改写得越彻底，越像自己的跨语种借用为何让作者的越界感与制度的查重同时归零

学 术 治 理 · RESEARCH INTEGRITY

改写得越彻底，越像自己的跨语种借用为何让作者的越界感与制度的查重同时归零

少 敏

SDE 学员专栏 · 约八千字 · 发表于2026年7月17日

摘　要

2026年7月，一篇2019年通过的硕士学位论文因9处与境外某篇期刊论文文字重合、未标注引用而被撤销学位。本文不重做事实调查，不推断当事人动机，只追问一个裁决不回答的问题：这9处穿过了导师指导、论文评阅、答辩、学位审核与查重系统，还穿过了一次正式调查——**五道关口是怎么同时失手的？**本文提出：写作者判断“这算不算我的”所用的量，与查重系统所测的量，是同一个量——**改写量**，即文字相对于来源的字面变形幅度。在单语写作里它相当准；但存在一类**归零操作**，能把字面变形推到满量程而让知识债务分毫未动，翻译是第一种，生成式模型重述是第二种。归零操作一旦介入，作者的越界感与制度的查重同时读出清白——这不是两次失灵，是一块表的一次读数在两处显示。由此得出：**N 道测同一个量的关口，防御力等于 1 道**；关口是否多重，不看数量，看是否正交。文章与瑞士奶酪模型（孔洞碰巧对齐，而非按设计永久共线）、补丁写作研究（越生疏越留痕，方向相反）、多手问题、越轨的常态化、心理安全与戴明的过程控制划出六条分界线，并给出与改写量正交的那个维度——**坐标**：你在哪一页读到它。据此把“可追溯性缺口”重新定义为可计算的数（1 减去来源可复核率），并给出四条证伪路径。

关键词：改写量　归零操作　共线关口　坐标　可追溯性缺口　跨语种借用　补丁写作

引言

九处，和它穿过的每一道关口

2026年7月13日晚，中国人民大学发布情况通报：经核查，涉事硕士学位论文有9处与境外某篇期刊论文存在文字重合，相关内容未标注引用、未列明参考文献；依据《高等学校预防与处理学术不端行为办法》与《中华人民共和国学位法》，学校认定构成学术不端，决定撤销其硕士学位。当事人随即公开表示接受处理并致歉。八天前的首轮通报，结论还只是"注释不规范"。

从纪律处分的角度看，事情已经完整：有可核验的文本重合，有规范依据，有调查与申辩程序，有正式裁决。本文不重做事实调查，不推翻正式认定，也不讨论未披露的内心动机。

但有一个问题，裁决没有回答，也不该由裁决回答：**这9处穿过了什么**。

它穿过了导师指导，穿过了论文评阅，穿过了答辩，穿过了学位审核，穿过了查重系统。2019年，它带着这五道关口的背书成为一个硕士学位。七年后，它还穿过了一次正式调查——那次调查看过它，并且给出了"不规范"。

对这种"多重关口同时失手"，现成的解释有三种：有人失职；有人包庇；责任在分工中被稀释了，每个人都以为别人会查。前两种需要假设动机。第三种——责任稀释——不需要，所以它通常被认为是最稳妥的解释。

本文认为第三种也不够。**它假设每个人只尽了部分责。可如果每个人都完整地尽了责呢？**

本文的判断是：这些关口看起来是五道，实际维度极低——**它们测的是同一个量**。而这个量，在跨过语言边界时会自动归零。于是不是五次失灵，是**一次失灵在五处显示**。

这个解释有一个不寻常的性质，也是本文选择它的原因：**它不需要假设任何人的动机**。 无论一段文字是被有意搬运还是无意丢失来源，仪表的读数都一样清白。所以"它如何穿过制度"这个问题，可以独立于"当事人想了什么"来回答——而后者，公开材料本来也回答不了。

全文七节。第一节命名那个人人都在用、却从没被命名的量。第二节说明翻译如何让它归零。第三节推出本文的核心结论：作者的自查与制度的查重共用这一块表，所以它们不是两道防线。第四节与最近的祖先——补丁写作与第二语言写作研究——划清界线。第五节与三种组织解释划清界线。第六节给出与该量正交的那个维度。第七节写明这套判断如何可能失败。

需要先声明一句，它决定了本文的分寸：**本案是本文的机缘，不是本文的证据**。 支持本文机制的材料在第二节，来自别处。

第一节

改写量：一个人人都在用、从没被命名的量

一个写作者判断"这段话算不算我的"，用的是什么？

不是法条。多数人写作时并不会调出《著作权法》或学位规章。用的是一种体感，而这种体感有一个相当稳定的刻度：**我改动了多少**。

刻度的三档，每个写过论文的人都熟悉。**照抄**：改动为零，体感是偷，明确越界。**换词式转述**：改动少许，体感是心虚，知道该标。**消化后用自己的话表达**：改动很大，体感是"我干了活，这是我的"。

本文把这个被反复调用、却从没被命名的量称为**改写量**——一段文字相对于其来源的字面变形幅度。

必须先说改写量的好话，否则后面的批评会落空。**在单语写作里，它相当准**。 因为在同一种语言里，改写量与"我是否重构了这个思想"高度相关：你要把一段话改得面目全非而又不出错，你得先吃透它。改不动，通常是因为没懂；改得动，通常是因为懂了。所以体感"改得多＝更是我的"，在单语环境里大致成立，甚至可以说是一条经济而有效的经验法则。

但它测的从来不是归属权。**它测的是字面变形**。 在单语环境里，字面变形与知识重构高度相关，所以两者的区别不显影——相关不是等同，而单语环境从来不提供把两者分开的机会。

这条经验法则要失效，需要一个特殊的输入：一种能把字面变形推到最大、同时让知识债务分毫未动的操作。本文称这类操作为**归零操作**。

这类操作存在，而且第一种就是学术写作的常规步骤。

第二节

翻译让仪表满量程

把一段英文译成通顺的中文，会发生什么？

字面变形：百分之百。 没有一个词幸存，没有一个句法结构幸存。查重系统在这两段文字之间找不到任何字面重合——不是因为它不够灵敏，是因为字面上确实没有重合。

劳动：真实。 好的翻译很难。你要读懂，要在中文里找到没有现成对应的说法，要重排语序，要处理术语，要让它读起来像中文而不是译制腔。这是几个小时的实活。

体感：满量程。"这个中文句子是我造的。"

而且这个体感**没有读错**。那个中文句子确实是你造的——不是别人写的，词是你选的，句子是你排的。改写量这块表尽职地测量了它该测的东西，并给出了一个准确的读数。

表没有坏。是它测的那个量，在跨过语言边界时与归属权脱钩了。

于是出现一个在单语环境里不可能出现的组合：**字面变形 100%，知识债务 100%**。 前者是你的，后者是别人的，而那块表只测前者。

这不是思辨。跨语种抄袭检测这个技术领域里，有一项关于学者实践与态度的调查给出了一个数字：**只有 36.25% 的学生认为，把外文片段翻译后放进自己的作品构成抄袭**。 近三分之二的人不把它体验为越界。

这个数字值得停一下。它不是"三分之二的人不诚实"——同一批人，你问他们照抄中文文献算不算抄袭，答案会压倒性地一致。他们的道德感没有问题，**是他们手里的表在这个输入下读出了清白**。

但必须立刻交代这个数字的处境，而它比"文献存在分歧"更糟。另一项针对西班牙莱里达大学 1150 名一年级学生的调查，给出的不只是一个不同的数字，而是一个反向的排序：认为翻译后使用构成抄袭的占 82.1%，认为复制粘贴构成抄袭的只有 69.3%，转述 68.3%。在那个人群里，**翻译被认作抄袭的比例高于照抄**——这与本文的单调预测（改写量越高、越界感越低）正好相反。

本文不打算把它解释掉。它意味着两件事，两件都得说。

第一，36.25% 不能算作本文的支持。两项调查方向相反，而本文没有材料判断哪一个更接近中文与英文之间的实情。所以这个机制的经验地位只能写作"待检验"，不能写作"已获支持"。本文的重量因此整个压在第八节的检验二上——那个同材料三臂设计之所以是必需的而不是锦上添花的，正因为现有的态度数据自相矛盾。

第二，这个反向排序本身是一条该追的线索。两项调查有一个可能相关的差别：语言对。莱里达的学生所设想的翻译，多半发生在西班牙语、加泰罗尼亚语与英语之间——同源词密集，语序相近，译文与原文的字面距离远小于中文与英文之间。若本文的机制成立，语言对越近，改写量归零得越不彻底，越界感就该残留得越多，82.1% 正落在这个方向上。但本文没有掌握两项调查各自的人群构成与问法细节，所以这是一个可以去测的猜想，不是一个解释——而它要测的东西，恰好是第八节的检验三。

这里必须立刻画一条线，本文的分寸全在这条线上：**本文不主张涉事作者是这个机制的实例**。 那需要草稿、笔记、译稿、访谈——公开材料一样也没有，而从终稿反推写作时的心理状态，正是本文批评的那种论证。本文主张的只是：这个机制存在，它有独立于本案的证据，而且它足以解释制度侧的谜题。至于当事人是有意搬运还是无意丢失，本文不知道，也不需要知道——**两种情况下，表的读数完全一样**。

第三节

两道防线，一个仪表

现在把视线从作者移到制度。

查重系统测什么？**字面重合度**。

也就是说——查重与作者的体感，**测的是同一个量**。

这不是巧合，而是设计。查重之所以能被信任、能被写进流程、能出一个人人接受的百分数，正是因为它把人人都有的那个体感自动化了、客观化了、可比较了。**它是那块表的机器版**。

后果是一个结构性的结论：**它们不是串联的两道关口，是并联的同一道**。 一个能让表归零的输入，会让两个读数同时清白——作者感到自己没越界，系统报告没有重合。这不是两次独立的失灵，是一块表的一次读数，在两个地方显示。

而且不止两处。把五道关口逐一拆开看它们实际测的量：

导师看到的是成熟的表述——改写量高 → 读作"消化过了"

评阅人看到的是流畅的论证——改写量高 → 读作"写作能力强"

答辩现场检验的是应答——测的是当场的掌握，不测来源的坐标

学院拿到的是查重报告——表的机器版

首轮调查启动的是常规核查——同一台机器，同一个量

五道关口，最多两个维度。四道共线于改写量；只有答辩测的是另一个量——现场掌握度。而这一道恰恰最有教益：**它不共线，它照样没用**。一个熟练的写作者当场谈论一段他译过的文字，会谈得头头是道——他确实掌握了那个内容，掌握得和自己写的一样好。多出来的这个维度，投影在"这段话是谁的"上，仍然是零。

于是原本的直觉可以升级成一句更准也更严的话：**关口的防御力，不追随关口数量，也不单纯追随维度数量——它追随这些关口张成的量空间里，是否存在一个能分辨该威胁的量**。五道关口张成两维（改写量、现场掌握度），而"知识债务归谁"在这两个方向上的投影都是零。加第六道测改写量的关口，维度不变；加第六道测某个不相干量的关口，维度加一，防御力仍然不变。前一种情形本文称为**共线关口**，后一种称为**离题的正交**——两者都是签字增加而防御不增加，只是失效的方式不同。

这个结论顺带解释了那个最刺眼的现象：为什么首轮调查看过它，还给出了"不规范"？因为**首轮调查也站在同一条线上**。它不是不够严格，它是在同一个维度上又量了一遍——量出来的当然还是清白。二次调查之所以不同，不是因为态度变了，而是因为"新线索"来自这条线之外：它给出的不是一个更低的重复率，是一个**具体的定位**——哪一篇，哪一段。

这是本文的枢纽，也是第六节的入口：**打破共线的，从来不是更严格的同一个量，是另一个量**。

第四节

补丁留下补丁，翻译不留

本文最近的祖先，是写作研究里一条已经走了三十年的路，必须切干净，否则本文只是它的一个脚注。

1993年，霍华德（Rebecca Moore Howard）为一种现象造了一个词：**补丁写作**（patchwriting）——从原文抄下来，然后删掉几个词、改动语法结构、塞进一对一的同义词替换。她的关键主张不是"这是抄袭"，而是相反：**这是学习阶段的一种策略，不是不诚实**。 一个刚进入某个领域的写作者，还没有能力用自己的话重述，于是贴着原文挪。1995年她把这个主张扩展为对"学术死刑"的整体质疑。

这条路后来走得很实。佩科拉里（Pecorari, 2003）发现，第二语言写作者即使在博士论文里也补丁写作。石（Shi, 2004）发现，中国学生用英文做摘要时，照搬的词串比母语者更长。罗伊格（Roig, 2001）甚至发现，让心理学教授去概括一段陌生领域的复杂文本，22% 的人会补丁写作。

这条路研究的是同一件事：**能力不足的人，怎样在文本上留下痕迹**。 补丁写作的每一个特征——删词、换同义词、微调语法——都是**改写量不足**的表现。改写量不足，痕迹就留下来了，所以它可以被看见、被研究、被写成三十年的文献。

本案是这条路的反方向。

不是用弱势语言写作，是用**母语**写作。读的是英文，写的是中文。输出语言的能力不是不足，是充分——甚至过剩。没有补丁，没有删词的疤，没有同义词替换的僵硬。中文是漂亮的、原创感十足的中文。**改写量不是不足，是满量程**。

于是两条路的机制正好倒转：

补丁写作：能力不足 → 改写量低 → 留下痕迹 → 可检出 → 有文献

跨语种转换：能力充分 → 改写量满 → 抹除痕迹 → 难检出 → 无文献

霍华德的补丁写作是**能力不足的症状**，会随训练消失；跨语种归零是**能力充分的产物**，越熟练越彻底。前者训练可以治，后者训练只会加重——你把一个人的翻译能力训练得越好，他留下的痕迹越少。

这里还有一个值得说出口的观察。为什么补丁写作有三十年文献，而跨语种归零几乎没有？**因为文献偏向可检出的案例——而不可检出的案例，按定义就没有数据**。 这不是研究者的疏忽，是机制自己造成的：一个抹除全部痕迹的现象，不会在痕迹的研究里现身。第二节那两个相反的数字之所以仍然值得看，正因为它们绕开了这一点——它们不是检测研究，是**态度调查**。它问的不是"我们抓到了多少"，是"人们感觉这算不算"。在一个抹除痕迹的机制面前，问感觉，可能是唯一还开着的窗。

第五节

与三种组织解释分界

组织研究与安全科学里有四个成熟的解释框架，看上去都能覆盖本案。四个都不够，而不够的地方各不相同。

瑞士奶酪模型（Reason，1990）。这是离本文最近、因而最危险的祖先——它讲的正是"多层防御何时失效"。里森把系统的防御画成一叠瑞士奶酪：每片都有孔洞（每层都有弱点），事故发生在**所有孔洞碰巧对齐**的那一刻。这个模型统治了航空、核电与医疗安全三十年。本文的"共线关口"看上去只是它的一个中文说法。

差别落在一个词上：**碰巧**。里森的孔洞是随机的、漂移的、时开时合的——正因为随机，他的解药"多加一层"在他自己的模型里成立：每加一片，全部孔洞同时对齐的概率就再降一个数量级。本文的关口不是碰巧对齐的。**它们按设计永久共线**：因为测的是同一个量，那就不是五片上各自随机分布的五个斑点，而是同一个孔在五片上的五次投影。概率在这里不起作用——一个能让改写量归零的输入，让五片的孔**必然**对齐，每一次都对齐。

于是解药也反了。在里森的模型里加一层，防御力上升；在本文的情形里加一层，防御力**一点也不上升**。安全科学后来把"层与层其实并不独立"这件事叫做共模失效，并且知道它是瑞士奶酪模型最常见也最危险的误用——就此而言，本文可以被读作共模失效在学术审查上的一个实例。而本文多出来的那一点，恰恰是共模失效理论给不出的：**它说得出那个共同的模是什么**。共模失效告诉你"层可能不独立"，却不告诉你在任何一个具体系统里那个被共享的量叫什么名字——因此它只能在事故之后回头指认。本文给出的是一个可以预先指认的量：改写量。能命名它，才能在事故发生前算出哪几道关口其实是同一道。

多手问题（Thompson, 1980）。一件坏事经过许多人之手，每个人只贡献一小部分，于是谁也不为整体负责，责任在分工中稀释。它的解药是明确分责。本文的情形更奇怪：**没有人只尽了部分责**。导师看了，评阅人看了，答辩问了，学院查了，调查组查了——每一个人都完整地履行了自己那一环的职责，而且履行得没有瑕疵。**不是责任被稀释了，是责任在同一个维度上被完整履行了五次**。因此，多手问题的解药在这里失效：分责再明确，分的仍是同一个维度；把一条线切成五段，它还是一条线。

越轨的常态化（Vaughan, 1996）。挑战者号的机制是：一个群体在时间中反复接受微小的偏差读数，异常慢慢变成了正常。它需要三个条件——一个群体，共同的数据，逐步的过程。学位论文一样都不满足：写作是单人的，审查是分散的（导师、评阅人、答辩委员素不相识，看不到彼此的判断），而且**没有"逐步"可言**。挑战者号的工程师们**看见过**异常读数，只是一次次接受了它；这里从头到尾**没有出现过异常读数**。不是异常被常态化了，是异常从未显示。

心理安全（Edmondson，1996）。埃德蒙森那个著名的发现——报错更多的医疗团队，实际表现反而更好，因为差别不在犯错率而在敢不敢报——解决的是**隐瞒**。本文处理的是它前面一格：**报什么**。如果作者手里的表读数清白，他没有东西可报；如果导师的表读数清白，他没有东西可问。**再安全的环境，也不会让人报告他没看见的东西**。心理安全是必要的，但它是第二步；第一步是让东西可见。

最后一刀留给最疼的那个——戴明。《走出危机》第三条：停止依靠检验来达成质量。这句话是"从事后追责转向全过程治理"这个整个纲领的祖师爷，也是本案最容易被套上的解释：终点查重不够，要把核验前移到过程里。

但全面质量管理的过程控制有一个前提：**过程变量可测**。你能在生产线上量出温度、压力、公差，所以你能在缺陷形成前干预。本文的情形是**可测性本身的缺失**——不是没有量，是量的那个东西在特定输入下与要防的那件事脱钩了。

所以这里有一句必须说的、对本文自己的纲领也成立的话：**过程核验如果仍然只测改写量，它只是把同一块表往前搬**。在开题问"你改写够了吗"，在中期问"你改写够了吗"，在预答辩再问一遍——三次都清白，三次都没用。戴明的增量在"过程 vs 终点"；本文的增量在**"维度 vs 数量"**。这两件事经常被当成一件，而它们不是：把一道共线的关口前移，它仍然是共线的。

第六节

正交的那个维度：坐标

如果加关口无用，加什么有用？

需要一个与改写量**正交**的量——一个不会因为改写量上升而下降的量。

这个量存在，而且平凡得让人失望：**坐标**。

你读的是哪一篇？哪个版本？第几页？哪一句？译法的依据是什么？

坐标的关键性质只有一条，但它是决定性的：**改写量再大，坐标也不变**。 你把一段话译得再漂亮，把它重铸得再彻底，它在原文里的位置一个字都不会移动。**改写是对文本做的事，坐标是关于文本来源的事——后者不在前者的作用域里**。 所以坐标是唯一一个不随改写量归零的量。

由此得到核验的原则：**不问改写量，问坐标**。

"你这段改写得够不够"——一个熟练的写作者永远答得出来，而且答得漂亮。"你在哪一页读到它的"——改写得再彻底也答不出来，除非他在那一页读过。

这条原则回收了原本只凭直觉成立的那些做法，并且第一次给了它们理由：

为什么过程材料有用？ 不是因为"留痕"是一种行政美德，是因为**过程材料就是坐标的载体**——阅读记录里的页码、译稿旁的原句、版本之间保留的临时标注，全都是与改写量正交的东西。要求学生保留它们，不是要求他们证明自己勤奋，是要求他们保留那个查重查不出的维度。

为什么答辩现场追问核心文献有用？ 因为"这个判断你在哪读到的"这个问题，无法用改写能力回答。

三道关口因此获得了一个原则，而不只是一个清单：来源核验、过程核验、责任核验必须测三**个彼此正交、且至少有一个是坐标**的量——只正交不够，那会得到离题的正交；否则它是一道关口做了三遍。来源核验测坐标的可复核性；过程核验测坐标在时间中的保存度；责任核验测的则是元层面的东西——**每个节点是否被指派了不同的量**。多一道签字不增加维度；多一个量才增加维度。

这也让原本含糊的"可追溯性缺口"变成一个可以算出来的数。定义：**抽取一篇论文的 N 个关键论点，能在合理成本内返回具体来源坐标的比例，即来源可复核率；1 减去它，就是可追溯性缺口**。 这不是修辞，是一个可以在任何一篇论文上、由任何一个第三方、用一个下午算出来的数。它与重复率正交：一篇重复率 2% 的论文，可复核率可以是 30%。

最后一条推论，直接对着本案的那个特征：**跨语种材料应触发坐标强化，而不是重复率强化**。 面对一份大量使用外文文献的稿子，加大查重力度是加大一个已经归零的量。该要的是四样与改写量完全正交的东西：**原文出处、页码、原句、译法说明**。 改写得再彻底，这四样也答不出来——而这，恰恰是它们的全部价值。

第七节

另一台归零机器

翻译不是唯一的归零操作，只是第一种。第二种已经在每一张书桌上。

把一篇英文论文交给生成式模型，让它"用中文梳理要点并展开"，输出是什么？流畅的、原创感十足的中文。字面重合：零。知识债务：全部在源头。

它比翻译更彻底。翻译至少保留了原文句子的骨架——语序、句法、论证的推进节奏，多少能看出来源的形状。模型重述连骨架都换掉：它可以把一段演绎重排成归纳，把三个句子并成一个，把术语替换成同义的表达。**改写量不只是满量程，是溢出了量程**。

而体感这一头更麻烦。译者至少经手了那个转换——他坐了几个小时，他知道自己在处理别人的句子。使用模型的人**连这个都没有**。他既没有"我抄了"的感觉，也没有"我译了"的感觉；他有的是"我让它整理了一下"的感觉。**改写量满，而且连劳动感这个最后的提示都不需要了**。

于是"人工智能检测分数"作为对策，就落进了本文诊断的那个陷阱：它是**架在同一根轴上的第三台仪器**。它测的仍然是文本自身的字面与统计特征——测得更精巧，但测的还是那个量。把它加进流程，只是让共线的关口从五道变成六道。

本文的原则在这里一个字也不用改，只需要重申：**问坐标**。

不论一段文字经过多少次转换、由人还是由机器转换、转换得多么彻底，"这个判断你在哪一页读到的"这个问题的答案都不变。这就是坐标相对于一切检测分数的长处：**检测分数随技术进步而失效，坐标不会**。 每出现一种新的归零操作，检测方就要重新训练、重新追赶，而且永远慢一步；坐标这个量则根本不在那场竞赛里——它不在文本上，它在文本之外。

这也给出一个可以直接观察的预测，写在下一节的检验四里：如果本文是对的，人工智能辅助写作普及之后，**重复率与可追溯性缺口应当反向移动**——重复率越来越好看，缺口越来越大。

第八节

这套判断如何可能失败

一套关于"看不见"的理论，最方便的活法是永远正确：查出来了，说明维度够；没查出来，说明维度不够。本文拒绝这条活路。以下写明它在什么情形下应当被判为错。

检验一 · 维度而非数量（核心检验）。 本文的核心命题是"防御力追随测量维度数，而非关口数量"。它可以这样检验：对一批高校的学位论文审查流程逐道关口编码，标出它实际测的量（字面重合度／坐标可复核性／论证接口／现场掌握度）。命题预测：**在关口数量相同时，维度数更高的流程，其问题的发现时点显著更早；而在维度数相同时，增加关口数量对发现时点没有改善**。 若数据显示发现时点只追随关口数量、与维度数无关，本文核心命题为伪。

检验二 · 越界感（成本最低，可以立刻做）。 命题预测：写作者的标注意愿追随改写量，而非知识债务。设计：同一份外文材料，被试分三组——(a) 直接引用原文，(b) 中文转述，(c) 译为中文后使用。三组的知识债务完全相同。预测：**(c) 组的标注意愿显著低于 (a)(b) 组**。

第二节的两个数字在这一点上互相矛盾（36.25% 对 82.1%，且后者的排序与本文预测相反），这正是检验二不可省的理由。而且即使它们一致，本文也不会把它们当作证明：它们是态度调查，不是同材料对照，被试面对的是一个抽象问句而非一段具体文本，而人在抽象问句上的回答与在具体文本上的行为经常不一致。它们是线索，不是结论。**若 (c) 组与 (b) 组的标注意愿没有显著差异，则"改写量仪表"这个假说为伪**，本文的第一、二、三节一起垮掉。

检验三 · 检出延迟。 命题预测：跨语种借用的检出应系统性地晚于同语种借用，且延迟幅度随改写量归零的程度上升——语言对的距离越大（比如中英之间远于中日之间），延迟越长。可在撤稿与更正记录上做。**若两者的检出延迟没有系统差异，或延迟与语言对距离无关，机制为伪**。

检验四 · 反向移动。 命题预测：重复率与可追溯性缺口是两个正交的量，因此在归零操作普及时应当**反向移动**——人工智能辅助写作越普及，抽样论文的重复率应当越低，而来源可复核率也应当越低（缺口越大）。这可以在同一批院系的连续年份样本上直接测。**若两者同向移动——重复率下降的同时可复核率上升——则本文的正交性主张为伪**，第六节整节垮掉。

边界声明，四条，一条也不能省。

其一，**本文不推断当事人的动机**。机制解释的力量恰恰在于它不需要动机假设——有意与无意，表的读数一样。这也意味着本文对"她是否故意"这个问题**没有立场**，因为它不在本文的作用域内。

其二，**仪表读数清白不是免责**。署名意味着对来源真实性负首要责任。解释一个人为什么可能感觉不到越界，与认定他确实越了界，是两件同时成立的事——正式认定管已发生的行为，机制解释管下一次怎样更早发现。**本文一个字也不用来削弱前者**。

其三，**不存在不可归零的表**。坐标也可以被伪造：页码可以编，原句可以拼。坐标的价值不是绝对可靠，只是**与改写量正交**——它防的是这一种失灵，不是所有失灵。任何声称找到了万能判准的说法，都应当被怀疑，包括本文的。

其四，**这是单案例的机制研究**。它揭示一种可能的机制，不估计这种机制在全部学位论文中的发生率。而且它把本案当机缘不当证据——支持机制的材料在别处，本案只提供了一个足够清楚的谜题。

结语

表读数清白，不等于清白

回到那9处。

它们穿过了五道关口，不是因为哪一道失了职。恰恰相反：每一道都尽了职，而且用的是同一块表——那块表在跨过语言边界的那一刻，归了零。

这个解释不减轻任何人的责任。作者对署名文本的来源真实性负首要责任；**表读数清白，从来不是合规的证明**。但它改变了着力点：如果你的对策是再加一道关口，而那道关口看的还是同一块表，**防御力一点也不会增加**。签字会变多，维度不会变。

成熟的制度不是关口更多，是表更多。

所以，下一次面对一段漂亮的中文，值得问的不是"这改写得够不够"——一个好写手永远答得上来。

值得问的是：

你在哪一页读到它的？

参考文献

新华社. 中国人民大学决定撤销蒋方舟硕士学位[EB/OL]. (2026-07-13).

全国人民代表大会常务委员会. 中华人民共和国学位法[Z]. 2024-04-26.

中华人民共和国教育部. 高等学校预防与处理学术不端行为办法：教育部令第40号[Z]. 2016-06-16.

中华人民共和国教育部. 学位论文作假行为处理办法：教育部令第34号[Z]. 2012-11-13.

Howard, Rebecca Moore. "A Plagiarism Pentimento." Journal of Teaching Writing, 1993, 11(3): 233–246.（补丁写作的提出）

Howard, Rebecca Moore. "Plagiarisms, Authorships, and the Academic Death Penalty." College English, 1995, 57(7): 788–806.

Pecorari, Diane. "Good and Original: Plagiarism and Patchwriting in Academic Second-Language Writing." Journal of Second Language Writing, 2003, 12(4): 317–345.

Shi, Ling. "Textual Borrowing in Second-Language Writing." Written Communication, 2004, 21(2): 171–200.

Roig, Miguel. "Plagiarism and Paraphrasing Criteria of College and University Professors." Ethics & Behavior, 2001, 11(3): 307–323.

Potthast, Martin, Alberto Barrón-Cedeño, Benno Stein & Paolo Rosso. "Cross-Language Plagiarism Detection." Language Resources and Evaluation, 2011, 45(1): 45–62.

Barrón-Cedeño, Alberto. "On the Mono- and Cross-Language Detection of Text Re-Use and Plagiarism." PhD thesis, Universitat Politècnica de València, 2012.（本文所引 36.25% 的态度调查数据出处）

Olivia-Dumitrina, Nechita, Montserrat Casanovas & Yolanda Capdevila. "Academic Writing and the Internet: Cyber-Plagiarism amongst University Students." Journal of New Approaches in Educational Research, 2019, 8(2): 112–125.（本文所引 82.1% 的对立数据出处）

Reason, James. Human Error. Cambridge: Cambridge University Press, 1990.（瑞士奶酪模型的提出）

Thompson, Dennis F. "Moral Responsibility of Public Officials: The Problem of Many Hands." American Political Science Review, 1980, 74(4): 905–916.

Vaughan, Diane. The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA. Chicago: University of Chicago Press, 1996.

Edmondson, Amy C. "Learning from Mistakes Is Easier Said Than Done: Group and Organizational Influences on the Detection and Correction of Human Error." Journal of Applied Behavioral Science, 1996, 32(1): 5–28.

Deming, W. Edwards. Out of the Crisis. Cambridge, MA: MIT Press, 1986.

Strathern, Marilyn. "'Improving Ratings': Audit in the British University System." European Review, 1997, 5(3): 305–321.

· 全 文 完 ·