

道德判断中的理性与情感研究

郭 晓

(浙江大学 科技与法律研究中心, 浙江 杭州 310008)

摘要: 在道德判断过程中,“理性”成为被过度演绎的神话,情感被过度忽视。情感具有重要的道德意义:道德判断的实际过程中我们高度依赖进化获得的、情感相关的“自动驾驶”模式,情感有助于防止我们成为“理性的傻瓜”。道德推理内在地包含了“合‘理性’”与“‘合理’性”两种模式,前者指逻辑严密性而后者指实践适用性。由于不完全信息、有限理性和决策成本三方面的限制,人类的多数决策都只具有“合理”性而非合“理性”,这对道德判断具有重要影响。

关键词: 道德判断;理性的神话;合“理性”;“合理”性;行动的逻辑;解释的逻辑;实验哲学

中图分类号: B82-054 **文献标识码:** A **文章编号:** 1674-5639 (2020) 01-0078-05

DOI: 10.14091/j.cnki.kmxyxb.2020.01.012

Rationality and Emotion in Moral Judgment

GUO Liang

(Center for Science, Technology and Law, Zhejiang University, Hangzhou, Zhejiang, China 310008)

Abstract: In moral judgment, “rationality” becomes an over-interpreted myth, and emotions are over-ignored. Emotions have important moral significance; in the actual process of moral judgment, we are highly dependent on the “autonomous driving” model obtained by evolution and related to emotions. Moral reasoning inherently includes two types: “rationality” and “reasonability”. The former refers to logical rigor and the latter means practical applicability. Due to the limitations of incomplete information, bounded rationality, and decision-making cost, most human decisions are only “reasonable” rather than “rational”, which has an important impact on moral judgment.

Key words: moral judgment; the myth of rationality; rationality; reasonability; logic of action; logic of interpretation; experimental philosophy

一、“理性”的神话

“理性”(Rationality)经常被哲学家视为一种值得称道的品质。在与道德判断相关的扶手椅哲学(Armchair Philosophy)中,诸如“理性主体的存在是道德上负责的前提条件”之类的命题甚多;某些哲学提议甚至试图剥夺“非理性主体”被人道地对待的权利——请注意,这些“非理性主体”历史上一度包括奴隶、黑人和妇女。直到今天,这种对理性的过度推崇依然在哲学中占有特别的地

位;这迫使我不得不写下如下的警句:

“理性”是哲学中被过度演绎的神话。

神话本身就意味着某种“非理性”。试图刺杀希特勒的那个勇敢的德国哲学家迪特里希·朋霍费尔(Dietrich Bonhoeffer)对人类的理性能力表示了质疑。他指出,“令人震惊的是有理性的人(Die Vernunftigen)的无能为力,他们既没有能力看到邪恶的深渊,也没有能力看到神圣的深渊。他们怀着最善良的意愿,自以为能够借助一点理性,把解体的木骨架重新接合。他们视力不济,却还想公道

收稿日期:2019-06-18

基金项目:国家社科基金青年项目“道德判断的实验哲学研究”(15CZX017);浙江大学争创优秀博士学位论文资助(201502)。

作者简介:郭晓(1986—),男,湖北郧阳人,博士,博士后,主要从事道德判断、实验哲学等研究。

地对待双方，结果在猛烈相接的两股暴力之间被碾碎，落得一事无成。”^[1]事实上，这段话用来描述当代许多互相冲突的政治观点也是合适的。朋霍费尔对人类理性的批判令人印象深刻，他洞察到了“理性”本身的无力。

马特·里德利（Matt Ridley）则从进化的角度指出，“社会并不是由理性的人类创造的，它是随着我们的本性一起进化的结果。和我们的身体是基因的产物一样，社会也是我们基因的产物。……一定要设法走出自身的局限，回顾一下我们这个有着各种缺陷的物种。”^{[2]XII-XIII}马特·里德利继续指出，“正如哥白尼和达尔文一样，威廉姆斯和汉密尔顿给了妄自尊大的人类一记重击，让他们羞愧难当。人类不仅只是动物的一种，而且也是自私基因委员会用完即扔的玩物和工具。”^{[2]9}他清晰地指明了人类的“内在生物性”。政治经济学家阿玛蒂亚·森（Amartya Sen）则将此种目光短浅而又自私自利的人称为“理性的傻子”^{[2]115}。

不过，正如扶手椅哲学家们所强调的，“理性”本身是无害的，有害的只是对理性的滥用。然而，这种滥用如今已经到了一种无以复加的程度，以至于我们本处的批评看起来像是一种情感的宣泄，一种矫枉过正。不过，让理性回归它应在的位置，或许这本身就是一种基本的“理性”。一种“纯粹理性”就像人类心智中没有天敌的外来物种，这一“入侵者”霸占了太多的生态位导致了其他一些有价值的人类能力——例如情感和“实践理性”的濒危。因此加以“清算”也并不为过。

事实上我们的确有——至少现在，有更多的理由，去怀疑人类这种“逻辑理性”的“合理性”。在我博士论文答辩的当天，在中国浙江，人类“理性的丧钟”即将敲响。这倒不是因为我的论文对逻辑理性统治的质疑，而是因为，2.0版的AlphaGo将彻底战胜人类最杰出的棋手柯杰（KE jie），成为“围棋之神”^①；这将在第一个人类引以为傲的、变化超出了 10^{81} 这个宇宙原子数目上限的领域宣告人类“理性的失败”，未来还会有更多的人类“专属”

领域、以更快的速度被占领、统治。与其继续“敝帚自珍”，倒不如给理性以客观的评判、正确的定位，与此前被遮蔽的人类其他天赋——比如“情感”一起各得其所、各安其位为妙。

情感常常被视为“理性”特别是“逻辑理性”的“对立面”，它在责任归因和道德判断中被忽视的程度令人诧异——尽管历史上并不是这个样子的。亚当·弗格森（Adam Ferguson）早就指出，“情感是与被认作善或恶的事物相关的一种精神状态。”^[3]而“善/恶”则是重要的道德价值判断。罗莎琳德·赫斯特豪斯（Rosalind Hursthouse）认为，“一种恰当的情感观念”是这样的：“在‘善’与‘恶’最一般、最基本的意义上，情感包含着关于它们的观念或想象（或想法或理解），把它们当做追求和回避的正式对象。”^{[4]123}大卫·休谟（David Hume）则认为，情感是一种“基本的存在”：“情感是一种原始的存在，或者也可以说是存在的一个变异，并不包含有任何表象的性质，使它成为其他任何存在物或者变异的一个复本。”^{[5]453}情感最重要作用在于“防止我们成为理智的傻瓜”^{[2]105}。亚里士多德主义者也会支持“情感在道德上具有重要意义”的主张，^{[4]120}罗莎琳德·赫斯特豪斯指出，由于亚里士多德关于人类理性的论述使得情感能够成为理性的一部分，“从而在定义完整的美德时发挥了恰当的功能”，亚里士多德及其主义者在情感的道德意义问题上具有超越康德和休谟的优势之处。“亚里士多德论证的优势之处在于他对人类理性的论述，该论述允许情感参与到理性之中，而不仅仅是纯粹动物式的非理性冲动，从而情感可以在详细说明完整美德的过程中扮演它们恰当的角色”^{[4]132-133}。罗伯特·特里弗斯（Robert Trivers）同样指出，“情感在我们内心的精打细算和外在的行为之间居中调停。”特里弗斯说出了关于人类的一个可怕真相：“人类情感更像一个实行互惠原则的社会动物所使用的工具包，经过了精心的修饰。”^[6]哲学家关于情感的认知展示了情感丰富多维的面像。

情感作用的一个重要特点在于其“即时性”。

①正如我们事后所知的，柯洁的确以0-3完败于2017年5月17日的AlphaGo Master。不过很快，AlphaGo Zero在训练3天后以100-0战胜了打败李世石的AlphaGo LEE，用21天战胜了打败柯洁的AlphaGo Master，成为“围棋之神”。正如“棋士柯洁”在新浪微博所感叹的：“一个纯净、纯粹自我学习的Alphago是最强的……对于Alphago的自我进步来讲……人类太多余了。”

阿格妮丝·赫勒 (Agnes Heller) 指出,“某些行动只允许短暂的思考。这是从道德视角评估行动的习惯成为关键的地方。习惯培养着像闪电一样运转的直觉:它照亮预期行动的黑暗角落;道德决定和行动几乎是同时的。”^[7]事实上,在道德判断的过程中,我们的确高度依赖我们人类进化获得的自动驾驶模式——系统 1^① (system 1)。弗兰西斯·哈奇森 (Francis Hutcheson) 指出,“正如每一个人通过密切的观察和反省会使他自己确信的那样,存在着一种认可某些感情和随着这些感情而产生的行为的天赋的、直接的决断;或者说对于行为之中当下显现的卓越存在着一种天赋的感官,而无需涉及任何我们通过其他感官或凭借推理可知觉到的品性。”^{[8]57}他将这种决断称为“感官” (sense) 或“本能” (instinct), 强调了它与情感的密切关系以及直接性。弗兰西斯·哈奇森甚至认为,“德性,正如前面一卷所证明,主要由各种感情所组成。”^{[8]214}这一天才的猜想得到了当代迅猛发展的认知神经科学的不少经验证据的支持。在责任归因和道德判断中,情感快速而直接地“解决问题”,得出结论。

情感还具有对于理性的“调节”作用。情感“改变了履行承诺难题带来的回报,把理性的算计中没有考虑的长远代价都放到目前来考虑。怒火会吓阻那些违背承诺的人;良心谴责会让欺骗他人的人感到痛苦不堪;嫉妒代表着私利;鄙视会赢来尊重;丢面子就是种惩罚;同情心引出互惠互利的同情心。”^{[2]113}情感与美德之间则有着更为广泛和历史悠久的联系。“在特定的情境中体验特定的情感具有内在的道德价值,而不仅是工具价值或其他类型的内在价值。我们可以说,体验该情感之所以有‘内在的道德价值’,就是因为情感体现了美德。……还有另一种方式来表达同样的事实:由于情感反应包含着正确的理性内容,因此,情感具有内在的道德价值。”^{[4]121}事实上,过度强调思辨理性会带来许多不良的后果。例如,仅凭理性思辨难以发现在责任归因过程中“‘部位’与‘责任’的相关度大于‘控制性’”“‘部位’对责任归因的解释力也强于‘控制性’”等重要的实验哲学经验事

实。^[9]换言之,理论的构建难以真正解释和有效预测道德事实。

在理性与情感关系的问题上,大卫·休谟显然颇具远见。休谟的主要观点是:(1) 理性和情感的作用不同;(2) 理性不能作为判别情感的标尺。在其《人性论》(我们最好不要忘记该书副标题是“在精神科学中采用实验推理方法的一个尝试”)中,休谟指出,“反对我们情感的那个原则不能就是理性,而只是在不恰当的意义下被称为理性。……理性是并且也应该是情感的奴隶,除了服务和服从情感之外,再不能有任何其他的职务。”^{[5]453}因为“理性的作用在于发现真或伪”,但“我们的情感、意志和行为是不能有那种符合或不符合关系的;它们是原始的事实或实在,本身圆满自足,并不参照其他的情感、意志和行为。因此,它们就不可能被断定为真或伪的,违反理性或者符合于理性。”^{[5]498-499}理性“在严格的哲学意义下”只有两个方式能影响我们的行为:(1)“它把成为某种情感的确当的对象的某种东西的存在告诉我们,因而刺激起那种情感来”;(2)“它发现出因果的联系,因而给我们提供了发挥某种情感的手段”。因此,情感可以是“活动原则的源泉”而理性则不可能——“理性是完全不活动的。”^{[5]498-499}另一方面休谟又指出,情感“不合理”的条件则只有两种:(1) 建立于一个虚妄的假设上;(2) 选择了不足以达到预定目的的手段时。因此,休谟总结认为,“理性和情感永远不能互相对立,或是争夺对于意志和行为的统治权。”^{[5]454-455}——情感始终“领导”理性。

二、两种“理性”概念:“合‘理性’”与“‘合理’性”

我们一般要求道德推理具有“合理性”,然而即使对“合理性”本身,也存在一种概念上的混淆。^②当单独使用“合理性”一词的时候一般理解

①“系统 1”及与之相对应的“系统 2”是由基思·斯坦诺维奇和理查德·韦斯特率先提出的,随着卡尼曼的《思考,快与慢》这两个术语也逐渐为普通公众所熟悉。

②实际上,道德判断是一个复杂的过程,涉及“原因”与“理由”“解释”与“辩护”等基本概念的区分,以及两种不同的逻辑。

为“合‘理性’”，然而这并非该术语的全部，它也内在地蕴含着“‘合理’性”。本文以“归纳推理”为例展示合“理性”与“合理”性的异同。作为“休谟问题”的核心，归纳推理的“合理性”是一个一直受到广泛关注的有趣问题，不仅哲学学者，其他领域的研究者也对此表现出了极大的兴趣。可以说，归纳推理的合理性问题是一个跨学科的基础性问题，在探询归纳所得结论的有效性时人们难以绕开。在此我想表明，区分合“理性”与“合理”性对这一问题的最终解决是十分有益的。

休谟问题是否有终极解？这取决于我们如何理解“合理性”。事实上，许多争论的发生，都是由于混淆了“合理性”的两个层面所致。这两个层面是：(1) 合“理性”，(2) “合理”性。这一区分既涉及数学归纳与心理归纳问题，也涉及理论上的“确证度”与实践中的“有效性”。一旦我们正确地厘清了这个容易混淆的概念，休谟问题也就有了一种新的可能的解决路径。如果我们追溯归纳推理的辩护历史，容易发现主要的辩护都是从逻辑“可靠性”即“合‘理性’”而非“‘合理’性”角度展开的。由于用于归纳的单称命题和作为规律的全称命题不对称等因素，归纳推理难以在逻辑体系内得到证明，换言之，归纳推理不能严格地合“理性”——对此卡尔·波普尔（Karl Popper）已有著名且令人信服的论证（——并基于此发展出了他一整套证伪主义的“科学发现的逻辑”），汉斯·赖欣巴哈（Hans Reichenbach）也曾指出，就真理合理性而言，归纳法得不到辩护——归纳法是不“合‘理性’”的。中国的逻辑哲学家陈波在《休谟问题和金岳霖的回答》中同样指出，“归纳问题在逻辑上无解”，的确如此。然而，就逻辑体系之外实践层面的合理性而言，归纳法可以获得有效辩护；归纳法具有“‘合理’性”——某种“现实必然性”和可接受性。在逻辑体系之外得到辩护的不是逻辑层面的“合‘理性’”而是实践层面的“‘合理’性”，实质概念不同。我们难以用心理归纳的“合理”性（实践适应性和适用性）来为“归纳推理”的合“理性”（逻辑严密性）作辩护，它们实际是两个概念，只是由于共享名字而易被混淆。^[10]

归纳法或许属于一种特殊的人类心智“适应

器”（Adaptation）——可能是通过自然选择进化而来、用于解决与人类生存相关问题的适配器。对历史时间上人类绝大多数问题而言，归纳推理已经足够了！这一适配器可以有效地、经济地、可靠地解决许多需要根据过往经验预测未来的问题——正如人类语言已经被可靠的证据证明是具备适配器所有特征的“一种卓越超群的适配器”一样。这种进化心理学的视角可以对解决“休谟问题”提供一种有力支撑。归纳法也典型体现了合“理性”与“合理”性各自的特点和冲突。

关于两种“合理性”，另一个著名的案例是“囚徒困境”（prisoner's dilemma）。在经典的博弈案例囚徒困境中，合“理性”与“合理”性之间的差异得到了充分展示。以某一玩家A为例，如果他的搭档B没有认罪，那么A的出卖行为就可以让A无罪释放并且获得一小笔奖金；反过来如果A的搭档B出卖了A，那么A的最佳选择仍然是出卖他的搭档B，否则可能遭受更严厉的惩罚。“总之，尽管双方进行相互合作确实是最佳的选择，但是不管搭档选择哪一种行为，两个人的合理决策都应该是招供和背叛。”^[11]（黑体为笔者所强调）“最佳的选择”是合“理性”的选择，而“合理决策”只是具有“合理”性的选择，是缺乏“上帝视角”人类个体的最佳决策。

有理由相信，人类生活中的大多数决策可能都只具有“合理”性而非合“理性”。主要原因有三：(1) 多数情况下我们处于不完全信息（incomplete information）状态，难以做出最优决策；(2) 由于人们的理性是有限的，即使面对完全信息我们也经常做出“不完全决策”。例如，由于“启发式”的存在，人们在决策时倾向于使用比较容易评价的线索来辅助判断——正如有些审稿专家和读者可能首先依据文本的参考文献数量、范围和质量来评估文本的质量，然后才考察内容本身，这就是“启发式”。两种典型的启发式是“代表性启发式”（representativeness heuristic）和“可得性启发式”（availability heuristic）；(3) 作为最优选择的合“理性”也需要成本，当决策的代价甚至大于其所带来的收益时，最合“理性”的决策就是放弃对这样一种合“理性”本身的执念。比如在著名的“布里丹之驴”（buridan's ass）情境中。这头最终

被饿死的可怜驴子显然也是一头被合“理性”误导了的“哲学标准驴”(philosophical standard ass, PSA)^{①[12]},因为它完全可以“合理”性地选择面前的任何一垛草^②——先吃饱了再管哪一堆草更好吃留作增肥的宵夜,反正都是自己的!最好不要忘记,没有任何选择是无需代价的!有时为了“最佳决策”所付出的代价本身可能超过标的价值;在工商管理中,决策成本(cost of decision making)早已有了成熟的计量方案,难以理解伦理学中还有“唯理性论”的“霸占”。

人类的“理性病毒”跨物种传播,感染了一头无辜的驴子,我对此深感痛心。不过鉴于走兽并不像认死理的哲学家那么傻,可以推测“理性病”并非一种人畜共患病(zoonose),因而应该并没有因“理性病毒”致死的动物,因此也就私心窃慰。

三、作为事后建构的“理由”

那么,“理性”与道德推理又是一种什么关系呢?乔纳森·戴维·海特(Jonathan David Haidt)通过实证指出,“道德推理通常只是一种事后的建构,产生于已经达成的判断之后”^[13]。他据此提出了与“理性模型”对立的一个富有生命力和解释力的“社会直觉模型”来替代传统滥用理性的道德推理模型;“社会直觉模型”强调社会和文化的影响而不再强调个体的“理性”推理;其事实基础在于,道德判断基于快速、自动化的评价——“道德直觉”。

那种“滥用理性”的倾向在“责任归因”和“道德判断”的过程中也反复出现,是一个需要正视的问题。一个潜在的原因可能是,当我们讲“责任归因”的时候,更多地是在进行一种社会性活动:我们试图达成一致或者说服对方,此时我们是“旁观者”而非“行动者”;旁观者需要一种能够实现“主体间一致性”的方案——“说理”,但

说理的“理由”对行动者而言可能并不是必需的——理由往往是奢侈的。二者遵从各自不同的逻辑——“行动的逻辑”与“解释的逻辑”。因此,当我们以“事后”的标准来回顾时,“行动者”看起来就显得莫名其妙甚至格格不入了。

这就是道德判断中与“理性”有关的一个真相。

[参考文献]

- [1] 迪特里希·朋霍费尔. 伦理学 [M]. 胡其鼎, 译. 上海: 上海人民出版社, 2007: 73.
- [2] 马特·里德利. 美德的起源: 人类本能与协作的进化 [M]. 吴礼敬, 译. 北京: 机械工业出版社, 2015.
- [3] 亚当·弗格森. 道德哲学原理 [M]. 孙飞宇, 田耕, 译. 上海: 上海人民出版社, 2005: 31.
- [4] 罗莎琳德·赫斯特豪斯. 美德伦理学 [M]. 李义天, 译. 南京: 译林出版社, 2016.
- [5] 大卫·休谟. 人性论 [M]. 石碧球, 译. 北京: 中国社会科学出版社, 2009.
- [6] TRIVERS R L. The Evolution of Reciprocal Altruism [J]. Quarterly Review of Biology, 1971 (1): 35-57.
- [7] 阿格妮丝·赫勒. 道德哲学 [M]. 王秀敏, 译. 北京: 北京大学出版社, 2003: 199.
- [8] 弗兰西斯·哈奇森. 道德哲学体系: 上册 [M]. 江畅, 舒红跃, 宋伟, 译. 杭州: 浙江大学出版社, 2010.
- [9] 郭晓. “选择的可能”与“控制的祛魅”——一个实验哲学报告 [J]. 浙江学刊, 2019 (2): 70-76.
- [10] 郭晓. 辨析归纳推理的“合理性” [N]. 中国社会科学报, 2015-06-29 (5).
- [11] 巴斯 D M. 进化心理学 [M]. 熊哲宏, 张勇, 晏倩, 译. 上海: 华东师范大学出版社, 2007: 289.
- [12] 郭晓. 责任归因的实验哲学研究 [D]. 杭州: 浙江大学, 2017.
- [13] HAIDT J. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment [J]. Psychological Review, 2001, 108 (4): 814-834.

①笔者构造了“哲学标准驴”以及与之相应的“哲学标准驴”概念,以期清晰地展示部分经典哲学观点的理论预设。

②布里丹之驴的最初版本是: 这头可怜的驴在等距离的一垛干草和一桶水之间饥渴至死(真残忍!), 亚里士多德也曾“使人”同等地饥渴以强调“理性”。对此驴而言,这可不是“中道”的理性,这简直是有毒的理性。