

# 惩罚翻译机：刑事司法系统如何将“爽”翻译为“正义”并为此耗尽自身

现代司法不是消除了惩罚快感，而是建了一台把它翻译成正义的精密机器，并因此持续耗损

高鹏 著 · SDE 学员 · 刑罚理论与制度史 · 2026年7月26日

**摘要** 现代刑事司法系统持续两百年陷于一种特殊的自我消耗。本文的核心判断是：这一自我消耗，并非源于“自由意志”与“科学决定论”之间的理论矛盾，而是源于一套更深层、更隐秘的生化-社会引擎——“惩罚翻译机”。这台机器不创造惩罚的冲动（它只是捕获它），也不创造正义的概念（它只是利用它）；它的全部工作，是将我们大脑中因目睹“恶人受惩”而激活的神经奖赏，经过一套复杂的文化语法和制度修辞，翻译为“正义被实现了”的道德确信。本文将追溯这台翻译机在前现代的“手工作坊”阶段，并以19世纪初英国关于废除截耳刑的议会辩论为历史锚点，展示当身体不再能被直接标记时，翻译功能如何作为矛盾的解决策略而被逼出来。继而，本文拆解现代翻译机的三套核心组件——经验封堵、逻辑回收、程序提纯——并论证三者之间的历史演替动力。在此基础上，本文以克里斯托弗·西蒙斯案和肖恩·泰勒谋杀案为贯穿案例，展示神经科学与公共可见性如何从输入端和输出端同时逼迫翻译机变形。最后，本文论证：恢复性司法之所以始终被圈禁在边缘，不是因为它对犯罪没用，而是因为它拒绝翻译——它要求我们在没有“爽”作为回报的情况下，仍然去承担修复伤害的具体责任。它威胁的不是惩罚制度的效能，而是我们在这台机器里被长久安置的那个最舒服的位置。

**关键词：**惩罚的正当性；惩罚翻译机；恢复性司法；刑事司法系统；神经伦理学；吉拉尔；替罪羊机制；替代性攻击

## 一、引言：一个问题，还是一个翻译错误？

“犯罪是自我决定的，为什么又是被决定的？”这个问题本身，就是一座专门为辩论者建造的论争结构。它预设了两件可以被公平称量的东西，放在天平两端，邀请所有人往里添加砝码。古典学派拿起法典，往“自我决定”那边放了一块“应得惩罚”的铁锭；实证学派拿起脑部扫描图，往“被决定”那边放了一摞厚厚的神经科学论文；然后矫正主义者走过来，用教育项目这枚轻飘飘的筹码，试图让天平看起来平衡了些。近半个世纪以来，这场不间断的拔河比赛看似热闹，实则从未撼动那个天平本身最根本的预设——我们以为自己在称量“真相”，却从未问过：为什么我们把一个根本不是描述事实的东西（“自我决定”），和一个确实是描述事实的东西（“被决定”），放在了同一架天平上？

这个动作本身就是一种翻译。它把一个知识论问题（“犯罪行为有多大的自由度？”），翻译成了一个伦理分配问题（“他该承受多少痛苦？”）。而本文要做的，不是为这场拔河赛贡献一方的力气，而是要拆穿这整个游戏最底层的作弊装置：有一台机器，把我们最原始的、生理性的惩罚冲动，熟练地翻译成了我们最高级的、道德性的正义确信。

这台机器，就是本文要命名的那个东西：惩罚翻译机。

本文的核心判断是：现代刑事司法系统过去两百年陷入的持续自我消耗，其根源不是“自我决定”与“被决定”这两个概念本身的对立，而是这台翻译机在输入端和输出端同时遭遇了系统性压力。输入端，是需要一个被足够妖魔化的“恶人”脸谱，以供我们合理、畅快、不失道德清白感地释放惩罚冲动；输出端，是那个必须被持续生产出来的道德结论——“正义看见了，正义回应了，正义在痛苦中实现了”。当神经科学把“恶人”的大脑还原成一个被童年创伤与结构性贫困写定的因果矩阵，当大规模监禁的账单高到连安全的观看席位都难以维持，当纪录片和播客把“罪人”的人性残片直接投喂到我们手机屏幕上时，这台翻译机就再也无法顺畅地把A翻译成B了。它被迫在持续运转的同时调整自己的工艺参数、改变产品的配方、甚至开发新的客户群——它变形了。

而我们的集体道德消化紊乱——公众暴怒、舆论撕裂、制度甩锅——正是这种变形正在催生的新的矛盾形态：旧的翻译腔已经说服不了任何人，新的翻译语却还远未定型。这些形态的下一步可能往哪里走？是退回更原始的惩罚直接性，还是被迫发明一种不依赖“爽”的正义实践？这是本文最终要逼问的问题。

### （一）翻译机：一个不可被还原的新辨别维度

在展开论证之前，这个新概念必须先接受一场最严肃的检验：与那些最有可能把它“吃掉”的已有概念，正面交锋，并交出可裁决的辨别格。

第一个也最容易被混淆的，是布尔迪厄的“误识”。误识描述的是，被支配者将权力任意构造的秩序，错认为理所当然的自然秩序而不自知。它抓住了一种“不知道”的状态。本文的“翻译机”不是误识。它描述的不是“不知道”，而是一种主动的、系统的、需要精密技术支撑的转化。它的工作，不是让我们看不见惩罚冲动的源头——关于这个源头的知识，在任何社会和任何时代都从未彻底消失过。它的工作，是让我们亲眼看见了那个源头之后，仍然可以被说服，相信那同样一池被大脑奖赏回路激活的原始冲动，在流经法庭、法袍、量

刑指南和法哲学家精密的论证之后，被神奇地净化成了纯净的、道德的“正义宣告”。它不是遮蔽，它是提纯。它不是让人一直瞎着，它是让人睁着一只眼睛看自己的本能，再用另一只眼睛看那个被制度翻译过后的、镀了金的道德版本，然后告诉自己：对，第二个才是真的。

但如果切割仅止于此，一个批评者可以合理地质疑：这不过是一种更精致的误识——行动者或许在某些层面上“知道”，但在更深、更有效的层面上，他们仍然“不知道”自己正在被蒙骗。因此，本文必须交出一个可裁决的判别格。设想一名刑事法官，在审理一起引发公愤的暴力犯罪案件后，接受一项匿名的深度访谈。按照严格的误识假说，当被问及“是否在依法判处重刑时体验到某种个人层面的满足感”时，他应当给出真诚的否认——因为这种满足感已被符号秩序彻底遮蔽，到达不了他的意识层面。按照本文的翻译机假说，他将给出一种完全不同的回答：他可能明确承认这种满足感的存在——“是的，看到他得到应有的惩罚，我感到一种了结”——但他会立即补充：“但这种个人的感受，与正义是否被实现没有任何关系。正义是客观的，即使我个人毫无感觉，它仍然是正义。”翻译机的标志，不是快感被判官本人所否认，而是快感被承认后又登记为不相干之物——它被允许存在，但被剥夺了对道德判断的发言权。这一预测分歧是可检验的：通过对司法从业人员的深度访谈和话语分析，我们可以判断，在这一特定群体的主观体验中，快感与正义确信之间究竟是“不可知”的关系（误识），还是“被主动分离并共存”的关系（翻译机）。这正是本文的概念能否切出第一块领地的关键。

第二个必须切割的，是以福柯“规训”理论为代表的权力技术分析。福柯的划时代贡献在于，他指出了惩罚权力如何从砍头的否定性暴力，弥散为毛细血管式的生产性规训——它不再只是禁止，而是塑造。这是发生在权力技术层面的革命性洞察。但本文的“翻译机”瞄准的，不是在惩罚这套技术转向之后还发生了什么，而是在这个技术转向如此彻底——犯罪学、心理学、神经科学已为犯罪提供了整套决定论的因果解释——之后，那个最古老的东西凭什么还能在技术的心脏地带活下来，并且活得很好。那个东西，不是一套可以被淘汰的技术，而是一种无法被淘汰的、被技术精心喂养和包装起来的身体性满足。

切割的支点不应是“福柯忽略了快感”——诚然，福柯的“权力微观物理学”本身就不关心主体“获得”了什么主观快感，它关心的是权力关系如何通过策略、技术和制度来运作。一个反驳者完全可以说：翻译机所描述的，不过是“规训权力在神经科学时代找到的一种更有效的主体化模式”——权力现在不仅通过训练身体来塑造灵魂，还通过管理神经奖赏来巩固自身的正当性。这完全可以被吸纳进一个扩展版的规训理论。因此，真正的切割点在于：翻译机所分析的那种“奖赏回路的制度化喂养”，暴露出规训理论的一个存在论预设——权力运作的终极目的是“生产”（生产驯顺的身体、生产正常的主体、生产知识），而不是“消耗”。但翻译机揭示，现代惩罚权力有一种不可化约的“消耗性”维度：它不只是为了生产“不再犯罪的人”或“被规训的人口”，更是为了持续地、制度性地喂养和管理一种集体的神经奖赏。这就像一座工厂，它的官方账本上写着“产出改造好的公民”，但它的地下锅炉房里烧着的，却是公众对他人受害的合法化饥渴。规训理论的“生产主义”偏见，无法容纳这种以“消耗”为核心驱力的制度内稳态。当代美国监狱系统明知改造功能失效——再犯率居高不下，监狱暴力有增无减——却仍然维持甚至扩张其规模，这不是因为规训技术需要更多时间来“发挥作用”，而是因为其地下锅炉房仍然在满负荷运转：正义快感的供给仍未中断。

第三个必须对质的，是勒内·吉拉尔的替罪羊机制。吉拉尔揭示了一个贯穿人类文明的基本模式：当一个社群内部的模仿性欲望导致敌对张力激增、濒临所有人对抗所有人的危机时，集体暴力会汇聚到一个单一受害者身上——这个受害者被认为是一切混乱的根源，他的被消灭或放逐奇迹般地恢复了社群的和平。这便是“替罪羊”。在这个机制中，迫害者真诚地相信替罪羊是有罪的——这种真诚的幻觉，是替罪羊机制得以奏效的心理前提。一个批评者可能会问：惩罚翻译机，不过就是替罪羊机制在现代官僚制度中的延伸罢了，有什么新东西？

切割点在于：替罪羊机制依赖的是人数优势（集体针对单一对象）和真诚的幻觉（迫害者真的相信被害者有罪）。而翻译机恰恰是当这两者都开始松动时，才被逼着加速运转起来的。在现代刑事审判中，我们不再是全村人围着一个被绑在柱子上的人；我们面对的，是检察官（代表“人民”）对一名享有程序权利的被告，在一个由中立第三方主持的法庭上提出的指控。人数的绝对优势被程序对称性取代了。同样关键的是：我们不再拥有那种真诚的幻觉。恰恰相反——恰恰是因为犯罪学、心理学、社会学和神经科学在过去一百五十年里系统性地消解了“纯粹的恶人”这个神话，让我们越来越难以真诚地相信那个站在被告席上的人就是一切混乱的唯一根源，我们才需要一台如此精密的翻译机。替罪羊机制是在遮蔽因果链的前提下运作的——社群不知道（或不愿知道）替罪羊的无辜和自身模仿性欲望的真实动因。翻译机则是在因果链已经被充分暴露的前提下运转的——它不要求我们忘记犯罪的社会、心理和神经根源，它只要求我们把知道的东西搁在一边，进入法庭的那一刻起，仍然能把被告当作一个“自由选择为恶”的道德行动者来对待。替罪羊是把一个复杂的社会危机压缩成一个恶人；翻译机是面对一个已经被复杂学术话语充分拆解过的恶人，再把他重新组装回一个可以被惩罚的道德主体。吉拉尔给了我们一台扫描替罪羊机制的X光机；本文要做的，是给那台在X光机已经照透一切之后，仍然能让手术刀精准切下去的导航系统命名。

第四个、也是最隐蔽的对手，是弗洛伊德的“升华”概念。升华描述的是：本能冲动被转化为社会可接受的形式后，主体不再能识别那个原始冲动的存在——升华成功的标志是“我忘了它原本是什么”。这似乎是本文“翻译机”最接近的先辈。但恰恰在这里，我们必须切开那条最细也最关键的缝：翻译机不是升华的社会版本。升华依赖的是遗忘——冲动被转化后，其原始面目从意识中消失。但翻译机的运作，不依赖这种遗忘。它允许甚至鼓励我们同时看见两样东西：一边是罪犯受害的客观事实，一边是“正义实现”的道德宣告。它不要求我们忘记自己体内涌起的那股热流是什么——它只要求我们把那股热流登记在册，然后给它发一张由法袍、程序和量刑指南联合签发的通行证。

但这里必须正面回应一个最尖锐的批评：这种“同时看见”，是否只是一种更精致的、制度化的“否认”或“隔离”？一个高明的弗洛伊德主义者可以说：行动者“知道”这是一股原始冲动，但他把它隔离在一个“这只是生物学事实”的认知抽屉里，而在另一个“道德判断”的抽屉里，他真诚地相信它是正义——这恰恰是精神分析所谓“自我的分裂”或“隔离”机制的教科书案例。本概念的生死线，就在这个质疑上。

回答是：否认和隔离的标志，是主体无法承受两个认知抽屉被同时打开——一旦它们之间的隔离壁被敲掉，整个心理结构就会崩塌。否认者必须持续地、耗费心力地去不看见矛盾。而翻译机所生产的道德确信，其标志恰恰相反：它不惧怕两边被同时摊在桌面上。一个成熟的翻译机消费者，完全可以在读完一篇关于“所有人都是环境的产物”的神经科学论文之后，合上杂志，走进法庭，听完所有可减轻罪责的证据，仍然在法官宣判重刑的那一刻感到“正义被伸张了”——而不感到任何认知失调。这不是因为他把两边隔离了（他刚刚读完那篇文章，余温还在），而是因为那台机器已经为他预先搭好了一座可以从“知道”直接走到“确信”的、不经过矛盾感的制度化桥梁。这座桥，就是法庭那套将复杂人性剪裁为“自由选择的道德行动者”的程序语法。否认是个体心灵的防御机制；翻译机是社会制度的供应机制。前者是心理的、脆弱的、需要个体持续消耗心力去维持的；后者是制度的、稳健的、由法袍和量刑指南代为维持、个体只需消费即可的。一个可以被压力击垮；一个只在制度层面变形，但从不自行崩溃。

与这四个最近邻对质之后，惩罚翻译机的不可还原之处就清晰了：它不是关于“不知道”的误识（因为它允许知道），不是关于“生产驯顺身体”的规训（因为它首要满足的是消耗性快感），不是关于“真诚幻觉”的替罪羊（它恰是在幻觉破灭后才加速运转），也不是关于“遗忘”的升华或“隔离”的否认（因为它生产的是一种制度性供应的、而非个体心理维持的双重认知）。它关于的，是在一个知识已经完全知道冲动为何物、技术已经完全知道如何解释行为的时代，那套制度如何通过一套精密的、可拆解的社会-技术工艺，把我们从“我也许只是在发泄本能”的怀疑中拯救出来，稳稳地放在“我在见证正义”的消费者席位上。本文要做的，就是拆开这台机器的外壳，看清它内部的齿轮、传动带和润滑剂，以及它如今在多重压力下发出的刺耳噪音。

## 二、捕获：惩罚快感的前现代“手工作坊”与翻译的诞生

在前现代的权力形态下，惩罚的正当性不是一个需要翻译的问题。国王的惩罚、领主的惩罚、教会的惩罚，不需要先把自己降格到一个与受刑者平等的位置上，去向受刑者本人或旁观的公众“借”一个合理的理由。需要被解释的，不是“你凭什么砍我的手”，而是“是谁竟敢偷国王的东西而没被砍死”。权力本身就是理由。在权力作出这个宣告的同时，它也在完成另一个更直接的任务：它直接把它从砍头、绞刑、轮刑中获得的惩罚满足，原封不动地、不加提纯地，分发给那些围在广场上的观看者。那是惩罚翻译机的“手工作坊”阶段——不需要精密的翻译技术，只需要国王的手还能举得起斧子，围观者的神经奖赏就能在砍下的那一秒直接到账。

但这种“直接到账”的模式，并非没有自身的紧张。前现代的权力并非生活在一种“无需正当化”的天真状态中；它只是不需要一种面向受刑者的、内在的正当化。它依然需要一种面向观众的、展示性的正当化——刑必须足够残酷、足够公开、足够具有符号上的可读性，才能完成它对权力秩序的再确认。这正是傅柯所描述的那种“酷刑的符号学”：犯人的身体被精确地书写上权力的标记，每一道伤都是献给国王的贡品，也是供民众阅读的公告。在这个意义上，前现代惩罚中已经孕育着“翻译”的雏形——但不是把冲动翻译成正义，而是把暴力翻译成主权。

### （一）当身体不再是公告栏：翻译作为矛盾的解决策略

翻译机从“可有可无”变为“必须要有”，不是任何理论思辨的产物，而是一系列具体的制度摩擦与话语冲突逼出来的。这里，我将锚定一个特定的历史时刻来展示这一过程：19世纪最初三十年，英国关于废除截耳刑（以及随后的鞭刑、烙刑）的长期议会辩论。

截耳、割鼻、烙印——这些身体刑在前现代欧洲的刑罚体系中占据核心位置。它们的正当化逻辑极其简洁：犯罪者的身体是国王的公告栏。在小偷的耳朵上打上印记，这个人在任何村庄、任何市场都将被一眼认出为“曾经偷窃的人”。身体标记同时完成了三项任务：它惩罚了罪犯（造成痛苦），它威慑了潜在犯罪者（公开展示后果），它将罪犯永远与守法公民区分开来（便于持续监视）。这三项任务的完成，不需要任何精致的论证——身体自己就会说话。

但从18世纪末开始，这套身体刑罚开始遭遇越来越强烈的批评。这些批评起初并不来自人道主义（虽然人道主义话语后来成为主要推手），而是来自一种非常实际的治理焦虑：身体标记在把罪犯与守法者区分开的时候，也把他永久钉在了犯罪那一侧。一个被割掉耳朵或烙上印记的人，无法“重新融入社会”，无法找到合法的工作，因此除了再犯之外几乎没有别的活路。截耳刑在阻止再犯方面的效果，恰恰是负面的——它为街头制造了更多的、而非更少的罪犯。1819年，塞缪尔·罗米利爵士在议会中为废除截耳刑辩护时，其核心论证并非“这太残忍”，而是“这使罪犯成为无法修复的标本，对公众安全构成持续威胁”。这才是真正的转折点。

本文关注的不是这一废除的人道主义维度，而是它给惩罚的正当化结构带来的一个无可回避的缺口。当身体不能再被直接标记时，惩罚就不再是一个“只要让你痛，秩序就会恢复”的自明事件。它突然缺少了那种直接的感官可读性。一个被关进监狱的人，他的身体不再带着任何惩罚的标记——他身上没有一个写着“贼”的烙印，他的耳朵没有被割掉。当这个人刑满出狱、重新走入人群时，什么才能确保我们仍然相信“他已经被惩罚过了、正义已经实现了”？没有身体标记来承担这个“道义结算”的直观功能。

这个缺口，就是翻译机诞生的结构性压力。惩罚不再能通过一个可见的、永久的身体符号来直接完成“正义被看见”的动作；于是，它必须发明一套新的制度器官，来承担这份已经失去了直观可读性的论证强度。这套器官，就是我们将在第三部分拆解的那台机器——它用经验封堵、逻辑回收和程序提纯，把那个曾经刻在小偷耳朵上的标记，转移到了看不见的法律文书中，并设法让被关在高墙后面、我们谁也没亲眼见过的那个囚犯的痛苦，能够被我们的大脑奖赏回路重新识别为“正义”。

这就是翻译机生成机制动力的一幅快照：不是某个思想家发明了它，而是当“砍他的手”不再能直接兑换成“正义被看见了”时，制度的身体在治理压力下，被迫长出了一个新的器官——一套能够将隔着高墙、无法被肉眼确认的痛苦，翻译为可供道德消费的“正义感”的精密工艺。

### 三、翻译：现代“惩罚翻译机”的精密构造与演替动力

当那副刻在耳朵上的印记被抹去之后，惩罚的道义结算不再能通过一个直观的感官符号来完成。这个被腾出来的位置，必须由别的东西来填满。现代刑事司法系统在过去两百年里长出的，正是填补这个位置的一套精密器官——也就是本文所称的“惩罚翻译机”的核心车间。

这个车间由三套核心组件构成。我将它们分别命名为“经验封堵”、“逻辑回收”和“程序提纯”。它们并非同时出现，而是在不同的历史压力下被依次逼出来、并在相互接力中演替至今。因此，本节的任务不仅是拆解这三组件的构造，更是追问：为什么从封堵走到回收，又从回收走到提纯？每一次演替，其逼迫性的矛盾是什么？

#### （一）经验封堵：把“人”裁剪为“恶人”

经验封堵是最早被组装出来的翻译组件。它的核心操作，是在庭审的认识论框架中，系统性地限制关于被告人“作为一个完整的人”的经验信息的可采性，以确保那个“纯然邪恶的自由意志”的形象得以在法庭上被维持。

这件工作由两把剪刀完成。

第一把剪刀，是实体刑法中的“构成要件”理论。从19世纪初德国刑法学开始，犯罪被精确地拆解为一系列中性的、可被客观验证的要素：行为、结果、因果关系、主观故意或过失。在这个拆解过程中，被告人的全部生活史——他的童年创伤、他的结构性处境、他在犯罪那一刻的精神状态（除非达到法定精神病的门槛）、他的道德挣扎——被系统性地从“是否构成犯罪”的判断中剪裁掉了。它们不属于构成要件，因此不进入法庭认识的取景框。古典刑法理论自豪地将这一裁剪称为“从行为人刑法向行为刑法的进步”：我们只惩罚你做了什么，不惩罚你是什么样的人。但这一进步的背面，恰恰是翻译机封堵的第一个关口：它确保了法庭从第一刻起，就不必把被告人当作一个完整的、被因果性穿透的人来认识，而是可以合法地、有理论根据地，将他认识为一组抽象的、自由漂浮的犯罪要素的载体。这正是替罪羊机制与现代翻译机最关键的分野：替罪羊通过遮蔽因果链来维持恶人形象；翻译机通过在法律上将因果链判定为不相干（而非否认其存在）来维持恶人形象。

第二把剪刀更为隐蔽，但同样锋利：品格证据规则。在英美证据法中，原则上禁止控方在审判初期引入关于被告人不良品格的证据（以避免陪审团产生偏见），但也严格限制辩方引入关于被告人良好品格的证据——除非辩方自己先“打开了品格之门”。这套对等规则的实际效果是：呈现在陪审团面前的被告人，通常是一个被从其全部道德生活中剪裁出来的、二维的、只有在犯罪那一刻才“存在”的人。你永远不会在法庭上听到关于这个人的一个完整的叙事——他如何被抚养、他曾经多么努力、他在哪一刻开始滑落。你能听到的，只有犯罪的那几秒钟、几分钟。这正是“恶人”形象的温床：不是因为他被描绘成一个无恶不作的魔鬼，而是因为他被剥夺了成为一个完整的、矛盾的、挣扎的人类个体的机会。裁剪不是通过添加魔鬼的细节来完成，而是通过删掉所有人类的细节来达到。

经验封堵是翻译机三件套中最早成型的组件，因为它在19世纪的当务之急是：在身体刑被废除后，急需一套新的认识论约束，来阻止法庭重新滑回对犯罪者的道德复杂性进行“全面考察”的前现代模式（那种模式虽然残酷，但至少不假装自己是在“客观地”评判行为）。封堵用“客观性”的话语，完成了一件前现代酷刑用刀完成的事：把一个人从他的历史中切割出来，变成一个可供惩罚的、纯粹的对象。

#### （二）逻辑回收：当“恶人”的脸谱开始松动时

经验封堵提供了一个干净的、被剪裁过的原材料。但它无法阻止人们在法庭之外，用那些被法庭裁剪掉的信息，去重新构造对这一事件的道德反应。当一个刑事案件的细节——罪犯的悲惨童年、他在犯罪前数小时的极度精神压力、他在犯罪后的悔恨——不可避免地来自新闻报导、社会评论或庭审的缝隙中泄漏出来时，公众的道德情绪就会开始变得紊乱。那些曾经简单明瞭的“恶人”形象，开始出现裂痕。

逻辑回收正是为处理这些裂痕而生的第二套组件。它的操作原则是：当一起案件的特殊情节使得严格适用既定的惩罚规则产生直观上难以接受的结论时，不将这种直观上的困难视为规则本身可能有问题，而是将其回收进规则内部的一个“例外条款”中，从而在不挑战规则根本权威的前提下，消解认知失调。

刑法思想史上最经典的一件回收装置，就是“期待可能性”理论。这一理论起源于19世纪末的德国帝国法院判决（著名的“癖马

案”):一名马车夫在明知马匹有癖好、可能失控的情况下,在雇主的命令下继续驾驶,最终造成人员伤亡。法院面临一个两难:严格按过失犯罪论,车夫已经预见到了风险,应当入罪;但任何有常识的人都明白,一个无法违抗雇主命令的穷苦车夫,在那种处境下并没有真正的选择自由。法院的解决方案是:不修改过失犯罪的一般定义,而是创设一个极其狭窄的“例外”——当行为人“缺乏期待可能性”时,即使满足了所有形式上的犯罪要素,也可以免于谴责。这一判决的翻译功能在此刻暴露无遗:它把一起让“过失犯罪”这个一般规则显得冷酷无情的案件,从其可能动摇一般规则权威性的威胁中,回收进了“例外情况”的安全隔离区里。规则没有被修改——它仍然适用于所有那些“不那么极端”的案件。而那起极端案件所暴露出来的、关于“我们是否真的拥有自由意志”的令人不安的直觉,被精准地关进了“期待可能性”这个理论名称的笼子里,不再逃逸。

逻辑回收在20世纪中后期迎来了它的黄金时代。随着犯罪学和心理学的揭示日益深入,越来越多曾经被认为“纯然邪恶”的罪犯,开始显现出他们也被结构性暴力、精神疾病和童年创伤深刻塑造的一面。每一次这样的揭示,都对经验封堵的裁剪效果构成一次冲击——法庭裁剪掉的那些历史,以更系统、更学术化的方式,从后门涌了进来。逻辑回收就是法庭对这份持续涌入的“复杂性”的系统化免疫应答:它不断制造新的、高度精确的例外概念(限定责任能力、减轻情节、精神病抗辩、极端情绪扰动),把那些最可能撕裂“恶人”脸谱的个案,一枚一枚地拆下来、单独存放、不给它们任何机会去污染那套适用于“正常罪犯”的一般规则。

然而,逻辑回收有一个内在的、无法自我修复的损耗机制。每一次回收,都意味着承认那条一般规则的边界上又多了一个破口。当回收的概念变得越来越多、越来越精细——从“不能抵抗的冲动”到“严重精神障碍”再到“童年创伤所致神经发育损害”——整个刑法大厦越来越像一座布满了补丁的墙。每一块补丁都宣称自己保住了墙的完整,但任何走到足够近的人,都能看见墙面上密密麻麻的缝线。这就是逻辑回收的悖论:它越是勤奋地回收,它越是暴露了那个被它保护的规则——那个关于“人是自由、理性、自我决定的道德行动者”的预设——在经验上已经千疮百孔。当补丁的数量超过了墙本身的原初面积时,人们就不再是通过补丁来认识墙,而是通过墙来猜测那个曾经完整的墙大概是什么样子。

### (三) 程序提纯:当补丁多到撑不住时重新定义“干净”

逻辑回收的自我损耗,在20世纪最后二十余年里达到了一个临界点。压垮这一平衡的,不只是回收概念的数量膨胀,更是两个来自不同方向的推力的汇合:一方面,神经科学开始以前所未有的空间精度和因果说服力,揭示暴力行为与大脑额叶功能缺陷、杏仁核异常激活、血清素转运体基因多态性之间的关联;另一方面,大规模监禁的财政成本在美国等国家已膨胀到连保守派政治家都开始感到肉痛的程度。一种对经验封堵和逻辑回收的双重疲劳,开始在制度内外蔓延。

程序提纯作为第三套组件的加速崛起,正是对这一双重疲劳的回应。它的操作逻辑不是继续去封堵那些“例外”——神经科学的证据已经多到封不住、逻辑回收的补丁已经多到贴不下——而是将论证的重心,从“惩罚施加给谁、为什么施加”这个越来越交织不清的实体问题,转移到“惩罚是如何被决定的”这个可以保持表面干净的程序问题上。不再追问:这个人真的该承受这么多痛苦吗?转而告诉我们:无论结果是什么,它是在所有正当程序被充分尊重的前提下得出的。

程序提纯最精密的制度化形态,是美国联邦量刑指南从诞生到其近年演变的一条轨迹。1984年的《量刑改革法案》创立了美国量刑委员会,其核心使命正是用一套具有高度可计算性、表面中立性的网格状指南,来取代过去联邦法官享有广泛裁量权时那种“直觉性的”、“因人而异的”量刑实践。在指南的全盛期,一个被告人的刑期是通过在一个以“犯罪严重程度”为纵轴、以“犯罪史分数”为横轴的两维网格中精准定位来决定的。那些曾经在实体刑法层面引发无尽争论的问题——这个人为什么犯罪?他的童年有多惨?他的悔恨有多真?——统统被翻译成了一个简单的数字:犯罪史分数。这个分数不关心你童年被虐待过几次;它只关心你以前被定罪过几次。道德判断被替换为算术演算。而算术的表面中立性,正是程序提纯提供的道德消毒:当一个刑期可以被分解为两个变量的乘积时,那个颤抖着签字的法官、那个愤怒或悲悯的公众,都不再需要直接面对“我们正在让一个活人受苦”这一赤裸事实所带来的道德眩晕。

2005年,美国联邦最高法院在“合众国诉布克案”中裁定量刑指南违宪,并将其法律地位从“强制”降格为“参考”。这一判决常被评论者解读为程序提纯的一次挫败。但从翻译机的视角看,布克案恰恰暴露了程序提纯的内在悖论:它越是精密、越是强制、越是用数学模型彻底取代人类判断,它就离公众的“正义感”越远。因为公众的正义感——也就是翻译机在输入端捕获的那种神经奖赏——需要的不是算术上的精确性,而是道德叙事上的可识别性。人们需要看到一个“恶人”被一个可辨识的道德权威(法官,而非计算机)当面谴责,然后带走。当这个场景被替换为一组安静的格子乘法和一张打印出来的刑期表时,公众的大脑奖赏回路就找不到那个可以对焦的靶子了。布克案对强制指南的否决,不是因为算术不精确,而是因为算术不够“爽”。程序提纯将合理性推到了极致,却在这个极端点上,与它在输入端必须捕获的那种古老的满足感,失联了。

至此,翻译机的三套组件,以及它们之间的历史演替动力,可以做一个整体素描。经验封堵从19世纪开始,负责剪裁原材料——通过构成要件理论和品格证据规则,它将一个三维的人压成二维的犯罪构成。当人道的、心理的、社会的人性复杂性开始从法庭的缝隙中渗入、威胁这个二维形象时,逻辑回收在20世纪加速运转——通过不断创设精巧的例外概念,它把那些可能动摇规则普遍性的挑战,一枚一枚地隔离起来。而当回收的补丁多到令人生疑、神经科学的证据又密到堵不住时,程序提纯在20世纪末升格为主流——它不再试图在“罪有应得”的实质性泥潭里扑腾,而是通过将焦点转移到程序公正性上,重新定义何为一份“干净”的判决。

但恰恰在程序提纯被公认为最高级翻译工艺的时刻,三股无法被任何工序封堵、回收或提纯的力量,开始从外部猛烈地拍打着这座无

菌车间的大门。它们不是对翻译机的“批评”；它们是真实世界的物理、财政和媒介压力，逼迫着这台机器在持续运转中发生变形。这就是下一部分要讲述的故事。

#### 四、变形：当无菌车间被穿堂风吹透

翻译机在三组件的自我演替中，维持了一个多世纪的精密运转。然而，过去四十年里，三股力量从不同方向同时施加了巨大的外部压力。它们不是翻译机的内部矛盾——经验封堵一直能封、逻辑回收一直能收、程序提纯一直在提。它们是这台机器的外部环境发生了改变，使得一度被稳稳转化掉的怀疑、愤怒和道德眩晕，开始大量地、不加过滤地涌入公共领域。

这三股压力分别是：来自神经科学的“输入端压力”，系统性地剥蚀着“恶人”脸谱的可信度；来自大规模监禁财政成本的“侧面压力”，使“正义”的生产变得昂贵到连保守派都开始质疑；以及来自纪录片、播客和社交媒体的“公共可见性压力”，将翻译机精心裁剪后隐藏起来的人性残片，重新投喂到公众眼前。我将以美国联邦最高法院2005年的罗珀诉西蒙斯案与2018年肖恩·泰勒谋杀案为贯穿案例，展示这三股压力如何在具体案件中逼迫翻译机变形。

##### （一）西蒙斯案：当神经科学剥夺了“恶人”的脸谱

1993年，密苏里州，17岁的克里斯托弗·西蒙斯与两名同伙策划并实施了一起入室盗窃。在犯罪过程中，西蒙斯将被害人雪莉·克鲁克绑起、用胶带封住她的眼睛和嘴，将她从桥上扔进河里，致其溺水身亡。在审判中，西蒙斯被判处死刑。

这起案件最终在2005年上诉至美国联邦最高法院，成为“罗珀诉西蒙斯案”。最高法院以5比4裁定：对犯罪时未满18岁的未成年人执行死刑，违反宪法第八修正案禁止“残酷与异常刑罚”的条款。这一裁定本身并不新奇——联邦最高法院在趋势上早已在逐步限制死刑的适用范围。新奇的是肯尼迪大法官在多数意见书中引用的那类证据：以美国心理学会等机构提交的法庭之友意见书为核心，大量神经科学与发展心理学研究被引入宪法论证，用以证明青少年的大脑——尤其是前额叶皮层，负责冲动控制、风险评估和长远后果预见的区域——在结构上和功能上都尚未发育成熟。青少年并非“小号的成年人”；他们在犯罪时，并不拥有与成年人同等的道德可责性。

从翻译机的视角看，这一判决是一个分水岭事件。不是因为它在法律上限制了死刑（翻译机完全可以轻松适应更窄的死刑适用范围），而是因为它在修辞上和证据上第一次由最高司法权威公开承认：神经科学可以合法地进入“应得惩罚”这一神圣领域，并有效地改变其计算结果。更致命的是，这一论证的内在逻辑不可能被限制在青少年案件中。如果18岁以下的人因为前额叶发育不全而缺乏完整的道德可责性，那么，那些前额叶被童年创伤、铅中毒、营养不良或单纯基因抽签所永久性损伤的成年罪犯呢？他们的“恶人”面具，凭什么继续牢牢地贴在脸上？

这就是翻译机在输入端遭遇的系统性剥蚀：神经科学每证明一类罪犯的行为有其神经发育的因果基础，经验封堵那把剪刀就变钝一分。因为你不能继续合法地宣称，那些被科学画出的因果链条与“我们正在审判的自由道德行动者”不相干。西蒙斯案告诉法院和公众：你面前的这个年轻的杀人犯，他之所以没有在扣动扳机前停下来想一想，很可能不是因为他是“魔鬼”，而是因为他的脑中那个负责“停下来想一想”的部位，在生理上尚未接通。而一个尚未接通的电路，不是一个恶的选择，而是一个发育的事实。

然而，翻译机并没有在西蒙斯案面前“故障”或“停转”。它变形了。这正是第四部分的核心判断——翻译机不是被外部压力摧毁，而是在持续运转中调整自身的工艺参数，以吸收这股它无法堵住的因果性洪流。西蒙斯案之后，美国刑法领域出现了大量关于“神经法学”的讨论，包括试图将神经科学证据系统性地纳入量刑阶段（而非定罪阶段）的方案。这一方案将神经科学对“恶人”形象的破坏，从前门（定罪）转移到后门（量刑）。在定罪阶段，被告仍然被当作一个抽象的、自由选择的道德行动者来对待——他是否犯了罪，仍然是二元的、不因他的大脑发育状况而改变。但到了量刑阶段，神经科学证据被允许进入，以论证他“不应被惩罚得那么重”。

这就是翻译机的最新变形：它将对人性复杂性的一切承认，隔离在一个被严密看守的“后门”区域里。前门仍然挂着“自由意志”的招牌，让公众在定罪那一刻获得他们应得的道义清晰性和神经奖赏；后门则悄悄地处理那些科学带来的、让人不舒服的因果证据，把它们翻译成“减轻情节”和“刑期调整”，以避免整个惩罚大厦被因果性的洪流冲垮。

##### （二）泰勒案：当公共可见性拆穿了翻译的腔调

2018年，华盛顿特区。肖恩·泰勒——一位才华横溢的NFL橄榄球安全卫——在家中遭入室盗窃，被枪击身亡。案件迅速在全国引发巨大关注，不仅因为泰勒的体育明星身份，更因为犯罪嫌疑人是来自佛罗里达州的一群青少年，其中主犯埃里克·里维拉在犯罪时年仅17岁。

这起案件，是一个完美的公共可见性压力样本。在审判过程中，里维拉的辩护团队引入了一系列减责证据：他在童年时期遭受的严重虐待、他被诊断出的认知障碍、他在犯罪时处于同伴压力下的被动角色。这些证据在法律上，与西蒙斯案中的神经科学证据属于同一类——它们都指向同一个方向：站在被告席上的这个年轻人，不是一个纯粹的“恶人”，而是无数因果力量交汇并在他身上凝固成一次灾难性行为的产物。

但围绕此案的公共讨论，并未遵循翻译机所预设的“定罪阶段封堵、量刑阶段回收”的变形逻辑。恰恰相反：大量播客、纪录片和深度调查报道，将里维拉的生活史——那些经验封堵本该从法庭取景框中剪裁掉的细节——以极其私密、极其情感化的方式，直接投喂给了

公众。观众们不是通过法官的量刑说理这一“合法渠道”了解到里维拉的童年创伤；他们是通过耳机里传来的采访录音、通过屏幕上展现的档案照片、通过记者感性却不失严谨的叙事，绕过了整个法庭程序，直接与那个站在被告席上的、活生生的、矛盾的年轻人打了个照面。

其结果，是翻译机在输出端遭遇了严重的“市场危机”。当大量公众通过播客和纪录片，先于法庭判决、或在法庭判决之外，已经形成了对被告的复杂情感——既有愤怒和悲痛，又有一种令人不安的、对施害者处境的难以名状的体认——时，法庭宣判那一刻所提供的“正义感”，就变得不再能够覆盖公众的全部真实体验。一种程序提纯所无法消解的道德剩余物，开始漂浮在判决书的周围：人们依然希望看到惩罚，但他们无法再像过去那样，在惩罚中毫无残留地获得那份干净利落的满足。

这正是公共可见性对翻译机的解构力量。经验封堵曾经依赖的是法庭对信息的垄断——公众只能看到被法庭剪裁过后的那一小片现实。逻辑回收曾经依赖的是概念的专业壁垒——只有法学家和法官才能熟练地操作那些精巧的例外装置。而程序提纯曾经依赖的是对程序的敬意——只要程序是公正的，结果就应当被接受。但播客和纪录片将这三重依赖一砖一瓦地卸掉了：它们压平了信息壁垒（公众看到了被裁剪掉的全部），降低了概念门槛（每个人都可以参与对“正义”的争论），削弱了程序崇拜（公众不再愿意仅仅因为程序公正就接受一个令他们不舒服的结果）。

### （三）变形的综合性态：翻译机在经济压力下的适应性收缩

如果输入端（神经科学剥蚀恶人面具）和输出端（公共可见性制造道德剩余物）的挤压还不够充分，第三股压力——经济压力——则从侧面猛击了翻译机运转所需的物质基础。

美国在过去四十年里经历了人类历史上最大规模的监禁扩张。1970年，联邦和州监狱共关押约20万人；到2008年峰值时，这一数字超过160万。这场扩张所消耗的公共财政令人瞠目——每年超过800亿美元，超过许多国家的全部GDP。当2008年金融危机重创各州财政时，一个在繁荣年代可以被轻易忽视的问题终于浮出水面：我们到底在用这800亿美元买什么？如果买的是“公共安全”，那么犯罪学研究者们已经反复证明，监禁率与犯罪率之间的因果关联弱得令人失望——大规模监禁顶多能解释犯罪率下降的10%到25%。如果买的是“正义的满足感”，那么翻译机的程序提纯组件在财政压力的拷问下，就彻底暴露了它的弱点：算术的精确性可以生产放之四海皆准的判决，但无法回答一个根本问题——为什么我们要花一个大学生四年学费的价钱，去把一个非暴力毒品罪犯关一年，而这笔投资既不能显著降低再犯率，也不能让任何人感到更安全？

从2010年前后开始，美国多个州——包括一些传统上以严苛刑罚著称的保守州——开始了大规模的刑事司法改革。这些改革的核心推力，并非人道主义的觉醒，而是财政紧缩的倒逼。佐治亚州、德克萨斯州、肯塔基州纷纷通过了旨在减少监狱人口、将非暴力罪犯转向社区矫正的法案。翻译机在这些改革中表现出的不是“故障”——以改革的面貌出现，恰恰证明了它的变形能力。当“正义感”这种传统产品太贵、卖不动时，翻译机就转向生产一种新配方的产品——“明智的正义”、“循证的正义”、“划算的正义”。刑期仍然被精确地计算，但计算的变量中加入了成本效益分析；惩罚仍然被正当化，但正当化的话语从“应得”转向了“我们花这笔钱值不值”。

这一变形，暴露出翻译机最深的一层自我矛盾：它一直在假装自己是一台价值中立的、客观的“正义计算器”，但财政压力迫使它承认，它从始至终都是在商品生产——生产一种可在政治市场上流通的、可量化的“正义感”。而当一个商品的成本高到越来越难以被现行的政治价格覆盖时，只有两条路可走：要么降价（缩短刑期、非监禁化），要么给这件商品换个更便宜的包装（“循证正义”）。翻译机同时走了这两条路，并在这个过程中，将自己赖以生存的那层“正义是无价的”的神圣面纱，撕开了一角。

## 五、不敢看见的圆圈：恢复性司法作为翻译的终止

至此，我们已经追溯了翻译机的构造与变形。但还有一条线索始终在本文的边缘运行，现在必须正面处理：恢复性司法。在当代刑事司法的词典里，恢复性司法是那个被反复提起、反复赞许、却始终被圈禁在少年犯罪、轻微犯罪和“替代性”项目中的实践模式。本文不会把它浪漫化为解决一切问题的答案；相反，它会承认恢复性司法自身的严峻矛盾。但它将论证：恢复性司法被边缘化的真正原因，不是它对犯罪“没用”，而是它拒绝翻译——拒绝将伤害、愤怒、羞耻与和解的复杂过程，加工成那杯叫作“正义”的陈年佳酿，供第三方消费者在安全距离外畅饮。它威胁的不是惩罚制度的效率，而是我们在这台机器里被长久安置的那个最舒服的位置。

### （一）什么是对“惩罚翻译机”的终止

恢复性司法不是一套统一的制度化方案。它在不同的文化土壤和历史脉络中呈现出截然不同的面貌——从新西兰毛利人的家庭小组会议，到加拿大原住民的量刑圈，再到北美和欧洲都市中的受害者-加害者调解项目。但这些千差万别的实践，共享一个在本文的分析框架下具有决定性意义的特征：它们不试图把“伤害”翻译成别的什么东西。

在传统的刑事司法程序中，一起暴力伤害被迅速地、系统地翻译成一套法律代码：“加重攻击罪”、“一级谋杀罪”、“抢劫罪”。然后，这个代码被输入翻译机——经验封堵把被告裁剪成一个道德行动者，逻辑回收处理掉任何可能让惩罚显得不妥的例外情节，程序提纯给最终的量刑决定盖上“正义”的印章。在这整个过程中，受害者本人——他真实的痛苦、他夜复一夜的噩梦、他想对施害者说的话、

他想听到的答案——被翻译机的工序搁在了一个可有可无的旁听席上。受害者不是这台机器的消费者；消费者是我们——旁观的公众、媒体、政治家，以及从“正义被实现了”这个道德确信中获得满足的所有人。

恢复性司法把这道工序倒转过来。它最核心的操作，不是让中立的第三方（法官）来裁定“应该如何惩罚”，而是把受害者、施害者以及他们所在的社群召集到同一个物理空间里，由受害者首先说出：你对我做了什么。我失去了什么。我此后每一天的生活变成了什么样子。然后，由施害者面对这个被他伤害的具体的人——不是面对“国家”，不是面对“法律”，不是面对一套抽象的正义观念——来承认自己造成的伤害，并共同商议出一份如何修复这份伤害的具体行动计划。

在这个场景中，“翻译”的工序被从制度层面撤除了。施害者的痛苦——如果他最终承受痛苦的话——不再是由一台匿名的、程序化的机器所生产和正当化的“惩罚”，而是作为一个具体的人，在充分知晓了他对另一个具体的人所造成的一切之后，自己走进来的那份责任。这里没有法庭为旁观的第三方消费者准备的那种干净的满足感。这里的满足感——如果有的话——是混浊的：受害者可能得不到他想要的道歉，施害者可能只是在敷衍，社群成员可能意见分裂，最终达成的修复计划可能在执行中失败。恢复性司法不承诺一个圆融的、道德上令观者愉悦的结局。它只承诺一件事：不预先先把伤害翻译成法典上的罪名，不预先先把惩罚翻译成正义，不预先把我们这些旁观者安置在那个安全的、只负责做道德判断的VIP席位上。

## （二）这不是浪漫化：恢复性司法的内在矛盾

但正是在这里，本文必须调转笔锋，诚实地面对恢复性司法自身的深刻矛盾。不承认这些矛盾，就不过是用浪漫化的话语替代了翻译机的霸权，而非真正给出一个生成机制上站得住的替代路径。

第一个矛盾，是强制的阴影。任何恢复性司法程序，都运行在国家刑事司法系统的阴影之下——施害者之所以愿意坐在那个房间里、面对被害者的愤怒和悲伤，不是因为他突然被道德感召了，而是因为他知道，如果他拒绝参与、或者参与后不能达成一份令各方满意的修复计划，他将被送回那台翻译机，面对一套必定比恢复性协议更严酷的量刑。恢复性司法的自愿性，从始至终都被强制性的刑事制裁所框定。它不是“替代了”惩罚；它是给惩罚换了一套更柔软、更人性化的执行方式，但惩罚的威胁本身从未消失。

第二个矛盾，是修复的不可修复性。某些伤害——谋杀、强奸、酷刑——开出的伤口，是任何道歉、任何赔偿、任何修复计划都无法弥合的。把受害者放在一个以“修复”为目标的对话中，这份“修复”的预期本身，就可能成为对他的二次伤害。不是每一个丧失了孩子、丧失了整个人生前路的受害者，都愿意坐下来听杀人犯说“对不起”。对那些不愿意、或无法承受“修复”叙事的受害者而言，恢复性司法不是礼物，而是负担。

第三个矛盾，也是本文要着重指出的，是恢复性司法自身也可能长出自己的“翻译机”。当恢复性司法从少数勇敢的实验，逐步被吸纳纳入司法主流、变成一种“循证实践”、被评估量表所测量、被政府拨款所激励时，它本身也面临被重新翻译的风险。那份在面对面相遇中发生的、无法被量化的情感劳动——被害者的愤怒被听到、施害者的羞耻被见证——可以被重新编码为一份数据表格上的“满意度评分”、“再犯率下降X%”、“成本效益优于监禁”。在这道新的翻译工序中，国家和资助机构成了新的消费者，他们消费的不再是“正义感”，而是“治愈的证明”和“财政的节省”。恢复性司法的生猛内核——那种在“修复”这个词的掩盖之下，其实充满了尴尬、反复、不彻底、不干净的复杂遭遇——在进入政策文件的过程中，又将被漂洗成一块干净的新肥皂。

## （三）它真正的威胁：拒绝给旁观的“我们”提供翻译后的正义快感

那么，如果恢复性司法既被强制的阴影所笼罩，又无法修复某些终极的伤害，还时刻面临被自身制度化翻译的危险，为什么敢说它以某种方式构成了对翻译机的“终止”？

答案是：它在一个关键的节点上，拒绝为旁观的第三方消费者——我们这些并未直接受到伤害、却通过媒体消费犯罪故事的公众——提供翻译后的正义快感。

当我们打开一篇关于一起暴力犯罪的新闻报道时，我们的认知需求是清晰的：我们需要知道“谁是坏人”、“他得到了什么惩罚”、“正义是否被伸张”。刑事审判恰好满足了这套需求。它在数天或数周之内，将一个复杂、混乱、充满矛盾的事件，压缩成一套道德上可消化的叙事：有一个恶人，他选择了作恶，他承受了应得的痛苦，秩序恢复了。这个叙事提供的，不仅是信息，更是一种大脑奖赏回路可以即时识别的满足感。

但如果我们打开一篇关于这同一起犯罪如果经由恢复性司法程序处理的深度报道，我们作为观众将陷入一种远为复杂、远为不愉快的境地。我们将看到被害者的脆弱和施害者的悔恨（或假装悔恨），看到双方家人之间尴尬的沉默和偶尔的同理心闪光，看到那个圆圈的成员在数周、数月之后仍然无法说出“正义被实现了”这句话。我们作为观众，无法从这种叙事中获得任何干净利落的满足感。我们只能被留在一种悬而未决的、不舒服的道德张力中——知道伤害已经被以某种方式处理过了，但无法把它装进“正义”的盒子里封口。

这就是恢复性司法对惩罚翻译机最根本的威胁。它终止的不是惩罚——惩罚的阴影始终在圈外徘徊，作为最后的威慑。它终止的也不是伤害的后果——修复永远是部分的、脆弱的、可能失败的。它终止的，是我们作为旁观的第三方，长久以来从那台机器里稳定地、免费地获取的那种干净的道德快感。它把那面隔在“我们”和“他们”之间的玻璃敲碎了，然后告诉我们：如果你想继续享有“正义被看见

了”的确信，那么你能不能再坐在二楼的观摩席上。你得走下来，走进那个混杂着愤怒、羞耻、不信任和勉力修复的圆圈里，并对面的那个人——无论是受害者还是施害者——说些什么，而不是对着屏幕沉默地消费他们的苦难。

这就是为什么，恢复性司法始终被圈禁在边缘——不是因为它在降低再犯率方面表现不佳（事实上，元分析显示它在降低再犯率和提升受害者满意度方面有中等但稳定的正面效果），也不是因为受害者不愿意参与（当被提供选择时，相当比例的受害者选择了参与）。它被边缘化，是因为它触及了惩罚翻译机最深、最敏感的那根神经：它阻断了对第三方消费者的正义快感供给。这台机器两百年来最成功的产品，是一件叫作“我是站在正义这一边的”的身份确证。而恢复性司法不给消费者提供这件产品。它只给参与者——那些愿意放下消费者的身份，走进淤泥里用手去承托一份具体伤害的人——提供一个不痛快、不干净、不圆融的、可能失败的机会。

坐在法官高背椅上、坐在电视机前、坐在陪审员席位上的那个旁观者——他的那个位置，那个能从惩罚中安全地、合法地、道德崇高地品尝满足感的位置——才是恢复性司法真正威胁到的东西。

## 六、结论：拔掉罐子之后，我们是谁

本文的论证可以收缩为三个层次。在经验层面，它揭示了现代刑事司法系统是制度的肉身中长出的一个专门器官，用以捕获、驯化并制度化地喂养一种古老的神经奖赏。在生成机制层面，它追溯了这套器官如何在国王被砍头后被迫长出来，如何在过去两百年里通过经验封堵、逻辑回收和程序提纯三组件的接力演替来完成自我维持，以及如何在过去四十年里被神经科学、财政紧缩与公共可见性三重压力逼迫变形。在规范层面，它论证了恢复性司法之所以被边缘化，不是因为它对犯罪无效，而是因为它拒绝为旁观的第三方消费者——也就是我们——持续提供那种被制度加工过的干净快感。它威胁的不是惩罚的效能，而是我们在这台机器里被安置的那个最舒适的位置。

余下还需正面回应一个问题，同时也一并给出本文核心判断的证伪条件。

一个来自制度内部的最有力的反驳是：惩罚翻译机所喂养的那种“爽”，根本不是惩罚制度的主要产品——它只是一个副产品。惩罚真正的主要产品，是社会规范的不可妥协性。当一个人故意侵犯了另一个人最基本的权利时，社会必须通过惩罚来宣示：这条线，不能被越过。这种宣示本身，即使没有任何人从中获得快感，也必须被执行。快感，只是我们大脑在见证了规范的不可妥协性被肯定之后，产生的一个正向心理反馈——它从属于、并必须从属于那个更高阶的社会规范宣示功能。

这个反驳是严肃的，因为它直指本文论证中一个可能被读出的逻辑跳跃：将“过程中产生了一个副产品”与“这台机器以生产这个副产品为核心目的”混为一谈。它要求我们区分一台机器的副产物与它的存在理由。

本文的回应是：我们判断一台制度机器“实际上在做什么”的依据，不是它为自己宣称的存在理由，而是它在面临可观察的压力时实际做出的决策序列。如果惩罚的主产品是社会规范的不可妥协性，那么当大规模监禁被经验研究反复证明对降低犯罪率仅有微弱效果、对彰显规范力量也已因过度使用而贬值的时刻，这台机器应当能够轻松地、理性地削减监禁规模，而不会面临内部任何身份认同的抵抗。因为“不可妥协性”本身不绑定于任何具体的刑罚形式或惩罚量——它完全可以被转移去支撑一套更温和、更有效、更经济的规范宣示体系。然而，我们在过去二十年的美国刑事司法改革中所实际观察到的，却是一个恰恰相反的图景。每一次削减监禁规模的尝试——从加州的三振出局法改革，到联邦量刑指南的降格——都遭遇了来自检察官协会、警察工会、部分被害人团体和大量政治家的激烈抵抗。他们的抵抗话语，极少采用“这会导致犯罪率飙升”的经验论证（因为数据不支持），而是几乎千篇一律地诉诸同一个强烈的道德直觉：“这对受害者和公众的正义感是一种背叛”。

正义感——不是在哲学讨论中被精确定义的那个“正义”，而是在公众大脑中实实在在被体验到的那股热流，那种在见到“恶人终于得到了报应”时从背侧纹状体和内侧前额叶皮层涌出的满足——正是抵挡改革的那道最坚固的堤坝。如果快感只是一份无足轻重的副产品，它就不可能构成如此稳定、如此跨政治光谱的抵抗核心。它之所以能构成堤坝，是因为它从来就不是副产品——它是这台机器用来润滑其全部政治运转的最核心的燃料。规范不可妥协性的宣示，是这台机器公开展示的商标；但对那份古老神经奖赏的制度化管理和稳定供给，才是维持这台机器持续运转和抵抗变革压力的地下发动机。减少惩罚，意味着减少“正义感”这种公共产品的供应量。这才是改革阻力的真正的动力源头。而这正是本文的判断：翻译机的主产品，正是它自己宣称的那个“副产品”。

最后，本文必须交出它承诺的那张表——核心判断的证伪条件。

本文的核心判断是：现代刑事司法系统内部运行着一套将惩罚冲动翻译为道德确信的、具有功能自主性的制度化装置；该装置的主产品是翻译后的“正义感”，而非其宣称的“规范维护”；该装置的运作变形是当代刑事司法系统诸多自我消耗现象（而不仅仅是外部批评）的深层动因。

这一判断可被下列条件下的观察所证伪：第一，如果能够证明，当刑事司法系统面临财政压力时，其削减惩罚规模的主要阻力来源并非公众对“正义感”损失的道德抗议，而是独立的、不依赖于公众情绪的、仅出于规范完整性的制度性惰性——也就是说，如果能够观察到，警察工会、检察官协会和被害人团体的游说活动中，不出现诉诸“正义感”的道德话语，而仅以技术性和程序性理由为依据。第二，如果能够证明，恢复性司法在降低再犯率和提高受害者满意度方面的优越性一旦被公众充分知晓后，公众主动选择恢复性司法而非传统惩

罚的比例出现大幅、持续上升——这就反驳了“恢复性司法被边缘化是因为它阻断了正义快感供给”的子假说。第三，也是最严格的一条：如果能够证明，在完全没有“受害者”可被指认的纯粹“无被害人犯罪”（如某些毒品犯罪或自愿性交易）的执法中，公众对那些被判刑者所承受的痛苦毫无任何道德满足感的神经或行为迹象，并且执法机构在这些案件中没有援引“向社会输送正义”的道德修辞来为其耗费的公共资源进行辩护。

如果上述三项中的任意一项被充分的经验证据所证实，本文的核心判断就需要被重大的修正乃至放弃。如果三项同时被证伪，本文的核心判断即为错误。

但在此之前——在这台机器继续运转、继续将我们的古老冲动翻译为当下确信的時刻——本文所提供的，不是一份如何修好它的操作手册，而是一面被架在机器旁边的镜子。这面镜子不求全面，但求不被任何翻译的工序再次擦干净。它要留在那里，映出我们每一个在看新闻时、在读判决书时、在为某个判决拍手称快时，不愿对自己承认的那件事：我们审判，不只是因为正义必须被完成。我们审判，也因为那份在正义之名下尝到的快感，实在太难戒断。

这面镜子最终要问的，是一个比“正义应如何被定义”更难的问题。当惩罚不再能够被翻译为正义，当我们作为旁观的消费者不再能够从他人的痛苦中获得那份干净的、道德的确信时，我们还能不能、还愿不愿意，去做那些伤害发生之后必须有人去做的事？修复、安抚、承认、陪伴、约束——这些不会让大脑分泌奖赏的、沉闷的、耗竭的劳动——我们将它们交给谁？如果我们不再能躲在翻译机的无菌车间里，如果我们必须亲手去触碰那些被伤害玷污了的生活，我们是谁？我们能与那些我们曾称之为“恶人”、而现在必须改称为“伤害了我们的人”的人，共同在一个圆圈里坐多久，而不起身离席、逃回那个给我们准备了最舒服的旁听席的法庭？

拔掉罐子之后，我们是谁？本文不回答这个问题。它只负责把那罐子带到光下，让人看见它一直在被我们喝得津津有味的，到底是什么。

## 参考文献

Bourdieu, P. (1980). *Le sens pratique*. Paris: Les Éditions de Minuit.

de Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–1258.

Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Paris: Gallimard.

Freud, S. (1905). *Drei Abhandlungen zur Sexualtheorie*. Leipzig und Wien: Franz Deuticke.

Girard, R. (1972). *La violence et le sacré*. Paris: Grasset.

Roper v. Simmons, 543 U.S. 551 (2005).

United States v. Booker, 543 U.S. 220 (2005).

## 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：与惩罚社会学传统的硬边界

本文最近的三个对手全未出场：涂尔干论惩罚作为集体情感的表达与团结的重申，尼采在《道德的谱系》中论惩罚的快感谱系与债务隐喻，加兰论晚期现代的控制文化如何重新配置惩罚的表达功能。不打这三家，“惩罚翻译机”极易被读成尼采命题加涂尔干功能论的复述。分野在机制类别：涂尔干的惩罚表达既有的集体情感，尼采的惩罚释放既有的残忍冲动——两者的情感都先于制度存在，制度只是表达或释放它；本文主张的是翻译，即制度把一种不可宣之于口的输入改写为一种可宣之于口的输出，而这个改写过程会消耗制度自身。可判别面：表达论与释放论预测惩罚强度随集体情感强度变化；翻译论预测的是另一件事——当输入与输出之间的落差扩大时，制度的自我损耗（程序膨胀、教义复杂化、正当化文本长度）应上升，而惩罚强度可以不变。后者是本文独有的预测。

### 增补二：翻译机的可编码形式

翻译若不可测，本文就只是一个有力的隐喻。三项指标：正当化密度——单位判决中用于说明惩罚理由的文本长度，及其历时变化；落差指标——公众对该类案件的惩罚意愿与实际判决之间的差值；损耗指标——该类案件的程序步骤数与平均审理周期。翻译论的预测是三者同向上升，且第三项滞后于前两项。三项均可从公开司法统计与文书取值。

### 增补三：恢复性司法作为翻译终止的限界

正文把恢复性司法处理为翻译的终止——不再把惩罚快感翻译成正义，而是让当事人直接相对。这一判断须加限界：恢复性司法同样可以被制度化为一套新的翻译程序（会议脚本、调解员培训、成效指标），届时它终止的只是旧翻译，同时开启了新翻译。因此本文不能推出恢复性司法必然是出路，只能推出一条更窄的判据——若某种实践一旦被赋予可考核的成效指标就失去其效力，那说明它原本承担的正是不可被翻译的功能。这条判据同时适用于本文自己所推荐的一切方向。

### 增补四：与作者已发表工作的分野

作者已发表《司法缄默：法律系统道德感知的禁言机制》与《法度的负片》，两篇处理法律系统对某类内容的排除与不言。本文处理的不是排除而是转换——那个不可宣之于口的东西并未被禁言，它被改写成了另一种话语并继续运转。分野可一句话说清：缄默是不说，翻译是换个说法说。两者可以叠加，且叠加处正是本文最有价值的推论——一个系统可以一边对某种输入保持缄默，一边靠翻译它来维持运转，而这正是它耗尽自身的方式。

### 编辑增补所依据的核验文献

Émile Durkheim, *De la division du travail social*, Félix Alcan, 1893.

Friedrich Nietzsche, *Zur Genealogie der Moral*, C. G. Naumann, 1887.

David Garland, *Punishment and Modern Society*, University of Chicago Press, 1990.

Michel Foucault, *Surveiller et punir: Naissance de la prison*, Gallimard, 1975.

Nils Christie, “Conflicts as Property,” *The British Journal of Criminology*, 17(1), 1977.