

证身与未经见证的选择

一个不可单方面归零的承诺结构，如何在无人见证的让步中逐步凝固

高鹏 著 · SDE 学员 · 刑法理论与承诺结构 · 2026年7月26日

摘要 刑事归责体系将“自我决定”视为定罪的前提，却从未追问这个前提本身是如何在一个人身上生成的。本文的核心判断是：真正的自主选择，不是个体在私密意识中完成的内部事件，而是一种必须在被他人见证的条件下才能完整发生的社会-认知合成物。当一个人在做出重大行为之前，将其理由、方式和失败预案以语言形式提交给一个真实的他者，并被该他者追问、检验和记住时，选择才从私密的冲动转化为撤回成本不可单方面归零的承诺——这一过程，本文称之为“证身”。证身所生成的，不是一种可被个体独立携带的认知技能或惯习，而是一种依附于特定见证关系的、不能被自我欺骗攻破的承诺结构。未经见证的选择，在生成层面的意义上不是完整的选择：行动者的冲动未经“被另一个人追问并记住”这道工序，就直接滑入了行为输出。本文通过建构这两个互相界定的概念工具，论证了刑事归责的根本困境不在于“这个人有没有抽象的选择能力”——大多数犯罪者具备这种能力——而在于社会是否在他最需要被见证的成长期，为他提供了让这种能力从潜在形式转化为可实际操作形式的证身关系。在与内化理论、布迪厄的惯习理论、法兰克福的二阶欲望等最接近的理论对手逐一正面区分之后，本文证明“证身”切开了一个既有理论装不下的辨别面：它生成的不是可被个体原子化携带的心理资产，而是一种撤回成本不可由主体单方面归零的关系性承诺。在此基础上，本文提出“生成性归责”框架，要求在量刑中将被告成长过程中“证身关系的剥夺”程度纳入责任程度的裁量，并对这一框架的司法适用路径、证据标准与可能面临的滥用的具体操作难题做出了初步建构。全文的核心主张是：判决的正当性，不应只看行为那一刻的门槛式能力，而应追溯这种能力是否曾被允许真正发生。这一判断的证伪条件在于：如果有严格设计的大规模追踪数据表明，那种包含持续追问和容错回路的成人见证关系对高风险青少年的犯罪率没有显著的独立降低效应，则本文的全部理论建构将被推翻。

关键词：证身；未经见证的选择；自我决定；犯罪归责；生成性归责；生成机制

——论刑事归责中“自我决定”的生成机制条件

引言：法庭上那扇从未被建造的門

十八岁的阿远站在刑事法庭的被告席上。案卷记载的事实很清楚：他在夜市与人发生口角，对方骂了他一句很难听的话，旁边的人笑出声来，他从自己摊位下面抽出那根平时用来支雨棚的钢管，对着对方后脑勺一棍子砸下去。目击证人说，他砸完之后站在原地愣了两三秒，脸上没有表情，像是一个人突然发现自己已经跨过了一道门，正在回头找门在哪里。

检察官说：你本可以不这么做。你的手没有被绑住。你当时还有其他选择。法官最终宣告他有罪。法槌落下。[1]

这个宣告能够成立的逻辑前提，是每一间刑事法庭里都被反复言说却很少被追问的那句话：一个人本可以不这么做。法律不需要他当时有多么清醒、多么深思熟虑，它只需要确认那条通往“不做”的道路在那一刻没有被彻底堵死。

然而，如果我们把视线从法庭那扇厚重的木门移开，回到阿远走进被告席之前的十八年生命历程，我们开始看到一件令人不安的事。五岁，母亲上夜班时把他反锁在出租屋里——不是心狠，是怕他半夜自己跑出去，而楼道灯是坏的。八岁，他在班里从不说活，老师在家访时说“这孩子特别乖”——没有人觉得一个八岁从不说活的孩子有什么不对，因为“安静”在这个评价系统里不意味着退缩，它意味着安全。十一岁，他被高年级学生拦住要钱，第一次偷了母亲的钱，被母亲用衣架抽了一顿，从头到尾没有一个人问过他“你为什么要偷”——他从这次挨打里学到的东西，不是偷钱不对，而是下次要拿得更隐蔽。十四岁，团伙里一个年纪稍大的成员递给他一把折叠刀，问他要不要跟去堵一个人。

他在接刀之前，有没有过一秒钟的犹豫？当然有。但那一秒钟的思想——如果那能被称作思想的话——不是“我该如何应对这个风险”的审慎计算，而是“我如果表现出怕，他们以后就不带我玩了”的恐惧。那一秒钟里，没有一个他者——没有一个会追问“你确定要这么做？万一对方比你狠，你从哪里退”的人——在他和那把刀之间撑开过一道审视的缝隙。

我们现在回到法庭上那个法槌落下的瞬间。法庭说：阿远，你本可以选择不打。但当我们把他十八年生命里所有这些从未发生过的事——那些从未出现的追问、从未被给予的迟疑空间、从未被打开过的容错回路——一件一件摊开在法庭的大理石台面上，那个“本可以”的“本”字，究竟是安放在哪一块地基上的？

这里发生的不只是贫穷，不只是教育缺失，不只是社区衰败。这些都被反复讨论过了。本文要追问的是一个更深、更隐蔽、更少被触及的问题：选择这件事本身，需要什么样的社会条件才能完整地发生？

刑事归责体系对这个问题的沉默，不是偶然的疏忽。它根植于归责体系的一个深层预设：选择能力被当作一种门槛——一个人一旦跨过了最低限度的精神能力门槛（能辨认为行为性质、能控制行为），他的一切行为就都被推定为“自我决定”的产物，他对这些行为负有完全责任。至于这个能力是如何在他身上生成的，他曾经被置于什么样的条件之下、那些条件是否允许他把这种能力从一纸抽象潜能转化为

一套实际可操作的内在装置——这些问题，归责体系既不追问，也没有为它们预留概念空间。

这不是归责体系的“缝隙”——用这个词，等于把这个体系想象成一个原本完整、后来破损的东西。归责体系没有破损。它被完好地建成了它被设计成的那个样子：一个只问能力门槛、不问能力生成的装置。它是在一个特定的智识传统中被发生出来的——那个传统假设个体是原子化的理性载体，其选择能力是某种自然禀赋，只要未被明确的精神疾病所剥夺，就始终在线。本文要做的，不是诊断这个体系的病灶，而是追问：当这个体系把“自我决定”默认为一种门槛式能力时，它系统性地排除了什么？

这扇被排除在外的门，并不是没有人伸手推过。法哲学内部已经积累了一系列朝向它的推力。哈特（Hart，1968）在追问一个人“本可以不这么做”需要什么条件时，不仅要求他具备一般的行动能力，还要求他拥有“运用这些能力的公平机会”。达夫（Duff，2001）的应答性责任理论则进一步论证，将一个人作为道德主体召唤到法律面前，这项实践本身就预设了此人已被纳入一个“能够理解规范、并能够按规范引导自身行为”的道德共同体之中。神经法学对自由意志的讨论——以格林与科恩（Greene & Cohen，2004）的论文为代表——从另一个方向推进了这一追问：如果人类行为最终可以被还原为神经机制，那么“自由意志”本身就是一个幻觉。在司法实践层面，澳大利亚Bugmy案（2013）和加拿大Ipeelee案（2012）已经将“社会不利背景”纳入量刑裁量，承认极端的社会剥夺可以降低被告的道德可责性。在犯罪学领域，劳布和桑普森（Laub & Sampson，2003）的年龄分级理论论证了社会纽带对犯罪轨迹的关键塑造作用，莫菲特（Moffitt，1993）的发展分类学则将犯罪者区分为“终身持续型”与“青春期限定型”，前者往往植根于早年的认知与自控能力缺陷。

然而，这些进路各自留下了一道他们自己未能命名的缺口。哈特提出了“公平机会”，却没有追问这一机会在个体生命史中具体需要哪些社会互动条件才能兑现——“机会”如果只意味着外部障碍的消除，那它仍然只是一道敞开的门，却没有回答为什么有些人走到了门前却迈不出去。达夫洞察到道德共同体是责任归因的构成性前提，但他没有具体展开：被纳入这个共同体，在认知发生的意义上需要经历什么样的微观互动？阿远在名义上从未被逐出这个共同体——他上过学、有过户籍、甚至上过道法课——但那些从未有人问过他“你刚才为什么要那么做”的时刻，那些他刚要开口就被一句“你还敢顶嘴”堵回去的时刻，那些他在心里翻来覆去想了一整夜、第二天却发现没有任何人记得他昨天差点出事的时刻，合在一起构成的，难道不是在认知发生的意义上从未被真正纳入过这个共同体吗？如果“纳入”只意味着法律身份的赋予，而不包括那些让规范真正被理解的微观互动，那么这个概念本身就需要被重新审视。

神经法学的还原论进路从一开始就悬置了一个它无法回答的问题：即便自由意志在形而上学层面是幻觉，为什么有些人在行为那一刻能够暂停下来审视，而另一些人则不能？司法实践中的“社会不利背景”裁量，则停留于将剥夺当作外部环境压力来对待，未能从认知机制上阐明这些压力具体如何削弱了一个人的审视能力——这种能力不是被压力直接“压垮”的，而是因为成长环境中缺乏某种特定类型的互动而从未被完整地建构出来。发展犯罪学的范式在实证层面揭示了早年风险因素与犯罪轨迹的强关联，但它将“成人支持关系”当作一个一般性的保护因子来处理，没有从理论上辨析这种关系中的哪一个特定维度——追问？拉住？容错？——才是真正承担生成性功能要素。

本文要做的工作，不是在这些进路之外另立山头，而是进入它们之间的那片未开垦地带，补上他们触及了却未能命名的那一笔：“能够选择”与“真正地选择过”之间的那道缝隙，只有在一种特定的社会关系之中才能被弥合。哈特的“公平机会”、达夫的“道德共同体”、发展犯罪学所说的“成人支持关系”，都可以被视为朝这个方向探出的触角——但它们都缺少一层追问：在这些机会、共同体和关系里，究竟是哪一个特定的操作，在认知发生的意义上，把一种抽象的选择潜能转化为了一次真实的、撤回成本不为零的选择？

本文把这个操作称作“证身”。这个概念的提出，不是为了在学术市场里再增加一个新词，而是为了给一道无人命名的缝隙——那扇在太多人的童年里从未被建造的门——一个可以被讨论的名字。它将在下文中，在与它的理论近邻们逐一正面碰撞的过程中，一层一层地展露出它的全部结构。

一、从“内化”到“证身”：一次不可还原的切割

本章的任务不是给“证身”下定义，而是通过它与它最危险的近邻——内化理论——的正面碰撞，逼出它不得不单独存在的那个剩余物。如果证身说的一切都能被“外部指令内化为内部声音”这个经典图景完全涵摄，那么它就没有任何理由作为一个独立概念存在。

1.1 内化理论的经典图景及其盲区

维果茨基（Vygotsky，1978）的内化理论、米德（Mead，1934）的“一般化他人”以及精神分析的“超我”思想（Freud，1923），共享一个基本图景：儿童在外部互动中接收来自他者的指令、评价和期待，这些外部的声音逐渐被“内化”为一套内部的自我监控系统。按照这套图景，一个被舅舅追问过“你想过万一打输了怎么办”的孩子，在日后面对相似的冲突情境时，舅舅的声音会在他的意识中响起，形成一道内部的制动。这正是从维果茨基到当代社会认知理论一脉相承的叙事：社会互动被“内化”为心理工具，心理工具又在关键时刻被“调用”为行为调控。

我必须承认，当我在本文的早期构思阶段描述证身时——他者追问、被记住、在关键时刻在脑中响起——我自己听着都觉得这几乎就是内化的一个更华丽的修辞版本。如果证身就是内化，那本文的全部工作就是在给一个上世纪就已经被研究透了的老概念换上一件新外

套。这是本文面临的最严肃的学术挑战。

但是，内化理论有一个它自己从未正面处理的盲区，而阿远这样的案例恰恰把这个盲区照得雪亮。这个盲区是：内化的内容是可以被自我欺骗所篡改的，而且这种篡改恰恰发生在它最需要被调用的那个时刻。

让我把这一点讲清楚。假设阿远在十三岁那年，学校里有一位他很尊敬的老师在课后单独对他说过一句话：“遇事先想三秒，实在气不过，去找你信得过的人说。”假设他听进去了——在那一刻，他真的觉得这句话有道理。这句话被内化了吗？按照任何版本的内化理论，它都被内化了。老师的外部声音，进入了他的内部自我监控系统。

现在来到十四岁，那把折叠刀递到他面前的时刻。那一秒钟里，老师的那句话能被调用出来吗？答案很可能是不能。不是因为那句话没有被存储——道法课上被反复灌输的“暴力不能解决问题”他也都记得——而是因为，在那一秒钟里，调取这些内容意味着要面对一个他整个生命都在逃避的认知痛苦：承认自己在害怕，承认自己其实并不想去，承认自己接刀只是因为怕被同伴抛弃。这种自我承认，恰恰是内化内容得以被激活的前提。而这个前提，在一个从未有人当着他的面说过“你怕也没关系”的成长环境里，恰恰是他最不具备的东西。

更致命的是内化的第二重不可靠性：它不是被动消解，而是可以被主动篡改。一个人在自己内心对自己说过的话，可以在事后被重新解释、被遗忘、被合理化、被歪曲成方便自己的版本。那个被内化了的“遇事先想三秒”，在十四岁的阿远那里完全可能被改装成“只要想清楚对方有多欠揍，动手就是对的”。他自己甚至意识不到这种改装已经发生了——因为大脑在认知失调面前，第一反应不是面对真相，而是降低失调感。你无法让一个内化了的指令去纠正它自己的被篡改；那等于让一个被贿赂了的警察去查他自己的案子。

现在，让我们对比另一种情况。假设十三岁那年，老师不只是说了一句话。假设在那次课后谈话中，老师问了阿远一个问题：“如果下次有人激你，你最怕自己做出什么？”阿远想了很久，说了一个答案。老师没有评价这个答案，只是说：“好，我记住了。下次我再见到你，我会问你，你有没有做到这件事。”然后阿远知道——不是以为，是确凿地知道——老师的脑子里记着一个他给出的承诺。一年后，那把刀递过来的时候，他脑子里响起的不是一句可以被篡改的话，而是一个外在于他、不受他控制、他无法单方面改写的活体记忆：老师记得。老师下次见到他的时候，会问。

这就是证身与内化之间那道不可逾越的鸿沟。内化可以被自我欺骗所侵蚀，但证身在关系现场制造的撤回成本，不能由主体自己单方面归零。你可以在自己脑子里篡改“我当初说了什么”。你可以在漫长的自我合理化中把当初的承诺歪曲到你需要的形状。但你无法走进老师的脑子，删除他记住的那个版本。你无法在那个时刻对自己说“老师肯定会理解”——因为老师当初什么都没说，他只是记住了。而你知道他记住了。这个“他知道我知道”的结构本身，就在冲动和行动之间制造了一道独立的制动间隙——不是靠提高理性，而是靠改变认知架构：把选择从一个私密的、可以被自我欺骗封闭的回路里，硬拽到一个不可撤回的关系事件中。

1.2 证身的本质：一个不可单方面归零的承诺结构

现在，我可以更精确地说出证身与内化之间那个不可还原的“剩余物”是什么了。内化处理的是指令从外部向内部的迁移，它的输出是一个被个体存储在记忆中的心理内容。证身处理的是一件完全不同的事：在冲动与行动之间的那道缝隙里，它植入的不是一个心理内容，而是一个关系的认知事实——你确凿地知道，有另一个人记得你的选择，并且你对他的那个记忆版本没有编辑权限。

这个认知事实一旦被建立，它就对自我欺骗构成了一道结构性的屏障。这不是因为那个他者会来惩罚你——他可能永远不会主动来追你的承诺。而是因为，只要你还在乎他对你的看法（而一个曾经认真追问过你、并且记住了你的回答的人，恰恰是在表示他对你的在乎），你就无法把这件事当成没发生过。你可以说服自己“我当时没想清楚”，但你知道他脑子里存着的那个版本不是这样的。你可以说服自己“情况变了”，但你知道他可能会问“哪里变了”。你可以选择再也不见他，但只要你还可能见到他，那道缝隙就还在。这道缝隙不是来自恐惧，不是来自道德教育，不是来自理性计算。它来自一个最简单的事实：撤回的成本不是零，因为另一个人替你记住了你当初说的话。

从认知机制的层面看，证身与当代认知科学中关于“自我距离感”（self-distancing）的研究有一处可以沟通的界面。这一系列实验范式表明，当一个人被要求以旁观者视角审视自身处境时，其判断质量会显著提升——因为虚构的第三人称视角强制他将事件从体验流中抽离，赋予其对象性。但这个范式的局限正在于：那个“虚构的第三人称”是安全的距离，是可以被自我欺骗轻易绕过的。被试可以在想象中采纳“旁观者”的视角，但他完全可以按照对自己有利的方式来想象那个旁观者会怎么看。他甚至可以想象出一个会原谅他一切行为的旁观者。而证身所引入的那个他者，不是一个可以被自我随意采纳或抛弃的虚构视角，而是一个我有实际利害关系在其面前保持可信度的、活生生的人——尤其是，一个我不能替他决定他会怎么看我的人。

这个“不能替他决定他怎么看我”的结构，就是证身比内化多出来的那块硬骨。内化的内容可以由主体单方面编辑；但证身所制造的那个“他知道我知道”的回路，一旦闭合，就不能由任何一方单方面断开。那个他者可以忘记——但那不是主体能控制的；那个他者可以死去——但那也不是主体能控制的。在忘记或死去之前，承诺是锁定的。这正是证身与内化在工程上最本质的差异：内化是一个可以被自我欺骗攻破的缓存，而证身是一道不可单方面撤除的关系性封条。

1.3 证身与法兰克福二阶欲望的区分

在进一步展开证身的生成机制之前，还有一位近邻需要在此一并切割——它在外观上与证身非常相似，都涉及对一阶冲动的超越。

法兰克福（Frankfurt, 1971）关于“二阶欲望”的经典论述刻画了这样一个过程：一个人退回到自己的内心，对自己此刻的一阶欲望形成一个评判性态度——例如，一个想吸毒的人，同时希望自己不再有想吸毒的欲望。这个二阶动作即使从未被任何人知晓，它在哲学上也可以成立。由此，一个人的自由不在于他能实现自己的一阶欲望，而在于他能对自己的一阶欲望形成意志的认同或不认同。

这个框架在哲学上自治，但当一个法哲学家把它拿到刑事法庭上去解释为什么有些人在被激怒时“能够”停下来时，它就撞上了一堵它自己看不见的墙。一个从未在与他人的对话中被追问过“你刚才那股冲动的理由到底是什么”的人，他是否可能自动启动法兰克福所描述的那种二阶审视？答案很可能是不能——而且不是因为他的“意志”弱，而是因为这个二阶审视的启动开关，根本就不在他自己的内部。

法兰克福的“二阶欲望”是一种可以被个体在孤独中完成的内心动作。但本文所要论证的恰好相反：对自身冲动形成距离化的审视，这种能力本身，是在一种特定类型的关系中被反复训练出来的，而且它的启动，在很长一段时间内，依赖那个关系中的他者的稳定在场。没有人追问“你为什么那么想冲上去”，一个人就不可能自动学会在自己的冲动面前停下来问自己这个问题。这不是意志力的缺失，这是审视习惯的生成失败。证身与法兰克福的二阶欲望，看似都涉及对一阶冲动的超越，实则分别指涉两个完全不同的辨别维度：一个是内在反思结构的层级——你在几层上反思自己；一个是关系性见证的生成——你的那几层结构，有没有被另一个人的声音撑开过。

二、证身的生成机制与操作结构

经过第一章的对质，证身作为一个不可被内化还原的独立概念，它的轮廓已经露了出来。但它内部到底是什么结构？它是怎么运行的？本章将从正面展开它的操作机制，并在这个过程中证明它同样不能还原为布迪厄的“惯习”。

2.1 证身的三步操作

证身在操作上包含三个不可再拆的步骤。一个完整的证身事件，三步缺一不可。

第一步：提交。行动者在做出一个即将产生重要后果的行动之前，将该行动的理由、方式和失败预案——以语言的形式——提交给一个真实的他者。这不是在心里默念，不是在微信上发一条可以撤回的消息，而是用口头语言、面对面地说出来。为什么必须是口头语言？因为口头语言不可撤回。说出来的话，收不回去。对方听到了，就记住了。而你清楚地知道对方听到了。这个“你清楚的知道”就是第一步制造的那个认知事实：从此刻起，你的选择不再是一件只有你知道的私事。

这一步的发生需要两个前提条件。第一个条件是，这个他者必须是一个你不愿意在他面前掉链子的人——不是怕他，而是在意他对你的评价。一个让你害怕的人，你只会对他说他认为你想听的话；一个你完全不在乎的人，你根本不会真正地把选择提交给他。第二个条件是，这个他者必须是一个你不会因此被他伤害的人——他不会拿你提交的脆弱去羞辱你、威胁你、或者在日后翻旧账来攻击你。这两个条件合在一起，描述的是一种极其特定的人际距离：近到足以让你在意他的评价，又远到足以让你不被恐惧所胁迫。在阿远的整个童年里，他遇到的所有成年人对他的关系——母亲、老师、高年级学生——无一例外的，要么是近到只有惩罚而没有追问，要么是远到只有评价而没有记住。

第二步：被追问。他者在接收行动者的提交之后，并不简单地点头或批判。他追问。追问的内容不是“你怎么能这么想”（那是道德审判），而是“你想过最坏的情况是什么吗？”“万一你判断错了，你最怕发生什么？”“如果那个情况真的出现了，你第一个求救的人是谁？”这些追问的本质，是把行动者的视角从他此刻沉浸的行为中拽出来，强制他看到自己行动可能带来的后果全景，尤其是那些他因为冲动而自动忽略的后果。

这一步在认知上的操作，与单纯的“告诫”有根本区别。告诫是单向的——我告诉你你会犯什么错。追问是双向的——我逼你自己说出来。自己说出来的内容和别人告诉你的内容，在大脑里的编码深度完全不同。前者是你主动建构的认知产物，它留下了你自己的判断痕迹；后者是你被动接收的外部信息，它随时可以被归类为“他知道什么，他又不懂我的情况”。证身的追问，不是在给他增加信息，而是在逼他进行一种他自发不会进行的认知动作：从体验主体临时切换为后果评估者。

第三步：被记住。他者在追问结束之后，不做评判，不说“我相信你能做到”，不说“你太让我失望了”。他只说一句最朴素的话：“我记住了。”然后他真的记住。下次见面的时候，他可能会问、也可能不会问，但行动者知道——他脑子里的那个版本还在。

这第三步是证身全部效力的最终落脚点。如果没有被记住，提交和追问就只是一次性的对话实验，事后可以被当作“当时只是随口说说”而被无害化处理。只有加上“被记住”，整个事件才获得了一个不可随时间衰减的结构：它从一个“当时有过的一次谈话”，变成了一段“仍在进行中的、每次见面都可能被调阅的活体承诺”。撤回成本的正是不为零，正是从这一步开始。它不是被他者惩罚的恐惧——

那是外部约束，不是生成机制的生成——而是“他知道我的选择，而我不能替他删除”这个纯粹认知事实的持续存在。

2.2 证身为什么不能还原为惯习

以上三步操作，如果被反复执行多次（不是一次，是几十次、上百次，在几年时间里），会不会只是在一个人身上培育出了一种“遇事先说理由”的审慎性惯习？如果是这样，那么证身不过是布迪厄（Bourdieu, 1977）意义上的惯习理论的一个特例——在一个特定场域中，通过反复的身体与语言规训，塑造出一个特定的认知倾向性。

这个质疑必须被正面回应，因为它直接指涉了证身概念最根本的那个主张：证身生成的，不是可被个体独立携带的惯习，而是一种关系资产。两者之间的差异体现在三个关键维度。

第一个维度：可迁移性。惯习的一个核心特征是，它一旦形成，就可以被主体跨情境地携带和调用。一个在严格学校纪律中被养成的学生对规则的尊重惯习，会在家庭生活、职业场域乃至消费行为中一再复制自身。但证身所培育的那种在冲动面前停下来审视的习惯，不遵循这种迁移逻辑。它的启动，从根本上依赖于那个特定的他者——至少是被那个他者记住的那种背景性的在场感——的持续存在。当那个他者退出、死亡或不复记忆时，那种“我选之前必须把理由说清楚”的内驱力可能迅速衰减。这不是惯习的消解（惯习不会因为单个他人离场而消解），而是证身能力特有的结构退化：证身从来不是一种可以独立于见证关系而持有的个体技能，它更像一根必须在某个人那里存有底本的活态纽带。那根纽带一旦断了，单纯的惯习残余——那些曾经被追问过的行为模式——尚且可能保留，但那种“撤回成本不为零”的承诺感，会当场失效。一个被严格教导过的人，即使离开了学校，他的“守时”惯习仍然运作；一个被长期证身的人，一旦那个他者从他的生命中消失，他的“被见证感”会随之蒸发。这是两个完全不同种类的心理构造。

第二个维度：生成的机制。惯习的生成方式主要是被动的、前语言的身体规训——它不是通过语言追问来运作的，而是通过身体在特定条件下的反复暴露来沉淀的。布迪厄笔下那些阿尔及利亚农民的时间惯习，或者法国中产阶级的文化消费惯习，都不是被人用语言追问出来的，而是被环境中的物质条件和他人行为的默默期待刻进身体里的。证身则完全发生在语言互动层面：它要求主体将选择以命题的形式陈述出来——理由、方式、失败预案——接受追问，并在他者的记忆中存档。这不是身体规训，这是承诺制造。而承诺制造和惯习沉淀，在工程上是两套工序：前者的产物依赖于特定他者的记忆和在场；后者的产物可以被个体原子化地携带。一个可以用身体记住的东西（惯习），一个必须在关系中才能维持的东西（证身）——它们根本不是同一种存续方式。

第三个维度：规范性的方向。惯习可以朝向任何方向生长——它既能朝向审慎栽培，也能朝向暴力栽培。布迪厄笔下那些在街头被反复规训的年轻帮派成员，同样形成了某种“遇事先动手”的惯习——这个惯习同样是被一个特定场域中的反复互动刻进他们身体的。惯习理论本身不区分这些方向：它不追问一个惯习是不是“好的”，它只解释它为什么能够在那里生成。但证身这个概念，内在地包含了一个规范性的方向：一个真正的证身关系，必须包含“对方可以拒绝这个选择、而不会因此被关系毁灭”的容错空间。也就是说，当一个少年在证身关系中说“我不想打”——那个追问他的成人，不能因为他给出了一个“不够硬”的回答而收回对他的认可。如果他没有这个容错空间，那么他所谓的“提交选择”其实只是在猜对方想听什么答案——那不是证身，那是表演。这是证身与惯习在规范性上的根本区别：惯习不必包含“退出权”，证身必须包含。

因此，布迪厄的惯习理论，完全可以描述一个在帮派中被反复规训的少年，如何在面对冲突时自动地“选择”了最暴力的回击方式——因为这个惯习已经在无数次身体实践和同伴期待中被刻进了他的行为模式。但在本文的生成机制框架中，他的“选择”恰恰是被伪装的未经见证的选择——他的行动理由从来没有被提交给一个要求他解释、并允许他拒绝而不毁掉他的关系的人。惯习理论能解释他为什么那样“自动”，但它无法区分这种自动与另一种自动——被证身关系按住过一秒的自动——之间在生成机制上的本质差异。正是这个差异，构成了证身在归责问题上的辨别力：不是你有没有惯习，而是你那个惯习在被情境激活之前，有没有被一双真实的耳朵按住过至少一次。[2]

2.3 证身的操作性定义

现在，在对证身做了两轮正面碰撞——与内化（它最危险的近邻）和惯习——之后，这个概念的内在结构已经被一层一层地剥了出来。作为对这整个发生过程的事后总结，我们现在可以给出它的操作性概括：

证身，是主体在做出一个将产生重要后果的行为之前，将该行为的理由、方式和“万一失败的第一个信号和退路”，以口头语言形式提交给一个真实的其他者，并被该他者追问、检验和记住的关系事件。这个过程使选择从私人的、可被自我欺骗地撤回的念头，转变为被见证的、撤回成本不可由主体单方面归零的承诺。

这个定义里的每一个限定词——口头、真实、追问、检验、记住——都不是修辞，而是对一道独立工序的命名。缺少任何一道工序，整个机制都会在关键时刻失效。口头不可撤回，真实的其他者提供不可篡改的记忆，追问强制认知切换，检验防止虚假承诺，记住使撤回成本随时间持续存在。它们合在一起，构成了证身作为一个关系事件的完整工程结构。

三、未经见证的选择：从让步到凝固的生成轨迹

如果证身是让选择完整发生的条件，那么它的缺失会制造出什么？这不是一个抽象的逻辑反推。它要求我回到阿远的那十八年，不是作为一种修辞，而是作为一个认知生成机制的案例，实实在在地展示：一种“不审视、不迟疑、被冲动直接推过门槛”的行为策略，是如何在特定的关系生态环境中，从一次次具体的矛盾处理中，被一步一步塑形、加固、最终凝固为稳定的反应模式的。

3.1 八岁：沉默作为安全策略的生成

八岁的阿远，在班里从不说话。老师在家访时说“特别乖”。我们需要暂时搁置后见之明——搁置我们对十四岁接刀的那个时刻的全部知识——回到八岁的那个教室，去看他每天面对的是什么。

小学课堂上，提问是一个危险的时刻。被叫起来的学生，要么答对（获得短暂的认可），要么答错（在全班面前丢脸），要么说不知道（老师皱眉，同学笑）。对于大多数孩子来说，这些风险是可以承受的——因为他们从前的经验告诉他们：答错了，老师最多说一句“坐下再想想”；说不知道，最多被说一句“上课有没有认真听”。这些惩罚在他们那里只是短暂的、低强度的，不值得用一个全盘回避策略来应对全程。

但在阿远的经验系统里，这些后果不是这样编码的。五岁时他被反锁在家里，母亲上夜班前没有留下一句话解释为什么要锁他——他是在被锁住之后才自己推理出“妈怕我跑出去”的。这件事的认知后果，不是他学会了“外面危险”，而是他学会了一件事：母亲对他的行为的全部反应，都是在他不知道理由的情况下强加于他的。他不会被告知“为什么”，他也不会被询问“你怎么想”。他唯一能做的，是在母亲可能做出反应之前，先把自己的行为压到最低可见。不被看见，就不会被反应。不反应，就不会引发反应。

当这个策略被迁移到课堂上，它的形态就是“从不说话”。从不在课堂主动举手——不是不懂，是举手意味着把自己暴露在老师的评价之下，而评价的后果是不可预知的。从不跟同学发生冲突——不是不生气，是从五岁起他的行为管理系统就不是“如何减少惩罚”，而是“如何不变成目标”。一个孩子在这个阶段学会了沉默，不是因为沉默有益，而是因为在那个关系生态里，任何将自己暴露在他人视线下的行为都意味着不可控的风险。“安静”，从老师的评价视角看是一次顺利的社会化；从阿远的认知生成轨迹看，是一次成功的适应——他找到了在这个环境里不被伤害的最低暴露策略。

3.2 十一岁：解释通道的闭合

偷了母亲的钱被母亲用衣架抽了一顿。从头到尾没有一个人问过他为什么要偷。这件事在日常语言里会被描述为“母亲没有试图理解他”。但本文要借助这个案例讲清楚一个更精确的机制：这不是“理解”的问题，这是在冲动与行为之间，那条可以被语言撑开的审视通道，从未被建造。

一个孩子偷钱，在行为主义的账本上是“通过惩罚来消除不良行为”。但在一个生成机制的账本上，这件事的后果远不止于此。当一个人在做了一个行为之后，被直接给予身体痛苦，而没有任何人问他“你当时是怎么想的”，他学到的东西不是“偷钱会痛所以不该偷”，而是一个更深层的元认知规则：我做出选择之后，唯一会发生的事情是身体后果，而不是我的理由需要被陈述。陈述理由这个动作本身，在他整个十一岁之前的人生里，从未被任何人要求过。他不知道可以这样做。他更不知道，当他学会陈述理由的时候，理由可以在冲动和行动之间撑开一道审视的缝隙。

这正是未经见证的选择在生成初期的关键形态。它不是“无法控制冲动”的病症，而是在一个特定的关系生态中，一种被反复强化、逐渐固化的认知习惯：选择是私人的事，它不需要被口头提交。当这个习惯在十一岁时已经凝固成形，十四岁时他面对那把刀，他的整个认知系统里都不存在“把这件事告诉某个人”这个操作选项。他不是选择不去求助。他是从未有过“求助”这个按钮。

3.3 十四岁：“不审视”作为高风险环境的精确适应

十四岁，团伙里那个年纪稍大的成员递过折叠刀，问他要不要跟去堵一个人。这一刻最容易犯的错误，是把这个情境解释为“同伴压力”，然后用社会心理学里现成的“从众”“服从权威”来打发改。但本文的要求是要打到更细。

在接刀的那一秒之前，阿远面临的是一个极其具体的情境判断：如果我表现出怕，他们会怎么反应？他脑子里跑过的不是一堆理论假设，而是一长串他已经亲身经历过的、或者亲眼见过的记忆片断——只要你在圈子面前露出一丝犹豫，他们以后就不再叫你；只要你不被认定是“不敢”，下次他们商量事的时候你就没份；而如果你连这些人也没份了，你就什么人都没了。他判断得没错。在那个生态里，表现恐惧的后果，就是被踢出唯一接纳过他的群体。他从小就知道被排除意味着什么——被锁在空房间里，灯是坏的。他不会再用任何方式回到那个状态。

所以他接刀。不审视——不是他缺乏审视的能力，而是从八岁起，审视与迟疑对他的身体来说就是危险的信号。他在长期的关系环境中经历了无数次的确认：每次他迟疑，就被推开；每次他表现示弱，就被攻击。他的整个行为管理系统，是由一个内置的最高优先级规则

驱动的：不要表现出你在犹豫。这个规则不是他选的——是那个生态把他选出来的。

现在把这十四年的轨迹连起来看。他五岁学会不暴露自己，八岁学会不公开说话，十一岁学会不解释自己的行为，十四岁学会不迟疑。每一层策略，都不是从天而降的恶习，而是对上一层环境矛盾的合理让步——每一次让步都降低了当下的风险，但也进一步巩固了那条路径：面对冲突时，不审视、不陈述、不求助。这不是一个“缺陷”的累积，这是一套策略在特定生态中反复被验证有效之后、自然沉淀成的稳定形态。当法学家事后说“他本可以不这么做”的时候，他们看到的不是这个人，他们看到的是一张定格在行为时刻的快照。这张快照里没有八岁、没有十一岁、没有那一条从不敢暴露到不敢迟疑的路径。它只有一个已经凝固了的、不可能在那一刻被单独打开的形态。从生成机制来看，未经见证的选择不是一次选择的失败，而是一条生成路径的终点。

这里需要回应一个预判性追问。批评者会问：即便阿远的认知策略是适应环境的产物，但他在法庭上仍然是一个具备刑事责任能力的人——他能够辨认行为的性质，他能够控制行为。你说的这一切，难道不是在为他的脱罪吗？这个问题触及了本文全部论证与刑事归责体系之间最根本的那个分界点。我不是在说阿远没有刑事责任能力——按照现行法的门槛式标准，他的确有。我是在说，这个标准本身不问的事情，恰恰是阿远这十八年里最核心的那个认知事实：他的“能够控制”，是一块被安装了但没有被接电的开关。法庭假定开关在盒子里就是功能完整的，但证身关系的缺失意味着从开关到行为的那段电线从未被铺设。这不是精神病的辩护，这是对“责任”本身的生成机制前提进行追问。他不是没有罪责，但他的罪责必须包含社会在他最需要被见证的成长期，系统性地没有为他铺设那段电线这一事实。具体如何在量刑中处理这一事实，将在本文第五部分展开。

四、选择如何被生成：一个正面的生成机制叙事

前三章的论证在逻辑上是完整的，但它只完成了一半的生成机制任务。生成机制不是哀悼缺失，不是指出一个东西没有生成——它最深的力道，在于展示一个新结构如何从矛盾的结算里长出来。如果本文的论证只停留在“未经见证的选择是一条被环境失败所凝固的路径”，那它就只有创伤的解剖，而没有关于愈合如何可能的知识。这一章，必须补上那另一半：证身是如何在一个同样身处高风险的少年的生命里，生成出一种在他的既往经验中本不可能产生的审视性制动装置的。

以下是阿杰的叙述。他不是某个真实个案的如实记录——他是基于我国司法社工、校园辅导实践和青少年犯罪研究中反复出现的有效干预模式而建构的理想类型。他的生命史中的每一个转折点，都对应于证身机制中一道具体的操作工序。

4.1 十四岁之前：阿杰与阿远的共同起点

阿杰出生的年份和阿远相同，他母亲的夜班时间表和阿远母亲的时间表相差无几。小学五年级那年，阿杰也是班里不说话的那一个，也在老师家访时得到过“这孩子特别乖”的评价。十二岁，他也被高年级学生拦住要钱，也偷过一次母亲的钱，也被发现、被揍了一顿——同样没有人问他“你为什么要偷”。

到十三岁，阿杰在社会关系上和阿远一样，已经站在那个临界点的边缘：他已经开始和一些比他自己年纪稍大的、有打架经验的孩子在一起。他的行为管理策略也和阿远一样，倾向于在冲突面前不暴露、不解释、不求助。如果没有意外，他会在十四岁的某一天，被递上一把刀，然后被问要不要跟去堵一个人。

4.2 篮球教练的第一问：提交的机会被创造出来

意外发生在他十三岁那年的暑假。社区里有一个公益篮球训练营，免费的。阿杰去的原因不是想打篮球，是去那里可以正当地不在家待着——家里太热，而且母亲那天不上夜班，会在家，在家意味着被盯着，被盯着意味着随时可能因为某件事被骂。他对篮球没有兴趣，但篮球场上有其他人，比一个人待在街上安全。

教练老周不是一个专业的心理咨询师。他只是一个打了二十年篮球、见过太多这种孩子的中年人。他的训练方法和其他教练没什么两样。但有一件事他做得不同：每次训练结束之后，他会留十分钟，让队员围成一圈，每个人说一件今天自己做过的、自己觉得做得不好的事。没有惩罚，没有公开批评，没有记名。说完，老周只会点一下头，然后下一个。

前五次，阿杰每次都说“我没有”。老周从不追问，只说“行”。第六次，阿杰说“我今天训练不认真”。老周点了一下头。这是阿杰生命里第一次，在一个成人面前，主动说出了一件自己做得不好的事。而且那个成人没有骂他，没有打他，也没有对他失望——他只是点了一下头，记住了。下个周末训练开始前，老周把阿杰叫到一边，说了一句话：“上次你说你训练不认真。这周有没有更认真一点？”阿杰愣了一下。他没想到有人真的记住了。这是他生命里第一次体验到：我把自己的事说给一个人听，他会记住。这不是内化的开端——因为老周什么道理都没讲。这是一段证身关系的第一次微小的闭合：提交→被记住→被调阅。这个回路闭合的那一刻，阿杰开始拥有了一个他从前从未有过的东西：一个他在意的成人的记忆里，存着他自己的口头提交的承诺。那个记忆不在他的脑子里，不受他的篡改，不由他的遗忘所控制。

4.3 冲突时刻：证身在冲动面前撑开的那半秒钟

一个月后，阿杰在场外的街道上被两个比他大的少年拦住，要他“借点钱”。阿杰的第一反应——身体反应，十三年来被反复训练出来的反应——是跑。但跑不掉，两个人在两头堵着他。第二反应是打。他已经摸到了地上的半块砖。

就在他手碰到那半块砖的同一瞬间，他的脑子里闪过一个画面：老周的脸。不是因为老周教过他“打架不对”——老周从来没见过这句话。而是因为上周的训练结束后，他曾经在那一圈的轮流发言里说过一句话：“我这周差点在外面跟人动手，我忍住了。”老周当时没说什么，但点了头。现在老周脑子里存着那句话。如果他今天动手了，下次训练，老周会不会问他：“你上次不是说忍住了吗？”他不知道自己到时候怎么回答。他甚至不确定老周是不是还记得。但他在那一刻发现，自己希望老周记得。而且更重要的，他无法自己走进老周的脑子删除那句“我忍住了”。他不能控制老周记住什么、忘记什么。他想起了老周的脸。那个他可能会在它面前掉链子的脸。那个他不会对它有所恐惧、但不想在它面前失信的脸。

就是这半秒钟。他扔了砖头，从侧面冲出一个缺口，跑了。

如果当时老周只是曾经对他说过一句“打架不对”，那句话在他摸到砖头的此刻早就被冲动的热浪淹没了。内化的内容是可以被冲动的热浪淹没的。但老周在他脑子里植入的不是一句话。老周植入的是一个关系的认知事实：有一个人，既不是来抓他、也不是来打他，但他恰好是阿杰在所有成年人里唯一不想在他面前掉链子的那个；这个人的脑子里存着一份阿杰自己口头提交的、关于“不动手”的记录；阿杰对这份记录没有删除权限。就是这个外在于他的、他无法控制的记忆存根，在冲动和行动之间，替他撑开了他十三年来从未自动撑开过的那半秒钟。

4.4 容错：违约之后，关系没有断裂

事情的后续同样关键。阿杰跑掉了，但他咽不下这口气。两个星期后，他在另一个地方堵了那两个少年中的一个，动了手，把人打伤了——鼻骨骨折。他自己被带到了派出所。

老周知道这件事的时候，是第二天——阿杰的母亲打电话给训练营，说孩子出事了。老周去了派出所。他见到阿杰的时候，阿杰的第一反应是把脸转过去——他不想被看见，他不知道怎么面对。但老周走过去，蹲下来，问的第一句话不是“你为什么要打架”。他问的是：“你受伤没有？”

阿杰告诉本文作者——在一次后来的访谈中——那一秒钟他脑子里全空了。他之前演练的所有解释全部作废。他本来准备好了硬的、软的、推卸责任的、卖惨的，一共四套说辞。但老周问的是他有没有受伤。这个他从来没给别人说过的、他不敢给别人说的、连他自己都觉得丢脸的“怕”——那天晚上，他对着老周全说了。说完之后他等着被骂。老周没有骂他。老周只是说：“我听到你说怕了。那下次，你在怕的时候，能不能先给我打个电话？”

他没有说“你不应该打架”，没有说“我对你很失望”，没有说“你辜负了我的信任”。他问的是：下次，在同一个怕出现的时候，你能不能做一个你之前没做过的动作——在出手之前，先打一个电话？这不是要求他停止打架。这是在他已经知道他会怕的情况下，给他提供了一个新的、在怕和出手之间的操作选项。而他之所以能接受这个选项，不是因为它听起来很有道理，而是因为这个选项的提供者，是那个他刚刚在他面前失信、却没有被他收回认可的成人。

这就是证身操作中那道最关键的容错回路：违约之后，关系没有断裂。没有这道回路，证身就只是一种谎言检测器——你只能在我面前说正确的话，说错就完了。而只有有了这道回路，证身才真正成为生成机制的脚手架：它允许他在违约之后不销毁关系，从而让那种“下一次，在违约之前先做出另一个动作”的能力，在不被恐惧支配的条件下，被慢慢长出来。阿杰后来没有成为犯罪者。他在那次事件之后被老周介绍到了另一个社区的公益项目里做志愿者，再后来学了汽修，现在在一个修理厂工作。他不完美，他仍然有时会冲动，但他拥有了一个他在十三岁时完全没有的东西：在冲动面前，他会想起有一通电话可以先打。那通电话，是他的大脑里一根以前从未被铺设的神经通路。它不是从内化里长出来的，它是从一段被见证的关系里，被那许多次的“我说了→他记住了→下次他问了我→我做错了→他没有抛弃我”的循环，一根一根地焊接出来的。

4.5 证身不制造天使，它制造半秒钟

最后必须澄清一件事，以免这个案例被误读为某种温情主义的道德童话。老周不是圣人，证身不是万能药，阿杰也没有变成一个从不犯错的人。他仍然会愤怒，仍然在某些情境下想要动手。证身在他身上制造的，不是一个完美的道德人格，而是在冲动和动手之间的那半秒钟制动。这半秒钟，对于阿杰来说，从来没有被任何人在他的生命里安装过；对于阿杰来说，是被老周那几百次的追问、点头、记住和不抛弃，一条突触一条突触地焊接出来的。

本文主张的全部规范性的落脚点，就是这半秒钟。刑事归责体系的追问，不应该只停在法庭上的那一刻——他有还是没有那半秒钟？它应该追问的是：在那半秒钟还没有被需要的那些年里，社会给这个人焊接那条制动装置的机会了吗？如果没给，那么在量刑的时候，这

个“没给”本身，是不是应该被算进去？这个追问，就是本文接下来要建构的“生成性归责”框架的核心。

五、论证的边界：对最有力反驳的正面回应

在将本文的核心判断推入法律实践之前，必须先让它经受几轮最应该被请进来对质的攻击。一个好的判断，是在撞过这些硬反驳之后还能站住的东西。

5.1 “惩罚本身就是对未经见证的首次公共见证”——为什么这个论点不能成立

这是本文会遇到的最凶险、也最有深度的反驳。它的经典版本——在某些黑格尔主义法哲学的当代诠释中可以找到其雏形——是这样表述的：当社会通过其惩罚机构，以公开的、仪式化的方式宣告“你所做的是错的，你本可以不这么做”的时候，它恰恰完成了那些在被告童年中缺席的成人从未为他做过的事——它见证了他。如果童年缺失了“证身”，那么法庭的审判本身，不就是社会为他补上的、那场迟到的、对他的选择进行公开追问和命名的仪式吗？惩罚不是对证身缺失的漠视，而是对它的首次公共修复。

这是一个真正有分量的立论。它不容许我简单地宣布“为时已晚”了事。我必须论证：为什么事后公共见证在生成机制的意义上与事前证身是两种本质不同的事件，以至于前者永远不能替代后者的缺失。

答案是：两者的运作对象不同。事前证身作用的操作对象，是冲动与行动之间的那半秒缝隙；事后惩罚的见证，作用的操作对象是行动已经完成之后的行为后果的公共定性。二者看起来都涉及“对选择的审视”，但实际上分别作用在认知系统的两个完全不同的时间窗口上。事前证身之所以能生成审视能力，不是因为它向行动者传递了“打架是错的”这一信息，而是因为它在冲动正在发生的那个时刻，从内部为他植入了一个他无法单方面消除的外部记忆存根——一个他知道另一个人脑子里存着的、关于他曾经口头承诺过的“不动手”的版本。这个机制在事后无法被复制，不是因为事后太晚——从道德教导的角度来看，事后永远不晚——而是因为事后他不再处于冲动之中。审视能力的对象是冲击本身，它必须在冲击正在运行的那一刻被激活。拿事后的惩罚去替代事前的证身，等于拿驾照考试去替代驾驶过程中的刹车干预。

更进一步，惩罚作为公共见证，恰恰是在被告最不信任外部权威的時刻发生的。阿远站在法庭上，他所面对的不是一个像老周那样的、在几百次容错中被反复验证过不会伤害他的成人，而是一个穿着法袍的陌生人，其全部功能就是对他做出有罪宣告。在这种权力落差下，法官的“追问”不可能产生证身的心理效应——因为它不具备证身成立的那个前提条件：被告在这个他者面前保持可信度的意愿，是自发产生的，而不是被法庭的强制管辖权所制造的。法庭可以强迫他站在那里，可以强迫他听到判决，但它不能强迫他在意法官对他的看法。而这个“在意”，恰恰是证身回路的第一个节点。

因此，惩罚作为事后公共见证，与事前证身之间，存在一道结构性的、不可跨越的时间差：前者是事后定性，后者是干预过程。前者永远无法替代后者，不是因为后者更“好”、更“温和”、更“人道”——不是这些价值判断——而是因为后者的工程对象是那个转瞬即逝的、在冲动运行之中才会短暂开启的认知窗口。一旦窗口关闭，事后输入任何信息都无法再作用于那个已经关闭的窗口内部。这个论证不只是对黑格尔主义反驳的回应，它也设定了“生成性归责”概念的精确边界：它不是主张“社会没有做好所以被告无罪”，而是主张“社会未能在认知发生的时间窗口内完成它本可以完成的那道焊接，因此归责的程度应当相应扣减”。它的落脚点不是犯罪之前的免责，而是量刑之中的比例。

5.2 证身是否可以被策略性操控？

第二个预判性追问来自刑事辩护的实务视角：如果“证身关系的剥夺”成为一项可以减轻责任的量刑因素，那么辩护方是否会开始策略性地寻找或制造被告“童年没有被见证”的证据？一个犯罪者是否可以辩称——无论真假——“没有任何人见证过我的选择”，来谋求更轻的刑罚？这不是理论上的杞人忧天。任何新的量刑情节的引入，都面临被泛化和滥用的风险。

对此，本文的回应不是给出一个杜绝滥用的一揽子方案——那不是一篇理论论文能做到的事。但需要明确两点。第一，并非任何“没有人管过我”都可以构成证身剥夺。本文对证身的界定是高度结构化的：它必须涉及一个特定种类的、包含了提交、追问、记住和容错四道工序的关系事件。日常体验中的“缺乏父母关爱”不必然等同于证身剥夺。一个被严格管教但从未被给予追问空间的孩子，和一个在放任中被忽视的孩子，可能都缺乏证身关系，但缺乏的机制不同；反之，一些在艰苦环境中长大的人，仍然有可能在某个非亲属成人（老师、教练、邻居、社工）那里经历过有限的证身时刻。因此，“证身剥夺”在证据法上不是一个简单的是非判断，而是一个需要对被告成长过程中的关系结构进行具体评估的任务。

第二，这恰恰是为什么本文主张将证身剥夺作为量刑情节、而非责任排除事由。如果“未经见证”被当作责任排除事由，那么一旦被证明，被告就可以完全免责——这将制造巨大的滥用激励。但如果它的法律效果只是量刑阶段的减让，那么法官可以通过社会调查报告来评估其具体程度和可信度，并赋予其相应的权重，而不必进入“有或无”的三元开关。这种阶梯式的处理，可以部分地缓解滥用的担忧。

5.3 证身关系是否可能有害？

第三个反驳是：如果证身关系发生在一个封闭的、反社会的群体内部——例如一个帮派的“导师”对少年进行追问、记住和容错——它是不是也会生成同样强度的承诺结构？如果是，那么证身这个概念不就失去了它在本文中的规范性的方向了吗？

这个反驳迫使我把一个此前隐含的前提挑明为明确的理论判断：证身不只是一种形式化的社会操作，它内部嵌有一个不可剥离的规范性的条件——容错回路中的“退出权”。在一个帮派“师徒”关系中，一个少年也可能被追问“你确定要跟去吗？万一翻车了你怎么办”，也可能被要求记住承诺，甚至也可能在第一次违背后被允许继续留在帮派。但关键的区别在于：如果他最终决定退出——他拒绝了那把刀，他说“我不想跟去”，帮派的他者会怎么反应？如果他的退出会导致他被当作“叛徒”来惩罚、被排斥、被攻击，那么他此前每一次对“选择”的提交，都不是真正的证身。那只是政治审查——他只是在不断地确认自己对帮派意志的服从。

真正的证身，在规范性的内核上，要求容错回路包含一个真实的退出选项：对方可以拒绝这个行为，而不会因此被关系毁灭。没有这个退出权，就没有“撤回成本不可单方面归零”的真实性——因为他撤回承诺的代价是失去一切，那其实等价于他从未有过选择。这正是证身这个概念不能被纯粹形式化使用的根由：它的生成力，不仅依赖于追问和记住，更依赖于那个追问者是否为被追问者保留了一个可以不遵从、却仍被接纳的空间。阿杰在老周面前第一次说出“我怕”的时候，老周没有因此收回对他的认可。如果老周当时露出了哪怕一丝轻蔑，整个证身回路就会在那个节点上断掉——跟帮派的回路没有区别。

这意味着，证身这一概念的运用本身就需要带有规范性的判断：不是任何一段追问-记住的关系都能通过“证身”的标准。只有那些包含了真实容错空间的，才具备本文所描述的那种生成审视性制动装置的功能。这同时设定了这个概念在实际法律适用中的一个操作要点：在评估被告成长过程中是否存在证身关系时，调查关注的不能仅仅是“有没有一个成人管过他”，而必须进一步细究：那种关系的结构，是否允许被告在其成长的关键时刻，行使过真实的“不遵从而被接纳”的体验。

六、从哲学判断到法律操作：“生成性归责”框架的初步建构

前五章完成的是证身作为一个生成机制判断的全部理论论证。但这个判断如果只停留在法哲学层面，它只是一个精彩的智识练习，而成不了一篇能改变法学讨论格局的论文。本章必须迈出那一步：把哲学洞见翻译成法学家和法官能用的操作性框架。

6.1 “生成性归责”的基本结构

本文提出“生成性归责”这一概念，以区别于现行刑事归责体系的“门槛式归责”。生成性归责的核心主张可以概括为这样一个递进结构：

第一层：能力门槛判断（维持现行法标准）。被告在行为时是否具备辨认行为性质和依照辨认控制行为的能力？如果答案是否定的，则适用现行法中的无责任能力或限制责任能力条款。在这个层面上，证身剥夺不产生任何影响——它不是一个精神病学判断，不当被用于否定能力门槛本身。

第二层：责任程度的裁量（引入本文的新变量）。在确认被告具备门槛式能力的前提下，量刑时应当进一步考虑一个被现行法体系系统性忽视的问题：这份能力，在其成长的关键时期，是否曾被给予足够的生成条件？具体而言，法官应当考察被告在早期成长过程中，是否经历过那种包含了追问、记住、容错和退出权的证身关系。如果这种关系在被告最需要它的那些年里系统地缺失，那么被告虽然在行为那一刻“能够选择”，但他那种“能够”是未经充分锻造的——他的选择能力没有经过那道让冲动在行动之前被审视的社会工序。这一事实，应当构成减轻其责任程度的一项独立考量因素。

第三层：社会亏欠的司法承认。在量刑理由的宣告中，法官不应只对被告说“你本可以不这么做”，而应当同时对那个从未为被告建造那扇门的共同体，做出公开的司法陈述：本庭注意到，被告在其成长的关键时期，没有被提供让他的选择能力得以完整生成的必要社会条件。这不是脱罪，这是在量刑的比例性裁量中，承认社会亏欠是责任分配的一个真实变量。判决不仅是对被告行为的回应，也是共同体对其自身失职的司法记录。

6.2 法教义学的安置：它是什么性质的量刑情节

现行刑法典中的量刑情节可以分为法定情节和酌定情节。从比较法的视角来看，“社会不利背景”在澳大利亚Bugmy案（2013）中被明确承认为减轻量刑的因素，加拿大Ipeelee案（2012）更进一步指出，对于原住民被告，其遭受的系统性歧视和代际创伤是量刑时必须纳入考量的结构性因素。这些判例确立了一个原则性的洞见：一个人的道德可责性，不仅取决于他在行为那一刻做了什么，还取决于他是在什么样的条件下被塑造成他做出这个行为时的那个人的。

证身剥夺情节在刑法定性上，最恰当的定位是作为一种特定形态的“社会不利背景”——但不是笼统的贫困或教育缺失，而是对一种特定的、与选择能力之生成直接相关的社会条件的剥夺。它作为一种酌定量刑情节，可以在量刑阶段由辩护方提出，由法院依据社会调查

报告来评估其存在与否及其程度，并据此在法定刑幅度内适当调整刑期。在法教义学上，它的规范依据可以落脚在现行刑法关于“犯罪情节”的概括性条款中——那种“虽然法律没有明确列举，但从罪责原则中可以推导出来”的减轻因素。

它不同于现有的“精神障碍”减轻情节——后者处理的是能力缺陷，而它处理的是能力未被生成。它也不同于“被害人过错”——后者指向的是对方行为的归责影响，而它指向的是被告自身生成史中的社会亏欠。因此，它不会与既有条款发生体系冲突，而是为量刑分类增加了一个新的、迄今被忽视的考量维度。

6.3 操作上的难题：如何评估“证身关系的剥夺”？

这是本章最难的部分。一个法官，在案头堆积如山的案卷里，如何去评估一个被告在二十年前是否经历过“被追问并记住的承诺”？这不是一个可以事后用心理检测量表来回答的问题——没有任何量表可以穿透二十年的记忆衰减和自我叙述的不可靠性。本文对这个问题不能假装有一个完美的技术方案，但可以指明一条原则性的调查路径。

社会调查报告的扩展。现行量刑程序中已广泛使用的社会调查报告，通常包括被告的成长经历、家庭背景、教育程度、前科情况。在引入证身剥夺情节之后，这份报告需要增加一个专项评估：在被告的童年和青少年时期，是否存在一个持续的、包含了追问、记住和容错等特定操作的关系事件链？这不能只靠被告自己的叙述（他有明显的动机夸大自己的缺失），也不能只靠亲属的证言（他们同样有动机要么淡化、要么甩锅给他人）。评估需要寻找的是那种不容易被单方面伪造的、可交叉验证的踪迹——学校老师的家访记录中有没有提到过这个孩子的沉默、退缩和“拒绝沟通”？社区工作者或社工的存档中有没有关于这个家庭“拒绝外部干预”或“对外封闭”的记录？少年时期的同伴那里有没有一致的证言说“他从来不跟大人说真话”？

这种评估当然不可能追求自然科学的精确。但量刑从来就不是一门精确科学——它本来就是在有一个有弹性的幅度内，对各种加重和减轻因素进行综合权衡的实践理性。心理健康评估、品格证据、再犯风险评估，这些早已被纳入量刑考量的因素，没有一个是精确量化的。困难不是不做的理由；只要这个问题是真实的，操作上的困难就只是一个需要逐步解决的工程问题，而不是一个应该回避的智识借口。

6.4 防止滥用的结构性限制

最后，必须为这个新情节设置防止滥用的几道结构性闸门。第一：证身剥夺只在严重暴力犯罪和长期犯罪模式中才产生实质性的量刑减让——在这些案件中，“选择能力的生成失败”与“行为的严重性”之间的关系是最直接、最可辨的。在轻微犯罪中，它可以被提及但不产生独立的减让效应。第二：证身剥夺不能用被告的自我陈述作为唯一证据。必须达到一定的证明标准——至少要能展示出存在一个合理的推论，即在其成长期，那种包含了追问、记住和容错的关系事件是系统性缺失的，而不仅仅是“不够好”的家庭环境。第三：不适用于被告本人曾经主动拒绝或破坏已有的证身关系的情形——如果有可靠证据表明，在其成长过程中确实有人曾经试图提供这种关系，但被告以持续性暴力和敌意将其推开，则证身剥夺的主张不应成立。这三重限制的设计意图是：让证身剥夺情节只保护那些真正是因为社会的系统性忽视而从未被给予选择之生成机会的人，而不保护仅仅把童年缺失当作脱罪工具的策略主义者。

6.5 一个必须承认的操作局限

在结束本章之前，还有一个重要的限定必须做出。本文在此提出的“生成性归责”框架，在司法适用的技术细节上是不完备的。它的完整建构——包括社会调查的具体指标、证明责任的分配标准、与现行量刑指南的衔接方式——需要另一项独立的、以实证法学方法为主导的研究来完成。本文的贡献在于命名了那道缝隙，并给出了它在量刑裁量中应当被考量的原则性理由。论证的工程学完善，需要法律实践中渐进的、以个案积累为基础的发展。一个生成机制的论文，应当有勇气在展开其规范性的后果的同时，诚实地标出它自己尚未抵达的边界——这不是学术的不自信，而是对司法系统复杂性的尊重。

结论：法槌落下之前，那扇门是否被建造过

本文的理论旅程是从一间刑事法庭开始的。现在让我们回到那里。

阿远的法槌已经落下。但在这篇论文的核心判断里，那个法槌的正当性，不能只取决于他在被人激怒的那一刻是否有精神医学意义上的能力去选择不挥出那根钢管。一个真正完整的归责，应当追问的是更早的一步：在他还没有站上被告席的那些年里，在他还是那个被反锁在出租屋里的五岁孩子、那个被老师当作“特别乖”的沉默的八岁学生、那个用身体挨打来结束第一次偷窃的十一岁少年、那个用不犹豫来保护自己不被踢出唯一接纳他的群体的十四岁男孩的时候——这整个共同体，有没有在他和被激怒的那个瞬间之间，为他建造过哪怕一扇可以被推开的门？

刑事归责体系今天只做法槌落下之后的事。它把“自我决定”当作一个给定的前提，像验收一台出厂设备那样检验它：开关在盒子里吗？在。那就通过。但本文用了整整六章的论证来证明一件事：那根从开关通往行为输出的电线，不是在出厂时就自动接好的。它是在一

个人成长的几十年里，被那些追问过他、记住过他、容错过他、没有在他失信时抛弃过他的他者，一毫米一毫米地焊接出来的。如果那条电线从未被铺设过，法庭就没有任何理由假定它在那一刻会奇迹般地自己接上。

这绝不是主张“社会是犯罪的根源”那样的简单因果论。那太粗疏了。本文做出的，是一个窄得多的、在操作上可被具体指认的判断：选择能力的生成，需要一种特定的社会工序。这道工序的缺失，与犯罪的发生之间，存在一个可以被理论命名并在实证上被检验的因果界面。证身，就是这道工序的名字。

最后，按照一个经得起反驳的理论应当具备的品格，我在此诚实地写下这个判断的证伪条件。本文的核心因果命题是：包含追问、记住和容错的成人见证关系，在青少年的成长期对其后的严重暴力犯罪具有独立的降低效应——这个效应的发生机制，不是通过一般性的社会支持或道德教育，而是通过在冲动和行动之间植入不可单方面归零的承诺结构。这个命题的证伪条件是：如果有严格设计的大规模追踪数据表明，这种特定类型的见证关系——当其他已知保护因子（家庭社会经济地位、学校教育质量、同伴影响、心理强制干预）被控制之后——对严重暴力犯罪率没有显著的独立效应，则本文的全部理论建构将被推翻。

一个理论写出自己的死法，不是因为它不自信——恰恰相反，是因为它对自己的判断足够认真，以至于它愿意在一张所有人都能看见的检验台上，接受它可能失败的结果。这是生成机制最深的风骨：它不怕被检验，它只怕自己什么都说得通。

注释

[1] 本文引述的阿远、阿杰两个案例，均系基于公开裁判文书、青少年犯罪研究和司法社工实践中的典型情形建构的匿名化、合成性思维示例，不指涉任何真实个案。其中的人名、地点与具体情节均经过虚构、简化与时间线压缩处理，仅作理论辨析之用，不具有实证数据的性质。特此说明。

[2] 本文对布迪厄惯习理论的批评，严格限定于“惯习可脱离原始生成场域而被个体原子化携带”这一默认前提。这并不减损惯习理论在解释社会结构再生产、符号暴力及阶层惯习传递等问题上的穿透力。相反，本文提出的“证身”概念可以视为在微观生成层面，对惯习如何在特定关系中被最初“雕刻”的一种生成机制补充。

参考文献

- Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Bugmy v The Queen [2013] HCA 37.
- Duff, R. A. (2001). *Punishment, Communication, and Community*. Oxford: Oxford University Press.
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, 68(1), 5–20.
- Freud, S. (1923). The Ego and the Id. *Standard Edition*, Vol. 19. London: Hogarth Press.
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London B*, 359(1451), 1775–1785.
- Hart, H. L. A. (1968). *Punishment and Responsibility: Essays in the Philosophy of Law*. Oxford: Clarendon Press.
- Laub, J. H., & Sampson, R. J. (2003). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard University Press.
- Mead, G. H. (1934). *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674–701.
- R v Ipeelee [2012] SCC 13.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

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：与承诺理论、叙事同一性的硬边界

正文切开了法兰克福的二阶欲望与布迪厄的惯习，但漏了两个更近的对手。斯坎伦式的承诺理论论证承诺之所以约束人，在于它在他人身上产生了合理期待——其约束力依赖见证者；本文的证身恰恰主张一种未经任何人见证仍然生效的凝固，因此它不能被承诺理论覆盖，反而构成对后者的一个反例类型。麦金泰尔与利科的叙事同一性论证自我通过讲述而获得连续性——其载体是叙述；本

文的证身发生在讲述之前，且当事人往往讲不出来。两条分野共享一个可判别面：若所有证身个案在充分的叙述条件下都能被当事人自己复述为一段有连贯理由的历史，则叙事同一性足够；若那些最硬的凝固恰恰是当事人说不出来的，本文成立。

增补二：证身的操作化收紧

正文已给操作性定义，此处收紧到可编码。三项：不可归零性——个体在明确得知该承诺不再有任何外部约束（对方已不知情、不再在意、或已不存在）之后，行为是否仍受其约束；见证缺失度——该承诺形成时是否有任何他人知晓，及其后是否被任何人确认过；撤回代价——若撤回，个体所报告的损失是外部后果还是自我形状的改变。第三项是关键：若撤回代价全部指向外部后果，则该承诺是策略性的，不是证身。

增补三：可证伪的检验设计

核心命题——存在一类无人见证却不可单方面归零的承诺结构——可通过一个反事实设计检验。选取一批个体，识别其成长期中的无见证让步事件（由回溯访谈与可佐证材料交叉确认），并在实验情境中提供一次“该承诺已完全失效、无人会知道你违反”的明确解除条件。承诺理论预测：解除条件成立后行为约束显著减弱；本文预测：一部分个体的约束不减弱，且这部分人应集中在让步发生早、重复次数多、当时解释通道闭合的那一组。若两组无差异，或差异可由一般尽责性人格特质完全解释，本文即被证伪。

增补四：与同批《审判前的塌陷》的分野

作者同批已发表《审判前的塌陷：一个可归还主体的结构性消亡》，两篇位置紧邻且方向互补。那篇主张认知基础设施从未被安装，因此“你本可以不这么做”这句话落不进去；本文主张一种确实被安装、却在无人见证中凝固成特定形状的承诺结构。两者共同界定了责任语言的两种失效方式：一种是无从落地，一种是落在了一个已被写定的形状上。这一分工须明确，否则读者会把两篇读成同一个减责论证——而它们在归责后果上恰恰相反：前者动摇归责前提，后者要求归责看得更长。

编辑增补所依据的核验文献

T. M. Scanlon, *What We Owe to Each Other*, Harvard University Press, 1998.

Harry G. Frankfurt, “Freedom of the Will and the Concept of a Person,” *The Journal of Philosophy*, 68(1), 1971.

Paul Ricœur, *Soi-même comme un autre*, Éditions du Seuil, 1990.

Alasdair MacIntyre, *After Virtue*, University of Notre Dame Press, 1981.

Pierre Bourdieu, *Le sens pratique*, Éditions de Minuit, 1980.