

互利、对等与公平^{〔*〕}

——罗尔斯正义理论的休谟因素

张国清¹, 高礼杰²

(1. 浙江大学 哲学学院, 浙江 杭州 310012;

2. 西南政法大学 马克思主义学院, 重庆 401120)

〔摘要〕在阐述公平正义理论时,罗尔斯自称是社会契约传统理论的继承者,却把霍布斯排除在这个传统理论之外;他以古典效用论为主要批评对象,却把古典效用论者休谟排除在批评对象之外;他声称其理论“颇具康德色彩”,却把休谟在《人性论》中讨论契约要件时提出的对等概念引入正义理论,用来限定差别原则的适用范围。罗尔斯批评者大多忽视霍布斯互利概念和休谟对等概念之间的本质差异,忽视休谟对等概念在罗尔斯建构正义理论中扮演的重要角色。通过评估罗尔斯正义理论的休谟因素,有助于我们更好地把握他的理论。

〔关键词〕罗尔斯;休谟;康德;互惠;对等;差别原则

DOI:10.3969/j.issn.1002-1698.2022.04.004

美国政治哲学家约翰·罗尔斯在建构公平正义理论时,为了区分“公平”和“正义”概念,引入“互利(mutual advantage)”“互惠(mutual benefit)”和“对等(reciprocity)”概念。他用“对等”来描述处于初始位置上并且站在无知之幕后理性的个体的初始契约关系,表示那些个体参与社会合作的理由:“差别原则表达了某种对等观念(a conception of reciprocity)。它是一种互惠原则(a principle of mutual benefit)……处境较优者承认,每个人的福利都依赖于某种社会合作,没有它,任何人都不可能有令人满意的生活……他们

还承认,只有在合作条款合理的情况下,他们才能期待所有人皆愿意合作。”^{〔1〕}1971年,几乎在《正义论》发表的同时,罗尔斯专门写过《对等正义》(justice as reciprocity)一文,表示“公平和正义是两个不同的概念,但它们有一个根本的共同点。本人称之为对等概念(a concept of reciprocity)”^{〔2〕}。

为了创立公平正义理论,罗尔斯借鉴了四个先行假设,它们分别是霍布斯互利假设、洛克自利假设、休谟互惠对等假设和康德道德理性假设。罗尔斯表示,在初始位置上,自由而平等的

作者简介:张国清,浙江大学哲学学院教授、博士生导师,主要从事政治哲学研究;高礼杰,哲学博士,西南政法大学马克思主义学院讲师,主要从事政治哲学和法哲学研究。

〔*〕本文系国家社会科学基金后期资助重点项目“《正义论》评注”(21FZX008)的阶段性成果。

人们将达成休谟式“对等”契约,而非霍布斯式“互利”契约。“互利”和“互惠”经常被人混用,罗尔斯本人偶尔也会如此。两者表面含义相近,但实际含义相差甚远。与“互利”相比,“互惠”与“对等”的含义更加相似,罗尔斯经常并用“互惠”和“对等”。“对等正义(justice as reciprocity)”在证明“公平正义(justice as fairness)”中占有一席之地。

尽管罗尔斯本人忽视“对等”和“互利”的差异,但是这个差异对理解罗尔斯正义理论是至关重要的。它是罗尔斯正义理论区别于传统契约论而成为新契约论的重要依据,也是该理论对效用论既有批判又有借鉴的重要依据。如果我们对“互惠”概念作休谟式解读,亦即解读为“对等”,而非罗尔斯本人一度承认的霍布斯式解读,亦即解读为“互利”,那么罗尔斯正义理论的霍布斯因素将削弱,休谟因素将增强。尽管罗尔斯本人和众多批评者否认罗尔斯和休谟正义理论的这种内在联系,但是罗尔斯正义理论的确存在休谟起源。要想准确理解罗尔斯正义理论,那么其休谟起源不应当被忽视。

依据纳斯鲍姆、巴里、吉巴德等人分析,罗尔斯“对等”概念与霍布斯“互利”概念不相匹配。罗尔斯本人也承认这一点。^[3]在罗尔斯正义理论中,“对等”概念表示存在一种人人生而平等的倾向,足以令差别原则区分于效用论的有限效用原则,以突出正义原则的公平属性。^[4]不过,在纳斯鲍姆和科恩看来,罗尔斯借助于“对等”概念强调的平等,不是全面平等,而是片面平等。^[5]巴里则表示,罗尔斯支撑“对等”概念的平等动机和博弈手段,本质上源自两种不同的权益观念,二者无法完美地融贯起来。^[6]

批评者大多忽视罗尔斯“对等”概念的休谟起源。他们主要从霍布斯互利假设、洛克自利假设和康德道德理性假设来解读罗尔斯正义理论。“互利”“自利”和“道德理性”而非“互惠对等”概念成为理解正义原则的主要入口。假如这种诠释成立,那么罗尔斯为消解公平与效率的张力

而构想的初始位置和无知之幕,就丧失了必要性和有效性,由此演绎的正义原则也丧失了正当性。

因此,一方面,罗尔斯批评者(甚至包括罗尔斯本人)对“reciprocity”概念作霍布斯式解读是错的,那个错误进一步导致批评者对罗尔斯公平正义理论的误解。并且,罗尔斯本人是造成那个错误的原因之一。另一方面,不仅因为休谟的古典效用论者身份,而且因为罗尔斯正义理论明确把古典效用论作为主要批评对象,虽然个别批评者看到罗尔斯“对等”概念与休谟“对等”概念存在紧密联系,但是他们未能站在休谟的立场来解读罗尔斯的新契约论,即公平正义理论。罗尔斯本人及其追随者都没有认真思考过对正义理论作休谟式解读的可能性。

笔者认为,正如休谟在知识论方面对康德产生致命影响一样,他在道德心理学方面对罗尔斯产生致命影响。罗尔斯在接受《哈佛哲学评论》访谈时流露了那种影响的致命性:他表示《正义论》最令人向往的是第三部分,即讨论道德心理学的部分,那是他原本打算用其余生作深入研究的。而它恰好是其在休谟身上用功最勤的部分。赫尔曼透露了那种影响的一些蛛丝马迹:“对休谟的广泛讨论,尤其是对《人性论》第2卷的广泛讨论,一直是康德讲座的组成部分……罗尔斯细心而详尽地考查了休谟的道德心理学(尤其关注休谟对理性直觉论的批判),从《人性论》原著中小心翼翼地提取了休谟对慎思和实践理性所作的丰富的、层次分明的探讨。休谟的观点既是强有力的又是原创性的。不过,罗尔斯还是阐明了建立在欲望基础上的对慎思所做的探讨的局限性,并指出这促成了对于康德的唯理论道德心理学是至关重要的‘依赖于原理的欲望’在对实践理性的探讨上的至关重要性。”^[7]因为罗尔斯只承认康德道德心理学的至关重要性,不承认自己推崇的道德心理学受到休谟的决定性影响。时至今日,这种影响的重要性仍然没有受到政治哲学界的认真评估。

有鉴于此,本文旨在讨论休谟和罗尔斯的

“对等”概念,揭示罗尔斯正义理论的休谟因素和康德因素,评估它们对于正确理解罗尔斯正义理论的重要性。我们将(1)探索霍布斯“互利”概念、休谟“对等”概念和罗尔斯“对等”概念的联系;(2)证明罗尔斯“对等”概念存在休谟的思想起源,罗尔斯对等正义观念有着休谟的思想来源;(3)评估罗尔斯正义理论的休谟因素和康德因素,展示罗尔斯对休谟效用正义理论借鉴多于批评的事实,表明罗尔斯在主观上试图划清公平正义理论同效用正义理论的界线,但在客观上难以摆脱效用正义理论对自己的深远影响。

一、“互利”和“对等”的观念史追溯

在近代西方政治哲学中,“互利”(mutual advantage)概念源于霍布斯,“对等”(reciprocity)概念则源于休谟。古典效用论者推崇互利和互惠,社会契约论者也不反对互利和互惠,他们都没有严格区分两者。休谟是一个例外,认为只有“对等”概念才能更好地表达平等契约的本质。

罗尔斯在《正义论》中也没有严格地区分“互利”(mutual advantage)和“对等”(reciprocity)概念。但是,两者的区别明显且非常重要。他后来在《道德哲学史讲义》中明确提到休谟的“对等”正义观念,从中可以清晰看到罗尔斯“对等”概念在休谟对等契约论中的起源。

首先,从词义上讲,“互利”和“对等”具有不同的含义,用“互利”来解读“reciprocity”将对罗尔斯正义理论产生误解。在罗尔斯著作的中文译本中,英文单词“reciprocity”有时译成“互利”(mutual advantage),有时译为“互惠”(mutual benefit),^[8]个别译者译为“相互性”^[9]或“对等”。^[10]在国际政治和国际贸易中,“reciprocity”一词首先强调的不是契约各方寻求的“利益”(interests),而是契约各方承诺的“平等”(equality)与“公平”(fairness)。罗尔斯在《万民法》中也谈到过“reciprocity”准则在国际贸易领域的运用。他之所以使用“reciprocity”一词,正是因为它凸现了平等观念。如果“reciprocity”被译为

“互利”“互惠”或“相互性”,没能充分体现该词着重强调的“平等”与“公平”观念。只有汉语词汇“对等”或“对等性”才能准确表达在罗尔斯正义理论中的“reciprocity”的确切含义。

其次,在西方政治思想史上,“互利”和“对等”反映着不同的契约论传统。国际法把“reciprocity”理解为带来互利互惠的“对等”,在实践上是可行的。但是,在罗尔斯正义理论中,“互利”和“对等”存在着根本差异。它不只表现为强调平等的程度差异,而且代表了两个完全不同的对待契约的传统。在那两个传统中,一个可以追溯到霍布斯,另一个可以追溯到休谟。

“互利”观念体现了霍布斯式契约论传统,允许达成不平等的契约,甚至奴隶制度也被视为“互利”的契约形式。“理性提出了便利的安全条款,人们可以据此达成协议。”^[11]霍布斯式契约论包含两层含义:(1)协议的直接动机是单纯的,仅在于对作为首要利益的安全的确认;(2)参与协议的民众可以获得较之自然状态更好的处境,但处境的类型并非既定一种。

霍布斯式契约论会导致两个难题:(1)如果有人安全前提下,为了更多利益而违背初始社会契约,那么社会无法提供任何惩罚和规制手段强制其遵循已与他人签订的协议,此为“搭便车难题”;(2)如果人们沦为奴隶比死亡更好,那么签订卖身协议似乎既能提升奴隶主实际控制奴隶的能力,又能提升奴隶本身的生存处境,此为“不平等难题”。霍布斯没有对这两种违背当下道德直觉的状况提出辩护。因此,霍布斯式协议顶多算是一种休战协议,而不是平等协议。

按照霍布斯式互利契约论,除了“正义观念下的社会合作”,它不排斥“强力观念下的社会合作”“等级观念下的社会合作”等等。罗尔斯显然不会接受霍布斯式契约,因为它是一种替各种奴役制度作辩护的强力观念下的社会合作。然而,在霍布斯赤裸裸的“互利”契约中,此类情感是找不到合法来源的。那种“互利”契约需要利维坦这样的至高强权者才能得到维持。因此,

霍布斯式契约是迫不得已的,是一种契约方基于各自利益最大化的“盟下之约”。主人和奴隶之间达成的协议是极其不平等的,等于奴隶的“社会性死亡”。“如果奴隶不再属于一个共同体,如果在其主人以外他没有任何社会存在,那么他还是什么呢?在几乎所有奴隶制社会里,最初的回应是,奴隶被定义为在社会意义上死亡的人。”^[12]在互利契约中,权力有着严格的等级序列,权力的本质是统治。利维坦或绝对君主假设表明,人际平等是不可能的。因此,按照霍布斯式契约,真正的平等的心甘情愿的社会合作是不可能的。

相比之下,“对等”的前提是成员资格平等。订立契约各方为了平等分享共同价值、责任和负担而作出社会合作的努力。论证心甘情愿的社会合作的可能性,也即证成某种良序社会的可能性,是罗尔斯一生追求的学术目标。这也是罗尔斯的新契约论区别于传统契约论的重要方面。我们从中可知,以“互利”为目的的霍布斯式契约论背离了罗尔斯公平正义理论的道德目标。

就契约的初始动机而言,休谟同霍布斯类似。在休谟看来,自利和对自己父母、爱人、子女的偏爱是人之情。与霍布斯不同的是,休谟不认为,处于人人偏私的自然状态,人们能顺利地达成协议,而进入普遍合作的社会政治状态。因为人人偏私的天性如此根深蒂固,以至于在得到驯化之前,就会出现霍布斯“互利”契约论的“搭便车难题”,进而导致合作的失败。因此,“我们不可能希望在未受教化的自然状态中给这种不利条件找到一种补救方法;我们也不能希望,人类心灵中有任何一个自然的原则,能够控制那些偏私的感情,并使我们克服由我们的外界条件所发生的那些诱惑”^[13]

休谟认为,在直接利己背景下,建立普遍的社会合作是不可能的。但在家庭和小型社会中,建立合作是可能的。原因在于,在家庭和小型社会中,人们共享某些关于道德和共同行为的情感和评价。随着生活范围的扩大,关于道德和共同

行为的共享评价,逐渐转化为内化规则。“他们最初只是由于利益考虑,才一般地并在每个特殊例子下被诱导了以这些规则加于自己的身上,并加以遵守;而且在社会最初成立的时候,这个动机也就是足够强有力的。”^[14]也正是到了这个阶段,人们才会因为某个未加谋面的人触犯了规则而感到不安,才会产生羞辱感。于是,人的自然德性需要人为德性即正义来规制和调节。

休谟认为,在共享情感背景下,“协议只是一般的共同利益感觉;这种感觉是社会全体成员互相表示出来的,并且诱导他们以某些规则来调整他们的行为”^[15]这样一来,对道德行为的评价就有两个来源,一个来源是自然的,是人的自利;另一个来源是人为的,是人们共享的道德评价。“自私是建立正义的原始动机;而对于公益的同情是那种德所引起的道德赞许的来源。”^[16]由此可见,关于契约的霍布斯传统与休谟传统的关键差异在于,后者在“自利的偏好”与“互利的契约”之间,增加了一个关键阶段——“正义的观念”。只有在大社会里,对等正义超越互利道德,成为规范人们行为的准则,凸现其作为人为正义(artificial justice)的本质特性。

根据人为对等正义观念形成的社会契约,既杜绝了“搭便车”行为,也杜绝了违反道德直觉的协议规则。休谟认为,国家与家庭在某种程度上具有相似性。霍布斯契约论拒绝承认这一政治经验,显然存在以下问题:如果人们仅仅追求自身利益是一条心理学法则,他们就不可能具有一种有效的(按照效用原则规定的)正义感。这种正义感是政治事实的唯一合理解释。欠缺正义感是霍布斯契约论的一大致命软肋。这是罗尔斯用休谟尤其是康德的正义感理论改造传统契约论的重要原因,也是罗尔斯把霍布斯排除在他想要继承的契约论传统之先驱者名单之外的原因。

最后,罗尔斯在霍布斯和休谟之间作出了选择,他的“对等”概念直接来自休谟的“对等”概念。罗尔斯在契约的直接动机问题上站在休谟

一边。不同于霍布斯,罗尔斯否认正义的道德原则单纯产生于纯粹理性选择,否认正义原则是为了提升个体利益而设计的。罗尔斯提出的初始位置和无知之幕后安排有两个假设前提,一是良序社会,二是有限知识,实际上把正义背景和正义感放在契约条件下的社会选择行为之前。只有拥有正义感和道德判断力的个体,才会产生真正的社会合作,才能超越单纯的互利原则,接受公平正义且互惠对等的差别原则。这是罗尔斯对民主社会的个体即通情达理者提出的道德要求。由此不难理解,罗尔斯何以会认为《正义论》中最重要的部分是关于正义感的讨论,并非两个正义原则的筛选论证。

休谟和康德都深入讨论过正义感这个话题。出于推崇道德自律,康德把正义感只赋予少数人。休谟则认为,正义感是人人具有的普通情感,不需要特别的道德训示。正义感不仅必要,而且能够仅仅通过经验形成。这正是休谟比康德更加出色的地方。罗尔斯认为,具体途径包含经验主义效用论(休谟、西季威克)和经验的发展心理学(密尔、弗洛伊德、皮亚杰和科尔伯格)两种传统学说。无论正义感形成的具体途径如何,各种学说各有其合理成分,结合起来似乎更能把问题讲清楚。最终,罗尔斯得出一个非自利式结论:“大多数传统学说认为,至少在某种程度上,人性是这样的:如果我们生活在正义制度之下并从中受益,我们就渴望做事合乎正义。如果真的如此,那么一个正义观念在心理上合乎人心之所向。此外,如果事实证明,渴望做事合乎正义也调节着理性人生规划,那么做事合乎正义是我们拥有的权益(善好)的组成部分。在这种情况下,正义观念和善好观念是兼容的,这个理论作为一个整体是同余的(congruent)。”^[17]

二、罗尔斯正义理论的休谟因素

罗尔斯在《正义论》最后部分用很大篇幅来证明正义和善好的同余关系,想要把休谟所持的作为人为正义的对等观念和康德所持的作为定

言令式的正当观念调和起来。正如弗雷曼指出,罗尔斯把对等观念与分配正义原则联结起来:“为了让它显得令人信服,罗尔斯不得不把平等的公民资格的条件同分配正义原则联系起来,那个原则表达了一种对等观念(a conception of reciprocity)。”^[18]因此,对等观念既涉及平等的公民资格,又涉及分配正义,主要是差别原则。对等观念在罗尔斯提出差别原则中扮演着决定性作用。这个观念的首要贡献者是休谟。在分配正义和差别原则方面,休谟给罗尔斯的启示要多于其他契约论者。

休谟是主张“自利优先兼顾利他”的道德哲学家,罗尔斯为了论证公平正义理论而提出的人性假设也以自利作为起点。最后罗尔斯也和休谟一样肯定“对等恰好处于自利和利他之间”。^[19]与休谟的主张相比,霍布斯式契约不太符合罗尔斯的基本立场和理论意图。不过,罗尔斯在《正义论》中明确提出,自己的正义理论是对洛克、卢梭和康德创设的契约论的提升和改进,其中没有提到休谟。^[20]一般认为,罗尔斯正义理论同康德道德哲学的联系更为紧密,他本人表示公平正义理论“颇具康德色彩(highly Kantian in nature)”。^[21]罗尔斯在该书第40节就公平正义作出了康德式解释,表示其契约论是在道德心理学和社会选择理论视角下对康德契约论的补充。罗尔斯把“对等”契约视为康德“互尊和自尊(mutual respect and self-esteem)”的天然关联。^[22]问题在于,我们何以能够对罗尔斯的“对等”契约论作出休谟式而非康德式解读?换言之,罗尔斯的“对等”契约论在何种意义上是康德的,在何种意义上是休谟的,两者是否兼容?

在罗尔斯看来,康德伦理学在四个层面对“对等”契约理论提供支持。(1)道德原则本来就是理性个人选择的对象,而不是说个人的理性选择对象仅仅只是自利。^[23](2)所选择的道德原则对社会而言是基本结构层面的原则,而并非具体的实践原则。罗尔斯正义理论把社会基本结构作为首要主体与康德学说非常相近。^[24](3)康

德伦理学暗含着初始位置和无知之幕假设。^[25] 罗尔斯关于初始位置和无知之幕的设想,排除了理性个人在道德原则选择中基于“他律”而行动的可能性。因此,在对等的动机下签订契约选择的正义原则,本身就是一种定言令式(绝对命令)。^[26] 康德对自律行动的保证仅仅是自律原则本身。(4)相互冷淡或互不关切(mutually disinterested)动机假定同康德“自律”观念具有相似性。这一立场使罗尔斯满足其有限假定的初衷。根据这一理由,他拒绝承诺某种普遍利益的利他主义作为契约的动机。^[27]

不过,罗尔斯也对康德伦理学和契约论提出两点批评:(1)康德的“自律”对伦理学而言是一个过于狭隘的基础,它不包含人们在“对等”契约中选择的正义原则的前提,即关于正义背景的基本事实。^[28] (2)康德式自律契约论并非只依据纯粹理智的自由,事实上其中广泛存在深层次的二元论,包括形式与内容、理性与欲望、本体与现象等。^[29] 罗尔斯认为,康德以纯粹形式论证契约论是失败的。康德的纯粹形式隐含二元矛盾对立,承认康德的假设不啻于承认社会合作的动机既是经验的又是先验的,既是内容的又是形式的。准确来讲,罗尔斯在正义的前提问题上坚持“奥卡姆剃刀”,主张最小程度依赖假设条件和直觉断言,以便获得科学的普遍原则。^[30] 出于这一考虑,他断言道德意识具有稳固的支点,这些支点是历史演变和个体发生的结果,是我们必须依赖的内容,但否认这些支点能够获得理性或经验的终极基础。^[31]

也就是说,罗尔斯在道德意识问题上双双否定了经验和先验的人类学进路,而只保留了道德意识的社会学进路。虽然他认同康德“自尊和互尊”式契约下的道德原则所建构的社会毫无疑问具有正义性,认同康德珍视的人具有自由、平等、理性的特性。“康德认为,当个体的行动原则被选作这个人作为自由平等理性生命之本性最充分的可能表达时,他就是在自律地行动。”^[32] 但是罗尔斯并不认为自由平等的理性生命必须是

订立契约的前提,也不认为康德对他律的消解必然走向自律。解决康德问题的途径在于,在正面评价自由平等理性的人类社会状况的同时,用经验的道德社会学取代先验哲学人类学。罗尔斯由此与康德产生巨大分歧:罗尔斯使用“从自我偏好出发的论证,而不是实质性论证”。^[33] “自律”在罗尔斯那里更接近于认知原则,而非道德原则。斯坎伦很好地继承了罗尔斯的衣钵,阐明了一种非实质性道德立场:人们在回答道德问题时,务必以某种方式考虑他者利益。因此,以下判断是成立的:“康德的观点不承认普遍道德事实,但是罗尔斯正义理论却依赖于它们。”^[34] 也就是说,罗尔斯的论证依赖于普遍认可的经验事实,而不依赖于相互冲突的实质性价值判断。道德主体作为理性主体在道德认知的立场上是“自律的”,而在经验的道德判断立场上则“公然是他律的”。^[35] 罗尔斯的以下见解是显著地反康德的:“没有必要援引神学学说或形而上学理论来支持其原则,也没有必要想象另一个世界,来补偿和矫正在这个世界里两个正义原则允许的不平等。正义观念要么必须经由我们所知的生活条件得到证明,要么根本不能得到证明。”^[36] 由此,罗尔斯便回到了休谟。

罗尔斯曾经这样评价休谟的契约论:“这种契约观念是复杂的。我保证不了我一直能够公正地对待休谟的观点。我认为,他正在讨论的这些东西,他在第9-11段——实际上是在这一整节中——作了更多讨论的这些东西,是《人性论》中比较精彩的部分。在这里,其明察秋毫而鞭辟入里的运思力量得到了鲜明而简洁的表达。”^[37] 罗尔斯关于契约稳定性的论述进一步支撑了上述观点。稳定性是任何道德概念都试图追求的特质:“即便一种正义观念在其他方面具有吸引力,但如果它的道德心理学原则不能刺激人们产生按照这种观念行动的欲望,那么这种观念就有严重的缺陷。”^[38] 在此,罗尔斯把有效的道德心理学原则视为稳定性的来源。这一观点有两层含义:第一,稳定性本身暗示,订立社会契

约的各方是存在个性差异的个体,正义原则不可能是某个直觉判断或先天概念的直接展开,正义社会也不可能是一个抽象的非历史的社会。第二,符合道德心理学的行为原则应当具有普遍优先的吸引力,否则相互冲突的正义原则无法形成稳固的社会契约。罗尔斯的对等契约不可能来自于纯粹的实践理性和彻底的自利,也不可能来自关于人类道德本质的实质性判断。针对稳定性所提出的难题,罗尔斯再次发现了休谟:“我认为,休谟是看到如下情况的第一个人:在一个小型社会里,自然约束力能够引领人们以正义的契约为荣。于是我们便拥有了关于契约的极其重要的进一步特征:自我强制性和稳定性。”^[39]

正义是人为德性。罗尔斯接受休谟的这一核心正义观念。他表示,“休谟把正义看作与自然德性相反的人为德性。这个话题对其自然虔敬来说是至关重要的:他试图证明,以我们在世上的处境和我们对社会的依赖为条件,道德和我们的道德实践是我们的人性的表现。”^[40]因此,对等契约实际上体现了两种对立的哲学谱系:(1)正义和道德在休谟那里是约定俗成的产物;(2)正义和道德在康德那里则来自于理想的道德王国。^[41]罗尔斯对两种传统的倚重是不同的。这一节的讨论强化了上一节提出的见解:在平等的道德心理学动机方面,罗尔斯显然更倾向于休谟,而不是康德。

三、结 论

罗尔斯断定,公平正义理论“颇具康德色彩”,^[42]但是,从前面考证可知,那个理论也留有厚重的休谟痕迹。休谟因素和康德因素至少平分秋色。也许,休谟的对等契约观启发了罗尔斯关于正义理论起点的思考。休谟提出“正义是人为德性”假设,它引导着罗尔斯关于“正义的首要德性”的思考。康德的道德理性观,尤其是“人是目的”的见解,则引导着罗尔斯关于正义社会制度的目标的思考。这一点从《正义论》的结束语中可以得到佐证:“人若能明心见性,即可

洞悉万物,以是观之,其所言所行,皆雅致而从容,克己而有度。”^[43]不过,假如我们只从霍布斯开始一直到康德的传统契约论视角来解读罗尔斯正义理论,那么,我们将忽视罗尔斯从休谟身上吸取的重要思想资源。与学术界看重的也是罗尔斯本人承认的洛克、卢梭和康德的传统社会契约论相比,那是以“对等”概念为核心的罗尔斯正义理论的另一个重要起源。我们称之为休谟因素。休谟因素的最大作用,就是通过对等概念限制差别原则的适用范围。

首先,罗尔斯的《正义论》主题清晰地隐藏于休谟的《人性论》之中。《正义论》的主题是,通过揭示以自由平等和公平正义为基本价值的社会基本结构,探讨社会合作的可能性。通过发展由洛克、卢梭、康德所代表的传统契约理论,对正义作出一种系统解释。罗尔斯这样评价《人性论》中关于自然德性、人为德性和人为正义的讨论:“休谟的难题是:发轫于原始自然状态且仅仅具有自然德性的人类如何最终抵达了他们还拥有刚才提及的那些特性的人为德性的文明阶段?”^[44]换言之,休谟关注的课题是,人的自然德性是如何升华为人为德性的,人类是如何从野蛮走向文明的?休谟对自然德性、人为德性、人为正义和对等正义等的探讨为罗尔斯完成自己的学术目标开辟了道路。

其次,罗尔斯直接从休谟那里借鉴了“对等”概念,“对等正义”成为“公平正义”的平行理论。罗尔斯著作《正义论》发表于1971年,他的论文《对等正义》(justice as reciprocity)也完成于同一年,后者收录于1999年出版的《罗尔斯论文全集》之中。笔者的疑问是,罗尔斯为什么几乎在提出“公平正义”(justice as fairness)理论的同时又提出不太为人注意的“对等正义”观念?笔者推测,在提出公平正义理论之前,罗尔斯最早接受且最为欣赏的正义观念是“对等正义”观念。而“对等正义”观念的原创者是休谟。在近代西方道德哲学史上,休谟把“对等”概念引入道德哲学,“对等正义”思想起源于休谟。公平

正义理论正是从休谟“对等”概念和“对等正义”思想中得到灵感的。如果《正义论》以专章方式讨论“对等正义”，那么，罗尔斯正义理论的原初性将大打折扣。罗尔斯有意回避这一点，而以论文形式专门讨论“对等正义”。所以，尽管罗尔斯从一开始就借鉴休谟的“对等”(reciprocity)概念，但他在两个论著中皆有意误导人们对“对等”概念的理解，把“reciprocity”解释为“mutual benefit”(互惠)甚至“mutual advantage”(互利)，导致许多批评者从霍布斯式契约论而不是休谟式契约论来解读那个概念，进而误解公平正义理论。就正义理论被误解而言，罗尔斯本人的做法也是重要原因之一。

再次，除了康德，休谟是最受罗尔斯喜爱的哲学家。罗尔斯的《道德哲学史讲义》从休谟道德哲学开始讲起，他用5讲篇幅来解读休谟道德，用10讲篇幅讲康德道德哲学。他在《政治哲学史讲义》中用了2讲篇幅来讲解休谟政治哲学，但没有康德政治哲学讲座。罗尔斯毫不掩饰自己对休谟的偏爱：休谟是罗尔斯在两大讲义中作专题讲座的唯一哲学家。当然，这并不意味着，罗尔斯对休谟的偏爱超过对康德的偏爱。但是，这些数据至少说明，除了康德道德哲学和政治哲学，休谟道德哲学和政治哲学是罗尔斯关注最多、讲得最多、用力最勤的研究对象。

最后，在道德哲学和政治哲学上，正像斯密和密尔一样，休谟是效用论者。但是罗尔斯对休谟道德与政治理论，继承多于批判，他在批评边沁、斯密、西季威克、埃奇沃思和庇古等人的古典效用论时，把休谟的效用论排除在外。“休谟所推崇的那种效用论，不符合本人的目的；事实上，它严格说来算不得是效用论的。”^[45]罗尔斯之所以作出这样的评价，是因为他在休谟那里看到了自己正在追寻的东西。在休谟对等契约论和罗尔斯公平正义理论之间存在着连续性，罗尔斯正义理论存在休谟因素，罗尔斯有意掩盖了那个因素。与霍布斯版和康德版正义理论相比，休谟版正义理论更接近罗尔斯本人的学术目标。尽管

罗尔斯从未承认过这一点，然而事实就是如此。

注释：

[1][17][20][21][22][23][24][25][26][28][29][30][32][36][38][42][43][45] Rawls, J., *A Theory of Justice*, Revised Edition, Cambridge, MA: The Belknap Press of Harvard University Press, 1999, pp. 88, 399, 10, xviii, 225, 221, 222, 121, 222, 221, 226, 39, 222, 398, 398, xviii, 512, 28.

[2] Rawls, J., *Collected Papers*, Cambridge, MA: Harvard University Press, 1999, p. 190.

[3][27] Rawls, J., *Political Liberalism*, New York: Columbia University Press, 1996, pp. 17, 16.

[4][美]罗尔斯：《作为公平的正义——正义新论》，姚大志译，上海：上海三联书店，2002年，第149页。

[5][美]玛莎·C·纳斯鲍姆：《正义的前沿》，陈文娟等译，北京：中国人民大学出版社，2016年，第11-12页；[英]G·A·科恩：《拯救正义与平等》，陈伟译，上海：复旦大学出版社，2014年，第16页。

[6][33] Barry, B., *Justice as Impartiality*, New York: Oxford University Press, 1995, pp. 66, 9.

[7][37][40][44] Rawls, J., *Lectures on the History of Moral Philosophy*, Barbara Hermann (ed.), Cambridge, MA: Harvard University Press, 2000, pp. xv, 36, 51, 51.

[8][美]约翰·罗尔斯：《正义论》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年，第491页。

[9][美]约翰·罗尔斯：《罗尔斯论文全集》，陈肖生等译，长春：吉林出版集团有限责任公司，2012年，第774页。

[10][39][美]罗尔斯：《道德哲学史讲义》，张国清译，上海：上海三联书店，2003年，第82、86页。

[11][英]霍布斯：《利维坦》，黎思复、黎廷弼译，北京：商务印书馆，1985年，第97页。

[12] Patterson, O., *Slavery and Social Death: A Comparative Study*, Cambridge, MA: Harvard University Press, 1982, p. 38.

[13][14][15][16][英]休谟：《人性论》，关文译译，北京：商务印书馆，1980年，第528、539、530、540页。

[18][美]弗雷曼：《罗尔斯》，张国清译，北京：华夏出版社，2013年，第13页。

[19] Rawls, J., *Justice as Fairness: A Restatement*, The Belknap Press of Harvard University Press, p. 77.

[31][美]涛慕思·博格：《康德、罗尔斯与全球正义》，刘莘、徐向东译，上海：上海译文出版社，2010年，第179页。

[34][41][美]迈克尔·J·桑德尔：《自由主义与正义的局限》，万俊人等译，南京：译林出版社，2001年，第50、44-46页。

[35] Scanlon, T. M., *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, 1998, p. 6.

[责任编辑：汪家耀]