

正义论

〔美〕约翰·罗尔斯著

何怀宏 何包钢 廖申白译



外国伦理学名著译丛

70707

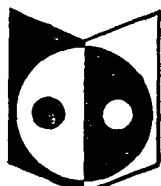
外国伦理学名著译丛

罗国杰 郑文林主编

正义论

[美] 约翰·罗尔斯著

何怀宏 何包钢 廖申白译



中国社会科学出版社

责任编辑：苏晓离
责任校对：李 晓
封面设计：谭国民
版式设计：王丹丹

正 义 论

Zhengyi lun

〔美〕约翰·罗尔斯著
何怀宏 何包钢 廖申白译

中国社会科学出版社 出版
发 行
社 考 古 居 经销
北京邮电学院印刷厂 印刷

850×1168 毫米 32 开本 20.75 印张 3 插页 508 千字

1988 年 3 月第 1 版 1988 年 3 月第 1 次印刷

印数 1—35 000 册

ISBN 7-5004-0244-9/B·45 定价：4.80 元

John Rawls
A THEORY OF JUSTICE

The Belknap Press of
Harvard University Press
Cambridge, Massachusetts

1971

根据美国哈佛大学出版社1971年版译出

27/100/09

外国伦理学名著译丛序

出版一套外国伦理学名著译丛，我们素有此愿，但碍于各种原因，却一直未能实现。近几年，一些同志已经努力译出了一批国外伦理学专著，但毕竟力量分散，也难见系统。现在承中国社会科学出版社的同志大力支持，这套丛书得以陆续出版，这是一件值得拍手称庆的好事。

中国历来号称文明古国，礼义之邦，伦理思想一向发达，特别是值此加强建设社会主义精神文明的时期，伦理学更有勃兴之势。为了迎接和促进伦理学的发展和繁荣，当然最重要的是研究当前我国社会主义社会的道德现象，按照党和人民的要求，探讨建设社会主义精神文明、提高全社会的道德水平和人们的精神境界的规律和方法，这是我们的主要着力点。但是，也有必要整理我国历史上留传下来的丰富的道德文化遗产，有必要借鉴国外的从古典到当代的各种独特的伦理思想成果，这是我们不能忽视的两翼。

这一套外国伦理学名著译丛，我们力求选入已有定论的古典名著和有较大影响的当代专著，包括较好的伦理学史和教科书；在翻译上，则力求做到文从字顺，不走原意。我们不仅希望伦理学专业的同志，也希望其他研究领域的同志来参加这一工作。本着贵精勿滥的原则，准备一年先出两本，积数年之功想必会做

出较大的成绩。

国外的伦理思想所产生的社会历史条件，所处理的道德问题和依凭的价值观念，跟我国目前的情况均有不同，所以，一番批判改造和消化吸收的功夫自然是不可少的。我们的读者，一定能以马克思主义的立场、观点、方法，带着中国的问题去阅读这些书，并从中得到正反两方面的启发借鉴，这也正是我们出这套丛书的希望所寄。

罗国杰 郑文林

译者前言

约翰·罗尔斯 (John Rawls 1921—)，美国著名哲学家、伦理学家，1921年生于马里兰州的巴尔的摩，1943年毕业于普林斯顿大学，1950年在该校获博士学位，以后相继在普林斯顿大学 (1950—1952年)、康奈尔大学 (1953—1959年)、麻省理工学院 (1960—1962年) 和哈佛大学 (1962—) 任教，现为哈佛大学哲学系教授。

罗尔斯著作不多，但在西方学术界影响甚大。1951年发表的《用于伦理学的一种决定程序的纲要》是他的初试之作。以后他专注于社会正义问题，并潜心构筑一种理想性质的正义理论，陆续发表了《作为公平的正义》(1958年)、《宪法的自由和正义的观念》(1963年)、《正义感》(1963年)、《非暴力反抗的辩护》(1966年)、《分配的正义》(1967年)、《分配的正义：一些补充》(1968年)等文。除了写作这些公开发表的论文，他还着手撰写《正义论》一书，前后三易其稿，最后又专门用了一段时间，在斯坦福的高级研究中心完成了此书，于1971年正式出版发行。

所以，《正义论》一书可以说是罗尔斯积近二十年的努力思考的一部心血之作，它集罗尔斯思想之大成，把罗尔斯十多年来所发表的论文中表达的思想发展成为一个严密的条理一贯的体系——即一种继承西方契约论的传统，试图代替现行功利主义

的、有关社会基本结构的正义理论。

《正义论》出版之后，更准确地说，还在它出版之前，其中的思想就已通过罗尔斯所发表的文章，产生了广泛的反响，引起了热烈的讨论。《正义论》的成书既是一种体系化的努力，又试图对各种已提出的反对意见进行辩驳或修正自身，这就自然使它在理论逻辑上更臻完善。所以，此书一出，很快就被誉为“二次大战后伦理学、政治哲学领域中最重要理论著作”，甚至被认为将列入经典之林。一般大学的哲学、政治、法律等有关学科都把它列为最重要的必读书之一，作为“标准的精神食粮”，许多大学还开设了专门讲解这本书的课程。报刊上发表了许多评论文章，出版了一些专门的评论文集和辅助性读物，并召开了讨论这本书的各种规模的学术讨论会。对于一位在世的且并非老资格的哲学家来说，这种盛誉和影响确实还是不多见的。而且，这种影响还波及到美国以外的其它一些西方国家，在那儿也激起了热烈的反响。

在我们看来，《正义论》一书产生如此广泛的影响可能有以下几个原因：首先是罗尔斯所研究的社会正义问题涉及到广泛的领域。他是作为一个伦理学家从道德的角度来研究社会的基本结构的，即研究社会基本结构在分配基本的权利和义务、决定社会合理的利益或负担之划分方面的正义问题。然而由于这一对象的性质，他在学科上就必然要涉及到伦理学、政治学、法学、社会学、经济学等许多领域，而且，他所研究的问题又关系到对每个人来说都是至关重要的切身利益，因此，他的讨论就不仅为伦理学而且为其它一些学术领域的学者所注目，甚至为一般群众所关心。

其次，罗尔斯酝酿和写作《正义论》的年代，在美国正是一个动荡不安的年代。在五十年代，美国外有朝鲜战争，内有麦卡锡掀起的反共喧嚣等；到六十年代情况愈演愈烈，在涉外方面有

古巴导弹危机、越南战争；在国内则有此起彼伏、如火如荼的争取民权运动，黑人抗暴斗争，校园学生运动，与豪富相对而言的贫困现象也成为令人瞩目的问题。美国社会正处在一种危机之中，处在一个亟须调整关系的关口，而罗尔斯《正义论》中所探讨的平等自由、公正机会、分配份额、差别原则等问题，恰以一种虚拟或抽象的方式提出了一些解决问题的建议或希望。照罗尔斯的说法，他的理论是理想性质的，不涉及任何现实的制度和政策，探讨范围仅限于一种“法律被严格服从的状况”限于一个“组织良好的社会”，因而他的理论又被人称为一种乌托邦理论。但是，这并不是那种老式的、真诚幻想和期待的乌托邦理论，而宁可说是一种证明方式和标准，一种想为非理想的正义理论提供基础的尝试。罗尔斯认为，正义论可分成两部分，第一部分即理想部分，确立了那些在有利的环境下一个组织良好的社会的原则，即那些处理人类生活中不可避免的自然限制和历史偶然因素的原则；第二部分即非理想部分，则面对现实，主要由解决不正义问题的原则组成。他主要考虑的是理想部分，然而，他认为理想理论是非理想理论的基础，理想的正义要为怎样对待现实的不正义提供指导。所以，在他的著作中实际上体现着一种高度的虚拟性和强烈的现实性的结合，他是有感而发的，但所发并非一定通过直接诉诸现实政治问题的形式。相反，有时思辨的程度愈高，倒愈能表现时代的面貌。因此，他的思想作为他所处的动荡时代、他所处的美国社会的一种折射乃至聚光，自然会引起许多人的注目和反应。

第三个原因涉及到罗尔斯理论的一些特点。二十世纪以来，英美伦理学乃至整个哲学一直是由实证和分析传统占主导地位的。伦理学家们大都专注于从形式方面探讨道德陈述及命令的语义和逻辑关系，而不太关心紧迫的现实道德问题，自然也不齿于构筑那种形而上的、绝对的伦理学体系。这样，在某种程度上，伦理

学实际上变成了道德方面的逻辑学和认识论，以致被人讥为“冷冰冰的伦理学”。这种现象在哲学的其它一些领域内也同样存在。而罗尔斯的《正义论》的出版，则在某种意义上标志着哲学、伦理学潮流的一个重要转折：即由形式的问题转到实质性的问题，由怀疑和否定转到试图重新肯定，由实证的分析转到思辨的概括。这个转变在某种意义上可以说是对十九世纪及其以前的古典的非怀疑的哲学、伦理学传统的复归，是对康德、密尔等所代表的哲学传统的复归。罗尔斯明确地谈到：道德理论是一种描述我们的道德能力的企图，正义论即为描述我们的正义感的一种企图，与描述我们的语法感需要一种语法理论相类似，描述我们的正义感也需要涉及原则和理论结构，所以，不能高估定义与意义分析，它们在道德理论中并不占有中心地位，而是要随基本理论的兴衰而兴衰，而且，实质性地解释道德观念，反倒有益于意义分析。他说他“希望强调研究实质性道德观念的中心地位。”^①但是，罗尔斯又吸收了分析哲学的某些成果，例如，他在构造其正义论体系时努力避免独断的倾向，谨慎小心地进行逻辑、语言方面的推敲，仔细琢磨证明的方式，确立自己的有限目标，对一些重大和根本的问题存而不论，以明智审慎来代替道德结论等等。罗尔斯正义论的上述特点引起了学术界乃至一般公众的广泛兴趣。实际上，这种在哲学领域中转向实质性问题的趋势在美国迄今仍在继续，并有加强之势。

罗尔斯理论的另一个特点是：在他的作为公平的正义理论中，表现出一种试图达到全面和综合的倾向，从而使得他的理论具有巨大的理论上的伸缩余地 and 回旋空间，以致具有相当不同倾向的理论家都能从他的著作中各取所需地证明自己的观点或找到抨击的对象。他的正义论既可以满足那些仍缅怀和执着于构造某种永恒正义理论的人的思辨兴趣，又可以以为那些焦灼地面对社会

^① 见《正义论》中译本第48页。

现实中的严重不正义而绞尽脑汁的人提供某些理论根据或启发；既可以说通过强调他的两个正义原则的平等主义倾向和展示社会的理想状态，而为抵制和反对政府政策的新老左派提供了某种支持，又可以说通过强调人人都享有平等的自由权利、强调设计社会基本结构要考虑到稳定性和可行性，而为政府和右派做出某些辩护和建议。一些人认为他属于自由主义一派，另一些人则又认为他属于平等主义一派。但无庸讳言，他是试图为他所处的美国民主社会提供一个合适的、能最广泛地为人接受的道德基础；他试图发掘这一社会的活力，建立这一社会的良性循环；他直率地承认，他的正义论要通过一种反复比较、互相修正，达到与这一社会所流行的、人们所考虑和推重的正义判断接近一致的状态，并且把这种“反思的平衡”（reflective equilibrium）作为证明他的正义论的一种方式。

当然，使罗尔斯的正义论产生巨大影响的最重要原因，还是他所提出的基本观点的具体内容和理论深度，这正是我们下面所要涉及的。

二

“正义”（justice）一词的使用由来已久。在亚里士多德那里，它主要用于人的行为。然而，在近现代的西方思想家那里，

“正义”的概念越来越多地被专门用作评价社会制度的一种道德标准，被看作社会制度的首要价值。罗尔斯则更明确地规定，在他的正义论中，正义的对象是社会的基本结构——即用来分配公民的基本权利和义务、划分由社会合作产生的利益和负担的主要制度。他认为：人们的不同生活前景受到政治体制和一般的经济、社会条件的限制和影响，也受到人们出生伊始所具有的不平等的社会地位和自然秉赋的深刻而持久的影响，然而这种不平

等却是个人无法自我选择的。因此，这些最初的不平等就成为正义原则的最初应用对象。换言之，正义原则要通过调节主要的社会制度，来从全社会的角度处理这种出发点方面的不平等，尽量排除社会历史和自然方面的偶然任意因素对于人们生活前景的影响。

为此，罗尔斯通过进一步概括以洛克、卢梭、康德为代表的契约论、使之上升到更高的抽象水平而提出了他的“作为公平的正义”理论。在此，契约的目标并非是选择建立某一特殊的制度或进入某一特定的社会，而是选择确立一种指导社会基本结构设计的根本道德原则（正义原则）。罗尔斯的契约论是完全与社会历史分开的。他认为，订立契约的“原初状态”（original position）纯粹是一假设的状态，一种思辨的设计，对它可以有各种旨在引出不同结论的不同解释；我们可以合理地设置原初状态的条件，使一个人任何时候都能进入这种假设状态，模拟各方进行合理的推理而作出对正义原则的选择。这些选择是在无知之幕后（the veil of ignorance）后进行的。原初状态中相互冷淡的各方除了有关社会理论的一般知识，不知道任何有关个人和所处社会的特殊信息。这时，各方运用游戏理论中的最大的最小值规则（maximin rule）是恰当的，即选择那种其最坏结果相比于其它选择对象的最坏结果来说是最好结果的选择对象。这样，这一规则马上就排除了功利主义的选择对象，因为功利主义在产生最大利益总额（或平均数）的前提下容许对一部分人的平等自由的严重侵犯。罗尔斯认为：各方将选择的原则是处在一种“词典式序列”（lexical order）^①中的两个正义原则，第一个原则是平等自由的原则，第二个原则是机会的公正平等原则和差别原则的结合。其中，第一个原则优先于第二个原则，而第二个原则中的

^① “词典式序列”即一种编辑词典时的次序安排：比如，只有列举完所有需列举的以A为首字母的单词，才能考虑以B为首字母的单词。

机会公正平等原则又优先于差别原则。这两个原则的要义是平等地分配各种基本权利和义务，同时尽量平等地分配社会合作所产生的利益和负担，坚持各种职务和地位平等地向所有人开放，只允许那种能给最少受惠者带来补偿利益的不平等分配，任何人或团体除非以一种有利于最少受惠者的方式谋利，否则就不能获得一种比他人更好的生活。所谓“作为公平的正义”即意味着正义原则是在一种公平的原初状态中被一致同意的，或者说，意味着社会合作条件是在公平的条件下一致同意的，所达到的是一公平的契约，所产生的也将是一公平的结果。罗尔斯的正义论确实具有一种平等主义的倾向。

以上只是对罗尔斯正义论的一个极其简略的勾画，为了使读者对这一正义论的最主要观念获得一个大概的印象。具体说来，这一正义论分为两个部分：一是对原初状态的解释和对其间要选择的正义原则的概述，一是对实际选择正义原则的论证。

我们先扼要地阐述一下罗尔斯认为在原初状态中将被各方选择的两个正义原则的内容。罗尔斯认为，这两个正义原则是一个更一般的正义观的一个专门方面，这个一般的正义观是：

所有的社会基本价值（或者说基本善）——自由和机会，收入和财富、自尊的基础——都要平等地分配，除非对其中一种或所有价值的一种不平等分配合乎每一个人的利益。

而体现这一正义观的两个正义原则通过几次过渡性的陈述而达到的最后陈述则是：

第一个正义原则：每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（平等自由原则）。

第二个正义原则：社会的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（差别原则）；（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（机会的公正平等原则）。①

这两个正义原则暗示着社会基本结构的两大部分，一是有关公民的政治权利部分，一是有关社会和经济利益的部分，第一个原则要处理前一方面的问题，第二个原则则要处理后一方面的问题。

我们要注意这样一个转换：即从一般正义观的“合乎每一个人的利益”到两个正义原则的最后陈述的“合乎最少受惠者的最大利益”，这是理解罗尔斯正义论的一个关键。罗尔斯实际上总是从最少受惠者的地位来看待和衡量任何一种不平等，换言之，他的理论反映了一种对最少受惠者的偏爱，一种尽力想通过某种补偿或再分配使一个社会的所有成员都处于一种平等的地位的愿望。罗尔斯挑选出一种特殊地位（最少受惠者的地位），这不仅是理论上的简化所需要的（论证“合乎每一个人的利益”要复杂得多），而且使两个正义原则不仅仅是上述一般正义观的简单展开，更是前者的深化和倾向化，使罗尔斯的正义理论成为一种独具特色的理论。

要解释清楚“合乎最少受惠者的最大利益”，还有两件理论上的工作要做，这就是：（1）怎样鉴定最少受惠者的地位？（2）怎样衡量人们的利益，或者说合法期望的水平？对于前者，罗尔斯认为：每个人都占据两种地位，一是平等公民的地位，一是在收入和财富分配中的地位（假定权力与财富通常结为一体）；这样，确定最少受惠者可通过选择某一特定社会地位（如非熟练工人），或按达不到中等收入水平的一半的标准来进行。至于对人们的合法期望水平的衡量，罗尔斯则认为期望即等于基本社会善的指标，如果说善就是理性欲望的满足，那么基本的社会善就是一个理性人无论他想要别的什么善他都想要的善。这些

① 参见本书第2编第46节。

基本的社会善包括自由和机会、收入和财富以及自尊的基础。

但是，我们马上注意到一个情况：即我们所看到的正义原则不是一个，而是两个，不是象功利主义、至善论那样只有一个单独的最后标准（功利或至善），而是在某种程度上类似于容有一批不可再追溯的原则的直觉主义；而直觉主义的正义论按罗尔斯的话来说，“只是半个正义观”，因为它停留在一批最后标准上就认为它们是不可追溯的而逡巡不前了，不去寻找那种唯一的根本标准，这在理论上是不彻底的，且易走向相对主义。罗尔斯认为：他的“作为公平的正义理论”并不能完全排除对直觉的依赖，但他试图从几个方面来限制直觉，其中最重要的就是提出一种词典式序列来正视两个原则孰先孰后的优先性问题。他认为：第一原则优先于第二原则，第二原则中的公平机会又优先于差别原则，只有在充分满足了前一原则的情况下才能考虑后一原则，这样，就有了两个优先规则：

第一个优先规则（自由的优先性）：两个正义原则应以词典式次序排列，因此，自由只能为了自由的缘故而被限制。这有两种情形：（1）一种不够广泛的自由必须加强由所有人分享的完整自由体系；（2）一种不够平等的自由必须可以为那些拥有较少自由的公民所接受。

第二个优先规则（正义对效率和福利的优先）：第二个正义原则以一种词典式次序优先于效率原则和最大限度地追求利益总额的原则，公平机会又优先于差别原则，这也有两种情形：（1）一种机会的不平等必须扩展那些机会较少者的机会；（2）一种过高的储存率必须最终减轻承受这一重负的人们的负担。

实际上，在这两个优先的后面还蕴含着第三个也是最重要的优先，即正当（right）对善的优先。

正当与善这两个概念可以说是伦理学的两个基本概念，它们之间的关系也就成为伦理学的一个主要问题，西方伦理思想史上

目的论与义务论两大流派的分野就与此有关。目的论认为善是独立于正当的，是更优先的，是我们据以判断事物正当与否的根本标准（一种目的性标准）；正当则依赖于善，是最大限度增加善或符合善的东西，而依对善的解释不同，就有各种各样的目的论，如功利主义、快乐主义、至善论等等。义务论则与目的论相反，认为正当是独立于善的，是更优先的，康德就是义务论的一个突出代表。罗尔斯认为他的作为公平的正义理论也是一种非目的论意义上的义务论，同样强调正当对善的独立性和优先性，这在两个优先规则中已经表现得很明显了，譬如他强调自由的优先性，强调由正义所保障的自由权利决不受制于政治的交易和社会利益的权衡，不管这种有损于自由的交易多么有利或将带来的社会利益多么大；自由只能为了自由本身的缘故而被限制。

正义在某种意义上可以说是正当的一个子范畴，或者说，正义即是应用于社会制度时的正当。按罗尔斯的说法，伦理学必须包括正义论，而正义总是意味着某种平等，这等于是说，设计一种正义的社会制度就是要使其最大限度地实现平等。罗尔斯的两个正义原则也确实透露出这样一种平等乃至平均主义的倾向；他认为他的差别原则达到补偿原则的某种目的，即给那些出身和天赋较低的人以某种补偿，缩小以至拉平他们与出身和天赋较高的人们在出发点方面的差距。在他看来，天赋不是道德上应得的，应当把个人的天赋看成是一种社会的共同资产，虽然自然资质的分布只是一个中性的事实，但社会制度怎么对待和处理它们却表现出正义与否的性质。他反复申明这两个原则决不会导致一个英才统治的社会，不会导致一个差别悬殊的社会，甚至不无天真地表现出这样一种预期：倘若始终遵循这两个原则的话，未来社会的人不仅将在制度形式上保证平等，而且能够接近事实上的平等。

以上只是阐述罗尔斯的一些主要结论，而不是他的论证。罗尔斯费力最多的也许还是提供证明和使论据完善化。

罗尔斯的正义论主要是通过契约论，或更具体地说，是通过原初状态来证明的。按照罗尔斯的契约观点，证明实际上是一种深思熟虑的设计，即通过假设原初状态的各种合理的和能被普遍接受的限制条件，详细描述这一把各种对选择正义原则有意义的约束联为一体的原初状态，从而达到选择他的两个正义原则而非功利原则等。换言之，对两个正义原则的证明在于合理地说明它们将在原初状态中被一致同意地选择。证明的另一途径则是将两个正义原则与我们所考虑和推重的直觉性和常识性的正义判断和准则进行对照和平衡。

对订立契约的最初状况可以有各种不同的解释，罗尔斯把他提出来用以论证他的两个正义原则的一种对最初状况的解释称之为“原初状态”，这一解释的要点如下：

选择的对象——罗尔斯认为，难以给原初状态中的各方列出所有可供选择的对象，而只能提供一个简短的表格，表格中的选择对象包括：①两个正义原则；②古典目的论观点（古典功利主义和平均功利主义、至善论等）；③直觉主义观念；④利己主义观念；⑤混合观念。他主要比较两个正义原则与功利主义。

正义的客观环境——各方所处的环境客观上存在着一种中等程度的匮乏，不可能完全满足人们的欲望和要求，而人人又都想得到较大的一份利益；人们的利益有一致的方面，又有冲突的方面，从而使人们的合作既有可能也有必要。

各方的主观动机和理性——各方都有自己的合理生活计划或者说不同的善的观念，虽然他们不知道其内容和细节。他们相互冷淡，不受仁爱或怨恨、妒忌的推动，既不自利也不利他，他们只是想寻求自己尽可能高的收益而不考虑去损害对方。他们以建立在社会基本善基础上的统一期望和对或然性的客观解释来采取达到目的的有效手段。罗尔斯特别强调客观环境中的中等匮乏条件和主观环境中的相互冷淡条件，认为只要相互冷淡的个人对中