



## 理论文库

### 作为公平的正义：三种论辩

[美]托马斯·博格 叶明雅 译

2009-07-31

**【内容摘要】** 在罗尔斯提出并论述“作为公平的正义”理论后，引起广泛的学术争议和讨论。本文讨论了罗尔斯的政治哲学与自由至上主义、社群主义和康德哲学的关系。政治哲学这些不同思想派别之间的争论，使我们可以更深刻地理解罗尔斯的理论定位和思想贡献，理解其正义理论的思想 and 道义基础。而罗尔斯本人的世界观在更深和更广泛的意义上则是康德式的。

**【关键词】** 罗尔斯 自由至上主义 社群主义 康德

罗尔斯的《正义论》复活了政治哲学，激起人们在哲学、政治学、经济学和法学期刊上发表成千上万篇文章，这已是老生常谈。该著作已经被译成28种文字，仅英文版就售出了40万册之多（仍然保持着哈佛大学出版社的记录），并且在发展中国家找到了许多支持者，这也是人所共知。我提及这些事实，只是想由此过渡到围绕罗尔斯正义观念的三个重要的实质性辩论。

#### 一、 罗尔斯与自由至上主义

就在《正义论》出版三年以后，另一位哈佛大学哲学家诺齐克(Robert Nozick, 1938-2002)出版了与此观点对立的著作：《无政府、国家与乌托邦》。诺齐克对罗尔斯的批评之核心围绕着三个至今仍然发生诸多论辩的主要观点。

首先，诺齐克反对的是把罗尔斯的正义判断以结果论的方式引导到分配方案这一点。他发现与此相关的两个问题：为了建立任何分配方案，一个人必须推翻现存的财产权。而且，即使在这样彻底的干涉以后，还需要进一步的再分配干预，以便维护所要求的分配方案。诺齐克的结论是，像罗尔斯这样的正义观念要求极度干预主义的政府，正如在福利国家时代所典型表现的，这样的政府持续地干涉公民的自由协议。任何尊重人的自由的人，都必须拒绝任何承诺某些分配方案的正义观念。

作为对这些观念的一种替代，诺齐克提出了一种历史的正义观念。根据这种观念，一种分配方式的公正性不是取决于它接近某种理想方案的程度，而取决于它是否以一种道德上可接受的方式去发展。一个人不能从分配方案的形态中确认它是否公正。因为每一种可能的分配方案都可以用公正的方式（例如，通过一种无人欺骗的自愿的扑克游戏规则）来发展。诺齐克不仅把这种观念应用于财产的分配，而且还应用于基本权利的分配。照此来看，奴隶制只有当人们在强迫或欺骗之下成为奴隶时才是不正义的，而不是当他们自愿地放弃自己的自由权（例如在紧急情况下以此来换得某种帮助）的时候。

这一在美国赢得相当一些人支持的批评意见，包含了一种根本性的误解。作为公平的正义谈论的不是政府官员或其他人应当如何在既定的规则之下干预人们的交易，以便改善由这些交易所产生的基本益品的分配。问题在于，罗尔斯的观念针对的是这些规则本身的设计。通过提出这一问题，罗尔斯否定了诺齐克的批

评之预设前提，即我们已经知道哪些规则应该支配社会合作，以及那些现有的财产持有方式是在道义上得到证成的。罗尔斯提议，我们在将产生各种切实可行规则的分配方案的基础上来选择这些规则（我们社会基本结构的设计）。一旦确立了这些规则，就把它们当作是纯粹程序正义的一个实例来对待，恰如诺齐克会喜欢的：这些规则事先为人所知，它们所产生的分配方式都是公正的，因而防止政府的各种特设性干预或再分配方案。罗尔斯的观念把关于分配方案的信息运用于评估和构建一个社会的制度——决不干涉其运作。

这样，罗尔斯就能回应诺齐克说，我们能够接受公民们在现有规则之下自由交易的结果（而不是通过政府干预来“纠正”这些结果），而仍然可以避免诺齐克愿意赞同的可怕结局。我们避免这样的结果，不是通过干预或再分配，而是通过仔细地阐述那些支配公民交易、从而规定由此产生的分配方式的规则。一种在所产生的分配方案的基础上设计出来的制度秩序将不会导致诺齐克所赞同的可怕结局。

在罗尔斯看来，哪些规则最适合于某个特定时期的某种特定的社会，这在某种程度上是个经验的问题。这些规则也许包含固定的税收和补贴，也许安排政府在经济生活中的某种重要的角色。例如，在一个拥有产权（property - owning）的民主制度中，教育有可能是公共资金支持的，以便确保大众都能获得教育。这样一来，通过税收，富裕的公民对这一教育制度的贡献多于其从中所获得的，贫穷的公民则相反。这种再分配即使是规则所固有的，难道不仍然是一种再分配吗？

这里容易落入这样一种幻觉，即这样的社会存在两类不同的运作规则：给予人们所应得的东西、他们的毛收入的规则，以及实行再分配的纠正规则。然而，这种区分是随意的，认为毛收入反映了人们在深层意义上应得的东西这一观念也是错误的。税收制度对于毛收入有一种反馈效应，这一事实可说明上述两点。之所以如此，乃是因为在作职业决定时，人们对于自己在各种职业中能获得的净收入很敏感。例如，受医疗职业吸引的人们，会根据医生的净收入与护士的净收入之比来决定是否接受担任医生所需要的比较艰苦的训练。因此，税收制度越是累进，出现在市场上的这两种职业的毛收入间的比值就越高。举例说明：在一种无税收或固定税率的制度下，2：1的毛收入比率将导致2：1的净收入比率，让我们假定，这导致医生与护士数量的有效组合。如果采用一种级距曲线很陡的累进税率，则2：1的毛收入比率会导致低得多的净收入比，这将导致护士占多数的低效率的人员混合比。但是，这种混合比实际上不会发生，因为医疗职业的劳动力市场将通过把毛收入比率提高到大于2：1来适应很陡的累进税，从而使净收入比例接近2：1。因此，如果认为很陡的累进税率制度下的毛收入比显示了医生与护士的贡献比，那将是错误的。

毛收入也受一个社会的制度秩序的许多其他特性影响，这些特性与它们所偿付的贡献的“真实价值”没有关系。各类职业人士的毛收入部分地取决于社会各阶层的公民得以从事这些职业的难易程度。房地产的毛收入取决于支配房地产所有制的规则。继承者的毛收入受到遗产继承规则和税率的影响。从事各种行业（旅游、赌博、卖淫、娱乐、医疗、金融等等）的毛收入也受到对这些行业规约或管制的方式影响。

对于所有这些体制性选择的反思还带来了再分配的第三种含义。比如，某人认同某些特定的制度设计，但也许会抱怨说，与此相关，某些现存的制度秩序的再分配方式使得投资银行家得到较低的收入，而使穷人的孩子获得较好的学校教育。但是，这种抱怨可能并无说服力，因为任何一种可设想的制度设计都可以籍此去批驳其他的制度设计。相对于事物在其他制度设计下的分配方式而言，每一种制度设计都会“再分配”事物。如此理解之后，指责再分配是一种干涉，就失去了批判的力量。因此，为了取得进展，我们必须准确地聚焦于罗尔斯所提出的问题：相对于其他可能的制度设计，一个人如何能够证成对某种制度设计的选择？通过考察每一个候选设计将会产生的分配方案，罗尔斯提出了自己的理由。这一提议并不会因为与再分配有关而被轻易驳倒。

然而，诺齐克提出了针对此的第二个批评意见，这再度显示他认为自己的极端放任自由主义的设计是优越的。第二个批评援引了一种说法，即应当让社会的每个成员（尽可能多地）获取他或她自己贡献的价值。当然，这一原则需要得到证成。但是，首先必须澄清它的含义。一个人贡献的价值究竟指什么？例如，我们应该说，那些一生都不能工作的人的贡献为零，还是应该说，他们对运用于生产中的自然资源（最终由全体人类所拥有）做出了部分贡献？当然，阐述这一原则还有许多不同的方式，这一事实本身就大大地削弱了下述说法，即贡献的观念可以充当罗尔斯正义观的一个自然或显而易见的竞争者。

有人也许会认为，可以通过诉诸一种社会契约论的说法来克服这种不确定性。在诺齐克的第二种批评意见中，他论辩道，如果看起来有理由认可参与者们事前会协议同意的那种社会秩序的设计，那么，必须以很



不同的方式来设想这种假定的协议。罗尔斯所提出的协议是单方面的和不公平的：无才能的人获得了对他们最佳的设计，而有才能的人则必须忍受这种设计，即便他们自己单独做(把那些无才能的人排除在外)会好得多。针对这种设计，诺齐克坚持认为，如果这两个群体在一个社会秩序中进行合作，他们的合作就应当对双方都有利。为了确保这一点，一个人就必须把假设的契约情形设想为各方意识到他们的自然禀赋的情况。这会保护有才能者的利益，其方式是，当所提议的合作生产所得的份额比他们单独工作时获益要小时，可以允许他们拒绝合作。无才能的人的利益也以类似方式得到保障，尽管是在低得多的水平上。双方之间的谈判因而就会只是有关合作产品的剩余，即超出两个群体单独工作时所能够获得的价值部分。

罗尔斯对于这一异议可以通过探讨想象的非合作的情形来回应，这种情形被认为对于谈判是一种潜在的威胁。这是一种不包含任何道德约束的虚构情形，其中的一个群体能够以谋杀或奴役制来威胁另外一个群体吗？如果是，那么，给理性的各方针对这种底线所达成的协议附加任何有关我们应当如何构建自己共存条件的道德意义，就显得是一种冒犯。抑或这是一种考虑到了某些道德规则(也许是指限制暴力的使用和对自然资源的开采)的虚构情形？如果是，那么这些规则需要得到阐述和证成——这就意味着在提出一种正义观念的思想实验里已经预设了一种初步的正义观念的前提存在了。

而且，即使可以比较充分地描述这种虚构的情形，仍然不可能(即便是非常近似地)弄清楚有才能者和无才能者在其中的处境好坏。相关的虚拟情形与实际相距甚远。最后，罗尔斯还可以指出，对这两个群体间假设契约的设想是主观随意的。如果正义要求每个群体的待遇应该不会比他们单独工作时要差，那为什么他们应当是上述这样的两个群体？在所提议的替代的思想实验中，其他的分类组合都可以形成一种合作联盟，要求合作后的待遇不比单独工作时差。如果每一种可以设想的此类要求对协议设置了限制，那么(相当不同于遥远的虚拟情形)就不可能把握将会达成怎样的协议，因为可能的合作联盟的数量实在是太多了。为了说明这一点，设想一个由100人组成的相当小的社会，并且只考虑这样的场景，即把这些人一分为二，每50个人组成一个群体。即便是这个极端简单化的情形，我们仍然要考虑这50个人中数以千亿计的可能的合作联盟。

诺齐克对罗尔斯第三个重要的批评(也是被社群主义者桑德尔所采用的)与罗尔斯对其差别原则的一个误导性阐述有关。罗尔斯写道，这一原则事实上代表了“一种把自然才能的分布在某些方面当作是公共资产的协议”①、“一种社会资产”②或“一种集体资产”③。这听起来好像是，罗尔斯提出把自然禀赋看作是公共财产——好像他赞同重新分配器官，赞同强行征召有才能的人去强迫劳动或至少是对他们课以特别高的税费。诺齐克宣称，鉴于罗尔斯认为“从道德的观点看，最初的自然资产的获得……[是]任意的”④，以及“我们在天资的分布上所处的位置并不是我们应得的”⑤，他为了自治，就必须支持诸如此类的措施，或者至少是允许之。

在回应这种反对意见时，我们应该首先谨记罗尔斯关于正义的公共标准只是集中于社会基本益品的分配，而忽略所有关于自然禀赋的信息。因此，他的公共标准并没有为器官的再分配提供直接的支持。如果器官再分配或强迫劳动的方案能够促进经济收入、从而提高最低层民众的社会经济地位，那么在这一语境下，罗尔斯的标准也许提供了间接的支持。但是，诸如此类的方案由于违反了(在词典式次序上优先的)第一条正义原则，因而会遭到禁止。同样的道理也适用于诺齐克认为罗尔斯的理论会支持对特别有才能的成人征收“人头税”。这样的税收会强迫有才能的人进入产出更高的职业，也许可以提高最低层的社会经济地位。但是，这样会侵犯他们由第一条正义原则所保障的自由选择职业的权利⑥，并且由于剥夺了有才能的人从事其他人可以从事的低收入职业的机会而侵犯了机会的形式平等。

如果罗尔斯的公共标准明确地排除了诺齐克所预期的那种制度安排，那我们如何能够解释罗尔斯对于差别原则的古怪评论呢？关于上面所引的论述的一个主要观点是，社会应该从其成员中才能的分布和多样性中得益。⑦为了实现这一点，也许可以通过经济秩序设计，让有才能的人与较低才能的人相比可以为同样的努力获得更多的报酬，这样一来，让那些持有罕见或卓有成效的才能的人比拥有中等、一般或低经济效益的才能的人获得更多的报偿。诺齐克会完全同意这一得到广泛拥护的主张。实际上，我们知道的每一种经济秩序都允许报酬上的这种差别，这往往是通过市场机制，该机制有利于吸引各种有才能者从事更有成效的工作。罗尔斯理论的独特之处在于，他要求报酬上的这种差别应当安排得让处于最低社会经济地位的人得到最大的好处。大致来说，这意味着各种类型的工作的净收入应当受到这样的原则的影响(比如通过税收和补贴)，它让最低层民众获得尽可能高的净收入。

这样一来，罗尔斯式的社会中某些特殊种类的工作该如何获得报酬，这取决于许多经验因素。有多少公民拥有做类似的工作所必需的才能？与其他工作相比，这些公民优先选择这一种工作的倾向性有多高？各种可能的税收和补贴对于这些公民选择这种工作、对于他们工作的品质会有怎样的影响？这种类型的工作对于社会生产有什么贡献？对某些特殊才能所得的报酬，将取决于这种才能能够让其拥有者从事的那些类型的工作的经验因素。

有人也许会反对这样一种经济秩序，即某些才能赋予其持有者以某种特殊的价值，这种价值必须反映在报酬分配当中。（诺齐克自然不会支持这种反对意见，他自己的放任自由主义经济秩序同样是这种意见的攻击对象）对于这种反对意见，罗尔斯的回答是，才能在道德上是随意的，不应得到奖赏的。这并不意味着一个人对自己的才能不拥有权利，或者说该把才能没收或征用于公共善事。这也不意味着特殊才能的行使不应该得到报酬。这只是意味着没有一种才能可以赋予其拥有者以特殊奖赏的权利。才能在道德上是无意义的；才能与其拥有者的道德价值没有什么关系。因此，可以这样来安排对于施展才能的奖赏制度，即采取看起来是促进公共善事的最佳设计方式来进行，而差别原则将公共善事与最低社会经济地位的人相挂钩。

## 二、 罗尔斯与社群主义

自由至上主义建立在洛克和霍布斯的遗产之上，它与罗尔斯立场的不同主要在于它支持一种放任自由主义的市场经济。而美国的社群主义则是一个比较散漫的知识分子运动（受黑格尔、亚里士多德和经院哲学的影响），包含了一系列不同的实践-体制建议。这一运动的某些代表（麦金太尔[Alasdair MacIntyre]）属于政治的右翼，其他人（昂格尔[Roberto Unger]、巴伯[Benjamin Barber]）属于政治左翼，还有一些人（沃尔泽[Michael Walzer]、泰勒[Charles Taylor]、桑德尔[Michael Sandel]）则发展了另一些版本的自由主义，与罗尔斯自己的自由主义差别并不太大。

这里我将把讨论限于社群主义者专门批评罗尔斯的两点上。第一点是桑德尔批评罗尔斯时所持的核心观点（此前威廉斯[Bernard Williams]也以比较一般的方式提了出来）。罗尔斯要求社会应该首先致力于朝向一种正义的基本结构，这种主张预设了一种有关人的特定的“形而上学的”或“认识论的”观念，而桑德尔认为这是错误的、在道德上不自洽的。桑德尔通过讨论罗尔斯从原初境况这一设置得出其社会理想的方式，来支持这一核心反对意见：

原初境况是证成过程的支点，因为它是所有证成的必经之途。……一种正义理论中的一端发出的东西，必须与另一端一种关于人的理论，或者更准确地说，关于道德主体的理论相呼应。从一个方向透过原初境况的棱镜，我们看到了两个正义原则；从另一个方向我们看到了自己的反射映象。<sup>⑧</sup>

桑德尔发现，罗尔斯观念中暗含的道德主体的理论认为，人们是希望按照正义规则与他人共处的，并且为了实现这一目的而情愿削减或放弃所有其他的目标、野心、理想、成员资格、友谊和承诺信念。这样的人并没有“构成性”的承诺信念——放弃这些信念将是不可思议的。他们总能再造自身，发现新的目标、理想和忠诚对象。

人们几乎难以指摘桑德尔对这一道德主体形象的否弃。但是，罗尔斯并没有主张支持这一形象。他的正义的公共标准及其所证成的基本结构并不是喜怒无常的神明，不会出人意料地要求人们做出这种或那种牺牲。一个正义社会的公民当然能够拥有构成性承诺信念。然而，他们将以某些思想来形成这样的承诺信念：他们不会深深地卷入与黑社会成员的关系，不会与一名职业杀手结婚，并且会认为友谊有可能对他们提出的道德要求仍然处于道义的范围以内。因此，一个理想的正义社会并不意味着一群肤浅的、互不往来的公民。

罗尔斯的确不允许原初境况中的各方知道所有关于其委托人的具体承诺信念和深层奋斗目标的信息。然而，这并不是说这些构成性归属特性没有多大意义。其理由是双重的：个人是在一种特定的社会情境中获得这些归属特性的，这在很大程度上由其社会的正义观念以及由这种观念所规定的制度秩序所决定。由于公民的构成性归属特性是在社会结构当中形成的，而这些结构的性质有待于原初境况中的人们去达成协议，所以就不能预设这些结构是各方在审慎思考时已经知道的东西。另一个理由是，罗尔斯寻求一种可以向具有不同归属特性的人们进行证成的正义公共标准和与此相称的社会体制。他因而把公共标准的选择模式定为完全不取决于任何具体的归属特性。

尽管各方并不知道他们所代表的人们的具体归属特性，罗尔斯仍然告诉他们，人们通常深深地依恋这些



对他们必不可少的归属特性：“公民们在其个人事务中或者在其联合体的内部生活中……也许拥有其归属特性和爱的对象，他们认为这是他们不会或不能离开的东西；他们也许认为，如果不通过某些宗教或哲学的信念和承诺来理解自己，那将是不可思议的。”⑨罗尔斯并不把这种信念约束看作是对其批评者的让步。他在对基本自由权特别是良心和结社自由的词典式优先性的证明中需要这种信念约束。[10]因此，罗尔斯可以回答桑德尔说，他的理想社会的确为公民们的深度信念承诺留下了空间——事实上是比其他理想社会选项更大的空间。

可以肯定，甚至在最好的自由社会，在公民们的深度承诺信念与其正义责任之间也会存在冲突。在某些情况下，所要求的诉诸正义的解决方案会过于严苛。（这并不意味着一个人违反正义的责任就得到了辩护；正如威廉斯所说，某些情形是在证成范围之外的。）一个正义的社会能够在这种冲突中持续存在。只要绝大多数公民通过支持其基本结构来尊重自己的正义责任，它的正义性、合法性和稳定性就能得到保障。

桑德尔要求理想社会的公民们能得到一种共享的关于公共善的理念的鼓舞，这种理念要比对社会基本结构的正义性的共同信念更丰富。但是，他对这一理念的内容所说甚少。他明显地渴求社群——不是渴求罗尔斯所接纳的一个社会内部的那些局部社群（如“协会”），而是一个包含了所有社会成员的社群。按照罗尔斯的观点，这种理想必须围绕合理的多元主义来运行。一个民主的社会造就了大量相互竞争和不相容的价值观；要让一个社会的成员都接受同样内容丰富的社群价值观，如果没有相当大的政府管辖和权力回归，那是无法实现的。在现时代，一个由自由社会组成的庞大社群是完全不可能的。

大部分社群主义者以不同的方式提出的另一种批评意见，是以经验的社会学命题为基础的。在一个罗尔斯式起源的自由社会中，无论留给局部社群和深度承诺信念多大的空间，它们都不会繁荣起来。这一命题也得到了下述说法的支持，即在现代自由社会，不再存在任何真实的宗教承诺或深度宗教信仰的个人。友谊成了肤浅的东西，失去了如此多的价值，以致许多人宁愿一个人独自打保龄球。婚姻成了便捷的短暂伙伴关系，人们只要还能相处，就待在一起，当兴趣和爱好改变时，就换新的伙伴。我们今天生活在这样的社会里，人们在其中不再真的信奉任何事情，也许对正义的信奉除外，更多的是信奉他们自己的权利，他们也许总是会坚持捍卫这些权利，无论他人为此付出的代价有多大。

由于对现代文化的这种批判一般总伴随着某种怀乡情绪，这里应当指出的是，尽管以前的那些时代是不同的，它们也并不必定比现在好。100年前的确离婚者不多，但这也意味着妇女们经常不得不无法言表地痛苦忍受虐待她们的丈夫。的确也有些怀有更深层宗教信仰的个人，但他们也往往受到教士、牧师的剥削和欺骗。这并不是否认一个自由的社会秩序也许包含了某种价值观的丧失。正如罗尔斯（在以赛亚·伯林之后）乐于指出的，“不存在无损失的社会秩序”[11]。在每一个社会，某些生活形式将会完全消失，许多其他形式将无法繁盛。面对这样的事实，我们被迫作出抉择。罗尔斯认为，在我们现有的理解状况的基础上，我们应该至少为我们的社会选择一种平等主义的自由秩序。

这并不是让步说，罗尔斯在支持这种选择时，愿意接受有价值的生活形式很大的损失。罗尔斯决不是现状的辩护士，而是坚信他所预想的这个社会在社群价值方面可以比现存的自称为自由的社会做得好得多。通过维护政治自由权的公平价值，他的社会将吸引贫穷的公民们进入关于正义和公共善的联合公众慎议。通过维护机会的公平平等，这个社会将极大地改善社会流动性。通过满足差别原则，它将减少现有的工资率上的不平等，从而增加可以支配的自由时间，特别是对人口中较贫穷的阶层。将这些原则恰当地制度化，罗尔斯能够寄希望于公民之间真正的互惠和公民友好意识，这又反过来支持洪堡关于各种社会联盟的总社会联盟思想。[12]

### 三、 罗尔斯与康德

罗尔斯将其正义观表述为具有重要的康德主义特征，这些偶尔的努力除了让人（尤其在德国学者当中）不以为然外，也吸引了诸多关注。从一些杰出前辈那里，罗尔斯博采众长，将他们的思想吸纳入自己的理论。康德在这些前辈中占据了一个特殊位置。无论在讲座还是在出版物里，罗尔斯都大量地提到康德，而不是其他人。罗尔斯一直在不停地重新思考并重新阐述他跟康德之间的理论联系。

罗尔斯对前人思想的借鉴，是通过将作者的思想从其写作的语境剥离出来。例如：他从亚里士多德那里

得出亚里士多德原则，从霍布斯那里得出假设的与历史无关的社会契约理念，从洛克那里得出关于自由主义宽容的思想，从休谟那里得出正义环境，从卢梭那里得出关于民主和道德教育的理念，从边沁和马克思那里得出对于社会制度的关注，还从马克思那里找出对尤其是政治自由的（公平）价值的兴趣，从密尔那里得出对于思想和良心自由的论证，从西季威克那里激发出他的反思平衡的灵感，最后，从哈特那里得出对许多重要概念所作的区分，特别是对自然责任(natural duty)与制度义务(institutional obligation)的区分。

然而，从康德那里，罗尔斯不仅仅是学习和借鉴。康德的作品(有趣的是，与其说是后期的《法的权利科学》，还不如说是《道德形而上学原理》和《实践理性批判》中的道德哲学)决定性地启发了罗尔斯自己的研究方案。接下来，我们就对罗尔斯和康德关系的几个方面作一番细密的考察。

罗尔斯第一次具体阐述他跟康德的关系是在《正义论》的第40节。他在此处对自己的理论提出了一个“康德式解释”。其基本思想是：原初境况中各方的慎思与一个具有善良意志的个人的慎思是一致的，后者用康德的定言命令(categorical imperative)来检测自己的行动准则。当然，这种一致性并不是严格的。康德描述了有关行动如何产生否定结论（拒绝某些行动准则）的一元逻辑的反思。罗尔斯则简略地描述了有关导出一个积极结论（应当满足的一个具体的正义标准）的对制度安排的集体慎思。尽管如此，罗尔斯试图（用一个更加技术性的工具，在更精确的程度上）准确地表明（按照他的解释）康德也想表达的内容：像这样的理性的个人如何做出道德决定。像康德一样，他“开始于这样的想法，即道德原则是理性选择的目标”[13]。

最终要表明的康德式解释是：“对原初境况的描述类似于本体自我的观点，它意指自由而平等的理性存在物。……于是，原初境况可以被看成是，在一个经验理论框架内对康德的自律观和定言命令所做的程序性解释。”[14]罗尔斯试图在两个方面将对原初境况的设置与康德的自律概念联系起来。“我相信，康德认为，当一个人所选择的行动原则是作为对其自由而理性之存在物本质的最充分可能的表达时，他是在自律地行动……现在，无知之幕后原初境况中的人们无法得到那些能使他们选择他律原则的知识。”[15]此外，“动机所做的相互无利益牵扯的(mutual disinterest)假设，类似于康德的自律概念”[16]。

罗尔斯还试图将他的正义原则与定言命令关联起来：“正义原则还类似于定言命令。因为，康德是借助定言命令来理解一个运用到其本质是自由而平等之理性存在物的人身上的行为原则的。……对两条正义原则的论证并没有假设各方抱有具体的目标，而只是假设他们欲求某些基本的益品。这些都是一个人所合理欲求的，无论他想要其他什么东西。因此，鉴于人之本性，对它们的欲求是理性人的一部分……”[17]正如奥利弗·约翰逊(Oliver Johnson)最早表明的，这一类比是失败的。这尤其明显地见于罗尔斯的这段话：“我的提议是，我们认为原初境况在重要的方面与本体自我看待世界的观点相类似。作为本体自我的各方，拥有充分的自由去选择他们所希望的任何原则；但是他们还具有用这种选择自由将其本性表达为理智王国自由而平等之成员的欲望……”[18]于是，罗尔斯想说，当我们用在原初境况中将要选择的正义标准来调整自己的行为时，我们由此表明，像各方当事人一样，我们是作为理想型的自由而自律的本体自我而受到激励和对待的。但是，这一说法与罗尔斯从狭义的经济意义上将各方的合理性描述为“手段—目的”的合理性的做法并不一致。[19]对尽可能多实现各种目的的手段之欲求，与将一个人的本性彰显为可知世界中有理性的成员的欲求，并不是同一个欲望的两个不同方面。追求既定目标时的审慎的理性并不是康德所指的理性，尽管罗尔斯在这一段中，使用了两种意义上的“理性的”。（罗尔斯在《正义论》的其他地方表明自己非常清楚“理性的(rational)”与“合理的(reasonable)”之间的区别，并对此做了明晰的区分。[20]原初境况中的各方是进行最大限度计算的理性主体，而非自由、自律的本体意义上的主体。

人们可以尝试利用斯蒂芬·达沃尔(Stephen Darwall)所建议的对康德的一种修正性解读来了解这一批评。达沃尔为《正义论》第40节的结论做了辩护：“通过他们在原初境况中会接受的方式来行动，人们展现出他们的自由、他们对于自然和社会之偶然因素的独立性。”[21]但是，他并不赞成罗尔斯得出这一结论的方式，也没有把各方表达为自由、自律和理性的。达沃尔所作的修正——后被罗尔斯采用并做了进一步发展[22]——将各方只是合理的自律与现有或理想社会中实际的公民的充分自律区别了开来。它的基本思想是：在展望一个正义的社会时，我们受康德的指导，不应该允许自己受任何个人的爱好或利益的影响。我们可以这样来遵守此条禁令，即询问我们自己，如果忽略自己的具体倾向，我们会如何确定正义。罗尔斯可能这样来回答：我们所展望的正义社会将使它的成员有能力实现他们作为人的利益（这可以通过对成功实现个人选



择的具体的善观念的兴趣表现出来)。

但是，这一回答真的拯救了他与康德之间的关联吗？诚然，利用原初境况之思想实验来确定正义，由于无知之幕的缘故，确实将我们与自身的偏好及个人的利益切断了开来。但是，我们的反思仍然是与一般人的需要和利益联系在一起的，《正义论》所规定的第三个基本利益以及罗尔斯的基本益品的观念反映出了这一点。于是，人们可以指责作为公平的正义是康德意义上的他律的一个范例——并不是利己主义者的他律（这些人的正义观受其个人欲望的影响），而是结果论者的他律，其正义观由所有人的欲望造就。

针对这项有关他律的指责，有三点可以为罗尔斯的观念作辩护。第一，在实现需要和利益时，利己主义与不偏不倚的关注之间存在重要的区别。原初境况中的各方所追求的是一个所有公民都勤奋努力的社会，因为这是每一个人确保自身（或其委托人）过上好生活的唯一途径。人们因此也许会说，各方的动机是利己主义（或特殊主义）的。但是，进行思想实验的我们却不能这样说。我们并不是从利己的动机来认可各方在原初境况中将要达成的协议。相反，我们之所以认可它，乃是因为，它不偏不倚地考虑了所有公民尤其是境遇最差者的利益，从而契合了我们经过思量后对于公正的判断。这种对他人利益的公正无私的关心，是康德所认可的。例如，他主张，我们在道德上都有责任来促进他人尤其是最需要帮助的人的幸福。于是，当我们预备帮助别人而询问自己何种帮助将最能实现他们的需要和利益时，我们并没有与康德相抵触——即使我们将这一问题作为“自私的”假设性问题提出来：他们将合理地希望我们为他们做些什么？类似地，当我们着手组建自己的社会而询问自己，何种社会秩序将最能实现其公民的需要和利益时，我们也没有与康德相抵触——即使我们将这一问题作为“自私的”假设性问题提出：这些公民将合理地希望从他们的社会中得到些什么？正如罗尔斯可能提出的那样，尽管处在我们的思想实验中的慎思的各方只有理性自律，但是，进行这项实验并在行动中尊重其结果（不考虑我们个人的利益和价值观）的我们，却由此而展现了我们充分的自律。

[23]第二，一方面是关注爱好（幸福）的满足，另一方面是关注对需要和利益的实现或根据一种善观念而追求成功，这两者之间存在着一个重要的区别。诚然，一些公民也许宁可牺牲社会的基本益品而追求他们自身的幸福。但是，许多公民将追求截然不同的善观念，或许是知识和文化，或许是爱、家庭和友谊，或许是环境保护，或许是体育锦标，或许是艺术成就，或许是某些宗教项目。于是，一个罗尔斯式的社会之要旨不在于满足公民的爱好，让他们快乐，而是使他们有能力获得他们自由选择的目标（该目标的内容和质量由他们自己负责）。一言以蔽之，罗尔斯式的社会秩序的要旨不在于个人的幸福，而是他们的自由。这一点当然反映在基本自由权及其优先性上，但是，它也反映在其他的社会基本益品上，这些益品被看成是公民实现自己自由选择的几乎任何目标的通用手段。

第三，虽然罗尔斯的基本益品是适用于当前的人类世界，而不是任何可想象的理性存在物的世界，然而，人们还是可以轻而易举地给予这样的辩护，即，它们展示了一个更加抽象的定言命令：正义要求社会制度应当安排得让其参与者在最大程度上实现其自由选择的目标，而理性存在物之间的交往活动就应当由这样的制度来组织。这一步（罗尔斯本人很可能弃而不用）将把罗尔斯的观念表述为一个运用到人类世界具体条件中去的一个更加普遍的正义观念。

在我们得出康德式解释是可辩护的这一结论之前，我们必须讨论最后一点重要的差异，这是由奥特弗雷德·赫费(Otfried Hoffe)第一次突出强调的。康德在他的实践哲学中区分了权利学说和美德学说，前者关注社会的恰当组织方式，后者关涉个人的道德行为。在后一个领域内，动机扮演了重要的角色：一个行为是否具有道德价值，主体是否具有善良意志和自律，均由动机来判断。然而，在权利学说中，动机却被排除了。这里惟一关注的是外在的自由，更确切地说，是人们相互保障了外在自由的领域。为了保证这一点，人们必须遵守法律对其行为所施加的限制，但是，他们如此顺从的动机就与法律无关了。于是，这个差异在于：尽管康德将对动机的讨论排除在权利学说之外，罗尔斯还是长篇大论地对它作了讨论，尤其是这样一个问题：正义观及它所证成的制度安排，是否以及如何能够在公民身上产生出一个有效的正义感和政治美德。

罗尔斯之所以对动机感兴趣，乃是由于他认为，政治哲学的任务是界定一个不仅是正义的而且还能在世上长存下去的社会秩序。我不想说，康德对这个问题不感兴趣。但是，除了建议一个历久长存的社会秩序需要依靠国家的警力以及武力的威胁和运用之外，康德并没有对此给予过多的思考。罗尔斯发现这个回答在两个方面是不足的：只依靠武力不但在经验上不充分，在道德上也是可被拒绝的。

罗尔斯所看到的这种经验上的不充分，并非出自一般的违法行径；对违法行径，罗尔斯也愿意依靠警力

来解决。问题在于：那些在民主社会中占据有影响地位的人，有可能通过官方的和合法的政治程序来影响制度安排，从而有利于他们自己的利益和价值观。凭借多出的200年历史经验，罗尔斯比康德更清楚地看到这样的危险，即政治的（包括宗教或哲学的）党派或联盟之间由规则支配的竞争，有可能堕落为完全围绕竞争规则本身展开的倾注全力的权力斗争。

罗尔斯比康德更为严肃地对待这一危险，他可能还受到另一事实的触动，即他所展望的社会的平等主义特征尤其容易受到这个问题的威胁。康德（名声不佳地）认为，他所认可的公民平等与收入及后天的财富和特权的最大不平等是完全一致的。相比之下，罗尔斯要求制度安排应当设计得能让处于社会经济地位最低的人得到最优的安排。比康德更严重的是，罗尔斯因此而受到一个来自他的理想社会的历久问题的挑战：由于这个社会在政治上最有影响力的公民（他们在自然禀赋、教育和经济手段上占据优势）将运用自己优越的影响力来冲淡和损害（符合基本政治自由的公平价值、机会的公平平等和差别原则的）社会秩序中的平等主义要素，这个社会难道不会注定要失败吗？罗尔斯的否定性回答依赖于公民拥有一个有效的正义感，这种正义感使他们在道德上恪守作为公平的正义，以及这一观念所选择的制度安排。

此外，罗尔斯发现，一个公正的社会秩序应该不仅靠制裁来维持，还应该并主要地依靠一个被共同正义观所激励的广泛的正义感而支持，如此，在道德上才是可取的。当然，这并不意味着，这样一种正义感应该通过法律来强制；那些仅仅是因为害怕惩罚而纳税的人不应该因此而受到惩罚。它仅仅意味着，我们有理由选择一个依靠在公民之间激发一种广泛有效的正义感来维持的社会，而不是通过一个威胁和引诱的狡猾机制来维系的社会。也许在一个恶魔社会中也有某些运行很好的可以接受的制度性安排设计。但是，有一个不需要用恶魔就运行很好的社会秩序，才是更可取的，因为它在其公民身上激发了他们在政治行为中展示正义、公平和教养的道德愿望。

虽然罗尔斯本人并没有讨论他与康德之间的所谓的分歧，但是，我们有理由承认它，并把它看成是有助于罗尔斯的一点。罗尔斯试图提供一个康德式的解释，而不是对于康德的解释。如果罗尔斯的观念能表明在重要的方面符合康德的理念，而且即使有分歧也是有好理由的，那么，这样一个康德式解释就成功了。

在《正义论》面世之后，罗尔斯就他跟康德的关系做了些增添和删除。康德出现在罗尔斯70年代一篇论文的标题中，以及他的三次杜威讲座（这是罗尔斯在《正义论》和《政治自由主义》之间最重要的出版物）的标题里。这些杜威讲座的后来版本（发表在《政治自由主义》里）删去了康德的名字，因为罗尔斯从“康德式”建构主义走向了“政治”建构主义。

我们可以将这些改变理解如下。一方面，罗尔斯希望他的观念与康德实践哲学之间呈现出一种不对称的联系。他让好的康德主义者确信，他的观念是对康德价值和方法的可靠而引人注目的发展。但是，他同时也想让他其他读者相信，他们可以不需成为康德主义者就可以接受他的观念。这是对罗尔斯观念的康德式解释，但是，它并不意味着，就不存在其他对罗尔斯观念的可靠解释——例如天主教的解释。尤其是罗尔斯想对康德的两个理论保持中立：他不想肯定或否定人们应该努力实现意志自律。此类有关“是什么使得人生值得过”的问题，应该留给公民个人去从他们认为合适的角度来回答。（为了公平起见，有人肯定会在这里补充说，康德会在他的政治哲学里说同样的话——只是罗尔斯没有认识到的这一点。与罗尔斯不同的是，康德在他的政治哲学之外发展了一个道德哲学。但是，这并不意味着，前者就依赖于后者，或是后者的一个组成部分。）此外，罗尔斯并不想肯定或否定，道德价值和原则在人类理性中具有它们的源泉。他的正义观应该也可以被这样一些公民所接受，即他们认为，最佳生活是最令上帝满意的生活，或者相信，道德价值和原则在它们的存在或有效性上独立于我们。

然而，另一方面，康德的实践哲学对罗尔斯来说不仅仅是诸多广包性世界观中的又一个，而且，对他个人来说，也是它们当中最重要的一个。我会认为，与罗尔斯的政治哲学相比，罗尔斯本人的世界观在更深和更广泛的意义上是康德式的。

注释：

① John Rawls, *A Theory of Justice*, revised edition, Cambridge, MA: Harvard University Press 1999, p. 87.

② *Ibid.*, p. 92.



- ③ Ibid., p. 156.
- ④ Ibid., p. 274.
- ⑤ Ibid., p. 89.
- ⑥ Ibid., pp. 228, 335.
- ⑦ John Rawls, *Justice as Fairness: A Restatement*, ed. Erin Kelly, Cambridge, MA: Harvard University Press 2001, pp. 75 - 76.
- ⑧ Michael Sandel, *Liberalism and the Limits of Justice*, Cambridge: Cambridge University Press 1982, pp. 47 - 48.
- ⑨ John Rawls, *Collected Papers*, ed. Samuel Freeman, Cambridge, MA: Harvard University Press 1999, pp. 331 - 332.
- [10] John Rawls, *A Theory of Justice*, § 32; John Rawls, *Political Liberalism*, New York: Columbia University Press 1993, pp. 310 - 315; John Rawls, *Justice as Fairness: A Restatement*, pp. 104 - 105.
- [11] John Rawls, *Collected Papers*, p. 462.
- [12] John Rawls, *A Theory of Justice*, § 79; John Rawls, *Political Liberalism*, pp. 320 - 323.
- [13] John Rawls, *A Theory of Justice*, p. 221.
- [14] Ibid., pp. 225 - 226.
- [15] Ibid., p. 222.
- [16] Ibid., p. 223.
- [17] Ibid., pp. 222 - 223.
- [18] Ibid., p. 225.
- [19] Ibid., p. 125.
- [20] Ibid., pp. 16 - 17.
- [21] Ibid., p. 225.
- [22] John Rawls, *Collected Papers*, pp. 303 - 322.
- [23] John Rawls, *Political Liberalism*, pp. 72 - 79. [ZK]]

(来源: 《马克思主义与现实》2009年第3期) (译者单位: 南京大学政治学系)

(网络编辑: 胡毅)

上一条  
下一条

重叠共识对政治秩序的公共证成

[网络链接](#) | [联系我们](#)

版权所有: 中央编译局

地址: 北京市西城区西斜街36号 邮编: 100032

设计制作: 文献信息部信息技术处