

抄得越像， 坏得越快

决定一次模仿是养出本事还是掏空信誉的，不是它抄了多少，是抄完之后谁挨打

作者 王德生 + Claude · 约两万字 · 22 页 · 三种读法 · 发表于 2026 年 7 月 28 日

提 要

照着别人做，三个领域给了三个撞车的答案：演化生物学说模仿是寄生，不付代价的模仿者一多，信号对所有人失效——包括真货；文化演化说连不理解的部分也高保真照抄，正是人类累积文明的引擎，越忠实越好；而洗稿分析说，前两家都在谈"抄得多像"，可相似度从一开始就不是那个变量。前两家在同一个量上给出相反符号，第三家说这个量无效。本文指出三家共同隐去的变量：**模仿者会不会因为这件事做错，而承担与被模仿者同样性质的后果**——本文称之为**跟着挨打**。挨打的模仿加固信号（抄得越全越好，多余的会被自己的失败磨掉）；不挨打的模仿稀释信号（**抄得越像，死得越快**）。文章给出三分判据、一个可估的崩溃点 $p^*=1-c/L$ 、"两头夹击"的处境诊断，论证加强查验在算术上必输、崩溃后受损最重的是货真价实者、以及为什么"做得更好"救不了自己。附与昂贵信号理论等六个近邻的划界、五条做法、六个位置、八条证伪与三条写死到二〇三〇年的赌注。

关键词 跟着挨打 · 贝氏拟态与缪氏拟态 · 保真度的符号翻转 · 崩溃点 $p^*=1-c/L$ · 两头夹击 · 后果链而非成本

本 篇 由 哪 三 方 撞 成

按栏目规矩挑选：分属不同领域、且在同一个动作上正面打架。这一次的冲突是本栏最硬的一种形态——**前双方在同一个变量（模仿得多像）上给出相反符号，第三方说这个变量根本不该用**。三方各有硬证据：拟态有野外实验数据，过度模仿有跨物种跨年跨文化的重复验证，洗稿有每天都在发生的现场。所以问题不是谁错了，而是它们共同隐去了什么。

- 频率依赖的贝氏拟态 · 演化生物学 · Finkbeiner 等, Proc. R. Soc. B 285:20172786, 2018; 另见 Pfennig 珊瑚蛇系列
哥斯达黎加模仿者与被模仿者数量相当时拟态崩溃、捕食者更易攻击；厄瓜多尔被模仿者多出八十四倍时两者同受保护。**信号能否成立，取决于不付代价的模仿者占多大比例。**
- 过度模仿与高保真复制 · 文化演化与发展心理学 · Horner & Whiten 2005; McGuigan 等, Br. J. Psychol. 2011; PLOS ONE 2016; Child Dev. Perspect. 2025
人类儿童（黑猩猩不会）会把因果上无关的动作也照抄，且该倾向随年龄增强、成人更甚。**只抄目标不抄动作的"聪明"做法反而不支持文化累积。**
- 洗稿的本体论 · 知识生产 · 王德生，站内《论文抄袭专栏》之十
洗稿系统地扔掉字句、保留骨架；而字句正是查重唯一测的东西。**查重对洗稿的失效不是灵敏度不够，是构造性的——你可以原样搬走整条路径，同时改掉每一个字。**

本栏此前多篇处理的是**个体或一对关系**：之十二问委托之后你还验不验得了，之十三问支持有没有留出力位，之十六问一件成品身上留不留得下判断。本篇第一次把单位换成**群体**——同一个动作，单个来看无害，占到某个比例就使一整个信号失效。所以它与之十六尤其需要分开：那一篇说弃径在物上不留痕（载体问题，个体层面）；本篇说凭证的语义在被一群人消耗（比例问题，群体层面）。两篇可对读——之十六给的是“怎么看出这一件有没有判断”，本篇给的是“这一类凭证还能撑多久”。

序·三个不肯让步的答案

照着别人做——这件事有一个很朴素的名字：模仿。它可能是人类做得最多的一件事。

而关于它，三个领域给出了三个互不相容的判断。不是侧重不同，是**正面撞车**：其中两家在同一个变量上给出了相反的符号，第三家说这个变量根本不该用。

第一家：模仿是寄生，而且有一个崩溃点

来自演化生物学。有一类无毒的蝴蝶，长得跟有毒的那种几乎一样。捕食的鸟吃过一次有毒的，就学会避开这个花纹；于是无毒的那种搭了顺风车——它不用付出有毒的代价，却拿到了有毒的保护。这叫贝氏拟态。

关键的部分在后面。二〇一八年一项野外实验给出了硬数据：研究者在两个地方放了人造蝴蝶模型。在哥斯达黎加，有毒的那种和模仿它的那种数量差不多——结果拟态**崩溃了**，鸟更容易去攻击那个花纹。而在厄瓜多尔，有毒的那种比模仿者多出八十四倍——结果两边都受到同样的保护。

这个结果说的是一件很硬的事：**一个信号能不能成立，取决于不付代价的模仿者占多大比例**。超过某个比例，那个花纹就不再意味着“有毒”，于是它对所有人都失效——包括那些真的有毒的。

结论：**模仿是寄生。模仿者一多，信号对所有人失效，而且受害最重的是原来那个。**

第二家：连没用的一起抄，才有文明

来自文化演化与发展心理学。

有一个做了二十年的实验：给一个透明的盒子，里面有奖励。示范者先用棍子敲盒子顶上几下——这几下**完全没用**，透明的盒子让你一眼就能看出它跟打开毫无关系——然后再从侧面把奖励取出来。

黑猩猩看完之后，跳过没用的那几下，直接从侧面取。

而人类的孩子会把没用的那几下**一起照抄**。这个现象叫过度模仿。更让人意外的是后来的发现：这个倾向**随年龄增长而增强**——成人比孩子抄得更彻底、效率更低。而且它在没有社会压力、当事人不知道自己在被观察的自然情境里同样出现。还有研究发现，过度模仿的孩子会把那些不必要的动作当成**规矩**，并且去要求第三个人也照做。

这在文化演化里被当作一个正面的东西：只抄目标、不抄动作的那种学习（黑猩猩那种），被认为**不支持累积**；而高保真的、连不理解的部分也一起复制的模仿，被认为正是人类文化能一代代往上累的前提。

结论：**不理解也照抄、连没用的一起抄，是文明累积的引擎。越忠实越好。**

第三家：像不像，根本不是那个变量

来自知识生产。

有一种抄袭叫洗稿：把一篇文章的论证骨架整体搬走——问题从哪切入、先立什么后驳什么、用哪几步把读者从疑问带到结论——然后换掉全部术语、全部案例，甚至换掉领域，把每一个句子打散重排，直到字句重合度降到检测线以下。

送进查重系统，干干净净。送到任何一位没读过原文的评审面前，逻辑自洽，挑不出证据。可任何一个读过原文的人，一眼就看出：这篇文章的骨头是从那里搬来的。

关键的判断是：查重对洗稿的失效**不是灵敏度不够，是构造性的**。查重测的是字句的重合，而剽窃发生在路径那一层——**你可以原样搬走整条路径，同时改掉每一个字。**

结论：**前两家都在谈"抄了多少、抄得多像"，而相似度从一开始就不是那个变量。**

三个答案，一个动作

领域	对"照着做"的判断	"像"这个变量
演化生物学	寄生。模仿者过多则信号崩溃	越像越糟
文化演化	引擎。人类靠它累积文明	越像越好
知识生产	问错了。被拿走的是路径不是字句	这个变量无效

请看第三列。前两家在同一个量上给出了**相反的符号**，而第三家说这个量根本测不到要害。

三家各自都有硬证据：拟态实验有野外数据，过度模仿有跨物种、跨年龄、跨文化的重复验证，洗稿那一条有每天都在发生的现场。所以这不是谁错了的问题。

这篇文章要做的，是找出那个它们三家都没有测、而一旦放进来三家就能同时成立的变量。

为什么这三家会在这几年同时被翻出来

还要交代一句，否则前面那张对照表像是三家不够聪明。

三家各自的结论在它们自己的场景里都成立了几十年，没有出过问题。它们之所以现在互相打架，是因为一个**此前从不变的量开始变了**。

在拟态那里，模仿者的比例由生态决定，变化以世代计，慢得可以当常数。在文化演化那里，学习者随后要真的用它做事——这在人类历史的绝大部分时间里没有例外，所以它是背景不是变量。在洗稿那里，改写一篇文章的骨架仍然需要相当的功夫，所以稀释是缓慢的。

三家各自把一个当时确实不动的东西当成了背景。而这几年，这个背景动了，而且动得很快。**于是三家同时开始不好用，而且是以互相矛盾的方式不好用。**

这也是本文之所以可能有一点用处的全部理由：它不比三家中的任何一家聪明，它只是站在一个三家都开始失灵的时刻，因而有机会看见那个被它们共同当作背景的东西。**这种机会不是本事，是运气——它属于时间，不属于任何人。**

第一章·先把"模仿"拆开

三家打得起来，一部分原因是"模仿"这个词把几件不同的事装在了一起。先拆开，后面才好说。

三个可以分开的问题

抄了什么。是表面的样子（花纹、字句、笔法），还是内部的做法（步骤、骨架、路径）。洗稿那一家的气，就花在指出这两者可以完全分离。

抄得多像。保真度。过度模仿研究测的是这个；拟态实验里的"不完美拟态"测的也是这个。

抄完之后干什么。——这一栏，三家都没有系统地看。

而本文要说的是：**决定一次模仿的性质的，是第三栏。**前两栏都只是它的表现。

先把结论摆出来

为了让后面几章有个靶子，先把话说在前面：

一次模仿是把一个信号加固还是稀释，不取决于它抄了什么，也不取决于它抄得多像，而取决于——模仿者接下来会不会因为这件事做错，而承担和被模仿者同样的后果。

会承担的，本文叫**跟着挨打**的模仿。不承担的，叫**不挨打**的模仿。

这两种在外面看起来完全一样：都是照着做，都可以做得很像，都可能出于最好的动机。而它们对那个被模仿的信号，作用方向相反。

第二章·蝴蝶那一侧：一个可以算的崩溃点

为什么这个例子这么有用

无毒蝴蝶模仿有毒蝴蝶，这件事之所以值得反复看，是因为它把一件在人类社会里含混不清的事，做成了一个**数量关系**。

那个花纹是一个信号，它的含义是"吃了我会难受"。这个含义不是天然的，是鸟一次次吃到苦头学来的。也就是说：**这个信号的可靠性，是由被吃过的那些有毒蝴蝶用命换来的。**

无毒的模仿者做了什么？它拿到了这个含义带来的保护，而它**永远不会为这个含义添一分钱**——它不可能变得有毒，它的整个生存策略就是不付那份成本。

于是这里有一个非常干净的结构：**信号的可靠性是一个被反复充值的账户，而模仿者只提不存。**

崩溃点是真的存在的

更要紧的是那个数。

如果模仿者很少，鸟遇到这个花纹时绝大多数时候确实会中招，于是它继续避开——信号维持，模仿者搭到了车，被模仿者也没损失多少。

如果模仿者变多，鸟遇到这个花纹时越来越多次没事，于是它开始试探，再往后就不再避开了——**信号失效，模仿者失去保护，而被模仿者也失去了保护。**

那项野外实验测到的正是这两种情况：模仿者与被模仿者数量相当的地方，拟态崩溃，捕食更多；被模仿者多出八十四倍的地方，两边都安全。

所以这不是一个"有没有坏蝴蝶"的问题，是一个比例问题。单个模仿者不构成任何威胁；比例过线，整个信号一起完蛋。

这一点后面第八章要用到，因为它直接决定了一整类治理办法为什么注定失败。

还有一个方向相反的例子，非常关键

生物学里还有另一种拟态：几种**都有毒**的蝴蝶，长得越来越像。这叫**缪氏拟态**。

它的结构完全不同：每一种都真的有毒，所以每一种都在为这个信号**充值**。它们长得越像，鸟需要学习的次数就越少，每一种付出的伤亡也越少。**这里，越像越好——对所有参与者都好。**

把两种拟态并排放，这一整篇文章的骨架就出来了：

	贝氏拟态	缪氏拟态
模仿者本身	无毒，不承担后果	有毒，承担同样后果
越像的结果	信号越快崩溃	信号越快稳固
对被模仿者	有害	有利

同一个动作，同一个保真度，两个相反的方向；而区分它们的，恰恰是模仿者会不会跟着挨那一下。

生物学早就把这两类分开了，而且分得很清楚。但它把这个区分留在了"有毒/无毒"这个具体的生物学事实上，没有把它抽出来当作一个一般的判据。这一步，是本文要做的。

第三章·孩子那一侧：连没用的一起抄，为什么反而对

一个看起来很蠢的行为

回到那个透明盒子。孩子明明看得见敲顶部那几下没用，他还是照抄。成人抄得更彻底。黑猩猩反而"聪明"——它跳过没用的部分，直奔目标。

按常识，孩子这是笨。按文化演化的说法，这恰恰是人类最要紧的一项本事。

理由是这样的：一个学习者面对一套别人的做法，通常**看不出**哪些步骤是必要的。透明盒子是实验室里的简化，现实里的绝大多数做法都不透明——为什么鞣制皮革要加这一步、为什么这种植物要泡三次水才能吃、为什么这个焊接要停两秒。这些步骤里，有的确实没用，有的**看起来没用而实际上性命攸关**。

如果每一代人都像黑猩猩那样只抄自己看得懂的部分，那么每一代都会把上一代辛苦攒下的、暂时讲不出道理的那部分丢掉。**文化就永远累积不起来，只能在每一代人能自己想明白的深度上反复。**

而人类的做法是：先全抄，包括看不懂的。**先保住，再慢慢弄明白。**

但这里有一个所有人都没细说的前提

过度模仿之所以是好事，有一个前提，而这个前提在实验室里被隐去了，因为它太理所当然。

那个前提是：**照抄的那个孩子，接下来要用这套做法去真的把事情做成。**

他要真的打开那个盒子，真的吃到那个东西。如果他多抄的那几下让他慢了、累了、错过了什么，他会承担；如果他少抄的那一步是关键，他会失败。**他把自己整个放进了这套做法的后果里。**

正因为如此，他多抄的那些没用的动作是**安全的**——它们的成本很低（多敲几下而已），而漏抄的风险很高（可能漏掉性命攸关的一步）。在这样的收益结构下，“宁可多抄”是一个非常正确的策略。

而且更要紧的是：**这些多余的动作会被后面的使用慢慢淘汰。**一代代人用下去，成本高又确实没用的部分会被磨掉，看起来没用而实际有用的部分会被留住——因为留不住的那些人，做不成事。

过度模仿之所以是引擎，不是因为高保真本身有魔力，是因为它后面接着一会筛掉多余、留下必要的后果链。

把这个前提抽掉试试

现在设想：那个孩子照抄了全套动作，但他**永远不需要真的打开那个盒子**。他抄这套动作，是为了在别人面前表演“我会开这个盒子”。

这时候会发生什么？

多余的动作不会被淘汰——因为没有任何使用来淘汰它们。必要的动作也不会被格外看重——因为反正不用真开。**整套动作会原封不动地传下去，而且传得非常忠实，因为忠实正是表演所要求的。**

而“会做这套动作”这个信号，会慢慢不再对应“能打开盒子”这件事。

这不是一个假想。它是仪式、是形式主义、是“合规动作”的生成机制，也是本文接下来要谈的一切的原型。

第四章·洗稿那一侧：为什么“像”这个变量测不到要害

它做对的那一步

洗稿这件事之所以在理论上重要，是因为它是一个极其干净的实验：**它系统地把字句相似度抹到接近于零，而“抄袭发生了”这个判断丝毫不动摇。**

这就等于做了一次对照实验，把“相似度”这个变量单独拿掉，看结论变不变。结论没变。所以相似度不是那个变量。

而被拿走的是什么呢？是**论证的骨架**——问题从哪切入、先立什么后驳什么、用哪几步把读者从疑问带到结论。这条路径是原作者一步步走出来的，而它**不留任何字句指纹**：同一条路可以用一千种措辞说出来。

这一条如何顶住前两家

把这个判断拿去对前两家，会发现它同时切中了两边。

对生物学那一家：拟态实验测的是花纹的相似度。可如果“相似度”不是那个变量，那么整套拟态研究是不是也测错了？

这里要为生物学说一句公道话：在拟态那个场景里，**花纹就是全部**——鸟能看到的只有花纹，没有别的层。所以在那个场景里，测相似度是对的。洗稿那一条真正指出的是：**当信号有“表面”和“内部”两层时，测表面就会失灵**——而人类社会的绝大多数信号都是两层的。

对文化演化那一家：过度模仿研究测的是“抄了多少个动作”。可按洗稿的判断，抄动作和抄“为什么这样做”完全是两件事，而实验只测了前者。

这一击相当有力：**那个把没用的动作也照抄的孩子，他抄到的是动作，不是判断。**他不知道哪几下有用。他之所以最后能长出判断，靠的不是模仿本身，是后面那一长串“自己去开盒子”的过程。

第三家自己的盲区

洗稿那一条也有它没走到的地方，这里要说清楚，否则三家就不是平等地被对待。

它指出了被拿走的是路径，并且主张治理应该从“比对字句”转向“检验路径是否被独立地生成”。可它停在了一个个案例的层面：这一篇是不是抄的。

它没有问一个更靠上的问题：**当这样的事情足够多时，会发生什么？**——而这正是拟态实验回答的那一问。一个洗稿者不构成什么；当洗稿者占到某个比例时，“发表过论文”这个信号本身会不会失效？

三家各自都差一步，而三家差的不是同一步。合在一起，那一步才补得上。

把三家的证据强度分开说

最后为三家各自的证据强度划一条线，免得本文借它们的可信度借得太顺手——按本文自己的判断，这正是一次不挨打的模仿最容易做的事。

拟态那一条最硬。它有野外实验、有两个地点的对照、有可重复的方法，而且结论方向清楚。它的限度是：它讲的是花纹这种**单层信号**，人类社会的信号大多是两层的，所以本文借它的是**比例一崩溃这个结构**，不是它的具体数值。八十四倍这个数字属于那两种蝴蝶，不属于任何别的地方。

过度模仿那一条证据很多，但解释仍有争议。现象本身被反复重复出来了——跨年龄、跨文化、跨情境；而“它是累积文化的前提”这个解释是一种被广泛接受的说法，不是被证明的结论。也有研究者认为它主要是社会性动机（要合群、要被认可）而不是学习策略。本文借的是**现象**与那个被隐去的前提，不依赖于哪一派解释成立。

洗稿那一条是分析，不是实证。它没有数据，它做的是概念上的辨析——指出查重的失效是构造性的。这一类论证的强度取决于它的推理，而不取决于样本量。本文借它的是那个辨析本身。

把三种强度摆出来，是为了让读者知道本文哪一部分是站在硬地上、哪一部分是站在推理上。**第九章那个“必输”的判断**，主要依赖拟态那条硬证据加上算术；而**第十二章那个“全局跳变”**，主要依赖推理，硬证据要等几年。

第五章·撞出来的那一条：谁跟着挨打

把三家的缺口对齐

把三家没走到的地方并排放：

- **生物学**有了阈值和方向，但它把“有毒/无毒”当作一个具体的生物学事实，没有把它抽成一般判据。
- **文化演化**有了保真度的价值，但它隐去了那个前提——学习者随后要真的用它做事。
- **洗稿**有了层级的辨析，但它停在个案，没有走到群体。

三个缺口拼在一起，恰好指向同一样东西：**模仿者与被模仿者，是不是站在同一条后果链上。**

正式说法

跟着挨打：一次模仿之后，模仿者会不会因为这件事做错，而承担与被模仿者同样性质的后果。

"同样性质"这四个字要说清楚，它不是指同样大小。一个学徒做砸了不会像师傅那样赔上招牌，但他会做砸——他面对的是同一件事的失败，而不是另一件事的失败。

反过来，一个洗稿者也会承担后果：被抓到会身败名裂。但那是**另一件事**的后果——是"被发现抄袭"的后果，不是"这条论证路径走不通"的后果。他永远不会因为那条路径本身错了而受损，因为他不打算真的顺着它走下去。

这个区分是本文的全部要害，值得再说一遍：

跟着挨打，指的是被同一件事的失败所惩罚，不是被"抄袭"这件事的暴露所惩罚。前者是内在的，后者是外加的。

三家为什么同时成立

放进这个变量，三家的结论一次对齐：

无毒的模仿者不挨打——它不可能变得有毒，永远不会因为"这个花纹不对应毒"而受损（在信号崩溃之前）。所以它稀释信号，且越像稀释越快。生物学是对的。

那个开盒子的孩子挨打——他要真的打开盒子。所以他的高保真是安全的，多余的会被后续磨掉。文化演化是对的。

洗稿者不挨打——他的产出不会被那条路径原本要面对的检验：没有人会顺着他的文章去复现、去证伪、去拿它盖桥。所以查重（测相似度）测不到要害，而要害在别处。洗稿那一条也是对的。

三家没有一家需要退让。它们只是各自看到了同一个变量的一个侧面。

为什么这一条三家单独都到不了

生物学到不了，因为在它的场景里"挨不挨打"完全等同于"有没有毒"——一个生化事实。它没有理由把它抽象成一个跨领域的关系。

文化演化到不了，因为它的实验设计里，学习者总是要去用那套做法（不然实验没法计分）。那个前提被固定住了，所以看不见它是一个变量。

洗稿到不了，因为它处理的是单个文本的归属问题，而"跟着挨打"是一个关于**后续**的性质——它要看这个模仿者接下来去哪儿。一篇文章躺在那里，看不出这件事。

而三家合起来就逼出来了：**有阈值（生物）、有前提（文化）、有层级（洗稿）**——三样一凑，那个被共同隐去的变量就浮出来了。

第六章·是三种模仿，不是两种

上一章把模仿分成了挨打的和不挨打的。这个二分够用来看清方向，但不够精确——因为它会把一大批完全无害的模仿错划进坏的那一格。这一章把它切准。

第三个条件：有没有在声称

一个人跟着菜谱做菜，他不挨任何打——做砸了顶多今晚难吃。按上一章的二分，他属于"不挨打的模仿者"，应该在稀释"会做这道菜"这个信号。

可他显然没有。为什么？

因为他**没有在向任何人声称什么**。他没有拿这道菜去换机会、去应聘厨师、去当作能力的凭证。他只是做了顿饭。

于是加一个条件。三个条件合起来，才是完整的判据：

类型	照着做	拿它声称	跟着挨打	对信号的作用
纯学习	是	否	—	无影响
学徒	是	是	是	加固
稀释者	是	是	否	稀释

只有第三行是问题。而它需要**两个条件同时成立**：既拿它去声称，又不承担后果。缺任何一个，都不构成稀释。

这个三分为什么要紧

第一，它挡住了一个很坏的用法。没有这个三分，本文很容易被用来指责一切学习、一切参考、一切"站在别人肩上"。而按三分，绝大多数模仿落在第一行，与本文无关。

第二，它指出了最省力的干预点。要把一个稀释者变成学徒，有两条路：让他挨打，或者**让他不要声称**。后者常常便宜得多。

一个学生用工具写完作业，如果他交上去当作自己的能力证明，他在稀释；如果他明确标注"这一段是工具写的，我做的是这三处判断"，他就退回了第一行——**而他实际做的事一模一样**。

这解释了披露为什么有用，也解释了披露**为什么不是道德问题**：它不是诚实与否，它是把一个动作从第三行移回第一行的机械操作。

第四种，以及为什么它不成立

细心的读者会问：那"挨打但不声称"呢？

这一类存在，而且相当重要：一个人默默照着做、默默承担后果、从不拿出去说。他不影响任何信号，但他在**为自己积累**那样东西。

这一类之所以不单列，是因为它对信号的作用为零，与第一行相同。但它值得被点出来，因为它是唯一一种**先长本事、再去声称**的路径——而它在今天正在变得罕见，原因很简单：**声称的成本降到了几乎为零**，于是没有人有理由先等一等。

一个立刻可用的自查

把这一章收成一句：

你拿它去说事了吗？说错了你担吗？两个都是"是"，你在加固；第一个是、第二个否，你在稀释；第一个否，你什么也没影响。

第七章·为什么"像"的符号会翻转

同一条公式的两端

本文最反直觉的一条推论，是关于保真度的。

常识认为，模仿得像不像，本身有一个固定的道德色彩：抄得越像越可耻，或者学得越像越用功。两种说法都很流行，而且互相矛盾。

按本文，两种说法都对，只是各自成立于不同的一侧：

模仿者挨打时，保真度是优点：他抄得越全，越不容易漏掉那些看起来没用而实际关键的部分；而多余的部分会被他自己的后果慢慢磨掉。

模仿者不挨打时，保真度是加速器：他抄得越像，接收者越难把他和被模仿者区分开，那个信号也就失效得越快。

换句话说，"像"这个量本身没有符号，符号是后果链给的。

一个立刻的推论：好的假货比坏的假货更危险

这一条值得单独拎出来，因为它与所有人的直觉相反。

我们本能地认为，一个粗劣的仿冒品危害小，一个精致的仿冒品危害大——这个直觉是对的，但通常给的理由是错的。常见的理由是"精致的更容易骗到人，卖得更贵"。

按本文，真正的理由要深一层：**粗劣的仿冒会被接收者学会区分，于是信号得以保全；精致的仿冒无法被区分，于是信号本身被消耗。**粗劣的仿冒者伤害的是买他的那几个人，精致的仿冒者伤害的是这个信号的所有使用者，包括所有真货。

拟态实验里那个"不完美拟态"的细节正好对上：**不完美的模仿者在被模仿者稀少时会被优先吃掉**——鸟能区分，于是信号还在工作。

另一个推论：越是内行越难识别

这一条也反直觉，而且很实用。

我们默认内行更容易识破。在"抄了什么"这个层面上，这是对的——读过原文的人一眼就看出洗稿。

可在"会不会挨打"这个层面上，内行没有任何优势，因为**这件事不在作品上**。它是一个关于将来的性质：这个人接下来会不会被同一件事的失败所惩罚。你把作品翻烂了也看不出来。

所以那句常听到的"这个我一看就知道真假"，在旧时代成立，是因为旧代表表面与内里高度绑定；一旦模仿的保真度上去，内行的优势会先消失，而**他自己往往最后一个知道**。

第八章·这不是昂贵信号理论

一个新说法最要紧的一关，是它顶不顶得住"这不就是那个已有的概念吗"。这一章逐个划界。如果任何一个既有概念能一比一替换掉"跟着挨打"，本文就白写了。

与昂贵信号理论

这是最近的邻居，必须先处理。

经济学与生物学都有一套成熟的说法：一个信号之所以可信，是因为发出它的成本很高，而且这个成本对不同的人不一样——高质量的人发这个信号便宜，低质量的人发这个信号贵，于是低质量的人不发，信号就分离出了真假。文凭之所以有用，按这套说法，是因为读下来对聪明人比对不聪明的人容易。

本文与它的差别，可以用两个反例说清楚。

反例一：成本极低却可靠的信号。一位外科医生在手术记录上写下自己的判断——写字这件事的成本接近于零，任何人都写得出。可这个记录相当可靠。为什么？因为他**接下来要面对这个病人**。写错了，出问题的是他手里的这条命，追责的是他的执照。**成本不在写的那一刻，在后面。**

反例二：成本极高却不可靠的信号。一份耗资巨大、动用几十人几个月的评估报告，如果没有任何人会真的按它去做决定、也没有人会在事后核对它对不对——它的成本高得惊人，可靠性接近于零。**贵，但不挨打。**

两个反例合起来说明：

可靠性的来源不是成本，是后果。成本只是后果的一个特例——当成本恰好在发信号的那一刻结清时，两者重合；而在绝大多数场合，它们并不重合。

昂贵信号理论测的是**付出**，本文测的是**暴露**。付出可以是一次性的、可以被计算的、可以被外部第三方补贴掉；暴露是持续的、条件依赖的、无法被转移的。

与柠檬市场

二手车市场那套说法：买方分不出好车坏车，于是只肯出平均价，于是好车退出，于是市场只剩坏车。

差别有两处。

第一，机制不同。柠檬市场的病因是信息不对称——卖方知道，买方不知道。本文说的稀释在**双方都诚实、都不隐瞒**的情况下照样发生：一个学徒诚实地照抄师傅的做法，没有任何欺骗，如果他此后从不独立承担一件活，那个"跟这位师傅学过"的信号照样在被稀释。

第二，本文给出了方向性预测。柠檬市场只说"会变坏"，本文说**越像坏得越快**，而且有一个可以估的比例阈值。这是两个不同强度的说法。

与古德哈特定律

一个指标一旦成为目标，就不再是好指标——因为人们会去优化那个指标本身。

本文与它的分界，和柠檬市场那一条同源，但更彻底：**古德哈特需要有人在博弈，本文不需要。**

那只无毒蝴蝶没有在博弈，它甚至不知道有个信号存在。那个照抄了全套动作却从不真的开盒子的人，可能完全出于崇敬。**稀释是一个自动发生的统计后果，不需要任何人有意图。**

这一条很要紧，因为它决定了对策的方向：针对古德哈特的办法是**反博弈**（换指标、多指标、抽查），而这些办法对本文说的稀释**几乎无效**——因为没有人**在博弈**。

与公地悲剧

这是本文最愿意承认的一个近邻：信号的可靠性确实像一个公共池，模仿者**确实**在提取。

但有一个决定性的差别：**公地是可以被补充的**。牧场休耕会长回来，鱼群禁渔会恢复。所以公地悲剧的标准解法是产权化、配额、共同治理——都建立在“只要停止过度提取，资源会回来”这个前提上。

而这个信号池**无法被单方面补充**。被模仿者再努力也补不回来，因为——**他的努力恰恰又提供了新的可被模仿的材料**。做得越好，可抄的越多。这是一个公地悲剧里没有的、也不可能有的性质：**补充资源的动作，同时在扩大提取面**。

与知识产权

版权、专利、署名规范——这一整套处理的是**归属**：这东西是谁的，谁有权用。

本文处理的是**存亡**：这个信号还成不成立。两者可以完全脱钩。

一门完全公开、明确鼓励所有人自由使用的公开课，在授权上百分之百干净；而如果它培养出的人绝大多数从不进入这门手艺的后果里，“上过这门课”这个信号照样会失效。**合法的模仿一样稀释**。这一条决定了知识产权那一整套工具对这个问题基本无能为力。

与过度模仿研究自身

最后一个划界，也是最微妙的：本文与它借用的那套研究之间的差别。

过度模仿研究把高保真复制当作一项**纯粹的适应优势**，因为它的全部实验都在一个学习者随后必须真的把事做成的设置里进行。它没有理由去问：如果学习者永远不需要真的做成，会怎样。

而这恰恰是现在的处境。本文对那套研究的增量是：它把一个被固定的实验条件，还原成了一个变量。

第九章·比例：为什么加强查验注定输

两个层面的错配

拟态实验最有价值的一点，是它把这件事变成了**比例**问题：崩不崩溃，取决于不挨打的模仿者占多大份额。

而人类社会应对这类问题的全部工具，几乎都在**个体**层面：查重、鉴定、背景核查、AI 检测、面试追问。它们的动作单位是“抓出这一个”。

这里有一个层面上的错配，而且是致命的：

抓出个体，只在抓得足够多时才影响比例。而如果模仿的成本远低于查验的成本，这一仗在算术上就已经输了。

洗稿那一条已经从另一个方向证明了同一件事：查重对洗稿的失效是**构造性**的，不是灵敏度问题。加强灵敏度不解决构造问题。

为什么这一仗的算术从一开始就是输的

把三个数字放在一起看：

造一份仿制品的成本——正在快速趋近于零。

查验一份的成本——基本没变，因为查验需要人，而且需要懂行的人。

被模仿者做出真东西的成本——完全没变。

三个数里，只有第一个在动，而它在往下掉。这意味着比例会一路往上，而查验的总量受限于懂行的人数。

这不是悲观，这是一个可以算的判断。它推出的结论是：**凡是主要依靠"识别出假的"来维持的信号，都在倒计时**。剩下的问题只是它还有多久。

那还能做什么

只剩另一个方向：**不去减少模仿者，而是把模仿者接进后果链**。

这句话的意思是：不要试图分辨谁是照着做的，而是让所有声称的人都进入"如果这件事错了，你会承担"的位置。一旦进入，模仿这件事就自动失去了危害——因为一个会跟着挨打的模仿者，要么真的学会，要么被后果淘汰。**他不再是稀释者，他成了充值者**。

这一转向的关键在于：它**不需要区分真假**。而这正是它能赢的原因——所有需要区分真假的办法都输在成本上，而这个办法绕开了区分。

第十三章会给具体做法。

第十章·一个可以估的数

前面说崩溃是比例问题。这一章把它做成一个可以估的东西——因为一个只能定性说的判断，用处有限。

两个数

需要的只有两个量。

第一个是纯度，记作 p ：在所有拿这个信号去声称的人里面，跟着挨打的占多大比例。

第二个是接收者的处境，由两件事决定：他**绕开这个信号自己去验一次**要花多少（试探成本 c ），以及他**信错了一次要赔多少**（被骗代价 L ）。

接收者什么时候会放弃这个信号

接收者的算盘很朴素：继续信这个信号，他有 $(1-p)$ 的概率吃亏，期望损失是 $(1-p) \times L$ ；不信而自己去验，他确定要花 c 。

于是当 $(1-p) \times L$ 超过 c 的时候，他就该改成自己去验。把这个不等式反过来写，得到那个临界纯度：

$$p^* = 1 - c / L$$

纯度低于 p^* ，这个信号对这个接收者就不再值得用。它没有变成假的，它只是不再比自己去看一眼更划算——而这正是它作为信号的死亡。

三条能直接用的推论

推论一：信错了代价越大的领域，要求的纯度越高。

L 越大， c/L 越小， p^* 越接近 1。所以外科、结构工程、飞行、司法这些"信错了会死人"的领域，对纯度的要求近乎苛刻——这不是行业保守，是算术。它们必须把 p 维持在极高的位置，否则接收者会立刻放弃这个信号。而维持高 p 的唯一办法，就是执照、追责、可撤销——全都是"让持证者跟着挨打"的装置。

反过来，一个信错了没什么代价的领域（比如挑一家餐馆）， L 很小， p^* 很低，所以它可以容忍大量不挨打的模仿者而信号照样能用。这解释了为什么点评类信号被水军污染多年仍未彻底失效，而一个医生执照的含金量一旦被怀疑就是大事。

推论二：验证越便宜，信号越早被放弃。

c 越小， p^* 越高。当自己去验一次几乎不花钱时， p^* 趋近于 1——意味着只要有一点点不纯净，接收者就该放弃这个信号，直接自己看。

这一条解释了一个容易被误读的现象：这几年"现场做一件小的"在到处回归，通常被解读成"大家变严格了"。其实不是。是验证的成本掉下去了，于是绕开信号变得划算。信号没有被修好，它是被绕过去了。

推论三：现在是两头夹击。

这是本章最要紧的一条。把前面几章的判断放进这两个数里，会看到两个变量正在朝相反方向同时移动：

- 做一份像样东西的成本在掉 → 不挨打的声称者变多 → p 在下降。
- 自己去验一次的成本也在掉 → p^* 在上升。

一边是实际纯度在往下走，一边是要求的纯度在往上抬。两条线相向而行，交叉点来得比任何一边单独变化都快。

这就是为什么这一轮的信号贬值给人的感觉是"忽然之间全垮了"——不是某个信号出了事，是所有以"做出一份东西"为形式的信号，同时被两头夹住。

这个数怎么粗估

不需要精确，能分出量级就够用。三个问题：

估 p ：在拿这个凭证去声称的人里，有多少人此后会因为这件事做错而实际承担后果？——不是"理论上会被追责"，是"实际发生过吗"。查一查过去三年这个领域有没有人因此被撤销、被追回、被记录在案。一个都没有，那 p 大概率很低。

估 c ：让我自己验一次要花多少？半小时的现场测试？一个月的试用期？如果这个数在下降，你手上的信号在加速贬值，与它自身的质量无关。

估 L ：我信错一次要赔多少？这个数决定了你该不该费力去维护这个信号，还是干脆改用直接验证。

一个诚实的限定

这三个量都不好精确测，而且上面那个式子是一个粗糙的一阶近似——它假设接收者算得清、假设每次判断独立、假设 L 是一个可比的数。这些假设都只是大致成立。

它的用处不在于算出小数点，而在于它把"这个凭证还行不行"这个含糊的感觉，变成了三个可以分别去查的问题。分不出小数点没关系，能分出"这个信号还有五年还是已经过去了"就足够改变决定。

第十一章·崩溃之后，谁损失最大

一个不太好接受的答案

直觉上，一个信号失效之后，受损的是被骗的人。这没错，但不是最重的。

拟态实验给的答案很干脆：**崩溃时，捕食者更容易攻击那个花纹——包括真正有毒的那些。**

翻成人的处境：当一个信号被稀释到失效，**损失最重的是那些原本靠它被认出来的、货真价实的人。**

骗子在崩溃前后的处境没差多少：崩溃前他冒充，崩溃后他去冒充下一个信号。真正做出东西的人不一样——他失去的是**他唯一一件能证明自己的东西**，而他并没有第二件。

"人人都能做出好东西"这句话的直接后果，不是好东西变多，是做得出好东西的人无法再被认出来。

这解释了一种到处可见的憋闷

这几年有一种情绪在很多行业里蔓延，而当事人常常说不清它是什么。它不是"我被抢了饭碗"——很多人的活并没有少。它更像是：**我做的事和别人做的事，从外面看不出区别了。**

按本文，这不是心理问题，是一个可以被指认的结构性事实：他所依赖的那个信号已经越过了阈值。他的本事没有减少一分，只是**承载这份本事的那条通道被塞满了。**

这也解释了为什么"你要提高自己"这类劝告在这里格外无力——问题不在他那一头。他再提高，那条通道照样是塞的。

为什么被模仿者不能靠更努力自救

这里有一个残酷的、也是本文与公地悲剧最不同的地方，值得单独说。

面对信号被稀释，货真价实的那一方最自然的反应是：我做得更好、更深、更难，把他们甩开。

这个反应在直觉上无懈可击，在结构上却是失效的，原因有两层。

第一层：他做出来的每一样新东西，同时是新的可抄材料。他把手法推进一步，这一步很快也会被照着做。他跑得越快，提供的模板越多。这是公地悲剧里不存在的性质——**牧民多养一头羊不会让草长得更容易被吃，而他做得更好确实让模仿变得更容易。**

第二层：接收者分辨不出"更好"。按第十一章，在保真度足够高的地方，接收者本来就区分不了；而"更深一层"这种差别，恰恰是最难从成品上被外行读出来的那一类。**他为之付出的额外努力，落在接收者看不见的那一段上。**

所以"做得更好"这条路不是不该走——它对做事本身当然有意义——但它**解决不了信号问题**。信号问题只能在信号那一层解决：要么提高 p （把声称者接进后果），要么降低 c （让验证变便宜，主动绕过信号）。

值得注意的是，第二条对被模仿者其实是有利的，而且他可以单方面推动：**主动提出"我们现场做一件"，比任何自我证明都有效。**他不是**在证明自己**，他是在把接收者带到一个 c 极低的场景里——而在那个场景里，纯度不再要紧。

一个被忽略的次生后果

还有一层，比上面更麻烦。

信号崩溃之后，接收者不会就此放弃判断，他会去找**替代信号**。而替代信号通常更差——更依赖关系、更依赖出身、更依赖已经出名这件事本身。

因为可靠的信号需要时间去建立，而崩溃是很快。在新的可靠信号长出来之前，人们只能退回那些**虽然不精确但难以伪造**的东西：我认识他、他是某某带出来的、他已经有名了。

所以信号崩溃的直接后果不是无序，是复辟。它把一个曾经靠"你做得出来"来分配机会的领域，推回到靠"你是谁认识谁"来分配。

这一条相当重要，因为它说明了这件事的利害不止于诚信：**可靠的能力信号是流动性的基础设施**，它一垮，向上流动的通道跟着窄。

第十二章 · 一次全局跳变

不是很多个危机，是同一个变量动了一次

把上面的框架推到底，会得到本文最大的一个判断。

这几年，很多领域同时在喊同一句话：文凭不管用了、作品集不管用了、论文数量不管用了、推荐信不管用了、口试也不管用了。这些通常被当成很多个独立的问题，各自归因于各自领域的毛病。

按本文，它们是**同一个变量的一次跳变**。

那个变量是：**做出一份像样东西的成本，与真的具备那份本事的成本，两者的比值。**

凡是靠"做出来很难"来承载的信号，它的寿命就等于这个比值降到某个水平所需要的时间。而生成式系统做的事，是在几乎所有这类信号上**同时**把这个比值压低了一大截。

所以这不是若干个领域各自出了问题，是一次全局的、同步的跳变。它解释了为什么这些抱怨在同一年里一起出现，也解释了为什么每个领域各自的改革都不太管用——**因为动的不是那个领域的变量。**

一份清单

把这个比值下降最快的几类信号列出来，可以当作一份优先级表：

信号	它原本承载什么	现在的处境
一份写得好的文档	此人能把事情想清楚	比值已经接近零
一个作品集	此人能做出这个水准	下降极快
一段能跑的代码	此人懂这个系统	下降极快
一篇论文	此人独立走过一条论证路径	洗稿早已侵蚀，现在加速
一封推荐信	某个可信的人愿意为他背书	取决于写信人是否跟着挨打
一次现场解题	此人当场能做	比值几乎没变
长期共事的口碑	此人在真实后果里做过很多次	比值几乎没变

请看最后两行。它们之所以没变，不是因为技术还没追上，而是因为它们**本来就不是"做出一件东西"这种形式的信号**——它们直接测的是"这个人在后果里的表现"。

这给出了一个预测：**接下来会被重新重视起来的，一定是这两类**。而这个预测已经在发生——现场做题、试用期、共事推荐的权重正在各处回升。本文的贡献不是预言它，是**说清它为什么**：它们不是更严格，是它们测的东西不在被压低的那个比值上。

历史上这件事发生过好几次

把这一章读成"我们这一代遇上了前所未有的事"，会读偏。这个比值一直在动，只是过去慢，而且每一次都留下了可以借鉴的东西。

印刷术之后。手抄本时代，一本书的存在本身就意味着有人花了几个月抄它，因而值得抄。印刷把这个成本压到极低，"有一本书讲这个"不再意味着任何东西。后来长出来的替代信号是**出版社与编辑**——一个会为它出的书承担声誉后果的中间人。请注意这个替代物的形状：**它正是一条被人工接上的后果链**。

照相术之后。"一幅画得很像的肖像"曾经意味着画师的本事，照相把这个成本压到接近零。绘画没有消失，它转向了照相做不到的地方；而**"像"这个信号本身被放弃了**——没有人再试图去鉴定一幅画像不像。

复印与模板之后。一份格式规整的文件曾经意味着有人认真做过，复印机和模板让它一文不值。替代物是**签字与用印**——同样是把一个具体的人接进后果。

三次的应对方式其实只有两种：**要么人工接一条后果链（出版社、签字），要么放弃这个信号改用别的（不再看像不像）**。没有第三种。而这两种，正是第十四章那五条做法的两个方向。

这一段也给了一个不太乐观的提示：三次里没有一次是靠**加强检测**解决的。检测在每一次都被尝试过，也在每一次都输了。

一件必须说清的事：这不是技术的错

把这一章读成"生成式系统毁了信誉"是读错了。

那个比值一直在下降，只是过去慢。印刷术、照相、录音、复印机、模板、外包写作——每一次都把某一类"做出来很难"变得容易一些，每一次都有一批信号失效，每一次也都有新的信号长出来。

这一次的特别之处只有两点：**快**，以及**同时**。快，让新的信号来不及长出来；同时，让人无处可退——你不能从一个失效的领域跳到另一个还好的领域，因为它们一起在动。

为什么会有人心甘情愿地进入后果

第十三、十四章会给一堆"把后果接回去"的做法。这里先处理一个必然会被问到的问题：**凭什么有人愿意？**

如果不挨打是划算的，那么一个理性的声称者应当尽力避免进入后果。按这个推理，上面所有做法都推不动。

可现实里有大量人主动往里走：医生考执照、工程师签字盖章、作者写下可被推翻的预测、师傅让徒弟署自己的名。为什么？

答案在第十一章那条推论里：**信号一旦崩溃，损失最重的是货真价实的那一方**。所以对一个真有本事的人来说，进入后果不是牺牲，是**他把自己从一堆无法区分的人里摘出来的唯一办法**。他愿意挨打，恰恰因为他觉得自己挨不上。

这一条给出了一个很实用的读法：

面对两个人，一个主动提出"如果错了我担着"，一个避开这句话——你不需要判断谁的本事大。愿意进入后果这件事本身，就是当下最难伪造的信号。

它难伪造，不是因为说这句话有成本（说说很便宜），而是因为说了之后要兑现，而兑现的时点在将来、由事实决定、不由他决定。一个没本事的人可以说这句话，但他会在第一次兑现时被清出去。

于是有一个略带讽刺的收束：在一个所有"做出来的东西"都不再可信的时代，最可靠的信号退回到了最古老的那一种——一个人肯不肯为这件事担着。这不是道德的回归，是算术的结果：它是唯一一个仿制成本没有被压低的東西。

第十三章·修复的方向：把后果接回去

诊断完了给做法。五条，全都指向同一件事：不去分辨谁是照着做的，而是让所有声称的人都进入后果。

一·当场做一件小的

不看作品集，看他在你面前做一件小的、真的、会有结果的事。

要点是理解它为什么有效。常见的解释是"现场没法作弊"——这个解释很快会过时。真正的原因是：当场做一件事，是在当场造出一条后果链。他做的这一件，对错就在眼前，由他自己承担。

所以设计的关键不是防作弊，是让那件事有真实的对错。一道有标准答案的题不算；一件必须当场取舍、而且取舍会立刻显出后果的事才算。

二·把署名和后续绑在一起

不是查一篇东西是不是自己写的，是让署名者对这篇东西此后的问题负责——被引用后发现错了要撤、被采用后出了事要担、结论被推翻要留在记录里。

这一条会立刻改变洗稿的收益结构：洗稿的全部好处，建立在"拿走路径而不承担这条路径的风险"上。一旦风险跟着署名走，搬一条自己没走过的路就变成了一件危险的事——因为他不知道这条路在哪里会断。

三·让推荐人跟着挨打

推荐信之所以正在失效，不是因为写信的人不诚实，是因为写信这件事对写信人没有任何后续：推荐错了，他不损失什么。

修法很简单，而且历史上存在过：让推荐记录可追溯、可回看。一个人推荐过的十个人后来怎么样，是可以被看到的。不需要惩罚，只需要可见——可见本身就是一条后果链。

四·学徒必须有自己那份活

照着师傅做没有任何问题，问题在于他有没有一件署自己名、后果落在自己头上的活。

这一条正是第三章那个前提的制度化：过度模仿之所以是引擎，是因为后面接着一条会筛选的后果链。没有那条链的照抄，是纯粹的稀释——哪怕抄的人一片诚心。

可检验的迹象：一个跟了三年的人，有没有一件事是"这件砸了算他的"。没有，那三年他一直在稀释他师傅的招牌。

五·给你自己的东西装上一个会回来找你的赌注

这一条是给做事的人自己的。

在你交出去的东西里，写一句会回来找你的话：一个可以被检验的预测、一个明确的失败条件、一个"如果几年后没出现某件事，我这套说法就是错的"。

它的作用不是让别人信你，是**把你自己接进后果链**。一份带着这样一句话的东西，和一份不带的东西，在稀释与充值这两端上是相反的——哪怕内容一模一样。

五条的共同结构：没有一条是在检测模仿，全都是在**制造后果**。这正是第九章那一转向的落实——所有需要区分真假的办法都输在成本上，而制造后果绕开了区分。

第十四章 · 六个位置各一件事

招人的

把"看材料"的权重往下压，把"当场做一件小的"和"试用期里的一件真活"往上提。理由不是材料会作假，是**材料所承载的那个比值已经掉下去了**。

带徒弟的

让他照抄，越全越好——但必须给他一件署他名的活。**这两件事要同时做，缺一件都不成立**。只给抄不给活，是在稀释你自己的招牌；只给活不给抄，是让他从零开始，累积不起来。

做评审的

把力气从"这是不是抄的"移到"这个结论如果错了，谁会知道、谁会承担"。前一个问题的成本在上升且注定输；后一个问题几乎不花钱，而且它**不需要区分真假**。

发文章、发作品的

在里面留一个会回来找你的赌注。这是你在一个信号普遍贬值的环境里，**唯一还能自己给自己加的东西**——因为它不依赖任何人相信你，它只依赖时间。

做平台、做认证的

这是最要紧的一个位置，因为它决定了比例。

一个认证如果只测"能不能做出这样一份东西"，它就是在批量生产不挨打的模仿者，而且是它自己在稀释自己发出的那个证书。**认证的可靠性，等于持证者进入后果链的比例**。要保住它，就得让持证这件事与后续的表现挂钩——可撤销、可追溯、有记录。

用别人东西的人

照着别人的做法做事，本身没有任何问题——这是学习的正路。要问自己的只有一句：**这件事做错了，我会不会承担？**

会 → 尽管抄，越全越好。不会 → 那你正在稀释你所抄的那个人的信誉，哪怕你满怀敬意，哪怕你注明了出处。

第十五章 · 这条判断不适用的地方

一 · 大量模仿本来就该不挨打

学一门外语时照着母语者说，读一本书时跟着作者的思路走，看菜谱做菜——这些模仿里，模仿者确实不承担被模仿者的后果，而这完全没有问题。

因为这里**没有一个被消耗的信号**。没有人靠"会说这句话"来获得机会分配，所以不存在稀释。

本文只在一个条件下成立：**这个被模仿的东西同时是一个用于分配机会的信号**。不是信号，就不适用。

二 · 有些信号本来就该失效

不是所有信号都值得保。一个原本只在保护某个圈子的门槛，被大量模仿者冲垮，这是好事——那正是它该有的下场。

本文没有主张保住一切现存信号。它只主张**看清这件事的机制**，因为看清之后才谈得上选择保哪个、放哪个。

三 · 后果链本身可能是坏的

这一条最要紧。

"让人承担后果"听起来天然正确，但一条设计糟糕的后果链会造出比稀释更坏的东西：不敢做事、只做安全的事、出事互相推。之前那些讲压力与支持的讨论已经说过——**后果落在人身上还是落在事情上，是两码事**。

本文要的是**事情的后果**：这件事做错了会不会显出来、会不会回到做的人手里。不是**人的后果**：做错了会不会被处分。前者让人学会，后者让人不敢动。

四 · 这不是在反对模仿

本文从头到尾没有主张少模仿。第三章说得很清楚：连不理解的部分也照抄，是人类能累积文明的原因，而黑猩猩那种"只抄看得懂的"聪明做法，恰恰累积不起来。

本文主张的是**模仿要配一条后果链**，不是**少模仿**。这两件事的距离，正好是这一整篇文章。

第十六章 · 怎样才算我说错了

下面八条，任何一条成立，本文都该被推翻。

1. **保真度的符号不翻转**。如果在两类环境里（模仿者随后要独立承担 vs 不承担）比较模仿保真度与信号可靠性的关系，两者的相关方向相同，那么第七章那条核心推论不成立，本文垮掉。**这一条在师徒制与认证培训之间可以直接比对**。
2. **没有阈值**。如果人类社会的能力信号不表现出比例阈值现象——即可靠性随不挨打的模仿者比例连续平滑下降、不存在快速崩溃的区间——那么第九章的整套论证失去依据。
3. **崩溃后被模仿者不受损**。如果一个信号失效后，原本靠它被识别的高水平者的境遇没有变差，那么第十一章错，而那一章是本文最重的一条实践主张。
4. **加强查验有效**。如果在某个领域里，持续加大检测投入确实把不挨打的模仿者比例压回了阈值以下并维持住，那么第九章"算术上必输"就错了。**这一条有大量现成案例可查**。

5. **成本能替代后果**。如果存在一批可靠性很高、而其发出者明确不承担任何后续后果的信号，且它们的可靠性可以完全由发出成本解释，那么第八章与昂贵信号理论的划界失败，“跟着挨打”就只是“成本”的换皮。
6. **诚实的模仿者不稀释**。如果完全诚实、注明出处、无任何博弈意图的模仿者，即使不进入后果链也不降低信号可靠性，那么本文与古德哈特的划界失败，本文只是古德哈特的一个特例。
7. **那个比值没有同时下降**。如果各领域信号失效的时间点是分散的、各有各的原因，与“仿制成本／真做成本”这个比值的变化无关，那么第十二章那个“全局跳变”的判断没有依据。**这一条可以用近三年的时间序列直接查。**
8. **五条做法无效**。如果一个组织把评审重心从“查真伪”转到“接后果”（第十一章那五条），若干周期后在识别高水平者上的准确率没有提高，那么第十三章的处方是空的。

另补一条**本文不主张的**：本文没有主张模仿有害、没有主张加强惩罚、也没有主张技术进步是坏事。它只主张一件事——一次模仿是加固还是稀释一个信号，由模仿者是否被同一件事的失败所惩罚决定；而这个变量既不是相似度，也不是成本，也不是意图。其余全部内容都是从这一句推出来的。

第十七章·这个栏目自己就是不挨打的模仿

最后必须交代一件事，因为不交代，上面全部内容就有一个漏洞。

这篇文章，以及它所在的这个栏目，正是本文所描述的那种不挨打的模仿。

它做的事是：把三个领域各自辛苦建立起来的东西——野外拟态实验、二十年的过度模仿研究、一篇关于洗稿的分析——拿过来，重新组织成一个新的说法。它用了它们的可信度，而它**不进入它们任何一条后果链**：

- 如果那项拟态实验的结论被推翻，做这个实验的人要面对同行的质疑、要撤、要重做。**本文什么也不会发生。**
- 如果过度模仿的整套解释被证明是另一回事，那个领域的几十位研究者要重写他们的教科书章节。**本文什么也不会发生。**
- 如果洗稿那篇文章的核心判断站不住，它的作者要在自己那个栏目里承担。**本文什么也不会发生。**

按第七章那条推论，情况还要更糟一层：**这篇文章写得越像样、组织得越顺、读起来越有分量，它稀释得越快**——因为它越难被读者与那些真正在各自领域里挨打的工作区分开。

这不是一句自谦。它是本文自己的判据用在自己身上的直接结果，而且结果不利。

那么补救是什么

按第十三章第五条，唯一的补救是**给自己装一个会回来找你的赌注**。所以这里写死三条：

1. **到二〇三〇年**，如果主流的招聘与学术评审中，“当场做一件真活”与“可追溯的推荐记录”这两项的权重没有明显上升，而“看材料”仍占主导，那么本文第十二章那个预测错了。
2. **到二〇三〇年**，如果反抄袭与内容检测领域的投入仍在增长、且从业者普遍认为这条路走得通，那么本文第九章“算术上必输”的判断与整个领域的判断相反，而我这一边更可能是错的。
3. 如果本栏此后各篇的证伪条件，**到期后没有任何一位读者回来追问**——那么这些赌注就没有构成后果链，本章的补救是无效的，本栏仍然是一台稀释机。

第三条是最诚实的一条：一个赌注要成为后果链，需要有人来收账。而这件事不由写的人决定。

还有一栏账

三套来源、九个判断、三十六次两两对撞，留下十一条站得住的，其余二十五条被判定为强行联系或换皮重述，作废。作废的理由中，最值得留下的一条是：曾试图用"成本"来统一三家（即把本文的判断说成昂贵信号理论的一个推广），推到第八章那两个反例时垮掉——**成本可以为零而可靠，可以极高而不可靠**。正是这条弃案，逼出了"后果"这个词。

尾声

把两万字收成一句话：

照着别人做，是养出本事还是掏空信誉，不看它抄了什么、也不看它抄得多像，只看一件事——这件事做错了，抄的人跟不跟着挨那一下。

跟着挨的，抄得越全越好：多余的部分会被他自己的失败磨掉，必要的部分会被留住，而那个"跟这位学的"的信号会因为他而更结实。

不跟着挨的，抄得越像越坏：他和真货越难分，那个信号就越快地不再意味着任何东西。而到它彻底不意味着任何东西的那一天，损失最重的不是被骗的人，是那些原本靠它被认出来的、货真价实的人——他们的本事一分没少，只是再也没有一条路把这件事说给别人听。

这件事现在之所以急，不是因为人心变坏了。是因为"做出一份像样东西"和"真的具备那份本事"之间的成本差，在几乎所有地方同时被压薄了一大截。这不是很多个危机，是同一个数字动了一次。

而出路不在于把假的挑出来——那一仗在算术上已经输了。出路在于另一件成本低得多的事：**让所有声称的人，都站到会挨打的位置上去**。它不需要你分辨谁真谁假，这正是它还能赢的原因。

如果你读到这里正想着"这个说法挺有道理"——那么请对你手上正在做的那件事问一句：它如果错了，谁会知道，谁会承担？如果答案是"没有人"，那么无论你做得多好、多用心、多问心无愧，你都在稀释你正在使用的那个招牌。