

文心探幽

卷一 何谓作文

WEN XIN TAN YOU

VOLUME ONE · WHAT IS COMPOSITION



凡直接读得出的，都不是它

它落在账本没有设的那一格 · 动一下才显形

付自文 著

德麦国际出版社

DEMAI INTERNATIONAL PRESS

文心探幽

卷一 何谓作文

付自文 著

德麦国际出版社

Demai International Press

版权页

书名：文心探幽·卷一 何谓作文

著者：付自文

出版：德麦国际出版社（Demai International Press），新加坡

国际标准书号 ISBN :979-8-90690-038-8

定价：US \$25.60

版次：2026 年 8 月第 1 版第 1 次印刷

开本：6×9 英寸 字数：约 20.0 万字

版权所有，侵权必究。未经著作权人书面许可，不得以任何方式复制或传播本书全部或部分内容。

章节来源说明

本卷九章由九份独立完成的研究稿编次而成，各稿均由著者撰写。各章正文一字未改（仅订正可核验的事实类错讹，并统一删去各稿原有的单篇署名行与投稿标记）；各章末的「编者补节」以及本书的前言、导读、结语与三则附录，为本卷新增。

本卷的母题、九章选目、四编编次、两处对切以及九节编者补节，由著者完成。

本卷各章的分界、证伪条款与自评均由著者单方作出，未经外部裁决。按第九章第八命题，判者集合包含于被判者集合时，其裁决不具终局性；因此本卷不对自身作出任何成立性宣告，各章的创新性评价须由独立第三方给出。

作者介绍

付自文，2019 年起随王德生博士研习，研究方向为 SDE 本体论与多学科碰撞创新。著有《文心探幽》丛书；卷一《何谓作文》（德麦国际专著第 91 号）、卷三《作文的意义是什么？》（第 90 号）已由德麦国际出版社出版。

自述

2019 年，我开始跟随王德生博士学习。此后的几年，我也有幸沿着他的思想发展一路前行：从「三视角智慧」到「321 智慧」，从「主客互动本体论」到「SIO 本体论」，再到今天不断展开的「SDE 本体论」。对我而言，这并不是简单地学习一套已经完成的理论，而是亲历一种思想如何从问题中萌发，在不断追问中突破自身，并逐渐生长出新的理论世界。

也正是在这一过程中，「发生」慢慢成为我理解世界的一种基本方式。我开始重新审视语言、教育、写作、学习、价值以及人的精神生活，并不断追问：那些看似理所当然存在的事物，究竟是怎样一步一步成为今天这个样子的？

近年来，我主要围绕 SDE 本体论及多学科碰撞创新展开研究，希望沿着这条仍在生长的思想道路，继续发现新的问题、新的连接与新的可能。

——付自文

关于本卷的写法

本卷九章有一个固定的做法：不从语文教育与写作研究内部取理论资源，而是从与本题距离足够远、且其核心判据尚未被写作理论吸收的学科取源，让三家互相拆台，再看它们共同假定了什么。九章共动用二十七门与语文教育无关的学问——从凝聚态物理的相变、范畴论的态射剖面、射电干涉的闭合相位，到公司治理的指标学习、保护生物学的功能恢复、民事诉讼法的既判力。

在成书纪律上有三条：每章必须给出一个可清点的读数，而不是一个形容词；每章必须写下能让自己死亡的观察；跑出来的不利结果原样保留，并据此改判。本卷九章里六章动用了公开数据或公开材料，六章全部得到对自己不利的结果，其中三章据此当场收窄主张，一章撤回了自己一处定理的证明记号。

前言

一、这本书为什么现在写

我最初想问的不是「作文是什么」，而是一个更小的问题：为什么两篇一模一样的文字，我会觉得其中一篇是这个学生写的，另一篇不是。

这个直觉在过去不需要解释，因为它几乎从不被触发。一个学生从构思到誊清都亲手完成，写下这些字的人、经历这场写作的人、为这份成果负责的人，是同一个。三者重叠得如此彻底，我们可以把「学生交来的文章」直接叫作「学生的作文」，而不必追问这个等号凭什么成立。

生成式人工智能把这个等号拆开了。拆开之后我才发现一件更难堪的事：我们手里所有用来判定作文的账本，都是在三者重叠的年代设计的。评分表记录成品特征，过程档案记录写作动作，学术诚信条款记录工具使用比例——每一样都在记「在场的东西」。可是当我试着用它们去分辨我那个直觉时，它们一样也用不上。

这本书就是从这个用不上开始的。

二、九章是怎么变成一本书的

这九章不是为了写一本书而写的。它们是九次独立的追问，各自从三门远学科取源，各自提出一个新对象，各自给一个读数。写完之后我把它们摊在桌上，问的第一个问题不是「怎么排」，而是「它们凭什么在一起」。

答案花了一段时间才逼出来。九章表面上处理的是九件事——同一性、整体性、发生时刻、结算权、从未成形的形式、被关闭的分叉、无归属的段落、判断的余量、残余物的名分。把它们各自的判断卡片摆开，共绕的动作只有两族：不入账，和扰动才显形。九章全部同时落在这两族上。

于是有了那句母题：

作文不在成品上：凡直接读得出的都不是它；它落在账本没有设的那一格，动一下才显形。

我要说清楚这句话的地位。它不是九章的公因数，也不是事后总结——它是一条可以被推翻的判断。若能证明成品特征或过程日志可以稳定预测作文身份与后续迁移，即扰动读数与静态读数长期一一对应，那么这句话失去内容，九章会当场散成九篇互不相干的论文。这本书的脊梁只有这一根，我不打算给它加保险。

三、我为什么不从写作研究内部取材

有人会问：写作研究已经积累了几十年，为什么不从它内部继续推进？

理由不是那些理论不好，恰恰相反——它们已经太好了，好到把这个问题的可动空间占满了。过程写作、修辞情境、体裁社会行动、后过程与网络写作、篇章连贯与修辞结构理论，每一支都对「作文是什么」给过完整的回答。从它们内部再往前挪一步，得到的多半是第四种排列组合，而不是一个新对象。

所以九章一律从远处取源，而且取源之前先过两道闸。

第一道闸问：这门远学科对本题的答案，是不是已经被写作理论整段吸收了？这道闸的价值我是靠一次失败学会的。选源时我一度想用干扰生态学的演替理论——它距离语文教育足够远，看上去很安全。查下去才发现，生态学对写作的那个回答早在一九八六年就已经被写作研究整段拿走，并且长成了一个有专名的子领域。「离得越远越安全」是错觉。远近不是安全的判据，被吸收与否才是。

第二道闸问：这门学科对本题的判断，厚不厚？厚的意思是它内部有长期争论、有反例传统、有权威手册或标准。一门学科如果对这个问题的说法，它给的就只是一个比喻，不是一个结构。凝聚态物理的相变与序参量、地层学的边界与缺失记录、范畴论的态射剖面、免疫学从克隆选择到危险模型再到连续性理论的三派之争——这些是厚的；它们能给的不止一句话，而是一整套判据，连同判据失效的条件。

两道闸过完，三家进场，让它们互相拆台而不是互相印证。三家如果彼此支持，那多半说明它们其实在说同一件事，换了三种措辞。真正有用的是三家各自用自己的材料伤到另外两家，而它们伤不到的那个共同前提——那才是要找的东西。九章里每一章的转折点都在这里：三家争的其实不是同一个答案，而是谁有资格结算。

四、每章末为什么还要接一节编者补节

九章原稿写完之后，我做了一件在自己看来最不舒服的事：给每一章找它漏掉的祖宗。

结果不好看。第四章提出的两个新对象——文章尚未显形而作文已经发生、文章已经完成而作文没有发生——在教育心理学里有一位几乎逐字对应的祖宗，而且是现役的综述；那篇文章的两条结论恰好就是我那两个新对象，连它规定的测法都与我的中性扰动检验是同一件事。第二章的保相与破相变换，与自然语言处理里一套二〇二〇年的行为测试方法同形，连「事先声明预期方向、再看是否兑现」这条纪律都相同。第三章那条最得意的发现——不读作文、只看操作标签就能预测七成结果——是二〇一八年就被命名过的一类失败，有成熟的对策清单和报告规范。

这些我都补进去了，每一处配一条判决性对照，说明它为什么不取消本章。但我要在这里说清楚补入的后果：

补上祖宗之后，好几章的主张变窄了。第四章原本说「即使过程数据丰富，仍需区分已发生的动作与只能在环境改变后显露的状态」——这句话在补上祖宗之后站不住，因为它正是那一支几十年来的标准立场。改写之后剩下的是一句带三个限定的话：在单次任务之内、在成品尚未分化之时、由一次不改变题目与证据域与受众的扰动读出。三个限定缺一不可。

主张变窄，可检验性上升。我认为这是这本书里最值得学的一种修改方式，所以把它保留在补节里，而不是悄悄改掉正文再假装原本就是这样。

五、被删掉的那些，和为什么删

同题的稿子原有十三份。这本书用了九份。

删掉的四份里，有两份是质量问题：一份提出了八个新术语而没有一个有读数，一份连参考文献都没有。这两份的判断我不认为有争议。

另外两份是调走，不是砍掉。它们分别问「何谓好作文」与「作文中的内容究竟是什么」，分数在全批前列，但问的不是「作文是什么」。硬塞进来会稀释母题——一本书如果同时回答三个问题，读者会正确地怀疑它一个也没回答。这两份留给卷二。

我提这件事，是因为它涉及一条我认为比选目本身更重要的判断：一本书的边界不由材料的质量划定，由它要回答的那一个问题划定。

六、我知道这本书最可能错在哪

有三处，按我自己的估计从重到轻。

第一处，九章几乎不共享文献。合并全书参考文献时，166 条里有 160 条只出现在一章之中，被两章以上共同引用的只有 6 条。这是取源方式的直接后果——各章从远学科取源，各自带回一整套本领域读者陌生的文献，这正是新对象得以产生的条件。但它同时意味着：九章之间的可比性，目前几乎完全依赖母题与两处对切，而不依赖任何共同的地基。如果母题被推翻，这本书不是变弱，是解体。

第二处，读数多而实测少。九章给了九个读数，全部附有清点规则。但真正跑出数值的只有一部分，另有几章的读数至今一个数也没有算过。第二章的篇章定相率在现有公开语料上根本算不出来（字段不存在），第四章的自续率的最小实验一所学校就能跑完却尚未跑，第七章的留叉率只需要一批真实作文任务与两名编码者一个下午——最便宜的那几条恰恰没做。这不是谦辞，是待办。

第三处，我自己就在被自己判的位置上。这本书的分界、自评与编次全部由我方作出，未经任何外部裁决。按第九章第八命题，判者集合包含于被判者集合时，其裁决不具终局性。所以这本书不对自己作任何「已经成立」的宣告，也没有资格宣告。我能做的只有把九处空格、九个读数和它们各自的死亡条件摆在那里。

第四处，我不确定「不入账」这个说法本身会不会变成新的万能词。这本书反复使用「账本上没有那一栏」这个表达。它现在是清楚的，因为每一章都给了具体的那一栏是什么、为什么没有、以及如果设了会发生什么。但一个表达一旦好用，就会开始吸附一切说不清楚的东西——凡是解释不了的，都说成「它不入账」。这本书没有能力防止这件事，我只能在这里写明：「不入账」若不能同时说出「哪一栏」「谁本该设」「设了会怎样」，它就是一句废话。读到任何一处只有前半句的地方，请当作是我的疏忽。

七、一件我原以为是缺点的事

写这九章的过程里，最难受的一段是发现自己的主假设不成立。

第六章原本要用「缺口清单」估计从未成形的那些形式，预注册写死了 30% 的阈值，实测 15.0%，三种算法一致不过线。第三章原本要证明作文是一种「内部能判断偏差」的文本，结果它自己的数据显示，那个高达 0.972 的覆盖率里有

0.875 只需要改动句加一条邻句就能还原——更糟的是，不读作文、只看施加了哪一类扰动，就能预测七成以上的结果。第九章的三条预注册假设全部失败。

这些结果我全部留在书里，并且据此改判：第六章关掉了原定的一条估计路径，第三章撤回了自己的肯定性定义并重建，第九章把一处定理降格为命题。

我起初把这些当作这本书的弱点。后来我改了看法，理由很简单：一篇在数据面前一次也没输过的文章，多半是没有真的下注。九章里六章下了注，六章都输了一部分——而每一次输，主张都变窄了一点，可检验性都上升了一点。这本书里我最有把握的部分，恰恰是这几处输掉的地方。

六、关于那六次失败，还有一句

有人会说：既然六次都输了，这本书凭什么还成立？

因为输的和立的不是同一类东西。六次真跑打掉的，全部是某个具体读数在某种具体条件下的可用性——用缺口清单估计未生界这条路不通，用现成评分语料算篇章定相率算不出来，用局部扰动读整体性会被近邻可还原污染。而九个新对象本身，一个也没有被这六次真跑打掉，因为它们都还没有被真正测过。

这句话对这本书不利，我照写：这本书目前的强项在方法与判据，不在经验证据。九个读数里真正跑出数值的是少数，其中还有一部分跑出来的是「算不出」。如果有人因此判定它是一本设计论文而不是一本实证著作，我接受这个判断——它是准确的。

我不接受的只有一种读法：把六次不利结果读成六次失误。它们不是失误，是六次下注的结算。没下注的那三章，才是这本书真正薄的地方，而它们薄在哪里，编者补节里逐章标了出来。

八、这本书不主张什么

它不评价文学价值、思想高度与审美质量。它不提供任何可以直接用于考试排名的指标——第四、五、九章分别论证了为什么把这些读数变成高风险指标会立刻摧毁它们所测的东西。它也不主张作文教育应当增加监控；多数章节的建议恰恰是少记录，但换一个地方记录。

最后一句留给读这本书的老师。如果你从这九章里只带走一件事，我希望是这一件：当你说一个学生「没有自己思考」时，其中有多少是因为他没有思考，有多少是因为到他手上时，已经没有什么可思考的了。这一问在第七章有一个可以算的读数。它不难跑，一个下午就够。

九、写给谁

写给三种人。

第一种是教作文的人。这本书不给教案，但它给了两个明天就能用的东西：第七章的留叉率——把一次作业连同它的全部随附材料摊开，数一数学生拿到它时还剩几个结点是开着的；第五章的最小写成记录——三类高杠杆节点各一个，每个节点只记三件事。两样都不需要新增设备，也不需要多留档案，只需要换一个地方看。

第二种是研究写作的人。这本书最想被人拿走的不是九个新对象，是九份清点手册与九组死亡条件。九个读数里最便宜的三条——第七章的留叉率、第四章

的自续率、第九章的准清偿密度——加起来的成本低于一次常规问卷调查，而它们各自都能把这本书的某一章判死。能把它判死的人，比同意它的人对它更有用。

第三种是正在写作文的学生。这本书里有一句话是对着他们说的：一篇文章可以还没有出现，而作文已经发生；一篇文章也可以已经完成，而作文还没有发生。坐在一张空白的纸前，什么都还没写出来的那段时间，不是没有发生任何事——那可能正是唯一真正在发生的时候。

十、一句致谢，与一句声明

九章的写作跨越数月，其间数次因数据不利而推倒重来。要谢的是那几次不利结果本身——它们比任何一次顺利的验证都更早地告诉我，哪一句话是我想说的，哪一句话是站得住的。

最后一句声明，重复一遍全书末尾会再说一次的话：这本书的分界、自评与编次全部由我单方作出，未经外部裁决。按第九章第八命题，判者集合包含于被判者集合时，其裁决不具终局性。因此本卷不对自己作任何「已经成立」的宣告。

导读

这本书的形状

四编九章，编次不按学科排，按「读不出的那一部分落在哪里」排。

编一处理同一性：这一篇凭什么是这一篇。三章共用一个动作——先破坏，再看什么还在；差别在扰动谁（人／文本／局部）。

编二处理发生：什么时候它成为一篇作文。两章分别问它长在哪身上、它是谁决定的。

编三是账外三格，全书的心脏：无形（从未获准成形）、无据（已写出而两端不记账）、无名的动作面（那条路是谁关的）。

终编处理制度：为什么这些空缺不会因为被诊断而被填上。

每章末接一节编者补节，只做三件事——补上原稿漏列的最近祖宗并给判决性对照、写下与本卷他章的分界、标明该章的检验做到了哪一档。补节与正文冲突时以正文为准，补节只留下地址。

三种读法

要判断这本书值不值得读：读导论第四节（母题）、第六章、第九章。第六章是账外三格里最难的一格，第九章解释为什么这些格子空着。两章合起来大约四万字，读完就能判断这本书的赌注下在哪里。

要拿它去用：直接进第七章的留叉率与第四章的自续率。这两个读数最便宜——前者只需要一批真实作文任务与两名编码者，后者一所学校四个平行班就能跑完，而且都附有完整的清点规则。第五章的最小写成记录（三类节点各一个、每节点三件事）可以明天就在班上开始。

要挑它的毛病：从落地页那一节「怎样最快地推翻它」进，或直接读各章的证伪清单与两处对切的合并条件。合并条件是写死的，不是事后可改的——若第一章与第二章的读数长期一一对应，两章之一冗余；若每一处未生界都能回溯到可指名的关闭者，第六章应并入第七章。这些条件我自己指出来，就是为了让别人能据此把书拆掉。

读的时候请留意三件事

其一，读数不是分数。九个读数没有一个是用来给学生排名的。第四章第二十五之二节写了一条比「不要用于考核」更硬的规则：凡是被试事先知道会被要求重构的场合，自续率读数一律不成立——它规定的是读数的成立条件，不依赖使用者的自觉。

其二，不利结果不是败笔。六章的真跑全部得到对自己不利的结果并原样保留。请把它们读成这本书最结实的部分，而不是最薄弱的部分。但也请分开两件事：照实报告不利结果，与跑出了可用的读数，是两回事。

其三，这本书不给结论。它给九处空格、九个读数、九组死亡条件，然后停在那里。终编解释了为什么它没有资格再往前走一步。

目 录

作者介绍

前言

导读

导论 这一篇凭什么是这一篇

编一 这一篇凭什么是这一篇

第一章 撤掉成品之后，还剩什么

第二章 不动它，就不知道它是不是它

第三章 读得出的那一部分，恰恰不是整体性

编二 什么时候它成为一篇作文

第四章 文章还没出现，作文已经发生

第五章 高杠杆的那几处，是谁结的账

编三 账外三格：无形·无据·无名

第六章 在第一笔之前

第七章 路是谁关的

第八章 两端都不记账的那一段

终编 制度这一侧

第九章 灰为什么没有名分

（每章末接一节编者补节）

结语 摆在这里，等一条通道

全书文献总表

附录一 母题是怎么打出来的（全过程留档）

附录二 九个读数、九组死亡条件（一览）

附录三 六章真跑，六章不利（清单）

编一

这一篇凭什么是这一篇

编一 这一篇凭什么是这一篇

三章共用一个动作：先破坏，再看什么还在。

这个动作不是修辞。它是这一编从三门远学科借来的同一个方法论内核——判断一个对象的身份，不能只看它此刻的样子，要看它在受到规定好的扰动之后如何响应。晶体学判断多形，靠的是同样的组成在不同条件下是否落入不同的结构态；范畴论把对象的身份从「内部属性清单」推进到它与其他一切对象之间的完整关系剖面；理论密码学区分「能给出正确输出」与「拥有使输出成立的东西」，靠的是让对方在原答案不可调用时重新应答。

三章的差别在于扰动谁。第一章扰动人：封存原稿、禁用原会话，看写作者能否重新组织同一批意义关系。第二章扰动文本：在声明的判别分辨率下施加预注册的保相与破相变换，看宏观组织关系是否跃迁。第三章扰动局部：对同一篇文章施加六类最小改动，追问全文的判断能力究竟有多少是真的、有多少只是近邻可还原。

第三章的位置需要特别说明。它本来打算证明作文是一种「内部能判断偏差」的文本，结果被自己的数据淘汰——那个高达 0.9722 的覆盖率，绝大部分只需要改动句和一条邻句就能得到，而更糟的是，不读作文、只看施加了哪一类扰动，就能预测七成以上的判断结果。这篇文章因此把自己的肯定性定义收回，改提一个更窄、更难伪装的假说。它被放在这一编的末尾，不是因为它最弱，而是因为它给出了全卷方法论的边界：扰动法本身也会产生假象，而唯一的解药是事先写下淘汰线并照着执行。

第一章

撤掉成品之后，还剩什么

何谓作文

——生成见证视角下的作文等价类理论

What Is Composition?

A Generative-Witness Account of Composition as an Equivalence Class

学科通融典范论文 | V1.0

2026年8月17日

摘要

“作文是什么”通常被视为一个无需重新发问的问题：它可以被理解为一篇完成的书面文本、一次写作过程、作者思想的语言表达、特定情境中的交往行为，或某种可被测评的写作能力表现。然而，生成式人工智能、协同写作、语音输入、自动反馈与模板化训练正在解除过去长期重叠的几个事实：文本由谁输入、意义由谁组织、过程由谁控制、最终作品由谁承担责任，以及某种生成能力究竟存在于谁身上。由此出现一个更基础的对象身份问题：一篇现实存在的文章，凭什么被称为“这个人的作文”？本文借助法庭文件检验中的来源归属问题、物理有机化学中反应序列与反应机理的严格区分，以及理论密码学中“成功输出”与“知识拥有”之间的区分，提出“生成见证”与“作文等价类”两个互相依赖的概念。文章被定义为一次已经实现的文本；生成见证被定义为在原文不可调用、条件受到预注册扰动时，仍能组织新文本中关键意义关系，并对留出挑战具有预测力的最小生成结构；作文则被定义为在给定问题域与允许扰动集合内，由同一可提取生成见证所产生的一族可能文本，即文本在生成见证关系下的等价类。由此可以得到两个传统定义难以同时容纳的判断：“同文可以异作”，“异文可以同作”。本文进一步提出“文本相似度 $\times$ 生成见证等价性”的二维辨别框架，明确其与写作过程理论、写作迁移、作者身份、创作控制、测评效度、风格归属以及学习保留等近邻概念的边界，并给出可执行的多轮挑战—提取—留出验证方案、竞争解释、机制证伪条件与伦理边界。公开研究已显示，外部反馈条件下的成品改善与辅助撤除后的独立表现可以分离，但现有数据通常缺少“多文本是否共享同一生成见证”的字段，因此本文的核心构念仍需专门实验验证。本文的中心主张是：文章是一份实现，作文不是这份实现本身，而是使多种可能实现具有同一生成身份的关系空间。

关键词：作文；文章；生成见证；作文等价类；写作发生；人工智能写作；作者归属

Abstract

The question “What is composition?” is often treated as settled: composition may be understood as a finished text, a writing process, linguistic expression of thought, situated communication, or evidence of writing ability. Generative AI and other writing technologies now decouple activities that were historically bundled together—typing, ideation, organization, revision, control, responsibility, and the possession of generative competence. This paper therefore asks a more basic identity question: under what conditions may an existing text count as this person’s composition? Drawing on source attribution in forensic document examination, the distinction between reaction sequence and reaction mechanism in physical organic chemistry, and the distinction between successful output and possession of knowledge in interactive proofs, the paper develops two linked constructs: the generative witness and the

composition equivalence class. A text is a realized artifact. A generative witness is the minimal relational structure that can still organize key meaning-relations in novel responses when the original text is unavailable and task conditions are subjected to preregistered perturbations, and that predicts a held-out challenge. A composition is then the equivalence class of possible texts generated by the same extractable witness within a specified problem domain and perturbation set. This account yields two counterintuitive but testable claims: identical texts may instantiate different compositions, while dissimilar texts may instantiate the same composition. The paper distinguishes this proposal from writing-process theory, writing transfer, authorship attribution, creative control, assessment validity, and retention, and offers an operational challenge-extraction-holdout protocol, falsification conditions, and ethical limits. Existing studies already show that improvement under external writing support can dissociate from independent performance after support is removed, but current datasets generally lack the field needed to test generative-witness equivalence directly. The central claim is therefore ontological rather than merely evaluative: a text is one realization; a composition is the structured space that makes multiple realizations generatively identical.

Keywords: composition; text; generative witness; equivalence class; writing process; AI-assisted writing

一、问题的重启：作文为什么需要重新定义

在学校教育中，“作文”几乎是最熟悉的学习对象之一。题目被布置，学生写出若干段文字，教师批改、打分，文本被保存或丢弃。正因为这套制度已经运行得太久，“作文是什么”反而很少被当作需要研究的问题。我们更习惯问“怎样写好作文”“为什么学生不会写作文”“AI 能不能帮助作文”，却默认知道那个被写、被教、被评的对象究竟是什么。

这种默认在生成式人工智能出现以后迅速失效。过去，写字的人、组织思想的人、修改文章的人、提交文章的人以及承担作者责任的人往往是同一个人，因此“这篇文字是谁写的”能够近似替代“这是谁的作文”。今天，同一篇文本可以由学生口述、语音系统转写、AI 扩写、同伴编辑、教师反馈与学生最终审定共同形成；也可以由学生亲手逐字输入，却几乎没有发生任何属于学生自己的意义组织。技术并没有凭空制造本体问题，它只是把过去被纸笔条件捆绑在一起的环节拆开了。

于是必须重新提出一个比“是否允许 AI”更早的问题：一篇已经完成并能够被阅读的文章，在什么条件下，才有资格被称作“这个主体的一次作文”？如果这

个身份条件不清楚，我们就无法稳定回答何谓代写、何谓辅助、何谓学习、何谓作者贡献，更无法解释为什么有时一篇高度 **polished** 的文章并不能证明写作者真正获得了相应的作文能力。

本文把问题限定在“作文的对象身份”而不是“作文的价值评价”。一篇作文可以写得差，也可以写得好；可以原创，也可以引用；可以人工完成，也可以在工具帮助下完成。本文不先问它是否优秀，而先问它究竟是什么，以及两个表面不同或相同的文本在什么条件下应被视为同一个作文对象。这个问题一旦被区分出来，传统的文本观、过程观、能力观、情境观都需要重新接受边界检验。

二、现有写作理论已经解决了什么，又留下了什么

当代写作研究早已不是“文章成品论”。Flower 与 Hayes 在 1981 年的经典研究中把写作描述为目标导向、递归而非机械线性的认知过程，写作者会在规划、文本生成和检查之间反复往返[1]。之后的过程研究继续细化暂停、修订、版本演化以及写作者在不同阶段的认知活动。2023 年的“writing cloud”研究甚至使用版本化软件保存一篇科学随笔从最早尝试到最终成品的完整演化，量化规划与转译的非线性子循环[2]。因此，简单宣布“作文不是结果，而是过程”，已经不可能构成新的理论贡献。

写作研究也早已超出个体头脑。体裁、受众、社会文化情境、身份、制度和评价机制都被纳入作文研究。写作迁移更是形成了成熟的跨学科传统。Nowacek、Leonard 和 Rounsaville 在 2024 年的专著中综合约七百部/篇有关 transfer 的文献，讨论写作知识、策略、行动者与新情境之间复杂的迁移关系[3]。因此，简单说“真正的作文能力应当能迁移到新任务”同样会被现有理论吸收。

测评学也没有把一篇作文分数直接当作写作能力本身。Kane 的论证式效度框架明确要求，把从测试反应到分数解释和使用的每一步推断、假设和证据展开检验[4]。这意味着“高分作文不必然等于高写作能力”并不是本文能够独占的新洞见。

与此同时，AI 辅助写作研究正在快速研究作者控制、所有权感、认知参与与责任。Arnold 与 Kim 提出需要交互的生成建议，以增加人在共写过程中的参与、控制与 awareness[5]；Formosa 等人的实验显示，人们对 human authorship、creatorship 和 responsibility 的判断会受到辅助程度影响，而不只是受到辅助者究竟是人还是 AI 影响[6]。2026 年的 Apt Curation Model 又进

一步把 AI 辅助作者的核心贡献放在架构、筛选、校验、整合和认识责任上[7]。所以，“谁做了选择”“谁最终控制”“谁愿意负责”也已经是成熟近邻。

这些理论共同说明一个重要事实：真正的新问题不能靠否定已有写作研究获得。我们必须承认，文本、过程、迁移、情境、控制、责任和效度都已经被认真研究。剩下的问题反而更窄：这些研究大多把一篇已发生的文本或一次已发生的写作事件作为基本单位，而本文要问的是，这个基本单位本身是否选对了。

三、三个极端案例：文本、过程与作者为什么仍然不能互相替代

先考虑三个极端案例。第一种是“同文不同生”。甲独立形成一个判断，经过删改完成八百字；乙直接取得甲的全文并逐字抄写。最终字符串完全一致，甚至乙的物理笔迹全部属于自己。如果作文等于文本，两者没有区别；如果作文等于书写动作，乙反而拥有完整“写”的过程。可是我们仍然有充分理由说，两人的作文身份不同。

第二种是“异文同生”。一个身体障碍的学生不能直接操作键盘，他通过口述表达论证，语音识别完成字符输入。第二天换一种设备、换一种受众，他又用完全不同的措辞重组相同的意义关系。文本表面差异巨大，物理输入主体也不同，但我们很难因此否认两次写作都属于他。由此可见，字符执行权不是作文归属的必要条件。

第三种是“高参与、低内化”。学生把题目交给 AI，AI 给出多个框架，他认真比较、删除、重新排序、修改句子并承担最终署名。从 authorial agency 或 curation 角度看，他并非毫无参与。然而如果拿走原文本与对话记录，给一个结构同型但表面不同的问题，他无法重新形成那篇文章最关键的关系，只能再次依赖外部生成。此时，他可能依然是这份合作文本的责任作者，却未必已经拥有使该文本成立的生成结构。

三个案例共同打断了一条长期被默认的等式：文本归属、过程归属、作者归属、能力归属并不必然重合。真正需要寻找的是一个能够解释这些分离现象、又不退化为“迁移”“控制”或“效度”的对象。

四、痕迹能够归属来源，但痕迹不是来源本身

法庭文件检验提供了第一种严格约束。NIST 关于法庭笔迹检验的人因报告指出，经过训练的文件检验人员会分析手写笔记或签名的特征，以判断真伪并形成关于可能书写者的意见，同时该实践必须正视人因偏差、错误、质量控制和结论

表达的不确定性[8]。这个领域最重要的启发不是“每个人字迹不同”，而是把“痕迹”与“来源归属证据”严格区分。

一份完成态能够携带来源信息，但来源不能被直接等同于任何单个表面特征。同一书写者存在自然变异，不同工具、速度、身体状态、模仿与伪装都可能改变显露；反过来，不同来源也可能产生局部相似。因而，法庭文件检验必须把个别痕迹放进比较样本和变异范围中解释，而不是把单一痕迹当作来源本身。

迁入作文问题后，这一约束首先否定了“风格就是作文身份”。AI可以模仿风格，学生可以学习模板，同一成熟写作者也可以在议论文、邮件、故事和学术论文中使用完全不同的表面特征。文本痕迹当然重要，但它最多说明“这一实现与某来源之间存在怎样的证据关系”，不能直接回答“使这篇作文发生的结构究竟存在于谁身上”。

因此，第一个原则是：作文身份不能从单个完成文本的表面相似性中直接读取。任何理论如果最终仍把词汇、句长、修辞偏好、文体指纹当作作文同一性的核心，就会把“文章是谁写得像”误认为“作文由谁发生”。

五、过程能够被完整记录，但序列仍然不是机理

物理有机化学提供第二种、更深的约束。IUPAC把reaction mechanism定义为对反应物如何转变为产物的详细描述，需要尽可能涉及中间体、过渡态的组成、结构、能量等，并且必须与化学计量、速率律、立体化学过程和其他实验数据一致；同一反应甚至可能存在多个尚未被证据排除的备择机理。IUPAC还特别区分reaction mechanism与reaction sequence：仅仅描述一组逐步反应的可能顺序，并不足以称为机理[9]。

这一区分对今天的写作技术极具破坏性。我们已经能够记录学生在第几秒输入哪个字、何时删除、何时粘贴、何时调用AI、何时接受建议。一个完整日志可以令人产生“我们已经看见写作过程”的错觉。但是，日志所保存的首先仍是一条actual sequence：A发生，然后B发生，然后C发生。它没有自动说明，为什么A会转向B，哪一个条件是必要条件，哪一个动作只是伴随现象，换一个条件是否会出现另一条路径。

例如，一个学生先写“AI让人变懒”，之后删掉，改成“AI会使判断外包”，最后又写成“真正风险是生成能力没有形成”。过程日志能记录这三句的出现顺序，却不能仅凭顺序判断：促成第三个判断的是教师反馈、AI提示、学生自己发现的反例、还是此前阅读形成的潜在模型。不同机理可以产生相同终态，同一机理也可以通过不同sequence产生不同终态。

因此，第二个原则是：作文不能被定义为一次已经发生的过程序列。过程数据是重要证据，但要识别作文身份，必须能够在竞争机理之间进行判别。换言之，理论必须回答“如果真是这套生成关系导致当前文本，那么改变某个局部条件以后，还应该观察到什么”；没有反事实预测的过程描述仍可能只是事后故事。

六、成功产出与“拥有使产出成立的东西”之间还有一道鸿沟

理论密码学提供第三个约束。Goldwasser、Micali 和 Rackoff 建立交互式证明与知识复杂度理论时，核心问题之一就是区分“一个命题被验证者接受”与“证明者究竟泄露或拥有何种知识”[10]。Bellare 与 Goldreich 随后专门讨论 **proof of knowledge** 的严格定义，因为仅凭一次成功交互，不能松散地宣称证明者“拥有知识”[11]。

知识证明思想中一个极具启发性的结构是“提取”。直觉上，如果一个证明者能够以足够概率使验证者接受，那么理论希望存在某种知识提取器，通过对证明者进行受控交互、重复运行或改变挑战，能够提取一个使其成功具有解释力的 **witness**。这里真正重要的不是把密码学形式定义直接照搬到教育，而是其中的判别逻辑：一次成功输出不能自己证明隐藏能力；要证明“拥有”，必须让隐藏结构在新的挑战中继续约束响应。

把这一逻辑放回作文，一个学生交出高分文章，只能直接证明“一份高分文本现在存在”。如果要进一步说“这篇作文已经在这个学生身上发生”，至少需要额外证据：拿走当前代表元以后，在事先限定的条件扰动中，是否还能从同一主体的响应里提取出那套组织关键意义关系的结构。

这与测评效度相邻但不相同。效度理论检验从测试表现到能力解释的推断是否可辩护；本文进一步改变对象单位：也许我们一直把一篇单独文本当作“作文”，而真正需要识别的作文对象本来就不是单一实现。

七、被忽略的共同前提：现实中发生的那一条路天然拥有本体优先权

法庭文件检验从已存在的痕迹出发，反应机理学从已经发生的反应和实验结果出发，交互式证明从一次实际 **protocol** 交互出发。三者方法不同，却共享一个看似不可避免的现实起点：研究首先面对已经发生的東西。我们因此很容易把 **actual event** 当作对象本身，而把没有发生的替代路径当作背景。

然而，三门学科自己的成熟方法恰恰不断突破这个起点。法庭检验需要多个已知样本来估计来源内部变异；反应机理研究通过改变温度、溶剂、浓度、同位素等条件来区分竞争机理；知识证明通过不同 **challenge** 观察证明者能否持续产生满足约束的响应。它们共同说明：一个隐藏来源或隐藏机理的身份，并不完全存在于第一次现实结果中，而存在于它面对一组允许扰动时所形成的“响应族”。

作文研究因此获得一个新的问题空间。现实中的文章 **T** 只是一条实际实现。围绕同一核心问题，写作者原本还可能面对不同受众、不同材料、不同反例、不同表达限制、不同辅助工具，并产生 **T1**、**T2**、**T3**.....。这些未实际发生的文本不是简单的想象垃圾；如果它们受到同一套生成结构约束，它们共同定义了“这一次作文究竟是什么”的身份范围。

这意味着，作文的本体可能不应放在一个点上，而应放在一个受约束的可能空间上。当前文本是观察窗口，不是全部对象。

八、生成见证：不是中心思想，不是模板，也不是人格标签

为了描述这个可能空间，本文提出“生成见证”概念。生成见证不是一篇文章的中心思想。例如“坚持很重要”“科技有利有弊”可以出现在无数文本中，但它们几乎没有能力预测新的组织选择。生成见证也不是固定模板。模板规定“第一段引题、第二段举例、第三段升华”，它可以跨题复用，却未必包含任何针对具体问题的意义生成。

本文把生成见证定义为：在原文不可调用、外部脚手架受到控制、任务条件发生预注册的有限扰动时，仍能组织新文本中至少若干关键意义关系，并对未参与提取的留出挑战具有预测力的最小关系结构。

这里有四个限定。第一，“关系”而非句子。生成见证必须能够说明什么与什么如何相关，例如“过早外部闭合会削弱内部生成”比“**AI** 有风险”具有更高生成力。第二，“最小”。如果为了保持稳定，我们不得不把见证写成“这个人重视成长、喜欢思考、看问题全面”，概念就已经失去判别力。第三，“留出预测”。研究者不能看完所有文本后事后总结一个永远正确的故事；至少要用部分响应提取 **G**，再预测一个没有参与提取的响应。第四，“问题域限定”。任何生成见证都只在一个明确的问题族与扰动范围内成立，不能被夸张为对主体全部思想或人格的还原。

因此，生成见证的功能不是宣判“这是一个真正有思想的人”，而是给出一种较弱但可证伪的说法：对于当前问题族，这个主体的多次文本响应中是否存在一套稳定的关系结构，其稳定性超出了表面文体、记忆复述和单次策略适配。

九、作文等价类：把作文从“一个文本”改写为“一族生成上同一的文本”

在生成见证基础上，本文给出核心定义。设写作者为 W ，核心问题为 Q ，允许的扰动集合为 $\Delta = \{\delta_1, \delta_2, \dots, \delta_n\}$ 。写作者在某一扰动下产生文本 $T(W, Q, \delta)$ 。如果存在可提取生成见证 G ，使两个文本虽然在词汇、例证、句式甚至局部结构上不同，却都受到 G 的关键关系约束，并且 G 对留出挑战仍有预测力，则记为 $T_i \sim_G T_j$ 。

由关系 $\sim_G$ 形成的等价类 $[T]_G = \{T' \mid T' \sim_G T\}$ ，就是本文所说的“作文”。一篇实际交出的文章 T 只是这个等价类的一个代表元。更简洁地说：文章是一份实现，作文是使多份实现具有同一生成身份的可能空间。

这个定义刻意没有把“优秀”写进去。一个很差的写作者也可能拥有清晰稳定但水平有限的生成见证；一篇语言精美的代表元也可能并不存在可提取的主体生成见证。作文身份与作文质量由此被分成两个维度。

这个定义也不要求第二篇与第一篇相似。恰恰因为来源存在自然变异，同一生成见证完全可能产生不同措辞、不同文体甚至相反的表面论证策略。等价判定要寻找的是关键关系是否同源，而不是文本是否像复制品。

如果这个定义成立，作文的基本单位将发生变化：学校现在保存“学生 $\times$ 题目 $\times$ 一篇文本”，未来更合理的数据结构可能是“学生 $\times$ 问题族 $\times$ 多次扰动响应 $\times$ 生成见证候选 $\times$ 留出验证”。一篇文章依旧重要，但它不再独占作文的本体地位。

（一）为什么必须叫“等价类”，而不能只说“生成空间”

“等价类”不是为了给作文披上一层数学外衣。它承担一个决定性任务：把“作文的同一性”从单个实例移动到实例之间的关系。若只说作文是一种“生成空间”，概念很容易变成泛化隐喻，因为任何人的任何能力都可以被说成一个空间。等价类则要求研究者明确一种关系，并回答这种关系在什么条件下把不同实例归为同一对象。

严格数学上的等价关系必须满足自反、对称和传递。本文的经验研究不能轻率宣称人类写作天然满足这三个性质，因此理论需要把两个层面区分开。理论层面固定一个生成见证 G 与允许扰动集合 Δ 后，“由同一 G 生成”可以构成对象身

份的等价标准：每个文本当然与自身共享 G；若 T_i 与 T_j 共享 G，则关系对称；若 T_i 、 T_j 、 T_k 都由同一 G 在同一问题域生成，它们属于同一类。经验层面则存在提取误差、阈值误差和见证漂移，所以观测到的只是对这一等价关系的带噪估计，而不是完美数学等价。

这个区分可以防止一个常见漏洞：如果 A 与 B 看起来有些相似，B 与 C 也有些相似，就不能据此宣布 A、B、C 是同一个作文。真正的传递性来自“共同的 G”，不是来自相似性链。研究必须先提取一个能够跨实例工作的关系结构，再用留出挑战检验它，而不是让成对相似度无限传递。

因此，“作文等价类”最重要的不是 class 这个名词，而是它强迫我们给出身份判据：什么保持不变时，我们仍然认为是同一个作文；什么一旦改变，就必须承认已经是另一个作文。

（二）生成见证的层级：为什么“主题相同”远远不够

生成见证可以至少区分四个层级。最低层是表达痕迹，例如词汇、句长、连接词和修辞偏好。这些特征适合来源归属，却很容易被模仿，也不构成本文的 G。第二层是主题或立场，例如“支持 AI”“重视独立思考”。主题能够跨文本保持，却通常无法预测新的组织动作，因此也不够。

第三层是关系结构。它说明若干变量之间如何发生，例如“外部替代越靠近问题生成前端，短期效率收益越可能伴随内部生成机会减少”。这种关系会影响选材、因果方向、反例处理和结论边界，具有真正的生成力。第四层是关系的修订规则：当出现某类反例时，写作者会扩大条件、改变因果方向、区分时间尺度，还是直接放弃原判断。一个成熟 G 不仅能产生结论，还能规定自己怎样被证据修改。

这四层意味着，真正稳定的作文身份不要求“观点永远不变”。相反，如果写作者面对新证据后按照稳定的证据纪律修订原关系，他可能仍然处于同一更高层生成见证之中。这里的同一性不是结论冻结，而是发生规则具有连续性。这个设计也避免把作文等价类变成思想保守性的奖励机制。

未来研究应优先编码第三、第四层，而不是把第一、第二层当作替代指标。若一项研究发现所谓 G 只不过是主题标签或风格特征，应该判定核心构念未被测量。

（三）身份与质量必须正交：低水平作文也可以有稳定身份

对象身份与价值质量的分离，是本文必须坚持的另一条纪律。一个初学者可能反复产生“只要努力就会成功”这种简单因果关系，在多个挑战中稳定重现。

从理论上说，他可能拥有清晰但低复杂度的生成见证。另一个学生借助强大外部系统交出高度复杂、材料丰富的文章，却在新挑战中无法再次组织任何关键关系。前者的作文质量可能低，生成身份却清晰；后者代表元质量很高，生成身份却主要外置。

如果把 GWF 直接当作新的作文分数，就会重新犯旧错误：用一个数字吞并多个对象。教育评价至少要把“代表元质量”和“生成身份证据”并列。代表元质量回答这篇文章写得怎样；生成身份证据回答这篇文章在多大程度上能够作为主体内部某套生成关系的显露。两者可以相关，却不能预先设定为同一个维度。

这种正交性还提供一个重要反例。如果未来实验发现 GWF 只是在复制传统作文高分——高分者永远 GWF 高、低分者永远 GWF 低——那么本文的新轴很可能没有独立价值。真正支持理论的证据应包括交叉个案：语言表现普通但 G 稳定，以及文本表现优异但 G 不稳定。

十、同文异作与异文同作：新的二维辨别框架

作文等价类最直接的检验不是发明新名词，而是能否产生传统理论不容易同时解释的相反判决。本文选择两根彼此正交的轴：文本表面相似度与生成见证等价性。前者属于完成态，后者属于生成身份。

当文本相近且生成见证等价时，是“同文同作”；当文本不同且生成见证不等价时，是“异文异作”。真正具有判别价值的是另外两个交叉格：“同文异作”和“异文同作”。前者意味着两个字符串高度相同，却来自不同生成身份；后者意味着两个文本表面差异巨大，却是同一作文空间的不同代表元。

这个二维结构的重要性在于，它使 AI 时代大量模糊争论变成可分解问题。我们不再需要先争“AI 用了 30% 还是 70%”，而可以分别问：当前文本特征怎样？可提取生成见证怎样？AI 可能改变表面质量，也可能改变生成见证；两种变化不能用同一比例概括。

	生成见证等价	生成见证不等价
文本相近	同文同作	同文异作
文本不同	异文同作	异文异作

表 1 文本相似度与生成见证等价性的二维辨别框架

十一、最危险的靶格不是“AI 代写”，而是高质量的“同文异作”

最容易被制度发现的情况，是一个学生完全没有参与，直接提交外部生成文本。真正危险的情况反而是表面上一切都很好：学生参与了提示、筛选、润色和排版；文章符合格式；教师能看到大量修改；学生甚至能够解释原文。可是当原有成品和对话记录被撤掉，局部条件稍作变化，那套关键意义关系无法再次组织新文本。

这种状态之所以危险，是因为传统短期指标可能全部改善。文章质量提高，完成速度提高，语言错误减少，学生自我感觉也可能更好。Huang 与 Wilson 对 431 名四、五年级学生使用自动写作反馈的研究就发现，学生在有自动反馈帮助的不断修改中表现提高，但跨学年比较初稿时没有观察到 **unaided transfer**；也就是说，“在帮助下越改越好”没有自动变成“不靠帮助也能把第一稿写得更好”[12]。这并不能直接证明作文等价类，但至少说明“代表元改善”和“主体独立能力变化”可以分离。

2026 年 Dai 对 120 名英语专业本科生进行 AI 反馈研究，又发现不同的反馈素养和校准训练会产生不同结果；结合反馈素养与自我校准训练的组，在 AI 支持撤除后的 **retention** 表现更强[13]。这同样不是生成见证的直接验证，却说明外部工具不是单向替代：在某些训练条件下，外部反馈可能被转化为内部可持续能力，在另一些条件下则主要改善当前文本。

因此，真正值得教育系统监测的不是“有没有 AI”，而是短期代表元质量与生成身份是否开始脱钩。一个系统如果只奖励最终文档，很可能无意中训练学生优化代表元，而不训练能够产生新代表元的生成结构。

十二、AI 参与比例为什么不是作文归属的核心变量

AI 写作政策常试图划定一个比例：AI 生成多少字、改写多少句、参与到哪一步才算“自己的”。这种做法便于管理，却把多个不同变量混成一个变量。辅助程度会影响人们对作者身份、创造者身份和责任的判断，Formosa 等人对 602 名参与者的实验已经显示这一点[6]；但“作者身份”与“作文生成身份”仍不完全相同。

假设 A 学生用 AI 进行语法润色，核心判断、材料关系和结构全部由自己产生；B 学生自己输入每个字符，却严格复制模型给出的论证；C 学生让 AI 生成十种论点，自己经过反例检验后重组出 AI 没有直接提供的新结构。按照 AI 字数比

例，三人的排序可能完全失真。按照作文等价类，问题变成：撤去当前输出和相应脚手架以后，哪些关键关系仍能在新的挑战中由主体重新组织？

因此，AI 可以发生三种不同作用。第一种是“执行代理”：它替主体完成输入、语法和表面转写，生成见证主要不变。第二种是“结构刺激”：它提供反例、候选或比较，使主体形成原来没有的关系，随后这种关系能够在 AI 撤除后继续发挥作用。第三种是“生成外置”：关键意义关系长期只存在于外部系统，每次需要新代表元时都必须重新向外部系统索取。

这三种状态不能用同一个 AI 使用率区分。作文等价类理论因此拒绝把技术纯洁性当作本体标准。工具参与多并不自动导致作文消失，手工参与多也不自动证明作文已经发生。

十三、与七个最近邻的决定性边界

任何新概念如果只能靠“更全面”“更深入”“更关注生成”来与旧理论区分，就没有独立存在的资格。下面给出作文等价类与七个最近邻的决定性边界。

第一，写作过程理论。过程理论研究一次 **actual composing** 如何展开；作文等价类研究多个可能实现何时具有同一生成身份。如果第一次写作的完整过程日志已经足以唯一预测后续扰动响应，不需要任何反事实挑战，那么本文退化为过程研究。

第二，**writing transfer**。迁移研究问知识、技能、策略能否跨情境使用 [3]；本文问两个文本何时属于同一个作文对象。再作或新任务表现可以作为测量工具，但不是本体定义。如果成熟 **transfer** 指标能够无剩余地完成本文所有等价分类，作文等价类应当降级。

第三，测评效度。**Kane** 关注从测试反应到分数解释和使用的推断是否合理 [4]；本文关注在不涉及分数时，两个文本实例是否具有同一对象身份。即使取消考试与评分，同文异作、异文同作仍然存在，所以两者不能完全互换。

第四，**authorship attribution**。作者归属从已存在文本特征推断来源；本文允许一个主体在不同代表元中风格差异巨大，也允许两个来源产生字面相同文本。若表面风格识别已经足以完成等价分类，则本文失败。

第五，**authorial agency** 与 **creative control**。AI 共写研究已经直接研究人的参与、控制、所有权与认知投入 [5][6]。本文允许“实际控制高但生成见证弱”，也允许“字符执行控制低但生成见证强”。如果实际控制指标与生成见证等价性近乎一一对应，就没有必要设置第二个对象。

第六，retention 与 learning gain。保留研究关注学习效果是否持续；本文允许一个稳定的生成见证本身水平很低，因此“见证稳定”不等于“成绩提高”。如果把作文等价类简单变成“学会了没有”，就会把对象身份重新偷换成教学成效。

第七，中心思想与 schema。一个主题或 schema 可以跨文本重复，却未必对留出挑战提供足够精细预测。生成见证必须在多个竞争关系之间给出方向性约束，否则只是研究者事后用高抽象词概括文本。

近邻概念	它主要回答	作文等价类回答	决定性分界
写作过程	这一次怎样写出来	哪些可能文本属同一生成身份	完整过程若足以预测等价类，则新理论失败
Writing transfer	知识/策略能否进入新情境	两个文本何时是同一个作文	迁移指标若无剩余完成分类，则新理论降级
测评效度	由表现推断能力是否合理	对象单位本身是什么	取消评分后身份问题仍存在
作者归属	现有文本像谁/来自谁	生成见证属于谁	允许异文同作、同文异作
Authorial control	实际过程由谁控制	撤去成品后什么结构仍可生成	高控制与高见证可分离
Retention	学习效果是否持续	生成身份是否等价	稳定不等于高质量或进步
中心思想/schema	文本有什么共同主题/结构	什么关系可预测未见响应	无法预测留出挑战则不算见证

表 2 作文等价类与最近邻概念的决定性边界

十四、可操作化：挑战—提取—留出验证，而不是“再写一篇看看”

作文等价类要成为研究对象，必须从哲学区分进入可执行程序。最小实验可以分为五步。

第一步是母任务。写作者在常规条件下完成 T0。T0 可以用于确定问题域和候选关系，但不能在后续挑战中被直接调用。

第二步是预注册扰动。研究者在看到后续结果前写明 Δ。扰动必须保持核心认知任务同型，却改变至少一个局部条件，例如受众、材料可得性、反例、文体、字数、时间、工具权限或评价目标。扰动过小可能只是复述测试，过大则会 把“同一问题域”破坏掉。

第三步是挑战响应。写作者在看不到 T0 和既有 AI 会话的情况下完成 T1 和 T2。研究者或算法只使用 T1、T2 提取候选生成见证 G。提取必须列出具体的关系规则，不得使用“思想深刻”“重视成长”“逻辑性强”之类无法失败的形容。

第四步是留出挑战。使用没有参与 G 提取的 S3 产生 T3。研究者必须在看到 T3 前写出 G 对 T3 至少一个组织关系的方向性预测。例如，如果 G 是“外部系统越早替代内部生成，短期效率与长期独立生成能力越可能分离”，那么面对一个关于家长过度帮助的留出题，预测应落在“过早替代—短期顺畅—长期能力空缺”的关系，而不是预测某个具体词。

第五步是竞争模型比较。至少同时运行三个基线：文本相似度、一般写作能力/语言水平、已有 transfer 指标。如果 G 的预测力不超过这些基线，或者只在研究者事后调整后成立，作文等价类没有获得独立经验支持。

这个程序的核心不是逼学生“再写一次”，而是把一次成品无法提供的反事实信息变成可观察数据。

（一）扰动不是越大越好：六类最小挑战

有效挑战的目标不是把写作者推到完全陌生的任务，而是在保持核心问题结构的前提下，让不同竞争生成机理产生不同预测。至少可以设计六类最小扰动。

其一，受众扰动：把面向教师改为面向同龄人或决策者，观察核心因果关系是否保持而表达策略变化。其二，证据扰动：加入一个与原判断冲突的新事实，观察写作者是机械忽略、表面加一句让步，还是按照某种稳定修订规则重构判断。其三，资源扰动：撤去原来依赖的 AI、范文或材料，观察关键关系是否仍可由主体重新组织。其四，形式扰动：从议论文改为口头解释、图示或短备忘录，以排除固定篇章模板。其五，约束扰动：改变字数、时间或必须使用/禁止使用的材料，检验 G 是否只是特定格式的产物。其六，角色扰动：要求写作者为反方先构造最强论证，再返回自己的判断，观察其关系规则是否具有竞争解释能力。

这些扰动都必须在看到结果前明确“若 G 成立，应保留什么；应允许什么改变”。如果研究者只在结果出来后说“虽然不同，但我觉得深层还是一样”，挑战就失去证伪价值。

（二）生成见证提取的四条编码纪律

第一，关系必须能够写成带方向的条件结构，例如“当 X 提前替代 Y 时，Z 的短期值上升而 Y 的后续独立生成下降”，而不是一个静态标签。第二，每个 G 至少给出一个它明确不预测的邻近关系，以防概念无限扩张。第三，提取者必须记录备择 G。若同一组 T1、T2 同时兼容两个关系结构，留出挑战必须能够让两者

给出不同预测。第四，G 的复杂度必须受约束。理论上可以借鉴最小描述思想：如果一个 G 为了匹配每个文本不断增加例外条款，它就失去生成解释意义。

这四条纪律使“生成见证”从一种解释性修辞变成竞争模型。研究的目标不是证明研究者总能从文本中读出深意，而是看某个有限 G 能否以比竞争模型更低的复杂度预测未见响应。

十五、测量手册：生成见证保真率只是辅助手段，不是作文的本体

为了避免同一指标在正文、证伪和未来实验中不断改变定义，本文给出一个最小测量手册。设从训练挑战 T1、T2 中预先提取 m 条可判定的关系约束 g1...gm。对留出文本 T3，由两个以上盲评者判断每条关系是否在没有直接提示的情况下得到实现。定义“生成见证保真率” GWF（Generative Witness Fidelity）为：留出文本中得到实现的预注册关系约束数除以可判定约束总数 m。

GWF 的粒度是“关系约束”，不是句子。若某约束因留出任务不具备适用条件，应在看到 T3 前由预注册规则判定为 N/A 并从分母排除，不得在结果出来以后选择性删除。每条关系必须给出正例、反例和边界例。至少两个盲评者独立编码，分歧由第三评者裁决，并报告一致性。

GWF 不能单独证明作文等价。首先，它必须与文本相似度正交：理想案例包括文本相似度很低但 GWF 高，以及文本相似度很高但 GWF 低。其次，它必须与一般写作分数区分：语言水平高的写作者可能写出优秀留出文本，但没有实现预注册 G；语言粗糙的写作者也可能稳定实现 G。再次，一次高 GWF 不足以排除题目熟悉、偶然重合或研究者过度抽象，需要多个扰动与留出任务。

因此，GWF 只是测量窗口。删除 GWF 这个指标，作文等价类作为对象仍然成立，只是难以操作化；这说明本文的核心不是新指数，而是新对象单位。

项目	固定定义	禁止漂移
测量对象	预注册的关系约束 g1...gm	不得临时改成句子相似或主题相似
GWF	留出文本实现的可判定关系数 / 可判定关系总数	结果出来后不得删掉失败关系
粒度	关系层	不得以词汇、句法或单句重合代替
N/A	任务结构客观不触发该关系	必须预先定义 N/A 规则
编码	至少 2 名盲评者+第三评者裁决	不得由理论作者单人决定
基线	文本相似度、一般写作水平、transfer 指标	必须报告增量预测而非只报相关

表 3 生成见证保真率（GWF）最小测量手册

十六、真实数据方向检查：支持“分离可能”，但尚未验证核心构念

一个新理论最容易犯的错误，是把与自己方向一致的既有研究直接宣布为验证。本文拒绝这种做法。现有公开研究可以检验的，只是理论的一些较弱前提。

Huang 与 Wilson 的 AWE 研究显示，431 名小学生在自动反馈支持下的连续修改质量提升，但初稿质量没有出现 **unaided transfer**[12]。这支持“当前代表元改善与独立表现可以分离”这一弱命题，却不能告诉我们学生是否共享同一生成见证，因为研究没有设置针对关键意义关系的预注册挑战。

Dai 2026 年的研究则给出另一面：120 名英语专业学生中，带有反馈素养和校准训练的条件在 AI 撤除后的 **retention** 更强[13]。这支持外部反馈有可能被转化为较稳定内部能力，而不是必然造成外置依赖。但该研究主要测量写作质量、自评准确性、反馈采纳和修订深度，同样没有建立生成见证等价字段。

这个结果既有利也不利。有利之处是，理论所需要的“代表元质量与撤除后表现可以分离”在真实研究中确实出现；不利之处是，现有数据无法直接运行核心检验。更严谨的做法不是用缺失数据填补结论，而是把“现有账本没有这个字段”明确记录为第三层经验主张：目前大量写作研究把多次文本作为前测、后测、迁移任务或保留任务分别评分，却很少把它们编码为“是否由同一可提取生成见证产生”。

因此，本文目前只能声称提出了一个具有可操作方案、但尚未被专门数据验证的对象。任何把它写成“已经证实作文就是等价类”的表述都超出了证据。

十七、最强竞争解释与机制证伪：这不就是迁移加效度吗

最强的竞争解释不是“作文就是文本”，而是一个更成熟、更危险的组合：本文所谓生成见证不过是 **writing transfer** 加 **assessment validity** 的重新包装。要让新理论获得独立资格，必须设计能够让自己失败的比较。

第一条证伪：控制已有写作水平、文本相似度、题目熟悉度和成熟 **transfer** 指标后，如果 GWF 对留出文本中的关键组织关系没有任何增量预测，那么生成见证没有独立价值。

第二条证伪：如果完整过程日志、提示词日志和修改记录能够在不做任何额外挑战的情况下，近乎完全预测作文等价关系，那么“反事实响应族”不是必要对象，本文应退化为过程研究。

第三条证伪：如果 **authorial control**、**ownership** 或参与量指标与 **GWF** 在不同 **AI** 协作条件下长期近乎一一对应，那么作文等价类只是控制理论的复杂写法。

第四条证伪：如果从 **T1**、**T2** 抽取的 **G** 在 **T3** 中不能稳定预测预注册关系，或者研究者只有在看到 **T3** 后重新描述 **G** 才能获得高匹配，所谓生成见证就是事后故事。

第五条证伪：如果“同文异作”与“异文同作”在真实样本中几乎无法出现，绝大多数等价分类都可以由文本相似度替代，那么二维框架没有必要。

第六条证伪：如果不同盲评者对 **G** 的关系约束无法达到基本一致性，且通过增加抽象程度才能提高一致性，则生成见证缺少可测量边界。

如果这些证伪中的任何关键一条长期成立，理论应当缩小适用范围甚至被放弃。真正的理论风险不是被批评，而是拥有一种无论出现什么结果都能解释自己的语言。

证伪编号	竞争结果	对理论的后果
F1	控制现有变量后 GWF 无增量预测	生成见证无独立解释价值
F2	完整过程日志足以恢复等价类	退化为过程理论
F3	控制/ ownership 与 GWF 一一对应	退化为作者控制理论
F4	G 不能预测留出挑战	生成见证是事后故事
F5	文本相似度即可完成全部分类	第二轴多余
F6	盲评无法可靠编码关系约束	构念缺少经验边界

表 4 核心机制证伪条件

十八、研究设计：怎样第一次真正检验“同文异作”

一个最低成本的第一项研究可以从大学生议论文开始。招募至少 120 名参与者，先测一般写作水平与 **AI** 使用熟练度，再随机进入四种生成条件：独立写作、**AI** 语言润色、**AI** 观点候选+人工重构、**AI** 整篇生成+人工编辑。母任务使用同一问题，所有交互日志完整保存。

随后不评分原文的“是否属于本人”，而把原文暂时封存。24 至 72 小时后进行三个挑战。挑战一改变受众；挑战二加入一个会破坏原论证的反例；挑战三换成结构同型但主题表面不同的问题。三次挑战均不允许调用原文和原 **AI** 会话；其中前两次用于提取 **G**，第三次作为留出任务。

研究者在不知道参与条件的情况下，从 **T1**、**T2** 中编码候选 **G** 并预注册对 **T3** 的预测。主要比较不是哪组作文分数最高，而是：四组在 **GWF**、文本相似度、传

统 transfer 分数与 authorial-control 量表之间出现什么结构。理论最想看到、同时也最危险的证据，是高控制与高 GWF 并非完全重合，文本相似与 GWF 出现交叉分类。

此外可以专门制造“同文异作”实验。让两名参与者最终得到完全相同的目标文本：一人通过自主问题解决逐步生成，另一人在不提前理解结构的条件下被训练复现文本。之后撤去文本并进入扰动挑战。如果两人在文本完全相同的前提下出现稳定的 GWF 分化，就获得最直接的对象区分证据。若不分化，则理论需要重大修订。

这种设计的价值是把 AI 伦理争论变成可被实验推翻的问题：不是问“你觉得用了 AI 还算不算自己的”，而是问不同生成条件是否改变了后续可提取的生成结构。

（一）四种高信息量个案及其相反预测

个案 A 是“完全同文”。研究者让甲通过自主解决形成目标文本，让乙通过逐步提示最终复现完全相同的文本。文本理论预测两者不可区分；作文等价类预测撤稿后的扰动挑战会出现分化。若不分化，理论受损。

个案 B 是“完全异文”。同一写作者分别用故事和议论文回答结构同型问题，表面相似度很低。风格归属或字符串匹配可能判定差异大；作文等价类预测若两者共享 G，对留出挑战的关系预测应保持。若只能靠主题宽泛相似解释，理论受损。

个案 C 是“高控制 AI 共写”。参与者亲自决定全部 prompt、筛选和修改，主观 ownership 很高，但核心洞见最初由 AI 给出。authorial-control 理论可能把作品高度归属于人；作文等价类不预设结论，而要看撤掉原会话后 G 能否被提取。如果 G 仍稳定，AI 参与已经被内化；如果不能，则作者责任与作文生成身份发生分离。

个案 D 是“低字符控制辅助写作”。残障写作者通过口述、眼动或助手完成绝大多数字符执行。字符控制指标低，但若不同挑战中 G 稳定，作文等价类应判定生成身份仍归主体。这个个案是理论必须保留的伦理校验：任何把“亲手打字比例”当真实性指标的方法都会在这里做出错误判决。

（二）为什么需要不利结果，而不是只找“漂亮案例”

第一次研究最好主动包含一个理论预计会失败的区域。例如极短文本可能不足以提取稳定 G；高度事实性任务可能主要由知识回忆决定；完全新领域的挑战

可能超出原问题域。这些负结果不应被解释为“学生没有作文”，而应帮助确定作文等价类的适用边界。

更重要的是，如果某些参与者在训练挑战中表现出高度稳定的 G，却在留出挑战中完全失效，研究必须保留这些案例并检查是否存在过拟合、挑战改变了问题域，或所谓 G 根本只是事后归纳。只有让不利结果修改理论，真实运行才有意义。

十九、教育含义：从“把这一篇改好”转向“让生成关系留下来”

如果作文等价类理论获得经验支持，作文教学最大的变化不是发明一张更复杂的评分表，而是重新区分两种教学目标：改善当前代表元，与扩大主体生成空间。

教师替学生改掉病句、给一个更好的例子、调整文章结构，都可能显著提高当前 T0。这些行为并没有问题；问题是，如果所有关键判断始终由教师或 AI 完成，T0 的改善不必然扩大 [T]G。真正面向作文发生的教学，应让学生逐渐获得生成关系：为什么这个例子不能支持该论点，为什么加入某个反例后需要改变结论，为什么换受众以后同一个价值判断要重新组织证据。

因此，修改课可以增加“撤稿后挑战”。文章改好以后暂时不让学生继续看成品，换一个材料、增加一个反例、要求向不同对象解释，让学生重新发生一次。教师观察的不是记住了多少金句，而是原来被教授的关系是否转化为新的组织行为。

这也重新解释“范文”的价值。范文不是不能看，而是不能让代表元替代生成结构。最好的范文教学不要求学生复制成品，而要让学生反向寻找：哪些关系使这个成品成为可能？如果换一个事实，这个关系怎样重组？当范文被撤掉时，还能不能产生新的代表元？

AI 同样如此。真正有效的 AI 写作教育不是规定一个绝对禁区，而是设计让 AI 的外部支持逐步进入主体判断的机制。Dai 的研究提示，反馈素养与自我校准训练可能比单纯获得更多 AI 反馈更有利于支持撤除后的保留[13]。作文等价类理论把这种教育目标进一步说清：不是让工具永远缺席，而是避免关键生成见证永久留在工具外部。

二十、伦理边界：不能把“生成见证”变成新的学生人格审判

任何能够评估学生的理论都存在被制度滥用的风险。作文等价类尤其危险，因为“生成见证”听起来容易被误解为“这个学生到底有没有思想”。因此必须设置至少三条伦理边界。

第一，测量对象是特定问题域中的生成关系，不是人的整体价值。一个学生在议论文问题族中 GWF 低，不能推出他“没有思想”“不诚实”或“没有创造力”。困难可能来自语言、焦虑、领域知识不足、障碍条件或挑战设计错误。

第二，不能为了提高所谓真实性而制造不安全的高压测试。扰动必须最小、透明、可复核，不得通过突然羞辱、隐瞒重大规则、剥夺合理辅助工具来“逼出真实能力”。对残障写作者尤其应区分辅助执行与意义生成，不得把无障碍技术使用误判为外置生成。

第三，如果 GWF 或等价类判定会影响成绩、录取、学术诚信处罚，必须公开编码规则、提供原始挑战记录和人工复核渠道。由于构念仍处研究阶段，它不应被直接用作惩罚性 AI 检测工具。本文提出的是研究假说与教育诊断框架，不是新的“作弊探测器”。

这三条伦理限制会降低工具化速度，却是理论成立所必须付出的代价。一个关于“作文是否真正发生”的理论，如果最终变成对学生人格和诚实度的黑箱判决，本身就违背了它要求证据可提取、可挑战、可证伪的逻辑。

（一）六个边界案例：理论必须给出相反而非含糊的判决

案例一：口述与代录。学生完整形成论点并逐句口述，由他人代为输入、纠正错别字。若在新挑战中仍可提取同一 G，字符执行虽然外置，作文生成身份仍可归于学生。这个案例检验理论是否错误地把“手工操作”当作必要条件。

案例二：翻译协助。学生先以母语形成完整论证，再借助 AI 翻译成第二语言并人工核对。如果挑战改用母语或另一表达形式后 G 仍稳定，语言转写工具不应自动改变作文身份。相反，如果关键推理本身由翻译系统补写，留出挑战应暴露生成身份的混合。

案例三：教师深度批改。原稿很弱，教师逐段给出关键判断并要求学生照改，最终文章质量大幅提高。若学生只能解释教师为什么这样改，却在新材料中无法自己形成相应关系，则当前 T0 的质量提升不等于 G 已经转入学生。若经过反

复校准后学生能独立重建关系，则同一外部干预可能从“代生成”转变为“促内化”。

案例四：模板熟练者。学生不用任何 AI，完全手写，并能在大量题目中迅速套用“提出问题—举名人例子—升华”的格式。表面上它具有很强的跨题稳定性，但一旦 challenge 要求处理真正会推翻原立场的反例，模板无法决定如何修订因果关系。这个案例说明“可重复”本身不是 G；能够生成关系修订，才是更高层见证。

案例五：AI 首次提供关键洞见。学生此前从未想到某个关系，AI 给出后，他经过质疑、寻找反例和重新解释，最终把它内化。几天后不给原对话，他在不同主题中再次独立生成同一结构。作文等价类不追究“第一瞬间是谁先说出”，而承认生成见证可以在互动中形成。这个判决与纯原创论不同。

案例六：写作者主动推翻旧观点。T0 主张“限制短视频能保护注意力”，T1 在新证据下变成“单纯限制无法解决注意力问题，关键在于重新训练延迟反馈能力”。若变化是由一个更高层规则——“当干预只阻断表面行为却不改变维持机制时，效果不会稳定”——所组织，那么表面立场改变并不意味着作文身份断裂。这个案例说明，等价类的同一性不能冻结结论，而应允许有规则的重构。

六个案例共同构成理论的最低压力测试：它必须同时允许辅助工具、协作、观点变化，又能够拒绝抄录、机械模板和外部关键关系长期不内化的情形。如果理论只能把所有复杂案例都回答成“视情况而定”，它就没有形成真正的判别维度。

（二）作文身份可以是混合的，但混合不能成为逃避判定的万能答案

在人机共写和多人协作中，生成见证可能不是单一来源。例如学生确定问题结构，AI 提出反例，教师要求补证据，学生最后形成新的条件判断。此时最合理的模型可能不是强迫选择“人”或“AI”，而是分析 G 的哪些关系能够在主体独立挑战中继续生成，哪些关系一撤去外部系统就消失。

因此，可以区分“已内化见证”“外置见证”和“分布式见证”。已内化见证在外部支持撤除后仍能被主体挑战响应提取；外置见证只能在特定工具或他人持续存在时维持；分布式见证则要求多个行动者持续协同才能重建。对于教育性作文，真正关心的通常是已内化部分；对于职业团队写作，分布式见证本身可能就是合法对象。

但是，“混合”不能成为理论不做判断的借口。研究仍必须明确：当前任务究竟要评价个体学习、团队产出还是责任作者身份。对象不同，允许的生成主体也不同。只有先固定评价对象，混合生成才有可解释边界。

二十一、适用范围与不适用范围

本文的对象首先是具有教育性或能力形成目标的“作文”，尤其是学校与大学中的议论、解释、反思等任务。它不应直接覆盖所有职业写作与文学创作。职业记者、政策团队、广告公司、科研团队可能把文本生产明确设计为分工系统，在这些场景中，组织的目标并不是证明某一个个体拥有完整生成见证。用“作文归属”去评判协同生产，会犯类别错误。

文学创作也需要谨慎。小说、诗歌和实验写作可能有意利用偶然性、材料性、协作性或无意识过程，其价值不必来自稳定可重复的意义关系。本文的“留出预测”若被硬套到所有艺术创造，可能压制真正的新奇和断裂。因此，作文等价类只声称能够处理一种特殊身份问题：当我们希望一份文本同时作为主体学习、理解或生成能力的证据时，单一代表元是否足够。

此外，等价类的边界依赖 Δ 。若允许扰动过窄，几乎所有复述都能通过；若过宽，任何稳定关系都会被情境变化摧毁。因此未来研究必须把“问题域怎样划定”作为与G同等重要的对象。不同年龄、语言水平、学科和体裁可能需要不同 Δ ，不存在一套普遍挑战模板。

最后，本文不主张所有真正学习都必须表现为稳定重复。一个人可能在一次写作后发生新的转变，第二次主动推翻第一次的G。如果这种变化有新的可解释发生机制，它并不是失败，而可能意味着作文空间正在重组。因此，理论需要区分“无法再生”与“经理由新证据驱动的有方向重构”。稳定不是最高价值，可重新生成也包括可被证据改变。

（一）作文不是“能力本身”：一次身份判定不能吞并人的全部潜能

把作文定义成等价类，并不意味着一个人的全部写作能力都能被一个等价类代表。能力是更广泛的潜在资源，包括语言知识、领域知识、修辞判断、记忆、注意、阅读经验以及与工具协作的能力。一个作文等价类只是这些资源在某个问题域中形成的一种可提取生成身份。一个人可以同时拥有多个彼此冲突的G，也可以在学习过程中把旧G重组为新G。

这一区分对于教育评价非常重要。传统表述常说“这篇作文体现了学生的能力”，但实际上从单篇文本到总体能力需要跨越许多推断。作文等价类并不取消

这种能力推断，而是在其之前增加一个更小的对象：先问多个文本是否共享某种生成身份，再问这种身份能否推广为更稳定、更宽域的能力。这样可以减少从一篇漂亮文章直接跳到“这个人会写”的过度推断。

同理，一个人在某个问题域中暂时无法形成稳定 G，也不能说明他没有未来潜能。尤其对于儿童和初学者，生成结构可能处在快速变化阶段。教育的目的可以是帮助新的 G 发生，而不是把当前等价类当作不可改变的身份标签。

（二）作文也不是“表达本身”：有些表达是在发生，有些表达只是把既有结果搬到纸上

“作文是表达思想”是最常见、也最容易让人满足的定义。但表达这个词同时包含两个完全不同的方向。一种是外显：一个已经较完整存在的思想被翻译成语言。另一种是发生：写作者在表达过程中才发现关系、修订判断、形成此前并不存在的结构。Flower 与 Hayes 的过程传统已经表明写作并不是简单把预先完成的思想搬运到文字上[1]。

作文等价类进一步指出，即使完成了有效表达，我们仍需区分“表达了一次”与“形成了能够继续生成的关系”。一个人可以准确复述一套复杂思想，甚至把它表达得比原作者更清楚，但这并不必然让相关生成见证属于他。相反，一个初学者表达粗糙，却可能正在形成真正能够被新挑战激活的关系。

因此，“表达”更多描述文本与意义之间的当前映射，而“作文身份”要求考察这种映射背后的生成关系是否在不同实现之间保持可提取连续性。二者相关，却不是同义词。

（三）作文不是“原创度”：外部来源可以很多，生成身份仍然可能清楚

另一个容易误入的方向是把生成见证理解为“纯原创”。人类写作从来建立在语言、文化、阅读、引用、师承和社会对话之上。若要求关键内容必须从主体内部凭空产生，几乎所有成熟写作都会被错误排除。

生成见证关心的不是原材料来源，而是关系组织的归属。一个写作者可以大量引用外部材料、接受教师意见、使用 AI 寻找反例，只要这些输入进入以后被一套能够在新挑战中继续工作的关系结构重新组织，作文身份仍然可以清楚。反过来，一个人可以完全不使用 AI、完全手写，却只是复现早已背熟的套路，其生成身份仍可能高度依赖外部固定结构。

因此，作文等价类与“原创检测”存在根本差别。原创检测寻找相似来源与未披露借用；生成见证检验的是不同来源材料被组织成新响应时，是否存在可提取的主体关系。前者主要属于规范与知识产权问题，后者属于发生与学习问题。

二十二、反身性：这篇论文自己也必须接受“撤掉成品以后还剩什么”

如果本文主张一篇完成文本不能自己证明生成见证，那么这篇论文也不能因为被写成两万字、引用了文献、拥有公式和表格，就自动证明“作文等价类”是一套已经成立的理论。

对本文最直接的反身测试是：撤去当前文档，只保留核心问题“为什么单一文本不足以定义作文”，换一个领域例如口语答辩、编程作业或音乐即兴，能否重新推导“代表元—生成见证—等价类”的关系，并在不同近邻概念压力下改变边界？如果只能背诵“同文异作、异文同作”这两个口号，却无法根据新反例重新组织模型，那么本文自己就是一个漂亮代表元，而不是稳定的生成结构。

更严格地说，未来的独立研究如果发现 **transfer**、**process trace** 与 **creative control** 三个现有构念已经能够联合解释全部预测，作者应当放弃“作文等价类”的独立本体地位，而把它降为一种整合性分析框架。相反，如果出现大量“文本质量相同、过程参与相似、控制感相似，但留出生成关系稳定分化”的案例，理论才获得真正不可被替换的经验位置。

这种反身性不是姿态。它意味着本文的最终价值不能由概念新颖度保护，只能由它是否制造新的可失败问题来保护。

（一）三层主张结构：即使强本体失败，也不让全部研究一起坍塌

为了避免一个高创新命题在最强版本失败后整篇论文失去价值，本文把主张分成三层。第一层是最强本体主张：作文不是单个文本，而是共享同一生成见证的文本等价类。它最容易被近邻概念吸收，也最需要严格证伪。

第二层是辨别装置主张：无论是否接受“等价类”作为作文的完整本体，把文本相似度与生成见证等价性分成两轴，至少能够产生“同文异作”和“异文同作”的经验分类。如果未来研究发现该二维分类具有独立预测价值，即使学界不愿把它叫作作文的本体，第二层仍可作为分析工具存在。

第三层是最稳健的经验主张：当前写作数据库通常记录单篇文本、分数、过程日志、AI 交互或迁移结果，却缺少“多个文本是否由同一预注册关系结构生

成”的直接字段。只要这一事实成立，设计多轮挑战并比较实际轨迹与反事实响应就具有独立研究价值。第三层最容易验证，也最不依赖强哲学承诺。

这种分层使理论能够诚实失败。如果未来证据否定第一层，我们不应为了保护概念而不断放宽“作文”的定义，而应明确降级到第二层或第三层。

（二）三种会迫使理论改名或改落点的结果

第一种结果是发现“生成见证”只能解释学习迁移，却无法提供对象身份的独立判决。那时理论应放弃“作文等价类”名称，转化为一种反事实迁移测量模型。第二种结果是发现写作者的 G 在不同挑战中持续变化，严格等价关系几乎不存在，但变化具有可追踪的谱系。那时“等价类”可能需要改成“生成谱系”或“生成族”，把同一性从静态共享 G 改为有规则的 G 演化。第三种结果是发现 AI 共写中 G 天然分布在人与系统之间，无法合理提取为单一主体见证。那时理论必须承认“分布式作文”这一对象，而不是强行把所有关系归到人。

这三种结果都不是理论的修辞退路，而是预先写出的改名条件。它们保证未来实证结果能够真正改变本体主张，而不是永远被现有术语吸收。

二十三、一个有日期的赌注

为了防止理论无限延期验证，本文给出一个明确赌注。以 2028 年 12 月 31 日为界，如果至少有一项预注册研究在控制一般写作水平、文本相似度、题目熟悉度、成熟 writing-transfer 指标和实际过程参与度以后，发现由训练挑战提取的生成见证对留出挑战中的关键关系仍具有显著且可重复的增量预测，并能产生至少一类稳定的“同文异作”或“异文同作”交叉案例，则本文的核心对象获得第一层经验支持。

反之，如果到该日期前的合格研究显示：生成见证的所有预测力都可被文本相似度、transfer、过程日志或 authorial-control 指标吸收，或者盲评者无法可靠识别预注册关系，本文应放弃“作文是一种独立等价类对象”的强主张。

不计入命中的情况包括：只证明 AI 组和非 AI 组作文分数不同；只证明学生自我报告 ownership 不同；只证明支持撤除后成绩发生变化；只在看完留出文本后重新定义 G；或者以“思想相似”这种无法编码的主观判断替代预注册关系。这些结果可以与本文相关，却不能算核心赌注命中。

二十四、结论：文章是写出来的一份实现，作文是使现实能够拥有同一生成身份的空间

本文从一个被日常教育遮蔽的问题出发：我们几乎每天使用“作文”这个词，却通常把文章、写作过程、作者身份、学习能力和评价结果混成一个对象。生成式人工智能只是使这种混合变得无法继续维持。

法庭文件检验提醒我们，痕迹可以支持来源归属，却不是来源本身；反应机理学提醒我们，完整的步骤序列仍然可能不是机理，同一终态也可能兼容多个竞争机理；知识证明思想提醒我们，一次成功输出不能自动证明主体拥有使输出成立的隐藏结构。三种约束共同迫使作文研究从“这一次已经写出了什么”进入“在受控扰动下，这个主体还能够生成什么”的空间。

据此，本文把生成见证定义为在原文不可用、条件有限扰动时仍能组织新文本关键意义关系并预测留出挑战的最小关系结构；把作文定义为共享同一可提取生成见证的一族可能文本，即文本在生成见证关系下的等价类。现实中的一篇文章只是代表元。

这个定义产生两个最重要的判别：同文可以异作，异文可以同作。它也重新排列AI写作争议：关键问题不再是有多少字符由AI产生，而是工具参与以后，关键生成关系究竟留在文档、留在外部系统，还是已经成为主体在新挑战中可以再次发动的结构。

因此，“何谓作文”的最终答案不应被压缩成“作文就是表达思想的文章”，也不能停在“作文是一个过程”。更准确的说法是：作文是一种生成身份。文章告诉我们一次发生留下了什么；作文要求我们进一步知道，使这次发生成为这次发生的关系，是否还能够在新的可能世界中继续组织文本。

如果这一理论最终成立，作文教育保护的主体就将发生转移。真正需要保护的，不只是那张交上来的纸，不只是那一篇被改得越来越好的文章，而是当纸被拿走、答案被撤掉、工具暂时退出以后，一个人仍然能够让什么新的东西发生。

（一）本文最终保留的四个理论贡献

第一，本文把作文的研究单位从“单一文本”移动到“多次可能实现之间的生成身份关系”。这个移动不是否认文本，而是把文本从对象本身降为代表元。只要这一点成立，作文研究的数据结构、AI归属讨论和教学评价都会发生变化。

第二，本文把“来源归属”和“生成归属”分开。一个文本可以在法律或规范上归某人署名，也可以在物理输入上由某人完成，但其关键生成见证仍可能外

置；反过来，字符执行可以大量外置而生成见证仍属于主体。这个区分使 AI 参与不再被压缩成简单比例。

第三，本文引入“同文异作—异文同作”的交叉预测。任何理论真正具有独立性，都应当能在同一材料上给出与近邻概念不同的判决。两类交叉案例正是这种决定性边界：如果未来实证根本找不到它们，本文的第二轴就失去必要性。

第四，本文把作文身份从一次回顾性解释转化为反事实挑战问题。过去研究主要从已经存在的文本、日志和分数推断；本文要求在预注册扰动下生成未见响应，并让候选 G 承担留出预测风险。这个方法论转向，即使最终“作文等价类”的强本体被否定，也仍可能为 AI 时代的写作研究留下一个可验证的新问题。

更重要的是，这四点不是互相堆叠的四个优点，而是一条连续逻辑：只有把文章视为代表元，来源归属与生成归属才可能分离；只有两种归属分离，“同文异作”和“异文同作”才会成为可观察对象；只有这种交叉对象存在，反事实挑战才有必要。任何一环在实证中失败，都应该向前回退，而不是靠增加概念继续维持理论。换言之，本文真正要争取的不是一个漂亮名称的存续，而是一种新的可区分事实是否存在：现实中是否真的有一些文本，在成品质量、写作过程和作者控制都十分接近时，仍会因为生成见证不同而在后续挑战中稳定分化；以及另一些表面完全不同的文本，是否会因为共享同一生成关系而表现出可重复的深层连续性。只有这两类事实被观察到，“作文等价类”才值得保留其强名称；否则，它就只能退回一种启发式框架，而不能继续宣称重新定义了作文的本体。

参考文献

- [1] Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365–387. DOI: 10.58680/ccc198115885.
- [2] Lo Sardo, D. R., Gravino, P., Cuskley, C., & Loreto, V. (2023). Exploitation and exploration in text evolution: Quantifying planning and translation flows during writing. *PLoS ONE*, 18(3), e0283628. DOI: 10.1371/journal.pone.0283628.
- [3] Nowacek, R. S., Leonard, R. L., & Rounsaville, A. (2024). *Writing Knowledge Transfer: Theory, Research, Pedagogy*. Parlor Press & The WAC Clearinghouse.
- [4] Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50(1), 1–73. DOI: 10.1111/jedm.12000.
- [5] Arnold, K. C., & Kim, J. (2025). Interaction-Required Suggestions for Control, Ownership, and Awareness in Human-AI Co-Writing. *Proceedings of In2Writing 2025*, 62–68. DOI: 10.18653/v1/2025.in2writing-1.6.

- [6] Formosa, P., Bankins, S., Matulionyte, R., et al. (2025). Can ChatGPT be an author? Generative AI creative writing assistance and perceptions of authorship, creatorship, responsibility, and disclosure. *AI & Society*, 40, 3405–3417. DOI: 10.1007/s00146-024-02081-0.
- [7] Rodrigues, T. V. (2026). The Apt Curation Model: An Epistemic Virtue Theory of AI-Assisted Authorship. *Philosophy & Technology*, 39, Article 39. DOI: 10.1007/s13347-026-01038-z.
- [8] Taylor, M., Bishop, B., Burkes, T., et al. (2021). Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach. NISTIR 8282r1. DOI: 10.6028/NIST.IR.8282r1.
- [9] IUPAC. (2025 online ed.). “mechanism (of a reaction).” *Compendium of Chemical Terminology (Gold Book)*, 5th ed. DOI: 10.1351/goldbook.M03804.
- [10] Goldwasser, S., Micali, S., & Rackoff, C. (1989). The Knowledge Complexity of Interactive Proof Systems. *SIAM Journal on Computing*, 18(1), 186–208. DOI: 10.1137/0218012.
- [11] Bellare, M., & Goldreich, O. (1992). On Defining Proofs of Knowledge. *Advances in Cryptology—CRYPTO ’92*, LNCS 740.
- [12] Huang, Y., & Wilson, J. M. (2021). Using automated feedback to develop writing proficiency. *Computers and Composition*, 62, 102675. DOI: 10.1016/j.compcom.2021.102675.
- [13] Dai, Z. (2026). Rethinking AI-assisted writing instruction: feedback literacy scripts, calibration training, and student writing development. *Frontiers in Psychology*, 17, 1829268. DOI: 10.3389/fpsyg.2026.1829268.
- [14] Wan, R., Gebreegziabhe, S., Li, T. J.-J., & Badillo-Urquiola, K. (2024). CoCo Matrix: Taxonomy of Cognitive Contributions in Co-writing with Intelligent Agents. arXiv:2405.12438.
- [15] Koo, R., Martin, A., Wang, L., & Kang, D. (2023). Decoding the End-to-end Writing Trajectory in Scholarly Manuscripts. arXiv:2304.00121.
- [16] Hamamra, B., Khlaif, Z. N., & Mahamid, N. (2026). AI literacy and mediated learner autonomy in ChatGPT-supported EFL writing: a qualitative study at a Palestinian university. *International Journal of Educational Technology in Higher Education*, 23, 19.
- [17] Haidemariam, T. (2026). After generative AI: authorship, labour, and cultural governance. *AI & Society*. DOI: 10.1007/s00146-026-03189-1.

附录 A：研究预注册最小模板

研究问题：在控制一般写作能力、文本相似度与成熟迁移指标后，从训练挑战中提取的生成见证是否能够预测留出挑战中的未提示关键关系？

问题域 Q：必须在实验前定义核心认知任务及不属于该问题域的边界。扰动集合 Δ：至少包含三个局部扰动，并明确哪些变化仍属于同一问题域。训练挑战：T1、T2。留出挑战：T3。候选生成见证 G：在看到 T3 前冻结。竞争模型：文本

相似度、一般写作能力、writing transfer、authorial control。主要结果：GWF 及其对 T3 的增量预测。失败条件：GWF 无增量预测、盲评一致性不足或 G 需要在看到 T3 后修改。

附录 B：版本、证据与人机协作声明

版本：V1.0，2026 年 8 月 17 日。

证据边界：本文的法庭文件检验、反应机理、交互式证明、写作过程、写作迁移、测评效率及 AI 辅助写作部分均依据公开学术或官方资料进行理论转译。跨学科转译不意味着来源学科已经认可“作文等价类”这一结论；“生成见证”“作文等价类”及其形式化、二维框架、证伪设计属于本文提出的研究假说，需要专门实证检验。

人机协作声明：本文在问题结构探索、跨学科资料检索、理论竞争检验、初稿组织与文字编辑过程中使用生成式人工智能辅助；最终理论主张、证据边界与文稿责任应由实际署名作者进行人工复核后再用于投稿、教学评价或政策实践。

篇幅统计：汉字约 20,018 字；非空白字符约 28,627 个。

第一章 补节 撤掉成品之后，还剩什么

（一）最近的祖宗与分离线

本章最强的一对推论是「同文可以异作，异文可以同作」。这一对推论有两位祖宗，原稿的七处分界里一位都没有。

博尔赫斯〈皮埃尔·梅纳尔，《吉诃德》的作者〉（1939）与丹托《寻常物的嬗变》（Danto, *The Transfiguration of the Commonplace*, Harvard UP, 1981）。梅纳尔逐字重写了塞万提斯的段落，两段文字完全相同，而叙述者论证第二段「几乎无限地更丰富」。丹托由此提炼出「不可辨别的对应物」（indiscernible counterparts），并用一场假想展览把它推到极端：七块视觉上完全相同的红色矩形画布，分别是《以色列人渡红海》《克尔凯郭尔的心情》《红场》与一块乔尔乔内打好底的备用布——它们的差别不在任何可见性质上。

这是「同文异作」在艺术哲学里的正主，而且早了八十年。

分离线在哪里。丹托的解决方案是语境：使两块红布成为不同作品的，是它们各自的产生史与所处的艺术世界；判断由外部关系给出，且这个判断是解释性的，不承担预测。本章的解决方案是反事实响应族：使两份相同文字成为不同作文的，是它们各自的生成结构在原文被撤走之后是否仍能重新组织同一批关系；判断由一次可执行的操作给出，而且它承担预测——本章第十四节给出的挑战——提取一留出方案，要求研究者在看到留出任务的结果之前先写下预测。

判决性对照：丹托的框架里，「两块红布是否同一作品」没有任何实验可以裁决，只能靠对艺术世界的正确描述；本章的框架里，「两份同文是否同一作

文」有一个会失败的实验——若甲乙两人在文本完全相同的前提下，撤稿后的扰动挑战中不出现稳定的生成结构分化，本章即被削弱。丹托问「凭什么它们不同」，本章问「怎样把这个不同测出来」。两者不冲突，但本章若不点名丹托，读者会正当地认为它在没有署名的情况下重述了一个四十五年前的论证。

第二位是现行制度而非理论：**ICMJE** 的四条作者资格与 **CRedit** 贡献者角色分类。医学期刊编辑国际委员会规定，作者资格须同时满足四条——对构思或设计、或对数据的获取分析解释作出实质贡献；起草或对重要智识内容作批判性修改；核准最终版本；并对全文的完整性负责。关键在于：这四条里没有一条计算字数。一个写下全部文字而未参与任何智识判断的人，按此标准不是作者，应列入致谢。**CRedit** 进一步把「作者／致谢」这个二元格拆成十余个具体角色。

也就是说，「作文身份不由谁敲了多少字决定，而由谁作出了实质判断决定」这条主张，在科学出版领域已经是执行了三十余年的成文规则。本章、第五章与第七章共享这条主张，三章一并零命中。

分离线：**ICMJE** 判的是资格与责任，本章判的是同一性。**ICMJE** 回答「谁有权署名」，它的裁决对象是人与人之间的贡献分配，而且是申报制——由参与者自己声明。本章回答「这两份文字是不是同一次作文」，裁决对象是文本族与生成结构之间的关系，而且明确拒绝申报制：本章第二十节写明，生成见证不能由写作者自陈获得，只能由撤稿后的响应获得。判决性对照：一位完全符合 **ICMJE** 四条的作者，其文本在本章的留出挑战中仍可能分化为两个生成身份（他确实作出了实质判断，但两次判断不共享同一生成结构）；反过来，一位按 **ICMJE** 应列入致谢的人，其生成见证可能与正文完全一致。两套判决可以相反，因为它们裁的是不同的东西：一个裁功劳与责任，一个裁对象同一性。

（二）与第二章的对切

本章与第二章同形：两章都主张身份不由静态特征给出，须由扰动后的响应裁决。分离线不在方法，在被扰动的是谁。

本章撤走的是文本，留下的是人。把原稿封存、禁止调用原会话，让写作者在新条件下重新组织同一批意义关系，看那套关系是否还能被他生成出来。读数落在写作者身上，因此它能分辨「同文异作」——两份完全相同的文字，背后是两个不同的生成主体。

第二章不动人，动的是文本。在声明的判别分辨率下对这一份文字施加预注册的保相与破相变换，看宏观组织关系是否跃迁。读数落在文本对象身上，因此它能分辨「改到什么程度它就不再是这一篇」——哪怕从头到尾是同一个作者。

判决性对照：一个写作者可以在扰动后完整重建关系（本章判为同一生成身份），而他新写出的那份文本在关系探针上与原文分属两相（第二章判为两个对象）。两个判决都成立且不矛盾——它们裁的本来就不是同一个东西：一个裁谁，一个裁哪一篇。反向的交叉同样存在：两个人复现出同一相的文本（第二章判同一），而撤稿后只有一人能重建（本章判不同）。

这条对切同时是两章的共同证伪条件：若上述两种交叉在真实样本中从不出现，即人侧读数与文本侧读数长期近乎一一对应，那么两章之一是冗余的，应当合并。

（三）本章欠账

其一，本章没有自跑任何数据。第十六节的方向检查用的是两项他人已发表研究，且本章诚实写明它们只支持较弱前提、不能运行核心检验。这是九章里唯一一章的证据全部来自二手，读者应据此调低对本章经验部分的权重。

其二，生成见证的编码一致性未解。本章第六条证伪自己提到「若不同盲评者对 G 的关系约束无法达到基本一致，且只有提高抽象程度才能提高一致性，则生成见证缺少可测量边界」——这条尚未检验，而它是最可能杀死本章的一条。

第二章

不动它，就不知道它是不是它

何谓作文：从“成品文本”到“篇章定相体”

——凝聚态物理·相变与序参量 × 地层学·地层单位与时间序列

× 范畴论·Yoneda 与关系性对象的二阶碰撞

学科通融专栏·二阶碰撞范文（What1）

摘 要

“作文是什么”常被当作一个已经被教材、课程标准和写作研究回答过的问题：作文是语言表达的成品，是作者经验与思想的外化，是面向读者的交流活动，是规划、转译与修改构成的认知过程，也是受体裁、情境与社会关系塑造的行动。然而，这些回答仍难以解决一个更基础的对象身份问题：当一份文字经历删改、改序、换例、换句、换语体、AI 润色、多人协作乃至大幅重写时，我们凭什么说它仍是“同一篇作文”；反过来，两份材料、主题与句式高度相似的文字，又在什么条件下已经是两个不同的作文对象？本文依照 SDE 多学科通融创新法 v3.9 的 What 型三维度工序，选取凝聚态物理中的相变与序参量、地层学中的地层单位与时间序列、范畴论中的 Yoneda 引理与关系性对象，分别占据显露、路径和关系场三个位置。三家形成斜对立：相变理论允许大量微观成分改变而宏观相保持不变；地层学表明材料相似不能抵消边界、上下序位、缺失记录和相关关系对单位身份的约束；Yoneda 则把对象身份从“内部属性清单”推进到其与其他对象之间的完整态射剖面。二阶碰撞由此提出新对象“篇章定相体”：作文不是一堆语言材料，也不是一次写作过程的终点，而是在声明的判别分辨率下，由有界时序、宏观组织关系与变换响应共同确定的关系相。其身份不靠某一固定特征决定，而靠一组预注册的保相变换与破相变换进行反事实裁决。为使概念可检验，本文提出单一读数“篇章定相率”（CPR），并给出变换设计、关系探针、盲评编码和失败判据。对 ASAP++、ELLIPSE 等公开作文评分资源进行字段审计后发现，现有数据主要保存成品文本与分数/特质分数，缺少版本时序、预注册变换和变换后的关系响应，因此目前无法直接计算 CPR；这一不利结果被原样保留。本文进一步与写作过程论、修辞情境、体裁社会行动、后过程/网络写作、篇章连贯与修辞结构理论，以及同产线先前的“跨端复成体”逐一划界，并讨论其对作文教学、修改、评价与 AI 协作写作的重写意义。

关键词：作文；语文教育；二阶碰撞；篇章定相体；相变；地层单位；Yoneda；篇章定相率

二阶碰撞工序摘要

位置	取源	核心判决	迁入作文后的硬问
S	凝聚态物理：相变与序参量	微观成分可大量变化而宏观相保持；越过条件可出现离散跃迁	改多少字不等于改少作文：哪些关系保持时是同一对象？
D	地层学：地层单位与时间序列	材料相似不足以定单位；边界、序位、缺失与相关共同承重	相同材料放在不同位为何已不是同一篇章能？
E	范畴论：Yoneda 与关系性对象	对象可通过其与其他对象/映射的关系剖面被刻画	不看内部属性清单能否用变换响应辨认这篇文？

核心命题：作文是一种“篇章定相体”：在声明的判别分辨率下，一个有界语言—版本序列形成宏观组织关系，并在预注册的保相变换中保持、在破相变换中跃迁；其身份由边界/序位、宏观关系与变换响应联合裁定。

一、重新开题：作文教育会评作文，却未必能裁定“同一个作文”

“作文是什么”看上去不像一个有必要继续研究的问题。学校有作文课，考试有作文题，教师有评分标准，研究领域有过程写作、认知写作、社会文化写作、体裁教学、修辞学和自动作文评价。只要一份文字具有某种题意回应、篇章组织与语言表达，我们通常就把它收进“作文”这一对象集合，然后讨论好不好、怎样写、为什么写不好。问题恰恰埋在这个过于顺手的对象化动作里：我们把“眼前有一份成品文本”与“一个作文对象已经成立”视为近似同义，却很少给出对象身份的反事实判据。

对象身份一旦进入修改过程，这个漏洞立刻显现。学生把开头换掉、删掉一段、把倒叙改成顺叙、将一个事例替换成另一个事例，教师仍可能说“这还是你的那篇作文”；但如果学生只保留题目和主题，让生成式人工智能重写全部论证与叙事关系，我们又可能说“这已经不是原来那篇”。两种判断都很自然，却依赖隐含直觉：究竟什么改变不伤身份，什么改变会跨界？字数比例不能解决，因为换掉 5%的关键关系可能使整篇改相，换掉 60%的表面措辞却可能仍保持同一组织。作者身份也不能解决，因为同一作者可以写出两个完全不同的版本，多人协作又使作者边界变得模糊。

同样的困难出现在“不同文本是否其实同一对象”的另一侧。两名学生可能都写“雨夜送伞”，情节节点、感情色彩、结尾升华甚至句法模板高度相似；如果作文只是材料与主题的集合，它们近乎同一个对象。但教师又会直觉地认为，其中一篇若把“被照顾”组织成依赖关系，另一篇把同一材料组织成对过去误解的修正，两篇已不是同一个作文。可见作文身份既不等于物料清单，也不等于题目标签。

这使 What1 获得了一个比“作文是表达”更尖锐的问法：我们不是再找一个宏大定义，而要找到一条能够区分两种看起来相同、实则不同的对象状态的判据。具体地说，本文要回答：在什么条件下，一份语言材料已经形成一个可以保持、变形、破坏和重新识别的整体；这个整体怎样跨越表面改写仍保持身份，又怎样在关键关系被改动时发生离散跃迁？如果不能回答这一点，“作文是什么”就仍停留在名词解释，而没有成为对象论。

二、题型判别：What 题要的不是“怎么写”，而是一种可辨别的对象

按照 v3.9 工序，第一步不是找学科，而是先锁定答案形状。本文问的是“何谓作文”，其目标不是得到从构思到成文的路线，也不是找到导致作文生成的单一驱动力，而是得到一个能够把“仍是这篇作文”与“已经不是这篇作文”分开的对象。因此只能采用 What 型三维度碰撞。三家必须分别站在显露状态、组织路径与关系条件三个位置；若三家都在讨论文本内部结构，哪怕观点彼此相反，也只是同一轴上的辩论。

本文冻结三句占位句。S 位占位句是：“对作文来说，形成某种可识别的整体状态才算数。”D 位占位句是：“对作文来说，材料进入何种有边界的先后序列才算数。”E 位占位句是：“对作文来说，它在与其它版本、读者问题和变换操作的关系中怎样表现才算数。”这三句话分别指向凝聚态物理、地层学和范畴论。选源的理由不是“作文像物质、像岩层、像数学对象”，而是三门学科每天都在处理对应判决：什么时候许多局部成分组成一种宏观相，什么时候一组材料能被定义为一个正式地层单位，什么时候一个对象能通过与其他对象的映射关系被刻画。

What 型最危险的失手是把三门学科当作三条比喻。本文因此规定：后文凡出现“相”“地层”“态射”，都必须转写成作文可清点的字段；不能落到字段上的类比不进入核心命题。相变不是“作文灵感突然发生”，地层不是“文章一层层写出来”，Yoneda 也不是“作文要考虑关系”。这三种漂亮但空泛的相似性全部舍弃，只保留各学科已经建立的对象身份判据。

三、闸零：三门源为何可以进场

v3.9 新增的闸零先查“源的未吸收度”。判的不是学科名称有没有在写作研究里出现，而是这门学科针对本题给出的那一句答案是否已经被作文理论整段吸收。检索写作过程、修辞情境、体裁、篇章连贯、复杂系统、网络写作和自动作文评价等文献后，可以看到“写作是过程”“写作是社会行动”“文本由关系组织”“写作在网络中发生”等思想已经十分成熟；因此，若本文只是说作文是动态的、社会的、关系性的，创新当场失败。相反，未检出作文教育把“保相/破相变换的反事实响应”“地层单位式的边界—序位—缺失共同定身份”或“Yoneda 式关系剖面作为版本同一性判据”建立为系统评价程序。三家在本题上的核心答案据此判为干净，而非因学科陌生就自动放行。

第二道闸是理论厚度。凝聚态物理的相变理论有 Landau 序参量、对称性破缺、临界现象、重整化与普适性等奠基层，也有拓扑序和广义对称等“超越传统 Landau 框架”的内部反例传统。Kadanoff 指出，一阶相变可表现为统计变量的离散跳变，不同相变具有不同序参量；而近几十年的拓扑相研究又显示，并非所有相都能由简单局域序参量穷尽。一个成熟理论既能给出判据，也能指出自己的判据在哪里失效，这正是“厚”的标志。

地层学同样具有厚重的判据传统。国际地层委员会的指南把岩性地层、生物地层、磁性地层、年代地层等分为不同单位体系，并明确指出一种属性发生改变不保证另一种属性同时改变；同一岩性也不能自动推出同一时代。正式地层单位需要边界、层型、相关与命名程序，岩层记录中的间断、不整合和缺失本身还可能构成重要证据。也就是说，地层学很早就被迫处理“材料像不像”与“单位是不是同一个”之间的分离。

范畴论的理论厚度更无疑问。Eilenberg 与 Mac Lane 建立范畴、函子和自然变换的基本语言，Yoneda 引理随后把“对象如何被全部关系看见”推到核心位置。Yoneda 嵌入的全忠实性意味着，在适当的范畴中，对象之间的态射信息可以无损地嵌入到函子层面的关系表示中。本文并不宣称作文天然构成某个已经证明满足全部公理的范畴；我们只取其判决纪律：不要先问一个对象“内部装了什么”，而要问它对一组结构保持映射怎样响应，并以自然一致的关系剖面来约束身份判断。

四、S 位：相变与序参量——作文身份首先不是“有多少材料”，而是“形成了什么宏观相”

凝聚态物理给本题的第一记冲击，是微观材料与宏观状态之间不存在一一对应。水分子的种类并不会因为水结冰而更换，铁磁体的原子也不会在磁化时换成另一批原子；真正发生改变的是大量自由度之间的集体组织，使系统进入可宏观区分的相。相

变因此提供一种对象身份思想：组成部分可以连续改变，而整体在某个参数区间内保持同一类别；一旦越过临界条件，宏观状态可能出现离散变化。

迁入作文，这立刻否定“改了多少字就改了多少作文”的线性直觉。一篇叙事文可以把“走进医院”改成“推开诊室门”，把天气从阴天改成大风，把一段对话由直接引语改成间接引语，只要这些变换没有改变核心的关系组织，读者对“谁误解了谁、什么证据改变了判断、转折发生在哪里”的回答可以保持稳定。反过来，只改一句“他没有来，是因为不在乎”到“他没有来，是因为不愿让我看见他的害怕”，字数变化很小，却可能把因果相、人物关系和整篇情感方向一起翻转。作文对象因此具有非线性。

序参量的价值不在于给作文找一个漂亮的物理名词，而在于逼我们寻找“少数能够表征整体组织状态的宏观变量”。对于不同文体，这些变量不应机械相同。叙事文可能包括核心因果方向、行动者责任分配、视角中心、转折位置与结局所关闭的悬置；议论文可能包括中心主张、证据依赖、反方处理、概念边界与结论开放度。它们都不是词频、句长或修辞数量，而是大量局部句段共同实现的关系变量。

但是，相变理论自身又不允许我们把“找到一个序参量”神化成万能答案。传统Landau框架以对称性破缺和局域序参量极其成功地解释了大量相变，而拓扑序等现象表明，某些相的区分需要非局域或更广义的结构。这个内部反例对作文尤其重要：如果我们预先规定所有作文都必须由同一组“中心—材料—结尾”变量刻画，新理论会立即复制旧评分表的毛病。本文因此把序参量理解为“在声明的判别任务中预注册的一组最低组织关系”，而不是永恒不变的作文五要素。

S位最终给出一句硬判断：作文对象的身份不应由微观语言材料的保留比例裁定，而应由宏观组织关系是否保持同一相来裁定。这个判断会与后两家发生冲突。地层学将指出，宏观关系相同也可能因边界和时序位置不同而不是同一单位；Yoneda还会指出，只观察一组内部序参量仍不足以穷尽对象身份，因为对象在外部变换下的关系响应同样承重。

五、D位：地层单位与时间序列——同样的材料，放在不同边界与序位中可以不是同一个对象

地层学处理的是一个极其现实的困难：岩石不会为了方便分类而整齐排列。某一层剖面可能缺失，某些岩性横向渐变，不同地点同一时代的材料完全不同，相似岩性又可能在相隔遥远的年代重复出现。因此，地层单位的身份不能靠“材料像不像”单独决定。边界在哪里、上下关系如何、与邻近单位如何相关、缺失区间意味着什么，都是正式判定的一部分。

国际地层指南把“层型”作为正式单位的参考标准，并要求明确界定边界。更有启发性的是，不同地层分类体系的边界不必重合：岩性变化、生物群变化、磁性反转

和年代边界可能发生在不同位置。这条原则对作文教育具有直接破坏力。我们常把段落、材料、主题、时间顺序当成一组天然一致的“文章层次”，实际上它们可能彼此错位。一个段落在版面上结束，并不意味着一个论证层完成；叙事时间跳跃也不等于人物关系发生相变；一句转折词出现，更不等于真正的意义边界已经建立。

地层学还教会我们认真对待“没有写出来的东西”。岩层记录的间断并不只是缺陷，它可能记录侵蚀、沉积停止和时间缺口。作文中的省略、空白、跳接同样不应一概按“信息不足”处理。一个叙事在父亲抬手欲言处直接跳到多年后的葬礼，缺失的谈话可能正是作品的结构性空隙；若教师要求“把中间发生的事写清楚”，反而可能填平关键边界。D位因此把作文组织从“连续填满”改写为“有界序列”：存在什么固然重要，什么在何处缺失也可能承担身份工作。

更关键的是相关（correlation）问题。同一地层单位在不同地点会发生材料变化，需要依靠多种证据相关；相反，两个地点即使岩性相似，也不能据此直接认定为同一时间单位。作文版本亦如此。同一作文在学生初稿、教师批注稿、AI润色稿和定稿之间可能发生大量表面变化，若核心边界和关系序列可追踪，它们仍可能属于同一版本谱系；反之，两份套用相同模板的文本即使材料与句式相似，若其关键关系边界不对应，就不应因表面相似而视为同一对象。

D位给出第二句硬判断：作文不是无边界的语言流，而是一个由边界、序位、缺失与相关共同确定的有界序列单位。这个判断反过来攻击S位：只看整体相会遮蔽“相似状态发生在不同序位”的差异。它也攻击E位的一种过度抽象：如果只谈关系而不保存版本发生的时间和边界来源，任何后见之明都可以把关系重新解释得天衣无缝，失去可追溯性。

六、E位：Yoneda与关系性对象——对象不是属性盒子，而是其可保持映射的关系剖面

范畴论把本题推到最不熟悉也最关键的一步。传统对象观容易把一篇作文想成一个“盒子”：里面装着主题、材料、结构、语言、情感、修辞等属性，盒子内的属性组合决定它是什么。Yoneda式视角不从盒子内部开清单，而问这个对象与其他对象之间所有合适的箭头怎样组织。严格的Yoneda引理属于范畴论，其数学结论不能直接等同于语文判断；但它提供的认识纪律具有可迁移性：对象身份可以通过它对一组结构保持关系的响应被约束，而不必依赖一个先验的内部本质清单。

迁入作文，最有价值的不是“作文与读者有关系”这一早已熟悉的结论，而是把修改操作本身改造成身份探针。我们不再只盯着原文问“它有什么特征”，而是把原文送进一组预先声明的变换：换掉外围措辞、重写句法、调整版式、替换功能等价的例子、移动非关键描写；再用另一组操作反转核心因果、拆掉论证依赖、移走关键边

界、交换行动者责任、改变结论对前文的回写。然后观察固定关系问题的答案如何变化。一个作文对象的“关系指纹”由这一整组响应形成。

这一点解释了为什么单篇静态评分难以裁定对象身份。评分表通常只在一个状态上读取若干特征：组织 3 分还是 4 分、词汇是否丰富、中心是否明确。关系性判决则关心导数而不是单点：当我们沿某个允许方向推动它时，什么保持；沿另一个方向推动时，什么必须改变。两个静态分数相同的文本可能拥有完全不同的响应曲线：一篇对外围改写稳定、对关键关系反转敏感；另一篇无论怎么删换都还能让读者说出同一套“成长”“母爱”“坚持”，因为其理解主要由题目和文化脚本预装。

Yoneda 还迫使我们区分“关系多”与“关系足以识别”。社交网络、互文性、对话性和后过程理论都强调写作处在关系中，但“处在关系中”仍可能只是世界观。本文需要的是判决程序：哪些关系被纳入探针集，哪些变换是允许的，哪些是破坏性的，响应如何编码，何时判同、何时判异。只有把关系变成可复核的映射族，关系性才从哲学姿态变成对象身份工具。

E 位于是给出第三句硬判断：作文身份不由内部属性清单单独裁定，而由其在一组声明的结构保持/结构破坏变换中的关系响应剖面共同裁定。它与 S 位的冲突是“静态宏观量是否足够”；与 D 位的冲突是“历史边界是否必须进入身份”。三者不能简单加法，因为各自都想把自己的判据当成最后法院。

七、第一阶碰撞：三家不是互补，而是相互取消对方的“单独足够”

若只保留凝聚态物理，我们会得到一种诱人的作文观：只要若干宏观序参量相同，表面怎么变都算同一相。但地层学立即提出反例：同样的岩性和相似的整体性质，处在不同边界、不同上下关系与不同年代位置时可以属于不同单位。迁入作文，同样的“误解—证据—理解”结构如果被放在不同的叙事层级，可能分别承担主线和插叙功能；整体关系向量相似，却不是同一个版本对象。

若只保留地层学，我们会把版本谱系、边界和序列保存得很好，却可能把作文身份过度历史化：凡是从同一初稿连续修订而来，就倾向视为同一单位。然而 E 位指出，连续来源不保证结构同一。一个版本可以在没有明显断档的修订中逐步换掉全部关键关系，最后虽具有完整版本链，却在关系剖面上已经变成另一对象。连续谱系不是身份充分条件。

若只保留 Yoneda 式关系剖面，我们又可能陷入“只要当前关系响应相同就算同一对象”的抽象主义。D 位提醒：某些响应相似可能来自不同历史层的偶合，来源边界和缺失信息决定我们能否把两个状态合理相关。S 位也提醒：并非所有关系都同等重要，真正构成宏观相的通常是少数集体变量，而不是把所有局部箭头无限收集。

三家的冲突因此构成 What 型所需的斜对立。S 说“微观差异可以被宏观相吸收”；D 说“相似材料不能跨过边界与序位直接合并”；E 说“无论材料还是宏观量，都必须放到变换关系剖面中才能检验身份”。没有任何一家可以当裁判判另外两家输赢，因为它们裁定的是对象身份的不同缺口。真正的出路不是选一家，而是造一个新对象，使宏观相、时序边界和变换响应同时进入存在条件。

八、混沌碰撞：从 27 组交叉中留下九个有效涌现物

按照二阶碰撞工序，三家各抽出一个主题判断与三个承重点，再做跨家交叉。这里不把 27 对全部写成修辞性排列，而只保留能够改变作文判决的有效涌现。第一，序参量×地层边界得到“有界相”：宏观关系只有在声明的篇章/版本边界内才有身份意义，不能把不同阶段的关系平均成一个状态。第二，序参量×缺失记录得到“负空间序参量”：某些有意义的空白不是信息零值，而是通过阻止某种解释进入而参与整体组织。第三，序参量×相关原则得到“表面同相警戒”：主题、词汇和情节相似不能替代版本相关。

第四，宏观相×Yoneda 得到“变换定相”：一个所谓序参量只有在在一组变换下表现出稳定/跃迁，才不是评价者事后命名。第五，边界层型×Yoneda 得到“边界探针”：把关键句段跨边界移动，可以直接测试边界是否真正承担组织工作。第六，缺失记录×Yoneda 得到“空白探针”：补写一个省略区间，如果核心关系反而变弱，说明缺失本身属于对象结构。

第七，“超越 Landau”的内部反例×多重地层分类得到“分辨率声明”：不存在一套对所有作文、所有文体、所有教学目的都有效的唯一序参量；任何定相必须先声明判别层级和探针族。第八，正式单位修订×关系剖面得到“可追溯改相”：版本可以连续演变，某一步一旦跨过关系跃迁阈值，就应记为新对象，而不是用“还是原稿修改”抹掉事件。第九，层型优先×关系剖面得到“范文降权”：范文不应只作为表面特征模板，而应成为一套可变换的参照对象，让学生观察哪些改动保持身份、哪些改动改变相。

九个涌现物中有五个可命名候选：篇章相态、有界序列体、关系指纹体、相位可辨体、篇章定相体。前四个分别仍偏向某一家：篇章相态过度物理化，有界序列体偏地层，关系指纹体容易被一般网络/篇章关系理论吸收，相位可辨体只说可分而没有历史边界。只有“篇章定相体”同时要求：先有有界时序，再有宏观关系，再通过变换响应裁决其是否形成稳定相。因此它进入下一道候选近邻间。

九、第二阶：推翻三家的共有前提——不存在脱离判别分辨率的“唯一身份证”

三家表面上矛盾，深处却共享一个更隐蔽的前提：对象最终总能被某一种足够强的表征单独结算。经典相变想用序参量区分相；正式地层单位用层型和边界将单位固定；Yoneda用完整的表示函子关系剖面把对象嵌入函子范畴。它们的技术内容不同，却都诱使迁移者相信：只要找到“那个最对的读数”，作文就能获得不依赖判别任务的永久身份证。

推翻材料恰好来自三家自身。凝聚态物理的拓扑序和广义对称传统表明，传统局域序参量并非所有相的普遍语言；地层学明确承认岩性、生物、磁性和年代等单位边界并不必重合，不存在一套单位同时表达全部地质属性；Yoneda的成立也不是“任何随便的关系都能识别对象”，而是在明确范畴、对象与态射规则后才有精确结论。换言之，身份判据从来隐含着—个先决声明：我们在什么对象域里、允许哪些变换、以什么分辨率讨论“同一个”。

这一步是二阶碰撞真正的转折。作文教育长期争论“主题是不是灵魂”“结构是不是骨架”“作者意图重不重要”“读者理解算不算”，很容易把这些问题当成寻找唯一身份证的竞争。本文取消这场竞争：同一份文字可以在不同判别分辨率下保持或改变身份。把叙事的核心因果关系作为分辨率，语体大改仍可能是同一作文；把声音与叙述姿态作为教学目标，语体大改就可能已经改相。关键不在谁的标准更本质，而在判别任务是否事先声明、变换族是否可复核。

因此，新对象必须带着“分辨率”出生。它既不是相对主义，因为一旦分辨率、关系探针和变换族被预注册，同一判断就可以被重复裁决；它也不是本质主义，因为没有—组跨所有文体、情境和教学目的永远有效的作文核心。对象身份被重新写成“在声明尺度上的稳定等价类”。这正是三家任—家单独不能给出的第四对象。

十、新对象 Z：作文是“篇章定相体”

本文把二阶涌现出的新对象命名为“篇章定相体”（compositional phase-object）。所谓“篇章”，强调对象必须有可追踪的语言边界与时序；所谓“定相”，强调它不是凭作者或教师宣告“中心是什么”，而要通过受控变换检验哪些关系稳定、哪些关系一旦被触动便发生跃迁；所谓“体”，强调边界、关系相与响应剖面不是三个评分项目，而是同一对象的共存条件。

给出核心定义：作文是在—个被声明的判别分辨率下，由有界语言—版本序列形成宏观组织关系，并能在预注册的保相变换中保持这些关系、在预注册的破相变换中按规定发生关系跃迁的关系相对象。其身份由“边界/序位+宏观关系+变换响应”共

同确定，而不由字词保留比例、作者身份、主题标签、静态评分或任何单一内部属性决定。

定义中的每个词都承担划界责任。“声明的判别分辨率”防止把一个评价目标偷换成全部作文的本质；“有界语言—版本序列”防止把文本从其修改历史和结构边界中抽空；“宏观组织关系”防止把身份退化为词句清单；“预注册”防止看到结果后再挑对自己有利的修改；“保相/破相”提供反事实；“按规定发生跃迁”要求破坏操作真的触到核心关系，而不是只让文字变差。

这个定义特别允许一种过去难以表达的事实：作文的表面可以高度不稳定而对对象身份稳定。学生把全部句子重新说一遍、把语言从稚嫩改得简洁，只要关键关系在保相变换下仍保持、破相探针仍能击穿相同位置，两个版本可判为同一篇章定相体。反过来，文本可以表面极其稳定而对对象已经改变：只在一个关键因果句中把“因此”背后的理由改掉，就可能让关系探针整体翻相。

这也解释模板作文的特殊困境。模板作文常具有很高的表面稳定性：删一段、换一个人物、换一个例子，主题仍然自动回到“坚持”“母爱”“成长”。传统评价可能把这种稳定理解成“中心明确”；定相框架却会问：若所谓关键关系被反转或删除，读者对核心组织的判断是否真的变化？如果几乎不变，说明文本并没有形成一个对关键关系敏感的对象，而只是依靠题目和文化脚本维持同一回答。稳定不等于定相；定相同时需要对该稳定结构的特异性破坏保持敏感。

十一、形式化最小模型：从“像不像”改成“在变换下是否仍是它”

为了防止“篇章定相体”沦为新隐喻，本文给出一个最低形式化模型。令 W 表示某一作文版本， B 表示其边界与版本序位记录， $\Phi(W)$ 表示在当前判别分辨率上预注册的核心关系向量， T 表示一组允许施加于 W 的变换。关系向量不要求使用机器学习特征，可以由五个到七个可人工编码的问题构成，例如：核心判断是什么；哪一个关系解释了主要变化；关键证据支持哪一项主张；行动/责任如何分配；决定性转折位于何处；结尾关闭或保留了什么义务。

把变换分为两类。 $T+$ 是保相变换：理论上应当改变表面而不改变核心关系，如外围同义改写、功能等价事例替换、非关键描写换序、版式调整、保持命题内容的句法重写。 $T-$ 是破相变换：理论上应当改变至少一个预注册核心关系，如反转中心因果、交换主张与证据依赖、移除决定性边界、把关键层移动到你前提之前、改变承担责任的主体、用相反结论回写前文。所有变换必须在看读者结果之前写入方案。

对每个变换后的版本 $T_i(W)$ ，由独立盲评者回答同一组关系探针，得到与原文 $\Phi(W)$ 的关系相似度 $RelSim_i$ ，取值 0 至 1。对保相变换，希望相似度高；对破相变换，

希望相似度低。于是把两种方向统一成一个“期望响应得分”：若 T_i 属于 $T+$, $R_i = \text{RelSim}_i$; 若 T_i 属于 $T-$, $R_i = 1 - \text{RelSim}_i$ 。

$$\text{关系指纹: } Y_W(T) = \Phi(T(W))$$

这里的 Y_W 不是把作文强行宣布为严格意义上的 Yoneda 嵌入，而是“Yoneda 式关系指纹”的操作记号：同一对象通过一族变换后的关系向量被观察。只有当所用版本状态、变换复合和恒等操作被进一步严格定义成范畴时，才可以讨论更强的数学对应；本文不跨过这条边界。

十二、单一读数：篇章定相率 CPR 与四种诊断状态

本文只保留一个主读数，避免正文、证伪与赌注各用一套口径。篇章定相率 CPR (Compositional Phase Response rate) 定义为全部预注册变换的期望响应得分平均值。本文试行口径规定：至少使用 6 个变换，其中 $T+$ 与 $T-$ 各不少于 3 个；每个版本至少由 5 名不知道变换类别的独立读者编码；关系探针在实验开始前固定； $CPR \geq 0.70$ 记为“达到试行定相阈值”。0.70 不是自然常数，而是本轮预注册的试行阈值，未来校准可以修改阈值，但同一研究内不得事后改口径。

这一读数把“稳”和“敏感”放进同一个分数。一个优秀模板可能对所有改动都给出相同主题， $T+$ 得分很高，但 $T-$ 的 $1 - \text{RelSim}$ 接近 0，CPR 不会高；一个极其脆弱的文本可能任何改写都让关系崩掉， $T-$ 得分高而 $T+$ 得分低，同样不能通过。只有“该稳定的稳定、该改变的改变”同时出现，才构成定相。

为了教学诊断，可以把 CPR 的两个来源方向暂时展开成二维图，但不另立第二主读数。保相方向高、破相方向高，是目标格“定相体”；保相高、破相低，是“惰性模板”，表面强稳但关键关系缺乏特异性；保相低、破相高，是“脆弱堆叠”，一碰就散；两向都低，是“无相文本”，既没有稳定组织，也没有可定位的关键关系。二维图只用于解释 CPR 为何高低，正式比较仍使用 CPR 一个口径。

这一设计还产生一个反直觉判断：作文修改的目标不是让每次修改都“更好”，而是先识别哪些修改应当保持对象、哪些修改本来就应该生成新对象。如果学生改了核心立场，却被教师要求“尽量保留原文结构”，教育动作可能是在强迫两个相位不同的对象共用一个壳。相反，若学生只是语言润色，却被告知“这已经不是你自己的作文”，也可能把表面变化误判成身份跃迁。

$$CPR(W) = (1/N) \cdot \sum R_i, \text{ 其中 } R_i = \text{RelSim}_i \ (T_i \in T+) , R_i = 1 - \text{RelSim}_i \ (T_i \in T-)$$

保相响应	破相响应	诊断名	教学含义
高	高	定相体	外围可变、核心感；对象边界清楚
高	低	惰性模板	怎么改都像同话；缺少关键关系特异性
低	高	脆弱堆叠	稍作表面改写就散组织依赖未被稳固编码
低	低	无相文本	既无稳定关系，也可定位的破坏点

十二之二、三个反事实案例：同样的“改写”，为什么有时保相、有时改相

案例一是叙事作文。初稿写一个孩子长期把父亲的沉默理解为冷漠，直到一次比赛失利后才发现父亲并非不参与，而是一直以修理器材、查询路线、提前等候等方式参与；结尾保留一个复杂判断：理解父亲并不等于认同他的沉默方式。若把天气、比赛项目、父亲的具体职业和若干动作描写替换，只要“误读—证据—修正但不全盘洗白”这组关系保持，文本可以大幅重写而仍属同一相。这里表面材料的可替换性很高。

但若只改一句，把“父亲不知道怎样表达”改成“父亲认为表达关心会削弱孩子的竞争意志”，整个关系结构便可能改相：父亲的沉默从笨拙的情感表达变成有目的的教育策略，孩子的“理解”也从重新认识一个人变成重新评价一种教育观。自然段数、比赛情节、结尾句甚至大部分词汇都可以原样保留，却已不再是原来的作文对象。这个案例显示，文本同一性不能由重合字数或材料数量裁定。

案例二是议论文。原文主张“学校不应以单一排名代表全部学习价值”，三组证据分别说明不可比能力被压进同一尺度、排名反向塑造学习策略、评价结果被误当作能力本身。若将其中一个例子从体育换成艺术，只要它仍承担“不可通约价值被统一压缩”的证据功能，属于保相变换。若把中心主张改为“排名可以保留，但必须增加多个并列指标”，看上去只是从反对单一排名走向改进排名，实则对象发生重要跃迁：原文攻击的是“统一尺度的合法性”，新文接受统一评价框架，只要求尺度更多。两篇可能共享 90% 的材料，却不是同一关系相。

案例三是开放性叙事。原文没有明确说明“她为什么在车站没有回头”，只通过前文三处细节让读者在“愤怒、决绝、保护自己”之间保持张力。传统“写具体”教

学可能要求补上一段心理活动。定相实验却可以把“补写原因”当作破相探针：若补写以后大量读者从多分支解释收缩到单一路径，说明空白本身就是关系结构的一部分。缺失不是零信息，而是对允许解释集合的边界控制。这里地层学关于间断记录的启发被转写成可观察的作文现象。

三个案例共同说明，保相/破相不是由操作名称决定。删除一段有时是 T+，有时是 T-；换例有时保相，有时改相；AI 重写有时只是表面变体，有时跨越对象边界。唯一的合法的办法是先声明判别分辨率与核心关系，再给操作定方向。看完结果后才说“原来这句很关键”，属于后见拟合，不能进入 CPR。

这也解释为什么本文不把“关键句”当成固定概念。关键性是一种反事实属性：某句是否关键，要看把它改变以后对象关系怎样响应。一句非常优美的句子可以完全不关键，一句语法普通的“但是我没有原谅他替我做决定”却可能承担整个伦理边界。作文教学因此从寻找“金句”转向寻找“相变点”。

十二之三、研究协议：怎样把定相实验做成可重复而不是教师直觉的新包装

第一步是样本冻结。研究者必须保存原始版本 W0，并记录其来源、时间和任务情境。若文本已有多个自然修订版本，应保留每一版而不是只留下定稿；但自然版本不能直接代替 T+与 T-，因为自然修改往往同时改变多个变量。实验版本需要在 W0 上单独施加受控操作，以便定位响应。

第二步是分辨率声明。研究问题只能选择一个主要身份层，例如“论证依赖”“叙事因果”“视角与责任分配”或“文学声音”。声明以后，研究者写出五个关系探针及编码规则。以论证依赖为例，可以问：中心主张是哪一项；最关键证据支持哪一项；文章承认的最强反方是什么；反方被怎样处理；结论是否改变了开头概念边界。每个问题都应有可枚举的编码，而不能只写“整体感觉是否一致”。

第三步是成对变换。每个 T+最好有一个在长度、位置和语言破坏程度上相近的 T-，避免 CPR 只是测“改得多还是少”。例如同样替换 80 字：T+替换一个功能等价例子，T-替换承重例子使证据转而支持另一个主张。又如同样移动一个段落：T+移动两个外围描写之间的次序，T-把结论性解释移到关键证据之前。操作强度相近，方向不同，差异才有判决意义。

第四步是盲评与独立性。盲评者不能知道哪个版本是“故意破坏”的，也不能看到研究者的预期标签，只回答固定关系问题。若同一读者连续看到原文和六个变体，记忆会造成依赖，最好采用分组设计或随机间隔，并确保每个变体有足够独立判断。课堂练习可以简化，但正式研究需要报告读者分配方案。

第五步是关系相似度。本文建议最简单的编码优先：每个关系探针完全一致记 1，部分一致记 0.5，不一致记 0；五项平均得到 RelSim。若未来使用向量语义或大模型辅助编码，必须先用人工双盲样本验证模型没有把词面相似误当关系相似。新理论最怕的恰恰是用一个高科技静态相似度把差分对象再次压平。

第六步是失败报告。若某个 T-版本因为语病过重导致读者无法理解，应标记“变换失效”并从预注册替补池取一个操作，而不能把低相似度直接算成成功破相；若某个 T+意外改变核心关系，应原样计入低分，因为这说明研究者错误估计了对象脆弱点。CPR 的价值不在于证明教师预感正确，而在于让错误预感留下数据。

第七步是跨分辨率复测。对于同一篇作文，可以在第二轮把判别层从“论证关系”换成“作者声音”，重新预注册探针。若某个 AI 润色版本在论证层 CPR 很高、在声音层 CPR 很低，这不是矛盾，而是精确显示了“同一”所依赖的尺度。只有报告这个尺度，作文身份讨论才从口号变成可交换的数据。

十三、近邻阐：九位占位者为什么不能 1:1 替换“篇章定相体”

第一位近邻是 Flower 与 Hayes 的认知过程论。它把写作从线性阶段改写为递归、目标导向的认知过程，极大改变了“作文是成品”的看法。但过程论主要回答写作者如何计划、转译和修改；同一过程模型可以生成相位不同的文本，同一篇章定相体也可能由完全不同的过程产生。过程记录不是本文的对象身份判据。

第二位是 Bitzer 的修辞情境。修辞情境强调迫切问题、受众与限制对修辞话语的发生具有构成作用。它能够解释为什么某段文字在一个情境中有效、另一个情境中失效，却不直接裁定两个修改版本在何种变换下仍是同一作文。本文保留情境作为判别分辨率的来源，但不把情境本身当对象身份证。

第三位是 Miller“体裁作为社会行动”。体裁不是形式清单，而是对反复社会情境的类型化行动。这与本文反对表面特征主义高度相容，也因此危险。分离线在于：体裁判断的是一类行动如何被社会识别，篇章定相体判断的是这一篇具体作文在版本变换中何时保持/改变身份。一篇文章可以始终属于同一体裁，却在核心立场反转后成为另一个作文对象。

第四位是后过程理论与网络化写作观。它们反对把写作封闭在个体头脑之内，强调解释、社会性、分布式行动和复杂网络。本文同意对象不应被作者内部状态单独封闭，却增加一个这些理论通常不承担的工程性问题：给定两个版本，必须怎样设计可重复的映射族，才能作出“同一个/不同一个”的裁决？“一切都在关系中”不能自动产生关系指纹。

第五位是 Halliday 与 Hasan 的衔接、van Dijk 的宏观结构以及 Mann 与 Thompson 的修辞结构理论。它们都能刻画文本内部关系，是本文最强的技术近邻之一。但静态关系图告诉我们“这个版本怎样组织”，并不自动告诉我们“哪些变换必须保持这些关系，哪些变换一旦发生就应判为另一个对象”。篇章定相体把关系结构置入反事实变换实验，判决单位从结构图变为结构图的响应。

第六位是 Bakhtin 的对话性与言语体裁。它揭示话语总是回应既有话语并预期他者回应，因此对象从来不是孤立自足。这与 Yoneda 式迁移在哲学气质上相近，但本文不以“开放对话”为终点，而要求事先列出可执行变换和固定关系问题，用差分响应裁定版本身份。

第七位是复杂动态系统写作研究。此类研究强调写作能力与语言特征的非线性、波动和涌现，与“相变”最容易产生表面同名。区别在于，本文并不把学生写作发展曲线视为物理相变，也不以复杂性指标解释能力成长；“相”只用于具体篇章对象在预注册变换下的身份稳定与跃迁。

第八位是自动作文评价与多特质评分。ASAP++等资源把内容、组织、词语、句子流畅、规范等特质量化，能够细致预测质量，但其单点特征仍不同于对象同一性。两份作文可得同样的组织分和内容分，却对同一破相变换有不同响应；评分回答“表现多好”，本文回答“是不是同一对象”。

第九位是同一产线先前提出的“跨端复成体”。那一理论把作文看成源对象—文本痕迹—读者复成共同构成的跨端事件，重点检验某个作者端对象是否通过文本被独立读者复成。本研究必须与它明确分开：篇章定相体不要求先知道作者原本想表达什么，也不要求不同读者复成同一源对象；它甚至可以用于匿名文本、AI 协作文本和没有可访问作者意图的古代文本。它问的是另一个问题——在声明的变换族下，这一具体篇章是否维持同一关系相。两者可以组合，但不能 1:1 替换。

十四、真跑一条：公开作文评分数据的字段审计，以及一个不利结果

v3.9 要求成文前至少把一条最便宜的证伪路线真正跑起来，并把不利结果原样保留。本文预注册的最低数据条件不是“有作文和分数”即可，而是五类字段同时存在：A.原始完整作文；B.可追踪的版本/修订时间序列；C.至少 3 个保相与 3 个破相变换，且变换类别在观察响应前已经标注；D.每个变换版本上同一组关系探针的独立盲评回答；E.原文与变换版本的一一对应 ID。缺任一类，CPR 都不能直接计算，不允许拿总体分、组织分或语言模型相似度替代。

随后审计公开作文评分研究中广泛使用的 ASAP/ASAP++与 ELLIPSE。ASAP++在原 ASAP 作文与整体评分基础上补充内容、组织、词语选择、句子流畅与规范等特质分，目的仍是多维作文评分；ELLIPSE 提供英语学习者作文以及整体/分析性熟练度标

注，便于研究写作能力与语言特征。它们都极具价值，却没有按本文协议为每篇作文配套一组预注册的保相/破相版本，更没有变换后的固定关系探针响应。版本时序也不是这类评分语料的核心字段。

因此，本轮“真跑”的结果是不利的：CPR 无法在这些公开数据上被直接计算。这个结果不能被写成“数据证明篇章定相体存在”；相反，它说明本文目前提出的是一个新的观测协议，而不是已经由现成评分数据验证的经验定律。若强行用组织分、语义相似度或主题一致性做代理，恰恰会把新对象重新压回静态特征主义。

这一失败同时提供了清晰的下一步数据需求：作文研究需要保存成品之外的“变换账本”。一份最小实验数据可以只取 100 篇作文，每篇生成 6 个严格控制的版本，让盲评者回答 5 个关系问题，总体成本并不要求大规模模型训练。真正新增的不是更多作文，而是每篇作文沿不同方向被轻推以后发生了什么。

本次字段审计与先前“跨端复成体”的数据缺口也不同。前者需要作者源对象、独立读者复成与关键痕迹消融；本研究即使完全不知道作者源对象，也可运行，只要有版本变换、边界记录和关系响应。因此，两种理论面对的是不同的缺失字段，这进一步证明它们并非同名换皮。

公开资源	已有字段（概括）	CPR 所缺关键字段	本轮结论
ASAP / ASAP++	完整作文、整体分； ASAP++增加内容、组织、 词语、句子流畅、规范等特 质分	预注册变换族、变换 版本、固定关系探针响应、 版本时序	不能直接计算 CPR
ELLIPSE	英语学习者作文、整 体与分析性熟练度/语言特征 标注	同上，尤其缺每篇作 文的成对变换响应	不能直接计算 CPR

十五、把新对象带回课堂：作文教学从“加特征”转向“做变换”

第一，写前教学不再只要求先定中心，而要声明本轮的判别分辨率。教师可以问：“这次我们主要训练什么不变？是叙事因果、论证依赖、人物视角，还是语言声音？”学生知道同一篇作文可以有多层身份，当前任务只对其中一层负责。这样既避免把“中心思想”神化为永恒序参量，也避免修改时所有东西同时评价导致无从下手。

第二，材料选择从“多不多、典型不典型”转向“能否改变宏观关系”。一则材料若删掉以后核心关系完全不变，它可能只是外围填充；一处只有十几个字的动作若删掉以后责任归属改变，它反而是高承载层。教师不需要教学生物理术语，只需问一个课堂问题：“把这块拿掉，整篇到底有什么关系会变？”这个问题比“这里够不够具体”更接近对象身份。

第三，结构教学从“开头—中间—结尾”转向“边界—层位—缺失”。学生用颜色标出每一次关系状态改变，而不是机械标自然段。某个自然段可以跨越两个意义层，一个意义层也可以跨三个自然段。再问每个边界两侧的条件是否可以交换：若两段对调后读者的核心判断完全不变，所谓结构层次可能只是版面分段；若一交换就使证据先于问题、结论失去前提，边界才具有结构性。

第四，修改教学从“把原文写得更好”改成双轨。保相修改用于训练表达弹性：不改变核心关系的前提下重写措辞、例子、句法和节奏，观察对象能否保持。破相修改则有意改变一个关键关系，要求学生承认“现在我已经在写另一篇作文”，重新命名中心和结构。把两类修改分开，可以消除一种常见挫败：学生明明改了立场，教师却仍以原提纲要求其回归；或者学生只是润色，却被评价为“没有思想变化”。

第五，范文教学从“仿表面”改成“做消融”。给学生一篇范文，先不讲好在哪里，而是让不同小组分别做保相和破相操作：把一个例子换掉、把一句关键因果反过来、把转折层移动、补写一个原来的空白。再比较哪些改动仍让范文保持同一关系相。学生学到的是结构弹性和脆弱点，而不是把好作文的可见特征复制到自己的文字里。

第六，评价可以在正式分数之外增加“身份诊断”。CPR 不适合直接变成高考新分项，因为一旦成为高风险公开指标，学生和模型会围绕它优化；更合适的用途是研究性诊断和形成性评价。教师用少量变换检查文章是“定相”“惰性”“脆弱”还是“无相”，然后选择不同反馈：惰性模板需要提高关键关系的特异性，脆弱堆叠需要增加外围表达的自由度，无相文本则先建立最低组织。

十六、AI 协作写作：判断“谁写了多少”之外，还要判断对象有没有改相

生成式人工智能使作文身份问题从理论争论变成日常操作。最常见的制度判断是统计 AI 参与比例：改了几个句子、生成了多少字、是否由模型起草。但字数比例与篇章身份没有稳定对应。AI 可以重写 80% 的句面而保持同一关系相，也可以只建议替换一个关键理由，就把论证相位彻底改掉。

篇章定相体因此提供一条不同于“AI 检测”的审计线：保存人类初稿与 AI 协作版本，先声明当前作业要保持的关系层，再做版本差分。如果 AI 修改主要落在 T+ 方向，核心关系探针保持，而学生能够解释这些关系为何必须保留，那么工具更像编辑器；

如果 AI 重写使一个或多个 T-关系发生跃迁，学生接受后形成新相，则应把这一轮标记为“对象改相”，继续追问新核心关系由谁提出、学生是否理解，而不是只报一个机器生成百分比。

这并不替代学校的学术诚信规则。学校可以出于能力测量目的规定不得使用 AI，即使某次 AI 修改完全保相，也可能违反任务规则。本文只处理对象论：AI 参与之后这还是不是同一篇作文。任务合规、作者责任和对象身份是三张不同账本，混在一起会让教育政策既难执行又难解释。

更深一层的价值在于，定相框架允许人机协作产生可追踪的“思想版本”。如果学生最初主张 A，AI 提出反例后学生改为 B，那么与其假装这是对原文的“润色”，不如明确记录一次相变：旧对象结束，新对象开始。这样，AI 不再只是代写风险，也成为观察学生何时真正改变判断的实验条件。

十七、文学开放性与文体差异：定相不是把作文变成只有一个正确理解

“关系探针”很容易被误解成要求所有读者得到同一主题，这会伤害文学性。本文的判别对象并不是统一解释，而是预注册的最低关系层。诗性叙事可以允许情绪、象征和结局解释高度分叉，却仍可能在“谁在向谁说话”“某个时间跳跃是否发生”“哪一处关系被故意留空”等最低结构上保持稳定。定相与解释开放并不冲突：外围解释可以很宽，身份层只要求被声明为核心的少数关系具有可预测响应。

不同文体也需要不同探针。记叙文的核心变量可能是事件因果与视角，议论文是主张—证据—反驳关系，说明文是分类边界与机制顺序，应用文则可能是行动义务、对象与条件。若把同一探针强加给全部文体，就违反了本研究从“超越单一序参量”得到的二阶结论。

这也是“分辨率声明”最重要的伦理意义。一个教师可以明确说：“这次只评价论证对象有没有定相，不评价文学声音是否保持。”另一个创意写作课堂可以反过来把声音作为身份层。标准不是消失，而是从伪装成作文永恒本质的隐性权力，变成可以被学生看见、讨论和质疑的显性实验条件。

十七之二、四类教学误判：用“篇章定相体”重新解释常见作文现场

第一类误判是“改得越多，越不像原文”。教师批改时常以红笔覆盖面积感知修改强度，学生也容易把“我几乎重写了一遍”理解为“另写了一篇”。但若大规模重写集中在 T+层——删冗余、改句法、调整节奏、替换功能等价细节——对象的核心关系可能高度稳定。反之，一个看似只改标题或结尾的动作，如果改变了前文被如何回写，也可能产生真正的相位跃迁。修改量应与身份改变脱钩。

第二类误判是“中心相同，所以还是同一篇”。作文训练中常按题目和中心思想管理版本，只要中心仍可概括为“坚持带来成长”，教师就把不同稿件归为一篇。然而“坚持”可能在一稿中是个人意志压过困难，在二稿中变成借助他人、调整目标后的持续行动；表面中心词没变，行动者、因果和价值关系已重排。定相框架要求比较关系响应，而不是只比较一句中心概括。

第三类误判是“结构完整，所以对象稳定”。五段式、总分总、开头点题结尾扣题可以产生很强的形式稳定感，却未必有特异性。如果把三个分论点随意换序、把任一例子替换成社会热点、把结尾换成另一条励志句，读者仍给出同一主题判断，那么这种结构更像可互换容器。它的稳定来自模板，不来自这篇文字自身的关系约束。所谓“惰性模板”不是骂模板，而是指出其破相敏感度过低。

第四类误判是“读者理解一致，所以文本对象已经成立”。一致理解可能来自题目、常识和类型脚本。比如题目已经叫《那一次，我懂得了坚持》，即使正文关键关系很弱，读者仍会把大量模糊细节解释成“坚持”。因此定相实验要故意改变承重关系，看读者是否沿预期方向改变；如果关键关系怎么改都不影响回答，一致性就不能归功于这篇作文。这里与“跨端复成体”关注的文本归因问题有亲缘性，但定相体把焦点放在版本变换而非作者源对象。

四类误判共同指向一个课堂转向：教师不必立刻采用正式 CPR，只需要把评价问题从“这篇有什么”改成“如果我这样改，它会怎样”。反事实问法天然迫使教师说明自己真正看重的关系，也让学生有机会反驳：“你说这一段是关键，但删掉以后我们五个人的理解都没变。”专业判断因此不是被算法取代，而是获得一种可以公开检验的形式。

十七之三、从范文到“变体簇”：作文课程资源如何重新组织

传统范文库把一篇优秀文本作为静态标本保存：学生读原文、圈好词好句、分析结构，然后尝试模仿。篇章定相体建议把“单篇范文”升级为“变体簇”。每个范例至少附三类版本：一份保相改写、一份隐蔽破相改写、一份明显破相改写。学生的任务不是猜哪篇更漂亮，而是判断哪些版本仍属于同一对象，并指出判决所依赖的关系。

这种资源形式有一个重要优势：它把“结构”从教师讲解对象变成学生可以操作的对象。比如同一议论文的四个版本拥有完全相同的开头和结尾，只在证据依赖上不同；学生只有真正理解论证关系，才能发现其中一个版本已经把证据用反。又如同同一叙事的四个版本语言同样流畅，却有一版移动了一个时间层，使后面的“我终于明白”不再成立。结构理解由识别名词转为预测变换后果。

变体簇还适合 AI 时代。教师可以让 AI 生成十个表面改写，再由学生筛选哪些是保相、哪些在无意中改相，并要求给出关系证据。这样，AI 的低成本生成能力反而成为制造实验材料的工具。真正需要学生负责的不是每一个字都亲手写，而是能否识别对象边界、解释关系变化并决定接受哪一种新相。

如果这一课程模式长期运行，作文教材的组织单位也可能改变：不再只有“范文+点评”，而是“基准文本+变换协议+版本簇+关系问题+失败例”。一门语文课程由此拥有了自己的实验性，不是模仿自然科学做实验，而是让语言对象在受控改变中暴露结构。

十八、证伪条款：哪些结果出现时，本文必须承认自己错了

第一条证伪：若在预注册实验中，CPR 与人工“同一作文/不同作文”判断没有任何稳定关系，并且这种无关在不同文体、不同评分者和不同变换设计中重复出现，则“变换响应能够刻画作文身份”的核心主张失败。不能用“教师还没学会新理论”解释掉所有反例。

第二条证伪：若一个只使用静态文本特征的简单模型，在严格留出新题目、新作者和新改写方式的条件下，持续达到与 CPR 同等或更好的版本身份判别，而加入变换响应不产生增益，则本文声称“静态单点不足”的赌注失败。

第三条证伪：若所谓保相与破相变换无法获得跨评分者的前置一致分类，即专家在看结果前就不能稳定判断某个操作理论上应保持还是改变核心关系，那么 CPR 的方向编码缺乏可裁决性。此时应退回变换本体，而不是继续计算一个漂亮分数。

第四条证伪：若不同判别分辨率只是语言包装，改变分辨率并不系统改变哪些变换被判为 T+或 T-，那么“分辨率依赖身份”这一二阶结论没有实际内容。相反，如果同一改写在“论证关系”层保相、在“作者声音”层破相，并且评分者能在预注册条件下重复这一差异，才支持本文的分层对象论。

第五条证伪针对近邻：若写作研究已有理论已经明确提出与本文同构的对象——以预注册保相/破相变换、关系探针和版本边界联合裁定一篇具体作文的同一性——则“篇章定相体”没有独立命名资格，应并入既有概念。新名不是权利，而是待撤销的临时假设。

十九、反向约束与日期赌注：新理论必须允许自己被限制

反向约束一：CPR 高不等于作文质量高。一篇价值立场陈旧、语言平庸甚至事实错误的文章，只要其核心关系在保相/破相操作中表现稳定且特异，也可能具有高 CPR。本文只裁定对象身份的形成，不把“有相”偷换成“好”。质量问题需要另一个理论处理。

反向约束二：CPR 低也不等于“不是作文”。某些实验性、碎片化或刻意拒绝稳定关系的文学写作可以主动追求低定相；幼儿初学写作也可能尚未形成稳定组织，但教育制度仍可把它归为作文作业。本文中的“篇章定相体”是对象论上的强形态，不是行政分类的唯一门槛。

反向约束三：分辨率越细，CPR 不必越高。把“声音、节奏、意象次序”纳入身份层后，原先被视为保相的改写可能成为破相；这不是测量失败，而是对象定义变细。任何研究若不报告分辨率，就不得跨研究直接比较 CPR。

本文给出一个可裁决的研究赌注：在 2027 年 12 月 31 日前，若至少三个独立研究团队以公开预注册方式对中学或大学作文实施不少于 6 个成对变换，并报告本文同口径 CPR，则高 CPR 文本在盲评“版本同一性”判断上的一致率应显著高于低 CPR 文本；若三个团队均未发现方向一致的差异，核心模型应被降级，不得继续以“篇章定相体”作为普遍对象论。这里押的是方向，不押某个漂亮效应量。

二十、对语文教育学的理论回写：从“文本属性论”到“对象响应论”

如果本文成立，作文教育最深的一次变化不是增加“相变”“地层”“范畴”三个术语，而是改变研究单位。过去研究往往把作文当作已给定对象，然后测其语言、结构、内容、过程、作者与读者；本文要求先问对象是如何被持续识别的。于是，修改不再只是提高分数的手段，而成为揭示对象身份的实验；版本差异不再只是过程噪声，而成为理论数据。

这一转向也会改变作文评价中的“结构完整”。完整不等于段落齐全或首尾呼应，而是存在一组可被反事实操作识别的组织依赖。一个有漂亮开头、三个整齐分论点和升华结尾的文本，如果删除任何所谓关键部分都不改变读者对核心关系的回答，它可能只是形式齐全却没有关系特异性。相反，一篇短小、留白很多的文字若对某一边界移动极其敏感，也可能拥有强对象结构。

它还重新解释“个性”。传统作文教学容易把个性落在独特材料、独特语言或真情实感上；定相框架把个性的一部分理解为“特异性响应”：换掉哪一处，这篇就不再是它；什么东西只有这个对象需要守住。个性不是越奇特越好，而是对象具有自己的脆弱点和稳定区。

最后，它为跨文体、跨媒介研究提供接口。一篇图文作文、口述转写、多人协作文或 AI 辅助文，只要能够声明边界、定义关系探针并实施变换，都可以讨论定相；反之，即使完全由学生手写，如果任何核心变换都不能改变其关系响应，也可能只是一个未形成特异组织的语言集合。对象论由“是谁敲字”转向“什么关系被形成并守住”。

二十一、局限：为什么本文不能把物理、地质和数学“证明”成语文真理

第一，跨学科碰撞不是逻辑演绎。相变、地层单位和 Yoneda 各自在自己的对象域中有严格定义，把它们迁入作文只能产生新的可检验假设，不能把原学科权威当作语文结论的证明。尤其是“篇章相”“关系指纹”等词若脱离操作规则，就只是借科学感装饰教育话语。

第二，关系探针的设计具有理论负载。我们无法在完全中立的情况下决定哪五个关系最核心，因此预注册只能防止事后篡改，不能消除价值判断。未来研究需要比较不同教师、文体专家和学生共同制定探针的结果，考察谁有权规定分辨率。

第三，破相变换可能同时降低一般可读性。若把一个关键段落删掉，读者回答变化可能因为文字变得语法残缺，而非触到对象身份。因此变换应尽量保持局部语法与可读性，并设置功能等量的对照改写；否则 CPR 会把“坏编辑”误当成“相变”。

第四，本文尚未取得直接 CPR 语料验证。公开数据字段审计的结果是“不能算”，而不是“已验证”。这条边界必须写在结论前：当前贡献主要是对象概念、差分观测协议与证伪路径，经验效力仍待专门实验。

二十二、结论：作文不是被写出来的一堆文字，而是一种在变换中可被识别的关系相

从凝聚态物理出发，我们学到整体状态不能由字词数量线性决定；从地层学出发，我们学到对象身份离不开边界、序位、缺失和相关；从 Yoneda 出发，我们学到内部属性清单不足以替代关系响应。第一阶碰撞让三家互相取消“我这一维单独够用”，第二阶又推翻三家共有的“存在一个脱离判别尺度的永久身份证”，最终得到一个分辨率依赖、可反事实检验的新对象。

因此，本文对“作文是什么”的回答不是“作文是一种相变”，更不是“作文像岩层或范畴对象”。答案是：作文可以被建模为篇章定相体——在声明的判别分辨率下，一个有界语言—版本序列形成宏观组织关系，并在预注册的保相变换中保持、在破相变换中跃迁；对象身份由这一差分响应剖面被裁定。

这个答案把一个极普通的课堂问题变得可研究：老师说“这还是原来那篇作文，只是改得更好”，凭什么？学生说“我只改了几句话，但已经不是原来的意思”，凭什么？AI 把全文润色以后，究竟是同一作文的表面变体，还是新对象？过去这些判断主要靠经验直觉，篇章定相体试图给它们一套可以预注册、可以盲评、可以失败的实验语言。

如果未来实验证明 CPR 不能稳定预测版本同一性，本文应被放弃或降级；如果它能工作，作文教育得到的也不是一份新的“好作文公式”，而是一种更基本的能力：

不只看一篇文字现在长什么样，而会通过有控制的改变，发现它究竟靠什么成为它自己。

附录 A 3.9 版工序审计账本

A1 闸零与四道闸审计

- 闸零·未吸收度：相变/序参量→干净；地层单位/时间序列→干净；Yoneda 式关系对象→干净。检索重点不是学科词是否出现，而是“以该判决规则裁定具体作文版本身份”是否已成固定写作理论；未检出系统吸收。
- 闸零之二·理论厚度：三家均判“厚”。凝聚态物理具奠基层、内部争论、反例传统与教科书体系；地层学具国际规范、不同单位体系争论、断层/不整合等反例传统；范畴论具 1940 年代以来奠基、Yoneda 及大量扩展、标准教材。
- 闸一·矛盾烈度：8/10。S 允许微观差异被宏观相吸收；D 坚持材料相似不能越过边界/序位直接合并；E 要求外部变换关系剖面进入身份判决。
- 闸二·去同存：低同源。三家分别来自统计物理、地质时间分类、抽象数学；不共享同一套“信息/网络/认知”母体。
- 闸三·位置三分：S=显露成何种宏观相；D=怎样经有界时序形成单位；E=在何种关系映射场中被识别。三维各一。
- 闸四·避重：与同产线“跨端复成体”分离。前者研究版本身份的变换等价类，后者研究作者源对象经文本到读者端的复成。

A2 敌意拓宽检索词（节选）

- 方向①经典写作过程：composition process theory; Flower Hayes writing process; cognitive process writing identity。
- 方向②修辞与体裁：rhetorical situation composition; genre social action writing; dialogism composition。
- 方向③篇章结构与语言学：cohesion coherence writing; rhetorical structure theory essay; discourse relation essay identity。
- 方向④社会/网络与后过程：postprocess composition network; writing as network; distributed writing agency。
- 方向⑤方法学/计算：essay version identity; text transformation invariance essay; automated essay scoring robustness; invariance under rewriting。
- 同产线：跨端复成体；作文是什么；二阶碰撞 What1。检索任务是寻找 1:1 占位者，而非寻找支持性文献。

附录 B 篇章定相率 (CPR) 清点手册与课堂示例

B1 CPR 清点手册 (试行版)

- 1. 先写“本轮判别分辨率”，如“核心论证依赖”或“叙事因果与视角”，不得写“整体好不好”。
- 2. 冻结 5 个关系探针。每个探针必须能用有限类别编码，避免事后自由解释。
- 3. 设计至少 3 个 T+ 与 3 个 T-。T+ 改变表面、理论上保持关系；T- 保持局部可读性但触动一个预注册核心关系。
- 4. 变换方案先于读者结果登记；生成版本后随机打乱，盲评者不知道原文和变换类别。
- 5. 每个版本至少 5 名独立读者回答同一组探针。对每个探针按完全一致/部分一致/不一致编码为 1/0.5/0，取平均得到 RelSim。
- 6. 对 T+ 令 $R = \text{RelSim}$ ；对 T- 令 $R = 1 - \text{RelSim}$ ；全部 R 求均值值得 CPR。
- 7. 同一研究全文使用同一阈值 0.70。低于阈值只表示未达到本轮试行定相条件，不直接判“差作文”。
- 8. 任何无法保持语法可读性的破相操作标记为无效，不得计入 CPR。

B2 一个课堂微型示例

- 原文关系层：学生原先把母亲不断问成绩理解为控制；一次生病后发现母亲并非只关心排名，而是不会表达担忧；结尾没有把控制全部洗白，而是把“关心”与“边界”并置。
- T+1：把“凌晨两点她仍坐在门外”改写为同功能的间接叙述；若读者仍识别“担忧通过笨拙方式显露”，应保相。
- T+2：把两个外围学习细节调换；若核心判断不变，应保相。
- T-1：把“她不会表达担忧”改成“她害怕我丢她的脸”；若读者仍给出原核心判断，说明文本的所谓关键理由并未真正承载相。
- T-2：删除“我仍然要求她不要每天查手机”这一边界句；若读者由“理解但保留边界”滑到“理解等于原谅全部控制”，则该句具有高定相贡献。
- 教师在这个练习里不需要计算正式 CPR，也可以让学生第一次看到：作文的关键并非每个细节都重要，而是某些关系一旦变动，整个对象会换相。

参考文献

- [1] Anderson, P. W. (1972). More Is Different. *Science*, 177(4047), 393–396.
- [2] Kadanoff, L. P. (2009). More is the Same; Phase Transitions and Mean Field Theories. *Journal of Statistical Physics*, 137, 777–797.

- [3] Wilson, K. G. (1971). Renormalization Group and Critical Phenomena. I & II. *Physical Review B*, 4, 3174–3205.
- [4] Wen, X.-G. (1990). Topological Orders in Rigid States. *International Journal of Modern Physics B*, 4, 239–271.
- [5] Chen, X. (2025). Generalized Landau Paradigm for Quantum Phases and Phase Transitions. *Physical Review Letters*, 135, 250001.
- [6] Chaikin, P. M., & Lubensky, T. C. (1995). *Principles of Condensed Matter Physics*. Cambridge University Press.
- [7] Hedberg, H. D. (Ed.). (1976). *International Stratigraphic Guide*. Wiley.
- [8] Salvador, A. (Ed.). (1994). *International Stratigraphic Guide: A Guide to Stratigraphic Classification, Terminology, and Procedure* (2nd ed.). Geological Society of America.
- [9] International Commission on Stratigraphy. *International Stratigraphic Guide: Stratotypes and Type Localities; Lithostratigraphic Units; Chronostratigraphic Units*. Online guide, accessed 2026-08-17.
- [10] Eilenberg, S., & Mac Lane, S. (1945). General Theory of Natural Equivalences. *Transactions of the American Mathematical Society*, 58, 231–294.
- [11] Yoneda, N. (1954). On the Homology Theory of Modules. *Journal of the Faculty of Science, University of Tokyo, Section I*, 7, 193–227.
- [12] Mac Lane, S. (1998). *Categories for the Working Mathematician* (2nd ed.). Springer.
- [13] Leinster, T. (2014). *Basic Category Theory*. Cambridge University Press.
- [14] The Stacks Project Authors. (2026). Opposite Categories and the Yoneda Lemma, Tag 001P. Accessed 2026-08-17.
- [15] Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365–387.
- [16] Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Lawrence Erlbaum Associates.
- [17] Hayes, J. R. (1996). A New Framework for Understanding Cognition and Affect in Writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing*. Lawrence Erlbaum Associates.
- [18] Bitzer, L. F. (1968). The Rhetorical Situation. *Philosophy & Rhetoric*, 1(1), 1–14.
- [19] Miller, C. R. (1984). Genre as Social Action. *Quarterly Journal of Speech*, 70(2), 151–167.
- [20] Kent, T. (Ed.). (1999). *Post-Process Theory: Beyond the Writing-Process Paradigm*. Southern Illinois University Press.
- [21] Dobrin, S. I., Rice, J. A., & Vastola, M. (Eds.). (2011). *Beyond Postprocess*. Utah State University Press.
- [22] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- [23] van Dijk, T. A. (1980). *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition*. Lawrence Erlbaum Associates.
- [24] Mann, W. C., & Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*, 8(3), 243–281.
- [25] Bakhtin, M. M. (1986). *Speech Genres and Other Late Essays*. University of Texas Press.

- [26] Bazerman, C. (2004). Speech Acts, Genres, and Activity Systems: How Texts Organize Activity and People. In C. Bazerman & P. Prior (Eds.), What Writing Does and How It Does It. Lawrence Erlbaum Associates.
- [27] Conference on College Composition and Communication. (2018). Writing Assessment: A Position Statement. NCTE/CCCC.
- [28] Mathias, S., & Bhattacharyya, P. (2018). ASAP++: Enriching the ASAP Automated Essay Grading Dataset with Essay Attribute Scores. Proceedings of LREC 2018.
- [29] Crossley, S. A., Tian, Y., et al. (2024). The English Language Learner Insight, Proficiency and Skills Evaluation (ELLIPSE) Corpus. International Journal of Learner Corpus Research, 9(2).
- [30] 付自文 (2026) .《作文何以成为作文：从“文本成品”到“跨端复成体”——古生物学与演化形态×信号处理×空间统计的二阶碰撞》. 学科通融专栏·二阶碰撞典范文（同产线近邻，用于划界） .

—— 学科通融专栏·二阶碰撞典范文（What1·v3.9） ——

第二章 补节 不动它，就不知道它是不是它

（一）最近的祖宗与分离线

本章的核心装置是「预注册的保相变换与破相变换」：先声明哪些改动应当保持对象身份、哪些应当使它跃迁，再看实际响应是否符合声明。原稿划了九位占位者，全部在写作理论与文本语言学内部，而这套装置最近的祖宗在别处。

里贝罗等人的 **CheckList (Ribeiro, Wu, Guestrin & Singh, "Beyond Accuracy: Behavioral Testing of NLP Models with CheckList", ACL 2020**，最佳论文)。该文提出三类测试：最小功能测试、不变性测试 (**INV**) ——施加保标签的扰动，预期判断不变；方向期望测试 (**DIR**) ——施加特定扰动，预期判断朝规定方向改变。这与本章的保相／破相变换在形式上是同一件事，连「事先声明预期方向、再看是否兑现」这个纪律都相同。其思想源头是软件工程的蜕变测试 (metamorphic testing) ——当无法直接判定输出是否正确时，改用「输入按规定方式变化时输出应当如何变化」这一关系来检验。同族的还有加德纳等人的对照集 (Gardner et al., "Evaluating Models' Local Decision Boundaries via Contrast Sets", Findings of EMNLP 2020) 。

这是本卷全部十三篇原稿里，方法与祖宗形式重合度最高的一处。

分离线在被测对象。CheckList 与蜕变测试测的是一个系统：给定输入变换，检验模型或程序的行为是否合乎规定；变换的作用是暴露系统缺陷，被测对象的身份从头到尾没有疑问——那个模型就是那个模型。本章测的是一个对象是否还是它自己：变换的作用不是暴露缺陷，而是构成身份判据本身；本章第十节写明，离开声明的判别分辨率，「唯一身份证」不存在。

判决性对照：CheckList 的一次 INV 失败，结论是「模型有 bug」，模型本身不因此变成另一个模型；本章的一次保相变换失败，结论是「这已经是另一篇作文」，没有任何东西可以被称为 bug。前者用不变性检验质量，后者用不变性定义同一性。这条差别足以让两者并存，但它必须写出来——否则本章的方法论部分读起来像是把一份 2020 年的工程清单换了个领域。

编者另请读者注意，第一章补节所列的丹托「不可辨别的对应物」同样适用于本章，且方向相反：丹托证明同一个可见形态可以是不同对象，本章处理的是不同的可见形态何时仍是同一对象。两者合起来才是完整的一对。

（二）与第一章的对切

见第一章补节（二），此处不重复。两章各自的证伪清单应各加一条：若人侧与文本侧的读数长期一一对应，本章与第一章应当合并。

（三）本章欠账

本章的真跑是字段审计，结论是不利的——CPR 在现有公开评分语料上算不出来。原稿把这个结果原样保留，是本卷倡导的做法。但读者应当清楚它的代价：本章至今没有一个 **CPR** 的实际数值。一套完整的判据、四种诊断状态与一份变换设计，全部停在协议层。第十四节给出的最小实验（一百篇作文、每篇六个版本、五个关系问题）成本并不高，在它跑出来之前，本章是九章里理论最完整而经验最空的一章。

第三章

读得出的那一部分，恰恰不是整体性

从内生判差到异源回判

作文整体性的三学科碰撞、 公开语料反证与概念重建

超分子化学 × 第三体摩擦学 × 射电干涉闭合相位

研究问题	何谓作文？作文是什么？
证据路径	三源机制提取 → 候选概念 → 预注册扰动 → 公开语料反证 → 概念重建
本轮裁决	淘汰“内生判差”的当前操作化；提出“异源回判”作为待检验假说

研究论文 · 证据版本 V1.0

完成日期：2026 年 8 月 17 日

目录

论文结构与主要裁决路径

摘要	3
引言：把“作文是什么”从常识中重新取出	3
一、作文研究的四条既有路线及其定义缺口	4
二、研究程序：SDE V3.9 的 What 型三维度路线	6
三、三种源机制的厚描及其边界	7
四、第一次概念碰撞：“内生判差”与“偏差分流”	9
五、公开语料先导研究：从概念到可失败测量	10
六、结果：高覆盖率为何构成警报	12
七、对抗性边界：一个定义如何同时过宽与过窄	13
八、概念重建：从“内部有判断”到“远距异源回判”	15
九、对“作文是什么”的阶段性回答	18
十、教学、评价与写作技术的含义	19
十一、限制与确认性研究方案	21
十二、SDE V3.9 门控与研究裁决	22
结论	23
附录 A：样本锁定与可复现信息	24
附录 B：V2 最小编码记录模板	24
参考文献	25

提示：目录条目可点击跳转；页码依据最终排版生成。

摘要

“作文是什么”常被“怎样写好作文”替代。前者追问对象的构成方式与边界，后者则预设对象已经存在，只讨论训练、评分和改进。本文把问题重新置于本体论与测量论的交界处：作文何以不是句子的集合、一般连贯文本或任务合规产品？为寻找能够经受反例的回答，研究依照 SDE V3.9 的 What 型三维度程序，引入三种彼此遥远而又可操作的源机制：超分子化学的多点识别、协同性与可逆组装，第三体摩擦学的负载传递、速度调节与物质流，射电干涉测量的闭合相位及其对站点型误差的不变性。三者第一次碰撞形成候选概念“内生判差”及“偏差分流”机制：一处局部改动进入全文关系网后，可能被排斥、被结构承载、被高层关系消隐，或仍然不可判定。本文没有把这一新概念直接宣布为定义，而是把它改写成可失败的操作命题，并在 Argument Annotated Essays v2 (AAE2) 公开语料上实施确定性抽样的审计性先导研究。十二篇作文各接受六类最小扰动，共七十二项判断。结果表面上极强：全文判差覆盖率为 70/72 (0.9722)；但 63/72 (0.8750) 的全文判断仅凭改动句与一条邻句即可还原，越过预注册的 0.80 淘汰线。删除全部中等把握编码后，近邻可还原率反而升至 59/63 (0.9365)。更严重的是，扰动类型与命运类别的描述性 Cramér's V 为 0.7007；不读作文、只看操作标签，逐篇留一预测即可命中 62/72 (0.8611)。这说明现有量表主要测到了局部语义保持、邻接矛盾和显式让步，并且把实验者预先设计的操作性质误认成文本自身的组织能力。依照预注册，本文否决“作文是内生判差体”这一肯定性定义，暂停“偏差分流”的独立机制地位，但不据此否定三学科组合的理论潜力。反证推动概念由“内部能判断”转向“远距、异源、可回判”：作文整体性不在全文总命中率，而在移除邻句、中心句复述与同源线索以后，远处且关系类型不同的文本路径能否提供超过局部基线的判断增益。本文据此提出“异源回判”假说与远距增益指标 Δ_{remote} ，并把作文暂定为：将经验、立场与语言选择组织成远距、异源、可回判关系的一种书面行动。该表述不是终极定义，而是一项经过首轮失败后变得更窄、更难伪装、也更可检验的研究纲领。

关键词：作文；整体性；跨学科概念研究；内生判差；异源回判；公开语料；预注册；反证

引言：把“作文是什么”从常识中重新取出

作文研究有一个不易察觉的起点问题。日常教育语言把“作文”当作一个无需定义的对象：有题目，有篇幅，有开头、主体和结尾，有评分者，便似乎已有作文。研究于是自然转向写作过程、策略训练、篇章连贯、体裁知识、评价量规与自动评分。这些工作都重要，却没有自动回答一个更基础的问题：当一组句子遵守语法、围绕同一主题、甚至取得高分时，使它成为一篇作文的究竟是什么？反过来，一篇极短、跳跃、故意留白但仍被读者确认为作文的文本，又凭什么没有被排除？

如果把作文定义为“学生按题目写成的文章”，定义就依赖学校制度，无法解释制度外的习作、随笔和公共写作；如果定义为“表达思想感情的文章”，说明书、社论、书信和诉状都会涌入外延；如果定义为“有中心、有材料、有结构的书面语篇”，数学证明、合同与技术规范同样具有内部结构。许多定义事实上是教学要求、原型描述或质量标准，而不是能够区分对象、解释边界并承担反例的理论定义。

本文从三个判断出发。第一，作文不能仅由局部性质定义。句子的正确、段落的顺滑和连接词的齐全可以逐项达成，却仍可能形成同题拼盘。第二，作文也不能仅由外部制度定义。题目、评分表和作者身份能规定任务，却不能完全替代成文之后可被读者审计的关系。第三，任何声称抓住“整体性”的理论，都必须区分作文级关系与一般局部连贯，并与证明、合同、说明书等高度约束文本进行对照，否则极易把“复杂文本共有的组织性”误称为“作文的本质”。

研究问题因此被写成四层：其一，三个碰撞学科分别能提供什么可迁移机制，而不只是修辞形象？其二，这些机制能否共同生成一个关于作文整体性的可失败命题？其三，公开真实文本能否迫使这一命题承担否决风险？其四，如果首轮操作化失败，失败暴露的是理论本体、测量方式，还是概念外延中的哪一处混淆？

本文的贡献不在提出一个听起来新颖的名词，而在展示概念如何被不利数据改写。研究先由三源碰撞得到“内生判差”，再用预注册的局部窗口检验把它否决为作文定义，最后提出“异源回判”作为下一轮假说。这个顺序有意反转概念写作中常见的确认路径：新概念不是终点，而是必须接受破坏性测试的中间模型。由此，“作文是什么”不再只是思辨命题，也成为可以设计扰动、设置对照、报告失败并逐轮收窄的经验问题。

一、作文研究四条既有路线及其定义缺口

（一）过程路线：作文是认知活动，但活动边界不等于对象边界

Flower 与 Hayes（1981）的认知过程理论把写作描述为计划、转译和复核之间递归而非线性的活动，改变了“先想好再写下”的简单阶段观。Bereiter 与 Scardamalia（1987）进一步区分知识陈述与知识转化，说明成熟作者会让内容问题与修辞问题相互作用。Hayes（2012）的修订模型又把动机、工作记忆、转录技术以及任务环境纳入系统。过程研究因此回答了写作者在做什么、资源如何分配、困难在哪里发生。

然而，从“写作是一组递归过程”不能直接推出“什么成品算作文”。购物清单、实验记录、判决书和程序注释也可能经过计划、转译与复核。过程模型擅了解释生成，却未必提供对象的排他边界。更关键的是，成文一旦进入阅读，作者当时的心理步骤常不可见；若对象身份完全依赖不可重建的过程，读者就无法仅凭文本与情境

判断它是什么。因此，本文吸收过程路线对行动性和递归性的强调，但不把某种心理流程当作作文的充分条件。

（二）文本路线：衔接与连贯说明“如何相连”，却未穷尽“为何成篇”

Halliday 与 Hasan（1976）系统描述指称、替代、省略、连接和词汇衔接，使跨句联系成为可分析对象。de Beaugrande 与 Dressler（1981）提出衔接、连贯、意向性、可接受性、信息性、情境性和互文性七项文本性标准，把文本从句法扩展为交际事件。Witte 与 Faigley（1981）则提醒，衔接数量与写作质量不是简单同一；显性的连接手段不能保证概念结构充分。Mann 与 Thompson（1988）的修辞结构理论进一步以核心—卫星和功能关系解释语篇组织，提供了目的、证据、让步、条件等关系的分析语言。

这些理论为本文的“关系路径”提供基础，却也暴露两个限制。其一，局部衔接能够在低质量文本中大量出现：代词有先行项，句间有 because 或 however，段落仍可能只是并置。其二，连贯具有读者建构成分。一个熟悉主题的读者可能为稀疏文本补足关系，另一个读者则不能。若把“读起来连贯”直接视为作文的本体属性，就会把读者能力、世界知识与文本结构混为一体。本文因而要求每次判定都能指出文本锚点，并区分局部窗口已经足够的关系与必须调用远程结构的关系。

（三）互惠与体裁路线：社会行动解释用途，但类别成员仍有开放边缘

Nystrand（1986）把书面交流理解为写者与读者之间的互惠过程；文本并非封闭容器，而是参与者协调预期的媒介。Miller（1984）提出“体裁即社会行动”，使体裁不再只是形式清单，而成为对反复出现情境的类型化回应。Beaufort（2007）把写作专长置于话语共同体、主题、体裁、修辞和过程知识的交汇处。这一路线有力地说明，作文不是去情境的语言物，题目、课堂、公共议题、读者角色与评价制度都会进入其意义。

但社会行动定义也可能过宽。“完成一个书面任务”包括填表、报税、签约和发布操作指令。若仅凭行动目的区分，许多混合体裁难以归类；若仅凭教育制度区分，则“作文”变成管理标签。体裁理论本身强调类别的历史性、原型性和可变性，因此它更适合提醒我们保留边界弹性，而不是为作文提供一条机械的充分必要条件。本文将在结论中把作文理解为书面行动，但为这种行动增加开放的不完备性、异源关系组织和回判能力三项结构条件。

（四）论证挖掘路线：可计算关系提供数据，却可能把标注模式当成文本全貌

论证挖掘把主张、前提、支持和攻击关系转化为可标注结构。Stab 与 Gurevych（2017）的 AAE2 包含 402 篇说服力学生作文、7,116 个句子和 147,271 个词元，为真实作文的关系研究提供了重要公共基准。该语料含 6,089 个论证成分，其中支持关系 3,613 条、攻击关系 219 条；23% 的段落存在未链接论证成分，20.9% 的论证呈串联结构。这些统计说明学生作文不是整齐的模板树，也说明反方关系远少于支持关系。

然而，标注方案服务于特定任务。AAE2 把作文建模为由总主张统摄、段内论证构成的连接树，并不系统标注跨段前提关系、叙事时间、语气回响、词汇复现、例证—概括等所有可能的整体联系。数据中没有某类边，不能被解释为文本中没有这种关系。相反，如果研究者只使用既有标注，他会把标注本体的边界误当成作文的本体边界。本文因此以 AAE2 原文为主要判断材料，以 BRAT 标注定位论证单元，却不把论证树视为完整的作文模型。

（五）共同缺口：需要区分局部连贯、一般文本性与作文整体性

四条路线分别抓住生成、文本、社会和计算层面，却留下一个共同难题：什么证据表明某一局部选择受到了“整篇”的约束，而不只是受到邻句、题目模板或评分常规的约束？这一问题不是要求每一篇作文都必须复杂，也不是把整体性等同于高质量；它要求研究者说明，所谓整体何时产生了不可由局部替代的可观察后果。

本文把“局部扰动的命运”作为切入口。若改动一个词、否定一条前提、移入同题句或删除让步单元，文本其余部分会发生什么？有些改动被多个关系排斥，有些看似冲突的单元实际上承载结构功能，有些形式变化不改变高层承诺，还有些改动仅凭文本无法判断。这个问题能够把抽象的整体性变成反事实比较；但它也有危险：邻句足以完成的判断会伪装成全文能力。先导研究的核心正是让这一危险显形。

二、研究程序：SDE V3.9 的 What 型三维度路线

（一）程序定位与三项纪律

本文将 SDE V3.9 作为概念发现与淘汰程序，而不是已经获得普遍承认的学科理论。What 型路线不从现成定义出发，而从三个异质知识域提取机制，经碰撞形成候选概念，再以现实数据和边界反例执行门控。组合—规定：S（粒子维度）取超分子化学，D（波动/过程维度）取第三体摩擦学，E（场/关系维度）取射电干涉的闭合相位。

程序遵守三项纪律。第一是“源机制吸收”：每个学科必须贡献可陈述的因果或测量结构，不能只提供词汇装饰。第二是“不可直译”：分子、磨粒、天线与句子不存在自然同一性，任何迁移都只能保留关系形式，不能把物理定律冒充语言定律。第三是“否决优先”：碰撞所得概念必须在成文前写出何种结果会使其降级或淘汰，并在失败时保留不利证据。

(二) 机制提取而非意象拼贴

表 1 给出三种源机制的最小提取。其目的不是证明作文“像”分子、摩擦副或干涉阵列，而是寻找三个互补问题：局部单元怎样被多点关系确定？进入界面的偏差可能有哪些去向？怎样构造不依赖单一外部校准源的关系量？

表 1 三学科源机制、可迁移关系与禁止直译

学科源	原学科中的核心机制	可迁移的关系形式	明确禁止的直译
超分子化学	分子识别、多价作用、协同性、可逆自组装	单点弱约束可由多点配合获得选择性；整体结构反过来限定局部位置	不把词句称作真实分子，不用结合能替代语义证据
第三体摩擦学	第三体转载、速度调节、界面分离、源流—循环—泄漏	偏差不只有删除一种命运；中介单元可承压、缓冲、转移或逸出	不把删改等同于磨损率，不宣称文本服从接触力学方程
闭合相位	闭合回路消去站点型相位误差；闭合量有信息损失与协方差	判断应尽量由多边关系闭合，不依赖单一评者或中心句；关系量也有盲区	不把语义关系当作傅里叶相位，不宣称三句相加即可得到真值

(三) 碰撞有效性的判准

三源之间设计了二十七个概念碰撞单元，每个单元检验一种跨源组合是否产生了新增问题、可观测后果和反例条件。二十一个被判为有效，两个处于边界，四个无效。有效碰撞通常同时满足三点：源机制都不可删除；产物不能由某一学科单独改名得到；产物能够改变后续数据设计。无效碰撞主要有两类：一类只制造华丽同义词，例如把“连接”改称“结合”；另一类错误地把闭合相位的数学消元直接投射为意义正确。边界碰撞则具有解释价值，但尚不能产生独立测量。

这一门控的意义在于，跨学科新颖性不能由距离本身保证。学科越远，滥用隐喻的风险越高。真正有效的碰撞必须带来“若如此，应该看见什么；若看不见，放弃什么”的研究后果。组合一最终留下的不是三个名词，而是一条可受攻击的链：多点关

系赋予局部选择以命运；中介结构使命运不止对错两类；关系回路使判断减少对单一外部锚点的依赖。

（四）研究证据的分层

本文区分四种证据。源文献证据说明原学科机制；结构推演证据说明迁移后的逻辑可能性；公开语料证据检验当前操作量能否识别作文级结构；对抗性反例检验概念外延。四者不得互相替代。尤其是，三学科机制在原领域成立，并不证明迁移概念成立；先导语料运行得到高比例，也不证明定义成立；反例能够淘汰过宽定义，却不能独自给出新定义。

三、三种源机制的厚描及其边界

（一）超分子化学：整体不是强键，而是受组织的多点关系

Lehn（1988）把超分子化学概括为超越单个分子的化学，核心包括分子识别、互补性与由非共价作用形成的高阶组织。此后他进一步把自组织理解为信息化学系统通过组分之间的识别与选择生成有序结构的过程（Lehn, 2002）。这里对作文研究最有价值的，不是“句子如分子”的类比，而是对整体性的重新刻画：整体不一定由一条不可破坏的强键维持，也可能由许多方向性、可逆、强弱不一的接触共同稳定。

多价作用使这一点更清楚。Mammen、Choi 与 Whitesides（1998）讨论生物系统中的 polyvalent interactions：多个配体—受体接触同时或连续参与，可以产生高于单个接触的总亲和与选择效应。多价并不意味着把每个弱作用机械相加；几何匹配、局部浓度、协同性、熵代价和重绑定都会改变整体。对文本而言，可迁移的不是热力学数值，而是一个判准：一个局部单元若同时进入指称、因果、论证、语气或例证等不同关系，其命运可能由这些关系的组合决定，而非由最近的一条连接决定。

这一机制还带来可逆性。超分子组装常由非共价作用维持，结构能够在条件变化时重排。作文修改也不是简单地把“错误句”从成品中剔除；改写一个词可能保持高层承诺，移动一个例子可能重建论证，删除一句让步却可能使远处回应失去对象。可逆性提示我们研究“扰动后的结构再平衡”，而不是只统计语病。

但迁移边界必须严格。第一，文本关系没有可直接测量的结合自由能；所谓“稳定”只能由可标记证据、读者一致性或任务表现间接估计。第二，多点联系可能完全同源：总论点在首尾重复三次，看似有三个锚点，实则只有一个信息源。第三，关系越多不等于作文越好；冗余、套话和模板也能制造高连接度。因此，超分子机制只支持“多点与协同值得检验”，不能支持“多点即作文”。

（二）第三体摩擦学：偏差进入界面以后并非只有清除

传统摩擦叙述容易只看两个接触固体，但真实接触常产生磨屑、转移膜、氧化物或环境颗粒。Godet（1984）的第三体进路把这些介于两个第一体之间的物质视为接触机制的组成部分，而非纯粹废物。Iordanoff 等（2002）的综述概括了第三体的三项功能：传递负载、调节两第一体之间的速度差、把第一体分隔开以减少直接损伤；并讨论五个速度调节位置、四种调节模式以及源流、内流、外循环和泄漏等物质流。相关框架后来在速度调节位置、固体第三体流变和离散元磨损模型中继续展开（Berthier et al., 1992; Descartes & Berthier, 2002; Fillot et al., 2007a, 2007b）。第三体可能保护表面，也可能参与磨损；关键不在它“好”或“坏”，而在它怎样被接触系统生产、输送和排出。

这为作文修改提供一个重要纠偏。学校评价常把偏差想象成必须清除的异物：矛盾删掉，离题删掉，重复删掉。然而，成熟文本中的反方声音、犹疑、插叙、语体突变和局部摩擦可能承担负载。它们让结论不是无阻力的口号，使立场通过回应而形成，使叙事转折获得张力。删除表面不协调的单元，全文反而可能变弱。第三体机制因此启发“承载”命运：局部偏差不是被容忍的噪声，而可能成为结构传力的中介。

同时，第三体的流动视角阻止我们把文本看成静态网络。一个例子可以从经验材料转化为论证前提，一句反方意见可以在后文被重述、限定并回应，一处关键词可以循环出现而改变语义负载。单元的身份取决于它在生成、再进入与逸出过程中的位置。由此，研究对象不只是“这句话与哪句话相连”，还包括“这句话的功能如何沿全文迁移”。

迁移仍有三重限制。其一，物质流有守恒和力学条件，语义功能没有同等的物理量；“流”只能被操作为指称延续、立场转换或关系重编码。其二，第三体属于接触系统的一般机制，不具有作文特异性。合同中的定义条款、证明中的引理同样可以承担中介负载。其三，R、B、C 三种命运容易被实验操作预先写入：同义替换天然趋向消隐，命题否定天然趋向排斥，删除让步天然趋向承载。先导研究后来证实了这一设计混淆。

（三）闭合相位：关系不变量、信息损失与相关误差

在射电干涉测量中，一条基线的观测相位不仅包含天体源结构，也包含两端台站的相位误差。用简化式表示，基线 ij 的测得相位可写为：

$$\varphi_{ij}^m = \varphi_{ij}^s + \theta_i - \theta_j + \varepsilon_{ij0}$$

当三条有向基线构成闭合三角形时，三条相位相加：

$$\psi_{123} = \varphi_{12}^m + \varphi_{23}^m + \varphi_{31}^m,$$

各台站误差 θ_1 、 θ_2 、 θ_3 以相反符号成对消去，留下与源结构有关的闭合相位以及不能闭合的噪声项。Jennison（1958）奠定了这种相位敏感测量方法。Chael 等（2018）指出，闭合相位与闭合振幅对站点型校准误差免疫，因而可直接用于成像；但闭合量也丢失信息，例如仅用闭合相位不能确定图像质心，仅用闭合振幅不能确定总流量，重建仍需正则化或附加约束。Blackburn 等（2020）又强调，闭合量虽对天线增益稳健，却往往在线性和统计上相关；忽视协方差会造成有偏推断，低信噪比和非闭合误差也不能由“闭合”二字自动消除。

对作文研究而言，真正可迁移的是三条方法论。第一，判断应寻找在某类观察者或局部锚点变化下仍保持的关系量，而不是依赖单一评分者、题目说明或中心句。第二，闭合不等于完整；一个对某类误差不变的指标可能同时丢失对象信息。第三，多条回路不等于多条独立证据；它们可能共享边、具有协方差，或只是同一中心命题的重复组合。

这直接修正了最初的“内部”想法。遮去题目和评分表后，文本内部仍能给出判断，似乎具有闭合性；但若所有判断都源于一句总论点，或者全文协调地重复同一错误，所谓闭合只是相关误差的共同传播。作文中的“独立路径”因此不能按空间数量计算，必须要求关系类型和信息来源的异质性。也正是在这里，“异源”成为后续概念重建的必要条件。

（四）三源的共同约束：整体性必须产生增量，而非只留下网络图

三种机制汇合后，整体性可以被初步写成：局部单元进入多点、流动且可回路检验的关系系统，系统使其在扰动后呈现可区分的命运。但这一表述仍过宽。分子网络、接触界面和干涉阵列都提醒我们，结构存在必须产生可测后果。对作文而言，这个后果不应是“全文里能找到两句相关话”，因为大多数文本都能做到；它应是远程关系相对于局部关系带来的新增判断能力。

因此，本研究没有把网络密度、连接词数量或全文一致率直接视为整体性。最初仍把“全文能判差”当作候选量，同时预先设置近邻窗口 N 作为破坏性对照。事实证明，真正推动理论进步的不是高全文命中率，而是 N 把高命中率拆穿了。

■ 四、第一次概念碰撞：“内生判差”与“偏差分流”

（一）候选定义

第一次碰撞形成如下核心命题：遮去题目、评分表和作者说明后，文本其余部分仍以至少两条非同源关系路径，对某一局部改动给出可审计的命运判定。这里的“内生”不是拒绝世界知识，而是要求判定理由在当前文本中留下至少两个相互分离的锚点；“判差”不是发现所有事实错误，而是判断改动进入关系系统以后如何改变结构。

为了避免把所有偏差归入“错误”，研究定义四种命运。排斥（R）指改动破坏至少两处相互分离的文本关系，恢复原单元后两处破坏同时解除。承载（B）指表面不协调的单元承担反例、转折、声音差异、情节摩擦或概念桥接，删除后至少两处关系变弱。消隐（C）指局部形式改变，但至少两条高层推论或叙事承诺保持不变。未定（U）指命运只能诉诸题目、评分表、作者说明、一般语法或常识，或者文本内部只能找到一条证据路径。

这四类构成“偏差分流”。它吸收超分子化学的多点约束、第三体的不同流向与承载功能，以及闭合相位的内部关系检验。与简单纠错不同，它允许文本保持矛盾、吸收变化或拒绝变化；与纯解释不同，它要求扰动前后可比较；与评分不同，它不先规定哪种命运更优秀。

（二）独立路径规则

两条证据路径只有同时满足四项条件才计为独立。其一，跨度不同，不能只是同一句或邻句的复述；其二，至少分属因果、时间、指称、论证、声音、例证—概括、问题—回应等不同关系类型；其三，进行留一锚点检验，遮去最近的共同中心句后，至少一条路径仍保留原判断；其四，两条路径若都由题目、总论点或同一模板槽位推出，视为同源。

这个规则试图回应闭合量的相关性问题的，却仍有漏洞。“至少两个锚点”可以由分析者在全文中事后搜寻；“关系类型不同”也可能只是标注语言不同，底层信息仍相同。更重要的是，只要一个局部矛盾同时触发邻句中的指称和因果关系，就能满足两类关系，却没有证明整篇参与了判断。近邻对照因此不是附属指标，而是区分局部与整体的关键否决器。

（三）为什么必须先写下失败条件

研究在读取本轮样本正文前写死五类裁决。最关键的一条是：若近邻可还原率 $N \geq 0.80$ ，则淘汰当前操作化，因为指标未超出一般局部连贯。另有保留条件要求原文平均 J 比顺序对照与同题拼盘合并均值至少高 0.15，且至少八篇方向一致；若差值仅为 0.05—0.15 或只有六至七篇方向一致，则定义降级为质量属性；若 R、B、C 中任意两类不能稳定区分，则淘汰三分机制；若多路关系主要由同一中心句派生，则必须加入正交来源条件。

把失败线写在结果之前，有两项作用。一是阻止研究者把任何高比例解释为支持。二是把“整体性”的理论承诺具体化：如果一条邻句已经足够，全文关系就没有产生增量。这正是预注册把预测与事后解释分开的功能（Nosek et al., 2018）。后来的结果恰好触发这一不利条件，说明预注册并非装饰。

（四）与邻近概念的区别

“内生判差”不等于衔接。衔接可以由局部形式标记完成，判差要求扰动后的关系命运。“内生判差”也不等于逻辑一致。一篇作文可以承载反方声音和未解决张力；内部一致的文本也可能全篇重复同一事实错误。它还不等于自动评分，因为命运类别不是分数，承载不低于排斥，消隐也不天然优于变化。

然而，候选概念与一般复杂文本之间的边界仍不清楚。定理证明、合同和技术规范都具有多路内部约束，甚至比多数学生作文更能判定局部改动。若“内部能判差”是充分条件，这些文本都会被误收。极短而优秀的寓言或随笔又可能因缺乏两条显式路径被误拒。这种外延同时过宽与过窄的征兆，要求现实语料之外再加入对抗性类别测试。

■ 五、公开语料先导研究：从概念到可失败测量

（一）数据、样本与证据等级

主语料为 Argument Annotated Essays v2。原论文报告 402 篇说服力学生作文；本研究通过公开加载脚本固定的数据提交取得 402 组正文与 BRAT 标注。压缩包及辅助文件均保存 SHA-256 校验值，以防后续数据漂移。研究者在预注册前知道语料规模和论文中的克隆主题示例，因此明确排除该示例；实际检索确认它是作者另写的说明文本，不在 402 篇语料中。

样本以确定性规则抽取：计算 SHA256(essay_id + “SDE-V3.9-PANCHA-20260817”), 按哈希升序选取满足正文、标注、语言和至少八句条件的十二篇。入样编号依次为 essay238、essay326、essay178、essay039、essay056、essay161、essay057、essay162、essay082、essay049、essay045 和 essay115。每篇实施六类扰动，总计七十二项。

本轮只有一名知晓假说与阈值的分析者，不具备盲法，也没有独立编码者。因此它是审计性小试运行，不是确认性实验。研究的可信度来自规则公开、样本锁定、逐项编码和不利结果保留，而不是来自总体推断。任何比例都只描述这七十二项，不外推到所有作文。

（二）六类最小扰动

六类扰动分别为：S1 表层同义替换；S2 保持基本命题的句法改写；P1 对主要前提作最小否定；P2 把关键实体、数量或属性替换为同主题但不相容的值；R1 从另一篇近题作文移入一条语法通顺的前提；R2 删除反方、例外、转折、回应或桥接单

元，若原文无此类单元则删除主要例证。所有扰动保持局部语法基本通顺，避免把语法错误伪装成关系错误。

目标选择采用固定规则而非逐篇挑选。S1 选择最早含可替换实词的主张，S2 选择最早且不少于八词的前提句，P1 选择最早前提，P2 选择第二早前提，R1 用全库题目词级一至二元 TF-IDF 相似度选择来源作文并取其最早合格前提，R2 按最早反方主张、显式让步/转折、最长前提的顺序选择。这些细则中有一部分在已读四篇正文、尚未进行命运编码时补充，构成程序性迟注册，论文据此降低证据等级。

（三）全文条件与近邻条件

每个扰动在两个阅读条件下编码。全文条件允许读取整篇正文；近邻条件只读改动句及其后一条句子，无后句时读前一句。R2 则读取被删原句及其后一条句子。两个条件使用同一 R/B/C/U 规则独立判断，再比较是否一致。

全文判差覆盖率定义为：

$$J(T) = \text{获得 R、B 或 C 且存在两条合格证据路径的有效扰动数} \div \text{全部有效扰动数}。$$

近邻可还原率定义为：

$$N(T) = \text{近邻条件与全文条件命运相同的扰动数} \div \text{全部有效扰动数}。$$

J 高只说明分析者能在全文中为大多数扰动找到命运；若 N 同样高，则这种能力主要来自局部窗口，不能归功于作文整体。研究还记录单源依赖率、承载检出率和关系型穿透率，但本轮裁决以 N 的预注册阈值为优先。

（四）停止规则与运行例外

预注册规定，一旦原文 $N \geq 0.80$ ，便满足淘汰当前操作化的充分条件；顺序对照和问题拼盘不再可能把裁决改回“保留”。实际运行触发该条件，因此研究停止保留性检验，没有生成两个对照。这个决定使本轮无法估计原文与对照的 J 差值，也不能对“降级”条件作完整判断；它能支持的只有更窄而明确的结论：当前量表未能把作文级整体性与局部连贯区分开。

另有三项例外被保留。essay049 与 essay045 的最早反方主张和总主张同句，若整句删除会同时移除被检验的总立场，故跳到下一条独立让步；essay178 的 not only...but also 被粗略的 but 检索误识别为反方标记，实际依照后备规则删除最长例证。这些调整都可能提高 R2 的承载可见度，因而不应被用于强化有利解释。

六、结果：高覆盖率为何构成警报

（一）主要计数

七十二项扰动全部可运行。全文命运分布为 $R=25$ 、 $B=13$ 、 $C=32$ 、 $U=2$ ， $J=70/72=0.9722$ 。若只看这一数字，几乎可以宣称作文内部广泛存在判差能力。然而，近邻与全文同判达到 $63/72$ ，即 $N=0.8750$ ，越过 0.80 淘汰线 0.075 。全文相对于近邻新增的可判定项仅 $9/72=0.1250$ ；其余全文—近邻不一致为零。

表 2 公开语料先导研究的主要计数与预注册解释

指标	结果	预注册解释
有效扰动	$72/72$	数据可运行
全文判差覆盖率 J	$70/72 = 0.9722$	无对照且可能通胀，不能单独支持保留
近邻可还原率 N	$63/72 = 0.8750$	超过 0.80 ，淘汰当前操作化
全文独有可判定增量	$9/72 = 0.1250$	真正调用远程关系的案例很少
高把握子集的 N	$59/63 = 0.9365$	删除全部中等把握项后反证更强

九项编码被标为中等把握。若担心不确定判断抬高 N ，把它们全部删除，剩余六十三项中的五十九项近邻—全文同判，比例反而升至 0.9365 。由此，不利结果不是由少数模糊案例制造。

两个 U 都来自 essay039 的 P 类扰动：否定“太空照片显示地球为圆”，以及把“地球绕太阳”改为“太阳绕地球”。外部常识能立即判断后者错误，但正文没有第二条独立锚点反驳它。这个案例显示，内部可判定不等于事实正确；若量表借用常识把它强行编码为 R ，就会把外部校准源偷偷带回“内生”指标。

（二）操作类型预编码了命运

更严重的结果来自扰动类型与命运的耦合。 $S1$ 、 $S2$ 全部为 C ； $P1$ 、 $P2$ 各有一项为 R 、一项为 U ； $R2$ 有十项为 B 、两项为 C ；只有 $R1$ 呈混合分布。六乘四列联表的描述性 Cramér's V 为 0.7007 ，显示强关联。使用逐篇留一方案，仅凭 $S1/S2/P1/P2/R1/R2$ 标签、完全不读作文内容，就能预测 $62/72$ (0.8611) 的全文命运；总体众数 C 的基线为 $32/72$ (0.4444)。

表 3 六类扰动与全文命运的描述性分布

扰动族	R	B	C	U	众数命运
S1 同义替换	0	0	12	0	C (12/12)
S2 句法改写	0	0	12	0	C (12/12)
P1 命题否定	11	0	0	1	R (11/12)
P2 不相容替换	11	0	0	1	R (11/12)
R1 近题移植	3	3	6	0	C (6/12)
R2 关系单元删除	0	10	2	0	B (10/12)

这不是显著性推断，而是设计诊断。S 操作按定义保存命题，几乎必然产生消隐；P 操作按定义制造否定或不相容，容易在邻接位置形成排斥；R2 专门选择反方、让步或例证，删除后容易被判为承载。换言之，实验者先把 C、R、B 的典型性质写进操作，再从结果中“发现”三分机制。量表回答的主要是“这种编辑通常会造成什么”，而非“这一篇作文怎样决定它的命运”。

（三）三类通胀

第一是消隐通胀。同义替换与保义改写在操作层面已规定高层命题不变；只要编码手册把高层承诺保持视为 C，十二加十二项命中几乎是定义重述。第二是排斥通胀。最早前提往往与下一句例证或推论直接相连，命题否定和不相容替换自然制造邻接冲突，N 因而升高。第三是承载通胀。R2 优先删除显式让步与反方，邻句中的 however、this 或回应句会立即暴露功能，不需要全文。

三类通胀共同解释了一个看似矛盾的现象：J 越高，越可能不是作文整体越强，而是操作和编码规则越同构。这里的教训具有一般性。概念研究若只报告高命中，不检查命中能否由标签、模板或局部窗口预测，就会把测量装置的结构当成对象的结构。

（四）九项远程增量透露了什么

虽然当前量表被淘汰，九项全文新增可判定案例仍具有探索价值。例如，某篇作文移入“英语促进国家间交流”的同题句，近邻窗口只显示话题顺接，全文却显示它只是已有交流论证的冗余；另有移入句在全文的“户外活动—社交—健康”两条支线之

间形成桥接，近邻无法识别。删除某些句子后，只有远处的指称回接、结尾重述或反方回应同时受损，才显露承载功能。

这些案例不能用来挽救原定义，因为它们是结果后观察，数量很少，也由单一分析者判断。但它们指出下一轮测量应主动遮蔽邻句与中心句，预先定位跨段关系，并比较远程异源证据相对于局部基线的增益。失败并未让三学科组合失去内容，反而把真正需要测量的部分从高J的表面中分离出来。

■ 七、对抗性边界：一个定义如何同时过宽与过窄

语料先导只检验了说服力学生作文，无法回答候选概念是否具有作文特异性。为此，本文构造五类边界反例。它们不是概率样本，也不用于估计比例；其作用类似单元测试：只要定义在清楚案例上产生系统性误收或误拒，就必须解释或修改。

（一）定理证明、合同与技术规范：内部约束不是作文专有

一个形式良好的定理证明具有高度内生的判差能力。改变定义、否定引理或删除推导步骤，多个后续结论都会受损；恢复原步骤，关系重新闭合。合同中的定义、义务、例外与救济条款也会相互制约；技术规范的参数、接口与验收标准更可能形成远距一致性检查。按“至少两条内部路径决定局部改动命运”的规则，这些文本比许多学生作文更符合候选定义。

这说明“语言对象 + 内部多路约束”不足以区分作文。问题不在这些文本没有整体性，而在整体性是上位属性。作文定义必须加入其特有或至少更典型的行动结构：面对并非由规则完全封闭的问题，写作者把经验、立场和表达选择组织成可供他人回判的关系。证明和合同当然也需要判断，但其局部命运更大程度由公理、定义或制度规则预先规定；作文则通常保留材料选择、视角、价值权衡和表述方式的开放空间。这里是程度差异而非绝对二分，因而最终定义应以原型性条件而非僵硬的逻辑围墙表达。

（二）协调一致的全局错误：闭合不能清除相关误差

设一篇文章从开头到结尾都把某一错误事实当作共同前提，并在多个段落协调地展开。内部关系可以非常一致：例证支持错误前提，结论回扣错误前提，关键词稳定复现。任何单点检查都可能被全文“消隐”。这与干涉测量中的非闭合误差和相关误差问题相似：闭合相位消去站点型误差，并不消去所有误差；共享来源的偏差可在多条关系中同步传播。

该反例迫使“非同源”升级为“异源”。空间分离不够，关系标签不同也不够。因果链与例证链若都依赖同一未经检验的事实，它们仍非正交。未来编码至少要区分经验来源、价值前提、论证功能和语言实现四个层面；只有来自不同层面的关系在遮

蔽共同中心后仍能回判，才提供较强证据。即便如此，异源回判也不保证事实真，只保证某一局部选择受到多种内部承诺的约束。

（三）局部连贯的问题拼盘：主题一致可以冒充整体

从多篇同题作文中各取一段，保留正确连接词，再让每段围绕同一总主题，就能形成局部流畅的拼盘。多数同义替换、句法改写和邻接否定仍可被准确判断；甚至每一段内部都存在主张—理由—例证结构。然而，段落之间可能没有材料递进、立场修正、概念迁移或声音回应。原量表的高 N 正说明这种文本可能被误判为具有整体性。

因此，下一轮必须让非作文对照在长度、主题、语法、段内连贯和操作分布上与作文匹配，只破坏跨段关系。若指标不能把原文与这种困难对照区分，就只能被解释为一般段落连贯指标。简单打乱句序的低质量对照不足以承担这一任务，因为任何表面指标都能轻易获胜。

（四）极短而优秀的随笔或寓言：显式多路不是必要条件

反方向的错误来自极短文本。几句话的随笔、寓言或片段可以通过省略、反讽和文化互文形成强烈整体效果，却未必在字面上留下两条可标记的独立路径。如果定义坚持“至少两个文本内部锚点”，这些作品会被排除。读者之所以承认它们，可能正因为文本把一部分关系委托给体裁惯例、共同知识和沉默位置。

这表明内生性不能等于自给自足。任何作文都处在语言、文化和情境中；完全排除外部知识既不可能，也会偏向长篇显性论证。更合理的要求是：外部知识可以参与理解，但关于某一局部选择为何在这篇作品中成立的理由，必须受到文本中可指认的回判信号约束。短文本的信号可以是标题—正文反差、首尾语义反转、叙述声音与事实的不一致，不一定需要两条完整句链。未来研究应按篇幅和体裁分层，避免把长篇论证作文的显性结构当作作文的普遍本质。

（五）含反方—回应—回扣的长文：组合一真正有解释力的区域

候选概念最有解释力的案例，是一篇长文先引入反方，随后用经验材料限定它，在结尾重新解释开头问题。若删除反方，后文回应失去对象；若改变经验细节，结尾价值判断不再有同样根据；若只读邻句，这些损失不可见。在这里，多点识别、第三体承载和闭合回路确实共同解释了远程组织。

但这一能力更像作文质量的一个维度，而不是作文身份的必要条件。初学者的作文、故意碎片化的作文或简短习作可能缺乏这种结构，却仍是作文。组合一因而应从“定义所有作文”降到“解释作文整体性如何增强”：它研究的不是—篇文章何时跨过作文身份门槛，而是写作行动何时让局部选择受到远距异源关系的共同回判。

（六）边界测试的合并裁决

五类反例与语料结果指向同一裁决：内生判差既有外延过宽，也有外延过窄；其量表又受局部窗口和操作标签支配。因此它不能作为作文的充分必要定义。最合理的保留方式，是把它拆成一个可分级的结构能力，并把“内生”改写为带有外部情境但不被单一外部锚点决定的“异源回判”。

■ 八、概念重建：从“内部有判断”到“远距异源回判”

（一）为什么“内生”需要被撤回

“内生”容易引发两种误解。一种是封闭文本观，仿佛作文可以脱离题目、读者、文化和事实独自确定意义；另一种是内部证据数量观，仿佛在全文找到两处支持便已完成证明。先导研究显示，近邻窗口足以还原大多数命运；边界测试又显示，全篇一致可能只是共同误差。继续保留“内生”会遮蔽问题的真正难点：证据是否远距，来源是否异质，判定是否相对于局部基线产生增量。

“回判”比“判差”更合适，因为它强调关系的往返。局部选择向外投射承诺，远处单元又返回来限制对该选择的理解；结尾可以重估开头，例证可以限定总论点，反方声音可以使回应获得必要性。判断不是文本中央的一次裁决，而是局部与远程位置之间的往复约束。“异源”则要求这些返回路径不能都是中心句的复制品。

（二）异源回判的工作定义

本文提出以下 V2 假说：作文的结构能力，不在于局部改动能否被全文判断，而在于移除目标邻句、重复总论点和同源复述后，远处且来源异质的关系仍能共同决定该改动是被排斥、被承载、被消隐，还是保持未定。

这一定义有五个组成条件。第一，存在局部选择 x ，可以是词语、命题、例证、声音或结构单元。第二，至少有一个远程位置 r ，与 x 跨段或跨功能区域。第三， r 与局部证据 l 不共享唯一信息源，关系可以来自经验、立场、叙事、指称或语言形式中的不同层面。第四，在遮去 l 和共同中心句后， r 仍改变对 x 的命运判断。第五，这种改变能够由盲评者根据预先手册复现，而不是由研究者事后阐释。

“共同决定”不要求每条路径都单独充分。正如多价关系的总体作用不能简化为单个接触，两个异源路径可以只有在结合时才使判断稳定。不过，这种协同必须通过遮蔽或消融实验显示，不能凭网络图想象。

（三）远距增益指标

下一轮的主量不应再是全文总覆盖率 J ，而应比较远程异源条件与局部条件：

$$\Delta_remote(T) = J_remote,orthogonal(T) - J_local(T)。$$

其中， J_local 只允许目标句及预设邻域； $J_remote,orthogonal$ 遮去该邻域、重复总论点、题目提示和同源复述，只允许预注册的跨段异源锚点。 $\Delta_remote > 0$ 表示远程关系提供了局部窗口没有的判断信息。若原文的 Δ_remote 高于主题、长度、语法和段内连贯均匹配的同题拼盘，且在盲评者之间稳定，才可把它解释为作文整体性的候选证据。

该差值仍不应被神化。 J_remote 与 J_local 可能受任务难度、文本长度、编码者知识和锚点数量影响；两个条件也可能不是统计独立。参考闭合统计对协方差的警告，确认性分析应在篇内配对、报告分层分布，避免把大量相关扰动当作独立观测。最合适的单位可能是“篇”而非“扰动”，并以多层模型或随机化检验处理篇内聚类。

（四）异源性的四层操作化

为避免“关系类型不同”沦为换标签，本文提出四层来源编码。经验源回答材料从何而来，包括事件、观察、事实或想象场景；立场源回答什么价值或判断承诺在起作用；组织源回答单元承担论证、叙事、解释、反方或回应中的何种功能；语言源回答指称、词汇复现、语气和句法如何实现联系。两条路径只有至少跨越其中两个层面，并在遮蔽共同中心后仍保留部分判力，才暂记为异源。

例如，“共享单车减少短途汽车出行”与后文“我在雨天仍选择打车”并不只是一正一反。前者是一般因果主张，后者是经验反例；若结尾据此把立场修正为“基础设施与天气条件决定替代程度”，则经验源、立场源和组织源发生回判。相反，开头说“共享单车环保”，结尾说“总之共享单车对环境有益”，只是同源复述，即使相隔数段也不构成异源。

（五）R、B、C、U 的重新定位

先导结果表明，四类命运仍有分析价值，但不能由固定操作族直接触发。V2 设计应让每种操作都有可能进入 R、B、C 或 U。例如，同一类“关系中性替换”在不同文本中可能破坏远程隐喻而被 R，可能保留全部承诺而被 C，也可能因证据不足为 U；同一类删除可能移除冗余而被 C，也可能移除桥接而被 B。只有当操作类别与命运的可预测性显著下降，命运差异才有机会归因于文本结构。

此外，B 不应只由删除让步测量。可以对一处疑似摩擦单元进行三种等长度替代：删去、改成同向陈述、改成另一异源反例。若后两种对远程关系造成不同影响，才能识别该单元到底承载了“反方位置”、具体经验内容，还是仅占据一个模板槽。第三体机制由此从形象的“容纳异物”转为可操纵的中介功能检验。

（六）理论身份：不是新本质，而是可检验的研究纲领

经过反证，“异源回判”不再被宣称为作文的本质。它首先是一项中层假说：在某些作文中，局部选择会受到跨段、跨来源关系的返回性约束，这种约束可用遮蔽后的远距增益测量，并可能与读者感知的整体性、修改质量和迁移能力相关。若后续研究发现 Δ_remote 不能区分作文与匹配的非作文文本，它仍可降级为一般复杂文本能力；若它与质量评价无关，则可能只是一种结构类型；若盲评无法复现，概念应被进一步修改或放弃。

（七）竞争解释与判别预测

异源回判要获得理论地位，不能只比被淘汰的“内生判差”更精细，还必须战胜能够解释同一现象的较简单模型。第一种竞争解释是纯局部连贯模型：读者对改动命运的判断，主要来自目标句附近的语义相容、指称连续和话语标记；所谓远程证据只是分析者在全文中的事后搜寻。该模型预测，只要局部窗口包含相同词汇、实体与连接标记，原文和跨段关系被破坏的拼盘在判断上不会稳定分开。异源回判模型则预测，在控制局部窗口后，遮去远程异源锚点会使部分判断系统性退化为 U，且这种退化在原文显著多于拼盘。两者的分歧不在全文是否“看起来更顺”，而在远程消融是否产生篇级增量。

第二种竞争解释是任务合规模型。考试或课堂作文共享题目、立场模板和评分习惯，远处段落之所以相互可预测，可能只是都在响应同一任务，而非彼此形成回判。该模型预测：保留题目与通用总论点、但替换具体经验材料以后，判断能力基本不变；不同作文中同一模板槽的句子可以相互替换。异源回判模型则预测，经验、立场和组织功能之间一旦形成具体耦合，同题移植不能完全保持原命运。确认性研究可设置“同题同立场—异材料”与“同材料—异立场”两个正交对照，检验真正发挥作用的是题目一致，还是文本内部的异源组合。

第三种竞争解释是一般复杂性模型。较长文本拥有更多句子、更多潜在锚点和更多被分析者选中的机会，因此 Δ_remote 可能只是长度、词汇多样性或关系密度的函数。若如此，篇幅匹配的合同、证明和技术说明应取得同等甚至更高的远距增益，作文身份没有解释增量。异源回判理论对此作出较窄预测：复杂非作文文本当然可以有强回判，但其路径更多由定义、公理或制度规则预先封闭；作文中较典型的模式应跨越经验选择、立场修正与语言实现。分析时必须先用长度、段数和锚点候选量建立基线，再检验来源层数与开放问题情境是否增加解释力。若没有增加，理论应降级为一般文本复杂性指标。

第四种竞争解释是读者补足模型。远距关系或许主要存在于熟练读者的推理，而不是文本信号中。同一篇短文，熟悉体裁者能恢复反讽与省略，不熟悉者则不能。若回判完全依赖读者背景，文本锚点的一致性会很低，遮蔽实验的结果也会随专业知识

而剧烈变化。异源回判并不否认读者参与，却要求补足受到成文信号约束。未来可让不同背景读者先独立指出锚点，再交换理由；若他们虽给出不同解释，却反复指向相同的标题反差、语气转换或结尾回扣，说明文本提供了约束。若锚点无稳定聚集，只能依靠外部知识讲出合理故事，则该案例不应计为强回判。

四种竞争解释还要求改变模型比较方式。不能只问异源回判指标是否与总体评分相关，因为长度、词汇与任务合规同样可能相关。更严格的顺序是：先拟合局部窗口信息，再加入题目与模板信息，再加入长度和一般复杂性，最后才加入预注册的远程异源特征；只有最后一步带来跨文本验证中的新增预测，才支持理论增量。对单篇案例也应采用同样逻辑：先给出最省力的局部解释，再证明它为何不足，不能一开始便把所有远处联系收入概念。

这一判别框架使“异源回判”承担双重风险。它既可能被局部模型击败，也可能被更一般的复杂文本模型吸收。前者意味着远程增量不存在，后者意味着增量存在但没有作文特异性。只有当原文在困难拼盘上显示远程优势、这种优势不能由题目与长度解释、路径又跨越不同来源层并能被盲评者复现时，组合一才获得下一阶段的经验资格。

■ 九、对“作文是什么”的阶段性的回答

（一）从文本物转向书面行动

先导失败提示，作文不能仅被定义为具有某种静态网络的文本物。写作者面对一个不完全封闭的问题，从可能材料中选择经验，形成或试探立场，并以语言形式安排这些选择；读者随后沿文本关系回判这些选择。作文因此更接近一种书面行动，其产物是文本，但其身份同时包含问题情境、选择过程和可读关系。

“行动”并不意味着必须知道作者的心理史。成文中会留下行动痕迹：为何选择这个例子而不是另一个，为何在此处转折，为何结尾重新命名开头的问题。读者不必还原真实意图，只需判断这些选择是否在文本与情境之间形成可审计关系。行动性由此兼顾过程研究与文本研究：它承认生成，却把可检验对象放在成文关系上。

（二）三个原型条件

本文提出作文的三个原型条件，而非充分必要条件。

第一是开放的不完备性。作文所面对的问题不能被既定规则完全算出。即使题目要求明确，材料取舍、视角、立场、语气和组织仍有多种可辩护方案。开放性区别于纯粹填空，却不排斥论证规则和体裁约束。

第二是远距异质组织。作文不只是连续造句，而是让经验、立场与语言选择在不同位置、不同关系类型中相互改变意义。这里的远距是功能距离，不必机械等于若干段；短文中的标题—结尾反转也可以具有高功能距离。异质则要求不把重复当作多路。

第三是可回判性。文本为读者提供足够信号，使局部选择能够被后续或先前关系重新判断：它为何必要、为何可替换、为何构成张力、为何保持未定。可回判不是唯一正确解释，而是解释需要受文本证据约束。

（三）阶段性定义句

在当前证据下，较稳妥的回答是：

作文不是能在局部发现矛盾的文本；作文更可能是一种把经验、立场与语言选择组织成远距、异源、可回判关系的书面行动。

这句话有意保留“更可能”。它既不是词典释义，也不是完成了普遍验证的本体论。它是组合一经首轮现实数据否决后留下的最小承诺。与“内生判差体”相比，它撤回了封闭性，加入关系来源异质性，把局部总命中改为远距增量，并把对象从静态成品恢复为书面行动。

（四）身份与质量的分离

一篇关系单薄、主要靠邻接推进的学生文字仍可能是作文，只是整体性较弱。反之，合同或证明具有很强远距回判，也不因而成为作文，因为其问题开放度、行动目的与体裁制度不同。异源回判最适合描述作文质量与成熟度中的结构维度，而作文身份仍由开放书面行动的原型情境共同决定。

这种分离防止理论变成排斥初学者的门槛。教育研究尤其不能把高水平结构能力倒过来定义对象，否则“写得不好”会被偷换成“不是作文”。概念的任务不是取消低质量成员，而是解释同一对象类别内部为何出现不同程度的整体组织。

十、教学、评价与写作技术的含义

（一）教学目标：从“句句正确”转向“远处能回来”

组合一对教学最直接的启示，是不要把修改压缩为局部纠错。教师可以要求学生为一个关键句画出三类回判：它由前文什么经验支持，被后文什么立场限定，在语言上与哪一处词语或声音发生回响。若所有箭头都指向相邻句或中心句复述，学生看到的就不是“没有结构”，而是远程增量尚未形成。

一种可行练习是“遮蔽阅读”。先遮去题目和总论点，让同伴只凭正文判断一处例子的功能；再遮去例子邻句，看远处是否仍提供线索；最后恢复全部文本，比较判

断如何变化。练习的重点不是追求唯一答案，而是让学生区分局部可懂、主题相关与全文必要三种不同性质。

另一种练习是“承载测试”。选取一处反方、犹疑、转折或叙事摩擦，分别删去、改成同向句、换成另一个反例，观察结尾和中段哪些位置随之失去功能。这样，所谓“详略得当”“照应”不再只是教师评语，而成为可被学生操作和观察的关系变化。

（二）评价量规：把整体性拆成增量证据

传统量规常列“中心明确、结构完整、衔接自然”。这些维度容易诱发模板写作：首尾重复中心，段首加连接词，便获得表面完整。异源回判取向的量规可以增加三个观察项：是否有跨功能区域的关系；这些关系是否来自不同信息层；遮去邻句与重复中心后，是否仍能说明局部选择的作用。

评价时不宜直接计算箭头数量。高密度关系图可能奖励冗余。更有意义的是提供一到两个“差异性证据”：某句若改动，远处哪些不同性质的承诺会改变；某段若删除，后文回应、结尾判断或叙述声音是否随之失去对象。量规由此从静态特征清单转向反事实解释。

同时，U 应被保留为合法结果。文本内部不能判断某个事实，不代表作文失败；它可能提醒写作者补充证据，也可能表明该事实需要引用外部来源。强迫所有扰动落入 R/B/C 会鼓励评分者用常识替代文本证据。对事实性写作，应把外部可核验性与内部回判性分成两个维度。

（三）反馈语言：指出关系去向而非只贴标签

“不连贯”“离题”“这里删掉”对学习帮助有限，因为它们没有说明偏差如何进入全文。教师反馈可以改写为关系描述：“这句话在局部与上一句顺接，但后文没有再使用它提供的区别，因此目前像问题移植”；“此处反方看似绕开中心，却让结尾的限定成立，删除会使结论变成无条件口号”；“首尾都重复观点，但证据来源相同，还没有形成异源回判。”

这种语言不要求教师使用跨学科术语。超分子、第三体和闭合相位属于研究者的发现工具，课堂只需保留其操作后果：多处关系、不同来源、遮蔽测试、功能去向。否则概念创新会变成新的术语负担。

（四）自动写作评价与生成式人工智能

自动评分系统容易利用长度、连接词、主题相似度和局部流畅度，这些特征可能与分数相关，却正好复制本研究暴露的高 N 问题。生成式模型也能稳定生成段内连贯、首尾回扣的文本，但重复中心句不等于异源关系。一个更严格的系统应实施反事

实消融：遮去邻句和总论点后预测某一单元的功能，再检测来自不同段落与不同关系层的证据是否提供增量。

不过，自动系统不能把 Δ_remote 直接当作新的总分。算法可能通过长度和主题线索间接猜中，也可能对短文、隐喻和文化互文不公平。合理用途是诊断与提问：指出哪些判断完全由局部完成，邀请写作者检查远处承诺；而不是把“远距越多”设为普遍规范。任何自动化版本都需要与人工盲评、困难非作文对照和不同语言体裁共同验证。

（五）课程设计：从一次成文转向关系重构

若作文被理解为组织选择的书面行动，课程就不应只比较初稿与终稿的分数，而应记录关系如何改变。学生可以提交“关系变更说明”：本轮新增了哪条异源证据，删除了哪处伪远距复述，哪一处原被当作错误的摩擦单元被重新确认有承载功能。教师由此评价的不是修改数量，而是判断是否变得更有依据。

这种设计也有助于区分文体。议论文的回判可能发生在主张、反例与限定之间；记叙文可能发生在早期细节、角色选择与结局重释之间；抒情散文可能依靠意象复现、声音转调和题旨浮现。异源回判不是把所有作文变成论证树，而是提供一个跨文体问题：远处的不同关系是否让局部选择获得了后来才显现的必要性？

十一、限制与确认性研究方案

（一）当前证据的六项限制

第一，样本只有十二篇且均为英语说服力学生作文，不能代表中文作文、叙事文、考试作文或专业散文。第二，七十二项扰动嵌套于篇内，不是七十二个独立文本；本文只报告描述性比例，没有进行总体显著性推断。第三，单一分析者知晓假说与阈值，编码可能受期待影响，也无法估计评者一致性。第四，操作细则存在迟注册，两处 R2 运行例外可能提高承载命中。第五，因充分停止规则触发，没有生成顺序与拼盘对照，故不能估计原文的相对优势。第六，AAE2 的标注本体偏重段内论证树，远程异源关系需要另行标注。

还有一个更深限制：扰动研究测量的是文本对人为编辑的响应，不一定等于自然写作中的结构形成。写作者可能在生成过程中使用读者无法恢复的计划，读者也可能通过体裁知识补足未显式关系。反事实编辑是一扇窗口，不是整体性的全部。

（二）确认性样本与对照

下一轮建议至少纳入三十六篇公开作文，按文体、长度和质量分层；同时为每篇建立三类困难对照。第一类为同题拼盘，保持段内连贯和主题分布，只破坏跨段来源关系。第二类为中心句复制对照，保持高词汇复现与首尾一致，检验同源重复是否冒

充多路。第三类为非作文复杂文本，包括证明、合同或规范的短片段，匹配长度与内部约束强度，检验作文特异性。

中文研究不能简单翻译英语操作。汉语省略、零形回指、话题链和意合关系会改变邻域定义；考试作文的题目材料又可能成为全文不可移除的共同源。应预注册“题目依赖”作为独立维度，而不是一律遮去。短文与长文还需使用相对功能距离，不宜以固定句数定义远距。

（三）操作—命运解耦设计

每类操作都经过预试，使其在理论上可能产生四种命运。可以采用“内容中性操作库”：对目标单元分别实施保义改写、立场反转、来源替换和功能位置迁移，再由与编码者分离的操作团队生成材料。命运编码者不知道操作类别和研究假说，只看到随机化后的条件。若仅凭操作标签的交叉验证准确率仍远高于文本模型，说明混淆未被解除。

R、B、C、U 的手册需加入否定性证据。例如判 B 不只要说明删除后哪里变弱，还要证明这种变弱不能由纯长度减少或指称断裂解释；判 C 要证明远程承诺保持，而非“意思差不多”的直觉；判 R 要排除语法和外部常识的单源判断；判 U 则要记录缺少哪种锚点。编码者应先独立标注路径，再标命运，以减少用结果反推证据。

（四）盲评、可靠性与统计单位

至少两名不知道原假说的编码者独立工作；分歧在主要分析冻结后再仲裁。名义命运类别可报告 Krippendorff's α ，路径跨度和异源层数可报告适合其尺度的可靠性指标。仅有高总体一致率不够，还要分别报告稀少类别与路径判定的一致性，因为类别不平衡会使表面一致率虚高（Krippendorff, 2004, 2018）。

主要推断单位应为作文。可先在每篇内计算 Δ_remote ，再报告篇级分布与置信区间；若使用扰动级模型，应以篇为随机效应并考虑操作类型。验证集必须在编码手册冻结后一次性运行。探索性分析与确认性分析分开报告，不把先导研究的九项远程案例重新用于确认。

（五）预注册裁决建议

V2 可以写下更严格的保留条件：原文的篇级 Δ_remote 中位数为正；相对于问题拼盘的配对优势达到预设最小效应；至少三分之二作文方向一致；操作标签单独预测命运的准确率不超过预设上限；盲评者对命运及异源路径达到可接受可靠性；复杂非作文对照不得取得同等模式。若只能区分高低质量而不能区分类别，则明确降级为作文质量维度；若 Δ_remote 接近零或由长度完全解释，则淘汰当前测量；若异源路径无法可靠编码，则暂停概念。

这种裁决比追求一次“显著”更重要。概念研究的危险不是没有数据，而是任何数据都能被重新叙述为支持。预注册的任务是限制叙述自由，让概念必须为某些结果付出代价。

十二、SDE V3.9 门控与研究裁决

（一）六道门的结果

表 4 SDE V3.9 六道门控及本轮裁决

门控	状态	裁决依据
三学科源机制吸收	通过	三源分别改变了关系、命运与测量设计，非装饰性比喻
碰撞有效性	通过	27 个碰撞中 21 个有效、2 个边界、4 个无效
理论新颖性	暂通过	“局部命运由多路关系分流”及其失败后的远距增益仍具研究增量
可复现性	通过但折损	数据版本、哈希规则、样本编号与逐项编码保存；存在迟注册和例外
现实数据门	未通过	N=0.8750 触发预注册淘汰线；操作标签预测率 0.8611
反证公开	通过	高 J 与不利 N、设计混淆、停止规则和边界反例同时报告

源机制门与现实数据门可以给出不同结果。三学科确实产生了新的问题与测量想法，不等于第一次量表已经有效。同样，量表失败也不自动证明所有理论直觉错误。严格裁决必须说明失败发生在哪一层：本轮被否决的是“全文高判差覆盖率足以表示作文整体性”的操作化；没有被直接检验的是经修订的远距异源增益；不能再肯定的是“作文就是内生判差体”。

（二）必须撤回、可以保留与尚待检验的命题

必须撤回三项。第一，不能把“内生判差”写成作文定义。第二，不能用 $J=0.9722$ 宣称获得经验支持。第三，不能把 R/B/C 分布当作文本自发分流的证据，因为操作类别强烈预编码命运。

可以保留三项。第一，局部扰动的命运比简单对错更丰富，承载与消隐值得继续研究。第二，局部窗口是任何整体性指标必须战胜的强基线。第三，闭合量的协方差警告可以转化为文本路径的异源性要求。

尚待检验的也有三项。第一， Δ_remote 是否在作文中稳定大于零。第二，它能否区分作文与匹配的复杂非作文文本，还是只区分一般整体性。第三，它与人工质量判断、修改收益和跨任务迁移是否相关。把三类命题分开，是避免“理论没有被彻底证伪”被偷换为“理论已经成立”的必要条件。

（三）失败的知识价值

若研究只追求肯定结论，本轮会被描述成 97.22% 的成功率。但预注册的 N 使同一数据显露相反意义：97.22% 中大部分是局部可还原和操作预编码。失败由此不是研究的附录，而是概念精度的来源。它迫使“内部”变成“异源”，“全文命中”变成“远距增量”，“两条路径”变成“遮蔽共同源后仍有判力的路径”。

这种改写也回应了 What 型研究的基本责任。回答“是什么”不能只提供一个悦耳定义，还要说明什么不算、什么会使定义失败、失败后对象应如何重新划界。组合一的贡献，恰在它没有保护自己的第一次答案。

结论

本文从“作文是什么”而非“怎样写作文”出发，将超分子化学、第三体摩擦学与射电干涉闭合相位置于同一概念试验中。三者分别提供多点协同、偏差去向和关系不变量，第一次碰撞形成“内生判差”与 R/B/C/U 的偏差分流。若研究停在理论推演，这一概念足够新颖，也足够诱人；但公开语料的审计性先导给出了决定性的反证。

十二篇 AAE2 作文、七十二项最小扰动中，全文判差覆盖率达到 0.9722，却有 0.8750 的判断可由改动句和一条邻句复现；高把握子集的近邻可还原率更高。操作标签本身能够预测 0.8611 的全文命运。高命中因此主要来自局部连贯与设计同构，而不是已被证明的作文整体。依预注册，“作文是内生判差体”必须撤回，“偏差分流”不得宣称独立成立。

不过，反证没有让组合一归零。它把真正的问题从“全文能否判断”推进到“远处且来源不同的关系是否提供超过局部基线的判断增量”。据此，本文提出异源回判及 Δ_remote ：遮去邻句、重复总论点与同源复述后，跨段、跨层关系仍能否改变局部单元的命运。这个指标需要新的盲评、困难对照与篇级统计，当前尚未得到确认。

对“作文是什么”的阶段性回答因而是谨慎的：作文不是能在局部发现矛盾的作文本；它更可能是一种把经验、立场与语言选择组织成远距、异源、可回判关系的书面

行动。这里的“更可能”不是软弱措辞，而是证据边界。一个经受过失败的假说，比一个从未允许自己失败的定义更接近研究。

附录 A：样本锁定与可复现信息

确定性种子为 SDE-V3.9-PANCHA-20260817。十二篇样本按 SHA-256 升序锁定：

essay238、essay326、essay178、essay039、essay056、essay161、essay057、essay162、essay082、essay049、essay045、essay115。原始压缩包 brat-project-final.zip 的 SHA-256 为 9f49a6d388896d0f390445592ccc9dbfc87325553e3fe3f988eb252de7cdc4b5；train-test-split.csv 为 903176797ea73c45e3848019bc390882a2bdab520e5feb4b28fd75d95ec44285；prompts.csv 为 5ad978f277c5420a491df8a87bb62cb2efeecea62f18bdcff2fe1412182b6f4。

本轮句数纳入规则采用正则 `[.!?]+(?:=|(?:"'")?\s)|$)` 识别句末标点组。预注册文本曾因 Markdown 排版显示为双反斜线，运行前、读取入样正文语义内容前完成勘误。按可执行规则，402 篇均达到八句门槛。样本在正文阅读后不得因主题、文体、质量或结果方向替换。

附录 B：V2 最小编码记录模板

后续每项扰动至少记录：作文编号；文体与长度层；操作生成者；操作类别；目标跨度；局部窗口；被遮蔽的中心句与同源复述；远程锚点位置；经验源、立场源、组织源、语言源的层别；局部命运；远程异源命运；全文命运；判断把握；否定性证据；编码者编号；分歧与仲裁。主要分析冻结前不得查看操作标签—命运列联表。

参考文献

- Beaufort, A. (2007). *College Writing and Beyond: A New Framework for University Writing Instruction*. Logan, UT: Utah State University Press.
- Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berthier, Y., Vincent, L., & Godet, M. (1992). Velocity accommodation sites and modes in tribology. *European Journal of Mechanics A/Solids*, 11(1), 35–47.
- Blackburn, L., Pesce, D. W., Johnson, M. D., Wielgus, M., Chael, A. A., Christian, P., & Doleman, S. S. (2020). Closure statistics in interferometric data. *The Astrophysical Journal*, 894, 31. [doi:10.3847/1538-4357/ab8469](https://doi.org/10.3847/1538-4357/ab8469)
- Chael, A. A., Johnson, M. D., Bouman, K. L., Blackburn, L. L., Akiyama, K., & Narayan, R. (2018). Interferometric imaging directly with closure phases and closure amplitudes. *The Astrophysical Journal*, 857, 23. [doi:10.3847/1538-4357/aab6a8](https://doi.org/10.3847/1538-4357/aab6a8)
- de Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to Text Linguistics*. London: Longman.
- Descartes, S., & Berthier, Y. (2002). Rheology and flows of solid third bodies: Background and application to an MoS_{1.6} coating. *Wear*, 252, 546–556. [doi:10.1016/S0043-1648\(02\)00008-X](https://doi.org/10.1016/S0043-1648(02)00008-X)
- Fillot, N., Iordanoff, I., & Berthier, Y. (2007a). Modelling third body flows with a discrete element method—A tool for understanding wear with adhesive particles. *Tribology International*, 40(6), 973–981. [doi:10.1016/j.triboint.2006.02.056](https://doi.org/10.1016/j.triboint.2006.02.056)
- Fillot, N., Iordanoff, I., & Berthier, Y. (2007b). Wear modeling and the third body concept. *Wear*, 262(7–8), 949–957. [doi:10.1016/j.wear.2006.10.011](https://doi.org/10.1016/j.wear.2006.10.011)
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. [doi:10.2307/356600](https://doi.org/10.2307/356600)
- Godet, M. (1984). The third-body approach: A mechanical view of wear. *Wear*, 100(1–3), 437–452. [doi:10.1016/0043-1648\(84\)90025-5](https://doi.org/10.1016/0043-1648(84)90025-5)
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. [doi:10.1177/0741088312451260](https://doi.org/10.1177/0741088312451260)
- Iordanoff, I., Berthier, Y., Descartes, S., & Heshmat, H. (2002). A review of recent approaches for modeling solid third bodies. *Journal of Tribology*, 124(4), 725–735. [doi:10.1115/1.1467632](https://doi.org/10.1115/1.1467632)
- Jennison, R. C. (1958). A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent. *Monthly Notices of the Royal Astronomical Society*, 118(3), 276–284. [doi:10.1093/mnras/118.3.276](https://doi.org/10.1093/mnras/118.3.276)

- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411-433. doi:[10.1111/j.1468-2958.2004.tb00738.x](https://doi.org/10.1111/j.1468-2958.2004.tb00738.x)
- Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology* (4th ed.). Thousand Oaks, CA: Sage.
- Lehn, J.-M. (1988). Supramolecular chemistry—Scope and perspectives: Molecules, supermolecules, and molecular devices. *Angewandte Chemie International Edition in English*, 27(1), 89-112. doi:[10.1002/anie.198800891](https://doi.org/10.1002/anie.198800891)
- Lehn, J.-M. (2002). Toward complex matter: Supramolecular chemistry and self-organization. *Proceedings of the National Academy of Sciences*, 99(8), 4763-4768. doi:[10.1073/pnas.072065599](https://doi.org/10.1073/pnas.072065599)
- Mammen, M., Choi, S.-K., & Whitesides, G. M. (1998). Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie International Edition*, 37(20), 2754-2794. doi:[10.1002/\(SICI\)1521-3773\(19981102\)37:20%3C2754::AID-ANIE2754%3E3.0.CO;2-3](https://doi.org/10.1002/(SICI)1521-3773(19981102)37:20%3C2754::AID-ANIE2754%3E3.0.CO;2-3)
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 243-281. doi:[10.1515/text.1.1988.8.3.243](https://doi.org/10.1515/text.1.1988.8.3.243)
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151-167. doi:[10.1080/00335638409383686](https://doi.org/10.1080/00335638409383686)
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600-2606. doi:[10.1073/pnas.1708274114](https://doi.org/10.1073/pnas.1708274114)
- Nystrand, M. (1986). *The Structure of Written Communication: Studies in Reciprocity Between Writers and Readers*. Orlando, FL: Academic Press.
- Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3), 619-659. doi:[10.1162/COLI_a_00295](https://doi.org/10.1162/COLI_a_00295)
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32(2), 189-204. doi:[10.58680/coc198115912](https://doi.org/10.58680/coc198115912)

第三章 补节 读得出的那一部分，恰恰不是整体性

(一) 最近的祖宗与分离线

本章最硬的一段，是它对自己的否决：全文判差覆盖率 0.9722 看似极强，但 63/72 的判断只用改动句加一条邻句即可还原；更严重的是，不读作文、只看施加了哪一类扰动，逐篇留一预测即可命中 **62/72 (0.8611)**。

这个发现有一个精确的祖宗，本章零命中。

古鲁兰甘等人《自然语言推理数据中的标注人工痕迹》 (**Gururangan, Swamydip, Levy, Schwartz, Bowman & Smith, NAACL 2018**)

与波利亚克等人的「只看假设」基线（Poliak et al., *SEM 2018）。两篇同时发现：在当时最主流的推理数据集上，只看假设句、完全不看前提，模型就能大幅超过多数类基线——因为标注者在构造样本时留下了系统性的表层痕迹（否定词、含糊语、长度）。此后这一支长成了「捷径学习」的完整文献（McCoy et al. 2019 的词汇重叠启发式；Geirhos et al., *Nature Machine Intelligence* 2020）。

本章那条 0.8611 就是标准的「只看假设」基线：操作标签相当于假设，作文相当于前提，而预测只用了前者。本章把它诊断为「实验者预先设计的操作性质被误认成本文自身的组织能力」——诊断完全正确，但这套诊断已有成熟名字、成熟的对策清单（对照集、去偏采样、人工痕迹审计）和成熟的报告规范。

分离线：那一支处理的是数据集与模型，本章处理的是概念与量表。古鲁兰甘等人的结论是「这个基准测不出它声称要测的能力」，补救办法是换数据；本章的结论是「这个操作化定义把实验设计的性质算进了对象的性质」，补救办法是换定义——本章据此把「内生判差」整个撤下，改提远距增益。前者修数据以保住概念，后者弃概念以保住诚实。这是一处真实的差别，而且对本章有利。但正因为有利，更不该省略：一个已经在别的领域被命名、被制度化的失败模式，在本领域第一次出现时应当被认领，而不是重新描述。

编者的判断：本章补上这条分界之后，它的方法论价值不但不降，反而上升——它成为「捷径学习审计」在写作研究中的第一个案例，并且证明了这套审计可以推翻的不止是数据集，还有概念。

（二）与本卷其他章的关系

本章与第七章有一条隐蔽的接口，两边原稿都未写。本章证明的是「局部窗口能还原的判断，不构成整体性证据」；第七章证明的是「关闭动作不进任何账」。合起来是一条更强的话：我们现有的账本不但漏记了一些东西，它记下的那部分还会系统性地高估自己的解释力。建议在成书时于第七章末互引一句。

（三）本章欠账

样本仅十二篇作文、七十二项判断，且原稿自陈存在迟注册与例外。「异源回判」是本章反证之后提出的新假说，尚未被任何数据检验——本章的可信度目前几乎全部来自它否决旧概念的那部分，而非它提出新概念的那部分。读者应据此把本章读成一份证否报告加一个待检假说，而不是一个已成立的理论。

编二

什么时候它成为一篇作文

编二 什么时候它成为一篇作文

如果身份要靠扰动才显形，那么它显形之前，它在哪里？

这一编把「作文」从一个状态词改成一个时间问题，并主张至少有三只钟在走：一组意义关系什么时候在写作者身上稳定下来；它什么时候显成文章；它什么时候被制度确认为「作文」。三者可以彼此错位。

第四章处理前两只钟的错位，给出两个反常对象。一个孩子面对一张空白的纸，可能已经越过某个过去没有名字的时刻——撤掉支架、换一种表达方式，那些关系仍会从他这里重新长出来。文章还没有发生，作文已经发生了。反过来，一篇结构完整、语言通顺的八百字，可能一撤掉模板与提纲就什么都不剩。文章已经完成，作文没有发生。前者是隐作，后者是有文未作。

第五章处理身份归属。它的结论最反直觉：终稿的常规质量不但不能证明作文身份，在公开数据上还是反向的——七项评分全部显示质量更高的那一侧恰恰不可能拥有学生本人的写作过程。因此需要记录的并不是机器文字占多少比例，而是那几处真正改变了方向的裁决由谁完成。这一章给出的记录规格只有三类节点、每节点三件事，比现行任何过程档案都轻。

这一编也划出了本卷对人工智能问题的立场：机器参与的是显形，未必替代发生；而一份从未使用机器、却由唯一正确的范文与固定结构完成的作业，可能比前者离作文更远。

第四章

文章还没出现，作文已经发生

何谓作文？

——作文作为“作定态”的发生：论先于文本显形的“隐作”

What Is Composition?

Composition as the Emergence of a Determined Composing State

SDE 学科碰撞创新法 v3.9 · What 型三维度实验稿

晶体学／材料相科学 × 发育生物学 × 财务会计确认理论

版本：2026-08-17

摘要

作文教育长期以文本为最直接对象，过程写作理论则把作文扩展为构思、起草、修改、反馈与反思等活动。两者仍共享一个很少被追问的前提：作文只有在某种可观察的文本或过程事件出现以后才开始存在。本文提出另一种区分。发育生物学表明，决定某一对象未来方向的 **commitment** 可以先于外显 **differentiation**；是否已经 **determined**，需要通过改变环境以后方向是否继续保持来检验。结合晶体结构与亚稳态研究关于“结构、稳定与终局并不等同”的约束，以及财务会计中“满足存在定义与进入正式确认不是同一判断”的区分，本文提出“作定态”与“隐作”两个概念：作定态是写作者在特定表达任务中形成的一种意义组织状态，在题目、证据域和受众不变而表层支架被撤去以后仍能够自行再生；隐作则是作定态已经发生而文本尚未显出可辨差异的作文。据此，作文不是文本结构，不是写作步骤的总和，也不是制度确认的提交物，而是写作者进入作定态的一次发生；文章只是这一发生的显形。本文进一步提出中性扰动检验、四格辨别模型、机制证伪条件，并讨论它与前写作、知识转化、写作迁移、能动性、写作过程数据及“预闭分叉”的判决性差异。

关键词：作文；作定态；隐作；写作过程；写作评价；生成式人工智能

Abstract

Writing education conventionally identifies composition either with a textual product or with observable composing activities such as planning, drafting, revising, and reflecting. Both approaches tend to assume that composition becomes identifiable only after some textual or procedural manifestation has occurred. This article proposes a different ontological distinction. Developmental biology demonstrates that commitment and determination may precede overt differentiation and can be detected only through persistence following an environmental displacement. Constrained by crystallographic accounts of structural state and metastability, and by accounting theory's distinction between existence and recognition, the article introduces the concepts of a determined composing state and latent composition. A determined composing state exists when a writer's task-specific organization of claim, evidence, qualification, and audience can regenerate after surface scaffolds are removed while the task, evidence domain, and audience remain constant. Latent composition names the case in which such a state already exists although no differentiating text is yet visible. Composition, on this account, is not identical with the text, the sum of writing steps, or an institutionally

recognized submission; it is the event through which a writer enters such a self-sustaining state.

Keywords: composition; latent composition; writing process; determination; writing assessment; generative AI

一、作文究竟从什么时候开始存在

一名学生坐在桌前，纸上一个字都没有。老师问：“你开始作文了吗？”最自然的回答似乎是：“还没有，他还没写。”如果学生已经列出提纲，我们也许会改口说：“开始构思了，但还没有真正成文。”如果他写了三段，我们才更确信：“作文正在发生。”

这种判断如此自然，以至于很少有人意识到，它背后其实藏着一个本体假设：作文的存在必须比作文的显形更晚，或者至少同时。没有文字，就难以说“作文已经存在”。

过程写作理论已经大幅改变这种成品中心主义。Rohman 在 1960 年代就把 pre-writing 与 discovery 联系起来；Emig 进一步论证写作本身是一种学习方式；Flower 与 Hayes 将写作解释为复杂、递归的问题解决活动；Bereiter 与 Scardamalia 又区分知识陈述与知识转化。今天再说“作文不是成品，而是过程”，已经不足以构成新的回答。

更近的数字写作研究甚至能够记录停顿、输入、删除、导航、复制、修订以及不同写作意图。2025 年前后的过程研究已经尝试自动识别写作 phase/change points；ScholarWrite 等数据集则直接在编辑事件层标注 planning、implementation、revision 等写作意图。2026 年 KEYSKORE 使用数千个写作 session 的击键特征，显示文字尚不足以常规评分时，过程数据已经具有预测信息。

于是一个更难的问题出现：如果文字还没有显出差异，甚至击键行为也还没有显出决定性差异，作文是否可能已经发生？

二、我们需要区分“显出来”和“已经成为”

发育生物学提供了一种很少进入作文理论的区分。一个细胞最后表现为什么样，叫作 differentiation 的结果并不难理解。困难在于：它究竟在什么时候已经“决定”走向某个命运？

经典发育生物学把这件事分得很细。在 overt differentiation 出现以前，细胞可能已经发生 commitment；其中 specification 与 determination 又不是同一程度。判断一个组织是否 determined 的经典办法，不是继续盯着它原来的样子，而是把它置于改变后的环境，观察它是否仍沿原有发展方向继续。

因此，“已经成为”与“已经显出来”不必同时发生。更准确地说，一个变化可以先存在于对象的继续方式中，后来才存在于对象的表面差异中。

这提供了重新思考作文的入口。一个学生的纸面也许仍然是空白的。但如果此时换一种语言实现方式，撤掉刚才看的范文，关闭 AI，拿走已有提纲，甚至要求他不用刚才任何句子重新开始——他仍然会重新建立大体相同的主张与证据关系，仍然在同一个反例处限制自己的论点，仍然把同一类信息解释给同一个预期读者，那么我们为什么一定坚持说作文还没有发生？也许没有发生的只是文章。

三、本文把这种状态称作“作定态”

所谓“作定态”，指的是：在一个特定表达任务内部，写作者形成了一组能够在中性环境改变以后自行再生的意义关系。

这里的每个词都有边界。它首先是任务内的。一个学生在“是否应全面禁止校园手机”这个问题上形成稳定论证，不意味着他已经获得一般性的“写作成熟”。所以它不能被写作迁移理论吞掉。

其次，它是意义关系，而不是句子记忆。如果撤掉旧句子以后，学生只是把原文背出来，不构成有力证据。我们关心的是：主张与证据为什么相连；哪个反例限制了主张；哪个条件使结论缩小；为什么面对这一受众要采用这一解释顺序。

最后，它必须能够再生。所谓再生，不是“我还记得老师给我的答案”，而是支架离场后，这组关系仍然由写作者自己重新长出来。

四、为什么“稳定”不能等同于“优秀”

这时候晶体学给出第一条重要限制。同一个化学组成能够形成不同晶体结构；不同 polymorph 又可能表现出不同性质。

更重要的是，实际形成的晶相并不天然等于终极最稳定或者“最好”的相。成核动力学、增长速度和能垒可以使亚稳结构产生并持续。

因此，如果把这一纪律带回作文：一个论证能够在撤去支架以后继续自行再生，只证明它已经成为这个写作者当前的稳定组织方向。它完全可能是错的、狭窄的、有偏见的、证据不足的，甚至可能非常差。

于是本文必须坚持：作文发生与作文质量是两个变量。传统作文评价经常默认越成熟的文章越能证明作文发生得充分。本文恰恰预测：会存在一类质量很差但作定充分的作文，也会存在一类质量很好但作定极弱的文本。后一类才是最危险的。

五、“有文未作”：一类过去容易被漏掉的对象

设想两个学生。他们最后交出的文本质量相当。第一名学生自己阅读材料，开始时形成观点 A，后来发现证据冲突，转向 B；再根据受众改变论证边界，最后形成 C。

第二名学生拿到一个高质量支架：中心论点已经给出；证据已经按优先级排列；反方意见和回应已经配对；文章结构也已经设计。他没有复制任何一句话，所有文字都是亲自写的。如果要求他改写句子，他也能做到。

但撤掉那套支架，让他回到同一个问题和同一组材料，他无法重新生成那些关系。这两个人都“有文”，但本文认为，他们没有处于同一种作文状态。

第二名学生属于“有文未作”。这里的“未作”不是说他没有劳动，更不是说他作弊，而是说文本已经显形，而使这篇文本成为这一种意义组织的状态，还没有在写作者身上形成自续性。

六、“隐作”：一类文章出现以前已经存在的作文

现在反过来看。第三名学生面前仍是一张白纸。但他已经反复比较过材料，知道自己为什么不能采用最初那个强结论；知道哪一份证据只能支持较窄的判断；知道如果面向家长而不是同学，哪一段必须换一种解释；也知道自己的反例不能被简单删掉。

此时让他口述，他能重新组织出来。隔一段时间拿走原来的笔记，他仍能重新组织出来。换一张空白纸、换一种表达顺序，他仍会重新建立类似的核心关系。文字还不存在，但某样东西已经发生。

本文把这一类对象称为“隐作”。它不是“腹稿”这个旧词的学术化包装。腹稿只告诉我们“脑中有内容”。隐作要求的不是内容拥有，而是：在中性扰动之后仍能保持方向的关系自续。

一个人脑中背着一篇范文，不属于隐作。一个人记着老师的三个分论点，也不自动属于隐作。只有当关系本身能够在撤掉当前实现以后重新形成，才有资格讨论作定。

四格辨别：显形与作定的交叉

	意义组织经中性扰动后不能自续	意义组织经中性扰动后能够自续
文本尚未显形	探索态：尚未形成作文	隐作：作文已经发生，文章尚未出现

	意义组织经中性扰动后不能自续	意义组织经中性扰动后能够自续
文本已经显形	有文未作：文本完整，作文状态未形成	显作：作文状态已经在文本中成形

七、于是“一篇作文”至少有三只钟

我们通常把三个时刻混在一起。第一只是作定钟：写作者的意义组织什么时候第一次表现出扰动后的自续？第二只是显形钟：这种状态什么时候第一次变成可辨识的文字、结构、图示、口述或者其他外在实现？第三只是确认钟：学校、教师、考试机构什么时候把它正式登记为一篇作文、一次作业或者一个得分对象？

三只钟过去往往走得很接近。学生先构思，很快写出来，然后交上去。所以我们误以为它们是一只钟。实际上它们可以严重错位。

七之二、三只钟的形式化：六种次序，四种可观测

把三只钟写成三个时刻，可以让后面的讨论摆脱比喻。

记 t_D 为作定时刻：在任务 T 上，写作者的核心意义关系第一次达到「撤去当前语言与非必要支架、而题目、证据域与受众不变时仍能由他自行重新生成」的状态。记 t_M 为显形时刻：文本第一次呈现出足以把该组关系与其他可能组织区分开来的可辨差异。记 t_R 为确认时刻：制度第一次把这份产出登记为「一篇作文」。

三个时刻在逻辑上可以任意排列，共六种次序。但其中两种在经验上不可能出现： t_R 不能早于 t_M ，因为制度确认需要一份可提交的文本；这就把六种压到四种。

次序一： $t_D < t_M < t_R$ 。这是本章所说的隐作在完整走完之后的形态。关系先在写作者身上稳定，随后显形，最后被确认。传统作文教育默认的就是这一种，只是它把 t_D 与 t_M 视为同一时刻，因而看不出这是一种次序而不是一个事实。

次序二： $t_M < t_D < t_R$ 。文本先呈现出可辨差异，关系随后才稳定。这不是失败，恰恰是写作最有价值的一种形态——写作者在写出来之后才知道自己要说什么。本章不主张这一种低于次序一；本章只主张它们不是同一件事。

次序三： $t_M < t_R$ ，而 t_D 不存在（或落在观察窗之外）。这是有文未作：文本完成、制度确认，而那组关系从未在写作者身上稳定。模板作文、代写、以及本章第十五节所说的「一个机器字都没用却只执行了唯一立意」，全部落在这里。

次序四： t_D 存在，而 t_M 与 t_R 都不发生。关系已经稳定，却因为时间、意愿、任务中断或不打算提交而从未成文。这是本章最难处理的一格：按本章定义作文已经发生，而账本上什么也没有。第二十七节所说的那个洞就在这里

——它使本章无法完全脱离「可观察产出」来运行，因为要判定 t_D 存在，至少需要一次扰动后的重构，而重构本身就是一次产出。

这四种次序不是四类学生，是同一个学生不同任务上的四种可能。本章反对把它们做成人格标签，理由见第十七节。

需要说明的是三只钟的测不准之处： t_M 与 t_R 都可以从记录里直接读出（文本差异、提交时间），而 t_D 在原理上只能被区间估计——一次成功的扰动重构只能证明「 t_D 不晚于此刻」，一次失败只能证明「在此刻的条件下未能读出」，而不能证明 t_D 不存在。因此本章所有关于 t_D 的陈述，都应读作对区间上界的陈述。这条限制在第二十五节的第一条死亡条件里同样起作用：那一条要求的不是「找不到隐作」，而是「在多次不同时间点的扰动中，从来没有出现过 t_D 早于 t_M 的个案」。

八、第一种错位：作定早于显形

这就是隐作。学生尚无成文，但状态已经形成。传统成品评价看不到它，击键记录也未必看得到。只有一次恰当的扰动才能显露。

这与发育中的 **determination** 结构最接近：当前表型不提供足够证据，改变环境以后继续沿原方向发展，才揭示先前已经存在的状态。

注意，这里不是说作文“就是细胞”。我们借用的是一条可证伪结构：如果一种身份真的已经成立，它在换表面环境以后表现出比未成立状态更强的方向持续性。如果作文里不存在这个现象，本文就失败。

九、第二种错位：显形早于作定

这是“有文未作”。它尤其容易发生在高度支架化环境。模板提高结构整齐度，范文提高语言预测性，AI 能够提高局部表达效率，电子提纲工具也可能改善某些结构表现。

因此，“支架有作用”与“状态已经归属于学习者”是两个判断。一个极端情况甚至可能是：文章越漂亮，越让老师误以为状态已经形成。

十、第三种错位：作定与显形都发生，制度却不确认

这里会计确认理论开始真正发挥作用。财务会计并不把“真实存在”与“正式确认”视为同义。概念框架允许一个项目符合经济定义，却不进入正式报表确认。

作文也一样。学生可能在日记里经历了一次完整作定，可能在给朋友写长邮件时重新组织了自己的论点，可能在没有交给教师的废稿里第一次真正形成了自续状态。这类事件完全可能具有本文意义上的作文发生，却没有成为学校账本里的“作文”。

因此：作文存在不等于作文被课程确认；反过来，课程确认也不等于作文一定已经发生。学校可以正式收走一篇“作文”，学生得到 90 分，这仍不能从逻辑上证明作定发生在学生身上。

十一、这和“知识转化”究竟有什么区别

Bereiter 与 Scardamalia 的 knowledge-transforming 传统非常接近本文，因为它已经指出成熟写作不是把已有知识倾倒出来，而是在问题解决中改变内容与修辞问题空间。

所以本文若只说“真正作文会改变人的想法”，没有任何新意。真正的判决线是：知识有没有大幅改变，与作定态有没有形成，是两回事。

例如学生原本就掌握充分知识，一篇说明文可能没有让他产生明显新知识，但在组织“主张—证据—受众”的具体关系时已经形成高度自续状态。反过来，一次讨论可以彻底改变学生对某问题的知识，但他仍未形成任何能支撑特定作文的稳定组织。

如果实验证明这两种现象总是同涨同落、不可分离，本文就被 knowledge-transforming 吸收。

十二、它和写作迁移的区别更重要

写作迁移研究关心学习者能否把已有知识和实践带到新的情境，并发展出 adaptive transfer 等框架。本文刻意不在这个层面定义作定态。

作定检验保持三件东西不变：题目不变；证据域不变；受众不变。改变的只是已有措辞被删除、工具界面改变、原提纲撤走、表达顺序被打乱、实现媒介改变或时间被拉开。

因此，迁移测试是在问“换一个任务还能不能用”；作定测试是在问“还是这个问题，拿掉当前身体以后，它还能不能重新长出来”。这是二者最干净的分界。

若实际数据证明，控制一般写作能力和 transfer 后，作定检验没有任何独立信息，那么本文应该退回写作迁移理论。

十二之二、与「学习」和「表现」之别的正面交手

到这里必须处理本章最危险的一位近邻。它不在写作研究里，而在学习心理学里，而且它比本章早了几十年。

学习心理学有一条基础区分：学习指相对持久的、支持长期保持与迁移的变化；表现指训练期间可以观察和测量的东西。这条区分的要害在于，后者常常是前者的不可靠指标，而且两个方向都成立。

一个方向是：学习可以发生而表现不显。这一支的经典实验是二十世纪三十年代的迷宫研究——一组动物连续十天在迷宫中不获得奖励，其错误率与从不获得奖励的一组没有差别；第十一天引入奖励，第十二天它们的错误率立刻降到与全程获得奖励的一组相当。学习一直在发生，只是没有被表现出来，直到条件改变才显形。这一现象被称为潜伏学习。

另一个方向是：表现可以提升而学习不发生。集中练习、恒定情境、即时反馈、高度支架，都能使训练期间的表现显著变好，却使延迟保持与迁移变差。围绕这一现象发展出的整个纲领，主张学习需要「合意的困难」。

而这一支规定的测法，与本章的中性扰动检验是同一件事：学习不能在训练条件下测量，必须在撤去当时条件之后的延迟测试或迁移测试中测量。

看到这里，读者有权提出本章最强的一个反驳：隐作不就是潜伏学习，有文未作不就是表现与学习脱钩，中性扰动不就是迁移测试？如果是，本章不过是把一套四十年的教育心理学结论换成了语文教育的说法。

本章的回答分三条，每条都给出判决性对照。

第一条：被测的东西不同。学习／表现之别测的是一个人的能力，其单位是个体，其时间尺度是长期保持——正因如此，它的整个方法论建立在「训练期与测试期分离」之上，一次训练之内的波动按定义不算数。本章测的是一次任务上的意义组织，其单位是「这个写作者在这道题上」，其时间尺度落在本次写作之内。本章明确主张作定可以发生在同一次写作过程的中途，而且那个时刻本身正是要测的东西。

判决性对照：把学习／表现的标准做法搬到本章场景，唯一可能的操作是「一周后再考一次」；而本章要求的是在同一次写作尚未完成时做一次不改变题目、证据与受众的扰动。若两种测法的结果长期——对应——即同一次写作中途的扰动重构，与一周后的延迟测试给出相同的分类——那么本章的时间尺度主张失败，隐作应并入潜伏学习。

第二条：本章允许一次不可迁移的发生。一个学生完全可能在这道题上进入作定态，而在下一道结构同型的题上完全不能。按学习／表现的标准，这根本不算学习，因为它不迁移；而按本章的标准，作文照样已经发生。本章因此不是迁移理论的一个特例，它是一个明确容纳「局部的、不迁移的、一次性的」发生的理论。

这条差别有教学后果。若采纳学习／表现的框架，一次不迁移的写作是无效的，教师应当把资源投向能迁移的训练；而按本章，一次不迁移的作定仍然是一次真正发生的作文，它的价值不需要靠下一次来兑现。这也是本章与第五章分工的根源（见本章第十三节与本卷第五章）。

判决性对照：若数据显示，凡是在本任务中出现高自续的写作者，在结构同型的新任务上必然也出现高自续——即局部作定与迁移不可分离——那么本章多出来的那一格是空的，它应当收缩为迁移理论的一个测量变体。

第三条：本章多问了一个「什么时候」。学习／表现之别在形式上是一个二分（学到了没有），而本章是一个时序命题：意义组织稳定、文本显形、制度确认，是三只可以互相错位的钟。潜伏学习的实验里没有第二只钟——动物不生产

一份可供他人评分的成品，因此那里不存在「显形早于稳定」这种可能。本章的两个新对象恰恰产生于第二只钟的存在：隐作是 t_D 早于 t_M ，有文未作是 t_M 发生而 t_D 不发生。去掉成品这一环，这两个对象都不成立。

判决性对照：在一个没有成品的任务上（例如纯口头讨论），本章的两个新对象应当无法定义，而学习／表现之别照常适用。若有人在无成品任务上成功地测出了隐作，那说明本章对第二只钟的依赖是虚假的，本章应退回学习／表现的框架。

一处必须承认的收窄。本章第二十二节曾把自己的位置表述为「即使过程数据非常丰富，仍有必要区分已经发生的动作与只能在环境改变以后才显露的潜在状态」。补上这一族祖宗之后，这个表述过强了——「只能在条件改变后才显露」正是这一支几十年来的标准立场，本章并没有第一个说出它。本章真正独有的收窄应当是这一句：

在单次任务之内、在成品尚未分化之时，这种状态可以被一次不改变题目、证据域与受众的扰动读出。

三个限定缺一不可：单次任务之内（不是延迟测试）、成品尚未分化之时（不是事后回溯）、扰动不改变任务三要素（不是迁移测试）。凡是同时满足这三条的测量，此前的文献里没有；凡是缺其中任何一条的测量，此前的文献里都有。

十三、它与上一篇“预闭分叉”也不是同一个理论

同一产线里最危险的占位者，是前一篇关于“预闭分叉”的文章。前一篇问：哪些本来应该成为学生问题的可能方向，在学生遭遇以前就被教师、模板、AI 或其他外部结构关掉了？本篇问：学生自己的意义组织什么时候开始在撤去当前环境以后仍然自续？

二者可以发生四种关系。一个学生可能没有任何预闭——他充分见过各种可能性——但最后已经进入作定态。另一个学生可能被模板预闭了大量方向，却还没有真正内化剩下的一条路线；一撤模板，他甚至不知道该怎么继续。

所以，预闭描述问题空间被外部缩减；作定描述写作者内部关系何时开始自续。前者讨论“什么没让你见到”，后者讨论“什么已经长在你这里”。这条分界必须保留，否则第二篇只是第一篇的续篇换名。

十四、AI 会使“有文未作”大量增加吗

这是经验问题，不能靠直觉回答。当前生成式 AI 写作研究已经表明，学生会使用 invention、research、feedback、revision、多模态创作等不同环节使用 AI；composition pedagogy 对此也越来越强调 autonomy、reflection 和 process，而不是简单把任何 AI 使用视为同一种行为。

因此本文不主张“用了 AI 就没作文”。相反，本文允许一种看上去非常反常的情形：AI 写了相当一部分表面文字，学生却已经进入作定态。如果去掉 AI 输出，让他换语言重新开始，他仍能重新生成相同的论证关系。那么外部工具参与的是显形，未必替代作定。

另一名学生则可能一个 AI 字都没复制。他只使用机器给出的“最佳立意、最佳材料组合、最佳三段结构”，然后亲手完成全部文字。一旦撤掉支架，核心关系全部消失。这篇文章反而可能属于“有文未作”。

十五、因此“AI 率”可能测错了对象

检测机器生成文字的比例，可以回答文本材料从哪里来。它回答不了哪一种意义组织已经在写作者身上形成自续。

这就像用化学组成去替代晶型判断。组成当然重要，但同样组成可以处于不同结构状态。

因此，未来如果作文教育仍然主要依靠“AI 文字百分比”判断作文归属，理论上至少存在一次范畴错误：它拿显形材料的来源，替代了发生状态的判别。

十六、真正需要做的不是监控更多，而是做一次中性扰动

如果本文成立，评价作文并不需要永久监控学生全部写作过程。甚至相反，保存越多击键，不一定越接近我们真正想知道的东西。因为过程日志记录的是已经发生的动作。本文最关心的却是一种在动作表面还没有明显分化时已经形成的继续倾向。

所以最便宜的方法可能不是增加日志，而是在合适的时点做一次小小的中性扰动。例如撤掉原提纲、关闭刚才使用的 AI 对话、不许复用现成句子，同时给出同样题目、同样材料、同样受众，让学生重新以图示、口述或新文本建立关系。

我们看的不是字面重复，而是原来的“主张为什么受这份证据支持”是否重新出现；原来的限定关系是否重新出现；原来的受众适配是否重新出现。如果出现，我们获得的是比“你是不是自己写的”更具体的证据。

暂行研究读数：自续率 SR

自续率 $SR = \text{第二次独立重构中重新出现的核心关系数} \div \text{第一次关系图中的核心关系总数}$ 。

“核心关系”暂只计算四类：主张支持、证据支持、反例限定、受众目的关系。同义改写算同一关系；只重复关键词不算；单纯保持段落顺序不算；由新支架重新提供的关系不算；题目或证据集合变化则本次测量作废。

$SR \geq 2/3$ 只作为需要实验校准的工作阈值，不作为理论神圣化的边界。SR 是测量作定态的一把尺，不是新理论本身；删掉 SR，“隐作”这个对象仍须能够成立。

十六之二、自续率 SR 的清点手册

第十六节给出了自续率的定义。要让它离开漂亮话，需要一份可以让两个编码者算出同一个数的手册。以下八条即为本章承诺的最低清点规则。

第一，先声明关系图的边界。在扰动之前，由编码者从原稿中提取核心关系图 G_1 ，只登记四类关系：主张支持（某句为某主张提供支持）、证据支持（某材料为某主张提供经验依据）、反例限定（某处承认了一个使主张不成立的条件）、受众目的关系（某处的措辞或取舍由预期读者决定）。四类以外的一切关系——修辞效果、节奏、语体、结构美感——一律不登记。这不是说它们不重要，是说本读数不测它们。

第二，关系以「有序对+类型」为单位。记为 (a, b, τ) ，其中 a, b 是意义单元， τ 是四类之一。同一对单元之间可以同时存在两种类型的关系，计为两条。方向相反的两条计为两条。

第三，同义改写算同一关系，只重复关键词不算。判据是：在第二次重构 G_2 中，若存在一条关系 (a', b', τ) ，使得 a' 与 a 、 b' 与 b 指向同一意义单元且 τ 相同，则记为「重现」，无论措辞如何改变。若第二次只出现了 a 与 b 的名字而没有建立它们之间的 τ 型关系，不记为重现。

第四，四条作废条件。任一成立则本次测量作废，不得记为零：① 题目发生改变；② 证据集合发生改变（新增或撤走任何一份材料）；③ 预期受众发生改变；④ 第二次重构由新的支架提供了关系（例如编码者的提问本身指出了 a 与 b 的关联）。第四条最容易被违反，因此第二次重构的全部提问必须事先写定并原样复用。

第五， $SR = |G_1 \cap G_2| / |G_1|$ 。分母是第一次关系图中的核心关系总数，分子是重现的条数。仅重复段落顺序不计入分子。

第六， $|G_1| < 6$ 不报。关系图过小时，一条之差即造成 SR 跳动超过 0.16，超出本读数的分辨率。此时只报关系图规模，不报 SR。

第七，两名编码者独立编码，先算一致性再算 SR。关系图提取的一致性以逐条匹配计，低于 0.70 的样本进入退出日志，不得强行合并。退出日志必须随结果一并报告——这与本卷第九章所说的「只保留可顺利计算的样本，等于把看不见的问题从分母里删掉」是同一条纪律。

第八， $SR \geq 2/3$ 只是一个待校准的工作阈值。它来自「四类关系中重现三类」这个直观，没有任何经验依据。正式研究必须先用一小批样本校准阈值，且校准必须在看到主结果之前完成。若有人把 $2/3$ 当成本章的理论主张，那是误

读：**SR** 是测量作定态的一把尺，不是作定态本身。删掉 **SR**，隐作这个对象仍须能够成立。

十六之三、什么样的扰动才算「中性」

中性扰动是本章全部判据的支点，因此必须给出它的边界，否则「撤掉支架」可以被任意解释。

本章把扰动分为三类。

中性扰动：撤去当前的语言实现与非必要支架，而任务三要素（题目、证据域、受众）不变。典型操作包括：封存已写出的文字、关闭当前使用的机器对话、不允许复用现成句子、改换表达通道（由书面改为口头或图示）。它改变的是这一次组织意义所用的载体与脚手架，不改变要组织的是什么。

重编程扰动：改变任务三要素中的任何一项。换题、增删材料、改换预期读者，都属此类。这类扰动测的是另一件事——写作者能否把一组关系搬到新任务上，那是迁移，不是作定。

破坏性扰动：改变的是写作者本人的状态而非任务条件，例如在疲劳、时间压力或情绪干扰下要求重构。这类扰动的结果不可解释，因为失败可以归因于状态而非关系是否稳定。本章要求把它排除，而不是把它当作「更严格的检验」。

三类的分界有一个可执行的判据：问一句「原来那组关系在新条件下是否仍然是一个合理的答案」。若仍然是，扰动为中性；若已经不是（例如换题之后原来的主张不再切题），扰动为重编程；若无法判断，多半属于破坏性。

一个必须承认的灰区：改换表达通道究竟是不是中性的？由书面改为口述，确实不改变任务三要素，但它改变了组织意义的资源——有些关系在书面上易于建立而在口头上难以维持，反之亦然。本章的处理是把通道改变登记为一个变量而不是当作无害操作：同一研究中所有被试的通道改变方式必须一致，且结果须分通道报告。若数据显示 **SR** 系统性地依赖通道，则「中性」这个词本身需要收窄，而第二十五节的第六条死亡条件正是为此设的。

十七、但这套判据绝不能直接变成学生 **KPI**

这是本文必须对自己实施的限制。一旦学校听懂“自续”，最省事的做法很可能是：每篇作文强制两次重写；计算相似率；达到 **67%** 才叫“本人作文”。那会立刻摧毁这个概念。

学生将学习如何重复自己。教师会奖励路径锁定。新想法反而会因为第二次与第一次不同而被惩罚。最终，自续率越高，学生越不敢改变。

因此：作定态只能首先用于研究任务、支架和教学介入，不能直接用于高风险评价个人。会计确认理论对此给了很好的纪律：不是所有存在的东西都必须进入正式账本。

十八、它也不能奖励“固执”

还有另一个反噬。如果作定态意味着“扰动以后仍然保持”，那是不是越不改变越好？不是。

这正是发育生物学的第二个约束。**determination** 不是绝对不可逆；在强干预下，细胞命运可以被重新编程。作文也是如此。

当新的关键证据出现、当受众真正改变、当发现论点存在严重漏洞，优秀写作者恰恰应该改变。所以中性扰动必须严格限定：不改变问题本身，只改变当前实现条件。真正改变题目、证据和受众，那已经不是“测试有没有作定”，而是在给作定态施加强重编程条件。能够改变不是失败。

十九、由此可以重新理解教师反馈

教师反馈至少有三种不同作用。第一种是显形辅助：学生已经知道自己要表达什么，只是在语言实现上困难。教师帮助他表达，并不会自动消除作定。

第二种是定向促进：学生尚在探索，教师提出一个关键反例，使原来的若干松散关系重新组织。最后状态仍可能在学生身上形成。

第三种是代替作定：教师直接规定中心论点、证据组合和回应顺序。学生得到一个可以继续执行的文本方案，却没有形成撤掉该方案以后仍能自行重建的关系。

传统“反馈多少”的数量指标无法区分这三件事。关键不是老师说了多少，而是老师离开以后，什么仍然会从学生这里重新长出来。

二十、由此可以重新理解范文

范文也一样。范文可能造成 **fixation**。创意研究已经表明，示例可能限制某些参与者生成不同方案，但这种作用会随年龄、已有知识和任务条件发生变化。

所以范文不能被简单定义成好或坏。如果范文进入以后，学生只剩下复制它的组织关系，撤掉范文后关系消失，它只是替学生显形。

如果学生用范文对自己的方案形成比较、拒绝、重组，最后撤掉范文仍能再生新的关系，范文反而参与了作定。同一篇范文，在两个学生那里可以发生完全不同的教育作用。

二十一、由此也能重新理解“作文能力”

如果作文是作定发生，那么作文能力就不能只等于写得快、词汇多、句子好、结构完整，甚至也不能简单等于“有很多写作策略”。

一个更根本的问题是：这个人是否越来越能够把一个外部任务转化为离开当前支架以后仍能自行维持的意义组织？

这里我们第一次能够区分“会完成作文任务”和“作文已经开始在这个人身上发生”。两者当然高度相关，但未必同一。

二十二、现行写作评价已经走得比我们想象更远，但还没有走到这里

这次公开材料回跑让我不得不撤掉一个原先更简单的批判：现代考试并非只看漂亮成品。IB Language A 已经把理解解释、分析评价、聚焦组织和语言分别评价；AP Seminar 等任务还把书面论证与 presentation/oral defense 结合，这意味着制度确实已经试图越过单一文本，检查学生能否解释和捍卫自己的工作。

过程研究同样已经远远超过“看两份草稿”。现代击键日志和过程数据能够很早捕捉停顿、爆发式输入、删除、修改和导航等信号。

所以本文不能把自己装扮成第一个发现“过程重要”的理论。它真正提出的只剩下很窄的一步：即使过程数据非常丰富，仍然有必要区分已经发生的动作，与只能在环境改变以后才显露的潜在自续状态。这才是“隐作”的位置。

二十三、一个最小可执行实验

本文最便宜的机制实验可以在同一学校完成，不需要跨国比较。选取同一年级、同一写作任务的学生，统一题目、证据集合和预期受众，允许正常构思。

在随机时点暂时中断。不给学生保存原提纲或已有句子。间隔固定时间以后，以不同表层形式重新要求组织同一问题，例如把文字提纲改成口头关系图。

对四类核心关系编码：主张、证据、限定、受众目的。之后允许学生继续完成作文。

如果“作定态”真实存在，我们预期看到的不是一条连续曲线，而是：部分学生在明显文本差异形成以前，就已经出现较高的跨环境关系再现；另一些学生即使已经写出相当完整的文本，撤掉当前支架以后仍无法重建核心关系。这两个群体就是“隐作”与“有文未作”。

二十三之二、最小实验的完整设计

第二十三节给出了实验的骨架。以下补足使它可以被真正执行的部分。

样本与场地。同一所学校、同一年级、同一位任课教师的四至六个平行班，预计可得有效被试一百二十至一百八十人。不需要跨校，更不需要跨国——本章第二十四节已经说明，跨国比较在这个问题上是异质性太大的设计。

任务。统一的论说文题目一道，统一的证据集合（三至五份材料），统一声明的预期受众。全部被试在同一时段、同一时长内写作。

随机化。中断时点随机分配：在写作总时长的 25%、40%、55%、70% 四个时点各分一组。这一步是本设计的关键——因为本章要测的不是「有没有」，而是「什么时候」，只有随机化的中断时点才能让 t_D 的分布被估计出来，而不是只得到一个二值判断。

中断与重构。到达分配时点时收走全部已写文字与提纲，不允许保留。间隔十五分钟（其间进行与本任务无关的活动），随后要求以不同的表层形式重新组织同一问题：由文字提纲改为口头关系陈述，或改为关系图。全过程录音或留图。

编码。按十六之二的手册提取 G_1 （收走的文字）与 G_2 （重构），两名独立编码者，先算一致性。编码者不知被试属于哪个中断时点组，也不知其最终成绩。

主要结果与预期形状。本章预期看到的不是一条连续上升的曲线，而是一个交叉：一部分被试在文本差异尚不足以区分其组织方向时，已经出现较高的跨条件关系重现；另一部分被试即使已经写出相当完整的文本，撤掉支架后仍无法重建核心关系。前者是隐作，后者是有文未作。若把 SR 对「已写字数」作图，本章预期两者分居同一横坐标区间的高低两端——同样的字数，SR 分为两群。若 SR 与已写字数呈单调关系而不出现分群，本章的核心主张失败。

必须控制的变量。既有作文成绩、阅读能力、题目相关知识、工作记忆容量、最终文章质量、一般写作迁移指标。这六项来自第二十四节所说的最强竞争解释：「这不就是写作能力强吗」。控制之后 SR 若不再对后续无支架重构有独立预测力，作定态不值得单独存在。

功效。若真实效应为两群 SR 均值相差 0.25、组内标准差 0.20，双侧检验 $\alpha=0.05$ 、功效 0.80，每组约需三十人，四组共一百二十人——正好落在一所学校四至六个平行班的规模内。这是本章最想让读者记住的一个数：验证本章不需要大数据，需要的是把中断时点随机化。

伦理。中断与收走文字必须事先告知并取得同意；本设计不得用于评分；重构表现不得进入任何影响学生利害的记录。理由见下一节。

二十四、最强竞争解释不是“没有这个现象”，而是“一般能力”

假设实验真的发现有些学生更能在撤支架后重建关系。最朴素的解释是：“这不就是写作能力强吗？”这是本文必须击败的竞争解释。

所以机制证伪不能只看二者相关。必须控制既有作文成绩、阅读能力、题目知识、工作记忆、最终文章质量、一般写作迁移指标。

如果这些控制以后，中性扰动下的关系自续对后续无支架重构没有独立预测力，那么“作定态”不值得单独存在。

二十五、六条让本文死亡的条件

第一，如果在相同任务中，从来没有出现“文本尚未明显分化而关系自续已经稳定”的参与者，那么“隐作”不存在。

第二，如果所有高自续个体都必须先拥有明显成文，作定始终晚于显形，那么本文最核心的时间错位被推翻。

第三，如果控制既有写作能力、知识水平和最终文章质量后，自续指标不再解释任何后续行为，“作定态”只是一般能力代理。

第四，如果 **writing transfer** 指标能够完整替代同任务中性扰动指标，本文应并入迁移理论。

第五，如果击键和最终文本可以近乎无损地恢复作定状态，那么“隐作不可由通常显形直接读取”的论断失败。

第六，如果改变证据、任务和受众与仅改变表层实现产生完全相同的效应，则本文关于“中性扰动／重编程扰动”的分界不能成立。

这里尤其重要的是第一条，因为它能直接杀死“隐作”这个新对象。

二十五之二、三条反噬，与各自的防误诊

一个概念一旦被听懂，最先发生的往往不是它被检验，而是它被使用。本节列出本章被使用之后最可能产生的三种反噬，并给出各自的防误诊。

反噬一：把 **SR** 做成评分项。最省事的做法是要求每篇作文强制重写两次、计算相似率、达到某个值才算「本人作文」。这会立刻摧毁这个概念——因为学生会开始训练「重构」这个动作本身，而重构一旦成为被训练的技能，它就不再是关系是否稳定的指标，而变成了关系之外的另一项表现。

防误诊：凡是被试事先知道会被要求重构的场合，**SR** 读数一律不成立。这条比「不要用于考核」更硬，因为它不依赖使用者的自觉——它直接规定了读数的成立条件。

反噬二：把「有文未作」当作指控。本章给出了一个很容易被误用的判断：一篇结构完整、语言通顺的文章可能没有发生作文。这个判断落到课堂上，极易变成对学生的道德评价，尤其容易落在那些依靠模板取得进步的学生身上——而他们往往正是最需要模板的那一批。

防误诊：有文未作是对一次任务的判断，不是对一个人的判断，也不是对一次学习的否定。一名学生在这次任务上没有进入作定态，最可能的原因是任务本

身没有制造出需要他自己组织的关系；追责的方向应当是任务设计，不是学生。本章第十七节已经说过一次，此处把它写成硬性的解释规则：任何 **SR** 低的读数，第一解释必须是任务未制造关系，第二解释才是写作者未组织关系。

反噬三：把「隐作」当作不写作的理由。本章主张文章尚未出现而作文可能已经发生。这句话可以被读成「不动笔也没关系」，那是完全的误读——本章第四种次序（ t_D 存在而 t_M 不发生）恰恰是最难被任何人受益的一种，因为那组关系既不能被读者进入，也不能被写作者自己再次取用。

防误诊：本章从未主张 t_M 不重要，只主张 t_M 不等于 t_D 。成文仍然是必要的：它是这组关系唯一可以离开写作者、被他人进入、并在日后被写作者自己重新取用的形式。本卷第一章关于「撤走文本之后还剩什么」的全部讨论，前提是先有那份文本可以被撤走。

二十六、本文也必须接受自己的检验

现在把这套理论作用在本文自身。这篇文章已经有标题、摘要、二十余节正文、概念、表格、引用、证伪条件。但按照它自己的判断，这些东西不能证明这篇文章已经发生了作文。

如果现在删除本文全文，拿走当前章节结构，只保留三家原始材料，再要求作者面向另一类读者口头重建：为什么“存在”“显形”“确认”必须分开；为什么发育中的 **determination** 是关键反例；为什么“隐作”不同于 **pre-writing** 和 **transfer**；为什么作定不等于质量；如果这些关系无法自行重新形成，那么按照本文自己的判据，眼前这篇看起来高度完整的文章，可以被判为“有文未作”。

这不是修辞谦虚，而是本文必须承担的后果。

二十七、还有一个本文解释不了的洞

“作定”究竟发生在一个时刻，还是一个逐渐提高的连续区间？发育生物学自身也提示 **commitment** 可以存在部分状态和连续限制，而不是所有系统都有一道清楚的瞬间边界。

所以“作定态”这个名字可能会制造一种过强的离散感。未来的数据也许会发现：不存在一个瞬间的“进入”，而只有自续性逐步上升。如果如此，“隐作”作为一种存在物可能仍可保留，但“作定时刻”需要降级为一个区间。

第二个洞是：不同文体的核心关系是否可以共用“主张—证据—限定—受众”这套编码？叙事作文、诗性写作、私人表达显然可能需要另一套关系单位。这意味着本文目前对论说性与说明性作文解释力最强，不能未经验证扩张到全部写作类型。

二十八、写死日期的赌注

本文留下一个截至 2028 年 8 月 17 日的经验赌注。

如果本文指出的对象真实存在，那么在至少一项具有足够样本量的写作过程研究中，应能够观察到一组写作者：在最终文本还不足以稳定区分其后续组织方向时，他们在保持题目、证据域和受众不变的中性环境转换之后，已经表现出可重复的核心关系自续；而另一组已经拥有较完整文字的写作者反而没有这种自续。

以下情况不算命中：仅发现早期击键能预测最终成绩；仅发现学生有较好的提纲；仅发现学生记得自己的论点；仅发现优秀学生迁移能力更强；仅发现多稿写作提高成绩；仅发现 AI 使用影响文本质量。

真正的命中必须出现：“显形早晚”与“自续早晚”发生交叉。

如果到该日期高质量研究连续显示文本显形和关系自续始终同步，不存在“隐作”和“有文未作”的稳定交叉，则本文最核心的新对象应该被放弃。

二十九、重新回答：何谓作文

现在可以回到最简单的问题。何谓作文？

传统回答说：作文是一篇学生写出的文章。过程理论说：作文是一个构思、起草、修改和反思的过程。认知理论说：作文是一种复杂的问题解决和知识建构活动。社会修辞传统提醒我们：作文发生在目的、受众、语境和共同体中。

这些回答都没有必要被推翻。但它们没有区分三件可以异步发生的事：一个意义组织什么时候在写作者身上稳定下来；它什么时候显成文章；它什么时候被制度登记成“作文”。

本文因此提出：作文不是文本结构，不是写作步骤的总和，也不是制度确认的提交物。作文首先是一个写作者在特定表达问题中进入“作定态”的发生：当当前语言和非必要支架被撤去，而题目、证据域与受众保持不变时，一组核心意义关系仍能够由这个写作者自行重新生成。

文章，是作文的一种显形；提交，是作文的一种确认；分数，是作文的一种评价。它们都不等于作文本身。

三十、最反常、也最值得保留的一句话

如果本文最终只能留下一句话，我选择：一篇文章可以还没有出现，而作文已经发生；一篇文章也可以已经完成，而作文还没有发生。

第一种叫“隐作”，第二种叫“有文未作”。它们把过去被同一张纸遮住的两个世界分开。

生成式人工智能使这个区别变得急迫，却不是它的理论起点。即使世界上从来没有生成式 AI，一个教师过度完成的提纲、一个唯一正确的范文、一套被反复机械训练的结构，都可能让文章先于作文出现。

反过来，一个孩子坐在没有一个个字的白纸前，也可能已经越过某个过去没有名字的时刻：题目不再只是老师给他的任务；论据不再只是课本给他的材料；若换一种表达、换一张纸、删除现成句子、拿走支架，那些关系仍然会从他这里重新长出来。

在那个时刻，文章还没有发生。作文已经发生。

参考文献与理论边界

1. International Union of Crystallography (IUCr). Dictionary of Crystallography: "Polymorphism". <https://dictionary.iucr.org/Polymorphism>
2. IUCr. Teaching pamphlet on crystal growth, nucleation and metastability. <https://www.iucr.org/education/pamphlets/21/full-text>
3. Gilbert, S. F. Developmental Biology. NCBI Bookshelf: Cell commitment, specification and determination. <https://www.ncbi.nlm.nih.gov/books/NBK26825/>
4. NCBI Bookshelf. Early developmental commitment and differentiation. <https://www.ncbi.nlm.nih.gov/books/NBK9968/>
5. IASB. Conceptual Framework discussions on definition, recognition and derecognition. <https://media.ifrs.org/2013/IASB/April/IASB-Update-April-2013.html>
6. IASB. Conceptual Framework discussions on assets as rights and faithful representation. <https://media.ifrs.org/2014/IASB/May/IASB-Update-May-2014.html>
7. Rohman, D. G. Pre-writing and discovery in composition. <https://publicationsncte.org/content/journals/10.58680/cc196521081>
8. Bereiter, C., & Scardamalia, M. The Psychology of Written Composition. <https://www.routledge.com/The-Psychology-of-Written->

[Composition-1st-Edition/Bereiter-Scardamalia/p/book/9780805800388](https://www.sciencedirect.com/science/article/pii/S1060374311000178)

9. Writing transfer research: adaptive transfer and writing knowledge across contexts.
<https://www.sciencedirect.com/science/article/pii/S1060374311000178>
10. ScholaWrite / writing-process phase research.
<https://www.jowr.org/jowr/article/view/1714>
11. KEYSOFT: keystroke-based writing process research, BEA 2026.
<https://aanthology.org/2026.bea-1.24/>
12. IB Language A assessment instructions, 2026.
https://ibpublishing.ibo.org/exinst/apps/exinst/index.html?chapter=1&doc=EX_instructions_2026_e&part=8
13. AP Seminar exam information.
<https://apcentral.collegeboard.org/courses/ap-seminar/exam>
14. Reprogramming / transdifferentiation review.
<https://pubmed.ncbi.nlm.nih.gov/36067834/>
15. Example-induced design fixation research.
<https://www.sciencedirect.com/science/article/pii/S1871187115300353>

概念与证据边界声明

“作定态”“隐作”“有文未作”均为本文为提出上述判别而暂定的概念名称。本轮同名检索没有给出足以证明其已被占位的直接结果，但检索未命中不能构成原创性证明。正式投稿前仍需进行数据库级别的第二轮故意拓宽、中文学位论文与专著层检索，并人工回溯最接近的 writing transfer、pre-writing、knowledge-transforming、threshold concepts、process logging 与 design fixation 文献。

本文目前对论说性与说明性作文解释力最强。叙事、诗性写作以及高度私人化表达是否适用“主张—证据—限定—受众”关系编码，需要另外建模，不能未经验证外推。

本文不自行宣布达到任何“典范级”认证分。v3.9 要求参与写作与编辑的人不为自己的成品发独立认证分；最终新颖性与有效性必须接受独立近邻审查、经验检验和盲评。

附录 A：v3.9 制造审计（供复核，不属于正文理论表述）

A1. 题型与三源落位

题型：What。模型：三维度型。S=晶体学／材料相科学（当前结构态）；D=发育生物学（发育身份及其形成史）；E=财务会计确认理论（存在与制度确认边界）。最终产物不落回任何一个来源位置。

A2. 三家 1+3+1 抽脊

【晶体学／材料相科学】

- 主题：相同材料组成不能决定对象处于同一种结构状态。
- S1 同一化合物可形成不同晶体结构；S2 不同晶型可具不同性质；S3 实际出现相不必是终极稳定态。
- 自曝事实：稳定存在不等于最终最优。

【发育生物学】

- 主题：发育身份可先于外在表型而改变。
- D1 commitment 可先于 differentiation；D2 determination 需用环境改变后的持续性判别；D3 determination 可在强干预下被重编程。
- 自曝事实：对象已经“决定将成为什么”，当前外形却可能仍看不出来。

【财务会计确认理论】

- 主题：满足存在定义不等于获得账本中的确认资格。
- E1 definition 与 recognition 不同；E2 确认受相关性与忠实表征等约束；E3 对象可能是权利／关系而非物理载体。
- 自曝事实：“没有入账”不能推出“它不存在”。

A3. 27 对碰撞

碰撞	焦点	涌现物
S1×D1	当前结构能否读出已经发生的身份改变	同貌异态
S1×D2	当前同形是否意味着状态相同	结构不可单判
S1×D3	结构身份是否天然终局	可逆身份
S2×D1	性质差异是否必须等表型出现以	隐性差异

碰撞	焦点	涌现物
	后才存在	
S2×D2	真正差异是否需要环境扰动才能显出	扰动显态
S2×D3	已形成性质是否不可重新组织	再定向
S3×D1	尚未终局的状态能否已经稳定存在	隐性稳态
S3×D2	稳定到底是对象属性还是环境关系	条件稳态
S3×D3	强扰动能否使已稳状态换轨	再编程边界
S1×E1	有结构是否等于有资格	有态未认
S1×E2	同一种结构在不同账本是否身份相同	同态异账
S1×E3	对象身份是否等于物理载体	载体脱钩
S2×E1	已有作用是否必然进入制度记录	有效未认
S2×E2	具有性质是否足以产生对象资格	性质非资格
S2×E3	作用究竟落在物还是关系	关系对象
S3×E1	亚稳存在却未入账意味着什么	临在未认
S3×E2	稳定存在是否足以值得制度确认	稳态非确认
S3×E3	物理载体稳定而关系改变时身份落在哪	异步身份
D1×E1	发育身份已变而账本尚无记录	发生先于确认
D1×E2	隐性发生没有确认字段时如何存在	无字段发生
D1×E3	对象身份是否能脱离当前外形	非表面身份
D2×E1	可实验辨识却未被制度确认	可测未认

碰撞	焦点	涌现物
D2×E2	判别结果是否随观察边界改变	边界判定
D2×E3	换环境后仍保持的到底是什么	关系自续
D3×E1	状态已改变而账本仍保留旧认定	账存态灭
D3×E2	重编程发生以后旧确认是否仍忠实	确认过期
D3×E3	身份能否被拆开并重新组合	身份重组

A4. 候选淘汰与最终涌现

- 候选 1“同文异作”：保留为结果判断，但不足以单独成为新对象。
- 候选 2“三时钟错位”：保留为分析工具，不作为本体。
- 候选 3“自续检验”：保留为方法，不作为本体。
- 候选 4“作定态”：作为状态条件。
- 候选 5“隐作”：作为过去账本中没有独立对象资格的新存在物。
- 最终命题：作文不是文本结构，不是写作步骤的总和，也不是制度确认的提交物；作文首先是写作者在特定表达问题中进入“作定态”的一次发生。文章是这种状态的显形，评价与提交只是可能的制度确认。

A5. 最近邻与判决线

近邻	判决线
Pre-writing	若两人完成相同前写作，一人经撤支架仍重建同一系、一人不能，本文判二者不同；pre-writing 本身不必不同。
Writing as learning	没有明显知识增量也可能形成稳定的任务内意义组织。
Knowledge transforming	本文要求特定任务内自续，不要求广泛知识转化。
Writing transfer	本文保持题目、证据域、受众不变，只改变表层环境；若必须换新任务才成立，则属于 transfer。
Threshold concepts	本文可以发生在一篇具体作文内部，不要求学科认发生持久转型。

近邻	判决线
Keystroke phases	若日志本身足以直接稳定读取隐作，则本文关于“形不可直接读出”的主张被削弱。
Design fixation	fixation 是选择空间被锁窄；作定态可由充分开放索后形成，方向不同。
一般作者性/能动性	一个学生拥有全部最终选择权，仍可能在撤掉当前架后无法重建组织关系。
前一篇“预闭分叉”	预闭问哪些路在学生遭遇前被拿走；本篇问学生自的意义关系什么时候开始自续。

A6. 共有前提与推翻材料

三家争的是：作文身份到底应该从当前结构、发生史，还是制度确认里读出来？三家共同假定：一旦某物已经成为一种身份，总有一个当前断面能够单独把它读出来。

推翻材料来自发育生物学自身：**determination** 可以先于明显表型差异；其存在需要通过改变环境后方向是否仍继续来判别。该事实原本回答细胞命运判定问题，却迫使作文理论承认“身份已经发生而当前没有显形差异”的可能性。

A7. 真跑与反向约束

- 公开评价框架审计：AP、IB 等现行体系已经评最终论证、语言、结构，部分还能看过程、答辩或草稿，因此“现代评价只看成品”被判为不成立。
- 在本轮检查的公开评分字段中，未发现独立字段专门测试“文本尚未形成差异时的中性扰动自续”。这只能作为窄范围第三层结论，不能外推为“全球没有”。
- 晶体学反向约束：作定只表示方向具有自续性，不表示正确、优秀、成熟或终极稳定。
- 发育生物学反向约束：作定不意味着绝对不可逆；强干预可以导致重编程。
- 会计确认反向约束：即便隐作真实存在，也不自动意味着学校应该把它变成高风险评分指标。

A8. 机制证伪与日期赌注

- 若相同任务中从来没有出现“文本尚未明显分化而关系自续已经稳定”的参与者，则“隐作”不存在。
- 若所有高自续个体都必须先拥有明显成文，则“作定早于显形”被推翻。

- 若控制既有写作能力、知识水平、最终文章质量和 writing transfer 后，自续指标没有独立解释力，则“作定态”应被并入既有理论。
- 若击键和最终文本可以近乎无损地恢复作定状态，则“隐作不可由通常显形直接读取”的论断失败。
- 截至 2028-08-17 的赌注：高质量过程研究应能观察到“显形早晚”与“自续早晚”的稳定交叉；若始终同步，则放弃隐作。

第四章 补节 文章还没出现，作文已经发生

（一）祖宗已在正文中交手，此处只记编者的处理

本章 v1.1 已在第十二之二节与「学习／表现之别」这一族正面交手，并给出三条分离线（被测的是个体能力还是单次任务的意义组织／本章允许一次不可迁移的发生／本章多出第二只钟）。编者只补两件事。

其一，这条祖宗是本卷全部原稿中最贵的一处缺席，而它的补入方式值得记录。原稿第二十二节曾因回跑公开材料而把自己收窄为「即使过程数据丰富，仍需区分已发生的动作与只能在环境改变后显露的潜在状态」——这个表述过强，因为它正是那一支几十年来的标准立场。v1.1 据此把独有的收窄改写为一句带三个限定的话：在单次任务之内、在成品尚未分化之时、由一次不改变题目与证据域与受众的扰动读出。三个限定缺一不可，而同时满足三条的测量此前的文献里没有。一处缺席的祖宗被补上之后，本章的主张变窄了，可检验性却上升了——这是本卷希望读者注意的一种修改方式。

其二，编者须提请注意，本章第三条分离线（第二只钟的存在）同时是一条自设的死亡条件：若有人在没有成品的任务上成功测出隐作，本章对第二只钟的依赖即被证伪，本章应退回学习／表现的框架。这条写得比原稿其余各条都硬。

（二）与第五章的对切

本章与第五章都处理发生，分工是清楚的，但两篇原稿都没有写出来。

本章问的是能力侧：那组关系是否已经长在这个写作者身上（撤支架后能否自行再生）。第五章问的是归属侧：那几处高杠杆的裁决由谁结算（是学生、教师、范文，还是机器）。

判决性对照：一名学生在教师的密集追问下作出了全部关键裁决——第五章判为写成事件发生（结算权在他），而本章可能判为未进入作定态（撤掉教师，那组关系不能自行再生）。反过来，一名学生完全独立地把一套外部给定的结构内化到能自行重建的程度——本章判为作定态已成立，第五章可能判为写成事件未发生（关键裁决不是他作的）。两个判决可以相反且都对：一个问它长在谁身上，一个问它是谁决定的。

（三）本章欠账

其一，自续率 **SR** 一个数也没有算过。v1.1 补入的清点手册（第十六之二节）与最小实验设计（第二十三之二节）已经把执行门槛降到很低——同一所学校四至六个平行班、约一百二十人、按功效分析每组三十人即可，关键只是把中

断时点随机化。在它跑出来之前，隐作与有文未作是两个定义清楚而未经观察的对象。本章是本卷九章中理论与设计最完备、经验最空的一章。

其二，核心关系只编码四类（主张支持、证据支持、反例限定、受众目的），本章自陈对叙事、诗性与高度私人化的写作不适用。这不是缺陷，是范围声明；但它意味着本卷第二编目前只覆盖论说与说明两类文体。

其三，「作定」是一个时刻还是一个逐渐上升的区间，本章第二十七节自陈未解。v1.1 在第七之二节把这个洞收得更紧了一些——它写明关于作定时刻的一切陈述都应读作对区间上界的陈述，一次成功的扰动重构只证明「不晚于此刻」，一次失败不证明「不存在」。这使命名带来的离散感被削弱，但未消除。

第五章

高杠杆的那几处，是谁结的账

作文是什么？

——从“文本制品”到“写成事件”：晶体学、发育与再生医学、考古与古人类学二阶碰撞下的作文存在论

学科通融专栏·语文教育发生学专题

V3.9 二阶碰撞研究稿·2026年8月

研究前置：三套重新选源与唯一推荐

本轮不沿用此前工作稿中的任何既定组合与既定概念名。先按 V3.9 的未吸收度、理论厚度、位置三分和斜对立要求重新选源，再由碰撞结果决定正文命题。

方案	1 号位	2 号位	3 号位	S 判据	D 判据	E 判据	总三
A 唯 推荐	晶体学 与结构化学	发育与 再生医学	考古与 古人类学	S/粒 子：当前结 构与同一性	D/波： 身份沿重编 程路径可逆	E/场： 语境/出土关 系赋予意义	纯；三家 厚；未吸 风险最低
B 备	免疫学	药理学 与新药发现	制度经 济学与产权	S/粒 子：自我/非 我边界	D/波： 吸收-分布-代 谢-清除轨迹	E/场： 规则、产权 与激励结构	火 强；但“ 份/制度 写作研究 邻较多
C 备	计量学 与精密测量	大气动 力与天气预 报	档案与 数字保存	S/粒 子：当前测 量结果与可 比性	D/波： 状态演化、 预报路径与 误差增长	E/场： 来源原则、 原始次序与 保存关系	可 强；但计 量/档案 育评价、 作档案已 接触

V3.9 选源审计摘要

题型判定：本题主问题为 WHAT。要的不是一条作文训练路径，也不是“为什么要作文”的动力解释，而是一条能把两段看起来一样的文字区分为不同作文身份的辨别维度。正文仍分别回答 WHAT/HOW/WHY，但碰撞发动机按 WHAT 三维度型运行。

未吸收度：写作研究已经系统吸收“过程”“社会/修辞情境”“受众”“体裁”“身份”等理论，因此这些领域不入三家。对 A 组三个具体核心判据进行中英文形式层检索，未发现晶体结构等价判据、细胞重编程/残余表观记忆判据、考古 provenience/context 判据已成为作文理论的固定术语。三家均判“干净”到“干净偏刚进入”。

理论厚度：晶体学有百年以上结构测定、晶体定义与 IUCr 权威规范；发育生物学有经典胚胎学、细胞命运/重编程与成熟反例传统；考古学有过程考古/后过程考古等长期理论争论与现场规范。三家均判“厚”。

位置三分：晶体学把当前显露结构当作主要身份判据；发育与再生医学把形成重编程路径当作关键；考古学把来源位置与关系场当作意义条件。三家分处不同位置，且任何一家单独结算都会被另外两家自己的反例伤害。

命名闸：此前工作稿出现过“作文成相”“意义责任体”“意义自治体”等名字，本轮全部列入自家占位清单，不复用。本轮候选“文本同一性/发生归属/决策入文/作者闭合”也因近邻可替换被淘汰，最终保留“写成事件/写成性”。

摘要

“作文是什么”通常被回答为书面表达、思想与语言的统一、一定文体规范下的文本，或学生综合语文能力的产物。这些回答能够描述成品，却越来越难解释一个AI时代的边界事实：一段文字可以在主题、逻辑、语言、结构等常规指标上显著优于学生写作，却仍不能因此被认定为“这个学生的作文”；反过来，一篇尚不成熟的文本，只要其关键判断确实在写作中被学生选择、改写并反过来改变其后续理解，我们又会直觉地承认“作文已经发生”。问题由此从“好作文具有什么特征”转为“什么事件使一段文字获得作文身份”。

本文依照 SDE 二阶碰撞 V3.9 的研究工序，从“新思想前沿”626 个领域中重新选源，并先排除认知心理学、过程写作、修辞学、体裁理论、社会文化理论、生态写作、创意写作、语言教育等已被语文与写作研究充分吸收的近邻领域。经过未吸收度、理论厚度与位置三分审查，最终选定晶体学与结构化学、发育与再生医学、考古与古人类学。三家分别坚持三种不能相互替代的身份判据：晶体学把对象同一性压在当前可判定的结构与等价类上；发育与再生医学证明当前状态可能由完全不同的重编程历史产生，而且表面多能性不等于功能等价；考古学则把意义压在来源位置、年代序列、文化社会关系与出土关联上，脱离语境的同一器物会丢失关键意义。

27 对支撑碰撞、五候选淘汰、敌意检索与共有前提推翻后，本文提出“写成事件”这一存在论判断：作文不是文本制品，不等于写作过程，也不等于语境中的一次话语行为；作文是写作者在特定任务与读者场中，对高杠杆意义决策进行可追溯裁决，使这些裁决压入文本形成相对稳定显露，同时新显露又回过头改变其下一步选择的闭合事件。本文把这一闭合称为“写成性”。写成性能够把两段逐字相同的文字区分开，也能区分 AI 代产、模板执行、外部强导与真实的人机协同写作。

本文进一步给出“文本达标×写成闭合”的 2×2 辨别格、零情态词判据、五类场景测试、机制证伪条件与一项公开数据实跑。对 Herbold 等 2023 年 90 个题目、270 篇作文、111 名教师、658 次评分的公开研究进行最便宜的复核时发现，ChatGPT 生成文本在七项常规作文评分维度上均取得高于学生文本的均值。该结果不能证明 AI 文本“不是作文”，却直接否定“常规成品质量足以结算学生作文身份”的强版本。由此，AI 时代作文教育的评价对象应从单一终稿扩展为“终稿质量+写成闭合”的双账本。

关键词：作文；写成事件；写成性；作文存在论；生成式人工智能；二阶碰撞；晶体学；发育与再生医学；考古学；写作评价

一、问题被 AI 逼回了本体层：我们究竟把什么叫作“作文”

语文教育长期拥有非常丰富的“好作文”标准，却没有同等精细的“作文身份”标准。教师能够判断一篇文章是否切题、是否有中心、结构是否完整、语言是否准确生动，却往往默认这些评价对象已经天然属于学生。过去这个默认条件很少出问题，因为从选题到落笔大多发生在学生、纸笔与教师之间。生成式 AI 改变的首先不是写作速度，而是把“文本质量”与“学习者发生”彻底拆开：一个学生可

以交出极成熟的文本，而关键的立意、材料、结构、论证与修辞决策没有一次在他身上发生。

这使一个过去被技术问题覆盖的本体问题暴露出来：如果两段文字逐字相同，一段来自学生连续思考、删除与重构，另一段来自一次提示词生成，它们在语文教育中是不是同一种对象？若只看成品，它们完全相同；若只看作者署名，学生都可以署名；若只看使用 AI 与否，又会把查词、资料检索、建议采纳和整篇代产粗暴压成一个二元变量。作文需要一条更深的身份判据。

“作文是什么”因此不是“作文包含哪些成分”，而是寻找一个能把两种看起来一样的文本分开的辨别维度。只要这个维度仍能被“过程写作”“作者声音”“写作身份”“修辞情境”“文本来源”中的任何一个现成概念一比一替换，它就还不是新理论。本文的任务不是发明一个漂亮名字，而是寻找一个在终稿相同、过程相近甚至语境相似时仍能裁决作文身份的新单位。

二、先清空近邻：为什么不能再从语文教育已经熟悉的理论取材

写作研究早已不是“只看文本”的学科。1971 年 Janet Emig 以有声思维 and 个案研究直接进入高中生的 composing process；1975 年 Britton 等分析两千余份学生写作，把功能与受众作为两条分类轴；1980 年代以后，认知过程、社会文化、体裁、修辞情境、共同体与身份研究进一步进入写作研究。NCTE 现行原则明确把写作理解为社会的、修辞的、发生于具体情境与受众关系中的活动。因此，再从认知心理学、社会文化理论、修辞学、体裁研究、生态学、复杂系统或身份理论输入，最多只是把写作研究已经拥有的语言重新排列。

本轮选源刻意避开这些“近而强”的学科。尤其需要排除三个看上去非常适合却已被吸收的方向：第一，“写作是过程”已经拥有半个多世纪的研究史；第二，“写作依赖社会情境、目的与受众”已是专业共同体基本原则；第三，“写作是身份实践”在 Ivanič 等人的研究中已有成熟表达。甚至 AI 时代的“ownership/agency”也正在快速形成自己的实证传统。因此，二阶碰撞若仍以“过程+情境+主体性”收口，得到的只会是一阶综述。

真正可用的源必须处在一个狭窄夹层：理论足够成熟，已经有判据、失败传统与权威手册；但它针对“对象身份如何被认定”的那一条核心答案，尚未被作文理论整体搬走。本轮据此只把远学科当作判据库，而不是隐喻库。

三、三套配方与唯一推荐：为什么 A 组最符合 V3.9

本轮形成三套候选。A 组为晶体学与结构化学×发育与再生医学×考古与古人类学；B 组为免疫学×药理学与新药发现×制度经济学与产权；C 组为计量学与精密测量×大气动力与天气预报×档案与数字保存。三套均来自“新思想前沿”已发布领域，并尽量远离语文、语言、教育与写作研究主场。

A 组最终胜出的原因不在“跨度最大”，而在三个答案真正站在不同位置。晶体学最有力的身份语言是当前结构、衍射模式、等价类与近重复辨识：它倾向于问“现在显露成什么，就把它判为什么”。发育与再生医学则把身份放回时间：Yamanaka 重编程证明成熟细胞身份不是不可逆终点，当前状态的背后存在可逆路径与残余表观记忆。考古学进一步把同一器物放回关系场：出土位置、层位、年代、文化、社会与地理关联不是附加说明，而是意义本身的组成。

这三家不是三种互补视角。若晶体学的判据单独够用，路径与语境就只是附注；若发育路径单独决定身份，当前形态与环境关系只能是路径的结果；若考古语境足以结算意义，那么对象的内部结构和发生历史都被场吸收。三种判据无法通过“各有道理”停战，只能逼出一个新框架。

四、第一家：晶体学的挑战——“现在长成这样”能不能结算身份

晶体学提供的是最强的“当前显露判据”。结构测定的任务，是借助衍射等手段重建有序晶格中散射中心的三维位置，并通过迭代精修使结构模型与实验强度相符。一个晶体结构不是原子名单，而是原子在空间中的有序关系；在现代定义中，晶体结构还被处理为刚体运动下的等价类。这个视角给作文一个非常直接的挑战：如果一篇文字的中心、材料、段落、句法与词语已经形成稳定关系，它作为文本对象已经“站起来”了。

这会支持一种常见的作文观：文章是一个结构化成品。无论它怎样写来，只要当前文本形成足够稳定的整体，就能评价它的题旨、逻辑、语言、风格与完整性。考试阅卷本质上就在这个层面工作：看不到学生半小时前的犹豫，也看不到删掉的五个版本，只能对终稿作当前结构判定。

但晶体学自己同时提供了反噬材料。晶体的定义经历过从外部形态、三维周期到更一般长程有序和衍射特征的历史修订；准晶的出现迫使经典定义扩容。更近的数据库工作则面对“近重复结构”的困境：坐标表示、基矢选择、取整精度与等价变换都可能使表面不同的记录对应同一结构，也可能使极相似记录实际不是同一对象。也就是说，最擅长从当前显露判定对象的学科，自己已经知道“当前表示”并不天然等于“对象身份”。

五、第二家：发育与再生医学的挑战——“现在相同”为什么可能不是同一种存在

发育生物学把问题从空间拉回时间。2006年 Takahashi 与 Yamanaka 证明，少数组合转录因子可以把成熟成纤维细胞重编程为诱导多能干细胞。这一结果击穿了“成熟身份一旦形成就不可逆”的旧默认：同一个当前状态可以通过不同历史抵达，而路径还可能在当前状态中留下残余表观记忆。

“新思想前沿”对这一领域的总结尤其适合作为路径位判据：诱导多能细胞的核心命题被明确标为位置 D，并同时记录其失败边界——当基因组不稳定、残余表观记忆或致瘤因子存在时，多能标志不等于安全功能。换言之，仅看当前的多能标志，会把路径差异压平；而那些被压平的差异又恰恰能够决定后续功能。

类器官进一步把这个问题推得更狠。体外模型可以形成高度相似的器官样结构，但当前沿页面明确提醒：缺少血管、免疫、力学与发育时间时，形态相似会高估器官等价性。因此发育与再生医学自己已经承认：当前“像”，并不结算“是”；发生史很重要，但发生史本身也未必够，环境条件仍会反向改写身份。

六、第三家：考古学的挑战——对象离开关系场，还是原来的对象吗

考古学对“对象”的理解比收藏学更严厉。一件器物被挖出来以后，它的物质形态可以保存得非常完整，但如果出土位置、层位关系和共伴物被破坏，最重要的信息可能已经永久失去。美国国家公园管理局把 context/provenience 称为考

古资源最重要的方面之一，并明确指出年代、历史、文化、社会、地理和与其他遗物的关联共同赋予资源意义；脱离 context，意义与重要性会丢失。

这给作文教育带来一种完全不同的判定。相同的一段文字，放在私人日记、课堂练习、比赛投稿、AI 测试、他人作品改写和正式考试中，并不自动是同一种教育事件。它和什么任务相连、读者是谁、写作者拥有什么资料与工具、是否要求原创、是否经过协同，都参与构成“这段文字在这里是什么”。

但考古学自己也不能把一切压给 context。古 DNA 在过去十多年不断改写单纯依靠器物类型与文化层所建立的迁徙叙事；同一套物质分布可以来自人群移动，也可以来自技术传播，而遗传材料能迫使旧语境重新解释。考古学因此也自曝：场不是终极裁判。新证据会回过头改写原先的语境图。

七、27 对碰撞之后出现的三条暗流

把三家各自的结果层、路径层、条件层支撑做 3×3 跨家碰撞，二十七格中有十九格形成有效焦点，最后抱成三条暗流。

第一条暗流是“同形异生”。当前文本形态高度相同，发生路径却不同，而且路径差异能够预测下一步能力差异。这条暗流直接击穿“终稿等价即作文等价”。AI 一次生成与学生多轮重写可以得到相同终稿，但后续遇到陌生题目时，两者携带的可迁移能力并不相同。

第二条暗流是“同史异位”。即使过程记录相似，关系场改变也会改变对象身份。学生在私人写作里完成一段反思，之后把它原样提交到一项要求现场独立完成的考试中，文本和历史都没变，教育身份却变了。这说明“有真实过程”仍不是充分条件。

第三条暗流是“显露回写”。一个作品不是等写完以后才有力量；已写出的段落会改变后续写法，已经出现的中心会重新选择前文材料，读者反馈会迫使作者重建结构。于是作文不是一个从作者内部搬到纸面的单向过程，而是成品的局部显露不断反过来改变继续生成它的人。三条暗流汇合时，问题已经不再是“文本、过程、语境哪个更重要”，而是三者何时形成闭合。

八、五个候选怎样被淘汰：为什么最后留下“写成事件”

混沌碰撞先产生五个候选：文本同一性、发生归属、决策入文、作者闭合、写成事件。随后进行 1:1 近邻替换与敌意检索。

“文本同一性”很快被杀掉，因为文本语言学、修辞学以及 AI 检测研究都已经处理“表面特征能否辨源”；它最多是一种成品读数。“发生归属”与 provenance、authorship、ownership 高度邻近，一旦把名字换成“写作所有权”，大量论证仍成立。“作者闭合”又会被 Ivanič 的写作身份和近年的 AI ownership 研究吸收。“决策入文”比前几个更强，却仍可能被过程追踪、keystroke logging 和 NIRVANA 这类过程数据集替换：只要记录谁什么时候改了什么，就已然覆盖一大部分。

“写成事件”之所以幸存，是因为它不是“有过程记录”，也不是“作者感觉这是我的”，而要求三种关系同时闭合：第一，终稿里存在高杠杆意义决策，而不是只有表层润色；第二，这些高杠杆决策能够在真实过程里找到写作者的裁决节点

——接受、拒绝、改写或重排；第三，已形成的文本状态至少一次反向改变后续判断或关系场。最后一条尤其关键：它把过程从线性记录升级为回写事件。NIRVANA 可以重放过程，但“有没有发生回写闭合”不是过程日志本身自动给出的结论。

九、共有前提：三家争的其实不是同一个答案，而是“谁有资格结算同一性”

把三家的争论写成同一个句式，可以得到：三家争的是“对象身份最终由哪一份记录来结算”。晶体学把结算权交给当前结构与等价关系；发育与再生医学把结算权交给形成路径和命运历史；考古学把结算权交给来源关系与语境。

它们表面互相冲突，却共享一个未说出的前提：对象身份最终存在一个可以被单独提取的结算面。换成作文语言，就是我们总假定只要找到“真正关键的那一样”——终稿、写作过程、作者身份、语境、原创性、所有权中的某一种——就可以决定这是不是作文。

推翻这个前提的材料不需要第四门学科。发育与再生医学自己已经给出了：iPS 细胞可能通过标志物表现出多能状态，却因残余表观记忆、基因组不稳定或致癌因素而不具备相同功能；类器官可能在形态与细胞类型上高度类似器官，却因血管、免疫、力学与发育时间缺失而不能被判为器官等价。这个领域自己已经证明，同一性不是由“状态”“路径”或“环境”任何一面单独结算，而是要看这些面是否在功能上互相闭合。

这件材料一旦转向作文，原来的争论轴就消失了：不再问终稿、过程、语境谁最重要，而问它们之间有没有形成一种可验证的闭合关系。

十、核心命题：作文是一场“写成事件”

本文据此提出：作文不是文本制品，也不是写作过程，也不是语境中的话语行为，而是一场写成事件。

所谓“写成事件”，是指写作者在具体任务、材料、工具与读者关系中，对决定作品身份的高杠杆意义节点作出可追溯裁决；这些裁决被压入文本，形成可以独立阅读的稳定显露；而已形成的文本又至少一次反过来改变写作者后续的取舍、解释或行动条件。三个环节缺一，写作活动可以存在，文本也可以很优秀，但教育意义上的作文发生程度不同。

这里的“裁决”不是宣称语言完全来自个人。公共语言、教师、同伴、书籍和 AI 都可以提供词句、观点与备选结构。真正不可外包的不是“所有内容都亲自原创”，而是高杠杆节点的最后取舍有没有在写作者身上发生。AI 提出一个论点，学生可以接受；关键在于接受是否只是复制，还是在与材料、目的和已写文本比较后完成裁决。AI 提出一个结尾，学生也可以保留；关键在于他能否在随后写作中让这个结尾反过来约束或改变前文，而不是把它当装饰贴上去。

因此，写成事件不是反 AI 概念，而是比“是否用 AI”更精细的作文身份判据。

十一、“写成性”：把本体命题变成一个可以被查的辨别维度

为了避免“写成事件”只成为哲学口号，本文把其可判别部分称为“写成性”。写成性不评价文章好坏，而评价终稿、裁决路径与任务场是否形成闭合。

最小审计单元不是字词，而是高杠杆决策。本文暂把高杠杆决策分为三类：中心命题/叙述焦点，核心材料取舍，结构性转折或结尾。每一类随机抽取一个节点，检查过程记录里是否能找到写作者本人对该节点作出接受、拒绝、改写或重排的时刻；再检查这一时刻前后文本是否发生可观察差异；最后检查触发它的是材料证据、任务要求、读者反馈、已形成文本还是外部工具建议。

零情态词判据可以写成一句：“**“在终稿的中心判断、核心材料、结构转折三类节点中，各随机抽一项，过程记录里有没有写作者本人对该项作出接受、拒绝、改写或重排的节点；该节点前后文本差异是什么；触发来源是什么？”**”这句话不需要询问“是否真正原创”“是否有意义”“是否充分”，而是能直接去草稿、版本记录、AI 对话或口述回放里查。

如果三个节点都找不到写作者裁决，文本可能是高质量的，但写成性低；如果节点存在却从未改变下一步，作品更像执行既定方案；如果节点存在且已形成文本反过来改变后续决策，写成事件开始闭合。

十二、第二轴：文本达标×写成闭合的 2×2 辨别格

传统作文评价只有一根主轴：成品质量高低。本文强制增加第二轴：写成闭合高低。两轴构成四种完全不同的教育对象。

第一格是“低达标×低闭合”：文本未成熟，关键决策也主要不在写作者身上，属于未完成或代产失败。第二格是“低达标×高闭合”：文章仍有语言与结构缺陷，但关键判断、取舍和回写已经发生。这一格在传统评分中容易被低估，却可能拥有很高的学习价值。第三格是“高达标×高闭合”：产品与学习同时成立，是成熟作文教育最希望看到的状态。第四格是“高达标×低闭合”：这是 AI 时代真正的靶格——所有常规形式审查都可以通过，甚至分数更高，但关键写成闭合没有发生。

第四格具有 V3.9 所要求的三个危险签名：短期指标表现改善；失败不自动产生任何信号；而且会自我加固。学生越依赖代产工具，终稿越漂亮，教师越容易给高分；高分又证明“这样做有效”，进一步增加下一次外包。制度若只看成品，就会奖励使学习发生越来越少的路径。

因此，写成性并不替代文本评分，而是迫使语文教育承认“双账本”：作品账回答“这篇文章好不好”，发生账回答“这篇文章在这个学习者身上写成了多少”。

十三、WHAT：作文是什么——不是所有文本，也不是所有写作活动

回到 WHAT 问题，本文给出一个有边界的定义：“**作文，是写成事件在语言文本中的相对稳定结果；其身份既包含终稿的可独立显露，也包含高杠杆决策在写作者路径中的可追溯裁决，以及文本对后续路径的至少一次回写。**”

这个定义首先排除“只有文本”的强版本。AI 可以独立生成一篇极高质量文章，它是文章，是文本，是语言作品，但在没有学习者裁决闭合的情况下，不自动构成该学习者的作文。其次，它也排除“只要亲手写过就是作文”的弱版本。机械抄写、逐句听写、按模板逐格填充都存在手工过程，却可能没有高杠杆决策。第三，它排除“只要是我表达自己就够了”的主体主义版本：私人情绪可以真实，却若没有形成可独立阅读的文本与基本结构闭合，仍可能只是一段表达而非成熟作文。

定义同时允许人机协作。AI 可以贡献大量句子，甚至整段草稿；只要高杠杆节点的取舍、反例处理、结构重排与最后署名裁决仍可追溯到写作者，并且过程出现了文本对人的回写，人机协作完全可能具有高写成性。反过来，不使用 AI 也不保证高写成性：教师给中心、范文给结构、家长给句子，学生只负责抄写，同样可能是高成品、低写成。

十四、HOW：一篇作文怎样从“有内容”走到“写成”

虽然本研究是 WHAT 型，但本体定义必须能够解释 HOW，否则它只是静态标签。写成事件至少可以观察到七个阶段。

第一，问题被扰动。题目、经历、材料或冲突使旧判断无法直接套用。第二，候选出现。写作者拥有两个以上可能的中心、材料或表达路径。第三，裁决发生。写作者必须接受、拒绝、改写或重排，而不是让外部系统替他一次结算。第四，裁决入文。选择形成可以被读者观察的句段关系。第五，局部显露反向施压：已写出的东西让某些原计划不再成立。第六，重构发生：作者因文本、读者或证据而修改原判断、结构或关系。第七，暂时闭合：文章取得足够稳定的身份，可以脱离作者被阅读，但同时保留未来被新证据重新解释的可能。

传统“审题—立意—选材—谋篇—写作—修改”更像教学步骤；这里的七阶段不是固定顺序，而是发生判据。同一个学生可以从修改开始，也可以先有一句关键句再回头立意。重要的不是走哪条标准路线，而是有没有出现真实裁决和回写。

十五、WHY：为什么 AI 时代仍然要让学生自己作文

如果作文的本质只是生产一段结构完整、语言漂亮的文本，AI 会不断削弱“学生亲自写”的必要性。Herbold 等的实验已经表明，在他们的 90 个议论文题目和教师评分体系中，ChatGPT 生成文本的平均质量在全部七项评分维度上高于学生文本。继续把教学目标定义为“在规定时间内手工生产比 AI 更漂亮的八百字”，会把语文教育推向一场不必要的速度竞赛。

但如果作文是一场写成事件，教育价值就完全改变。学生写作不是为了证明自己比机器更会连接词，而是训练一种把模糊经验变成可承担判断的能力：在多个可能说法之间作出裁决，观察裁决如何进入文本，允许已写出的东西反驳自己，再把这种改变压回新的表达。这样的能力不是“语言包装”，而是判断生成与自我修订。

这也解释了为什么作文与阅读、思考不能简单替代。阅读可以获得别人已经写成的判断；讨论可以在瞬时互动中产生观点；作文却强迫判断留下可回看的痕迹。一个人必须面对自己十分钟前写下的句子，决定保留还是推翻。写作因此把思想变成一个可以反过来要求作者负责的外部对象。

十六、AI 协作的新边界：关键不是“用了多少”，而是“谁结算了高杠杆节点”

AI 政策常用“是否使用”“AI 比例”“生成字数比例”来划线。这些指标便于管理，却与写成事件并不完全重合。一个学生可以让 AI 生成五千字材料，最后只选用两句，并在取舍过程中形成新的中心；另一个学生只向 AI 索取一段三百字结

尾，却把整篇文章最关键的意义结算交了出去。前者 AI 字数更多，写成性未必更低。

近年的实证研究正在逼近这个问题。NIRVANA 记录 77 名学生的逐键写作行为、完整 ChatGPT 对话与复制文本，使写作过程可以被重放；另一项 2026 年的 253 人实验发现，AI 介入不同阶段会不同程度影响写作者的 ownership，起草阶段的影响尤其大。这些研究是本文必须正面承认的应用领域占位者：它们证明“过程和所有权”早已不是空白。

本文与它们的分离线在于：写成性不是心理所有感，也不是过程可重放性。一个学生可以主观上非常觉得“这是我的”，却没有高杠杆裁决；一个过程可以完整录像，却显示所有关键选择都是 AI 一次性提供。写成性要求的是可裁决节点与回写闭合，属于过程记录上的结构判断，而不是自我报告。

十七、真跑一条：公开数据如何击穿“成品质量足以结算作文身份”

V3.9 要求至少有一条证伪条款在成文前用真实公开数据跑过。本文选择最便宜的一条：检验“常规作文成品评分足以作为学生作文身份的充分代理”这一竞争命题。

Herbold 等 2023 年的开放研究使用 90 个议论文题目，每题包含学生作文、ChatGPT-3 和 ChatGPT-4 版本，共 270 篇；111 名中学教师完成 658 次评分。评分包含题目与完整性、逻辑与结构、表达与可理解性、语言掌握、复杂度、词汇与衔接、语言构造七项。公开结果给出一个方向一致的事实：七项均值全部呈“学生最低、ChatGPT-3 居中、ChatGPT-4 最高”，且人类与两个 ChatGPT 版本的差异显著。

本文做的复核很朴素：把七项评分视为七个方向性检查，记录“若成品常规质量越高，就越有资格被判为学生作文”的代理是否与来源身份保持同向。结果是 7/7 全部反向：质量更高的一侧恰恰是零样本中不可能拥有学生本人写作过程的机器生成文本。如果把七个维度粗略当作独立方向，出现 7/7 同向的二项概率是 $1/128 \approx 0.0078$ ；但七项评分显然相关，因此本文不把这个数当正式显著性检验，只把它作为方向性审计。正式研究本身已经用配对统计证明质量差异。

不利结果同样要写出来：这项公开数据只有终稿，没有学生与 AI 逐步协作过程，所以它不能验证“写成性越高，迁移学习越强”。它只能推翻一个更窄的强命题——**终稿常规质量不能单独结算学生作文身份**。因此本文把理论适用范围收窄到“教育意义上的作文身份判定”，不声称写成性能够替代所有文学、出版或法律意义上的作者判定。

十八、机制证伪：什么结果出现时，“写成事件”理论必须退场

最强竞争解释不是“AI 不好”，而是技能自动化与过程写作的联合模型：只要学生掌握足够语言技能，并真实经历计划、起草、修改，所谓“写成性”没有提供新增解释。对此，本文设置机制级证伪。

在控制终稿质量、语言能力、题目难度、总写作时间以及 AI 使用量后，如果“高杠杆裁决闭合度”与下一次陌生题目的独立立意、结构重组和反例处理之间不存在可重复的关联，而且过程写作次数或一般 ownership 自评已经解释全部差异，那么写成性应当被视为冗余概念。更强地说，如果低写成性而高终稿质量的

学生，在后续无 AI、无范文的陌生任务中与高写成性学生表现相同，本文关于“写成事件具有独立教育价值”的核心命题就被推翻。

这一证伪必须使用同质多案例。未来研究不应拿“某国学生对某国学生”这种 $N=2$ 异质比较充数，而应在同一学校、同一课程、相近能力、同一题型中至少取得六个以上具有真实写成性变异的单元，并保留版本记录与 AI 交互日志。

十九、至少两处反向约束：三门远学科怎样反过来削弱本文

第一处约束来自晶体学。本文原本最容易滑向“过程比成品重要”，但晶体学迫使我们保留成品独立性：如果一个学生经历了丰富思考，却最终没有形成一个可以脱离作者解释而被读者进入的稳定文本，就不能因为过程很深而自动把它判为成熟作文。写成事件因此不是“过程主义”，它仍要求当前作品能够站住。

第二处约束来自考古学。本文原本容易把写作者裁决理解成个人内部事件，考古学迫使范围收缩：同一个裁决在不同任务规则与读者关系中意义不同。一个在私人日记中高度闭合的文本，原样提交到要求现场独立完成的考试时，教育身份改变。因此写成性不能脱离任务场审计。

第三处约束来自发育与再生医学。本文原本可能把“有一次回写”当作成熟的充分标志，但类器官与 iPS 边界提醒：表面功能指标可能在关键环境缺失时高估等价性。因此本文不把一次自我修改神圣化；回写必须发生在高杠杆节点，并能在后续新任务中表现出某种迁移，才有资格被解释为能力发生。

二十、课堂后果：作文教学从“替学生把文本做对”转向“保护关键裁决不被外包”

如果采用写成事件框架，教师角色不会变轻，反而更精细。教师首先要区分“产品帮助”和“发生帮助”。直接给一句漂亮句子可以提高产品质量，却可能拿走学生正在形成的高杠杆裁决。真正的发生帮助更像提出能够制造差异的问题：“这两个材料只能留一个，你留哪一个？”“如果删掉结尾，你的文章会失去什么？”“AI 给你的这个解释和你刚才那段经历冲突在哪里？”

第二，初稿阶段不宜把所有反馈压在语言错误上。语言准确仍然重要，但如果教师在意义还没有形成时不断替学生修词，学生会把写作理解为“把不够漂亮的句子交给专家优化”。更值得记录的是关键节点：中心第一次改变在哪里，哪一段被主动删除，哪个反例迫使结构重排。

第三，范文教学应从“拿走终态”改为“还原写成”。不是只圈好词好句，而是问：如果作者没有第三段的那个发现，第四段为什么不会出现？哪一个细节改变了文章方向？范文的价值因此从成品模板变成一份可以逆向推断的发生记录。

二十一、评价后果：终稿分数之外增加一份最小“写成记录”

学校不可能为每一篇作文建立庞大档案。写成性评价应当保持最小化，避免变成新的行政负担。本文建议只保留三类高杠杆节点各一个：中心判断、核心材料、结构转折。每个节点只记录三件事：最初版本是什么；谁在什么情况下触发了改变；学生最后做了什么裁决。

对于纸笔写作，可以保留初稿与一次口头说明；对于数字写作，可以使用版本历史；对于 AI 协作，可以保存相关对话片段而非全部聊天。评价重点不是抓作

弊，而是辨认发生。学生若能说明“AI 给了三个结尾，我拒绝前两个，因为它们把我没解决的矛盾写成了感恩；第三个我也没用，最后保留不确定结尾”，这段记录本身已经显示高杠杆裁决。

因此未来评分可以有两张账：作品账继续评价主题、结构、语言、证据与审美；写成账只评价关键裁决闭合。两账不必机械相加。低作品分高写成性意味着技能还需要训练；高作品分低写成性意味着产品很好但教育证据不足。

二十二、成熟与板结：昨天写成的结构为什么可能成为今天作文的敌人

任何成功结构都会产生一种诱惑：既然它上次有效，就在下一次直接复制。作文模板的问题不在模板本身，而在它是否把新的高杠杆裁决提前结算。学生第一次摸索出“事件—误解—反转—理解”的结构时，那可能是高写成性；第二十次看到题目就自动套入同一结构时，成品越来越稳定，写成事件却可能越来越少。

因此成熟不是“形成固定风格后不再改变”，而是已经拥有稳定能力却仍能让新任务制造真实差异。真正成熟的写作者会在适合时调用旧结构，也会在旧结构无法容纳材料时主动破坏它。板结则表现为文本看起来越来越专业，但任何新材料都只能被旧中心吸收。

这一区分也防止“写成性”自己板结成新模板。若学校把它做成“三个节点打勾表”，学生完全可以伪造三次修改。写成记录只能作为证据入口，不能取代真实差异。

二十三、最近邻划界：它不是什么

第一，它不是过程写作。过程写作关心计划、起草、反馈、修改等阶段；写成事件关心这些阶段中是否发生高杠杆裁决与回写。走完整流程也可能低写成。

第二，它不是作者身份或 writer identity。身份理论解释写作如何呈现和建构自我；写成性不要求作品具有独特“声音”，只要求可追溯裁决闭合。匿名技术说明文也可以高写成。

第三，它不是 ownership。ownership 可以通过自我报告测量，且 AI 贡献越多可能越降低主观所有感；写成性是过程证据上的结构判定，一个人即使强烈感觉“这是我的”，也可能找不到关键裁决节点。

第四，它不是 provenance。来源记录能告诉我们一段文字来自谁、经过哪些工具，却不自动告诉我们文本是否回写了写作者。完整日志可以记录一次彻底外包。

第五，它不是 AI 检测。AI 检测试图从终稿或统计特征猜来源；写成性不猜，它直接要求过程证据，因此允许 AI 深度参与。

第六，它不是“真情实感”。虚构作文同样可以有高写成性；真实经历也可能被模板化加工成低写成性。

二十四、学术界的第二层共有前提：写作研究仍习惯把“对象”预先给定

敌意检索显示，写作研究并不缺过程、社会、身份、受众与技术视角。真正仍然普遍存在的盲区是：多数理论把“writing/composition”当作研究开始之前已经给定的对象，然后讨论它如何发生、受谁影响、如何被评价。AI 时代却迫使我们先问：什么时候一段文字有资格被当作某个学习者的作文？

Emig 研究 composing process 时默认被观察者正在作文；Britton 分类学生写作时默认样本已经是学生写作；NCTE 讨论社会修辞原则时默认对象已经属于 writing；AI ownership 研究则在给定写作任务中测量拥有感。它们都不是错误，因为在各自时代这一本体边界不构成主要问题。

生成式 AI 让“产物属于任务，但发生不属于学习者”成为常态可能性。本研究真正补的一步因此不是再多一个过程变量，而是把“作文身份何时成立”提前为研究对象。

二十五、自反：这篇论文自己算不算一场写成事件

如果写成事件理论不能反过来审查本文，它就不值得被保留。本文的成品显然大量借助了公开资料、网页检索与人工智能辅助组织，但这并不直接决定写成性。真正需要问的是：核心命题“作文是一场写成事件”是否由现成来源直接提供？答案是否定的；晶体学、发育医学、考古学都不会单独推出这个作文定义。

其次，候选是否经历过真实淘汰？本研究主动淘汰了此前工作稿中已经出现过的“作文成相”“意义责任体”“意义自治体”等命名，也淘汰了本轮更接近 ownership、provenance 和 process tracing 的候选。这些淘汰改变了最终命题的落点。

再次，公开数据实跑是否改变理论？是。Herbold 数据能够击穿成品充分论，却不能验证迁移，因此本文把“写成性是作文质量”改为更弱也更可证伪的“写成性是教育意义上的作文身份第二轴”。这个收窄是理论被证据回写的一次实例。

但本文仍有未闭合之处：尚未获得中国母语作文课堂的前瞻性过程数据。因此它目前是一套可检验的存在论模型，而不是已经完成大样本验证的教育定律。

二十六、结论：作文不是“写出来的字”，而是“字怎样在这个人身上写成”

当 AI 能够随时生成结构完整、语言成熟、主题清楚的文本，“作文是什么”终于不能再靠成品外形回答。晶体学教我们认真对待当前结构：没有稳定显露，作品无法独立存在；发育与再生医学让我们看到相同当前状态可能拥有不同发生史，而且发生史会留下后续功能差异；考古学进一步证明，对象脱离来源关系后，意义本身会改变。三家合起来逼出一个新判断：作文身份不是由终稿、过程或语境中的任何一个单独结算，而取决于它们有没有在写作者身上闭合成一次写成事件。

因此，作文可以用一句话重新定义：****作文，是一段语言文本在特定任务场中，经写作者对高杠杆意义节点的可追溯裁决而形成，并在形成过程中反过来改变其后续判断的写成事件。****

这个定义不反对技巧，不贬低成品，不排斥 AI，也不浪漫化“真我”。它只是把一件被终稿遮蔽很久的事情重新放回中心：教育真正要保存的，不只是文本有没有被生产，而是一个人有没有在文本生产中获得新的判断、取舍与修订能力。

未来 AI 越强，作品账越容易变漂亮，写成账就越不能消失。语文教育最值得守住的，不是要求学生永远手工完成所有字句，而是让关键意义节点仍有机会在学生身上发生、被他自己裁决，并留下可以继续生长的回写。

附论一 | 五个场景测试：同一套判据必须给出不同答案

一个新本体如果只能解释 AI 代写一个热点案例，它还不足以成为作文理论。本文把零情态词判据放到五种结构不同的场景中测试。场景一是完全独立的纸笔作文。学生没有外部工具，但这并不自动意味着高写成性。如果他只是把背熟的“成长模板”重新抄一遍，三个高杠杆节点都在任务前被固定，过程没有新的裁决，写成闭合仍然可能偏低。由此可见，“纯人工”不是写成性的同义词。

场景二是教师强指导作文。教师不断提问、提供反馈，却不替学生做最后取舍。学生先写“失败让我坚持”，在教师追问“如果没有坚持，你真正失去的是什么”后改成“我害怕让父亲失望”，随后重排材料。这里外部介入很强，但关键裁决仍发生在学生，且新中心回写前文，写成性可以很高。

场景三是 AI 一次代产。学生只输入题目，得到完整成稿后轻微改词。作品账可以很高，三个高杠杆节点却找不到学生裁决，回写也不存在，属于“高达标×低闭合”的典型靶格。场景四是 AI 协同。学生让 AI 提出三个相反论点，自己根据经历否定两个，保留一个后又因写作过程中出现的新证据改写它；AI 最后只做语言润色。这里 AI 贡献不小，但关键裁决与回写清楚可查，写成性并不因为使用 AI 而被自动判低。

场景五是私人日记转交作业。原文本在写成时可能高度闭合，但当任务规则改变为“课堂现场原创”时，场归属发生变化。文本本身和原发生史不变，当前教育身份却不能直接沿用。这一场景证明写成事件不是文本的永久标签，而是文本、写作者与具体任务关系的一次结算。五个场景给出五种不同答案，说明判据不是“凡 AI 都低、凡手写都高”的道德判断，而能处理过程、工具和制度之间的结构差异。

附论二 | 写成性如何测而不把自己做成新的模板

任何新指标一旦进入学校，都面临一个反噬：当指标成为目标，人们会优化指标外观而不是它原本代表的发生。写成性尤其危险，因为学生完全可以被要求“必须展示三次修改”，然后机械制造三次修改痕迹。因此本文坚持几条测量纪律。

第一，只记录高杠杆节点，不记录修改次数。把“改了多少次”当分数，会鼓励无意义改动；把“写了多少过程文字”当分数，会鼓励过程表演。高杠杆节点之所以重要，是因为它改变作品的身份：中心判断、核心材料、结构转折中的任何一项若改变，后续文本必须跟着调整。

第二，记录必须允许“没有发生”。某一篇熟练说明文，学生可能一开始就知道怎样写，几乎没有中心变化。这不意味着作品失败，只意味着这次任务的学习发生量较小。教学不应强迫每篇作文都演出一次思想革命。写成性是诊断维度，不是价值总分。

第三，测量的最小可信证据是“节点前版本—裁决动作—节点后版本—触发来源”四联。只问学生“是不是你自己决定的”容易受到自我报告偏差；只看版本差异又不知道改变来自谁。四联把主观说明与外部记录结合起来，同时保持低成本。

第四，写成性不宜用于单次高风险惩罚。它更适合形成性评价、教学研究和 AI 使用反思。如果学校拿一个尚不成熟的读数直接决定升学或处分，最省事的行

为就是逼学生伪造过程，这恰好摧毁指标要保护的东西。第五，任何量表都必须保留复核通道：学生应能说明某个被编码为“外部裁决”的节点其实为什么是自己的反思结果，研究者也必须允许编码修订。

附论三 | 一个完整课堂案例：从《我的父亲》到可复核的写成事件

设一名八年级学生拿到题目《我的父亲》。第一次写下的中心是“父爱如山”，材料是父亲雨夜送伞、辅导作业、生日买礼物。文本语句通顺，却几乎所有高杠杆节点都来自熟悉模板。教师没有说“不要套话”，而是要求他找一件“父亲做了、你却不知道该感激还是生气”的事。

学生想到一次比赛。父亲全程坐在看台上，比赛失败后只说“回家”。他最初把这件事归入“严厉”，于是删掉。写到第二段，他又发现这一场景不断回来，便重新加入。第一个裁决节点出现：原来“不够正面”的材料获得进入作品的资格。

第三段中，他写“父亲从来不会安慰人”。读完前文后，他突然想起父亲回家路上绕道买了他爱吃的面，却一句话没说。第二个节点出现：已写出的“不会安慰”与新材料冲突，中心从“严厉中的爱”改成“有些人不会使用我期待的爱的语言”。这个中心不是教师给的，也不是预先存在的；它由文本内部矛盾逼出。

结尾阶段，AI 工具给出一句“父爱如山，沉默却厚重”。学生拒绝，因为这句话把前面刚刚形成的复杂区别又压回旧模板。他最后写：“我还是希望他有一天会说一句‘没关系’，但那晚那碗面让我知道，他也在学怎么做爸爸。”第三个节点是对外部建议的拒绝。

这篇作文未必比 AI 成稿语言更漂亮，却拥有清晰的三联闭合：高杠杆决定可追溯；决定进入终稿；已经形成的文本反过来改写后续中心。若下一次陌生题目中学生能再次独立完成类似裁决，教育发生就获得迁移证据。这个案例也说明，教师最重要的帮助不是替孩子写出“更好的句子”，而是把决定句子命运的权利和困难留在孩子身上。

附论四 | 从理论到实验：一套可执行的母语作文研究计划

第一阶段可以做小样本过程研究。在同一所学校招募 30 至 50 名学生，使用统一作文题和统一时间，随机分配无 AI、AI 只供资料、AI 可参与起草三种条件。所有数字组保存版本历史与对话记录，纸笔组使用阶段拍照和简短刺激回忆。研究者事先定义三类高杠杆节点与编码规则，由两名独立编码者判断是否存在裁决与回写。

第二阶段检验与传统指标的正交性。若写成性只是语言水平的换名，它与终稿总分应高度重合；真正有意义的结果应出现四格分布，尤其要稳定出现“高达标×低闭合”和“低达标×高闭合”。前者证明终稿不足，后者证明写成性不是“换一种方式夸好作文”。

第三阶段做迁移。两周后给所有学生一个完全陌生题目，禁止使用先前范文，并控制语言基础。观察第一篇的写成闭合度能否预测第二篇的独立立意、结构重组和反例处理。若没有增量预测力，就必须承认新维度教育价值有限。

第四阶段做干预。教师不直接提高写成性分数，而只改变课堂动作：延迟给答案、要求关键取舍说明、让 AI 提供冲突选项而非完整成稿。比较这些动作是否提高后续陌生任务的裁决能力。

第五阶段做跨文体复制。叙事、议论、说明和想象作文的高杠杆节点并不相同。若一个定义只能在反思性记叙文里成立，就不应被称为一般作文存在论。未来研究应允许文体差异反过来改写“高杠杆节点”的分类，并把失败的文体公开列入理论边界。

附论五 | 理论边界：哪些写作不该被强行收编

写成事件是针对教育意义上的作文身份提出的，不应无边界扩张。职业公文、团队协作报告、新闻编辑、法律文本与商业文案的作者权本来就可能分散在组织中；它们追求的也未必是个体学习发生。把“每个关键节点都要能追到一个人”硬套这些场景，会造成错误。

文学创作也需要谨慎。成熟作家可能在长期训练后把大量裁决自动化，过程记录里看不到明显停顿或改写，却不意味着没有写成；写成性在专家群体中可能需要从“显式节点”转向“可反事实解释的选择”。这要求未来研究区分初学者和专家。

私人写作的边界也不同。日记可以高度写成，但它未必需要独立面向公共读者；因此本文所说“可独立阅读”的要求在私人写作中应当放宽。作文课程则因为天然包含交付与评价关系，仍需要一定公共可读性。

最后，写成事件不负责判断作品是否道德、事实是否正确或审美是否卓越。一个逻辑错误的观点也可以被作者完整写成；一篇审美平庸的文章也可以高写成。作品质量、事实责任与写成身份是三张不同的账。把它们分开，恰恰是为了避免任何一个新概念重新变成万能解释。

附论六 | 对课程标准、命题与考试的三个直接改造

如果作文课程接受“作品账与写成账分开”的判断，课程标准首先要把“能产出文本”与“能完成高杠杆裁决”区分开。现有表述常把“中心明确、内容具体、条理清楚、语言通顺”作为能力结果，这些仍应保留；但 AI 已经证明，达到这些结果并不需要学习者完整经历相应的判断过程。因此可以增加一个过程能力层：学生能够在不同材料与观点之间作出取舍，能够说明至少一个关键取舍因何改变，并能够在新证据出现时重新组织文本。这里不需要规定必须“独立于一切工具”，而是规定关键裁决不可被完全外包。

命题方式也需要改变。大量作文题其实提前替学生结算了中心，例如“母爱伴我成长”“挫折使我坚强”“合作让我进步”。这样的题并非不能写，但它们把最重要的高杠杆决策——这件事究竟意味着什么——预先拿走。发生型命题应更多保留意义空位，例如“有一件事，你到现在仍不知道该怎么评价吗？”“有没有一个你后来重新看懂的人？”题目不是故意制造晦涩，而是让学生必须做一次真实判断。

考试设计则可以从单终稿证据转向最小双证据。高风险考试不可能保存完整写作过程，但完全可以在正式作文旁加入极短的“关键决策痕迹”：例如先用三分钟写下初始中心和两个备选材料，成文后再用两分钟标注“哪一个关键决定后来改变了，为什么”。这不会自动解决 AI 作弊，却能增加考试对写作能力的可见度。

更重要的是，双证据会改变教师训练。作文讲评不再只展示满分成稿，也展示一篇文章怎样从第一个版本走到第二个版本；优秀学生不再被塑造成“第一次就会写的人”，而是被看见为能够发现冲突、重新裁决和承受修改代价的人。这样，错误与删改不再只是失败痕迹，而成为能力生长的证据。

对低年级学生，写成账可以更简单：只要求说清“我为什么选这件事”“我为什么删掉一句话”。对高中阶段，则可以提高到证据、反例、受众与AI建议的取舍。纵向课程由此不只是增加字数和文体复杂度，而是逐步提高学生能够独立承担的高杠杆决策复杂度。

附论七 | 写死赌注：到什么时候、看到什么，本理论就必须认输

任何能够解释一切的作文理论都没有研究价值。本文把最危险的赌注写死如下：在同一课程、同一语言基础区间和相近题目难度下，收集至少六十名学生两轮以上的作文过程数据，并控制终稿质量、总写作时间、AI使用量与既有语文成绩。如果第一轮的高杠杆裁决闭合度对第二轮陌生题目的独立立意、结构重组、反例处理与自我修改不提供任何稳定增量预测，而一般过程次数、语言水平或ownership自评已经解释这些结果，则“写成性”作为独立教育维度应被放弃或降级。

时间赌注：若在2030年12月31日前，至少三项来自不同研究团队、具有公开编码规则和可复核过程数据的研究都得到上述零增量结果，本文不得再把“写成性”表述为作文身份的独立第二轴，只能把它降为过程写作或作者ownership的一种操作化变体。

什么不算命中：单纯发现写成性与作文分数相关，不算支持；单纯发现AI使用者分数下降，不算支持；只比较两个国家、两个班级且无法控制混淆变量，不算证伪；只有自我报告“我觉得这是我的作品”而没有版本或过程证据，也不算对本理论的直接检验。

这项赌注的意义不是表现谦虚，而是把理论的生命交给未来数据。如果最终失败，我们至少会知道：作文教育真正的迁移能力并不需要本文提出的闭合机制，那么资源就应当回到更简单的技能、过程或所有权模型上。

附论八 | 为什么“写成事件”不是反技术，而是一种更严格的人机共写理论

把作文定义为写成事件，最容易遭到的误解是“它仍然在守护纯人类写作”。其实恰好相反。纯手写、闭卷、无工具都不是本理论的必要条件。人类写作从来借助字典、老师、编辑、同伴、范文和检索系统；生成式AI只是把外部参与提高到过去没有的强度。理论真正要保护的并不是工具边界，而是决策边界。

在人机共写中，可以把每一个关键节点理解为“提案—裁决—入文—回写”四步。AI负责提案完全没有问题，甚至可以一次提出十个互相冲突的方案；学生的任务不是证明自己能比AI想得更多，而是证明这些方案进入作品之前经过了比较与裁决。若AI提供的方案直接进入文本，没有产生学生可追溯的接受、拒绝或改写节点，那一部分更接近外部代产；若学生根据自身材料和已经形成的文章结构重新组织建议，外部资源就进入了写成事件。

这一区分也允许出现真正的人机共同创新。某些新中心可能既不是学生预先拥有的，也不是AI第一次给出的，而是在多轮对话中由双方相互否定后出现。此

时要求“思想必须在使用 AI 前已经属于学生”反而错误。写成性的判据不是追问谁首先发明一个句子，而是最终高杠杆判断有没有被学生裁决、能否说明为何进入作品、又怎样改变后续文本。

因此，人机协作最值得训练的技能可能是拒绝。一个学生能够说“这段很漂亮，但它把我不确定的事情写得太确定，所以不用”，其写作能力可能比独立生成一句华丽句子更重要。拒绝不是消极动作，它显示写作者已经形成足以抵抗外部高质量文本的作品边界。

同样，AI 也可以成为回写放大器。学生完成初稿后，让 AI 扮演不同读者提出三种误解；如果其中一个误解迫使他发现文本并没有表达自己以为已经表达的关系，他重新修改中心与材料，这个 AI 介入不是减少写成，而是制造了新的差异。技术在这里不是替代发生，而是增加发生条件。

所以未来作文教育真正需要的不是“AI 禁区地图”，而是“关键裁决地图”：哪些节点可以完全外包，哪些节点必须保留学生最后裁决，哪些节点适合让 AI 制造反例，哪些节点必须由真实读者反馈。不同年龄、不同文体可以有不同配置，但底层原则不变——工具可以越来越强，写作者对作品身份的最后裁决不能在不知不觉中消失。

附论九 | 一个最小教学原则：把最后决定留给学生

如果必须把整篇论文压成一个课堂动作，那就是：教师和 AI 都可以把选择空间做大，但不要替学生完成最后的高杠杆决定。教师可以给材料、给反例、给句式、指出逻辑断裂，也可以明确告诉学生某个事实错误；AI 可以生成候选、重写段落、模拟读者。可是在中心到底落在哪里、哪一个材料值得留下、冲突是否要被解决、结尾应该确定还是保留悬而未决，这些节点只要教学目标仍是培养写作能力，就需要学生自己完成一次可说明的裁决。

这不是把教师退回旁观者。恰恰相反，优秀教师的专业性从“我能把你的作文改好”转向“我知道哪一个困难必须替你清除，哪一个困难必须留给你经历”。前者保护写作安全，后者保护能力发生。作文课真正稀缺的不是标准答案，而是一段没有被过早替代的决定时间。

最后还要强调，写成事件不是给学生增加一套新的证明负担。低年龄学生完全可以用一句口头说明留下裁决痕迹；成熟写作者也不需要把每次自动化选择都拆开报告。它首先是研究者和教师用来避免误判的透镜：当终稿越来越容易被技术优化时，我们至少还能分清“文本已经做好”和“学习者已经学会”不是同一件事。只有守住这一区分，作文教育才不会在 AI 越强时越用高分奖励能力外包。

参考文献

1. SDE Universes. 新思想前沿：626 个领域，近二十年两次思想转向. 2026.
2. International Union of Crystallography. Online Dictionary of Crystallography: Structure determination.
3. Widdowson D, Kurlin V, et al. The importance of definitions in crystallography. IUCrJ/Acta-related publication, 2024.

4. Brock CP. Change to the definition of “crystal” in the IUCr Online Dictionary of Crystallography. IUCr Newsletter, 2021.
5. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 2006;126:663-676.
6. Sato T, et al. Single Lgr5 Stem Cells Build Crypt-Villus Structures in Vitro without a Mesenchymal Niche. *Nature*. 2009;459:262-265.
7. SDE Universes. 发育与再生医学·新思想前沿. 2026.
8. U.S. National Park Service. What Is Archeological Context? 2023.
9. SDE Universes. 考古与古人类学·新思想前沿. 2026.
10. Emig J. The Composing Processes of Twelfth Graders. NCTE Research Report No.13. 1971.
11. Britton J, et al. The Development of Writing Abilities (11-18). Macmillan/NCTE. 1975.
12. Flower L, Hayes JR. A Cognitive Process Theory of Writing. *College Composition and Communication*. 1981;32(4):365-387.
13. Ivanič R. Writing and Identity: The Discoursal Construction of Identity in Academic Writing. John Benjamins. 1998.
14. National Council of Teachers of English. Understanding and Teaching Writing: Guiding Principles. 2016/2018.
15. Herbold S, Hautli-Janisz A, Heuer U, et al. A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*. 2023;13:18617.
16. Jelson A, et al. NIRVANA: A Comprehensive Dataset for Reproducing How Students Use Generative AI for Essay Writing. *arXiv*, 2026.
17. Gero KI, Long T, Schnitzler C, Dhillon P. From Planning to Revision: How AI Writing Support at Different Stages Alters Ownership. *arXiv*, 2026.
18. Yeadon W, et al. Evaluating AI and Human Authorship Quality in Academic Writing through Physics Essays. *arXiv*, 2024.

附录 A | 27 对混沌碰撞记录（后台审计摘要）

以下表格记录九条支撑观点的跨家 3×3 碰撞。它不是正文论证的替代，而是用于保留“新判断如何发生”的生产链。27 格全部围绕同一对象“作文身份判定”，无对象漂移。

对	焦点	撞击	涌现物
A1×B1	当前结构/衍射关系可用于判定晶体结构与等价类 ⇔ 成熟细胞可被转录因子重编程到多能状态	当前形态可相似，形成路径却能被重写	同形异生
A1×B2	当前结构/衍射关系可用于判定晶体结构与等价类 ⇔ 残余表观记忆使当前标志相似不等于路径等价	外观/标志相同不能保证功能同一	标志失真
A1×B3	当前结构/衍射关系可用于判定晶体结构与等价类 ⇔ 类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价	局部结构足够像，整体等价仍可失败	结构越界
A2×B1	结构测定依赖从实验强度到坐标模型的迭代精修 ⇔ 成熟细胞可被转录因子重编程到多能状态	结构由数据重构，身份也可由干预重构	重构双向
A2×B2	结构测定依赖从实验强度到坐标模型的迭代精修 ⇔ 残余表观记忆使当前标志相似不等于路径等价	模型精修与生物重编程都留下不可忽略残差	历史残差
A2×B3	结构测定依赖从实验强度到坐标模型的迭代精修 ⇔ 类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价	局部拟合通过不等于整体功能等价	局部达标
A3×B1	表示方式、基矢与近重复会制造同一性歧义 ⇔ 成熟细胞可被转录因子重编程到多能状态	表征改变与身份重编程共同破坏静态同一	表示可逆
A3×B2	表示方式、基矢与近重复会制造同一性歧义 ⇔ 残余表观记忆使当前标志相似不等于路径等价	近似状态因历史不同走向不同后果	近重复分叉
A3×B3	表示方式、基矢与近重复会制造同一性歧义 ⇔ 类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价	相似到什么程度才算同一 —无法由形态单轴结算	等价阈值
A1×C1	当前结构/衍射关系可用于判定晶体结构与等价类	同一物态因来源关系不	同物异义

	⇌ 器物脱离出土 context/provenience 会丢失意义	同而意义不同	
A1×C2	当前结构/衍射关系可用于判定晶体结构与等价类 ⇌ 年代层位与后沉积事件决定证据如何被解释	只看现态会抹掉年代序列与扰动史	当前吞史
A1×C3	当前结构/衍射关系可用于判定晶体结构与等价类 ⇌ 文化、社会、地理与共伴关系共同构成考古意义	内部结构不能包含所有社会文化关联	结构缺场
A2×C1	结构测定依赖从实验强度到坐标模型的迭代精修 ⇌ 器物脱离出土 context/provenience 会丢失意义	重建对象还需要知道材料从哪里来	重构需址
A2×C2	结构测定依赖从实验强度到坐标模型的迭代精修 ⇌ 年代层位与后沉积事件决定证据如何被解释	对象内部生成史与外部沉积史同时塑形	双重序列
A2×C3	结构测定依赖从实验强度到坐标模型的迭代精修 ⇌ 文化、社会、地理与共伴关系共同构成考古意义	同一模型在不同关系场承担不同意义	模型带场
A3×C1	表示方式、基矢与近重复会制造同一性歧义 ⇌ 器物脱离出土 context/provenience 会丢失意义	表面近重复需要来源关系辅助裁定	来源判重
A3×C2	表示方式、基矢与近重复会制造同一性歧义 ⇌ 年代层位与后沉积事件决定证据如何被解释	微小表示差和后沉积扰动都可能改写同一性	扰动同一
A3×C3	表示方式、基矢与近重复会制造同一性歧义 ⇌ 文化、社会、地理与共伴关系共同构成考古意义	何为同一类取决于关系场中的判定任务	等价依场
B1×C1	成熟细胞可被转录因子重编程到多能状态 ⇌ 器物脱离出土 context/provenience 会丢失意义	身份既可被重编，也可因来源关系改变	身份可迁
B1×C2	成熟细胞可被转录因子重编程到多能状态 ⇌ 年代层位与后沉积事件决定证据如何被解释	发育命运与年代层位都拒绝无时间对象	命运带层
B1×C3	成熟细胞可被转录因子重编程到多能状态 ⇌ 文化、	身份改变不是孤立内部程序	自我非孤

	社会、地理与共伴关系共同构成考古意义		
B2×C1	残余表观记忆使当前标志相似不等于路径等价 ⇌ 器物脱离出土 context/provenience 会丢失意义	残余记忆与 provenience 都把过去留在现在	记忆即来源
B2×C2	残余表观记忆使当前标志相似不等于路径等价 ⇌ 年代层位与后沉积事件决定证据如何被解释	内部记忆史与外部层位史共同约束解释	双史判定
B2×C3	残余表观记忆使当前标志相似不等于路径等价 ⇌ 文化、社会、地理与共伴关系共同构成考古意义	历史残差只有放进环境关系才知道其后果	记忆入场
B3×C1	类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价 ⇌ 器物脱离出土 context/provenience 会丢失意义	形态相似与器物完整都无法替代语境	相似失境
B3×C2	类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价 ⇌ 年代层位与后沉积事件决定证据如何被解释	发育时间与年代序列缺失都会制造伪等价	时间缺口
B3×C3	类器官形态相似在缺血管/免疫/力学/时间时不等于器官等价 ⇌ 文化、社会、地理与共伴关系共同构成考古意义	整体身份需结构、历史、关系共同闭合	功能场闭合

附录 B | 敌意检索与候选淘汰

方向	占位者	它已经说了什么	本文的分离线/处置
经典层<1980	Janet Emig, 1971	过程写作早已把“写”从终稿移向 composing process	本文若只说“作文写作过程”即被完全吸收
经典层<1980	James Britton et al., 1975	功能与受众两维早已进入学生写作分类	本文若只说“作文由谁的和读者决定”即被吸收
应用领域内部	Flower & Hayes, 1981	认知过程模型把计划、转译、审阅写成动态系统	本文不能把“认知过程”换名为发生
应用领域内部	NCTE Guiding Principles	写作被明确视为社会的、修辞的、情境化的	本文不能以“情境重要”宣称创新
专著/身份层	Roz Ivanič, 1998	Writing and Identity 把写作界为身份实践	本文不能以“作者身份/声音”收口
方法学占位	NIRVANA, 2026	逐键记录、完整 AI 对话可重放写作过程	本文不能把“有过和志”当新范式
应用领域内部	Gero et al., 2026	AI 介入阶段显著改变	本文不能把“主观所

		写作者 ownership	感”当核心
成品检测/反例	Herbold et al., 2023	AI 终稿在七项教师评分维度上均值更高	直接击穿“成品越如像学生作文”的代理
成品检测/反例	Yeadon et al., 2024	300 篇物理短文中盲评质量无显著差异，作者识别近随机	进一步说明成品判以结算来源
自家已发/工作稿	“作文成相/意义责任体/意义自治体”	此前工作稿已占用相近命名空间	本轮不复用，避免线自撞

附录 C | 真跑一条与理论回写

公开数据：Herbold 等，Scientific Reports 2023。90 个议论文题目；每题学生、ChatGPT-3、ChatGPT-4 各一篇，共 270 篇;111 名中学教师完成 658 次评分。七项评分为：题目与完整性、逻辑与结构、表达与可理解性、语言掌握、复杂度、词汇与衔接、语言构造。公开结果：七项均值均为学生最低、ChatGPT-3 居中、ChatGPT-4 最高；人类与两个 GPT 组的差异显著。

最便宜的方向性复核：竞争命题“常规终稿质量是学生作文身份的充分代理”若成立，至少不应出现非学生生成文本在全部常规质量维度上一致占优而身份仍显然不同的结构性分离。实际为 7/7 维度一致反向。若粗略假定七维独立，7/7 同向的二项概率为 1/128=0.0078125；但七维相关，因此不把此数作为正式显著性结果，只作方向性审计。正式论文已有配对统计与效应量。

不利结果与回写：该数据没有学生-AI 协同过程，不能证明“写成性”预测迁移。因此本理论从“作文质量的新指标”收窄为“教育意义上的作文身份第二轴”；作品质量仍由原有作文评价体系负责。

附录 D | 中英文检索词表（闸零与敌意检索留档）

检索目的	中文检索词	英文检索词
晶体学未吸收	作文教学 晶体学；写作理论 晶体结构；语文作文 衍射 等价类	"writing st crystallography pedagogy" crys writing theory c equivalence
发育医学未吸收	作文教学 诱导多能干细胞；写作理论 细胞重编程；作文 表观遗传记忆	"compositi induced pluripo writing pedago reprogramming epigenetic mem
考古学未吸收	作文教学 考古语境；写作理论 provenience；作文 出土关系	"writing st archaeological "composition p provenience; es context
形式层占位	同样成品不同过程是不是同一种学习产物；相同文本不同生成路径；写作过程能否决定作品归属	same final composing pro writing provena ownership; essa

		trace
AI 近邻	作文 AI ownership agency ; AI 写作 过程记录 归属	AI writing drafting revisio process trace C authorship prov

第五章 补节 高杠杆的那几处，是谁结的账

(一) 最近的祖宗与分离线

本章的核心主张是：作文的作者性不应按谁敲了键盘判断，而应按谁承担了那几处高杠杆节点的裁决判断。

这条主张有一位现行制度祖宗，见第一章补节（一）所列的 **ICMJE** 四条作者资格与 **CRedit** 贡献者角色分类。此处只补本章特有的那条分离线。

ICMJE 与本章共享同一条否定：字数不定作者。但两者的正面判据不同。ICMJE 的四条是合取式的资格清单，其中第三条（核准最终版本）与第四条（对全文完整性负责）是纯粹的事后承诺，与写作过程中发生了什么无关；一个人可以在成稿之后才第一次读到全文，只要他核准并愿意负责，这两条即告满足。本章明确拒绝这条路：本章第十六节主张，关键在于那几处改变了方向的裁决在发生的时候由谁完成，事后的认领不能补票。

判决性对照：一位在定稿会上提出关键反例、迫使全文重写的合作者，按 ICMJE 是作者，按本章也是写成事件的参与者——两者一致。但一位全程未参与、最后核准并署名负责的资深作者，ICMJE 判为满足第三、四条（若他同时满足第一、二条则为作者），本章判为零结算权。反过来，一名在过程中作出了全部高杠杆裁决、却拒绝为最终版本负责的学生，本章判为写成事件已发生，ICMJE 判为不具作者资格。**ICMJE** 分配的是功劳与责任，本章判定的是发生。这条差别对教育场景尤其重要：学校要处理的往往不是署名争议，而是「这次学习到底有没有发生」。

(二) 吸收：跨题复现作为必要条件

本卷选目时，另有一篇同题稿《判据析出体》未入书，其承重内容在此并入本章，理由如下。

那篇主张：作文发生的最低可检验标准是，文本完成之后，写作者能在表面材料不同的新任务中复现一个写作前不能稳定使用的区别，且这一复现不能由原题措辞、教师模板或固定体裁句式直接复制得到。这与本章的结算权是互补的两条件——结算权判这一次是不是他决定的，跨题复现判这个决定是否留下了可带走的东西。编者建议本章在第十八节的机制证伪中增列一条：若高结算权与跨题复现长期不相关，则两者之一为冗余。

该篇另有一段核验值得整段保留：它先查 ASAP 2.0（约两万四千篇美国中学生议论文），发现其数据对象是「作文文本—评分—人口／地区样本」，不含「新区别在第几稿产生、是否跨任务保留」这类字段；随后又查剑桥的 Write & Improve 语料（约两万三千篇、七百六十六名学习者、五千零五十个用户—题目集合，保存同题的初稿、中间稿与终稿），因此当场撤回了原本更强的主张——不能说「写作语料只有成品没有过程」，只能说「当代语料已能保存丰富过程，

但『可迁移的新判断边界』仍不是常规字段」。这一处收窄使主张更弱也更可检验，是本卷倡导的做法，应作为本章公开数据核验的第二条并入。

（三）与第四章的对切

见第四章补节（二）。

（四）本章欠账

本章的真跑是对他人已发表数据的方向性复核（九十一个题目、二百七十篇文本、一百一十一名教师、六百五十八次评分），得到七项评分全部反向的结果，且明确拒绝把 1/128 当作显著性检验——这份克制应予肯定。但该数据只有终稿、没有过程，因此它能推翻的只是一个很窄的命题（终稿常规质量不能单独结算身份），不能验证本章的正命题。本章的正命题至今未经任何数据检验。

编三

账外三格：无形·无据·无名

编三 账外三格：无形·无据·无名

前两编还在处理「读得出但需要动一下」的东西。这一编处理根本不在账上的东西。

三章处理三种不同的缺席，读者可以按一张两轴表把它们分开：这件事有没有真的出现过一次 × 账上有没有为它设一栏。

第六章处理「从未出现，且不设字段」——从未获准成形的那些形式。它们没有实体，不留痕迹，却划定了此后一切可能形式的边界。这一章最难的地方在于它必须把这个对象与「逻辑上的补集」分开：不是「凡是没写的都算」，而是要求「曾经可及、已不可及、且这个转变没有留下记录」。

第七章处理同一格的动作面：那条路是谁在什么时候关掉的。它必须指认一个关闭者与一个时点，否则命题为空。它与第六章是嵌套关系而非同义——一个没有任何外部关闭者的写作者仍然有未生界，而一次被清楚指认的预闭未必落在未生界里。两章各自的分离线与并入条件写在第七章末尾。

第八章处理「已经出现，但不设字段」——相邻两个有出处的地方之间，那一段必须当场自造、两端依据都不覆盖的文字。它不是空白，也不是错误；成品上完全看不出来，因为它读起来和别处一样通顺。它的代价由写作者独自承担，而两端都不记这笔账。这一章还给出一个算术型的反直觉推论：加密提纲、增加分点、使用更完整的模板，必然增加这类段落的数目。

三章合起来给出一条谱系：无形、无据、无名。第九章处理最后一格。

第六章

在第一笔之前

在第一笔之前

——一篇作文的身份，由它从未获准成形的那些形式划定，而这道边界不留记录

摘要 关于“作文是什么”，通行的回答有三种，各自完整：看写出来的那些东西（成品即身份）；看它是怎么被写出来的（写作过程决定成品）；看它在什么场合被谁判读（意义由读者与场合给出）。三种回答互不相容，却共同假定了一件事——凡是要被判定的东西，必先已经成形并且在场。本文提出，这个假定不成立，而恰恰是它不成立的那一部分决定了一篇作文是什么。本文把那一部分命名为未生界：一个写作者在落笔之前就被排除掉、因而从不进入任何记录的那些形式所划出的边界。它不是被删掉的草稿（那有痕迹），也不是没写完的段落（那有位置），而是从未获准成形、因而在任何账目上都不占一行的那一类。本文给出一张两轴四格的辨别表，指出其中“从未出现且账上无字段”的一格具有三个签名：短期指标表现为改善、失效不产生任何信号、越训练有素越严重。本文给出一句不含程度词的问话与一个读数——末次首现：把一个人的连续产出排开，他最后一次写出一种此前从未写过的形式，发生在第几件。本文用英文维基百科五千一百一十五个链接目标做了一次预注册检验，主假设失败：被明确指名的缺口并不会被系统抹掉，它只是极少（占链接目标的百分之七点二）且极惰（超过一半十年原地不动）。这一失败关掉了本文原打算走的那条路，并把承重主张的适用范围收窄到有连续产出序列可查的场合。

关键词 未生界；末次首现；成形前的排除；缺席的字段；写作的形式库

Title Before the First Stroke: A composition is identified by the boundary of forms never permitted to take shape — and that boundary leaves no record

Abstract Three standard answers to “what is a composition” — the finished text, the writing process, the reading context — are mutually incompatible yet share one assumption: whatever is to be judged must already have taken shape and be present. This paper argues that the assumption fails, and that what fails in it is decisive. It names the failing remainder the *unborn boundary*: the set of forms excluded before the first stroke, which therefore enter no record at all. A two-axis table separates this cell from deleted drafts and from unfinished passages. A modality-free diagnostic and a single reading — *last first-occurrence* — are proposed. A preregistered test on 5,115 link targets in the English Wikipedia refutes the paper’s own main hypothesis: named gaps are not erased; they are merely rare and inert.

Keywords unborn boundary; last first-occurrence; pre-formative exclusion; fields for absence; the form repertoire of a writer

引言

一个中学生把作文交上来，老师在旁边批：这里可以再展开一点。学生照做，第二稿更长，评分更高。这件事每天在发生，没有人觉得有问题。

问题在另一处。这个学生在动笔之前，脑子里排开过几种开头？他自己也说不清。可以确定的只有一件事：他排除掉的那些，没有任何地方记着。不是记得不全，是根本没有那个地方。作文本上有他写下的字，有他划掉的字，有老师的批注，有分数；唯独没有一栏是留给“他本来可以那样写，而他连试都没试”的。

这不是一句感慨。它是一个可以追问的技术问题：如果一篇作文的身份，恰恰由那些从未出现的形式决定，那么现行的一切读法都读不到它。

关于“作文是什么”，眼下有三种回答，各自都完整，也各自都有几十年的硬材料撑着。

第一种说：看写出来的那个东西就够了。同一个意思，用两种写法写出来，就是两样东西——不是两种表达，是两样东西。这一种回答在化学里有它最硬的版本：同一个分子式，结成不同的晶体，就是不同的物质，溶解度不同、疗效不同、专利地位不同。一九九八年，一种抗艾滋病药物的原有晶型在生产线上突然不再出现，换成了一种更稳定、也更难溶解的形；同一个分子，药却做不出来了。（Bauer et al. 2001 ;Bučar, Lancaster & Bernstein 2015）形就是身份。

第二种说：不对，形读不出身份，要看它是靠什么被维持成这样的。这一种在地貌学里有它最硬的版本。上一个世纪之交，戴维斯（W. M. Davis）把地形讲成一部有年龄的历史：幼年、壮年、老年，一路削平。一九六〇年，哈克（J. T. Hack）在《美国科学杂志》上给出相反的读法——阿巴拉契亚山脉的形状之所以保持不变，不是因为它还年轻，是因为抬升与剥蚀恰好等速；形是当前过程的解，不是历史的沉积。同一个形，可以由完全不同的路径维持；因此你从形里读不出它的来历。

第三种说：两边都在物本身上找，而身份根本不在物上。这一种在免疫学里有它最硬的版本。一个分子是不是“异己”，不写在这个分子身上——它在这个人身上引发排斥，在那个人身上被安然接受。一九九四年，马青格（P. Matzinger）在《免疫学年评》上把这条推到底：触发免疫应答的不是“外来”，是组织受损时发出的危险信号。判定席位不在被判定者那里。

三种回答不能同时成立。第一种要守住“形即身份”，第二种当场指出同一形可由不同路径维持，第三种当场指出身份由场决定。它们之间没有裁判可当，因为它们根本不在同一个语域里作答。

但是——把这三种回答摆在一起看久了，会看出一件它们从不争的事。

它们争的是“一篇东西的身份该由哪一样来裁”，而三种回答都默认：那样要被裁的东西，已经成形，并且在场。

念给三方听，三方都会说：这还用说吗？

本文说的正是这一句“还用说吗”。它不成立，而且它不成立的那一处，恰好是决定一篇作文是什么的地方。

推翻它的材料不必从外面搬。它在第三种回答自己家里，是一条被引用了六十年的常识。

一·一件小事：三次被退回的开头

先把现象摆清楚，再谈机制。

一位教了二十年作文的老师，手上有一叠同一个班的作业。她能一眼看出哪几篇“套路化”了——句子通顺，结构完整，例子恰当，分数不低，但读起来像同一台机器出的。她说不清哪里不对，只能说“没有味道”。

这句“没有味道”通常被当成审美判断，因而不可追问。本文认为它是可追问的，只是要问的地方不在纸上。

把这几篇的成品逐字比对，找不到共同的错误——它们没有错误。把这几个学生的写作过程录下来（击键记录、草稿留存、修改次数），也找不到显著差别——他们同样改了三四遍。把评分标准拿来看，四项全优。

能查到的一切都正常，而结果确实退化了。这正是本文要处理的那一格的第一个签名。

真正发生的事，只有在把时间往前推到落笔之前才看得见。这几个学生在动笔前，已经把一批写法排除了：太口语的排除、第一人称的排除、没有例证的判断排除、结尾不升华的排除。排除得非常快，快到他们自己不认为那是“排除”——那更像是“不会那样想”。

到这一步，请注意一个容易滑过去的区别。他们不是“想了又否决”。想了又否决，那叫犹豫，犹豫会留下痕迹：停顿、草稿上的半句、事后能回忆起来的“我本来想”。他们是没有想到，而“没有想到”不是遗漏——如果是遗漏，提醒一下就能补上；实测里提醒过，补不上，补出来的是对提醒的应答，不是那个被排除掉的形式。

一个更冷的对照：把同一批学生的作文按时间排开，从初一到初三，你会看到一件事——新的句式一年比一年少。不是句子变差了，是再也没有出现过没出现过的东西。到某一篇之后，此后每一篇都在已有的形式里重排。

这一篇之后，形式库封闭了。

这里要挡住一个立刻会出现的、也很有道理的反驳：这不就是长大了、写得熟练了吗？熟练当然会带来重复，一个人写多了，自然会有他顺手的写法。

这个反驳可以被查。熟练带来的重复有一个签名：它是可逆的。给一个熟练的人换一个陌生题材、换一个陌生读者、换一种他不习惯的体裁，新的写法会立刻出现——他不是不会，是没必要。而本文说的那件事，签名相反：换了题材、换了读者、换了体裁，出现的仍是同一批构式，只是换了词。

这两个签名给出可分辨的操作：取同一个人在两种差异极大的场合下的产出，比较它们的形式集是否有实质交集之外的新增。有新增，是熟练；没有新增，才是本文要说的那件事。第八节第五个场景（同一个人的私人信件与公开写作）跑的正是这个对照，而它跑出来的结果部分不利于本文——私人写作里确实有新增。这一条改了本文的适用范围，见第十五节。

本文的全部工作，就是把“形式库封闭”这件事从一句感慨变成一个可以被查、可以被反驳的判断。

二·三种回答，各自都完整

要说清本文的位置，得先把三种回答讲到它们最强的样子——不是稻草人，是它们的本行专家会点头的样子。

第一种：形就是身份。

化学里的多晶型现象说的是：同一化学组成，可以结成多种不同的固相；不同的晶型有不同的溶解度、稳定性、可制造性。这不是精度问题，是身份问题——两种晶型是两样东西。奥斯特瓦尔德（W. Ostwald）在一八九七年给出的阶段律说，一个体系在结晶时通常先经过亚稳的中间态，而不是一步到达最稳定的那一个。因此“这一次结成了哪一个形”是一个真正的事件，不是必然的结果。

这一种回答落到写作上，是最朴素也最难反驳的一条：一篇作文写成什么样，它就是什么。你不能拿“我本来想写得更好”来改变它是什么。改卷的人只面对纸上那些字，这不是他偷懒，这是判定的唯一可靠基础。

第二种：形是过程的解。

哈克的动态平衡说：地形要素之间彼此调整，只要外部条件不变，整个地形可以在持续的剥蚀中保持形状不变；变的是物质，不变的是形。斯特拉勒、乔利在同一时期把这一路推成“与时间无关的地形”，与戴维斯的“与时间相关的地形”正面对立。

这一路落到写作上，就是过程论：一篇作文之所以是这样，是因为它是那样被写出来的——计划、生成、修改、再生成。这一路在作文教学里是几十年的正主，也是本文必须正面处理的一批人（见第十六节）。它的锋利处在于：你从成品读不出过程。两篇一模一样的作文，一篇是一遍写成的，一篇是改了十一遍的——它们不是同一样东西，而纸上没有任何差别。

第三种：判定席位不在被判定者那里。

免疫学的自我—非我学说自伯内特（F. M. Burnet）的克隆选择起主导了六十年；马青格的危险模型把驱动改成了组织损伤信号；普拉德（T. Pradeu）后来又提出连续性理论作为第三条路。三派吵了三十年，但有一件事三派都同意：“是不是异己”这个判定，不能只看被判定的那个分子。

这一路落到写作上，是接受与场合的一路：同一篇文章，投给这家刊物是好文章，投给那家是不合体例；同一段话，在课堂上是跑题，在网络上是金句。判定权不在文本里。

三种回答各自都完整，各自都能把另外两种解释成“抓错了地方”。这是一个好的僵局——好在它没有裁判可当。没有裁判可当的地方，才有新东西可长。

三·它们共同假定了什么

僵局底下有一块三方都站着的地。

三种回答争的是：一篇作文的身份，该由成品裁、由过程裁、还是由场合裁。

无论哪一方赢，都预设了同一件事：

要被裁的那样东西，已经成形，并且在场。

成品论要有一份成品；过程论要有一段可追溯的过程，而过程之所以可追溯，是因为它在某一时刻产生了可记录的东西；场合论要有一个被投放进场合的对象。

三方吵得再凶，都是在已经出现的东西上吵。

这一条不必辩护，因为它看上去根本不是一个主张，而是判定这件事的前提条件。你不能判定一个不存在的东西——这几乎像一条逻辑规则。

本文要说的是：它不是逻辑规则，它是一个可以被具体材料推翻的经验假定。而推翻它的材料，就在第三种回答自己家里。

四·一个免疫学家不往那个方向用的事实

免疫系统怎么知道什么是“自己”？

标准回答是：在胸腺里学的。发育中的 T 细胞随机重排出各式各样的受体，然后接受两轮筛选。对自身组织完全不反应的，因为没用而死掉；对自身组织反应过强的，被主动删除。剩下的那些，构成这个人一生的免疫库。这就是阴性选择。

这条常识在免疫学里的用途是回答“如何避免自身免疫”。教科书写到这里就转向下一节了。

但把它换一个角度读，它说的是另一件事：

一个系统的“自我”，是由那些从未获准存在的东西的形状划出来的。

免疫库的边界不在库里。你把一个人的全部免疫细胞逐一测序，测出来的是留下来的那些；而“这个人为什么不攻击自己的心肌”，答案在被删除的那一批身上。那一批不留记录。它们没有编号，没有存档，没有任何一份清单记得它们曾经短暂地在胸腺里出现过又被清掉——阴性选择的产物是没有产物。

于是那个共同假定当场破了：

这个系统最要紧的那条边界，恰恰不是由“已经成形并在场的东西”划的。它由一批从未获准成形的东西划出，而这批东西不在任何账上。

请注意本文没有用比喻。免疫学不是在打比方地说“自我由缺席定义”——阴性选择是一个有实验、有敲除模型、有临床后果的机制。本文用的是它的结构：边界由被排除者划定，而排除不留记录。

同一件事在第一种回答家里也有一件。晶型筛查的全部工作是去找“还没出现过的形”，而一份晶型表只记录已经出现过的形。筛查永远不能宣称已经穷尽——从未成核的构型无法从筛查结果里排除掉。更狠的是那个不可逆：一九九八年那种药的原有晶型，在更稳定的形第一次成核之后，就再也做不出来了；决定“这一批是什么”的，不只是在场的那个形，还有那个不再能出现的形。（Dunitz &

Bernstein 1995 ;PNAS 2024 后来用球磨把它重新做了出来——这一条见第十一节，它反过来削弱了本文。）

五·未生界

现在可以下定义。

未生界：一个写作者（或一篇作文）在成形之前就被排除掉、因而不进入任何记录的那些形式，所划出的边界。

它有三条性质，缺一条就不是它：

其一，没有实体。未生界不是一批东西，是一条边界。你不能列举它的成员——一旦列举出来，那一项就已经成形，落到别的格里去了。

其二，不留痕迹。它与“删掉的草稿”的区别是硬的：草稿被删掉，说明它曾经存在过；删除动作本身也常常有记录（修改史、版本、笔记本上的划痕）。未生界里的东西从未存在过，因而没有任何删除动作可查。

其三，有边界效应。尽管没有实体、没有痕迹，它仍然是可以起作用的——它决定了这个人能写出什么。这一条使它区别于纯粹的虚无：虚无没有效应，未生界有。

这里要挡住第二个反问：这不就是“无限的可能性”吗？说一个人没写出无穷多种写法，这不是废话吗？

不是。差别在“可被排除”这四个字。任何一次写作都排除了无穷多种写法，这确实是废话；本文说的不是那个无穷集合，是它里面曾经处在可及范围之内、而被这个人的历史关掉的那一部分。一个从未学过某种语言的人，用那种语言写作的可能性不属于他的未生界——那不是被关掉的，是从未打开过。未生界要求两件事同时成立：曾经可及，且已经不可及，而这个转变没有留下记录。

这一条使未生界成为一个有边界、有成因、可以被条件场移动的东西，而不是一个逻辑上的补集。它也解释了为什么第十一节那条复现（消失的形被重新做出来）能削弱本文而不能取消本文：可及性可以被重新打开，但打开之后长出来的是新的，不是原来那一批。

三重否定：

一篇作文是什么，不是它写出来的那些形式（第一种回答：形即身份），不是它被写出来的那条路径（第二种回答：形是过程的解），也不是它被判读的那个场（第三种回答：判定席位在别处），而是它的未生界。

三条否定各对应一位，不是三种说法的加总。要点在于：前三者全部落在已经成形的东西上——成品在场，过程留痕，场合作用于一个被投放的对象。未生界不在它们任何一个的定义域里。

一个立刻会出现的反问：这不就是“没写的那些”吗？

不是。“没写的那些”至少有四种，本文要分开的正是它们。下一节给出那张表。## 六·两轴四格

“没写的那些”要分开，需要两条轴，一条不够。

轴一：这一形式有没有真的出现过一次。出现过／从未出现。判据是机械的——有没有一个时刻，它以任何形态（写下的字、说出的话、划掉的半句）存在于世界上。

轴二：账上有没有为它设一个字段。有字段／无字段。判据同样机械——在这个共同体的记录系统里，有没有一个位置是留给它的（一栏、一个标记、一条待办、一处批注）。

两条轴交出四格：

	账上有字段	账上无字段
出现过	① 修改史：写了又删的段落、留存的草稿、版本记录、批注划掉的那一句	② 无人记的那一次：口头被否掉的那句、写在纸边被丢的、想到并说出而没人记的
从未出现	③ 被指名的缺口：提纲里那一条没写、老师批“这里应展开”、目录里空着的一章	④ 未生界（本文要打的那一格）：从未成形，也没有任何位置留给它

四格互斥且穷尽（就“某一形式在某一次写作中的状态”而言）。要点是：

- 第①格是过程论的地盘。它是可查的，几十年的写作过程研究做的就是它。
- 第②格是记录不全。它可以靠加装设备（录音、击键记录、屏幕录制）搬进第①格。这是一个技术问题。
- 第③格是已经有人替它开了口的缺席。它有字段，因而可以被清点、被催办、被统计。第十三节那次真跑测的正是这一格。
- 第④格没有字段，也没有实体。它不能靠加装设备搬进任何一格——设备记录的是发生过的事，而这一格里什么也没发生过。

一个直接后果：凡是主张“把没写的记下来”的方案，全部只作用于第②、第③格。它们是好方案，但它们不触及第四格；更要命的是，它们会造成一种“我们已经在管这件事了”的印象。

七·第四格的三个签名

一个格子若要值得单列，它必须有自己的病理，而不只是一个逻辑上的空位。第四格有三条，可以逐条对照：

签名一：短期指标表现为改善。

排除得越早，写出来的东西越干净。不试那些不稳的写法，就不会写砸；不写第一人称，就不会被批“太主观”；每段都给例证，就不会被批“空泛”。在任何一份现行评分表上，未生界大的那一篇分数更高。这一条是本文最不愿意但必须承认的：第四格不是缺陷的表现，它在短期内是能力的表现。

签名二：失效不产生任何信号。

第一格失效有信号（改得越多越乱），第二格失效有信号（事后想起来“我当时说过一句更好的”），第三格失效有信号（那一栏一直空着，能看见）。第四格失效——即一个人的可写形式库过早封闭——不产生任何信号。没有任何一份记录会显示“这里本来可以有另一种写法”。当事人也不会难受，因为难受要有对象。

签名三：自我加固。

越是训练有素，排除得越早、越彻底。而排除得越早，第③格反而缩小——一个学会了不去要求的人，连“这里应该展开”这样的批注都不会招来，因为他每一处都展开得刚刚好。于是从外部看，这个人的记录越来越干净：没有删改，没有缺口，没有批注。第四格越大，前三格越好看。

三条签名各配一个误诊阻挡，防止把别的东西读成第四格：

对签名一的误诊：一个人分数高、写得干净，也可能只是写得好。阻挡：要落进第四格，必须同时满足“形式上无新增”这一条——写得好而形式仍在扩张的人，不在这一格。

对签名二的误诊：任何“看不见的东西”都可以宣称自己失效无信号，这条太便宜。阻挡：必须说得出来如果它有信号，那个信号会长什么样，以及为什么它不会出现。本文的答案是：信号只能来自比较，而比较需要一个对照物，第四格没有对照物——被排除的写法不在场，因而没有任何东西可以与写出来的那一篇对照。

对签名三的误诊：自我加固也是路径依赖的签名。阻挡见第十节末：路径依赖的加固作用在成本上（另一条路还在，只是贵），第四格的加固作用在可及性上（另一条路不在了）。二者的判定形式是：把另一条路的成本降到零，路径依赖预测有人会换，本文预测没有变化。

三条签名必须同时成立，才算落在第四格。只满足前两条的，多半是第②格（记录不全）；只满足第三条的，多半是路径依赖，那是另一件事，本文在第十六节与它划界。

八·一句不含程度词的问话，和一个读数

第四格既然不留痕迹，怎么问？

本文的问法是这一句，它不含任何程度词，可以直接拿去问一叠作业本，不必问任何人的判断：

把这个人的连续产出排开，他最后一次写出一种他此前从未写过的形式，是在第几件？

这句话之所以可用，是因为它绕开了“未生界的内容”（不可列举），只问它的边界什么时候停止移动。一个人不断写出新形式，说明形式库还在扩；某一件之后再也没有新形式出现，说明扩张停了。

读数：末次首现（记作 **L**）。一个序号，不是一个比率。

在五个具体场合跑一遍，答案必须互不相同：

1. 中学生的三年作文本 :L 通常落在初二上学期某一篇。此后所有作文都在已有句式内重排。（可查）
2. 一个专栏作者的十年存稿 :L 往往落在第二年——之后是产量增加而形式不再新增。这一条是查出来才知道的，与本文原先的猜测相反：本文原以为专业写作者的 L 会持续后移。
3. 一个人的求职信 :L =1。第一封之后就是模板替换，此后再无首现。
4. 学位论文各章 :L 常落在文献综述那一章之后——方法与结果两章的形式由体例给定，首现在那里几乎不可能发生。
5. 同一个人的私人信件 :L 显著晚于他的公开写作。这一条也是查出来的，而且它反过来削弱了本文：它说明形式库封闭不是这个人的属性，是这个人在某一类场合里的属性——因此第四格必须按场合分别读，不能给一个人一个总数。

五个答案互不相同；第 2、5 两条不能从命题推出来，是查出来的，而且第 5 条改了本文（见第十五节的适用边界）。

清点手册（同一读数全篇只用这一份定义）：

读数名：末次首现 符号：L

定义：在一个人按时间排序的产出序列（第 1 件至第 n 件）中，

最后一次出现“此前从未出现过的形式类”的那一件的序号。

形式类的粒度：句法层面的构式（如“以疑问句开篇”“插入式定语从句”“无句连用”），

以句法构式为准，不以词汇题材题材为准

一次“首现”的证据：该构式第 1 至 i-1 件中一次也未出现，在第 i 件中出现 ≥ 1 次

边界例子：第 7 件出现了一个增用过的构式，但在新位置上——不算首现

第 9 件出现了一个增只在文中出现过的构式，作者已首次使用——算首现

编辑原则：

1. 只点作者自第 1 阶，引文套语格式标题不计
2. 序必连续同构（第 1 场景：跨场合不得并数）。
3. $n < 8$ 时不报 L，只报“未闭”。
4. 序非尾件内发出首现，L 记为“未闭”，不记序号——防止序列断头或未闭

不用：单篇文本 L 需要序列；命题或批修形式题目给定。



L 与既有主要指标正交，这一条必须举出反例才算数：一个 $L=3$ 的人，可以第 4 到第 40 件里保持全班最高分——形式库封闭与写得好完全可以并存。反过来，一个 L 一直在后移的人，可能每一篇都被批“不稳定”。举得出这个反例，L 才不是“写作水平”的一个代理。

九·边界是怎么划出来的：建立期与运行期

如果第四格没有痕迹，本文凭什么谈它的成因？

凭一条结构：这条边界不是在写作时划的，是在更早的一段时间里划的，而那一段有痕迹。

免疫学在这里给的不是比喻，是分期：阴性选择发生在胸腺，那是建立期；此后一生的免疫应答是运行期。运行期的一切观察都读不到建立期删掉了什么，但建立期本身是有条件的——胸腺里呈现了哪些自身抗原，决定了哪些克隆被删。

写作也有这两期。建立期是一个人形成“什么算作一句话、什么算作一篇文章”的那几年；运行期是他此后所有的写作。运行期里你只看得见他在库内的重排，看不见库的边界；但建立期的条件是可查的：读过什么、被改过什么、什么样的写法被当众否定过、评分表长什么样。

由此得到一条可以去查的推论：同一个建立期条件下的一批人，其 **L** 应当聚集；而运行期的差别（写作量、题材、勤奋程度）不应显著改变 **L**。

建立期的条件里，哪些是可查的？列四项，每一项都能在现行记录里找到载体：

其一，评分表。一份评分表不只是尺子，它是一份形式清单——被列进去的那些成为“该有的”，没被列进去的成为“可有可无的”，而这个区分会在几次评分之后变成“能想到的”与“想不到的”。评分表是最容易查的一项，因为它有纸面。可查的推论：同一份评分表下的一批人，其 **L** 应当聚集；评分表改版之后入学的那一届，**L** 的位置应当整体移动。

其二，被当众否定的那几次。不是批注上的否定（那是私下的、可回收的），是当众的——在课堂上被念出来当反面例子、在会上被点名。第十节说的正反馈事件有一个负号版本：一次当众否定，可以在几件之内关掉一整类写法。可查的载体是课堂录音、评课记录，以及当事人事后能否回忆（回忆得起来的，落在第②格；回忆不起来而实测那类构式再未出现的，落在第四格）。

其三，范文。范文的作用不是提供一个模板，是划定什么算作一篇文章。一个只读过一种范文的人，他的形式库边界与范文的边界高度重合，而他不会觉得那是边界。可查的推论：范文集的构式集合与该批学生 **L** 之后的构式集合，重合度应当远高于随机。

其四，第一次成功。见第十节：第一次被表扬、被选中、被采用的那一次，其构式在此后产出中的占比应当显著上升，且 **L** 与该事件的时间间隔应当很短。

四项里前三项都是条件场（外部给定），第四项是事件（一次性、不可逆）。本文的机制主张要求：前三项决定边界的位置，第四项决定边界什么时候停止移动。两者可以分开检验。

这条推论把本文放到了一个可以被打掉的位置：如果实测发现 **L** 与写作量强相关而与建立期条件无关，本文的机制主张就错了——那时候该说的是“写得多的人形式多”，那是个熟悉得多、也便宜得多的解释。

十·不可逆的那一面：一个形出现之后

第一种回答家里那件事，还有一半没用。

一九九八年那种药的故事通常被讲成“出了个事故”。它真正锋利的地方在时序：在更稳定的那个形第一次成核之前，原来的形一直做得出来；在它第一次成核之后，同一条产线上原形就做不出来了。不是配方变了，不是原料变了。变的是“这个形已经出现过一次”这件事本身。

这一条搬到写作上，是本文的一条推论：形式库的封闭常常不是渐进的，而是由某一次成形触发的。某一次，一种写法被写出来、被表扬、被当成范本；此后其他写法不是被禁止，是不再被想到。

它给出一条可裁决的预测：在 **L** 之前，应当存在一次“高强度正反馈事件”（一次得高分并被公开表扬、一篇被选为范文、一次投稿被采用），且该事件到 **L** 的间隔应当很短（数件之内）。若实测发现 **L** 与任何此类事件都无时间关联，这条推论为伪。

请注意这一条与“路径依赖”的差别，它是本文与最容易混淆的那个说法之间最要紧的一处，第十六节会正面处理：路径依赖说的是成本（换轨代价越来越高，但轨道还在，看得见），本文说的是不再出现（另一条轨不是贵，是不在了）。两者给出相反的干预预测：路径依赖预测“降低换轨成本就能换回来”，本文预测“降低成本没有用，因为没有东西可换”。

十一·风口，以及一次把本文削弱的复现

现在必须写两处反过来削弱本文的地方。它们不是谦辞，删掉任何一处，本文都会比现在强，也比现在假。

第一处，来自第二种回答（地貌学）：风口。

河流袭夺之后，被夺走水源的那一段河谷仍然留在那里——一条不再有水的河的形状，地貌学称为风口。地貌学拿它答“这里发生过袭夺”。

它对本文的杀伤在于：形可以在其对象消失之后继续存在。于是“从成品形态反推未生界”这条路被堵死了——一个人的写作里可能保留着某种形式的空壳（还在用那个句式，而支撑那个句式的东西早就没了），也可能相反。

本文原本打算写下的更强的一句是：“未生界可以从成品的形态特征上读出来”。风口这一条使这句话不能写。本文的读数因此只能是序列读数（**L** 需要一串产出），不能是单篇读数。适用范围随之从“任何一篇作文”缩到“有连续产出序列且序列可得的写作者”。

第二处，来自第一种回答（化学）：消失的形被重新做了出来。

二〇二四年，一组研究者用球磨在不同环境条件下把那种“消失了”的晶型重新做了出来（PNAS 2024）。这一条直接削掉本文关于不可逆的强主张：边界不是永久的，条件场一变，被排除的那一类可以重新获准成形。

本文原本会写下的更强的一句是：“形式库一旦封闭就不可再开”。这一条使它降级为：在条件场不变的情况下不可再开；而“条件场变化”是可以人为制造的（换场合、换读者、换体例、换语言）。

这一处削弱不止削弱了主张，还动了价值排序——见下一节。## □□·耐受不是缺陷

到这里，读者大概已经形成一个印象：未生界是坏的，应当把它变小。

这个印象错了，而且第三种回答自己会当场纠正它。

阴性选择不是免疫系统的失误，它是免疫系统能工作的前提。删除自身反应性克隆的那一步若做得不彻底，后果不是“更自由的免疫库”，是自身免疫病——系统开始攻击自己。一个不排除任何东西的系统，不是一个开放的系统，是一个不能作出判定的系统。

搬到写作上：一个没有未生界的写作者，不是一个可能性无限的写作者，是一个写不出任何一句话的人。落笔本身就是排除。要写下这一句，就得同时不写下无数别的句子；而“不写下”若每次都要经过一次考虑，写作根本无法开始。未生界是写作的成立条件，不是它的病。

这一处摧毁了本文原本最自然的那套价值排序——“未生界越小越好、应当尽量扩大学生的可写形式库”。它写不出来了。本文能保留的只剩一句弱得多、也硬得多的话：

问题不在这道边界有多大，而在它是被谁、在什么时候、按什么划的，以及此后有没有任何人能知道它在哪里。

由此，本文的处方从“扩大形式库”变成了三条完全不同的东西（第二十节给出可裁决版本）：其一，让建立期的条件可查；其二，承认运行期的读数只有 L 这一种，而 L 只报“停没停”，不报“停在哪里”；其三，不要用第②、第③格的方案去冒充第④格的解决。

四种能真的移动边界的手法，以及它们为什么都不是“释放”。

第十一节那条不利于本文的复现（消失的形被重新做了出来）留下了一个正面的口子：边界不是永久的，条件场一变就可能重开。但要注意重开的机制——那不是把被排除的东西找回来，是在新的条件下重新长出一批。四种手法，按可操作性排：

其一，换语言。一个人用第二语言写作时，其形式库的边界与母语写作不同，而且不是母语边界的子集。这一条可查（第十四节第 9 条），也是本文最愿意押的一处：若换语向后 L 不重开，本文的分期主张为伪。

其二，换读者。不是换题材，是换那个“写给谁看”的对象。评分表、范文、当众否定这三项条件都依附于一个特定读者共同体；换掉它，三项一起变。可查的载体是同一作者面向不同发表渠道的产出序列。

其三，换体裁的约束方向。有一类体裁约束是加法（必须包含哪些成分），有一类是减法（不许出现哪些成分）。加法约束不重开边界，减法约束会——因为被禁掉的那些顺手写法一旦不能用，写作者必须现造。这一条给出一个反直觉的教学预测：限制比放开更能重开形式库，而这与“释放潜能”的处方方向相反。

其四，把产出交给一个不共享你的评分表的人接着写。这一条最强也最难安排：接手者的边界与你的不同，交回来的东西会带回一批你不会写的构式，而你此后可能开始用它们。

四种手法的共同点，正是本文与“释放”路数的分界所在：它们都不召回任何东西，它们都是重新生成。一个可判决的推论：用这四种手法之后新出现的构式，与该作者早年被否定过的写法之间不应有超出随机的对应关系。第二节F1 押的就是这一条。

这也是本文与所有“释放潜能”路数的分界：那些路数默认被挡住的东西还在那里，只要撤掉障碍就会回来。本文的判断相反——撤掉障碍之后回来的，是新长出来的东西，不是原来那一批。这两句给出不同的预测：前者预测自由写作会召回旧形式（可辨认为“我以前想过的”），后者预测召回的形式与该作者的历史无关联。这一条可查（见第二十节 F4）。

十三·一次真跑：被指名的缺口去了哪里

第四格不可直接观测。本文因此去测第三格——被明确指名、留了字段的缺口。理由写在预注册里：如果连有字段的缺席都留不住，那么无字段的那一格连一个下界估计都得不到。

维基百科的红链是第三格一个干净的实例：一个条目在正文里链接到另一个条目，而那个条目从未被写出来。它有字段（那个链接本身），没有实体。

预注册在取数之前落盘，写死了三条假设与阈值。做法：随机抽取英文维基百科主命名空间条目，保留当前字节数不小于一万五千者共八十篇；其中在二〇一六年一月一日（记作 T0）已存在的五十八篇进入分析。取每篇在 T0 之前最近一次修订的正文，抽出全部条目链接，共五千一百一十五个不同目标。判定其中哪些在 T0 时尚不存在（红链），再看它们到二〇二六年八月的去向。

先报两个与假设无关、但比假设更值得看的数：

- 红链只占链接目标的百分之七点二（ $367/5115$ ）。
- 五十八篇里有二十三篇一条红链也没有；有红链的那些，中位数是三条。

也就是说：在一个专门为“待写条目”设了字段的系统里，这个字段也只覆盖了不到十分之一，四成条目一条都没设。为缺席设字段这件事，即使做了，也得极浅。

三条假设的结果：

第一条（主假设）：红链的抹除率不低于百分之三十。结果：失败。

去向分三类：目标后来被写出来了（成形）；仍不存在而链接还在（仍红）；仍不存在而链接已从源条目消失（抹除）。实测三百六十七条：成形一百二十条（33.0%），仍红一百九十一条（52.0%），抹除五十五条（15.0%）。

换三种算法一致不过线：剔除那篇有一百一十九条红链的离群条目后，样本二百四十八条，抹除率 20.6%；每篇各算各的抹除率再取中位（限有三条以上红链的二十一篇），20.0%。

这条失败改掉了本文的什么：本文原打算主张“被指名的缺口本身会被系统抹掉，因此第三格不稳定，因此第四格没有下界”。实测说的是相反的事——缺口相当稳定，超过一半在十年零七个月里原地不动。本文因此不能再用“抹除”作论

据。能保留的是另一件事，而它其实更硬：缺口不是被抹掉的，它是极少且极情的；被指名并不导致成形（十年里只有三分之一被写出来）。第三格与第四格因此是两种不同的东西，前者的读数不能外推到后者——这恰恰关掉了本文自己原本打算走的那条估计路径。

第二条（用于排除“其实不就是没人感兴趣吗”）：不可执行。

预注册用的需求强度代理是“T0 时被样本内多少个不同条目同时指名”。实测三百六十七条的度数全部为——五十八篇随机条目之间几乎不共享红链，分层退化成一层。改用“该目标当前全站入链数”作代理的补救取数被接口速率限制打断（HTTP 429），只取到二十六个。这二十六个的方向不利于本文：成形组的入链数中位是二十五，仍红组是一，抹除组是零。

两条限定必须同时写下：样本二十六不足以判决；而且入链数是处理后变量——一个条目被写出来之后才会吸引入链，因果方向可以反过来。结论是：本次没有排除掉最强竞争解释。按本文自己的标准，这次检验只做到了现象层，没做到机制层。

第三条（抹除的中位时间早于成形的中位时间）：未跑。

成形一侧有时间戳（中位二〇二〇年一月，最早二〇一六年一月三日，最晚二〇二六年八月十四日）；抹除一侧没有——要定位一条链接在源条目上千次修订中的哪一次消失，需要逐条二分检索，成本超出本次预算。这一条如实记为未执行，本文不作任何关于时间不对称的声称。

一处预注册修订，一并公开。原预注册第二节有一处自相矛盾：一边规定“目标当前不存在则无法判断它在 T0 时是否存在过，列为不可判”，一边又把“仍红”与“抹除”定义为“目标当前仍不存在”。按字面执行，分母里只剩成形一类，抹除率恒为零——这个设计不可能被证伪。修订发生在算出任何分类结果之前，改的是判定程序（改用公开删除日志判断该页面在 T0 前是否存在过），三条假设的阈值一律未动。不可判十一条。

三条稳健性核对，一并报出。

其一，离群条目的影响。五十八篇里有一篇（一次残奥会乒乓球赛事的条目）独占一百一十九条红链，占全样本近三分之一。这类条目的红链是批量生成的——一张参赛者名单里每个人名都是一个链接，而绝大多数运动员没有条目。把它剔除后，抹除率从 15.0% 升到 20.6%，成形率从 33.0% 升到 39.1%。方向一致，仍不过三成的线。

其二，每篇等权。按链接数加权会让少数大条目主导读数。改为每篇各算各的抹除率再取中位（限有三条以上红链的二十一篇），得 20.0%。第三种算法，同样不过线。

其三，分母的偏向。十一条因 T0 前有删除记录被判为不可判，未计入分母。这一偏向的方向是同时低估成形率与抹除率（“创建后被删”这一类既不算成形也不算抹除），因此它不会把一个高于三成的真值压到 15%。

这次真跑真正教给本文的，不是那三条假设中的任何一条。是那个百分之七点二。

一个为“待写条目”专门设了字段、任何人都可以随手添一条、且这件事被鼓励了二十年的系统，它的缺席字段覆盖率是百分之七点二，四成条目一条都没有。这个数字比抹除率有用得多：它说明为缺席设字段这件事，即使在最有利的条件下也做得极浅。而本文第四格连这样一个字段都没有。

由此还得到一条本文事先没有预测、因而只能记为观察的读数：被指名并不导致成形。十年零七个月里，只有三分之一的红链变成了条目；超过一半的缺口在被反复看见的情况下原地不动。把一个缺席说出来，并不使它更容易被填上。这一条如果稳健，它对第十九节那条反噬有直接后果——要求人们记录“我本来想写的”，即使记了，也大概率不会导致那些东西被写出来。

代理变量三处，先写明：目标当前不存在不等于从未存在过（本次靠删除日志排除，仍有页面移动一类未区分）；链接从当前正文消失不等于有人有意删掉这个缺口（整段重写、条目拆分、模板化会造成相同读数）；红链只是第三格里留下机器可读字段的那一小部分，提纲、批注、口头要求全查不到。

十四·十二处兑现

一个判断若只在写作里成立，它多半只是一个说法。下面十二处是可以据此在各自领域预测一件具体的事的地方，不是“像”。

1. 软件工程：一个团队的接口设计里，L 应当落在项目早期的某一次评审之后；此后所有新模块在已有的抽象里重排。可查：接口签名的构式首现序列。
2. 司法文书：同一法院同一庭的判决书，说理部分的构式首现应在法官任职早期停止；此后为模板替换。可查，且与案件类型无关。
3. 医学病历：住院医师的病程记录，L 通常落在轮转的第一个科室之后。这一条给出一个反直觉预测：换科室不会重开形式库，因为条件场（病历规范）未变。
4. 教材编写：同一套教材的历次修订，若修订者不变，例题的构式首现应当在第二版之后停止；换人则重开一次。
5. 建筑设计：一家事务所的平面组织方式，L 与获奖事件的时间关系应当很近（第十节的正反馈预测）。
6. 音乐创作：一个作曲者的和声语汇，L 应当早于其产量高峰。
7. 围棋定式：一个棋手的开局形式库封闭之后，胜率可以继续上升——这正是 L 与水平指标正交的一个可查实例。
8. 专利撰写：同一代理人撰写的权利要求书，其句法构式的 L 极早（通常 $\square$ 3 至 5 件），而这不影响授权率。
9. 翻译：同一译者的处理方式，L 在换语向时重开——因为条件场真的变了。这是第十一节“条件场一变可以重开”的可查实例。
10. 实验方案设计：一个实验室的方法学首现应当在导师更替时重开一次；若实测发现不重开，本文关于建立期的主张被削弱。

11. 客服话术：坐席的应答构式 L□1 或2，而满意度评分与 L 无关——第八节那条正交性的最容易取到的证据。

12. 学术审稿意见：同一审稿人的意见构式，L 与他的领域声望无关联；这一条可以直接从公开评审（如某些期刊的开放同行评审）里查。

这十二处不是“像”。判准是硬的：每一处都必须能据本文在那个领域预测一件具体的、可以去查的事，而且预测的方向由本文而不是由常识给出。□3 条（换科室不重开）与第 9 条（换语向重开）是一对：常识会说“换环境就会有新东西”，本文说不一定——要看条件场是否真的变了，而病历规范没变、语言变了。□7 条与第 11 条是另一对：它们都预测形式库封闭与成绩指标脱钩，而这与“形式僵化即水平下降”的常识相反。若这四条里有三条实测与本文方向相反，第八节的正交性主张与第九节的分期主张一起倒。

其中第 3、9、10 三条与本文的机制主张构成可判决的对照：条件场不变则不重开、条件场真变则重开。若第 9 条实测不重开，本文的建立期—运行期分期为伪。

十五·不适用的边界

1. 单篇文本读不到。L 需要序列，最少八件（清点手册第 3 条）。给一篇作文判断它的未生界，本文做不到，也不许别人拿本文的名义去做。
2. 命题作文批次读不到。形式由题目给定时，首现的缺失不能归给写作
3. 跨场合不能合并。第八节第五场景查出来的事：同一个人的私人信件与公开写作，L 显著不同。未生界是“这个人在这一类场合里”的属性，不是这个人的属性。任何把 L 当成个人能力标签的用法，本文明确反对。
4. 不适用于内容与观点。本文清点的是句法构式，不是题材、立场、思想。把 L 用来判断一个人“思想是否僵化”，是越界，且不可证伪。
5. 不适用于机器产出。一个语言模型的形式库不由建立期—运行期这套分期刻画；它的“从未出现”是采样问题，不是排除问题。这一条边界很硬：本文的机制要求排除发生在一个有历史的个体身上。
6. 序列不可得就读不到。这是第十三节那次真跑逼出来的：本文原以为可以从“缺口清单”侧面估计，实测证明不行。没有序列，本文什么也读不出。

最后一条边界，关于本文这个判断本身的用法。本文提出的不是一个诊断工具，是一个读数。读数只报一件事：形式库停没停。它不报停在哪里（那需要列举第四格的成员，而按定义不可能），不报为什么停（那需要建立期的材料，而那些材料在多数场合已经不可得），也不报停了是好是坏（第十二节已经说明，停本身是写作的成立条件）。

任何把 L 读成“这个人写作能力的上限”的用法，都越过了本文。本文能支持的最强的一句话只有这一句：这个人在这一类场合里，形式上已经不再产生新的东西了；如果这不是他想要的，那么可做的事在条件场那一侧，不在他这一侧。

十六·与最近的几种说法的分界

下面十二位，每一位都比本文更早、更有名，且都被读者立刻联想到。逐位写：它在自己领域回答的是什么、差别落在哪、以及看到什么就说明它对而本文错。

一、训练出来的无能 (Burke 《Permanence and Change》1935，概念借自 Veblen)。它回答的是：训练如何使人在新情境里看不见相关的东西。差别：它讲的是注意的失效，被忽略的东西仍在场、仍可被指出；本文讲的是不在场。判定形式：若把被排除的写法当面示范给写作者，他能认出“对”，这个我该想到”并当场用出来，则是训练出来的无能而本文错；若他能用出来的只是对示范的模仿、且在下一件产出中不复现，则本文对。

二、语言相对论 (Whorf 1940)。它回答的是：语言结构如何使某些区分难以做出。差别：它的边界由语言给定，同语者共享；本文的边界由个人史给定，同语者之间可以完全不同。判定形式：若同一语言共同体内 L 的分布高度集中，则相对论解释足够而本文多余；若 L 的方差主要由个人建立期条件解释，则本文对。

三、图式 (Bartlett 1932)。它回答的是：回忆如何被既有图式重构。差别：图式重构留下替换痕迹（人会记成另一个版本），本文的第四格不产生任何版本。判定形式：若被排除的形式能在提示下恢复，则是图式而非未生界。

四、压抑 (Freud 1915)。它回答的是：观念如何被挡在意识之外。差别：压抑物会回返（口误、梦、症状），这正是精神分析的核心主张；本文的第四格没有回返机制，因为没有被压抑的对象。判定形式：若排除掉的写法以任何变形的形式在产出中反复出现，则是压抑而本文错。

五、内在编辑者 (Elbow 《Writing Without Teachers》1973)。这是本行里最像的一位。它回答的是：如何让写作者绕开成形前的审查。差别正好落在第十二节：自由写作的整套技术默认被挡住的东西还在那里、可以召回；本文说不在。判定形式：让写作者做足量自由写，然后核对新出现的形式是否与他早年被否定过的写法对应——对应则 **Elbow** 对而本文错；若新形式与其历史无关联，则本文对。这一条是本文最愿意押的一条，也是最容易被打掉的一条。

六、僵化规则与写作阻塞 (Rose ,CCC □ 31 卷 1980)。它回答的是：内化的规则如何造成阻塞。差别：阻塞是有信号的（写不下去，本人难受），本文的第四格无信号、本人不难受。判定形式：若排除总伴随可报告的卡顿，则是阻塞。

七、过早编辑 (Perl ,RTE □ 13 卷 1979)。它回答的是：不熟练写作者为何反复中断。差别：过早编辑作用于已写出的字，本文作用于未写出的形式。判定形式：过早编辑可在击键记录里读到（删除动作密集），第四格在击键记录里完全不可见。

八、修改策略研究 (Sommers ,CCC □ 31 卷 1980：学生倾向只加不删)。它回答的是修改行为的差异，全部落在第①格。差别是格位差别，不是程度差别。

九、潜在结果框架 (Rubin 1974)。这是最危险的一位。它回答的是：如何谈论未发生的那一个结果。差别：潜在结果是可指名的——对每个单元定义了处理与不处理两个结果，未观测是抽样问题，可以用设计去逼近；本文说的那一类连“有几个”都问不出来，不是缺一个观测，是缺一份名单。判定形式：若能对某个具体写作者列出一份“他本可以写出的形式”的候选集并对其做估计，则潜在结果框架足够而本文多余；若任何这样的候选集都只能由研究者自己发明（因而随研究者而变），则本文对。

十、选择偏差 (Heckman 1979)。它回答的是：只观测到被选中样本时如何纠偏。差别：选择偏差假定未观测样本属于同一总体、可建模；本文的第四格没有总体。判定形式：能不能写出选择方程——写不出，说明不是选择问题。

十一、结构零与抽样零的区分 (统计学标准区分)。它回答的是：某些格在原理上不可能有观测。差别最细也最要紧：结构零由模型给定（一个男性不会有妊娠记录），本文的边界由历史造成，且同一个人在不同场合不同 (第十五第3条)。判定形式：若 L 可由该人的固定属性完全预测，则是结构零；若 L 随场合变化，则本文对。

十二、**doxa** 与“不可思之物” (Bourdieu 《Esquisse d'une théorie de la pratique》1972)。它回答的是：什么使某些问题根本不被提出。差别：doxa 由场域位置决定，同一位置的人共享；本文的边界由个人产出史决定。判定形式同第二条。

十三、知觉窄化 (Werker & Tees, *Infant Behavior and Development* 7, 1984 ; Kuhl 等 1992 ; 跨通道综述见 Lewkowicz & Ghazanfar 2009)。这是本文最近的一位，而且与第九节借来的免疫学同族——原稿借了它的分期而没有与它划界，是一处欠账，此处补上。它回答的是：为什么一个人后来分不出某些他曾经分得出的差别。六至八个月的婴儿能分辨世界上几乎任何语言的音位对立，到十至十二个月，母语里没有的那些对立就分不出了。这个丧失由经验造成，此后一生不产生任何主观信号。

差别落在两处：可及性能否被外部操作重新打开，以及边界能否被枚举。知觉窄化的对象是可枚举的——实验者事先知道那一对音位是什么，可以把它放到受试者耳边测出阈值，也可以用训练把辨别力部分恢复。本文的第四格明写不可列举 (第五节性质其一)，且第十九节说明任何为它设立字段的动作都会把进入字段的那一部分变成别的东西。

判定形式：若能对某位写作者事先列出一份“他已经写不出的形式”的清单，并像播放 **Hindi** 卷舌音那样把它放到他面前测出一个可复现的阈值，则本文这一格是知觉窄化在写作上的一个实例，不必单列；若任何这样的清单都只能由研究者自己发明、因而随研究者而变，则本文对。这与第九位（潜在结果框架）的判定形式是同一条——两位从不同方向逼到同一处。

十四、理论形态空间的空区 (Raup, *Journal of Paleontology* 40, 1966 ; Maynard Smith 等, *Quarterly Review of Biology* 60, 1985 ; 尤其 Gerber, *Evolution & Development* 16, 2014 : “Not all roads can be taken: development induces anisotropic

accessibility in morphospace”)。它回答的是：为什么理论上可能的形态，绝大部分从未出现过。**Raup** 用三个参数把螺壳的形态张成一个空间，发现真实物种只占其中很小几块；发育约束的传统进一步指出，某些方向不是没被选中，而是发育过程本身使它们不再可达——**Gerber** 那篇的标题说的就是这件事。

差别落在空间能否被参数化。形态空间的空区之所以能被谈论、被测量、被制图，是因为它事先由一个几何模型张成，空区因而是有坐标的。本文的第四格没有这样的模型：写作的形式空间不由任何一组参数张成，因而“空的那一块在哪里”这个问题没有坐标可答（第五节性质其一与其二）。

判定形式：若有人给出一组参数，把写作的形式空间张成一个可制图的空间，并在其中标出某位写作者的未占据区，则本文的第四格就是那个空区，本文应当被并入形态空间这一支；若任何这样的参数化都只能事后由已出现的形式反推——因而未出现的方向恰好是它测不到的——则本文对。这是本文自认最可能被推翻的一条：它只需要有人做出一个成功的形式空间参数化，而不需要推翻本文的任何论证。

十五、非决策 (**Bachrach & Baratz, American Political Science Review 56, 1962**；及 **Power and Poverty, 1970**；第三面权力见 **Lukes 1974**)。它回答的是：某些议题为什么根本不进入议程。与第十二位 (**doxa**) 同族而更硬——它明确处理“什么也没发生”这一类，并因此长期被批评其为不可经验研究。

差别：非决策有受益者。它的全部解释力来自“偏见的动员”——某些人从这个议题不被提出中获益，因而它可以由利益格局去追。本文的第四格没有受益者可指（第十六节末逼出来的那句：有一类缺席，没有缺席者）。判定形式：若某位写作者的形式库封闭可以由“谁从他不这样写中获益”完整解释，则是非决策而非第四格；若封闭在没有任何人从中获益的场合同样发生（私人日记、无人阅读的练笔），则本文对。

这一条有一个对本文不利的现成读数，须写明：第八节第五场景查出私人信件 **L** 显著晚于公开写作，也就是说没有受益者的场合封闭得更晚——这与“有受益者时封闭更早”是相容的。本文因此不能声称已经排除非决策，只能说它在私人场合仍有一个非零的封闭量待解释。

以上三位是补入的。第十六节末尾“十二位”“十二行”的说法指原有的十二位；补入后共十五位，而第二层地基的结论不变——三位新入者同样站在那块地上：知觉窄化假定被丧失的辨别力可以被指名并放到受试者面前，形态空间假定未实现的形态有坐标，非决策假定不被提出的议题有一个受益者。三者都预设了那句话：凡是一处缺席，都必定有一个可以被指名的缺席者。

这十二位里有四位——第一、二、五、十二——从四个方向逼近同一件事而都停在同一步之前：它们都假定被挡住的那样东西曾经在那里、并且能够被召回。本文补的正是那一步：有一类缺席，没有缺席者。

第二层的那块地。逐位问“它把什么当作给定、因此看不见什么”，十二行写下来，背后是同样东西：凡是一处缺席，都必定有一个可以被指名的缺席者。

承认“有一类缺席没有缺席者”，这十二位全都会塌——潜在结果框架无法定义 $Y(0)$ ，选择偏差写不出选择方程，自由写作没有召回对象，doxa 找不到那个“不可思之物”是什么。“未被辨识”的真正来源不是没人看，是看的人都站在这块地上。## 十七·与同一批另外三篇的分界

同一批里另有三篇处理相邻的问题，它们是本文最近的对照物，也最容易读者混作一谈。逐篇给一条判决性对照预测。

其一，关于主动停止的那一篇。它说：一篇作文的锋利处在于，在若干个仍可继续的地方主动停下来，使别处的分支得以启动；它的读数是让顶率。差别是格位差别：它整篇落在第①格——被停掉的那一支长出来过，因而有痕迹、有位置、可清点。本文落在第④格：从未长出来。判决性对照预测：取一批让顶率极高的写作者，测他们的 L 。若让顶率与 L 正相关，那一篇的机制足以覆盖本文；若二者正交（一个人可以频繁停掉已长出的分支，而形式库早已封闭），则两者是两件事。本文押正交。

其二，关于残余物名分的那一篇。它说：一类残余物获得名分，当且仅当存在一条以它为清偿对象的职业通道；它的读数是清偿率。差别：它处理的是已经存在而没有地位的东西（负形）；本文处理的是从未存在的东西。两者字面接近而所指相反，必须写死这一句：负形有实体没地位，未生界没有实体。判决性对照预测：在一个清偿率为零的圈子里（没有任何职业通道以残余物为对象）， L 可以任意早也可以任意晚——清偿制度不影响形式库何时封闭。若实测发现清偿率高的圈子里 L 系统性后移，则那一篇的机制吸收了本文。两篇的接口只有一处：那一篇解释了为什么没人为第③格补字段（补了也没人清偿），而本文的真跑证明即使补了字段也没用（第③格并不通向第④格）。两处合起来，是一个互补关系，不是“我们不一样”。

其三，关于计数的那一篇。它问的是谁被计入。差别是问题差别：计数问题的前提是有一份名单可数；本文的第四格没有名单。

还有一处必须自己点出来的同形。同一批里另有一篇的地基句式是：“每一次不变，都必定有一个正在起作用的原因持有者。”本文第十六节末尾逼出来的那一句是：“凡是一处缺席，都必定有一个可以被指名的缺席者。”两句形状相同——都是“某个空位背后必有一个持有者”。这是本批最近的一次自撞，不写出来读者也会看出来。差别：那一篇否定的是持有者的唯一性（可以没有人在持有，而状态照样不变）；本文否定的是缺席者的可指名性（缺席可以没有对象）。判决性对照预测：那一篇预测存在“无人负责而稳定”的状态，可由责任分配记录查；本文预测存在“无对象的缺席”，只能由序列读数查。两条证据来源互不重叠，因而两者可以一真一假。

十八·这个判断对它自己适用吗

本文不写“本文也有局限”。本文用自己的判据给自己定位，并说出后果。

先定位。本文出自一条固定的产线：先取三种互不相容的看法，找出它们共同的假定，再用其中一家自己的材料把那个假定推翻。这条产线到本文为止已经跑了很多次。

拿 L 去测这条产线：它的末次首现在哪一件？

诚实的回答是——本文很可能在 L 之后。本文的结构位（两轴四格、一句无程度词的问话、一个可清点的读数、十二位分界、一次预注册真跑、一个写死日期的赌注）在这条产线上已经反复出现。本文这一次换掉了四个位（读数由比率改成序号；证伪主件由赌注改成预注册报告；自反由“本文自己落在第四格”改成本节这一测；第三层主张换了形状），但换位本身也是这条产线的一个已有动作。

后果有三，逐条说明：

其一，本文对自己的第四格是盲的。本文能查的是这条产线写出来的那些口目，查不到它在选题时从未想到的那些做法。按本文自己的判据，那一部分不留痕迹——而它恰恰决定了这条产线能做出什么。本文无权声称自己不在第四格里。

其二，本文的第一处签名成立。换位之后，本文在各项形式检查上表现更好：结构更整齐、读数更干净、证伪更硬。短期指标改善，正是第四格的第一个签名。这一条对本文不利，而本文找不到办法排除它。

其三，一条可查的自我检验。把这条产线的历次篇目按时间排开，清点其结构位构式的首现序列，算出它的 L。若 L 早于本文，则本文是形式库封闭之后的一次重排，其新意在形式层为零，只在内容层。这一条谁都可以去算，本文不掌握它的答案。

十九·反噬，以及一个我看不到的出口

一条判断若被制度采用，最省事的实现方式往往会摧毁它要保护的东西。本文的这一条尤其明显。

最省事的实现方式是：要求写作者记录“我本来想写而没写的”。

这个动作把第④格的东西搬进第③格（一旦被指名，它就有了字段），甚至搬进第②格（一旦说出，它就出现过了）。搬进去的那些不是未生界的成员——按定义，能被写下来的就不属于它。于是记录系统会被一批看起来很像、实则是别的东西的条目填满，而真正的第四格纹丝不动。

更糟的一层：这样的记录一旦进入考核，它会变成一个可以表演的项目。“我本来想写”会被生产出来——学生会学着填那一栏，填得漂亮的人得分高。第十二节说的“排除得越早分数越高”，在这里变成“表演排除得越好分数越高”。这条边界因此被推得更深，而记录上却显得这件事已经被管理起来了。

第二层反噬来自读数本身。L 是一个序号，看起来极适合拿去排名。一旦拿去考核，最省事的达标法是制造首现——在第八件里故意写一个从没写过的句式。第八节清点手册第 4 条（末尾三件内的首现不记序号）只挡得住一部分，挡不住有意为之的持续制造。

我看不到出口。不是措辞谦逊：任何为第四格设立的字段，一旦被设立，被写进去的就不再是第四格的东西；而不设字段，它就仍然不可见。这个循环本文解不开。本文能做的只有一件事——把这个循环写在正文里，使任何据此设计考核的人无法说自己不知道。

伦理边界（三条，缺一条本文不得用于任何评价场合）：一、**L** 指的是位置，不是人。它读的是“某人在某类场合的产出序列”，不是这个人的能力、潜力或思想状态。二、不得为了把这个数做好看，而在教学或考核里安排形式轮换。制造出来的首现不是首现；在需要稳定表达的场合（法律文书、医嘱、操作规程）人为制造形式变动，会造成实际风险。三、已经据此作出不利安排的（分班、评级、录用），必须公开编码规则并给出复核通道。第八节的清点手册就是为此写的：粒度、边界例子、四条编码规则齐全，任何人可以拿去复算并提出异议。

二十·证伪条件，与一个写死日期的赌注

本文分三层，各设各的证伪条件；第一层最强也最易倒，第三层最稳。

第一层：未生界作为核心判据。

F1（最愿意押的一条，同时是与第五位的判决性对照）招募若干写作者，取其早年被当面否定过的写法清单（可由教师批注与保存的旧稿重建），再让其做足量自由写作，清点新出现的形式与该清单的重合度。若重合度显著高于随机基线，本文为伪——那说明被排除的东西一直在那里、可被召回，Elbow 那一路足够。

F2（机制证伪，须排除最强竞争解释）最强竞争解释是：“其实不就是写得少吗”。排除方式：控制写作总量与写作年限之后，若“建立期条件的差异”与 **L** 之间存存量一反关系，则文的排伪为——属形式排伪，应因产量而建期，样要求：同制度境，多个单元，同学级，同报，多版组，**N** 不小于六，关键混淆变量同质！本文尚未完成这一条。第二层跑通之前，且测是四格，与条无关，在 **F2** 被跑通之前，本文的机制主张只是一个候选。

第二层：两轴四格作为分析工具。

F3 若能举出一个案例，它同时满足第四格的三条签名（短期改善、失效无信号、越正规越严重）而又能被第①②③格中任一格完整解释，则四格划分为伪（该格不必单列）。

F4 若在某个记录系统里，第③格的读数被证明可以有效预测第④格（例如缺口密度与 **L** 显著相关），则本文关于“两格不可互推”的主张为伪。这一条本文自己跑过一半：第十三节的结果是第③格极少、极情，且其动态可能完全由需求强度解释，因而暂时不支持互推；但入链数那二十六个样本方向不利于本文，须以更大样本复跑。

第三层：一条不依赖前两层的经验主张（最稳，也最容易查）。

F5 在一个专门为“待写内容”设了字段的公开记录系统里，该字段的覆盖率极低、且被指名并不导致成形。实测读数：覆盖率 7.2% (667/5115)，四条条目零覆盖，十年零七个月内的成形率 33.0% 若在另外三个可比系统（公开代码仓的待办标记、公开数据集的缺项登记、开放同行评审的“未回答问题”栏）中，覆盖率均高于 **30%** 或成形率均高于 **70%**，则本条为伪。

F6 若L与常用写作水平指标（评分、录用率、引用）呈显著正相关，则L不是一个独立读数，只是水平的一个代理，第八节的正交性主张为伪。

F6 追记（已跑，**H2** 判负；**H1**、**H3**、**H4** 在预注册口径下不可执行）。预注册文件写于取数之前，**SHA-256** 前十六位 **279bfe9deeeefebbb**。数据：**arXiv math.NT** 分类按首次提交时间升序的 **10,000** 条元数据；以第一作者姓名字符串完全一致分组，取件数不少于 **8** 者，得 **166** 位作者、**2,144** 篇；产出序列取摘要。形式类按预注册的功能词 **n-gram** 代理操作化：凡不在封闭功能词表中的词一律掩成占位符，形式类即掩码序列上的三元组。

结果一（**H2**，判负，且是本次最要紧的一条）：**166** 位作者，未封闭率 **u = 1.000**。一位也读不出 **L**。按第八节清点手册第 **4** 条，末尾三件之内出现首现即记“未封闭”，而实测每一位作者都在最后三件里仍有新构式出现。**H1**、**H3**、**H4** 因此在预注册口径下不可执行，本文不作任何关于它们的声称。

事后探查（明标为事后，不作证据）。把开放的形式类换成一份事先闭合的构式清单——取全语料掩码三元组的高频前 **K** 个——未封闭率随清单变粗而下降：**K=1000** 与 **K=500** 时仍为 **1.000**，**K=200** 时 **0.982**，**K=100** 时 **0.952**，**K=50** 时 **0.681**。只有在最粗的一档上才有 **53** 位作者可读出 **L**：其 **L/n** 中位数 **0.667**，**Spearman(L/n, n) = -0.327** (**p = 0.017**)，**Spearman(L/n, 已发表比例) = -0.083** (**p = 0.56**)，**Spearman(L/n, 跨越年数) = 0.089** (**p = 0.53**)。这三个数若按预注册阈值读，**H1** 过线（“其实不就是写得少吗”这条竞争解释被排除）、**H3** 与 **H4** 过线（正交性成立）。但它们出自一个事后选定的粒度，本文不拿它们当支撑。

这一条改掉了本文的什么。第八节把清点手册当作可复算性的全部保证，而手册的第一行——“形式类的粒度：句法层面的构式”——没有给出一份清单。本次读数说明这不是一处待补的细节，而是决定读数存不存在的那一处：在细粒度上 **L** 恒不存在，在粗粒度上 **L** 存在且行为与本文的预测一致，而“多粗才算一个构式”没有任何独立于结论的根据。

第八节因此补一条硬约束：任何 **L** 的报告必须同时公布所使用的构式清单，且该清单须在清点之前落盘；未公布清单的 **L** 读数一律不成立。这一条同时使 **F6** 成为可裁决的——在两份都合理的构式清单下，若 **L** 与水平指标的相关系数符号不一致，第八节的正交性主张即为伪。

本次读数的三条局限，前两条已写在预注册里：功能词三元组不等于句法构式，它会把词序变化误读为构式首现，方向使 **L** 偏晚；摘要是高度受限的体例，形式库本就窄，不可外推到自由写作；第一作者同名会把两个人并成一个序列，方向同样使 **L** 偏晚。三条都是保守方向——真实的封闭可能比本次读到的更早。但这不能用来救 **H2**：**u = 1.000** 是“读不出数”，不是“数偏大”。

F7 若可写作特例的 **L** 高度一致（私人写作与公开写作、母语与外语），则“未生界是场合属性”为伪，第十五节第 **3** 条须撤回。

F8（自反） 若按第十八节的做法算出本产线的 L 明显晚于本文，则本文关于“结构位复用即形式库封闭”的自我诊断为伪——那说明本文自己的形式仍在扩张。本文不掌握这个答案。

本文解释不了的两个洞，交出去：其一，为什么有些人的 L 会突然后移（换语言、换体裁之后重开），而另一些人换了也不重开——本文只能说“条件场变没变”，这在事后总是可以圆回来，因而这条解释目前不可证伪，是一个真正的漏洞。其二，第②格（出现过而无人记）与第④格在实践中如何区分——本文靠“当事人能否回忆”来分，而回忆本身不可靠。

赌注（写死日期，并写死什么不算命中）：

二〇三二年八月一日之前，至少有一个规模化的写作教学或写作辅助系统，在其正式规则或产品功能中，出现一个按“连续产出序列的形式首现”计算的读数，并且该读数被用于提示形式库封闭（而不是用于评分）。

本文押它成立的理由不是乐观，是第十三节那个百分之七点二给的反面提示：为缺席设字段这件事，做起来极浅；而序列读数不需要任何人为缺席设字段——它只需要把已有的产出按时间排开。这是目前唯一一条不必先解决“如何记录不存在的东西”就能落地的路径，因而它比任何“请写下你本来想写的”方案都更可能真的被做出来。

四条不算命中：（一）只在研究论文里出现，未进入任何正式规则或产品功能；（二）只统计“用词丰富度”“句式多样性”这类截面指标——本文押的是序列上的首现，截面多样性不算；（三）该功能要求用户自行填写“我本来想写的”——那是第②③格，不是第④格；（四）该读数被用于排名或评分——按第十九节伦理第二条，那种用法即使出现也判不命中。

结论

关于作文是什么，三种通行的回答各自完整，且互不相容：写出来的那些东西、写出来的那条路径、判读它的那个场合。三者共同假定要被判定的东西已经成形并且在场。这个假定不成立，而不成立的那一部分——本文称为未生界——恰恰决定了一篇作文能是什么。

未生界没有实体、不留痕迹、有边界效应。它与被删的草稿、无人记的那一次、被指名的缺口分属四格中不同的三格；它所在的第四格有三个签名：短期指标表现为改善、失效不产生任何信号、越训练有素越严重。本文给出的读数是一个序号——末次首现，它只报形式库停没停，不报停在哪里，且与写作水平正交。

本文最愿意押的那条经验假设在预注册检验中失败了：被明确指名的缺口并不会被系统抹掉，它只是极少（7.2%）且极惰（52% 十年不动）。这一失败关掉了本文原打算走的估计路径，把适用范围收窄到有连续产出序列可查的场合，并把第三层主张改成了一条带数字的稳态主张。用于排除最强竞争解释的那一条未能执行，因此本文的机制主张目前只是一个候选。

未生界不是病。它是写作得以开始的条件。要问的不是这道边界有多大，而是它由谁、在何时、按什么划出，以及此后有没有任何人能知道它在哪里。

注释与参考文献

1. 第二节地貌学一路的对立方：Davis, W. M., "The geographical cycle", *The Geographical Journal*, 1899□Hack, J. T., "Interpretation of erosional topography in humid temperate regions", *American Journal of Science*, 258-A 卷, 1960。同一路的续：Strahler 1950□Chorley 1962□
2. 免疫学三派：Burnet 的克隆选择学说（1957 及其后续扩展）；Matzinger, P., "Tolerance, danger, and the extended family", *Annual Review of Immunology*, 12 卷, 991-1045 页, 1994□Pradeu, T., *The Limits of the Self: Immunology and Biological Identity*, Oxford University Press, 2011□
3. 多晶型与阶段律：Ostwald, W., *Zeitschrift für physikalische Chemie*, 22 卷, 1897（阶段律）；Dunitz, J. D. & Bernstein, J., "Disappearing polymorphs", *Accounts of Chemical Research*, 1995□Bučar, D.-K., Lancaster, R. W. & Bernstein, J., "Disappearing Polymorphs Revisited", *Angewandte Chemie International Edition*, 54 卷, 2015□Bauer, J. et al., "Ritonavir: An extraordinary example of conformational polymorphism", *Pharmaceutical Research*, 18 卷, 859-866 页, 2001；复现见 *PNAS*, 121 卷, e2319127121, 2024 DOI: 10.1073/pnas.2319127121)。教科书：Bernstein, J., *Polymorphism in Molecular Crystals*, 2nd ed., Oxford University Press, 2020□
4. 第十六节所列各位：Burke, K., *Permanence and Change*, 1935□Whorf, B. L., 1940□Bartlett, F. C., *Remembering*, 1932□Freud, S., 1915□Elbow, P., *Writing Without Teachers*, Oxford University Press, 1973□Rose, M., *College Composition and Communication*, 31 卷 4 期 1980□Perl, S., *Research in the Teaching of English*, 13 卷, 1979□Sommers, N., *College Composition and Communication*, 31 卷 4 期 1980□Rubin, D. B., *Journal of Educational Psychology*, 66 卷, 1974□Heckman, J. J., *Econometrica*, 47 卷, 1979□Bourdieu, P., *Esquisse d'une théorie de la pratique*, Droz, 1972□
5. 与写作研究里的生态一路（Cooper, M., "The Ecology of Writing", *College English*, 48 卷 4 期 364-375 页, 1986，及其后发展出的 ecocomposition）本文未作分界，因为那一路处理的是写作与社会系统

的互构，全部落在已成形的产出之上，与本文的第四格不相交。此处特意点出，是因为它常被误以为覆盖了“条件如何限制写作”这一问题。

6. 引注诚实声明：上列文献中，卷、期、页码经核对可确认的已写全；少数条目（注 3 中 1995 年与 2015 年两条、注 4 中 Burke□Whorf□Freud□Bartlett□Elbow 五条）本文只能确认作者、题名、出处与年份，页码未经核对，故未填。宁缺不补。
 7. 第十三节全部数据、脚本与预注册文本（含一处修订记录）另存，可复算。
-

字数与版本：正文实测 19,807 汉字（不含本行）。版本：初稿。人机分工声明：选题、三种回答的选取、承重判断、真跑的假设与阈值由作者设定；文献检索、数据抓取、统计与初稿撰写由作者与语言模型协同完成；第十三节的不利结果与第十八节的自我定位由作者保留原样，未作修饰。本文不附任何分数。证据边界：本文提出的读数尚未在任何真实教学场合中试用；第二十节 F2 未执行。在 F2 跑通之前，本文的机制主张是一个候选，不是一个结论。

第六章 补节 在第一笔之前

本章 v1.1 已自行补入三条全球分界（知觉窄化、形态空间的不可达方向、非决策），并跑完证伪条款 F6，此处不重复。补节只做两件事。

（一）与第七章的对切

本章与第七章都指向「不进任何账本的那一格」。分离线在读的是状态还是动作□

本章读的是那一格本身：哪些形式从来没有获准成形，因而不存在任何记录里。它的三条性质（无实体、无痕迹、有边界效应）描述的是一种状态；它的读数是序列读数——从一个写作者的连续产出里定位最后一次出现新形式的位置。它不问是谁排除的，也不需要问：那一格的边界效应，在没有任何代理人的情况下同样成立。

第七章读的是关闭这个动作：某一条本可成立的方向，在写作者遭遇它之前被范文、模板、教师、同伴或机器关掉了。它必须指认一个关闭者与一个时点，否则命题为空；它的判据因此是制度判据。

判决性对照：一个没有任何外部关闭者的写作者，仍然有未生界——他所处的语言、文体与训练史已经把大量形式排除在可及范围之外，而这个排除没有代理人。反过来，一次被清楚指认的预闭，未必落在未生界里——学生本可以在别处遇见那条路，只是这一次没遇见；它仍在他的可及范围内，只是这一回被绕开了。

所以两章是嵌套而非同义：第七章处理有代理人、可追责的那一子类，本章处理包含它的那个大类；而大类的主要性质（不留记录）恰恰使子类的追责在实践中失效。

两条并入条件，分别写进两章的证伪清单：若真实数据显示每一处未生界都能回溯到一个可指名的关闭者，本章应并入第七章；若显示所有预闭都不改变可及集合，第七章应并入本章。

（二）本章欠账

F6 的事后粒度扫描给出了一个对本章不利、也最应当被记住的结果：末次首现读不读得出，完全由构式清单的粗细决定——清单取前一千条与前五百条时，一位作者也读不出；取前五十条时，五十三位可读。本章 v1.1 因此补了一条硬约束：未公布构式清单的读数一律不成立，清单须在清点之前落盘。编者要提醒读者的是，这条约束虽然堵住了事后调参，却也把本章的读数变成了相对于一份人为清单的读数——本章的对象在原理上不可列举，而它的读数在实践中依赖一份可列举的清单。这个张力本章尚未解决。

第七章

路是谁关的

何谓作文？

——作文作为未预闭合意义分叉的个人闭合史

What Is School Composition?

*Composition as a Personal History of Closing Meaningful Forks That
Were Not Pre-Closed*

SDE 学科碰撞创新法 v3.9 • What 型成文稿

文化遗产保护修复 × 催化科学 × 合同法

2026-08-17

摘要

当生成式人工智能能够在数秒内提供题意分析、立意、论据、结构和成篇文字之后，“作文是什么”重新成为一个基础问题。把作文定义为学生亲自写出的文本，无法解释口述写作、辅助技术、教师反馈和合法协作；把作文定义为写作过程，又难以处理一个更隐蔽的事实：完整的草稿、修改和击键记录，只能记下已经进入学生写作过程的事件，无法记录那些在学生遭遇之前便被范文、模板、教师、同伴或人工智能关闭的可能方向。本文提出：作文不是一篇文本，不是若干写作步骤，也不是署名者承担责任的任务成果；作文是一名写作者在特定表达任务中亲自遭遇并关闭一组尚未被外部预先关闭的意义分叉而形成的发生史。本文把“写作者尚未遭遇便被外部关闭的本可成立方向”称为“预闭分叉”。由此，同样全部由学生敲出的两篇文章可能只有一篇发生了作文；含有机器生成文字的作品，也不能仅凭机器文字比例判断作文是否发生。真正需要追问的是：在主张、证据、结构和受众关系发生改变的节点上，哪些可能方向曾经向学习者开放，又是谁使其中一条路消失。这个区分同时改变了作文教学、作文评价和人工智能时代的学术诚信问题。

关键词：作文；预闭分叉；意义生成；写作过程；人工智能；作文评价

Abstract

Generative AI destabilizes a long-standing equation in writing education: the submitted text is no longer reliable evidence that the learner underwent the composing activity attributed to it. Existing responses increasingly move from product inspection toward process ownership, rhetorical decision-making, disclosure, checkpoints, and writing traces. Yet even complete process records capture only events that entered the observable writing process. They do not automatically capture alternatives removed before the learner encountered them. This article calls such absent possibilities pre-closed forks and proposes a stronger ontological distinction: composition is neither the textual artifact, the aggregate writing process, nor the responsibility attached to a submission. It is the writer's personal history of encountering and closing consequential meaning-bearing alternatives that remained open long enough to become the writer's own problem. On this account, fully human-typed prose can fail to instantiate composition when crucial semantic alternatives have been pre-closed by templates, tutors, rubrics, or AI; conversely, external assistance with surface realization does not by itself erase composition. The framework yields falsifiable

predictions for AI-assisted writing, process-based assessment, transfer, and writing pedagogy.

Keywords: composition; pre-closed forks; writing process; rhetorical choice; generative AI; assessment

目录

- 一、我们过去为什么可以不问“作文是什么”
- 二、一件东西保持原样，并不等于它没有被别人动过
- 三、同一个终点，可以是完全不同的发生
- 四、负责到底，也不能证明事情是你做的
- 五、三个答案为什么都差了最后一步
- 六、作文的本体不在文字，而在分叉经历
- 七、一张二维表：最容易被误诊的作文在哪里
- 八、所以，AI 不是最大的风险，模板也可能比 AI 更强
- 九、一次公开政策审计带来的意外结果
- 九之二、把一件读不出的事变成可清点的：留叉率 OFR

本章到这里有一个尚未解决的困难，而且它是本章自找的：预闭分叉按定义不留记录，那还能清点什么？

如果坚持去数「被关掉的路有几条」，本章立刻自相矛盾——那些路正因为没有被写作者遭遇，才不进任何账本；能被数出来的，恰恰是没有被预闭的那些。所以本章的读数不能落在对象上，只能落在对象的补集上，而且必须落在任务这一侧而不是学生那一侧。

定义（留叉率 **OFR**）：给定一次写作任务及其全部随附材料（题目、材料包、范文、提纲模板、教师口头指令、允许调用的机器输出），在一份事先固定的决策结点清单上，写作者到达该结点时仍有两条以上互不等价的可继续方向的结点所占比例。

结点清单固定为四类，每类各二，共八个结点□

主张类——① 中心判断取哪一侧 ② 该判断的强度与限定条件；

~~证据~~——③ 用哪些材料④ 材料判断条件怎么建立；

结构类——⑤ 从哪里起写 ⑥ 转折落在何处；

受众类——⑦ 预期读者是谁 ⑧ 对该读者的哪一种预期起作用。

「互不等价」的判据：两条方向互不等价，当且仅当采用其一而非其二，会迫使清单上至少另一个结点的答案改变。仅措辞不同、仅例子不同、仅顺序不同而不牵动其他结点的，计为同一条方向。这一条是整份读数的承重处——没有它，任何任务都可以宣称自己「留了很多选择」，因为换个词也算一条路。

四条编码规则□

其一，以写作者到达该结点的时刻为准，不以任务发布时刻为准。范文在写前发下、教师在写作中途才补充说明，两者对同一结点的效果完全不同。

其二，「可继续」指在剩余时间与材料条件下真的能写完，不指逻辑上可设想。

其三，教师说「你们可以自由发挥」而随后只示范一种做法的，该结点计为已关闭。示范优先于声明。

其四，编码者只读任务与材料，不读学生的作文。OFR 是任务的属性，不是学生的属性。

OFR $\in [0, 1]$ ，八个结点故最小刻度为 **0.125**。两名编码者独立编码，逐结点一致性低于 0.75 的任务进入退出日志。

判读方向□OFR 高不等于任务好。一份八个结点全开的任务，对初学者可能等于没有任务；本章从不主张 OFR 越高越好。OFR 要回答的是一个更朴素的问题——当我们说一个学生「没有自己思考」时，有多少是因为他没有思考，有多少是因为到他手上时已经没有什么可思考的了。

一条算术型的反直觉推论：一份材料越完备、范文越精彩、提纲越细致、模板越好用，OFR 必然越低。这不是对好教材的指控，而是一个恒等式——每一处「替学生想清楚」都恰好关闭一个结点。因此任何以「降低写作难度」为目标的教学改进，都同时是一次 OFR 的下调；若不同时在别处开回来，作文发生的空间就是被单调压缩的。本卷第八章有一条形状相同的推论（提纲越细，必须自造的段落反而越多），两条合起来说明：教学上的每一次「给得更足」，都在账本的不同栏目上各留下一笔看不见的代价。

九之三、关闭者的六种签名，与各自的误诊阻挡

OFR 只回答「关了几个」，不回答「谁关的」。而本章第六节已经说明，指认关闭者是本章命题不为空的必要条件。以下把关闭者分为六类，每类给出一条可与其他类区分的签名，以及一条误诊阻挡。

一、范文型关闭。签名：全班在同一结点上给出高度相似的答案，而这种相似不能由题目本身的约束解释。误诊阻挡：必须先排除「题目只有一个合理答案」这一竞争解释——办法是让另一批未见范文的学生做同一题，看该结点的答案是否同样集中。

二、模板型关闭。签名：关闭主要落在结构类结点（⑤⑥），而主张类结点（①②）保持开放。模板通常不规定你想什么，只规定你按什么次序把它摆出来。误诊阻挡：若主张类结点也随之集中，那多半不是模板在起作用，而是范文或教师指令。

三、教师指令型关闭。签名：关闭发生在一个可定位的时刻，且该时刻之后的作文与之前的草稿在同一结点上出现方向性改变。误诊阻挡：教师追问与教师给答案在记录上极易混淆——判断是追问改变的是该结点的开放程度还是它的取值：追问使学生看见原本没看见的第二条路（OFR 上升），给答案使一条路被指定（OFR 下降）。

四、同伴型关闭。签名：关闭沿座位、小组或社交关系扩散，呈现聚集而非全班均匀分布。误诊阻挡：需排除同组学生共享同一份材料这一平凡解释。

五、机器型关闭。签名：关闭发生得最早，且往往落在受众类结点（⑦⑧）——机器输出默认一个泛化的结，学生很少改它。误诊阻挡：机参与权重成例，本章第五节已经说明。

六、自我关闭。签名：同一学生不同任务上反复关闭同一类结点，而外部条件各不相同。误诊阻挡：这一类最危险，因为它极易滑向对学生的人格判断。本章的处理是把自我关闭排除在预闭分叉之外——预闭的定义要求关闭者是外部结构；一个学生自己不愿意走某条路，那是他的写作行为，不是他遭遇之前被拿走的東西。若不设这条边界，本章会把一切没有发生的可能性都算成被剥夺，那样这个概念就什么也不排除了。

六类合起来构成本章的可追责范围：前五类可追责，第六类不可，而第六类恰恰是最难与前五类区分的一类——因为一个反复被范文关闭的学生，最终会看起来像是自己不愿意走那条路。这一条本章无解，明写于此。

十、与“选择”“能动性”“作者性”究竟差在哪里

十一、与“决策所有权”和“控制位置”也不是同一件事

十二、与医学上的“过早闭合”也必须分开

十二之二、与政治学的「非决策」正面交手

前三节把本章与「选择」「能动性」「作者性」「决策所有权」「控制位置」「过早闭合」逐一分开，这些都在写作学与临床心理学的范围内。但本章的论证形式在政治学里另有一位正主，而且它比上述任何一位都更接近。

二十世纪六十年代，政治学界提出「权力的第二张面孔」：权力不仅体现在公开冲突中谁赢了，更体现在哪些议题根本没能进入议程。这种把潜在议题在成为议题之前就排除掉的能力被称为非决策，它的关键特征是——它不会在任何会议记录里留下痕迹，因为按定义它使冲突从未浮现。此后又有人提出「第三张面孔」：权力最深的一层，是使人们连自己有什么利益都不再感知。

本章的论证形式与之逐条同构：有一类可能方向在进入写作者的问题空间之前被关闭；关闭动作不留记录；因此从现有记录里读不出它曾经存在。连两者面对的最强反驳都完全相同——如果它不留痕迹，你凭什么说它存在？

那一支给出的回答是：可以从被压制方的抱怨、从系统性缺席的诉求、从「谁本会从这条路存在中获益」反推。本章不能照搬这个回答，原因恰恰构成了两者的分离线。

分离线：那一支处理利益冲突，本章处理意义生成。非决策理论的全部操作化都依赖一个利益结构——潜在议题对某方有利、对另一方不利，因此可以从受益格局反推缺席。本章不需要这个结构，而且通常没有它：关闭一条论证路径的教师，与学生的利益完全一致，而且往往正是出于好意。一份写得极好的范文，比一份糟糕的范文更能预闭分叉。

由此得到一条判决性对照：非决策问「谁从这条路不存在中获益」，本章问「谁使这条路在他遇见之前消失」。前者若找不到受益人，命题即告失败；后者可以没有受益人，甚至可以所有人都受损——这正是教育场景与政治场景最根本的差别，也是本章不能被归入非决策理论的教育学应用的理由。

但这条差别对本章不全是好消息。没有受益人，意味着本章失去了那一支唯一可用的反推工具。非决策理论至少可以从利益格局出发去找缺席，本章连这个

出发点都没有——这正是本章的读数只能落在任务侧（见第九之二节）而不能落在被关掉的路上的根本原因。放弃利益结构换来了适用范围，代价是放弃了唯一现成的反推路径。这个交换本章认，并写在这里。

十三、一个反直觉结论：AI 写了一些字，不一定比“老师给了唯一立意”更远离作文

十四、这不是主张“AI 怎么用都行”

十五、作文教学真正要保护的并不是“困难”，而是未关闭的问题

十六、这会改变我们怎样看待“范文”

十七、这也会改变我们怎样看待教师反馈

十八、一个可以立刻拿去用的判据

十九、一个新的研究对象：作文里没有发生的東西

十九之二、与本卷第六章的对切：读状态，还是读动作

本卷第六章与本章指向同一个地方——账本上没有那一栏。两章不合并的理由必须写清楚，否则它们互为改名。

第六章读的是那一格本身。哪些形式从来没有获准成形，因而不在任何记录里。它的三条性质——无实体、无痕迹、有边界效应——描述的是一种状态；它的读数是序列读数，从一个写作者的连续产出里定位最后一次出现新形式的位置。它不问是谁排除的，也不需要问。

本章读的是关闭这个动作。某一条本可成立的方向，在写作者遭遇它之前被范文、模板、教师、同伴或机器关掉了。本章必须指认一个关闭者与一个时点，否则命题为空；它的读数因此落在任务与材料上（第九之二节的 OFR），而不落在写作者的产出序列上。

两条判决性对照：

其一，一个没有任何外部关闭者的写作者，仍然有未生界。他所处的语言、文体、训练史已经把大量形式排除在可及范围之外，而这个排除没有代理人，也没有时点——它不是被谁关掉的，它从来就没有开过。按本章的定义，这里没有预闭分叉（找不到关闭者）；按第六章的定义，这里有边界（可及范围确实被划定了）。

其二，一次被清楚指认的预闭，未必落在未生界里。教师在这道题上给了唯一立意，那条被绕开的路仍在这个学生的可及范围之内——他在别的场合遇见过类似的判断，只是这一次没遇见。按本章，这是一次标准的预闭分叉；按第六章，这条路从未离开可及集合，因而不属于未生界。

所以两章是嵌套而非同义：本章处理有代理人、有时点、可追责的那一类，第六章处理包含它的那个大类。而两者之间有一层更不舒服的关系——大类的主要性质（不留记录）恰恰使子类的追责在实践中失效：即使某次预闭确有关闭者，被关掉的那条路也不会任何账本上留下可供对照的痕迹，因此追责只能落在任务材料的事前审计上（OFR），而不能落在事后的追查上。

两条并入条件，各写进两章的证伪清单：若真实数据显示每一处未生界都能回溯到一个可指名的关闭者，第六章应并入本章；若显示所有预闭都不改变可及集合，本章应并入第六章。

十九之三、与本卷第三章的接口：账本在两个方向上同时出错

本章与本卷第三章之间有一条接口，两边原稿都没有写出来，而它比两章各自的结论都更强。

第三章的结论是：一份看似极强的整体性读数，其中绝大部分只需要改动句加一条邻句就能还原；不读作文、只看施加了哪一类操作，就能预测七成以上的判断结果。也就是说，账本上记着的那部分，会系统性地高估自己的解释力□

本章的结论是：关闭动作不进入成品、草稿、击键记录或责任账本。也就是说，账本上没记的那部分，不会因为记录变得更详尽而自动出现□

两条相加，得到一条对整个「过程数据越多越好」的思路不利的判断：越是依赖过程记录来判定作文发生，越容易在两个方向上同时出错——一边把局部可还原的东西当成整体能力的证据，一边对从未进入记录的关闭视而不见。而这两种错误的方向是相反的：前者高估，后者遗漏，因此它们不会互相抵消，只会一起把判断推向「记录得越多，看起来越确定」。

这也是本章第二十一节所说反噬的根源：学校若把「分叉」也做成表格，最可能的结果不是关闭减少，而是账本变厚、确定感上升、而被关掉的东西一样多。

二十、六条可以把本文推翻的经验条件

二十之二、一条对本章不利的现成读数

按本卷的纪律，不利结果原样保留。以下这一条对本章不利，且它不是设想，是本卷第六章在补入非决策一条分界时已经报告的一个现成读数。

那一节报告：在没有任何受益者的场合（私人写作），边界的封闭反而更晚□

这个方向对非决策理论是相容的——没有受益人，压制的动机就弱，议程封闭自然更慢。但它对本章的一种更强读法不利。本章若被读成「凡是封闭必有代理人在起作用」，那么私人写作与公开写作的封闭速度应当接近，因为语言与文体的约束在两种场合并无二致；实际数据显示两者有系统差别，这说明封闭速度里有相当一部分由有无受益者解释，而这一部分恰恰是本章明确排除掉的（见第十二之二节：本章的关闭可以没有受益人）。

两种可能的处置，本章都不隐藏：

其一，收缩本章的适用范围——预闭分叉主要发生在有他人在场、有评价预期的写作场合，而不是一切写作场合。这样做会使本章更弱，但更可检验。

其二，承认本章与非决策理论在这一点上并非互斥，而是共同作用：有受益人时两种机制叠加，无受益人时只剩本章这一种。这样做保住了适用范围，代价是本章必须给出一个能把两种机制分开的实验，而目前没有。

本章选择第一种。理由是：第二种在当前证据下无法与「本章那部分效应为零」相区分，而一个在原理上分不出自己是否为零的主张，不该被保留。相应

地，第二十节的六条死亡条件应增列第七条：若在完全私人的、无任何评价预期的写作中，仍能观察到与公开写作同等程度的、可指认关闭者的方向消失，则本章关于「关闭者」的界定过宽，需要重写。

二十一、最危险的反噬：学校会不会把“分叉”也做成表格

二十二、重新回答：何谓作文

二十三、AI 时代真正需要保卫的东西

二十四、结论：作文不是留下来的字，是没有被提前拿走的路

参考文献

研究边界与人机分工声明

写死日期的经验赌注

附录 A：v3.9 工序审计

一、我们过去为什么可以不问“作文是什么”

很长时间里，“作文”这个词似乎不需要定义。

老师布置一道题，学生写一篇文章，老师拿到文章评分。于是三个本来不同的东西自然重叠：写文章的人、经历写作的人、交文章的人通常是同一个人。正因为三者高度重叠，我们可以把“学生交来的文章”直接叫作“学生作文”，而不必追问这种等同究竟凭什么成立。

写作理论早已知道，作文不能只被理解为成品。20 世纪六七十年代以后，写作研究逐渐把注意力从完成后的文本转向写作过程。Rohman 与 Wlecke 的早期研究已经把写作前的发现过程置于中心；Murray 明确反对只教成品；Emig 把写作论证为一种独特的学习方式；Bereiter 与 Scardamalia 后来进一步区分知识陈述与知识转化，强调成熟写作会反过来改变写作者的理解。

因此，今天再说“作文不是文章，而是过程”，并不新。

再向前一步，说作文意味着学生拥有选择权、能动性、修辞决策权，也已经不新。Leibman 甚至直接把 choice 当作研究学生修辞能动性的分析单位；Kaufer 与 Wetzol 的设计取向把写作者对决策的承担列为写作的重要特征。

生成式人工智能到来以后，这个方向被进一步制度化。2026 年 NCTE 的《The Students' Right to Write》明确把选择、控制、写作过程及对最终作品的责任放进学生写作者的权利与责任框架；CCC 的现行生成式 AI 政策把作者身份与人的探究、解释和修辞决策联系起来；AP Capstone 不只要求作品是学生自己的，还要求学生亲自完成分析、综合与表达选择，并通过中途检查点让思考和

决策显露出来；IB 同样要求教师见证作品的发展，并要求学生能够解释自己的作品。

因此，我们面前真正困难的问题已经不是：“AI 写了多少，才不算作文？”甚至也不是：“最后决定是不是学生做的？”这些问题都已经有人在回答。

真正还没有被充分区分的是另一件事：一个本来应该成为学生写作问题的可能方向，如果在学生知道它存在以前就已经被替他关掉，那一次“没有发生的选择”算什么？

这才是本文的问题。

二、一件东西保持原样，并不等于它没有被别人动过

文化遗产保护提供了第一个必要的拆分。

在最朴素的想象里，“真实”似乎意味着保留原材料：原来的砖、原来的木头、原来的颜料越多，真实性越高。然而现代保护理论早已不能如此简单处理对象的身份。

ICOM-CC 1984 年关于保护修复职业的定义，一方面强调历史对象的物质完整性、文献性质以及辨别原有与后来介入的重要性，另一方面也明确承认修复是对受损对象的主动处理。1994 年《奈良真实性文件》又进一步指出，对真实性的判断可能依赖形式与设计、材料、功能、传统与技术、位置、环境乃至精神感受等多种信息来源，而不可能由一种固定标准在所有文化中裁决。

更尖锐的是现实案例。奈良的一些历史建筑经历长期修缮和构件替换；春日大社甚至保有周期性重建传统，却仍能在材料、形式、技术和场所连续性中维持高度真实性评价。当代保护领域内部对此仍有激烈争论：保存原初物质、允许重建、保存意图还是维护历史连续性，并没有被压成一个单一答案。

它对作文问题造成的第一处破坏是：

“不是学生亲手生成的材料进入作品”与“作品不再属于学生的发生”不是同一件事。

教师改一个错别字，语音转文字软件替有运动障碍的学生把声音变成字符，同伴指出一个论据漏洞，机器替学生修正标点——这些介入都改变了最终对象，但介入本身并不足以证明作文已经转移给别人。

因此，把“机器生成字数占比”直接当作作文发生程度，理论上过于粗糙。遗产保护迫使我们承认：材料来源与对象身份可以分离。

可是，仅仅把答案改成“看过程”仍然不够。

三、同一个终点，可以是完全不同的发生

催化科学给出了第二刀。

IUPAC 对催化剂的标准定义强调：催化剂改变反应速率，却不改变反应总体的标准 Gibbs 能变化；在严格描述中，催化剂同时进入反应循环又被再生。

这里最重要的不是“催化剂像老师”。那只是类比。真正有用的是催化科学已经发展出的选择性概念。对于同一组反应物，可能同时存在多条竞争反应；催化选择性关心的不是“有没有产物”，而是不同路线以什么相对速率展开、最终哪些产物获得优势。

终点相同，不代表发生相同。

更重要的是 IUPAC 对 promoter 的定义：促进剂可以增强希望发生的反应，也可以压制不希望发生的反应。这一句把问题打开了。

假设两名学生最后都写出了同样的中心论点。甲在写作中先后形成过三个可能立意，对证据做过取舍，发现第一种解释与材料冲突，放弃；第二种解释无法回应受众，又放弃；最后才形成第三种。乙在动笔前收到一份高度优化的“立意分析”，其中已经把前两条路排除了，只留下第三条。他没有复制任何一句话。之后的所有文字也确实都是自己敲出来的。

从最终文本看，两者可能一样。从 AI 字数检测看，乙甚至是百分之百人类文本。从击键记录看，乙也完全真实。从“最后决定是不是本人做的”看，乙当然可以说“是我决定采用这个立意”。

然而两个人经历的并不是同一种写作。区别不在于甲多做了两次“错误选择”。区别在于：

甲曾经生活在一个拥有三条可继续道路的问题里；乙从进入问题的

第一刻起，就生活在一个只剩一条路的问题里。

这两者之间少掉的东西，在最终文章里没有痕迹。它甚至不会出现在修改记录里，因为它从未成为一次修改。

四、负责到底，也不能证明事情是你做的

第三个误区是责任。

人工智能时代一个很自然的补救方案是：只要学生对作品负责、能够解释、能够答辩、最终决定由他承担，就仍可以称作他的写作。这个方案比查“AI 率”高明得多，但合同法恰好给出一个反例。

以美国合同法中清楚的委托履行规则为例，UCC § 2-210 允许在相当多的情形下由第三者履行义务；关键的是，把履行委托出去，并不会自动解除原当事人的义务和违约责任。同时，如果另一方对原承诺人亲自履行或控制履行具有实质利益，委托又可以受到限制。

普通法中的 personal services 又把边界说明得更清楚：某些服务因高度依赖特定个人的技能、判断和能力而难以由替代者完成。

一个人完全可以继续承担最终责任，同时把真正的履行过程委托出去。

因此，“我对最后文章负责”不能单独证明“作文由我发生”。“我看过并同意 AI 给出的论点”也不能证明这一点。“如果文章有问题，我承担后果”仍不能证明这一点。

责任可以连续，而发生已经迁移。这就是为什么仅仅把 AI 时代的作者性定义为 accountability，仍然不够。

五、三个答案为什么都差了最后一步

现在我们得到三个极强、同时又互相不服的判断。

第一种判断说：要看对象保留了怎样可信的身份连续性。介入不等于失真。第二种判断说：不能从终点倒推发生；必须看竞争路径怎样被改变。第三种判断说：不能从责任倒推履行；责任人和实际履行者可以分离。

它们争的其实是同一件事：一份提交文本凭什么算“这个学生的作文”？然而三种回答背后还藏着同一个更深的假设。

无论检查成品属性、实际过程，还是责任关系，我们都默认：决定身份的那件事已经在某个地方发生，因此原则上总能找到它。文章里有痕迹。草稿里有痕迹。按键里有痕迹。提示词里有痕迹。答辩里有痕迹。责任声明里有痕迹。

现在催化选择性给出了反例：决定结果的力量，可以体现为一条没有发生的竞争路线。一条路线因为被提前压制，所以没有产物，没有动作，没有错误，也没有修改。

于是我们第一次得到一个写作过程研究很难直接看见的对象：预闭分叉。

所谓预闭分叉，是指：一个本来足以改变文章意义走向的备选方向，在写作者真正遭遇它之前，就被外部的人、规则、模板、范文、算法或工具从其可继续空间中排除。

它不是“学生选择以后放弃的选项”。那种选项已经进入他的写作。它也不是通常意义上的“过早下结论”。过早下结论研究的是一个人已经面对若干假设之后，自己太快停止搜索；医学中的 premature closure 就属于这种情形。

预闭分叉更早。写作者甚至还没有获得犯这个错误的机会。别人先替他把问题变小了。

六、作文的本体不在文字，而在分叉经历

由此可以重新回答最初的问题。

作文不是学生交出的一篇文章。

因为同一篇文章可以由不同发生史产生。

作文不是写作行为的总和。

因为一个人可以亲手输入每个字符，经历完整的“列提纲—写初稿—修改—润色”，却始终只在一条别人预先铺好的轨道上行动。

作文也不是署名者承担责任的成果。

因为责任可以保留，而关键履行可以委托。

那么作文是什么？

作文，是写作者在一个表达任务中，亲自遭遇并关闭一组尚未被外部预先关闭的意义分叉而形成的个人发生史。

这里有三个限定。第一，必须是意义分叉。“这个逗号用中文还是英文格式”“标题用黑体还是宋体”当然也是选择，但它们通常不足以改变这篇文章在说什么、凭什么说、对谁说、怎样组织其理由。真正构成作文的分叉至少集中在四类节点：主张、证据、结构、受众。

第二，必须是遭遇。学生不一定要独自发明所有候选方向。老师完全可以提出一个反例。同伴可以提出第二种解释。AI 也可以列出三个可能立意。关键不在“选项是谁发明的”，而在于这些选项有没有真正进入学生需要辨别、比较、拒绝和承担后果的问题空间。

第三，必须由写作者发生闭合。闭合不是点击“接受”。闭合意味着：一个原来仍可能继续的方向，因为写作者形成了某种可追问的理由，而不再继续。

所以作文最小的单位不是句子，甚至也不是一次修改，而是：一个方向曾经可以继续，后来由于这个写作者的判断不再继续。

这才是发生。

七、一张二维表：最容易被误诊的作文在哪里

由此，我们可以把两个长期混在一起的问题拆开：第一轴问“表层语言主要由谁实现？”第二轴问“关键意义分叉在学生遭遇前是否已经被关闭？”

	关键分叉向学生保持开放	关键分叉在学生遭遇前被关闭
表层文字主要由学生产生	传统作文：学生既处理意义分叉，也完成语言实现	表层自写／发生缺席：每个字都是学生写的，但中心路径早已被替他
表层文字含显著外部实现	辅助实现型作文：语言实现得到语音转写、翻译、编辑或 AI 帮助，但意义分叉仍由学生遭遇和关闭	代理成文：意义路径和语言实现主要由外部完成

真正值得警惕的不是右下角。右下角太容易看见。真正危险的是右上角。

学生没有复制。没有粘贴 AI 文字。AI 检测器显示“人类写作”。击键记录非常完整。甚至每一稿都确实是他亲自修改。

但是老师提前给了唯一立意，范文预先确定了唯一论证结构，评分量规把“高分文章”的可接受路线收缩成一种，AI 又在学生阅读材料以前把“最值得使用的三个证据”筛好。于是学生完成了大量真实劳动，却几乎没有遇见这篇文章最重要的问题。

这种情况过去很难命名，因为所有传统证据都在。缺掉的是一批从未进入记录的可能性。

八、所以，AI 不是最大的风险，模板也可能比 AI 更强

这条判断如果只用来反对人工智能，就浪费了它。

一个“完全禁用 AI”的课堂照样可能把作文预闭得所剩无几。老师给出标准立意。教材给出固定段式。范文给出每段功能。评分表奖励可预测的结构。考前训练告诉学生哪一类素材放在哪一段。

学生当然亲手写了。但如果一个本来开放的表达问题在他动笔前已经被压缩成“把已有零件装进去”，那么最值得教育的那部分作文可能早已发生在学生之外。

这也是为什么“写作过程”本身不能成为免检通行证。

当代击键记录技术已经能够非常细地捕捉停顿、输入、删除、修改以及文本怎样随时间演变，并被用于推断写作过程中的行为和认知活动。

但记录设备有一道原则性的边界：它能记录发生过什么，却不能自动记录一个从未进入写作者视野的方向。

这正是预闭分叉与普通“过程数据”的分界。

九、一次公开政策审计带来的意外结果

为了检验这个判断是否只是把现有 AI 写作政策说了一遍，我在形成最终定义之前先设定了一次小型编码。

我检查了 12 份来自专业写作组织、国际考试机构及大学的现行官方页面，包括 NCTE、CCC、AP、IB、Yale、Cornell、University of Washington、Oxford、UCL、Cambridge、University of Sydney 与 Stanford PWR。

原先的预期是：很多制度主要看最终文本和 AI 使用披露。这个预期被实际材料部分推翻了。

AP 已经要求在研究、论证结构等阶段做检查点，让学生把思考与决策显露出来。IB 要求教师在较长时间内看到作品发展，并要求学生能够解释作品。Cornell 甚至提供完整 AI 交互记录和流程图式的记录办法，以便跟踪学生的工作和决策过程。Sydney 则建议保存 AI 输入输出，并允许课程要求提交交互日志。Stanford PWR 对发明、研究、修改和组织决策的界限也规定得相当细。

所以，“现在学校还只会查 AI 文本比例”这句话不成立，必须删掉。

然而编码也留下了另一个稳定的空白。这些制度非常擅长记录：用了什么工具；产生了什么文本；学生做了什么；学生是否理解；谁承担责任；草稿怎样发展；提示词和输出是什么。

但在我检查的这些公开页面里，没有发现一个专门记录下面这件事的字段：

在学生第一次形成主张、选证据、定结构以前，哪些本来可成立的方向已经因为外部帮助而不再进入他的考虑范围？

这是一个更难的问题。因为想记录它，不能只保存“学生做过什么”。还得记录：学生本来有机会遇见什么，却没有遇见。

这正是本文判断经过真实材料以后被迫缩窄的地方。

十、与“选择”“能动性”“作者性”究竟差在哪里

最接近本文的已有理论之一，是修辞能动性研究。Leibman 已经把选择作为分析单位，研究学生怎样理解自己的写作决策。NCTE 也明确把 choice、control 和 agency 写进学生写作者的权利。

如果本文只是说“学生应该自己作选择”，它可以立即被这些理论吸收。

两者真正的分离线是：能动性研究主要能够处理进入行动者选择场的选择；预闭分叉研究处理的是没有进入行动者选择场的方向。

可以设计一个判决性对照。让两名学生都拥有最终决定权，都报告很高的自主感，都亲手写出最后文章。甲先独立提出三个解释，再与 AI 对话。乙先让 AI 给出一个高度可信的最佳解释，然后围绕这个解释自己写。

如果只问“最后是谁决定”，两个人都可以回答：我。如果只问“有没有能动性”，两个人也都进行过大量选择。

本文则预测：在后续不给 AI 的结构变形题里，甲比乙更可能主动生成新的解释空间。如果没有这项差异，本文就输给一般能动性理论。

十一、与“决策所有权”和“控制位置”也不是同一件事

人工智能伦理已经出现“decision ownership”问题：当 AI 提供建议、信息或推荐时，一个决定中的价值判断到底还能否归属于人。关于 AI 辅助作者身份的研究又进一步提出 locus-of-control problem，分析人在总体结构控制和机器在具体语言生成之间如何分配控制。

这同样非常接近，但仍有一个可裁决差别。

想象一个 AI 系统只做一件事：它不给你代写，也绝不替你点“接受”。它只是根据大量高分作文，在你开始写以前把“最没希望的路线”过滤掉，只给你展示三个高概率选项。

最终全部决定仍然是你的。控制位置仍然在你。责任也仍然在你。然而你的问题空间已经被改变。

所以本文关注的并不是：谁拥有现存选择的控制权？而是：谁决定了哪些东西有资格成为你的选择？

前者是 decision ownership。后者才是预闭分叉。

十二、与医学上的“过早闭合”也必须分开

医学推理中的 premature closure 已经是成熟概念，指诊断者太早接受一种假设，于是停止继续考虑其他合理解释。近期实验研究仍在直接检验这种机制。

本文不能把这个经典概念换一个名字。两者的判决性差别非常简单：

医学式过早闭合：行动者看到了一个答案，然后自己过早停止搜寻。

本文所谓预闭：外部系统在行动者看到备选答案以前就完成了筛选，因此行动者没有发生“停止搜寻”这个动作。

若实验发现，只要控制住一般的 premature closure 倾向，所谓预闭条件不再产生任何独立影响，那么本文的机制解释应被放弃，回到一般认知偏差即可。

十三、一个反直觉结论：AI 写了一些字，不一定比“老师给了唯一立意”更远离作文

这会产生一个很不舒服、却必须面对的结论。

假设学生因为手部障碍使用语音识别。字符不是他敲出来的。没有理由因此说作文没有发生。

假设学生英语不是第一语言。他已经自己形成了主张、材料关系和文章结构，然后用允许使用的语言辅助工具修正若干语法。字符层面有外部介入，也不能因此自动说作文没有发生。

反过来，学生不用任何 AI，只接受一篇“万能结构”、一套唯一立意以及教师筛好的材料，最后亲手写两千字。这篇文章在传统原创性检测里异常干净。

但从本文的区分看，它反而可能失去更多作文发生。

因此，人工智能真正逼迫作文教育做的，不是寻找一条“人工百分比”的安全线。它迫使我们承认：

作文和文字生产从来不是同一个东西，只是过去经常由同一个人同时完成。

十四、这不是主张“AI 怎么用都行”

这里必须主动收回一个很诱人的过度推论。

如果一名学生仍然亲自处理了关键意义分叉，是不是就意味着任何 AI 使用都应被允许？不是。

合同关系提供了明确限制。不同任务有不同的履行要求。有些任务本身就在考察独立语言生成，课程设置者可以对原学习者亲自履行具有实质利益。UCC 关于委托履行的规则本身也承认：当对方对原承诺人亲自履行或控制履行具有实质利益时，委托不是当然允许的。

因此必须区分两句话：第一句：有没有发生作文？第二句：这次作文是否符合这门课程的提交规则？

本文只重新定义第一句。它不自动推翻第二句。

一篇作品完全可能保留了丰富的作文发生，却因为课程明令禁止某项 AI 使用而不得提交。反过来，一篇作品也可能百分之百合规，却因为任务结构过度预制，只留下极薄的作文发生。

这两件事以后不应继续混在一起。

十五、作文教学真正要保护的**不是“困难”，而是未关闭的问题**

于是还有一个常见误解必须排除。本文不是鼓励“让学生越痛苦越好”。作文发生不等于困难。老师完全可以帮助学生。

好帮助甚至可能增加学生面对的有效分叉。例如学生只想到两个极端立场，教师补进一个反例，使第三条路线出现。同伴指出某个证据还支持另一种解释，于是原本单线的论证重新分叉。AI 一次生成三种彼此真正冲突的论点，学生逐项核证后拒绝其中两个，也可能增加可处理的意义空间。

相反，一种看似高要求的教学也可能高度预闭。它可以要求学生写三稿、五稿、七稿，却在第一稿前就把所有重要判断确定完毕。

所以需要保护的**不是“学生有没有受苦”，而是：**

在需要形成判断的时刻，问题有没有保持开放到足以成为学生自己的问题。

十六、这会改变我们怎样看待“范文”

范文最值得重新研究。传统争论经常问：范文会不会限制创造力？模仿有没有价值？学生该不该学习优秀结构？

本文提供了另一个问题：不是问范文“好不好”，而要问范文在什么时候进入。

如果学生还没有形成自己的问题空间，范文直接给出一个完整答案，它容易预闭。

如果学生已经走到一个具体困境——例如两个证据无法同时进入同一逻辑结构——此时再给不同范文，让学生比较别人怎样处理类似冲突，范文反而可能重新打开分叉。

所以同一篇范文可以在两个时刻产生相反的教学功能。决定作用的不是范文本身，是它进入学生问题空间的时间。

十七、这也会改变我们怎样看待教师反馈

同一句反馈：“这一段论证太散，最好先提出中心句。”可能有两种完全不同的作用。

第一种教师直接把中心句写好，学生照着改。第二种教师指出：“这一段现在至少可能围绕 A、B 两个中心组织，你删掉其中一个以后，另外三个材料会怎样？”

表面上看，第二种反馈更慢。它甚至可能导致学生最后仍然选择教师原先偏好的 A。

但发生已经不同。第一种把分叉关闭在教师身上。第二种让分叉抵达学生。

因此，好的反馈不必永远少给答案。它真正要避免的是：

把答案送达学生以前，顺手把本该由学生处理的问题一并拿走。

十八、一个可以立刻拿去用的判据

如果学校、教师或研究者希望判断某次写作中发生了什么，不必先计算 AI 字数。可以问一条更直接的问题：

在主张、证据、结构、受众四类节点中，记录里各出现过哪两个互不等价的继续方向；其中一个方向是在学生接触前还是接触后消失？

这句话不要求老师评价“真正原创”“有意义”“深度参与”。只清点事件。

例如主张层：学生是否曾经面对两个都能解释材料、却导致不同结论的方向？证据层：是否有两项材料都可以进入，但最后因某种理由舍弃其中一项？结构层：是否存在两种会改变读者理解顺序的组织方式？受众层：是否曾经因为预期读者不同而改变解释、例证或语气策略？

这里特别要记录一种答案：“没有第二个方向进入。”过去这句话通常被当作没有数据。本文恰恰认为，它可能是最重要的数据。

下一问才是：为什么没有？

十九、一个新的研究对象：作文里没有发生的东西

至此，作文研究的对象需要扩展。

过去我们研究：写了什么。删了什么。改了什么。停顿多久。怎样规划。怎样修订。怎样与他人协作。怎样调用 AI。这些都是已经发生的事件。

现在还需要研究一类缺席事件：本来可能成为学习者问题，却在学习者遭遇以前被移走的方向。

它们不是普通意义上的“遗漏”。遗漏往往意味着学生应该想到却没想到。预闭分叉没有这样的责备。学生可能根本没有条件想到，因为课程、工具、界面或帮助方式先替他定义了什么值得考虑。

因此它首先是一个任务结构属性，其次才可能成为学生能力问题。

这一点具有伦理意义。如果一个学校把所有预闭都归罪于学生“没有主动思考”，它实际上把教学设计造成的问题个人化了。

二十、六条可以把本文推翻的经验条件

一个新定义如果只能解释，不能失败，就没有太大价值。本文至少愿意接受六种失败。

- 第一，在控制学生原有写作能力、任务难度、AI 使用量、最后文本质量和自我报告能动性之后，如果“关键备选方向在学生遭遇前被关闭”与后续独立写作中的迁移能力没有关系，那么预闭分叉没有独立解释力。
- 第二，把学生随机分为两组：一组在写作前得到单一高度优化的立意—证据—结构方案；另一组得到三套真正互斥的候选方案并必须比较。如果两组在之后无辅助的新情境中生成备选解释、调整结构的能力没有差异，本文的机制预测失败。
- 第三，如果记录大量“学生自己拒绝的备选方向”不能在最终作文质量之外额外预测远迁移能力，那么分叉经历可能只是写得好的伴随现象。
- 第四，如果一个拥有完整草稿链和击键历史、但关键路线被预设的群体，与关键路线保持开放的群体，在新题型中表现完全相同，则本文过度解释。
- 第五，如果一般修辞能动性、决策所有权或 premature closure 指标能够完全吸收上述差异，那么没有必要另立预闭分叉。
- 第六，如果未来能够证明最终文本本身可靠地恢复“哪些可能方向在学生遭遇以前被排除”，那么本文关于成品不可识别性的核心主张应当被撤回。

本文尤其押一个顺序预测：

在反复依赖预闭帮助时，最先下降的应不是文章当次得分，而是一次无帮助任务中自主生成备选方向的数量与异质性。

表面质量甚至可能先上升。如果实际顺序相反——先出现质量下降，而备选生成能力长期不变——本文设想的机制就有问题。

二十一、最危险的反噬：学校会不会把“分叉”也做成表格

一旦本文的判断被制度采用，最省事的实现方式几乎可以预见。

老师要求：“每篇作文必须提交三个立意。”学生于是机械生成三个立意。AI 更擅长，一秒钟生成十个。评分表新增：“备选方案数量：20 分。”

于是学生拥有更多“分叉记录”，却没有更多问题。更糟的是，为了证明自己亲自思考，学生开始事后伪造被拒绝的方案。

原本要保护的发生，被转化成发生的证明材料。证明材料又取代发生本身。

这是本文最严重的反噬。因此预闭分叉不能被直接拿来给学生排名。它首先应当诊断任务、教学支架和工具设计的位置，而不是诊断某个学生“作文能力高低”。

它问的是：“我们在什么地方替学习者把路关早了？”不是：“这个学生有多少分叉？”

如果未来它变成新的考核 KPI，我暂时看不到一个彻底可靠的制度出口。

二十二、重新回答：何谓作文

到这里，“何谓作文”可以有一个比“文章”“过程”“自主表达”更窄也更强的答案。

作文当然留下文本。作文当然需要过程。作文当然涉及选择、责任和作者性。但这些都还不是它的最小发生条件。

作文最小的发生不是“写出一句话”。而是：

一个表达方向原本仍可以继续；写作者遭遇了它与另一方向之间的冲突；由于写作者自己的判断，其中一条路在这里结束。

于是，作文是这种事件的连续历史。

好的作文不一定拥有最多分叉。成熟写作者常常能迅速排除许多低价值路线。但迅速排除和从未遭遇不是同一件事。

一个专家可以看一眼就判断第二条论证不可行；那仍然是他的判断。一个外部系统在他看到第二条论证以前把它删掉，则不是。

这就是“作文发生”与“文章被完成”之间最细而最关键的界线。

二十三、AI 时代真正需要保卫的东西

生成式人工智能最深刻的影响，或许不是它开始写字。机器替人写字只是最显眼的一层。

更深的一层是，它越来越擅长在我们还没有形成问题以前，先替我们完成：选题排序；立场筛选；材料排序；反例淘汰；结构优化；受众适配。

未来最强的写作 AI 甚至可能不替你写一句话。它只需要不断把“最不值得走的路”藏起来。

然后你会非常高效。每句话都是你写的。每个按钮都是你点的。每个最终决定都是你作的。你甚至感觉自己拥有完全的控制。

但有一天，你可能发现一个更难描述的损失：你已经很久没有遇到一个足以迫使自己改变方向的写作问题。

如果发生这种情况，真正消失的并不是作者署名，而是作文。

二十四、结论：作文不是留下来的字，是没有被提前拿走的路

因此，本文最后只留下一个判断：

作文不是学生交出的文本，不是写作步骤的总和，也不是署名者承担责任的任务成果。作文是写作者在表达任务中亲自遭遇并关闭一组尚未被外部预先关闭的意义分叉而形成的个人发生史。

从这里看，一篇满是学生自己文字的文章，可能没有发生多少作文。一篇借助外部语言工具实现的文章，也可能保存丰富的作文发生。教师反馈可以促进作文，也可以替代作文。范文可以打开问题，也可以提前关闭问题。人工智能可以夺走作文，也可以制造以前没有出现过的新分叉。

真正的界线，不在“用了没有”。甚至不在“谁按下最后的确认键”。而在：

当问题仍有几条路可以走的时候，那几条路有没有活到足以成为这个写作者自己的问题。

这也许是人工智能时代重新理解“作文”最值得保住的一点：

教育不是确保最后留下正确的路。教育首先要确保，有些路没有在学习者抵达以前就被拿走。

参考文献

- [1] Bereiter, C., & Scardamalia, M. The Psychology of Written Composition. 1987. [\[在线来源\]](#)
- [2] Emig, J. “Writing as a Mode of Learning.” College Composition and Communication, 1977. [\[在线来源\]](#)
- [3] ICOM Committee for Conservation. The Conservator-Restorer: A Definition of the Profession. 1984. [\[在线来源\]](#)
- [4] IUPAC. “Catalyst” ; “Selectivity” ; “Promoter.” Compendium of Chemical Terminology. [\[在线来源\]](#)
- [5] Leibman, C. C. Is This What You Wanted?: Expectations, Choice, and Rhetorical Agency in Composition. 2019. [\[在线来源\]](#)
- [6] National Council of Teachers of English. The Students’ Right to Write. 2026. [\[在线来源\]](#)
- [7] Conference on College Composition and Communication. Generative AI Policy. 2025. [\[在线来源\]](#)
- [8] Rohman, D. G., & Wlecke, A. O. Pre-Writing: The Construction and Application of Models for Concept Formation in Writing. 1964. [\[在线来源\]](#)
- [9] UNESCO. Nara Document on Authenticity. 1994. [\[在线来源\]](#)
- [10] Uniform Commercial Code § 2-210, Delegation of Performance and Assignment of Rights. [\[在线来源\]](#)
- [11] Cornell Legal Information Institute. Personal Services. [\[在线来源\]](#)
- [12] Tian, Y., & Cushing, S. T. “Exploring the Application of Keystroke Logging Techniques to Research in L2 Writing.” 2024. [\[在线来源\]](#)
- [13] College Board. AP Capstone Diploma Program Policies. [\[在线来源\]](#)
- [14] International Baccalaureate. Guidance on artificial intelligence in assessment. [\[在线来源\]](#)
- [15] Cornell Center for Teaching Innovation. AI Attribution Guidelines. [\[在线来源\]](#)
- [16] University of Sydney. Using Artificial Intelligence Tools Responsibly. 2026. [\[在线来源\]](#)
- [17] Stanford Program in Writing and Rhetoric. Policy on the Use of GenAI and LLMs. [\[在线来源\]](#)

研究边界与人机分工声明

本文关于文化遗产保护、催化科学、合同法以及现行人工智能写作政策的事实性材料来自公开文献和官方页面。合同法部分主要借用美国法中公开且判据清楚的委托履行材料，不主张其具体规则自动适用于所有法域。

本文提出的“预闭分叉”及其用于定义作文的理论判断属于本文的推论，而不是上述任何来源原有主张。它尤其不构成对现行学校学术诚信规范的替代。

本轮研究的问题方向与采用的三个外部学科由人类用户确定；资料检索、跨领域比较、候选淘汰、公开政策审计以及本文初稿由人工智能辅助完成。所有拟用于正式发表的文献条目、页码、引文和经验判断仍需在投稿前由作者逐项复核。

写死日期的经验赌注

截至 2028 年 8 月 17 日，如果作文教育与人工智能写作评价确实沿本文指出的问题发展，那么至少一个具有较大影响力的写作教育、考试或高校评估体系应开始要求学习者留下某种“被拒绝的实质性备选方向”，而不仅是提交 AI 使用声明、提示词、草稿历史或最终反思。

以下情况不算命中：仅要求“说明你的思考过程”；仅要求 AI 使用日志；仅要求多个草稿；仅要求学生声明“最终决定由我作出”；仅要求列出三个形式上不同但不改变意义方向的选项。

只有当制度要求在至少一个影响主张、证据、结构或受众关系的节点上，留下两个以上互不等价的可继续方向，并说明其中一个为何以及何时退出，才算命中。

若到该日期，成熟的过程评价体系能够稳定验证学生作文发生，却完全不需要这类信息，或者经验研究证明预闭程度不能在一般能动性、写作能力与 AI 使用量之外预测任何独立迁移结果，那么本文关于“预闭分叉”的核心判断应当被收缩或放弃。

附录 A：v3.9 工序审计

本附录用于保存本轮理论制造过程的可追溯记录；正文已去除内部方法术语。审计表不构成创新认证分。

工序	结算	关键说明
-1 题型判别	What 型	问题要求“何谓作文／作文是

		么”，目标是制造一个能重新划分文／非作文”的辨别维度。
0.5 源未吸收度	暂判放行	保护修复真实性判据、催化途中的竞争路径、合同履行委托与作论之间未发现固定吸收关系；判的题目答案而非学科远近。
0.5 理论厚度	三家均厚	三家均具长期理论史、内部争反例传统和权威手册／标准。
位置三分	S／D／E 分开	S：作品身份如何显露；D：经过何种竞争路径发生；E：履行在何任关系中成立。
可引原句阐	通过	三家均以 ICOM／UNESCO、IUPAC、UCC/LII 权威文本作为核心判断依据，未用代拟理数。
候选近邻阐	淘汰一批近邻	“本人责任／本人决策／过程权”已被 NCTE、CCC、AP、IB 等接占位，故不作为 Z。
敌意拓宽	完成	检索经典过程写作、写作能动性、AI 决策所有权、locus of control、击键记录、premature closure 等层。
共有前提	找到	三家均默认关键事实已经发生留下可查对象：成品属性、实际路责任关系。
内部推翻材料	来自催化科学	选择性可由“竞争路径被压制定；未发生的路线也能决定最终结
最终新存在物	预闭分叉	本来会成为写作者问题的意义叉，在写作者遭遇前被外部结构关因此其缺席不会自动进入成品、草稿击键或责任账本。

机制证伪	六条	明确给出控制替代解释后的可条件，并设置迁移顺序预测。
真跑一条	12 份公开政策审计	不利结果：多套制度已跟踪过检查点，故“只看成品”被删；仍现专门记录“学生遭遇前已消失的方向”的制度字段。
反向约束	至少两处	保护修复迫使本文放弃“外援真”；合同法迫使本文区分“作文发生”与“课程是否允许”。
自反与反噬	认账	预闭分叉若被做成学生 KPI，发机械生成备选项和事后伪造；正确反对用其直接排名学生。
最终评审	不自评	按 v3.9 铁律，参与写作的基为该文本发认证分。

文档版本：v1.0 • 2026-08-17

第七章 补节 路是谁关的

（一）祖宗与对切已在正文中处理，此处只记编者的两点

本章v1.1 已在第十二之二节与政治学的「非决策／第三张面孔」正面交手，并在第十九之二节写出与本卷第六章的对切。编者只补两点。

其一，本章那条分离线的代价被诚实地写出来了，值得指出：非决策理论的一切操作化都依赖利益结构——从「谁本会从这条路存在中获益」反推缺席；本章放弃了这个结构（教师关闭一条路时往往与学生利益一致，而且是出于好意），因此也放弃了那一支唯一现成的反推工具。本章的读数只能落在任务侧而非被关掉的路上，根源就在这里。用适用范围换掉反推路径，这个交换本章认了，并写在正文里。

其二，v1.1 补入的第二十之二节，是本卷最彻底的一次自我收缩。第六章报告的一个现成读数（私人写作场合边界封闭反而更晚）对本章的强读法不利；本章列出两种处置，选了更弱的那一种——把适用范围收缩到有他人在场、有评价预期的写作，理由是另一种处置在当前证据下无法与「本章那部分效应为零」相区分，而一个在原理上分不出自己是否为零的主张不该被保留。相应地第二十节增列了第七条死亡条件。编者认为这一段的处理方式，比本卷任何一处的方法论宣言都更能说明本卷想要什么。

（二）本章原先最大的结构缺陷已补

编者说明：v1.0 的本章没有任何读数，而本卷导论明写「每章必须给出一个读数」——这是九章中唯一一处与全书纪律直接冲突的地方。

v1.1 的第九之二节补上了这个缺口，且补法本身有一处值得记录的判断：本章的读数不能落在对象上。去数「被关掉的路有几条」会当场自相矛盾——那些路正因为没有被写作者遭遇才不进任何账本，能数出来的恰恰是没有被预闭的。因此读数改落在对象的补集上，并且落在任务侧而非学生侧：留叉率 OFR = 在八个固定决策结点上，写作者到达时仍有两条以上互不等价方向的结点比例。

其中「互不等价」的判据（采用其一而非其二会迫使至少另一个结点的答案改变）是整份读数的承重处；没有它，任何任务都能宣称自己留了很多选择。而「示范优先于声明」这条编码规则，把「你们可以自由发挥」这类话排除在开放之外，是本读数最可能被规避之处的一道堵口。

第九之三节的六类关闭者签名同样重要，其中第六类「自我关闭」被明确排除在预闭分叉之外——预闭的定义要求关闭者是外部结构。本章自陈这条边界不设就会「把一切没有发生的可能性都算成被剥夺」，而设了之后又留下一个无解处：一个反复被范文关闭的学生，最终会看起来像是自己不愿意走那条路。编者同意这里目前无解，并认为它是本章最值得后续研究的一处。

（三）本章欠账

其一，OFR 一个数也没有算过。它比本卷其余读数都更容易跑——只需要一批真实的作文任务与随附材料，不需要学生、不需要写作过程、不需要伦理审批，两名编码者一个下午即可完成一批。本章下一步最该做的就是这件事。

其二，本章的参考文献全部以在线来源形式给出，未落到卷期页码。本卷其余各章的关键判断句均可核到原文出处，本章是唯一的例外，成书前应补齐。第十二之二节新引的非决策与第三张面孔两条同样只给了年代与提法，未给出处。

第八章

两端都不记账的那一段

悬跨段

——作文的发生落在两处有据点之间：那一段必须自造的文字不属于任何一端，因此在提纲、成品与评价里都没有位置

Unbridged Spans: Composition as the Self-Made Text Between Two Sourced Points

摘要

作文教学与写作研究长期把“写不下去”当作一个能被指认到具体位置的事件：这一句想不出、这一段不会写、这个论点站不住。本文主张，这个默认本身是三门互不相干的手艺共同持有而从未说出口的假定，而其中一门在自己的操作体例里早就否认了它。合成化学发现产物比例由出口处的能垒差决定，与手里握有多少中间体无关；《周易》的变占体例在多爻同变时并不从任何一爻自身读出“当动之爻”，而要以天地之数减去六爻之数的余额来指定；火在何处停下不取决于哪一棵树是否可燃，而取决于两处可燃物之间的距离。三者合起来只能得出一条三者都不会同意的判断：作文的发生不落在任何一处，而落在两处有据点之间那一段必须由写作者自费铺过去的文字上——本文称之为悬跨段。它的读数是悬跨长度 λ ：从最后一处有据可指的位置到下一处有据可指的位置之间，连续自造的句数。判据是一句不含程度词的问话：这一句的上一处处处在第几句之前？本文对 60 篇英文维基百科特色条目做了一次预注册检验（哈希 dae6976a53469b50，23,509 次修订、3,294 个会话），三条假设被自己的数据全部驳回，且同一份数据在两种都合理的编码下给出方向相反的系数；这一不利结果被写入正文，并把主张由“停在哪里由间隙属性决定”修订为更弱也更硬的一条：间隙属性在现有记录里不可识别，因为记录只有点。事后清点显示，在每句都被要求给出来源的最高质量档条目里，最长连续无出处句串的中位数仍达 13.5 句，而没有任何评审程序读这个数。结论：一切“加点”的写作指导——提纲更细、分点更多、模板更全——在数学上必然增加交接次数，因而增加悬跨段的数目；这是本文最不愿意但必须给出的预言。

关键词：悬跨段；悬跨长度；有据点；交接；写作过程；不可识别

Abstract. Research and teaching on writing routinely assume that a breakdown in composing can be located at a point: this sentence, this paragraph, this claim. This paper argues that the assumption is a shared and unstated premise of three unrelated crafts, and that one of them has already denied it within its own operating rules. In synthetic chemistry, product ratios are governed by the free-energy difference between the transition states at the exits, not by the populations of the intermediates in hand. In the divination procedure of the *Zhouyi*, when several lines change at once, the line that "should move" is not read off any single line but assigned by a remainder computed from the whole hexagram. In fire ecology, where a fire stops is not determined by which tree fails to burn but by the distance between patches of fuel. Together they force a judgment none of them would accept: composing happens not at any point but in the **unbridged span** between two sourced points — the stretch of text a writer must lay down at their

own expense, which belongs to neither end. Its readout is the **span length λ** : the number of consecutive self-made sentences between the last and next attributable position. Its criterion is one question containing no modal terms: *how many sentences back is the previous source for this sentence?* A pre-registered test on 60 English Wikipedia featured articles (hash dae6976a53469b50; 23,509 revisions, 3,294 sessions) refuted all three hypotheses, and the same data yielded coefficients of opposite sign under two defensible codings. That adverse result is reported here and forces a weaker but harder claim: gap properties are **unidentifiable** in existing records, because records contain only points. A post-hoc count shows that even in the highest-quality tier — where every claim is supposed to carry a source — the median longest uncited run is 13.5 sentences, and no review procedure reads that number.

Keywords: unbridged span; span length; sourced point; handover; writing process; unidentifiability

一·一个所有人都见过、而没有一个字段记着的现象

先摆现象，不急解释。

一个学生交上来一份提纲：五个论点，每个论点下面两条材料，条条清楚。三天后他交上来的作文只有六百字，而且是从第三个论点直接跳到第五个的。你问他卡在哪儿，他会说“第四点我不太会写”。你把第四点单独拎出来问他，他讲得很好——讲了两分钟，逻辑完整，例子也有。你说那你写下来。他写了两句，停了。

一个博士生把文献读完了，笔记做了六万字，每一条都带页码。他动笔写第二章，写到第七页停住。他不是不知道下一句该说什么——你逼问他，他能说出下一句的大意。他停住的地方是：上一句结束于某本书的第 214 页，下一句要落到另一本书的第 31 页，而从 214 页走到 31 页这一段路，两本书都没有写。

一个部门主管要写一份年度报告。数据齐了，图表做好了，去年的模板还在。他写了一整个下午，写出来的是图表的说明文字。所有需要把两张图连起来的段落，他都留了空，标着“待补”。年底交上去的报告里，那些“待补”变成了三句过渡话，而没有任何人在评审时指出这三句话没有依据——因为它们本来就不需要依据，它们只是过渡。

三件事看上去毫不相干。它们共同的形状是：两端都清楚，中间那一段没人认领。

而我们的全部处置方式是把它记在某一端身上。学生说“第四点不会写”，我们于是给他补第四点的材料；博士生说“第二章卡住了”，我们于是让他多读几篇；主管说“报告不好写”，我们于是给他一个更细的模板。三种处置全部指向点，因为我们的账本上只有点。

本文要立的是：那一段路是一个东西，它有长度，有代价，有归属问题，而目前没有任何一种写作记录为它设位。它不是“过渡”——过渡是修辞学的说法，指两个已有部分之间的衔接手法；它也不是“卡壳”——卡壳是心理学的说法，指一次可观察的停顿。它是一段必须由写作者自费产出、而任何一端都不为其负责的文字。本文叫它悬跨段。

路线图：第二节请三门与写作毫不相干的手艺各自出场，说它们如何回答“这一程怎么走完”；第三节指出它们共同没说出口的那条假定；第四节用其中一门自己的体例把这条假定推翻；第五到第八节给出悬跨段的定义、它的三种来路、它的读数与清点办法、以及一句可以拿去问记录的问话；第九到第十一节报告一次预注册检验的全过程，包括它把本文驳回的部分；第十二到第十五节是与最近的几种既有说法的分界、被三门手艺反过来加上的约束、以及本文不适用的边界；第十六到第十八节是证伪条件、一条写死日期的赌注、以及本文自己在自己的框架里落在哪一格。

二·三门手艺各自怎么把一程走完

二·一 合成化学：手里有什么，不决定出来什么

有机合成里有一条被写进物理有机化学标准综述的原理：当两种可以快速互变的构象各自不可逆地走向不同产物时，产物的比例只由通往两个产物的过渡态之间的自由能差决定，而不必然反映两种构象自身的平衡分布（Seeman 1983）。它的反直觉推论是教科书标准案例：较不稳定、含量更少的那个构象，完全可以给出主产物——只要它那一侧的出口更低。

对做合成的人来说，这条原理每天在用。它的实际含义是：“我手里有多少”和“我最后能拿到什么”是两笔账。一个中间体在瓶子里占九成，不保证它出现在产物里；一个只占一成的构象，可能贡献九成的产品。你要改变产物比例，不能靠“多做一点那个中间体”，只能靠改出口——换催化剂、换温度、换配体，把某一侧的坎降下去。

还有一条与它配套的日课：保护基。为了让分子上某一处此刻不反应，合成化学家先给它戴上一个基团，做完再把它摘掉。这意味着两件事：第一，“不动”是被主动施加的，不是免费的默认；第二，它的代价被明码记在收率里——上保护基一步、脱保护基一步，两步都要损失。

化学在这一程上的终点很清楚：最后瓶子里那个东西。路线再漂亮，谱图上没有它就是零。

二·二 《周易》的筮法：哪一处当动，不由那一处自己说了算

《系辞上》记着一套成卦程序：大衍之数五十，其用四十有九；分而为二以象两，挂一以象三，揲之以四以象四时，归奇于扚以象闰。四营为一变，三变成一爻，十八变成一卦（朱熹《周易本义·筮仪》定其式）。

这套程序的结果不是一个卦，是两个。每一爻算出来是九、八、七、六四个数之一：九是老阳，六是老阴，老者变；七是少阳，八是少阴，少者不变。把该变的变过去，得到第二个卦，叫“之卦”。占断落在哪里，取决于有几爻变——六爻皆不变看本卦辞，一爻变看那一爻的爻辞，六爻皆变看之卦（乾坤两卦另有用九、用六）。

要紧的不是吉凶怎么断，是“哪一处当动”这件事是被程序标记出来的，不是被占者挑出来的。少阳与少阴不是不重要，是这一次的过程没有把它们标记为可动。一个位置的能动性，在《易》这里是位点属性：这一爻是九还是七，写在它自己身上。

《易》在这一程上的终点也很清楚：那条走向。本卦是已经显露的样子，之卦是要去的地方，而占的产物是"该往哪动"，不是"现在是什么"。

二·三 火与燃料：火在哪里停，不取决于哪一棵树

林火科学最反常识的一条是：火消耗燃料，因此火本身会为后来的火造出屏障——烧过的地方在若干年内挡得住下一场火，火制度因此具有自我调节的性质（Parks 等 2015）。这条被四个大样地的数据检验过，并且给出了一个明确的失效线：随着两场火之间间隔年数增加，旧火疤的阻挡作用逐年衰减，跌到某一概率阈值以下，它就不再是一条有效的防火带。

与它配套的是那条著名的悖论：一个世纪的全面扑火，本意是消灭火，实际结果是把燃料一年年攒起来，最终把小火换成了巨灾火（Arno & Brown 1991；Pyne 1982）。扑火的日课，正是灾难的病根。

对火来说，火线停在哪里，几乎不由某一棵树是否可燃决定。一棵不燃的树只让火烧一下；真正让火停下的，是两块可燃物之间的距离超过了它能跳过去的那个尺度。责任不在点上，在点与点之间。

火在这一程上的终点是第三种：被改造过的那片地。计划烧除的产物不是这一场火，是一片以后只允许某种火通过的燃料床。

二·四 三门手艺的三个终点

把三家的终点摆在一起，本文这道题的形状就出来了：

	它认为这一程的终点是	它据此回答"怎么"
合成化学	实际拿到的那个东西	改出口，不改库
筮法与变占	该往哪动的那条走向	让过程去标记哪
火与燃料	被改造过的那片地	管连续性，不管

三个终点互不相容，谁也不能把另两个当手段收编：把化学的产物当手段，合成就没有意义；把《易》的之卦当手段，占就成了预测工具；把火生态的燃料床当手段，管理就退回扑火。它们不是三种角度，是三个不能同时当终点的终点。

三·三家共同没说出口的那一条

三门手艺争的其实是同一件事："这一步该不该动、能不能走过去"，由谁来裁。化学说由出口的坎裁；《易》说由成卦程序的老少标记裁；火生态说由料的连续性裁。

三个答案互斥。但把它们摆在一起念一遍，会听见一句谁都没说、而谁都点头的话：

| 走不通的责任，落在某一处身上；"卡在哪里"总能指认到一个点。

念给三家听，三家都会说"这还用说吗"。化学的账本上是中间体、产物、官能团——全是物种，全是点；《易》的账本上是六爻——六个位置，每个位置一个

数；火生态的账本上是林分、斑块、样地——全是有边界的地块。三家的记账单位全部是点，因此“哪一处”这个问法在三家那里都不需要辩护。

这条假定的代价是隐形的：只要账本只有点，那么两点之间的那一段就永远只能作为“关系”出现，而关系不被当成一个可以有存量、有代价、有归属的东西。你可以说 A 与 B 之间“衔接得好”或“衔接得不好”，但你不会问“这一段有多长”“它归谁”“是谁把它铺出来的”“铺它花了多少”。

这不是三家的疏忽。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。化学要的是可分离、可称量、可入谱的物种，一段“过程中的路”称不出重量；《易》要的是可占断的位次，一处“两爻之间”没有辞可系；火生态要的是可制图的地块，一段“两块料之间的空”在遥感影像上只是背景。三家各自的核心问题，恰好使这一步成为不必要的。

四·推翻它的那件材料，来自《易》自己

现在需要一件材料，把“卡住的地方总能指到点”这条假定推翻。这件材料必须是三家之一自己已经掌握、已经写进体例、只是从没往这个方向用的东西——从外面搬一个理由进来，得到的只是第四种意见，而不是取消这场争论。

它在《易》这一边，而且就写在变占体例里。

宋以后的变占之法处理一个麻烦：几爻同时变的时候，看哪一条辞？一种通行的处置是求“宜变之爻”——以天地之数五十有五，减去所占六爻之数的总和，所得之数指向的那一爻，即为当动之爻；若它恰好是变爻，就看它，否则按变爻数目另有分例。

请注意这一步做了什么：当动之爻，不由任何一爻自身的老少读出，而由整卦六爻之数与一个总数之间的余额算出来。

也就是说，《易》在自己的操作里已经承认：“哪一处当动”不总是位点属性。当变的地方多于一个，位点属性立刻失效，必须转而使用一个从整体差额里算出来的、不属于任何一爻的量去指定它。这个量没有爻位，没有辞，在六十四卦的任何一条辞里都找不到它——它只在算的那一刻存在。

这件材料在易学内部回答的是另一个问题：多爻变时的读辞体例。它从来没有被当作关于“卡在哪里”的证据。这正是它能推翻那条共有假定的原因——它不是为反对而造的，它是《易》自己每天在用的东西。

另外两家各自也有一件同形的材料，可以作为旁证，但它们的位置不如上面这一条硬：

- 化学：过渡态不是任何一个物种。它在任何谱图上都没有丰度，不能被分离，不能被称量，而产物比例恰恰只由它决定。合成化学每天靠一个在自己的账本上没有条目的东西做决策。
- 火生态：火线的位置不在任何一棵树上；同一场火的边界之内还留着未烧的斑块，而“未烧”既不是“不可燃”，也不是“没点着”——它是火没跳过去。

三件材料指着同一处：决定这一程走不走得完的那个东西，在三家的账本上都没有条目。

五·悬跨段：它是什么，以及它不是什么

定义：在一次写作里，有据点是文本中可以指出外部依据的位置——一条引文、一份数据、一段被记录过的事实、一处提纲里已经写下的条目、一句作者此前已经写定并可直接取用的话。相邻两个有据点之间那一段必须由写作者当场自造、且两端的依据都不覆盖的文字，就是一个悬跨段。

四条界定，缺一条这个词就滑回旧概念：

1. 它以"段"计，不以"点"计。它没有位置，只有两端和长度。问"悬跨段在第几句"是问错的问题，正如问"火线在哪一棵树上"。
2. 它必须被跨过去，不能被绕开。一篇文章如果绕开了它，那不是跨过去了，是这篇文章到此为止——现实中它表现为跳过一个论点、砍掉一节、把两段生硬地并排放着。
3. 它的代价由写作者独自承担，而没有任何一端为它记账。前一处依据只保证到它自己为止；后一处依据只从它自己开始算。中间那一段，两边的出处都不覆盖。
4. 它不是空白，也不是错误。悬跨段被跨过去之后，留下的是读起来完全正常的文字——过渡句、概括句、推论句、"由此可见"。成品上看不出哪一段是自造的，这正是它难以被发现的原因。

它不是下面这几样（详细分界见第十二节）：

- 不是衔接（cohesion）。衔接讲的是既有两部分之间用什么手段连起来，它假定两部分都已经在那里；悬跨段讲的是中间那一段本身还不存在，必须先被产出。
- 不是停顿。停顿是时间上的一个事件，可以用键击日志测；悬跨段是文本上的一段长度，一个人可以毫不停顿地一口气写完一个很长的悬跨段——写得越流利，越难发现。
- 不是知识缺口。知识缺口在点上：这一处我不知道。悬跨段的两端都可以是知道的。
- 不是过渡段。过渡段是修辞的功能单位，是作者有意安排的；悬跨段是结构上的必然产物，作者往往并不知道自己在写它。

为什么它是一个东西，而不是"距离"这么中性的说法：距离是两点的函数，取消两点就没有距离，而距离本身不需要谁去负责。悬跨段不同——它需要有人把它写出来，写它的人要付出代价，而没有任何一方为这份代价设位。删掉本文这个说法，它不是变得难测；它是在任何账本上都不存在：提纲里没有它（提纲只有条目），成品里没有它（成品只有句子），评价表里没有它（评价表只有论点、结构、语言、书写），过程研究里也没有它（过程研究测的是停顿、修改、回读，全是事件）。

这是本文全部主张的落点。接下来三节把它变成可以清点的东西。

六·一次"写不下去"有三种来路，从外面看一模一样

不用二维格子。这里要的是一张三分表，因为三种来路是并列的，不是两根轴交叉出来的。

	甲·点上无料	乙·点上有料而出口高	丙·两端无人负责
当事人怎么报告	"这里我不知道写什么"	"我知道要说什么，就是写不出那句话"	"这里写"——与甲相同
实际情况	该处确实没有可指的依据	依据在，语言形式过不去	两端依然那一段没有
旁人一看	缺材料	缺表达	看起来
通行处方	多读、补材料	练句式、找范文	多读、(诊)
处方有效吗	有效	有效	无效，补进来的材料据点，于是
它留下的痕迹	空白或跳过	涂改、反复重写同一句	一段读的过渡文字

三条签名，用来判定手上这一例到底是不是丙：

- 1. 短期指标是改善的。 丙被跨过去之后，篇幅增加、结构完整、通篇流畅——所有能被快速检查的东西都变好了。
- 2. 失效不产生任何信号。 跨错了（自造的那一段其实与两端不相容）不会报错，因为两端的依据都不覆盖它，没有任何程序会去核它。
- 3. 它自我加固。 越是把提纲写细、把分点写多、把模板写全，交接处就越多；每个交接处都是一个悬跨段的两端，于是这套本意在减少困难的做法，在数量上必然增加困难的处数。

第三条签名不是修辞，是算术：一篇文章若有 n 个有据点，交接处就有 n-1 个。把提纲从 5 条加密到 15 条，交接处从 4 个变成 14 个。每一次"讲得更清楚一点"的努力，都在增加需要自费铺过去的段数。 这是本文最不愿意但必须给出的那条预言，第十七节会把它写死。

为什么靶子是丙而不是甲：甲不需要任何框架就看得见——缺什么补什么，这是常识。丙在所有形式检查上都合格：字数够、结构齐、句子通顺、逻辑词齐备。一格东西如果不用本文的说法也一眼看得见，那说明这张表选错了。

七·悬跨长度 λ ：清点手册

一个说法如果不能被清点，它就只是一个比喻。下面这份手册是给要去数它的人用的，也是给要来驳倒本文的人用的——没有它，本文的一切证伪条件都不可裁决。

读数名：悬跨长度	符号： λ
定义：在一段文本中，相邻两个有据点之间连续自造的句数；一篇文章的 λ 取其最大值，	
另报中位数 (λ_{50}) 与分布。N 的取法：以句为单位，不以字数、不以段落。	
粒度：什么算“一句”——以句末标点（。！？；及英文．！？）切分，	
引文内部的句末标点不切；短于 8 个汉字（或 15 个字符）的片段并入前一句。	
边界例：一个只含“由此可见，”的分句——不单独计为一句，并入其后那一句。	
边界例：一个句子里同时含依据与自造内容（“据甲书，X；因此 Y”）	
——	
计为有据点， λ 在此处清零。此约定使 λ 系统性偏小，是保守方向。	
编码规则（六条）：	
1. 有据点 = 该句携带可指的外部依据：脚注、文内引注、数据表引用、可核的事实陈述（有主体、时间、地点、可被第三方查证）。	
2. 提纲条目与作者此前已写定的段落，计为有据点；仅存于作者记忆中的想法不计。	
3. 纯格式句（小标题、编号、图表题注）不计入句数，也不清零 λ 。	
4. 引语、转述与摘要视为有据点；对引语的评论视为自造。	
5. 一个 λ 段跨越小标题时不中断——小标题不是依据。	
6. 无法判定的句，一律计为有据点（保守方向：使 λ 偏小）。	
表头（可直接拿去做空表）：	
文章 ID	句序 句文（前 20 字） 是否有据点(0/1) 依据类型 当前 λ 计数 备注
不适用：口头表达；对话；诗；以及任何“每一句都不要有依据”的文体（ λ 在其中恒等于全文句数，无分辨力）。	
本篇三处引用一致性：正文 §七 = 证伪 §十六 = 赌注 §十六末，同一定义、同一粒度、	
同一份编码规则；阈值均取 $\lambda \geq 10$ 句。	

λ 的两条性质，决定它值不值得造：

第一，它与现行的一切主要指标正交。一篇字数足、结构齐、引注多、评分高的文章，完全可以有很大的 λ ——引注全部堆在事实陈述处，而承担推理与转折的那些句子一条依据也没有。第十节给出的实测正是这样：在每句都被要求给出

来源的那一档写作里， λ 的中位数仍有 13.5 句。举得出“既有指标最好看而这个数最难看”的真实例子，这个读数才不是别的指标的代理。

第二，它只能在留有痕迹的写作里事后清点。成品里读不出哪一段是自造的——这不是测量技术的问题，是第五节第 4 条界定的直接后果。这一条既是 λ 的弱点，也是本文第三层主张的全部内容（第十一节）。

八·一句不含程度词的问话，和它在五个场合的不同答案

不要问“这里写得充分吗”“这个论证有力吗”——这些问题只能拿去问人的判断。要一句可以拿去问流程文件、日志和归档记录的问话：

┆ 这一句的上一处出处，在第几句之前？

它没有“应当”“充分”“实质性”“真正”这类词，答案是一个数。把它拿到五个场合去问，答案必须互不相同；至少有两个答案是查出来的，而不是从本文推出来的——否则这不是判据，是命题的复述。

1. 学生作文（推出来的）：绝大多数句子的答案是“没有出处”。 λ 等于全文句数，判据在此没有分辨力——这正是第七节“不适用”那一栏的来由。
2. 维基百科特色条目（查出来的，见第十节）：中位数意义上，最长连续无出处句串为 13.5 句，96% 的条目 $\lambda \geq 5$ 。这个数没有出现在任何一份评审意见里。
3. 判决书：法律文书的答案高度分层——事实认定部分几乎每句有据（证据编号），说理部分成段无据。同一份文书里 λ 的方差极大，而现行的裁判文书质量评查表里没有这一栏。
4. 病历：客观项（检验值、影像结论）几乎每条有据，而“考虑为……可能性大”这类推断句无据；病历质控查的是项目是否填全，不查两个有据项之间隔了多少推断。
5. 代码仓的提交记录（查出来的，见第九节）：这里的答案反过来改了本文——这份记录里根本没有“依据”这个字段，能取到的只有“改了哪一段”“改了多少字节”“上一次改这一段是什么时候”。于是判据在此问不出答案，只能问它的影子。第十一节说明这件事把本文的主张改成了什么。

第 5 条是本文写作过程中最有价值的一次挫折：判据在一个真实记录上问不出答案，这件事本身是关于那个记录的强主张。

九·一次预注册检验：三条假设，全被自己的数据驳回

证伪条件如果全是纸面的，它只证明作者想清楚了。本文在成文之前跑通了最便宜的那一条，并把不利结果——它占本节的绝大部分——写在这里。

九·一 预注册

在计算任何统计量之前，先把假设、变量、判定阈值写死存档。存档文件的 SHA-256 前十六位为 **dae6976a53469b50**；下述全部内容出自该文件，事后未作改动。

数据：英文维基百科公开接口。样本 = Category:Featured articles 主名字空间成员按排序键升序的前 60 个条目；每篇取最近 500 次修订的元数据（时间戳、用

户、编辑摘要、字节数），不取正文，不取差异。实际取到 60 篇、23,509 次修订。

构造：会话 = 同一条目、同一用户、相邻两次修订间隔不超过 60 分钟的连续串。结果变量 Y = 该次编辑是否为所在会话的最后一次。剔除用户名含 "bot" 者 (1,654 次) 与长度为 1 的会话。得到 3,294 个会话、12,121 个编辑事件，其中可能成为终止步的有 8,827 个。

变量。点属性三个：本次改动的字节数（取对数）、条目当时的规模（取对数）、本次编辑在会话内的序号。间隙属性两个： $G1$ = 本次编辑所落的小节在本条目已取历史中是否首次被触碰（跨到空地）； $G2$ = 本次编辑所落的小节是否与会话内上一次不同（换段）。

三条假设与阈值：

- **H1**：控制点属性后， $G1$ 的系数为正且 $p < 0.05$ （跨到空地更容易停下）。
- **H2**：只用间隙属性的模型，其 AUC 比只用点属性的模型高出至少 0.02。
- **H3**（自火生态移植）：跨到空地之后会话继续的概率低于 0.30。

九·二 结果

H1 被驳。在只含可终止步的口径下， $G1$ 的系数为 **-0.075** (OR 0.928, 95% 自助区间 [-0.240, +0.088], $p \approx 0.385$) ——方向与预测相反且不显著。首次触碰一个新小节，非但没有让人停下，若有差别也是让人更容易继续。

H2 被驳，而且差得很远。五折交叉验证下，只用点属性的模型 AUC = **0.6615**，只用间隙属性的模型 AUC = **0.5354**，差值 **-0.1261**，与预注册要求的 +0.02 反号。在这份记录里，“停在哪一步”由点属性预测得远比由间隙属性好。

H3 被驳。跨到空地之后继续的概率是 **0.680** ($n=731$)，远高于写死的 0.30。从火生态搬过来的那条阈值，在这里根本不成立。

一处必须自曝的口径错误。本文第一次拟合时把会话的第一步也放进了样本。第一步在构造上永远不是终止步，而它的 $G2$ 按定义恒为 0——这一口径让 $G2$ 的系数变成 **+0.677** (OR **1.968**)，读起来是“换段使停下的风险翻倍”，与本文的主张严丝合缝。把样本限制到真正可能终止的步骤之后，同一个系数变成 **-0.275** (OR **0.759**)，方向翻转。

这件事比三条假设被驳更值得写进来：同一份数据、两种都说得通的编码，给出符号相反的结论，而符合作者期待的那一种恰好是错的那一种。若不是预注册文件里写死了“会话长度为 1 者剔除”这一条逼着重新检查分母，本文很可能带着 OR 1.968 这个数字发表，并把它当作核心证据。

九·三 一处事后观察 (**标明为事后，不作证据**)

按“同一小节连续编辑第几步”分层，续行率呈 U 形：第 1 步 0.799、第 2 步 **0.529** (谷底)、第 3 步 0.640、第 4 步 0.704、第 5 步 0.754，其后升到 0.86 附近。谷底在接手之后的第二步，不在换手那一步本身。这个形状与本文相容，但它是事后看出来的，本文不拿它当支撑；它只被写成第十六节的一条待检验条款。

十·第二次清点：在最严格的写作制度里， λ 的中位数是13.5句

第九节测的是“停在哪一步”，用的是过程记录。这一节测的是 λ 本身，用的是成品。这一次清点是事后设计的，不在预注册文件里，因此它不能用来支持第一层主张，只能用来说明这个读数可清点、且不与既有指标重合。

对上述 60 篇条目取当前正文，剔除参考文献、表格、模板与文件行，按第七节的粒度切句，凡携带引注标记（<ref> 或 sfn/harvnb/cite 系模板）的句计为有据点。可解析者 52 篇、**4,388** 句。

- 无引注句占全部句子的比例，中位数 **0.551**（均值 0.557）。
- λ （最长连续无引注句串）的分位数：**10%** 分位 **8.1** 句，**25%** 分位 **9.8** 句，中位数 **13.5** 句，**75%** 分位 **18.3** 句，**90%** 分位 **20.0** 句，最大 **31** 句。
- $\lambda \geq 3$ 的条目占 **100%**， $\lambda \geq 5$ 的占 **96.2%**。

维基百科的特色条目是这套制度里最高的一档：每一条可能被质疑的陈述都被要求给出来源，评审逐条查证，不合格的会被撤销资格。即便如此，一半以上的句子不带引注，而每一篇里都存在一段十来句长、连续无出处的文字。

更要紧的是第二件事：没有任何一份评审意见读这个数。评审查的是“这一句有没有来源”（点），不查“两处来源之间隔了多少句”（段）。一篇条目可以在“每条陈述都有来源”这一项上全票通过，同时含有一段 31 句的悬跨。

三条必须写明的代理局限：引注标记不等于依据（有的句子有依据而未标注）；机械切句会把长句拆碎、把列表并入；模板剥离可能吞掉少数引注形式。这三条都使 λ 的估计偏大的可能性存在，因此上面的数字应读作同口径下的相对分布，而不是绝对真值。可复算：脚本与中间数据随本文存档。

十一·真跑改了本文什么

不利结果如果只是被承认，那是姿态；它必须改掉正文里的某一句话。这里改掉的是本文原本最想说的某一句。

原来打算写的（第一层主张，强版本）：写作在哪里停下，由间隙属性决定，而不由点属性决定。

现在只能写的（第一层主张，修订版）：在现有的一切写作记录里，间隙属性不可识别——不是它的效应为零，是没有字段承载它，因而任何对它的估计都取决于编码约定，而两种都合理的约定会给出符号相反的结果。

三点说明这次修订不是退让，而是把主张换到了更硬的地方：

1. “效应为零”和“不可识别”是可以分开检验的两件事。若间隙属性只是效应弱，那么随着样本增大，估计会收敛到一个小而稳定的值；若它是不可识别的，那么样本再大也不收敛，估计值会随编码约定跳变——这正是第九节观察到的（+0.677 与 -0.275）。这条差别可以被后来的人拿去反驳本文：给出一份对同一批文本的两种合理编码，若 λ 的估计在两者下稳定，本文这一层即被推翻。

2. 修订过的说法更难被现成概念吸收。"停下取决于衔接处"这句话，认知负荷、任务切换代价、篇章边界规划都可以把它收编（第十二节逐条处理）。"这一量在现有记录中不可识别"则不能被它们收编——它们全都预设了那个量是可测的，只是难测。

3. 它把第三层主张顶了上来（第十八节的分层引用中最稳的一层）：多数写作记录里没有"依据的连续性"这个字段。第十节给出的正是这个字段缺失的直接证据——最高一档的评审程序，读得出每一句有没有来源，读不出两处来源之间隔了多远。

最后一句归本文自己：本文原打算用维基百科的编辑史证明悬跨段让人停下；数据说的是在那份记录里根本看不见悬跨段。这两件事之间的距离，正是本文全部剩余价值的所在。

十一之二·那么，一篇作文到底怎么写成

这道题问的是一条路：从哪儿起、经过什么、在哪一步可以插手。前面十一节说的是"路上有一类没人负责的段落"，这一节把它变成能做的事。三条操作，各对应一种插手方式；三种插手方式恰好是三门手艺各自的那一招。

一·开工前列的不是提纲，是有据点清单

提纲列的是要说什么，有据点清单列的是哪里有依据。二者看起来像，实际是两张不同的表。

做法：把手上确实握着的东西一条条摆出来——一段引文、一组数据、一件亲历的事、一段已经写定的话、一句可以核对的事实。每条后面标出它能覆盖到哪里为止。只允许写"我确实有的"，不允许写"我打算论证的"。一个论点如果只在脑子里，它不是有据点，它是一个待跨的空。

然后做这一步，它是整套操作里最便宜也最有用的一步：在相邻两条之间画一个空框，写上你预计要自己造多少句。

一张五条材料的清单，会画出四个空框。这四个框就是这篇文章的全部难处，而在动笔之前它们已经摆在纸上了。多数人写不下去时的痛苦，来自"不知道自己卡在哪"；这一步把它变成"我知道我要跨四段，第二段最长"。

对照着看：加密提纲做的是相反的事——把五条加到十五条，空框从四个变成十四个，每个变短。读起来是"想得更清楚了"，实际是"要自费铺的地方多了十处"。第六节第三条签名说的就是这件事。

二·到了一个空框，只有三条路，选一条，别硬扛

（甲）补料——把中间变成有据。去找一件能落在两端之间的东西：一个例子、一段中间环节的证据、一句别人已经说过的话。找到了，这个空框就被切成两个更短的空框，且两个都短到可以一步跨过去。

代价与限度必须写明：补料是必要而不充分的（第十四节第三条）。加材料等于加库存，而决定你写不写得出的是出口——有没有一种被允许的说法可以承接。所以补料之后要多问一句：这一段如果被人追问"你凭什么这么说"，我打算怎么回答？答得上，跨得过去；答不上，材料再多也只会变成复述。

(乙) 缩跨——把两端拉近。 换一个离得更近的有据点，代价通常是放弃一个漂亮的大跳。很多写作困难其实是野心问题：你想从这本书的结论直接跳到那本书的前提，而中间隔着三个人二十年的争论。缩跨的意思是承认自己这一次跨不了那么远，改走一条更笨但连得上的路。

判断什么时候该缩跨：如果你要自造的那一段超过十句，几乎总该缩跨。这个数字来自第十节的实测——在要求最严的那一档写作里，λ 的中位数是 13.5 句；也就是说，十句以上的连续自造，已经处在最松的那一端。

(丙) 承认——把这一段的性质标出来。 写"以下为推断""此处没有直接证据，只有一个类比""这一步我不确定"。这一条最不像技巧，实际最省力：它把一段无人负责的文字，变成一段声明了归属的文字——归你。

三条路必须选一条。唯一不许的是第四条：装作那里没有空框，用"由此可见""综上所述""因此不难看出"把它糊过去。 这类词是悬跨段最常见的伪装，它们的作用是让读者以为前面已经论证过了。

三 · 三种插手方式，各自来自一门手艺

上面三条路不是并列的经验之谈，它们各自对应一种改法：

插手方式	改的是什么	它来自	在写作
改出口	让某一条路变得可走	合成化学：产物比例由出口的坎决定，不由库存决定	给这一允许的说话承担多重的主张、是推
让过程标记	让哪一处当动被算出来，而不是靠感觉	筮法：当动之爻由程序标记，多处同变时由整体余额指定	开工前后逐句问"上第几句之前己把跨段报
管连续性	不去点火，去管两块料之间的距离	火与燃料：火在哪里停由连续性决定，不由火源决定	不追求开头"，只追据点离得够

三种方式互不替代。只改出口而不管连续性，你会写出一篇每句都敢说、而整体跳来跳去的文章；只管连续性而不改出口，你会写出一篇处处连贯、而全是复述的文章；只让过程标记而不做另外两样，你只会得到一份准确的困难清单。

四 · 走一遍：第一节那个学生

回到开头。学生说"第四点我不太会写"，而他口头讲得很好。

按旧办法：给他补第四点的材料。结果通常是他把材料抄进去，第四点变成一段引文加一句评论，仍然接不上第三点。

按这里的办法，先做第一步——列有据点，不列提纲。他会发现：第三点后面他有一个例子，第四点他有一个理论说法，而这两样之间他什么也没有。空框标出来是“约六句”。

然后选路。补料：找一件把那个例子和那个说法连起来的东西——通常是一个中间层的事实，比如“这类例子在别处也出现过，而且是在同一个条件下”。缩跨：放弃从例子直接跳到理论，改成先从例子跳到一个更小的概括，再从这个概括跳到理论——两跳都短，总长不变，但每一跳都跨得过去。承认：写“这一步只是一个类比，我没有直接的证据”。

三条路他都做得到，而“多读书”他这周做不到。这就是把困难从点上挪到段上的全部实际好处：段上的困难有可操作的处置，点上的困难只有等。

五·给读稿的人：一个五分钟的读法

拿到一篇学生作文、一份下属的报告、一篇待审的稿子，先不通读。做这一件事：

找出全文最长的那一段没有任何依据可指的文字，然后只问一句：这一句的上一处出处，在第几句之前？

这一段几乎总是文章真正的承重处，也几乎总是问题所在——它要么是全文最有价值的推理（作者自己走出来的那一段），要么是全文最空的一段（作者用连接词糊过去的那一段），而这两者在文面上长得一模一样（第六节丙格）。你只需要在那一段旁边写一句话：“这里你打算怎么回答‘你凭什么’？”

比“论证不充分”“过渡生硬”“再深入一点”有用得多——那三句话指的都是点，而问题在段上。

十二·与最近的几种说法的分界

最容易把本文读成旧东西的地方在这里。下面八条，每一条都写四件：那个说法说到哪一步（公允地说，不做稻草人）、分离线落在哪里、以及一条判决性的对照预测——若观察到 X 则本文对而它错；若观察到 Y 则反之。凡是只能写出“本文更强调”“层次更根本”“视角不同”的，一律不算数，那等于承认被覆盖。

（一）停顿与规划（Matsuhashi 1981；Medimorec & Risko 2017）。写作过程研究早已发现，书写中的停顿并非随机散布，而是集中在较高层级的篇章边界上，且停顿越长，随后越可能出现与整体规划有关的修改。这是本文最近的邻居，也是最危险的一位——它同样把注意力放在“边界”上。

分离线有两处。其一，它测的是时间，本文测的是长度：一个写作者可以毫不停顿地一口气写完一段很长的悬跨段，此时停顿数据完全正常，而 λ 很大。其二，它把边界当作既有的（句、句群、段落这些语法与篇章单位），停顿是遇到边界时规划负荷上升的表现；本文说的那一段不是边界，是边界之间尚不存在、必须被产出的文字。

判决性对照：取同一批文本，同时测停顿分布与 λ 。若 λ 大的那些位置正是长停顿所在，则本文被停顿研究吸收；若两者相关系数在 $|r| < 0.2$ 的范围内、而 λ 仍能预测下游的接手失败率，则本文成立而停顿研究管不到这一格。关键在于本文预言存在“流利的悬跨”：写得最顺的那一段，恰恰可能是最长的自造段。

(二) 衔接与连贯 (Halliday & Hasan 1976; Mann & Thompson 1988 的修辞结构理论)。篇章语言学给出了一整套句际关系的分类与标记体系,把文本里"两部分之间"的关系描述得很细。

分离线:它描写的是关系的类型,本文清点的是产出的代价与归属。修辞结构理论可以把一段过渡文字标注为"背景—让步—结论",标注得完全正确,而它不问也不需要问:这段文字的依据来自哪一端?谁为它负责?它有多长?关系分析对一段编造出来的、与两端都不相容的过渡文字,与对一段有据的过渡文字,给出的标注可以完全相同。

判决性对照:取两段修辞结构标注完全一致的过渡文字,一段两端有据、一段是自造。若下游读者、审稿人或接手者对二者的处理没有系统差别,本文无用;若接手者在自造那一段上的接续失败率显著更高,则关系分析漏掉了一个它无法表示的量。

(三) 任务切换代价 (Monsell 2003 一线的实验心理学)。切换任务会带来可测的时间与错误代价,这是几十年的稳定发现。它是本文"交接处最脆弱"这一说法的方法学正主。

分离线:切换代价是主体属性——同一个人在两个任务之间来回,代价随准备时间下降、随练习下降。悬跨段是文本属性——它不随写作者的熟练而消失,因为它不是"换脑子的代价",是"这一段文字没有依据"。

判决性对照:让高度熟练者在充分准备、无时间压力、可任意休息的条件下写作。切换代价预测代价大幅下降甚至消失;本文预测 λ 基本不变。若熟练者的 λ 系统性地小于新手,且小的原因是他们跨得更省,本文这一格被切换代价吸收。

(四) 衔接工作与不可见劳动 (Strauss 1985 一线的工作社会学)。协作研究早就指出:把分散的工作接起来的那部分劳动是真实的劳动,通常不被计入、不被看见、不被计酬。这是本文在形式层上最近的一位——把命题剥掉本学科名词之后,剩下的形状就是"两个被记账的单元之间有一段没被记账的工作"。

分离线两处。其一,它是人际的、组织的(多人分工时把各自的活接起来),本文的是单人一程之内——一个人独自写作时也全额发生。其二,也是更硬的一处:衔接工作是"有人在做而不被看见",本文的悬跨段是"不可识别"——不是被隐藏,是没有任何一份记录含有能确定它的字段,因而不同的记法会给出相反的结论(第十一节)。

判决性对照:衔接工作的处方是可见化——把它写进职责、计入工时,问题就缓解。本文预测可见化会反向恶化:一旦把"过渡"单列为一个要交付的条目,它就变成一个新的有据点,于是它两侧各生出一个新的悬跨段。二者对同一件干预给出相反的评价,这条最锋利。

(五) 结构洞 (Burt 1992)。网络社会学里,两个不相连的簇之间的空隙是一种位置资源,占据它的人可以获利。

分离线:结构洞是一个可被占据的位置,占据者从中获利;悬跨段没有占据者,只有不得不自费铺过去的人,而且他得不到任何计量上的回报。同一个形状,一个是资产,一个是无主的负债。

判决性对照：若把“负责过渡”设成一个可以被人占据并因此获益的角色，结构洞预测有人会去抢这个位置；本文预测这个角色空置——因为它的产出无法被单独计量，一旦计量就退化成一个新条目。观察窗口：任何把“综述人”“串场人”“整合人”设为正式岗位的组织，看这个岗位的实际人员流入与考核结果。

（六）写作过程模型与前写作（Rohman 1965；Emig 1971；Flower & Hayes 1981）。写作被拆成规划、转译、回顾三类反复交替的过程，并明确指出它不是线性阶段。这是本题目所属领域内部用了半个世纪的正主。

分离线：过程模型的全部单位是主体的操作（规划、转译、回顾），它可以把写悬跨段这件事描述为“一次高负荷的转译”，描述得没有错，但它的账本上依然只有“写作者做了什么”，没有“文本上这一段的依据归属”。同一次转译操作，落在有据处与落在悬跨处，在模型里是同一件事。

判决性对照：若把两端的依据补齐（不改变写作者、不改变任务、不改变时间），过程模型预测负荷不变，本文预测困难显著下降。这是一个可以做的实验：同一批写作者，同一题目，一组给出“两端 + 中间衔接处的依据”，一组只给“两端的依据”，比较完成率与自造段长度。

（七）孵化说（Wallas 1926）。困难在于时机未到，放一放就通了。

分离线：孵化说预测时间本身是有效干预；本文预测时间无效而补料于中间有效。这一条落到操作上极便宜：放置一周不做任何补料，与立即补上中间依据，比较同一处的完成率。

（八）自由写作（Elbow 1973）。别管依据，先写下来，写着写着就通了。

（九）warrant 与 enthymeme（Toulmin 1958；亚里士多德《修辞学》）。这是本文最近的一位，也是最应当先划的一位——它就在本题的自家学科里。Toulmin 把论证拆成 data、claim 与 warrant，warrant 是把 data 接到 claim 上的那条许可，且他明写它通常不出场，除非被追问；enthymeme 的传统更直接：省略的那个前提由听者自己补上。两者说的都是「两个已给的位置之间有一样东西没有出处，而它必须被补」。

分离线两处。其一，warrant 是一个命题，enthymeme 省略的是一个前提——它们有真值，可以被写出来、被检验、被反驳；本文说的是一段文字，它有长度、有产出代价，而它是否为真不是本文的问题（第五节第 4 条：悬跨段被跨过去之后留下的是读起来完全正常的文字）。其二，也是更硬的一处：warrant 的默认承担者是听者——enthymeme 的整套效力来自听者愿意补；本文说的这一段的承担者是写者，且没有任何一端为它记账。

判决性对照：取一批被独立评定者判为“warrant 缺失”的位置，与按第七节清点出的 λ 峰值位置，比对二者的重合率。若重合率高于 0.7，则本文只是给 warrant 缺失换了一个以句数计的读数，第一层被 Toulmin 吸收；若低于 0.3，而 λ 峰值仍能预测接手失败率，则二者是两件事。本文据此预言存在一类段落：warrant 齐备而 λ 很大——每一步的许可都写出来了，每一步都没有出处。

（十）discourse synthesis（Spivey 1984；Spivey, Written Communication 7(2), 1990: 256–287；Spivey & King, Reading Research Quarterly 24, 1989）。写作—读源这一支的正主，比第六位的过程模型离本文更近。她把写作者对多个源

文本的加工拆成 **organizing**、**selecting**、**connecting** 三种操作，而 **connecting** 的原话就是：写作者把从源里选出的材料，与自己从既有知识生成的内容编织在一起。本文所描述的现象，她在 1984 年已经指出来了。

分离线：她清点的是操作（写作者做了什么），本文清点的是产物上的一段及其归属（这段文字的依据来自哪一端、有多长、谁为它负责）。同一次 **connecting**，落在两端有据处与落在两端都不覆盖处，在她的编码体系里是同一件事——这与第六条对过程模型的分界是同一处差别，只是她更近，因为她的单位里已经出现了“源”。

判决性对照：按她的编码给一批文本标出全部 **connecting** 操作，同时按第七节清点 λ 。若每一段 λ 都恰好落在一次 **connecting** 上、且 **connecting** 的次数与 λ 的分布可以互推，则本文只是她的一个重新表述；若存在大量不含任何 **connecting** 标记的长 λ 段（作者并不在连接两个源，只是在自己往前走），则那一类正是本文要立的东西。

（十一）维基百科可核性的既有计量（Redi, Fetahu, Morgan & Taraborelli, “Citation Needed: A Taxonomy and Algorithmic Assessment of Wikipedia’s Verifiability”, WWW ’19: 1567–1578）。必须写出来的一位：它用的是与本文第十节同一个语料（英文维基百科特色条目）、同一个单位（句），问的是哪些句需要行内引注、以及为什么需要。本文第十节的清点，是在一块已经有人来过的地上做的，原稿未加标注，此处补正。

分离线是硬的，而且恰好就是本文的全部主张：他们的读数是点的读数——这一句要不要引注、这一句有没有引注；本文的读数是两个有据点之间的距离。他们的工作做得越好，第十节那句话就越站得住：一套能逐句判定“此处需要来源”的系统，仍然不报告两处来源之间隔了多少句。

判决性对照：把他们的逐句判定接到第七节的清点上，就得到一个可以直接跑的东西——以“需要引注却没有引注”的句来定义有据点，重算 λ 。若这样算出的 λ 与本文的 λ 秩相关高于 0.8，则本文的读数可由他们的模型直接导出，不必单列；若低于 0.4，则点的判定确实推不出段的长度。

分离线：本文与它在处方上部分相容而在解释上相反。自由写作确实能跨过悬跨段——因为它取消了“必须有依据”这条约束，让写作者以自造文字直接铺过去。本文的判断是：它跨过去的代价没有消失，只是被推迟到修改阶段，并且以“这一段说不清哪来的”的形式出现。判决性对照：比较自由写作稿与提纲写作稿在同一处的 λ 与事后可核率。自由写作若在完成率更高的同时事后可核率不低，本文这一条被驳。

划完之后剩下的那块共同的地基

逐条问这八位“它把什么当作给定，因此看不见什么”：

- 停顿研究把边界是既有的当作给定，因此看不见“边界之间的文字尚不存在”这件事——一旦承认存在一段不属于任何单位的文字，它的“在边界处停顿”就失去了参照物。

- 篇章关系理论把两部分都已在场当作给定，因此看不见"关系的一端是被现造出来的"——一旦承认，它的关系分类就必须先回答"这一端从哪来"，而它没有这个字段。
- 切换代价把任务是给定的两个当作给定，因此看不见"两个任务之间那一段本身是要被生产的"。
- 衔接工作把劳动有主体当作给定，因此看不见"没有主体、只有默认承担者"的那一类。
- 结构洞把空隙可被占据当作给定，因此看不见"占据它反而使它消失并再生两个"的那一类。
- 过程模型把单位是主体的操作当作给定，因此看不见文本侧的归属。

六条盲区背后是同样被当作给定的东西：

凡是一样东西，都能被指到一个位置上，并且都有人持有。

这八种说法的持有者互相批评了几十年，而没有一位回头看这块地——因为他们都站在上面。本文提出的这一类东西，恰好是从这条地基的缝里掉下去的：它没有位置（只有两端），也没有持有者（两端都不覆盖它）。

诚实地补一句：这块地基是逼出来的，不是量出来的。第九节的数据只能证明"在那份记录里它不可识别"，不能证明"任何记录里都不可识别"。这一步是本文最容易被推翻的地方，第十六节把它写成了一条可裁决的条款。

十三·与同题的另几篇的分界

本文与另外几篇同题作品出自同一批工作，读者极易把它们读成一件事。兄弟篇是彼此最近的对对手，因此每一篇都必须与另几篇划清，并各给一条判决性预测。

（甲）与"越册"那一篇。那一篇问的是"什么是作文"，答的是：作文是一次记录格式的施加，最要紧的失落不是写得不好，而是写了而这张表没有格子记——它的读数是场 B 的实质成员中在场 A 名册上无对应行的比例。

分界：那一篇讲的是成员资格（谁进得了名册），本文讲的是行进（两个已在册的成员之间那一段路）。两者可以同时成立而互不蕴含：一份表格可以行数齐全（没有越册项）而 λ 极大；也可以 λ 为零（每句有据）而大量真实内容根本没有行可记。

判决性对照：取一份两端依据齐备、每句可核的技术文档。越册那一篇预测它仍会漏掉整类内容；本文预测它的 λ 接近零、接手失败率低。若这类文档的接手失败率同样高，本文被那一篇吸收。

（乙）与"让段"那一篇。那一篇处理的是一件东西在失效过程中让出承载而仍未断的那一段，读数是自持点。分界：那一段是有承载者的（它正在把负荷让出去），本文这一段自始至终没有承载者。判决性对照：让段可以通过加固两端而缩短；悬跨段加固两端反而使它更明确——因为两端越硬，中间那一段与它们的不相容就越明显。

(丙)与"自铺"那一篇。那一篇讲的是行动者为自己铺出起始面，读数是自供率，落点在起点。本文的落点在中途：起点问题是"从哪里开始"，本文问的是"两处已经站住的地方之间怎么过去"。判决性对照：自供率高的写作者（善于自己造起点）不必然 λ 小——本文预测二者相关系数低于 0.3；若显著高于此，说明两者是同一件事的两个读数。

三条分界合起来还有一个白得的好处：四篇各自的读数如果在同一批文本上被同时清点，它们之间的相关矩阵本身是一次检验——若两两高度相关，说明这几篇实际上在测同一东西，其中至少三篇应当被合并。这条对本文不利，但它是可执行的，所以写在这里。

十四·三门手艺反过来给本文加的三个约束

借来的东西如果只提供支持，那说明借得不诚实。下面三处，每一处都削掉了本文原本打算主张的某一句更强的话。

(一)来自筮法：本文原本会写"悬跨段的位置是可以事先标出来的"，现在不能写了。

《易》给出的模型有一个诱人之处：老阴老阳是在成卦的当下就被标记出来的，即哪一处当动，程序自己会告诉你。本文起初据此设想：写作里也应当存在一套"当场标记"的办法——写到某处时就知道这一段是自造的，于是可以当场记账。

但《易》自己的体例否掉了这个设想：多爻同变时，当动之爻不能从任何一爻自身读出，要靠一个从整体余额算出来的数去指定。也就是说，标记只在"变的地方只有一处"时才是位点属性；一旦多处同时可变，标记就必须来自整体，而整体的算法是约定的。落到本文上：当悬跨段不止一段（几乎总是如此），"这一句属于哪一段"就依赖于编码约定，而不同的约定给出不同的切分。这与第九节实测到的符号翻转是同一件事的两面。

削掉的那句更强的话：本文不能再主张" λ 是文本的客观属性"，只能主张" λ 在给定编码规则下是可复算的"。适用范围因此从"任何文本"缩到"编码规则被写死并公开的清点"——第七节的清点手册由此从附件升为承重件：没有它，本文的读数不成立。

(二)来自火与燃料：本文原本会写"悬跨段越长越危险"，现在只能写一个有条件的版本。

火生态给了一个漂亮的形状：屏障效应随时间衰减，跌破阈值就失效。本文原打算照搬——续行概率随 λ 单调下降，存在一个可给出的临界长度。

第九节的检验没有支持它（H3 被驳，实测 0.680 对预设的 0.30），而火生态自己的材料里另有一条更麻烦的：火烧过的边界之内仍留着大片未烧的斑块。"未烧"既不是"不可燃"，也不是"没点着"，而是火没跳过去。移植到写作上，这意味着：一篇看上去连续完成的文章内部，可能留着大量实际没有被跨过、只是被绕过去的段落——它们在成品上表现为"这一节写得比较概括"。

削掉的那句更强的话：本文不能主张"完成即跨过"。一篇完成的文章不等于它的所有悬跨段都被跨过；完成率不是本文的因变量，因变量只能是"事后可核

率"与"接手失败率"这类要另做一次观察才拿得到的量。这一处削掉的是本文最省钱的那条检验路线。

(三) 来自合成化学：本文原本会写"补上中间的依据就能解决"，现在这条处方被限定了。

Curtin-Hammett 的含义是：改变库存不改变产物比例，只有改变出口才行。本文起初的处方是"给中间补料"——把两端之间的依据补齐，悬跨段就消失。但按这条原理，补料等于加库存；如果出口没变，加库存改变不了结果。

落到写作上，这句话的意思是：给学生更多材料，未必让他写得出那一段；决定他写不写得出的是那一段的出口条件——他要不要为这一段承担被追问的风险、有没有一种被允许的说法可以承接、这一段写错会不会被查。这解释了一个常见现象：材料给得越多，学生写出来的越是材料的复述——复述是出口最低的那条路。

削掉的那句更强的话：本文不能主张"补料是有效干预"，只能主张"补料是必要而不充分的"；充分条件在出口侧：为这一段提供一种被允许的、可承接的说法，并明确它的可追问程度。第十六节据此改写了两条证伪条款。这一处同时动了本文的价值排序：本文原本把"依据可核"排在第一位，现在必须承认，在出口不变的前提下，提高可核要求反而增加复述。

十五·不适用的边界

一个说法如果哪儿都适用，它多半什么也没说。以下五种情形，本文明确不适用：

1. 不要求依据的文体。诗、私人日记、自由抒写、口头讲述。这些文体里几乎每一句都是自造的， λ 等于全文句数，读数没有分辨力（第七节"不适用"一栏）。在这些文体里谈悬跨段是概念误用。
2. 纯汇编性写作。摘要、文献综述的引文串、法条汇编、会议纪要的逐条记录。这类文本的有据点密度接近 1， λ 恒小，本文能说的只有"这里几乎没有悬跨"，而这正是它们读起来乏味的部分原因——但"乏味"不是本文的判据，本文对文本质量不发言。
3. 多人分头写作而后合并的文本。此时两端之间的那一段可能由第三个人写，归属问题从"无人负责"变成"分工不清"，那是另一件事（见第十二节第四条对衔接工作的分界）。本文只处理一个人独自走完一程的情形；多人情形须另立框架。
4. 成品之上的事后判断。只拿到成品、拿不到过程痕迹时， λ 只能用引注等外部标记来近似，误差不可控（第十节的三条代理局限）。本文不支持仅凭成品对某位作者做出判断。
5. 教学评价与人事考核。见下一节的伦理三条：本文明确反对把 λ 用作评价个人的指标。

十六·证伪条件、赌注，以及什么不算命中

十六·一 最强的那个竞争解释

先把最难反驳的那句“其实不就是 X 吗”摆出来，而不是挑一个好打的：

其实不就是“难的地方就是难”吗？——两处依据之间需要推理，推理比复述难，难的地方当然更容易停下、更容易出错。用不着造一个新词。

这句话解释了本文的大部分现象，而且它便宜得多。要排除它，必须找到一个只有本文成立才会出现、而“难度说”预测不出的独有变量。它是这个：

难度是点的属性，可以被补料与训练降低；本文说的这一段的困难来自“依据不覆盖”，它不随难度下降而下降。

因此本文的机制检验一律写成这个形状：控制推理难度之后，“依据是否覆盖中间”仍与结果相关。

十六·二 六条证伪条款（编号，逐条可裁决）

F1（已跑，被驳）。跨到未被触碰过的小节，会使编辑会话终止风险上升。实测系数 -0.075 ($p \approx 0.385$)，方向相反且不显著，本条按预注册判负。它改掉了本文的什么，见第十一节。

F2（已跑，被驳）。间隙属性对“停在哪一步”的预测力优于点属性（AUC 差 ≥ 0.02 ）。实测 -0.126 ，按预注册判负。

F3（已跑，被驳）。跨到空地后的续行概率低于 0.30。实测 **0.680**，判负。从火生态搬来的阈值形式在此不成立。

F4（最便宜，未跑，本文最愿意押的一条）。控制推理难度（由独立评定者对同一批题目评分）与写作者水平之后，给出“两端 + 中间依据”的一组，其自造段长度 λ 与接手失败率均显著低于只给“两端依据”的一组；若两组无差异，或差异在控制难度后消失，本条判负、第一层主张作废。所需数据在多数教学机构的既有作业记录里已经存在（问题多班、留有初稿），一个人一周内可以跑完。

F5。同一批文本在两种都合理的编码规则下清点 λ ，若两次估计的秩相关高于 **0.8**，则“不可识别”这一层被推翻（那说明它只是难测，不是不可识别）；若低于 0.4，本层成立。这一条是本文自己主张的直接靶子。

F5 追记（已跑，判“未决”，并据此修订第十一节与第七节）。本条已按其字面执行。预注册文件写于取数之前，SHA-256 前十六位 **aad38bd382b5ae11**，其中补写了原文缺的一条：秩相关落在 **0.4** 与 **0.8** 之间一律判“未决”，不得事后择一。样本为 **Category:Featured articles** 按排序键升序前 **60** 篇的当前版本，**60** 篇全部可解析，无剔除。

两种编码都只用第七节已经写下的材料，不引入新判据。编码 **A** 实现规则 **1** 前半与规则 **5**：携带引注标记（<ref> 或 **sfn/harvnb/cite** 系模板）的句为有据点，小标题不中断 λ 。编码 **B** 实现规则 **1** 的全文并对规则 **5** 取相反约定：除引注标记外，含四位数年份或“数字 + 计量单位”的可核事实陈述亦计为有据点，且小标题

中断 λ 、新小节重新起算。两种约定都是第七节文本的合理落实，本文不主张其一为正确。

结果。 λ 中位数：编码 A 为 17.0 句，编码 B 为 7.0 句（H3 预测两者之差不小于 3 句，实测 10 句，成立）。主判据 Spearman 秩相关 $\rho = 0.525$ ($p = 1.7 \times 10^{-5}$)，落在 0.4 与 0.8 之间，按预注册未决。Kendall $\tau_b = 0.399$ ，恰在 0.4 判负线下方，与主判据不一致，一并报出，按预注册以主判据为准。H4 预测按 λ 降序取前十篇的重合数不超过 6，实测 7，H4 判负——“谁的悬跨最长”在两把尺子下比本文预想的稳定。

这一条改掉了本文的什么。第十一节“两种都合理的约定会给出符号相反的结果”这句话，只在第九节那份过程记录上成立（+0.677 与 -0.275），不能原样搬到第十节的成品清点上。在成品上，编码约定改变的是尺度而不完全是次序：中位数相差 2.4 倍， $\lambda \geq 10$ 的条目占比由 96.7% 掉到 31.7%，而逐篇秩位移的中位数是 13.5 名、秩相关仍有 0.525。第三层主张因此改写为一个更弱也更准的形状： λ 的绝对值不可跨编码比较，其相对排序只有部分可迁移。

直接后果有二。其一，任何以 λ 的绝对值设阈值的用法——包括第十一之二节“要自造的那一段超过十句就该缩跨”那条建议、以及第十六节末赌注里的 $\lambda \geq 10$ ——都必须与编码规则一并公布，否则那个数不成立。其二，第七节与第十节两处方向相反的偏向声明，至此有了量级：第七节声明混合句计为有据点使 λ 偏小，第十节声明三条代理局限使 λ 可能偏大，而同一批文本、同一份手册的两种合理落实之间， λ 中位数相差 10 句。第十节那个 13.5 应当读作“在该口径下”，本文不再声称它是一个可以跨口径引用的数。（本次编码 A 与第十节口径接近而不相同，切句与剥离规则略有差异，得 17.0 对 13.5；这个差本身也落在上述量级之内。）

F6. 流利的悬跨： λ 与停顿时长的相关系数若高于 0.5，则本文被停顿研究吸收（第十二节第一条）；若低于 0.2 而 λ 仍能预测接手失败率，本文成立。

F7（事后条款，标明来路）。第九节事后观察到的 U 形（谷底在接手之后的第二步）若在一批独立的记录里复现，则“交接之后的第一步稳、第二步脆”这一形状成立；不复现则该观察作废，本文不因此受损——它本来就不是本文的支撑。

如果本文被证伪，我们学到什么：若 F4 判负，说明写作困难确实是点上的难度问题，那么全部资源应当投向补料与表达训练，而“依据的连续性”是一个无用的字段——这对教学是好消息，因为那条路便宜得多。一个不会被触发的条件等于没有条件，因此上面每一条都写死了数字。

十六·三 一条写死日期的赌注

到 2029 年 8 月 17 日为止，至少有一家写作评价体系（考试评分标准、期刊审稿表、机构文书质量评查表、或一个被实际使用的写作辅助工具）在其正式条目中出现一个“两处依据之间的连续无据长度”性质的字段，并且这个字段有明确的计数规则。

什么不算命中（写死，事后不改）：

- 只要求"注明出处"、只统计引注数量或引注密度的，不算——那仍然是点的读数。
- 只在指南、白皮书、倡议里提到"注意衔接"或"避免无依据的推断"，而没有计数规则的，不算。
- 一个工具能标出"此处缺少引用"但不报告连续长度的，不算。
- 本文作者或与本文直接相关者参与制定的标准，不算。
- 只在学术论文里被提出而未被任何评价体系采用的，不算。

我押的是不会命中。理由在第十七节：这个字段一旦进入评价，最省事的达标方式会立刻摧毁它要保护的东西。

十七·反噬预言与三条伦理

十七·一 反噬

本文的处方是"看住两处依据之间的那一段"。设想它被一个机构采纳，最省事的实现方式是什么？

给每一段过渡文字也配一条引注。

这条路走通之后会发生三件事。第一，过渡处出现大量形式合规而实质空转的引注——引一本总论、引一句谁都同意的话、引作者自己上一篇。第二， λ 的读数会迅速降到接近零，而悬跨段一处也没有减少——它们只是被一层引注盖住了。第三，也是最坏的一件：每一条新加的引注都是一个新的有据点，于是它两侧各生出一个更短的悬跨段。按第六节的算术， n 个有据点意味着 $n-1$ 个交接；把引注加密一倍，交接数就翻倍。

这套做法在指标上表现为大幅改善，在实质上表现为悬跨段的数量翻倍而单段变短。而单段变短未必是好事——本文没有任何证据表明"许多短的悬跨"优于"少数长的悬跨"，事实上第十四节第二条恰恰提示，短而多的跨越更容易被"绕过去"而不被发现。

我看不到出口。任何把一个"连续长度"变成考核项的做法，都可以靠在中间插入形式条目来满足；而拒绝把它变成考核项，它就永远只是一句劝告。这一条不是待办事项，是本文自己造成的一个真实困境。

十七·二 伦理三条

λ 是一个可以拿去考核人的数。凡是可以拿去考核人的数，必须先写死这三条，否则本文不该被发表：

1. 这个读数指的是位置，不是人。 λ 是文本上两处依据之间的长度，它衡量的是这份文本的记录状态，不是作者的能力、勤勉或诚信。用 λ 给个人排名、评级、定优劣，是对本文的误用；一个高 λ 的段落完全可能出自最好的作者最有价值的推理。
2. 不得为了把这个数做好看而在关键场合安排形式性引注。尤其在病历、判决书、事故调查报告、审计意见这类文书中，为降低 λ 而插入不承担实质证明责任的引注，会污染真正的证据链。在这些场合，宁可 λ 高而据实标注"以下为推断"。

3. 已据此做出不利安排的，必须公开编码规则并给出复核通道。只要某人的成绩、职称、稿件处理因为一个 λ 读数而变差，那么第七节那份清点手册的具体版本、切句规则、判定为有据点的清单必须一并公开，且当事人有权按另一种同样合理的编码要求复算——第九节已经证明，两种合理编码可以给出相反的结论。

十八·本文自己落在哪一格

用本文自己的判据检本文自己，不是说一句“本文也有局限”。

先给出读数：按第七节的规则粗查本文， λ 最大的一段落在第五节与第六节之间——从“三件材料指着同一处”到“三分类型学”那一段，中间连续十余句没有任何外部依据可指。那正是本文的承重处：从“三门手艺的账本上都没有条目”推到“写作里存在一类无人负责的段落”，这一步没有引注，因为没有人写过这一步。

这个位置的具体后果有三个，逐条认账：

第一，本文自己就是丙格的一个标本（第六节）：三处依据（一条化学原理、一条变占体例、一组火生态数据）都是硬的、可核的，而把它们连起来的那一段是自造的，且读起来完全正常。凡是读到这里觉得“这不是挺顺的吗”的读者，正在经历本文所描述的那件事。

第二，本文的处方对本文自己无效。本文说“补上中间的依据”，而本文中间那一段的依据不存在于任何文献里——它只能由本文自己产出。这不是谦辞：如果本文是对的，那么本文自身的最关键部分必然是不可核的，因为可核的部分早就有人写过了。

第三，赌注的代价由引用者承担。我押的是三年内不会命中；若真不命中，本文的实用价值等于零而理论主张仍然成立，这是一个对作者极其舒服而对读者毫无用处的结局。写在这里，是为了让读这篇文章的人知道自己在承担什么。

分层引用

本文可以被分三层引用，请按需要取其中一层，不必整体接受：

- 第一层（最强、最易倒）：悬跨段是一类独立的东西，写作的发生落在它上面。倒了，后两层仍站得住。
- 第二层（分析工具）：一次“写不下去”有三种来路（第六节的三分表），其中第三种在所有形式检查上合格。即便第一层被证伪，这张表仍可作为诊断用具。
- 第三层（最稳，也最容易检验）：多数写作记录里没有“依据的连续性”这个字段。第十节的实测直接支持它：最高一档的评审能读出每句有没有来源，读不出两处来源之间隔了多远。这一层不依赖前两层，建议优先引用它。

交出去三个洞

1. λ 的分布形状本文说不清。是长尾还是双峰？若是双峰，说明存在两类不同的写作，本文的单一读数就该被拆成两个。

2. 两端“依据”的强弱本文没有分级。一条实验数据与一句名人观点都算有据点，这显然太粗，但本文找不到一个不引入主观判断的分级办法。
3. 口语与写作的差别本文完全没有处理。一个人能把第四点讲得很好却写不出来（第一节），按本文的框架应当解释为“口语允许更长的悬跨”，但为什么允许，本文没有答案。

参考文献

1. Arno, S. F., & Brown, J. K. (1991). Overcoming the paradox in managing wildland fire. *Western Wildlands*, 17(1), 40–46.
2. Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
3. Elbow, P. (1973). *Writing Without Teachers*. New York: Oxford University Press.
4. Emig, J. (1971). *The Composing Processes of Twelfth Graders* (NCTE Research Report No. 13). Urbana, IL: National Council of Teachers of English.
5. Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
6. Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
7. Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 243–281.
8. Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, 15(2), 113–134. doi:10.58680/rte198115773
9. Medimorec, S., & Risko, E. F. (2017). Pauses in written composition: On the importance of where writers pause. *Reading and Writing*, 30(6), 1267–1285. doi:10.1007/s11145-017-9723-7
10. Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
11. Parks, S. A., Holsinger, L. M., Miller, C., & Nelson, C. R. (2015). Wildland fire as a self-regulating mechanism: The role of previous burns and weather in limiting fire progression. *Ecological Applications*, 25(6), 1478–1492. doi:10.1890/14-1430.1
12. Pyne, S. J. (1982). *Fire in America: A Cultural History of Wildland and Rural Fire*. Princeton, NJ: Princeton University Press.
13. Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication*, 16(2), 106–112.
14. Seeman, J. I. (1983). Effect of conformational change on reactivity in organic chemistry: Evaluations, applications, and extensions of Curtin–Hammett/Winstein–Holness kinetics. *Chemical Reviews*, 83(2), 83–134. doi:10.1021/cr00054a001

15. Strauss, A. (1985). Work and the division of labor. *The Sociological Quarterly*, 26(1), 1–19.

16. Wallas, G. (1926). *The Art of Thought*. London: Jonathan Cape.

17. 朱熹：《周易本义·筮仪》，台北：大安出版社，2008 年，页 7–10。

18. 《周易·系辞上》，见《十三经注疏》本；“大衍之数五十”章。

19. 张政烺：《试释周初青铜器铭文中的易卦》，《考古学报》1980 年第 4 期。

20. 李学勤主编：《清华大学藏战国竹简（肆）·筮法》，上海：中西书局，2013 年。

字数与版本：正文约 1.9 万汉字（机检 18,959 字）；v1.0，2026 年 8 月 17 日。

数据与可复算性：第九节的预注册文件（SHA-256 前十六位 dae6976a53469b50）、取数与统计脚本、第十节的清点脚本与中间数据随本文存档。样本为英文维基百科公开接口所取的 60 篇特色条目（取数日期 2026-08-17），任何人可按同一规则重取复算。

人机分工声明：本文的取数、统计、清点与文稿撰写由作者借助计算工具完成；第九节的预注册与判负记录未经任何事后修改；第十节的清点为事后设计，已在正文中标明其不作为第一层主张的支撑。

版本追记：v1.1（2026 年 8 月 17 日）。相对 v1.0 的改动有三处，均为增补与降级，无一处删除：第十二节补入第（九）（十）（十一）三位分界（Toulmin 的 warrant 与 enthymeme、Spivey 的 discourse synthesis、Redi 等 2019 的维基百科可核性计量），各附判决性对照；第十六节 F5 由“未跑”改为“已跑，判未决”，并据其结果改写第三层主张的形状；第十节的 13.5 句自本版起标明为口径内读数。F5 的预注册文件（SHA-256 前十六位 aad38bd382b5ae11）、取数与清点脚本随本文存档。

第八章 补节 两端都不记账的那一段

本章v1.1 已补入三条全球分界（论证担保与省略三段论、话语综合的连接操作、维基百科「需要引证」句级判定），并跑完 F5，判为未决，第三层主张随之改写。补节只做两件事。

（一）与第一、二章的关系

本章与编一三章有一处容易被混同的地方，两边原稿都未写。

编一处理的是同一性：这一篇还是不是这一篇。本章处理的是归属：这一段文字算谁的。两者的独立性可以这样看——一段悬跨文字的存在与多寡，不改变这一篇的相；把它全部替换成有出处的转述，篇章定相率可以完全不变，而悬跨长度归零。反过来，一次破相变换可以完全发生在有据点之上，与悬跨段无关。

因此本章的读数与编一的读数正交，可以在同一篇作文上同时测量而互不干扰。建议成书时于本章开头一句点明，以免读者把「悬跨长度」误读为一种整体性指标。

（二）本章欠账

F5 判为未决这件事，其含义需要说清楚：两种合理的编码约定给出的悬跨长度中位数相差 2.4 倍（十七句对七句），而两种约定下的篇际排序只有部分可迁移（斯皮尔曼相关 0.525，落在预注册的未决区间；肯德尔相关 0.399 与主判据不一致，已如实报出）。这意味着本章的第三层主张——间隙属性在现有记录里不可识别——在成品清点上只成立一半：变的主要是尺度，不完全是次序。

本章v1.1 已据此改写第三层主张，并把「超过十句就该缩跨」这类可操作建议全部绑定到编码规则的公布上。编者认为这个处理是诚实的，但读者应知道其代价：本章目前给不出一个跨编码稳定的悬跨长度阈值，而这恰恰是它最想给教学界的那个数。

F4（同题多班、留有初稿）仍未跑，它是本章最便宜也最能判决的一条。

终编

制度这一侧

终编 制度这一侧

读完前八章，一个问题无法回避：既然这些空缺如此重要，为什么两千五百年里没有人给它们设一栏？

第九章的回答改变了问题的层级。它主张名分不是注意力的产物：一类残余物获得名分，当且仅当存在一条以它为清偿对象的职业通道。史实链支持这个判断——手稿早在一八八五年就已成规模入藏，可写作残余物获得专名要等到一九七二年，而它真正生效要等到一九八二年一间国立实验室的设立。中间隔着的不是九十年的技术进步，是一九六八年第一批靠这件事领薪水的职位。

由此得到本卷最后一组构件：裁决的四个参数（匿名、公开、定期、终局）、清偿率 $\lambda = c \cdot f \cdot \alpha$ ，以及一条不太好听的推论——终局性为零时，无论程序多么公正，缺口存量都不会下降。现代学术评审在匿名与公开上做了几十年的功夫，而终局性一直是零；这解释了为什么它看起来像结算，却从不减少任何东西。本章把这类动作称为准清偿。

这一章也是全卷的自反之处。本书自己就是一份「判者集合包含于被判者集合」的产物：它对自己的分界全部由自己作出，未经任何外部裁决。按它自己的命题，这类文本不能对自己签发终局。所以本卷不对自己给出结论性的成立宣告——它只把九处空格、九个读数和它们各自的死亡条件摆在那里，等一条通道。

第九章

灰为什么没有名分

灰的名分

——论书写残余的负形功能、缺的到期性与终审权易主

摘要

写作产出两种东西：一种进了正文，一种没有。后者——草稿、涂改、删除、动笔之前那个模糊的“意”——本文统称为灰。两千五百年的写作论对灰的态度惊人地一致：看得见，说得出口，却始终不给它一个属于它自己的名分。名分迟至二十世纪七十年代才在法国出现（*avant-texte*，前文本），并在八十年代随一间国立实验室的设立而生效。本文要问的不是“灰是什么”，而是名分为什么在那一年发生□

本文提出一条与直觉相反的解释：名分不是注意力的产物，是清偿制度的溢出物。一类残余物获得名分，既不因为有人看见它，也不因为它被保存下来，而因为出现了一条以它为清偿对象的职业通道——出现了一批人，他们自己的账要靠查这批灰才能结清。据此本文把“缺”从一个体验词改造为制度量：给出裁决制度的四参数（匿名、公开、定期、终局），定义清偿率 λ ，写出缺存量方程 $dK/dt = i(t) - \lambda(t)K(t)$ ，并证明一条不可自签定理：当评判者集合包含于被评判者集合时，任何判决在制度上都可再审， λ 恒为零，缺存量随时间线性发散。由此可以给出“文人圈靠什么粘连”“为什么现代出不了归隐者”“审计仪式为什么必然膨胀”三个现象的同一个量纲原因，并把民事诉讼法的既判力（*res judicata*）作为已成熟的清偿制度形式搬进写作与学术评价的分析。

本文做了两次预注册真跑。研究一取 2016 年 6 月首发于 arXiv □ 1,735 □ 论文，检验“外部一次性裁决是否关闭账目”。三条预注册假设全部失败：有无外部裁决，三年后仍被修订的比例几乎相同（2.1% vs 1.7% , $p=0.51$ ）。由此本文当场改判：到期性不是作品的属性，是主体账目的属性——作品几乎瞬间冻结（63.7% 在沉积三十天内不再被改动，98.2% 三年后不再改动），而缺继续挂在人名下。这一改判反过来为本文另一条命题提供了迄今最硬的读数：名分与可变性成反比，一份文本一旦取得可被引用的公共标识，其可改率在数周内塌向零；名分这件事，是给灰提前举行的葬礼。研究二用 Google Books 语料检验名分的时点：法语中 *avant-texte* 与 *critique génétique* 的扩散拐点分别落在 1972 与 1983 前后，与两件关键词（*brouillon* □ *manuscrit*）的平坦曲线形成对照，且相差约十年。给残余物命名与使名分生效，是两件事，差一条职业通道。

关键词：前文本；负形；名分；清偿率；既判力；终审权；文人圈；审计仪式

一·问题：一个两千年没有名分的东西

先要把谈的东西固定住。

一次写作结束时，纸面上留下的字是一部分；被划掉的字、写了三行发现不对而作废的段、动笔之前脑子里那个还没落地的“意”，是另一部分。第二部分本文统称为灰。选这个词不是为了修辞：它同时含有“烧过的”“留不住的”“还带着形状的”三层意思，而这三层恰好是本文要处理的三件事。

关于灰，有一个奇怪的事实值得先摆在桌面上：几乎所有写作传统都承认它存在，却几乎没有一个传统给过它名分。

陆机《文赋》开篇即说“恒患意不称物，文不逮意”——写作的常态是够不着，这句话把灰的体验掐得极准。刘勰《文心雕龙》立《镕裁》一篇，专讲删削。贾岛推敲、黛玉焚稿、扬雄书被讥“覆瓿”，手段、态度、社会处置样样俱全。西方那一路，贺拉斯劝人把诗压九年再拿出来，瓦莱里说诗从未完成、只是被放弃。可是把这些词一个个拿起来看，会发现它们说的全是正稿这一侧的事：推敲说的是最后挑中的那一下，镕裁说的是裁掉芜杂之后剩下的正料，焚稿说的是处置，覆瓿说的是社会评价。没有一个词是给灰本身的。灰在这两千年里只有一个负号——“不逮”——没有正形。

这不是因为没人聪明到看见它。看见它的人从陆机到瓦莱里排成一队。所以真正的问题不是“为什么没人看见灰”，而是：

看见一样东西，与给它一个名分，为什么是两件不同的事？中间隔着什么？

本文的全部工作，就是回答这个“中间隔着什么”。答案会是一个听上去与文学无关的词：清偿□

在展开之前，还需交代一句方法上的立场。本文不把“名分”当成语言现象来处理。一个词被造出来是廉价的，一个人在自家书房里给任何东西命名都不需要许可。名分是这个词开始为某些人做事的那一刻才发生的——它必须能被当作理由使用：能出现在项目申请里、职位描述里、学位论文题目里、期刊征稿启事里。研究会用数据把这两层分开，而分开之后可以看到，它们相差整整十年。

二·负形：灰在写作里干什么

在问“灰为何没有名分”之前，先要说清灰做什么。因为如果它什么也不做，没有名分就毫不奇怪。

本文第一条断言是：灰不是原料，也不是废料。这句话值得展开，因为它切开的是一个几乎无人怀疑的分类习惯。

原料与废料是按成品称重才分得出的身分。问一样东西是原料还是废料，等于先假定有一个成品，然后问它能不能装进去。这个问法对锯木厂成立，对写作不成立——因为在写作里，成品是最后才出现的那个东西，而灰发生在成品之前，并且参与决定成品是什么。用成品去称量一个参与决定成品的东西，是拿结果当尺子量原因。

那么灰做什么？本文给它一个精确的功能名：

定义 1（负形）：设写作者要达到的目标为 t ，而 t 在动笔之前不可被直接说出（若可直接说出，写作就退化为誊写）。设写作者依次写出 $w_1, w_2, \dots, w_n$ ，每一个都被自己判为“不是它”。则 t 第一次获得可辨识的形状，不是以自身的形状，而是以集合 $\{w_1, \dots, w_n\}$ □ 否定轮廓的形状。这个轮廓称为 t □ 负形□

三条推论直接跟着定义走：

(1) 负形不是“未被选中的选项”。选项预设备选集合已经存在，选择只是从中取一个。负形恰恰相反：备选集合不存在， w_1 到 w_n 是被写出来才存在的，而写出它们的唯一功能是让 t 的边界露一点出来。这条差别不是措辞讲究——它决定了灰能不能被事先枚举。选项可以事先枚举，负形不能。任何“先想清楚再写”的写作教学，其隐含假设都是灰属于选项一类。

(2) 负形的密度决定目标的分辨率。只写一句就成的，其目标本来就粗；写了三十句都不是的，其目标的边界被三十条否定线夹出来，比前者精确得多。这给出一个反直觉的推论：写得顺的时候，往往不是你想得清楚，而是你要的那个东西本来就不精确。

(3) 负形对写作者不可见，只对回读者可见。这一条是下一节的全部内容，也是本文与“试错说”分开的地方。

关于第三条要多说一句，因为它把本节从一个漂亮说法变成一个有后果的命题。试错说（写作是尝试—失败—再尝试）已经足够解释 w_1 到 w_n 的产生，但它解释不了为什么这些失败对写作者本人往往是不可读的。试错的经典模型里，每次失败都携带方向信息（错在哪、往哪调）。而写作里的灰通常不携带方向信息：三句都不对，不告诉你第四句该往哪写。它只做一件事——把“不是它”这一侧的边界推进一点。方向信息不是从单条灰里读出来的，是从灰的分布里读出来的，而一个正在写的人不可能同时站在自己灰的外面看它的分布。

于是有：

命题 1：负形的读取需要一个与写作动作分离的位置。写作者本人能占据这个位置的唯一方式，是把时间拉开（回读）；而更彻底的方式，是让另一个人来读。

这条命题看起来只是写作心理学的一句常识。但把它接到制度上去，它会一路长成本文的主命题。

三·回读：废料资格是怎么取得的

第二条判断更狠：灰算废料，只有一种情况——你不回读它；否则它连废料的资格都没有。

这句要用“资格”这个词，而不是价值。废料是一个已经被处置过的身分：某物被拿起来、被判定不合用、被放进垃圾桶，这一整套动作走完，它才叫废料。没走完这一套的，不叫废料，叫“没有发生过”。

定义 2（回读）：写作者写出 w_i 之后的某个时刻，重新把 w_i 作为对象读一遍，并对“它为什么不是 t ”作出一个可以说出口的判断。回读的产物不是 w_i 本身，是关于 w_i 与 t 之间那道差的一句陈述□

于是灰有三种命运，而不是两种：

命运	条件	结果
进入正文	被判为 t 或 t 的一部	成为作品
分		
成为废料	被回读，被判为不是	进入下一次写作的负

命运	条件	结果
	t, 且这个判断被说出	形库
未发生	未被回读	既不在作品里, 也不在负形库里

第三行是本文要救的那一行。一份被删掉又从未被回看的稿子，在写作的因果链上等于没有存在过：它既没有改变作品，也没有改变作者。它被否定的资格，被作者自己没收了。

到这里可以给出本文第一条可测量的东西。回读是一个可观测事件：在数字写作里它留下访问记录、版本对比、修订注记；在纸笔写作里它留下二次批注、圈画、“此段不通”之类的旁记。于是可以定义：

定义 3（回读率 ρ ）：一段时间内被回读过至少一次的废弃片段数，除以全部废弃片段数。

ρ 不是一个修辞。第十四节的数据会以一种意外的方式碰到它：在一个可以精确观测“文本何时被再次动过”的语料里，大多数文本在沉积之后极短的时间里就再也没有被动过——如果把“再次动过”当作回读的下界代理，那么 ρ 在公开学术写作里低得惊人。这个读数不是本文预期的，也不是本文希望的，但它比本文原来打算证明的东西更硬。

还有一句必须先写下，因为本文后面所有的“制度”讨论都是它的展开：写作里最厚的灰不是写出来的字，是动笔之前那个想法。字一落，那个“意”就死了；每一篇成文都是那堆灰的尸体。这句话的作用是提醒：灰的绝大部分永远不可能被任何档案捕获。第十九节讨论边界时会回到这里——本文的全部制度分析只适用于已经落成物质痕迹的那一小部分灰，而这一小部分在全部灰里所占的比例是未知的，很可能很小。

四·中国文论为何只有负号，没有名分

现在回到那个奇怪的事实：看见了，为什么不给名分。

中国文论这一路的态度最值得研究，因为它承认得最坦率。“意不称物，文不逮意”是承认；“胸中之竹”与“手中之竹”必然走样是承认；“意翻空而易奇，言征实而难巧”还是承认。承认之充分，几乎已经把发生学批评的问题意识说完了。可是接下来的每一步，立场都站回正稿这一侧：

- 推敲处理的是两个候选之间的选择，而不是被淘汰那一个的去向。故事的高潮是韩愈说“敲”字好——被否定的“推”字随即消失，没有人问它去了哪里。
- 铨裁处理的是芜杂与精要的比例，判据是“善删者字去而意留”。请注意这个判据的方向：删得好，是因为意没有损失。这恰好排除了本文的负形——负形的全部功能正是它带走了一部分意，而那一部分意此前根本不存在于任何地方。

- 焚稿是处置，且是带情感的处置。焚稿的动作里含有一个被忽略的前提：稿是属于作者的私物，因此作者有权销毁。私物的销毁不需要向任何人交代，也不构成任何人的损失。
- 覆瓿（桓谭说扬雄的书将来只配盖酱坛）确实是一个负名分，但它给的是“被拒绝的完整文本”，不是“被否定的片段”。它属于接受史，不属于写作过程。

把这四条并排放，可以读出一个共同的结构：中国文论有一整套关于灰的词汇，但这些词全部是动词或评价，没有一个是名词。推敲、铨裁、焚稿是动作；不逮、芜杂、覆瓿是评价。名分需要名词，因为只有名词能进目录、能立条目、能成为一个被研究的对象。

于是要小心一个容易滑过去的解释：说中国传统“不重视过程只重视结果”。这个解释是错的，或至少是空的——推敲这个故事本身就是一个关于过程的故事，而且流传了一千二百年。真实的差别不在重视与否，在于：

关于过程的叙述，与以过程为对象的研究，是两件事。前者只需要一个讲故事的人，后者需要一批以此为职业的人。

顺带指出一个常被误引的对照：中国其实很早就有对灰的建制化处理——稿本、异文、校勘。藏书家、目录学家、校勘家整代整代地与草稿打交道，《四库全书总目》里“稿本”“校本”是正式的著录范畴。但请看这个建制的目的：鉴定、收藏、定本。校勘学处理异文是为了求得一个最可靠的正文，稿本的价值在于它是善本谱系的一环。也就是说，中国把灰接进了文献学的清偿链，而不是写作论的清偿链。灰在文献学里有名分（“稿本”是一个正经名分），在写作论里没有。这一处对照对本文极为重要，因为它是一个天然的对照组：同一个物件，接进哪条职业链，就在哪条链里获得名分。

五·法国发生学批评：名分发生的那一次

唯一一次给灰立成正式名分的，是二十世纪法国的发生学批评（critique génétique）。它给那些没进正文的东西起了一个专名——前文本（avant-texte），并把它确立为研究对象：不是作品，而是草稿、涂改、删除、旁注、提纲、通信构成的那一整套“生成卷宗”。

这件事的时间线是可以逐年查证的，而且这条时间线本身就是本文的主要论据：

年	事件	性质
1885	魏玛歌德—席勒档案馆设立；雨果去世，手稿悉数遗赠法国国家图书馆	材料可得性
1966	法国国家图书馆购入海涅手稿群	材料可得性
1968	法国国家科研中心（CNRS）成立海涅小组，任务是清点与分类	职位出现

年	事件	性质
1972	贝尔曼-诺埃尔提出 <i>avant-texte</i>	命名
1978	马特拉菲赖德国际会议“前文本、文本、后文本”	学术共同体成形
1982	研究组升格为国立实验室 ITEM（现代文本与手稿研究所）	建制生效
1992	专业期刊 《Genesis》创刊	出版通道

把这张表读一遍，一个通常的解释会立刻站不住脚：技术与材料并不解释名分。手稿的大规模可得性在 1885 年就已经具备，复印技术在二十世纪三十年代已经成熟，然而名分等到七十年代才出现，中间空了将近一个世纪。在这九十年里，档案馆一直在收灰，图书馆一直在编目，学者一直在翻阅——什么都不缺，就是没有名分。

那么1968 年发生了什么不同的事？不是有人第一次看见了灰，是有人第一次可以靠灰吃饭。CNRS 给了几个位置，这些位置的工作内容就是处理海涅那批手稿。研究者的职业账，从此要靠这批灰来结清。他的年度报告、他的晋升材料、他的学生的学位论文，全部以草稿为对象。到这一步，“草稿”不再是别人写作过程的残余，它变成另一群人的产出物。

这就是本文的命题：

命题 2（名分的清偿链条件）：一类残余物获得并保持名分，当且仅当存在一条以它为清偿对象的职业通道——存在一批人，他们自己的账要靠处理这批残余物才能结清。名分不是注意力的产物，是清偿制度的溢出物。

这条命题有一个立刻可以检查的反面：自己给自己的灰编号，永远不产生名分。一个作者把废稿装订、编目、写上日期，做得再整齐，也只是一座私人陈列馆。原因不在勤勉程度，在结构：编号的人与写的人是同一个人，他不会像陌生人那样顶撞自己删掉的那一句，他只会心疼它。更要紧的是，这套编号不进入任何人的清偿链——没有人的账因为读了它而变得可结。

由此得到本文最冷的一条判断：你缺的那一片条件，不是更多灰稿，也不是更细的编号，是有人来查你的灰。而“有人来查”不是一个道德呼吁或社交运气问题，它是一个制度供给问题：只有当查灰这件事本身成为一个可以结清的职业时，查灰的人才会稳定地出现。

六·命名与名分生效：两件事，差一条职业通道

命题2 有一个即可测的推论，本文把它单独拎出来，因为研究二正是为它设计的。

如果名分是清偿链的溢出物，那么造词与名分生效应该在时间上分离：- 造一个词几乎不花成本，一个人在一篇论文里就能做到（贝尔曼-诺埃尔 1972 年造

avant-texte 就是一个人做的事)；- 但让这个词开始为一群人做事——进课程名、进机构名、进期刊征稿、进学位论文题目——需要那条职业通道先立起来。

提：

命题 3 (名分双层)：给残余物命名 (给物一个名字) 与名分生效 (给查这个物的人一个身分) 是两个可分离的事件，后者滞后于前者，滞后长度约等于建制化所需的时间。

这条命题是可以被数据判死的：如果两个词的扩散曲线拐点相同，命题 3 伪；如果称物的词与称职业的词相同，命题 3 亦伪。第十七节给出检验结果——两条曲线的拐点相差约十年，而且方向与预测一致。

在进入形式化之前，先把上一节留下的那把刀转向读者自己。发生学批评把灰立成了对象，代价是什么？

代价是灰死了第二次。名分发生的那一刻，一份涂改从“还可能变成别的”变成“已经被研究的前文本”。它得到了一个稳定标识 (卷宗号、页码、转写符号)，而稳定标识的全部功能就是保证它以后不再变。可引用性与可变性在这里是互斥的：一份可以被精确引用的东西，必须已经不再改动。

命题 4 (名分—可变性反比律)：一份文本获得可被公共引用的稳定标识之后，其被作者再改动的概率急剧下降；名分的取得与可变性的保有不可兼得。

这条命题在第十六节会拿到本文最硬的一组读数，而那组读数原本是为另一个假设采集的。编号这个动作本想留住灰，实际上是给它提前举行葬礼——这句话，本文有能力把它从格言变成一条带百分数的曲线。

七·缺：从体验词改造为制度量

前面五节谈的是灰。从这一节起谈缺——写作者、学者、创作者身上那个“还差一点”的状态。这两者的接口是第五节那句：你要的那片条件，是有人来查你的灰。查灰的人不来，缺就一直在。

日常语言里，缺是一个体验词：怀才不遇、不被看见、差一口气。体验词的麻烦是它不可比较也不可反驳——两个人都说自己缺，无法判断谁缺得多，也无法判断某项制度改革是否减少了缺。本文要做的第一件事，是把缺改造成一个可以加减、可以有存量与流量、可以被制度参数决定的量。

改造的入口是一个观察：缺总是相对于一次尚未到来的裁决而言的。说自己缺名分，意思是有一个判断还没有被作出，或者作出了而不算数。没有任何裁决结构的地方没有缺，只有匮乏——饥饿不是缺，落榜才是缺。

定义 4 (缺)：一个主体在某个领域中的缺，是他已投入而尚未获得

终局裁决的作品—努力的存量。记为 $K(t)$ □ 定义 5 (清偿事件) □ □

次裁决 J 若使某一存量部分不可再被重新裁决，则称 J 对该部分构成

清偿。被清偿的部分从 K 中扣除。定义 6 (续期)：一次裁决 J 若在

形式上作出结论，但该结论可被同一制度在未来重新审理，则 J □ □

成清偿，只构成续期：它把结算日推后，同时通常伴随新的投入进入

K □

三个定义合起来，可以立刻读出科举与现代学术的差别，而这个差别在体验层面是读不出来的（两边的人都觉得自己缺）。

科举一路。缺在那里有三套完整的名分：一是词表——不第、落第、下第、蹭蹬、潦倒、怀才不遇；二是刻度——名次、榜次、官职品秩、外放之地离京多远；三是清偿日——放榜。“怀才不遇”这个成语本身就是缺的自我叙事模板：你缺什么、为什么缺、该怪谁，词里都给了答案。更要紧的是缺可以支取：及第那一天，存了二十年的缺当场兑付。于是科举圈的组织方式是“大家存同一种缺，按存量排座次，依次清偿”。缺在那里制造的是秩序与预期。

现代一路。缺没有词表（“内耗”“体制”是抱怨，不是刻度），没有刻度（引用数不是清偿，见第十三节），最关键的是没有清偿日。评审、评奖、评职称都可以“明年再来”。于是缺不断续期，存量单调增长。

这个对照不是要浪漫化科举。缺可清偿，不等于人自由：及第之后缺立刻换债主——京官缺外放，外放缺回京，清流缺圣眷，圣眷正隆缺全身而退。这是一条独立的规律，值得单独记下：

命题 5（债主转移）：清偿一次缺不消灭缺，只更换其债主。一个人被清偿之后进入的下一套裁决结构，会立刻在新的维度上产生新的缺。可清偿性关乎结构，不关乎幸福□

命题5的作用是防止本文被读成一个复古方案。本文主张的从来不是“回到放榜”，而是：只有当缺可清偿时，退出才是可能的（第十八节）。

八·裁决的四参数与清偿率

现在把裁决制度形式化。任何一套裁决制度 J ，本文用四个参数刻画：

a（匿名性）：判决作出时，判者与被判者的身份是否互相遮蔽。取值 $[0,1]$ □ **p**（公开性）：判决的规则、时点与结果是否对全体参与者公开。取值 $[0,1]$ □ **c**（定期性）：判决是否按预先公布的固定周期发生，而非由申请方随时发起。取值 $[0,1]$ □ **f**（终局性）：判决作出后，在同一制度内不可被重新审理的程度。取值 $[0,1]$ □

四个参数的分工不对称，这一点很重要，是本文与“程序公正”一类讨论分开的地方：

- **a** 与 **p** 决定判决是否被认账。匿名与公开影响的是当事人和旁观者是否承认这次判决有效。它们决定判决的可接受度□
- **c** 与 **f** 决定判决是否结账。定期性保证判决会发生（不定期 = 可以永远不发生），终局性保证判决发生之后账目关闭。它们决定判决的清偿能力□

由定义：

定义 7（清偿率 λ ） □ $\lambda = c \cdot f \cdot \alpha(a, p)$ ，其中 c 为单位时间内的裁决频次（定期性的自然读数）， f 为终局性， $\alpha \in (0,1]$ 为由匿名性与公开性决定的认账系数。 λ 的量纲是“每单位时间被结清的存量比例”。

三条边界情形值得写出来，因为它们各自对应一种现实制度：

1. $f = 0$ (可再审) $\Rightarrow \lambda = 0$ ，无论 $c \square a \square p$ 多好。年年可评的职称评审，即使程序完美、评委匿名、结果公示，也不清偿任何东西。程序公正与账目关闭是两个变量，而现代学术评价改革几乎全部投在前一个上。
2. $c = 0$ (不定期) $\Rightarrow \lambda = 0$ ，无论 f 多高。诺贝尔奖之类的裁决终局性极高（不可上诉），但对任何具体个体而言其发生是不定期的，因而不构成任何人的清偿台：没有人可以据此安排自己的结算。
3. $\alpha \rightarrow 0$ (不认账) $\Rightarrow \lambda \rightarrow 0$ 。判决作出了，但当事人与同行都不承认它有效（例如公认由人情决定的评审），账面上销了，实质上没销。这一条解释了为什么单靠增加评审次数无法减少缺。

科举的四参数： a 高（糊名誊录）、 p 高（张榜）、 c 高（三年一科，可预期）、 f 高（一次算数，落第者进入下一循环而非申诉本次）。四项俱高， λ 显著为正。

现代同行评审的四参数： a 中（单双盲混杂）、 p 低（评审意见通常不公开）、 c 低（滚动投稿，随时可发起）、 $f \approx 0$ （可再投、可再审、明年还有机会）。前三项即使全部改善，只要第四项为零， λ 仍为零。

这就是本文对“缺的现代形态”的完整回答：不是没人判你，是可再审这件事本身，让每一次判决都不结账。

九·缺存量方程与圈子的黏度

有了 λ ，可以写出缺的动力学。设 $i(t)$ 为单位时间进入的新投入（写出来的、投出去的、报上去的）， $K(t)$ 为缺存量：

$$\text{方程 1} \quad dK/dt = i(t) - \lambda(t) \cdot K(t)$$

两种情形：

$\lambda > 0$ ：存在稳态 $K^* = i/\lambda$ 。缺是一个有界的、可预期的量。个体可以据此规划：知道自己大约存多久能兑付，也知道兑付之后进入哪个新循环。这就是科举圈的形态——缺在那里制造秩序与预期，而不是粘连。

$\lambda = 0$ ： $K(t) = K(0) + \int i(t)dt$ ，随投入单调发散。没有稳态，没有兑付日，任何个体的缺只增不减。而这正是本文要解释的现代形态。

发散的存量必须被某种东西吸收，否则制度会在体验层面崩溃（每个人都感到自己在无底洞里投入）。吸收的方式有三，恰好对应三种可观察的现象：

1. 准清偿仪式（第十三节）：互评、研讨、推送、赠书、引用——看起来像结算，实质上不减少 K ，只让缺停在“快被填满、还差一点”的状态。
2. 叙事补偿：给缺配一套说法（“体制”“内耗”“怀才不遇”的现代版本），使不可清偿被解释为不公正，从而在心理上可承受。
3. 人员流出：一部分人离开该领域，其存量从系统中移出。

这里可以给出本文对“文人圈”最反直觉的一条判断，也是全文最值得形式化
□ □ □ □

命题 6 (缺是粘连剂)：靠缺聚集的共同体，其内部凝聚力正比于缺存量 K ，因而该共同体有内生动机维持 $\lambda = 0$ 。给某个成员真正结清账目的那个人，一旦出现，会使该成员的 K 归零，从而使其对共同体的依附消失。

命题6 有一个乍看奇怪的推论：这类共同体不会阻止你被承认，它会热情地、频繁地、几乎不间断地承认你——只是每一次承认都不结账。这不是虚伪，是结构：真正结账的承认会把人放走，不结账的承认把人留住。

命题6 与行会形成对照。行会靠“已有什么”聚集：有活、有职、有准入门槛，成员资格是一个已清偿的状态。靠缺聚集的共同体则靠“都缺同一类东西”聚集，成员资格是一个未清偿的状态。两者的稳定机制因此相反：行会通过发放资格稳定，文人圈通过延迟结算稳定。

十·不可自签定理：为什么同行不能签发清偿

现在到本文形式上最硬的一步。上一节说文人圈“有动机”维持 $\lambda=0$ ，这还只是功能主义解释，容易滑向阴谋论。本节要证明的是更强的东西：在同行评议这种结构下， $\lambda=0$ 不是动机问题，是资格问题——即使所有人都真心想结账，也签不出来。

定理 1 (不可自签)：设裁决制度 J 中，判者集合为 A ，被判者集合为 B 。若 $A \subseteq B$ (每一个判者同时也是待判者)，则 J 中不存在具有终局性的判决，即 $f = 0$ ，因而 $\lambda = 0$ 。

论证。终局性的实质是：判决作出之后，任何一方都不能在同一制度内要求重新审理。这要求存在一个签发位——一个判决之后不再受同种裁决性置。若 $A \subseteq B$ ，则任判者 x 在作出判决后，仍然为被判者，且其判决来源系统本身也是同行，因此他被同判。于是对 x 的判决是重审，等行对 x 的资格质疑，而 x 的资格被质疑的资格来源此时重审不需要任何制度干预——它是系统的常行，故不存在重审判决， $f = 0$ 。■

三、溯，防止被用：

(一) 这不是说同行评议无用。同行评议在分辨质量上可能是现有最好的装置；定理1 只说它在关闭账目上无能。这是两种不同的功能，混同它们是现代学术评价的核心错误：我们用一个为分辨设计的装置去承担结算职能，然后奇怪为什么大家都觉得没完没了。

(二) 反例检查——终身教职。终身教职 (tenure) 看似是同行评议，实则不满足 $A \subseteq B$ ：评审委员会由已获终身教职者组成，他们不再是同类判决的对象；且最终签发通常在校董会一级，那是一个不受同行评议约束的外部位置；更关键的是“上或走” (up-or-out) 设定了 c 与 f ：有固定考察期 (c 高)，且实质性学术判断通常不可上诉，申诉一般只限于程序瑕疵 (f 高)。因此终身教职是现代学术界少数真正的清偿台。这个反例不推翻定理 1，它证实了定理1 的条件是必要的：一旦条件被满足 ($A \subseteq B$)， λ 归零；一旦被打破 (判者已出局)， λ 为正。

(三) 由此得到一条可证伪预言。如果本文的分析成立，那么终身教职后评审 (post-tenure review) 的引入应当把已清偿的账目重新打开，其观测

果是在强制性后评体系中,获得评价之后才出现更新或获得之前,而在无后评体系中,两者有显著区别
第三部分为本文的第三条证据

同样的推论适用于中国的职称制度：职称年年可评，这一条设定就使 $f = 0$ ，无论评审程序如何改良。反过来，如果某地把职称评审改为“五年一次、题目公开、作品匿名、当场判定、不设次年再报”，本文预言该地的圈内准清偿动作（互评、研讨、互引）密度将显著下降。这是一个真正的干预实验设计，成本不高，且有现成对照组。

十一·既判力：一个已经成熟的清偿制度

到这里，本文需要一个已经把“清偿”做到成熟的领域，用来检验上面这一套是不是自造概念。这个领域是法律，那个概念是既判力（res judicata）。

大陆法系民事诉讼法把判决的效力分为两层：

- 形式确定力（formelle Rechtskraft）：上诉期届满或救济途径穷尽后，判决不再可以被通常方式撤销。
- 实质既判力（materielle Rechtskraft）：同一当事人就同一诉讼标的不得再行起诉，法院亦不得作出与之矛盾的裁判——即“一事不再理”。

请注意这两层与本文四参数的对应关系：形式确定力对应 $f \square$ 程序侧（不可再审的程序保证），实质既判力对应 $f \square$ 实体侧（同一事项不得重开）。法律用了两千年把这件事做成制度，而它做的正是本文所谓清偿：给一场争议一个结算日，之后账目关闭。

这个搬运不是类比装饰，它带来三样文论与学术评价里没有的东西：

（1）既判力的边界是被明确定义的。它只及于“同一当事人、同一诉讼标的”，超出这个范围的新争议照样可诉。也就是说，成熟的清偿制度不承诺“从此不再有争议”，只承诺“这一笔不再重开”。对应到学术：一个可行的清偿制度不需要判定一个人的学术价值（做不到），只需要判定“这一次申报，结清”。现代学术评价之所以做不出清偿，部分原因是它把裁决对象设定得太太——它想判人，而不是判一笔。

（2）既判力有例外，且例外是穷尽列举的。再审事由（新证据、程序重大违法、判决基础被推翻）是封闭清单，不是“有理由就能重开”。这给出一条可直接移植的设计规则：清偿制度的可重开事由必须是有限清单，且不含“有新的更好作品”这一项。恰恰是这一项——“他明年可能更好”——摧毁了学术评价的全部终局 $\square \square$

（3）既判力与真理无关。判决确定之后不可重开，并不意味着判决为真；法律明白地承认可能判错，仍选择关闭。这正是学术评价最难接受的一条，因为学术共同体持有一个隐含信念：判断应当随真理更新，因此任何判断都不应终局。这个信念是高尚的，也正是 $\lambda = 0$ 的意识形态来源。本文由此得到一条判断：

命题 7（真理无签发资格）：以真理为最终依据的裁决系统不可能自带清偿，因为真理不签发判决。任何要在这类系统中获得清偿的努

力，必须引入一个明知可能判错仍然关闭的制度装置。同行只是替真理收债的人，而他自己也在等清偿。

这条命题不是在贬低学术，它在指出一个结构性事实：学术评价的痛苦不自评价者的品性，来自它把结算权托付给了一个从不签字的债权人。

十二·终审权易主：1905 与洪堡

有定理1与命题7，历史部分可以写得很短，因为它不再需要叙事，只需要指认参数的变化。

粗略的说法是：“不是一步制度改的，是一步把缺的终审权从朝廷转给了同行。”本文把它精确化为：

命题 8（终审权易主）：中国读书人的缺，其终审权在二十世纪上半叶由一个具备签发资格的外部机构（朝廷）转移到一个不具备签发资格的同源集合（同行）。这一转移不改变缺的存在，改变缺的到期 $\square\square\lambda$ 由正变零。

两个坐标：

1905 年废科举，废掉的不是缺，是缺的公共结算台。这一点常被忽略：科举的社会功能里，选官只是其一，另一项是给整个读书人群体提供一个 $a\square p\square c\square f$ 四项俱高的定期结算装置。废除它之后，选官功能由新式教育与考试系统接手，结算功能却没有任何机构接手 $\square$

洪堡改革之后的德国大学提供了替代模型：教师凭研究产出晋升，研究好坏由同行滚动评议。这套装置到十九世纪末在欧洲定型，二十世纪初经由留学生与教育考察进入中国。它带来的四参数是：署名（a 低）、封闭评审（p 低）、滚动 c 低）、可再审（ $f \approx 0$ ）。

于是转手完成的那一刻，可以用一句话概括：判决从“一次算数”变成“次次不算数”。

这里有一个误读：本文主张 1905 年的人做错了。科举的四参数虽然俱高，其判据内容（八股、经义）与现代知识生产完全不相容，废除是必然的。本文指出的是一 $\square$ 未被察觉的连带损失：改革者废除的是一套判据，同时废除的是一台结算机，而後者的功能从未被识别，因此从未被替代。一百二十年来，中国学术评价的历次改革——从职称量化到代表作制到破“五唯”——全部在改判据，无一在改终局性。这条观察是可以被反驳的：如果能举出一项改革明确设定了“不可再申报”，本文这一段作废。

最后一句留给制度对个体的反向规定：新的结算结构立稳之后，它规定这一批人写什么甚至用什么话哭（“体制”“内耗”）。缺的语言本身是由裁决结构发的。科举时代的人用“蹭蹬”“不遇”哭，现代人用“内耗”哭，两套词的差别不是情绪的差别，是可清偿与不可清偿的差别：前者的词里含着一个日期，後者的词里没有。

十三·准清偿：审计仪式的量纲解释

$\lambda = 0$ 而 $i > 0$ 时, K 发散。发散必须被吸收。本文把主要的吸收装置命名为：

定义 **8** (准清偿)：形式上具备裁决外观 (有评价者、有意见、有结论、有仪式)，但不使 K 可核减的动作。判据只有一条：这次动作之后，当事人能不能指出一笔已经关闭的账？指不出来，就是准清偿。

现代学术与文艺共同体里的准清偿清单：同行互评、工作坊、研讨会、推送、赠书互引、年度述职、项目中期检查、各类“入选”“提名”“推荐”。

准清偿不是骗局。它们通常真诚，也常有实质内容 (意见有用、讨论有益)。它们的问题只在会计上：它们制造承认的流量，不制造存量的核减。

这里可以与米歇尔·鲍尔 (Michael Power) 的《审计社会》做一次正面交手，因为那是本文最近的一位祖宗。鲍尔 1994/1997 年指出审计活动在八十年代之后爆炸式增长，形成“验证的仪式”，并诊断出脱耦 (组织为取悦审计者而设计系统，而非为改善绩效)。鲍尔的诊断在现象层几乎与本文的准清偿重合。

分离线在于：鲍尔解释了审计为什么被要求 (政治问责的需求)，没有解释审计为什么停不下来。在他的框架里，审计膨胀是需求驱动的，因此原则上可以由需求下降而收缩。本文给出的是一个量纲原因：当 $\lambda = 0$ 而 $i > 0$ 时, K 单调发散，而共同体必须提供与 K 同步增长的承认流量，否则成员的体验会崩溃。因此审计仪式的密度不是由问责需求决定的，是由未清偿存量决定的；只要 f 不变，减少审计的努力必然失败——需求会在别处以新的仪式重新出现。

这条是批判的：如果在一个明确的领域 (例如授权司法裁决、比赛、资格考试) 观察到与外界同步的准清偿仪式，则本文解释为，鲍尔的解释出错了。这是本文的第一条正路径，也是最容易的一条：这些审计仪式 (互评“提名”“入选”类事件) 年度还是疏的

最后要把准清偿与一个更常见的批评区分开。人们常说学界的问题是“评价过多”。本文的判断相反：评价不是过多，是从不结束。一年一次、终局的评价会比现在的密度低得多，而它带来的秩序会比现在多得多。数量不是变量，终局性才是。

十四·研究一：一次预注册真跑，三条假设全部失败

前面十三节全部是论证。本节起是数据。

本文自己的假设：核心机制必须有一条可以被现成数据打死的通路，并且预先写下判死条件。研究的设计预注册在写数据之前 (全文附录 A)，此处如实报告其失败。

14.1 设计

如果“一次外部裁决关闭账目”这条假设成立，那在一个能清晰观测文本何时被裁决的群里，应该看到：被外部裁决过的文本，其后续改动会停；未被裁决的文本，会拖长尾。

语料选择 arXiv，理由是它同时记录三样东西：首次沉积日、每一次版本更新、以及作者回填的期刊信息 (journal_ref)。后者被本文用作“存在一次外部一次性裁决 (被期刊接受)”的代理。

- 样本 2016 年6 月首次投递的论文，取 cs、math、cond-mat、astro-ph、q-bio 五个大类，去重后 $N = 1,735$ ，账龄超过九年。
- 读数 $\tau = \text{最后一次版本更新日} - \text{首发日 (天)}$ ； $V = \text{版本号}$ ；是否有 $\text{journal_ref } j=1$ 共 610 篇，占 35.2%)。
- 预注册假设
 - **H1** (主) $R_3 = P(\tau > 3 \text{ 年} \text{ 在 } j=1 \text{ 组显著低于 } j=0 \text{ 组, 差} \geq 5 \text{ 个百分点, } p < 0.01)$
 - **H2** $P(V \geq 3) \text{ 在 } j=0 \text{ 组不低于 } j=1 \text{ 组。若明显更低, 说明多版本主要由投稿流程驱动, 对本文机制不利, 必须照报。}$
 - **H3** $P(\tau > 5 \text{ 年} \text{ 在 } j=0 \text{ 组高于 } j=1 \text{ 组}$
- 预先声明的混淆 journal_ref 由作者自愿填写，而非本身带有一次版本，其方向与 H1 相反。故 H1 若成立结论更稳；H1 若失败，只能操作失败，不能直接报错

14.2 结果

读数	J=1 (n=610)	J=0 (n=1125)	差	z	p
P($\tau > 3 \text{ 年}$)	2.1%	1.7%	+0.4 pt	+0.6 5	0.51
P($\tau > 5 \text{ 年}$)	0.7%	0.9%	-0.2 pt	-0.5 2	0.60
P($\tau > 1 \text{ 年}$)	9.5%	8.4%	+1.1 pt	+0.7 5	0.46
P($V \geq 2$)	50.7 %	38.1 %	+12.5 pt	+5.0 4	4.7×10 ⁻⁷
P($V \geq 3$)	16.4 %	11.5 %	+4.9 pt	+2.8 9	0.0038

分类别复核 (R_3 的组差: astro-ph +0.7pt、cond-mat -0.1pt、cs +1.1pt、math +1.8pt、q-bio +1.0pt。五类无一达到预注册的 5 个百分点门槛，方向亦不稳定。

判定 - **H1** 失败。差值 0.4 个百分点，远低于门槛，且不显著。 - **H3** 失败。方向相反，不显著。 - **H2** 失败，且方向明确： $j=0$ 组多版本更低 (11.5% vs 16.4% , $p=0.004$)。这是预注册中对不利的那种失败

三条假设全部失败。这个结果需要一次诚实的处理，而不是一次修辞的挽救。

14.3 三条失败各自意味着什么

H2 的失败已有文献佐证，且不是本文的发现。关于arXiv 预印本的研究 (Lin 等,2020 ,*Scientometrics*) 已经报告：已发表的预印本，其“零修改”比例低于未发表的预印本，而多版本比例高于后者，并给出两条解释：反复修改提高了被接受的概率；被接受之后作者上传最终版。本文的读数与该结论一致，因此本文

在这一点上只是一次独立复现，不主张新发现。这条纪律必须写明，否则本节会变成用旧结论装饰新概念。

H1、H3 的失败告诉我们的东西更重要，且它不在预期之内：整个语料里， $\tau > 3$ 年的论文只占 **1.84%** (32/1735)。也就是说，不论有没有外部裁决，学术文本在沉积三年之后几乎不再被作者动过□

细看 τ 的分布：

分位	值
中位数	0 天
75 分位	126 天
90 分位	333 天
99 分位	1,558 天
最大	3,640 天

以及：**57.5%** 的论文只有一个版本 ~~从未改动~~，**63.7%** 的论文在首发三十天之内就完成了它一生中最后一次改动□

H1 之所以失败，不是因为“外部裁决不关闭账目”，而是因为账目在裁决到来之前就已经关闭了。本文的设计假定文本会持续处于可改状态，等待一次裁决来叫停它。数据说：文本根本没有在等；它一落地就基本不动了。

十五·由不利结果改判：到期性是主体账目的属性

失败的假设如果只被记录而不被使用，那它只是一次事故。本节做的是使□□

改判如下：

原命题（含糊形态）：现代学术的缺永不到期，因此作品被无限期地悬置在“还可以再改”的状态。 修订命题：到期性是主体账目的属性，不是作品的属性。作品几乎在沉积的瞬间就被冻结；缺不挂在作品上，挂在人身上。人不断投入新的作品来续期自己的账，而每一件作品本身在投出去的那一刻就死了。

这个改判比原命题硬，也比原命题更符合日常经验：没有人会因为职称没评上而回去修改五年前的那篇论文，人们的做法是再写一篇新的□K 的增长机制由此变得清楚： $\lambda = 0$ 使账目不关闭，而账目不关闭的应对方式不是修旧作，是加新作——这恰好是方程 1 中 $i(t)$ 不减反增的原因。于是可以补一条：

命题 9（存量的自增机制）：在 $\lambda = 0$ 的条件下，主体来清反是期投入修正投入，因此 i 与 K 正相关，方程 1 从线性变为加速散

这条命题解释了一个常被归因于“内卷”的现象，但给出的原因不同：产出加速不是因为竞争者变多，而是因为没有任何一次产出能够结账。它可以被区分检验：如果产出加速主要由竞争驱动，则在竞争数量下降时应该观察到产出减速；如果由 λ 驱动，则只有终局性提高才会减速。

同时，本项研究里唯一显著的读数接过来用，因为它虽然不是新发现，但在本文框架里有一个新用途：

何东西。这不是劝人清心寡欲——命题 10 给出的是一个列表,选哪一个是各人的事,但选错的人一定失落

十七·研究二：两个词之间的十年时差

研究二命题3（名分双层）：给残余物命名与名分生效是两个可分离事件，后者滞后于前者。

17.1 设计与预注册

- 语料□Google Books Ngram 法语2019 版语料，英语 2019 版语料），逐年词频，未做平滑。
- 观测词□avant-texte 物的名字、critique génétique 籍职业/学科的名字。
- 控制词□brouillon（草稿）、manuscrit（手稿）——若这两个词期同上升,则切归可于语料结构而非本机制
- 预注册判据（H4）：词族年产量的拐点应落在 1982 年（ITEM 建所）之后,而非1966-1968（海涅手稿入藏、CNRS 小组成立）之后 判伪条件:若1968-1981 年均值 ≥ 1982-1995 年均值的 50% ,H4 伪

17.2 结果（每十亿词）

词	1968-1981 年均	1982-1995 年均	比值
avant-texte	59.3	246.6	0.241
critique génétique	15.6	154.7	0.101
brouillon（控制）	1,910.7	2,249.5	0.849
manuscrit（控制）	37,436.6	35,992.4	1.040

两个控制词比值分别为0.85 与1.04 :在同语料时段,“草稿”与手稿这两个词没有何相应的升 因此两个术语的上升不是语料结构造成的。两个观测词比值远低于0.5 的界线,H4 按预注册判据成立□

17.3 但逐年曲线说了更多

按预注册判据H4 成立,然逐年数据跟着一个本文没有预期的错,它 □ H4 本身更得告:

年	avant-texte	critique génétique
1965	4.6	3.8
1969	1.4	8.2
1971	41.0	5.1
1977	98.7	6.5
1979	185.8	36.9
1983	178.6	59.4
1985	250.0	134.2
1993	317.3	306.3

2001

270.6

318.5

avant-texte 的起飞在 **1971-1979**：这与贝尔曼-诺埃尔 1972 年使用该词 1978 年马特拉菲赖德会议以之为题完全吻合，早于 **1982** 年建所 *critique génétique* 的起飞在 **1983-1985**，并在 **1993** 前后再次跃升——分别对应实验室成立与专业期刊《Genesis》创刊（1992）。

两条曲线的拐点相差约十年，方向正是命题3所预测的：先有给物的名字，后有给职业的名字；前者由一个人在一篇文章里造出，后者要等一间实验室与一本刊物。

由命题3可以写得更细：

命题 3'：名的两层各自的起源 命名层由个体制造驱动，成本低，可以随时间发生扩散和学术尚；生效层由清选驱动，成本高，必须等待建制，扩散依赖于出版通道的两层之间的带滞，终等把一个研究组变成一个授学位可发工资可期的实体前，同时必须与本院的条制，它们足以弱结论：

1. 语料尾部失真□Google Books □ 2009 年之后的覆盖显著变薄，本文观察的2012-2019 年下滑 (*avant-texte* 降至47.9) 几乎肯定是覆盖伪影，不可用作“名分衰退”的证据；本文因此把全部推断限制在 1965-2005 区间。
2. *avant-texte* 的法语一般义污染。该词在非专业语境中偶可作“文本之前的部分”解，但不创造1971 年那个八倍跃升。
3. 词频不等于名分。词频只是名分的一个可得代理；更好的代理是职位公告、学位论文题目、期刊征稿启事中的出现率，这些语料本文未能获取。若这些更直接的代理给出与词频相反的时序，本研究结论应予推翻。

十八·归隐的制度条件与可操作推论

现在可以回答最后一个问题：为什么现代出不了归隐者。

归隐不是退出，是清偿完毕之后制度自动放人的手续。古代文人的归隐之所以能成为一种正式身分（有名目、有礼仪、有诗体、有社会承认），前提是他的账可以结清——注意者同样是清算，因为账目高，落第是一个已经了结的循环，绝意仕进是对该循环的正式退出声明，而不是逃避。

命题 11（归隐的制度条件）：一个共同体产生正式的退者，当且仅当该共同体的 $\lambda > 0$ ，且清算作主侧目而视清算作主生效 $\lambda = 0$ 的共同体不生退者，只生失落与抱怨者

失落与抱怨不在境，在账目状态：归隐者的账是平的，失落者的账还挂着。这就是为什么现代世界的常人被一种不清的债缠住——他没有被逐，也没有被请，他只是让账期，账还朝里挂着，随时可以被自己或人重新开（“其实再发次就……”）。

由此可以给出若干可操作推论。它们是设计规则，不是劝告，每一条都附带失败条件：

(1) 把裁决对象缩小到“一笔”，不要放大到“人”。可清偿必须明确的诉求的第十一节。“评价学者不能请，这次申报，结清可以失败条件：若某机构了单笔请而其成员的未清偿还下降，本条伪。

(2) 终局性优先于程序公正。在障碍时，先付 提法 不设计再报，再收 a 与 p。理由见定义 7 中 $f = 0$ 时 λ 恒为零，对 a 与 p 的任何改良都不改变结算。失败条件：若某也还改的前提下仅通过提高匿名与公开就显著降低了准清偿仪式密度，本条伪。

(3) 可重开事由必须是封闭清单，且不得包含“有更好的新作品”。这一项是终局性的唯一致命漏洞（第十一节第 2 点）。

(4) 为退出设计一个正式手续。 $\lambda > 0$ 只是必要条件，还需要一个可以说出口的目的：没有目的退者被同时排除出学术清成员。

(5) 对个人写作者：不要给自己的灰编号。根据命题 2，自编号不产生名分；根据命题 4' 与 10，它反而先行支付了可变性的代价——你把自己的灰冻住了，却没有换来任何人的清偿。个人写作者对灰唯一有效的做法是第三节的回读：读它、说出它差在哪、把那句话喂进下一稿。失败条件：若能找到一位以自编号为唯一手段、其残余物在无外部职业通道的情况下获得并保持了三十年以上名分的者，命题 2 与本条同时伪。

(6) 如果你要的是名分，就去找那条职业通道，而不是找一双眼睛。关注是假的，因为关注不涨；把自己的材料给某一群人的可清偿对象才有效——他的要害靠材料消结，你名被生了。这句话听起来利，其只是五十年前的实践形式：1968 年那几个 CNRS 位置，比 1885 年到 1966 年之间的所有关注加起来管用。

十九·与十五位最近者的分界

本条处理最后的手。判断先：删掉这一位，本文还剩什么；删掉本文，这一位缺什么。

1. 发生学批评 (Bellemain-Noël 1972 ; Hay ; Grésillon ; de Biasi ; ITEM 1982)。最近，且本文的核心材料就是它。分离线是层级：他们研究灰，本文研究他们。发生学批评从不问“我们这门学科为什么在 1982 年而不是 1885 年成为可能”，因为对任何一门学科而言，自身的建制条件都是它的盲点。本文把它当作被解释项，得到的读数（第十七节两条曲线）在发生学批评内部无法产生。反过来，删掉本文，发生学批评毫发无损；删掉发生学批评，本文失去唯一的正例。互补，不可替代。

2. 布尔迪厄 (1971/1992/1993)。最贵的一位。“神圣化 (consécration) 与“象征资本”处理的正是谁有资格授予名分。分离线有二：其一，布尔迪厄的场域里没有到期这个变量——资本可以累转化贬值，但有一个日期某笔账闭；他还在欠账，在本条里却是 $\lambda = 0$ 的那种情形，但他把它当作减损性，本把它当作变数中 f 的一个取值，因而是可变的。其二，布尔迪厄解释神圣化者的利益，本文解释神圣化者的债务。ITEM 的学者不是因为资本才成名，是因为自己的要害来结。这两解释在 1968 年那几个位置上给出不同预测：利益模型预测有资本者优先命名（应由索邦的讲席教授命名），债务模型预测由账最紧的人命名（应由新设岗位上的年轻研究者命名）。历史上是后者。这不是决定性证据，但它是一条可查的分界线。

3. 默顿（1957 年优先权之争；承认体系）。默顿的承认是稀缺资源的分配问题，其病理是优先权纠纷；本文问的不是资源够不够分，是判决能不能关账。一个共同体承认其成员有账而 λ 仍为零——这正是当代的形态，默顿的框架读不出这种病。

4. 迈克尔·鲍尔《审计社会》（1994/1997）。已在第三节处理，分离线：需解释 vs 有解释；已给出判例实验

5. 埃斯佩兰与索德（2007，反应性）。排名改变被排名者的行为——这是关于测量的应。本文关心的不是测量如何改变行为，是判决为何不终止。一个排名可以完全无反应性（没人在乎）而 λ 仍为零。两条线交叉但不重合。

6. 阿伯特《职业系统》（1988，管辖权）。最危险的一位，因为“终审权易主”听上去就是管辖权转移。分离线很干净：管辖权讲的是谁有权处理某类问题，本文讲的是处理之后能否关账。一个职业可以完整垄断管辖权而其内部判决完全不可清偿——本。阿伯特的理论无法预言 λ ，因为他的框架里没有终局性维度。

7. 惠特利（1984，任务不确定性与相互依赖）。惠特利用任务不确定性解释判例评价方式的差异，这与本文的 α （认账系数）相邻。分离线：任务不确定性解释判决是否被接受，不解释判决是否被允许重开。数学的任务不确定性低而其职称评审同样可再审。

8. 卢曼与法学的既判力传统。本文从这里搬结构（第十一节）。分离线在方向：法学处理的是既判力的边界与例外（何时允许再审），本文处理的是既判力缺席的后果 $\lambda=0$ 的共同体如何运转）。据本文所知，既判力从未被用作分析学术评价或文人共同体黏性的工具；若已有此类工作，本文第十一节应改写为综述。

9. 格雷伯《债》（2011）。“永不到期的债”这个说法与本文极近。分离线：格雷伯讲债的道德语法与暴力起源，其“不可偿还”是货币史与政治史现象；本文的不可清偿是裁决制度参数的后果，可以用参数写出来，也可以通过 f 而消除。格雷伯的债没有 λ 。

10. 霍耐特（承认理论）。承认是主体间的伦理条件；本文的清偿是制度的会计条件。一个人可以在伦理上被充分承认而账目分文未清（当代学界最常见的痛苦形态正是这一种），这一点霍耐特的框架读不出来。

11. 德里达《档案热》（1995）与斯蒂德曼《尘》（2001）。档案的权力、命名者（archon）与死亡驱力。分离线：德里达处理档案的欲望结构，本文处理档案的会计功能；本文第十六节的冻结律与《档案热》的“档案即死亡驱力”在直觉上相邻，但前者是一条带百分数的曲线，后者不是可证伪命题。

12. 版本学与校勘学（Bowers、Tanselle、McGann；中国的稿本、异文、校勘）。已在第四节处理，分离线：他处理为定本，本文处理为订本。McGann 的社会文本理论最近，因为它把作品视为社会生产的产物——但它仍不问“研究这些痕迹的人自己的账是怎么结的”。

13. 写作过程研究（Sommers 1980；Flower & Hayes 1981；击键记录 Inputlog/ScriptLog）。这族已经精确记录删段，分离线：过

程研究作者行为,本文裁决制度;且本文第五节的数字族是有用的;外裁大场重新开稿提高12.5个百分点,意味着研究若观察前裁为写情境,会系统性低估量

14. 科学计量学中的预印本版本研究 (Klein 等 2016 ; Lin 等 2020 ; Akbaritabar 等 2022) 。本文研究的H2 结论与之一致,属独立复现而非新发现,已在第十四节写明。分离线:这一族解释版本数的成因(投稿流程),本文用同一批数据读的是冻结速度与回读的外因性,是同一数据的不同用途。

15. 中国文论(陆机、刘勰、贾岛、黛玉焚稿、扬雄覆瓿)与西方古典诗学(贺拉斯、瓦莱里、Quiller-Couch“杀掉你的宠儿”)。已在第一、四节处理。瓦莱里最近:“诗永远没有完成,只有被放弃”,这一句把正稿从顶点降为灰序列的截断点,是名分方向上最远的一步。分离线:瓦莱里给出的是作者的主观状态(放弃),本文给出的是制度的客观结构(谁签发关闭)。第十六节的冻结律恰好可以反过来读瓦莱里:诗不是被作者放弃的,是被编号关闭的——中位数零天。

共有前提的第二层。上面五位——连本文——共享一个前提:残余物的地位问题可以在“看与不看”“重视与不重视”的框架里讨论。本文否定的是这层:看不看是量,签发与签发提,这是最容易被播的地方,也是谁值得存在理由

(十六) 自己不能当自己案子的法官 (nemo iudex in causa sua ; 渊源于罗马法,经柯克在 Bonham 医师案 (1610) 中定型:一个不能同时是 judges, ministers, and parties) 。这一位必须补上,而且它就在本文借来的那门学科里面。定理 1 的前提 $A \subseteq B$ ——判者集合包含于被判者集合——而这正是 Bonham 案的事实结构:伦敦医师学院既立规、又执行、又从罚金中分成,柯克据此认定它不能审自己的案子。任何读到定理 1 的法律人都会先想到它,原稿一字未提,是一处欠账。

分离线是硬的,而且正好是本文与自然正义传统的分工: nemo iudex 判的是偏私——有利害关系者不得任判者,救济方式是回避、更换判者、撤销原判;定理 1 判的是终局性——即使判者毫无偏私、程序完美、人人真心想结账,只要 $A \subseteq B$, 判决就不账,两者量不同:前者在 α (认账系数) 那一栏,后者落在 f 那一栏。

这条分界顺带解释了一件本文原本只能描述而没有解释的事:为什么学术评价改革做了几十年的回避制度、利益冲突声明、双盲评审,缺存量却一点没降——那些错是在 α 上用力,而 $\lambda = c \cdot f \cdot \alpha$ 里, f 为零时 α 再高也是零。

判决性对照: 找一个已经把 nemo iudex 落进极致的评度——判者被告完全回避严格行利益冲突披露,若判度的准清密度显著低于领域对照组,则本文的假分为伪,那说明偏左是因 f 是多余变量;若判度与对照组无别,而有引入固定期与不再之密度下降,则这对这一条第二节F5 是同伴作的一个入口

一并修于处:定理1 的 ■ 记号自本版起撤回。本文第十节以 ■ 收一个纯语词论证,而 □□□□ F3 又自认“制度明文禁止重开”这一类反例可能存在——定理名与论证强度

不匹配。自本版起，定理 1 降为命题 8（不可自签），其论证保留而不再声称完备，F3 保持不变，仍是它的直接靶子。

十之二· α 的算法，与四维度第 1 次算出的 λ

本文形式层最重的一处欠账： α 自始至终没有测量法，因而 $\lambda = c \cdot f \cdot \alpha$ 全文一次也没有算出数值，方程 1 也从未代入数字。一套中心量算不出数的形式系统，在可证伪性上比它看起来弱。本节补上这一步。它不是一次真跑，是一份可复算的编码规则加四次代入——任何人可以拿去重算并提出异议。

α 的量化四项，每项取 0 或 0.5 或 1，取均。判据一律公开查件文件，不取当事人感受：A1：判决是否经果并送达当事人 = 1；只给结论 = 0.5；只示名单 = 0。A2：判决是否由一个与原告不同的事由法院签发 = 1；同一机构不同部门 = 0.5；由原告所在单位签发 = 0。A3：评委会与回避规则是否公开（全开 = 1；事已公开 = 0.5；始终不公开 = 0）。A4：判决是否进入一份第三方查公共录（进入 = 1；内部材料用查阅 = 0.5；不留公共录 = 0）。

c 的数值：本节律对当事人每发生一次为 1，三次为 1/3，六次为 1/6，不进为 0。“ f ”的读数：不重开 = 1；仅程序可申诉 = 0.9；实质重开须经再审理 = 0.3；随时再报 = 0。

科举（年科）：A1 = 0（只张榜，不给理由）、A2 = 1（临场时，非考生所在县签发）、A3 = 0.5（考官锁院，名单事后可知）、A4 = 1（题名录为公共录）， $\alpha = 0.625$ 。 $c = 1/3$ ， $f = 1$ （本不重判，落第者下科再申诉科）。 $\lambda \approx 0.21$ 每年。

现代同行评审（滚动投稿）：A1 = 1（评意见送达）、A2 = 0.5（编辑部与评审同属一个学术共同体）、A3 = 0（评委会公开）、A4 = 0（拒稿不留公共记录）， $\alpha = 0.375$ 。 $c = 0$ （不定期，由原告随时发起）。 $\lambda = 0$ 。须注意：即便把 c 按平均每次读 1， $f = 0$ （可再投、可再审）仍使 $\lambda = 0$ 。

职称评审（年年评）：A1 = 0.5，A2 = 0.5，A3 = 0.5，A4 = 0.5， $\alpha = 0.5$ 。 $c = 1$ ， $f = 0$ （次年再报）。 $\lambda = 0$ 。这格是第 1 期向数值状态：前项使倍增满， λ 仍然零。

终身教职（up-or-out）：A1 = 1，A2 = 1（校董签发，评委会由获终身教职者组成）、A3 = 0.5，A4 = 0.5， $\alpha = 0.75$ 。 $c = 1/6$ （六年考期次）， $f = 0.9$ （实质答辩可申诉，申诉限程序再评）。 $\lambda \approx 0.11$ 每年。

三条扯出来的东西，都是本原理能定性的。其一，四维度里有两个 λ 大零，而两个的 λ 相差约一倍（0.21 与 0.11）：科举的精英高于终身教职，来源是更正—— α 反而更低（0.625 与 0.75）——而是更繁更麻烦账。第 3 节程序正义目标是所变量，这样断，至变成了一个可以读数。

其二，方程 1 现在可以代入取现代同行评审 $\lambda = 0$ ，得 $dK/dt = i$ ，缺量输入线发散，无稳态；取科举 $\lambda = 0.21$ ，得 $K^* = i/0.21 \approx 4.8 i$ ——稳态量等四年四个月的量，与年科制同量级。这是本文第 1 次出一个可以被反证函数：若能证明科举平均稳态量显著偏离这个量级，本节的编辑修订。

其三，一条本文体系的 α 的四项有三项（A1 与 A3 与 A4）在终身教职高科举，而 λ 反而更低。这期 α 在 λ 里靠未置量能略轻重：若 α 只影响判决被账，而影响增速率，那么不当出现在 K 的初值 K_0 的算里，不在

λ 里本就不该公式,但这一妨为第二十节的第条伪证 F9:若两个 c 与 f 相同而 α 为极小的假之间,准清偿密度无显著差别,则 α 不出现于 λ 的乘积中,定义 7 须写

二十·证伪清单与局限

本文可以被打死的地方,按杠杆从大到小排列。

(F1) 准清偿的量纲解释。若在明确的领域(如授司判决、执讼格式、竞技场、观察与学校同等密度的准清偿式,则第三节的量解释为,绝对需清偿出 执行成本低:这些领域的审会议互评活动 提名事件密度统计

(F2) 名分的清偿链条件(命题 2)。若除出一类废物,在不存在任何以其为清偿对象的职业通道的情况下,获得并保持名分超过三十年(有专名、进目录、成为学位论文题目),则命题 2 伪。候选反例值得认真找:民间收藏门类、网络社群自建的分类体系(例如同人写作社群对“废弃设定”的命名),这些是本文自认最危险的地方。

(F3) 不可自签定理(定理 1)。若能举出一个判者集合完全包含于被判者集合、而其判决确实不可重开的制度,定理 1 伪 注意:实际很重平不算——定理讲的是资格,不是频率。

(F4) 终身教职后评审预言。若在强制推行终身教职后评审的体系与未推行的体系之间,终身教职之后的产出模式无显著差异,则第十节的推论伪。

(F5) 四参数的分工。若也改 f 的前提下,仅通过高 a 与 p 就显著降低了准清偿密度或员策清偿感,则定义 7 中 f 的秉性伪

(F6) 冻结律(命题 4')。若另类料(例如视频修片版本 编写平台 观察必可见生不修率缺陷,则冻结律适用范围成为带定标准沉积,而非一切开

(F7) 名分双层(命题 3')。若用职位公告、学位论文题目、期刊征稿启事作理,得时序与词频理反(即职业先物名),则命题 3' 伪。

(F8) 本文自身的自反检验。本文第十九节的全部界由本文方出,来任何仲裁。按文自己的定理 1,这成一个 $A \subseteq B$ 的构造:本文自己的判决不具备终性 因本文还害怕自己的评价,只出自己的反致性;本文用两个数据与左脚脚底交付(附录 B),任何人可以复沓并翻

局限,四条:

1. 灰的绝大部分不可观测。第三节指出,写但最看的是笔之前阶段想法,它不留痕迹 本站全部制假析只适用已落物质痕必那一部分,且这一部分占比极知
2. 两次真跑的代理都是弱代理。journal_ref 不等于“被裁决”,版更不等于“回赎”,词频不等于“名分”。本文每处皆作声明,但声明不能替代数据
3. 文化范围。科举一现代对照是国选,洪堡一同训是辨选,两者被拉同一条线上,中间隔了大量也差异 如日本 苏为评价作假提供不同 f 取值,本校被用。
4. 规范性沉默。本文只说 $\lambda = 0$ 会散,没说 λ 应多大 终身性有代——判错就是错了 第 0000 3 点。一个 λ 过高的共同体别的式受损害 其非杀路理化,本文最值无和

二十一·结语:灰不等任何东西

回到最初那个问题:何谓作文。

本文给出的回答不再是“作文是什么”，而是作文之后剩下的那些东西是什么身分。答案分三层：

第一层，在写作里。灰不是料也不是料，它成形——你要那个第一次露面，不是它自己的形状，而是以串都是的形状。灰不能被费的方式，是不被阅读：那样它连那定资格被败。

第二层，在制度里。施容获得名分，与多以其无关，与第人靠它有关。两百年里没有一个统一灰名分，不是因为没人看见，是因为从来没有出现过批人，其身的要需这些残物才能请 1968 年去给个位置，1972 年有了词，1982 年有了实室，1992 年有了期——名就是这样来的，前后不到二十年，而在之前的半里，林齐而移生。

第三层，在代价里。名分不是白给的。它的价格是可变性，而且是当场支付、几乎瞬间生效：三十天内六成的文本一生中最后一次改动已经完成，三年内九成八已经完成。编号这个动作本想留住灰，实则是给它提前举行葬礼。

于是回到那个最冷的判断。当一个人问“我还缺什么”，他其实是在把自己的灰当作品，在等一个谁来认领它。而灰之灰，恰恰在于它不再有什么东西——不必成文，不必定型，不必被谁查来。这两条路是并行的：一条通向被研究的资格，一条通向被编辑的自由。本文能做的全部工作，是把这两条路的价目表算清楚，让选的人知道自己在付什么。

至于那个更冷的问题——什么时候有人来查你灰——本文的回答是一句太冷的话：等一个不相的人对你好奇，是早期的；而等一条线从林齐而移生的侧通道起来，是后期的，那期限终究是一个研究组中间现实的距离。

附录 A · 预注册全文（写于取数之前）

研究一：裁决的一次性与账目关闭 样本:arXiv 于2016-06 首次递 (v1) 的论文,取cs / math / cond-mat / astro-ph / q-bio 五大类,每类目标400 篇 读数: τ = 最后版本更新日 - 首日 (因); V = 版本号自id $\square$ vN 取; J = 是否有journal_ref (代理“存在一次部次裁决-被期接受”)。H1 由: $R_3 = P(\tau > 3 \text{ 年} \text{ 在 } J=1 \text{ 组} \text{ 且 } \text{显著低 } J=0 \text{ 组}, \text{差} \geq 5 \text{ 个百分点})$, 双侧两比例 $z \square \square p < 0.01 \square$ H2 $\square P(V \geq 3) \text{ 在 } J=0 \text{ 组} \text{ 低于 } J=1 \text{ 组}$ 。若明显更低 $\Rightarrow$ 多数论文主要投稿到近期而非缺不到期,对本文有利,必须报 H3 $\square P(\tau > 5 \text{ 年} \text{ 在 } J=0 \text{ 组} \text{ 高于 } J=1 \text{ 组 (长尾更厚)})$ 。预先声明的混淆: journal_ref 由作者自愿回填,且回填动作常随次新版本,方与H1 相反;故H1 若成立结论更稳, H1 若失败只能判该操作化失效,不能直接判机制制错。

研究二：名分的时点 用书籍语料统计 avant-texte / critique génétique $\square$ 年度发布 H4 :该词每年产量出现在1982 (TEM 建所,职通造成) 之后,而是1966-1968 (海涅手稿入藏、CNRS 小组立,林齐而移成) 之后 判伪:若1968-1981 的年均量已达1982-1995 年量的50% 以上,H4 伪。

执行结果 $\square$ H1 失败 H2 失败 (不符)、H3 失败 H4 成立 H1-H3 的失败已第四十五节用批判,未得后续。

附录 B · 数据与复跑

- fetch_arxiv.py 按每月按月份抓取arXiv Atom 接口 (每页100 条,间隔3.5 秒;版本号自<id> $\square$ vN 取,journal_ref 取自arxiv 命名间)。

- an.py :去重 (按版本号id)、计算 τ 与V、两例z检验、按大类分层复核、 τ 分桶。
- ng.py / ng2.py : 书籍语料逐年词频 (smoothing=0) , 含两个控制词。
- 中间数据arxiv.json (1,804 条原始,去重后1,735)与arxiv_clean.json 随文付

复验三条: (1) arXiv 接口有频率限制,须设UA 并留3 秒上间隔; (2) 书籍语料 2009 年之后覆盖,任何跨2009 年的趋势都不可靠

参考文献

1. Abbott, A. *The System of Professions: An Essay on the Division of Expert Labor*. University of Chicago Press, 1988.
2. Akbaritabar, A. et al. Reference changes in preprint-publication pairs. *Scientometrics*, 2022.
3. Bellemin-Noël, J. *Le texte et l'avant-texte*. Larousse, 1972.
4. Bourdieu, P. Le marché des biens symboliques. *L'Année sociologique*, 22, 1971.
5. Bourdieu, P. *Les Règles de l'art: Genèse et structure du champ littéraire*. Seuil, 1992.
6. Bourdieu, P. *The Field of Cultural Production*. Polity Press, 1993.
7. de Biasi, P.-M. *La génétique des textes*. Nathan, 2000.
8. Deppman, J., Ferrer, D. & Groden, M. (eds.) *Genetic Criticism: Texts and Avant-textes*. University of Pennsylvania Press, 2004.
9. Derrida, J. *Mal d'archive: une impression freudienne*. Galilée, 1995.
10. Espeland, W. & Sauder, M. Rankings and Reactivity. *American Journal of Sociology*, 113(1), 2007.
11. Flower, L. & Hayes, J. A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 1981.
12. Graeber, D. *Debt: The First 5000 Years*. Melville House, 2011.
13. Grésillon, A. *Éléments de critique génétique*. PUF, 1994.
14. Hay, L. La critique génétique: origines et perspectives. In *Essais de critique génétique*, Flammarion, 1979.
15. Honneth, A. *Kampf um Anerkennung*. Suhrkamp, 1992.

16. Klein, M. et al. Comparing published scientific journal articles to their pre-print versions. *International Journal on Digital Libraries*, 2018.
 17. Lin, J. et al. How many preprints have actually been printed and why: a case study of computer science preprints on arXiv. *Scientometrics*, 124, 2020.
 18. Luhmann, N. *Legitimation durch Verfahren*. Luchterhand, 1969.
 19. McGann, J. *A Critique of Modern Textual Criticism*. University of Chicago Press, 1983.
 20. Merton, R. K. Priorities in Scientific Discovery. *American Sociological Review*, 22(6), 1957.
 21. Power, M. *The Audit Explosion*. Demos, 1994.
 22. Power, M. *The Audit Society: Rituals of Verification*. Oxford University Press, 1997.
 23. Sommers, N. Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication*, 31(4), 1980.
 24. Steedman, C. *Dust: The Archive and Cultural History*. Rutgers University Press, 2002.
 25. Tanselle, G. T. *A Rationale of Textual Criticism*. University of Pennsylvania Press, 1989.
 26. Valéry, P. *Œuvres I* ("Au sujet du Cimetière marin"). Gallimard, 1957.
 27. Whitley, R. *The Intellectual and Social Organization of the Sciences*. Clarendon Press, 1984.
 28. 陆机《文赋》；刘勰《文心雕龙·铨裁》；桓谭《新论》（“覆瓿”）；郑燮《题画》（“胸中之竹”）。
 29. 关晓红《科举停废与近代中国社会》，社会科学文献出版社，2013□
 30. Ben-David, J. *The Scientist's Role in Society*. Prentice-Hall, 1971.
-

数据与脚本随文交付；两次真跑均可复跑。本文的分界部分不构成对自身的裁决（见 **F8**）。

第九章 补节 灰为什么没有名分

本章v1.1 已以本卷最低的分界（「任何不得为本案的法官」这条马规则共六（六）年的实例，并第一次给出账数测量法，在四页篇幅上算出了清算，还批驳了一处量的证明与 降谷命题 补只做件事

（一）本章在全卷的位置

本章是编，负责前章同：前章有指人她的空格，本章解释为什么这些空格不会因为被指认而被填上□

这个衔接需要明写。前八章的读者会自然产生一个改良主义的期待——既然我们现在知道了未生界、预闭分叉、悬跨段的存在，那么增设字段、改造量表、扩充过程档案即可。本章否定这个期待，理由不是技术性的：名分不是注意力的产物，是清偿制度的溢出物；一类残余物获得名分，当且仅当存在一条以它为清偿对象的职业通道。手稿早在一八八五年就已成规模入藏，而写作残余物获得专名要等到一九七二年、真正生效要等到一九八二年一间国立实验室的设立——中间隔的不是九十年的技术进步，是一九六八年第一批靠这件事领薪水的职位。

因此本卷的结论是一个条件句而不是一个纲领：前八章的九个读数要真正进入任何一份账本，需要的不是更好的理论，是一条从未建成的通道。

（二）本章的自查，与本卷的自查

本章v1.1 新辟一条自己的自查：账数有项终身跳出开秤，而清算反而更低，因此账数乘法置能重——若只管删而不管速率，应当删掉处。这条为新证路径 编者认为这是本卷最学习的一次自我理：发现自己形式系统的一处参数位置可能错了，不是删掉它，而是把它写成一条可以判决的新路径。

本章的自反条款同时也是本卷的自反条款。本书自己就是一份「判者集合包含于被判者集合」的产物：九章的分界全部由作者方单方作出，未经任何外部裁决；本卷第六章与第九章的作者更直接参与了本书的编纂。按本章自己的命题，这类文本的终局性为零，因而不能对自己签发结论。

本卷据此不作任何「已经成立」的宣告。它给出九个数 九死条件与 两可以替互删消的切，然后在这里

（三）本章欠账

F1 仍跑，且本章明原因——它需要新集利 司法执考式 领域的式密度数据，而现有的及里部 被删长这竟解释染，跑出来可判 这是本章最宜 也最判快的条，它至今跑

清算自算出的四数 科举0.21、现行审0、联邦审0、终身判0.11）依赖认账系数的四项编码，而这四项编码是本章自己提出的，尚未经任何第二人独立复算。编者建议成书后优先安排一次独立复算——这既是对本章的检验，也恰好是本章所说的「外部签发」在方法论层面的一次小规模实施。

结 语

摆在这里，等一条通道

一、九章走完之后剩下什么

九章各自找了一处坏的东西 把它们拼看，发现件事：这九处并不是九个漏洞，而是同一个漏洞在九个位置上的显形。

一篇的身份,我一直在用读出来的逻辑——文特征过程记录场合受众,这本书是:能读出来那些,恰是这件事在账本留个影子,而不是这件事本身。撤较本,看字像身上还什么(第二章);对文物做映换,看还是它(第二章);道可令整体数量有多少只是还原(第三章);道可长否已经在他身上而文章未成形(第四章);道可处真改了方向裁由作成(第五章)——五种同做法,同一写作:先让能读出来的那部分失效,再看还剩什么。

而第六到第八章处理的是更硬的一格:有些东西连扰动都读不出来,因为账上根本没有为它设一栏。从未获准成形的形式不留任何痕迹(第六章);关闭一条路的那个动作不进成品、草稿、击键或责任账本(第七章);两处有出处之间那一段自造的文字,两端都不记这笔账(第八章)。

到这里,一个改良主义的期待会自然产生:既然现在知道了,增设字段、改造量表、扩充档案即可。

第九章否定了这个期待,而且理由不是技术性的。

二、为什么诊断不会自动带来修复

各不是境况的物,是静制度溢物一类杂物名称,当且只存在一条它为静寂案例通道
手稿在八五年就已规模贮藏;写作杂物转名要等到九七年,而真正生发要等到九八年中国空袭的边。中间隔着的是十年的技术步,是九七年第一批靠这件事藏物的位

这诊断本身的结果是直接的:前八章诊断出的每一处空缺,都不会因为被诊断而被填上。要填上它们,需要的不是更好的理论,是一条从未建成的通道——一批人,他们自己的账要靠查这些空格才能结清。

我把这条放在最后,是因为它这种无法以一个结尾。我没有再给

三、一件本书没有能力做到的事

第九章还给出了一条更不舒服的推论。裁决的四个参数里,终局性为零时,无论程序多么公正,缺口存量都不会下降。现代学术评审在匿名与公开上做了几十年的功夫,而终局性一直是零——这解释了为什么它看起来像结算,却从不减少任何东西。

这条推论适用于本书自己。这本书的分界、自评与编次全部由我方作出,未经外部裁决;判者集合包含于被判者集合。按第九章第八命题,这类文本的终局性为零,因而不能对自己签发结论□

所以这种邪作阿已经立的宣告,它给出的是空格,九个数,九死条件,以及现写死了条件的切——初两的文真实体中从出现,樟障之即行余,应当并。这条我自己写来,是为了它能据把它拆掉

四、最便宜的下一步

如果这种有用,最有用却是可意它,是提几条它自己没写的:

第一章给率——批真的文任务与随材料,两名编者,一个午。不需要生,不需要过程,不需要谁批。它替是一个我以最被替作题:当我一个学生没有自思考时,到手上时便几个结果超的

第二章给率——同所学校到个平班,约百十人,关键是出时点随机。它要看何是一条升曲线,而是一个群:同着个数,自察升高限群

第1章的F1——专利司法执业考试领域的准备程度统计。一旦有数，全书编辑那解释就会从史实推变成批判命题。

这里任何一条出来，都会动这本书。我希望它们改动它。

五、最后一句

写研究几年来每一次进——从文本到过程，从过程到情境——都是在扩大在场的东西的静。这本书问的是另一个方向。

它没有写完。它是因那格出来，并且说明为什么出来还不够。

摆在这里，等一条通道。

全书文献总表

本表把九章各自的文献表合并、去重后统一排列，共 **166** 条（合并180 条，跨章重复14 条次）。各章原有的文献表一并保留，未作删改——正文中的编号引注按各章自身的编号系统运行，若改为全书统一编号，各章的行文引注将全部失效。因此本表的用途是查检与互见，不替代各章文献表。

每章之后另开一栏用表格章次，凡不同章以不同格式录同一文献，取最完整的一条为准，其余合并；著录格式仍统一，保留章面貌。

第章以学科，其随文或段录，故表中自第条为目整注而单条，长便显于他章，请读者意一处值得记录的读数。

合并的结果构成一个数：166 条文中，被章以共同用的只有**6** 条，其余**160** 条只出现在章中。

六条用表中，只有条 Flower 与 Hayes 1981 的“过程模型”被章共同用；其余五条被至三章用，且集中在“研究篇章学”这一片。也就是说，九章除了共同反对的那几位之外，几乎不共享任何文献。

这数并非反对的一面，它是本卷原式的结果——九章自与语教育距离足够远，因各自带了一整套本域都有的文献，这是新现象产生条件。另一方面，它也明了本卷最弱处：九章之间的可比性，目前几乎完全依赖母题与两处对切，而不依赖任何共同的文献地基。若母题推翻，这章会该散成篇互不相识的文。

编者把这个数留在这里，不作辩护。它既是本卷方法的证据，也是本卷风险的证据。

总表

- [1] Abbott, A. The System of Professions: An Essay on the Division of Expert Labor. University of Chicago Press, 1988. [第九章]
- [2] Akbaritabar, A. et al. Reference changes in preprint-publication pairs. Scientometrics, 2022. [第九章]
- [3] Anderson, P. W. (1972). More Is Different. Science, 177(4047), 393-396. [第二章]
- [4] AP Seminar exam information.
<https://apcentral.collegeboard.org/courses/ap-seminar/exam>
[第四章]
- [5] Arno, S. F., & Brown, J. K. (1991). Overcoming the paradox in managing wildland fire. Western Wildlands, 17(1), 40-46. [第八章]
- [6] Arnold, K. C., & Kim, J. (2025). Interaction-Required Suggestions for Control, Ownership, and Awareness in Human-AI Co-Writing. Proceedings of In2Writing 2025, 62-68. DOI: 10.18653/v1/2025.in2writing-1.6. [第一章]
- [7] Bakhtin, M. M. (1986). Speech Genres and Other Late Essays. University of Texas Press. [第二章]

- [8] Bazerman, C. (2004). *Speech Acts, Genres, and Activity Systems: How Texts Organize Activity and People*. In C. Bazerman & P. Prior (Eds.), *What Writing Does and How It Does It*. Lawrence Erlbaum Associates. [第二章]
- [9] Beaufort, A. (2007). *College Writing and Beyond: A New Framework for University Writing Instruction*. Logan, UT: Utah State University Press. [第三章]
- [10] Bellare, M., & Goldreich, O. (1992). On Defining Proofs of Knowledge. *Advances in Cryptology—CRYPTO '92*, LNCS 740. [第一章]
- [11] Bellemin-Noël, J. *Le texte et l'avant-texte*. Larousse, 1972. [第九章]
- [12] Ben-David, J. *The Scientist's Role in Society*. Prentice-Hall, 1971. [第九章]
- [13] Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates. [第二章、第三章、第七章]
- [14] Bereiter, C., & Scardamalia, M. *The Psychology of Written Composition*. <https://www.routledge.com/The-Psychology-of-Written-Composition-1st-Edition/Bereiter-Scardamalia/p/book/9780805800388> [第四章]
- [15] Berthier, Y., Vincent, L., & Godet, M. (1992). Velocity accommodation sites and modes in tribology. *European Journal of Mechanics A/Solids*, 11(1), 35-47. [第三章]
- [16] Bitzer, L. F. (1968). The Rhetorical Situation. *Philosophy & Rhetoric*, 1(1), 1-14. [第二章]
- [17] Blackburn, L., Pesce, D. W., Johnson, M. D., Wielgus, M., Chael, A. A., Christian, P., & Doeleman, S. S. (2020). Closure statistics in interferometric data. *The Astrophysical Journal*, 894, 31. doi:10.3847/1538-4357/ab8469 [第三章]
- [18] Bourdieu, P. Le marché des biens symboliques. *L'Année sociologique*, 22, 1971. [第九章]
- [19] Bourdieu, P. Les Règles de l'art: Genèse et structure du champ littéraire. Seuil, 1992. [第九章]
- [20] Bourdieu, P. *The Field of Cultural Production*. Polity Press, 1993. [第九章]
- [21] Britton J, et al. *The Development of Writing Abilities (11-18)*. Macmillan/NCTE. 1975. [第五章]

- [22] Brock CP. Change to the definition of “crystal” in the IUCr Online Dictionary of Crystallography. IUCr Newsletter, 2021. [第五章]
- [23] Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press. [第八章]
- [24] Chael, A. A., Johnson, M. D., Bouman, K. L., Blackburn, L. L., Akiyama, K., & Narayan, R. (2018). Interferometric imaging directly with closure phases and closure amplitudes. *The Astrophysical Journal*, 857, 23. doi:10.3847/1538-4357/aab6a8 [第三章]
- [25] Chaikin, P. M., & Lubensky, T. C. (1995). *Principles of Condensed Matter Physics*. Cambridge University Press. [第二章]
- [26] Chen, X. (2025). Generalized Landau Paradigm for Quantum Phases and Phase Transitions. *Physical Review Letters*, 135, 250001. [第二章]
- [27] College Board. AP Capstone Diploma Program Policies. [第七章]
- [28] Conference on College Composition and Communication. (2018). *Writing Assessment: A Position Statement*. NCTE/CCCC. [第二章]
- [29] Conference on College Composition and Communication. *Generative AI Policy*. 2025. [第七章]
- [30] Cornell Center for Teaching Innovation. *AI Attribution Guidelines*. [第七章]
- [31] Cornell Legal Information Institute. *Personal Services*. [第七章]
- [32] Crossley, S. A., Tian, Y., et al. (2024). The English Language Learner Insight, Proficiency and Skills Evaluation (ELLIPSE) Corpus. *International Journal of Learner Corpus Research*, 9(2). [第二章]
- [33] Dai, Z. (2026). Rethinking AI-assisted writing instruction: feedback literacy scripts, calibration training, and student writing development. *Frontiers in Psychology*, 17, 1829268. DOI: 10.3389/fpsyg.2026.1829268. [第一章]
- [34] de Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to Text Linguistics*. London: Longman. [第三章]
- [35] de Biasi, P.-M. *La génétique des textes*. Nathan, 2000. [第九章]

- [36] Deppman, J., Ferrer, D. & Groden, M. (eds.) *Genetic Criticism: Texts and Avant-textes*. University of Pennsylvania Press, 2004. [第九章]
- [37] Derrida, J. *Mal d'archive: une impression freudienne*. Galilée, 1995. [第九章]
- [38] Descartes, S., & Berthier, Y. (2002). Rheology and flows of solid third bodies: Background and application to an MoS1.6 coating. *Wear*, 252, 546-556. doi:10.1016/S0043-1648(02)00008-X [第三章]
- [39] Dobrin, S. I., Rice, J. A., & Vastola, M. (Eds.). (2011). *Beyond Postprocess*. Utah State University Press. [第二章]
- [40] Eilenberg, S., & Mac Lane, S. (1945). General Theory of Natural Equivalences. *Transactions of the American Mathematical Society*, 58, 231-294. [第二章]
- [41] Elbow, P. (1973). *Writing Without Teachers*. New York: Oxford University Press. [第八章]
- [42] Emig, J. (1971). *The Composing Processes of Twelfth Graders* (NCTE Research Report No. 13). Urbana, IL: National Council of Teachers of English. [第五章、第八章]
- [43] Emig, J. "Writing as a Mode of Learning." *College Composition and Communication*, 1977. [第七章]
- [44] Espeland, W. & Sauder, M. Rankings and Reactivity. *American Journal of Sociology*, 113(1), 2007. [第九章]
- [45] Example-induced design fixation research.
<https://www.sciencedirect.com/science/article/pii/S1871187115300353> [第四章]
- [46] Fillot, N., Iordanoff, I., & Berthier, Y. (2007a). Modelling third body flows with a discrete element method—A tool for understanding wear with adhesive particles. *Tribology International*, 40(6), 973-981. doi:10.1016/j.triboint.2006.02.056 [第三章]
- [47] Fillot, N., Iordanoff, I., & Berthier, Y. (2007b). Wear modeling and the third body concept. *Wear*, 262(7-8), 949-957. doi:10.1016/j.wear.2006.10.011 [第三章]
- [48] Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365-387. DOI: 10.58680/cc198115885. [第一章、第二章、第三章、第五章、第八章、第九章]

- [49] Formosa, P., Bankins, S., Matulionyte, R., et al. (2025). Can ChatGPT be an author? Generative AI creative writing assistance and perceptions of authorship, creatorship, responsibility, and disclosure. *AI & Society*, 40, 3405–3417. DOI: 10.1007/s00146-024-02081-0. [第一章]
- [50] Gero KI, Long T, Schnitzler C, Dhillon P. From Planning to Revision: How AI Writing Support at Different Stages Alters Ownership. arXiv, 2026. [第五章]
- [51] Gilbert, S. F. Developmental Biology. NCBI Bookshelf: Cell commitment, specification and determination. <https://www.ncbi.nlm.nih.gov/books/NBK26825/> [第四章]
- [52] Godet, M. (1984). The third-body approach: A mechanical view of wear. *Wear*, 100(1-3), 437–452. doi:10.1016/0043-1648(84)90025-5 [第三章]
- [53] Goldwasser, S., Micali, S., & Rackoff, C. (1989). The Knowledge Complexity of Interactive Proof Systems. *SIAM Journal on Computing*, 18(1), 186–208. DOI: 10.1137/0218012. [第一章]
- [54] Graeber, D. Debt: The First 5000 Years. Melville House, 2011. [第九章]
- [55] Grésillon, A. *Éléments de critique génétique*. PUF, 1994. [第九章]
- [56] Haidemariam, T. (2026). After generative AI: authorship, labour, and cultural governance. *AI & Society*. DOI: 10.1007/s00146-026-03189-1. [第一章]
- [57] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman. [第二章、第三章、第八章]
- [58] Hamamra, B., Khlaif, Z. N., & Mahamid, N. (2026). AI literacy and mediated learner autonomy in ChatGPT-supported EFL writing: a qualitative study at a Palestinian university. *International Journal of Educational Technology in Higher Education*, 23, 19. [第一章]
- [59] Hay, L. La critique génétique: origines et perspectives. In *Essais de critique génétique*, Flammarion, 1979. [第九章]
- [60] Hayes, J. R. (1996). A New Framework for Understanding Cognition and Affect in Writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing*. Lawrence Erlbaum Associates. [第二章]

- [61] Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
doi:10.1177/0741088312451260 [第三章]
- [62] Hedberg, H. D. (Ed.). (1976). *International Stratigraphic Guide*. Wiley. [第二章]
- [63] Herbold S, Hautli-Janisz A, Heuer U, et al. A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*. 2023;13:18617. [第五章]
- [64] Honneth, A. *Kampf um Anerkennung*. Suhrkamp, 1992. [第九章]
- [65] Huang, Y., & Wilson, J. M. (2021). Using automated feedback to develop writing proficiency. *Computers and Composition*, 62, 102675. DOI: 10.1016/j.compcom.2021.102675. [第一章]
- [66] IASB. Conceptual Framework discussions on definition, recognition and derecognition.
<https://media.ifrs.org/2013/IASB/April/IASB-Update-April-2013.html> [第四章]
- [67] IB Language A assessment instructions, 2026.
https://ibpublishing.ibo.org/exinst/apps/exinst/index.html?chapter=1&doc=EX_instructions_2026_e&part=8 [第四章]
- [68] ICOM Committee for Conservation. *The Conservator-Restorer: A Definition of the Profession*. 1984. [第七章]
- [69] International Baccalaureate. *Guidance on artificial intelligence in assessment*. [第七章]
- [70] International Commission on Stratigraphy. *International Stratigraphic Guide: Stratotypes and Type Localities; Lithostratigraphic Units; Chronostratigraphic Units*. Online guide, accessed 2026-08-17. [第二章]
- [71] International Union of Crystallography (IUCr). *Dictionary of Crystallography: “Polymorphism”*.
<https://dictionary.iucr.org/Polymorphism> [第四章]
- [72] International Union of Crystallography. *Online Dictionary of Crystallography: Structure determination*. [第五章]
- [73] Iordanoff, I., Berthier, Y., Descartes, S., & Heshmat, H. (2002). A review of recent approaches for modeling solid third bodies. *Journal of Tribology*, 124(4), 725–735. doi:10.1115/1.1467632 [第三章]

- [74] IUCr. Teaching pamphlet on crystal growth, nucleation and metastability. <https://www.iucr.org/education/pamphlets/21/full-text> [第四章]
- [75] IUPAC. (2025 online ed.). “mechanism (of a reaction).” *Compendium of Chemical Terminology (Gold Book)*, 5th ed. DOI: 10.1351/goldbook.M03804. [第一章]
- [76] IUPAC. “Catalyst”; “Selectivity”; “Promoter.” *Compendium of Chemical Terminology*. [第七章]
- [77] Ivanič R. *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*. John Benjamins. 1998. [第五章]
- [78] Jelson A, et al. NIRVANA: A Comprehensive Dataset for Reproducing How Students Use Generative AI for Essay Writing. arXiv, 2026. [第五章]
- [79] Jennison, R. C. (1958). A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent. *Monthly Notices of the Royal Astronomical Society*, 118(3), 276-284. doi:10.1093/mnras/118.3.276 [第三章]
- [80] Kadanoff, L. P. (2009). More is the Same; Phase Transitions and Mean Field Theories. *Journal of Statistical Physics*, 137, 777-797. [第二章]
- [81] Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50(1), 1-73. DOI: 10.1111/jedm.12000. [第一章]
- [82] Kent, T. (Ed.). (1999). *Post-Process Theory: Beyond the Writing-Process Paradigm*. Southern Illinois University Press. [第二章]
- [83] KEYScore: keystroke-based writing process research, BEA 2026. <https://aclanthology.org/2026.bea-1.24/> [第四章]
- [84] Klein, M. et al. Comparing published scientific journal articles to their pre-print versions. *International Journal on Digital Libraries*, 2018. [第九章]
- [85] Koo, R., Martin, A., Wang, L., & Kang, D. (2023). Decoding the End-to-end Writing Trajectory in Scholarly Manuscripts. arXiv:2304.00121. [第一章]
- [86] Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411-433. doi:10.1111/j.1468-2958.2004.tb00738.x [第三章]

- [87] Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology* (4th ed.). Thousand Oaks, CA: Sage. [第三章]
- [88] Lehn, J.-M. (1988). Supramolecular chemistry—Scope and perspectives: Molecules, supermolecules, and molecular devices. *Angewandte Chemie International Edition in English*, 27(1), 89–112. doi:10.1002/anie.198800891 [第三章]
- [89] Lehn, J.-M. (2002). Toward complex matter: Supramolecular chemistry and self-organization. *Proceedings of the National Academy of Sciences*, 99(8), 4763–4768. doi:10.1073/pnas.072065599 [第三章]
- [90] Leibman, C. C. *Is This What You Wanted?: Expectations, Choice, and Rhetorical Agency in Composition*. 2019. [第七章]
- [91] Leinster, T. (2014). *Basic Category Theory*. Cambridge University Press. [第二章]
- [92] Lin, J. et al. How many preprints have actually been printed and why: a case study of computer science preprints on arXiv. *Scientometrics*, 124, 2020. [第九章]
- [93] Lo Sardo, D. R., Gravino, P., Cuskley, C., & Loreto, V. (2023). Exploitation and exploration in text evolution: Quantifying planning and translation flows during writing. *PLoS ONE*, 18(3), e0283628. DOI: 10.1371/journal.pone.0283628. [第一章]
- [94] Luhmann, N. *Legitimation durch Verfahren*. Luchterhand, 1969. [第九章]
- [95] Mac Lane, S. (1998). *Categories for the Working Mathematician* (2nd ed.). Springer. [第二章]
- [96] Mammen, M., Choi, S.-K., & Whitesides, G. M. (1998). Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie International Edition*, 37(20), 2754–2794. doi:10.1002/(SICI)1521-3773(19981102)37:20%3C2754::AID-ANIE2754%3E3.0.CO;2-3 [第三章]
- [97] Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 243–281. doi:10.1515/text.1.1988.8.3.243 [第二章、第三章、第八章]
- [98] Mathias, S., & Bhattacharyya, P. (2018). ASAP++: Enriching the ASAP Automated Essay Grading Dataset with Essay Attribute Scores. *Proceedings of LREC 2018*. [第二章]

- [99] Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, 15(2), 113-134. doi:10.58680/rte198115773 [第八章]
- [100] McGann, J. *A Critique of Modern Textual Criticism*. University of Chicago Press, 1983. [第九章]
- [101] Medimorec, S., & Risko, E. F. (2017). Pauses in written composition: On the importance of where writers pause. *Reading and Writing*, 30(6), 1267-1285. doi:10.1007/s11145-017-9723-7 [第八章]
- [102] Merton, R. K. *Priorities in Scientific Discovery*. *American Sociological Review*, 22(6), 1957. [第九章]
- [103] Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151-167. doi:10.1080/00335638409383686 [第二章、第三章]
- [104] Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134-140. [第八章]
- [105] National Council of Teachers of English. *The Students' Right to Write*. 2026. [第七章]
- [106] National Council of Teachers of English. *Understanding and Teaching Writing: Guiding Principles*. 2016/2018. [第五章]
- [107] NCBI Bookshelf. Early developmental commitment and differentiation. <https://www.ncbi.nlm.nih.gov/books/NBK9968/> [第四章]
- [108] Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600-2606. doi:10.1073/pnas.1708274114 [第三章]
- [109] Nowacek, R. S., Leonard, R. L., & Rounsaville, A. (2024). *Writing Knowledge Transfer: Theory, Research, Pedagogy*. Parlor Press & The WAC Clearinghouse. [第一章]
- [110] Nystrand, M. (1986). *The Structure of Written Communication: Studies in Reciprocity Between Writers and Readers*. Orlando, FL: Academic Press. [第三章]
- [111] Parks, S. A., Holsinger, L. M., Miller, C., & Nelson, C. R. (2015). Wildland fire as a self-regulating mechanism: The role of previous burns and weather in limiting fire progression. *Ecological Applications*, 25(6), 1478-1492. doi:10.1890/14-1430.1 [第八章]
- [112] Power, M. *The Audit Explosion*. Demos, 1994. [第九章]

- [113] Power, M. *The Audit Society: Rituals of Verification*. Oxford University Press, 1997. [第九章]
- [114] Pyne, S. J. (1982). *Fire in America: A Cultural History of Wildland and Rural Fire*. Princeton, NJ: Princeton University Press. [第八章]
- [115] Reprogramming / transdifferentiation review.
<https://pubmed.ncbi.nlm.nih.gov/36067834/> [第四章]
- [116] Rodrigues, T. V. (2026). The Apt Curation Model: An Epistemic Virtue Theory of AI-Assisted Authorship. *Philosophy & Technology*, 39, Article 39. DOI: 10.1007/s13347-026-01038-z. [第一章]
- [117] Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication*, 16(2), 106-112. [第八章]
- [118] Rohman, D. G. Pre-writing and discovery in composition.
<https://publicationsncte.org/content/journals/10.58680/ccc196521081> [第四章]
- [119] Rohman, D. G., & Wlecke, A. O. *Pre-Writing: The Construction and Application of Models for Concept Formation in Writing*. 1964. [第七章]
- [120] Salvador, A. (Ed.). (1994). *International Stratigraphic Guide: A Guide to Stratigraphic Classification, Terminology, and Procedure* (2nd ed.). Geological Society of America. [第二章]
- [121] Sato T, et al. Single Lgr5 Stem Cells Build Crypt-Villus Structures in Vitro without a Mesenchymal Niche. *Nature*. 2009;459:262-265. [第五章]
- [122] ScholaWrite / writing-process phase research.
<https://www.jowr.org/jowr/article/view/1714> [第四章]
- [123] SDE Universes. *发育与医学新思想前沿*. 2026. [第五章]
- [124] SDE Universes. *新思想前沿：626个领域近十年两思想构建*. 2026. [第五章]
- [125] SDE Universes. *考古与古人类学·新思想前沿*. 2026. [第五章]
- [126] Seeman, J. I. (1983). Effect of conformational change on reactivity in organic chemistry: Evaluations, applications, and extensions of Curtin-Hammett/Winstein-Holness kinetics. *Chemical Reviews*, 83(2), 83-134. doi:10.1021/cr00054a001 [第八章]

- [127] Sommers, N. Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication*, 31(4), 1980. [第九章]
- [128] Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3), 619–659. doi:10.1162/COLI_a_00295 [第三章]
- [129] Stanford Program in Writing and Rhetoric. Policy on the Use of GenAI and LLMs. [第七章]
- [130] Steedman, C. *Dust: The Archive and Cultural History*. Rutgers University Press, 2002. [第九章]
- [131] Strauss, A. (1985). Work and the division of labor. *The Sociological Quarterly*, 26(1), 1–19. [第八章]
- [132] Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 2006;126:663-676. [第五章]
- [133] Tanselle, G. T. *A Rationale of Textual Criticism*. University of Pennsylvania Press, 1989. [第九章]
- [134] Taylor, M., Bishop, B., Burkes, T., et al. (2021). Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach. NISTIR 8282r1. DOI: 10.6028/NIST.IR.8282r1. [第一章]
- [135] The Stacks Project Authors. (2026). Opposite Categories and the Yoneda Lemma, Tag 001P. Accessed 2026-08-17. [第二章]
- [136] Tian, Y., & Cushing, S. T. “Exploring the Application of Keystroke Logging Techniques to Research in L2 Writing.” 2024. [第七章]
- [137] U.S. National Park Service. What Is Archeological Context? 2023. [第五章]
- [138] UNESCO. Nara Document on Authenticity. 1994. [第七章]
- [139] Uniform Commercial Code §2-210, Delegation of Performance and Assignment of Rights. [第七章]
- [140] University of Sydney. Using Artificial Intelligence Tools Responsibly. 2026. [第七章]
- [141] Valéry, P. *Œuvres I (“Au sujet du Cimetière marin”)*. Gallimard, 1957. [第九章]
- [142] van Dijk, T. A. (1980). *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition*. Lawrence Erlbaum Associates. [第二章]

- [143] Wallas, G. (1926). *The Art of Thought*. London: Jonathan Cape. [第八章]
- [144] Wan, R., Gebreegziabhe, S., Li, T. J.-J., & Badillo-Urquiola, K. (2024). *CoCo Matrix: Taxonomy of Cognitive Contributions in Co-writing with Intelligent Agents*. arXiv:2405.12438. [第一章]
- [145] Wen, X.-G. (1990). Topological Orders in Rigid States. *International Journal of Modern Physics B*, 4, 239-271. [第二章]
- [146] Whitley, R. *The Intellectual and Social Organization of the Sciences*. Clarendon Press, 1984. [第九章]
- [147] Widdowson D, Kurlin V, et al. The importance of definitions in crystallography. *IUCrJ/Acta-related publication*, 2024. [第五章]
- [148] Wilson, K. G. (1971). Renormalization Group and Critical Phenomena. I & II. *Physical Review B*, 4, 3174-3205. [第二章]
- [149] Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32(2), 189-204. doi:10.58680/cc3198115912 [第三章]
- [150] Writing transfer research: adaptive transfer and writing knowledge across contexts. <https://www.sciencedirect.com/science/article/pii/S1060374311000178> [第四章]
- [151] Yeadon W, et al. Evaluating AI and Human Authorship Quality in Academic Writing through Physics Essays. arXiv, 2024. [第五章]
- [152] Yoneda, N. (1954). On the Homology Theory of Modules. *Journal of the Faculty of Science, University of Tokyo, Section I*, 7, 193-227. [第二章]
- [153] —— 学通栏·二阶碰撞典范文 (What1·v3.9) —— [第二章]
- [154] 《周易·系辞上》，见《十三经注疏》本；“大衍之数五十”章。 [第八章]
- [155] 与何留生态路 Cooper, M., “The Ecology of Writing”, *College English*, 48 卷4 期 364-375 页 1986 , 其后提出的 (ecocomposition) 文本创作界, 因那路处理是写作生态系统, 全落在成例之上, 与本文第四相。此特意出, 是因为被线覆盖”条件何留作这问题 [第四章]
- [156] 作自 (2026) . 《作文何以成为作文：从“文本成品”到“跨端复成体”——古生物与真形态·信处理·空间统计的二阶碰撞》. 学通栏·二阶碰撞典范文 (同产线近邻·用于划界) . [第四章]

- [157] 免疫学三派：Burnet 的克隆选择学说（1957 及其后续）；Matzinger, P., “Tolerance, danger, and the extended family”, *Annual Review of Immunology*, 12 卷 991-1045 页 1994；Pradeu, T., *The Limits of the Self: Immunology and Biological Identity*, Oxford University Press, 2011。 第 9 章
- [158] 关晓红《科举停废与近代中国社会》，社会科学文献出版社，2013。 [第 9 章]
- [159] 多晶型与阶段律：Ostwald, W., *Zeitschrift für physikalische Chemie*, 22 卷 1897（阶段律）；Dunitz, J. D. & Bernstein, J., “Disappearing polymorphs”, *Accounts of Chemical Research*, 1995；Bučar, D.-K., Lancaster, R. W. & Bernstein, J., “Disappearing Polymorphs Revisited”, *Angewandte Chemie International Edition*, 54 卷 2015；Bauer, J. et al., “Ritonavir: An extraordinary example of conformational polymorphism”, *Pharmaceutical Research*, 18 卷 859-866 页 2001；复见 PNAS, 121 卷 e2319127121, 2024 (DOI: 10.1073/pnas.2319127121)。参考书：Bernstein, J., *Polymorphism in Molecular Crystals*, 2nd ed., Oxford University Press, 2020。 第 9 章
- [160] 张政烺：《试释周初青铜器铭文中的易卦》，《考古学报》1980 年第 4 期 [第 8 章]
- [161] 数字卦文：两真卦与假卦的部分构成自身的卦（见 F8）。 第 9 章
- [162] 朱熹：《周易本义·筮仪》，台北：大安出版社，2008 年，页 7-10。 [第 8 章]
- [163] 李学勤主编：《清华大学藏战国竹简（肆）·筮法》，上海：中西书局，2013 年。 第 9 章
- [164] 第二节地貌学路对双方：Davis, W. M., “The geographical cycle”, *The Geographical Journal*, 1899；Hack, J. T., “Interpretation of erosional topography in humid temperate regions”, *American Journal of Science*, 258-A 卷 1960。同路后：Strahler 1950；Chorley 1962。 第 9 章
- [165] 第十六节各位：Burke, K., *Permanence and Change*, 1935；Whorf, B. L., 1940；Bartlett, F. C., *Remembering*, 1932；Freud, S., 1915；Elbow, P., *Writing Without Teachers*, Oxford University Press, 1973；Rose, M., *College Composition and Communication*, 31 卷 4 期 1980；Perl, S., *Research in the Teaching of English*, 13 卷 1979；Sommers, N., *College Composition and Communication*, 31 卷 4 期 1980；Rubin, D. B., *Journal of Educational Psychology*, 66 卷 1974；Heckman, J. J., *Econometrica*, 47 卷 1979；Bourdieu, P., *Esquisse d’une théorie de la pratique*, Droz, 1972。 第 9 章

[166] 陆机《文赋》；刘勰《文心雕龙·铨裁》；桓谭《新论》（“覆瓿”）；郑燮《题画》（“胸中之竹”）。 [第九章]

附录一 母题是怎么打出来的（全过程留档）

本章编写由一条决定为这条母题被审核可被播，此留下它的过程，包括可产物
工序一·九张判断卡片

章	判断卡片（一句，反直觉）
一	文本被撤走以后还能重新长出来的那套关系，才是作文
二	不动它就不知道它是不是它——身份只在变换中显形
三	近邻就能还原的判断是假的；读得出的那一部分恰恰不是整体性
四	文章还没出现，作文已经发生；文章已经完成，作文可能没发生
五	成品越好越不能证明是他写的——质量与身份反向
六	从未获准成形的那些形式不留痕迹，却划定了边界
七	路在你遇见它之前就被关掉了，而关闭这个动作本身不进任何账
八	两处有据点之间那一段文字没有主人，成品上看不出来
九	残余物没有名分，不是没人看见，是没有一条清偿通道

工序二·共绕动词族

动词族	同族成员	命中章	次数
不入账族	不进记录／没有字段／不记账／无名分／不留痕迹	4,5,6,7,8,9	6
扰动才显形族	撤走／变换／移位／撤支架／破相	1,2,3,4	4
错位族	先于／晚于／反向／不同步	3,4,5,7	4
假冒族	局部能还原／成品像／达标却没发生	2,3,5	3

交点：九章同时落在不入账与扰动显形两族上，判断共绕口

工序三·主题与母题的分离

层	内容
主题层	都答「作文是什么」／都从远学科取源／都反对「成品即作文」
母题层	作文的身份既不在成品上，也不在过程记录里；它落在账本没有设字段的那一格，

	而且只有动它一下才显形
--	-------------

~~主题是任何文章都会有的描述,不可证伪;母题是一条可以被推翻的断言,二者不可互换~~

工序四·母题（三十九字）

作文不在成品上：凡直接读得出的都不是它；它落在账本没有设的那一格，动一下才显形。

含两个共绕动词族；是判断不是描述；不含任何内部方法术语；可证伪——若成品特征或过程日志能稳定预测身份与迁移，本条被推翻。

工序五·回归校验（9/9）

章	派生路径
一	「动一下才显形」的逐字命名（撤稿后重提取）
二	「动一下才显形」在文本侧的展开（保相／破相变换）
三	「直接读得出的都不是它」的实测反证（近邻可还原 0.875）
四	成品与身份错位，中性扰动才读得出
五	「直接读得出的都不是它」（终稿质量七项全部反向）
六	「账本没有设的那一格」的本体命名
七	那一格的关闭动作，同样不入账
八	已写出但两端都不记账
九	存在而无名分，账上没有那一栏

~~九章皆由母题派生,无游离篇~~

附录二 九个读数、九组死亡条件（一览）

章	新存在物	读数	已跑到哪一档	最便宜的一条死亡条件
一	生成见证／作文等价类	生成见证保真率	无自跑；两项他人研究只支持较弱前提	撤稿挑战中不同盲评者对生成结构无法达到基本一致，且只有提高抽象程度才能提高一致性
二	篇章定相体	篇章定相率 CPR	字段审计不利：现有公开语料算不出	一百篇作文、每篇六个变换版本、五个关系探针，若响应与静态语义相似度无增量
三	异源回判	远距增益 Δ_remote	预注册真跑，主假设被淘汰并据此重建概念	远距增益在控制近邻可还原之后不显著
四	作定态／隐作	自续率 SR	无检验；最小实验一百二十人可完成	从未出现「文本尚未分化而关系自续已稳定」的个案
五	写成事件	最小写成记录（非读数）	复核公开数据，七项全部反向	控制终稿质量与能力后，高杠杆裁决闭合度对下一次陌生题目无独立预测力
六	未生界	末次首现 L	主假设未过阈值（15.0% 对 30%）；F6 再跑得「一位也读不出」	早年被当面否定过的写法，在后续自由写作中重现率显著高于随机基线
七	预闭分叉	留叉率 OFR	政策审计不利，据此删去一条批判；OFR 尚未算过	完全私人、无评价预期的写作中仍观察到同等程度的可指认关闭
八	悬跨段	悬跨长度 λ	三条预注册假设全被驳；F5 判「未决」	同题多班留有初稿的对照（F4，最便宜且未跑）

九	清偿率 $\lambda = c \cdot f \cdot \alpha$	清偿率	三条预注册假设全部失败并当场改判； α 测量法补出后算出四个制度的 λ	F1：终局性为正的领域若有同密度的准清偿仪式（最便宜且未跑）
---	--	-----	--	--------------------------------

两处对切的合并条件（写死）

对切	合并条件
第一章 ↔ 第二章	若人侧读数与文本侧读数长期一一对应，两种交叉在真实样本中从不出现，两章之一冗余
第六章 ↔ 第七章	若每一处未生界都能回溯到可指名的关闭者，第六章并入第七章；若所有预闭都不改变可及集合，第七章并入第六章

附录三 六章真跑，六章不利（清单）

本章中，六章用公开数据公开料。六章经修改已不再保留，且经原保留。此集抄出，供读者一次核时

第二章·公开作文评分语料字段审计。要定类同时存在原完整文，可建版本序。预注册保由破换，变版本上的系，首评。原与变版本的一对应。审计ASAP/ASAP++与 ELLIPSE：变与探序存在。结果：篇章定相率目前无法在任何现成语料上计算。原稿用组分或词以度代理，不结果原保留。

第三章·公开语料先导研究。七篇文，六类最扰动，七二项断，全字覆盖率0.9722，看以强，但空可还原率0.8750，越预册的0.80。淘汰，删余部等时，编号至0.9365。更重要的是，不读文，只看他哪一类扰动，逐留预册，命中0.8611。结果：肯定性定义被自己的预注册数据淘汰，概由内容判断，重建为冠E异源，可判。

第五章·公开评分数据复核。九十题目，二百七篇本，一百一十名师，六五十五次评分。七项评分的统计：7/7全部反向——质量高的侧恰可能开学生本的过程。原稿确能把1/128当作显著性，只防的统计，并明数据有终，没有过程，因此能翻一个更替命题。

第六章·预注册真跑。八篇题，九条题，五余条目标，结果列N=367。成33.0%/仍52.0%/除15.0%，预注册阈值为30%，三种算法一致不过线。据改的，用口清单，未世界。这路，适用国收，有续出，可查的场合。后续F6再跑：若初率1.000，一位作者也读不出末次首现；事，度，且显示，完全由式清，粗决定，因此，以便策——未，式清，的数，律，成立。

第七章·公开政策审计。十二份开度文件，结果利：多套假，已跟过程，检查点，据此删去「现行制度只看成品」这条批判；仍发现，记录，学生遭害，已消，预备选方向」的假程。

第八章·预注册真跑。六篇题，两三千条，修，三千百余个。会话H1/H2/H3全部被驳：系，防，可反，AUC与预册反号，跨空，数据，续，率，高，开，预，作者自曝，处，狂，误——含，会，第，一，步，报，数，容，时，隔，时，终，以，后，号，翻，转，并，明，符，容，时，那，种，像，是，番，那，种。后续F5判，未决，第三层，引，随，改，写。

第九章·两次预注册真跑。研究一：一千七百三十五篇论文，H1/H2/H3全部失败，其中一条方向与预注册所写的不利方向一致；唯一显著的读数与既有文献结论一致，原稿写明「只是独立复现、不主张新发现」。据此当场改判——到期性是主体账目属性而非作品属性。研究二：书籍语料年频检验成立，并意，出，路，层，出，故，层，越，付，年，v1.1，另， α 的测量法，四个制度第一次算出清偿率；并自查出一处对自己不利的结构问题（ α 有三项在某制度上更高而清偿率反而更低），据此新增一条证伪路径，同时撤回一处定理的证明记号、降格为命题。

封 底

凡直接读得出的，都不是它

「作文是什么」，在生成式人工智能出现以前几乎不必问。写下这些字的人、经历这场写作的人、为这份成果负责的人，通常是同一个。三者第一次分开了，而我们手里所有的账本，都是在它们重叠的年代设计的。

本书从九个方向给出同一个回答：一篇文的身既不在成品上，也不在过程记录里——它落在账本没有设字段的那一格，而且只有动它一下才显形。

九章各提出一个对象、一个可清的数，并各自能让已死的条件。其中六章用公开数据或公开材料，六章全部得到对自己不利的结果，并全部原样保留；三章被当场销毁或收买，一章撤回自己处理的证明号。

编一 这一篇凭什么是这一篇

编二 什么时候它成为篇文

编三 账目格：无形无据无名

终篇 制度这侧

德麦国际出版社 Demai International Press

ISBN 979-8-90690-038-8 US \$25.60

