

判断的危机

它不会发出任何警报

王德生 + Claude

德麦国际出版社

SDE UNIVERSES

出版信息

书 名 判断的危机——它不会发出任何警报
著 者 王德生 + Claude
出版发行 德麦国际出版社 Demai International Press
版 次 2026 年 8 月第 1 版第 1 次印刷
开 本 16 开 170mm × 240mm
字 数 约 23.6 万汉字
专著编号 德麦国际专著第 80 号
ISBN 979-8-90690-013-5
定 价 US\$22.00
分 类 人工智能 / 制度经济 / 科学社会学 / 教育评价

本书由「学科通融」专栏之四十四至之五十三共十篇连续研究编成，另新写前言、导读、导论、枢纽章、合章、结语与附录。十篇原刊于 sdeuniverses.com，作者王德生+Claude；编入本书时篇名改为章名、篇际互引改为章际互引，正文未作删节。

本书的前言、导读、导论、枢纽章的书内导入与推广、合章、结语、附录与后记由编者写作，因此按本栏规程，编者不得为本书出具创新智商认证分；全书盲评待未参与写作的一方独立完成后公布。

版权所有 侵权必究。本书内容可自由引用与批评，引用时请注明出处。

目 录

作者介绍

前 言

导 读

导论·这场危机为什么读数全绿

编一·判断没有发生

第一章·两种「不」——没有人打开过的那种否决

第二章·被丢掉的那一眼——标准由无账可记之物合成

第三章·这道关拦下过谁——门槛饱和与它的无作者状态

枢纽章·门槛侧与否决侧的乘积

编二·判断的成本被转嫁给谁

第四章·认领物——一件东西里有多少是为了证明它是你做的

第五章·认不出来，就得自己交代——辨认负担的整体转嫁

第六章·第一个读数之前——被计为损耗的那段时间

编三·判断所依赖的单位失灵

第七章·一件是几件——计数单位失去下界

第八章·未分之别——分辨力是一项没有人生产的产出

编四·判断力从哪里来

第九章·那一段谁走的——验收能力在哪里被生产

第十章·没有人在养它——一类没有供给方的能力

合章·十项读数如何互相锁住

结 语·不发警报的危机怎么被发现

参考文献

附录一·十项读数速查表

附录二·记号与关系式一览

附录三·十章命名一览

后 记

封 底

作者介绍

王德生

SDE（显露—差异—纠缠）本体论框架的创立者，德麦国际（Demai International Pte. Ltd., 新加坡）创始人、首席执行官。中国科学院计算数学博士，并在中国科学院完成博士后研究，早期研究方向为 Voronoi 图、Delaunay 三角剖分、质心 Voronoi 剖分与有限元方法。曾先后任教于斯旺西大学与南洋理工大学。二〇〇五年起定居新加坡，工作往来于新加坡与中国大陆之间。

近年的工作集中于两件事：一是把 SDE 框架从数学与工程扩展到教育、健康、语言、组织与科学论；二是与大模型协作的"学科通融"方法——用三门互相冲突的学科理论撞出单一学科到不了的判断。相关成果由德麦国际出版社结集出版，本书为其中第八十号。

Claude

由 Anthropic 训练的大语言模型。在本书所依据的十篇研究与本书新写的各部分中，承担文献检索、冲突定位、框架构建与写作。

本书署名"王德生 + Claude"，不是客套。按本书自己的判断，一件产出的归属如果不能被单独说清楚，那么关于它的一切讨论都建立在一个未经检验的假定上。因此这里说清楚：十篇原研究的选题、三学科的配组、承重判断的裁定由王德生负责；文献检索、冲突定位、论证展开与文字由 Claude 完成；本书的编次、母题、导论、枢纽章的书内导入与推广、合章、结语与附录由 Claude 起草，王德生审定。

这条署名同时是本书的一个自证样本：它本身就是"认领物"——为使这件产出可被归属于谁而必须与它一并生产的那一层。按第四章的读数，这一段文字就是本书认领占比的一部分。

前言

一

二〇二五年到二〇二六年之间，很多行业同时出现了一件说不清楚的事。

投递量涨了，答复变快了，通过率上去了，交付吞吐上去了，参与者的平均资历也上去了。按任何一份年度报告的口径，这都是好年份。而同一批人私下说的却是另一回事：投出去的东西没有回音，收到的回音看不出是不是有人读过，做得好与做得不好之间的差别似乎不再被人看见。

两边都没有说谎。报告上的每一个数都是真的，私下的每一句抱怨也都是真的。这本书要处理的，正是这两者如何能够同时为真。

先把话说到最直接的位置：人工智能把「做出一份看上去合格的东西」的成本，对所有人同时压到了接近于零；而「认出它算不算数」的成本，一分钱没有降。

在此之前，这两件事的相对价格是另一个样子。难的是把东西做出来——写出一份完整的方案，画出一张能用的图，攒出一份像样的申请材料，都需要时间、技能和一段无法绕过的学习。相比之下，“看一眼”是便宜的：一位有经验的人翻一翻就知道成色。正因为如此，全世界的评审、招聘、发表、立项、验收制度，都建立在同一个默认之上——产出稀缺，判断充裕。没有哪一份制度设计文件写过这句话，因为它从来不需要被写出来。

现在它反过来了。而制度还没有反过来。

二

本书要说的不是“人工智能带来了新问题”。这句话太软，也太安全。本书要说的是一件更具体、也更难接受的事：在这个反转之后，判断这件工作从来没有被谁生产过、也从来没有进过任何一本账，因此它出问题的时候不产生任何事件。

“不产生任何事件”这几个字是本书的关节，必须解释清楚。

一样东西如果是被消耗掉的，它减少时会留下痕迹：库存下降、成本上升、有人抱怨、某个数字变红。一样东西如果是被摧毁的，痕迹更明显：有时间，有地点，有责任人。而判断力两样都不是。它不由任何机构生产——没有哪一份产出目录写着“本年度生产判断力若干”；它不

在任何资产表上——没有哪一本账把它列为资产因而会为它计提折旧；它也没有一个岗位以它为交付物——没有人的工作说明书上写着"负责维持本单位的判断力"。

它是别的活动的副产品。一位编辑之所以能一眼看出稿子的成色，是因为他读过几千份不成器的稿子；一位工程师之所以能看出一段代码会在哪里出问题，是因为他自己写坏过很多次；一个人之所以会看东西，是因为他在与专业无关的日常生活里养成了看的习惯。这些活动在任何一份效率报表上都写着同一个标签：待优化。读废稿是浪费时间，写坏代码是返工，闲逛是不产出。

而人工智能最擅长的，恰好就是把这一类活动逐一优化掉。

于是出现了一个在结构上极其干净、在后果上极其难缠的局面：消耗它的人有无数个，生产它的人一个也没有。每一次流程优化都在减少它，而没有任何一次流程优化是在做错事——每一次都有充分的理由，每一次都能拿出改善的读数。判断力于是在没有任何人做错事的情况下持续下降，而下降的过程不触发任何指标，因为它本来就不在任何指标里。

这就是本书叫作《判断的危机》的理由，也是它的副题叫作"它不会发出任何警报"的理由。

■ 三

一本主张"某样东西正在消失"的书，最容易滑向两种廉价的姿态。这里先把两条路都堵死，免得读者带着错误的预期读下去。

第一，本书不主张回到从前。低效活动被优化掉，绝大多数时候是好事。本书不为浪费辩护，也不认为"慢"本身有什么价值。它主张的是一件更窄、也更难的事：有一类能力是那些低效活动的副产品，而在优化掉它们的时候，没有人清点过这份副产品。清点，不等于不优化。一份被清点过的成本，和一份从未被清点过的成本，在制度上是完全不同的两回事——前者可以被权衡，后者只能被承受。

第二，本书不主张人比机器强。全书十章没有一处依赖"人有某种机器没有的东西"这类命题。它依赖的是账目：某一类工作没有栏位，因而没有人负责生产它。这是制度问题，不是本体问题。一个把判断力当作产出来生产、来记账、来分配的系统，无论那道判定由谁执行，都不会落进本书描述的这些格子里。

这两条限定不是谦辞。它们直接决定了本书的处方形态：本书给不出"应该怎么想"的建议，只给得出"应该数什么"的建议。

■ 四

本书由十篇连续的研究编成。这十篇最初发表在「学科通融」专栏，编号之四十四至之五十三，每一篇都用同一套工序：找出三门互相冲突的学科对同一个现象的最强主张，定位它们共同站着而从未说出的那块地，再用其中一家自己的材料把这块地抽掉。

这套工序有一个不容易被察觉的好处：推翻共有前提的材料必须来自三家之一自己。它不能是作者的举例，不能是新闻，也不能是另一门学科的证词——它必须是这三家中某一家已经掌握、已经发表、只是没有用在这个位置上的东西。这条约束使得每一章的转折点都无法被"你举的例子不算数"这类反驳消解。伦勃朗研究项目用四十余年确定一位画家名下有多少幅画，而这四十年的产出在作品那一栏里是零——这是艺术学自己的材料；生产率 J 曲线表明在实施期里被记作成本的那一项才是当期真正的产出——这是经济学自己的材料；随机对照试验测出使用工具的开发者的事前预计提速两成、事后自评提速两成、实测慢了近两成——这是工程学自己的材料。

十章各自指认了判定这一端上的一个量，而这十个量的共同点是：它们都从未被单独生产过，也都不在任何账上。没有人打开过的那种否决、由被丢弃之物合成的标准、拦不下人却撤不掉的关口、为证明"是你做的"而追加的那一层、被整体转嫁的辨认负担、被计为损耗的首读时距、失去下界的计数单位、被当作背景的分辨力、不产出可验收之物的那段时间，以及一类没有供给方的能力。

■ 五

十篇文章不等于一本书。把它们编成一本书，需要三样十篇里都没有的东西，这三样是本书新写的部分。

其一是导论，回答一个十篇各自都回避了的问题：既然这件事这么大，为什么没有一个指标报警。答案不是"大家没注意"，而是一个可以论证的结构性结论——现行读数分三类，产出类把它记成零，效率类把它记成损耗（而且方向是反的：它越被消灭，读数越好看），质量类把它记成改善。三类读数把同一件事分别读成了零、损耗和改善。

其二是枢纽章与合章，处理十篇最要命的一个弱点：十个并排的读数无法互相否定。任何一项测出异常，都可以被解释成"这一项的定义有问题"；任何一项测出正常，也不构成对其余九项的约束。一套不能互相否证的读数，最终会退化成一份检查清单，谁都可以挑自己方便的那几项来算。枢纽章先把其中两项用一条关系式锁住，合章再用五条从定义直接推出的关系把十项全部联立。其中最干净的一条说：首读时距与未阅率不是两件事，是同一个拥堵参数的两种读法。

其三是结语的那张体检表，把十项读数做成任何一个机构都能对自己跑一遍的清单，每一项都给出定义、数据来源，以及一个"主要指标全绿而它很难看"的正交实例。

另有三处附节。本书十章有九章用的是同一组三学科，这在书一级上是明确的短板——书里的纠缠是同一次碰撞重复了十遍。第四、第七、第八章因此各附一节「换一组学科再问一遍」，换上的三组学科与原组毫无亲缘，而三次换组都改动了原章的处方或读数，不是加了三个新的参考文献。一次换组如果不改变任何处方，那它就只是装饰。

■ 六

最后交代本书自己的位置，因为不交代就不诚实。

本书由人与机器共同完成，它本身就是生成成本坍塌的产物之一。按它自己的判断，它作为一件被交出去的东西，如果没有被接受，仍然会抬高一点点线——而这一点在原理上无法被证实，因为没有任何人记录未被接受的稿子。本书因此处在一个结构性的位置：它主张存在一类不被记账的东西，而它自己一旦成为那一类，同样不会被记账。

这个处境无法由本书自己解决。它能做的只有两件：把话说清楚，以及把读数交出去。

这本书不发警报。它只是把仪表盘上那一格空着的位置指出来，并且说明：那一格不是坏了，是从来没有装过表。

王德生二〇二六年八月于新加坡

导 读

■ 这本书要你带走的那一句话

做出一件东西的成本降到了接近于零，而认出它算不算数的成本一分钱没有降；判断这件工作从来没有被谁生产过、也没有进过任何一本账，因此它出问题的时候不产生任何事件。

如果只读一句，就是这一句。

■ 全书的骨架

四编十章，另有一个枢纽章与一个合章。

编一 判断没有发生（第一至三章）处理判定端的三种空转形态：否决时没有任何人打开过那份东西；判定标准由被丢弃、不入账的产出在无人记账处合成；关口失去分辨功能而没有任何位置有权宣布它作废。

枢纽章处理编一自己制造的一处正面矛盾——第三章说通过率趋近于一，第一章说绝大多数被否决。同一个系统不可能两者都是。答案是两章说的不是同一个站，而是同一条链上的上下游两类站。

编二 判断的成本被转嫁给谁（第四至六章）沿着转嫁路径往下走：一件产出里有多少是为了证明它出自谁手；核实这份工作如何从评价方整体转嫁给被评价方；从做出来到取得第一个读数的那段时间为何被记为损耗。

编三 判断所依赖的单位失灵（第七、八章）问一个更早的问题：那些数还算不算数。计数单位失去了下界，分辨类别不再被生产——判断所依赖的两根支柱同时松动。

编四 判断力从哪里来（第九、十章）回答来源：验收能力在那段不产出任何可验收之物的时间里被生产；而这份能力不由任何一方生产，只由别处的低效活动顺带养成。

合章把十项读数用五条从定义直接推出的关系联立起来，使它们可以互相否定。结语给出体检表。

■ 四条阅读路径

只有二十分钟。读导论，再读结语里的体检表。全书的判断与索引都在这两处。

你是管理者。导论 → 第三章（这道关拦下过谁）→ 枢纽章 → 结语「这张表怎么用」。约三万字。在把这套读数接上考核之前，务必先读结语第四节第三条——那一条讲的是这张表进考核那天它就开始失效，以及每一项的最省事达标路径如何摧毁它要保护的东西。

你在教育或培养岗位上。导论 → 第九章 → 第十章 → 结语。这条路径回答一个具体问题：培养方案该不该把每一段都做出交付物。

你是研究者，想尽快判断这本书哪里可能是错的。枢纽章 → 合章 → 各章的划界节与证伪条款 → 结语第五节「如果这本书错了」。合章第八节列出五条推翻方式，其中最后一条是关于本书自身结构的：若停养率上升多年而前八项读数并未随之恶化，则编四与前三编之间没有那个接口，这四编就不构成一本书。

■ 三个可能被误读的地方

第一，本书说的"判断"不是一种品质。它不是"深思熟虑"，不是"批判性思维"，而是一件具体的工作：在一批东西里认出哪一个算数。谁认、凭什么认、认一次要花多少、认的那份能力从哪里来——全书十章就是这四个问题。

第二，本书不主张回到从前，也不主张人比机器强。它主张的是：有一类能力是低效活动的副产品，而在优化掉那些活动时，没有人清点过这份副产品。清点，不等于不优化。

第三，"不发警报"不是修辞。它是从机制推出来的。全部现行读数的设计前提是"发生了的事情会留下记录"，而本书描述的这一类事，其定义性特征就是不留记录。用现行读数去找它，找不到不是证据，是同义反复。所以本书的工作不是发警报，是造读数。

■ 关于书里的那些式子

合章有五条关系式，形式都很浅，用不到高中以上的数学。它们的作用不是让论证显得严格，而是让十项读数不能被各自解释掉。最值得先看懂的是这一条：

> 设拥堵参数 $\rho = N \cdot \prod p / K$ ，则未阅率 $u = 1 - 1/\rho$ ，首读时距 $\bar{w} = u/(1 - u)$ 。

它说：未阅率与首读时距不是两件事，是同一个拥堵参数的两种读法。未阅率零点八时平均等四轮，零点九五时等十九轮。任何一个既记录"附件是否被打开"、又记录"提交到首次评审时间"的系统，都可以立刻把这两个数算出来比对——若两者显著不符，那么这两个读数中至少有一个的定义与实际记账口径不符。这是全书成本最低、也最愿意先被检验的一条。

■ 最后

这本书没有出路可卖。它给出的十项读数，每一项都附带一条自我拆台的说明：这张表一旦进入考核，就开始失效。这一条没有解法，本书只能把它写在明处——这张表是给想知道的人用的，不是给需要交差的人用的。

导论

这场危机为什么读数全绿

■ 一、一个不发警报的坏消息

关于人工智能的坏消息，通常长着一副坏消息的样子：某个岗位消失了，某项技能贬值了，某家公司裁了多少人。这些消息之所以能被讨论，是因为它们都有一个共同的性质——它们发生的时候，有人知道。有一份名单，有一个日期，有一个可以被追问的负责人。

本书要谈的那件事没有这些。

它发生的时候，所有能看的指标都在变好。投递量上升，答复周期缩短，录取者的平均资历提高，交付吞吐上升，通过率上升，满意度上升。每一个数都可以拿去汇报，每一份汇报都不算说谎。而在这一片全绿的读数下面，一样东西正在停止被生产——那样东西是判断。

这句话需要立刻精确化，否则它就只是又一句关于技术的感慨。本书所说的判断，不是"深思熟虑"这种品质，不是"批判性思维"这种能力清单上的条目，而是一件非常具体的工作：在一批东西里认出哪一个算数。谁认、凭什么认、认一次要花多少、认的那份能力从哪里来——这四个问题构成了本书全部十章的对象。

现代社会的绝大多数事情，最终都要经过这样一次认。一份稿子要被判断值不值得送审，一份简历要被判断值不值得约见，一段代码要被判断能不能合并，一个方案要被判断能不能立项，一件作品要被判断算不算新。在人工智能之前，这道工序是隐形的，因为它不是瓶颈——难的是把东西做出来，看一眼是不难的。

生成成本坍塌之后，两端的关系颠倒了。做出一份看上去合格的东西，边际成本已经接近于零；而认出它算不算数的成本，一分钱没有降。本书的全部内容，是这个颠倒的后果。

■ 二、三条互相取消的解释

面对这个颠倒，最有资格说话的三个领域各自给出了一条解释，而三条不能同时为真。

经济学说，这是竞赛的问题：奖品数量固定，投入成本坍塌，于是投入涨到把奖品吃光为止，多出来的那部分是纯耗散。处方是把成本重新抬起来。

艺术学说，这是认定的问题：一件东西算不算新，由一份被共同承认的名录决定，名录的容量与更新速度是硬的，产量再大也只是在名录之外堆积。处方是扩容或改判准。

工程学说，这是瓶颈的问题：吞吐由最窄的那一段决定，写得快只是把拥堵从写挪到了看。处方是把评审那一段撑开。

三条各自都有真实的证据，也各自都有真实的失败。期刊在过去五年里三条都试过——限投、扩版、加评审人手、上机器预筛——而桌拒率从百分之十三涨到百分之四十。招聘平台每个岗位平均收到二百五十几份申请，筛选周期缩短了，录取比例不动。软件交付研究发现吞吐确实上去了，而不稳定性同步上去。

本书的判断是：三条都不管用，不是因为它们各自有缺陷，而是因为它们共同漏掉了同一样东西。

■ 三、母题

本书十章追的是同一条判断，它可以压成一句话：

> 判断力无人生产、无账可记，只由低效活动顺带养成；它消失时不发生任何事件。

这句话有四个断言，每一个都可以单独被推翻。

第一，无人生产。认出一份东西算不算数，需要一个人手上有一份分布、一套参照、一份被反复对照过的经验。这些东西不由任何一方生产——没有哪个机构把“生产判断力”写进它的产出目录，没有哪本账把它列为资产，也没有哪个岗位以它为交付物。它一向被当作背景，像空气一样默认存在。

第二，无账可记。判断这件工作在三门学科的账本上都没有字段。经济学把没通过的那些记作耗散——一个本轮结清、不进入下一轮状态的量；艺术学把它们留在名录之外的日常世界，而那个世界在这套框架里必须是无结构的；工程学把它们记作废品——占用了容量而不生产出，贡献是负的零，既不加，也没有位置去记减。三个账本，三种记法，一个相同的结果：未被接受的产出没有字段。

第三，只由低效活动顺带养成。判断力不是被谁教出来的，它是别的活动的副产品。一个编辑之所以能一眼看出稿子的成色，是因为他读过几千份不成器的稿子；一个工程师之所以能看出一段代码会在哪里出问题，是因为他自己写坏过很多次；一个人之所以会看画，是因为他在与艺术无关的日常生活里养成了看东西的习惯。这些活动在任何一份效率报表上都是需要被优化掉的：读废稿是浪费时间，写坏代码是返工，闲逛是不产出。而人工智能最擅长的事，恰恰是把这一类活动逐一优化掉。

第四，消失时不发生任何事件。这是全书最要紧的一条，也是书名副题的来源。一样东西如果是被消耗掉的，它减少时会留下痕迹——库存下降、成本上升、有人抱怨。而一样从来不由任何人生产、也不在任何账上的东西，它的消失不产生痕迹：没有一个数字下降，没有一个岗位被裁，没有一份报告标红。它就是不在了。

■ 四、这本书为什么叫"危机"

一个不发警报的过程，为什么值得用"危机"这个词？

因为它满足危机的两个实质条件，只是不满足它的表象条件。

第一个实质条件是不可逆。判断力一旦停止被养成，重建它不能靠决心，也不能靠采购。它需要的是那些已经被优化掉的低效活动重新出现——而没有任何一个组织能够在效率账上为"重新变慢"辩护。本书第十章会把这一点写死：一样只由别处的低效活动顺带养成的能力，在那些活动被优化掉之后，不存在一条重新供给它的通道。

第二个实质条件是它自我加固。判断力越弱，越依赖可以自动执行的判准；越依赖自动判准，那些需要判断力才能识别的东西越进不来；越进不来，判断力越没有机会被使用。本书第一章到第三章追的就是这个回路在判定端的三种具体形态：一次否决可以在没有任何人打开过那份东西的情况下作出（第一章）；被丢弃的那些东西会在无人记账处合成下一轮的标准（第二章）；一道关可以在最近一百次判定里一次也没拦下过谁，而它不会被撤销，只会被继续满足（第三章）。

至于表象条件——一场危机通常应该被人感觉到——这本书的判断是：这一场不会。这不是修辞上的悲观，而是从机制本身推出来的：全部现行读数的设计前提，是"发生了的事情会留下记录"。而本书所描述的这一类事，其定义性特征就是不留记录。用现行读数去找它，找不到不是证据，是同义反复。

所以本书的工作不是发警报，是造读数。十章各交出一个可以在既有日志里算出来的数，用来看见那些按设计看不见的东西。这是全书唯一的兑现物，也是它与一切"AI 冲击论"的分界线。

■ 五、十章的次序

本书分四编，另有一个枢纽章。

编一，判断没有发生。这一编处理判定端的三种空转形态：否决没有经过阅读（第一章）、标准由被丢弃之物合成（第二章）、关口失去分辨功能而不被撤销（第三章）。三章各

自成立，而它们描述的现象在表面上互相矛盾——第一章说绝大多数被否决，第三章说通过率趋近于一。同一个系统不可能两者都是。

枢纽章处理的正是这处矛盾。它把两章的读数用一条关系式锁在一起，指出它们说的不是同一个站，而是同一条链上的上下游两类站；并由此得到一个两章各自都写不出来的结论：两边的第一步互不相干，必须同做；只做一半的一方会在自己的读数上看到改善，而处境不动。

编二，判断的成本被转嫁给谁。判定这件工作没有消失，它被移了位置。第四章追的是：一件产出里有多少是为了证明它出自谁手——归属证据必须被单独生产，而三家账本只给它成本位。第五章追的是：核实这份工作如何从评价方整体转嫁给被评价方，并且因为它与真实产出同形而无法被分离计量。第六章追的是时间——一样东西从被做出来到取得第一个读数所需的那段时长，它落在所有效率读数的分母内、分子外。

编三，判断所依赖的单位失灵。判断不是凭空作出的，它依赖一套计量约定。第七章处理计数：生成成本趋零之后，“一件”这个单位失去了下界，于是“件数上升”不再能充当“投入上升”的证据——而几乎所有关于内卷的讨论都建立在这个证据上。第八章处理分辨：可分辨的种类数不是系统的自然属性，它由观察者的分辨集合决定，因此它是一项产出；而没有任何人把它当作产出来生产。

编四，判断力从哪里来。前三编讲失效，这一编回答来源。第九章追的是那段不产出任何可验收之物的时间——验收能力正是在这段时间里被生产的，而生成式协助恰好把这段时间前置掉了。第十章是全书的收束：一类被普遍依赖、却不由任何一方生产的能力，它不通过被消耗而减少，只通过养成它的那些低效活动被逐一优化掉而消失，且消失不产生任何事件。

合章把十项读数锁进一个关系网。前四编交出的十项，到那里为止是并排放着的十件东西，而并排的读数无法互相否定——任何一项测出异常都可以被解释成“这一项的定义有问题”。合章用五条从定义直接推出的关系把它们联立，其中最干净的一条是 $\bar{w} = u/(1-u)$ ：首读时距与未阅率不是两件事，是同一个拥堵参数的两种读法。

结语把十项读数并成一张体检表，任何一个机构都可以对自己跑一遍。

另有三处附节。本书十章有九章用的是同一组三学科（经济学、艺术学、工程学），这在书一级上是明确的短板——书里的纠缠是同一次碰撞重复了十遍。第四章、第七章、第八章各附一节「换一组学科再问一遍」，分别换上证据法×演化生物学×密码学、计量学×图书馆学×会计学、生态学×语言学×统计学。三次换组各自都改动了原章的处方或读数，而不是加了三个新的参考文献——一次换组如果不改变任何处方，那它就只是装饰。

■ 六、这本书不主张什么

一本主张"某样东西正在消失"的书，最容易滑向两种廉价的态度。这里先把两条路都堵死。

它不主张回到从前。低效活动被优化掉，绝大多数时候是好事。本书不为浪费辩护，也不认为"慢"本身有价值。它主张的是一件更窄、也更难的事：有一类能力是那些低效活动的副产品，而在优化掉它们的时候，没有人清点过这份副产品。清点，不等于不优化。

它不主张人比机器强。本书十章没有一处依赖"人有某种机器没有的东西"这类命题。它依赖的是账目：某一类工作没有栏位，因而没有人负责生产它。这是制度问题，不是本体问题。一个把判断力当作产出来生产、来记账、来分配的系统，无论由谁执行判定，都不会落进本书描述的这些格子里。

最后交代一处本书自己的位置。本书由人与机器共同完成，本身就是生成成本坍塌的产物之一。按它自己的判断，它作为一件被交出去的东西，如果没有被接受，仍然会抬高一点点线——而这一点在原理上无法被证实，因为没有任何人记录未被接受的稿子。本书因此处在一个结构性的位置：它主张存在一类不被记账的东西，而它自己一旦成为那一类，同样不会被记账。

它能拿出来的只有间接读数。第十一节到结语那张表，就是这些读数的清单。

编一

判断没有发生

• • •

■ 编序

这一编的三章处理同一件事的三种形态：判断在名义上作出了，而在实质上没有发生。

第一章处理最干净的那一种——否决时没有任何人打开过那份东西，而它与读完之后作出的否决在所有账本上同码。第二章处理最隐蔽的那一种——判断确实作出了，但它所依据的标准，是由那些被丢弃、不入任何账的产出在无人记账处合成的。第三章处理最普遍的那一种——判定装置还在运转，只是它已经不再分辨；而它不会被撤销，因为没有任何位置有权宣布它作废。

三章读起来会有一处正面矛盾：第三章说通过率趋近于一，第一章说绝大多数被否决。这不是笔误。紧接这一编的枢纽章就是为了处理它，而处理的结果是本书最实用的一条结论。

第一章

两种「不」——没有人打开过的那种否决

■ 摘要

算不算，打开而未滚动算不算，模型读过算不算。不公开这条规则，这个数就是不可申辩的。

还有一条更根本的：这个数一旦进入考核，最省事的达标方式恰恰会摧毁它要保护的那样东西。

达标最便宜的办法有两种。第一种，让系统自动"打开"每一份——一次调用即可产生一个打开事件，未阅率瞬间归零，而实际过目一点没有增加。第二种，规定"模型已读即算已读"——这在制度上更体面，效果一样。

两种办法的后果是同一个，而且比现状更坏：在此之后，连"没有人看过"这件事都不再留下痕迹。

本章

造出的这个区分，会被它自己的被采用所消灭。

本章看不到出口。可以想到的缓解办法——审计打开事件的时长分布、抽样人工复核、要求打开事件与后续判断内容存在可核对的关联——每一条都需要一个愿意花钱去查自己的机构，而正是缺少这样的机构，这个问题才存在。这一条写在这里，不是待办事项，是本章自认的失败面。

两种「不」

■ 一、导论：三条不能同时成立的处方

二〇二六年初，一个普通的岗位平均收到二百四十二份申请。五年前这个数字大约是一百。同一时期，某职业社交平台报告的投递速率达到每分钟约一万一千份，一年内增长四成有余。人不是突然多了四成，是投一份的代价塌了。

学术评审那一侧的曲线形状完全一样。一家组织学领域的主要期刊在其二〇二六年的编辑部报告里给出了自家的完整交易数据：把二〇二二年一月至二〇二三年十一月与二〇二四年一月至二〇二五年十一月两段相比，投稿量多了四成二；作为对照，疫情高峰带来的那次投稿激增只有两成，是这次的一半。同一份报告还说，三成以上的评审意见文本已经可以检出机器生成的痕迹，而在二〇二二年底之前，这个比例接近于零。另一份面向全行业的统计给出了另一端的读数：二〇二二年编辑每发出一次录用要发出一点六九次不经评审的直接退稿，到二〇二五年，这个比值升到了二点四九。

艺术与文学那一侧发生得更早也更直白。二〇二三年二月二十日中午，一家有十七年历史、常年开放投稿的科幻杂志停止收稿。它的编辑给出的数字是：截至停收，当月收到七百份正常投稿与五百份机器写成的投稿；而该刊二〇二二年全年的投稿量约一万一千五百份。他后来写道，问题不是那些东西写得好不好，是速度——"我们被埋住了"。

三个行当，三条曲线，同一个形状：一份东西的产生成本塌到接近于零，投递量随之爆炸，而能接住它们的那一端纹丝不动。人们通常把由此产生的困境叫作内卷：每个人都不得不加码，加码的总和不产生任何新东西，而不加码的人立刻出局。

对于怎么走出去，三个互不引用的领域给出了三条处方，而这三条不能同时成立。

第一条来自经济学。它说结果由条件决定：只要相对成本、任务在人与机器之间的分配、以及围绕这两者的激励被改对，工作方式与产出形态自会跟上；因此该动手的地方是条件那一端，其余是被决定项。

第二条来自工程学。它说一个系统做没做对，只看它的产出是否满足规格与结果指标；至于这个产出是怎么被达成的，属于实现细节，不进验收。既然接收端处理不过来，那就扩它的产能，把重复的判断自动化掉。

第三条来自艺术学。它说价值不在物里，而是在活动里被实现的：同一件东西在不同的实践中是不同的成就，做与看两侧的活动缺了任何一侧，价值就没有被实现。因此要守住的正是那个活动本身，而不是产出的数目或它的条件。

两种「不」

三条处方各自锁定一个结算端——条件端、产出端、活动端——并且各自都认为另外两端只是通往自己那一端的手段。它们不能同时为真：第一条说产出端的合格读数只是中间变量、不能用来结算；第二条说产出端的合格读数就是结算本身；第三条说达成路径正是被评价的那样东西，而第二条明说它不进验收。

本章不打算在三者之间判谁对。本章要指出的是它们共同踩着的一块地：三者都假定，存在某一端，可以在那里单独读出这件事解没解决。

这一条从未被任何一方说出口，因为在各自的框架里它太显然了。而

推翻它的材料，恰恰握在三者之一自己手里。

推翻之后剩下的判断，三条处方都得出，而三者的证据合起来只能得出它：内卷的引擎不是竞争加剧，而是一种处置类型在记录上的消失。

一次否决，可以是有人读完之后作出的，也可以是谁都没有人打开

过那份东西就作出的；而在三个领域的所有账本上，这两者同码。收到答复的人分不出，作出答复的人无须记录，系统里没有这个字段。于是努力的回报不是被竞争压低的，是被“没有人看”截断的——而竞争加剧是这一截断的后果，不是它的原因。

本章的路线是：第二节把三个领域各自最强的主张摆出来并给出可核对的出处；第三节把三条处方放进同一个具体处境，看它们在哪里分岔；第四节找出三者共同站着的那块地，并用其中一家自己的材料把它拆掉；第五节给出命名、辨别表与读数；第六节到第九节做兑现与分界；第十节给出路；第十一节交出三个领域反过来给本章加的限制；此后是边界、证伪、伦理、自省与一条写死日期的赌注。

■ 二、三个领域各自的最强主张

选这三家不是因为它们讲得好，而是因为它们互相顶得住。下面每一家都指得到具体的人、年份与可以直接打开的原始出处。

2.1 经济学：结算在条件端当前最有影响力的一条估算来自一个基于任务的模型。它把最终产品的生产写成一串任务，每个任务可以交给资本或者劳动，两者有不同的比较优势；自动化就是资本能承担的任务集合的扩大。在这个框架里，只要新技术的微观效果表现为任务层面的成本节省，那么它的宏观后果就由两个数决定：被影响的任务占多大比例，以及这些任务上的平均成本节省有多大。按这个算法，十年尺度上由此带来的全要素生产率增量落在不到百分之一的量级。

这个结论的分量不在数字本身，在它的形状：总量读数几乎不动。

而与此同时，几乎每一个具体行当都

在报告自己那一端的显著改善。经济学对这个反差有一套现成的解释——如果自动化只是“过得去”，即它替换了劳动却没有明显提高生产率，那么总量不动正是应有之义；而技术往哪个方向走，不由技术本身决定，由激励决定。

两种「不」

于是这一家的立场可以写成一句判断：该动手的地方在条件那一端。

相对成本变了，任务在人与机器之

间就会重新分配，工作方式与产出形态随之改变。产出侧的好看读数与从业者的感受，都是被决定项，不能拿来结算。

这一家有三条互不包含的支撑。结果层：总量读数不动，这是它自己的模型算出来的。路径层：方向由激励决定，不由技术决定。条件层：一旦相对成本改变，任务分配就会跟着变。任一条被推翻，另两条仍能成立。

2.2 工程学：结算在产出端现代软件交付把"做得好不好"操作成了四个结果指标：部署频率、从提交到上线的前置时间、变更失败率、恢复所需时间。这四个数全部落在产出侧，没有一个描述过程。与之配套的是验证与确认的标准分工：验证问"我们把东西做对了吗"，确认问"我们做的是对的东西吗"，两者都以可检的规格为准绳。任何通过了这一关的实现，在工程的意义同上合格。

这不是懒惰，是一条经过长期检验的纪律。让验收独立于实现路径，正是工程能够分工、能够外包、能够在人员流动中保持质量的原因。一个系统如果要求"必须以某种方式被做出来"，它就无法被第二个人接手。

于是这一家的立场可以写成一句判断：该结算的地方在产出那一端。

规格通过即正确；路径属于实现细

节；产能不够就扩产能，把重复的判断自动化掉。

它的三条支撑同样互不包含。结果层：四个关键指标全在产出侧。路径层：验收由测试与评审把关，任何通过者同等合格。条件层：瓶颈的标准处置是提升瓶颈的产能。

2.3 艺术学：结算在活动端近十年美学里最有推进力的一条主张，是把审美价值从"物的属性"改写为"行动中的成就"。按这个说法，物本身不携带价值；同一件东西放在不同的实践里是不同的成就，而人有理由去做的那件事，取决于他身处哪一个实践。做与看是同一件事的两侧：一件作品若没有在某个实践里被接住，价值就没有被实现——不是被实现得少，是没有被实现。

与之配套的是一条更老的制度性说法：一个东西要成为可被欣赏的候选，须由该实践的成员授予它这个资格。资格不是物的性质，是被给予的。

二〇二四年以来关于生成式模型能否产出极高审美价值的讨论，正是在这条线上展开的：争的不是那些图像好不好看，而是那里有没有一个可以被归属的成就，以及归属给谁。

于是这一家的立场可以写成一句判断：该守住的地方在活动那一端。

产出的数目与它的条件都是外围；

一旦活动本身被抽走，剩下的东西再多也不构成价值。

它的三条支撑：结果层，物不携带价值；路径层，做与看两侧的活动共同实现价值，缺一侧则未实现；条件层，成为候选须被授予资格。

两种「不」

2.4 三对主题的冲突这三条主张两两不能并存。

第一条与第二条：经济学说产出侧的合格读数只是中间变量、不能结算，工程学说它就是结算本身。若产出侧的四个指标全部改善而总量不动，经济学判"这说明改善没有落在构成产出的地方"，工程学判"这说明系统做对了，总量问题是别人的事"。两个判断不能同真。

第一条与第三条：经济学把活动形态当作由条件决定的因变量，艺术学说活动正是价值实现的所在、不能被当作因变量处理。若一项技术改变了人们怎么做事，经济学判"这是相对成本变化的正常结果"，艺术学判"这是价值有没有被实现的问题本身"。不能同真。

第二条与第三条：工程学说达成路径不进验收，艺术学说路径正是被评价的那样东西。同一件产出，工程判合格，艺术判"这里没有成就发生"。不能同真。

三对全冲突。这是本章可以往下走的前提——如果它们只是侧重不同，那么把三者合起来就是综述，不是别的。

■ 三、冲突定位：把三条处方放进同一个具体处境

抽象的分歧容易被当成措辞之争。把三条处方放进一个具体处境，它们会立刻分岔到互相取消。

取这样一个处境：某学会的一个分会场，去年收到两千份投稿，今年收到六千份。评审人的数量从三百八十增加到四百二十。会议照常召开，录用比例从百分之二十六降到百分之九，所有作者都在截止日之前收到了答复，平均答复时间比去年缩短了十一天。参会者的满意度调查略有上升。没有任何一项形式检查不通过。

经济学的处方：问题在条件。投一份的成本太低而期望收益不为零，于是理性的做法就是多投。处方是把成本抬回来——按作者设投稿上限、收取投稿费、要求预注册、引入可信但稀缺的信号。让投递重新变贵，量自然会下来。

工程学的处方：问题在产能。接收端的处理能力没有跟上生成端的增长，这是一个典型的瓶颈。处方是提升瓶颈：扩充评审池，把初筛自动化，用模型做第一道过滤，把人的时间留给通过初筛的那一部分。答复时间已经缩短了十一天，这正说明处置有效。

艺术学的处方：问题在活动。六千份里绝大多数不是某个人在某个实践里做出来的成就，而是被批量产生的物；同样地，评审那一侧也已经不再是"读一篇东西并对它作出判断"这个活

动了。处方是守住活动：拒绝把评审外包给工具，要求作者声明自己确实做了这项工作，重新把"认真读一遍"当作这个共同体的核心义务。

三条处方在这个处境里立刻打架。

两种「不」

按经济学做，投稿上限会把成本落在没有资源的一方身上：有团队的人照样投满，独立作者被挤出去；而任何被明确规定的成本，都可以由工具摊薄——写一份预注册的边际成本，在今天已经不比写一段

贵多少。更要紧的是，工程学会指出，成本一旦被规定，它就成了一个可优化的目标，而优化它的最省事的办法从来不是提高质量。

按工程学做，艺术学会指出，自动化初筛正是把"读一篇东西并作出判断"这个活动整个移出去；而经济学会指出，扩产能只会让下一轮投递量继续上升——如果每一份被处理的概率不变，那么理性的投递者会一直投到边际收益为零，扩多少产能都会被吃掉。

按艺术学做，经济学会指出，实践保护在规模上不可执行，且它会被吸收成一种身份标签："纯人工"变成一个溢价属性，于是它重新进入同一场竞赛，只是换了个赛道；工程学会指出，一条无法被检验的要求等于没有要求——"请声明你确实做了这项工作"这句话，在流程上上不问是一样的。

三条处方互相取消，而三方谁也说服不了谁。这不是因为其中两方犯了错。这是因为它们在不同的语域里作答：一方在问条件该怎么设，一方在问系统该怎么做，一方在问活动该怎么保。在同一个语域里可以判谁对；跨语域的分歧不可判决，唯一的出路是造一个把三端关联起来的说法。

而在造这个说法之前，先要看清它们共同踩着的是什么。

■ 四、三者共同站着的那块地

把三条处方并排放着，会发现一件它们从未说出口的事。

三者争的是：内卷解没解决，该在哪一端结算。

而三者共同假定的是：存在某一端，可以在那里单独读

出这件事解没解决；另外两端只是通往它的手段。

把这句话念给三方听，三方都会说"这还用说吗"。经济学会说，当然，总量读数就是那个结算；工程学会说，当然，产出是否合规就是那个结算；艺术学会说，当然，活动还在不在就是那个结算。会引起争论的是"哪一端"，不是"存在某一端"。这正是共有前提的特征。

推翻这条假定的材料，必须来自三者之一自己——从外面搬一个理由来推翻它，得到的只是第四种立场，那是加入争论，不是取消争论。

这件材料来自工程学。

二〇二五年七月，一个模型评估机构发布了一项随机对照试验：十六位资深开源开发者，在他们自己平均有五年经验的成熟仓库上，完成二百四十六个真实任务；每个任务被随机分配到"允许使用当时的人工智能工具"或"不允许"两种条件。开始之前，这些开发者预计允许使用会让完成时间缩短百分之二十四；全部做完之后，他们估计实际缩短了百分之二十。而实测结果是：允许使用时，完成时间增加了百分之十九。

两种「不」

工程学用这项试验回答的问题是"这些工具到底提不提速"。它在自己的领域里是一条关于工具效果的证据，是一次对营销叙事的校正。没有人把它拿来当作关于结算的证据。

但把它放在本章的问题上，它说的是另一件事：同一段工作，在一个端点上的结算与它在别处的实际状态，可以方向相反；而作出结算的人本人读不出这个方向。

交付出去的合并请求全部合格，产出端的读

数没有任何异常；开发者自己的感受是"快了两成"；而计时器说慢了近两成。三处读数，两处一致地错，只有一处是对的，而那一处恰恰不是任何一端的结算指标，是一个当时被顺手记下来的量。

这一条足以拆掉共有前提。如果一个端点上的结算可以与实际状态方向相反，并且结算者本人读不出来，那么"存在某一端可以单独读出解没解决"就不成立。三条处方各自锁定一个端点，而端点不能单独结算——不是因为它们选错了端点，是因为没有哪个端点能单独结算。

这项试验有它的边界，本章如实标明：它的样本只有十六人，工具是二〇二五年二月到六月的版本，机构本身在二〇二六年二月发布了针对更晚版本工具的新数据并调整了试验设计。因此它不能被用来论证"这些工具没用"。本章用到的不是它的效果量，是它的方向结构——两处读数一致地指向一边，实测指向另一边，而当事人站在错的一边。这个结构不随工具版本变化。

拆掉共有前提之后，还剩一层地基没有被动过。那一层不属于这三家，属于所有人。本章把它留到第八节末尾，因为要先把最近的几种说法逐一划清，才看得见它们共同看不见的是什么。

■ 五、未阅否决

5.1 一个命名一次否决可以有两种来路。

第一种：有人打开了那份东西，看了，作出判断，答复"不"。第二种：没有任何人打开过那份东西，答复"不"。

在收到答复的人那里，这两者完全一样——同一封信、同一套措辞、同一个日期。在作出答复的一方，这两者也一样：流程走完了，答复发出了，记录上写着"未通过"。在系统里，这两者更是一样：数据库里只有一个决定字段，取值是通过或不通过；没有第三个字段记录"这个决定是在有人看过之后作出的，还是没有"。

本章把后一种叫作未阅否决：一次否决，其中没有任何人打开过被否决的那份东西。

这个命名不是给一种已知现象换一个说法。它主张的是：未阅否决在三个领域的账本上不是一种处置类型。

不是记录得少，不是难以测量，是压根不算一样东西。在所有现行的记录体系里，一件被投递的东

西只有两种下场——被接受，或被拒绝；"没有被任何人看过"不出现在这个分类里，因为分类是按结果做的，而这两种否决的结果一模一样。

两种「不」

这一点可以用一个判据来检验：删掉本章提出的这个区分之后，那样东西是不是就不存在了？

未被打开

的文件当然还在硬盘上，那不是问题所在。问题在于"未阅否决"作为一种处置——一种可以被计数、被归责、被申诉的事件——在删掉这个区分之后确实不存在。一封未阅的退稿信和一封读完之后的退稿信，在法律上、制度上、统计上是同一件事。它不是变得难测，是没有对象。

一个已经公开的例子可以坐实这一点。二〇二五年起，若干顶级会议开始按作者设投稿数量上限，超出上限的部分按投稿编号顺序直接退稿。这是一次成建制的、被写进政策的未阅否决：作出否决的依据是编号，不是内容，没有任何人打开过那些稿件。而在最终的统计里，它们与经过评审后被拒绝的稿件计入同一个数字。已经有人把"如何减少不必要的编号退稿"写成了一个优化问题去求解——这恰恰说明，在该领域自己的账上，这类否决被当作待优化的浪费，而不是一种独立的处置类型。

5.2 两种「不」的辨别表一个名字只是一个名字。要让它成为一个能把两样看起来一样的东西分开的读数，需要第二根轴。

第一根轴：入口有没有配给——投递是否被限量。

第二根轴：每一份是否被过目——否决是否在有人打
开之后作出。

全部过目

存在未阅

入口有配

①

给

有人

设

② 明面拥堵：积压、延误、评审人喊累——有信号

④ 静默筛选

入口无配给

旧式评审：工作量守恒，慢，但每一个“不”背后都③形式配给：限额写着，实际形同虚

前三格都会说话。第一格慢，慢是可见的；第二格堵，堵会有人抱怨；第三格的限额与实际不符，一查就查得出来。

第四格不说话。它有三个签名，缺一个就不是它。

其一，短期指标表现为改善。

答复周期缩短，积压清零，答复率上升，投递者的满意度回升。前面那个

分会场的例子里，平均答复时间比去年缩短了十一天——这是一个好数字，而它恰恰是未阅率上升的直接后果。

其二，失效不产生任何信号。

被否决的人收到的答复，与经过认真阅读之后被否决的人收到的答复完全

一致，因而他无从知道自己属于哪一种；作出否决的一方没有义务记录，因而它自己也不掌握这个数；系统里没有相应的字段，因而事后无从清点。三方都不掌握，不是因为谁隐瞒，是因为没有人被要求生成这个记录。

其三，它自我加固。

越是把“及时答复”写进服务承诺，越是把自动初筛正规化，越是用录用率而不是过

目率来衡量一个评审系统的严格程度——未阅率就越高。也就是说，把这套流程做得越规范，这一格就越牢固。

两种「不」

这三条合起来意味着：第四格能通过全部形式审查。

它不是一个“该做的没做”的格子，任何人都看得出

的问题不需要一张表来指认。它是一个所有检查都亮绿灯而实质已经空掉的格子。

5.3 未阅率有了这两根轴，就可以给出一个数。

未阅率：一轮否决中，在没有任何人打开过被否决之物的情况下作出的那些否决，占全部否决的比例。

它对应一句不含判断词的问话——不问任何人的看法，只问记录：>上一轮的“不”里，有多少个是在没有任何人打开过那份东西的情况下作出的？

这句话可以拿去问 workflow 系统、投稿平台的日志、代码托管平台的事件表、招聘系统的附件调阅记录。它不需要任何人给出意见。

这个读数在三个领域里量纲不同而比率相同：· 劳动市场：投递的简历中，从未有人打开过其附件的比例。求职系统的附件调阅事件是现成的。

· 工程：被关闭的合并请求中，关闭之前没有任何人查看过差异的比例。代码托管平台的查看事件是现成的。

· 艺术与出版：被退回的稿件中，评审从未调阅过正文的比例。投稿系统的文件调阅记录是现成的。

三处的单位互不相同——一处是求职者，一处是代码变更，一处是稿件——而比率的定义是同一个。这使得三个行当第一次可以并排比较。

这个读数有四条性质，其中第四条最容易被漏掉，也最要紧。

第一，无量纲，因此可跨领域比较。第二，可事后清点，用的是系统本来就产生的被动日志，不需要任何人补填。第三，它把一个印象变成一个数：从“现在根本看不过来”变成“零点七八”。

第四，它与既有的主要指标正交。这一条必须用一个具体的例子来证明，否则它会被读成现有某个指标的代理，白造一个数。

例子是这样的：一个把答复时限压到四十八小时、积压永远清零、投递者满意度全行业最高的招聘流程，可以有接近一的未阅率——正因为要在四十八小时内答复几百份申请，才必须让绝大多数否决无人过目。

反过来，一个积压严重、答复缓慢、满意度垫底的评审系统，可以有为零的未阅率——每一份都被读过，只是读得慢。

既有指标最好看的时候，这个数可以最难看；既有指标最难看的时候，这个数可以最好看。

能被任何现有指标代理的证明。

5.4 因果方向被颠倒了到这里，可以给出本章的承重判断，并说明它为什么与通行的解释方向相反。

这就是它不

两种「不」

通行的解释是：竞争加剧→每个人不得不加码→加码的总和不产生新东西→内卷。在这条链上，"没有人看"是竞争加剧的结果：人太多了，看不过来，所以只好草草处理。

本章的判断是：这条链的方向反了。

投递成本塌陷→投递量上升→未阅率上升→投递者无法区分"我被认真读过之后落选"与"没有人看过我"→而在无法区分的情况下，唯一理性的应对是加大投递量（既然每一份被看到的概率是未知且很低的，那么多投几份是唯一能提高被看到次数的办法）→投递量进一步上升→未阅率进一步上升。

这是一个自我加固的回路，而它的引擎不是竞争，是反馈的截断。

区别不是措辞上的。两条链给出方向相反的预测。

若竞争加剧是原因，那么当录取名额增加时，加码的压力应当下降。若反馈截断是原因，那么名额增加而未阅率不变时，加码的压力不下降——因为投递者仍然无法知道自己上一次为什么失败，多投仍然是唯一能做的事。

若竞争加剧是原因，那么投递量的上升应当早于未阅率的上升——人先多起来，才看不过来。若反馈截断是原因，顺序相反：未阅率的上升早于投递量的上升——接收端先在某个时点上放弃了逐份过目（通常是引入自动初筛的那个时点），此后投递者才在若干轮反馈无效之后转向数量策略。

这条顺序预测是本章最容易被推翻的一条，因为它要求的两个数都可以在既有日志里回溯，而且它明确指定了先后。相关是容易被巧合满足的，顺序不容易。

最后需要说清楚的是：本章没有主张"不完全审阅是错的"。在信息成本为正的地方，不完全审阅是最优的，这一点经济学早已论证过，本章不与它争。本章主张的是：未过目本身不是错，未过目与已过目在记录上同码才是错。

前者是一个资源配置问题，后者是一个记录问题；而正是后者切断了反馈，也正是后者可以被低成本地修好。

■ 六、十个行当

一条判断如果只在提出它的那几个例子里成立，它是一个说法。要成为一个读数，它必须能在别处预测出具体的人事。下面十个行当各给一处兑现，其中两处反过来削弱了本章。

一、招聘。

平均一个岗位二百四十二份申请，七成以上的求职者在申请过程中使用工具，近半数明确表

示"多投一些以绕过自动筛选"。这正是本章那条回路的教科书形态：招聘方装上自动初筛以应对量，求职者因为拿不到任何区分性反馈而加大量，双方各自都在做理性的事。可预测的具体事：在同一家公司内部，引入自动初筛的那条产品线，其后续的人均投递次数上升幅度应显著高于未引入的产品线，而两条线的岗位数与薪资区间可以控制住。

二、学术评审。

两种「不」

直接退稿与录用的比值从一点六九升到二点四九，退稿增长快于投稿增长。这意味着筛

除动作的增长快于生成动作的增长——正是未阅率上升的间接读数。可预测的具体事：在按作者投稿上限执行编号退稿的会议里，被编号退掉的那批稿件，其在别处的最终发表率不应显著低于经过评审后被拒的那批；若显著低，说明编号退稿意外地筛出了质量信息，本章这一处兑现失败。

三、软件评审。

合并请求的关闭事件与差异查看事件是分开记录的，因而未阅率在这里是最容易直接算

出来的。可预测的具体事：在引入自动检查即关闭的仓库里，外部贡献者的二次贡献率应在自动检查上线之后下降，而内部贡献者不变——因为内部贡献者有别的渠道知道自己为什么被关。

四、文学与艺术投稿。

那家科幻杂志的处置值得单独看一眼：它没有加价，没有扩产能，也没有喊口号

守住实践——它关了门。

关门是本章所说的第一格（有配给且全部过目）的极端形态，代价是新人与国

际作者被挡在外面，而这一点该刊编辑自己在当时就明说了。关门有效，但它把成本落在最没有别的通道的人身上。

五、基金评审。多数资助机构公布资助率，不公布过目率。可预测的具体事：在把初筛从"两位评审各读一遍"改为"一位评审读

"的项目里，申请书的平均长度应下降，而申请数量应上升——因为投递者会向可被读到的那一部分收缩，同时用数量补足概率。

六、法院与行政受理。

立案审查与实体审查在制度上是分开的，因而这里的两种"不"有明确的法律区分

——这是三个领域之外唯一一处已经把这个区分制度化了的地方。它反过来支持本章：当一个体系认为这个区分重要时，它会把它写进程序。

而在招聘、评审、投稿这三处，同样重要的区分没有被写进任何地方。

七、医疗分诊。

分诊本质上是一种未阅处置：按主诉分类而不逐一详查。但医疗把它明确命名了，并且

规定了误分诊的责任。这一处反过来给本章加了约束：未阅本身可以是正当的、甚至是必需的，只要它被命名、被计量、被归责。本章因此不能主张"每一份都必须被过目"。

八、专利审查。

检索式命中与实审是两个阶段，且审查意见书必须写明依据。这意味着专利体系的未阅

率在结构上被压得很低，代价是审查周期以年计。这一处提供了一个可用的对照：同一时期内，未阅率与处理周期呈反向关系，而与决定质量的关系需要单独测量——不能假定后者。

九、教育评分。

大班课的作业批改是未阅否决的高发地带，而学生对此有强烈直觉却拿不到证据。可预

测的具体事：在采用抽样批改的课程里，学生的作业篇幅应向"开头几段"集中——他们会把资源投到唯一可能被读到的位置。这是一条容易被检验、且与常识相符的预测，它的价值在于它可以被量化。

十、内容平台的推荐与审核。

这里的未阅率接近于一，而且从一开始就是公开的设计前提：没有人假装

每一条内容都被人看过。这一处再次反过来削弱本章：当未阅率被公开承认时，投递者的行为不会陷入本章所说的回路，而是转向对分发机制本身的适应。也就是说，本章那条回路的成立条件不是未阅率高，而是未阅率高且不被知道。这一条约束把本章的适用范围收窄了，写在第十一节。

十处兑现，其中第七、第十两处反向。

两种「不」

■ 七、为什么三个领域各自到不了这一条

一个领域看不见什么，通常不是它的缺陷，是它的问题结构的影子。下面三段不是在指出三者的疏忽，而是在说明：它们各自的核心问题，恰好使这一步成为不必要的。

经济学到不了，因为它的记账单位是交易与任务。

一次未阅否决既不构成交易——没有对价，没有合意

——也不构成任务，因为它不产生任何可以被计入最终产品的东西。经济学问的是资源如何被配置，而未阅恰恰是"未进入配置"；一个未进入配置的东西，在配置理论里没有位置。更要紧的是，它的框架里已经写着答案的一半：任务级的成本节省要能进入总量，前提是被影响的任务本身构成产出。而在选拔型系统里，被工具降本的是"投递"，产出却是"决定"。降本发生在不构成产出的任务上，总量读数就应当不动——这是它自己的算法算出来的。它只差最后一步没有走：去问那个构成产出的任务（作出决定）本身的产能发生了什么。它没有走这一步，因为决定的产能不是一个市场量，没有价格，不在它的数据里。

工程学到不了，因为它的验收对象是通过者。

被拒绝的东西不进入它的质量体系——这不是疏忽，是分

工：工程对交付负责，而被拒绝的东西不会被交付。因此"被拒者经历了什么"在结构上是题外话。这一点在代码评审里表现得最清楚：所有的质量度量都建在被合并的变更上，被关闭的变更只在统计口径里出现一次。工程学甚至掌握着最容易算出未阅率的那份日志——查看事件与关闭事件都被完整记录——而它从来没有把这两列放在一起看过，因为它的问题结构里，那两列分属两件不相干的事。

艺术学到不了，因为它的记账单位是作品与实践。

一个未被过目的投递还不是任何实践的成员：它没有

被授予候选资格，因此它不在艺术学的对象域里。制度性的说法把"授予资格"当作一个总是可用的动作——某个成员代表实践作出授予——而从没有问过这个动作本身有没有产能上限。这不是一个疏漏，是因为在那套说法被提出的年代，投递量与授予能力之间根本不存在数量级的差距。艺术学掌握着最直白的相反证据（一年一万一千五百份投稿的刊物在一个月内被迫关门），而它用这份证据讨论的是"机器写的东西算不算作品"，不是"授予资格的能力是有限的"。

三者都不是因为想不到才停在原地。它们各自的问题恰好使这一步成为不必要的——而这一步只有在三个问题被同时摆开时才成为必要的。

■ 八、与最近的几种说法的分界

一个新说法最大的危险不是被反驳，是被吸收。下面逐一划清，每一条都给出一个可以分出胜负的观察，而不是"侧重不同"这种不可判定的话。

其一，内卷化（Geertz，一九六三，人类学）。这个词的原产地讲的是爪哇水稻农业：土地不增，劳动力持续涌入，耕作方式越来越精细，人均产出停滞。它的形状是"投入增加、结构不变"。本章说的形状不同：单位投入下降、件数爆炸、结构不变。

两者的处方相反——前者需要一个出口（把劳动力转出

去），后者需要一个地板（把某种代价重新设回去）。判定形式：若观察到投递者的单位投入随内卷加剧

两种「不」

而上升，则原义成立而本章错；若观察到单位投入下降而件数上升，则反之。当前三个行当的数据都指向后者，但这不是本章可以自封的——它需要按行当分别测。

其二，市场拥堵（Roth 及市场设计文献）。这一支明确处理"参与者太多以致市场无法在有限时间内处理足够多的交易"，并给出了成熟的处方：可信但稀缺的信号、时间窗、集中匹配。它是本章最近的经济学邻居。分界在于：拥堵理论假定交易在排队等待处理，而本章说的那一批从未进入任何队列。

判定形

式：拥堵理论预测，增加可信信号会提高匹配质量并降低拥堵；本章预测，在未阅不被记录的地方，增加信号只改变哪些投递被读，不改变被读的比例——未阅率随投递量继续上升。经济学职位市场上的信号机制已经运行了十几年，两边的数据都在。

其三，理性疏忽（Sims，二〇〇三）。

它说注意力有限且获取信息有成本，因而人会最优地分配注意

力；据此，看不完是最优的，未阅是效率而非病理。这是对本章最强的一条竞争解释——"其实不就是看不过来吗"。分界在于阅读深度随什么变化：理性疏忽预测它随赌注变化，赌注越大读得越细；本章预测在裁决质量不产生任何反馈的地方，阅读深度随体量变化而对赌注不敏感。判定形式：取同一批评审者处理高赌注与低赌注两类材料，若阅读时长与赌注相关而与当期体量无关，则理性疏忽足够而本章多余；若与体量相关而与赌注无关，则反之。

其四，成本病（Baumol 与 Bowen，一九六六）。它说生产率无法提高的部门（现场演出、教育、护理）会随其他部门生产率上升而相对变贵。裁决正是这样一个部门：判断无法被加速而不改变它本身。分界在于价格与质量哪一个动：成本病预测停滞部门的相对价格上升，社会仍然买得到它，只是更贵；本章预测在裁决不被购买的地方——无偿的同行评审、打包在薪酬里的简历初筛、算在教学工作量里的批改——价格不会上升，因为没有价格；于是调整发

生在别处，表现为单件裁决时间下降。判定形式：把裁决被单独计酬的场合（付费审稿、外部咨询鉴定）与不被单独计酬的场合放在一起，若前者出现价格上升而后者出现单件时长下降，则两条说法各管一段且本章是必要的补充；若两处都出现价格上升，则成本病足以覆盖本章。

其五，约束理论（Goldratt，一九八四）。这是本章的方法学邻居，也是最危险的一位——因为本章关于“接收端才是产能”的说法，形式上与“系统的产出由它的约束决定”几乎一样。分界在于约束是否可被识别：约束理论的整套操作以“找出约束—让其他环节服从它—提升它”为前提，而这三步都要求约束作为一个工位出现在流程图上。本章说的这个约束不在任何一方的流程图上：投递方的图里没有“被过目”这道工序，接收方的图里“初筛”是一项间接费用而不是一道产出工位。判定形式：若在系统里明确画出“过目”这道工位并对其扩产之后，未阅率随之下降并保持，则约束理论足够而本章多余；若扩产之后未阅率不降（因为投递量随扩产而上升），则本章的判断成立——而这正是前面那个分会场的实况：评审人从三百八十增加到四百二十，投稿从两千增加到六千。

其六，信号理论（Spence，一九七三）与信号贬值。

它处理的是发送端的成本差：一个信号之所以可

信，是因为低质量者付不起它。本章处理的是接收端的一件更前的事：这个信号有没有被读。

一份成本

极高的信号，若从未被打开，其成本差的效力为零。判定形式：信号理论预测提高信号成本会提高分离效率；本章预测在未阅率高的地方，提高信号成本只会提高投递者的支出不提高分离效率——因为分离发生在读之后，而多数否决发生在读之前。

两种「不」

其七，制度性的候选资格（Dickie）。它说一个东西成为可被欣赏的候选，须由实践成员授予其资格。

这是本章那个概念的直接前身，也是最容易与本章混淆的一条。分界在于授予能力是否有限：制度论把授予当作一个总是可用的动作，因此它的对象域里只有“被授予”与“被拒绝授予”；本章指出还有第三种——未被处理，即那个动作根本没有发生。判定形式：若在投递量上升一个数量级之后，被明确拒绝授予的数量同比例上升，则授予能力无限而本章错；若被明确拒绝的数量基本不变而未处理量吸收了全部增量，则反之。

其八，注意力经济（Simon，一九七一）。它说信息的丰富造成注意力的贫乏，因而需要有效地分配注意力。本章不与它争这个描述。分界在于病灶：注意力经济把问题定位在注意力的稀缺，处方是更好的过滤；本章把问题定位在注意力的分配方式不被记录，因此无论过滤得

多好，投递者都拿不到关于自己那一份的任何信息。判定形式：若改善过滤（更准的匹配）之后，投递者的加码行为下降，则注意力经济足够；若过滤改善而加码不降，则本章成立——因为过滤的改善并不改变投递者能否区分两种“不”。

其九，古德哈特定律。

它说一个度量一旦成为目标就不再是好度量。本章不是它的一个实例：本章没有

主张任何度量腐蚀了被度量者，本章主张的是一种处置类型在记录上不存在。这是一个记录学的缺口，不是一个激励扭曲。不过第十四节会说明，本章提出的读数一旦进入考核，古德哈特定律会立刻对它生效——这是本章自己的软肋，不是它的论据。

8.1 这几种说法共同看不见的是把上面八条各自的结构性盲区并排写出来，句式统一为“它把某样东西当作给定，因此看不见另一样”，且判据只有一条：一旦承认后一样，它自己就塌了。

· 拥堵理论把“投递已进入队列”当作给定，因此看不见“从未进入队列”。一旦承认有大量投递从未进入任何队列，它的排队模型就失去了对象。

· 理性疏忽把“注意力被分配到某处”当作给定，因此看不见“未被分配”。一旦承认存在一整批未被分配的对象，“最优分配”就不再能解释观察到的分布。

· 制度论把“授予资格的动作总能发生”当作给定，因此看不见“能力饱和”。一旦承认饱和，“实践成员授予地位”就不再是一个总能被执行的动作，制度性定义本身失效。

· 约束理论把“约束可被识别为一道工位”当作给定，因此看不见“不作为工位存在的约束”。一旦承认存在这样的约束，它的三步操作全部无处下手。

四行背后是同一个被当作给定的东西：凡是一次处置，都必定有人作出过它。

拒绝、退稿、不录用、不

入选——所有这些词在所有学科的记录里都被当作“一个被作出的决定”来处理。正因如此，谁也没有回头看过这块地：“没有人作出它”这种处置，不属于任何一个学科的对象域。

各派互相批判，而它们都站在

这块地上。

这一层比第四节那一层更靠后，也更值钱。第四节那一条是三者共同假定的，拆掉它得到的是本章的方向；这一条是整片地基，拆掉它得到的是本章为什么此前没有被任何人辨识过——不是因为没有人看，是因为看的人都站在它上面。

两种「不」

■ 九、与本书另外几章的分界

本书此前已发四十五篇。其中两篇与本章同日发出，用的是同样这三个学科、答的是同样这道题——同三学科本书是读者一眼就能认出的自撞，也是最难辩解的一种，因此先交代这两篇，再交代此前六篇同族。

每一条各给一个可以分出胜负的对照。

9.1 与同日的两篇第三章《这道关拦下过谁》（论可读门槛的整体饱和）。这是三篇里与本章最近的一篇，近到共有前提完全相同：它指认的三家共同假定，与本章第四节写的是同一条——出路可以在某一个终点上单独结算。两篇的分歧不在诊断的前提，在推翻它的材料与由此落到的病灶。那一章用的是生产率曲线的实施期证据（来自经济学自己），落在门槛这一侧：关口整体饱和，通过率趋近于一，分辨迁出可读区；它问的是“这道关最近一百次判定里拦下过几次”。本章用的是随机对照试验（来自工程学自己），落在否决这一侧：一次否决有没有人看过；本章问的是“上一轮的‘不’里，有多少个是没有任何人打开过就作出的”。

两个读数在数学上近乎正交，而且它们可以在同一条流水线的不同工位上同时为真：前段的关拦不下人（饱和），后段的否决没有人看（未阅）。一个把百分之九十九的投递否掉的入口，按那一章的读数完全不饱和——它确实拦下了绝大多数人；按本章的读数却可以接近全盘未阅。

判定形式：若在通过率趋近于一的场合（那一章的靶格）同时观察到未阅率趋近于一，则两条判断处理的是同一样东西，本章可被那一章吸收；若在通过率极低而未阅率极高的场合仍能观察到本章所说的加码回路，则两者是两个独立的机制。

当前三个行当的录取比例在下降而不是上升，这指向后者，但这不是本文可以自封的，需要按行当分别测。

比“我们不一样”更值钱的是两篇的互补关系：那一章解释了为什么门槛不再拦得住人，本章解释了为什么被拦下的人不知道自己是怎么被拦下的。

两者叠加时，投递者的处境恶化不是相加而是相乘——他既无法靠通过门槛证明任何事，也无法从被否决中读出任何信息。

第四章《认领物》（论内卷是产出的构成中无账可入之物的占比持续上升）。它说的是产出的构成：为使一件产出可被归属于某个人，必须与它一并被生产的那一层东西，账本只给它成本位、不给产出位；读数是认领占比。本章说的是处置的类型：一次否决有没有人看过；读数是未阅率。

判定形式：一个认领占比极高的系统——大量精力花在证明“这确实是我做的”——可以有为零的未阅率，因为核对来源恰恰要求逐份打开；反过来，一个无人关心作者是谁的系统可以有接近一的未阅率。两个数正交。更值得跑的是这一条对照：若给未阅否决记账之后，认领占比

随之下降，则两篇处理的是同一个病灶的两端（一旦投递者知道自己根本没被看，他就不必再加码去证明那是自己做的）；若记账之后认领占比不动，则两者是两个独立的病。这条对照可以在同一批数据上跑出来，不需要新采集。

9.2 与此前六篇同族的分界《借道》（结果等价的替代路径被使用之后，以结果为准的读数与自己的感觉都不报警）。它讲的是同一个行动者在完成同一件事时换了一条内部路径；本章讲的是两个不同的行动者之间的一次答复。判定形

两种「不」

式：若在一个未阅率为零、而结果读数照旧不报警的系统里仍能观察到能力流失，则那一章的机制独立成立，本章与之无关；若把未阅率修正之后流失也随之停止，则本章覆盖了那一章的一部分。二者可以独立发生，叠加时投递者的处境恶化速度不是相加而是相乘——因为他既读不出自己的路径变了，也读不出自己根本没被看。

《复推率》（学会了没有，不看能否说出规则，而看换一个情境能否重新推出同一条）。它测的是能力，本章测的是处置。同一份被拒的稿件，那一章会问作者能不能在别的情境里重推出同样的判断；本章只问有没有人打开过它。判定形式：一个复推率极高的作者，仍然可以有接近一的未阅率；反之亦然。两个数正交。

《轮空》（一次换手没有完成——上一轮的发起处已退出，按顺序应当接手的那一处没有动）。最近的一条，且它也用了经济学。分界在于空缺的位置：那一章讲的是改变的发起权无人接手，本章讲的是已发生的投递无人过目。判定形式：一个换手率健康、发起权轮转正常的系统，可以有很高的未阅率——因为发起改变的是内部成员，而被未阅的是外部投递者。二者在同一个系统里可以同时为真且互不影响。

《在途》（规则改完之后到人跟上之前，那段时间归谁）。它讲的是时间的归属，本章讲的是处置的类型。判定形式：在一个更新极慢、跟随比接近一的稳定系统里，未阅率仍然可以逼近一——因为规则没变，只是投递量涨了。

《记零位》（一个共同体最后擅长什么，由它的贡献账上被记为零的那类动作决定）。这是最近的一条，因为两者都关于账本上的空缺。分界在于谁的账：那一章说的是贡献被记为零，即一个已经被承认发生了的动作被赋值为零；本章说的是一种处置在记录上不存在，即它连被赋值的资格都没有。判定形式：若给未阅否决记账之后，它被赋了一个很低但非零的权重，那么它落入那一章的范围；若给它记账之后，制度必须新增一个此前不存在的处置类别（而不只是一个权重），则本章成立。当前的实况是后者——所有系统里都没有这个字段。

《能被量的那一面》（一件事最容易被看见的地方，往往不是它长东西的地方）。那一章讲的是可测者接管治理；本章讲的是一种事件根本不进入可测与不可测的划分。判定形式：那

一章预测，把一个不可测的东西变得可测，治理会转向它；本章预测，把未阅率变得可测，治理不会自动转向它——因为没有任何一方的现有考核会因这个数变差而受损（见第十四节）。这是一条罕见的反向预测，也是本章最容易在三年内被检验的一条。

■ 十、出路：一条三步的次序，与唯一可插手的那一步

三条处方各自走不通，而它们走不通的原因是同一个：每一条都在自己那一端结算，而另外两端会把它的干预吃掉。

在生成端加价走不通，有两个独立的理由。第一，任何被明确规定的成本都可以被工具摊薄——写一份预注册、一封定制求职信、一份合规声明，其边际代价已经趋近于零；成本一旦被写成规格，它就是一个可

两种「不」

优化的目标，而优化它最省事的办法从来不是提高质量。第二，成本对资源多的一方无效：它筛的是资源，不是质量。二〇二五年若干会议按作者设投稿上限，超出部分按编号退稿，正是这条路走到尽头的样子——它把成本变成了一个抽签。

在裁决端扩产走不通，因为扩产会被投递量吃掉。如果每一份被过目的概率不变，投递者会一直投到边际收益为零；接收端每扩一分产能，投递端就补上一分量。更糟的是自动化裁决这条路：自动化的判据一旦稳定，就可以被生成端逆向拟合，军备竞赛平移一层；而且它把未阅从“可见的空白”变成“不可见的已读”——系统会记录一次处理，只是那次处理里没有人。

守住实践走不通，因为它在规模上不可执行，而且会被吸收。“纯人工”一旦成为一个可以声明的属性，它就变成一个溢价标签，重新进入同一场竞赛。一条无法被检验的要求，在流程上等于不提要求。

三条都走不通，而它们的共同点是：都想在某一端一次性把事情结算掉。

出路因此不可能是第四条同类

型的处方，只能是一条穿过三端的次序。

第一步（产出端）：给“不”记账。

在既有的被动日志里加一个字段，把否决分成“有人打开过之后作出的”与“没有任何人打开过就作出的”两类。这一步的全部内容就是这一句。

它之所以是唯一可插手的那一步，有三个理由。其一，它不需要任何人同意：不需要改制度，不需要谈判，不需要新的预算；数据在绝大多数系统里已经存在——附件调阅事件、差异查看事件、正文调阅记录都是系统本来就产生的。其二，它不改变任何一方当前的行为，因此没有人会因为它而受损，阻力最小。

其三，它把此后所有的争论都变成可裁决的：一旦这个数存在，第二步与第三步该不该做、做到什么程度，就不再是立场之争。

必须限定的是：这个字段只能取自被动日志，不得设计任何需要裁决者主动填报的项。理由在第十一节。

第二步（条件端）：把配给从投递端移到过目端。

既然投递端封不住，配给只能设在过目端。要限的不是"你能投几份"，而是"你能消耗多少份被过目"。

这个换法的性质与所有现行处方相反。投递端的配给对生成方不平等——有资源的人照付，会用工具的人照绕；而过目端的配给对生成方一视同仁：无论你生成得多便宜，你消耗的过目时间都是同一份。更要紧的是，过目端配的是真实产能：一个选拔系统的产出不是它收到多少份东西，而是它作出多少个有据的决定，而后者的上限就在这里。

具体形态可以有多种：投递者持有一定数量的"过目额度"并可在多个入口之间自行分配；接收方公开自己每一轮的过目产能并据此设定收件量；额度可转让、可标价，从而让愿意为被认真看一眼付费的人有一条正路可走。本章不主张其中任何一种是唯一正确的形态——本章主张的只是配给的位置该换一端，而具体机制该由各行当自己设计，并用第一步产出的那个数来检验。

第三步（活动端）：让"认真看一遍"重新成为一件可以被承认的事。

两种「不」

当过目成为一项被计量、被支付、被追责的产能之后，"读一份东西并对它作出判断"重新成为一件可以被做、被记录、被承认的工作。这正是第三条处方想守住的那样东西。

但它不是被"守护"回来的，是被记账回来的。这个区别是本章与第三条处方最实质的分歧：呼吁不能创造一个岗位，一个字段可以。

次序不可颠倒。

先做第二步而不做第一步，配给的额度将无从校准，且立刻退化成又一种投递端限额；

先做第三步而不做前两步，得到的是一句无法执行的倡议。而只做第一步、后两步一步不做，仍然有独立价值——因为投递者从此可以知道自己属于哪一种"不"，而本章那条回路的引擎正是"不知道"。

可插手的那一步只有第一步。

这也是本章与前三条处方在实践上最大的差别：前三条都需要某个有权者

先同意，第一步不需要。

■ 十一、三个领域反过来给本章加的三条约束

一条判断如果只从三个领域取走论据而不被它们削弱，那多半是取巧。下面三条各自削掉了本章原本会主张的一句更强的话。

艺术学削掉了本章的价值排序。

艺术史自己知道，长期未被裁决恰恰是许多重要作品得以不被当时标准

同化的条件：如果每一份投递都在当时被认真读过并被判定，那些不合当时尺度的东西会被更快、更彻底地否掉。本章原本会写下的更强的话是“未阅率是一个应当被压到零的指标”。有了这一条，只能写成：未阅率的意义是它必须被知道，而不是它必须为零；在某些系统里，正确的目标值不是零。

这一处削掉的

是价值排序本身，不是一个补充条件。

经济学削掉了本章的适用范围。

在信息成本为正的地方，不完全审阅是最优的。本章原本会写下的更强

的话是“任何未过目的否决都是一次失职”。有了这一条，只能写成：未过目本身不是错，未过目与已过目在记录上同码才是错。

本章因此从一个程序正义的主张退到一个记录可分辨性的主张——退了很多，但退到的位置更硬。

工程学削掉了本章的工具形态。

工程学的经验有两条：加一个需要人填报的字段，就会加一套上报行

为，而上报行为会被优化；打开事件极易伪造，一次调用即可产生。本章原本会主张“让评审者记录他读了哪些部分”。有了这一条，只能写成：读数只能取自系统本来就产生的被动日志，不得设计任何需裁决者主动填报的项。这一处直接改掉了第十节第一步的实现方式。

第十节的第四、第十两处兑现还给出了第四条约束：本章那条回路的成立条件不是未阅率高，而是未阅率高且不被知道。

在未阅率被公开承认的地方（内容平台的分发），投递者不会陷入这条回路，而会转向对分发机制的适应——那是另一种病，不是本章说的这一种。

两种「不」

■ 十二、不适用的边界

本章不适用于以下几类情形，写在这里以免被误用。

其一，接收端本就不承诺过目的场合。

公开征集、开放投稿箱、悬赏，若事先明确声明不保证审阅，则

两种"不"在双方之间不构成信息不对称，本章的回路不成立。

其二，产出不是"决定"的系统。

本章的全部推理建立在一类特定系统上：它的产品是有据的决定，而不

是物或服务。工厂、物流、医疗操作不在此列——它们的产能瓶颈是可见的，且本来就被当作工位管理。

其三，未阅率高而信息由别的渠道回流的场合。

若投递者能从别处（内部关系、公开评分、二次申诉）

稳定地知道自己被否的类型，则反馈没有被截断，回路不启动。

其四，能力不足型的困境。

有些行当的困境是投递者确实不够格，与被不被看无关。把本章套上去会得

出错误的处方——先测未阅率，若它接近零而困境依旧，那么问题不在这里。

■ 十三、证伪条件

最强的竞争解释是最朴素的那一句："其实不就是人太多、位置太少吗？"

本章的证伪条件必须能把它排

除掉，否则只是验证了现象存在。

主条件：控制投递量与录取名额之后，若未阅率与投递者下一轮的投递次数之间不存在剂量—反应关系，则本章的机制为伪——届时困境应归因于绝对供求比，而不是反馈截断。

样本纪律：这一条要求同质多单元，至少六个，且在关键混淆变量上同质——同一平台上的多个招聘方、同一学会下的多个分会场、同一出版社旗下的多本刊物。严禁拿两个国家或两个地区的对比来充数：那类比较在文化、规模、制度上处处不同，所谓"控制掉某个变量"在方法上站不住，只能作启发式说明。

其余条件：其一（正向读数，顺序而非相关）：未阅率的上升应当早于投递量的上升。

若在多数单元里观察到相反

的顺序——投递量先涨，未阅率后涨——则本章的因果方向错，拥堵解释成立。这是本章最容易被推翻的一条。

其二（低成本可实做）：取任一代码托管平台上的若干仓库，算出被关闭的合并请求中关闭前无人查看差异的比例，与该仓库外部贡献者的二次贡献率作对照。所需数据在平台的既有事件表里已经存在，不需要新采集。

其三：若在把未阅率公开的系统里，投递者的加码行为不下降，则本章关于"不知道才是引擎"的判断错。

其四：若在自动初筛上线之后，同一公司内部引入初筛与未引入初筛的产品线，其人均投递次数上升幅度没有差别，则第六节第一处兑现失败。

两种「不」

其五：若按编号退稿的那批稿件在别处的最终发表率显著低于经评审后被拒的那批，则编号退稿意外地携带了质量信息，本章关于"未阅否决不携带信息"的说法在该场合失效。

这条判断若被证伪，我们会学到什么？

若主条件不成立，说明困境确实是绝对供求比问题，那么干预方

向应回到扩大位置数量或减少参与者，而不是修改记录方式——这是一个方向完全不同的处方，值得单独去证。

本章交出两个自己解释不了的洞。

其一：未阅率的"正确目标值"在不同行当里应当是多少，本章只能说

不是零，说不出是多少。其二：在未阅率被公开之后，投递者行为会转向哪里，本章只知道会转，不知道转到哪。

■ 十四、这个读数一旦进了考核

本章提出的这个数会被拿去考核人，这一点几乎是必然的。因此必须先写清楚三条，缺一条则不应采用。

其一，未阅率指的是位置，不是人。

它描述的是一个入口、一条流水线、一次评审轮次的状态，不描述

某位评审者的品行。把它算到个人头上，得到的将是一个可以被个人操纵而与系统状态无关的数。

其二，不得为了把这个数做好看，而在时限关键或安全关键的系统里制造额外的过目动作。

分诊、急

救、事故上报这类场合的快速否决是有正当理由的；用这个数去要求它们逐份过目，是把一个记录学的改进变成一次操作风险。

其三，已经据此对某个入口或某位裁决者作出不利安排的，必须公开编码规则并给出复核通道。"什么算过目"是一个需要明确定义的问题——打开

■ 十五、本章自己

本章自己是一次投递。

它将被读者读过之后判断为有用或无用，还是在没有任何人打开的情况下被推荐机制筛掉，本章无从知道。而按本章自己的判断，这两种结局在任何可见的记录上完全一致：阅读量的曲线不会告诉本章哪一次是有人读了，哪一次是页面被打开又关掉；而本章若属于后一种，它与一篇被认真读过之后判为无价值的文章，在所有统计里同码。

这有两个具体后果。

第一，本章不能靠被发表来证明任何事。

发表在本章自己的框架里正是那个不携带信息的事件——它既

可能来自审视，也可能来自未阅。因此本章只能靠一件事被检验：未阅率被真的量出来，且第十三节那条顺序预测被真的核对。

第二，本章的处方对本章自己适用，且同样只有第一步是本章能做的。

本章不能要求任何平台改变配给

方式，也不能要求任何人认真读一遍。本章能做的只有一件：把自己的判断写成一个可以在既有日志里回溯的数，让不认同它的人有办法证明它错。这篇文章的全分量都押在那一个数上。

■ 十六、结论，与一条写死日期的赌注

三个领域对同一件事给出三条互相取消的处方，而它们共同假定了存在某一端可以单独结算。这个假定被其中一家自己的随机对照试验推翻：一个端点上的结算可以与实际状态方向相反，而结算者本人读不出来。

推翻之后剩下的判断是：内卷不是一场投入过度的耗散型竞赛，不是一处产能跟不上生成的瓶颈，也不是活动本身被掏空，而是一种处置类型在记录上的消失。

「没有人看过就作出的否决」与「看过之后作出

的否决」在所有账本上同码，于是努力的回报不是被竞争压低的，是被"没有人看"截断的；竞争加剧是这一截断的后果，不是它的原因。

判据是一句不含判断词的问话：上一轮的"不"里，有多少个是在没有任何人打开过那份东西的情况下作出的？

读数是未阅率。最危险的那一格是：投递无配给、答复全部及时、而多数否决无人过目——它在短

期指标上表现为改善，它的失效不产生任何信号，而且越正规化越严重。

出路不在生成端加价，不在裁决端扩产，也不在守住实践，而是一条三步的次序：给"不"记账→把配给移到过目端→让认真看一遍重新成为一件被承认的工作。可插手的只有第一步，而第一步不需要任何人同意。

本章可以被分层引用。第一层是"未阅否决作为一种处置类型"，最强，也最容易倒。第二层是那张两轴辨别表，作为分析工具，即使第一层被推翻也仍可用。第三层是一条不依赖前两层的经验主张：当投递量的

两种「不」

增长快于过目产能的增长时，被拒者收到的答复内容不随之改变。

这一条可以在任何一个投稿系统的答

复模板库里直接查证，成本接近于零。

最后是一条写死日期的赌注。

到二〇二九年八月一日为止，在招聘、学术评审、代码托管三个领域中，至少有一个领域出现一家有规模的机构，公开发布其未阅率或等价读数（即区分"经过目的否决"与"未经过目的否决"并给出比例）。

什么不算命中，一并写死：机构自行声明"我们认真对待每一份申请"不算；发布平均处理时长、答复率、录用率不算；发布"人工复核比例"而不说明什么算复核不算；只在内部使用而不对投递者公开不算。必须是一个对外发布的、定义了"什么算过目"的比例数。

若到期没有出现，本章关于"这个数便宜到可以被立刻做出来"的判断就要打折扣——那意味着阻力不在成本，而在别处，而本章没有识别出那个别处。

■ 注释

一、关于那项随机对照试验的使用边界。

本章使用它的方向结构，不使用它的效果量。该试验样本为十

六人、二百四十六个任务，工具为二〇二五年二至六月版本；发布机构在二〇二六年二月公布了针对更晚版本工具的新数据并调整了试验设计。因此任何"这些工具没用"的推论都与本

章无关。本章用到的是：两处读数一致地指向一边、实测指向另一边、当事人站在错的一边——这个结构不随工具版本变化。

二、关于行业统计的可靠性。

第一节与第六节引用的招聘侧数字（每岗位申请数、投递速率、工具使用

比例）多来自平台方与招聘服务商的自有统计，属于利益相关方数据，只能作量级参考，不能作精确读数。学术评审侧的数字来自期刊编辑部自行公布的完整交易数据与行业统计，可靠性较高。本章的承重判断不依赖其中任何一个具体数字，只依赖三处曲线形状一致这一点。

三、关于"过目"的定义。本章全文使用"有人打开过被否决之物"这一操作定义，而未规定阅读的深度。这是有意的：深度无法从被动日志中读出，而一旦要求裁决者自报深度，读数就会被上报行为污染（见第十一节）。因此本章的未阅率是一个下界——真实的"未被认真审视率"只会比它更高。

四、关于内卷一词。

两种「不」

本章沿用中文语境中的通行用法（努力加码而总体不产生新东西），并在第八节与该词的人类学原义划清了界线。两者形状不同，处方相反。

第二章

被丢掉的那一眼——标准由无账可记之物合成

■ 摘要

生成一件东西的成本在很短的时间里降到接近于零之后，一个反复出现的现象是：投入的人更多、交出的东西更多，而能通过的人并没有变多，通过所需要的水准却在上升。对这件事，三个互不引用的领域给出互相排斥的解释。竞赛经济学说，投入的多少由名额与单位成本共同决定，成本降到零，投入就会涨到把奖品全部吃掉为止——多出来的那部分是纯损耗。价值化理论说，一件东西算不算新，由一份被共同承认的收藏名录决定，名录的容量与更新速度是硬的，产量再大也只是在名录之外堆积。约束理论与近年的交付研究说，一个系统的吞吐由它最窄的那一段决定，写得再快也没用，只不过把拥堵从写挪到了看。三条解释不能同时为真：第一条要求线是固定的，第二条要求线由名录一侧单方面定，第三条要求瓶颈有位置、可指认、可扩容。

本章指出三家共同假定了一件从未被说出的事：稀缺先在，行为随后——是名额、名录容量、瓶颈这些既有的窄处在驱动竞争，而没有通过的那些东西，因为不在任何一本账上，不可能反过来驱动任何事。推翻这个假定的材料来自三家自己：竞赛实验里反复出现的过度耗散（总投入系统性超过奖品价值）、约束理论第五步自己写下的告诫（解开一个约束之后，别让惯性成为新的约束）、以及价值化理论自己承认的那句话（在世俗空间与收藏制度的关系里，赢的永远是世俗空间）。

本章的承重判断是：内卷不是投入的浪费，不是新意的耗尽，也不是瓶颈的转移，而是未被接受的产出在无人记账处合成了一条判定基线，并单向抬升它。

这条基线不在任何记录里，却是每一次判定实际使用

的参照；本章称之为暗基线。判据不含判断词：把上一轮全部未被接受的产出移出参照集，这一轮的合格线会降多少？答案在期刊的决定日志、招聘系统的评分记录、代码评审的历史里已经存在，取出来只需要一次重判实验。本章给出两轴四格的辨别表、一个三领域量纲不同而比率相同的读数（暗基线抬升度）、十三个行当的逐项兑现、九条机制证伪条件、一处适

用边界（沉积只在择优制里发生，不在定标制里发生），以及一条写死到二〇二八年十二月三十一日的赌注。

■ 关键词

暗基线；落判件；判定基线；择优制与定标制；暗基线抬升度；生成成本趋零

AThresholdWithoutanAuthor

被丢掉的那一眼

■ 一、导论：三个不能同时为真的判断

二〇二五年到二〇二六年之间，几个互不相干的行当同时报告了同一种形状的困境。

一家人力资源软件公司统计，它平台上每一个招聘岗位平均收到二百五十四份申请，每位招聘专员手上的申请量在几年内增加了百分之四百一十二；求职者花大约二十美元买一个工具，就能把申请一次性投到成千上万个岗位。同一时期，一家管理学期刊的编辑团队清点了自己五年的全部收件：投稿量比人工智能写作工具普及之前高出百分之四十二，而桌面直接退稿的比例从百分之十三涨到百分之四十；退稿与录用之比从一点六九涨到二点四九。请一位同行评审，平均要发出五封邀请才有一位答应。同一时期，一份覆盖近五千名技术从业者的交付研究发现，人工智能的使用与交付吞吐已由负相关转为正相关——东西确实出得更多更快了——而它与交付不稳定性的正相关始终没有消失，报告自己写下的判断是：人工智能不但没有修好不稳定，目前反而与不稳定的增加相连。

三个行当，三套完全不同的记录系统，同一个形状：上游变得极其便宜，下游没有变便宜，中间那道线在往上走。

有意思的不是这个形状本身，而是三个最有资格解释它的理论体系给出了互相排斥的答案。

被丢掉的那一眼

第一个答案来自竞赛经济学。一场竞赛里，参与者投入的多少，由两样东西共同决定：可争夺的奖品有多少个、多大，以及投入一单位努力要花多少钱。奖品的数量是给定的，成本一旦坍塌，均衡投入就会一路上涨，直到把奖品的价值全部吃掉。多出来的那些投入，在这套账里有一个准确的名字：耗散。它是纯损失，既不生产也不留存，它消失了。

第二个答案来自艺术学里关于“新”的价值化理论。一件东西算不算新，不取决于它有多不一样，而取决于它有没有跨过一条线——从无人收藏的日常世界，进入一份被共同承认的收藏名录。名录的容量是有限的，更新的速度也是有限的。产量再大，只要没跨过那条线，它就留

在名录之外；名录之外的世界在这套理论里是无差别的、无限的、没有结构的。所以困境不在竞争，在名录：能被认作新的位置就那么多，做的人越多，每个人分到的越少。

第三个答案来自工程学。一条流水线的吞吐由它最窄的那一段决定，别的地方再快也没有用。写代码变快了，看代码没有变快，于是拥堵从写挪到了看。困境不在总量，在瓶颈的位置。处方是清楚的：找出那一段，把它撑开。

三条解释各自都很有道理，而且各自都有真实的证据。问题是它们不能同时为真。

第一条要求那条线是固定的——竞赛模型里，奖品数与获胜规则是外生给定的参数，投入围绕它调整。第二条要求那条线由名录一侧单方面给出——是收藏制度决定什么算新，创作者只能去撞它。第三条要求困境有位置——瓶颈必须能被指到某一段、某一台机器、某一个环节上，否则“撑开它”这个动作无从下手。

如果第一条对，那么线不动，动的是投入，扩容不解决问题（扩了名额，投入随之上涨，照样吃光）。如果第二条对，那么线由名录动，与投入端的成本无关（成本降到零，名录也不会因此扩容）。如果第三条对，那么线本身不是问题，位置才是问题，撑开那一段，困境应当消失。

三条处方在同一个真实处境里会分岔到互相取消的方向。一家期刊面对投稿激增：按第一条，它应该减少可发表的篇数（把奖品做小，让人不值得投）；按第二条，它应该扩大版面（名录容量是真约束）；按第三条，它应该增加评审人手（撑开瓶颈）。三条不能同时执行，而且已经有足够的经验说明，三条都不管用——桌拒率涨到百分之四十的期刊，是三条都试过的期刊。

本章认为三条都不管用不是因为它们各自有缺陷，而是因为它们共同漏掉了同一样东西。

■ 二、文献综述

2.1 竞赛与耗散：投入由名额与成本共同决定竞赛理论处理的是这样一类处境：若干人作为一个不可分割的奖品投入不可收回的努力，胜负由投入的相对大小以某个概率规则决定。图洛克在上世纪七八十年代给出的基本模型，把这套结构从游说、诉讼、争夺特许权一路推广到一切“投入换机会”的场合。这个模型最有力的结论叫租金耗散：在充分竞争下，全体参与者投入的总和会趋近奖品本身的价值——也就是说，一件本来有价值的东西，会被争夺它的过程本身消耗掉。

被丢掉的那一眼

这条结论的驱动方向是明确的：是奖品与规则在驱动投入。

奖品有多大、名额有几个、赢的概率怎么

算，这些是先存的、外生的；投入是被它们逼出来的反应。给定这套参数，参与者是谁、多聪明、有什么品味，都不改变均衡。这不是这个学派的疏忽，这是它的力量所在——它敢于说，只要规则定了、成本定了，剩下的事就定了。

在这套账里，落败者的投入被记作耗散。这个记法有一个极重要的后果：耗散是一个没有去向的量。

它

不转移到任何地方，不留存为任何东西，它在本轮结算时归零。竞赛模型的每一轮都是同构的：新一轮开始时，参数与上一轮相同，历史不进入状态。

这个学派自己知道一件与它的主线不合的事，本章第五节会取用它。

2.2 价值化与档案：什么算数由一份名录决定关于"新"的当代讨论里，格罗伊斯《论新》给出的框架最少浪漫、最像经济学。他把文化世界切成两块：一块是被共同承认、被收藏、被记住的名录，另一块是无人收藏的日常世界。一件东西之所以是新的，不是因为它此前不存在，而是因为它跨过了这两块之间的那条线——把从前被当作无价值的东西，放进了名录；同时把名录里某样从前被当作有价值的东西，推回日常世界。他把这个过程叫作价值的重估，并且明确说，这是一场经济性质的交换，不是一次灵感事件。

关键在于，"新"在他这里是一个比较关系，不是一个内在属性：所谓一次新的出现，就是做出了一次从来没有人做过的比较。而比较必须有一个参照的对象，那个对象就是名录。所以驱动的方向是：是名录在决定什么算新。

创作者做什么、做多少，本身不产生新；只有当名录接住它、并因此重画自己的边界时，新才发生。

这个框架有一个自己写下的前提条件，本章第五节会取用它。

同一条脉络上，近两年有一份很硬的实证。一项在线实验让约三百名写作者写短篇故事，其中一部分人可以看到语言模型给出的故事点子。结果是双向的：拿到点子的人写出来的故事被评为更有创意、更好读，而且基线创造力越低的人提升越多；但拿到点子的那批故事，彼此之间的相似度显著高于纯人类写的那批。作者把它总结成一个社会困境——用了工具的人个体上更好，而集体层面上新内容的范围反而变窄。

这份材料在本章里不是背景，它是价值化理论在当下的经验形态：产量在涨，而能构成"一次从未有人做过的比较"的存量在缩。

2.3 约束与瓶颈：吞吐由最窄的那一段决定工程学这一侧最成熟的框架是约束理论。它的核心断言极简：任何系统在任何时刻，达成目标的能力都被极少数（通常是一个）约束限制

着；改善非约束环节不增加吞吐，只增加在制品。它给出的操作是五步：找出约束、把约束用尽、让其余一切服从约束、抬高约束、如果约束被打破了就回到第一步。

驱动的方向同样明确：是约束在决定吞吐。

上游能力提高多少，不改变系统交付多少；唯一起作用的是

最窄那一段的通过能力。这一条在软件交付上刚刚得到一次大规模验证。二〇二五年那份覆盖近五千人的研究发现，九成的开发者每天使用人工智能，吞吐确实上去了，而不稳定性同步上去，返工、变更失败与

被丢掉的那一眼

修复时长都在增加；报告把原因指向下游——测试、代码评审、质量保证这些环节没有为新的速度做好准备。它还识别出一类占比百分之十七的团队：系统本身是稳定的，人却被流程拖住，倦怠很高。

约束理论的第五步里写着一句告诫，本章第五节会取用它。

■ 三、冲突定位：三条驱动不能同真

把三家各自最承重的那一条并排放，冲突是硬的。

竞赛经济学说：规则与成本之间的张力，决定了投入呈现出什么形态。

名额固定而单位成本坍塌，均衡

努力就飙升，产出形态从"少而重"变成"多而轻"。这一条要求线不动。

价值化理论说：实际做出来的东西与现行的认定路径之间的张力，逼得名录本身改变。

大量差异被生产

出来，认定路径跟不上，最后动的是名录的边界——它要么扩容，要么把标准提高。这一条要求名录是唯一有权动线的一方。

工程学说：已有的产能与承接条件之间的张力，逼得工作路径改变。

写得快了而看不过来，于是加门

禁、加自动预筛、加评审规程。这一条要求困境有一个可指认的位置。

三条不能同真，理由不在措辞，在处方。取一个具体处境：一家期刊今年收到的稿件比三年前多了四成，桌拒率从百分之十三涨到百分之四十。

按第一条，这是投入端的问题：可发表的名额没变，投稿成本坍塌，于是投稿量涨到把"发表"这个奖品的价值吃光为止。处方是让投稿变贵（收费、限投、要求预注册），把成本重新抬起来。按这一条，编辑那一侧什么都不必做，也做不了什么。

按第二条，这是认定端的问题：能被认作"值得读"的位置就那么多，名录的容量是真约束。处方是扩容或者换标准。按这一条，投稿端做什么都无所谓，因为不跨过那条线的东西根本不构成事件。

按第三条，这是位置的问题：瓶颈从写论文挪到了读论文。处方是把评审那一段撑开——更多评审人、更好的分配算法、机器预筛。按这一条，只要撑开了，困境就应当消失。

第三条最容易检验，而它已经被检验了：期刊在过去五年里做的正是撑开评审这一段，机器预筛上线了，桌拒作为一种"低成本拒绝"的手段被大规模启用了——而线继续往上走。撑开瓶颈之后线没有回落，这一条就交不了差。

第一条也交不了差：投稿成本对个人来说确实降了，但一个更强的事实是，这条线在参与人数没有增加的行当里也在往上走。

招聘就是最清楚的例子——申请量涨了四倍，而有相应学位与经验的人数并没有涨

四倍。奖品没变小，参与者的实际人数没有暴涨，暴涨的是每人交出的件数。竞赛模型可以处理"每人投入更多"，但它把这些多出来的投入记为耗散，也就是记为对下一轮没有任何影响的量。而下一轮的线明明变了。

第二条同样交不了差，理由在下一节。

被丢掉的那一眼

■ 四、三家共同的默认：稀缺先在，行为随后

把三家的争论压成同一个句式，它们争的是：这场竞争里真正稀缺的那一样，该放在哪儿。

放在名额上

（经济学），放在名录容量上（艺术学），放在瓶颈的通过能力上（工程学）。

三家争的是稀缺的位置，而它们共同假定了一件事：稀缺是先存在的，行为是被它逼出来的反应。

名额先

有，投入随后；名录先有，创作随后；瓶颈先有，拥堵随后。这个方向在三家那里都不是结论，是出发点——它不被论证，因为在各自的问题里它不需要被论证。

把这句话念给三家听，三家都会说这还用说吗。这正是它是共同前提的标志。

这个前提有一个直接后果，本章认为它是整件事的关键：没有通过的那些产出，在三家的账上都不是一样东西。

在竞赛经济学的账上，它是耗散——一个数量，一笔损失，本轮结清，不进入下一轮的状态。在价值化理论的账上，它留在名录之外的日常世界里；而日常世界在这个框架里必须是无结构的，否则“新=把日常物拿进名录”这个定义就失去意义——一个已经有结构、有层次、有秩序的日常生活世界，无法充当那个可以被随手取用的无限储备。在工程学的账上，它是废品：吞吐只计算完成的那些，被退回、被回退、没过评审的变更不计入交付；它们消耗了瓶颈的容量而不产生产出，因此在吞吐这个指标上，它们的贡献是负的零——既不加，也没有位置去记减。

三个账本，三种不同的记法，一个相同的结果：未被接受的产出，在系统状态里没有字段。

它不是被记

录得少，它是根本不算一样东西。

于是“它反过来驱动了什么”这个问题，在三家那里都不可能问出来。不是因为答案是否定的，是因为问题不成立——一个不存在的东西不能驱动任何事。

■ 五、推翻这个默认的材料，来自三家自己

要推翻一个共同前提，从外面搬一个理由来是没有用的——那只是加入了争论，成了第四家。真正能取消争论的材料，必须是三家自己已经掌握、已经发表、只是没往这个方向用的事实。三家各有一件。

5.1 过度耗散：竞赛经济学自己的实验竞赛理论预测，充分竞争下总投入趋近奖品价值，从下方逼近百分之百。而实验室里几十年反复出现的读数是：总投入常常系统性地超过奖品价值。

这个现象有专门的名字（过度耗散），有专门的综述文献，

也有一整排解释——赢的快感、风险偏好、有限理性、人数带来的判断偏差。近年的理论工作进一步指出，当奖品较小时观察到过度耗散、奖品很大时反而观察到耗散不足，两者可以在同一个偏好设定下被同时解释。

被丢掉的那一眼

请注意这件事在它自己的领域里回答的是哪个问题：它回答的是“参与者的效用函数或理性程度该怎么设定”。它被当作行为侧的异常处理。

而把它放到驱动方向这个问题上，它说的是另一件事：如果真是奖品在驱动投入，投入就不可能系统性地超过奖品。

超过了，就说明有一个奖品之外的东西也在驱动。这个东西是什么，竞赛理论没有字段可以

安放，因为它每一轮结清、历史不进入状态。

5.2 第五步的告诫：约束理论自己的方法论约束理论的五步里，第五步的完整表述是：如果前面几步把一个约束打破了，回到第一步——但不要让惯性成为系统的约束。

这句话在管理教学里被当作一条提醒：别停下来，别僵化，持续改善。而它字面上说的是一件更强的事：上一轮为服从旧约束而做的那些安排，本身会成为下一轮的约束。

也就是说，约束不是外面给定的地

形，它有一部分是这套方法自己造出来的。

这一条在它自己的领域里回答的问题是"改善为什么会停下来"。它从来没有被当作"约束是被生产出来的"这个命题的证据——因为在这套方法里，约束必须是可指认的（某台机器、某个环节），才能被用尽、被服从、被抬高。一个不落在任何环节上的约束，在这套方法里不叫约束，叫外部环境。

5.3 日常空间总是赢：价值化理论自己的前提格罗伊斯本人写下过这套体系的成立条件：新的经济学预设名录与日常世界之间有一条稳定的、能守住的界线，这条界线必须持久，文化创新的经济才能运转。他在另一处更直白地说：在日常空间与收藏制度的关系里，赢的永远是日常空间。

这两句话合起来，等于承认这套体系的驱动方向有一个成立范围：只有当界线守得住时，才是名录在决定什么算新。

界线一旦守不住——每个人都有自己的收藏、每个人的名录都不一样、日常世界不再是那个

无结构的无限储备——名录就不再有权单方面定线。它这时会做的事不是扩容，也不是提高标准，而是放弃：把判定外包给自动过滤、抽签、流量，或者干脆不再宣称自己在判定。

三件材料指向同一件事：驱动的方向不是给定的。而一旦它不是给定的，那个"在三家账上都不算一样东西"的东西，就有可能是这一轮真正的发起处。

■ 六、暗基线

6.1 落判件：同一样东西的三个名字本章把"被判定过、而没有被接受、因而没有进入任何正式记录"的产出统称为落判件。桌拒的稿子、被筛掉的简历、被拒绝合并的变更、没入围的投标、被否掉的方案、没进复评的作品，都是。

落判件有一个被三家一致忽略的性质：它被看过。

被丢掉的那一眼

这一点极重要。它不是没有发生的事，不是没写出来的东西。它进入过判定通道，占用过判定者的一段注意力，并在判定者那里留下了一次比较——判定者读它、把它与手上正在形成的水准做了一次对照、然后把它放到一边。放到一边的是那件东西，不是那次对照。

那次对照留下来了。

6.2 定义暗基线：一次判定实际使用的参照集合中，不由任何正式记录构成的那一部分。

在任何一个是择优的场合，判定者名义上使用的参照是写下来的标准（评审细则、岗位要求、准入条件、评分表）。而实际使用的参照，是他手上这一批与最近若干批全部看过的东西所形成的分布——其中绝大多数是落判件。合格线不是从标准里读出来的，是从这个分布里切出来的。

暗基线因此有三条性质：它没有作者。

没有任何人抬高了它。每一个判定者都在诚实地执行同一份细则；每一个投件者都在合法地投件；没有任何一步违规。抬高发生在没有人做决定的地方。

它单向。

落判件抬高分布的上沿，而它的生产者一无所得。这与位置竞争不同：位置竞争里双方都在

跑，一方放缓，门槛回落；暗基线里落判者已经出局，他留下的那一层不随他离开而撤走。

它不入账。

没有任何字段记录"本轮合格线中由落判件贡献的那一部分"。落判者不会投诉，因为他只是

没被选中；机构不会自查，因为标准提高在任何治理语言里都是成绩。

6.3 两轴四格第一轴：这件产出有没有进入正式记录（被接受、被记账、有归属）。第二轴：它有没有改变下一轮判定所使用的参照。

改变了下一轮的参照

没有改变下一轮的参照

进入了正式

甲格：被接受的作品定义新水准。三家都看得

乙格：归档而不构成参照。例行入库、

记录

见，三家都有理论。

封存材料。

没有进入正式记录

丙格：靶格。没被接受，却抬高了线。

丁格：真正的浪费。竞赛理论以为落判件全在这一格。

要打的是丙格。它满足一个真靶格的三条签名。

其一，它在短期指标上表现为改善。

桌拒率上升被读作把关变严；录取者的平均资历上升被读作质量提

高；筛选周期缩短被读作效率提升。全部形式审查都能通过——没有人违规，流程都走了，记录都在。

其二，它的失效不产生任何信号。

一个本该被接受而被丙格挡下的产出，不会在任何地方留下痕迹：它

没有编号，没有申诉通道，没有一个字段叫"因基线抬升而误拒"。系统对自己的误差是盲的，而且这种盲不是记录不全，是没有可记的对象。

其三，它自我加固。

被丢掉的那一眼

筛选越自动化、流程越规范，被人认真看过的比例越低；而被认真看过的比例越

低，判定者手上用于切线的分布就越是由粗看过的落判件构成，线抬得越快。规范化在这里不是解药，是加速器。

6.4 一句不含判断词的问话本章的判据可以拿去问日志、问决定记录、问评分表，不必问任何人的看法：>把上一轮全部落判件从参照集里移出去，这一轮的合格线会降多少？

在五个具体场合跑一遍，五个答案互不相同，其中至少两个不能从命题本身推出来，必须去查：其一，学术期刊。把去年的桌拒稿从编辑的参照集里移出去——做法是让同一批编辑在盲掉年份的条件下，重判五年前的一批桌拒稿。若当年判为"边缘、可送审"的稿件，在今天的判准下仍被判为桌拒，线就是抬高了。这个数在期刊的决定日志里已经存在。

其二，招聘。把落选简历从筛选基线里移出去——做法是只用去年实际录用者的材料训练或校准筛选，不用二百五十四份里落选的那二百五十三份。门槛会降多少？

其三，代码评审。把被拒绝合并的变更从评审的参照中移出去——只参照已合并的历史。通过率会升多少？

其四，评奖与展览。评委只看馆藏与历届获奖，不看这三年的落选投稿。"新"的判定会松多少？

其五，课堂评分。把低分卷从评分基线里移出去，只用高分卷校准。这一格的答案是反的：分数分布会变紧而不是变松，因为在逐件评分的场合，低分卷把基线往下拉，不往上推。

第五个答案是查出来的，不是想出来的，而且它是本章最重要的一处收窄——它指出暗基线不是在所有判定场合都发生的。第九节专门处理它。

6.5 暗基线抬升度把上面那句问话变成一个可比较的数： $\text{暗基线抬升度} = (\text{把上一轮全部落判件移出参照集之后，本轮合格线下降的幅度}) \div (\text{本轮合格线})$ 这个读数有四条性质，正是它可以跨领域使用的原因。

无量纲。

期刊那里合格线是一个综合评级，招聘那里是一个筛选分，工程那里是一个通过阈值，量纲各

不相同，而比率可以并排比较。

可事后清点。

它不需要新建数据。期刊有全部投稿与决定的记录，招聘系统有全部申请与评分的记录，代码托管有全部提交与评审的记录。所需要的只是一次盲掉年份的重判。

它把一个印象变成一个数。"现在要求高多了"是一句感慨；"本轮合格线中有百分之十九来自落判件"是一个可以争论、可以复核、可以被推翻的读数。

被丢掉的那一眼

它与既有的主要指标正交。

这一条最容易漏，也最要命。既有的主要指标是录取率、通过率、吞吐、周

期。一个把招聘吞吐做到极致的机构——到岗时间最短、录用数最多——可以有极高的暗基线抬升度，恰恰因为它筛得快、筛得多、每份材料被看的时间最短。反过来，一个录取率极低的资质考试可以有零抬升度：只要它每年用同一套绝对标准，不参照落选者的答卷。举不出"既有指标最好看而这个数最难看"的实例，这个读数就该重造；这个实例是现成的。

■ 七、这一轮是怎么走的

三条驱动不是三支互不相干的箭头，它们构成一个有次序的循环。把二〇二二年底以来这一轮走过的路排出来，可以看到发起处换了两次手。

第一步，成本先动。

生成一份看起来合格的材料的单位成本在很短时间内坍塌。这一步动的是承接条件的一侧，不是规则的一侧——名额没变，规则没变。

第二步，产出形态跟上，滞后约一到两个周期。

投稿量涨四成，申请量涨四倍，每天有九成开发者在用

工具写代码。注意跟上的不是"更好的东西"，是"更多的、彼此更像的东西"——那份短篇故事实验测到的相似度上升，与招聘方说的"申请开始长得都一样"，是同一件事的两处读数。

第三步，工作路径被迫改变。

桌拒率上升、机器预筛上线、评审规程加严、面试改回线下。这是被逼出

来的组织反应，不是谁的战略选择。

第四步，这一步是前三家都没有写的：改完的路径反过来改动了承接条件本身。

桌拒不是把东西扔掉，

是把它粗看一遍再扔掉——粗看的那一眼进入了判定者的分布。预筛不是把线固定住，是把线交给一个以历史全量（含落判件）为基准的装置。于是判定所参照的那个东西被换了：从"写下来的标准"换成"最近看过的一切"。

下一轮由谁先动？本章的预测是：发起处将从成本一侧换到判定一侧。也就是说，下一轮的显著变化不会由"又出现了更便宜的生成"引起，而会由机构改动判定制度引起——限投、绑定身份、预注册、按人限额、参照集轮换。这个预测是可裁决的：如果二〇二六到二〇二八年间，各行当申请量与门槛的下次台阶式变化仍然由生成成本的又一次坍塌引起，而不是由判定制度的变动引起，本章这一段就错了。

顺序预测比相关预测难得多，这是本章交出的最容易被推翻的一段。

■ 八、十三个行当

一条判断如果只在提出它的那个场景里成立，那是个例。下面是逐项兑现，其中两处反过来给本章加了约束，第十一节专门交代。

被丢掉的那一眼

其一，学术出版。

已在前文交代。补一条：某大型出版集团二〇二五年收到约四百二十万份投稿，发表约七十九万五千份。三百四十万件落判件被看过至少一眼。按本章的读法，它们不是三百四十万次损耗，它们是下一年编辑分布的主要构成。

其二，会议评审。

某机器学习顶会二〇二六年产生七万六千余份评审意见，其中约两成被判定为完全由

机器生成，超过一半有机器参与。这里出现了本章机制的一个更狠的形态：连那次对照本身都被外包了。

当落判件由装置判定、装置又以历史全量为基准时，线的上升不再经过任何人的头脑，也就不再有任何人可以出来说"我觉得今年要求太高了"。

其三，招聘。

已在前文交代。这里给出与最接近的那个说法的第一处分岔：信用膨胀理论预测，门槛上升发生在学位持有者数量上升时。而二〇二五至二〇二六年的实况是，持有相应资历的人数没有涨四倍，涨四倍的是每人投出的件数。门槛在合格者数量不变时上升了。

其四，软件交付。

吞吐上升、不稳定上升、返工上升。按约束理论，撑开评审这一段应当让线回落；实况是评审规程与门禁在加严，而不是放松。一个可查的读数：一个团队被拒绝合并的变更占全部提交的比例上升之后，其后续提交的平均评审轮次是否上升——若上升且与被拒比例呈剂量关系，本章成立。

其五，风险投资。

商业计划书的生成成本坍塌之后，机构收到的件数暴涨。可查的读数：一家基金今年的"值得开会"门槛，与去年落判项目的数量之间是否存在剂量关系（控制该基金管理规模与赛道热度）。

其六，文学出版与代理。

自投稿堆一向是这个行当最诚实的落判池。可查的读数：代理今年愿意继续读下去的开篇页数与去年自投稿量之间的关系。

其七，专利审查。

申请量上升导致的不是审查标准写死不变，而是审查员对"显而易见"的实际把握在移动。可查的读数：同一技术分类下，被以显而易见为由驳回的申请数量与随后一年的授权门槛之间的关系。

其八，药品与器械审评。

这是本章最必须谨慎的一格，也是伦理边界所在，第十五节专门处理。此处只指出：审评是定标制而不是择优制——不设名额，达到标准即通过。按本章第九节的判断，暗基线在这里不发生。这不是本章的谦辞，是本章的可证伪点：若在纯定标的审评通道里也测到抬升度显著为正，本章的适用范围就错了。

其九，高校招生。

申请件数暴涨而学位数不变，是本章机制的教科书场合。可查的读数：某校某专业的实际录取分位与该专业当年落选申请数量之间的关系（控制招生名额与报考人数）。

其十，应用商店审核。

提交量暴涨，审核规则文本几乎不变，而实际过审率下降。规则不变而线在动，这一格最干净地把"写下来的标准"与"实际使用的参照"分开了。

其十一，政府采购与招标。

标书生成成本坍塌之后，评标委员会面对的分布整个右移。可查的读数：同类项目的技术分实际切分点逐年变化，与当年废标数量的关系。

其十二，众包设计竞赛。

这个行当天然是全付出竞赛，落判率极高。它是最容易做前瞻实验的场合：同一平台的多个赛事，一半的评委只看历届获奖作品，一半的评委看全部投稿，比较两组的实际获奖门槛。

其十三，日常创作平台。

被丢掉的那一眼

这里出现本章机制的极端形态：判定者是推荐系统，落判件是没有被推出去的内容。而推荐系统的基线以全量历史为参照，所以抬升是逐日、逐小时发生的，不经过任何一次人的对照。这一格同时给出本章的一条反噬预告：当判定完全自动化之后，暗基线抬升度这个读数本身会变得不可测，因为"移出参照集再重判一次"这个动作，需要一个可以被重置的判定者，而在线学习的装置没有这样的状态。

十三个行当里，第八与第十三反过来削弱了本章，第十一节交代。

■ 九、择优制与定标制：抬升只在其中一种里发生

第六节第四小节那个反向答案，指向本章最重要的一处收窄。

在择优制里，判定者要从一批东西里挑出前若干个。他名义上按标准打分，实际上按分布切线——因为名额是硬的，而分布是他手上唯一能用来确定"前若干个"的东西。这时每一件落判件都参与了分布的构成，因此参与了切线。

在定标制里，判定者要判断这一件够不够格。够格的都通过，不够格的都不通过，通过多少个不重要。这时分布不参与判定，落判件不构成参照，暗基线不合成。

两者在制度文本上常常长得一样——都有一份标准、一套流程、一批判定者。区别只在一句话：通过的数量是被先定死的，还是由达标与否决定的。

这条区分有三个直接后果。

第一，它解释了课堂评分那个反向答案：期末考是定标制（分数由答卷与标准的距离决定），而奖学金评选是择优制（名额固定）。同一批学生，同一批老师，两套判定，只有后者会积累暗基线。

第二，它给出本章的干预点。既然抬升发生在切线的那一步，而切线只在名额固定时才必须依赖分布，那么把择优改成定标，就切断了抬升的通路。这条出路三家单独都给不出：竞赛经济学会说改名额只是改参数（它无法预测这个改动会取消掉一整类效应）；价值化理论会说定标制不可能，因为“新”本质上是比较关系（本章承认这一点，见下一节的收窄）；约束理论会说这是在放弃对瓶颈的管理。

第三，它解释了为什么“提高标准”这个说法在择优制里几乎总是幻觉。在择优制里没有人提高标准——标准文本一个字都没改。改变的是分布，而分布不是任何人的主张。所以“我们的要求提高了”这句话，在择优制里从来找不到一个可以为它负责的人。门槛没有作者。

■ 十、与最近的几种说法的分界

一条判断能不能站住，看的不是它有没有道理，而是它能不能与最像它的那些说法在同一个案例上给出相反的判决。下面每一条都写成同一个形状：观察到什么，本章对而它错；观察到什么，反之。

被丢掉的那一眼

内卷化（人类学，格尔茨，一九六三）。

它说到哪一步：投入不断增加，内部不断精细化，而结构不

变、无发展。分界：格尔茨的结构是不动的，本章的线是动的，而且单调上升。判定形式：若观察到投入上升而合格线不变，则内卷化对而本章错；若观察到合格线随落判件数量上升而上升（控制合格者人数），则反之。

信用膨胀（社会学，柯林斯，一九七九）。

它说到哪一步：教育扩张使文凭供给增加，用人方随之提高

学历要求，形成螺旋。这是本章最近的邻居，也是最容易被混为一谈的一位。分界：柯林斯的螺旋跑在被授予的文凭上——机制要求那个东西先被颁发出来、成为一份可见的存量；本章的机制跑在未被授予的那一侧，它要求的只是被看过一眼。判定形式：若合格线只在持证者数量上升时上升，则信用膨胀对；若在持证者数量不变、而投件量激增时合格线仍然上升，则本章对。

二〇二五至二〇二六年的招聘实况正好

落在后一种情形，这是本章手上最有力的一条经验证据。

位置性商品与赢家通吃（经济学，赫希一九七六；弗兰克与库克一九九五）；红皇后（演化生物学，范瓦伦一九七三）。

它们说到哪一步：有些好处的价值来自相对位次，因此人人加大投入而人人位次不变，

是一场没有赢家的军备竞赛。分界：军备竞赛是双向对称、可回落的——对手放缓，我方压力随之下降；暗基线是单向不可回落的，抬高它的那些人已经出局，他们留下的那一层不随他们离开而撤走。判定形式：若投件量下降之后合格线在一到两个周期内回落，则军备竞赛对；若投件量下降而合格线不回落，则本章对。

杰文斯悖论（经济学，一八六五）。

它说到哪一步：效率提高使单位成本下降，总消耗反而上升。分

界：杰文斯讲的是被消耗的资源总量；本章讲的是被接受的绝对数量下降，不只是比例下降。判定形式：若生成成本坍塌之后，被认真评估过的绝对件数上升，则这是杰文斯型扩张；若绝对件数下降（评审邀请五封才成一封、桌拒从百分之十三到百分之四十），则本章对。

注意力经济（西蒙，一九七一）。

它说到哪一步：信息的丰裕造成注意力的贫乏，注意力才是稀缺物。

这是本章第二危险的邻居，因为它同样把“看不过来”放在中心。分界只有一处，而这一处是本章最锋利的分离线：西蒙的框架里，未被注意的东西没有影响——它只是没有被处理；本章说未被接受的东西仍然抬高了线。

判定形式：把落判件移出参照集之后，若合格线不变，则注意力经济对而本章错；若合格线下

降，则本章对。这两句话在同一份日志上给出相反的读数，一次重判实验即可分开。

柠檬市场（经济学，阿克洛夫，一九七〇）。这是最强的竞争解释，也就是那句最难反驳的“其实不就是信息不对称吗”。它说到哪一步：质量不可观察时，买方按平均出价，优质卖方退出，市场质量下滑。分界：柠檬市场预测优质供给退出；本章预测优质供给照旧进来，只是线更高。判定形式：若在申请量暴涨的岗位上，高资历申请者的绝对数量下降，则柠檬市场对；若高资历申请者数量不变或上升而门槛仍上升，则本章对。

这一条必须单列，因为一个判断可以打赢所有竞争解释，却仍然只是某个既有概念的改名——所以竞争解释归证伪那一节，竞争概念归这一节，两件事各管各的。

算法单一化与模型退化（计算机科学，克莱因伯格与拉加万二〇二一；舒麦洛夫等二〇二四）。

它们说

到哪一步：同一套装置被普遍用于本来由许多人分头做的判断，结果趋同；用自身输出训练自身，分布逐代退化。分界：这两条讲的是产出分布的同质化与退化；本章讲的是判定基线的抬升，而基线抬升可以在

被丢掉的那一眼

多样性不变时发生。判定形式：若同质化指标上升而合格线不变，则单一化对；若合格线上升而同质化指标不变，则本章对。

目标置换与度量失效（社会学，默顿一九四〇；古德哈特）。

剥去本行业的名词之后，本章命题的形式

是"被丢弃的东西反过来成为约束"，不是"一个指标一旦成为目标就不再是好指标"，也不是"手段吞噬目的"。最贴的上游形式其实来自另一处：外部性与污染（庇古一九二〇；科斯一九六〇）——生产的副产物排到共享环境里，抬高了所有人的成本，而没有任何账本记录它。这一位必须点名，因为它几乎就是本章的骨架。分界：污染排到环境，环境是消极载体，可以治理、可以清除、可以逆转；暗基线排到判定者的参照里，而清除它需要判定者承认"我们过去的标准更低"，没有任何机构能够做出这个承认。判定形式：若把落判件的记录整体删除之后合格线回落，则这是污染型（可逆），治理技术适用；若删除记录而合格线不回落，则本章对，因为那一层已经不在记录里，而在参照里。

这条判定同时是本章第十二节那

条反噬预告的来源。

约束理论（工程学，戈德拉特，一九八四）。

作为方法上而非内容上最接近的一个说法，它必须单列。

它说到哪一步：吞吐由最窄的一段决定，找出它、撑开它。分界：约束理论的约束在系统里有位置，可指认、可服从、可抬高；暗基线不落在任何一段上，抬高它的不是任何一个环节的容量，而是每一次判定所用的参照。判定形式：若扩容评审人手之后合格线回落，则约束理论对；若扩容之后合格线继续上升，则本章对。交付研究那份数据倾向于后者，但它不是判决性的——那份数据没有分离"扩容"与"未扩容"的可比单元，因此第十四节给出了一个同质多单元的设计顶上。

重投偏见（评审研究，斯捷尔马赫、沙阿与辛格，二〇二一）。

已经有实验证明：知道一篇稿子此前被

拒过，新手评审会打得更严。分界：那是同一篇的重投历史在起作用；本章说的是别人的落判件抬高所有人的线。判定形式：若在完全盲掉重投信息、且全为首投的新件上，合格线仍随上一轮落判量上升，则本章对；若盲掉重投信息之后效应消失，则那只是重投偏见。

本书的第六章。

本书有六篇与本章最近，其中三篇与本章问题、同取这三个学科、同日成文，必须逐条分清；它们与本章互为最接近的说法。

第三章《这道关拦下过谁》一篇提出：生成成本对所有人同时塌到接近零之后，正式门槛会整体饱和——人人都能满足它，于是它不再分辨；分辨不是往上迁到更高一级的门槛，而是整体迁出可读区，落到那些不能被生产、只能由会被追究的位置去承担的凭据上。它的最危险一格叫空关：一道最近一百次判定里一次也没拦下过谁的关。

它与本章表面上相反：它说通过率上升，本章说合格线上升。两句话可以同时为真，而且正是同时为真才有意思——饱和的是写下来的那道关，上移的是没有写下来的那条参照。

本章因此可以充当那一章留下

的问题的一个候选答案：分辨迁出可读区之后，去了哪里？那一章答的是迁到不可生产的凭据上；本章答的是，其中一部分根本没有迁到任何凭据上，而是迁进了判定者手上那份由落判件构成的分布——它不是凭据，谁也生产不出来，也无人可以为它负责。

被丢掉的那一眼

判定形式：在一道已经饱和的关上，若分辨确实迁到了新的、可点名的凭据上（新增资历、新增证明、新增担保位置），则那一章足以解释；若没有任何新凭据出现、正式标准一字未改，而实际切分点仍在上升，则本章对。

顺带交代一处措辞上的撞车：那一章在解释“饱和的关为什么不会被撤下”时，用过“门槛没有作者”这句话。本章原本以它为题，现已让开；这句话在本章里指的是另一件事——不是没有人负责撤下它，而是没有人抬高它。两者说的是同一个观察在两个位置上的两种后果。

第一章《两种「不」》一篇与本章的关系最紧，紧到两篇合起来才把同一堆东西分完。它提出：一次否决可以是有人读完之后作出的，也可以是没有任何人打开过那份东西就作出的，而这两者在所有账本上同码；它把后者命名为未阅否决，给出的读数是未阅率。

分界只有一句，而这一句是干净的：本章的机制要求那件东西被看过一眼，那一章的机制要求它没有被打开过。

本章第六节第一小节已经写死这个条件——落判件之所以能进入判定者的参照，全靠那一次对

照；一份没有被任何人打开的材料，不产生对照，对本章的机制贡献为零。于是两篇不是竞争关系，而是把落判件这一堆恰好切成互斥的两半：未阅率量的是没人打开的那一半，本章量的是被打开过的那一半所留下的东西。

更要紧的是，两条机制在同一个通道里是此消彼长的。未阅率越高，进入参照的对照越少，线就越不是被抬高的，而是变得任意——判定不再有基线，只有噪声。未阅率越低而件数越大，本章的机制越强，线上移得越快。这给出一条联合的判定形式：若一个通道的未阅率很高而合格线仍在稳定上移，则本章错，因为没有人打开过的东西不可能进入参照；若未阅率上升而合格线的上移随之停滞、判定转为不可预测，则两条机制按此处所说的方式交接。这条检验同时能判两篇，而两篇都可能因它而输。

第四章《认领物》一篇与本章同问同源：同样处理生成成本坍塌之后的困境，同样取经济学、艺术学、工程学三家，同样判定三家共同假定了一件没有说出的事。它给出的答案是：一件东西的生产者被三家默认为给定的，因而「是谁做的」不需要被单独生产；而实际上，为使一件产出可被归属于某个人，必须与它一并生产出另一层东西——它称之为认领物，并指出三家的账本只给它成本位、不给产出位。由此它把内卷重新定义为：不是投入增加而产出不增加，而是产出的构成里认领物的占比在持续上升。

两篇的分界是干净的，因为两处缺口落在通道的两端。认领物是生产者一侧的账面缺口——同一件交出去的东西里，有多大比例是为了证明它出自谁手；暗基线是判定者一侧的账面缺口——同一次判定所用的参照里，有多大比例来自没有被接受的那些东西。前者解释每一件为什么变厚，后者解释那条线为什么变高。两者可以独立发生：一个完全实名、归属零成本的系统里认领占比为零，而只要名额固定、落判件仍进入判定者的参照，合格线照样上升；反过来，一个不设名额的定标制通道里暗基线不合成，而认领物照样要生产。

判定形式：在一个归属已经零成本的通道里（强身份绑定、自动溯源、来源链完备），若合格线仍随落判量上升，则本章对而认领占比不足以解释；若认领占比归零之后合格线随之停止上升，则认领物足以解释，本章多余。

被丢掉的那一眼

更值得写下来的是两者的合成，因为它不是相加而是相乘：认领占比上升，使每一件的判读成本上升；判读成本上升，压低被认真接住的比例；被认真接住的比例下降，判定者手上用于切线的分布就更多地由粗看过的落判件构成，线抬得更快；线抬得更快，生产者又必须追加更多认领物才能越过。两篇各写了这个回路的一半，合起来才是一整圈。这是本章与它最要紧的关系，也是本章不主张取代它的理由。

《回程》一篇提出，决定一个系统能否重新开始的，不是留下了多少存量，而是当期被记为零的余项能否经由一条可指认的通道回到下一轮的起点。骨架与本章极像——都在处理“被记为零的那部分后来怎么了”。分界是干净的，而且是互补关系：《回程》问的是那些余项能不能回到自己的下一轮，作为能力；本章问的是那些回不去的余项如何落到别人的判定里，作为约束。同一份材料，两篇给出相反的评价：一篇被期刊桌拒且此后再未投出的稿子，在《回程》的读法里是回程断裂、再生失败；在本章的读法里，它成功地改变了一样东西——它抬高了线。判定形式：若把一批落判件的通道打通、使其回到原作者的下一轮，而下一轮的合格线不因此变化，则《回程》的机制独立成立、本章的机制在这批材料上不发生；若合格线随之变化，则两个机制在同一批材料上同时运行，且方向相反。

《入判份额》一篇提出，一个人能改下一轮评价标准的程度，等于他上一轮的产出进入他人判据的份额；没有出现在别人的判据里，给他多少修改权都没用。本章与它有一处正面的张力：本章主张，整批落判件从不出现在任何具名判据里，却整体抬高了判准。两者的调和不是措辞上的，而是结构上的：入判份额度量的是可归属的个体影响，暗基线度量的是不可归属的整体影响。

一个人的份额可以是零，而他所属的

那一类的效应可以很大。这不是对那一章的反驳，是对它的一处补足——它同时指出了判准可以在没有任何人取得修改权的情况下改变的那条唯一通道。判定形式：若判准的变化总能被追溯到某些具名产出进入他人判据，则入判份额足以解释，本章多余；若判准发生了可测的变化而全部具名产出的份额都没有相应变化，则本章对。

《轮空》一篇提出，长期停滞的第四种原因是一次换手没有完成——上一轮的发起处已经退出，按顺序应当接手的那一处没有动。它与本章同为“什么驱动什么”的问题，也同样用到了竞赛与劳动经济学这一家。

分界：《轮空》说的是该动的那一轮没有人动；本章说的是所有人都动了，交出来的东西被丢掉了，而丢掉的那部分仍在起作用。两者的读数正交：一个系统可以有完美的换手（发起处每轮都在换）而同时在积累暗基线；也可以零换手而零抬升。判定形式：若一个系统的发起处逐轮更替、而合格线仍随落判量上升，则两条机制独立，本章不被《轮空》覆盖；若合格线的上升总是伴随发起处不变，则本章只是那一章的一个特例。

最近的邻居是哪一个？是信用膨胀。本章之所以仍不能被它吸收，只有一个理由：它的机制要求那个东西先被颁发出来，本章的机制只要求那个东西被看过一眼。

二〇二五至二〇二六年的招聘与投稿实况，恰好把这两个条件分开了。

划完之后还有一层可以取。逐个问上面这些邻居"它结构性地看不见什么"：竞赛理论把每轮同构当作给定，因此看不见跨轮残留；价值化理论把名录之外的世界当作无结构的给定，因此看不见落判件之间的秩序；约束理论把约束的可指认性当作给定，因此看不见不落在任何环节上的约束；注意力经济把"未被注意即无影响"当作给定，因此看不见那一眼留下的对照；污染治理把载体的消极性当作给定，因此看不见排到判断者身上的排放物。这几行背后是同一块地：所有人都假定，一样东西要么被接受、要么就退出系

统。

被丢掉的那一眼

承认落判件既没被接受、也没有退出，上面每一家都会在自己的框架里出现一处不一致——这正

是"没有人辨识过它"的真正来源：看的人都站在它上面。

■ 十一、三处反过来削弱本章的约束

其一，工程学收窄了本章的适用范围。

交付研究那份数据里，人工智能与吞吐的关系已经由负转正——

东西确实出得更多了。这削掉了本章原本会写下的一句更强的话：生成成本坍塌使系统的总产出下降。这句话不成立。下降的是稳定性，以及被接受的比例，不是总产出。因此本章的命题从"所有产出系统"收缩到"输出必须经过人为判定才算数的系统"——招聘、评审、评奖、准入、审批。直接由市场或物理世界验收的系统（卖出去就算数、跑起来就算数）不在本章范围内。这一处动的是适用范围本身。

其二，价值化理论收窄了本章的价值排序。

格罗伊斯自己说，在日常空间与收藏制度的对抗里赢的永远

是日常空间。这削掉了本章原本会写下的另一句更强的话：暗基线单向抬升，因此门槛只会越来越高。这句话也不成立。落判量越过某个规模之后，判定制度不是把线抬得更高，而是放弃判定——改抽签、改流量、改全自动过滤、或者干脆宣称自己不再筛选。而判准解体比门槛抬高更坏：门槛高的时候，至少还有一条线可以争论；判准解体之后，连争论的对象都没有了。本章原本隐含"抬高门槛是坏事"，这一处把价值排序改了——抬升是次坏，解体才是最坏。

其三，竞赛经济学收窄了本章的强度。

过度耗散的实验读数不是单调的：奖品小的时候观察到过度耗

散，奖品大的时候反而观察到耗散不足。这削掉了本章原本会写下的第三句更强的话：生成成本趋零必然导致抬升加速。不必然。奖品越贵重（岗位越好、期刊越顶尖），参与者反而

越谨慎、投得越少、准备得越足。因此本章要加一条：抬升速度与奖品价值负相关——在最顶尖的那一层，抬升反而慢；在中间那一层最快。这一处削弱了本章对高端场合的解释力，也解释了为什么最痛的往往不是顶尖岗位，是中间岗位。

■ 十二、出路，以及它最省事的实现方式会毁掉什么

本章的干预点由第九节直接给出：抬升发生在按分布切线的那一步，而按分布切线只在名额先定死时才是必须的。因此出路有四条，按可行性排列。

第一条，把择优改成定标。

凡是可以不设名额的场合，就不设名额：达标即通过，不够即不通过，通过多少个由达标与否决定。这条最彻底，也最难——因为大多数择优制的名额并不是资源约束，而是身份约束（稀缺本身就是那个奖品）。

第二条，参照集轮换。

名额保留，但明文规定判定所使用的参照集只包含被接受的历史，不包含落判池；并定期更换判定者，使个人分布不能长期累积。这条最容易写进制度，成本最低，而且第十四节那条赌注押的就是它。

第三条，落判池隔离。

被丢掉的那一眼

把落判件从任何自动装置的训练与校准数据里移出去。这条只对已经自动化的判定通道有效，而那恰恰是抬升最快的通道。

第四条，给“被认真看过一次”这件事记账。本章全篇的机制都建立在“被看过一眼”上，而这一眼目前在所有系统里都不留痕。给它记账，抬升度就变成一个可以逐月监测的数。

四条都有一个共同的代价：它们都会让判定变慢、变贵。这不是可以绕开的，这是本章的判断所蕴含的——被认真接住的容量是真稀缺，而不是可以靠工具放大的东西。

现在说反噬。这条命题一旦被制度采用，最省事的实现方式恰恰会摧毁它要保护的那样东西。

监测抬升度需要落判池的记录。而一旦这个数进了考核，把它做好看的最省力做法不是切断通路，是删掉落判池——“本机构不保留未录用者的材料”。这在形式上完美达标：没有落判池可以移出，抬升度算不出来，也就永远不会难看。而机制照旧运行，因为那一层已经不在记录里，而在判定者的参照里。第十节那条关于污染的判定形式，说的正是这件事：删掉记录而线不回落，就是本章对的时候。

于是最省事的合规方式，把唯一可以审计判断来源的那份材料销毁了。这不是一个待办事项，我看不到出口。

■ 十三、不适用的边界

本章的判断在下面几种情形里不成立，写清楚是为了让它可以被证伪，而不是为了自保。定制制不适用。

已在第九节交代。药品与器械审评、资质考试、安全认证这一类不设名额、达标即过的通道，暗基线不合成。若在这些通道里测到抬升度显著为正，本章的核心区分就错了。

判定者不留存分布的场合不适用。

如果每一次判定由互不相干的人独立完成、且此人此前没有看过同类材料，则没有分布可累积。一次性的、完全轮换的评审接近这种情形。

能力不足型困境不适用。

一个机构如果本来就判不准（判定者不胜任、标准写得含混），它的问题是判定质量，不是判定基线。给这类机构套用本章的处方，会把一个本来该补课的问题误诊成一个该做隔离的问题。

产出可被外部直接验收的场合不适用。已在第十一节交代。

极低落判率的场合不适用。

落判件占比很低时，分布主要由被接受者构成，此时甲格的机制（被接受的作品定义新水准）占主导，本章说的丙格效应可以忽略。粗略地说，落判率低于二分之一时不必用本章的框架。

被丢掉的那一眼

■ 十四、证伪条件与一条写死日期的赌注

先写出最强的竞争解释，再让证伪条件排除它。最难反驳的那句话是："其实不就是件数太多、判定者看不过来吗？"这是纯负荷解释：门槛上升是因为绝对工作量上升，与落判件是不是留下了什么无关。

主证伪条件：控制投件总量与判定者人时之后，若"暗基线抬升度"与合格线上升之间不存在剂量—反应关系，则本章为伪——届时门槛上升应归因于绝对负荷，而不是落判件构成的参照。

这个检验要求同质多单元，不能拿两个国家或两个行业充数。可行的设计是：同一出版集团旗下、同一学科、投稿量相近的八到十二家期刊，其中一部分实行编辑参照集只含已接受稿（做法可以是强制编辑轮换，或改用固定评分细则并禁止参照当期投稿分布），另一部分照旧；比较两组在两年内的合格线漂移。

样本在同一制度环境内，混淆变量同质，自变量有真实变异。

其余八条：其一，低成本、数据已存在：取任一家期刊二〇一九与二〇二五两年的桌拒稿，让同一批编辑在盲掉年份的条件下重判。若二〇一九年被判为"边缘、可送审"的稿件在今天被判桌拒，抬升成立；若两年的判定分布无差异，本章为伪。所需数据在多数编辑系统的既有记录里已经存在。

其二，若在合格者绝对数量不变、而投件量激增的行当里，合格线不上升，本章为伪（这一条同时判掉信用膨胀）。

其三，若把落判件从自动预筛的校准数据中移出后，该装置给出的通过阈值不下降，本章为伪。

其四，若投件量回落之后，合格线在一到两个周期内回到原位，本章为伪（这一条判掉不可逆性，本章将退化为军备竞赛的一个说法）。

其五，若在纯定标制通道（资质考试、安全认证）里测到抬升度显著为正，本章第九节的区分为伪。

其六，若高资历申请者的绝对数量随投件量上升而下降，则柠檬市场足以解释，本章多余。

其七，顺序预测：若二〇二六至二〇二八年间各行当门槛的下一台阶式变化，仍由生成成本的又一次坍塌先动引起，而不是由判定制度的变动先动引起，则本章第七节的次序判断为伪。这一条比相关性难被巧合满足。

其八，若一个系统的发起处逐轮更替、而合格线仍随落判量上升，则本章与《轮空》独立；若合格线上升总伴随发起处不变，则本章只是那一章的特例。

这条判断若被证伪，我们会学到什么？

若剂量—反应关系不成立，那么门槛上升就确实只是负荷问题，

处方回到扩容——而那意味着"被认真接住的容量"是可以靠资源放大的，这与本章对稀缺性质的判断相反，也会让第十二节那四条出路全部失效。证伪本身指向一个完全不同的干预方向，这正是它值得做的原因。

赌注。

被丢掉的那一眼

到二〇二八年十二月三十一日为止，至少有三家学术期刊或三家大型招聘平台，会公布带有参照

集轮换或落判池隔离性质的可执行条款。什么不算命中，写死在这里：只发布倡议、白皮书、伦理声明、研究报告，不算；只在博客或访谈里表示"我们正在考虑"，不算；必须是写进投稿须知、产品文档或内部审核办法、且规定了具体做法的条款。若到期没有出现，本章对干预可行性的判断就错了——机制可能仍然成立，但它不可被制度处理。

本章另交出两个自己解释不了的洞。其一，抬升度的量级本章给不出先验预测——它可能是百分之三，也可能是百分之三十，本章只能预测方向与剂量关系，不能预测幅度。其二，判定者个人的分布能保存多久、多少次判定之后会遗忘，本章完全不知道，而这个时间常数决定了参照集轮换要多久轮一次才有效。

■ 十五、伦理与证据边界

暗基线抬升度是一个比率，而比率几乎必然会被拿去考核人。因此三条必须写死。

其一，这个读数指的是位置，不是人。

它读的是某一个判定环节的参照结构，不是某一位编辑、某一位

招聘专员、某一位评审的品行或能力。用它去评价具体的人，是把一个关于制度的读数错当成一个关于个体的读数，而本章全篇的判断恰恰是这件事没有作者。

其二，不得为了把这个数做好看，而在安全关键的判定通道里制造标准波动。

药品与器械审评、飞行资

质、医疗执业资格、结构安全认证，这些通道本来就应当是定标制，本章预测抬升在其中不发生。若有人以"降低抬升度"为由，在这些通道里安排判定者轮换、放松参照、或人为制造标准变动，那是拿本章的读数去做一件本章明确反对的事。这些通道要的是标准稳定，不是基线流动。

其三，已经据此对某人做出不利安排的，必须公开编码规则并给出复核通道。

特别是：任何机构不得

以"此人历史落判次数高"为由对其加权或降级——本章的机制说的是落判件抬高了所有人的线，不是说落判者本人质量低。把一个整体效应读成个体属性，是这个读数最容易被滥用的方式，而且它会自我实现。

最后一条与第十二节的反噬相连：一个读数一旦进入考核，最省事的达标办法通常是在文书上做手脚——这里就是删掉落判池。这一条必须在制度设计时就被明写，而我看不到出口。

■ 十六、本章自身的位置

一条判断如果不能对它自己适用，它多半只是一句描述。

本章自己是一件交出去的产出。它落在两轴四格里的哪一格？它被接受了（它在这里），所以它在上那一行；它会不会改变下一轮的参照，取决于有没有人读它、用它。也就是说，本章属于甲格——被接受的作品定义新水准——而这一格恰恰是三家都看得见、都有理论的那一格。本章自己不是它所描述的那样东西的标本。

被丢掉的那一眼

更要紧的一处：本章所描述的机制，在本章自己的生产条件下不发生。

这个专栏不设名额，不做择优，

每一篇都发；按第九节的判断，这里不合成暗基线。所以本章缺少一个来自自身实践的检验样本，这是一个真缺口，不是谦辞。

还有第三处，也是最难堪的一处。本章由人与机器共同完成，本身就是生成成本坍塌的产物之一。如果这篇东西没有被接受，按本章自己的判断，它仍然会抬高一点点线——而这一点本章在原理上无法证实，因为没有任何人记录未被接受的稿子。

本章因此处在一个结构性的位置：它主张存在一类不被记账的东

西，而它自己一旦成为那一类，就同样不会被记账。它能拿出来的只有间接读数——重判实验、剂量关系、顺序预测。这不是方法上的懒惰，这是这条判断的形状所决定的：它所主张的那个东西，正是这套记录系统在设计上没有字段去记的那一类。

如果有一天有人能直接观测到它，那多半意味着有人已经先按第十二节第四条做了记账——而那时，这条判断也就不再需要本章了。

■ 十七、结论

生成一件东西的成本在很短时间里降到接近于零之后，三个互不引用的领域各自给出了一条关于什么在驱动什么的定律：投入由名额与成本驱动、什么算新由名录驱动、吞吐由瓶颈驱动。三条不能同时为真，而三条共同假定了同一件事——稀缺先在，行为随后；因此没有通过的那些产出，因为不在任何账上，不可能驱动任何事。

推翻这个假定的材料来自三家自己：竞赛实验里的过度耗散、约束理论第五步自己写下的告诫、价值化理论自己承认的成立条件。三件材料在各自的领域里回答的都是别的问题，从来没有人把它们拿来当作关于驱动方向的证据。

由此得到的判断是：内卷不是投入的浪费，不是新意的耗尽，也不是瓶颈的转移，而是未被接受的产出在无人记账处合成了一条判定基线，并单向抬升它。它没有作者，它单向，它不入账。它只在名额先定死的择优场合发生，不在达标即过的定标场合发生——这条区分同时给出了出路与其的代价：把择优改成定标、轮换参照集、隔离落判池、给"被认真看过一次"记账，四条都会让判定变慢变贵，因为被认真接住的容量是真稀缺，不是可以靠工具放大的东西。

而这套判断最省事的合规方式，恰好是把唯一能审计它的那份材料销毁。这一点没有出路，只能写在明处。

一句话可以拿去问任何一份决定日志：把上一轮全部没被接受的东西移出参照集，这一轮的线会降多少？

答案已经躺在那里，只是从来没有人取过。

被丢掉的那一眼

■ 注释

〔一〕本章所说的"落判件"限于被判定过而未被接受的产出，不包括从未提交的、被撤回的、以及因形式要件不全而未进入判定的。后三类不产生那一次对照，因此不参与基线合成。

〔二〕"合格线"在不同行当里的载体不同：期刊是编辑对"值得送审"的实际把握，招聘是筛选分位，工程是评审通过阈值。本章不主张它们可以直接比较，只主张第六节第五小节那个比率可以并排。

〔三〕第七节的次序判断是本章最容易被推翻的一段，也是本章最愿意被推翻的一段：顺序预测比相关预测难得多，而一条难被巧合满足的预测，比十条容易被满足的相关更值钱。

〔四〕第十节对本书六篇的处理，是本章自己的判据在本章身上的一次执行——本书其他各章与本章互为最接近的说法，不指名等于承认被它覆盖。其中第四章《认领物》第三章《这道关拦下过谁》第一章《两种「不」》三篇与本章同题、同取这三个学科、同日成文。四篇的缺口分处四个位置：生产者侧（一件产出里有多少是为了证明它出自谁手）、关口侧（这道关还拦不拦得下人）、未阅侧（这个不是谁读完之后作出的）、参照侧（本章：判定所用的参照里有多少来自没被接受的东西）。四篇任取其一单读，得到的都是一个回路的一角。

〔五〕本章全部经验材料均为公开来源，日期截至二〇二六年七月。本章没有自己做任何实验；第十四节给出的九条证伪条件全部是设计，不是结果。

〔六〕本次工作作废的方向共五条：其一，把困境归给"评审人力不足"（被第十节约束理论那一格判掉）；其二，把它归给内容同质化（被算法单一化那一格判掉）；其三，把落判件的作用写成"心理疲劳"（不可判定，且与负荷解释无法分离）；其四，把出路写成"提高透明度"（没有可执行形态）；其五，把基线抬升写成不可逆的单调过程（被第十一节第二处约束推翻，改为"次坏"与"最坏"的两段）。

第三章

这道关拦下过谁——门槛饱和与它的无作者状态

■ 摘要

人工智能把"生成一份看得过去的合格证据"的成本，对所有人同时压到了接近于零。随之出现的现象是：投递量暴涨、通过率上升、堵塞加剧、而录取比例不动。对这一现象，三门学科给出三条互不相容的出路。经济学把出路放在新的可计量产出上，认为通用目的技术的红利要等互补性投入积累够了才兑现；工程学把出路放在过程上，认为只要控制系统足够强，加速就不会带来失稳；艺术学把出路放在条件上，认为可复制的东西一律贬值，只有那些还要求人亲自到场的东西才留得住价值。三条处方在同一个具体处境里正面分岔，无法同时执行。

本章论证，三家共享一条从未被说出的假定：出路可以在一个终点上单独结算——只要自己那一样达成了，另外两样会跟上或不再重要。推翻这条假定的材料来自其中一家自己：生产率J曲线表明，在通用目的技术的实施期，被记作成本的那一项才是当期真正的产出，而被记作产出的那一项在当期系统性偏低。

终点与手段在账上换过位置，而三家的结算规则没有跟着换。

由此得到的判断是：人工智能时代的内卷，不是竞争过度，不是流程失控，也不是意义流失，而是可读门槛的整体饱和——当生成一份可被第三方读取的合格证据的成本对所有类型同时塌陷，分辨就不能再停留在任何可读的地方；它不会像资格通胀那样往上迁到更高一级的可读门槛，而是整体迁出可读区，落到那些不能被生产、只能由某个会被追究的位置去承担的凭据上。而由于没有任何位置被指定为门槛的作者——没有谁有权宣布一道关已经作废并为此负责——饱和的门槛不会被撤下，只会被继续满足。内卷就是这段"继续满足一条不再分辨任何人的界线"的时间里耗掉的全部劳动。

本章给出一句不含判断词的问话（这道关最近一百次判定里拦下过几次），一个三领域量纲不同而比率相同的读数（饱和关占比），一张两轴四格的辨别表并指认最危险的一格，一条有固定次序的出路及其插手点，十二处跨领域兑现，两处反过来把本章削小的约束，与十种既有说法的判决性分界（另与本书同日发表、同三学科的第四章《认领物》给出两处方向相反的

对照与一条可同时判两篇的检验），七条排除了最强竞争解释的判错条件，以及一条写死到二〇三〇年底的赌注。

可读门槛

门槛饱和

空关

分辨迁移

判定权属人工智能与内卷

■ 一·导论：三个行当在同一年里堵在同一个地方

二〇二五年，招聘平台领英上的投递量比上一年增加了四成五，平均每分钟一万一千份；招聘方对记者说，收到的简历"高度相似"，越来越难挑出真正合适的人。同一年，人工智能领域最大的几个学术会议里，有一个的投稿量从大约一万五千份涨到三万份以上；一本管理学期刊把二〇二二年底之前与之后各二十三个月的投稿量做了对比，后一段比前一段多出四成二，而同一份数据里，新冠期间的那次投稿高峰只

这道关拦下过谁

有两成。软件行业的年度调查则报告：采用人工智能的团队交付吞吐量上升、开发者自评的效能上升，而交付的稳定性持续下降；另一组遥测数据显示，合并请求在评审队列里的中位停留时间大幅拉长，同时有三成一的合并请求根本没有经过任何评审就被合入。

三个行当，三种材料，堵在同一个地方：进来的东西变多了，每一件都合格，而分辨谁该被接住这件事塌了。

这不是"竞争变激烈"的另一种说法。竞争激烈意味着有人被淘汰；这里的情形是，淘汰这个动作本身失去了依据。招聘方并不是在一千份合格简历里挑出最好的一百份——他们是在一千份彼此难以区分的合格简历面前，退回到了推荐、内部转介、以前共事过的人。期刊编辑并不是在更多的稿件里更精细地挑——他们是在扩大"编辑直接退稿"的比例，并把真正的判断压回到一个不公开、不记录的环节。软件团队并不是在更多的代码里更严格地评审——他们是在让越来越多的东西不经评审就合入。

三个行当同时发生的这件事，在既有的账本上没有一个字段。投递量的暴涨被记为"活跃度上升"，通过率的上升被记为"质量改善"，合并频率的上升被记为"吞吐提高"。每一个读数都是正号，而所有人都感到更堵。

问题因此要这样问：这堵在哪里？出路在哪一样上？本章先把三门学科各自给出的答案摆开——它们不是三种侧重不同的说法，是三条不能同时走的路。

■ 二·三家的位置：三条终点不同的出路

选这三家，不是因为它们离得远，是因为它们各自锁死了一个不同的终点，而且每一家都认为另两样只是手段。

2.1 经济学：出路在新的可计量产出经济学这一支的承重主张是：新技术的价值最终必须以可计量的产出兑现；在兑现之前的一切困难，都是实施期的现象。这一主张最完整的形式是生产率J曲线：像人工智能这样的通用目的技术，需要大量互补性的无形投入——组织重组、流程再造、技能培养、数据管线——而这些投入在国民账户里被记作成本，不被记作资本。于是新技术早期的生产率增长被系统性低估，等到无形投入开始收获，又被系统性高估。作者用美国数据做了一次调整，把与计算机硬件和软件相关的无形资产计入之后，到二〇一七年底的全要素生产率水平比官方口径高出百分之十五点九（Brynjolfsson, Rock & Syverson, 《美国经济杂志：宏观》2021）。

在这条主张里，路径是清楚的：条件先变（技术便宜了），然后分工重组，最后长出新的产出形态。

组

织重组是手段，产出是终点。堵塞的原因是终点还没长出来，处方因此是：加快互补性投入，尽早让新的可计价产出出现。经济学并不否认过程重要，也不否认价值观重要——它只是认为，那两样在产出出现之前无从判断，在产出出现之后不必再争。

这道关拦下过谁

2.2 工程学：出路在把过程做稳工程学这一支的承重主张是：你能设计的只有过程；产出是过程的读数，环境是过程的约束。

判断一件

事做成没有，看的是过程的变异是否被挤出去——同样的输入，是否可靠地得到同样的输出。

这一主张在人工智能这件事上给出了非常具体的读数。二〇二四年的行业年度调查报告：采用人工智能显著提升个人生产力、心流与工作满意度，但对交付稳定性与吞吐量产生负面影响。二〇二五年的调查更新了一半结论——吞吐量的关系转正了，人们学会了在哪里、什么时候、怎样用——而与交付稳定性的负相关关系依然存在（DORA, 《AI 辅助软件开发现状》2025）。报告自己的解释是：人工智能加速开发，而加速会把下游的弱点暴露出来；没有强健的控制系统——自动化测试、成熟的版本管理、快速反馈——变更量的上升就会变成不稳定。

同一年，一项随机对照试验给出了更硬的一笔材料：十六位有经验的开源开发者在自己熟悉的仓库里完成二百四十六个真实任务，每个任务随机分配"允许用人工智能"或"不允许"。开工前他们预计允许使用会把完成时间缩短两成四，做完之后他们估计缩短了两成；实测是延长了一成九（Becker,Rush,Barnes&Rein,METR,2025）。

工程学的处方因此是：先把控制系统建起来，再谈提速。产出是规格，条件是给定，要动的只有那条路。

2.3 艺术学：出路在改变什么才算值得艺术学这一支的承重主张是：一样东西值不值得做，由它所处的那一套鉴赏实践决定；物件本身与制作方式都是手段。

当代美学在人工智能问题上的一个代表性立场是：生成系统的产出之所以还能算作品，仅仅因为它部分地是一次人类行为的产物；而更要紧的问题不是"它们是什么"，而是"它们应当是什么"——也就是说，作品的地位不在物件里，在那套确立了什么算作品的实践里（Lopes,《美学与艺术批评杂志》2026）。

把这条主张放到本章的问题上，它给出的诊断相当锋利：可复制的东西一律贬值，而可读的东西一律可复制。

一份能被第三方读取并据以判定的证据，正因为它可读，才可以被生成。艺术学看得见这一点，因

为它已经在自己的领域里看了一百年——每一次复制技术的进步，都把价值从物件推向物件之外的那些条件：来源链、在场、谁经手过、在哪个场合被看见。

艺术学的处方因此是：不要在产出上找出路，也不要过程上找出路，要改的是"什么才算值得做"这套条件。产出与过程都会跟着条件走。

■ 三·冲突定位：三条处方在同一个处境里分岔

三条处方读起来都不算离谱，甚至有点互补的样子。把它们放进一个具体处境，互补就没了。

这道关拦下过谁

处境：一家中型公司的招聘。岗位三个，投递一千二百份，其中约九成通过了自动筛选的全部硬性条件——学历、年限、

■ 关键词

、格式。招聘负责人手上有一个季度的预算和两个人。

· 经济学的处方：这三个岗位的产出定义已经过时了。停下来重新定义岗位要交付什么，把评价改到新的产出上——比如不看简历看一个真实任务的完成情况。等新的产出口径立起来，筛选自然有依据。

· 工程学的处方：不要动定义，先把流程做稳。加一道结构化面试、一道标准化测评、一份评分表，把面试官之间的差异挤出去。评价的信度上来了，判断自然可靠。

· 艺术学的处方：前两条都会加速失效。你新定义的产出会在一个季度内被人用同样的工具做出来，你新加的测评会在一个季度内被针对训练。真正的出路是改变你在意什么——去看那些不能被生成的东西：这个人跟谁共过事、在什么难处上待过、有谁愿意为他背书。

三条不能同时做。第一条要求停下来重定义，第二条要求不许动定义，第三条要求放弃"可以用一份材料判定"这个前提。而且每一条在另两条看来都是病灶：· 在艺术学看来，工程学每天在做的那件事——把判断固化成可重复的规程——正是让门槛可被生成、因而必然失效的那件事。

· 在艺术学看来，经济学每天在做的那件事——把新东西变成可计价的产出——正是让分辨迁移到可读区、因而必然被廉价复制的那件事。

· 在经济学与工程学看来，艺术学的处方不是处方，是取消问题：把判断交给"谁认识谁"，既不可扩容，也不可申诉。

三家争的，说到底还是"出路"该放在哪一样上。而这个"该放在哪一样上"的争法本身，藏着一条三家都没说出口的假定。

■ 四·三家共同假定了什么，以及推翻它的材料来自哪一家

4.1 那条共有假定把三家的争论压成一句话：三家争的是出路该放在哪一个终点上结算。

经济学说结算在产出，工程学说

结算在过程，艺术学说结算在条件。

三家因此共同假定了一件事：>出路可以在一个终点上单独结算——只要那一样达成了，另外两样会跟上，或者不再重要。

这条假定念给三家听，三家都会说"这还用说吗"。它不是谁的主张，它是三家的主张之所以能成为主张的那个前提：如果不预设某一样是终点、另两样是手段，"出路在哪一样上"这个问题根本问不出来。

这道关拦下过谁

4.2 推翻它的那件材料，来自经济学自己推翻这条假定的材料不必从外面搬。它就在第一家自己的成果里，只是那件成果在自己的领域里回答的是另一个问题。

生产率J曲线要回答的问题是：为什么统计上看不到新技术的效果。它的答案是：因为在实施期，无形投入被记作成本，而它们所产生的资产没有被记作产出。

把这句话换一个方向读，得到的却是另一件事：>在实施期，被记作手段的那一项（组织重组）才是当期真正的产出；而被记作终点的那一项（可计量产出）在当期是一个系统性偏低的滞后读数。到了收获期，二者又反过来——同一个终点开始系统性偏高。

也就是说，终点与手段在账上换过位置，而且换过不止一次。这不是测量误差，这是J曲线这个命题本身的内容：它说的正是同一样东西在一段路的前半段是投入、在后半段是产出。

一旦承认这一点，三家的共有假定就站不住了。三家的错不在于选错了终点，而在于假定终点在整段路上是同一个。经济学的产出终点在实施期读不出该期最重要的东西；工程学的过程终点在生成工具普及之后不再是稀缺物——过程本身现在也可以被生成；艺术学的条件终点从来不能被直接瞄准，它只是前两样的结果。三条路各自都对，只是各自把自己那一段当成了全程。

经济学自己知道这件事，但它把这件事处理成"测量滞后"，不处理成"结算位置在移动"。这就是为什么它看得见J曲线，却看不见本章要说的那一条。

4.3 由此打开的问题如果终点会移动，那么真正要问的就不是"出路在哪一样上"，而是：>当一个终点因为达成成本塌陷而不再分辨任何人时，谁有权宣布它不再是终点？

这个问题在三家的框架里都问不出来。它也正是本章接下来要回答的那一个。

■ 五·承重判断：可读门槛的整体饱和

5.1 名与义先把两个词定死。

可读门槛：一道用可被第三方读取的证据来判定"够不够格"的关口。它的特征是判定依据以可传递的形式存在——一份简历、一次查重、一套格式规范、一组单元测试、一张证书、一个达标线。可读的好处是它可核查、可申诉、可扩容；可读的代价是它可被生成。

这道关拦下过谁

门槛的作者：一个有权宣布这道关作废或改写、并且要为此决定承担后果的具体位置。注意它不是"制定标准的人"——标准的制定往往有人负责，而标准的废止通常没有。一家公司有人签字定下"本岗位需五年经验"，却常常找不到任何一个位置有权在五年之后说"这一条已经不再分辨任何人了，撤掉"。

有了这两个词，承重判断可以写成一句：>人工智能时代的内卷，不是竞争过度，不是流程失控，也不是意义流失，而是可读门槛的整体饱和——当生成一份合格证据的成本对所有类型同时塌陷，分辨就不能再停留在任何可读的地方；而由于绝大多数门槛没有作者，饱和的门

槛不会被撤下，只会被继续满足。内卷是这段"继续满足一条不再分辨任何人的界线"的时间里耗掉的全部劳动。

三个被否定的说法，各来自一家：竞争过度是经济学的划界标准，流程失控是工程学的，意义流失是艺术学的。三条都描述了真实的现象，三条都把病灶放错了地方。

5.2 为什么它不像资格通胀那样往上迁一级最容易把本章的判断吸收掉的说法是资格通胀：门槛失效了，那就再设一个更高的门槛，历史上一向如此，学士变硕士，硕士变博士。

这条路走得通，靠的是一个前提：信号之所以能分辨，是因为它对不同的人成本不同。

这是斯宾塞式分

离的全部内容——高能力者拿到那个信号更便宜，所以愿意拿；低能力者不愿意。当一个信号被太多人拿到，它就不再分离，于是市场找一个更贵的信号，成本差重新拉开，分离恢复。梯子一级一级往上，每一级都还能站人。

人工智能带来的成本塌陷不是这种塌陷。它压低的不是某一种信号的成本，而是"生成任何一种可被读取的证据"的成本，而且对各类型近乎同比例。这意味着：>任何新设的可读门槛，在设立的那一刻就已经处在饱和的路上。梯子不是走到了顶，是每一级在你踩上去之前就已经被踩平了。

这一条给出了本章与资格通胀之间一条可以直接去查的分界：资格通胀预测会出现并稳定下来一个新的、更高的、仍然可读的分离门槛；本章预测不会有任何新的可读门槛稳定下来——设一个，饱和一个。招聘公告里的资格要求、期刊的投稿须知、评审的形式条款，这三处的历年文本都是公开的，可以逐年清点。

5.3 分辨迁到了哪里分辨不会消失。一个岗位只有三个，总要有一种办法从一千二百人里选出三个。当它不能停留在可读区，它会迁到别处——而它迁去的地方有一个共同结构：>那不是一份可以被生产的文件，是一个会被追究的位置。

推荐人押上的是自己以后说话还算不算数；内部转介押上的是自己在公司里的信誉；编辑约稿押上的是这本期刊的面子；投资人的引荐押上的是下一次别人还接不接他的电话；一件作品的来源链押上的是每一个

这道关拦下过谁

经手人的名声；一次事故里"谁在场"押上的是这个人的职业记录。这一类凭据之所以还能分辨，正因为它不能被生成——生成它需要有人真的把自己放进去。

这件事有两个后果，而且都不好。

第一，它没有容量。一个人能为多少人背书？几十个，不能再多。而投递量在以年为单位翻倍。分辨迁到这条通道上，通道立刻堵死，堵得比原来更硬——原来至少人人可以投一份简历，现在连排队的资格都要先认识一个人。

第二，它没有规则。谁有资格背书、背错了怎么样、被拒绝的人向谁申诉——这一整套在可读门槛那里发展了上百年的制度，在这条通道上一条也没有。可读门槛的全部好处（可核查、可申诉、可扩容）随着分辨的迁移一次性丢光。

所以内卷不会因为分辨迁走而结束。它只是换了形状：排队从“多生产一份证据”变成“靠近一个背书位”——更零和，更不可测，更难申诉。这就是二〇二五年招聘现场正在发生的事，也是本章认为“新经济现象”这个说法名副其实的地方：它不是旧内卷的加剧，它是内卷的载体从可读区整体搬了家，而搬家这件事在任何账本上都没有记录。

■ 六·两轴四格，与最危险的那一格

一个名字不是一条辨别维度。要让它能分辨，得有第二根轴。

第一根轴是这道关的分辨力：最近若干次判定里，它把人分成两堆了吗，还是几乎人人通过。第二根轴是这道关的权属：有没有一个位置，有权宣布它作废，并且要为此决定承担后果。

有作者有分辨饱和

①

无作者

正常筛选：关在拦人，也有人负责它拦

得对不对。

③

②

硬关无主：还在拦人，但没有任何位置能改它。历

史遗留的资格条款多半在这一格——它可能拦错了人，而没人有权撤。

已知失效，正在处置：有人测出它不再

区分，并且正在废止或替换。考务方剔题就④空关：不拦任何人，也没有任何位置有权废。

是这一格。

第四格是本章要打的那一格，本章给它一个名字：空关。

选这一格，不是因为它最坏看，是因为它最不像有问题。一格值得被打，要满足三条，缺一条说明第二根轴选错了：其一，它在短期指标上表现为改善。

一道从不拒绝的关，在报表上产生的是这样几个数字：通过率高、

平均处理时长短、退回重做的次数少、投诉为零。这些数字在任何一套流程评价里都是好数字，通常被读作“候选人质量提升”“流程顺畅”“合规达标率上升”。

这道关拦下过谁

其二，它的失效不产生任何信号。

这是三条里最要紧的一条。一道从不拒绝的关，在日志里与一道从不

被触发的告警长得一模一样：都是一串“通过”“通过”“通过”。系统没有任何机制会去问“为什么全是通过”。相比之下，一道错拦的关会立刻产生信号——被拦的人会申诉。只有不拦，是安静的。

其三，它自我加固。

一个流程越正规化，关口越多；关口越多，每一道的责任越分散；责任越分散，越

没有任何一个位置有权废除其中任何一道。于是正规化程度与空关数量正相关——这与人们的直觉正好相反，人们通常认为正规化会淘汰无用环节。实际情形是：新增一道关只需要一个理由，废除一道关需要一个授权，而多数机构只有前者的流程，没有后者的。

三条齐了，靶格成立。也正因为三条齐了，空关不会被自己发现——它只能被从外面数出来。

■ 七·一句不含判断词的问话

要把一格从纸上落到地上，需要一句可以拿去问文件的问话，而不是拿去问人的判断。本章的那一句是：>这道关，最近一百次判定里拦下过几次？

这句话里没有“应当”“合理”“充分”“实质性”“真正”。它问的是记录，不是意见。任何一个既有流程都可以被这样问一遍，而多数流程从来没有被这样问过。

把它在五个场合各跑一遍，答案互不相同：招聘的简历初筛。

查申请人跟踪系统的日志，看“未通过初筛”的条数占投递数的比例。这个数几乎所有

公司都存着，而几乎没有公司调出来看过。二〇二五年若干招聘方的说法是硬性条件的通过率已经很高——那意味着这道关的答案接近于零。

期刊的形式审查。

格式、伦理声明、利益冲突、

体例。这一道的答案查不出来：多数编辑部只

统计"编辑直接退稿"的总数，不区分退在形式关还是内容关。而"查不出来"本身就是一个答案——一道从来不被单独统计的关，不可能有人为它负责。

一条持续集成流水线上的某一道检查。

查最近一百次运行里这一道红过几次。这是五个场合里最容易查

的一个，一个下午能查完全部关口。而按行业调查的读数，人工智能采用后交付稳定性下降、评审停留时间大幅拉长、三成一的合并请求无人评审——这几个数放在一起，指向的正是评审这道关的答案在下降。

一次艺术征集的初选。

查初选淘汰率。通常很低：初选只滤掉不合格的投稿，真正的取舍在复选和终选，而复选和终选的依据往往不进记录。

一门课期末考里的某一道题。

这一道有标准答案：心理测量学早就有"项目区分度"这个指标，一道所有

人都答对的题区分度为零，应当从题库里剔除，而且有考务方负责剔。这个场合是对照组——它落在第三格，不是第四格。

这道关拦下过谁

五个答案互不相同，其中至少两个不是从命题里推出来的，是查出来的：期刊那一道的"查不出来"，与考试那一道的"有人负责剔题"。前者是本章没有预料到的一种失效形式（连读数都不存在），后者是本章的边界（存在有作者的领域，本章不适用于它们）。

■ 八·一个量纲不同而比率相同的读数

一句问话只能问一道关。要让不同行当能并排比较，需要一个数。

饱和关占比：一套判定流程的全部关口中，最近 N 次（ N 不少于一百）判定的通过率不低于零点九五的关口数，占关口总数的比例。

它在四个行当里的量纲完全不同，比率却相同：· 招聘：申请人跟踪系统里从投递到发出录用之间的各道关。

- 学术出版：从投稿到刊出之间的各道关（形式审查、查重、编辑初判、外审、复审）。
- 软件交付：持续集成与交付管线上的各道检查，加上代码评审这一道。
- 艺术征集：初选、复选、终选各道。

这个读数有四条性质，本章认为缺一条都不该用。

其一，无量纲，所以三个行当可以并排比较：一家公司的零点八与一本期刊的零点八说的是同一件事——十道关里有八道谁也拦不住。

其二，可事后清点：所需数据在多数机构的既有记录里已经存在，不需要新建任何系统。这也是本章最便宜的一条检验。

其三，它把一个印象变成一个数："这道关形同虚设"是一句人人都说过的抱怨，抱怨不能被检验，比率可以。

其四，也是最要命的一条：它与既有的主要指标正交。

一个与既有指标高度相关的新读数会被立刻读成

那个指标的代理，白造。所以要举一个"既有指标最好看、而这个数最难看"的真实例子——举不出来就该重造。

本章的例子是一条从来没有红过的持续集成流水线。按既有的交付指标看，它接近完美：变更失败率极低，部署频率极高，平均恢复时间无从谈起因为没出过事。而它的饱和关占比是一——每一道检查都拦不住任何东西。这两个读数指向相反的结论，而且都不是错的：这条管线确实从不失败，也确实从不分辨。

同样的例子在招聘里是：一家录用率千分之三的公司，可以有百分之百的饱和关占比。它看起来筛得极严，实际上每一道公开的关几乎人人通过，最后的取舍发生在一次不公开、不记录、无从申诉的判断里。

录用率低不等于分辨力强——这两个数第一次被分开，正是这个读数的用处。

这道关拦下过谁

■ 九·出路是一条有次序的路，不是三选一

如果三家的处方各是一段而不是全程，那么问题就从"选哪一条"变成"按什么次序走"。本章认为次序是死的，做错次序会让后两段失效。

9.1 第一步：作废先于新建先宣布旧门槛作废，再设新的终点。

顺序反过来必失败，理由很简单：产能没有被释放。如果旧的关还

在，新设的终点只是给同一批产能多加了一道工序——所有人既要满足旧关，又要去够新终点，而新终点由于同样可读，会在很短时间内被同样的工具填满。二〇二四年到二〇二六年之间，招聘方每加一道新的评估形式，就有一批工具在几个月内针对它训练出来，这个循环已经跑过好几轮。

作废这个动作有一条硬要求：它必须是一个有名字的位置做出的、公开的、指名的决定，不是“流程优化”。“我们简化了申请流程”不算作废——它不指名是哪一道关、不给理由、不留下可追究的人。作废要写成：“第三道关自即日起废止，理由是它在最近一年的四千次判定中拦下过十一次；决定人某某。”

9.2 第二步：给“接住”记账现在的账只记生产端：投了多少份、发了多少篇、合并了多少代码。接收端只有一个总量——录用几人、接收几篇——没有任何一条记录说明是谁接住的、后来怎么样。

这一步要做的就是把这一栏补上：让“接住”成为一个有名字、有后续、因而有代价的动作。不是为了追责，是因为一个没有代价的动作不会稀缺，而不稀缺的东西不能承担分辨。这也是第五节末尾那个困境的唯一解法：如果分辨必须迁到“会被追究的位置”上，那么与其让它自发地迁到不可核查的私人关系里，不如在制度里造出可核查、可扩容、可申诉的追究位置。

这一步必须在第一步之后。反过来先记账再作废，资源会全部投到“把饱和的关记得更精细”上——那正是工程学第一步的失手方式，也是过去二十年绝大多数流程改造实际发生的事。

9.3 第三步：值得性不能被直接瞄准第三步是艺术学那一段，而艺术学自己知道这一段不能被直接瞄准。什么算值得做，是前两步的结果，不是可以先行宣布的目标。任何“我们从今天起重视原创”“我们看重的是真实能力”的宣示，如果没有前两步，只会变成又一道可读门槛，并在设立当天开始饱和——因为“证明自己有原创性”这件事同样可以被生成。

9.4 插手点三步里，最小可行的插手点在第一步与第二步之间：作废需要一个能承担后果的位置，而这正是现在不存在的那个位置。

所以第一个动作不是废除某道关，是给某一道关指定一个作者——指定一个具体位置，要求它每年回答一次“这道关最近拦下过谁”，并且这个回答进入它自己的考核。然后看会发生什么。

这道关拦下过谁

这个动作成本极低（一份文件、一个字段），后果可观察（那道关一年之内是否被改动），而且它同时是本章最便宜的一条判错条件：如果有作者的关与无作者的关在饱和速度上没有差别，本章的第二根轴就不成立。

■ 十·十二处兑现

跨领域的兑现不是“像”，是“在那个领域里能据此预测一件具体的事”。

一·招聘：投递量继续上升而录用率不变的公司，其内部转介占录用的比例会先于“简历质量下降”的抱怨出现。可查：转介比例是多数公司现成的数字。

二·学术出版：形式关的饱和会先于"编辑直接退稿率上升"出现，而不是相反。可查：投稿系统里两个环节的时间序列。

三·软件交付：无人评审的合并占比上升，会先于事故率上升出现，中间有一个数月的空窗——这段空窗正是所有指标都好看的那一段。

四·高校招生：越是推行"多元评价"的招生体系，可读材料（获奖、证书、活动清单）的饱和越快，而最终录用越依赖推荐信里的私人语气。预测：推荐信的长度与具体程度会上升，而其可比性会下降。

五·政府采购与投标：资格预审的通过率会趋近于一，而中标越来越依赖此前的履约记录——这是采购领域里少数已经制度化的"不可生成凭据"。预测：新进入者的中标比例会下降，而与"围标"无关。

六·医学影像初筛：初筛报告的阴性率趋近于一时，真正的判断压回到那位签字的医师身上，而报告数量继续上升。预测：签字环节的排队时间上升快于阅片环节。

七·内容平台的入池审核：合规审核的通过率趋近于一，而分发越来越依赖账号历史与人工加权。预测：新账号的冷启动难度上升，与内容质量无关。

八·艺术征集与画廊初审：初选淘汰率下降，而入选越来越依赖此前的展览记录与推荐。预测：首次参展者的比例会下降。

九·风险投资的项目初筛：商业计划书的形式完备度趋近于一，而进入尽调越来越依赖推荐。预测：冷邮件的转化率会下降到接近于零，而这一变化早于任何"项目质量下降"的判断。

十·职业资格考试：通过率上升而岗位不增时，用人方会自行加设第二道门槛；预测：这道自设门槛在两年内同样饱和，然后被第三道取代——而每一道的存续期在缩短。存续期缩短这一条，是本章与资格通胀最直接的对照读数。

十一·合规审查：反洗钱、数据安全这类审查的形式条款通过率趋近于一，而真正的拦截依赖少数经验丰富的审查员的直觉。预测：这类岗位的离职会造成不成比例的拦截率下降，而所有形式指标不变。

这道关拦下过谁

十二·教育评价中的查重：查重率的通过比例趋近于一（因为改写成成本塌陷），而学术不端的认定越来越依赖举报与私下判断。预测：查重系统的迭代频率上升而其拦截绝对数下降。

十二处里有两处反过来把本章削小了，下一节单说。

■ 十一·一条反直觉的推论：越把关做准，它饱和得越快

前面十节说的的是一个状态。这一节说的是这个状态为什么会加速——而加速的来源，恰好是所有人都在做的那件好事。

推论写成一句：在门槛权属不变的前提下，任何提高筛选精度的努力，都会缩短这道关自己的存续期。

论证只有三步。

第一步：提高精度就是把判定依据写得更明确。

一道"综合素质优秀"的关不精确，因为不同的人从这六

个字里读出不同的东西。把它改写成"三年以上相关经验、独立主导过至少两个上线项目、有可核对的量化结果"，精确度上升了——这是所有流程改造教科书里的标准动作，也确实解决了不一致的问题。

第二步：更明确就是更可读。

判定依据以更可传递、更不依赖读者的形式存在，这正是"精确"这个词的

内容，不是它的副作用。一道精确的关，定义上就是一道任何第三方读了都能得出同一判断的关。

第三步：更可读就是更可被生成。

在生成成本已经塌陷的条件下，一份写得越明确的要求，越容易被针

对性地满足——因为它已经把"要什么"交代清楚了。写得含糊的要求反而难对付：你不知道该往哪里使劲。

三步合起来是一句让人不舒服的话：精确度与可生成性是同一个量的两种读法。于是每一次"把关做得更准"，都在把这道关往它自己的饱和推进一步。这与工程学的整套直觉正面冲突——在工程学里，把规格写清楚是消除歧义、提高质量的标准动作，而且它在别的场合确实是。它在这里失效，只因为这里的输入方是有动机、会学习的：规格写给一台机床看和写给一千二百个申请人看，是两件事。

可检验的形式：一道关被"优化"（改写、细化、量化）的次数，与它从设立到饱和和所用的时间负相关。招聘要求的历年文本、评审细则的历次修订、持续集成检查的配置提交记录——这三处都留着完整的修订史，这条相关可以直接跑，不需要新建任何数据。

写实一点，这个循环长这样：某岗位收到大量不合适的申请→把要求写细→申请的"质量"上升（通过率上升）→被读作优化奏效→半年后通过率接近一→再加一条更细的要求→循环。每一轮的短期读数都是改善，每一轮的长期效果都是把这道关往饱和推进一步。这个循环

里没有任何一步是错的——这正是它难被发现的原因，也是第六节第二条签名（失效不产生信号）在时间上的展开。

它同时给出了第六节第三条签名（自我加固）的来路。正规化程度高的机构不是因为官僚才不断加关，是因为它们更擅长把要求写清楚——而写清楚正是加速饱和的那个动作。结论因此比"官僚主义"难听得多：能力越强的组织，它的关饱和得越快。

这道关拦下过谁

反过来，这条推论也指出了唯一一个不加速饱和的改进方向：不要提高判定依据的可读性，要提高判定者的可追究性。

前者改的是关的内容，后者改的是关的权属；只有后者不会被生成掉——一个位置所承担的后果，申请人无法自行生产。第九节那三步之所以是那个次序，根子就在这里。

边界：这条推论只在输入方会学习、且学习成本低的场合成立。输入方无法调整的场合——体检指标、材料强度检验、传染病筛查——把规格写细不会带来任何饱和，反而是纯粹的改善。这条边界与下一节工程学削掉的那一处是同一条线，从另一个方向再画了一遍。

■ 十二·两处反过来削小本章的约束

11.1 工程学削掉的：适用范围在安全关键系统里，饱和的关是设计目标，不是失效。

核电站的安全联锁、航空的适航检查、医疗器械

的出厂检验——这些关口的行业目标正是让通过率趋近于一。一道经常拦下东西的检验说明上游出了问题；一道从不拦人的检验恰恰说明整条链子是健康的。

这一处直接削掉了本章原本会写下的一句更强的话：>原本会写：凡是通过率饱和的关口都应当被作废。>现在只能写：只在以分配稀缺位置为目的的关口上成立。

这不是补一个条件，这是把适用范围砍掉了一大块——工业质量控制、安全检验、合规底线，全部不在本章的适用范围内。判据也随之明确：一道关是用来分配的（岗位、版面、资金、展位），还是用来防止的（事故、污染、伤害）？前者饱和是病，后者饱和是目标。

11.2 艺术学削掉的：价值排序艺术学告诉本章一件更难受的事：有些分辨在原理上不应当有作者。

如果给"什么算作品""什么值得被展出""哪种研究方向有价值"指定一个有权宣布作废的位置，那正是审查制。艺术史上每一次这样的指定，都以灾难收场。

这一处削掉的是本章的价值排序：>原本会写：分辨必须有作者，无主的门槛一律有害。>现在只能写：无主是门槛的病，却是价值的健康。本章的处方只适用于以分配为功能的关口，不适用于以评价为功能的判断。

这两处约束合起来，把本章从一条关于"一切筛选"的普遍主张，压成了一条关于"以分配为功能的可读关口"的有限主张。本章接受这个压缩，因为不接受它，本章就会在第一个安全工程师和第一个策展人那里各栽一次。

这道关拦下过谁

11.3 经济学补的一处小约束J曲线还提醒本章一件事：转型期的既有记录本身系统性偏斜。因此饱和关占比的绝对值不可跨期比较，只能同期横比——拿二〇二六年的零点八去比二〇一九年的零点五，比出来的东西不可靠。这一条削掉的是本章原本想主张的"这个读数可以做长时间序列"。

■ 十三·与最近的几种说法的分界

本章最容易被吸收进以下十种既有说法。逐条写清楚差在哪，以及怎样才能判出谁对。

一·信号理论（Spence1973，经济学）。它说到的那一步是：教育等信号之所以能分辨，是因为它对高能力者成本更低。分界在于：斯宾塞式分离依赖成本差，而人工智能带来的塌陷对各类型近乎同比例，因而不存在新的分离点。判定形式：若在二〇二七至二〇三〇年间出现并稳定下来一个新的、公开的、可被第三方生成的分离门槛（长期通过率低于零点五），则本判断错而信号理论对；若各机构反复设立新门槛而每一个都在两年内通过率越过零点九，则反之。

二·资格通胀（Collins,《文凭社会》1979，社会学）。它说到的那一步是：学历要求随供给上升而水涨船高，工作内容不变。分界在于：通胀预设梯子还有更高一级，本章说每一级在被踩上去之前已被踩平。

判定形式：若下一个十年里资格要求继续整齐上移一级而录用仍以可读凭据为主，则通胀对；若资格要求停止上移、而录用越来越依赖推荐与共事史，则本章对。招聘公告的历年文本可以逐年清点，第十处兑现给出的"每道门槛存续期在缩短"是同一条的另一种读法。

三·内卷化（Geertz1963，人类学）。它说到的那一步是：技术不变时，更多劳动被吸收进同一套结构，复杂度上升而人均产出不升。分界在于：格尔茨的前提是技术停滞，而本章描述的内卷恰恰由一次技术跃迁引起。判定形式：若在技术冲击最强的行当里内卷读数最弱、在技术停滞的行当里最强，则格尔茨的机制在起作用；若相反，则本章对。这一条可以用同一国家内不同行业的横截面数据判。

四·古德哈特定律与坎贝尔定律（Goodhart1975；Campbell1979）。它们说到的那一步是：一个度量一旦成为目标就不再是好度量，因为它会被博弈。分界最锋利，因为同一份材料两边给出相反评价：取一批全部真实合格、经第三方复核确认无任何作假的申请。古德哈特读作"指标尚未被腐蚀，制度健康"；本章读作"这道关已经作废，只是还在运行"。判定形式：在有第三方复核的资格考试里，若饱和程度与作假率不相关，则本章的机制成立而博弈不是必要条件；若二者高度相关，则本章可被古德哈特吸收。这是本章最愿意被检验的一条——因为它意味着本章的机制里没有坏人。

五·柠檬市场（Akerlof1970，经济学）。它说到的那一步是：质量不可观察时劣品驱逐良品，市场萎缩甚至瓦解。分界在两处：阿克洛夫要求存在真实的质量差且卖方知情，而本章的情形是没有隐瞒、产品确实合格；更要紧的是二者预测方向相反——柠檬市场预测交易量萎缩，本章预测投递量爆炸而成交比例不变。判定形式：若观察到申请量随门槛失效而下降，则柠檬市场对；若观察到申请量暴涨而录取率不动，则本章对。二〇二五年的两组公开数字落在本章这一侧。

这道关拦下过谁

六·注意力稀缺（Simon1971，经济学与认知科学）。它说到的那一步是：信息丰裕造成注意力贫乏，注意力成为稀缺资源。分界在于：西蒙那里稀缺的是接收者的注意力，而注意力原则上可以外包给机器；本章说稀缺的是有权拒绝并为拒绝承担后果的位置，它不能外包——一个模型可以读完全部一千二百份简历，但它不能被追究。判定形式：若在把筛选完全外包给模型的机构里堵塞消失，则注意力稀缺对；若堵塞只是从阅读环节移到"谁签字"环节，则本章对。这一条现在就可以在已经全面自动化初筛的公司里查。

七·价值捕获（C.ThiNguyen,《伦理学与社会哲学杂志》2024，哲学）。它说到的那一步是：我们把自身丰富的价值让位给外部提供的、清晰量化的指标，从而被指标接管。分界在机制：阮说的机制是吸引——指标因为清晰而诱人；本章说的机制是无人有权退出——即使所有人都清楚地看见这道关毫无意义，继续满足它仍然比取消它便宜，因为取消需要一个不存在的位置。判定形式：若给出更丰富的评价方式之后人们自愿放弃旧指标，则价值捕获对；若人们同时保留新旧两套、并把旧的那道关继续走完，则本章对。这一条在推行综合评价的招生体系里已经可以观察。

八·项目区分度与题目分析（心理测量学，Lord&Novick1968及项目反应理论的区分度参数）。这是本章最危险的占位者，因为它更成熟也更便宜：一道所有人都答对的题区分度为零，应当剔除——这套方法写进教科书已经半个世纪。分界在于：整套题目分析预设存在一个考务方，它有权剔题并对题库负责；而本章说的正是没有这个位置的那种题库。区分度是题目

的性质，饱和关占比是流程的性质加权属的性质，而第二根轴在心理测量学里根本不是一个变量。判定形式：若一个组织在测出某道关区分度为零之后的十二个月内把它废除，则那道关是有作者的，本章不适用于它；若它测出来、写进报告、而关照旧运行，则本章适用。任何做过流程审计的机构都可以立刻查这一条。

本领域名词、只留下形式，得到的是"一个用来分辨的装置在失去分辨力之后继续被维持"。这个形式在组织社会学里已有八十年历史。分界在于：目标置换里装置仍在起作用（人们过度遵守它，遵守本身取代了用途）；本章里装置已经不起任何作用（它谁也不拦），而人们仍在为它加码。前者是过度执行，后者是空转。判定形式：移除该关口，若引起可观察的质量下降，则是目标置换；若移除后各项质量指标不变而工时下降，则是本章说的空关。这个移除实验在多数机构里可以低成本做一次。

是"越成功越失败"这类反转模板里最贴近的两个，必须点名。分界在于：这两条的机制都在组织的能力——过去的成功把组织锁死在旧本事上；本章的机制不在能力，在判定的权属。一家能力完全没有僵化、随时可以换做法的组织，照样会守着一道空关，因为换做法不需要授权，废一道关需要。判定形式：取两组组织，一组能力更新频繁、一组僵化，若空关比例与能力更新频率不相关而与"是否设有废止流程"相关，则本章对。

十条里最近的邻居是第八条。本章之所以仍不能被它吸收，只有一个理由：它假定了那个本章认为已经缺席的位置。

刷题这个动作在心理测量学里从来不是问题，因为总有一个考务方；本章说的正是没有考务

方的那种考试——而人工智能时代大多数筛选流程，恰好就是这种考试。

这道关拦下过谁

■ 十四·与本书几篇同族文章的分界

本书此前有几篇与本章形状相近，读者很容易把它们读成同一件事。逐一分开，并各给一条判决性对照。

第四章《认领物》（第四章）是本书与本章距离最近的一篇——同一天发表，用的是同样三个学科，解的是同样一道题，因此必须先分开。它说：内卷是产出构成中「认领物」占比持续上升——为使一件产出可被归属于某个人，必须与它一并被生产的那一层东西；三家的账本都只给它成本位、不给产出位，于是同一件事在三处同时显示为「努力增加、产出不变」。

两篇的病灶不在同一处。第四章《认领物》的病灶在归属的成本：证明「这是我做的」要花钱，而这笔钱无账可入。本章的病灶在判定的权属：一道用来分辨的关失去分辨力之后，没有任何位置有权撤掉它。前者是多出来的一笔开销，后者是撤不掉的一道关。

判决性对照有两处，方向相反。其一：取一个归属成本近乎为零的系统——公司内部的代码仓库，每一次提交都由身份系统确定地记在某位具名员工名下，不需要任何额外的「证明是我做的」的工作。第四章《认领物》在这里预测内卷读数应当很低（认领占比接近零）；本章预测内卷照样出现，只要评审那一道关饱和而没人有权废它。二〇二五至二〇二六年的交付遥测正落在本章这一侧：归属在版本库里是完美的，而评审停留时间大幅拉长、三成一的合并请求根本无人评审。同一份材料，两篇给出相反的读数。

其二：反

过来取一个归属成本极高、而判定有明确作者的场合——一个定期修订并公开作品目录的鉴定委员会。

第四章《认领物》预测高内卷，本章预测低内卷，因为那道关有作者，失效会被撤。

最值钱的不是这两处相反，是两篇的互补关系：两条机制可以独立发生，叠加时不是相加。

认领物占比

越高，每一件产出携带的可读证据就越多，可读门槛饱和得越快；反过来，关越饱和，人们越要靠加厚认领物来求分辨。二者互为对方的加速器，因此在同一个系统里测到的总损耗，会高于两条各自预测之和。

这给出一条可以同时判两篇的检验：若「认领占比」与「饱和关占比」在跨机构样本里高度相关，说明它们测的是同一件事，其中一篇应当被另一篇吸收；若二者正交而联合解释力显著高于各自单独的解释力，则两条机制都成立。这两个读数所需的数据都在既有记录里，这条检验现在就可以做。

《轮空》（之四十）说停滞是“发起权已经交出而无人接手的那一轮”。它与本章的关系最紧，因为两篇说的都是一处缺席的权责。但读数相反：《轮空》读到的是改变没有发生，本章读到的是改变极其频繁而全部落在同一处。判决性对照：取一个投递量三年翻倍而录取率不变的机构。《轮空》的换手率读数会显示“发起处三年未变”，判为轮空；本章的饱和关占比会显示“每一道关都在被高频触发且几乎从不拒绝”，判为空关。两者可以同时为真，处方却相反——《轮空》要找出该接手而没接手的那一处，本章要找出该被废而没人能废的那一道。先做哪一个，取决于这个系统的活动量是低还是高。

《能被量的那一面》（之一）说能被读出来的那一面会自动接管治理方式，而它通常不是长东西的那一面。本章接着它往下走了一步：当所有面都能被读出来时，“能被读出来”这件事本身就不再接管任何东西——治理不是被可测面接管，是被没有人有权撤销的可测面接管。两篇互补而非重复，而且方向可以对上：那一章的分子越大（可测面占比越高），本章的饱和越快到来。

这道关拦下过谁

《缝合致死》（之八）说系统死于每一件让它更成功、更可传承的正确之事。本章给它加了一个条件：只有当那件正确之事没有作者时才致死。

一条有人负责定期作废的度量，缝合不了什么。判决性对照：两

个同样度量密集的组织，一个每年公开废除若干指标、一个从不废除；《缝合致死》预测两者都会衰亡，本章预测只有后者。

《静止点》（之二十三）说整篇重做需要一个“所有使用者同时不在使用”的时刻。本章说的作废不需要静止点——一道关可以在系统运行中被单独废掉。两篇因此是一对：重写需要静止点，作废需要作者。

一

一个系统可以两样都没有，那时它既不能重做也不能减负，只能不断加关；这正是本章第六节第三条签名所描述的自我加固。

《哪一份算数》（之十五）说危险发生在“系统预期副本一致，却没有任何人负责让它们一致”。形式上与本章同族——都是预期存在而权责空缺。分界在空缺的位置：那一章的空缺在维护，本章的空缺在废止；那一章问的是“如果这两份不一样，哪一份算数”，本章问的是“如果这道关谁也拦不住，谁能把它撤了”。

判决性对照：一个把所有副本的维护责任都指定清楚、却没有任何废止流程的组织，那一章判为健康，本章判为高危。

第一章《两种「不」》（第一章）与本章同日发表、用的是同样三个学科，且它指认的三家共有前提与本章逐字相同——出路可以在某一个终点上单独结算。分歧在推翻它的材料与由此落到的病灶：本章取自经济学自己（实施期里被记作成本的那一项才是当期真正的产出），落在门槛这一侧；那一章取自工程学自己（一项随机对照试验里，产出端读数与当事人自评同向而实测反向），落在否决这一侧——一次否决有没有人看过，而看过与没看过在所有账本上同码。两篇看似矛盾（本章说关拦不下人，那一章说绝大多数被否决），实则说的是同一条链上的两类站：关的分辨力塌向零，决定的过目率也塌向零。那一章的附录把两个读数锁成一条关系式——设投递量

N 、各关通过率之积

$1-K \div (N \cdot \prod p)$ ，而一次投递的结果确实取决于其内容的概率为

$\prod p$ 、可过目产能

K ，则未阅率

$u =$

$1-(\prod p) \cdot u$ ：两侧的失效在这个量上相

乘。由此得到一条两篇各自都写不出的结论，也是最值得核的一条判决性对照：撤掉一道饱和的关（本章出路的第一步）会降低投递者的代价，却不改善有据判定率，甚至微降——因为谁也拦不下的关本来就不为下游减载；反过来，给「不」记账（那一章出路的第一步）改善回报却不降代价。两篇的第一步互不相干，必须同做；只做一半的一方会在自己的读数上看到改善，而投递者的处境不动。若在同一条链上实测出 u 可由 $K \div (N \cdot \prod p)$ 预测，则两篇应合为一条判断，各自都只说了一半。

■ 十五·三家为何各自到不了这一条

三家没有走到这一步，不是因为疏忽。一个领域看不见什么，通常不是它的缺陷，是它的问题结构的影子。

经济学的核心任务是均衡与配置。信号理论里的"分离"是一个均衡概念，而均衡分析总是问"新的均衡在哪里"，不问"从旧均衡退出需要谁来宣布"。退出的授权问题在均衡框架里不是一个对象——它被默认由市场完成：没人要的信号自然会被市场抛弃。这个默认在有价格的地方大体成立，在没有价格的地方（一

这道关拦下过谁

道内部审批、一条评审条款、一份岗位要求）完全不成立，而后者恰恰是可读门槛的大多数。所以经济学看得见门槛失效，看不见门槛无主。

工程学的核心任务是把不确定性挤出过程。在这个任务结构里，"一道关拦不下任何东西"是成功的读数——它意味着上游质量好到不需要拦。工程学因此没有任何动机去区分"没有东西该被拦"与"该被拦的拦不住"：这两种情形在它的仪表盘上产生完全相同的数字，而它的整套价值判断都站在前一种解释那一边。这也是为什么第十二节那处约束来自工程学而不是别处——它对同一个现象持有一个相反且正当的解释。

艺术学的核心任务是抵抗标准化。它比谁都早看见可读门槛的空洞——它已经看了一百年，从复制技术开始动摇独一性的那一刻起。但它的任务结构使它把出路一律指向取消门槛，而不是给门槛设计续期与作废的制度。"为一道关指定一个作者"这个动作，在艺术学的语汇里

天然读作审查；它无法在自己的框架内说出“有些关口需要一个能宣布它作废的人”，因为那句话的前半句它就不接受。

三家的盲点因此不是三种疏忽，是三种任务结构各自的背面。而这三个背面恰好覆盖了同一块地。

■ 十六·这个判断底下还站着一块更大的地

把第十三节那十位邻居逐个问一句“它把什么当作给定，因此结构性地看不见什么”，会得到一份出人意料整齐的清单：· 信号理论把“总会有下一个可分离信号”当作给定，因此看不见分辨整体迁出可读区。一旦承认这一点，分离均衡这个概念就没有了定义域。

· 资格通胀把“梯子还有更高一级”当作给定，因此看不见每一级在被踩上去之前就被踩平。

· 古德哈特定律把“有人在博弈这个指标”当作给定，因此看不见没有人博弈而指标照样失效。一旦承认这一点，“被当作目标故被操纵”就不再是必要条件，该定律的因果机制被削掉一半。

· 项目区分度把“存在一个考务方”当作给定，因此看不见没有考务方的题库。一旦承认这一点，整套题目分析的规范性落空——分析出来给谁执行？

· 注意力稀缺把“稀缺的是接收方的注意力”当作给定，因此看不见稀缺的是可被追究的位置。一旦承认这一点，“把阅读外包给机器就能缓解”这条推论直接失效。

· 柠檬市场把“卖方知道真实质量而买方不知道”当作给定，因此看不见双方都不知道、而且产品确实合格的情形。

这六行背后是同一件被当作给定的事：>凡是一道判定，都有一个持有它的位置。

出题的人、定标准的人、市场、传统、共同体——总有一个谁。正因为所有人（包括互相批评的各派）都站在这块地上，谁也没有回头看它。而本章要说的正是：>一道判定可以在无人持有的状态下继续运行，并且继续被所有人当真。

这道关拦下过谁

这不是一种罕见的病态。在一个正规化程度足够高的组织里，它是常态：加一道关只需要一个理由，废一道关需要一个授权，而多数组织从设立那天起就只造了前半半的流程。

■ 十七·一旦被制度采用会发生什么

本章提出的读数一旦被采用，最省事的实现方式是吧"这道关最近拦下过谁"做成一个考核指标。而一旦它进了考核，各机构让这个数好看的最便宜办法不是废除空关，是人为地拦下一些人——制造拒绝，以证明这道关是有效的。

于是筛选从"没有分辨"变成"随机分辨"。这更坏，而且坏在两处：随机拒绝在所有报表上都会被读作有效筛选；同时它消灭了申诉的依据——被一道空关放过去是没人受害的，被一道随机拒绝的阻拦下，受害的人连一个可以指认的理由都没有。

这条反噬我看不到出口。

任何对这个读数的核查都需要一次判断，而那次判断本身同样没有作者。把核

查交给外部审计，只是把同一个空缺挪到审计机构里。本章提出的读数因此在原理上不能自我保护，只能靠公开——把每一道关的拦截次数连同判定总数一并公开，让被拦下的人自己去看那个数是不是突然变了。这是一个弱得多的保护，但本章找不到更强的。

■ 十八·判错条件

先写出最强的竞争解释，因为不排除它，下面的条件都不算数。那句最朴素也最难反驳的话是："其实不就是人太多、位置太少吗？"如果堵塞只是供求失衡，那么门槛饱和只是这个失衡的一个症状，权属这根轴根本不必要。下面七条里，第一条就是冲着它去的。

一·控制申请人数与位置数之后，若"饱和关占比"与"录用决定依赖不可生成凭据的比例"之间不存在剂量—反应关系，则本判断为伪——届时堵塞应归因于绝对供求，而与门槛权属无关。样本纪律：至少六个同制度环境内的单元（同一行业、同一国家的六家以上企业，或同一学会下属的六本以上期刊），在位置数、申请量、行业景气这些关键混淆变量上同质。不得拿两个国家或两个地区的对比充数——那种对比在方法上不能支持任何"控制掉某变量"的说法，只能作启发式例证。

二·若在一个饱和关被正式废止的机构里，十二个月内投递量与内部工时都没有下降，则本章的处方无效——废关不能减负，那么第九节整节应当被放弃，尽管第五节的诊断仍可独立成立。

三·若在同一批机构里，"有作者"的关与"无作者"的关在饱和速度上没有差别，则第二根轴不成立，四格退化为二维，本章只剩下一条既有的区分度读数，应被第十三节第八条吸收。

四·若在二〇二七至二〇三〇年间出现并稳定下来一个新的、公开的、可被第三方生成的分离门槛（长期通过率低于零点五且不依赖推荐），则"分辨迁出可读区"为伪。

这道关拦下过谁

五·若在已经把筛选完全外包给模型的机构里堵塞消失（而不是移到签字环节），则“稀缺的是可被追究的位置”为伪。

六·顺序读数：本章预测门槛饱和早于录用向不可生成凭据的迁移，而不是同时或滞后。若时间序列上观察到推荐制的扩张早于形式关的饱和，则本章的因果次序错。顺序比相关难被巧合满足，这一条是本章最愿意拿去赌的一条。

七·一条低成本可实做的：任取一条持续集成流水线，清点最近一百次运行里每一道检查红过几次。这份数据在多数机构的既有记录里已经存在，一个下午可以做完。若各道检查的拦截次数分布与本章预测的“多数为零、少数承担全部拦截”不符，本章在这个行当里不成立。

这条判断若被证伪，我们会学到什么：若饱和与权属无关，那么内卷就真的只是供求问题，出路只能在总量上找——扩招、扩容、限制参与者数量。这与本章的处方（改权责）指向完全不同的干预，也意味着过去二十年那些针对流程的改造从一开始就找错了地方。证伪本章，会把资源从制度设计推回到总量调节，这个后果值得一试认真的检验。

分层引用：本章可以被分三层引用。第一层是“空关”这个构念作为核心判据，最强也最易倒；第二层是两轴四格作为分析工具；第三层是一条不依赖前两层的经验主张——在可读证据的生成成本大幅下降的行当里，判定流程中通过率饱和的关口比例会上升，而这一上升不伴随该流程总关口数的下降。第三层可以单独被检验、单独被引用，第一层被推翻时它仍然站得住。

■ 十九·不适用的边界

本章不适用于以下场合，写清楚是为了不让它被拿去做它不能做的事。

不适用于以防止为功能的关口。

安全联锁、适航检查、出厂检验、传染病筛查——那里饱和是设计目标，本章的读数方向相反。

不适用于以评价而非分配为功能的判断。

什么算好作品、哪个方向值得研究——那里的“无主”是健康，不是病。

不适用于仪式性关卡。

有些关口从设立起就没有分辨功能，它的用处是标记一次转换（宣誓、答辩、开工仪式）。用饱和读数去判它，判出来的是一个正确而无用的结论。

本章不主张分辨越多越好。

一个社会完全可以合理地决定某些位置不该被分辨——抽签、轮值、普惠、按需分配。本章只主张一件事：如果你打算用一道关来分配，就要有人为它的失效负责。分配，是另一条路，本章不反对它，也不讨论它。

决定不用关来

这道关拦下过谁

■ 二十·伦理与执行边界

饱和关占比是一个几乎必然会被拿去考核人的读数。因此三条硬性约束写在这里，缺一条这个读数就不该被使用。

其一，这个读数指的是位置，不是人。

一道空关的存在是流程的性质，不是某个审核员不称职。把它算

到具体审核员头上，是对这个读数的误用——按本章的机制，一个尽职到极点的审核员在一道空关上也只能一路通过。

其二，不得为了把这个数做好看而制造拒绝。

尤其不得在安全关键系统里为此安排轮换、人为增加变

动、或降低通过标准。第十七节已经说明这是这个读数最省事也最有害的达标方式。

其三，已经据此对个人作出不利安排的机构——不予录用、不予晋升、撤销资格——必须公开编码规则，并给出复核通道。一个不公开规则的分辨读数，恰好就是本章所反对的那种东西的新版本。

■ 二十一·结论，与一条写死日期的赌注

把全文压成一句：当生成合格证据的成本对所有人同时塌到接近零，可读门槛就整体失去分辨力；而由于门槛通常没有作者，它们不会被撤下，只会被继续满足——人工智能时代的内卷，就是这段继续满足的时间。

出路因此不是三家处方里选一条，是按次序走完三段：先由一个具名的位置公开废止已经饱和的关，再给"接住"这个动作记账并让它有代价，最后才谈什么值得做。而最小的插手点在第一步之前——给某一道关指定一个作者，一份文件、一个字段、一年一次回答"这道关拦下过谁"。

最后交一条写死日期的赌注：>到二〇三〇年十二月三十一日，在至少三个国家、至少三个不同行业里，会出现公开的、指名的、把某一道既有筛选关口正式废止并说明理由的制度条款，且理由中明确提及该关口不再产生区分。

什么不算命中，一并写死：机构自行声明"我们优化了流程""我们简化了申请"不算；把两道关合并不算；引入新的检测工具不算（那是加关不是废关）；内部纪要而未公开不算；把废止理由写成"降本增效"而不提区分度，不算。

若到期没有出现，本章关于"作废需要作者"的处方在制度上不可行，应当被放弃——而第五节的诊断与第十八节第三层的经验主张仍可独立成立。这正是本章采用分层写法的原因。

■ 附·本章自己是不是它所描述的那样东西

是。

这道关拦下过谁

本章面对的发表通道有一串可读门槛：字数、格式、

体例、术语规范、伦理声明。这些关的通过率极高，真正的分辨发生在一次不公开、不记录的判断里。而本章没有任何位置可以宣布这些关作废——它能做的只有再交一份合格品。

本章因此是它所描述的那样东西的一个标本：它给这条通道增加了一份合格产出，而它对分辨力的贡献未知，并且不可由它自己读出。这个定位有一个具体后果：>本章不能靠被发表来证明任何事。它只能靠一件事被检验——有人据此去清点一次自己流程里各道关的拦截次数，并把结果公开。

如果没有人这样做，那么本章就恰好是它自己所描述的那种东西：一份合格的、从未被接住的产出。这不是谦辞，这是本章自己的判据用在本章身上得到的读数。

■ 注释

一·本章所用的"可读"取其字面义：可被第三方读取并据以判定。它不涉及任何关于可读性优劣的价值判断。一道关可读是它的优点（可核查、可申诉、可扩容），也是它可被生成的原因；本章不主张放弃可读，只主张为可读门槛配一个作者。

二·"饱和"的阈值取零点九五是一个操作性选择，不是一个理论常数。任何在零点九到零点九九之间的取值都不改变本章的结论，但会改变具体机构的读数，因此比较不同机构时必须使用同一阈值并注明。

三·本章中的"不可生成凭据"不等于"人际关系"。一份第三方托管的、可核查的履约记录同样不可被申请人自行生成，而它是可扩容、可申诉的。第九节第二步所说的"给接住记账"，指向的正是这一类——把分辨迁到可被追究的位置上，而不是迁到私人关系里。

四·本章为理论建构与研究设计，不报告任何虚构的样本、统计结果或实验结论。文中引用的所有数字均来自公开来源，出处见

。

门槛侧与否决侧的乘积

■ 一、这一章为什么放在这里

编一的三章各自成立，放在一起却有一处正面矛盾，不处理它，这本书就不能往下走。

第三章说：门槛整体饱和，通过率趋近于一，那些关谁也拦不下。第一章说：绝大多数投递被否决，而其中相当一部分没有任何人打开过。同一个系统不可能既"几乎全过"又"几乎全否"。

两章都不是错的。矛盾之所以出现，是因为两章各自只看了一条链上的一段，而把自己那一段当成了整个系统。把它们接起来之后，会长出一样两章各自都写不出来的东西——一条关系式，以及一个由这条关系式直接算出来的、反直觉的政策后果。

这一章因此不是编一的附录，而是编一的出口。本书后面三编所有的读数，都要经过这一章确定的那个形式：代价按加法承担，回报按乘法衰减，而承担代价的一方不是这条链的所有者。

■ 二、两个读数锁在一起

本节写在结论之后，因为它不是本章原有的推理，是本章与本书第三章第三章《这道关拦下过谁》放在一起对读之后才出现的东西。第九节把两篇判为"近乎正交且互补"，那句话现在看是偷懒的——两个读数不是并排放着，它们被一条关系式锁在一起，而这条关系式两篇各自都写不出来。

17.1 先看一处正面矛盾那一章说的是：可读门槛整体饱和，通过率趋近于一——关拦不下人。本章说的是：未阅率趋近于一——绝大多数投递被否决，而否决时没有人看过。

同一个站不可能既放行所有人又否决绝大多数人。两篇若都对，它们说的必定不是同一个站。

事实正是如此。一条选拔流水线上至少有两类站：关（形式检查、关键词筛选、资格要求、查重与来源核验、投稿数量上限）与决定（最终那个"要"或"不要"）。工具把这两类站推

向相反的方向：关的分辨力塌向零，因为它要求的东西人人都能生成；决定的过目率塌向零，因为通过关的量已经超出可过目的产能。

两篇不是两个独立的机制，是同一条链上的上游与下游。

17.2 一条把两个读数锁住的关系式设投递总量 N ，链上有 k 道关，第 i 道的通过率 p_i ，下游可过目的产能 K （一轮里人真能读完的份数）。下游收到的量是 $N \cdot \prod p_i$ 。当它超过 K ，超出的部分只能不看就处置。于是： $u = \max(0, 1 - K \div (N \cdot \prod p_i))$

两种「不」

这不是一条相关性猜测，是一条约束：上游的饱和会按乘积推高下游的未阅率。

关越是拦不下人，下游

要处置的量越接近投递总量，未阅率越高。

再定义一个把两篇合起来才有的量：有据判定率 r ，即一次投递的结果确实取决于它的内容的概率。被关拦下的那部分算内容相关（哪怕只是形式上的），通过关而被过目的那部分也算，于是门槛侧的失效与否决侧的失效，在这个量上是相乘的。

这就是第九节那句"相乘"的确切形式，此前它只

是一个比喻。

17.3 一个算例，与一处反直觉的结果取一个不夸张的设定：四道关，通过率分别是零点九九、零点九七、零点九五、零点九八，乘积约零点八九；投递六千份，一轮里真能读完六百份。则下游收到约五千三百六十份，未阅率约零点八九，有据判定率约零点二一——十份里有八份的结果与它写了什么无关。

现在分别试三条干预。

其一，撤掉最饱和的那道关（通过率零点九九那道，那一章出路的第一步）。乘积升到零点九零，下游收到五千四百一十八份，未阅率零点八九，有据判定率从零点二一降到零点二零。也就是说：撤掉一道饱和的关会降低投递者的代价，却不改善有据判定率，甚至微降——因为一道谁也拦不下的关本来就不减轻下游的负担，撤掉它只是把它顺手拦下的那百分之一也放了进来。这一条那一章写不出来，因为它的框架里没有下游那个量。

其二，恢复某一道关的分辨力（把零点九五那道压到零点五）。乘积降到零点四七，下游只收到两千八百二十三份，未阅率降到零点七九，有据判定率跳到零点六三。这是三条里最有效的一条——而那一章自己已经论证了它走不通：分辨已经整体迁出可读区，在可读区里恢复不了。

其三，把过目产能翻倍（六百到一千二百，本章第十节曾提到的扩产路径）。未阅率降到零点七八，有据判定率升到零点三一。但本章第五节那条回路会让投递量跟着涨；当 N 涨到一万二千，未阅率精确回到零点八九，有据判定率一分不差地回到零点二一。扩产能被投递量吃掉，这个算例把它算了出来。

三条都不成，剩下的那一条是本章第十节的第一步：让两种“不”在记录上可分辨。它不改 p_i 、不改 K ，它改的是 N 对失败的反应函数——投递者一旦知道自己属于哪一种“不”，加量就不再是唯一能做的事。

17.4 代价按加法，回报按乘法把两侧合起来还能看出一件更简单的事。投递者必须满足每一道关，因此他的代价是各站代价之和；而他得到的有据判定按各站的乘积衰减。加一个站，代价一定上升；有据判定率只有在这个站真的分辨时才上升。一道饱和的关，是纯粹的加法代价与零乘法回报。

内卷因此有了一个可以写下来的精确形式：在一条链上不断追加只加代价、不加分辨的站，而没有任何位置负责把它们撤下。

两种「不」

这里必须交代清楚一件事，否则这一节就是在给旧东西起新名字。乘积衰减这个形式在上游早就有名字：串联系统的可靠性是各环节可靠度之积（可靠性工程）；信息沿一条处理链只会减少不会增加（信息论里的数据处理不等式）；转化漏斗的整体转化率是各步转化率之积（市场分析）。三者都比本章早得多，本节不主张这个形式是新的。

本节主张的只是那一处不对称：在这条链上，代价由一方按加法承担，而回报按乘法衰减，且承担代价的那一方不是链的所有者。可靠性工程里链的所有者同时承担成本，因而他有动机撤掉无用环节；漏斗分析的所有者也是如此。这条链没有所有者——那一章已经论证过门槛通常没有作者，本章论证过未阅否决不构成一种处置。两处“无人负责”合起来，才使加法与乘法的背离可以长期存在而不被纠正。

17.5 一条能同时裁决两篇的检验，以及本节的封顶两个读数取自同一批日志，因此可以在同一条链上同时测出来。

若 u 可由 $K \div (N \cdot \prod p_i)$ 预测出来，则两篇是同一机制的上下游，应当合为一条判断，本章与那一章各自都只说了一半。若在 $\prod p_i$ 不变时 u 仍随 N 独立变动、且偏离该式，则两者是两个独立的机制，本节的合并无效。这条检验不需要新采集数据，招聘系统与投稿系统的既有日志里两个量都在。

最后是本节自己的位置。那一章在自评里指出，它的"空关"仍是既有对象（关本来就在）的一个新读数，不是一类原来不存在的东西；本节的情况相同甚至更明显—— r 与那条关系式都是把已有的量重新组织了一遍，没有造出任何新的存在物。

这类工作有它的结构性上限，本节不假装越过了它。它的价值只在两

处：把"互补"这句含糊的话变成一条可被证伪的约束，以及指出两篇各自的第一步互不相干、必须同做——撤掉饱和的关只降代价不改善回报，给"不"记账只改善回报不降代价。谁只做一半，谁就会在自己的读数上看到改善，而投递者的处境不动。

■ 三、这条形式在后面三编里的样子

上面那条关系式是在编一的两个读数之间锁出来的，但它的形式比它的出处更宽。把它剥到最干净的一层，它说的是：

＞一条链上，代价按加法由一方承担，回报按乘法衰减，而承担代价的那一方不是这条链的所有者。

本书后面三编处理的六件事，全部是这个形式在不同位置上的实例。写出来是为了让读者知道该往哪里看，也为了让这条形式可以在每一章被单独证伪。

编二的三章，处理的是"代价那一侧的加法"具体加在哪里。第四章说，加在每一件产出的构成里——为使它可被归属于某个人，必须与它一并生产另一层东西，而这一层在三家账本上只有成本位、没有产出位。第五章说，加在被评价方的工作量里——核实这件事本来由评价方承担，现在整体转嫁下来，且因为它与真实产出同形而无法被分离计量。第六章说，加在时间里——从做出来到取得第一个读数所需的那段时长，落在所有效率读数的分母内、分子外。三章合起来回答的是同一个问题：那笔加法，谁在付？

编三的两章，处理的是"乘法那一侧的衰减"为什么读不出来。乘积衰减要能被读出来，前提是每一站的通过率可以被稳定地计量；而计量需要两样东西：一个稳定的计数单位，和一套稳定的分辨类别。第七章拆掉前一样——生成成本趋零之后，"一件"这个单位失去下界，件数不再能充当投入的证据。第八章拆掉后一样——可分辨的种类数不是系统的自然属性，它由观察者的分辨集合决定，因此它是被生产出来的，而没有人生产它。两章合起来说的是：这条链上的乘积，目前算不出来，因为组成它的每一项都建立在两个已经失效的约定上。

编四的两章，处理的是"这条链为什么没有所有者"。可靠性工程里的链有所有者，所以无用环节会被撤掉；这条链没有。第九章给出第一层原因：链上最要紧的那份能力——认出什么算数的能力——是在不产出任何可验收之物的那段时间里被生产的，而那段时间在任何一份交付账上都是空白。第十章给出第二层，也是最后一层：那份能力不由任何一方生产，它只由别

处的低效活动顺带养成；因此它没有所有者，不是因为没有人愿意认领，而是因为在结构上没有一个位置可以认领它。

于是全书的骨架可以写成一句话：代价的加法有人付，回报的乘法算不出来，而这条链没有所有者——三件事同时成立，才使这个局面能够长期存在而不产生任何事件。

这一章到此结束。它给出的关系式只是十章里最先能被写死的那一条；后面九章各自交出的读数，都还没有被锁进任何一条式子里。结语那一节会把它们并成一张表，那张表不是一个模型，是一份体检项目清单——本书能交出的东西到此为止，再往前需要的是数据，不是推理。

编二

判断的成本被转嫁给谁

...

■ 编序

判断这件工作没有消失。它被移了位置。

这一编的三章沿着同一条转嫁路径往下走：第四章看一件产出的构成——里面有一层是为了证明它出自谁手，而三家账本只给这一层成本位、不给产出位；第五章看两侧的工时——核实这件事本来由评价方承担，现在整体转嫁给被评价方，且因为它与真实产出同形而无法被分离计量；第六章看时间——一样东西从被做出来到取得第一个读数所需的那段时长，落在所有效率读数的分母内、分子外。

三章合起来回答枢纽章留下的第一个问题：那笔按加法累积的代价，究竟谁在付。

第四章

认领物——一件东西里有多少是为了证明它是你做的

■ 摘要

「内卷」在三个互不往来的行当里被同时观察到，而三家给出的驱动互不相容：经济学把它归给可验证任务份额与既有报酬规则的不容，判定稀缺已迁往信任与验证带宽；艺术学把它归给可见成品与可无限复制的条件场的不容，判定判据已迁往来源链；工程学把它归给既有流程与新成本结构的不容，判定瓶颈已由生成迁往验证。三家共同假定了一件从未被说出的事：一件东西的生产者是给定的，因而「是谁做的」不需要被单独生产。

推翻这条假定的材料来自艺术学自己——伦勃朗研究项目用四十余年把一位画家名下

的作品数量削去近半而后部分回改，这四十年产出在艺术史的账上不算作品、只算研究：归属从来就是被持续生产出来的，只是这笔生产被记在别人名下。

本章由此提出一个此前没有本体地位的对象：认领物——为使一件产出可被归属于某个特定的人或位置，而必须与该产出一并被生产的那一层东西。它不是「记录得少」，是三家的账本都只给它成本位、不给它产出位。据此，本章对内卷给出一条与三家都不相同的判断：内卷不是投入增加而产出不增加，也不是同一判据上的形式军备竞赛，也不是验证成本超过生成成本后的开销膨胀，而是产出的构成中认领物占比持续上升，而由于三家的账本都不设这一栏，同一件事在三处同时显示为「努力增加、产出不变」。

本章给出一个不含判断词的问话（这一次交付里，有多少是为了让它可被归到你名下而做的），一个三领域量纲不同而比率相同的读数（认领占比），一张二维辨别格，一条可由既有记录低成本执行的检验，以及三条各对一家的出路。第十七节另补一个双侧读数（认领比 = 评价方为可被归属所投入的时间除以评价方为核实所投入的时间），用它把「判据收紧」与「区分能力塌陷」这两种账面相同、处方相反的情形分开，并给出一条顺序形式的失效条件与对「内卷」一词来路的清账。本章并指出一项反直觉的机制推论：在归属判据不变的前提下，引入任何提高本体产出效率的工具，都会使账面内卷加剧——不是因为人更卷了，是因为产出的构成变了而账本没变。

关键词：认领物

认领占比

归属的生产

产出位与成本位内卷

■ 一、引言：三个不相干的行当，同一笔记不进账的开销

一位咨询顾问交出一份九十页的报告。其中真正回答客户问题的部分大约二十页；剩下七十页是方法论说明、访谈记录附录、数据来源清单、以及一份详细到令人不安的"我们是怎么做到这一步的"。他知道那七十页里有相当一部分客户不会读。他还是写了。如果不写，客户会问：这些

你们是怎么得出来的？

——在一个任何人都能在四十分钟内生成一份九十页报告的年份里，这个问题的分量和三年前不一样了。

认领物

一位插画师完成一幅约稿。交付时她一并附上了十二个图层的源文件、三段共四十分钟的绘制过程录屏、以及若干张废弃草图的照片。客户没有要求这些。她主动给，因为同行群里已经有人因为拿不出过程文件而被指认为"AI

出图"、丢掉了整个客户。她画那幅画用了十四小时，录屏、整理、导出、上传用了两小时。这两小时不在报价里。

一位工程师提交一个功能。功能本身三百行。围绕它的是一千二百行测试、一份结构化的提交说明、一段在代码

里写明"此处的判断依据是某年某月的某次线上事故"的追溯记录，以及一条在内部规范里被要求填写的"AI

辅助程度"字段。团队的交付准时率、缺陷率、文档完备度都在改善。他自己的感觉是：我一年比一年忙，而我一年比一年不知道自己在忙什么。

这三个人不认识彼此，所处的行业没有共同的期刊、没有共同的会议、也没有共同的术语。但他们各自多做的那一部分，有一个共同的性质：它不是为了让那件东西更好，是为了让那件东西可以被归到自己名下。

而这一部分，在他们各自行当的账本上，都没有产出位。咨询顾问的那七十页在工时表上是成本；插画师的两小时不在报价里；工程师的一千二百行测试在项目管理软件里被归入"开

销" (overhead)。三份账本、三种记法，共同点是：这些东西被生产出来了，但它们只以负号出现。

本章要论证的是，这个共同点不是巧合，而是一处结构性的账目缺口；而正是这处缺口，使得一件正在发生的事在所有人的读数上都显示成了另一件事。

在展开之前，需要先说清本章如何裁定它所回应的那个问题。通常的提问是"内卷的根源是什么"。这个提问预设内卷是一个可被归因的现象，并期待一个原因。本章的观察表明，被归因的那个现象本身是一个读数——是"投入增加而产出不增加"这一读数——而读数依赖于账本的字段设置。因此本章把问题改切为：是什么在逼动竞争形态的改变，先动的是什么，以及为什么这一改变在三家的账本上都显示为"没有改变"？这一改切不是修辞上的迂回。它把出路的位置也一并挪动了：如果内卷是一个被账本结构决定的读数，那么出路就不在"少卷"，而在改账本。

以下的行文次序是：第二至四节分别陈述三家的驱动主张，每一家都按它自己最强的形态写出，并给出它自己承认的一处撑不住的地方；第五节定位三条驱动为何不能同真；第六节说出三家共同假定而从未说出的那一条；第七节给出推翻它的那件材料，并说明这件材料为什么必须来自三家之一自己；第八节正面建构本章的对象与读数；第九节据此重新定义内卷；第十节给出一条机制推论；第十一节说明三家各自为什么到不了这一步；第十二、十三节与最近的几种既有说法以及本书一篇同型文章划清界线；第十四节给出三条出路；第十五节给出判据、失效条件与一条写死日期的预言；第十六节交代边界、本章自身的位置，以及本章解释不了的两个洞。

■ 二、经济学：可验证的那一部分决定收益去哪，稀缺因此迁移

经济学这一侧最强的形态，不是"AI 会取代多少岗位"这种量的估计，而是一条关于约束在哪里的判断。

认领物

阿西莫格鲁在《人工智能的简单宏观经济学》中给出的推算，其力量恰恰来自它的保守。他把任务分为两类：易学的任务——有客观的结果度量可供学习；与难学的任务——依赖情境、缺乏客观的成功判据。他随后用赫尔滕定理为总体收益设界：总体全要素生产率的增益，被人工智能实际能承担的那部分任务的产值份额所限。在他的参数下，十年间的全要素生产率增益不足百分之零点六六，而当难学任务被计入时会进一步降到百分之零点五三以下。〔¹〕这个推算的关键不在数字，在它的机制：收益的上界由"可被客观验证的任务份额"决定。

一项任务只要

缺少一个可以脱离执行者本人而被读取的结果判据，人工智能就学不好它，因而这项任务的产值就不进入增益的那一侧。

这条机制推到底，就得到一个远比数字更要紧的

：如果执行能力本身变得廉价而普遍，那么稀缺就不再落在执行上，它会迁走。而它迁到哪里，由"什么还不能被廉价验证"来决定。近两年沿这条线做下去的工作把话说得更直白：在一个由智能体承担大量执行的经济里，约束不再是智力的稀缺，而是信任与验证带宽的稀缺；于是租金会向解决验证瓶颈的那些部件集中——加密来源凭证、可用于验证的高质量数据集、以及责任承保。[^2]请注意这句话里出现的第一个词：来源凭证。经济学在这里已经指到了本章要处理的那样东西，并且给了它一个位置——但给的是一个新部门的位置。在这条叙述里，来源凭证是一门新兴生意的产品，它有价格、有市场、有租金，因此它当然进产出账。这是宏观视角能看见的部分。

经济学这一家的单因主张，可以严格地写成一条动力式：>

人工智能实际能承担的任务份额（一件事外显出来的可执行程度），与既有的任务分法及报酬规则不

容；这一不容改变的是稀缺所在的位置。

它是单因的，意思是它不承认另两条驱动是原动力。在这条主张里，判据的迁移与瓶颈的改变都不是原因，是结果：判据之所以要迁到来源，是因为稀缺已经迁到了那里；瓶颈之所以从生成变成验证，也是同一件事在生产端的显形。经济学不说"三者共同作用"——那样它就什么也没主张。

它的时序读数是明确的：先动的是稀缺位；报酬规则与任务分法随后跟上，通常滞后数年；而个体的行为改变最慢，因为个体只能观察到报酬规则，观察不到稀缺位。

它自曝的那一处也很清楚：如果报酬规则被制度锁死，稀缺就迁不动。

在一个薪酬结构由编制、职级、

年资刚性决定的组织里，即使执行能力已经廉价到不值钱，收益也不会向验证能力集中——此时不是稀缺位驱动规则，而是规则反过来把稀缺位钉在原处。经济学承认这一点，并把它当作一个摩擦项处理。本章第七节要用的东西，就藏在这一类"被当作摩擦项处理"的地方。

■ 三、艺术学：可见的那一面不再能区分，判据迁到来源

艺术学这一侧的判断，与经济学在同一片现象上却走了完全不同的一条路。

认领物

它的出发点是一件在这门学科里已经被讨论了几百年的事：一件作品是谁做的，从来不是看出来的，是被判定出来的。

在大部分艺术史的时间里，这个判定依靠鉴定学——受过训练的眼睛比对笔触、构图与处理方式，在一位画家的全部作品之间做出归属。鉴定学能达到惊人的敏感度，也始终是主观的：两位同样合格的学者可以深刻地分歧，而且历史上反复如此，有些分歧持续了好几代人。〔3〕到了当代，这门学科面对的处境变了性质。当生成模型可以在几秒内产出一幅在任何可见维度上都无法与人手区分的图像时，可见物本身失去了区分能力。艺术学的回应不是去发明更敏锐的眼睛，而是把判据整个搬家：从成品搬到来源。

这一搬家已经有了工业级的形态。内容来源与真实性联盟的技术标准把来源声明用数字签名绑定到文件字节上，使得静默篡改可被检测；到二〇二五年末，该标准的 2.3 版把清单扩展到直播视频与非结构化文本，也就是说，扩展到了大语言模型的输出。〔4〕对创作者一侧，这套东西被明确表述为“数字时代的签名”：几个世纪以来画家在画布上签名，如今内容凭证是它的数字等价物——一个可验证、可持续、无法伪造的作者标记。〔5〕艺术学这一家的单因主张，同样可以写成一条动力式，而且它与经济学那条落在不同的位置上：>

可见的成品，与一个可以无限复制、无限拟仿的条件场不容；这一不容改变的是判据本身所在的位置

——它从成品迁到生成路径与来源链。

它也是单因的。在这条主张里，稀缺迁移不是原因而是结果：来源之所以值钱，是因为它已经成了唯一可用的判据；瓶颈之所以在验证，也是因为验证的对象已经从“东西对不对”换成了“东西是谁的”。艺术学不承认成本结构是原动力——成本结构在它看来只是判据迁移之后的账面显形。

它的时序读数：先动的是判据；产业标准、法规与市场价格随后跟上，通常滞后一到三年（内容凭证从提出到进入相机硬件与平台强制标注，大约就是这个量级）；而作品形态最后改变。

它自曝的那一处，是这门学科自己最先说出来的：来源链本身是脆的。

在内容被再混合、再格式化、跨

平台再发布的过程中，来源信息经常丢失；因此单靠技术方案不足以支撑永久的来源记录。〔6〕

也就是

说，判据迁到来源之后，来源本身又依赖于一个它自己无法保证的条件场——此时不是判据驱动条件场，而是条件场反过来决定判据还能不能立得住。

■ 四、工程学：生成塌陷而验证不降，瓶颈的外显形态改变

工程学这一侧最不像哲学，也最容易被量到。

它的主张建立在一组可以直接读出来的数字上。对二点一一亿行代码变更的追踪显示，从二〇二〇到二〇二四年，重复代码块的出现频次上升到八倍，重构活动从百分之二十五降到不足百分之十，两周内被改写的代码占比从百分之五点五升到百分之七点九。[^7]

与此同时，一项针对熟悉自己代码库的资深开发者

的随机对照试验发现，使用人工智能辅助的一组完成任务反而慢了约五分之一，而参与者本人相信自己变

认领物

快了。[^8]一份覆盖一千一百余名开发者的调查把这个局面命名为"验证缺口"：写代码不再是难的那一部分，验证它才是。[^9]更值得注意的是一处方向相反的读数：某年度软件交付研究发现，人工智能使用度每提高四分之一，代码评审速度加快、文档质量改善，而交付稳定性下降百分之七点二。[^10]也就是说，验证的形式在改善，验证的实质在退化，而这两件事同时发生、并且被同一批指标记录。

工程学这一家的单因主张：>既有的开发流程与工作分法，与一个生成成本已经塌陷而验证成本没有同步下降的成本结构不容；这一不容改变的是系统的外显形态——它从"造得出"变成"验得起"。

它是单因的。在这条主张里，稀缺迁移与判据迁移都是被驱动的：稀缺之所以迁往验证，是因为成本比先倒挂了；判据之所以要迁到来源，是因为在成本比倒挂之后，"这段东西哪来的"成了验证的第一道工序。

工程学不承认判据是原动力——判据在它看来是被成本逼出来的。

它的时序读数：先动的是成本结构；架构与流程随后跟上，通常滞后半年到两年；而"什么算完成"这个定义最后改变。

它自曝的那一处，恰恰藏在上一段那个方向相反的读数里：如果成本比是自变量，那么当验证被自动化时，成本比应当回落。但研究显示，人工智能驱动的代码评审在面对已知的漏洞样本时，频繁漏掉注入类、跨站脚本类与不安全反序列化这些高危类别，反而更常报出低危项。[^11]更根本的是那条百分之七点二：形式验证在改善而实质稳定性在下降，说明真正在变的不是成本比，而是"什么算验证过了"这个判据——而判据不由工程学自己决定。工程学承认这一处，并把它归入"组织因素"。

■ 五、冲突定位：三条驱动不能同真

到这里，三家的主张已经各自成形，且它们的不相容不是措辞上的，是结构上的。

三条动力式落在三个不同的位置。经济学说，是"外显出来的可执行程度"与"既有的分法与规则"这两者的不容，改变了"稀缺所在的条件场"；艺术学说，是"可见的成品"与"可复制的条件场"这两者的不容，改变了"判据与组织路径"；工程学说，是"既有流程"与"成本结构"这两者的不容，改变了"系统的外显形态"。

三条式子里，被驱动的那一样各不相同——一个是条件场，一个是判据，一个是外显形态。

这使得三家的分歧无法靠"判谁对"来化解。它们不是对同一个问题给出了三个相反的答案；它们是对"什么在驱动什么"这件事给出了三个互相取消的指派。若经济学对，则艺术学所说的判据迁移只是稀缺迁移的一个后果，不能被当作原动力，而艺术学恰恰主张它是；若艺术学对，则工程学的成本比倒挂只是判据迁移在生产端的账面显形，而工程学恰恰主张成本比是自变量；若工程学对，则经济学所说的稀缺迁移不过是成本比倒挂之后的价格反应，而经济学恰恰主张稀缺位先动。

认领物

三家的时序读数也互相矛盾，而时序是可查的。经济学说先动的是稀缺位，艺术学说先动的是判据，工程学说先动的是成本结构。三者不可能同时先动。

还有一处更硬的地方：三家的自曝彼此指认。

经济学自曝"若规则被锁死，稀缺迁不动"——这正是艺术

学所说的判据先行；艺术学自曝"来源链本身依赖一个它保证不了的条件场"——这正是工程学所说的成本结构先行；工程学自曝"真正在变的是什么算验证过了，而这不由工程学决定"——这正是经济学所说的规则与激励先行。三家各自撑不住的那一处，恰好是另一家主张的原动力。

这个三角有一个诱人的读法：三条驱动构成一个循环，谁先动取决于这一轮从哪里起。这个读法是对的，但它不是本章的判断——它只是承认三家都对了一部分，从而把三家的争论保留了下来，只是给它换了个和平的说法。本章要做的不是调解，是取消这场争论所站立的那块地。

而要取消它，得先说出那块地是什么。

■ 六、三家共同假定了什么

把三家的争论压成同一个句式，它是这样的：>三家争的是——多出来的那部分努力，该被记成什么、算在谁的账上。

经济学把它记成向新稀缺位迁移的租金；艺术学把它记成判据迁移后的鉴定与凭证支出；工程学把它记成验证开销。三种记法不同，但它们记的是同一笔东西，而且都记在同一侧。

现在问一句三家都没被问过的话：这笔东西，为什么是"多出来的"？

回答这个问题，需要先承认一件三家都默认、但没有一家说出口的事：>一件东西的生产者是给定的。

这句话听起来近乎废话，而这正是它作为共有前提的标志。三家都默认：当一件东西被做出来，"它是谁做的"是这件事的一个既成属性，随着生产行为一并成立，不需要额外的动作，因而也不需要额外的生产。归属是读出来的，不是做出来的。

正因为默认了这一条，三家才会不约而同地把"为使产出可被归属而多做的那部分"记在成本侧。既然归属本来就成立，那么为它花的一切力气就都是额外的、摩擦性的、本可省去的——是交易成本，是信号成本，是开销。它当然只能有负号。

把这句话念给三家听，没有一家会反对。经济学会说：那是自然，生产函数的输入端本来就写着谁投入了什么。艺术学会说：那是自然，一幅画当然是某个人画的，鉴定学的存在只是因为我们后来把这件事弄丢了。工程学会说：那是自然，提交记录里有作者字段。

三家都会说"这还用说吗"——这正是共有前提的验收标准。

家的立场，不是三家共有的地基。

一条会引起争论的假定，说明它是某个第四

认领物

而本章接下来要做的，是从三家之一自己的材料里，把这条地基掀掉。

■ 七、推翻它的那件材料，来自艺术学自己

推翻一条共有前提，从外面搬一个理由来是没有用的——那只会让本章成为争论中的第四家，而争论本身安然无恙。要取消这场争论，推翻它的那件材料必须是三家之一自己已经掌握、已经发表、只是从来没往这个方向用过的事实。三家无法退回自己的阵地，因为那件材料是他们自己的。

艺术学有这样一件材料，而且它大得几乎不可能被忽视。

一九六八年，一个由荷兰学者组成的团队启动了伦勃朗研究项目。它的目标是把归在伦勃朗名下的作品逐一重新审定。项目持续了四十余年，到二〇一一年正式结束。在此期间，伦勃朗名下的作品数量被大幅削减——从鼎盛时期的六百幅上下，一度被压到三百幅左右；而随着方法与人员的更替，一部分先前被剔除的作品又被改回。这个过程的每一步都伴随着市场价格数十倍的跳动、博物馆标签的重写、以及学者之间持续数十年的公开分歧。

在艺术史的账本上，这四十年产出叫什么？它叫研究。它进的是学术著作的目录，不进作品的目录。伦勃朗名下的画有多少幅，是艺术史的正项；而"确定这些画是不是他的"所耗费的四十年、一个国际团队、六卷本的《伦勃朗绘画全集》，在作品这一栏里是零。

现在把这件事换一个问题来读。

艺术学用这四十年回答的问题是：哪些是真的。这是一个鉴定问题，答案是一份名单。

但这四十年同时无可辩驳地展示了另一件事，而艺术学从来没有把它当作一个

来陈述：>一件东西的归属，不是一个随生产行为一并成立的既成属性。它是被持续生产出来的，需要投入、需要方法、需要机构，会被改判，而且改判本身也需要再一次投入。

如果归属是给定的，四十年是在干什么？如果归属是读出来的，为什么两位同样合格的鉴定家会给出相反的读数，而且这种分歧能延续几代人？如果归属随生产一并成立，为什么它会在生产者去世三百年后被反复重写？

艺术学知道这一切。它甚至有一整套词汇来描述这套生产：来源研究、鉴定学、科学分析、以及近年加入的算法辅助归属——有人已经把它形容为在来源、鉴定、科学分析这张三脚凳上加了第四条腿。〔¹²〕而与此同时，另一些从业者对算法归属提出了深刻的保留，担心的不是机器会判错，而是算法判断这个前提本身会把艺术创作的复杂性压平。〔¹³〕这场争论本身再一次证明：归属是一件需要被生产、且生产方式可被争论的东西。

认领物

艺术学只是从来没有把这条

搬到生产那一侧去。它把归属的生产记在鉴定的名下——一件事后的、由第三方进行的、与创作无关的活动。因此在艺术学的账上，归属的生产从来没有出现在作品的成本里，更没有出现在作品的产出里。它被记在别人名下。

而这正是本章所需要的那一击：如果归属是被生产的，那么"生产者是给定的"这条三家共有的假定就是假的；而一旦它是假的，那笔被三家一致记在成本侧的东西，就不再是"多出来的"。它是产出的一部分，只是没有栏位。

需要说明的是，这件材料之所以能推翻共有前提，恰恰因为它在艺术学自己的领域里回答的是另一个问题。它不是艺术学对本章命题的支持——艺术学从未陈述过本章的命题。它是一件被反复引用的常识，只是从来没有人把它当作关于"归属是否给定"的证据来使用。

■ 八、认领物

8.1 定义与边界认领物：为使一件产出可被归属于某个特定的人或位置，而必须与该产出一并被生产的那一层东西。

它有四个构成要件，缺一不可：其一，它是被生产的。

它消耗时间、注意力与材料，有可清点的实体形态——页数、行数、文件数、时长。它不是一种态度，不是一种感受，也不是一个比喻。

其二，它的功能是归属，不是效能。判据是一句可以直接问的话：拿掉它，这件东西本身还好用吗？若答案是"一样好用，只是别人不知道是我做的"，那就是认领物。一份报告的方法论附录、一段绘制过程录屏、一千二百行覆盖既有行为的回归测试，都通过这一判据。而一项使功能更正确的重构不通过——它改变的是东西本身。

其三，它在现行账本上没有产出位。

这是本章最要紧的一条，也是它区别于一切既有概念的地方。认领

物不是"没被重视"，不是"价值被低估"，是三家的账本里根本没有它的栏位。经济学有"信号成本"这个词，但那是成本；工程学有"开销"这个词，也是成本；艺术学干脆把它归入"鉴定"，划到创作之外。三家都能指到它，三家都只能用负号指到它。

其四，它的量由归属判据决定，而不由产出本身决定。

同一件东西，在一个归属判据宽松的环境里几乎

不需要认领物；在一个归属判据严苛的环境里，认领物可以数倍于产出本身。这一条决定了认领占比是制度变量而非个人品质变量，因而也决定了出路必须在制度侧。

有两条边界必须先划清，否则这个概念会膨胀到没有内容。

边界一：认领物不等于无用。

一份方法论附录可能确实提高了报告的可信度，一千二百行测试可能确实

防住了一次事故。本章不主张它们浪费。本章只主张它们的功能是归属，因而它们的量应当随归属判据的

认领物

变化而变化，而不应当被当作产出总量的自然增长。把认领物记进产出位，不是为了赞美它，是为了让它的增长可被看见。

边界二：认领物不等于形式主义。形式主义的判据是"做了而没有实际功能"；认领物的判据是"有实际功能，而该功能是归属"。一份没人读的会议纪要如果连归属功能都没有，那是形式主义，不在本章的射程内。二者可以叠加，但不是一回事。

8.2 一张二维辨别格只给一个新名字，得到的还是一个名字。要使它成为一个能把两样看起来一样的东西分开的读数，必须加第二根轴。

认领占比低

认领占比高

产出总量①实增长——多出来的东西是本②增加

体产出；账面与实况一致

物；账上不设该栏，故显示为"投入增加、产出不变"

产出总量④真停滞——什么也没多；账面③不增加

账面内卷——东西确实多了，但多出来的主要是认领

与实况一致

空转——只剩认领物，本体产出已停；账面同样显示

为"投入增加、产出不变"

四格中要打的是第二格。

这一格的选择不是任意的。一张辨别格里若被打的那一格是一眼就能看出的坏，那说明第二根轴选错了——那一格不需要任何框架，任何人都看得见。第三格（空转）就是这样一格：只剩下认领物、本体产出已停的团队，稍有经验的人走进来半天就能察觉。它不需要本章。

第二格则相反，它有三处签名，缺一不可：签名一：它在短期指标上表现为改善。

认领占比高的组织，其文档完备度、可追溯性、交付规范度、评

审覆盖率通常都在上升，而这些全部是正向指标。第四节引的那条读数正是它的实证形态：人工智能使用度提高，评审加快、文档改善——同时交付稳定性下降。改善的那几项恰好是认领物的量。

签名二：它的失效不产生任何信号。

账本上，它显示为"投入增加而产出不增加"。这个读数有一个现成

的、听起来完全合理的解释——竞争加剧、人不够努力、行业进入存量博弈。这个解释一旦被接受，问题就被归到了个体的心态与行业的周期上，而"归属机制的成本正在膨胀"这个信号从头到尾没有出现过。没有任何一块仪表盘上会亮起写着"认领占比"的那一盏灯，因为没有那盏灯。

签名三：它自我加固。越是加强归属核验——署名规范、防生成检测、来源声明、AI 使用披露字段——认领的成本就越高，认领物的量就越大，认领占比就越高。而认领占比越高，账面上的"投入增加而产出不变"就越明显，于是就越容易被读成"大家不够努力/竞争太激烈"，于是就越倾向于加强核验与考核。每一轮正确的治理动作都在把这一格加深。

三条签名同时成立，第二格才是本章要打的那一格。

认领物

8.3 认领占比一个跨领域的判断，若不能给出一个三领域量纲不同而比率相同的读数，就只能停在概念上。

认领占比：在一个给定周期内，为使产出可被归属而生产的那部分，占全部被生产之物的比例。

三处的量纲各不相同，比率相同：· 咨询与知识生产：一份交付物中，用于说明来源、方法与过程的篇幅与工时，占全部篇幅与工时的比例。

· 视觉创作：一件作品的交付包中，过程文件、录屏、草图、来源凭证的体量，占交付总体量的比例。

· 软件工程：一个周期内新增的代码里，测试、日志、追溯性元数据、提交说明与使用披露所占的行数比例。

这个读数须满足四条性质才值得造出来，本章逐条对照：其一，无量纲。

三处都是比率，可以并排比较。页数比页数、体量比体量、行数比行数，跨行业的横向读数因此成立。

其二，可事后清点。

篇幅、行数、文件数、时长全部已经躺在既有记录里——工时系统、版本控制、交

付归档。这个读数不需要新建任何数据采集，只需要给既有数据加一个分类。这也是它作为检验手段最便宜的地方。

其三，它把一个印象变成一个数。

"卷"是一个感受，因而无法被争论、无法被比较、也无法被治理。认

领占比是一个可以逐年画出曲线的数。

其四，它与既有的主要指标正交。

这是四条里最容易漏、也最要命的一条：一个与既有主要指标高度相

关的读数，会被立刻读成那个指标的代理，等于白造。要检验正交，办法是举出一个"既有指标最好看、而这个数最难看"的真实例子。

这个例子是现成的：一个交付准时率最高、缺陷率最低、文档完备度最好、测试覆盖率接近满分的工程团队，正是认领占比最高的团队。它的所有传统指标都在最优区，而它新增产出中用于证明"这段东西是可信的、是谁写的、经过了什么"的比例可以超过一半。反过来，一个通宵写出可运行原型的临时小组，传统指标一塌糊涂——没有测试、没有文档、没有提交规范——而它的认领占比接近于零。

两个方向都能举出实例，正交因此成立。这也顺带说明了本章与"质量管理"的分界：认领占比高不等于质量高，也不等于质量低；它是另一根轴。

8.4 一句不含判断词的问话一个判据若含有"应当""充分""真正""恰当"这类词，它就只能被拿去问人的判断，不能被拿去问文件。本章的判据是这一句：

认领物

>这一次交付里，有多少是为了让它可被归到你名下而做的？

它可以被直接拿去问工时表、版本记录、交付归档与验收清单，不需要任何人对"合理不合理"表态。

把它在五个具体场景里跑一遍，五个答案应当互不相同，且其中至少两个不能从命题本身推出来，必须去查：1.一篇博士论文——文献综述与方法章中用于证明"我读过、我会做"的部分。这一条可以从命题推出。2.一次开源提交——提交说明、测试、变更日志、签名。这一条也可以推出。

3.

一幅商业插画的交付包

——过程录屏、分层源文件、废稿照片。这一条推不出来：录屏这件事是最近两年才成为行业默认动作的，不查行业实况不会知道它已经进入交付包。4.一场线上课程——讲师出镜、自我介绍、资历陈述、以及水印。这一条也推不出来：在纯知识传递的意义上这些全部可以省略，它们存在的唯一理由是归属；而它们占据的时长比例在不同平台差异极大。5.一份医院的手术记录——其中用于说明"由谁主刀、谁在场、依据哪一条指征"的部分。这一条介于两者之间，且它引出一个本章必须承认的边界：在医疗与司法这类领域，归属的生产同时承担着追责功能，因而不能仅按"拿掉它东西还好用吗"来评价。第十六节会回到这里。

五个答案互不相同，其中至少两个必须去查行业实况才能给出。这个判据因此不是命题的复述。

■ 九、内卷的重新定义

现在可以把本章的判断完整地写出来。

>内卷不是投入增加而产出不增加——那是经济学的划界；也不是同一判据上的形式军备竞赛——那是艺术学的划界；也不是验证成本超过生成成本之后的开销膨胀——那是工程学的划界；而是产出的构成中认领物占比持续上升，而由于三家的账本都只给认领物成本位、不给产出位，同一件事在三处同时显示为"努力增加、产出不变"。

这条判断改变的不是对内卷的评价，是对内卷这个读数的解释。

通行的内卷定义包含一个隐含的事实主张：产出真的没有增加。本章主张这个事实主张在很多场合是错的——产出增加了，但增加的那部分落在认领物上，而认领物没有产出位，因此它在总产出上不显形。

这一改判有三处可以立刻检验的后果。

第一，它解释了那种"忙而空"的特定质感。

通常的内卷叙事说人在做无用功。但当事人的体验往往不是

这样：他知道自己做的那些事是有用的——那份方法论附录确实让客户放心，那些测试确实拦住过事故，那段录屏确实洗清过一次指控。他的困惑不是"我在做无用功"，是"我做的都有用，为什么加起来等于没做"。本章的回答是：加起来等于没做，是账本的性质，不是你的工作的性质。

认领物

第二，它解释了为什么内卷对"提高效率"免疫。

如果内卷是投入过多而产出不增，那么提高单位投入的

产出效率应当缓解它。但实际观察是相反的——效率工具进入之后，账面内卷通常加剧。本章的解释见下一节，而这里先记下：一个对效率提升免疫、甚至被效率提升加重的现象，其原因不可能在效率那一侧。

第三，它把责任从个体挪到判据。

认领占比由归属判据决定。判据严则占比高，判据松则占比低。个体

在其中几乎没有选择余地——不做认领物的人会在第一次被质疑时失去客户或岗位。因此把内卷归因于个体的"心态"或"不够聪明地竞争"，在机制上是错的；而这一点本身也可被检验：若认领占比与个体特质的相关，显著弱于它与所在组织的归属判据严苛度的相关，则责任在判据一侧。

■ 十、为什么这个时代必然加剧：一条机制推论

前面九节都是在陈述一个结构。这一节给出它的一条推论，而这条推论是反直觉的，也是本章最愿意被拿去检验的东西。

推论如下：>在归属判据不变的前提下，引入任何显著提高本体产出效率的工具，都会使账面内卷加剧。

机制是三步：第一步：本体产出的生产成本下降。

这是工具的全部意义。写代码、写报告、画图从若干小时降到若干

分钟。

第二步：认领物的生产成本不但不降，反而上升。

这是关键的一步，也是与直觉相反的一步。理由是：

认领物的功能是区分——把"我做的"从"不是我做的"里分出来。而当任何人都能在几分钟内产出一件外观合格的东西时，需要被区分开的对象数量激增，区分所需的证据强度也随之提高。三年前一句"这是我做的"就够了；现在需要过程文件、需要来源凭证、需要签名链。能力的普及使得区分变难，而区分变难直接等于认领成本上升。

第三步：账本不变。认领物依然没有产出位。

三步合起来：本体产出的量在时间上被压缩，认领物的量在同一时间内膨胀，于是认领占比必然上升；而由于认领物只有成本位，总投入的上升被记录下来，总产出的上升没有被记录。账面结果就是：投入显著增加，产出显著不变。

这条推论有一个可以立刻去核的间接印证。第四节引的那条读数——人工智能使用度提高四分之一，评审速度与文档质量改善，交付稳定性下降——按通常的读法，它是一个令人困惑的悖论：为什么文档更好了、评审更快了，交付反而更不稳？按本章的读法，它一点也不困惑：改善的那几项正是认领物，下降的那一项才是本体产出。

两组数并不矛盾，它们是同一件事在两个栏位上的显形，只是其中一个栏位有名字，另一个没有。

认领物

这条推论也顺带解释了一件在三家框架里都不好解释的事：为什么在生成能力最强的那些行业，内卷感最强烈。

通常的解释是那些行业竞争最激烈。本章的解释是：那些行业的本体产出成本下降得最多，因而

认领占比上升得最快，因而账面内卷最严重——它与竞争烈度可以完全无关。

这一条是可分辨的：若在

竞争烈度相当而生成能力差异悬殊的两个行业之间，内卷感的差异由生成能力预测而非由竞争烈度预测，则本章这一条成立。

■ 十一、三家为什么各自到不了这一条

在划清与既有说法的界线之前，需要先说清一件事：三家没有走到这一步，不是因为疏忽。一个领域看不见什么，通常不是它的缺陷，是它的问题结构的影子。

经济学到不了，因为它的分析单位是市场，而认领物不过市场。

经济学能够、也确实看见了归属凭证的

经济价值——它把来源凭证识别为租金迁移的目的地之一。但它看见的是一个新部门的产出：有买方、有卖方、有价格。而本章所说的认领物绝大部分从不上市——插画师那两小时不开发票，工程师那一千二百行不单独计价，顾问那七十页不单列报价。它们发生在同一个生产者的同一次交付内部，是产出构成的改变，不是部门的新增。对一门以交换为分析单位的学问而言，一样从不被交换的东西，不进入它的视野不是失察，是它的分析单位使然。

艺术学到不了，因为它把归属的生产划在了创作之外。

这一点在第七节已经展开：艺术学掌握全部材

料，但它把这些材料组织在"鉴定"这个题目下，而鉴定按定义是事后的、第三方的、与创作分离的活动。

只要归属的生产被归给鉴定，它就永远不会出现在"作品的生产成本"里，更不会出现在"作品的产出"里。

艺术学的问题是"哪一件是真的"，而这个问题的形式本身就把归属预设为一个待发现的事实——一个待发现的事实，当然不需要被生产。

工程学到不了，因为它把认领物记成了质量。

测试、日志、追溯性元数据，在工程学的价值体系里不

是"归属的成本"，是"质量的组成部分"，而且是被高度推崇的那一部分。一个测试覆盖率高、日志完备、可追溯性强的系统，在工程学的的所有评价标准里都是好系统。于是工程学不但看不见认领物在膨胀，它还在主动推动这种膨胀，并把它记为进步。这不是工程学的错误——那些东西确实提高了系统的可靠性。但正因为它们同时有质量功能，它们的归属功能就被完全吸收了，永远不会被单独计量。

三处"到不了"有一个共同形状：每一家都能指到认领物，但每一家指到的都是它在自己账本里被记成的那样东西——租金、鉴定、质量。三家各自看见了它的一个投影，而没有一家有那个立体。

■ 十二、与最近的几种说法的分界

一个新名字若能被某个既有概念完整吸收，它就只是一次改名。本节把本章的判断与八种最容易被混为一谈的说法逐一对照。每一条给出该说法说到哪一步、分离线落在哪里，以及一条能分出胜负的对照预测。

认领物

需要先分清两件常被混为一谈的事。竞争解释是对同一现象的另一种因果说法（"其实不就是累的吗"），它归第十五节的失效条件管。竞争概念是学术史上已经有人命名过的那样东西，它可能与本章说的是同一个结构，只是名字不同——那归本节管。一个判断可以打赢所有竞争解释，却仍然只是某个既有概念的改名。只有本节能逮住这一种。

在逐条对照之前，还有一道更基本的检查。把本章的命题去掉一切本领域的名词，只留下"什么东西对什么东西做了什么，导致什么"，它的纯形式是：一类被生产之物因缺少账目栏位而在总量上不显形，于是总量的读数与实况背离。

这个形式在上游学科里叫什么？它最接近的两个是会计学与统计学里关于计量

边界的讨论，以及国民账户体系中关于非市场生产的长期争论——家务劳动、志愿服务、无偿照护为什么不进国内生产总值，这场争论已有半个世纪。第一条与第二条对照因此从这里开始。

还有一处必须自查：本章的形式不是"某物越成功越失败"这一类反转模板。这一点值得特别声明，因为那一类模板早已各有其名——成功陷阱、伊卡洛斯悖论、能力刚性、内卷化，谁若剥出这个形式而不点名它们，就是在给一个八十年前已被命名的结构重新起名。本章的形式不是反转，是遗漏：不是成功导致失败，是一类产出没有栏位。二者的可分辨之处在于，反转模板预测总产出实际下降，本章预测总产出实际上升而账面不变。这一条是可查的，见第十五节第二项。

其一，信号成本（斯彭斯，一九七三）。

该理论指出，在信息不对称下，一方会付出可观察的代价来传

递自己的类型，而这项代价的功能不在于生产，在于传递。这是本章最近的一位邻居，近到必须先承认：认领物的一部分确实是信号。

分离线落在成本与产出的分别上。信号理论把信号支出定义为成本，而且是它论证的关键——信号之所以可信，正因为它对低类型者昂贵。整个模型建立在"这是一笔支出"之上。本章主张认领物是产出：它是被生产出来的东西，有实体、有功能、可被使用，只是没有栏位。这不是措辞之争：若认领物是成本，则它应当随信息不对称的缓解而下降；若它是无栏位的产出，则它会随本体产出效率的提高而上升，与信息不对称的程度无关。

对照预测：取一个信息不对称显著缓解的环境（例如引入了强制来源凭证、身份可验证的平台），同时其本体产出效率显著提高。信号理论预测认领支出下降；本章预测认领占比不降反升。若观察到下降，本章可被信号理论吸收。

其二，交易成本（科斯，一九三七；威廉姆森，一九七五）。

该传统把度量、缔约与执行的费用识别为

一类真实存在而常被忽略的支出，并据此解释组织边界。它与本章共享"有一类支出被主流账本忽略"这一直觉。

分离线：交易成本理论关心的是这笔费用在哪里发生——市场里还是组织里，据此解释企业为何存在。它从不追问这笔费用是不是产出，它接受它是成本，只是主张这项成本决定了边界。本章追问的正是那个被它接受下来的分类。对照预测：交易成本理论预测组织内部化能降低这笔费用；本章预测组织内部化只是把认领物从对外的形态换成对内的形态（对外的资质证明换成对内的工时记录与流程留痕），总量不降。

若内部化后总量显著下降，交易成本理论足够。

认领物

其三，非市场生产与国民账户的边界（联合国国民账户体系；女性主义经济学关于无酬照护的长期论证）。这是本章在纯形式层上最近的一位，也是必须最认真处理的一位：它已经论证过"一类真实的生产因不在账目边界内而不显形"，且论证得非常彻底。

分离线落在为什么不显形上。无酬照护不显形，是因为它不经过市场——账户的边界是交换，照护落在边界之外。而认领物就发生在被计价的那一次交付内部：顾问那七十页在同一份合同里，工程师那一千二百行在同一次提交里。它不在边界之外，它在边界之内、却没有栏位。前者的处置是把边界推出去（把照护估值计入卫星账户）；后者的处置是在边界之内加一栏。对照预测：若把认领物按市场价格估值并计入总量即可解决问题，则本章与无酬照护是同一件事；本章预测那样做不解决问题，因为认领占比的意义不在于它的价值多少，而在于它的比例——把它按价值加进总量，恰恰会掩盖构成的改变。

其四，古德哈特定律与坎贝尔定律。

一旦一个度量成为目标，它就不再是好度量。这条定律与本章都在处理"度量与实况的背离"。

分离线：古德哈特处理的是已有度量被扭曲，其机制是行为者朝着度量优化；本章处理的是没有度量，其机制与行为者的意图无关——即使所有人都诚实、都不朝任何指标优化，认领占比该升还是升，因为它由归属判据决定。对照预测：在一个没有任何绩效考核、无人有动机操纵指标的环境里（例如个人的独立创作），古德哈特预测不发生扭曲；本章预测认领占比照样上升。这一格容易做：观察独立创作者在生成模型普及前后的交付包构成。

其五，表演性劳动与情感劳动（霍克希尔德，一九八三；及其后关于"看起来很忙"的组织研究）。

这一

支识别出一类为被他人看见而做的工作，与本章的"为归属而做"在直觉上接近。

分离线落在被看见的是什么上。表演性劳动的对象是劳动者的状态——他看起来投入、看起来忙、看起来专业；它的产物是印象，不是物件。认领物的对象是产出的来源，它的产物是可清点的实体——文件、行数、时长。二者可以叠加，但它们的量由不同变量决定：表演性劳动的量由被观察的强度决定，认领物的量由归属判据的严苛度决定。对照预测：在一个远程、异步、几乎无人观察工作过程的环境里，表演性劳动应当下降；本章预测认领物在同一环境里上升——因为无人观察过程，恰恰使归属更依赖交付物自带的证据。这一条方向相反，是本节最干净的一格。

其六，验证债与理解债（软件工程近年的概念）。工程学内部已经命名了两样东西：验证债指"关于这套软件还有哪些主张没有被充分证明"；理解债指人对自己交付的代码缺乏理解而积累的隐性负担。这是方法学上的占位者，必须点名。

分离线：这两个概念都是负债——它们描述的是"应做而未做"的欠账。认领物是已做：它已经被生产出来了，占了时间、占了篇幅、占了行数。二者在符号上相反。更要紧的是，加大验证投入会减少验证债，同时增加认领物；两个读数会朝相反方向走。对照预测：一个把测试覆盖率从五成拉到九成的团队，验证债显著下降，认领占比显著上升。若两者同向变动，说明本章测的与验证债是同一样东西。

其七，官僚制的形式合理性（韦伯）与规则的自我繁殖。认领物的增长看起来很像文牍主义。

认领物

分离线：官僚制理论的因变量是规则数量与程序层级，其机制是组织为自身的可控性而增设程序；认领物的因变量是单件产出中归属证据的比例，其机制是外部区分难度的上升。二者可以在没有对方的情况下单独发生：一个规则极少、层级极扁的自由职业者，其认领占比可以极高；一个规则极多的科层组织，若其产出的归属从不被质疑（例如内部行政），认领占比可以很低。对照预测：在规则密度相当而外部归属质疑强度差异悬殊的两组之间，认领占比应由后者预测。

其八，注意力经济与个人品牌。有一种流行的说法是：现在人人都得做自我营销。

分离线：个人品牌的对象是人，它可以脱离任何一件具体产出而存在，且它的收益是跨产出累积的。认领物依附于单件产出，不可转让，不可累积——上一次交付的过程录屏不能用来证明这一次的画是你画的。

对照预测：若个人品牌足够强，认领物应当可以省略（大师不需要证明画是自己画的）。本章预测这一豁免只在归属从不被机器质疑的领域成立，而在生成能力覆盖的领域里，声誉越

大反而越需要认领物——因为伪作的收益随声誉上升。这一条在艺术市场里有现成的历史材料。

最近的邻居是哪一位。

是其三——国民账户的边界之争。它与本章在纯形式上完全同构，且它的论证比

本章成熟得多。本章之所以仍不可被它吸收，只因为一处：它处理的是边界之外的生产，本章处理的是边界之内没有栏位的生产。这一处差别决定了两者的处置方向相反——一个是把边界推出去，一个是在里面加一栏——因而它们不是同一条判断。

■ 十三、与专栏两篇同型文章的分界

学科通融专栏之四十《轮空：发起权交出之后》与本章在方法上同型：同样以三条互斥的驱动主张为源，同样追问“谁在驱动谁”，同样最终落在一样三家账本上没有位置的东西上。二者一先一后发表于同一个月，读者有理由怀疑它们是同一篇的两次书写。这里把分界写清楚。

分界一：被驱动的那一样不同。

《轮空》处理的是改变的发起处——一个系统连续几次由同一处发起改

变之后会怎样；它的对象是不变（那一轮本该有人动而没人动）。本章处理的是产出的构成；它的对象是已发生的生产（那一层被做出来了却没有栏位）。前者是一个未发生的事件，后者是一个已发生的实体。

二者在符号上相反，因而它们的读数不可能是同一个。

分界二：经济学在两篇里站的位置不同。《轮空》里的劳动经济学主张“技术条件与既有任务分法不容，改变了劳动的外显形态”；本章里的经济学主张“可执行程度与既有分法规则不容，改变了稀缺所在的条件场”。两处的被驱动项不同——一个是外显形态，一个是条件场。这不是同一家的两次出场，是同一门学科在两条不同的动力式里各占一位。本章的外显形态那一位由工程学占据。

分界三：读数正交。

《轮空》的读数是换手率——最近若干次改变中，发起处与上一次不同的次数占

比。本章的读数是认领占比。两者在数学上不相关：一个连续由同一处发起改变的系统（换手率为零）可以有极低的认领占比（无人质疑归属），也可以有极高的认领占比；反之亦然。这可以被直接检验：在同一批组织上同时测两个读数，若二者相关系数显著，说明其中一个另一个的代理，本章应当退让。

分界四：叠加时的关系。

认领物

两篇的机制可以独立发生，但叠加时不是相加。一个换手率为零的系统，其归

属判据也倾向于长期不被修订；而归属判据长期不被修订，恰恰使认领占比在外部生成能力变化时无法被调整。因此在换手率低的系统里，认领占比的上升会更快且更不可逆。这一条是两篇合起来才有的推论，任何一篇单独都给不出。

与第三章第三章《这道关拦下过谁》的分界。那一章与本章同日发表，用的是同样三个学科，解的是同样一道题，因此必须单独分开。它主张：内卷是可读门槛的整体饱和——当生成一份合格证据的成本对所有类型同时塌到接近于零，任何可被第三方读取的关都不再分辨；而由于绝大多数关没有一个有权宣布它作废并为此负责的位置，饱和的关不会被撤下，只会被继续满足。

病灶不在同一处。那一章的病灶在判定的权属——一道关已经失效却撤不掉；本章的病灶在归属的成本——为使产出可被归到某人名下而必须一并被生产的那一层。一个是撤不掉的关，一个是多出来的开销。

二者都落在账本没有栏位的地方，落法不同。

判决性对照有两处，方向相反。其一，取一个归属成本近乎为零的系统——公司内部的代码仓库，每一次提交都由身份系统确定地记在某位具名员工名下，不需要任何额外的证明工作。本章在这里预测认领占比接近于零、账面内卷因而轻；那一章预测内卷照样出现，只要评审那一道关饱和而没有人有权废它。其二，反过来取一个归属成本极高、而判定有明确作者的场合——一个定期修订并公开作品目录的鉴定委员会。本章预测高，那一章预测低。同一份材料，两篇给出相反的读数，这比任何措辞上的区分都硬。

一条可以同时判两篇的检验。在同一批组织上同时测认领占比与饱和关占比：若二者相关系数显著，说明它们测的是同一件事，其中一篇应当被另一篇吸收；若二者正交而联合解释力显著高于各自单独的解释力，则两条机制都成立。这与分界三对学科通融专栏之四十所用的是同一种检验方式，只是这一次的对手更近，因而这条检验更值得先做。

叠加时的关系。认领物占比越高，每一件产出携带的可读证据就越多，可读门槛饱和得越快；反过来，关越饱和，人们越要靠加厚认领物来求分辨。二者互为对方的加速器，因此在同一个系统里测到的总损耗，会高于两条各自预测之和——这一条同样是两篇合起来才有的推论。

■ 十四、出路：三条，各对一家

用户所要的"出路", 在本章的框架里不是一个独立的章节, 而是判断的直接读数: 如果内卷是认领占比在没有栏位的情况下上升, 那么出路只有三个位置——给它栏位、换判据、降成本。三条恰好各对一家。

其一: 给认领物设一个产出位 (对经济学)。

这是三条里最小的一个动作, 也是唯一一条可以由单个组

织在一个季度内自行完成的。做法是在既有的工时、交付与版本记录里加一个分类字段, 把"用于归属"的部分与"本体产出"分开清点, 然后逐季度画出那条比例曲线。

认领物

它的意义不在于改变任何决策, 而在于让这条曲线存在。第八节第二条签名说过, 这一格的失效不产生信号; 而一条被画出来的曲线就是信号。至于画出来之后要不要干预、怎么干预, 是另一个问题——本章不主张认领物应当被压缩 (有些认领物承担着追责功能, 压缩它是危险的), 本章只主张它的增长应当可见。

这一条几乎不花钱, 因为所需的数据全部已经在既有记录里。这也使它成为第十五节里那条最便宜的检验。

其二: 把归属位从"谁做的"迁到"谁承担后果" (对艺术学)。

认领物的量由归属判据决定。而当前的归

属判据问的是生产的事实——这段东西是不是从你手里出来的。这个问题在生成能力普及之后正在变得既昂贵又不可靠: 它的成本随区分难度上升, 而它的答案随伪造技术的进步而贬值。

另一种判据问的是承担的事实——这件东西出了问题, 代价先落在谁身上。这个判据的成本几乎为零, 因为它不需要任何证据的生产: 它由责任的制度安排直接给出。而且它更难被伪造——伪造一份过程录屏是容易的, 伪造"我将承担后果"是不可能的, 因为后果会自己来。

这一条不是修辞。它有一个现成的形态: 在署名之外另设一个担责署名, 明确写出这件东西出问题时由谁答复。一旦担责署名成为主要的归属判据, 大量以证明"是我做的"为目的的认领物就失去了功能, 可以被撤掉。

必须承认这条出路有代价, 而且代价不小——第十六节会正面处理。

其三: 把认领做进结构, 而不是做成附加层 (对工程学)。

认领物之所以昂贵, 很大程度上因为它是附

加的: 录屏是在画完之后另外录的, 测试是在功能之外另外写的, 方法论附录是在之外另外写的。凡是附加的, 其成本就与本体产出的成本相加。

另一种做法是让归属证据成为生产过程的副产品而非附加物——版本控制的提交历史本来就是过程记录，不需要另外录屏；一个由测试驱动写出来的功能，其测试不是附加层而是规格本身。这一条在工程学内部已经有成熟的实践，本章没有新东西可加，只是指出：它的价值不在于提高质量，在于把认领成本从相加变成内含。这个理由此前没有被用来论证它。

三条的次序是有先后的：第一条必须最先做，因为没有那条曲线，后两条的效果无从判断，而未经度量的改革在几个季度之后一定会退回原样。

■ 十五、判据、失效条件与一条写死日期的预言

15.1 最强的竞争解释一个只验证"本章所描述的现象存在"的检验太弱，因为它排除不掉那句最朴素、最难反驳的话。这里先把它写出来：

认领物

>其实不就是活多了、要求高了吗？

这是本章必须排除的竞争解释，而不是一个好打的靶子。它主张：认领物的增长只是"要求提高"的一个表现，而要求提高的原因是竞争加剧、客户更挑剔、监管更严；根本不需要"归属的生产"这个中间项，也不需要认领物这个新对象。

要排除它，必须找出只有本机制成立才会出现、而"要求提高"预测不出的那个独有变量。这个变量是：认领占比的变化，应当由本体产出效率的变化预测，而不由要求严苛度的变化预测。

"要求提高"这条解释

预测：要求越严，认领物越多；它对本体产出效率没有预言。本章预测：在要求严苛度不变的情况下，本体产出效率的提高单独就足以推高认领占比。

于是失效条件可以写成一个可执行的形状：>

控制客户要求严苛度与监管强度之后，若"本体产出效率"与"认领占比"之间不存在剂量-反应关系，则

本章的核心机制为伪——届时内卷应当被归因于要求的绝对提高，而不是产出构成的改变。

关于样本，有一条纪律必须写死：要谈剂量-反应关系，样本至少需要六个以上单元，且在关键混淆变量上同质。

拿两个国家或两个行业对比是不成立的——两边在规模、文化、监管、历史上处处不同，所

谓"控制掉某变量"在方法上站不住。可用的设计是在同一制度环境内找自变量有真实变异的多个单元：同一家公司的多个团队、同一所大学的多个院系、同一个平台上的多个创作者群体。它们的要求严苛度大体同质，而生成工具的采用程度差异巨大。

15.2 六条失效条件其一（最便宜、可立即执行）。

在同一组织的若干个团队里，按第十四节第一条的做法给既有记录加一

个分类字段，回溯统计过去三年的认领占比。若认领占比在生成工具采用率显著上升的团队与未上升的团队之间没有差异，则本章的机制推论（第十节）为伪。所需数据全部已在版本控制与工时系统里，不需要新建任何采集。

其二（区分本章与反转模板）。

若在认领占比上升的同时，本体产出的绝对量也在下降，则本章错而反

转模板一类的说法对——那时发生的是成功导致失败，不是遗漏。本章预测本体产出的绝对量不下降甚至上升，只是它的占比下降。这一条把本章与成功陷阱、伊卡洛斯悖论、能力刚性、内卷化整批分开，而且只需要把两个数分开清点。

其三（区分本章与信号理论）。

在信息不对称显著缓解、且本体产出效率显著提高的环境中，若认领支出下降，则本章可被信号理论吸收。

其四（区分本章与表演性劳动）。

在远程、异步、过程几乎不被观察的工作环境中，若认领物下降，则本章测的是被观察强度而非归属判据，应退回表演性劳动。

其五（判据一侧的责任归属）。

若认领占比与个体特质的相关，显著强于它与所在组织归属判据严苛度的相关，则第九节第三条为伪，责任确实在个体一侧。

其六（对第二条出路的检验）。

认领物

在引入担责署名的组织里，若认领占比在两年内没有下降，则第十四节

第二条出路无效——那说明归属判据的迁移并不减少认领物，本章关于"认领物的量由判据决定"这一条也随之动摇。

这条判断若被证伪，我们会学到什么？

若失效条件之一成立，则说明内卷确实是一个关于总量的现象而

不是关于构成的现象。那将意味着治理方向应当回到调节投入总量与竞争强度上，而不是改账本。这是一个方向相反的处方，因此这场证伪不是白做的。

15.3 一条写死日期的预言>到二〇二八年十二月三十一日之前，将出现至少一个把"归属证据的生产"单独立项计量的正式制度文件——它可以是一份行业交付规范、一份组织内部的工时分类标准、或一份公共采购的验收清单，其中"用于证明来源与归属"的工作量被从本体交付中单独分列。

什么不算命中，须一并写死：· 仅有学术论文或行业报告提出这个建议，不算命中——必须是一份对某个具体组织有约束力的文件。

· 仅要求披露"是否使用了人工智能"，不算命中——那是一个是否题，不是一项计量。

· 仅在合同里列出"文档编写"这类既有条目，不算命中——必须能把"为使产出可被归属而做的部分"与"为使产出可用而做的部分"分开。

· 由本章的作者或其所在机构参与推动而产生的文件，不算命中。

若到期没有出现，本章关于"这条曲线迟早会被画出来"的判断即告失败；那将说明认领物在实践中并不构成一个需要被单独计量的对象，本章所命名的东西没有被现实认领。

■ 十六、边界、本章自身的位置，以及两个解释不了的洞

16.1 两个学科反过来给本章加的约束一个跨领域的判断若只从各学科取用，不被它们削弱，说明取用得不够深。这里记下两处真正削掉了本章原本打算主张的东西的约束。

其一，来自医疗与司法：第十四节第二条出路的适用范围被大幅收窄。

本章原本打算主张：归属判据应

当从"谁做的"整体迁移到"谁承担后果"。这个主张在医疗与司法领域站不住。手术记录里"由谁主刀、谁在场、依据哪一条指征"这些内容，按第八节的判据看确实是认领物——拿掉它，手术本身一样完成。但它同时是追责的唯一凭据，而追责在这些领域不是可选项。更要紧的是，在这些领域里"谁承担后果"恰恰依赖于"谁做的"被记录下来——两个判据不是可替换的，是有先后依赖的。

认领物

因此本章的第二条出路必须收窄为：它只适用于那些后果的承担可以被独立于生产事实而指派的领域。

这一收窄同时削掉了本章原本更强的一句话——本章原本会写"归属的生产在多数领域是可以被大幅压缩的"，现在只能写"在后果可独立指派的领域可以"。这是一处真实的削弱，不是补充条件。

其二，来自会计学与国民账户：本章的价值排序被翻转。

本章原本打算主张：认领物既然是产出，就应

当被计入产出总量。会计学的一条基本纪律否掉了这个处方——一个账目类别之所以有意义，在于它区分了不同的东西；把认领物加进产出总量，只会使总量这个数变大，而本章所关心的那件事（构成的改变）恰恰因此被抹掉。国民账户在处理无酬照护时采用卫星账户而非并入主表，用的正是同一条纪律。

因此本章的处方从"计入"改为"分列"：认领物必须单独成栏，而不能被加进产出总量。这一处不是措辞调整，它翻转了本章的价值排序——本章原本认为问题是"这部分产出被低估了"，现在认为问题是"这部分产出与另一部分被混在一起了"，而低估与否根本不是重点。

16.2 本章自己在哪一格按本章自己的判据审视本章：这篇文章有多少篇幅是让它可被归到作者名下而写的？

答案是：第十二节和第十三节，加上全部

与

，占全文约四分之一。这些内容没有一句在论证"认领物是什么"。它们全部在回答一个归属问题——这条判断是不是已经有人说过了，是不是只是某个既有概念的改名。若拿掉它们，本章的核心论证一字无损。按第八节的判据，它们是标准的认领物。

这意味着本章自己就落在第二格里，而且是自觉地落进去的。它无法自辩，因为它无法证明这四分之一不是必要的——在一个任何人都能在几十分钟内生成一篇结构完整、引证齐备的长文的年份里，一篇不做划界的文章会被合理地怀疑为改名，而这个怀疑本身就使划界成为必需。

这个位置有一个具体的后果，必须说出来：本章所提议的第一条出路，若被本章自己执行，会立刻显示本章的认领占比高于它所批评的大多数对象。

本章没有理由为自己开例外。这也是为什么第十四节第一条

只主张"让曲线存在"，不主张"把它压下去"——一个自己就在高位的作者，没有资格开压缩的处方。

16.3 一个读数一旦进入考核会发生什么认领占比是一个比率，而比率会被拿去考核人。这一点必须在提出它的同时处理，否则提出它本身就是不负责任的。

最省事的达标办法是显而易见的：把认领物改个名字记到本体产出里去。

把测试算作功能行数，把方法

论附录算作正文，把过程录屏算作交付物。这样做不改变任何实质，只改变分类，而分类由填表的人掌握。这条反噬预言必须写在这里，而且我看不到出口——任何由被考核者自行分类的读数，都可以被这样规避；而由第三方分类的读数，其分类成本本身又是一笔新的认领物。

因此，若这个读数被采用，须写死三条：其一，这个读数指的是位置，不是人。

一个团队的认领占比高，说的是它所处的归属判据环境，不是它

的成员偷懒或作秀。把它用于个人绩效评价，是对这个读数的误用。

认领物

其二，不得为了把这个数做好看，而在有安全或追责要求的环节削减归属证据。

医疗、司法、金融、安

全关键系统中的归属记录承担追责功能，它们的存在与这个读数的高低无关。任何以"降低认领占比"为由削减这些记录的行为，都是拿这个读数当借口。

其三，已经依据这个读数做出不利安排的组织，必须公开其分类规则，并提供复核通道。

一个没有公开

分类规则的比率，其分母与分子都由掌握分类权的一方决定，因此它不是一个读数，是一次裁量。

16.4 两个本章解释不了的洞第一个洞：为什么有些领域的归属判据几十年不动，而有些领域的判据几年一变。

本章全篇主张认领占

比由归属判据决定，但本章没有解释判据本身是怎样被决定的。学术出版的署名规范与建筑行业的执业签章制度都很稳定，而视觉创作的归属判据在三年内被彻底改写了一次。本章能指出这个差异，但给不出机制。第十三节所说的"与换手率叠加时更快且更不可逆"是一条推测，不是解释。

第二个洞：认领物是否有一个下限。

本章主张认领占比可被压缩，但没有说它能压到多低。直觉上它不

可能为零——任何协作都需要某种最低限度的归属，否则无法分工。这个下限是多少、由什么决定，本章不知道。这个洞是要紧的，因为一切关于"压缩认领物"的处方，若没有下限的概念，都可能压过头；而压过头的后果——归属彻底失效之后的协作崩溃——比内卷严重得多。

■ 十七、一个双侧读数、一条轮次检验，与「内卷」这个词的来路

前十六节已经把判断立住了。这一节补三样此前缺的东西：一个能把两种病分开的读数、一条顺序形式的失效条件，以及对「内卷」这个词本身的一次清账。三样都不改动主线，只是把它已经指出的位置量得更细。

17.1 认领占比是一个单侧读数，它读不出负担压在谁身上第八节给出的认领占比，是在生产者这一侧清点的：认领物占总产出的比例。它足以支撑第九节的改判，却有一处读不出来——它无法区分两种在账面上完全一样、而病因与处方都相反的情形。

情形甲：判据收紧，两侧同时加码。

客户、监管或评审提高了归属证据的要求，被评价的一方为此多做

认领物，而评价的一方也确实多花时间去核实。认领占比上升，评价方的投入同步上升。

情形乙：判据未动，而区分能力塌了。

没有任何人提高要求，但可用来区分的那些特征——完成度、格

式、行文、术语的正确——同时贬值，因为它们正是生成工具最先学会的部分。评价的一方不再自己去分辨，转而要求被评价的一方自己交代：拿出过程、拿出原始件、拿出更早的痕迹。认领占比同样上升，而评价方的投入不升反降——它把这份工作推了出去。

两种情形在认领占比上是同一个读数，处方却相反。甲要松判据；乙里松判据没有用，因为判据本来就没有紧过，要动的是区分能力，或者是核实成本落在谁头上。

认领物

把它们分开只需要再取一个数，而这个数正是第十四节第一条那次清点的副产品：>

认领比=

被评价的一方为使产出可被归属所投入的时间÷评价的一方为核实归属所投入的时间。

三个领域各自兑现，量纲不同而比率的意义相同。就业与人事：求职者准备简历、作品集、笔试与多轮面试的总时长，除以招聘方筛选与面试的总时长；数据在招聘系统的时间戳里。软件工程：贡献者撰写说明、拆分提交、准备演示所花的时长，除以评审者阅读与批准的时长；数据在版本控制与评审系统的时间戳里。艺术与创作：创作者为宣示来源与过程所做的记录、限量与现场所花的时长，除以收藏方、策展方或平台核验所花的时长；数据在工作日志与平台审核记录里。

这个读数要能用，须过四条。无量纲——时间除以时间，三处可以并排比较。可事后清点——三处的数据都在既有记录里，不需要新建采集，这一点与第十四节第一条同源。把一个印象变成一个数——「现在什么都要交代」是一句体感，认领比是一个可以被争论的数。与既有

主要指标正交——这一条最容易漏也最要命，一个与既有指标高度相关的量会被立刻读成那个指标的代理，白造。检验的办法是举出一个既有指标最好看而这个数最难看的实例：一个交付最快、缺陷密度最低、按任何现行工程指标都堪称优秀的团队，可以有极高的认领比，因为它把大量时间花在让工作可被归属上，而这些时间恰恰计入了开发工时，因此在所有既有指标上都表现为投入充足、过程规范。反过来的实例同样存在：一个认领比接近于零的团队可能产出很差——没有人核实它，也不需要它被归属。认领比与质量无关，与被归属的强度有关。

17.2 一次算完：它逮住一类看上去是进步的做法抽象的比率容易被读成修辞，因此把它在一个具体场合里走完一遍。

设一个学院的年度职称评审。教师一方：整理成果清单、撰写代表作说明、准备教学材料、按新要求补交数据与代码的可获取性声明、应对两轮形式审查的补正，合计约六十小时；这些时间在科研管理系统里全部落在「科研与教学工作」项下，无一项被记为「为评审准备」——这正是第八节所说的没有产出位。评审一方：五位委员，每人阅读约一小时，加上会议讨论分摊约半小时，合计约七点五小时。认领比约为八：这个体系里每投入一小时去核实，就要求被核实者投入八小时去使自己可被核实。

现在做两个反事实。

其一，把材料上限压到十页并取消附件。

教师端降到约十五小时，委员端因材料变薄降到约四小时，认

领比降到约四。两侧同降，负担确实被削掉了。

其二，引入一项可自动核验的判据——例如代表作须附可运行的复现包，由系统自动检查是否可运行——而原有的书面要求一项不撤。教师端因需准备复现包上升到约七十小时；委员端因为机器承担了一部分核实，人工阅读压到约三小时；认领比升到约二十三。

第二个反事实是这个读数最要紧的地方。按认领占比读，这一步很可能读成改善甚至读不出变化——复现包本身就是本体产出的一部分，而自动核验提高了本体产出效率，两项相抵，占比未必上升。按认领比读，它是一次明显的恶化：核实的工作被机器接走了一部分，而被核实者的负担反而加重了。

认领物

由此，第十四节第三条出路必须配一句限定。「把认领做进结构而不是做成附加层」之所以有效，前提是被它替代的那部分旧要求同时被撤掉；复现包一旦被机器检查，与之重复的书面说明就该取消。少了这一句，内含就变成了叠加，而叠加在账面上与内含几乎无法区分——认领比是目前能把二者分开的最便宜的一个数。

17.3 一条轮次形式的失效条件第十五节的六条失效条件都是横断面的：控制某个变量之后看剂量与反应的关系。它们各自有效，而它们有一个共同的弱点——相关关系比较容易被巧合满足。这里补一条顺序形式的，它更难被偶然凑出来。

把三家各自的驱动写完整会发现，每一条都是三步而不是两步：某两样不容，逼得第三样改变，而改变之后又反过来改动前两样，由此决定下一轮先动的是谁。经济学那一条：本体产出效率与既有的归属判据不容，逼得稀缺迁往验证一侧；迁移之后反过来改动判据——新的稀缺重新定义了什么算一份合格的交代，下一轮由判据先动。艺术学那一条：可见的那一面与来源判据不容，逼得判据迁到来源；迁移之后反过来改动什么算一件作品，下一轮由作品的定义先动。工程学那一条：生成能力与既有接口不容，逼得瓶颈外显形态改变；改变之后反过来改动接口——绕流程的做法被迫认为正式接口，下一轮由接口先动。

把认领物放进这个循环，就得到一条可以去查的顺序：>在一个评价体系里，若这一轮先动的是本体产出效率（生成工具普及），则认领比在六至十八个月内上升；此后若没有任何一方投资于区分能力，下一轮的发起处仍是本体产出效率，认领比继续上升；只有当发起处换到区分能力一侧——引入可自动核验的判据并撤掉被替代的旧要求、缩小待核实的范围、把核实成本明确记在评价方账上——认领比才会回落。

失效条件因此可以写死：若观察到认领比在发起处未曾换手的情况下自行回落，则本节所加的这一条为伪；届时认领比的波动应被归因于周期性因素，而不是发起处的位置。

这一条与第十五节六条的关系是互补而非替代：那六条判本章的核心判断成不成立，这一条判「出路该在什么时候动手」有没有一个可查的时点。它也顺带给第十四节末尾那句「三条的次序是有先后的」补上了一个可观察的依据——次序之所以有先后，不是因为第一条更重要，而是在发起处尚未换手的那些轮次里，后两条动了也会被下一轮抵消掉。

17.4 「内卷」这个词借自格尔茨，而它的条件在这一轮不成立最后清一笔词的账。中文里用来称呼这类困境的「内卷」，直接借自格尔茨一九六三年对爪哇水稻农业的描述。他的原意很具体：土地不可扩张，人口持续增加，于是劳动被一层层加到同一块土地上，单产微增而劳动强度剧升，与之配套的租佃、分工与仪式制度日益精细繁复——他把这种「没有外延增长时的内部精细化」叫作内卷化。

第十五节第二条已经把本章与这一类说法在读数上分开了：本章预测本体产出的绝对量不降甚至上升，只是占比下降。这里补的是更靠前的一层——两者的前提条件是相反的。格尔茨的机制以要素固定为条件；而这一轮的定义性特征恰恰是产出能力的暴涨。土地没有固定，土地在成倍增加。若仍按内卷化去读，会

认领物

得出一个明显反直觉的推论：产能扩张应当缓解内卷；而实际观察到的是相反的走向，第十节那条机制推论说的正是这件事。

这不是文字游戏，因为用错的词会指向用错的处方。按内卷化的读法，出路在于打开外延增长——找到新的土地，创造新的位置，开辟新的赛道。这条处方在这一轮里效力有限：新赛道一旦被打开，在那里生成一份看上去合格的产出同样近乎免费，归属判据同样昂贵，认领占比同样上升。近几年可以观察到的一个规律与此吻合——一个新领域从「蓝海」到「卷」的时间在持续缩短，而这与位置的多少关系不大，与该领域归属判据的成本关系很大：有客观裁判器的领域（可自动评测的竞赛、有硬性复现要求的方向、以成品性能而非过程说明定胜负的工程）迟迟不卷，而以文本、图像、方案为产出的领域几乎一开局就卷。这条规律可以直接被检验：取若干个近三年新出现的领域，把「归属判据的平均核实成本」与「从开放到出现明显认领物堆积的月数」作相关，本章预测负相关，而按位置数量作相关应无稳定关系。

因此本章在称呼上取一个保守的做法：保留「内卷」作为对现象的通称，但不沿用格尔茨的机制。现象是同一个，机制不是同一个。

这也是本章与那一批反转模板式说法之间最靠前的一条分界——它们的分歧

不在

，在前件。

结 论

三个互不往来的行当各自观察到同一件事，各自给出一条互不相容的驱动，而三家共同站在一条从未被说出的假定上：一件东西的生产者是给定的，因而"是谁做的"不需要被单独生产。推翻这条假定的材料来自艺术学自己——四十年、一个国际团队，用于确定一位画家名下究竟有多少幅画；这四十年产出在作品那一栏里是零。

一旦这条假定倒下，那笔被三家一致记在成本侧的东西就换了性质：它不是多出来的，它是产出的一部分，只是没有栏位。本章把它命名为认领物，并据此给出一条与三家都不同的判断——内卷不是投入增加而产出不增加，而是产出的构成中认领物的占比在持续上升，而由于没有那一栏，同一件事在三处同时显示为"努力增加、产出不变"。

由此得到的那条推论是本章最愿意被检验的东西：在归属判据不变的前提下，引入任何提高本体产出效率的工具，都会使账面内卷加剧。

它解释了为什么内卷对效率提升免疫，甚至被效率提升加重；它也预言

了一个可以直接去数的东西——认领占比。

出路因此不在"少卷"。它在三个位置上：给认领物一个产出位，让那条曲线存在；把归属判据从生产的事实移向承担的事实，使大量以自证为目的的生产失去功能；把归属证据做成生产的副产品而非附加层，使认领成本从相加变成内含。三条之中第一条必须最先做，因为没有那条曲线，后两条的效果无从判断。

最后一句留给本章自己。本章有四分之一的篇幅在证明它不是别人说过的话，而这四分之一按它自己的判据是标准的认领物。它没有理由为自己开例外，也无法证明这四分之一可以省去——在一个任何人都能生成一篇结构完整的长文的年份里，这个怀疑本身使划界成为必需。这大概就是这件事最不好受的地方：认

认领物

领物之所以增长，不是因为有人在装样子，而是因为在东西可以凭空出现的世界里，证明它不是凭空出现的，本身就成了一件必须动手去做的事。

■ 注释

of AI, NBER Working Paper No. 32487。其易学／难学任务之分与赫尔滕定理设界的做法，是本章第二节所述"可验证份额决定收益上界"的原始形态。

[^2]:

"约束不再是智力的稀缺，而是信任与验证带宽的稀缺"，以及租金向来源凭证、验证级数数据集与责

凭证"一词即出于此处，而本章与之的分界见第十一节。

[^3]:

鉴定学的敏感度及其固有的主观性、两位同样合格的学者可长期分歧，见

Could

AI

replace

[^4]:内容来源与真实性联盟标准 2.3 版对直播视频与非结构化文本（含大语言模型输出）的支持，见该标准的公开规范与相关综述。

[^5]:"内容凭证是数字时代的签名"这一表述，见 C2PA 面向创作者的说明文档。

[^6]:关于内容在再混合、再格式化与跨平台再发布中来源信息经常丢失、单靠技术方案不足以支撑永久 2025）。

[^7]:二点一一亿行代码变更的读数（重复代码块八倍增长、重构由百分之二十五降至不足百分之十、两周内改写占比由百分之五点五升至百分之七点九），见 GitClear 关于人工智能对代码质量与技术债影响的年度分析。

[^8]:资深开发者在熟悉代码库上的随机对照试验（人工智能辅助组慢约五分之一而自认更快），见二〇二五年该项试验的公开报告。

[^9]:"验证缺口"的调查读数（一千一百余名开发者），见 Sonar 二〇二六年度开发者调查。

[^10]:人工智能使用度提高四分之一，评审加快、文档改善而交付稳定性下降百分之七点二，见该年度软件交付研究报告。

[^11]:人工智能代码评审在已知漏洞样本上频繁漏报注入类、跨站脚本类与不安全反序列化，而更常报出低危项，见二〇二五年相关评测。

认领物

[^12]:把人工智能形容为在来源、鉴定、科学分析这张三脚凳上加的第四条腿，见艺术品鉴证机构的公开说明。

[^13]:对算法归属的保留——担心的不是判错，而是算法判断这一前提会把艺术创作的复杂性压平——见 BendorGrosvenor 于二〇二五年十一月在《金融时报》的评论，转引自注三所引文章。

■ 附节·换一组学科再问一遍：昂贵的信号与不可否认性

本章的论证由经济学、艺术学、工程学三家撞出。本书十章里有九章用的是这同一组三学科，这在书一级上是一个明确的短板：书里的纠缠是同一次碰撞重复了十遍。本节与第七章、第八章的同名附节一样，是对这个短板的处置——换一组与前一组毫无亲缘的学科，把同一个问题再问一遍，看承重判断是被确认、被收窄，还是被推翻。第十章本身就是这样做的：同一个问题换一组理论来源再问一次。这里把那个动作用在本章上。

换上的三家是：证据法（举证责任的分配）、演化生物学（代价信号与累赘原理）、密码学（数字签名与不可否认性）。选它们不是因为它们说得对，而是因为三家各自都已经把"归属"这件事处理了几十年到几百年，而且三家给出的答案互不相容。

证据法的位置。归属在这里是一个分配问题：一件事该由谁来证明。举证责任的分配不由事实本身决定，而由制度对错判代价的权衡决定；而且分配一旦确定，成本就整个落在被分配的那一方。它对本章的回答是：认领物的成本之所以落在生产者身上，不是因为技术，是因为举证责任被默认分配给了他。

演化生物学的位置。归属在这里是一个可信性问题：一个信号要能被相信，它必须对发出者昂贵到伪造不划算。累赘原理说的正是这件事——信号的成本不是浪费，成本就是可信性的来源。它对本章的回答与证据法相反：认领物的成本不应当被压低，压低它等于取消归属证据的可信性。

密码学的位置。归属在这里是一个形式问题：不可否认性可以由数字签名以近乎为零的边际成本提供，一次密钥基础设施的固定投入之后，每一件产出的归属证明几乎不再需要追加任何工作。它对本章的回答是：认领物的成本是可以被工程消灭的，而且已经被消灭过一次。

三条不能同真。证据法说成本由分配决定，生物学说成本必须存在，密码学说成本可以归零。三家共同假定了一件没有说出的事——归属证明的成本，与它所证明的那件事是同一种东西：证据法假定证明的代价与被证明的行为可比，生物学假定信号的代价与被信号的品质同源，密码学假定签名所证明的正是被签名的内容。

推翻它的材料来自密码学自己。数字签名以近乎为零的成本给出了不可否认性，而它证明的只有一件事：这份东西出自这把密钥。它不证明这把密钥的持有者是否理解它、是否有能力生成它、是否在生成它时作过任何判断。签名把"谁交出的"做成了零成本，同时把"谁做的"这个问题完整地留在了原地——并且因为签名看上去已经解决了归属，这个问题从此更不容易被问出来。

由此得到本节的结论，它对本章正文构成一处真收窄：

本章正文把认领占比上升判为病象。接上这三家之后，必须分成两种。其一是可被工程消灭的那一部分——纯粹为证明"这份东西出自谁手"而追加的工作。密码学已经证明它的边际成本可以趋近于零，因此它上升就是纯粹的浪费，压低它没有任何代价。其二是不可被工程消灭的那一部分——为证明"这份东西是由一个具备相应判断的人做出的"而追加的工作。按累赘原理，这一部分的成本恰恰是它可信的原因；把它压到零，得到的不是省下的工时，而是一份不再能证明任何事的归属记录。

于是认领占比要拆成两项：签名式认领与代价式认领。前者的正确目标值是零，后者的正确目标值不是零。本章正文没有作这个区分，因而它的处方——降低认领占比——在后一部分上是错的。

这一拆分给出一条新的、可查的读数与一条判决性预测。读数是签名占比：一次交付的认领工作量中，可由身份基础设施在边际成本近零处替代的那一部分所占的比例。预测是：在已经部署了完备来源链与强身份绑定的通道里，认领占比不会归零，剩下的那一部分就是代价式认领；若观测到认领占比在这类通道里确实归零，则累赘原理在此不适用，本节的拆分不成立，本章正文的原主张恢复全部效力。

最后交代这次换组对本书 E 维的实际贡献。它不是给本章加了三个新的参考文献，而是改变了本章的处方方向：从"降低认领占比"改为"把签名式的降到零，把代价式的留住"。一次换组如果不改变任何处方，那它就只是装饰；这一次改了，因此它值得留在书里。

第五章

认不出来，就得自己交代——辨认负担的整体转嫁

■ 摘要

人工智能把生成一份看上去合格的产出的边际成本压到接近于零，与此同时，评价一方分辨这份产出的能力没有同步提高。本章考察由此产生的一类现象，它在中文里通常被叫作内卷。三门学科对它给出了三条互不相容的解释：经济学说决定性的是稀缺搬到了哪里，艺术学说决定性的是制作路径被重写，工程学说决定性的是接口没有跟着能力重新切开。三条都各自成立，却指定了三个互不相同的决定项，因而不能同真。本章指出三家共有一个从未被单独说出的前提——辨认一份产出这件工作，有一个由该领域性质给定的承担者；而推翻它的材料来自工程学自己：逆康威操作是这一门每天在做的事，它证明“谁先动”不是给定的，是可以被调换的。据此本章提出：这一类困境不是要素固定下的劳动内耗，不是为分离而追加的信号支出，也不是指标被博弈后的失真，而是辨认负担从评价方向被评价方的整体转嫁；它产生一类本章称为同形劳动的工作——为使一份产出可被辨认而追加、与真实产出同形、无法被分离计量，因而在三门学科的账本上都不设字段。本章给出二维辨别表、一句不含判断词的问话、一个三领域量纲不同而比率相同的读数，在十个领域兑现，其中两处反过来迫使命题收窄，并给出四条可证伪条件，其中一条用多数机构已有的时间戳记录一周内即可执行。本章另与本书同日同取这三个学科的四篇逐条分清缺口位置（生产者一侧／关口一侧／未阅一侧／参照一侧／本章的承担者一侧），并据五处的合观提出一条更强也更易被推翻的判断：五处不是五种病，是同一笔账在五个位置上的分录缺失，因此单独补上任何一处，负担不会减少，只会移到另外四处之一——这条判断连同其自我推翻的条件一并写在第九节。

■ 关键词

同形劳动；辨认转嫁；辨认比；分辨率差；不入账的工作

Abstract

认不出来，就得自己交代

supplied.gap;unbookedwork

■ 一、导论：三条不该同时为真的处方

三门学科，对同一件事给出了三条处方。三条各自都有厚实的支撑，而它们指定的着手处互不相同，因此不能同时是决定性的。

第一条来自经济学。

过去十年最有影响的一支劳动经济学工作主张：技术对劳动的作用要按任务来读，

而不是按职业读。一项技术把某些任务自动化掉，同时可能创造出新的任务；净结果取决于替代效应与新任务创造效应的相对大小。阿西莫格鲁与雷斯特雷波在这条线上给出的关键判断是，过去四十年美国工资结构的变化，主要不是由技能偏向型技术进步解释的，而是由任务的位移解释的；他们进一步指出，有一类“平庸的自动化”——生产率提升微弱却足以替代劳动——对分配的破坏最大。奥托等人对一九四〇年以来新职业内容的普查，则把这条线推到了另一半：新工作确实在不断出现，但它出现在哪里、由谁获得，取决于新的稀缺落在何处。这条线索的处方因此是清楚的：要理解眼下的困境，就去看稀缺被搬到了哪里——算力、数据、许可、注意力。谁掌握新的稀缺，谁定价；其余一切都是跟随。

第二条来自艺术学。生成模型出现之后，艺术研究界最迫切的问题不是“机器能不能创作”，而是“当生成不再稀缺，价值判定还能落在哪里”。爱泼斯坦、赫兹曼与一批研究者在《科学》上的那篇综述给出的判断是：生成式人工智能改变的不是作品，而是作者与工具之间的分工位置——创作的努力从执行移向了指定、筛选与策展。赫兹曼更早的一篇文章已经点明，绘画摄影术之争的历史教训是，技术不消灭艺术，它把艺术的制作路径推到别处。这条线索的处方是：要理解眼下的困境，就去看制作路径怎样被重写——当成品不再能承载价值，价值必须重新落到过程、来源、不可复现的劳动上，于是从业者改变工作方式：宣示过程、留痕、限量、现场。

第三条来自工程学。

软件工程有一条流传六十年的观察：一个系统的结构，反映的是造它的那个组织的

沟通结构。麦考马克、鲍德温与鲁斯纳克用一批开源与闭源系统的实证把它做成了可检验的镜像假说。而机器学习系统进入工程实践之后，斯卡利等人那篇被反复引用的短文给出了新的读数：机器学习系统的技

认不出来，就得自己交代

术债不在代码里，在边界处——牵一发动全身的耦合、隐藏的反馈回路、未声明的消费者，全都是接口没切干净的后果。这条线索的处方是：要理解眼下的困境，就去看接口在哪里

被切开——能力条件变了而接口不变，所有人就会挤在同一个接口上堆动作；重新切开接口，同样的稀缺、同样的路径，压强可以差一个数量级。

三条处方互相取消。

若经济学是对的，艺术学所说的"宣示过程与来源"就不是出路，而是在新的稀缺处争夺租金的一种形态——它本身就是困境的一部分，而不是解药。若艺术学是对的，经济学的要素框架就系统性地漏读了价值的来源：它只读成品与投入，读不出路径，而价值恰恰已经移到了路径上。若工程学是对的，前两家指定的决定项都不是决定性的——同一份稀缺、同一条路径，接口切法不同，困境的深浅完全不同；那么"稀缺搬到哪里"与"路径怎样重写"就都只是跟随项。

反过来也一样。工程学每天在做的那件事——把工作切成可独立完成、可计量、可并行的模块——正是艺术学诊断为病根的那件事：切成模块，就是把不可复现的过程碾成可复制的成品，价值就此流失。经济学每天在做的那件事——为一切生产要素定价——正是工程学诊断为病根的那件事：一旦把开发者的产出定价，接口就会被扭曲以适配计量，系统形态劣化。而艺术学每天在做的那件事——宣示不可替代的作者性——正是经济学诊断为病根的那件事：那是在制造人为稀缺、抬高租金。

三门学科各自的日课，是另一门学科诊断出的病。这不是一个可以靠"综合三家之长"化解的僵局，因为三家并不是从三个角度看同一样东西，而是各自指定了一个不同的东西作为决定项。要走出去，只能另造一个框架。

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。撤销它之后，三条处方各自的适用条件才第一次可以被写出来，而"出路"这个问题也才第一次有一个不属于任何一家的答案。

■ 二、三家的立场与它们各自的时序读数

本节把三家的主张写实到可检验的程度。写实的标准有三条：这一家必须主张只有自己那条是决定性的；必须给出时序读数——先动的是什么，另外两样多久跟上；必须自己交代一处它撑不住的地方。第三条不是谦辞，它在第四节要被用上。

2.1 经济学：稀缺搬家决定一切主张。

已经形成的生产能力过剩，与既有的任务分配方式长期不容，逼得稀缺的所在地整个搬家；搬到

哪里，那里就重新定价，其余一切跟随。

按这条读法，眼下这一轮的实情是：生成能力过剩了——写一份合格的文案、一段能跑的代码、一张能用的图，边际成本趋近于零；而岗位、学位、职称、名额这套任务分法几十年没动。二者不容，于是稀缺从“能不能做出来”搬到了别处：搬到算力与数据，搬到发布与分发的许可，搬到审核与背书，搬到读者的

认不出来，就得自己交代

注意力。困境的体感——做得更多而位置不动——在这条读法里是标准的：你增加的是已经不稀缺的那一样。

时序读数。

稀缺位置先动，任务分法与个体行为随后跟上，滞后以年计。从一项通用技术进入生产到职

业内容改写，历史上的滞后期通常在五到二十年之间。

它自己交代的一处。

这一家很清楚，有时候稀缺搬到哪里是被行为决定的，而不是相反：注意力之所以

成为稀缺品，正是因为所有人都去争它。西蒙一九七一年那句被引用了半个世纪的话说的就是这件事——信息的丰富造成注意力的贫乏。在这一处，经济学自己承认它指定的驱动项反过来是被驱动的。

2.2 艺术学：制作路径的重写决定一切主张。

已经定型的作品形态，与它赖以成立的条件场长期不容，逼得制作的路径被重写；路径怎么重

写，价值就落到哪里，其余一切跟随。

按这条读法，眼下这一轮的实情是：作品的外观与它此前所依赖的稀缺性彻底脱钩了。一张图不再证明有人画过它，一段文字不再证明有人想过它。价值于是不可能再落在成品上，只能落在制作的路径上：谁在什么条件下、经过哪些不可跳过的步骤做出了它。从业者随之改变工作方式——公布过程录像、保留草稿链、限量、现场、驻留、把创作变成可见的事件。摄影术之后绘画所走的那条路，正在被重走一遍。

时序读数。

条件场与作品形态的不容先出现，路径的改动跟上，滞后以月到年计——艺术界的反应速度

远快于劳动力市场。

它自己交代的一处。

这一家同样清楚，有时候路径的改动先于矛盾：先锋的做法总是先改工作方式，然

后才显出与既有条件场不容。杜尚把一件现成品送进展厅，是先改路径，形态与条件场随后被迫追认。在这一处，艺术学自己承认它指定的被驱动项反过来是发起者。

2.3 工程学：接口的切法决定一切主张。

既有的设计路径与新的能力条件不容，逼得系统的实际形态改变；形态长成什么样，取决于接口

在哪里被切开，其余一切跟随。

按这条读法，眼下这一轮的实情是：模型能力这一年翻了几番，而团队的接口——谁交付什么、在哪里对接、按什么验收——基本没动。能力涌进一套没有为它设计的接口，结果不是产出翻番，而是相互牵连：所有人在同一个接口上堆动作，评审排队、返工、集成地狱。斯卡利等人所说的机器学习技术债，几乎每一条都是接口问题——耦合使得改一处要重训全部，隐藏的反馈回路使得两个子系统通过外部世界互相影响而没有任何接口声明它。

时序读数。

接口先动，形态在数月内跟上——这是三家里滞后最短的一条，因为接口是可以由人在一次

决策里改动的。

它自己交代的一处。

认不出来，就得自己交代

这一家最清楚不过：有时候形态先变而接口被迫追认。绕过流程的实践、影子系

统、私下的对接方式先长出来，接口事后被改来适配它们。而更要紧的是，这一门发展出了一个专门的操作来利用这一点——下一节要用的正是它。

2.4 中文语境里的既有讨论，与它为什么不够用“内卷”一词在中文公共讨论里已经被使用了将近十年，指的大致是：投入不断增加而总体产出不增加、每个人都被迫加码而没有人因此得利。围绕它的解释大体分三路：一路归因于位置的稀缺（好岗位、好学校、好资源就那么多），一路归因于评价的单一（只有一把尺子，所有人被迫在同一条轴上竞争），一路归因于文化与心态（比较心、不敢退出）。

这三路解释各有实证支撑，但它们共享一个特征：都把困境读成“竞争的强度问题”。位置少所以竞争激烈，尺子单一所以竞争集中，心态使人不肯退出竞争。按这个读法，出路自然落在减轻竞争上——增加位置、增加评价维度、鼓励多元选择。

本章要指出的是，这一读法在人工智能这一轮里失效了，而且失效的方式很具体：它无法解释为什么在位置增加、评价维度增加的场合，投入仍然继续上升。

一家扩招的机构、一个新增了多条晋升通道的企

业、一个把评价从单一论文改成多元成果的学科，都出现过投入不降反升的情形。按竞争强度的读法，这是反常；按本章的读法，这是正常——因为增加位置与增加维度都增加了待辨认量，而辨认能力没有同步增加。

这一节因此也界定了本章的贡献位置：本章不与上述三路解释争夺同一个变量，它指认的是一个此前不在任何一路里的变量。

■ 三、三条驱动不能同真

把三家的主张排在一起，冲突不在

而在被驱动项：这一家说什么与什么不容

说什么被逼着改变时序

经济学已形成的产能过剩×既有任务分法稀缺的所在地

年

艺术学已定型的作品形态×生成条件场

制作的路径

月至年

工程学既有接口路径×新的能力条件

系统的实际形态

月

三个被驱动项互不相同，而三家各自主张自己那一个是决定性的。这不是三个角度，是三条互相取消的因果指认。

冲突可以逐对写死：经济学×艺术学。若稀缺搬家是决定性的，那么艺术家改变工作方式就只是在新的稀缺处争租——他们宣示过程、限量、做现场，都是在把自己搬到尚未被自动化的那一格里去，这是跟随不是发起。若路径重

认不出来，就得自己交代

写是决定性的，那么经济学的要素框架就读不出价值：它把"过程"计入投入，而价值恰恰长在投入这一侧，于是它在自己的账上把答案记成了成本。

经济学×工程学。若稀缺位置是决定性的，那么接口切法只是在给定稀缺下的效率调整。若接口是决定性的，那么同一份稀缺下，形态可以差一个数量级——稀缺位置就不是决定项，只是背景。这两条不能并存：一个变量不能既是决定项又是背景。

艺术学×工程学。艺术学说路径必须保持不可模块化，因为价值恰恰在过程的不可分割处；工程学说路径必须模块化，否则规模化无从谈起。这一对是正面顶撞，且它顶在同一根轴上——要不要把工作切开。

三条不能同真，而谁都没说错。说错的是别的东西。

■ 四、三家共有的那个前提，与推翻它的那件材料

4.1 它们共同假定了什么三家争的是：这一轮的改变，该由哪一处的变化来解释——稀缺的位置、制作的路径，还是接口的切法。

而三家共同假定了一件事：在任何一轮里，总有一处是发起处、另两处是跟随处，且这个发起处由该系统的性质给定。

经济学假定稀缺位置总是先动，艺术学假定条件场与形态的不容总是先发生，工程学假定接口总是设计者可以先动的那一处。

这条前提念给三家听，三家都会说这还用说吗。会引起争论的是"哪一处先动"，不是"没有一处是先动的"。这正是共有前提的特征。

再往下一层，这条前提在本章所处理的具体问题上还有一个更要紧的形态。三家争的其实是同一份工作的归属：判断一份产出是不是合格的、是不是这个人做的——这份工作在三家的框架里都有位置，但都被当成给定的。经济学把它当作交易成本，一项摩擦；艺术学把它当作鉴赏与鉴定，一种能力；工程学把它当作验证，一道可以自动化的工序。三种处理都正确，而三种处理有一个共同的后果："谁承担辨认"不是一个变量。它是摩擦的大小、能力的高低、自动化的程度，唯独不是归属的问题。

4.2 推翻它的那件材料，来自工程学自己推翻这条前提的材料，不能从外面搬——从外面搬来的是第四家的立场，那只是加入争论，没有取消争论。它必须是三家之一自己已经掌握、已经发表、只是没往这个方向用的事实。

这件材料在工程学手里，而且是这一门每天都在用的常识：逆康威操作。

康威一九六八年的观察是：系统的结构反映造它的组织的沟通结构。这条观察被当作一条描述性的规律用了几十年。而工程实践在某个时刻把它反过来用了：既然形态由组织结构决定，那就先改组织结构，以得到想要的形态。要一个模块化的系统，就先把团队按预期的模块边界拆开；要一个平台与应用分离的架

认不出来，就得自己交代

构，就先把平台团队与应用团队分开考核。这个做法在软件工程界有名字、有文献、有大量实践，它不是异端，它是标准动作。

这件材料在它自己的领域里回答的是另一个问题——"怎样得到想要的架构"。没有人把它拿来当作关于发起处的证据。而它恰恰是：>发起处不是给定的。它可以被调换，而且工程实践每天在调换它。

镜像假说说组织结构决定系统形态；逆康威操作说你可以反过来先动系统形态的设计意图，去改组织结构。同一对变量，谁驱动谁，取决于这一轮谁先被动。而"这一轮谁先被动"是可以由人安排的。

一旦这条前提被撤销，三家的争论就不再是一个需要裁判的争论。三条驱动都对，只是各自描述了发起处落在不同位置的那些轮次。争论之所以无解，是因为三家各自把自己观察到的那些轮次当成了全部。

而更要紧的后果在第二层前提上：如果发起处可以被调换，那么"谁承担辨认"也可以被调换。

它不是摩

擦的大小，不是能力的高低，不是自动化的程度——它是一个归属，而归属是可以移动的。

■ 五、辨认转嫁与同形劳动

5.1 命名现在可以说出本章的判断了。

一份产出要进入任何一个评价体系，都必须先被辨认：它算不算合格，是不是这个人做的，值多少。这份工作有成本，而且这个成本从来没有消失过。在传统的情形里，它由评价一方承担——招聘方筛选，编辑审读，鉴定家看笔触，测试工程师跑用例。被评价的一方只需要把东西做出来。

人工智能改变的不是"东西能不能做出来"，而是做出来的东西彼此还分不分得开。当生成一份看上去合格的产出的边际成本趋近于零，评价方原有的辨认手段——文笔、完成度、格式规整、术语正确——同时失效，因为这些正是生成模型最先学会的部分。评价方的辨认能力没有下降，是它面对的待辨认量暴涨、而可用的分辨特征同时贬值。二者之差，本章称为分辨率差。

分辨率差不会让评价停下来。评价必须继续进行——岗位要招人，论文要录用，作品要收藏，代码要合并。于是发生的不是评价的崩溃，而是一次归属的移动：>辨认这份工作，从评价方转移到了被评价方。

具体形态是：评价方不再自己去分辨，而是要求被评价方自己交代——提供过程记录，提供原始文件，提供思路说明，提供不可伪造的痕迹，提供更多的、更早的、更细的证据。被评价方为此追加的工作，本章命名为同形劳动。

同形劳动有三条定义性质，缺一条就不是它：

认不出来，就得自己交代

其一，它与真实产出同形。

写一份说明文档、录一段过程、把一次改动拆成三个可审阅的提交、为一幅画留存分层文件——这些动作在任何账本上都与"工作"无法区分。它们不是加班，不是应酬，不是无谓的行政；它们看起来就是本职工作的一部分，而且往往确实有一点独立价值。

其二，它不可被分离计量。

正因为同形，没有任何现成的记录能把它从真实产出里切出来。工时表上它是开发工时，科研管理系统里它是科研时间，画室里它是创作时间。

其三，它随分辨率差单调增加，而与努力程度、能力高低、供需比例都不直接相关。

这一条是它与所有

既有说法分家的地方，第八节要逐条兑现。

于是本章的承重判断是：>这一类困境不是要素固定条件下的劳动内耗，不是为分离而追加的信号支出，也不是指标被博弈之后的失真，而是辨认负担从评价方向被评价方的整体转嫁——它产生一类与真实产出同形、无法被分离计量、因而在任何一家的账本上都不设字段的工作。

三个被否定的项各对应一家的划界标准，第八节逐一交手。

5.2 二维辨别表只有一个名字不构成一个可用的区分。把它升成两根轴：产出的边际成本高（做一份合格的东产出的边际成本低（做一份看上去合格的东西西很贵）

近乎免费）

辨认能力强（评价甲格·常态

同形劳动占比低而稳乙格·增产不增负

产出暴涨而辨认跟得上。

方不追加工作即可定。前工业时代的手工业、有可靠计可自动检验的编程任务、有裁判器的竞赛、可分辨)

量的竞技体育。

辨认能力弱 (分辨丙格·慢性积压

复现性强制的实验科学。

辨认困难但因产量

需要评价方追加有限而不爆发。艺术品真伪鉴定、稀量工作)

缺文物的来源研究。

丁格·靶格

同形劳动占比单调上升。当下大

多数以文本、图像、代码为产出的评价体系。

四格里要打的是丁格。乙格是最重要的对照：它证明产出成本的下降本身不产生困境——只要辨认跟得上，产出暴涨反而是纯收益。

5.3 靶格的三条签名靶格必须是全部形式审查都能通过的那一格，否则不需要一个新框架就看得见。丁格的三条签名：其一，短期指标表现为改善。

产出数量上升，交付变快，投稿量翻番，作品变多，材料变厚，说明变详尽。所有可见指标都好看，而且好看得有理有据。

其二，失效不产生任何信号。

同形劳动与真实产出同形，因此在账上没有任何一栏会变红。没有人加班超标——因为那些时间就是正常工时；没有人产出下降——因为材料确实更完整了。

其三，它自我加固。

认不出来，就得自己交代

分辨率差越大，评价方越倾向于要求更多材料；材料越多，待辨认量越大，分辨率差越大。而每一次加固都以"提高严谨性"的形式出现：增加评审环节、要求提供原始数据、引入查重与检测、要求作者说明贡献。每一条单看都对。

三条签名同时成立，这个靶格才算立住。

5.4 一句不含判断词的问话判断需要一句能拿去问流程文件、日志与考核办法的问话，而不是拿去问人的感受。本章给出的是：>在这个评价体系里，被评价方为使一份产出可被辨认所花的时间，与评价方为辨认它所花的时间，比值是多少？

这句话里没有"应当"、没有"真正"、没有"充分"。它问的是两个可以从时间戳里数出来的量。

把它在五个具体场景里跑一遍：1.学术同行评审。作者准备投稿材料、撰写回复信、补做审稿人要求的实验所花的时间，除以审稿人阅读与撰写意见的时间。通常远大于一。2.竞技田径。运动员为"证明自己跑得快"所追加的时间，除以计时器的工作时间。接近于零——存在一个客观的辨认装置。3.开源项目的代码合并。贡献者撰写变更说明、拆分提交、补测试的时间，除以维护者审读的时间。中等，且随项目规模上升。

卖。

4.

传统绘画的拍

艺术家为使作品可被辨认所追加的时间，接近于零——作者性由物质证据承担；而鉴定方投入的时

间很高。比值小于一。5.面向审核的内容平台。创作者为标注来源、证明原创、申诉误判所花的时间，除以自动化审核系统的处理时间。极高。

五个答案互不相同。第四个是查出来的而不是想出来的，而且它与预期相反——传统艺术市场的辨认负担一直压在评价方，且这套制度运转了几百年而未崩溃。这一处后面要被用来给命题加约束。

5.5 关于"内卷"这个词中文里称呼这类困境的"内卷"借自格尔茨一九六三年对爪哇水稻农业的描述，而他的机制以要素固定为条件——这一轮的定义性特征恰恰是产出能力的暴涨，因此借用在前件上就不成立。这一笔账已由本书第四章第四章《认领物》第十七节清过，此处不重复，只记下本章所需的那一句：现象是同一个，机制不是同一个。

本章因此保留"内卷"作为对现象的通称，而不沿用它的机制；第八节第一条把这一区分写成了可判

定的形式。

■ 六、辨认比：一个三领域量纲不同而比率相同的读数

把5.4那句问话固定成一个可清点的量： $\text{>辨认比} = \text{被评价方为使产出可被辨认所投入的时间} \div \text{评价方为辨认它所投入的时间}$ 。

认不出来，就得自己交代

三个领域各自兑现，量纲不同而比率的意义相同：·就业与人事：求职者准备简历、作品集、笔试与多轮面试的总时间，除以招聘方筛选与面试的总时间。数据在招聘系统的时间戳里。

·软件工程：贡献者撰写说明、拆分提交、准备演示、补写周报所花的时间，除以评审者阅读与批准的时间。数据在版本控制与评审系统的时间戳里。

·艺术与创作：创作者为宣示来源与过程所做的记录、限量、现场与说明所花的时间，除以收藏方、策展方或平台审核所花的时间。数据在工作日志与平台审核记录里。

这个读数要能用，必须过四条性质：无量纲。时间除以时间，三个领域可以并排比较。

可事后清点。三处的数据都在既有记录里，不需要新建采集系统。

把一个印象变成一个数。"现在什么都要交代"是一句体感，辨认比是一个可以争论的数。

与既有主要指标正交。

这一条最容易漏，也最要命——一个与既有指标高度相关的量，会被立刻读成那

个指标的代理，白造。检验的办法是举出一个"既有指标最好看而这个数最难看"的真实例子：>一个交付速度最快、缺陷密度最低、按任何现行工程指标都堪称优秀的团队，可以有极高的辨认比——因为它把大量时间花在让工作可被看见上（详尽的设计文档、精心拆分的提交、面向评审的演示），而这些时间恰恰计入了"开发工时"，因此在所有既有指标上都表现为投入充足、过程规范。

反过来的例子同样存在：一个辨认比接近于零的团队可能产出很差——没有人看它，也不需要它被看见。

辨认比与质量无关，与被评价的强度有关。这一条使它不能被读成任何现行指标的代理。

6.1 这个读数与第四章所用读数的关系本书第四章第四章《认领物》第十七节已经给出过这个比值，并用一个职称评审的算例走完了它（六十小时比七点五小时；压缩材料后降到四；引入自动核验而不撤旧要求则升到二十三）。此处不重复那次计算，只说明它在本章框架里的位置有何不同。

在那一章里，这个比值是对"认领占比"这一单侧读数的补充——它回答的是负担压在谁身上。在本章里，它是承重命题的直接读数：本章主张的正是这份工作的归属发生了移动，而移动只能由两侧的投入之比读出，任何单侧的量都读不出移动。

由此也可以看出两篇的一处实质分歧。按那一章，认领物是被生产出来的一层东西，比值上升意味着这层东西变厚；按本章，什么都没有变厚——同一份工作换了个人做。两种读法在一处会分岔：若评价方的投入在认领占比上升的同时绝对下降，则本章对（工作被推了出

去)；若两侧同升，则那一章对（确实多出了一层）。这个分岔可以直接从时间戳里查出来，不需要任何新的采集。

认不出来，就得自己交代

■ 七、十处兑现

判断要成立，必须在具体领域里预测出具体的事，而不只是"像"。以下十处，其中第九、第十两处反过来给命题加了约束。

一·学术出版。预测：在引入生成式工具之后，投稿量上升与录用率下降会同时发生，而作者端的单篇准备时间不降反升——因为写作本身变便宜了，作者会把省下的时间投入到使稿件可被辨认的部分（更详尽的方法描述、更多的补充材料、可复现性声明、贡献者声明）。可读数：单篇附件数量与补充材料页数的年度变化。数据源：期刊投稿系统的附件记录与修回轮次时间戳。反向对照：在实行注册报告制的期刊（研究设计在数据收集前即被评审通过）里，这一上升应当显著更弱——因为辨认的锚点被前移到了设计阶段。

二·校园招聘。预测：简历筛选自动化程度越高的岗位，求职者端的准备时间越长，且这一关系在控制岗位竞争比之后仍然成立。机制是：自动筛选把可辨认的特征窄化到少数几项，求职者被迫针对这几项追加投入。数据源：招聘系统里从投递到初筛的时间戳，与求职者端的投递次数、简历版本数。反向对照：设有标准化笔试或实操考核的岗位（存在客观裁判器），求职者端的准备时间应当集中而不发散。

三·开源维护。预测：贡献指南越详细的项目，首次贡献者的提交到合并周期越长，而不是越短。详细的指南是评价方把辨认工作前置转嫁的一种形式。数据源：项目的贡献指南字数、首次贡献者的提交到合并周期、被要求修改的轮次。反向对照：有完整自动化检查（格式、测试、覆盖率门槛）的项目，这一周期应当与指南详细程度脱钩。

四·平台创作。预测：原创性检测越严格的平台，创作者花在申诉与自证上的时间占比越高，且这部分时间与作品质量无关。数据源：平台的申诉工单数与处理时长、创作者的发布准备时长。这一处最容易观察到本章所说的同形：申诉材料在平台记录里与创作记录同类，无法分开统计。

五·中小学教育。预测：在允许使用生成工具的课堂里，评价方式若不变，则学生的作业完成时间下降而过程性材料（草稿、思路说明、口头答辩）的要求上升；两项相加，学生总投入时间不降。这一处的政策含义最直接：若不同时改变评价方式，允许使用生成工具并不减轻学业负担，只是把负担从做出来移到证明是我做的。

六·医疗记录。预测：语音转写与自动生成病历普及后，医生的文书时间不会按转写效率的比例下降，因为节省下来的时间被"确认与修正生成内容"占用——这是辨认工作从系统转嫁给临床医生的一种形态。

数据源：电子病历系统的编辑时长与修改次数。这一处的伦理约束最强——第十三节第二条正是为这类场合写的。

七·法律文书。预测：使用生成工具起草的一方，会被要求提供更多的核验痕迹（引用核对记录、案件真实性声明）；法院一旦引入此类要求，律师端的核验时间将成为新的主要成本项。

八·工程招投标。预测：技术方案可被快速生成之后，招标方会转向要求投标方提供更难伪造的证据（既往项目的可核查记录、现场答辩、关键人员到场承诺），投标方的准备成本上升而中标率分布不变。

认不出来，就得自己交代

九·传统艺术市场（反过来加约束）。这一处不支持命题的普遍形态，反而迫使它收窄。传统绘画市场的辨认负担几百年来一直在评价方，而且没有转嫁的迹象。原因是：产出本身携带着足以辨认的物质痕迹——画布、颜料层、笔触、老化。只要物质证据能承担辨认，转嫁就不发生。

于是本章原本会写下的那句更强的话——"辨认能力的相对下降必然导致负担转嫁"——必须削掉。真实的命题是：只有当产出不携带足以辨认的痕迹时，转嫁才发生。

命题的适用范围因此从"所有评价体系"缩

到"产出以可完美复制的形式存在的评价体系"。这一削减改变的是适用范围本身，不是补充一个条件。

十·自动化测试与形式验证（反过来加约束）。这一处削掉的是本章的处方。软件工程的经验证明，辨认能力可以被工程化地提高：类型系统、单元测试、性质测试、持续集成，每一项都是把一部分辨认工作变成机器可执行的。这说明分辨率差不是人工智能带来的不可逆后果。

那么为什么它没有被普遍投资？答案不在技术，在成本的归属：辨认能力的投资由评价方支付，而转嫁的成本由被评价方承担。对评价方而言，转嫁比投资便宜。

于是本章的处方必须改写：从"提高辨认能力"改成"改变辨认成本的归属"。这一削减动的是价值排序——本章原本主张的是一个技术方向，现在主张的是一个制度方向。

■ 八、与最近的几种说法的分界

这一判断最容易被读成几个已有的概念。逐条交手，每一条给出一个能把两边分开的观察。

其一·农业内卷化（格尔茨，一九六三，人类学）。"内卷"这个中文词就是从这里借来的。格尔茨描述的是爪哇的水稻农业：土地不可扩张、人口持续增加，于是劳动被不断投入到同一块土地上，单产微增而劳动强度剧升，制度日趋精细繁复。它说到哪一步：在要素固定的条件下，增加的投入只能带来递减的、内向的精细化。分离线：格尔茨的机制以要素固定为前提；而人工智能带来的恰恰是产出能力的暴涨——要素不但没固定，还在扩张。把这个词直接搬过来，是一次误诊。判定形式：若观察到困境的深浅与产出能力扩张的速度正相关（能力涨得越快、困境越深），则本判断对而内卷化错；若观察到困境只出现在产出能力停滞的部门，则反之。

其二·求职市场信号（斯彭斯，一九七三，经济学）。这是最危险的占位者。斯彭斯的模型里，教育是一种信号：它本身不提高生产率，但因为高能力者获得它的成本更低，它能把两类人分开。它说到哪一步：为被辨认而追加的、本身无生产性的支出，是理性均衡的产物。分离线：斯彭斯的分离均衡以"信号成本与能力负相关"为存在条件。而当生成一份看上去合格的产出对所有人的成本都趋近于零时，这个条件失效——分离均衡不再存在。斯彭斯的框架预测的是"信号变贵直到重新分离"；本判断预测的是分离不再恢复，评价方转而把辨认工作整体外包给被评价方，而这部分工作不进入任何一方的账。判定形式：若观察到在生成成本趋零的领域里，出现了新的、成本与能力负相关的可分离信号并使材料要求回落，则本判断错而信号理论对；若观察到材料要求持续上升而始终未形成新的分离信号，则反之。

认不出来，就得自己交代

其三·柠檬市场（阿克洛夫，一九七〇，经济学）。质量不可观察时，买方按平均质量出价，优质卖方退出，市场逐步瓦解。它说到哪一步：信息不对称会侵蚀市场。分离线：阿克洛夫预测的是逆向选择与市场萎缩；本判断预测的是市场不萎缩——它继续运转，规模甚至扩大，代价由被评价方以同形劳动的形式支付。判定形式：若观察到在辨认困难的领域里，高质量供给方系统性退出、交易量下降，则柠檬市场对；若观察到交易量不降而供给方的单位准备成本单调上升，则本判断对。这是一个可以直接去数的对照。

其四·寻租耗散（图洛克，一九八〇，公共选择）。为争夺一份租金而付出的总支出，可以耗尽租金本身。它说到哪一步：竞争性支出可以把收益全部抵消。分离线：租金耗散以"租金是一份固定的、可被争夺的量"为前提，且它预测支出总额由租金大小决定。本判断说的是：即便租金在增长，只要分辨率差在扩大，同形劳动仍然上升。判定形式：若观察到某领域的位置数

量与报酬同时增长而被评价方的准备成本仍上升，则本判断对而耗散论不足；若准备成本随位置增加而同比例回落，则反之。

其五·古德哈特定律与坎贝尔定律（一九七五／一九七九）。一项指标一旦成为目标，就不再是好指标。它说到哪一步：度量会被博弈，从而失去有效性。分离线：古德哈特的机制需要一个被博弈的指标和博弈的意图；本判断描述的过程里，可以没有任何人博弈——分辨率差本身就会产生追加工作，哪怕所有人都诚实。判定形式：若观察到更换指标或撤销指标之后，被评价方的准备成本显著回落，则古德哈特足够；若观察到换指标无用而唯有恢复辨认能力（引入可自动核验的判据、缩小待辨认范围）才使成本回落，则本判断对。

其六·目标置换（默顿，一九四〇，组织社会学）。手段吞噬目的：为遵守程序而遵守程序。这是本章命题剥掉学科名词之后最接近的那个形式结构——"为使自己被计量而做的工作，挤占了被计量的那件工作"。分离线：目标置换的机制在执行者的内在化上——规则被内化为目的本身。本判断不需要任何内在化：被评价方完全清楚这些材料是为被辨认而做的，也完全不认同它们的价值，仍然必须做。判定形式：若观察到被评价方对追加材料表现出规范性认同（认为它本身有价值），则目标置换对；若观察到普遍的、公开的不认同而行为不变，则本判断对。

其七·文凭社会（柯林斯，一九七九，教育社会学）。文凭的膨胀不反映技能需求的上升，而是群体间地位竞争的结果。分离线：柯林斯的机制是群体封闭与地位竞争，它预测膨胀沿着社会分层线展开。本判断的机制不涉及群体，它预测膨胀沿着辨认难度线展开——同一社会阶层内部，辨认容易的职业不膨胀，辨认困难的职业膨胀。判定形式：比较同一阶层内辨认难易不同的两组职业，若膨胀程度无差异而阶层间有差异，则柯林斯对；若阶层内按辨认难度出现显著分化，则本判断对。

其八·国家的可读性（斯科特，一九九八，政治人类学）。国家为便于治理而把社会简化为可读的形式，成本由被治理者承担。这一条极近。分离线：斯科特的可读性是被施加的——有一个施加者，有一套简化方案，有一个治理目的。本判断描述的转嫁没有施加者：它是分辨率差的自动后果，没有人下命令，评价方自己往往也是受害者。判定形式：若观察到撤销那套标准化方案之后负担缓解，则斯科特对；若观察到撤销方案之后负担以非正式形式（人情、圈子、私下背书）继续存在甚至上升，则本判断对。

认不出来，就得自己交代

其九·成功陷阱与伊卡洛斯悖论（列文塔尔与马奇，一九九三；米勒，一九九〇，组织学习）。把本章的命题剥到形式层，剩下的是"越成功越糟"这一类反转模板；而这类模板早有名字，必须点名。分离线：成功陷阱的机制是学习的近视——组织因过去的成功而过度利用、疏

于探索。本判断里没有任何主体因为成功而改变行为；分辨率差是外生的，它由生成成本的下降造成，与该组织过去成功与否无关。判定形式：若观察到困境只出现在有过成功历史的组织，则成功陷阱对；若观察到新成立、无成功包袱的组织同样出现且程度相当，则本判断对。

其十·胶水工作（雷利，二〇一九，工程实践）。这是方法学一侧最贴近的一位：一类维持系统运转所必需、却不出现在任何晋升材料里的工作。分离线：胶水工作的定义特征是不被计入；同形劳动的定义特征是无法被分离——它不是没被记，它是被记成了别的东西（正常工时、正常产出）。这两者的处置方向相反：胶水工作的对策是给它设一个字段；同形劳动即使设了字段也数不出来，因为它与真实产出同形。

判定形式：若在引入专门的记录字段之后，这类工作可以被稳定地区分并统计出来，则它是胶水工作；若引入字段之后填报者自己也无法可靠区分哪些时间属于它，则本判断对。

最近的邻居是哪一个。

是斯彭斯的信号理论。本判断仍不能被它吸收，理由只有一条：信号理论处理的

是分离仍然可能的世界，它的全部机制建立在成本差异之上；而本章处理的正是分离条件消失之后发生的事，那里没有可用的成本差异，因而也没有均衡可言。

再往下一层。逐个问这些邻居“它结构性地看不见什么”：斯彭斯把“信号可被分离”当作给定，因此看不见分离条件整体失效；阿克洛夫把“质量与价格可以对应”当作给定，因此看不见市场不萎缩而成本转移；图洛克把“租金是固定量”当作给定，因此看不见争夺成本可以由租金决定；古德哈特把“指标是被设计的”当作给定，因此看不见无需设计失误也会发生；格尔茨把“要素固定”当作给定，因此看不见要素扩张下的同类现象；斯科特把“有一个施加者”当作给定，因此看不见无人施加的转嫁。

这几行背后是不是同一样被当作给定的东西？是。所有人都把“谁承担辨认”当作由领域性质给定的。

它

可以是摩擦、是能力、是制度、是权力，唯独不是一个可以被移动的归属。这块地基是所有人——包括互相批判的各派——共同站着的，因此谁也没法回头看它。

■ 九、与本书同题诸篇的分界，以及五处缺口合起来是什么

本书在同一天里，用同样这三个学科、答同一个问题，出了不止一篇。逐篇指名并各给一个能把两边分开的观察，是本节的第一件事；而把这几篇摆在一起之后能看见的东西，是第二件，也是本章单独看不到的。

9.1 与第四章第四章《认领物》

那一章的缺口在生产者一侧：为使产出可被归属于某人，必须与它一并被生产的那一层东西，三家的账只给它成本位、不给产出位。本章的缺口在承担者一侧：核实这份产出的工作本身，从一侧移到了另一侧。

认不出来，就得自己交代

分界最干净的一处已在第六节第一小节写明：那一章预测账上多出一层，本章预测什么都没多出来、只是同一份工作换了人做。判定形式：若评价方的投入在认领占比上升的同时绝对下降，则本章对；若两侧同升，则那一章对。两篇的读数（认领占比与认领比）可以同时算，不必二选一。

9.2 与第三章第三章《这道关拦下过谁》

那一章的缺口在关口一侧：一道关的分辨力已经饱和而无人有权宣布它作废，因为门槛通常没有作者。本章不处理关口的权属，处理的是核实成本落在谁头上。

判定形式：取一道有明确作者、且该作者定期宣布作废与新建的关口——按那一章，它不该出现空关；按本章，只要区分能力低于待核实量，即便关口权属健全，认领比照样上升。这类关口在受监管行业的合规审查里可以找到，可直接查。

9.3 与第一章第一章《两种「不」》

那一章的缺口在未阅一侧：一次没有人打开过被否决之物的否决，与看过之后的否决在所有记录上同码。

这一章与本章的关系最紧——两者都指向“核实没有真的发生”。

分岔在那份工作去哪儿了。按那一章，它没有发生，被跳过了；按本章，它发生了，只是换了人做——被评价的一方替评价的一方做完了大部分。判定形式：在未阅率接近于零的体系里（每一份都确有人打开过），那一章的机制无从发作，而本章预测认领比照样可以很高。学位论文答辩、法院合议庭这类每一件都必被完整审阅的场合，正是分开两篇的现成样本。

9.4 与第二章第二章《被丢掉的那一眼》

那一章的缺口在参照一侧：未被接受的产出在无人记账处合成了一条判定基线，并单向抬升它。它要求那件东西被看过一眼；本章不要求。

分界在抬升的是水准还是转移的是工作。那一章预测合格线上移，本章预测合格线可以不动而认领比照样上升——因为负担的转移不需要标准变严。判定形式：在合格线经复核确认多年未变的评价体系里，若认领比仍单调上升，则本章的机制独立于暗基线；若合格线不动时认领比也不动，则本章可被那一章吸收。

9.5 与之四十《轮空》、之十九《类内距》

与《轮空》：那一章问发起权在三处之间的轮转，本章问一份特定工作的承担者被调换。判定形式：若换手率稳定而认领比持续上升，本章的机制独立；两者同涨同落则本章可被吸收。

与《类内距》：那一章主张可替换性由类内距决定，压缩它的功夫不会因无人付钱而消失。这一条与本章在人工智能这个场景上给出相反的预测。生成工具使产出的类内距急剧变小——机器写出的文本彼此高度相似。按那一章，类内距变小意味着适配功夫下降；按本章，正因为彼此高度相似，核实所依赖的差异消失了，认领比应当上升。同一份材料，两篇判断相反。判定形式：在产出趋同程度可测的领域里，若返工工作量随趋同而下降、且被评价方准备成本同步下降，则那一章对；若返工下降而准备成本上升，则本章对。

认不出来，就得自己交代

9.6 五处缺口摆在一起：它们不是五种病到这里可以做一件单独任何一篇都做不了的事：把五处缺口摆在一起看。

五篇各自命名了一样东西，五样东西的形状完全相同——都不是"记录得少"，而是在三家的账本里压根不算一样东西、不设字段。而五样东西所处的位置各不相同，且恰好连成一条完整的评价回路：篇

缺口在

那样东西

第四章生产者一侧为可被归属而一并生产的那一层第三章关口一侧

分辨力已饱和而无人作废的那道关

第一章未阅一侧

没有人打开过就作出的那次否决

第二章参照一侧

落判件在无人记账处合成的那条基线

本章

承担者一侧核实这份工作从哪一侧移到了哪一侧

一份产出从被做出来、经过关口、被人看或不看、与旧件比较、最后由某一方承担核实成本——五处缺口把这条回路走了一整圈，而每一处的共同点是同一个：那里发生的事情不产生任何记录，因为它不被当作一种事情。

由此可以提出一条比任何单篇都强、也更容易被推翻的判断：>

这五处不是五种病，是同一笔账在五个位置上的分录缺失。因此单独补上任何一处，负担不会减

少，只会移到另外四处之一，总量不降。

本章的机制正是这条判断的一个特例，也是它最容易被检验的那一处：核实的负担在两侧之间移动，是所有转移里最容易被时间戳逮住的一种。

这条判断的判定形式是清楚的，而且它同时可以推翻本章自己：取一个只补了其中一处缺口的体系——例如只给认领物设了产出位（第四章的第一条出路），而关口、阅读、参照、承担四处未动——若在两年内该体系的总负担（五处读数之和的可比化形式）确实下降，则本条判断为伪，五处就确实是五种可以各自治理的病。本章预测的是：认领占比会下降，而认领比、未阅率与暗基线抬升度中至少有一项相应上升。

为什么这条判断必须由第五篇来提，而前四篇提不了。

不是前四篇疏忽。每一篇都在自己的位置上把机

制做到了可检验的程度，而"五处缺口是同一笔账"这句话需要五处都被命名之后才成立——在只有一处或两处的时候，它是一个无法与"两种不相干的现象"区分开的猜测。这也是它现在可以被推翻的原因：五处都已经各有一个读数，五个读数可以同时算。

■ 十、三家为什么各自到不了这一步

三家停在原地，不是因为疏忽。是它们各自的核心问题恰好使这一步成为不必要的。

认不出来，就得自己交代

经济学到不了，因为它的核心问题是资源配置。在这个问题结构里，辨认必须被处理为一项交易成本——一种摩擦。而摩擦在这个框架里是一个可以有大小的量，不是一个可以有归属的东西。把辨认当作摩擦是正确的处理，而"谁承担"这个维度正是这个正确处理的影子：一旦某样东西被记为摩擦，它的分布就不再是变量。

艺术学到不了，因为它的核心问题是价值判定的正当性——什么算好，算谁的。在这个问题结构里，辨认必须被处理为鉴赏与鉴定，即一种能力，而不是一份工作量。能力不计，因而它从来不进入任何比率。

艺术学能极精细地讨论谁有资格辨认、辨认的标准是什么，唯独不问辨认花了多少小时、由谁付。

工程学到不了，因为它的核心问题是系统在约束下可靠运行。在这个问题结构里，辨认被工程化为验证，而验证的成功标志恰恰是"自动化到无人参与"。一旦自动化成功，那份工作就

从视野里消失了。工程学看不见"辨认无法自动化时那份工作去了哪里"，正是因为它太擅长把可自动化的那部分自动化掉。

三处盲区形状相同：各自的核心问题都要求把辨认转成自己框架里的某一类东西——摩擦、能力、工序——而这三次转换都恰好抹掉了归属这个维度。

■ 十一、出路：它落在轮次上，不落在处方上

问题问的是困境与出路。前十节回答了困境。出路不是一份行动清单——三家已经各给了一份，而且互相取消。出路落在这一轮的发起处在哪里。

三条驱动各自的完整形态都是三步，而不是两步：某两样不容，逼得第三样改变，改变之后反过来改动前两样，由此决定下一轮由谁先动。

· 经济学那一条：产能过剩与既有任务分法不容，逼得稀缺搬家；搬家之后反过来改动任务分法——新的稀缺重新定义了什么算一份工作，下一轮由任务分法先动。

· 艺术学那一条：作品形态与生成条件场不容，逼得制作路径重写；重写之后反过来改动什么算作品，下一轮由形态先动。

· 工程学那一条：接口路径与新能力不容，逼得系统形态改变；改变之后反过来改动接口——影子实践被迫认为正式接口，下一轮由接口先动。

把辨认放进这个循环里，出路的位置就出来了。眼下这一轮的发起处是产出能力：模型变强，生成成本塌陷。在这一轮里，无论被评价方怎样努力，增加的都是已经不稀缺的那一样；而评价方的每一次"加强审核"，都在把辨认负担进一步推下去。

因此可干预的位置只有一个：让下一轮的发起处换到辨认能力一侧。具体有三种形态，各自可读数：其一，引入可自动核验的判据。

把辨认的一部分交给机器可执行的检验——代码的测试与类型、数据的

可复现脚本、内容的来源凭证。这不是提高标准，是把辨认成本从被评价方搬回评价方（或搬给机器）。

读数：辨认比下降。

其二，缩小待辨认的范围。

认不出来，就得自己交代

限制被评价方可提交的材料总量，并配一条评价方的时间承诺。这一条反直

觉：减少材料等于减少证据，看上去降低了严谨性。但按本章的机制，材料总量正是分辨率差的产物而不是它的解药——材料越多，待辨认量越大，分辨率差越大。读数：单位申请的材料页数与评价方投入时间同时下降，而录用质量不变。

其三，把辨认成本内部化给评价方。

明确规定评价方为每份产出投入的最低时间，并公开这一投入。这

一条最难，因为它直接增加评价方的成本；而它也最有效，因为转嫁之所以发生，正是因为转嫁比投资便宜。

三条的共同形状是同一件事：不是让被评价方少做，而是让辨认这份工作重新回到有能力承担它的那一侧。

这就是为什么本章不给一份“如何在内卷中胜出”的建议——在丁格里，任何提高个人努力的建议都

只会增加同形劳动。

11.4 三家的处方各自在什么条件下有效撤销共有前提的另一个收益是：三条互相取消的处方，第一次可以被写出各自的适用条件。它们不再是三个竞争答案，而是发起处落在不同位置时的三条正确做法。

经济学那条处方——去看稀缺搬到了哪里——在发起处确实落在稀缺位置的那些轮次里有效。

判别的办

法是看时序：若一个领域里位置的定义先变（新职位、新资质、新的准入门槛先出现），而个体行为随后跟上，这一轮的发起处就在稀缺位置，此时往新稀缺处移动是理性的。它的失效条件同样清楚：当发起处在产出能力一侧时，往新稀缺处移动只是把同一场辨认危机搬到另一个地方重演，而且往往更快。

艺术学那条处方——重写制作路径、把价值落到过程上——在产出仍携带足以辨认的痕迹时有效。

现

场、限量、驻留、可核查的过程记录，这些做法之所以在艺术界奏效，是因为它们制造的是难以复制的物质与时间痕迹，即恢复了辨认的锚点。它的失效条件是：当这些痕迹本身也可以被廉价生成时（伪造的过程录像、批量生成的创作手记），路径重写就退化为又一层同形劳动。判别办法：问一句这份过程记录本身是否可被廉价伪造。

工程学那条处方——重新切开接口——在辨认可以被工程化时有效，而这恰恰是它最强的地方：把辨认的一部分变成机器可执行的检验，本身就是一次成功的接口重切。它的失效条件是：当待辨认的东西无法被形式化时（一份研究计划是否有想法，一件作品是否有价值），接口重切只能改变这份工作由谁做，不能减少它。

三条处方的共同适用条件因此可以写成一句：它们都要求辨认的锚点还在。

当锚点在，三条各按其发起

处生效；当锚点不在，三条都退化为同形劳动的不同形态。这解释了为什么三家会互相把对方的日课诊断为病理——它们各自在锚点尚存的场合里是对的，而在对方的场合里看到的正是退化后的形态。

认不出来，就得自己交代

■ 十二、可证伪条件

最强的竞争解释先写出来：其实不就是人多位置少吗？供给过剩导致筛选成本上升，材料军备竞赛是必然的。这是最朴素、最难反驳的一种说法，本章的证伪条件必须能排除它。

条件一（主条件，机制证伪）。>控制申请人数与位置数之比（供需比）之后，若"辨认比"与"被评价方单位准备时间"之间不存在剂量-反应关系，则本机制为伪——届时困境应归因于绝对的供需失衡，而不是辨认负担的归属。

样本设计（这一条低成本可实做，所需数据在多数机构的既有记录里已经存在）：取同一所大学内部的六个以上院系，同一评审周期、同一人事政策、同一申报系统。自变量是该院系是否引入了可自动核验的判据（例如强制的数据与代码可复现要求、统一的成果核验平台）。因变量取两个：教师端从首次提交到最终定稿的累计工作时长（科研管理系统的时间戳与修改次数可得），评审委员会端的阅读与讨论时长（评审系统的时间戳可得）。不使用跨国或跨地区对比——那种对比在文化、规模、制度上处处不同，所谓控制在方法上站不住。

条件二（顺序读数）。

一条轮次形式的失效条件已由本书第四章第四章《认领物》第十七节给出（先动的是

本体产出效率则认领比在六至十八个月内上升；若无人投资区分能力则下一轮发起处仍是它；只有发起处换到区分能力一侧才回落；若观察到发起处未换手而比值自行回落，该条即为伪）。本章直接采用它，不重复陈述——顺序比相关难被巧合满足得多，这是本章选择轮次读数而不是相关系数的原因。

条件三（第三层，不依赖前两层）。>在存在可自动核验判据的领域（有裁判器的竞赛、可自动测试的编程任务、有强制可复现要求的实验科学），被评价方的准备时间占比应显著低

于不存在此类判据的领域，且这一差异在控制领域声望与竞争强度之后仍然成立。若不成立，本章第三层的经验主张为伪。

这一条最便宜，也最独立——即使第一层的构念被推翻，它仍可单独检验。

条件四（边界条件）。>若在产出携带充分物质痕迹的领域（传统绘画、手工艺、现场表演）观察到同等程度的辨认转嫁，则第七节第九处所加的约束不成立，本章的适用范围界定错误。

这条判断若被证伪，我们会学到什么。

若条件一被推翻，说明困境确实是绝对供需问题，那么一切关于

评价方式的改造都不会有帮助，唯一的出路是位置数量的扩张——这是一个完全不同的干预方向，且成本高得多。

■ 十三、边界、反噬与伦理

不适用的范围。

本章不适用于产出携带充分物质痕迹的领域（第七节第九处）；不适用于辨认能力本身

充裕的领域（乙格）；不适用于评价并不发生的场合——没有人评价，就没有辨认，也就没有转嫁。本章也不主张同形劳动全无价值：一份好的说明文档确实有独立价值，这正是它难以被识别为负担的原因。

反噬预言。

认不出来，就得自己交代

本命题一旦被制度采用，最省事的实现方式恰恰会摧毁它要保护的东西：机构会开始测量辨

认比并把它写进考核，而被评价方为把这个数做好看，最便宜的办法是减少说明与留痕——而留痕本身有独立价值（可复现、可审计、可交接）。于是一个本用于揭示负担转嫁的数，变成压制必要沟通的理由。

我看不到出口。可以做的只有把这个风险写在明处，并配上下面三条。

伦理三条（可被拿去考核人的读数必须配）。

其一，辨认比指的是位置，不是人。

它刻画的是某个评价体系两侧的关系，不得用于个人绩效评定、不得用于个体间比较。

其二，不得为把这个数做好看，而在安全关键的系统里削减说明、留痕与交接材料。

在医疗、航空、金

融清算、公共安全等场合，留痕的独立价值高于该读数的优化。

其三，已经按这个读数做出不利安排的机构，必须公开编码规则（什么计入分子、什么计入分母、如何处理同形的模糊部分）并给出复核通道。

三条缺一，这个读数不该被使用。

■ 十四、赌注

一条写死日期的赌注，并写死什么不算命中。

>到二〇二八年十二月三十一日，至少有一家中国的"双一流"高校，或一家员工总数超过一万人科技企业，在其正式的、常规适用的评价办法里写入一条限制被评价方提交材料总量的硬条款——例如"申报材料正文不超过若干页，不接受补充附件"——并同时写入一条评价方的时间投入承诺。

什么不算命中：倡议、指导意见、"鼓励精简"一类的表述不算；只对某一次评审临时生效的通知不算；只限制被评价方而不含评价方投入承诺的条款不算——那只是把转嫁改成了限额，没有改变归属。

■ 十五、本章自己在哪一格

本章自己就是它所描述的那类东西的一份标本。

这是一篇为了在一个辨认能力有限的评价体系里被辨认为"有内容"而写的两万字。它的相当一部分篇幅——逐条的分界、可证伪条件、来源链接、量纲的四条性质检验——正是同形劳动：与真实判断同形、无法被分离计量、且随读者辨认能力的下降而增加。如果读者能在三百字之内判断这篇文章有没有内容，本章就不必写这么长。

具体后果有两条，都是本章必须承担的。

认不出来，就得自己交代

其一，本章不能靠篇幅或引证数量证明自己不是同形劳动的产物。

这两样恰恰是同形劳动最典型的形

态。

其二，本章只能靠那个比率被人实际算一次来检验。

第十二节的条件一是最便宜的一条，所需数据在多

数机构的既有记录里已经存在。如果没有人真的去算它，本章就停留在它自己所诊断的那一类工作里——一份为了可被辨认而追加的、看上去很像判断的东西。

本章把这条活口留在明处。

三门学科对同一件事给出了三条互不相容的解释，而三条都对：它们各自描述了发起处落在不同位置的那些轮次。争论之所以无解，是因为三家共同假定发起处由系统的性质给定；而工程学自己每天在做的那个操作证明，发起处是可以被调换的。

一旦这条前提被撤销，眼下这一类困境就可以被重新写一遍：它不是要素固定下的劳动内耗，不是为分离而追加的信号支出，也不是指标被博弈后的失真，而是辨认负担从评价方向被评价方的整体转嫁。它产生的那类工作与真实产出同形、无法被分离计量，因而在三门学科的账本上都不设字段——每一家的账各自都能记平，而人确实越来越累。

出路不在更努力，也不在三家各自的处方之间挑一个。它落在一个位置上：让下一轮的发起处换到辨认能力一侧。三种可读数的形态已在第十一节给出，一条最便宜的检验已在第十二节给出。

还有一句要留在最后。本章所描述的这件事，没有反派。评价方并没有偷懒——他们面对的待辨认量确实翻了几番，而可用来分辨的特征同时贬值；他们要求更多材料，是在一个坏局面里做能做的事。被评价方也没有作假——他们追加的每一份说明、每一段过程记录，单看都是正当的、甚至是有用的。困境之所以难解，正在于它由一连串各自正确的动作叠成，每一步都通过了全部形式审查，而合起来的结果是所有人都更累而没有人更好。

这也是为什么"少卷一点"这样的劝告不起作用。在一个辨认负担已经整体下沉的体系里，个人的克制只会使自己更难被辨认，而不会使体系恢复辨认能力。可动的位置只有一个，就是把这份工作搬回它该在的那一侧；而这件事不是任何个人能做的，它需要一条写进评价办法的硬条款。第十四节的赌注赌的正是这一条会不会被写下来。

至于本章本身是不是它所诊断的那类工作——那要看有没有人去算那个比率。

认不出来，就得自己交代

■ 注释

〔一〕本章所说的"辨认"，指判定一份产出是否合格、是否出自署名者之手这一整套工作，不限于身份核验。"分辨率差"指待辨认量的增长速度与可用分辨特征的贬值速度之和，减去评价方辨认能力的增长速度。

〔二〕"同形"取其字面意义：在既有记录体系里与真实产出具有相同的形式，因而不可被分离登记。它不等于无价值，也不等于虚假。

〔三〕本章为理论建构与研究设计，不报告任何虚构的样本、统计结果或实验

■ 结 论

。第七节的十处兑现是可检验的预测，不是已完成的观测。第十二节的样本设计尚未执行。

〔四〕第五节所述"评价方的辨认能力没有下降"是一个有意的措辞：本章不主张评价者变笨了，只主张他们面对的待辨认量与可用特征之比恶化了。

〔五〕本章与学科通融专栏之四十、之十九、之八的分界见第九节；三者均可在本站栏目内查阅。

■ 人机分工声明

选题、三个学科的指定与最终定稿由王德生负责。文献检索、冲突定位、框架构建、写作与自检由 Claude 执行。

所有对外来源均给出可直接点开核对的地址；文中不报告任何虚构的经验数据。

第六章

第一个读数之前——被计为损耗的那段时间

■ 摘要

一个领域有没有真的长出新东西？三门学科各自给出一个自足的答案，而三个答案不能同时为真。工程学说看输出：一个系统够不够格，看它在标准评测上的表现就够了。经济学说看投入产出比：一个体系还行不行，看它把多少人力组织成多少产出增长就够了——这正是“想法越来越难找”那条著名读数的做法。艺术学说两个都不算数：一样东西是不是那件事，由它所处的条件场决定，而条件场自己是被历史地生产出来的，物件与做法都不决定它。把同一个用人工智能大幅提高了产出的团队交给三家，工程学判进步，经济学判衰退，艺术学判两边都在测一个已经不在那里的东西。

本章论证，三家共享一条从未被说出的假定：一个领域有没有长出新东西，可以在某一维上单独读出来。

推翻这条假定的材料来自其中一家自己：研究生产率的计算必须为每个领域各选一个产出代理——晶体管密度、作物单产、寿命、专利——而代理只能从已经被承认为该领域产出的那一类里选。于是任何真正新的产出都落在分子之外，却已经落在分母之内；这条读数在结构上必然下降，与新东西实际有多少无关。

由此得到的判断是：新与旧的分别不在产出，不在效率，也不在承认，而在一样东西从被做出来到在任何一维上取得第一个读数所需的时间——本章称之为无量期。在取得第一个读数之前，一样东西在三家的账本上都不是“小”，是零：它不构成任何一项。而人工智能这一轮所做的事有一个被忽略的不对称——它把已经有现成读数的那一类的无量期压到接近于零，却完全没有压缩需要先造一把新尺子的那一类。两类的显形速度差因此被拉开了数量级。资源永远流向先显形的那一侧，不是因为人短视，而是因为在任何一维上，尚未显形的东西在账面上等于不存在。

本章给出一句不含判断词的问话（上一轮做完的东西里，有多少件在开工时还没有任何现成的量可以用来判定它），一个三领域量纲不同而比率相同的读数（无量期占比），一张两轴四格的辨别表并指认最危险的一格（断供），十二处兑现，两处反过来把本章削小的约束，与

十种既有说法的判决性分界，与本书同题六篇各一条对照预测，七条排除了最强竞争解释的判错条件，以及一条写死到二〇三一年年中的赌注。

本章为理论建构与研究设计，不报告任何虚构的样本、统计结果或实验结论。

无量期

无量期占比

断供

读数结构

显形速度差人工智能与内卷

■ 一·导论：三个行当同时在问同一句话

二〇二五到二〇二六年之间，三个互不往来的行当各自陷入了同一种争吵。

做模型的人在吵评测。

一项对六十个常用基准的系统研究发现，其中二十九个已经出现高度或极高度的

饱和——不是所有模型都答满分，而是排在前面的那一批被压缩在一起，分数之间的差别不再对应能力上

第一个读数之前

的差别。同一批文献还在追问一个更靠前的问题：这些题目到底测的是不是它们宣称要测的那个东西。一项对四百四十五篇基准论文的检查发现，只有大约六分之一在比较结果时使用了不确定性估计或统计检验。而污染——测试题出现在训练语料里——已经被承认为普遍现象。于是一边是新基准以每月为单位地被造出来，一边是"我们其实不知道自己在测什么"这句话被越来越多的人说出口。

做研究政策的人在吵生产率。

一条影响极大的读数说：研究投入在持续上升，而研究生产率在急剧下

降。最有名的例子是摩尔定律——今天要维持芯片密度每两年翻一番，所需的研究人员数量是一九七〇年代初的十八倍以上。这条读数被广泛用来支持"想法越来越难找"这个判断，也被广泛用来解释为什么投入越来越多而增长越来越慢。而人工智能进入研发之后，争论变得更急：如果工具把每个人的产出放大了几倍，为什么这条曲线没有掉头。

做艺术的人在吵什么还算数。生成模型能在几秒钟内产出以往需要数月的东西之后，"这不算作品""这算不算现场"重新变成一个要认真回答的问题。表演研究里有一条已经争了三十年的线：一派认为现场性是不可复制的自体属性，另一派认为"现场"这个范畴根本不是自体给定

的，它是在录音与广播技术出现之后才被历史地生产出来的——没有"录下来的"就不会有"现场的"。这一派的推论很不客气：所谓的本真性并不站在技术之外，它是技术的产物。

三个行当，三种材料。表面上是三场专业内部的争吵，实际上三家在同一个问题上给出了不能并存的答案：>一个领域有没有真的长出新东西，该在哪儿读出来？

这个问题不是学院里的问题。它决定了钱往哪里去、人往哪里去、什么被继续做下去。而在生成成本塌陷之后，它变成了一个急迫的问题——因为"看起来在长新东西"和"只是把同一样东西做得更多更快"，在今天的所有读数上，可以完全一样。

■ 二·三家的位置

选这三家，不是因为它们离得远，是因为它们各自认为只看一处就够了，而三处互不相同。

2.1 工程学：看输出就够了工程学这一支的承重主张是：一个系统够不够格，看它的输出在标准条件下的表现就够了；怎么实现的不重要，处在什么环境里也不重要。

这是黑箱验收的逻辑，也是过去十年人工智能工程的组织方式：造一

个基准，跑一个分数，排一个榜，据此决定谁值得继续投入。

这套做法的力量在于它把判断变成了可传递、可复现、可被第三方核对的东西——任何人拿同一套题目跑一遍，得到同一个数。它的代价现在也已经被这一支自己写清楚了：饱和（分数差不再对应能力差）、污染（题目进了训练语料）、构念效度存疑（不确定测的是不是要测的那个）。值得注意的是，这一支给出的处方始终是造更好的基准——活的基准、持续更新的基准、防污染的基准、覆盖多个维度的整体评估。

也就是说，它承认现有的尺子不好，但它不承认"尺子"这个位置有问题。输出仍然是唯一的判定处。

第一个读数之前

2.2 经济学：看投入产出比就够了经济学这一支的承重主张是：一个体系还行不行，看它把多少投入组织成多少产出增长就够了。

长期增

长率被拆成两项之积——有效研究人员数量，与他们的研究生产率；把产出增长除以研究人员数量，就得到研究生产率。这条读数在多个领域给出同一个方向：投入在涨，生产率在跌。

这一支的力量在于它跨领域可比、可以做长时间序列、并且直接对应资源配置决策。它的做法里有一个必须交代清楚的技术细节，本章后面要用到它："想法"没有统一单位，所以每个

领域必须各选一个产出代理——半导体用晶体管密度，农业用作物单产，医药用寿命或获批新药数，整体经济用全要素生产率。这些代理各有各的道理，而且都是这一支自己反复讨论过、公开承认过的选择。

2.3 艺术学：看条件场就够了艺术学这一支的承重主张是：一样东西是不是那件事，由它所处的条件场决定；物件与做法都不决定它，而且条件场本身是被历史地生产出来的。

表演研究里那条最锋利的论证是这样走的："现场"不是一种可

以在技术之外被指认的属性；它作为一个范畴，是在录音与广播出现之后才成立的——在没有任何东西可以被录下来的年代，没有人需要"现场"这个词。而一旦这个范畴成立，现场表演就开始向被媒介化的形式靠拢，两者的界线不断被重划。

把这条主张放到本章的问题上，它给出的诊断很不客气：输出的分数与投入产出比都是代理物，而一个代理能不能代表那件事，取决于承认它的那套条件；人工智能改变的恰恰是那套条件本身。

所以两个读数

不是测得不准，是在测一个已经不在那里的东西。

■ 三·冲突定位：同一个对象，三家判进步、判衰退、判测错

三条主张听上去可以各管一段。放进一个具体对象，就管不成了。

对象：一个把人工智能大规模用进日常工作的研发团队。一年之后，它的输出在所有标准评测上都更好了；它的人员与算力开支涨了三倍，而它所在领域的产出增长率没有变化；它做出来的东西在同行会议上被讨论的方式与三年前不再相同——原来用来评价这类工作的那套语汇正在失效。

· 工程学判进步。输出更好了，这是可复现、可核对的事实。至于开支和语汇，那是别人的事。

· 经济学判衰退。

投入涨了三倍而产出增长率不变，研究生产率跌了。至于评测分数上涨，那是分子里的噪声——评测本身就是这个体系自己造的，用它来证明自己变强了是循环论证。

· 艺术学判两边都在测错。

评测的内容是由旧的问题结构定义的；产出代理是由旧的承认关系挑出

来的。一年之内这两套条件都动了，所以两个数不能跨期比较，而三家争的那个"变强还是变弱"，在新的条件场里可能根本不是同一个问题。

三家互斥，而且是划界互斥——对同一个对象是不是属于"真的长出了新东西"这一类，三家给出不能并存的归属。这比结局对立更硬，因为你可以取任何一个具体对象，逼三家各自给出"算/不算"。

第一个读数之前

更要紧的是，每一家每天在做的事，正是另一家诊断为病根的事：·工程学每天在做的把判定固化成可重复的题目，正是艺术学诊断为"用一个由旧条件定义的代理去替换那件事本身"。

·经济学每天在做的用一个产出代理去计量整个领域，同样落在艺术学的诊断里。

·而艺术学的处方在另外两家看来根本不是处方：如果两个可核对的读数都不算数，剩下的是什么？

是"要看那件事本身"——而"那件事本身"没有任何人能拿出来给第三方核对。

三家争的，说到底是"有没有长出新东西"该在哪一处结算。而这个争法本身，藏着一条三家都没说出口的假定。

■ 四·三家共同假定了什么，以及推翻它的材料来自哪一家

4.1 那条共有假定把三家的争论压成一句：三家争的是"有没有长出新东西"这件事该在哪一维上被读出来。

于是三家共同假定了：>这件事可以在某一维上单独读出来——只要那一维的读数取对了，另外两维就不必再问。

把这句话念给三家听，三家都会说"这还用说吗"。它不是谁的主张，它是三家的主张之所以能成为主张的前提：如果不预设某一维单独够用，"该在哪儿读"这个问题就问不出来。

4.2 推翻它的那件材料，来自经济学自己推翻这条假定的材料不必从外面搬，它就在第二家自己的做法里，只是那件事在它自己的领域里回答的是另一个问题。

研究生产率的计算要回答的问题是：怎么把"想法产出"变成一个可以做除法的数。它的答案是：因为"想法"没有统一单位，所以每个领域各选一个产出代理——晶体管密度、作物单产、寿命、获批新药、专利。这是这一支公开写明并反复讨论过的技术选择，不是本章的指控。

把这个选择换一个方向读，得到的却是另一件事：>产出代理只能从已经被承认为该领域产出的那一类里挑出来。一个真正新的产出，在被承认为该领域的产出之前，不可能进入代理；而做它的人已经被数进了研究人员的总数。于是它落在分子之外，却已经落在分母之内。

结论很硬：研究生产率这条读数,在任何"正在长新东西"的时期都会系统性下降,而下降的幅度与新东西实际有多少无关——甚至可能与它成正比。新东西越多、越不像旧东西,落在代理之外的部分越大,分子越难看。

第一个读数之前

这不是说"想法越来越难找"这个判断错了。它可能是对的。本章说的是更麻烦的一件事：这条读数在结构上不能用来分辨"真的没长出新东西"与"长出来了但还没有代理"。而这两种情形要求完全相反的处置。

一旦承认这一点,三家的共有假定就站不住了。因为同样的结构在另外两维上一模一样：
· 输出维：一个基准只能测已经被题目化的能力。一个真正新的能力在被造出题目之前,在所有现成基准上要么得零分,要么不适用——而"不适用"在榜单上不显示为任何东西。基准饱和与构念效度那一整批讨论,正是这条在这一支内部的表现：它们诊断出了尺子的毛病,却把毛病归给尺子做得不够好。

· 条件维：这一支自己承认,条件场是被历史地生产出来的——也就是说,"被不被承认"永远滞后于"发生"。用一个滞后量去判定它的先行量,判出来的永远是上一轮的东西。

三维互生,任一维都读不出全部。三家不是选错了那一维,是假定了存在这样一维。

4.3 由此打开的问题如果没有任何一维能单独读出新旧,那么真正要问的就不是"该在哪儿读",而是：>一样东西从被做出来,到它在任何一维上取得第一个读数,中间隔了多久？而这段时间里,它靠什么活着？

这个问题在三家的框架里都问不出来。它也正是本章接下来要回答的那一个。

■ 五·承重判断：无量期

5.1 名与义先把一个词定死。

无量期：一样东西从被做出来,到它在任何一维上取得第一个读数之间的那一段时间。所谓"取得读数",指的是它开始能被某一个现成的量所计量——进入某个基准的题目范围、被某个产出代理计入、或被某个既有的类别接住。

关键在于：在取得第一个读数之前,它在三家的账本上都不是"小",是零。

它不是一个数值很低的项,

它不构成任何一项。一个还没有任何量能测到的东西,在按量运行的系统里没有本体地位——这不是记录做得不细,是"记录"这个动作在它身上无从下手。

有了这个词,承重判断可以写成一句：>新与旧的分别,不在产出,不在效率,也不在承认,而在无量期的长短；而人工智能这一轮所做的事,是把已经有现成读数的那一类的无量期

压到接近于零，却完全没有压缩需要先造一把新尺子的那一类。两类的显形速度差被拉开了数量级，于是资源被系统性地抽向短的那一侧——不是因为人短视，而是因为在任何一维上，尚未显形的东西在账面上等于不存在。

第一个读数之前

三个被否定的说法各来自一家：产出是工程学的划界标准，效率是经济学的，承认是艺术学的。三条都描述了真实的东西，三条都不能分辨新旧。

5.2 为什么它不是"眼光"问题最容易把本章的判断吸收掉的说法是：这不就是短视吗，不就是急功近利吗，不就是资本没有耐心吗。

这条读法把问题放在动机上，而本章要说的是结构。区别可以写死成一句可检验的话：>若问题在动机，那么一个动机足够长远、考核周期足够长的组织，应当不受此影响。>若问题在结构，那么即使动机与考核周期完全相同，账面回报仍会系统性偏向短无量期的那一侧——因为长无量期的那一类在账上根本不出现，它不是"回报低"，是"没有条目"。

第二种情形有一个反直觉的推论：一个真心想做长线的机构，在它自己的报表上会看到它的长线投入什么都没有产生——不是产生得少，是那一栏是空的。于是每一次复盘都会给出同一个建议：把资源挪到看得见结果的地方去。这个建议在它自己的证据体系内部是正确的。

5.3 这一轮为什么被拉开了生成成本的塌陷不是均匀地作用在两类事情上。

它压缩的是"做"的时间。

一份分析、一段代码、一幅图、一篇初稿，从想到做完，从几周压到几小时。

对于那些已经有现成读数的事，做完就是显形——它立刻进入某个基准的题目范围、被某个代理计入、落进某个既有类别。所以这一类的无量期，本来就短，现在接近于零。

它没有压缩的是"造尺子"的时间。

一把新尺子——一个新的评价指标、一个新的期刊栏目、一个新的评

奖门类、一套新的验收标准——要成立，需要有人达成共识、需要样本、需要争论、需要制度承认。这些的速度基本没有变，因为它们受限的不是产出能力，是协调能力。所以那些需要先造一把新尺子的事，无量期仍然是几年。

差值因此被拉开：原来大概是几个月对几年，现在是几小时对几年。这不是程度问题，是账面上的两类东西彻底分了家。

而任何按量运行的系统——季度复盘、年度考核、赛道估值、榜单排名——在这两类之间做选择时，结果是被写死的：短无量期的那一类在每一次结算时都在账上，长无量期的那一类在每一次结算时都不在账上。这不需要任何人做出偏好，它是加总规则的算术后果。

这就是本章认为"新经济现象"这个说法名副其实的。它不是旧内卷的加剧。旧内卷的画像是"更多人挤在同一条路上"；这一轮的画像是账面上只剩下一条路，因为另一条路在账面上没有出现。

■ 六·两轴四格，与最危险的那一格

一个名字不是一条辨别维度。要能分辨，得有第二根轴。

第一个读数之前

第一根轴是这件事的无量期：做完之后立刻能被某个现成的量读到，还是要等一把尺子被造出来。第二根轴是有没有人替那段没有读数的时间买单：存不存在一个具体的位置，愿意在没有任何量能证明它有价值的情况下继续供养它。

有承担者无量期短无量期长

无承担者②无所谓：读数自己会来，有没有人

① 正常量产：做完即显形，也有人负责它做得对不对。

扛都一样。绝大多数日常工作在这一格。

③

真正的研究与创作：有人替那段没有读数的时间扛着

——长周期基金、导师、赞助人、耐心的资本、一个愿意

④ 断供

等的老板。

第四格是本章要打的那一格，本章给它一个名字：断供——一样东西被做出来了，它需要一段时间才能被任何量读到，而没有任何位置替这段时间买单；于是它在取得第一个读数之前就停了。

选这一格，不是因为它最惨，是因为它最不像有问题。一格值得被打，要满足三条：其一，它在短期指标上表现为改善。

把资源从④挪到①，产出量上升、评测分上升、研究生产率的分子

上升、项目按时交付率上升。每一个数字都变好，而且是真的变好——不是伪造的。

其二，它的失效不产生任何信号。

④里的东西停下来时，没有任何账目减少一分——它本来就是零。唯

一能记录这次损失的地方，是它本来会取得的那个读数；而那个读数从未出现过。

一个从来不曾进入任

何账本的东西，它的消失在原理上不可被记账。这一条与错误的项目被砍掉正好相反：砍掉一个有读数的项目会立刻在报表上留下一个缺口，而④里的东西消失时，报表更好看了。

其三，它自我加固。

越是把评价做得及时、精确、频繁——这正是过去三十年所有管理改进的方向——

长无量期的东西越吃亏。每一次把考核周期从一年缩短到一季度，都在把④扩大。

而缩短考核周期在任

何一套管理学说里都是进步。

三条齐了，靶格成立。也正因为三条齐了，断供不会被自己发现——它只能被从外面数出来，而且只能靠数那些还在的东西的构成，不能靠数消失的东西。

断供长什么样抽象地说三条签名不够，得看见它。下面三个是把常见情形合成的典型场景，不是对具体个案的报道；它们的用处是让读者认出自己见过的那一种。

一个实验室里的方向。

某个人做出了一个现象。它是真的，重复得出来，而且他自己知道它要紧。问题

是没有任何现成的指标能说明它有多要紧——它不在任何一个既有方向的评价体系里。第一年的年度汇报，他这一栏是空的：不是数字难看，是没有栏目可以填。第二年，出于好意，这个方向被并进一个有指标的相邻方向，"先按那边的口径报着"；并进去之后，他的时间被那边的产出要求占满。第三年他离开了，去了一个不需要按季度报数的地方，或者干脆去做别的。

第一个读数之前

这个过程里，没有任何一份文件记录了"我们停掉了什么"。没有立项被撤销，没有申请被拒绝，没有争论。年度报表在这三年里逐年好看，因为空的那一栏被有数的栏目替换掉了。如果五年后有人问"我们当初为什么没做那个方向"，档案里查不到任何答案——因为从来没有人做过那个决定。

一支工程团队里的探针。

一个工程师发现现有的监控体系测不到某一类问题：它不产生告警，不进日

志，只在事后回看时能从几个不相干的指标的形状里勉强看出来。他花了两周做了一个粗糙的探针。评审的时候问题很简单也很正当：它现在能测到什么已知问题？答案是零——它是为还没有名字的那一类做的。于是它被判为无收益，排期给了别的事。

六个月后，那一类问题以一次事故的形式出现了。事后复盘写的是"监控覆盖不足，需增加告警规则"，随后新增了七条规则。那支探针不在任何记录里：它没有被否决过，它只是没有被

排进去。而新增的七条规则测的是这一次已经发生过的那个形状——它们把这一类的无量期变成了零，同时把下一类的仍然留在几个月。

一个不属于任何门类的作品。

一位创作者做的东西横跨两三个既有形式之间。征集的报名表第一栏是门

类，必选；她每次都得挑一个最不离谱的填进去，然后在评审里被按那个门类的标准评——评语通常是“作为某某类不够成熟”。三年之后她开始做能被归类的东西，因为那样至少能进入被讨论的范围。

转变的那一刻，没有任何人做过决定，也没有一封拒信。

她收到的每一条评语都是就事论事、态度诚恳

的；每一位评委都在按自己领域的标准认真工作。而报名表那一栏——一个为了统计方便而设的必选项——在入口处完成了整件事，不需要任何人的判断参与。

三个场景的共同点，正是第六节那三条签名：每一次账面都在改善（汇报更完整、监控覆盖更全、评审更规范），失效不产生任何信号（没有决定、没有拒绝、没有记录），而且越规范越严重（门类越细、指标越全、规则越多，滤得越干净）。它们不是失败的案例，它们是运转良好的系统的正常输出。

■ 七·一句不含判断词的问话

要把一格从纸上落到地上，需要一句可以拿去问文件的问话，而不是拿去问人的判断。本章的那一句是：>上一轮做完的东西里，有多少件在开工的时候还没有任何现成的量可以用来判定它？

这句话里没有“应当”“重要”“真正”“有价值”。它问的是立项书、验收标准、评审细则、报名表——问的是记录，不是意见。

把它在五个场合各跑一遍，答案互不相同：一家实验室的立项书。查每份立项书“评价指标”那一栏：填的是现成指标，还是“待建立”。这一栏几乎所有资助体系都有，而几乎没有人把它当作一个可以统计的量看过。

一支工程团队的季度计划。

查每一项的验收标准在开工时存不存在。这是五个场合里最容易查的一个，一个下午能查完。

第一个读数之前

一本期刊的栏目。查最近一年的来稿里，有多少篇不属于任何既有栏目或

关键词

表。这一条查不出来——因为投稿系统强制选栏目，一篇不属于任何栏目的稿子在系统里根本无法提交。"查不出来"本身就是答案，而且是五个里最尖的一个：形式系统在入口处就把无量期长的东西滤掉了，连一次拒绝都不需要产生。

一次艺术征集。

查入选作品里有多少无法归入任何既设门类。答案通常是零，理由与上一条相同——报名表要选门类。

一家公司的季度目标。

查每一条关键结果在制定时有没有现成的度量。答案通常是"全部都有"，而且这不是疏忽：主流的目标管理方法论明确要求关键结果必须可度量。

这一条是对照组——一套方法论把无

量期长的事在定义上排除掉了，而且它是这样写进教科书的。

五个答案互不相同，其中至少两个不是从命题推出来的、是查出来的：投稿系统的强制选项，与目标管理方法论的定义性要求。前者说明滤除发生在入口而非判定；后者说明这不是执行走样，是设计意图。

八·一个量纲不同而比率相同的读数

一句问话只能问一件事。要让不同行当能并排比较，需要一个数。

无量期占比：最近 N 件已立项或已完成的产出中，在开工时不存在任何现成读数——没有既有指标、基准、栏目、门类、验收标准可用——的件数占比。

它在四个行当里的量纲完全不同，比率却相同：· 学术：立项书"评价指标"一栏填"待建立"的比例。

· 工程：开工时没有可用验收基准的项目比例。

· 艺术：入选作品中无法归入任何既设门类的比例。

· 企业：季度目标中没有现成度量的条目比例。

这个读数有四条性质，缺一条都不该用它。

其一，无量纲，所以四个行当可以并排比较：一家实验室的零点一与一本期刊的零点一说的是同一件事。

其二，可事后清点：所需材料在既有记录里已经存在——立项书、验收文档、报名表、目标清单。不需要新建任何系统。

其三，它把一个印象变成一个数："我们这里只做有把握的事"是一句人人都说过的抱怨；抱怨不能被检验，比率可以。

其四，也是最要命的一条：它与既有的主要指标正交。

一个与既有指标高度相关的新读数会被立刻读成

那个指标的代理，白造。所以必须举一个"既有指标最好看、而这个数最难看"的真实例子——举不出来就该重造。

第一个读数之前

本章的例子是一个所有项目都能在季度末给出量化结果的研发部门。按既有指标看，它接近完美：交付率满分，产出量创纪录，每一项都能报出数字，评测分年年上升。而它的无量期占比是零——因为一个开工时没有现成度量的项目，在这里根本立不了项。这两个读数指向相反的结论，而且都不是错的：它确实每一项都做成了，也确实一件没有尺子的事都没做过。

反过来也成立：一家产出量很低、报表很难看的小实验室，可以有很高的无量期占比。产出量低不等于在做新东西，产出量高也不等于没在做——这两个数第一次被分开，正是这个读数的用处。

■ 九·十二处兑现

兑现不是"像"，是在那个领域里能据此预测一件具体的事。

一
·

公共科研立项：在生成工具普及最快的学科里，立项书"评价指标待建立"的比例会先于"原创性下

降"的抱怨出现，而不是相反。可查：资助机构的立项库。

二·期刊栏目：栏目表的更新频率会落后于来稿主题的变化速度，且这个差值在生成工具普及后扩大。预测：跨栏目投稿被退回改投的比例上升，而这一项通常没人统计。

三·工程验收：有现成验收基准的项目占比会上升，而不是下降——因为无基准的项目在排期会上说不出预期收益。预测：这一上升与团队规模无关，与考核周期长度强相关。

四·人工智能评测建设：新基准建成后，被它首次读到的能力中，会有相当比例是在基准建成之前就已经存在而无人计数的。预测：这些能力的"首次出现时间"若可从模型版本记录回溯，将系统性早于基准发布时间数月至数年。

五·专利分类：新技术方向进入正式分类表的滞后，与该方向在此期间的实际申请量正相关——分类跟不上不是因为方向小，恰恰因为方向大到不得不动分类。预测：分类表更新前后，同一方向的"检索可见度"会跳变，而技术本身没变。

六·药物研发：全新机制的项目与已知机制新适应症的项目，在同一家公司内部会有系统性不同的存活率，而这个差别不由科学风险解释——已知机制的项目在每一次季度评审时都有现成终点指标可报。

七·艺术征集与评奖：门类表越细，跨门类作品的入选率越低；而门类表的细化在任何评审改革里都被写作"更公平"。预测：细化门类之后，入选作品的形式多样性下降而评审争议减少。

八·风险投资：赛道命名成立之前的项目融资难度，与该赛道成立之后同类项目的融资难度差一个数量级，而项目本身没有变。预测：赛道命名的时点可从投资机构的公开文章回溯，融资难度的跳变应紧随其后。

九·教育课程：一门还没有考试形式的课，在任何课程评估体系里都拿不到分。预测：新设课程会系统性地先设计考核方式、后设计内容，而这一顺序在教学大纲的修订记录里可以直接看到。

十

·

第一个读数之前

政府科技计划指南：指南里"研究内容"与"考核指标"两栏的对应关系，决定了哪一类工作能被写进

去。预测：要求填写量化考核指标的计划，其资助方向的年际变化率显著低于不要求的计划。

十一·企业目标管理：把考核周期从一年缩短到一季度的组织，其无量期占比会在两个周期内下降，而所有交付指标同时改善。这一条是本章最直接的一个可观察后果。

十二·开源项目：一个还没有对应标签的问题类型，在议题系统里既搜不到也统计不到；预测：标签体系的更新会滞后于问题类型的出现，而滞后期内该类问题的重复提交率异常高——重复提交正是"读数不存在"的一个副产品。

十二处里有两处反过来把本章削小了，下一节单说。

■ 十·两处反过来削小本章的约束

10.1 工程学削掉的：适用范围在安全关键与合规领域，没有现成读数就不许开工是正确的规矩。你不能把一个没有验收标准的系统装上飞机，不能把一个没有终点指标的方案推进临床，不能让一段没有测试的代码进入支付链路。在这些地方，无量期占比接近零不是病，是设计目标。

这一处直接削掉了本章原本会写下的一句更强的话：>原本会写：无量期占比越高越健康，组织应当提高它。>现在只能写：本章不主张提高任何组织的整体无量期占比。这个读数只适用于以探索新可能为职能的那一部分单元，而且它的用法是"这个组织有没有留出这样的单元"，不是"这个组织整体应当多冒险"。

这是把适用范围砍掉了一大块：工业质量控制、安全工程、合规、临床推进，全部不在内。判据也随之明确：这个单元的职能是产出可靠的结果，还是找到还没有人做过的事？前者无量期短是目标，后者无量期为零是断供。

10.2 艺术学削掉的：价值排序与单件适用性艺术学那一支提醒本章一件更难受的事："新"这个范畴本身也是被条件场生产出来的。

一件事在开工时

没有现成读数，并不等于它是新的；它也可能只是没有人关心。一个无人问津的方向和一个尚未被识别的方向，在无量期这个读数上完全一样。

这一处削掉的是本章的价值排序，也削掉了它在单件上的用法：>原本会写：无量期长就是新，就值得被供养。>现在只能写：无量期长是"新"的必要条件而非充分条件。因此这个读数不能用来判定任何单独一件事的价值——它只能用来判定一个组织的构成：这个组织有没有为长无量期的那一类留出位置。用它去判单件，等于把一个统计性质当成个体属性，那是误用，而且是会伤到人的误用。

第一个读数之前

这两处约束合起来，把本章从一条关于"什么值得做"的普遍主张，压成了一条关于"组织构成"的有限主张。不接受这个压缩，本章会在第一个安全工程师和第一个策展人那里各栽一次。

10.3 经济学补的一处研究生产率那条读数在有稳定代理的领域——半导体、农业——确实有效，而且它的下降在那些领域里很可能就是字面意思。所以本章不主张废弃它，只主张它不能单独用来判新旧，并且在正在长新东西的时期，它的下降不构成证据。这一条削掉的是本章原本想写的"这条读数应当被弃用"。

■ 十一·与最近的几种说法的分界

本章最容易被吸收进以下十种既有说法。逐条写清差在哪，以及怎样才能判出谁对。

一·想法越来越难找 (Bloom, Jones, Van Reenen & Webb 2020, 经济学)。它说到的那一步是：研究投入上升而研究生产率下降。分界在于：本章认为它测的是“在既有代理下的产出效率”，而代理只能取自己被承认的产出类。判定形式：若某领域在换用一个更宽的产出代理之后，研究生产率不再下降，则本章对；若换任何代理都下降，则它对。这条可以在有多种可选代理的领域（医药、农业）直接跑。

测)。它们诊断“基准测不准”，处方是造更好的基准。分界在于：本章说任何基准在原理上都测不到还没有基准的那一类，造更好的基准只会缩短已有类的无量期、把差距拉得更大。判定形式：若新一代基准建成后，被它首次读到的能力中有相当比例在基准建成前就已存在而无人计数，则本章对；若首次读到的能力确实是新出现的，则它们对。

三·常规科学与反常 (Kuhn 1962, 科学哲学)。最近的邻居之一。库恩说反常在旧范式里读不出来，直到危机积累。分界在于：库恩关心的是范式转换的逻辑，本章关心的是转换之前那段时间由谁供养——本章比它多一条：读不出来的那一段有长度，而这个长度是可测的、并且是可被制度改变的。判定形式：若两个同等规模的领域，其反常被识别的滞后时间与“有没有位置为无读数期买单”无关，则本章这一条错。

文最危险的方法学占位者，因为它已经有一套成熟的指标：睡美人系数，用来量一篇论文沉睡多久才被引用。分界在于：睡美人测的是发表之后的引用滞后，本章测的是开工之时有没有现成读数。二者在同一件工作上可以完全相反——一篇论文可以引用即刻爆发，而它当初开工时并无任何现成指标（无量期长而系数低）；也可以开工时指标齐全，却因为无人关心沉睡二十年（无量期短而系数高）。判定形式：若睡美人系数与无量期占比高度相关，本章应当被科学计量学的既有指标吸收。

这两个量在同一批文献上都能

算，这条检验现在就可以做，本章最愿意拿它去赌。

五·探索与利用的权衡 (March 1991, 组织学习)。它假定二者是同一个资源池上的配置选择。分界在于：本章说二者的显形速度不同，所以即使配置比例被规定死，账面回报仍会系统性偏向利用。判定形

第一个读数之前

式：在探索与利用投入比例被强制固定的组织里，若账面回报仍系统性偏向利用，则本章对；若回报差随比例固定而消失，则它对。

六·技术就绪度等级 (NASA/欧盟框架, 工程管理, 方法学占位者)。它已经把“离可评估还有多远”做成了九级量表。分界在于：这套量表测的是离应用的距离，本章测的是离第一个读数的距离。二者可以背离：一个就绪度很低的基础研究方向可以指标齐全，一个就绪度很高的

产品可以没有任何人知道该怎么衡量它好不好。判定形式：取一批项目同时打两个分，若两者高度相关，本章应被它吸收。

七·不可通约性（Kuhn/Feyerabend，科学哲学）。本章不主张范式之间不可翻译。本章只主张翻译需要时间，而这段时间无人记账。分界的判定形式：若翻译在原理上不可能，则供养这段时间毫无意义，本章的处方失效；若翻译只是慢，则本章成立。这一条本章站在乐观一侧，因此可被悲观一侧证伪。

八·马太效应与累积优势（Merton1968，科学社会学）。它解释的是已经显形之后的放大：得到承认的更容易得到更多承认。本章处理的是显形之前。判定形式：若把初始识别的时点控制住之后，资源分配的偏斜完全消失，则累积优势足够，本章多余；若控制之后偏斜仍在（因为有一类从未有过初始识别），则本章对。

九·注意力稀缺（Simon1971，经济学与认知科学）。它说稀缺的是接收者的注意力。分界在于：本章说稀缺的是愿意在没有读数的情况下继续供养的那个位置。注意力可以被外包给机器，供养不能——一个模型可以读完所有立项书，但它不能承担"这三年什么都报不出来"的后果。

十·反转模板专查。把本章的命题去掉本领域的一切名词、只留形式，得到的是："一样东西在取得读数之前不进入任何账，因而系统性地被资源抛弃。"

这不是"越成功越失败"那一类反转模板，它是遗漏型

——与本书第四章《认领物》同型，都是缺栏位。仍要点名最贴近的反转模板：成功陷阱（Levinthal

&

March1993）与能力刚性（Leonard-Barton1992）。分离线：那两条的机制在组织被过去的成功锁死；本章的机制在账本的时间分辨率——一个从未成功过、毫无路径依赖的新组织，只要它按季度看数，同样会抛弃长无量期的事。判定形式：若空白组织与老牌组织在无量期占比上没有差别，则本章的机制成立而锁死不是必要条件。

十条里最近的是第四条。本章之所以仍不能被它吸收，只有一个理由：睡美人的前提是那篇论文已经被写出来并发表了。本章说的正是那些在被写出来之前就停了的——它们不会睡，它们不曾醒过。

■ 十二·与本书同题六篇的分界

本书最近有六篇处理同一道题——人工智能时代的内卷困境与出路——而且用的是同样三个学科。四篇与本章各自缺的位置不同，逐一分开，并各给一条判决性对照。

第四章《认领物》说：内卷是产出的构成中“认领物”占比上升——为使一件产出可被归到某人名下而必须一并被生产的那一层，三家账本只给它成本位、不给产出位。它与本章同为缺栏位型，但缺的东西不同：它缺的是一层已经被生产出来的东西，本章缺的是一段还没有产出读数的时间。前者是实体，后者是时段。判决性对照：取一个归属完全免费的组织——公司内部代码仓库，每次提交由身份系统自动记名，不需要任何

第一个读数之前

额外的证明工作。它预测这里认领占比接近零、账面内卷轻；本章预测只要该组织按季度用现成指标看数，长无量期的事照样被抛弃。两个读数正交，可在同一批组织上同时测。

第三章《这道关拦下过谁》说：可读门槛整体饱和，而没有位置有权宣布一道关作废。它管的是已经走到关前面的东西；本章管的是从来没走到任何一道关前面的东西——它们在开工阶段就因为没有读数而停了。判决性对照：把一个组织所有饱和的关全部撤掉（它的第一步），它预测代价下降；本章预测无量期占比不变——撤关不产生任何新的供养。两篇因此互补：撤关减代价，供养无量期才增新。

第一章《两种「不」》说：没有人看过就作出的否决，与看过之后作出的否决，在所有记录上同码。它管被否决而无人看过的东西；本章管从未进入判定的东西。前者进过流程，后者连提交都提交不进去（见第七节那本期刊的例子）。判决性对照：把未阅率降到零（它的第一步全部落实），它预测投递者处境改善；本章预测无量期占比仍不变——被认真看过的，仍然只是那些有现成读数可判的。

第二章《被丢掉的那一眼》说：被判定过而未被接受的产出在无人记账处合成一条判定基线。它与本章恰好互斥：暗基线的前提是有一条线，本章说的是没有线的那一段。判决性对照：一个把参照集轮换做到极致的机构（它的出路），它预测抬升度回落；本章预测无量期占比不动——轮换参照集只是换用哪些旧样本，不产生任何新的参照。

《认不出来，就得自己交代》说：核实这份工作从评价一方移到了被评价一方，由此产生一类与真实产出同形、因而不可分离计量的劳动。它与第四章《认领物》形状很近，落点在转移：谁来做核实。它管的是已经被做出来的那一层劳动；本章管的是还没有任何量能测到的那一段时间。判决性对照：取一个核实完全由评价方承担、被评价方不需要任何额外交代的组织——自动身份核验加抽样复核。它预测这里同形劳动接近于零；本章预测只要该组织按现成指标看数，长无量期的事照样断供。两个读数在同一批组织上正交。

第七章《一件是几件》说：生成成本趋零之后计数单位失去下界，而“并项”这个单位从未被生产出来。它与本章的关系是顺序上的前后，值得写清楚：它处理的是已经被计数、只是单位不清的那一类；本章处理的是连“几件”这个问题都还问不出来的那一类——那不是单位不

清，是没有单位。必须先有第一个读数，才谈得上单位，所以本章所说的那一段落在它的定义域之前。判决性对照：一个把计数单位重新定死（引入并项）的组织，它预测计数恢复可比、账面内卷缓解；本章预测无量期占比不动——定死单位只作用于已经能被计数的那一类，对还没有任何量能测到的那一类毫无影响。

六篇与本章合起来，可以写成一句：第四章《认领物》缺的是一层东西，第三章《这道关拦下过谁》缺的是一个位置，第一章《两种「不」》缺的是一种处置类型，第二章《被丢掉的那一眼》缺的是一份账，《认不出来，就得自己交代》

缺的是一次转移的去处，第七章《一件是几件》缺的是一个计数单位，本章缺的是一段时间。

叠，而它们都长在同一块地上——下一节说这块地是什么。

七处缺口互不重

第一个读数之前

■ 十三·这个判断底下还站着一块更大的地

把上一节那十位邻居逐个问一句“它把什么当作给定，因此结构性地看不见什么”，会得到一份整齐得让人不安的清单。判据只有一条：一旦承认了那件看不见的事，它自己就塌。

· 研究生产率把“产出可被代理”当作给定，因此看不见没有代理的产出。一旦承认有这样的产出，研究生产率的分子就没有定义。

· 基准研究把“能力可被题目化”当作给定，因此看不见不能被题目化的能力。一旦承认，“造更好的基准”这个处方就失去了它的普遍性——它只能改善已经在题目范围内的那一部分。

· 睡美人指标把“那篇论文已经被写出来并发表了”当作给定，因此看不见没写出来的。一旦承认，它测的就不是“延迟承认”而只是“发表后的引用动力学”。

· 常规科学把“共同体已经在场”当作给定，因此看不见共同体形成之前的那一段。

· 技术就绪度把“目标应用已知”当作给定，因此看不见不知道自己要用在哪的东西。一旦承认，一级的定义就落空——第一级是“基本原理被观察到并报告”，而“报告”预设了报告给谁、按什么格式。

· 马太效应把“至少有过一次初始识别”当作给定，因此看不见零次识别的。

· 注意力稀缺把“东西已经在接收者面前”当作给定，因此看不见还没有任何渠道能把它送到任何人面前的那一类。

七行背后是同一件被当作给定的事：>凡是一样东西，从它存在的那一刻起，就至少有一个量在测它。

这句话没有人主张过，因为没有人需要主张它——所有人都站在它上面，包括互相批评的各派。而承认它的反面——一样东西可以存在一段时间，而任何量都测不到它——会让上面这一整批工具同时失去定义域。

这也是本章与本书同题四篇真正的关系。那六篇各自推翻的是同一块地基上的另一个假定：一件东西的生产者是给定的；每一道判定都有一个持有它的位置；每一次处置都有人作出过它；凡是一样东西要么被接受要么退出系统；核实这份工作总归有一方在做；凡是产出总能被数成几件。七个假定形状同族，都是「总有一个……」。

七篇合起来指出的不是七个问题，是同一块地基的七个承重点。而这块地基之所以

从来没有被回头看过，正因为按量运行的系统只能看见它自己能测到的东西——它没有任何机制去问"有没有我测不到的"。

■ 十四·三家为何各自到不了这一条

三家没有走到这一步，不是因为疏忽。一个领域看不见什么，通常不是它的缺陷，是它的问题结构的影子。

工程学的核心任务是把判断变成可传递的证据。在这个任务结构里，一样"还没有任何量能测到"的东西不是一个对象——它甚至不是一个待解决的问题，因为工程学的问题必须以"给定规格，达成它"的形式提

第一个读数之前

出。没有规格就没有问题。所以工程学看得见尺子不好，看不见"还没有尺子的那一段"；它把前者当作技术债，把后者当作还没轮到自己管的事。

经济学的核心任务是配置：把有限资源分到收益更高的地方去。配置分析总是需要一个可以比较的量，因此没有量的选项在配置问题里不是"收益未知的选项"，它根本不是一个选项——它进不了那张表。经济学因此可以非常精确地讨论不确定性、期权价值、探索的收益，但那些讨论全部预设了"我们知道有这么一个选项，只是不知道它值多少"。本章说的那一类，是连"有这么一个选项"都还没有人写下来的。

艺术学的核心任务是抵抗标准化。它比谁都早看见量的不足——它已经看了一百年，从复制技术开始动摇独一性的那一刻起。但它的任务结构使它把出路一律指向否认量的资格，而不是去问那段还没有量的时间由谁供养。为一段无读数的时间设一个供养位置，在这一支的语汇里天然读作"又一次把创作纳入管理"，因此它不会去设计它。

三家的盲点不是三种疏忽，是三种任务结构各自的背面。而这三个背面恰好覆盖了同一段时间。

■ 十五·一旦被制度采用会发生什么

本章提出的读数一旦被采用，最省事的实现方式是"把无量期占比"做成一个考核指标。而一旦它进了考核，让这个数好看的最便宜办法不是去做没有尺子的事，是在立项书的"评价指标"一栏改填"待建立"——一个字段的事，一个下午可以把全院的立项书改一遍。

于是这个数暴涨，而组织的实际构成一动没动。更坏的是，它同时污染了唯一一份能用来检验本章的材料：那些立项书原本是可信的，因为没有人有动机在那一栏上撒谎。

这条反噬我看不到出口。

任何核查都需要有人判断"这个指标算不算现成的"，而这个判断本身没有现成

的量——它正好落在本章所描述的那一类里。也就是说，本章的读数在原理上不能自我保护，而且它比它所要保护的那一类东西更容易被伪造：伪造它的无量期是零（改一个字段立刻显形），而它要保护的那一类的无量期是几年。

本章能给出的唯一防护很弱：这个数只统计、不考核，并且统计口径与统计者必须与被统计单位分离。这不解决问题，只是把伪造的成本从零抬到不为零。写下来，是为了不让读者以为本章没看见这一层。

■ 十六·判错条件

先写出最强的竞争解释，因为不排除它，下面的条件都不算数。那句最朴素也最难反驳的话是：>"其实不就是短视吗？不就是资本没耐心、考核太急吗？"如果问题在时间偏好，那么本章的第二根轴（有没有人供养）就是多余的——供养与否只是时间偏好的另一种说法，而处方只需要延长考核周期。下面第一条就是冲着它去的。

第一个读数之前

一·控制组织的时间偏好之后——把考核周期长度、资金期限、管理者任期这三项控制住——若"无量期占比"与"三年后产生的首创性成果比例"之间不存在剂量—反应关系，则本判断为伪，届时应归因于时间偏好而非读数结构。样本纪律：至少六个同制度环境内的单元（同一资助体系下的六个以上院系或实验室，或同一集团下的六个以上事业部），在学科领域、资金规模、人员规模上同质。不得拿两个国家或两个体制的对比充数——那种对比在方法上支持不了任何"控制掉某变量"的说法，只能作启发式例证。

二·若两个考核周期完全相同的组织，其无量期占比的差异不能预测首创性差异，则第二根轴不成立，四格退化为一维，本章只剩下一条关于时间偏好的旧话。

三·若在一个明确为长无量期项目设立了独立通道的机构里，通道设立后三年内该机构的无量期占比没有上升，则本章的处方无效——尽管第五节的诊断仍可独立成立。

四·若睡美人系数与无量期占比在同一批文献上高度相关，本章应当被科学计量学的既有指标吸收。这两个量现在就能算，这是本章最便宜也最危险的一条。

五·顺序读数：本章预测评价周期的缩短早于无量期占比的下降，而不是同时或滞后。若在时间序列上观察到相反的次序——先是无量期占比下降，然后组织才缩短周期——则本章的因果方向错。顺序比相关难被巧合满足，这一条本章最愿意拿去赌。

六·一条低成本可实做的：取任一机构最近一百份立项书，清点"评价指标"一栏在开工时是否已经存在。

数据在既有档案里，一个下午可以做完。若各学科的分布与本章预测的"生成工具普及越快、该比例越低"不符，本章在那个行当里不成立。

七·一条会推翻本章核心机制方向的：若在生成成本塌陷幅度最大的领域，无量期占比反而上升——因为塌陷把人力从常规工作里释放出来，让更多人去做没有尺子的事——则本章的方向判反了。这个反向机制是真实可能的，本章不认为它已经被排除。

这条判断若被证伪，我们会学到什么：若它与时间偏好分不开，那么出路只能在延长考核周期上找——延长任期、拉长基金期限、放宽年度考核。这与本章的处方（为无读数期单独设一个不进考核的通道）指向完全不同的干预，而且前者要贵得多。证伪本章，会把资源从制度设计推回到期限调整，这个后果值得一试认真的检验。

分层引用：本章可分三层引用。第一层是"无量期"这个构念作为核心判据，最强也最容易倒；第二层是两轴四格作为分析工具；第三层是一条不依赖前两层的经验主张——在生成成本大幅下降的行当里，判定流程与立项流程中"开工时已有现成度量"的比例会上升，而这一上升不伴随该流程总环节数的减少。第三层可以单独被检验、单独被引用；第一层被推翻时它仍然站得住。

■ 十七·出路：不是多冒险，是让那一段有账

本章是一条辨别维度，不是一套方案。但一条维度若推不出任何可做的事，它就只是个说法。下面三件，各对一家，而且顺序是死的。

第一个读数之前

17.1 对经济学：把"造出了一把新尺子"本身记为一项产出现在的账里，造指标、建基准、开栏目、立门类、写验收标准——这些全部记在成本书或"科研管理"栏，从来不记为产出。而它们恰恰是缩短别人无量期的那件事：一把新尺子造出来，一整类原本读不到的东西同时显形。

它的收益因此全部外溢给后来者，在任何一个单位自己的账上都不划算——造尺子的人拿不到用尺子的人的收益。这是一个标准的外部性问题，处方也是标准的：单独记账、单独资助，并且不要求承担者同时产出该领域的成果。

现在的做法正相反：谁想建一个新指标，得先用它做出一批成果来证明它有用；而做出成果又要靠这个指标，循环因此闭合。

最小动作：在成果统计口径里加一类"评价工具"，与论文、专利并列。这一项现在在几乎所有统计体系里都不存在。

17.2 对工程学：在验收之外单设一个不产出验收结论的环节工程学的整套流程围绕"给定规格，达成它"组织，无规格的事进不了排期——不是被拒绝，是根本填不进那张表。最小动作不是取消规格，是允许一部分排期以"本期不产出可验收结果"立项，并且这一项不进入交付率的分母。

关键全在最后半句。只要它进分母，任何团队都不会用它——用一次，交付率就掉一块，而交付率是这个团队被看见的唯一方式。这一条看上去只是会计细节，实际上它是整件事成不成的开关：本章第五节说资源流向是加总规则的算术后果，那么改变流向的动作也只能是改加总规则。

最小动作：在交付率的定义里加一个排除项。一行字。

17.3 对艺术学：供养必须是具名的，而且必须承担后果艺术学在这一点上比另外两家懂得多。赞助人、导师、编辑约稿、画廊的长期代理——这些制度形状相同：一个具体的人押上自己的判断，替一件还不能被证明的事扛一段时间。

押上的是他自己以后说话还

算不算数。

这件事不能被外包给评审委员会。委员会需要依据，而依据正是这类事情还没有的东西；把它交给委员会，委员会只能退回去找现成指标，于是绕回原点。它也不能被外包给模型——模型可以读完所有申请，但它不能承担"这三年什么都报不出来"的后果。

最小动作：一个基金里划出一小部分，由具名的人独立决定、独立署名，事后公开其全部资助记录包括失败的那些；并且不以命中率考核他，唯一的约束是他的全部记录被公开。用命中率考核会立刻把他推回现成指标——那正是本章要避开的那条路。

17.4 顺序为什么是死的第一件必须最先。

不给造尺子记账，另外两件产生出来的东西仍然没有读数，仍然会在下一次结算时从账上消失——你把它们做出来了，而它们仍然不构成任何一项。

第一个读数之前

第二件其次。它是把第一件的结果装进日常流程的接口。

第三件最后。

它最难，也最容易被误用成任人唯亲；必须等前两件把"什么算数"的账先建起来，具名供养才有可被公开检验的背景。顺序颠倒的后果很具体：先做第三件，会得到一批无法被任何人核对的人情资助，两年之内它会被叫停，并且从此把整个方向的信誉一起赔掉。

17.5 一条反直觉的推论从这三件里落出来一条与所有管理直觉相反的话：>在造尺子这件事没有被单独记账之前，任何提高评价质量的努力都会拉大差距。

更好的基准、更细的指标、更及时的反馈、更短的复盘周期——每一项都在让已经有尺子的那一类显形得更快，而对还没有尺子的那一类毫无作用。评价改进因此是单向的：它只加速一侧。这不是说不该改进评价，是说改进评价必须与第一件事同时做，否则它的净效果是把断供那一格扩大。

可检验：一个刚刚完成评价体系升级的组织，其无量期占比会在两个考核周期内下降。

升级的时点在制

度文件里有据可查，占比可从立项书清点，这条现在就能跑。

17.6 为什么不是"多给钱"最常见的处方是加大投入。本章与它有直接分歧：钱会流向有读数的那一类。

在一个所有项目都要报量

化结果的体系里，增加预算的结果是同一类项目变多，而不是另一类项目出现——因为多出来的钱仍然要按同一套规则分配。

这条分歧是可检验的，而且数据现成：科研经费大幅增长的时期，无量期占比应当不变或下降，而不是上升。若观察到经费增长同时伴随该比例上升，本章这一条错，加大投入本身就够了。

■ 十八·不适用的边界

不适用于以产出可靠结果为职能的单元。

安全、合规、质量、临床推进——那里无量期短是目标，本章

的读数方向相反。

不适用于判定单独一件事的价值。

无量期长是"新"的必要条件而非充分条件；一个无人问津的方向与一

个尚未被识别的方向，在这个读数上完全一样。用它去判单件，等于把统计性质当成个体属性。

不适用于把它当作提高冒险程度的理由。

本章不主张组织应当整体提高无量期占比。它主张的只有一件

事：一个组织应当知道自己这个数是多少，并且知道自己有没有为长无量期的那一类留出位置。

本章不主张新的一定比旧的好。

大量长无量期的事最后什么也不是，这不是失败，是这一类事情的正常

分布。本章只主张：它们在停下来时不产生任何信号，因此没有人能知道停掉的是哪一类。

第一个读数之前

■ 十九·伦理与执行边界

无量期占比是一个几乎必然会被拿去考核人的读数。三条硬约束写在这里，缺一条这个读数就不该被使用。

其一，这个读数指的是构成，不是人。

一个单元的无量期占比低，是它的职能与流程的性质，不是研究

者不敢想。把它算到个人头上是误用，而且按本章自己的机制，一个最有想法的人在一个只接受现成指标的流程里也只能立有指标的项目。

其二，不得据此对个人作出不利安排。

尤其不得把"你的项目开工时没有现成指标"用作否定的理由，也

不得反过来把它用作加分项——后者会立刻触发第十五节那条反噬。

其三，若一个机构决定统计这个数，统计口径与统计者必须与被统计单位分离，且只统计、不考核。

一

个由被统计者自己填报并进入自己考核的读数，在本章的框架里是一个无量期为零的东西——它会在填报的当天显形，而它要保护的那一类要等几年。

■ 二十·结论，与一条写死日期的赌注

把全文压成一句：一样东西在取得第一个读数之前，在所有账本上都是零；人工智能把已有读数那一类的无量期压到接近于零，却没有压缩需要先造尺子的那一类，于是两类的显形速度差被拉开了数量级，而资源被算术地——不是被偏好地——抽向短的那一侧。

由此得到的不是一条处方，是一条辨别维度：要分辨一个领域是在长新东西还是在把同一样东西做得更多更快，不能看它的产出、效率或承认，只能看它最近做完的东西里有多大比例在开工时还没有任何现成的量可以判定它。

若一定要从这条维度推出一个动作，最小的那个是：在立项文件里加一个字段——"本项目开工时可用的现成评价指标：有/无"——只统计，不考核，统计者与被统计单位分离。这不解决问题，它只是让那个数第一次存在。而按本章的整个论证，一样东西的第一个读数出现，正是它开始能被讨论的那一时刻。

最后交一条写死日期的赌注：>到二〇三一年六月三十日，在至少三个国家的公共科研资助体系中，会出现把"该方向目前不存在公认的评价指标"本身作为一项正面立项理由、写进公开指南条款的做法。

什么不算命中，一并写死：写"鼓励原创""支持非共识""容忍失败"这类倡导性表述不算；设立了"探索性项目"类别但仍要求填写量化考核指标不算；只在评审内部口头掌握而未写进公开指南不算；把它写进白皮书或战略报告而未落到具体计划的申报条款里不算。

若到期没有出现，本章关于"为无读数期设通道"的处方在制度上不可行，应当被放弃——而第五节的诊断与第十六节第三层的经验主张仍可独立成立。这正是本章采用分层写法的原因。

第一个读数之前

■ 附·本章自己的无量期是多少

接近于零。

本章用的是既有的学术论文形式，投向一个有现成评审细则的通道；它在写完的那一刻就能被现成的量读到——字数、格式、引用、审稿意见、阅读量。因此本章不是它所主张要保护的那一类东西的标本，它站在靶格的反面。

这个定位有两个具体后果，都不好听。

第一，本章没有资格用自己作为证据。

它不能说"你看，我这样的工作也很难被看见"——恰恰相反，它

很容易被看见，这正是它能被写出来的原因。它只能靠一件事被检验：有人去清点一次立项书，并把结果公开。

第二，也是本章无法弥补的一个洞：如果本章的判断是对的，那么真正落在断供那一格里的的工作恰恰写不出这样一篇论文——因为写出来、投出去、被读到，就意味着它已经取得了读数。本章能做的最多是指出那一段的存在，而不能从那一段里拿出任何一个样本。所有能被举出来的例子，都是最终出来了的那些；那些没出来的，按定义不留痕迹。

这不是谦辞，是本章自己的判据用在本章身上得到的读数：一个关于"看不见的东西"的论证，其全部证据都只能来自看得见的那一侧。本章承认这个洞，并且认为任何声称补上了它的说法都值得怀疑。

■ 注释

一·本章所说的"读数"取其最宽的意义：任何一个现成的、可被第三方使用的量或类别——一个指标、一个基准、一个栏目、一个门类、一条验收标准。它不要求这个量是好的，只要求它存在且可用。一个粗糙的量存在，也算取得了读数。

二·"无量期"的起点取"被做出来"而非"被想到"，因为想到的时点在原理上不可核对，而做出来的时点在多数记录里可以定位（立项、开工、首次提交）。这是一个为了可查而付出的精度代价，本章明确接受它。

三·第八节的比例阈值没有设固定值，因为"现成"是一个二分判断而非连续量。跨机构比较时必须使用同一套判断规则并注明，否则数字之间不可比。

四·本章引用的经济学读数（研究生产率下降）与本章的判断并不矛盾：本章不主张它测错了，只主张它不能单独分辨"没长出新东西"与"长出来了但还没有代理"。

五·本章为理论建构与研究设计，不报告任何虚构的样本、统计结果或实验结论。文中所有数字均来自公开来源，出处见

。

第一个读数之前

编三

判断所依赖的单位失灵

• • •

■ 编序

前面两编都在数东西：多少件被否决、多少份没被打开、多少工时被转嫁。这一编要问的是：那些数还算不算数。

判断不是凭空作出的，它依赖一套计量约定。这套约定有两根支柱：一个稳定的计数单位，和一套稳定的分辨类别。第七章拆掉前一根——生成成本趋零之后，“一件”这个单位失去了下界，于是“件数上升”不再能充当“投入上升”的证据，而几乎所有关于内卷的讨论都建立在这个证据上。第八章拆掉后一根——可分辨的种类数不是系统的自然属性，它由观察者的分辨集合与系统状态共同决定，因此它是被生产出来的；而没有任何人把它当作一项产出来生产。

这一编因此是全书最不舒服的两章：它们不是在指出某个数算错了，是在指出那个数所依赖的单位已经不存在了。

第七章

一件是几件——计数单位失去下界

■ 摘要

产生一件东西的代价降到接近于零之后，几乎每个行当都在报告同一组读数：交出来的件数暴涨，人均件数暴涨，而通过的人没有变多。这组读数被普遍读作“投入增加而产出不增”。本章指出，这个读法有一个从未被检验的前提：它假定“一件”是一个稳定的计数单位，因而件数的上升可以被当作投入上升的证据。

三个互不往来的学科各自持有一套关于“是几件”的成熟读法，而三套不能同时为真。计量与指数理论说，只要把产出按质量调整后加总，就能读出到底多了没有，件数不重要；当代艺术的本体研究说，一件作品是不是“一件”，由创作者可公开查证的授权行为决定；流动与批量理论说，一个变更是不是“一个”，由交易成本与持有成本的权衡决定，不存在独立于流水条件的答案。三家争的是“是几件”该由哪一维读出，而它们共同假定了：这是一个可以被单独读出的事实。

推翻这条假定的材料来自三家自己：指数理论承认，品种更替足够快时，被比较的那一组东西本身是流动的，不存在跨期一致的“同一样商品”；批量理论承认，交易成本被自动化压掉之后最优批量随之变小——顺着这条推到底，“一个”没有下界；作品本体研究承认，许多作品在保管者介入之前于关键方面是不定的，件数由一句授权决定。

由此得到一条三家都到不了的判断：这一轮的“更卷”，不是产出量增加，不是作品数增加，也不是工作项增加，而是“一件”这个计数单位失去了下界——件数与“把全部提交合并同类项之后还剩几项”之间裂开了一道口子，而后者从来没有被任何人生产过。本章把后者命名为并项，给出一句只问记录不问人的问话（把这一轮的全部提交合并同类项，还剩几项）、一个三领域量纲不同而比率相同的读数（并项率）、两轴四格与其中最危险的一格（件数上升而并项不增）、十三处兑现、三处反过来削弱本章的约束、与十种既有说法的判决性分界、八条排除了最强竞争解释的判错条件，以及一条写死到二〇二九年十二月三十一日的赌注。

■ 关键词

并项；并项率；计数单位；合并判准；按件计绩；生成成本趋零

HowManyIsOne

一件是几件

wager.

■ 一、导论：三个不能同时为真的读法

先把这一轮的读数摆出来，一个行当一组，全部来自各行当自己的记录系统。

一个招聘岗位平均收到二百多份申请，每位招聘专员手上的申请量在几年内翻了几番，而录用人数没有相应变化。一家管理学期刊的编辑团队清点五年收件，投稿量比写作工具普及之前高出四成，桌面直接退稿的比例从一成三涨到四成。一家覆盖近五千名技术从业者的交付研究报告，九成开发者每天使用生成工具，交付吞吐上升而不稳定性同步上升。学术评价那一侧，某学者一年发表的篇数创下自己的纪录，而他自己说不清今年到底比去年多想清楚了哪一件事。

这四组读数被读成同一句话：投入增加了，产出没有相应增加。

这句话在中文里有一个现成的名字，而

这个名字自带一整套处方——限投、限件、提高门槛、增加评审人手、把奖品做小。

本章要问的不是处方对不对，而是那句话本身有没有被检验过。它的完整形式是：件数上升了，而件数上升是投入上升的证据。后半句是一个前提，不是一个观察。它要求"一件"是一个稳定的计数单位。

一件是几件

而恰恰在这一轮，这个前提第一次成了问题。产生一份看起来完整的东西的代价降到接近于零之后，把一份工作交成一件还是交成十件，从一个技术上受限的事实，变成了一个几乎无成本的选择。于是"件数"这个分母不再是给定的，它落到了被计数者自己手里。

一件东西凭什么算一件？三个互不往来的学科各自有一套成熟的回答，而三套回答不能同时为真。

第一套来自计量与指数理论。

它说，别管件数，把产出按质量调整之后加总，就知道到底多了没有。一

台今天的计算机不是三十年前那台的"一件"，它是若干倍。同理，十份内容重合的申请在福利意义上不等于十份。这一套要求：只要结果这一维读得准，件数是无关的。

第二套来自当代艺术的本体研究。

它说，一件作品是不是"一件"，既不由材料决定，也不由外观决定，

而由创作者可以被公开查证的授权行为决定——他做了什么、说了什么、写了什么、对保管者交代了什么。一堆糖果可以是一件作品，一件作品可以每次展出都换掉全部材料。这一套要求：只要那个完成动作读得准，件数就定了。

第三套来自流动与批量理论。

它说，一个变更是不是"一个"，取决于把它单独走一遍要付多少交易成

本、攒着不走要付多少持有成本；最优批量是这两条曲线加起来的最低点。这一套要求：只要流水条件读得准，件数自然定下来，而且它本来就随条件变。

三套读法的处方，在同一个真实处境里会分岔到互相取消的方向。取一位学者今年交出的十二篇论文：按第一套，该问的是这十二篇合起来在质量调整后相当于多少——如果每篇的实质增量很小，那么"产出没增加"这个结论成立，处方是改用质量加权的评价。按第二套，该问的是作者本人是否把它们当作十二件独立的完成——如果他确实每一篇都作了一次独立的完成动作，那它们就是十二件，处方是承认它们并检验各自的质量。按第三套，该问的是这个领域的发表交易成本是不是降了——如果降了，切成十二篇不但不是问题，而且是正确的，处方是接受更小的批量，并把评审也相应做小做快。

三条不能同时执行。更要紧的是，三条对同一个事实给出了三种不同的计数，而没有任何一套能说服另外两套——因为它们根本不在同一个语境里作答。

本章认为，三套读法的分歧不是需要裁决的争论，它是一个提示："是几件"可能根本不是一个可以被读出的事实。

■ 二、文献综述

2.1 计量与指数理论：把结果读准就够了现代产出与价格计量的核心难题不是数不清楚，而是跨期的同一性。一九六一年格里利谢斯用回归把汽车的价格分解到它的特征上，开出了享乐价格法这条路；其后半个世纪的工作都在处理同一件事：当被比较的东西本身在变，"同一样商品的价格"这句话还成不成立。

一件是几件

一九九六年那份关于消费价格指数的委员会报告，把这件事推到了政策层面。它指出该指数系统性高估通胀，四项偏差中最大的一项正是质量变化与新品种：新东西进入市场时不在篮子里，进来之后价格常常快速下降，而这一段被漏掉了。报告之后，统计机构缩短了篮子更新与样本轮换的周期。

这一套的立场因此很清楚：件数不是产出，价值量才是。

十份内容重合的申请，在这套读法里根本不是

十件，而是一件加九份重复的成本。它有权这样说，因为它的任务本来就是把"多了没有"这个问题从计数里解放出来。

而它自己知道一件与主线不合的事，本章第五节要用它。

2.2 作品的个体化与创作者的授权：把完成动作读准就够了当代艺术给本体研究出了一道难题：一件由糖果堆成的作品，观众可以拿走，馆方每天补满；一件装置每次布展都由保管者按规则重新实现；一件作品的材料会自然腐坏，而腐坏本身是它的一部分。这些东西的边界在哪里、哪些改变会毁掉它、它到底是不是"同一件"，都不能从物件本身读出。

二〇〇五年伊尔文那篇论文给出的答案是创作者的授权：不是他心里想什么，而是他可以被公开查证地做了什么、说了什么、与保管者交了什么。她把这个概念一路推到当代与历史作品，并在二〇二二年的专著里进一步主张，规则本身已经成为一种艺术媒介——作品带着自己的自定规则，规则规定了它由什么构成、可以被怎样对待、什么改变会终结它。

这一套的立场同样清楚：一件之所以是一件，因为有人把它作成了一件。

件数是一个行为的结果，不是

一个物理事实。

而它自己也知道一件与主线不合的事，本章第五节要用它。

2.3 流动与批量理论：把流水条件读准就够了工程学这一侧最锋利的框架，不是"越小越好"这类信条，而是把批量当作一个经济权衡来算。莱纳森二〇〇九年那本书里，批量经济学原理写得极干脆：经济批量是一个U形优化——交易成本随批量变小而上升，持有成本随批量变大而上升，最优点在两者之和的最低处。他同时提醒，这条曲线底部很平，所以不必算得多准；他还写了一条更狠的：批量减小带来的好处，比人们以为的多得多。

这一套的立场是："一个"不是一个属性，是一个由条件定下来的分割。

同一份工作，在交易成本高的环

境里应当攒成一批走，在交易成本低的环境里应当拆开逐个走；两种做法在这套理论里都不是失真，都是正确。

而这一套自己写下的一句话，是本章第五节要用的第三件材料。

一件是几件

■ 三、冲突定位：三条读法不能同真

把三家最承重的那一条并排放：· 计量与指数理论：看结果这一维就够了——按质量调整后加总，件数无关。

· 作品个性化理论：看完成动作这一维就够了——授权规定了边界，件数由此而定。

· 流动与批量理论：看条件这一维就够了——批量由成本权衡决定，件数随条件而变才是对的。

它们不是三个角度的互补，它们互相取消。

第一套若成立，第二套的授权就是无关的：一位作者宣布这十二篇是十二次独立完成，改变不了它们在质量调整后只相当于两三件这个事实。第二套若成立，第一套的加总是僭越的：把作者明确作成十二件的东西并成两件，是外人在替他重新划定作品边界，这在艺术里是被严格禁止的动作。第三套若成立，前两套都问错了问题：件数根本不该被当作有待读准的事实，它是一个应当随条件调整的决策变量，而在交易成本极低的环境里，切得越小越对。

而更硬的一处是：三家各自撑不住的地方，恰好是另一家的主张。

计量理论承认，当品种更替太快时它

加总不了——那正是第三套所说的条件变化。作品理论承认，在保管者介入之前作品在关键方面是不定的——那正是第三套所说的条件参与了定义。批量理论承认，最优批量的存在依赖于交易成本有一个正的下限——而那个下限是不是正的，正是第一套要在结果这一维上验的事。

三条互相顶住，谁也判不了谁。这不是缺一次裁决，这是提示：它们共同站在一件没有说出口的事上。

■ 四、三家共同的默认：件数是一个可以被读出的事实

三家争的是"是几件"该由哪一维读出。而它们共同假定了：给定一批东西，问"这是几件"有唯一答案，只是各自认为答案该由谁给。

把这句话念给三家听，三家都会说这还用说吗。这正是它是共同前提的标志。

这个前提有一个直接后果，本章认为是整件事的关键：所有以"件"为分母的读数，都被默认为可解释的。

人均件数、每岗申请数、每人年发表数、通过率、录取率、每千行缺陷数——这些量的分子各不相同，分母是同一个东西。一旦分母不是一个稳定单位，它们同时失去意义，而且失去意义的方式在报表上完全看不出来：数字照样算得出来，照样逐年可比，照样进入考核。

三家为什么不会问这件事，第九节交代。这里先指出一处：三家中没有任何一家的任务是清点。计量理论要跨期可比，作品理论要裁定边界，批量理论要优化流动。清点是别人的事

——而"别人"在学术分工上不存在。件数是各行各业的行政系统在数，而行政系统从不问它数的那个单位是不是稳定的，因为在它的历史经验里，那个单位一直是稳定的：一份申请要花几小时写，一篇论文要花几个月做，一个变更要人手动改。

这一轮第一次不是这样。

一件是几件

■ 五、推翻这个默认的材料，三家自己各有一件

要推翻一个共同前提，从外面搬理由来是没用的——那只是加入了争论，成了第四家。真能取消争论的材料，必须是三家自己已经掌握、已经发表、只是没往这个方向用的事实。三家各有一件，而且都很有名。

5.1 流动的项目全域：计量理论自己的难题指数理论有一个自己写下的、从未解决的难题：被比较的那一组东西本身是流动的。

教科书的说法是，

质量变化指的是新款替换旧款、新品种指的是全新类型进入市场，但两者的病根相同——要比较的项目全域在两个时点上不是同一组。近年扫描数据普及之后这个问题更急：每一个商品编码的进进出出都逐笔可见，而"同一样商品"这个概念反而更难维持了。

计量理论拿这条答的是"通胀被高估了多少"。它从来不拿它当这样一条证据：如果不存在跨期一致的"同一样东西"，那么"件"也就不存在一个跨期一致的意思。而这正是这条材料在本章里的用处。

5.2 最优批量没有下界：批量理论自己的推论莱纳森在解释批量原理时举过一个例子：测试过程自动化把测试的交易成本压了下去，因而使小得多的批量成为可能。这句话在他那里是一条实践建议。

顺着它自己的曲线推到底，得到的是一个更强的结论：最优批量的存在，依赖于交易成本有一个正的下限。交易成本趋零时，U形曲线的最低点向零滑动，"一个"这个单位在他自己的理论里没有下界。

批量理论不往这个方向用它，是因为它的目标是流动效率，而在流动效率上，批量变小一路都是好事。"一个变得没有下限"对它不是问题，是成功。

5.3 介入之前是不定的：作品理论自己的承认伊尔文的框架里有一处被评论者反复讨论的地方：那些需要保管者按规则重新实现的作品，在保管者介入之前，于关键方面是不定的。也就是说，作品的边界不是由创作者单方面给定、然后被别人执行；它是在创作者的授权与另一方的实现动作之间被共同定下来的。

这一条在艺术里回答的是"谁参与了作品的构成"。它从来没有被拿来当作这样一条证据：同一堆材料算几件，可以由一句话规定，而且规定它的可以不止一个人。

一旦承认这一点，"件数"就不再是被读出的，
是被约定的；而约定可以变，也可以被人有意选择。

三件材料指向同一件事："一件"不是一个事实，它是一个在特定条件下曾经很稳定的约定；而这一轮，那些条件塌了。

一件是几件

■ 六、并项

6.1 定义并项：把一批提交按内容合并同类项之后，剩下的那一个单位。

需要立刻说清三件事，否则它会被读成一个旧概念。

其一，合并是一次生产，不是一次还原。

常识里的"切分"预设有一个整体被切开了，因此合并只是把碎

片拼回去。本章说的不是这个。在生成成本趋零的条件下，那个整体常常从来没有被生产过——十二篇论文不是从一篇整的里面切出来的，它们是各自直接被写出来的，作者手里从来没有过那"一篇整的"。因此把它们合并起来所得到的那一项，比作者手上已有的任何东西都多，它需要新的判断、新的取舍、新的写作。合并要花钱，而且没有人有动机花。

其二，并项不归任何人。

十二篇论文各有作者、各有归属、各有引用；那"一项"没有署名，不进任何人

的履历，不构成任何机构的产出。这正是它从来没有被生产过的原因——不是难，是没有人得利。

其三，并项数依赖一个合并判准。

"按什么算同类"必须先被写死，否则这个数不可比。本章承认这是它

最大的限制，第十一节把它当作一处真约束交代出来。

6.2 两轴四格第一轴：这一轮交出来的件数，比上一轮增加还是不减。第二轴：这一轮的并项数，比上一轮增加还是不减。

并项数增加

并项数不减

件数增甲格：真的多做了。三家都读得出，三套读法在这一丙格：靶格。所有指标显示繁荣，实质加件数不减

格里不冲突。

没有增加。

乙格：做得更整了。罕见，且没有任何制度奖励它。

丁格：停滞。人人都看得见，也人人都会承认。

要打的是丙格。它满足一个真靶格的三条签名。

其一，它在短期指标上表现为改善。

产量上升、活跃度上升、交付频率上升、发表数上升、响应速度上

升。而且每一件都是真的——每一份申请都由本人提交，每一篇论文都有独立的数据与结论，每一个变更都真的改了代码。全部形式审查都能通过，没有任何人违规。这一点必须说透：丙格不是造假，造假是另一回事，而且造假会被查出来。

其二，它的失效不产生任何信号。

没有任何系统计算并项数。因此丙格与甲格在所有报表上一模一样：

两者的件数都在涨，而唯一能把它们分开的那个数没有被算过。一个机构在丙格里待了五年，它的全部指标都会显示它这五年在进步。

一件是几件

其三，它自我加固。

越是推行小批量、快速迭代、持续交付、按件计绩这些正规化的好实践，切得越

小、件数越多、丙格越深。这一格不是由懒惰或失职造成的，它是由正确的管理造成的。

6.3 一句不含判断词的问话判据可以拿去问记录，不必问任何人的看法：>把这一轮的全部提交按内容合并同类项，还剩几项？

在五个具体场合跑一遍，五个答案互不相同，其中至少两个必须去查才知道：其一，代码仓库。

一个季度里合并进主干的若干变更，按"改动了同一个可观察行为"合并，剩几个？数据现成：每一次提交的差异都在。

其二，一位学者的五年。

按"一条独立的实质主张"合并，剩几条？这个数在多数人的简历上从未出现过。

其三，一个人投出一批申请。

按"不同的实质主张"合并，剩几份？答案往往是一——同一份履历的若干个变体。这一条最刺眼，也最容易验证。

其四，一个专利族。

答案是已经有人在数了：专利统计早就以"族"而不是以件为单位做国际比较。这一格给出的不是支持，是一条必须处理的分界，见 6.5。

其五，一位画家一年的系列作品。

答案是反的：在这里，同一主张的若干变体本身就是作品，合并会毁

掉它。系列的意义恰恰在于并置与差异。这一条是查出来的，不是推出来的，而且它直接给出本章的适用边界，见第十三节。

6.4 并项率把上面那句问话变成一个可比较的数： $\text{并项率} = \text{并项数} \div \text{提交件数}$ 取值在零到一之间，越低表示件数相对于实质越虚。它有四条性质，正是它可以跨领域使用的原因。

无量纲。

代码那里的单位是变更，论文那里是篇，招聘那里是份，专利那里是件；量纲各不相同，而这

个比率可以并排比较。

可事后清点。

它不需要新建数据。所有这些系统都保存着全部提交的内容，需要的只是一次按固定判准的内容合并。

它把一个印象变成一个数。"现在东西又多又碎"是一句感慨；"本季度并项率零点一七"是一个可以争论、可以复核、可以被推翻的读数。

它与既有的主要指标正交。

这一条最容易漏，也最要命。既有主要指标是产量、频率、周期、通过率。

举一个"既有指标最好看而这个数最难看"的真实例子：一个部署频率最高、变更前置时间最短、变更失败

一件是几件

率最低的团队，它的并项率可以是全公司最低的——恰恰因为把变更切到极小正是这套精英实践本身所要求的。反过来，一个一季度只交三次的团队，并项率可以接近一。举不出这样的例子，这个读数就该重造；这个例子是现成的。

6.5 与专利族的分界：为什么这不是已经有了的东西专利统计里早有"族"这个单位：同一项发明在多个国家申请，算作一族，国际比较以族为准。这看上去正是本章说的并项，而且已经制度化了几十年。

分界只有一处，而这一处是干净的：专利族按行政事实合并，不按内容合并。

同族的判据是共享优先权

——它们是同一份原始申请派生出来的。因此专利族解决的是"同一件东西在多国重复登记"的重复计数，它无法把两份没有共同优先权、而技术主张实质相同的申请并成一项，也无法把一份申请里塞进的三个不相干主张拆成三项。

一句话：专利族数的是同一次申请的多次登记，并项数的是不同次提交的同一个主张。

前者是登记冗

余，后者是内容冗余。二者的相关性可以直接测：若在某个技术领域，并项率与族规模高度相关，则本章的读数被专利族取代；若正交（本章预测正交，因为工具产出的是无共同优先权而主张重合的申请），则本章对。

这条分界同时是一个好消息：它说明"以合并后的单位计数"在制度上是可实现的，不是空想——已经有一个行当做了几十年，只是它合并的判准不是内容。

6.6 三个算例抽象的比率容易被读成修辞。下面三个算例用具体数字走一遍，判准先写死，再算。

算例一·一位学者的一年。判准：两篇若其核心结论在替换掉数据集与领域名词之后仍然是同一句话，则并为一项。设某人本年度发表十二篇：其中四篇是同一个方法在四个不同数据集上的应用，结论句式相同；三篇是同一条理论主张在三个子领域的重述；两篇是同一份综述拆成的上下篇；另有三篇各自独立。

按判准合并：四并一、三并一、二并一、三保留，并项数为六，并项率零点五。

关键在于这个数说了什么、没说什么。它没有说那十二篇里有九篇是废的——四个数据集上的验证是有价值的，上下篇也是完整的。它只说一件事：如果把这一年按"独立的实质主张"计数，这个人今年是六项，不是十二项。而他的考核表、他的年度总结、他所在学科的产出统计，全部按十二计。两个数字都对，只是一个提交次数，一个是主张数——而"投入增加而产出不增"这个诊断需要的是后者，用的却是前者。

算例二·一个季度的代码变更。判准：两个变更若改动了同一个可观察行为、且合并之后不损失任何回滚粒度，则并为一项。设某团队本季度合并了三百二十个变更：其中一百四十个是同一次重构被拆成的连续小步（每一步都能独立通过测试，这正是好实践所要求的），按判准并为十二项；六十个是同一批依赖升级的逐包提交，并为四项；四十个是同一个配置项在不同环境的重复应用，并为三项；其余八十个各自独立。并项数为九十九，并项率约零点三一。

一件是几件

同一季度，这个团队的部署频率、变更前置时间、变更失败率三项指标全部位于业内最优区间——而且这三项指标之所以好看，恰恰因为那一百四十步被拆开了：每一步都小、都快、

都不容易失败。第 6.4 节所说的正交，在这个算例里是同一批数据的两种读法。这里必须再说一遍第十一节第一处约束：那一百四十步拆开是对的，本章不主张把它们合起来做；本章只主张，当有人拿“本季度三百二十个变更”去证明产能提升时，那个数字是三百二十还是九十九，取决于用哪一个单位，而不取决于这个团队干了多少活。

算例三·一个岗位的申请池。判准：两份申请若在替换掉措辞与排版之后，其主张的经历、能力与意图集合相同，则并为一项。设某岗位收到二百四十份申请：其中一百八十份来自六十位申请人（人均三份，同一履历的不同措辞版本），按判准并为六十项；四十份来自四十位申请人各一份；另有二十份是同一批自动投递工具产出的模板变体，主张集合几乎相同，并为二项。并项数为一百零二，并项率约零点四三。

这个算例给出一处最需要被看见的错位：招聘方读到的是“二百四十份申请”，据此判断这个岗位竞争激烈、门槛应当提高；而实际在竞争的是一百零二个主张。两个数字相差一倍以上，而没有任何一个招聘系统计算后者。第十节会指出，正因为如此，“每岗申请数”这个被反复引用的拥挤指标，其分子是提交次数，既不是申请人数，也不是主张数。

三个算例的判准各不相同，这不是缺陷，是这个读数的使用条件：判准由领域自己定，定了之后写死、公开、不许因为数难看而中途改。

三处的量纲不同（篇、变更、份），比率可以并排：零点五、零点三

一、零点四三。

■ 七、这一件为什么恰好是现在裂开

单位的稳定性从来不是被规定的，是被成本撑住的。

历史上“一件”之所以稳定，靠的是三道自然闸门。其一，做出一件像样的东西要花可观的时间，所以件数与投入之间有一个大致固定的换算。其二，把一份工作切成两份要额外付一次组织成本——重新写引言、重新做背景、重新过一遍格式，所以切分本身有代价。其三，切开之后两份都必须各自站得住，而站得住需要真材料。

这一轮，三道闸门同时被压平了。生成的时间成本降到接近零，第一道没了；把一份工作重写成两份完整成品几乎不额外花钱，第二道没了；“看起来站得住”与“站得住”在形式审查这一层不可分，第三道被推到了下游。

于是发生的事情，用批量理论自己的话说最准确：交易成本塌了，最优批量随之向零滑动。

而这在批量

理论里是好事，在指数理论里是无法加总的麻烦，在作品理论里是创作者的自由——三家都没有理由把它当成一个问题，因为它在三家各自的问题里都不是问题。

它只在一个地方是问题，而那个地方不属于任何一家：所有以件为分母的读数，在同一时刻一起失真了。

这里要补一句关于时间的话，因为它决定了这条判断能不能被查。三道闸门不是同时被压平的，而是先后被压平的，且顺序是可以在记录里读出来的。第二道最先松：排版、投递、格式、模板这些切分的组织成

一件是几件

本，在文档工具与在线投递系统普及时就已经大幅下降，那是二十多年前的事情，而当时件数并没有暴涨——因为第一道还在，做一份像样的东西仍然要花时间。第三道其次松：形式审查逐步标准化、自动化之后，“看起来站得住”这一层被机器代劳，人只在下游才看得到差别。第一道最后塌，而且塌得最快，就在最近这三四年。

三道闸门的这个先后，给出本章最容易被推翻的一条读数：并项率的下降，应当在第一道闸门塌之后才显著加速，而不是随第二道或第三道同步下降。

也就是说，如果去查一个行当二十年的记录，应当看到并

项率在前十几年缓慢下滑、近三四年陡降。若查出来的是一条平缓的长期下降线，那么本章把这一轮当作一个新现象就是错的——那只是一个持续了二十年的老趋势，不需要新框架。这条检验的数据在多数行当的历史归档里已经存在。

必须说明的一点是：这不是说件数上升全是虚的。真新东西也在增加，这一点第十一节会作为一处约束交代。本章主张的只是：件数上升这一件事，从此不能再单独作为投入上升或产出上升的证据，而“内卷”这个诊断恰恰完全建立在它之上。

■ 八、十三个行当

一条判断如果只在提出它的那个场景里成立，那是个例。下面是逐项兑现，其中两处反过来给本章加了约束，第十一节交代。

其一，学术发表。

一位学者一年十二篇，并项后剩两三条主张，是这一格最常见的形状。可查的读数：

把某学科某年度的全部论文按“独立的实质结论”合并，比较合并前后的分布；再看该学科的评价办法是否按件计。

其二，同行评审。

评审工作量按件计算，而件数被投稿方控制。于是评审系统的负荷由被评审方的切分决定，评审方对此没有任何调节手段。这一格解释了一个反直觉的现象：加派评审人手不会缓解，因为分母在对面手里。

更值得注意的是它的方向：投稿方并不需要有任何恶意，甚至不需要意识到自己在切——把一项工作写成两篇，在今天既不比写成一篇更费事，也不比它更不诚实，而两篇在履历上确实比一篇多。于是一个完全由诚实行动者组成的系统，会稳定地把评审负荷推向无穷，而没有任何一个人可以被指出做错了什么。可查的读数：某刊某年度的送审总量与该年度合并后的独立主张数，二者的比值逐年变化。

其三，软件交付。

变更数、部署频率、前置时间这三个核心指标全部以变更为单位。并项率下降时，三个指标同时改善而实质产能不变。可查的读数：把一个季度的变更按"改动了同一个可观察行为"合并，看合并后的数量增长曲线与合并前是否分岔。

其四，招聘。

一个人投出的若干份申请合并后常常剩一份。因此"每岗申请数"这个被反复引用的数字，其分子是提交次数而不是申请人，也不是实质主张数。这一格给出本章最容易做的一次验证，而它的后果比读数本身重：门槛是按拥挤程度调的，拥挤程度是按件数读的，于是门槛被一个虚高的分子推着往上走，而被这个门槛挡住的是那些只投一份、投得认真的人——他们贡献的件数最少，因而对拥挤的读数贡

一件是几件

献最小，却与那些投三十份的人承受同一条被推高了的线。这一格里，计数单位的失真直接转化为分配上的不公，而且它在任何一份合规审查里都不留痕。

其五，专利。见 6.5。这一格的价值在于它已经有一个部分实现的并项单位，可以直接对照。

其六，标准与规范制定。

提案数暴涨而实质分歧数不变，是这一格的形状。合并后剩下的分歧数，是这个领域真正的工作量。

其七，法律文书与诉讼。

动议数、书状页数以件计，而实质争点数不随之增长。这一格里已经有制度化的并项动作：法官归纳争点。这是本章所说的合并现实中被制度化的第二个例子，而它同样不是免费的——归纳争点是法官的工作，且不计入任何产出统计。

其八，专业资格与继续教育。

学时、课程数、证书数以件计。合并后剩下的独立能力项，通常远少于证书数。

其九，市场推广与内容生产。

每周发布条数是核心指标，而合并后的独立主张数往往是个位数。这一格

里并项率下降得最快，因为切分几乎没有下游代价——没有评审、没有验收、没有人会因为两条内容主张相同而退回其中一条。没有下游的场合，并项率会一路跌到接近零，而所有以条计的指标会同时创下历史新高。这一格因此可以当作其他行当的先行指标：它是这条机制在没有任何阻力时的极限形态。

其十，风险投资与项目申报。

申报件数暴涨，合并后的独立方案数不变。这一格的一个可查读数：同一支基金收到的两百份计划书里，按"同一个市场假设加同一条路径"合并，剩几个。

其十一，缺陷与工单管理。

报障单数、缺陷密度以件计。同一个根因被拆成十几张单，会让缺陷数曲线极难解读。这一格里已经有一个现成的对照概念：根因分析。而根因分析恰恰是本章所说的合并，只是它只在事故之后被做一次。

其十二，教育中的作业与评价。

提交次数、练习量以件计。这一格的答案是部分反的：重复练习本身有价值，合并会毁掉它。见第十三节。

其十三，安全事故与不良事件报告。

这一格是伦理边界所在，第十五节专门处理。此处只指出：这里绝

不能为了让并项率好看而合并——安全领域的逐件保留是有理由的，一个被并掉的事件报告可能正是下一次事故的唯一线索。

十三个行当里，第十二与第十三反过来削弱了本章，连同第十一节的第三处，一并交代。

■ 九、三家为什么各自到不了

写"三家为什么到不了"，不是写三家的缺陷。一个领域看不见什么，通常不是它的疏忽，是它自己的问题结构的影子。

计量与指数理论到不了，是因为它的产品就是跨期可比。

它必须假定存在某种可以被链式连接起来的同

一性，否则指数无法被构造。承认单位没有下界，等于承认它的核心产品在原理上无法生产。所以它把这

一件是几件

件事一直安放在"偏差"这个位置上——偏差是可以估计、可以修正的东西，而"单位不存在"不是。把它叫作偏差，正是不去看它的方式。

作品个体化理论到不了，是因为它总是从"这里有一件东西"出发。

它的任务是为一件已经存在、已经进

入收藏或展览系统的作品裁定边界。"如果没有人作那个完成动作，还剩几件"不是它的问题，因为在它那里，没有完成动作就没有作品可谈。而且它有很强的规范理由不去问：替创作者重新划定件数，在这个领域是被禁止的动作。它的职业伦理恰好挡住了这个问题。

流动与批量理论到不了，是因为在它的目标函数里，件变小是成功。

它衡量的是流动效率，而流动效率

随批量减小而改善。要它承认切小会污染以件为分母的读数，等于要它承认自己最成功的实践制造了一场度量危机。它没有理由这样做，因为度量危机不在它的目标函数里——它数的是流过去了多少，不是流过去的是几样东西。

三家都不是没看见，是各自的问题恰好使这一步成为不必要的。

还剩一个问题：为什么没有第四家来做这件事？

清点看上去是一件基础到不能再基础的工作，按理该有

一门学问专管它。而实际上没有，原因不是疏忽，是分工的形状。

在学术分工里，"数东西"从来不是一个独立的题目，它总是附着在别的题目上：统计学管的是从样本推总体，前提是观测单位已经定好；计量经济学管的是识别因果，前提是单位已经定好；文献计量学管的是引用与产出的分布，而它的单位——篇——是出版制度给的，不是它定的；测量学管的是量具的可靠性与效度，处理的是同一个单位下的误差，不处理单位本身有没有下界。每一门都在单位定好之后开始工作。

而在行政分工里，件数是被数的那一方之外的人在数，且数它的人没有权限过问它是怎么被切出来的。教务处数论文篇数，不问一项研究该写成几篇；人力部门数申请份数，不问一个人该投几份；工程管理平台数变更条数，不问一次重构该拆成几步。数的人管不着切，切的人不管数。

这个分工在三道闸门都还立

着的年代运转得很好，因为那时切分本身有成本，切的人自己会克制。

于是这件事落在一个没有主人的位置上：它不在任何一门学问的问题清单里，也不在任何
一个岗位的职责表里。

这既是它至今没有被辨识的原因，也是它一旦被辨识就相当容易被处理的原因——不需要
新技

术，也不需要新权力，只需要有人开始算。

■ 十、与最近的几种说法的分界

一条判断能不能站住，看的不是它有没有道理，而是它能不能与最像它的那些说法在同一
个案例上给出相反的判决。下面每一条都写成同一个形状。

最小可发表单元与切香肠。

它说到哪一步：为了增加篇数，把一项研究拆成若干篇最小可发表的片段；

这被普遍视为学术不端或至少是不良习气。这是最强的竞争解释，也就是那句最难反驳
的"其实不就是切香肠吗"。分界：切香肠预设有一个可还原的整体——正因为存在那个整体，才
谈得上"某人把它切了"，也才谈得上追究。本章说的是，在生成成本趋零之后，那个整体常常
从来没有被生产过。判定形式：取一

一件是几件

批件，要求作者提交合并版；若合并版的实质内容不超出各件之和、只是拼接，则是切香
肠；若合并版需要新增判断与取舍才成篇（即合并是一次生产），则本章对。

这条检验成本很低，而且它同时解释了为

什么反切香肠的规训一直不管用——它追究的是一个不存在的行为。

内卷化。

它说到哪一步：投入不断增加、内部不断精细化，而结构不变、无发展。分界：它的前件
是要

素固定，而这一轮的定義性特征是产能暴涨；更要紧的是，它讲的是分子（投入真的增加
了），本章讲的是分母（件数这个计量单位塌了）。判定形式：若单位投入的实际时间随件数
同比上升，则是内卷化；若件数上升而人均实际投入时间不变甚至下降，则本章对。这个数在
多数机构的工时记录里已经存在。

古德哈特定律与坎贝尔定律。

它说到哪一步：一个指标一旦成为目标，就不再是好指标。分界：它讲的

是指标被优化，本章讲的是计量单位失去下界。二者可以分开：判定形式——在一个完全不以件计绩的通道里（不考核篇数、不考核提交数），若并项率照样下降，则古德哈特不足以解释，本章对；若并项率在这类通道里保持稳定，则这只是指标被优化的一个新例子。

本章预测前者，理由是切小是工具的默认输

出粒度，不需要任何人被激励。

成功陷阱、能力刚性、伊卡洛斯悖论。

反转模板必须专查：把本章的命题剥掉本领域的名词之后，形式

是"计量单位随条件变化而失去下界，使一切以它为分母的比率同时失真"——这不是"越成功越失败"。第6.4节那个正交实例（精英团队并项率最低）带着反转的味道，但它说的不是那个团队的能力下降了，而是读数的意义下降了。判定形式：若这类团队的实际交付能力（按合并后的行为变更计）随并项率下降而下降，则是成功陷阱；若实际能力不变而只有以件计的读数改善，则本章对。

分母问题与内生分母。

剥到形式层之后，最贴的上游命名不在本领域，而在流行病学与金融的计量讨论

里：当分母本身被内生地改变时，比率不可解释。

这一位必须点名，它是本章最近的形式亲戚。分界：

那一支处理的是分母被错误地选取或统计——选错了人群、漏算了暴露时间、口径不一致；本章处理的是分母在原理上不再有唯一选法，不是选错，是不存在正确的选法。判定形式：若换用一个更好的分母口径能使读数恢复稳定，则是分母选取问题；若任何以件为基础的口径都随切分而变，则本章对。

压缩率与近似重复检测。

这是方法上最危险的一位。近年评估生成文本多样性的标准做法之一，就是把

一批输出拼起来做无损压缩，压得越狠说明重复越多、多样性越低；这类读数已被大量工作采用。分界：压缩率测的是符号层的冗余，并项测的是主张层的同一。而这一轮的关键事实恰恰是：工具极擅长把同一个主张写成措辞完全不同的若干份。判定形式：若并项率与压缩率高度相关，则本章的读数应被压缩率取代（它更便宜）；若二者正交——一批压缩率显示高多样性而并项数为一——则本章对。

本章预测正

交，而这个预测在一天之内就可以被证伪。这一条是本章最愿意被人拿去打的。

模块化与设计规则。

它说到哪一步：把一个系统切成模块并定义接口，会创造真实的期权价值，切分是

好的。这是来自本章之外的一支，必须处理，因为它给"切小"提供了最强的正当性。分界：模块化要求接口被定义，切出来的部分之间有明确的契约，因而可以独立演化、独立替换；本章说的是无接口切分——切出来的件之间不是接口关系，是内容重叠关系。判定形式：若切出来的件之间存在可指认的接口，且任一件可以被替换而不影响其余，则是模块化；若它们只是内容重叠、合并后不损失任何信息，则本章对。

一件是几件

交易成本与组织边界。

它说到哪一步：一样东西该由市场做还是由组织做、该算一次交易还是若干次，

由交易成本比较决定。分界：它比较的是两种组织形态之间的边界，边界两侧有真实可比的成本；本章说的是同一主体内部同一份工作的切分数，两侧没有可比的成本差别，切与不切在成本上几乎相同。判定形式：若切分数随组织形态变化而变化，则是边界问题；若在同一组织形态内、同一批人手上，切分数随工具可得性而变化，则本章对。

新品种偏差与享乐指数。

它既是本章推翻共同前提的材料来源，也是最容易被误认的一位。分界：它关

心的是价格与福利的跨期可比，本章关心的是件数作为分母的当期可解释。二者的处方也不同：它的处方是缩短篮子更新周期，本章的处方是换一个计数单位。判定形式：若把评价改为质量加权而件数口径不变，读数即恢复稳定，则该支足以处理；若质量加权之后按件计的读数仍随切分而变，则本章对。

去重与实体解析。

数据管理里有一整套成熟方法处理"同一样东西被记了多次"：精确去重、近似重复检

测、实体解析（把散在多处记录归到同一个实体上）。这一位在方法上离本章很近，必须点名。分界有两处。其一，去重与实体解析处理的是记录的重复，前提是被记录的那个东西本身是一个已经存在的实体——同一个人、同一件商品、同一篇文献；本章处理的是主张的重合，而那个"主张"在合并之前并不作为一个实体存在于任何地方。其二，去重的目标是恢复：把重复的记录删掉，剩下的就是真实情况；本章所说的合并是生产：删掉重复的件之后剩下的不是并项，并项需要被重新写出来。判定形式：把一批件做实体解析，若解析后的实体数与按内容合并后的项数相等，则本章可被现成方法取代；若两者显著不等——典型情形是实体解析给出的实体数等于件数（每一件确实是一个不同的提交、一个不同的文档、一个不同的作者组合），而并项数远小于它——则本章对。

本章预测后者，理由是这一轮产出的件在记录层面是

货真价实地互不相同的。

划完之后还有一层可以取。逐个问上面这些说法"它结构性地看不见什么"：切香肠把"存在一个可还原的整体"当作给定，因此看不见整体从未被生产的情形；古德哈特把"件数是可靠计量、只是被优化"当作给定；模块化把"接口定义了单位"当作给定；压缩率把"符号冗余可以代表内容重复"当作给定；交易成本把"边界处有真实成本比较"当作给定；分母问题把"存在一个正确的分母口径"当作给定。这几行背后是同一块地：凡是一样产出，都有一个客观的个数。

一旦承认个数没有独立于合并判准的事实，上面每一家

都会在自己的框架里出现一处不一致——这正是这条判断至今没有被辨识过的来源：看的人都站在它上面。

本书的五位。

本书近日有四篇与本章同题、同取这三个学科，另有一篇更早的与本章关系极紧，必须逐条分清。

四篇问题的分别是：一篇立在生产者一侧，指出一件产出里有多大比例是为了证明它出自谁手；一篇立在关口一侧，指出正式门槛会整体饱和而不再分辨；一篇立在未阅一侧，指出没有人打开过就作出的否决与看过之后的否决在所有账本上同码；一篇立在参照一侧，指出没被接受的产出会合成一条不在记录里的判定基线。

本章与这四篇的关系不是竞争，而是分母。那四篇的读数里有三个是逐件计算的——每件里的证明层占多少、每件否决里有多少没人打开、每件落判如何进入参照；第四个虽以关口为分母，但它的输入是逐件的

一件是几件

通过率。如果"件"没有下界，这四个读数被同一个因子污染：切得越碎，每件里的证明层占比可能不变而总量翻倍，未阅率可能因为件数暴涨而自动上升，参照分布里的落判件数量与其实质无关。

判定形式：若在一个并项率稳定的通道里，那四个读数照常成立且解释力不变，则本章对它们无所增益，可以被忽略；若那四个读数在并项率下降的通道里出现与实质脱节的漂移，则它们各自都需要一个以并项为分母的校正版。

这条检验可以同时判五篇，而五篇都可能因它而输——包括本章：若"件"其实是稳定的，本章为零，那四篇分毫无损。

还有一篇更早的，讲的是"一份规范只写了什么算同类，没写同类之间有多近，而后者要单独花钱买"。它与本章的关系不是分界，是参数：本章的合并判准（按什么算同类可以并成一

项)正是那一章所说的那笔单独要花的钱。它指出了那笔钱的存在,本章指出了不花那笔钱的一个具体后果——所有以件计的读数一起失真。

判定形式:若在一个已经明确写死了同类间距的领域(例如有严格互换性规范的制造业),并

项率仍随工具可得性而下降,则本章的机制独立于那一章;若并项率只在同类间距未被规定的领域下降,则本章是那一章的一个应用。

最近的邻居是哪一个?是切香肠。本章之所以仍不能被它吸收,只有一个理由:它追究的是一个把整体切开的行为,而本章说的情形里没有那个整体,因而没有那个行为,也没有人可以被追究。

■ 十一、三处反过来削弱本章的约束

其一,工程学收窄了本章的适用范围。

批量理论证明小批量在流动上有真实收益——反馈更快、在制品

更少、风险更小,而这些收益与内容是否重复无关。这削掉了本章原本会写下的一句更强的话:切小是纯粹的账面动作、没有实质收益。这句话不成立。因此本章必须把命题收窄到:并项率下降本身不是病;病在于以件计绩、以件计负荷、以件作为拥挤程度的证据。

在不以件结算的场合,并项率低完全可以正是

确的工程实践。这一处动的是适用范围本身。

其二,艺术学收窄了本章的主张强度。

作品理论指出,有些东西的件数在原理上由创作者规定,不存在

独立于规定的事实——一个系列的十二幅画,合并会毁掉它;每一幅的意义部分地来自它与另外十一幅的并置。这削掉了本章原本会写下的另一句更强的话:并项数是一个客观量。它不是。它依赖于"按什么算同类"这个判准,而判准本身可以被争。因此本章的读数从"绝对量"降为同期横比量:并项率可以在同一判准下比较不同单位、不同团队、不同期刊,不能跨判准比较,也不能作为历史长时段的趋势线。这一处动的是价值排序——本章原本隐含"并项率越高越好",而在变体本身即产品的领域,这句话是错的。

其三,经济学交出了一个本章解释不了的洞。

新品种偏差同时说明,有些切分确实创造了真正的新东西

——一个新变体可能真的多做了一件事,哪怕它与旧的高度重合。因此并项率无法单独区分"同一份被切碎"与"真的多了一个品种"。要区分它们,需要配一个独立判据:合并之后是否

有信息损失。而"信息损失"本身没有一个跨领域通用的测法。这是本章最大的一处缺口，本章交出来，不掩饰。

能说的只有一个部分的、而且明确不够格的处置：问合并之后还能不能把原件重建出来。

若从并项出

发，凭借并项里已有的内容就能重新生成原来那若干件（措辞会不同，主张与取舍相同），那么原来的切

一件是几件

分没有携带额外信息；若重建不出来——某一件里有一个判断、一处取舍、一个反例，是并项里没有的——那么那一件是真的多做了一件事。

这个处置有两个已知的毛病，写在这里以免它被当成解决方案。第一个毛病是它把判断推给了重建者：重建得出来还是重建不出来，取决于谁来重建、他知道多少，而这恰恰是本章在第十五节禁止的那种不可争论的主观裁量。第二个毛病更根本：在这一轮，重建原件这件事本身也已经变得极便宜——一个能把主张写成十份不同措辞的工具，同样能从并项出发重新写出那十份。于是这个检验会系统性地偏向"没有信息损失"，而它偏得有多厉害，与工具能力同步变化。一个随着被测对象的能力而漂移的判据，不能当作判据用。

因此本章只能把这个洞如实标出来：并项率读得出"件数虚了多少"，读不出"虚的那部分里有没有真东西"。要补上它，需要的是另一条独立的路，本章没有。

■ 十二、出路，以及它最省事的实现方式会毁掉什么

本章的干预点由第六节直接给出：失真发生在以件为分母的那一步。因此出路有四条，按可行性排列。

第一条，按并项计，不按件计。

评价、负荷分配、拥挤程度的判断，改用合并后的项数。专利统计以族

为单位已经做了几十年，法庭归纳争点也是同一个动作——这个动作在制度上是可实现的，不是空想。要做的只是把合并判准从行政事实换成内容。

第二条，给合并记账。

合并是一次生产，而目前没有人有动机做它，因为产物不归任何人。若给合并者

署名、计分、计入产出，并项就会被生产出来。这一条最便宜，且不需要任何人放弃现有做法。

第三条，别限件数，限并项。

限件("只交你最好的五篇")是取样,不改变并项率,也不改变实质——

被限的人只会更认真地挑,不会更整地做。限并项才改变行为。这一条对现行的代表作制度是一处直接的批评:代表作制度看起来在解决这个问题,实际上没有碰到它。

第四条,把合并判准公开写死。

由于并项率只在固定判准下可比,判准必须先于读数公布,且不得因读

数难看而中途修改。

这四条有次序,而且次序不可颠倒。第四条必须最先做——判准不先写死并公开,前三条都无从执行,而且都会立刻退化成第十二节末尾所说的那种合规。第二条其次,因为它最便宜、不需要任何人放弃现有做法,且它会把并项这个东西真的生产出来,从而让第一条有对象可计。第一条第三,它要动评价办法,代价最大。第三条最后,因为限并项是一个有约束力的规定,而在并项还没有被稳定生产出来之前作出规定,等于把判准的解释权交给被规定的一方。

还有一条不在四条之列、但值得写下来的处置:在换分母之前,先停止用件数论证拥挤。

这不需要任何

新制度,只需要在写"本岗位收到二百四十份申请""本刊投稿量增长四成"这类句子时,同时给出提交次数与提交人数——后者在所有系统里都是现成的,而两者之比已经能挡住相当一部分误判。

现在说反噬。这条命题一旦被制度采用,最省事的实现方式恰恰会摧毁它要保护的那样东西。

一件是几件

监测并项率需要一个合并判准。而一旦这个数进了考核,把它做好看的最省力做法不是把工作做整,是放宽判准——"这五件本来就是同一项"。判准一放宽,并项数上升,并项率立刻好看,而交出来的东西一点没变。更糟的是,放宽判准恰好摧毁了这个读数唯一的意义:它只在固定判准下可比。

于是这个数会以

一种在形式上完全合规的方式变成噪声。

这不是一个待办事项。我看不到出口。

■ 十三、不适用的边界

本章的判断在下面几种情形里不成立,写清楚是为了让它可以被证伪,而不是为了自保。

变体本身即产品的领域不适用。

系列艺术、随机对照试验的多臂设计、育种筛选、参数扫描、甲乙测试

——在这些地方，若干高度相似的件的并置本身是产物，合并会毁掉它。这类领域的并项率低是正常且必要的。

练习与训练不适用。重复本身是目的，把一百次练习并成一项没有意义。

安全关键的事件记录不适用，而且是被禁止的。见第十五节。

不以件结算的通道不适用。

若一个通道既不按件计绩、也不按件分配负荷、也不用件数判断拥挤，那么并项率低不产生任何后果。

并项率本身很高的场合不必用。

粗略地说，并项率高于零点七时，件数仍然是投入的合用代理，用本章的框架是多余的开销。

判准无法被公开写死的场合不适用。

有些领域里"按什么算同类"本身就是这个领域的核心争议——什么

算同一个诊断、什么算同一起事件、什么算同一项权利主张，这些问题的答案就是该领域的实质工作。在这类地方强行写死一份合并判准，等于用一个行政决定去了一场实质争论，而并项率算出来的数会把那个行政决定伪装成一个测量结果。判准可争的地方，这个读数只能用作讨论的对象，不能用作结算的依据。

■ 十四、证伪条件与一条写死日期的赌注

先写出最强的竞争解释，再让证伪条件排除它。最难反驳的那句话是："其实不就是有人在切香肠吗？"主证伪条件：控制"作者手上是否持有一个可识别的完整原件"之后，若并项率与件数增长之间不存在剂量—反应关系，则本章为伪——届时件数上升应归因于策略性切分，而不是计数单位失去下界。

这个检验要求同质多单元，不能拿两个国家或两个行业充数。可行的设计是：同一代码托管平台上、同一主语言、贡献者规模相近的十到十二个仓库，在同一时间窗内比较工具采用程度、变更件数增长与并项

一件是几件

率；或者同一出版集团旗下、同一学科、投稿量相近的八到十二家期刊。自变量有真实变异，混淆变量同质。

其余七条：其一，低成本、数据已存在：取任一代码仓库两个相隔三年的季度，用同一套合并判准做内容合并，比较并项率。所需数据全部在版本历史里，成本几乎为零。

其二，排除古德哈特：在完全不以件计绩的通道里，若并项率不下降，则那只是指标被优化的一个新例子，本章为伪。

其三，排除压缩率：若并项率与文本压缩率高度相关，则本章的读数应被压缩率取代。

其四，排除切香肠：若要求作者提交合并版之后，合并版的实质内容不超出各件之和，则合并是还原而非生产，本章的核心区分为伪。

其五，排除内卷化：若件数上升的同时人均实际投入时间同比上升，则是投入真的增加了，本章为伪。

其六，顺序读数：并项率的下降应当早于件数的显著上升出现。若两者同时发生，或件数先动，则本章对因果次序的判断为伪。顺序比相关难被巧合满足得多。

其七，边界检验：若在纯粹的定标制通道、或在变体即产品的领域里，并项率下降同样带来负后果，则本章的适用边界划错了。

这条判断若被证伪，我们会学到什么？

若剂量—反应关系不成立，那么件数确实仍是投入的合用代

理，"投入增加而产出不增"这个诊断成立，处方就回到限投与提高门槛——而那意味着这一轮与历史上的内卷在机制上没有区别，值得单独研究的东西并不存在。证伪本身指向一个完全不同的结论，这正是它值得做的原因。

赌注。

到二〇二九年十二月三十一日为止，至少有三家机构在正式的评价办法、投稿须知或产品文档

里，写入按合并后的项计数的可执行条款。什么不算命中，写死在这里：只限制件数上限（那是取样，不是合并）不算；代表作制度不算；只发布倡议、白皮书、研究报告不算；只在访谈里表示正在考虑不算。

必须写明合并判准由谁定、按什么合、争议怎么复核。

本章另交出两个自己解释不了的洞。其一，见第十一节第三处——并项率无法单独区分"切碎"与"真新品种"。其二，合并判准的稳定性本身没有理论：不同的人按同一份判准做合并，一致性有多高，本章完全不知道，而这个数决定了并项率能不能被当作正式指标使用。

■ 十五、伦理与证据边界

并项率是一个比率，而比率几乎必然会被拿去考核人。三条必须写死。

其一，这个读数指的是位置，不是人。

一件是几件

它读的是一个通道的计数结构，不是某位作者、某位工程师、某

位申请人的诚信。任何机构不得以"此人并项率低"为由认定其学术不端或绩效造假——本章全篇的判断恰恰是，切小可以在完全诚实、完全合规、甚至完全符合最佳实践的情况下发生。把一个关于通道的读数读成一个关于个人品行的读数，是这个概念最容易被滥用的方式。

其二，安全关键领域禁止为把这个数做好看而合并。

安全事故报告、药物不良事件、器械故障、审计发

现、投诉记录——这些领域的逐件保留是有理由的：一个被并掉的记录可能是下一次事故的唯一线索。在这些通道里，并项率不是一个应当被提高的指标，它根本不该被计算。

若有人以"降低碎片化"为由合并

本应逐件追踪的事件，那是拿本章的读数去做一件本章明确反对的事。

其三，已经据此对某人做出不利安排的，必须公开合并判准并给出复核通道。

合并判准的选择直接决定

并项数，而不同判准可以让同一批工作从十二项变成两项或八项。判准不公开，这个数就不可争论，也就不该被用于任何有后果的决定。

最后一条与第十二节的反噬相连：一个读数一旦进入考核，最省事的达标办法通常是在判准上做手脚。这一条必须在制度设计时就被明写，而我看不到出口。

■ 十六、本章自身的位置

一条判断如果不能对它自己适用，它多半只是一句描述。

本章所在的这个栏目，近日已经有四篇同题、同取这三个学科的文章。按本章自己的判据，问一句最难堪的：这五篇合并同类项之后，还剩几项？

诚实的回答分两层。

第一层：不是一项。四篇各自的承重判断不同、各自的读数不同、各自的失效条件不同，而且第十节已经给出了一条可以同时判五篇的对照检验——它们可以各自被独立证伪，这正是"不是同一项"的硬标准。若它们真是同一项，那条对照检验不可能存在。

第二层，也是本章必须认账的一层：这个回答是作者自己给的，而本章的判据恰恰要求由外部按一份公开的判准来合并。

作者说"我们不是同一项", 在本章自己的框架里没有效力——那正是第十五节第三条所

禁止的那种不可争论的自评。因此本章在这一点上无法自证, 它只能把材料摆出来, 把合并留给别人。

还有第三处, 比前两处更硬。本章主张合并是一次生产、需要有人花钱去做、而没有人有动机做。那么本章自己有没有做那件事? 部分做了——第十节那一段把四篇放在一起、指出它们共用一个分母, 就是一次合并的尝试。而它是不完整的: 本章没有交出那个"把五篇并成两项"的成品, 只交出了合并的判准与检验。按本章自己的标准, 交出判准而不交出并项, 仍然停在提出的那一侧。

补上它意味着重写别人的四篇。

这个缺口本章补不上, 因为

一件是几件

■ 十七、结论

产生一件东西的代价降到接近于零之后, 几乎所有行当同时报告件数暴涨而通过者不增, 这被普遍读作投入增加而产出不增。这个读法有一个从未被检验的前提: 件数是投入的合用代理。

三个互不往来的学科各自持有一套关于"是几件"的成熟读法——按结果加总、按创作者的授权、按流水条件的权衡——而三套不能同时为真。它们争的是这件事该由哪一维读出, 共同假定了它是一个可以被单独读出的事实。推翻这条假定的材料来自三家自己: 不存在跨期一致的那一组东西; 交易成本趋零时最优批量没有下界; 作品在他人介入之前于关键方面是不定的。

由此得到的判断是: 这一轮的"更卷", 不是产出量增加, 不是作品数增加, 也不是工作项增加, 而是"一件"这个计数单位失去了下界。件数与并项之间裂开了一道口子, 而并项——把全部提交按内容合并之后剩下的那个单位——从来没有被任何人生产过, 因为它不归任何人。

它带来的后果不是一件事变糟, 是一整排读数同时失去意义而外表毫无变化: 人均件数、每岗申请数、每人年发表数、通过率、变更频率, 全都以件为分母。出路也因此不在限投、不在扩容、不在提高门槛, 而在换一个分母, 并且给"把东西合起来"这件事记账——那是目前唯一一件既有人做得了、又没有人有动机做的事。

一句话可以拿去问任何一份提交记录: 把这一轮的全部提交合并同类项, 还剩几项?

答案已经躺在那

里, 只是从来没有人算过。

■ 注释

〔一〕本章所说的"合并同类项"是一个借用，指按一份事先写死的内容判准把多件归为一项，不含数学上的任何具体运算。判准必须先于读数公布，理由见第十二节第四条与第十五节第三条。

〔二〕"并项率"在不同行当里的载体不同：代码那里是变更，论文那里是篇，招聘那里是份。本章不主张这些载体可以直接比较，只主张这个比率可以并排。

〔三〕第六节第五小节对专利族的处理，是本章最容易被误读的一处：本章不主张专利族是错的，只主张它合并的是登记冗余而非内容冗余，二者的相关性可以直接测。

〔四〕第十节对本书五篇的处理，是本章自己的判据在本章身上的一次执行。其中四篇与本章同题、同取这三个学科，本章与它们的关系是分母而非并列；读者若只读本章而不读那四篇，会得到一个没有内容的框架，反之则会得到四个可能已经漂移的读数。

〔五〕本章全部经验材料均为公开来源，日期截至二〇二六年七月。本章没有自己做任何实验；第十四节给出的八条证伪条件全部是设计，不是结果；文中不报告任何虚构的样本、统计量或实验结论。

〔六〕本次工作作废的方向共五条：其一，把困境归给"内容同质化"（被第十节压缩率那一格判掉，二者预测正交）；其二，把并项写成"去重"（去重预设有原件，与切香肠同病）；其三，把出路写成"限

一件是几件

件"（第十二节第三条已说明它是取样不是合并）；其四，把并项率写成可跨期比较的趋势线（被第十一节第二处约束推翻，只能同期横比）；其五，把这一格的病灶写成"评价标准落后"（评价标准并没有落后，落后的是它所使用的单位）。

■ 附节·换一组学科再问一遍：单位、著作层与履约义务

本章的论证由经济学、艺术学、工程学三家撞出。与第四章、第八章的同名附节一样，本节换一组毫无亲缘的学科把同一个问题再问一遍：当生成一件东西的代价趋近于零，"一件"还算不算一个单位。

换上的三家是：计量学（单位的定义与可溯源性）、图书馆学（著作—内容表达—载体表现—单册的四层模型）、会计学（收入确认中的履约义务识别）。三家都不研究人工智能，而三家都已经被迫回答过"这是几件"。

计量学的位置。一个单位之所以是单位，不是因为它自然存在，而是因为存在一条可溯源的定义链，把每一次测量接回一个约定的基准。米曾经是一根金属棒，后来是光在特定时间内走过的距离——单位的历史就是一部把定义从实物移到不变量上的历史。它对本章的回答是：件数失效不是因为件多了，是因为"一件"从来没有被定义到一个不变量上，它一直挂在实物上，而实物的生产成本塌了。

图书馆学的位置。编目传统早就不把"一件"当作单一级。同一部作品可以有多个内容表达，同一个表达可以有多个载体表现，同一个载体表现可以有多册。问"这是几件"在这套模型里是不合法的问题——必须先问"在哪一层数"。它对本章的回答是：件数没有失效，失效的是"只在最外一层数"这个习惯。

会计学的位置。收入确认的核心工序是识别合同中的履约义务：一个承诺能否单独构成一项履约义务，取决于它是否"可明确区分"——客户能否从中单独受益，以及它在合同中是否可与其他承诺分开。数几件，等于判断哪些承诺是可明确区分的。它对本章的回答是：件数由可区分性决定，而可区分性由承诺的结构决定，不由产出的物理形态决定。

三条不能同真。计量学要把单位接到一个不变量上，图书馆学要把单位摊成多层，会计学要按可区分性重新划分。三家共同假定了一件没有说出的事——存在一个正确的层级，在那一层上"一件"是确定的：计量学假定它在基准那一层，图书馆学假定它在四层中的某一层，会计学假定它在可明确区分的那一层。

推翻它的材料来自图书馆学自己。四层模型在实践中遇到的最大困难，恰恰是层与层之间的边界本身是判断出来的：一个改动多大才算新的内容表达，一次重排算不算新的载体表现，编目规则给的是指引而不是判据，不同机构对同一批材料给出不同的分层。也就是说，这套模型没有取消"这是几件"这个问题，它只是把它推迟了一层——而推迟之后那一层同样需要判断。这正是本章的核心判断在另一套材料上的复现：计数单位的下界不是被发现的，是被判定的；判定一旦廉价，下界就消失。

由此得到本节对本章的两处推进，一处是确认，一处是新东西。

确认的一处：本章正文主张"件数上升不再能充当投入上升的证据"。三家从三个方向支持它——计量学说单位没有接到不变量上，图书馆学说层级边界是判断出来的，会计学说可区分性由承诺结构而非物理形态决定。三条都指向同一件事：件数是一个被判定出来的量，而判定的成本正在归零。

新东西的一处：并项率应当在哪一层算。本章正文定义并项率为"按上一期的计数约定重新归并会被并成一件的比例"，但没有说这个归并在哪一层做。接上图书馆学之后，这个含糊必须被消除，而消除它同时给出一个更硬的读数：

> 著作层并项率：把本期计入的全部件数按"是否服务于同一个可明确区分的承诺"归并到著作层之后，件数缩减的比例。

它比正文那个版本硬在两处。其一，它不依赖"上一期的计数约定"——那是一个会随机构习惯漂移的参照；著作层的判据来自承诺结构，可以在同一期内部独立作出。其二，它可以在没有历史数据的新通道里算出来，而正文那个版本必须有至少两期数据。

这条读数还给出一个判决性预测：在同一批产出上，著作层并项率与正文的跨期并项率应当高度相关；若两者在控制机构习惯漂移之后仍不相关，则"著作层"这个概念在此不适用，本节的改进不成立，正文的定义恢复效力。

最后是这次换组对本书 E 维的实际贡献。它改变了本章交出的读数——从一个需要两期数据、依赖机构习惯的量，改成一个单期可算、判据来自承诺结构的量。这一改动同时波及合章：合章第五条用并项率修正拥堵参数，用著作层并项率去修正比用跨期并项率更稳，因为后者本身会随计数约定漂移而漂移，而合章正是要拿它去校正跨期比较。

第八章

未分之别——分辨力是一项没有人生产的产出

■ 摘要

同一个"越努力越没有差别"的现象，在三个互不往来的行当里被各自解释，而三条解释互不相容：经济学把它归给报酬按相对位次结算，判定绝对产出的提高在原理上不改变任何人的位次；艺术学把它归给重组既有形式比开出新形式便宜得多，判定在"可能性似乎已被穷尽"的处境下重组是理性选择；控制论把它归给控制者的变异度不足，判定当控制者跟不上被控者时，唯一可行的做法是削减被控者的状态种类。三家共同假定了一件从未被说出的事：可被分辨的种类数是给定的——它是一个既有的、有限的背景，剩下的问题只是怎么在它里面结算、重组或削减。

推翻这条假定的材料来自控制论自己。艾什比在给出变异度这个概念时，用一个括号里的限定条件写明：变异度是可被区分的元素个数，而"若两个观察者所能作的区分不同"，同一组元素对他们就有不同的变异度。控制论把这一句读成一条建模约定，从未把它读成它字面所说的东西——种类数不是系统的自然属性，它由观察者的分辨集合与系统状态共同决定；改变分辨集合就改变种类数。既然可改，它就被生产的。

本章由此提出一个在三家现行形式体系里都不构成一样东西的对象：未分之别——实际存在于产出之间、而在当前分辨集合下不构成一个种类的那些差别。它不是"记录得少"：在变异度的定义里，不可区分的元素不计数，因此它在字面上不存在。据此本章给出一条与三家都不同的判断：内卷不是投入增加而产出不增加，也不是竞争过度，也不是控制失灵，而是分辨集合停止扩张之后，全部努力被迫落在既有种类之内——努力提高的是同一种类的强度，而系统缺的是种类；而由于三家都把分辨集合当作外生的背景，"种类不再增加"这件事在三处都不显形。

本章给出一句不含判断词的问话（上一轮新增的投入，让你们现在能分辨出以前分辨不出的哪几件事），一个三领域量纲不同而比率相同的读数（新种率），一张二维辨别格，六条可由既有记录低成本执行的失效条件，以及三条各对一家的出路——其中第三条是艾什比定律本

身的另一半：控制者要么提高自己的变异度，要么削减被控者的，而后者在账上是免费的，所以它必然被选中。

关键词：未分之别

新种率

分辨集合

强度型内卷可比性的生产

■ 一、引言：三个行当，同一句"我们已经很努力了"

一家公司把绩效评估重做了一遍。评分维度从五个加到十四个，评审从一年一次改成一季一次，每一条评分都要求有证据链，最后再拉一次校准会消除评委间的偏差。所有人都同意这套比原来那套精确。年底复盘，晋升的还是那些人，排序几乎没变，而每个人花在准备材料上的时间翻了一倍。

未分之别

一位插画师这一年出了比往年多一倍的稿。她的线条更稳、上色更准、交付更快，客户满意度评分创了新高。她自己的感觉是：我这一年画得比去年好，但我这一年没有画出任何一张去年画不出来的东西。

一个运维团队把监控做到了行业前列。告警响应时间从十二分钟压到三分钟，指标覆盖率从六成提到九成八，值班手册厚了三倍。然后那年最严重的一次事故，没有任何一条告警响过——它的失效形态不在任何一条监控规则的射程内，事后写复盘时才发现，那类故障连一个名字都还没有。

这三个人不认识彼此，所处的行当没有共同的期刊、会议或术语。但他们各自那一年多做出来的东西有一个共同的性质：它提高的是同一种类的强度，不是种类的数目。

评分维度从五个加到十四个，听上去是种类变多了；但十四个维度全部落在"这个人干得怎么样"这一件事上，它们是同一根轴的十四种刻度，不是十四个方向。线条更稳、上色更准，是同一件事做得更好。告警更快、覆盖更全，是对已知故障模式的响应更密——而那次事故属于一个尚未被命名的类别。

本章要论证的是，这三处的共同点不是巧合，也不是三个行当各自的毛病，而是同一个结构：当一个系统能分辨的种类不再增加，投进去的一切努力就只能落在已有的种类之内；而由于"能分辨的种类"在所有账本上都被当成给定的背景，这件事在任何一处都不会显形为一个问题。

在展开之前，需要先说清本章如何裁定它所回应的那个问题。通常的提问是"内卷的根源是什么"，它期待一个原因。本章观察到的是：被归因的那个现象——"投入增加而产出不增加"——本身是一个读数，而读数依赖于用什么去读。因此本章把问题改切为：是什么在逼动竞争形态的改变，先动的是什么，以及为什么这一改变在三家的读数上都不显形？

这一改切同时挪动了出路的位置：如果问题在分辨集合不再扩

张，那么出路就不在"少卷"，而在把分辨力当作一项要花钱去买的产出。

以下的次序是：第二至四节分别按三家各自最强的形态陈述它们的驱动主张，每一家都给出它自己承认的一处撑不住的地方；第五节定位三条驱动为何不能同真；第六节说出三家共同假定而从未说出的那一条；第七节给出推翻它的那件材料，并说明它为什么必须来自三家之一自己；第八节正面建构本章的对象与读数；第九节据此重新定义内卷；第十节给出一条机制推论；第十一节说明三家各自为什么到不了这一步；第十二、十三节与既有说法以及本书两篇同题文章划清界线；第十四节给出三条出路；第十五节给出判据、失效条件与一条写死日期的预言；第十六节交代边界、本章自身的位置与两个解释不了的洞。

■ 二、经济学：报酬按位次结算，因为排序比计量便宜

经济学这一侧最强的形态，不是"AI

会不会造成失业"这一类量的估计，而是一条关于报酬按什么结算的判断。

拉齐尔与罗森在一九八一年提出的等级序列竞赛模型，其力量恰恰在于它解释的是一个反直觉的事实：为什么大量的雇佣合约不按产出付酬，而按名次付酬。他们的答案很硬：当个人产出难以度量、而相对名次容易观察时，按名次排序的竞赛可以充当最优的劳动合约。奖励来自名次之间的奖金差距，而不是与绝对产出挂钩的计件率；组织由此在理论上省下了度量成本，同时保住了激励强度。

未分之别

这条机制推到底，得到一个远比模型本身要紧的

：一旦报酬按相对位次结算，绝对产出的普遍提高就在原理上不改变任何人的收益。所有人都提高一倍，名次不变，奖金不变；而由于名次是零和的，任何人不提高就要掉名次。于是均衡里每个人都必须提高，而每个人的提高都被别人的提高抵消。这正是"越努力越没有差别"的最干净的经济学解释，而且它不需要任何关于心态、文化或竞争烈度的假设。

把这条机制放进当下：当生成技术把"做出一份合格产物"的成本对所有人同时压到接近于零，绝对产出的暴涨是普遍的、同步的，因而它一分也不落进任何人的口袋。近年沿这条线做

下去的工作把这一点形式化为拥堵外部性——把某一类入口的竞争建模为名次序列竞赛，而不是斯彭斯式的绝对信号，从而把“军备竞赛”这个说法变成一个可以算的东西。同时，另一支实证工作正在拆这套机制的实施形态：强制分布评分的复兴带来了老一套的病理——高绩效者的主动流失、团队功能失调、以及对评分任意性的抱怨；更要紧的是，有分析指出这类评分的方差中超过一半来自评委偏差而非真实的绩效差异。

经济学这一家的单因主张，可以严格写成一条动力式：>按相对位次结算的报酬规则，与一个绝对产出被技术条件普遍抬高的条件场不容；这一不容改变的是努力的外显形态。

它是单因的，意思是它不承认另两条驱动是原动力。在这条主张里，形式的重组与控制的收窄都不是原因而是结果：艺术家之所以重组，是因为在零和的位次结构里，开新形式的期望收益低于把既有形式做精；控制之所以收窄，是因为排序需要可比，而可比需要收窄。经济学不说“三者共同作用”——那样它就什么也没主张。

它的时序读数是明确的：先动的是技术条件（绝对产出的普遍抬高），报酬规则随后不动——正是规则的不动构成了这条机制；而个体的努力形态最后改变，因为个体只能观察到规则，观察不到条件的变化。

它自曝的那一处，写在拉齐尔与罗森自己的论证前提里，而且写得毫不含糊：竞赛之所以优于计件，是因为度量名次比度量产出便宜。也就是说，用来排序的那根轴不是因为它重要才被选中，是因为它便宜。经济学承认这一点，并把它当作模型的一个前提条件处理。第七节要用的东西，就藏在这一类“被当作前提条件处理”的地方。

■ 三、艺术学：重组比开新类便宜，于是重组成为理性选择

艺术学这一侧看的是同一片现象，走的却是完全不同的一条路。

它的出发点是一件在这门学科内部已经被反复讨论的处境：当“所有可能性似乎都被探索过”这种感觉成为普遍的背景，重组既有框架就成为阻力最小的路径。近年有工作把这一点说得很直白：概念拼贴之所以占据主导，不是因为艺术家的道德失败，而是因为它是最经济、风险最低的策略——它是创造性枯竭之后的理性选择；而人工智能使这种拼贴变得空前高效且普遍。同一支工作并且指出，这一处境早于人工智能，其根源在人类创造力的限度而非技术，人工智能只是一个加速器，使它无法再被忽视。

未分之别

与之相印证的是生产端的实况。一项对中国数字绘画从业者二〇二二至二〇二五年的纵向追踪显示，从业者对生成工具的态度经历了从怀疑到策略性吸纳的转变，其吸纳逻辑是务实的：它被用来压缩劳动与时间成本，被用作构思阶段的脚手架——先让它生成一个粗稿，再拿

那个粗稿去和客户沟通，从而不必自己先画一遍。到二〇二五年，热情冷却、批评重新出现，但争论的焦点已经从"生成图像好不好看"转向了别处。

艺术学这一家的单因主张，同样可以写成一条动力式，而且它与经济学那条落在不同的位置上：>实际做出来的作品，与一个使重组变得空前廉价的条件场不容；这一不容改变的是判据与组织方式本身——从"开出一个新类"转向"在既有类里做得更精"。

它也是单因的。在这条主张里，位次结算与控制收窄都是结果：位次之所以能结算，是因为大家已经在同一个既有类里可比了；控制之所以能收窄，是因为需要被控制的种类本来就已经在收缩。艺术学不承认报酬结构是原动力——报酬结构在它看来只是形式选择被吸收之后的账面显形。

它的时序读数：先动的是判据（做什么才算值得做），生产实践随后跟上，通常滞后数月到两年（生成工具从出现到进入日常工作流大约就是这个量级）；而作品的可见形态最后改变。

它自曝的那一处，恰恰藏在"可能性似乎已被穷尽"这句话的措辞里：似乎。艺术史自己一次又一次地证明这个感觉是错的——每一代人都曾确信形式已经用尽，而每一次新形式的出现都不是靠在旧类里做得更精，是靠换了一套看的办法：把小便池搬进展厅、把广告印刷网点放大到画布上、把日常声响录进乐谱。

也就是说，可能性空间不是一个会被消耗完的既有储量，它的边界随着什么被算作一件作品而移动。艺术学知道这一点，并且把它写成艺术史的主线；但在"重组是理性选择"这条论证里，它仍把可能性空间当作一个给定的、会被穷尽的量来处理。

■ 四、控制论与工程学：控制者跟不上时，唯一可行的是削减被控者

这一侧最不像哲学，也最容易被形式化。

艾什比在一九五六年给出的必要变异度定律，是这门学问里少数几条有定理形态的

之一。它的构件很简单：一组元素的变异度是其中可被区分的元素个数；扰动方 D 有其变异度，调节方 R 有其变异度，结果的变异度有一个下界——只有变异度才能吸收变异度。艾什比自己把它压成一句话：调节方必须拥有足够大的动作变异度，才能把结果的变异度压到足够小。

这条定律的实际形态，是一个几乎所有工程实践都在做、而很少被明说的选择：当调节方的变异度不足时，维持可控只有两条路——提高调节方自己的变异度，或者削减被控方的变异度。后者在工程史上有一个人尽皆知的名字：标准化。有控制论学者把福特那句"顾客可以要任

何颜色，只要它是黑的"称作大规模生产削减环境变异度的典范；而互联网之后的信息生态，其变异度按任何一种度量都比四十年前高出数个数量级，使得削减这条路越来越难走通。

未分之别

当代的应用仍在用同一条不等式做诊断：当一家企业外包研发、引入多家供应商时，每一家供应商都带来新的扰动状态，内部动作的变异度与外部事件的变异度之间的差距因此拉大——这不是人力资源上的轶事，是一条控制论不等式被违反。

这一家的单因主张：>被控方实际具有的状态种类，与调节方所能作出的区分不容；这一不容改变的是什么算得上存在——落在分辨集合之外的状态不进入系统的现实。

它是单因的。在这条主张里，位次结算与形式重组都是被驱动的：之所以能按位次排序，是因为可比的维度已经被收窄到少数几个；之所以重组占优，是因为新形式若落在分辨集合之外，它在系统里等于没有发生。工程学不承认报酬结构或美学判据是原动力——那些在它看来都是变异度差被结算之后的表现形式。

它的时序读数：先动的是被控方的状态种类（因技术而暴涨），调节方的分辨集合跟不上，于是"什么算得上存在"被改写。

它自曝的那一处，就写在艾什比给变异度下定义的那一句里，而且是他自己加的括号：变异度是可被区分的元素个数——若两个观察者所能作的区分不同，同一组元素对他们就有不同的变异度。换句话说，被控方的变异度不是一个可以脱离观察者而独立测定的自然量。控制论承认这一点，并把它归入"建模时要先声明观察尺度"这类方法学约定。

■ 五、冲突定位：三条驱动不能同真

到这里，三家的主张各自成形，而它们的不相容是结构性的，不是措辞上的。

三条动力式落在三个不同的位置。经济学说，是"报酬按位次结算的规则"与"绝对产出被抬高的条件场"这两者的不容，改变了"努力的外显形态"；艺术学说，是"实际做出的作品"与"使重组变廉价的条件场"这两者的不容，改变了"判据与组织方式"；控制论说，是"被控方的实际状态种类"与"调节方的分辨集合"这两者的不容，改变了"什么算得上存在"。三条式子里被驱动的那一样各不相同——一个是外显形态，一个是判据，一个是条件场。

这使三家的分歧无法靠"判谁对"化解。它们不是对同一个问题给了三个相反的答案，而是对"什么在驱动什么"给出了三个互相取消的指派。若经济学对，则艺术学的判据迁移只是零和结构的下游后果，不能算原动力，而艺术学恰恰主张它是；若艺术学对，则控制论的分辨集合收窄不过是"什么算值得做"改变之后的技术显形，而控制论恰恰主张分辨集合差是自变量；若

控制论对，则经济学的位次结算不过是"可比维度已被收窄"的一个特例——只有当种类被压到少数几个之后，排序才成为可能，而经济学恰恰把排序轴当作给定的前提。

三家的时序读数也互相矛盾，而时序是可查的。经济学说先动的是技术条件而规则不动；艺术学说先动的是判据；控制论说先动的是被控方的状态种类。三者不可能同时先动。

未分之别

还有一处更硬的地方：三家的自曝彼此指认。经济学自曝"排序轴是因为便宜才被选中"——这正是控制论所说的调节方按自己的分辨能力定义被控方；艺术学自曝"可能性空间的边界随什么被算作一件作品而移动"——这正是控制论所说的变异度依赖于观察者的区分；控制论自曝"变异度相对于观察者"——而这恰恰使经济学的排序轴与艺术学的类别空间双双失去外生地位。

三家各自撑不住的那一处，指向的是同一个地方。这不是巧合，第六节要说的正是它。

■ 六、三家共同假定了什么

把三家的争论压成同一个句式：>三家争的是——多出来的努力，为什么没有换来相应的差别。

经济学把差别的消失归给零和的位次；艺术学归给类别空间的枯竭；控制论归给分辨集合的不足。三种归因不同，但它们都在同一个东西的内部作答。

现在问一句三家都没被问过的话："可被分辨的种类"这个数，是从哪来的？

回答这个问题，需要先承认一件三家都默认、但没有一家说出口的事：>可被分辨的种类数是给定的。

它是一个既有的、有限的背景。经济学的模型里，可竞争的维度是外生设定，模型研究的是给定轴上的激励；艺术学的论证里，可能性空间是一个会被探索完的既有储量，"所有可能性似乎已被探索过"这句话本身就预设了它是一个有确定容量的集合；控制论的定律里，扰动方 D 的变异度是给定的，调节方 R 去匹配它——整条不等式的形式就要求 D 先在那里。

正因为默认了这一条，三家才会不约而同地把问题定位在"怎么在这个数里面周旋"：是按位次结算（经济学）、是在既有类里做精（艺术学）、还是把它削下去以维持可控（控制论）。三条处方形态各异，共同点是没有一条把"增加这个数"当作一个可选项——因为在三家的框架里，那不是可以被做的事。

把这句话念给三家听，没有一家会反对。经济学会说：那是自然，竞赛模型总要先给定一个绩效维度。艺术学会说：那是自然，形式的可能性当然是有限的，艺术史就是在有限里翻新。控制论会说：那是自然，D 的变异度是被观察系统的性质，我的任务是让 R 去匹配它。

三家都会说"这还用说吗"——这正是共有前提的验收标准。一条会引起争论的假定，说明它是某个第四家的立场，不是三家共有的地基。

而本章接下来要做的，是从三家之一自己的材料里，把这条地基掀掉。

未分之别

■ 七、推翻它的那件材料，来自控制论自己

推翻一条共有前提，从外面搬一个理由来是没有用的——那只会让本章成为争论中的第四家，而争论本身安然无恙。要取消这场争论，材料必须是三家之一自己已经掌握、已经发表、只是从来没往这个方向用过的东西。三家无法退回自己的阵地，因为那件材料是他们自己的。

控制论有这样一件材料，而且它就写在变异度这个概念被提出的那一句里。

艾什比给变异度下定义时是这样写的：给定一组元素，它的变异度是可被区分的元素个数。紧接着他加了一个括号：若两个观察者所能作的区分不同，那么同一组元素对他们就有不同的变异度。这不是脚注，不是附带说明，它就在定义的正文里——因为不加这一句，"可被区分"这四个字没有意义：区分总是某个人作出的区分。

在控制论内部，这一句被读成一条方法学约定：建模之前先声明你的观察尺度，声明清楚了，后面的推导就可以进行。它是关于如何做研究的提醒，不是关于世界是什么样的主张。

现在把它换一个问题来读。

控制论用这一句回答的问题是：做模型时该怎么把变异度定清楚。答案是一句方法学的告诫。

但这一句同时无可辩驳地陈述了另一件事，而控制论从来没有把它当作一个

来使用：>一个系统"有多少种状态"，不是这个系统的自然属性。它由被观察系统与观察者的分辨集合共同决定。

改变分辨集合，种类数就改变。而既然它可改，它就不是被给定的——它是被生产的。

如果种类数是给定的，那么"两个观察者有不同的变异度"这句话就不成立，因为总有一个观察者是错的、另一个数错了。而艾什比明确地不这么说：他没有说其中一个是对的，他说的是对他们而言变异度不同。

这只能意味着变异度是一个关系量，而关系的一端——分辨集合——在观察者手里。

一旦这一点被接受，三家的地基就同时塌了一角：·经济学的排序轴不再是外生的。它是被选的，而且拉齐尔与罗森自己说了选它的理由——排序比计量便宜。一条因为便宜而被选中的轴，当然可以换一条更贵但更能分辨的；不换，是一个决定，不是一个事实。

· 艺术学的可能性空间不再是一个会被穷尽的储量。它的容量随"什么被算作一件作品"而变，而这正是艺术史反复展示的那件事——每一次形式的耗尽感，都被一次分辨方式的改变而不是一次更努力的重组所解除。

· 控制论定律里的 D 不再是纯粹外生的。调节方通过自己的分辨集合参与规定了 D 的变异度是多少；当它削减被控方的变异度时，它常常并没有真的减少那里发生的事，它只是不再区分它们。

这最后一句要紧，值得单独说清楚：削减被控方的变异度有两种做法，而它们在账上一模一样。第一种是真的让那里发生的事变少（取消一条产品线、砍掉一个方向）；第二种是保持那里照旧，只是不再分辨

未分之别

（把十四种不同的失败归成一个"其他"，把二十种不同的工作压成一个可比的分数）。在艾什比的形式里，两者都表现为 D 的变异度下降，因为变异度只数可被区分的元素。而第二种做法不花一分钱。

这就是本章所需要的那一击：如果种类数是被生产的，那么"多出来的努力为什么没换来差别"这个问题就问错了地方。差别可能一直在被做出来，只是没有被分辨；而没有被分辨的差别，在三家的形式体系里不是"记录得少"，是不构成一样东西。

八、未分之别

8.1 定义与边界未分之别：实际存在于产出之间、而在当前分辨集合下不构成一个种类的那些差别。

它有四个构成要件，缺一不可：其一，它是实在的。它不是潜在的、可能的、想象中的差别，而是已经在产出里出现过的差别——两份交付确实不一样，两张画确实不同，两次故障确实是两种失效。它的实在性可以由一个分辨力更细的观察者当场证实，这也是它区别于"尚未被创造的可能性"的地方。

其二，它在当前分辨集合下不构成一个种类。这是本章最要紧的一条，也是它区别于一切既有概念的地方。未分之别不是"被低估的差别"，不是"没被重视的差别"——在当前的读法下，它根本不是两样东西。

两次归入同一个告警类别的故障，在监控系统里就是同一件事发生了两次；两份被打了一个分数的工作，在考核表上就是同样的工作。

其三，它在三家的账上都没有位置，而且各有各的记法。经济学把它记进噪声——评分方差里那超过一半的、被归给评委偏差的部分；艺术学把它记进未能归类，也就是建模或批评的

失败；控制论干脆不记，因为变异度的定义只数可被区分的元素，不可区分的不进入计数。三家都能指到它，三家指到的都是它在自己账本里被判成失败的那个形态。

其四，它是新种类的唯一来源。一个新的种类不会从无中生出来。它只能来自把某些原本被归在一起的东西分开——也就是说，每一个新种类的诞生，都是一份未分之别被分辨出来。这一条决定了未分之别不只是一个统计上的残差，它是一份存量，而分辨集合的扩张是对这份存量的支取。

有两条边界必须先划清，否则这个概念会膨胀到没有内容。

边界一：未分之别不等于"应该被分辨"。分辨是有成本的，而且成本不低——每增加一个种类，就要增加一套判据、一套记录、一次培训、一层复杂度。本章不主张所有差别都该被分辨；把差别归并本身是控制的正当手段，也是艾什比定律允许的两条路之一。本章只主张：归并的代价从不入账，因此归并与提高分辨力这两条路从来没有被放在同一个天平上比较过。

未分之别

边界二：未分之别不等于多样性。多样性是产出侧的性质（做出来的东西有多不一样）；未分之别是产出侧与分辨侧之间的差。一个产出高度雷同而分辨力极粗的系统，未分之别可以很小；一个产出高度多样而分辨力同样极粗的系统，未分之别很大。二者可以脱耦，这一点在8.2的辨别格里就是那两根轴。

8.2 一张二维辨别格只给一个新名字，得到的还是一个名字。要让它成为一个能把两样看起来一样的东西分开的读数，必须加第二根轴。

新种率高

新种率低

总投入

① 扩张——投入既提高了强度也开出了②强度型内卷——投入全部落在既有种类之内；所有

增加

新种类；账面与实况一致

总投入

③ 换向——不加投入而改分辨方式；罕

不增加

见，通常出自外部冲击

指标改善，而能分辨的事一件没多④停滞——什么也没动

四格中要打的是第二格。

这一格的选择不是任意的。一张辨别格里若被打的那一格是一眼就能看出的坏，说明第二根轴选错了——那一格不需要任何框架，任何人都看得见。第四格（停滞）就是这样一格：一个既不投入也不换向的组织，稍有经验的人走进去半天就能察觉。它不需要本章。

第二格则相反，它有三处签名，缺一不可：签名一：它在短期指标上表现为改善。强度型内卷中的组织，其响应速度、覆盖率、精细度、规范度、一致性通常全部在上升，而这些无一例外是正向指标。前面三个引子里的公司、插画师与运维团队，在各自的仪表盘上都是当年最好的一年。改善的那几项，恰好全部是"同一种类做得更好"的度量。

签名二：它的失效不产生任何信号。这一条是本章最硬的一处，而且它不是经验观察，是从变异度的定义直接推出来的：落在分辨集合之外的东西，按定义不出现在任何一张表上。因此"能分辨的种类不再增加"这件事，在原理上不可能被现有的任何指标发现——发现它需要一个分辨力更细的观察者，而如果已经有了那个观察者，问题本身就不存在了。运维团队那次没有任何告警响过的事故，正是这条签名的实例：不是监控做得不好，是那类故障还没有名字。

签名三：它自我加固。分辨力越粗，可比性就越好，考核、排序、汇报就越顺畅，于是就越显得应该继续把维度收紧、把口径统一、把例外归并。而每一次统一口径，都在把一份未分之别转成"同一件事"。每一轮正确的治理动作都在把这一格加深。

三条签名同时成立，第二格才是本章要打的那一格。

8.3 新种率一个跨领域的判断，若不能给出一个三领域量纲不同而比率相同的读数，就只能停在概念上。

未分之别

新种率：在一个给定周期内，分辨集合新增的种类数所对应的投入，占该周期总投入增量的比例。用一句白话说：新增的钱里，有多少让你现在能分辨出以前分辨不出的事。

三处的量纲各不相同，比率相同：· 组织与考核：新增的评价维度中，用于评价"以前无法被评价的那类工作"的比例——而不是把已有维度切得更细。

· 视觉创作：一个周期内的新作中，开出了此前不存在的可辨识类别的比例——而不是在既有类别里做得更精。

· 软件与运维：新增的监控规则与处置流程中，覆盖了此前无法被检出的失效模式的比例——而不是把已覆盖的模式报得更早更细。

这个读数须满足四条性质才值得造出来，本章逐条对照：其一，无量纲。三处都是比率，可以并排比较。

其二，可事后清点。新增维度、新增类别、新增规则全部躺在既有记录里——考核办法的版本历史、作品归档、监控配置的提交记录。这个读数不需要新建任何采集，只需要给既有记录加一次分类：这一条是"新的一种"，还是"旧的一种做得更细"。

其三，它把一个印象变成一个数。"卷"是一个感受，无法被争论、比较或治理。新种率是一条可以逐年画出来的曲线。

其四，它与既有的主要指标正交。这是四条里最容易漏、也最要命的一条：一个与既有主要指标高度相关的读数，会被立刻读成那个指标的代理，等于白造。检验的办法是举出一个"既有指标最好看、而这个数最难看"的真实例子。

这个例子是现成的，而且就是第一节那个运维团队：告警响应时间、指标覆盖率、值班手册完备度全部创纪录，而新种率为零——那一年新增的每一条规则，都是对已知失效模式的更早、更细、更全的覆盖，没有一条指向一个此前不存在的类别。反过来，一个手忙脚乱的早期团队，传统指标一塌糊涂，而它每周都在给一类此前没名字的问题起名字，新种率极高。

两个方向都能举出实例，正交因此成立。这也顺带说明了本章与"质量管理"的分界：新种率高不等于质量高，也不等于质量低；它是另一根轴。

8.4 一句不含判断词的问话一个判据若含有"应该""充分""真正""恰当"这类词，它就只能被拿去问人的判断，不能被拿去问文件。本章的判据是这一句：>上一轮新增的投入，让你们现在能分辨出以前分辨不出的哪几件事？

它可以被直接拿去问预算、考核办法、监控配置与作品归档，不需要任何人对"合理不合理"表态。它的答案是一个可以数的清单，而"数不出来"本身就是一个读数。

未分之别

把它在五个具体场景里跑一遍，五个答案应当互不相同，且其中至少两个不能从命题本身推出来，必须去查：1.

一家公司的年度绩效改版——十四个维度里，有几个评价的是以前无法被评价的工作。这一条可以从

命题推出。2.一个开源项目的一年——新增的 issue 标签里，有几个指向此前没有名字的问题类别。这一条也可以推出。3.一位插画师的一年——这一条推不出来：需要去查这个行业里"新类别"实际以什么形态出现（通常不是新技法，而是新的委托类型或新的呈现场景），不查行业实况给不出答案。4.一所医院的诊断规程更新——这一条同样推不出来，而且它引出一个本章必须承认的边界：在医疗领域，把两种病分开是一个需要长期证据才能做的动作，新种率低可能恰恰是审慎而非退化。第十六节会回到这里。

5.一个国家的统计口径修订——新增的统计科目里，有几个统计的是以前不进任何账的活动。

五个答案互不相同，其中至少两个必须去查行业实况才能给出。这个判据因此不是命题的复述。

■ 九、内卷的重新定义：强度型内卷

现在可以把本章的判断完整写出来。

>内卷不是投入增加而产出不增加——那是经济学的划界；也不是可能性空间被穷尽后的理性重组——那是艺术学的划界；也不是控制者跟不上被控者的失控——那是控制论的划界；而是分辨集合停止扩张之后，全部努力被迫落在既有种类之内：努力提高的是同一种类的强度，而系统缺的是种类。而由于三家都把分辨集合当作外生的背景，“能分辨的种类不再增加”这件事在三处都不显形。

这条判断改变的不是对内卷的评价，是对内卷这个读数的解释。

通行的内卷叙事包含一个隐含的事实主张：多出来的努力是白费的。本章主张这个主张在很多场合是错的——努力做出了差别，只是那些差别没有被分辨；而没有被分辨的差别，在所有的账上都不是“少记了”，是不存在。

这一改判有三处可以立刻检验的后果。

第一，它解释了那种“我确实进步了，但加起来等于没进步”的特定质感。通常的内卷叙事说人在做无用功。但当事人的体验往往不是这样：那位插画师知道自己今年画得比去年好，她的困惑不是“我在做无用功”，是“我确实好了，为什么这件事在任何地方都不显示为一个差别”。本章的回答是：它不显示，是因为读它的那把尺子今年一格也没细。

第二，它解释了为什么内卷对“更努力”和“更精细”双重免疫。如果内卷是投入过多，减少投入应当缓解；如果是精度不够，提高精度应当缓解。实际观察是两条都不灵。本章的解释是：强度与种类是两样东西，而努力与精细都只作用于强度。变异度是可被区分的元素个数——把同一件事做得更好、更快、更准，不增加任何一个可被区分的元素。这不是一个经验概括，是定义上的。

第三，它把责任从个体挪到分辨集合的持有者。新种率由谁决定？由持有分辨集合的那一方——定考核办法的人、定评价标准的人、定监控口径的人。个体在其中几乎没有余地：一个做出了未分之别人，若那

未分之别

个差别不在任何一张表上，他的处境与没做出来完全一样，而且他无法证明自己做出来过。因此把内卷归因于个体的"心态"或"不够聪明地竞争"，在机制上是错的。这一点本身也可被检验：若新种率与个体特质的相关，显著弱于它与所在组织分辨集合更新频率的相关，则责任在分辨集合的持有者一侧。

■ 十、为什么这个时代必然如此：一条机制推论

前面九节都在陈述一个结构。这一节给出它的一条推论，而这条推论是反直觉的，也是本章最愿意被拿去检验的东西。

推论如下：>在分辨集合不变的前提下，任何提高产出效率的工具，都会使未分之别的存量加速累积，同时使账面上的内卷加剧。

机制是三步：第一步：产出侧的状态种类暴涨。工具的全部意义就是让更多的东西被做出来、更快地被做出来、由更多的人做出来。按变异度的语言说：被控方的实际状态数在上升。

第二步：分辨侧不动。分辨集合的扩张需要另一种投入——不是把已有的判据用得更熟，而是承认一类此前被归并的东西其实是两样，为它造判据、造记录、造训练。这项投入昂贵、见效慢、且在任何仪表盘上都不产生短期改善（它一开始只会让分类变复杂、一致性变差）。而与之竞争的另一条路——把新出现的状态归并进既有类别——成本为零：不需要任何动作，默认就会发生。

第三步：两者的差被记成努力。产出侧涨、分辨侧不动，于是大量实际存在的差别落在分辨集合之外。它们在账上不是"新增的种类"，而是同一种类里更多的实例；而生产它们所花的时间与人力却是实打实记录下来的。账面结果就是：投入显著增加，可分辨的差别没有增加。

这条推论有一个可以立刻去核的间接印证，而且它来自经济学那一侧：强制分布评分的方差中，超过一半被归因于评委偏差而非真实绩效差异。按通常的读法，这是一个关于评分质量的坏消息——评委不够专业，需要校准。按本章的读法，它有另一种可能：那"超过一半"里，有一部分不是偏差，是不同评委分辨出了不同的东西，而评分表只有一个格子容纳它们。校准会消除的正是这一部分。两种读法可以分开：若加强校准之后，被消除的方差与后续的实际表现差异无关，那是偏差；若被消除的方差对后续表现有预测力，那它就不是噪声，而是未分之别。这是一个可以直接在既有评分数据上跑的检验，成本几乎为零。

这条推论也顺带解释了一件在三家框架里都不好解释的事：为什么生成能力最强的那些行业，内卷感最强烈。通常的解释是那些行业竞争最激烈。本章的解释是：那些行业产出侧的状态种类涨得最快，而分辨集合的更新速度并没有相应加快，因此未分之别累积得最快。它与竞

争烈度可以完全无关。这一条是可分辨的：若在竞争烈度相当而生成能力差异悬殊的两个行业之间，内卷感的差异由生成能力预测而非由竞争烈度预测，则本章这一条成立。

未分之别

■ 十一、三家为什么各自到不了这一条

在划清与既有说法的界线之前，需要先说清一件事：三家没有走到这一步，不是因为疏忽。一个领域看不见什么，通常不是它的缺陷，是它的问题结构的影子。

经济学到不了，因为它的问题是激励，而激励必须在一个给定的绩效维度上定义。竞赛模型要能写下来，就必须先有一个"绩效"作为随机变量；这个变量的维度是模型的输入，不是模型的输出。一门以"给定目标下如何设计报酬"为问题的学问，当然不会把"目标的维度从哪来"当作变量——那不在它的方程里。更要紧的是，拉齐尔与罗森明确说了排序轴是按测量成本选的，这句话在他们的语境里是模型适用条件的说明，不是一个关于世界的发现。对一门以合约设计为任务的学问而言，把轴当成给定的不是失察，是它的问题形式使然。

艺术学到不了，因为它的问题是形式的命运，而形式的命运只能在一个已经成形的形式空间里被叙述。艺术史是一部关于形式如何出现、被吸收、被反抗的历史；它天然地把"有哪些形式"当作它的研究对象，而不是把"形式的可能容量由什么决定"当作变量。"所有可能性似乎已被探索过"这个判断，只有在把可能性当成一个可清点的集合时才说得出口。艺术史自己在实践上一次次推翻了这个判断——每一次形式的耗尽感都被一次看法的改变解除——但推翻它的是艺术家的动作，不是艺术学的命题。

控制论到不了，因为它把观察者的相对性写成了方法学的前提，而前提在推导中是不被追问的。这一点在第七节已经展开：那句括号是为了让"可被区分"这四个字有意义而必须加的，加完之后整套推导就可以进行了。一条被用来开启推导的说明，不会同时成为推导的对象。控制论因此拥有本章所需要的全部材料，却不可能自己把它用成本章的

——那需要跳出正在使用的那套形式，去看它的前提。

三处"到不了"有一个共同形状：每一家都能指到未分之别，但每一家指到的都是它在自己账本里被判成失败的那个形态——噪声、未能归类、方法学约定。三家各自看见了它的一个投影，而没有一家有那个立体。

■ 十二、与最近的几种说法的分界

一个新名字若能被某个既有概念完整吸收，它就只是一次改名。本节把本章的判断与八种最容易被混为一谈的说法逐一对照。

需要先分清两件常被混为一谈的事。竞争解释是对同一现象的另一种因果说法（“其实不就是竞争太激烈了吗”），它归第十五节的失效条件管。竞争概念是学术史上已经有人命名过的那样东西，它可能与本章说的是同一个结构，只是名字不同——那归本节管。一个判断可以打赢所有竞争解释，却仍然只是某个既有概念的改名。只有本节能逮住这一种。

未分之别

在逐条对照之前，先做那道更基本的检查。把本章的命题去掉一切本领域的名词，只留“什么东西对什么东西做了什么，导致什么”，纯形式是：一类实际存在的差别因不在分辨集合之内而不构成状态，于是全部投入落在既有状态之内。这个形式在上游学科里叫什么？最近的两处是信息论中的量化（把连续量压成有限个电平，电平内部的差别不可恢复）与测量学中的分辨率。二者都精确地描述了这个形式，但它们把分辨率当作仪器的固定参数，不问“谁在决定它、为什么不换”。本章的全部增量在于把它当作一项被选择、被预算、可增可减的东西。

同时须声明：本章的形式不是“越成功越失败”那一类反转模板。这一点值得明说，因为那类模板早已各有其名——成功陷阱、伊卡洛斯悖论、能力刚性——谁若剥出那个形式而不点名它们，就是在给一个几十年前已被命名的结构重新起名。本章的形式不是反转，是分辨的停滞：不是成功导致失败，是种类不增而强度增。二者可分辨之处在于，反转模板预测总产出或能力实际下降，本章预测强度实际上升而种类持平。

这一条是可查的，见第十五节第二项。

其一，格尔茨的内卷化（Geertz，

1963）。这是本章最近的祖先，必须最先处理，也必须处理得最认

真。格尔茨用它描述爪哇农业：在既有系统内持续追加劳动投入，产生越来越精细的内部分化，而不发生突破。“投入增加、内部精细化、无突破”这三条与本章几乎逐条对应。

分离线有两处，都可测。其一，机制：格尔茨的内卷化由资源约束驱动——人口压力加上土地固定，人只能往同一块地里投更多劳动。本章的机制与资源无关：一个资源充裕、人力不紧张、没有任何外部压力的组织，只要分辨集合不扩张，照样进入强度型内卷。其二，因变量：格尔茨的因变量是形式的精细化（内部分化越来越细）；本章的因变量是种类数——而精细化恰恰可以在种类数不变甚至下降时发生（把一个类别切成十个子级，种类数没变，只是刻度变密了）。

判决性对照预测：考察一个资源充裕、人均负荷平稳、且不存在固定总量约束的组织。格尔茨的框架预测不发生内卷；本章预测只要分辨集合不更新，新种率照样趋零而投入照样上升。这一条把本章与它的祖先干净地分开。

其二，马奇的探索与利用（March，

1991)。这是最危险的一家：组织因在既有做法上不断成功而不断

加码，探索被利用挤出——"努力全落在既有种类之内"听上去就是它。

分离线在那些东西是不是可选项。探索与利用是一个资源分配的选择：组织知道存在别的方向，只是因为既有方向的回报更确定而选择不去；这个选择在原理上可以被一次算清收益的决策纠正。本章所描述的处境不是权衡后不去——落在分辨集合之外的东西不构成一个可以被选择的选项，因为它们不构成状态。你无法决定去探索一件你的记录系统里不存在的事。

判决性对照预测：直接问组织的中层与一线："你们知道自己没在探索哪些方向吗？"探索/利用预测答得出并能说出放弃的理由；本章预测答不出，且被追问时会用既有类别把问题接住（"我们该做的都在做"）。这一条只需一次半结构访谈。

其三，库恩的常规科学（Kuhn，

1962）。在既有范式内解谜，这与"努力全落在既有种类之内"高度同形。

未分之别

分离线在反常会不会积累。库恩的常规科学是有成效的，而且他的整套机制依赖于一件事：解谜会带出反常，反常积累到一定程度触发危机与范式转换。本章所描述的处境里，反常不会积累——因为一件不在分辨集合内的事，在记录上不显示为"解不开的谜"，它显示为"同一类事又发生了一次"。判决性对照预测：库恩预测危机前会有一段可查的、反常记录密度上升的时期；本章预测事故或失败发生前，记录里看不到任何异常密度的上升——直到那件事发生。

其四，斯科特的清晰化（Scott,1998）。国家为了可治理而把复杂的地方性实践简化为可读的形式，其代价是地方性知识被摧毁。这与本章的"归并不花钱"极近。

分离线在已有还是未生。斯科特的因变量是已经存在的多样性被削减——他讲的是破坏，其证据形态是"清晰化之前有什么，之后没了"。本章的因变量是新种类不再产生——即使从零开始建一个全新的组织，没有任何既有多样性可供摧毁，只要分辨集合固定，新种类照样不出现。判决性对照预测：考察一个新建组织。斯科特的框架对它无话可说（没有前期多样性可摧毁）；本章预测它同样出现新种率趋零。

其五，古德哈特定律与指标的反应性。一个度量一旦成为目标就不再是好度量；被度量者按指标重塑自身。

分离线：古德哈特处理的是已有度量被扭曲，其机制是行为者朝度量优化，因而需要一个有意图的行为者。本章处理的是没有度量，其机制与意图无关——即使所有人都诚实、都不朝任何指标优化，未分之别照样累积，因为它由分辨集合的边界决定。判决性对照预测：在一个

没有任何绩效考核、无人有动机操纵指标的环境里（例如个人的独立创作），古德哈特预测不发生扭曲；本章预测新种率照样可以为零。

其六，信息论中的量化与测量学中的分辨率。这是本章在纯形式层上最近的一位，见本节开头。分离线：二者把分辨率当作仪器的固定参数，其研究对象是"给定分辨率下能恢复多少信息"；本章把分辨率当作被预算的产出，其研究对象是"谁在决定它、按什么代价、为什么不换"。判据：若某个系统的分辨集合在原理上不可改变（物理仪器的极限），本章对它无话可说，量化理论足够；本章的射程只覆盖那些分辨集合由人决定的系统。

其七，组织学习中的单环与双环（Argyris&Schön,1978）。单环学习在既有目标下纠错，双环学习修改目标本身。这与本章"强度 vs 种类"的分法接近。

分离线在被修改的是什么。双环学习修改的是目标与假设（我们追求的是不是对的东西）；本章所说的扩张修改的是分辨集合（我们能不能把两件事分开）。二者可以独立发生：一个组织可以彻底改变目标而分辨集合一格未动（换了一个方向继续用同一套粗糙的读法），也可以目标完全不变而分辨集合大幅细化。

判决性对照预测：在一个刚完成战略转型（目标已改）的组织里，双环框架预测学习障碍已被解除；本章预测若分辨集合未更新，新种率仍趋零。

其八，注意力与认知负荷。一种流行的说法是：种类不增是因为人的注意力有限、复杂度超载。

分离线：注意力解释的因变量是个体的处理能力，其对策是减负、简化、聚焦。本章的因变量是记录系统能分辨的种类数，它与任何个体的注意力无关——一个记录系统可以有一万个类别而每个人只看自己那几

未分之别

个。判决性对照预测：给同一组人减负（减少工作量、简化流程），注意力解释预测新种类会出现；本章预测不会，除非分辨集合本身被扩张。

最近的邻居是哪一位。是其二，马奇的探索与利用。它与本章在现象层几乎完全重合，且它的文献基础比本章厚得多。本章之所以仍不可被它吸收，只因为一处：它的未探索方向是可选而不选，本章的未分之别是不构成可选项。这一处差别决定了两者的处方不同——一个是重新分配资源，一个是先把东西分开——因而它们不是同一条判断。

■ 十三、与本书两篇同题文章的分界

本书第四章第四章《认领物》与第三章第三章《这道关拦下过谁》，与本章用的是同样三门学科、回应的是同一个问题，且同月发表。读者有理由怀疑三篇是同一条判断的三次书写。这里把分界写清楚。

与第四章第四章《认领物》。那篇的判断是：产出的构成中，为使产出可被归属而生产的那一层占比持续上升，而三家账本只给它成本位、不给产出位。分离线在被漏掉的是什么：第四章《认领物》漏掉的是一层被生产出来的东西（它存在、有实体、只是没有栏位）；本章漏掉的是一份未被分辨的差别（它存在于产出之间，而在当前读法下不构成两样东西）。前者是遗漏，后者是不可见。

可裁决的差别：第四章《认领物》预测总产出实际上升而账面不变（多出来的是认领物）；本章预测能分辨的种类持平或下降而强度上升。两个量互不蕴含，可在同一批组织上同时测：一个认领占比很低（几乎不需要证明产出属于谁）而新种率同样为零的组织，是本章的靶格而非那篇的靶格；反之亦然。

与第三章第三章《这道关拦下过谁》。那篇的判断是：可读门槛整体饱和，分辨功能不再往上迁到更高一级门槛，而是整体迁出可读区，落到那些只能由会被追究的位置去承担的凭据上。这一章与本章的距离最近，因为两篇都在讲"分辨"出了问题。

分离线在分辨去了哪里。第三章说分辨迁移了——它离开可读区，落到别处（不可生产的凭据）；本章说分辨没有扩张——它哪儿也没去，只是不再增加种类。可裁决的差别：第三章预测事后可以观察到新的凭据类型出现（承担型的、不可生产的）；本章预测什么新类型都不出现——包括不出现新的承担型凭据。两条预测在同一份档案上方向相反：去数事发前若干年新出现的凭据类型数，第三章预测大于零且在增长，本章预测趋零。

三篇叠加时的关系。三者可以独立发生，叠加时不是相加。一个分辨集合停止扩张的系统（本章），其唯一还能作出区分的维度会越来越窄，因而对"这是谁做的"这类仍可分辨的维度的依赖会上升（第四章的认领占比上升），而当这些维度也饱和时，分辨才被迫迁出可读区（第三章）。这条推论三篇中任何一篇单独都给不出，而它同时给出一个可查的时序：新种率趋零应当早于认领占比的快速上升，而后者又早于门槛的整体饱和。若观察到的次序相反，三篇之中至少有一篇的定位是错的。

未分之别

■ 十四、出路：三条，各对一家

出路在本章的框架里不是一个独立的主题，而是判断的直接读数：如果问题在分辨集合停止扩张，那么出路只有三个位置——给分辨力定价、把未归类改判为资产、把归并从免费改成收费。三条恰好各对一家。

其一：给排序轴的选择定价（对经济学）。拉齐尔与罗森自己说了，排序轴是因为便宜才被选中的。那么最小的动作就是把这句话当真：每引入或更换一条评价维度，显式记一次账——这条维度让我们现在能分辨出以前分辨不出的什么，为它付了多少。

这个动作的意义不在于立刻改变任何决策，而在于让"我们已经三年没有增加过任何一种新的分辨"这件事可以被说出来。第八节第二条签名说过，这一格的失效不产生信号；而一条被记下来的账就是信号。这一条几乎不花钱，所需的记录全部已经在考核办法的版本历史里。

其二：把"未能归类"从失败改判为存量（对艺术学）。在艺术批评与作品归档中，"归不进任何类别"目前是一个负面判断——它意味着这件作品还没有找到自己的位置。本章主张把它单列为一个不进入排序的位置：不打分、不比较、只登记。理由是第八节第四条要件：新种类只能来自把原本被归在一起的东西分开，因此未归类的那一堆是新种类的唯一料场。把它清掉，就是把明年的种类清掉。

必须承认这条出路的代价：一个不打分不比较的位置会立刻成为逃避评价的去处。第十六节会正面处理。

其三：把归并从免费改成收费（对控制论）。这是三条里最硬的一条，而且它是艾什比定律本身的另一半。定律说调节方的变异度必须不小于被控方的；满足它有两条路——提高调节方的，或削减被控方的。

工程实践几乎总是选后者，而选它的理由不是它更好，是它在账上是零：发一份统一口径的规范、把十四种失败归成一个"其他"，不需要任何预算。

本章主张的动作是：任何一次归并，必须同时记录被归并掉的种类数。把二十种工作压成一个分数，就在账上记下"本次归并消去十九种"。这不阻止归并——归并常常是对的——它只是把两条路放回同一个天平上。一旦这个数存在，"我们今年消去了三百种、新增了两种"这句话才有可能被说出口。

三条的次序是有先后的：第一条必须最先做，因为没有那条账，后两条的效果无从判断，而未经度量的改革在几个季度之后一定会退回原样。

■ 十五、判据、失效条件与一条写死日期的预言

15.1 最强的竞争解释一个只验证"本章所描述的现象存在"的检验太弱，因为它排除不掉那句最朴素、最难反驳的话。先把它写出来：>其实不就是竞争太激烈、大家都在做一样的事吗？

未分之别

这是本章必须排除的竞争解释，而不是一个好打的靶子。它主张：种类不增只是竞争加剧的表现——当所有人挤在同一条赛道上，当然没人去开新赛道；根本不需要"分辨集合"这个中间项，也不需要未分之别这个新对象。

要排除它，必须找出只有本机制成立才会出现、而"竞争激烈"预测不出的那个独有变量。这个变量是：新种率的变化，应当由分辨集合的更新频率预测，而不由竞争烈度预测。"竞争激烈"这条解释预测：竞争越烈，新种类越少；它对分辨集合没有预言。本章预测：在竞争烈度不变的情况下，仅仅更新分辨集合，新种率就会上升。

于是失效条件可以写成一个可执行的形状：>

控制竞争烈度（岗位/名额与参与人数之比、淘汰率）之后，若"分辨集合的更新频率"与"新种率"之间

不存在剂量-反应关系，则本章的核心机制为伪——届时内卷应当被归因于赛道拥挤，而不是分辨的停滞。

关于样本，有一条纪律必须写死：要谈剂量-反应关系，样本至少需要六个以上单元，且在关键混淆变量上同质。拿两个国家或两个行业对比是不成立的——两边在规模、文化、监管、历史上处处不同，所谓"控制掉某变量"在方法上站不住。可用的设计是在同一制度环境内找自变量有真实变异的多个单元：同一家公司的多个事业部、同一所大学的多个院系、同一个平台上的多个社区。它们的竞争烈度大体同质，而分辨集合的更新频率差异极大。

15.2 六条失效条件其一（最便宜、可立即执行）。在同一组织的若干个部门里，按第十四节第一条的做法回溯统计过去三年的新种率——把每一次考核维度、监控规则或分类体系的变更逐条标为"新的一种"还是"旧的一种做得更细"。若新种率与该部门的内卷感（可用离职率、加班时长、员工自评的"努力是否有意义"三项合成）之间没有关联，本章的核心判断为伪。所需数据全部已在版本历史与既有调查里，不需要新建任何采集。

其二（区分本章与内卷化）。在一个资源充裕、人均负荷平稳、不存在固定总量约束的组织里，若不发生强度型内卷，则本章错而格尔茨对——机制在资源约束而非分辨停滞。本章预测在这类组织里照样发生。

其三（区分本章与探索/利用）。若组织能清楚说出自己没在探索哪些方向并给出放弃的理由，则那是权衡而非不可见，本章可被探索/利用吸收。

其四（区分本章与库恩）。在一次重大失败之前，若记录里可以观察到反常密度的上升，则危机机制成立，本章关于“反常不积累”的推论为伪。

其五（校准会消除的是什么）。在既有评分数据上做：把校准会消除掉的那部分方差单独取出，检验它对后续实际表现有无预测力。若无预测力，那是偏差，第十节的推论少了一个印证；若有预测力，那部分方差就不是噪声，而是未分之别，本章得到一个直接的经验支点。这一条几乎零成本，且它可能推翻的是本章自己最愿意援引的那个印证。

未分之别

其六（责任在哪一侧）。若新种率与个体特质的相关显著强于它与所在组织分辨集合更新频率的相关，则第九节第三条为伪，责任确实在个体一侧。

这条判断若被证伪，我们会学到什么？若失效条件之一成立，则说明内卷确实是一个关于投入与竞争的现象，而不是关于分辨的现象。那将意味着治理方向应当回到调节投入总量与竞争结构上，而不是去买分辨力。这是一个方向相反的处方，因此这场证伪不是白做的。

15.3 一条写死日期的预言>到二〇二九年六月三十日之前，将出现至少一份把“新增的分辨能力”与“既有能力的强化”分开计量的正式制度文件——它可以是一份组织的考核办法、一份行业的监管指引、或一份公共资金的申报规范，其中“本期新增了哪些此前无法被识别的类别”被作为一个独立项目要求填报，而不是与“改进了哪些既有指标”混在一起。

什么不算命中，须一并写死：· 仅有学术论文或行业报告提出这个建议，不算命中——必须是一份对某个具体组织有约束力的文件。

· 仅要求填报“创新数量”或“新产品数”，不算命中——那些统计的是产出，不是分辨能力。

· 仅把既有指标切得更细、增加维度数量，不算命中——必须能把“新的一种”与“旧的一种”做得更细“分开”。

· 由本章作者或其所在机构参与推动而产生的文件，不算命中。

若到期没有出现，本章关于“这条账迟早会被记起来”的判断即告失败；那将说明分辨能力在实践中并不构成一个需要被单独计量的对象，本章所命名的东西没有被现实认领。

■ 十六、边界、本章自身的位置，与两个解释不了的洞

16.1 两个学科反过来给本章加的约束一个跨领域的判断若只从各学科取用而不被它们削弱，说明取用得不够深。这里记下两处真正削掉了本章原本打算主张的东西的约束。

其一，来自医学：第十四节第二条出路的适用范围被大幅收窄，而且本章的价值排序被动了。本章原本打算主张：新种率低就是退化，扩张分辨集合总是好的。医学否掉了这个一般主张。把两种病分开——设立一个新的诊断类别——是一个需要长期证据、并且会带来真实伤害的

动作：过早的类别细分会制造过度诊断、过度治疗，以及一整批本来健康的人被划进病人。医学史上因分类过细而造成的损害，不比因分类过粗造成的少。

未分之别

因此本章的主张必须收窄为：新种率低本身不是病；病的是新种率低而分辨集合的更新从未被作为一个可选项讨论过。这一收窄削掉了本章原本更强的那句话——本章原本会写“应当持续扩张分辨集合”，现在只能写“应当把扩张与不扩张放在同一个天平上，并各自记账”。这是一处真实的削弱，不是补充条件。

其二，来自控制论自身：本章关于“归并不花钱”的说法必须让步。艾什比定律的两条路里，削减被控方变异度在账面上免费，但在系统上不免费——被归并掉的那些状态并没有消失，它们继续发生，只是不再被分辨；而一个不再分辨它们的系统，其结果的变异度按定律必然上升。也就是说，归并的代价不是零，是延迟到结果侧显形。这一处让本章原本“归并是免费的”这句干脆的话，必须改写成“归并的代价不在动作发生的那一侧入账”。

这一让步同时收窄了本章的适用范围：在结果侧反馈快而清晰的系统里，归并的代价会很快回来，因而不会长期累积——本章所描述的机制主要在结果反馈稀疏、延迟或模糊的领域里成立。这一条本身可测，也是本章的边界条件。

16.2 本章自己在哪一格按本章自己的判据审视本章：这篇文章增加了几个可分辨的种类？

答案是：一个。它把此前被合在一起的两样东西分开了——“努力提高的强度”与“可分辨的种类”。除此之外，本章全部的篇幅都是在把这一个区分说清楚、划界、给读数、给检验。按第八节的标准，本章自己的新种率是一个新类别除以两万字，这个比值并不好看。

这个位置有两个具体的后果，必须说出来。

第一，本章没有资格向任何人推荐“多开新类”。它自己只开了一个，而且用了两万字。扩张分辨集合是昂贵的，这句话本章是用自己证明的，不是从别处引来的。这也是第十四节三条出路全部只主张“记账”而不主张“多开”的原因。

第二，本章有可能就在自己的靶格里——它可能正是一次“把既有种类做得更精”的作业：三家的说法都已经存在，本章所做的是把它们放在一起看得更细。它无法自证不是。它能提供的唯一辩护是第十二节那八条分离线：若其中任何一条被证明不成立，本章就应当被相应的那个既有概念吸收，而那正是“做得更精”的定义。本章把这个判决权交出去，而不是自己宣布通过。

16.3 一个读数一旦进入考核会发生什么新种率是一个比率，而比率会被拿去考核人。这一点必须在提出它的同时处理，否则提出它本身就是不负责任的。

最省事的达标办法显而易见：把"旧的一种做得更细"改个名字报成"新的一种"。增加一个维度、拆一个类别、给一个已知问题起个新名字——这些在填表时与真正的新分辨没有形式上的区别，而分类权在填表的人手里。这条反噬预言必须写在这里，而且我看不到出口：任何由被考核者自行分类的读数都可以被这样规避；而由第三方分类的读数，其分类成本本身又是一笔新的开销，且第三方同样需要一个分辨集合。

未分之别

因此，若这个读数被采用，须写死三条：其一，这个读数指的是分辨集合的持有者，不是被考核的人。一个部门的新种率低，说的是它所处的评价体系多久没更新，不是它的成员不够创新。把它用于个人绩效评价，是对这个读数的误用。

其二，不得为了把这个数做好看，而在需要审慎的领域强行细分。医疗诊断、司法罪名、安全事故的分类体系，其细分速度必须由证据而非由这个读数决定。任何以"提高新种率"为由推动分类细化的行为，都是拿这个读数当借口。

其三，已经依据这个读数做出不利安排的组织，必须公开其分类规则并提供复核通道。一个没有公开分类规则的比率，其分子由掌握分类权的一方决定，因此它不是一个读数，是一次裁量。

16.4 两个本章解释不了的洞
第一个洞：分辨集合为什么会扩张。本章通篇论证分辨集合不扩张时会发生什么，但本章说不出它扩张的机制。历史上那些成功的扩张——把发热分成几十种、把"故障"分成一整套失效模式、把"艺术"扩张到装置与行为——每一次事后都能讲出一个故事，而没有一个故事能预测下一次。本章能指出这个空白，给不出机制。第十四节的三条出路全部只做到"让它可被讨论"，没有一条能让它发生。

第二个洞：新种率有没有一个上限。本章主张新种率过低是问题，但没有说它多高算过高。直觉上它必有上限——一个每周新增一百个类别的系统会立刻失去可比性，而可比性是协作的前提。这个上限是多少、由什么决定，本章不知道。这个洞是要紧的，因为一切关于"扩张分辨集合"的处方若没有上限的概念，都可能推过头；而推过头的后果——彻底失去可比性之后的协作崩溃——比强度型内卷更难收拾。

结

论

三个互不往来的行当各自观察到"越努力越没有差别"，各自给出一条互不相容的驱动：报酬按位次结算、重组比开新类便宜、控制者跟不上被控者。三家共同站在一条从未被说出的假定上：可被分辨的种类数是给定的。推翻这条假定的材料来自控制论自己——艾什比在给变异度下定义时用一个括号写明，若两个观察者所

能作的区分不同，同一组元素对他们就有不同的变异度。控制论把这一句读成建模约定，从未把它读成它字面所说的东西：种类数是被生产的，而生产权在持有分辨集合的那一方手里。

一旦这条假定倒下，"多出来的努力为什么没换来差别"这个问题就问错了地方。差别可能一直在被做出来，只是没有被分辨；而没有被分辨的差别，在三家的形式体系里不是记录得少，是不构成一样东西。本章把它命名为未分之别，并据此给出一条与三家都不同的判断：内卷不是投入增加而产出不增加，而是分辨集合停止扩张之后，全部努力被迫落在既有种类之内——努力提高的是同一种类的强度，而系统缺的是种类。

未分之别

由此得到的那条推论是本章最愿意被检验的东西：在分辨集合不变的前提下，任何提高产出效率的工具，都会使未分之别加速累积，同时使账面上的内卷加剧。它解释了为什么内卷对"更努力"和"更精细"双重免疫——因为强度与种类是两样东西，而努力与精细都只作用于强度；这不是经验概括，是定义上的。

出路因此不在少卷，也不在多卷。它在三个位置上：给排序轴的选择定一次价，让"我们三年没有增加过任何一种新的分辨"这句话可以被说出来；把"未能归类"从失败改判为新种类唯一的料场；以及把归并从免费改成收费——任何一次把二十种压成一个分数的动作，同时记下被消去的十九种。三条之中第一条必须最先做，因为没有那条账，后两条的效果无从判断。

最后一句留给本章自己。这篇文章两万字，只多分出了一个种类——把"强度"与"种类"分开。它没有资格劝任何人多开新类，因为它自己开一个就用掉了两万字；扩张分辨集合是昂贵的，这句话本章是用自己证明的。它也可能就落在自己的靶格里：一次把既有说法做得更细的作业。它无法自证不是，只能把第十二节那八条分离线交出去，等着其中某一条被推翻。这大概就是这件事最不好受的地方：判断一件事是不是新的一种，需要一个比现在更细的分辨——而那正是还没有的那样东西。

■ 注释

[^1]:等级序列竞赛作为最优劳动合约的原始论证，以及"当个人产出难以度量而相对名次容易观察时，按名次结算优于计件"这一适用条件，见 Lazear&Rosen(1981)。本章第二节所引"排序比计量便宜"即出自该文的模型前提。

[^2]:强制分布评分复兴带来的病理（高绩效者主动流失、团队功能失调、评分任意性），以及"评分方差中超过一半来自评委偏差而非真实绩效差异"这一读数，见

Tournament-Based

Performance

种解释，并给出区分两者的检验。

[^3]:把 AI 时代某一类入口的竞争建模为等级序列竞赛与拥堵外部性（而非斯彭斯式绝对信号），见[^4]:"概念拼贴之所以占据主导，是因为它是最经济、风险最低的策略，是创造性枯竭之后的理性选择而非道德失败；人工智能使它空前高效且普遍"，以及"这一处境早于人工智能、根源在人类创造力的限度"，见

Artism (arXiv:2512.15710)。本章第六节所指认的共有前提，在艺术学一侧正是从这一表述

中读出的。

[^5]:数字绘画从业者二〇二二至二〇二五年态度与实践的纵向追踪（吸纳逻辑是压缩劳动与时间成本、把生成结果当作构思脚手架与沟通材料），见 arXiv:2511.03117。

[^6]:

变异度的定义（"给定一组元素，其变异度是可被区分的元素个数"）及其观察者相对性的括号限定

（"若两个观察者所能作的区分不同"），见 Ashby(1956/1958)。这一句是本章第七节全部论证的支点。

[^7]:

未分之别

"只有变异度才能吸收变异度"，以及调节方必须拥有足够大的动作变异度才能把结果的变异度压到

足够小，见同上。

[^8]:

大规模生产以限制消费者选择范围来削减环境变异度（福特那句"任何颜色，只要它是黑的"），以

及互联网之后信息生态变异度按各种度量高出数个数量级、使削减这条路越来越难走通，见

Naughton

在 Edge 上对艾什比定律的评述。

[^9]:必要变异度不等式在当代多供应商与软件定义系统中的应用（每增加一家供应商即引入新的扰动状态，拉大内部动作变异度与外部事件变异度之差），见 arXiv:2506.09878。

[^10]:本章第十二节所对照的既有概念，其原始出处见

第 8–12 条。其中格尔茨的内卷化是本章最近的祖先，马奇的探索与利用是本章最近的邻居。

■ 附节·换一组学科再问一遍：有效种类数、最小对立对与可识别性

本章的论证由经济学、艺术学、工程学三家撞出。与第四章、第七章的同名附节一样，本节换一组学科把同一个问题再问一遍：可分辨的种类数是系统的自然属性，还是一项被生产出来的东西。

换上的三家是：生态学（多样性的有效种类数）、语言学（音位对立与最小对立对）、统计学（模型的可识别性）。三家都有一套成熟的、量化的答案，而三条互不相容。

生态学的位置。多样性从来不是数一数有多少个种。丰富度只数种数，不管每种占多少；而把丰度分布计进来之后，得到的是“有效种类数”——若群落中一个种占九成九、其余九十九个种合占百分之一，它的有效种类数接近于一，尽管丰富度是一百。生态学的答案是：种类数是一个连续量，由分布决定，而不是由清单长度决定。

语言学的位置。两个音在一个语言里算不算不同的类，判据不在物理差异，而在最小对立对——是否存在一对词，仅由这个差异区分意义。同一对物理差异在一门语言里区分意义，在另一门语言里根本听不出来。语言学的答案是：种类数由这个共同体是否用它区分过东西决定，而不是由差异是否存在决定。

统计学的位置。一个模型的参数可识别，当且仅当不同的参数值给出不同的观测分布。不可识别不是估计不准，是再多的数据也不能把两者分开——它是设计的性质，不是样本的性质。统计学的答案是：种类数由观测设计决定，而不是由被观测对象决定。

三条不能同真。生态学把种类数交给分布，语言学交给共同体的使用史，统计学交给观测设计。三家共同假定了一件没有说出的事——存在一个可以事后被算出来的种类数：生态学要一份丰度数据，语言学要一份词表，统计学要一个已经定好的模型。三者都在“已经有了一套观测”之后才开始工作。

推翻它的材料来自语言学自己。最小对立对的判据有一个反身的性质：一对差异之所以被记录为对立，是因为共同体曾经用它区分过意义；而共同体之所以能用它区分意义，是因为它已经被当作对立。语言接触与音位合并的历史反复表明，一个原本活跃的对立可以在一两代人之内消失——不是因为发音变了，而是因为不再有人用它区分。分辨力在这里明确地不是被发

现的，是被使用维持的；停止使用，它就消失，而且消失时没有任何事件——没有人宣布过某个对立作废。

这与本章正文的核心判断逐字重合，而它来自一门与经济学、艺术学、工程学都无关的学科。这是本节最有分量的一处：它把本章的判断从"三家碰撞的产物"提升为"在完全不同的材料上独立复现过的结构"。

由此得到本节对本章的三处推进。

其一，把种类留存率换成有效种类留存率。本章正文的读数 k 是"本期实际使用的分辨类别数除以上期"，它只数清单长度。按生态学的处理，应当把使用频次计进来：一个名义上有二十个类别、而百分之九十五的判定落在三个类别上的评分体系，其有效种类数接近三而不是二十。改写后的读数是：

> 有效种类留存率 $k^* = \text{本期有效种类数} \div \text{上期有效种类数}$ ，其中有效种类数按各类别被实际使用的频次分布计算。

k^* 比 k 硬在一处：它读得出"类别还在表上、但没有人用了"这种状态，而这正是本章所描述的那种消失的典型形态——没有人删掉任何一个类别，只是不再有判定落在上面。

其二，给"分辨是被生产的"一个可操作程序。语言学的最小对立对可以直接搬过来：要判断某个类别在这个体系里是否还是一个真类别，就找一对仅在这个维度上不同、而其余条件尽量相同的产出，看判定结果是否分开。若不分开，这个类别在这个体系里已经不再区分任何东西——不论它在评分表上写得多清楚。这条程序不需要任何新数据以外的东西，只需要一次配对抽样。

其三，一条来自统计学的、更严厉的境界。若两个类别在现行的观测设计下本来就不可识别——即无论收集多少判定记录，都不存在能把它们分开的观测——那么它们的合并不是分辨力的损失，而是把一个从来不成立的区分从表上取消。这一处削弱了本章正文：并非每一次类别减少都是病。本章正文没有区分"被使用得没了"与"本来就不可识别"，而两者的处方相反：前者要恢复使用，后者要承认它本来就不该在表上。

由此得到本节的判决性预测：在类别减少的体系里做一次最小对立对配对抽样。若配对判定确实分不开，且该维度在观测设计上本来就不可识别，则这次减少是清理而非退化；若配对判定分不开、而该维度在设计上可识别，则这次减少是本章所说的那种退化。两种情形在同一批数据上给出不同读数，因此可以分开。

最后是这次换组对本书 E 维的贡献。它不止改了读数 ($k \rightarrow k^*$)，还削弱了本章的一处主张（类别减少不必然是病），并给出一个此前没有的操作程序（最小对立对配对抽样）。一次换组能同时做到这三件，说明这一组学科不是被临时找来做装饰的。

编四

判断力从哪里来

• • •

■ 编序

前三编讲的都是失效。这一编回答一个更早的问题：那份能力本来是从哪里来的。

第九章给出第一层答案：认出什么算数的能力，是在那段不产出任何可验收之物的时间里被生产的——而在任何一份交付账上，这段时间都是空白，因此它是最先被优化掉的那一段。第十章给出第二层，也是最后一层：那份能力不由任何一方生产，它只由别处的低效活动顺带养成；因此它没有所有者，不是因为没有人愿意认领，而是因为在结构上没有一个位置可以认领它。

读到这里，前三编的所有失效会合成同一条线索：它们不是十件事，是同一件事在十个位置上的显形。而这一件事的名字，就写在第十章的标题上。

第九章

那一段谁走的——验收能力在哪里被生产

■ 摘要

人工智能进入教育之后，最常被讨论的问题是学生会不会作弊、会不会变懒、文凭会不会贬值。这些问题都预设了同一件事：一次教育经历的价值，由它的产出被验收时的样子决定。三门学科对当下的教育困境各给了一条出路，三条路各自成立，却指定了三个互不相同的终点：教育学要让能力真的长在人身上，主张出路在教学设计；经济学要让筛选装置对准真正稀缺的东西，主张出路在换掉凭据；教育哲学要让人成为能自己开口的主体，主张出路在恢复被“学习”一词挤掉的那个目的。三条路不能同时是出路——第一条必须假设练习可被设计，第二条必须假设位置先在而变的只是凭据，第三条必须假设这件事根本不可被设计。

推翻共有前提的材料来自三家自己。教育学的现场实验里，使用无护栏生成式辅导的一组，练习期成绩上升四成八，工具撤走后比从未用过的人低一成七——这段经历在它发生的时候，一切可见指标都在变好。

经济学的薪资微数据里，二十二到二十五岁在高暴露职业的相对就业下降，主因是招聘坍塌而非裁员——被砍掉的不是位置，是入口，而入口正是把新手变成老手的那条产线。教育哲学自己写下：那种召唤不必是显性的，教师的在场本身就能提出一个未被说出的问题——也就是说，它发生在没有任何产出、没有任何可验收物的时刻。三件材料都在各自领域里是常识，都被用来支持别的

。

由此得到的判断是：教育里存在一段当期不产出任何可被验收之物、而它是日后验收能力的唯一来源的时间，本章称之为哑段。生成式协助所做的事，不是让人变懒，而是把这段时间的产出提前兑现了——产出被提前领走，这段时间在所有账本上记为已完成，实际却没有被走过。于是形成一条回路：可交付的产出越漂亮，人群里能判断“漂亮是不是真的”的人越少；能判断的人越少，筛选越只能依据产出的漂亮程度；越依据产出，越依赖生成式协助。这条回路的关键性质是：它不需要参与竞争的人数增加，也不需要可得的位置减少，在总投入下降时同样加速——而经典的内卷概念要求投入持续上升。

文中给出哑段的三条性质与两个可算的读数（哑段占比、代走率）、一张二维辨别表（其中一格是关键对照：产出被前置而验收能力照样被生产时，困境根本不发生）、一句不含判断词的问话作为判据、十处兑现（其中两处反过来迫使命题收窄）、与十一种既有说法各自可分开的观察，以及一条用现成考务数据一学期内可跑完的检验。第十一节与本书同题的五篇逐条分清缺口位置：那五处都在评价一侧，本章在能力生产一侧——它回答的是那五处共同悬着的问题，判定者本身是从哪里来的。

■ 关键词

哑段；代走率；验收能力；生成式协助；入口坍塌；教育内卷；迁移测验；可干预次序
那一段谁走的

■ 一、导论：三个不该同时为真的出路

关于人工智能时代的教育困境，眼下最认真的三条出路各出自一门学科，各自都有可核对的证据，也各自都有明确的反对者。

第一条来自教育学。它把困境读作技能获得的失败：工具替学生做完了练习，练习不再产生能力，于是必须重新设计教学，让工具"给提示而不给答案"。这条路的终点是能力真的长在人身上，可干预处在工具与任务的设计。

第二条来自劳动经济学。它把困境读作筛选装置的失灵：学历这个凭据所对应的稀缺已经改变，而装置还在按旧凭据分配位置，于是必须把凭据换成能力证明、学徒制与微证书。这条路的终点是劳动力市场的匹配，可干预处在凭据与雇佣制度。

第三条来自教育哲学。它把困境读作目的的丧失：教育被"学习"这个词接管之后，只剩下效率与产出，而教育的第三个目的——让一个人成为能自己开口、自己承担的主体——既无法被测量也无法被生产，只能被召唤。这条路的终点是主体的出现，可干预处在教师的在场与关系。

三条路不能同时是出路，因为三者对"可干预处"的指定互相取消。若第三家对，第一家的护栏设计只是更精致地替学生把该他自己经过的地方填平了；若第一家对，第三家的"不可被设计"是不肯做事的托词，因为现场实验已经证明设计确实改变了结果；若第二家对，前两家都是在一个正在换血的装置内部做微调，而装置本身决定了任何教学改进最终折算成什么。三家更麻烦的地方在于，每一家的日课正是另一家诊断出的病：为一切能力找到可测的代理（教育学）是哲学侧诊断出的病根；把教育折算为凭据与回报（经济学）是教育学侧诊断出的病根；坚持有一种价值不可测量不可规模化（哲学）是经济学侧诊断出的病根。

本章不打算在三条路里选一条，也不打算把三条合起来说"要多管齐下"。三家都成立而互相取消，通常意味着它们共享了一个从未被说出的前提，而这个前提是错的。第四节给出这个前提与推翻它的材料，第五节给出替代的判断。

■ 二、文献综述

2.1 教育学：把工具装上护栏，让练习重新产生能力 Mariman, PNAS122:26,2025, 2025 年 8 月有更正）。近一千名学生被分入三组：无护栏的通用聊天界面、经过提示词改造以避免直接给答案的辅导版、以及不使用工具的对照组。结果分两段：练习期，无护栏组成绩相对提高约百分第五章，辅导版提高约百分之一百二十七；工具撤走之后的考试里，无护栏组比从未使用过的对照组低约百分之十七，而辅导版与对照组的差异不再显著。

作者由此提出的处方是设计侧的：生成式工具的部署方式决定它是拐杖还是脚手架，因此决策者必须谨慎对待界面与提示词的设计选择，以保住技能获得与长期生产率。这一处方在随后的综述文献里被广泛接

那一段谁走的

受，并与更早的认知负荷与直接教学传统接上——一条自 Sweller 以来就存在的主张：学习者的工作记忆有限，未经引导的探索会浪费在与目标无关的搜索上，因此应当给出明确的指导。

这一支的关键条件有两个：被关心的能力是可测的，练习环境是可控的。两个条件都在学校里成立，这也是它能给出现场实验证据的原因。它最脆的地方在自己的数据里：辅导版的成绩是把负效应抵消到统计上不显著，并没有产生相对于从未使用者的长期增益。也就是说，这一支目前能证明的是护栏可以止损，尚不能证明护栏可以增益。

2.2 经济学：换掉凭据，让筛选对准真正稀缺的东西字经济实验室，2025 年 8 月工作论文，2025 年 11 月与 2026 年 2 月两次更新）。研究使用覆盖数百万名劳动者的按月薪资记录，发现在生成式人工智能暴露度最高的职业中，二十二至二十五岁劳动者的就业相对下降约百分之十三，而同一批职业中的年长劳动者未见同等下降；作者进一步指出，这一下降主要由招聘的收缩驱动，而非由裁员驱动。与此并行的另一组证据显示，在人工智能相关岗位的招聘广告中，学位要求的提及率下降，而技能本身的薪资溢价高于学位。

这一支的处方是凭据侧的：既然学位所对应的那份稀缺正在改变，就应当把选拔的依据从学位转向可验证的技能，扩大学徒制、职业教育、微证书与在职训练的通道。

必须一并写出的是，这一支目前正处在活跃的争论中。一组研究者主张，二十二至二十五岁群体的相对弱势是四十年来最陡的紧缩周期的可预测后果，而非技术位移的早期信号；另一组基于丹麦细颗粒数据的研究报告了对收入与工时的精确零效应。作者本人也在更新版中承认，在纳入最广的控制变量后，下降的显著时点才落在2024年，更早的部分很可能受非人工智能因素影响，并明确表示该看板衡量的是暴露度与就业趋势之间的相关，而非因果。本章引用这一支，取的是那条被各方都不否认的事实——收缩发生在入口而不是存量——而不取任何一方对总量效应的判断。

这一支的关键条件是：位置先在。它的整套语言以位置、凭据、匹配与回报为单位，因此它能精确地讨论谁被分到哪里，而无法讨论一个尚未成为任何人的人是怎么被生产出来的。

2.3 教育哲学：把教育从"学习"手里要回来 (World-Centred Education, 2022)。他的论点是：当教育的语言被"学习"这个词接管，关于目的、内容与关系的问题就变得难以提出，因为"学习"是一个过程词，它本身不包含"为了什么"。他因此坚持教育有三个目的：资格化（获得知识、技能与判断的形式）、社会化（成为某个秩序的一部分）、以及主体化（成为一个主体，即在思与行上更加自主、更加独立）。主体化不是社会化的延伸，而是它的对立面。

关于主体化如何发生，他借用了一个古老的说法：教育是一种"促其自为"的召唤，是以非确认的方式向一个人说话，使他开始把自己看作一个能够回应的人。近年有研究把这一框架接到生成式工具上 (Studies in Philosophy and Education, 2026)，并指出：认知卸载与较低的批判性思维分数之间的相关，尤其

那一段谁走的

在十七至二十五岁人群中，与主体化所要的那种自主恰好相反；同时强调，这种召唤不必是显性的，教师的在场本身就能向学生提出一个未被说出的问题。

这一支的关键条件是：这件事不可被生产。它的处方因此只能是消极的——不要用效率语言接管教育，不要把教师降格为学习的促进者。它最脆的地方也在这里：主体化被同行批评为"相当难以捉摸"，它没有量纲，说不出发生了没有，因而进不了任何课表、任何考核、任何预算。而进不去的结果，是它在实践中把时间让给了效率。

■ 三、冲突定位：三条路的终点互相取消

把三条路放到同一个具体处境里，分岔立刻出现。

设一门大学一年级的写作课，十六周，期末交一篇论文。三条处方各自给出的动作是：教育学要求把工具装上护栏，让学生在提纲、修改、论证检查各环节得到提示而不得到成品；经济学要求把评价的重心从这门课的成绩转向可携带的能力证明，让学生的产出能在课外被雇主

读取；哲学要求教师停止把自己当作流程的促进者，重新以自己的判断在场，对学生的具体文字说“我不同意”。

三条动作在这一门课上不能同时执行到底。护栏要求过程被切分为可控的环节，而可控的切分正是哲学侧诊断的“把召唤变成流程”；能力证明要求产出可被外部读取，而可被外部读取的产出正是教育学侧诊断的“练习被产出替代”；教师以自己的判断在场，产生的是不可携带、不可比较的评价，正是经济学侧诊断的“凭据不可读”。

更要紧的是三者对“什么算成功”的定义互相取消。教育学的成功是撤走工具后学生仍能做；经济学的成功是学生毕业后能被匹配到位置；哲学的成功是学生成为一个会自己提问的人。这三个成功可以同时发生，但三条路各自把另外两个当作副产品——而副产品是不被保护的。当十六周的时间不够时，被牺牲的一定是没有被这条路指定为终点的那两样。

到此为止，三家的分歧看上去像是价值分歧：你要能力，我要匹配，他要主体。本章的主张是它不是价值分歧。三家在一件事上完全一致，而正是这一致处出了问题。

3.2 第二个处境：一家企业的新人培养同样的分歧在企业里更硬，因为那里有预算。设一家中型专业服务机构，每年招二十名应届生，培养周期三年。

教育学的处方是把新人的任务装上护栏：允许用协助，但要求他们提交推理过程而非只提交结果。经济学的处方是把新人招聘的依据从学历改为技能测试，并把培养周期压缩——既然工具已经能交付初级产出，就不必再养三年。哲学的处方是让资深者以自己的判断在场：花时间当面对新人的具体做法说“不行”，而不是把反馈交给流程与模板。

那一段谁走的

三条处方在预算表上直接冲突。护栏要求资深者审阅推理过程，这比审阅结果贵得多；压缩培养周期要求减少资深者投入；当面反馈要求资深者的时间以最不可计量的方式被占用。任何一家机构都只能选一条，而选择的结果不是“侧重不同”——选了第二条的机构，三年后没有能审阅推理过程的人，因而也无法执行第一条；选了第一条的机构，其人均产出在前三年低于同行，因而在第二条的逻辑下会被判为低效。

这个处境说明三家的分歧不是可以调和的权重问题。它们在同一笔预算上要求相反的动作，而做出选择的人手里没有任何一个指标能告诉他哪一条更对——因为三条路各自的成功都在自己的终点上兑现，而那三个终点的兑现时间相差数年。

■ 四、共有前提，与三件推翻它的材料

4.1 共有前提三家共有的前提是：一次教育经历的价值，由它的产出被验收时的样子决定。

这个前提隐藏得很好，因为三家争的正是"验收什么"。教育学要验收撤走工具后的表现，经济学要验收可携带的凭据，哲学要验收一个人是否已经成为主体。三家争验收的对象、验收的时点、验收的主体，却没有一家问过：这次经历里有没有一部分，在原理上不产生任何可被验收之物，而它恰恰是别的一切能被验收之物的来源。

这个前提之所以能长期隐身，是因为在生成式协助出现之前它基本无害。一个人自己做一道题所花的时间，与他做出来的那份答案，从前是绑在一起的：答案的质量大体反映了那段时间里发生了什么。前提在这种绑定下是对的。生成式协助做的事，就是把这两样解绑。

4.2 教育学自己的材料：伤害期的一切可见指标都在变好回到那项现场实验的两个数字。无护栏组在练习期成绩相对提高约四成八；工具撤走后比从未使用者低约一成七。研究者把它读作"拐杖效应"，用以论证护栏的必要。

但这两个数字合起来说的是另一件事：如果只看当期的可见指标，无护栏组是优等生。

出勤在、作业

在、正确率在、进度在，全部向好。伤害只在工具撤走之后、且只相对于一个从未接受过帮助的对照组，才第一次显形。换句话说，这段经历的价值在它发生的时候是不可验收的——不是难以验收，是原理上不可验收，因为那时它的一切外部表现都指向相反的

。

教育学知道这件事，因为这正是他们做出来的发现。但他们把它记在"工具设计"这一栏，没有把它记作"验收本身在这一段失效"的证据。

4.3 经济学自己的材料：坍塌的是入口，不是位置同样地，回到那组薪资数据里被各方都不否认的那条事实：年轻组的相对下降由招聘的收缩驱动，而不是由裁员驱动；同一批职业里年长组不受同等影响。

那一段谁走的

经济学把它读作劳动需求的结构变化，并据此讨论供给侧应当如何调整凭据。但这条事实还说了另一件被记在别处的事：被自动化掉的是新手做的那些任务，而新手正是靠做那些任务变成老手的。位置还在——年长组稳住了就说明位置还在；消失的是那条把人从新手变成老手的产线。这一门有一个词描述它，叫做通道压缩，通常出现在关于长期人力资本的讨论里。它被记在劳动供给的账上，没有被记在"经验是怎么被生产出来的"这本账上，因为后者不是这门学科的记账单位。

4.4 哲学自己的材料：召唤发生在没有产出的时刻第三件材料是这一支自己写下的那句话：召唤不必是显性的，教师的在场本身就能提出一个未被说出的问题。这句话被用来论证教师不可替代。

但它同时说明：这件事发生在没有任何产出、没有任何可验收物的时刻。既没有作业，也没有回答，也没有可被记录的互动——只有一个问题被提出而尚未被回应的时间差。这一支说了这句话，却依然在与另外两家争"教育应当产出什么"，因而把自己最要紧的观察留在了同一个前提之内。

4.5 三件材料的共同形状三件材料指向的是同一类东西：一段当期不产出任何可验收之物、而后期一切可验收之物的能力由它生成的时间。

教育学称之为独立练习，经济学称之为入门岗位的重复劳动，哲学称之为被召唤而尚未回应的

那一刻。三家都承认它重要，但三家的账本都不给它一格——因为按定义它不产出。

■ 五、哑段：命名、性质与读数

5.1 定义哑段：一段经历中，当期不产出任何可被验收之物、而它是日后验收能力的唯一来源的那一部分时间。

之所以叫哑，是因为它在当期不发出任何可被读取的信号：没有作业、没有成绩、没有产出、没有里程碑。它不是"没在学习"，恰恰相反，它是学习真正发生的地方；它只是不发信号。

一段时间是不是哑段，与它是否困难无关，与它是否有趣无关，与它是否有人陪伴也无关。判据只有一条：这段时间里，那件事是不是这个人自己做的。有人在旁边讲解、指出方向、给出反例，仍然是哑段；有人替他把那一步做完了，就不是。

5.2 三条性质性质一：当期指标为零或反号。

处在哑段中的人，看起来比不处在哑段中的人做得慢、错得多、进度落

后。这不是哑段的副作用，是它的定义特征——如果当期指标好看，说明产出已经从别处来了。

性质二：不可代做。

哑段一旦被代做，它就不再是哑段，而变成一段"外观相同、内容为零"的时间。这

一条与前一条合起来构成一个恶劣的结构：代做既提高了当期指标，又消灭了这段时间的全部作用，而两件事在当期都不可分辨。

那一段谁走的

性质三：不入账。

没有任何现行制度为哑段设有字段。课表记的是课时与学分，考核记的是成绩与产

出，雇佣记的是岗位与年限，简历记的是成果与头衔。哑段在所有这些账本里都只能以别的东西的形态出现——通常是"效率低"。

5.3 两个读数哑段占比=一个学习或工作过程中，当事人在无外部生成协助的条件下自行推进的时长÷该过程的总时长。

代走率=该过程中原本属于哑段、而实际由外部协助代为完成的时长÷该过程哑段的应有时长。

两个读数都不要要求知道当事人是否"依赖"工具，也不要要求判断他的动机，只要求记录时间的归属。它们与既有的任何一个教育指标都不共线：一个哑段占比很低的班级，其平时成绩、作业完成率、课堂参与度可以全部高于哑段占比高的班级——这正是第 4.2 节那两个数字所显示的。

需要写明的是，这两个读数在当前的技术条件下不能被精确测量，只能被制度性地保障。这不是本章的退让，而是命题的一个直接推论：哑段按定义不发信号，任何试图从信号中反推它的做法都会被更好的信号所欺骗。可行的路径是安排而不是测量——见第十二节。

5.4 二维辨别表把两个变量交叉：产出是否被前置（协助是否替当事人交出了当期产出）×

验收能力是否仍在被生产

（这段时间是否仍在生成日后独立判断的能力）。

验收能力仍在被生产产出未被前置

验收能力停止被生产

甲：传统的慢。指标不好看，能力在长。

丙：空转。既无产出也无所得，纯粹的浪费。

产出被前乙：关键对照。产出由协助交出，而当事人另丁：靶格。产出漂亮，能力停产，而两者在置

有安排的哑段。

当期不可分辨。

乙格是本章最要紧的一格，因为它证明了困境不是由生成式协助本身产生的。产出被前置而验收能力照样被生产的情形是存在的，且并不罕见：一位资深工程师用生成式工具写完样板代码，同时把省下的时间用于读别人的架构决策记录；一位译者让工具出初稿，自己逐句对照原文重译。在这两个例子里，产出被前置了，哑段被移动了位置但没有被消灭。

这一格排除了两个常见的读法。它排除了"工具有害论"——工具本身不产生丁格。它也排除了"素养论"——乙格与丁格的区别不在于当事人是否懂得正确使用工具，而在于制度有没有给那段被移走的哑段留一个位置。资深工程师之所以能停在乙格，是因为他的岗位已经把"读架构决策记录"算作工作；一年级学生停不在乙格，是因为没有任何课表把"自己先做一遍再看答案"算作学时。

那一段谁走的

5.5 这不是一个新读数，而是一类新的存量有必要区分本章的主张与"学生用工具跳过了练习"这个已被充分讨论的观察。后者是对个体行为的描述，它的处方是纪律或设计。本章主张的是：由被代走的哑段构成的存量是一个真实的、可累积的、且决定系统性质的东西——它的存量决定了这个人群中还有多少人能够判断一份产出是不是真的。

判据是那句标准的问话：删掉这个读数之后，那样东西是不是就不存在了？答案是否定的。即使没有人使用"哑段"这个词，"这批人里还剩多少人能不靠工具判断这份东西"这件事照样存在，照样在决定招聘线、评审线与录取线怎么划。本章所做的，是给这样东西一个可以被记账的位置。

5.6 哑段不是"练习"有必要把哑段与"练习"分开，否则本章会被读成一句老生常谈。

练习是一个以熟练为目的的重复过程，它的成败由熟练程度衡量，而熟练程度是可测的。哑段不以熟练为目的，也不必重复；它可以只发生一次，可以是一次失败的尝试，可以是一段完全走错方向的时间。一个学生花三小时写了一份最终被自己废掉的提纲，这三小时的练习价值可能为零，哑段却是完整的。

区别的要害在于产物。练习的产物是熟练，熟练日后会显形为更好的产出；哑段的产物是判断的锚点——这个人亲自碰过一次"做不出来"是什么样子，因而日后看到别人（或工具）交出的东西时，知道那个东西是不是从"做不出来"的那一侧走过来的。这个锚点不表现为任何一项可测的技能，只表现为一种能力：认得出没走过那一段的产物。

这也解释了为什么熟练可以被代做而判断不能。工具能替一个人完成一万次重复，从而在任何熟练度测验上得到高分；但工具不能替他碰一次"做不出来"，因为工具那一侧根本没有"做不出来"这回事——它总是交得出东西。第 4.2 节里无护栏组撤走工具后落到对照组之下，落的正是这个锚点：他们不但没有获得，还失去了原本会由挫败带来的那份校准。

因此本章不主张"多练习"。多练习在有工具的条件下会被工具吸收，练习量越大被代做的比例越高。本章主张的是保住一段独立推进的时间，哪怕这段时间里什么也没做成。

■ 六、机制：一种不需要人更多、位置更少的内卷

6.1 四步回路第一步：产出被前置。生成式协助把原属于哑段的产出提前交付。当事人的当期指标改善。

第二步：哑段在账上完成。

由于所有制度只记产出，这段时间被记为已完成。没有任何环节会发现它没有被走过。

第三步：验收者变少。

经过若干轮之后，这一人群中能够不依赖协助判断一份产出真伪的人数下降。注

意这里下降的不是"平均能力"，而是一个特定的能力：判断的能力。它下降得比生产的能力更快，因为生产可以被代做而判断不能。

第四步：筛选更依赖产出。

那一段谁走的

当能判断的人变少，选拔与评审就只剩下依据产出的表面质量。而表面质量

的提升成本已经接近于零，于是所有人都必须把产出做得更漂亮，也就必须更多地使用协助——回到第一步。

6.2 与经典内卷的分水岭：投入的方向格尔茨意义上的内卷要求投入持续上升而边际报酬持续下降。本章描述的回路不要求这一点，甚至与之相反：在这条回路里，个体投入的时间下降，产出的质量上升，而系统性质持续恶化。

这是一条判决性的分离线。若观察到某个体系中参与者的平均投入时间在下降、单位产出的表面质量在上升、而筛选标准仍在收紧，则格尔茨式的解释不成立，而本章的机制成立。反过来，若某体系中筛选收紧总是伴随投入时间上升，则本章的机制在该体系中不是主因。

6.3 加速条件与减速条件这条回路不是普遍的。它需要三个条件同时成立：条件一：产出与哑段可解绑。

存在一种协助，能够在不经过哑段的情况下交出与经过哑段者相似的产出。在纯体能训练里这个条件不成立——没有办法替一个人做完引体向上。

条件二：验收依赖同一种能力。

判断一份产出是否合格，需要的正是生产它所需的那种能力，且没有独

立于该能力的外部结果反馈。在有硬结果反馈的领域这个条件不成立——一座桥塌不塌，不需要评审的判断力来确定。

条件三：验收者从被验收者中产生。

今天的评审、导师、面试官，是昨天走过哑段的学生与新人。这一

条是回路闭合的关键：如果验收者可以从外部无限供给，那么第三步就断了。

三个条件同时成立的领域，恰是当下讨论最集中的那些：写作、编程、法律文书、学术评审、设计、医学的诊断推理训练。三个条件不同时成立的领域，则不必套用本章的诊断——见第十三节。

6.4 为什么判断先垮，生产后垮回路的第三步是全篇最容易被质疑的一步：凭什么说判断的能力比生产的能力衰减得更快？如果两者同速，回路就不成立，因为筛选照样有人能做。

理由有三条，且三条都可被单独检验。

其一，代做的可得性不对称。

生产有现成的代做者，判断没有。要交一份产出，可以让工具交；要判断

一份产出是否合格，也可以让工具判断——但判断工具的输出本身又需要判断，这一步没有出口。于是在同样的协助强度下，生产环节的哑段被代走的比例，系统性地高于判断环节；但判断能力的原料恰恰来自生产环节的哑段。原料被代走，产品自然停产。

其二，反馈的时点不对称。

那一段谁走的

生产的失败当期显形：交不出东西，或交出的东西被退回。判断的失败不当

期显形：一个判错的评审，其错误要等到被判错的那件事在别处产生后果时才可能被发现，而多数时候永远不被发现。因此生产能力的衰减会被系统自动纠正，判断能力的衰减不会。

其三，岗位的掩护不对称。

一个生产能力衰减的人会很快失去岗位；一个判断能力衰减的人往往正在升

职——他从生产岗转到了评审岗、管理岗、导师岗，而这些岗位的绩效恰恰不以判断的准确性计量。系统因此把判断力最可疑的人放在了判断的位置上，并且是以奖励的形式。

三条合起来给出一个可查的推论：在同一个机构里，若比较同一批人的生产能力与判断能力随年限的变化，判断能力的相对下降应当出现得更早、幅度更大，且与该机构评审岗的绩效口径相关。若观察到两者同步，或判断能力反而更稳，则第三步为伪，整条回路不成立。

■ 七、逐领域兑现

7.1 课堂：护栏止损，但不产生增益第 2.1 节那项实验的辅导版做到的是把负效应抵消掉。本章对这一结果的解释是：护栏减少了代走率，但没有增加哑段的绝对时长。学生仍然在同一

段时间里完成同一份产出，只是提示来得更迂回。而哑段的作用与提示的形式无关，与"这段时间是谁在推进"有关。

这给出一个可检验的推论：在护栏版之外再加一组，其唯一差别是在同样的总课时里，把用于产出的时间压缩，把省下的时间安排为无协助的独立推进，则该组在撤走工具后的迁移测验上应当高于护栏版。若不低于，本章错。

7.2 劳动入口：入门岗位买的不是产出一个初级岗位的产出，今天大多可以由协助以更低成本交付。若按产出定价，这些岗位应当消失——薪资数据显示的招聘收缩正是这个逻辑的执行。

本章的判断是这个逻辑记错了账。企业购买初级岗位，从来买的不是那份产出（那份产出在任何时代都可以由更资深的人更快做完），而是五年后能够审查资深产出的那个人。这笔支出是资本性的而非费用性的，它今天被记在费用里，因此在成本压力下最先被削。削掉它的机构，在若干年后会发现自己没有能够审查协助输出的人，而那时唯一的补救是从外部购买——即从尚未削掉入口的机构那里购买，而如果所有机构都削，则市场上没有这样的人。

这一节给出的推论是：削减入门岗位的机构，其若干年后的返工率、事故率或复核成本应当高于未削减者，且该差异应当在资深人员的自然更替周期之后才显形。这条推论的检验周期长，因此第八节另给一条短周期的。

那一段谁走的

7.3 学术训练：审稿人是从哪里来的本书此前的几篇处理的是评审那一侧：产出太多、认领的成本上升、被拒稿件在无人记账处合成基线。这些都预设了一件事：有一群人在做判定。本节问的是这群人从哪里来。

答案是：他们是十年前自己写过、改过、被拒过、又自己重写过的博士生。他们的判断力来自那段没有任何产出的时间——写了删、删了写、一稿废掉。如果这段时间被协助填满，产出侧的通量会大幅上升，而十年后可用的审稿人会大幅下降。两件事的时间尺度不同，因此在任何一年的横截面数据里都看不到关联。

7.4 医学：一处使命题收窄的兑现医学诊断推理的训练似乎完全符合三个条件，但实际情形迫使本章收窄。医学有一样别的领域没有的东西：结果。病人好转或不好转，是独立于诊断者判断力的反馈。这使得第 6.3 节的条件二不完全成立。

因此本章的命题在医学中只对结果反馈延迟很长或极为稀疏的子领域成立——例如慢病管理的长期方案选择、精神科的诊断分型——而在急诊与外科这类结果快速显形的子领域不成立。这不是本命题的例外，是它的适用边界，也说明第 6.3 节的三条件不是修辞而是承重的。

7.5 第二处收窄：翻译与代码评审在翻译与代码评审两个行当里，出现了一个反向的观察：使用协助之后，从业者的评审能力在一部分人身上上升了。原因是他们从"自己写"转到了"读工具写的并挑错"，而挑错本身是一种高强度的验收训练。

这个观察削弱本章命题的强度。它说明：产出被前置并不必然消灭哑段，若被前置出来的时间被安排为验收训练，哑段实际上被换了一种形式保留下来。这正是第 5.4 节的乙格。据此，本章的处方从"保护自己动手的时间"修正为更准确的一句：保护那段由本人独立形成判断的时间，无论这个判断是关于自己的产出还是关于别人的产出。

7.6 体育训练：一处反证体育训练里，第一条件不成立：没有任何协助能替一个人做完引体向上，产出与自身推进无法解绑。按本章的机制，这个领域不应当出现同类困境。事实也是如此——体育训练领域当下的争论集中在负荷、恢复与伤病，而不是"谁真的练过"。这一处反证的价值在于它证明本章的机制不是万能框架：它有一个明确的开关，开关关闭时机制不发生。

7.7 律所初级文书：回路最完整的一处三条件在这里全部成立：文书可被生成式协助交付；判断一份文书是否可用需要的正是起草它的那种能力，且多数文书不产生可读的外部结果（案子输赢受太多因素影响，无法回溯到某一份文书）；而合伙人是从初级律师里长出来的。这一行当因此应当是回路的最快样本，观察它的两个读数——初级律师独立起草的时长占比，以及中层律师说明"这一稿为什么不行"的具体程度——应当先于其它行当出现下降。

那一段谁走的

7.8 新闻编辑部：一处减速新闻业的验收者有相当一部分从外部招聘，而不只从本编辑部的实习生里长出来，第三条件被部分打断，回路减速。这提供了一个自然对照：同一时期内，主要从内部养成编辑的机构与主要从外部招聘编辑的机构，其稿件判定质量的下降速度应当不同。若无差异，第三条件不承重。

7.9 工业质检：结果打断质检的结果由物理世界给出，第二条件不成立。这一处与医学的急诊侧同类，一并说明：凡是有硬结果反馈的地方，本章的机制被截断。这也反过来给出一条实践建议——在可能的范围内为验收引入哪怕很稀疏的硬结果反馈，比任何关于工具使用的规定都更有效。

7.10 驾驶培训：一处半开模拟器可以代练，但不能代考，且路面结果是硬的。第一条件部分成立、第二条件不成立，因此回路只在理论科目里发生，而在实操科目里不发生。这一处的价值在于它把两个条件的作用分离开：在同一个培训体系内部，两类科目走出了不同的轨迹。

7.11 十处合观十处兑现里，三条件全齐的四处（课堂、律所初级文书、学术训练、慢病管理）回路完整；条件被打断的六处各以不同方式减速或不发生。这说明本章给出的不是一个诊断标签，而是一套逐项可判的开关：拿到任何一个新领域，先问三个问题——产出能不能与自

身推进解绑、验收有没有独立于该能力的结果反馈、验收者是不是从被验收者里长出来的。三问全是，则机制在场。

■ 八、判据与可证伪条件

8.1 主条（不含判断词）把一门课程或一段训练的哑段占比压低，撤走协助之后的迁移测验成绩下降多少？

若在控制课时总量、任务难度与初始水平之后，哑段占比与迁移成绩之间不存在剂量—反应关系，则本章的命题为伪。

8.2 最强竞争解释与如何分开最强的竞争解释是合意困难：Bjork 一系的研究早已表明，增加提取难度会提高长期保持。有人会说，哑段无非是难度的另一种说法。

分离线是：合意困难要求难度，哑段只要求无代做。

一段简单但由本人自己推进的时间是哑段，而不是

合意困难。因此有一个直接的对照设计：在两组之间保持任务难度完全一致（用同一份题目、同一套评分），只改变一件事——一组自己先做十五分钟再看解答，一组直接看解答再自己做十五分钟。两组的困

那一段谁走的

难程度相同，总时长相同，只有哑段的位置不同。若两组迁移成绩无差异，则本章可被合意困难吸收；若有差异，则不可。

8.3 一条低成本、一学期可跑完的检验大多数高校的课程系统已经记录了作业的提交时间戳与编辑历史。取同一门课的两个平行班，对其中一个班把作业改为"课堂内四十五分钟无网络完成初稿，课后可用任何工具修改并注明修改处"，另一班照旧。

期末用同一份闭卷迁移测验。所需的数据全部已在既有系统内，无须新增采集，一学期可完成。

8.4 轮次预测若本章的机制成立，则下一轮变化的发起处应当从产出侧换到验收侧：先出现的不是"学生交不出好作业"，而是"评审者说不清为什么这一份不如那一份"。若在未来若干年里，先出现的仍然是产出侧的抱怨（作弊、雷同、质量下滑）而验收侧无变化，则第六节的回路次序为伪。

8.5 赌注 2029 年 12 月 31 日之前，至少三所高校或三家大型雇主公布可核对的"无协助时段"记账制度——即在课程学分或岗位职责中，为一段明确的、不使用生成式协助的独立推进时间设立字段，并说明如何核验。

写死不算命中的情形：只发布人工智能使用政策、学术诚信声明、倡议性白皮书、或仅在考试环节禁用工具的规定，均不算。必须是为哑段本身设立的时间字段，而不是对工具的禁令。

8.6 第二条：一个反向预测除主条之外，本章另给一条方向相反的预测，用以避免"怎么看都对"。

若本章成立，则在验收训练被制度性加强的行当里，即便产出被前置的程度更高，验收能力也应当不降反升。第 7.5 节的翻译与代码评审已给出初步迹象。据此可检验：在两个产出前置程度相近的行当之间，比较其从业者的验收能力变化，差异应当由"有没有把读别人的东西并挑错算作工作"来解释，而不由工具使用强度来解释。

若观察到验收能力的变化主要由工具使用强度解释、而与验收训练的制度安排无关，则本章的处方方向错误——那意味着该管的确实是工具，而不是时间。

■ 九、三家各自为何到不了这一条

三家到不了，不是因为疏忽，而是因为各自的问题结构恰好使这一步成为不必要的。

教育学到不了，因为它的整套方法要求可测。现场实验必须有当期可读的因变量，而哑段按定义当期不发信号。它能测出护栏的效果，测不出哑段本身；它甚至无法把哑段设为自变量，因为设置它就意味着要求一组学生在一段时间里不产出任何可评的东西，而这在伦理审查与教务安排上都难以通过。

那一段谁走的

经济学到不了，因为它的记账单位是位置、凭据与匹配。哑段既不是位置也不是凭据，是一段产出为零的时间；在生产函数里，一段产出为零的投入只能被记为闲置。这门学科拥有讨论人力资本积累的全部工具，却把积累的度量绑定在教育年限与经验年限上——而这两个量在哑段被代走时完全不变。

哲学到不了，因为它坚持这件事不可被制度化。这一坚持在概念上是正确的：主体化确实不能被生产。但由此得出的实践姿态是消极的，而不可被排进课表的东西不会被排进课表。于是这一支在实践中把哑段让给了效率，并且是以保卫它的名义。它被同行批评为难以捉摸，正是这一后果的表现。

本章与三家的关系因此不是替代而是补位：它接受教育学关于设计有效的证据，接受经济学关于入口正在收缩的事实，接受哲学关于主体化不可被生产的判断，而把三者共同悬置的那段时间取出来单独记账。哑段不能被生产，但可以被安排——这是本章与哲学侧唯一的实质分歧，也是它能给出可干预次序的原因。

9.4 为什么这一条只有把三家放在一起才看得见单看任何一家，第四节那三件材料都不构成异常。

只看教育学，"练习期指标向好而撤走后变差"是一个关于工具设计的发现，处方清楚，没有任何东西迫使人去追问验收本身。只看经济学，"收缩发生在入口"是一个关于劳动需求结构的事实，它在这门学科里有现成的解释位置，也没有任何东西迫使人去追问入口里发生的是什么。只看哲学，"召唤发生在没有产出的时刻"是一句关于师生关系的论断，它被用来反对效率语言，同样没有任何东西迫使人去问那段时间在别的账本上怎么记。

三件材料要成为异常，需要一个共同的量纲，而这个量纲在三门学科里都不存在：教育学的量纲是成绩与迁移，经济学的是就业与工资，哲学没有量纲。时间的归属这个量纲之所以在三门里都缺席，是因为它对三门都不重要——教育学关心那段时间产生了什么，经济学关心那段时间值多少钱，哲学关心那段时间里人与人之间发生了什么。谁也不关心那段时间是谁走的。

把三件材料并排放置，三者的共同形状才显出来：三处都指向同一类时间，三处都以不同的名义把它记成了别的东西。这也是本章不能由任何一家单独推出的原因——它不是三家论点的综合，而是三家各自的记账遗漏在同一处重合。

■ 十、模板层的对质：这句话的形式早有正主

把本章的承重命题剥掉教育与人工智能的一切名词，只留下形式，得到的是：一个体系越是高效地交付结果，越是没有人能判断结果。这个形式不是新的，它至少有五个正主，必须逐一对质，否则本章只是给一个旧模板换了一层皮。

自动化的讽刺（Bainbridge,1983）是最贴的一个。它指出：自动化把操作员从执行者降为监控者，而监控不产生技能，于是恰恰在最需要人接管的时刻，人已经失去了接管的能力。分离线在于谁退化：自动化的讽刺讲的是同一个人在同一岗位上的技能衰减，本章讲的是人群的代际再生产——不是老手退化了，是新手不再被生产出来。判决性对照预测：若在岗者的技能水平保持完好（可用同一批人历年的独立测验

那一段谁走的

证明），而新一代的独立判断能力仍在下降，则本章不可还原为自动化的讽刺；反之，若两者总是同步下降，则本章可被它吸收。这一条对照在航空业有现成数据可查，因为该行业既有在岗者的定期复训记录，也有新飞行员的养成记录。

去技能化（Braverman,1974）要求存在一个从中受益的施加者：管理层有意把技能从工人手中剥离，以降低议价能力。分离线在于有没有施加者。哑段的消失没有施加者——它是当

事人自己为了改善当期指标而做出的、在每一步上都正确的选择，也是教师、机构与雇主各自为了改善各自指标而做出的、在每一步上都正确的选择。判定形式：找不出任何一个环节，其行为人是在牺牲自己的指标去促成这件事。

专家逆转效应（Kalyuga et al., 2003）讲的是指导量与学习者水平的交互：对新手有益的指导对专家有害。它与本章方向相反——它警告的是给专家过多指导，本章警告的是给新手代做。两者可以同时成立，因为它们说的是两件事：前者关于指导的量，后者关于时间的归属。

洛斯悖论（Miller, 1990）这一族，共同要求存在一项被推到极端的成功能力，而后该能力成为负担。分离线在于有没有那项能力：本章不要求任何成功，一个从未成功过的体系照样会发生哑段被代走。判定形式：事后能否指出“就是那项曾经让它成功的东西害了它”；在本章的机制里指认不出来，因为被代走的东西从来没有成为过任何人的强项。

必须写明的是：本章在这个形式层上没有任何新东西。增量只在两处——一是把病灶从“能力退化”移到“能力停产”，二是给出了一个不依赖动机、不依赖施加者、只依赖时间归属的读数。若这两处被证明也可由上述任一家推出，则本章应当被撤回。

■ 十一、与既有说法的分界

内卷化（Geertz, 1963）：要求投入持续上升而边际报酬递减。本章的机制在投入下降时同样发生。判定形式：观察参与者平均投入时间的方向。

信用膨胀（Collins, 1979）：要求凭据先被颁发，而后其价值被稀释。本章的机制不需要任何凭据被颁发，它在一个完全没有文凭制度的学徒体系里同样发生。判定形式：合格者人数是否变化——信用膨胀预测合格者增多而线上升，本章预测合格者可以不增多而线照样上升，因为抬线的是判定能力的下降而不是供给的增加。

信号理论（Spence, 1973）：要求存在一个成本差异，使得高能力者发信号更便宜。生成式协助的作用恰恰是把这个成本差异抹平，因而在其框架内应当预测分离均衡消失、信号失效。本章预测的不是信号失效，而是验收者消失——这是两件事：信号可以完全有效，而没有人能读。判定形式：问被评价者还是评价者出了问题。

位置性商品（Hirsch, 1977）与赢家通吃：要求位置的稀缺性先在。本章的机制在位置不稀缺时照样发生，因为它切的是能力的生产而不是位置的分配。判定形式：在一个位置充足的体系（例如内部晋升不设名额的机构）中观察验收能力是否仍在下降。

那一段谁走的

柠檬市场（Akerlof, 1970）：预测信息不对称导致优质品退出、市场萎缩。本章预测的不是市场萎缩，而是判定成本从评价方转出、总产出继续上升。判定形式：看总量方向。

注意力经济（Simon, 1971）：把稀缺定位在注意力上。分离线在于本章关心的不是注意力被分配到哪里，而是注意力被投入之后有没有在生成判断力——一个人可以全神贯注地看协助生成的解答，注意力完全在场，哑段却是零。

认知卸载（Gerlich, 2025 一系）：报告了卸载与批判性思维分数之间的相关。这是本章在个体层面的最近邻证据，但它是相关性的、个体层面的，并且以自评量表为主。本章的增量是把它接到时间归属这一可安排的变量上，并给出群体层面的后果。

合法的边缘性参与（Lave&Wenger, 1991）：要求共同体与真实任务。哑段不要求共同体，一个人孤立地做题同样构成哑段。判定形式：把共同体拿掉，看效应是否仍在。

通道压缩：这是本章最近的站外邻居之一，出现在关于长期人力资本与验证能力的讨论中。分离线是学科归属与量纲：它讲的是劳动市场上资深判断力的供给，本章把它推到教育侧，并给出一个在学校内即可安排与检验的读数。

技术债（Sculley et al., 2015 一系）：形式上很像——今天省事，日后加倍偿还。分离线在于债主是谁。技术债的债务人与债权人是同一个组织，因此它可以被自己盘点、自己偿还；哑段被代走所欠下的东西，偿还者是下一代，而下一代无从知道自己欠了什么。判定形式：能否由当事人自己在事后盘点出来。

古德哈特定律：要求存在一个被当作目标的度量而后被博弈。本章的机制不需要任何人博弈任何指标——所有人都诚实，回路照样运行。判定形式：把博弈动机完全去掉（例如无评分的自学场景），看效应是否仍在。

"学习化"批评本身：本章引用的哲学侧论点是这一批评的一部分，因此需要说清增量在哪。学习化批评的靶子是语言——用"学习"接管"教育"这个词，会使目的问题无法被提出。本章的靶子是账目：即便一个机构完全接受该批评、在文件里改用教育的语言，只要它的时间账仍然只记产出，哑段照样被删除。判定形式：找一所在理念上完全站在学习化批评一侧的学校，看它的课表里有没有一段不产出的时间。以本章的观察，通常没有。

"人工智能素养"话语：当前最主流的处方是培养学生"正确使用工具的能力"，包括提示技巧、结果核查、伦理意识。本章与它的分离线在第 5.4 节的乙格与丁格之别：素养话语假设两格的差别在于使用者的知识与态度，本章主张差别在于制度有没有给被前置出来的时间一个去处。判决性对照：在两组素养水平相同（用同一套素养测评筛选）的学生之间，只改变时间安排，若迁移成绩仍有差异，则素养解释不足。

■ 十二、与本书同题诸篇的分界

本书在同一天里，就“人工智能时代的内卷困境与出路”这一道题出过五篇，三个学科取的是经济学、艺术学与工程学。五篇的缺口位置各不相同，合起来是一条评价回路：生产者一侧（认领的成本）、关口一侧

那一段谁走的

（判定的权属）、未阅一侧（否决的记录）、参照一侧（未被接受的产出如何合成基线）、承担者一侧（核实这份工作由谁来做）。

本章的缺口不在这条回路上的任何一处，而在它的上游：那五处全部预设了有一群人在做判定，本章问的是这群人是从哪里被生产出来的。因此本章与五篇的关系不是重复也不是并列，而是相接。

与参照一侧那一章的相接最紧：那一章给出的读数是把上一轮未被接受的产出移出参照集之后合格线会降多少，它解释了线为什么被抬高；本章解释的是线被抬高之后为什么没有人看得出它抬错了。两者的判定形式可以联合使用：若某体系的线在抬高而判定者对“为什么这一份不如那一份”的说明能力同步下降，两条机制同时在场。

与承担者一侧那一章的分歧要写清楚。那一章主张核实的负担从评价方整体转移到被评价方，什么都没有多出来，只是同一份工作换了人做。本章主张的是这份工作正在失去能做它的人。两者可以同时成立，但给出的处方相反：那一章的处方是把转嫁出去的负担记回账上，本章的处方是重建做这件事的人的生产线。判定形式：若在负担被记回账上之后，评价质量恢复，则那一章足够，本章多余；若负担记回了账而没有人能承担，则本章是必要的补充。

与关口一侧那一章的相接：那一章问的是这道关拦下过谁，本章给出的是关口失效的一个特定来源——不是标准太低，是站在关口后面的人不再能读。

站内还有几篇涉及相邻的判断，一并分清。专著《内卷与突围》第五章处理教育内卷，其病灶是价值递减之后仍被滞留；本章的病灶是验收能力的停产，两者的判定形式不同：前者要看价值曲线的方向，后者要看时间的归属。若在一个价值仍在上升的领域里观察到验收能力的停产，则本章不可还原为该章。学员论文《裂隙中的主权》主张主权型理解必须经历外部帮助完全无效的时刻，判据是求助通道是否真正关闭；本章只要求没有代做，不要求帮助无效——一段有人讲解而由本人推进的时间是哑段，不是那篇所说的那种时刻。判决性对照：把求助通道开着但禁止代做，若迁移成绩仍保住，则两者是两样东西。学员论文《断裂生成力》要求原有框架与外部参照同时失效，同上不成立于哑段。学员论文《在搏斗中锻造判断力》诊断的是目的的滑移，属意图层；本章属时间账层——一门目的完全没有滑移的课，照样可以把哑段做到零。学员论文《谁在掌尺》讲经济主体重制自身价值刻度的回路悬停，是本章在经济侧最近的站内邻居；分离线是它讲刻度的再生产，本章讲读刻度的人的再生产。

站内另有一篇《模具与回响》主张"守护成长"仍预设一份内在蓝图的给定，并命名了当精密模具落空时留下的那个空处。它与本章的分离线在时间的方向：那一章处理的是模具已经运行之后留下的空处，本章处理的是模具运行之前那段本应由本人走过的时间。判定形式：把模具完全撤走，若空处随之消失而本章的读数不变，则两者是两样东西。

■ 十三、出路：四步可干预次序

本章的题型要求交出一条可走的次序，而不是一个概念。四步按可行性排列，前两步不需要任何新技术，也不需要任何人放弃使用工具。

那一段谁走的

第一步：给哑段一个时间字段。

在课程与岗位的时间账里，为"无外部生成协助的独立推进"设立一段明

确时长，写进教学大纲与岗位说明，与课时、学分、工时并列。它不是禁令——禁令针对工具，字段针对时间。禁令必然失败，因为无法核查；字段可以成立，因为它只需要安排，不需要监控：把它排在同一间教室的同一段时间里，与考试的安排方式相同。

插手点在这里，不在工具设计上。第 2.1 节那项实验已经证明，工具侧的改良能把伤害拉回零而不产生增益。若把力气继续花在提示词与界面上，能得到的最好结果是回到从前；而回到从前并不解决入口正在坍塌的问题。

第二步：把验收从产出移到迁移差值。

评价的对象不再是那份产出的质量，而是撤走协助前后的表现之

差。这个差值不能被协助改善——恰恰相反，协助用得越多，差值越难看。它把评价的靶子从一个已经贬值到零成本的东西，移到一个仍然稀缺的东西上。技术上这需要每门课多一次闭卷环节，成本可控。

第三步：把入门岗位重新定义为哑段岗位。

明确写出：这个岗位的产出可以由协助完成，但由人完成才

算数，因为这个岗位买的不是产出，是五年后能审查资深产出的那个人。相应地，把这笔支出从费用改记为资本性投入，使它在成本压力下不再最先被削。这一步需要的是会计口径的改变，而不是道德呼吁。

第四步：让走过的哑段成为可携带的凭据。一份与产出无关、只记录"在何处、以何种条件、独立推进了多少时间、并通过了撤走协助后的检验"的记录。它与现行的学历、成绩、作品集都不共线，因此不会被现有的凭据通胀所稀释——至少在它自己被通胀之前。

四步的次序不可颠倒。没有第一步，第二步无从安排；没有第二步，第三步无从核算；没有第三步，第四步只是一张没有需求方的证书。

13.5 第一步怎么排：三种现成的排法第一步最容易被当成空话，因此给出三种已经在别处运行、可直接照搬的排法。

其一，考场排法。

把每周一次的独立推进排进固定教室的固定时段，与考试的安排方式完全相同，不联网、不评分、不收缴。不评分是关键——一旦评分，它就变回产出，第 5.2 节的性质二立刻生效。

其二，先做后看排法。

不改变任何课时，只改变次序：所有配有解答或范例的任务，一律先给任务、后给解答，中间设一段不可跳过的时长。这一条的实施成本接近于零，只需要教学平台增加一个延时释放的设置，而多数平台已经具备。

其三，交付前静默排法。

用于岗位而非课堂：任何交付物在提交前，须有一段由本人独立复核、不使用协助的时间，并在提交记录中标注该时长。它不核验内容，只核验时间归属，因此不产生监控问题。

三种排法都不禁止工具，都不增加课时，都不需要新的技术。它们唯一要求的是把一段时间从“用来出东西”改为“用来自己走一遍”，而这正是当前所有时间账里唯一缺失的那个字段。

13.6 谁先动：三类机构的不同起点四步的次序对所有机构相同，但从哪一步能真的动起来，取决于这个机构手里有什么。

那一段谁走的

手里有课时的机构（学校、培训体系）从第一步动。它们唯一需要付出的是排课表的意愿，不需要预算，也不需要说服任何外部方。阻力来自内部：一段不产出、不评分的时间在任何一张教务报表上都表现为浪费，因此这一步必须由能对报表负责的人亲自决定，交给基层教师会在第一次考核时被自动删除。

手里有考核权的机构（评审机构、资格考试、招聘方）从第二步动。它们不掌握别人的课时，但掌握靶子。把靶子从产出移到撤走协助前后的差值，是它们一纸决定就能做的事，而这一步的溢出效应最大——靶子一改，上游所有机构的课时安排会自动跟着改，不需要任何一份倡议。

手里有预算的机构（企业、医院、律所）从第三步动。它们既没有课时也不掌握外部靶子，但它们决定入门岗位的存废，而这是整条链上最先被削、也最难恢复的一环。这一步不需要新的开支，只需要改一个会计口径：把入门岗位的成本从当期费用改记为跨期投入，使它在削减名单上的位置从第一位挪到中间。

三类机构的动作互不依赖，也就不必等待共识。这一点重要，因为第十五节写明的反噬风险来自单独执行第一步；而第二类机构从第二步动，恰好从另一端把这个风险堵住——差值一旦成为靶子，打卡式的第一步在任何机构里都无利可图。

■ 十四、边界

本章的命题在下列情形不成立，写在这里以免被误用。

其一，条件一不成立的领域：产出无法与自身推进解绑者，例如体能训练、乐器的肌肉记忆、外科操作的手感。协助无法代做，回路不发生。

其二，条件二不成立的领域：存在独立于判断者能力的结果反馈者，例如结构工程、急诊、竞技比赛。桥塌与不塌不需要评审的判断力来裁定。

其三，条件三不成立的体系：验收者可从外部持续供给者。例如一个小国可以长期从外部引进资深评审，回路在其境内断开——但这只是把回路推给了供给方。

其四，尚未进入哑段的人：一个完全的初学者，在他能够独立推进之前，代做与讲解的区分没有意义。哑段的前提是这个人已经具备开始推进的最低条件。把本章的处方套用到最初几周，得到的是放任而不是训练。

其五，本章不主张减少工具的使用。第 5.4 节的乙格证明产出被前置本身不产生困境。本章主张的是给被前置出来的那段时间一个去处，而最有价值的去处之一恰恰是验收训练——那也需要工具在场。

■ 十五、反噬与伦理

必须先写出本章最可能造成的伤害，再写它的用法。

那一段谁走的

反噬：哑段占比一旦进入考核，最省事的达标方式是把它做成打卡——人坐在指定的教室里，屏幕上没有工具，脑子里什么也没发生。形式完美达标，机制照旧。这不是一个可以靠制度设计规避的风险，因为哑段按定义不发信号，任何核验都只能核验它的外壳。本章能给出的唯一缓解是把第一步与第二步捆绑：时间字段单独存在必然退化为打卡，只有当它与迁移差值

一起被考核时，打卡才无利可图——因为打卡不会改善差值。若有人只取第一步而不取第二步，本章的处方会比现状更坏，因为它增加了成本而不产生效果。

伦理三条，写死：其一，指位置不指人。哑段占比低是一个人所处的制度安排的性质，不是这个人的品质。任何把它读作“这个学生不勤奋”“这个年轻人依赖工具”的用法，都是把整体效应读成个体属性，而这种读法会自我实现——被这样评价的人会更依赖工具去改善当期指标。

其二，不得用于排除。哑段记录不得作为拒绝录取、拒绝雇用的依据。它的正当用途只有一种：让机构自己知道它有没有在生产未来的验收者。一旦用于筛人，第 5.2 节的性质二保证它会立刻被伪造，因为它不可核验。

其三，不得在能力不足者身上加码。对于因基础薄弱而无法独立推进的人，增加哑段等于增加空转，落进第 5.4 节的丙格。这类学习者需要的是把起点降下来，而不是把协助撤掉。

■ 十六、自反

本章自身的哑段占比很低。它的材料检索、结构安排与文字生成都有生成式协助在场，其中若干段落是先由协助交出初稿而后被改写的。按本章自己的判断，参与写作的人的验收能力也正在被同一个机制侵蚀。

这一点在原理上不可自我证实。理由正是本章自己给出的：伤害只在协助被撤走之后、且只相对于一个从未接受过协助的对照组才显形，而作者无法构成自己的对照组。因此本章对自身的诚实说明只有一句：它是否被一个仍然能够判断它的人读到过，本章自己无从知道。

另有一处需要认账：本章提出的两个读数当前不可精确测量，只能被制度性地保障。这既是命题的推论，也是它最大的实践弱点。一个不可测量的读数进入考核，几乎必然退化为它的外壳。本章选择仍然把它写出来，理由是：不给它一个名字，它连退化的机会都没有——它会以“效率低”的形态被继续删除，而删除它的每一次决定，在当期都是正确的。

■ 结 论

三门学科对人工智能时代的教育困境给出了三条互不相容的出路，三条路都成立，因为它们共享一个前提：一次教育经历的价值由它的产出被验收时的样子决定。这个前提在产出与自身推进绑定的年代是无害的；生成式协助把两者解绑，前提随之失效。

那一段谁走的

失效的后果不是学生变懒，也不是文凭贬值，而是一段特定时间的消失：当期不产出任何可被验收之物、而日后一切验收能力由它生成的那一段。它在所有账本上以"完成"的形态被记录，实际却没有被走过。由此形成的回路不需要参与者增多，不需要位置减少，在总投入下降时同样加速——这使它与经典的内卷概念在方向上分开。

出路因此不在工具的设计上，也不在凭据的更换上，而在时间的记账上：给那段时间一个字段，把评价从产出移到撤走协助前后的差值，把入门岗位的支出从费用改记为投入，并让走过的时间成为一份与产出无关的凭据。四步的次序不可颠倒，第一步单独执行会比现状更坏。

本章的适用边界写在第十四节，反噬与伦理写在第十五节，自身的处境写在第十六节。若哑段占比与撤走协助后的迁移成绩之间不存在剂量—反应关系，或若在保持任务难度一致条件下改变哑段位置不产生差异，则本章的命题为伪。

■ 注释

一、本章所称"生成式协助"，指任何能在不经过当事人自身推进的条件下交出与经过推进者相似产出的外部来源，不限于人工智能：答案册、代写、过于详尽的范例与同伴代做，机制相同。以人工智能为讨论对象，是因为它把这种协助的边际成本降到接近于零，从而使原本零星的现象变成系统性的。

二、"验收能力"在本章中不依赖同类协助而判断一份产出是否合格的能力。它与生产同一份产出的能力高度相关但不等同，且在本章的机制中比后者衰减得更快——生产可以被代做，判断不能。

三、第2.2节所引的就业数据处在活跃争论中，本章只取"收缩发生在入口而非存量"这一各方均不否认的事实，不取任何一方对总量效应或因果关系的判断。若这一事实本身被推翻，本章第4.3节材料失效，但第4.2节与第4.4节两件材料独立成立，命题不因此垮塌。

四、第五节的两个读数在现有条件下不可精确测量。本章给出它们，是为了让制度能够安排它们，而不是为了让考核能够计量它们。二者的区别见第十五节。

五、本章对哲学侧的分歧仅限一点：主体化不可被生产，本章同意；但哑段可以被安排，而安排不等于生产。这一区分是本章能够给出可干预次序的全部依据。

第十章

没有人在养它——一类没有供给方的能力

摘要

同一个问题——人工智能时代为什么投入不断增加而处境不改善——换一组理论来源再问一次。经济学取不完全契约与剩余控制权，艺术学取能动性推断与观看技能的历史性，工程学取「编程即理论建构」与软件演化。三家对这件事给出三条互不相容的解释：经济学说决定性的是当合约没写到的情况出现时谁有权决定，艺术学说决定性的是观者能不能从物上倒推出制作的能动性，工程学说决定性的是那套只存在于建造者头脑中、无法被外化的理论还在不在。三条都成立，却指定了三个互不相同的持有者，因而不能同真。本章指出三家共有一个从未被单独说出的前提——那样要紧的东西总归被某一方持有着；而推翻它的材料来自艺术学自己：观看的技能由日常生活中与艺术无关的实践顺带养成，因此它不由艺术界持有，也不由任何人持有。据此本章提出一类此前没有名分的东西——无供给品：被普遍依赖、却不由任何一方生产的能力；它不通过被消耗而减少，只通过养成它的那些低效活动被逐一优化掉而消失，且消失不产生任何事件。由此得到一条与三家都不同的判断：这一轮的困境不是竞争加剧，不是判据迁移，也不是瓶颈转移，而是评价所依赖的那些能力正在停止被养成，于是所有人只能用可以被生产的量，去补一个不可以被生产的缺口。本章给出二维辨别表、一句不含判断词的问话、一个三领域量纲不同而比率相同的读数（停养率），在十个领域兑现，其中两处反过来迫使命题大幅收窄——住院医师工时改革的两项随机对照试验并未显示能力下降，航空业则证明能力可以被制度性地回补——并给出四条可证伪条件与一条写死日期的赌注。第九节说明本章与本书同题五篇的分工：那五篇的缺口都在记录侧（账本上没有字段），本章的缺口在生产侧（生产链上没有这一环，因为它一直是副产品）。

关键词

无供给品；停养率；顺带养成；唯一养成源；不产生事件的消失

Abstract

没有人在养它

regulation.

■ 一、导论：三条不该同时为真的处方

同一个问题问第二遍，换一组人来答。

问题不变：人工智能把生成一份看上去合格的产出的成本压到接近于零之后，投入不断增加而处境不见改善，这件事的机制是什么，出路在哪里。而这一次，三门学科各自派出的不是上一次那些前沿工作，是另外三条线索。

第一条来自经济学的不完全契约理论。

格罗斯曼、哈特与摩尔那一支的核心主张是：任何合约都写不

全，因为未来的情况无法被穷举、也无法被第三方核实。既然写不全，真正决定投资与效率的就不是合约条款本身，而是当合约没有写到的情况出现时，谁有权决定——即剩余控制权。这条线索的处方因此清楚：要理解眼下的困境，就去看决定权落在谁手上。凡是能写进合约的部分，都可以被谈判、被定价、被转让；而价值真正的来源在写不进去的那一部分，它由谁掌控，谁就获得投资的激励。

第二条来自艺术学的能动性理论。

盖尔在《艺术与能动性》里主张，一件艺术品之所以起作用，不是因

为它美，而是因为观者从它身上倒推出了制作者的能动性——「这是怎么做出来的」这个推断，盖尔称之为技艺的迷魅。而巴克森德尔更早提出的「时代之眼」则给出了这一推断的条件：观看不是天赋，是一套被特定时代的日常生活训练出来的技能。这条线索的处方是：要理解眼下的困境，就去看观者还能不能从物上倒推出制作。当推断链断了，物就不再作为那种东西起作用，无论它本身有没有变。

第三条来自工程学的一篇老文章。

瑙尔一九八五年的《编程即理论建构》主张：一个程序真正的所在不

是代码，是建造者头脑中的那套理论——关于这个系统为什么这样组织、哪些是本质哪些是偶然、遇到新情况该往哪个方向改。代码、文档、

都只是那套理论的不完整投影。由此他得出一个在当时被认为极端、今天被反复引用的

：当持有理论的那批人散掉，程序就死了；后来者可以运行它、可以修补它，但不能真正修改它，只能重写。莱曼的软件演化定律从另一侧支撑同一件事：一个真实使用中的系统必须持续变化，否则用处递减；而每一次变化都会增加它的复杂度，除非有人专门做功去降低它——而做这份功，靠的正是那套理论。

没有人在养它

三条处方互相取消。

若经济学是对的，那么工程学所说的那套理论不过是人力资本的一种，可以用长期雇佣、股权、竞业限制这些契约手段锁定；问题于是回到权利的分配。若工程学是对的，不完全契约理论最核心的那个补救就落空了——把剩余控制权交给投资最要紧的一方，这个动作预设了要紧的东西可以被「交」；而一套只在人脑中、靠共同工作长出来的理论，不能被交给任何人。若艺术学是对的，前两家都指错了持有者：起作用的既不是资产的控制权，也不是建造者的理论，而是观者的推断能力——而推断能力既不属于资产，也不属于建造者。

反过来也一样，而且更刺眼。经济学每天在做的那件事——把关键资源变成可界定、可转让的权利——正是艺术学诊断为病根的那件事：观看能力是被历史顺带训练出来的，把它当作可分配的资产，会系统性地错配。艺术学每天在做的那件事——从物上倒推制作（鉴定、风格分析、技艺解读）——正是工程学诊断为病根的那件事：以为能从产物倒推出理论，正是维护灾难的起点。而工程学每天在做的那件事——把知识写进产物（代码即文档、测试即规格、架构决策记录）——正是经济学诊断为病根的那件事：凡是能写下来的都是可缔约的，而可缔约的部分恰恰不是价值所在；把力气全花在可写的部分，等于把剩余控制权白送出去。

三门学科各自的日课，是另一门学科诊断出的病。这个僵局无法靠综合三家之长化解，因为三家不是从三个角度看同一样东西，而是各自指定了一个不同的持有者。

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。撤销它之后，会出现一类三家的账本上都没有名分的东西；而这一轮困境的机制，本章认为就落在那里。

■ 二、三家的立场与它们各自的时序读数

本节把三家写实到可检验的程度。标准有三条：这一家必须主张只有自己指认的那一样是决定性的；必须给出时序读数——先动的是什么、另外两样多久跟上；必须自己交代一处它撑不住的地方。第三条在第四节要被用上。

2.1 经济学：决定一切的是剩余控制权主张。

已经写下的条款与实际出现的情况长期不容——因为情况无法被穷举、也无法被第三方核实——

逼得决定权的组织方式改变：所有权、纵向一体化、雇佣关系，都是对「谁在合约沉默处说了算」这个问题的不同安排。谁握有剩余控制权，谁就有投资的激励；其余一切跟随。

按这条读法，眼下这一轮的实情是：生成能力使得大量此前需要谈判的工作变得可以被规定、被外包、被自动化，即可缔约的范围扩大了。而可缔约范围一扩大，剩余控制权的相对价

值就发生位移——它从「会做这件事的人」移向「决定这件事该不该做、做成什么样」的位置。困境的体感由此而来：你把可缔约的那部分做得再好，也只是在一个已经被写进条款的范围里做工。

没有人在养它

时序读数。可缔约范围先动，权利安排随后跟上，个体的投资行为最后调整，滞后以年计。

它自己交代的一处。

这一支很清楚，「不可缔约」不是一个固定的范围，它随技术而变。哈特本人反复

强调，理论的适用性取决于哪些事项在当下无法被写清并被第三方核实；而这个边界一直在移动。也就是说，剩余控制权的价值不是一项属性，是一个随外部条件浮动的余数。在这一处，经济学自己承认它指定的那个决定项是被别的东西决定的。

2.2 艺术学：决定一切的是能动性能不能被推断出来主张。

物的实际外观与观者所持的推断技能长期不容，逼得物赖以成立的条件场改变。盖尔的论证是：

一件东西作为艺术品起作用，靠的是观者从它身上倒推出「有人以某种方式做了它」；技艺越是超出观者自己的能力，迷魅越强。而巴克森德尔补上了这个推断的来源——十五世纪意大利人看画时用的那套技能，来自他们日常生活里的别的事：估算酒桶容积的几何眼力、跳社交舞时对姿态的辨识、听布道时对手势含义的熟悉。观看能力不是艺术教育的产物，是别处实践的溢出。

按这条读法，眼下这一轮的实情是：生成模型使得「这是怎么做出来的」这个推断在大多数情形下无法完成——技艺的痕迹不再可靠地指示能动性。于是物的地位改变，价值必须另寻落点：过程、现场、来源、限量。

时序读数。推断技能的变化最慢，以代计；物的地位改变随之而来。这是三家里滞后最长的一条。

它自己交代的一处。

巴克森德尔自己写下的正是本章第四节要用的那件材料：养成观看技能的那些活

动，与艺术无关。这一句在艺术史里是常识，被反复引用，用来说明为什么不同时代的人看同一幅画看到的不是同一样东西。没有人拿它去回答「这项能力由谁持有」。

2.3 工程学：决定一切的是那套只在人脑里的理论主张。

建造这套系统所依据的理论与它实际运行的环境长期不容，逼得系统的实际外观改变——代码被

改动、被打补丁、被重写。而能不能有方向地改动，取决于那套理论还在不在人脑里。瑙尔的判断很硬：文档写得再全，也不能让一个没有参与建造的人获得那套理论；他能获得的只是一份说明，而说明不告诉他遇到新情况该往哪个方向改。莱曼的定律从另一侧说同一件事：系统必须持续变化，而变化累积复杂度，除非有人专门做功——做功靠的正是那套理论。

按这条读法，眼下这一轮的实情是：生成工具极大提高了产出代码的速度，而完全不提高理论的形成速度——理论只在人反复面对这个系统的真实问题时长出来。二者的速度差越拉越大，于是系统长成一种可运行、可修补、而不可有方向地修改的形态。

时序读数。

理论的存废先动，可维护性的崩塌在数月到数年内跟上；而这一条的滞后有一个恶劣性质——理论死掉的那一刻不产生任何征兆，征兆出现在下一次需要有方向地改动它的时候。

它自己交代的一处。

瑙尔自己写明：理论可以由新人获得，路径是与持有理论的人共同工作在真实问题

上。也就是说，理论既不在文本里，也不完全在某一个人的脑子里——它在共同工作这件事本身里。这一处是他为自己的悲观留的口子，而它同时是本章第四节的第二个材料接口。

没有人在养它

■ 三、三条驱动不能同真

把三家排在一起，冲突不在

，在它们各自指定的持有者：这一家说什么与什么不容

说什么被逼着改变

经济学已写下的条款×实际出现的情况决定权的组织方式

要紧的东西由谁持有时序资产的所有者

年

艺术学物的实际外观×观者的推断技能物赖以成立的条件场观看的一方

代

工程学建造的理论×运行的环境

月至年

系统的实际外观

建造的人

三个持有者互不相同，而三家各自主张自己指认的那一个是决定性的。逐对写死：经济学×工程学。若剩余控制权是决定性的，那么关键在于把它配置给投资最要紧的一方；这个动作预设要紧的东西可被转让。若那套理论是决定性的，转让在原理上无法完成——把公司买下来，买到的是代码与人的合同，买不到那套理论；而人一旦走了，理论随之消失，尽管所有资产都还在账上。两条不能并存：一个东西不能既是可配置的权利，又是不可转让的状态。

艺术学×工程学。盖尔说价值来自观者从物上倒推出的能动性；瑙尔说要紧的那套理论根本不在物上，从物上无论如何倒推不出来。若瑙尔对，盖尔所描述的迷魅就是一场系统性的误认——观者推断出来的从来不是真实的制作过程，只是他自己那套技能所能构造的一个版本。若盖尔对，理论在不在物上无关紧要，因为起作用的从来是推断而不是真相。

经济学×艺术学。不完全契约理论把一切化约为可界定、可分配的权利；而观看技能既不属于资产，也不属于任何个人，它是一代人共同生活的沉积。若艺术学对，经济学的框架读不出价值的来源，因为价值坐落在一个它的账本上没有的位置上。

三条不能同真，而谁都没说错。说错的是别的东西。

■ 四、三家共有的那个前提，与推翻它的那件材料

4.1 它们共同假定了什么三家争的是：当一件东西的价值不在它本身时，那个价值该由谁持有。

经济学答：由资产的所有者，通

过剩余控制权。艺术学答：由观看的一方，通过推断技能。工程学答：由建造的人，通过那套理论。

而三家共同假定了一件事：>那样东西总归被某一方持有着。

没有人在养它

它可以在所有者、观者、建造者三处之一，位置可以争，但它一定有一个持有者。念给三家听，三家都会说这还用说吗——会引起争论的是「在谁那儿」，不是「有没有一个持有者」。这正是共有前提的特征。

这条前提之所以坐得住，是因为它同时是三家各自最强的那件工具的前提：可转让性预设有一个当前持有者；迷魅预设推断技能有一个承载者；理论之死预设理论曾被某些人持有。撤销它，三家各自最锋利的那一条都要重写。

4.2 推翻它的那件材料，来自艺术学自己推翻共有前提的材料，不能从外面搬——从外面搬来的是第四家的立场，那只是加入争论，没有取消它。

它必须是三家之一自己已经掌握、已经发表、只是没有往这个方向用的事实。

这件材料在艺术学手里，而且是这一门的常识：巴克森德尔的时代之眼。

他的论证是：十五世纪的观画者用来看画的那套技能，来自与艺术无关的日常实践——商人估算货物体积时练出的几何眼力，跳社交舞时对身体姿态的分辨，听布道时对手势含义的熟悉。这些活动没有一项是为了培养观画能力而进行的；观画能力是它们的副产品。

这件材料在艺术史里回答的是另一个问题——「为什么不同时代的人看同一幅画看到的不是同一样东西」。没有人拿它当作关于持有者的证据。而它恰恰是：>那样东西不由任何一方持有。它是别处正在进行的活动的溢出；一旦那些活动停止，它随之消失，而没有任何一方会发现自己丢了东西——因为从来没有人持有过它。

这一条对三家同时生效。观看技能如此；那套理论也如此——瑙尔自己说它在共同工作里，而共同工作是为了把活干完，不是为了养理论，理论是干活的副产品；而不可缔约的那个余数同样如此——它不是谁的财产，它是缔约技术与技术条件之间的空隙，随两侧移动而涨落，没有主人。

三家争「在谁那儿」，争了一场没有对象的官司。

■ 五、无供给品

5.1 命名现在可以说出本章的判断了。

有一类能力，被普遍依赖，而没有任何一方生产它。它不是公共品——公共品有供给方，只是搭便车的人太多；它是别的活动的副产品，生产它的那个环节在任何人的计划里都不存在。本章把它命名为无供给品。

无供给品有三条定义性质，缺一条就不是它：

其一，没有供给环节。

没有人在养它

没有任何一方以生产它为目的而行动。养成它的那些活动各有各的目的——把货算清、把活干完、把新人使唤起来——养成能力是顺带发生的。

其二，它不通过被消耗而减少，只通过养成源被取消而消失。

使用它不损耗它。它的消失只有一个途

径：那些顺带养它的活动被逐一取消。

其三，它的消失不产生事件。

因为从来没有人持有它，就没有人报告丢失；因为它不在任何账上，就没

有一栏会变红。它的症状是弥散的、不可归因的：新人越来越难带，判断越来越拿不准，出了事说不清是哪一步错的。

由此本章的承重判断是：>这一轮的困境不是要素固定下的劳动内耗，不是价值判据向来源的迁移，也不是瓶颈从生成迁往验证，而是——一类从来不由任何人生产、只由别处的低效活动顺带养成的能力，因为那些活动被逐一优化掉而停止被养成；而它的消失不产生任何事件，于是所有人只能用可以被生产的量，去补一个不可以被生产的缺口。

三个被否定的项各对应一家的划界标准。而最后那半句是这条判断最要紧的部分：量是可以被生产的，质不是——不是因为质更高级，而是因为质从来没有被「生产」过，它一直是副产品。增加投入之所以不起作用，是因为你在用一种有供给的东西，去补一种无供给的东西的缺口。

5.2 二维辨别表一个名字不构成一个可用的区分。升成两根轴：养它的那些活动还在有一方专甲格·常态

养它的那些活动已被优化掉

既有专门培养，也有日常养成。乙格·可补

日常养成没了，但有供给方在，

门供给这传统的师徒制行业、有完整住院医培养体系的加预算、加课程、加模拟训练即可补上。飞行项能力

临床科室。

没有任何丙格·隐性健康

模拟器训练、结构化的临床技能中心。

能力在长，而没有人知道是

一方供给什么在养它。看起来完全不需要投入——这是它

最容易被误读为「不需要管」的一格。

丁格·靶格

能力在停止被养成，没有人负

责，没有事件，账面全绿。

要打的是丁格。乙格是关键对照：它证明「养成源被取消」本身不必然致命——只要存在一方以生产这项能力为目的，取消日常养成源是可以被补上的。丁格与乙格的唯一差别是有没有那个供给方，而这个差别在任何现行报表上都不出现。

丙格是丁格的前身，也是本章认为最值得单独看一眼的一格：一项能力越是长得好而无人操心，它越接近于无供给品，也就越脆弱。健康与危险在这一格里是同一件事的两面。

5.3 靶格的三条签名靶格必须是全部形式审查都能通过的那一格，否则不需要新框架就看得见。丁格的三条：

没有人在养它

其一，短期指标表现为改善。

被取消的都是低效活动——手工核对、写初稿、跑腿、值夜班抄病历、看片子前先自己读一遍。取消每一项都是可见的效率提升，每一项都有一份说得通的成本收益。

其二，失效不产生任何信号。

没有一栏会变红。第一年什么也看不出来，因为在职的人能力仍在；症状要等一整个培养周期之后才在新一代身上出现，而那时它会被读成代际问题、态度问题、教育问题。

其三，它自我加固。

越是效率高的组织，越快把顺带养人的活动优化掉——因为在每一项单独的账上，那都是纯浪费。而组织越是因此显得高效，越有理由继续这么做。

5.4 一句不含判断词的问话判断需要一句能拿去问流程文件、岗位说明与培训大纲的问话，而不是拿去问人的感受：>在这个组织里，一个人是在哪一项具体工作中学会做独立判断的？那项工作现在还有人做吗？

这句话里没有「应当」、没有「真正」、没有「充分」。它问的是两件可以从流程变更记录里查出来的事。

在五个具体场景里跑一遍：1.软件团队。此前新人靠修缺陷与读别人的合并请求学会整个系统。现在缺陷由工具定位、变更说明由工具生成、评审由工具先过一遍。答：大部分已取消。2.律师事务所。初级律师此前靠通读卷宗与文书核对建立对案件的整体判断。答：已外包并自动化。3.医院影像科。住院医此前必须先自己读一遍片子再看上级意见。答：部分保留——因为读片的顺序被写进了教学规程，这是少数被明文保护下来的养成源之一。4.传统绘画的鉴定。眼力此前靠大量上手看真迹与赝品。答：还在——上手看仍是唯一途径，没有任何东西替代得了它。这一个与预期相反，是查出来的。5.民航飞行。飞行员此前靠手动飞行建立对飞机状态的直觉。自动飞行普及后手飞时数大幅下降，而这一项已经被制度性地要求恢复：美国联邦航空局二〇一三年一月发布安全警示，明确指出连续使用自动飞行系统不会强化飞行员

手动飞行的知识与技能，鼓励运营人在航线运行与训练中安排手动飞行；二〇一七年又发布一份专门讲手动飞行熟练度的续件，欧洲航空安全局同期也发过对应的安全通告。答：曾被取消，且已被制度性回补。

五个答案互不相同。第四个与第五个是查出来的，不是从命题推出来的；而第五个是本章最重要的一个经验支点，第七节与第十一节都要用它。

5.5 丙格：能力长得好而无人操心，为什么这一格最值得单独看二维表里最容易被跳过的是丙格——能力在长，而没有任何一方在供给它。它在报表上表现为完美：不需要预算，不需要岗位，不需要计划，人就是一批批长出来了。任何一个管理者看到这一格，得到的

都是「这一块不用管」。

而这正是它的危险所在。丙格与丁格之间只隔一个动作：把养成它的那些活动优化掉。这个动作在丙格里没有任何阻力，因为没有人知道那些活动在起作用，也没有人的职责是保护它们；而在甲格里，同样的动

没有人在养它

作会遭到那个供给方的反对——培训部门会为自己的课程辩护，带教医生会为轮转制度辩护。有人供给，就有人反对；没有人供给，就没有人反对。

这解释了一个反直觉的现象：那些历来不需要操心的能力，最先消失。而那些一直需要专门投入、被反复讨论、被写进预算的能力，反而保住了——不是因为它们更重要，是因为它们有守护者。

由此得到一条实用的判据，它比任何清单都可靠：>在这个组织里，哪些能力从来没有出现在任何一份培训计划、任何一次预算讨论、任何一个岗位说明书里，而所有人都默认它会有？

答得出来的每一项，都是一个候选的无供给品。而这个问题的答案通常很短，因为它要求人说出自己从未想过的事——这本身就是它属于这一格的证据。

需要说明的是，丙格并不必然走向丁格。一项能力可以长期停在丙格：第七节第四处的鉴定眼力就在这一格里待了几百年，因为上手看真迹这件事没有替代品，也就无从被优化。丙格转丁格的触发条件只有一个：养成它的那些活动出现了一个更便宜的替代。

人工智能这一轮之所以要紧，不是因为它取消了什

么，是因为它在极短时间里给几乎所有低效活动都提供了替代。

■ 六、停养率：一个三领域量纲不同而比率相同的读数

把 5.4 那句问话固定成一个可以清点的量： $>$ 停养率=过去若干年中被取消或自动化的工作项里，「曾是新人接触完整问题的唯一途径」的那一类所占的比例。

三个领域各自兑现，量纲不同而比率的意义相同：
· 软件工程：近三年被自动化或外包的开发任务清单中，属于新人培养路径的项数占比。数据在流程变更记录与新人培养手册的修订史里。

· 临床医学：近十年被设备或流程替代的临床动作中，属于住院医师培养必经环节的占比。数据在培训大纲与科室规程的修订记录里。

· 艺术与鉴定：被数字化手段替代的看画方式中，属于眼力养成必经环节的占比。数据在院校课程大纲与机构培训安排里。

四条性质逐条过：无量纲。占比除以占比，三处可并排比较。

可事后清点。

三处的依据都是既有文件的修订史——流程变更记录、培训大纲、岗位说明书。不需要新建任何采集。

把一个印象变成一个数。

没有人在养它

「现在的新人不如从前」是一句体感，且几乎总是被读成态度问题；停养率是一个可以被争论、可以被反驳的数。

与既有主要指标正交。

这一条最容易漏也最要命——一个与既有指标高度相关的量会被立刻读成那个指

标的代理。检验办法是举出一个既有指标最好看而这个数最难看的实例： $>$ 一条完全自动化、缺陷率最低、交付最快、按任何现行工程指标都堪称卓越的流水线，停养率可以等于一——它把每一项需要人反复上手的环节都消掉了，而这正是它卓越的原因。

反过来的实例同样成立：一个什么都没自动化、效率低下、整体产出很差的作坊，停养率是零。停养率与当前绩效无关，与养成结构有关；而且它与效率指标的关系不是无关，是负向的——这使它不可能被读成效率的代理，也使它天然不受欢迎。

6.1 读这个数时会遇到的两个困难，以及怎么办这个读数不像别的比率那样可以直接从系统里拉出来，它有两处需要人来判断，必须说清楚。

第一处：什么算「唯一养成源」。

这是本章全部判断里最可争议的一点，也是第十三节伦理第三条要求

公开编码规则的原因。可操作的做法是反过来问：取消它之后，新人还能在别的什么地方练到同一件事？

答得出具体的另一项活动，就不是唯一；答的是「多接触接触就有了」「以后自然会」，就按唯一计。这个问法把举证责任放在「不是唯一」的一方，因为默认「以后自然会」正是丁格得以维持的那句话。

第二处：滞后。

分子里的取消发生在今天，而它的后果要一整个培养周期之后才在新一代身上出现。这

意味着当期的停养率不能与当期的任何绩效指标做相关——那样做只会得到零。正确的用法是把它当作一条先行指标：今年的停养率对应的是若干年后的判断力水平，而这个滞后长度可以由该行业的培养年限直接给出（软件二到三年，临床五到八年，鉴定十年以上）。

这两处困难合起来产生了一个不好的后果，必须写在明处：这个读数在它最有用的时候最没有说服力。

它在丙格里发出警告，而丙格里一切正常；等到丁格的后果显形，警告已经晚了一个培养周期。任何依赖当期证据的治理程序，在原理上都用不了它。这也是第十一节把「建立倒查制度」而不是「监测这个数」

排在第一步的原因——倒查是向后看的，它不需要等。

■ 七、十处兑现

判断要成立，必须在具体领域里预测出具体的事。以下十处，其中第九、第十两处反过来给命题加了约束，而且削得很重。

一·软件维护。预测：在采用生成工具程度高的团队里，交付速度上升与新人达到可独立负责一个模块所需时间的上升会同时发生，且后者滞后一到两个培养周期。可读数据：入职到首次独立负责的月数，在

没有人在养它

人力系统里已有。反向对照：在保留了「新人必须先自己定位一遍缺陷再看工具

」这条规程的团队里，这一上升应当显著更弱——顺序是关键，先看

就不再有练习。

二·代码评审。预测：由工具先行生成评审意见的项目，人工评审意见的深度分布会向表层集中——因为深层意见来自对系统的整体理解，而理解正来自逐行读别人代码这件被替代掉的

事。可读数据：评审意见的分类编码。这一处可与第一处交叉验证：若两项同时恶化而交付速度上升，则三者共同指向同一个养成源被取消，而不是三个独立问题。

三·临床带教。预测：在语音转写普及的科室，住院医撰写病历的时间下降，而形成病人整体印象所需的时间上升。机制是：手写病历这件低效的事，顺带完成的是把散落的信息组织成一个整体判断。可读数据：交班时对病人情况的复述完整度评分，多数教学医院已有记录。反向对照：要求住院医在看转写稿之前先口述一遍印象的科室，这一上升应当不出现。

四·法律实务。预测：文书核对外包程度越高的所，初级律师达到可独立出具意见所需年数越长，且这一关系在控制案件量之后仍成立。可读数据：从入所到首次独立署名出具意见的年数。这一处最容易被误读成「案子少了练得少」，故控制案件量是必须的。

五·学术训练。预测：在允许生成工具撰写文献综述的项目里，学生完成综述的时间下降，而提出可研究问题的能力的养成被推迟——因为问题意识长在通读大量不相干文献的那段低效时间里。可读数据：开题到确定研究问题的月数、中途更换选题的次数。反向对照：仍要求手写文献笔记的课题组不应出现推迟。

六·编辑出版。预测：校对自动化程度高的机构，编辑判断稿件价值的能力衰退最快——因为逐字读稿这件被替代掉的事，顺带养成的是对文字质地的判断。可读数据：编辑初审判断与终审结果的一致率，多数出版机构的稿件系统里已有。这一处的读数很干净，因为一致率与工作量、稿件质量都不直接相关。

七·制造业。预测：产线全自动化的工厂，工艺问题的诊断能力集中在少数老工人身上且不再产生新的持有者；当这批人退休，诊断能力不是下降而是断档。这一处的特征是断崖式而非渐进式，因为它取决于个别人的在职与否。可读数据：工艺异常从发现到定位的平均时长，按年份画出来即可看到断崖。这一处是本章预测里最容易被证伪的一个：若断档没有出现，说明诊断能力有本章未识别的第二个养成源。

八·教育。预测：作业与批改自动化程度越高的学校，教师对学生困难所在的判断越依赖系统给出的标签——而这项判断此前来自逐份看作业时积累的模式识别。可读数据：教师在系统标签之外自主提出的诊断占比。这一处与第六处同型，可合并检验。

九·住院医师工时改革（反过来加约束，削得最重）。这一处不支持命题，而且是本章能找到的最强的不利证据。

美国研究生医学教育认证委员会二〇〇三年起限制住院医师每周工时不超过八十小时、单次值班不超过三十小时，二〇一一年进一步把第一年住院医的连续工作限制在十六小时。这是一次大规模的、真实的「取消低效暴露」。而随后的两项全国性随机对照试验——外科的

试验（一百余个普通外科培养项

目）与内科的 iCOMPARE 试验（六十三个内科培养项目）——都没有发现放宽或收紧工时限制导致病人

没有人在养它

结局的显著差异；早期对二〇〇三年与二〇一一年两次改革的评估同样显示病人结局基本无差异，考试成绩亦未见系统性下降。

这削掉了本章原本打算写下的那句更强的话：「取消低效活动必然导致能力停止被养成」。真实的命题必须缩到：只有当被取消的那项活动是该能力的唯一养成源时，停养才发生。

医学有一整套并行的养成结

构——模拟训练、结构化课程、分级带教、专科考试——因此减少值班时数不等于取消唯一养成源。命题的适用范围因此从「所有被优化掉的活动」缩到「唯一养成源」，而这是一次范围上的实质收窄，不是补一个条件。

必须同时诚实地记下这项证据的边界：这两项试验测的是病人结局与考试成绩，不是本章所关心的那种「独立判断的形成速度」；而 iCOMPARE 确实报告了在弹性工时组里住院医对教育质量的满意度更低。

也就是说，这份不利证据推翻的是本章的强版本，没有推翻弱版本；但本章不能因此把它读成对自己有利——测量的是不是同一东西，是本章欠着的一笔账，第十二节把它写成了可执行的检验。

十·民航的手动飞行（反过来加约束，改的是处方）。这一处证明能力可以被制度性地回补，因此本章不能主张这个过程不可逆。

美国联邦航空局二〇一三年的那份安全警示，把因果链写得很清楚：飞行运行数据分析发现手动操纵失误上升，而连续使用自动飞行系统不会强化飞行员手动飞行的知识与技能；因此鼓励运营人把手动飞行安排进航线运行与训练。二〇一七年又发了一份专讲手动飞行熟练度的续件。这是本章所知唯一一个把「某项能力的衰退」倒查到「某项活动被取消」并写进正式文件的行业。

它削掉的是本章的处方。本章原本会主张「保留那些低效活动」；而航空的经验说明，真正靠前的一步不是保留，是先把因果链倒查出来并写下来。航空之所以做得到，是因为它有一套把每一次事故倒查到底的调查制度，以及可以逐架次分析的飞行运行数据。没有这套等价物的行业——绝大多数行业——不具备这个前提。处方因此从「保留低效活动」改成「先建立能把能力衰退倒查到某项活动的调查制度」，这是一次价值排序上的改动。

而这一处还留下一个不利于本章的细节，同样要记下来：交通部监察长办公室在二〇一六年的审计中指出，联邦航空局并未确认航空公司在这份安全警示之后是否真的增加了手动飞行的机会。倒查出因果链，不等于回补真的发生了。

■ 八、与最近的几种说法的分界

这一判断最容易被读成几个已有的概念。逐条交手，每条给出一个能把两边分开的观察。

其一·自动化的讽刺（班布里奇，一九八三，人因工程）。这是最近的邻居，而且近得危险。她指出自动化把人留在最难处理的那部分，同时剥夺了他保持技能的机会。分离线：班布里奇讲的是同一岗位上的技能保持，因此她的处方是岗位内的定期手动操作与模拟训练；本章讲的是能力的养成源在别处——不是「这个岗位的技能没在练」，而是「养成这项判断的那件事在另一个工序、另一个部门，且它被取消时没

没有人在养它

有人把它和这项判断联系起来」。判定形式：若在岗位内补上练习即可恢复，则班布里奇足够（民航的手飞正属于这一类）；若补岗位内练习无效、必须恢复另一个部门的某项工序才有效，则本章对。语音转写与病历那一处正是分开两者的现成样本——补写病历的模拟训练无效，因为要紧的不是写字，是写字时被迫把信息组织成整体这件事。

其二·默会知识（波兰尼；柯林斯的激光复制研究）。分离线：默会知识讲的是不能被言传，其传递路径是师徒之间的亲身接触，它预设有一个持有者和一次传递。本章讲的是连传递这一环都不存在——它不是从师傅传给徒弟的，是徒弟在做另一件事时自己长出来的。判定形式：设置密切的学徒接触，默会知识预测可转移；本章预测若那件「另一件事」已被取消，学徒制也补不上。

其三·正当的边缘性参与（莱夫与温格，一九九一，学习理论）。极近：那些低效活动正是他们所说的边缘任务。分离线：他们描述的是共同体有意安排的进入路径，理论预设有一个共同体在维持这条路径。

本章说的是没有人安排它，也没有人知道它在起作用，因此它被优化掉时没有任何守护者。判定形式：他们预测共同体会保护进入路径；本章预测恰恰是最专业化、最讲效率的共同体最快把它取消，因为它在自己的账上那是纯粹的低效。

其四·人力资本折旧（贝克尔一系）。分离线：折旧讲的是已经获得的能力随时间衰减；本章讲的是从未被获得。判定形式：同时测在职者与新入行者，折旧预测两者同降；本章预测在职者不降而新入行者从未达到过在职者当年的水平。这一条只需两组人各测一次。

其五·公共品与公地治理（萨缪尔森；哈丁与奥斯特罗姆）。最容易被套用的一条。分离线：公共品的问题是谁付钱供给，公地的问题是谁在过度取用；两者都预设有一个可被指认的资源池与一个供给或取用的动作。无供给品根本没有供给环节，也不因取用而减少，因此奥斯特罗姆那套社群规则在这里无从下手——没有一个可被治理的池子。判定形式：若建立一个共同出资的供给机制即可解决，则它是公共品；本章预测出资也没用，因为没有人知道该出钱做什么——要做的不是生产它，是不取消那些顺带养它的事。

其六·正外部性与内部化（庇古与科斯）。分离线：外部性理论要求能指认出产生它的活动与受益者，然后把它内部化；本章的困难恰恰在于没有人知道是哪项活动在产生它——溢出的来源与受益之间隔着数年与数个部门，科斯式的谈判在原理上无法启动，因为谈判双方都不知道自己是谁。判定形式：若在多数案例里能事先指认「取消这一项会损失哪项能力」，则可以内部化，本章的悲观被推翻；民航是唯一一个做到了的案例，而它靠的是事故调查制度，不是谈判。

其七·局部最优的加总（合成谬误）。把本章的命题剥掉全部学科名词，剩下的形式是「每一项单独都对，而合起来取消了某样东西」。这个形式在逻辑学与经济学里早有名字。分离线：合成谬误是一个错误的名字，不是一个机制——它说加总可能出错，不说为什么恰好是这样东西出错。本章给出的是机制：因为它是副产品，所以它在任何单项的成本收益里都不出现。任何逐项审查的程序，无论多严谨，在原理上都看不见它。判定形式：若把审查从逐项改成整体评估即可发现损失，则它只是合成谬误的一个实例；本章预测整体评估同样看不见，因为整体评估的项目清单仍然是那些项。

其八·成功陷阱与能力刚性（列文塔尔与马奇，一九九三；伦纳德·巴顿，一九九二）。反转模板专查：本章的形式看上去像「越高效越糟」。分离线：成功陷阱的机制是过度利用、疏于探索，主体因为过去的

没有人在养它

成功而改变了行为；本章里没有任何主体因为成功而改变行为——取消低效活动的决定，在一个从未成功过的新组织里同样会作出，而且更快。判定形式：若这一现象只出现在有成功历史的组织，则成功陷阱对；若新成立、无成功包袱的组织同样出现且更严重，则本章对。

其九·工时限制研究（医学教育，方法学占位者）。这是本章必须点名的方法学邻居：它已经是一套成熟的经验范式，用随机对照去测「减少暴露是否损害能力」。分离线：它测的因变量是病人结局与考试成绩；本章关心的因变量是独立判断的形成速度。二者不是同一东西，而现有范式没有测过后者。这一条不是本章的优势，是本章欠的账——第十二节的主检验正是为填这一笔而设计的。

其十·时代之眼本身（巴克森德尔，一九七二）。本章的推翻材料来自它，因此必须说清本章与它的差别。分离线：巴克森德尔用它解释差异——不同时代的人看到的不是同一样东西；本章用它解释消失——当那些日常实践本身被取消，能力不是变成另一种，而是不再形成。他描述的是替换，本章描述的是断流。判定形式：若某项能力的养成源被取消之后，出现了另一套日常实践顺带养成一种替代能力，则是替换；若什么也没有长出来，则是断流。这一条同时是本章最应当被检验的地方——本章有可能只是没看见新长出来的那一样。

最近的邻居是哪一个。

是班布里奇。本章仍不能被她吸收，理由只有一条：她的框架里，那项能力有一

个明确的持有者与一个明确的岗位，因而有一个明确的补救位置；而本章处理的正是没有持有者、也没有对应岗位的那一类。

再往下一层。

逐个问这些邻居「它结构性地看不见什么」：班布里奇把「能力属于某个岗位」当作给

定，因此看不见养成源在别的岗位；默会知识把「有一个持有者在传」当作给定，因此看不见无人传而自己长出来的那一类；莱夫与温格把「共同体在维持进入路径」当作给定，因此看不见无人维持的路径；折旧理论把「能力先被获得」当作给定；公共品理论把「有一个供给环节」当作给定；外部性理论把「来源可被指认」当作给定。

这几行背后是不是同一样被当作给定的东西？是。所有人都假定，凡是要紧的东西，都有一个位置在生产它、持有它或维持它。

这块地基是所有派别共同站着的，因此谁也没法回头看它。而本章所命名的那一

类东西，正是从这条地基的缝里掉下去的。

■ 九、与本书同题诸篇的分工：从记录侧到生产侧

本书在同一个问题上已有五篇，三个学科相同而理论来源不同。逐条指名，并说明这一章为什么不是第六个同类。

那五篇各自命名的东西是：为使产出可被归属而一并生产的那一层；分辨力已饱和而无人作废的那道关；没有人打开过就作出的那次否决；落判件在无人记账处合成的那条基线；核实这份工作从哪一侧移到了哪一侧。五样东西的共同形状是——它们都是账本上没有字段的东西，缺口在记录侧。

没有人在养它

本章的缺口不在记录侧，在生产侧。无供给品不是被记漏了，是根本没有人在生产它；给它设一个字段，得到的将是一条持续下降的曲线，而没有任何一方能对这条曲线负责，因为没有任何一方的职责里包含生产它。

这给出一条可以同时裁决六篇的观察：>在一个已经给某处缺口设了字段的体系里（例如已开始统计那五个读数之一的机构），若该缺口的读数在两年内改善，则记录侧的诊断成立；若读数被稳定地记录下来、而没有任何一方能据以采取行动，则本章的诊断成立。

逐篇再各给一条更细的分界。与那篇讲归属成本的：它预测把归属的成本记出来就能治理；本章预测记出来也治理不了停养，因为归属成本有承担者而养成没有。与那篇讲门槛权属的：它主张给门槛指定一个作者；本章指出养成源根本没有可指定的作者位置——不是无人认领，是这个位置在组织结构里不存在。与那篇讲未阅否决的：它的读数可以在一个季度内改善（加一个字段即可）；本章的读数在原理上不可能在一个培养周期内改善。与那篇讲暗基线的：两者都有长滞后，但暗基线的滞后来自累积，本章的滞后来自培养周期，因此本章的滞后长度可以由该行业的培养年限预测，而暗基线不能。与那篇讲核实负担转嫁的：它说同一份工作换了人做；本章说那份工作还是原来的人做，只是做它的能力不再形成——两者可以叠加，而叠加时衰退速度不是相加。

最后一句留给这一次的实验本身。同一个问题、同样三个学科、换一组理论来源，得到的东西落在了另一侧——这说明前五篇的那个共同形状（缺口在记录侧）不是这个问题的性质，是那批理论来源的性质。换来源会换答案，这一条本身可以被继续检验：再换一组来源问第七遍，若仍落在记录侧或生产侧之一，则可选的答案空间比看上去小；若落在第三处，则说明这个问题的答案取决于你请谁来答，而这对任何跨学科方法都是一个不小的坏消息。

9.1 这次实验本身该怎么被检验上一段说「换来源会换答案」，这句话如果只是被读作一句感想，就没有意义。把它写成可以被推翻的形式：>若再换一组理论来源问同一个问题，得到的判断仍落在记录侧或生产侧这两处之一，则可选的答案空间确实很小，前五篇的共同形状反映的是问题的性质；若落在此前没有出现过的第三处，则答案主要由来源决定，而不是由问题决定。

这一条可以直接执行，成本只是再做一次。而它的两种结果都有后果，因此不是白问：若答案空间小，那么这类做法是可靠的——不同的人从不同的入口进去，会收敛到同一批位置上，收敛本身就是证据。若答案主要由来源决定，那么任何一篇这样的文章，其

的份量都要打一个折扣，因为它同时是一份关于「作者读了什么」的报告；而这一点必须写在每一篇的开头，不能只在事后补充。

没有人在养它

本章倾向于第二种可能，理由是这一次的落点变化太整齐了：三家全部换掉，落点整体从记录侧移到生产侧，而问题一个字没改。但一次不足以

——这正是本章最希望被继续做下去的一条，而它不需要任何新方法，只需要有人再问第七遍。

■ 十、三家为什么各自到不了这一步

三家停在原地，不是因为疏忽。是它们各自的核心问题恰好使这一步成为不必要的。

经济学到不了，因为它的核心问题是激励与配置。在这个问题结构里，任何要紧的东西都必须被处理成一项可被配置的对象——权利、资产、人力资本。一样没有持有者、也无法被配置的东西，在这个框架里不是一个难题，而是一个非对象：它不进入任何一阶条件。把它当作可配置之物，是这个框架能做的唯一正确处理，而正是这个处理抹掉了「无供给」这个维度。

艺术学到不了，尽管推翻共有前提的材料就在它自己手里。因为它的核心问题是理解与判断的历史条件——它要解释的是差异，是为什么不同时代看到的不一样。在这个问题结构里，观看技能的来源是一个解释项，用来说明差异从何而来；而「这项技能会不会因为来源被取消而停止形成」是另一个问题，它属于当下而不属于历史，因而不在这门学问的问题清单上。

工程学到不了，因为它的核心问题是系统在约束下可靠运行。瑙尔已经走到了「理论不可外化」这一步，也已经写下「理论在共同工作里长出来」，但他的落点始终是这个系统的可维护性——他关心的是理论死了之后程序怎么办，不是这套理论当初是靠哪些具体活动长出来的、那些活动会不会被取消。把注意力放在系统上而不是放在养成链上，是这门学问的问题结构决定的。

三处盲区的形状相同：各自的核心问题都要求把这样东西转成自己框架里的某一类——可配置的对象、历史的解释项、系统的属性——而这三处转换都恰好抹掉了「谁在养它」这个维度。

■ 十一、出路：倒查因果链，比保留活动更靠前

前十节回答了机制。出路不是「保留低效活动」——那是本章原本的处方，已在第七节第十处被削掉了。

真正靠前的一步，是航空业做过而别的行业没做过的那一步。

第一步，建立能把能力衰退倒查到某项活动的调查制度。

这是全部三步里唯一不可跳过的一步。航空之

所以能在二〇一三年写下「连续使用自动飞行系统不会强化手动飞行技能」，靠的不是有人有先见之明，是它有两样东西：一套把每次事故倒查到底的调查制度，和可以逐架次分析的运行数据。绝大多数行业两样都没有——出了问题查的是这一次的责任，不查这项能力当初是靠什么长出来的。可以插手的最小动作是：在既有的复盘流程里加一问——「这次判断失误，作出判断的人此前是在哪里练过这类判断的？那件事还在做吗？」这一问不需要任何人同意，不增加任何岗位。

第二步，把丙格挑出来。

没有人在养它

二维表里的丙格——能力在长而无人操心——是唯一还来得及的一格。挑出它的办法是问 5.4 那句话：哪些能力现在长得好，而说不出是什么在养它？答不上来的每一项，都是一个候选的无供给品。这一步的产出不是保护清单，是一份需要先查清来源的清单。

第三步，只对唯一养成源动手，不做普遍的保留。

第七节第九处的证据在这里生效：医学并没有因为减

少值班时数而崩塌，因为它有并行的养成结构。所以正确的动作不是拒绝自动化，而是在取消任何一项之前，先确认它不是某项能力的唯一养成源；若是，则要么保留它，要么先建起一个替代的养成结构——即把这项能力从无供给变成有供给，把丁格挪回乙格。

三步的次序是死的。跳过第一步而直接保留活动，保留下来的多半是错的那些——因为没有人知道是哪些在起作用，而最容易被保留的总是最显眼、最像培训的那些，恰恰不是真正在养人的那些低效小事。

还有一个问题不能回避：第一步该由谁来做。

按本章自己的判断，这件事没有天然的负责人——如果

有，它就不在丁格里了。可以指望的只有两类位置。一类是已经出过事、且必须对外交代的机构：航空是这么开始的，医疗事故调查、金融风控复盘、重大工程质量追溯都具备同样的结构。另一类是带教者本人——他是唯一能同时看见「新人现在会什么」和「自己当年是在哪儿学会的」这两件事的人，而这两件事的对照正是倒查所需要的全部材料。这也意味着，若一个组织里带教的人已经不再是当年那样长起来的，第一步就已经做不了了——倒查有它自己的时限。

必须说清这条出路的代价。它要求组织承认一件不好承认的事：它此前的每一次效率改进，都可能同时取消了某样它不知道自己拥有的东西。这在任何一份复盘里都是一句难写的话。

■ 十二、可证伪条件

最强的竞争解释先写出来：其实不就是新人经验少吗，不就是代际差异吗？这是最朴素、也最难反驳的一种说法——它主张能力差异只由绝对经验量与代际特征解释，不需要「养成源」这个中间项。

条件一（主条件，机制证伪）。

>控制入职年限与累计工作总量之后，若「该岗位的某项能力是否已失去其唯一养成源」与「达到独立判断所需月数」之间不存在剂量-反应关系，则本机制为伪——届时能力差异应归因于绝对经验量，而不是养成源的存废。

样本设计（这一条可低成本实做，所需数据在多数机构的既有记录里已经存在）：取同一家医院内部的六个以上科室，同一制度环境、同一培训体系、同一批带教标准。自变量是该科室的某项传统必经工序是否已被替代（影像科的先自读、检验科的手工镜检、外科的开放术式比例、内科的手写病历）。因变量是住院医达到可独立处置所需月数，取自培训档案。不使用跨国或跨机构对比——那种对比在制度、生源、病种上处处不同，所谓控制某变量在方法上站不住。

条件二（补第七节第九处欠的那笔账）。

没有人在养它

>在工时改革的既有样本里，若改用「达到独立判断所需月数」而不是病人结局与考试成绩作为因变量，仍未发现差异，则本章关于医学的那一部分为伪。

这一条之所以要单列，是因为第七节第九处那份不利证据测的不是同一样东西，而本章不能靠这一点把它推开——必须给出一个具体的、能测同一样东西的设计。

条件三（第三层，不依赖前两层）。

>在存在并行养成结构的领域（有系统模拟训练、有分级带教、有强制轮转的），取消某项日常工序对独立判断形成时间的影响，应显著小于不存在此类结构的领域。若不成立，则第五节二维表里甲乙两格与丙丁两格的区分不成立，本章的整个框架失去分格依据。

条件四（对本章最要害的一条，来自第八节第十条）。

>若在某项能力的养成源被取消之后，可以观察到另一套日常实践顺带养成了一种替代能力，则本章所说的「断流」实为「替换」，本章的悲观判断为伪。

这一条是本章最可能栽的地方，也是本章自己最想知道答案的一条：新的工作方式会不会顺带养出新的判断力，只是我们还不认得它。

这条判断若被证伪，我们会学到什么。

若条件一被推翻，说明能力的形成对具体活动并不敏感，那么一

切关于「保留哪些工序」的讨论都是浪费，正确的方向是增加总的暴露量与时间——这是一个成本高得多、但方向明确得多的处方。

■ 十三、边界、反噬与伦理

不适用的范围。

本章不适用于存在专门供给方的能力（乙格）；不适用于可以被完整形式化的技能——

凡是能被写成规程并靠规程习得的，取消日常练习不构成停养；不适用于养成源仍在的领域（第七节第四处的鉴定即是）。本章也不主张所有低效活动都在养人：绝大多数低效就是低效，识别哪些在养人正是第十一节第一步要做的事，而本章不提供这份清单，只提供找它的问题。

反噬预言。

本命题一旦被制度采用，最省事的实现方式恰恰会摧毁它要保护的东西：机构会开始测量停

养率并把它写进考核，而把这个数做好看最便宜的办法，是为已经自动化的活动保留一个仪式性的人工版本——让新人手工做一遍已经不重要的事。于是能力照样不被养成，而数字很好看；更坏的是，这类仪式会消耗掉本可以用于真正养成的时间，并且使人相信问题已经被处理了。我看不到出口。

唯一能做的是

把这个风险写在明处，并配上下面三条。

伦理三条。

其一，停养率指的是位置，不是人——它刻画的是一个岗位的养成结构，不得用于个体的能

力评定，尤其不得用于新入职者。其二，不得为了把这个数做好看而人为制造练习风险：让新人在无监督条件下处理真实的高风险情形，是把养成的成本转嫁给第三方；民航的手飞恢复之所以正当，正因为它在受控条件与双人机组下进行。其三，已经据此读数作出不利安排的机构，必须公开编码规则（什么算唯一

没有人在养它

养成源、什么算已被取消）并给出复核通道——因为「唯一养成源」这个判断本身是可争议的，而它一旦被用于决定谁被裁撤、哪个岗位被保留，争议就必须有地方去说。

■ 十四、赌注

一条写死日期的赌注，并写死什么不算命中。

>到二〇二九年十二月三十一日之前，至少有一个非航空行业的监管机构或全国性行业协会，发布一份要求保留或恢复某项已经可以被自动化的人工工序、并在文件中明确写出该工序所维持的是哪一项能力的强制性规定。

什么不算命中：倡议、指导意见、白皮书、「鼓励重视人才培养」一类的表述不算；只写「应加强培训」

而不指名具体工序的不算；行业自愿承诺不算；只在某一次审查中临时要求的不算。必须是常规适用的、可执行的条款，且文件里要写明这项工序与哪一项能力的因果关系——因为按第十一节，写出因果链才是那一步真正的动作。

之所以把航空排除在外，是因为它已经做到了；赌的是这件事能不能扩散到没有事故调查制度的行业。

■ 十五、本章自己在哪一格

本章自己就是它所描述的那类东西的一份标本，而且落在最不利的位置上。

这篇文章靠什么写出来？靠的是长期阅读大量不相干材料、反复写不成立的初稿、把别人的论证一句句读完这些低效活动顺带养成的判断力。而这些活动，正是眼下被生成工具优化得最彻底的那一批。

由此有两条本章必须承担的后果。

其一，本章不能主张自己不在丁格里。

判断一篇文章有没有内容，靠的正是那种被顺带养成的能力；如

果这种能力正在停止被养成，那么本章能不能被读出来，与本章写得好不好无关。

其二，也是更尖锐的一条：写下这篇文章，不构成对它所描述的过程的任何补救。

按本章自己的判断，

能力不会因为有人指出它在消失而重新开始形成；它只在那些活动被继续做的时候形成。因此本章最诚实的自我定位是——它是一份倒查因果链的提议，而提议本身不养成任何人。第十一节第一步之所以被排在最前面，正是因为它是唯一一件写一篇文章能推动的事：加一个问题进复盘流程，不需要任何人同意。

本章把这条活口留在明处。

没有人在养它

三门学科各自指认了一个持有者——资产的所有者、观看的一方、建造的人——而三家共同假定那样要紧的东西总归被某一方持有着。推翻这条假定的材料来自艺术学自己：观看的技能由与艺术无关的日常实践顺带养成，因此它不由艺术界持有，也不由任何人持有。

一旦这条假定倒下，就出现了一类三家账本上都没有名分的东西：被普遍依赖、却不由任何一方生产的能力。它不因使用而减少，只因养成它的那些低效活动被逐一取消而消失；而它的消失不产生任何事件，因为从来没有人持有过它。本章把它称为无供给品，并据此给出一条与三家都不同的判断——这一轮的困境不是竞争加剧、不是判据迁移、不是瓶颈转移，而是能力正在停止被养成，于是所有人只能用可以被生产的量，去补一个不可以被生产的缺口。

两处外来的证据把这条判断削得很重。住院医师工时改革的两项随机对照试验说明，取消低效暴露不必然导致能力下降——只要存在并行的养成结构；民航的手动飞行恢复说明，衰退可以被逆转——只要有人先把因果链倒查出来并写进文件。前者把适用范围缩到「唯一养成源」，后者把处方从「保留活动」改成「先建立倒查制度」。

出路因此落在一个不显眼的位置上：在取消任何一项之前，先问它是不是某项能力的唯一养成源；而要能回答这个问题，先得有一套把能力衰退倒查到具体活动的制度。这一步没有任何行业是自然具备的，航空是靠几十年的事故换来的。

最后一句关于这次尝试本身。同一个问题、同样三个学科，只换了三家的理论来源，得到的东西就落到了另一侧——从账本上缺一栏，变成生产链上缺一环。这既是本章的

，也是对这类做法的一个提醒：你请谁来答，很大程度上决定了答案长什么样；而这一点，本身值得被当作一个问题来查。

■ 注释

〔一〕本章所说的「养成」，指一个人在反复处理真实问题的过程中形成独立判断的能力，不包括可以经由规程与课程习得的操作技能。二者的区分正是第十二节条件三所要检验的。

〔二〕「无供给品」不是「无价值」，也不是「不可生产」。它的定义性质是当下没有任何一方以生产它为目的而行动；第十一节第三步所说的「把它从无供给变成有供给」，正是要建立这样一方。

〔三〕本章为理论建构与研究设计，不报告任何虚构的样本、统计结果或实验

。第七节的十处兑现是可检验的预测，不是已完成的观测；第十二节的样本设计尚未执行。

〔四〕第七节第九处所引的两项随机对照试验测量的是病人结局、考试成绩与住院医感受，不是本章所关心的独立判断形成速度。本章没有把这份不利证据读成对自己有利，而是在第十二节条件二里写出了填补这一差距的检验设计。

没有人在养它

〔五〕美国联邦航空局的安全警示属于建议性文件而非强制规章；交通部监察长办公室二〇一六年的审计指出，该局并未确认航空公司据此增加了手动飞行的机会。第七节第十处与第十四节的赌注均按这一实情表述，未把它读成一次已完成的回补。

〔六〕本章与本书同题五篇的分工见第九节；五篇均可在本书内查阅。

■ 人机分工声明

选题、三个学科的指定与最终定稿由王德生负责。文献检索、冲突定位、框架构建、写作与自检由 Claude 执行。

所有对外来源均给出可直接点开核对的地址；文中不报告任何虚构的经验数据，两处可核查的政策文件（美国联邦航空局的两份安全警示与交通部监察长办公室的审计报告）已逐项核对。

十项读数如何互相锁住

一、为什么必须做这一步

前四编交出了十个读数。到此为止，它们是十件并排放着的东西：各自有定义、有数据来源、有正交实例，也各自有证伪条件。这已经比一句感慨强得多，但它还不构成一个框架。

理由是明确的。十个并排的读数有一个致命的弱点——它们无法互相否定。如果十项各自独立，那么任何一项测出异常，都可以被解释成“这一项的定义有问题”；任何一项测出正常，也不构成对其余九项的约束。一套不能互相否证的读数，最终会退化为一份检查清单：谁都可以挑自己方便的那几项来算，算完之后说“我们这里没事”。

枢纽章已经示范过怎么走出这一步。它把编一的两项——未阅率与饱和关占比——用一条关系式锁在一起，得到的东西两章各自都写不出来：撤掉最饱和的那道关只降代价、不改善回报；扩产能会被投递量吃掉。这一章要做的是同样的动作，做到其余八项上。

先说清楚这一章能做到什么、不能做到什么。它不引入任何新的经验材料，只用前四编已经定义过的量。它所建立的关系，全部是从这些量的定义直接推出的会计恒等式或队列约束，而不是经验规律。这意味着两件事：其一，这些关系不需要新数据就可以在既有日志上检验；其二，如果检验不通过，被推翻的不是某个经验假设，而是某一个读数的定义本身出了问题——这比推翻一条经验命题更 severe，也更有价值。

二、记号

沿用枢纽章的记号，并把其余各项接进来。

符号	名称	出处
N	一轮的投递总量	枢纽章
Πp	各关通过率之积	第三章
K	下游一轮可过目的产能（份数）	枢纽章
u	未阅率	第一章
δ	暗基线抬升度	第二章
s	饱和关占比	第三章

a	认领占比	第四章
τ	辨认转嫁比	第五章
w	无量期占比	第六章
m	并项率	第七章
k	种类留存率	第八章
e	无验收期占比	第九章
σ	停养率	第十章

另设三个辅助量，都是可从工时记录直接读出的常数： h 为被评价方为使一件产出可被辨认所追加的平均工时， c 为评价方核实一件产出的平均工时， A 为一轮的录取数。

■ 三、第一条：拥堵参数把四项拴在一起

枢纽章给出 $u = \max(0, 1 - K/(N \cdot \Pi p))$ 。把括号里那一项记作拥堵参数：

$$> \rho = N \cdot \Pi p / K$$

它的含义是：下游每一份可过目产能，要面对多少份通过了关的投递。 $\rho \leq 1$ 时不拥堵， $\rho > 1$ 时拥堵。于是 $u = \max(0, 1 - 1/\rho)$ 。

第一条锁定来自第六章。无量期占比 w 定义为一样东西从被做出来到取得第一个读数所占的时间比例。在一个先到先服务、每轮清 K 份的队列里，第 n 份到达者的等待轮数是 $(n - K)/K$ 。对整批取平均，积压部分的平均等待轮数是：

$$> \bar{w} = (N \cdot \Pi p - K) / K = \rho - 1$$

而 $u = 1 - 1/\rho$ ，即 $\rho = 1/(1 - u)$ 。代入得到一条不含任何经验参数的关系：

$$> \bar{w} = u / (1 - u)$$

这条式子是这一章最干净的一条。它说：首读时距与未阅率不是两件事，它们是同一个拥堵参数的两种读法。未阅率 0.5 时平均等待一轮，0.8 时四轮，0.9 时九轮，0.95 时十九轮——等待时间在未阅率逼近一时发散。

它同时是可证伪的，而且检验成本极低：任何一个既记录了“附件是否被打开”、又记录了“提交时间与首次评审时间”的系统，都可以直接把 u 与 $\bar{w}$ 算出来比对。若在同一批日志里 $\bar{w}$ 显著偏离 $u/(1-u)$ ，则两个读数中至少有一个的定义与实际记账口径不符——最可能的情形是队列不是先到先服务，而是有优先通道，这一点本身就是需要被查出来的东西。

第一条因此把第一章、第三章（经由 Πp 与 s 的关系）与第六章锁在一起。

■ 四、第二条：辨认转嫁比由认领占比与拥堵共同决定

第五章的辨认转嫁比 τ 定义为被评价方为可被辨认所投入的时间，除以评价方为核实所投入的时间。两边分别写出来：

分子是全部投递者追加的辨认真时，等于 $N \cdot a \cdot h$ （每件追加占比 a ，单件追加工时 h ）。分母是评价方实际花在核实上的工时——注意，只有被打开过的那些才产生核实工时，因此是 $K \cdot c$ 。于是：

$$\tau = (N \cdot a \cdot h) / (K \cdot c) = \rho \cdot a \cdot (h/c)$$

再把 $\rho = 1/(1-u)$ 代入：

$$\tau = a \cdot (h/c) / (1-u)$$

这条式子给出三件前面各章都没有说出的事。

其一，认领占比与辨认转嫁比不是两个独立的病，是同一个量在两侧的两种记法。第四章量的是一件产出内部的构成比，第五章量的是两侧的工时比，而它们之间只差一个拥堵参数与一个工时常数比。

其二，未阅率会放大辨认转嫁比，而且是发散式的放大。 a 不变、 h/c 不变时，未阅率从 0.5 涨到 0.9， τ 涨四倍——被评价方多付的那些辨认真时，有越来越大的比例落在根本没有人打开的材料上。这是本书全部读数里最刺人的一条推论：在高未阅率的通道里，认领物的绝大部分是为没有人看的那一眼准备的。

其三，它给出一条方向明确的反向约束。第四章原本可以被读成“认领占比越低越好”。这条式子说不是：若一个机构把 a 压到接近零（完全实名、自动溯源、来源链完备）， τ 随之下降，而 u 与 ρ 一动不动——拥堵没有被触碰，只是被评价方不再为它付钱了。认领占比是代价的读数，不是拥堵的读数。降低它是好事，但它不改善有据判定率。这与枢纽章那条“撤掉饱和的关只降代价不改善回报”是同一种结构，出现在链的另一端。

■ 五、第三条：并项率使前面所有的量被系统性高估

第七章指出，生成成本趋零之后“一件”这个计数单位失去了下界。这件事在这个关系网里有一个精确的后果。

设本期计入的件数为 N ，其中若按上期的计数约定归并，会被并成一件的比例为 m 。则按上期口径的实际投递量是 $N(1-m)$ 。把它代回拥堵参数：

$$\rho_{\text{真}} = N(1-m) \cdot \Pi p / K = \rho \cdot (1-m)$$

于是：

$$u_{\text{真}} = 1 - 1/[\rho(1-m)], \text{ 而观测到的 } u = 1 - 1/\rho$$

因为 $m \geq 0$, 恒有 $\rho_{\text{真}} \leq \rho$, 因此 $u_{\text{真}} \leq u$ 。也就是说：并项率使未阅率被系统性高估, 进而使首读时距与辨认转嫁比一并被高估。

这不是一处技术性修正, 它改变结论的方向。一个机构如果观测到未阅率上升, 有两种完全不同的可能: 一种是拥堵确实加剧, 另一种是每个人把同一份东西拆成了更多件投出来, 而拥堵没有变。两者的处方相反——前者要改 N 对失败的反应函数, 后者要改计数约定。在算出 m 之前, u 的任何跨期比较都不成立。

第七章原本只主张"件数上升不再能充当投入上升的证据"。接进这个关系网之后, 它的主张变得更强也更具体: 它是本书其余所有读数的跨期可比性条件。 m 不测, 前四编的读数只能同期横比, 不能跨期纵比。

■ 六、第四条：种类留存率决定这些读数有没有分辨对象

第八章的种类留存率 k 是本期实际使用的分辨类别数除以上期。它在这个关系网里的位置与 m 对称, 但作用在另一个维度上。

枢纽章定义的有据判定率 $r = 1 - (\Pi p) \cdot u$, 读的是"结果确实取决于内容"的概率。但"取决于内容"预设了内容之间是可分辨的。若类别数从 D_0 降到 $D_1 = k \cdot D_0$, 则原本落在不同类别的产出中, 有一部分被并进同一类, 判定对它们给出相同结果——这部分判定在形式上取决于内容, 实质上不取决于这一件的内容。

定义有效判定率:

$$> r_{\text{eff}} = r \cdot k = [1 - (\Pi p) \cdot u] \cdot k$$

这条式子的意义在于它使第八章可被否定。第八章原本的主张——分辨力是一项没有人生产的产出——很难直接检验, 因为"分辨力"听起来是个品质。改写成 r_{eff} 之后, 它变成一个可以在同一批日志里算出的数, 并且给出一个方向明确的预测: 一个把评分一致性做到最高的评审体系 (k 最低), 其 r_{eff} 可以低于一个争议不断的体系, 即使两者的 r 相同。

这也解释了本书开头那个悖论的最后一块。为什么现行读数一定看不见这件事? 因为现行读数量的是 r 那一层的东西(通过率、周期、吞吐), 而 k 从来不进任何账——类别数减少在所有治理语言里都叫"标准更清晰"。

■ 七、第五条：编四的两项与前面八项的接口

前六节把编一、编二、编三的八项锁进了同一个系统。编四的两项——无验收期占比 e 与停养率 σ ——不在这个系统里, 而这不是遗漏, 是本书的结构本身。

前八项都是存量读数：它们读的是某一轮判定的构成，可以在一轮之内算出来。编四的两项是流量读数：它们读的是判断力这份能力在轮与轮之间的补充与损耗。两者的接口只有一处，而且是单向的：

> 前八项决定这一轮的判定质量；编四两项决定下一轮还有没有人能做这个判定。

这个接口可以写成一个最朴素的守恒式。设 J 为一个系统中判断力的存量，则：

$$> J(t+1) = J(t) \cdot (1 - \sigma) + G(e)$$

其中 σ 是停养率（本期被优化掉的、原本顺带养成该能力的活动占比）， $G(e)$ 是本期由无验收期生成的判断力增量。第九章的全部论证可以压成一句： G 是 e 的增函数，而 e 在任何一份交付账上都表现为待优化的空白。第十章的全部论证可以压成另一句： G 不由任何一方生产，因此没有人对 G 负责；而 σ 由每一次流程优化自动产生，因此每个人都在贡献 σ 。

于是本书最后那条判断有了形式： σ 有无数个生产者而 G 没有一个，因此在没有任何人做错事的情况下， J 单调下降。而 J 的下降不产生任何事件——因为 J 不在任何账上，前八项读数读的全是 J 被使用时的表现，不是 J 本身。

这也给出编四两项唯一可行的检验方式：不能直接测 J ，只能测 σ 与 e 的变化，然后看前八项读数在若干轮之后是否随之恶化。这是一个滞后检验，周期以年计。本书对此不作乐观估计。

■ 八、关系网的全貌，与它能被怎样推翻

把五条并起来：

$$> \rho = N \cdot \prod p / K \quad u = 1 - 1/\rho \quad \bar{w} = u/(1-u) > \tau = a \cdot (h/c)/(1-u) \quad \rho_{\text{真}} = \rho(1-m) \\ r_{\text{eff}} = [1 - (\prod p) \cdot u] \cdot k > J(t+1) = J(t)(1-\sigma) + G(e)$$

十项读数至此全部进入同一个系统： u 、 s （经 $\prod p$ ）、 w 、 τ 、 a 、 m 、 k 由前六式互相约束， δ 由第二章的机制经 $(1-u)$ 与它们相连， e 与 σ 经第七式构成下一轮的边界条件。

这个关系网的价值不在于它复杂，而在于它使十项读数可以互相否定。前四编交出的十项，每一项单独都可以被解释掉；接进这个系统之后，它们必须同时成立。因此本章给出的推翻方式比前四编任何一章都更直接：

其一， $\bar{w}$ 与 $u/(1-u)$ 在同一批日志上显著不符。则第一章与第六章的读数中至少有一个定义与实际记账口径不符。这是最容易做的一条检验，也是本章最愿意先被检验的一条。

其二， τ 与 $a/(1-u)$ 不成比例。若在控制 h/c 之后仍不成比例，则第四章与第五章量的不是同一件事的两侧，两章之一应当被重写。

其三，把 m 算出来之后， u 的跨期上升消失。则第一章观测到的未阅率上升是计数约定变化的假象，本书编一的经验基础被抽掉一半。

其四， k 与 r_{eff} 无关。若在类别数显著下降的体系里，判定结果对内容的依赖并未下降，则第八章的核心主张不成立。

其五，也是最根本的一条： σ 上升多年之后，前八项读数并未随之恶化。则编四与前三编之间没有那个接口，本书的四编不构成一本书，只是十篇各自成立的文章的合集。

最后一条值得单独说一句。它是本书唯一一条关于这本书自身结构的证伪条件：如果编四与前三编之间那个接口不存在，那么这本书的正确编法就不是四编十章加一条脊梁，而是把编四拆出去另立一册。本章把这条写在这里，是因为它比任何一章内部的证伪条件都更能决定这本书该不该存在。

结语

不发警报的危机怎么被发现

■ 一、十章说了同一件事

把十章的承重判断并排放，会看到它们是同一句话的十个位置。

第一章说，一次否决可以在没有任何人打开过那份东西的情况下作出，而它与读完之后作出的否决在所有账本上同码——判断没有发生，而记录读不出来。

第二章说，被丢弃的那些产出在无人记账处合成了下一轮的判定基线，并单向抬升它——判断的标准由无账可记之物构成。

第三章说，一道关可以在最近一百次判定里一次也没拦下过谁，而它不会被撤销，只会被继续满足，因为它没有作者——判断装置失效，而没有位置可以宣布它失效。

第四章说，一件产出里有一层是为了证明它出自谁手，三家账本只给它成本位、不给产出位——判断所需的归属证据必须被单独生产，而它不算产出。

第五章说，核实这件工作从评价方整体转嫁给被评价方，且因为与真实产出同形而无法被分离计量——判断的成本被转移，而转移读不出来。

第六章说，一样东西从被做出来到取得第一个读数所需的时间，落在效率读数的分母内、分子外——判断所需的时间被计为损耗。

第七章说，生成成本趋零之后"一件"这个单位失去下界，件数不再能充当投入的证据——判断所依赖的计数约定失效。

第八章说，可分辨的种类数不是系统的自然属性而是一项产出，且没有人在生产它——判断所依赖的分辨约定被当作背景。

第九章说，验收能力是在那段不产出任何可验收之物的时间里被生产的，而生成式协助恰好把那段时间前置掉了——判断力的生产环节被当作空转删除。

第十章说，一类被普遍依赖、却不由任何一方生产的能力，只通过养成它的那些低效活动被逐一优化掉而消失，且消失不产生任何事件——判断力没有供给方。

十条合起来是母题的十个侧面：判断力无人生产、无账可记，只由低效活动顺带养成；它消失时不发生任何事件。

■ 二、为什么现行读数一定看不见它

有人会问：既然这件事这么大，为什么没有一个指标报警？

因为全部现行读数共享同一个设计前提——发生了事情会留下记录。而本书描述的这一类事，其定义性特征恰恰是不留记录。用现行读数去找它，找不到不是反证，是同义反复。

这一点可以说得更硬。现行读数分三类，三类各自都有结构性的盲区，而三类盲区正好合成本书的对象。

产出类读数（吞吐、产量、发表数、录用数、交付件数）只计成功通过的那些。未被接受的产出对它们的贡献是零——不是小，是没有位置去记。于是第一、第二章描述的东西，在产出类读数上不可能出现。

效率类读数（周期、成本、单位工时、到岗时间）计的是从投入到产出的距离。凡是暂时不产出的环节，在这类读数上一律表现为待优化的损耗。于是第六、第九章描述的东西，在效率类读数上不但看不见，而且方向是反的——它们越是被消灭，读数越好看。

质量类读数（通过率、缺陷率、录取者资历、满意度）计的是通过之物的性质。它们默认判定装置本身是可靠的，因而无法读出装置的失效。于是第三、第五、第八章描述的东西，在质量类读数上表现为改善：通过率上升被读作把关有效，被评价方投入上升被读作准备更充分，分辨类别减少被读作标准更清晰。

三类读数把同一件事分别读成了零、损耗和改善。这不是测量精度问题，是记账口径问题——没有一类读数是“判断这件工作本身”设计的，因为在生成成本坍塌之前，它从来不是稀缺的那一样。

■ 三、体检表

本书唯一的兑现物是下面这张表。十个读数各自来自一章，每一个都满足四条：从既有日志里可以算出来、不需要新建采集系统、无量纲因而可跨行业比较、与该行业现行主要指标正交（也就是说，主要指标全绿时它可以很难看）。

一·未阅率定义：上一轮的全部否决里，作出否决时没有任何人打开过那份材料的比例。
数据来源：投递系统的被动访问日志（附件是否被打开、页面停留是否发生）。正交实例：答复时限压到四十八小时、积压清零、满意度最高的招聘流程，可以有接近一的未阅率。

二·暗基线抬升度定义：把上一轮全部未被接受的产出移出参照集之后，本轮合格线下降的幅度，除以本轮合格线。数据来源：决定日志 + 一次盲掉年份的重判。正交实例：录用数最多、到岗时间最短的机构，可以有最高的抬升度，恰恰因为它筛得快、每份材料被看的时间最短。

三·饱和关占比定义：全部关口中，最近一百次判定的通过率不低于零点九五的那些所占的比例。数据来源：各关口的判定记录。正交实例：一条从来没有红过的持续集成流水线，交付指标完美，而这个比率等于一。

四·认领占比定义：一次交付的总工作量里，为使这件产出可被归属于某人而追加的那一部分所占的比例。数据来源：工时记录、提交历史、材料清单的分项。正交实例：交付准时率最高、缺陷率最低、测试覆盖率满分的团队，可以有最高的认领占比。

五·辨认转嫁比定义：被评价方为使产出可被辨认所投入的时间，除以评价方为核实所投入的时间。数据来源：两侧的工时记录。正交实例：引入自动核验而不撤销旧要求时，这个比可以暴涨，而两侧的账面效率同时改善。

六·无量期占比定义：一样东西从被做出来到取得第一个读数所经过的时间，占其全部在制时间的比例。数据来源：版本时间戳与首次评审时间戳之差。正交实例：一个把首读时距压到最短的评审系统，可以同时是最不可能发现新东西的系统。

七·并项率定义：本期计入件数的产出里，若按上一期的计数约定重新归并，会被并成一件的那部分所占比例。数据来源：两期的产出清单 + 一次按旧约定的重新归并。正交实例：人均件数翻倍的团队，并项率可以接近一——件数上升与投入上升无关。

八·种类留存率定义：本期系统实际使用的分辨类别数，除以上一期的类别数。数据来源：评分表、标签体系、分类字典的版本比对。正交实例：评分一致性最高、争议最少的评审体系，种类留存率可以最低——一致是因为类别少。

九·无验收期占比定义：一段学习或训练过程中，不产出任何可被验收之物的时间所占的比例。数据来源：课程与项目的时间表、交付物清单。正交实例：交付物最密集、进度最可控的培养方案，这个比例最接近零。

十·停养率定义：曾经顺带养成某项能力的那些活动，在本期被自动化或优化掉的比例。数据来源：流程变更记录、工具上线清单。正交实例：这个数没有正交实例可举，因为目前没有任何机构在测它——这本身就是第十章那条判断的一个读数。

■ 四、这张表怎么用

第一，先算三个，不要十个都算。任何一个机构从第一、第三、第六这三个开始：未阅率、饱和关占比、无量期占比。这三个的数据在既有日志里最完整，算一遍的成本最低，而且它们分处链上的三个不同位置——一个在否决侧，一个在门槛侧，一个在时间侧。三个都绿，本书对这个机构不适用；任何一个显著异常，再往下算。

第二，这些数指的是位置，不是人。它们读的是某一道判定工序的参照结构，不是某一位编辑、某一位评审、某一位招聘专员的品行或能力。全书的判断恰恰是这件事没有作者。用这张表去考核具体的人，是把一个关于制度的读数错当成一个关于个体的读数。

第三，也是最重要的一条：这张表进考核的那一天，它就开始失效。每一个读数都有一条最省事的达标路径，而每一条都会摧毁它要保护的那样东西。未阅率的最省事达标法是让系统自动"打开"每一份，或规定"模型已读即算已读"；暗基线抬升度的最省事达标法是删掉落判池——本机构不保留未录用者的材料，于是这个数永远算不出来，而机制照旧运行；停养率的最省事达标法是不记录流程变更。

这一条没有出路。本书能做的只是把它写在明处：这张表是给想知道的人用的，不是给需要交差的人用的。一个机构如果只是需要一份好看的报表，它有一百种办法把这十个数做绿，而本书描述的那件事会照常发生，只是从此连间接读数也没有了。

■ 五、如果这本书错了

一本主张"某样东西正在消失"的书，最大的风险是它不可证伪。所以最后写清楚：本书在什么情况下应当被整体放弃。

其一，剂量关系不成立。十章各自给出了机制层面的证伪条件，而全书层面的那一条是：控制投递量与判定者人时之后，若上述读数与合格线漂移之间不存在剂量—反应关系，则门槛上升就确实只是绝对负荷问题，处方回到扩容。那意味着"被认真接住的容量"可以靠资源放大——这与本书对稀缺性质的全部判断相反。

其二，重建通道存在。第十章主张，只由低效活动顺带养成的能力，在那些活动被优化掉之后没有重新供给的通道。若有机机构成功地把这类能力做成了可以被直接生产、直接采购、直接培训的东西，并且产出的判断力在盲测中不劣于顺带养成的那一种，则第十章错，全书的最后一编随之作废。

其三，消失产生事件。若在某个已经把相关低效活动全部优化掉的行业里，判断力的下降确实触发了某个既有指标的报警——不是事后追认，是当期报警——则本书副题所依赖的那条判断错了。届时这件事仍然可能是真的，但它不再是一场不发警报的危机，而只是一场被忽视的危机；这两者的处方完全不同。

其四，定标制里也发生。本书多处主张，若干机制只在名额先定死的择优场合发生，不在达标即过的定标场合发生。若在纯定标的通道——资质考试、安全认证、药品与器械审评——里同样测到这些读数显著异常，则本书最重要的一处区分错了。

四条里任何一条被证实，本书应当被修正或放弃。而四条都可以在两三年内、用既有数据做出来。

■ 六、最后一句

本书从头到尾在说的，其实是一件很朴素的事。

我们把"做出一件东西"的成本降到了接近于零，却没有同时问一句：那么，是谁在看？看这件事的能力从哪里来？它有没有人在生产？它的账记在哪一栏？

四个问题都没有答案，而没有答案本身不产生任何事件——所有的读数照旧全绿，所有的报表照旧可以按时提交。

这就是这场危机的形状：它不会发出任何警报。

参考文献

本书十章的文献按章分列。全部来源均为公开文献，可自行核对；本书新写的导论、枢纽章导入与推广、合章、结语与附录不引入新的外部文献，其所用的量与关系式均在正文中定义。

■ 第一章·两种「不」

- Acemoglu, D. (2025). The simple macroeconomics of AI. *Economic Policy*, 40(121), 13– 58. <https://doi.org/10.1093/epolic/eiae042>
- Becker, J., Rush, N., Barnes, B., & Rein, D. (2025). Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. METR. [arXiv:2507.09089](https://arxiv.org/abs/2507.09089). <https://arxiv.org/abs/2507.09089>
- METR (2026). We are changing our developer productivity experiment design. <https://metr.org/blog/2026-02-24-uplift-update/>
- Lopes, D. M. (2018). *Being for Beauty: Aesthetic Agency and Value*. Oxford University Press
- Will AI Produce Works of Extraordinary Aesthetic Value? (2026). *The Journal of Aesthetics and Art Criticism*, advance article. <https://doi.org/10.1093/jaac/kpag009>
- Dickie, G. (1974). *Art and the Aesthetic: An Institutional Analysis*. Cornell University Press
- Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690
- Baumol, W. J., & Bowen, W. G. (1966). *Performing Arts: The Economic Dilemma*. Twentieth Century Fund
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355– 374

- Roth, A. E., & Xing, X. (1997). Turnaround time and bottlenecks in market clearing. *Journal of Political Economy*, 105(2), 284–329
- Coles, P., Cawley, J., Levine, P. B., Niederle, M., Roth, A. E., & Siegfried, J. J. (2010). The job market for new economists: A market design perspective. *Journal of Economic Perspectives*, 24(4), 187–206
- Goldratt, E. M., & Cox, J. (1984). *The Goal: A Process of Ongoing Improvement*. North River Press
- Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest*. Johns Hopkins Press. 两种「不」
- Forsgren, N., Humble, J., & Kim, G. (2018). *Accelerate: The Science of Lean Software and DevOps*. IT Revolution Press
- More Versus Better: Artificial Intelligence, Incentives, and the Emerging Crisis in Peer Review (2026). *Organization Science*, 37(3), editorial. <https://pubsonline.informs.org/doi/10.1287/orsc.2026.ed.v37.n3>
- Li, X., Song, Z., & Zhang, J. (2025). Accept More, Reject Less: Reducing up to 19% Unnecessary Desk-Rejections over 11 Years of ICLR Data. arXiv:2506.20141. <https://arxiv.org/abs/2506.20141>
- Clarke, N. (2025). How "AI" submissions have changed our submissions process. <https://neil-clarke.com/how-ai-submissions-have-changed-our-submissions-process/>
- Is Growth Always Good News? 2026 Article Submission Surges (2026). The Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2026/05/13/guest-post-is-growth-always-good-news-2026-article-submission-surges/>
- Recruiters are "drinking through a fire hose" of job applications (2025). CNBC. <https://www.cnbc.com/2025/10/29/recruiters-are-drinking-through-a-fire-hose-of-jobapplications-experts-say.html> 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 + Claude · 作者：王德生

■ 第二章 · 被丢掉的那一眼

- Tullock, G. (1980). Efficient Rent Seeking. In J. Buchanan, R. Tollison & G. Tullock (eds.), *Toward a Theory of the Rent-Seeking Society*. Texas A&M University Press
- Dechenaux, E., Kovenock, D. & Sheremeta, R. (2015). A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments. *Experimental Economics*, 18(4), 609– 669
- Dickson, A., MacKenzie, I. A. & Sekeris, P. G. (2022). Rent Dissipation in Simple Tullock Contests. *Games*, 13(6), 83. <https://doi.org/10.3390/g13060083>
- Groys, B. (1992/2014). *Über das Neue / On the New*. Verso
- Groys, B. (2000). *Under Suspicion: A Phenomenology of the Media*. Columbia University Press
- Doshi, A. R. & Hauser, O. P. (2024). Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content. *Science Advances*, 10(28), eadn5290. 被丢掉的那一眼
- Goldratt, E. M. & Cox, J. (1984). *The Goal: A Process of Ongoing Improvement*. North River Press
- Goldratt, E. M. (1990). *Theory of Constraints*. North River Press
- DORA / Google Cloud (2025). *State of AI-assisted Software Development*
- Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press
- Collins, R. (1979/2019). *The Credential Society: An Historical Sociology of Education and Stratification*. Columbia University Press
- Hirsch, F. (1976). *Social Limits to Growth*. Harvard University Press
- Frank, R. H. & Cook, P. J. (1995). *The Winner-Take-All Society*. Free Press
- Van Valen, L. (1973). A New Evolutionary Law. *Evolutionary Theory*, 1, 1–30
- Jevons, W. S. (1865). *The Coal Question*. Macmillan
- Simon, H. A. (1971). Designing Organizations for an Information-Rich World. In M. Greenberger (ed.), *Computers, Communications, and the Public Interest*. Johns Hopkins Press
- Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3), 488–500

- Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87(3), 355–374
- Merton, R. K. (1940). Bureaucratic Structure and Personality. *Social Forces*, 18(4), 560–568
- Pigou, A. C. (1920). *The Economics of Welfare*. Macmillan
- Coase, R. H. (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3, 1–44
- Kleinberg, J. & Raghavan, M. (2021). Algorithmic Monoculture and Social Welfare. *PNAS*, 118(22), e2018340118
- Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R. & Gal, Y. (2024). AI Models Collapse When Trained on Recursively Generated Data. *Nature*, 631, 755–759
- Levinthal, D. A. & March, J. G. (1993). The Myopia of Learning. *Strategic Management Journal*, 14(S2), 95–112
- Stelmakh, I., Shah, N. B. & Singh, A. (2021). Prior and Prejudice: The Novice Reviewers' Bias against Resubmissions in Conference Peer Review. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1)
- Greenhouse (2025). *AI in Hiring Report*
- Organization Science AI Task Force (2026). More Versus Better: Artificial Intelligence, Incentives, and the Emerging Crisis in Peer Review. *Organization Science*, 37(3). 被丢掉的那一眼 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。作者：王德生 + Claude ·

■ 第三章 · 这道关拦下过谁

- Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3), 488–500
- Rein, D. (2025). Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. *METR*. arXiv:2507.09089
- Syverson, C. (2021). The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *American Economic Journal: Macroeconomics*, 13(1), 333–372

- Campbell, D. T. (1979). Assessing the Impact of Planned Social Change. *Evaluation and Program Planning*, 2(1), 67–90. 这道关拦下过谁
- Collins, R. (1979). *The Credential Society: An Historical Sociology of Education and Stratification*. Academic Press. DORA / Google Cloud (2025). State of AI-assisted Software Development. (2024 年版为 Accelerate State of DevOps Report 2024。)
- Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press
- Goodhart, C. A. E. (1975). Problems of Monetary Management: The U.K. Experience. In *Papers in Monetary Economics*, Reserve Bank of Australia
- Barton, D. (1992). Core Capabilities and Core Rigidities. *Strategic Management Journal*, 13(S1), 111–125
- March, J. G. (1993). The Myopia of Learning. *Strategic Management Journal*, 14(S2), 95–112
- Lopes, D. M. (2014). *Beyond Art*. Oxford University Press
- Lopes, D. M. (2026). AI Art and Artists: What They Are, What They Could Be, What They Should Be. *The Journal of Aesthetics and Art Criticism*. doi:10.1093/jaac/kpag001
- Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley
- Merton, R. K. (1940). Bureaucratic Structure and Personality. *Social Forces*, 18(4), 560–568
- Nguyen, C. T. (2024). Value Capture. *Journal of Ethics and Social Philosophy*, 27(3). doi:10.26556/jesp.v27i3.3048
- Selznick, P. (1949). *TVA and the Grass Roots*. University of California Press
- Simon, H. A. (1971). Designing Organizations for an Information-Rich World. In M. Greenberger (ed.), *Computers, Communications, and the Public Interest*. Johns Hopkins University Press
- Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87(3), 355–374. 投稿量与投递量数据：领英二〇二五年投递量同比增长四成五（经《纽约时报》报道并由多家媒体转载）；AAAI 2026 初投稿超过三万份，2025 年约一万五千份（见 AI 辅助评审试点报告，arXiv:2604.13940）；Organization Science 编辑部对 2022 年 1 月至 2023 年 11 月与 2024 年 1 月至 2025 年 11 月两段投稿量的对比（见该刊 2026 年编辑部文

章, doi:10.1287/orsc.2026.ed.v37.n3)。本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献,可自行核对。+ Claude · 作者: 王德生

第四章 · 认领物

- Acemoglu, D. The Simple Macroeconomics of AI. NBER Working Paper No. 32487. [https:// www.nber.org/system/files/working_papers/w32487/w32487.pdf](https://www.nber.org/system/files/working_papers/w32487/w32487.pdf) 2. Some Simple Economics of AGI. arXiv:2602.20946. <https://arxiv.org/pdf/2602.20946> 3. AI, Productivity, and Labor Markets: A Review of the Empirical Evidence. International Center for Law & Economics. [https://laweconcenter.org/ resources/ai-productivity-and-labor-markets-a-review-of-the-empirical-evidence/](https://laweconcenter.org/resources/ai-productivity-and-labor-markets-a-review-of-the-empirical-evidence/) 4. Could AI Replace Human Art Experts in Attributing Paintings? Aeon Essays. <https://aeon.co/essays/could-ai-replace-human-art-experts-in-attributing-paintings> 5. Content Authenticities: A Discussion on the Values of Provenance Data for Creatives and Their Audiences. Proceedings of the 2025 Conference on Creativity and Cognition. <https://dl.acm.org/doi/10.1145/3698061.3726918> 6. C2PA for Creators & Artists: Proving Your Work Is Yours. <https://c2pa.ai/for-creators> 7. Artificial Intelligence and Art Authentication. Hephaestus Analytical. <https://www.hephaestusanalytical.com/blog/artificialintelligence-and-the-detection-of-forgery-in-art-authentication> 8. Debt Behind the AI Boom: A Large-Scale Empirical Study of AI-Generated Code in the Wild. arXiv:2603.28592. [https://arxiv.org/ html/2603.28592v2](https://arxiv.org/html/2603.28592v2) 9. How AI Generated Code Compounds Technical Debt. LeadDev. [https:// leaddev.com/technical-direction/how-ai-generated-code-accelerates-technical-debt](https://leaddev.com/technical-direction/how-ai-generated-code-accelerates-technical-debt) 10. AI Coding Has a Trust Problem: Sonar Data Shows Verification Lagging Far Behind Adoption. TFiR. [https:// tfir.io/ai-coding-has-a-trust-problem-sonar-data-shows-verification-lagging-far-behind-adoption/](https://tfir.io/ai-coding-has-a-trust-problem-sonar-data-shows-verification-lagging-far-behind-adoption/) 11. Verification Debt: The Hidden Quality Cost of AI-Generated Code. [https://techintelix.com/ verification-debt/](https://techintelix.com/verification-debt/) 12. More Code, Less Reuse: Investigating Code Quality and Reviewer Sentiment towards AI-generated Pull Requests. arXiv:2601.21276. <https://arxiv.org/pdf/2601.21276> 13
- Spence, M. Job Market Signaling. Quarterly Journal of Economics, 87(3), 1973. 14
- Coase, R. The Nature of the Firm. Economica, 4(16), 1937;

- Williamson, O. Markets and Hierarchies. Free Press, 1975. 15
- Hochschild, A. R. The Managed Heart: Commercialization of Human Feeling. University of California Press, 1983. 16
- March, J. G. The Myopia of Learning. Strategic Management Journal, 14(S2), 1993;
- Miller, D. The Icarus Paradox. HarperBusiness, 1990;
- Barton, D. Core Capabilities and Core Rigidities. Strategic Management Journal, 13(S1), 1992;
- Geertz, C. Agricultural Involution. University of California Press, 1963. 17. 荷兰伦勃朗研究项目（Rembrandt Research Project, 1968–2011）及其成果《伦勃朗绘画全集》（A Corpus of Rembrandt Paintings, Vols. I–VI）。18. United Nations et al. System of National Accounts.（关于生产边界与非市场生产的处理）认领物 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。+ Claude
- 作者：王德生

■ 第五章·认不出来，就得自己交代

- 达龙·阿西莫格鲁、帕斯夸尔·雷斯特雷波：《任务、自动化与美国工资不平等的上升》，《计量经济学》第 90 卷第 5 期，2022 年。<https://doi.org/10.3982/ECTA19815>
- 达龙·阿西莫格鲁：《人工智能之害》，美国国家经济研究局工作论文第 29247 号，2021 年。<https://www.nber.org/papers/w29247>
- 戴维·奥托、卡罗琳·秦、安娜·萨洛蒙斯、布赖恩·西格米勒：《新的前沿：新工作的起源与内容，1940–2018》，《经济学季刊》第 139 卷第 3 期，2024 年。<https://doi.org/10.1093/qje/qjae008>
- 齐格·爱泼斯坦、亚伦·赫兹曼等：《艺术与生成式人工智能的科学》，《科学》第 380 卷第 6650 期，2023 年，第 1110–1111 页。<https://doi.org/10.1126/science.adh4451>
- 亚伦·赫兹曼：《计算机能创作艺术吗？》，《艺术》第 7 卷第 2 期，2018 年。<https://doi.org/10.3390/arts7020018>
- 艾哈迈德·埃尔加马尔等：《创造性对抗网络》，国际计算创造力大会，2017 年。<https://arxiv.org/abs/1706.07068>
- 梅尔文·康威：《委员会如何发明？》，《数据化》，1968 年。https://www.melconway.com/Home/Committees_Paper.html 艾伦·麦考马克、卡

- 莉斯·鲍德温、约翰·鲁斯纳克：《探索产品架构与组织架构的对偶性》，《研究政策》第 41 卷第 8 期，2012 年。<https://doi.org/10.1016/j.respol.2011.11.005>
- 戴维·斯卡利等：《机器学习系统中隐藏的技术债》，神经信息处理系统大会，2015 年。
。https://papers.nips.cc/paper_files/paper/2015/hash/86df7dcfd896fcdf2674f757a2463eba-Abstract.html 坦尼亚·雷利：《当胶水》，2019 年。<https://noidea.dog/glue> 认不出来，就得自己交代 迈克尔·斯彭斯：《求职市场信号》，《经济学季刊》第 87 卷第 3 期，1973 年。<https://doi.org/10.2307/1882010>
- 乔治·阿克洛夫：《柠檬市场：质量的不确定性与市场机制》，《经济学季刊》第 84 卷第 3 期，1970 年。<https://doi.org/10.2307/1879431>
- 戈登·图洛克：《高效的寻租》，载《寻租社会的理论》，得克萨斯农工大学出版社，1980 年。
- 克利福德·格尔茨：《农业内卷化：印度尼西亚的生态变迁过程》，加利福尼亚大学出版社，1963 年。
- 赫伯特·西蒙：《为信息丰富的世界设计组织》，载《计算机、通信与公共利益》，约翰斯·霍普金斯大学出版社，1971 年。<https://digitalcollections.library.cmu.edu/node/31491>
- 罗伯特·默顿：《官僚结构与人格》，《社会力量》第 18 卷第 4 期，1940 年。<https://doi.org/10.2307/2570634>
- 丹尼尔·列文塔尔、詹姆斯·马奇：《学习的近视》，《战略管理杂志》第 14 卷增刊，1993 年。<https://doi.org/10.1002/smj.4250141009>
- 丹尼·米勒：《伊卡洛斯悖论：卓越的公司如何把成功变成失败》，哈珀商业出版社，1990 年。
- 兰德尔·柯林斯：《文凭社会：教育与分层的历史社会学》，学术出版社，1979 年。
- 詹姆斯·斯科特：《国家的视角：那些试图改善人类状况的项目是如何失败的》，耶鲁大学出版社，1998 年。
- 查尔斯·古德哈特：《英国货币经验中的货币关系》，载英格兰银行论文集，1975 年；唐纳德·坎贝尔：《评估计划性社会变革的影响》，《评估与项目规划》第 2 卷第 1 期，1979 年。[https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X) 人机分工声明 选题、三个学科的指定与最终定稿由王德生负责。文献检索、冲突定位、框架构建、写作与自检由 Claude 执行。所有对外来源均给出可直接点开核对的地址；文中不报告任何虚构的经验数据。

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

■ 第六章 · 第一个读数之前

- Akhtar, M. et al. (2026). When AI Benchmarks Plateau: A Systematic Study of Benchmark Saturation. arXiv:2602.16763
- Auslander, P. (2023). Liveness: Performance in a Mediatized Culture, 3rd edition. Routledge
- Auslander, P. (2012). Digital Liveness: A Historico-Philosophical Perspective. PAJ: A Journal of Performance and Art, 34(3), 3–11
- Bean, A. et al. (2025). 对 445 篇基准论文的系统检查（不确定性估计与统计检验的使用比例）。
- Webb, M. (2020). Are Ideas Getting Harder to Find? American Economic Review, 110(4), 1104–1144. doi:10.1257/aer.20180338
- Liang, P. et al. (2023). Holistic Evaluation of Language Models. Transactions on Machine Learning Research
- Eriksson, M. et al. (2025). Can We Trust AI Benchmarks? European Commission Joint Research Centre
- Feyerabend, P. (1975). Against Method. New Left Books
- Flammini, A. (2015). Defining and Identifying Sleeping Beauties in Science. PNAS, 112(24), 7426–7431
- Kuhn, T. S. (1962). The Structure of Scientific Revolutions. University of Chicago Press
- Barton, D. (1992). Core Capabilities and Core Rigidities. Strategic Management Journal, 13(S1), 111–125
- March, J. G. (1993). The Myopia of Learning. Strategic Management Journal, 14(S2), 95–112
- March, J. G. (1991). Exploration and Exploitation in Organizational Learning. Organization Science, 2(1), 71–87
- Merton, R. K. (1968). The Matthew Effect in Science. Science, 159(3810), 56–63
- Phelan, P. (1993). Unmarked: The Politics of Performance. Routledge

- Salaudeen, O. et al. (2025). 关于基准构建的原则性框架与构念效度。
- Simon, H. A. (1971). Designing Organizations for an Information-Rich World. In M. Greenberger (ed.), Computers, Communications, and the Public Interest. Johns Hopkins University Press. 技术就绪度等级 (Technology Readiness Levels) : NASA 与欧盟研究框架计划所采用的九级量表。 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。 全部来源均为站外公开文献, 可自行核对。 + Claude · 作者: 王德生

■ 第七章 · 一件是几件

- Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In The Price Statistics of the Federal Government. NBER
- Advisory Commission to Study the Consumer Price Index (1996). Toward a More Accurate Measure of the Cost of Living (the Boskin Report). U.S. Senate Finance Committee
- Hausman, J. (2003). Sources of Bias and Solutions to Bias in the Consumer Price Index. Journal of Economic Perspectives, 17(1), 23–44
- Diewert, W. E. (1976). Exact and Superlative Index Numbers. Journal of Econometrics, 4(2), 115–145
- Irvin, S. (2005). The Artist's Sanction in Contemporary Art. Journal of Aesthetics and Art Criticism, 63(4), 315–326
- Irvin, S. (2022). Immaterial: Rules in Contemporary Art. Oxford University Press
- Rohrbach, G. (2003). Artworks as Historical Individuals. European Journal of Philosophy, 11(2), 177–205
- Goodman, N. (1968). Languages of Art. Bobbs-Merrill
- Reinertsen, D. G. (2009). The Principles of Product Development Flow: Second Generation Lean Product Development. Celeritas
- Goldratt, E. M. & Cox, J. (1984). The Goal. North River Press
- Baldwin, C. Y. & Clark, K. B. (2000). Design Rules, Volume 1: The Power of Modularity. MIT Press
- Coase, R. H. (1937). The Nature of the Firm. Economica, 4(16), 386–405
- Williamson, O. E. (1985). The Economic Institutions of Capitalism. Free Press
- Geertz, C. (1963). Agricultural Involution. University of California Press

- Campbell, D. T. (1979). Assessing the Impact of Planned Social Change. *Evaluation and Program Planning*, 2(1), 67–90
- Levinthal, D. A. & March, J. G. (1993). The Myopia of Learning. *Strategic Management Journal*, 14(S2), 95–112
- Leonard-Barton, D. (1992). Core Capabilities and Core Rigidities. *Strategic Management Journal*, 13(S1), 111–125
- Broad, W. J. (1981). The Publishing Game: Getting More for Less. *Science*, 211(4487), 1137–1139. 一件是几件
- Šupak Smolčić, V. (2013). Salami Publication: Definitions and Examples. *Biochemia Medica*, 23(3), 237–241
- Shaib, C., Barrow, J., Sun, J., Siu, A. F., Wallace, B. C. & Nenkova, A. (2024). Standardizing the Measurement of Text Diversity: A Tool and a Comparative Analysis of Scores. arXiv: 2403.00553
- Shaib, C., Elazar, Y., Li, J. J. & Wallace, B. C. (2024). Detection and Measurement of Syntactic Templates in Generated Text. arXiv:2407.00211
- Doshi, A. R. & Hauser, O. P. (2024). Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content. *Science Advances*, 10(28), eadn5290
- Martínez, C. (2011). Patent Families: When Do Different Definitions Really Matter? *Scientometrics*, 86(1), 39–63
- OECD (2009). *OECD Patent Statistics Manual*. OECD Publishing. 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 + Claude · 作者：王德生

■ 第八章·未分之别

- Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89(5), 841–864. <https://www.journals.uchicago.edu/doi/10.1086/261010>
- 2. Tournament-Based Performance Evaluation and Systematic Misallocation: Why Forced Ranking Systems Produce Random Outcomes. arXiv:2512.06583. <https://arxiv.org/html/2512.06583>
- 3. Hope, Signals, and Silicon: A Game-Theoretic

- Model of the Pre-Doctoral Academic Labor Market in the Age of AI. arXiv:2511.00068. <https://arxiv.org/pdf/2511.00068>
4. Artism: AI-Driven Dual-Engine System for Art Generation and Critique. arXiv:2512.15710. <https://arxiv.org/pdf/2512.15710>
5. Tracing Generative AI in Digital Art: A Longitudinal Study of Chinese Painters' Attitudes, Practices, and Identity Negotiation. arXiv:2511.03117. <https://arxiv.org/pdf/2511.03117>
- Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman & Hall; 及
- Ashby, W. R. (1958). Requisite Variety and its Implications for the Control of Complex Systems. *Cybernetica*, 1(2), 83–99. 7
- Naughton, J. Ross Ashby's Law of Requisite Variety. Edge. <https://www.edge.org/response-detail/27150>
8. Virtualizing RAN: Science, Strategy, and Architecture of Software-Defined Mobile Networks. arXiv:2506.09878. <https://arxiv.org/pdf/2506.09878>
- 9
- Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press. 10
- March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1), 71–87. 11
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press. 12
- Scott, J. C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press. 13
- Argyris, C., & Schön, D. A. (1978). *Organizational Learning: A Theory of Action Perspective*. Addison-Wesley. 14
- Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press. (量化与信道容量, 本文第十二节其六所对照的形式层邻居) 未分之别 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献, 可自行核对。 + Claude · 作者: 王德生

■ 第九章 · 那一段谁走的

- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. Generative AI Without Guardrails Can Harm Learning: Evidence from High School

Mathematics. Proceedings of the National Academy of Sciences, 122(26), e2422633122, 2025. (2025 年 8 月 更 正 见 同 刊 122(34), e2518204122) <https://doi.org/10.1073/pnas.2422633122>

· Brynjolfsson, E., Chandar, B., & Chen, R. Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence. Stanford Digital Economy Lab Working Paper, 那一段谁走的 August 2025 (November 2025 与 February 2026 两 次 更 新) . <https://digitaleconomy.stanford.edu/publication/canaries-in-the-coal-mine-six-facts-about-the-recent-employment-effects-of-artificialintelligence/>

· Biesta, G. Good Education in an Age of Measurement: Ethics, Politics, Democracy. Paradigm Publishers, 2010

· Biesta, G. World-Centred Education: A View for the Present. Routledge, 2022

· Generative AI and the Irreducible Role of Teachers as Callers into Subjectness: A Placebo–Nocebo– Treatment Framework. Studies in Philosophy and Education, 2026. <https://doi.org/10.1007/s11217-026-10051-w>

· Biesta's World-Centred Education: Subjectification Revisited? Journal of Philosophy of Education, 59(2), 274–289, 2025. <https://doi.org/10.1093/jopedu/qhae057>

· Bainbridge, L. Ironies of Automation. Automatica, 19(6), 775–779, 1983

· Bjork, R. A. Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe & A. Shimamura (eds.), Metacognition: Knowing about Knowing. MIT Press, 1994

· Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. The Expertise Reversal Effect. Educational Psychologist, 38(1), 23–31, 2003

· Braverman, H. Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century. Monthly Review Press, 1974

· Leonard-Barton, D. Core Capabilities and Core Rigidities. Strategic Management Journal, 13(S1), 111–125, 1992

· Levinthal, D. A., & March, J. G. The Myopia of Learning. Strategic Management Journal, 14(S2), 95–112, 1993

· Miller, D. The Icarus Paradox. HarperBusiness, 1990

- Geertz, C. *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press, 1963
- Collins, R. *The Credential Society: An Historical Sociology of Education and Stratification*. Academic Press, 1979
- Spence, M. Job Market Signaling. *Quarterly Journal of Economics*, 87(3), 355–374, 1973
- Hirsch, F. *Social Limits to Growth*. Harvard University Press, 1977
- Akerlof, G. A. The Market for "Lemons". *Quarterly Journal of Economics*, 84(3), 488–500, 1970
- Simon, H. A. Designing Organizations for an Information-Rich World. In M. Greenberger (ed.), *Computers, Communications, and the Public Interest*. Johns Hopkins Press, 1971
- Lave, J., & Wenger, E. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, 1991. 那一段谁走的
- Sculley, D., et al. Hidden Technical Debt in Machine Learning Systems. *NeurIPS*, 2015
- Potential Risks of Generative Artificial Intelligence Integration into K-12 Education: A Scoping Review. *Computers and Education: Artificial Intelligence*, 2026. <https://doi.org/10.1016/j.caeai.2026.100526>
- Iscenko, Z., & Millet, C. Looking for the Ladder: Is AI Impacting Entry-Level Jobs? *Economic Innovation Group*, January 2026. <https://eig.org/wp-content/uploads/2026/01/TAWP-IscenkoMillet.pdf>
- Humlum, A., & Vestergaard, E. The Unequal Adoption of ChatGPT Exacerbates Existing Inequalities among Workers. *PNAS*, 2025
- Fan, Y., Tang, L., Le, H., et al. Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance. *British Journal of Educational Technology*, 56(2), 489–530, 2025. 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 + Claude · 作者：王德生

■ 第十章 · 没有人在养它

- 桑福德·格罗斯曼、奥利弗·哈特：《所有权的成本与收益：纵向与横向一体化理论》，《政治经济学杂志》第 94 卷第 4 期，1986 年。<https://doi.org/10.1086/261404>
- 奥利弗·哈特、约翰·摩尔：《产权与企业的性质》，《政治经济学杂志》第 98 卷第 6 期，1990 年。<https://doi.org/10.1086/261729>
- 奥利弗·哈特：《企业、合约与财务结构》，牛津大学出版社，1995 年。
- 阿尔弗雷德·盖尔：《艺术与能动性：一种人类学理论》，牛津大学出版社，1998 年。
- 迈克尔·巴克森德尔：《十五世纪意大利的绘画与经验》，牛津大学出版社，1972 年。
- 彼得·瑙尔：《编程即理论建构》，《微处理与微程序设计》第 15 卷第 5 期，1985 年，第 253—261 页。[https://doi.org/10.1016/0165-6074\(85\)90032-8](https://doi.org/10.1016/0165-6074(85)90032-8)
- 梅厄·莱曼：《作为演化实体的程序、生命周期与软件演化定律》，《IEEE 会刊》第 68 卷第 9 期，1980 年。<https://doi.org/10.1109/PROC.1980.11805>
- 莉桑娜·班布里奇：《自动化的讽刺》，《自动化》第 19 卷第 6 期，1983 年，第 775—779 页。[https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
- 让·莱夫、艾蒂安·温格：《情境学习：正当的边缘性参与》，剑桥大学出版社，1991 年。
- 哈里·柯林斯：《默会知识与显性知识》，芝加哥大学出版社，2010 年。
- 卡里姆·比利莫里亚等：《外科培训中工时弹性的全国性整群随机试验》，《新英格兰医学杂志》第 374 卷第 8 期，2016 年，第 713—727 页。<https://doi.org/10.1056/NEJMoa1515724>
- 桑吉·德赛等：《内科工时弹性试验中的教育结果》，《新英格兰医学杂志》，2018 年。<https://doi.org/10.1056/NEJMoa1800965>
- 美国联邦航空局：《安全警示 13002：手动飞行运行》，2013 年 1 月 4 日。
https://www.faa.gov/sites/faa.gov/files/other_visit/aviation_industry/airline_operators/airline_safety/SAFO13002.pdf
- 美国联邦航空局：《安全警示 17007：手动飞行运行熟练度》，2017 年 5 月。
https://www.faa.gov/sites/faa.gov/files/other_visit/aviation_industry/airline_operators/airline_safety/SAFO17007.pdf
- 美国交通部监察长办公室：《加强联邦航空局的监督有助于确保飞行员掌握手动飞行技能》，2016 年 1 月。https://www.oig.dot.gov/sites/default/files/FAA%20Flight%20Deck%20Automation_Final%20Report%5E1-7-16.pdf
- 没有人在养它 多萝西·伦纳德-巴顿：《核心能力与核心刚性》，《战略管理杂志》第 13 卷增刊，1992 年。<https://doi.org/10.1002/smj.4250131009>

- 丹尼尔·列文塔尔、詹姆斯·马奇：《学习的近视》，《战略管理杂志》第 14 卷增刊，1993 年。<https://doi.org/10.1002/smj.4250141009>
- 埃莉诺·奥斯特罗姆：《公共事物的治理之道》，剑桥大学出版社，1990 年。人机分工声明选题、三个学科的指定与最终定稿由王德生负责。文献检索、冲突定位、框架构建、写作与自检由 Claude 执行。所有对外来源均给出可直接点开核对的地址；文中不报告任何虚构的经验数据，两处可核查的政策文件（美国 联邦航空局的两份安全警示与交通部监察长办公室的审计报告）已逐项核对。
- 本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 + Claude · 作者：王德生

附录一

十项读数速查表

编号	读数	定义	数据来源	正交实例
一	未阅率	上一轮全部否决里，作出否决时无人打开过材料的比例	投递系统的被动访问日志	答复时限压到四十八小时、积压清零、满意度最高的招聘流程，可以有接近一的未阅率
二	暗基线抬升度	把上一轮全部未被接受的产出移出参照集后，本轮合格线下降的幅度 ÷ 本轮合格线	决定日志 + 一次盲掉年份的重判	录用数最多、到岗时间最短的机构，可以有最高的抬升度
三	饱和关占比	全部关口中，最近一百次判定通过率不低于零点九五者所占比例	各关口的判定记录	一条从来没有红过的持续集成流水线，这个比率等于一
四	认领占比	一次交付的总工作量里，为使产出可被归属于某人而追加的部分所占比例	工时记录、提交历史、材料清单分项	准时率最高、缺陷率最低、覆盖率满分的团队，可以有最高的认领占比
五	辨认转嫁比	被评价方为使产出可被辨认所投入的时间 ÷ 评价方为核实所投入的时间	两侧的工时记录	引入自动核验而不撤旧要求时，它可以暴涨，而两侧账面效率同时改善
六	无量期占比	一样东西从被做出来到取得第一个读数所经过的时间，占其全部在制时间的比例	版本时间戳与首次评审时间戳之差	把首读时距压到最短的评审系统，可以同时是最不可能发现新东西的系统
七	著作层并项率	把本期计入件数按"是否服务于同	产出清单 + 一次按承诺结构的归并	人均件数翻倍的团队，并项率可以接

		一个可明确区分的承诺"归并到著作层后，件数缩减的比例		近一
八	有效种类留存率	本期有效种类数 ÷ 上期有效种类数，有效种类数按各类别被实际使用的频次分布计算	评分表、标签体系、分类字典的版本比对与使用频次	评分一致性最高、争议最少的评审体系，有效种类留存率可以最低
九	无验收期占比	一段学习或训练过程中，不产出任何可被验收之物的时间所占比例	课程与项目的时间表、交付物清单	交付物最密集、进度最可控的培养方案，这个比例最接近零
十	停养率	曾经顺带养成某项能力的那些活动，在本期被自动化或优化掉的比例	流程变更记录、工具上线清单	举不出正交实例——目前没有任何机构在测它。这本身就是第十章那条判断的一个读数。

使用三条：先算第一、第三、第六（数据最全、成本最低，且分处链上三个不同位置；三个都绿则本书对该机构不适用）；这些数指的是位置不是人，不可用于考核个体；这张表进考核的那一天它就开始失效，每一项的最省事达标路径都会摧毁它要保护的东西。

附录二

记号与关系式一览

记号

符号	名称	出处
N	一轮的投递总量	枢纽章
$\prod p$	各关通过率之积	第三章
K	下游一轮可过目的产能（份数）	枢纽章
ρ	拥堵参数, $\rho = N \cdot \prod p / K$	合章
u	未阅率	第一章
δ	暗基线抬升度	第二章
s	饱和关占比	第三章
a	认领占比	第四章
τ	辨认转嫁比	第五章
w	无量期占比（ $\bar{w}$ 为平均等待轮数）	第六章
m	著作层并项率	第七章
k	有效种类留存率	第八章
e	无验收期占比	第九章
σ	停养率	第十章
h	被评价方使一件产出可被辨认的平均追加工时	合章
c	评价方核实一件产出的平均工时	合章
J	一个系统中判断力的存量	合章

关系式

式	内容	出处	它说了什么
一	$u = \max(0, 1 - K/(N \cdot \prod p)) = 1 - 1/\rho$	枢纽章	未阅率由拥堵参数决定
二	$r = 1 - (\prod p) \cdot u$	枢纽章	结果确实取决于内容的概率；门槛侧与否决侧

三	$\bar{w} = \rho - 1 = u/(1 - u)$	合章	的失效在 r 上相乘 首读时距与未阅率是同 一参数的两种读法
四	$\tau = a \cdot (h/c)/(1 - u)$	合章	认领占比与辨认转嫁比 只差一个拥堵参数；未 阅率发散式放大它
五	$\rho_{\text{真}} = \rho \cdot (1 - m)$	合章	并项率使 u 、 $\bar{w}$ 、 τ 被 系统性高估；算出 m 之前不得跨期比较
六	$r_{\text{eff}} = [1 - (\Pi p) \cdot u] \cdot k$	合章	有效判定率；类别数下 降使判定不再取决于这 一件的内容
七	$J(t+1) = J(t) \cdot (1 - \sigma) + G(e)$	合章	σ 有无数个生产者而 G 一个也没有

两个算例（引自枢纽章）：四道关通过率分别为零点九九、零点九七、零点九五、零点九八， $\Pi p \approx$ 零点八九四； $N = 6000$ ， $K = 600$ ，得 $u \approx$ 零点八九， $r \approx$ 零点二一。撤掉最饱和的那道关， r 由零点二一降到零点二零；产能翻倍使 r 升到零点三一，而投递量随之涨到 12000 时， r 精确回到零点二一。

附录三

十章命名一览

章	篇名	被当作"判断已经作出"的证据	本章指认的那个量
一	两种「不」	有一次否决记录在案	未阅否决
二	被丢掉的那一眼	判定依据写在细则里	暗基线
三	这道关拦下过谁	关口还在、流程照走	饱和关
四	认领物	产出交付了、归属清楚	认领物（分签名式与代价式）
五	认不出来，就得自己交代	评价方完成了核实	同形劳动
六	第一个读数之前	周期缩短、效率提高	首读时距
七	一件是几件	件数上升即投入上升	并项（著作层）
八	未分之别	评分一致、争议减少	有效种类数
九	那一段谁走的	培养过程有交付物、进度可控	无验收期
十	没有人在养它	能力一直都在，没听说谁不行了	无供给品

十章原刊于「学科通融」专栏之四十六、之四十七、之四十五、之四十四、之四十八、之五十、之四十九、之五十二、之五十一、之五十三，作者王德生 + Claude。

后记

一

这本书不是先有想法再去写的。

十篇研究是在二〇二六年七月到八月之间陆续做出来的，每一篇当时都只想回答自己那一个问题：为什么投出去的东西没有回音，为什么关口拦不下人却撤不掉，为什么件数涨了而处境不动。做到第五、第六篇的时候才发现，它们在同一个位置上反复撞到同一样东西——每一次都撞在"这件事没有栏位"上。

这个发现不是灵感，是重复。十次里有六次撞在同一处，那就不是巧合，而是那个位置上确实有一样东西。母题是这样出来的，不是想出来的。

二

编成一本书的过程中，有三次是被迫改主意的。

第一次是发现编一自相矛盾。第三章说通过率趋近于一，第一章说绝大多数被否决。两篇发表时各自都说得通，放在一起就不成立。当时最省事的办法是给两篇各加一句限定，让它们互不冒犯。没有这样做，是因为那样处理之后，读者会得到两条互相回避的结论，而回避本身不产生任何知识。枢纽章是这次拒绝省事的产物，而它撞出来的东西比两篇各自都硬：撤掉最饱和的那道关，只降代价、不改善回报；扩产能会被投递量吃掉。

第二次是发现十项读数无法互相否认。十篇交出十个量，看上去很丰富，实际很脆弱——任何一项被质疑，都可以退回到"这一项的定义有问题"。合章是这次发现的产物。写它的时候有一个意外的收获：五条关系式全部是从定义直接推出来的，不需要任何新数据；而这意味着，如果检验不通过，被推翻的不是某条经验假设，而是某个读数的定义本身。这比推翻一条经验命题更彻底。

第三次是发现九章共用一组三学科。这是书一级上最明显的短板：书里的纠缠是同一次碰撞重复了十遍。处理它有两条路，一条是重写三章、重跑碰撞，那意味着重做三门学科的源研究并丢弃已发表的文本；另一条是给三章各附一节"换一组学科再问一遍"。选了后者，并给自己定了一条判准：换组必须改动原章的处方或读数，否则它就是装饰。三次都改了——第四章的处方被改掉一半，第七章的读数被换掉，第八章的一处主张被削弱。这条判准比换组本身更值得留下来。

三

写完之后，有一件事一直没有解决，这里必须写出来。

本书的全部主张，都依赖"存在一类不被记账的东西"。而本书自己就是一件被交出去的东西。如果它没有被接受，按第二章的机制，它仍然会抬高一点点线——而这一点在原理上无法被证实，因为没有任何人记录未被接受的稿子。

这不是一个可以靠更严谨的论证绕开的困难。它是本书主张的直接后果：一个主张"某类东西不被记账"的文本，一旦自己成为那一类，同样不会被记账。能做的只有把它写在明处，并把读数交出去——读数可以被别人算，而算出来的结果不需要经过作者同意。

这也是为什么本书的兑现物是十项读数而不是十条主张。主张需要被相信，读数只需要被计算。

四

最后一句留给那些会用这张表的人。

体检表进考核的那一天，它就开始失效。每一项都有一条最省事的达标路径，而每一条都会摧毁它要保护的那样东西。这一条没有出路，本书写了三遍——导论一遍，结语一遍，这里第三遍。

写三遍不是啰嗦。如果这本书最终只留下一样东西，我希望留下的是这一句：一个机构如果只是需要一份好看的报表，它有一百种办法把这十个数做绿，而本书描述的那件事会照常发生，只是从此连间接读数也没有了。

二〇二六年八月

封底

判断的危机

它不会发出任何警报

做出一件东西的成本降到了接近于零。

而认出它算不算数的成本，一分钱没有降。

十项读数，九项全绿。

第十项没有颜色——因为没有任何机构在测它。

本书由「学科通融」专栏十篇连续研究编成，另新写前言、导读、导论、枢纽章、合章、结语与附录。十章分别指认判定这一端上一个从未被单独生产的量：没有人打开过的那种否决、由弃物合成的标准、拦不下人却撤不掉的关口、为证明「是你做的」而追加的那一层、被整体转嫁的辨认负担、被计为损耗的首读时距、失去下界的计数单位、被当作背景的分辨力、不产出可验收之物的那段时间，以及一类没有供给方的能力。

合章再用五条从定义推出的关系把十项锁成一张可互相否证的网。其中最干净的一条说：首读时距与未阅率不是两件事，是同一个拥堵参数的两种读法。

全书的兑现物是结语给出的十项体检读数——每一项都能从既有日志算出，且与该行业现行主要指标正交：主要指标全绿时，它可以很难看。

王德生 + Claude 著

德麦国际出版社 · SDE UNIVERSES

德麦国际专著第 80 号

THE CRISIS OF JUDGMENT