



伦理困境中的黄金法则何以可能

马梁英

【内容摘要】 亲缘、国家、种族等将人类划分成了不同群体，使行动者往往只维护自身和群体成员的利益与道德形象，践行着有限度的道德准则。然而，作为理想社会的基石，“黄金法则”要求设身处地为他人着想，强调突破边界的道德力量。以黄金法则求同存异的一致性要求为起点，道德心理学提供了当下实现黄金法则的可能路径。但相关实验结果表明，有限的认知资源与情感直觉的驱动，对人们的道德判断与行为构成威胁，并与自我控制和理性推理的约束力量形成对抗，构成黄金法则的日常危机。而通过打破自我中心主义的道德教育，拓展社会群体边界的道德认同，塑造超越族群身份的认同方案，人们有望走出实验室，更为自主地构建广阔的道德关怀圈，推动公正而平等的人类命运共同体构建。

【关键词】 黄金法则 伦理困境 日常危机 道德心理学 社会认同

【作者】 马梁英，北京大学社会学系博士研究生。（北京 100080）

【基金项目】 国家社会科学基金一般项目“宗教和灵性心理学的跨学科研究”（20BZJ004）

作为经久不衰的道德议题，“黄金法则”要求个体本着珍视与自己无二的心意来对待他者，无论血亲、友人、陌生人乃至仇敌。在儒家语境下，它被表述为“己所不欲，勿施于人”，而在基督教中则是“你希望别人怎么对待你，你就怎么对待别人”。《辞海》中将“道”和“德”分别定义为“理想的人格与道德图景”和“立身依据与行为准则”，而当道德要求被冠以“黄金”之称时，无疑更加具有超凡的崇高色彩。依据黄金法则，个人的道德实践要顺应自身关怀亲者的道德情感，和自己宣称秉持的价值理念相一致，不因交往对象和外在条件而改变。这条道德法则可以被追溯至不同时代与不同文化传统，其奠定了不同共同体的道德根基。

当下，分化冲突的社会现实正在不断挤压着道德共识的边界，使黄金法则看似愈发难以企及。人们习惯于宽容自己和“自己人”的道德污点，对“外人”的过失则给予严苛的批判与社会制裁，或是对他人的痛苦漠不关心，因先入为主的偏见而认为其罪有应得。即便黄金法则简明地告诫社会成员应笃信公平、公正与良善，在复杂的个体认知过程与集体成员身份



的驱使下，越来越多的人仍然被卷入了道德危机。他们追随特定道德权威，以正统意识形态划定道德标准适用的清晰边界，将不道德的外群排斥在外。换言之，倘若遵循无边界的黄金法则，行动者反而可能无法被认定为一个有道德的、荣耀的群体成员，随之面临丧失成员资格的风险。

面对日益极化的世界与壁垒分明的群体，理解黄金法则的道德力量日益紧迫。现有关涉黄金法则的研究主要提供了三类视角：首先是对黄金法则所代表的伦理体系与宗教传统的哲学反思，试图厘清道德规范的抽象定义与多种道德类型所导致的分歧；其次是借助情境材料构建道德困境，观察人们如何受制于关系与文化等因素的道德特殊性的影响，心理学对此进行了大量探索；此外还有对黄金法则适用对象的拓展，譬如技术造就的道德行动者应当遵循何种道德范式；等等。相较于对行动者思维过程及差异的强调和拓展，黄金法则的实现策略则未被明确提及。

笔者认为，持续拓展或细化道德不一致场景的道德心理学研究，构成探索黄金法则实现策略的关键。本文旨在明确黄金法则在道德心理学实验研究中的相关性，将黄金法则带回现实生活经验，揭示这些呈现行动者有限注意力与关怀的研究到底划定了怎样的道德视野。通过捕捉那些不可避免的道德偏离、扭曲道德标准以及区别对待道德对象的时刻，我们能够进一步明确黄金法则所面临的现实挑战，倡导一种协调个体自我认知与群体认同的应对策略。概而言之，本文试图回答以下问题，黄金法则提出了怎样的道德要求？通往黄金法则所面临的阻碍有哪些？在克服情境与个体限制的基础上，人们如何能超越造成隔阂的群体边界，使黄金法则真正落地生根，发挥其道德提升的使命？

黄金法则的一致性要求

不同文化中“黄金法则”的表达差异，意味着它可能提供了不同的要求选项：儒家传统中的“勿施于人”规定了“不要做什么”，而基督教的“爱人如己”则要求“做什么”。当“不要做什么”被视作定向防御时，更具主动性的后者有着更为消极的“白金法则”之称。在另一视角下，根据对别人行为的预期而主动行动，也可能被解读为要求他人回报的利己式道德，这时“己所不欲，勿施于人”要求内省的人格修养，又展现了胜过基督教不求回报的非功利性道德观念。事实上，抽象层面上的高下之争并不重要，因为“己所不欲，勿施于人”具备不局限于个体的内涵，它能够引导各层级践行角色使命，兼具天下主义的开放性，因而被钱穆视为一种信任道德之力的最大功利。^①而对基督教语境下的黄金法则而言，仅以互惠关系概之，也忽视了其内在张力。利科(Paul Ricoeur)认为，“你希望别人怎么对待你，你就怎么对待别人”，其宣扬的等效逻辑下的双边正义，与“爱你的仇敌”所展现的单向的、富余的至上之爱，共同构成了基督教伦理的一体两面。^②

可见，不同文化传统下的黄金法则都强调为人处世的一致性，要求主体的道德理念不受情境与行为主体的影响，在不同社会环境下的人际交往中持之以恒地保持自控与内省，并在更为宏观的水平上，不因譬如性别、国籍、种族、信仰等立场差异而对外群成员变换标准。黄金法则的一致性可以被进一步解读为两条准则：一是，人们应对得起自己的良心，也就是自己的道德信仰；二是，评价应不偏不倚，不受对方身份影响。^③

一致性要求产生了最简洁的黄金法则执行步骤——存异而求同。人们首先要觉知自我与他人

① 钱穆：《从中国历史来看中国民族性及中国文化》，北京：中华书局，2016年。

② Paul Ricoeur, "The Golden Rule: Exegetical and Theological Perplexities," *New Testament Studies*, vol. 36, no. 3, 1990.

③ Harry J. Gensler, *Ethics and the Golden Rule*, New York: Routledge, 2013.



的差异,以及不同群体间主导原则和规范的差异,但不以这种差异或熟悉程度确证优越与劣等之分。同时,行动者承认且尊重他人与自己有着一样的不被伤害的意愿,由此不必艰难考量遵守道德的必要性,担忧为他人和社会付出能否带来回报,因为一致性要求意味着他者行为的可预测性。个体以推己及人的同理心与他人产生联系,感知自身情绪变化并换位思考,以此来保障与促进道德行为,减轻道德权衡的负担,实现差异下的和谐共处。

黄金法则一致性要求有理想化之嫌,然而其可行性并不必然诉诸神明的至高权威,或是圣贤的仁恕之道,人类的进化历程早已验证了这一适应性进化法则。自然选择虽偏爱资源争夺的胜者,人类行为中却普遍存在着互惠利他的倾向,这使人们在严酷与不可预测的环境风险中得以生存。人们会帮助并非血亲的他者,对他者伸出援手从而树立道德声望,进而使自己也能得到其他人的帮助。避免不道德行为的适应性机制也展现出黄金法则的逻辑:如果人们搭便车为自己谋取好处,^①或是伤害他者,^②则可能遭到同样的报复,为此付出高昂代价。黄金法则兼具鼓励与强制的效用,既能带来生存优势与社会认可,同时也促使人们执行惩罚以维持公平。

不过,黄金法则的一致性要求与道德的不一致问题并不是光谱的两极,而是呈现出一种不对称的关系。从根本上来说,自然选择鼓励的是群体内合作。人们为黄金法则所拓展的边界,可能至多延展到自身所属族群而非全人类。黄金法则中作为一致性起点的“己”不单单是个体自我,它还是在群体中占据一席之地的成员。这使得为了维护自身利益或袒护所属群体而实施有违一致性的道德观的行为,成为一种直觉性的、快速而自动的选择,人们需要顺应看似不言自明的共同价值观。而想要超越血缘与群体界限践行道德行为准则,则首先需要识别出植根于文化传统,或是存在于意识形态、人际关系与个人认知中的偏见与敌意,然后进行批判性分析。后者要求控制情感冲动,克服普遍接受的群体内的立身依据,这无疑不是高效实用的手段。黄金法则包含身份的不确定性,要求不同立场的社会成员之间实现差异的相容,而非预设无原则的统一,因此它离不开深思熟虑且缓慢的处理过程。

需要说明的是,本文对黄金法则的讨论并不旨在勾画一个独立于伦理困境之外的、超凡崇高的道德场域,而是要首先识别这种一致性与不一致性之间不对称的内在张力。这意味着仅仅意识到存在偏离道德规范,还不足以鼓励人们打破自我中心的利益考量,给予不同立场的个体更慷慨的共情。换言之,唯有先明确导致道德缺陷的因素,才能依据一致的道德标准探索检验自身道德行为、提升道德水平的可能路径。

自我控制的日常危机及监控机制

在成文的社会规章制度之外,道德规范发挥着无形的约束作用,是一个社会文明程度的关键象征。强调自律、尊重与公正的黄金法则,是任何社会都试图追求的道德理想,它有益于维系家庭亲密关系,促进职场中的团队合作与提高工作效率,在社交场合建立良好的人际关系,亦是建设和谐社会的基石。出于对个人声誉的考虑和对社会排斥的担忧,社会成员显然很难公然抵抗黄金法则,宣称自己不能对他者给予一视同仁的理解与关怀。不过,人们并不认为自身德性是外界强制规约的结果,而是相信无论过去与未来,不管面对怎样的伦理困境,自己都能自觉自发地表现出胜于大多数普通人的道德行为,不会见利忘义、损人利己。^③显然,这是美化后的理想图景。

① Ernst Fehr and Urs Fischbacher, "Third-Party Punishment and Social Norms," *Evolution and Human Behavior*, vol. 25, no. 2, 2004.

② Robert L. Trivers, "The Evolution of Reciprocal Altruism," *The Quarterly Review of Biology*, vol. 46, no. 1, 1971.

③ Nicholas Epley and David Dunning, "Feeling 'Holier Than Thou': Are Self-Serving Assessments Produced by Errors in Self- or Social Prediction?" *Journal of Personality and Social Psychology*, vol. 79, no. 6, 2000.

拥有善良、公正而无私奉献的道德观念的人们，其现实道德行为与“心中的道德律”的偏差，呈现了道德的日常形式中难以被人察觉的弹性。

日常情境中，自我控制能够保障黄金法则的两个关键过程不至陷入失灵：一是根据一致的文化规范发起行动的过程；二是对自己进行监督抑制，避免在即时刺激下变换标准。心理学实验中所运用的研究技术致力于更好地检验人们无意识的、难以表达或有意隐藏的心理偏好和行为倾向，从而使我们得以探索行动者执行道德标准时自我监控的意志与实践过程。例如，在一项研究中，进行内隐联结测验（Implicit Association Test）的白人被要求与黑人谈论种族问题后，再进行斯特鲁普测试（Stroop Test）。实验发现，在内隐联结测验中对黑人无意识偏见越强的研究参与者，越是能在跨种族谈话中更好地控制自己的行为以压制种族歧视。然而，因为先前集中与消耗了较多注意力来抑制不一致的行为反应，他们在随后执行功能的任务中便表现不佳。^①这意味着通过自我控制来避免冒犯他人并且保持道德形象的努力，可能在未来适得其反——即在无暇深思熟虑的时候，偏见与歧视将难以控制地浮现出来，成为威胁正常人际互动的不良行为。

人们倾向于认为道德行为展示了人的本性，不同道德行为下潜在的价值观差异是难以弥合且令人生厌的根本分歧，但自我控制不仅是个体层面上的差异，同一个体在不同场景下也会产生区别。践行道德理念的能力实际如同体力一样有其限度。越是试图维护道德高尚的自我观，越难以持续遵循黄金法则的理想标准。自我控制能力的有限效用，被比喻为一个“取之有竭”的资源池，或是劳作后乏力的身体肌肉。伴随对行为调控能量的消耗，自我耗竭（ego depletion）的人们会变得疲于应付后续的道德情境，原本遵循道德的人们也难以抵抗撒谎和作弊的诱惑。可以预想，在高度占用认知资源的工作场所，自我控制能力的耗竭很可能加剧催生了造假欺诈与损公肥私问题，将对社会造成难以估量的经济损失。^②对自我控制的要求不仅是实验情境下认知负荷的结果，更是一种日常性的挑战。而当其叠加经济上的弱势地位时，短期压力与忧虑会演变为一种长期自控能力的稀缺与沉重的心智负担，使人容易犯错，反应变慢。^③不过，当意识到自我耗竭的潜在风险，借助于认知资源分配优先次序的调整，加以外界的监控，个体的自我控制能力仍然有望发挥其对不道德行为的抵制。

然而，黄金法则的日常危机除了自我监控的有心无力，还有实际运作中的临时性和任意性问题，这比起类似体力下降的自我耗竭来说更难以察觉。鉴于物理污秽与道德污点，以及洁净和道德上“清白”的关系，许多学者展开有关卫生清洁程度与道德判断双向关系的讨论，认为这可能扮演着合理化社会排斥或更极端的种族清洗的角色。实验发现，通过刺激对清洁的感知，人们的道德概念也随之被启动。同时，激活清洁感与研究参与者的主体视角，他们会因清洁感作出更为严厉的道德判断。^④比如，如果要求参与者洗手，则不仅有助于减轻负向情绪，降低实验参与者对他者进行道德谴责的可能性，也能净化修复道德良善的自我概念，于是减少了人们进行补偿性利他行为的意愿。^⑤因此，维持积极的道德自我观未必会推动人们弥补过往行为过失，或者持续作出更多善行，遵循个人道德价值标准反而可能为将来放纵自我进行不道德行为提供道德许可，造成道德行为的前后不一致。以隐喻形式呈现的表述，称德性卑劣者“令人作呕”，试图掩藏个人生命历程中的“道德污点”，以及因处理复杂任务而感到“心力交瘁”，其呈现了人们与道德相关的价值判断与行为决策的真实过程，而不仅仅是简单的联想与象征意义。

聚焦于省内缺席区域的无意识的研究，在看似无序的碎片化不一致中试图发掘背后有序的运作规则，具体点明了与人们道德心理现象相关的情境因素。但对个人反应的监控，究竟能多大程

① Jennifer A. et al., "When Prejudice Does Not Pay: Effects of Interracial Contact on Executive Function," *Psychological Science*, vol. 14, no. 3, 2003.

② Francesca Gino et al., "Unable to Resist Temptation: How Self-Control Depletion Promotes Unethical Behavior," *Organizational Behavior and Human Decision Processes*, vol. 115, no. 2, 2011.

③ 塞德希尔·穆来纳森、埃尔德·沙菲尔：《稀缺：我们是如何陷入贫穷与忙碌的》，魏薇、龙志勇译，杭州：浙江人民出版社，2014年。

④ Chen Bo Zhong et al., "A Clean Self Can Render Harsh Moral Judgment," *Journal of Experimental Social Psychology*, vol. 46, no. 5, 2010.

⑤ Chen Bo Zhong et al., "Washing Away Your Sins: Threatened Morality and Physical Cleansing," *Science*, vol. 313, no. 5792, 2006.



- ① Kendall J. Eskine et al., "A Bad Taste in the Mouth: Gustatory Disgust Influences Moral Judgment," *Psychological Science*, vol. 22, no. 3, 2011.
- ② Cinzia Cecchetto et al., "Body Odors (Even When Masked) Make You More Emotional: Behavioral and Neural Insights," *Scientific Reports*, vol. 9, no. 1, 2019.

- ③ Jonathan Haidt, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review*, vol. 108, no. 4, 2001.
- ④ Joshua D. Greene et al., "An Fmri Investigation of Emotional Engagement in Moral Judgment," *Science*, vol. 293, no. 5537, 2001.

度上避免由感知觉刺激导致的认知与行为改变,研究者们却无法断言。具身体验与道德表现的作用方向、中介因素以及影响效应,是作用于眼前暂时的特定道德行为还是超越情境的长期一般态度,这些都难以确证。比如,陪审员是否需要避免吃太苦的食物,防止诱发道德厌恶导致的裁决不公?^①又比如,在嗅觉领域,清洁的气味指向“干净的灵魂”,但是,如果男性的体味能使女性在面对道德困境时更倾向于作出避免伤害的道义选择,^②那么保持卫生的建议就变得模棱两可。因为启动的运作机制并未被系统建构起来,行动者无法简单依据某种特质观念而监控和调整行为倾向。

总体而言,黄金法则的日常危机导向以对抗性的努力来维持统一标准,运用意志力以控制、压抑乃至矫饰个人偏好,调整引发道德不一致性的因素来避免问题出现。然而,这种效用能否长期持续尚不可知。由于具体文化语境和日常情境下的影响因素充满不确定性,实验室研究也不能列出所有风险清单来全部施以针对性治疗,这就可能导致人们在调整行为倾向时适得其反。事实上,人们不仅需要竭力在环境干扰下战略性地进行表演,接受保障当下行为合乎道德规范的审视,还需要在时间流动与情境变换中获得持续的道德导航。否则,即便有自我或第三方监控的影响,面对群体认同边界稳定而强大的力量时亦可能被消除,因为即便是最为符合外部道德标准的个体,所具备的也不过是内群偏好的社会性。尽管大多道德提升建议处于个体德性与社会文化两端,面向中间互动情境的实验考察仍说明,黄金法则的道德要求尚不充分,人们可能首先需要从离黄金法则最为遥远的、对不道德行为的预防做起。

情感捷径的风险与协调

伦理道德在中国几千年传统文化中扮演着权威角色,从治世之道到人伦关系,道德既是官方律令所强制的秩序,也是主体德性所自然显现的实践。当下道德法则已经不再是评价社会主体的唯一权威,个体价值由市场逻辑所断定,与物质财富和社会地位紧密相连,人们被驱使追求权力与经济成功。拜金主义、享乐主义与个人主义被视为精神文明建设的主要挑战,相关道德失范行为被描述为个体沉湎于追逐肤浅的情感价值,因冲动与短视而放弃公共利益。前文所述,日常生活中的常见刺激,在不知不觉间动摇着人们视之为难以改变的道德本性。人类往往引以为傲的自身理性,使自身区别于动物而掌握美德的罗盘,然而,掀起“情感革命”的学者们试图证明,个体行动者的道德判断是“情感的狗与理性的尾巴”,^③理性的尾巴显然构不成牵制的力量,需要与情感协调运作来发挥其积极效用。

一个长期未被解决的哲学难题在于,当人们面对两个相似道德困境时,会表现出不一致的反映。比如,在电车难题中更倾向于让电车转而驶向一个人,而在人行桥困境中却难以接受牺牲一个人的生命来拯救其余五个人。格林(Joshua Greene)等对人们在道德困境中作出反应时进行脑部扫描,发现近身伤害行为所带来的冲突,与在有一定空间距离的、非直接接触的困境中大脑的激活区域不同:后者与受控认知相关,从而使人愿意为更多人牺牲一个人的生命;而前者则揭示了情感的参与,导致人们拒绝伤害一个无辜者来拯救其他人。^④这说明,人们面对道德困境时的权衡与选择,可能直接取决于情感之上不假思索的判断,而不是通过公正、理性的推理作出深思熟虑的选择。这种情感反应既可以带来格林等人实验中对伤害无辜者的抵触,也会导致对不曾伤害他人行为的排斥,人们倾向于选择道德化那些冒犯了社会规范但无害的行为。

这一研究表明,道德判断依据的是自身感受,而非是否存在实际伤害后果。对符合不伤害(no harm)规范的行为,人们的反应是伤害这种“无辜者”。道德推理在事后才发生,而直觉的道德判断可能经不起道德推理的审视。^①

既然情感捷径是人们道德思维中节省认知资源的反应,一种情感策略考虑便不再是如何提升人们运用理性进行公正推理的能力,而是注重情感的效用及其可塑性。道德生物改良(moral bioenhancement),包括影响神经递质水平的药物、无创脑刺激与基因技术等,就是利用节省认知资源的情感反应来增加亲社会情绪、抵触伤害行为。如此改良情感的正当性实际来自于对生理疾病及其导致的道德问题的考察。不同于对不道德的“恶”的判定,这类道德问题被解释为非有意的失控,例如前额叶皮层损伤会导致行为与认知背离,使患者作出不顾他人安全的道德越轨行为;特定共情基因的问题导致个体不够宜人的性格乃至残酷的行为,被界定为是一种共情侵蚀(empathy erosion)的神经功能障碍。^②当向来受到贬低的激情与欲望遭受过度节制与毁灭,作为理性机器的个体与社会反而陷入了功能失调,将彻底走向黄金法则的反面,毫无心理负担地对更多人施加伤害。由此,利用道德生物改良,便可以因对人体缺陷的功能性修复以及预防个人出现反社会行为的正当性而受到辩护。

但这些技术其实将黄金法则降为了一种由专业人士定义的固定情感算法,无视社会角色与情境,将一致性要求替换为确定的价值,将公正的理性推理替换为最低风险的安全。现代人的傲慢——即相信历史前进必然伴随着人们道德观念的改善——在这里成为仅对有知者适用的信条。道德生物改良所述美好未来的愿景,建立于其他理性能力被高估、无法正确运用情感的普通民众的可塑性之上。毫无疑问,没有任何权威(即便是作为硬科学的神经科学)能够承担起建立统一规范道德的任务,遑论对其进行改良。

理性推理与情感直觉作为道德判断中的两项关键心理机制,实际需要匹配具体任务而进行调用。直觉反应凸显了人们作为文化性动物的本质,那些来自遗传、文化和个人经验的情绪化行为反射与直觉,亦即“自动模式”的认知过程,它们高效亦节省认知资源,使人们在日常生活中无须反复论证内化的价值信念,或是为行为习惯寻找合适的理由。同时,人们在日常生活中面对的通常是一般性的道德难题,不是“纵使天崩地裂,也要正义昭彰”的那种呼唤抛下任何利益算计与冲破外界压迫性束缚的挑战,因此在社会化进程中习得且默认的概念与规则足以成为指导道德实践的理想资源。直觉的情境敏感性使其发挥着一种“无人监督”的风险警示作用,为认知负荷过载的人提供直接的指引,纠正理性思考中所调用的违背个人深层价值观的社会偏见。

而理性被认为应当在道德决策中占据上风的情境,则可能涉及群际争端、公共决策与陌生的道德议题。直觉的重要性尽管得到承认,但面对存在道德差异与利益竞争的外群体,人们熟练掌握的内群规范显然鼓励的只是对内群同伴的感同身受,拒绝对道德圈外的人们给予关怀,这使得理性反思仍必不可少。^③不过,也有实验发现,置身于要求多角度公正思考的“无知之幕”下,普通人无须特殊训练和激励,便可能倾向于以公正的、最大化福利的思维进行推理,并且克服情感冲突——哪怕可能造成自己或内群成员的牺牲,也为社会治理提供了积极的激励手段。^④面对时下激增的争议性道德问题,尤其是当遇到与道德直觉相冲突的陌生道德议题时,个体应当选择秉持认知无奇迹原则(No Cognitive Miracles Principle),反思自身冲动从何而来,根据有意识的、深思熟虑的推理来进行判断,而非纯然顺从于自己可能被轻易激活的、来自文化与道德部落的直觉。^⑤

① Jonathan Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion*, New York: Vintage, 2012.

② Simon Baron-Cohen, *The Science of Evil: On Empathy and the Origins of Cruelty*, New York: Basic Books, 2012.

③ Joshua D. Greene, *Moral Tribes: Emotion, Reason, and the Gap between Us and Them*, New York: Penguin, 2013.

④ Karen Huang, Joshua D. Greene et al., “Veil-of-Ignorance Reasoning Favors the Greater Good,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 48, 2019.

⑤ Joshua D. Greene, “Beyond Point-and-Shoot Morality: Why Cognitive(Neuro) Science Matters for Ethics,” *Ethics*, vol. 124, no. 4, 2014.



重写社会认同的边界

在长达三年的全球公共卫生危机之后，当今全球秩序尚未完全重回正轨，社会认同跨越不同国家与群体的凝聚力面临着不断冲击。俄乌战争与巴以冲突加剧了地缘对抗格局的紧张态势，高温与污染等气候问题威胁着全球人口健康与生存资源状况，而全球经济复苏缓慢、极端贫困的消除进程也陷入停滞。在皮尤研究中心 2023 年的一项全球态度调查中，大多数公众都感到自己与本国或当地社群的人们关系更为紧密，而与全球其他地区的人们则相对疏远；更多受访者认为自己的国家应当减少对国际事务的参与，优先专注本国问题；当被问及国家在追求自身利益的同时是否需要顾及他国利益，半数受访者将本国福祉置于优先地位，即便这会招致其他国家的强烈反对。

可见，尽管对人类团结采取行动以应对挑战的呼声日益高涨，公众对于世界道德的关注同时在不断削弱。国家内部乃至不同国家之间面临意识形态的强烈分歧，核心是将同理心与道德责任扩展到关系松散乃至遥远广阔的群体，还是仅保护紧密联结的小圈子——它最小收缩至具备亲缘关系的家庭，最大拓展到自己的族群。这使得对黄金法则的讨论，最终都不可避免地回到对道德主体所属群体成员资格的检验。这是因为一致性要求的本质在于，他者作为个体的生存值得道德的对待，而非具备特定群体成员资格才使他们进入道德标准的适用范围之内。

美德是可教的，与此同时，“也许没什么比严肃的美德教育更危险的事”。^①在皮亚杰(Jean Piaget)与科尔伯格(Lawrence Kohlberg)两位学者提出的道德发展阶段论中，儿童与成人之间的道德差异并不能归因于知识多寡或社会化程度高低，而是因为两者处于不同的道德发展阶段。道德发展学说的起点和终点都在于公正地对待自己与他人的黄金法则：幼童可能已经拥有了第六阶段的道德观念，只不过尚处于洞穴内的阴影之中；而在道德水平发展的最高阶段，人们无须设身处地地考虑关于特定社会位置的期望与义务的“具体的金科玉律”，^②普遍化的道德原则已化于心中。而实现这种最高道德水平的方式，则在于一种打破自我中心的道德教育。

在科尔伯格看来，儿童作为哲学家，只应听从最优秀的、唯受自身理性统治的哲人的道德指引，道德不能来自外界直接强加的所谓“新知”以及儿童对给定规则的“内化”。因此，与对长者权威的服从相反，他提倡营造参与式民主氛围、鼓励多元思想的学校教育。德育的基本问题是个人的非中心化，这并非所谓的个人主义向集体主义的复归，实际上人们需要超越的正是围绕国家、种族、阶级而形成的精细的自我中心状态及潜在无意识偏见。通过民主讨论与决策营造积极的道德气氛，培育学生的多元视角与包容的国际态度，使学生意识到道德情境的复杂性与既定标准的矛盾，并自由探索没有现成解决方案的问题，道德教育有望提升他们的道德发展阶段，达到批判性的信息接收与识别，在团体中真正运用个人最高标准的道德推理。^③

认知发展论者对道德发展阶段与道德教育的讨论进一步超越道德与不道德、善与恶的简单二元划分，强调连续、发展和动态变化的进程而非站在单一性与多元性的任何一极，认为个人在社会关系中通过能动性与创造性的发挥，接受经验刺激不断进行修正与重构，探索是否存在某种普适性公理，以及与其他诸多特殊“真理”间是何种联系，^④从而成就完备、独特而多样的自我，不断攀登至更高的道德水平。这种“自我”不能被简单固化为刺激—反射的产物，而是在不同阶段应有所差别。当人们抵达以普适伦理准则为取向的第六阶段，所依据的就是黄金法则这样的抽

① Lawrence Kohlberg, *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*, New York: Routledge, 2013, p. 36.

② L·科尔伯格：《道德发展心理学：道德阶段的本质与确证》，郭本禹等译，上海：华东师范大学出版社，2004 年，第 175 页。

③ 柯尔伯格：《道德教育的哲学》，魏贤超等译，杭州：浙江教育出版社，2000 年，第 262—340 页。

④ 威廉·詹姆斯：《实用主义：某些旧思想方法的新名称》，李步楼译，北京：商务印书馆，2009 年。

象伦理准则，它代表着至上的美德，要求公正、平等、互惠地对待所有人，尊重每个个体的尊严与内在价值。

既然人的道德倾向不只是固有的私人信念，还是外部世界的支持力量、某个社会与特定同胞族群的表现形式，那么不可避免地，“我们”自然也可以通过个人与社会定位的变化来加以调整。因此，另一条路径便是从社会认同论出发，试图展现群体认同搭建更大共同体的潜能。“群体自我”（the group self），即由人们生来所属的性别、种族或国籍所界定的群体身份，或人们对自我选择的特定职业、宗教或政治群体的认同，产生了基于群体成员身份的内群偏好与外群敌意。^①群体认同具有社会疗愈效应（social cure effects），有益于促进个体心理健康，激发个体对自我能动性的感知，而个体亦能对他人的利益视如自身利益般负责，互相提供社会支持。^②而对于经受痛苦的外群成员，人们倾向于选择漠然无视，更遑论给予帮助。因此，研究者提出调整社会分类的策略，期望让亲社会行为的内群身份为更多人所共享，亦即以统摄性的认同将外群纳入内群，重新定位内群积极归因，从而减少群际偏见，发挥群体的积极社会力量。

对于如何调整分类以实现涵盖更大范围的社会认同，使社会成员成为积极的道德行动者，已有学者们提出一些方案。比如，共同内群认同模型（Common Ingroup Identity Model）提出同化与双重认同两种途径，希望通过激活共享的更高层次的成员资格，或使少数群体表征双重身份认同，以达到更具包容性的内群归属和更和谐的群际关系。^③而对道德认同加以概念化并进行测量的研究认为，道德认同由于较其他社会认同分类更具道德性的价值关联，具有使人超越亲缘与地缘关系而关怀更广泛的人类群体的可能。^④道德认同这种围绕一系列道德特质形成的自我概念，能够扩展定义内外群的心理边界，以更高的道德认同搭建更为包容的道德尊重圈。启动道德认同，能够中和其他道德调节机制的负面效应。

社会认同作为一种激励因素，其力量来自唤醒行动者内部群体自我的推动，因而重写社会认同方案试图将“我”与“我们”、“我们”与“他们”之间的张力，化解为“我们”和“我们”的关切。这种努力将简洁性的偏好转向分类本身，而非化简道德本身的复杂性。尽管研究者并未就如何进行自我归类、如何对他人进行分类达成认同共识，它已经不再以一种秩序井然的道德秩序，或者某种普遍适用的道德根基为前提。

简而言之，实现黄金法则的一致性要求，至关重要的一点是实现不同群体独特性基础上承认的联合。

当今世界需要黄金法则

当今世界正前所未有地需要黄金法则。

一方面，全球化背景下文化多样性的发展与种族融合的提升，并未阻止不同群体间仇视情绪的加剧，差异与竞争依旧催化着更多国际冲突的可能。崇高德性的另一面便是与之相伴而生的卑劣的罪恶，道德部落主义仍旧是实现黄金法则所面临的巨大阻碍。对此，中国提出了构建人类命运共同体以弥合群际偏见、跨越“我者”与“他者”道德壁垒的倡议，彰显着消解全球文明冲突的愿景。^⑤方文指出，一种真正处于至高地位的分类范畴，本身就拒斥任何多样而破碎的分类线索，代表着人类命运共同体的成员资格，它呼唤人们为“承认所有人类生命同等禀赋珍贵价值和尊严基础”的本真分类（authentic categorization）而抗争。而从本真分类出发，以年龄为基准的至简分

① Naomi Ellemers, “The Group Self,” *Science*, vol. 336, no. 6083, 2012.

② Katharine H. Greenaway et al., “From ‘We’ to ‘Me’: Group Identification Enhances Perceived Personal Control with Consequences for Health and Well-Being,” *Journal of Personality and Social Psychology*, vol. 109, no. 1, 2015.

③ John F. Dovidio et al., “Understanding Bias toward Latinos: Discrimination, Dimensions of Difference, and Experience of Exclusion,” *Journal of Social Issues*, vol. 66, no. 1, 2010.

④ Karl Aquino et al., “The Self-Importance of Moral Identity,” *Journal of Personality and Social Psychology*, vol. 83, no. 6, 2002.

⑤ 刘德斌：《“文明冲突”的预言与国际关系的演变》，《探索与争鸣》2023年第11期。



① 方文：《认同政治学的本真分类方案》，《当代美国评论》2019年第2期。

② John T Jost, *A Theory of System Justification*, Cambridge: Harvard University Press, 2020.

③ Julian W. Fernando et al., "Functions of Utopia: How Utopian Thinking Motivates Societal Engagement," *Personality and Social Psychology Bulletin*, vol. 44, no. 5, 2018.

④ 约翰·杜威：《杜威全集·晚期著作：1925—1953》第9卷，王新生、朱剑虹、沈诗懿译，上海：华东师范大学出版社，2015年。

类可能是相对公平的行动策略。^①

另一方面，在共同体内部，教育、住房、医疗资源分配的公平性问题仍未解决，维持现状而非改变社会分类的倾向仍占据主流，弱势群体受无意识内群偏见的影响，认为更有才能的优势群体成员理应享有更丰厚的资本。^②弱势群体沦为内群不公的同谋，服从于一小撮权威代理人——无论是宣称代表公众意志的能力卓越的官方机构，或是垄断了艰深科学知识解释权的知识精英，实际都使个人情绪为议程设置所牵引，而非依据更高的理性来规避错误的道德决断。尽管这看似有益于维持系统稳定，实际上不仅会导致歧视、社会排斥与污名的延续，还将使人们减少消除不平等的努力。对此，引导对黄金法则这一古老价值的向往，激发对理想社会德行、智慧和友爱的追求，则有可能激励普通社会成员对公民参与和社会变革持有更开放的态度。^③

此外，数字化转型的趋势不仅正在复制现实的道德难题，也造成了更多不确定性。新的道德主体如人工智能，可能被赋予自主进行道德决策的能力，他们应当依从何种道德标准尚未可知，而虚拟世界如元宇宙的发展，也可能在社会秩序与社会关系的重建中加剧伦理困境。置身于发展和改革的世界道德图景之中，人们通过对新的具体事实的观察与概括，不断地设计更好或更坏的道德，而黄金法则为道德行动者提供了这样一个出发点：对于层出不穷的道德问题，并不必然存在善恶分明的道德主体与非黑即白的答案，个体还需要警惕集体的激情或权威机构的训令，认识到自身是自己可靠的道德领航者。

黄金法则的界定与要求仍是一个充满论辩与抗争的现代议题，也即这种抽象的道德准则，究竟怎样才能真正植根于现实生活，使个体的知与行、个体与他者以及不同群体之间达到最大程度的和合统一？以往研究对个人修养与共同生活准绳的讨论，简要划分出不言自明的道德内容与动机种类，这被对应到实验中有限的困境与游戏设置，再将抽象法则的不确定性转移到道德行动者身上，因为人们总是冒着犯下道德错误的风险。但如果以黄金法则为线索回溯，这些试图穷尽道德问题情境的研究实际扮演了“建设性的否定”角色：尽管人们需要抵御自我所受到的日常限制，但仍应当赋予其自主性以超越某种权威的关键位置，发挥其早在孩提时代便具禀赋的黄金法则的纯粹力量，并且通过扩大社会认同、拓展道德边界而延伸道德义务，自然而然地在群体自我的指导下对广阔群体进行包容性的道德判断。就此而言，塑造超越族群身份的认同方案，相较于对复杂领域中道德多元现象的划分，更具实现黄金法则的优势。

黄金法则所被寄予的期许，在杜威对人们冠以“上帝”之名的概念描述中早已提及，“在一个心烦意乱的时代，急需这样的观念。它能聚合现在涣散的志趣和能量；它能指导行动，而且能产生情绪热度和理智之光”。^④作为在动态过程与具体实在之中被激活和证实的“厚概念”，或者说作为一种在进化挑战、日常生活与社会文化中被定位的过程性语言，未来黄金法则还需要在公共领域下被更多讨论，在至高社会理想的认同视角下，不断探索有助于维系社会共识、提升社会团结与在全球化语境下实践黄金法则的可能路径，而非将黄金法则仅仅停留于严格控制的实验室设计之中。

编辑 李梅高原