岗位之外的安排提供，不可能由岗位自己长出来。其三 · 安全关键系统的合理 v 值低，且低是对的。 见第十一节第三条与第十六节第二条。其四 · 高 v 不等于好。 高 v 可能读出的是失控、无标准、或纯粹的浪费。v 是一个方向读数，不是一个质量分数；单独看它会得出荒谬的结论。要判断一个系统的处境，必须同时看第七节那两根轴。

十八、证伪条件

先写出最强的竞争解释，那句最朴素、最难反驳的话："其实不就是练得少了吗？不就是投入不足吗？"若真是如此，那么决定应变能力的应当是总投入时间与总练习量，而不是无要求时段的自发率；把投入加上去，能力就该回来。第九节那份座舱实测已经给了这条解释一次打击（手上那部分同样用得少却保持得好），但那只是一个案例，不足以排除它。第二强的竞争解释是"其实不就是探索不足吗"，第十二节第二条给了它一条分离线。核心证伪条件写成这个形状：> 控制总投入时间、总任务量与探索预算之后，若无要求时段的自发变异率 v 与该系统对一类全新要求的响应速度之间不存在剂量—反应关系，则本章的第一层主张为伪；届时应归因于总投入而非无要求时段的变异。样本纪律写死：同一机构、同一工种的 N≥8 个班组或团队，在人员构成、任务量、技术条件上同质，自变量（无要求时段的记录）有真实变异。不得拿两个国家或两个行业作对比充数——那种比较两边处处不同，所谓"控制掉某变量"在方法上站不住，只能作启发式说明。另有八条，其中第三条最便宜，所需数据在多数机构里已经存在：1. 顺序检验：v 的下降应当早于应变能力的下降出现，且早若干个周期。若两者同时出现或 v 滞后，本章作为一个早期读数的全部价值即告消失。 2. 分层检验：与职能相关的那部分变异应当预测响应，与职能无关的那部分不应当。若两者预测力相当，则第十六节第三条的分层是多余的，v 的定义须重写。 3. 最便宜的一条（第三层主张）：在同一组织内的一组同类团队中，若无工单期的自发提交占比与该团队应对一次结构性新要求的用时无关联，则第三层主张为伪。所需资料：版本库的一条查询 + 一次事件复盘记录。 4. 座舱那一处的直接检验：在巡航段自发做非要求认知活动的飞行员，其"用脑"技能的保持应显著优于不做的人，而两组的"用手"技能不应有差别。若两组在两类技能上都无差别，第九节的解释为伪。 5. 要求掩蔽与余量掩蔽的分离：在一组同型设备上，若全部退化特征读数一致而无故障期自发试验批占比不同的那些设备，其应对新工况的表现无系统性差异，则第八节的区分不成立。 6. 无奖惩条件的必要性：若带奖惩的"自由探索时间"与无奖惩的自发变异在预测响应上没有差别，则第五节第四条判定条件多余，本章的反噬预言同时失效——那将是一个好消息。 7. 形状检验：响应与 v 的关系若呈现单峰（有一个最优值，两侧都差），则本章被临界性一支吸收；若单调，则本章对。8. 解耦区的存在：若找不到任何一个 v 高而响应低的真实系统，则第十六节第一条的让步是多余的——那反而使本章更强，但也说明本章对物理学那条限制的搬用是错的。 9. 强制不一致的退路：若认知科学那一支公开承认"落差"包含不可由当下观测定义的反事实落差，则第十五节所指的不一致不成立——那时它只是变弱，不是自相矛盾，本章对它的那一处主张作废。若本章被证伪，我们会学到什么：若 v 与应变能力确实无关，那么"能力在无人要求时被决定"这个说法就是错的，全部注意力应当回到投入与配置这一层——而这将是一个正面的结论，因为它会把大量资源从"给系统留白"这类结构改造上收回来，投向更直接的地方。

十九、反噬预言

本章一旦被制度采用，最省事的实现方式是：每周划出两小时"自由探索时间"，要求各团队在这段时间里做与日常不同的事，并提交记录。那一刻，本章的定义前提被销毁。被划出来的两小时不是无要求时段——它有起止、有要求、有交付物、有考核。 在这两小时里测到的一切，按本章自己的定义，一条也不算数。而机构会得到一份很好看的报表，上面写着 v 从零升到了百分之四十。比这更省事的一种是：把日常工作中本来就存在的偏差重新分类，全部记成"自发变异"。这一步连两小时都不必划，改一改字段名即可。唯一能想到的减缓办法是：只做匿名的、无奖惩的、事后的清点，且清点结果不进入任何个人或团队的档案，只作为机构层面的一个诊断数。但这个办法自己也有出口——一个不进任何档案的数没有人会去认真采集；而一旦有人认真采集，被采集者就会知道有人在看，那一刻它就又成了一次要求。我看不到出口。 这不是一个待办事项，是本章交出去的一个洞。它的结构是：任何使 v 可核查的装置，都构成一次要求；而不可核查的 v 不能进入任何决策。 本章能给的底线只有一条：验收必须看无要求时段里已经发生过的事的事后记录，不看是否设立了"自由探索"制度。有制度而事后清点为零，与没有制度是同一格。

二十、自反：本章自己在哪一格

本章提出的判据，对本章自己适用吗？适用，而且结果不好看。本章是在有要求的条件下写出来的：有一条明确的指令，有一个栏目，有一个交稿的时刻，有三门指定的学科。按第五节的五条判定条件，本章写作期间的 v 是零——它的每一个动作都在响应一个要求。后果具体有三条：第一，本章属于第七节那张表里的乙格，而不是甲格。 它的现行表现（论证、划界、兑现处数）可以做得很好，而这一切都是被要求的产物。本章因此只能诊断，不能示范：一篇在要求下写成的东西，说不出无要求时会写出什么。第二，本章最可能的命运正是它自己的反噬。 一个提出"要留出无要求时段"的文本，最容易被做成的东西就是一项制度——每周两小时，提交记录。本章预言了这件事，而预言它并不能阻止它，因为阻止它需要的正是本章没有的那个东西：一个不被要求的场合来消化它。第三，本章对它自己的这个判断不能靠讨论来修正。 讨论会产生引用，引用不会改动本章任何一行。本章能给自己安排的唯一一次真正的变动，是下一节那条写死日期的赌注——它把一个可能的负项固定在一个不可撤销的时间点上。最后一处更难看的：本章这套读数一旦被采用，第一批被读出 v＝零的对象就是学术论文与它的生产制度——选题响应基金指南，写作响应栏目要求，发表响应审稿意见，每一步都有一个人或一个程序在等着收东西。本章没有理由认为自己例外。

二十一、结论，与一条写死日期的赌注

一个系统还能不能被改变，读不出来的原因不是仪器不够精，是全部仪器都装在有要求的时候。而决定这件事的那个量，只在没有人要求它的时候存在。三门学科给了三条互不相容的读法——看误差、看响应、看余量——而三条共同假定：能力只在被调用时才存在。物理学自己的一条关系式说，那个只在不被调用时存在的量，与被调用时能拿出的反应，是同一个量。把这个量命名为静息变异，把它的比率记作 v，就得到一张两轴四格。要打的是"表现合格而 v 为零"那一格：它在全部现行指标上最好看，它的失效不产生任何信号，而且它越正规化越严重。它不是失职的产物，是尽职的产物。而这条判断的根，比学习更深一层：可靠性要求把变异消灭，可改变性要求把变异留住。 二者要求相反方向的处置，而且用的是同一笔预算——同一套变更管制、同一套考核、同一段时间。一个系统被同时要求做到两件事，而它一次只能做到其中一件。这不是一个可以靠更好的设计消解的张力，只能靠在空间上把两者分开：给可靠性一套受管的场合，给可改变性一套隔离的、无要求的、事后有记录的场合，并承认后者的产出在当期一定看起来像浪费。赌注，写死日期：> 到二〇三二年十二月三十一日为止，至少有三家高可靠性行业的机构（航空、电力、医疗、制造任一）在其能力维持标准中，把"无任务时段的自发活动"设为一项独立于培训学时与事件率的指标，并公开其年度读数。写死什么不算命中：1. 仅设立"创新时间""自由探索时间"一类有起止、有交付、有考核的制度安排； 2. 仅在质量报告或社会责任报告中自述"鼓励员工创新"而无可核对的事后清点； 3. 仅统计培训学时、演练次数或提案数量； 4. 与其他改革打包实施，无法分离； 5. 只公布制度存在与否，不公布事后清点的实际读数； 6. 设立了场合而年度实际清点为零——有制度而零发生，与没有制度是同一格。最后交出两个本章解释不了的洞。其一：第五节第五条要求静息变异必须有记录才计入，而第十节第二处的住院医师表明，大量真实的静息变异恰恰发生在无人记录的地方；本章的读数在那里读出零，而那个零是错的。本章没有办法修正它——任何补救措施都是去建立记录，而建立记录就是建立要求。其二：本章说清楚了在什么时候测，没有说清楚为什么有的系统在无人要求时仍然在动，而有的不动。第十六节第一条的"解耦区"给了一个物理学上的形状，但那个形状怎么在一个组织里被识别出来，本章给不出办法，只给了一句粗筛。这是本章最大的空缺，也是它眼下最不该被拿去做决定的原因。

注释

[1] 暗室问题是对这一支被提出得最多的一条反驳：若最小化预测误差是唯一驱动，最优策略应是躲进一间毫无刺激的暗室。见 Friston, Thornton & Clark（2012）与 Sun & Firestone（2020）。[2] 见 Seth 等（2020）对上一条的答复：期望自由能可分解为工具性项

与认识论项，后者从形式中自然出现而非外加，且只在对时间上延展的序列作推断时才出现。[3] 涨落与响应之间这条关系的经典表述见 Kubo（1966）；用平衡涨落的时间关联算输运系数的一族公式通称 Green–Kubo 关系。[4] 关于该关系在玻璃态与老化体系中的破缺及"有效温度"，综述见 Crisanti & Ritort（2003）。[5] 反面证据见 Jabbari-Farouji 等（2008）：在硬球胶体玻璃与 Laponite 凝胶中，跨若干频率量级未观察到破缺。两方的争论至今未决；本章取对自己不利的一支作约束。[6] Casner, Geven, Recker & Schooler（2014）。十六名航线飞行员在波音七四七模拟机上飞常规与非常规科目，系统变化自动化水平。结论是仪表扫视与操纵控制这类技能保持良好，而与手动飞行相关的认知技能的保持取决于飞行员是否仍在主动监督自动化。研究者的公开表述是：对飞行员"用手做"的事可以少担心一点，对他们"用脑做"的事应当多担心一点。

第十章

选择听谁，本身就是一次判断

论鉴别资质的独立性，及证言依赖在何处失效

学科入口　美学 × 伦理学 × 知识论

摘要

一般证言知识论把默认接受他人之言视为理性的常态，而美学与伦理学中各有一条长期存在的禁令与之对立：审美判断须基于亲身经验（熟识原则），道德信念不应仅凭他人告知而形成（道德证言悲观论）。这两条禁令近年同时遭到强攻——审美一侧出现了第一部系统的乐观论专著，并有学者称熟识原则今日已近乎被公认失败；知识论一侧则出现了更彻底的主张，认为「认知自主」根本不是一种德性，应当被「智性互赖」取代。三家的冲突不是侧重之别：若知识论一侧正确，则另两条禁令是把一个误导性理想抬成了规范；若伦理一侧正确，则互赖的处方正在系统性地生产不再寻求理解的人；若美学乐观论一侧正确，则前两者都把一条社会性的表演规范误诊成了认识论规范。本章提出，三家共同未加追问的，是鉴别资质与执行资质之间的关系。本章将某领域中「判断谁在此事上可靠」的能力称为鉴别资质，将「自己做出该领域判断」的能力称为执行资质，并主张：一个领域能否安全地依赖证言，取决于该领域的鉴别资质是否独立于执行资质。独立者，外行可凭外部凭据识别专家，依赖因而既理性又安全；不独立者，除以自己的同类判断衡量对方之外别无他法，于是选择听谁本身就是该判断的一次完整行使，依赖并未省去判断，只是把它移到了一个不被记录的位置。本章进一步论证该独立性的成因：外部凭据之所以在事实领域可用，是因为那里存在不预设该领域判断的结果反馈，且此类反馈的读取不需要专业能力；而在审美与道德领域不存在这样的反馈，凭据只能由同类判断者互授，故无法被外行锚定。由此得到四项推论：智性互赖的处方在不独立领域是循环的；依赖在此类领域会自我锁死（不是动机下降而是能力下降）；聚合多人意见不构成解决；以及——审美乐观论所举的描述性证据，全部落在「无下游使用」的一侧，因而不是反例而是本判据所预测的正常情形。文末给出与五个既有说明的辨别、九个领域的兑现、人工智能条件下的新处境、六项反驳的回应与六条可证伪条件。

关键词

鉴别资质；执行资质；独立性判据；证言依赖；熟识原则；道德证言悲观论；智性互赖；专家识别

一、问题：一条禁令，三种互不相容的处理

1.1 一个不对称

在绝大多数事情上，我们靠别人告诉我们而知道。地球的年龄、某种药的剂量、去年的通货膨胀率、这座桥能承多重——这些信念的保证几乎全部来自他人，而没有人认为这有什么不妥。当代知识论主流的判断是：证言与知觉、记忆并列，是一种基本的知识来源；默认接受一个看起来可靠的证言者，是理性的常态而非退让。可是有两个领域一直被当作例外。一个是审美。按照沃尔海姆一九八〇年提出、此后被反复援引的熟识原则，审美价值的判断必须基于对其对象的第一手经验，除极窄的限度外不可在人与人之间传递。你不能仅凭朋友说「那幅画很好」就形成「那幅画很好」这个判断——你可以相信她说的是真话，可以据此决定去看，但你不能因此就把那个判断算作你自己的。另一个是道德。霍普金斯于二〇〇七年把这一立场命名为道德证言悲观论：仅凭他人告知而形成道德信念，在某种意义上是有缺陷的。希尔斯给出的最有影响的解释是：证言可以传递足以构成道德知识的保证，却传递不了道德理解——那种把握「为什么如此」的状态；而道德理解是若干道德成就（如有德性地行动、行动具有道德价值）的必要条件。于是出现一个需要解释的不对称：为什么偏偏这两个领域不能依赖证言？

1.2 三家的最新进展，与它们的正面冲突

这一不对称在最近数年里同时受到三个方向的推动，而三个方向互不相容。其一，知识论一侧不但为依赖辩护，而且否认「自主」是一种德性。 在关于认知自主的专题讨论中，列维主张「智性自主」是对某种执行管理德性的误导性命名：鉴于我们的社会处境与深度认知依赖，更合适的名称是智性互赖——一种在延迟他人与自己推理之间恰当权衡的德性。他进一步表示，一旦把互赖理解妥当，就完全不需要再设一个自主的德性。与之对峙的格林则主张保留自主，把它理解为在「与他人一同思考最易被带偏」的那些场合鼓励独立思考的德性；卡沃尔随后加入，三方在二〇二五年形成了一场公开往还。这一支还带着经验证据：关于「自己做研究」的讨论普遍指出，坚持自行判断在若干领域可靠地把人带离真理。其二，伦理一侧的悲观论给出了一个更硬的动态机制。 传统悲观论是静态的——延迟于证言时你缺少理解。而近年的推进是动态的：以证言了结一个问题，会使人此后更少有理由去寻求更广的理解，从而使他持续地处于缺乏理解的状态。与此同时，关于道德理解本身的研究把它进一步系于道德领会以及

主体在情感与动机上的投入。这两步合起来，使悲观论从「你少了一样东西」变成「你少了那样东西，并且因此更不容易再得到它」。其三，美学一侧的乐观论正在把那条禁令整个拆掉。罗布森于二〇二二年出版了该题目上的第一部专著，系统地为审美证言辩护。其中最要紧的一步是方法论上的：他要求悲观论者交出一个说明——是什么使审美（以及味觉）这两个领域独具这种熟识要求；找不到这样的说明，那条禁令就只是一组直觉。他还给出大量描述性证据：我们在失传作品、自然景观之美、经时间检验的审美共识等场合大量依赖审美证言，且这些依赖看起来完全正当。里格尔在一次讲演中说，熟识原则曾经根深蒂固，而今日几乎已成共识认为它失败了。在解释那条禁令的来源时，罗布森早期提出的一条路径是信号规范：作出审美断言即在标示自己的智力、技艺与感受力，凭二手证言而断言之错在于虚报自己。阮氏对此提出了一个尖锐的反例——在完全私下、不向任何人标示的场合，那种「不对」的直觉反而更强。

1.3 冲突的形态

把三者并置，冲突的烈度可以精确陈述。

学科	对「能不能靠别人告诉你」的判断	若它成立，另两家如何
知识论 · 反自主一支	能，而且坚持自己想反而更糟；不需要「自主」这个德性	另两条禁令是把误导性理想抬成规范
伦理 · 悲观论新版	能得到信念，得不到理解，且此后更不会去找	互赖的处方在系统性地生产不再求解的人
美学 · 乐观论	能；那条禁令的来源不是认识论而是社会表演	前两家把一条表演规范误诊成了认识论规范

三方各自有可核查的支撑：知识论一侧有社会认知的大量研究与「自己做研究」的经验代价；伦理一侧有一条可陈述的动态机制；美学一侧有一部专著、一批描述性反例，以及一个方法论上的强要求。因此问题不在于谁错，而在于三者共同没有追问什么。

1.4 本章的主张

本章提出，三者共同未加追问的，是一个比「能不能传递」更靠前的问题：> 在这个领域里，你凭什么认为那个人在这件事上靠得住？而这个「凭什么」，是否可以被一个不具备该领域判断力的人使用？本章将「判断谁在此事上可靠」的能力称为鉴别资质，将「自己做出该领域的判断」的能力称为执行资质，并主张一条分界：> 一个领域能否安全地依赖证言，取决于该领域的鉴别资质是否独立于执行资质。在独立的领域，外行可凭外部凭据识别可靠者，依赖既理性又安全；在不独立的领域，除了用自己的同类判断去衡量对方，别无他法——于是选择

听谁，本身已经是那个判断的一次完整行使。依赖在这里并没有省去判断，它只是把判断移到了一个不被记录、也不被自己察觉的位置。第二节完成不可还原性校验，第三节界定概念并给出独立性的成因，第四节给出可操作判据，第五节给出自我锁死机制，第六节以此重新解释三家的证据，第七节在九个领域兑现，第八节处理人工智能条件下的新处境，第九节回应六项反驳，第十节给出边界、局限与可证伪条件。

二、不可还原性校验：与五个既有说明的正面辨别

一个新判据若能被既有说明一比一替换，它就没有提供增益。本节与五个最强的邻近说明逐一交锋。

2.1 与希尔斯的「理解不可传」

这是距离最近的一个，也是本章最需要说清的一处。希尔斯的说明处理的是接收端：证言给了你信念与保证，没有给你「为什么」的把握，而后者是若干道德成就的必要条件。本章处理的是上游：在你接收之前，你凭什么选中这一位证言者。两处差别是实质性的。第一，解释范围不同。理解不可传这一条若成立，它同样适用于大量事实领域——医生的诊断迁移能力、工程师对新结构的判断，都要求理解而非仅仅知识，而我们并不因此认为依赖医生与工程师是有缺陷的。希尔斯的说明因此过度概括，这也是文献中对它的一项标准批评：它蕴含许多直觉上无可指摘的信念形成方式（如溯因、归谬、直觉）同样令人不安。本章的判据不遇到这个问题：医学与工程属于鉴别可外部化的领域，故依赖在那里是安全的，尽管那里同样要求理解。第二，给出的对策不同。若问题在接收端缺理解，对策是补理解（去学、去想、去要求解释）。若问题在上游的鉴别，补理解不解决问题——因为你仍然要先选中一位值得去理解其理由的人，而那次选择已经用掉了你要外包的那个能力。

2.2 与阮氏的自主原则

自主原则主张：人应当通过运用自己的官能与能力来得出自己的审美判断。它是一条规范性主张——你应当如此。本章的主张是结构性的：在鉴别不独立的领域，你没有别的办法，这不是应当不应当的问题。这一差别有两个后果。其一，本章不需要为「自主为何有价值」作辩护，因而不受「自主是不是一个好理想」这场争论的牵连——即便列维一侧完全正确、自主不是德性，本章的分界依然成立，因为它讲的不是德性而是可行性。其二，本章能够解释自主原则为何在事实领域不适用，而无须诉诸「那里的价值不同」这类说法：那里的鉴别可以外部化，所以运用自己的官能不是唯一通路。

2.3 与列维的智性互赖

本章不反对互赖。深度依赖确是常态，聪明地延迟确是一种可欲的能力，「自己做研究」确有可观的经验代价。本章提供的是它的适用条件。互赖的核心动作是「延迟得聪明」——对证言者在该领域的胜任度、善意、以及内外群体中支持与反对的比例保持敏感。而这一动作在鉴别不独立的领域是循环的：要判断某人在审美或道德上是否胜任，除了用你自己的审美或道德判断去衡量他，没有第二条路。于是互赖的处方在这些领域要求的，正是它试图替你省下的那样东西。需要强调：这不是对互赖的反驳，而是一条边界。互赖在鉴别可外部化的领域是完全正确的，而那是绝大多数领域。

2.4 与麦格拉思的道德专家难题

麦格拉思指出，我们无法在不预设自己的道德判断的情况下识别道德专家，并以此构成对道德实在论的一个难题：若存在道德事实，为何在这一领域不存在我们能够识别并加以依赖的专家？这是本章最近的一位盟友，也因此最需要划清。本章的增量在两处。其一，它不是道德的特殊难题。麦格拉思在道德领域观察到的结构，同样出现在审美领域，并且——如第七节所示——出现在一整类领域中。把它当作道德的特异现象，会误导人去它的本体论中寻找原因。其二，本章给出了该结构的成因（见 3.3），而成因是关于验收结构的，不是关于对象本性的。这一点削弱了它对道德实在论的威胁：无法识别专家，未必说明没有事实，可能只说明没有不预设该判断的结果反馈。

2.5 与戈德曼的「外行如何评估专家」

戈德曼一九九一年的经典处理，正是试图把鉴别外部化。他列出外行可用的若干判据：论证的表现、其他专家的一致同意、专业履历、利益偏倚、以及过往的成功记录。本章对它的一击是精确的：这几条判据中，只有最后一条能起锚定作用，而它恰恰在不独立领域不存在。- 论证的表现：外行只能评估论证的表面质量（流畅、自信、看似有条理），而这与实质可靠性可以脱钩，这一点在本领域内早有讨论。- 其他专家的同意：把鉴别问题推给了「谁算专家」，构成回退一步的同一问题。- 专业履历：其效力最终取决于颁发履历的机构是否可靠，而这又需要锚定。- 利益偏倚：能剔除若干明显不可靠者，不能在剩下的人中排序。- 过往的成功记录：唯一可以由外行直接读取而不预设专业能力的项——病人活没活、桥塌没塌、预测中没中。

于是本章的判据可以被表述为对戈德曼名单的一次收束：鉴别的独立性，等价于第五条判据是否可用。

2.6 与群体智慧／意见聚合

一条常见的实践建议是：不确定时多听几个人的，取其中位或多数。这一建议的成立要求两个条件：个体判断相互独立，以及存在一条被认可的聚合规则。在不独立领域，第二个条件本身就是问题的重演——选择哪一条聚合规则、把谁算进这个池子，同样需要鉴别。把十个人的审美判断平均，与听其中一人的，之间的差别取决于这十个人是怎么被选进来的；而选人的那一步，正是本章所说的那次判断。

2.7 校验结论

五个说明中，两个（希尔斯、阮氏）处理的是接收端或规范面而非上游的可行性；一个（列维）与本章相容而缺一条边界；一个（麦格拉思）观察到了同一结构但把它当作道德的特异现象且未给成因；一个（戈德曼）正是要外部化鉴别，而本章指出其名单中唯一有锚定力的一项在此类领域缺席。没有任何一个能覆盖「鉴别资质是否独立于执行资质」这一辨别面。

三、概念界定与独立性的成因

3.1 两种资质

执行资质：在某领域自己作出该领域判断的能力。会看病、会算结构、会判断一幅画好不好、会判断一件事该不该做。鉴别资质：判断某个他人在该领域是否可靠的能力。二者在概念上分离，这一点常被忽略。一个完全不懂心脏外科的人，可以相当可靠地判断某位心脏外科医生是否值得信任——他凭执照、专科认证、所在机构、同行转诊、并发症率。他的鉴别资质高，执行资质为零。这种分离不是偶然的便利，而是分工得以可能的条件。若鉴别必须以执行为前提，则一切专业分工都将无从建立：你要找一位可靠的电工，就得先会电工。

3.2 独立与不独立

据此可作一条分界。鉴别独立的领域：存在一批外部凭据，使不具备执行资质者也能可靠地排序候选证言者。医学、工程、航空、税法、气象预报属于此类。鉴别不独立的领域：不存在这样的凭据；判断某人在此事上是否可靠，只能以自己的同类判断去衡量他。审美判断、道德判断属于此类，并且——第七节将表明——不止于此。在不独立的领域，「依赖证言」这一描述是误导的。你并没有把判断交出去，你做了一次判断：你判断了这个人在这件事上值得听。而这次判断动用的正是你在该领域的执行资质。因此：> 在鉴别不独立的领域，选择听谁本身就是那个判断的一次完整行使。依赖没有省去判断，只是把它移到了一个不被记录的位置。

3.3 独立性从何而来

这是本章的关键一步，也是对罗布森那个方法论要求的正面回答——他要一个说明，说明是什么使某些领域独具熟识要求。本章的回答是：> 外部凭据之所以在某些领域可用，是因为那里存在不预设该领域判断的结果反馈，且此类反馈的读取不需要该领域的执行资质。拆开来看有两个条件，缺一不可。条件一：存在独立于该判断的结果。 一台手术之后病人活没活，这件事的成立不依赖于任何人对该手术的医学评价；一座桥塌没塌，不依赖于任何人对该设计的工程学评价。条件二：该结果的读取不需要专业能力。 病人是否活着、桥是否塌了、飞机是否落地、预报的雨是否下了——这些外行都读得出。正是这一条使凭据可以被外行锚定：执照制度、专科认证、机构声誉，其可靠性最终由这些外行也能读取的结果所校准。现在看审美与道德。审美：一件作品「好不好」没有独立于审美判断的验收。销量、奖项、经时间的存留都是同类判断的聚合，而不是外部结果——它们由具备（或自称具备）审美判断的人产出。因此条件一不成立。道德：一个决定「对不对」同样没有独立于道德判断的验收。后果可以被观察（有人受伤、有人得益），但从后果到「这个决定是对的」这一步，本身就是一次道德判断。因此条件一同样不成立。于是熟识原则与道德证言悲观论所捕捉的那件事，可以被推导出来，而不必被当作一组原始直觉：在这两个领域中，鉴别不可外部化；而在鉴别不可外部化的领域，「靠别人告诉你」这一操作在结构上不能省去你自己的判断。

3.4 三点澄清

其一，这不是怀疑论。 本章不主张在这些领域里无法获得知识，也不主张他人的意见没有价值。他人的审美与道德意见极有价值——但其价值的兑现方式不是替代你的判断，而是给你一个去看、去想的方向。这一点与本章所属栏目此前几篇的立场一致，此处不展开。其二，独立性是程度问题。 少数领域完全独立（结果硬碰硬且外行可读），少数完全不独立，多数居中。第七节给出一条居中的谱系。其三，不独立不等于不可学。 执行资质可以增长，而它增长的同时鉴别资质随之增长——因为在这类领域两者是同一样东西。这正是第五节自我锁死机制的另一面：它同样意味着唯一的出路是提高执行资质，而这条出路是通的，只是不能被外包。

四、判据及其操作化

4.1 一问

把第三节收成一个可以当场提出的问题：> 我凭什么认为这个人在这件事上靠得住——而这个理由，一个完全不具备该领域判断力的人能不能用？能用 → 鉴别独立，依赖安全。 不能用（那个理由本身要求你已经会判断这件事） → 鉴别不独立，你正在做的不是外包判断，而是行使判断。

4.2 三层答案

实际使用时，答案通常落在三层之一，而三层的处置完全不同。第一层：可锚定的外部凭据。「他有执照，那家医院的并发症率公开可查，出了事有人被吊照。」——这个理由完全不依赖我懂不懂医。鉴别独立。第二层：同类判断的聚合。「他在这个圈子里公认有品味」「大家都说他判断力好」。——这个理由把鉴别推给了一群人，而那群人是谁、他们凭什么，仍然需要我判断。鉴别不独立，尽管它看起来像第一层。这一层最危险，因为它在形式上模仿了第一层。第三层：我自己的同类判断。「我读过他写的东西，他说的那些我自己核对过，对得上。」——这个理由诚实地承认了它动用了我的执行资质。鉴别不独立，但这是正确的形态：它至少没有伪装。第二层与第三层的分别值得强调：两者都不独立，但第三层是在行使判断，第二层是在误以为自己没有在行使判断。后者的代价见第五节。

4.3 一个更快的粗筛

三层要一个个想，有一个更快的办法：> 这个领域里，有没有人因为判断错了而承担过可被外人读出的后果？有——且这类事件在该领域是常规的、可查的 → 大概率独立。 没有，或者「判断错了」这件事本身只能由同行认定 → 不独立。这条粗筛之所以比一问更快，是因为它不问你的理由，只问已经发生过什么。而人对自己所处领域的独立性判断，几乎总比实际情况乐观一格。

4.5 三个实例走一遍

判据的价值全在能否当场用。此处以三个日常场合各走一遍。场合一：选一位牙医。「我凭什么认为她靠得住？」——她有执照，这家诊所有可查的投诉记录，做坏了会被追责，而且我自己的牙好没好，几个月内我就知道。这几条理由，一个完全不懂口腔医学的人全都能用。鉴别独立，放心依赖。场合二：选一位理财顾问。「我凭什么？」——他有资格证，他所在的机构有名。但这两条的锚在哪里？他的过往业绩要放在多长的周期上读、如何与市场基准分离、他所承担的风险有多大——这些的读取本身需要专业能力。条件一成立而条件二不成立，因而其独立性远低于表面。这解释了一个长期存在的经验事实：该领域凭据齐全而事故不断，且事故往往在数年后才显形。正确的处置不是放弃依赖，而是承认自己的鉴别在这里是弱的，并据此限制单次委托的规模——这是本章框架给出的、与「找一个更权威的机构」方向不同的建议。场合三：听一位前辈说「你该往那个方向做」。「我凭什么认为他在这件事上靠得住？」——他做出过成绩。可他做出成绩的是另一个时代的另一个方向，而这条建议要在十年后才被检验，届时检验它的人是他这一代的同行。我能给出的唯一诚实理由是：我读过他判断过的一些事，其中有几件我自己后来也想明白了，对得上。这条理由不懂行的人用不了。鉴别

不独立——而这意味着，我采纳这条建议这一步，已经是我自己的一次方向判断，不是一次省力。三个场合的形式完全相同（找一位更懂的人问），而按判据分属三格。这正是本章认为该判据有用的理由：形式相同的动作，其性质并不相同，而形式是我们默认用来分类的东西。

4.4 两个必须防的误用

其一，不要用它去否定专业分工。本章的判据在绝大多数领域给出的答案是「独立」，即依赖安全。它的适用面是窄的，不是宽的。其二，不要用它为拒绝倾听作辩护。判据说的是「你选择听谁这一步动用了你自己的判断」，不是「不要听」。恰恰相反：既然那一步已经动用了你的判断，那么听得越多、看得越具体，那一步就做得越好。判据反对的是以为自己没在判断，不是判断本身。

五、机制：不独立领域中的自我锁死

5.1 与伦理一侧动态机制的关系

第一节提到，悲观论近年的推进是动态的：以证言了结一个问题，会使人此后更少有理由去寻求更广的理解。那是一条动机层面的机制。本章给出的是一条能力层面的机制，两者独立且叠加。

5.2 三步

第一步。在不独立领域，我依赖某人的判断。这一次依赖是有代价的，但代价不在当期显现——我得到了一个可用的结论。第二步。由于在这类领域鉴别资质与执行资质是同一样东西，我每一次以「听他的」代替「自己看」，就少了一次运用该资质的机会。执行资质随之停滞，鉴别资质同步停滞。第三步。鉴别资质停滞，使我下一次更难判断该听谁——而这恰恰增加了我依赖的必要，因为我更没有把握自己看。于是依赖加深。三步构成一个闭环，且每一步都局部合理：第一步是明智的省力，第二步不产生任何可见的损失，第三步是对自己能力不足的诚实反应。

5.3 为什么它不被察觉

这一机制的隐蔽性有一个具体的来源，它同时解释了 4.2 第二层为何最危险。在依赖加深的过程中，我的判断的外在表现在改善而不是恶化：我引用的是更可靠的来源，我的意见与公认的意见更加一致，我犯明显错误的次数下降。按任何以「与公认意见的偏离度」为指标的评估，我在进步。而唯一在退化的那样东西——「在一个尚无公认意见的新情形里，我自己能不能判断」——只在遭遇新情形时才显形，而那时通常已经晚了。

5.4 与「过度概括」批评的关系

对悲观论的一项标准批评是过度概括：若依赖证言令人不安，则许多无可指摘的信念形成方式（溯因、归谬、直觉）同样应当令人不安。本章的机制不遇到这一批评，因为它并不主张依赖本身令人不安。它主张的是一个条件句：在鉴别不独立的领域，依赖会同时消耗鉴别的能力。 在鉴别独立的领域，同样的依赖没有这个后果——我依赖医生一辈子，我鉴别医生的能力（读执照、读并发症率、听转诊）丝毫不受影响，因为那个能力从一开始就不是医学能力。这正是本章判据的价值所在：它把一条模糊的不安，收束成一个有明确适用条件的机制。

5.5 一个可能的出路，及其代价

这一机制并非无解，而它的解法形状值得点出，因为它与常见建议不同。由于在不独立领域鉴别资质与执行资质是同一样东西，唯一有效的干预是运用它——而不是提高对它的认识、不是学更多的判断标准、不是找更好的老师。这三样都属于「先得判断该学什么、该信哪套标准、该跟谁」，因而重演同一个循环。有效的动作只有一个形状：在有下游使用的场合，先自己作出判断并记下来，再去看别人怎么说。 顺序是关键。先看别人的，你自己那一次判断就不会发生，而不发生的判断不产生任何资质。其代价是实在的：这样做的当期产出更差，且这一更差是可测量的，而其收益不可测量且延迟。这使它在任何以当期产出计量的安排中都处于劣势。本章对此没有对策——第八节末对同一困难的表述是一样的，因为它本来就是同一个困难。

六、以此重新解释三家的证据

一个判据的力量，不在于它另起炉灶，而在于它能否把三家已有的证据重新安置。本节逐一处理。

6.1 知识论一侧：「自己做研究」的经验代价

这一支最有力的经验支撑是：坚持自行判断在若干领域可靠地把人带离真理，阴谋论与各类抗拒专业意见的现象是标准例证。按本章的判据，这些例证全部落在鉴别独立的领域：疫苗是否安全有效、气候是否在变暖、某种疗法是否有效——这些领域都存在不预设该领域判断的结果反馈，且外行可读（人有没有病、冰有没有化、病人有没有好）。在这些领域坚持自主，确实是把一个可外部化的鉴别硬做成执行，而且做不好。因此这些证据并不支持「自主一般地是坏理想」，它们支持一个更窄的结论：在鉴别可外部化的领域，坚持自己判断是低效且危险的。 而这正是本章所主张的。还有一层值得点出：抗拒专业意见者的自我描述往往正是「我自

己做了研究」——也就是说，自主是一件可以被表演的事。表演的自主与实际的执行资质在外观上不易区分，而这一点使得以「他们坚持自主」为由来反对自主的论证需要额外小心。

6.2 伦理一侧：延迟的自我巩固

如 5.1 所述，本章的能力机制与该动态机制独立且叠加。此处补充一点两者的分工：- 动机机制解释了为什么人不去补（既然问题已了结，寻求更广理解的理由就减少了）。- 能力机制解释了为什么即使去补也更难补（用于判断该向谁学、哪种解释是好解释的那个资质，本身已经停滞）。

两者相加，解释了一个常被观察到的现象：在这类领域，长期依赖者一旦决心自己判断，往往不是从零开始，而是从一个被扭曲的起点开始——他会把「更符合他所依赖过的那个来源」误当作「更正确」。

6.3 美学一侧：乐观论的描述性证据

这是本章最要紧的一次重新安置。罗布森举出的描述性证据是有力的：我们确实在若干场合大量依赖审美证言而不觉有何不妥——已经失传、无从亲见的作品；未曾到过的自然景观之美；经时间检验的审美共识。本章的主张是：这些例子有一个未被注意的共同点——它们全部落在「没有下游使用」的一侧。- 失传的作品：我永远不会亲见它，因而这个判断不会被我用于任何后续的观看、比较或选择。- 未到过的景观：同上，直到我真的去。- 经时间检验的共识（如某位大师的卓越）：这类判断在我的实际审美生活中几乎不承担任何工作——它不决定我今晚看什么、不影响我对眼前这幅画的反应。

而在有下游使用的场合——我要挑一件作品带回家、我要决定把三年时间投进哪一种手艺、我要判断眼前这位学生的作品是否值得鼓励——同样的依赖立刻显出问题，而这正是悲观论直觉最强的那些场合。于是乐观论的证据不是本章的反例，而是本章所预测的正常情形：在鉴别不独立的领域，当一个判断没有下游使用时，鉴别失败不产生代价，因而看不出问题。这一重新安置有一个可检验的后果，见第十节第三条。

6.4 私下情形与阮氏的反例

阮氏对信号解释的反例是：在完全私下、不向任何人标示的场合，那种「不对」的直觉反而更强。信号解释无法处理这一点，因为私下场合没有信号。本章的判据可以处理它，而且给出的解释相当直接：私下场合正是下游使用最纯粹的场合。 我私下依据别人的判断决定挂哪幅画、听哪张唱片，这个判断将持续作用于我此后的全部感受与选择——它有最长的下游。而公

开断言反倒常常是一次性的社交动作，其下游可能为零。按 6.3 的框架，下游越长，鉴别失败的代价越大，「不对」的直觉因而越强。这与阮氏观察到的方向一致。

6.5 三家为何各自到不了这一条

把三家的处境并排，可以看清这条分界为何没有被它们中的任何一家提出。知识论一侧到不了，因为它的问题意识是辩护性的：面对广泛的抗拒专业意见的现象，它要论证依赖是理性的。而它援引的全部案例恰好落在鉴别可外部化的一侧——这不是选择偏差，是那些案例正是当时的公共问题所在。在一个全部样本都独立的样本集里，独立性不会成为变量。伦理一侧到不了，因为它的问题意识是接收端的：既然可以传递知识，为什么还不安？它因此把全部注意力放在「传过来的是什么、没传过来的是什么」上。上游那一步——你怎么选中了这位证言者——在它的问题构造里根本不出现，因为在思想实验中那位证言者总是被规定为「道德上可靠的人」。规定掉的东西不会成为变量，而本章的整个论点正在被规定掉的那一步上。美学一侧到不了，因为它的任务是拆除：它要证明那条禁令没有独立支撑。而拆除者的方法是举反例，反例越多越好。它因此不会去问「这些反例有没有共同点」——一旦去问，就等于承认那条禁令在某个范围内成立，而那正是它要否定的。6.3 所做的，正是它没有动机去做的那一步。三家合起来才逼出这条分界：一家给了「可外部化的凭据长什么样」，一家给了「不安在哪里」，一家给了「必须交出一个说明」这个方法论要求。三样一凑，那个被共同规定掉的变量就浮上来了。

七、十二个领域的兑现

一条判据若只在提出它的领域内成立，它就只是那个领域的经验。本节在十二处兑现，并按独立性排出一条谱系。

7.1 高度独立的一端

急诊医学。 结果硬碰硬且外行可读（人活没活）。执照与追责制度由这些结果锚定。依赖完全安全，且坚持自己判断是危险的。结构工程。 同上。桥塌与不塌不由工程学评价决定。天气预报。 预测在短期内被现实裁决，且裁决外行可读。这使得预报机构的可信度可以被完全不懂气象的人排序。

7.2 中段

税务与法律实务。 部分独立：申报是否被驳回、案子是否胜诉，是外行可读的结果；但「这个方案是否稳妥」在很长时间内没有结果反馈，且很多风险终身不显形。故其独立性随事项的结果显形速度而变。投资。 表面上极度独立（赚没赚钱是硬结果），实际上受噪声与周期

长度的严重干扰，短期结果几乎不携带信息。这是一个外部结果存在但读取需要专业能力的典型——条件一成立而条件二不成立，故独立性远低于表面。临床判断中的疑难部分。 常规诊疗高度独立；而在结果延迟数年、且多因素交织的部分（慢性病管理、生活方式建议），结果反馈的读取开始需要专业能力，独立性随之下降。

7.3 低度独立的一端

审美判断。 如 3.3 所述。道德判断。 如 3.3 所述。补一点：道德领域存在一个特殊的复杂性——某些道德判断有外行可读的后果（政策造成的死亡人数），但从后果到「这是错的」这一步仍是道德判断，故条件一仍不成立。「什么问题值得做」这一类判断。 学术方向的选择、研究问题的取舍、一个组织该往哪走。这一类的独立性极低，因为其结果反馈的周期长于职业生涯，且「这个方向是对的」的认定由同行作出。本章认为这是被严重低估的一格——它在形式上被当作专业判断（因而可依赖专家），而在结构上属于不独立领域。同行评审。 属混合型，且其两部分的独立性差距极大。「数据是否造假、方法是否有误」这一部分接近独立——存在可核查的事实，且核查不必预设对该方向价值的判断。而「这个问题是否重要、这个方向是否有前途」这一部分属于最低度独立的一格：它的结果反馈周期长于评审者的职业生涯，且其认定完全由同行作出。同一份评审意见里并置着独立性相差一个数量级的两类判断，而制度以同一套程序处理它们。人才选拔。 「能不能做这件具体的事」可以通过当场做一件小的来锚定，接近独立；「这个人将来能走多远」则完全不独立——它没有可读的近期结果，且其认定由已被选拔出来的人作出。选拔制度的可靠性因而随其权重在两者之间的分配而变。育儿与教育中的方向判断。 「这个孩子该往哪个方向走」是本章谱系中独立性最低的一格之一：结果显形以十年计，读取需要判断，且几乎所有可用的凭据（升学、评奖）都是同类判断的聚合。这一格之所以值得单列，是因为它是最多人每天都在其中依赖他人证言、而几乎无人意识到其鉴别不可外部化的一格。

7.4 谱系的用法

把十二处并排，可得一条实践上的读法：独立性随「结果显形速度」与「结果可被外行读取的程度」两者共同下降。前者决定凭据有没有机会被校准，后者决定校准能否不预设专业能力。这也给出一个警示：一个领域可以在自己不知不觉间从独立滑向不独立——例如当其结果的显形周期被拉长，或当其验收被改为由同行评定。这是一种在制度改革中相当常见、而几乎从不被记入的变化。

八、人工智能条件下的新处境

8.1 为什么这一分界现在才要紧

前述分界在原理上一直成立，但在实践中长期不构成日常问题，原因是：在不独立领域获取高质量证言的成本很高。要得到一位有品味的人对你手上这幅画的判断，你得认识他、约到他、让他愿意认真看。成本本身限制了依赖的频次。这一成本刚刚接近于零。而且——这是要紧的一点——新的证言来源在两类领域中都能产出高质量输出。

8.2 两类领域中的不同后果

在鉴别独立的领域，机器输出的可靠性仍可被外部锚定：代码跑不跑、诊断对不对、预测中没中。使用者不需要具备执行资质就能校准自己的信任程度。这是好消息，且是这项技术最扎实的收益。在鉴别不独立的领域，情况相反。机器给出的审美判断、价值权衡、方向建议，其可靠性无法被不具备该项执行资质的人评估；而它恰恰最令人满意，因为它按使用者已有的偏好优化。于是得到一个可检验的预测：> 在鉴别不独立的领域，机器辅助会先提高使用者的满意度与产出的表面质量，随后降低其执行资质，而使用者无法从自己的体验中察觉这一下降——因为按 5.3，退化的那一项只在新情形中显形。

8.3 一个不对称的建议

由此得到一条与常见建议方向不同的操作原则：- 在鉴别独立的领域，尽量依赖，并把力气花在维护外部凭据（可核查的成功记录、追责链条）上。- 在鉴别不独立的领域，依赖之前先自己做一遍——不是为了产出更好，产出多半更差；而是为了保住那个用来判断该不该采纳的能力。

第二条的成本是实在的，而它的收益不在当期。这使它在任何以当期产出计量的安排中都处于劣势——这一点本章没有对策，只能指出。

8.4 一个可以现在就做的检验

上述预测不需要等待多年才能被检验。一项低成本的设计如下。在某个鉴别不独立的领域（如某类作品的评判）中，取两组水平相近的从业者。实验组在作出自己的判断之前先获取机器的判断；对照组先自己判断，再看机器的判断，并记录两者的差异。两组的产出质量在短期内很可能实验组更高。关键的观测项不是产出，而是在一批刻意构造的新情形上的表现——那些机器与既有共识都没有现成答案的情形。若 8.2 的预测成立，则实验组在这一项上的表现应当随时间相对下降，而两组在自评的把握程度上不应出现相应差异。「产出更好而自评不变，唯独在新情形上落后」这一组合，是本章机制的特征指纹；若观察不到它，则 8.2 落空。

九、正面回应六项反驳

9.1 「这只是把熟识原则换个说法」

不是。熟识原则说的是审美判断必须基于第一手经验，它是一条关于信念来源的规范，且被批评为无法解释大量正当依赖的案例。本章说的是：在某些领域，鉴别不可外部化，因而选择听谁已经用掉了执行资质。这是一条关于可行性结构的描述，不是规范；它不禁止依赖，它指出依赖在这些领域没有省去它看起来省去的东西。两者的可检验后果也不同：熟识原则预测一切二手审美判断都有缺陷，本章预测只有有下游使用的那些才显出缺陷——而 6.3 表明后者与观察更相符。

9.2 「审美与道德领域也有专家，我们也确实能识别他们」

我们确实能，而且本章并不否认。问题在于凭什么。细看任何一次「我认得出谁有品味」的实际过程，其依据都落在 4.2 的第三层：我读过他说的，我自己核对过，对得上。这恰恰是以自己的执行资质去衡量。它有效——但它不是外部化，它是行使。而落在第二层（圈子公认）的那些识别，其可靠性完全取决于那个圈子，而评估那个圈子仍需同一种能力。历史上审美与道德共识的大规模转向（对某些艺术门类的重估、对某些制度的道德重判）正是这一层不可锚定的证据。

9.3 「结果反馈在道德领域是存在的：看后果就行」

这是最强的一项反驳，需要认真处理。后果确实可以被观察，而且往往外行可读（有多少人受害、有多少人得益）。但从「发生了这些后果」到「那个决定是错的」这一步，需要一次道德判断——需要决定哪些后果算数、如何在受损者与受益者之间权衡、是否存在与后果无关的约束。而这一步正是我们要外包的那一步。这与医学的对照很清楚：从「病人死了」到「这次手术的判断是错的」这一步，可以由医学专业内部以外行也能核查的方式作出（并发症率、同类手术的基线、尸检）。而从「这项政策造成了这些损失」到「这项决定是错的」，没有对应的可核查步骤——不同的道德立场会从同一组后果读出相反的结论。因此条件一在道德领域确实不成立，尽管后果可见。

9.4 「那么这一切是否只是说：价值判断不能外包？」

不是，而且这一简化会丢掉本章最有用的部分。「价值不能外包」是一条老话，它的问题是给不出边界，因而在实践中要么被无限扩张（凡涉及价值皆不可依赖，于是寸步难行），要么被当作套话。本章给的是一条可操作的分界，而这条分界不与「是不是价值判断」重合。7.3 给出的第三格——「什么问题值得做」——在形式上被当作专业判断，很少被当作价值判断，而

它按本章的判据属于低度独立。反过来，某些明显带价值色彩的判断（如某项安全标准的取舍）由于其结果显形快且外行可读，独立性相当高。判据看的是验收结构，不是内容的性质。

9.5 「若本章成立，我们凭什么相信本章」

这是本章自反的入口，也是唯一一项本章无法完全化解的反驳。「哪些领域的鉴别可外部化」这个元问题，其自身的鉴别是否可外部化？部分可以：本章的判据部分依赖可核查的事实（某领域有没有外行可读的结果反馈、有没有人因判断错误而承担过可被外人读出的后果），这些是第三方可查的。但整体不完全可以——判断本章这条分界是否切中要害，仍要用到读者自己的判断力。因此本章不能靠权威被接受。它只能靠使用被检验：拿第四节那一问去问你手上正在依赖的三件事，看它是否分出了你原本分不出的东西。分出来了，它对你成立；分不出，那它对你就是一句空话，无论其论证多么严密。

9.6 「凭据可以被制造出来：为什么不给审美与道德也建一套」

这是一项建设性的反驳，值得认真对待，因为它指向本章框架的一个真实的行动方向。若鉴别的独立性来自可锚定的凭据，而凭据的可锚定性来自外行可读的结果反馈，那么原则上可以为任何领域建造凭据——只要能找到一种外行可读的结果。历史上确有先例：医学的执照与并发症统计不是自然存在的，它们是被造出来的；天气预报的可信度评级同样如此。本章的回答分两层。第一层：确有可造的部分，而且值得造。 在审美与道德领域中，存在若干可被外行读取的次级结果：一位评审此前推荐过的作品，后来被更广的时间检验如何；一位顾问此前给出的建议，其当事人后来是否需要推翻。把这类记录变成可查的，确实能提高该领域的独立性。本章第七节所说的「独立性随结果显形速度与可读程度而变」，正蕴含这一条：改变这两个量就改变了独立性。第二层：但这类凭据的锚仍是同类判断。 「后来被时间检验如何」这句话里的「检验」，仍由具备该领域判断的人作出。这与并发症率不同——病人是否死亡不由医学判断决定。因此这类凭据能把独立性从极低提到中等，不能提到医学那一档。这一分层给出一条实践建议，也给出一项限度：值得为不独立领域建造可查记录，但不应期待它能把该领域变成可安全外包的领域；而在建造过程中最要防的，是第二层被当成第一层来用——即把同类判断的聚合当作外部锚定。按 4.2，那正是三层中最危险的一层。

十、边界、局限与可证伪条件

10.1 三条不适用边界

其一，紧急与高危场合按独立领域处置。 即便某项判断带有价值成分，若时间不允许且已有成熟凭据，正确的动作是依赖。在此处引入本章框架，会为拖延提供一个听起来有洞察力的

理由。其二，尚不具备任何执行资质者应当大量依赖。本章的机制以「存在可被消耗的执行资质」为前提。一个刚入门者没有可消耗的东西，他的正确路径是大量吸收他人的判断。分界应在他开始对某些判断产生自己的不同意见之后才启用。其三，本章不为拒绝倾听背书。见 4.4。

10.2 四项局限

局限一：独立性是连续谱，本章以三分呈现。第七节的谱系已表明多数领域居中，而本章的判据在中段给出的指示较弱。局限二：本章未处理类型的迁移。一个领域可从独立滑向不独立（7.4 末），反之亦可（当某领域建立起可核查的结果记录时）。这一迁移过程本章未着一字，而它可能是最具政策含义的部分。局限三：「下游使用」这一概念未被充分刻画。6.3 依赖它作出重新安置，而本章只给了例示，没有给出判定某个判断有无下游使用的独立标准。局限四：本章只讨论了个体依赖，未讨论制度依赖。一个组织如何在不独立领域组织其判断（评审、评奖、方向决策），是本章框架的自然延伸，而本章未展开。

10.3 六条可证伪条件

· 其一，一问不可操作。若若干具专业背景者分别用第四节那一问对同一批领域作独立／不独立分类，结果无法达成一致，则该判据不可用，本章的实践主张随之落空。此项成本极低。

· 其二，成因说明不成立。若能举出一个鉴别可外部化、而其中不存在「不预设该领域判断且外行可读的结果反馈」的领域，则 3.3 的成因说明错误。

· 其三，下游使用与直觉强度无关。若在受控设计中，对有下游使用与无下游使用的审美依赖，人们的「不对」直觉强度没有系统差异，则 6.3 的重新安置失败，罗布森的描述性证据重新构成对悲观论的独立反驳。此项可用问卷设计直接检验。

· 其四，鉴别与执行在不独立领域可分离。若能展示一种方法，使不具备审美或道德执行资质者可靠地排序该领域的证言者，则本章的核心分界不成立。

· 其五，自我锁死不发生。若在不独立领域长期依赖他人判断的人群，其执行资质未见相对下降，则第五节的机制错误。

· 其六，人工智能条件下的预测落空。若在鉴别不独立的领域，长期使用机器辅助者的执行资质未下降，或其下降可被使用者自行察觉，则 8.2 的预测错误。

10.4 本章不主张什么

本章不主张少依赖他人，不主张自主是一种德性，不主张审美与道德判断无法达成知识。它只主张一件事：一个领域能否安全地依赖证言，取决于该领域的鉴别资质是否独立于执行资

质；而在不独立的领域，选择听谁本身已经是那个判断的一次完整行使。其余全部内容皆由此推出。

10.5 被本章的整理动作扫掉的部分

按本章自身的立场，一次整洁的整理会腾出解释的位置，故此处原样列出被扫掉者：- 三家的内部分歧远比本章的呈现复杂。知识论一侧并非只有反自主一支，格林与卡沃尔的辩护是活的且有力的；伦理一侧的乐观论者（斯利瓦、博伊德等）有系统的回应，本章只在 5.4 处理了其中一条；美学一侧对熟识原则的辩护近年仍在出版，里格尔所说的「近乎共识」是一位与会者的表述而非计量结果。- 本章对戈德曼名单的处理（2.5）压缩得相当厉害，其中「论证的表现」一条在文献中有比本章所述精细得多的讨论。- 本章所举的三个「审美乐观论例子皆无下游使用」的案例是本章自己的归类，而罗布森本人并未以此方式呈现它们；这一归类是否穷尽了他的例子，本章未作检验。- 本章没有处理集体层面的鉴别。一个共同体如何在不独立领域维持其判断质量（学派的传承、行会的师承、评审团的构成），是一个与个体层面机制不同的问题，本章的框架对它只给了一个方向而未展开（见 9.6 第一层）。

十一、结论

本章的结论可收为三句。其一，证言依赖的分界不在内容的种类，而在验收的结构。审美与道德之所以被当作例外，不是因为它们涉及价值，而是因为在这两个领域中不存在不预设该领域判断、且外行可读的结果反馈；缺了这一条，凭据无法被锚定，鉴别便无法外部化。其二，在鉴别不独立的领域，依赖并未省去判断。选择听谁本身就是那个判断的一次完整行使。真正的风险因而不是「依赖」，而是以为自己没有在判断——这正是同类判断的聚合（「大家都说他好」）之所以是三层中最危险一层的原因。其三，这一分界现在才成为日常问题。在不独立领域获取高质量证言的成本刚刚接近于零，而在该领域中，输出的可靠性无法被不具备该能力者评估，且它按使用者的偏好优化因而最令人满意。由此得到的操作原则是不对称的：在鉴别独立的领域尽量依赖，并把力气用于维护可核查的结果记录；在鉴别不独立的领域，依赖之前先自己做一遍——不是为了产出更好，而是为了保住那个用来判断该不该采纳的能力。最后需要指出本章最易被扭曲的一处：这条判据的正当用法是施于自己正在依赖的事，而非用以质疑他人的依赖。「你凭什么认为他靠得住」这一问可以使任何一次引用停摆，而提问者不承担任何代价。其有效的用法只有一种：对着自己手上正在依赖的三件事各问一遍，看其中有几件，你给出的理由是一个不懂行的人也能用的。

合章

十项读数如何互相锁住

不引入新材料，只把十章各自的读数接进枢纽章那七条关系，并给出五条可以同时推翻它们的方式

位置　本章排在第四编之后、结语之前。它与枢纽章的分工是：枢纽章从三条判断推出七条关系，本章反过来，把十章的读数逐一放回这七条关系里，看它们是否彼此约束。一本书与十篇文章的差别，就在这一步做不做得成。

一、十项读数先列一遍

章	读数	一句话判据	归入
一	复推率 ρ	换一个表面不同、结构相同的情境，他会不会推出同一条（至少测两次）	R2
二	内外表述落差	该领域对同行说的与对使用者说的，是不是同一件事	R6 的上游；R7 的通解（第六之三节）
三	影子正本占比	「这几份应该一样吗」多数答应该，而「哪一份算数」答案分歧	R3 的第三种形态（第六之二节）
四	借清率 b	相邻处清晰度下降量 ÷ 目标处上升量	R5
五	剥夺落差 Δ	临时封掉替代路径，结果跌多少、多久回得来	R1、R3、R4
六	带责分叉数	单位工时内，路真的分开、他知道并被允许走、后果落在他本人身上的位置有几处	R3
七	类内距	合格样本在使用者关心的维度上的离散度	R6 的对偶

八	记零位	「发现了一个错误」记在谁名下、算多少；能否跟人走、进晋升材料、被第三方核对	R3 的第二条独立路径
九	静息变异率 v	无任务、无考核、无故障的时段里，自发做出的与默认做法不同的事占多少	R5 的代理
十	鉴别独立性	我凭什么认为他靠得住——这个理由，不懂行的人能不能用	R7

十项里有六项在本书之前互不相识：它们由六个不同的三学科组撞出来，作者写第七章时不知道会有第九章。本章要检验的正是这一点——它们是不是同一件事在十个位置上的显形，还是十件恰好被放进一本书的事。

二、第一处锁：Δ 与 ρ 是同一个量的两次测量

R2 给出 $\rho = \rho_{max}(1-\lambda)$，R4 给出 Δ 是 λ 的唯一读法。消去 λ：

$$\Delta / y_0 = 1 - \rho/\rho_{max}$$

于是复推率与剥夺落差不是两项独立的指标，而是同一件事的两种测法：一种在正常运行中做（换情境让他重推），一种要停机做（封路量跌多少）。 前者便宜而侵入学习者，后者昂贵而侵入系统。

这条给出一个可以直接执行的核对：在同一批人身上同时测 ρ 与 Δ，两者应当负相关，且斜率由 ρ_{max} 定。 若测出两者不相关，则第一章的复推率与第五章的剥夺落差测的不是同一样东西，本书第二编的两章不能放在同一编里。

这是本章成本最低、也最容易推翻全书的一次核对。 它需要的只是一组学习者、两次测量、一次封路。

三、第二处锁：两条独立的路走到「否定不入账」

R3 说凡不进入产物的都在退，而只有进入产物的被记录。第六章与第八章各自到达了这一条，而且互相不知情。

· 第六章：被删除的是岔路不是步骤。岗位在、工序在、工作量甚至上升，减少的只是那些「本来可以走另一条」的位置。岔路不进入产物。

· 第八章：被记为零的那类动作有一个共同特征——它的成果是一个否定，是某件事没有发生，因而无法计量、易被稀释、无法向第三方证明。否定不进入产物。

「岔路」与「否定」在形式上是同一样东西：一条没有被走的路，与一个没有发生的事故，都只能由它们不留下的东西来指认。

这一处锁给出一条推论，两章各自都推不出来：

> 凡以产物为结算单位的记账法，必然同时漏记岔路与否定；而这两样合起来，正是一个共同体在下一轮所需要的全部辨认能力。

第八章的航空实例是这条推论目前唯一一处被真的做出来的解：它没有去给「发现」记功，它给「报告」记了账——因为报告有一个可确定的时点与一份可比对的内容，而发现没有。把这条办法翻译到第六章的题域上，得到一个此前没有被提出过的处方：不要试图记录「他做了一次选择」，去记录「他提交了一次不采纳」。 前者无法核对，后者有时点、有内容、有第三方。

第六章明写现行法规要求人能够不采纳系统输出，却既不要求记录否决率，也不分配否决的责任。第八章给出的正是补上这一栏的办法。两章隔着认知教育、伦理人类学、法哲学与专利法、作坊研究，接口在这里对上了。

四、第三处锁：类内距是覆盖型量表在产物侧的对偶

R6 说覆盖型量表给均浅者加分。第七章说的是产物侧的同一件事：一份规范只写了什么算同类，没写同类之间有多近；压缩类内距的那笔功夫不会因为无人付钱而消失，只会换个地方发生。

把两条对起来：

	人的一侧	产物的一侧
量表只写了什么	及格线（格）	边界（合格判定）
量表没写什么	峰高与坑深（锐度 κ）	同类之间有多近（类内距）
没写的那一样去了哪	落在使用现场，表现为「他证书齐全但上不了手」	落在使用侧，表现为返工、胶合代码、排练次数
谁付这笔钱	用人方，且不入账	使用方，且不入账

于是 R6 与第七章是同一条判断的两个投影：规范化把「合格」写清楚，同时把「有多好」从账上删掉；而 $a \to 1$ 使合格产物的供给趋于无限，于是全部压力转移到那个已经被删掉的量上。

由此得到一条本书之外可以立即使用的推论：当一个行当里合格产物突然变得随手可得，它的返工率、上手时间、胶合成本会上升，而它的合格率会同时上升。两条曲线同向背离，是这件事在产物侧最早、最便宜的信号。第七章第 7.3 节那条不需要新数据的证伪条件（把规范条款数与胶合代码量对齐）在 a 上升的年份里应当出现一次可见的转折；这一点第七章没有预言，是本章加上去的，因而它也是本章自己的一条赌注。

五、第四处锁：静息变异是均浅唯一的无侵入读法

R5 说均浅者的 κ 没有定义，因为峰与坑都不在。这带来一个测量上的死结：一切基于「让他做点什么」的测量，都要求先给出任务，而任务一旦给出，a 就在起作用——他可以取用。

第九章正好处理这个死结。它的判断是：那个只在「什么也没要求」时才存在的量，与「被要求时能出多大反应」，是同一个量；而三门学科都在花钱把它降到零。

接起来：

> v（静息变异率）是 κ 在无任务条件下的唯一代理，因为它是唯一一个不给任务就能读的量。

这条锁有一个不受欢迎的推论。第九章已经指出一种最省事的作弊：划出两小时叫「自由探索」，有场地、有交付物、有考核——在这两小时里测到的一切，按定义一条也不算数。合章要补的是：在 a 高的环境里，这种作弊几乎不可避免，因为一旦有交付物，取用就是理性的。于是 v 的可测性本身随 a 下降。

这与 R4 是同一个形状：读出这件事的唯一办法，其成本与失效风险都随 a 上升。本书三次遇到这个形状（Δ、v、κ），三次都不是巧合——它们都是在测「没有发生的那一半」。

六、第五处锁：追认链闭合

第二章的判断是：一个领域采取哪种处置，主要不由对象性质决定，而由该领域下游使用的刚性需求决定，而这一选择随后被表述为一项关于对象本性的发现。

把它放到 $a \to 1$ 之后：

1. 下游使用的刚性需求，现在有相当一部分由供给侧（可随时取用的产物）塑造；2.「知识本来是什么样子」这一表述，据第二章，由下游需求追认；3. 于是关于知识本性的表述，正在被供给方的形状追认。

第二章的第四问——「如果下游需求改变了，这个立场会不会跟着变」——在这里有一个当下的应用对象：「知识就是可检索、可复述、可即时调用的东西」这个说法，若它在 a 低的年代不成立、在 a 高的年代成立，则按第二章的判据，它高度可疑地是一次追认，而不是一项发现。

这一条把全书的书名问回给了书名。「答案随时可得之后」，我们对「知道」的定义已经改了一次，而这次改动会被表述为我们终于看清了知识本来的样子。第二章给的四问，是本书能提供的唯一一件用来核对这次改动的工具。

六之二、第六处锁：影子正本是 R3 在同一性上的形态，而不是它的一个例子

第一节那张表把第三章的影子正本挂在「R3 的一个特例」上。这是十项里接得最松的两处之一，本节把它接紧——而接紧之后它不再是特例，是 R3 的第三种形态。

R3 说：凡进入产物的都留下，凡不进入产物的都在退，而只有前者被记录。前面已经指认了两样不进入产物的东西——岔路（第六章）与否定（第八章）。第三章指认第三样：「哪一份算数」这个答案本身。

它符合同样的形式。这个答案不是任何一份文件的内容，不在任何一份副本里；它存在于每个人心里，而每个人心里的不是同一份。第三章那句「它有全部的效力，没有任何的位置」，正是「不进入产物」的字面表述。于是：

> 岔路、否定、算数的那一份——三样都是关于产物的事实，而三样都不在产物里。

接到可及性上，得到一条第三章写不出、而 R3 也推不出的判断。第三章说影子正本这些年在变多，理由是数字化使副本廉价。这个理由现在不够了：副本廉价只解释份数上升，不解释预期为什么不跟着调整。补上的是 a：

> a 上升使「取到一份合格产物」这个动作的成本趋于零，因而每一次需要时都可以现取一份，而不必回到那一份。于是副本的产生频率上升，而每一份的持有时间下降——「哪一份算数」这个问题被问到的次数反而减少了。

一个人不会去问一份他五分钟前才取来、用完就丢的东西算不算数。预期一致的那个预期，因此不是被撤销的，是被绕过的——它连被问一次的机会都没有了。第三章说从未被表述的预期连撤销的对象都找不到；这里加一句：a 高的环境使它连被表述的场合都消失。

判决性对照：第三章预测影子正本随副本份数上升而上升；本节预测在 a 很高的通道里，份数继续上升而「哪一份算数」的分歧度先升后降——不是因为对齐了，是因为没有人再持有任何一份足够久，久到能形成一个关于它的预期。两条曲线在中段分开，一个学期的问卷即可分辨。这一条第三章没有说过，本书也没有测过。

六之三、第七处锁：R7 的自锁，第二章早就写下了它的通解

第一节那张表把第二章挂在「R6 的上游」上。这是另一处接得松的地方，而它其实可以接到全书最末端的那条关系上去。

R7 说：鉴别独立性低的领域里，依赖的安全性取决于 ρ；而依赖本身使 λ 上升、ρ 下降。这是一个自锁：被依赖的那一样越好用，评判它好不好用的能力越少，因而连「该不该换」这个问题也逐渐无法被提出。

自锁的形式是：一个立场的维持条件，由使用它所带来的后果供给。而这正是第二章第七节第四问所定义的那个形状——「如果下游需求改变了，这个立场会不会跟着变」。第二章处理的是本体论立场，本节处理的是依赖关系，两者是同一个结构的两次出现：

	第二章	R7
被维持的东西	一项关于对象本性的表述	对某个来源的依赖
维持它的力	下游使用的刚性需求	使用它所降低的那项鉴别能力
为什么难以察觉	被表述为一项发现	被体验为一次省力
第二章的解	四问，尤以第四问可事前预测	同样的四问，改问依赖

于是第二章的四问是 R7 自锁的通解，只需换一个宾语。逐条改写如下：

1. 相反面在实务上是否可行？　如果明天起不能使用这个来源，这项工作还能不能进行？若完全不能，那么「它可靠」这个判断高度可疑地是被需求决定的，与它是否真的可靠无关。2. 这个立场是什么时候形成的？　对它的信任是在有能力核对的时期形成的，还是在已经无力核对之后形成的？后者是自锁已经闭合的标志。3. 内外表述是否一致？　对同行说的（「它当然有出错的时候」）与对使用现场说的（直接采用），是不是同一件事？第二章说内外不一致是追认物的稳定特征。4. 需求变了它会不会跟着变？　若这项工作的验收标准改成需要出示推导过程，对这个来源的信任会不会立刻改变？会，则那份信任一直是被需求维持的。

第四问在这里最有用，因为它可以事前问，不必等到出事。而 R7 的麻烦恰恰是它不产生事件——第十章已经写明，鉴别不独立的领域里错误在数年后才显形，且显形时由同一批人认定。

这一处锁的代价要写清楚。把第二章的工具搬到 R7 上，等于承认本书对自锁没有给出新的解，只是指出旧工具适用。这与第二章、第十章各自的自陈是一致的：两章都没有主张能解开它，只主张能问出它。合章不该在这里比它的材料更乐观。

七、五条推翻方式

十项读数被锁在一起，代价是它们可以被一起推翻。下列五条，任何一条成立，本书相应部分作废。

其一（针对 R1–R2 的核心）　在某个 $0 < a < 1$ 且无人为封路的环境里，λ 长期稳定在明显低于一的水平。$\Longrightarrow$ 枢纽章的递推形式错，全书脊梁断。

其二（针对第二处锁）　在同一批人身上同时测复推率 ρ 与剥夺落差 Δ，两者不相关。$\Longrightarrow$ 第一章与第五章测的不是同一样东西，第二编不成编。

其三（针对 R6 与第四处锁）　在一份分格的覆盖型量表上，深学者系统性地高于均浅者；或在 a 明显上升的年份里，某行当的合格率与返工率同向下降。$\Longrightarrow$ 推论三错，第三编与枢纽章第 7 节作废。

其四（针对第五处锁）　在一个 a 很高的环境里，静息变异率 v 被无侵入地、稳定地测了出来，且与该群体的复推率不相关。$\Longrightarrow$ v 不是 κ 的代理，第九章与枢纽章第 6 节之间没有接口。

其五（针对本书自身的结构）　a 在多年内明显上升，而上述四项代理（ρ、Δ、v、κ）没有一项下降。$\Longrightarrow$ 本书全部十章各自可能仍然为真，但它们不是同一件事，这本书不该被装成一本书，应当拆回十篇。

第五条是本章最重要的一条。它不针对任何一章的内容，只针对把这十章放在一起这个动作。一本由十篇已发表文章装成的书，唯一需要单独辩护的就是这个动作；把辩护写成一条可以判它错的条件，是本书能做到的最诚实的形式。

八、一句必须写在这里的话

本章用七处锁把十项读数联立，这件事本身有一个代价：锁得越紧，越像一个自洽的系统，而自洽是最容易被误认为正确的性质。

七条关系里，只有 R1、R3、R4 有实测支持，且都是别人为别的目的做的测量。R2、R5、R6、R7 是从定义推出来的。第 6 节那条追认链，本身就是第二章那把刀可以砍向本书的方向：如果「知识是每次被重新推出来的事件」这一表述，恰好为本书所处的位置

（一个由人定题、由机器成文的写作方式）提供辩护，那么按第二章的四问，它也高度可疑地是一次追认。

本书没有办法自己回答这个问题。它能做的只有一件：把四问原样留在第二章，不作任何削弱。

结　语　一张十项的体检表，和它到期作废的那一天

一、十章说的是同一件事

十章由十个不同的三学科组撞出来，写作时互不知情。放在一起之后，它们说的是同一句话：

> 生成一份合格产物的成本趋近于零之后，被取消的不是知识，是重新推出知识的那一段；而那一段里长出来的全部东西，在现行的记账法里一栏也没有。

这句话不需要任何人做错事。第六章说明得最清楚：删除的顺序不由该不该删决定，而由可结算性决定，而可结算性与后果重要性是两条正交的轴。可结算的先被删，因为它便宜、干净、可以写进方案。于是最先消失的，恰好是最容易被安排掉的，而不是最不重要的。

二、为什么现行读数一定看不见它

三类读数，三种看不见的方式：

产物类记零。借道对结果读数恒为零（第五章）；岔路不进入产物（第六章）；核验的成果是一个否定（第八章）。凡以产物为结算单位，这三样一律不出现。

效率类记为改善。借道更省力（第五章），取消岔路减少差错（第六章），规范加细提高合格率（第七章），静息变异被记作失误率、噪声与非计划偏移（第九章）。这四样在效率账上全部记入贷方。

能力类记为进步。覆盖型量表上，峰在封顶处失效而坑照常扣分，深学者得分不高于均浅者（枢纽章推论三）。一场把峰与坑一起抹平的变化，在量表上读出来是覆盖面扩大。

三类读数各自都是对的，各自都在诚实地测它所定义的东西。问题不在测得不准，在这件事恰好落在三类读数的定义之外。

三、十项读数的体检表

结语的兑现物是下面这张表。它不是本书的观点，是本书十章各自给出的判据的汇编，每一项都注明数据从哪里来，以及一个与它正交的实例——正交是指那个实例里该项读数为真而其余九项未必。

#	读数	怎么问	数据从哪来	正交实例

1	复推率 ρ	换一个表面不同、结构相同的情境，他会不会推出同一条（至少两次）	一次教研或一次带教，成本一小时	一位能把规则背得极清楚而换个题面就推不出的人：ρ 低而其余九项正常
2	内外表述落差	这个领域对同行说的与对使用者说的，是不是同一件事	专业文献与使用手册各一份	新约文本批评：内部把构造性写进定义，使用现场不承认
3	影子正本占比	「这几份应该一样吗」多数答应该，而「哪一份算数」答案分歧	分头问四五个持有者，半小时	一个只有一份文件、而每个人心里的「最新版」不同的小组
4	借清率 b	相邻处清晰度下降量 ÷ 目标处上升量	相邻题域的配对测试	母语范畴锐化与非母语对立分辨损失：同一次搬运的两端
5	剥夺落差 Δ	临时封掉替代路径，结果跌多少、多久回得来	需要停机，成本最高	术后单腿跳远：距离比 97%、膝关节做功比 69%
6	带责分叉数	单位工时内，路真的分开、他知道并被允许走、后果落在他本人身上的位置有几处	岗位描述数据库（如 O*NET）现成可算	自动化程度与决策频次相关 0.042、与决策自由度 −0.199
7	类内距	合格样本在使用者关心的维度上的离散度	返工率、胶合代码量、上手时间，多数已有记录	规范条款数上升而胶合量不降的接口标准
8	记零位	「发现了一个错误」记在谁名下、算多少；能否跟人走、进晋升材料、被第三方核对	一份晋升材料模板即可判	同行评议：不记名、不可援引、不进履历
9	静息变异率 v	无任务、无考核、无故障的时段里，自发做出的与默认做法不同的事占多少	需事后清点，且极易被作弊	自动化提高后手飞技能保持完好，退化的是整段无人要求他做的那部分
10	鉴别独立性	我凭什么认为他靠得住 —— 这个理	一次自问，零成本	理财顾问：凭据齐全、事故不断、且

	由，不懂行的人能不能用	数年后才显形

第 9 项目前没有任何机构在常规测量。这不是本书的疏漏，它是第九章那条判断的一个读数：三门学科都为它设了字段，而三门学科都在花钱把它降到零。

四、怎么用这张表

三件事，按顺序做。

先测第 1 与第 6 项。两项都便宜，都不需要停机，且它们分别在人的一侧与岗位的一侧。若两项同时偏低而其余读数全绿，本书描述的情形大概率正在发生。

第 5 项要趁早测。枢纽章推论二说明了原因：剥夺落差的测量成本随可及性单调上升，而它的诊断价值也随之上升。它应当在可及性还低的时候先测一次留作基线，而这件事在任何组织里都不会被提出来，因为提出它的时候所有读数都还是绿的。

第 8 项最先能改。第八章从航空业带回来的那条办法是本书收集到的唯一一件真的被做出来过的解：不去给「发现」记功，给「报告」记账——因为报告有一个可确定的时点与一份可比对的内容，而发现没有。这条办法可以直接搬到第六章的题域上：不要试图记录「他做了一次选择」，去记录「他提交了一次不采纳」。

五、这张表进考核那天就开始失效

必须写在这里：这十项一旦被写进考核，它们立刻开始失效。

第九章已经写明了最省事的作弊：划出两小时叫自由探索，有场地、有交付物、有考核，然后报表上写着静息变异率从零升到了百分之四十。第八章写明了另一种：账目一开，记录的就不再是核验，而是关于核验的声称。第三章写明了第三种：只补一半比不补更糟。

这不是执行不力，是这些读数的性质。它们测的都是「没有发生的那一半」，而任何把没有发生的事变成一项考核指标的做法，都会立刻制造出它的表演版本。本书没有解决这个问题，也不认为自己解决了。它能给出的最诚实的建议是：这张表适合用来自查，不适合用来考核别人；一旦它出现在别人的考核表上，它测到的就已经不是它定义的东西。

六、如果这本书错了

四条，任何一条成立，相应部分作废：

其一，某个可及性大于零的环境里，借道率长期稳定在明显低于一的水平且无人为封路——枢纽章的递推形式错。

其二，同一批人身上复推率与剥夺落差不相关——第一章与第五章测的不是同一样东西。

其三，覆盖型量表上深学者系统性地高于均浅者——推论三错，第三编与枢纽章第 7 节作废。

其四，可及性多年明显上升而四项代理没有一项下降——十章各自可能仍然为真，但它们不是同一件事，这本书不该被装成一本书。

七、最后一句

这本书从头到尾没有说答案随时可得是一件坏事。它说的只有一件：在那件事发生之后，我们用来判断自己有没有损失的全部工具，都会告诉我们没有。

参考文献

本表按章合并自各章原有文献表，未作增删。

第一章　复推率

· 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

· 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

· 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

· 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.

· Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379-424.

· 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.

· Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1-38.

第二章　我们说的「本来的样子」，是我们造出来的

Epp, E. J. (1999). The multivalence of the term "original text" in New Testament textual criticism. *Harvard Theological Review*, 92(3), 245–281.

Epp, E. J. (2007). It's all about variants: A variant-conscious approach to New Testament textual criticism. *Harvard Theological Review*, 100, 275–308.

Gibney, E. (2025). Physicists disagree wildly on what quantum mechanics says about reality, Nature survey shows. *Nature*, 30 July 2025.

Greg, W. W. (1950). The rationale of copy-text. *Studies in Bibliography*, 3, 19–36.

McGann, J. J. (1983). *A Critique of Modern Textual Criticism*. University of Chicago Press.

McKenzie, D. F. (1986). *Bibliography and the Sociology of Texts*. British Library.

Mink, G. (2011). Contamination, coherence, and coincidence in textual transmission. In K. Wachtel & M. W. Holmes (Eds.), *The Textual History of the Greek New Testament: Changing Views in Contemporary Research*. Society of Biblical Literature.

Parker, D. C. (1997). *The Living Text of the Gospels*. Cambridge University Press.

Thompson, A., & Taylor, N. (Eds.). (2006). *Hamlet* and *Hamlet: The Texts of 1603 and 1623*. Arden Shakespeare, Third Series.

Wachtel, K., & Holmes, M. W. (Eds.). (2011). *The Textual History of the Greek New Testament: Changing Views in Contemporary Research*. Society of Biblical Literature.

第三章　哪一份算数

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), 1531–1545.

Hunt, A., & Thomas, D. (1999). *The Pragmatic Programmer: From Journeyman to Master*. Addison-Wesley.

Knight, J. C., & Leveson, N. G. (1986). An experimental evaluation of the assumption of independence in multiversion programming. *IEEE Transactions on Software Engineering*, SE-12(1), 96–109.

Lynch, M., & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science*, 290(5494), 1151–1155.

Ohno, S. (1970). *Evolution by Gene Duplication*. Springer-Verlag.

Reich, V., & Rosenthal, D. S. H. (2001). LOCKSS: A permanent web publishing and access system. *D-Lib Magazine*, 7(6).

United Nations. (1969). *Vienna Convention on the Law of Treaties*, Article 33.

字数：正文约 1.67 万汉字 · 版本 v1.0 · 2026-08-01 · 作者 王德生 + Claude

人机分工声明：本章的三家源材料均为站外既有文献（见注释与参考文献）；两条轴的设定、四格表、一分钟判据、以及第十九节的自我定位由本次写作产出。全文由机器一次写成，人类确定选题与验收标准。第十九节所述的那次重复事件属实，且尚未被处置。

第四章　借清

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*, 3(4), 128–135.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521–3526.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, 11(1), 23–63.

Francis, T. (1960). On the doctrine of original antigenic sin. *Proceedings of the American Philosophical Society*, 104(6), 572–578.

Cobey, S., & Hensley, S. E. (2017). Immune history and influenza virus susceptibility. *Current Opinion in Virology*, 22, 105–111.

Richards, B. A., & Frankland, P. W. (2017). The persistence and transience of memory. *Neuron*, 94(6), 1071–1084.

Tononi, G., & Cirelli, C. (2014). Sleep and the price of plasticity. *Neuron*, 81(1), 12–34.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories. *Perception & Psychophysics*, 50(2), 93–107.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination. *Proceedings of the National Academy of Sciences*, 88(11), 4966–4970.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal*, 13(S1), 111–125.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, 108(3), 652–661.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55–81.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67–82.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

第五章　借道

Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way. In *Psychology and the Real World*. Worth.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444.

Coplan, J. A. (2002). Ballet dancer's turnout and its relationship to self-reported injury. *Journal of Orthopaedic & Sports Physical Therapy*, 32(11), 579–584.

Dahmani, L., & Bohbot, V. D. (2020). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports*, 10, 6310.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *PNAS*, 98(24), 13763–13768.

Geirhos, R., Jacobsen, J.-H., Michaelis, C., Zemel, R., Brendel, W., Bethge, M., & Wichmann, F. A. (2020). Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2, 665–673.

Hodges, P. W., & Tucker, K. (2011). Moving differently in pain: A new theory to explain the adaptation to pain. *Pain*, 152(3 Suppl), S90–S98.

Kaufmann, J.-E., Nelissen, R. G. H. H., Exner-Grave, E., & Gademan, M. G. J. (2021). Does forced or compensated turnout lead to musculoskeletal injuries in dancers? A systematic review on the complexity of causes. *Journal of Biomechanics*, 114, 110084.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, 139(4), 665–682.

Koriat, A. (1997). Monitoring one's own knowledge during study. *Journal of Experimental Psychology: General*, 126(4), 349–370.

Kotsifaki, A., Korakakis, V., Graham-Smith, P., Sideris, V., & Whiteley, R. (2022). Single leg hop for distance symmetry masks lower limb biomechanics. *British Journal of Sports Medicine*, 56(5), 249–256.

McAllister, M. J., Blair, R. L., & Donelan, J. M. (2021). Energy optimization during walking involves implicit processing. *Journal of Experimental Biology*, 224(17), jeb242655.

Schachter, J. (1974). An error in error analysis. *Language Learning*, 24(2), 205–214.

Scholz, J. P., & Schöner, G. (1999). The uncontrolled manifold concept. *Experimental Brain Research*, 126(3), 289–306.

Selinger, J. C., O'Connor, S. M., Wong, J. D., & Donelan, J. M. (2015). Humans can continuously optimize energetic cost during walking. *Current Biology*, 25(18), 2452–2456.

Taub, E., Miller, N. E., Novack, T. A., et al. (1993). Technique to improve chronic motor deficit after stroke. *Archives of Physical Medicine and Rehabilitation*, 74(4), 347–354.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11), 1226–1235.

Tucker, A. L., & Edmondson, A. C. (2003). Why hospitals don't learn from failures. *California Management Review*, 45(2), 55–72.

Wellsandt, E., Failla, M. J., & Snyder-Mackler, L. (2017). Limb symmetry indexes can overestimate knee function after anterior cruciate ligament injury. *Journal of Orthopaedic & Sports Physical Therapy*, 47(5), 334–338.

Wong, J. D., Selinger, J. C., & Donelan, J. M. (2019). Is natural variability in gait sufficient to initiate spontaneous energy optimization in human walking? *Journal of Neurophysiology*, 121(5), 1848–1855.

Wren, T. A. L., Mueske, N. M., Brophy, C. H., & Pace, J. L. (2018). Hop distance symmetry does not indicate normal landing biomechanics in adolescent athletes with recent anterior cruciate ligament reconstruction. *Journal of Orthopaedic & Sports Physical Therapy*, 48(8), 622–629.

第六章　步骤留下，岔路删除

〔一〕 National Center for O*NET Development. O*NET 29.1 Database. https://www.onetcenter.org/database.html

〔二〕Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. Quarterly Journal of Economics, 118(4), 1279–1333.

〔三〕 Autor, D. H. (2015). Why Are There Still So Many Jobs? Journal of Economic Perspectives, 29(3), 3–30.

〔四〕Bainbridge, L. (1983). Ironies of Automation. Automatica, 19(6), 775–779.

〔五〕 Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. Human Factors, 39(2), 230–253.

〔六〕 Endsley, M. R. (2017). From Here to Autonomy: Lessons Learned from Human–Automation Research. Human Factors, 59(1), 5–27.

〔七〕Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The Retention of Manual Flying Skills in the Automated Cockpit. Human Factors, 56(8), 1506–1516.

〔八〕Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, Articles 12 and 14. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

〔九〕Iyengar, S. S., & Lepper, M. R. (2000). When Choice is Demotivating. Journal of Personality and Social Psychology, 79(6), 995–1006.

〔十〕Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press.

〔十一〕Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits. Psychological Inquiry, 11(4), 227–268.

〔十二〕Sen, A. (1999). Development as Freedom. Oxford University Press.

〔十三〕Braverman, H. (1974). Labor and Monopoly Capital. Monthly Review Press.

〔十四〕Zuboff, S. (1988). In the Age of the Smart Machine. Basic Books.

〔十五〕Scott, J. C. (1998). Seeing Like a State. Yale University Press.

〔十六〕Frankfurt, H. G. (1969). Alternate Possibilities and Moral Responsibility. Journal of Philosophy, 66(23), 829–839.

〔十七〕Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. Administrative Science Quarterly, 44(2), 350–383.

〔十八〕陈晓艳《接缝思维：不可通约处如何成为生产性的认知立足点》，SDE Universes 学员专栏。https://sdeuniverses.com/students/chen-xiaoyan/seam-thinking/

〔十九〕张琼《伦理空转项：中国伦理人类学田野采集的诊断》，SDE Universes 学员专栏。https://sdeuniverses.com/students/zhang-qiong/ethical-idling/

〔二十〕高鹏《使人成为主体的规则：论禁止性规则与期待型规则所安装的两种主体结构》，SDE Universes 学员专栏。https://sdeuniverses.com/students/gao-peng/rules-that-make-subjects/

第七章　类内距

Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(3), 261–290.

Chen, M., & Wang, W. S.-Y. (1975). Sound change: Actuation and implementation. *Language*, 51(2), 255–281.

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. *Annual Review of Sociology*, 24, 313–343.

Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Bobbs-Merrill.

Kiparsky, P. (2016). Labov, sound change, and phonological theory. *Journal of Sociolinguistics*, 20(4), 464–488.

Krishnamurti, Bh. (1978). Areal and lexical diffusion of sound change: Evidence from Dravidian. *Language*, 54(1), 1–20.

Labov, W. (1981). Resolving the Neogrammarian controversy. *Language*, 57(2), 267–308.

Labov, W. (1994). *Principles of Linguistic Change, Volume 1: Internal Factors*. Blackwell.

Mansoor, E. M. (1961). Selective assembly — its analysis and applications. *International Journal of Production Research*, 1(1), 13–24.

Wang, W. S.-Y. (1969). Competing changes as a cause of residue. *Language*, 45(1), 9–25.

Colosimo, B. M., et al. (2018). An economic decision model for selective assembly. *International Journal of Production Economics*.（选择装配与传统互换装配的成本比较模型）

第八章　记零位

· Burroughs Wellcome Co. v. Barr Laboratories, Inc., 40 F.3d 1223 (Fed. Cir. 1994).（构思作为发明完成的时点；付诸实施与常规协助不构成发明人资格） law.justia.com

· Rembrandt Research Project. *A Corpus of Rembrandt Paintings*, Vols. I–VI (1982–2015). Springer / Stichting Foundation Rembrandt Research Project.（归属的系统性重审与自我修订） rembrandtresearchproject.org

· Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 592-604.

· Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379-424.

· Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology*, 93(1), 59-69.

· Aviation Safety Reporting System (NASA/FAA).（不追责的自愿报告制度：给报告记账而非给发现记功） asrs.arc.nasa.gov

· CVE Program, MITRE.（可复现缺陷的编号与归属体系，以及它的灌水筛除成本） cve.org

第九章　没有人要求的时候

1. Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall.

2. Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The retention of manual flying skills in the automated cockpit. *Human Factors*, 56(8), 1506–1516. https://doi.org/10.1177/0018720814535628

3. Crisanti, A., & Ritort, F. (2003). Violation of the fluctuation–dissipation theorem in glassy systems: basic notions and the numerical evidence. *Journal of Physics A: Mathematical and General*, 36(21), R181–R290. https://doi.org/10.1088/0305-4470/36/21/201

4. Cyert, R. M., & March, J. G. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

5. Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *PNAS*, 98(24), 13763–13768.

6. Friston, K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, 3, 130.

7. Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Farnham: Ashgate.

8. Jabbari-Farouji, S., Mizuno, D., Atakhorrami, M., MacKintosh, F. C., Schmidt, C. F., Eiser, E., Wegdam, G. H., & Bonn, D. (2008). Effective temperatures from the fluctuation–dissipation measurements in soft glassy materials. *Europhysics Letters*, 84(2), 20006. https://doi.org/10.1209/0295-5075/84/20006

9. Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.

10. Kubo, R. (1966). The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1), 255–284. https://doi.org/10.1088/0034-4885/29/1/306

11. March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87.

12. Seth, A. K., Millidge, B., Buckley, C. L., & Tschantz, A. (2020). Curious inferences: Reply to Sun and Firestone on the dark room problem. *Trends in Cognitive Sciences*, 24(9), 681–683.

13. Sun, Z., & Firestone, C. (2020). The dark room problem. *Trends in Cognitive Sciences*, 24(5), 346–348. https://doi.org/10.1016/j.tics.2020.02.006

14. Wagner, A. (2011). *The Origins of Evolutionary Innovations*. Oxford: Oxford University Press.

15. Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, 17(2), 312–321. https://doi.org/10.1038/nn.3616

第一家主张驱动一切改变的只有预测与实际之间的落差，落差归零即模型收敛、不再改变是正确的；第二家主张可被推动的程度由响应函数单独决定，不施加扰动就无从判断；第三家主张能不能承受未来的要求由当下余量与所处工况单独决定，且不该为了知道好不好而去动它。三条在同一个具体处境里给出互相取消的处置，且各自的日课正是另一家诊断出的病：压低误差是第一家的日课，而第三家的余量学说指出误差被压到零的系统其退化不产生任何信号；施加扰动测响应是第二家的日课，而非破坏性检测这一整套东西存在的理由，正是「为了知道它好不好而去动它」本身不可接受；留足余量是第三家的日课，而第一家会说余量正是让落差归零的装置——余量越足，更新越少。　三家均为站外的公开文献，链接直达原始出处，可自行核对。

返回每日必读 →

第十章　选择听谁，本身就是一次判断

· Belkoniene, M. (2025). Moral understanding: From virtue to knowledge. *Noûs*. https://onlinelibrary.wiley.com/doi/full/10.1111/nous.12508

· Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1), 85–110.

· Green, A. (2025). Epistemic authenticity, skilled interactions, and strategic parallel processing: A dialogue with Levy and Kawall. *Social Epistemology Review and Reply Collective*. https://social-epistemology.com/2025/09/15/epistemic-authenticity-skilled-interactions-and-strategic-parallel-processing-a-dialogue-with-levy-and-kawall-adam-green/

· Hills, A. (2009). Moral testimony and moral epistemology. *Ethics*, 120(1), 94–127.

· Hopkins, R. (2007). What is wrong with moral testimony? *Philosophy and Phenomenological Research*, 74(3), 611–634.

· Kawall, J. (2025). Finding work for epistemic autonomy: A reply to Green and to Levy. *Social Epistemology Review and Reply Collective*. https://social-epistemology.com/2025/05/07/finding-work-for-epistemic-autonomy-a-reply-to-green-and-to-levy-jason-kawall/

· Levy, N. (2024). Against intellectual autonomy: Social animals need social virtues. In *The Philosophy of Epistemic Autonomy* (special issue). *Social Epistemology*. https://www.tandfonline.com/doi/full/10.1080/02691728.2024.2335623

· McGrath, S. (2011). Skepticism about moral expertise as a puzzle for moral realism. *The Journal of Philosophy*, 108(3), 111–137.

· Nguyen, C. T. (2020). Autonomy and aesthetic engagement. *Mind*, 129(516), 1127–1156.

· Robson, J. (2022). *Aesthetic Testimony: An Optimistic Approach*. Oxford University Press. 评论见 https://ndpr.nd.edu/reviews/aesthetic-testimony-an-optimistic-approach/

· Robson, J., & Wallbank, R. (2025). Aesthetic testimony. In *The Stanford Encyclopedia of Philosophy* (Fall 2025 ed.). https://plato.stanford.edu/archives/fall2025/entries/aesthetic-testimony/

· Wollheim, R. (1980). *Art and Its Objects* (2nd ed.). Cambridge University Press.

· 关于证言在发展理解中作用的悲观论新论证，见 *Australasian Journal of Philosophy*, 103(3), 2025. https://www.tandfonline.com/doi/full/10.1080/00048402.2024.2421277

若知识论一侧正确（认知自主根本不是一种德性，应由智性互赖取代），则美学的熟识原则与伦理的道德证言悲观论是把一个误导性理想抬成了规范；若伦理一侧正确（以证言了结一个问题会使人此后更少去寻求理解），则互赖的处方正在系统性地生产不再求解的人；若美学

乐观论一侧正确（那条禁令的来源不是认识论而是社会表演，且熟识原则今日已近乎被公认失败），则前两家都把一条表演规范误诊成了认识论规范。　三家均为站外的公开文献，链接直达原始出处，可自行核对。

附录一　十项读数速查卡

每卡四行：问什么 · 从哪取数 · 什么算异常 · 最容易的作弊。可脱离本书使用。

卡一　复推率 ρ（第一章）问：换一个表面不同、结构相同的情境，他会不会推出同一条？至少测两次。取数：一次带教或教研，一小时。异常：说得清楚而推不出来。注意——说得清楚是复述的证据，不是复推的证据。作弊：只测一次；一次推对可能只是记住了。

卡二　内外表述落差（第二章）问：这个领域对同行说的，与对使用者说的，是不是同一件事？取数：专业文献一份，使用手册或对外说明一份。异常：专业内部承认某样东西是构造的，使用现场当作给定的。作弊：把对外说明也写得很专业，落差消失在文本上而非事实上。

卡三　影子正本占比（第三章）问：分头问持有者两句——「这几份应该一样吗」「如果不一样，哪一份算数」。取数：四五个人，分开问，半小时。记原话，不要归纳。异常：第一句多数答应该，第二句答案分歧，或出现「应该都一样吧」这种反问。作弊：当众问。当众问会收敛到一个答案，而分歧本身才是读数。

卡四　借清率 b（第四章）问：目标处清晰度上升的同时，相邻处下降了多少？取数：相邻题域的配对测试；「近」的口径必须先声明。异常：只测到上升，没有人去测下降——多数场合的实际状况。作弊：把「近」定义得足够远，坑就落在测量范围之外。

卡五　剥夺落差 Δ（第五章）问：临时封掉替代路径，结果跌多少、多久回得来？取数：需要停机，全表中成本最高。应在可及性还低时先测一次留基线。异常：跌幅大且恢复慢。注意第五章实测：恢复不是更快，是更慢。作弊：封路时间太短，只测到一次性适应，测不到路径退化。

卡六　带责分叉数（第六章）问：单位工时内，有几处满足三条——路真的分开、他知道并被允许走、选错的后果落在他本人身上且不至于将他废掉？取数：岗位描述数据库现成可算；O*NET 一类公开库即可。异常：工作量与步骤数不变或上升，而分叉数下降。作弊：把「可以提意见」算作分叉。意见不被采纳且无后果的，不算。

卡七　类内距（第七章）问：合格样本在使用者关心的维度上有多离散？取数：返工率、胶合代码量、上手时间、排练次数——多数已有记录，只是从未放在同一张表上。异常：规范条款数上升而适配工作量不降。作弊：把「使用者关心的维度」改成规范里已经写了的维度。

卡八　记零位（第八章）问：「发现了一个错误」在正式记录里记在谁名下、算多少？三追问：能跟人走吗、进晋升材料吗、能被第三方核对吗？取数：一份晋升材料模板即可判。异

常：三追问全否。作弊：给「发现」记功。记录的将不再是核验，而是关于核验的声称——去给「报告」记账。

卡九　静息变异率 v（第九章）问：无任务、无考核、无故障的时段里，自发做出的与默认做法不同的事占多少？取数：事后清点，无法预约。异常：为零。注意为零通常被记作可靠性良好。作弊：划出两小时叫自由探索，有场地、有交付物、有考核——按定义，那两小时里测到的一条也不算数。

卡十　鉴别独立性（第十章）问：我凭什么认为他在这件事上靠得住——而这个理由，一个完全不具备该领域判断力的人能不能用？取数：一次自问，零成本。粗筛更快：这个领域里，有没有人因为判断错了而承担过可被外人读出的后果？异常：理由落在第二层（「大家都说他判断力好」）。这一层最危险，因为它在形式上模仿了第一层。作弊：无。这一项没有可作弊之处，只有「对自己所处领域的独立性判断几乎总比实际乐观一格」这一条系统偏差。

附录二　记号与七条关系式

记号

记号	名称	定义	章
a	可及性	面对一个具体任务，一份合格产物由外部即时供给的概率	枢纽章
λ	借道率	达成结果时走替代路径而不自己重推的比例	第五章
k	可塑性系数	使用依赖的可塑性强度，k > 0	第五章
ρ	复推率	换表面不同、结构相同的情境，推出同一条的可靠性	第一章
ρ_max	复推上限	完全不借道条件下可达到的复推可靠性	枢纽章
y	产物型读数	任何以产物为形式的读数	第五章
Δ	剥夺落差	y(封路) − y(不封路)	第五章
b	借清率	相邻处清晰度下降量 ÷ 目标处上升量	第四章
κ	锐度	峰高 ÷ 邻坑深	枢纽章
v	静息变异率	无任务、无考核、无故障时段里自发偏离默认做法的比例	第九章
C	覆盖型量表得分	分格量表上及格格数	枢纽章

七条关系式

R1　$\lambda(t+1) = \lambda(t) + k \cdot a \cdot (1 - \lambda(t)) \implies \lambda(t) = 1 - (1 - \lambda_0)(1 - k\,a)^{t} \to 1$　只要 a > 0，λ 单调升至一；a 只决定快慢，不决定终点。

R2　$\rho = \rho_{max} \cdot (1 - \lambda) \implies \rho \to 0$　复推率不是被消耗的存量，是每次使用时被重新生产的能力；停止生产不等于折旧。

R3　$\partial y/\partial \lambda = 0$　产物型读数与借道率正交。实测支持两处：术后单腿跳远距离比 97% 而膝关节做功比 69%；O*NET n=879 自动化与决策频次 r=0.042、与决策自由度 r=−0.199。

R4　$\Delta \equiv y(\text{封路}) - y(\text{不封路})$，且 $\partial(\text{成本 of } \Delta)/\partial a > 0$　唯一读法，且价格与诊断价值同向随 a 上升。实测：随机化压掉 Δ 的 92%，代价 y 掉 13 个百分点，仅当替代路径失效概率 > 46% 时划算。

R5　$b \propto (1-\lambda) \implies b \to 0$，$\kappa$ = 峰高/邻坑深 在极限处无定义　命名：均浅。κ 不是低，是不存在——这正是它读不出来的原因。

R6　C(深学者) − C(均浅者) = 0 − #{坑落在格上} ⩽ 0　峰在封顶处失效而坑照常扣分。覆盖型量表给损失方加分。

R7　鉴别独立性低的领域里，依赖的安全性取决于 ρ；而依赖使 λ 升、ρ 降。　自锁，且不产生任何事件。

两个算例

算例一：适度使用不是一个状态。取 $k = 0.1$，$\lambda_0 = 0$。$a = 0.9$ 时，λ 到 0.95 需 $t \approx 30$；$a = 0.3$ 时，$t \approx 98$。把可及性压到三分之一，只把到达同一终点的时间拉长约三点三倍。在一段三十年的职业生涯里，这个差别不改变结局。要改变结局，需要的不是降低 a，是存在一个把 λ 往回拉的力；R1 里没有这一项，因为现行读数里没有它。

算例二：覆盖型量表上的净惩罚。设一份量表有 20 格。均浅者每格恰好及格，得 20。深学者在 3 格有峰（各得 1 格，与及格者相同）、在 2 格有坑（低于及格线），得 18。差为 −2。峰的高度对总分毫无贡献，坑的深度也不影响扣分幅度——只要落进格里，扣一格。这解释了为什么每一次「提高考试覆盖面」的改革都在无意中加大对深学的惩罚：格数越多，坑被格子接住的概率越大。

附录三　十章命名一览

十个对象，十个三学科组，以及每一个命名所解决的问题。

章	命名的对象	三学科组	它解决的问题
一	复推率	教育心理学 × 运动学习 × 组织例程研究	「学会了没有」在规则的前后差之外，还能读什么
二	追认（下游需求对本体论立场的）	圣经神学 × 物理学 × 文学	一个够不着的原物，三种处置为什么被表述为三项发现
三	影子正本	软件工程 × 数字保存与条约法 × 演化遗传学	副本的危险不由份数决定，由什么决定
四	借清率	机器学习 × 免疫学 × 神经科学	学会一件事时，对没在学的东西发生了什么
五	借道	运动控制 × 古典舞技术与舞蹈医学 × 学习心理学	结果等价的路径替换，为什么读数与感觉都不报警
六	带责分叉	认知教育 × 伦理人类学 × 法哲学	被取消的选择为什么在全部现行读数上不可见
七	类内距	语言学 × 工程学 × 艺术学	规范写得更细，为什么不降低适配工作量
八	记零位	学习科学 × 专利发明人认定 × 作坊归属研究	一个共同体最后擅长什么，由什么决定
九	静息变异	认知科学 × 物理学 × 工程学	「没有问题」与「已经不会变了」怎么分开
十	鉴别资质的独立性	美学 × 伦理学 × 知识论	找一个更懂的人问，什么时候省去了判断
枢纽	均浅 · 锐度 κ	（从上述定义推出，不引入新学科）	借清停摆之后，知识分布变成什么形状

十组学科无一重复。 这一点是刻意的，理由见前言第三节：一本由同一组学科撞十次的书，它的书一级结论只是那一次碰撞的重复；十组各撞一次，才使「十章说的是同一件事」成

为一条可以被推翻的主张，而不是一个设计上的必然。合章第七节第五条正是这条主张的判错条件。

后　记

一

这本书有三处是我在写作过程中被迫改了主意的地方，写下来，因为它们比结论更能说明这套做法的实际状况。

第一处是书名。 最初拟的几个都是判断句——《知识不是存量》一类。改成一个时间状语，是因为写到枢纽章时发现全书说的其实不是一个判断，是一个条件：当一份合格产物随时可以从外部取到，此后发生的事。判断可以被同意或反对，条件只能被核对。书名改了，全书的自我要求也跟着变严了一层。

第二处是选篇。 原定的十篇里有一篇是关于「课堂自己造出来、随后成为自己条件的那类东西」的。查证时发现它已经进了另一本书。换上来的是关于类内距的那一篇，而这一换意外地补上了一个缺口：原来的十篇全部在人的一侧，换过之后第三编有了产物的一侧，合章第四处锁才有可能成立。一次因为撞车而被迫做的替换，改善了结构。 这件事我不打算解释成运气之外的东西。

第三处是枢纽章第七节。 我原来写的是「覆盖型量表读不出这场损失」。推导做完之后发现不止如此——峰在封顶处失效而坑照常扣分，净效应是深学者得分更低。读不出是中性的，加分是有方向的。 这两句话之间隔着一条我原本没有预料到的推论，而它是全书唯一一条我真心希望被人拿去推翻的。

二

第二件要说的，是这本书自己的处境。

书里反复说，一件事若不进入产物，就不会被记录。而这本书本身是一件产物：它有页数、有字数、有出版信息，可以被清点。它所主张的那些不入账的东西——重推的那一段、被搬空的近邻、没有走的岔路、没有发生的事故——一样也不在这本书里。读者拿到的是十章文本，而不是产生这十章的那些过程。

这不是一句谦辞。按本书自己的框架，这本书对读者而言正是一份「随时可得的合格产物」。读它省下的那一段，恰好是它要谈的那一段。唯一能避开这个循环的读法，是把结语那张表拿去自己测一遍——不是测别人，是测自己。附录一那十张卡是为这个用途写的。

三

第三件，把「这张表进考核那天就开始失效」这句话在书里写第三遍。

结语写过一次，附录一每张卡的最后一行写过一次。这里再写一次，因为在我的经验里，一本书里被引用最多的部分总是那张表，而被略过最多的总是关于那张表怎么会坏掉的说明。

十项读数测的都是「没有发生的那一半」。任何把没有发生的事变成考核指标的做法，都会立刻制造出它的表演版本——第九章那两小时的自由探索、第八章那份关于核验的声称、第三章那个只补了一半的对齐。这本书没有解决这个问题。它给出的建议只有一句：这张表适合自查，不适合考核别人。

四

最后。

我今年五十五岁，做过一段计算数学，教过书，现在做的事与那两样都不太像。前言里那位年轻同事说「昨天那个我查过了」的时候，我第一反应是想纠正他。这本书是那个反应被推迟了一年之后的结果——推迟的这一年里，我大部分时间在做的事是搞清楚我凭什么纠正他。

到现在也没有完全搞清楚。书里给出的是一张可以测的表，不是一个可以下的判断。这大概是这类问题上，一个人能做到的全部。

王德生二〇二六年八月于新加坡

十　句

这一辑怎么用　下面十句取自全书，每句附二百字左右的解释。它们不是摘要——摘要要求覆盖，这十句只求把最容易被读软的地方钉住。每句后面那段解释，主要用来说明这句话不是在说什么，因为本书被误读的方式比被反对的方式多。

一

知识不是被持有的存量，是每次被重新推出来的事件。

这是全书的第一块地基，来自第一章。它不是一个关于记忆的心理学主张，也不是"知识需要不断更新"的另一种说法——那句话仍然把知识当成存货，只是承认存货会过期。这里说的更硬：从来就没有那个存货。 教育心理学、运动学习与组织例程研究在"学会了没有"上互相取消，而三家共有的前提是存在一条被放在某处、可以在两个时点之间取出来比对的规则。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。接受这一条，"他拥有这方面的知识"这句话就指不到任何东西——不是因为它不礼貌，是因为它没有指称对象。全书其余部分都是这一条的后果。

二

贬值的不是知识，是「拥有」这本账。

这句话把"AI 来了知识还有没有价值"这个问法整个换掉。原问法预设知识是库存、可及性是一次降价，于是所有回答都在争比例：哪些货还值钱、哪些不值了。本书说争比例没有意义，因为账本记错了对象。一份合格产物随时可取，改变的不是知识的价格，是重新推那一段还发不发生；而那一段从来不在账上——没有哪本账为"这个人这次是自己推出来的"设一栏，也没有哪份验收标准区分一份产物是被推出来的还是被取来的。所以这不是一场贬值，是一次记账口径的失效：账面上一栏也不会少，而下一次推所需要的东西已经不在了。

三

凡进入产物的都留下，凡不进入产物的都在退，而只有前者被记录。

这是十章会合的那一句，也是全书最容易被读成套话的一句。它之所以不是套话，因为它有两组互不相干的实测支撑：术后运动员单腿跳远的距离比是百分之九十七，而同一次测量里膝关节做功比是百分之六十九；八百七十九个职业的岗位描述上，自动化程度与决策频次相关 0.042，与决策自由度相关 −0.199。一条测的是人的膝盖，一条测的是岗位说明书，两个学科互不知情，而它们是同一句话的两次显影。这句话的力量不在它听起来对，在于它已经被两处不该相关的数据分别撞见过。

四

「适度使用」不是一个状态，是一段速度。

这是枢纽章那条递推式的日常译法。设可及性为一份合格产物由外部即时供给的概率，借道率的演化是 λ(t+1) = λ(t) + k·a·(1 − λ(t))，解出来 λ 单调升至一：只要可及性严格大于零，终点就是同一个，可及性只决定到得多快。把它从零点九压到零点三，到达同一水平所需的轮数从约三十变成约九十八——在一段三十年的职业生涯尺度上，这个差别不改变结局。所以"控制使用量"这类处方在形式上是无效的：要让它停在中途，必须存在一个把借道率往回拉的力，而本书的主张是现行记录方式里这个力一项也没有。有人指出代码评审就是这个力；那么命题就该弱化为"这个力正在消失"——弱得多，也诚实得多。

五

复推率不是被消耗的存量，是每次使用时被重新生产的能力；停止生产它，它不是慢慢变少，是不再被生产。

这一句是第一句在动力学上的兑现，也是本书与"技能退化"这个通俗说法的分界线。退化预设一条折旧曲线，可以谈半衰期、谈保养、谈定期复训——所有这些安排都假定那样东西存在着并在缓慢损耗。这里没有折旧，只有停产。一项每次使用时被重新生产的能力，在停止使用的那一刻不是开始变少，而是不再被生产；此后测到的任何"还剩多少"，测的都是别的东西。这个差别决定了处方的形状：对折旧品该做的是保养，对停产品唯一能做的是重新开工，而重新开工需要一个理由，而理由在所有读数都是绿的时候不会出现。

六

y 高与 ρ 低不是一对需要权衡的目标，它们是同一个 λ 的两种读法。

这是全书最反直觉的一处，也最容易被读成"要在效率和能力之间取平衡"。不是。权衡的意思是两个量此消彼长、可以调节配比；这里说的是两个量由同一个参数生成，因而不存在配比可调。产物型读数对借道率的偏导恒为零——一个能区分路径的指标就不再是结果指标，这是定义上的事，不是精度上的事。于是在借道率从零走到一的全过程里，只看产物的组织看到的是一条持续向好的曲线。它不是疏于监测；它监测的那一样，正是这件事唯一不会留下痕迹的地方。 想在两者间"取平衡"的人，手上没有可调的旋钮。

七

唯一能读出这件事的动作，恰恰在最需要它的时候最贵。

剥夺落差是唯一读法：临时封掉替代路径，量结果跌多少、多久回得来。而它的价格随可及性单调上升——可用的替代通道越多，封路要挡的口子越多；借道率越接近一，封路后跌得越深，因为不借那条路已经很久没有被走过。诊断价值与价格同向上升，于是这项测量总是在"还不必做"的时候便宜、在"必须做"的时候贵得做不起。由此得到一条不受欢迎的实践含义：剥夺落差应当在可及性还低的时候先测一次留作基线。 而这件事在任何组织里都不会被提出来——提出它的那一刻，所有读数都还是绿的。本书三次遇到这个形状（剥夺落差、静息变异、锐度），三次都是在测"没有发生的那一半"。

八

在任何覆盖型量表上，衡量知识的通行办法不只是读不出这场损失，它给损失的那一方加分。

这是全书的承重判断，也是最便宜就能被推翻的一条。把一个领域切成若干格、每格一条及格线、总分是及格格数——几乎所有考试、认证、能力矩阵、岗位胜任力模型都是这个形状。在这样一份量表上，峰在封顶处失效（一格及格就是满格，深学者在有峰的那一格与刚好及格的人同分），而坑照常扣分（被搬空的近邻若落在格上，扣一格）。净效应不大于零。这解释了三件此前各自被单独解释的事：为什么随时可得的答案在标准化评估上表现为进步；为什么"什么都懂一点"在选拔中长期占优而在使用中长期落空；为什么每一次"提高考试覆盖面"的改革都在无意中加大对深学的惩罚——格数越多，坑被格子接住的概率越大。本书没有做这次检验，也不打算把没做的事说成做过的事。

九

均浅者的锐度不是低，是没有定义。

这一句必须与"样样通样样松"分开，否则全书唯一造出来的那样东西就只是一句老话的新名字。老话按能力的水平分类——他哪一样都不精；均浅按分布的形状分类——他哪一处也没有被搬运过，因而既没有峰，也没有坑。一位在三个领域各练五年、样样中上的人，按老话是典型的样样通，按本书有三个峰三片坑、锐度有定义且不低：两说在同一个人身上给出相反的判定。 更要紧的是，均浅不是贬语——在覆盖面上均浅者往往知道得更多，在覆盖型量表上得分更高。锐度在他身上是零除以零：这不是测量精度不够，是那两个量在那里不存在，因而一切"把它测得更准"的努力都无从下手。

十

> 在那件事发生之后，我们用来判断自己有没有损失的全部工具，都会告诉我们没有。

这是全书的最后一句，也是它唯一一次直接说出自己在担心什么。三类读数三种看不见的方式：产物类把借道、岔路与核验一律记零；效率类把它们记为改善——更省力、更少差错、更高合格率；能力类在覆盖型量表上把峰与坑一起抹平，读出来是覆盖面扩大。三类各自都在诚实地测它所定义的东西，问题不在测得不准，在这件事恰好落在三类读数的定义之外。所以这本书没有交出一个警报，只交出一张体检表；而它同时写明，这张表一旦进了考核就开始失效，因为它测的是"没有发生的那一半"，而任何把没有发生的事变成指标的做法都会立刻制造出它的表演版本。这张表适合自查，不适合考核别人。

十句之外必须补的一句

上面十句里，只有第三、第四、第七句有实测支撑，且都是别人为别的目的做的测量；第五、第六、第八、第九句是从定义推出来的。金句这种形式天然把推导与观测拉平——它们读起来一样确定。读到这一页为止，全书的真实证据状况是：七条关系里三条有实测，四条只有形式推导与逐条可判错的条件。 把这句放在十句之后，是为了不让这一辑成为全书最不诚实的部分。

封　底

答案随时可得之后

贬值的不是知识。

是「拥有知识」这本账。

—

一份合格产物随时可以从外部取到之后，被取消的不是知识，是重新推出知识的那一段；而那一段里长出来的全部东西，现行的记账法一栏也没有。

十个不同学科组撞出来的十条判断，说的是同一句话：

凡进入产物的都留下，凡不进入产物的都在退，而只有前者被记录。

—

术后单腿跳远的距离比是百分之九十七，同一次测量里膝关节做功比是百分之六十九。八百七十九个职业里，自动化程度与决策频次相关 0.042，与决策自由度相关 −0.199。

步骤留下，岔路删除。

—

本书唯一造出来的那样东西叫均浅：既无峰亦无坑的知识分布。它的锐度不是低，是没有定义。

由它推出的那条判断是：

在任何覆盖型量表上，衡量知识的通行办法会给损失的那一方加分。

一

书末给出一张十项读数的体检表，以及它进考核那天就开始失效的理由。

德麦国际出版社 · 专著第 81 号