



论康德对罗尔斯正义理论的影响

——兼谈哲学史对于当代哲学研究的启发意义

杨云飞

摘要: 康德的道德哲学对罗尔斯的正义理论具有重要影响。首先,罗尔斯证明正义原则的纯粹程序正义方法正是对康德定言命令程序的运用与发挥。其次,罗尔斯正义论的人性论基础,即自由平等的人的理念,来源于康德通过定言命令的目的公式和自律公式所表达出来的理性存在者的概念。最后,罗尔斯正当与善一致的论证以康德德福一致的论证为原型,这两个论证的结构体现了康德与罗尔斯所共享的正当优先于善的理论模式。康德对罗尔斯的深刻影响表明:哲学史作为启发当代哲学思考的源泉,具有重要意义。

关键词: 作为公平的正义; 定言命令; 自由平等的理性存在者; 正当优先于善

罗尔斯的正义理论是20世纪政治哲学的重要成就,其主要思想来源是康德的道德哲学。作为一名自觉的康德主义者,罗尔斯在不同的场合谈到过康德对自己的影响,这一点也被罗尔斯的学生和后来的研究者反复地论述过了^①。但是在一个简短的篇幅中比较系统地讨论康德对罗尔斯正义理论的影响的工作仍不多见。我相信,这样一种总结性的工作是必要的。本文将系统地论述康德的道德哲学对罗尔斯正义理论的影响。这种影响体现在论证方法、基本概念和理论模式三个方面。借助康德对罗尔斯的深刻影响这一个案,我希望不仅体现出康德哲学在当代的生命力,而且能在一定程度上展现出哲学史所具有的激发新思想的意义。

一、定言命令程序与纯粹的程序正义

罗尔斯的正义原则包括两个,其基本表述如下:“第一个原则:每一个人对与所有人所拥有的最广泛平等的基本自由体系相容的类似自由体系都应有一种平等的权利。第二个原则:社会和经济的不平等应这样安排,使它们:①在与正义的储存原则一致的情况下,适合于最少受惠者的最大利益;并且②依系于在机会公平平等的条件下职务和地位向所有人开放。”^②通常第一原则被简单地称为“平等的自由原则”,而第二原则的两个部分则分

^① 罗尔斯本人的陈述集中体现在他的《正义论》(何怀宏等译,中国社会科学出版社2009年)第40节、两篇重要论文“Kantian Constructivism in Moral Theory”和“Themes in Kant’s Moral Philosophy”(均收录于 *Collected Papers*, ed. by Samuel Freeman, Harvard University Press 2000)以及 *Lectures on the History of Moral Philosophy* (ed. by Barbara Herman, Harvard University Press 2000)中的康德讲座部分。罗尔斯几个弟子的总结工作也值得注意,比如 Onora O’Neill 的论文“Constructivism in Kant and Rawls” (*The Cambridge Companion to John Rawls*, ed. by Samuel Freeman, Cambridge University Press 2003, pp. 347~367)和 Thomas Pogge 的著作 *John Rawls* (Oxford University Press 2007)中的相关论述。国内也有不少学者做了这方面的工作,代表性的论文有丛占修的《康德道德哲学对罗尔斯正义理论的影响》(载《道德与文明》2011年第2期,第54~57页)和乔洪武等人的《罗尔斯对康德伦理学的建构主义诠释》(载《哲学研究》2011年第11期,第98~104页)等。罗尔斯的论著是本文最主要的引证对象,为简化起见,凡引上述论著,笔者将只列出作者、名称和页码,不再注明具体出版信息。

^② 罗尔斯:《正义论》,第237页。

别被称为“差别原则”和“公平的机会平等原则”。罗尔斯认为,在诸原则之间存在着“词典式的”先后次序,即第一原则优先于第二原则,而第二原则中的机会公平原则优先于差别原则。对这两个著名的正义原则,罗尔斯是在一种假设状态即原初状态中,并借助无知之幕的设置来加以证明的。其基本构想是:假设所有人在一种公平的初始条件下,为了更好地进行社会合作,理性地协议以分配彼此的权益与负担,最后以一致同意的方式建立基本原则。罗尔斯认为,在无知之幕的设置下,人们不清楚自己个人的和社会的具体信息,只能运用最大最小值的规则进行推理,从而对他们来说最合理的选择就是以两个正义原则作为调节社会基本结构的规则。这是罗尔斯正义理论的核心部分。

罗尔斯用“作为公平的正义”这个术语来描述自己的正义观。这意味着:在公平的初始条件下以一致同意的方式达成的契约具有正义性。鉴于正义原则来自理性人之间的契约,罗尔斯声称自己的正义理论是卢梭等人的社会契约论更抽象的形式。这种论证方式满足了纯粹的程序正义的要求:只要一种合理的程序被恰当地执行了,由此产生的任何结果都是正义的。在一篇著名的论文中,罗尔斯也把自己的正义论称作“康德式的建构主义”,其含义是:通过一种建构程序建立从特定的人格概念到正义原则之间的恰当联系^①。通过“康德式的”这个定语,罗尔斯清晰地表达了自己的正义理论的来源与特点。简言之,罗尔斯的正义理论来自对康德定言命令学说的阐释与发挥。他的建构主义的两个内核,建构程序与特定的人格概念,正是来自定言命令几个主要公式的变形与运用。

在此,先简单地回顾一下康德定言命令的基本内容是有必要的。康德对定言命令的各种表述主要出自《道德形而上学奠基》一书,具体包括如下几个公式:(1)普遍法则公式:你要仅仅按照你同时也能够愿意它成为一条普遍法则的那个准则去行动;(2)自然法则公式:你要这样行动,就像你行动的准则应当通过你的意志成为普遍的自然法则一样;(3)人性公式:你要这样行动,把不论是你的人格中的人性,还是任何其他的人格中的人性,任何时候都同时用作目的,而决不能只是用作手段;(4)自律公式:只是这样去行动,这个意志能够通过其准则把自己同时看作普遍立法的;以及(5)由自律公式而衍生出来的目的王国公式:任何一个理性存在者都必须这样行动,就好像它通过自己的准则任何时候都是普遍的目的王国中的一个立法成员一样^②。

在此,我无意于讨论诸公式之间的关系这一个在国内外康德研究界争议很大的问题,我将主要关注罗尔斯对上述公式的解读,并说明罗尔斯的正义理论或建构主义的基本内容正是源自这些公式的义理;而且,我的着眼点只是康德如何影响了罗尔斯的正义理论,所以,我对罗尔斯的康德解释的重构将仅以揭示这一点为目的。罗尔斯的康德解释有一个很大的特点,那就是区分了道德法则、定言命令和定言命令的程序。一般的研究者通常只注意到了康德自己明确阐释过的道德法则与定言命令的区分,即道德法则适用于一切理性存在者,而定言命令只适用于人类这样的有限的理性存在者。罗尔斯则进一步严格区分了定言命令本身和定言命令的程序:前者指普遍法则公式这个一般的表达形式,后者指自然法则公式、人性公式和自律公式所体现出来的道德检验程序。定言命令程序是一种检验一个行动的主观准则是否能够成为一种普适道德原则的机制。按罗尔斯的分析,这种机制包括四个步骤。借助于对自然法则公式的阐释,罗尔斯详尽地解释了这些步骤。

按照罗尔斯的分析,自然法则公式所体现的定言命令程序包括四个步骤:(1)对一个理性、合理而真诚的行动者而言,他的行动准则如下:在条件C之下为了产生事态Y,我得去做行动X;(2)把上述准则加以普遍化:在条件C之下为了产生Y,每个人都得去做X;(3)把第二步中的规范转化为自然法则:在条件C之下为了产生Y,每个人(就像遵照一条自然法则一样)总是做X;(4)我们把第三步中的自然法则加到现存的诸自然法则之上,并尽可能测算,一旦这个新加入的自然法则能够展现出自己的效应,自然的秩序将会是怎么样^③。这个新产生的自然秩序就是与步骤1中的准则相联系的一个社会世界,即按照该准则进行了调整的社会世界。我们是否应当按照该准则来行动,取决于两个条件是否得到满

^①John Rawls. *Kantian Constructivism in Moral Theory*, p. 304.

^②Immanuel Kant. *Grundlegung zur Metaphysik der Sitten*, *Kants Werke, Band IV*. Berlin: Walter de Gruyter 1968, pp. 421~438.

^③John Rawls. *Themes in Kant's Moral Philosophy*, pp. 499~500.

足:首先,当我们把自己看做上述社会世界的一员时,该准则是否是可执行的;其次,假如我们属于该世界,我们必须能够想要并且认可这样一个世界^①。如果我们不能同时满足这两条件,那么我们就应该采纳这个准则。哪怕对于我们来说,这个准则是有利可图的。

借助对自然法则公式的程序性解释,罗尔斯发掘出了康德定言命令学说中形式的要素:即任何行为的准则必须符合普遍性或一贯性的要求。凡不能通过普遍性检验的准则,即无法被所有人遵循,或者由此而来的世界不被我们认可的准则,必须被排除。以康德举过的那个著名的例子来说,为了借到钱而在明知自己不打算偿还的情况下做出还钱的虚假承诺,之所以在道德上不可接受,正是因为无法通过定言命令程序的检验。在把为了获益而说谎的准则按照罗尔斯上述的四个步骤进行考察后,我们会发现,在那个所有人都按此准则行动的社会里,一旦这种准则的效应充分地实现出来后,就不会有任何人相信其他人的承诺,从而使得该准则是自我取消的。

在罗尔斯看来,上述四个步骤的程序性解释不仅适用于自然法则公式,同样适用于人性公式与自律公式,这三个公式都是定言命令程序,它们之间并没有客观的差别。当然,定言命令的程序性功能,在自然法则公式上体现得最为鲜明^②。这些公式作为定言命令的程序而言,仍具有主观的差别。具体地说,三个公式采用了不同的视角来看定言命令程序,从不同角度体现出检验的功能:自然法则公式是从我们作为服从道德要求的行动者的角度来看我们的准则是否在道德上可被允许;人性公式把一切人性看做受我们的行动影响的受动者,表明人性本身必须被当做目的来看待;自律公式则回到了行动者的视角,但不是作为服从法则的、而是自己建立普遍法则的行动者,或目的王国中的立法者^③。

借助这种程序性的检验机制,我们主观的、不符合道德要求的欲求和目的就受到了限制。通过这种限制,我们的行动准则所体现的是我们的理性的一贯性。这个程序的特点在于考虑了我们的处境,即人类生活的条件和我们作为有各种需求的自然存在者的情况,并使之受到道德原则的约束^④。罗尔斯对独立而普适的定言命令程序的解释产生了这样的效应:借助一种单纯的程序或形式、而无需采纳任何实质性的观念或标准,就可以判定结果是否合乎道德的要求。这就是说:有一种正确的程序,按照这种程序,可以产生出正当的结果。这正是罗尔斯所阐发的纯粹程序正义方法的精髓。

如前所述,纯粹的程序正义在罗尔斯那里主要指从原初状态出发的论证:“原初状态的观念旨在建立一种公平的程序,以使任何被一致同意的原则都将是正义的。”^⑤显然,康德的定言命令程序起着类似的作用。这两者的相似性说明罗尔斯在方法问题上直接受到了康德的启发。借助纯粹程序正义的观念,罗尔斯把传统契约论提高到更抽象的水平。原初契约的目标是适用于社会基本结构的正义原则,这些原则是理性人会在一种公平的原初状态中一致同意的。正义论由此成为广义的合理选择理论的一个部分。罗尔斯认为,契约理论的程序为正义观念的比较研究提供了一种普遍的分析方法:我们也可用这种方法来证明功利主义的正义观^⑥。现在,我们可以很好地理解罗尔斯为什么要说,康德把道德哲学变成了一门研究合理选择的原则及结果的学问^⑦,正是康德定言命令程序启发了罗尔斯建立一种程序性方法来讨论正义问题。

罗尔斯提出,他把康德那里个人选择的原则扩展为集体选择的原则,并认为这是他与康德的一个主要差异^⑧。我认为,这个差异也许没有罗尔斯设想的那么明显,因为在康德那里已经包含了罗尔斯所要表达的东西。康德的普遍化检验无疑有着集体选择的意味:要站在全社会的立场来看某个准则是否能成为普遍原则。此外,康德还在其他场合明确地表达过这样的意思:真正理性的思考方式不仅仅要从自

① John Rawls. *Themes in Kant's Moral Philosophy*, p. 500.

② John Rawls. *Lectures on the History of Moral Philosophy*, p. 183.

③ John Rawls. *Lectures on the History of Moral Philosophy*, p. 183.

④ John Rawls. *Themes in Kant's Moral Philosophy*, p. 498

⑤ 罗尔斯:《正义论》,第105页。

⑥ 罗尔斯:《正义论》,第93~94页。罗尔斯用契约论方法证明平均功利原则的论证在《正义论》第28节,中译本第125~128页。

⑦ 罗尔斯:《正义论》,第198页。

⑧ 罗尔斯:《正义论》,第202页。

己角度思考,还要站在他人的位置上从普遍的角度思考,最终才能始终一贯地思考^①。这足以表明,在康德那里,个人选择与集体选择是一致的。罗尔斯的贡献毋宁说是:用一种契约程序的方式明确地把这建立为一种道德推理的模式。

二、自由平等的理性存在者的概念:从康德到罗尔斯

按照罗尔斯对于建构主义的界定,除了一种建构程序,还包括一种特定的人的观念:自由平等的理性存在者的概念。只有借助于这样一种自我理解,原初状态下的人才会选择平等的自由原则为最优先的正义规则。平等的自由原则真正地体现了人的这一本性。不仅如此,差别原则所包含的互惠的观念,也正是基于人的平等才是可确证的。本节要说明的是:罗尔斯对人的这种基本理解,完全来自康德对于理性存在者的规定。从原则上说,康德与罗尔斯的理性存在者的概念是一致的。

在罗尔斯看来,康德那里的自由平等的理性存在者的概念主要表现在定言命令的人性公式、自律公式以及实践理性的双重含义之中。让我们从人性公式开始:“你要这样行动,把不论是你的人格中的人性,还是任何其他的人格中的人性,任何时候都同时用作目的,而决不能只是用作手段。”^②人性公式表达了同等地对待所有人“人格中的人性”的道德要求。罗尔斯把“人格中的人性”理解为:处于自然之中的理性人的各种能力与力量,既包括使人们拥有善良意志和品格成为可能的道德人格力量,也包括由文化(艺术与科学)发达而产生的能力与技巧,即自然的能力^③。这意味着,按照人性公式的要求,我们应尊重其他人在自然能力的和道德禀赋双重意义上自我实现的要求。每个人的自我实现要求具有同等的权利,没有任何理由去剥夺这种权利。即便以社会总体利益的名义,也不能牺牲他人的基本权益。

再看自律公式:“只是这样去行动,这个意志能够通过其准则把自己同时看作普遍立法的。”^④自律公式的要害在于自我立法,即只服从自己建立的普遍法则。这种自我立法恰好体现了人的自由,因为人们所服从的法则并非外在的,而是自己基于理性的一贯性而建立的。如同很多研究者所提出的那样,罗尔斯也提醒读者康德的这样一种自由观念来源于卢梭。罗尔斯认为:康德的主要目标是加深和证明卢梭的观点,即自由就是按照我们给予自己的法则而行动,而这导致的是一种互尊和自尊的理论学^⑤。罗尔斯对康德的解释无疑是准确的。无论是康德在《道德形而上学奠基》中区分自律与他律的原则,即前者是按照内在的自我立法的一贯性而行动,而后者受外在的对象决定;还是在《实践理性批判》中对形式原则与质料原则的划分,即自律原则是形式法则,而他律的原则都是追求幸福的质料性的原则,都显示出他的确是从自我立法的角度来理解自由的。

自律原则在康德的道德哲学中是纯粹实践理性的最高原则,是纯粹实践理性自身一贯性与立法资格的体现。除了这种作为普遍的自我立法机能的纯粹实践理性,康德还区分出了一般的(经验性的)实践理性,后者指为实现某个一般欲求目标而进行理性计算的机能,相当于工具性的理性。两者的差别在于:纯粹实践理性只以自身的一贯性为目的,只涉及准则的合法则性,一般的实践理性则致力于获得某个外在对象或实现某种外在目的。康德认为,前者的最终目标是道德,后者追求的则是幸福。

按照罗尔斯的重构,如果我们把康德的自由平等的理性存在者的概念与定言命令程序结合起来,就得到了康德的道德建构主义。罗尔斯这样描述这种建构主义的本质特征:“正当和正义的诸第一原则被看做由一个建构的程序(定言命令程序)所规定的,而该程序的形式与结构反映了(mirrors)我们既理性又合理的、自由的道德人格。”^⑥罗尔斯认为,在康德的建构主义中,被建构起来的是具体的道德规则,即各种具体的义务,如《道德形而上学》中提出来的各种法权义务和德行义务,建构者是定言命令程序,程

① 康德:《实用人类学》,邓晓芒译,上海人民出版社2002年,第94页;还可参看康德:《判断力批判》,邓晓芒译,杨祖陶校,人民出版社2002年,第136~137页。

② Immanuel Kant, *Grundlegung zur Metaphysik der Sitten*, p. 429.

③ John Rawls, *Lectures on the History of Moral Philosophy*, pp. 187~188.

④ Immanuel Kant, *Grundlegung zur Metaphysik der Sitten*, p. 434.

⑤ 罗尔斯:《正义论》,201页。

⑥ John Rawls, *Themes in Kant's Moral Philosophy*, p. 512.

序本身是展现出来而非被建构出来的,但程序有其基础,即自由平等的道德人的概念^①。

罗尔斯的理性存在者的概念和相应的建构主义,几乎完全再现了康德的相应理论。罗尔斯反复援引康德人性公式来说明平等和互惠的概念。把人当做目的,就是按照人们会在原初状态中同意的原则来对待他们,或者说道德人具有平等的代表权;这为人们之间的互惠和相互尊重奠定了基础,因为没有人会同意自己仅仅被当做手段^②。按照罗尔斯这种契约论的解释,我们对于从人性公式出发所确证的人的平等有了更加清晰的理解。对于理性存在者的自由性状,罗尔斯同样发挥了康德的自律概念:当一个人自律地行动时,表现了这个人真正的自由,实现了人的本性。在罗尔斯的语境中,按照两个正义原则来行动,恰恰体现了人们的自由,因为他们行动的原则独立于任何自然的、社会的偶然因素^③。

与康德区分纯粹实践理性与一般实践理性相对应,罗尔斯提出了理性(reason)与合理性(rationality)两个概念,在晚期的著作中罗尔斯也称之为两种道德能力(two moral powers)。前者指一种拥有正义感的能力,即理解、应用和践行(而不是仅仅服从)政治正义原则的能力,而这些政治正义的原则规定了公平的社会合作条款;后者指一种拥有善观念的能力,即拥有、修正和合理地追求善观念的能力^④。这两种能力被罗尔斯看做是任何一个公民所拥有的最低限度的道德能力,由此他们才可能从事社会合作,作为平等的公民参与社会生活^⑤。

现在我们有了罗尔斯建构主义的另一个要素:自由平等理性而又合理的人的概念。把这一概念与前述的纯粹程序正义相结合,就得到了罗尔斯自己的建构主义:由一种建构程序和一种特定的人格概念相结合而产生出正义的第一原则。具体地说,就是从原初状态出发的程序性协议机制,从自由平等的理性存在者的概念出发,建构起两个正义原则。我们可以清晰地看到康德与罗尔斯的建构主义在结构和基础观念方面的相似性。就这样,借助于对康德定言命令学说的阐释与发挥,罗尔斯建立了自己的正义理论的内核:正义原则及其证明,或者说康德式的建构主义。罗尔斯建构主义的方法与观念正好对应着定言命令的程序与含义。在此意义上,罗尔斯的正义论即是对康德定言命令学说的合理发展;反过来说,康德的定言命令学说就是罗尔斯正义论的原型。

当然,我们不能忽略两种建构主义的差别。对此,罗尔斯在《政治自由主义》中有着经典的论述:康德的建构主义是统合性的即适用于包含政治活动在一切生活领域,其自律概念是通过实践理性而自我构成的,建基于先验的观念论这一形而上学预设,其最终目标是建立理论理性与实践理性的统一性;而罗尔斯的建构主义仅仅是政治的,相应的自律概念也只关联着政治价值秩序,不关涉任何形而上学理论,目标只是揭示正义问题证明的公共基础^⑥。可见,两种建构主义最主要的差异在于罗尔斯把建构主义的理论基础与运用范围均限制在了政治领域。

我想进一步说明的是:这些差异植根于康德与罗尔斯所阐释的理性概念的不同含义。康德那里的纯粹实践理性,像罗尔斯说的,具有自我构成的特点,只关注自我立法的一贯性,从根本上说只具有纯形式的规定。这种形式性使理性的原则普通适用于一切实践领域。而罗尔斯的理性概念或者说正义感的能力虽说沿袭了康德式的一贯性的意义,却还有着内容上的规定,即任何一个理性人必然会需要的以自由权利、机会和财富为最主要内容的基本社会善。这些内容原则上与政治生活中各种权益的分配相应。这是罗尔斯对于康德很重要的一个偏离。

最后值得一提的是:康德在目的王国公式中所提到的目的王国,作为道德理想,构成了罗尔斯心目中理想社会即良序社会的原型。结合这一点,我们可以充分地理解罗尔斯自己的总结了:“原初状态可以被看成是在经验论的框架内对康德的自律概念和定言命令的一个程序性解释。调节目的王国的原则也就是将在原初状态中被选择的原则,而且,对原初状态的描述使我们能够解释这样一种意义:即按照

① John Rawls. *Lectures on the History of Moral Philosophy*, pp. 24~327

② 罗尔斯:《正义论》,第138~139页。

③ 罗尔斯:《正义论》,第198页。

④ 罗尔斯:《作为公平的正义——正义新论》,姚大志译,上海三联书店2002年,第31页。

⑤ 罗尔斯:《作为公平的正义——正义新论》,第33页。

⑥ John Rawls. *Political Liberalism*. (expanded edition). Columbia University Press 1996, pp. 99~101.

哲学原则去行动表现了我们作为自由、平等的理性人的本质。”^①

三、罗尔斯正义理论的论证结构与理论模式

罗尔斯的正义理论,除了其内核,即正义原则及其证明,还有一个很重要的部分,那就是关于善的理论。特别是罗尔斯对于正当与善的一致性的论证,是其正义理论不可忽视的一个重要部分。所以,我们还应将目光投向《正义论》第三编以“目的”为名称的内容。在那里,罗尔斯对于他的正义观及相应的正义社会是否稳定的问题进行了论述。稳定性问题指的是作为公平的正义是否能够产生出足够的自我支持的力量。罗尔斯也把这称为“承诺的强度”的问题,即人们是否能维系这样的契约。这本身也构成对两个正义原则的非正式证明。在本节中,我想表明的是:罗尔斯正当与善一致性的论证,其论证结构的原型就是康德对德行与幸福相一致的论证;这种结构体现了康德式的正当优先于善的理论模式。

让我们从康德的论证开始。康德在建立了道德法则的有效性之后,首先遇到的是如何激发人们的道德动机的问题。这是因为,道德价值的本质固然取决于由纯粹实践理性的规律即道德律直接规定意志,但是人毕竟不是纯然理性的存在者,而是有感性的一面,所以仍需要借助某种情感才能真正地驱动人的行动。康德提出,这种情感就是对道德法则的敬重感。道德法则本身会在人身上唤起一种对道德法则的纯粹性和崇高性表示敬重的情感,这种情感吸引着我们出于义务而行动。

在说明了感性的动机之后,道德的性质算是完全确立了,康德现在面对的是德行与幸福不一致的问题。按照道德律的要求去行动的人是有德行的,德行是我们能实现的至上的善,但这对于人来说还不是完满的善。这是因为,人作为有限的理性存在者,具有各种感性的欲求,渴望追求幸福。幸福是人的自然目的。只有德行与幸福合乎比例的结合,才是圆满的善,康德称之为至善。问题在于:从追求幸福产生出德行是不可能的,因为幸福在于各种爱好得到满足,这只会导致意志的他律;反过来,从德行意向产生幸福,似乎同样是不同能的,因为德行在于独立于各种爱好而按照自律的法则行动。如何解决德行与幸福不一致这个难题?康德提出要立足于现象与本体的区分,沿着从德行到幸福的道路,把德行作为至善的首要条件,而把幸福作为随之必然产生出来的结果,有可能实现德福一致。康德还提出了三个纯粹实践理性的悬设来保证幸福与德行合比例的结合。这三个悬设是:意志自由、灵魂不朽和上帝存有,分别作为道德之所以可能的条件、在生命延续中达到道德完善的条件和使德行与幸福按比例结合的条件。康德论证的关键在于:确立德行对于幸福的优先性,在德行的前提下实现至善。

接着来看罗尔斯对于正当与善的一致性的论证^②。我认为,罗尔斯对正义观的稳定性论证包括两个部分:第一,作为公平的正义这一正义观会产生相应的正义感;第二,按正义感的倾向(即按照正义原则去行动的有效欲望)去行动,是和个人的善一致的,或者说与个人生活计划是一致的。第一部分是第二部分的前提,合起来构成正义观自我支持的完整论证:正当与善是一致的。对于这个论证,首先值得注意的两点是:第一,罗尔斯谈的是良序社会、即所有人公共地按照某种正义原则行动的社会中的一致性问题;第二,罗尔斯设定的是一个比较弱的论证目标:证明两个正义原则比其他正义观,尤其是比功利主义,更具有稳定性。只要证明了两个正义原则在稳定性方面更具有优势,论证就完成了。

罗尔斯论证的第一个步骤是通过道德心理学来完成的。罗尔斯提出,人的道德发展有三个阶段:权威的道德、社团的道德与原则的道德。与之相应的三条心理学法则是:(1)通过正当的家庭生活可使孩子对关爱他的父母产生出爱;(2)通过恰当的社团生活,人们会发展出对他人的友爱与信任的关系;(3)在一种正义的社会制度下,当一个人认识到自己和所爱的人是制度的受益者时,他会产生相应的正义感^③。按照这些心理法则,罗尔斯认为,他的两个正义原则,与功利主义相比,更能激发出人们的正义感。这是因为,基于契约的制度更加关心所有人的善,保障着每个人的平等自由,保证我们的权利主张不会因总体善的名义而被忽略或践踏。作为公平的正义体现了互惠的社会合作原则,给予了人们的自

^① 罗尔斯:《正义论》,第201~202页。

^② 对这个论证,姚大志教授作了非常深入和细致的研究工作,可参看他的著作《罗尔斯》(长春出版社2011年)第161页及以下。

^③ 罗尔斯:《正义论》,第429~430页。

尊以较大的支持^①。一个公开承认功利原则的社会,却总会使某些人(弱势者)的自尊受到伤害。所以,按两个正义原则建立起来的社会更易于使其公民产生正义感。

罗尔斯论证的第二部分是要证明:按照正义感行动与追求个人的善相一致。尤其是对于那些企图搭便车的人,一致性的问题变得格外突出。罗尔斯需要证明的是:对这样的人来说,按照正义原则去行动也是合理的。罗尔斯采用了三个论据来证明这一点^②。第一个理由是:在一个良序社会中,正义原则是公共的,那些企图搭便车的人不仅会为此付出很高的心理代价,而且无法确定自己的欺骗策略是不是会伤害到亲近的人,所以会更倾向于在各种事务上保持正义感。第二,按照运用越复杂的能力完成越精致的活动会带来越多快乐的亚里士多德主义原则,在一个良序社会中生活是一种极大的善,因为通过正义感的协调而更和谐地与他人合作,可以使人们更好地发展自己的潜能,享受丰富的生活。第三,最重要的是,正义的行动体现了我们作为自由平等的理性存在者的本性。按照正义原则行动的欲望,与表达我们自由本性的欲望,在实践上是一回事。基于上述理由,无论是我们为了获得更好的生活,还是为了实现自己的本性,我们都倾向于按照正义感的要求来调节自己的目标。

罗尔斯借助正义感的产生和按照正义感行动是善的这两个步骤的论证,证明了正当与善的一致性。在整个论证中,罗尔斯的基本立场是:正义原则是先决条件,始终具有规定作用,个体目的应符合正义的要求。从整个论证来看,罗尔斯的一致性论证对应着康德德福一致的论证。可以说,康德的论证是罗尔斯论证的原型。康德的原型论证是这样的两个步骤:道德法则产生敬重这种动机,由敬重道德律的动机而行动与追求幸福相一致。而罗尔斯论证的两个步骤正好是与之对应的:正义原则产生正义感,按正义感行动与个人追求善一致。换言之,罗尔斯的第一个论证所涉及的正义原则如何产生正义感这个问题,实际上源自康德在确立道德法则之后去发掘道德情感的问题。第二个论证中,康德的思路是:幸福的确可能与德行相冲突,但是如果以德行作为先决条件,再辅以某些悬设,幸福可以和德行一致;而罗尔斯的思路是:个人的生活计划可能与正义原则相冲突,这时候个人目标应当受正义原则的约束,在此前提下,两者可以达到一致。

这种论证结构上的对应实有其深层次的根基,即正当优先于善的理论模式。这种模式,在康德那里,意味着道德法则及其要求对于幸福的优先性,任何人追求幸福的行动,必须以满足德行的要求为先决条件,道德的要求任何时候都不以利益为转移。在罗尔斯那里意味着首先需要确定一种正当或正义的原则,这种原则优先于任何个人欲求的满足或合理的生活计划,后者必须受前者的约束。任何违背正当原则追求个人目的实现的做法都是不正义的。一言以蔽之,“无论我们的具体环境如何,我们的生活方式必须始终独立地达到与正义原则的一致。”^③

除了康德和罗尔斯的义务论哲学中正当优先于善的模式,在道德哲学中,还存在着另一种理论模式,即善优先于正当的模式。各种目的论的伦理学即属于后一种模式,其典型的形态是功利主义。按照这种模式,善可以独立地被定义,如功利主义把善界定为合理欲望的满足,并最终还原为快乐的感觉。这种快乐本身就是功利的衡量标准,不需要预设任何正当原则。在此基础上,功利主义再以能使总功利或平均功利达到最大的原则作为正当的标准,从而在善之上建立正当原则^④。

很容易看出,功利主义的理论模式,出于先行界定善的需要,非常依赖于自然事实与人类生活中的偶然因素,如人们的各种偏好。而在义务论的模式中,一种基于理性的原则始终是在先的。这是一种理性存在者可以超越各种偶然因素、甚至超越时代地采纳的思想模式,罗尔斯称之为一种“永恒的观点”^⑤。这体现了两类理论的出发点截然不同:理性的必然要求与自然欲求的满足。正当优先于善的理论模式从根本上是被理性自身的结构所规定的。在康德那里,纯粹实践理性优先于一般实践理性,而在罗尔斯那里,理性优先于合理性。罗尔斯对于康德式的实践理性之统一做出了清晰的解释:理性的东西

① 罗尔斯:《正义论》,第394页。

② 罗尔斯:《正义论》,第451及以下诸页。

③ 罗尔斯:《正义论》,第355页。

④ 罗尔斯:《正义论》,第20页。

⑤ 罗尔斯:《正义论》,第464~465页。

(the reasonable)为合理的东西(the rational)塑形并绝对地限制了后者^①。

四、启示：哲学史的当代意义

无论是罗尔斯正义理论的基本内核(道德建构主义或者说正义原则之证明),还是正当与善的一致性论证及相应的正当优先于善的理论模式,都脱胎于康德道德哲学的基本思想。我希望上面的论述已经表明了这一点。

当然,康德与罗尔斯之间也有着不容忽视的、多方面的重要差异。其中最突出的应该说有两点。第一,康德的纯粹实践理性概念的意义在于自我立法的一贯性、普遍性;罗尔斯与此相对应的理性(reason)概念的主要含义固然也是建立与执行原则的一贯性,但其所规定的具体对象则包含了自由、机会与财富等在罗尔斯看来任何理性人都需要的基本社会善。这显示出罗尔斯的理性概念与康德相比,涵盖内容更广。其中的某些内容,尤其是财富,在康德那里,属于构成幸福的要素,并不属于道德的领域。这是康德与罗尔斯最大的差异。从一个严格康德主义的立场看,罗尔斯的理性概念和相应的基本社会善的内容似乎带有一定的功利主义色彩,甚至显得混淆了康德所严厉区分的道德与幸福两个领域。

指责罗尔斯所固有的英美经验主义的文化先见在这里使得罗尔斯无意识地偏离了康德的轨道,似乎是一种自然的做法。但我觉得,这是罗尔斯为了扬弃康德的二元论而有意识地这样做的结果。这种二元论包括现象与本体、德行与幸福、形式与质料等诸多形式。这些二元论,在很多现代哲学家看来,似乎很难与大量的自然科学、社会科学的经验性研究成果相协调。罗尔斯显然注意到了这个问题。出于反思平衡的考虑,为了在原则和各种现实的经过了思虑的判断(considered judgements)之间达到协调一致,罗尔斯对康德哲学的很多内容作了经验理论范围内的重新考虑。理性概念即是其中之一。罗尔斯消除了康德的理性概念的本体基础,使得康德那里具有古典形而上学的玄妙色彩、相对而言难以证明的内容被取消了,理性概念可直接与经验科学接轨。这种做法在《政治自由主义》中更加明显:为了适合合理多元论的现实,罗尔斯把自己的自由主义理论限定为政治的,而非形而上学的或整全的(comprehensive)。罗尔斯的理性人的概念因此很接近于社会科学中的理性人的概念。

第二,与第一点相关联,罗尔斯正义论的很多具体操作方式也接近于现代社会科学:比如罗尔斯是从社会基本制度的角度来谈正义原则并把人类的生活条件纳入考虑;再如罗尔斯在具体证明两个正义原则的有效性时,选择的推论规则是经济学中采用的某些博弈论规则或数学规则。这一切与康德先天的理性演绎的方式颇为不同。甚至某些部分,比如在哲学研究中运用数学方法,对康德来说,是不可接受的。康德在《纯粹理性批判》中明确地提出:哲学方法与数学方法具有原则上的差异性^②。

即便有上述重要的差别,罗尔斯的正义理论极大地受益于康德的道德哲学,仍是一个不争的事实。罗尔斯的正义理论是康德式的义务论伦理学的当代形态。我相信,上文的描述与论证已经提供了一个很好的个案。此个案典型地表明了:古典哲学家的基本观念与方法可以深刻地影响当代主流哲学家的思考。由这个典型,我想提出一个比较一般的命题:古典哲学家作为我们灵感的源泉,对当代哲学思考具有重要意义。这也是罗尔斯本人反复强调的观念:对于哲学史上的伟大人物或传统应表示足够的尊敬^③。当然,康德—罗尔斯的关联绝不是孤证。仍以康德哲学为例,即便不考虑深受其影响的欧陆哲学^④,仅以19世纪末叶以来的英美哲学来说,所启发的远不止罗尔斯一人,查尔斯·皮尔斯、彼得·斯特劳森、理查德·黑尔等响亮的名字都可以列入这个证据的清单。

① John Rawls, *Lectures on the History of Moral Philosophy*, pp. 230~232.

② 康德:《纯粹理性批判》,A712-738/B740-766,邓晓芒译,杨祖陶校,人民出版社2004年,第552~569页。

③ 在纪念德雷本(Burton Dreben)的文章里,罗尔斯充分地表达了这一点。他甚至有些夸张地写道:“不管我们多么努力地尝试着去掌握他们的思想,所有那些伟大的人物……在某种程度上都是高不可攀的。就康德来说,这种距离常常对我显得格外地不可企及。”转引自 *Lectures on the History of Moral Philosophy*, p. X VII.

④ 康德对19世纪以来的欧陆哲学具有极其深刻的影响,大概是不刊之论。无论是从费希特到黑格尔(也许可以把叔本华也包括进来)的德国古典哲学,还是新康德主义、现象学等各主要思潮,无不以康德哲学为其直接的思想源泉。在此我还想提供一个有趣的例证:在牛津大学的哲学课程设置中,把19世纪以来的欧陆哲学课程一概称之为“后康德哲学”,隐隐有一种主张现代欧陆哲学思潮源于康德哲学的意味。

我相信,这可视作哲学反思与研究的一个特点:传统始终具有重要意义。就这一点来说,哲学明显不同于科学^①。当代的物理学家恐怕不必以牛顿的《自然哲学的数学原理》作为自己灵感和方法的来源,通晓欧几里得的《几何原本》也不会是一位几何学家的主要工作。而在哲学家那里,哲学史却显然是灵感的源泉。这大概是源于哲学的独特性。哲学的问题常常是古老的,甚至是永恒的。我们很难说柏拉图、孔子和黑格尔曾经面对过的问题,对于当代哲学家就不再有意义。哲学的方法往往只是在哲学家头脑中进行的抽象的理论反思,如概念构造、逻辑推论、思想实验等。这与具体实证科学在实验室中进行检测,或者以实际的田野调查与分析为内容,构成了鲜明的反差。哲学的目标是人的反身的系统的自我理解——这似乎是一个不那么容易完成的目标。也许,从根本上说,就像罗素所声明的那样,哲学的研究领域就像是科学和宗教之间的无人地带,其问题永远不太可能获得确定的问答^②。

我相信,正是哲学的独特性,可以解释为何历史上多次挑战形而上学、以科学替代哲学的尝试始终没能取得成功。从早期现代“人是机器”的主张,到19世纪中后期至20世纪前半叶的实证主义潮流,直到当代自然主义倾向,对于哲学传统的挑战,从未停息,但从未真正取得过胜利。我倾向于认为,这种挑战在未来也不会取得成功。一个理由就是上面所阐释的哲学的独特性。而康德对罗尔斯的影响与启发以及类似的大量个案所表明的哲学史的当代意义,同样是重要的理由。

当然,我们也不能忽视事情的另一面:对于当代哲学和科学新成果的吸收、新方法的借鉴,也是当前哲学研究的内在维度。重视哲学史及哲学经典,从来不意味着忽视哲学与科学的前沿。罗尔斯的工作同样典型地体现了这一点。罗尔斯对于康德二元论的扬弃、对于社会科学尤其是经济学方法的使用即是最好的例证。如果我们仍以康德的道德哲学研究为例,除了对于康德思想的阐释、辩护或批评,我们同样需要关注康德在当代所启发的思想,如罗尔斯的正义理论,如各种形态的新义务论。我们还应该关注那些与康德主义伦理学不同的工作路径,典型的如近年来兴起的进化论伦理学。即便如某些进化论伦理学家自己所提出的,进化论伦理学原则上只是描述性的理论,并不直接构成对各种规范伦理学的反驳或能替代规范伦理学的研究^③。但我们还是要问:这些描述性的工作与康德的道德哲学相容吗?如果它们真的构成了对康德的挑战,持康德主义立场的哲学家该如何回应这类挑战?

我的结论是这样的:我们当然要做前沿的研究,但同时我们也有充分的理由对任何忽视哲学传统与哲学经典的研究表示警惕。这是上述康德—罗尔斯思想关联的个案研究对我们的启示。最后,容许我套用康德那个著名的鸽子的比喻:一旦脱离了空气(哲学史与哲学经典),鸽子(真正严肃的哲学研究)恐怕是不能自由飞翔的^④。我想说的是:能经得起时间的考验、能作为人类精神生活的真正养料而沉淀下来的哲学理论,通常是各个伟大哲学传统的回声。

● 作者简介:杨云飞,武汉大学哲学学院讲师,哲学博士;湖北 武汉 40072。

● 基金项目:武汉大学“70后”计划自主科研项目(275931);美国坦普顿基金会(John Templeton Foundation)资助项目“康德哲学中的道德与宗教信仰”

● 责任编辑:涂文迁

①对于与之相关但涵盖更广的问题,即自然科学、社会科学与人文科学之间的差异问题,当代美国著名的心理学家 Jerome Kagan 做了深入的研究。读者可参看其 *The Three Culture: Natural Sciences, Social Sciences and Humanities in the 21st Century* (Cambridge University Press 2009)一书。在第1~5页,Kagan对三种科学的差异有一个总体的比较,他强调了过去的文本、事件和历史对于包括哲学在内的人文科学的重要意义。

②罗素:《西方哲学史》,何兆武、李约瑟译,商务印书馆1963年,第11~13页。

③Elliot Sober, David Wilson. "Summary of *Unto Others*", in Leonard Katz(eds.) *Evolutionary Origins of Morality*. Imprint Academic 2000, pp. 204~205.

④康德:《纯粹理性批判》,A5-B9,中译本第7页。