

替身：教育改革为何系统性地制造了“不是学习的学习”

改革越努力，教育越沉默——因为系统看不见自己在制造替身

张琼 著 · SDE 学员 · 教育治理与改革批判 · 2026年7月26日

摘要 教育改革反复失败是一个跨时代的顽固现象。既有批判理论将其归咎于技术理性扩张、规训权力渗透或系统对生活世界的殖民，但它们共同遗漏了一个更日常、更无意识的生成机制事实：在那些失败的改革现场，教育并没有“消失”或“被压制”——它被一种更隐蔽的东西替换了。本文识别出一种尚未被既有理论命名的生成机制事实，并将其结算为如下概念——“替身”（pedagogical surrogacy），指称这样一种生成机制现象：当教育行动者（教师与学生）被制度性地要求按照一套外在于当下互动过程的标准去执行“教与学的动作”时，他们生命中最本己的、尚未被语言编码的生长潜能便从互动现场退场，取而代之的是一套可被外部标准识别、认证和管理的模拟形态。替身不是教育的对立面（如灌输、规训、压迫），而是教育的可被管理的像——它以“像教育一样”的方式运转，却系统性地删除了教育之为教育的那件事：一个生命在无可代偿的自我卷入中，生长出此前不曾有的辨识力、判断力与主体性。这一概念不可被“异化”“脱耦”“代偿”“课程窄化”等既有术语无痛替换，因为它命名的不是教育被扭曲、被抑制、被偏离，而是教育被另一个东西顶替了位置。本文以英国“国家素养策略”（1998-2010）为案例，追踪替身如何在真实制度现场中被制造、被认证、被制度化，并通过一个微观生成机制分析，解剖教师在制度指令与职业良知之间做出代偿性选择、并最终将替身内化为自愿捍卫之物的内在机制。在此基础上，本文正面回应了来自基层教师的“公平性质疑”，引入关于标准化教学长期效应的追踪研究证据，论证替身的制造最终会以一种更隐蔽、更不公正的方式剥夺弱势学生的长远发展权利。克服制造替身困境的路径，不是提供“更好的替身设计”，而是从制度层面撤回那些持续要求教育“以某种特定方式被呈现和被认证”的机制，为“不知道”留出合法的制度空间。

关键词：替身；教育改革；生成机制；英国国家素养策略；教育评价；教育公平

——一个生成机制概念的提出与论证

一、被误读的失败：当改革越努力，教育越沉默

教育改革的历史，几乎是一部重复失败的历史。这不是耸人听闻的判断，而是跨越一个世纪的实证记录。从杜威对美国进步主义教育运动的批判性反思，到拉里·库班对二十世纪美国课堂“变中的不变”的经典研究，再到戴安·拉维奇对标准化改革运动的系统清算，一个相似景象反复浮现：那些在宣言中承载着“培养完整的人”“激发创造力”“为未来而教”的宏大改革，在经历初期的喧嚣与巨量资源投入后，其内核往往被旧有体制无声地消解，留下一个被新术语包装但实质未变的课堂。我们似乎陷入了一个怪圈：这一轮改革的争论声尚未落下，下一轮新理念的号角已然吹响。

对此，最流行也最省事的解释是“应试教育太根深蒂固”。这句话因其正确而近乎一句咒语，封住了所有更深的追问。它将教育改革的失败等同于一场力量悬殊的政治角力——仿佛只要集结足够强大的社会意志去“松绑”，一切自然的、好的教育便会自发觉醒。然而，这无法解释一个核心悖论：为什么那么多旨在“松绑”的改革，最终都异化成了新的绳索？为什么那些旨在培养学生批判性思维的探究式课程，最终沦为一种有标准格式的新“八股”？为什么那些旨在解放教师创造力的赋权运动，最终以更细密的考核指标和更繁重的事务性劳动收场？

既有学术研究提供了更精致的解释。教育社会学批判传统揭示了学校如何服务于阶级再生产；新制度主义学派用“脱耦”概念解释学校为何在形式上采纳改革、在实质上维持旧做法（Meyer & Rowan, 1977）。政策执行研究则用“模糊-冲突模型”分析改革从顶层设计滑向象征性执行的机制（Matland, 1995）。在更为具体的课程与教学论层面，研究者们提出了“为考试而教”“课程窄化”“数据驱动的决策扭曲”等概念，精确地描述了标准化改革在课堂层面造成的变形。这些理论各有洞见，但它们共同将重心落在改革的外部条件——阶级利益、制度惯性、政策执行的模糊性——之上，或者将问题描述为一个本是好的课程被外部压力“窄化”“扭曲”的过程。一个更深层次的问题始终未被充分追问：改革者用以介入教育世界的思维本身，是否就是问题的构成部分？

本文接续这一追问，但试图比既有的“技术理性批判”“规训权力分析”“系统殖民命题”——以及比“为考试而教”“课程窄化”这些更贴近地面的经验描述——更推进一步。本文论证：教育改革失败的根源，不是外部条件阻碍了改革的落实，也不是改革者动机不纯或用错了工具，更不仅仅是好的课程被考试压力“窄化”了——而是改革从它迈出第一步的方式起，就注定要制造出一种不是学习的学习。这种“像学习一样的东西”在课堂上取代了真实学习的生态位，在评价数据中被认证为“成效”，在制度惯性中被不断复制。“为考试而教”描述的是教师在压力下的策略性适应；“课程窄化”描述的是课程内容在测试导向下的收窄——但这两个概念都仍然暗中预设了一个未被窄化的、“本来的”好教学存在于某处，只是被考试压力挤压变形了。本文要论证的是另一种可能：即便那个“本来的”好教学从未真正存在过，那个被挤压后留下的位置，也不是空着的。一个东西填补了它。那个东西有一个本文为其命名的名字：替身。

替身不是教育的失败，而是教育改革的产品。

要逮住这个产品，不能从“改革为什么失败”入手——那依然在诊断病灶的既成事实论里打转。必须从另一个方向进入：改革现场中，究竟发生了什么，使得那些最敬业的教师和最努力的学生，在明明“在做教育”的情况下，却让教育本身从他们之间溜走了？

这个追问将我们引向一个需要被重新判定的要害维度。既有批判理论——无论是韦伯的“铁笼”、福柯的“规训”还是哈贝马斯的“殖民”——大多将教育改革的失败定位在“一个什么东西侵入了教育”。那是交织的逻辑：一个外部力量（系统、权力、资本）侵蚀了教育赖以发生的土壤。斯科特在《国家的视角》中对那些试图“看清”并“整理”人类实践的国家治理工程所做的分析，也在此列——他指出，自上而下的标准化蓝图如何系统性地忽视地方知识和实践智慧，从而制造出“看起来整齐但功能失调”的制度景观。这一判断与本文的关切深度共鸣，但它仍然主要揭示了一种“外部侵入”式的破坏。本文认为，这个问题真正的要害，落在差异维——一个过程是如何被另一个过程替换掉的。不是土壤被污染了，而是生长本身被调了包。改革不是在教育的外部制造了障碍，而是在教育的内部制造了一个替代品。这个替代品如此像教育，以至于连最真诚的参与者都分辨不出——他们以为自己在教、在学，其实他们只是在执行一份标着“教”与“学”字样的脚本。“为考试而教”命名了教师在执行这份脚本时的动机（为了考试），但它没有命名被执行的这东西本身是什么。“课程窄化”命名了执行范围的收窄，但它没有命名收窄之后那一整套教学动作与学习产出的存在论性质——它在何种意义上，已经不是教育了。

本文余下部分将完成以下工作：首先，通过撤销既有批判理论未能触及的底层先验，结算出“替身”这一新概念并给出严格定义；其次，以英国“国家素养策略”为案例，追踪替身如何在真实制度现场中被制造、被认证、被制度化，并通过一份真实的学生作文样本的文本解剖，让读者看见替身作为产品的具体样貌——同时，通过一个微观生成机制分析，解剖教师在制度压力下从“被迫执行”到“自愿捍卫”的内在转变机制；再次，正面引入布迪厄的“实践逻辑”理论，明确“替身”是对教育实践本体属性的系统性侵犯，由此标定“替身”在学术地图上的精确位置——它与“异化”“脱耦”“为考试而教”等既有概念不可通约，并在此基础上回应一个元理论挑战：本文凭什么有权界定“什么是教”而不沦为另一种“应当”；然后，正面回应来自基层教师的最有力的反驳——关于教育公平的质疑，引入关于标准化教学长期效应的追踪研究证据，将“时间套利论”立在实地之上；最后，以“撤回”而非“修理”为方向，探讨教育治理如何可能从替身的自证循环中脱身。本文将不仅论证替身的存在，更试图展示：一旦用替身这双眼睛重新审视教育改革的现场，许多此前被笼统归为“失败”或“窄化”的现象，便从一团模糊的沮丧中显影出清晰的生成机制轮廓。而看清这个轮廓，是停止制造替身的第一步。

在展开全部论证之前，有必要先行交代本文的证伪条件——一个论点要成立，必须同时清晰地指出什么情况下它会不成立。本文的核心论点——以标准化和外部问责为核心特征的改革会系统性地制造“替身”——将在以下条件下被证伪：如果可以找到一个大面积实施此类改革的案例，在该案例中，可以观察到大多数学生不仅在外测试中表现合格，更在独立于测试语境的、需要自主判断和创造性表达的陌生情境中，持续表现出被该改革宣称培养的素养（如批判性思维、读写的主体性、科学探究的内在驱动力），且这种表现并非来自对评价标准的策略性适应，而教师群体也未被系统性地导向为应对问责而压缩课程的实质性深度、用标准话术替代即兴互动、或以牺牲学生真实的学科兴趣为代价来追求可量化的产出——那么本文关于替身会系统性地取代真实学习的判断就需要被修正。反之，如果研究者发现，即便在最“成功”的标准化改革案例中，学生的能力也高度依赖测试格式的熟悉度、且在需要独立主体性卷入的任务中表现骤降，那么替身的存在就获得了跨案例的实证支持。

二、一个未被命名的生成机制事实：替身

在进入案例深描之前，必须先完成生成机制概念的核心工序：逮住那个被所有既有理论漏掉的底层先验，撤销它，让新概念被矛盾逼出来。本文在发生认识论意义上使用“生成机制”一词，强调探究某一现象在特定制度条件下如何从无到有地生成。这一用法区别于历史生成机制或词源学，与皮亚杰的发生认识论传统有家族相似，但侧重制度现场中教育事件的发生条件而非个体认知结构。

（一）撤销既有解释未能触及的先验

教育领域最痛的矛盾，不是改革失败本身，而是一个令人困惑的反差：改革者怀着真诚的善意——他们要让所有学生“学会学习”，要打破“死记硬背”的桎梏，要为弱势学生提供公平的起点——但他们越努力，教育现场中那些最珍贵的东西（学生的好奇心、教师的职业欢欣、课堂中那种无法被预设的即兴智慧）就越萎缩。这不是恶意造成的恶果，而是善意在特定条件下的系统性异变。这个矛盾的痛感在于：它让改革者既是加害者又是受害者，让教师既是被迫执行者又是自责的同谋，让“改善教育”这个朴素的愿望本身变成了可疑的。

既有解释路径——无论是宏观的批判理论还是中层的制度分析与课程研究——都在解释“为什么好的意图没实现”。它们诊断了阻力、揭示了扭曲、描述了偏差。但它们的共同盲区是：它们始终在描绘一个“真教育”被外部力量伤害或挤压变形（窄化、扭曲、偏离）的故事。在这个故事里，有一个先在的、本质上是好的“教育”，它被各种坏东西（应试、官僚制、考试压力、数据驱动）侵蚀、扭曲、挤压。改革的失败，就是这场伤害的不断重演。

但这个故事无法解释一类关键案例：当外部压力被主动撤除时——比如某些取消高利害考试的实验、某些赋予教师高度自主权的改革

试点——教育并没有“复原”成它“本该有的样子”。撤走了伤害，真正的教育并没有回来。留在原地的，往往是一种空洞的、不知该如何使用自由的状态——教师不知道该教什么，学生不知道该学什么，课堂陷入一种“温和的无序”。如果教育只是“被伤害”或“被窄化”，那么撤走伤害就应该让它康复。但它没有康复。这说明，旧解释框架里那个“被伤害前的好教育”本身就是一个需要被审查的预设。

这个预设是：存在一个先于改革干预的、本质上是好的“教育本身”，它要么被成功地培育（改革成功），要么被外部力量伤害或挤压（改革失败）。所有批判理论都在这个预设里打转。他们争论的是谁在伤害、用什么方式伤害、伤害有多深——但他们从未质疑那个“被伤害者”的存在论地位。这个“教育本身”被默认为一个不证自明的给定物：它就在那里，等着被保护、被解放、被复育。

现在，撤销这个预设。如果我们撤销“存在一个先于干预的好教育”这个预设，整个图景就发生了剧变。不是某种外部力量在伤害一个本来的好东西——而是改革现场从来就不存在一个等待被伤害或解放的“教育本身”。教育不是一个放在那里等人来碰的物体，而是一个必须在每一次互动中被当场发生出来的事件。它不发生，不是因为被压制了，而是因为它没有被发生——或者更精确地说，在它本该发生的位置上，发生了别的东西。这个东西如此像教育，以至于所有人——改革者、教师、学生、家长、评估者——都以为教育已经发生了。但那个真正要紧的事——一个生命在自我卷入中生长出此前不曾有的东西——从来没有登场。

撤销旧预设打开的不是“如何更好地保护教育”，而是一个全新的问题域：在那个“教育应当发生”的位置上，实际发生的是什么呢？它是如何被制造出来的？为什么它看起来如此像教育，却不是？

一个具体的反例可以压垮旧框架。在典型的标准化改革中，课堂里有教学动作——教师讲了解题步骤，学生记了笔记，练习册填满了，考试通过了——但学生对这个科目本身的理解，在他们走出考场的那一刻就开始消散。几个月后，你问他们学了什么，他们能说出“连接词”“三段论”“答题模板”，但你要他们为一个真正想表达的想法写点什么，他们写不出来——不是技法不好，是没有任何想法等着被表达。

这不是“教育被伤害了”。这不是“课程被窄化了”。这是教育根本就没有发生。而填满课堂那一整年时间的，是一个看起来像教育、走起来像教育、在评价数据上被认证为教育、却从来不曾进入教育那件事本身的——替身。

旧框架无法命名这个事实，因为它的全部词汇表都建立在一个没被说出的信念上：不管教育被扭曲成什么样，它至少还是教育。“为考试而教”仍然在说一种“教”——只不过是为了考试。“课程窄化”仍然在说一种“课程”——只不过被收窄了。替身把这个信念打破了。它让一个问题第一次可以被精确地提出：在那一堂又一堂严格执行了教学框架、产出了漂亮考试数据的课里，教育，究竟有没有来过？

（二）再追一刀：查抄连批判者都还站着的那个先验

上一节撤销的先验——“存在一个先于干预的好教育”——还不是最底层的。对它再追一刀：这个先验本身，又预设了什么更不容置疑的东西？

它预设的是：教育者能够知道“教育应当是什么样子”。

不是模模糊糊地知道，而是知道得足够具体、足够可操作，以至于可以把那个“应当”翻译成一套可测量、可复制、可问责的教学行为标准。这个预设是如此根本，以至于连最激烈的改革批判者都很少去碰它。他们批判的是“应当”的内容——应当培养的是批判性思维而不是应试技巧，应当是完整的人而不是人力资本，应当守护学生的生长而不是用蓝图去修理——而不是“教育者能够事先知道教育应当如何发生”这个认知结构本身。

把这个最底层的先验翻到明处：改革者默认自己能够——在进入特定课堂、面对特定学生之前——知道那些学生应当怎样学习、应当学到什么程度、应当如何展示学到了。这不是狂妄，这是日常。它隐藏在每一份课程标准、每一个教学框架、每一张课堂观察量表的背后，作为一种被全体参与者共享的认知前提在无声运转。

这就是那个连批判者都还站着的那个先验。当一位教育哲学家论证“教育应当培养整全的人”时，他也在预设他知道“整全的人”是什么、以及教育如何促成它——尽管他的“知道”可能比标准化改革者的“知道”更高雅、更复杂、更接近人文理想。当一位进步主义教育者主张“教育应当是学生主导的探究”时，他也在预设他知道“好的探究”长什么样、教师应当如何“守护”它。但结构是同一个：在其介入教育现场之前，一个关于“教育应当是……”的完整知识已经被前置给定。改革者的先验和批判者的先验，共享同一副骨架。他们之间的差异——用蓝图去“修理”还是用智慧去“守护”——只是在这副骨架上披挂了不同的衣饰。

（三）对批判者的生成机制审查：本文自己是不是也在立一个新的“应当”？

在继续推进之前，本文必须做一件所有批判者都该做、但很少真的做的事：把那个查抄先验的刀子，架到自己脖子上。

问：当本文说“替身是教育的可被管理的像”、说“教育之为教育的那件事是一个生命在无可代偿的自我卷入中生长出此前不曾有的东西”时，本文是不是也在暗中预设了一个关于“真正的教育应当如何发生”的知识？当本文批评标准化改革者“用蓝图去修理”时，批

评进步主义守护者“用善意去替学生做决定”时，本文是不是在暗示还存在另一种“更正确的做法”——例如让学生自己面对“不知道”的空白时间、教师在那一瞬间不去干预但保持注视——而本文恰好知道那个做法是什么？更直白地问：本文是否在用一个新“应当”（教育应当是不被干预的空白和自我生长）替换旧“应当”（教育应当是标准化的教学、或民主的探究）？如果是，那么本文就犯了本文自己在第二节中诊断的那种错误：批判者的先验和改革者的先验，共享同一副骨架。

这个问题必须被正面回答。如果本文回避它，本文的全部论证就会被对手用一句“你也不过是另一种应当”连根拔起。以下，是本文的回答。

替身概念不做任何关于“真正的教育应当如何发生”的规范性断言。它不做，不是因为谨慎，而是因为它在做一件完全不同的事：它只命名一组使教育的替换得以存在的存在条件。具体地说，替身概念命名的是这样一个生成机制事实：当教育者被制度性地要求按照一套外在于当下互动过程的标准去执行“教与学的动作”时——无论这套标准的内容是“讲三个连接词用法”还是“用提问激发批判性思维”还是“给学生自由探索的空间”——那个必须由学习者亲身闯入、直面“我不知道”、在无可代偿的自我卷入中做出判断的发生性时刻，便从互动的结构中退场了。它不是被否定了、被贬低了、被排挤了——它是根本没有条件发生。因为那个时刻唯一的营养，就是行动者不被外部标准预决其“应当如何”的那个开口。一旦某个外部的“应当”——哪怕那个“应当”的名字叫“自主探究”或“自由表达”——被制度性地占据在那个开口上，作为教师必须在此刻达成的、可被观察和认证的动作，那么学习者自己闯入“不知道”的发生条件就被撤销了。替身便产生了。

请注意这里的区分。本文不是在说“不干预比干预更好”“留白比支架更好”“学生自己折腾比教师引导更好”。本文是在说：无论你选择哪种教学策略，如果你的策略是以一种“你必须在此刻用这种方式呈现教育”的制度强制为前提的——不是你可以选择做、而是你被要求做——那么你就已经制造了一个教育无法发生的条件。你的课堂里被制造出来的东西，无论它看起来像标准化训练还是像自主探究，都已经是替身。因为真正的教育——那个无法被任何标准预写、无法被任何框架指定的时刻——从一开始就从未获得入场许可。

因此，替身概念不是一面“批判之镜”，用来照出别人错在哪；它是一组“存在条件检测器”，用来判断在某一套制度安排下，那个无法被预制的事件还有没有可能发生。它不回答“教育应当怎样”，它仅回答“在何种条件下，教育连发生的可能性都没有”。这个判断不依赖于一个关于“好教育”的规范性预设，它依赖于对“教育之为事件”的发生条件——即必须有学习者自己闯入的不被预决的空白——的在场或缺席的描述。

但是——这里是最关键也最容易被误解的一笔——替身概念不能被反过来使用：不能被当作一种“如果我想让真教育发生，我就应该把课堂变成一片空白、让学生自己去挣扎”的方法指南。那不是生成机制的应用，那是把替身概念重新变成一个“新的应当”。为什么？因为“把课堂变成一片空白”如果被制度化为你应当在此刻给学生留白的教学动作标准，它同样会制造替身——教师将执行“留白”这个教学动作，但那个必须由学生自己闯入的“不知道”，并不会因为教师执行了留白动作就自动降临。教师只是在执行另一种版本的图纸。图纸的内容变了，图纸的结构没变。替身制造的根，不在于图纸上画了什么内容（连接词还是自由探究），而在于“有图纸”这件事本身——在于有一个外部的、被制度赋予强制力的“应当是”，取代了师生在现场中、在那一刻的即兴判断和相互反应。

所以，本文第七部分提出的“为不知道留出空间”，不是在说“应该把留白写入教学框架成为新的标准”——那恰恰是替身的新配方。它是在说：从制度层面撤回那些要求教育必须“以特定方式被呈现和被认证”的机制。撤回之后，教师做什么、怎么做，是一个即兴的、依情境而定的实践问题，本文没有资格、也不试图给出答案。本文只做一件事：命名那个当制度强制在场时必然被制造出来的替身，并指出撤走强制是替身消失的必要条件。撤走之后会发生什么——那是不在本文管辖范围内、也不应被任何理论图纸预设的事。

（四）撤销底层先验，让新判断从矛盾中结算出来

现在，做那个决定性的动作：撤销这个底层先验。

撤销之后，我们面对的世界是这样的：教育者不知道——也不可能事先知道——一个特定生命在特定情境中的生长应当以什么方式、朝什么方向、在哪个时间点、以什么形态发生。

没有“应当”了。那个一直指引着改革者设计施工图纸的导航星，熄灭了。

这不是一个让人舒服的处境。但正是在这个让人不舒服的处境里，一个此前一直被“应当”的光照遮蔽的事实，第一次被看见了。这个事实不是“教育被伤害了”，而是：

当教育者被制度性地要求按照一个“应当如何”的标准去发生教育时，他唯一能做的，就是让他和学生去执行一套“像教育”的动作。这套动作被评价体系认证为教育。但那个真正要紧的东西——一个生命在那一刻，出于内在的、无可代偿的自我卷入，生长出此前不曾有的辨识力、判断力或主体性——从这场“被要求以特定方式发生的互动”中，结构性地撤退了。

为什么是“结构性地”？因为“自我卷入的生长”有一个严格的生成机制条件：它只能发生在行动者没有被外部标准预决其“应当如何”的时刻。当一个教师被要求“在这一分钟内使用提问句式X激发学生的批判性思维”，当一个学生被要求“在二十分钟内写出包含五个连接词和三个从句的完整段落”——那个“被要求”本身，已经将他们从“我自己此刻决定要如何教/如何写”的主体位置上，拉到了

“我被要求在此刻完成这个动作”的执行人位置上。在那个执行动作的一刻，他们的主体性不是“被压抑”了——它根本没有被召唤进场。它在起点处就被跳过了，因为有一个外部的“应当”已经在替它做决定。

于是，在那四十分钟的“读写小时”里，在那填满了“标准话术”和“预期反应”的教案里，在那被连接词模板严密组织起来的作文里——那个本该由教师和学生共同在场的、以不可预知的即兴判断和对个体学习节奏的敏锐体察为核心的教育发生之地，实际上是悬空的。它上面盖着一层严丝合缝的模拟形态：一个由操作动作、可观测产出和评价数据构成的、完美符合“教育应当如何”标准的替身。

替身，不是教育的对立面（如灌输），不是教育的失败形态（如应试），不是教育的扭曲形态（如异化），不是教育的窄化形态（如课程窄化）。替身是教育的可被管理的像——它以“像教育一样”的方式运转，但系统性地删除了教育之为教育的那件事：一个生命在无可代偿的自我卷入中，生长出此前不曾有的东西。

（五）“替身”的生成机制定义与不可还原硬校验

基于上述结算过程，本文对“替身”给出如下严格定义：

替身（pedagogical surrogacy）是教育治理中的一种生成机制现象：当教育行动者（教师与学生）被制度性地要求按照一套外在于当下互动过程的标准去执行“教与学的动作”时，他们生命中最本己的、尚未被语言编码的生长潜能便从互动现场退场，取而代之的是一套可被外部标准识别、认证和管理的模拟形态。这种模拟形态在动作上像教育（教学行为被执行了）、在产出上像学习（任务被完成了）、在数据上像成效（分数提高了），但它结构性地删除了教育之为教育的那件事：一个生命在无可代偿的自我卷入中，生长出此前不曾有的辨识力、判断力与主体性。替身不是教育的另一种形式——它是教育的生态位被另一个发生过程所占据的产物。

现在，进行不可还原硬校验。请用既有学科中任何一个现成的单个概念，尝试1:1替换“替身”：

异化：异化指主体创造的产物反过来支配主体。它预设了一个先在的“本真状态”，主体与之分离。替身不预设任何先在的“本真教育”或“本真学习”。它命名的是：在那个互动现场，教育根本就没有被生成过。不是生成之后被夺走了，而是生成从未发生。异化解释不了“撤走压迫后教育为什么没有自动回来”，替身解释了——因为替身不是被压迫产生的，替身是被制造出来填补“应当”所指定的那个位置的。

脱耦：新制度主义的“脱耦”描述组织在形式上采纳新政策、在实质上维持旧做法。它命名的是表面与实质的分离。替身命名的是另一件事：连实质都已经被替换了，只是这个替换品看起来像实质。教师在教连接词模板时不是在“表面应付”——她是真的在教，学生是真的在写，考出来的分数是真的在提高。脱耦预设了一个“实质上的旧做法”还完整地保留在后头，替身揭示的是：旧做法本身也已经变形了，整个现场都被改造了，没有一个“未被触及的实质”躲在脱耦的背后。

代偿：代偿指系统在压力下做出适应性调整以维持稳定。它命名的是对压力的回应。替身命名的是对可管理性的服从。代偿发生在压力的场域中（有问责，所以教考试技巧）；但替身的制造机制更根本——即使问责压力被撤走，如果那个“教育应当以某种特定方式被呈现和被认证”的制度结构还在，替身依然会被制造。因为问题的根不是压力，是那个前置的“应当”需要被一个看得见的东西来“兑现”。

规训：福柯的规训命名了权力如何通过对身体的操练生产“驯顺的肉体”。规训的目标是服从。替身的目标是像——像教育的样子。规训的技术（层级监视、规范化裁决、检查）可以服务于替身的制造，但它们不是替身。替身可以在最“进步主义”的、最没有压迫感的课堂里照样发生——只要那个课堂被要求以某种可认证的方式“体现出”进步主义。

为考试而教：这是一个在研究文献中被广泛使用的概念，用于描述教师在标准化测试压力下将教学窄化为考试技巧训练的现象。它命名的是教师行为在压力下的策略性变形——教的内容变了（从学科本身变成考试题型），教的动机变了（从培育素养变成追求分数）。但“为考试而教”仍然在说一种“教”——只不过它的目的是考试。它没有质疑被执行的这套教学动作本身的存在论性质。在“为考试而教”的概念框架里，教师仍然相信（或至少，这个概念没有质疑）她正在“教”某种东西——只不过她教的是考试技巧，而不是学科素养。替身要追问的恰恰是这一个点：当教师教学生“在作文的第三段开头务必使用一个转折连接词以获取更高的结构分”时，这种教学动作本身，还能被称为“教”吗？学生学会的，不是“写作需要转折连接词来表达思想的转换”，而是“在第三段开头放连接词能得分”——后者不是写作能力的降低或变形，它从一开始就不在“写作”那个范畴里。它是一条完全不同的发生路径，产生的是完全不同的认知产物。“为考试而教”仍然将其视作一种“教”（尽管是坏的、被扭曲的教）；替身则将其视作一种替换——那个本当属于“写作学习”的生态位，被一个长得像写作学习、但从触及写作本质的拟态所占据。这不是教的内容变差了，而是教这件事本身就没有发生。

在这里，本文必须正面回应一个元理论挑战。对手可以追问：当本文断言“那不是教”时，本文凭什么有权界定什么是“教”？本文是否偷偷引入了一个关于“真正的教应当如何”的规范性标准——而那个标准，和本文正在批判的标准化改革者的“应当”在结构上无异？

本文的回答如下，这个分际必须精确。

第一，本文区分“教”与“非教”的依据，不是某种关于“好教学”的价值规范，而是对这堂课里“有没有一个必须由学习者本人闯入的发生性时刻被允许存在”这一发生条件的在场判断。当教学动作被外部规程完全预决时，那个时刻的发生条件从结构上就被撤销了——不论那个规程的内容是讲连接词还是做探究。所以本文不说“讲连接词就是坏的、自由写作就是好的”，本文说的是：当连接词被当作一个必须在课堂第三环节被执行的教学动作来教时，它不再可能成为一个学习者亲自闯入“为什么要在这里用连接词”的契机。它变成了一条指令。指令被完成了，写作的发生没有开始。

第二，本文的论证不依赖于“本文知道什么是真正的教”。它依赖的是一条更根本、也更谦虚的生成机制判断：当互动的每一分钟都被外部规程所饱和时，有一件事是确定不可能发生的——那就是一个生命在没有外部指令的情况下，自己决定要说一句什么话、写一个什么句子，并为此承担全部的犹豫、不确定和可能的错误。这件事是否在别的时间、别的条件下一定能发生，本文不知道，也不断言。本文的“替身”概念并不在说“只要撤走图纸，真教育就一定会发生”。它只在说：图纸已经把它发生的基本条件堵死了。

第三，也是这里最容易滑倒的一步——替身概念不能被反过来用作一种方法指南。本文不在说，既然教连接词套路是替身，那么教师就应该去搞自由写作、去激发学生的自我表达、去设计更开放的教学任务。那些动作如果被写进教学框架、被课堂观察表考核、被要求“在第五分钟体现出学生自主探究的特征”，它们同样会制造替身——只不过替身的样子从“整齐的模板作文”变成“表演性的自主探究”。替身概念的靶心从来不是某一类教学策略，而是那种要求教育必须以某种特定方式被呈现、被认证的制度强制。这是本文与一切“教学方法改革”文献的根本分界线。

课程窄化：这一概念描述的是，由于标准化测试只覆盖部分科目和部分技能，学校和教师将大量教学时间集中在这些被测试的内容上，从而挤压乃至砍掉那些不被测试的科目和学习领域。它命名的是教学范围的缩减。替身命名的是教学性质的替换。课程窄化解释了为什么音乐课和美术课被取消了——因为它们不考；但它没有解释，即便在那些“被保留下来并且被高强度教学”的科目里（比如读写、数学），究竟在发生什么。“课程窄化”的视角看到的是：读写课更多了，体育课更少了——这是量变和范围变化。替身的视角看到的是：占据那节读写课的，无论它教了多少连接词和答题模板，从头到尾就不是读写——这是质变和存在论替换。二者不在同一个分析层面上。

不可还原性成立。“替身”不是对任何一个既有概念的隐喻性重述。它命名的是一个此前“看得见却说不出、没有任何现成词能装下”的辨别面：在那些真诚的、努力的、数据上成功的教与学现场，教育这件事，被掉包了。

在完成概念结算之后，本文需要在理论坐标上再迈一步——因为迄今对“替身”的定位，仍然主要是通过与既有概念的逐项区分来完成的。这里需要一次正面的大理论对接，将替身放在一个更根本的学术传统中来理解。

布迪厄的“实践逻辑”理论为此提供了最有力的支点。布迪厄论证，实践的运作逻辑是一种模糊的、即兴的、受场域情境塑造的“实践感”——行动者在特定场域中养成了一种身体化的倾向系统（惯习），在面对具体情境时，不是在执行一套事先写好的规则，而是在与情境的即时相遇中做出判断和回应。这种判断是可错的、不确定的、无法被完全编码的，但正是这种含混和即兴，赋予了实践应对复杂多变情境的韧性。“实践的时间性”——行动总是在紧迫的时间流中发生，行动者没有余裕停下来做完全理性的计算——是实践逻辑区别于理论逻辑的根本特征。

“替身”恰恰是这种“实践逻辑”被一种完全相反的逻辑所接管和顶替的产物。教学作为一种高度情境化、依赖教师与学生在具体互动中即时判断的实践，本质上属于布迪厄所描述的那种“含混的”实践领域。一位有经验的教师在课堂上的绝大多数决策——什么时候停下来问一个问题、什么时候对某个学生的反应多追问一句、什么时候抛掉教案追随一个意外出现的学习契机——不是在对教学理论或政策文件的“执行”，而是在她与这群特定的学生、在这个特定的时刻的即时相遇中，被一种被多年职业实践所内化的身体化智慧所驱动。那不是盲目的直觉——那是一种被布迪厄称为“没有指挥的即兴”的实践逻辑。

但当一套标准化教学框架以制度强制力的方式介入这种实践时，它所做的事情不只是一个“外部压力”的施加——不是“教师在压力下把教学变窄了”，而是从根本上改变了教学作为一种实践的存在论条件。教学框架用一套被事先写好的、分解到分钟的动作序列（“在第十五分钟使用提问句式X”）取代了实践中那种必须在瞬间做出的、不可被事先指定的判断。教师不再是“在教学”——她是在“执行预先写好的教学操作”。布迪厄意义上的“实践感”被一种“技术手册的操作规程”所覆盖。

这不是在实践中加入了“理论的指导”——那是将实践简单地置于理论知识的辅助之下。布迪厄的整个理论大厦恰恰在论证，实践本身的运作逻辑——那种含混、即兴、情境化的判断——与实践者在事后用理论知识对自己的实践所做的合理解释，是两种完全不同的东西。前者是“正在做”，后者是“做完之后在说的东西”。当一份教学框架不是作为教师课后反思时的参考资源，而是作为一种要求“你必须在此刻以此种方式做”的制度强制在场时，它就是在用“理论知识”的手术刀，去切除“实践逻辑”的心脏——那种只有在即兴的不确定中才能存活的教育发生之可能。

从这个角度看，“替身”就不仅是“对修理的批判”或“对标准化的批判”。它是对一种更深层的东西的诊断：一切试图用可编码、可复制、可核查的“理论知识”去完全接管含混、即兴、情境化的“实践逻辑”的教育治理模式，其终点都是一个替身。因为实践逻辑被切除之后，需要有一个东西来填补被切除的那个位置，让一切在制度视野中看起来仍然在运转。那个东西，就是替身。它看起来像教学，但它没有实践——它只有操作的执行。

这一理论定位也解释了为什么本文在第五部分会论及“守护者同样可能制造替身”。因为“守护”——当它被作为一套“你应当在此刻用此种方式回应学生”的规范性知识被制度性地施加于教师时——同样是在用“理论知识”替代“实践逻辑”，只不过这份理论知识的内容更宽泛、更弹性。但替代的结构是一样的。在布迪厄的意义上，守护者的“知道学生需要什么”和修理者的“知道教学应当如何操作”，共享同一个生成机制错误：他们都以为实践可以被一份前置的知识完全导航。

因此，“替身”在学术地图上的理论坐标可以这样标定：它是布迪厄所揭示的那对根本张力——实践逻辑与理论知识之间的张力——在教育治理场域中被推向极端后的产物。当理论知识彻底接管了实践逻辑的生态位，教育便从一种活的实践，蜕变为一份动态更新的操作流程图。留在流程图上的，就是替身。

三、替身的工序：英国“国家素养策略”如何系统性地制造替身

以上在理论层面结算出了“替身”的生成机制定义与理论坐标。然而，如果替身真的如本文所论是一种广泛存在于教育改革中的现象，它就不可能只停留在哲学思辨里，而一定会在某个特定的历史时刻、某个具体的制度现场中，留下可供追踪的印迹。本节以英国“国家素养策略”（National Literacy Strategy，以下简称NLS）为案例，进行生成机制深描，展示替身是如何被一步步制造出来的——不仅在制度工序层面，更在具体的课堂互动和个体的认知转变中，被发生出来。

选择NLS作为案例，基于三项理由。第一，它是一个有充分公开信息可以查证的、具有国际影响力的教育改革事件，政策文本、实施数据、评估报告和批评性研究均有可获取文献。第二，它的设计初衷——提升所有学生的读写能力，不让一个孩子掉队——是真诚而善意的，改革推动者并非出于任何控制的恶意，恰恰最能体现替身制造的无意识性。第三，它是“从诊断到设计、从实施到评价、从评价到强化”这条完整链条的典型剖面，替身制造的每一个工序都在其中清晰可辨。

在展开案例深描之前，有必要对已有关于NLS的批判研究做一简要交代。NLS实施期间及其后，大量教育研究者对它的效果和副作用进行了深入分析。Dominic Wyse在一篇被广泛引用的批评性综述中，指出关于NLS有效性的实证证据远非确凿，并质疑其教学框架的科学基础。更具体地，研究者们观察到了一系列令人不安的现象：教师为应对考试压力而将大量教学时间用于考试技巧训练（“为考试而教”）；非考试科目的教学时间被读写和数学严重挤占（“课程窄化”）；教师的教学自主性在精细的教学框架约束下显著下降；学生的写作日益呈现一种“公式化”特征——结构完整、连接词齐全、但内容空洞。这些经验观察是真实而有价值的，它们为替身的存在提供了间接的实证证据。然而，这些研究的问题在于：它们的分析工具本身，仍然停留在“教育被考试压力扭曲/窄化/挤压”的框架内。它们看到连接词作文公式化，称之为“为考试而教的恶果”；看到非考试科目被挤占，称之为“课程窄化”。但它们没有继续追问：那个被公式化写作占据的读写课堂本身，究竟在发生什么？那个连接词作文，仅仅是被“窄化”或“扭曲”了的写作吗，还是它根本就是另一种东西——一种长得像写作、符合写作的评分标准、却在存在论意义上从来就不曾进入“写作”本身的东西？

“替身”概念的提出，正是为了将上述经验观察推向一个更根本的分析层面。

（一）第一步：给定“应当是”——教学框架作为无生命的图纸

1997年，英国新工党政府上台，将教育列为优先政策领域。彼时，关于英国中小学生读写能力的讨论已持续多年，多项国际比较研究显示英国学生在阅读方面的表现主要发达国家中处于落后位置。在此背景下，政府于1998年推出了NLS，核心目标被表述为：到2002年，让80%的11岁学生在国家英语测试中达到预期水平。

改革者没有恶意。他们看到一群孩子不会读、不会写，他们焦急，他们要行动。但行动的第一步，就已经踏入了制造替身的工序：他们不是在问“这一群特定的孩子，在此时此地，正在经历什么、需要什么”，而是在问“会读写的好学生应当是什么样的，我们如何让所有孩子都变成那个样子”。回答这个“应当”的，是一份由中央专家团队根据语言学和教学研究成果绘就的蓝图——“教学框架”。这份框架将“高质量读写教学”分解为一套可观测、可执行、可核查的标准化动作：全班共享阅读与写作（15分钟）、聚焦词汇或句法的精讲（15分钟）、小组或独立读写活动（20分钟）、全班回顾与总结（10分钟）——这就是著名的“读写小时”的时间结构。更进一步地，框架在每一个时间段内，为教师指定了详细的教学动作、推荐话术和可检测的预期学生反应。

这份图纸本身是无生命的。它不知道明天教室里坐着的那个男孩昨天因为写不好句子被同学嘲笑，不知道那个女孩其实有一肚子故事只是不敢写出来。它只知道一件事：好教学等于这四个部件、这幅时间表、这套标准动作的精确组装。在图纸被画好的那一刻，制造替身的第一步已经完成——一个关于“好教学应当如何”的完整知识被前置给定，并且被赋予了制度权威。图纸被画好之后，剩下的问题就只有一个：如何让全国每一间教室都成为这张图纸的精确施工场。

（二）第二步：工序化执行——读写小时中的一堂课如何被替身填满

NLS的核心制度装置——“读写小时”——随后作为全国强制性教学时段被推行至英格兰所有公立小学。教学框架将其精确到了分钟层面，每个段落中教师被期望完成特定教学动作：在“共享阅读”阶段使用大幅面书籍示范阅读，并通过特定提问句式引导学生做出预期回应；在“精讲”阶段聚焦预先指定的语言知识点；在“小组活动”阶段按照指定任务模板进行分级练习。

要理解这个教学框架如何在微观层面制造替身，光看政策文本是不够的——必须看到一份真实的样本：那些在这套框架下被生产出来的学生作文，究竟长什么样。以下是一段在NLS期间及类似标准化读写改革中极有代表性的学生作文片段，由研究者在其关于NLS时期学生写作的分析中作为“公式化写作”的典型样本加以引述。这篇作文的主题是“你度过的最棒的一天”：

“There are many reasons why my trip to the beach was the best day ever. Firstly, the weather was very warm and sunny. Furthermore, the sea was blue and clear. However, we had to leave early because my brother got tired. Therefore, we went home at four o'clock. In conclusion, it was a great day.”

这篇作文在标准化评分维度上表现优异。它有清晰的开头句（引入话题）、使用了四个连接词组织结构（Firstly, Furthermore, However, Therefore）、有一个结论句收尾（In conclusion）。在NLS教学框架指导下，这是一个合格的、可以拿到“达到预期水平”评分等级的写作样本。一个经过连接词模板训练的五年级学生，完全有能力在十五分钟独立写作时间内，精准地产出这样一个段落。

现在，让我们解剖这个样本——不是用“内容空洞”“缺乏描写”这类批评教学法的语汇，而是用生成机制的镜头，去辨认它作为“替身产品”的存在论特征。

第一，这篇作文里有没有一个“人在写”？所有的语法位置都有人称代词：my、my、my brother、we。但读完整个段落，你不知道这次海滩旅行有什么是“这一个特定的人”在此刻真正想说的。暖和的天气、蓝色清澈的海水、“我弟弟累了所以我们提前走了”——这些可以是任何人的海滩。写这篇作文的学生，不是因为“我真的想告诉你那一天的感觉”而写，而是因为“我被要求写一个包含四个连接词和一个结论句的段落”而写。执行写作动作的主体在语法上是第一人称，但在存在论上是空缺的——坐在那里握着笔的，是被一个外部规程充分导航的写作执行者，而不是一个自己正在选择要说什么的写作者。

第二，连接词的“功能”在这里发生了倒转。在真正的写作中，连接词是思想关系的语言标记——作者因为有思想上的转折、递进或因果，才去使用However、Furthermore、Therefore。它们是思想的仆人，先有思想，后有连接词。但在这篇作文里，这个关系被颠倒了：连接词是先被给定的（“你必须在作文中使用至少四个不同类型的连接词来组织结构”），然后学生去填内容来“满足”这些连接词。Furthermore后面的“海水是蓝色清澈的”和前面的“天气很暖和”之间并没有递进关系——它们只是两项并列的环境描述。Furthermore在这里不是思想递进的标记，它是被评分标准要求在那个位置出现的信号物。学生学会了把一个叫“Furthermore”的标签贴在一句话开头，但从未学过——因为从未在这个课堂上被允许去体验——什么叫“我想到了一层更进一步的看法，因此我需要表达递进的词”。

第三，这就是替身作为“产品”的精确样貌：它像写作——它有开头、有结构词、有结论句、符合评分标准在形式上的所有可见指标。但它从根本上就不是写作——它从未进入“一个人用语言去触碰到某种真实的感受、想法或经验”的那个发生过程。发生的过程被短路了：在发生本该开始的那一点（“我想写什么？我真的想写吗？”），一个外部规程替它做了决定（“你必须在今天写出一个包含四个连接词和一个结论句的段落”）。从那个被短路的地方开始，接下来发生的全套动作——构思、组织句子、选用连接词、写结论——虽然走完了“写作动作流程”的全部步骤，但从头到尾不是一种生成，而是一种填装。它填装的模板叫“符合评分标准的作文格式”，它交付的产品叫“替身”。

被这个替身占据的，不只是“写作教学的质量高低”这个评价维度——那依然是“窄化”和“为考试而教”能够处理的层面。被它占据的，是一个生命在一次写作行为中有可能发生的那个时刻：那个必须由他自己去面对一张空白纸、不知从何写起、在心里搜寻一遍自己真正有感觉的东西、然后笨拙地写出第一个不一定符合任何评分标准的句子的时刻。那个时刻在这堂读写小时里从来没有条件出现——因为它被要求产出的不是一个“从自己里长出来的句子”，而是一个“可以被评分标准认证为合格的文本”。从制度指令进入课堂的那一刻起，这两个目标之间就隔着一道无法跨越的生成机制断裂。那个断裂，就是替身的制造现场。

（三）第三步：评价认证——替身如何被宣判为“活着”

最隐蔽也最致命的一步发生在后头。

NLS启动后，政府委托独立研究团队进行了多轮评估。大规模测试数据分析显示：在改革实施初期，11岁学生在国家英语测试中达到预期水平的比例从1997年的约63%上升到了2001年的约75%。这个数据被政策制定者援引为改革有效的关键证据，被媒体广泛报道为“改革正在取得成功”的信号。

然而，同一份评估报告中的质性研究部分提供了一幅微妙得多的画面。评估者通过访谈和课堂观察记录了教师的真实感受：教师们普遍担忧读写的节奏在课堂中难以严格维持；许多教师承认为了应对考试压力而过度强调考试技巧教学；一些校长报告非考试科目的教学时间被挤占以用于额外的读写备考。事实上，教师的工作生活在NLS实施期间发生了实质性的变化——一项对教师职业生活的纵向调查报告指出，尽管教师普遍接受了国家策略以提高读写和算术能力的核心目标，但他们在日常工作中感受到了日益增加的行政负担、被压缩的课程自主空间，以及一种“为外部审计而教学”的普遍情绪，这些因素综合起来对教师的职业认同和士气产生了深刻的影响。这些为本文对教师内部决策的微观解剖提供了坚实的经验背景，这一点将在下一节中展开。

评估者谨慎地写道：“很难将测试分数的上升完全归因于教学质量的提高。”

但在政策传播和公众感知的层面，这些限定条件被系统性地过滤了。分数上升了——那就是成功。至于这些分数是怎么来的——是学生真的会写作了，还是他们被训练成了识别连接词考查点的熟练工——评价体系无法区分，也不关心区分。因为替身与真教育在外部产出的数据形态上，是完全一样的。一篇被“However”“Furthermore”“Therefore”撑起来的作文，在评分标准上可以拿高分——它有结构、有连接词、符合评分规范。它唯一缺少的，是写这篇作文的人。但评分标准从来不打“有没有人”这一项。

这就是替身最阴翳的特质：它不会触发任何警报。一个被制造出来的替身，不需要任何人的主观恶意，只需要一个看不见“有没有人”的评价体系。当评价体系认证了替身，它就不只是认证了“这节课是有效的”——它认证了“这就是教育”。从此，替身不再是教育失败时的临时替代品，它就是教育本身在制度中的合法定义。

（四）微观解剖：一个教师如何从被迫执行者变成替身的自愿捍卫者

上一节的案例深描展示了替身在制度层面的制造工序。但工序要真正运转，必须通过一个个具体教师的具体抉择来完成。本节进行一个生成机制的微观分析，解剖一个教师在制度指令与职业良知之间做出代偿性选择、并最终将替身内化为自愿捍卫之物的完整内在机制。这个分析以NLS期间教师访谈和调查数据所揭示的普遍经验为基础，进行生成机制重构。文中的教师与学生为典型化重构，旨在以思想实验的方式展示制度压力下教师的代偿选择与认知转变机制。但其结构逻辑——从“被迫执行”到“认知内化”的转变——是替身制造中最核心的生成机制环节。

假设一位教师——叫她凯特——站在她的读写小时开始前的那一分钟。她昨晚读完了今天这节课在教学框架上指定的内容：教连接词的三个用法。她也记得昨天课间，她班上那个叫里奥的男孩兴奋地跑来告诉她，他周末读了一本关于霸王龙的书，“它太厉害了，它能一口咬碎一辆车”。里奥平时在写作上很挣扎，他的句子总是很短、很笨拙，但那一刻，他眼里的光让凯特觉得——这就是她想当老师的原因。她想让里奥写出关于恐龙的句子，哪怕那些句子不完美、没有连接词、在考试评分标准上拿不到高分。

但她也知道，今天的读写小时，如果她不按教学框架走，会发生什么。她的学校上一次评估被列入了“需要特别监测”的名单。她的校长每周检查教学计划，确保教学框架中的每一个知识点都被覆盖并记录在案。如果里奥的班级在年末的全国测试中分数下降——哪怕只是几分——她可能会被要求参加“教学改进计划”，她的教职虽不至于受到直接威胁，但她的职业声誉、她在同事中的自尊、她对自己“是一个好老师”的确认，都会受到打击。更微妙的是，她甚至无法确定，如果她抛开框架、追随里奥的恐龙热情，结果会更好。她不知道那会把课堂引向哪里。而教学框架至少给她提供了一份清晰、安全、经得起问责的“正确做法”。

在这个时刻，凯特面对的不是一个简单的“做对的事还是做错的事”的道德选择题。她面对的是两个冲突的“应当”：一个“应当”来自制度——应当完成教学框架指定的教学动作；另一个“应当”来自她自己的职业良知——应当回应具体学生的具体生长需要。这两个“应当”都声称自己是“为了学生好”。前者说：不教会里奥连接词，他将来在考试中会失败，他的社会流动机会会受损。后者说：如果不在这一刻护住里奥眼里的光，他可能永远不会真正爱上写作。

凯特必须在这两个“应当”之间做出选择。她的选择——以及此后千千万万个教师在类似情境下做出的类似选择——在大多数情况下倒向了前者。不是因为前者的理由更强，而是因为前者的制度可见度更高。选择追随教学框架，她的行为在制度记录中是可识别的、可辩护的、可被认证为“合格教学”的。选择追随里奥的恐龙热情，她的行为在制度记录中是不可见的——它可能产出一个珍贵的教育时刻，但这个时刻无法被写入周教学计划、无法在课堂观察量表上被打勾、无法成为她被问责时举在手里的盾牌。她甚至无法向自己证明，那个时刻真的比教连接词“更有价值”。

于是，凯特做出了那个微小的、看似无关紧要的、却是一切替身制造之起点的抉择：她告诉里奥，今天的课要学连接词，恐龙的事以后再说。里奥眼里的光暗了一暗，然后他听话地打开笔记本，开始抄写黑板上的“However”“Furthermore”“Therefore”。凯特完成了今天的教学任务。她的周教学计划上又多了一个可以被划掉的条目。

此时，替身被制造出来了。但——这是最微妙的——此时凯特仍然知道自己在做什么。她会对自己说：我不得不这样，等考试过了，我再去真正教他。替身在她心里仍然是一件她“被迫制造”但“并不认同”的东西。这是替身制造完整链条的第一步：被迫执行。

但这一步不是终点。接下来发生的事，才是替身制造得以持续、复制、强化的真正的生成机制齿轮：凯特被制度反馈慢慢地从“被迫执行者”改造成“替身的自愿捍卫者”。

第二个场景发生在几个月后。凯特的班级在年末全国测试中取得了比上一年更好的成绩，特别是在写作的结构组织维度上进步明显——这正是教学框架中连词教学所瞄准的评分点。校长在全校教师大会上表扬了她：“凯特的班级在写作结构分上提高了百分之十二，这说明她的教学方法非常有效。”地区教育顾问来学校做课堂观摩时，在她的课上看到学生整齐划一地使用“Firstly—Furthermore—However—Therefore”四个连接词完成段落写作，在评估表上打了高分，并邀请她在下一次全区教学经验分享会上介绍她的“连接词教学法”。

凯特感到一种她很久没有感受到的东西——职业上的被认可。在推行NLS的这几年里，她一直在“做教学框架要求我做的事”和“做

我觉得真正有意义的事”之间感到撕裂。但现在，制度告诉她：你做的那些事是有效的。你班上的法蒂玛，那个来自刚移民到英国的家庭、英语起初说不利落的女孩，因为学会了连接词模板而在写作考试中拿到了“达到预期水平”的评分。这是制度给出的“成功”的证据。

然后，那个关键的认知转变发生了。凯特回想起几个月前的那一天——她为了教连接词，打断了里奥关于恐龙的热情。那时她感到内疚。但现在，在新的制度反馈的累积作用下，她开始修正自己对这个记忆的理解。她对自己说：

“你看，我当初差点做了什么——差点为了那些看不见的东西，耽误了他们的考试。里奥喜欢恐龙，恐龙不会帮他在测试中得分。他的家庭不会在课后给他补习，他只能靠我。如果我当时心软了，让他去写什么恐龙——他的作文里就不会有这四个连接词，他就拿不到这个分数。幸好我按框架走了。”

在这个时刻，替身不再是一个凯特“被迫制造”的东西。它变成了她主动选择并愿意捍卫的东西。她不再把连接词教学看作一种在压力下的生存策略，而是把它内化为自己的职业判断——“对于这些孩子来说，这就是最好的教学”。此后，当一位新入职的年轻同事私下问她“你觉得我们教的这些东西真能帮助学生学会写作吗”，凯特不再犹豫或辩解。她用真诚的、基于经验的语气告诉那位新同事：

“我知道这看起来有点机械。但你得想一想，你班上那些孩子，他们毕业后是要去参加考试的。他们没有中产阶级孩子家里那些书和对话。你给他们连接词模板，就是给他们一个武器——让他们在考场上不被认出来是‘不会写的’。等他们进了中学、拿到了资格，再谈自由写作。这是对他们负责。”

这段话的每一个字都是真诚的。凯特不是在自欺欺人。她确实相信自己做的事情是对的。但正是在这种真诚中，替身的制造完成了闭环。替身的制造者，同时成为替身的第一位、也是最忠实的一位信奉者。她不再有能力从外部审视连接词教学——因为审视所需要的认知框架，已经在制度反馈的持续强化下，被“替身就是教育”这个新的认知框架覆盖了。她曾经感觉到的那个矛盾——制度要求与职业良知之间的撕裂——并没有消失。它在认知层面被“解决”了：不是通过制度改变自身，而是通过她从内部修改了自己对“什么是好教学”的判断标准。

这个微观分析揭示了一个生成机制上的关键机制：替身的制造，不发生在教师“想制造替身”的时刻，而发生在教师在认知上已经接受了“教学框架是正确的教学应当如何”这一前提的时刻。那个接受，不是某一刻突然的精神投降，而是一次又一次的“制度反馈—自我合理化”循环逐渐蚕食和重新校准她判断标准的过程。一旦凯特在认知上完成了这个转变，她就不再有能力识别替身了。她看到的，是她的学生在考试中取得了进步，她的教学方法被同事借鉴，她在职业上感到被认可。她看到的全部，都是真实的——数据是真实的，分数上升是真实的，法蒂玛的作文拿到了“达到预期水平”的评分是真实的。她看不到的是，那些东西之所以真实，正是因为它们本身就是替身的产物，而替身已经被她内化为判断“真实”的依据本身。

在布迪厄的意义上，凯特的“实践感”被制度重写了。原本，她作为一个有经验的教师，在面对里奥这样的学生时，能够凭着多年形成的那种即兴的、身体化的职业智慧，感知到他此刻需要的可能不是一个连接词，而是一个被认可的热情。但教学框架和问责制度对她日复一日的塑造，逐渐将她从一个即兴的判断者，改造成了一个“教学框架操作的熟练执行者”。她的身体化倾向系统——惯习——被新的制度场域重新塑形：她不再是那个能在课堂上与学生一起进入“不知道”的教师，而是成为一个能精准执行“读写小时四个环节”并产出可量化结果的高效操作者。她不是在被迫服从制度——她已经是制度在她自己生命里长出的新器官。这个器官的功能，就是稳定地、高效地、真诚地制造替身。

四、替身的自证循环：为什么系统看不见自己在制造替身

如果替身只是某一个特定改革的意外副产品，那么它的危害是有限的。真正的问题在于：替身具有一种极强的自证倾向——一旦它在某个改革现场被制造出来并被评价体系认证，它就能自我复制、自我强化、自我合理化，形成一个难以从内部打破的死循环。如上一节微观分析所示，这个循环在个体层面（凯特）和制度层面（评价体系、政策强化）之间存在一个接口：制度反馈不断为个体提供“替身是有效的”的证据，个体在认知上逐步接受并将替身内化为自己的职业判断，从而更忠实地执行制造替身的制度要求，产出更多的替身数据——这些数据又成为下一轮制度强化的依据。

（一）回路第一段：干预制造替身

改革者带着“应当是”的蓝图进入教育现场。在NLS中，这就是那份精确到分钟的教学框架。教师和学生面对这套外部规程，做出的是理性的适应性回应。如前所述，一位NLS期间的教师在调查中这样描述她的处境：政策文件一方面要求她根据学生的需要灵活调整教学，另一方面又要求她必须确保覆盖教学框架中的所有知识点。在学校的问责压力下——尤其是当学校被列入需要特别监测的名单时——她被要求提交周教学计划，详细证明她如何在每一节课里落实教学框架的要求。结果，她不得不把真正能激发学生兴趣的阅读材料放到一边，因为它们与当天必须教的语法点无关。

这位教师不是不敬业。她是太敬业了——敬业到在系统施加的要求与她对自己职业意义的感知之间，她已经打了一场持续几个学年

的、徒劳的内在战争。她最终选择的生存策略——按系统要求的格式来教，在评价中获得合格考核，同时心里为“真正的教育”留一个被搁置的角落——就是替身的制造在个体生命层面的微观发生。学生学会的不是写作，他们学会的是在连接词的位置放连接词；教师交付的不是教学，她交付的是在教案检查中合格的备课记录。那个本应在师生之间发生的东西——一个生命在另一个生命的注视下，用语言去触碰到某种真实——从一开始就没有被条件允许发生。它不是被批评了、被贬低了、被排挤了——它是根本没有机会进场，因为那四十分钟的每一个缝隙都已经被图纸填满了。

（二）回路第二段：替身被误读为成效

替身一旦进入评价体系，就发生了一次致命的信息过滤。量化评价——在NLS中是标准化测试的分数与等级——只能捕捉替身中被编码的那一面：作文中连接词的出现频率提高了，测试总分上升了。替身的代价——学生从未在写作中真正进入自我表达的状态，教师对职业意义的感受正在萎缩——完全无法进入这套以数字和等级为媒介的评价语言。

这不是一个主观欺骗的过程。教育部门官员、评估机构研究者、学校管理者中的大多数人，都是真诚地相信这些数据指向了读写教学质量的提升。但真诚地相信，并不妨碍这种相信本身是替身的产物：因为替身被设计的初衷，就是要在可观测的产出维度上看起来像教育。它在这一点上成功了，非常成功。数据上升是真实的——就像一个人造花的颜色和形状是真实的一样。问题不在于数据造假，问题在于数据无法识别它所测量的是花还是人造花。

（三）回路第三段：被合法化的强化

一旦替身被官方评价体系认证为有效，制造替身的元逻辑便被激活了。因为系统看到测试分数的上升——它无法将其解释为“教师在为生存而代偿”，它的认知范畴里没有这个选项。它只能将其解释为“教学框架正在发挥作用”。而当它看到分数虽然上升但仍未完全达标时，它也只能将其解释为“教学框架还不够精确”。

在这两种解释的逻辑终点，结论都是同一个：需要更细致的诊断、更精确的教学指引、更多的数据采集点。于是，在2006年——NLS实施八年后、测试分数进入平台期之时——英国政府推出了“革新版读写教学框架”。这个版本保留了读写小时的基本结构与时间分段逻辑，但将知识点更细致地分层归类，将教学建议更精确地匹配到每一个评分等级——一份在旧版替身基础上设计出来的、更精细的新版替身制造图纸。

从制度逻辑内部看，这不是非理性的顽固，而是完全理性的管理优化。强化干预是所有可用选项中，唯一与“我已经知道教育应当是什么样”这个先验兼容的选项。承认自己不知道——承认那份蓝图可能从一开始就不应该存在——从来不在制度理性备选项的清单里。

（四）回路第四段：替身制度化

最隐蔽也最具破坏性的一步发生在最后。当新版更精细的框架被引入教学现场时，它所干预的对象，已经不再是改革最初试图改变的那个“原始教学现场”。它的对象，是经过了八年旧版替身改造的、已经被“三段论模板”和“连接词伪装写作”浸透了的课堂。新一轮改革的设计，是在前一轮替身所产生的数据和分析基础上完成的——它依据的测试结果本身就是替身的产物，它参照的教学难点本身也是上一轮模具挤压后的变形。

因此，新框架不仅没能退回到替身被制造之前的那个节点去重新思考，反而在替身的断层上继续施工——在已经被框定的写作教学上增加更多框定，在已经被格式化的课堂上引入更细密的格式。旧的替身被当作“常态”锁进制度结构，新的替身在此基础上继续生长。至此，“干预制造替身→替身被认证为成效→认证合法化强化干预→强化干预制造更精细的替身”这条完整的自证回路，就闭合了。它不再需要任何外部的推力，因为它在自己体内产生维持自己运转的全部动力。

一个系统，当它用替身来证明自己有效时，就已经不再具备发现替身的能力了。不是因为它的成员愚蠢或虚伪，而是因为替身占据了“有效”这个判断本身的位置。你无法用替身的标准去检测替身——就像一个用假钞验钞机的人，永远不可能发现假钞。

值得注意的是，替身的自证循环不仅在单个制度体系内部运转，它还可以通过跨国的政策学习与政策转移机制向外扩散。当一个国家的标准化改革在评价数据上“取得成功”时（分数上升），替身产品——可量化的测试分数上升——便作为“政策成功的故事”被其他国家的改革者援引和移植。而这些“成功故事”的传播，进一步强化了那个底层先验在跨国范围内的制度合法性：既然那个国家用标准化教学框架提升了读写成绩，那么我们也应当这样做。跨国政策学习在这个过程中将“替身被误读为成效”的逻辑在更大的地理尺度上复制——这正是“脱耦”或“窄化”等概念无法捕捉的分析维度，因为它们只描述单一制度内的变形，而不追踪替身如何通过数据化的替身产出在制度之间自我传播。

五、“守护”的陷阱：为什么连批判者也无法摆脱制造替身

在完成案例深描与回路分析之后，本节将替身与最有深度的一类批判进路进行区分，以确立替身概念不可还原的最后一道防线。这不仅是一个学术正名的需要，更触及一个关键问题：为什么此前的批判理论——尽管在很多方面比标准化改革深刻得多——同样无法摆脱替身制造的陷阱？

（一）从“修理”到“守护”：批判走在正确的路上，但没有彻底

在教育批判的传统中，有一类论证已经将靶心从“改革的外部阻力”移向了“改革者的认知操作图式”本身。这类论证的核心判断是：标准化改革者的根本错误，不在于他们用错了教学方法，而在于他们对待教育的方式本身就是错的——他们用修理一张桌子的方式去对待教育。桌子是一个制成品，它的结构和功能在设计阶段就已经被完全给定，修理者的全部智慧在于诊断故障、拆解部件、按图纸替换，最终恢复它的预设功能。但教育不是一个制成品——它是一个生命在不断生长的过程。把教育当桌子修，就是在用一套预设的蓝图（“应当是”）去强制一个活的过程符合一个死的模板。这是一种根本层面别的错位。

这个判断已经比“应试教育害人”“制度惯性阻碍改革”等流行诊断深刻得多。它不再问“什么在阻碍改革”，而是问“改革本身的方式是否就是问题的构成部分”。它已经走到了那扇门前。

但它停在了门前。这类批判的有效性，依赖于一个隐藏的对照项：既然按照蓝图“修理”是错的，那么正确的做法应该是“守护”——像园丁照料一棵树的生长那样。园丁不去预设树应当长成什么形状；他去阅读树的生长倾向，为它提供适宜的土壤、水分、光照，守护它按照自己内在的倾向去生长。在教育哲学的传统中，诺丁斯的关怀伦理对“关心”作为一种教育本质的论述、范梅南对“教学机智”作为实践智慧的肯定，都在言说一种不同于技术理性操控的教育姿态——一种更柔和的、更贴近学生生命节奏的“守护”。这些论述极其重要，它们为那些在被标准化管理所磨损的教师提供了一种道义上的支撑，告诉他们：教育不是、也不应当是修理。

然而，守护的姿态尽管比修理柔和得多、谦逊得多，但它仍然——在最深层、最不易察觉的层面——站在那个本文第二节所揭露的底层先验上。因为它仍然预设：守护者知道被守护的生命需要什么。

园丁知道树需要水、阳光、肥料。这是真的。但一个教育者知道一个正在学习写作的孩子“需要什么”吗？如果他以为他知道——他知道那个孩子需要“被激发写作兴趣”“需要接触到好的文本范例”“需要在反馈中感到安全”——那么他仍然带着一份“应当是”进入了那个孩子的成长现场，只不过这份“应当是”比标准化改革者的那份更宽泛、更有弹性、更贴近“教育规律”。但它仍然是一份被前置给定的关于“好”的知识。当教育者以这份知识为基础去“守护”学生时，他已经在替那个学生决定了：你应当朝着这个方向生长。

这里有一个极细微但致命的分寸。照料一棵树是一回事，照料一个人的生成是另一回事。树没有主体性——或者说，树的主体性不构成对园丁认知框架的挑战。树不会因为园丁以为它需要某种肥料，而改变自己对土壤的真实反应。但人会。一个学生，当老师以“我要守护你的读写能力生长”的姿态介入时，他会回应——他会按照老师期待的方向做出表现，他会写出老师喜欢看的文字，他甚至会在老师的守护中感到温暖和安全。但这些文字里有没有他真正想说的话，他从来不曾问过自己——因为在“被守护”的整个过程中，从来没有人告诉他：你可以不知道你要写什么，你可以坐在那发呆、挣扎、自我怀疑，然后从那个不舒服的地方，自己长出第一个真正属于你的句子。守护者太想护住他了，反而替他把那个最关键的、必须由他自己去闯过的“不知道”的阶段给跳过了。

这就是“守护”的陷阱：当一个教育者被“守护”的善意装备起来、进入教育现场时，他永远倾向于把学生的“困难”解读为“需要被移除的障碍”——于是他去提供支架、策略、安慰、引导。这里需要一个精确的区分，以避免本文对“守护”的批评滑向对“提供支架”这个具体教学动作的否定——那恰恰是本文反对的“立新的应当”。问题不在于教师“提供了支架”这个动作本身。问题在于：当“提供支架”成为一种被制度性的前置知识（“你应当在此刻用此种方式提供支架”）所指导和审查的动作时，它就从一位教师在即兴判断中可能选择也可能不选择做的实践性回应，变成了一个必须被执行的操作。执行这个操作本身，挤占了教师去观察、去等待、去在那一刻判断“这个学生究竟需要什么”的空间——甚至，挤占了教师去承认“这一刻我其实不知道他需要什么”的可能。替身的制造，发生在这个制度性的认知结构层面——它让教师无法觉察到“学生此刻可能不需要支架，他可能只需要被留在那片空白里”这种可能性的在场。本文批评的不是“守护”作为一种教育方式，而是“守护”作为一种被前置给定从而封闭了上述觉察的认知框架。因此，本文的“撤回”主张也不意味着“教师不应提供支架”——它只意味着，“提供支架”不应当作为一套制度性的“你应当”被前置给定。

（二）“替身”与既有的“修理批判”的分界线

至此可以精确地标定“替身”与这类“以修理为靶心”的批判进路的分界线。

这类批判命名的是一种错误的方式——用修理制成品的思维去对待生成之物。它的理论重心在“改革者”身上——诊断的是改革者认知工具箱的错配。它仍然暗示了一种正确的替代方案（“照料”“守护”“园丁的智慧”）。

“替身”命名的是一种错误的产物——在“修理”与“守护”这两种方式之下，都可能被制造出来的那个不是教育却被当作教育的东西。它的理论重心在“教育现场”身上——命名的是那个本该由真实生长占据、却被模拟形态顶替了的生态位。

更关键的区别在于：这类批判暗示“守护”是比“修理”更正确的做法，而“替身”概念揭示的是：那些被认为是正确的替代方案——用更宽的、更弹性的、更贴近“教育规律”的“应当是”去“守护”——同样可能在制造替身。只是材料不同、工序不同、替身在数据上呈现的面貌不同。修理者制造的替身，叫“高效课堂”，特征是整齐、可控、产出明确；守护者制造的替身，叫“有温度的课堂”，特征是民主、对话、学生主导——但如果那里的学生依然从未经历过“被抛入自我怀疑、并在无人可代劳的情况下独自做出判断”的那一刻，如果那里的师生互动依然在温柔地绕开“我不知道你该长成什么样”的那片深渊，那么守护者制造的，同样是替身——一个更优雅、

更让人舒服、更难以被识破的替身。

替身概念的不可还原性，在此获得了最终的确立：它不与“修理”对立。它与“有人在现场真正地生长了”对立。这个对立，比修理与守护的对立更根本。因为修理和守护都是教育者的动作，替身和生长都是教育现场的结果——而一个动作是好是坏，不取决于它看起来像哪一种，而取决于它最终在另一个生命中结算出了哪一样。

六、最诚实的反驳：替身难道不是一种必要的“公平”吗？

在完成理论的建构与案例的深描之后，本文必须直面一个最有力、也最诚实的反驳。这个反驳不是来自理论对手的思辨攻击，而是来自一位站在教育第一线的实践者。假设这位教师——叫她萨拉——在读完本文后，坦诚地提出以下质疑：

“我理解你说的‘替身’。我甚至在我的课堂上看见过它。但你告诉我，我的学校在伦敦东部，我班上有三十五个孩子，其中超过三分之一来自非英语母语家庭，有些孩子的父母不怎么识字。我每天只有五十分钟的读写课时间。我必须有一个教学框架，必须有一些清晰可测的目标，否则我连最基本的公平教学都无法保证。我承认，我可能让里奥眼里的光暗了下去，我可能让一些孩子的自我卷入打了折扣。但那个来自刚移民到这个国家的家庭、连英语都说不利落的法蒂玛，她需要连接词。她需要有人明确地告诉她，一篇合格的作文长什么样，考试会怎么评分。因为她的父母无法在这方面帮她，而她的中产阶级同学在家里就能获得这些知识。我教她连接词模板，让她在考试中拿到还过得去的分数，让她有资格进入下一阶段的学习——这难道不是一种公平？如果按照你的逻辑，撤回这些框架，给‘不知道’留出空间——在那些‘不知道’的空白时间里，谁去帮法蒂玛？她的学习机会难道不是被那片空白进一步剥夺了吗？你是不是在为一个个精英班级里那些本来就爱读书、本来就敢写作的孩子辩护？”

这个反驳直指本文在政治伦理与实践可行性层面最要害的问题：教育公平。如果替身概念的逻辑终点是要求撤回外部标准、要求教育者为“不知道”留出空间，那么这一逻辑是否——无论其本意多么良善——最终会以牺牲弱势学生为代价？在回应这个反驳之前，必须先承认它的重量。这不是一个可以被理论优雅地绕过去的问题。它触及的是每一个真实的教育者每天都必须在其中做出抉择的道德张力：一边是系统性保障基本学业标准的责任，另一边是对每个个体生命生长节奏的尊重——这两者之间，没有一个可以被提前写好的完美平衡点。

以下，本文试图给出一个回应。这个回应不是要“驳倒”萨拉的质疑——她的质疑中那些坚实的成分，任何一个诚实的教育研究者都无法、也不应该试图驳倒。本文试图做的，是论证：替身的制造，最终会以一种更隐蔽、更不公正的方式，剥夺萨拉试图保护的那些弱势学生的长远发展权利。同时，本文将提供实证证据来夯实一个关键前提，并在最后援引那些已经在不制造替身的前提下有效服务于弱势学生的真实实践，证明“撤回替身制造机制”并不是用遥远的未来许诺换取当下可见的损失。

（一）替身如何制造一种更隐蔽的不公平：超越逻辑推演的证据

萨拉的论证基于一个前提：教法蒂玛连接词模板，是给了她一种实际有用的能力——一种她可以用来通过考试、从而获得进一步教育机会的能力。这个前提不完全错。在短期内，在当前的制度安排下，法蒂玛确实从连接词模板中获得了分数上的收益。然而，问题在于：这个收益的保质期有多长？更重要的是，当法蒂玛凭借这些模板通过了考试、进入了中等教育乃至高等教育的下一阶段时，等待她的是什么？

关于标准化应试训练对后续学习能力的长期影响，有多项教育追踪研究提供了值得认真对待的证据。在美国，已有研究考察了“为考试而教”对学生写作能力发展的深层影响。一项基于肯塔基州写作组合评价体系的长期实证研究发现，虽然学生的写作分数在短期内可能因针对性的考试技巧教学而有所提升，但这种提升对学生在需要真实读写能力的后续任务——包括大学阶段的学术写作——中的表现“几乎没有预测力”。更令人警惕的是，多项关于“写作教学与考试准备”的研究综述指出，过度聚焦于标准化写作测试格式的教学，会系统性地压缩学生在真实写作中练习构思、修改和自主表达的机会，从而在长期内削弱而非增强学生的写作能力。这种削弱，对于弱势学生而言代价尤甚，因为他们不像中产阶级同侪那样可以在校外获得阅读和写作的文化资源来弥补校内替身化教学所造成的练习空白。

另一组证据来自对“技能偏向型课程”（即以教授可测试技能为目标的课程）长期效应的研究。这些研究发现，虽然技能训练在短期内可以提高测试分数，但学生在需要高阶思维、批判性写作和独立判断的学术任务中，并没有表现出相应的能力提升。分数上升了，但可迁移的素养并没有同步增长。这就是说，短期分数的提升，并不能有效地转化为长期学术能力的增长。

这些证据共同指向一个本文称为“时间套利”的核心机制：替身在今天加给法蒂玛一个可测量的短期收益（考试及格、分数上升），但以此为代价，剥夺了她在未来真正需要读写能力时——在中等教育的学术写作中、在高等教育的研究报告中、在任何不按“五个连接词加一个结论句”格式来评分的真实表达情境中——能够胜任的那些更深层的能力的发展机会。套利的成本，即那个长期能力断层，在时间上被延后了；而当它最终暴露时，法蒂玛早已过了基础教育阶段，不再有任何补救机会。

更关键的是，这种套利成本不成比例地落在了弱势学生身上。中产阶级家庭的孩子，在替身之外有其他资源可以部分弥补——家庭藏书、父母的教育参与、校外文化体验——来维持某种程度的真实读写接触和社会资本积累。而法蒂玛没有这些资源。对她而言，学校提供

的读写教育，就是她所能获得的全部。如果这全部的教育，都被替身所占据，那么她在理论上获得了一个公平的分，却在实质上被剥夺了最后一次获得真实读写能力的机会。她带着那个“达到预期水平”的评分标签离开了小学，但在下一个没有连接词模板可依傍的学术情境里，她发现自己的中产阶级同学——那些可能在更宽松的写作环境中练习过表达自己——能写出有观点、有论证的文字，而她只能回忆起“第三段开头要用However”。那个替身制造的公平，在此刻露出了它的底色：它不是在缩小差距，它是在把差距从一个可见的场域（当下的考试分数）转移到了一个不可见但更具决定性的场域（未来的真实能力），同时在转移过程中让差距进一步扩大了——因为那些有校外资源的学生，在替身之外仍然触碰过真实，而法蒂玛触碰过的全部，就是替身。

（二）“不知道”的空间不是特权者的奢侈品，而是公平的前提

萨拉的担忧中还有一个深层预设需要被检视：她将“不知道”的空白时间视为一种风险——它可能被那些“本来就会”的学生占据，而让那些“本来就不会”的学生更加落后。这个担忧是真实的，但它忽略了一个关键的生成机制事实：真正能改变一个弱势学生学习轨迹的，往往不是另一个设计得更好的教学框架，而是教师在被允许“不知道”的瞬间，对那个学生真正的、不可被图纸预判的学习状态所做出的即兴回应。

法蒂玛需要的可能不是——或者说，不仅仅是——更多的连接词训练。她需要的，可能是一个有空间的课堂——一个没有被“必须在十五分钟内覆盖三个连接词用法”所完全填满的课堂——在那里，教师有时间和制度许可去观察她、去问她“你想写的是什么”、去在她写出的那个笨拙但真实的句子里找到那一点微光并告诉她“这里，我看到了你”。这种回应，无法被写进教学框架，无法被提前写进教案，无法被纳入课堂观察量表。它只能发生在一个教师没有被迫在那一刻优先考虑“我有没有完成教学框架指定的动作”的时候。而一个被“应当是”全面管制的课堂——无论它管制的是一套标准化的连接词模板，还是一套进步主义的探究步骤——都系统性地消灭了这种回应得以发生的条件。在这类课堂里，教师的目光已经被图纸预设的“应当”所占据；她没有余裕去看那个不在图纸上的、法蒂玛写出的笨拙句子里藏着的那一点真实。

（三）短期公平如何可能：不制造替身也能有效服务弱势学生的实践

萨拉还会追问：“你说替身最终损害了法蒂玛，但如果你让我撤回标准化框架、为‘不知道’留出空间，法蒂玛在五年级结束时连基本的读写能力都没有，她怎么进入下一阶段？你的‘长期公平’需要一个‘短期可存活’的路径来落地——否则你的论证就是在用一个遥远的未来许诺，换取当下可见的损失。”

这是一个完全正当的追问。本文不试图给出一个新的教学框架（那是本文立场禁止的），但可以指出：存在一些已经经过实践检验的教育生态，它们在不依赖替身制造机制的前提下，同样——甚至更好地——保障了弱势学生的读写能力发展。

芬兰的基础教育体系提供了一个有参照价值的案例。芬兰在国家层面取消了义务教育阶段的外部标准化测试，代之以教师主导的描述性反馈评价。教师不按一份全国统一的“读写教学框架”去执行每一课的教学动作，而是在课程纲要的宽泛指引下，依据对自己学生具体学习状态的长期观察，即兴地设计教学。在这个体系里，芬兰学生的读写能力在国际比较中长期名列前茅，而且校际差异和家庭背景对学生表现的影响显著低于多数发达国家。这一事实表明：撤回替身制造机制（标准化测试、教学动作框架、外部问责），并不会必然导致弱势学生的读写能力下降——恰恰相反，在一个给予教师专业自主和制度信任的生态里，教师反而更有余裕去回应那些具体的、无法被图纸预判的学生需要。

另一个值得注意的实践来自美国基于表现的评价改革运动中涌现的某些学校。在这些学校，毕业评审不再仅凭标准化测试分数，而是要求学生完成一份包括学术论文、科学探究项目、数学问题建模和公开答辩在内的“作品集评审”。这些学校的教师花了大量时间与学生进行一对一的写作辅导——不是教他们如何在一篇作文里塞进四个连接词，而是和他们反复讨论他们究竟想写什么、哪一段的论证还不清楚、如何找到更合适的证据。这种教学不制造替身——因为每一次写作，学生都是从自己真正想问的问题出发，并在教师的个性化反馈中不断修改和深化。而这类学校的学生，包括那些来自低收入家庭的弱势学生，在大学阶段的学业表现和持续就学率均显著优于接受传统标准化教学的同侪。

芬兰的表现评价体系及其学生素养长期优异表现的证据、以及上述基于表现的评价改革实践的跟踪研究，为本文的论证提供了一个关键的支点：撤回替身制造机制，不是一种“美好但不现实”的理论想象，而是在某些制度条件下已经被部分实现的教育现实。本文不是要求萨拉在一夜之间推翻整个英国教育体系、变成芬兰，而是请她看见一个她此前可能从未见过的可能性：法蒂玛需要的不是替身，她需要的恰恰是一个不被替身占据的课堂——在那个课堂里，有一个没有被迫在某一分分钟内必须讲完连接词第三个用法的老师，愿意在她写出的那个笨拙但真实的句子上停下来，告诉她，这里，你想说的话，我听到了。

七、从替身中撤出：教育治理的减法转向

如果说理解“替身”的存在是第一步，那么第二步就是：面对这个事实，我们怎么办？本文的最后一部分试图回应这个问题，但必须采取一种不自毁诊断立场的方式。替身概念的逻辑直接禁止了“设计一套新方案来取代旧方案”——因为任何“方案”本身，就是一个替身制造装置。它一定预设了对“应当如何”的某种知识，一定要求执行者按照特定方式去呈现教育，一定会催生新的替身来满足评估要

求。所以，真正的问题不是“如何设计更好的替身”，而是：如何从制度层面逐步撤回那些持续不断制造替身的机制？

（一）撤回“应当是蓝图”的制度装置

制造替身的第一工序，是给定一份“应当是”的图纸。那么，撤回的起点，就是拆解那些让这份图纸获得制度权威的装置。具体的制度动作至少包括以下几种。

大幅降低高利害标准化测试在整个评价体系中的权重。这不是说取消所有考试，而是改变考试的功能定位——从终结性的、用于贴标分等的外部裁判，后撤为服务于教与学内部诊断的信息来源。芬兰在国家层面取消义务教育阶段外部标准化测试、代之以教师主导的描述性反馈评价体系的实践，为此提供了一个虽然不能被简单复制但方向明确的参照。当一个课堂不再需要为一个外部测试的结果负责时，那个测试所携带的“应当是”蓝图就失去了强制力——教师和学生就少了一个必须制造替身的理由。

将评价权从单一的中央机构下沉至学校和教师共同体，并赋予过程性、叙事性的评价以制度权威。如果一间学校可以用对学生长期学习历程的持续记录、对典型作品的累积性评审、对学习品质的质性描述来向外界证明自己的教育品质，那么它就不需要让每一节课都产出可被即时量化的替身产品。美国一批公立学校在基于表现的评价联盟中采用毕业论文答辩和作品集评审作为毕业要求的实践，提供了这样一种替代性评价制度的存在证明——它证明了评价可以被设计为一种学习的内在组成部分，而非学习的替身制造终端。

但需要坦率承认的是：改变评价制度只是减少了制造替身的压力，并没有消除替身制造的全部根源。因为即便外部的“应当是”压力被完全撤走，教师自己头脑中仍然可能有一套“应当是”——关于什么是好学生、什么是好作文、什么是好课堂的内化标准。只要这套标准还在他走进课堂之前替他决定了他和学生应该怎么互动，替身的制造就不会停止。所以制度撤回只是前提，不是全部。

（二）为“不知道”留出制度空间

如果替身制造的根因在于改革者（以及教师、家长、以及整个社会）对“教育应当是什么”的前置知识，那么对抗替身的最根本动作，不是提供另一种关于“教育应当如何”的理论——那只会制造新的替身。最根本的动作是：在制度设计中，为“我不知道”创造合法的、被保护的、不被问责侵蚀的空间。

这意味着什么？意味着在课程安排中，为那些无法被预先写进教案的东西留出不被侵占的时间——不设具体教学目标的自由阅读时间、不要求产出的探究时间、不允许教师干预的学生自主讨论时间。意味着在教师评价中，不把“每一节课都有清晰可测的目标”当作优质教学的黄金标准——而是允许、甚至鼓励教师说出“这一节课我本来准备了一套方案，但学生提出了一个我没想到的问题，我们花了一整节课去探讨，我不知道他们从中学到了什么具体知识点，但我看到他们在思考”。意味着在教师培养中，不把教师训练成“掌握了全套应当如何”的专家，而是把他们从“必须总是知道该怎么做”的焦虑中解放出来，让他们敢于在学生面前承认“我也不知道，我们一起来想想”。

这是一个极其困难的要求。因为它不是在要求制度多做些什么，而是在要求制度少做一些事——少给一份指导、少设一个指标、少做一轮督查。在一个以“作为”证明自身合法性的治理文化中，“少做”比“多做”难得多。多做是在展示能力，少做是在暴露无知。而教育治理中最昂贵的无知，恰恰是对自身“不知道”的不承认。这种不承认，不表现为一句说出口的“我全知道”，而表现为一系列制度化的行为：编订越来越细的课程标准、设计越来越精密的课堂观察量表、制定越来越多层级的数据采集指标——这一切都在说同一句话：我们不需要承认不知道，因为我们可以测量一切。

撤回这些制度动作，不是放弃对教育质量的责任，而是将责任从“确保教育以特定方式被呈现”重新定义为“确保教育没有被某种特定方式阻碍其发生”。这两种责任之间的差异，是替身制造与替身撤回之间的分水岭。

（三）长期方向：不是推演，是生态演化

本文不试图提供一个“替身零制造”的乌托邦方案——那本身就是一个替身。本文试图做的，是改变看教育治理问题的那双眼睛。当我们用“替身”这双眼睛去看时，出现在视野里的就不再是“如何设计更有效的改革”，而是“一个具体的制度装置（一条课程标准、一套评价指标、一份课堂观察表）是否存在——无论它有多好的意图——持续地要求教育以某种特定方式被呈现、被证明、被认证”。如果答案是肯定的，那么它就是一个替身制造机。它越是精细、越是科学、越是获得高分，它制造的替身就越是难以被觉察。

这导向一个看似消极实则艰难的方向：不是去推演一场更合理的改革，而是去培育一种不依赖替身就能自我运转的教育生态韧性。这意味着信任那些分散的、局部的、由具体教师和具体学生在具体互动中自发生成的教育实践，信任它们在没有统一施工图纸的情况下所具有的自我更新能力，信任那些没有外部标准认证的教育发生——一个学生在某节课上因为一个没有被写进教案的追问而真正理解了什么，一个教师在一次没有评价指标的对话中重新感受到了职业的意义。这些事在数据上不可见。但它们就是教育发生时的样子。而制造替身的那整套制度逻辑，恰恰是因为看不见这些事——或者更准确地说，因为不信任这些事——才需要用替身来替代它们。

本文的最终主张，因此不是任何具体的政策方案。它是一种方向的转换：从不断追问“我们还能做什么来改进教育”，转向持续追问“我们在做的哪些事，正在以改进之名，系统性地制造着不是教育的东西。”这场转向不是一次性完成的，而是一场没有终点的、持续的

自我审查。它的核心能力，不是知道应该做什么，而是辨认出自己正在制造替身的那一刻——并停下来。

八、结论：承认那座看不见的替身

本文的核心工作，是提出并论证了一个此前被所有教育改革研究遗漏的生成机制概念——替身。它指称的不是教育的失败、扭曲、窄化或异化，而是一个更隐蔽的事实：在那些真诚、努力、数据上成功的教与学现场，教育本身被另一个东西替换了。那个替代品——替身——以像教育一样的方式运转：教学动作被执行了，学习任务被完成了，测试分数上升了。但它结构性地删除了教育之为教育的那一件事：一个生命在无可代偿的自我卷入中，生长出此前不曾有的东西。

替身被制造的根源，不是改革者的恶意，不是制度惯性，甚至不是技术理性或规训权力。它是一个被所有批判者和改革者共同分享的、从未被真正撤销的预设：教育者能够——而且应当——在他进入教育现场之前，知道教育应当以何种方式、朝何种方向、在何种标准下发生。正是这个预设，使得一切改革——无论是以蓝图去“修理”的标准化运动，还是以智慧去“守护”的进步主义实践——都天然携带着制造替身的倾向。因为只要这个预设成立，改革者就必须为自己的知识寻找一个可以被看见、被测量、被认证的“兑现”——替身，就是那个兑现。英国国家素养策略的案例为此提供了经验的血肉：从一张精确到分钟的教学蓝图，到一份符合评分标准但主体缺席的学生作文样本，到一个教师在制度反馈中逐渐从被迫执行者变为替身自愿捍卫者的认知转变，再到一个自我强化、自我合理化的制度性自证循环——替身制造的每一个工序，都在这个案例中获得了具体的、可追踪的形态。而针对基层教师公平性质疑的回应则表明：替身的制造不仅没有真正服务于公平，反而以一种时间套利的方式——有追踪研究证据支撑的方式——将不公平从可见的考试分数场域转移到了不可见但更具决定性的真实能力场域，最终不成比例地损害了它声称要保护的弱势学生。芬兰和基于表现的评价联盟的实践则提供了证据：撤回替身制造机制，并不会必然导致弱势学生掉队；相反，在不制造替身的生态里，弱势学生更有可能获得那些真正能支撑其长远发展的读写能力。

替身概念的理论贡献，不在于提供了一种“比既有理论更全面”的解释——那依然是在既成事实论的框架里做加法。它的贡献在于：它从撤销一个被所有既有理论共享的底层先验开始，让一个新的事实第一次被看见。这个事实不是“教育被伤害的方式比我们想象的更复杂”，而是“在那个我们一直以为是教育被伤害的位置上，占据着的，是另一个东西”。这个发现不增加既有知识的数量，而是改变看教育的眼光本身——它让那种熟悉的、难以名状的不对劲第一次有了名字，也让以往被“课程窄化”“为考试而教”“脱耦”等概念统摄的描述有了一个更根本的生成机制锚点。

本文的论证可以被考验。如果在某个以标准化和问责为特征的改革案例中，可以观察到学生的读写能力不仅在测试中表现优异，更在独立于测试语境的、需要自主判断和创造性表达的陌生情境中持续表现出这些素养，且教师群体能普遍维持其教学中的创造性、自主判断力和对学生个体差异的敏锐回应，而未被系统性地导向为应对问责而压缩课程的实质性深度、用标准话术替代即兴互动、或以牺牲学生真实的学科兴趣为代价来追求可量化的产出——那么本文关于替身会系统性地取代真实学习的核心判断就需要被修正。如果在某个案例中，来自弱势背景的学生在接受了以标准化测试为导向的教学之后，在后续的教育阶段中并未出现本文论证的“能力断层”——即他们能够与其中产阶级同侪一样，在需要独立判断和创造性表达的学术任务中胜任——那么本文关于替身制造之公平代价的论证就需要被重新检视。本文的论点不是不可错的——它要求自己处于可被否決的风险之中。

但直到那样的反例被扎实地提供之前，替身的幽灵将始终盘旋在每一场以“知道应当如何”为起点的教育治理上空。从替身中撤出，不是学会一种新的教育方法，也不是找到那个“正确的”应当去替代“错误的”应当。它是承认——承认我们有太多次，在那些以为教育已经发生的时刻，其实只在照料着一座动作齐全、数据漂亮、唯独没有人的替身。而那个本应那一刻用自己的声音说出第一句真话的生命，就在离替身不远的地方，独自沉默着。

后续的研究，如果用“替身”这双眼睛去看，至少可以在以下方向上推进。第一，对更多样化的教育改革案例进行替身制造的比较生成机制分析——在东亚的高考驱动型改革、北欧的现象教学实验、美国基于共同核心州立标准的改革等不同的制度土壤中，替身的制造工序与呈现形态有何异同？第二，对“撤回了替身制造机制”的教育现场进行长期追踪——那些被赋予了“不知道”的制度空间的学校和教师，实际发生了什么样的教育？他们的学生在后续的学习生涯中，是否表现出了不同于替身制造环境下的能力结构与学习品质？第三，也是最具挑战性的方向：替身概念能否被迁移到教育之外的治理领域？在医疗、社会福利、公共政策绩效评估等领域，是否同样存在着那种以“可被管理的像”去兑现“应当如何”之预设的替身制造机制？这些追问，是替身概念打开的新的问题空间。它们的答案，不是本文能够给出的——但那正是本文的贡献所在：让这些问题，第一次能够被精确地提出。

参考文献

- Cuban, L. (1984). *How Teachers Taught: Constancy and Change in American Classrooms, 1890-1980*. New York: Longman.
- Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher*, 19(1), 3-13.
- Dewey, J. (1938). *Experience and Education*. New York: Macmillan.
- Earl, L., Levin, B., Leithwood, K., Fullan, M., Watson, N., Torrance, N., Jantzi, D., & Mascal, B. (2001). *Watching and learning 2: OISE/UT*

- evaluation of the implementation of the National Literacy and Numeracy Strategies. Department for Education and Skills.
- Earl, L., Watson, N., Levin, B., Leithwood, K., Fullan, M., Torrance, N., Jantzi, D., Mascall, B., & Volpe, L. (2003). Watching and learning 3: Final report of the external evaluation of England's National Literacy and Numeracy Strategies. Department for Education and Skills.
- Foucault, M. (1977). Discipline and Punish: The Birth of the Prison (A. Sheridan, Trans.). New York: Pantheon Books. (Original work published 1975)
- Habermas, J. (1987). The theory of communicative action, volume 2: Lifeworld and system: A critique of functionalist reason (T. McCarthy, Trans.). Beacon Press.
- MacBeath, J., & Galton, M. (2008). A life in teaching? The impact of government policy on teachers' working lives. National Union of Teachers.
- Matland, R. E. (1995). Synthesizing the implementation literature: The ambiguity-conflict model of policy implementation. Journal of Public Administration Research and Theory, 5(2), 145–174.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340–363.
- Ravitch, D. (2010). The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education. New York: Basic Books.
- Scott, J. C. (1998). Seeing like a state: How certain schemes to improve the human condition have failed. Yale University Press.
- Weber, M. (1978). Economy and society: An outline of interpretive sociology (G. Roth & C. Wittich, Eds.). University of California Press.
- Wyse, D. (2003). The National Literacy Strategy: A critical review of empirical evidence. British Educational Research Journal, 29(6), 903–916.

深化增补：被放跑的最近邻、操作化与证伪

以下四节为编辑增补，与作者正文分开标注。

增补一：与表演性、指标失效律的硬边界

本文最近的手不在教育学而在测量与治理研究：坎贝尔定律与古德哈特定律已论证指标一旦成为目标即失去其作为指标的效力，而表演性研究进一步论证被考核者会生产出符合指标形状的活动。这些是“替身”最直接的近邻，正文未处理。分野在替换的彻底性：指标失效律描述的是活动被扭曲——学习仍在发生，只是偏向了可测的部分；本文主张的是替换——出现了一种形似学习而实际不是学习的活动，它不是学习的偏斜版本，而是另一种东西。可判别面：若所有替身活动在撤除考核后能够回落为正常学习，则扭曲模型足够；若替身在考核撤除后仍作为默认活动形态存续（因为教师与学生都已不再掌握另一种做法），本文成立。这一条也是本文最该被检验的地方。

增补二：替身的可编码形式

替身若不可识别，本文的核心断言就无法与“教得不好”区分。三项判据：可迁移性缺失——该活动所产出的能力在稍微改变的情境中完全不可用；自主发起率——学生在无要求情况下主动进行该活动的频次（真学习会有残余，替身趋近于零）；教师可辨识度——教师能否指出这节课上哪些环节是为了证明而做、哪些是为了理解而做。三项同时出现才判定为替身；只有第一项，可能只是浅层学习。

增补三：可证伪的检验设计

核心命题——改革在追求可见成效的过程中系统性地生产替身——可通过一次自然实验检验。选取因政策变动而取消或大幅削弱某项考核的地区，比较取消前后该学科活动的三项判据变化。扭曲模型预测三项同时改善；本文预测第二与第三项改善滞后甚至不改善，因为替身一旦成为默认形态，撤除压力并不自动恢复原来的做法。若取消考核后三项迅速同向恢复，本文即被证伪。这一设计的价值在于它不需要新的数据采集，只需找到已发生的政策断点。

增补四：与同批《自油火场》的分野

作者同批已发表《自油火场：教育改革作为一种自我滋养系统》，两篇同处教育改革批判，必须切开。那篇解释的是失败为何持续扩张——机制是压制物被转化为下一轮改革的合法性燃料，分析单位是改革周期；本篇解释的是改革在单次执行中生产了什么——机制是替换，分析单位是课堂活动。两者构成宏观与微观两层，且可以互为证据：替身正是被自油火场识别为“亟需加强”的那

类对象。这一衔接关系值得明说，它使两篇互相支撑而非互相重复。

编辑增补所依据的核验文献

Donald T. Campbell, “Assessing the Impact of Planned Social Change,” *Evaluation and Program Planning*, 2(1), 1979.

Marilyn Strathern (ed.), *Audit Cultures*, Routledge, 2000.

Stephen J. Ball, “The Teacher's Soul and the Terrors of Performativity,” *Journal of Education Policy*, 18(2), 2003.

Michael Power, *The Audit Society*, Oxford University Press, 1997.

Gert Biesta, *Good Education in an Age of Measurement*, Paradigm Publishers, 2010.