

文 心 探 幽

卷 二

作文如何发生？

看下一步由谁决定

每一次被判为成功的结算·都在无声处收回决定权

付自文 著

德麦国际出版社

文心探幽·卷二：作文如何发生？

付自文 著

德麦国际出版社 · Demai International Press

ISBN 979-8-90690-037-1

定价 US \$23.40

开本 6 × 9 英寸

2026 年 8 月第 1 版

版权所有 · 侵权必究

著者简介

付自文 SDE 学员。

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。

在此后的几年里，他沿着这条思想线一路前行：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」。对他而言，这并不是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」慢慢成为他理解世界的一种基本方式——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？

著有《文心探幽》丛书。丛书现有三卷，各问一个不同的问题：

- 卷一《何谓作文》（德麦国际出版社，专著第 91 号，2026）——问一篇作文的身份落在哪里。四编九章，约 19.4 万字。
- 卷二《作文如何发生？》——即本卷，问从没有到有中间发生了什么。
- 卷三《作文的意义是什么？》（德麦国际出版社，专著第 90 号，2026）——问为什么它在本轮结不清。四编九章，约 19.3 万字。

丛书以「不入账」为共同的一半：卷一读身份，卷二读发生，卷三读结算。三卷各自独立成书，可分开读，也共用一条纪律——每章给出一个可清点的读数，每个读数配一条能杀死它的观察，不利结果原样保留在正文里。

本卷的取源方式，是把三门与语文教育距离很远的学科放在同一个问题上相撞，取它们各自最硬的一条判断，看它们在什么地方共同失效；新对象从共同失效处长出来，而不从任何一门学科的现成结论里搬。九章合起来动用了二十四门学科：成核动力学、免疫学的动力学校对、分布式共识、结晶学的阶段律、药理学的给予节律、组织学习、形态发生、交通流、机制设计、高分子的可复启链端、形式化验证、细胞力学转导、埋藏学、第三体摩擦学、位置信息、仲裁的终局性、风险与可靠性工程、化学反应工程、公司治理、保护生物学、软件安全与隐私工程、发育生物学的转决定、交通工程的容量跌落、科学社会学的多重发现。

前 言

一·这本书是从一件小事开始的

一位初中语文老师在教研组里说过一句话，后来在那个组里流传了很久：

这孩子现在写什么都像作文，就是不像他自己写的。

她说的是一个从初一下学期开始每周一篇周记、两年没断过的学生。素材本记满了，好词好句抄了三大本，范文读过两百篇以上。到初三，他的作文分数没有下降，多数时候还略有上升。

这句话之所以在教研组里流传，是因为所有人都听得懂它，而没有一个人能把它写进任何一张表格。分数栏里没有它的位置，评语栏里也没有——那孩子的每一次作文都合格，很多次还不止合格。老师能感到有一样东西在减少，但她手上的所有仪表都显示正常。

本卷九章，是九种把这句话变成可清点的东西的尝试。

九章的作者互不知情，取的学科天差地别，读数各不相同。但九章最后指向同一个位置：那件在减少的东西，是决定下一步的权力。

二·三种现成的答案，与它们共同的那块地基

「作文如何发生」当然不是没有人答过。有三套现成的答案，而且这三套互不相容。

步骤论说，发生就是走完次序：审题、立意、列提纲、写开头、展开、修改、润色。它的长处是可教、可查、可评分。它的危险是把「顺序」当成了「发生」——同一个顺序可以由完全不同的机制产生。甲的提纲只是一个暂时假设，写到第三段被反例逼着改写中心；乙的提纲是一张不可撤回的施工图，后文只负责填空。日志上两人都是「规划—输入—修订」。

过程论说，写作不是线性誊录，而是规划、转译、回读相互嵌套的递归活动，目标会在写作中改变。这一套已经把步骤论拆掉了，今天说「作文是过程不是结果」已经不是新判断，是研究常识。但它解释不了另一件事：「有过程」为什么不等于「过程在长出下一步」。一个人可以按预先写好的提纲连续誊写，段落越来越完整，过程日志非常丰富，却几乎没有任何局部产物反过来改写后面的行动。

能力论说，写不出来是能力、输入或条件三者之一的不足，于是分别在成品端、过程端、条件端下药。它解释不了本卷开头那三个场面——那三处的能力指标全部正常，甚至上升。

三套底下共享着同一块地基，而且因为它们忙于互相批评，谁也没有回头看它：

凡发生过的推进，都会在产出或记录上留下量。

写了多少字、改了多少次、用了多长时间、拿了多少分、交没交、写得怎样。发生被默认为可清点的，而清点的口径是量。三套于是分工明确：步骤论去记次序，过程论去记事件，能力论去记水平。它们分工的前提，是那笔账总在三处之一。

本卷九章共同否定的，正是这一层。它们说：有一类推进与一类损失，三处都不记，而且不是疏漏。结果处不记，因为它合格；过程处不记，因为它按时；水平处不记，因为当事人自己也说不上哪里不对——这一次他确实做到了。这一类东西之所以不进账，恰恰是因为它发生在一次每一项都达标的通过里，而所有的记账规则都是围绕「哪一项没达标」设计的。

三·一句话，与它的两半

全书的脊梁是一句话：

作文如何发生，看下一步由谁决定；每一次被判为成功的结算，都在无声处收回决定权。

它有两半，两半都要分开说，因为分开之后各自都可以被推翻。

肯定的一半：发生就是决定权的交接。

不看写了多少，看下一步由谁决定——由外部的题目、模板、教师提示、机器补全决定，还是由已经写出的这一小块自己决定。第一章把这一刻命名为续写权移交，并要求它经得起反事实检验：把候选核删掉，后面若干个独立选择是不是跟着一起变。第三章把同一件事推到更窄的刀口上——内容一个字不改、可见性一点不减，只把那段文字从可编辑的正文位置移到只读的侧栏，随后一段时间里操作的分布会不会变。第四章与第五章分别读它向前走（下一段在还谁的账）与向后走（退出的账被谁接了过去）。

否定的一半：收回发生在被判为成功的地方，而且不报错。

这一半才是反直觉的那一半。常识认为写得越顺、评价越好，说明能力在长。本卷说，被判为成功的那一刻，恰恰是决定权被收回的时刻。

第二章的收回发生在受表扬处：最先成形的那段合格文字被批改、被打分、被读给全班、被抄进素材本、被当作下一届的范文——于是它成为环境中新的现成形态，下一轮更偏。第七章的收回发生在宣布完成处：宣布完成的那把尺子，只看见它被设计去看见的东西。第八章的收回发生在判准被学会处：一套判准一旦被公开、被教会、被用于优化，它就不能再证明它原本要证明的东西。第九章的收回发生在全部交付处：交上去了，分数正常，记录完整，而此后同类写作的通行能力被下调。

两半合起来才是全书。只有肯定的一半，本卷退回成过程研究的一个变种；只有否定的一半，本卷退回成一句关于评价制度的牢骚。

四·为什么是这九章，为什么按这个次序

九章排成四编，走的是一条权力的路线。

编一，权第一次落到文本身上。三章各用一种办法证明纸面骗人：删掉学生自己说「我有了」的那一点，看后面的选择变不变（第一章）；逐句追问这一句在这一次开始之前是否已经以成句的形态存在（第二章）；内容不动、只改这段文字在工作区里的位置（第三章）。第三章排在编末，因为它的刀口

最窄，窄到很容易被误读成一项编辑器界面研究——必须先由前两章把「纸面不可信」立住。

编二，权在文本内部怎么走。只有两章，因为它们是一笔账的两端：一章向前读，一章向后读。

编三，权交回去的那一刻。写完就是把权交回去——交回给文本自己、交回给读者、交回给判准。三章问同一个动作「结算」的三个侧面：账内还欠什么、这次结清把什么挤出了账、这本账本身被作品改了多少。

终编，权被收回的那一次。前面八章都还在「这一次写作」的范围里，终编讲下一次。而它讲的那一次收回，在所有记录上都是一个正项——交了、合格、按时。

这里有一个编者的选择要交代。第九章的检验三条假设全部失败，作者当场撤回了跨域主张，并给机制加了第三个条件。把一章「主要成果是一次失败」的稿子放在全书压轴，不是谦虚的姿态，是本卷体例的直接后果：如果一本讲「代价不进账」的书，把自己失败的那一次藏起来，它就在做自己批评的那件事。

五·九章互不引用，这既是它们的产生条件，也是本卷最脆弱的地方

有一件关于这本书的事实，编者认为读者应当在读正文之前就知道：九章之间几乎不互相引用。

把九章的文献表合起来，去重之后是二百零五条。其中被两章以上共同引用的，只有二十五组；剩下一百八十条，各自只出现在一章里。唯一被五章共引的是一篇写作认知的经典论文，其余几条也全部挤在写作研究与篇章语言学的一小片地上。

这不是疏忽，是取源方式的直接后果。

九章各自的做法，是把三门与语文教育距离很远的学科放到同一个问题上相撞，取它们各自最硬的一条判断，看它们在什么地方共同失效。第一章取的是成核动力学、免疫学的动力学校对和分布式共识；第二章取的是结晶学的阶段律、药理学的给予节律和组织学习；第三章取的是形态发生、交通流和机制设计；第九章取的是发育生物学的转决定、交通工程的容量跌落和科学社会学的多重发现。九章合起来动用了二十四门学科。

正因为各章从很远的地方带回本领域陌生的材料，它们才可能长出本领域原来没有的对象。如果九章共引一大批文献，那恰恰说明它们在同一片地上打转，谁也没有真的走远。从这个角度看，一百八十条各只出现一次的文献，是本卷取源纪律奏效的证据。

但同一件事换个方向看，就是本卷最脆弱的地方：

九章的可比性，目前完全依赖那一句母题与三处对切，不依赖任何共同的文献地基。母题一旦被推翻，九章立刻散成九篇不相干的论文。

这句话编者写在这里，不是为了显得诚实，而是因为它规定了读者应当怎样检验这本书。检验一本论文集，通常是逐篇挑毛病；检验本卷，最有效的一刀不在任何一章身上——在那句母题上。如果能证明「发生就是决定权的交接」这一半是空的，或者证明「被判为成功处的收回」根本不存在，那么九章即便每一章都成立，这本书也不再是一本书。

这也是本卷为什么要在每一编前加一篇编前语、在三对章之间硬切三刀、并且给每一章都配一节编者补节的原因。那些编者件不是装饰，它们是这本书唯一的承重结构。抽掉它们，剩下的是九篇独立论文；而九篇独立论文，构不成一个回答。

六·本卷不主张什么

一本提出九个读数的书，最容易被拿去做成九张评分表。所以有些话必须写在最前面。

不主张这九个读数越高越好。第三章明写拒绝质量单调性——高质量的预成写作可以有很低的读数，混乱的探索也可以有很高的读数；第二章的数据里，被评为最好的那一批先析率不低反略高；第六章撤回了「闭合率越高作文越好」；第八章明确要求先满足任务的最低合法边界，任何余量都不能让跑题的文本越级成为好作文。

不主张任何一个读数可以用来排名学生。第二章与第三章各有一条明文禁令，本卷把它提到前面：这些读数一旦进入奖惩，被测行为立刻改变，而最省事的达标方式恰恰摧毁读数的真实性。第八章说得更狠——任何「余量」指标一旦拿来考核学生，就会遭遇它自己批判的那个问题。

不主张预成执行是低级的。公文、法律文书、医嘱、安全规程、标准化报告里，高先析率与低端位是正当的，往往是必要的。第二章为此写死了五处不适用。

不主张本卷已经被证明。九章合起来是一份可执行的研究纲领，不是一份结论清单。九章里带真实统计量的是四章，其中两章的核心假设在数据面前失败过。

不主张本卷自己是从容通过的产物。按第九章的三条件——有不可推迟的截止、中间有过拥塞、最后被全部交付——本卷九章各自都是一次硬通的产物。它的未复流清单列在终编的编前语里，共三条，编者不为它们辩护。

七·这本书希望被怎样对待

不是被同意。

九章每一章都自带一条能杀死自己的观察，四章写死了到期日期的赌注，最早的一条到 2027 年年底，最晚的到 2029 年 8 月。这些赌注不是修辞——每一条都写明了「什么不算命中」，堵住了事后改阈值、事后换样本、事后把原量表换个说法重问一遍这三条退路。

其中获取成本最低的一条，在第九章：一名教师连续三年保留同一批学生的作文题材清单就够了。不需要任何新工具。如果那份清单显示题材范围在三年内单调收缩、从无回升，第九章对；如果停滞之后普遍出现跃升，第九章错，而且是被一份最普通的教研材料判错的。

一本书能提的最高要求，大概就是这个：请先去把它推翻。如果推不翻，再谈它说对了什么。

八·体例与凡例

四条，读正文之前知道即可。

一、每章一读数，每读数一条死亡条件。九章各自提出一个可清点的读数，并写明什么样的观察会让它失败。读数与死亡条件都汇在附录一。没有读数的章不入本卷——这条在选章时就是硬门槛。

二、不利结果原样保留在正文里。不移入附录，不作缺陷说明。九章里有四处修订是被自己的数据逼出来的：第三章删掉了整条第三层主张，第九章撤回了跨域主张，第二章收窄了适用区间，第六章撤回了「读数越高作文越好」。四处全部留在原位。

三、编者补节不改正文一个字。每章之后接一节编者补节，只补漏掉的祖宗、切开与本卷别处的重叠、写明该章没说出口的一处短板。凡分离线，一律附一条能让该章失败的观察。编者不替作者辩护。

四、各章文献表保留原编号，卷末另有汇总。九章原为独立论文，著录体例不一，正文的编号引注依各章自身编号运行。各章编号仅在本章内有效；卷末的汇总表按著者一年份归并，每条标出它在第几章出现。全书体例的统一列为再版事项。

导 读

给不同的读者的四条路

本卷九章各自独立成篇，都可以单独读。但九章之间有三处正面对切、一条贯穿的母题和九个互不相同的读数，按不同的目的进入，路线并不一样。

第一条·只有半小时

读导论第四节（母题的两半）与终编前语（本卷的自反与未复流清单）。前者告诉你这本书赌的是什么，后者告诉你它自己认为哪里最虚。两节合起来约六千字，足以判断值不值得往下读。

第二条·语文教师

按 第一章 → 第二章 → 第六章 走。

第一章给你一个课堂上一分钟能做的动作：当学生说「我有了」，不要问「中心是什么」，而问「把你现在最想保留的这一点拿掉，接下来三件你本来要写的东西会变哪几件？」——若答「那我就换个开头，后面照写」，说明这个开头还没有取得续写权。

第二章给你一条与常识相反的处方：诱导期内既不自由写作，也不列提纲，什么可交付的东西都不产出。这一条会让很多人不舒服，第二章自己也写明了五处不适用，读的时候请连边界一起读。

第六章给你一张可以带进修改课的账本：不再笼统说「这里再具体一点」，而是问「第二段提出了『不是没努力，而是目标被别人定义』，后面哪一句真正处理了『别人怎样定义』？如果没有，这是一笔核心欠账，补它之前不要加新材料。」

第三条·研究者

按 第五章 → 第三章 → 第四章 走，然后回到对切一。

第五章是全卷因果识别设计得最干净的一章——四臂随机、六个替代解释各配对照组，只是一次也没有被执行。第三章是全卷刀口最窄的一章，且它的第三臂设计恰好可以用来拆分自己的混淆。第四章给了最完整的一套编码纪律（冻结、双标注、不可判定单列、证伪优先）。

三章读完再看对切一：第三章与第五章都主张「一字未改的文字可以已经变了」，而两章的自变量在不同的地方。那一节给了双向的判决性对照，两章各有一条会因此失败的路。

第四条·关心评价与制度的人

按 第七章 → 第八章 → 第九章 走，然后回到对切二。

第七章说，写好作文最危险的时刻不一定是失败的时候，而可能恰恰是某一项已经做得很好、系统因此允许我们宣布「这一轮完成」的时候。第八章说，一套判准一旦被公开、教会、用于优化，它就不再能证明它原本要证明的东西。第九章说，一次被记为成功的交付会扣掉此后的通行能力，而这个减项不进任何记录。

三章合起来是一条完整的制度批评，但三章都没有把它写成一句口号，因为三章都被自己的数据修剪过。

三处对切怎么读

对切不是调和。编者把两章逼到同一张表上，写出一个能让其中一方失败的观察，然后交给读者与后来者去跑。编者不预判胜负。

- 对切一（第三章 ↔ 第五章）：文本不变，变的是哪一种角色——工作角色还是篇章角色。给了双向交叉的判决性对照。
- 对切二（第六章 ↔ 第七章）：对同一个动作（结算）的相反估价。三种结局都写死了，包括「两章其实是同一个量的两种说法，本卷应当合并成一章并承认多设了一个概念」。

- 对切三（第二章 ↔ 第九章）：一个没有解的两难——调用现成形态会消耗原料，不调用则更容易撞上截止。编者拒绝在这里给折中的时长建议，理由写在那一节里。

编者补节怎么读

每章之后接一节编者补节，只做三件事：补上该章漏掉的祖宗并写出分离线；切开该章与本卷别处的重叠；写明该章自己没有说出口的一处短板。凡分离线，一律附一条能让该章失败的观察。

补节不改动正文一个字，也不替作者辩护。如果你读完某一章觉得「这不是早就有人说过吗」，那句话很可能已经写在该章的补节里，并且附了一条能判决的实验。

一句提醒

本卷九个读数，没有一个是成绩单。第二章与第三章各有一条明文禁令，第八章说得最直接：任何「余量」指标一旦拿来考核学生，就会遭遇它自己批判的那个问题。

如果你打算把书里任何一个读数带进考场或评优表，请先读该章的「反噬」一节，再决定。

数据·证据·人机分工声明

本卷九章由九篇独立研究改编而成。为免各章重复声明、口径不一，编者把原分散在各章末尾的数据说明、证据边界与人机分工声明合并于此，各章末尾此后只留文献。

一·数据与可复算

本卷九章中，有四章带有在成文之前完成的预注册统计，其数据全部来自公开可获取的资源，任何人可按各章所载的编码规则与阈值复算：

第二章使用一份公开的中学生作文语料，取其前四万行，剔除清洗后汉字少于一百的篇目，最终样本三万九千六百七十二篇；统计用秩相关与非参数检验。

第三章使用一份公开的英文写作键击语料的任务级元数据，样本一千零六名参与者；模型与阈值在取数之前写定。

第六章使用一份公开的学生作文评分语料的同一题组，按人工总分分层各取两篇，共十二篇，做探索性编码；该编码为单编码者试跑，不作为量表效度证据。

第九章使用六个公开源码仓库的完整提交历史，纳入文件级改动四万二千一百四十六次；仓库选择规则在取数之前写定，不按星标数、不按活跃度筛选，也不在看过结果之后更换仓库。

各章数据的具体地址、字段与清洗规则载于对应章节的方法节与附录。

二·预注册

上述四章的预注册文本均先于统计写就并存档，事后未作修改。凡使用回溯数据或事后分析者，各章正文均已明确标注为事后分析，不与预注册结果混报。

三·不利结果

本卷的体例是：不利结果原样保留在正文里，不移入附录，不作缺陷说明。

第三章的预注册二次项方向与预期相反且不显著，作者据此删掉了整条第三层主张；第九章的三条假设全部未通过，作者据此撤回了跨域主张并给机制增加了第三个条件；第二章的分布诊断暴露了读数在高值区失去分辨力，作者据此把适用区间收窄；第六章的探索性编码出现天花板，作者据此撤回了「读数越高作文越好」。

四处修订全部保留在各章正文中，编者一处未删。

四·证据边界

本卷各章对其取源学科的叙述依据公开学术资料。跨学科转译不意味着来源学科已经认可这些概念在写作研究中的用法；各章提出的新对象与读数属于研究假说，不是已获确证的结论。

各章的读数均为机制读数，不是质量评分。本卷明确禁止把其中任何一个读数直接用于学生排名、评优或处罚。一旦这类读数进入奖惩，被测行为立刻改变，最省事的达标方式恰恰摧毁读数的真实性——这一点第二章与第三章各有专节论证。

五·人机分工

选题、取源学科的选定、成文方向与全部判断的取舍由著者决定。文献核对、语料获取、统计计算、初稿撰写与文档排版使用了生成式人工智能辅助。各章预注册文本均在运行统计之前写定。最终的理论主张、引文准确性、研究伦理与对外使用责任由署名著者承担；本卷对其中的全部错误负责。

章节来源说明

本卷九章由九篇独立论文改编而成，九稿均由著者撰写。各章保留原稿的节次编号、读数定义、预注册记录与真跑结果；编者未改动正文一字。

每章之后接一节「编者补节」，由编者撰写，只做三件事：补上该章漏掉的祖宗并写出分离线；切开该章与本卷别处的重叠；写明该章自己没有说出口的一处短板。凡分离线，一律附一条能让该章失败的观察。

三处对切分别接在第五章、第七章与第九章之后。对切不做调和，只把两章逼到同一张表上，写出一个能让其中一方失败的观察，交给读者与后来者去跑。编者不预判胜负。

本卷各章的分界与自评均由著者单方作出，未经外部裁决。按第九章的判断，一次由本人签发的成立性宣告不构成终局，故本卷不对自身作出成立性宣告。

九个读数

每一个都有清点规则，每一个都有一条能杀死它的观察

章	读数	一句话定义	死亡条件
一	CDR6 续写依赖率	删掉候选核后，六个独立语义决策里有几个发生类别性改变	删核不改变任何下游决策
二	ϕ 先析率	产出中此前已以成句形态存在的比重（按八字连续串）	提示「别用你以前写过的说法」后 ϕ 显著下降
三	E_T 端位效应	同一段文字由可编辑嵌入态移到只读侧栏后，随后 120 秒事件转移分布的散度	两种条件下的分布与置换零分布不可分
四	GTR 续生转移率	承接 = 1 且改场 = 1 且转位 = 1 的新增单位占比	打乱核心段序而读者的下一步判断不变
五	ZRM 零字面角色迁移率	字面不变而篇章角色改变的对齐单元占比	迁移率不高于对齐与标注噪声
六	ECR 退场闭合率	已结清或合法限界的加权续写义务占比	作者退场后读者的等待点与账本不重合
七	余损迁移序列	本轮通过后余损落在另外两域的哪一域，决定下一轮终点	余损从不影响下一轮焦点
八	CFR 判准回写余量	一篇作文回写了多少判它的判准	判准在作品前后不变
九	R 复流期	一次硬通之后重新以不排队方式通过所需时间（可取 ∞ ）	停滞段之后普遍出现跃升

目 录

著者简介	3
前 言	3
导 读	8
数据·证据·人机分工声明	10
章节来源说明	11
九个读数	13
导 论	15

编一 权第一次落到文本身上	18
第一章 作文怎样启动 自持起写核与续写权移交	21
第二章 先析 最先成形的那段合格文字如何使另一段永远不再成形	50
第三章 作文如何发生 局部终点转为下一轮条件的端位转换	79
编二 权在文本内部怎么走	108
第四章 作文如何持续长下去 续生位转移	109
第五章 未改之句如何被改写 作文发生的稿际余载机制	132
对切一 第三章 ↔ 第五章：文本不变，变的是哪一种角色	166
编三 权交回去的那一刻	167
第六章 作文怎样完成 退场闭合态	168
第七章 如何写好作文 结算余损与追余写作	194
对切二 第六章 ↔ 第七章：对同一个动作的相反估价	222
第八章 何谓一篇作文的好 判准回写余量	223
终编 权被收回的那一次	256
第九章 通扣 一次全部交上去的写作如何在记录之外扣掉此后的通行能力	257
对切三 第二章 ↔ 第九章：一个没有解的两难	285
结 语 账外的那一栏	286
参考书目 卷末统一文献表	287
附 录	303
附录一 九个读数总表	
附录二 九个新对象的名与英文对应	
附录三 写死日期的赌注一览	

附录四 三处对切索引

附录五 本卷的未复流清单

附录六 编者补节所补的祖宗一览

※ 每章之后接一节「编者补节」，未单列于目录。页码为终稿实测。

导论

一·一样东西一直在换手

卷一问的是「这一篇凭什么是这一篇」，答案落在身份上。卷二问的是另一件事：从没有到有，中间到底发生了什么。

先看三个都不报错的场面。

一个学生每周一篇周记写了两年，素材本记满，范文读过两百篇。分数没有下降，多数时候还略有上升。他的语文老师说了句话：这孩子现在写什么都像作文，就是不像他自己写的。

一个班的交稿时间戳，初二上学期分布在布置后第二天到第六天之间，形状接近均匀；到初三上学期，同一批学生、同一个老师，提交时刻高度集中在截止前六小时以内。分数没有下降，完成率两个学期都是百分之百。

一位教师连续三年上作文课，评价一直很好。第一年带学生写「给十年后的自己」，第二年写「我家阳台上的一盆植物」，第三年还是这两个题目。没有人问过她为什么不换。

三件事有同一个形状：一件东西在减少，而所有的仪表都显示正常。分数在涨，完成率满格，听课本上记的是「环节完整、指导到位」。

减少的是什么？本卷九章从九个方向回答同一句：减少的是决定下一步的权力。

这个说法需要立刻变窄，否则它只是一句漂亮话。九章各自把它压成一个可清点的东西：第一章问下一步的选择是不是由已经写出的那个约束核逼出来的；第二章问这一次写出的句子在这一次之前是不是已经以成句的形态存在；第三章问同一段文字换一个工作位置之后，后续操作的分布会不会变；第四章问下一段是在还前文的哪一笔账；第五章问被删掉的那句话的职责被谁接了过去；第六章问这一篇还欠着哪些没结的账；第七章问这一轮宣布成功之后损失被挤到了哪里；第八章问判它的那把尺子被它改了多少；第九章问这一次交上去之后，此后同类写作的通行能力被扣掉了多少。

九个问题，九个读数。它们问的是同一件事的九个位置：权在谁手里，正在往哪里走，以及在哪一步被悄悄收回。

二·三种现行答案，与它们共同的那个假定

「作文如何发生」不是没有答案，而是有三套，且这三套互不相容。

第一套是步骤论。审题、立意、列提纲、写开头、展开、修改、润色。它的好处是可教、可查、可评分。它把发生等同于次序：前一步完成，后一步接手。

第二套是过程论。写作不是线性誊录，是规划、转译、回读相互嵌套的递归活动；目标会在写作中改变，成熟写作者在内容问题空间与修辞问题空间之间往返。这一套已经把步骤论拆掉了，「作文不是结果而是过程」今天不是新判断，是研究常识。

第三套是能力论。写不出来是能力、输入或条件三者之一的不足，于是分别在成品端、过程端、条件端下药：多读、多写、多改，或者改善题目、评价与反馈。

三套各有根据，也各有解释不了的东西。步骤论解释不了「同一个顺序可以由完全不同的机制产生」——甲的提纲是一个暂时假设，写到第三段被反例逼着改写中心；乙的提纲是一张不可撤回的施工图，后文只负责填空。日志上两人都是「规划—输入—修订」。过程论解释不了「有过程」为什么不等于「过程在长出下一步」——一个人可以按预先写好的提纲连续誊写，段落越来越完整，日志非常丰富，却几乎没有任何局部产物反过来改写后面的行动。能力论解释不了本卷开头那三个场面：那三处的能力指标全部正常甚至上升。

三套底下共享着同一个假定，而这个假定谁也没有回头看：

凡发生过的推进，都会在产出或记录上留下量。

写了多少字、改了多少次、用了多长时间、拿了多少分、交没交、写得怎样——发生被默认为可清点，而清点的口径是量。于是三套分工明确：步骤论去记次序，过程论去记事件，能力论去记水平。它们分工的前提是那笔账总在三处之一。

本卷九章共同否定的，正是这一层。它们说：有一类推进与一类损失，三处都不记，而且不是疏漏。结果处不记，因为它合格；过程处不记，因为它按时；水平处不记，因为当事人自己也说不上哪里不对——这一次他确实做到了。这一类东西之所以不进账，恰恰是因为它发生在一次每一项都达标的通过里，而所有的记账规则都是围绕「哪一项没达标」设计的。

三·九章共同的上游：已写出的东西会成为条件

必须先把一位祖宗请到台前，否则本卷会被误读成一次集体的重新发明。

「已经写出的文本反过来成为下一步的生成条件」——这句话不是本卷提出的。写作过程研究在把已写出的文本（the text produced so far）放进任务环境的那一刻，就已经写下了它。此后四十年，围绕它积累了大量工作：写作者在停顿中重读它、用它评估、也用它生成新内容；重读与文本质量的相关

早有报告；键击记录方法学能逐键还原它的生长；「写着写着才发现要写什么」也已经有了实测量表与读数。

本卷九章里，只有第一章正面引了这一支。这是一处编者必须补上的欠账，而且不能靠脚注补——它规定了整卷的位置。

补上之后，本卷的主张要相应地压窄，压成一句：

上游已经证明已写出的文本会成为条件；本卷九章各自要多拿的，是这个「成为条件」的那一刻在哪里、由什么读数指认、以及它在什么情况下不发生。

九章各自多拿的那一小块地，可以逐条写出来：

章	上游已有	本章多拿的那一块
一	已写文本进入任务环境	它什么时候开始自己产生并排除下一步（续写权移交的时刻）
二	已写文本会被重新调用	它同时吃掉了另一段的成形条件（调用有代价）
三	已写文本影响后续	影响的是工作角色而不只是内容可及性（内容不变、位置变，后续就变）
四	已写文本改变后续目标	改变以逐单位可核验的义务转移的形式发生
五	已写文本会被修改	没被修改的那些也在改变，改变的是职责
六	修订会趋于收敛	收敛的判据是作者可以退出，不是作者觉得写完了
七	多轮改进会提高质量	每一轮成功都把损失挤到未被本轮标准计入的域
八	判准用来评价作品	作品会回写判准
九	一次写作会影响下一次	影响是一个减项，且它挂在被记为成功的那一次上

这张表就是本卷对上游的全部欠债与全部增益。凡本卷主张超出这张表右栏的，编者视为越界。

四·母题的两半

全书的脊梁是一句话：

作文如何发生，看下一步由谁决定；每一次被判为成功的结算，都在无声处收回决定权。

它有两半，两半都要拿出来单说，因为分开之后各自都可以被证伪。

肯定的一半：发生就是决定权的交接。

这一半规定了观察对象。不看写了多少，看下一步由谁决定：由外部题目、模板、教师提示、机器补全决定，还是由已经写出的这一小块自己决定。第一章把这一刻命名为续写权移交，并要求它可以被反事实检验——把候选核删掉，后面若干个独立选择是不是跟着一起变。第三章把同一件事推到更窄的刀口上：内容一个字不改、可见性一点不减，只把那段文字从可编辑的正文位

置移到只读的侧栏，随后一段时间的操作分布会不会变。第四章与第五章分别读它向前走（下一段在还谁的账）和向后走（退出的账被谁接了过去）。

这一半的死亡条件很明确：如果决定权的交接无论怎么干预都读不出差别，肯定的一半就是空的。第三章为此写死了一个日期赌注。

否定的一半：收回发生在被判为成功的地方，且不报错。

这一半才是反直觉的那一半。常识认为写得越顺、评价越好，说明能力在长。本卷说，被判为成功的那一刻，恰恰是决定权被收回的时刻。

第二章的收回发生在受表扬处：最先成形的那段合格文字被批改、被打分、被读给全班、被抄进素材本、被当作下一届的范文，于是它成为环境中新的现成形态，下一轮更偏。第七章的收回发生在宣布完成处：宣布完成的那把尺子，只看见它被设计去看见的东西，损失从没设字段的那本账上逃逸。第九章的收回发生在全部交付处：交上去了，分数正常，记录完整，而此后同类写作的通行能力被下调。第一章的伪启动、第三章的惰性成形说的是同一格的另一面——形式全部合格而权从未交出。

这一半的死亡条件同样明确：如果在一批被判为成功的写作之后，同类写作的内生依赖率、可选题材范围、可达路径分布都不下降，否定的一半就是伪的。

两半合起来才是母题。只有肯定的一半，本卷退回成过程研究的一个变种；只有否定的一半，本卷退回成一句关于评价制度的牢骚。

五·四编的次序，以及为什么收在「被收回的那一次」

编一 权第一次落到文本身上（第一、二、三章）。先讲权从外面落到里面：什么时候作文开始自己出题。三章各用一种办法证明纸面骗人——删掉候选核看下游变不变（第一章）、逐句追问这句在这一次之前是否已经成句（第二章）、只改工作位置不改内容（第三章）。第三章排在编末，因为它的刀口最窄，必须先由前两章把「纸面不可信」立住，读者才不会把它当成一项界面研究。

编二 权在文本内部怎么走（第四、五章）。权已经落到文本身上之后，它在文本内部怎么流动。第四章向前读——下一段的资格由前文欠下的账决定，不由主题相关决定；第五章向后读——退出的那句话的职责被一个一字未改的句子接了过去。一前一后，构成同一条账目的两端。

编三 权交回去的那一刻（第六、七、八章）。写作总要结束，结束就是把权交回去：交回给文本自己、交回给读者、交回给判准。三章问同一个动作——结算——的三个侧面：账内还欠什么（第六章）、这次结清把什么挤出了账（第七章）、这本账本身被作品改了多少（第八章）。

终编 权被收回的那一次（第九章）。全书收在这里的理由，与卷一收在制度上的理由是同一个：前面八章讲权怎么走、怎么交回去，都还在「这一次写作」的范围里；终编讲下一次。而它讲的那一次收回，在所有记录上都是一个正项——交了、合格、按时。

这里要说明一处编者的选择。第九章的真跑三条假设全部失败，作者当场撤回了跨域主张，并把机制修订为需要第三个条件（不可推迟的截止）。把一章「主要成果是一次失败」的稿子放在全书压轴，不是谦虚的姿态，是本卷体例的直接后果：如果一本讲「代价不进账」的书，把自己失败的那一次藏起来，它就在做自己批评的那件事。

六·本卷纪律

一、每章一个读数，每个读数一条能杀死它的观察。

章	读数	一句话定义	死亡条件
一	CDR6 续写依赖率	删掉候选核后，六个独立语义决策里有几个发生类别性改变	删核不改变任何下游决策
二	ϕ 先析率	产出中此前已以成句形态存在的比重（按八字连续串）	提示「别用你以前写过的说法」后 ϕ 显著下降
三	E_T 端位效应	同一段文字由可编辑嵌入态移到只读侧栏后，随后 120 秒事件转移分布的散度	两种条件下的分布与置换零分布不可分
四	GTR 续生转移率	承接 = 1 且改场 = 1 且转位 = 1 的新增单位占比	打乱核心段序而读者的下一步判断不变
五	ZRM 零字面角色迁移率	字面不变而篇章角色改变的对齐单元占比	迁移率不高于对齐与标注噪声
六	ECR 退场闭合率	已结清或合法限界的加权续写义务占比	作者退场后读者的等待点与账本不重合
七	余损迁移序列	本轮通过后余损落在另外两域的哪一域，决定下一轮终点	余损从不影响下一轮焦点
八	CFR 判准回写余量	一篇作文回写了多少判它的判准	判准在作品前后不变
九	R 复流期	一次硬通之后重新以不排队方式通过所需时间（可取 ∞ ）	停滞段之后普遍出现跃升

九章九个读数，每一个都有清点规则，每一个都有一条能杀死它的观察。这是本卷相对卷一的形制升级：卷一给出的是九个新对象，卷二给出的是九台可清点的仪表。

二、不利结果原样保留。本卷九章里有真实统计量的是四章。其中：第三章的预注册二次项方向与预期相反且不显著，作者据此删掉了整条第三层主张；第九章的三条假设全部未通过，作者据此撤回了跨域主张并给机制加了第三个条件；第二章的分布诊断暴露了读数在高值区失去分辨力，作者据此把适用区间收窄；第六章的探索性编码出现天花板，作者据此撤回了「读数越高作文越好」。四处修订全部保留在正文里，编者一处未删。这是体例，不是缺陷说明。

三、三处对切。本卷三对章之间存在正面重叠，编者不做调和，只把它们逼到同一张表上让数据裁决（见后附三处对切）。

四、本卷不主张什么。

1. 不主张这九个读数越高越好。第三章明写拒绝质量单调性——高质量的预成写作可以有很低的 E_T，混乱的探索也可以有很高的 E_T；第二章的数据里，被评为最好的那一批先析率不低反略高；第六章撤回了「ECR 越高作文越好」。九章没有一章把自己的读数当成绩单。
2. 不主张任何一个读数可以用来排名学生。第二章与第三章各有一条明文禁令，编者把它提到导论：这些读数一旦进入奖惩，被测行为立刻改变，最省事的达标方式恰恰摧毁读数的真实性。
3. 不主张预成执行是低级的。公文、法律文书、医嘱、安全规程、标准化报告里，高先析率与低端位是正当的，往往是必要的。第二章为此写死了五处不适用。
4. 不主张本卷已被证明。九章合起来是一份可执行的研究纲领，不是一份结论清单。
5. 不主张本卷自己是从容通过的产物。按第九章的三条件——有不可推迟的截止、中间有过拥塞、最后被全部交付——本卷九章各自都是一次硬通的产物。其未复流清单列在终编编前语。

编一

权第一次落到文本身上

共同动作：把「开始了没有」从看纸面改成看依赖。

这一编三章处理的是同一个转折：写作从被外面推着走，变成被已经写出的东西逼着走。三章都不接受纸面证据。

第一章的办法是删核：把学生说「我有了」的那一点拿掉，问接下来本来要写的三件事会变哪几件。若答「那我换个开头，后面照写」，说明这个开头尚未取得续写权。它由此得到一个反常预言——作文的启动可以晚于第一字，也可以早于第一字。

第二章的办法是追来源：这一次写出来的句子，在这一次开始之前，没有以成句的形态出现过。它由此发现，最先成形的那段合格文字不是白来的——它和另一段共用同一批原料，成形即意味着另一段不再成形。

第三章的办法是换位：内容一字不改、可见性一点不减，只把那段文字从可编辑的正文移到只读的侧栏。这一刀专治「不就是因为看得见吗」这条最强的竞争解释。

为什么第三章排在编末。它的刀口最窄，窄到很容易被误读成一项编辑器界面研究。必须先由第一章立住「形式合格不等于启动」、由第二章立住「合格文字有代价」，读者才会明白：被移到侧栏的那一段文字，正是前两章反复讨论的那个东西。

这一编内部有一处必须先讲的重叠。第一章的伪启动（表面痕迹稳定而下游依赖为零）与第三章的惰性成形（局部文本稳定而无因果再入）是同一格的两种测法，不是两个概念。本卷取第一章的「伪启动」为正名，因为它的定

义最严——形式全部合格而无一后续选择依赖它；第三章处改为引用，并说明视角差：第一章看下游依赖，第三章看操作分布。

还有一处跨章的动作重叠。第二章的第四步（先析物写回环境，成为下一次的现成形态）与第三章的第四步（产物写回环境）是同一个动作的两种读法：第二章读它消耗了什么，第三章读它改变了后续可行操作的分布。两章互不引用，编者在第二章末补一句划开，不另起对切。

第一章 作文怎样启动 ——自持起写核与续写权移交

第一字不是开始：作文从“没有”到“有”的自持起写核模型

——物理化学成核 × 免疫动力学校对 × 分布式共识的二阶碰撞（3.9 版）

学科通融专栏 · 语文教育学专题 · 2026 年 8 月

要

一篇作文究竟在什么时候开始存在？传统作文教学常把“开始”放在审题、选材、构思、列提纲或落下第一字上；过程写作理论则证明，写作并非产品之前的一次性准备，而是规划、转写与修订不断递归的过程。然而，这些理论主要回答“写作者在做哪些活动”，提供一个可以把“正在准备写”与“作文已经自己开始生长”分开的边界判据。本文依 SDE 多学科通融创新法 3.9 版，将问题判为 4v 题，并从新思想前沿中选取物理化学、免疫学、分布式系统三个远距成熟领域，分别抽取经典成核、动力学校对与共识提交三条路径的启动判据。三家分别把启动结算在局部稳定核、连续校验路径与共享承诺场上；碰撞后发现，它们共同把“启动”误作某个固定终点，而其自身的失败材料又显示：任何单一终点都可能产生假启动。本文据此提出“续写权移交”与“自持起写核”两个相连的新模型。作文启动，不发生在第一想法或第一字出现之时，而发生在一个最小约束簇第一次取得续写权之时——此后，已经形成的作文能够自主、筛选并协调下一步选择，外部题目、教师、模板或 AI 不再承担逐步驱动。为把该判断变成可裁决命题，本文提出“续写依赖度 R6”及反事实删核实验，建立表面痕迹稳定性 × 下游依赖性的二维判别格，并以 KLiCKe、写作云数据和 ScholaWrite 三个公开过程数据源进行字段审计。审计结果是不利的：现有公开数据均不能直接计算核心读数，因为缺少“早期选择—后续选择”的语义依赖边和变量分支。本文因此不声称模型已获经验确证，而将这一字段缺口转化为可执行的实验方案与课堂观察协议。该模型的教育含义是：帮助启动作文，不应等同于替学生提供更多开头、素材和步骤，而应帮助其形成一个能够逼出下一步、排除下一步并承受修改的内生约束系统。关键词：作文启动；写作过程；自持起写核；续写权；成核；动力学校对；分布式共识；二阶碰撞

核心命题：作文启动，不是第一想法出现，也不是第一字落下，而是“续写权”第一次从外部提示移交给正在形成的作文自身；取得这种续写权的最小约束闭环，本文称为“自持起写核”。

一、问题提出：我们知道怎样写，却不知道“什么时候算已经开始写”

语文课堂里有一种极常见却很少被理论单独命名的困境：学生已经读完题目，圈了关键词，老师也带着全班讨论了素材，甚至纸上已经出现了一个漂亮开头，但学生仍会抬头问：“然后呢？”另一名学生可能长时间没有落笔，只在草稿纸上改动几个词；一旦真正开始写，后面的段落却连续涌出。若把第一名学生判作“已经启动”，把第二名学生判作“尚未启动”，我们的判据其实只是把可见文字当成了作文存在的证明。教学经验不断提醒我们，这个证明并不可靠。第一字可以早于真正的生成，沉默也可以晚于真正的生成。

写作研究对此并非一无所知。Rohman（1965）把 pre-writing 称为发现阶段；Emig（1971）通过有声构思研究真实写作者；Murray（1972）要求把写作从产品转为过程；Flower 与 Hayes（1981）进一步把规划、转写、复查写成递归的认知过程。Bereiter 与 Scardamalia（1987）区分知识陈述与知识转化，Galbraith（1999）强调文字生产本身能够生成新的认识。到键击记录、版本记录与写作云研究，研究者甚至可以逐秒看到停顿、插入、删除、回返和计划—转写之间的切换。问题因此不是“过程有没有被研究”，而是一个更窄、更基础的问题：这些过程中哪一个事件，使系统从“尚未有一篇作文”跨到“已经有一篇正在生长的作文”？

这个边界不能由最终质量倒推。低质量作文也会真实启动，高质量成品也可能由模板、逐句提示或 AI 连续外包而来。它也不能由努力强度决定：学生可能想了二十分钟仍在漂浮，也可能被一个小细节突然牵住，三十秒后就进入稳定生成。更不能由“有没有计划”决定：计划可能只是老师给出的外部路线图，学生每走一步仍需再次询问。真正需要解释的是：某一刻之前，下一步主要由外部继续推动；某一刻之后，已经形成的东西开始反过来规定下一步。本文把这条边界称为“作文启动边界”。

因此，本文的核心问题不是“为什么有人不愿写”，也不是“怎样提高写作动机”，更不是“怎样写一个好开头”。它问的是发生学意义上的 How：一篇作文怎样从没有变成有？如果把作文看成一个能持续生成后续选择的过程，那么启动必须是一条可描述的发生路径，而不是一个心理标签。

二、方法：3.9 版如何把“启动”从熟悉理论中重新挖出来

本研究采用 SDE 多学科通融创新法 3.9 版作为选源和碰撞程序。按照该版本的题型纪律，“怎样从没有变成有”属于 How 题，要求的不是一个静态定义，也不是一个单因解释，而是一条从当前状态走向新状态的路径。How 型在 3.9 版中按三个不同焦点组织来源：有的路径把稳定对象作为结算点，有的把生成序列作为结算点，有的把关系场的共同状态作为结算点。三家必须各占一个焦点，否则只是同一轴上的学术辩论。

在正式碰撞之前，本文先执行两项前置检查。第一项是“未吸收度”：不是问学科名字离作文远不远，而是问我们准备从它那里拿来的那条具体答案，是否已经进入写作理论的固定术语、教材章节或标准手册。检索以“写作/作文+成核”“写作/作文+kinetic proofreading”“写作/作文+consensus/Raft”等学科词检索，同时把命题剥到形式层，改搜“一个局部候选何时算真正开始”“连续校验后才允许进入下一阶段”“局部完成何时变为系统承诺”等大白话结构。现有写作研究确有 prewriting、planning、goal setting、knowledge transforming、writing as discovery、keystroke bursts 等近邻，但未发现经典成核的临界核判据、免疫动力学校对的串行误识别抑制判据或分布式共识的提交判据成为作文教育的固定理论部件。故三源均判为“干净”，但这一判断只表示“未发现已被标准化吸收”，不等于证明历史上无人做过零散类比。

第二项是“理论厚度”。成核理论有 Gibbs 热力学、Becker-Döring 动力学、经典与非经典成核的长期争论及稳定的失效条件；动力学校对从 Hopfield、Ninio 到 McKeithan 及其后 T 细胞抗原辨识研究，已经形成关于等待时间、步骤数、能耗、反馈和灵敏度—特异性权衡的成熟传统；分布式共识从 FLP 不可能性、Paxos、PBFT 到 Raft，有完整故障模型、安全性、活性与成员变更理论。三家均能给出奠基层、内部争论、反例传统和权威教材，故均判“厚”。

三门学科的焦点安排如下：物理化学的成核把“出现一个足以继续长大的局部核”作为终点，属于焦点 S 路径；免疫动力学校对把“候选信号连续走完一串时间与状态检验”作为终点，属于焦点 D 路径；分布式共识把“一个局部提议成为系统可承诺、不可由单节点任意撤回的共同状态”作为终点，属于焦点 E 路径。三者不是为了凑出三个比喻，而是对同一个“何时才允许宣布启动”给出三个互不相容的结算时点。

三、第一条路径：物理化学成核——新东西先以一个很小的“核”开始

相变最有启发性的地方，在于宏观新相出现之前，体系中早已不断产生微小涨落。过饱和蒸汽里会偶然出现分子团簇，过冷液体里会出现短暂的局部有序；绝大多数团簇很快消散。经典成核理论把这个问题写成自由能竞争：形成体相新相可以获得体积自由能收益，但制造界面要支付表面能成本。小团簇表面积相对过大，容易回缩；当团簇越过临界尺度，继续长大的方向开始占优（Becker & Döring, 1935; Vekilov, 2010）。

这给“作文启动”提供了一个极强的第一判据：不是出现任何文字都叫开始，而是某个局部结构开始具有“继续长大比消失更有方向性”的时候，才算新对象真正起头。一个学生写下“有一天阳光很好”可能只是涨落；删掉它，下面写什么完全不变。另一名学生写下“外公把那条太小的鱼又放回水里”，这一个细节也许马上牵出外公的性格、孩子当时的困惑、之后对“占有”的重新理解。后者虽然只有一句，却已经像临界核一样，让未来文本的某些方向变得更可能，让另一些方向变得不自然。

成核路径因此把启动时点推得很早：宏观作文尚未成形，一个很小的局部就可能已经取得发生资格。若一定等到提纲完整、段落成形甚至中心明确才宣布“开始”，就像一定等晶体长到肉眼可见才承认成核，理论上已经把最关键的相变起点漏掉。这个观点对课堂有直接挑战：教师常用“先把结构想完整再写”保护学生，成核却允许先出现一个小而有方向性的实体，再由它组织后续。

但成核理论自己的边界也阻止我们把“找到一句好话”神化。临界附近仍有明显随机性，经典模型在许多体系中会被多步成核、非球形团簇、非经典前驱体和复杂反应坐标修正。所谓“核”不能只按大小识别，更不能把一切后来保留下来的句子事后封为起点。它真正贡献的不是“作文像晶体”这个类比，而是一个可反驳的时序主张：启动可以发生在宏观成品之前，而且启动点应当能在后续生长方向上留下非对称影响。

四、第二条路径：免疫动力学校对——不能因为“像”就过早启动

如果成核理论鼓励我们尽早承认一个局部核，动力学校对恰好从相反方向发出警告。Hopfield（1974）和 Ninio 提出的动力学校对思想说明，高特

异性识别不能只依靠一次结合的平衡差异；系统可以通过耗能的中间步骤和时间延迟，让错误底物更有机会在完成流程前脱落。McKeithan（1995）把这一结构引入T细胞受体信号转导：配体与受体的结合持续时间会影响信号是否有机会完成连续步骤，从而放大正确与错误配体之间的辨别。

迁入写作启动，核心不在“免疫系统会检查”，而在于它否定一个直觉：只要某个候选够稳定、够吸引人，就应立即让它掌权。学生在作文开始阶段经常出现“假核”：一句押韵的金句、一个网络热梗、一个很煽情的故事、一个老师刚讲过的立意都可能迅速占住纸面。它们像一个高亲和但错误的配体，短时间内很有存在感，却未必能够通过后续问题：它真的与题目张力相连吗？它能产生新的材料选择吗？它和另一段经验能建立关系吗？它允许反例进入吗？如果这些关口一碰就散，那么越早围绕它扩写，返工成本越高。

于是动力学校对给出第二个启动标准：一个候选不应因“出现”或“暂时稳定”就被宣布为作文起点；它至少要连续改变几次真实选择，并在这些选择中存活。这里的“校对”不是教师评分，也不是把学生困在无穷自检中，而是时间性的：只有能够在后续选择不断到来时保持组织作用的候选，才从灵感升级为路径。与成核相比，这一标准明显把启动时点向后推。成核说“核越过阈值就可以算开始”，动力学校对说“过早宣布开始正是系统产生假阳性的方式”。

这条路径也有自己的代价。校验步骤越多，特异性可以提高，但启动速度会下降，弱而真实的信号可能被错杀。免疫学由此反向限制本文：作文教学不能把“启动判据”变成新的繁琐审题表。若学生必须先回答十几个问题才能写第一句，我们只是把假启动风险换成了启动抑制。后文提出的起写核必须足够小，且校验必须嵌在真实续写选择中，而不能变成额外考试。

五、第三条路径：分布式共识——一个节点“写下了”不等于系统已经承诺

分布式系统面对另一种启动难题：一个节点已经做出决定，为什么系统还不能说这件事已经成立？网络会延迟、分区，节点会宕机，消息会重发，不同副本可能短暂拥有不同日志。FLP 结果说明在完全异步且可能发生故障的模型里，确定性共识存在根本限制；Paxos、PBFT 以及 Raft 等协议通过明确安全条件、领导者、日志复制、法定人数与提交规则，在可接受假设下让多个副本形成一致历史（Fischer, Lynch, & Paterson, 1985; Lamport, 1998; Castro & Liskov, 1999; Ongaro & Ousterhout, 2014）。

Raft 最适合说明这里的直觉差异：领导者把一个新条目追加到自己的日志，并不意味着该条目已经成为系统历史。它还要复制到足够多节点，并满足提交与安全约束；在此之前，领导者故障或任期变化都可能使局部记录无法进入最终一致历史。换句话说，“已经发生过一次局部操作”与“系统对它形成了承诺”是两个不同状态。SDE 新思想前沿也把分布式系统中的确定性执行、全局顺序与一致性问题放在关系场的 E 位讨论，强调局部动作不能自动代表整体状态。

这对作文启动提出第三次推迟。学生可能已有一个稳定想法，也可能围绕它连续写了几句，但若这些句子彼此仍是孤岛，每一段都可以换成另一套而不牵动其他地方，那么作文作为整体仍没有形成“承诺场”。真正的启动至少要出现两处以上的下游决定被同一个早期约束共同牵住：例如“外公放鱼”不仅决定写哪件事，还改变叙述顺序、观察角度和结尾判断。此时删掉那个早期约束，后面不是少一句，而是一串选择都要重新协商。

分布式共识因此把“开始”从个人脑内的确信、局部文本的稳定和单条生成路径，推进到跨位置的共同承诺。它与成核形成最直接冲突：成核允许少量局部组织先于宏观扩展而取得新相资格；共识则认为，局部状态在没有足够关系承诺时仍可回滚。它与动力学校对也冲突：一个候选即使在单一路径中通过许多步骤，也可能只是在一个节点内部自治，尚未让其他文本决策接受同一历史。

六、三家真正争的是什么：启动时点应该由谁来裁

把三条路径放到同一个学生身上，冲突就清楚了。学生看到“等待”这个题目，先想到小时候在医院门口等母亲，这是局部候选；他连续追问“我当时为什么害怕”“我记得什么声音”“母亲出来时说了什么”，这个候选通过了多轮内部检验；随后他决定全文都从“门”这个空间意象组织，第二段写急诊室门、第三段写多年后自己替孩子守门，结尾不再抽象赞美母爱，而回到“门开之前，人不知道自己在等什么”。三个理论会在三个不同时间点喊“开始”。

成核理论会在第一个候选开始持续长大时宣布启动：那个医院门口的记忆已经不是随机素材。动力学校对会等它经过几轮选择仍不解体再宣布启动：很多鲜明记忆只是诱人的假信号。共识理论则会继续等待，直到叙述顺序、意象、判断等多个局部决策共同接受它的约束。三家因此不是互补地贡献“稳

定、过程、关系”三个维度，而是在争一件非常具体的事：到底哪一个时点之后，我们才可以停止把学生视为“还需要别人告诉下一步”的人。

若只选其中一家，都会产生可预测的误判。只用成核标准，漂亮但孤立的开头会被当成真启动；只用校对标准，学生可能在一个局部反复打磨，却没有形成跨段依赖；只用共识标准，则可能把已经开始自我生长的早期作品误判为“还没有”，迫使教师继续介入，反而把新生的内生路径打断。真正的答案不能只是把三个条件并列成“稳定+校验+共识”，因为那仍然默认最后一个条件是终点、前两个是手段。How 题需要解释为什么终点会交换，以及系统何时不再依靠同一个外部驱动。

七、共有前提的坍塌：三家都默认“启动=达到一个固定终点”

三条路径表面上争得很凶，底下却共享一个不易察觉的前提：启动是一件可以由某一个固定终点单独结算的事。成核把稳定实体当终点，动力学校对把完成序列当终点，共识把共同承诺当终点；只要自己的终点达到，另外两样就退为条件或手段。这个前提正好符合 How 型最危险的结构：把一条发生路写成“走到某处就完了”。

推翻这条前提的材料并不需要从第四门学科搬来。成核理论自己已经告诉我们，经典的“一个临界核”并非所有体系唯一的反应坐标；两步成核、前驱体、团簇合并以及界面结构变化都会使“对象先稳定”与“路径先组织”互换次序。处在临界顶端的团簇也并非获得绝对不死身份，它仍有继续长大或回落的概率。这意味着成核领域自己已经承认：一个被称作终点的稳定对象，其发生资格依赖路径与环境，而且会反过来重写路径。

分布式系统给出另一面：所谓“提交”不是一个独立按钮，而是某段日志、复制路径、任期、法定人数和成员配置共同构成的结果。成员变更时甚至需要重叠多数来防止旧配置与新配置各自形成合法多数。也就是说，场的结算规则本身会被路径和对象变化重写。动力学校对同样不是单纯的“多过几关”，因为反馈、共刺激和信号网络可以改变每一关的有效阈值。三家的内部事实共同取消了“固定终点”的安全感。

于是问题转向：如果启动不是到达 S、D 或 E 中的某一个，那么三者何时组成了一个不可还原的新状态？本文的答案是：当三者第一次形成相互回写，使下一步生成的主要驱动力不再需要每次从外部重新注入。换言之，启动不是某个部件完成，而是续写权发生了移交。

八、涌现一：续写权移交——“下一步由谁决定”比“写了多少”更接近启动

“续写权”指的是对下一步写作选择拥有主要生成与排除作用的约束来源。作文尚未启动时，学生可以有题目、有素材、有开头，但每一步都需要新的外部指令：老师问一句，他答一句；范文给一段，他仿一段；AI 给一个建议，他继续一个建议。看起来文本在增长，真正决定“下一步是什么”的权力仍在系统外部。

作文启动以后，情况发生结构变化。已经形成的某个局部约束开始自己出题：如果我把“门”作为组织意象，那么下一段不能随便插入与门无关的旅游见闻；如果我把“等待中人会改写自己对时间的理解”作为判断，那么素材选择必须能制造这种改变；如果第一段采用孩子视角，后面突然出现全知评论就会产生张力。作者当然仍可修改甚至推翻这些选择，但新选择必须回应已有文本。正在形成的作文第一次从“被写的对象”变成“参与决定怎样继续写的对象”。

因此，本文提出一个比“写下第一字”更严格也更宽松的启动判据。它更严格，因为一百字模板可能仍没有取得续写权；它更宽松，因为一个尚未落笔的内部关系也可能已经取得续写权——作者虽然在沉默，但他已经知道下一步有哪些可能、哪些不可能，后续选择开始由这个关系网产生。由此得到一个反常预测：作文启动可以晚于第一字，也可以早于第一字。

“续写权移交”也把教师支持重新定位。外部支架在启动前承担的是点火功能；启动后如果仍逐句提供下一步，就会继续占有续写权，使学生无法检验自己的作文是否已经能够自持。真正高质量的支架不是永远给更多提示，而是在一个内生约束形成后逐步撤掉外部驱动，看系统能否自行产生下一步。

九、涌现二：自持起写核——作文从无到有的最小发生单位

为了避免“续写权”停留在漂亮比喻上，本文把取得续写权的最小约束簇命名为“自持起写核”。它不是一句话，不等于中心思想，也不是提前写好的提纲。它可以由一句话、一个意象、一组关系、一个矛盾、一个叙述视角或一条尚未语言化的判断构成。它的身份不靠表面形态，而靠下游依赖：拿掉它，后面的若干独立选择会一起改变。

一个自持起写核至少包含三种功能。第一是对象功能：它有足够稳定的边界，让后续选择知道自己在回应什么；这来自成核。第二是路径功能：它在

连续选择中经受“留／弃、前／后、详／略、说／不说”等真实分叉，而不是只在脑内被赞美；这来自动力学校对。第三是场功能：它让至少两个原本可独立做出的决策开始相互依赖，例如内容选择与叙述顺序同时受它牵引；这来自共识。三者不是三个静态指标，而是一个最小闭环：核约束路径，路径招募关系场，关系场再回写并修正核。

这一定义有一个重要后果：自持不等于固定。真正的核可以改变文字形态，甚至删除最初那句话，只要组织角色被后继结构接管。学生第一句“我一直不喜欢等待”后来被删掉，全文改成三个“门”的场景；如果“等待改变人对时间的关系”仍持续组织材料，那么核没有死亡，只是换了载体。反过来，一句开头从头到尾一字未改，如果后面的选择与它无关，它只是持久痕迹，不是起写核。

因此，作文从“没有”到“有”的最小转折，不是纸面上从0字到1字，而是系统中第一次出现一个能同时被后续选择使用、检验并回写的约束闭环。只要这个闭环尚未出现，文本再长也可能只是外力推动的堆积；一旦它出现，作文哪怕只有很少文字，也已经具有继续生成的内生条件。

十、二维辨别格：最危险的不是没开头，而是“有开头、没启动”

为了让新概念能够区分现实案例，本文用两条彼此独立的轴建立二维辨别格。横轴是“表面痕迹稳定性”：早期出现的文字或提纲是否在后续版本中持续保留；纵轴是“下游依赖性”：后续独立语义选择是否因该早期约束而改变。传统课堂最容易看见横轴，却几乎不测纵轴。

四格中，低稳定+低依赖是“漂浮候选”：学生不断写了又删，没有任何候选真正组织后续。高稳定+高依赖是“显性启动”：早期表达既保留，又不断牵引后文。低稳定+高依赖是“隐性启动”：文字表面反复改写，但同一组织关系通过不同表达持续约束后文，这一格解释了为什么一些看似“老在改开头”的成熟写作者其实已经进入深层生成。最值得警惕的是高稳定+低依赖——“伪启动”：开头、中心句甚至提纲都稳定存在，却没有任何后续选择真正依赖它。

“伪启动”恰好是课堂形式检查最容易放行的格子。学生写了四段，每段都有主题句；老师问中心思想，学生也能复述；但把第一段或所谓中心句换成另一套套话，后文几乎不需要改。这样的文章可以很完整，却仍靠题目惯例、段落模板和教师提示逐步推进。本文不是说它“不算作文成品”，而是说

它没有提供“学生的作文系统已完成内生启动”的证据。这一划界对 AI 时代尤其重要：外部模型可以极低成本维持文本增长，使“字数增加”与“作者系统启动”彻底脱钩。

十一、可清点判据：续写依赖率 CDR6 与删核分支实验

理论若不能被错，就只是解释性修辞。本文因此提出一个研究版读数“续写依赖率”（Continuation Dependency Rate, CDR6）。它不用于给学生打分，而用于实验性判断某个早期候选 K 是否已经取得下游组织作用。定义如下：固定候选 K 出现之后的六个独立语义决策作为观察窗，在不增加新外部提示的条件下，建立原分支与“删除或替换 K”的反事实分支；若某个后续决策在两分支中发生类别性改变，则记 1，否则记 0。CDR6 = 改变的决策数 ÷ 6。

“独立语义决策”不等于六次按键。本文将其编码为四类：内容决策（写哪件事/哪一证据）、关系决策（前后、因果、转折、详略）、立场决策（对对象采取何种判断或距离）、表达决策（叙述视角、核心意象、关键措辞的选择）。连续修改同一个词只算一次决策；纯错别字修正不进入分母。研究版 v1.0 把“起写核候选”设为 $CDR6 \geq 2/6$ ，且两次改变必须跨至少两类决策，同时至少出现一次回写：后续选择迫使 K 本身被重写、缩窄或重新确认。这个阈值是可被实验推翻的操作约定，不是本体常数。

为什么选择六个决策而不是六十个？因为启动是早期边界，窗口过长会把后期结构稳定混入判定；为什么最低两次而不是一次？一次改变可能只是句法牵连或局部搭配；跨两类变化才开始显示“关系场招募”。更重要的是，CDR6 刻意不奖励“开头一字未改”。如果 K 被改写但依赖关系持续，仍可以得到高读数；如果 K 原封不动但下游选择不受影响，读数仍低。

课堂上不需要完整做两份作文。教师可以用微型删核实验：当学生说“我有了”时，不问“中心是什么”，而问“把你现在最想保留的这一点拿掉，接下来三件你本来要写的东西会变哪几件？”若学生只能说“那我就换个开头，后面照写”，说明这个开头尚未取得续写权；若学生发现素材、顺序甚至立场都要一起变化，说明一个真正的依赖核已经出现。

十二、故意拓宽：这个想法是不是早就有人说过

“作文自己逼出下一步”很容易让人产生新鲜感，但写作研究史已经非常厚，任何关于开始、发现、计划和生成的主张都必须先假定自己多半是旧东

西。本文按经典层、作文教育内部、方法学、跨学科专著和同产线五个方向进行敌意检索。结果不是“没有近邻”，而是存在一整圈危险占位者；新概念只有在与最危险者做 1:1 替换后仍有不可替代部分，才有继续存在的理由。

方向	最危险占位者	1:1 替换后的判决
层	Wallas, 1926: 准备—孵化—顿悟—验证	已把“突然开始”放入创造过程，但不拂作文内部的下游依赖判据
实务/经典	Rohman, 1965: Pre-Writing	正面占位“第一字之前也属于写作”；不替代反事实依赖边界
实务/经典	Emig, 1971: 真实写作过程/有声构思	占位“过程而非成品”，但不裁定哪一刻统取得续写权
实务/经典	Murray, 1972: Process not Product	占位递归过程观，不提供启动事件的可选边界
人知	Flower & Hayes, 1981: Planning/Translating/Reviewing	最危险近邻之一；规划可外源给定，CD求内生反事实依赖
人知	Bereiter & Scardamalia, 1987: Knowledge Transforming	占位高水平写作中的问题空间互动，但不于最早启动闭环
人知	Galbraith, 1999: Knowledge-Constituting	占位“写着写着产生思想”；本文只保留“最早取得续写权的边界”
学	Leijten & Van Waes, 2013: 键击日志/burst/停顿	可测输出节律，不等于早期选择对后续连的因果依赖
学	Lo Sardo et al., 2023: writing cloud/exploration-flow	可测探索—转写切换；没有删核反事实，1:1 替代 CDR6
料专著	Maturana & Varela, 1980: Autopoiesis	“自持”上位概念；缺作文启动的最小决单位和删核判据

敌意拓宽的结论不是“无人说过类似话”，恰恰相反：写作研究已经把发现、规划、递归、知识生成和过程测量占得非常满。本文能够留下的只是一条更窄的缝——把“启动”定义为某个早期约束取得可反事实检验的下游组织权，并以外部驱动能否撤除作为功能后果。任何后续研究若证明这条缝可以被 planning、knowledge transforming 或现有 burst 指标 1:1 替换，新概念应当撤销。

十三、与最危险近邻的 1:1 替换：为什么不是“预写”“构思”“发现”或“流”

最危险的正面占位者是 Rohman (1965) 的 pre-writing。它已经明确把落笔前的发现视为写作过程的一部分，因此本文绝不能靠“启动早于第一字”取得创新。如果把“自持起写核”整体替换为“预写”，论证却会在关键处崩掉：预写描述一个阶段或活动集合，列提纲、自由联想、画思维图都可以属于预写；本文要求的是一个反事实依赖结构。一个学生可以做了二十分钟预写，却没有任何早期选择能改变后六个语义决策；也可以完全不做显式预写，却在第一段中形成高依赖核。因此两者不是强弱关系，而是分类轴不同。

Flower 与 Hayes (1981) 的“规划”同样危险。他们已经指出写作由层级目标引导，规划可以发生在任何时点。若本文只是说“有了目标后文字会受约束”，那就是规划理论换名。分离线在于：目标可以来自教师、任务说明或模板，并长期稳定存在，却仍不产生内生续写权。CDR6 测的是某个正在形成的作文内部选择对后续选择的反事实作用，而不是作者是否拥有一个显式目标。若外部给定目标能完整替代 K 而后续不变，本文判该核尚未成立。

Galbraith (1999) 与 Baaijen 和 Galbraith (2018) 的“写作即发现/知识构成”更接近本文。他们强调思想可以在文字生产中形成，而非先有完整思想再翻译成文字。这已经占掉“文本会反过来改变思想”这一大片区域。本文保留下来的狭窄位置是启动边界：知识构成可以发生在写作已经启动很久以后，也可以偶发而不形成持续依赖；自持起写核要求一个可定位的最早闭环，并通过删核分支决定后续多类选择是否一起改变。

方法学上，键击“burst”、停顿分析、写作云中的 exploration/translation flow 也都可能被误认为启动判据。它们记录的是输出节律或版本变化密度。流畅连续输入可以来自抄写、模板填充或 AI 建议后的转录，未必说明续写权已经内生；反过来，一个真正的起写核可能伴随长停顿和大量回改。若 CDR6 在控制首键时延、平均键击速率、burst 长度与规划时间后不再解释任何独特差异，本文就应该被这些成熟方法吸收。

跨学科上，“自持”还会撞上自创生 (autopoiesis) 与自催化集合。它们都描述系统如何维持或产生自身，但对象层级过宽：一个自持生命系统或反应网络并不提供“删掉哪一个早期写作决定，后六个语义选择会怎样改变”的操作判据。本文若无法在反事实依赖图上给出比这些上位概念更窄的预测，也没有资格保留新名。

十四、真跑一条：三个公开写作过程数据集为何仍看不见“启动”

按照 3.9 版的证伪纪律，本文不能只写“以后可以研究”。成文前先对三个公开或公开描述充分的写作过程数据资源做字段审计。预注册问题不是“数据能不能证明本文”，而是更低成本的一问：现有记录是否已经包含直接计算自持起写核所需的四类字段？四类字段在审计前写死为：A，时间化版本或键击；B，第一字之前的内部选择记录；C，早期选择到后续语义选择的依赖边；D，对早期候选做删除/替换的受控反事实分支。只有四项同时存在，才允许声称该数据集能够直接裁决本文核心机制。

KLICKe 记录约五千篇议论文的逐键、鼠标操作、时间戳、光标位置并配有整体质量评分（Tian, Crossley, & Van Waes, 2025），A 项充分；但其公开说明没有 B、C、D。Lo Sardo 等人的写作云保存版本演化，并能量化 exploration 与 translation flow，A 项充分；其设计也没有在每个早期选择后建立语义依赖边或反事实删核分支。ScholaWrite 更进一步，对五篇完整学术论文约六万余次文本变化附加认知意图标注，但公开描述仍没有“把某个早期约束移除后再从同一点继续写”的分支实验。结果为 0/3 数据资源同时具备四字段。

这是一个对本文不利但重要的结果。它意味着我们现在不能拿公开数据算出“真实作文平均在第几分钟启动”，也不能声称 CDR6 与作文成绩之间已有显著关系。此前若直接用停顿、首键时延或保留率做替代，会把代理变量偷偷当成新对象，违反单一口径。本文因此主动缩小主张：当前证据只支持“这个边界可以被操作化，但现有主流过程日志缺少直接裁决字段”。第三层经验主张由此变成一个数据工程要求：未来写作过程语料若要研究启动，至少要增加语义决策标注、依赖边和小规模反事实分支。

这个失败反过来修正了模型。原稿曾设想仅从“早期文本后来是否保留”识别起写核；字段审计后这一条被删除，因为 KLICKe 一类日志完全可以计算保留率，但保留率无法区分“模板开头一直没删”与“早期约束真正组织后文”。因此，反事实下游依赖而非表面保留被提升为主读数。

公开资源	A 时间化过程	B 前文字段	C 语义依赖边	D 反事实分支	结论
Ke（约 5000 篇文）	是：逐键/鼠标/时间戳/光标	否	否	否	不能直接计算
Writing	是：版本演化/编辑距	否	否	否	不能直接计算

公开资源	A 时间化过程	B 前文字段	C 语义依赖边	D 反事实分支	结论
公开资源 (61 名科研写	离				
Write (5 篇完 文, 约 6.2 万次文 本)	是: 文本变化+认知意 图标注	未覆盖完整前文字段	否	否	不能直接计算

十五、六条可证伪预测：怎样让“启动”真的可能被判错

第一条是时点预测。控制题目知识、打字速度、总规划时间和首键时延后，若 CDR6 首次达到研究版阈值的时点与随后三分钟的自主语义决策产出之间没有额外关联，而首键时点已经解释全部差异，则“续写权移交”不优于传统可见起点，核心模型应被削弱。

第二条是删核预测。在同一写作者、同一题目下，于候选 K 出现后随机进入“保留 K”与“替换 K”两个微分支。如果高 CDR6 候选被替换后，后续内容、关系、立场和表达决策并不比低 CDR6 候选发生更多协同变化，那么 CDR6 没有测到组织依赖，本文判据失败。

第三条是支架撤除预测。教师或 AI 在启动前提供外部提示会显著提高继续写的概率；当自持起写核形成后，撤除同强度提示造成的短时产出下降应明显减小。如果提示撤除的影响与核形成前后无交互，所谓“续写权移交”只是一般写作熟练度的改名。

第四条是伪启动预测。给学生一段固定高质量开头或结构模板，应提高表面痕迹稳定性，却不必提高 CDR6；若模板开头稳定度与 CDR6 几乎总是同升同降，且不存在“高稳定—低依赖”案例，二维辨别格的靶格为空，本文最重要的划界将失去经验意义。

第五条是早于第一字预测。允许学生在不落笔时用口头或卡片做内部关系选择；若在首键之前已经出现可反事实检验的依赖结构，那么后续落笔应表现出更高连续性。若所有可识别的依赖都只在第一字之后出现，则本文“启动可以早于第一字”的强命题被否定，但“可以晚于第一字”仍可独立保留。

第六条是竞争解释预测。最强竞争解释是“其实就是规划质量”。控制独立评分的计划完整度、目标明确度和题目知识后，如果 CDR6 对支架撤除后的自主续写没有增量预测，或二者高度共线到无法构成判决性反例，那么应优先使用成熟的规划概念，不再保留“自持起写核”这一名称。

十六、课堂实践：教师怎样帮助作文“点着”，而不是替学生一直烧

如果本文判断成立，作文启动教学首先要改变提问。传统提问常是“你准备写什么”“中心是什么”“第一段怎么开”，这些问题容易把一个声明当成结构。新的提问是“你现在这个选择，逼得后面哪两件事必须跟着变？”这句话直接测试下游依赖。学生若说不出，不代表他没有能力，只表示候选仍处在涨落区；教师此时应继续扩大可选差异，而不是马上替他定中心。

第二，教学要允许“亚临界时间”。成核理论给本文的第一条反向约束是：并非越早锁定越好。过早把第一个想法封成中心，会把探索阶段的正常涨落当成错误。课堂可以给学生三到五分钟生成三个不同核候选，每个只写一句“如果选它，接下来必须出现什么、不能出现什么”。目的不是选择最漂亮的一句，而是找哪一个候选真正具有生成后续差异的能力。

第三，校验必须嵌在续写中，而非增加审题清单。动力学校对提醒我们，辨别需要时间，但步骤越多也越可能错杀弱信号。教师可以只设置两关：第一关问“它能不能排除至少一种看似也能写的选择”；第二关问“它能不能招来一个原本没想到的选择”。既能排除又能招募，说明候选开始成为约束；若只能复述题意，仍在外部任务层。

第四，要做“小撤架”。当学生似乎已经启动，教师停止给新的内容提示两三分钟，只保留安全和时间边界。观察学生是否能够从已有文本自行提出下一步问题、做选择和回改。若一撤架就停，不应把它解释为懒惰；更可能是续写权仍掌握在教师手中。支架需要退回到“帮助形成核”，而不是继续提供下一句。

第五，把修改从“启动失败”改写为“核的回写”。很多学生害怕删开头，是因为课堂把稳定痕迹当作已经开始的证明。本文恰好预测，真正的核常在后续关系形成后被重写。教师可以保存版本，问“你虽然删了第一句，什么东西没有丢？”让学生学习区分文字载体与组织角色。这样的修订不是推倒重来，而是闭环第一次完成。

第六，评价记录应增加过程字段，而不是再加一个主观分。最低成本的记录不是要求教师给“启动能力”打1—5分，而是每篇作文只留两个时点：学生第一次说“我知道接下来怎么写了”的时点，以及教师第一次撤除内容提示后仍连续做出三个语义选择的时点。它们不是理论的最终测量，但可以让课堂开始看见此前完全没有被记录的边界。

十七、AI 时代的特殊风险：成品已经“有了”，学生的作文却可能从未启动

生成式 AI 使这个问题从理论细节变成教育治理问题。过去学生若没有启动，纸面通常也长不起来；今天外部系统可以持续提供题意拆解、素材、提纲、开头、过渡、例证和结尾，直到一篇完整文本出现。产品层的“有”与学习者系统层的“有”第一次可以大规模分离。若评价只看最终作文，我们无法知道学生是在借 AI 完成一个已经由自己启动的写作过程，还是从第一步到最后一步都没有取得续写权。

本文并不因此推出“AI 越少越好”。在启动前，AI 完全可以作为候选制造器，帮助学生扩大差异；在校验阶段，它可以扮演反例提供者，逼学生发现某个候选是否只是套话；在共识阶段，它甚至可以提出“如果删掉这一点，后文哪些地方必须改”的反事实问题。真正的边界是：AI 的帮助有没有在某个时点退出“下一步内容的供应者”角色，把选择权交给正在形成的作文—作者系统。

这也解释一种常见悖论：AI 给出的建议越顺，学生越可能“一路写完却从未开始”。外部模型持续占有续写权时，文本增长速度会升高，停顿会减少，表面过程反而更像“启动成功”。因此，若本文成立，未来 AI 写作教育最重要的过程指标之一不是“AI 用了多少”，而是“在哪个时点之后，即使 AI 停止提供内容，学生仍能由已有约束继续产生和筛选下一步”。

十八、反向约束与适用边界：这个模型不能被拿去强迫所有人“快速点火”

第一条反向约束来自成核：亚临界涨落不是浪费，而是新相选择的必要空间。本文因此撤回“越早形成起写核越好”的潜在强命题。对探索性、文学性和复杂议论文，过早追求高 CDR6 可能制造路径锁定，使学生围绕第一个可行核过度开发，牺牲后来更好的方向。教育目标应是“能形成并能重新成核”，而不是一次点火后永不改变。

第二条反向约束来自动力校对：更严格的校验提高错误排除，却增加时间与认知成本。对于低龄学生、二语写作者、书写自动化不足者或限时考试，要求显式做反事实分支可能本身占用工作记忆。CDR6 首先是研究工具，不得直接变成评分 KPI。课堂简化版只能用一两个问题粗测，不能把每个选择都编码成表。

第三条约束来自分布式系统：共识追求安全，但创新并不总需要多数。一个新颖而脆弱的写作方向可能暂时只在一个局部节点存在，如果课堂要求“所有材料都马上围绕中心”，反而会杀死尚未找到关系的异质片段。因此“招募多个决策”不是要求全文整齐划一，而是要求至少出现可追踪依赖；保留少数未承诺区域，可能正是后续第二次成核的来源。

本文还不适用于纯抄写、听写、机械转录等任务，因为这些任务的续写权本来就合法地位于外部源；也不直接等同于合作写作中的群体共识，后者还涉及人际权力和版本治理。对于意识流、自动写作等刻意降低显式控制的体裁，起写核可能表现为节奏、声音或局部句法约束，而非主题性内容，编码规则需要调整。

十九、自反检验：本文自己有没有形成“自持起写核”

若本文的概念只能解释学生而不能解释自身，它就缺乏最基本的自反性。本文最早的候选其实是“作文从第一句开始”。成核理论进入后，这个候选改成“从第一处能持续长大的局部结构开始”；动力学校对它进一步加入“必须在连续选择中存活”；分布式共识又迫使我们承认“局部存活仍不够，必须让多个后续决策接受同一约束”。如果删掉“续写权移交”这一约束，后面的二维辨别格、CDR6、AI分析与课堂撤架方案都必须一起重写，因此它确实对多个下游决定形成依赖。

但这并不证明本文理论为真，只证明本文本身符合它提出的形式。更危险的是概念一旦进入教学制度，最省事的实施方式会变成“每篇作文必须在五分钟内找到起写核，CDR6 低于某值扣分”。这恰好会摧毁模型要保护的亚临界探索。对此本文没有一个可以彻底封死的制度出口：任何可量化判据都可能被 KPI 化。能做的只有明确伦理限制——CDR6 只用于研究和诊断位置，不用于排名学生；课堂记录优先服务教师撤架时点，不服务高风险评价；任何以提高 CDR6 为目的的强制模板都应被视为对构念的污染。

二十、六条完整启动路径：同一个“开始”，可以从不同入口发生

3.9 版 How 型最重要的修正之一，是不能把方法写成唯一的线性次序。S、D、E 三种焦点之间理论上存在六条完整路径；同一个作文启动问题，学生可能从不同入口进入，最后形成同一种“续写权移交”。如果本文只保留最顺手的 $S \rightarrow D \rightarrow E$ ，就会把成核学科的偏好误当成作文的普遍规律。下

面六条路径不是六种写作风格，而是六种“谁先取得临时发起权、谁随后接管、最后哪一维把上一轮结果结算并回写”的发生次序。

第一条， $S \rightarrow D \rightarrow E$ ，可称“细节起核路径”。学生先被一个具体对象抓住：一只被放回水里的小鱼、一扇迟迟不开的门、一只磨破的鞋。这个对象先获得稳定边界，随后作者围绕它连续试写、筛掉不相干材料，最后多个段落、判断与表达共同接受它的约束。很多记叙文、散文和经验性作文从这里启动。教学上最有效的动作不是马上问中心，而是追问：“如果保留这个细节，接下来哪两件事必须跟着改变？”

第二条， $D \rightarrow S \rightarrow E$ ，可称“试写凝核路径”。作者一开始没有稳定主题或材料，只通过自由写、列句子、变换视角、连续问答等动作推进。路径先发生，某个反复出现的句法、冲突或意象在操作中逐渐凝成对象，随后再招募全文关系。它解释了为什么有些学生必须“先写废话”才会真正开始。对这类学生，强迫先列完整提纲会切断本来有效的启动路径。教师更应允许一段低风险试写，然后回看哪些东西在多次变化中仍然回来。

第三条， $S \rightarrow E \rightarrow D$ ，可称“对象成场路径”。一个局部对象刚出现，就迅速把多个未排序的材料吸到同一关系场中，路径反而最后才被整理出来。例如学生想到“爷爷每天把药盒放在门口”这一细节，立刻联到忘记吃药、出门上班、家人电话提醒、老年人的自主与被照料等多个关系；此时还没有段落次序，但“什么属于这篇文章”已经发生变化。随后作者才选择叙述路线。这类启动常被课堂误判为“思路乱”，因为E先于D。

第四条， $E \rightarrow S \rightarrow D$ ，可称“关系压核路径”。作者先感到一种关系张力，却暂时找不到承载它的对象。比如“我想写的是照顾一个人时，爱和控制很难分开”，这个关系场先存在；随后它压缩到一个可写对象——父亲偷偷把降压药藏进口袋、孩子把手机从老人手里拿走——最后围绕该对象生成具体路径。议论文、关系主题和复杂生活写作经常如此。若教师要求先找“典型事例”再谈关系，可能把真正的启动次序倒置。

第五条， $D \rightarrow E \rightarrow S$ ，可称“路径聚场路径”。作者先做一连串探索动作：比较两个经历、写两种结尾、问三次“为什么”，这些动作逐渐形成一个稳定关系场，最后才落成一个可以命名的中心对象或判断。科研写作、说明性写作以及“写着写着才知道自己在说什么”的作文常见此型。它与知识构成理论相邻，却多一步启动判据：只有当后续选择开始对这一关系场产生依赖，最终对象才不是事后总结。

第六条， $E \rightarrow D \rightarrow S$ ，可称“约束寻核路径”。任务或现实先给出强关系约束，例如“必须同时回应两个彼此冲突的观点”“必须给不同读者都能接受的解释”。作者从这个场出发，尝试多条组织路线，最后找到一个最小对象作为核。考试中的材料作文、辩论式写作、公共写作常从这里发生。此时题目并非简单外部刺激；关键在于作者是否把外部关系约束转译成自己的内部选择规则。只要仍逐条等待教师解释， E 尚未真正进入作者—文本系统。

六条路径共同说明，作文启动没有唯一“正确顺序”。真正不变的是闭环条件：无论从哪一维出发，某一局部稳定、某一生成序列和某一关系场最终都必须发生相互回写，使下一步不再只能由外部继续供应。教学因此不能把“先审题—再立意—再选材—再列提纲—再写”的单一路线升级为发生学定律；它只能是一种对部分学生有效的外部脚手架。高水平启动教学的目标，是帮助教师识别学生此刻在哪一条路径上，并把缺失的下一维补进来。

二十一、四个典型课堂案例：同样“卡住”，其实卡在不同的启动位置

案例一是“有开头，无续写”。七年级学生面对“那一次，我长大了”，很快写出：“成长像一场春雨，悄无声息地滋润着我们的心田。”这句话保持十分钟没有删除，看起来已经进入作文。但问他下一段写什么，他说“老师你给我一个例子吧”。教师给出“照顾生病的妈妈”，他就写妈妈；教师改说“第一次独自坐车”，他也能改写。无论替换哪一个开头，后文都不需要协同变化。按传统表面指标，这是快速启动；按本文二维格，它属于高痕迹稳定、低下游依赖的“伪启动”。真正需要的不是再教三个开头，而是让学生找到一个能够排除其他材料的具体关系，例如“长大不是第一次独立做事，而是第一次发现自己的选择会让别人轻松一点”。一旦这个约束形成，素材与结尾会同时变化。

案例二是“没写字，已经启动”。一名学生拿到“声音”题目后五分钟没有落笔，只写了三个词：“厨房—凌晨—塑料袋”。教师以为他在发呆。实际上他已经决定写母亲夜班回来后在厨房极轻地收拾东西，文章不能直接赞美母爱，只通过“尽量不发出声音”的动作让读者看到两个人彼此怕吵醒对方；因此他排除了学校广播、雨声、音乐等其他材料，也决定全文采用第一人称现场听觉。虽然纸面正文仍为零，这三个关系选择已共同约束后续。若此时教师因为焦虑给他一篇“难忘的声音”范文，反而可能抢回续写权，破坏已经形成的隐性起写核。

案例三是“写了很多，核一直换”。学生围绕“选择”连续写了三百字：先写选文理科，又换成选朋友，再转到买鞋，字数不断增加但各段互不依赖。教师若只看产量，会认为学生进入流畅写作；实际上系统仍在亚临界探索。这里最好的干预不是要求删掉前文，而是把三段同时摆开，问：“哪一个矛盾在三个选择里都出现？”如果学生发现自己真正关注的是“别人替我选时我反而轻松，但后来会怨别人”，这个关系可能在已有材料之间形成 E，然后通过 $D \rightarrow E \rightarrow S$ 或 $E \rightarrow S \rightarrow D$ 路径凝成新核。前面的三百字不是失败，它们是为成核提供涨落的材料池。

案例四是“AI 让文本一直有，学生的系统一直没有”。学生输入题目，AI 给出三点提纲；学生选第二点，AI 再给开头；学生把开头贴入文档，再问“第二段怎么接”；AI 给例子后，他又问“怎么升华”。最终成文语言顺畅。若突然关闭 AI，学生无法说明下一步还能改哪里，也无法回答“删掉第二段后其他哪两处必须变化”。这里的问题不是使用 AI 本身，而是外部系统持续占有每一个微小决策的续写权。相反，另一名学生可能让 AI 给出十个素材，自己否决九个，选中一个后重新安排三个场景，并拒绝 AI 建议的结尾，因为它破坏了自己的叙述视角。后者的 AI 使用量可能更多，却已经发生续写权内生化的。

四个案例说明，教师面对“卡住”时需要先诊断位置。若是候选太少，增加差异；若是候选很多但没有稳定对象，帮助比较；若是对象稳定但没有路径，让它通过两个真实选择关；若是局部路径顺畅却没有关系招募，要求做删核对照；若核已经形成，则最重要的动作可能是停止帮助。把所有“卡住”都解释为没素材、没方法或没动力，只会让外部提示越来越密，最终把原本应该转移给作文自身的续写权长期留在教师手里。

二十二、一个可执行的实验方案：怎样真正测到“作文启动时刻”

未来验证本文最直接的研究，不需要昂贵脑成像，而可以采用“版本记录+微分支”的课堂实验。样本可从同一年级、同一题型的学生中招募，避免把年龄、题目知识和体裁混成启动效应。每名学生使用能够保存版本与操作时间的写作界面；系统只记录文本变化，不自动给内容建议。写作开始前记录题目知识、打字速度和自我效能，作为竞争解释的控制变量。

第一阶段让学生自然写作。每当系统检测到一个相对稳定的语义单位持续一段时间，或学生主动标记“这点我想保留”，就把它列为候选 K，但不马

上判定。研究界面随后在不打断主文的情况下复制一个临时分支：A 分支保留 K，B 分支将 K 删除或替换为功能中性的占位提示。学生在两个分支各做六个后续语义决策，每个分支时间短、顺序随机，以降低疲劳和记忆效应。之后回到主文继续，不要求学生真的保留实验分支。

第二阶段由双人盲码把后续变化编码成内容、关系、立场、表达四类，并判断两分支同一序号决策是否发生类别性改变。分歧由第三名编码者裁决，报告一致性。由此得到每个候选的 CDR6 与回写事件。每名学生可能有多个候选，最早达到研究版阈值的时点暂记为 t^* 。关键是 t^* 不由最终成绩定义，也不由研究者事后挑“最像中心”的句子；它完全由预注册的分支规则产生。

第三阶段测试功能后果。在 t^* 前后各安排一次“撤除内容支架”窗口：如果实验原先允许教师只回答“下一步可以想哪些方向”，此时停止提供任何内容候选三分钟，仅保留时间与操作帮助。记录六个语义决策能否连续发生、外部求助次数、停顿结构以及后续是否大规模回退。本文预测， t^* 之后撤架的破坏应显著小于 t^* 之前。若没有这种差异，续写权所谓“移交”就缺少功能意义。

第四阶段比较竞争模型。以首键时延、显式规划时长、计划质量、前 100 字保留率、键击 burst 长度、平均停顿和最终作文分数为基线，检验 CDR6 是否提供增量解释。这里特别要防止把“会写的人什么都好”当机制：最好使用同一写作者内的多个题目，或在训练前后做重复测量，把稳定的个体写作能力尽量消掉。若 CDR6 只是在高水平写作者身上更高，却不能在同一人不同作文中定位启动差异，则它更像能力指标而非启动指标。

第五阶段必须保留失败样本。若某些学生整篇都没有候选达到阈值，不能为了统计方便把最后一个候选强行当起点；应记为“观察窗内未检出”。若有学生首字前已经通过口头卡片形成强依赖，则 t^* 可以为负值（相对首键时点）；若学生在最后一段才第一次形成闭环，也要允许 t^* 很晚。只有如此，我们才能真正检验“启动早于或晚于第一字”的强预测，而不是预先把启动定义在第一字之后。

二十三、从启动到持续：为什么“已经点着”仍然可能熄火

本文研究启动，而不是整篇作文如何完成，因此必须给概念划一道时间边界。自持起写核形成只说明系统获得了短程内生续写能力，不保证它一定完成，更不保证成品质量高。成核后的晶体仍可能遇到资源不足或竞争相；免疫

激活后仍可能被负反馈终止；分布式系统达成某次提交也不等于以后永不故障。作文同样可能在已经启动后因知识耗尽、情绪中断、时间压力、工作记忆超载或新证据冲击而熄火。

因此，“启动困难”与“维持困难”应当分开诊断。一个学生每次都需要老师帮他找第一步，但一旦找到就能一路写完，主要问题在启动；另一名学学生总能快速找到有力切口，却写到第二段就散掉，可能是核的下游资源不足、关系场过窄或维持机制有问题。若把两者都叫“不会写作文”，干预就会互相伤害。前者需要形成最小自持约束，后者可能需要扩展材料、重建结构或重新成核。

同样，修改不必被视为旧核失败。成熟写作可能经历多次“再成核”：第一轮核提供足够动力使作者进入文本，后续新证据使旧关系场失稳，系统重新寻找更高层约束。真正值得研究的是起写核能否把作者带到下一个可发生位置，而不是它是否从头到尾保持原样。这一点也限制了 CDR6 的用途：它只能判早期短窗中的组织依赖，不能替代整篇文章的结构质量、创造性或思想深度评价。

从课程设计看，这意味着“会启动”可以成为独立于“会完成”的微能力。教师可以设计五分钟短任务，只训练从题目到自持核，而不要求交完整作文；也可以设计“二次点火”任务，故意在中途加入一个与原计划冲突的新材料，让学生观察旧核如何解体、新核如何形成。这样，学生学到的不是一种固定开头术，而是对发生过程的可迁移感觉：什么时候我还在等外界推我，什么时候已有的文字开始推我，什么时候旧的推力已经失效，需要重新发生。

二十四、第二层共有前提：写作研究已经研究“过程”， 却仍缺少“过程何时取得自驱”的字段

第一层共有前提来自三门外部学科：它们都倾向把启动结算在某个焦点终点。按照 3.9 版方法，还需要检查第二层——作文研究自身的近邻理论是否共同站在另一块没有被争论的地基上。把 prewriting、认知过程模型、知识转化、知识构成、键击记录和写作云并排以后，可以看到一个微妙交集：它们对“写作由哪些活动组成、这些活动如何递归、作者何时停顿或回返”极其敏感，却大多默认只要活动已经发生，研究对象就是一个正在进行的写作过程。换句话说，它们记录活动的种类与时序，却很少把“这些活动的下一步究竟由外部任务继续驱动，还是由已经形成的文本内部约束驱动”单列成一个字段。

这一盲区不是任何理论的错误。Flower—Hayes 模型的任务是解释认知过程如何组织，不需要为“作文从零到一”另设法定出生证；键击记录的任务是高分辨率保存操作，不需要知道某个删除动作在语义上依赖哪一个早期选择；知识构成理论关心写作怎样改变认识，也不必回答第一次改变是否已经形成可持续闭环。正因为每一传统的问题结构都合理，“续写权”才会落在它们共同账本之外：不是没人观察写作早期，而是现有记录把“活动发生”与“活动获得内生驱动力”放在同一栏里。

这个第二层共有前提可以用一个极端例子检验。设想两名学生拥有逐键完全相同的前 100 字和相似停顿。A 学生每写一句都在等待教师下一条提示；B 学生已经形成一个内部约束，每次停顿是在比较两个与该约束冲突的选择。仅凭键击时间、字数和最终文本，二者可以高度相似；从学习发生看，二者却处在不同状态。若研究数据库没有“外部内容提示事件”“候选约束身份”“后续决策依赖边”这些字段，两种状态就会被强制合并。本文的贡献因此不只是给旧过程增加一个新名词，而是要求过程研究把驱动权的归属纳入数据结构。

第二层盲区也解释为什么作文教学长期偏爱可见的启动仪式：审题完成、提纲写好、第一段写出、规定时间内达到一定字数。这些指标便宜、可管理、可统一，却只记录“动作有没有出现”。在大班课堂里，教师很难追踪每个学生的反事实依赖，所以外部形式自然成为代理。代理一旦稳定，又会反过来塑造教学：为了让所有人尽快出现可见开头，教师提供统一句式；为了让提纲可检查，要求先填固定结构；学生因此更快产出痕迹，却更晚甚至从未取得续写权。这里形成一个值得进一步检验的制度性反转：越追求可见的快速启动，越可能延长真正启动对外部驱动的依赖。

这条制度性推论不能由本文目前的数据证明，因此只作为下一层待检验命题。可行研究可以比较两类课堂：一类以首段完成、提纲完整等可见指标催促启动；另一类允许较长候选期，并在形成依赖后撤除内容提示。若第二类学生的首键更慢，却在撤架后表现出更早的自主连续选择，就支持“可见启动速度与内生启动速度可分离”；若两者总是同向，本文关于制度代理的推论就应撤回。

二十五、结论：作文真正的第一步，是“已有的东西开始要求下一步”

本文从一个极小的问题出发：一篇作文究竟怎样启动。写作研究已经拥有丰富的过程模型，但“开始存在”的边界常被第一字、构思、计划、流畅输出或成品完成替代。物理化学成核把启动推到局部稳定核，免疫动力学校对拒绝过早承认候选，分布式共识又拒绝把单节点完成等同系统承诺。三者碰撞之后，固定终点的共同前提被取消，真正的变化落在“下一步由谁驱动”。

据此，本文提出：作文启动是续写权的第一次内生移交。一个最小约束簇一旦能够在连续真实选择中存活，招募至少两个不同类型的后续决策，并被这些决策反向修正，作文就形成“自持起写核”。从这一刻起，题目、教师、模板和 AI 仍可提供资源，却不再需要逐步供应下一步；正在形成的作文开始成为自己的条件。

这个判断把两种长期混在一起的情况分开：有文字而未启动，与无文字而已启动。它也把“写不出来”的干预从继续灌输素材，转向寻找一个能够产生差异、排除差异并形成下游依赖的最小核。公开数据字段审计同时提醒我们，目前还没有足够过程数据直接验证这条理论；这不是需要遮掩的缺点，而是下一步研究必须补上的测量对象。

所以，“第一句话怎么写”不是作文启动的最根本问题。更根本的问题是：学生写下或想出的这个东西，能不能开始反过来规定接下来什么可以写、什么不能写、什么必须被重写。第一字只是痕迹；真正的第一步，是已有的东西第一次开始要求下一步。

如果这一判断在后续实验中站得住，它还会改变语文教育学对“写作困难”的最小诊断单位。过去我们常在学生身上寻找缺陷：词汇少、素材少、思路窄、动力低、结构意识弱；本文则把其中一部分困难重新定位为发生链没有闭合。学生并非什么都没有，而是已有的素材、词句、判断之间尚未形成能够接管下一步的依赖结构。这样一来，帮助也不必总以“再给一点东西”为默认动作。教师可以减少替学生继续生产内容，转而帮助他看见哪个已有选择最有可能获得组织权，并在它刚刚取得续写权时及时退出。这种退出不是放任，而是检验教育是否真正把生成能力交回学生。更进一步，作文课也可以把“启动”单独拿出来训练：不要求四十分钟写完一篇，而只要求十分钟内经历候选、校验、招募与撤架，记录哪个选择第一次开始逼出下一步。学生一旦能够识别自己的启动方式，就不必把每次卡住都理解成能力不足，而能判断自己此刻缺的是对象、路径还是关系场。

附录 A CDR6 清点手册（研究版 v1.0）

- 1. 读数名：续写依赖率；符号：CDR6。
- 2. 定义：候选 K 出现后，取不增加新外部内容提示条件下的后六个“独立语义决策”。建立保留 K 与删除/等功能替换 K 的两个分支；在第 i 个决策上，若两分支发生类别性改变，记 di=1，否则 di=0。CDR6=(d1+...+d6)/6。
- 3. 粒度：一次独立语义决策指内容、关系、立场、表达四类之一发生可辨别选择。连续修改同一词不重复计数；拼写、标点、输入法纠错不进入分母；同一原因导致的一串机械改动合并为一次。
- 4. 研究版起写核候选阈值：CDR6≥2/6，且改变至少跨两类决策，并出现至少一次“回写事件”（后续选择导致 K 被重写、缩窄、重新确认或由等功能载体接管）。全文、证伪条款与赌注均使用这一口径。
- 5. 外部提示控制：观察窗内教师、同伴或 AI 不得新增具体内容候选；可以提醒时间、任务边界和安全规则。若发生内容提示，本窗作废，重新取窗。
- 6. 不适用：抄写、听写、逐句翻译、必须按固定字段填报的文本，以及无法合法进行分支实验的高风险评价情境。

附录 B 3.9 版选源体检摘要

来源	本题取用的核心判据	未吸收度	理论厚度	How 焦点	处置
化学·成核	局部涨落越过临界条件后，继续生长获得方向优势	干净	厚	S（稳定对象）	放行
学·动力学校对	候选须在时间中连续通过多步检验，过早承认会放大误识别	干净	厚	D（校验路径）	放行
式系统·共识提交	局部记录不等于系统承诺；满足复制/安全条件后才可视为共同历史	干净	厚	E（共享承诺场）	放行

检索词表（节选）：作文启动 成核／writing initiation nucleation／composition critical nucleus；作文 动力学校对／writing kinetic proofreading；作文 共识算法／composition consensus algorithm／writing Raft；形式层检索：first local choice becomes self-sustaining

writing、sequential validation before writing begins、local writing decision becomes shared commitment、when does a text start generating its own next step。判“干净”仅表示按上述口径未见进入写作理论固定术语、教材章节或标准手册，不宣称绝对历史空白。

附录 C 写死日期的经验赌注

截至 2028 年 12 月 31 日，如果出现至少两项独立、可复算的写作过程研究，在控制题目知识、规划时间、打字速度、首键时延和整体写作水平后，均发现按本附录 A 同一口径计算的 CDR6 对“撤除外部内容提示后的连续六次语义决策”没有增量预测，并且 planning/goal quality 能够 1:1 替代 CDR6 而不损失任何判决性预测，则本文关于“续写权移交是独立启动边界”的核心命题判伪。仅发现 CDR6 与最终作文分数不相关不算命中，因为本文没有把最终质量定义为启动；仅发现某些作家不用显式提纲也不算命中；改变 CDR6 定义、阈值或分母后得到的结果也不算对本赌注的直接检验。

参考文献

1. Becker, R., & Döring, W. (1935). Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Annalen der Physik*, 416(8), 719–752.
2. Vekilov, P. G. (2010). Nucleation. *Crystal Growth & Design*, 10(12), 5007–5019.
3. Oxtoby, D. W. (1992). Homogeneous nucleation: Theory and experiment. *Journal of Physics: Condensed Matter*, 4, 7627–7650.
4. Hopfield, J. J. (1974). Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences*, 71(10), 4135–4139.
5. Ninio, J. (1975). Kinetic amplification of enzyme discrimination. *Biochimie*, 57, 587–595.
6. McKeithan, T. W. (1995). Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences*, 92(11), 5042–5046.
7. Altan-Bonnet, G., & Germain, R. N. (2005). Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biology*, 3(11), e356.
8. Fischer, M. J., Lynch, N. A., & Paterson, M. S. (1985). Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2), 374–382.
9. Lamport, L. (1998). The part-time parliament. *ACM Transactions on Computer Systems*, 16(2), 133–169.

10. Castro, M., & Liskov, B. (1999). Practical Byzantine fault tolerance. *Proceedings of OSDI*, 173–186.
11. Ongaro, D., & Ousterhout, J. (2014). In search of an understandable consensus algorithm. *Proceedings of the 2014 USENIX Annual Technical Conference*, 305–319.
12. Tanenbaum, A. S., & Van Steen, M. (2023). *Distributed Systems* (4th ed.). Pearson.
13. Rohman, D. G. (1965). Pre-Writing: The stage of discovery in the writing process. *College Composition and Communication*, 16(2), 106–112.
14. Emig, J. (1971). *The Composing Processes of Twelfth Graders*. National Council of Teachers of English.
15. Murray, D. M. (1972). Teach writing as a process not product. *The Leaflet*, 71(3), 11–14.
16. Elbow, P. (1973). *Writing Without Teachers*. Oxford University Press.
17. Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
18. Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Lawrence Erlbaum Associates.
19. 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.
20. Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing*. Lawrence Erlbaum Associates.
21. Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing What to Write: Conceptual Processes in Text Production* (pp. 139–160). Amsterdam University Press.
22. Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3), 199–223.
23. Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392.
24. Lo Sardo, D. R., Gravino, P., Cuskley, C., & Loreto, V. (2023). Exploitation and exploration in text evolution: Quantifying planning and translation flows during writing. *arXiv:2302.03645*.
25. Tian, Y., Crossley, S. A., & Van Waes, L. (2025). The KLiCKe corpus: Keystroke logging in compositions for knowledge evaluation. *Journal of Writing Research*, 17(1), 299–336.

26. Wang, L., Lee, M., Volkov, R., Chau, L. T., & Kang, D. (2025). ScholaWrite: A dataset of end-to-end scholarly writing process. arXiv:2502.02904.

27. Wallas, G. (1926). The Art of Thought. Harcourt, Brace and Company.

28. Maturana, H. R., & Varela, F. J. (1980). Autopoiesis and Cognition: The Realization of the Living. D. Reidel.

29. 中华人民共和国教育部. (2022). 义务教育语文课程标准（2022 年版）. 北京师范大学出版社.

30. SDE Universes. (2026). 新思想前沿：物理化学.
<https://sdeuniverses.com/frontier/physical-chemistry/>

31. SDE Universes. (2026). 新思想前沿：免疫学.
<https://sdeuniverses.com/frontier/immunology/>

32. SDE Universes. (2026). 新思想前沿：分布式系统与区块链.
<https://sdeuniverses.com/frontier/distributed-systems/>

33. 付自文、Claude. (2026). SDE 多学科通融创新法：从 0 到 1 不是灵感，是一门可以重复的手艺. SDE Universes, 学科通融方法论特稿.

编者补节 · 第一章

一 · 本章是九章里唯一没有欠上游那笔账的

导论第三节已经说明，「已写出的文本会成为下一步的条件」是写作过程研究四十年的现成结论，本卷九章共同的上游是它。本章是唯一一章正面与它交手的：第十二节的十位占位者表里，把「写着写着产生思想」这一支单列一行，并写出了分离线——那一支占位的是写作产生新知识，本章只保留最早取得续写权的那一条边界。

编者把这一处标出来，不是表扬，是为了给其余各章立一个格式：祖宗要请进正文的划界表，不能放在脚注里。后面八节的补节全部按本章这一格式写。

二 · 需要补的两位

（一）自催化闭合

本章的核心动作是「一个最小约束簇第一次取得续写权，此后系统能够自己产生、筛除并协调下一步」。这个形状在生物化学与复杂系统研究里有一位很近的上游：自催化集合（Kauffman, S. A. 1986, *Autocatalytic sets of proteins*, Journal of Theoretical Biology, 119(1), 1-24），即一组分子在集合内部互为催化剂，使整个集合的再生产不再依赖外部特定输入。

本章正文里「自催化」一词出现过一次，无出处。这不合适——它恰好是本章最想说的这件事的现成名字。

分离线。自催化集合讲的是闭合（closure）：集合内每一个成员的生成都被集合内某个成员催化。本章讲的是移交（handover）：外部驱动仍然可以存在，只是下一步的主要生成与排除作用已经落到内部约束核上。二者的差别有一处可判决的后果——自催化集合一旦闭合，撤走外部催化剂系统仍能维持；而本章的续写依赖率允许外部提示继续存在，只要求下游选择的依赖关系指向内部核。

判定形式。在候选核已形成之后，完全撤走一切外部提示（教师、范文、机器补全），观察写作是否继续推进。若继续推进，本章测到的是闭合，应改用自催化的语言，本章的读数是它的一个弱化版本；若外部提示撤走后推进停止、而删核实验仍显示高依赖，则本章测到的确实是移交而非闭合，二者不可互换。

（二）写作作为约束满足

Sharples, M. (1999). **How We Write: Writing as Creative Design**. Routledge. 把写作处理成一个在约束之间往返的设计活动：写作者不断在「构思空间」与「外化」之间切换，而约束既来自任务，也来自已经写下的东西。

分离线。那一支讲约束从哪里来、如何被满足；本章讲约束什么时候开始自己生产新的约束。前者不要求区分外生约束与内生约束，本章的全部分量都在这一区分上。

判定形式。若把任务约束、教师约束与已写文本约束同等编码后，能够无剩余地预测删核实验的结果，则本章的「内生」一词没有额外解释力，应并回约束满足；若删核实验的结果只能由已写文本这一类约束解释，本章成立。

三·定名裁定：本章的「伪启动」为正名

「表面痕迹稳定而下游依赖为零」这一格，在本卷被两章各自命名过：本章称伪启动，第三章称惰性成形。

编者裁定：以本章的「伪启动」为正名，理由是它的定义最严——不仅要求表面稳定，还要求没有任何一个后续独立语义选择依赖它，并配了删核反事实作为判据。第三章处改为引用本章并说明视角差：本章看下游依赖，第三章看操作分布；二者是同一格的两种测法，不是两个概念。

读者若在第三章遇到「惰性成形」，应回到本节的定义。

四·编者提示：本章最强的一条与最弱的一条

最强的一条是那个反常预言：作文的启动可以晚于第一字，也可以早于第一字。它同时对抗两种常识——「动笔了就是开始了」和「想清楚才算开

始」。这一条也使本章可以处理一类现有研究几乎处理不了的人：文字表面反复改写而同一组织关系持续约束后文的写作者（本章的「隐性启动」格）。

最弱的一条是三个公开过程数据集的字段审计全部得出「不能直接计算」。本章把这一失败原样保留并转成实验方案，做法是对的；但读者应当明白，本章的核心读数至今没有在任何真实数据上被算出过一次。本章的本钱在设计，不在读数。

第二章

先析

——最先成形的那段合格文字如何使另一段永远不再成形

Table of Contents

先析

——一次写作里最先成形的那段合格文字，如何使另一段永远不再成形

摘要

作文教学长期把“写不出来”当作能力、输入或条件三者之一的不足，并据此在成品端、过程端、条件端分别下药。本文提出一个不同的判断：一次写作的失败，不是没能把内容组织成文，不是输入或暴露不足，也不是外部条件不许，而是在这一次写作里最先成形的那段合格文字消耗掉了同一批原料，并把自身写回环境成为下一次的现成形态，使另一段本来会成形的文字失去成形条件。本文把这段最先成形的产物称为先析，把它在一次产出中所占的比重称为先析率（ ϕ ）。材料取自三处互不相容的既有结论：结晶学关于亚稳相先于稳定相析出、以及一个已经量产的晶型在更稳定晶型出现后再也做不出来的记录；药理学关于同一物质的效应由给予的节律而非总量决定的记录；组织学习关于较差程序因先获成功而挤掉较好程序的记录。判据写成一句不含程度词的问话：这一次写出来的句子，在这一次开始之前，有没有以成句的形态出现过。本文用一份公开的中学生作文语料（39,672 篇）预注册并实跑了其中最便宜的一条检验：先析率与评分的秩相关为 +0.059，效应量可忽略而方向为正；最高分组中有 10.2% 的篇目先析率落在全样本第九十百分位以上。这一结果同时证实了该读数与现行评分正交，也暴露了它的一处粒度缺陷，本文据此收窄了适用区间。结论是：作文的发生取决于成形之前那段不产出可交付文字的间隔，而这段间隔在现行的任何一份作文记录里都没有字段。

关键词：先析；先析率；诱导期；接种史；写作发生；作文评价

Title: Pre-Precipitation: How the First Well-Formed Passage in an Act of Writing Forecloses Another

Abstract: Writing pedagogy has long treated “being unable to write” as a deficit of ability, input, or external conditions, and has accordingly intervened at the product end, the process end, or the condition end. This paper advances a different claim: the failure of an act of writing is not an inability to organize content into text, not insufficient input or exposure, and not the absence of enabling conditions, but rather that the first well-formed passage produced in that act consumes the same stock of raw material and writes itself back into the environment as ready-made form for the next act, thereby depriving another passage—one that would otherwise have taken shape—of the conditions for taking shape. This first-formed product is termed **pre-precipitate**, and its share of a given output is the **pre-precipitation ratio** (ϕ). The argument draws on three mutually incompatible established findings: from crystallography, Ostwald’s rule of stages and the record of a marketed polymorph that could no longer be manufactured once a more stable form appeared; from pharmacology, the finding that the effect of a substance is determined by the pattern of delivery rather than the amount delivered; from organizational learning, the finding that an inferior procedure that succeeds early crowds out a superior one. The criterion is stated as one question containing no modal terms: did the sentence written this time exist in sentence form before this time began. A preregistered analysis of a public corpus of 39,672 Chinese secondary-school essays yields a Spearman correlation of +0.059 between ϕ and rating—negligible in magnitude, positive in sign—with 10.2% of top-rated essays falling above the ninetieth percentile of ϕ . This result confirms the orthogonality of the measure to existing scoring while exposing a granularity defect, on the basis of which the paper narrows its range of application.

Keywords: pre-precipitation; pre-precipitation ratio; induction period; seeding history; genesis of writing; essay assessment

一、一个不报错的失败

先看一个在中学语文教研里被反复描述、却始终没有得到解释的现象。

一个初二学生，从初一下学期开始每周一篇周记，两年不断。素材本记满了，好词好句抄了三大本，范文读过两百篇以上。到初三，他的作文分数没有下降，多数时候还略有上升。但他的语文老师说了一句在教研组里流传很广的话：这孩子现在写什么都像作文，就是不像他自己写的。

再看第二个场景。一位高中语文教师做了一个对照：同一个班，同一道题，一半学生先读三篇下水文再写，另一半直接写。读过下水文的那一半，篇均字数多出约两成，句子更整齐，结构更完整，分数分布整体右移。而在两周后的另一道完全不同的题目上，读过下水文的那一半里，出现“整段与两周前几乎相同”的比例明显更高。教师的记录里没有任何一栏能容纳这件事——她只能在教案上写一句“迁移不够”。

第三个场景更极端。一位写作班老师发现，那些从三年级起就上作文班、写得又快又稳的孩子，到了初中，如果给他们一个从来没有教过的题目，比如“写一件你不愿意告诉任何人的小事”，他们交白卷的比例，高于从来没上过作文班的孩子。

这三件事有一个共同的形状：失败不产生任何信号。第一个学生的分数没降，第二个班的均分上升，第三批孩子在所有常规题目上表现更好。任何一个现行的检查环节——评分、批改、抽测、教研听课——都不会在这三处报错。一件东西没有长出来，而所有的仪表都显示正常。

本文要问的就是这一件事：一次写作到底是怎样发生的，以及那个“没有长出来”的东西，去哪儿了。

需要先把问题的形状说清楚。本文问的不是“作文是什么”，也不是“为什么要写作文”。本文问的是一条路：从什么状态出发，经过什么，到达一段成形的文字，以及这条路上的哪一步还插得上手。答案因此不会是一个定义，而是一段次序，外加一个可以在次序中间下手的位置。

二、三种现行的结算法

现行的解释不是没有，而是有三套，且这三套互不相容。它们分别在三个不同的端点上结算“作文发生了没有”。

2.1 成品端：成形即发生

第一套来自写作认知研究的主流。它把一次写作的成败结算在产出物上：内容能不能被组织成一篇结构完整的文字。这一套最完整的表述是知识陈述与知识

转化的二分 (Bereiter & Scardamalia, 1987) ——新手把想到的内容依次说出来, 成熟写作者则在内容问题空间与修辞问题空间之间往返, 写作过程本身改造了要写的东西。按这一套, 前面第一个学生的问题很清楚: 他停在陈述, 没有进入转化。处方也很清楚: 教转化。

这一套的划界标准是成没成形。没成形就是没发生; 成了形, 那就是发生了, 剩下的只是好坏之别。

2.2 过程端: 暴露决定结局

第二套把结算挪到过程上: 决定结局的不是有没有产出, 是这个过程本身的形状。这一套在药理学里有一个远比写作研究更硬的表述。Belchetz 等人在恒河猴身上做的实验给出一句可以逐字引用的结论: 这一现象“归因于促性腺激素释放激素被给予的方式, 而不是垂体细胞所暴露于的这一激素的量” (Belchetz, Plant, Nakai, Keogh & Knobil, 1978, *Science*, 202: 631–633)。同一个分子, 每小时脉冲给一次, 垂体恢复分泌; 连续恒速输注, 同样的总量甚至更大的总量, 垂体反而脱敏、下调, 分泌被压住。

这一套的划界标准是暴露够不够、节律对不对。不够就加量, 不对就改节律。它对第二个场景的解释是: 下水文这种输入是持续式的、高浓度的, 短期抬升, 长期下调。

这一套自己还留了一件更狠的材料。心律失常抑制试验 (CAST) 中, 用于抑制室性早搏的药物成功地改善了那个中间指标, 而接受治疗者的死亡率高于安慰剂组 (Echt et al., 1991, *New England Journal of Medicine*, 324: 781–788)。药理学自己承认: 一个被改善的中间读数, 可以与硬结局反向。

2.3 条件端: 条件决定产出

第三套把结算再往外挪一层: 个人做什么不重要, 重要的是他所在的那个环境奖励什么、记录什么、允许什么。组织学习研究给出的表述同样可以逐字引用: 能力陷阱发生于“一个较差程序上的良好表现使组织在该程序上积累更多经验, 从而使它在较好程序上的经验始终不足以让使用它变得有回报” (Levitt & March, 1988, *Annual Review of Sociology*, 14: 319–338)。March 后来把这件事写成利用对探索的挤出: 回报更近、更确定、方差更小的那一侧, 会系统性地吃掉另一侧 (March, 1991, *Organization Science*, 2: 71–87)。

这一套的划界标准是条件给不给。条件不到位, 就没有产出; 要改, 就改激励与记录。它对第三个场景的解释是: 作文班奖励的是稳定合格, 不奖励第一次写不好的东西, 孩子于是不去碰没有把握的题目。

2.4 结晶学的那一套，以及它为什么必须被算作第四种材料

上面三套里，第一套的划界标准最接近日常直觉，也最不容易被质疑：成了形就是发生了。而在一门与写作毫无关系的学问里，这条直觉早已被自己的材料推翻过一次。

结晶学处理的是同一件事的物质版本：一份过饱和的溶液里，什么时候、以什么形态、析出什么。Ostwald 在 1897 年给出的阶段律说：一个系统离开原有状态时，通常不会直接到达最稳定的那个状态，而是先到达能量上最接近的那个亚稳状态 (Ostwald, 1897, *Zeitschrift für Physikalische Chemie*, 22: 289–330)。也就是说，最先析出的那一相，通常不是最稳定的那一相。

这一条在工业上有极高的代价，因而被反复研究。而结晶学自己还掌握着一件更彻底的材料，本文第四节要专门处理它。

2.5 三者不能同真

把三条并排放，就能看出它们不是互补而是互斥。

第一套说：成形即发生。第二套说：成形与否不是判据，暴露的节律才是一——一次成形的产出完全可以是脱敏过程中的产物。第三套说：这两者都不是判据，是条件在决定，个体的路径与产物都是条件的影子。第四套（结晶学）说得更客气：最先成形的那一个，恰恰通常是不该被当作结局的那一个。

三方不能靠“判谁对”来化解，因为它们不在同一个语域里作答。第一套谈的是产物的属性，第二套谈的是过程的形状，第三套谈的是环境的账目。要把它们放在一起，只能另造一个框架。

三、三家共同假定了什么

三家争的是：“作文发生了没有”这件事，该在哪一端结清——成品端、过程端，还是条件端。

它们共同假定了一件从未被说出口的事：

这三端里总有一端能够一次结清；一次成功的产出，就是一次已经完成的发生。

这句话念给三家听，三家都会说这还用说吗。第一套当然认为一篇写成的文章就是一次完成；第二套认为只要暴露曲线对了，产物就是那条曲线的兑现；第三套认为只要条件到位，产出就会出现且算数。三家争的是在哪一端记这笔账，没有一家怀疑过这笔账能记完。

分歧全在归属，共识全在可结清性。而本文接下来要做的，是从三家之一自己的材料里，把这条共识拆掉。

四、一个已经量产的晶型，再也做不出来

1998 年夏天，雅培公司的抗艾滋病药物利托那韦（商品名 Norvir）出了一件在制药工业史上被反复引用的事。这个药已经上市两年，工艺稳定，批批合

格。忽然之间，生产线上开始析出一种此前从未见过的晶型；它的溶解度远低于原有晶型，做出来的胶囊不能用。产品被迫撤市，公司损失以亿美元计。

技术表征后来发表在《药物研究》上：利托那韦存在构象多晶型，两种晶格的溶解性质显著不同；那个新出现的、更稳定的晶型（II 型），其成核在能量上是不利的，除非溶液高度过饱和；1998 年那次事件是“高度过饱和的溶液”与“由一个降解产物引发的非均相成核”两件事恰好同时出现的结果

(Bauer, Spanton, Henry, Quick, Dziki, Porter & Morris, 2001, *Pharmaceutical Research*, 18: 859–866)。

真正要紧的不是 II 型出现了，而是接下来发生的事：同一套配方、同一个工厂、同一批工人、同一份工艺文件，I 型再也做不出来了。工厂被 II 型的晶种污染了。工艺开发团队后来把处理这件事的过程单独写成一篇文章

(Chemburkar et al., 2000, *Organic Process Research & Development*, 4: 413–417)。这类现象在结晶学里有一个专门的名字——消失的多晶型

(Dunitz & Bernstein, 1995, *Accounts of Chemical Research*, 28: 193–200)：一种曾经能被稳定制备的晶型，在更稳定的晶型在某个实验室或某条产线上出现之后，就再也制备不出来了，哪怕操作没有任何改变。

现在把这件材料放回第三节那条共有前提上。

I 型在 1996 年的那一次成功制备，是一次不折不扣的成品端结清：产物合格，检验通过，上市销售。但 1998 年之后，这次结清被追溯性地取消了——不是说那批药变质了，而是说“我们有能力做出 I 型”这个判断作废了。取消它的不是产物本身，也不是工艺过程，而是环境的一段历史：这个环境里已经出现过 II 型，并且这段历史无法从环境中撤销。

于是三端里没有一端能单独结清：

- 成品端结不清，因为一次成功的产出是否可重复，取决于环境的接种史，而接种史还在被不断写入；
- 过程端结不清，因为同一条工艺曲线在被接种前后给出完全不同的产物；
- 条件端也结不清，因为环境不是外生的背景——它正是被此前那些产物一次次写成现在这个样子的。

这件材料的分量在于它的来路。它不是本文从外面搬来的一个理由，而是结晶学自己已经掌握、已经发表、被反复引用的常识；它在结晶学自己的领域里回答的是另一个问题——为什么工艺不可复现，而不是作文如何发生。也正因为如此，三家无法退回自己的阵地：这件材料是其中一家自己的。## 五、先析

把上一节那件事从物质搬回文字，需要先立一个名字。

一次写作开始时，写作者手里有一批还没有形态的东西：一些印象、一句听来的话、一个说不清的别扭、一段还没找到词的感觉。这批东西是这一次写作的原料，而它的量是有限的——有限，不是因为人的容量有限，而是因为它们只在这一次的这个状态下同时在场；写完这一次，人换了状态，同一批东西不会再以同样的方式聚在一起。

在这批原料上，可能长出不止一种形态。有一种形态省力、可靠、随时可以取用：它由已经在环境里存在过的成句形态构成——背过的句子、上次用过的开头、范文里的结构、老师给的句式、这两年写过八十遍的那个抒情段。另一种形态费力、慢、成不成不一定：它没有现成模子，需要在这一次里长出来。

本文把前一种形态称为先析。

定义写死：先析，指一次写作中先于其他形态成形、由此前已存在的成句形态构成、并因其成形而占用了同一批原料的那部分产物。

这个定义有三个承重点，缺一个就不是先析：

1. 先——它在时间上先成形，而不是在质量上更差。先析物往往是这一次产出里最像样的部分。
2. 析——它是从同一批原料里析出来的，不是外加的。抄一段网上的话贴进去不是先析，那是抄袭；先析物是写作者自己写出来的，只不过写出来的是环境已经给好的那个形态。
3. 占——它成形之后，同一批原料不再能长出另一种形态。这一条是全部分量所在，也是最难被现有框架容纳的一条。

需要立刻处理一个同形词。冶金学里有“先共析相”（proeutectoid phase）这个标准术语，指相图上先于共析反应析出的那一相。本文借用的不是它。先共析是平衡态下的相图概念，它的“先”由温度与成分决定，且它不消耗后续相的成核条件；本文说的先析是非平衡下的成核选择，“先”由过饱和度与接种史决定，且它的全部要害正在于消耗。两者同源而不同所指，本文此后不再混用。

再处理一个更容易被读成同义的说法：套话。套话是一个关于质量的判断——这句话陈旧、空洞、不动人。先析是一个关于次序与消耗的判断——这句话先成形，且它的成形关掉了另一句。一段先析物完全可以在质量上很好：一个准确的比喻、一个漂亮的转折、一段真诚的抒情，只要它在这一次之前已经以成句的形态存在过，它就是先析物。质量判断与先析判断在同一段文字上可以给出相反的评价，这正是本文与套话批评的分界线所在。

六、四种落位与那条临界曲线

有了先析，就可以把“一次写作会长出什么”排成一张图。两根轴：

- 接种密度：这一次写作开始之前，写作者可以直接取用的现成成句形态的密度。刚读完三篇范文，接种密度高；连着两周没读过任何同类文字，接种密度低。
- 诱导期 τ ：从开始想这件事，到第一次产出一段可交付的合格文字，中间隔了多久。铃响就下笔，诱导期近乎零；想了三天才写第一句，诱导期是三天。

这两根轴分出四种落位。下表是本文的辨别装置，它替代了“写得好／写得不好”这一根轴。

落位	接种密度	诱导期	外观（当场看得见的）	短期读数	失效信号	可逆性
自发相	低	长	慢、不整齐、有几处别扭而准确的地方	篇均字数偏低，评分离散度大	有：写不完、跑题、交不上	高：这一次不成，下一次原料仍在
先析相	高	短	快、整齐、结构完整、像样	篇均字数高，评分稳定偏上	无	低：原料已被占用，且产物写回环境中
滞留	高	长	迟迟不下笔，落笔仍是现成形态	交稿延迟，成稿与先析相难以区分	弱：只表现为拖延	中：延长诱导期未换来另一相
未析	低	短	什么也没长出来	白卷、三行、“不知道写什么”	有：一眼可见	高：条件未被消耗

四格之间还有一条线。结晶学的经验是：成核所需的诱导期，随环境中已有晶种的密度上升而非线性地变长——高度过饱和才能压过不利的成核势垒，而接种一旦发生，这个门槛对另一相来说反而更高。落到写作上，这条临界诱导期 τ^* （要长出自发相，诱导期至少需要多长）随接种密度上升而快速变长。它的实际后果是：同一个学生，读得越多、练得越勤，他要长出一段不是先析物的文字，需要的空白就越长，而学校给他的空白只会越来越短。

现在指出本文要打的那一格：先析相。

这一格必须满足三条签名，缺一条就说明第二根轴选错了：

签名一：短期指标表现为改善。先析相区的产出更快、更整齐、更完整；篇均字数、结构完整度、分数分布全部好看。第九节的实践数据会给出一个更硬的版本：先析率与评分的秩相关为正。

签名二：失效不产生任何信号。这一格的产物形式合格，因此评分不报错、批改不报错、抽测不报错、教研听课不报错。没有长出来的那一相不出现在任何一份记录上——它不是被记为零，它根本不设字段。

签名三：自我加固。每一段先析物写出来之后都被批改、被表扬、被抄进素材本、被当作范文读给下一届。它于是成为环境中新的现成形态，接种密度上

升， τ^* 变长，下一次更难长出自发相。这一条与前两条一起，构成一个不需要任何人失职的闭环。

三条签名之外，还要说清楚这一格为什么需要一个框架才看得见。若把第二根轴设成“写得好／写得不好”，先析相区会整体落进“写得好”那一侧——它不需要任何人来揭发，因为它没有做错任何事。只有把轴换成“这段文字此前存在过没有”，这一格才从“好”里被分出来。

七、一句不带程度词的问话

判断一格是否成立，最后要落到一句可以拿去问记录、而不是拿去问人的判断的问话。本文的这句是：

这一次写出来的句子，在这一次开始之前，有没有以成句的形态出现过？

这句话里没有“应当”“真正”“充分”“实质性”“合理”这一类词。它不问这段文字好不好，不问学生用不用功，不问老师尽不尽责。它只问一件在原则上可清点的事。

把它拿到五个不同的场合去问，答案各不相同，而其中有两个答案是查出来的，且反过来改了本文。

场合一：高考考场作文。现行的作文题干标准表述是“不要套作，不得抄袭”，阅卷细则中通行的处置是：套作者基础等级在三等以下（含三等）给分，发展等级不给分；抄袭须提供被抄文章或线索，经确认后基础等级在四等给分。（检索日期 2026 年 8 月 17 日，检索词“高考作文 评分标准 套作 宿构”；本文取得的是各地阅卷细则的公开转载文本，未取得考试机构的正式印发件，此处只能算到“公开可见的细则文本”这一级。）这一套条款处理的是整篇层面的两个二值判断，由阅卷者当场认定，不产生任何逐句的来源记录。答案是：制度上有一个关于来源的开关，没有一个关于来源的刻度。

场合二：中学生周记。多数情况下没有任何可查的记录。老师凭印象说“这段你上次写过”。答案是：既无开关也无刻度，全靠个人记忆。

场合三：博士学位论文的文献综述。有查重系统，按连续相同字串的长度设阈值，逐句给出重复率。但这个系统只比对他人的文本；一个作者把自己三年前写过的段落原样搬进来，系统不报，制度上也不认为这是问题。答案是：有刻度，但刻度只朝外，不朝内。

场合四：企业周报与项目汇报。制度上要求先析：有模板库、有上季度的措辞、有必须沿用的口径。先析率高在这里是合规的表现，先析率低反而会被要求返工。答案是：刻度存在且方向相反，高者为佳。

场合五：律师起诉状与医嘱。同上，而且更强：使用经过审判确认或临床验证的既有表述，本身具有安全价值。一个“有创意”的起诉状是执业事故。答案是：先析是正当且必要的。

后两个答案是本文没有预料到、查过之后才知道的，它们直接改了本文的适用范围。第十一节会把这处让步写清楚。

八、先析率 ϕ ：一个三处量纲不同而比率相同的读数

判据要能被清点，就得有一个数。本文给出的是先析率。

先析率 ϕ = 一次产出中，此前已以成句形态存在过的部分所占的比重。

它在三个领域里的量纲完全不同，而算式相同：

- 结晶学：先于稳定相析出的亚稳相，占析出总量的质量分数。
- 药理学：进入体循环之前即被代谢掉的部分占给药总量的比例（首过损失）；或在持续给药下调之后，仍可被动员的受体占原有受体的比例。
- 写作：一次产出的文字中，此前已以成句形态存在过的部分占该次产出的比重。

三处量纲不同，比率相同。这个读数要成立，必须同时具备四条性质，本文逐条兑现：

1. 无量纲——三个领域可以并排比较。
2. 可事后清点——不需要在写作现场观察，事后拿产出物比对即可。
3. 把一个印象变成一个数——“这孩子写什么都像作文”这句话，此前只是印象。
4. 与既有主要指标正交——这一条最容易漏、也最要命。作文这个领域的既有主要指标是分数。若 ϕ 与分数高度相关，它会立刻被读成分数的一个代理，白造。第九节的实跑给出了这一条的直接兑现： ϕ 与评分的秩相关为 +0.059，效应量可忽略；且最高分组中有 10.2% 的篇目 ϕ 落在全样本第九十百分位以上——分数最好看而这个数最难看的例子是真实存在的，且并不罕见。

清点手册

一个读数没有清点手册就不是读数。以下四项在本文正文、证伪条款、末节赌注三处引用同一份定义、同一个阈值、同一套编码规则。

读数名：先析率 符号： ϕ

定义：一次产出中，被“此前已以成句形态存在过”的字串所覆盖的字数，

除以该次产出的总字数。

粒度：以连续 8 个汉字为最小成句单位。

恰在边界上的例子：一个七字的四字成语加三字短语组合

(如"守得云开见月明", 7 字),

不计入; 同一句里出现两次的同一个八字串, 按位置只计一次。

编码规则 (六条):

1. 只保留汉字, 标点、数字、拉丁字母、空白一律删除。
2. 以原始句子为单位切分, 字串不跨句边界。
3. 判定"此前已存在"的证据源必须在测量前写死: 本人此前的写作、指定的范文集、教材、素材本, 四者取并集。
4. 篇内位置重复只计一次。
5. 先做一道前置筛: 整篇近似复制 ($\phi \geq 0.5$) 者剔除, 另案处理。

6. 未通过前置筛的篇目不进入 ϕ 的任何统计, 也不得用它作任何评价。

表头: 篇目 ID | 总字数 | 被覆盖字数 | ϕ | 是否通过前置筛
| 证据源版本号

不适用: 以可靠交付为目的的写作 (公文、法律文书、医嘱、安全规程、

标准化报告), 这些场合的高 ϕ 是正当的, ϕ 在其中不作评价用。

三处一致性自查: 正文 §8 = 证伪 §17 = 赌注 §21 一致

九、三万九千篇作文: 一次真跑, 以及一条不利的结果

证伪条款不能全是纸面的。本文在成文之前, 用一份公开可复算的语料跑通了其中最便宜的一条, 并把结果——包括对本文不利的部分——写在这里。

9.1 预注册

统计运行之前, 假设、变量、阈值、编码规则已写死存档, 事后未作任何修改。

数据: GitHub 公开仓库 `cnunlp/Chinese-Essay-Dataset-For-Pre-Training`, 随 Song 等人的论文发布 (Song, Zhang, Fu, Liu, Liu & Cheng, 2020, *Proceedings of EMNLP 2020*, 6723–6733)。全库 93,002 篇中学生作文, 字段含句子列表、评分 (四档)、文体 (六类)、年级。按文件行序取前 40,000 篇 (容器内存 3GB, 全量索引跑不动, 该规则在跑数前写死); 剔除清洗后汉字少于 100 的篇目, 最终 $N = 39,672$ 。

代理变量声明: 本次计算的是 ϕ 的一个下界代理 $\hat{\phi}$ 。语料没有时间戳, 也没有作者身份, 因此无法判断"这一次之前"。 $\hat{\phi}$ 只能测"在同一语料环境中另有

他人也写出过同样的八字串”，它把共同来源（成语、名句、教材例句）与本人复用混在一起。 ϕ 的核心定义在这份公开记录里跑不动，因为字段不存在。这一点本身就是本文第三层主张的证据，第十七节第 6 条据此设了一条最便宜的证伪。

三条预注册假设：

- H1（正交性）： ϕ 与评分的 Spearman ρ 满足 $|\rho| < 0.20$ 。若 $\rho \leq -0.20$ 且 $p < 0.01$ ，则“与既有主要指标正交”这条性质受挫，必须修订。
- H2（上尾存在性）：最高分组中， $\phi \geq$ 全样本第 90 百分位的篇目占比 $> 5\%$ 。
- H3（文体差异）：议论文的 ϕ 中位数高于叙事文。

9.2 结果

项目	读数
样本量	39,672 篇
ϕ 中位数	0.0736
ϕ 均值	0.1442
ϕ 分位 (25/50/75/90/95/99)	0.038 / 0.074 / 0.131 / 0.273 / 0.940 / 1.000
H1： ϕ 与评分秩相关	$\rho = +0.059$, $p = 2.1 \times 10^{-32}$ (10.2% ($>5\%$), 通过)
H2：最高分组中 $\phi \geq p90$ 的占比	0.0891 / 0.0716, $p = 6.2 \times 10^{-48}$ (方向一致, 通过)
H3：议论文／叙事文 ϕ 中位数	$\rho = 0.036$ (长文更易撞句这一解释不成立)
混淆检查： ϕ 与篇长秩相关	
按评分分组的读数（评分 0 为最低档，3 为最高档）：	

评分档	篇数	ϕ 中位数	$\phi \geq$ 全样本 p90 的占比
0 (最低)	27,664	0.0716	8.9%
1	3,320	0.0737	9.9%
2	8,452	0.0809	13.8%
3 (最高)	236	0.0699	10.2%

三条假设全部通过。其中 H1 的意义最大：先析率与评分不构成可用的代理关系，且方向为正。那句最朴素的竞争解释——“其实不就是差生才用套话吗”——在这份数据上不成立：秩相关是正的，写得被评为更好的那一批，先析率不低反略高。这条结果同时兑现了第八节所要求的第四条性质。

9.3 一条不利的结果，以及它改了本文的什么

ϕ 的分布不是单峰的。第 95 百分位是 0.940，第 99 百分位是 1.000——有 5.2% (2,074 篇) 的篇目，其八字串几乎全部能在语料的别处找到。

第一反应是：这些是重复收录的同一篇。为查证，本文另做了一次事后计算（明确标注为事后分析，非预注册）：把每篇清洗后的全文取哈希，逐字完全相同的篇目只有 88 篇，占 0.22%。

这两个数放在一起，给出一个本文没有预料到的结果：存在一批文本，它逐字不同于任何一篇已有的作文，而它的每一个八字连续串都不是新的。这批文本占样本的 5%，是逐字重复者的二十四倍。

这件事对本文有利也有不利。有利的一面是：现行的抄袭与套作条款抓不到这一类——它们判的是整篇的同一性，而这批文本在整篇层面上是“原创”的。不利的一面更要紧：本文的读数在这个区间里失去了分辨力。 ϕ 达到 0.94 以上时，本文无法区分三种情形：一次极端的先析、一次经过改写的整篇转载、以及同一篇范文被多个学生分别默写后的变体。三者的机制完全不同，而读数相同。

因此本文当场做了两处修订，并把它们写进了第八节的清点手册：

1. 加一道前置筛： $\phi \geq 0.5$ 的篇目一律剔除，另案处理。理由是这一区间混入了机制完全不同的对象。 ϕ 的适用区间因此收窄为 $[0, 0.5)$ 。
2. 正交性的表述收紧： $\rho = +0.059$ 在这个样本量下 p 极小，因此不能说“ ϕ 与评分无关”，只能说“效应量可忽略”。事后剔除 $\phi \geq 0.5$ 的篇目后重算， $\rho = +0.041$ ，方向与量级都稳定。本文此后一律用“效应量可忽略”这一表述。

还有一处必须交代的短板：本次实跑用的是跨作者的共享作为“此前已存在”的代理，而定义要的是同一作者在这一次之前。这两件事的差距不是一个技术细节，它使本次实跑只能证实读数的正交性与上尾的存在性，不能证实先析的因果机制。机制那一条要靠第十七节的第 1、3、4、5 条，而那四条都需要现有记录里不存在的字段。

（本节全部数据与代码可复算：数据来自上述公开仓库，编码规则见第八节清点手册，统计用 Spearman 秩相关与 Mann-Whitney U 检验。）## 十、那条路，以及哪一步还插得上手

到这里可以回答开头那个问题了：作文如何发生。

答案不是一个定义，是一段次序。它有四步，且这四步的顺序不能换。

第一步，环境先动。一次写作开始之前，环境已经把它接种史交给了写作者：他读过什么、被表扬过什么、上周写过什么、这个班里流行什么开头。这一段历史不是背景，它决定了这一次哪些形态可以被直接取用、哪些形态需要从头长。接种密度是这一次写作的初始条件，而不是它的外部环境。

第二步，节律决定过饱和。原料的积累不是靠总量，是靠形状。持续的、高浓度的输入——每天一篇范文、每周一次仿写训练——会把系统推到一种奇怪的状态：现成形态密度极高，而未成形的原料反而少，因为每一次接触都会顺手把一批模糊的东西固定成已有的说法。间断的输入则相反：接触之后有一段不产出的时间，模糊的东西留在模糊状态里继续积累。药理学那句话可以逐字搬过来——起作用的是给予的方式，不是暴露的量。

第三步，成核选相。到了要写的时候，长出哪一相，取决于两件事的比较：过饱和度是否高到足以越过自发相的成核门槛，以及诱导期是否长到足以让它越过。接种密度高时，先析相的门槛几乎为零——它随时可以成形；自发相的门槛却因为接种而变得更高。于是在绝大多数场合，先成形的是先析相。这不是选择，是次序。

第四步，产物写回环境。这一步是全段次序里最容易被漏掉、也最要命的一步。先析物写出来之后不会停在纸上：它被批改、被打分、被读给全班、被抄进素材本、被当作下一届的范文。它成为环境中新的现成形态，接种密度上升，下一轮的第三步因此更偏。回写是这条路成为一个循环、而不是一支箭头的原因。

四步之后，插手的位置就清楚了，而且它不在人们习惯下手的两个地方。

不在第三步的产物端。在成品上下功夫——批改、评讲、修改、重写——是当前投入最大的一处，而它同时是接种密度最高的一处。一次精细的评讲会把若干个“好的说法”当场种进全班的环境。它抬高短期读数，同时抬高 τ^* 。

不在第一步的总量端。加大阅读量是第二大的投入。按本文的次序，增加输入总量同时增加两样东西：未成形的原料，和现成的形态。哪一样增加得更快，取决于输入的节律。持续式的高密度输入，增加现成形态更快。

在第二步。可插手的位置是节律与诱导期，而具体的动作只有一条，且这条动作在形式上是一个禁令：

在过饱和度尚未足以支持自发相之前，不产出可交付的合格文字。

落到操作上，它是一段有明确长度、且明确禁止产出成品的时间：知道题目之后，先不写任何“能交的东西”——不写提纲，不写开头，不写“先随便写点”，也不看任何问题的现成文字。这段时间里可以做的是另一类事：把还没有词的东西说给自己听、找具体的细节、回到那件事本身。这段时间的长度不是固定值，它随接种密度上升而变长。

这条动作与两种流行做法方向相反，必须点明。

其一是自由写作：不停笔地写，不管好坏，先写出来再说 (Elbow, 1973)。按本文的次序，不停笔地写是持续式给予加上允许先析——它在接种密度低的人身上确实能长出自发相，在接种密度高的人身上则会以最快的速度把先析物倒出来，并且倒得很痛快。

其二是先列提纲。提纲是最典型的先析物：它由现成形态构成（总分总、三个分论点、首尾呼应），它先成形，且它一旦成形，后面的文字只能在它划定的格子里长。

本文因此给出一条与两者都不同的处方：诱导期内既不自由写作，也不列提纲，而是什么可交付的东西都不产出。

十一、三处反过来的约束

一个判断如果只会往一个方向长，多半是描述而不是机制。本文有三处被材料反过来削掉的地方，逐处写清削掉了什么。

第一处：先析在一整类写作里是正当的。第七节的场合四与场合五是查出来才知道的：企业周报、法律文书、医嘱、安全规程，这些场合的制度要求高先析率，而且要求得有道理——使用经过验证的既有表述，是可靠性的来源。若没有这一处，本文原本会写下的更强的那句话是：“先析是写作的一种普遍病理。”现在这句话不成立了。本文的适用范围因此收缩为：以形成一个此前不存在的判断为目的的写作。凡以可靠交付为目的的写作，本文的全部结论都不适用， ϕ 在其中不作评价用。这一处动的是适用范围。

第二处：留白本身不产生任何东西。药理学那条材料反过来给本文加了一道硬约束。脉冲给药之所以有效，是因为受体存在一个下调之后的恢复时间常数；如果没有这个常数，间断给予就只是给得更少。落到写作上：若“过饱和和度”不会因为等待而回升，只会因为新的接触而回升，那么单纯的空白就不产生自发相，它只是把不成核推后——第六节那张表里的“滞留”格，正是这种情形的落位。若没有这一处，本文原本会写下的更强的那句话是：“只要不写、留出时间，另一相就会长出来。”现在只能写成：诱导期只有在原料仍在积累时才有意义；空等把未析变成滞留，不把滞留变成自发相。这一处动的是机制的完整性——本文必须承认，它给出的干预是必要条件，不是充分条件。

第三处：本文的处方在现行制度里不可能被采纳。组织学习那条材料反过来说：探索的收益是长尾的、延迟的，而利用的收益是近的、确定的；任何以短周期结算的组织都会系统性地选择后者。一个学期结算一次的评价制度，不会采纳一条“这段时间不产出任何可交付文字”的安排——不是因为它不理解，是因为这条安排在它的结算周期内只呈现为成本。若没有这一处，本文原本会写下的更强的那句话是：“学校应当在作文教学中设置诱导期。”现在这句话被删掉了。本文能主张的只剩一件：在不以短周期结算的场合——家庭、个人写作、师徒式的长期指导、不计入评分的写作——这条动作可以执行；在以学期结算的场合，提出它等于提出一件不会被做的事。这一处动的是价值排序：本文不再向制度提建议。

十二、十二处兑现

一个跨领域的判断，价值不在于它“像”别的领域的什么，而在于能不能据它在那个领域里预测一件具体的、可以去查的事。以下十二处，每一处都写成一条可查的预测；其中若干处在该领域里已有现成数据可以回溯检验。

一、制药结晶工艺。若一条产线上出现了更稳定的晶型，则同一企业中与该产线共用人员、器具或通风的其他产线，在此后六个月内出现同一晶型的概率，显著高于未共用者。这条预测有一个现成的对照：利托那韦事件之后，处理办法正是隔离产线、专人专服、更换器具——如果接种只是巧合而不是机制，这些措施就是白做的。

二、内分泌治疗。用促性腺激素释放激素激动剂做长期抑制治疗后，停药到轴功能恢复所需的时间，应随持续给药时长非线性变长，而不是线性变长。本文预测的是与临界诱导期同样的形状：接种越久，恢复所需的空白越不成比例地长。

三、育种。对某一性状连续多代强选育的群体，在被要求对一个新性状作出响应时，其响应速率低于未经强选育的近缘群体，且这个差距不能靠加大选择强度补回来。这里的“原料”是群体中的遗传变异，先析物是已经被固定下来的那个性状。育种记录中有现成的多代数据可查。

四、软件工程。一个功能在早期被复制粘贴形成一份“能用的实现”之后，该功能被彻底重写的概率随复制次数下降；而且这个下降早于缺陷率的上升出现。顺序是这条预测的要害：若重写率的下降与缺陷率的上升同时发生，那更像是资源问题；若下降在前，说明可重写性在缺陷出现之前就已经被消耗掉了。

五、音乐即兴。即兴演奏者在上台前听过同一曲目的著名版本，则当晚乐句中与该版本重合的比例上升；本文进一步预测，与他自己上一场重合的比例也同时上升。第二条比第一条难得多——它意味着接种不只是提供了模仿对象，而是压低了这一晚长出新乐句的可能。

六、法律文书。某份判决的说理段落被上级法院引用之后，同一法院后续同类判决中该段落的复用率上升，而新说理的出现率下降；两者的时间差可测。法院裁判文书是公开的，这条可以直接回溯检验。

七、组织例程。一次成功的应急处置被写成预案之后，下一次同类事件中偏离预案的处置减少；且偏离的减少早于处置效果的下降。这条预测把预案的收益与代价分开了：预案本身没有错，错的是它同时是一次接种。

八、方言与说法的消失。某个说法被写进教材或字典之后，同一社区中表达同一意思的其他说法，其消失速度加快——比未被收录的义位更快。语言志与方言调查有多个时间断面的数据可比。

九、生成模型训练。递归训练导致分布尾部消失这件事已有实证。本文在此之上加一条更细的预测：尾部消失的速度，随“模型生成物被人类采纳并写回公开语料”的比例上升而加快，而不只是随训练代数上升。若尾部消失只与代数有关而与人类采纳率无关，那么本文关于回写的那一步就不适用于这个领域。

十、医学教育。标准化病人考试引入之后，学生问诊语句的模板化程度上升，而考试分数同时上升；本文预测，模板化的上升早于“漏掉罕见主诉”的出现。这条与第四条共用同一个形状：短期读数改善在先，失效在后，而中间那段时间里没有任何仪表报警。

十一、家常烹饪。预制调味包普及之后，同一菜系家庭做法中“自配比例”的分布方差缩小，而消费者对成品的评分不降甚至上升。方差缩小而评分不降，正是本文那张表里先析相区的读数形状。

十二、科研写作。某个领域出现一篇被高频引用的范式论文之后，该领域后续论文方法部分的措辞趋同；本文预测趋同早于新方法出现率的下降。这一条可以用任一学科的全文语料直接跑，所需字段（发表年份、方法节文本）全都存在。

十二处之外要补一句：这些兑现里，第四、六、七、十、十二五条都是顺序预测（谁早于谁），而不是相关预测。这是刻意的。相关很容易被巧合满足，先后不容易。若这五条中有三条以上的时间顺序与本文预测相反，本文第一层主张即使在写作领域内成立，也说明它不是一个跨领域的结构，只是一个关于写作的经验概括。

十三、与最近的几种说法的分界

本文的判断与下面十种既有说法最容易被混为一谈。逐条给出各自说到哪一步、分界落在哪里，以及一条能把两者分开的观察。

其一，知识陈述与知识转化（Bereiter & Scardamalia, 1987，写作认知研究）。它说到哪一步：新手把想到的依次说出来，成熟写作者在内容与修辞两个问题空间之间往返，写作过程改造了要写的东西；写不出好文章，是因为缺少转化机制。分离线：这一套把写作者手里的内容当作给定的存量，转化机制作用于这个存量之上；本文说存量会被产出本身消耗，而且被消耗掉的部分不进入任何账。判定形式：给同一批学生安排“先口头讲一遍再写”的条件——若口头讲述提高了后续书面产出的量与质，则转化不足这一解释对而本文错；若口头讲述提高了产出量而同时提高了书面产出中的先析率（因为口头那一次已经是一次先析），则本文对。这一条可用现有课堂条件在一学期内做完。

其二，写作作为知识构成的过程（Galbraith, 1999, in Torrance & Galbraith eds., *Knowing What to Write*, Amsterdam University Press）。它说到哪一步：写作本身构成知识，而非仅仅传递既有知识；对某些写作者而言，先列提纲反而减少新想法的产生。这是十位里离本文最近的一位，方向也最一致。分离线：Galbraith 把提纲的抑制作用归于表述方式（预先的规划取代了构成过程），因此换一种规划方式即可缓解；本文把它归于存量消耗与接种，因此换规划方式无用，只有不产出可交付文字才有用。判定形式：比较三组——列提纲组、用非线性方式（如概念图、关键词云）规划组、不产出任何可交付文字组。若前两组的差异大于后两组，则 Galbraith 对；若非线性规划组与提纲组同样抬高先析率、只有第三组降低，则本文对。

其三，定势效应（Luchins, 1942, *Psychological Monographs*, 54(6), whole no. 248）。它说到哪一步：连续用同一个复杂解法成功之后，被试在一个可用更简单方法解决的题目上仍然沿用原解法；给出“不要盲目”的提示

后，机械解法的使用率降到五成左右。分离线：定势是注意与可及性的问题，一句提示就能解除一半；本文说的是条件被消耗的问题，提示解除不了。判定形式：在写作任务前给出明确提示——“注意不要用你以前写过的说法”。若先析率因此显著下降，则这是定势，本文错；若提示只改变表层措辞而先析率（按第八节编码规则，以八字串计）不降，则本文对。这一条极便宜，一个班一节课可做。

其四，功能固着（Duncker, 1945, *Psychological Monographs*, 58(5), whole no. 270）。它说到哪一步：一个物体先前被用于某个功能之后，被试更难看出它的另一种用法。分离线：功能固着说的是同一个对象的另一种用法被遮蔽，遮蔽者与被遮蔽者共用一个载体；本文说的是同一批原料只能长出其中一种形态，两种形态并不共用载体，而是竞争同一份存量。判定形式：若把先前写过的说法从写作者眼前彻底移除（不给素材本、不许回看旧作），固着解释预测先析率下降；本文预测不变，因为消耗发生在这一次的产出中，不发生在回看中。

其五，能力陷阱与利用挤出探索（Levitt & March, 1988；March, 1991）。它说到哪一步：较差程序上的早期成功使经验向它集中，较好程序因经验不足而始终不划算；组织因此在利用一侧越陷越深。分离线：能力陷阱是一个关于选择与回报的机制，它预测改变激励结构可以改变选择；本文说的是一个不由激励决定的次序——即便回报完全倒过来，只要接种密度高且诱导期短，先成形的仍然是先析相。判定形式：设计一个奖励反转的条件（明确宣布“用过的说法不给分，新长出来的说法即使不成熟也给分”）。若先析率随之显著下降，则这是能力陷阱；若激励反转只让学生把先析物做同义改写而八字串层面的先析率不降，则本文对。

其六，公式化语言与预制语块（Wray, 2002, *Formulaic Language and the Lexicon*, Cambridge University Press；Biber, Johansson, Leech, Conrad & Finegan, 1999, *Longman Grammar of Spoken and Written English*）。它说到哪一步：语言使用中存在大量整体存储、整体调用的多词单位，它们是流利的基础而非缺陷。分离线：这一套把预制单位当作能力的构成部分，它越多越流利；本文不否认这一点，但指出预制单位的调用在一次以形成新判断为目的的写作中会占用同一批原料。同一段文字，语块研究判为流利的证据，本文判为先析物。这是本文与它之间最锋利的一条：两边对同一份材料给出相反的评价。判定形式：取一批高语块密度的文本，若其在“作者事后能否说出该段所指的具体经验”上与低语块密度文本无差异，则语块研究对；若高语块密度段落系统性地对应不出具体经验，则本文对。

其七，自由写作与前写作（Elbow, 1973, *Writing Without Teachers*, Oxford University Press；Rohman, 1965, *College Composition and Communication*, 16: 106–112）。它说到哪一步：写作困难来自过早的编

辑与评判，解法是先不加评判地写、或在写之前专门安排一个发现阶段。分离线：两者都把“先产出”当作解药——自由写作要求不停笔地产出，前写作要求在写之前产出探索性文字；本文把“先产出”当作病因，主张诱导期内不产出任何可交付的东西。判定形式：问题分三组，自由写作组、前写作组、不产出组。若前两组的先析率不高于第三组，则本文错；若前两组显著更高，则本文对。这一条与其三、其五共用一个实验平台，可一次做完。

其八，击键记录方法学 (Leijten & Van Waes, 2013, *Written*

Communication, 30: 358–392)。这是方法学上离本文最近的一位：它能逐键记录写作过程，包括所有被删除的内容与停顿。分离线：击键记录记的是落到键盘上之后的一切；本文关心的那一类东西——起了头、在心里成了句、又被自己否决、从未落键——在击键记录里不存在。判定形式：若把击键记录的删除量与停顿时长作为自变量，能够预测本文的先析率，则本文的读数被它覆盖；若两者不相关（即写得顺、删得少的人先析率也可以很高），则本文的读数不可被它取代。这一条可用现有的击键记录数据集直接回溯检验。

其九，套作与宿构条款（中国作文阅卷实务）。它说到哪一步：整篇沿用已有构思、对原材料简单扩写、或对已有文章简单改写者，基础等级降档、发展等级不给分。这是本领域内部用了几十年的实务正主，必须正面处理。分离线有三处：条款判整篇，本文判逐句；条款是二值，本文是连续量；条款只处理沿用他人，本文把沿用自己算得同样重。判定形式：第九节的实践已经给出一条——样本中有 5.2% 的篇目其八字串几乎全部在别处出现过，而逐字完全相同者只有 0.22%。这批文本在条款下是原创的，在本文的读数下接近上界。若这批文本经人工复核后被判定为整篇改写型抄袭，则条款覆盖本文；若其中相当一部分被判定为“确系学生自己写的、只是每一句都不是新的”，则本文对。

其十，目标置换 (Merton, 1940, *Social Forces*, 18: 560–568)。把本文的命题剥掉写作这个领域的全部名词，剩下的形式是：一个系统为了产出可交付的东西，消耗掉了它据以产出的那样东西，并把消耗的痕迹写成了成绩。这个纯形式的结构在上游有名字：手段吞噬目的，即目标置换；相邻的还有成功陷阱 (Levinthal & March, 1993, *Strategic Management Journal*, 14(S2): 95–112) 与内卷化 (Geertz, 1963, *Agricultural Involution*, University of California Press)。分离线：目标置换讲的是规则与目的之间的替换，它的载体是组织成员对规则的遵从；本文讲的是原料存量的消耗，它不需要任何人遵从任何规则，一个完全自主的写作者独自在家也会遭遇它。判定形式：在没有任何外部规则与评价的条件下（自己写日记，无人阅读，无人评分），若先析率仍随接种密度上升，则目标置换解释不适用而本文成立；若先析率在无规则条件下消失，则本文只是目标置换的一个场景。

另有一位来自完全不同的领域，必须请进来。模型坍缩：生成模型在递归使用自身生成的数据训练时，会出现不可逆的缺陷，原始内容分布的尾部消失

(Shumailov, Shumaylov, Zhao, Papernot, Anderson & Gal, 2024, *Nature*, 631: 755–759)。这一条与本文第十节第四步（产物写回环境）几乎是同一个结构，且它是这个结构目前最强的实证。分离线：模型坍缩发生在群体与多代的层面，需要多轮回灌才显现，且它描述的是分布尾部的消失；本文说的事发生在一次写作内部，第一次就发生，不需要任何递归。判定形式：取一位接种密度接近零的写作者（从未受过写作训练、也很少读同类文本），给他一次写作任务。若先析仍然发生——即他产出的第一段合格文字仍由此前已存在的成句形态构成，且它关掉了另一段——则本文对，且模型坍缩只是它在多代尺度上的一个特例；若在零接种条件下先析不发生，则本文是模型坍缩的一个下位现象。

十条之外，还有三位本文尚未正面交手，在此指名请进，以便后来者从这三个方向检验：Anderson 的产生式复合与技能自动化（skill automatization，认知技能获得研究）；Sperber 与 Wilson 之后的关联理论关于“最省力的解释先到达”的论证；以及外文语境中的 writer’ s block 研究传统（原题 *Writer’ s Block: The Cognitive Dimension*, Rose, 1984）。第三位尤其危险：它与本文处理的是同一批现象，而它的解释落在规则的僵化上，与本文的存量消耗不同，值得一次正面较量。

十四、一块所有人都站着的地基

把上面十条邻居各自看不见的东西列出来，会发现它们指向同一处。

- 知识陈述与转化的二分，把“写作者手里的内容是一个给定的存量”当作给定，因此看不见“产出会消耗存量”。一旦承认后者，陈述与转化的二分就塌了——因为陈述本身改变了可供转化的东西。
- 知识构成模型把“写作产生新知识”当作给定，因此看不见“先产生的那一批会关闭后产生的那一批”。一旦承认，“多写就多出”这条就塌了。
- 定势效应把“可及性是注意的问题”当作给定，因此看不见“可及性可以是存量的问题”。一旦承认，提示即可解除这条就塌了。
- 能力陷阱把“选择可以被激励改变”当作给定，因此看不见不由激励决定的那一类锁定。一旦承认，组织学习的可干预性前提就塌了。
- 套作条款把“沿用是一个可以被认定的行为”当作给定，因此看不见“沿用可以在逐句层面发生而在整篇层面不可认定”。一旦承认，条款的可执行性就塌了。

五行背后是同样被当作给定的东西：

没有长出来的那一样东西，仍然是可以长出来的——它只是没被选中、没被注意到、没被奖励。

未发生者被当作待发生者。这是所有人共同站着的那块地，互相批判的各派都站在上面，因此谁也没有回头看它。教学的全部干预都建立在它上面：只要方法对、动机足、条件好，那个没写出来的东西就会写出来。

本文说的是另一件事：未发生者里有一部分已经不可能发生。不是没被选中，是它赖以成形的条件已被先析物耗掉，且接种史已被写入而无法撤销。这一类东西在任何一份作文记录里都不设字段——它不是被记为零，它压根不算一样东西。删掉本文这个读数，它不是变得难测，是不存在。

这也是本文与前面十位的根本分处：他们都在问“为什么它没有被选出来”，本文问的是“它还在不在候选之列”。

十五、与同栏几篇的分界

本栏此前已有数篇涉及相邻结构或相同学科，逐篇指名，并各给一条能把两者分开的观察。

与《熄形》的分界。那一篇处理的是“机会充足而一次也不成形”，读数是空机次；它用的也是结晶学的可成形性一线。两者是互补关系而非竞争关系：熄形计的是零产出的次数，本文计的是产出中被先析物覆盖的比重。数学上二者近乎正交：一个学生一学期可以空机次为零（每次都写出来了）而先析率极高；也可以空机次很高而每次真正写出来的都是自发相。判定：若某样本空机次为零而先析率在 p_{90} 以上，则本文适用而《熄形》不适用；若空机次很高而先析率在中位数以下，则反之。第六节表中的“未析”格属《熄形》，“先析相”格属本文。

与《阙用》的分界。那一篇处理的是可用集的退出——某样东西的末次使用时刻之后就不再被调用。本文处理的是可用集的占用：先析物不退出，它一直可用，而且正因为一直可用才关掉了别的。判定：取同一份材料，若某个说法长期不再被使用，属《阙用》；若某个说法被高频使用且其高频使用与另一类说法的消失同时发生，属本文。

与《垫指》的分界。那一篇的核心是不留痕的动作——做了而没有任何记录。本文的先析物留痕充分：它被批改、被打分、被抄录。本文关心的不留痕者是另一样东西——被它关掉的那一相。判定：若某样本的问题在于动作不进记录，属《垫指》；若动作全部进了记录而记录本身不设“这段占用了什么”的字段，属本文。

与《己资》的分界。那一篇借的是自创不得确认这条会计规则，讲自己长出来的东西不得入账。本文讲的是相反方向的一件事：自己长出来的东西不但入了账，而且入账之后成为下一轮的现成形态。两篇合起来给出一个完整的两难：不入账则不被承认，入账则成为接种。判定：若某样本的问题在于产物无法被确认，属《己资》；若产物被顺利确认且确认行为本身抬高了后续的先析率，属本文。

与本栏用过同样学科的几篇的分界。结晶学此前出现在数篇之中，本文取的是它的另一条线——阶段律与消失的多晶型，而非成核门槛；药学前出现过

两次，本文取的是给予节律与首过损失，而非剂量—效应窗口；管理学此前出现过数次，本文取的是能力陷阱与利用—探索，而非协同成本。三处文献线互不重叠，且三门在本文中所处的位置与此前各篇均不相同。## 十六、不适用的边界

本文的判断有五处明确不适用，逐条写死，以免被拿去用在它管不着的地方。

其一，以可靠交付为目的的写作。公文、法律文书、医嘱、安全规程、标准化报告、技术手册。这些场合的高先析率是正当的，且往往是必要的。在这些场合使用本文的读数去评价人，是误用。

其二，学习一门新语言或新文体的初期。第二语言学习者、刚接触某种文体的写作者，其产出必然高度依赖预制形态，且这是获得该形态的必经之路。本文的判断以“写作者已经具备该文体的基本形式能力”为前提；不具备时，先析不是问题，是台阶。

其三，极短的产出。一条批注、一句评语、一段两百字以内的说明。八字串的统计在这个长度上不稳定，且这类产出本来就复用于常态。本文的读数在总字数低于三百汉字时不予采信。

其四， $\phi \geq 0.5$ 的区间。第九节的实践证明本文的读数在这个区间失去分辨力，无法区分极端先析、改写型转载与集体默写。这一区间的对象须另案处理，不进入本文的任何统计与任何评价。

其五，把本文用于个人品性判断。先析率高不等于偷懒、不等于不诚实、不等于缺乏才华。第九节的数据里，被评为最好的那一批的先析率不低反略高。用这个数去判断一个人是什么样的人，是本文明确拒绝的用法。

十七、证伪条件

本文的主张分三层，各层的证伪条件不同。第一层最强也最易倒，第三层最稳且最便宜。分层的用处是：第一层被推翻时，后两层仍站得住。

- 第一层：先析是一个真实的机制——最先成形的合格文字消耗同一批原料并关闭另一相。
- 第二层：第六节那张四格表与临界诱导期曲线是一个有用的分析工具。
- 第三层：现行的作文记录里没有任何字段记录一次产出中有多少来自此前已成句的材料，因此“越练越写不出”这件事在制度上不可观测。

最强的竞争解释必须先写出来，不许挑一个好打的当靶子。那句最朴素、最难反驳的话是：其实不就是学生偷懒、老师逼出来的套路化吗？这是一个关于动机与制度的解释，它能解释本文列举的三个场景中的每一个，而且它比本文简单得多。下面的条款必须在控制掉它之后仍然成立。

条款一（剂量—反应，机制主件）。在同一所学校、同一年级、同一套评价制度、同一位教师的六个及以上班级内（同质多单元，不得用两省、两校、两

国的对比充数），测量各班的接种密度（可直接取用的现成成句形态密度：范文投放量、素材本共享程度、模板讲评次数）与先析率。控制学生的写作投入时间与教师的评价严格度之后，若接种密度与先析率之间不存在剂量—反应关系，则本文第一层为伪——届时“越练越写不出”应归因于动机与制度，而非成核条件。

条款二（正向读数，写成顺序而非相关）。在一个持续两学年的追踪中，先析率的上升早于产出量与评分的下降出现，且早出的时间间隔可测。若三者同时变化，或先析率的变化滞后于评分的变化，则本文的因果次序为伪。顺序预测比相关难得多，一个巧合很难同时满足方向与先后。

条款三（不可逆性，本文最愿意押的一条）。撤走全部范文、模板与讲评中的现成说法之后的十二周内，若先析率自发回落到撤除前基线的 80% 以下，则“接种史不可逆”这条为伪。本文的预测是：不回落，或回落极慢；要回落，必须更换题材域或语域（写一件此前从未被要求写过的事），相当于换一个未被接种的环境。这一条与利托那韦那个工厂的形状完全相同。

条款四（诱导期干预）。随机分配“知道题目后 N 小时内不产出任何可交付文字”的条件。若诱导期延长而先析率不下降，则第十节给出的插手位置为伪。需要同时记录第六节表中“滞留”格的出现率——若延长诱导期只是把未析变成滞留，那也算这条不成立。

条款五（节律优于总量）。总输入量严格相等，一组为持续式（每日投放），一组为脉冲式（隔日集中投放且投放后有明确的不产出间隔）。若两组的先析率无差异，则“节律优于总量”这条为伪，本文将退回到一个只关于总量的说法，而那个说法已被第二节的第二套占满。

条款六（最便宜的一条，可当场执行）。取任意一份现行的作文评价记录格式——教师批改记录、校本评价量表、区级抽测编码表、作文教学平台的数据字段。若其中已存在“逐句来源”或等价字段（要求按句标注该句此前是否已成句），则本文第三层主张为伪。这条所需的数据在多数学校的既有记录里就有，一个人一周内可以查完二十所学校。本文在第七节已经查过一处（高考阅卷细则），结果是只有整篇层面的二值开关，没有逐句刻度。

条款七（语料条款，本文已部分执行）。在一份带时间戳与作者身份的写作语料中，若同一作者的先析率随写作次数单调下降，则“自我加固”这条为伪。本文第九节的实践没能执行这一条，因为公开语料不含时间戳与作者身份——这个字段的缺失本身就是第三层主张的证据，也是本条至今无法裁决的原因。

条款八（生成式工具条款）。使用大语言模型辅助的写作，其先析率应接近上界且与评分保持正交。若模型辅助显著降低先析率（即模型帮助写作者产出了更多此前未以成句形态存在过的文字），则本文对“接种”的界定为伪——届时应当承认，现成形态的可得性与先析之间没有本文所说的那种关系。

若本文被证伪，能学到什么。若条款一倒了，说明写作困难确实主要是动机与制度问题，那么资源应当全部投向评价制度改革，而不是投向教学过程的节律。若条款三倒了（接种可逆），那么本文最不安的那个推论——某些孩子已经不可能写出某类文字——就不成立，这是一个好消息，且它意味着现行的大量投入并没有造成不可逆的损失。这两种情形下的干预方向截然不同，因此这几条值得真去做。

十八、反噬与伦理

反噬预言。先析率一旦进入考核，最省事的达标方式不是降低先析，而是把先析物做同义改写：换词、换句式、拆句、调换语序，使八字串层面的重复消失，而结构与来源分毫未变。这会产出一种形式上不同、实质上相同的伪自发行，并且训练出一种比现在更难被清点的先析。更糟的是，这种改写能力本身会被当作“语言表达能力提升”来教。

我看不到出口。任何基于文本表面的清点都可以被表面改写击穿；而基于来源的清点需要记录写作者接触过的全部文本，那是不可接受的监控。本文没有第三条路。这一条不是待办事项，是本文自己造出来的一个漏洞。

伦理三条（凡产物是一个可拿去考核人的读数，必须写死）。

其一，这个读数指的是位置，不是人。先析率描述的是一次产出的结构，不是写作者的品性、努力程度或诚信状况。它不得作为学术不端的认定依据，不得进入学生的品行评价，不得用于分班、评优或任何对个人不利的安排的单独依据。

其二，不得为了把这个数做好看，而在需要可靠交付的场合强制降低先析。公文、法律文书、医嘱、安全规程的高先析率是安全裕度的一部分。在这些场合以“提高原创性”为名要求改写既有表述，是拿安全换指标。

其三，已经按这个读数做出不利安排的，必须公开编码规则并给出复核通道。具体包括：公布证据源版本号（用哪一批范文、哪一份素材本作为“此前已存在”的判据）、公布前置筛的阈值（ $\phi \geq 0.5$ 剔除）、以及允许被评价者对逐句判定提出异议并要求人工复核。缺少前置筛的公布尤其危险：它会把整篇改写型抄袭者与高度模板化的诚实写作者混为一类。

十九、本文自己落在哪里

用本文的判据检查本文自己，以及检查提出这道题的那个指令。

本文的先析率无法清点。它的成句史包括两样不可查的东西：一是生成它的模型所接触过的全部文本，二是本栏此前已发的一百余篇同类文章。按第八节的编码规则，“此前已存在的成句形态”的证据源必须在测量前写死；对本文而言，这个证据源在原则上不可枚举。本文因此不能测自己。

但可以定位。按第六节那张表，本文落在先析相区，而且落得很深：接种密度极高（三家的材料、十位邻居的概念、本栏此前的全部篇目都在环境里），诱导期极短（从选定三门学科到成文，间隔以小时计）。它的三条签名齐全：产

出快且整齐（签名一）；没有任何环节会因为它是先析物而报错（签名二）；它一旦发表，它的读数名、它的四格表、它的那句问话都会成为下一篇的现成形态（签名三）。

这个定位有三个具体后果，必须说出来：

后果一，本文不能主张自己是自发表。它能主张的只有一件事：它给出了一条自己也过不了的判据。一个判据的价值不在于提出者过得了它。

后果二，本文可以被更省力的相取代，而且取代之后不可逆。若某个更简单的读数出现——比如某个可以由现成工具一键计算、且与分数相关性更高的指标——那个读数会成为这个环境里更稳定的相。此后本文这一相将不再有人重制，正如那个工厂再也做不出 I 型。本文对此没有防御手段，这正是本文自己的判断所预言的。

后果三，提出这道题的那个指令本身也在被检验。“从一百门学科里选三门去撞一个难题”这个做法，是一种高接种密度下的写作：一百门学科的清单、若干条选源规则、一套成文格式，全部是现成形态。按本文的判断，这套做法产出的东西会越来越快、越来越整齐，而它关掉的那一相——一个不按这套做法也会想到的判断——不会出现在任何一份记录上。本文没有办法证明自己不是这种情况。能做的只有一件：把这句话写在这里，让它成为后来者检验这条产线的入口。

二十、三个洞

本文解释不了下面三件事，如实交出。

洞一：为什么有的人不被接种。存在一批写作者，长期处在极高的接种密度里（编辑、语文教师、职业写手），先析率却很低。本文没有解释这件事。可能的方向是他们的原料补充速率远高于常人，也可能是他们的“证据源”结构不同——但这两条本文都拿不出材料。

洞二：诱导期的长度取决于什么。本文只说 τ^* 随接种密度上升而非线性变长，没说清这条曲线的形状，也没说清它对不同的人是否是同一条。若这条曲线在个体之间差异极大，第十节的处方就无法给出可执行的时长，只能给出一个方向。

洞三：口头与书面之间的关系。第十三节其一那条判定形式假设“口头讲一遍”也是一次先析。这个假设本文没有验证过。若口头产出与书面产出共用同一批原料，那么整个第十节的次序要重写，因为最常见的写前活动（讨论、口头汇报、师生问答）就不构成接种，本文的干预范围会大幅缩小。

二十一、一个写死日期的赌注

本文押的是第三层主张，因为它最稳，且它的裁决不依赖前两层。

赌注：截至 2029 年 8 月 31 日，不会出现任何一份面向中学作文的评价标准或教学记录格式，正式设立“逐句来源”或等价字段

——即要求写作者或评阅者按句标注该句在这一次写作之前是否已以成句形态存在——并在至少一个地级市范围内实际执行满一个完整学年。

什么不算命中，逐条写死：

1. 只要求“不得套作、不得抄袭”的条款不算。那是整篇层面的二值开关，第九节的数据已经证明它抓不到本文说的那一类。
2. 只做整篇查重的系统不算，无论它的阈值多细。逐句标注与整篇比对是两件事。
3. 只在研究项目、课题实验或个别学校试点、未进入常规教学记录格式的不算。
4. 由机构自行声明“我们已关注写作来源”而没有逐句编码规则与可复核的判定标准的不算。
5. 只针对他人文本、不把作者自己此前写过的内容计入的不算——第七节场合三已经指出，这个方向的刻度早就存在，而它恰好漏掉了本文关心的那一半。

若在此日期之前上述条件被满足，本文第三层主张判负。若判负，本文的实际收获比赢更大：那意味着这件事已经变得可观测，而本文全部论证的目的就是让它可观测。

参考文献

- Bauer, J., Spanton, S., Henry, R., Quick, J., Dziki, W., Porter, W., & Morris, J. (2001). Ritonavir: An extraordinary example of conformational polymorphism. *Pharmaceutical Research*, 18(6), 859–866. doi:10.1023/A:1011052932607
- Belchetz, P. E., Plant, T. M., Nakai, Y., Keogh, E. J., & Knobil, E. (1978). Hypophysial responses to continuous and intermittent delivery of hypothalamic gonadotropin-releasing hormone. *Science*, 202(4368), 631–633. doi:10.1126/science.100883
- Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Chemburkar, S. R., Bauer, J., Deming, K., Spiwek, H., Patel, K., Morris, J., ... McFarland, K. (2000). Dealing with the impact of ritonavir polymorphs on the late stages of bulk drug process development. *Organic Process Research & Development*, 4(5), 413–417. doi:10.1021/op000023y

Dunker, K. (1945). On problem-solving. *Psychological Monographs*, 58(5), whole no. 270.

Dunitz, J. D., & Bernstein, J. (1995). Disappearing polymorphs. *Accounts of Chemical Research*, 28(4), 193–200. doi:10.1021/ar00052a005

Echt, D. S., Liebson, P. R., Mitchell, L. B., Peters, R. W., Obias-Manno, D., Barker, A. H., ... Greene, H. L. (1991). Mortality and morbidity in patients receiving encainide, flecainide, or placebo: The Cardiac Arrhythmia Suppression Trial. *New England Journal of Medicine*, 324(12), 781–788.

Elbow, P. (1973). *Writing Without Teachers*. New York: Oxford University Press.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing What to Write: Conceptual Processes in Text Production* (pp. 139–160). Amsterdam: Amsterdam University Press.

Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. Berkeley: University of California Press.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, 14(S2), 95–112.

Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319–338.

doi:10.1146/annurev.so.14.080188.001535

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, 54(6), whole no. 248.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87.

Merton, R. K. (1940). Bureaucratic structure and personality. *Social Forces*, 18(4), 560–568.

Ostwald, W. (1897). Studien über die Bildung und Umwandlung fester Körper. *Zeitschrift für Physikalische Chemie*, 22, 289–330.

Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication*, 16(2), 106–112.

Rose, M. (1984). *Writer's Block: The Cognitive Dimension*. Carbondale: Southern Illinois University Press.

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631(8022), 755–759. doi:10.1038/s41586-024-07566-y

Song, W., Zhang, K., Fu, R., Liu, L., Liu, T., & Cheng, M. (2020). Multi-stage pre-training for automated Chinese essay scoring. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6723–6733).

doi:10.18653/v1/2020.emnlp-main.546

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

数据来源：cnunlp/Chinese-Essay-Dataset-For-Pre-Training (GitHub 公开仓库，随 Song et al., 2020 发布)，本文使用其 pretrain_essays.json 的前 40,000 行，剔除后 N = 39,672。第七节关于阅卷细则的清点，检索日期

2026 年 8 月 17 日，检索词“高考作文 评分标准 套作 宿构”，取得的是各地
阅卷细则的公开转载文本，未取得考试机构的正式印发件。

版本与字数：初稿，正文 20,303 汉字（不计参考文献与声明）。第九节的统计在成文前完成，预注册文本先于统计写就并存档，事后未作修改。

人机分工声明：选题、三门学科的选定、成文方向与全部判断的取舍由作者决定；文献核对、语料下载与统计计算、初稿撰写由人工智能助理完成；第九节的预注册文本在运行统计之前写定；本文对其中的全部错误负责。文中未标注出处的经验描述（第一节三个场景）为教研场景的概括性描述，不作为证据使用，仅用于说明问题的形状。

编者补节 · 第二章

一 · 与第三章的动作重叠：先切一句

本章第十节的第四步（先析物写回环境，成为下一次的现成形态）与第三章的第四步（产物写回环境）是同一个动作的两种读法。两章互不引用，编者在此切开。

- 本章读它消耗了什么。写回环境的后果是接种密度上升，下一轮里自发相的门槛因此更高——它抢走的是原料与成形条件。
- 第三章读它改变了什么分布。写回环境的后果是这段文字取得了新的功能位置，随后可行操作的分布因此改变——它改变的是下一步能做什么。

判定形式。取一段刚刚成形的合格文字，做两件事：其一，把它从写作者的可见范围内移走（不再作为环境中的现成形态），其二，把它保留可见但改变其工作位置。若先析率只对前者敏感、端位效应只对后者敏感，两章各自成立；若两个读数对同一种干预同等敏感，二者需要合并。

二 · 需要补的一位：上游那笔账

本章第十三节的十位占位者里，已经正面处理了「写作本身构成知识」这一支，并给出了一条相当锋利的分离线——那一支把提纲的抑制作用归于表述方式（预先的规划取代了构成过程），因此换一种规划方式即可缓解；本章把它归于存量消耗与接种，因此换规划方式无用，只有不产出可交付文字才有用。这条分离线是本卷所有分界里最能被一次课堂实验裁决的一条。

但本章没有处理这一支的上游：已写出的文本（the text produced so far）进入任务环境这一构念（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*, Erlbaum），以及围绕它的重读

—生成研究 (Pianko, S. 1979 关于写作者回看已写内容以决定下一步写什么; Leijten, M., & Van Waes, L. 2013, *Keystroke logging in writing research*, Written Communication, 30(3), 358–392) 。

分离线。那一支说已写文本会被回看、会被用来决定下一步; 本章说被回看的那一部分同时在消耗使别的部分得以成形的条件。前者是资源, 后者是代价。

判定形式。若把重读量、重读位置与重读时长全部纳入模型后, 先析率不再有额外解释力, 本章应并回重读研究; 若先析率在重读量受控之后仍能预测后续的自发相产出, 本章成立。

三·编者提示: 本章最硬的一处与最空的一处

最硬的一处是第九节那个意外发现: 语料中有百分之五点二的篇目, 其每一个八字连续串都能在别处找到, 而逐字完全相同者只有百分之零点二二——前者是后者的二十四倍。这批文本在现行的整篇查重与套作条款下是「原创的」, 在本章的读数下接近上界。这一个数据点比本章任何一段论证都更能说明它为什么值得存在。

最空的一处是作者自己在同一节末尾写明了的: 「存量消耗」这个机制本身没有独立读数。先析率测的是产物的构成, 不是存量, 更不是存量的耗减。本章全部因果主张都靠尚未执行的预测撑着。读者应当把本章读成一个读数加一套待检验的因果假说, 而不是一套已经成立的机制。

编者另指出处正文未标的连带后果: 既然存量不可测, 本章第十节那条处方(诱导期内不产出任何可交付文字)目前无法给出时长——作者自己在第二十节把这一点列为「洞二」。这条处方在落地之前, 只有方向, 没有刻度。

第三章

作文如何发生

——局部终点转为下一轮条件的端位转换

作文如何发生

——局部终点转为下一轮条件的“端位转换”理论

摘要

作文研究已经把规划、转译、修订、递归、问题解决、修辞情境和社会文化条件纳入“过程”，但“有过程”仍不等于解释了过程怎样不断产生下一步。本文把发育生物学的形态发生、交通流理论与机制设计放在同一问题上，提出“端位转换”假说：作文的发生不是预成结构的展开，不是思想流的连续通行，也不是规则场对行为的单向塑造，而是局部形态、生成路径与约束场在连续轮次中互换功能角色；一次的局部终点转成下一轮的条件、路径或规则。本文以“同一段已稳定文字保持可见，但从正文可编辑嵌入态移到不可编辑侧栏后，后续120秒操作序列分布是否改变”为零情态词判据，并给出稳定性×因果再入的二维辨别格。利用公开 KUPA-KEYS 1,006 篇作文元数据预注册真跑后，简单“活动强度存在中间最优”的代理假说未获支持，迫使理论放弃事件数量代理，转向状态与角色转换。文章进一步与递归写作、知识转换、反馈控制、行动中反思、认识性行动等逐一给出判决性分离，并提出随机干预实验与2028年底可裁决赌注。

关键词：作文发生；端位转换；形态发生；交通流；机制设计；键盘日志

Abstract

Writing research has long treated composing as recursive, goal-directed, situated, and socially mediated, yet describing a process does not by itself explain how one local outcome becomes the cause of the next move. This article develops an endpoint-role conversion account by bringing morphogenesis, traffic-flow theory, and mechanism design into a single mechanistic problem. The central claim is triple-negative: composition does not occur merely by unfolding a preformed structure, by continuously transmitting ideas through a channel, or by letting a fixed rule environment shape behavior. It occurs when local forms, generation routes, and constraint fields repeatedly exchange functional roles across successive rounds, so that a temporary endpoint becomes a condition, route, or rule for what follows. The proposed zero-modal test keeps a stabilized text segment fully visible while moving it from

an editable position in the main document to a non-editable side panel, and asks whether the distribution of operations over the next 120 seconds changes. A preregistered proxy test using metadata from 1,006 public KUPA-KEYS essays did not support a simple inverted-U relation between process activity and human scores; the adverse result rules out event counts as a proxy for the proposed mechanism and shifts measurement toward state and role transitions. The theory is separated from recursive process models, knowledge transforming, feedback control, reflection-in-action, epistemic action, and distributed cognition by discriminating predictions, and it ends with a preregistrable randomized intervention and a dated falsification bet.

Keywords: composition process; endpoint-role conversion; morphogenesis; traffic flow; mechanism design; keystroke logging

一、问题的重启：我们已经知道“写作是过程”，却仍不知道过程怎样长出下一步

如果问一个学生“作文是怎么写出来的”，最常见的回答是一条步骤：审题、立意、列提纲、写开头、展开、修改。现代写作研究当然早已比这复杂得多。Flower 与 Hayes 把写作描述成目标导向而且递归的认知过程，规划、文本生成与检查可以嵌套和往返；后来的研究继续加入工作记忆、动机、修辞情境、社会文化环境、身份、反馈与工具。于是，“作文不是结果而是过程”已经不是一个新判断，而是研究常识。[10][24]

但“过程”这个词有一个隐蔽的宽度。它可以告诉我们一连串事件发生过，却不自动告诉我们，为什么刚才这一步会改变下一步可做的事情。完整键盘日志能够记录某人在第几秒输入、删除、停顿、移动光标；过程挖掘能够从事件日志还原常见路径；写作者自己也能回忆“先想了这个，再删了那个”。这些材料都很重要，却仍可能停在事件序列。一个事件 B 出现在 A 之后，不等于 A 使 B 变得可能，更不等于 A 在 B 出现以前改变了自己的功能角色。

因此，“作文如何发生”需要比“有哪些步骤”“哪些步骤会递归”更窄的一问：在一次真实写作中，某个已经出现的局部结果，什么时候不再只是结果，而开始成为下一轮生成的条件？如果这个转换不存在，写作也可以持续：一个人可以按预先写好的提纲连续誊写，段落越来越完整，字符越来越多，过程日志也非常丰富，却几乎没有任何局部产物反过来改写后面的行动。

相反，一个并不成熟的句子有时会突然暴露问题，使作者改变证据、受众形象、论证路径乃至题目本身。两种写作都有“过程”，但它们的发生结构不同。

本文把这个差异作为独立研究对象。它不试图否定既有过程理论，也不把一切好写作都归结为一个新名词。更谨慎的目标是提出一种可以被随机干预检验的机制：局部文本、当前路径和约束场之间，是否存在“本轮终点转为下一轮条件”的功能角色变换。若这个对象最终能被递归、知识转换、反馈或外部记忆无剩余地解释，本文就应当撤回；若在控制这些近邻后仍能观察到独特效应，它才拥有独立理论地位。

还可以从“同一步骤、不同发生”反推问题。两名学生都先列提纲再写正文，日志上都显示规划—输入—修订。甲的提纲只是一个暂时假设，写到第三段时反例迫使他改写中心；乙的提纲是一张不可撤回的施工图，后文只负责填空。若研究把“列提纲”编码成同一个 planning event，就会错过两者最关键的差别：甲的局部结果具有重新分配后续行动权的能力，乙的结果只完成既定任务。这说明发生机制不能只靠阶段名识别。

这个问题也改变了“写作能力”的研究姿势。能力通常被当作写作者拥有的稳定资源，而端位问题首先是事件级的：同一个人在不同任务里可以切换发生方式。写熟悉公文时，他可能低端位、高效率；研究陌生理论时，他可能高端位、低短期效率。若把端位效应人格化成“这个人会不会写”，会把机制变量误当主体属性。因此本文后文所有预测尽量使用同一写作者内的随机比较，而不是用“高手与新手”直接做本体区分。

二、为什么完整步骤仍然可能不是发生机制

学校中的作文步骤天然诱发一种流水线想象：每一步有固定输入与输出，前一步完成，后一步接手。它的好处是可教、可查、可评分；它的危险是把“顺序”当成“发生”。同一个顺序可以由完全不同的机制产生。一个熟练写作者可能在动笔前已经把整篇文章想得很清楚，实际输入只是执行；另一个人可能边写边发现论证中的冲突，表面同样经历“起草—修订”，但修订正在重组问题。事件标签相同，因果结构并不相同。

这一区分在今天尤其重要。生成式人工智能可以一次给出完整段落，也可以作为局部提问工具进入写作者的探索。如果研究只记录“谁输入了多少字符”“修改了几次”“调用了多少次 AI”，就会把活动量当作发生量。一个学生可能对 AI 段落做了二十次表面修改，却仍沿着外部给定的结构前进；另

一个学生只保留一句自己写出的矛盾，却因为这句话改变了后续全部检索和结构。次数更少的一方，反而可能发生了更深的内部转向。

写作研究早已知道步骤是递归的，这正是本文必须面对的第一位近邻。递归说明过程可以回去，但“回去”仍是一种运动描述。我们进一步关心：回去以后，原来的局部产物还是原来的对象吗？它可能从“要完成的句子”变成“限制下一段的证据要求”，从“刚找到的论证路径”变成“迫使题目重写的条件”，从“外部评分规则”变成写作者通过实践重新定义的内部规则。发生机制的关键不只是路径有没有回头，而是对象在回头时换了什么角色。

如果这一区分成立，那么过程研究的观测单位也需要变化。事件日志中的“插入、删除、暂停、移动”仍然有用，但它们只是动作层。机制层还需要一个字段：某个局部状态在当前轮次承担什么功能，它在下一轮是否换位。正是这一缺项，把本文带到三个看似与作文很远、但都长期研究“结果—路径—条件”关系的学科。

从因果识别角度看，步骤日志还有一个更隐蔽的问题：它通常是“观察已发生路径”，而机制要求比较“如果局部结果不再占据同一工作位置，后续会怎样”。只有加入这种反事实，才能把伴随事件和真正改变可行域的状态分开。例如作者写完一句定义后停顿十秒，停顿可能是因为思考定义的后果，也可能只是喝水。单凭时序，我们不知道定义是否改变了后续。把定义原文保持可见却移出主文档，才第一次改变了它的操作角色而尽量不改变语义信息。

因此，本文不把更精细的日志本身当作答案。传感器分辨率无限提高，仍可能只得到更高清的相关。真正的升级是从“记录更多”转向“改变一个理论上承重的变量”。这也是为什么后文把侧栏干预放在中心：它不是最自然的教学操作，却是一个能把外部记忆、递归次数和局部内容尽量固定、只动工作角色的实验刀口。

三、第一条约束：形态可以后生——来自形态发生学

形态发生研究的是一个极基础的问题：一个有组织的整体形态怎样从局部相互作用中出现。Turing 1952 年的经典论文用反应—扩散说明，局部化学物质的反应与扩散足以产生空间模式。[1] 这条传统后来被机械力、细胞黏附、组织几何、形态素反馈等大量工作扩展；现代形态发生并不把胚胎看成照着一张完整蓝图执行的机器，而关注局部规则、物质流、力学边界和信号怎样共同把形态“做出来”。[2][3]

迁入作文问题，最先被削弱的是“完整结构必须先”。一篇文章当然可能先有详细提纲再执行，但这不是作文发生的普遍必要条件。局部判断、句

法选择、例证之间的相互作用，可以让一个此前没有明确存在的结构逐渐成形。写作者常有这样的经验：最初只知道一个不满意的地方，写两段以后才发现真正的中心；原先以为是三个并列理由，写到第二个时却发现它其实是第一个的反例，于是整篇结构改变。这里的结构不是被“表达”出来，而是在表达活动中才获得自己的边界。

然而，把作文说成“像胚胎一样自组织”仍然只是类比。形态发生真正带来的约束在于：局部产物一旦形成，会改变随后发生的条件。组织的几何、张力和受体状态会反馈到信号与运动上；形态不是到站后被摆在橱窗里的成品。[2][3] 如果把这个事实转回写作，一个已经出现的段落也可能不是“已完成部分”，而是新的力学边界：它压缩某些后续选择，放大另一些矛盾，改变证据需要，甚至改变作者认为自己正在回答的问题。

这条约束同时防止一种浪漫化：自组织不等于好。胚胎发育存在失调和畸形，局部规则可以生成稳定却不合功能的形态。作文亦然。一个结构可以非常自然地“长出来”，但论证仍可能错误。本文因此只讨论发生方式，不把后生结构自动等同于高质量。

形态发生还提醒我们区分“预先约束”与“预先图纸”。没有完整蓝图不等于没有边界。胚胎发育受到材料属性、几何、基因调控和环境条件约束；作文同样不可能在真空中自组织。题目、语言知识、体裁经验和事实证据都限制可形成的形态。端位理论因此不支持“随便写就会自己长成好文章”，它只取消“最终结构必须在动笔前以同一形式存在”的必要性。

这个区别在教学上很重要。很多“列提纲”争论把问题做成二选一：要么先规划，要么自由生成。形态发生视角提供第三种可能——初始约束可以很强，但最终形态仍在局部互动中出现。教师可以要求证据标准和问题边界，却不必把段落序列冻结；写作者也可以有初始中心，却允许局部材料把它改写。预测上，这种“有边界但可变形”的任务不必比完全自由任务有更多修改次数，却应更容易出现局部产物对下一轮路径的可测再入。

四、第二条约束：活动数量不是流动状态——来自交通流理论

交通流理论提供了另一种困难。Richards 1956 年把车辆近似为连续流体密度，并用速度—密度关系研究公路上的冲击波；与 Lighthill-Whitham 的工作共同构成经典运动波传统。[4][5] 后来的交通研究不断提醒我们：路上“有很多车”不是一个足够的状态描述。相同车辆数可以处于自由流、临界状态或拥堵；局部瓶颈失稳后，排队可以向上游回溢，当前拥堵还可能降低后续实际放行能力。

这对写作过程研究有直接的警告。击键很多、停顿很多、删除很多、修改很多，都不能单独被叫作“发生充分”。活动量就像车数。真正决定路径的是状态关系：局部材料密度、可用语言化容量、当前规则、光标与文档结构、已写文本带来的约束，以及它们之间是否出现瓶颈。一个人可以高速连续输入却只是抄写；也可以停顿很久，因为一个已经写出的句子正在迫使他重做问题。若只把速度慢叫“堵”，会把思考与拥堵混为一谈。

交通流更有价值的一点是“结果会改写容量”。拥堵不是一个结束标签；队列回溢会占据上游空间，容量下降会让后来进入的车辆面对不同道路。放到作文中，这意味着一个局部结果的意义不在于它是否“完成”，而在于它形成后有没有改变下一轮的可行操作集合。写出一句承诺“本文将证明 X”以后，后面可合法使用的证据与结构已经改变；写出一个无法解释的反例以后，旧路径的容量可能突然下降，作者不得不绕行。

这同时给本文加上一道反向约束：状态转换越多不代表作文越好。交通中的 stop-and-go 正是高频变化却低效的例子。一个写作者不断推翻自己、频繁换题、持续重构，也可能只是振荡。后文提出的端位效应因此只是机制读数，不是绩效分数，更不能直接成为学生考核指标。

交通理论还帮助区分“局部慢”与“系统堵”。高速公路上一次减速并不等于拥堵；关键是扰动是否传播、队列是否回溢、有效容量是否改变。作文里一个长停顿同样不能直接叫 writer's block。真正接近机制的信号是：一个局部困难是否改变了上游/下游的行动分布，例如反复回到前文、资料检索突然增加、后续追加显著减少，或者既定段落被迫重开。换言之，状态要由关系结构而不是单一时长定义。

这一点也解释了为什么 KUPA 的不利结果不是偶然插曲。我们最初用“每词事件数相对抄写基线”近似过程强度，等于用车数近似交通状态。统计不支持倒 U 型之后，理论被迫承认代理变量错层。以后即使完整日志显示某些高分作者修订更多，也不能据此反推端位机制；必须识别“某个局部结果出现前后，路径的转移概率是否改变”，最好再用随机角色干预检验这种改变是否具有因果性。

五、第三约束：规则场也会被结果改写——来自机制设计

机制设计把经济学常见的问题方向反过来：不是在既定规则下预测人会怎么行动，而是先问希望实现什么结果，再寻找什么规则能在信息不完整、参与者有自身策略的条件下把行为导向该结果。2007 年经济学奖把 Hurwicz、Maskin 和 Myerson 的工作概括为机制设计理论的基础。[6] 之

后鲁棒机制设计、动态机制设计与市场设计继续处理信息不确定、时间变化和制度外部环境。[7][8]

它对作文最有力的提醒是：所谓“写作环境”不是装饰性的背景。题目、受众、评分标准、可用资料、工具权限、时间、教师反馈方式，都会改变局部选择的收益与成本。如果只告诉学生“下一步写什么”，教育者是在直接控制动作；如果改变规则，让某些问题必须被证据回应、让反例拥有进入结构的权利、让草稿允许撤回，写作者可能自行产生不同的路径。

但成熟的市场设计又反过来破坏“规则是最终答案”的幻想。Roth 把 market design and maintenance 连在一起讨论，理由很直白：市场从不孤立运行，参与者会适应规则，环境会变化，设计出来的机制需要维护。[9] 一个机制成功运行以后，会创造新的行为模式与外部条件；规则的结果反过来成为下次规则设计的输入。规则场因此也不是作文发生链上最后一个不动的背景。

把这三条约束放在一起，冲突就出现了：形态发生不要求完整目标先在；机制设计从目标反推规则；交通流则告诉我们，即便目标和规则都“正确”，路径仍可能因容量与传播动力学失败。任何一家都不能单独结算作文如何发生。更重要的是，三家自己的成熟理论都承认各自所谓的结果会重新进入下一轮。这个共同事实将推翻一个比“线性写作”更隐蔽的前提。

机制设计的另一个启发是把“写作策略”与“写作制度”分开。学生被要求“多修改”是一条动作建议；系统只在提交前允许一次修改，是规则；评分只奖励格式一致，会改变所有局部选择的收益。相同写作者在不同规则下可能呈现完全不同的端位结构。因此，若研究只观察个人认知而忽略界面权限、截止时间、评分方式和 AI 工具的可调用规则，就可能把制度造成的固定路径误诊成个人不会生成。

但机制设计也迫使理论面对策略性反应。假如实验参与者知道研究者奖励“更多端位转换”，他可以故意把段落来回搬动、频繁重写目标，制造高 E_T 的表面痕迹。真正的机制读数必须尽量让参与者不知道哪一种转换被奖励，或者让研究指标只在后台用于科学检验而不进入即时反馈。这一要求不是伦理附录，而是理论内部约束：被测规则会改变行为，任何公开考核都可能把测量对象重新设计掉。

六、被三种理论共同隐藏的前提：终点被当作一次性终点

三种理论表面上争的是三件不同的东西。形态发生把整体形态视为需要解释的结果；交通流把实际流动状态与路径视为结果；机制设计把能够实施目

标的规则结构视为设计产物。若把它们用来解释作文，最自然的做法也是各取一段：先设计环境，再让思想流动，最后形成文本。这样看似跨学科，实际上仍保留了一条流水线。

流水线背后的共同前提是：一次生成轮次总有一个可以固定归属的终点。某一样一旦成为结果，另外两样就退成手段；结果被保存、评价或提交，下一轮另起炉灶。这一前提甚至不要求流程线性。流程可以递归，只要每次往返仍把“阶段”和“终点”当作稳定角色即可。

真正破坏这个前提的材料恰恰来自三家自身。形态发生已经知道形成的组织几何和力学边界会改变后续信号与运动；交通流已经知道拥堵结果会改变有效容量和未来通行条件；机制设计已经知道规则运行会改变参与者行为与市场环境，设计必须维护。三家都在自己的成熟研究中经历同一件事：本轮终点没有退出因果链。

因此要问的不是“作文到底以结构、路径还是规则为中心”，而是“这三种角色何时换位”。如果刚写出的段落只作为成品累积，它是结果；如果这段文字随后限制证据选择，它变成条件；如果它诱发新的检索、删改与结构重排，它成为路径入口；如果它迫使作者重写题目或评价标准，它甚至参与改造规则场。一个对象没有变化，变化的是它在发生链里的位置。

“固定终点”之所以难被看见，是因为我们的语言本身倾向给对象固定名词：这是“段落”，这是“计划”，这是“环境”。名词一旦固定，角色也被悄悄固定。可在真实发生中，一段文字首先是作者想得到的产物，随后成为必须遵守的承诺，再后来可能变成需要推翻的旧规则；一个提纲首先是环境约束，之后变成被新证据否定的历史痕迹。若编码系统坚持给它一个永久标签，角色转换只能以“异常”出现。

端位理论因此不是再增加一个阶段，而是要求研究把“角色”做成时间变量。它和阶段模型的关系类似于“位置”与“身份”的区别：同一对象可以移动位置而仍是同一对象。研究应记录的不只是文本单元 t 在时刻 n 存在，还要记录 t 在 n 、 $n+1$ 、 $n+2$ 中分别承担什么功能。这个要求若最终无法带来增量预测，就没有必要；但若可以，它会改变写作过程数据的最小表结构。

七、端位转换：作文发生的最小机制假说

本文把上述角色变化称为“端位转换”。“端”不是全文的最终结尾，而是某一局部生成轮次的临时结算点；“位”是这个局部状态在下一轮承担的功能角色。端位转换发生时，第 n 轮中作为结果出现的局部状态，在第 $n+1$

轮转成条件、路径入口或规则约束，并使后续可行操作的分布发生可检测变化。

于是本文的承重判断可以写成一个三重否定：作文的发生不是预成结构的展开，不是思想流的连续通行，也不是规则场对行为的单向塑造，而是局部形态、生成路径与约束场在连续轮次中发生端位转换。这里的“不是”都不是全盘否定。作文完全可以使用预先结构，也必须有信息流，也受到规则影响；本文只否定它们任何一个可以单独充当充分解释。

端位转换与“反馈”最容易混淆。反馈的最低形式是输出信息返回系统；端位转换则要求返回后的对象改变功能角色。一个拼写检查器把错误信号反馈给作者，作者改正拼写，这当然有反馈，但原句未必改变后续论证路径。相反，一个反例句子即使没有显式评分信号，也可能让作者放弃整条论证路线。前者反馈强而端位转换弱，后者未必存在固定参照的反馈，却可能有强端位转换。

它与“递归”也不同。递归描述阶段可以重新进入：规划之后可回到规划，修订后可继续生成。端位转换问的是同一个局部产物是否换角色。一个人可以机械地每段都“写—检查—改”十遍，递归很多，却始终执行同一外部提纲；另一个人只在一个关键句后改变一次证据和中心，递归次数不多，却发生了明显角色转换。

这个定义还故意不把意识、反思和学习写成必要条件。写作者可能清楚地“这一句让我换了思路”，也可能只是光标突然回到前文、长停顿后打开资料、删掉原提纲而本人说不清原因。只要可控制的工作角色改变能稳定地改变后续行动分布，就值得作为机制研究。反之，即使写作者主观觉得“我想通了”，若行为路径不受局部产物位置影响，端位机制也没有得到支持。

为了避免概念无限扩张，端位转换还有三个排除条件。第一，纯信息回看不算：作者重读上一段但后续行动分布未改变，只是访问，不是转换。第二，纯表面编辑不算：把“然而”改成“但是”但后续路径不变，不因为发生了修订就算转换。第三，研究者事后解释不算：只有当角色差异能够通过干预、时序或预注册编码产生可重复读数，才进入经验层。

反过来，端位转换不要求局部文本一定被修改。一个句子可以一字不动，却因为它提出的承诺使作者停止原计划、打开新资料、重写后面三段；它的文本形态稳定，功能角色却已经改变。这一点把“修订量”与“再入强度”彻底分开，也使理论能处理一些最重要但在版本 diff 中几乎不可见的发生：文本不变，路径变了。

八、稳定不等于发生：一个二维辨别格

要让新概念不退化成一句漂亮话，需要第二根轴。本文选择“局部文本是否稳定”和“它是否因果再入下一轮”两轴。稳定性回答表面形态有没有暂时落定；因果再入回答这个形态是否改变后续操作分布。两轴交叉后，最重要的不是理想格，而是一个容易被误判为成功的格：稳定但不再入。

所谓“惰性成形”，不是低质量文本，也不是作弊。它指局部文本已经稳定、甚至越来越流畅，但它没有改变下一轮路径，后续主要由此前外部计划、模板或既定规则驱动。预先完整构思后的誊写就可能属于这一类；一个AI连续按固定大纲补全的段落也可能属于这一类。它们可以很优秀，但其发生方式与边写边重构不同。

这个格之所以值得单独识别，是因为传统成品评价很难发现它。第一，它在短期指标上经常表现为改善：字数增长、语言流畅、结构完整。第二，它的失效没有报警声：文本仍在继续，写作者也未必感到卡住。第三，它会被制度自我加固：如果评分只看成品，最快的策略就是让局部产物尽快封闭，减少重新打开它们的机会。于是越“高效”的文本生产，越可能把发生链压成执行链。

与之相对，“端位转换”格要求局部文本既形成了暂时稳定，又重新进入下一轮。稳定是必要的，因为一个完全游离的念头尚未形成可识别的局部终点；再入也是必要的，因为没有再入，它只是沉积。这个二维格使“写得多”“改得多”“写得好”都不能直接等同于发生。

“预形改道”这一格也值得保留。草稿碎片、关键词或口头备注可能尚未形成稳定文本，却已经改变后续行动。它说明端位转换并不等于“稳定之后才有认知作用”；二维格只是把稳定局部文本作为核心实验的可控对象。若未来研究能对临时外化做可靠干预，理论可以扩展到未稳定状态，但不能为了覆盖更多现象而先把所有外化都叫端位转换。

“游离流”则避免把探索失败道德化。一个人在材料中漫游、不断试句、尚未形成稳定单元，并不一定低效；他可能处在需要高探索的前期。二维格只是描述当前状态，不能预设每次写作都应快速进入右下角。真正有价值的问题是状态怎样转换、什么条件使某一格长期滞留，以及不同任务对四格的需求是否不同。

	下一轮未产生可检测因果再入	下一轮产生可检测因果再入
部文本尚未稳定	游离流：探索尚未结晶，路径继续漂移	预形改道：临时外化虽未稳定，却已改变后行动
部文本已经稳定	惰性成形：文本稳定增长，但后续仍由既定计	端位转换：局部完成态成为下一轮条件、路

九、怎样测：保持内容可见，只改变它的工作角色

最强竞争解释是外部记忆。任何人都可以说：前文留在正文里当然会影响后文，因为作者看得见；把前文拿走，记忆负担增加，操作自然变化。这一解释如果不被控制，端位转换就会退化成外部认知或认识性行动的一个版本。

因此核心实验不把文字隐藏。设一个局部文本单元已经稳定，例如连续至少 10 个英文词在 60 秒内未被编辑且光标已经离开。系统随机进入两种条件。嵌入态：该段原样保留在主文档，可继续编辑、移动和引用；侧栏态：该段从主文档位置移出，但逐字显示在固定侧栏，正文原位置留下等长占位符。两边看到的语义和字面内容相同，改变的只是它是否仍是主工作区中的可操作对象。

随后记录 120 秒的事件转移：继续向后追加、局部修订、跨段回改、删除簇、远距离光标移动、停顿超过两秒、来源或工具调用。把同一写作者、同类题目、匹配文本长度下的两个条件比较，定义端位效应 E_T 为两种条件后续事件转移分布的 Jensen-Shannon divergence。统计显著性来自参与者内条件标签的随机置换，而不是人为规定“0.2 就算高”。

这个设计的零情态问句很简单：同一段已稳定文字保持可见，但从正文中的可编辑嵌入态移到不可编辑侧栏后，后续 120 秒的操作序列分布是否改变？它可以由日志回答，不需要问“这段文字是否真正影响了作者”“影响是否有意义”。如果两种条件无差异，端位机制在这个任务中没有读数；如果差异稳定存在，还必须继续排除递归阶段切换、知识变化、一般编辑便利性等竞争解释。

侧栏设计仍有残余混淆。例如从正文移到侧栏会改变视觉空间位置、光标距离和编辑便利性。正式实验因此应增加一个“正文内冻结态”：文字保留原位置、仍可见，但暂时不可编辑；再与“侧栏态”和“完整嵌入态”比较。若完整嵌入与正文冻结不同，说明可编辑性本身重要；若正文冻结与侧栏仍不同，说明空间/文档嵌入关系可能重要。理论必须允许这些因素把端位概念继续拆小。

此外， E_T 不是单个参与者一次窗口就能可靠估计的稳定人格分。Jensen-Shannon divergence 在短窗口和稀疏事件下会受采样噪声影响，应该通过多任务、多单元聚合并报告置信区间。研究者不能把“某学生 $E_T=0.18$ ”当作精确能力值，更不能跨平台比较不同事件编码下的绝对值。

单一口径只要求同一研究和同一论证中定义一致，不意味着一个未经校准的全球常模。

十、五种发生签名：同样叫“写作”，内部路线可以不同

第一种是“无转换”。抄写、按完整脚本誊写或高度程序化填充中，局部文本可以不断稳定，但下一步几乎完全由外部源或既有计划决定。KUPA-KEYS 专门设置了 copy-text 基线，这类数据提醒我们：丰富击键并不保证内部角色转换。[23]

第二种是“形态→路径”。Flower-Hayes 模型中，检查和评价可以触发新的规划或生成；现实写作中，一个已经写出的句子暴露歧义后，作者突然跳到前段修改定义，局部形态由结果变成路径入口。[10] 端位理论并不声称发现了“会回改”这件事，而要求进一步证明：相同内容只改变工作区角色，就会改变后续操作分布。

第三种是“路径→规则”。Bereiter、Burtis 与 Scardamalia 的研究显示，较成熟写作者在内容与修辞问题空间之间进行转换，解决一边的问题可能改变另一边的信念与原目标。[14] 这里最值得端位理论借用的不是“知识转换”这个名字，而是一条时序可能性：正在走的路径本身可以迫使规则重写。

第四种是“规则固定导致路径冻结”。Rose 对 writer's block 的经典研究指出，僵硬规则和不灵活计划会扼制语言生成。[11] 端位理论在此给出一个更具体的实验问题：规则相同的写作者中，那些局部产物不能反向改变路径的人是否更容易进入长期停滞？若答案是否定的，端位理论对 writer's block 的解释就没有增量。

第五种是“同形异端位”。在 AI 共写中，两个人可以保留字面完全相同的一段机器生成文本。甲把它当作按既定大纲填好的模块，下一步继续执行外部结构；乙因为其中一个意外论点改变检索、反例与标题。成品局部字符串相同，端位角色不同。这个案例说明，端位研究不能从最终文本表面直接读取，必须进入可干预的生成过程。

这五种签名的目的不是把作者分成五类人，而是让同一写作者在不同时间窗口中可被不同机制解释。一个论文作者上午可能在抄录实验方法，属于近乎无转换；下午写讨论时一个异常结果改变中心论点，出现形态→规则；晚上按期刊格式整理参考文献，又回到低转换执行。若把整篇文章平均成一个“写作风格”，这些阶段性机制会互相抵消。

因此未来研究更适合采用窗口级或单元级模型，再把窗口嵌套在写作者与任务中。这样既能尊重个体差异，也避免把机制人格化。端位理论最希望出

现的证据不是“高水平作者平均 E_T 更高”，而是同一个人在随机改变局部产物工作角色后，下一轮行动分布出现可重复的条件差异。

十一、一次不利的真跑：1,006 篇公开作文迫使理论放弃“数事件”

一套新理论如果只给未来实验开支票，很容易永远正确。本文因此在成文前用公开 KUPA-KEYS 数据做了一次低成本预注册真跑。Velentzas 等人 LREC-COLING 2024 发布了 1,006 名参与者的英文写作数据，每人完成抄写任务和自由作文任务；数据包含文本、逐事件键盘日志与元数据，作文由三位人工评分者给分。[23] 本次运行成功取得公开 META 表。

预注册没有直接声称能测端位转换，因为 META 只有任务级事件数、时间、词数和评分，没有逐事件角色字段。我们只检验一个较弱代理猜想：如果“过程活动强度”太低对应机械执行、太高对应拥堵，控制作文长度、作文时间、抄写速度、母语与题目后，活动强度与人工评分可能存在倒 U 型。活动强度定义为作文每词事件数对个人抄写每词事件数的对数比。事先写死：二次项必须为负、 $p < .05$ ，且峰值落在观测第 5–95 百分位内才算支持。

结果不支持。N=1,006；线性项 $\beta=0.0776$ ， $p=0.8406$ ；二次项 $\beta=+0.5627$ ， $p=0.1888$ ，不但不显著，方向还与预注册预期相反；计算峰值 -0.0690 也落在第 5 百分位 -0.0456 之外。按 process_surplus 四分位，人工均分约为 7.93、7.98、7.99、8.14，没有出现中间区间最好。

这不是“数据没测到所以理论仍对”，而是一处真正的删改。原先最容易写的第三层主张——“事件太少或太多都不好”——被取消。交通流理论反而帮助解释为什么这一代理粗糙：交通状态不能由车辆数量单独确定。同样，写作状态不能由击键、暂停、删除或修订次数直接确定。事件数量可以与字数、熟练度、设备习惯一起变化，却不告诉我们某个局部文本是否从结果转成下一轮条件。

本次运行也暴露了数据结构的缺口。KUPA-KEYS 的完整 TASK-2 日志包含时间戳、事件类型、光标位置和周期性全文快照，理论上可构造更细的稳定单元和时序分析；但本次环境未能稳定取得约 533MB 的完整事件表，因此没有把元数据代理冒充核心机制证据。本文最稳的经验结论反而变成：任务级活动量不足以测端位转换，未来需要逐事件状态与受控角色干预。

这个真跑还带来一个更具体的正面提醒。原数据论文报告作文事件数与评分存在中等正相关，而作文词数与评分相关更强。[23] 如果只看相关，很容易讲出“更活跃的过程带来更好作文”的故事；我们的控制模型与预注册二次

项却没有支持预想的中间最优。两种结果并不冲突，它们共同说明活动量缠绕着长度、熟练度和任务完成方式。任何把单个行为计数升级为机制的做法都需要特别谨慎。

公开数据的价值还在于它使失败可以被复算。本文保存了预注册文本、模型结果与四分位描述，而不是统计显著以后才选择故事。未来若取得完整 TASK-2，首先也应在看结果前冻结“稳定单元”“窗口长度”“事件类别”和匹配策略。否则角色转换这种高维对象极易被研究者自由度制造出来。

十二、与最近的十一种说法分开：必须给出相反预测

第一，递归写作过程。Flower-Hayes 已经说明规划、生成与检查可以递归、嵌套，甚至目标会变化。[10] 本文不以“会回去”自立。决定性分界是：匹配递归次数和阶段切换后，只把同一局部文本从正文嵌入态移到侧栏态，如果后续 120 秒的操作转移分布仍系统改变，端位转换有独立读数；若该干预无效且递归模型已经解释全部差异，本文应退化为过程理论的一种表述。

第二，知识转换。Bereiter 等人讨论内容问题空间与修辞问题空间的互动，解决修辞问题可引起内容、信念或原目标变化。[14] 本文允许端位转换而不发生知识更新。若前后知识测验与主张保持不变，但工作角色干预仍改变操作分布，端位转换成立而知识转换不是充分解释；反之，若控制知识变化后 E_T 消失，则本文应让位。

第三，reflection-in-action。Schön 关注实践者在情境“回话”中反思并重构行动。[18] 端位转换不把可报告反思列为必要条件。如果没有自述评价、没有显性“重新框定”，只改变文本的工作位置仍产生稳定 E_T，本文预测有机制作用，而强反思解释较难单独覆盖；如果所有效应都被可报告反思中介，Schön 更经济。

第四，认识性行动。Kirsh 与 Maglio 把改变外部世界以减少认知工作、揭示信息的动作称为 epistemic action。[19] 写下一行帮助记忆可以是认识性行动，但未必改变下一轮路径。若外化降低记忆负担却不改变后续转移分布，认识性行动存在、端位转换低；这条反例阻止本文吞并一切“写出来帮助思考”。

第五，反馈控制。Wiener 的控制论早已研究输出信息如何返回系统。[15] 如果端位转换只剩“输出又回来”，它没有新意。本文的分离线是固定参照与功能角色：在没有明确误差信号、甚至目标本身被改写的情况下，局部产物仍可从结果转成规则或路径。若所有 E_T 都能由固定参照的误差修正解释，控制论占位成功。

第六，修辞情境。Bitzer 把迫切性、受众和限制放在话语发生之前，情境“召唤”修辞反应。[13] 端位转换允许外部情境不变而内部有效情境被产物改写。若题目、受众、时间和资料全部固定，只改变前文工作角色仍改变后续路径，本文有剩余；若不存在这种内生变化，修辞情境足够。

第七，分布式认知。Hutchins 说明认知过程可以跨越个体、工具与文化活动系统。[20] 端位转换不判断认知边界属于谁。一支团队可以高度分布式地按固定模板生产文本，而每个局部模块从不改写后续规则，此时分布式认知高、端位转换低。

第八，稳定中间形态。Simon 1962 年用复杂系统中的稳定中间形态说明层级结构可以提高复杂对象生成的可行性。[16] 本文第二轴恰恰说明稳定不够。一个稳定段落只要不因果再下一轮，就仍属于惰性成形；如果稳定块只降低成本而不改变路径，Simon 的解释成立而端位转换不成立。

第九，Inputlog 等键盘日志研究。Leijten 与 Van Waes 提供了成熟的击键记录与可视化方法。[21] 它们能精细记录过程，却不自动给出因果角色。端位转换必须通过干预或足够强的准实验识别；“日志里出现修订”不能被当作端位证据。

第十，过程挖掘。van der Aalst 的 process mining 可以从事件日志发现过程模型、做一致性检查。[22] 端位转换与它是方法—机制关系：相同流程图可能对应不同的工作角色干预效应。若事件序列本身已能无剩余预测嵌入/侧栏干预结果，则没有必要另设端位变量。

第十一，同批研究提出的“作文等价类/生成见证”处理的是另一个问题：多个可能文本什么时候具有同一生成身份。[25] 本文问的是一次写作事件内部，局部产物怎样成为下一轮的条件。一个人可以拥有稳定生成见证，却在某次写作中按预成结构快速转写，E_T 很低；也可以在人机共写中产生很高的局部端位效应，却不能据此证明关键生成见证属于人。两者可以独立变化，因此 What 与 How 不互相吞并。

把这些近邻放在一起还有一个好处：它阻止“端位转换”向宏大总理论膨胀。递归写作已经占据阶段回返，知识转换已经占据问题空间互改，Schön 已经占据实践中的反思重构，Kirsh-Maglio 已经占据外部行动改变认知任务，Hutchins 已经占据认知跨人—工具分布。本文若把“写作会反过来改变思考”当贡献，几乎每一位都能取代它。

因此最终剩下的命题被压得很窄：同一局部产物的内容保持可见，只有它在工作区中的功能位置被改变，后续操作序列仍出现差异。这个干预如果失败，近邻理论不需要被推翻，它们反而获得胜利；如果成功，端位理论也只得

到一个新增变量，而不是宣布旧理论错误。这种“只能多拿一小块地”的结构，正是降低换皮风险的代价。

十三、第二层盲区：旧理论往往把“分析单元的角色”当成给定

逐一划界以后，可以看到一个比具体概念更深的共同点。过程理论把规划、生成、检查当作过程单元；知识转换把内容与修辞问题空间当作单元；反馈理论把输出、参照与调节当作单元；分布式认知讨论认知边界在哪里；过程挖掘把事件与活动当作单元。它们彼此并不一致，但大多在进入分析时已经决定“这是什么东西”。

端位转换提出的麻烦是：分析单元的功能身份本身可能在相邻轮次中改变。同一句已经写出的文本，第 n 轮可以是一个结果，第 $n+1$ 轮可以是约束；同一个“提纲”最初是规则，写到中段后可能只是被新证据推翻的旧形态；同一条路径跑出反例以后，可以变成下一轮评价规则。若研究者在编码前固定“这是结果”“这是环境”“这是过程”，这种角色转换会在数据结构里消失。

这不是说旧理论“没有讨论变化”。恰恰相反，它们讨论了大量变化。结构盲区的判断标准更严格：一旦承认同一个局部对象的功能类别必须作为随时间变化的变量，原有固定类别的统计与因果模型就需要增设一个状态维度。端位转换是否值得成为独立理论，最终取决于这个新增维度能否产生增量预测，而不是取决于命名新不新。

因此本文最重要的数据学主张也很克制：未来写作过程数据不只需要“事件类型+时间”，还需要“局部对象当前角色+角色转换”。最理想的方法不是靠研究者事后阅读猜角色，而是通过随机改变工作角色而保持内容可见，直接观察下一轮动作分布是否变化。

这种角色变量还带来一项建模上的后果。传统过程模型常用状态转移描述“规划→生成→修订”，而端位模型需要把对象状态和对象角色分开：文本片段可以在字面上不变，角色却从“产物”变成“约束”；规则文本可以字面不变，但对写作者而言从“必须执行”变成“待检验假设”。这意味着未来模型至少要有两层状态，一层记录可观察对象，一层记录它对下一步行动集合的功能关系。否则模型会把真正的角色转换错记成“同一状态继续存在”。

角色变量也可以帮助解释一个常见争论：到底是文本改变思想，还是思想改变文本？如果把二者当固定方向，就必须选一边；若允许端位转换，方向可以按轮次切换。某一轮是已有问题约束文本，下一轮是新文本暴露矛盾并改写问题，再下一轮又由新问题筛选证据。理论因此不需要宣称“文本比思想更

根本”，只需要提出一个可检验的次序：哪一个对象在什么时点获得了改变下一轮可行域的角色。

这个次序变量还可能重新解释“计划”的地位。计划并非只有有/无两档：它可以在早期是路径指示器，在局部文本出现后变成被比较的参照，若证据冲突再退化为历史记录。相反，一个最初只是随手写下的句子，也可能因为后续所有段落必须回应它而上升成新的规则。若只在研究开始时测一次“是否有计划”，会错过计划在事件内部的升降级。端位框架因此预测，真正有信息量的不是计划量，而是计划与局部产物之间的角色转移序列。

十四、十一个应用不是“类比”：每一处都必须带预测

第一，作文教学。若端位转换成立，教师不只教“先做什么后做什么”，还要安排让局部产物重新进入下一轮的问题。例如写完一个主张后，不是立即润色，而是强制它改变下一轮证据选择。预测是：只增加模板步骤不会稳定提高 E_T；改变局部产物的工作角色会。

第二，writer's block。Rose 的规则刚性解释可以进一步细化：一类卡顿不是材料少，而是局部产物始终无法取得改规则或改路径的权利。预测是：放松外部规则只有在局部产物真正能改写后续行动时才缓解这类卡顿；单纯给予更多时间未必有效。

第三，AI 共写。AI 生成比例不是核心变量。相同机器文字，如果只是填充外部大纲，端位效应可低；如果它使写作者重构证据、问题或下一提示，端位效应可高。预测是：字符来源比例与 E_T 可以显著分离。

第四，协同写作。团队成员交接的不只是文本，还可能是已形成文本对下一轮成员的约束。预测是：共享文档中一个成员的局部产物若改变下一成员的操作转移分布，跨人端位转换可被观察；若只是拼接模块则低。

第五，键盘日志研究。KUPA 真跑已经提示事件量不是状态。预测是：在控制总事件数、总时间与修订次数后，角色干预或角色转移序列比纯活动量更能区分发现式写作与抄写/固定脚本写作。

第六，自动写作支持。系统若只在停顿时补词，可能提高输出而减少端位转换；若系统能把局部矛盾重新呈现为下一轮约束，可能增加角色再入。预测是：同样减少停顿的两种工具，E_T 可以相反。

第七，无障碍写作。端位理论不把亲手键入当发生必要条件。口述、眼动、辅助输入只要允许局部产物反向改变下一轮路径，就可以有高端位效应。

预测是：字符执行主体与 E_T 不必相关，这避免把身体输入能力错误等同于写作发生。

第八，修订教学。传统修订常被理解为改正已有文本；端位转换把一类修订理解成重开后续可行域。预测是：只改局部措辞的修订可很多而 E_T 低；一次改变论证承诺并引起后续检索的修订次数少却 E_T 高。

第九，专家写作。专家有时能在写前形成完整结构，后续主要是执行；另一些研究性写作则在写中发现问题。预测是：高水平文本并不必然高 E_T。端位效应不是“专家分”，而是不同任务模式的发生指标。

第十，创意写作。小说和诗歌可能有意让材料、声音或意外句子反向改变人物与结构。预测是：探索性创作会出现多种角色换位，但频繁换位不等于质量高；优秀的极简写作也可能在少数关键点发生高幅度转换。

第十一，测评设计。若考试只看成品，惰性成形与端位转换可以获得相同分数。预测是：增加一次保持内容可见的工作角色干预，能把二者分开；但这种读数不应直接给学生奖惩，因为一旦进入考核，行为会被策略性制造。

这些应用还可以统一成一个更实用的诊断问句：“当前已经出现的東西，有没有权力改变下一步？”在教学中，这个“权力”来自任务规则；在 AI 共写中，来自人是否让机器产物改变下一提示和证据；在协同写作中，来自交接文档是否只是模块还是约束；在无碍写作中，来自口述产物是否能回到作者控制的工作循环。只有当这个问句能导出与现有方法不同的观察，跨学科才真正兑现。

同样，应用不能偷换成处方。对某些标准化任务，最好的设计恰恰是降低不必要的端位转换，让安全流程稳定执行；对探索性论文，则可能需要保留局部产物改写路径的余地。理论的价值不在给所有任务一个“高转换”方向，而在识别什么时候系统错误地把临时终点永久化，或反过来让应该稳定的规则不停漂移。

对 AI 产品设计，这种区分尤其重要。当前许多写作助手优化的是“下一句预测得更像成品”，这可能提升局部流畅，却把局部终点迅速封闭。另一类助手可以不直接补全文本，而是把刚写出的承诺、冲突和证据缺口转成下一轮可操作约束。两类系统可能获得相似成品分，却对应不同 E_T。若未来实验发现第二类工具只是增加操作复杂度而没有端位增量，端位理论在产品上的价值也应被撤回。

十五、三条反向约束：新理论不能把“发生”变成新崇拜

第一条来自形态发生：自组织没有价值保证。一个局部结构能够反向影响后续，不代表影响方向正确。错误前提也会自我强化，偏见也会把后续证据过滤成只支持自己。端位转换因此只回答“如何发生”，不回答“发生得好不好”。质量仍需证据、逻辑、伦理与修辞标准另外裁决。

第二条来自交通流：转换频率没有单调价值。交通中的高频刹停与恢复可以形成 stop-and-go wave，容量还可能在失稳后下降。写作中不停改题、跳段、撤回也可能把系统拖入振荡。本文因此拒绝用“端位转换次数”做绩效指标，E_T 只在控制条件下描述工作角色改变造成的因果效应。

第三条来自机制设计：任何测量一旦进奖惩，就改变被测行为。如果学校奖励“修订次数”“跨段移动次数”或某个端位分数，学生和 AI 系统都可以制造表面轨迹；最省事的达标方法可能恰恰摧毁读数的真实性。端位研究应优先采用隐藏随机条件、匿名研究或系统层诊断，不应直接把 E_T 排名到学生身上。

这三条约束共同改变本文的价值排序。原本很容易写成“越能让文本反过来塑造自己，作文越高级”，现在必须删掉。预成执行在新闻快讯、标准文书、紧急操作记录中可能是优点；低端位写作不等于低能力。理论的适用域是“我们要解释一个文本生成事件内部为什么会长出新的下一步”，不是给所有写作贴等级。

还有一个容易被忽略的边界：端位转换并不自动等于“创造”。在标准合同、医学记录或实验方法复写中，规则被刻意固定是安全条件，局部产物不应随意改写下一步。若把所有低端位过程都诊断成僵化，就会把可重复性和责任链破坏掉。真正需要端位空间的是那些任务目标本身允许、甚至要求发现新关系的场景。应用前必须先问：这个任务究竟是执行型、校准型还是发现型。

对教育而言，这意味着不能从“传统作文太模板化”直接推出“越开放越好”。完全没有稳定规则的课堂也可能让学生长时间停留在游离流，无法形成可被再次使用的局部结构。一个更可检验的设计是交替设置稳定与开放窗口：某些阶段冻结事实标准与伦理边界，某些阶段允许局部文本重写论证路径。若端位理论有教学意义，它应该预测这种“可稳定、可重开”的结构优于永远冻结或永远漂移，而不是只赞美开放。

十六、证伪：什么观察会让“端位转换”失去独立资格

第一条也是最硬的一条：控制文本可见性、语义内容、文本长度、题目与写作者以后，把已稳定局部文本从正文嵌入态随机移到逐字可见的侧栏态，若后续 120 秒的事件转移分布与嵌入态没有可重复差异，即 E_T 与参与者内置换零分布不可分，则该任务中的端位因果作用为零。若多种发现式写作任务都得到这一结果，核心理论应被判伪。

第二条针对外部记忆竞争解释。若“侧栏态”与“完全隐藏态”造成的影响等大、同向，而保持工作角色的其它操作不产生额外差异，则效应主要来自记忆与视觉可及性，端位角色不是必要变量。

第三条针对知识转换。实验前后测量作者对论题的主张、证据关系和知识结构；若控制这些变化后 E_T 消失，则所谓端位效应可能只是知识转换的行为影子。反之，知识不变而 E_T 存在，才说明工作角色本身有剩余。

第四条针对递归过程。把规划、生成、修订阶段切换次数、总修订量、总停顿时间纳入模型；若这些变量已经无剩余地预测嵌入/侧栏干预后的差异，端位变量没有必要。本文不能用“递归是形式、我们是机制”这种不可判定语言逃避。

第五条是时序正向读数。理论要求“局部状态先稳定，随后它的角色改变再影响下一轮”。如果高 E_T 案例中操作分布的改变系统性早于局部稳定单元出现，或者改变在侧栏干预之前已经发生，那么“终点转下一轮条件”的时序叙事是错的。

第六条是低成本日志检验。取得完整 KUPA-KEYS TASK-2 后，可先不做因果结论，只做预注册回溯：把稳定文本单元后的跨段光标移动、长停顿、修订转移与同一写作者匹配随机窗口比较。若完全没有时序富集，本文第三层经验主张受损；若有，也只能作为后续随机干预的筛选依据，不能越级当核心机制证明。

证伪设计还必须保证样本纪律。若要比任务类型的剂量或群体差异，不能拿两个完全不同班级、两个国家或两位作者做故事式对照。最低应在同一制度环境、同一平台和相似任务中获得多个可重复单元；核心随机干预更适合参与者内设计，因为它直接控制大量个体差异。端位效应若只能在跨群体相关中出现，却在参与者内随机中消失，理论应优先相信后者。

另外，机制检验应主动保存“无法编码”的窗口。某些写作发生在纸上、口述、脑内规划或外部资料交互中，键盘日志看不到全部路径。把不可见

部分强行推断成零，会制造虚假的低端位。数据集应区分“观察到无转换”和“缺少观测通道”，否则理论会把设备边界误写成认知边界。

十七、本文自身是不是端位转换的标本：一个不利结果怎样改变了理论

任何关于“结果会改写下一轮规则”的文章，都必须允许自己的结果改写自己。本文最具体的自反点不是一句“我们也有局限”，而是 KUPA-KEYS 真跑。执行统计之前，我们原本准备保留一个更强的交通流类命题：写作过程活动强度存在一个中间最佳区间，过低接近机械执行，过高接近拥堵。预注册结果没有支持它。

如果本文把这个结果解释成“代理不好，所以原理论不受影响”，就等于拒绝自己的机制。实际做法相反：二次项方向为正且不显著以后，“事件量—拥堵”整条论证被删除；交通流提供的概念从“数量阈值”收窄成“状态与容量关系”；核心读数从活动量转向工作角色干预。这是一次可以指认的“结果→规则”转换：经验结果成为下一轮理论边界。

但这仍不是对端位理论的自我证明。一个理论在写作过程中改过一次，不代表人类作文普遍如此。自反只能证明这套概念能够对自身生产过程提出可执行要求：不利结果必须改变后续允许说的话。真正的外部资格仍要交给随机干预、独立样本与竞争模型。

本文还保留两个解释不了的洞。第一，端位转换的“对象粒度”可能跨越词、句、段、论证结构，统一粒度是否存在尚不清楚。第二，角色转换可能发生在完全内部的心理表征中而没有可见文本，本文的工作区干预只能测外部可操作部分。第三，群体协同写作中，一个人的结果成为另一个人的条件时，角色归属与主体归属如何分开，需要另设模型。

自反还要求检查本文的概念命名是否反过来限制了后续研究。“端位转换”一旦被接受，研究者可能开始只寻找能被这个词识别的例子，把其它发生方式都当作噪声。这与本文批评固定角色的逻辑相冲突。更稳妥的使用方式是把它当竞争模型之一：每项研究同时放入递归、知识转换、外部记忆、任务规则和端位干预，允许数据决定哪一个解释有增量。

因此，本文最不希望看到的后续是“端位转换教学法”被迅速包装成固定七步。那会把一项关于终点角色可变的理论重新冻结成步骤。若它最终有教育价值，更可能表现为一组诊断与实验原则：保留局部产物重新进入下一轮的机会、区分再入与表面修订、允许不利结果修改规则，并在需要稳定执行的场景主动降低不必要转换。

结论：端位转换理论的有限主张

“作文如何发生”不能再被压缩成一条固定步骤，也不能因为过程理论已经证明写作递归，就认为发生机制已经穷尽。形态发生让我们看到整体结构可以后生，交通流让我们看到活动数量不能替代状态，机制设计让我们看到规则场也会被行为和环境反向改写。三条理论共同取消了一个隐蔽前提：本轮终点不是必须永久作为终点。

端位转换理论因此把研究单位放在角色而不是事件上：一个局部形态、路径或规则在第 n 轮完成某种结算以后，是否在第 $n+1$ 轮取得新的功能位置，并因此改变后续可行操作分布。稳定性与因果再入构成二维辨别，“惰性成形”解释了为什么成品可以持续改善而内部发生仍接近执行；“端位转换”则提供了一个可随机干预的机制候选。

这一候选目前没有被 KUPA-KEYS 的任务级活动量代理直接证明。相反，1,006 篇公开作文的预注册真跑否定了简单的活动强度 U 型，迫使本文把测量从“数事件”改成“识别状态与工作角色”。因此本文最重要的贡献如果存在，不是一句新的作文赞美语，而是一项更苛刻的数据要求：在保持内容可见的条件下改变局部产物的工作角色，并让它承担可失败的因果预测。

如果把这个结论放回日常作文，最简单的观察不是“我今天写了多少”，而是“刚才写出的东西有没有改变我接下来能做什么”。有时答案是没有：一段话只是执行预先知道的内容；有时答案是有：一句话突然使旧提纲失效，一个例子迫使论点收窄，一条证据让评价标准改变。端位转换试图给第二类事件一个不依赖主观感受、可以被日志和干预裁决的名字。它的意义不在于把第一类贬低，而在于让两种此前都被叫作“写作过程”的发生方式获得可区分的数据结构。

因此，本文最终保留的是一个有限命题：在部分写作任务中，局部文本不只是生成的结果，还会通过工作角色变化成为下一轮生成的因果条件；这种角色变化若独立存在，应能在内容可见性保持不变时被随机工作区干预捕捉。所有更大的主张——它是否提高质量、是否代表创造、是否属于某个主体、是否应该被教学奖励——都不由这一机制自动推出。把这些问题留在机制之外，是端位理论避免吞并整个作文研究的必要边界。

从研究程序看，这也意味着下一步不应该继续堆更多漂亮案例，而应优先做能够让理论失败的最小实验。只要一项设计能把同一局部文本的内容、可见性、写作者和任务固定，仅改变它在工作区中的功能位置，就能直接触碰本文最承重的变量。如果这种改变对下一轮行动没有稳定影响，我们就不需

要“端位转换”；如果有影响，再去研究不同方向、尺度与任务边界。与继续发明术语相比，这种实验更便宜，也更能决定本文究竟是一篇跨学科解释，还是一个可以进入作文科学的机制假说。

换句话说，本文不把“作文如何发生”的答案封成一条新流程。它把问题重新开放为可裁决的状态转换：局部形态何时被封闭，何时重新取得因果权；路径何时只是执行，何时反过来改写规则；规则何时提供稳定，何时必须被结果修订。只有这些转换能够被别人用同一实验重复观察，“如何发生”才从叙事解释进入机制研究。这也是本文最终愿意承担的最小科学承诺：先让机制能够失败，再谈它能解释多少作文。

十八、最后的赌注：到 2028 年底，什么结果算本文输了

为了避免“未来还要更多研究”这种无限延期，本文给出一个写死日期的赌注。截止 2028 年 12 月 31 日，如果公开文献中仍没有至少一项满足以下条件的研究： $N \geq 60$ ；参与者内随机；每位参与者至少完成多次同类论证写作任务；在局部文本稳定后，随机比较正文可编辑嵌入态与逐字可见、不可编辑侧栏态；两种条件保持文字内容和可见性一致；并以本文同一口径的 E_T 比较后续 120 秒事件转移分布——那么“端位转换”作为独立作文发生机制应降级为描述性框架。

命中也有严格条件。只比较高分和低分作文不算；只报告修订次数、停顿次数或字符数不算；把文字完全隐藏导致记忆负担变化不算；让模型生成看起来像“回写”的日志不算；没有参与者内随机、没有保持同一文字逐字可见、没有公开事件编码规则，也不算。至少一项研究还需要同时纳入递归阶段切换、知识变化与一般编辑便利性的竞争模型。

本文的正向预测是：在需要发现和重构的写作任务中，E_T 的参与者内分布应高于抄写或严格脚本执行任务；而且局部文本稳定时点应早于后续操作分布改变。更强的预测不是“E_T 越高分数越高”，本文明确拒绝这一点。高质量预成写作可以低 E_T，混乱探索也可以高 E_T。端位效应只解释发生方式，不承诺质量单调性。

如果到期没有这种增量效应，最合理的结论不是给概念换一个更模糊的名字，而是承认现有递归、知识转换、外部认知与反思理论已经足够。若效应存在，下一步也不是宣布“发现作文本质”，而是继续问：哪些任务、年龄、工具和协作结构下，哪一种端位方向真正发生，以及它什么时候帮助写作、什么时候制造振荡。理论只有在能够输的时候，才开始具有经验内容。

日期赌注还可以进一步量化失败的层级。若随机实验显示嵌入态与侧栏态存在差异，但差异完全由光标移动成本、可编辑便利性或视觉位置解释，端位转换作为认知—生成机制仍应失败；它最多成为界面效应。若差异只在一个特定编辑器存在，换到纸笔、语音或另一平台就消失，也不能宣称普遍作文机制。真正值得保留的是跨界面、跨任务仍能用“工作角色”而非具体按钮解释的稳定增量。

同样，如果研究发现 E_T 与写作质量毫无关系，本文不会因此自动失败，因为质量单调性从未被写入核心命题；但如果 E_T 与任何后续路径改变、目标重设、证据选择或结构重排都没有增量关系，只反映局部编辑操作，那么“发生”这个词就被用得太大，应改称界面内操作耦合。把失败层级提前写清，是为了防止未来每遇到反例就缩小定义直到永远正确。

截至赌注日期，还应至少出现一次独立团队的复现。若只有提出该理论的团队能通过专有编码得到 E_T，而外部团队使用公开数据与预注册规则无法复现，理论也应降级。一个关于作文发生的机制不应依赖只有作者自己会看的“特殊眼光”；它必须能被别人按同一清点手册打败或复现。

还有一种更彻底的失败：如果未来高分辨率日志显示写作者的下一步主要由内部不可见计划决定，而外部局部文本的位置、可编辑性和嵌入关系几乎不改变路径，那么端位转换最多是少数界面化写作的边缘现象。这样的结果会支持“外部文本只是内部生成结构的痕迹”而不是“局部产物成为下一轮因果条件”。本文不能用“真正的转换发生在脑内所以测不到”自救；一旦把核心机制放到不可观测内部，就失去了本文当前可证伪设计的价值。

反过来，即使实验支持端位效应，也仍需要回答方向与尺度。形态→路径、路径→规则、规则→形态是否有不同的时间常数？一句话、一个段落和全文结构的换位能否用同一指标描述？这些问题决定端位理论最终是一个统一机制，还是几个表面相似机制的总称。日期赌注只决定它是否值得继续存在，不提前保证它会长成一个大理论。

参考文献

- [1] Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B*, 237(641), 37–72.
<https://doi.org/10.1098/rstb.1952.0012>
- [2] Lecuit, T., Lenne, P.-F., & Munro, E. (2011). Force generation, transmission, and integration during cell and tissue morphogenesis. *Annual Review of Cell and Developmental Biology*, 27, 157–184.

- [3] Kicheva, A., & Briscoe, J. (2023). Control of tissue development by morphogens. *Annual Review of Cell and Developmental Biology*, 39, 91–121.
- [4] Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves II: A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A*, 229, 317–345.
- [5] Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, 4(1), 42–51. <https://doi.org/10.1287/opre.4.1.42>
- [6] The Nobel Prize. (2007). Mechanism design theory: Scientific background and popular information for the Sveriges Riksbank Prize in Economic Sciences.
- [7] Bergemann, D., & Morris, S. (2005). Robust mechanism design. *Econometrica*, 73(6), 1771–1813.
- [8] Bergemann, D., & Välimäki, J. (2019). Dynamic mechanism design: An introduction. *Journal of Economic Literature*, 57(2), 235–274.
- [9] Roth, A. E. (2023). Market Design and Maintenance. NBER Working Paper No. 31947.
- [10] Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- [11] Rose, M. (1980). Rigid rules, inflexible plans, and the stifling of language: A cognitivist analysis of writer's block. *College Composition and Communication*, 31(4), 389–401.
- [12] Perl, S. (1980). Understanding composing. *College Composition and Communication*, 31(4), 363–369.
- [13] Bitzer, L. F. (1968). The rhetorical situation. *Philosophy & Rhetoric*, 1(1), 1–14.
- [14] Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, 27, 261–278. [https://doi.org/10.1016/0749-596X\(88\)90054-X](https://doi.org/10.1016/0749-596X(88)90054-X)

- [15] Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press.
- [16] Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482.
- [17] Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28(2), 122–128.
- [18] Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books / Temple Smith.
- [19] Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513–549.
- [20] Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
- [21] Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392.
- [22] van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action* (2nd ed.). Springer.
- [23] Velentzas, G., Caines, A., Borgo, R., Pacquetet, E., Hamilton, C., Arnold, T., Nicholls, D., Buttery, P., Gaillat, T., Yannakoudakis, H., & Ballier, N. (2024). Logging Keystrokes in Writing by English Learners. *LREC-COLING 2024*, 10725–10746.
- [24] MacArthur, C. A., Graham, S., & Fitzgerald, J. (Eds.). (2025). *Handbook of Writing Research* (3rd ed.). Guilford Press.
- [25] 《何谓作文——生成见证视角下的作文等价类理论》（2026，同批内部研究稿）。

附录 A：E_T 清点手册与最小预注册

读数名：端位效应 E_T。定义：同一写作者、同类任务、匹配稳定局部文本单元下，正文可编辑嵌入态与逐字可见侧栏态在随后 120 秒七类事件转移分布之间的 Jensen–Shannon divergence。

粒度：一个“稳定局部文本单元×随机条件窗口”为一条记录。稳定单元在英文实验中预注册为连续 ≥10 词、连续 60 秒未被编辑、光标已经离开；中文或其它书写系统必须在看结果前给出等价字符/词单元规则。

事件编码：①向后追加；②局部修订；③跨段回改；④删除簇；⑤远距离光标移动；⑥停顿≥2 秒；⑦来源/工具调用。任何新增类别必须在主分析前冻结。

最小主分析：参与者内比较嵌入态与侧栏态 E_T；用条件标签随机置换形成零分布；同时报告递归阶段切换、总修订、知识前后测与编辑便利性控制。E_T 是机制读数，不转换成学生等级。

附录 B：证据边界、人机协作与版本

证据边界：本文对形态发生、交通流、机制设计、写作过程和认知科学近邻的叙述依据公开学术资料与权威来源。跨学科转译不意味着来源学科已经认可“端位转换”用于作文研究。端位转换、惰性成形、E_T 及其二维框架属于本文提出的研究假说。

真跑边界：本次 KUPA-KEYS 分析在统计前保存预注册，使用公开 META 数据并保留不利结果。该分析只检验任务级“过程活动强度”代理，不直接检验端位因果机制；完整逐事件日志在本次运行环境中未被稳定取得，因此核心结论没有以代理数据冒充验证。

伦理边界：键盘日志可能包含高度细粒度的行为信息。正式研究需取得知情同意、最小化原始日志保存，并优先在特征提取后删除不必要的原始内容。端位读数不得直接用于学生排名、处罚或安全关键岗位的强制轮换；若任何机构据此作出不利决定，应公开编码规则并提供人工复核通道。

人机协作声明：资料检索、跨学科比较、候选竞争检验、公开数据统计、初稿组织和文档排版使用了生成式人工智能辅助；最终理论主张、引文准确性、研究伦理与对外使用责任需由实际署名作者人工复核。

版本：V1.0, 2026-08-17。

一·被移到侧栏的那一段，正是那个上游构念

本章的核心实验是：内容一字不改、逐字可见，只把一段已稳定的文字从可编辑的正文位置移到不可编辑的侧栏，比较随后一段固定时间内的操作转移分布。

编者必须指出一件本章正文没有说的事：被移动的那个东西，在写作研究里已经有名字。它就是已写出的文本（the text produced so far），四十年前就被放进了写作的任务环境（Hayes 1996），此后一直是键击记录研究的主要观察对象（Leijten & Van Waes 2013）。本章引了这一支的方法学文献，却没有引它的构念本身，也没有把它列进第十二节那十一位占位者。

这是本章最贵的一处缺席。因为那一支不只是说「已写文本会影响后续」——它把已写文本定名为任务环境的一个组成部分，而本章的整个实验，正是在改变这个组成部分的一项属性。

分离线。那一支把已写出的文本当作任务环境中一个可被读取的信息源，它的研究问题是写作者何时读它、读多久、读了之后做什么。本章把它当作工作区中一个可被操作的对象，研究问题是它的可操作性本身是否构成一个独立变量。

判定形式（本章其实已经写好，只是没挂在这位祖宗名下）。本章的第三臂设计——正文内冻结态（保留原位置、仍可见、暂时不可编辑）——恰好就是这条分离线的实验实现：

- 若「完整嵌入」与「正文内冻结」无差别，则可编辑性不重要，本章测到的仍属于信息读取，应并回那一支。
- 若「正文内冻结」与「侧栏」仍有差别，则文档嵌入关系另有作用，本章的变量比那一支多一项。

编者建议本章在修订时把这三臂直接标注为对该构念的三种处理，本章的位置会因此立刻清楚。

二·定名：本章的「惰性成形」引第一章

「局部文本稳定而无因果再入」这一格，本卷以第一章的伪启动为正名（见第一章补节）。本章的「惰性成形」为同一格的另一种测法：本章看操作分布，第一章看下游依赖。读者遇到本章此词，应回到第一章的定义。

两者不是同义反复，因为两种测法可能分离：一段文字可以对下游语义选择有依赖（第一章判为已启动）而其工作位置的改变不引起操作分布变化（本章判为无端位效应）。这种分离若真实存在，说明「取得续写权」与「取得工作角色」是两件事，本卷两章都需要收窄。编者把这一条列为两章共同的观察项。

三·编者提示：本章的诚实是它的天花板

本章第十二节把话说了到这个地步：「本文若把『写作会反过来改变思考』当贡献，几乎每一位都能取代它」，并逐条列出递归写作、知识转换、实践中的反思、认识性行动、反馈控制、修辞情境、分布式认知、稳定中间形态、键击记录、过程挖掘各自已经占了哪一块地，最后把自己的剩余命题压成一句：同一局部产物的内容保持可见，只有它在工作区中的功能位置被改变，后续操作序列仍出现差异。

编者认为这是全卷最规矩的一次自我限定，也正因为如此，本章的不可还原性由作者自己封了顶。它明说自己「只能多拿一小块地」。读者应当按这个尺度读它：本章不是一套新的写作理论，是一个新增变量加一台可以判它死活的实验。

本章的真跑同样应当按这个尺度读。那次预注册检验的对象不是本章的机制，而是一个较弱的代理猜想（过程活动强度与评分的倒 U 关系）。结果不支持，二次项方向与预期相反。作者据此删掉了整条第三层主张，并把核心读数从活动量改为工作角色干预——这是一次真正的自我删改，但它没有为本章的核心机制提供任何正面证据。本章至今没有一次针对自身机制的数据。

编二

权在文本内部怎么走

共同动作：把「下一步写什么」从作者的选择改成账目的移动。

这一编只有两章，因为它们是一笔账的两端。

第四章向前读。它拒绝「相关即可续写」，要求下一个单位同时做完三件事：承接（结清一个在它出现之前就被冻结的旧义务）、改场（让「现在这篇文章在回答什么问题」发生可编码的改变）、转位（生成一个依赖于本次变化的新义务——删掉它，那个新问题就不该以同样的身份和优先级出现）。三件缺一，就不是持续生成。由此「材料堆积」第一次有了比「例子太多」更精确的定义：与主题有关、自身可能还很精彩，但既不还旧账，也不生成依赖性的新账。

第五章向后读。它问的是一个更奇怪的事：某句话被删掉之后，它原来承担的那份篇章职责去哪儿了。答案是它被别的句子接了过去，而那些句子一个字都没改。于是有了本卷最反直觉的一个测量对象——字面完全一致、角色已经不同。

两章合起来，「下一步写什么」不再是一个关于作者意图的问题，而是一个关于账目余额的问题。

这一编的共同短板，编者不隐瞒。两章都没有算出过一个数。第四章的预注册字段审计得出的是「现有语料不能直接计算」，第五章通篇是理论建构与实验设计，作者自陈没有完成任何一次数据采集。它们的本钱在设计而不在读数——第五章那个四臂随机的「同一现稿、不同历史」，是全卷设计得最干净的一个因果识别方案，只是还没有人跑过它。

第四章

作文如何持续长下去

——续生位转移

作文如何持续“长下去”：从材料追加到“续生位转移”——高分子科学 × 形式化方法与程序验证 × 细胞生物学的二阶碰撞

学科通融专栏 · 二阶碰撞典范文 (How 2)

研究类型：How (三焦点路径型) | 对象：作文启动后的持续生成 | 方法版本：SDE 多学科通融创新法 3.9

摘要：本文研究一篇作文启动以后怎样持续生成，以及怎样区分真正推进与同题材料堆积。依照 SDE 多学科通融创新法 3.9 的 How 型三焦点路径工序，本文从“新思想前沿”选取高分子科学中的活性/可控聚合、形式化方法中的翻译验证，以及细胞生物学中的 YAP/TAZ 力学转导。三家分别把持续性压在可复启生长端、逐次转换的可核验条件与当前状态对后续路径的回写上，由此形成斜对立。二阶碰撞提出新机制“续生位转移”：实质新增单位必须承接前文刚产生的未结生成义务，改变当前全文的问题场，并把一个依赖本次变化的新义务转移到下一位置。本文据此建立“承接一改场一转位”判据、承接×转位的 2×2 辨别格、续生转移率 GTR 及清点手册。对 ArgRewrite V.2 多稿修订语料进行预注册字段审计得到不利结果：现有数据虽有多稿文本、修订目的和操作标签，却缺少写作当时逐步冻结的前后续生位，因而目前

不能直接计算 GTR。本文据此将贡献限定为可实验、可证伪的持续生成理论，并提出课堂“续生三问”与段界回写卡。

关键词：作文教学；持续生成；二阶碰撞；续生位转移；写作过程；语文教育
二阶碰撞工序摘要

学科	How 路径	源内硬判断	单独成立时的盲点	本文中的作用
高分子科学	$S \rightarrow E \rightarrow D$	可逆单电子转移可把自由基聚合改造成链端可重启过程	“可继续”不等于“这一新增单位对整体有效”	提供“可重启生长端”位约束
形式化方法与程序验	$S \rightarrow D \rightarrow E$	为每次具体转换生成可检查条件，核验这一次是否保持语义	局部通过不等于未来还能生成；规范也可能被新状态改写	提供“逐次承接必须过闸”的结果核验
细胞生物学	$E \rightarrow D \rightarrow S$	基质刚度、形状和张力的经 YAP/TAZ 可改写增殖与分化	当前形态不是被动结果，而会回写后续路径；但不能单独归因	提供“写完这一段后，文状态发生什么改变”

一、引言：真正困难的不是“没有材料”，而是“下一段为什么非写它不可”

作文教学里有一种非常常见、也非常顽固的失败：学生并不是写不出来，而是越写越多以后，文章反而越来越不像一个正在生成的整体。第一段已经提出了一个人物、一件事或一个观点，第二段补一个例子，第三段再补一个例子，第四段加一句名言，第五段换一个相近材料。单看每一段，语法可以成立；单看主题，所有材料也都“有关”；甚至每一段前面都能加上“首先、其次、另外、最后”。可是读者读到任何一段以后，都很难回答：如果删掉这一段，后面为什么必须改变？如果把第二、三段对调，文章的生成路径为什么会破坏？如果再添十个相似例子，文章究竟会更深，还是只会更长？

这种失败说明，“有内容”与“能持续生成”不是一回事。材料库解决的是可调用资源不足；提纲解决的是事前安排；衔接词解决的是表面接续；主题句解决的是局部归属；篇章结构理论解决的是成品之间已经存在什么关系。它们都重要，却仍可能绕开一个更微观、更动态的问题：当 T_t 已经写在纸上以后， T_t 内部到底产生了什么，使下一步 m_{t+1} 不再是任意添加，而成为一个被前文逼出来、又会改变后文空间的动作？本文把这个“逼出来”的东西称为作文持续生成的待解释对象。

如果只把作文持续生成理解成“不断有想法”，那么思路活跃的学生理应最少堆料；现实却常相反：他可能能联想到很多材料，却不知道哪一个现在应该进入。若只理解成“围绕中心”，则凡与中心有关的材料都应当具有同等加入资格，但恰恰是这种同等资格制造了堆积。若只理解成“逻辑连贯”，我们又会把一个已经写完的成品结构倒投回生成过程，仿佛作者在每个时刻都已经知道未来的完整树形。真正需要解释的是一种局部生成责任如何一边被结清、一边转移的动态。

SDE 3.9 对 How 题的要求不是找三门学科各给一个技巧，而是让三条不同焦点的路径互相拆台：一条把“怎样继续”押在路径端，一条押在条件端，一条押在当前显露状态端；只有当任何一家都不能独自结算时，才可能涌现新的第三物。[1] 因此，本文不从作文教育内部已经高度熟悉的认知写作、修辞、

篇章语言学、读者反应或创意写作中再挑三套相邻理论，而从 SDE“新思想前沿”抽取高分子科学、形式化方法与程序验证、细胞生物学三家远距理论，专门研究“已经开始以后怎么长”。[2-4]

二、闸零：为什么这三门源可以进场，而“写作像生长”本身必须出场

3.9 版首先要求过“源的未吸收度”和“理论厚度”两道闸。本文以“作文/写作教学 + living polymerization”“composition pedagogy + translation validation”“writing education + YAP/TAZ”进行中英文公开检索，没有发现这三套具体理论被作文教育作为成熟理论直接吸收的证据。这里的结论只能写成“就本次公开检索范围属于干净源”，而不能写成绝对的“从未有人使用”。三家又都不是一条孤立比喻：活性/可控聚合有从 Szwarc 到金属催化活性自由基聚合的长期理论与实验谱系；形式化验证有规范、语义、反例、可信基和逐层验证体系；YAP/TAZ 力学转导也有明确的实验操纵、读数、边界和后续争论。因此三家均通过理论厚度闸。[5-10]

但同一轮闸零也找到了一个必须正面承认的危险占位者：Peter Elbow 在 1973 年《Writing Without Teachers》中已经明确使用 writing as growing 的有机生长隐喻，并把自由写作、生成、重写等过程放在“让文字长起来”的框架中。[13] 这意味着，如果本文最后只是说“作文不是拼装，而是像生命一样生长”“应该让文章自己长出来”，创新度直接归零。高分子科学也不能被浅表迁移成“句子像单体、段落像链条”；细胞生物学更不能被写成“文章像细胞会自组织”。这些都是同名比喻，不是新机制。

因此，本文给自己写下一条候选闸：新对象必须至少满足三个条件。第一，它能把“主题高度相关但只是堆料”和“主题相关且真正推进”分开；第二，它不能被“连贯”“递归”“知识转化”“未完成感”“生长隐喻”中任何一个旧词 1:1 替换；第三，它必须能够在写作过程中逐步清点，而不是只能等成品出来以后凭审美印象判断。若做不到这三点，三门远学科再漂亮，也只是替旧作文术语换皮。

三、第一条因果脊（D 焦点）：高分子科学把“增长”改写成“链端可复启”

高分子科学提供的第一条硬判断来自“活性”或“可控”聚合。Szwarc 在 1956 年提出 living polymers 的经典概念，活性聚合的核心并不是“聚合物会一直变长”，而是在适当条件下尽量避免终止与不可逆链转移，使已经形成的链端仍保留继续反应的能力。[5] 后来的金属催化活性自由基聚合把这件事推进到自由基体系：Kamigaito、Ando 与 Sawamoto 总结了金属催化循环如何通过可逆活化/失活降低活性自由基瞬时浓度，使链端能够被再次激活并继续增长。[6] SDE“高分子科学”页面将这一经典压成一句极有碰撞力的话：“可逆单电子转移可把自由基聚合改造成链端可复启过程”，并把它标为 D 位。[2]

这条理论对作文最有价值的地方，不是“链”这个形状，而是继续发生的资格被集中在一个可复启端上。一个体系中即使还有大量单体，如果链端已经

失活或发生不可逆终止，材料再多也不会自动变成同一条链的继续生长；反过来，只要端基身份被保留并能在合适条件下重新激活，增长可以暂停后继续。由此，作文持续生成第一次获得一个比“还有没有材料”更硬的类比：真正稀缺的不是材料总量，而是当前文本是否留下了一个属于这篇文章本身、能够被下一步接上的继续位置。

把这条路径写成 How 的 D 焦点，可以表示为 $S \rightarrow E \rightarrow D$ ：已经形成的当前链状态 S，在催化、配体、温度与可逆活化等条件 E 下，重新进入可延长的反应路径 D。迁入作文后， T_t 并不会因为“还没写够字数”就天然拥有后续路径；它必须留下一个可辨认的生成端。比如一篇议论文写到“手机扩大了人的选择，却也让人更难选择不回应”，真正的生成端不是“我还可以再举一个手机的例子”，而是刚刚出现的冲突：自由究竟是选择数量，还是保留拒绝即时回应的能力？这个尚未结清的差异，才有资格吸附下一段。

然而，高分子科学一旦独占结算权就会出问题。化学意义上的链端只需满足体系内的反应身份；作文中的“开放问题”却可以无限制造。学生完全可以在每段末尾故意抛出一个新问题，形成“那么我们又该如何看待……”，却让问题与前文主路径越来越远。这种文本看起来永远有下一步，实际上只是旁枝不断增殖。因此，可复启不足以证明同一作文正在生长。这正是第一家理论必须被第二家拆掉的地方。

四、第二条因果脊（E 焦点）：形式化方法把“接得上”改写成“这一次转换必须过闸”

形式化方法与程序验证提供另一种几乎相反的工程直觉：不要因为一个大工具过去表现可靠，就默认它这一次产生的结果仍然正确；对关键转换，可以把“相信整个系统”改成“检查这一次实际结果”。SDE 专栏用“翻译验证：不必证明整个优化器，只需核验每次结果”概括这一路线：不问成熟编译器是否永远正确，而问这一具体输入经过这一具体转换以后，输出是否与输入保持规定的语义关系。[3] Lopes 等提出 Alive，用可检查规则描述 LLVM 的局部优化，并自动证明规则正确或生成反例；他们将三百余条优化写入 Alive，发现其中八条存在错误。[8]

这一条对作文的迁移点非常明确：“下一段看起来能接”不能靠作者的整体自信结算，而要为这一次新增生成一个局部可检查义务。在作文中，所谓义务并不是形式逻辑中的语义等价，而是本研究预注册的一条生成约束：在揭示下一段以前，先冻结前文 T_t ，要求独立标注者写出“此刻最需要被回答、区分、解释、推进或兑现的 1—3 个问题”；再揭示新增单位 m_{t+1} ，检查它是否实质承接其中至少一项。如果 m_{t+1} 只是主题相同，却没有结清任何已冻结义务，它就不能因为“内容不错”被记作持续生成。

这条 How 路径是 $S \rightarrow D \rightarrow E$ ：从这一具体版本/结果 S 出发，沿本次转换关系 D，生成并落到一个可检查的条件边界 E。它与高分子路径的冲突很尖锐。高分子只问“端还活不活”；形式化验证要追问“你接上去的这一段，究竟承接了哪个已经存在的义务”。于是，一个很会“留扣子”的作文可能在 D 意义上充满活口，却在 E 意义上连续失败：新段落不断开新口，却从不还旧账。

但形式化方法若独占结算权，也会制造另一种僵死。第一，局部核验依赖规范；规范写错，证明越完整越可能只是把错误规范执行得更忠实。第二，一个段落完全可以把前文问题回答得非常清楚，却把所有张力一次性关闭，后文不再有任何必要性。此时“承接通过”只证明这一步没有跑偏，不证明文章还能继续生成。第三，文学叙事、抒情散文中的有效推进经常通过视角、节奏、意象、人物位置的改变发生，未必能化成一个固定命题表。因此，逐次过闸必须接受第三家的回写：新增以后，当前文本本身已经变了，下一步的规范不能永远照抄上一刻。

五、第三条因果脊（S 焦点）：细胞生物学把“当前形态”从结果改成下一步的输入

Dupont 等 2011 年关于 YAP/TAZ 力学转导的研究改变了一个重要边界：细胞并不只读取可溶性生长因子和遗传程序，也读取细胞外基质刚度、细胞铺展形状和细胞骨架张力等物理信息；YAP/TAZ 可以作为机械信号进入核内的中继，影响细胞生长和分化。[9] SDE“细胞生物学”专栏因此把硬判断写成：“基质刚度、形状和张力经 YAP/TAZ 核定位足以改写增殖与分化”，并把它放在 S 位，强调当前显露状态不是被动结局，而会进入下一轮命运结算。[4] Engler 等更早关于基质弹性与干细胞谱系倾向的工作，也为“环境力学条件能够进入命运选择”提供了经典实验背景。[10]

迁入作文后，这一条击穿一种非常深的线性直觉：我们常把“已经写出的段落”理解成过去动作的沉淀，把“写什么下一段”理解成作者脑中原计划的继续执行。若如此，提纲才是真正的因，正文只是逐格填充。但细胞力学的启发恰好相反：已经形成的局部形态会改变后续可行路径。作文写出一段以后，人物关系、证据负担、叙述视角、价值冲突、读者预期、论证尺度都可能被重新安排；因此下一段不是“原提纲的第 3 点”，而应首先读取当前文本已经变成什么。

这条 How 路径是 $E \rightarrow D \rightarrow S$ ：外部/局部条件场 E，经由力学转导路径 D，形成新的细胞状态 S。迁入作文，就是写作情境、前文结构与局部选择形成一个条件场，经本次表达动作，出现新的全文状态。比如前文只是讨论“手机是否增加自由”，某一段一旦引入“即时可达性”并区分“可以联系”和“必须随时回应”，全文对象就发生了尺度迁移：它不再主要是“设备功能多少”，而变成“谁有权占用谁的回应时间”。这一状态变化会重新排序后面哪些材料有资格进入。

细胞生物学同样必须自曝边界。YAP/TAZ 与力学输入的关联并不能在细胞密度、黏附配体、代谢状态等共同变化时被简单解释为单一力学因果；SDE 页面也明确把这一点列为失效边界。[4] 对作文而言，这意味着我们不能因为“写完某段后全文看起来不同”就把变化都归因于该段的续生作用。知识量、题目难度、语言熟练度、教师提示、段落长度等都可能共同造成变化。因此第三家只提供“状态必须回写”的约束，不能单独证明因果。

六、一阶对立：三家分别把“继续”交给不同的裁判

把三条因果脊并列以后，一阶冲突已经清楚。高分子科学说：只要生长端仍能复启，系统就保留继续延长的路径；形式化方法说：能继续不等于应该这样继续，本次转换必须对既有关系承担可检查责任；细胞生物学说：即使本次转换通过旧规范，转换之后的当前状态也已经回写条件，下一步不能继续使用完全不变的旧问题。三家争的不是词汇，而是谁拥有判决下一步合法性的最高权力。

若只站高分子一边，作文教学会滑向“保持开放、不断留悬念”。它可以解释为什么学生停住，却解释不了为什么“越写越岔”也同样是一种失败。若只站形式化验证一边，教学会滑向一张越来越细的段落核验表：上一段提问题，下一段回答；上一段提出观点，下一段给证据。它能减少跑题，却容易把文学表达和真正的观点转化压扁成固定协议。若只站细胞生物学一边，则容易滑向“根据现在的感觉往下写”，把任何局部变化都解释成自组织，最终又失去可裁决性。

一阶碰撞因此不能靠“综合三家优点”收场。所谓“既要开放、又要连贯、还要根据全文调整”只是三因素清单，没有产生新对象。3.9 版要求继续向下找三家的两层共有前提，并且推翻材料必须来自三家自身。

七、敌意拓宽：至少十个既有占位者已经覆盖了“过程、递归、未完成与连贯”

为了防止把旧写作理论重新命名为新概念，本文按“同名直系、同功能异名、上位理论、工程邻域、方法邻域”进行敌意拓宽。结果显示，这个问题的近邻极多，尤其不能靠“过程”“递归”“开放”“连贯”四个词自证创新。

者	它已经解释了什么	与本文最危险的重合	本文必须留下的分离线
arnik (1927)	未完成任务较易保持心理可及性	“未完成会持续占据注意”	记得/想回去不等于下一文本单有结构性承接义务
ankina (1928)	被打断任务具有重新开始倾向	“启动后会想继续”	行为复返不说明为什么某一段/另一段成为下一步
g (1971)	以过程研究真实写作者的 composing	“作文是时间中的过程”	本文只研究过程中“局部生成或如何转移”这一更窄对象
ay (1972)	写作应教过程而非只评成品	“写作是发现、重写”	过程立场不等于提供持续生成小事件判据
w (1973)	writing as growing、自由写作与生成	“写作会长”	本文拒绝生长隐喻本身，要求节点的承接—改场—转位
(1979)	非熟练写作者也表现出可辨识递归过程	“写作不是直线”	递归描述不等于下一步的义务有机制
er & Hayes (1981)	目标、计划、转译、评估等认知过程递归互动	“下一步受目标与问题空间影响”	本文把焦点从作者内部过程移外部冻结的文本边界义务
ter & Scardamalia (1987)	知识陈述与知识转化区分	“写作可反过来改变知识/问题表征”	知识转化是宏观模式；续生位逐新增单位的可核验事件
h & Thompson (1988) RST	描述文本跨度之间的修辞关系	“段落关系与整体连贯”	RST 多为成品关系分析；本文关系如何在时间中生成并转移责任
ewrite V.2 (2022)	标注跨稿修订目的、范围与操作	“可对写作变化做细粒度编码”	修订目的不记录写作当时前后位，正好暴露本文的新字段缺

这张敌意表最重要的结果不是“本文比前人更复杂”，而是把一个很大的创新空间主动砍掉。任何只说“写作是过程”“写作会生长”“写作是递归的”“段落之间要有关系”“未完成会促使继续”的版本，都已经有强占位者。[11-17,20-

22] 本文若有独立性，只能在更小的缝隙里成立：在一篇正在写的具体作文中，一个新增单位如何同时对过去负债、对现在改场、对未来转移负债。

七之二、27 对混沌碰撞后的三条暗流：持续生成不是“不断开口”，而是责任在三种位置间换手

按照 3.9 工序，真正参与碰撞的不是三个学科名称，而是三家各自抽出的结果层、路径层和条件层支撑。九条支撑跨家形成 27 对，逐对检查以后，大量配对只能得到一般类比，例如“系统需要条件”“过程会受环境影响”，这些都没有资格进入新理论。能够反复出现、并且彼此不能被任一家单独吸收的有效涌现，最终主要抱成三条暗流：生长资格必须被继承、每次继承必须被这一次结果重新核验、核验后的结果又会改变下一次继承资格。这三条不是三家原话的拼接，而是三家互相取消对方“单独够用”以后剩下的共同结构。

第一条暗流可称为“端—闸冲突”。高分子科学告诉我们，保留可复启端才能继续；形式化验证却迫使这个端每次都接受具体转换的检查。两者一撞，作文里的“还有话可说”就失去自动合法性：一个开放问题只有在它确实由当前文本产生、并且下一单位能说明自己承接了什么时，才是有效生长端。这里出现的不是“开放性+连贯性”，而是一个顺序约束：先冻结旧端，再揭示新单位，再判承接。若顺序反过来，研究者总能事后把任何新增解释成“原来需要它”，于是理论不可证伪。

第二条暗流可称为“闸—场冲突”。形式化验证倾向于让一个稳定规范裁决这一次转换，而细胞生物学提醒我们：这一次转换形成的新状态，会反过来改变下一轮条件。迁到作文以后，一个段落即使完全兑现了上一轮义务，也不能因此自动获得下一轮规范；成功的回答会改变问题。例如前一轮要求“解释为什么没有立刻回复”，回答若把原因从情绪改写成“无法确定哪个地方才算家”，下一轮就不能继续按“情绪原因”收集材料，而必须转向归属、身份与承诺。上一轮核验通过，本身就是下一轮核验口径发生变化的原因之一。

第三条暗流可称为“场—端冲突”。细胞状态回写意味着当前形态会改变后续路径，高分子科学则要求真正的继续仍需保留可复启端。二者相撞后，作文的“改场”不能停在“这一段使文章更丰富”：状态改变之后必须能够指出一个新的优先继续位置；否则所谓变化只是已经发生的结果，没有把生成责任带到未来。反过来，新问题若出现却并不依赖本次状态变化，就只是作者又开了一条支线。由此，三条暗流在同一时间箭头上闭合：旧端给本步以责任，本步过闸后改写当前场，当前场再生成新端。“续生位转移”正是这条闭环的对象化命名。

这一步也解释了为什么本文不把“生成”定义成信息量增加。信息量、材料量、句子数都可以单调上升，而责任链可以在第二段就中断。真正的持续生成具有一种局部不可交换性：当 m_{t+1} 承接 O_t 并制造 O_{t+1} 后，后一个单位 m_{t+2} 的合法性部分依赖 m_{t+1} 已经发生。若把 m_{t+1} 与 m_{t+2} 调换，至少一个承接关系应受损。材料堆积则常具有高交换性：多个例子只共同指向总主题，彼此并不生成对方的入场条件。这一可交换性差异后来成为本文最重要的反事实测试之一。

八、两层共有前提：三家都默认“下一步可以由一个特权终点单独结算”

第一层共有前提是：三家虽然选择了不同裁判，却都倾向把另外两项降为背景。高分子科学把可复启路径当作核心；形式化验证把可检查条件当作核心；细胞生物学把当前显露状态及其转导结果当作核心。三家都可以写成同一语法：只要 X 成立，另外两项作为条件或后果即可被带过去。这正符合 SDE 3.9 所说的“三路径共有前提”危险：把一个目标方向单独当成终点，其余只是手段。[1]

第二层共有前提更深：三家都默认“上一轮成功所用的判决口径，可以被相对稳定地带到下一轮”。活性聚合默认链端身份能在可逆循环中被保留；翻译验证默认有一组足以裁决这次转换的语义规范；YAP/TAZ 实验则必须固定足够多的细胞与环境条件，才能把力学变化和结果变化建立因果联系。可作文持续生成恰恰具有一种更麻烦的特征：成功写出这一段，会改变下一段必须满足的条件。也就是说，判决本身部分由前一步生成。

推翻材料仍来自三家内部。高分子科学自己承认端基保真会被不可逆终止、污染、传热与放大条件破坏；形式化方法自己承认验证只相对于规范、语义模型与可信基成立，核验器本身也会成为新信任点；细胞生物学自己承认力学信号与密度、配体、代谢等协变量纠缠，当前状态不是一眼可读的单因结果。[2-4] 因而任何一家都无法为作文提供“一个终极指标”。

九、二阶结算：作文持续生成的对象不是“材料流”，而是“续生位转移”

把三家重新结算后，本文得到的新对象不是“开放结尾”“段落钩子”或“主题链”，而是续生位（**generative continuation locus**）。所谓续生位，是当前文本边界上一个尚未结清、但已经被前文具体生成的关系义务：它可能是一处必须解释的因果缺口、一个刚被制造的概念区别、一项证据负担、人物行为后必须兑现的后果、叙事视角突然改变后必须重估的信息、一个意象反复后新出现的差异，或一个价值判断自相碰撞后必须重新划界的问题。

“位”不是物理位置，也不是段末一句话。它是下一步选择空间被前文压缩后形成的优先生成位置。同一段文字可以同时制造多个候选续生位，但它们有优先级；一个候选材料只有在承接高优先续生位时，才算沿本文的生成主路径进入。所谓“转移”，也不是把同一个问题原封不动拖到后面，而是这一轮承接成功以后，旧义务被部分或全部结清，当前全文状态发生改变，于是新的义务 O_{t+1} 在新状态中出现。

因此，作文“长下去”的最小事件可以写成：承接（**inherit/settle**）→改场（**field rewrite**）→转位（**transfer**）。承接回答“这一段在还谁的账”；改场回答“写完它以后，全文现在变成了什么”；转位回答“正因为发生了这个变化，下一步哪一个问题变得比以前更不可绕开”。三步缺一，持续生成的性质就改变。

这使“材料堆积”获得一个比“内容重复”“例子太多”更精确的定义：材料堆积是指新增单位与主题、人物或话题可能相关，甚至自身很精彩，但它没有承接当前优先续生位，也没有因承接而形成依赖性的新续生位。它给文本增加质量或体积，却没有完成生成责任的转移。一个段落可以非常有信息量，同时在持续生成意义上是“零推进”。

九之二、How 型的六条完整路径：同一个“续生位转移”可以从不同断点进入，但不能只做其中一步

三焦点选源并不意味着成文以后只剩三条互不相干的办法。How 型最终要把六条完整路径重新开放出来，因为真实课堂的断点并不总出现在同一个位置。对于一篇已经写到 T_t 的作文，教师首先要判断学生卡在“当前文本看不出问题”“知道问题但找不到下一动作”，还是“写了动作却不知道这一步是否真的改变全文”。不同断点决定从 S、D、E 中哪一处先进入；但无论从哪里进，完整干预都必须重新经过另外两处，否则只是在局部修补。

第一类从 S 出发。 $S \rightarrow E \rightarrow D$ 适合“写了一段却不知道下一段写什么”的学生：先读当前成形文本，冻结它刚制造的未结关系，再把这些关系写成下一步必须满足的条件，最后才选择材料、场景或论证动作。 $S \rightarrow D \rightarrow E$ 则适合学生已经有一个直觉动作，例如“我想举爷爷的例子”，教师先让他写出候选段，再追问这段到底改变了当前文本什么、承接了哪一笔旧账；若无法落到可核验条件，就不让“我觉得这个例子挺好”直接取得入场资格。两条都从当前文本出发，却分别把“条件”或“动作”放在中间。

第二类从 D 出发。 $D \rightarrow S \rightarrow E$ 适合写作者已经动笔但容易越写越散：允许他先完成一个候选新增，然后立即暂停，比较加入前后的全文状态，最后检查这个变化是否兑现旧续生位并形成新的边界。 $D \rightarrow E \rightarrow S$ 则更像写中自检：先提出一个候选动作，先用承接规则把明显的旁枝挡掉，通过以后再真正写入并观察全文状态是否被改写。前者容许“先写后判”，后者强调“先判再写”，二者可以针对不同写作熟练度进行实验比较。

第三类从 E 出发。 $E \rightarrow D \rightarrow S$ 适合教师已经清楚学生缺的是什么，例如“前一段提出了一个反例，但你还没有修改原观点”：先明确当前生成义务，再让学生选择能够兑现它的表达动作，最后看新文本状态是否真正改变。 $E \rightarrow S \rightarrow D$ 则适合修改教学：先用既有义务重读当前稿，定位哪一处状态与义务不匹配，再决定重写、删除、换序或补充的路径。它不要求学生把整篇推翻重写，而是把结构修改压缩到一个可定位断点。

六条路径的共同终点不是某一固定 S、D 或 E，而是恢复“承接一改场一转位”的循环。由此可以得到一条重要教学原则：不要把任何单一提示永久化。对所有学生永远追问“你下一段要写什么”会把路径固化在 D；永远追问“中心是什么”会把 E 变成固定规范；永远让学生“再读一遍自己刚写的”又可能停在 S。教师真正要做的是识别本轮哪个位置先失灵，从那一处进入，然后完成整条路径。

这一点也构成对未来教学实验的更严格要求。若“续生三问”有效，不应只表现在学生自评“更有思路”，而应能看到断点类型与干预路径的匹配效应：

承接缺失的文本接受 E 起点干预应比单纯补素材更有效；问题场不变的文本接受 S 比较干预应比再列提纲更有效；旁枝增殖的文本接受承接应降低无来源新问题。若所有断点对任何提示反应都一样，三焦点路径就没有提供真实区分力，理论应退回一般写作策略。

十、“承接—改场—转位”三步判据：把下一段从“相关”提高到“有生成责任”

（一）承接：下一步必须结清一个在揭示它之前就能冻结的旧义务

承接判据最关键的是时间顺序。不能先看到下一段，再倒推“原来前文需要它”，因为那会把任何结果合理化。研究或课堂编码时，应在 m_{t+1} 出现以前冻结 T_t ，让标注者独立写出此刻 1—3 个“最需要继续处理”的问题。问题不能只写“继续围绕主题”“再举例”，而要具体到可判决的关系，例如“为什么主人公明知会迟到仍没有跑”“前文把自定义成选择数，但刚出现的反例迫使定义怎么修改”“这个证据支持的是相关还是因果”。若下一单位对冻结列表没有实质回应，则承接记 0。

承接不是要求作者按读者猜测写唯一正确答案。多个不同下一段都可能承接同一续生位。比如“为什么主人公没有跑”可以用身体受伤、故意拖延、价值选择、叙述误导等不同路径回答。判据限制的是问题身份，不是答案内容。这样既保持开放写作的创造性，又不把任意联想都算成推进。

（二）改场：新增单位必须改变“现在这篇文章是什么问题”的局部状态

如果一段只把旧问题再说一遍，即使它补充了证据，也未必产生改场。改场要求至少一个可观察维度发生改变：对象范围被缩小或扩大、因果方向被改写、人物关系重新定位、叙述者可靠性改变、概念边界被重新划定、价值冲突增加一个不可兼容条件、证据负担从“有没有”转成“是否足够区分”、时间尺度发生转换等。换言之，读者在这一段以后，不应该仍然以完全相同的方式期待后文。

改场是本文与“论点+例子”流水线分离的关键。第二个、第三个同类例子可以让论据更充足，却未必改变问题场；它们可能是必要的证据累积，但若全文持续依赖这种同型添加，文章会进入“同一问题—同一关系—同一结论”的平移。真正生成通常意味着至少在若干关键节点上，文章改变了自己正在回答的问题。

（三）转位：本次成功不能只关账，还要让新的继续义务依赖性地产生

“转位”最容易被误解成故意不说完。本文恰恰反对用悬念替代机制。新续生位必须满足依赖性：如果删掉或替换 m_{t+1} ， O_{t+1} 就不应以同样强度、同样身份出现。比如前文只在讨论“手机功能多不多”，某段区分“可联系”和“被要求随时回应”以后，“这种要求来自设备还是来自群体规范”才成为新问题；若没有这一区分，后一个问题就不具有同样的必要性。

因此，一篇优秀文章也允许“承接成功但不再转位”的时刻——那就是合法闭合。本文研究的是“怎么继续长”，不是把无休止延长当成价值。文章到了

应该结束的位置，最后一个单位可能把最高优先续生位结清而不再制造新的生成责任，这应被记为闭合成功，而不是 GTR 失败。把“会继续”与“该结束”分开，正是为了避免活性聚合隐喻反客为主。

十一、2×2 辨别格：同样“有关”，为什么只有一格叫持续生成

为了使概念在课堂和研究中可快速使用，可以先不看改场强度，只用两个二元问题做第一轮筛查：本单位是否承接一个已冻结的既有续生位？本单位是否生成一个依赖于本次变化的新续生位？得到如下 2×2。随后再用“改场”作为真正续生格的必要复核。

	生成/转移下一续生位：是	生成/转移下一续生位：否
既有续生位：是	真正续生（再查“改场”）	合法闭合／局部结算
既有续生位：否	旁枝增殖／另起炉灶	材料堆积／平移追加

右下格是最典型的材料堆积：既没有还旧账，也没有生成依赖性新账，只是“和主题有关”。左下格是合法闭合：它不再继续，却完成了现有义务。右上格最危险，因为它表面非常有“创意”：一个新段落突然打开了新问题、新人物或新论点，但这个新口不是从前文生出来，而是作者为了继续写而另开支线。左上格才是持续生成候选；只有再通过“当前问题场确实被改写”这一关，才能记为一次完整续生位转移。

十二、统一账本：续生转移率 GTR 与清点手册

3.9 版要求新对象必须落到同一分母，不能一会儿用段落数、一会儿用教师印象、一会儿又用分数。本文预注册单一主读数 GTR（Generative Transfer Rate，续生转移率）：

$$\text{GTR} = \text{满足“承接=1 且 改场=1 且 转位=1”的实质新增单位数} \div \text{全部实质新增单位数}$$

分母中的“实质新增单位”以句群/段落为主，但可按研究设计统一到命题单元；只排除纯格式、错别字、标点或不改变命题内容的机械改写。合法闭合仍进入分母，但其“转位=0”，因此如果把整篇文章所有段落都计算，结尾会轻微拉低 GTR；为避免把正常结尾当失败，主分析应同时预注册“持续生成窗口”，即从启动段以后到首次被独立判定为整体闭合前的单位，最终闭合段只用于边界报告，不进入持续窗口的主分母。

清点手册分五步。第一步，冻结：在看到下一单位前保存 T_t 。第二步，列债：至少两名标注者独立列出当前 1—3 个最高优先续生位，并给出可判决表述。第三步，揭示并承接：呈现 m_{t+1} ，判断其主要功能是否结清冻结义务之一。第四步，改场：比较加入前后的“下一步候选问题排序”，若对象、关系、证据负担、叙事位置或价值冲突没有可编码变化，记 0。第五步，转位：标注新的 O_{t+1} ，并做依赖性问句——“若删去 m_{t+1} ，这个新问题还会以同样身份和优先级出现吗？”若会，说明它不是本次转位产生。

分歧不能靠讨论到一致后删除。若两名标注者对承接或转位冲突，第三名裁决；仍无法分类则保留“不可判定”标签，并留在总样本账本中单列，不从

分母偷偷移除。这样做的目的不是制造一个漂亮高分，而是让理论真正承担失败。

十二之二、从“相关材料”到“生成责任”：为什么续生链既不能守恒，也不能无限分叉

“责任转移”容易让人误以为作文里始终只有一笔债从前一段原封不动传到后一段。实际不是如此。一个续生位可以被一次新增完全结清，也可以被拆成两个更具体的问题；多个旧义务也可能被一个关键段落重新组织成一个更高阶的问题。因此，本文不要求 O_t 与 O_{t+1} 数量守恒，只要求新义务的身份能够追溯到这次承接与改场。这使理论允许真正的复杂化，却不允许无来源的无限分叉。

例如，一篇论“规则是否限制自由”的文章，前文可能同时留下“规则是谁制定的”“违反规则代价由谁承担”两个候选续生位。一个具体案例若揭示“群体默认规则从未被明文制定，却通过沉默惩罚执行”，它可能把两笔旧账合并成一个新问题：“没有明确制定者的规则，责任如何归属？”这里 O_{t+1} 数量减少，但问题场显著改变，仍是高质量转位。相反，若新增单位同时抛出“科技伦理、家庭教育、社会媒体、心理健康、法律治理”五个宏大议题，却没有哪一个由本次具体变化逼出，这种数量增加应判为旁枝增殖。

因此，课堂上不能把“每段必须留一个问题”当作固定模板。模板化以后，学生会学会在段尾机械写“那么，我们应该怎么办呢？”这种句子具有疑问形式，却没有生成来源。真正的 O_{t+1} 可以没有问号，也可以体现为证据负担、人物后果、概念重新划界或叙事期待。例如叙事中“父亲把门钥匙放回桌上”可能没有提出任何显性问题，但如果前文已经把“钥匙”写成是否被承认属于这个家的标记，这一动作会改变人物位置，并使“谁先开口”成为新的不可绕开关系。形式不是判据，依赖关系才是。

GTR 之所以坚持单一分母，也正是为了避免把“问题越多越好”变成新的量化迷信。主读数只问一个新增单位是否同时完成三件事，而不奖励它制造多少新口子。若未来研究希望分析分叉数量、问题深度或价值转化，可以另设副指标，但不能偷换 GTR 的分母，也不能把“更多新问题”直接当作“更强生成”。本机制要保护的是责任链的可追溯性，不是复杂性的表面繁荣。

从教师诊断角度看，这一约束带来三个不同处方。若承接=0，先回退到前文找丢失的旧账，而不是继续补素材；若承接=1、改场=0，说明学生已经会回答，却还没有把答案中的新差异逼出来，教师应追问“这个答案改变了什么”；若承接=1、改场=1、转位=0，则先判断是否已经到合法闭合，如果尚未完成任务，再问“这个变化让哪一件事第一次变得必须解释”。三个断点对应三种干预，避免一句笼统的“写具体一点/再深入一点”承担所有教学工作。

十三、从“围绕中心”到“段界回写”：课堂中的续生三问

作文课堂不可能要求中学生学习“YAP/TAZ”“翻译验证”或活性自由基聚合。跨学科理论只负责在研究端制造新机制；进入课堂以后必须压缩成低负担

操作。本文把“承接—改场—转位”翻译成三个学生可以在每个段界问自己的问题，称为“续生三问”。

第一问，承接问：前面刚刚逼出了什么还没解决的东西？学生只能写一个具体问题，不能写“继续写主题”。第二问，改场问：我这一段写完以后，事情/观点/人物关系和写之前有什么不同？若回答只能是“多了一个例子”，就提醒自己可能正在平移追加。第三问，转位问：正因为这一段成立，下一步什么变得更不能绕开？如果答案在写这一段之前就同样成立，说明没有真正转位。

教师可以把三问做成“段界回写卡”：每写完一个实质段落，不要求学生立刻修改语言，只在页边写 O_t （我承接了什么） $\rightarrow \Delta F$ （我改变了什么） $\rightarrow O_{t+1}$ （我把继续责任移到哪里）。初学阶段可以一段一卡；熟练后只在自己感觉“开始堆了”“不知道下一段写什么”的地方使用。这样，教学重点从“再想两个素材”转向“找回被上一段制造、但尚未结算的继续位置”。

这也改变了教师的讲评语言。过去教师常说“第二段和中心不够紧”“第三个例子可以删”“这里过渡不自然”。续生框架可以把评价变成可行动的问题：“第一段结尾已经把冲突变成‘想被看见还是怕被看见’，第二段为什么去写天气？它承接的是哪一笔账？”“你这个例子本身很好，但写完以后问题场没有变化；能不能把例子里与前一个例子不同的那一处差异逼出来，让它生成下一步？”这种反馈直接作用于生成机制，而不是只告诉学生成品哪里不好。

十四、机制演示一：同样有四个材料，一条是堆积，一条会自己逼出下一步

以议论文题目“手机让人更自由了吗？”为例。先看堆积版：第一段说手机让联系方便；第二段说移动支付方便；第三段说地图导航方便；第四段说网课方便；结尾说科技是一把双刃剑。四个材料都与“手机/便利”高度相关，也都能支持“手机扩大行动选择”的中心，但第二到第四段的顺序几乎可以任意互换。任何一段的存在都没有显著改变下一段的必要性。它的结构不是“生长”，而是同一槽位不断填充。

再看续生版。 T_1 写：“手机当然增加了选择：我可以随时联系任何人。但当‘可以联系’逐渐变成‘别人默认你应该立刻回’，选择变多为什么反而让人更难不回应？”此时冻结的 O_1 不是“再举手机例子”，而是“区分选择数量与拒绝即时回应的自由”。 T_2 写班级群里的已读、@与夜间消息：它承接 O_1 ，并把问题场从“功能便利”改成“时间自主权”；写完后 O_2 出现：“这种被占用究竟来自设备，还是来自群体规范？”

T_3 进一步比较“自己主动开通知”和“群体通过未回复追问制造压力”：它承接 O_2 ，把对象从设备转向规则来源，新的 O_3 变成“如果自由不是断联，而是保留可撤回的可达性，什么制度/习惯能让这种自由成立？” T_4 再提出“可撤回可达性”作为判据并回看前面案例。这里的四段也都可以找到材料，但材料不是先存在后排列；每一段都在还上一段的账，同时改写下一段的入场条件。

两版之间的差异不能用“第二版材料更高级”解释，因为完全可以使用同样的班级群、支付、导航材料做两版；关键在材料承担的关系不同。堆积版让

每个材料直接向总主题报到，形成“星形结构”；续生版让每个材料先承接刚刚形成的局部义务，再改变全文，形成“义务转移链”。这正是本文要从“内容好不好”中分离出的新变量。

十五、机制演示二：叙事文的持续生成不是“多发生几件事”，而是人物位置不断被前一事件改写

叙事文更能显示“改场”的必要。设题为“那一次，我没有立刻回答”。堆积式写法可以连续写：老师问我、同学问我、妈妈问我，我都犹豫；每一件事都证明“我不善表达”，但事件之间没有生成关系。即使三件事调换顺序，人物认知和关系结构仍基本不变。

续生式写法则不同。第一场，母亲电话里问“你还回来吗”，叙述者没有回答。这里产生 O_1 ：沉默是因为生气、害怕承诺，还是无法确定“回来”指哪一个家？第二场不是换一个“我又沉默”的例子，而是让父亲发来孩子们吃饭的照片；它迫使“回家”的对象分裂成多个归属，承接 O_1 并改写场。于是 O_2 变成：叙述者真正不敢回答的，是地点问题还是身份问题？第三场再让孩子发一句非常普通的“妈妈你什么时候来”，此时一句小话因为前两轮状态改变而获得不同重量，新的续生位落到“我能否在不假装一切没发生的情况下给出一个具体时间”。

这种叙事推进与“设置悬念”不同。悬念关注信息是否被延迟；续生位关注前一事件是否改变后一事件的意义条件。某个事实即使早就告诉读者，只要人物关系和价值位置被前一步改变，它仍可以成为真正的下一步。因而续生框架不要求文学写作机械地每段问一个问句，而要求每个关键场景对下一场景的意义资格产生影响。

十六、最强竞争解释：也许 GTR 只是“连贯性/写作水平”的另一个名字

任何新机制都必须先面对最强替代解释。这里最危险的解释是：高 GTR 的文章之所以看起来更好，只因为作者本来就写得更连贯、知识更多、语言更成熟；“续生位转移”只是把一般写作水平重新编码。若这个解释成立，本文没有增加机制，只多了一套术语。

因此实验必须控制至少四项：主题相关性、段落长度/信息量、一般语言质量、作者或作文总体水平。最关键的是设计转位断裂试验：保留相同主题词、相近长度、相近句法质量和同类材料，只替换一个被判定为承担转位功能的关键单位，使它仍“有关”、仍通顺，却不再改变下一步问题场。如果 GTR 真的捕捉独特机制，那么这种替换应选择性降低独立读者对“下一步最必要处理什么”的一致度，并增加后续段落的可交换性；若总体评分不变、下一步预测也不变，则续生位很可能只是连贯性的重复命名。

另一条机制独特预测来自“依赖性”。如果 O_{t+1} 真是由 m_{t+1} 转移出来，删除 m_{t+1} 后， O_{t+1} 的出现概率或优先级应显著下降；而一个仅由题目或常识预装的问题，即使删除该段也会照常出现。这个反事实差异把“作者/读者本来就会想到”与“这篇文章这一段逼出来”分开。

十七、反向约束：三门源如何反过来禁止我们滥用“续生位”

第一条反向约束来自高分子科学：不是所有“还能反应”的端都属于同一条受控生长路径。活性/可控聚合对终止、转移、端基保真和体系条件有严格要求；迁入作文意味着“还有问题可问”本身没有价值。若新问题只是为了续字数而另起一条无关支线，应记为旁枝增殖，而非转位。尤其在 AI 辅助写作中，模型极容易源源不断提出“还可以讨论”的角度；续生框架必须反过来限制这种无限扩展。

第二条反向约束来自形式化方法：核验器和规范本身不能神化。“续生三问”不是作文质量总评分，更不能成为新的标准答案协议。一篇诗性散文可以有低可判定 GTR，却在声音、节奏或意象组织上极强；一篇说明文也可能为了清晰采用并列展开，不需要每一小节都改变问题场。本文的判据只裁决“持续生成 vs. 材料追加”这一机制，不裁决文学价值、思想深度和语言美。

第三条反向约束来自细胞生物学：状态回写具有情境依赖，相关不等于单因。如果某一段之后学生突然写得更深，不能直接说是续生位造成；教师提问、背景知识、情绪投入、题材熟悉度都可能共同作用。未来实验必须把这些条件登记在案。三条反向约束共同防止一个新概念从“可检验机制”膨胀成万能写作哲学。

十八、边界条件：哪些文章不应被高 GTR 绑架

第一，并列说明与目录型文本。百科条目、产品比较、某些说明文的合理目标就是把平行维度并置；这类文本的“好”可能来自覆盖完整与分类互斥，而不是续生位高频转移。若作文任务要求“介绍三种传统节日习俗”，并列不是失败。本文只在任务本身需要发展、解释、转化、叙事推进或论证深化时具有较强适用性。

第二，抒情与碎片化写作。意象之间的推进可能不是命题式承接，而是音韵、视角、感官通道、象征密度与情绪位置的改写。对此，冻结续生位时必须允许“关系义务”而非只允许问句，例如“前一意象已经把‘冷’从天气改成关系距离，下一单位需要处理这种距离如何进入身体”。如果编码体系只能识别因果问答，就会系统性惩罚文学文本。

第三，合法停顿与留白。续生位可以被留给读者而不在文本内部继续兑现；短篇小说、散文、诗歌常有这种终止。本文的“合法闭合”格因此不能简单等同“所有问题都回答完”。闭合的判定应是：文本不再承担继续展开的内部责任，即使读者仍有开放解释。

第四，考试作文的外部字数压力。当学生被要求写满 800 字，系统可能被迫继续，但外部长度要求不是内部续生位。恰恰在这种场景里，GTR 能诊断一个常见现象：后 300 字仍然主题相关，却开始快速落入右下格。未来研究可以比较“内部续生位耗尽时点”与“机械凑字开始时点”是否一致。

十八之二、三种文体的判决性预测：同一机制不会要求所有文章长成同一种样子

如果“续生位转移”只是议论文的一套新术语，它很难成为作文持续生成的一般机制。本文因此给出三种文体的不同预测，而不是把同一模板强压过去。议论性文本中，续生位最常落在概念边界、证据负担、反例冲突与因果解释上。高 GTR 不要求每段“观点—例子—分析”格式一致，反而预测关键段落会不断改变论证所需要证明的东西：一个反例进入以后，原命题的适用范围要收缩；一个新区别成立以后，后续证据必须重新分组。若段落只是向同一中心不断提交独立证据，其可交换性会较高。

叙事性文本中，续生位更多落在人物位置、行动后果、信息权、关系承诺和读者预期。这里的“承接”不必表现为回答问题，而可以是让前一行行动留下的后果在下一场景中取得效力。判决性预测是：删掉一个真正承担转位的关键场景，不仅少了一件事，还会使后续场景的意义发生系统漂移；而删掉堆积事件，故事可能变短，却不改变人物后来为什么必须那样行动。换言之，事件重要性不能只用戏剧性衡量，而要看它是否改写后续场景的入场条件。

抒情/散文性文本中，续生位可能落在意象关系、感官通道、语气、观察尺度或自我位置的变化。一个意象再次出现并不自动算推进；若第二次出现只重复第一次的象征意义，它只是强化。只有当第一次建立的意义被新的情境扭转、压缩或重新命名，随后其他意象的解释条件也跟着变化，才构成改场与转位。因此，本机制预测高质量抒情发展常具有“同一意象反复出现、但每次回来都不再是原来的东西”的结构，而不是简单追求意象数量。

三种文体共享的不是外在结构，而是同一个反事实：如果拿掉这个单位，后面某个高优先关系是否失去它现在的必要性？议论文里可能失去一项证明责任，叙事文里可能失去人物行为的必然性，散文里可能失去某个意象转义的条件。若答案始终是否定的——后面照样可以原封不动继续——那么无论本段语言多漂亮，它在“持续生成”这一变量上都更接近可交换材料。

这也给不同年龄段留下可检验的发育预测。初学写作者可能主要表现为“主题相关性先建立、局部承接较弱”，即很多材料都能归到题目下但互相可换；随着写作能力提升，承接率可能先提高，学生开始知道上一段需要下一段做什么；再往后才是改场与转位增强，作者能够让已经写出的东西改变自己原先计划。若未来纵向数据不呈现任何类似层级，而 GTR 只与词汇量或篇幅同步上升，本文关于“生成责任”独立发展的判断就应被削弱。

十九、10.5 真跑一条：对 ArgRewrite V.2 做预注册字段审计，得到一个不漂亮但重要的负结果

3.9 版要求在成文前把新机制拿到真实公开对象上“跑一条”，并允许不利结果进入正文。本文选择公开的 ArgRewrite V.2，因为它不是简单成品评分库，而是专门记录议论文跨稿修订的语料。项目仓库与论文说明：语料包含 86 名写作者的三个草稿，共 258 个文本文件；有 172 个 Excel 修订标注文件，覆盖 Draft 1→2 与 Draft 2→3 两轮；还保存调查、专家反馈、标注核验与

各稿得分。[23] 其标注进一步给出句子索引、文本、跨稿对齐、修订目的和操作；修订目的可区分主张/观点、反驳/保留、理由/推理/支持、证据、一般内容发展、组织和精确性等，操作则区分增加、删除与修改，并有句内层面的修订元组。[23]

在打开语料字段说明以前，本文先预注册“直接计算 GTR 所需的最小字段”：A. 写作过程中的顺序中间状态 T_t ，而不只是完整 Draft 1/2/3；B. 在下一单位出现以前冻结的 O_t 候选；C. m_{t+1} 与哪一项 O_t 发生承接的链接；D. m_{t+1} 后新出现的 O_{t+1} 及其对 m_{t+1} 的依赖性；E. 至少能区分“合法闭合”和“没有转位”的边界标记。若缺 A—D 中任一项，主 GTR 不得用主题相关、修订目的或教师评分代理。

实际字段审计的结果是不利的：ArgRewrite V.2 有修订语义，却没有作文发生当时的续生义务链。三稿文本只能告诉我们某个句子在下一稿被增加、删除或修改；Revision Purpose 能告诉我们这次修订更接近证据、推理、反驳或组织；但它没有记录写作者在写入该句之前，前文当时生成了哪一个“必须继续处理”的 O_t ，也没有记录该句写完 O_{t+1} 如何重新排序。更重要的是，跨稿对齐反映的是 revision trajectory，不等于初稿内部从第一段到最后一段的实时生成轨迹。

因此，本次真跑不能给出一个漂亮的“高分作文 $GTR=0.73$ 、低分作文 $=0.41$ ”。如果现在硬算，就只能让研究者事后阅读完整文章，再倒推出所谓 O_t ；那会违反本文自己最重要的冻结原则。本文目前不能利用 ArgRewrite V.2 直接计算 GTR，也不能声称“续生位转移已经得到大样本验证”。这个失败必须保留，因为它准确指出了现有写作语料缺的不是更多分数，而是一种时间定向字段：前一步产生了什么生成责任，下一步结清了什么，又把责任移到哪里。

负结果同时给出一个可执行的数据建设方向：未来作文过程语料不必记录学生“脑中所有想法”，只需在段界或关键停顿点做轻量冻结——让学生或独立读者写下当前最必要的 1—3 个续生位，再继续写；系统保存版本差异。这样就能把不可见的“继续感”变成可审计链条，而不需要把写作还原成全程 think-aloud。

二十、可证伪方案：什么结果出现时，“续生位转移”应该撤回

第一条，预测不增益证伪。在控制主题相关性、语言质量、长度、年级与总体作文水平以后，GTR 若不能提高对后续段落顺序、不可交换性或下一步必要问题一致度的预测，说明它没有超越一般连贯性指标，应撤回独立机制地位。

第二条，消融不敏感证伪。对被标为高转位功能的单位做保持主题与语言质量的替换/删除，如果读者对下一续生位的判断不发生系统变化，说明所谓 O_{t+1} 并不依赖本次转位，而可能来自题目先验或常识脚本。

第三条，顺序不敏感证伪。若高 GTR 文本的若干核心段落随机对调后，读者理解、下一步预测与总体结构感几乎不变，则“义务转移链”并未对顺序形成真实约束，理论退化为普通主题聚类。

第四条，课堂干预无特异效应证伪。把学生随机分为“素材扩充”“普通提纲/衔接训练”“续生三问”三组，控制练习时间与教师反馈量。如果续生三问只提高字数或流畅度，而不降低右下格材料堆积、也不提高段落顺序依赖与问题场变化，则机制训练没有命中特异变量。

第五条，文体边界反噬证伪。若在需要持续发展的叙事与论证文本中，GTR 与专家对“推进/发展”的判断长期负相关，且这种反向不能由闭合、留白或文学碎片化解释，则编码定义本身有误，应重新划界。理论不是因为命名新就享有保留权。

二十之二、最小实验设计：不用等“完整作文大数据”，也能先让机制承担一次小规模失败

真正检验续生位不必一开始就建设数万篇作文。最小实验可以选取同一题目、同一年级、长度相近的 30—60 篇过程作文，在每个段界保存版本。每到一预注册段界，先不给标注者看下一段，只呈现 T_t ，让其列出 1—3 个最高优先续生位；随后揭示 m_{t+1} ，编码承接、改场和转位。为了避免“好文章看什么都像有道理”的光环效应，标注者不能看到教师总评分，也不能看到作者身份。主分析只使用预注册持续生成窗口，闭合段、不可判定段和机械修改分别报告。

同一批文本再制作一组“结构保真但转位破坏”的反事实版本：尽量保持字数、主题词、句法复杂度与局部事实，把承担转位的段落替换为一个主题相关却不改变问题场的平行材料。新的读者只看到其中一个版本，回答“下一步最需要处理什么”并对后续三段做顺序恢复。若续生位机制成立，原版应表现出更高的下一步判断收敛与顺序恢复，而两版的词汇丰富度和局部流畅度可以相近。这样能把机制从“文章整体更好”中拆出来。

第三个实验专门检验教学因果。学生随机进入三种等时训练：素材扩充组只学习快速联想与例证库；结构提示组学习常规提纲、主题句与过渡；续生组只学习段界回写卡。训练前后使用新题，不允许沿用课堂范文。若续生组只把作文写得更长而没有降低右下格堆积比例，理论失败；若它降低堆积却同时显著损伤任务完成度或导致大量故意留悬念，也说明干预反噬。只有当生成责任链改善、总体任务完成不下降、且效果在陌生题目中仍可观察，才值得进入更大规模教学研究。

这套最小设计保留一个重要原则：不要把理论自己造出的指标当作唯一成功标准。GTR 是机制读数，不是终极质量分。实验必须同时保留外部评价，如专家对推进、完整性、思想发展和文体適切性的盲评，以及学生完成时间和认知负担。若 GTR 上升而外部推进判断下降，不能宣布“学生只是还没适应新方法”；这种冲突本身就是反证材料。一个能够进入语文教育的理论，必须允许课堂结果迫使它缩小适用范围。

二十一、对作文教学的改写：从“给材料”转向“维护生成责任”

第一，备课单位改变。教师不再只准备“这个题有哪些素材”，而要准备“学生写到哪些典型状态时，会出现哪些有价值的续生位”。同一个题目可以

有完全不同的生长路径；教师的专业性不在于提供标准提纲，而在于识别学生当前文本真正产生了什么差异，并用最小提示帮助其看见。

第二，作文指导的时间点改变。传统指导大量发生在写前：审题、立意、选材、列提纲；本文的机制要求把一部分指导移到写中段界。因为续生位是由实际写出的文本生成的，写前不可能全部预知。一个好提纲只能给初始方向，不能替代每轮“写成以后状态变了”的回写。

第三，修改教学改变。修改不只是把表达变准确，也可以被理解为重新恢复被打断的续生链。学生常说“我这篇写散了”，教师可以让他逐段回溯：哪一段开始 O_t 没有被承接？哪个新材料开了旁枝？哪一处虽然回答了旧问题，却没有改场，导致后面只能继续堆同类证据？这样，“结构修改”从大幅重排变成可定位的断点修复。

第四，评价改变。现有评分常把内容、组织、语言、逻辑作为成品维度；续生位框架增加的不是第 N 个分项分数，而是一种过程诊断单位。两篇最终都很连贯的文章，可能一篇靠事前模板填充完成，一篇在写作中不断由已形成文本改写下一步。本文并不宣称后者必然更好，但如果教学目标包括让学生学会自主持续生成，那么两者必须能够被区分。

二十二、AI 辅助写作中的新风险：机器最擅长补材料，也最容易抹掉续生位

生成式 AI 让这个问题更紧迫。语言模型可以在任何段落后继续生成语法通顺、主题相关的下一段，因此“写不下去”表面上很容易解决。但模型的高可生成性恰恰可能遮蔽学生自己的续生位：当学生还没识别 O_t ，系统已经给出五个“可以继续写的方向”，写作就从前文生成转向候选菜单选择。结果是文本变长，作者对“为什么此刻写这一段”的控制反而下降。

续生框架提供一种比“禁止 AI”更细的接口原则：AI 在给下一段建议前，先要求学生或模型显式列出“前文刚产生的 1—3 个续生位”，再让候选建议逐一标注自己承接哪一项、会改写什么、将生成什么新义务。系统甚至可以故意提供一个“主题相关但不承接”的诱饵，让学生辨别材料堆积。AI 的角色由“续写器”变成“续生位显影器”。

更严格的实验可以比较两种 AI：A 组直接给 next paragraph；B 组只给 O_t 候选和承接检查，不代写正文。若 B 组最终文本的 GTR、独立续写能力与断点修复能力更高，而总耗时可接受，就支持“外部帮助应保护生成责任而非替代生成责任”的教学设计。这里的核心变量不是 AI 使用与否，而是谁拥有续生位的发现、承接和转移权。

二十三、理论意义：把“篇章结构”从静态关系推进到时间中的责任传递

篇章语言学和 RST 已经能精细描述成品中句段之间的关系：原因、证据、让步、背景、阐释、对比等。[17-19] 本文并不试图替代这些理论。它提出的是一个时间向度：当一个关系还没有完整出现时，前文如何先生成一种“待兑现关系”，下一单位如何把它兑现，并在兑现后产生新的待兑现关系。

换言之，成品结构是已经结算的账；续生位研究的是账如何一笔一笔被开出、结清和转移。

这个时间向度还能解释为什么“结构完整”与“生成感”可以分离。模板作文可以在写前拥有完整树形，写作只是把预设槽位填满；读者最终仍能看到清楚结构。另一篇文章可能在写作中真正经历问题转化，最终结构反而看起来很自然，甚至很简单。若研究只看成品，两者容易被合并；若关注续生链，生成机制就不同。

因此，本文对“作文结构”的最小改写不是增加一种结构类型，而是把结构从名词变成动词：结构不是材料被摆在什么位置，而是每一次位置变化如何改变下一次可进入的关系。这也是为什么本文选择 How 题而非 What 题：续生位不是作文的本体定义，它只解释已经启动的文本如何在时间中继续形成。

二十四、结论：作文真正“长下去”，增长的不是材料，而是可被转移的生成责任

本文从一个极普通的课堂现象出发：学生明明还有很多材料，却越写越像堆积。通过高分子科学、形式化方法与程序验证、细胞生物学三门远距理论的 How 型三焦点碰撞，本文把“持续生成”从三个旧直觉中分离出来：它不是材料储备增加，不是开放端永远存在，也不是局部连贯一次次通过。三家互相拆台后留下的第三物，是“续生位转移”。

这一机制的核心句可以压缩为：作文持续生成，是前一文本产生的生成义务被下一单位承接；承接使当前全文问题场发生可识别变化；这一变化再产生一个依赖于本次变化的新生成义务，并把继续责任转移到下一位置。于是，真正的生长不发生在“又多了多少字”，而发生在“下一步的必要性如何被上一轮重新制造”。

由此，“材料堆积”也不再只是教师的审美抱怨，而成为可判决失败：新增单位可以相关、正确、具体、甚至漂亮，却没有还当前的账，也没有把新账从本次变化中生出来。课堂上的“续生三问”、研究中的 GTR 与转位断裂试验，为这一差别提供了第一套可操作接口。

因此，本文宁可把结论写窄，也不把“持续生成”扩成一切好作文的总称。理论只负责解释启动以后生成责任怎样沿文本时间向前移动；至于文章最终是否深刻、优美、真诚、适切，仍需由其他评价系统分别裁决。这样的自限不是退让，而是保证新变量能够真正被实验击败。

从方法论上看，这篇文章的赌注也因此变得单一：它不赌“续生位”能解释所有作文质量，只赌一件更窄的事——当一篇文本确实在持续发展时，后一个实质单位对前一轮生成义务的承接、对当前问题场的改写以及对下一轮义务的依赖性转移，会提供一种不能被单纯主题相关性和表面连贯性完全替代的顺序信息。如果未来数据证明段落在控制这些旧变量以后仍高度可交换，或者续生编码不能改善对下一步必要关系的预测，本文的核心命题就应当撤回，而不是继续通过扩大定义保护自己。

更重要的是，本次公开语料真跑没有给理论送来一个好看的相关系数。ArgRewrite V.2 的字段审计表明，现有修订语料已经能够细分“改了什

么、为何改”，却仍缺少“写这一段之前什么必须被继续、写完以后什么才变得必须继续”的时间定向字段。[23] 因而本文的结论仍处在可检验理论阶段。它的下一步不是再写一篇更漂亮的解释，而是建设能记录续生义务链的过程语料，并允许数据把“续生位转移”推翻。只有到那时，“作文怎样长下去”才真正从教学经验进入可反驳的语文教育学问题。

参考文献

- [1] 付自文 + Claude. SDE 多学科通融创新法：从 0 到 1 不是灵感，是一门可以重复的手艺（3.9 方法框架）[EB/OL]. 学科通融·方法论特稿, 2026.
- [2] 付自文. 高分子科学／Polymer Science·新思想前沿[EB/OL]. SDE Universes, 2026.
- [3] 付自文. 形式化方法与程序验证·新思想前沿[EB/OL]. SDE Universes, 2026.
- [4] 付自文. 细胞生物学·新思想前沿[EB/OL]. SDE Universes, 2026.
- [5] Szwarc, M. “Living” Polymers. *Nature*, 1956, 178: 1168–1169. DOI: 10.1038/1781168a0.
- [6] Kamigaito, M., Ando, T., & Sawamoto, M. Metal-Catalyzed Living Radical Polymerization. *Chemical Reviews*, 2001, 101(12): 3689–3746. DOI: 10.1021/cr9901182.
- [7] Storey, R. F. Fundamental Aspects of Living Polymerization. In: *Controlled/Living Polymerizations*. Royal Society of Chemistry, 2013. DOI: 10.1039/9781849737425-00060.
- [8] Lopes, N. P., Menendez, D., Nagarakatte, S., & Regehr, J. Provably Correct Peephole Optimizations with Alive. *Proceedings of PLDI 2015*, 22–32. DOI: 10.1145/2737924.2737965.
- [9] Dupont, S., Morsut, L., Aragona, M., et al. Role of YAP/TAZ in Mechanotransduction. *Nature*, 2011, 474: 179–183. DOI: 10.1038/nature10137.
- [10] Engler, A. J., Sen, S., Sweeney, H. L., & Discher, D. E. Matrix Elasticity Directs Stem Cell Lineage Specification. *Cell*, 2006, 126: 677–689.
- [11] Emig, J. *The Composing Processes of Twelfth Graders*. NCTE Research Report No. 13. Urbana, IL: National Council of Teachers of English, 1971.
- [12] Murray, D. M. *Teach Writing as a Process Not Product*. The Leaflet, 1972, 71(3): 11–14.
- [13] Elbow, P. *Writing Without Teachers*. New York: Oxford University Press, 1973.
- [14] Perl, S. The Composing Processes of Unskilled College Writers. *Research in the Teaching of English*, 1979, 13(4): 317–336. DOI: 10.58680/rte201117867.
- [15] Flower, L., & Hayes, J. R. A Cognitive Process Theory of Writing. *College Composition and Communication*, 1981, 32(4): 365–387.
- [16] Bereiter, C., & Scardamalia, M. *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1987.

- [17] Mann, W. C., & Thompson, S. A. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*, 1988, 8(3): 243–281. DOI: 10.1515/text.1.1988.8.3.243.
- [18] Halliday, M. A. K., & Hasan, R. *Cohesion in English*. London: Longman, 1976.
- [19] van Dijk, T. A., & Kintsch, W. *Strategies of Discourse Comprehension*. New York: Academic Press, 1983.
- [20] Zeigarnik, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, 1927, 9: 1–85.
- [21] Ovsiankina, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychologische Forschung*, 1928, 11: 302–379.
- [22] Ghibellini, R., & Meier, B. A Meta-Analysis of the Zeigarnik and Ovsiankina Effects. *Humanities and Social Sciences Communications*, 2025.
- [23] Kashefi, O., Afrin, T., Dale, M., Olshefski, C., Godley, A., Litman, D., & Hwa, R. ArgRewrite V.2: An Annotated Argumentative Revisions Corpus. *Language Resources and Evaluation*, 2022, 56(3): 881–915. DOI: 10.1007/s10579-021-09567-z.
- [24] Afrin, T., Kashefi, O., Olshefski, C., Litman, D., Hwa, R., & Godley, A. Effective Interfaces for Student-Driven Revision Sessions for Argumentative Writing. *Proceedings of CHI 2021*.

附录 A：续生位编码的最小预注册表

1. 分析单位：预先规定为段落、句群或命题单元之一；不得看完结果后为追求高 GTR 改单位。
2. 冻结时点： mt_{+1} 出现之前保存 T_t ；若使用已有成品进行回溯，只能标为“事后模拟编码”，不得与前瞻冻结数据混报。
3. O_t 记录：两名标注者各列 1—3 个当前最优先生成义务；问题需包含具体对象和关系，不接受“继续主题”“继续论证”。
4. 承接判据： mt_{+1} 的主要新增内容是否对至少一个冻结 O_t 产生可辨认结算；仅词汇重合不算。
5. 改场判据：加入 mt_{+1} 后，下一步候选问题的对象、关系、尺度、证据负担、人物位置、叙述视角或价值冲突至少一项发生可编码改变。
6. 转位判据：存在 O_{t+1} ，且删除/替换 mt_{+1} 后该 O_{t+1} 的身份或优先级应明显下降。
7. 闭合：承接=1、转位=0 且独立判断认为文本已完成当前任务，可标“合法闭合”，不当作材料堆积。
8. 不可判定：保留在账本中，单列比例；不得为提高一致率从数据集中删除。
9. 主读数：GTR 只在预注册持续生成窗口计算；其他评分、字数、连贯性指标作为协变量或结果变量，不得替代主分母。

10. 证伪优先：先报告预测不增益、消融不敏感、顺序不敏感等失败，再报告支持性结果。

编者补节·第四章

一·需要补的一位：问题之下的问题

本章的核心对象是「续生位」：当前文本边界上一个尚未结清、但已经被前文具体生成的关系义务；下一个单位必须承接它、改变问题场、并转出一个依赖于本次变化的新义务。

这个机制在形式语用学里有一位正主，而本章一次未提：问题之下的问题 (Question Under Discussion, QUD ; Roberts, C. 1996/2012, *Information structure in discourse: Towards an integrated formal theory of pragmatics*, Semantics and Pragmatics, 5, article 6, doi:10.3765/sp.5.6) 。

它的原文机制与本章几乎逐条对应：话语被建模为一个围绕「当前正在被回答的问题」组织的记分板；一个问题入栈即构成回答它的承诺，若未被回答且未被判定为不可回答，它一直留在栈上；一个话语单位的相关性，由它是否处理当前问题决定。此外，这一支已经有面向篇章（而非对话）的标注实践——每一个句子被读成对一个多半是隐含的、由前文触发的问题的回答 (Büring 2003 ; De Kuthy et al. 2018 ; Riester 2019) 。

不补这一条，读者会当场把「续生位」还原成 QUD 栈，本章的新意归零。

分离线（三条，本章其实都已具备，只是没有对着这位写）。

1. 判据的严格程度不同。那一支的相关性判据是「是否处理当前问题」；本章要求承接、改场、转位三件同时成立，缺一不计入主读数。换言之，一个在那一支看来完全相关的续写单位，在本章可能被判为「合法闭合」或「平移追加」。
2. 冻结纪律不同。那一支的问题可以由分析者在读完整段文本之后重构；本章明令在下一单位出现之前冻结当前义务列表，并规定回溯性编码只能标为「事后模拟」，不得与前瞻数据混报。这条纪律是本章方法论上的主要投入。
3. 对象不同。那一支的栈是一个语义—语用对象（问题的集合）；本章的续生位是一个可被删除检验的文本事件——删掉本次新增单位，那个新问题是否还会以同样的身份与优先级出现。

判定形式。取同一批过程写作数据，一路按篇章 QUD 标注，一路按本章的三步判据编码。若两套编码的结果在预测下一步、预测顺序不可交换性上不可区分，本章应并回那一支，「续生位」只是 QUD 的一个别名；若本章编码在「改场」与「转位」两项上产生了那一支产生不了的区别，本章成立。

二·与第六章的一处共用格

本章的「合法闭合」（承接 = 1、转位 = 0，且独立判断认为文本已完成当前任务）与第六章的退场闭合是同一格。本章把它当作一个不该被误判为失败的边界情形；第六章把它当作研究对象本身，并给了整套义务账本与阈值。

编者不认为这构成重复：本章在持续生成窗口内工作，闭合段只用于边界报告，不进入主分母；第六章的整章都在闭合段里。但读者应当知道，两章在此处是接续关系——本章的窗口在哪里结束，第六章的工作就从哪里开始。第六章第二十二节已经给出了这个交接的可观察信号：结清义务数开始低于新激活义务数。编者建议读者把这句话读成两章之间的接口条款。

三·编者提示：本章的记号没有兑现

本章通篇使用了一套记号——当前文本状态、当前义务、下一个新增单位、新义务——并把三步判据挂在这套记号上。但这套记号始终没有被形式化：义务的身份如何界定、两笔义务何时算同一笔、义务的合并与拆分如何计数，正文只有散见的自然语言约定。清点手册解决了编码流程，没有解决记号本身。

后果是具体的：本章允许义务数量不守恒（一次承接可以把两笔旧账合并成一个新问题，也可以把一笔拆成两笔），却又要求新义务的身份「能够追溯到这次承接与改场」。在没有身份判据的情况下，这两条之间留着一个口子——编码者可以在事后决定两笔义务是不是同一笔，而这恰好是本章第十节反复防范的那种自由度。

本章的字段审计得出「现有语料不能直接计算」这一负结果，处理是对的；但读者应当明白，即便字段齐了，这个记号问题仍然要先解决。

第五章

未改之句如何被改写

—— 作文发生的稿际余载机制

未改之句如何被改写

作文发生的稿际余载机制

来自埋藏学、第三体摩擦学与位置信息理论的跨学科解释

“作文不是文字不断增加的结果，
而是篇章角色不断被重新分配的结果。”

理论建构稿·含可证伪的实证研究方案

2026年8月

摘要

“作文是如何发生的”通常被回答为计划、表达、回读与修订的递归过程，或被描述为若干草稿逐步接近成品的变化史。这些解释已经充分证明写作不是线性誊录，却仍留下一个不易被既有分析单位捕捉的问题：当某句话的字面形式完全未改时，它会不会因为别处的增删、移位与弃置而获得新的篇章角色？如果答案为是，那么作文的发生便不能只用新增了什么、删除了什么或改写了什么来说明，还必须解释未改文字为何也会被“结构性改写”。本文以结构映射而非表面比附为方法，将埋藏学关于记录由保存与破坏共同生成、死亡遗存反过来影响现存生态的认识，第三体摩擦学关于脱落物在接触界面承载、剪切、循环与排出的机制，以及形态发生素与位置信息理论关于位置依赖的命运分化和删减后的模式调节结合起来，提出“稿际余载”概念。稿际余载是文本成分退出当前可见稿后，其尚未解决的论证功能、语义承诺与位置约束在版本之间持续存在，并迫使其他文本成分重新分配篇章工作的效应。本文进一步提出“零字面角色迁移”：在相邻版本中保持字面不变的语句，由于关系场发生变化而从背景转为论点、从证据转为问题触发器、从解释转为反驳等。由此，作文被重新界定为：一次增删不再只改变被操作的文字，而开始释放稿际余载，使其他未被改动的语句发生位置重定和角色分化，并最终形成相互约束的篇章结构的过程。本文系统区分该命题与过程写作、修订研究、文本发生学、分布式认知、功能语言学修订研究及按键记录研究，给出稿际余载的来源、输运、吸收、再循环、排出与稳定化模型，提出零字面角色迁移率（ZRM）及“同一现稿、不同历史”的随机实验设计。本文的理论论证已经形成，但因尚未完成预注册实验，所有因果判断均保留为待检验命题。其意义在于把作文研究的最小解释单位从“被改动的字符串”推进到“跨版本迁移的篇章职责”，并为写作教学、过程评价和版本感知型人工智能写作工具提供新的观察框架。

关键词：作文发生；稿际余载；零字面角色迁移；修订；文本发生学；第三体；位置信息

How an Unchanged Sentence Is Rewritten

Interdraft Residual Load as a Mechanism of Composition

Abstract

Existing accounts of composing have convincingly rejected linear transcription and described writing as recursive planning, translating, reviewing, and revising. Yet they do not fully explain a peculiar event: a sentence may remain literally unchanged while its rhetorical role changes because material elsewhere has been added, deleted, moved, or abandoned. This article develops a theory of *interdraft residual load* through relational structure mapping across three distant fields: taphonomy, third-body tribology, and morphogen-based positional information. Interdraft residual load refers to the unresolved argumentative function, semantic commitment, and positional constraint that persist after a textual unit leaves the visible draft and that compel other units to take on redistributed discourse work. Its key observable consequence is *zero-literal rhetorical-role migration*: an aligned unit retains the same wording across versions but changes from, for example, background to claim, evidence to problem trigger, or explanation to rebuttal. Composition is therefore redefined as the process by which local textual operations release residual loads, relocate unchanged units, differentiate their roles, and stabilize a mutually constraining discourse pattern. The article distinguishes this proposal from process theory, revision studies, genetic criticism, distributed cognition, functional approaches to revision, and keystroke logging. It specifies a source-transport-absorption-recirculation-ejection model, proposes the Zero-Literal Role Migration rate (ZRM), and outlines a randomized “same current draft, different history” experiment. The paper is theoretical: it establishes a falsifiable research program but does not claim empirical confirmation. Its contribution is to shift the explanatory unit of writing research from modified strings to rhetorical obligations that migrate across versions.

Keywords: composing process; interdraft residual load; rhetorical-role migration; revision; genetic criticism; positional information

一、问题的提出：作文何时不再只是“字数增加”

作文的成品容易制造一种视觉错觉。读者面对的是一页在空间上同时展开的文字，于是倾向于把它理解为作者先有一个完整意思，再把意思依次安放到句子中的结果。过程写作研究已经有力地拆解了这种错觉。Emig 以个案和出声写作方法进入高中生的实际写作活动，Perl 揭示基础写作者的回返与递归，Flower 和 Hayes 则将写作描述为由计划、转译、回读等过程构成、受目标网络调节的复杂认知活动[1-4]。Sommers 进一步指出，成熟写作者不会把修订仅仅理解为词语替换，而会反复改变作品的形状和意义

[5]。自此，“作文不是直线生成的”已经不是有待争论的新命题，而是写作研究的共同起点。

然而，非线性并不自动等于机制解释。若研究仍以新增、删除、替换、移位的文字为唯一观察对象，那么分析虽然能够细致记录操作，却可能忽略操作之外的结构后果。例如，某段作为主张的文字被删除后，原本只承担例证功能的一句话可能在新稿中成为全文最醒目的问题入口；某个未改动的结论句，可能因为前面增加反方观点而从总结变成反驳；某句保持原样，但它与段落中心、读者预期及相邻句的关系全部改变。在这些情形中，被编辑器标亮的“差异”发生在甲处，真正承担新篇章工作的却可能是未被标亮的乙处。

这一现象带来三个层次的问题。第一个是描述问题：在版本对照中，如何命名“字面不变而功能改变”的单位？第二个是因果问题：为何别处的一次增删能够改变它的篇章命运？第三个是发生问题：一组句子究竟在什么时刻由并置材料转化为相互制约的作文结构？现有研究常把第三个问题回答为作者形成了目标、完成了计划、写出了初稿或进行了深层修订。这些答案各有根据，却大多把变化定位于作者内部的认知活动，或定位于可直接观察的修改事件。本文尝试把解释焦点移到二者之间：已经离开现稿的文本成分，其未完成的职责如何继续作用于当前稿，并通过其他文字重新落位。

本文的核心主张是：作文不是文字不断增加的结果，而是篇章角色不断被重新分配的结果。更精确地说，作文发生于一次增删不再只改变被操作的文字，而开始释放“稿际余载”，使其他未被改动的语句发生位置重定和角色分化，并最终形成相互约束的篇章结构之时。所谓“稿际”，强调作用跨越相邻版本或可恢复的写作状态；所谓“余载”，不是被删文字的神秘幽灵，也不等同于作者对旧句的清晰记忆，而是该文本成分曾经承担、退出后仍未得到处理的论证功能、语义承诺和位置约束。它可以被后续句子吸收，可以在多个位置间再循环，也可以因作者明确放弃有关问题而被排出。

这一定义有意避开“删除的文字仍然有用”之类过宽判断。被删内容当然可能影响作者，旧稿也当然可能影响新稿；文本发生学对此已有丰富描述。本文真正要求解释的，是一个更狭窄、因而更可检验的效应：当语句 u 在版本 V_t 与 V_{t+1} 中字面一致， u 的篇章角色 $R(u)$ 是否会由于 u 之外的变化

而改变；这种改变是否与某段先前被写出、随后退出现稿的历史有关；在控制当前稿、接触内容、词汇启动和版本可见性之后，这种历史效应是否仍然存在。只有当这些问题得到可重复的证据支持，“稿际余载”才不只是一个富有解释力的比喻，而能成为写作研究的有效构念。

因此，本文属于理论建构论文，而不是已经完成验证的实证报告。文章首先重审过程写作、修订研究、文本发生学、分布式认知与数字写作过程研究的解释成就和剩余空白；继而从埋藏学、第三体摩擦学和位置信息理论中抽取关系结构，建立稿际余载机制；随后通过三个构造性案例展示其可观察后果；最后提出指标、标注体系、随机实验和反证条件。本文所说“证明”只指概念推导能够形成内部一致且有区分力的命题，不指因果效应已经得到数据确认。

二、既有解释的成就与盲区

（一）从线性阶段到递归过程：写作研究已经解释了什么

早期过程研究的突破，在于把研究对象从完成的作文转向作文活动本身。Emig 的研究表明，写作者的实际行为不能被“构思—书写—誊清”的课堂程序充分代表[1]。Perl 对不熟练大学写作者的研究同样发现，写作包含反复回看题目、已写文本和自身“感觉到的意义”的循环，低水平表现并不意味着过程简单[2-3]。Flower 和 Hayes 的模型把任务环境、长期记忆和写作过程置于同一系统，强调计划、转译、回读可以相互嵌套，作者还会在写作中生成并修订目标[4]。Bereiter 与 Scardamalia 以“知识陈述”和“知识转化”的差异说明，成熟写作并非从记忆中依次提取内容，而是在内容问题空间与修辞问题空间之间往返，使解决一个问题的行动改变另一个问题[6]。Hayes 后来的模型继续加入工作记忆、动机、情感、身体和文本环境，扩大了写作系统的边界[7]。

这些理论足以否定“作文是脑内成品的外部抄写”。它们也容纳了写作历史的反馈：已经写出的文本会成为新的任务环境，作者回读它、评估它，再改变后续计划。由此可见，本文不能把“现有文本影响后续写作”宣称为新发现，也不能仅凭写作递归就证明稿际余载存在。过程模型的真正限制不在于它们否认反馈，而在于其核心变量多处于较高级别，例如目标、计划、记忆资源和控制过程。它们能够说明作者为何回头，却较少给出一

个可对齐到版本单位的机制，用来解释未改语句如何因远处删改而发生角色迁移。

换言之，过程理论主要回答“作者在做什么认知工作”，本文则追问“篇章工作在文本单位之间怎样迁移”。前者以写作者的活动组织为中心，后者以跨版本职责的分配为中心。二者不是替代关系：稿际余载若成立，应当被理解为过程理论在文本关系层面的一个中尺度机制。

（二）修订研究：对变化的精细分类为何仍不足够

修订研究最接近本文的问题。Sommers 把修订界定为写作过程中持续发生、由线索触发的一连串改变，并将增添、删除、替换和重排置于词、短语、句子和主题等层级考察[5]。Faigley 与 Witte 区分表层变化和意义变化，为跨稿比较提供了可操作的分类[8]。Fitzgerald 把修订概括为写作者在任何时点对计划或已写文本作出的改变，这一界定把前文本计划和成文后的编辑纳入同一范围[9]。此后的产品导向与过程导向研究不断细化修改的位置、时机、幅度、原因和效果。

数字环境使修订变得更易记录，也更难直观察觉。纸面手稿保留涂改痕迹，电子文档却常以一份“干净现稿”覆盖复杂历史。按键记录工具可以为每次按键、鼠标移动和光标位置附加时间戳。Leijten 与 Van Waes 指出，这类数据有助于观察停顿、回返与修改，但仅靠细粒度日志很难直接推断背后的认知过程，需与出声思考、眼动或回溯访谈结合[10]。Conijn 等构建的过程修订数据集包含 65 名学生和 7,120 次修订，为每次修订提供人工与自动特征；值得注意的是，数据集中“加工过程”和“触发因素”两个属性因现有日志无法取得而未被提取[11]。这恰好表明：记录“发生了哪次删除”不等于解释“删除后哪项职责去了哪里”。

近年的研究已经向功能层面推进。Bowen 与 Van Waes 将系统功能语言学 and 按键记录结合，分析修改如何发生在不同功能成分内部或之间，从而弥合过程与产品研究[12]。Mahlow、Ulasik 与 Tuggenner 提出“转换序列”“文本历史”和“句子历史”，将按键数据处理为可重建单句与全文生成轨迹的层级结构[13]。Conijn 等进一步开发了从日志中自动抽取包含删除、插入和替换内容的完整修订事件的方法[35]。Bekius 则把数字写作的按键轨迹编码为 TEI-XML，使文本发生学可以进入极细粒度的“纳米发生”层面[14]。这些工作已非常接近本文，却仍有关键差别：它们通常以

被修改的功能成分、发生过变换的句子或光标操作所在位置为起点；本文要测量的是未被修改单位的功能变化，以及这种变化同另一个单位的退出历史之间的关联。

可以把二者的差别概括为“事件中心”与“负载中心”。事件中心分析问：句子A经历了何种操作，意义如何改变？负载中心分析问：句子A退出后，原由A承担的篇章工作是否转移到字面未变的句子B？前一种研究为后一种研究提供了不可或缺的版本重建技术，但不能由修改事件本身推出职责迁移。若只统计删除次数、删改长度或最终保留率，一个作品可能表现为“删得很多”；只有观察其余句子的角色变化，才能判断删除是在排除无关材料，还是在重组全文的功能分工。

（三）文本发生学：保存“做过与撤销过”仍不等于说明作用机制

文本发生学把草稿、笔记、校样、增删与重写视为作品生成的一部分，而非通往定稿的废料。de Biasi 将草稿置于文本转化链条中，讨论不同发生文献在写作阶段中的功能类型[15]。Ferrer 认为，文本批评常试图从异文中建立文本，而文本发生学通过现实或潜在版本使“文本”重新不稳定[16]。Van Hulle 进一步把发生批评描述为一种带有时间维度的阅读策略：知道作品如何被制造，有助于理解它为何如此运作[17]。对数字原生写作而言，按键记录和版本历史又把以往不可见的微小操作带回研究视野[14,18]。

因此，“弃稿也属于作品历史”“删除并非无意义”“早期版本能解释定稿”都不能作为本文的新颖性依据。甚至“过去对后来具有持续作用”也已在发生学中被讨论。Ferrer 关于发生过程中的回溯作用与持续性的研究，已经提醒研究者：较晚的状态可能改写较早痕迹的意义，过去的选择也会在后续变化中保持效力[19]。

稿际余载与文本发生学的边界在于，前者不以重建某一作品的独特生成史为终点，而试图提出可跨作者、跨任务检验的机制变量。发生学可以证明某个删除段落曾经存在，并通过细读说明定稿仍与它构成解释关系；稿际余载研究还要继续问：这种关系能否被操作化？当两个写作者面对完全相同的当前文本，但其中一人曾写过又删去某段，另一人从未写过时，他们后续对未改句子的角色分配是否系统不同？若没有这一“同一现稿、不同

历史”的对照，任何解释都可能只是研究者从完整档案中作出的后见建构。

（四）分布式认知与写作物质性：为历史作用提供载体，却未指定职责流向

分布式认知反对把认知活动封闭在个体头脑内部。Hollan、Hutchins 与 Kirsh 强调，认知由内部资源和外部资源共同构成，信息在个人、社会组织和物质表征之间的传播与变换本身就是认知系统的结构[20]。Clark 与 Chalmers 以“主动外在论”说明，当环境部件在稳定耦合中发挥与内部过程相当的功能时，认知系统的边界可以超出皮肤[21]。写作研究据此可以把纸张、屏幕、提纲、便利贴、版本面板、搜索结果和合作者理解为写作认知的组成部分，而非被动容器。Piquette 与 Whitehouse 主编的《作为物质实践的书写》也从更长的历史尺度强调材料、表面、技术、方向与可见性如何参与意义生产[22]。

这些理论帮助回答稿际余载“凭什么持续”：它可能由工作记忆保持，也可能由撤销栈、版本历史、旁置笔记、屏幕残影、教师批注或口头讨论承载。可是，载体理论并不直接等于流向理论。知道旧段落保存在版本面板中，尚不能说明其原有的“为中心主张提供因果根据”这一职责会被哪个现存句子接管；知道认知分布于工具，亦不能预测某句将从证据转为主张。稿际余载因此把分布式认知作为输运条件，而把“篇章角色的重新分配”作为专门的解释对象。

（五）现有研究留下的窄缝

综合以上文献，可以把本文进入的窄缝表述为四个连续缺口。第一，过程理论承认递归反馈，却未把退出现稿的文本职责建模为可跨版本追踪的量。第二，修订研究精细记录被改单位，却较少把未改单位的功能变化设为主要因变量。第三，文本发生学保存并解释弃置痕迹，却通常不以随机对照识别“历史本身”的效应。第四，按键记录与句子历史技术能够重建过程，却仍缺少把版本历史、篇章功能和未改语句对齐的标注体系。

这个窄缝也规定了理论的最低门槛。稿际余载不能只换一种语言复述“写作是递归的”；不能把一切记忆、影响和互文都装入同一概念；不能因为某个最终文本与旧稿主题相似，就断言负载发生了迁移。它至少必须导出

一个现有事件统计不直接包含的可观察量，并提出一种可以被反驳的历史操纵。本文以“零字面角色迁移”承担这一任务。

三、跨学科结构映射：为何选择这三个碰撞学科

跨学科概念移植最常见的风险，是把相似词语误当作相同机制。稿件不是化石，删除不是材料磨损，句子也不是接受形态发生素的细胞。如果只凭“都有残余”“都有流动”“都有位置”建立类比，所得理论既无法检验，也容易用生动修辞遮蔽概念空缺。本文遵循 Gentner 的结构映射原则：跨域迁移优先对应关系系统而不是对象属性，只有在源域中相互制约的一组关系能在目标域中形成可检查的对应时，类比才具有推理价值 [23]。

据此，三个源学科分别提供不同而互补的机制部件。埋藏学提供“记录不是原貌，而是保存与破坏共同生成的选择性结果”，并提供“已死亡遗存反过来改变现存生态”的反馈关系；第三体摩擦学提供“从主体脱落的物质进入界面后，不只是废料，而会承载、剪切、循环、补给和排出”的中介流模型；位置信息理论提供“单位的命运由其在关系场中的位置决定，删去部分后其余单位能够重新读取位置并调节模式”的角色分化模型。三者的关系链分别回答：余载从何而来、如何在稿际运动、怎样改变未改句的命运。

（一）埋藏学：记录由保存与破坏共同制造

埋藏学研究生物遗存从生物圈进入岩石圈过程中经历的分解、搬运、埋藏、保存和破坏，以及这些过程如何改变研究者能够获得的信息。Behrensmeyer、Kidwell 与 Gastaldo 指出，埋藏学既用于评估化石样本的质量，也用于诊断生成沉积与化石记录的作用者、过程和环境，并重建有机物循环[24]。这一视角的关键不是“过去被残缺地保存”这一常识，而是：可见记录本身就是筛选机制的产物。缺失不在记录之外，缺失过程参与了记录的形状。

写作版本同样不是作者思想的透明切片。当前稿之所以呈现为现在的样子，既因为某些句子被写下和保留，也因为另一些句子被删除、移动、拆分、合并或从未完成。若把最终稿仅视为“剩下的文字”，删除就被当作信息归零；若采取埋藏学关系，当前稿应被视为生成、保存与破坏共同形

成的记录。研究者需要区分至少三层对象：一是曾经出现的文本材料，二是使其保存或退出的操作过程，三是这些过程给现稿的信息结构造成的偏向。由此，版本差异不只是通往真实意图的噪声，而是作文发生的构成性证据。

埋藏学还贡献一个更强的关系：埋藏反馈。Kidwell 与 Jablonski 讨论贝壳等死亡硬体在海底积累后如何改变底栖生物可利用的基底、空间和群落结构[25]。死去的部分没有恢复为生物，却以新的物质条件参与活系统。Behrensmeyer 后来把“埋藏反馈”概括为生物与死亡遗存之间影响现存生态的相互作用[26]。映射到写作，退出当前稿的段落不必以原句复活，仍可能通过留下一个论证空位、未兑现的定义、过早承诺的结论或已经形成的读者预期，改变余下语句的结构环境。这里的对应关系不是“删文等于死物”，而是“退出主要系统的成分通过改变界面条件影响仍在系统中的成分”。

同时，埋藏学提醒我们警惕时间混合。化石组合可能把不同时刻的遗存压缩在同一层位，形成时间平均[24,26]。一份看似同时存在的现稿，也可能把来自不同写作阶段、不同问题意识乃至不同读者反馈的句子压在同一页面。句子在屏幕上相邻，不代表它们在发生史上同时出生。稿际余载分析因此必须同时保留空间顺序和生成顺序，防止把后来的位置关系误认为从一开始就存在。

（二）第三体摩擦学：脱落物如何成为界面中的操作者

在两个物体接触和相对运动时，传统模型容易把注意力放在两个“第一体”的材料和载荷上。Godet 提出第三体路径，强调从接触表面脱落或外部引入的颗粒、转移膜等界面物质，是连接材料性质与实际摩擦磨损行为的“缺失环节”[27]。第三体既可能隔开第一体、承担部分载荷、容纳速度差和剪切，也可能聚集、破碎、再循环或被排出。它不是固定的第三件物体，而是一个由来源、状态和流量定义的界面层。

后续研究用“摩擦学回路”描述第三体的物质流。Renouf、Saulot 与 Berthier 区分源流 Q_s 、界面内部流 Q_i 、外部输入流 Q_e 、再循环流 Q_r 和最终离开接触区的磨损流 Q_w [28]。这一框架对写作最有启发的地方，不是把每个符号机械换成一种编辑动作，而是提供了连续的去向分类：从正文脱离的成分可能进入暂存区；暂存成分可以被重新带回；外部材料能够加入并

与旧材料混合；只有真正不再参与系统的部分才构成排出。由此，删除键不应被视为一个二值终点。它更像一扇把文字从可见正文推入另一种状态的门。

在本文模型中，“第三体”不直接等同于被删除的字符串。一个被删段落可能被作者彻底放弃，因而没有余载；相反，一句从未完整写出的计划也可能留下强烈的功能约束。更准确的映射对象是围绕退出成分形成的“稿际中介层”：其中包含作者记忆中的命题、版本系统中的旧文本、提纲上的结构位置、批注中的未决问题以及由删除造成的局部空位。稿际余载则是这个中介层在后续写作中能够承担和传递的篇章职责。

第三体模型还改变了对“摩擦”的理解。写作困难常被描述为作者缺少词汇、知识或动机；但许多停顿发生在材料过多、承诺冲突或同一段落承担过多任务之时。此时，阻力不是空无造成的，而可能由未被妥善排出的稿际材料造成。旧论点反复回流会使新结构难以稳定；删证据的职责无人接管，会迫使作者在多个位置补缀；外部输入的文献或人工智能建议若与原有问题不相容，也会增加界面剪切。第三体视角因此允许研究“写作阻力的物质史”，而不把所有困难还原为个体能力。

（三）位置信息：同样的文字为何会获得不同命运

Wolpert 的位置信息理论试图解释：具有相同遗传信息的细胞，如何依据在一个场中的相对位置形成不同结构。位置信息先于具体分化，细胞读取相对于参照点的位置，并以阈值等方式作出不同响应；当组织的一部分被移除或增加时，剩余部分可以重新获得位置信息并发生模式调节[29]。现代形态发生素研究进一步说明，信号的产生、运输和清除共同塑造时空分布，下游调控网络把连续信息转化为细胞命运，稳健性取决于梯度、反馈和尺度调节[30]。

写作中的句子当然没有生物感受器，但它的篇章功能同样不是字面形式的孤立属性。“研究发现学生修改得更少”可以是背景事实、核心证据、反常现象、方法局限或反方材料，取决于它与中心问题、前后语句和篇章目标的关系。一个语句的角色可以形式化为 $R(u \mid F_t)$ ，其中 u 是语句， F_t 是版本 t 的关系场，包括题目要求、中心主张、局部段落目标、相邻修辞关系、已建立的读者预期以及可调用的历史材料。即使 u 的字符序列不变，只要 F_t 变化， R 就可能改变。

位置信息理论为稿际余载补上“命运分化”的环节。删除一段文字不仅减少内容，还改变参照点、距离和关系强度；原来位于主张之后的句子可能失去“证据”位置，成为新的问题起点；原来夹在两个理由之间的句子被暴露到段首，可能承担框架功能。重要的是，角色迁移并非由物理顺序单独决定。句子即使仍在原页原行，只要中心主张被替换，它在论证网络中的相对位置也已经改变。因此，本文所说“位置”是关系位置，而不是页码坐标。

（四）三个源域的共同关系与二阶碰撞

三个学科表面相距甚远，却共享一个可抽象的关系结构：系统的现态不是活跃主体单独生成的；退出、脱落或被清除的成分会进入中介状态；中介状态通过流动改变局部关系场；场的改变使仍然存在的单位重新分化；最终结构则由保存、循环、吸收与排出共同稳定。埋藏学强调“缺失如何制造记录”，第三体摩擦学强调“脱落物如何在界面传递作用”，位置信息理论强调“关系场如何把同类单位分配到不同命运”。

二阶碰撞发生在三者相互校正之处。若只有埋藏学，理论容易停在“删除留下痕迹”；加入第三体，痕迹被改写为有来源和去向的中介流；再加入位置信息，流的后果不再只是旧内容回归，而是其他句子的角色分化。反过来，位置信息理论若没有埋藏学，容易把每个版本当作无历史的当前场；若没有第三体，也难解释已退出成分如何继续改变这个场。三个关系链闭合后，才得到本文的最小机制：退出产生余载，余载经稿际中介层输运，输运改变关系位置，位置变化导致未改语句的功能迁移。

这一闭合也限定了概念的适用范围。并非每一次删除都形成稿际余载，只有当退出成分曾承担或引发某项尚未解除的篇章职责时才可能形成；并非每次角色变化都由余载造成，当前稿中新输入的材料、读者反馈或作者目标改变也能直接重构关系场；并非所有余载都必须被吸收，作者可以显式撤销原承诺，使其从系统中排出。理论的任务不是把所有写作变化统一归因于一个概念，而是在多种原因中识别一条独立、可控制的历史路径。

四、核心概念：稿际余载与零字面角色迁移

（一）稿际余载的定义

稿际余载（interdraft residual load）是指：某一文本成分退出当前可见稿后，其尚未解决的论证功能、语义承诺和位置约束在版本之间持续存在，并对后续文本成分的选择、排列或篇章角色施加可观察影响。定义中的五个限定缺一不可。

第一，必须有可识别的“退出”。退出包括删除、剪切到暂存区、从正文移入注释、将明确段落降为提纲提示、用新表述覆盖旧表述等。纯粹未说出口的模糊意向虽然可能影响写作，但若无法与任何过程痕迹关联，不宜直接编码为稿际余载。第二，退出成分必须曾承担或激活某项篇章职责，例如提出中心主张、限定概念、提供根据、建立反方、承诺后文解释或设置段落位置。第三，这项职责在退出时没有被同时取消或完成。第四，职责通过记忆、工具或结构空位跨越至少一个版本状态。第五，后续影响需要落在可观察结果上，如未改句角色变化、补偿性增写、异常回返、结构移位或显式放弃。

“余载”包含三种相互关联但可分别标注的成分。论证功能负载指一个主张需要根据、一个反方需要回应、一个结论需要前提等功能义务。语义承诺负载指作者已经引入但尚未关闭的定义、因果关系、范围限定或价值立场。位置约束负载指某个单位的出现建立了参照点和顺序期待，使别的材料只能在若干关系位置中安放。例如，先写出“问题不在资源总量，而在分配节奏”会使后续数据倾向于被解释为节奏证据；即使这句后来被删，作者仍可能继续选择支持该区分的数据。三类负载经常共同出现，但理论上可以分离：一段证据被删可能主要留下功能空缺而不留下立场承诺；一条过早写下的结论则可能留下强语义承诺，却没有固定位置。

稿际余载不是心理能量，也不是可以直接称重的实体。它是一个关系性、反事实性的构念：如果退出成分从未出现，或其职责在退出时被完全撤销，后续文本的角色分配是否会不同？正因为它是反事实构念，单一作品的档案细读只能提供机制线索，不能独立完成因果识别。理论必须进入对照实验。

（二）载体、负载与痕迹的区分

概念清晰度要求把三个容易混淆的对象分开。载体是使历史可继续参与写作的媒介，包括工作记忆、旧文件、撤销记录、版本面板、提纲、旁注、截图、教师反馈和口头复述。负载是借这些媒介持续的篇章职责。痕迹是研究者能够观察到的证据，如一段旧文本、一次回看、一个删除动作或一个后续补偿句。

同一载体可以承载不同负载。例如版本历史保存了完整旧段落，但作者回看它时可能只取回“需要回应反方”这一职责，而不复用任何措辞。同一负载也可迁移到不同载体：最初由正文中的一句承载，删除后由作者记忆维持，稍后写进提纲，最终由新段落吸收。痕迹与负载之间则不存在一一对应。旧句仍在文件中，不代表它仍有作用；没有可见旧句，也不代表作用不存在。实证研究应通过多源数据建立推断链，而不能用“找到了旧稿”直接替代机制证据。

（三）稿际回路：来源、内部输运、外部输入、再循环与排出

借鉴第三体回路但避免物理同一化，本文把稿际余载的状态变化分为五类。来源流 *s* 来自正文成分的退出，它标记某项职责从可见文本中释放。内部输运 *I* 指该职责在作者记忆、现稿结构、提纲和版本系统之间转移。外部输入 *E* 指教师意见、同伴建议、资料文献或人工智能生成内容带来的新职责，它们可以与原有余载结合或冲突。再循环 *R* 指旧文字或旧职责重新进入现稿，例如原句恢复、改写后返回，或由另一句接管。排出 *w* 指作者显式撤销相关目标、改变论题范围或确认不再履行某项承诺，使该职责不再参与后续结构。

该回路不是守恒定律。写作中的职责会合并、分裂、放大和消失，不能假设来源负载等于吸收、循环与排出之和。但可把它作为分析账本：每当重要文本退出，研究者询问其职责是否被取消；若未取消，它暂由什么承载；下一次出现时以原句、改写句还是他句角色变化的形式出现；最终是被吸收还是被放弃。这个账本使“删除的东西似乎还在起作用”转化为一组可标注状态。

（四）零字面角色迁移

零字面角色迁移（zero-literal rhetorical-role migration）是稿际余载最关键的可观察后果。设 u_t 与 u_{t+1} 是相邻版本中经对齐确认的同一语句，若二者字符或规范化词元完全一致，即 $\text{Lex}(u_t) = \text{Lex}(u_{t+1})$ ，但盲法标注的篇章角色不同，即 $\text{Role}(u_t) \neq \text{Role}(u_{t+1})$ ，则记为一次零字面角色迁移。

“零字面”是一项严格测量约定，而不是关于意义完全不变的哲学判断。标点、空格、繁简体和引号样式可在预注册中规定是否规范化；只要发生实词替换、删加或语序变化，就不计入主要指标，可进入近零字面扩展指标。“角色”也必须预先限定。本文建议在议论文中采用八类：中心主张、分论点、根据或例证、论证保证或解释、反方主张、回应或反驳、情境框架、过渡或元话语。标注可以参考 Toulmin 论证模型、修辞结构理论和 PERSUADE 语料的论证话语单元，但不能简单照搬任何单一体系[31-33]。

零字面角色迁移率定义为：

$$\text{ZRM} = \frac{\text{发生篇章角色变化的字面不变对齐单元数}}{\text{全部字面不变对齐单元数}}。$$

ZRM 把分析焦点从“改了多少字”移向“多少未改文字被关系场重新解释”。一个版本对可能具有很低的编辑距离，却有较高 ZRM；也可能删改很多文字，却让其余句子保持原有职责，因而 ZRM 较低。前者更接近结构重组，后者可能只是材料更新。ZRM 不能取代修改量、质量评分或语义距离，但能增加一个既有指标通常遗漏的维度。

（五）作文发生的操作性定义

在理论层面，作文发生不是指敲下第一个字，也不是指文章达到提交长度，而是指材料之间出现了可持续的功能分工。操作上，可把一个局部写作区段判定为“结构发生事件”，需要同时满足三个条件：其一，出现源事件，即某项增删移位改变了关系场；其二，至少一个非目标单位在字面不变条件下发生角色迁移，或出现可与源职责对应的补偿性接管；其三，这种新分工在后续若干版本中保持，或成为新的修改参照点。

这一定义把“发生”理解为相变而非起点。若作者不断添加互不约束的材料，文字已发生，作文结构未必发生；若一次删除使剩余句子重新分工并

稳定为论证网络，即使总字数下降，作文反而更充分地发生。由此可解释成熟写作者为何常在删去关键段落时“突然找到文章”：决定性变化不是冗余减少本身，而是删除释放的职责迫使其他材料重新分化。

五、稿际余载如何运作：从退出到结构稳定

（一）第一环节：退出释放的不是内容残片，而是未决职责

稿际余载的源事件通常是删除，但删除本身没有固定意义。删除可能是排错、压缩、换词、撤销论点，也可能是把暂时无法安放的材料移出正文。同样长度的两次删除，后果可以完全不同。若作者删去一句重复例证，而原有证据链仍完整，这次操作大概率只减少冗余；若删去唯一解释主张与数据之间关系的论证保证，即使字数很少，也可能释放高强度功能负载。因此，余载强度首先取决于退出成分的关系中心性，而不是字符数量。

关系中心性可以从三个方面估计。其一是依赖数：有多少现存单位的解释依赖该成分。其二是不可替代性：同一职责是否已由别句承担。其三是承诺深度：该成分是否设定了全文的因果框架、概念边界或价值立场。一个孤立例子依赖少、替代多，余载可能很低；一句中心区分虽然短，却能决定多个段落被读成何种证据，其退出会产生较高余载。

源事件还包括移位和覆盖。移位并不让文字消失，但会让它原先维持的位置关系退出。例如，把段首的限定句移到结尾，原本由它承担的阅读框架在段首位置被释放；同一句在新位置可能承担总结。覆盖则常见于数字写作：作者选中旧句直接输入新句，旧字符立即不可见，却可能保留在撤销栈或记忆中。按键日志若只记录最终差分而不重建中间状态，容易把这种源事件压缩成一次替换，遗漏旧句职责短暂进入中介层的时段。

（二）第二环节：余载通过何种路径跨越版本

退出之后，职责要影响后续写作，必须存在输运路径。第一条是内部记忆路径。作者可能记得旧句的准确措辞，也可能只保留一个问题感，例如“这里还缺为什么”“前面那个反方没有回应”。后者更符合负载与载体的区分：文字形式消失，功能压力仍在。第二条是外部版本路径。撤销、修订记录、云端版本和旁置文件使旧材料保持可回访；它们延长余载寿命，也使原句再循环的成本降低。第三条是现稿空位路径。即使作者不再记得旧句，删除造成的衔接断裂、指代悬空或证据失配也能在回读时重

新触发同类职责。此时，余载并非由清晰的旧文本记忆传递，而由当前关系场的缺口再生。

第四条是社会路径。教师或同伴曾对旧段落提出的问题，可能在段落删除后继续作为评价标准存在。第五条是工具路径。搜索建议、自动补全和生成式人工智能会保存或重新提出先前被拒绝的框架，使外部输入与内部余载混合。这些路径可以并行：作者记得旧主张，又在版本面板中看见它，并受到教师批注提醒。实验研究必须操纵或测量路径可见性，否则难以判断观察到的效应来自工作记忆、外部检索还是当前缺口。

余载寿命也不是无限的。随着写作时间延长，未经回访的具体措辞会遗忘，新的目标可能取代旧承诺；但如果旧职责与高层目标相连，或被版本工具反复提示，它可能长期存在。由此可提出一个待检验预测：功能中心性比字面记忆更能预测余载寿命。作者可能无法复述被删主张，却仍持续围绕该主张组织证据；若只做旧句识别测试，会低估其结构影响。

（三）第三环节：关系场重排与位置重定

余载不会直接给某一句贴上新标签，而是先改变关系场。关系场至少包含四种参照。主题参照决定当前段落在回答什么问题；论证参照决定什么是主张、什么是支持、何处需要保证；读者参照决定哪些信息被视为已知或待解释；序列参照决定开端、展开、转折和收束的相对位置。当源成分退出后，一个或多个参照发生移动，仍在场的句子相对于参照的距离与方向随之改变。

位置重定有三种基本形式。第一是“中心暴露”：原本支持某主张的句子，在主张删除后成为局部最强命题，于是从证据或例证上升为分论点。第二是“空位接管”：某句与被删成分共享部分功能，在后续回读中被扩写或重新连接，从而接管其职责。第三是“极性反转”：加入反方或删除限定后，同一句从作者认可的主张变成需要被反驳的对象，或从结论变成让步。三者都可能在字面不变时发生。

关系场的变化可局部传播，也可跨段落传播。若删除只影响相邻指代，迁移可能限于两三句；若退出的是中心概念或全文主张，多个段落的证据资格都会改变。后者说明“全局修订”不一定表现为大面积改写。作者删掉一条中心句，其余材料几乎不动，但读者对所有段落的功能理解改变，这可能是编辑距离很小、结构距离很大的修订。

（四）第四环节：吸收、再循环、竞争与排出

关系场变化后，余载有四种主要去向。吸收是指现存或新写单位承担原职责。吸收可以是显性的，如在旧证据句前加“这说明”；也可以是隐性的，即句子不改，但因新位置和相邻关系被读作主张。再循环是旧文字以原样或近似形式返回正文。竞争发生在多个单位都可能接管同一职责，或同一单位同时承受多项职责时。竞争过强会产生段落“什么都像重点”的状态，常伴随反复移位和长停顿。排出则是作者改变问题、缩小范围或显式声明不讨论某项内容，从而取消职责。

成熟结构并不要求所有余载都被吸收。相反，有效写作的重要能力之一可能是识别哪些承诺应该退出系统。若作者舍不得放弃每个已形成的想法，余载持续再循环，会出现删而复写、换位不止和结论膨胀。所谓“杀死你的宝贝”之所以困难，不只是对漂亮句子的情感依恋，也可能因为漂亮句子承载的结构承诺尚未被明确解除。教学若只要求删除而不要求说明“它原先做什么、现在是否还需要”，就可能让学生在后续无意识地把同一负载重新写回。

（五）第五环节：稳定化与作文结构的涌现

当职责在文本单位之间形成相对稳定的分工，稿际回路进入稳定化。稳定并不意味着不再修改，而是后续修改以当前分工为参照：中心主张可以被识别，证据知道支持什么，反方知道针对什么，段落之间的顺序具有可解释理由。此时，未改句子的角色迁移率会在一段时间内下降，或迁移集中于局部而不再全局扩散。

稳定化具有路径依赖。两个字面完全相同的现稿，若来自不同历史，作者对哪些句子“本来应该做什么”的感受可能不同，因而下一步修改方向不同。这并不意味着历史决定一切；当前文本、任务要求和外部反馈都能覆盖旧路径。理论只预测，在其他条件相同且旧职责尚未排出时，历史会改变后续角色分配的概率。

由此，作文的发生可被理解为一种关系结构的涌现：起初材料的功能未定，增删让部分职责进入稿际中介层；中介层改变关系场，促使句子分化；分化后的句子又成为新的参照，约束下一轮增删。成品不是这条链的终点图像，而是某个时刻相对稳定的截面。

六、三个机制示例：看见“未改之句”的变化

以下案例均为理论构造，用于说明可标注现象，不是自然语料，也不构成实证证据。每个案例都刻意保持目标句字面不变，以排除“改词导致功能改变”这一更直接解释。

（一）删除中心句：证据变成问题触发器

设某稿的段落为：①“一座城市的韧性首先取决于基础设施能否持续运转。”②“2021年暴雨后，三个社区的非正式互助网络比专业救援更早恢复了基本供给。”③“因此，城市韧性也取决于关系网络能否迅速形成。”在这一版本中，①是初始主张，②是修正主张的反常证据，③是由证据推出的新主张。

作者随后删除①，②和③一字未改。新段落从②开始。②不再是针对“基础设施决定论”的反常证据，因为被反驳的主张已经退出现稿；它转而成为一个需要解释的开场事实，即“为何非正式网络更早恢复供给”的问题触发器。③则由补充性结论上升为段落中心主张。若作者后续增加关系网络的三个机制，②的角色变化会稳定下来。

这里的源负载包括“为基础设施观点提供限定”和“建立从设施到关系的转折”。删除①并未自动取消这两个职责。其一可能被排出，其二则由②的异常性和③的新中心位置共同吸收。版本差分只显示①被删；零字面角色标注则显示②从反常证据变为问题触发器，③从补充结论变为中心主张。作文结构的关键变化发生在未改句上。

（二）移走限定句：事实陈述变成反方主张

设段落原为：①“在资源极度不足的地区，线上教育仍可能是唯一可扩展的选择。”②“线上课程能够以更低边际成本覆盖更多学生。”③“但覆盖人数不等于形成持续学习。”此时①限定论域，②在限定范围内作为支持线上教育的事实，③提出评价标准的转折。

作者把①移到全文结尾作为政策保留意见，②和③字面不变。由于范围限定退出现段，②在当前局部不再被作者框定为特殊情境下的理由，而更像一个普遍的技术乐观主张；③于是从一般转折变为对②的直接反驳。②发生“根据—反方主张”迁移，③发生“限制—反驳”迁移。移位的字符只在①，篇章极性变化却落在②和③。

这一案例说明，稿际余载不仅由删除产生。①的限定职责从段首位置释放后，并未在当前段落被接管，因而②承受了超出原范围的普遍性，成为更容易被攻击的命题。若作者后来在②中加入“在资源不足地区”，角色迁移可能通过字面修改得到修复；若不修改，读者仍会以新关系解释它。

（三）删除人工智能建议：被拒绝的框架继续组织材料

设写作者询问生成式人工智能“短视频为何改变公共讨论”，系统建议用“注意力碎片化”作为中心解释。作者把生成段落贴入稿中，围绕它搜集了推荐算法、情绪传播和内容长度三组材料，随后认为该框架过于心理化，将整段删除。当前稿中只保留一句原有文字：“平台并不直接规定人们相信什么，却持续规定什么更容易被看见。”

如果作者此后仍把三组材料组织为“争夺有限注意力”，这句原本的媒介条件陈述可能在字面不变时成为中心论点。被删的人工智能措辞没有返回，但其“注意力配置”职责被原句吸收。另一位从未接触该建议的写作者，面对同一现稿，可能把这句当作算法权力的背景，并沿制度治理方向展开。两份现稿在某个时点完全相同，历史不同导致下一步角色分配可能不同。

这个案例同时提示伦理边界。不能仅凭最终结构与被删建议相似，就断定人工智能“实际写了”文章；相似可能来自共同知识，也可能由当前材料自然导出。要识别稿际余载，需要过程记录、接触控制和替代解释检验。理论关注的是被拒绝材料如何继续构成写作环境，而不是借历史数据扩大机器的作者权利。

七、与相邻概念的严格区分

（一）不等于修订、深层修订或全局修订

修订是更上位的活动范畴，稿际余载是其中一条可能机制。一次修订可以不产生余载，如修正错别字；一次余载效应也可以在没有立即可见编辑时显现，如作者回读后改变对未改句的理解。深层或全局修订按修改涉及的意义层级和篇章范围分类，零字面角色迁移则按“修改事件与功能后果是否落在同一单位”分类。二者可以重叠，但不是同义词。

Bowen 与 Van Waes 的功能性修订研究尤其需要区分[12]。该研究考察修改如何作用于主题—述位等功能成分，已经跨越纯形式统计。本文在其基础

上增加“非共址效应”：功能变化可以发生在未被修改的单位，而触发操作位于别处。若未来数据表明所有所谓零字面角色变化都能被功能性修订体系直接编码且无需历史变量，稿际余载的独立价值将显著下降。

（二）不等于文本发生学中的前文本或异文

前文本是由草稿、笔记等材料构成的发生档案对象；稿际余载是从这些对象与后续行为之间推断的机制构念。某个旧段落属于前文本，不代表它形成余载；某项余载也可能只有按键和停顿痕迹，未留下可读旧段。文本发生学强调作品个案的时间厚度，稿际余载研究则进一步寻求可比较指标和因果操纵。最理想的关系是互补：发生学细读帮助发现候选机制，实验研究检验可推广性，再由发生学解释具体作品中负载的语义内容。

（三）不等于互文性、影响研究或记忆启动

互文性讨论文本与既有话语的关系，影响研究追踪作者如何接收他者作品；稿际余载限于同一写作任务的过程历史，或至少要求来源事件可定位到当前作品的形成链。普通记忆启动是更强的替代解释：写过又删掉某段的人后续采用相关概念，可能只因近期接触提高了可及性。为此，实验必须设置“只阅读同段但未把它置入稿件”的接触控制，并测量词汇复用。如果“写入—删除”组与“只阅读”组无差异，且效应完全由相关词出现频率解释，就没有必要引入稿际余载。

理论预测的增量在于，负载效应应当超出词汇和主题启动，体现为职责定向。例如两组都记得“注意力”一词，但只有写入又删除组更倾向于让某句接管中心解释，或在特定结构空位进行补偿。只有这种角色层面的差异，才支持“稿件历史”而非一般接触史。

（四）不等于分布式认知或扩展心智

分布式认知说明写作过程可以跨越人、工具和表征，稿际余载说明其中一种信息——未决篇章职责——如何迁移并改变文本单位角色。前者提供系统边界，后者提供特定状态变量。版本面板是认知载体，不是余载本身；只有当它保存的历史被调用或通过当前缺口影响后续分工时，才进入机制链。

（五）不等于路径依赖或滞后效应的泛称

路径依赖表示系统结果受先前状态影响，稿际余载当然具有路径依赖性。但泛称不能告诉研究者追踪哪一类历史、何种中介和哪个结果。稿际余载限定源为文本文职的退出，限定中介为跨稿承载与关系场变化，限定结果为文本角色重分配。若不保留这些限定，任何“过去影响现在”都会被纳入，概念将失去反证可能。

八、可证伪的实证研究程序

（一）研究问题与假设

实证研究首先要回答四个问题。第一，在自然写作中，零字面角色迁移是否真实存在，频率是否高于标注误差和版本对齐误差？第二，重要文本退出是否提高后续未改单位发生角色迁移的概率？第三，在当前稿完全相同的条件下，不同写作历史是否导致不同的后续结构？第四，历史效应能否被普通接触、词汇启动、版本检索或当前文本缺口完全解释？

据此提出五个预注册假设。H1：与低关系中心性的删除相比，高中心性删除后的局部与全局 ZRM 更高。H2：在同一现稿条件下，“写入—删除”组的 ZRM 高于“从未接触”组。H3：在控制内容接触后，“写入—删除”组与“只阅读”组仍有差异，且差异主要体现在角色接管而非词汇复用。H4：可见版本历史会提高旧职责的再循环和定向吸收，因此“写入—删除且历史可见”组的效应强于历史不可见组。H5：若源职责在删除时被明确撤销，历史效应显著减弱；这区分“删除”与“排出”。

这些假设并非都必须为真。尤其 H4 可能出现反向结果：历史可见使作者直接恢复旧句，反而减少由未改句完成的零字面迁移。出现这种结果不会让整体理论立即失败，却会迫使研究者区分“原文再循环”和“他句吸收”两条路径。

（二）自然语料阶段：先证明现象，不急于证明原因

第一阶段可使用具有细粒度版本或按键记录的论说文语料。KLiCkE 语料包含约 5,000 篇美国成人写作的议论文及按键、鼠标、时间戳和光标信息，是估计现象基线的重要资源[34]。Conijn 等的修订数据集具有人工标注的修订属性，适合定位源事件[11]。然而，前者缺少跨版本篇章角色标注，后

者明确缺少“加工过程”和“触发因素”；PERSUADE 虽有大规模论证话语单元标注，却主要提供成品而非完整按键历史[33]。因此，现有公开语料难以单独完成检验，更适合开发对齐与候选筛选算法。

自然语料分析可以分四步进行。首先重建每隔固定事件数或固定时间的文本状态，既保存光标时间序列，也保存可读版本。其次以句子或基本话语单元为单位做跨版本对齐，严格识别字面不变单元。再次由不知版本历史的标注者分别判断各版本中单位角色，避免完整历史诱发后见解释。最后把角色变化与附近源事件连接，估计删除中心性、距离、时间间隔和最终去向。

第一阶段的成功标准不是得到漂亮案例，而是建立可靠测量。应报告对齐准确率、角色标注一致性、在随机打乱版本顺序时的伪迁移率，以及不同时间采样粒度对 ZRM 的影响。若盲标一致性无法达到可接受水平，或字面不变句的角色在独立标注者看来大多无法区分，理论的主要指标就需要重构。

（三）核心实验：“同一现稿、不同历史”

因果检验建议采用四组随机设计，计划样本不少于 240 名熟练程度分层的大学写作者，每组至少 60 人；正式样本量应在预实验取得效应与标注方差后，通过模拟功效分析确定。所有参与者围绕同一议题继续一份经过控制的短篇初稿。在关键时点，四组最终看到完全相同的当前稿，但此前经历不同。

第一组为“无接触组”：从未见过关键段落 P。第二组为“只阅读组”：阅读 P 并完成理解判断，但 P 从未进入其编辑窗口。第三组为“写入—删除组”：按要求把 P 输入稿件、围绕它完成一次局部连接，随后删除，且无法查看历史。第四组为“写入—删除—历史可见组”：经历与第三组相同，但可使用版本面板。为检验明确排出，可在第三、四组内再随机分配删除指令：一半只说“请删除此段”，另一半要求写下“本文不再主张 P，也不需要履行 P 引出的解释任务”。

关键在于当前状态匹配。实验平台在进入自由续写前应自动规范化文档，确保四组的字符、格式、光标位置和可见提示一致；历史可见性是唯一允许保留的界面差异。只阅读组控制语义和词汇接触，写入—删除组增加“曾作为本文组成部分并建立局部关系”的历史。续写阶段记录所有按

键、回看、停顿、版本调用和最终文本，并在结束后进行短时回溯访谈，但访谈问题不能先暴露“余载”假设。

实验应使用多个议题和多个 P 类型：中心主张、论证保证、关键反方、范围限定和例证，以避免效应只属于某一种文本。P 的字数、可读性和主题词频应匹配。研究者还可以设置“伪中心段”，其语言醒目但与当前论证关系弱，用来区分表面显著性与功能中心性。

（四）角色标注、主要指标与统计模型

标注单位建议采用可对齐的独立分句或基本话语单元。每个单位先标“是否字面不变”，再标八类篇章角色及其上位关系。标注者只看待判版本及必要上下文，不看参与者组别、删除事件和其他版本。至少两名标注者独立判断，第三名裁决；主要类别的 Cohen's κ 或 Krippendorff's α 目标值应在预注册中设定，例如不低于 0.75。无法判定的单位保留为缺失，不应强迫归类。

主要结果为个体 ZRM 和源事件后固定窗口内的局部 ZRM。次要结果包括：第一次全局回返的时间、旧职责被哪类单位吸收、原句再循环率、词汇复用率、结构质量盲评分、删除后停顿长度、版本面板调用次数以及回溯访谈中自发提及旧职责的比例。结构质量评分不能替代 ZRM，因为角色迁移可能改善也可能损害作文。

主要推断可使用混合效应逻辑回归，以每个字面不变单位是否迁移为因变量，历史组别、源类型、中心性、时间距离及排出指令为固定效应，参与者和题项为随机截距，必要时加入随机斜率。ZRM 作为比例可用二项模型而非直接做普通线性回归。角色迁移方向可使用多项模型，时间到第一次接管可用生存分析。写作熟练度、输入速度、工作记忆、内容熟悉度和主题立场作为预先规定的协变量，而不是见到结果后任意加入。

（五）替代解释与判别性预测

最重要的替代解释是普通启动。如果只阅读组与写入—删除组表现相同，历史效应可由内容接触说明。第二个替代解释是版本检索：历史可见组可能只是查回旧文。若效应仅在打开版本面板之后出现，并表现为原句复制而非未改句角色变化，结论应限定为外部检索，而不是更强的余载机制。第三个替代解释是当前缺口：四组虽然字符相同，但关键段落缺失本身就

会使所有人重新组织。无接触组提供了这一基线；只有组间增量才属于历史。

第四个替代解释是任务需求猜测。参与者可能认为研究者希望其使用刚刚删除内容。可采用掩蔽目的、多个填充任务和事后猜测问卷。第五个是标注后见偏差。盲标和打乱历史可以控制。第六个是一般投入差异：输入—删除组在任务上花了更多时间。可让无接触组和只阅读组完成时长、键击量相当的中性任务，并把总投入作为控制变量。

稿际余载的判别性预测不是“接触P的人更常写P的词”，而是“曾把P作为稿件组成部分的人，更可能在特定关系空位上让其他单位接管P的职责，即使不复用P的词”。如果这个增量不存在，理论应退回到启动或外部记忆的较简单解释。

（六）明确的反证条件

理论需要在失败时允许自己失败。以下结果将构成强反证。第一，经过可靠盲标后，零字面角色迁移频率不高于版本采样和标注噪声。第二，在当前稿匹配、内容接触与投入受控后，写入—删除历史不产生任何角色层面的增量。第三，所有差异均由旧词复用、直接版本查看或当前稿文本特征解释，加入历史变量不能改善预测。第四，源成分的功能中心性与后续迁移无关，而删除字符数等表面变量完全解释结果。第五，明确撤销职责与普通删除没有差别，说明所谓“负载”未表现出可区分的状态。

若只有部分假设失败，理论也应相应缩减。例如历史效应只在版本可见时存在，则稿际余载可能不是一般心理持续，而是版本工具支持的外部机制；效应只在中心主张删除后存在，则适用范围应限于高层结构；ZRM存在但不受历史影响，则它仍是有用的修订指标，却不能作为稿际余载证据。

九、理论贡献：重新回答“作文是如何发生的”

（一）从生成单位转向职责单位

传统过程研究常以计划、命题、句子、停顿或修改事件为分析单位。本文提出另一个单位：篇章职责。职责可以被一句承担，也可以分布在多句；可以从显性正文退出，又在另一个单位上重现。以职责为单位，研究者不

再只问“这句话后来怎样”，还问“这句话所做的工作后来由谁完成”。这使跨版本研究能够连接认知目标、文本关系和物质痕迹。

职责单位具有中尺度优势。它比“作者意图”更可观察，因为可以通过文本角色和操作历史推断；又比字符差分更接近作文结构，因为它关注主张、根据、限定与回应之间的关系。它还允许描述一对多和多对一迁移：被删中心句的负载可能分散到三个分论点，一组冗余例子也可能合并为一个论证保证。

（二）从“修订文本”转向“重分配关系”

修订常被理解为对已有文本作改变。稿际余载理论把改变进一步分成直接改变和关系改变。直接改变落在被操作字符上；关系改变落在整个场中，包括未改单位。由此，“改动范围”不再等同于选中范围。一个单句删除可以产生全局关系改变，一次大段替换也可能只更新材料而不重组功能。

这一区分有助于重新解释经验研究中修改量与质量提升关系不稳定的现象。若只按次数和字数测量，真正有效的关系重组可能被低估，反复表层改写却被高估。ZRM与职责去向可以作为补充指标，检验某次修订是否让篇章分工更明确，而不预设“改得多就是写得好”。

（三）把“删除”从负操作改写为结构操作

删除在编辑器中表现为字符消失，容易被归为负操作。本文并不浪漫化删除：许多删除只是损失，过早删除也可能破坏思考。但在机制上，删除可同时完成三件事：移走材料、释放职责、改变参照场。第三件事使其成为结构操作。删除的价值不在“少”，而在它是否促使更合理的职责吸收或明确排出。

这也解释为何保留完整版本历史具有研究价值，却未必总有写作价值。版本历史能让旧材料再循环，降低丢失恐惧；但过度可见也可能让已决定排出的职责不断返回。好的工具不应只提供“恢复删除”，还应帮助作者区分“文字已删除但职责未处理”与“文字和职责都已放弃”。

（四）作文发生的时间点不再是第一句，而是第一次稳定分化

若把作文视为字符串生成，它从第一键开始；若把作文视为职责网络，它在材料首次形成稳定互相约束时发生。这个时间点可能很早，例如一句问题与一句回答立即建立结构；也可能很晚，作者写了大量材料后，通过一

次关键删除才让全文分化。作文发生因此不是统一阶段，而是一系列局部结构事件，最终可能汇聚为全篇模式。

这一回答保留了过程理论的递归性，又增加了可观察的文本机制：写作目标通过材料进入稿件，材料的退出释放职责，职责迁移改变其他语句，新的角色分工反过来塑造目标。作文既发生在作者头脑中，也发生在版本之间的关系重排中。

十、教学、评价与人工智能写作工具的启示

（一）教学：让学生看见删除后的“职责债务”

许多写作教学要求学生保留草稿、展示修改痕迹或说明删改理由，但常把反思停留在“我删掉了重复内容”“我增加了例子”。稿际余载框架可把问题推进一步：被删内容原先承担什么职责？这项职责是否仍需要？若需要，现在由哪句话接管？若不需要，中心问题发生了什么变化？这种“职责审计”把修订从操作说明转为结构说明。

课堂上可以使用“影子档案”练习。学生每删除一个高中心性段落，只记录一句非措辞化说明，如“删除了唯一的反方”“取消了因果解释”“范围限定尚未安放”，而不必保留整段供反复粘贴。下一轮回读时，学生检查这些职责是被吸收、再循环还是排出。这样既利用历史，又降低漂亮旧句诱发回流的风险。

另一个练习是“未改句复审”。教师不要求学生寻找改得最多的地方，而让其在版本对照中抽取三句字面未变的文字，分别说明它们在旧稿和新稿中的工作。若学生发现一句从例子变成论点，就需要判断这种上升是否有足够根据；若一句从限定变成过渡，则检查原限定职责是否丢失。该练习直接训练关系阅读。

（二）评价：过程证据不应简化为勤奋分数

数字平台使教师容易获得总键击、总停顿、删除比例和版本数，但这些指标不应被直接转化为勤奋或诚信分数。写得熟练的人可能用较少操作完成深层重组，焦虑者可能留下大量无效回返。稿际余载提示评价者关注职责重分配的质量，而不是过程痕迹的数量。

更重要的是，过程数据具有隐私与权力风险。按键和版本历史可能暴露未成熟想法、情绪性文字、个人信息与被明确放弃的立场。研究和教学应遵循最小收集、知情同意、目的限定和可删除原则。被删文字曾出现，不意味着它应永久成为评价作者的证据；学术研究需要保存发生痕迹，伦理制度同时必须承认撤回权。

（三）版本感知型写作工具：从“恢复文字”升级到“提示职责”

现有版本工具擅长显示谁改了哪些字符，却很少提示结构后果。版本感知型助手可以在检测到高中心性删除后提出低干预问题：“这段原先是唯一的根据，现在哪一部分承担根据？”“删除该限定后，后文的结论范围是否扩大？”这种提示不替作者决定答案，只把潜在余载显性化。

技术上，工具需要结合版本对齐、论证单元识别和修辞关系分析。PERSUADE等语料证明大规模论证角色标注具有可行性[33]，按键记录和句子历史方法提供过程重建基础[10-14]。但自动系统的角色判断仍有不确定性，应以“候选提示”而非事实判决呈现，并允许作者关闭历史分析、删除记录或声明某项职责已经排出。

对于生成式人工智能，尤其要避免“拒绝无效”的设计。若系统不断把作者删除的建议重新生成，旧负载会被强制再循环，削弱作者改变问题的能力。更好的机制是让作者标记“保留问题、放弃表述”“放弃整个方向”或“以后再议”。前者允许职责迁移，后者应降低系统再次推荐同一框架的概率。

（四）人工智能时代的作者性：过程历史不是所有权算法

稿际余载可能揭示被拒绝的机器建议仍影响终稿，但这种影响不能直接换算为作者比例或抄袭判断。人类写作本就分布于阅读、交谈、笔记和工具，结构相似也不必然证明具体来源。研究可以分析外部输入如何进入稿际回路，却不应把每个历史接触自动转化为归属主张。

相反，理论更强调人的结构决策：作者可以吸收、改写、竞争、排出外部负载，并通过角色重分配形成不同作品。真正值得评价的不是文字是否曾由机器建议，而是作者是否能识别建议建立了哪些承诺，是否能在拒绝表

述时也决定保留或撤销其职责。这是一种比“用了多少机器文本”更细的过程素养。

十一、理论局限与研究边界

第一，跨学科映射始终存在过度实体化风险。“流”“负载”“场”和“稳定化”在本文中是关系模型，不表示写作服从物质守恒、扩散方程或生物阈值。除非未来数据支持量化形式，不应把分析账本写成自然定律。结构映射的价值在于导出问题和预测，不在于用源学科权威替目标学科作证。

第二，篇章角色具有解释性。同一句可以同时承担背景与证据，角色也会因读者目的不同而变化。八类体系是研究折中，不是文本本体。未来应比较单标签、多标签和连续中心性模型，并报告不同标注体系下 ZRM 是否稳健。若指标高度依赖某一套标签，理论的普适性需要降低。

第三，字面不变是严格但有限的窗口。许多真实职责迁移伴随小幅改词，主要 ZRM 会漏掉它们。研究可在验证零字面指标后建立“近零迁移”，用语义与编辑距离分层，但不应一开始就放宽到任何相似句，否则对齐和因果归属会失控。

第四，作者记忆与当前结构难以完全分离。即使平台把四组现稿规范为相同，写入一删除组仍拥有动作记忆、投入感和任务解释差异。多组接触控制、匹配活动和回溯访谈只能缩小替代解释，不能把生态写作压缩为纯实验变量。最可信的证据将来自实验、自然按键语料和发生学个案的三角互证。

第五，理论可能具有体裁边界。议论文的主张、根据和反驳相对可标，叙事、诗歌和碎片化数字写作的角色更开放。叙事中可改用叙事功能、视角、时间与动机角色，但不能未经检验直接沿用 ZRM 类别。不同语言的省略、主题推进和句界也会影响对齐。

第六，并非所有好作文都需要高 ZRM。高迁移可能表示深层重组，也可能表示结构不稳定；低迁移可能表示浅改，也可能表示作者早期计划清晰。ZRM 是过程描述量，不是质量分数。只有结合最终连贯性、论证质量和任务适切性，才能判断迁移是否有益。

第七，本文尚未越过实证门槛。理论从相邻文献的未解问题中导出可区分命题，也给出反证条件，但没有报告真实参与者数据。因而“稿际余载存在”“历史导致 ZRM 增加”都应保留为研究假设。本文的完成标志是问题被做成可失败的研究程序，而不是概念已经因其解释力获得事实地位。

十二、结论

回答“作文是如何发生的”，不能再满足于“作者先构思后写作”，也不能只重复“写作是递归的”。过程研究已经证明计划、表达和回读相互嵌套；修订研究已经分类增删替移；文本发生学已经保存并解释被尝试和被放弃的文字；按键记录已经把数字写作的细小动作重新变成数据。新的问题必须出现在这些成就之后：一次修改为何会改变未被修改的句子？

本文以埋藏学、第三体摩擦学和位置信息理论的关系结构为支架，提出稿际余载机制。退出的文本成分不仅从屏幕消失，还可能释放未决的论证功能、语义承诺和位置约束；这些职责借记忆、版本、提纲、结构空位和社会反馈跨稿输运，改变当前关系场；仍在稿中的语句由此重新定位和分化，产生字面不变而角色改变的零字面迁移；职责最终被吸收、再循环、竞争或排出，篇章结构在相对稳定的分工中形成。

这个模型给出的简洁答案是：作文不是文字不断增加的结果，而是篇章角色不断被重新分配的结果。更严格的答案是：作文发生于一次增删不再只改变被操作的文字，而开始使其他未改文字承担新的篇章职责，并让这种职责分配成为后续写作约束之时。

理论是否成立，取决于未来数据能否越过更简单的解释。研究者需要在同一现稿下操纵不同历史，控制内容接触与版本检索，盲法标注未改句角色，检验 ZRM 是否得到历史变量的增量解释。如果不能，稿际余载应被缩减或放弃；如果能够，作文研究将获得一个新的中尺度对象：不是脑内意图，不是屏上字符，而是在版本之间迁移、在句子之间重新落位的篇章职责。

参考文献

- [1] EMIG J. *The Composing Processes of Twelfth Graders*[M]. Urbana, IL: National Council of Teachers of English, 1971. ERIC ED058205.
- [2] PERL S. The composing processes of unskilled college writers[J]. *Research in the Teaching of English*, 1979, 13(4): 317-336.
- [3] PERL S. Understanding composing[J]. *College Composition and Communication*, 1980, 31(4): 363-369.
- [4] FLOWER L, HAYES J R. A cognitive process theory of writing[J]. *College Composition and Communication*, 1981, 32(4): 365-387.
- [5] SOMMERS N. Revision strategies of student writers and experienced adult writers[J]. *College Composition and Communication*, 1980, 31(4): 378-388. DOI: 10.2307/356588.
- [6] SCARDAMALIA M, BEREITER C. Knowledge telling and knowledge transforming in written composition[A]/ROSENBERG S, ed. *Advances in Applied Psycholinguistics, Vol. 2: Reading, Writing, and Language Learning*[C]. Cambridge: Cambridge University Press, 1987: 142-175.
- [7] HAYES J R. Modeling and remodeling writing[J]. *Written Communication*, 2012, 29(3): 369-388. DOI: 10.1177/0741088312451260.
- [8] FAIGLEY L, WITTE S. Analyzing revision[J]. *College Composition and Communication*, 1981, 32(4): 400-414. DOI: 10.2307/356602.
- [9] FITZGERALD J. Research on revision in writing[J]. *Review of Educational Research*, 1987, 57(4): 481-506. DOI: 10.3102/00346543057004481.
- [10] LEIJTEN M, VAN WAES L. Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes[J]. *Written Communication*, 2013, 30(3): 358-392. DOI: 10.1177/0741088313491692.
- [11] CONIJN R, DUX SPELTZ E, VAN ZAAANEN M, et al. A process-oriented dataset of revisions during writing[A]/*Proceedings of the Twelfth Language Resources and Evaluation Conference*[C]. Marseille: ELRA, 2020: 363-368. <https://aclanthology.org/2020.lrec-1.45/>.
- [12] BOWEN N, VAN WAES L. Exploring revisions in academic text: Closing the gap between process and product approaches in digital writing[J]. *Written Communication*, 2020, 37(3): 322-364. DOI: 10.1177/0741088320916508.
- [13] MAHLOW C, ULASIK M A, TUGGENER D. Extraction of transforming sequences and sentence histories from writing process data: A first step towards linguistic modeling of writing[J]. *Reading and Writing*, 2024, 37: 443-482. DOI: 10.1007/s11145-021-10234-6. (网络首发于 2022 年)
- [14] BEKIUS L. The reconstruction of the author's movement through the text, or how to encode keystroke logged writing processes in TEI-XML[J]. *Variants*, 2021(15-16): 3-43. DOI: 10.4000/variants.1245.
- [15] DE BIASI P-M, WASSENAAR I. What is a literary draft? Toward a functional typology of genetic documentation[J]. *Yale French Studies*, 1996(89): 26-58. DOI: 10.2307/2930337.
- [16] FERRER D. Genetic criticism with textual criticism: From variant to variation[J]. *Variants*, 2016(12-13): 57-64. DOI: 10.4000/variants.284.
- [17] VAN HULLE D. *Genetic Criticism: Tracing Creativity in Literature*[M]. Oxford: Oxford University Press, 2022. DOI: 10.1093/oso/9780192846792.001.0001.

- [18] VAN HULLE D. Dynamic facsimiles: Note on the transcription of born-digital works[J]. *Variants*, 2021(15-16). DOI: 10.4000/variants.1450.
- [19] FERRER D, CORCORAN M G. Clementis's cap: Retroaction and persistence in the genetic process[J]. *Yale French Studies*, 1996(89): 223-236. DOI: 10.2307/2930349.
- [20] HOLLAN J, HUTCHINS E, KIRSH D. Distributed cognition: Toward a new foundation for human-computer interaction research[J]. *ACM Transactions on Computer-Human Interaction*, 2000, 7(2): 174-196. DOI: 10.1145/353485.353487.
- [21] CLARK A, CHALMERS D. The extended mind[J]. *Analysis*, 1998, 58(1): 7-19. DOI: 10.1093/analys/58.1.7.
- [22] PIQUETTE K E, WHITEHOUSE R D, eds. *Writing as Material Practice: Substance, Surface and Medium*[M]. London: Ubiquity Press, 2013. DOI: 10.5334/bai.
- [23] GENTNER D. Structure-mapping: A theoretical framework for analogy[J]. *Cognitive Science*, 1983, 7(2): 155-170. DOI: 10.1207/s15516709cog0702_3.
- [24] BEHRENSMEYER A K, KIDWELL S M, GASTALDO R A. Taphonomy and paleobiology[J]. *Paleobiology*, 2000, 26(S4): 103-147. DOI: 10.1666/0094-8373(2000)26[103:TAP]2.0.CO;2.
- [25] KIDWELL S M, JABLONSKI D. Taphonomic feedback: Ecological consequences of shell accumulation[A]//TEVESZ M J S, MCCALL P L, eds. *Biotic Interactions in Recent and Fossil Benthic Communities*[C]. New York: Plenum Press, 1983: 195-248.
- [26] BEHRENSMEYER A K. Taphonomy[A]//ALDERTON D, ELIAS S A, eds. *Encyclopedia of Geology*[M]. 2nd ed. Oxford: Academic Press, 2021: 12-22.
<https://repository.si.edu/handle/10088/120162>.
- [27] GODET M. The third-body approach: A mechanical view of wear[J]. *Wear*, 1984, 100(1-3): 437-452. DOI: 10.1016/0043-1648(84)90025-5.
- [28] RENOUF M, SAULOT A, BERTHIER Y. Third body flow during wheel-rail interaction[A]//MOTA SOARES C A, et al., eds. *III European Conference on Computational Mechanics*[C]. Dordrecht: Springer, 2006. DOI: 10.1007/1-4020-5370-3_336.
- [29] WOLPERT L. Positional information and the spatial pattern of cellular differentiation[J]. *Journal of Theoretical Biology*, 1969, 25(1): 1-47. DOI: 10.1016/S0022-5193(69)80016-0.
- [30] KICHEVA A, BRISCOE J. Control of tissue development by morphogens[J]. *Annual Review of Cell and Developmental Biology*, 2023, 39: 91-121. DOI: 10.1146/annurev-cellbio-020823-011522.
- [31] TOULMIN S E. *The Uses of Argument*[M]. Cambridge: Cambridge University Press, 1958.
- [32] MANN W C, THOMPSON S A. Rhetorical Structure Theory: Toward a functional theory of text organization[J]. *Text*, 1988, 8(3): 243-281. DOI: 10.1515/text.1.1988.8.3.243.
- [33] CROSSLEY S A, BAFFOUR P, TIAN Y, et al. The persuasive essays for rating, selecting, and understanding argumentative and discourse elements (PERSUADE) corpus 1.0[J]. *Assessing Writing*, 2022, 54: 100667. DOI: 10.1016/j.asw.2022.100667.
- [34] TIAN Y, CROSSLEY S, VAN WAES L. The KLiCke corpus: Keystroke logging in compositions for knowledge evaluation[J]. *Journal of Writing Research*, 2025, 17(1): 23-60. DOI: 10.17239/jowr-2025.17.01.02.
- [35] CONIJN R, DUX SPELTZ E, CHUKHAREV-HUDILAINEN E. Automated extraction of revision events from keystroke data[J]. *Reading and Writing*, 2024, 37(2): 483-508. DOI: 10.1007/s11145-021-10222-w.

编者补节·第五章

一·需要补的两位

本章的观察对象是：语句在相邻版本中字面完全一致，而它的篇章角色改变了。这是本卷最反直觉的一个测量对象，编者认为它的窄口确实少有人正面占据。但有两位近邻必须请进来。

（一）修辞结构理论中的核—星之分

本章已经引了修辞结构理论（Mann, W. C., & Thompson, S. A. 1988, *Rhetorical Structure Theory: Toward a functional theory of text organization*, Text, 8(3), 243–281），用它作为角色标注体系的参考之一。但本章没有用到这一支最贴近自己的那条：核与星的区分（nucleus / satellite）——一个跨度是核还是星，不由它自身的内容决定，而由它在关系中的位置决定；删掉星，整体大致仍成立，删掉核，上层关系随之改变。

这几乎就是本章「关系中心性」与「中心暴露」的现成表述。

分离线。那一支处理的是成品内部的关系层级：在一份已完成的文本里，哪些跨度承担核、哪些承担星。本章处理的是跨版本的角色再分配：某一跨度的核星身份，是否因为别处的删除而改变。前者是静态标注，后者要求至少两个版本与一次对齐。

判定形式。对同一份文本的相邻两版分别做核星标注。若所有角色迁移都能被表述为「核星身份的改变」，本章的八类角色体系没有额外分辨力，应改用那一支的记号；若存在核星身份不变而本章判为角色迁移的单元（例如证据→问题触发，两者可以同为星），本章的体系更细。

（二）论证成分类别随上下文改变

在论证挖掘一支里，同一个句子被标为主张还是前提，本来就依赖于它在篇章中的位置与相邻成分（可参 Stab, C., & Gurevych, I. 2017, *Parsing argumentation structures in persuasive essays*, Computational Linguistics, 43(3)）。这一支实际上早就承认「同一串文字的论证角色不是它自带的属性」。

分离线。那一支承认角色依赖上下文，但它的任务是在一份给定文本上把角色标对；本章的任务是测量角色改变的发生率，并把它归因到一次特定的历史事件（某成分的退出）。前者视上下文依赖为标注难点，后者视之为研究对象。

判定形式。把同一份现稿交给标注者，一组知道版本历史，一组不知道。若两组标出的角色分布无差别、且都能重现本章的迁移，本章的历史归因成立；若只有知道历史的那一组标得出迁移，本章测到的是标注者的后见，不

是文本的性质。（本章的标注设计已经要求盲于版本历史，此处只是把它挂到这位近邻名下。）

（三）上游那笔账

同前两章，本章亦须补已写出的文本进入任务环境这一构念（Hayes 1996）。本章的「载体」一节列举了工作记忆、撤销栈、版本面板、旁注、教师批注等，实际上已经在做这一支的工作，只是没有点名。

二·编者提示：本章是本卷设计得最干净、也是最空的一章

最干净指的是第八节那个四臂随机设计：无接触、只阅读、写入—删除、写入—删除且历史可见。四组最终看到完全相同的当前稿，只有此前经历不同；平台在进入自由续写前自动规范化文档，使字符、格式、光标位置与可见提示一致，历史可见性是唯一允许保留的界面差异。六个替代解释（普通启动、版本检索、当前缺口、任务需求猜测、标注后见、投入差异）各自配了对照组或控制手段。这是全卷因果识别做得最彻底的一个方案。

最空指的是：它一次也没有被执行。本章通篇是理论建构与实验设计，作者在前篇即声明「本文属于理论建构论文，而不是已经完成验证的实证报告」，并把所有因果判断保留为待检验命题。这一声明是诚实的，编者原样保留。

两件事放在一起，读者应当得到的印象是：本章提供的不是一个发现，是一件工具与一份可以让它失败的说明书。本卷收它，是因为在九章之中，它是最容易被别人接手跑起来的一章——所需的仅是一个能记录版本与按键的写作平台，加两名盲于组别的标注者。

对切一·第三章 ↔ 第五章：文本不变，变的是哪一种角色

为什么必须切。两章都主张「一字未改的文字可以已经变了」，而且都把这一点当作自己最重要的发现。若不切开，读者会认为其中一章是另一章的换名。

分离线。

- 第五章读篇章角色。自变量是别处的增删移位；因变量是盲法标注者读出的角色（背景→主张、证据→问题触发、解释→反驳、结论→让步）。它测的是这份文本对读者呈现的功能分工。
- 第三章读工作角色。自变量是这一段自己在工作区中的位置（可编辑的正文嵌入态 vs 逐字可见但不可编辑的侧栏态）；因变量是随后一段固定时间内的操作转移分布。它测的是这段文字对写作者当下可做之事的约束。

一个在读者那一侧，一个在写作者那一侧；一个的自变量在别处，一个的自变量就在这一段自己身上。

双向判决性对照（列为两章共同的证伪条件）：

其一。把某一段已稳定的文字原样移到只读侧栏，别处一字不动。

- 第三章预测：随后的操作转移分布与嵌入态条件系统不同。
- 第五章预测：该段与其余各段的篇章角色不变——因为没有任何职责退出，关系场没有改变。
- 若此时角色迁移率也随之上升，第五章的「退出」定义过宽，必须收窄：它把「不能再编辑」也算成了一种退出，而那不是它定义的退出。

其二。删掉别处的中心主张，这一段的位置与可编辑性完全不动。

- 第五章预测：该段发生角色迁移（很可能是「中心暴露」——原本的例证成为局部最强命题）。
- 第三章预测：端位效应不变——这一段的工作角色一点没动。
- 若此时端位效应也变，第三章的「工作角色」混入了篇章角色，必须拆分：它测到的东西里有一部分其实来自内容关系的改变，而不是工作位置的改变。

编者按。这一处对切的价值不在于判谁赢，而在于它把两个都很窄的读数放在同一个实验平台上——一台键击记录设备加一套角色标注，可以同时产出两个读数。两章各自的最小实验都不便宜，合起来跑反而更省。

编三

权交回去的那一刻

共同动作：问同一个动作——结算——的三个侧面。

写完就是把权交回去。三章问的是同一次交接的三个位置。

第六章问账内：这一篇还欠着什么？它把「写完了」从作者的感觉改成一张可复核的账本——文本自己激活了哪些前向义务，其中哪些已结清或已合法限界，哪些还欠着。它的验收动作是作者退场：编码阶段作者不得解释，由不知道写作计划的读者去配对义务与结清位置。凡要靠作者口头说明才能连上的，一律记为未结。

第七章问账外：这一次结清，把什么挤了出去？它的判断是，写好作文最危险的时刻不一定是失败的时候，而可能恰恰是某一项已经做得很好、系统因此允许我们宣布「这一轮完成」的时候——因为宣布完成的那把尺子，只看见它被设计去看见的东西。

第八章问账本本身：判它的那把尺子，被它改了多少。前两章都默认判准在作品之外；第八章说判准会被作品回写。

这一编有一处编者必须交代的取源重叠。第六章与第七章都取用了形式化验证，第七章与第三章都取用了机制设计。三家学科在本卷里出现了两次以上。这不是重复取材，因为取的是不同的线：第六章取的是验证义务与可判定结果态，第七章取的是验证与确认之分（符合规格不等于满足真实用途）；第三章取的是多重均衡，第七章取的是激励相容与可测偏置。但编者仍要指出，第七章三家里有两家在本卷别处已经出现，这是它在取源独立性上不如其余各章的一处实情，也是编者把它夹在第六章与第八章之间的理由——让它在两侧的对照中被读，而不是单独被读。

第六章 作文怎样完成

——退场闭合态

作文不是“写尽”，而是获得退场许可： 从继续可写到整体闭合的“退场闭合态”

——形式化方法与程序验证 × 争议解决与仲裁 × 风险、安全与
可靠性工程的二阶碰撞

学科通融专栏·二阶碰撞典范文（How3）

研究类型：How（三焦点路径） | 研究对象：作文从持续可写到整体闭合的
完成机制

版本：v1.0 | 方法依据：SDE 多学科通融创新法 3.9 版 | 日期：2026 年 8 月
17 日

摘要

“作文怎样完成”通常被降格为“结尾怎样写”：点题、照应、升华、总结、留白，仿佛只要最后一段处理得漂亮，一篇文章便自然获得完成性。然而真实写作持续暴露一个更基础的困难：一篇作文即使已经有开头、主体和结尾，作者仍可能感觉“哪里没有交代完”；另一篇文章明明还可以再加一个故事、一个例子、一个反方观点，却已经具有不可随意增删的整体感。这说明“还能继续写”与“尚未完成”不是同一件事。本文依照 SDE 多学科通融创新法 3.9 版的 How 型三焦点路径要求，从新思想前沿中选择形式化方法与程序验证、争议解决与仲裁、风险、安全与可靠性工程三门尚未被作文教学固定吸收且理论体系成熟的学科进行二阶碰撞。三家分别把“完成”压在三个互不相同的终点上：形式化验证要求待证性质对应的验证义务获得判定，而不要求穷举全部可能状态；仲裁把终局性建立在请求边界、程序结清与裁决效力上，而不要求现实世

界再无争议；风险与安全工程则要求系统在操作者退出、环境变化和边界扰动下仍能保持可接受约束，而不是以“目前没有事故”冒充安全。三者互相冲突后共同推翻一个长期支配作文教学的隐含前提：完成必须等于内容耗尽，或者由某个局部终点——最后一句、中心点明、规定字数、写作时间结束——单独结算。本文据此提出新对象“退场闭合态”：在一个被冻结的篇章边界内，文本已经激活的必要续写义务均被结清或显式限界；作者退出后，独立读者无需作者补充即可维持核心关系；对结尾、例证与非承重文字做扰动时，不会重新暴露未结核心义务；此后若继续添加内容，新增文字主要是在创造新的义务，而非偿还旧义务。本文进一步提出单一读数“退场闭合率（Exit-Closure Ratio, ECR）”，给出义务识别、核心义务判定、结清链接与作者退出测试的清点手册，并在公开 ASAP 自动作文评分语料的 12 篇同题学生作文上进行探索性真跑。结果显示 ECR 与人工总分正相关，但在中高分段迅速饱和；一篇人工分数仅为 6 分的作文亦可达到 0.92，而一篇没有传统总结段的作文可以达到 1.00。这个不利结果反而划清了新概念的边界：退场闭合不是“好作文”评分器，而是独立回答“这篇文章是否已经完成”的过程判据。由此，作文教学可把“再写一点”改写为“还欠哪一笔”，把“写个结尾”改写为“作者能否退出”，把修改从无限润色改写为对未结义务的定向清算。

关键词：作文完成；写作教学；二阶碰撞；退场闭合态；续写义务；退场闭合率；程序验证；仲裁；可靠性工程

二阶碰撞工序摘要

本表只呈现成文所需的学术化结果，不在正文中使用 S/D/E 母体术语。How 型要求三家各锁一个不同焦点，最终产物不落在任何一家，而落在三条路径互相取消单独终点之后出现的新对象。

焦点	取源	本家“走到哪里才算走完”	抽取的理论脊	对作文完成的压力
结果焦点	形式化方法与程序验证 (第 352 号)	“走到可判定结果态才算走完”	性质导向验证、验证义务、未证明部分	把“写尽”改成“结清与心对象相关的必要义务”
程序焦点	争议解决与仲裁 (第 426 号)	“走到请求边界内获得终局才算走完”	请求边界、程序节点、终局、执行	把“世界无争议”改成“篇边界内有结算”
环境焦点	风险、安全与可靠性工程 (第 600 号)	“走到操作者退出后仍守住约束才算走完”	韧性、边界场景、开放环境、近失误	把“作者自觉写完”改成“作者退出仍稳定”

3.9 版闸零记录：未吸收度与理论厚度

学科	未吸收度	厚度	厚度依据	作文领域检索判定
形式化方法与程序验证	干净	厚	Hoare 逻辑；模型检查/定理证明/分离逻辑等长期分歧	检索“作文教学+程序验证/形式化验证”；未见完成判据成为作文理论定术语
争议解决与仲裁	刚进入（名称/资源层），答案级干净	厚	终局性 vs 司法审查；效率 vs 正当程序；执行 vs 形式裁决	法律写作/作文评分中可 ADR 或“仲裁”字样，但见“请求边界+终局性”作文完成判据
风险、安全与可靠性工程	干净	厚	故障链 vs 系统安全；Safety-I vs Safety-II；FMEA/FTA vs STPA/FRAM	“可靠性”进入写作评价指评分一致性，非复杂统开放环境的安全闭合逻辑

闸零结论：放行。仲裁被判为“刚进入（名称/资源层）”而非“已吸收”，因为写作与 ADR 的既有交叉处理的是法律写作训练、社会冲突或评分程序，尚未把“请求边界—程序终局—执行”这一答案固定为作文完成理论；因此成文期必须以“终局性”而非“冲突管理”作为抽取对象。

一、问题的重新提出：为什么“还能写”不能推出“还没写完”

作文教学最常见的完成信号来自外部：铃声响了、格子写满了、字数够了、最后一段已经出现“总而言之”，或者教师提醒“赶快收尾”。这些信号当然具有课堂管理价值，却都没有回答文章内部的完成条件。学生最困惑的情形恰恰发生在外部信号失灵以后：时间还剩十分钟，他不知道该再加一个例子还是停笔；草稿已经一千字，他仍觉得某处“虚”；文章只有七百字，却再加任何材

料都显得赘余。若没有内部判据，“完成”只能靠熟练者的感觉，而初学者只能模仿结尾样式。

把“还能继续写”等同于“尚未完成”，会把作文推向一个逻辑上无法收束的状态。任何经验都可以继续细化，任何论点都可以再找一个例子，任何叙事都可以向前追溯或向后延展，任何描写都可以再增加一个感官维度。可写空间原则上没有自然终点。一个作者若以“还有东西可加”为停笔否决条件，就永远无法获得完成。由此可见，作文完成所要终止的不是可能内容，而是某种更小、更硬、可以清点的“必要继续”。

反过来，“已经写了结尾”也不能推出“已经完成”。一篇议论文可以在最后重申中心论点，却仍有一个决定性的反例从未处理；一篇记叙文可以用景物描写首尾呼应，却让事物最关键的选择没有后果；一篇说明文可以用“综上所述”结束，却在正文里引入了一个术语而从未界定。结尾只是文本位置，完成则是整体状态。两者经常重合，但没有逻辑上的等价关系。

因此，本题不是“结尾有几种写法”，而是一道严格的 How 题：从一个已经能够持续生成、仍有大量可写可能性的草稿出发，经过什么次序，才能进入一个作者可以停止继续生成、文章又不会因为停止而坍塌的状态？更具体地说，我们需要找到三个东西：第一，尚未完成到底欠着什么；第二，这些欠账按什么次序被结清；第三，怎样证明作者此时可以退出，而不是只因为作者累了、时间到了或暂时想不到新材料。

这个重新表述也把 How3 与前面的作文启动、持续生成区分开来。启动研究的是“第一步怎样发生”，持续生成研究的是“下一步怎样不退化为堆积”，完成研究的则是“什么时候下一步不再是义务”。完成不是生成能力的反面；成熟的完成恰恰要求作者仍然有能力继续写，却知道此刻为什么不应再写。

二、3.9 版选源：先写“走到哪里才算走完”，再找每天替这句话做判断的行当

本文没有先从“哪些学科听起来和写作有关”出发，而是按 How 型三焦点路径先写三句占位句。第一句是“走到一个可以判定的结果态才算走完”；第二句是“走到一条程序获得终局、遗留事项有明确去处才算走完”；第三句是“走到操作者退出后仍能在真实环境中维持约束才算走完”。这三句话分别把焦点压在结果、路径和环境三个终点上。随后才在 SDE Universes “新思想前沿”626 个领域中寻找日常工作正是在处理这些终点的成熟领域。

筛选后的三家分别是第 352 号“形式化方法与程序验证”、第 426 号“争议解决与仲裁”、第 600 号“风险、安全与可靠性工程”。形式化验证近二十年的核心变化之一，是把“系统能运行”改成“哪些验证义务被真正判定、哪些仍超时、近似或未证明”；争议解决与仲裁把制度效果拆成进入、程序节点、裁决、执行等链条，而不是用“有裁决”覆盖所有未结事项；风险、安全与可靠性工程则从元件是否失效转向复杂系统在目标冲突、组织与环境变化中能否持续

守住可接受约束。三家都拥有成熟的理论争论、反例传统和专业手册，同时其核心判断并未成为语文作文教学的固定术语。

未吸收度的核验特别重要。形式化验证虽然偶尔以“验证”一词出现在语言评价、事实核查或自动评分研究中，但“性质导向可达性—验证义务—未证明部分”的整套判据并未进入作文完成理论；仲裁在考试评分流程中甚至存在“仲裁分数”这一同名词，法律写作课程也有人引入ADR，但那处理的是评分争议或法律训练，并非把“终局性与请求边界”作为学生作文的完成判据；可靠性一词在作文评价史上常与评分者一致性相关，却不是“系统在开放环境中维持安全约束”的可靠性工程理论。本文判定的是三家对本题的具体答案是否已被吸收，而不是学科名称是否在教育研究中出现过。

理论厚度也逐家过闸。形式化方法有Hoare逻辑、模型检查、定理证明、分离逻辑、程序细化等奠基层与长期争论；仲裁有程序自治、终局性、司法审查、执行、正当程序之间的成熟张力，并有UNCITRAL规则、示范法及长期判例传统；安全与可靠性工程内部则存在传统故障分析、系统理论事故模型、韧性工程、安全I/安全II、SOTIF等彼此竞争的路线。三家都不是刚出现的流行标签，因此能够提供“什么时候本家说法会失败”的反例，而不需要由本文替它们代写立场。

三家的门类距离也足够大：计算机科学中的形式验证、法学与争端制度、工程安全与可靠性分别拥有不同期刊、职业共同体、奠基人物和训练制度。它们会在高可信系统、合规与安全保证处发生跨引，但并不共享同一套母体语汇。更关键的是，它们对“完成”的终点不在同一位置上争“谁对”：证明可以完成而系统仍不安全，仲裁可以终局而事实争议仍存在，系统可以安全运行而某项形式性质从未被证明。正因为三者无法用一个裁判直接排序，碰撞才有生成新框架的必要。

三、第一条因果脊：形式化验证——完成不是穷尽状态，而是把与性质相关的义务判定掉

形式化方法给作文完成问题的第一个冲击，是它拒绝把“全部可能性都检查过”当作证明成立的必要条件。在有限状态系统里，朴素直觉会要求穷举所有状态；但状态空间很快指数爆炸。Bradley提出的IC3/PDR类性质导向方法改变了方向：从待证性质出发，学习足以阻断坏状态的归纳子句，最终得到不变量。决定性变化不只是算法更快，而是“完成”的对象被重新定义——不是把世界看完，而是把对当前性质必要的证明义务关掉。

这个转换迁入作文非常锋利。学生常把“文章还可以再补材料”理解成“我还没写完”，正像把“还有状态没枚举”理解成“性质还没证明”。若一篇议论文当前要成立的核心判断只依赖三个子主张，那么第四个、第五个例子属于可继续探索的状态，却未必属于当前验证义务。完成的第一步因此不是追问“还有什么可写”，而是追问“我已经让读者承担了哪些必须被兑现的期待”。

分离逻辑和bi-abduction进一步给出更像“欠账”的结构。程序分析不只问执行后得到了什么，还可以反推一个过程需要什么前置资源，并描述执行后

剩下什么。对作文来说，每一次重要的文本动作同样会生成前置要求和残余：提出“真正的问题不在努力，而在比较规则”之后，作者至少欠读者“比较规则是什么”和“为什么它比努力更决定结果”两笔；引入一个反例之后，又欠“这个反例改变了原论证的哪一层”一笔。材料不是中性的库存，材料一旦写进文章，就会创造新的解释责任。

形式验证还迫使我们区分“已证明”与“证明停在哪一层”。CompCert 一类验证编译器研究之所以重要，正因为高层程序即使被证明正确，未经验证的编译转换仍可能破坏性质。对应到作文，一段局部论证完全成立，不等于从这段到全文核心关系的传递已经成立；一个故事很动人，不等于故事与中心判断之间已经有承重链接。作文的“未完成”经常不是缺材料，而是证明链在层间断掉。

由此得到第一条占位判断：一篇作文走到“与核心对象有关的续写义务已经得到文本内判定，未判定部分被明确排除在当前边界之外”才算走完。这条判断直接反对“多写等于更完整”。它允许文本仍有无限可扩展世界，却要求当前自己激活的必要义务必须有去处。

四、第二条因果脊：仲裁——完成不是世界再无争议，而是请求边界内获得终局

仲裁给出的完成观与形式验证明显不同。一个裁决成为最终且具有约束力，并不意味着双方从此在事实、情绪、商业关系或道德评价上再无分歧；它意味着在被提交、被受理并经过程序处理的请求边界内，制度把“继续争”转化为“执行、补正、撤销审查或其他有明确入口的后续行动”。终局性首先是一种边界化的程序地位，而不是对现实真理的穷尽。

这对作文完成尤其关键，因为文章同样必须先有“请求边界”。一个八百字议论文讨论“竞争是否促进成长”，它不可能同时结清竞争心理、教育制度、劳动市场、体育伦理、家庭期望与历史文化的全部争议。若作者把“所有相关问题都解决”设为完成条件，文章注定失控。成熟的文章需要做的不是封闭世界，而是明确：本篇究竟承担哪一个判断，哪些争议被纳入，哪些只能留在边界之外。

UNCITRAL 2010 仲裁规则强调多方、追加、临时措施、指定机构等程序安排，说明一套争议解决机制若想真正运行，不能只等最后一纸裁决；进入、节点、证据、程序权利和执行必须形成链条。类似地，作文不能把所有“完整性”都外包给最后一段。开头提出的问题、正文建立的分歧、例证的解释、反方的处理、概念的界定，本身就是程序节点。最后一句只能结束已经被处理的事项，不能替前文补做从未发生的程序。

仲裁还提供一个非常有价值的反直觉：裁决之后仍可能存在补正、执行或有限审查，但这不取消裁决的终局性质。把这一点迁入写作，可以严格区分“完成”与“不可修改”。一篇已完成的作文仍可以润色句式、更换更准确的例子、增加风格层次，甚至在新的版本中扩大问题边界；这些后续变化不证明旧

版本当时没有完成。完成只是说：在当时冻结的请求边界内，不再存在必须依靠继续写作才能避免逻辑坍塌的事项。

由此得到第二条占位判断：一篇作文走到“由自己提出的核心请求、冲突与承诺在当前篇章边界内获得终局性，剩余事项都有明确的边界地位”才算走完。这条判断反对“文章必须把问题说尽”，也反对“留下开放性就等于没完成”。开放可以是经边界处理后的合法开放，欠账则是作者没有处理却假装已经终结的事项。

五、第三条因果脊：风险、安全与可靠性工程——完成必须经得起操作者退出与开放环境

风险、安全与可靠性工程给出的完成观再次换位。传统可靠性很容易把“部件没有坏”“测试没有出事故”当成成功，但近二十年的系统安全研究不断提醒：复杂系统的危险可以来自组件都正常时的交互，已知场景上的成功也不能代表未知边界场景。韧性工程因此把安全从“没有事故”改成预见、监测、响应、学习等持续适应能力；SOTIF 则明确把没有随机硬件故障时，由功能性不足和边界环境触发的危险纳入安全问题。

作文也存在“组件都正常而整体失效”的情形。每个段落都能单独读懂，每个例子都真实，每句话语法正确，最后却没有形成一个整体判断；或者作者在旁边解释时一切都说得通，一旦只留下文本，读者便不知道某个转折为什么发生。若完成判据只看作者在写作现场的自我感觉，就会把大量“现场可运行、离场即失效”的草稿误判为成品。

安全工程中最值得迁移的不是“风险”这个大词，而是操作者退出后的边界测试。工程系统要进入真实运行，就必须考虑环境变化、近失误和已知未知的场景；作文在交付后同样进入一个作者不能实时补丁的环境。读者不会自动拥有作者脑内的背景，不会知道“这件事我前面其实想过”，也不会把口头解释算入正文。作者退出不是写作之后的偶然事件，而是作文能否成为独立文本的验收条件。

这一家还与形式验证形成直接冲突：形式上所有局部义务都可能被作者自认为“已经处理”，但真实读者仍会在某个连接处失去方向；就像系统满足若干单项要求却在开放场景中出现危险。风险工程因此要求把“未发生问题”改成“在扰动下仍守住约束”。对作文而言，最便宜的扰动包括：去掉最后一段，核心关系是否立刻消失；删除一段非承重文字，意义是否反而更清楚；让不知道作者意图的同伴复述核心冲突，是否需要作者现场补充。

由此得到第三条占位判断：一篇作文走到“作者退出并接受最小扰动以后，文本仍能在独立读者端维持其核心约束”才算走完。这条判断反对把“我知道自己想说什么”当作文章已经说完，也反对把“现在没发现问题”当作整体稳定。

六、斜对立：三家为什么不能简单相加

如果把三家理解成“证明要严谨、程序要完整、系统要安全”，碰撞会迅速退化为一阶拼盘。真正的张力在于：每一家认为自己的终点足以结束，但另

外两家都能制造一个它单独无法处理的失败案例。形式验证可以宣布一组预先写下的性质全部证成，但风险工程会追问：性质有没有漏掉真实环境中的关键危险？仲裁可以在程序上获得终局，但形式验证会追问：某个决定性论证义务是否其实没有被处理，只是因程序期限而被关门？系统可以在当前环境中稳定运行，仲裁又会追问：你究竟以什么边界宣布“这部分已经不再由本案承担”？

把作文放进这组三角冲突，会出现三个常见伪完成。第一种是“证明型伪完成”：作者把中心论点和三个分论点都写齐，自认为逻辑完成，却没有让独立读者看见三者如何组成一个整体；第二种是“程序型伪完成”：按模板把开头、三段、结尾全部走完，仍有决定性承诺没有兑现；第三种是“运行型伪完成”：同桌读后说“挺顺的”，但因为双方共享题目、课堂范文和教师暗示，文本自身并没有承担全部解释工作。

三家的第一层共有前提是：完成可以由某一个终点单独结算——证明完成、程序完成或运行稳定。更深一层共有前提则是：停止行为本身不改变被检验对象。可是在作文中，作者是否退出恰恰改变了系统。作者在场时可以补充语境、解释指代、说明“我这里其实是反讽”；作者离场后，这些补丁全部失效。于是“能不能停止”不能在作者仍持续提供解释的条件下检验。

推翻这个共有前提的材料来自三家自身。形式验证承认规范缺口与可信基：证明只能证明被写入规范的东西；仲裁承认终局性只对受理边界内的请求生效；安全工程承认开放环境会制造未被测试覆盖的新情境。三家自己的反例共同说明：所谓“完成”从来不是把世界耗尽，而是在清楚边界、明确义务和离场条件下获得一种可撤回、可复核的终止资格。

由此，本文不把三家汇成“更完整的完成”，而是取消一个旧对象：“内容耗尽后的终点”并不存在。真正存在的是一种由边界、义务和离场稳定性共同生成的状态。它既不是最后一句的位置，也不是作者的心理满足，也不是教师的形式检查。

七、新对象 Z：退场闭合态——作文完成的对象不是“结尾”，而是“作者可以退出的整体”

本文将二阶碰撞后出现的新对象命名为退场闭合态（author-exit closure state）。它指：在一个已经冻结的篇章边界内，文本自身激活的必要前向义务都已获得结清或明确限界；作者退出以后，独立读者仍可以仅凭文本维持核心关系；对结尾、非承重例证或局部措辞进行最小扰动时，不会暴露新的核心欠账；继续添加内容若发生，主要是在扩大边界或创造新的义务，而不是偿还旧义务。

定义中的每一个限定词都有划界作用。“冻结边界”排除把全世界都装进一篇作文的要求；“文本自身激活”排除教师事后无限追加标准；“必要前向义务”排除所有可选的丰富化；“结清或限界”允许文章合法留白与开放结尾；“作者退出”排除口头补救；“独立读者”排除熟悉作者意图的同伴代写理解；“最小扰动”排除靠最后一句强行维持的脆弱完整；“新增内容主要创造新义务”则给停笔提供正向判据。

“续写义务”不是任何读者想知道但作者没写的东西，而是由当前文本动作产生、若不处理便会改变文章核心判决的前向责任。例如，作者写“真正导致我放弃的不是失败，而是成功之后的空白”，就至少激活“失败为何不是主因”“成功后的空白是什么”“它如何导致放弃”三项；若后文只讲失败多痛苦，第一句便成为未结义务。相反，作者没有交代人物午饭吃什么，通常不构成义务，因为这项缺失不改变核心关系。

退场闭合因此与“信息完整”相反。信息越多，可能激活的义务越多。一个新人物出现，需要身份与作用；一个新概念出现，需要界定与应用；一个新例子出现，需要解释它究竟支持哪一层。材料堆积为什么常让文章越来越难结束？不是因为篇幅变长，而是因为每次增加材料都可能提高“义务本金”。如果新增结清速度慢于新增义务速度，文章会越写越欠。

这个新对象还解释了成熟作者常说的“这篇已经到了，再写就散了”。这里的“散”不是字数变多，而是新内容需要打开一个尚未属于当前文章的争议分支，迫使原有边界重写。成熟停笔并非枯竭，而是识别到下一笔将从“结清旧账”切换为“开新账户”。这正是从持续生成跨入整体闭合的临界点。

八、从“义务”到可操作清单：一篇作文到底会欠下哪些续写债

为了避免“续写义务”变成一个漂亮比喻，本文把它压成课堂可编码的单位。第一类是主张义务：中心判断、分论点、因果断言、比较判断一旦写出，就欠支持或限定。第二类是冲突义务：文章主动建立“不是 A 而是 B”“看似 X 其实 Y”“一方面/另一方面”时，必须处理对立双方及转折依据。第三类是例证义务：故事、数据、人物或事实一旦作为证据进入，就必须说明它改变哪一个判断，不能让例子以“很感人/很具体”自我结算。

第四类是概念义务：关键词语若承担区分功能，就需要足以让读者知道“它不是什么”的边界。第五类是叙事义务：人物欲望、异常事件、承诺、危险或悬念一旦被赋予核心地位，就需要结果、转化或合法悬置。第六类是指向义务：代词、时间跳跃、视角切换、因果连接等若影响核心关系，文本必须提供足够锚点。第七类是反方义务：作者主动提出一个足以伤害自身主张的反对意见，就需要回答、让步或收窄，而不能只把反方写出来证明自己“全面”。

这些类型不是作文评分的新七维量表。编码时真正的单位仍是一笔具体义务，例如“解释为什么母亲的沉默不是冷漠而是保护”是一笔，而不是给“叙事义务”打几分。类型只用于帮助教师和学生发现欠账。每笔义务还必须标记“核心/局部”：移除或误解后会改变全文核心判断、人物关系或主要冲突的，记为核心义务；只影响局部理解但不改变整体判决的，记为局部义务。

一笔义务有三种合法状态。其一是结清：后文出现可定位的文字，使读者能处理这一期待。其二是限界：作者明确告诉读者这不是本篇要解决的事项，或者通过文体契约把它转为有意开放。例如“我不知道父亲为什么没有回头；我只知道那一次以后，我不再把沉默当作同意”，前一个问题可以被合法悬置，因为文章将判断边界转到“我如何理解沉默”。其三是未结：既未处理也未限界，却仍影响核心。

这个清单立刻改变修改提问。教师不再笼统说“这里再具体一点”“结尾再升华一下”，而可以问：“第二段提出了‘不是没努力，而是目标被别人定义’，后面哪一句真正处理了‘别人怎样定义’？如果没有，这是一笔核心欠账；补这笔之前不要加新材料。”学生第一次可以知道，自己需要继续写的不是“再来一点内容”，而是一个有名字、有来源、有结清位置的义务。

表 1 续写义务的最低编码类型

义务类型	触发规则	典型触发	结清判据
主张义务	中心判断、分论点、因果或比较一旦提出，需要支持或限定	“真正决定 X 的不是 A 而是 B”	后文分别处理 A 为何不足、B 为何决定，并建立证据链接
冲突义务	主动建立对立、转折或悖论，需处理冲突结构	“一方面……另一方面……”	双方都有去处，且说明为何并存/何者在何条件下占优
例证义务	例子作为证据进入后，需要解释支持哪一层判断	叙述一个人物事件	明确回写到分论点，或说明其作背景
概念义务	承担区分功能的关键词需要边界	“这里的自由不是随心所欲”	给出足以排除近邻的定义/判例/对照
叙事义务	核心欲望、异常、悬念、承诺、危险需有结果或合法悬置	“他答应第二天一定回来”	回来/不回来及其意义被处理或明确把悬置变成主题
指向义务	影响核心关系的指代、时间、视角变化需要锚点	“后来我才明白”	读者能定位“后来”“明白什么”“为何改变”
反方义务	足以伤害己方主张的反对意见一旦提出必须回应	“有人会说……”	反驳、让步或收窄主张，不忽略结算

九、单一读数：退场闭合率 ECR 及清点手册

本文用一个单一读数把新概念变成可证伪对象：退场闭合率（Exit-Closure Ratio, ECR）= 已经结清或合法限界的加权续写义务 / 文本已经激活的全部加权续写义务。全文只使用这一口径。核心义务权重为 2，局部义务权重为 1；“合法限界”与“结清”在分子中都计满，因为二者都使作者无需继续写作才能维持当前边界；完全未结计 0。若一笔义务只有部分处理，原则上不做 0.5 式主观打分，而拆成更小的可判定义务。

清点必须在看最终评价之前进行。第一步，冻结版本和篇章边界：记录题目、文体、当前版本、作者自报的核心判断或核心关系，但不允许作者解释正文之外的内容。第二步，从文本中逐段标记“向后提出要求”的句子或结构，生成义务 ID。第三步，为每笔义务标记核心/局部与触发位置。第四步，只在后文能定位到结清文本时建立“触发—结清”链接；若靠作者口头说明才能链接，记为未结。第五步，计算 ECR。

作者退出测试是 ECR 的必要编码条件而不是第二个指标。编码者必须在作者不参与的情形下完成义务识别与结清判断；对争议义务，可让至少两名读者独立标记再讨论达成规则。其目的不是追求绝对客观，而是防止作者在验收阶段把脑内意图补进文本。ECR 测的是“这篇文字是否把自己打开的账结掉”，不是“作者心里有没有答案”。

本文给出一个试用的教学阈值： $ECR \geq 0.90$ 且不存在未结核心义务，标记为“达到退场闭合”； $ECR < 0.90$ 或仍有任一未结核心义务，标记为“需要继续结清”。0.90 不是自然常数，而是未来研究应校准的初始门槛。真正需要守死的是口径：阈值可以通过数据修订，但同一研究、同一课堂实验中不能在正文、证伪条款和结果解释之间偷偷更换。

这个读数有两个重要的反向约束。第一，ECR 不能靠删掉所有复杂内容无限提高。若删掉一个核心义务的触发句同时改变了作者原本冻结的核心对象，属于“缩题重写”，必须建立新版本重新冻结边界，不能把它当成本版本的结清。第二，ECR 不能奖励模板化总结。若最后一段只是重复“所以我们要珍惜时间”，却没有处理正文里已激活的因果断点，它不会让那些义务进入分子。结尾只有在真实结清某笔账时才计入。

十、作文完成的实际路径：边界冻结 → 义务清算 → 作者退场 → 失败回写

退场闭合不是一个静态评分表，而是一条可执行路径。第一步是冻结边界。学生用一句不超过三十字的话写下“本篇最终必须让读者建立的关系”，并列出生篇主动不处理的两个相邻问题。这个动作不是先把“中心思想”写死，而是划出本轮结算范围。文章在写作中可以改变边界，但每次改变都要留下版本，否则作者会在成品出现以后倒填一个看似早就存在的目标。

第二步是拉出义务账本。草稿写完一轮后，不先润色语言，从第一段开始只找“这句话要求后文做什么”。每发现一笔义务就在页边写 ID，并在结清位置画回链。画不回链的先不要补内容，先判断它究竟是核心义务、局部义务，还是应该明确限界。很多材料堆积会在这一步暴露：学生原以为自己“内容很丰富”，账本却显示后半篇持续开新账、几乎不结账。

第三步是优先结核心债。核心义务必须先于局部修辞处理。若核心因果链缺一环，此时再增加景物描写、名人名言、第二个同类例子，只会扩大表面完成感而不改变 ECR。教师可以规定：一次修改只准处理一至两笔核心债，处理完再看局部。这样修改从“全文再来一遍”变成可完成的小任务。

第四步是作者退场测试。把稿件交给不知道作者原计划的同伴，只允许对方在文中划出“我认为文章在解决什么”“哪一处让我等待后文回答但最终没等到”。作者不能解释，只能记录。若读者指出的等待与义务账本重合，直接回写；若读者新提出的问题实际上不影响冻结的核心对象，则不强迫作者扩写，而要判断是否需要一句边界提示。

第五步是最小扰动。删除公式化结尾、一个非承重例子或一段华丽但可替换的描写，观察是否使核心义务重新变成未结。若删除最后一句就让全文中心首次出现，说明文章此前依赖末端补丁；若删除一整个例子 ECR 不变，说明该例子可能是可选材料，不是完成所需。扰动的目标不是“破坏文章”，而是找出完成性真正压在哪些文字上。

第六步是停笔判定。当 ECR 达到门槛、无核心欠账、作者退场后读者不需补丁，此时询问最后一个问题：“下一段若继续写，主要是在偿还哪一笔已有义务？”如果答不出具体的 ID，而新增内容将打开一个新的分支，则停笔。完成由此获得一个可说清的理由：不是“我没话了”，而是“当前账已结清；再写是在开新账”。

十一、为什么“结尾技巧”经常有效，却不是完成的本体

传统作文教学关于结尾的经验并非错误。总结、点题、照应、回扣题目、留白、景物收束等手法之所以常常有效，是因为它们可能承担最后一笔真实义务：把分散段落重新归入核心判断，结束叙事冲突，给前文意象一个新含义，或者明确边界。但这只能说明“结尾常常参与闭合”，不能推出“结尾就是闭合”。

最典型的错误是“末端代偿”。学生前文没有建立“父亲的沉默为何是爱”的证据链，最后突然写“那一刻我终于懂得了深沉的父爱”。阅卷者因为熟悉“父爱”题型，很容易用文化脚本替文章补上关系，于是成品看起来完成。退场闭合框架会把最后一句视为一个新的主张义务：它反而需要前文提供足够支撑。若删除最后一句，读者完全无法从前文建立同一关系，则结尾不是收束，而是唯一证据。

另一种情况恰好相反：文章没有传统总结段，却已经完成。三个分论点都被处理，反方已被限定，最后一个例子完成了对核心判断的回写，此时继续加“综上所述”只是在低信息密度地重复。许多成熟散文、小说片段和高质量议论都利用这种“停止即结尾”：不是因为作者忘了写最后一段，而是全文义务在前一句已经结清。

因此，课堂上更准确的问题不是“结尾用了哪一种方法”，而是“结尾到底结清了哪一笔”。如果一段结尾不能指向任何未结义务，它只是风格选择；如果它承担三四笔正文从未处理的核心债，它又承担过重。理想的结尾不是“越有力越好”，而是在篇章最后阶段以最低新增义务成本完成必要的最后结算。

十二、近邻划界：为什么“退场闭合态”不能被已有概念 1:1 替换

第一位近邻是结构完整。结构完整通常检查开端、发展、转折、结尾或论点—论据—结论等部件是否齐全；退场闭合检查的是这些部件激活的义务是否结清。一篇五段式作文可以结构齐全而核心因果未处理；一篇没有显式结论段的文章也可以义务全清。用“结构完整”替换退场闭合，会丢失义务账本与作者退出两个字段。

第二位近邻是连贯与衔接。Halliday 与 Hasan 以来的衔接研究以及文本语言学的 coherence 概念处理句段之间如何形成可理解关系。连贯是退场闭合的重要条件，却不是终止条件。一个连贯的草稿可以清楚地把问题越开越大，每一段都接得上，最后仍欠大量承诺；反之，某些风格化文本表面衔接很弱，却可能完成其核心义务。

第三位近邻是写作过程中的修改/修订。Sommers、Flower 与 Hayes 等研究把写作从线性“计划—写作—修改”改成递归过程，这对理解作文发生至关重要，但递归模型本身没有给出一个可清点的停止判据。它解释“作者如何继续改”，本文处理“哪一类继续修改已经不再是必要义务”。

第四位近邻是任务完成/切题。完成考试任务、达到字数、覆盖题目要求属于外部合规；退场闭合是文本内部生成的结算。学生可以完全切题却在自己提出的因果链上欠账，也可以写出一篇内部闭合但偏离考试任务的文章。两种失败必须分开，否则教师会把“合规”误当成“篇章完成”。

第五位近邻是叙事闭合、诗性闭合与“终结感”。Kermode 与 Barbara Herrnstein Smith 等传统认真研究“结尾如何产生终结感”，这是本文最危险的理论占位者。分离线在于：审美闭合关注时间、形式期待、主题回返与读者的结尾感，常与特定文类相关；退场闭合是一种跨叙事、说明、议论皆可编码的义务状态，并允许文本具有审美开放结尾，只要开放已经被合法限界。

第六位近邻是首尾呼应。它是一种结构同构，可以增强整体感，却可能只制造视觉上的环形。若开头问“为什么我害怕成功”，结尾重新出现同一句，却全文从未解释成功如何转成恐惧，呼应没有结清义务。反过来，文章完全不呼应也可以完成。

第七位近邻是“好作文”综合质量。内容丰富、语言准确、立意深刻、组织清晰、风格鲜明都可以影响质量，但“完成”只是其中一个更低维的门槛条件。一个论证平庸、语言粗糙的作文仍可能把自己开的账全部结清；一篇极具才华的草稿也可能故意或无意留下核心断裂。本文的真跑会把这个边界直接暴露出来。

第八位近邻是读者导向/reader-based prose。Flower 早已指出，作者私有思路不能直接等同于读者可理解文本。退场闭合继承“作者必须退出”的问题意识，但再推进一步：它不只问表达是否面向读者，而是给出“哪些前向义务在作者退出后仍未结”的账本和停笔条件。一个高度 reader-based 的文本可以十分清楚，却仍在最后打开一个没有处理的新论点。

十三、敌意拓宽：八位占位者如何围住新命题，又为什么没有把它吃掉

为了避免“换三个学科就宣布创新”，本文把近邻搜索向五个方向敌意拓宽。经典修辞方向包括亚里士多德《修辞学》第三卷关于 peroratio 的讨论，它明确处理演说如何在末端回顾、激发情感和强化记忆；文学理论方向包括 Kermode 关于终结感的讨论和 Smith 《Poetic Closure》对诗如何结束的系

统研究；作文过程方向包括 Flower 与 Hayes 的认知过程模型、Sommers 的修订研究、Bereiter 与 Scardamalia 的知识陈述/知识转换理论；文本方向包括 Halliday 与 Hasan 的衔接理论及 de Beaugrande 与 Dressler 的文本性标准；应用方向则包括当代作文教学中“点题、照应、总结、升华、开放式结尾”等实务。

这些占位者分别能解释新理论的一部分：末端技术、审美终结感、递归修改、读者适应、文本连贯、任务结构都不是空白。因此，若退场闭合只说“文章要有整体感”“结尾要服务全文”“修改要面向读者”，创新即失败。本文保留下来的不可替换字段只有三个同时成立：其一，文本动作会生成可清点的前向义务；其二，完成是这些义务在冻结边界内获得结清或限界，而非全世界意义耗尽；其三，完成判定必须在作者退出条件下进行，并允许用扰动检查闭合是否由文本自身承担。

1:1 替换测试也给出相同结果。把“退场闭合态”替换成“结构完整”，作者退出测试消失；替换成“连贯”，未结义务消失；替换成“结尾效果”，全文账本消失；替换成“任务完成”，内部生成义务消失；替换成“读者意识”，停止条件消失；替换成“叙事闭合”，跨文类与可清点指标消失。没有一个近邻能无损保留三个字段，因此新对象仍有剩余解释量。

同时，新理论必须主动承认一条限制：在高度实验性的文学写作中，“制造无法结清的义务”本身可能是作品策略。此时不能用学校议论文的闭合阈值直接判失败。退场闭合的适用对象是存在可冻结核心对象、具有教学或沟通责任的作文任务；若文体契约把无解、歧义或持续悬置本身设为核心对象，那么“合法限界”的编码规则必须随契约调整。这个限制不是退让，而是防止概念吞并所有写作。

十四、真跑一条：公开 ASAP 学生作文上的 12 篇探索性编码

3.9 版要求新概念不能只在纸面上可证伪，至少要用公开数据跑通一条最便宜的检验。本文选用 2012 年 Hewlett Foundation 自动学生评价奖 (ASAP) 公开训练语料的 essay set 1。该语料含人工专家评分，公开镜像训练集共 12,976 篇、8 个题组；本文只取同一题组，以避免不同题目与评分范围混杂。探索性样本按人工总分 2、4、6、8、10、12 各取前两篇，共 12 篇。此抽样不用于估计总体效应，只用于检查 ECR 能否实际编码以及它是否会立刻被“作文总分”吸收。

为了降低编码自由度，真跑只使用“显式或可直接恢复的前向义务”：中心判断及其列举理由、明确的对立结构、核心反问、作为论据进入的例证、段首承诺的 first/second/last 等。核心义务权重 2，局部义务权重 1；若后文有可定位的处理或清楚限界即结清。评分在编码之后才打开用于比较。由于本次只有单一编码者，结果只能视为方法学试跑，不能当作量表效度证据。

结果出现两个方向。第一，ECR 与人工总分总体同向，12 篇样本中的 Spearman 等级相关约为 0.93，说明低分极短文本往往同时存在明显未结义务；这符合直觉。第二，也是更重要的不利结果：ECR 在中高分段迅速达到天花板，人工 8—12 分的 6 篇全部达到 1.00；同时，一篇人工总分仅 6 分的作文仍达到 0.92 并且没有未结核心义务。换言之，闭合一旦达到，继续提高语言、内容、论证质量并不会继续提高 ECR。

另一个直接反例来自样本 ID 5：文本没有传统意义上独立的总结段，却把开头列举的三个理由依次处理完，末段结束于第三个理由的展开，按本研究编码 ECR 为 1.00。相反，样本 ID 14 具有非常显眼的“In conclusion”式收束，人工总分仅 6，但 ECR 仍达到 0.92。两例共同说明：结尾样式与退场闭合既不是充分关系，也不是必要关系。

这个结果对本文是有利中的不利。若 ECR 能够精准预测总分，反而说明它可能只是在换名重做综合作文质量；现在的天花板说明它测到的是一个更窄的状态变量。完成之后仍有巨大质量空间：词汇可以更准确，思想可以更深，证据可以更强，节奏可以更好。本文因此撤回任何“ECR 越高作文越好”的解释，只保留“ECR 用于判断是否仍存在必要续写义务”这一较窄命题。

表 2 ASAP essay set 1 的 12 篇探索性 ECR 编码

essay_id	人工总分	激活义务 (加权前单位)	已结清权重	ECR	未结核心义务
41	2	4	2.00	0.50	是
359	2	2	0.00	0.00	是
19	4	3	2.00	0.67	是
27	4	5	3.75	0.75	是
14	6	6	5.50	0.92	否
15	6	4	3.00	0.75	是
1	8	4	4.00	1.00	否
5	8	4	4.00	1.00	否
4	10	5	5.00	1.00	否
7	10	4	4.00	1.00	否
16	12	4	4.00	1.00	否
37	12	5	5.00	1.00	否

注：本表为探索性单编码者试跑，义务数与结清权重按本文清点手册人工编码；不用于报告总体效度。12篇中人工分与 ECR 的 Spearman 等级相关约为 0.93，但 8—12 分样本 ECR 全部达到 1.00，出现明显天花板。

十五、真跑的第二个不利发现：公开作文数据几乎不保存“完成过程”所需字段

对 ASAP 语料做字段审计还暴露出更根本的问题。数据保存了最终作文文本、题组、若干人工评分和部分分项分数，却没有草稿版本、作者何时决定停笔、作者在停笔前自报的未结问题、同伴读者的等待点、结尾删除后的理解变化等字段。也就是说，现有大规模作文数据主要把“最终成品”当作研究对象，而本题真正关心的“从可继续到闭合”发生在版本序列中。

这意味着 ECR 在单一最终文本上只能做“事后闭合审计”，不能直接验证因果命题“结清义务导致作者停笔”。若要真正检验 How3，需要新的过程数据：至少保存每次重大增删的时间戳、每轮义务账本、作者停笔判断、独立读者的未结标记以及后续是否重新开写。缺少这些字段时，任何关于“学生为什么在这里结束”的解释都只能是推断。

因此，本文把这一字段缺失原样写入理论边界，而不把最终分数冒充过程证据。未来研究最便宜的设计不是再收一万篇成品，而是在一间教室里让 30 名学生完成两轮写作：第一轮按常规教学，第二轮保留“边界—义务—结清—退场”日志。只要能够观察到未结核心义务是否在停笔前显著下降，以及退场测试失败是否预测后续必要修改，就能对本文的因果机制进行真正检验。

十六、由此重写作文教学：从“加材料”转向“结义务”

第一项改变发生在写后第一次诊断。传统诊断常从“内容够不够丰富”开始，教师一眼看见篇幅短就让学生“再举个例子”。退场闭合教学先问：现有文字已经开了哪些账？若一篇短文只有三笔义务且全部结清，增加材料不是必要修改；若一篇长文有十二笔核心义务只结了七笔，继续加新例子反而危险。篇幅从诊断起点降为资源约束。

第二项改变发生在结尾教学。教师可以继续教照应、点题、总结、留白，但每一种技巧都必须附加一句“它适合结哪类义务”。总结适合合并分散分论点，照应适合完成前文设置的意象或问题，留白适合把无法在当前边界内解决的事项合法限界。学生不再为了“像作文”而机械套结尾，而是先看最后还欠哪一笔。

第三项改变发生在修改顺序。先核心义务，后局部义务，最后语言润色。许多学生的修改无效，是因为他们在结构债没有清掉前就进入词句装修。新的顺序可以非常机械：红色标核心未结，蓝色标局部未结，绿色标已结；第一轮只准处理红色。这样“修改”第一次有明确结束点，而不是老师一句“再好好改改”。

第四项改变发生在同伴互评。同伴不需要给“文采几分”，只需要做两个工作：标出“这里让我等待后文回答什么”，以及“我读完后仍必须问作者什么”。

才能理解核心”。这两个动作比泛泛评价更容易训练，也更接近作者退出测试。互评者的差异本身可以成为数据：若五名读者都在同一处等待，说明那笔义务高度稳定；若只有一人提出，可能是个体兴趣而非文本债。

第五项改变发生在教师反馈语言。从“再具体些”改成“这句话开了哪笔账”；从“结尾没升华”改成“全文还有哪笔核心义务只能靠最后一句补”；从“内容不够”改成“目前缺的是新材料，还是已有材料与判断之间的结清链接”；从“不要跑题”改成“你是不是在中途扩大了请求边界”。反馈由形容词变成可执行操作。

十七、不同文体怎样进入退场闭合：同一个机制，不同义务形状

记叙文的核心义务多来自异常、选择与关系变化。一个人物被赋予欲望，读者便等待其行动或失败；一个动作被写成异常，读者便等待它为何异常；一个“我从此才知道”会激活认知变化的证据义务。记叙文不要求所有事件有结局，但核心事件必须完成结果、转化或合法悬置。开放结尾不是不结账，而是把“事件结果”转换成“读者知道为什么结果被悬置”。

议论文的核心义务主要来自主张、理由、反方与证据解释。学生最常见的欠账不是“没有例子”，而是例子与主张之间没有结清链接：写完名人故事后只说“由此可见坚持很重要”，但故事也同样可以证明天赋、运气、环境或资源。此时真正未结的是排除替代解释的义务。议论文完成的标志不是论据数量，而是当前核心主张不再依赖作者口头补充“这个例子为什么算我的证据”。

说明文的义务来自对象边界、分类标准、机制与例外。一旦文章说“有三种”，就创造数量承诺；说“A 不同于 B”，就创造区分承诺；说“原因在于”，就创造因果承诺。说明文的退场闭合尤其适合机械清点，因为很多义务由明示结构触发。

抒情散文与较开放写作的义务更需要谨慎。意象重复不必被还原成命题，矛盾情绪也可能被有意保留。此时编码应以作者冻结的“关系不变量”而不是明确论点为核心，例如“我对故乡既想离开又无法彻底离开”本身就是可维持的双重关系，文章不需要把它解决成单一态度。退场闭合要求的是这份复杂关系已被文本承担，而不是强迫它变简单。

由此可见，退场闭合并不把所有作文变成逻辑证明。它只提供一个跨文体的最低框架：你写出的重要动作会让读者期待后续；完成要求这些期待被处理、转化或限界。至于用论证、叙事、意象、节奏还是留白完成处理，仍由文体决定。

十八、AI 写作条件下，完成问题为什么变得更尖锐

生成式 AI 把“永远还能继续写”从理论困境变成按钮级现实。人类作者过去还会因为记忆、时间、词汇或精力不足而自然停下；模型几乎可以无限追加例子、扩展段落、提供反方、润色结尾。如果完成仍以“还有没有可加内容”为标准，AI 时代的作文原则上永远不能结束。完成必须从生成能力中独立出来。

AI 也更容易制造“形式完成”。模型熟悉五段式结构、总结句、首尾照应和升华模板，能在任何薄弱正文后快速生成一个看似完整的末段。若教师把结尾形式当完成信号，AI 会让伪完成变得更普遍。退场闭合的优势在于它不问“结尾像不像”，而问前文激活的每笔核心义务是否真的有结清链接；模型生成的新句如果只是重复中心，不会替未处理的因果债入账。

与此同时，AI 可以成为义务审计工具而不是代写工具。一个更有教育价值的提示词不是“帮我把结尾写得高级”，而是“只标出这篇草稿自己提出、但后文没有处理的核心问题；不要替我补写”。学生再决定哪些要结、哪些应限界。这样 AI 的生成能力被转为诊断能力，作者仍承担边界选择和结清工作。

在 AI 协作记录中，还应特别区分“AI 新增义务”。模型每加一个漂亮类比、名人故事、概念或反方，都会把新的前向责任带进文本。学生若无能力结清这些由工具生成的义务，文章会出现一种新型债务扩张：语言越来越成熟，闭合率反而下降。因而 AI 时代最重要的停止规则可能不是“别让 AI 写太多”，而是“任何新增材料先问它新开了几笔账”。

十九、可证伪条款与日期赌注：怎样让“退场闭合”真正承担失败风险

第一条证伪条款针对结尾必要性。在同题学生作文中，预先标出已达到 $ECR \geq 0.90$ 且无核心欠账的文本，删除其公式化总结段。如果多数文本在独立读者编码中出现新的核心未结义务，且 ECR 显著跌破阈值，则“结尾只是可能的结清工具而非完成本体”这一命题受损。若删除不改变 ECR，则支持完成性已经由前文承担。

第二条证伪条款针对义务清算的教学作用。随机班级采用常规“内容—结构—语言”修改与“义务账本”修改，两组时间相同。若义务组不能降低独立读者标出的核心等待点，或其必要修改次数没有减少，则本文关于“把继续写从内容生成改成义务清算”的教学优势应撤回。不能用学生觉得“方法很新颖”替代结果。

第三条证伪条款针对作者退出。同一篇草稿分别由作者在旁解释和作者完全离场两种条件让读者判断未结义务。若两条件下核心欠账识别几乎没有差异，则“作者在场会掩盖未完成”这一机制的强度被高估；反之，如果离场显著增加稳定出现的核心等待点，就说明作者补丁确实是伪完成来源。

第四条反向约束针对过度清算。如果使用义务账本的学生显著减少探索性材料、复杂反方和有意留白，导致作文思想复杂度下降，即使 ECR 上升也不能宣布教学成功。这说明“清债”压制了必要的开账。成熟写作需要先生成、再闭合，而不是从第一句开始避免一切新义务。How2 与 How3 必须保持次序：先允许文章长，再在适当阶段进入清算。

本文给出一个可写死日期的赌注：在 2027 年 12 月 31 日前，若至少三个独立班级、每组总样本不低于 90 篇的预注册实验显示，义务账本训练相较等

时长常规修改既不能降低独立读者标记的未结核心义务，也不能减少作者解释依赖，则本文关于“退场闭合可作为作文完成教学的独立操作变量”应当降级为描述性框架，而不再主张教学机制。

二十、理论含义：作文的“完成”是一种边界化的停止能力

退场闭合把作文完成从一个末端技巧问题改成了“停止能力”问题。不会完成的学生不一定不会写；有时恰恰因为他会持续生成，所以无法判断哪一步已经不再必要。写作教育长期重视“想得出、写得多、展开充分”，却较少训练“知道何时不再打开新分支”。而在任何复杂创造活动中，停止本身都是能力：研究要封题，设计要冻结版本，工程要交付，法律要终局，文章也需要获得退出条件。

这个观点还改变“整体”的含义。整体不是所有材料都有关联，也不是每段都围绕同一个词，而是一个有限边界内的义务网络能够自我结算。一个局部若加入后没有打开需要处理的新关系，它可能只是装饰；一个很短的句子若承担三笔核心结清，它可能是承重节点。整体因此可以被理解为“债权债务结构”而非“材料集合”。

更深一层，完成不是绝对封闭，而是边界化开放。文章可以承认未知、留下余味、允许多重解释，只要这些开放被安排在不破坏核心对象的位置。真正的未完成不是“还有未知”，而是文章一方面要求读者依据某个关系作判断，另一方面又没有给出足以维持这个关系的处理。开放是被承认的边界，欠账是被遗忘的责任。

于是，“还能写”与“还需写”终于被分开。前者属于可能空间，后者属于义务空间。可能空间可以无限，义务空间必须有限。作文完成发生在义务空间收敛而可能空间仍然开放的那一刻。这个判据比“字数够了”“结尾写了”“没话可说了”都更接近写作作为一种自主组织活动的本质。

二十一、三种最容易被误判的“完成现场”：少写、过写与假闭合

第一种现场是少写但已经完成。一个学生只写六百字，却从开头冻结了一个很小的对象：一次道歉为什么比解释更难。正文只处理“解释是在保护自己的正确，道歉是在承认对方受过伤”这一条差异，给出一个事件、一次反转和一个最终判断。教师若按“篇幅不够丰富”要求再增加童年回忆、名人名言和另一场冲突，新增材料会打开新的时间线和比较义务，反而稀释原有闭合。退场闭合提醒我们：短不是未完成的证据；只要当前核心义务已结清，继续写属于扩题而非补欠。

第二种现场是写得很多却一直未完成。学生为了证明“挫折让人成长”，连续写三位名人、两段自己的经历、一个社会新闻，每个例子都能单独说明“经历过困难”，却没有处理最关键的反例：为什么同样的挫折有时让人退缩、创伤或放弃？随着材料增加，文章实际上不断强化“挫折之后发生了变化”，但“为什么是成长而不是别的变化”这笔核心义务始终未结。此时越补同

类材料，越像在用资产规模掩盖负债结构。正确动作不是第五个例子，而是收窄主张或建立机制。

第三种现场是有漂亮结尾的假闭合。正文围绕“自由”写旅行、独处、离开竞争，最后用一句高度概括的话把自由定义成“忠于内心”。这个结尾在情绪上很完整，却新开了最大的一笔概念义务：为什么忠于内心就是自由？内心冲突与长期价值冲突时怎么办？如果正文从未建立这些区分，末段不是结清器，而是债务集中器。传统评分可能因为语言成熟而给予较高评价，但 How3 仍会判定“完成性不稳”。

还有第四种现场：故意开放但已经完成。记叙文最后停在父亲没有回答的沉默里，读者不知道父亲最终是否原谅“我”。如果全文核心对象本来就是“我第一次意识到，道歉不能迫使另一个人立刻原谅”，那么父亲是否原谅可以合法留白；相反，如果全文一直承诺要揭示“父亲究竟为什么多年不说话”，最后突然以沉默结束，就仍欠核心叙事义务。两篇都叫“开放式结尾”，完成状态却相反。

这些现场说明，教师真正需要识别的不是文章“像不像成品”，而是继续写作的必要性来自哪里。篇幅、结尾形式、情绪强度、语言成熟度都只能提供弱信号。退场闭合把判断压回一个更难伪装的问题：若现在停止，究竟有哪一笔由本文自己激活的核心义务会因为停止而无法结算？只有能指出具体欠账，“继续写”才有理论依据。

二十二、与 How1、How2 的接口：完成不是截断生长，而是停止“必要生长”

作文的三个 How 问题不能互相替代。启动阶段解决的是从无到有：如何让第一批具有方向性的文字出现；持续生成阶段解决的是从有到更多但不堆积：新材料如何改变既有结构并继续制造可承接的下一步；完成阶段则解决从“还能产生下一步”到“下一步不再必要”。如果把完成规则过早塞进启动阶段，学生会每写一句时都担心欠账，生成被审查压死；如果只教生成而没有完成规则，文章会无限扩张。

因此，成熟课堂应当允许“开账期”和“清账期”分离。开账期鼓励异常、联想、例证、反方、细节和新问题进入，不急着每笔立刻结清；清账期则冻结边界，停止无条件扩张，开始识别哪些账必须还、哪些可以删、哪些应明确移出本篇。这个节奏与创作中的发散—收敛相似，但退场闭合比一般“收敛”更具体：收敛不是少写，而是把注意力从生成新可能转向解除必要继续。

How2 与 How3 之间存在一个可观察的转折信号：新增一段文字带来的“结清义务数”开始低于“新激活义务数”。在持续生成早期，一段新材料通常既回应前文又建立新的有价值方向，净效果是结构生长；临近完成时，若每补一个例子都需要再解释一个概念、再引入一个背景、再处理一个反方，义务总量不降反升，说明文章已经接近边界，继续生成可能从生长转为蔓延。课堂不一定要把这一差额量化，但可以让学生每新增一段标记“还了几笔、又开了几笔”。

How3 也不能把“义务清零”理解成把文章磨成封闭说明书。真正的目标不是零开放，而是零未处理的必要开放。诗性余味、价值冲突、不可知的人心、未来尚未发生的结果都可以保留；关键是文本是否把这些开放的地位说清楚，使读者知道“作者有意把它留在这里”，而不是误以为自己漏读了一段。完成不是消除复杂性，而是给复杂性一个稳定的边界位置。

从这个接口看，作文能力可以重新描述为三种相继又可循环的能力：启动能力让对象出现，生长能力让对象通过新关系不断改写自己，退场能力则让作者在对象已经能够独立存在时停止占用它。一个真正成熟的作者并不是始终控制文本的人，而是在适当时刻能把文本交出去的人。How3 最终训练的，正是这种“知道什么时候应该离场”的写作判断。

二十三、结论：真正的停笔不是枯竭，而是下一笔已不再属于本篇的欠账

本文从“作文怎样完成”这一语文教育难题出发，拒绝把答案缩成结尾技巧。形式化方法与程序验证让我们看到：完成不需要穷尽全部状态，只需把与当前性质有关的验证义务判定掉；争议解决与仲裁让我们看到：终局不等于世界再无争议，而是被提交的请求在边界内获得结算；风险、安全与可靠性工程让我们看到：任何完成声明都必须经受操作者退出和环境扰动，不能用“目前没出事”替代稳定。

三家二阶碰撞后出现的新对象是“退场闭合态”。它把作文完成定义为一种作者可以合法退出的整体状态：当前文本激活的必要续写义务已结清或限界，独立读者无需作者补丁即可维持核心关系，最小扰动不会暴露新的核心欠账，继续添加内容主要会扩大边界或创造新义务。与此配套的退场闭合率 ECR，使“我觉得差不多了”第一次可以被拆成一张可复核账本。

公开学生作文的探索性真跑给出一个重要警告：ECR 并不是作文质量总分。它会在完成后迅速饱和，一篇普通甚至语言粗糙的作文也可能闭合；一篇高质量文本也可能在草稿阶段暂时不闭合。这个结果守住了本题的边界——How3 研究的是如何结束一轮写作，不是重新回答“什么是好作文”。

对课堂最直接的变化可以压成三句话：不要先问“还能加什么”，先问“还欠什么”；不要先问“结尾怎么升华”，先问“最后一段究竟结哪笔账”；不要用“作者说自己写完了”验收，先让作者退出。真正成熟的停笔不是作者无话可说，而是他仍有很多话可说，却能指出：从下一句话开始，我将不再是在完成这一篇，而是在开启另一篇。

附录 A 退场闭合清点手册（课堂/研究通用版）

【A1 冻结对象】先写版本号、题目、文体、作者希望读者最低建立的 1—3 条核心关系；再写“本篇主动不处理”的相邻问题。冻结完成前不编码 ECR。

【A2 找触发点】逐段只问一句：“这句话一写出来，读者接下来有权等待什么？”若答案会影响核心判决，登记一笔义务；纯粹想知道更多不算。

【A3 定核心级】删除该义务对应内容后，全文的中心判断、人物关系、主要冲突或对象边界是否改变？改变=核心（权重 2）；不改变但造成局部理解断裂=局部（权重 1）。

【A4 找结清链接】必须在文本内指到处理位置。作者说“我心里知道”“前面暗示过”不算；需要读者调用题型套路才能自动补上的也不算。

【A5 合法限界】若作者明确把问题排除在当前边界外，且排除不破坏核心对象，可记为结清；模糊逃避“这个问题很复杂以后再说”不自动算限界。

【A6 作者退出】编码阶段作者不得解释。至少一名不知写作计划的读者能够完成义务—结清配对；研究使用时建议双人独立编码并报告一致率。

【A7 算 ECR】 $ECR = \text{结清或合法限界的加权义务} / \text{全部已激活加权义务}$ 。试用门槛 0.90 且无未结核心义务。

【A8 做扰动】优先删公式化结尾、一个低承重例子或一段可替换描写；重算义务状态。若删结尾后旧债暴露，说明结尾在承重；若不变，结尾是风格而非闭合条件。

【A9 停笔】问“下一段将偿还哪一个已有义务 ID？”若没有已有 ID，而必须开新义务，当前版本获得停笔许可。

【A10 版本纪律】如果作者决定扩大题目、改变中心关系或主动加入新的核心分支，建立 v2 账本，不能把旧版 ECR 与新版直接混算。

附录 B 研究预注册建议表

项目	预注册内容	控制纪律
主要自变量	常规修改指导 vs 退场闭合义务账本指导	两组时长、教师接触量相同
主要结果	独立读者标记的未结核心义务数	编码者盲于组别与前测分数
次要结果	ECR；必要修改轮次；作者解释依赖	ECR 阈值在研究开始前写死
反向结果	思想复杂度、反方数量、探索性材料是否下降	防止“清债”把文章教窄
最小样本	建议总 $N \geq 90$ ，至少 3 个独立班级	以班级聚类处理标准误
失败判据	义务组不降低核心欠账且不减少解释依赖	失败结果必须原样报告

参考文献

- [1] 中华人民共和国教育部. (2022). 《义务教育语文课程标准（2022 年版）》. 北京师范大学出版社。
- [2] 付自文. (2026). 《SDE 多学科通融创新法（3.9 版）》及“新思想前沿”领域资料。SDE Universes。
- [3] SDE Universes. (2026). “形式化方法与程序验证”（新思想前沿，第 352 号）。
- [4] SDE Universes. (2026). “争议解决与仲裁”（新思想前沿，第 426 号）。
- [5] SDE Universes. (2026). “风险、安全与可靠性工程”（新思想前沿，第 600 号）。
- [6] Hoare, C. A. R. (1969). An Axiomatic Basis for Computer Programming. Communications of the ACM, 12(10), 576–580.
- [7] de Moura, L., & Bjørner, N. (2008). Z3: An Efficient SMT Solver. In TACAS 2008, LNCS 4963, 337–340.
- [8] Bradley, A. R. (2011). SAT-Based Model Checking without Unrolling. In VMCAI 2011, LNCS 6538, 70–87.
- [9] Calcagno, C., Distefano, D., O’ Hearn, P. W., & Yang, H. (2009). Compositional Shape Analysis by Means of Bi-Abduction. POPL, 289–300.
- [10] Leroy, X. (2009). Formal Verification of a Realistic Compiler. Communications of the ACM, 52(7), 107–115.
- [11] Klein, G., et al. (2009). seL4: Formal Verification of an OS Kernel. Proceedings of SOSP, 207–220.
- [12] Clarke, E. M., Grumberg, O., & Peled, D. A. (1999). Model Checking. MIT Press.
- [13] Baier, C., & Katoen, J.-P. (2008). Principles of Model Checking. MIT Press.
- [14] UNCITRAL. (2010). UNCITRAL Arbitration Rules (as revised in 2010). United Nations.
- [15] UNCITRAL. (1985/2006). Model Law on International Commercial Arbitration. United Nations.

- [16] United Nations. (1958). Convention on the Recognition and Enforcement of Foreign Arbitral Awards (New York Convention).
- [17] United Nations. (2019). United Nations Convention on International Settlement Agreements Resulting from Mediation (Singapore Convention on Mediation).
- [18] Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). Resilience Engineering: Concepts and Precepts. Ashgate.
- [19] Leveson, N. (2011). Engineering a Safer World: Systems Thinking Applied to Safety. MIT Press.
- [20] Hollnagel, E. (2014). Safety-I and Safety-II: The Past and Future of Safety Management. Ashgate.
- [21] ISO. (2022). ISO 21448:2022 Road Vehicles — Safety of the Intended Functionality.
- [22] Flower, L. (1979). Writer-Based Prose: A Cognitive Basis for Problems in Writing. *College English*, 41(1), 19–37.
- [23] Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365–387.
- [24] Sommers, N. (1980). Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication*, 31(4), 378–388.
- [25] Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Lawrence Erlbaum Associates.
- [26] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- [27] de Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to Text Linguistics*. Longman.
- [28] Kermode, F. (1967). *The Sense of an Ending: Studies in the Theory of Fiction*. Oxford University Press.
- [29] Smith, B. H. (1968). *Poetic Closure: A Study of How Poems End*. University of Chicago Press.

- [30] Aristotle. On Rhetoric: A Theory of Civic Discourse (2nd ed., G. A. Kennedy, Trans.). Oxford University Press, 2007. Book III, Chapter 19 (收束部分) .
- [31] Hewlett Foundation / Kaggle. (2012). Automated Student Assessment Prize (ASAP): Automated Essay Scoring dataset.
- [32] Turanga1. (retrieved 2026-08-17). Automated-Essay-Scoring: public mirror of ASAP training_set_rel3.tsv. GitHub.

网络资料核验入口（对应上表第 [1]—[5]、[14]、[21]、[32] 条；访问日期：2026-08-17）

<https://sdeuniverses.com/frontier/>

<https://sdeuniverses.com/frontier/formal-methods-program-verification/>

<https://sdeuniverses.com/frontier/dispute-resolution-arbitration/>

<https://sdeuniverses.com/frontier/risk-safety-reliability-engineering/>

<https://uncitral.un.org/en/texts/arbitration/contractualtexts/arbitration>

<https://www.iso.org/>

https://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html

<https://github.com/Turanga1/Automated-Essay-Scoring>

编者补节·第六章

一·需要补的三位：本章语汇的正主

本章把「作文完成」重写为一张账本：文本自己激活了哪些前向义务，其中哪些已结清或已合法限界，哪些还欠着；作者退场之后，独立读者仍可仅凭文本维持核心关系。整套语汇——义务、激活、结清、未结、账本——在论辩逻辑与对话计算里是四十年的现成术语，而本章一位未提。

（一）承诺存量

Hamblin, C. L. (1970). **Fallacies**. Methuen. 提出对话双方各有一个承诺存量（commitment store）：说出一个断言即把它计入自己的存量，此后要为它负责，除非依规则撤回。

（二）对话中的承诺

Walton, D., & Krabbe, E. C. W. (1995). **Commitment in Dialogue**. SUNY Press. 把承诺存量发展成一套对话类型的规范框架：不同的对话类型规定了承诺如何产生、如何转移、如何合法退出。

（三）话语义务

Traum, D. R., & Allen, J. F. (1994). Discourse obligations in dialogue processing. In **Proceedings of the 32nd Annual Meeting of the ACL**, 1–8. 直接把「义务」作为一种独立于信念、目标与意图的态度引入，用以解释问与答之间的连接：一个问题使对方负有回答的义务。

三位加起来，本章那句「作者写出一个中心判断，就欠下支持或限定」在上游是一句老话。

为什么本章绕开了它们（编者的判断）。本章的三家取源是形式化验证、争议解决与仲裁、风险安全与可靠性工程——它借的是法律的终局性与工程的边界稳定，而不是论辩的承诺。这不是疏忽，是取源路线决定的盲区：仲裁给了它「请求边界内获得终局」，却没有给它「承诺如何产生」。

分离线（本章可以站住，但必须写出来）。

1. 承诺者不同。那三位的承诺属于说话人——是某个主体欠下的，可以由主体撤回。本章的义务属于文本——由文本动作激活，作者不在场时仍然成立，且明令「作者说『我心里知道』不算结清」。
2. 结清的裁决者不同。那三位由对话规则与对方裁决；本章由不知道写作计划的独立读者裁决，并把「作者退场」写成编码的必要条件而非第二个指标。
3. 合法退出的机制不同。那三位的退出是撤回（retraction），要在对话中显式执行且可被对方追究；本章的退出是限界——作者在文本中明确把某问题排除在本篇边界之外，且排除不破坏核心对象。限界不是撤回：被限界的问题仍然被承认存在，只是不属于本篇。

判定形式。取一批作文，一路按承诺存量的规则编码（谁欠、欠什么、是否撤回），一路按本章的义务账本编码。若两套编码在「哪些篇目仍有未结核心项」上给出相同的名单，本章应并回承诺一支，退场闭合率只是承诺存量余额的换算；若两套名单系统不同——尤其若存在一批作文，其作者承诺已全部履行而文本仍有未结核心义务（或反之）——本章的「义务属于文本」这一条成立。

二·与第四章的接口

见第四章补节第二节：本章的工作从第四章的持续生成窗口结束处开始，交接信号是结清义务数开始低于新激活义务数。两章在「合法闭合」这一格上共用，不是重复。

三·也需要补的一位：问题之下的问题

与第四章同（见第四章补节第一节）。对本章而言，分离线更短：那一支的问题入栈即构成回答的承诺，本章的义务激活即构成结清或限界的责任——两句几乎同形。本章多出来的那一条是作者退场：那一支不要求提问者离场后问题仍然成立，本章要求义务在作者不在场时由独立读者读得出来。

四·编者提示：本章最重要的一次自我收窄

本章的探索性编码在十二篇公开学生作文上跑出两个方向。第一个方向符合直觉：退场闭合率与人工总分总体同向。第二个方向不符合，也更重要：闭合率在中高分段迅速触顶——人工评分较高的六篇全部达到 1.00，同时一篇人工总分仅中等的作文也达到 0.92 且无未结核心义务。

作者由此撤回了任何「闭合率越高作文越好」的解释，只保留「用于判断是否仍存在必要续写义务」这一较窄命题，并写了一句编者认为应当被反复引用的话：若这个读数能够精准预测总分，反而说明它可能只是在换名重做综合作文质量。

编者提醒读者注意这一撤回的连带后果：本章给出的阈值 0.90，不是一条质量线，是一条状态线。跨过它，只意味着这一轮的账已经结清，不意味着这一篇写得好。本卷第七章正是从这条线的另一侧接着往下问的（见对切二）。

第七章

如何写好作文

—— 结算余损与追余写作

如何写好作文？

——从“纠错”到“追余”：论结算余损与三焦点轮转 写作法

How to Write a Good Essay?

*From Error Correction to Residual Pursuit: Settlement Residuals and
Rotating-Focus Composition*

SDE 学科碰撞创新法 v3.9 · How 型三焦点路径实验稿

形式化验证 × 化学反应工程 × 机制设计

版本：2026-08-17

摘要

“如何写好作文”常被回答为一条固定流水线：审题、构思、起草、修改、润色；更精细的教学则把质量拆为内容、组织、语言、证据、受众等多项标准。本文指出，两种方案仍共享一个方法论前提：改进可以由某个稳定终点结算——只要成品更符合标准、生成路径更高效，或任务环境更合理，作文就沿同一方向趋近于“好”。本文以形式化验证、化学反应工程和机制设计为三条互相冲突的路径来源。形式化验证强调把性质写成可检验要求并证明实现满足要求；反应工程强调控制步骤、产品决定步骤、选择性与转化率之间可能并不重合；机制设计强调规则、私有信息、激励相容与多个均衡对结果的塑造。三者共同逼出一个被普通“纠错—再修改”模型遗漏的对象：某一轮已经按自身标准宣布成功，却在该标准没有记账的质量域中仍存在、或因这次成功而新产生的损失。本文将其称为“结算余损”。据此提出“追余写作法”：每轮只设一个可判定终点；一旦通过，立即检查未被本轮标准计入的另外两类记录；由最大且具有因果依据的结算余损夺取下一轮修改终点。写好作文因此不是把同一套标准反复拉高，而是让每一次局部成功暴露并转交它无法结算的质量损失，直到余损不再跨域迁移。本文以六套现行官方写作评价框架做公开字段审计，发现多维评分早已普遍存在，因此“评价太单一”的批判不成立；但这些框架仍主要以静态并列标准结算，未设置“成功触发换靶”的动态规则。本文进一步给出近邻划界、课堂操作、AI 场景、机制证伪、反噬预言及经验赌注。

关键词：作文；结算余损；追余写作；形式化验证；反应工程；机制设计；生成式人工智能

Abstract

The question “How can one write a good essay?” is often answered by a fixed sequence—understand the task, plan, draft, revise, polish—or by a multidimensional rubric covering content, organization, language, evidence, and audience. Both approaches usually assume that improvement can be settled at a stable endpoint. This article collides three mature traditions with different terminal foci: formal verification, chemical reaction engineering, and mechanism design. Formal verification asks whether specified properties are objectively

satisfied; reaction engineering shows that rate-controlling steps, product-determining steps, conversion, and selectivity need not coincide; mechanism design shows that rules and incentives may support multiple equilibria and place hard limits on attainable outcomes. Their conjunction reveals an object not adequately represented by ordinary error-correction models: a loss that remains outside the ledger of a round that has already been declared successful, or is newly displaced there by that very success. The article calls this a settlement residual. It proposes residual-pursuit composition: each round adopts one explicit terminal criterion; once it passes, the writer audits the other two domains, and the largest causally grounded residual determines the next terminal focus. Good writing is therefore treated not as repeated optimization against a fixed standard, but as a sequence in which each local success is forced to disclose what it failed to settle.

Keywords: composition; settlement residual; revision; formal verification; reaction engineering; mechanism design; generative AI

一、为什么“多改几遍”仍然没有回答 How

“如何写好作文”看起来比“作文是什么”更容易回答。教师可以给出一条熟悉的流程：审题，构思，列提纲，起草，修改，润色；也可以给出一张更细的量规：立意、证据、结构、语言、受众、规范分别提高。问题在于，这些方案描述了很多动作，却没有回答一个更严格的 How 问题：下一步究竟为什么是这一步，而不是另一种修改？一个学生现在最该改语言，还是证据？最该重写段落，还是换掉任务结构？最该继续打磨成品，还是停止打磨，先改变自己写这篇文章的方式？

现代写作研究当然早已超越“写完再改错”。1970 年代的过程写作、问题解决与 **writing-to-learn** 传统已经把构思、目标、计划、反馈和知识变化纳入写作本身；今天再说“作文要迭代”并不新。真正困难的是：当一次修改成功以后，凭什么继续沿着同一终点优化？

本文的核心判断由此产生：写好作文最危险的时刻，不一定是失败的时候，而可能恰恰是某一项已经做得很好、系统因此允许我们宣布“这一轮完成”的时候。因为宣布完成的那把尺子，只看见它被设计去看见的东西。

二、第一条路径：先把“好”写成可以失败的要求

形式化验证给出的路线，是所有写作建议中最容易被误解为“做量规”的一条。其真正力量不在于列清单，而在于要求一个系统的性质被表达得足够明确，以至于实现可以被客观证据确认或否认。NIST 将 **verification** 定义为以客观证据确认指定要求已经满足；与此同时，又把 **validation** 与特定 **intended use** 联系起来。NIST 的 **requirements verification** 页面还把完整、一致、正确、可修改、可排序、可追踪、无歧义、可理解和可验证列为需求质量需要检查的属性（NIST, 2021/2026；NIST SP 800-160, 2022）。

如果把这一纪律真正用于作文，第一步不是“写得更漂亮”，而是把这一篇文章必须成立的关系写成可失败命题。例如：中心主张必须能指出至少一组直接承担它的证据；删掉关键限定语以后结论会发生可识别的越界；每一个主体段必须承担一个不能被其他段落完全替代的论证功能；反例不能只被提及而不改变主张边界。

这条路径可写成：先建立要求与使用场 E，再把要求转化为检查、反例和对抗性测试 D，最后得到能够通过检验的文本 S。它的焦点是成品显露态。

但形式化验证自己同时留下第一道裂缝：实现符合 **specification**，并不自动证明 **specification** 表达了真正需要的 **intended use**。验证成功与有效性确认因此不是一回事。换到作文里，一篇文章可以极其准确地满足量规，却仍然没有回答真正的问题；甚至量规本身就on能遗漏最重要的质量。

三、第二条路径：不要平均修改，先找谁在控制结果

化学反应工程把 **How** 的焦点从最终成品移到实际生成路线。IUPAC 对 **reaction mechanism** 的定义要求描述从反应物到产物的过程、关键中间体和过渡状态；对 **rate-controlling step** 的定义则强调某一步的速率常数对总体反应速率具有最强控制。更关键的是，IUPAC 还单独定义 **product-determining step**：决定产品分布的步骤可能与 **rate-controlling step** 相同，也可能出现在更晚位置。

这意味着“最慢的一步”“最影响总速度的一步”和“最决定最后成为什么的一步”并不必然是同一步。MIT 的反应工程课程进一步把 **selectivity versus conversion**、反应器串并联、停留时间分布以及不同反应器如何改变选择性—转化率关系放在同一教学链条上。

作文中的对应问题不是“哪一步最差”，而是“哪一步现在拥有最大控制”。有时材料读取已经足够，真正控制全文的是把证据转换为理由的那一步；有时观点和证据都好，过早把第一版措辞固定下来才是控制步骤；有时语言表面是最明显的错误，但修语言几乎不改变总体质量，因为真正的产品决定步骤发生在论点边界。

这条路径从成品症状 S 出发，改变材料供给、时间、提示、工具、顺序等条件 E，最后重组实际生成路线 D。它的焦点是路径本身。

这条路线同样自带反噬：提高 **conversion** 不保证提高 **selectivity**；加快一个控制步骤之后，真正决定产品分布的步骤可能仍在别处。换句话说，“改进成功”本身可以把损失从一个指标挤到另一个指标。

四、第三条路径：如果规则奖励表面成功，人就会学会表面成功

机制设计把终点放在第三个地方。它不先问参与者“应该怎么做”，而是问：在现有规则、信息和激励下，理性参与者实际上会怎么做；产生的均衡是否符合目标；如果不符合，应当怎样重新设计规则。2007 年诺贝尔奖材料把这一分析概括为预测规则下的行为、按目标评价结果、再寻找兼顾行为后果的机制，并强调 incentive compatibility、implementation 与 private information。

放到作文教育，很多“学生为什么不深度思考”的问题会改写成机制问题。如果得分高度可预测地来自固定段式，学习固定段式就是理性策略；如果“引用数量”比证据功能更容易计分，引用数量会被优化；如果教师口头鼓励独立判断，却通过范文、模板和评分反馈不断奖励某一套可识别形态，那么学生并不需要被指责为“缺少创造性”，因为制度本身把模仿变成了高收益策略。

机制设计的完整路线是从目标结果 S 出发，预测真实行动路径 D，最后重构任务、信息、反馈和评价场 E。它的焦点是条件场。

然而机制设计自己的成熟理论也拒绝轻易乐观。一个机制可以同时存在理想均衡和劣质均衡；在某些信息与参与约束下，效率、真实报告和自愿参与不能同时完全实现。于是“把规则设计好”也不是终局。

五、三条路真正冲突的不是方法，而是“在哪里宣布完成”

形式化验证会说：先明确要求，成品通过可证伪检验才算完成。反应工程会说：终点检查太晚，真正需要优化的是决定结果的生成路线。机制设计又会说：连路线都不是最终层；只要规则允许伪装性最优，写作者就会沿激励结构适应。

三个答案表面上互相否定，底下却共享一个前提：一轮改进总可以由某个稳定终点结算。成品、路径、环境三者中，总有一个可以成为本轮“最后一站”，另外两个则退为手段、边界或条件。

问题恰恰在“结算”二字。形式化验证已经承认符合 *specification* 不等于 *intended use* 已经被满足；反应工程明确告诉我们 *rate-controlling step* 与 *product-determining step* 可以不同，*conversion* 与 *selectivity* 还可能发生权衡；机制设计则表明理想均衡可以与劣质均衡共存。三家都在自己的领域里知道同一件事：本轮最醒目的成功，可以把未解决的损失留在另一张账上。

六、一个过去没有被作文修改充分命名的对象：结算余损

本文把这一类对象称为“结算余损”。

结算余损，是指某一轮改进已经按该轮明确的通过条件被判定成功，但在未被该通过条件计入的质量域中仍然存在，或者由于该轮成功而新产生、转移或放大的损失。

它与普通“还有错误”不同。普通错误意味着当前标准已经能够看见问题，只是尚未修掉。结算余损恰恰出现在标准宣布成功之后：当前标准没有失败，因此它没有理由继续修改自己；损失只能从另一张账里出现。

它也与“副作用”不完全相同。副作用可以只是偶然伴随；结算余损要求能够说明，它为什么与本轮成功的优化方向存在因果关系或结构关系。把段落压缩得更紧以后，读者理解更快，却可能把必要限定压没；把任务量规写得极清楚以后，学生表现更稳定，却可能开始围绕可见条目进行策略性优化；把构思流程加速以后，完成率提高，却可能让产品决定步骤更早被第一版措辞锁住。

因此，写好作文不应只记录“这一轮修好了什么”，还必须记录“这一轮的成功把什么挤出了账本”。

七、三类结算余损

第一类是规外余损。文本已经满足公开标准，但公开标准没有覆盖真正的使用目的。例如一篇议论文拥有明确论点、充足引用、漂亮结构，却把一个根本不成立的前提当成默认；或者它完全符合考试量规，却无法让真实读者知道作者为什么相信这个判断。形式化通过制造了“完成感”，剩余损失藏在 *specification* 之外。

第二类是路外余损。一次路径优化明显提高了速度、完成度或某一局部产出，却把决定最终质量的控制转移到别处。学生通过模板迅速建立结构以后，下一控制点可能变成证据解释；AI 帮助迅速获得多个方案以后，下一控制点可能变成判断和舍弃；反复润色语言以后，下一控制点可能变成论点边界。

第三类是局外余损。任务规则成功地产生了可预测行为，但这个行为不是规则制定者真正想要的。学生会写出满足量规的“高分样子”，却把精力放在最便宜的可见信号；教师要求列三个反例，学生就学会机械地造三个反例；制度要求展示过程，过程痕迹本身又可能变成被表演的对象。

三类余损的共同点是：它们不是在失败处排队，而是在成功处逃逸。

八、追余写作法：每轮只设一个终点，成功以后立即换账

由此产生的写作方法不是一条固定“审题—构思—起草—修改”流程，而是一条由余损决定下一站的路线。本文称为“追余写作法”。它有五个动作。

第一，写死本轮唯一终点。一次只选择成品、路径、环境中的一个作为本轮结算焦点，并把通过条件写成可以失败的句子。比如本轮只检查“每个核心证据都改变了主张的可信程度”，而不同时追求华丽语言。

第二，只做能触及当前控制点的修改。不是平均修十处，而是明确哪一次改动在因果上最可能改变本轮终点。若改语言不会改变证据功能，就暂时不改语言。

第三，成功即开余损账。只要本轮通过条件满足，立刻停止继续优化当前焦点，转而查看另外两个焦点的记录：有没有新出现的损失，或者以前被当前失败遮住、现在才显露出来的损失。

第四，让余损夺靶。下一轮修改终点不按预设顺序，而由最大、最有因果依据、且不能被当前终点继续解释的余损决定。如果文章更清楚以后暴露出论证其实空洞，下一轮转向路径；如果路径变好以后发现学生在量规面前开始机械最优化，下一轮转向任务环境。

第五，连续两轮不迁移才允许停。一次“没有新余损”可能只是检测盲区。只有连续两轮的主要余损都没有跨焦点迁移，或者任务预算已经到

达明确边界，才把文本交付。这里的停止不是“完美”，而是当前可承担的残余风险已经稳定。

九、这不是固定的三步法：焦点顺序必须由余损决定

如果把追余写作误解为“先成品、再过程、再环境”三步法，新的方法立刻退化成另一套模板。三焦点只是可能的结算位置，不是固定顺序。

例如一名学生面对考试作文，第一轮最合理的焦点可能是 S：先确认是否真正回答题目。通过以后发现语言和结构都不错，但每个证据只是例子，没有发生推理，于是余损落到 D，第二轮改生成路径。第二轮之后学生理解变深，却发现评分量规长期奖励固定结构，使第三轮最好转向 E：调整练习规则，让“先产生两个互相冲突的解释再成文”成为练习条件。

另一名学生可能从 E 开始。AI 辅助写作规则混乱，使他所有注意力都耗在“哪些能用、哪些不能用”，先重设工具边界和披露规则；环境稳定以后，余损才显出真正的路径控制点。

因此最重要的不是三家给出三种技巧，而是把“下一步是谁”从教师经验判断改造成可追踪的余损迁移。

十、一张可实际使用的“余损台账”

追余写作不要求增加复杂监控。最小台账只记录六个字段：本轮焦点、本轮通过条件、关键修改、通过证据、另外两类记录中新出现或仍未解释的损失、下一轮焦点。

它故意不记录“总分提高多少”作为唯一主线，因为总分会把不同来源的损失重新压成一个数。真正有信息的是焦点怎样换手。

例如：第 1 轮焦点 S，标准是“每一主体段都能写成证据—理由—主张链”；通过后发现三条链全部来自老师给定模板，余损落在 E；第 2 轮焦点 E，撤掉模板并要求自定证据路径；通过后发现学生耗时暴涨且观点仍反复改写，余损落在 D；第 3 轮焦点 D，固定材料和受众，只比较两种生成顺序；通过后文章清楚度下降，余损又回到 S。这个 S→E→D→S 序列本身比“一共改了三稿”更能解释写作如何变好。

十一、五个具体场景：同样是“改作文”，下一步可以完全不同

场景一：高分应试文。文章已经完整满足公开量规，语言稳定，结构清楚，但读者换成真实受众以后，核心论据的说服力迅速消失。这不是继续加分项的问题，而是规外余损：公开 **specification** 没覆盖真实 **intended use**。下一轮不能继续润色，而要重开证据与受众关系。

场景二：论点混乱的初稿。教师发现学生并不是“不会结构”，而是在证据读完以后才临时决定主张，导致每增加一份材料就重写中心。此时路径控制点明显，先改 **D**；一旦主张稳定，再检查它是否造成新的选择性损失，例如为了稳定而排除了真正重要的反例。

场景三：**AI** 辅助写作。学校已经把可用工具、引用和披露写得很清楚，环境焦点通过；但学生的实际生成路径变成“先问 **AI** 最佳结构—再补个人措辞”。下一轮余损不在合规，而在 **D**：要测试没有输出模板时能否自行组织证据。

场景四：语言很弱但思想完整。若教师因为表面语法错误把所有精力投入 **S**，可能长期看不到文章已经具有稳定的论证结构。先使用语言辅助把噪声降到可读水平，成功以后再追余：辅助是否开始替代本应由学生完成的判断。

场景五：范文教学。范文使全班结构迅速改善，**S** 显著成功；成功后若所有文章开始拥有相同的论证路径，余损落到 **E** 或 **D**。下一轮不是换一篇更好的范文，而是改变规则：把范文变成相互冲突的多个方案，要求学生说明拒绝哪一种结构及原因。

十二、对 **AI** 时代最重要的改变：不要问“**AI** 用了多少”，先问余损移到哪里

生成式 **AI** 最大的教育风险之一，是它把“成功”变得非常便宜。清楚的结构、顺滑的语言、充足的例子、格式完整的引用，都可能迅速出现。若评价系统仍把这些显露特征当作主要结算终点，学生很容易在极短时间内通过 **S**。

但通过 S 不是本文意义上的坏事。它只是意味着现在必须开余损账。AI 让语言更清楚以后，真正的损失可能迁移到 D：哪些主张—证据关系仍由学生掌握？也可能迁移到 E：作业规则是否奖励“看起来像深度思考”的廉价代理？

因此合理的 AI 写作教学不应围绕一个永久的“人类比例阈值”，而应围绕焦点迁移。机器完成某一类显露工作以后，教师必须主动把下一轮目标移到机器最容易隐藏的余损处。反过来，如果学生在判断、证据组织和受众推理上已经形成稳定路径，让工具修正局部语言不应被自动解释为作文质量下降。

十三、敌意检索之后，这个方法与哪些成熟理论仍然不同

第一个危险近邻是过程写作。Murray 1972 年的“process not product”、Emig 1977 年的 writing-to-learn、Flower 与 Hayes 1977 年的问题解决模型，都已经证明写作不是一次性交付。本文不以“过程重要”作为新意。判决线是：普通过程模型允许多轮围绕同一组目标继续改进；追余写作规定，本轮一旦成功，当前终点必须暂时失去优先权，下一轮由成功以后在别处显露的损失夺靶。

第二个危险近邻是 Theory of Constraints。其五个 focusing steps 已经成熟地要求识别约束、利用、从属、提升，再在约束改变后回到第一步。若本文只是“找作文瓶颈”，应直接被 TOC 吸收。二者的分离线是：TOC 通常在固定系统目标下寻找限制 throughput 的约束；结算余损可以在当前约束已经被解除、总体 throughput 上升时出现，并且迫使我们改变下一轮究竟把什么当终点。conversion 上升而 selectivity 下降，就是典型的“成功但目标域换了”。

第三个近邻是 Goodhart/Campbell 型指标腐化。它们已经指出量化指标一旦承受决策压力会被操纵或失真。本文若只说“公开量规会被学生钻空子”，也没有新意。分离线是：结算余损即使没有任何策略性行为也会出现。例如学生完全诚实努力地提高篇章清楚度，仍可能因为压缩而丢掉限定；反应路径优化发生的选择性损失不需要一个会作弊的主体。

第四个近邻是 **double-loop learning**。双环学习会在结果不合时反思 **governing variables**，而追余写作最锋利的动作发生在结果已经合格时：成功本身触发对未结损失的搜索。

第五个近邻是 **multi-objective optimization**。它通常要求预先声明多个目标并寻找权衡或 **Pareto 解**；追余写作不假设教师能够事前知道所有质量目标。它利用局部成功作为探针，让没有进入原目标集的损失在下一轮获得对象资格。

第六个近邻是 **rubric-based formative assessment** 和 **assessment ecology**。现代写作评价早已可以是多维、可适应且关注环境。本文不反对多维量规；它只增加一个时序规则：维度不能永远平行摆放，某一维成功以后必须观察损失有没有迁到另一维。

第七、第八个近邻来自同一产线的前两篇。“预闭分叉”回答哪些可能性在学生遇到以前就被外部拿走；“隐作”回答作文状态是否可能先于文本显形。本文则回答怎样连续改进：即使没有预闭、也已经形成作定态，仍会出现一轮成功制造下一轮盲区的结算余损。三个概念互相可以独立成立。

十四、一次公开字段审计：多维评分已经存在，动态换靶尚未成为评分常规

为避免把“今天的评价只看一个分数”当作稻草人，本文在成文前预先设定了一次最便宜的公开字段审计。对象包括 **AP English Language**、**IB Language A**、**IELTS Writing**、**TOEFL Writing**、**Cambridge IGCSE First Language English** 以及新西兰 **e-asTTle/SMART** 写作评价公开材料。编码只问四件事：是否有明确的成品质量维度；是否把生成过程本身作为评分字段；是否把任务规则／激励结构作为学生作文评分字段；是否存在“某一维通过以后，由另一维余损自动取得下一轮优先权”的动态触发规则。

结果首先推翻了本文最初一个更简单的批判。多维评价早已非常普遍：**IB** 把理解解释、分析评价、焦点组织与语言分开；**IELTS** 使用 **task response**、**coherence and cohesion**、**lexical resource**、**grammatical range and accuracy**；**TOEFL** 明确关注观

点发展、组织、语言准确与任务目的；新西兰 e-asTTle 甚至把 Ideas、Structure and Language、Organisation、Vocabulary、Sentence Structure、Punctuation、Spelling 分项独立评分。不能再说主流作文评价“只有一把尺”。

但是在这批公开框架中，评分维度主要以静态并列方式存在。它们告诉评分者每一维怎么看，却没有把“某一维已经成功”设计成一种诊断事件：成功后强制检查未计入该维的质量损失，并由损失的位置决定下一轮修改焦点。这个结果只支持一个很窄的第三层判断——本文提出的是动态修改规则，而不是又一张多维量规；它不证明全球不存在任何相似课堂实践。

十五、一个零情态词的操作判据

为了让“结算余损”不退化成教师的感觉，可以把判据压成一句不含“真正、充分、合理、有意义”等评价词的问题：

“本轮通过条件全部满足后，未列入本轮通过条件的另外两类记录中新增了什么问题；下一轮首先修改的是哪一类？”

这里的三类记录分别是：成品记录——文本、读者反应、证据链、语言和结构；路径记录——草稿、停顿、修改序列、工具调用、素材组织；环境记录——题目、量规、范文、AI 规则、反馈方式和时间配置。

如果答案始终是“没有新增问题”，或者新增问题从不影响下一轮焦点，追余写作就没有比普通多轮修改多提供信息。

十六、如何研究它：不要先造总分，先看焦点迁移序列

本产线前两篇已经大量使用二维格和比率指标，这一篇刻意换掉结构。本文不首先发明一个“追余指数”，而是把读数写成序列。

一篇作文的研究记录可以长成： $S \rightarrow D \rightarrow S \rightarrow E \rightarrow D$ 。每一个箭头必须对应一次“当前终点通过—别处余损出现—下一轮改靶”。如果箭头只是教师随意换任务，就不算。

最早可用的经验量不是总分，而是“第一次余损迁移发生在哪一轮”“迁移以后被转移到哪一焦点”“下一轮是否确实降低了那笔余

损”。这种记录能够直接比较固定量规反复修改、TOC 式瓶颈修复和追余写作之间的过程差异。

十七、六条机制证伪：什么结果会迫使本文退场

第一，在控制写作时间、既有能力和反馈次数以后，追余组如果在盲评质量、无支架重写或后续任务上长期不优于固定量规迭代组，那么“成功后换靶”没有实质价值。

第二，如果所谓结算余损不能在本轮通过之前后显示差异，即它既不是被成功新制造、也不是因成功才显露，只是原来就存在的普通错误，那么新概念应并回一般错误诊断。

第三，如果 TOC 五步法对同样写作任务产生与追余法不可区分的焦点序列和结果，那么本文的“余损夺靶”不需要另立理论。

第四，如果只在学生知道评分标准、存在策略性适应时才出现结算余损，而在无策略性主体或匿名修改实验中完全消失，那么 Goodhart/Campbell 型解释足够，本文应缩减。

第五，如果多目标优化在事前列出内容、结构、语言、受众、证据等目标以后，能够完整预测每一次所谓“余损迁移”，且无需任何事后新增目标，那么本文“成功使新损失获得对象资格”的部分不成立。

第六，最便宜的同校交叉实验可以把同一批学生随机分到两种修订方式：A 组每轮按同一量规把所有维度同时提高；B 组每轮只设一个焦点，通过后审计另两域并由余损决定下一焦点。控制总时长与教师反馈量。若 B 组的迁移序列并不比 A 组更能预测最后的独立盲评、无量规重写与换受众重写，本文核心机制失败。

十八、反噬：一旦“追余”成为新量规，它会立刻背叛自己

这套方法最危险的制度化方式，是学校要求每篇作文必须填写“本轮结算余损三项”，然后把余损数量、焦点迁移次数或台账完整度计分。学生会很快学会制造漂亮的余损记录；AI 更容易为每一轮自动生成一套“我发现了新的问题”。

一旦发生这种事，“追余”会从诊断工具变成新的显露终点，新的局外余损随即产生：学生优化的是余损报告，而不是作文。这个反噬没有一个可以靠再加一条量规彻底解决的出口，因为任何公开指标都可能再次成为策略对象。

因此本文坚持三条伦理边界：余损记录首先用于评价任务、反馈和工具配置的位置，而不是给学生贴能力标签；不能为了获得更漂亮的迁移序列故意制造高风险、羞辱性或不必要的困难；若研究中真的把迁移记录用于影响学生成绩，必须公开编码规则，并提供人工复核与申诉路径。

十九、对教师最实用的改变：把“你还哪里不好”改成“这次变好把什么藏起来了”

传统反馈很容易从错误出发：哪里不清楚，哪里证据不足，哪里句子有问题。追余写作增加的并不是另一套错误清单，而是一种反馈时机。

当学生刚刚成功时，教师反而最应该提问。例如：“现在结构已经很清楚了，删掉这个结构模板，你的证据为什么还是按这个顺序？”“语言已经很顺了，哪一句限定因为压缩而消失？”“你已经完全满足量规了，如果没有这张量规，你仍然会做哪两个判断？”“AI 使用已经合规了，关闭这段对话以后哪三条关系还能自己重建？”

这种反馈不把成功当作终点，而把成功当作探针。它利用成功造成的对比，把原来混在一起的质量损失分离出来。

二十、重新回答：如何写好作文

到这里，可以给“如何写好作文”一个不同于固定流水线的回答。

写好作文，不是把同一套标准从低分改到高分；不是先把所有目标列全，然后一次性同时优化；也不是永远寻找一个最慢瓶颈。

写好作文，是每轮只让一个终点承担结算责任；当它真正通过以后，立刻拒绝把“通过”解释为完成，而去寻找这次成功没有结算、甚至被这次成功挤到别处的质量损失；然后让那笔余损夺取下一轮的终点。

因此最重要的修改问题不是“下一稿还要改什么”，而是：

本轮已经写好的东西，把哪一笔质量损失赶出了自己的账本？

如果这个问题能被持续记录，作文修改就不再只是围绕错误做清扫，而变成一次次重新决定“什么有资格成为下一轮终点”的活动。好的作文并非没有余损，而是在可用时间内，最重要的余损已经无法再通过一次局部成功悄悄迁到另一张账上。

这就是本文最后的判断：作文质量不是沿一条固定道路不断接近终点，而是在成品、生成路径与任务环境之间追踪失败如何被成功重新分配。写作者真正成熟的标志，不只是越来越会把一件事做好，而是越来越能在“已经做好”的瞬间，看见下一笔尚未结算的东西。

附表 1：追余写作最小台账

轮次	当前焦点	通过条件	关键干预	本轮结算	结算余损	下一焦点
	S 成品	每个核心证据都改变主张可信度	重写证据—理由链	通过	所有链都来自固定模板	E
	E 环境	撤掉唯一模板仍能产生两种结构方案	改为冲突范文+自选方案	通过	重构耗时暴涨，论点反复移动	D
	D 路径	固定材料下两次构思能稳定形成同一主张边界	改变构思顺序并延迟措辞	通过	文本变得难读、转折失去提示	S
	S 成品	读者能在无教师解释下复述论证边界	增加必要路标并恢复限定	通过	未见跨域主导余损	—
	S 复核	另一读者仍得到相同边界	盲读复核	通过	未见跨域主导余损	停止

附表 2：公开评价框架字段审计（2026-08-17）

体系	成品多维标准	过程作为评分字段	任务／激励场作为学生评分字段	成功触发换靶规则
English Language	是：论证／证据／修辞分析等成品表现	否（考试评分）	否	未见
Language A	是：理解解释、分析评价、焦点组织、语言	否（考试评分）	否	未见
S Writing	是：任务回应、连贯衔接、词汇、语法	否	否	未见
FL Writing	是：观点发展、组织、语言与目的	否	否	未见
bridge IGCSE FLE	是：清楚准确表达、风格、	公开页面未列为评分主轴	否	未见

体系	成品多维标准	过程作为评分字段	任务／激励场作为学生评分 字段	成功触发换靶规则
	受众等			
-asTTle / SMART	是：Ideas、Structure & Language、Organisation 等七项	否	否	未见

说明：本表是公开字段审计，不是全球文献穷尽检索。它只能支持“这些主流公开框架没有把成功触发换靶做成明确评分规则”这一窄结论；不能据此证明不存在教师个体实践、研究项目或未公开内部流程。

经验赌注：2028-08-17

本文留下一个写死日期的经验赌注：截至 2028 年 8 月 17 日，如果“结算余损”确实具有独立机制价值，那么至少应出现一项可复核的写作修订研究，在控制总写作时间、反馈量与既有能力后，发现“当前焦点通过以后发生的跨域余损迁移”能够预测下一轮最有效的修改位置，并且这种预测优于只依赖当前量规最低分、教师主观瓶颈判断或固定多维量规的方案。

以下情况不算命中：仅证明多稿写作有效；仅证明多维量规有效；仅证明学生会 对评分规则进行策略适应；仅证明某个写作瓶颈被修复后另一个瓶颈出现；仅报告相关而不记录“本轮先通过—随后别域余损出现—下一轮换靶”这一时序。若到该日期仍没有任何数据表明“成功之后的跨域余损”比普通错误、TOC 约束或 Goodhart 型指标失真多提供独立信息，本文应把“结算余损”降级为已有概念的写作应用。

参考文献与公开来源

说明：第 2、3、5-8、13-19 条为机构或标准文献的公开页面。各考试机构与标准组织的具体页面路径随年度更新而变，故按栏目层级著录并标注检索日期；条目内容以检索日当天的公开页面为准。

1. NIST (2022). 《工程化可信安全系统》（NIST SP 800-160v1r1）. 美国国家标准与技术研究院. <https://doi.org/10.6028/NIST.SP.800-160v1r1>
2. NIST 计算机安全资源中心. 术语表：verification; validation（定义取自 ISO 9000:2015）. <https://csrc.nist.gov/glossary>（检索日期：2026-08-17）

3. NIST 计算机安全资源中心. 需求验证相关工具与属性清单（完整、一致、正确、可修改、可排序、可追踪、无歧义、可理解、可验证）. <https://csrc.nist.gov/>（检索日期：2026-08-17）
4. IUPAC. “rate-controlling step”. 《化学术语汇编》（金皮书）第 5 版，在线版 5.0.0（2025）；原始来源 PAC 1994, p.1156. <https://doi.org/10.1351/goldbook.R05139>
5. IUPAC. “product-determining step”. 《化学术语汇编》（金皮书）第 5 版，在线版 5.0.0（2025）；原始来源 PAC 1994, p.1153. <https://goldbook.iupac.org/>（检索日期：2026-08-17）
6. IUPAC. “selectivity”. 《化学术语汇编》（金皮书）第 5 版，在线版 5.0.0（2025）；原始来源 PAC 1994/1996. <https://goldbook.iupac.org/>（检索日期：2026-08-17）
7. MIT OpenCourseWare (2007). 10.37 Chemical and Biological Reaction Engineering, Spring 2007. 麻省理工学院. <https://ocw.mit.edu/>（检索日期：2026-08-17）
8. 瑞典皇家科学院／NobelPrize.org (2007). Mechanism Design Theory: 科普说明与颁奖辞. <https://www.nobelprize.org/prizes/economic-sciences/2007/>（检索日期：2026-08-17）
9. Murray, D. M. (1972). Teach writing as a process not product. The Leaflet, 71(3), 11-14.
10. Emig, J. (1977). Writing as a mode of learning. College Composition and Communication, 28(2), 122-128. <https://doi.org/10.58680/ccc197716382>
11. Flower, L. S., & Hayes, J. R. (1977). Problem-solving strategies and the writing process. College English, 39(4), 449-461. <https://doi.org/10.2307/375768>
12. Campbell, D. T. (1976). Focal local indicators for social program evaluation. Social Indicators Research, 3, 237-256.
13. Theory of Constraints Institute. The five focusing steps（持续改善流程 POOGI）. <https://www.tocinstitute.org/>（检索日期：2026-08-17）
14. College Board. AP English Language and Composition: 考试说明与评分量规. <https://apcentral.collegeboard.org/>（检索日期：2026-08-17）
15. International Baccalaureate. Language A: 评审说明（2026 年版）. <https://www.ibo.org/>（检索日期：2026-08-17）
16. IELTS. Writing 分项评分标准与备考资料. <https://ielts.org/>（检索日期：2026-08-17）

17. ETS. TOEFL iBT Writing Section: 评分标准与说明.
<https://www.ets.org/toefl> (检索日期: 2026-08-17)
18. Cambridge International. Cambridge IGCSE First Language English (0500): 教学大纲概览. <https://www.cambridgeinternational.org/> (检索日期: 2026-08-17)
19. 新西兰教育部. e-asTTle: 写作如何评分. <https://www.education.govt.nz/> (检索日期: 2026-08-17)

附录 A：v3.9 全工序制造审计（内部）

本附录保留制造过程，正式正文不依赖这些内部标签。它用于复核本篇是否真的按 How 型三焦点路径运行，而不是把三个学科拼成三种写作技巧。

A1. 题型、闸零与三焦点路径

来源	How 完整路径	未吸收度初判	理论厚度	厚度依据
化验证	焦点 S：E→D→S	暂判干净	厚	NIST/ISO 长期 V&V、requirements, mal methods 传统
反应工程	焦点 D：S→E→D	暂判干净	厚	IUPAC 术语体系+反应工程成熟教材与反例传
设计	焦点 E：S→D→E	刚进入教育设计，作文理论暂判干净	厚	Hurwicz-Maskin-Myerson；激励相容、理论、不可能性结果

Gate Zero 检索的反向结果：writing studies 已大量吸收过程、问题解决、反馈、评价量规、游戏化与一般设计思想，因此这些不能承担本篇的新源；机制设计在教育领域已出现邻近应用，故不判“绝对干净”，而判“教育层刚进入、作文核心回答尚未形成固定传统”。

A2. 三家 1+3+1 抽脊

来源	主题判断	支撑 1	支撑 2	支撑 3	自己知道但不往作向用的事实
化验证	指定性质应被明确表达，并用客观证据确认实现是否满足。	要求应完整、一致、正确、可追踪、无歧义、可验证	verification 与 validation 分开	测试／证明必须针对明确性质	实现符合 specifica 不保证 specificati 等于 intended use
工程	最终结果由具体反应机制、控制步骤与反应器条件共同塑造。	rate-controlling step 对总速率控制最强	product-determining step 可与 rate-controlling step 不同	reactor choice 改变 selectivity vs conversion	conversion 的成功随 selectivity 损失制位置也会转移
设计	要实现目标，规则必须考虑参与者的私有信息和策略激励。	先预测规则下行为	按目标评价均衡结果	寻找 incentive-compatible / implementable mechanism	理想均衡可与劣质共存，且存在不可前边界

A3. 27 对碰撞结果

碰撞	焦点	展开	涌现物
R1	规格通过 vs 控制步骤	达到规格的局部成功可能没有触及实际控制点	伪完成

碰撞	焦点	展开	涌现物
R2	规格字段 vs 产品决定步	关键产品差异可能由规格未设字段的步骤决定	未设终点
R3	合规 vs 选择性权衡	一个指标通过可伴随另一质量指标恶化	结算余损
R1	验证/确认分离 vs 控制迁移	当前证明成功不代表下一条件下仍由同一步控制	条件换控
R2	intended use vs product distribution	使用目的可能要求不同于当前通过标准的产品分布	目标错位
R3	validation vs conversion	完成度提高可能降低真正用途所需选择性	高完成低适用
R1	对抗测试 vs 速率控制	测试最容易找到显性错误，控制步骤可能没有报错	无错瓶颈
R2	性质测试 vs 产品决定	真正决定成品类别的步骤可能不触发当前测试	静默决定
R3	多测试通过 vs 多目标权衡	所有已知测试通过仍可能留下未测试损失	规外损失
M1	明示规格 vs 策略行为	公开通过条件会改变参与者行为	规则反身
M2	可验证性 vs 激励相容	容易测的性质未必是最值得激励的性质	可测偏置
M3	通过 vs 多均衡	存在好结果不等于机制只产生好结果	合规多解
M1	intended use vs 给定目标	目标本身需由制度选择，验证不能独立决定	目标外生
M2	有效用途 vs 私有信息	用户真实需求可能因激励不被如实呈现	需求失真
M3	validation vs implementation	一次成功使用不证明制度稳态可重复	一次成功陷阱
M1	对抗测试 vs 策略适应	一旦测试规则公开，行为可绕过测试	测试被学习
M2	测试数据 vs 私有信息	可观察痕迹不足以恢复真实动机	记录盲区
M3	测试集合 vs 劣质均衡	通过所有测试的行为仍可能落在制度不想要的均衡	合规劣解
M1	控制步骤 vs 行为规则	规则变化本身会重排系统控制位置	规则换控
M2	控制系数 vs 激励相容	最大技术控制点不一定是最高可干预的人类行为点	可控错位
M3	瓶颈突破 vs 多均衡	解除一个约束后系统可进入另一劣质稳定态	成功换坏
M1	产品决定 vs 目标设计	制度定义什么产品被奖励，从而改写产品分布	奖励决定产物
M2	产品分布 vs 私有信息	表面产物不能单独读出行为者真实状	产物不识因

碰撞	焦点	展开	涌现物
		态	
M3	决定步骤 vs 实施结果	存在理想产品路径不保证实际均衡选中它	可行非实现
M1	选择性权衡 vs 多目标制度	任何单目标优化都可能把损失转到另一目标	目标迁损
M2	选择性 vs 激励约束	提高某一类产出可能降低另一类被真实表达的质量	诱导副产物
M3	conversion/selectivity vs 不可能性	最优不是“全部最好”，而是约束下未被结算的损失怎样分布	可达余损

有效涌现物远高于最低门槛。最锋利的三个是“结算余损”“规则换控”“一次成功陷阱”。后续候选阶段选择“结算余损”，因为另外两项均可被成熟系统理论直接吸收。

A4. 五候选与近邻闸

候选	50 字承重句	最危险占位者	处理
轮换法	每轮换一个终点	被 Theory of Constraints／一般迭代法高度占位	淘汰
规写作	防止学生围绕公开标准最优化	被 Goodhart/Campbell、assessment gaming 占位	淘汰
写作	失败时反思目标与规则	被 double-loop learning 占位	淘汰
标作文优化	同时考虑成品、路径、环境	被 multi-objective optimization／assessment ecology 占位	淘汰
余损＋追余写作	成功后追踪未被本轮结算的跨域损失，由余损决定下一终点	与 TOC、Goodhart、double-loop、multi-objective 均有可裁决分离线	保留

A5. 敌意拓宽五方向

方向	占位者	它回答什么	判决性分离线
层	Murray 1972	过程而非成品	成功触发跨域换靶并非其承重命题
层	Emig 1977	写作作为学习方式	不处理“局部成功把损失挤到另一算域”
层	Flower & Hayes 1977	写作问题解决策略	不要求当前终点一旦通过就失去价值权
层	Campbell 1976 / Goodhart 传统	指标承压导致失真	本文余损可在无策略行为时出现

方向	占位者	它回答什么	判决性分离线
领域	多维写作量规	并列评价多质量维度	本文不是加维度，而是规定维度顺序时序夺靶
领域	写作过程／形成性反馈	反复修订和反馈	可围绕同一目标循环；本文成功必须查别域
学	Theory of Constraints 五聚焦步骤	修复约束后再找新约束	TOC 在固定全局目标内换约束；允许成功迫使“终点”本身换域
学	Double-loop learning	反思 governing variables	其触发点通常是失配/失败；本文点是本轮成功
学	Multi-objective optimization	多目标权衡	目标集合事前给定；本文允许成功暴露目标集未设字段
层	Goldratt, The Goal (1984)	持续瓶颈改进	同 TOC 分离线
层	Argyris & Schön, Organizational Learning (1978)	单环／双环学习	同 double-loop 分离线
产线	《预闭分叉》	学生遇见前被拿走的可能性	回答 What, 不回答局部成功后的下一步路径
产线	《隐作》	文本显形前已形成的作文状态	回答 What, 不规定如何连续改进

A6. 共有前提、推翻材料与最终收敛

三家争的是：写好作文最终应由成品符合性质、生成路径被控制，还是任务与激励场被重构来结算。三家共同假定：至少存在一个稳定终点，可以让另外两者退为手段或约束；只要本轮终点成功，改进就可以在该终点内结算。

推翻材料取自反应工程自己：IUPAC 明确区分 rate-controlling step 与 product-determining step，后者可以与前者不同；MIT 反应工程课程又把 selectivity versus conversion 作为需要专门处理的权衡。也就是说，一个被判定为“改善总体速度／转化”的成功，不保证产品分布也改善；本轮成功可以把质量损失留在另一终点的账外。

最终新对象：结算余损。最终新路径：追余写作。其二阶点不在“综合三家”，而在于取消“某一个终点能够单独宣布完成”的共有前提，让每一次成功成为寻找下一笔未结损失的触发器。

A7. 真跑一条与不利结果

预注册问题：如果现行主流写作评价已经具有“成功触发换靶”的动态字段，那么本文方法很可能只是已有评价理论的重述。编码对象为 AP English Language、IB Language A、IELTS、TOEFL、Cambridge IGCSE First Language English、NZ e-asTTle/SMART 的公开官方材料。

不利结果：原先草稿想批评“主流写作评价是单一终点”，但实际资料显示多维评价早已普遍存在。IB、IELTS、e-asTTle 等均把多个维度分开处理。因此本文删除“单一量规”批判，把第三层主张缩窄为：主流公开框架主要是静态并列多维标准，尚未把“一维成功→审计别域余损→由余损决定下一轮”写成明确动态规则。

A8. 命名闸与同批检查

已扫描本产线前两篇核心命名：“预闭分叉、作定态、隐作、有文未作”等未与“结算余损、追余写作”发生同名或反指。外部检索中“residual loss”在机器学习、工程和生态补偿等领域有各自技术含义，因此正式中文概念不采用“残差损失”，避免被误读为数学损失函数；“结算余损”专指“当前终点已宣布成功之后，留在该终点账外的质量损失”。

A9. 不适用边界

- 极短、一次性、无修订预算的考试现场：追余法只能用于训练与事后复盘，不能假装有无限轮次。
- 纯创意／诗性写作：若无法给出本轮可失败的终点，形式化验证路径需要弱化，不能硬套论证作文标准。
- 初学者认知负荷极高时：一次只追一个余损，不能把三张账同时压给学生。
- 高风险评价：焦点迁移记录不直接用于给人打分；否则方法自身会触发局外余损。
- 任务目标本身非法、不伦理或有害时：追余法不能替目标提供正当性。

A10. 人机分工与证据边界

本轮问题、方法框架和“从全球知识网络按 v3.9 选源”的任务由用户提出；资料检索、跨领域候选生成、敌意拓宽、公开字段审计、概念收敛与初稿由 AI 辅助完成。本文关于形式化验证、反应工程、机制设计和官方写作评价的事实性材料来自文末所列公开来源；“结算余损”“追余写作”及其用于作文教育的推论属于本文提出的理论假设，不是三家来源的原有主张。正式发表前仍需由作者人工核验所有原文、页码和完整书目，并用真实学生样本执行预注册实验。

一 · 取源的一处实情，先交代

本章的三家取源是形式化验证、化学反应工程、机制设计。其中形式化验证与第六章重复，机制设计与第三章重复——本章三家里有两家在本卷别处已经出现。

编者核过取的线，确实不同：

学科	本卷别处取的线	本章取的线
形式化验证	第六章取验证义务与可判定结果态（走到能把与性质相关的义务判定掉才算走完）	本章取验证与确认之分（实现符合规格，不等于规格表达了真实用途）
机制设计	第三章取多重均衡（规则场也会被结果改写）	本章取激励相容与可测偏置（容易测的性质未必是最值得激励的性质）

两条线在各自学科内部确实是不同的问题。但编者不打算把这一点说成没有代价：在一本以「取三门远距离学科相撞」为体例的书里，一章的三家有两家已经被本卷别处用过，它在取源独立性上就是弱于其余各章。本卷的处置是编排——把本章夹在第六章与第八章之间，让它在两侧的对照中被读；而不是单独被读。读者若要单独评估本章，应当把这一条计入。

二 · 需要补的三位

本章的核心对象是结算余损：某一轮改进已按该轮明确的通过条件被判定成功，但在未被该条件计入的质量域中仍然存在、或因这次成功而新产生、转移、放大的损失。

本章第十三节已经处理了六位近邻（过程写作、约束理论、指标腐化、双环学习、多目标优化、量规评价），处理得都相当利落。尤其与指标腐化那一条的分离线是决定性的：结算余损在完全没有策略性行为时也会出现——诚实地提高篇章清楚度仍可能压掉必要的限定；反应路径优化带来的选择性损失不需要一个会作弊的主体。这一刀切得干净。

但有三位没有请进来，而且每一位都在本章的正下方。

（一）后果效度

Messick, S. (1989). Validity. In R. L. Linn (Ed.), **Educational Measurement** (3rd ed.). American Council on Education / Macmillan.

这一支把效度从「测量是否测到了它声称要测的东西」扩展到测量被使用之后的社会后果是否可接受：一次测量即便在传统意义上完全有效，它进入使用之后所引发的教学窄化、内容偏移与价值后果，仍然属于效度论证必须承担的部分。

这几乎就是本章那句「宣布完成的那把尺子只看见它被设计去看见的东西」的教育测量版本，而且早了三十多年。本章划了指标腐化却没有划它，编

者认为这是本章最贵的一处缺席——因为本章的三类余损里，局外余损（规则成功地产生了可预测行为，但那不是规则制定者真正想要的）几乎就是后果效度所说的那一类后果。

分离线（本章可以站住）。

1. 对象层级不同。那一支评的是一套测量工具的效度，它的论证单位是「这个测验／这套量规是否可以被这样使用」。本章处理的是一轮具体改进：这一次修改成功之后，损失去了哪里。前者是工具层的长期判断，后者是过程中的逐轮事件。
2. 本章允许没有工具。后果效度的全部讨论都以存在一套被使用的测量为前提。本章的规外余损与路外余损在没有任何测量、没有任何评分者的情况下同样成立——一个人独自写作、无人评分，把段落压紧使读者理解更快，仍然可能把必要限定压没。
3. 本章要求因果或结构关系。本章明写结算余损与「副作用」不同：它要求能够说明这笔损失为什么与本轮成功的优化方向存在因果或结构关系。后果效度不要求这一条，它容纳纯粹伴随的后果。

判定形式。在一个完全没有外部评价的写作场景（自己写日记，无人阅读，无人评分，也不使用任何量规）中，按本章的三域做余损审计。若仍能逐条指名余损且与本轮优化方向存在可说明的因果关系，本章成立且不可被后果效度覆盖；若在无测量条件下余损全部消失，本章应并回后果效度，成为它在写作教学中的一个应用。

（二）残余风险

本章在第八节用过「残余风险已经稳定」这个说法，正文一次也没有意识到，这是安全与风险管理标准里的一个正式术语：残余风险（residual risk）= 在风险处理措施实施之后仍然存留的风险（见 ISO/IEC Guide 51 与 ISO 31000 一系的定义）。

一个与自己的核心概念几乎同名的现成术语，出现在自己的正文里而未被辨认，是本章体例上的一处失手。编者补上，并给出分离线。

分离线。

4. 残余风险是处理之后剩下的——它的参照是一次已经施加的控制，被处理掉的部分与剩下的部分在同一本账上。本章的余损是被挤到另一本账上的——它的参照是本轮通过条件所没有登记的质量域。前者是同一维度上的剩余，后者是跨维度的转移。
5. 残余风险可以被接受。那一支的整套实务就是「把风险降到可接受水平，然后正式接受残余部分」，接受本身是一个有记录的决定。本章的余损不进入任何记录——正因为本轮标准没有失败，它没有理由继续修改自己。
6. 后果不同。残余风险的存在不改变下一轮的目标；本章的余损夺取下一轮的终点。

判定形式。对同一次改进，先按残余风险的口径清点（本轮标准之下未被降到零的那部分），再按本章口径清点（本轮标准之外新出现或被放大的那部分）。若两份清单重合，本章应改用现成术语；若本章清单里存在一批项目，它们在本轮标准的维度上读数为零或更好，本章成立——而这恰恰是本章第六节反复举的那类例子（把段落压得更紧以后，清楚度这一维读数更好了，必要限定却被压没）。

（三）未预期后果与外部性

本章第十三节把「目标置换」处理为纯形式上游的一位（在第二章的对应节里也处理过）。编者补两位与它相邻、而本章未提的：

- **Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American Sociological Review*, 1(6), 894-904.** —— 有目的的行动会产生行动者未曾预期的后果，其来源包括知识不足、当下利益压过长期后果、以及基本价值本身的自我否定。
- 外部性（经济学，Pigou 一系）：一项交易的成本或收益落到交易边界之外、不进入交易双方的核算。

分离线。编者认为这两位与本章的关系比本章已处理的那几位更紧，因为结算余损在形式上就是一次结算的外部性：损失落在结算边界之外，因此不进入这次结算的核算。本章要保住自己，靠的是三条：

7. 本章要求损失与本轮成功的优化方向有因果或结构关系，外部性不要求（外部性可以纯属伴随）。
8. 本章要求这笔损失夺取下一轮终点，即它必须进入下一轮的操作序列；外部性与未预期后果都不规定后续动作。
9. 本章处理的是同一个主体的连续轮次（写作者对自己的稿子），不是不同主体之间的成本转嫁。

判定形式。若在批案例中，被指认的余损全部可以还原为「行动者当时不知道」（知识不足），而不能还原为「本轮通过条件的构造方式」，本章应并回未预期后果一支；若余损可以由通过条件的构造预先推出——即知道了本轮拿什么当终点，就能预言损失会落到哪一域——本章成立。这一条是本章可执行性最高的一次检验，编者建议优先跑它。

三·编者提示：本章的枢纽没有刻度

本章的方法叫「追余写作」，五个动作里最关键的是第四个：由最大、最有因果依据、且不能被当前终点继续解释的余损，夺取下一轮的修改终点。

这一句里有三个待判定的量，正文一个也没有给出清点规则：

10. 「最大」怎么比。三类余损（规外、路外、局外）分属不同的质量域，本章没有给出跨域可比的尺度。一笔「必要限定被压没」与一笔「学生开始围绕可见条目做策略性优化」，谁更大？

11. 「有因果依据」怎么判。本章明确把这一条作为余损与「副作用」的分界，却没有给出判定程序——是要求可说明的机制，还是要求可观察的共变？
12. 「连续两轮不迁移才允许停」怎么检测。这条停止规则要求一个「没有新余损」的判定，而没有余损检测程序，这条规则无法执行；本章自己也承认「一次『没有新余损』可能只是检测盲区」。

本章刻意不造一个「追余指数」，改用序列读数（本轮终点落在哪一域→下一轮落在哪一域），并给出了理由：先看焦点迁移序列，不要先造总分。编者认为这个选择是可辩护的——本卷第六章的天花板正是「先造一个能拿满分的读数」的后果。但可辩护不等于已完成：序列读数同样需要一条规则来决定序列的下一项是什么，而这条规则目前落在编码者的判断上。

对读者的建议：本章可以按方法读，不宜按读数读。它在本卷九章中是唯一一章其读数尚不能被两名独立编码者一致产出的。

四·一处编者认可的做法

本章在成文前做了一次公开字段审计，对象是六套现行的官方写作评价框架。审计推翻了本章自己最初的批判——多维评分早已普遍存在，「主流作文评价只有一把尺」这句话不成立。作者因此把主张收窄为：缺的不是维度，是维度之间的时序夺靶规则。

编者把这一处单列出来，是因为它示范了一件事：一次审计如果只能证实自己，它就没有被真正执行。本章的审计改变了本章的主张，这是它值得被收进本卷的主要理由之一。

对切二·第六章 ↔ 第七章：对同一个动作的相反估价

为什么必须切。两章以同一个动作为轴——结算——却给出方向相反的价值判断。第六章教人怎样合法地宣布完成，第七章说宣布完成正是损失逃逸的那道口子。放在同一本书里而不切，读者会认为编者没读懂其中一章。

分离线。

- 第六章问账内的余额。它的对象是已经激活而尚未结清的义务。结清或合法限界之后，账平了，作者可以退场。
- 第七章问账外的新增。它的对象是本轮通过条件没有登记的那些质量域里，因为这次成功而仍然存在或新产生的损失。它明确区分自己与「还有错误」：普通错误意味着当前标准已经看得见问题，只是还没修；结算余损恰恰出现在标准宣布成功之后——当前标准没有失败，因此它没有理由继续修改自己。

一个测未结，一个测未记。

判决性对照。取一批退场闭合率达到 1.00、且经第六章的编码手册确认无未结核心义务的作文，交给三名不知道该批作文闭合状态的评阅者，按第七章的三域（成品、路径、环境）做余损审计。

- 若查不出任何可指名的余损 $\Rightarrow$ 第七章伪。「结算余损」只是「未结义务」的换名，第七章应并回第六章。
- 若查出的余损与第六章的未结义务清单完全重合 $\Rightarrow$ 两章是同一个量的两种说法，本卷应把它们合并成一章，并承认多设了一个概念。
- 若余损可以逐条指名、且与未结义务清单不重合 $\Rightarrow$ 两章各自成立，且第七章证明了一件第六章没能证明的事：闭合不等于无损。

这一处对切还有一个更麻烦的可能，编者一并写出。第六章的探索性编码已经发现，退场闭合率在中高分段迅速达到天花板——一批人工评分只有中等的作文也可以闭合。若第七章的余损审计恰好在这一批「已闭合而质量中等」的作文上收获最大，那么两章合起来给出的不是矛盾，而是一条本卷未曾预料的联合命题：闭合是一个低门槛的状态，而质量的余下部分全部落在闭合之后。若如此，本卷第六章的阈值 0.90 就不是「完成线」，而是「余损审计的起点线」。这条联合命题目前没有任何数据支持，编者只把它作为对照的第三种可能列出。

第八章

何谓一篇作文的好

—— 判准回写余量

何谓一篇作文的“好”：从 评分达标到“判准回写余 量”

——公司金融与治理 $\times$ 保护生物学与生物多样性 $\times$ 软件安全
与隐私工程的二阶碰撞

研究类型: What (三维度型: S / D / E) | 方法版本: SDE v3.9 | 研究对象: 作文评价中的“好”

摘要: 一篇作文何以被称为“好”, 似乎是语文教育中最熟悉、也最难重新发问的问题。长期以来, 内容充实、中心明确、结构完整、语言准确、表达生动、立意深刻等判准, 构成课堂教学、考试评分、范文生产与自动作文评价的共同词汇。问题在于: 这些判准一旦被公开、反复教学、用于训练并进入生成式人工智能的优化回路, 它们就不再只是外部测量作文的尺子, 而会反过来成为作文生产的目标。于是, 高度“像好作文”的文本与真正完成写作任务的文本可能发生系统性脱钩: 学生可以学习评分特征, 模板可以复制评分特征, AI 更可以低成本地同时满足大量显性特征。本文依照 SDE 多学科通融创新法 v3.9 的 What 型三维度工序, 从 SDE Universes “新思想前沿”选取三门与语文教育距离较远、且在闸零中未检出被作文评价系统吸收的学科: S 位的公司金融与治理、D 位的保护生物学与生物多样性、E 位的软件安全与隐私工程。公司治理研究显示, 治理指标公开并被市场学习后, 原先的异常收益关联可以消失, 说明“指标被采用”本身会回写指标含义; IUCN 绿色状态把“避免灭绝”与“恢复分布和生态功能”分开, 说明没有失败并不等于已经达到正向功能; 控制流完整性研究则要求在明确威胁模型、合法目标集合和豁免清单后判断系统是否安全, 说明正常条件下的表现不能替代对抗性边界中的成立。三家碰撞共同击穿作文评价中的一个隐含前提: 判准是稳定、外在且不会因被教学和优化而改变诊断意义的。本文由此提出新对象“判准回写余量”(Criterion-Feedback Residual, CFR): 在作文已满足最低任务合法边界后, 当显性评价判准已经公开并进入教学、练习与 AI 优化, 控制这些显性判准的达成程度, 文本在未用于训练和评分的保持性任务、变体情境与独立读者中仍能稳定产生的文体功能增量。本文进一步给出“显性判准达成度 $\times$ 判准回写余量”的 2×2 辨别格、CFR 研究读数与课堂简化版清点手册, 并以写作评价中的效度、washback、Goodhart 式指标异化、自动作文评分及 2026 年 AI 生成作文实证作为敌意近邻进行划界。结论是: 好作文不是评分项目

的总和，也不是脱离判准的神秘整体感；它是在最低合规之后，面对判准已经被学会、被模仿、被优化这一事实，仍保有可重复、可迁移、不可由表面合规完全替代的功能余量。

关键词：作文评价；语文教育；二阶碰撞；判准回写余量；评价回写；生成式人工智能；写作教学

Abstract: What makes an essay “good” becomes harder to define once evaluative criteria are no longer external to writing but are explicitly taught, practiced, optimized, and increasingly reproduced by generative AI. This paper uses the What-type S/D/E collision procedure of SDE v3.9 and brings together three distant fields selected from the SDE Universes Frontier: corporate finance and governance (S), conservation biology and biodiversity (D), and software security and privacy engineering (E). The first shows that a public indicator can lose its original predictive association after market participants learn and price it; the second separates mere avoidance of extinction from recovery of distribution and ecological function; the third evaluates security only within an explicit threat model and legal control-flow boundary. Their collision rejects the assumption that writing criteria remain diagnostically stable after they become production targets. The paper proposes Criterion-Feedback Residual (CFR): after an essay satisfies a minimum task-admissibility boundary and after explicit scoring criteria have entered teaching and optimization, CFR is the reproducible genre-specific functional gain that remains across held-out tasks, prompt variants, and independent readers after controlling for visible criterion compliance. A 2×2 diagnostic grid, an operational counting manual, falsification tests, and classroom implications are provided. Good writing, on this account, is not the sum of

rubric traits; it is the functional remainder that survives the feedback of the rubric into the production process.

Keywords: writing assessment; Chinese language education; criterion feedback; rubrics; generative AI; interdisciplinary collision

核心命题 好作文 ≠ 评分项目之和。好作文 = 最低任务合法边界成立 + 判准被公开并回写生成之后，仍可重复、可迁移、不可由显性合规完全解释的功能余量。

二阶碰撞源与位置摘要

位	SDE 前沿取源	可引核心判断	本文占位	闸零
\$	公司金融与治理：治理指数异常收益消失	指标公开并被市场学习后，原先的收益关联可消失；采纳行为会回写指标含义	对象显露/信号意义	干净；厚
0	保护生物学与生物多样性：IUCN Green Status	避免灭绝≠恢复；恢复包含分布、存续、生态功能、依赖和潜力	变化轨迹/正功能	干净；厚
生	软件安全与隐私工程：控制流完整性	正常运行≠对抗安全；须写明合法目标集合、威胁模型、豁免与绕过	条件场/鲁棒边界	干净；厚

一、问题不在“标准太少”，而在标准已经进入了作文的生成回路

“什么是好作文”通常被处理成列举题。教师会说内容真实、中心突出、材料具体、结构严谨、语言流畅；考试量表会把这些词拆成若干等级；写作软件会把其中一部分转成可以计算的特征；生成式人工智能则能够按要求同时执行“观点明确、层次清楚、过渡自然、句式有变化、结尾升华”等指令。列举越细，看起来越接近科学评价。然而，一个更困难的问题随之出现：如果学生已经知道评分表，如果教师围绕评分表反复训练，如果范文把高分特征稳定展示出来，如果 AI 可以把这些特征即时补齐，那么这些特征还是原来意义上的“好”的证据吗？

这不是说评分标准没有用。相反，没有共同判准，作文评价会退回私人品味；没有明确的任务要求，学生会面对不可见的权力。问题发生在第二步：教育评价的尺子会被被评价者看见，并被当成行动目标。一条原本用于识别质

量的特征，经过长期教学以后，可能变成可以单独生产的表面信号。五段式结构最初可能与清晰论证相关；当“五段式”本身成为训练目标，结构完整就可以在论证薄弱时独立存在。细节描写最初可能与具体经验相关；当“必须有动作、语言、心理、环境”成为操作口诀，细节数量就可以在经验贫乏时迅速增加。结尾的思想提升最初可能来自全文生成；当“最后一定升华”成为评分预期，一句抽象价值判断可以被贴到几乎任何故事之后。

生成式 AI 把这种长期存在的现象推到了极端。ETS 关于 e-rater 的说明显示，自动作文评分会利用内容词汇、词汇复杂度、语法与用法错误比例、风格、组织与发展等特征来预测写作表现；而 2026 年 Zhong 等对 AI 生成 GRE 作文的研究报告，AI 文本在不少语言相关特征上强于人类基线，同时 e-rater 对 AI 生成作文的评分系统性高于人类评分。这里最重要的并不是“机器不如人”，而是一个更基础的评价学事实：过去在人类写作样本中能够间接指向深层能力的特征，在生成机制改变之后，其诊断关系可以变弱。语法准确、组织整齐、词汇成熟仍是文本属性，但它们与推理深度、证据质量、思想生成之间不再保持原来的统计绑定。

因此，本文不准再增加第六个、第七个或第八个“好作文标准”。如果旧标准已经成为生成目标，再加一条标准，只会把新标准也送入同一个回路。真正需要解释的是一个二阶问题：当评价判准已经被公开、被教学、被训练、被模板化、被 AI 优化以后，“好”还能以什么方式成立？这意味着研究对象必须从“作文有哪些好特征”转向“判准进入生成回路之后，还剩下什么无法被判准本身解释的质量差异”。

这一转向也使本文与 What1 “作文是什么”的对象论分开。作文的身份可以讨论作者、文本与读者之间如何构成一个写作事件；但“好”是另一个问题：即便两篇文字都毫无疑问是作文，也都能被读者理解，它们仍可能在质量上有根本差异。本文的任务不是再次证明文本能够承载作者意图，而是寻找一种能够解释“同样高分、同样流畅、同样结构完整，为什么仍有一篇明显更值得称为好”的新判据。

二、What 型的答案必须是“一个东西”：把“好”从形容词改造成可裁决对象

SDE v3.9 首先要求题型判别。本文问的是“何谓一篇作文的好”，答案形状不是一条操作路线，也不是一个单一驱动力，而应当是一个可以被指认、比较和证伪的对象。因此采用 What 型三维度：S、D、E 三家不能都回答同

一个局部问题，而要分别占据对象显露、变化过程和成立环境三个位置。若三家只是共同证明“作文要有内容”“作文要真实”“作文要考虑读者”，那仍是三条赞成票，不会逼出新对象。

本文先冻结一个占位句：一篇作文的“好”，应当是在显性评价指标已经进入写作生成之后，仍不能被这些指标的达成所穷尽的质量。这句话暂时不算结论，它只用于寻找“日常就在处理同类困难”的远距学科。我们需要三类材料：第一类必须处理“指标被公开和采用以后，指标与对象关系会不会改变”；第二类必须处理“没有失败与真正达到正向功能之间有没有不可省略的距离”；第三类必须处理“一个对象在正常环境中通过检验，为什么仍不足以证明它在关键边界上成立”。

按照这一占位句，本文从SDE Universes“新思想前沿”中选出公司金融与治理、保护生物学与生物多样性、软件安全与隐私工程。它们不因为“作文像公司、像物种、像软件”而被选择；恰恰相反，类比越像越危险。选择它们的理由是三门学科各自在自己的硬问题中发展出了一套不能彼此替换的判决规则，而这些规则恰好对作文评价的共同前提形成斜向冲击。

三、闸零：三门学科为什么可以进场

3.9版在选源之前增加“源的未吸收度”和“理论厚度”两道闸。前者防止把语文教育早已吸收的理论重新包装为跨学科创新，后者防止使用只有新名词、没有争论史和反例传统的薄学科。本文的检索不是要证明世界上“从来没有任何人”把这些领域与写作并列，而是检查一个更严格也更可操作的问题：作文评价的主流理论是否已经系统吸收这三家的核心判决机制，使本文的碰撞可以被既有写作概念1:1替换。

对公司金融与治理，检索组合包括“corporate governance / governance index + writing assessment / essay rubric / 作文评价”，以及治理指数、市场学习、指标异常收益消失与写作评价的交叉词。检到的主要是“公司治理课程本身怎样用rubric批改学生作业”之类偶然同页，而未检到把“治理指标被市场学习后预测关联消失”作为作文质量理论的系统迁移，故判为“干净”。理论厚度判“厚”：其奠基层有治理指数与股东权利测量，内部有治理条款究竟代表什么、内生性如何处理的争论，反例有治理—收益关联在后续样本中消失，且已经形成稳定的公司治理、资产定价与公司金融研究传统。

对保护生物学与生物多样性，检索“IUCN Green Status / species recovery + writing assessment / essay / 作文评价”等组合，没有发现其“避免灭绝不等于恢复；恢复必须含分布、存续、生态功能及对保护行动的依赖”被语文教育系统吸收，判为“干净”。理论厚度判“厚”：从IUCN红色名录到绿色状态有明确标准、历史基线争议、功能定义争议、依赖与反事实评估，并已有全球标准和实际物种试评估。

对软件安全与隐私工程，检索“control-flow integrity / CFI + writing assessment / essay rubric / 作文评价”等组合，没有发现把“合法目标集合、威胁模型、豁免清单和对抗绕过”作为作文质量对象论的系统引入，判为“干净”。理论厚度同样为“厚”：CFI至少从Abadi等2005年的工作起形成长期研究链，既有细粒度与粗粒度的实现分歧，也有JIT-ROP、control-flow bending等反例传统，还有编译器、二进制重写、硬件支持与现实部署中的兼容性问题。

三家的S/D/E落位也不能互换。公司金融与治理占S位，因为它直接改变“一个可见分数还显露什么对象”的判断：同一个治理指数，在被市场广泛学习之前和之后，作为收益信号的含义不同。保护生物学占D位，因为绿色状态把物种从“没灭绝”推进到“在历史范围内可存续并履行功能”的恢复轨迹，并进一步追踪保护依赖、未来增益和恢复潜力。软件安全占E位，因为控制流完整性从不在真空中宣称“安全”，而是在攻击者能力、允许目标集合、运行环境和豁免条件被写明之后判断结论是否成立。三家分别卡住信号、轨迹、场，因此位置三分成立。

四、S位因果脊：公司治理告诉我们，公开的“好指标”会被对象学会

2003年，Gompers、Ishii与Metrick以24项公司治理条款构造治理指数，并报告基于强弱股东权利组合的交易策略在其样本期存在显著异常收益。这个发现的重要性不只在公司金融内部：它展示了一个常见的现代评价理想——把复杂质量压缩成可见指标，再利用指标预测更远处的真实结果。作文评分同样依赖这个理想。我们不会因为一篇作文“用了三个比喻”就说它好，而是假定某些可评分特征与更深的表达、思考和沟通质量保持稳定关系。

真正有碰撞价值的是后续变化。Bebchuk、Cohen与Wang把观察期延长后发现，治理指数与异常收益的关联在2000—2008年不再存在，并给出与“市场参与者学习了治理差异”一致的证据。原文摘要的关键句是治理指

数 “was no longer associated with abnormal returns during the period of 2000-2008”。这里发生的不是简单的测量误差，而是指标进入市场知识以后，市场行为改变，原先可利用的关系被吸收。指标被更多人知道，本应让评价更成熟；但恰恰因为更多人知道，指标原来的信息优势反而衰减。

把这条因果脊迁入作文评价，第一层冲击是“公开判准不只是信息，它还是干预”。教师把“中心明确”写进量表，学生就会在开头和结尾重复中心句；把“细节丰富”写进量表，学生会增加可计数的动作、语言、心理；把“高级词汇”作为高分迹象，训练与 AI 就会提高词汇复杂度。此时判准不再站在作文外面观察作文，而是进入作者的行动函数。评价制度越成熟、范例越公开、反馈越即时，文本越可能被塑造成评价者希望看见的形状。

第二层冲击更关键：**被塑造出来的特征仍然是真的，但它不再具有原来的诊断强度**。AI 生成的语法准确是真准确，结构清楚也是真清楚；问题不是这些属性“假”，而是它们对深层写作质量的证据权重改变了。正如治理指数在后期仍可与公司价值和经营表现相关，却不再产生原来的异常收益，同一作文特征在新生成环境中可能仍有价值，却不能继续承担过去那样强的替代性判断。

因此 S 位给出第一条反常识命题：**好作文不能定义为若干“长期与好相关”的可见特征，因为一旦这些特征被制度化生产为目标，它们与深层质量之间的关系本身就成为需要重新估计的变量**。一项判准越成功地进入教学，它越不能自动保持“未经教学时”的诊断意义。评价的成功可能侵蚀评价的证据结构。

五、D 位因果脊：保护生物学告诉我们，“没坏掉”与“真正好起来”之间隔着功能恢复

保护生物学长期以灭绝风险为核心对象，这是极其重要的工作，但风险下降并不等于恢复完成。一个物种可以依靠圈养、极小范围保护或持续人工干预避免灭绝，却仍然缺席于大部分历史分布区，也没有恢复原有生态功能。IUCN 绿色状态的转向恰恰把这种“成功中的失败”写进评价：评价不仅问物种是否还活着，还问它是否在空间单元中可存续、是否履行生态功能、目前状态多大程度依赖持续保护、未来还能恢复到哪里。

Akçakaya 等 2018 年的工作把这一方向概括为需要建立 “an objective and ambitious definition of species recovery”。2021 年 Grace 等对全球标准进行测试，进一步强调绿色状态将功能纳入物种恢复评估。这里有一个

对作文教育极有价值的结构：红色名录并没有因为绿色状态出现而“错误”。避免灭绝仍是必要目标，只是它回答的是“怎样不掉出底线”，而不是“怎样达到完整正功能”。

作文评价中大量指标其实也具有“红色名录”性质。跑题、病句、错别字、材料空泛、结构断裂、论据与观点无关，都是需要排除的风险；中心明确、层次完整、基本连贯则常常构成“避免失效”的条件。问题在于，我们经常把“风险降低”直接写成“质量提高”。一篇没有明显病句、结构完整、开头结尾照应、材料贴题的文章，确实可能已经避免了许多失败，但这些成功还不能证明它完成了更强的正向功能。

D 位迫使本文把“好”从错误的反义词中解放出来。坏作文的反面首先是“合格作文”，而不是自动等于“好作文”。好作文需要某种正功能完成：叙事需要让关系、变化和事件意义在文本推进中成立；说明需要使读者获得可迁移的对象模型；议论需要使主张、证据、理由与反条件形成可检验的推理结构；抒情和经验性写作需要把无法用通用套话替代的感受组织成可传递的形式。不同文体的功能不同，但它们共同具有一个特征：**功能不是“没有错误”的总和。**

绿色状态还贡献第二个更尖锐的变量——依赖。一个物种当前看似稳定，如果一撤去保护行动就迅速崩塌，那么“稳定”与“自主恢复”必须区分。迁入作文教育后，我们得到“判准依赖”：如果一篇文章的高质量表现只有在持续提供范文、句式、段落模板、AI 改写或明确评分提示时才存在，而一换题、一换文体、一撤掉即时提示就消失，那么当前高分包含了外部支架的贡献。支架并非坏事，教学本来就需要支架；但把支架支持下的表现直接当成学习者已拥有的稳定写作质量，会把“保护依赖”误写成“已经恢复”。

因此 D 位给出第二条命题：**好作文必须有正向功能读数，并且要把对评分脚手架、模板和即时优化的依赖单列，而不能用“没有明显失败”替代“已经产生功能”。**这使“好”第一次拥有了时间和迁移方向：它不是一次评分现场的静态状态，而是面对条件变化仍能保持多少功能。

六、E 位因果脊：软件安全告诉我们，正常运行从来不是鲁棒性的充分证据

软件安全中的控制流完整性（CFI）处理一个非常具体的问题：程序运行时的间接跳转，能否被限制在静态分析允许的合法目标集合内。Zhang 与 Sekar 在 2013 年的 USENIX Security 论文中指出，CFI 是一项低层安全属

性，其实施可以抵御大量控制流劫持攻击。SDE“新思想前沿”对这一转向的概括非常适合本文：不要只问攻击者“猜不猜得到地址”，而要问运行时跳转是否落在允许集合。

这里最值得迁移的不是“白名单”这个比喻，而是安全研究的判决方式。软件在普通测试中运行顺畅，不能推出它在攻击者主动寻找薄弱路径时仍然安全；一个防御机制写着“已启用”，也不能推出它足够强，因为目标集合的粒度、豁免模块、攻击者能力、依赖版本和运行环境都会改变可绕过空间。后续研究反复说明，粗粒度 CFI、地址随机化或某些现实妥协可以被攻击路径绕过。安全结论因此总是带着条件：在什么威胁模型下，禁止了哪些路径，放行了哪些路径，哪些例外必须单列。

作文评价的“正常运行”情境是什么？通常是学生面对熟悉题型、明确量表、有限时间、教师训练过的段落范式进行写作；评价者再按同一套已知判准打分。这样的环境当然有现实意义，但它天然偏向展示“被训练过的合法路径”。如果我们要判断更强意义的好，就需要加入保持性或对抗性条件：题目表面类型变化、材料组合变化、读者任务变化、评分提示撤去、允许使用 AI 但更换未公开功能测试，或者把两篇显性评分几乎相同的文章进行盲配对。此时，真正被检验的是文本与写作者是否只会走已标出的得分路径。

E 位还提醒我们，不应该用“越难骗越好”取代作文评价。安全系统需要合法目标集合，因为系统仍要工作；如果为了阻止攻击而禁止所有跳转，程序也失去功能。同理，为了防止学生迎合评分而把评价标准全部隐藏，会制造新的不公平与权力任意。本文因此把显性判准区分为两类：一类是最低任务合法边界，必须公开，例如是否回应题目、是否达到基本可理解、引用与事实规则、考试工具规则；另一类是用来区分“好到什么程度”的诊断特征，它们可以公开教学，但研究者必须承认公开本身会改变其证据权重，并以未用于优化的保持性任务重新校验。

由此 E 位给出第三条命题：**好作文不是在单一、友善、已知的评分环境中表现优异，而是在写明任务合法边界之后，其关键功能能在预设的变体和压力条件下保持；任何“好”的结论都必须同时写明它在什么情境下会失效。**这不是把写作变成攻防游戏，而是阻止我们把一次顺利通过当成普遍能力。

七、二阶碰撞：三家共同推翻的不是某条评分标准，而是“尺子不回写对象”的前提

三家第一次碰撞时似乎在说三件完全不同的事：金融谈市场学习，保护生物学谈物种恢复，软件安全谈攻击路径。如果只做一阶类比，可以得到三句常识——“评价指标会失真”“作文要有功能”“评分要考虑情境”。这些话都对，却不足以构成新理论。二阶碰撞要求继续追问：三家为什么都必须改变原来的评价方式？它们共同假定过什么，后来又由自己的材料推翻了什么？

第一层共有前提是：**可见指标与目标对象之间的映射在评价过程中保持稳定**。治理研究最初相信指数与收益关系可以在后续使用中延续；红色名录式风险读数容易被误读为整体保护成功；软件的“已加固/能运行”容易被当作安全状态。三家后续都发现，同一可见结果并不能自动保持同一含义。

第二层共有前提更隐蔽：**评价行为本身只是观察，不进入被评价对象的因果链**。这正是公司金融材料亲自推翻的地方。治理指标和研究结果被市场看到后，市场学习、定价与套利改变了后续关系；“公布指标”不再是观察之外的动作。迁入作文场景，量表、范文、点评和 AI 评分同样是干预：它们改变学生注意什么、练习什么、文本生成器优化什么。只要判准可以被行为主体感知，评价系统就不是外部相机，而是生产环境的一部分。

保护生物学随后给这个新问题补上方向：当旧指标因为被优化而越来越容易达标，我们不能只寻找更难达标的指标，而要把目标从“避免失败”推进到“恢复功能”。软件安全再补上边界：功能结论必须在明确的压力场中检验，不能只在训练环境里成立。于是三家共同逼出一个作文评价账本从未稳定记录的字段：**显性判准已经回写生成以后，文本还剩多少正向功能？**

这个字段既不是“创新性”，也不是“真情实感”，更不是“教师整体印象”。创新性可以被评分表公开并被策略性制造；真情实感很难直接测量，而且优秀虚构写作并不要求事实经历真实；整体印象把问题重新交还给评价者，却没有处理评价者本身也可能被范式训练。新字段必须能够在显性得分大致相同的文本之间继续产生可复核差异，并且这个差异要来自未被同一判准直接训练的功能任务。

因此本文把新对象命名为“判准回写余量”。“判准”指已经被公开并实际进入教学或评分的特征集合；“回写”指这些特征从测量端返回生产端，改变作者、教师、模板或 AI 的生成策略；“余量”不是神秘剩余，而是控制显性判准后，在保持性条件中仍可重复观测到的功能增量。这个名称刻意不用“真正质量”“深层质量”，因为后两者容易成为不可证伪的价值宣

言；“余量”要求我们说清楚：相对于哪套判准、在哪些未训练条件、用什么功能任务、剩下多少。

八、新对象 Z：判准回写余量（CFR）的严格定义

本文将判准回写余量（Criterion-Feedback Residual, CFR）定义为：对于一篇已进入任务合法集合的作文，在一套显性评价判准已经被公开、教学或用于生成优化之后，控制该作文对这些显性判准的达成水平，其在未用于训练与直接评分的保持性任务、题目变体和独立读者中，仍能稳定产生的文体功能增量。

定义中有五个不可删条件。第一，先有“任务合法集合”。跑题、无法理解、违反明确引用规则或未满足任务必需条件的文本不能仅凭某个神奇余量越级成为好作文。第二，判准必须是真的“已暴露”——写在评分表上、被教师明确教过、被训练软件反馈过、被 AI 提示词直接优化过，才叫进入回写。第三，要控制显性判准达成，不能把语言更流畅造成的优势冒充余量。第四，功能测试必须至少有一部分未被用于原评分训练，否则只是换一种方式重复旧量表。第五，余量必须可重复：至少跨两个独立读者组或两个变体任务方向一致，单一评价者的惊艳感不算稳定证据。

研究层面的一个可操作读数为“CFR 保持率”。先在同一题目、同一文体内，为目标作文寻找显性量表得分相近的匹配作文；再设计三项不直接写入原评分表的保持性任务。每一项任务让独立读者完成可计数结果，例如：议论文读后能否在新案例中正确应用其论证边界；说明文读后能否回答一个需要迁移而非原句复述的问题；叙事文读后能否在若干相近因果解释中稳定识别文本真正建立的关系。目标作文在某一任务上显著优于匹配基线记 1，否则记 0； $CFR\text{ 保持率} = \frac{\text{通过的保持性任务数}}{\text{全部有效保持性任务数}}$ 。

为遵守单一口径律，本文预注册研究阈值：三项保持性任务中至少通过两项，且两个独立读者组的方向一致，记为“CFR 阳性”；只通过一项记为“不确定”；三项均未通过记为“CFR 阴性”。这里的 2/3 不是宇宙常数，而是本文第一轮研究方案的裁决阈值，后续可通过数据校准，但同一研究中正文、证伪条款和结论必须使用同一个阈值。

若需要连续量，可进一步使用标准化残差形式。设显性判准向量为 C ，保持性功能表现为 F ，在历史人类样本上估计 $\hat{F} = g(C, T)$ ，其中 T 表示题目与文体；目标作文的功能残差 $r = F - \hat{F}$ 。在至少三个保持条件 k 中取 r 的中位数并标准化，得到 $CFR-z$ 。 $CFR-z$ 大于 0 表示该作文在同等显性判准水平下产生了高

于预期的功能表现，小于 0 则表示显性得分高于其保持性功能所支持的水平。课堂不需要计算 z 值，但研究可以用它区分“高分来自显性特征”与“高分之外仍有功能”。

这一定义故意允许 CFR 随评价制度变化。同一篇文章放在不同教育生态中，判准是否已经被广泛教会不同，其“余量”可以不同。有人会认为质量应该是文本固有属性，怎么能随环境变化？三家碰撞恰恰要求放弃这个假设：当评价标准已经成为生成过程的一部分，“好”的证据结构本来就与环境相关。CFR 不是宣布文本本体随人任意变化，而是承认我们用什么证据称它为“好”，必须考虑这些证据是否已经被系统性生产出来。

九、2×2 辨别格：把“高分”和“好”重新拆开

把横轴设为“显性判准达成度”，纵轴设为“判准回写余量”，可以得到四种过去经常被一个总分压扁的文本类型。

第一格是低显性达成、低 CFR，可称“基础失效”。它既没有满足任务与常规表达要求，也没有在保持性任务中产生额外功能，传统评价与本文评价基本一致。

第二格是高显性达成、低 CFR，可称“合规拟态”。这是本文最关心的隐藏失败：文章中心清楚、结构完整、词语成熟、段落均衡，甚至自动评分很高，但在新情境迁移、理由边界识别、机制理解或延迟读者任务中不优于同分基线。模板作文与部分 AI 高分文本可能落在这一格，但“AI 参与”本身不是判定条件；人类也完全可以写出合规拟态，AI 也可能帮助形成高 CFR 文本。

第三格是低显性达成、高 CFR，可称“异格有效”。它可能不符合某套既定量表，却在保持性任务中产生很强功能。例如一篇极简叙事没有典型的开端—发展—高潮—结尾，却让读者对人物关系形成高度稳定且可迁移的理解。教育评价不能因此自动给它高分：若其违反明确任务要求，仍应在任务合规上扣分。但这格提醒我们，量表可能发生构念漏失，不能把“量表未捕获”直接写成“文本没有质量”。

第四格是高显性达成、高 CFR，可称“稳健好文”。它既满足任务、语言和组织等显性要求，又在这些要求被控制后仍表现出可迁移的功能优势。本文所说的“好作文”主要指这一格，但需要再加一个最低合法边界：若文章事实欺骗、严重跑题或违反明示的任务规则，即使某项功能很强，也不能作为该任务中的好作文。

这个2×2最重要的用途不是制造新评分表，而是强迫评价者承认两个维度不可互相代替。传统总分主要集中在横轴；CFR提供纵轴。如果纵轴最终与横轴高度相关到几乎没有独立信息，本文理论应被削弱甚至放弃；如果出现稳定的高横低纵与低横高纵样本，则说明“好”确实存在一个被传统量表压缩掉的方向。

十、五场景压力测试：同一个“高分”为什么要给出不同判决

为了防止新概念只在抽象层面漂亮，本文预设五个场景，并要求CFR给出不同答案。

场景一：学生独立写作，教师事先只给任务要求，不给细化量表。文章显性得分高，保持性任务也高。这里高分与CFR都提供正证据，可判为稳健好文。场景二：学生接受一个月“高分作文模板训练”，结构、过渡、结尾和细节指标显著提高，但同分匹配后保持任务不升。这里应判“合规进步，但未证明CFR进步”，而不能把分数上涨全部解释为写作质量增长。

场景三：学生先形成完整观点、证据与反例，再使用AI改善语句、删冗和调整段落。最终显性分数上升，保持性功能也保持或提高。CFR不因AI介入自动归零；只要功能余量仍在，而且任务规则允许工具使用，就可以是好作文。场景四：学生只输入题目，让AI按评分表一次生成，再做少量换词。文本显性得分很高，但在新案例迁移和反条件识别中表现一般。此时应判高显性、低CFR。这里批评的不是“机器写”这个身份，而是评分信号与功能脱钩。

场景五：一篇风格非常陌生的文章在常规量表上组织分低，却在两组读者的保持任务中显著优于同题文本。此时不能直接宣布它是考试中的好作文，因为最低任务合法性和评分公平仍存在；但研究者必须把它记入“异格有效”，并检查量表是否漏掉一种真实功能。新理论若只奖励标准作文，就没有发现新维度；若一见怪异就奖励，也会退回审美浪漫主义。

其中至少两个场景受到外部资料反向修正。第一，Cambridge近年的写作评价讨论明确指出，评分标准既是质量的心理模型，也会变化，并强调评估后果与公平，因此本文不能声称“写作评价从未考虑评价的回写”；本文只保留更窄的主张：现有讨论通常把回写放在测验效度、教学后果或公平层面，而没有把“控制显性判准后剩余的文本功能”定义为作文之好的对象。第二，2026年AI生成作文研究说明，AI文本并非简单“质量低”：部分先进

模型的人类评分本身已经不低。于是本文不能把 CFR 写成“人类好、AI 坏”的身份判据，而必须坚持功能与条件判据。

十一、敌意拓宽：八类最危险的占位者为什么都不能 1:1 替换 CFR

新概念最容易失败的方式不是逻辑错误，而是早有人用另一个名字说过。本文因此主动把最危险的近邻摆进来。

第一类是分析性评分量表。它把“好”拆成内容、组织、语言、论证等维度，优点是透明、可训练、可提高评分一致性。但它通常把维度当作稳定证据，没有把“维度一旦公开后被策略性生产”设成对象变量。CFR 不是多加一个评分维度，而是询问现有维度进入生产后还有多少解释剩余。

第二类是整体评分。整体评分允许评价者对文章总体质量作综合判断，似乎天然能捕捉“量表之外”的东西。但整体感本身也会受到范文、考试文化、语言风格和评价者训练塑造，而且不能自动区分“我喜欢”与“功能保持”。CFR 要求匹配显性分数并使用独立保持任务，因此不是整体印象的技术化改名。

第三类是构念效度与构念覆盖。ETS 长期研究提醒自动评分必须覆盖写作构念，现代写作评价也反对把表层语言当作全部写作能力。这是极危险的占位者。分离线在于：构念效度的中心对象是“分数解释是否支持我们想做的推断”；CFR 把判决对象换成“文本在判准回写以后还保留多少功能”。前者可以发现量表漏测，后者还要求把判准暴露本身作为生成条件并做保持测试。

第四类是 washback 与后果效度。它们直接研究考试怎样改变教学和学习，确实已经承认评价不是中性的。这一近邻迫使本文大幅收窄创新主张。CFR 不声称发现“考试会反过来影响学生”，而是把这种影响继续推进一步：一旦 washback 已经发生，原先判准的诊断含义要重新估计；作文的“好”必须在控制被回写的特征后再裁决。

第五类是 Goodhart 定律与 Campbell 定律。它们指出指标成为目标后会被腐蚀，几乎与公司金融 S 位同脊。若本文只说“不要唯分数”，创新就失败。CFR 的新增结构在三处：一是为作文建立最低任务合法集合，防止反指标主义；二是引入保护生物学的正功能/依赖区分，要求测量“好起来”而非只识别指标失真；三是引入安全工程的保持性压力条件和豁免边界，使理论可以被课堂与研究实测。

第六类是写作过程评价与作品集评价。它们强调构思、修订、反馈和多次作品，避免一次成品代表全部能力。CFR 可以与过程数据结合，但两者不是同物：一个拥有完整过程记录的作品仍可能高度围绕已知量表优化；反之，一篇无法获取过程数据的匿名文本也可以研究其判准回写余量。

第七类是自动作文评分的对抗鲁棒性研究。此类工作测试评分器能否识别跑题、胡言、句序扰乱或 AI 生成文本，关注“评价器会不会被骗”。CFR 关注的是另一端：“文本即使让评分器满意，它在未训练的功能任务上还剩什么”。评价器鲁棒而文章 CFR 低、评价器脆弱而文章 CFR 高，两种组合都可能出现。

第八类是修辞效果、读者反应与真实写作。它们早已强调写作应面向真实受众和真实目的。这些理论提供重要前史，但常把效果放在具体读者—情境关系中。CFR 并不取代它们，而是增加一个二阶控制：当“面向受众”“有感染力”“有证据”等效果性判准也已经被教成套路时，仍需要用未参与训练的任务检验其保持性。

按 v3.9 的敌意拓宽要求，再向历史与应用内部寻找占位者：1961 年 Diederich、French 与 Carlton 对大量评价者判断因素的归纳，是现代写作量表维度化的重要前史；1970 年代 Campbell 关于社会指标被高压使用后易受腐化的论述，则提供指标回写的一般社会科学前史。应用领域内部，ETS 的 e-rater 构念覆盖与 2026 年 AI 作文研究、Cambridge 对写作评价公平与效度的最新讨论，都直接构成比抽象理论更危险的近邻。经过这些近邻后，本文不再把“指标被优化”“评分有后果”当创新点，只保留“以判准暴露为条件、以控制显性判准后的保持性功能残差定义好”的新对象。

十二、一次真实的反证跑：公开资料没有把本文杀死，但迫使它缩小了三次

为了避免“先写结论、再挑文献支持”，本文把公开资料检索设计成一条可以让新概念失败的反证跑。预注册的第一条杀死条件是：如果写作评价文献已经明确把作文质量定义为“在评分标准被公开并进入教学后，控制这些标准达成度，在未训练任务中剩余的功能表现”，那么 CFR 只是改名。检索结果没有发现完全同构定义，但发现 washback、后果效度、rubric 批评和 Goodhart 式讨论已经覆盖“标准会影响教学与行为”。因此本文删除了“过去评价把尺子当纯外部观察者”这种过强表述，改为“现有研究已经看见回

写，但通常没有把回写后的功能余量设为作文之好的对象”。这是第一处收窄。

第二条杀死条件是：如果 AI 生成文本在显性语言特征上更优时，人类对整体写作质量也同步同幅提高，那么“显性指标与功能脱钩”的现实必要性会减弱。Zhong 等 2026 年的结果相反：AI 生成作文在语法、用法、机械准确性等语言特征上表现强，而 e-rater 相对人类评分对 AI 文本给出更高分。这个结果支持“特征—深层判断关系随生成机制变化而重估”，但同时也显示先进模型的人类评分并不低。因此本文删除“AI 作文只是表面漂亮”的简单说法，改为：AI 是一种使旧特征权重发生变化的生成环境，具体文本仍需逐篇判定 CFR。这是第二处收窄。

第三条杀死条件针对“隐藏判准”。如果 CFR 要求评价者永远保留秘密标准，理论会与公平原则冲突。Cambridge 关于写作评价的讨论强调，透明标准有助于减轻学生对不可见目标的焦虑，评价标准应与任务、教学和时间条件相关。本文因此明确：最低任务边界必须公开，显性评价维度也可以公开；保持性任务不是教师秘而不宣的“真正标准”，而是研究和教学中轮换的、预先说明其存在但不把具体答案教成模板的迁移检查。CFR 反对的不是透明，而是把“透明后被学会的信号”继续当作未经回写的纯测量。这是第三处收窄。

这次反证跑还有一个不利结果：Goodhart 思路在写作教育中并非空白，近年来已有直接把公式化写作与 Goodhart 联系起来的讨论，2026 年的语言复杂度自动评价研究也明确提到指标可能被策略性操纵。由此，本文若把创新点放在“评分标准成为目标就会失真”，I 维应当判低。新对象必须依赖三家共同提供的额外结构：**回写之后测什么正功能、怎样区分支架依赖、在什么压力场中保持、如何对同分文本计算余量。**

十三、清点手册：如何在研究中真正测“判准回写余量”

为了让 CFR 不沦为新的漂亮词，本文给出最小清点手册。

第一步，确定评价制度版本。记录具体评分表、发布时间、学生何时看到、教师是否逐项教学、是否存在范文、是否使用自动反馈或 AI 辅助。没有“判准暴露史”，就不能谈回写。表头至少包括：文本编号、题目、文体、评分表版本、暴露时长、训练方式、AI/模板使用规则。

第二步，定义最低任务合法集合 L。L 只包含任务成立的硬边界，而不承担优秀程度排名。例如“回应指定材料”“达到基本可理解”“遵守引用与工

具规则”“满足规定文体的最低交流任务”。L 最好用 0/1 记录，并由两名评价者复核。把“文采好”“立意深”塞进 L，会让余量在进入测试前就被旧标准吞掉。

第三步，记录显性判准向量 C。若现有量表有内容、组织、语言、论证四维，就原样保留，不自行美化。研究中为每篇目标作文寻找 C 相近的匹配文本：总分差不超过量表总分的 5%，且核心维度不跨越一个等级。找不到匹配项的文章不进入“余量”比较，而应单列为不可匹配。

第四步，为文体建立三个保持性任务。保持性任务要满足四条：不直接重复评分表措辞；需要读者使用文本完成一个新动作；答案可以计数；任务难度对同题文章基本公平。议论文可用“新案例适用判断、最强反条件识别、证据—结论错配检测”；说明文可用“新情境预测、机制缺口定位、近似概念区分”；叙事文可用“因果关系辨别、人物选择在新情境中的一致性预测、关键细节消融后的关系识别”。

第五步，至少两个独立读者组，每组不少于 3 人。读者不知道原评分，也不知道哪篇是目标文本。每项保持任务按预注册答案或一致性规则计 0/1。若任务本身允许多个合理答案，应在正式实验前用一小批文章校准，不允许看到结果后再移动正确答案。

第六步，计算 CFR 保持率。目标作文在一项任务上的平均表现若比匹配组高至少一个预注册最小差（课堂探索可设 10 个百分点，正式研究应通过功效分析确定），记“通过”；三项中 ≥ 2 项通过且两读者组方向一致，CFR 阳性。若三项结果分裂或读者组方向相反，记“不确定”，禁止硬判。

第七步，记录依赖。把 AI 改写、教师逐句修改、固定模板、实时自动评分等支持作为 D 变量，而不是道德标签。可通过“撤支架变体”观察功能是否保持。CFR 高但高度依赖支架的作品可以被称为“高功能协作文本”，却不能自动推断学生已经独立形成同等写作能力。

第八步，保留失败与豁免。无法匹配、读者分歧过大、题目不适合迁移、保持任务与评分表实际上同义、文本太短导致任务无法执行，都要进入退出日志。只保留可顺利计算的样本，会像保护生物学漏掉未调查单元、软件安全漏掉豁免模块一样，把看不见的问题从分母里删除。

十四、课堂版不需要统计模型：教师只要把“教会判准”之后再多问四个问题

研究读数可以复杂，课堂不能把每篇作文都变成实验。CFR 对一线教学最有价值的部分，是把修改和评价的注意力从“再满足一个指标”转向“指标已经满足以后还有没有功能”。教师可以保留原有量表，但在讲评或二稿环节加入四个问题。

第一问：**如果把这篇文章最容易得分的部分换成同样合格的版本，文章还剩下什么不可替代的作用？**例如把漂亮开头换成普通开头、把三个高级词换成常用词、把标准升华句删掉。如果核心判断、读者理解或新情境迁移几乎不变，说明这些高分迹象更像装饰信号。

第二问：**换一个相邻问题，这篇文章建立的东西还能工作吗？**议论文不只是“这个题目有观点”，而是理由能否处理一个边界案例；说明文不只是“把原文说明白”，而是读者能否用文章给出的机制解释一个新例子；叙事文不只是“读懂发生了什么”，而是人物关系和选择逻辑能否支持对新场景的合理预测。

第三问：**同分作文之间，谁让读者完成了更多原量表没有直接要求的任务？**教师可以随机挑两篇总分相近的作文，遮住作者和分数，让同伴读者完成一个迁移问题。学生会第一次看到：评分表相同不等于文本效力相同。这比不断告诉学生“要有深度”更可操作，因为“深度”被转成具体的保持性任务。

第四问：**这次表现有多少是外部支架代做的？**这不是为了惩罚 AI 或教师帮助，而是为了决定下一步教学。如果文章在 AI 整理后 CFR 很高，说明人机协作产出了好文本；若目标是训练学生独立能力，就需要安排下一次在较少支架下重跑保持任务。如果文章在撤支架后仍保持功能，才说明能力开始转入学生自身。

这样处理后，评分量表不必被废除。量表负责最低规范、共同语言和可解释反馈；CFR 负责量表被教会之后的二阶校验。前者告诉学生“哪些东西值得学习”，后者防止教师把“学生终于学会了老师教的评分特征”误认为“写作的所有重要东西都已经出现”。

十五、作文教学的五个改写：从“把标准教满”到“教会后仍要留出未知测试”

第一，范文教学要从“找得分点”改成“做消融”。过去读范文常让学生圈好词好句、概括结构、总结开头结尾方法。CFR框架要求再做一步：删掉或替换一个显性得分点，看文章的关键功能是否下降。若一个被反复赞美的排比句删除后对理解、迁移、判断毫无影响，它可以是审美增益，却不应被误教成所有好作文的必要成分。

第二，修改教学要增加“保持性修改”。一稿按公开量表修改后，二稿不要立刻再按同一量表打分，而给一个相邻任务。比如议论文写“竞争是否促进成长”后，不再只问论点是否鲜明，而给一个“合作与竞争同时存在”的案例，让作者说明原理由在何处失效。能处理边界，说明论证结构真的形成；只能重复原中心句，说明高分可能停在题面。

第三，评价要保留“轮换测试槽”。教师可以公开告诉学生：课程会检查迁移、反条件、机制理解或叙事关系，但每次不使用完全相同的具体问题。这样既不制造秘密标准，又避免把单一保持任务本身迅速教成模板。软件安全的经验在这里非常重要：防御不是永远藏规则，而是不断把攻击面、目标集合与绕过路径重新纳入检查。

第四，AI辅助写作要从“允许/禁止”二元争论改成“支架依赖账本”。如果任务允许AI，教师记录AI承担了哪一层：想法生成、证据寻找、结构整理、语言润色、反例测试。最终文本可以很好，但学习者能力判断要通过撤支架后的保持任务另测。文本质量与作者能力是两个对象，不能用一把尺子同时裁决。

第五，作文评分研究要从“更高一致率”推进到“回写后校准”。一个自动评分器与教师相关达到很高，并不意味着特征权重永远稳定。只要生成环境发生变化——学生大规模使用某类模板、AI工具或特定训练课程——就应该重新估计显性特征与保持性功能之间的关系。未来自动作文评价需要版本化：不仅记录模型版本，也记录写作者生态版本。

十六、AI时代的新典型：为什么“把所有显性标准做到极致”反而使“好”更难判

在没有生成式AI时，很多评分特征之所以有用，是因为它们昂贵。一个学生能够持续保持语法准确、段落连贯、词汇多样、论证结构完整，通常意味

着他投入了相当的写作训练；这些表面特征因此可以作为深层能力的代理。AI 改变的不是“语言标准突然不重要”，而是生产成本结构：过去昂贵的迹象变得廉价，代理变量的证据权重就必须更新。

这与公司金融 S 位高度同构。一个公开的市场异象在被普遍认识以前可以带来超额收益；被普遍认识以后，市场参与者围绕它行动，收益空间被吸收。作文中的“高分迹象”被 AI 普及后，也可能从稀缺能力信号变成基础设施输出。教师若仍按旧权重解释，会把“工具能够稳定提供”误读为“学习者已独立形成”。

但 AI 也使 CFR 更容易实施。过去教师很难为每篇文章生成多个题目变体、反条件问题和读者任务；现在 AI 可以根据同一作文快速生成相邻案例、反例、概念辨别题，再由教师审核。也就是说，AI 既加速了判准回写，也降低了保持性评价的成本。真正的教育创新不应是与 AI 比赛谁更会识别“机器味”，而应重写评价单元：从一篇成品的显性特征，扩展到文本在变体任务中的功能保持。

这也解释一个常见误区：为了防 AI，有人主张不断增加“个性化”“独特经历”“个人声音”等评分维度。这些维度一旦公开，同样可以被提示工程优化；AI 可以制造拟真的个人细节、口语停顿与不完美句式。任何固定可见特征都有被学习的可能。CFR 因此不是最后一条永不失效的新标准，而是一种评价制度的自更新原则：**凡已成为生产目标的特征，都要降低其未经校验的证据权重，并用未参与该轮优化的功能任务重新验证。**

因此，AI 时代“好作文”的稀缺性从“能不能做出漂亮文本”转向“漂亮文本之外，是否还有能够在新条件中继续工作的结构”。这不必神秘化成人灵魂。AI 文本也可以拥有这样的结构；人类文本也可以没有。判决点落在作品的功能保持，而不是作者身份。

十七、反向约束：CFR 理论至少必须削弱自己三次

一个新理论如果只能扩大适用范围、不能说自己在哪里不该用，通常只是概念扩张。本文给出三条反向约束。

第一，CFR 不能取消透明评分。对高风险考试、未成年人课堂和存在权力差异的评价，任务要求、基本评分维度、学术诚信规则必须尽可能公开。保持性任务可以轮换，但“将存在迁移与保持检查”应被告知。若教师把 CFR 变成“我心里还有一套不告诉你的标准”，它违反公平边界，理论应用失败。

第二，CFR 不能把所有套路化都判坏。某些约定俗成的结构是语言共同体降低沟通成本的基础设施。新闻、学术论文、考试议论都有稳定格式。只要格式没有替代核心功能，使用模板不降低 CFR。相反，一个故意反模板、极端陌生的文本若让读者无法完成任务，也不能凭“原创”获得余量。本文的价值排序因此是：**功能保持高于形式新奇，任务合法高于反规则姿态。**

第三，CFR 不能直接等同学生能力。高 CFR 文本可能由多人协作、教师深度指导或 AI 共同完成。文本“好”与作者“会写”必须分开。若研究目的是能力评价，必须另加支架依赖与独立重测；若目的是作品评价，则无需追问每个词由谁敲出。这个区分尤其重要，否则理论会把关于作品质量的发现越界成对人的道德判断。

第四，CFR 不适合所有写作目的。纯私人日记、艺术实验、无意面向公共读者的自由写作，未必存在可统一的任务功能；某些文学文本的价值还可能通过长期历史阅读才出现。本文主要适用于语文教育中具有可陈述任务、可比较文本和教学评价目的的作文，不把文学价值整体纳入。

第五，CFR 并不保证评价成本更低。相反，它增加读者任务、匹配与变体设计，研究成本高于一次打分。因此课堂版只建议抽样使用：单元末、代表作、量表已训练成熟以后，用少量保持任务校验，不要求每篇作文都全套运行。若把 CFR 变成新的高频考核表，它自己也会进入回写并快速被套路化。

十八、伦理边界：任何“余量”指标一旦拿来考核学生，就会遭遇它自己批判的问题

CFR 具有自反危险：它一旦被学校写进正式评分，学生就会开始针对 CFR 训练，新的保持任务也会变成显性目标。理论不能假装自己免疫于回写。因此本文提出三条伦理条款。

其一，不把单次 CFR 结果用于高风险惩罚。它首先是研究与形成性教学工具，用于发现量表可能漏掉的质量，不适合作为一次考试中决定升学的秘密加分项。其二，任何用于学生评价的保持任务必须经过可理解性与群体公平检查，尤其避免把特定文化经验、家庭资源或阅读背景当作“迁移能力”。其三，必须给学生申诉与展示替代证据的通道；若某一保持任务误读了文章，学生可以通过解释、另一个变体任务或写作过程证据进行复核。

更深的伦理要求是定期让指标“退役”。当某个保持性任务在一个学校已经被反复练到成为固定套路，它就不再适合承担未训练测试的角色，应被移入显性教学库，并开发新的变体。这里的“隐藏”不是永久保密，而是时间上

的轮换：今天的未知测试，明天可以成为公开课程；课程进步以后，再创造下一轮迁移情境。评价系统因此从静态量表变成动态校准机制。

十九、自反检验：用 CFR 来评价本文自己

如果本文主张成立，它不能只靠“二阶碰撞、SDE、回写余量、控制流”等新词显得创新。把这些显性术语删掉，论文至少还要完成三项保持性任务。

第一项是解释一个旧量表解释困难的反例：为什么两篇内容、组织、语言分数相近的作文仍可能有稳定质量差异？CFR 给出的答案是，显性分数匹配后，两篇文本在未训练的功能任务中可能产生不同保持率。第二项是解释生成环境改变：为什么 AI 出现后，同一套语言特征的评分权重可能需要重估？CFR 通过“判准已进入生成回路”解释特征—深层质量关系变化。第三项是产生一个可被否证的研究设计：如果匹配显性分数后，CFR 保持任务无法提供任何稳定增量，或者所谓增量完全可以被已有评分维度重新预测，本文核心对象就没有独立必要。

按本文自己的 2×2 ，当前论文最多只能被放在“概念合法、CFR 待实证”的位置。它已经给出显性定义、操作阈值和公开资料反证跑，但还没有完成大规模中文学生作文的前瞻性实验。因此不能把“理论可测”写成“理论已被验证”。这是一条有意保留的不利结论，也是后续研究的入口。

二十、可证伪命题与写死赌注

本文提出四条可证伪命题。

命题 1：在显性评分维度匹配的作文对中，至少存在一部分文本在预注册保持性任务上表现稳定不同；若大样本中这些差异接近随机，CFR 作为独立对象失败。

命题 2：随着某一评价特征被更明确地公开教学和自动反馈，该特征与保持性功能之间的零阶相关可能上升、稳定或下降，但其“控制训练强度后的增量预测力”不会被保证恒定；若跨多年、跨不同生成工具和不同暴露强度都保持完全稳定，回写假设被显著削弱。

命题 3：单纯提高显性量表得分的干预，与提高 CFR 的干预可以分离。若任何能提高量表得分的教学都同比提高保持性任务，且从无高显性低 CFR 样本，本文 2×2 中的“合规拟态”格将被经验上清空。

命题4：在允许 AI 辅助的情境中，AI 参与比例不能单独预测 CFR 方向；若在控制题目、能力和使用方式后，AI 比例几乎决定 CFR 高低，那么本文“身份不作为第一判据”的边界需要重写。

写死赌注如下：**到2028年8月31日前**，若能获得至少三个学校、每校不少于100篇中文中学生议论文的匿名数据，使用统一显性量表并预注册三项保持性任务，本文预计会观察到至少10%的“高显性达成—CFR 阴性”样本；并且在解释独立读者迁移表现时，加入 CFR 变量应在交叉验证中提供超出显性评分维度的增量预测。什么不算命中：事后改变“高显性”的阈值、只挑 AI 文本、只用教师主观印象充当保持任务、把原量表换个说法重新问一遍，都不算支持本文。

二十一、中文语文教育可以立即启动的三项验证议程

第一项是“判准暴露纵向实验”。研究不需要故意降低任何学生的教学质量，而可以利用不同班级已经存在的评分表公开程度、范文训练强度与 AI 辅助方式差异，连续收集同一批学生三个时间点的作文。每次既按学校现行量表评分，又保留若干不写进评分表的保持性任务。若显性指标随训练稳定上升，而保持性功能先升后平、甚至下降，就说明我们确实观察到了“判准回写”与“功能增长”两条并不重合的曲线。反之，如果两者始终高度同步，CFR 作为独立对象的必要性会明显下降。这项研究的关键不是比较哪个班“更会考试”，而是直接追踪同一套判准进入生成回路后，指标的诊断含义是否发生漂移。

第二项是“等分匹配压力测试”。从同一题目中抽取一批在现行评分表上得分接近的文章，最好同时包含学生独立写作、教师深度修改、模板训练和合规使用生成式 AI 辅助的不同来源，但在盲评时隐藏来源。随后给独立读者实施本文预注册的三个保持性任务，并做关键段落或细节消融。若等分文章在 CFR 上出现稳定分层，尤其出现大量“高显性达成—低 CFR”的合规拟态格，就说明总分压缩掉了一个具有教育意义的差异；如果等分后差异极小，则本文的核心主张受到直接削弱。这里必须强调，来源不是判决依据：AI 文本、人工文本、模板文本都只能按同一套 CFR 程序接受检验，不能把技术身份偷偷写成质量标签。

第三项是“判准轮换与退休实验”。很多作文教学的默认做法是：发现旧评分点被套路化以后，再加一个新评分点，例如从“细节”加到“真情实感”，从“立意”加到“思辨”，最后评分表越来越长。CFR 提出另一种可

能：与其不断增项，不如把部分已经高度可优化的判准暂时退出显性训练，在保持最低任务边界的前提下，轮换未知保持性任务，观察学生能否把同一篇文章的功能迁移到新情境。如果一个班只有在老师预告“这次考反例”时才会写反例，而换成未预告的新案例适用判断就迅速失效，那么教学得到的可能是新的判准顺从，而非更稳健的写作功能。相反，如果经过一段训练以后，学生对未预告任务也越来越稳定，才更接近本文所谓的正向余量。

三项研究共同构成一条故意对本文不友好的验证路线：第一项问余量会不会随着判准暴露而与显性分数脱钩，第二项问同分文本能不能被余量重新分开，第三项问不再增加评分点时功能能不能迁移。如果三项都失败，最合理的结论不是继续发明更复杂的 CFR 版本，而是承认现有作文评分已经足以承担“好”的主要判决；只有当这些实验反复发现稳定、可复制、跨学校仍存在的剩余差异时，“判准回写余量”才有资格从一篇概念论文成长为语文教育的有效研究对象。

二十二、结论：好作文不是“标准的完成品”，而是标准被学会以后仍然有东西在工作

本文从一个看似朴素的问题出发：何谓一篇作文的“好”？传统回答不断增加正向属性，却较少处理一个二阶事实：评价标准会被教会，被练习，被模板化，被自动反馈，被生成式 AI 直接优化。一旦如此，“符合标准”与“标准原来试图指向的质量”之间就不能默认保持同一关系。

公司金融与治理让我们看到，公开指标会进入参与者学习并改变后续关联；保护生物学让我们看到，避免灭绝只是底线，恢复还必须包含分布、功能和依赖；软件安全让我们看到，正常运行与鲁棒安全之间必须隔着威胁模型、合法集合和绕过测试。三家并非分别给作文增加一条比喻，而是共同逼出一个新的评价单位：判准回写以后仍保留的功能余量。

因此，本文的核心定义可以压缩成一句话：一篇作文的“好”，是在满足最低任务合法边界之后，当显性评分判准已经被公开并反向塑造写作时，文本在未参与该轮优化的变体任务与独立读者中仍能稳定保持的文体功能余量。本文称之为“判准回写余量”（CFR）。

这个定义不会让教师放弃评分表，恰恰要求评分表更诚实：哪些是底线，哪些是当前有效的代理指标，哪些已经因为长期训练而需要降权，哪些功能必须去新的情境里验证。它也不会把 AI 排除在写作之外，而是把问题从“谁敲了这些字”推进到“这些字在标准被学会以后还完成了什么”。

如果这个理论最终成立，作文教育评价的核心动作将发生一次变化：从“把好拆成更多可教的项目”，转向“每当一个项目被成功教会，就主动检查它是否还在指向原来的东西”。真正成熟的评价制度不是拥有一张永远正确的好作文清单，而是拥有更新清单的能力。好作文也不再是“标准的完成品”，而是标准已经成为环境的一部分之后，仍然有结构、有功能、有余量在继续工作。

附录 A 3.9 版工序审计与可复核账本

A1 闸零检索词与未吸收度记录

原	检索词（节选）	检索结果	判定
公司金融与治理	"corporate governance" + "writing assessment"; governance index + essay rubric; 治理指数 + 作文评价	未检测到治理指标 “被学习后关联消失” 作为作文评价理论的系统引入	干净
保护生物学与生物多样性	"IUCN Green Status" + "writing assessment"; species recovery + essay; 绿色状态 + 作文评价	未检测到 “避免灭绝 ≠ 功能恢复/保护依赖” 作为作文评价对象论的系统引入	干净
软件安全与隐私工程	"control-flow integrity" + "writing assessment"; CFI + essay rubric; 控制流完整性 + 作文评价	未检测到合法目标集合/威胁模型/豁免机制被系统用于作文质量定义	干净
新概念命名	"判准回写余量"; "criterion-feedback residual"; "criterion feedback residual" writing	未检测到同名概念；检测到大量 washback/Goodhart 近邻，已在正文划界	名称暂过

A2 显性判准达成度 × 判准回写余量：2×2 辨别格

象限	名称	判决
低显性 / 低 CFR	基础失效	既未满足常规要求，也没有保持性功能
高显性 / 低 CFR	合规拟态	显性高分，但未训练功能不优于匹配基线
低显性 / 高 CFR	异格有效	量表可能漏测真实功能；仍受任务合法边界约束
高显性 / 高 CFR	稳健好文	显性合规与保持性功能同时成立

A3 三条因果脊 × 三条因果脊的 27 项碰撞账本

下表不把所有碰撞结果都提升为概念。其作用是显示新对象不是直接从任一家搬运，而是在重复交叉中由多个中间对象收敛而来。至少八个有效涌现物包括：合格—好二分、判准依赖、任务合法集合、合规拟态、回写后二阶评价、撤支架保持、文体功能模块、双层判准制、自反退役机制，最终收敛到“判准回写余量”。

脊 A	脊 B	碰撞判断	涌现物
1 治理指标原有效	D1 存活 ≠ 恢复	高分可是真的，但“高分”只证明避免部分失败，不证明正功能	合格—好二分
1 治理指标原有效	D2 功能进入恢复	原先代理指标要由更接近目标功能的读数校验	功能保持测试
1 治理指标原有效	D3 保护依赖	高分需区分自主持有与支架维持	判准依赖账本
1 治理指标原有效	E1 合法目标集合	评价先划最低合法边界，再谈质量	任务合法集合 L
1 治理指标原有效	E2 威胁模型	评分有效性必须注明生成环境与读者条件	评价环境版本
1 治理指标原有效	E3 豁免/绕过	高分系统中的异常路径必须进入分母	退出与豁免日志
2 发表后关联消失	D1 存活 ≠ 恢复	判准越被优化，越要从“没错”转向“有功能”	回写后二阶评价

脊 A	脊 B	碰撞判断	涌现物
2 发表后关联消失	D2 功能进入恢复	原指标被吃掉后，用末训练功能任务恢复诊断力	CFR 锥形
2 发表后关联消失	D3 保护依赖	训练成功可同时增加分数与依赖，需分开记账	高分依赖态
2 发表后关联消失	E1 合法目标集合	判准被学会不等于废除规则；底线规则仍公开保留	反反指标主义
2 发表后关联消失	E2 威胁模型	指标意义变化要在新生成环境中重校	生态版本校准
2 发表后关联消失	E3 豁免/绕过	模板与 AI 可能走规则允许却未想过的高分路径	合规拟态
3 采纳动作回写	D1 存活 ≠ 恢复	教学判准是干预，达标只是干预后的表型	判准回写
3 采纳动作回写	D2 功能进入恢复	回写之后必须测正功能剩余，而非再加一个表面指标	判准回写余量 Z
3 采纳动作回写	D3 保护依赖	撤掉提示/模板后的功能保持成为关键	撤支架保持
3 采纳动作回写	E1 合法目标集合	公开最低边界与动态诊断标准可同时存在	双层判准制
3 采纳动作回写	E2 威胁模型	同一文本的质量证据要随评价生态注明条件	条件化好
3 采纳动作回写	E3 豁免/绕过	任何新 CFR 指标被制度化后也会被优化，必须轮换	自反退役机制
1 存活 ≠ 恢复	E1 合法目标集合	底线合规是必要条件，但安全/质量不能停在底线	最低合法+正功能
1 存活 ≠ 恢复	E2 威胁模型	“没出错”只在当前条件成立，换条件需重测	保持性压力测试
1 存活 ≠ 恢复	E3 豁免/绕过	均值良好可能掩盖关键失败路径	尾部失败入账
2 功能进入恢复	E1 合法目标集合	功能必须在允许任务空间内实现，反规则不等于好	异格有效边界
2 功能进入恢复	E2 威胁模型	不同文体功能需要不同保持任务，不能一张总表	文体功能模块
2 功能进入恢复	E3 豁免/绕过	功能评价需主动寻找“看似达标但可绕过”的文本	敌意样本
3 保护依赖	E1 合法目标集合	支架允许与否属于任务规则，不能事后道德化	AI/支架合规分离
3 保护依赖	E2 威胁模型	撤支架是能力测试条件，不是所有作品质量条件	作品/能力双对象
3 保护依赖	E3 豁免/绕过	长期依赖可藏在一次成功后，需连续版本追踪	回写依赖曲线

A4 五场景裁决表

场景	显性达成	CFR	判型	解释
独立写作，少量显性训练	高	高	稳健好文	高分与保持功能相互支持
模板强化后量表分上涨	高	低	合规拟态	记录“合规进步”，不把全部涨分解释为质量增长
AI 润色但观点/证据由学生形成	高	高或不变	可为好文	AI 身份不直接取消 CFR
AI 按量表从题目直接生成	高	低/未知	需保持测试	不能由表面完备自动判好
异格文本量表分较低但迁移强	低	高	异格有效	触发量表构念漏失检查，不自动改考试分

A5 近邻占位者 1:1 替换测试

近邻	它已经解决什么	为何不能 1:1 替换 CFR
分析性评分量表	把质量拆成维度、等级，强调透明与一致	通常不把判准暴露后的策略性生成作为对象变量

左邻	它已经解决什么	为何不能 1:1 替换 CFR
整体评分	以整体印象评定作品水平	无显性控制与保持任务，难分整体感与功能余量
构念效度/构念覆盖	问分数是否支持对写作能力的解释	中心对象是分数解释；CFR 中心对象是回写后的文本功能残差
Washback/后果效度	研究考试对教学与学习的反向影响	已看见“回写”，但通常不把回写后余量定义为好作文
Goodhart/Campbell	指标成为目标后可能被腐蚀	一般原理；缺少作文的正功能、合法边界、支架依赖与保持读数
过程评价/作品集	扩大证据到构思、修订和多篇作品	完整过程仍可高度围绕已知量表优化；对象不同
AES 对抗鲁棒性	测试评分器对异常、跑题、攻击文本的稳健性	关注评分器会不会被骗；CFR 关注高分文本还剩什么功能
修辞效果/读者反应	强调受众、目的与实际效果	未必控制判断暴露，也未必测显性得分之外的保持性增量
真实性/真情实感	强调经验、声音或主体真实	难操作，且虚构与协作写作可好；CFR 不以作者身份裁决
原创性/创造力	强调新颖、非模板	新颖本身可失效；CFR 以功能保持高于形式新奇

A6 CFR 最小编码表头（研究版）

字段	含义	编码	规则
text_id	文本匿名编号	字符串	唯一且不可反推学生身份
task_id / genre	题目与文体	分类	同类匹配
rubric_version	显性量表版本	版本号	必须可追溯
exposure_level	判断暴露强度	0-3	0 未见；1 见过；2 系统教学；3 实时反馈/直接优化
	最低任务合法集合	0/1	L=0 不进入“好”判定
1...Cn	显性评分维度	原量表分	不得事后重权
support_type	支架类型	多选	教师/模板/AI/同伴/无
H1,H2,H3	三项保持任务	0/1 或比例	任务事先注册
CFR_rate	保持率	0-1	通过任务数/有效任务数
CFR_class	余量判型	阳性/不确定/阴性	≥2/3 且双读者组同向=阳性
exit_reason	退出原因	文本	不可匹配/任务无效/读者分歧等

参考文献

[1] Abadi, M., Bui, M., Erlingsson, U., & Ligatti, J. (2005). Control-flow integrity. Proceedings of the 12th ACM Conference on Computer and Communications Security, 340–353.

[2] Akçakaya, H. R., et al. (2018). Quantifying species recovery and conservation success to develop an IUCN Green List of Species. Conservation Biology, 32(5), 1128–1138.
<https://doi.org/10.1111/cobi.13112>

[3] Bebchuk, L. A., Cohen, A., & Ferrell, A. (2009). What matters in corporate governance? Review of Financial Studies, 22(2), 783–827. <https://doi.org/10.1093/rfs/hhn099>

- [4] Bebchuk, L. A., Cohen, A., & Wang, C. C. Y. (2013). Learning and the disappearing association between governance and returns. *Journal of Financial Economics*, 108(2), 323–348.
<https://doi.org/10.1016/j.jfineco.2012.10.004>
- [5] Campbell, D. T. (1976). Assessing the impact of planned social change. The Public Affairs Center, Dartmouth College.
- [6] Diederich, P. B., French, J. W., & Carlton, S. T. (1961). Factors in judgments of writing ability. ETS Research Bulletin RB-61-15. Educational Testing Service.
- [7] Flor, M., & Cahill, A. (2025). Automated scoring of open-ended written responses: Possibilities and challenges. In *Innovative Digital-Based International Large-Scale Assessments* (pp. 265–279). Springer. https://doi.org/10.1007/978-3-031-90951-1_11
- [8] Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics*, 118(1), 107–156. <https://doi.org/10.1162/00335530360535162>
- [9] Grace, M. K., et al. (2021). Testing a global standard for quantifying species recovery and assessing conservation impact. *Conservation Biology*. <https://doi.org/10.1111/cobi.13756>
- [10] Mace, G. M., Collar, N. J., Gaston, K. J., et al. (2008). Quantification of extinction risk: IUCN’ s system for classifying threatened species. *Conservation Biology*, 22(6), 1424–1442.
<https://doi.org/10.1111/j.1523-1739.2008.01044.x>
- [11] Odendahl, N., & Deane, P. (2018). Assessing the writing process: A review of current practice (Research Memorandum RM-18-10). Educational Testing Service.
- [12] Powers, D. E., Burstein, J., Chodorow, M., Fowles, M. E., & Kukich, K. (2014). Comparing the validity of automated and human essay scoring. In *The Research Foundation for the GRE revised General Test*. Educational Testing Service.
- [13] Quinlan, T., Higgins, D., & Wolff, S. (2009). Evaluating the construct-coverage of the e-rater scoring engine (RR-09-01).

- Educational Testing Service. <https://doi.org/10.1002/j.2333-8504.2009.tb02158.x>
- [14] Snow, K. Z., Monroe, F., Davi, L., Dmitrienko, A., Liebchen, C., & Sadeghi, A.-R. (2013). Just-in-time code reuse: On the effectiveness of fine-grained address space layout randomization. 2013 IEEE Symposium on Security and Privacy, 574–588.
- [15] Xu, X., Ghaffarinia, M., Wang, W., Hamlen, K. W., & Lin, Z. (2019). CONFIRM: Evaluating compatibility and relevance of control-flow integrity protections for modern software. 28th USENIX Security Symposium.
- [16] Zhang, M., & Sekar, R. (2013). Control flow integrity for COTS binaries. 22nd USENIX Security Symposium, 337–352.
- [17] Zhong, Y., Hao, J., Fauss, M., Li, C., & Wang, Y. (2026). AI-generated essays: Characteristics and implications on automated scoring and academic integrity. Educational Measurement: Issues and Practice, 45(1), e70013. <https://doi.org/10.1111/emip.70013>
- [18] 付自文. (2026). 公司金融与治理·新思想前沿. SDE Universes.
- [19] 付自文. (2026). 保护生物学与生物多样性·新思想前沿. SDE Universes.
- [20] 付自文. (2026). 软件安全与隐私工程·新思想前沿. SDE Universes.
- [21] SDE 多学科通融创新法 (v3.9) . (2026-08-16). 三领域·SDE 二阶碰撞出典范文 (内部方法文件) .

编者补节 · 第八章

一 · 需要补的一位：后果效度

本章的核心对象是判准回写余量：一套显性评价判准被公开、教学并进入生成优化之后，控制作文对这些判准的达成水平，它在未用于训练与直接评分的保持性任务中仍能稳定产生的功能增量。

本章第十一节的八类占位者里，已经点到了 washback（反拨效应）、Goodhart 式指标异化、量规批评与自动作文评分，处理得相当利落；

第十二节的反证跑更把「过去的评价把尺子当纯外部观察者」这句过强表述当场删掉，改为「现有研究已经看见回写，但通常没有把回写后的功能余量设为作文之好的对象」。这是一次做得很实的自我收窄。

但有一位没有被点名，而它恰恰是这一整片地的持有者：

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.). New York: American Council on Education / Macmillan.

这一支把效度从「测量是否测到了它声称要测的东西」扩展到测量被使用之后的后果是否可接受：一次测量即便在传统意义上完全有效，它进入使用之后所引发的教学窄化、内容偏移与价值后果，仍然属于效度论证必须承担的部分。本章正文里「后果效度」四个字出现过，却没有落到具体出处，也没有进入第十一节的划界表。

编者认为这是本章最贵的一处缺席——因为本章的整个问题意识（判准一旦公开并进入生成回路，它就不再只是尺子）正是这一支立起来的。

分离线（本章可以站住，三条）。

1. 论证的对象不同。那一支论证的是一套测量工具是否可以被这样使用，它的结论落在测量方案上（这份量表在这个用途下效度是否成立）。本章论证的是一篇具体作文是否好，它的结论落在文本上。前者裁工具，后者裁作品。
2. 本章要求一个可计数的读数。那一支给的是效度论证的框架与应当纳入的证据类型，不给单一读数。本章给了 CFR 保持率与 CFR-z，并预注册了阈值（三项保持性任务通过两项且两个独立读者组方向一致记为阳性），还写明这个阈值「不是宇宙常数」而是第一轮方案的裁决阈值。
3. 本章允许「好」随评价生态变化。那一支追问的是某次测量使用是否正当，不主张作品质量的证据结构本身随环境改变。本章明确接受这一后果：同一篇文章放在判准被教会程度不同的教育生态里，其余量可以不同——并且把这一条当作三家碰撞逼出来的结论，而不是缺陷。

判定形式。取一批显性量表得分匹配的作文对，先按那一支的路径做一次完整的效度论证（这套量表在这个用途下是否成立、后果是否可接受），再按本章的路径测 CFR 保持率。若效度论证的结论能够无余地预测哪一篇的保持率更高，本章应并回那一支，CFR 只是后果效度在作文上的一个操作化；若存在一批作文，其量表在效度论证上完全成立而两篇之间仍有稳定的保持率差异，本章的对象独立成立。

二·与第七章共用的一位，与两章的分工

本章与第七章（结算余损）共缺后果效度这一位，编者在两章各补一次，此处补分工。

- 第七章问的是一次改进：本轮宣布成功之后，损失被挤到了哪一域，并由它夺取下一轮的终点。它的时间单位是一轮修改。

- 本章问的是一篇成品：判准被学会之后，这一篇还剩下多少不能由显性达成解释的功能。它的时间单位是一次评价。

两章合起来给出一个本卷未曾预料的接续：第七章说每一次结算都在往账外挤东西，本章说挤剩下的那一部分正是「好」的所在。若两章都成立，「好」就不是判准的完成度，而是判准回写之后的剩余——这条联合命题目前没有数据支持，编者只把它作为两章之间的一个可检验方向列出。

三·编者提示：本章最实的一处与最虚的一处

最实的一处是第十二节那次反证跑。它预注册了三条杀死条件，逐条去查公开资料，三条都没有把本章杀死，但三条各逼出一次收窄：删掉「过去评价把尺子当纯外部观察者」的过强表述；删掉「AI 作文只是表面漂亮」的简单说法，改为「AI 是一种使旧特征权重发生变化的生成环境，具体文本仍需逐篇判定」；把「保持性任务」从教师秘而不宣的真正标准，改为预先说明其存在、但不把具体答案教成模板的轮换迁移检查——这一处把本章与「隐藏判准」的公平性质疑正面切开了。

更值得记的是这一节末尾那句作者自评：

本文若把创新点放在「评分标准成为目标就会失真」，不可还原性这一维应当判低。

编者读遍本卷九章，这是唯一一处作者主动指出自己在哪一维应当被扣分的地方。它同时给出了自救路线——新对象必须依赖三家共同提供的额外结构：回写之后测什么正功能、怎样区分支架依赖、在什么压力场中保持、如何对同分文本计算余量。这四问才是本章真正的贡献所在，读者应当按这四问读它，而不是按「指标会失真」读它。

最虚的一处，本章自己也写明了：那次反证跑是文献反证跑，不是数据跑。第十九节的自反检验按本章自己的 2×2 把本文放在「概念合法、CFR 待实证」这一格，并写道：不能把「理论可测」写成「理论已被验证」。

编者把这一条抄进本卷终编的未复流清单第一条——本卷收入本章时没有为它补一次数据检验，而是把它按原样收入。按第九章的定义，这正是一笔未复流。

四·本章在本卷中的位置

本章是本卷九章里唯一一篇不问「如何发生」而问「什么算好」的。编者仍把它收进来，理由写在编三的编前语里：前两章（第六、第七章）问的是结算这个动作的账内与账外，本章问的是账本本身被作品改了多少。三章合起来，「权交回去的那一刻」才是完整的——交回给文本自己、交回给读者、交回给判准。

读者若要单独评估本章，应当知道它的题域与其余八章不同；若要按本卷的母题读它，它承担的是那一句的后半——每一次被判为成功的结算，都在无

声处收回决定权——的制度侧：判准一旦被学会，它就不再能证明它原本要证明的东西。

终编

权被收回的那一次

共同动作：无。这一编只有一章，它的动作是回头看前面八章。

第九章处理的是全书最后一个位置：一次写作结束之后。

它给出的判断是：一次写作真正的代价不落在这一次，而落在此后。当一次写作进入「推进已经停止而截止时间照常逼近」的状态并最终被全部交付时，同一类写作此后的通行能力被下调，而下调不进入任何一份记录——因为所有记录登记的都是交没交、写得怎样。

三个承重点缺一不可：拥塞、不可推迟的截止、全部交付。少了截止是延期；少了交付是缺交（那是有信号的失败，另有一套办法处理）；少了拥塞是从容通过。

这一章的第三个条件是被数据逼出来的。作者原本主张这是一个跨领域的通用结构，并用六个公开源码仓库四万余次改动做了预注册检验，三条假设全部未通过。事后诊断发现，在那个领域里改动规模与前置停滞相互独立——本文假定的那个形状（推进停止、压力累积、最后一次推过去）根本不存在。作者由此没有换阈值、没有换仓库，而是撤回了跨域主张，把「不可推迟的截止」写进机制，并把全部兑现改写成带前件的形式：若该领域存在不可推迟的截止，则应观察到……；不满足前件的领域一律标注为不适用。

本卷的自反，写在这里。

按第九章自己的三条件检本卷：本卷九章各自都有一道不可推迟的截止（一次成稿）、中间都有过拥塞（三家材料互不相容处）、并且都被全部交付。因此本卷九章各自都是一次硬通的产物。

本卷的未复流清单，至少有以下三条：

1. 第八章的判准回写余量，其反证跑是文献反证跑，不是数据跑。本卷没有为它补一次数据检验，而是把它按原样收入。
2. 第四章与第五章一个数都没有算出来。本卷没有为它们跑最小实验，而是把设计按原样收入。
3. 第七章三家里有两家与本卷别处重复。本卷没有为它换源重做，而是用编排位置代替了取源独立性。

三条都是「一条本来要走的路，因为这一次的截止而被绕开」。按第九章的定义，它们此刻正是本卷的未复流项。编者把它们写在这里，不是为了免责——按第九章第二个证伪条款，承认成功并不改变收缩。

第九章 通扣

——一次全部交上去的写作如何在记录之外扣掉此后的通行能力

Table of Contents

通扣

——一次全部交上去的写作，如何在记录之外扣掉此后的通行能力

摘要

作文教学把“写不出来”当作一次性的事件来处理：这一次没写好，下一次补上。本文提出一个不同的判断：一次写作真正的代价不落在这一次，而落在此后——当一次写作进入“推进已经停止而截止时间照常逼近”的状态并最终被全部交付时，同一类写作此后的通行能力被下调，而下调不进入任何一份记录，因为所有记录登记的都是交没交、写得怎样。本文把这种“全部通过而通行能力被扣”的事件称为通扣，把一次拥塞消散之后重新以不排队方式通过所需的时间称为复流期（R），并预测多数情形下 R 不收敛。材料取自三处互不相容的既有结论：发育生物学关于命运一旦决定即不可逆、而转决定的方向与可达命运集随培养时长改变的记录；交通工程关于瓶颈激活后排队排放率比激活前低约一成、且不随排队消散而恢复的观测；科学社会学关于多重发现是常态而承认只给一个的统计。判据是一句可以拿去问记录的问话：这一次的产出，集中在布置之后的哪一段时间里。本文用六个公开源码仓库的完整提交历史

（42,146 次文件级改动）预注册并实跑了一条跨域检验，三条假设全部未通过；事后诊断表明改动规模与停滞时长在该领域相互独立（ $\rho = -0.007$ ， $p = 0.16$ ），本文据此撤回了跨域主张，并把机制修订为：通扣需要一个第三方约束——不可推迟的截止；没有截止的场合，拥塞只表现为延期，不表现为通扣。这一修订是本文全部结论中唯一由数据逼出来的一条。

关键词：通扣；复流期；通行方式；截止约束；写作发生；作文评价

Title: Through-Deduction: How an Essay Fully Handed In Deducts Future Throughput Outside the Record

Abstract: Writing pedagogy treats “being unable to write” as a one-off event to be made up for next time. This paper argues otherwise: the real cost of an act of writing falls not on that act but on those following it. When an act of writing enters a state in which progress has halted while the deadline continues to approach, and is

nonetheless fully delivered, the throughput of that class of writing is subsequently reduced—and the reduction enters no record, because records register only whether the work was submitted and how well it scored. This paper names the event **through-deduction**, defines the **recovery period** (R) as the time required to resume passage without queueing, and predicts that R does not converge in most cases. Three mutually incompatible established findings supply the material: from developmental biology, that fate once determined is irreversible, yet the direction and reachable set of transdetermination shift with culture duration; from traffic engineering, that queue discharge rates run about ten percent below pre-activation flow and do not recover when the queue clears; from the sociology of science, that multiple independent discovery is the norm while recognition goes to one. A preregistered cross-domain test on the complete commit histories of six public source repositories (42,146 file-level changes) failed on all three hypotheses; post-hoc diagnosis showed change size and preceding idle time to be independent ($\rho = -0.007$, $p = 0.16$). The paper therefore withdraws its cross-domain claim and revises the mechanism: **through-deduction requires a third-party constraint—a non-deferrable deadline; absent one, congestion manifests only as delay.**

Keywords: through-deduction; recovery period; mode of passage; deadline constraint; genesis of writing; essay assessment

一、三份记录都显示正常

从三份真实存在、且都可以调出来看的记录说起。

第一份是一个班的交稿时间戳。某中学用作文平台交作业，平台记录每一篇的提交时刻。初二上学期，全班的提交时刻分布在布置后的第二天到第六天之间，形状接近均匀。到初三上学期，同一批学生，同一个老师，提交时刻高度集中在截止前六小时以内，且有若干名学生的提交时刻在截止前十分钟。分数没有下降。老师的评语没有变化。教务系统里，这两个学期的作文完成率都是百分之百。

第二份是一名学生两年的作文本。每一篇都写完了，每一篇都交了。翻到后面会发现一件事：题材范围在缩小。初一时他写过外婆的病、写过转学、写过

一次和父亲的争吵；初三只剩三类——校园里的小事、读了一本书的感想、一次考试失利后的振作。没有任何一次是他“不会写”，每一次都写满了，分数还略有上升。没有任何一份记录登记过“他不再写某一类题材”这件事，因为记录里只有写过的那些。

第三份是一份教研组的听课本。组里有一位老师连续三年上作文课，评价一直很好。第一年她带学生写“给十年后的自己”，第二年写“我家阳台上的一盆植物”，第三年还是这两个题目。没有人问过她为什么不换——她的课评分很高，学生的作文分数很稳。听课本上记的是“环节完整、指导到位”。

三份记录有一个共同的形状：每一次都通过了，而通行的方式变了。从容通过与勉强通过在记录上是同一件事；从写十类题材到只写三类，在记录上是同一件事；从三个题目里选到只剩两个题目，在记录上是同一件事。所有的账都是平的，被改变的那个量不在账上。

本文要问的就是这一件事：一次写作到底是怎样发生的，以及一次“全都交上去了”的写作，扣掉了什么。

问的不是“作文是什么”，也不是“为什么要写作文”。问的是一条路——从哪儿起、经过什么、到达什么，以及在哪一步还插得上手。答案因此不是一个定义，是一段次序加一个位置。

二、通扣

先把名字立出来，后面所有的论证都挂在它上面。

一次写作可以用两种方式通过。

第一种是不排队的通过。布置之后，推进是连续的：想到、写、改、交。中间可能慢，但没有出现“停在那里而时间继续走”的状态。

第二种是排队之后的通过。中间出现过一段推进停止的时间——不是在想，是卡住了；而截止时间照常逼近。最后它还是被交上去了，可能是在截止前一晚，可能是靠一段用过的结构撑起来，也可能是靠把题目往一个熟悉的方向拗过去。交上去了，分数正常，记录完整。

本文把第二种通过所留下的、不进入任何记录的减项，称为通扣。

定义写死：通扣，指一次写作在推进已经停止而截止时间继续逼近的状态下被全部交付之后，同一类写作此后的通行能力被下调，而这一下调不出现在任何一次交付的记录里，也把本文与所有“通过即完成”的记账方式分了开：一次通过在账上是一个正项，而本文主张它同时是一个未被登记的负项，两者共用同一条记录。

三个承重点，缺一个就不是通扣：

1. 全部交付——它不是失败。没交、迟交、白卷都不算通扣，那些是有信号的失败，另有一套办法处理。通扣的前提是所有的账都平。

2. 通行能力被下调——被扣掉的不是这一次的质量，是此后同一类任务“不排队就能过去”的最大流率。它表现为：同类题目下一次更早开始拖、更容易卡住、更快退回到熟悉的写法。
3. 不入账——没有任何一份现行记录登记通行方式，因此这个减项在制度上不可见。

需要立刻区分两个容易混进来的说法。

一个是拖延。拖延描述的是行为的时间分布；通扣描述的是能力的一次不可逆下调。一个人可以次次拖到最后一晚而通行能力毫无变化（他的通行方式一直是排队，且稳定），也可以从不拖延而在某一次之后发生通扣。二者可以在同一个人身上一个有一个没有。

另一个是畏难。畏难是一个关于情绪与意愿的判断；通扣是一个关于通行方式的判断，它不问这个人怕不怕，只问他这一次是怎么过去的。第十三节会给出一条能把两者分开的观察。

还要挡住第三个说法：生疏。一个人很久不写某类题材，自然会写得吃力，这是熟练度的自然衰减，与本文无关。两者的分别在于方向：生疏是先不写、后写不动；通扣是先写不动、后不写。判定不难——查一次时间顺序即可。若那一类题材的最后一次交付恰恰落在可用时间的末段（拥堵的标志），此后再无该类题材，那是通扣；若它的最后一次交付从容而此后长期无机会写，此后写得吃力，那是生疏。第一节那本作文本里的三类退出，全部属于前一种。

最后交代命名的来路。“扣”取的是账目上的减项，“通”取的是通过——这个词要说的正是那件古怪的事：减项恰恰发生在一次成功的通过里。交通工程对同一个结构有一个成熟的名字（容量跌落），本文第四节会正面处理它，也会在第十四节说明为什么本文不能直接沿用那个名字。

三、三家各自把成败结在哪一端

现行的解释不止一套，而是三套，且互不相容。它们分别在三个端点上结算“一次写作到底成没成”。

3.1 落点端：命运被决定的那一刻

发育生物学处理的是同一件事的生物版本：一团还没有形态的细胞，什么时候、以什么方式，变成一个再也回不去的东西。

这门学问的判据极硬，而且是操作性的：移植。把一块组织搬到另一个位置去，看它长成什么——长成新位置该长的，说明它还没被决定；仍然长成原位该长的，说明它已经被决定了。这个判据自 1924 年的组织者实验起沿用至今（Spemann & Mangold, 1924, *Archiv für mikroskopische Anatomie und Entwicklungsmechanik*, 100: 599–638）。

按这一套，一次写作的成败结算在落点上：这段文字有没有成为一个已定的东西。没成，就是还没发生；成了，就是发生了，此后只是好坏之别。

这一套自己掌握着一件材料，本文第四节要用到：转决定。已经被判定为“决定了”的成虫盘细胞，在成年宿主体内长期培养、反复移植之后，会以低频率跳变成另一种命运（Hadorn, 1963, *Developmental Biology*, 7: 617–629；Hadorn et al., 1968, *Revue Suisse de Zoologie*, 75: 557–569）。更要紧的是这种跳变不是随机的：它发生在盘上特定的位置，方向固定、频率固定，而且培养的时长会改变它此后能跳到哪里。

3.2 流量端：瓶颈被激活的那一次

交通工程处理的是流：一条路上能过多少、什么时候过不去、过不去之后会怎样。

这门学问在本题上有一句可以逐字引用的结论。Cassidy 与 Bertini 对多伦多两处高速瓶颈的观测表明：车辆从排队中排出的平均速率，可以比排队形成之前测得的流量低约一成（Cassidy & Bertini, 1999, *Transportation Research Part B*, 33(1): 25–42）。这个现象在这门学问里有专名——容量跌落，且它逼出了一个制度性的麻烦：“通行能力”必须分两个值，一个是瓶颈激活前的自由流容量，一个是激活后的排队排放率（Hall & Agyemang-Duah, 1991, *Transportation Research Record*, 1320: 91–98；Banks, 1991, 同卷: 234–241）。

按这一套，成败结算在流量上：决定结局的不是终点，是这条路上有没有形成排队。

3.3 承认端：算不算数由制度裁定

科学社会学处理的是第三件事：一样东西做出来了，凭什么算是“你做出来的”。

Merton 用大量案例证明：多重发现——同一个东西被多人独立做出——是科学中的常态而不是例外（Merton, 1961, *Proceedings of the American Philosophical Society*, 105(5): 470–486）；而承认制度只给一个人，优先权之争因此不是个别人的品行问题，是制度的结构性产物（Merton, 1957, *American Sociological Review*, 22(6): 635–659）。他后来把这件事的后果写成马太效应：已获承认者更容易再获承认，未获承认者的同等工作则被系统性地低估（Merton, 1968, *Science*, 159(3810): 56–63）。

按这一套，成败结算在承认上：做没做出来不重要，被不被算作数才重要。

3.4 三者不能同真

三条摆开，就能看出它们不是分工，是互斥。

发育说：看落点——有没有成为一个已定的东西。交通说：落点不是判据，所有的车都到了终点，被改变的是路本身。科学社会学说：这两家争的“成没

成”根本不在对象里，在承认制度里——同一份东西可以既算数又不算数，取决于谁先被记上。

三家无法靠“判谁对”化解，因为它们不在同一个语域里作答：一个谈对象的属性，一个谈通道的状态，一个谈制度的账目。要把三者放在一起，只能另造一个框架。

四、都通过了，而容量小了

三家争的是：一次写作的成败该在哪一端结清——落点端、流量端，还是承认端。

它们共同假定了一件谁也没有说出口的事：

每一次通过都被完整记账。通过了就是通过了；代价要么当场显现，要么不存在。

这句话念给三家听，三家都会说这还用说吗。发育生物学当然认为一个已经分化完成的结构就是一次完成；交通工程认为通行量的账逐辆记录、分毫不差；科学社会学认为承认制度虽然不公平，但它至少记下了它记下的那些。三家争的是在哪一端记账，没有一家怀疑过账能记完。

推翻这条共识的材料，来自交通工程自己，而且是它每天都在用的那一件。

容量跌落的形状是这样的：瓶颈被激活的那一段时间里，没有任何一辆车没有通过。排队形成了，又消散了，所有的车都到了终点，通行量的账逐辆记录、分毫不差。而在排队消散之后，这条路的通行能力比激活前低了约一成，且不随排队消散而回升。被改变的那个量，不是任何一次通过的属性；它不出现在任何一辆车的记录里。

这件材料的分量有三处：

第一，它不是异常，是常量。交通工程为它专门发展了对策（下一节会讲到匝道控制），并且被迫在规范里承认通行能力有两个值。一门学问为一件事修改自己的基本量纲，说明这件事是核心工艺，不是边角案例。

第二，它在自己的领域里回答的是另一个问题。容量跌落在交通工程里回答的是“标定通行能力时该取哪个数”，不是“作文如何发生”。正因为如此，把它拿来当作关于共有前提的证据，是这门学问自己没有走过的那一步。

第三，三家因此都退不回自己的阵地。落点端结不清，因为所有的车都到了终点而路变了；流量端结不清，因为流量的账是平的；承认端也结不清，因为没有任何一次通过被拒绝承认——恰恰相反，每一次都被承认了。

于是那条共识垮掉的方式很具体：代价既没有当场显现，也不是不存在；它落在一个没有字段的量上。

五、五段通行序列

有了通扣，就可以把一次写作从布置到交付之间的状态排成一段序列。这不是四个格子，是五个依次发生的段落，而且只有第五段是不可逆的。

第一段·自由流。需求低于通行能力。推进连续，交付时刻分布在整个可用时间里。这一段的特征不是写得快，是写得动：卡住的时候可以换个角度接着走。

第二段·临界。需求接近通行能力。推进仍然连续，但已经没有余量：一次分心、一次改题、一次多出来的作业，就会把它推过去。这一段在记录上与第一段没有任何差别。

第三段·拥塞形成。推进停止，而截止时间继续走。这一段是全过程的关键，也是唯一一段有内部信号的：写作者自己知道自己卡住了。但这个信号不出任何记录——没有一份作文记录有“这一次卡了多久”这一栏。

第四段·全部通过。交了。可能靠一段用过的结构，可能靠把题目拗向熟悉的方向，可能靠一晚上硬撑。产物合格，分数正常，账是平的。

第五段·通扣。此后同一类写作的通行能力被下调。表现为下一次更早进入第三段，或者干脆不再选这一类题材。这一段没有任何信号，也没有任何记录。

四种通过方式的对照如下。这张表是本文的辨别装置，它替代了“写得好／写得不好”这一根轴。

通过方式	走到第几段	当场看得见的	记录上的读数	有没有信号	此后的通行能力
从容通过	一、二	交付时刻分散，题材范围不缩	完成、分数正常	不需要	不变
勉强通过	一至四，但第三段极短	集中在后段交付	完成、分数正常	弱：只表现为赶	略降，可恢复
硬通	一至五	与勉强通过看不出差别	完成、分数正常	无	下调且回升
不通过	停在三	迟交、白卷、求助	缺交、低分	有：一眼可见	不变（未被扣）

这张表最要紧的一行是第四行：不通过反而不留下通扣。一个孩子说“这个我写不出来”并且真的没交，他的通行能力没有被扣——被扣的是那个卡住了、最后还是硬撑着交上去、并且拿到了正常分数的孩子。这条推论违反直觉，而它正是本文与所有把“完成率”当作正指标的做法之间的分歧所在。

现在指出本文要打的那一格：硬通。

它必须满足两条签名，且每条都要配一条误诊阻挡——因为这两条签名很容易被读成别的东西。

签名一：它在所有短期读数上与“勉强通过”完全同形。完成率、分数、字数、按时率，四个指标都无法把它们分开。> 误诊阻挡：若你能从分数或字

数上看出差别，那看到的就不是硬通，是质量下降。硬通的定义要求所有形式审查通过。

签名二：它的代价落在下一次的题材选择上，而不是下一次的分数上。被扣掉的是“这一类还敢不敢碰”，不是“这一次写得好不好”。> 误诊阻挡：若下一次的分数明显下降，那多半是别的原因（生病、家庭变故、教师更换）。硬通的特征恰恰是分数稳定而可选范围收缩——第一节那本作文本里的题材从十类降到三类，就是它。

两条签名之外还要说明：为什么这一格需要一个框架才看得见。若第二根轴设成“完成／未完成”，硬通整体落在“完成”一侧；若设成“分数高／分数低”，硬通分散在两侧且无规律。只有把轴换成“这一次是以哪种方式通过的”，它才从“完成”里被分出来。最后要回答一个实操上必然会被问到的问题：勉强通过与硬通在当场看不出差别，那怎么分？

答案是不当场分，事后分，而且只用两个已经存在的字段。第一个字段是交付时刻相对可用时间的位置——它区分的是“有没有走到第三段”。第二个字段是此后同类题材还写不写——它区分的是“有没有走到第五段”。两个字段合起来才构成判定：交付在末段而此后该类题材仍在，是勉强通过；交付在末段而此后该类题材退出，是硬通。

这个判定有一个必须承认的代价：它必然是滞后的。硬通只能在它发生之后的一到两个学期里被认出来，当场无法预警。本文因此不主张“识别硬通”是可行的干预；可行的干预在第十节，它不需要识别，只需要在拥塞形成之前降低入流。一个只能事后认出来的读数，用途是核算，不是调度。

六、复流期 R 与未复流清单

判据要能被清点，就得有一个数。本文不给比率，给一个时长——因为通扣要说的事情是“回不回得去”，而回不回得去是一个关于时间的问题。

复流期 R = 从一次拥塞消散之日起，到同一类写作重新以不排队方式通过所经历的时间。若在观测窗内始终未恢复， R 记为未复流 ($R = \infty$)，并进入未复流清单。

R 在三个领域里的量纲不同而算式相同：

- 交通工程：从排队完全消散起，到该瓶颈的排放率回到激活前流量所需的时间。观测结果通常是不回，因此这一栏在实践中多为未复流。
- 发育生物学：一块被判定为已决定的组织，从被移入新环境起，到它能够转向另一命运所需的培养时长。Hadorn 的材料表明这个时长是实在的量，而且培养越久，可达的命运集合越大。
- 写作：某一类题材或文体，从一次硬通起，到该写作者重新能够不排队地写它所经历的时间。

三处量纲不同，读数相同。这个读数要成立，必须具备四条性质，本文逐条兑现：

1. 可事后清点——不需要在写作现场观察，只需要交付时间戳与题材标注。
2. 带 ∞ 值——多数情形下它不收敛，而“不收敛”本身是最要紧的读数。一个只能取有限值的量会把最重要的那一类对象排除在外。
3. 把一个印象变成一份清单——“他现在不写这类了”这句话，此前只是印象；未复流清单把它变成一份可以逐条核对的名录。
4. 与既有主要指标正交——作文这个领域的既有主要指标是分数与完成率。R 与两者的关系必须是可分离的：一个完成率百分之百、分数稳定上升的学生，其未复流清单可以很长。第一节那本作文本正是这样的实例。

需要说明本文为什么不给一个比率。比率会把未复流清单里的条目压成一个分母上的数，而这份清单的价值恰恰在于它是逐条的：哪一类题材退出了，是可以指名的；而一旦压成“题材收缩率 0.7”这样的数，指名就没了，复核也没了。这个读数的粒度就是它的内容。

清点手册

读数名：复流期 符号：R 单位：天（或“未复流”）

定义：从一次拥塞消散（该次作品交付之日）起，到同一类写作重新以

不排队方式通过之日所经历的天数；观测窗内未发生则记未复流。

配套清单：未复流清单——逐条列出该写作者在观测窗内退出的题材类别。

粒度：

· 什么算“一类写作”：以题材类别记，不以文体记，也不以题目记。

分类表须在观测开始前写死并公布（例如：家庭／校园／自然／读物／

社会事件／自我经历／虚构，七类）。

恰在边界上的例子：写“外婆生病时我读的一本书”——同时落在家庭与

读物两类，按写作者事后自述的主意所在归一类，归不了的记为跨类，

跨类不计入任何一类的复流判定。

- 什么算"拥塞形成"：推进停止而截止时间继续逼近。
可操作的代理：交付时刻落在可用时间的最后 20% 以内，且该次的可用时间长度不短于同批其他作品。
- 什么算"不排队通过"：交付时刻落在可用时间的前 60% 以内。
- 观测窗：不短于两个学期。窗内未复流即记 ∞ ，不外推。

编码规则（五条）：

1. 分类表在观测前写死，中途不得增删类别。
2. 缺交、迟交不进入本读数（它们是有信号的失败，另案处理）。
3. 同一类别在窗内多次硬通，只记首次，复流期从首次算起。
4. 教师更换、转学、休学期间的作品整段剔除，并在清单上注明。
5. 未复流清单必须逐条可指名，不得汇总成一个比率对外发布。

表头：学生编号 | 类别 | 首次硬通日期 | 复流日期 | R（天／未复流） | 剔除说明

不适用：单次写作任务；无固定截止时间的写作；题材由他人指定且

不允许选择的场合（第十七节详述）。

三处引用一致：正文 §6 = 证伪 §16 = 赌注 §20 一致

一处必须交代的取舍：为什么不把三家的判据合成一个综合指标

三家各有一个成熟判据：移植（发育）、排放率（交通）、优先权归属（承认）。把三者加权合成一个“写作健康度”是很自然的做法，本文不做，理由有两条。

其一，三者的失败定义互不可译。发育里的失败是长成了不该长的东西，交通里的失败是通不过去，承认里的失败是做出来了而不算数。把三种失败折算成一个数，需要一套折算率，而这套折算率没有任何一家能提供——它只能由做合成的人自己定，于是这个数的含义就是那个人的偏好。

其二，也是更要紧的一条：合成会把本文最要紧的那件事抹掉。通扣的全部分量在于它发生在所有分项都达标的时候。任何加权合成都会让达标的分项抬高总分，于是硬通在合成指标上表现为优秀。一个把靶格算成优秀的指标，比没有指标更坏。

这条取舍也适用于本文自己的读数：复流期与未复流清单不得与分数、完成率合成任何总分，理由同上。

七、一句可以拿去问记录的问话

判断一格是否成立，最后要落到一句能拿去问记录、而不是拿去问人的判断的话。本文的这句是：

这一次的产出，集中在布置之后的哪一段时间里？

这句话里没有“应当”“真正”“充分”“恰当”这类词。它不问写得好不好，不问学生用不用功，不问老师尽不尽责。它只问一件在多数场合已经有数据的事——交付时刻相对可用时间的位置。

把它拿到六个场合去问，答案各不相同，其中两个是查过之后才知道的，且反过来改了本文。

场合一：中学作文平台。有交付时间戳，有布置时间，因此这句话可以直接查。答案是：数据早就有了，而没有任何一份评价表读它。平台把时间戳用于统计按时率，按时率只有两个值。第一节那个班从“均匀分布”变成“集中在最后六小时”的过程，全程留在数据库里，没有被任何人看过一眼。

场合二：高考考场作文。可用时间被压成一次，所有人的产出必然集中在同一段里。通行方式的区分在这个场合被制度取消了。答案是：本文的判断在考场上不适用——不是测不准，是那里根本没有第一段与第三段之分。这一处是查过才明白的，它把本文的适用范围削掉了一整类场合。

场合三：博士学位论文。有开题、预答辩、送审的节点记录，可以查出停滞段；而且这个领域已经有一个字段——延期。答案是：本文所说的东西在这里有一半已经被登记了。这一处对本文不利，第十四节会说明它削掉了什么。

场合四：企业周报。周期固定、通行能力被制度设为常量。出现拥塞时的处理不是延期，是降低质量交付，且这件事在多数单位有记录（返工次数、被打回次数）。答案是：这个场合已经把硬通制度化了，并且给它配了信号。

场合五：律师文书与医疗记录。截止时间由外部程序设定，不可推迟；而这些场合以先例与既有表述为正当。答案是：硬通在这里是常态，且不构成问题——因为通行能力本来就不要求扩张。

场合六：无人催的个人写作。没有截止时间。答案是：这个场合根本不会出现第三段。推进停止就是停止，不会有“而截止时间继续逼近”这件事。这一处是第九节那次实跑逼出来的，它是本文机制的关键限定，也是本文全部结论里唯一被数据改动过的一处。

八、六个仓库、四万二千次改动：一次没跑成的检验

证伪条款不许全是纸面的。本文在成文之前用另一类公开可复算的记录跑通了一条跨域检验——之所以不用作文语料，是因为公开的作文语料没有交付时间

戳，而本文的判据要的正是时间戳。结果是三条假设全部未通过。以下按预注册原样报告。

8.1 预注册

统计运行之前，数据、单元、事件定义、假设与阈值已写死存档，事后未作修改。

数据：六个公开源码仓库的完整提交历史，以 bare 方式克隆自 github.com：psf/requests、pallets/click、tqdm/tqdm、pallets/flask、pytest-dev/pytest、scrapy/scrapy。选择规则在跑数前写死：纯 Python 主仓、单一上游、提交数在 2×10^3 至 2×10^4 之间、历史完整公开；不按星标数、不按活跃度筛，也不在看过结果之后更换仓库。

代理声明：本次检验的不是写作，是“一次在停滞之后被硬推过去的通过是否留下不回升的减项”这一结构在另一个领域成不成立。提交历史与写作过程的对应是结构上的，不是内容上的。

事件定义：对某个 .py 文件的一次改动记为一次通过；规模 s 为该次增删行数之和，间隔 d 为距上一次涉及该文件的通过的天数。硬通事件 = $s \geq$ 该仓库全部改动规模的第 95 百分位，且 $d \geq$ 该仓库间隔分布的第 75 百分位。

三条假设：H1（跌落）事件后 20 次通过的硬通率高于事件前 20 次；
H2（不复流）事件后第二、第三个窗口仍不低于事前水平的 110%；H3（顺序）硬通率的上升早于返修率的上升。

8.2 结果

项目	读数
纳入的文件级改动	42,146 次（六仓合计，限 .py，文件改动次数 ≥ 20 ）
满足事件点前后各有三个窗口的文件	28 个
事前窗口硬通率（均值）	0.0018
事件后第一、二、三窗口（均值）	0.0054 / 0.0036 / 0.0054
H1（Wilcoxon 符号秩）	统计量 2.5， $p = 0.32$ ；中位差为零 $\rightarrow$ 不成立
H2	三个窗口的中位数全部为 0，比较无意义 $\rightarrow$ 视同不成立
H3（符号检验）	硬通先行 6 例，返修先行 8 例， $p = 0.79 \rightarrow$ 不成立

三条全部未通过。按预注册的处理办法，本文当场撤回跨域主张。

8.3 事后诊断：为什么事件率接近零

预注册不允许改阈值，但允许追问失败的来路。以下为明确标注的事后分析。

事件定义要求两件事同时成立：改动规模大，且此前停滞久。在全部 42,146 次改动中，两者的秩相关为 $\rho = -0.007$ ， $p = 0.16$ ——相互独立。大改动之

前的间隔中位数是 6.93 天，其余改动是 7.03 天，几乎一样。两条件同时成立的比例是 1.21%，而按独立性预期正是 1.26%。

也就是说：在源码这个领域里，大改动并不跟在长停滞之后。本文假定的那个形状——推进停止、压力累积、最后一次推过去——在这里根本不存在。事件率因此低到任何 20 次的窗口都测不出东西；28 个文件的样本不是“证据不足”，是这个设计从一开始就没有可测的对象。

8.4 这一条改了本文的什么

这是本文全部结论中唯一由数据逼出来的修订，写在这里，也写进第九节。

失败的来路不是统计量太小，是这个领域缺一样东西：不可推迟的截止。一个源码仓库里的改动，可以永远不合并、可以挂在分支上三年、可以由别人接手。推进停止在那里的表现是延期，不是硬通——因为没有任何力量逼着它在某个时刻必须通过。

于是本文做了两处修订：

1. 撤回跨域主张。本文不再声称通扣是一个跨领域的通用结构。第十一节的兑现全部改写为“若该领域存在不可推迟的截止，则应观察到……”这种带前件的形式，凡不满足前件的领域一律标注为不适用。
2. 把截止约束写进机制。通扣不是“拥塞 → 减项”两步，是三件事同时在场：拥塞、不可推迟的截止、以及全部交付。少任何一件，得到的都是别的东西：少了截止是延期；少了交付是缺交（有信号，不留通扣）；少了拥塞是从容通过。

这处修订让本文的适用范围显著变窄，也让它变得更难被既有说法吸收——第十三节会说明，正是“截止约束”这一件把本文与滞后效应、习得性无助、马太效应三位最危险的邻居分开了。

（数据与算法可复算：六仓地址如上，事件定义与阈值见 8.1，统计用 Wilcoxon 符号秩检验与符号检验。预注册文本先于统计写就并存档。）

九、硬通需要一个第三方

上一节那次失败值得单独占一节，因为它把机制补完整了。

先看没有截止会怎样。一个人自己在写一篇没人催的东西，写到一半卡住了。他会怎样？他会放下。三个月后可能回来，也可能不回来。放下不产生通扣——因为他没有在卡住的状态下把它交出去，那批还没成形的东西留在没成形的状态里。这正是第八节那六个仓库的情形：分支挂着，没人合并，通行能力不受影响。

再看有截止会怎样。同一个人，同样卡住，而后天必须交。他有三条路：交白卷（有信号，不留通扣）；求助或改题（相当于换一条通道，也不留通扣）；

硬撑着交上去。第三条路是唯一留下通扣的那一条，也是学校里最被鼓励的那一条——它保住了完成率，保住了分数，保住了教师的教学记录。

所以本文的机制有三个条件，缺一不可：

拥塞 + 不可推迟的截止 + 全部交付 → 通扣。

这三件事在中小学作文里天天同时在场，而且是被制度组织起来同时在场的：每周一篇是拥塞的来源，交作业的期限是截止的来源，完成率是交付的来源。没有任何一环是失职，三环都是尽职的表现。

三条件缺一时各得到什么，值得逐个说清，因为这三种情形在学校里都有对应物：

缺截止，得到的是积压：那几篇一直没写完，堆在那里，学生自己知道，老师也知道。它难看，但它不扣通行能力。缺交付，得到的是缺交：一次可见的失败，有信号，会被处理，同样不扣通行能力。缺拥塞，得到的是从容通过：正常写完，什么也没发生。

三种情形都比硬通“难看”，而三种都不留减项。本文最不受欢迎的一条推论就在这里：在现行记录下，最好看的那一种通过方式，是唯一留下减项的那一种。

这也解释了第五节那张表里最反直觉的一行：不通过反而不留通扣。因为它缺了“全部交付”这一件。一个允许学生说“这次我写不出来”的班级，在完成率上会比隔壁班难看，而它的未复流清单会比隔壁班短。两份记录里，现在只有前一份被登记。

顺带指出一个必须承认的后果：本文因此不能主张“完成率是个坏指标”。完成率保住的东西是实在的——它拦住了放弃、拦住了脱落、拦住了失学。本文能主张的只有一件：完成率与通行能力是两个量，而现在只有一个被记账。

十、那条路，与匝道

现在可以回答开头那个问题了：作文如何发生。答案是一段四步的次序，顺序不能换。

第一步，承认制度先动。什么被要求、什么被记录、什么算数，决定了哪一类写作构成常态需求流。若记录只登记完成率与分数，那么“写完并写像样”就是唯一被计量的产出，需求流会被组织成一条持续满负荷的流。

第二步，流量组织。需求持续接近通行能力，余量被耗尽。此时任何一个扰动都会把它推过临界，拥塞形成。

第三步，落点。拥塞在截止的逼迫下被强行消散：交了，合格了，命运判定为“完成”。

第四步，回写。这一次硬通留下通扣：同类写作的通行能力下调，可选题材范围收缩。而这一步不进任何账，于是下一轮的第二步更容易过临界，通扣累积。回写是这条路成为一个循环、而不是一支箭头的原由。

四步之后，插手的位置就清楚了，而且它不在人们习惯下手的两个地方。

不在第三步。在产物上下功夫——批改、评讲、重写、限时训练——是投入最大的一处。它作用在拥塞已经形成之后，而这一段的全部结果就是把拥塞更快地消散掉，也就是更快地完成一次硬通。限时训练尤其如此：它以缩短可用时间为手段，把从容通过的余量直接削掉。

不在第一步的总量端。减少作文次数是第二常见的处方。按本文的次序，减少总量确实降低需求流，但它同时减少了处在自由流段的写作次数；若减少的那些恰好是学生本来写得动的题材，通行能力不升反降。

在第二步，而且动作是逆直觉的。交通工程对同一个位置有一套成熟做法：匝道控制——在瓶颈被激活之前主动限制入流，宁可让一部分车在匝道上等。它的依据正是容量跌落：激活后的排放率低于激活前的容量，所以拦住一部分入流反而提高总通过量。这不是推想，是被实测和一次真实的自然实验双向验证过的：Cassidy 与 Rudjanakanoknad 的现场试验表明，对合流处的匝道限流可以提高该处的通过量（2005, *Transportation Research Part B*, 39(10): 896-913）；而明尼苏达在 2000 年关停全部匝道控制信号的那次“控制假期”，则从反方向给出了对照（Levinson & Zhang, 2006, *Transportation Research Part A*, 40(10): 810-828）。

落到作文上，动作是这一条：

在拥塞形成之前，就把这一次写作从“必须交”的队列里撤出来。

具体的操作有三种形态，共同点是主动撤出，不是延期：① 布置时即声明本次可以交“写不出来的理由”而不是作文，并且这一交付同样计入完成；② 在写作者报告卡住时立即换题或换体裁，且不追补原题；③ 每学期预留固定比例的“可撤出额度”，由学生自主使用，用完为止。

三者与“延期”的差别是硬的：延期保留了那道截止，只是把它往后推，第三段仍在继续；撤出取消了那道截止，第三段就此结束。按第九节的三条件，取消截止即取消通扣的成立条件。

需要说清它与两种现行做法的关系，因为它们看上去很像。

分层作业是按能力给不同难度的题目。它降低的是需求的高度，不改变通行方式：一个被分到简单题的学生照样可能卡住、照样必须交。按本文的次序，分层作业作用在第一步（降低需求流），有用，但它不取消那道截止，因此拥塞一旦形成，硬通照样发生。

弹性作业是允许延期。它改变的是截止的时刻，不改变截止的存在。第三段在延期期间继续走——推进仍然停止，时间仍然逼近，只是逼近得慢一点。按第九节的三条件，延期不取消通扣的成立条件，只推迟它。

撤出与两者的差别是硬的：它取消的是这一次的交付要求本身。那批还没成形的东西因此留在没成形的状态里，而不是被压成一个合格的产物交出去。第十六节的条款四要求把撤出组与延期组分开报告，正是为了检验这个差别是不是真的。

十一、十二处兑现（带前件）

第八节撤回了跨域主张，因此以下每一处兑现都写成带前件的形式：若该领域存在不可推迟的截止且要求全部交付，则应观察到……凡不满足前件的领域，本文明确标注不适用。这一改动让兑现的数量没有减少，可裁决性却提高了——前件本身就是一条可查的条件。

一、急诊科的病历书写。前件满足（交班时限、病历时限、必须完成）。预测：一次在信息不足下被硬性完成的病历之后，同一医师此后对同类不确定病例的书写会更快收敛到模板表述，且这一收敛早于差错率上升出现。

二、法院的审限。前件满足。预测：在审限压力下完成的疑难案件，其后同一合议庭对同类案件的说理长度下降、援引来源集中，而结案率与被发回率均无变化。

三、飞行签派。前件满足（放行时限）。预测：一次在边缘天气下被硬性放行的航班之后，该签派席位对同类边缘条件的判断会更快收敛到保守或激进的一侧，而不是保持在需要逐次判断的中间状态。

四、期刊的修回期限。前件满足。预测：在修回期限内被硬性补完的实验，其后同一课题组提出新假设的多样性下降，而发表数量不降。

五、财务结账。前件满足（法定报出期）。预测：一次在资料不全下被硬性结出的报表之后，该科目此后的处理方式固定为那一次的处理方式，且不再随资料完备程度变化。

六、软件的发布窗口。前件满足当且仅当发布日期不可推迟。预测：在硬性发布日之前被强行合入的功能，其模块此后接受结构性重构的概率下降。注意：第八节的实跑正是在此失败的——那六个仓库没有不可推迟的发布日，前件不满足，因此那次失败与本条并不矛盾，反而是同一条判断的两面。这一处必须原样写出来。

七、译稿的交付期。前件满足。预测：一次在期限内硬性交出的难译段落之后，同一译者对同类难点的处理会固定为那一次的处理，而读者评价不变。

八、学位论文。前件部分满足（有延期通道）。预测：使用过延期通道的作者，其通扣显著低于未使用延期而硬性交付者，尽管后者在记录上更好看。这一条可用现有的研究生管理数据直接回溯，也是本文最容易被检验的一条跨域预测。

九、竞赛类写作与命题作文。前件满足且极强。预测：连续参加命题竞赛的学生，其自选题材的范围收缩速度快于不参加者，而竞赛成绩不下降。

十、新闻的截稿。前件满足。预测：一次在截稿前被硬性写完的复杂报道之后，该记者对同类题材的结构选择收敛，而稿件通过率不变。

十一、发育生物学的培养时长（反向兑现）。前件不满足——那里没有截止。本文因此预测一个相反的结果：培养时长越长，可达的命运集合越大，而不是越小。Hadorn 的材料正是这个方向。这一处是本文的自我约束：若在无截止的领域里也观察到可达集合收缩，本文第九节的三条件说就错了。

十二、无人催的个人写作（反向兑现）。前件不满足。预测：长期无截止的写作者，其题材范围不收缩，即便他长期卡在某几篇上写不完。若观察到无截止者的题材范围同样收缩，则本文的截止约束条款为伪。这条可用公开的博客与个人网站的长期存档直接检验。

十二处里有五处是顺序预测（谁早于谁），两处是反向兑现（预测本文自己不适用）。反向兑现比正向兑现更值钱：一个只会往一个方向长的判断，多半是描述而不是机制。再补一处本文自己也没想到的读法。第一节那三份记录里，第三份（同一位老师三年只用两个题目）说的其实不是学生，是教师的通扣。教师同样有拥塞（新题目要重新备课）、同样有不可推迟的截止（下周就要上这节课）、同样必须全部交付（课得上完）。按本文的三条件，教师的可选题目范围也会收缩，而听课本上记的是“环节完整、指导到位”——完全是硬通的签名。

这一读法把本文的适用对象从学生扩到了教师，也让第十七节的伦理条款变得更要紧：未复流清单一旦被拿去考核教师，最省事的达标办法就是多报几个题目名称，而不是真的换题。本文没有能力阻止这件事，只能把它写在这里。

十二、三处让步

三处被材料反过来削掉的地方，逐处写清削掉了什么。

第一处：完成率保住的东西是实在的，本文不能反对它。第九节已经指出，允许“这次写不出来”的班级在完成率上会难看。而完成率拦住的是放弃、脱落与失学，这些后果比通扣严重得多。若没有这一处，本文原本会写下的更强的那句话是：“完成率是一个坏指标，应当取消。”现在这句话删掉了。本文只能主张：完成率与通行能力是两个量，现在只有一个被记账；在两者冲突时该保哪一个，本文不提供答案。这一处动的是价值排序。

第二处：本文所说的东西，在一个领域里已经有字段了。第七节场合三是查过才知道的：学位论文管理里有“延期”这个字段，而延期正是本文所主张的“撤出”的一种制度化形态。若没有这一处，本文原本会写下的更强的那句话是：“没有任何制度登记通行方式。”现在只能写成：在有延期通道的场合，通行方式已经被部分登记；在中小学作文里没有。这一处削掉了本文第三层主张的普遍性，把它限定在基础教育阶段。

第三处：撤出这个动作本身可能被制度吞掉。第十节给出的三种操作，都要求由写作者主动使用，而组织学早已指出，任何自主额度在有排名的环境里会迅速变成新的评价项——用得更多的人被读成能力弱。若没有这一处，本文原本会写下的更强的那句话是：“给学生一个撤出额度，问题即可缓解。”现在只能写成：撤出额度只有在使用记录不进入任何比较的前提下才有效；一旦它可以被横向比较，它就变成另一道截止。这一处动的是干预的可行条件，也直接连着第十八节的反噬。

十三、与最近的几种说法分界

本文的判断与下面十种既有说法最容易被混为一谈。逐条给出各自说到哪一步、分界落在哪里，以及一条能把两者分开的观察。

其一，滞后效应 (Blanchard & Summers, 1986, *NBER*

Macroeconomics Annual, 1: 15–78)。这是十位里最危险的一位。它说到哪一步：一次衰退过后，失业率不回到衰退前的水平，均衡本身被那次冲击改变了。分离线：滞后效应把减项归给冲击的强度与持续时间，因此它预测足够长的扩张期可以修复；本文把减项归给在截止逼迫下的一次全部交付，因此它预测修复与时间长度无关，只与那道截止有没有被取消有关。判定形式：给两组同样卡过一次的学生，一组延长可用时间（相当于扩张期），一组取消该次交付要求（相当于取消截止）。若前者的复流期显著缩短，则这是滞后效应；若只有后者缩短，则本文对。

其二，习得性无助 (Seligman & Maier, 1967, *Journal of Experimental*

Psychology, 74(1): 1–9)。它说到哪一步：反复经历不可控的负面结果之后，个体在可控情境中也不再尝试。分离线：习得性无助以失败经验与不可控性为前提；本文的硬通全部是成功经验——交上去了，分数正常，而且往往被表扬。判定形式：给一组学生看他们自己的成绩记录，明确指出“你每一次都通过了”。若通行能力因此恢复，则这是无助感，本文错；若确认成功并不改变题材范围的收缩，则本文对。这一条极便宜，一次班会即可做。

其三，马太效应 (Merton, 1968, *Science*, 159(3810): 56–63)。它说到哪一步：已获承认者更易再获承认，未获承认者的同等工作被系统性低估。分离线：马太效应是分配机制，它作用于外部资源与注意的流向，预测改变承认规则即可改变分配；本文说的是通道本身的容量，它在承认完全公平的条件下同样发生——第一节那个学生每一次都被公平地承认了。判定形式：设一个匿名互评环境（承认完全对称），若题材范围仍然收缩，则本文对。

其四，能力陷阱与利用挤出探索。它说到哪一步：较差程序的早期成功使经验向它集中，探索始终不划算。分离线：能力陷阱以回报比较为机制，改变激励即可改变选择；本文的通扣发生在一次没有比较的通过里——那个孩子并没有在两个题材之间权衡，他是在截止前只剩一条路可走。判定形式：宣布“写没写过的题材加分、写熟悉题材不加分”。若范围随即扩张，则是能力陷阱；若激励改变而未复流清单不变，则本文对。

其五，写作焦虑（Daly & Miller, 1975, *Research in the Teaching of English*, 9(3): 242-249）。这是本领域内部用了几十年的实务正主，必须正面处理。它说到哪一步：对写作情境的焦虑会降低写作意愿与表现，且可用量表测量。分离线：焦虑是一个状态量，它随情境变化、可被降低；通扣是一个已完成的减项，它在焦虑水平恢复正常之后仍然存在。判定形式：对同一批学生同时测焦虑量表与未复流清单。若两者高度相关且焦虑下降时清单缩短，则本文被它覆盖；若存在一批焦虑水平正常而清单很长的学生，则本文对。第一节那个分数上升的学生正是这一类的候选。

其六，写作阻塞（Rose, 1984, *Writer's Block: The Cognitive Dimension*, Southern Illinois University Press）。它说到哪一步：写不下去的人往往被过于僵硬的规则支配，写得动的人规则更灵活。分离线：这一套把原因放在规则的僵硬上，因此松动规则即可解除；本文把原因放在一次已经发生的通过上，规则松动不改变已经被扣掉的那一份。判定形式：教给两组学生同样一套更灵活的写作规则。若两组的题材范围同等扩张，则本文错；若只有从未硬通过的那一组扩张，则本文对。

其七，学习的高原期（Bryan & Harter, 1897, *Psychological Review*, 4(1): 27-53）。它说到哪一步：技能获得中存在长期不进步的 plateau，之后可能出现跃升。分离线：高原期是暂时的停滞，其定义里就含着“之后会上去”；本文说的是不回升，且预测复流期在多数情形下取 ∞ 。判定形式：拉长观测窗到三年。若停滞段之后普遍出现跃升，则这是高原期；若一部分个体的题材范围在三年内单调收缩、从无回升，则本文对。这条正是本文最愿意接受检验的地方——它可以直接证伪本文。

其八，中断时间序列设计（方法学占位者）。这是方法上离本文最近的一位：在一个时间序列中标定一个干预点，比较前后的水平与斜率。分离线：这套方法要求干预点是外生且可观测的；本文的事件点（硬通）恰恰是不可观测的——记录里没有它，必须先用交付时刻的位置把它推断出来。判定形式：若用外生事件（换教师、换教材、疫情停课）作为断点就能解释题材范围的收缩，则本文不必要；若断点必须落在那些看不出异常的个体交付事件上，则本文的读数不可被这套方法取代。

其九，带回（磁滞，源出 Ewing 1885 的磁化研究，后被广泛借用）。剥掉写作这个领域的全部名词，本文的形式是：一个系统在一次被迫的越界之后，回不到原来的状态，而越界过程本身留下的记录是完整的。这个纯形式的结构在上游有名字——带回；相邻的还有路径依赖与临界迟滞。分离线：带回描述的是状态与驱动之间的回线，它不要求那次越界被记为成功；本文要求的正是“被记为成功”这一条，且全部分量都在这里。判定形式：找一次被记为失败的越界（缺交），若它同样留下减项，则本文只是带回的一个场景；若只有被记为成功的那些留下减项，则本文对。

其十，套作与宿构条款（阅卷实务）。它说到哪一步：整篇沿用已有构思者降档。分离线：条款处理的是产物的来源，本文处理的是通过的方式；一篇完全原创的作文可以是硬通的产物，一篇高度套用的作文也可以是从容通过的产物。判定形式：取一批被判定为套作的作文与一批未被判定的，比较两组的交付时刻分布。若两者无差异，则条款与本文测的是两个正交的量——那正是本文的主张。

另有三位本文尚未正面交手，在此指名请进：目标设置理论中关于截止期的实验传统（Locke 与 Latham 一系）；行为经济学关于最后期限效应与自设期限的研究（Ariely 与 Wertenbroch 一系）；以及组织研究中关于时间压力下的搜索收窄（threat-rigidity 一系）。第二位尤其危险：它已经证明自设期限优于外部期限，而这与本文第十节的处方方向一致，却是从完全不同的机制推出来的。

十四、一块没有人回头看的地基

把上一节十位邻居各自看不见的东西列出来，会发现它们指向同一处。

- 滞后效应把“减项由冲击造成”当作给定，因此看不见由一次成功造成的减项。一旦承认后者，它的核心工具——用冲击的强度与持续时间来解释均衡的移动——就失去了自变量。
- 习得性无助把“行为的收缩来自失败经验”当作给定，因此看不见收缩可以完全建立在成功经验上。一旦承认，它整套实验范式（给予不可控的负面刺激）就无法覆盖这一类现象。
- 马太效应把“承认是可以被记录的”当作给定，因此看不见没有被承认拒绝、却同样被扣掉的那一份。一旦承认，它关于分配不公的论证就只覆盖了问题的一半。
- 写作焦虑把“状态可以被量表捕捉”当作给定，因此看不见一个已经完成、不再有情绪表现的减项。一旦承认，量表就不再是这一类问题的入口。
- 中断时间序列把“干预点是外生且可观测的”当作给定，因此看不见事件点本身不进记录的那一类。一旦承认，它就必须先解决一个它不处理的问题：断点从哪里来。

五行背后是同样被当作给定的东西：

凡是发生过的代价，必定在某处留下过记录——要么记在结果上，要么记在过程上，要么记在当事人的状态上。

代价被默认为可记录的。这是所有人共同站着的那块地。互相批判的各派都站在上面，因此谁也没有回头看它：滞后效应去记结果，焦虑量表去记状态，时间序列去记过程，三家分工明确，而它们分工的前提是那笔账总在三处之一。

本文说的是另一件事：有一类代价，三处都不记，而且不是疏漏。结果处不记，因为它合格；过程处不记，因为它按时；状态处不记，因为当事人自己也说不上哪里不对——他这一次确实做到了。这一类代价之所以不进账，恰恰是因为它发生在一次每一项都达标的通过里，而所有的记账规则都是围绕“哪一项没达标”设计的。

这也是本文与前面十位的根本分处：他们都在问“这笔代价记在了哪里、记得对不对”，本文问的是“有没有一笔代价，恰恰因为一切正常而无处可记”。

需要如实补一句：这块地基不是本文推翻的，本文只是站到它外面看了一眼。上一节那十位里有四位——滞后效应、容量跌落、马太效应、习得性无助——是从四个不同方向逼近同一件事，而都差最后一步：它们各自都指出了“一次事件之后系统回不去”，却没有一位追问过那一次事件在记录上是什么样子。四位都默认它是一次可辨认的坏事。本文补的正是这一步：它是一次可辨认的好事。

十五、与同题三篇分界

本栏这道题已有三篇同族，本文必须逐篇指名，并各给一条可裁决的对照。

与《先析》的分界（同题、同为一条路）。那一篇说：最先形成的那段合格文字消耗同一批原料，关掉了另一段；读数是产出中此前已成句部分的比重。本文说：卡住之后被硬性交付的那一次，扣掉的是此后的通行能力；读数是复流期。两者是互补而非竞争——一个测的是这一次产出的内部构成，一个测的是这一次通过对下一次的内部影响。数学上近乎正交：一个从不排队、每次都从容通过的学生，其先析比重可以极高（他熟练地调用现成形态而毫不费力）；一个先析比重很低、每次都在硬长新东西的学生，恰恰最容易反复硬通。判定：若某样本的先析比重在高位而未复流清单为空，则《先析》适用而本文不适用；若先析比重在中位数以下而未复流清单很长，则反之。两篇合起来给出一个两难：调用现成形态会消耗原料，不调用则更容易撞上截止。

与《摊归》的分界（同栏 What 题）。那一篇处理的是“何谓作文”——一篇作文的归属里有多少份额并非本人所出，读数是入位事件。本文不问归属，只问通行。判定：若问题在于产物的份额归不了本人，属《摊归》；若份额清楚而此后的可写范围在收缩，属本文。

与《他承》的分界（同栏 Why 题）。那一篇处理的是“为什么需要作文”，落在支撑的转移上。本文不回答需不需要，只回答一次已经发生的写作扣掉了什么。两者不冲突，但共用一个前提：写作是一件有代价的事，而代价的账目不全。

与本栏用过交通流那一篇的分界。本栏此前有一篇借用过交通流的容量跌落，用以解释一种持续加码而不收敛的局面，读数是份额。本文与它有三处不同：那一篇把容量跌落当作动力机制来问“为什么不收敛”，本文把它当作通道状态来问“这一次是怎么过去的”；那一篇的读数是比率，本文的读数是时长且带 ∞ 值；最要紧的是，本文加了那一篇没有的第三个条件——不可推迟

的截止，而这一条是第八节那次失败逼出来的。判定：若在没有截止约束的场合观察到同样的不收敛，则那一篇的机制成立而本文的三条件说多余；若不收敛只出现在有截止的场合，则本文对。这一条对照是本文与自家最近的一篇之间唯一的分离线，不写出来，两篇会被读成同一个说法的两次表述。

十六、一个可当场执行的最小实验，与七条证伪条款

本文的主张分三层，各层证伪条件不同。第一层最强也最易倒；第三层最稳，且它的检验最便宜。

- 第一层：通扣是一个真实的机制——拥塞、不可推迟的截止、全部交付三者同时在场时，此后同类写作的通行能力被下调。
- 第二层：第五节那段五段序列与四种通过方式的对照是一个有用的分析工具。
- 第三层：在基础教育阶段的作文记录里，交付时刻数据普遍存在而无人读取；因此通行方式的变化在制度上可观测而未被观测。

最强的竞争解释必须先写出来：其实不就是学生越大越怕写、越写越应付吗？这是一个关于动机的解释，它能解释第一节的三份记录中的每一份，而且比本文简单得多。以下条款必须在控制掉它之后仍然成立。

最小实验（可在一个学期内、一所学校内完成）。同年级六个班（同质多单元，不用两校两地充数）。三个班保持现行做法；三个班给每位学生每学期两次“撤出额度”：报告卡住即可撤出该次写作，交一段说明代替，计入完成，且使用记录不进入任何比较。测量：两组的未复流清单长度、题材类别数、交付时刻分布、分数与完成率。控制变量：教师、教材、作业总量、家庭背景分布。若撤出组的未复流清单长度与对照组无差异，则本文第一层为伪。

条款一（剂量—反应，机制主件）。控制写作动机量表得分与作业总量后，若“硬通次数”与“未复流清单长度”之间不存在剂量—反应关系，则第一层为伪——届时收缩应归因于动机衰减而非通行方式。

条款二（正向读数，写成顺序）。交付时刻向后段集中的变化，早于题材类别数的下降出现，且间隔可测。若二者同时发生或次序相反，则本文的因果次序为伪。

条款三（截止条款，本文最愿意押的一条）。在没有不可推迟截止的写作场合（自愿投稿、无人催的个人写作、允许无限期延后的课题），若题材范围同样收缩，则第九节的三条件说为伪——届时截止不是必要条件，本文应退回到一个只关于拥塞的说法。第八节那次失败给这一条提供了第一份材料，方向与本文一致，但因为那是源码而不是写作，只能算旁证。

条款四（复流条款）。取消该次交付要求（而非延长期限）之后的一个学期内，若复流期不缩短，则第十节给出的插手位置为伪。必须同时报告延期组：若延期与撤出效果相同，则本文关于“延期保留截止”的区分不成立。

条款五（不通过条款，最反直觉的一条）。比较缺交组与硬通组。若缺交组的未复流清单不短于硬通组，则第五节那张表的第四行为伪，本文将失去它最反直觉也最有用的那条推论。

条款六（最便宜的一条，可当场执行）。取任意一所学校的作文平台数据，导出一学年的交付时刻与布置时刻。若已有任何一份评价表、教研记录或学情报告读取过这两个字段的相对位置，则第三层主张为伪。这条所需数据在多数使用作业平台的学校里已经存在，一个人一周内可以查完二十所。

条款七（三年窗口）。把观测窗拉到三年。若停滞段之后普遍出现题材范围的回升，则本文与高原期之说无法区分，第一层退化为一个关于暂时性停滞的描述。

条款八（分类表锁定条款，防自毁）。在同一份分类表版本下，若把类别数由七类细分为二十类之后，未复流清单长度的下降幅度超过 50%，则本读数对分类粒度过度敏感，第二层的分析工具为伪，必须重做分类规则。这一条是第十八节反噬预言的直接检验，也是本文唯一一条用来打自己的条款。

若本文被证伪能学到什么。若条款三倒了（无截止处同样收缩），说明问题出在拥塞本身，那么资源应当投向减少作业总量，而不是投向撤出机制。若条款五倒了（缺交同样留下减项），那么“允许写不出来”就不是解药，反而可能是另一种伤害——这两种情形下的干预方向截然相反，因此这几条值得真去做。

十七、不适用的边界

五处明确不适用，逐条写死。

其一，没有不可推迟截止的写作。个人写作、无人催的长期项目、允许无限期延后的课题。第九节已说明，这些场合的拥塞表现为延期，不表现为通报。

其二，可用时间被压成一次的场合。考场作文、限时竞赛。那里没有第一段与第三段之分，本文的判据无对象。

其三，题材由他人指定且不允许选择的写作。公文、指令性稿件、格式化报告。通扣的表现形式是可选范围收缩；范围本来就为一时，这个读数无处落地。

其四，观测窗短于两个学期的场合。复流期是一个带 ∞ 值的时长读数，短窗会把大量“尚未复流”误记为“未复流”。

其五，用于个人评价。未复流清单不是能力评价，更不是意志力评价。第一节那个学生的清单很长，而他的分数在上升。用这份清单去判断一个人行不行，是本文明确拒绝的用法。

十八、反噬与伦理

反噬预言。未复流清单一旦进入考核，最省事的达标方式不是让学生真的重写那些题材，而是把题材分类表改细——原来的七类拆成二十类，任何一次

写作都能落进一个“新类别”，清单于是自动变短。更糟的是，这种拆分能够以“精细化管理”的名义被正当化，而且它会同时让本文的读数失效。

我看不到出口。任何依赖分类表的读数都可以被分类表的改动击穿；而把分类表锁死则会让它无法适应不同学段与不同地区。本文能做的只有一件：把分类表的版本号写进表头（见第六节清点手册），使改动至少留下痕迹。这不是解决，是让作弊变得可查。

伦理三条（凡产物是可拿去考核人的读数，必须写死）。

其一，这个读数指的是通道，不是人。未复流清单描述的是某一类写作的通行状态，不是写作者的能力、意志或品性。它不得进入学生评价、不得用于分班、不得作为教师考核指标——尤其不得反过来考核教师，因为教师降低清单长度最省事的办法就是减少题材种类。

其二，撤出额度的使用记录不得横向比较。这一条是第十二节第三处让步的直接后果，也是整个干预成立的前提。一旦“谁用了几次撤出”可以被排名，撤出就变成另一道截止，干预当场失效并反向加重。

其三，已按这个读数做出任何安排的，必须公开分类表版本与编码规则，并给出复核通道。具体包括：公布类别定义与跨类处理办法、公布“拥塞形成”的时间阈值（可用时间的最后 20%）、允许被评价者对某一类别的判定提出异议。缺少分类表版本公布尤其危险——它使前后两次的清单长度不可比，而不可比的数最容易被拿来证明任何结论。

十九、两处说不清的地方

本文解释不了下面两件事，如实交出。

其一：为什么有的人反复硬通而通行能力不降。存在一批写作者，长期在极强的截止压力下工作——记者、编辑、连载作者——他们几乎每一次都是在拥塞状态下交付的，而题材范围不但不收缩，反而扩张。本文对此没有解释。可能的方向有两个：一是他们的通行能力本来就远高于需求，从未真正进入第三段；二是他们的题材分类粒度与学生不同，看上去的“扩张”其实是同一类的细分。本文两条都拿不出材料，也没有设计出能把它们分开的观察。

其二：撤出与放弃之间的界线在哪里。第十节的处方是“在拥塞形成之前把这次写作撤出必须交的队列”，第十六节又把“缺交”列为不留通扣的一类。那么反复撤出与干脆不写，区别在哪里？本文只能说出一个不够硬的差别：撤出是在通道仍然通畅时主动让出，不写是在通道已经堵死后被迫停下。而这两者在记录上同样是“没交”。也就是说，本文一边主张撤出，一边拿不出一个能把撤出与放弃分开的读数——这是本文自己的判据管不到的地方，也是第十七节把“用于个人评价”列为不适用的真正原因。

二十、本文自己在这条路上的位置，与一个写死日期的赌注

先用本文的判据检本文自己。

本文有一道不可推迟的截止（一次对话之内成稿），有拥塞（三家的材料互不相容，中间确实卡过），并且被全部交付。按第九节的三条件，本文是一次硬通的产物。

这个定位有三个具体后果，必须说出来：

后果一，本文的未复流清单不为空，且第一条就在第八节。那次跨域检验失败之后，本文没有重做一个能测的设计，而是把它写成了机制修订。这个处理是诚实的，但它同时是通扣的标准形态——一条本来要走的路，因为这一次的截止而被永久绕开了。若换一个没有截止的场合，本文应当做的是重新设计事件识别、把“截止压力”这个变量补进去再跑一次。本文没有做，此处记为未复流。

后果二，再用本文的判据检本栏这条问题家族。这道题已经出了三篇，每一篇都在一次对话内完成，每一篇都有截止、有拥塞、有全部交付。按本文自己的判断，这条产线的可选范围应当在收缩——具体表现为：三篇的结构位越来越像，取材的领域越来越集中在“某学科自己承认解决不了的那件事”这一类。这是一条可以由第三者检验的预测，检验材料就是这三篇本身。本文不能自证其否。

后果三，本文因此不主张自己是从容通过的产物。它能主张的只有一件事：它给出了一条自己也过不了的判据，并且把自己没跑成的那次检验原样留在正文里。

赌注。本文押第三层主张，因为它最稳，且裁决不依赖前两层。

截至 2029 年 8 月 31 日，不会出现任何一份面向中小学作文的评价表、教研记录格式或学情报告模板，把“交付时刻相对可用时间的位置”作为一项常规读数——即按篇记录该次作品的交付时刻处在布置到截止之间的哪一段，并在至少一个地级市范围内实际执行满一个完整学年。

什么不算命中，逐条写死：

1. 只统计按时率、缺交率的不算。那是二值开关，第一节那个班从均匀分布变成集中在最后六小时的整个过程，在按时率上是同一个数。
2. 作业平台已经采集了时间戳但没有任何评价表读取它的不算——采集不是读取，这恰恰是本文第三层主张所说的情形。
3. 只在课题、试点或个别学校使用而未进入常规记录格式的不算。
4. 只记录教师布置时间而不记录学生交付时刻的不算，反之亦然；两个字段缺一，相对位置就算不出来。
5. 以“关注学生写作状态”之类的表述写进文件而没有可清点字段的不算。

若在此日期之前上述条件被满足，本文第三层主张判负。判负比赢更值得：那意味着这件事已经变得可观测，而本文全部论证的目的正是让它可观测。

参考文献

- Banks, J. H. (1991). Two-capacity phenomenon at freeway bottlenecks: Some theoretical issues. *Transportation Research Record*, 1320, 234–241.
- Blanchard, O. J., & Summers, L. H. (1986). Hysteresis and the European unemployment problem. *NBER Macroeconomics Annual*, 1, 15–78.
- Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review*, 4(1), 27–53.
- Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1), 25–42. doi:10.1016/S0191-2615(98)00023-X
- Cassidy, M. J., & Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological*, 39(10), 896–913.
- Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, 9(3), 242–249.
- Hadorn, E. (1963). Differenzierungsleistungen wiederholt fragmentierter Teilstücke männlicher Genitalscheiben von *Drosophila melanogaster* nach Kultur in vivo. *Developmental Biology*, 7, 617–629.
- Hadorn, E., Hürlimann, R., Mindek, G., Schubiger, G., & Staub, M. (1968). Entwicklungsleistungen embryonaler Blasteme von *Drosophila* nach Kultur im Adultwirt. *Revue Suisse de Zoologie*, 75, 557–569.

Hall, F. L., & Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation Research Record*, 1320, 91–98.

Levinson, D., & Zhang, L. (2006). Ramp meters on trial: Evidence from the Twin Cities metering holiday. *Transportation Research Part A: Policy and Practice*, 40(10), 810–828.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22(6), 635–659.

Merton, R. K. (1961). Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society*, 105(5), 470–486.

Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63.

Rose, M. (1984). *Writer's Block: The Cognitive Dimension*. Carbondale: Southern Illinois University Press.

Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology*, 74(1), 1–9.

Spemann, H., & Mangold, H. (1924). Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikroskopische Anatomie und Entwicklungsmechanik*, 100, 599–638.

数据来源：psf/requests、pallets/click、tqdm/tqdm、pallets/flask、pytest-dev/pytest、scrapy/scrapy 六个公开仓库的完整提交历史（bare clone 自 github.com，克隆日期 2026 年 8 月 17 日）；纳入 .py 文件级改动 42,146 次。预注册文本先于统计写就并存档，事后未作修改。第一节三份记录为教研场景的概括性描述，不作为证据使用，仅用于说明问题的形状。

版本与字数：初稿，正文 19,958 汉字（不计参考文献与声明）。

人机分工声明：选题、三门学科的选定、成文方向与全部判断的取舍由作者决定；文献核对、数据获取与统计计算、初稿撰写由人工智能助理完成；第八节的预注册文本在运行统计之前写定，失败结果原样保留；本文对其中的全部错误负责。

编者补节·第九章

一·本章不需要补祖宗

本章第十三节逐条交手了十位近邻：滞后效应、习得性无助、马太效应、能力陷阱、写作焦虑、写作阻塞、学习高原期、中断时间序列、滞回、套作条款；每一条都给出了一个可裁决的对照，其中若干条便宜到一次班会或一节课即可执行。第十四节又把十位共同站着的那块地翻了出来（代价被默认为可记录的），并做了一次编者认为全卷最准的判读：滞后效应、容量跌落、马太效应、习得性无助四位从四个方向逼近同一件事，都差最后一步——都默认那一次事件是一件可辨认的坏事；本章补的是，它是一次可辨认的好事。

第二十节又主动请进三位尚未交手者（目标设置理论中的截止期实验、行为经济学的最后期限与自设期限、组织研究中时间压力下的搜索窄化），并指出其中第二位最危险——它已经证明自设期限优于外部期限，与本章的处方方向一致，却是从完全不同的机制推出来的。

编者核过这十三位，没有可补的第十四位。本章是本卷唯一一章不需要编者补祖宗的。

二·与第二章的接口：那个两难

见对切三。编者在此只补一句本章正文没有说的话：本章的处方（识别硬通、保留未复流清单、在有条件时取消那道截止）若被简化成「早点动笔、别拖」，会直接落进第二章的病灶——越早产出可交付的合格文字，先析率越高。本章与第二章互为对方的成本，这一点必须与本章的处方一起读。

三·本卷的自反落在这一章

本章第二十节用自己的判据检了自己：有不可推迟的截止、有拥塞、被全部交付，因此本章自身就是一次硬通的产物；其未复流清单的第一条就在第八节——那次跨域检验失败之后，作者没有重做一个能测的设计，而是把它写成了机制修订，「一条本来要走的路，因为这一次的截止而被永久绕开了」。

作者还做了一件更少见的事：预测自己这条产线的可选范围正在收缩（结构位越来越像、取材越来越集中于「某学科自己承认解决不了的那件事」这一类），并把这条产线的几篇本身交出去当检验材料，声明「本文不能自证其否」。

编者把这一动作扩展到全卷：本卷九章各自都是一次硬通的产物，本卷的未复流清单列在终编编前语（三条）。编者不为这三条辩护——按本章第十三节其二那条证伪条款，确认成功并不改变收缩。

四·编者提示：本章最愿意被证伪的地方，读者应当认领

本章自己指出，与学习高原期那一条的对照是它最愿意接受检验的地方：高原期的定义里就含着「之后会上去」，而本章预测不回升，且预测复流期在多数情形下取无穷。把观测窗拉长到三年即可裁决——若停滞段之后普遍出现跃升，这是高原期；若一部分个体的题材范围在三年内单调收缩、从无回升，本章对。

编者提请读者注意，这条检验不需要任何新工具：一名教师连续三年保留同一批学生的作文题材清单就够了。本卷九个读数里，这是获取成本最低的一个。它同时也是全卷最容易把作者判输的一个。

对切三·第二章 ↔ 第九章：一个没有解的两难

这一处的分离线是作者自己写出来的，编者只把它从第九章挪到两章之间，并补一句后果。

分离线（作者原有）。两个读数近乎正交：

- 一个从不排队、每次都从容通过的写作者，先析比重可以极高——他熟练地调用现成形态而毫不费力。
- 一个先析比重很低、每次都在硬长新东西的写作者，恰恰最容易反复硬通。

判定：若某样本的先析比重在高位而未复流清单为空，则第二章适用而第九章不适用；若先析比重在中位数以下而未复流清单很长，则反之。

合起来是一个两难：

调用现成形态会消耗原料，不调用则更容易撞上截止。

编者补的那一句：这个两难没有解，只有代价的分配。

第二章开出的处方是一条禁令——在过饱和度尚未足以支持自发相之前，不产出任何可交付的合格文字；诱导期内既不自由写作，也不列提纲。这条处方必然抬高第九章所说的硬通风险：诱导期越长，留给交付的时间越短，撞上不可推迟的截止的概率越大。

反过来，第九章的处方（识别硬通、保留未复流清单、在有条件时取消那道截止）如果被简化成「早点动笔、别拖」，就会直接落进第二章的病灶：越早产出可交付的合格文字，先析率越高。

两章互为对方的成本。这是本卷唯一一处由两个读数共同确定、而本卷给不出解的位置。编者拒绝在这里给出一个折中的时长建议（比如「诱导期占可用时间的三分之一」），理由是第二章明写诱导期长度随接种密度上升而非线性变长，且承认自己说不清这条曲线的形状——在曲线未知的情况下给出一个固定比例，是把本卷批评的那种做法再做一遍。

可执行的替代物只有一条：把两个读数同时挂在同一批写作者身上，先看联合分布长什么样，再谈处方。这需要一次同时记录逐句来源与交付时刻的采集，而这两个字段目前的作文记录里一个都没有——这恰好是第二章与第九章各自那个写死日期的赌注所押的东西。

结语·账外的那一栏

一·九章走到了同一个地方

九章的作者互不知情。取的学科从成核动力学到仲裁法，从药理学的给予节律到交通工程的容量跌落，隔得不能再远。读数从续写依赖率到复流期，量纲互不相通。九章之间几乎没有互相引用。

但九章走到了同一个地方。

第一章说，启动不是第一个字，是续写权第一次落到文本自己手里。第二章说，最先成形的那段合格文字吃掉了另一段的成形条件。第三章说，局部终点在下一轮转成条件，而稳定却不再入的文字是假推进。第四章说，下一段的资格由前文欠下的账决定。第五章说，被删掉的那句话把职责甩给了一个一字未改的句子。第六章说，写完不是没话可说，是作者可以退场而文本自己站得住。第七章说，一轮改进宣布成功的那一刻，损失被挤到没设字段的那本账上。第八章说，判准被学会之后，剩下的那部分才是「好」。第九章说，一次被记为成功的交付，扣掉了此后同类写作的通行能力。

九句话可以压成一句：发生就是决定权在换手；而每一次被判为成功的结算，都在无声处收回一点它。

二·九章共同站到了那块地基之外

导论说过，三种现成答案共享一块地基：凡发生过的推进，都会在产出或记录上留下量。

九章共同做的一件事，是站到这块地基外面看了一眼。

它们发现：结果处不记，因为它合格；过程处不记，因为它按时；水平处不记，因为当事人自己也说不上哪里不对。三处分工记账，而分工的前提是那笔账总在三处之一。有一类代价，恰恰因为一切正常而无处可记。

第九章把这一层说得最准。它指出滞后效应、容量跌落、马太效应、习得性无助这四位，是从四个不同方向逼近同一件事，而都差最后一步——它们都默认那一次事件是一件可辨认的坏事；本卷补的是，它是一次可辨认的好事。

第二章从另一个方向说了同一句。它把十位邻居各自看不见的东西列出来，指向同一处：未发生者被默认为待发生者——它只是没被选中、没被注意到、没被奖励，而不是已经不可能。于是十位邻居都在问「为什么它没有被选出来」，而第二章问的是：它还在不在候选之列。

三·这本书没有做到的三件事

按本卷的体例，这一节不能省。

第一，第四章与第五章一个数都没有算出来。两章给的是设计与纪律，不是读数。第五章那个四臂随机是全卷因果识别做得最彻底的方案，所需的不过是一个能记录版本与按键的写作平台加两名盲于组别的标注者——但它至今没有被任何人跑过一次。

第二，第八章的反证跑是文献反证跑，不是数据跑。它按自己的辨别格把自己放在「概念合法、待实证」这一格，并写道：不能把「理论可测」写成「理论已被验证」。本卷收它时没有为它补一次数据检验，而是按原样收入。

第三，第七章三家取源里有两家在本卷别处已经出现。取的线确实不同，编者核过并列了对照表，但在一本以「取三门远距离学科相撞」为体例的书里，这仍然是取源独立性上的一处实情。本卷用编排位置代替了取源独立性——把它夹在第六章与第八章之间读，而不是单独读。

三条都是「一条本来要走的路，因为这一次的截止而被绕开了」。按第九章的定义，它们此刻正是本卷的未复流项。

编者不为这三条辩护。因为按第九章第十三节那条证伪条款，确认成功并不改变收缩。

四·最后一件事

本卷九章各自都有不可推迟的截止，中间都有过拥塞，最后都被全部交付。按第九章自己的三条件，本卷就是一次硬通的产物。

这不是一句谦辞。它是一条可以被第三者检验的预测：如果第九章成立，那么这条写作路线的可选范围此刻正在收缩——具体表现为各章的结构位越来越像，取材越来越集中于「某学科自己承认解决不了的那件事」这一类。检验材料就是这九章本身。

本卷不能自证其否。

所以最后要说的只有一句，与前言里那句是同一句：请先去把它推翻。

九章里最便宜的那条检验不需要任何新工具——一名教师连续三年保留同一批学生的作文题材清单。如果那份清单显示题材范围在三年内单调收缩、从无回升，本卷的终编对；如果停滞之后普遍出现跃升，本卷的终编错，而且是被一份最普通的教研材料判错的。

一本书能得到的最好结局，未必是被接受。按本卷第二章与第九章各自的赌注条款，两章都写了同一句话：判负比赢更值得——那意味着这件事已经变得可观测，而全部论证的目的正是让它可观测。

卷末统一文献表

本表汇总本卷九章与编者补节的全部文献，按「首著者一年份」归并去重后，西文依姓氏字母、中文依笔画后置排列。每条之后的方括号标出它在本卷第几章出现（「补」指编者补节）。

各章末尾保留各章原有的文献表与编号；各章编号仅在本章内有效，本表为全卷汇总，不重编号。各章著录体例不一（方括号编号制、著者一年份制、题名格式标识制并存），本表统一取著者一年份形；体例的全书统一列为再版事项。

西文文献

- Abadi, M., Budiu, M., Erlingsson, U., & Ligatti, J. (2005). Control-flow integrity. *Proceedings of the 12th ACM Conference on Computer and Communications Security*, 340–353. [8]
- Afrin, T., Kashefi, O., Olshefski, C., Litman, D., Hwa, R., & Godley, A. Effective Interfaces for Student-Driven Revision Sessions for Argumentative Writing. *Proceedings of CHI 2021*. [4]
- Akçakaya, H. R., et al. (2018). Quantifying species recovery and conservation success to develop an IUCN Green List of Species. *Conservation Biology*, 32(5), 1128–1138. <https://doi.org/10.1111/cobi.13112> [8]
- Altan-Bonnet, G., & Germain, R. N. (2005). Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biology*, 3(11), e356. [1]
- Aristotle. *On Rhetoric: A Theory of Civic Discourse* (2nd ed., G. A. Kennedy, Trans.). Oxford University Press, 2007. Book III, Chapter 19 (收束部分). [6]
- Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3), 199–223. [1]
- Baier, C., & Katoen, J.-P. (2008). *Principles of Model Checking*. MIT Press. [6]
- Banks, J. H. (1991). Two-capacity phenomenon at freeway bottlenecks: Some theoretical issues. *Transportation Research Record*, 1320, 234–241. [9]
- Bauer, J., Spanton, S., Henry, R., Quick, J., Dziki, W., Porter, W., & Morris, J. (2001). Ritonavir: An extraordinary example of conformational polymorphism. *Pharmaceutical Research*, 18(6), 859–866. [doi:10.1023/A:1011052932607](https://doi.org/10.1023/A:1011052932607) [2]

- Bebchuk, L. A., Cohen, A., & Ferrell, A. (2009). What matters in corporate governance? *Review of Financial Studies*, 22(2), 783–827.
<https://doi.org/10.1093/rfs/hhn099> [8]
- Bebchuk, L. A., Cohen, A., & Wang, C. C. Y. (2013). Learning and the disappearing association between governance and returns. *Journal of Financial Economics*, 108(2), 323–348.
<https://doi.org/10.1016/j.jfineco.2012.10.004> [8]
- Becker, R., & Döring, W. (1935). Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Annalen der Physik*, 416(8), 719–752. [1]
- BEHRENSMEYER A K, KIDWELL S M, GASTALDO R A. Taphonomy and paleobiology[J]. *Paleobiology*, 2000, 26(S4): 103-147. DOI: 10.1666/0094-8373(2000)26[103:TAP]2.0.CO;2. [5]
- BEHRENSMEYER A K. Taphonomy[A]//ALDERTON D, ELIAS S A, eds. *Encyclopedia of Geology*[M]. 2nd ed. Oxford: Academic Press, 2021: 12-22.
<https://repository.si.edu/handle/10088/120162>. [5]
- BEKIUS L. The reconstruction of the author’ s movement through the text, or how to encode keystroke logged writing processes in TEI-XML[J]. *Variants*, 2021(15-16): 3-43. DOI: 10.4000/variants.1245. [5]
- Belchetz, P. E., Plant, T. M., Nakai, Y., Keogh, E. J., & Knobil, E. (1978). Hypophysial responses to continuous and intermittent delivery of hypothalamic gonadotropin-releasing hormone. *Science*, 202(4368), 631–633. doi:10.1126/science.100883 [2]
- Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates. [1·2·4·6]
- Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, 27, 261–278. [https://doi.org/10.1016/0749-596X\(88\)90054-X](https://doi.org/10.1016/0749-596X(88)90054-X) [3]
- Bergemann, D., & Morris, S. (2005). Robust mechanism design. *Econometrica*, 73(6), 1771–1813. [3]
- Bergemann, D., & Välimäki, J. (2019). Dynamic mechanism design: An introduction. *Journal of Economic Literature*, 57(2), 235–274. [3]
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman. [2]
- Bitzer, L. F. (1968). The rhetorical situation. *Philosophy & Rhetoric*, 1(1), 1–14.
 [3]

- Blanchard, O. J., & Summers, L. H. (1986). Hysteresis and the European unemployment problem. *NBER Macroeconomics Annual*, 1, 15–78. [9]
- BOWEN N, VAN WAES L. Exploring revisions in academic text: Closing the gap between process and product approaches in digital writing[J]. *Written Communication*, 2020, 37(3): 322-364. DOI: 10.1177/0741088320916508. [5]
- Bradley, A. R. (2011). SAT-Based Model Checking without Unrolling. In *VMCAI 2011*, LNCS 6538, 70–87. [6]
- Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review*, 4(1), 27–53. [9]
- Calcagno, C., Distefano, D., O’ Hearn, P. W., & Yang, H. (2009). Compositional Shape Analysis by Means of Bi-Abduction. *POPL*, 289–300. [6]
- Cambridge International. Cambridge IGCSE First Language English (0500) : 教学大纲概览. <https://www.cambridgeinternational.org/> (检索日期 : 2026-08-17) [7]
- Campbell, D. T. (1976). Focal local indicators for social program evaluation. *Social Indicators Research*, 3, 237–256. [7·8]
- Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1), 25–42. doi:10.1016/S0191-2615(98)00023-X [9]
- Cassidy, M. J., & Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological*, 39(10), 896–913. [9]
- Castro, M., & Liskov, B. (1999). Practical Byzantine fault tolerance. *Proceedings of OSDI*, 173–186. [1]
- Chemburkar, S. R., Bauer, J., Deming, K., Spiwek, H., Patel, K., Morris, J., ... MFarland, K. (2000). Dealing with the impact of ritonavir polymorphs on the late stages of bulk drug process development. *Organic Process Research & Development*, 4(5), 413–417. doi:10.1021/op000023y [2]
- CLARK A, CHALMERS D. The extended mind[J]. *Analysis*, 1998, 58(1): 7-19. DOI: 10.1093/analys/58.1.7. [5]
- Clarke, E. M., Grumberg, O., & Peled, D. A. (1999). *Model Checking*. MIT Press. [6]
- College Board. AP English Language and Composition : 考试说明与评分量规. <https://apcentral.collegeboard.org/> (检索日期 : 2026-08-17) [7]

- CONIJN R, DUX SPELTZ E, VAN ZAAANEN M, et al. A process-oriented dataset of revisions during writing[A]//Proceedings of the Twelfth Language Resources and Evaluation Conference[C]. Marseille: ELRA, 2020: 363-368. <https://aclanthology.org/2020.lrec-1.45/>. [5]
- CONIJN R, DUX SPELTZ E, CHUKHAREV-HUDILAINEN E. Automated extraction of revision events from keystroke data[J]. *Reading and Writing*, 2024, 37(2): 483-508. DOI: 10.1007/s11145-021-10222-w. [5]
- CROSSLEY S A, BAFFOUR P, TIAN Y, et al. The persuasive essays for rating, selecting, and understanding argumentative and discourse elements (PERSUADE) corpus 1.0[J]. *Assessing Writing*, 2022, 54: 100667. DOI: 10.1016/j.asw.2022.100667. [5]
- Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, 9(3), 242-249. [9]
- de Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to Text Linguistics*. Longman. [6]
- DE BIASI P-M, WASSENAAR I. What is a literary draft? Toward a functional typology of genetic documentation[J]. *Yale French Studies*, 1996(89): 26-58. DOI: 10.2307/2930337. [5]
- de Moura, L., & Bjørner, N. (2008). Z3: An Efficient SMT Solver. In *TACAS 2008*, LNCS 4963, 337-340. [6]
- Diederich, P. B., French, J. W., & Carlton, S. T. (1961). Factors in judgments of writing ability. *ETS Research Bulletin RB-61-15*. Educational Testing Service. [8]
- Duncker, K. (1945). On problem-solving. *Psychological Monographs*, 58(5), whole no. 270. [2]
- Dunitz, J. D., & Bernstein, J. (1995). Disappearing polymorphs. *Accounts of Chemical Research*, 28(4), 193-200. doi:10.1021/ar00052a005 [2]
- Dupont, S., Morsut, L., Aragona, M., et al. Role of YAP/TAZ in Mechanotransduction. *Nature*, 2011, 474: 179-183. DOI: 10.1038/nature10137. [4]
- Echt, D. S., Liebson, P. R., Mitchell, L. B., Peters, R. W., Obias-Manno, D., Barker, A. H., ... Greene, H. L. (1991). Mortality and morbidity in patients receiving encainide, flecainide, or placebo: The Cardiac Arrhythmia Suppression Trial. *New England Journal of Medicine*, 324(12), 781-788. [2]

- Elbow, P. (1973). *Writing Without Teachers*. New York: Oxford University Press. [1·2·4]
- Emig, J. *The Composing Processes of Twelfth Graders*. NCTE Research Report No. 13. Urbana, IL: National Council of Teachers of English, 1971. [1·4·5]
- Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28(2), 122–128. <https://doi.org/10.58680/ccc197716382> [3·7]
- Engler, A. J., Sen, S., Sweeney, H. L., & Discher, D. E. Matrix Elasticity Directs Stem Cell Lineage Specification. *Cell*, 2006, 126: 677–689. [4]
- ETS. TOEFL iBT Writing Section : 评分标准与说明.
<https://www.ets.org/toefl> (检索日期 : 2026-08-17) [7]
- FAIGLEY L, WITTE S. Analyzing revision[J]. *College Composition and Communication*, 1981, 32(4): 400-414. DOI: 10.2307/356602. [5]
- FERRER D, CORCORAN M G. Clementis’ s cap: Retroaction and persistence in the genetic process[J]. *Yale French Studies*, 1996(89): 223-236. DOI: 10.2307/2930349. [5]
- FERRER D. Genetic criticism with textual criticism: From variant to variation[J]. *Variants*, 2016(12-13): 57-64. DOI: 10.4000/variants.284. [5]
- Fischer, M. J., Lynch, N. A., & Paterson, M. S. (1985). Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2), 374–382. [1]
- FITZGERALD J. Research on revision in writing[J]. *Review of Educational Research*, 1987, 57(4): 481-506. DOI: 10.3102/00346543057004481. [5]
- Flor, M., & Cahill, A. (2025). Automated scoring of open-ended written responses: Possibilities and challenges. In *Innovative Digital-Based International Large-Scale Assessments* (pp. 265–279). Springer. https://doi.org/10.1007/978-3-031-90951-1_11 [8]
- Flower, L. S., & Hayes, J. R. (1977). Problem-solving strategies and the writing process. *College English*, 39(4), 449–461. <https://doi.org/10.2307/375768> [7]
- Flower, L. (1979). Writer-Based Prose: A Cognitive Basis for Problems in Writing. *College English*, 41(1), 19–37. [6]
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. [1·3·4·5·6]

- Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing What to Write: Conceptual Processes in Text Production* (pp. 139–160). Amsterdam: Amsterdam University Press. [1·2]
- Geertz, C. (1963). *Agricultural Involution: The Processes of Ecological Change in Indonesia*. Berkeley: University of California Press. [2]
- GENTNER D. Structure-mapping: A theoretical framework for analogy[J]. *Cognitive Science*, 1983, 7(2): 155-170. DOI: 10.1207/s15516709cog0702_3. [5]
- Ghibellini, R., & Meier, B. A Meta-Analysis of the Zeigarnik and Ovsiankina Effects. *Humanities and Social Sciences Communications*, 2025. [4]
- GODET M. The third-body approach: A mechanical view of wear[J]. *Wear*, 1984, 100(1-3): 437-452. DOI: 10.1016/0043-1648(84)90025-5. [5]
- Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics*, 118(1), 107–156. <https://doi.org/10.1162/00335530360535162> [8]
- Grace, M. K., et al. (2021). Testing a global standard for quantifying species recovery and assessing conservation impact. *Conservation Biology*. <https://doi.org/10.1111/cobi.13756> [8]
- Hadorn, E. (1963). Differenzierungsleistungen wiederholt fragmentierter Teilstücke männlicher Genitalscheiben von *Drosophila melanogaster* nach Kultur in vivo. *Developmental Biology*, 7, 617–629. [9]
- Hadorn, E., Hürlimann, R., Mindek, G., Schubiger, G., & Staub, M. (1968). Entwicklungsleistungen embryonaler Blasteme von *Drosophila* nach Kultur im Adultwirt. *Revue Suisse de Zoologie*, 75, 557–569. [9]
- Hall, F. L., & Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation Research Record*, 1320, 91–98. [9]
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman. [4·6]
- Hamblin, C. L. (1970). *Fallacies*. London: Methuen. [補]
- 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. [1·補]
- HAYES J R. Modeling and remodeling writing[J]. *Written Communication*, 2012, 29(3): 369-388. DOI: 10.1177/0741088312451260. [5]

- Hewlett Foundation / Kaggle. (2012). Automated Student Assessment Prize (ASAP): Automated Essay Scoring dataset. [6]
- Hoare, C. A. R. (1969). An Axiomatic Basis for Computer Programming. *Communications of the ACM*, 12(10), 576–580. [6]
- HOLLAN J, HUTCHINS E, KIRSH D. Distributed cognition: Toward a new foundation for human-computer interaction research[J]. *ACM Transactions on Computer-Human Interaction*, 2000, 7(2): 174-196. DOI: 10.1145/353485.353487. [5]
- Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). *Resilience Engineering: Concepts and Precepts*. Ashgate. [6]
- Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Ashgate. [6]
- Hopfield, J. J. (1974). Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences*, 71(10), 4135–4139. [1]
- <https://uncitral.un.org/en/texts/arbitration/contractualtexts/arbitration> [6]
- Hutchins, E. (1995). *Cognition in the Wild*. MIT Press. [3]
- IELTS. Writing 分项评分标准与备考资料. <https://ielts.org/> (检索日期：2026-08-17) [7]
- International Baccalaureate. Language A：评审说明（2026 年版）. <https://www.ibo.org/> (检索日期：2026-08-17) [7]
- ISO. (2022). ISO 21448:2022 Road Vehicles — Safety of the Intended Functionality. [6]
- IUPAC. “product-determining step”.《化学术语汇编》（金皮书）第 5 版，在线版 5.0.0（2025）；原始来源 PAC 1994, p.1153. <https://goldbook.iupac.org/> (检索日期：2026-08-17) [7]
- Kamigaito, M., Ando, T., & Sawamoto, M. Metal-Catalyzed Living Radical Polymerization. *Chemical Reviews*, 2001, 101(12): 3689–3746. DOI: 10.1021/cr9901182. [4]
- Kashefi, O., Afrin, T., Dale, M., Olshefski, C., Godley, A., Litman, D., & Hwa, R. ArgRewrite V.2: An Annotated Argumentative Revisions Corpus. *Language Resources and Evaluation*, 2022, 56(3): 881–915. DOI: 10.1007/s10579-021-09567-z. [4]
- Kauffman, S. A. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1), 1–24. [补]

- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing*. Lawrence Erlbaum Associates. [1]
- Kermode, F. (1967). *The Sense of an Ending: Studies in the Theory of Fiction*. Oxford University Press. [6]
- Kicheva, A., & Briscoe, J. (2023). Control of tissue development by morphogens. *Annual Review of Cell and Developmental Biology*, 39, 91–121. [3·5]
- KIDWELL S M, JABLONSKI D. Taphonomic feedback: Ecological consequences of shell accumulation[A]//TEVESZ M J S, MCCALL P L, eds. *Biotic Interactions in Recent and Fossil Benthic Communities*[C]. New York: Plenum Press, 1983: 195-248. [5]
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513–549. [3]
- Klein, G., et al. (2009). seL4: Formal Verification of an OS Kernel. *Proceedings of SOSp*, 207–220. [6]
- Lamport, L. (1998). The part-time parliament. *ACM Transactions on Computer Systems*, 16(2), 133–169. [1]
- Lecuit, T., Lenne, P.-F., & Munro, E. (2011). Force generation, transmission, and integration during cell and tissue morphogenesis. *Annual Review of Cell and Developmental Biology*, 27, 157–184. [3]
- Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392. [1·2·3·5·補]
- Leroy, X. (2009). Formal Verification of a Realistic Compiler. *Communications of the ACM*, 52(7), 107–115. [6]
- Leveson, N. (2011). *Engineering a Safer World: Systems Thinking Applied to Safety*. MIT Press. [6]
- Levinson, D., & Zhang, L. (2006). Ramp meters on trial: Evidence from the Twin Cities metering holiday. *Transportation Research Part A: Policy and Practice*, 40(10), 810–828. [9]
- Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, 14(S2), 95–112. [2]
- Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319–338. doi:10.1146/annurev.so.14.080188.001535 [2]

- Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves II: A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A*, 229, 317–345. [3]
- Lo Sardo, D. R., Gravino, P., Cuskley, C., & Loreto, V. (2023). Exploitation and exploration in text evolution: Quantifying planning and translation flows during writing. *arXiv:2302.03645*. [1]
- Lopes, N. P., Menendez, D., Nagarakatte, S., & Regehr, J. Provably Correct Peephole Optimizations with Alive. *Proceedings of PLDI 2015*, 22–32. DOI: 10.1145/2737924.2737965. [4]
- Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, 54(6), whole no. 248. [2]
- MacArthur, C. A., Graham, S., & Fitzgerald, J. (Eds.). (2025). *Handbook of Writing Research* (3rd ed.). Guilford Press. [3]
- Mace, G. M., Collar, N. J., Gaston, K. J., et al. (2008). Quantification of extinction risk: IUCN’ s system for classifying threatened species. *Conservation Biology*, 22(6), 1424–1442. <https://doi.org/10.1111/j.1523-1739.2008.01044.x> [8]
- MAHLOW C, ULASIK M A, TUGGENER D. Extraction of transforming sequences and sentence histories from writing process data: A first step towards linguistic modeling of writing[J]. *Reading and Writing*, 2024, 37: 443-482. DOI: 10.1007/s11145-021-10234-6. (网络首发于 2022 年) [5]
- Mann, W. C., & Thompson, S. A. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*, 1988, 8(3): 243–281. DOI: 10.1515/text.1.1988.8.3.243. [4·5·补]
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. [2]
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel. [1]
- McKeithan, T. W. (1995). Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences*, 92(11), 5042–5046. [1]
- Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American Sociological Review*, 1(6), 894–904. [补]
- Merton, R. K. (1940). Bureaucratic structure and personality. *Social Forces*, 18(4), 560–568. [2]

- Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22(6), 635–659. [9]
- Merton, R. K. (1961). Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society*, 105(5), 470–486. [9]
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63. [9]
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.). New York: American Council on Education / Macmillan. [补]
- MIT OpenCourseWare (2007). 10.37 Chemical and Biological Reaction Engineering, Spring 2007. 麻省理工学院. <https://ocw.mit.edu/> (检索日期：2026-08-17) [7]
- Murray, D. M. (1972). Teach writing as a process not product. *The Leaflet*, 71(3), 11–14. [1·4·7]
- Ninio, J. (1975). Kinetic amplification of enzyme discrimination. *Biochimie*, 57, 587–595. [1]
- NIST 计算机安全资源中心. 术语表：verification；validation（定义取自 ISO 9000:2015）. <https://csrc.nist.gov/glossary> (检索日期：2026-08-17) [7]
- NIST (2022). 《工程化可信安全系统》（NIST SP 800-160v1r1）. 美国国家标准与技术研究院. <https://doi.org/10.6028/NIST.SP.800-160v1r1> [7]
- NIST 计算机安全资源中心. 需求验证相关工具与属性清单（完整、一致、正确、可修改、可排序、可追踪、无歧义、可理解、可验证）. <https://csrc.nist.gov/> (检索日期：2026-08-17) [7]
- Odendahl, N., & Deane, P. (2018). Assessing the writing process: A review of current practice (Research Memorandum RM-18-10). *Educational Testing Service*. [8]
- Ongaro, D., & Ousterhout, J. (2014). In search of an understandable consensus algorithm. *Proceedings of the 2014 USENIX Annual Technical Conference*, 305–319. [1]
- Ostwald, W. (1897). Studien über die Bildung und Umwandlung fester Körper. *Zeitschrift für Physikalische Chemie*, 22, 289–330. [2]
- Ovsiankina, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychologische Forschung*, 1928, 11: 302–379. [4]
- Oxtoby, D. W. (1992). Homogeneous nucleation: Theory and experiment. *Journal of Physics: Condensed Matter*, 4, 7627–7650. [1]

- Perl, S. The Composing Processes of Unskilled College Writers. *Research in the Teaching of English*, 1979, 13(4): 317–336. DOI: 10.58680/rte201117867. [4·5]
- Perl, S. (1980). Understanding composing. *College Composition and Communication*, 31(4), 363–369. [3·5]
- Pianko, S. (1979). A description of the composing processes of college freshman writers. *Research in the Teaching of English*, 13(1). [補]
- PIQUETTE K E, WHITEHOUSE R D, eds. *Writing as Material Practice: Substance, Surface and Medium*[M]. London: Ubiquity Press, 2013. DOI: 10.5334/bai. [5]
- Powers, D. E., Burstein, J., Chodorow, M., Fowles, M. E., & Kukich, K. (2014). Comparing the validity of automated and human essay scoring. In *The Research Foundation for the GRE revised General Test. Educational Testing Service*. [8]
- Quinlan, T., Higgins, D., & Wolff, S. (2009). Evaluating the construct-coverage of the e-rater scoring engine (RR-09-01). *Educational Testing Service*. <https://doi.org/10.1002/j.2333-8504.2009.tb02158.x> [8]
- RENOUF M, SAULOT A, BERTHIER Y. Third body flow during wheel-rail interaction[A]//MOTA SOARES C A, et al., eds. *III European Conference on Computational Mechanics*[C]. Dordrecht: Springer, 2006. DOI: 10.1007/1-4020-5370-3_336. [5]
- Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, 4(1), 42–51. <https://doi.org/10.1287/opre.4.1.42> [3]
- Roberts, C. (2012 [1996]). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5, article 6. doi:10.3765/sp.5.6 [補]
- Rohman, D. G. (1965). Pre-Writing: The stage of discovery in the writing process. *College Composition and Communication*, 16(2), 106–112. [1·2]
- Rose, M. (1980). Rigid rules, inflexible plans, and the stifling of language: A cognitivist analysis of writer’ s block. *College Composition and Communication*, 31(4), 389–401. [3]
- Rose, M. (1984). *Writer’ s Block: The Cognitive Dimension*. Carbondale: Southern Illinois University Press. [2·9]
- Roth, A. E. (2023). *Market Design and Maintenance*. NBER Working Paper No. 31947. [3]

- SCARDAMALIA M, BEREITER C. Knowledge telling and knowledge transforming in written composition[A]//ROSENBERG S, ed. *Advances in Applied Psycholinguistics*, Vol. 2: Reading, Writing, and Language Learning[C]. Cambridge: Cambridge University Press, 1987: 142-175. [5]
- Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books / Temple Smith. [3]
- SDE 多学科通融创新法 (v3.9) . (2026-08-16). 三领域·SDE 二阶碰撞出典范文 (内部方法文件) . [8]
- Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology*, 74(1), 1–9. [9]
- Sharples, M. (1999). *How We Write: Writing as Creative Design*. London: Routledge. [补]
- Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631(8022), 755–759. doi:10.1038/s41586-024-07566-y [2]
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482. [3]
- Smith, B. H. (1968). *Poetic Closure: A Study of How Poems End*. University of Chicago Press. [6]
- Snow, K. Z., Monrose, F., Davi, L., Dmitrienko, A., Liebchen, C., & Sadeghi, A.-R. (2013). Just-in-time code reuse: On the effectiveness of fine-grained address space layout randomization. *2013 IEEE Symposium on Security and Privacy*, 574–588. [8]
- Sommers, N. (1980). Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication*, 31(4), 378–388. [5-6]
- Song, W., Zhang, K., Fu, R., Liu, L., Liu, T., & Cheng, M. (2020). Multi-stage pre-training for automated Chinese essay scoring. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6723–6733). doi:10.18653/v1/2020.emnlp-main.546 [2]
- Spemann, H., & Mangold, H. (1924). Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikroskopische Anatomie und Entwicklungsmechanik*, 100, 599–638. [9]
- Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3). [补]

- Storey, R. F. Fundamental Aspects of Living Polymerization. In: Controlled/Living Polymerizations. Royal Society of Chemistry, 2013. DOI: 10.1039/9781849737425-00060. [4]
- Szwarc, M. “Living” Polymers. *Nature*, 1956, 178: 1168–1169. DOI: 10.1038/1781168a0. [4]
- Tanenbaum, A. S., & Van Steen, M. (2023). *Distributed Systems* (4th ed.). Pearson. [1]
- The Nobel Prize. (2007). Mechanism design theory: Scientific background and popular information for the Sveriges Riksbank Prize in Economic Sciences. [3]
- Theory of Constraints Institute. The five focusing steps (持续改善流程 POOGI) . <https://www.tocinstitute.org/> (检索日期 : 2026-08-17) [7]
- Tian, Y., Crossley, S. A., & Van Waes, L. (2025). The KLiCKe corpus: Keystroke logging in compositions for knowledge evaluation. *Journal of Writing Research*, 17(1), 299–336. [1-5]
- TOULMIN S E. *The Uses of Argument*[M]. Cambridge: Cambridge University Press, 1958. [5]
- Traum, D. R., & Allen, J. F. (1994). Discourse obligations in dialogue processing. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, 1–8. [补]
- Turanga1. (retrieved 2026-08-17). *Automated-Essay-Scoring: public mirror of ASAP training_set_rel3.tsv*. GitHub. [6]
- Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B*, 237(641), 37–72. <https://doi.org/10.1098/rstb.1952.0012> [3]
- UNCITRAL. (1985/2006). *Model Law on International Commercial Arbitration*. United Nations. [6]
- UNCITRAL. (2010). *UNCITRAL Arbitration Rules (as revised in 2010)*. United Nations. [6]
- United Nations. (1958). *Convention on the Recognition and Enforcement of Foreign Arbitral Awards (New York Convention)*. [6]
- United Nations. (2019). *United Nations Convention on International Settlement Agreements Resulting from Mediation (Singapore Convention on Mediation)*. [6]

- van Dijk, T. A., & Kintsch, W. *Strategies of Discourse Comprehension*. New York: Academic Press, 1983. [4]
- van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action* (2nd ed.). Springer. [3]
- VAN HULLE D. Dynamic facsimiles: Note on the transcription of born-digital works[J]. *Variants*, 2021(15-16). DOI: 10.4000/variants.1450. [5]
- VAN HULLE D. *Genetic Criticism: Tracing Creativity in Literature*[M]. Oxford: Oxford University Press, 2022. DOI: 10.1093/oso/9780192846792.001.0001. [5]
- Vekilov, P. G. (2010). Nucleation. *Crystal Growth & Design*, 10(12), 5007–5019. [1]
- Velentzas, G., Caines, A., Borgo, R., Pacquetet, E., Hamilton, C., Arnold, T., Nicholls, D., Buttery, P., Gaillat, T., Yannakoudakis, H., & Ballier, N. (2024). Logging Keystrokes in Writing by English Learners. *LREC-COLING 2024*, 10725–10746. [3]
- Wallas, G. (1926). *The Art of Thought*. Harcourt, Brace and Company. [1]
- Walton, D., & Krabbe, E. C. W. (1995). *Commitment in Dialogue*. Albany: SUNY Press. [補]
- Wang, L., Lee, M., Volkov, R., Chau, L. T., & Kang, D. (2025). ScholaWrite: A dataset of end-to-end scholarly writing process. arXiv:2502.02904. [1]
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press. [3]
- WOLPERT L. Positional information and the spatial pattern of cellular differentiation[J]. *Journal of Theoretical Biology*, 1969, 25(1): 1-47. DOI: 10.1016/S0022-5193(69)80016-0. [5]
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press. [2]
- Xu, X., Ghaffarinia, M., Wang, W., Hamlen, K. W., & Lin, Z. (2019). CONFIRM: Evaluating compatibility and relevance of control-flow integrity protections for modern software. 28th USENIX Security Symposium. [8]
- Zeigarnik, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, 1927, 9: 1–85. [4]
- Zhang, M., & Sekar, R. (2013). Control flow integrity for COTS binaries. 22nd USENIX Security Symposium, 337–352. [8]

Zhong, Y., Hao, J., Fauss, M., Li, C., & Wang, Y. (2026). AI-generated essays: Characteristics and implications on automated scoring and academic integrity. *Educational Measurement: Issues and Practice*, 45(1), e70013. <https://doi.org/10.1111/emip.70013> [8]

中文及其他文献

《何谓作文——生成见证视角下的作文等价类理论》（2026，同批内部研究稿）。
〔3〕

中华人民共和国教育部. (2022). 《义务教育语文课程标准（2022年版）》. 北京师范大学出版社。〔1·6〕

付自文、Claude. (2026). SDE 多学科通融创新法：从 0 到 1 不是灵感，是一门可以重复的手艺. *SDE Universes*，学科通融方法论特稿。〔1·4〕

新西兰教育部. e-asTTle：写作如何评分. <https://www.education.govt.nz/>（检索日期：2026-08-17）〔7〕

瑞典皇家科学院／NobelPrize.org (2007). Mechanism Design Theory：科普说明与颁奖辞. <https://www.nobelprize.org/prizes/economic-sciences/2007/>（检索日期：2026-08-17）〔7〕

SDE Universes. (2026). “争议解决与仲裁”（新思想前沿，第 426 号）。〔6〕

SDE Universes. (2026). “形式化方法与程序验证”（新思想前沿，第 352 号）。
〔6〕

SDE Universes. (2026). 新思想前沿：免疫学.
<https://sdeuniverses.com/frontier/immunology/>〔1〕

SDE Universes. (2026). 新思想前沿：分布式系统与区块链.
<https://sdeuniverses.com/frontier/distributed-systems/>〔1〕

SDE Universes. (2026). 新思想前沿：物理化学.
<https://sdeuniverses.com/frontier/physical-chemistry/>〔1〕

SDE Universes. (2026). “风险、安全与可靠性工程”（新思想前沿，第 600 号）。
〔6〕

付自文. (2026). 《SDE 多学科通融创新法（3.9 版）》及“新思想前沿”领域资料. *SDE Universes*。〔6〕

付自文. (2026). 保护生物学与生物多样性·新思想前沿. *SDE Universes*。〔8〕

付自文. (2026). 公司金融与治理·新思想前沿. *SDE Universes*。〔8〕

付自文. (2026). 软件安全与隐私工程·新思想前沿. *SDE Universes*。〔8〕

付自文. 形式化方法与程序验证·新思想前沿[EB/OL]. *SDE Universes*, 2026。〔4〕

付自文. 细胞生物学·新思想前沿[EB/OL]. SDE Universes, 2026.〔4〕

付自文. 高分子科学／Polymer Science·新思想前沿[EB/OL]. SDE Universes, 2026.

〔4〕

附录

附录一·九个读数总表

章	读数	符号	清点单位	阈值／口径	死亡条件
一	续写依赖率	CDR6	候选核之后的六个独立语义决策	≥2/6 且跨两类决策，并含一次回写；阈值为可推翻的操作约定	删核不改变任何下游决策
二	先析率	ϕ	八字连续串	适用区间 [0, 0.5)；总字数低于三百汉字不采信	提示「别用你以前写过的说法」后 ϕ 显著下降
三	端位效应	E_T	稳定局部单元 × 随机条件窗口	七类事件在随后 120 秒的转移分布之散度；零分布由参与者内条件标签置换生成	两种条件下的分布与置换零分布不可分
四	续生转移率	GTR	实质新增单位（句群／段落／命题单元择一，事前固定）	承接 = 1 且改场 = 1 且转位 = 1；仅在预注册的持续生成窗口内计算	打乱核心段序而读者的下一步判断不变
五	零字面角色迁移率	ZRM	字面不变的对齐单元	八类篇章角色；标注一致性目标事前设定	迁移率不高于对齐与标注噪声
六	退场闭合率	ECR	加权续写义务（核心 2／局部 1）	试用门槛 0.90 且无未结核心义务；「不是自然常数」	作者退场后读者的等待点与账本不重合
七	余损迁移序列	—	一轮改进	序列读数，不设总分；本轮终点通过后余损落在另外两域的哪一域	余损从不影响下一轮焦点
八	判准回写余量	CFR	保持性任务（三项，不写入原评分表）	三项通过两项且两个独立读者组方向一致记为阳性；	判准在作品前后不变

				2/3「不是宇宙常数」	
九	复流期	R	一次硬通事件之后	重新以不排队方式通过所需时间，可取 ∞ ；配未复流清单	停滞段之后普遍出现跃升

※ 九个读数全部是机制读数，不是质量评分。本卷禁止把其中任何一个直接用于学生排名、评优或处罚。

附录二·九个新对象的名与英文对应

章	新对象	英文
一	自持起写核·续写权移交	Continuation Dependency Rate (读数)
二	先析·先析率·诱导期·接种史	pre-precipitation; pre-precipitation ratio; induction period; seeding history
三	端位转换·惰性成形	endpoint-role conversion
四	续生位	generative continuation locus
五	稿际余裁·零字面角色迁移	interdraft residual load; zero-literal rhetorical-role migration
六	退场闭合态	author-exit closure state
七	结算余损·追余写作	settlement residual; residual-pursuit composition
八	判准回写余量	Criterion-Feedback Residual
九	通扣·复流期	through-deduction; recovery period

附录三·写死日期的赌注一览

各章赌注均写明了「什么不算命中」，堵住事后改阈值、事后换样本、事后把原量表换个说法重问一遍这三条退路。

到期	章	赌的是什么
2027-12-31	六	退场闭合的经验条款
2028-12-31	一	启动判据的经验条款
2028-12-31	三	不会出现满足六项条件（参与者内随机、 $N\geq 60$ 、文字逐字可见、公开事件编码规则、纳入三类竞争模型、独

		立团队复现) 的研究 追余写作的经验条款
2028-08-17	七	
2028-08-31	八	三校各不少于 100 篇的匿名 数据下, 预计观察到至少 10% 的「高显性达成—CFR 阴性」样本
2029-08-31	二	不会出现设「逐句来源」字 段并在一个地级市执行满一 个完整学年的评价标准
2029-08-31	九	不会出现把「交付时刻相对 可用时间的位置」作为常规 读数并执行满一个完整学年 的评价表

最早到期的是第六章，2027 年年底。最便宜的检验在第九章——一名教师连续三年保留同一批学生的作文题材清单即可。

附录四 · 三处对切索引

对切	两章	分离线	判决方式	位置
一	三 ↔ 五	工作角色 vs 篇章角色；自变量在这一段自己身上 vs 在别处	双向交叉：只移侧栏别处不动／只删别处位置不动	第五章末
二	六 ↔ 七	账内未结 vs 账外未记	取一批闭合率 1.00 的作文做三域余损审计，三种结局均已写死	第七章末
三	二 ↔ 九	产出的内部构成 vs 这一次通过对下一次的影响	两读数近乎正交；合起来是一个没有解的两难	第九章末

附录五 · 本卷的未复流清单

按第九章的判据，本卷自身即一次硬通的产物。以下三条为本卷的未复流项，编者不作辩护。

- 1. 第八章的反证跑是文献反证跑，不是数据跑。 本卷未补数据检验，按原样收入。
- 2. 第四章与第五章一个数都没有算出来。 本卷未跑最小实验，按原样收入设计。
- 3. 第七章三家取源里有两家与本卷别处重复。 本卷未换源重做，以编排位置代替取源独立性。

附录六 · 编者补节所补的祖宗一览

章	所补	分离线的落点
---	----	--------

全书	已写出的文本进入任务环境这一构念	上游已证明产物成为条件；本卷各章多拿的是「那一刻在哪里、由什么读数指认、何时不发生」
一	自催化集合；写作作为约束满足	闭合 vs 移交；约束从哪来 vs 约束何时开始自产
二	上游那笔账	资源 vs 代价
三	同上（被移到侧栏的正是那个构念）	可被读取的信息源 vs 可被操作的对象
四	问题之下的问题	判据严格程度／冻结纪律／语义对象 vs 可删除检验的文本事件
五	修辞结构理论的核—星之分；论证成分类别随上下文改变	成品内部的静态标注 vs 跨版本的角色再分配
六	承诺存量谱系（三位）	承诺者是说话人 vs 文本；裁决者是对方 vs 独立读者；退出是撤回 vs 限界
七	后果效度；残余风险；未预期后果与外部性	同账剩余 vs 跨维转移；可被正式接受 vs 不进记录；不改下一轮目标 vs 夺取下一轮终点
八	后果效度	裁工具 vs 裁作品；不给单一读数 vs 给读数与阈值
九	（无。十三位近邻已在正文逐条交手）	—