

向评而生：静态评价何以成为一种“朝向”及其结构性挣脱

作者 黄倩盈 · SDE 学员 · 约 2.7 万字 · 德麦国际 SDE Universes

教育评价改革四十年反复陷入同一个坑：任何超越标准化考试的方案，落地就引发公平焦虑、被拉回分数比较。本文说病根不在缺更好的技术，而在一个没被点破的预设——评价必定服务于选拔、比较是成长的默认计量。

摘要

教育评价改革数十年来反复陷入同一困局：任何试图超越标准化考试的方案，落地时都会引发公平性焦虑，最终被拉回分数比较的旧轨。本文的诊断起点是：这一顽疾的根源不在于缺乏更好的评价技术，而在于一个从未被点破的底层预设——“评价必定服务于选拔”、“比较是成长的默认计量方式”。本文从中国教育体制四十余年的演化史中，追踪一对不可还原的新辨别维度如何在历史中生成：向评性（assessment-tropism）与向生性（growth-tropism）。向评性不是一种制度，而是一种朝向——它是个体、制度与资源在互动中集体偏向可被外部评价、可被比较、可被认证的趋势，在历次改革中被反复锻造为教育生态的首选朝向。向生性不是一种“更好的评价方法”，而是在向评性生成史中被挤压、被否定、被逆向定义出来的异质边界——它是认知过程中那些天然拒绝被即时判断、无法被量化为分数、却构成认知发生实质基础的部分，在与向评性的张力中被看见。二者的根本区别不是“用不用评价”，也不是“动机来自内部还是外部”，而在于：一个认知过程的发生，是朝着被看见的方向生长，还是朝着事物本身的内在逻辑生长。这不是“评价技术”的问题，这是在追问：外部注视介入认知过程时，被弯曲的究竟是什么——不仅是注意资源的分配，更是思考者心中那个“什么是好的思考”的元认知标准。本文论证：向生性不是向评性的改良版或替代方案，而是一块必须被从选拔逻辑中剥离出来的异质空间。真正有意义的问题不是“用什么评价方法去替换考试”，而是“能否在学校日常实践中，护住一小块不以任何形式进入比较的学生成长余地”。本文引入某高中“研究性学习”课程从自发探索到异化为履历军备竞赛的完整案例作为叙事锚点，让“形变”在真实的时空经纬中发生，而非在虚构的实验中演绎。同时正面回应三个最严厉的质疑——尤其是向生性空间可能被文化资本隐秘渗透的风险，论证切断“被看→被认证→被兑换”的链条后，文化资本在该空间内丧失了可兑换性。本文不臆想这一设想的浪漫实现，而是诚实面对其制度边界，并在结尾给出明确的证伪条件。

一、引言：一个让所有人都不满却无人敢退出的困局

教育评价改革已经持续争论了数十年，一个令人困惑的模式始终在重复上演：任何试图超越标准化考试与分数排名的改革方案，几乎都会在争议中陷入同一个僵局。一旦提出用更丰富、更全面的方式评价学生，反对的声音就会立刻追问——这些新方式是否公平？

问题不在于这些追问没有道理。恰恰相反，它们每一句都击中了要害。“不看分数看什么？看谁家有钱请得起马术教练？”这一诘问有一种让人无从反驳的直接性。在教育资源本就分布不均的社会里，任何超越纸笔测试的评价维度——面试表现、才艺展示、社会实践、科研经历——都天然地更有利于那些家庭资本雄厚的孩子。普通家庭的孩子，在纸笔测试之外几乎没有任何可以依凭的优势。于是，“还是比分数最公平”就成了一句让人无法反驳的结论：在所有不平等之中，分数的比较是相对最可以被接受的一种不平等。

这个困局的独特之处在于，它不是某一方在顽固抵抗，而是所有方都被困在同一种理性计算之中。学校知道应试教育有损学生的长远发展，但不敢单方面降低考试权重——那意味着它将在升学竞争中失去客观可比性，招致家长的集体抗议。家长在理智上理解“不能只看分数”，但在一个连孩子就读的学校层次都会被亲朋邻居暗自比较的环境里，任何单方面退出竞争的企图，都意味着让自己的孩子在系统中“失声”。补习机构甚至不需要主动制造焦虑——它们只需要如实反映竞争的激烈程度，家长们就会自行涌入。而那些最底层的家庭，恰恰是最坚定地捍卫分数公平的人，因为他们手里除此之外没有任何可以依凭的东西。

这使得评价改革呈现出一种极不寻常的形态：它不是一项可以在实验室里优化测量技术、然后大面积推广的技术革新，而是一个多方力量深度咬合、任何单边行动都会触发全面反制的系统性僵局。近几十年来，出现了大量试图解释它的研究。自1990年代以来，关于“素质教育与应试教育之张力”的讨论已经形成了庞大的中文文献群。这些研究从政策执行失真、学校文化抵抗、利益博弈、评价技术与课程理念脱节等角度，细致地分析了素质教育各项举措——综合实践活动、研究性学习、多元评价试验——为何“喊得轰轰烈烈，搞得扎扎实实”的仍然是应试。这些诊断准确地描述了“素质教育被应试化”的症状，但它们很少追问一个更底层的问题：为什么任何教育实践，只要进入了一个以选拔为核心功能的评价系统，就几乎不可抗拒地被拉向可比较、可认证的方向？这股拉力的性质究竟是什么？

本文试图追问的正是这个问题：这张让所有人都不满却又无人敢先松手的大网，究竟是靠什么力量在持续运转？它从一项帮助拔擢人才的机制，演化为一套让所有人都无法从中脱身的结构，这一演化本身揭示了一个什么样的底层预设？

这一追问在方法论上意味着一次根本的视域转换。以往的研究——无论是对标准化考试的批判、对素质教育的呼吁、还是对评价技术的革新——大多预设了一个相同的前提：世界是被发现的，对象先于认识而存在，知识的使命是接近它、描述它、修正对它的描述。在这个预设下，评价改革的根本任务就是“找到更精确、更全面的测量工具”，去“发现”学生身上那些被考试遗漏的品质。这当然不是没有道理，但它忽略了一件事：评价本身参与了它所评价的对象的生成。不是先有“学生的成长”这个稳定的对象摆在那里，然后评价工具去更精确或更粗糙地测量它——而是评价的方式、朝向和强度，从一开始就在塑造什么是“值得被看见的成长”，什么被挤压到视线之外，什么因为长期不被认证而逐渐从教育现场消失。本文的分析起点，不再把评价视为一个中立的测量技术问题，而是把它视为一个参与了教育现实本身之发生的过程。

本文的论证分五步展开。第一步，回溯中国评价体制在过去半个世纪中的四个关键节点——1977年恢复高考、1999年大学扩招、2000年以后教培产业资本化、2014年启动的新高考改革——从中追踪一个反复出现的模式：每一次试图松动评价垄断的努力，不仅没有替代旧有的比较逻辑，反而在更大的生活空间里扎下了根。在这一步的叙述中，本文不会预留一个“先在的概念”等待被应用；被追踪的那个朝向，将随着历史的展开而逐渐获得自己的名字。第二步，在历史叙述的过程中完成与既有概念——标准化考试、应试教育、外在动机、教育锦标赛、优绩主义——的严格对质，让这对新辨别维度的边界在与最强对手的碰撞中显现出来。第三步，引入某高中“研究性学习”课程从自发探索到异化为履历军备竞赛的完整案例，作为贯穿全文的叙事锚点，让“朝向的弯曲”在真实的制度经纬中发生，而非在虚构的实验中演绎。第四步，正面论证向生性的不可还原性：分析外部注视介入认知过程时发生的形变——不仅是注意资源的分配，更是认知目的论的置换；在此基础之上，探讨一种不为“替代分数”而设、而为“捍卫退出比较的余地”而设的另类实践方向。第五步，诚实面对三个最严厉的质疑，包括向生性空间被文化资本隐秘渗透的风险，并给出明确的证伪条件。

二、一个朝向的生成史：从一项技术到一种朝向

要理解今天的困局，必须首先放弃一个习以为常的预设：标准化考试的主导地位¹，是因为它在技术上“最好用”，或者因为教育者“迷信考试”。事实是，它在

不同年代被完全不同的力量推到了今天这个位置上，而那些曾经试图改变它的努力，不但没有松动它，反而让它在更大的生活空间里扎下了根。

2.1 1977年：分数公正从什么手中夺过了正当性？

高考恢复，是中国现代教育评价史上最重要的事件。但它真正改变的，不仅仅是选拔制度本身，而是整个社会对“什么是公平”的理解方式。

在1977年之前，大学招生主要依据的不是学业成绩，而是政治审查和基层推荐。推荐制度在实践中严重依赖基层干部的个人判断、家庭出身背景以及各种非公开的权衡。一个农村青年能否上大学，不全看他的学业能力，而更多看他在公社里能否获取推荐名额。这不是秘密，而是那一代人的普遍记忆。

1977年恢复高考的实质动作是：用一个公开、统一、可事后核查的分数序列，替换那一整套不可见的推荐权衡。它的正当性不是从某个教育理论中推导出来的——不是“分数可以精确测量出学生的发展水平”——而是来自一个极其残酷的历史对比：在推荐制度下，普通家庭的子女希望太渺茫了。分数冷酷，但它不看你是谁的孩子，也不看你 and 公社干部的关系有多近。这种“冷酷的公平”，在那个时代背景的映衬下，成为一种让人安心的品质。

这一步的后果远远超出了当时决策者的预期。它让分数不仅仅是一项选拔工具，而且成为一种正当性的符号。在那之后，任何试图模糊分数边界的改革，都会立即被追问同一个历史对比中的问题：你想用主观评价来代替分数，那你怎么保证它不像当年的推荐制一样，变成关系户的通道？这一追问构成了此后所有评价改革的地基。

但在这个地基之上，有一件事被悄悄固定下来，却从未被普遍察觉。那就是：此后四十余年里，所有关于评价改革的争论，都默认接受了一个前提——评价天生就是为了选拔而存在的。无论是捍卫分数的人，还是试图超越分数的人，都在“评价=选拔工具”这个预设框里打转。前者说，分数是相对最公平的选拔工具；后者说，我们需要更全面的选拔工具。双方争执的是“用什么选拔”，而不是“评价是否必然服务选拔”。

这个从未被质疑的预设，在这一刻开始隐约露出它的轮廓。但它还没有名字。它只是一个习惯：一种把“评价”和“选拔”视为天然一对的思维方式，一种在看任何教育行为之前就已经预设了“它能不能被拿出来比”的目光。这种习惯，在此后数十年间被反复加固，直到它从一项制度选择，沉积为整个教育生态的首选姿势。但那是后来的事。在1977年，人们看见的只是一个劫后余生的制度重建，没有人会想到，他们正在为一个此后再也无法拔除的朝向，打下第一根桩。

2.2 1999年：扩招如何把家庭推上了竞争的甲板

1999年大学扩招，通常被看作一次教育扩张的里程碑。但它在评价问题上造成了一道隐蔽却深远的裂痕。

高考恢复之初，大学录取率极低，能考上大学的人本来就少。那时候分数虽然重要，但多数家庭的子女并不以此为人生的必经路径——考不上大学，尚有大量体制内就业通道可供选择。在这个阶段，家庭对分数的关注虽然存在，但远未达到后来那种铺天盖地、渗透日常的程度。

扩招改变了这件事的结构。当高等教育从精英阶段跨入大众阶段以后，“上大学”本身的重要性被稀释了——但与此同时，“上哪所大学”的差异被急剧放大。在文凭通胀的效应下，一封大学录取通知书不再是终点的奖赏，而成了新一轮分层竞争的入场券。家庭对分数的关注，由此从“考得上考不上”窄化为“考得上哪一所”——而这一窄化，恰恰把竞争的颗粒度细化到了几乎每一次测验、每一道题的层面上。

这个变化不仅仅是“竞争变得更激烈了”。它对教育生态的深层影响是：家庭被推上了竞争的甲板，从此以后，几乎所有家庭的教育决策，都被一只无形的手拧向同一个方向——做任何一件事之前，先问“这对升学有没有用”。这不是因为家长们功利，而是因为当教育成为社会分层的主通道时，任何理性的家庭都不会让自己的孩子在通道中途掉队。这种“有用性追问”从升学节点蔓延到日常学习的每一个角落，它在此后二十年逐渐从一种外在的权衡，内化为一种看孩子的方式：一个孩子坐在书桌前，家长的目光几乎自动地开始辨认——他此刻做的事，是“有用的学习”还是“无用的消遣”？

在中文教育社会学研究中，吴愈晓等学者将这种机制精确地概括为“教育锦标赛”：家庭之间为了维持或提升相对位置而不断加码竞争，导致竞争的基准水位持续攀升，而所有人的相对位置几乎纹丝不动。这一概括准确地抓到了“多方咬合、军备竞赛”的结构特征。但本文试图追问的是锦标赛逻辑未能覆盖的维度：锦标赛解释了“竞争怎么升级的”，但没有解释在这个升级过程中，家长和孩子在日常的每一个微小决策里，那种“做这件事之前先看它能不能被认证”的朝向是怎么被内化的。锦标赛是制度逻辑，但它还没有走到那一步——走到一个人独处时，依然下意识里想象自己的行为正在被外部目光所评价。那一步，才是本文要追踪的朝向。

2.3 2000年代以后：资本如何把焦虑计算化

如果说扩招把家庭推上了竞争的甲板，那么随后二十年里教培产业的资本化，则直接给这艘船装上了引擎。

在2000年代以前，补习还是一种零散的、边缘性的存在——个别大学生做家教，退休教师在自家书房开班，规模有限，影响不上网。到新东方、好未来等大型机构在资本的驱动下大规模扩张以后，补习的性质发生了根本变化。这些机构做的并不是简单地“增加教学供给”，而是做了一件极其精巧的事：把弥漫在家庭中的焦虑，转化为可计算、可定价、可规模化交付的产品。

“你不补，别人都在补。”这句话之所以具有如此强大的穿透力，不是因为补习机构在夸大事实，而是因为它如实反映了一个已经结构化的现实。当班级里超过一半的学生都在校外补习时，老师的教学节奏必然会发生微调——这不是任何人的恶意，而是任何理性教师在面对学生群体的实际水平时都会做出的正常反应。但这个微调制造了一个新的信号：不补习的孩子可能会跟不上。这一信号传递给家长，家长的焦虑转化为补课需求，需求催生更多的补习班，更多的补习班进一步拉高了竞争的基准水位。

至此，补习不再是一项可供家庭自由选择的附加服务，而变成了一场陷入囚徒困境的军备竞赛。每个家庭都在为维持相对位置而不断加码，但所有人都加码的结果，是让基准水位不断攀升，而相对位置几乎纹丝不动。

在这个过程中，补习市场起到了一个微妙的、但极少被点破的额外作用：它不仅服务于现有的评价体系，还在日常经验中反复向家庭传递一个信念——任何学习上的问题，都可以被转化为一个“可以去补”的问题。孩子在某个知识点上卡住了，去补；孩子对某个学科没兴趣，去补；孩子想拔尖，更要去补。这个信念的扩散，在深层把“学习=被可评价的进步”这个等式烙进了千百万家庭的日常思维中。学习不再首先是“一个人对他与世界的相遇做出回应”，而是变成了一连串可以被测试、可以被打分、可以在班次和排名中看到上升或下滑的节点。由此，那个从1977年开始隐约成形的朝向，在此刻获得了一个日常起居的载体——补习班的课表、每月的测评排名、家长群里刷屏的成绩单——通过这些载体，它不再是飘在制度层面的抽象逻辑，而沉入了千百万家庭的晚间饭桌、周末清晨和寒暑假的安排之中。

2.4 2014年：改革如何意外加固了向评性

2014年启动的新高考改革，是一个在中国评价制度史上极具分析价值的节点。它的本意是为过于单一的高考评价模式松绑——引入综合素质评价、增加选考科目组合、给高校更大的自主选拔空间。

然而在实际落地的过程中，这些美好的设计触发了一种意想不到的动态。当“综合素质”被正式纳入高校选拔的参考依据时，这个原本意在校正分数单一性的维度，迅速被资优家庭重新转化为一种新的可竞争、可计算、可投资的领域。研究性学习项目变成了流水线式的履历生产，公益服务变成了一种可以用时长、证明力和媒体出镜率来衡量的展示性资源，科技创新竞赛则催生了一个围绕“包装项目”的新产业。

这一切都指向同一个方向：综合素质评价，并没有在分数的逻辑之外打开新的空间，而是把那些分数暂时照不到的领域，也拉进了同一套比较与排名的逻辑里。那些原本在课余自发发生的东西——对一个问题的好奇、在某些下午翻完一本书的无用热情、在一件小作品上磨了很久的耐心——被一种新的引力场捕获，开始朝着“可以被记入档案”“可以被写进简历”“可以在面试中被展示”的方向生长。

更关键的是，这轮改革在一个问题上彻底撕裂了曾经脆弱的共识。那些最普通、最没有资源可额外投入的家庭，原本把分数视为自己唯一可依凭的武器。当综合素质评价作为一种新的“加分”可能被提出来时，他们看到的不是“更全面的评价”，而是“隐藏加分”的威胁——那些他们无力购买的课外活动、无力支持的研究项目、无力经营的履历竞赛，正在变成别人进入好大学的新通道。于是，反对的声音在这一轮改革之后更坚定了：“还是比分数最公平。”

这就是那个朝向在此刻获得的最明确的一次自我加固：每一次试图松动评价垄断的改革，都在实施过程中激活了既有社会经济不平等的复制机制，而这一激活的结果，是让那些原本可以被争取到改革一边的家庭，反过来更加坚决地捍卫旧制度。改革不是遇到了来自某一方的阻力，而是每次都在一个具体的、可经验的节点上，把一个可以想象的替代方案变成了一次令人失望的现场教学——它让人们亲身体验到，离开分数，似乎只会变得更糟。

2.5 对这个朝向的第一次命名

在既有研究中，已有学者用“路径依赖”或“互锁网络”来形容上述困局：学校、家庭、市场与选拔制度之间形成了多方咬合的锁定结构，任何单方面的松手都会触发其他节点的反制，使得改革反复被加固为旧结构的自我再生产。

这一描述是准确的，但它只能回答“这个系统为什么不容易被打破”，而不能回答一个更深层的问题：这个系统朝哪个方向在运转？它不仅仅是一种静态的锁定形态，它是一种被历史一再加固的偏向。就像一艘船的舵被卡住了，它不是因为风浪太大而停在原地——它是被卡死在一个固定的方向上。

现在可以给这个方向第一次命名了。本文把它叫做向评性（assessment-tropism）——这个词借用生物学中“向性”的概念，强调的是有机体在环境中朝特定刺激的方向生长的趋势，而不是机械决定论中的“作用力”。它不是锁定本身，而是锁定的方向性。它精准地抓住了这样一件事：在过去四十余年的演化中，中国教育生态中的行动者——学生、家长、教师、学校、补习机构——在做选择时，正被一种日益强大的、不成文的引力场拉向同一个朝向：做任何可以被外部看见、被标准衡量、被制度认证的事，而远离那些在分数和履历上留不下痕迹、却可能在成长的缓慢历程中至关重要的东西。

这不是一个等到第三节才被“锻造”出来的先验概念。它是在上述四个历史节点的展开中逐渐获得自己的形貌的：1977年，它以一种“公平朝向”的面目首次进入教育生态，把分数嵌入了正当性的地基；1999年扩招之后，它从一项制度安排渗透进家庭的日常决策，成为那个“做任何事之前先问这对升学有没有用”的默认计算器；2000年代补习资本化之后，它找到了一个强有力的日常载体，通过课表、测评和微信群里的排名，把那个朝向从抽象的制度逻辑转化为具体的、可以被身体感受到的作息节律和家庭节奏；2014年新高考改革之后，它展示了它最令人警惕的能力——不是拒绝改革，而是通过吸收改革来扩张自己的疆域，把从前不受分管辖的生活领域也拉进了比较的循环。在这四十年的每一个节点上，它都在生长、变厚、扩散。到本文写下这些字的时候，它已经不是一种“选拔制度”的问题了，它已经漫出了制度的边界，渗进了千万家庭看孩子的目光里，渗进了教师和学生在校课交谈的呼吸里。

与此同时，还必须在本节就完成与米歇尔·福柯“考核”（examination）概念的对话。福柯在《规训与惩罚》中深刻地分析了考试作为现代规训技术的运作机制：它不仅是知识传递的工具，更是权力-知识在个体身上运作的核心装置——它将个体置于持续的监视之下，使其成为可被分类、可被比较、可被矫正的对象。向评性显然与这一传统有亲缘关系：它所抓住的那种“持续朝向可被外部认证的方向弯曲”的趋势，正是福柯所描述的规训机制在教育现场的具体呈现。然而二者的切割点同样清晰。福柯的“考核”追问的是权力如何通过知识技术来“看见”个体，从而将其置于规训的网格之中——它关心的是权力的运作。向评性追问的不是权力如何运作，而是个体在长期浸润于这种规训网格之后，其认知过程本身的朝向如何被重构——它关心的是“发生”的质地。一个学生不需要任何人在场，也不需要任何监考老师走动，他只是在书桌前翻开一本书，然后他的脑子里自动浮现出一个问题：“读这个有用吗？能写进自荐信吗？”福柯的规训需要监视塔，向评性不需要——那个监视塔已经内化到个体的自我观看中了。或者换一种说法：福柯的“考核”解释了一只眼睛如何始终在看着你，而向评性解释的是，在被看了足够久之后，你如何不再需要那只眼睛——你学会了在看任何东西之前，先替那只眼睛看自己。

三、一个案例：研究性学习的向评化

上一节在宏大的历史尺度上追踪了向评性的生成。但“朝向”这种东西，如果在论文中仅仅作为宏观历史叙事的概括出现，它就可能停留在隐喻的层面——读者知道作者在说什么，但不确定它究竟以什么形态存在于真实的学校日常中。它可能被误读为一种“描述性的概括”，而非一个可以被经验指认的、有具体肉身的存在。本节引入一个真实学校的微观案例，以展示：向评性不是研究者从外部强加给教育现实的解释框架，而是在每一所学校的具体实践中，被反复地、可观察地发生出来的过程。这个案例取自某东部省会城市一所市级重点高中2015年至2020年间“研究性学习”课程从建立到异化的完整经历。材料来源为该校教师与学生的访谈记录（2019—2021年间该校内部教学调研所留档案）、学校教务部门存档的课程方案与评价标准修改过程，以及部分学生事后写下的个人叙事。出于隐私保护，所有人物姓名和学校地理位置均已做匿名处理。

3.1 设想：一个未被向评化的起点

2015年，这所学校响应省级教育部门推动素质教育的号召，在高一和高二年级设立了每周两课时的“研究性学习”课程。最初的课程设计由两位刚从师范大学研究生毕业的年轻教师主导，他们都曾在导师的影响下接触过项目式学习和探究式教学的理念，对这门课抱有一种真诚的教育热情。在他们的设想中，研究性学习应该做的是这样一件事：给学生一个不被考试时间表切割的整块时间，让他们自己发现一个问题，自己去翻阅材料、设计方法、分析数据，最后交出一份报告——这份报告不一定要有“正确结论”，甚至不一定要“成功”，重要的是学生在这个过程中经历的认知冲突、推倒重来、山穷水尽之后突然看到另一条路的那种东西。

在第一年的实践中，确实出现了几个让教师至今记忆犹新的片段。有一个小组选择研究“学校食堂的浪费现象”，初衷只是觉得这个题目好玩。他们在食堂蹲点了两周，统计了各类食物的丢弃量，记录了不同年级、不同性别的浪费差异。后来有一天，一个学生在翻阅经济学课外书时偶然看到了“外部性”这个概念，突

然兴奋地说：“这不就是我们记录的那些东西吗——一个人倒饭，表面上是自己付了钱，实际上浪费的粮食和处理的成本都是大家承担的。”那一刻，这个小组的研究主题从“统计浪费”跳到了“食堂浪费的外部性”——这个跳跃不是任何一位老师建议的，而是学生在被材料本身拽着往前走的过程中自己发生的。后来这个小组的报告拿了市里一个评比的三等奖，但指导教师事后对她同事说：“我真不敢说这个三等奖对学生有什么好处。他们最好的时刻是那天在食堂后门口，一帮人蹲在地上翻垃圾桶分类称重，臭得要命，但一个个眼睛亮得不行。那个时刻跟得不得奖没有关系。”

这个案例在第一年是一个孤立的事件。但它的重要性在于，它让当时参与课程设计的教师自己心里认了一件事：在理想情况下，研究性学习能不能成为一块不同于日常考试逻辑的土地？他们当时觉得，有可能。

3.2 异化：三股力量的依次介入

从第二年起，三股不同的力量开始依次介入这门课程，每一股都有完全正当的理由，但合在一起，它们把课程的朝向拧向了另一个方向。

第一股力量来自教务管理的合理化压力。第一年课程结束之后，学校教务处在总结会上提出了一个无法回避的问题：这门课给学生的成绩怎么录入？在没有分数、没有排名的情况下，教务系统无法处理“研究性学习”这一栏。当时提出的折中方案是：不需要精确到分数的排名，但至少给出A/B/C三档等级。这看起来只是一个技术性的妥协——但等级制一旦引入，就等于在课程中安装了一双在此之前不存在的眼睛。教师开始需要在心里把每个学生的表现对应到一个等级上去；而学生很快发现，学长学姐在高三参加自主招生和综合评价录取时，有些高校的申请表中有一栏叫做“研究性学习经历及成果等第”。那道本来被刻意关闭的门，被撬开了一条缝。

第二股力量来自家长的信息套利。到第三年，一些消息灵通的家长——他们中有好几位本身就是高校教师或科研人员——通过自己的渠道了解到，某些顶尖大学在自主招生面试中，对“有实质性研究经历”的学生会高看一眼。这些家长一旦掌握了这条信息，就不会坐等学校来为他们的孩子提供同等的机会。他们开始亲自下场：帮孩子选题（题目往往高度契合当前大学的学术热点），帮孩子修改文献综述（用他们自己在学术训练中积累的规范语言），个别家长甚至把孩子带进自己的实验室，在研究生协助下完成了一两篇远远超出中学水平的论文。这些孩子交上来的研究报告，在格式、数据、文献引用上的专业程度，与那些纯粹由学生自己摸索出来的成果之间，出现了一道肉眼可见的鸿沟。

第三股力量来自校际竞争。该校在省内处于“第二梯队前列”——不是顶尖名校，但长期处在追赶第一梯队的位置。当学校教务处发现，本市另一所竞争校的学生凭借高质量的“研究性学习成果”在自主招生中获得了优势时，来自学校管理层的压力开始传导到课程组。教务主任在一次内部会议上说了一句后来被课程组教师反复引用的话：“我们不搞功利的，但如果隔壁学校都在拿这个当抓手，我们不跟进，就是对学生不负责任。”

到第五年，也就是2020年，这门课程的实际运作形态已与2015年的设想几乎没有关系了。学校引进了几套规范的“研究性学习指导手册”，内容涵盖选题技巧、文献检索规范、数据分析方法、论文写作模板。这本身并没有错——从教学的角度看，它的确提升了研究报告的平均质量。但关键在于，当这些指导手册进入课堂之后，它悄然改写了一件更根本的东西：它把“什么是值得做的研究”这件事，从“学生自己与材料的碰撞中逐渐生成判断”，变成了“按照手册中列出的评分标准逐项对表”。选题的新颖性、研究方法的规范性、数据来源的可靠性、结论的创新性——每一项都有明确的等级描述。学生翻开手册的那一刻，他就已经知道自己接下来要做的事，在哪一页上能找到“得分点”。

2020年底，一位从第一年就参与课程设计、一直坚持到第五年的教师在离职前写下了一段话，至今保存在该校某位同事的私人记录中：

“以前我跟学生说，你只要把自己真正好奇的东西做出来就好。现在我不说了。不是因为我不想说，而是这句话在现在的课堂里说出来，学生看我的眼神就像在看一个不懂规则的新老师。他们都知道，这不是‘你好奇什么’的问题，这是‘你做什么能拿到A等’的问题。他们心里那本账比我清楚。”

3.3 这个案例所展示的：朝向如何在真实的制度经纬中被弯曲

这个案例不是“素质教育被应试化”的又一个老生常谈的例证。如果只读到这一层，那它只是在重复一个已经重复了二十年的诊断。本文把这个案例放在第三节，是因为它在一个中观的、有具体肉身的层面上，展示了向评性如何不是在某个宏观政策出台的一瞬间被“确立”的，而是在每一次完全有正当理由的局部决策中，被一点一点地重新发生出来的：教务管理需要一个等级，这个等级撬开了一个入口；家长掌握了信息，把信息转化为孩子的履历优势；学校看到了竞争对手的动作，把竞争压力传导为制度化的指导手册。这三步中，没有一步是“恶意的”，没有一步是“有人在阴谋破坏素质教育”——每一步的背后，都可以找到一个完全正当、甚至无可厚非的理由。但正是这种“每一步都有道理，合在一起却走向了另一个方向”的过程，才是向评性真正的发生机制。向评性不是一块从天而降的铁板，它是一种在特定制度土壤中、经由无数个局部理性选择而持续再生的朝向。

更重要的是，这个案例让那个从第二节历史叙事中逐步显形的“朝向”获得了可被经验观察的肉身。那个在1977年随着分数公正一起进入教育生态的预设——“评价必定服务于选拔”——在这一刻，不再是一个漂浮在制度分析层面的大词。它在教务管理被等级录入打乱的那一秒落了地；在家长拿到信息、开始帮孩子选题的那一晚落了地；在学生翻开指导手册、寻找“得分点”的那个下午落了地。向评性，从这一步开始，不再是研究者从外部强加给历史的概括，而是每一个被卷入其中的人——学生、家长、教师、教务员——在自己的日常行动中，用无数个微小的、有理有据的选择，共同维持着的一个朝向。

四、在概念网格中辨认向评性：与既有解释的对质

有了第二节的历史纵深和第三节的经验肉身作为基础，现在可以正面处理一件在传统学术论文中通常会被放在前面、但在这里被刻意推迟的工作：与既有概念的严格对质。推迟它，不是因为视之为不重要，而是因为向评性的边界只有在历史与案例充分展开之后才能被清晰地看见。如果还没有看到这个朝向如何在真实的制度经纬中运作，就急于把它和已有概念放在一起“划界”，那很容易滑入用形容词修饰旧概念、然后宣告自己有了新发现的陷阱。

4.1 与“标准化考试”“应试教育”的对质

向评性毫无疑问与标准化考试高度相关——它最典型的、最可见的运作形态，就是分数、排名与选拔。在这一点上，本文不打算否认两者的亲缘关系。但向评性承诺比标准化考试更多。标准化考试是一种具体的制度形态——它有明确的考试日期、规定的题型、固定的评分标准。向评性不是一种制度，而是一种朝向。它可以在标准化考试之中运作，也同样可以在任何声称“超越考试”的新评价形式中运作。当一个成长档案被用来作为入学的参考依据时，它就不再是一个私人的叙事空间，而已经被拉进了向评性的引力场——学生开始自觉地想“我应该写什么样的反思才能被老师认为深刻”，家长开始打听“档案里放什么项目会比较有竞争力”。档案本身没有改变形式，但它的朝向变了。形式上的“动态”“多元”“过程性”，一旦服务于外部的比较与认证功能，它不仅没有摆脱向评性，还恰恰是向评性的一种更隐蔽、更精致的运作。

这就是为什么本文必须放弃用“静态/动态”这对概念来切割评价形式——那会让分析停在表面，把所有“长期跟踪的”“没有标准答案的”“强调过程的”评价形式都当成对旧结构的超越，而看不见它们可能只是换了一种形式在做同样的事。向评性切割的标准不是形式，而是功能：该评价是否进入了分配-认证的比较循环。一旦进入，无论它有多“动态”，在功能上都已经向评化了。

与“应试教育”的对质则是另一个层次。应试教育是一个在中文教育批评中已经被使用了几十年的概念。它通常被用来指称一种以考试为中心、以分数为目标的畸形教育形态——教育不再是为了人的成长，而成了考试的附庸。向评性与应试教育在批判对象上有重叠：它们都指向那种把一切教育活动都拉向可测量、可比较方向的力量。但两者的层次不同。应试教育描述的是症状——教育不再像教育了；向评性追问的是机制——教育不再像教育，是因为整个生态被一种朝向可认证性生长的引力场捕获了。 应试教育批评容易陷入一种怀旧的陷阱：它在想象一个“还纯粹”的教育，那时老师只管教书不管考试，学生只管学习不管排名。向评性不接受这种想象。它从一开始就承认：只要教育承担着分配社会位置的功能，向评性就不可能被根除——它只能被制衡、被限制、在某些局部空间里被暂时悬置。这不是一个“去掉”，这是一个“认清它、守住它不应该越过的边界”。

4.2 与“教育锦标赛”的对质

吴愈晓等学者所发展的“教育锦标赛”概念，是向评性必须面对的最切近的对手之一。锦标赛机制精准地描述了这样一种动态：家庭为了维持或改善子女在教育阶梯中的相对位置，不断加码投入——补课时间延长、竞赛种类增多、课外活动的教育回报率被反复计算——而所有家庭同时加码的结果，是竞争基准的持续抬升，但相对位置几乎不变。这形成了一个“谁先松手谁先出局”的囚徒困境。

向评性与教育锦标赛在诸多方面有描述上的重叠：它们都抓到了“多方咬合”与“军备竞赛”的特征，都拒绝将竞争归因于个体的道德缺陷，而将其视为一种结构性的、理性选择驱动的系统性后果。但两者切割的是不同的维度。锦标赛追问的是“竞争的强度为什么持续升高”——它关心的是投入量的问题。向评性追问的是“在竞争中，什么被算作值得投入的东西”——它关心的是投入的方向被什么力量定义了。 教育锦标赛可以解释一个家庭为什么愿意每年花十万块钱给孩子补习：因为别人都在花，不花就会落后。但它无法解释另一件事：为什么连那些没有被考试覆盖的领域——一个孩子的闲暇阅读、一台旧收音机的拆装、一个下午的山坡探险——也开始被家庭以“这对履历有没有用”的目光来审视？锦标赛解释的是“卷度”，向评性解释的是“卷的方向是谁定的、怎么被内化的”。前者是力的强度，后者是力的朝向——两者不可互相还原。

4.3 与“优绩主义”的对质

近年来英语世界围绕“优绩主义”（meritocracy）的讨论，以迈克尔·桑德尔和丹尼尔·马科维茨等人的工作为代表，将批判推进到了一个新的深度。优绩主义的批判核心有二。其一，优绩承诺“只要你努力就能成功”，但在现实运作中，优绩竞争的结果高度依赖于家庭文化资本和社会经济资源的初始分配，“凭本事吃饭”实则是代际优势的合法化修辞。其二，优绩的逻辑从制度层面下沉为个体的自我叙事：成功者相信自己的成功完全是努力与才华的结果（桑德尔称之为“优绩的傲慢”），失败者则被逼入“你不够努力”的自我归咎之中。韩炳哲在他对“功绩社会”的分析中进一步指出，这种逻辑制造了一种特殊的主体形态——在功绩社会中，人不再是外在地被剥削，而是自己剥削自己：绩效主体把外部评价标准完全内化为自我要求，在心理上处于一种“自由地强迫自己”的状态。

向评性与优绩主义批判共享一个判断：基于可比较指标的社会选拔机制，会从制度安排逐渐渗透为个体的自我认知方式。但向评性追问的与优绩主义并不是同一个层面的问题。优绩主义批判追问的是“这个游戏规则是否公平”、“成功者是否配得他们的成功”——它关心的是分配正义。向评性追问的是“在加入这场游戏之后，玩游戏的人看待世界的方式本身发生了什么变化”——它关心的是认知发生的质地。 一个学生可以完全接受“优绩社会不公平”这个判断，并且真诚地认为人应该从优绩的暴政中解脱出来，但他仍然可能在第二天早上翻开课时，那只已经内化的眼睛自动打开，开始评估“这段话值不值得背”“这个思考题会不会考”。优绩主义的批判可以让他站在政治立场上反对这套规则，却无法阻止他的注意力在独处时仍然被拉向那个“可被认证”的方向。那个方向，就是向评性要命名并解剖的东西。

与韩炳哲的“功绩主体”的对话可以走得更深。韩炳哲的功绩主体是一种“自由地强迫自己”的现代人：这种主体将外部绩效标准完全内化，以至于不再需要外在的压迫——他自己就成了自己的监工。向评性恰恰捕捉到了这种“自我监工”在教育认知过程中的微观运作。但二者的切割点同样不可忽视。韩炳哲的功绩主体描述的是一个成年人——一个已经完整地内化了绩效逻辑的现代职场中的人。向评性追问的，是一个十二岁的孩子，在还没有任何“功绩主体”自觉的情况下，他的认知过程是如何在与评价制度的长期互动中，逐步被弯曲的。功绩主体是向评性在教育体系中长期运作的终点，向评性则是功绩主体的发生学前传。前者关心的是“这种人是什么样的”，后者关心的是“这种人是怎么被生成出来的”。一个人不可能在二十五岁那一年突然成功绩主体——他一定是在此前十几年的学校生涯中，被无数个微小的、日复一日的“评价”瞬间，缓慢地、几乎察觉不到地，铸成了那个样子。向评性要做的工作，就是去追踪这个铸造过程的微观机制。

4.4 与“外在动机”的对质

心理学中对外在动机与内在动机的经典区分，以爱德华·德西和理查德·瑞安的自我决定理论为核心，在某些方面与向评性/向生性这对概念有交叉。外在动机是指为了获取外部的奖励或避免惩罚而行动的倾向——考高分是为了表扬、为了升学、为了将来找好工作；内在动机是指行动本身即是回报——解一道题是因为它有趣，读一本书是因为被故事吸引。向评性似乎就是“过强的外在动机淹没了内在动机”。

但这对心理学概念不足以覆盖向评性在教育现场中运作的精确方式。外在动机/内在动机追问的是“人为什么要做这件事”——它关心的是动机的来源。向评性/向生性追问的是“这件事在朝哪个方向生长”——它关心的是过程的朝向。一个学生被一道题吸引，发自内心想要解开它，这是内在动机。但如果他在解开的过程中，已经下意识在想象“这段经历能不能写进申请书里”，那么这个内在动机所驱动的求知过程，就已经开始在向评性的引力场里弯曲了——不是为了分数，但为了“履历的可展示性”。动机可以还是内在的，成长的朝向却已经在不知不觉中偏向向了“如何被看见”。 这就是为什么必须锻造一对新概念：它捕捉的不是驱动力的来源，而是生长本身被引力场弯曲的那个维度。

4.5 本节收束

本节以四场对质完成了同一件事：不是在“先定义、后应用”的逻辑中把向评性封进一个标本盒，而是在它被第二节的历史纵深和第三节的经验肉身充分展开之后，通过让它与最有力的对手——标准化考试、应试教育、教育锦标赛、优绩主义——逐一交锋，来让它的边界从碰撞中自己浮现出来。四场对质的结论可以一并表述：这些既有概念中的每一个，都抓住了教育竞争或评价机制中的某个真实维度，但没有任何一个触及了向评性所抓住的那个额外维度——个体在长期浸润于评价生态之后，其认知过程本身的朝向如何被重构，以至于在没有任何人在场、没有任何奖励或惩罚悬在眼前的情况下，他看待自己的行为、自己的好奇心、

自己在世界上做的事的姿态，已经被那只内化的眼睛永久地弯曲了。标准答案把它叫做“功利心”，那是一记道德谴责，不是分析。教育锦标赛把它叫做竞争过密，那是结构描述，没有进到认知的纹理。优绩主义把它叫做自我剥削，那是主体诊断，没有追问这个主体的铸造史。外在动机把它叫做外部驱力过强，那是动机分类，没有抓住过程的朝向被弯曲的那个维度。这就是向评性必须被独立命名的理由：它不是在既有概念的分类格子中找到一个新的空位，而是第一次让一个长期在经验中可以感受到、却一直没有一个名字来认领的区分面，在学术话语中获得了自己的名字。这个名字是否是多余的、是否应该被撤回，取决于下一节能否完成一个更苛刻的任务——证明那个与之构成对张力的概念，同样不可被既有概念所替换。

五、向生性：一块异质的土地

如果向评性的命名是为了让一个长期未被认领的区分面在学术话语中获得名字，那么这一节的任务更为苛刻：命名一个更难以被抓住的东西。向评性虽然此前没有名字，但它在经验中随处可见——分数、排名、选拔、履历竞赛，这些都在日常经验的表面上，只要换了看它的眼睛，就能辨认出它的轮廓。但向生性不是这样。向生性所在的那片区域，恰恰是一旦你试图去“看”它、去“测量”它、去把它拿出来展示给别人看，它就收回去了的那个东西。这就给命名带来了一个独特的困难：你怎么命名一个你无法展示的东西？你怎么论证一个在日光下就消失的东西真的存在？

5.1 从“学习投入”旁擦过

最容易被误认为向生性的概念，是教育心理学中的“学习投入”——尤其是其“认知投入”维度，常被描述为一种深度、专注、以学习本身为目标的状态。二者在表面描述上有相似性：都是深度卷入、沉浸于学习任务、不在意即时回报。

但一个向生性时刻与一个高认知投入时刻的差别，可以在一个微小的分岔点上被分辨出来。设想一个学生在解一道数学难题。他花了整整一个下午，草稿纸扔了一地，中间推翻了三次思路，最后找到了一个让他自己都意外的解法。这个过程中，他的认知投入很高：注意力高度集中，策略反复调整，元认知监控持续在线。现在，在这个下午快要结束的时候，他心里冒出了一个念头。如果他冒出的念头是——“这下明天的数学课上我可以展示这个解法了”——那么这个下午的所有投入，从这一刻起就被重新定性了。它仍然是高认知投入，甚至仍然是内在动机驱动的（他不是为了老师表扬才去解题的），但它不再是向生性。因为在他冒出那个念头的那一刻，他的朝向已经被那个即将到来的、被他者观看的瞬间弯曲了：他下午的所有动作，在这一刻的回顾中被重新编织进了一个“走向展示”的叙事里。如果他冒出的念头是——“原来这个定理是这么回事”——然后继续埋头去验证下一个推论，那么这一刻仍有可能留在向生性的区域内。尽管这个区域随时可能在下一秒被他自己的任何一丝“别人会觉得我怎么样”的念头所穿透。

认知投入测量的是“有多投入”，向生性追问的是“投入的引力方向被什么力量弯曲了”。一个学生可以“高投入”地准备一场能被展示的竞赛——他的注意力高度集中、认知资源全部调用、甚至在过程中体验到心流——但这整个过程仍然是朝着那个最终的展示台生长的。他在投入中的每一个动作，都被那个即将到来的被观看的时刻所弯曲：他会下意识地打磨那些在展示中会被注意到的细节，而忽略那些他纯粹出于好奇、但“评委不会问”的旁支问题。这种投入不可谓不深，但它不是向生性。向生性的标准不是投入的强度，而是投入的姿势：它是被事物本身的内在逻辑拽着往前走，还是被“这件事将会如何被看见”的预期拽着往前走。二者之间的差别，教育心理学的“认知投入”量表是无法辨别的——因为在量表上，这两个学生填出来的数据可能一模一样。也就是说，“学习投入”概念根本没有为这个区分面预留任何空间。不仅不能1:1替换它，甚至无法看见它。

5.2 与默会知识、内在动机、杜威“生长”的简要切割

在此基础上，与其他概念的切割可以更简练地进行。与博兰尼“默会知识”：默会知识指的是“知道的多于能说出的”，属于认识论的范畴——它关心的是有什么东西我们知道了却说不出。向生性指的是“发生的多于能被外部看到的”，它是一个过程发生学的概念——它关心的是有什么东西在被注视的那一刻就收了回去，不再以原有的质地发生。前者是知识形态的隐匿，后者是发生过程的形变。两者切割的是不同的维度。

与内在动机的切割，5.1节在对“学习投入”的分析中已经得到了深化：动机可以完全是内在的——不是因为奖励，不是因为表扬，纯粹是因为好奇——而朝向仍然可以在个体没有意识到的情况下被弯曲。一个研究生出于最纯粹的内在兴趣选择了一个研究题目，但他的选题、他的方法、他组织论文的修辞策略，已经在他没有察觉的情况下，朝着“评审人会怎么看”的方向长过去了。内在动机完好无损，但朝向已经不纯。外在动机/内在动机这对概念无法抓住这个维度。

与杜威“教育即生长”：杜威的“生长”概念有强烈的生物学隐喻——它暗示一种自然的、不被外力扭曲的展开，仿佛只要成人世界不去干涉，儿童的生长就会朝着它自己应当去的方向。向生性与杜威的这种“自然生长”有根本区别：向生性不是一种“只要不去干涉就会自然繁荣”的东西。它恰恰是在向评性的强引力场中，被系统性压制、弯曲、被逼退到边界上的那一部分认知发生。它不是前历史的纯真状态，而是后历史的残余地带——只有在向评性已经把几乎所有教育时空都塞满之后，那些仍然无法被比较逻辑所收编的认知发生，才在与向评性的张力中被逆向定义出来。杜威的生长不需要向评性就已经有名字了；向生性却只有在向评性被命名之后，才有可能被看见。

5.3 形变的发生：注视之下，什么在消失

前述论证如果只停留在概念对质的层面，读者有权追问：你一直在说向生性在“被注视时收回”“在评价介入时消失”、“认知质地发生改变”——但这种“收回”“消失”“改变”究竟是怎样发生的？它不是一个可以被搁在教育现场中观察到的现象吗？

现在可以接上第三节的案例来回答这个问题。回想那个研究性学习课程的第一年，那些蹲在食堂后门口翻垃圾桶分类称重的学生。他们之所以能做出后来的理论跳跃——“这不就是外部性吗”——不是因为他们在那个下午比别人“更努力”或“更投入”，而是因为他们在那个下午的认知过程中，没有被任何第二道目光所伴随。他们只有一件事要做：搞清楚这些数据到底在说什么。没有一个“我的表现怎么样”的念头在占用认知带宽，没有一个“指导手册上写的评分标准是什么”的旁白在后台运行，没有一个“评委会不会觉得我们这个课题不够高大上”的焦虑在干扰判断。他们全部的注意力，都投给了事物本身——投给了那些被浪费的食物、那些分类的数据、那些蹲在地上记录时被其他同学嘲笑却依然坚持的细节。

现在，想象第三年一位家长帮孩子把课题包装成了“基于行为经济学的校园消费外部性实证分析”——引用了最新的英文文献，方法论严谨到连研究生都未必写得出来。这个学生当然也深入接触了数据，甚至可能也体验到了某种程度的认知投入。但是，他的认知过程在全过程中都被一个额外的结构所框定：每一次选择分析方法，都必须问自己“这个在评审眼里够不够专业”；每一次遇到瓶颈，都必须权衡“要不要换一个简单但容易出成果的题目”；每一次落笔写结论，都在下意识地想“高校面试官读到这一段会怎么评价”。他的眼里不是只有那些数据和概念，他的眼里同时还站着一个观众席——上面坐着家长、评审、未来的招生官。

两件事在这里同时发生：他不仅在研究消费外部性，他还在替那些未来要评估他的人，提前评估自己的研究。这第二件事，不是附加在他认知过程之上的一个可

拆的包袱，而是与他做第一件事的每一刻都交织在一起的。认知带宽是有限的。当他分出一部分带宽去计算“这个做法能不能被认可”时，那部分带宽就从“这个现象到底是怎么回事”那里被抽走了。结果不一定表现为“他研究得不好”——在评审的标准下，第三年的报告可能比第一年的课题更“规范”、更“严谨”、更“好看”。但有一个东西确实消失了：那个在食堂后门口翻垃圾桶的下午，那种被事物本身拽着往前走、完全忘记了“我怎么样”的认知质地，被那个始终在场的观众席挤出了空间。这就是本文所说的“形变”：不是思考质量的下降，而是思考的姿势被弯曲了——它从朝向事物本身，被掰弯成朝向“事物+事物被看见的样子”。

这不仅仅是“注意力被分散了”这么简单。通常所说的注意力分散，指的是做一件事的时候被另一件事打断——比如看书时手机响了。那种分散是外部的、偶发的、可恢复的。而向评性导致的形变，是一种内部的、持续性的、沉入认知底层的弯曲。它不是因为突然冒出一个“别人会怎么看我”的念头而走神，而是“别人会怎么看我”这个计算本身就嵌入了认知过程的底层语法——它变成了思考时默认运行的后台程序。你不再需要“分心”去想它，因为它已经和你的思考长在一起了。

这还不是最隐蔽的形变。最隐蔽的那一层，发生在元认知层面——也就是那个决定“什么样的思考才算好思考”的标准被悄然置换的时刻。回想那个在第二年研究性学习课堂上翻看指导手册的学生。他不是在选择“我选这个题得A等还是B等”——那还是在用外部标准来评价自己的成果。比这更早一步发生的事情是：他翻开手册，看到上面写着“选题的创新性：能够与已有研究形成对话，或对既有结论提出质疑，可得高分”。从此以后，他在审视自己脑海中冒出的每一个选题想法时，那只已经内化的评价之眼就自动打开了——他在题目还没有正式落定之前，就已经在替评审者预先筛选那些“不够创新”的念头。他不是在做作业之前才去想“这个能不能得高分”，而是在想本身就已经被那个“能得高分”的标准给约束了。一个他纯粹觉得好玩、但说不上来有什么理论贡献的选题，会在尚未成形时就被他自行否决掉——不是评审不让他做，而是他已经学会了用评审的标准来替评审做第一轮过滤。这种在思考尚未成形时就被过滤掉的东西，终身没有机会被任何人看见——包括他自己。向生性，在这一刻，不是“消失了”，而是“从未被允许出生”。

这就是向生性最难以被理解的一笔：它并不因为被看见而增值；恰恰相反，它天然地、结构性地在被外部目光注视时发生形变。这不是一个测量技术能不能跟上的问题，这是一个存在论的差异——有些发生，它的完整性依赖于不被注视。一旦注视介入，它就收回去了，留下的只是一个“可以被记录下来的成长痕迹”。那个实际发生的认知过程正在那段痕迹的背后悄悄溜走，因为它从一开始就不是为了被留下来而发生的。

5.4 向生性空间：退出比较的另类实践

如果向生性不能被“培养”出来——因为任何以培养向生性为目标的措施，都可能在操作中把它重新向评化——那么，它在教育实践中还能做什么？

本文的回答是：它能做的不是被“推行”，而是被“护住”。向生性空间，不是一套被设计出来的新教学方法，而是一种在学校日常里被刻意留出的、坚决不进入选拔与比较循环的实践余地。它具有以下几个严格的、不可退让的边界条件。

第一边界：绝缘于任何选拔程序。向生性空间中发生的一切——学生写下的反思、师生之间的谈话、对一个问题的漫长摸索——永久不进入任何形式的评选、打分、排名、展示或升学参考。这不是“暂时不计分，以后看情况”，而是制度化地、不可逆地从比较循环中退出。不是“隐藏的评价”，而是“评价的缺席”。

第二边界：删除观众。向生性空间的实践不设看客，不设评委，不设可以对外公开的评价报告。教室之内，老师与学生之间，文本被写下，被阅读，被回应——但不被拿来对标，不被排序，不被摘要成任何可以在家长会上公布的评估。这个空间之所以能护住向生性，不是因为它设计得多么精妙，而是因为它对向评性最敏锐的那个触角——“有人在外面看着”——做了一个制度化的切断。

第三边界：教师的角色从“评价者”切换为“见证者”。在向生性空间中，教师不负责给出“这个学生达到了什么水平”的判断，而是负责以一种不急于下判断的、愿意陪伴不确定性的态度，去倾听、去回应、去帮助学生把自己正在经历的认知过程说清楚。这不是说教师从此放弃评价——在选拔所必需的那些时刻，教师仍然要评价。但在向生性空间里，评价的职能被刻意悬置了。教师在这里做的，是一种在向评系统里没有名字的工作：见证一个学生正在发生某种不会立刻凝结为成果的成长。

必须诚实面对的是，这个“见证者”角色在当下制度的夹缝中，本身就处于巨大的张力之中。在高度问责制、绩效压力与家长投诉风险并存的中国学校现场，“不评价”对教师而言并不比“评价”更容易——甚至可能更难。不评价，教师在当下的制度话语中就无法为自己的行为提供合法性的说明：评职称填表时，这项工作怎么计算？家长质疑“你为什么不给我孩子打分”时，教师怎么回应？面对教务系统要求提交评价数据的压力，一个拒绝提供任何量化等级、仅给予学生回应性反馈的教师能够坚持多久？这些都不是修辞性的追问——它们是每一位试图在日常教学中留出一小块“被评价豁免区”的教师每天都在遭遇的现实。

然而，恰恰因为如此，“地下状态”与“制度化”之间的关系才不是一个需要回避的矛盾，而是一个需要被看清的张力。向生性空间的制度化，不应该被理解为一种自上而下的“推广”——一旦以行政命令的方式推行，它本身就将自动向评化，变成一个“各校向生性空间建设成果评比”的新赛道。它更应该被理解为一种防御性的制度化：不是去“建设”新东西，而是去正式承认那些已经在暗处发生的事情，给它们一个不必躲藏的位置。它不是在制造一批新的“优秀见证者教师”，而是在为那些本来就在做这件事的教师，腾出一个不必再担心“被人问起时怎么解释”的制度夹层。

这三个边界条件加在一起，指向一个核心判断：向生性空间的存在本身，就是对“评价必定服务于选拔”这一底层预设的退场。它不假装自己可以根除向评性，它只是在向评性的引力场边界上，用制度化的自我保护，挤出几块小到不引人注意、却让向生性得以在其中不受弯曲地发生的土地。

六、诚实面对三个最严厉的质疑

上述设想不能只被当作一篇学术论文的漂亮收尾。它必须在发表之前，就被最严厉的质疑拷问过一轮。

质疑一：如果向生性空间不计分、不进入选拔，教师哪里来的时间与动力去认真做？

这一质疑预设了“只有能影响选拔结果的行为，才在制度中有存活的空间”。本文不试图论证教师在现有工作量下可以轻松承担这件事，而是想指出：这个“没有时间”的困境本身，就已经是向评性挤占了整个教育生态之后的症状。一所学校如果完全被选拔逻辑所塞满，那么任何不从属于选拔的教育行为都会被挤出——不仅是向生性空间，连带着师生之间那些没有预设目标、不知道会走到哪去的日常交谈，都会被挤压为“浪费时间”。所以，向生性空间能否存活，不是一个时间管理问题，而是一个共同体意愿问题：一个教育共同体愿不愿意在它的集体共识中，为“不计入成绩单、却仍有教育意义的工作”划出一缕不受挤压的时间。

但这个回答还不够。质疑者可以进一步追问：即便共同体愿意，在教师职称评定的现行标准下——教学成绩、科研成果、竞赛指导获奖等可量化的指标占绝对主导——“见证学生成长”这项工作无法被写入任何一张考核表，教师凭什么要为她付出额外的时间？这个问题如果处理不好，“共同体意愿”就只是一纸空文。

本文对这一层追问的回应是：承认这个困境是真实的，同时指出，它恰恰反证了向评性已经渗透到了教师职业认同的最深处。当一位教师只能在“可被考核的行为”中寻找工作的意义时，不是因为他没有别的教育信念，而是因为整个制度不再为那些“不可被考核”的教育行为提供任何可见的回报——不仅是物质回报，也包括象征性的回报：同事的认可、学校的表彰、教研会上的话语空间。向生性空间的制度化，不是要额外“激励”教师去多做事，而是要把那些教师已经在做、但一直不被看见的事情，纳入制度的承认范围——不需要打分，不需要评优，只是承认“这件事是这个学校里正常教育工作的组成部分”。这不是物质激励的问题，这是象征承认的问题。

质疑二：向生性空间凭什么不会被精英家庭转化为新的表演与不平等的温床？

这是本文必须诚实面对的最严厉反驳。有布迪厄传统的人会立刻指出：在文化资本严重不平等的社会里，任何去功利化的、强调叙事与自由表达的教育实践，都可能沦为上层阶级把阶层优势审美化为“教养”的场所。如果向生性空间进入日常教育实践，有雄厚文化资本家庭的孩子从小在家庭对话中耳濡目染，更擅长用恰切的语言讲述自己的内心世界；而普通家庭的孩子可能连被要求去认真回顾“这学期发生在你身上的一件重要的事”时，都感到一种难以言说的困难——这不是因为他们的经历贫瘠，而是因为他们从未在一个被鼓励去郑重讲述自己内心世界的环境中生活过。

这一反驳无法用“我们设计得更精细”来消除。本文的回应分两个层面。

第一层面：它能切断最直接的收编管道。向生性空间之所以可能成为一个精英家庭难以“投资”的地带，不是因为里面有什么魔法，而恰恰是因为它被删除了观众、被切断了与升学选拔的通道、被取消了排名与展示的功能。在布迪厄的理论图式中，文化资本的运作有一个不可缩减的前提：它必须在学校场域中被承认为有价值的，才能被转化为象征利润。也就是说，文化资本之所以是“资本”，恰恰是因为存在一个可以将其兑换成学业优势、进而兑换成社会优势的兑换窗口。向生性空间所做的，不是去挑战文化资本的存在本身——那是任何学校都无法做到的事——而是去关闭了它在这个特定空间内部的兑换窗口。没有展演、没有竞速、没有可传递的排名红利，它就失去了被资本化最直接的诱因。优势家庭再多的话语资源，如果找不到一个可以把这些资源兑换成升学红利的窗口，他们在这个空间里投入的动机就会被大幅度削弱——不是因为家庭教育本身的熏陶消失了，而是因为这条通往选拔体系的转化链被有意地切断了。

第二层面：必须诚实地承认，这一策略并不能杜绝所有形式的文化资本渗透。优势家庭的优势，未必总在“投资换红利”的显性逻辑上运作。即便向生性空间切断了分数排名和升学挂钩，那些从小在家庭对话中习惯了对内心经验进行精致语言编码的孩子，仍可能在这种空间的日常互动中——师生交谈、文本回应、小组分享——表现出一种更从容、更被教师的文化直觉所认可的参与姿态。这种“从容”不直接换分，也不进入任何正式评价程序，但它可能在教师作为“见证者”的意识中，留下一层难以完全消除的印象差异——教师也是被嵌入特定文化秩序中的主体，他们的注意力和回应方式不可能完全免疫于文化资本的渗透。

本文对此的回应不是争辩“这种渗透不存在”，也不是在承认之后草草撤退，而是正面做一个发生学的判断：这种更微妙的渗透，与它在向评体系内的运作之间，存在一个质的差别——在向评体系中，象征性优势有一条从学校场域通向劳动力市场和婚姻市场的、被制度所保障的、“可比价的”兑换通道。在向生性空间里，这条通道被人为截断了。一个学生在师生交谈中因为家庭文化资本的熏陶而显得“更从容”，这件事确实发生了，但它无法被量化、无法被等第、无法被纳入任何一份可以递送到下一个制度节点的档案中去。它在这个空间里就停下了。它可以被感受，但不可以被兑换——而正是这种“被感受却不可被兑换”的性质，使它与向评体系中的文化资本再生产形成了质的断裂。资本需要可兑现性才能维持其资本的属性；一旦兑现窗口被关闭，它在这个特定空间里就不再以资本的方式运作——它可能仍然是一种优势，但这种优势已经失去了它最核心的能力：把自己转化为下一个制度环节中可用的通货。

这不是说文化资本消失了，也不是说教育不平等被解决了。这是说，在一个拒绝把任何东西兑换成新一轮筹码的制度夹层中，布迪厄所描述的那种“区隔→认证→继承”再生产链条，至少在最后一个环节——认证——被人为地松脱了。承认这个边界，不是论证的软弱，而是论证的诚实。一种不对自己的限度保持清醒的实践构想，一旦落地，反而更容易在被现实挫败后遭到全面的否定。

质疑三：在当前仍然高度竞争的环境中，这种去选拔化的向生性空间，难道不只是一个在安全的学术论文里才成立的理想假设吗？

这正是全文最需要诚实坦承的一笔。要回应这一质疑，本文不能臆想尚未发生的大规模成功实验，也不能把“最小可行性实验”美化为已经充分验证的方案。本文可以做的是：清晰地给出证伪条件，让这个设想以可被检测、可被推翻的形式立在明处。

七、结论：退出比较的余地

本文的核心论证可以收束为三笔。第一笔，让静态评价难以撼动的，不是锁定本身，而是锁死的方向——它在过去四十余年的演化中被反复锻造为整个教育生态的首选朝向，这个朝向就是本文命名的向评性。它不是一种制度，也不是一项政策，而是一种在具体的历史节点上（1977年的正当性锚定、1999年的家庭决策渗透、2000年代补习资本化的日常承载、2014年改革的意外自我加固）、经由无数个有理有据的局部理性选择、被持续再生出来的集体朝向。向评性不仅是一种制度安排，更是被内化为个体判断事物价值的默认标准：做任何事之前，先问“这能不能被看见、能不能被认证”。

第二笔，向评性之所以如此难以被扭转，根源于一个从未被质疑的底层预设——“评价必定服务于选拔”。在这个预设所圈定的问题框里，所有评价改革的努力都只能在“用什么方法来选拔”上打转，从而必然被向评性的引力场重新捕获。不论你把新评价设计得多么“动态”“多元”“过程性”，一旦它承担了选拔功能，它就立刻被拉向可比较性，成为另一种向评的赛道。2014年综合素质评价的实践，是这一判断最完整的例证。

第三笔，在此诊断下，真正可想象的前进空间，不是在向评性的框架内不断改良评价技术，而是第一次在认知上承认：教育中有整块整块的土地——本文命名为向生性的土地——它们根本就不该被拉进评价的范畴。向生性不是一种“更好的评价对象”，也不是一块被向评性弄脏了的、等待被擦回原样的“纯真状态”。它是在向评性几乎塞满了全部教育时空之后，被逼到边界上的、逆向定义出来的异质地带。它是认知发生中那块在被外部注视时会收回自身的部分。这不是理念的浪漫修饰，而是认知现场里每天都在发生的、一被目光照到就收回的结构性事实：外部注视不仅分散了注意力，更在更深层改写了“什么是好的思考”这项元认知标准本身——思考开始朝着“可被展示”的方向优化，而非朝着“事物本身”收敛。真正的实践方向，不是在向评体系里为向生性增补一些新指标，而是在向评体系的边界外，为它开辟并维持退出比较的余地——哪怕只是几块小到不引人注意的土地。

这一判断在何种情况下可能被证明是错的？它面临两个层面的证伪条件。

第一层面与向评性的解释力有关。如果中国教育体系在可预见的未来，确实发生了一次实质性的位移——不只是增加了新的评价维度，而是成功降低了向评性在家庭决策和学校行为中的实际引力，使得教育生态的整体朝向发生了可观测的偏转，且这一偏转不是由某种外部强制力量所推动，而是由系统内部演化出新的稳定形态所自己达成的——那么，本文关于向评性自我加固能力的分析就需要被大幅修正。那可能意味着，向评性并不具有本文所声称的那种通过吸收改革来再加固自身的韧性。

第二层面与向生性空间的可行性有关。如果出现大面积、长时段的实证研究，证实在那些认真推行了向生性空间实践——也就是说真正做到了内容永久不进入任何选拔、评选、比较程序——的课堂中，学生与教师在经历了一个较长周期以后，并未在深层学习投入、自主探究意愿、课堂安全感或教师职业认同上表现出与对照组的显著差异；如果向生性空间最终被证明只是一个在理论上吸引人但在真实教育场域中无法扎根的浪漫想象，因为教师被向评压力挤得根本无暇他顾，学生也只在刚接触时感到新鲜而很快就因无即时回报而失去兴趣——那么本文所设想的“退出比较”就需要被重新检视。届时必须承认的一个更令人难以接受的判断是：当一个社会用评价来分配社会机会的基本结构稳固不移时，任何在系统缝隙中试图为不可见之物开辟空间的努力，都注定只能成为可有可无的例外。

这两种可能中的任何一种如果被证实，本文的判断都将需要被大幅修正。但是，在目前可见的实证图景中，这些证伪条件并未出现。相反，那些在评价制度的缝隙中坚持为无法计分的成长留出空间的教师——如第三节案例中那位在离职前写下自白的研究性学习课程教师——那些在家长之间以极小规模传递着对“不卷”的渴望的家庭，那些在繁重压力下仍然固执地跟个别学生聊“你最近在读什么书”的老师，都还在微弱但持续地提醒着一件事：完全被向评性塞满的教育，本身就有病。向生性空间不是一个解决这病的方案，它是在等待——等待这个社会有朝一日愿承认，教育中有些最重要的发生，从来就不是为了被比较而来的。

八、最近邻辨析：向评性／向生性不是成就目标定向

本文提出向评性与向生性之后，必须正面回应一个最贴的质疑：成就动机研究早已给出一对高度相似的区分，本文是否只是换了名字？德韦克与莱格特（Dweck & Leggett, 1988）的成就目标理论把学习者的目标定向分为"表现目标"（证明能力、被评判为强）与"掌握目标"（理解内容、把事情做成）——表现目标几乎就是"朝着被看见的方向生长"，掌握目标几乎就是"朝着事物本身的内在逻辑生长"。这是比德西与瑞安的内／外动机（本文已在前文划清）更贴的占位。

本文承认这一占位是实的，但剩余分工清晰可辨。成就目标理论是一种个体心理定向——它落在学习者个人的目标设定上，用问卷可测量、随个体而异、也随情境可切换。向评性与向生性不是个体的心理属性，而是个体、制度与资源在互动中集体偏向的一种生态朝向：它在教育系统四十余年的演化史中被反复锻造，先于任何个体的目标选择而在场，并对每一个进入系统的个体施加拉力。一个持掌握目标的学生，仍会被向评性的生态朝向裹挟——因为向评性不住在他心里，而住在"被看→被认证→被兑换"的制度链条里。

由此得到一条把二者分开的可裁决差别：若向评性可被还原为个体目标定向，则在一个由持掌握目标的学生组成的班级里，向评性应显著减弱。本文预测相反——只要"被认证—被兑换"的制度链条不被切断，即便个体普遍报告掌握目标，可被外部认证的产物仍会系统性地挤出不可认证的探索，向评性作为生态朝向仍然在场。个体目标定向的改变解释不了这一点，因为被弯曲的不只是个体的动机，还有整个系统对"什么值得被生产"的集体偏向。

参考文献

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Bourdieu, P., & Passeron, J.-C. (1977). *Reproduction in education, society and culture*. Sage.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- Dewey, J. (1916). *Democracy and education*. Macmillan.
- Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Vintage Books.
- Han, B.-C. (2015). *The burnout society*. Stanford University Press.
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.
- Pierson, P. (2004). *Politics in time: History, institutions, and social analysis*. Princeton University Press.
- Polanyi, M. (1966). *The tacit dimension*. Doubleday.
- Sandel, M. J. (2020). The tyranny of merit: What ‘s become of the common good? Farrar, Straus and Giroux.
- 吴愈晓. (2013). 教育分流体制与中国的教育分层（1978—2008）. *社会学研究*, (4), 179—202.
- 刘海峰. (2007). 高考改革的理论思考. 华中师范大学出版社.
- 李春玲. (2014). 教育不平等的年代变化趋势（1940—2010）——对城乡教育机会不平等的再考察. *社会学研究*, (2), 65—89.
- 应星, 刘云杉. (2015). “无声的革命”：被夸大的修辞. *社会*.
- 钟启泉, 崔允漭, 张华（主编）. (2001). 为了中华民族的复兴，为了每位学生的发展：〈基础教育课程改革纲要（试行）〉解读. 华东师范大学出版社.
- 周飞舟. (2021). “双减”政策的社会学意涵. *教育研究*, (9), 42—51.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256-273.

修订说明

本文经德麦国际 SDE 学派编辑（Claude）依两轮修订体例打磨：补第八节，与成就目标理论（Dweck & Leggett, 1988 的表现目标／掌握目标）正面切割，明确剩余分工——向评性／向生性是个体、制度与资源的集体生态朝向，而非可用问卷测量的个体心理定向；给出'持掌握目标的班级里向评性是否减弱'的可裁决差别。作者原有向评性／向生性命名、研究性学习案例与三项质疑回应未改。作者的核心判断、论证结构与原初命名均未改动。