

Globethics Repository



正义原则的证明方法 [Proof of Justice Principle]

This page was generated automatically upon download from the Globethics Repository.
More information on Globethics see <https://www.globethics.net>. Data and content policy of Globethics Repository see <https://repository.globethics.net/pages/policy>.

Item Type	Book chapter
Authors	何, 怀宏
Publisher	山东人民出版社
Rights	With permission of the license/copyright holder
Download date	2026-07-12 23:03:11
Link to Item	http://hdl.handle.net/20.500.12424/185122

何怀宏：正义原则的证明方法

何怀宏

罗尔斯的《正义论》整本书都可以说是证明。罗尔斯不遗余力所要从事的工作也是证明。类似的结论可能有不少人都能提出，甚至提法更形象、更新颖、更有号召力。但要更准确，更全面、尤其是更有理性上的说服力，则非一种艰巨的学术理论工作所莫能为。罗尔斯所要证明的并不是新奇的、革命性的理论，而在某种程度上是相当具有普遍性的道德常识，他也正是要寻求某种道德共识，但这并不就减少了证明的困难，甚至在某些方面可能更难，因为他不仅必须注意原则的微妙之处，必须在那些看似不起眼的岔路口分析日后将可能变得深远的分歧（如一种“作为公平的正义”的道德契约论与功利主义的分歧），还必须把自己的观点在“反省的平衡”上同时诉诸而又修正社会流行的道德常识。罗尔斯的证明有时看起来是相当“绕”和琐碎的，需要我们相当的耐心，但它却正可以限制那种根据印象和直觉的评论，以及言论上凭借机智和聪明的“信口开河”，这类评论和言论可能仍对学术有刺激作用，却终非学术的大道。而作为中国学者，我们尤其有必要培养这种耐心。一般的读者也许可以略过这一章不读，但学者却最好不要如此。总之，只要我们细心体会，我们将能够从罗尔斯的证明方法学到很多东西。

本章先从条件设计和推理过程两方面讨论罗尔斯正义原则的证明方法，并探讨了正义论与历史上的社会契约论的联系，这种联系即主要表现在把契约论作为一种证明方法。最后我们想讨论一下对理性方法的反省。

一、原初状态的设计

原初状态相当于社会契约论中的自然状态，但它并不是被设定为历史上的存在，而只是在思维中的存在，历史上的存在若作为社会与国家起源的普遍解释可能只有一种，而这种思维的设计却可以有多种，它很像一种思维的试验，设计者不时在天平上称称各个概念的份量，在试管里加进一些成分，然后再试试看是否会发生什么反应，是否能得到自己希望的结果，如果不理想，便改变一些条件再试，就这样，一直到获得合理和理想的结果为止。自然，在此我们看到试验者还是要受到一定价值目标（或理想结果，在罗尔斯这里常常是他所谓的“被考虑的判断”）影响和规定的，要不时把自己的设计及结果与这一价值目标对照，另外，整个试验本身的程序和条件设计，也应当是合理的，不至于相互抵消、相互矛盾，而是要相互支持、相互印证。

罗尔斯的原初状态也就是这样一种思维的试验和理性的设计，它的结论是开放的，容易有修正的，而不像在黑格尔的三段论体系中每一论点乃至概念都是作为整个体系中不可缺少的一环，作为绝对真理出现。在罗尔斯这里，目标只是要设计一种订立原初契约的状态，描述订约各方的信仰、利益、知识、理性、相互联系、所处环境、可供选择的对象等等各种条件，并把它们视为一体，统称之为“原初状态”。这样的设计肯定可以有多种，但是否它们之间就没有高下之分，没有较可取与较不可取之别呢？不是这样的。而衡量高下，可取与否的标准，在罗尔斯看来，就是要看它所假设的条件是否合理，是否能够被人们广泛地接受和同意，设计应当从共同享有和广泛承认的前提出发，而不是从较专门的、带明显倾向性的薄弱前提出发，这是罗尔斯的一个基本思想。每个假设条件本身都应当是自然的，看来有道理的，不与人们的直觉明显冲突。罗尔斯的原初状态的目标就是要将对可接受的正义原则有意义的约束联为一体，从而排列出一些主要的传统的社会观念的次序，并选择那看来最合理、优点最多的正义观作为社会基本结构的正义原则（当然不是照搬，而是要重新进行概括和改造），从而以此来决定各方所要达成的社会联合的基本合作条件。

现在我们就来考虑罗尔斯的原初状态，广义的原初状态亦包括其间的选择和推理过程，但我们在本节中仅限于静态的考察。

1. 正义的环境

现在我们来假设这样一群个人（记住我们总是在假设的层次上进行），他们关心和追求自

己的利益，现在由于碰到个人所克服不了的障碍（卢梭的理由），或看到社会合作将产生巨大的利益（罗尔斯的理由），而想联合为一个社会，现在假设他们要来确定联合的条件即订立某种契约，最好的设想当然是每个人都能得到他想要的一切，但这是不可能的，仅仅别人的存在就阻止了这一点。别人不会同意大家都一起来推进他的利益，实现他的有关什么是“好”的观念；那么现在降格以求，即要求别人都正当地行动，既不干涉和妨碍他，还承担某种合作将带来的义务和负担，而他却可以任意豁免自己，像逃票乘客一样免除自己的义务，别人无疑也不会同意这种联合的条件。那么，他们将选择什么样的原则作为他们联合的基本条件呢？

原初状态的设计就是为了解答这个问题。这样，所选择的正义原则就应当是每个人对别人的相应要求的最好回答，就意味着某种妥协或平衡。

正义原则将从正义的环境”中引出，或者说，它们作为正义问题的解决办法，将需要首先描述使正义问题得以产生，从而使正义原则成为必要的环境，换言之，那群欲进入一种社会联合的个人，他们是处在一种什么样的环境中呢？

“正义的环境”是这样规定的：在那里聚集着这样一些条件，使人类合作成为可能和必需。人类社会同时具有利益一致和利益冲突的特点，利益的一致在于人们的社会合作能使所有人都过一种比他们各自努力、单独生存所能过的生活更好的生活，利益冲突是由于人们谁也不会对怎样分配由他们的合作所产生的较大利益无动于衷，因为对他们所追求的目的来说，不管这些利益是什么，较大份额的利益总比较小份额的利益要好。这样就引出正义的必要性——需要一些原则来决定能够恰当划分出利益的社会安排。换句话说，利益的一致引出人类合作之必要性和可能性，而在合作中利益的冲突则更加强了正义原则的必要性。

但是，人们有利益一致的一面，人们有对一种互惠的合作的共同愿望虽然是引出人类合作之必要和可能性的一个前提，却还不是充足的条件。人类合作之所以可能而且必要，还存在一些主客观方面的条件，客观条件是：第一，众多的个人在确定的地理区域内生存，他们的身体和精神能力大致平等，差别不大，没有那种能凌驾和压倒所有人的超人，每个人都是易受攻击的，都有某种脆弱性：每个人的计划都易受到其他人的合力阻止而归于无效或失败。

第二，许多领域中都存在一种中等程度的匮乏，自然资源和其他资源并不是非常丰富以致合作的计划成为多余，同时条件也不是那样艰难，以致有成效的合作也终将失败。也就是说，人们的需求和可能的提供之间存在差距，但又不是极度匮乏。

至于涉及正义环境的主体方面的条件，罗尔斯认为：主体各方有大致相近的利益的需求，从而使相互有利的合作在他们中间成为可能；另一方面他们又各自都有自己的生活计划，这些歧异的计划（或者说“好”的观念）使他们抱有不同目标，造成利用自然与社会资源方面的冲突要求。在此，罗尔斯特别假设各方对别人的利益是不感兴趣的。即相互冷淡

（mutually disinterested）。各方珍重自己的“好”的观念，认为这一观念是值得追求的，它所提出的要求是应当满足的，但这并不意味着各方是利己主义者，只关心自己的某种利益和财富、权力、威望；而只说他们在此时对他人的利益是不感兴趣的，他们是否是利己主义者要依他们的“好”的观念（或生活计划）的具体内容而定，如果其内容是个人财富、权力和威望，那他们确实是利己主义者，然而他们的“好”的观念也可能是共同幸福或助人为乐。而这些具体内容在此都没有被规定。我们只知道这些人各自抱有自己的“好”的观念，各自都有自己合理的生活计划（长期计划），但并不知道这些观念或计划的内容。总之，相互冷淡的假定意味着各方一方面不是仁爱和无私的利他主义者，总是去考虑照顾别人的愿望和满足别人的“好”的观念（如此就不会有正义问题出现），另一方面他们又不是追求个人特殊利益的利己主义者——如此也达不成普遍同意的正义原则。“相互冷淡”可以说是正义环境的一种基本的动机假设。

另外，罗尔斯还假设了处在正义环境中的各方在知识、思想、判断方面是有缺点的，即他们的知识是不完全的，推理、记忆力和注意力总是受到限制，判断易被渴望、偏见和私心歪曲，正是这些缺点不仅造成了人们“好”的观念的歧异，而且引起了哲学、宗教信仰、政治和社会理论方面的一致。最后，罗尔斯还假定各方都知道他们是处在这样一种正义环境中，并把它看作是理所当然的，假定各方都表示出对下一代的关心，至此，背景的假设就基本完成。

为简化起见，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各方对中等匮乏条件下社会利益的划分提出了相互冲突的要求，正义的环境就算达到了。

在上述条件中，罗尔斯大致遵循了休谟、哈特和卢克斯的观点，哈特在其《法律的概念》中提出的人性观五要素是：(1)人的脆弱性；(2)大致的平等（指体力和智力）；(3)有限的利他主义；(4)有限的资源；(5)有限的理解。卢克斯在《政治学原理》中对人性的五点概括稍有不同：(1)某种相互作用；(2)某种共享价值；(3)不完全的无私；(4)易出错的判断；(5)不完全的信息。至于休谟，我们可以注意他的这样一段话：“正义只是起源于人的自私和有限的慷慨，以及自然为满足人类需要所准备的稀少的供应。”

2. 无知之幕

罗尔斯在原初状态中设计无知之幕的目的在于体现公平，使原初状态的条件和程序显得公平，从而所达到的结果（正义原则）也是公平的。因为，事实上各个人的情况是千差万别的，天赋、出身、地位、利益、志趣、愿望，都各不相同。在此，罗尔斯实际上提出了一个重要的假定前提（未经证明的）：在选择有关社会基本结构的正义原则时，任何人都不应当因其或社会背景的缘故而得利或收益，要使所选择的原则成为公平合理的，普遍同意的，就不能允许把原则剪裁得适合于个人的特殊情形，不能允许它们受到特殊的爱好、兴趣及个人价值观的影响。而且，在这方面我们不能指望订约各方作为公正的仲裁人出现，因而我们只能通过限制人们对自己特殊情况和利益的知识来达到这一点，从而使将选择的原则摆脱个人好恶的影响，避免无休止的纷争。同时，这也意味着排除各种各样的来自自然和社会的偶然因素的影响，把各方置于一种公平的条件之下，地位平等，被作为道德人同等地对待，换言之，世界上的偶然性必须通过调整最初契约状态的环境来纠正。以上这些，可以说都是提出无知之幕的理由。

那么，无知之幕的范围有多大呢？或者说各方不知道哪些事实或情况呢？

1. 无人知道他在社会中的地位，他的阶级出身，也不知道他的天生资质如何，他的体力和智力怎样，即他对个人自然天赋和社会出身一无所知。

2. 也没有人知道他的“好”的观念的具体内容，他的合理生活计划的特殊性，甚至不知道他的心理特征：像是否喜欢冒险、气质是乐观还是悲观，即对个人价值观念和气质性格的无知。

3. 各方也不知道这一社会的经济或正当状况，或者它能达到的文明和文化水平，以及他们所处的特定社会和所属的世代。

那么，他们知道一些什么呢？

1. 他们所知道的唯一特殊事实就是上文所说的，他们知道自己所处的社会是一种正义的环境；即在此人们的合作是可能和必须的，他们有利益一致的一面，又有利益冲突的一面；客观上存在中等程度的匮乏，主观上各方都关注自己的“好”的观念而对别人的利益不感兴趣。

2. 他们知道所有人类社会的一般事实，理解政治事务和经济理论原则，知道社会组织的基础和人的心理学法则。在一般的信息方面，没有任何限制，他们知道所有可以用于选择正义原则的一般事实。

这一设计看来十分怪异，不合常理，使人难于把握原初状态的含义。但罗尔斯认为在思维中进入这一状态并不困难（它当然不是一个事实——无论是过去有过还是未来可能有的事实），实际上这就意味着依据一种一般的观点，一种排除了特殊因素干扰的观点所进行的思考。在此，由于无知之幕后的各方都是同等理智和境况相似的，每个人都是被同样的论证说服的，实际上就是每个人为所有人选择，而所有人又都像一个人那样选择。各方的思考都是相似的，实际上就等于一种完全排除了个人因素的反思过程。如此也才可望达成一种一致同意的契约。正是无知之幕使一种对某一正义观的全体一致的选择成为可能，否则，订立契约的问题就将变得无比复杂，即使存在着一个答案，也无法达到它，或者只能在一些琐屑的问题上达到一致。

在此，我们注意到，这实际上已不是达成一种契约，而是达成一种普遍性的观点。对原则的选择已经不是一种个人的选择了，而是一种把所有人融为一体的一个抽象的理性的人的选择，这一选择通过无知之幕把所有人的特点隐去，而突出了人的工具理性和合理慎思。

罗尔斯认为，他的“相互冷淡+无知之幕”的假定胜过“仁爱+知识”的假定，因为不仅这种条件的结合迫使原初状态的每个人都考虑到别人的立场（即同时也为所有人选择），达到一种确定的一致同意的正义原则；而且，它还有简洁和清楚的优点，本身是一个弱条件。

3. 主体条件假设：自由、平等、有理性的各方

前面我们谈到了与正义环境直接相关的、选择各方主体方面的一些条件，如利益需求大致

相近，能力大致相等，相互冷淡等，现在我们继续讨论主体的其他条件假设，简单地说是这样的：各方是自由、平等、有理性的存在，或者说：各方必须是自由人、道德人和理性人。我们现在来分两步说明这一假设。

首先，各方是自由、平等的存在，这说的是一种地位。在此，自由比较好理解，如果没有自由的条件，各方不能按自己的心愿进行选择，也就不可能产生订立契约的行为了。平等的根据要复杂一些，为什么说各方享有平等的建议、推荐和选择权呢？在此是否有结论的因素渗入了前提呢？

我们前面已说过，罗尔斯认为平等的根据在于各方的道德人格，这种道德人格表现为两个方面的能力：第一，作为道德主体，他们具有一种形成他们的“好”的观念的能力，能够有一种前后一致的目的体系，有稳定的对他们自身的利益和幸福的理解；第二是他们有能力获得一种正义感，即有能力获得一种在正常情况下有效地理解、应用和实行所采用的任何正义原则的欲望。人的这两种道德人格能力就构成平等的基础，哪怕它们只是潜在的能力也是如此，而且它们可以说是唯一的基础，是一个充分的条件，原初状态中各方对选择正义原则的平等权利有这个条件作为根据就足够了，不再需要其他的根据。这种道德人格能力是使一个人成为平等的权利主体的充足条件。自然，对这种道德人格能力不能做狭义、严格的理解，而且我们谈的是潜能，不是实现。这种能力就可以说是人所共有的，几乎囊括了所有人。

这里值得注意的是这样两点：第一是对人们的“好”的观念、目标体系或生活计划并没有排出高下和优劣。在无知之幕后我们并不知道它们的具体内容，但即使撤去无知之幕，也不能排列人们的“好”的观念的次序，并据此有差别地对待持不同“好”观念的人。能够形成“好”的观念（加上起码的正义感能力）就足以构成平等对待的基础了，而无须考虑“好”观念的内容，第二，对正义感的要求也是起码的，这是指对一种最低限度的正义感能力的要求。

罗尔斯对各方是平等者的假设是为了表现原初状态的公平，这一点我们下文还要涉及，现在我们转到各方的理性的条件假设。

原初状态中的各方都是有理性的，他们的理性是和他们的“好”的观念结合在一起的，不仅体现为他们能形成自己的“好”的观念，而且能尽可能好地采取最有效的手段来实现他的“好”的观念，推进他的利益，如此，他们就要尽可能选择那最能推进他们利益的正义原则，但在无知之幕后，他们是不知道自己所理解的利益究竟含有什么具体内容的，然而，他们总得根据一种利益标准来进行衡量和选择，那么，他们是根据什么来进行衡量和选择的呢？罗尔斯的回答是根据“基本好”（或者说“基本的社会好”），这些“基本的社会好”（自由、机会、财富、自尊基础等）是无论他们的“好”的观念是什么内容都有助于这一“好”的观念的（至少不会妨碍），即使某人因为宗教或别的缘故对“好”的理解事实上不包括这些“基本好”（如不包括财富），他至少可以很方便地不要较多的财富，社会决不会强迫他要，一个人也不会因为一种较大的自由而受罪，这样，根据这种“基本好”的理论，以及道德心理学的一般事实（人们一般都希望得到较多的“基本好”以促进他们的“好”的观念的实现），他们就能够做出一种合理的决定和选择。

总之，各方的这种理性表现为一个人对他的目标有一种前后一贯的倾向，并能有次序地排列这些目标，并遵循那个满足他较多欲望并具有较大成功机会的计划，采取最有效的达到价值目标的手段。

在有关主体的理性和动机的假设中，罗尔斯还有两个特殊的假定值得一提。第一，他假定一个理性的人并不受嫉妒之累，假定原初状态中的各方是免除了嫉妒的，这样，各方关心的就是自身利益的绝对增加而不是相对增加，也就是说，不在乎他与别人的差距，而只是追求自身利益的最高值。他不会因知道别人拥有较多的“基本好”而烦躁恼怒，他们有他们自己珍视的、足以自为的生活计划，他们不会“以别人的损失会更大”的理由而放弃自己的某些利益，不会有“宁可我不好，但我要你更不好”的心理，他们对自身的价值有一种牢固的自信和坚定的追求。而这个假设和前述相互冷淡的假设是相合的，这意味着各方的立场是：一方面不受仁爱心或虚荣心的推动，另一方面也不受恶意、宿怨和嫉妒心的推动。他们既不想牺牲自己的利益也不想损害他人，不寻求相互亲密，也不会相互为敌。他们只是在努力为自己寻求一种尽可能高的绝对得分，而不去希望其他人的或高或低的得分，也不寻求最大限度地增加或减少自己的成功与别人的成功之间的差别。

另一个特殊假定是假定各方都互相知道所有各方都具有一种能建立正义感的能力，这是为

了保证严格的服从，是为了使将订立的契约（不论它是什么）真正有约束力，从而订约各方也就是在做一件严肃、郑重的事情，他们在选择原则之前，能够在对原则的理解方面互相信赖；在选择原则之后，能够在对原则的遵循方面互相信任。他们就知道，这样选择决不是儿戏，他们只要做出承诺，就必须履行它们，因为他们知道别人也将履行其承诺，自己的承诺不会是徒劳的。这就使订约或选择原则成为一件十分郑重、富有意义的事情，同时也意味着，他们要考虑承诺的强度，考虑自己是否能承受得起契约将带来的对自己的约束和义务，因为一旦承诺，就意味着要严格服从（这点在后面将成为选择罗尔斯的两个正义原则的一个论据）。

我们对罗尔斯所设计的原初状态的条件描述就此告一段落，那么，所有这些假设条件是为了达到一个什么目的呢？是为了引出罗尔斯的两个正义原则。这种引出同时也就是对这两个正义原则的契约论论证。而这种论证的核心就是要使原初状态体现出公平，只有在公平的原初状态下所选择的正义原则才是正确的原则。所以，我们说，原初状态的设计主要是为了体现出公平。

体现原初状态的公平的最重要因素有三个，这就是：(1)各方的自由和平等的地位；(2)各方在试图最有效地推进其“好”的观念的过程中所表现出来的工具理性(3)无知之幕。它们都意味着公平，意味着类似和平等，意味着意见一致的可能性。和以前的自然状态论者相比较，应该说第一个条件是大致相同的，仅在平等的基础方面罗尔斯提出了自己的解释；在有理性的条件假设中，罗尔斯明显地排除了理性中的价值因素，这就意味着排除了那些容易引起纷争的冲突的因素；最特异的假设是无知之幕，正是这一假设使原初状态完全脱掉了历史和现实性质而成为纯粹理性的虚拟，也正是这一假设最终排除了订约各方的特异性，使他们成为抽象的、一般的人，排除了一切会影响到原则选择的来自自然和社会的偶然因素，排除了一切会妨碍人们达到意见一致的冲突因素，这就使契约已完全不是现实的契约，订立契约的行为变成了对原则的选择，这种选择实际上已不是在各方之间进行，而是在一个人的脑子里进行。

在罗尔斯看来，只有这样，才能选择正确原则，只有从公平的原初状态，才能引出公平的结果——他的两个正义原则。各方不再能把原则剪裁得适合自己的特殊利益的根本保证是：他们根本就不知道自己的特殊利益是什么！所选择的原则将确保不受到特殊的爱好、志趣及个人价值观的影响。我们甚至可以说无知之幕使第一个和第二个条件都失去了意义了，因为这里实际上已不是许多人在商讨、建立、辩论和选择，而是一种普遍、抽象、无涉具体知识的理性在运行。

二、正义原则的择出

1. 第一次筛选：选择对象的表格

若是按照理论的普遍性的要求，在原初状态欲选择正义原则的各方面前，是应当呈现所有可能的正义观的，这方面是不应加以任何限制的，因为从观念上说，只有查看了所有可能的正义观之后做出的最后选择才具有最大的普遍性和合理性，才真正可以称之为是最可取的选择。但这样做显然有几个困难：一是这将使问题变得无限的复杂，因为实际上有无数的、难以竭尽的可能性，原则上可供选择的对象是不可穷尽的；其次对于那些对历史和现实没有透露过一点信息的未来正义观，我们无法描述其特征以便它们向原初状态各方呈现，我们说不出我们尚不知晓的正义观念的内容。

所以，罗尔斯对选择对象的提出实际上不能不依赖历史和传统。而且主要是西方的历史和传统，他只能提出一个有限的简要表格，把他继承和发展了契约论传统所提出的两个正义原则，与历史上另一些主要的传统正义观（如功利主义、至善论）进行比较，并认为两个正义原则优于这些为人熟知的观念，但是他承认他所进行的这种有限的比较是一种不能尽如人意的比较方式，他的论证是在一种较弱的基础上进行的，但是他觉得一下子找不到更好的方法。他期望以后人们能最终确定一种唯一最好的正义观的充分和必要条件。然后展示一种满足这些条件的观念。而在目前，他还仅限于挑选社会基本结构的某些可欲特征，使之具有自然的最大值的性质——如他的第一正义原则保证了各方面的最大平等自由；第二正义原则保证了最大的最低受惠值，从而看来是最理想的答案。

罗尔斯提供的选择对象的最初表格简要地说是这样的：

(1)处在一种词典序列中的两个正义原则；

(2)功利原则，包括古典功利原则和平均功利原则；
(3)至善原则，这是以亚里士多德为代表提出来的原则；
(4)利己主义原则，包括一般的利己主义和有特殊人称的利己主义(专制型和逃票型的)；
(5)直觉主义观念，包括平衡少数自明原则的直觉主义观念和平衡功利总额与平等权利、或平衡平均功利与补偿原则的直觉主义；

(6)混合观念，主要是指平均功利原则与两个正义原则的混合，即在承认罗尔斯第一正义原则的前提下，以平均功利原则（或修正过的平均功利原则）来代替差别原则。

上面列举的六种选择对象实际上只有前面四种是明确的、独立的，最具实质意义的。它们确实都在历史上很有影响。但是否它们就囊括了历史上的所有正义观呢？

罗尔斯认为：那些有条件的正义观，即那些随着时代和社会条件为转移的特殊正义观是要被排除在表格之外的，首先是因为各方是要选择那些在任何环境里都绝对有效的普遍原则，而不是某种具有强烈意识形态性的和时代性的特定正义观，故而这些特定正义观要被排除在外。

这样，我们就可以说，选择对象表格的提出实际上就意味着第一次筛选，因为，且不说它把那些具有鲜明时代性和有条件的正义观排除在外，而且它还是以西方文化为本位的，以作者自己的主观视野为转移的。

2.第二次筛选：正义原则的形式限制

然而，对上面提出的选择对象的最初表格，还要有进一步的限制。这种限制来自正义原则本身，或者更广义地说，来自正当观念本身（可以把正义看成是正当的一个子范畴，即应用于社会制度时的正当）。但这一说法并不意味着这些限制是来自对正当或正义概念的分析，是来自其定义，而是意味着来自其作用和功能。我们知道：正义原则的作用在于分配基本的权利义务和决定利益负担的划分，这种作用就提出了下述五个方面的限制要求：

(1)首先，正义原则应当是一般性质的，即要表达一般的性质和联系，而不涉及具体的个人或事物。因为这些原则作为社会基本结构的首要原则是绝对的，是在正义的环境里始终有效的，必须始终作为一个组织良好的社会的公开蓝图，任何时代的个人都一定能知道它们。这样，对这些原则的理解就不应要求有一种对偶然的特殊情况的知识，就不能参照特定的个人或集体。而原初状态中的各方，在无知之幕后实际上也是不可能辩认自身的，他们必须选择一种一般的原则。

(2)正义原则在应用中，也必须是普遍有效的，即它们适用于一切场合，一切个人，它们因道德人格而必然对每个人有效，这样我们就假定每个人都能理解这些原则，并在思考中运用它们，而假如一个原则，一旦所有人实行就是自相矛盾、自行取消的，它就要被排除。

一般性的条件限制和普遍性的条件限制的区别是：前者是指原则本身是高度概括的，不涉及具体的事和特定的人物，后者是指原则能普遍地应用于一切场合。

(3)正义原则还必须是公开的众所周知的，这自然是契约论的一种要求，他们要选择原则，要订立契约，原则或契约的结果对他们自然就是公开的。公开性条件的目的是使各方把各种正义观作为被公开承认和充分有效的社会生活道德法典来评价它们。

(4)正义原则必须排列各种冲突的利益要求的先后次序，这种次序一般应是相互衔接的，汇合起来就构成一个层次分明的体系。有序性的条件直接来自正义原则的作用：它要调整、安排各种冲突要求，恰当地划分权利、义务、利益、负担。

(5)正义原则的最后一个形式限制条件是终极性的条件，即它们应当是裁决实践推理的最后上诉法庭。它们说的话是最后的，决定性的。在它们背后没有更高的标准可以诉诸，作为社会制度的基本道德原则，法律、风俗和合理审慎的考虑最后都要由它们来做出最后决定。

总之，正义原则必须是这样的原则：它们具有一般的形式，普遍适用于一切场合，能够公开地作为排列各种冲突要求之次序的最后结论来接受。

罗尔斯概括出这五个形式限制条件的意义主要有二：首先，它们现在把各种形式的利己主义从选择表格中排除出去了，有两种特殊的利己主义：一种是要求别人都服务于他的利益的专制型利己主义；另一种是要求别人都履行正义义务而自己却可豁免的逃票型利己主义，这两种出现了特殊人称的利己主义都因不符合一般性的条件而被排除，还有一种一般的利己主义，即允许和承认所有人都可以如其所愿地推进他自己的利益，这种利己主义不违反一般性的条件，但却因它不可能排列各种冲突要求次序而被排除。

实际上，利己主义可以解释为一种不想订立契约的立场，在订约各方之间，选择利己主义意味着不存在可以达成协议的一致点。故而利己主义也许可以成为个人身体力行的道德原则，但却不能成为一种有关社会制度的道德原则，所以，与其说利己主义是可供选择的一种正义原则不如说它是对所有正义原则的一种拒斥和挑战。故而它们不能列入表格。所以，正义原则的形式限制条件的提出意味着进一步的筛选和排除。

3.第三次筛选：导出两个正义原则的推理及其论据

在罗尔斯提供的表格中，利己主义被排除之后，实质性的观念主要还剩下至善原则和功利原则，在罗尔斯看来，至善原则不是一个正义原则，不是一个政治原则，因为公民权利和利益的分配与人的不同目标或内在价值没有关系。这是一个典型的现代自由主义的命题，但我们不打算仔细探讨罗尔斯对至善论的批评，而主要考察他在比较两个正义原则与功利原则中得出的结论。

功利原则有两种形式，一种是边沁所代表的古典功利主义，其原则要求制度应安排得能最大限度地增加各有关代表人的利益的绝对总额，而另一种是以当代哈桑伊、布兰特等人为代表的平均功利主义，其原则要求制度不仅要最大限度地增加功利总额，而且要增加平均功利（按人分配的功利），亦即最大限度地增加各代表人的利益总额的百分比，在罗尔斯看来，在这两种功利原则中，从契约论的角度看，无疑平均原则要比古典原则更可取一些，因而最可能成为他的两个正义原则的竞争对象。下面我们就来在这种比较中观察原初状态各方选择正义原则的推理过程。

罗尔斯认为：原初状态中各方可能首先考虑接受一种平等原则——这一原则不仅包括自由权利的平等分配，而且也包括收入和财富的平等分配，然而，如果有一种不平等能使所有人都变得比最初的平等状况要好，为什么不接受这样一种不平等呢？这样一种不平等无疑可以鼓励一种积极的表现和努力，抵消训练的费用，带来社会的活力，改善所有人的条件，然而，这种不平等又必须加以限制，限制在一个能够为人们接受的范围内，这意味着这种不平等道德要限制在社会经济利益的范围内，而不允许在基本自由方面出现不平等，不允许以较大的经济利益来补偿较少的自由，不允许这两者之间的交换；其次，这种不平等即使在社会经济利益的领域里也要限制在一定范围内即要符合最少受惠者的最大利益。这就意味着，在基本权利的领域内，由于权利是平等的，最低值也就是最高值，而在收入和财富的领域内，必须明文规定一种最大的最小值，保证一种最高的社会最低生活水平，当然这里所说的最高值是与别的选择对象比较而言，这就把我们带进“最大最小值规则”（*maximin rule*）的思路。

最大最小值规则就是一种可用于在不确定条件下进行选择（*choice uncertainty*）的规则。在不确定条件下的选择有各种可供采用的选择规则，最大最小值规则只是其中的一种，虽然它并不是不确定条件下选择者的可靠指导，但它在一种具有某些专门特征的境况中特别有吸引力，而在罗尔斯看来，原初状态中的人恰恰处在一种适于采用最大最小值规则的境况之中。这种境况恰恰使他们寻求最大最小值的解决办法，而他的两个正义原则就将作为这样一种解决办法出现。

这种适于采用最大最小值规则的选择境况有三个主要特征：第一，选择者不考虑他的选择可能把他带入的各种环境的可能性，并且有对这种可能性不予考虑的理由，也就是说，他有理由把这种可能性的衡量或计算置之不理、弃而不论。这些进入各种环境的可能性（或概率）既然是我们所不知悉的，它们可以说是相等的。

第二，选择者抱一种保守的“好”的观念，即他主要是关心他有把握得到的最少收益是多少，而不大关心他能得到的最多收益会是多少，这个特征就主体而言的，而与此相联系的第四个特征是就对象而言的，即他面临的选择对象中有的确实会产生一种个人很难承受的后果。这三个特征结合起来构成一种特殊境况，处在这一境况中的人自然就倾向于采用最大最小值规则来进行选择。打个比方说，假如一个流亡者要选择进入三个国家，而这三个国家中都是由与他为敌的人来分配他的地位，这样他自然就要观察在这三个国家中哪个国家的最差地位相比于另两个国家的最差地位来说是最好的，他就要选择进入这个国家，以保证得到一个最好的最坏结果。当然，这里的三个特征并没有假设由敌人分配地位这一条件，而代之以对进入各种环境的可能性无法计算、选择者本人的保守气质和选择对象中有的有令人难于忍受的后果的条件，但最后达到的效果还是一样的：即采用最大最小值规则。

现在原初状态中的人恰恰也处在相似的境况中：

第一，无知之幕使各方不可能衡量或计算他们将进入的社会和他们将在这一社会中取得的地位的可能性（概率），他们就必须对这种计算置之不理；而且他们还不能光考虑自己，还要考虑后代，从而也就不能进行那种仅考虑自己利益的计算，而必须使他们对原则的选择在别人看来也是合理的。

第二，原初状态各方也是免除了冒险气质的，他们不想孤注一掷，以求得到最大利益，而是合理审慎地力求确保他们可以得到的最起码利益。而我们前面已经看到，两个正义原则是提供一种令人满意的最小值的。第三个特征就与功利主义有关了，即功利主义的选择对象在罗尔斯看来有一种使人几乎能接受的后果，即便不说它为了最大功利可能允许奴隶或农奴制存在，至少它会被人们用来为严重侵害人们的自由权利辩护。这方面甚至不必深究，因为如果我们面前还有别的方便的选择对象——两个正义原则，而这两个原则又明确保证了一种令人满意的最小值时，我们何必再做他求呢？

这是就功利原则总的精神而言，它有一种可能带来令人不堪忍受的后果的倾向，虽然功利主义者一般也都表示出对人们的自由平等权利和社会最低值的关怀，但他们并没有在其基本原则中直接申明这一道德观点，罗尔斯则认为他的公平正义理论既注重人类社会的一般事实（表现在原初状态的各方都知道一般知识）；同时又把道德理想直接植入了首要原则之中，从而能直接和明确地保证各方进入社会后的自由权利和合理的最小值，因而就是更可取的。

至于功利主义的特殊形式——平均功利主义，虽然按契约论的观点它要优于古典功利主义，但首先它并没有超越功利原则的基本框架，其次它还要遇到一些难以克服的困难：它要依赖一种对或然性的计算，依赖一种个人冒险气质的假定，而这些都是不可靠的。

按照契约论的观点，还可以通过正义原则的形式限制——公开性和终极性的条件给出两个正义原则优于功利原则的论据，这些论据也可置入最大最小值规则所暗示的结构中，即展示两个正义原则体现了在一种很不确定的状态中可望达到的最大最小值。

首先看终极性的条件，由于有关社会基本结构的正义原则的原初契约是最终和永久性的，其带来的结果是否能够承受就成为非常尖锐和严重的问题，人们是否能够在原则所规定的最坏情况下仍然尊重它们，就成为契约是否能够在所有环境里都坚持这些原则就要认真地考虑承诺的强度问题。而在这一方面，处在词典次序中的两个正义原则对功利原则来说显然占有明显的优势，它们保证各方不会因别人要享受的较大利益而损害自己的自由，这种牺牲是不容易承受的一个负担，而它恰恰是功利原则可以允许甚至赞成的。

其次是公开性等限制条件引出的论据。这涉及到人们的正义感和心理稳定性。有一条心理学法则是：人们倾向于热爱、珍惜和支持所有肯定他们自己的价值的东西，而公开申明了一种道德理想的两个正义原则恰恰肯定了所有人的自尊、自爱和自信。所有人的平等自由及其对利益的优先性公开地表示出人们的相互尊重，保证了人们的一种自我价值感，允诺了一种互利互惠的社会结构，从而也就增加了社会合作的有效性和稳定性。而功利原则却更为要求一种对别人利益的认同，要求某些人别人或整体的较大利益而放弃自己的利益，这作为一种私人的道德原则可以成立，而作为社会基本结构的道德原则却显然是一种过分的要求。

在罗尔斯看来，两个正义原则在这方面更深的哲学意义在于：它们表明了正如康德所说的，人们相互不把对方作为手段，而只是作为自在的目的来对待的意愿。而功利主义则是把人既作为目的也作为手段来对待；它强调人们的利益幸福是其原则的指数，强调每个人的福利都是同等重要的，这都可以说是把人作为目的；然而它又允许牺牲某些人的权利和利益以达到最大福利的目的，或者用来衡量一些较不利者的较低生活前景，这实际上就等于把这些人又作为达到目的的一种手段来对待。故而理性的天平在这一方面也要倾向于两个正义原则而非功利原则，表格上最后一个与两个正义原则并列作为候选者的功利原则也就要被排除。

以上所讨论的对两个正义原则的论证——原初状态的整个设计，选择对象表格的提出和形式限制，采用最大最小规则的选择过程，以及与终极性相关的承诺程度，与公开性相关的稳定性等——都是契约论性质的论证，那么，这种论证是否成功呢？罗尔斯对历史上的契约论究竟继承了一些什么因素，又在哪些方面有所发展呢？契约论能不能成为一种普遍的证明方法？能不能经由罗尔斯在当代复兴呢？这些就是我们下一节需要考虑的问题。

三、契约论作为一种证明方法

罗尔斯在他的第一篇探讨社会正义的文章“公平的正义”中，就谈到他受到了可追溯到古希腊智者的契约论思想的启发，但也指出他自己的理论与先前契约论的三点区别：第一是不涉及一般的人的动机理论，即不涉及一般的人性论，而只是作为一种推测的哲学解释来分析正义的概念；第二是比先前的契约论更抽象，各方不是要建立任何特殊社会或政体，也不是像游戏论中那样要确立个人致胜的策略，而是要共同接受可用于共同实践的某些道德原则，或者说找到最高层次的游戏的答案——正义原则；第三是非现实，非历史的情况，既然不是想提供一种对事实的解释，自然也就避开了说它虚伪、不真实的责难。

在《正义论》中，罗尔斯不再联系博弈论来谈论契约论了，他明确地指出，他的正义论是对契约论的继承和发展，并申明了他袭用契约论术语的理由，认为契约论术语的优点在于：

1.可使正义论成为合理选择理论的一部分，这样就可以为正义理论提供一条由理性的人们选择的原则，就可以科学地根据合理选择的各种模式或规则，来描述、分析乃至设计人们的选择过程和结果。

2.“契约”一词暗示着个人或团体的复数，而正义原则处理的正是若干个人或团体在分享利益时的冲突要求，而且，“契约”一词还暗示只有按所有各方都同意的原则来处理冲突要求才算恰当。

3.契约术语也表现了正义原则的公开性，正义原则作为原初契约结果，同时又是随后原则的指导原则是众所周知的。

4.契约论还有一种悠久的传统，以这一思考方式来表现人际关系有助于明确观念且符合自然的虔诚，即能够取得一种来自传统和习惯的适应和支持。

5.契约论除作为证明程序外，也可成为一种比较手段，即可以通过追溯对订约最初状态的各种不同解释而进行各种正义观的比较（如比较“公平的正义”与功利主义），这样，契约论程序也就为正义观念的比较研究提供了一种普遍的分析方法。

这里主要是就证明方法谈继承，我们后面将谈到罗尔斯在实质观念方面的继承。

至于发展，罗尔斯认为他的正义论进一步概括了洛克、卢梭和康德所代表的传统的社会契约论，使之上升到一种更高的抽象水平。在此，所谓“上升”、“更高”和“更抽象”的含义在于：

1.不把原初契约设想为要进入一种特殊社会或建立一种特殊政体；

2.原初契约的目标是要确立社会基本结构的道德原则；

3.这些原则是那些想促进他们自己利益的自由和有理性的人们将在一种公平的最初状态中，在一种无知之幕后选择和接受的，并以此作为他们联合与合作的基本条件；

4.这些原则(即原初契约论所选定的原则)将调节进一步的契约指定各种可行的社会合作的政府形式；

5.此即为“公平的正义”，它最接近于我们所推崇、所看重的正义判断和直觉信念，可以替代功利主义，构成一个民主社会的最恰当道德基础。

借用一个罗尔斯使用的、使正义原则应用于社会制度的“四个阶段序列”，我们也许可以更清楚地看到罗尔斯的正义论与先前契约论形式上的区别。这一序列是：

1.选择正义原则；

2.制宪；

3.立法；

4.执法。

我们可以说，先前的社会契约论注意的主要是第二阶段，即通过制定国家的根本大法来确定某一种特殊社会或建立某一种特殊政体；而历史上的统治契约论所注意的则还要靠后，它注意的主要是第四阶段即执法阶段，即在国家性质已确立的情况下考察君主与臣民双方执行最初契约的情况。这样，集中注意于在制宪、立法与执法之前对正义原则进行选择的罗尔斯的理论自然就可以说是更抽象、更概括了，而且，向前跨了一步不仅有更抽象、更概括、更具理性色彩的含义，也有更强调道德的意义。以前在社会契约论中隐涵和被认为是自明的道德原则（自然法、自然正义）在罗尔斯这里被公开地提出来作为探讨的主题，并且受到了严格的分析和检查，对要选择的正义原则也做出大量细致而缜密的论证。也正是因此，我们可以把罗尔斯的理论称之为道德契约论，以区别于前述的社会契约论与统治契约论。

自然法学说赋予契约理论的两个鲜明特征：一个是它的理性原则，一个是它的道德原则。现在我们看到，这两个特征更为突出和明显了，但我们还可以从论证方法和实质结论两个方面，进一步概括罗尔斯正义论的继承和发展。

在论证方法方面，罗尔斯的理论进一步继承和发展了理性主义精神，在此，理性的方法一方面表现得更抽象、更概括、更带虚拟性质、更直接地指向道德目标——正义原则；另一方面又表现得更谨慎、更精密、更少依赖直觉，在其运作过程中更有意识地试图摆脱价值的纠缠，罗尔斯的正义论继承了契约论的先验和演绎的理性精神，它不是经验事实的概括和总结，不是对历史现象的分析和解释，而是一种理性的推演，但这种理性的推演现在不是要推演出绝对的定律，而毋宁是一种设计、一种试验，容有修正、补充和改造。罗尔斯在设计原初状态中所表现出来的道德几何学，它不带有形而上学的本体论含义，它的目标是有限的，前提是尽量排除价值因素的，方法是严格借鉴现代科学的（如博弈论与微观经济学），所得出的结论也不是绝对意义上的。从这些方面都可以看出哲学的时代特征，看出罗尔斯所受到的分析哲学与科学哲学的影响。但尽管如此，罗尔斯并没有逸出契约论强调理性的大道。罗尔斯正义论对契约论的复兴在某种意义上也是对理性主义的复兴。

在实质结论方面，我们看到，罗尔斯理论的基本倾向是结合与调和自由与平等、既强调人的基本自由及其优先性，又强调人们对这些自由的平等权利，尤其是强调人们在社会地位利益方面也要尽量平等，即使有差别也要以符合最少受惠者的最大利益为准。罗尔斯在这两方面都对契约论有所继承和发展，在某种意义上，我们可以把他的理论看成是对洛克的维护自由的契约论和卢梭的渴望平等的契约论的调和，但我们在此宁愿更注意他与洛克和卢梭两人的共同点的联系，注意“公平正义”与自然权利理论的联系。

“自然权利”的另一个译法是“天赋人权”。是“天赋人权”还是“人赋人权”？这两种说法可以并不矛盾，“天赋人权”是表达一种价值观，一种具有道德意味的价值观，表示对“权利”有一种坚定的、毫不动摇的看法，表示这种“权利”有一种天然的、绝对不可剥夺的性质，剥夺它就意味着侵犯到道德、侵犯到人之为人的根本条件。而“人赋人权”则可以表达一种具体实现这一“权利”的途径和手段，这种“权利”当然要通过人的活动来实现，通过社会制度，通过法治来实现。在此意义上，“天赋人权”也可以说是“人赋人权”。

在罗尔斯的《正义论》中，很少出现“权利”的字眼，但是，他所说的“基本自由”实际上就等于人的基本权利，这些基本自由包括思想和信仰的自由、参政和交往的自由、人身和财产的自由等等。与洛克所说的个人享有的“生命、健康、自由和财产”的权利，与卢梭所说的“人是生而自由的”权利，以及后来《独立宣言》又加上的“追求幸福的权利”相去不远。而且，在此重要的还不是抽象地规定这些权利（大概现代国家的几乎每部宪法都规定了这些权利），而是要表明对这些权利要捍卫到什么程度，当这些权利与别的东西发生冲突时又采取什么态度，罗尔斯在其《正义论》正文开始时说过“正义是社会制度的首要价值”一段话，强调“权利”不能因为任何利益而被放弃。虽然这并不是作为定论出现的，而是作为有待探讨和证明的论点，作为一种直觉性的正义信念，一种将放在“反省的平衡”中与正义原则对照的“被考虑的判断”出现的，但我们看到，这些论点实际上也是结论，它不仅体现在罗尔斯的第一正义原则及其优先规则中，甚至也渗入了其前提，体现在他对原初状态的设计中。在那里，各方是自由、平等的个人，有同等的选择权利。罗尔斯不仅在原则对制度的应用中（即“反省的平衡”中），而且在原则的选择中（即原初状态条件的设置中），都诉诸了这种流行的道德信念和直觉，而这些道德信念和直觉与社会契约论有着不解之缘，它们正是在17世纪社会契约论的气氛中培植、繁荣和普及开来的，当然不止是社会契约论宣传了这些观念，但在社会正义的领域里，契约论曾经长期是占支配地位的理论，对英国光荣革命、美国革命和法国革命有直接的影响。权利在此不仅是在与权力对立的意义上使用，也常常在与功利对立的意义上使用，换言之，任何一个人的基本权利，不仅是有权力的政府或其他团体所不能任意剥夺的，也是不能以社会整体利益或大多数人利益之名而损害的，而按罗尔斯的优先规则，第一原则是优先于第二原则的，这就意味着，制度对个人权利的某种侵害行为，即使符合最少受惠者的最大利益，也仍然是不允许的。这就是“由正义所保障的权利决不受制于政治的交易或社会利益的权衡”的真实含义。在罗尔斯看来，这正是他的契约论性质的“公平的正义”理论胜过流行的功利主义的一个突出优点。

然而，罗尔斯强调的并不是这种对契约论的实质结论方面继承，而是对其证明方法的继

承。前一种继承是隐藏在“反省的平衡”，“被考虑的判断”的后面，因为300年前新鲜的权利观念，到现在实际上已成为大多数人默认和遵守的常识。但是，这类观念毕竟还是受到了挑战，甚至受到了严重威胁，如果麦金太尔的历史分析确有合理之处，那么可以说，为这类观念所建立的理性基础还不够牢固和可靠。否则，就很难解释为何道德相对主义、价值虚无主义在本世纪广泛流行。因此，重新证明这类观念，为这类观念建立起牢固而可靠的论据，就不仅仅是一种证明方法的问题了，它对捍卫实质性的观点也具有关键的意义。

那么，契约论，或者说经过改造和提升的契约论，在新的历史条件下，是否能够提供一种有效的证明方法呢？罗尔斯通过契约论的构架提出的证明是否成功呢？

德沃金对罗尔斯的契约论方法提出了批评。首先，他认为，罗尔斯提出的原初状态的论据，对于证明他的二个正义原则来说，看来并不是很必要的，甚至是多余的，因为罗尔斯还有一种证明，即把他建立的正义理论体系、基本观点和我们在日常生活和行为实践中事实上推重的、直觉的、无须反思（unreflective）的信念和准则进行比较和对照，这主要见于罗尔斯《正义论》的第二编，即把两个正义原则用于制度时的推论与我们日常流行的正义准则与基本和共享（或最多的人共享）的正义观进行对照比较，来回修正。这里可以说有两种平衡，上面说的是第一种平衡；还有一种平衡是理论本身的平衡，理论结构的平衡，用通俗的话说就是“自圆其说”，即一种理论体系其中各个论点都是相容的，相互支持和前后一贯的，不发生逻辑上的矛盾，它们综合起来，达成一种统一的、和谐的理论结构，这种结构就构成了某种合理性，就具有了某种说服力。整个理论体系就达到了一种平衡，一种是涉外的平衡，一种自身内部的平衡，这就是所谓“道德的一致性理论”（coherence theory of morality）或“一致性论据”。在德沃金看来，契约论的原初状态的论据即使不是多余的，也至少在书中占的比重过大，有喧宾夺主之嫌，罗尔斯使用经济方面的论据太多，过于繁复，不太成功。

关于“一致性”还可以再说几句。德沃金认为解释“一致性”有两种模式，一是自然模式，比方说在物理学中，把理论与客观存在的现象对照比较，看它们是否一致；一是结构模式，即把建立的理论与某些直觉信念比较，自身内部的各论点之间也互相比较，看它们是否达到了某种一致。后一种模式要经过道德直觉与道德理论之间的反复平衡，互相修正，才能达到一致。德沃金认为罗尔斯的平衡表面上看与自然模式有关，实际上是以从特殊信念到一般正义理论的推理的结构模式为前提。这样的结构模式所达到的东西，虽不构成绝对意义上的真理，但不见得对行为就没有一种指导意义，因为它还是基于某种合理性的。而且由于它要和共享或比较流行的实际道德观念平衡，就使它处在一种较为大众化的立场。

德沃金对罗尔斯理论方法的第二点批评是认为：用原初状态来证明两个正义原则本身是不充分的，很不够的。因为原初状态本身只是一个中介，一个中点，它后面还应有一个更深的理论作为依托。换言之，正义原则只能通过（through）而不能从（from）契约中产生，它们不能从原初状态中得到根本证明。德沃金实际上是要追溯到契约论的原初状态的公平性前提上去。即不是从原初状态，而是从罗尔斯为原初状态设置的自由平等条件中寻找正义原则的根据。在德沃金看来，还必须有一种更深的道德理论来证明罗尔斯对契约的用法，这一理论将解释为何原初状态对于正义论是一个有感染力的模式，那么，这一深层理论肯定不是如功利主义那样目的论的，罗尔斯自己认为原初状态是对康德的自律和绝对命令的义务理论的一种程序性解释，而德沃金认为罗尔斯的理论是一种基于权利的理论，这一点是从其采用契约论的论据而表明的。契约论本质上是一种权利理论，或者说是权利理论的自然发展，目的论和康德式义务论都不能利用契约论的方法，契约论重视个人权利，以个人为基础，它假定每个人都追逐他自己的利益，对集体决定有一种否决权，不得到他的同意就不能产生义务。所以正是因为有一种更深的权利理论在作为基础和依托，社会契约和原初状态才能得到说明，才能作为中介提供论证。

另外，德沃金认为：原初状态本身也是不确定的，设计得并不太好。那种虚拟的无知之幕后有些不可信之处，跟人们的实际选择处境相差很大。原初状态中的人的先定利益亦不同于当代人的利益，原初状态中的人是求稳妥和保险的，是态度保守的，但实际生活中的人却可能不是这样，而是更趋进取。所以，德沃金说：只有撤去无知之幕后能证明两个正义原则才是真证明。

德沃金一方面说罗尔斯的契约论论据是多余的，因为在其“公平的正义”理论中，实际上起着更重要、更有说服力的作用的论证是一种“一致性”的证明：理论在外部与人们的直觉信念的一致和理论自身内部的一致性；另一方面又说罗尔斯的契约论论据是不够的，因为所有这些论

据本身又还需要一种深层理论的支持，一种权利理论的支持。一方面说多余，另一方面又说不够，这种说法看起来是矛盾的，然而它真的是如此矛盾吗？德沃金并没有深言做出这种似乎矛盾的批评的理由，甚至没有明言这种矛盾，没有明确地觉察出这种矛盾，但我们却可以由此思路做一些有意义的引伸。

罗尔斯借重契约论作为论证方法的理由已见上述，他认为通过契约论可以利用当代的合理选择理论的规则；可以实现由竞争的个人向建立正义共识的社会的过渡；或以表明正义原则的公开性；可以作为一种把“公正的正义”与其他正义原则（主要是功利主义）进行比较的手段；最后还可以得到传统的支持，但我们说，契约论作为证明方法在罗尔斯那里最重要的意义是作为一种集合手段，即把基本的论据都集合为一体，这也就是他所说的，通过“原初状态”，把一个相当明显的选择问题，同合理公认的在选择道德原则所面临的条件统一到一个观念中来了。“原初状态”是从契约论的“自然状态”发展而来的，然而是在形式上剔除了一切历史、现实和经验因素的一种是最初状态。人们对这种最初状态的设计和解释还可以有其他方式，因此，面临的第一个问题就是：“原初状态”本身的设计是否合理？因为这里很可能面临一个麦金太尔所做的含有讥讽意义的批评：不但处于某种无知之幕后的原初状态中的有理性的人确实会选择如罗尔斯所主张的正义原则，而且只有处在此种原初状态的人才会选择此类原则。这就意味着，结论至少已经部分地包含在前提之中了，当然，罗尔斯会回答说，证明就在于结论和前提的互相支持，就在于整个体系的互相平衡、互相协调，他也确实是努力这样做了，正像一个批评者所言：《正义论》全书每一页都是在努力提供证明。但从上面我们已经看到：契约论所起的作用并不是提供一种根本的证明，而只是作为一种集合手段，一种容纳各种条件的手段，一种装载合理选择规则的手段，一种启动一系列明智推理的手段，一种与其他原则比较的手段。原初状态本身需要证明，而对它的证明现在主要是来自本身要通过它来证明的正义原则及其制度应用。因此，人们完全可以设想，他们不仅可以设计另外的最初状态，而且可以超出契约论的理论框架，以另外一种方式，另外一个名称来集合这些论据，这样，契约论的形式对正义原则的证明就可以说是多余的，即如果这种原初状态的设计并不十分合理，它所集合的论据并非是最有说服力的论据，或者有一些更为根本的论据不能装载到它的形式框架之内，那么，就可以对这一形式框架弃而不用。总之，说它多余，是在契约论仅仅作作为一种证明方法，甚至仅仅作作为一种集合论据的手段的意义上的，与契约论的实质结论无涉。

罗尔斯确实继承了契约论的某些实质结论、实质精神，然而，这方面在他的理论中经常是隐而不显的，他对契约论的强调主要表现于把它作为一种证明方法来应用。

至于说到契约论论据还是不够的，则是采取了一种更广阔的视角或进入了更深的层次。我们在第一章叙述了“契约”概念在历史上的广泛含义，它不仅是一个法律概念，而且还具有哲学乃至神学的涵义。我们还阐述了可以从“契约”概念引伸出来的某些逻辑推论。罗尔斯的实质结论无疑是与自然权利理论有某种联系的，但他并没有指明这种联系，而主要是把契约论作为一种方法论，在这种方法中又装载了可能太多的明智性质的推理和经济、数学方面的论据，这样反而容易模糊人们投向正义原则的最深根基的视线，正义原则的根肯定不是最终地扎在此处，而是延伸到人们对人生、对世界的整个看法之中。罗尔斯的理论表现了一位严谨的学者应当具有的谨慎和小心，但我们也许可以责备他过于谨慎小心了。他希望通过重新诉诸契约而得到一种传统的支持，但把契约论仅作为一种证明方法却会限制这种支持，甚至我们可以说，契约论遭到攻击最多的地方，最为人诟病之处恰恰是它的方法而非它的结论。罗尔斯当然在很大程度上改造了这种方法，这种方法也确实有一种富于启发性、甚至解放性的意义，它能使人以一种新的眼光重新审视社会制度的道德根基，使人在建立社会正义最初逻辑起点，也是道德起点上发言，它也提供了一种集合论据的手段，提供了一种作为证明的选择过程，但是，这种道德根基、道德起点还应当得到一种更深层、也更广泛的理论支持，得到一种具有实质意义的文化观、历史观、价值观的支持，它还可以有一种更大范围内的自成一体。在此意义上，仅仅作作为一种证明方法的契约论又确实是不够的了。

罗尔斯曾讲到过去对功利主义的批评只是零打碎敲的进行，只是否定性的批评而非建设性的工作，因此，他努力构造另外一种观点，构造一种具有同样明晰性和同样系统的正义论体系以取代功利主义。罗尔斯在这方面的努力是相当成功的，但我们还是可以说，在罗尔斯那里，还是没有能完全摆脱过于谨慎小心、自我局限、自我封闭的分析哲学的影响，还是有零打碎敲

的痕迹。罗尔斯因不满于功利主义而诉诸另一种传统——社会契约论，但正像仅仅对功利主义的批评不能取代功利主义一样，仅仅把契约论作为一种证明方法也不能充分挖掘契约论的潜力而复兴这一传统。而真正的问题也许并不在于复兴这种或那种思想传统，而在于复兴一种对人生和世界的全面关怀、恢复一种博大的胸怀，进行一种最高层次的哲学观照，这不是僭妄，不是非分之想，而正是后来者的使命。这种复兴是否取契约论的形式并不重要，历史完全可能给它另外一个名称，但一种伟大的综合确实是我们的期望。

四、“反思的平衡”

罗尔斯的证明方法从契约论中得到启发，而契约论在对社会制度的讨论上具有强烈的理性主义色彩，在这一节里我们将追溯这种理性方法在社会正义领域内的历史，并进而讨论罗尔斯的理性方法——“反思的平衡”，包括其理论渊源、与别种方法的比较，以及它在伦理学史上所处的地位。

契约和伦理的历史结合产生了统治契约论，而统治契约论考虑的尚是制度中人的伦理，到社会契约论才进一步发展到考虑制度本身的伦理，使契约论真正进入社会正义的领域。培根、笛卡尔、霍布士等人所生活的时代，至少对思想家个人来说是一个堪称幸运的时代，他们对人类的知识和理性都充满信心，这种普遍的信心一直持续到19世纪。不管社会契约论者三个伟大代表——霍布士、洛克、卢梭——在有关制度和政体的最后结论上是多么不同，也不管他们三个人在一般哲学，在认识论方面有何不同，他们在讨论社会制度的领域内都一致地推崇理性，都采用了理性主义的方法，但在运用这各种理性方法上却又存在差别，我们现在就来分析这些差别。

霍布士认为他的思考次序是模仿上帝创世的次序。上帝在一周中创世的次序是：光明、昼夜之别、苍穹、有感觉的生物、人，而紧接着创世之后就是诫命，所以他沉思的次序也是：理性（光明）、定义（区别）、空间（苍穹）、星球（天体）、可感觉的性质、人，等到人成长起来，就是服从命令。

霍布士推崇理性，这和当时对几何学的崇拜有关，在他之前的格劳秀斯和在他之后的斯宾诺莎，都借鉴了几何学方法来构建自己的哲学体系：即从最简明的一个或几个命题出发，演绎出一套复杂的理论来。但他又承认一切知识是抽象的经验，也就是说，在涉及知识的起源上，他是一个经验论者，然而，他又是一个唯理论者，认为只有推理能得到普遍的知识，察知事物运动的原因要靠理性的演绎，并且，他以数学方法来解释对理性的运用，认为理论思维也就是对观念，名称的加减。

因此，我们可以说霍布士的理性方法是一种推演的理性，演绎的理性，一种深受数学方法影响的理性。我们在《利维坦》中看到的正是这种方法：霍布士善于列举、解说定义，如他对各种自然法的列举和规定，这种列举和定义有时未免使人有烦琐之感。而他最擅长的还是推演，即从一个基本前提推演出一系列结论来，在这些推演过程中，确实有一种逻辑的明晰性和一贯性。但如果前提有错误，就很难避免结论的错误。我们知道，在社会契约论中，最初的起点都是自然状态，由于霍布士思考的事实前提实际上是一种人性的性恶论，他还把这种性恶论描绘成自然状态中“人对人是狼”的极其悲惨的图景，加上其价值前提是以生命的保存为至高无上，因此就逻辑地引出一种绝对放弃权利的契约和君主专制的政治结论来。

洛克在认识论上是个经验论者。在他的《人类理智论》中，他批判了天赋观念论，他认为并不存在全人类普遍同意的原则，无论是思辨原则还是实践原则（前者指数学、逻辑，后者指道德、宗教），而且组成原则的观念也非天赋。心灵又在此基础上加工构成复杂观念。真实的观念（真理）是同事物原型符合的观念。虽然有直觉的知识和论证的知识，但知识并不能超过我们的观念范围。

洛克的经验主义的认识论和他具有唯理主义色彩的政治学、伦理学并不是逻辑上严密一贯的，这中间存在着某种断裂。在认识论方面，洛克是个经验论者，而在政治学、伦理学方面，洛克却又表现出强烈的理性主义倾向。他所谈到的自然法实际上就是一些直觉的不证自明的普遍适用的原则。人的自然权利被看成是天赋的、不可争议的、不可剥夺的权利。洛克的认识论对于他的政治学、伦理学最有意义的是他对天赋观念论的批判，从这种批判中可以直接引伸出否定君权神授的结论来。这就是他在《政府论》上篇中所做的工作，他以其犀利的笔抨击了当

时的保王党费尔默爵士的君权神授论和世袭权力说。然而，在其对“普遍同意”说的批判与其自然法理论之间却存在着明显的逻辑矛盾，这种矛盾不仅存在于他的认识论与政治学之间，也存在于他的政治学内部。这也可以附带说明洛克的哲学是一种受欢迎的常识哲学，受欢迎的常识常常需要兼容并蓄，综合各种意见，以致容忍逻辑上的某些矛盾。所以，洛克的理论也往往可以由后来者从不同的方面加以发展。常识常意味着妥协，然而在政治领域中这一特点与其说是毛病，毋宁说是优点。

因此，我们说，洛克的理性方法是一种常识的理性，它将人们日常生活中已经流行或开始形成的某些直觉信念集合起来。把这些有关正义的经验常识转变为具有普遍必然性的原则，甚至把它们熔铸为一个体系，则非理性莫属。洛克的伟大意义就在于他敏锐地看到了当时正在人们生活中形成的，代表时代发展方向的某些有关社会正义的信念、准则，并把它们提升为明确的、具有普遍约束力的原则。理性并不因与常识结盟而变得低下，变得琐屑，相反，在社会伦理的领域里，理性必须借重常识，而且我们所说的常识是指正义方面的常识。借重常识也就意味着借重传统、诉诸大众，这对达成一个良好的共同体意义重大，因为任何一个良好有序的共同体都必须首先建立在有关什么是正义的共识的基础上。至于如何消弭逻辑矛盾，这正是要求发挥理性力量的地方，但没有理由贬低常识。

我们前面讲到过卢梭对情感的推崇。卢梭擅长的不是说服人，而是打动人，不是诉诸大脑，而是诉诸心灵，卢梭的论著从来不是平心静气的，而是充满着激情。诚然，在最冷静的笔调、最枯燥的行文和最抽象的概念后面都可能隐藏着一种感情，一种信念，一种价值评判。生命不可能对外界现象无动于衷，不可能完全担当一个纯粹的旁观者的角色，即使这常常是真理的一种要求。但是，激情在涉及个人追求时常常可以大胆地前行，就像在一个艺术家、一个自为的圣徒那里一样，而在涉及到社会事务、涉及到不仅自己，而且是许多别的人，尤其是在涉及到对千百万人影响巨大的社会制度的问题上发言时，却显然要抑制个人的情感，不能纯然让自己的生命体验中所感觉到的东西成为普遍的原则，不能纯然做情感的反应，而是要以一种更为普遍、更为客观的眼光看待问题，我们确实看到卢梭在自己的著作中常常是感情流溢、不加掩饰的、真诚热烈的，但我们还是可以说，《社会契约论》是他最具理性色彩的一部著作。

《社会契约论》是卢梭思考得最久的一部拟议中的长篇著作《政治制度论》的撮要。卢梭在这部著作中做了一种他很少做的使其思想体系化的工作。他首先讨论人类是怎样由自然状态过渡到政治状态的，以及公约的根本条件是什么，继而讨论制宪，讨论政治法即政治的形式，最后阐明巩固国家体制的方法，他的论证过程已如前述，其中也表现出一种逻辑的明晰，不乏许多有力的论据。卢梭也诉诸了人们在社会生活中某些常识性的正义直觉，他敏锐地认识到共同体的“善”对于所有生活在其中的人们的意义，同时又意识到直接民主制不可能完全实现，因为人不是神，“这样完美的政府不是人类份内的东西”，如果发展这样的思想萌芽，将可能构成对政治浪漫主义的有效抑制。卢梭特别注意一种在人们心中根深蒂固的信念：人不能在自己的同类遭受苦难时淡漠置之，尤其是在极少数人占有大量财富的时候，让大多数人因冻饿而死于沟壑，此无异于“率兽食人”。

但是，常识性的正义直觉不是单数，而是复数，要建立道德的普遍原则，尤其是社会政府的原则，就必须以种理性的态度，在各种常识性直觉之间进行冷静的分析和恰当的平衡。确实，一种同情，一种恻隐之心，可以说揭示了我们道德的源泉，我们的许多道德规范和原则之所以能起作用，能发挥效力就有赖于此，但道德的这种源头和道德的原则还不是一回事。如何实现这种同情、如何把那人心源头上一点十分宝贵的不忍之心化为制度的不忍，却显然要经过一种理性思考的途径。卢梭的理论看来还是受到了情感的过分影响，以致他在考虑平等时没有仔细地思考如何实现平等，首先是实现什么样的平等，其次是通过什么途径去实现平等，以及这种平等可以实现到什么程度。所以，我们说在卢梭《社会契约论》中所使用的理性方法主要还是一种受到情感影响的理性，一种具有浪漫色彩的理性。

康德从人同时属于现象界和理智界的著名命题出发，在《实用人类学》中依据经验进行分析，而在三大批判和《道德形而上学基础》中，他采用的一种先验理性的方法。如果说卢梭关注的是道德的源头，是恻隐和同情之心，那么，康德所关注的就是向人发出绝对命令的道德原则了，在康德看来，这些原则绝对不顾及人的情感，人的喜好，也不依行为的目的效果为转移，它们是先验地从人作为意志自由和有理性的存在本质中演绎出来的，是置于一种纯粹的实

践理性、道德理性的基础之上的。

黑格尔自然也是推崇一种思辨的理性的，然而在此，过程的概念被引入了，理性，或者说绝对理念被理解作为一种发展过程，自然与社会也被理解作为一种由它决定的发展过程，在这个意义上，我们也可以说黑格尔的理性是一种历史的理性（但并不能完全代表历史理性或思辨的历史理性）。但是在此，一种“辩证”的逻辑必然性是至高无上的，一切演变和发展都被纳入他所称的“正、反、合”的三段式，而且，这一发展有一顶点，它在哲学上是在黑格尔自己的哲学中，在政治上是在普鲁士国家形态中得到了最终的实现。

所以，我们说，社会正义理论在方法论上最容易受到来自两个方面的不利影响，一个是来自情感，来自政治上的浪漫主义；再一个就是来自理性本身，来自僭越的逻辑，来自对必然性近乎宿命论的强调。

我们现在又回到罗尔斯。我们在上一节考察了他对正义原则的契约论据的证明，在这一节中，我们要考察他在更广泛意义上而言的“反思的平衡”的证明，以及康德式证明，不过，为了了解他进行这些证明的背景，我们先得回顾一下20世纪英美伦理学的发展，看看由于理性害怕自身的僭越，退缩到了何等狭小的范围。这种谨慎小心，虽然作为对黑格尔主义的一种反拨是不无道理的，它也并不会对社会正义理论构成直接的损害，但我们仍旧可以说，由于限制了人们去探讨社会正义问题，因此它还是影响到了正义理论的发展。

本世纪英美伦理学大致经历了这样三个阶段，第一阶段是直觉主义，穆尔指出了传统伦理学一种“自然主义的谬误”，即想用自然的东西去定义非自然的东西，而有些东西从根本上说就是不可定义的，不可定义就意味着无法用推理来把握，无法解释和说明，而只能通过直觉去把握、去认识，这就限制了理性的范围，就在事实与价值之间划了一条界限，就意味着前此的大多数伦理学家都犯有这种“自然主义的谬误”。穆尔指出“善”就是这样一个不可定义的概念，后来的罗斯指出“正当”也是一个不可定义的概念，普里查德则认为在我们具有什么义务的问题上也不可进行论证。当然，穆尔等人并不认为“善”、“正当”、“应当”这些概念就没有意义，而只是认为它们无法定义，只能通过直觉去把握。

第二阶段是情感主义的阶段。艾耶尔在30年代轰动一时的《语言、真理和逻辑》一书中认为：伦理命题不是有意义，可证实的命题，它们只是在表达我们的感情，这一理论后来由史蒂文森发展和完善了。情感主义进一步限制了理性的范围，经过这样两个阶段，伦理学的理性基础就在很大程度上被动摇了，理性（尤指19世纪那种思辨和概括的理性）就几乎从伦理学中驱逐出去了，而这种对理性的限制恰恰又是出于一种寻求真理和确定性的动机，正是用理性批判、考察理性本身的结果。理性发现自身无力在价值领域中建立和证明真理：直觉主义者认为那些价值和正当与否的评价只是一种自明的、无法定义和论证的东西（但他们还是相信它们是可认识的），而情感主义则进一步坚持道德判断只是在表述我们的感情和喜好。这样，伦理学领域中还剩下什么东西具有绝对的确实性，还有什么工作是真正哲学的，发掘真理的工作呢？显然就只剩下对道德陈述的语言和逻辑所做的分析了。这样，理性就走到了对实质性道德问题表现出冷淡，仅仅对道德语言和逻辑进行分析的阶段，伦理学仅成为伦理学的认识论和逻辑学。

我们当然要注意到：这种冷淡是出自一种寻求真理、寻求确定性这样一种动机。一个真正的哲学家不会提出他本人也怀疑其真的命题，可是现在伦理命令、道德规范之类的命题恰恰就是这样一类让现在哲学家们生疑的命题。如果我感到我提出的规范只是表述我的情绪，但又苦于找不到我提出的规范的合理根据，那么，我也会对提出规范或解决方案持审慎态度。所以，在英国伦理学家海尔看来，欧洲大陆存在主义哲学家所提出的伦理命题就像一些形状各异、色彩斑斓的气球，一旦破裂就会发现其中一无所有。于是，真理或者说真正的严密性、确定性、明确性现在看来就只存在于形式之中、逻辑之中、语言之中、分析之中，这种分析虽然可能是琐屑的，但它至少是可靠的。

罗尔斯正是在这种哲学伦理学的气氛中开始他的哲学工作的，在他的早期论文中，也较多地表现出这种语义和逻辑分析的倾向。但在《正义论》中，他明确地对这种倾向提出一批评，认为它不应在伦理学中占据主导地位。他说：“无论如何，仅仅在逻辑的真理和定义上建立一种实质性的正义论显然是不可能的。对道德概念的分析和演绎（不管传统上怎样理解）是一个太薄弱的基础。必须允许道德哲学如其所愿地应用可能的假设和普遍的事实。”他说他希望强

调研究实质性道德观念的中心地位。

因此，罗尔斯的《正义论》本身就体现了一种转向：由形式的问题转向实质性的问题；由纯粹语言和逻辑的分析转向进行思辨的概括；由谨慎地限制理性转到大胆地在公共事务上运用理性。

从《正义论》整个体系来观察，罗尔斯所采用的理性证明方法是一种他称之为“反思的平衡”的方法。罗尔斯把正义论看作是一种想描述人们正义感的尝试。人们在社会生活中已经形成一些比较固定、比较流行的有关正义的常识性准则和信念，罗尔斯把那些在有利环境下形成的、比较可靠、在相当程度上被公认为正确的信念准则称之为“所考虑的判断”。但他认为，这些判断尽管相对可靠正确，但并没有理由认为我们的正义观能够准确地由大家所熟悉的常识性准则来表现其特征。对正义感的正确解释，肯定要涉及到理论结构，肯定需要建立一种理论框架，也就是说，需要通过一种思辨的理性，概括出一种包含有普遍原则的正义论体系来。

然而，这种建立就是一个“反思的平衡”的过程，“反思的平衡”也就是对正义理论的一种证明方法。在“公平的正义”理论通过被原初状态中的人一致选择（契约论据）而确立之后，还必须把它与人们的日常正义准则（“被考虑的判断”）进行对照，看是否能够达到一致，在这种对照过程中或者修改原则，或者纠正判断，总之要力求达到最后的一致或者说趋于平衡。

这就意味着，道德哲学是苏格拉底式的：在这一类似对话的过程中，我们现在所考虑的判断的原则被阐明，我们就可能想改变这些判断。也就是说，道德哲学将影响到我们的实践，将促使我们对原先的常识性准则进行进一步的反思，加以修正，甚至彻底改变，从而也就最终改造我们的行为，改造我们所建立的社会制度，这就是理论的巨大作用，就是理论的实践意义。

罗尔斯认为；他的理性证明方法有别于笛卡尔式的从某些自明原则推演的标准和规则体系的方法，也有别于以非道德概念来定义道德概念的“自然主义”的方法，而是使证明依赖于整个观念、依赖于这个观念在何种程度上适合于在“反思的平衡”中“所考虑的判断”，并把这些判断组织成一个系统。证明是许多思考的互相支持，是所有因素的调协一致。罗尔斯正义论体系三部分的互相联贯、互相支持正如前述，第一部分是提出基本理论的结构，提出两个正义原则，第二部分是把这些原则应用于制度时与“被考虑的判断”对照，第三部分则考虑“公平的正义”理论在联系价值目标时的可行性，而每一部分之中的各个论点也是努力做到调协一致。

罗尔斯也引康德的哲学来加强对自己的正义理论的证明。他认为，对他的平等自由的正义观可以做一种康德式的解释，这个解释建立在康德的自律概念基础上，康德伦理学的意义和力量不在于它强调原则的一般性和普遍性，而是在于它把道德原则视作理性选择的目标，并设想道德原则将被作为自由和平等的理性存在的人们一致同意。当人选择作为这一本性的最准确表现的道德原则时，他就是在自律地行动着，他所遵循的原则之所以被选择，不是因为他的社会地位或自然禀赋，也不能用他生活在其中的特殊社会以及他恰好需要的特定事物来解释，如果那样的话，就是在他律地行动了。原初状态中相互冷淡的动机假设是符合康德的自律概念的，它体现了选择的自由，即不以某种特殊方式来限制他们的目的，不对人们的愿望施加任何先验的限制。原初状态中的人最后所选择的两个正义原则也就是康德意义上的绝对命令，因为对它们的选择不以任何特殊目的和愿望为转移。不管我们具体目的是什么，正义原则都适用于我们的意义，按照正义原则行动也就是按照绝对命令行动。这说明在推出作为绝对命令的正义原则时，没有以任何偶然因素为前提。

于是，原初状态可以看成是对康德的自律和绝对命令观念的一个程序性解释，而且，对原初状态的描述使我们能够解释这样一种意义：按照这些正义原则去行动，表现了我们作为平等、自由的理性人的本质，我们之所以想按照正义原则去行动，是因为我们想成为具有一种选择自由的平等、自由的理性的存在物。罗尔斯认为他的理论与康德的自律伦理学只有两点值得注意的不同：一是他把作为一个本体自我的个人选择为一个主体的选择；二是他始终假设原初状态中的各方知道他们是服从于人类生活的各种条件。这些不同看来不是实质性的，而只是因为罗尔斯在此考虑的不是一种普遍适用于理性存在的伦理哲学，而是一种人类社会的正义理论。

我们从以上叙述可以看出罗尔斯对洛克的常识理性和对康德的哲学思辨理性的回溯，但是，罗尔斯还是吸收了分析哲学的某些成果，他在构成其正义论体系时努力避免独断倾向，明确地声明自己只是在进行一种虚拟的深思熟虑的理性设计，他所理解的证明是整个理论的互相

支持、前后一贯，是可以不断容有修补甚至局部推翻的，他细致地进行逻辑和语言方面的推敲，仔细琢磨证明方式，确定自己的有限目标，对有些重大和根本的问题存而不论，划上括弧，尽量提出一些弱的假定，力求实行某些简化的手段而突出重点，尽量以明智审慎来代替价值判断——这些都反映出分析哲学的影响。

总之，正像我们在实质结论方面看到罗尔斯努力调和自由与平等的冲突一样，他在理性方法上也表现出一种综合的倾向，他重振理性的信心，试图为正义论建立一个理性的基础，为此而调和分析哲学与思辨哲学，调和常识理性与先验理性，并在某种程度上做得相当成功，也正是这种颇为成功的综合和调和的倾向，使他的正义论产生了很大的影响，引起了热烈的讨论和批评，使今天西方的社会政治哲学很难绕过罗尔斯的理论。

（本文系何怀宏先生《公平的正义——解读正义论》一书的第五章）

来源：<http://www.law-thinker.com/show.asp?id=1368>

/