

“原初状态”中的“人”并不是形而上学的抽象

——兼论罗尔斯思想实验的实践价值

李牧今*

摘要：“原初状态”中所包含的个人观在罗尔斯的正义理论中颇受争议。有批评者认为，原初状态只是一个形而上学的假设，它所包含的个人观过于抽象与空洞。事实上，作为一个思想实验，原初状态旨在对正义原则的产生情景进行高度概括和形象模拟。在罗尔斯看来，正义原则应该被表达为自由的、平等的、“理性的”、“合理的”个体在公平对等条件下的建构，是决策程序的产物，它并不基于抽象的个人观，而是由公民的规范观念所决定。原初状态中各派的行动既由个体的一般欲望所驱动，也受其道德能力中的最高阶利益和终极目的所主导，这种个人观既不缺乏动机要素，也不缺乏对日常合理性欲求的承认。原初状态的约束虽然是理论假设的，但其中所体现与反映的内容却是我们内心的确信，是沉淀于公共文化中的共识。此外，原初状态启发与帮助着我们每个人认清自身的根本认同与全面需求，在其中所达成的正义原则不仅具有应然性，同时也具有实然性。

关键词：原初状态；无知之幕；思想实验；罗尔斯

Abstract: The concept of person contained in “original position” is quite controversial in Rawls’ s theory of justice. Some critics believe that the original

* 李牧今，北京大学中国政治学研究中心博士后。

position is just a metaphysical hypothesis, and the conception of person in it is too abstract and empty. In fact, as a thought experiment, its original position aims to highly generalize and simulate the scene how justice principles emerge. According to Rawls, principles of justice should be the construction of free, equal, “reasonable” and “rational” individuals under fair and reciprocal conditions. It is a product of decision-making process, and is not based on abstract concept of person, but determined by citizens’ normative concepts. The actions of the various parties in original position are not only driven by individuals’ general desires, but also by the highest-level interests and ultimate goals of their moral abilities. This conception of person lacks neither motivational elements nor rational desires. Although the constraints of original position are theoretically hypothetical, the contents embodied and reflected in it are our inner convictions, which are also consensus rooted in public culture. In addition, original position inspires and helps each of us to recognize our own fundamental identity and overall needs. The principles of justice reached in it are not only with desirability, but also with feasibility.

Key words: Original position, Veil of ignorance, Thought experiment, Rawls

“原初状态” (original position) 既是罗尔斯最受关注的思想实验之一, 也是他理论中受到最多误读与批评的内容。社群主义者, 如迈克尔·桑德尔 (Michael Sandel)、查尔斯·泰勒 (Charles Taylor)、罗纳德·德沃金 (Ronald Dworkin) 等认为, 充满形而上学约束的原初状态不过是一个抽象的思想实验, 由它得出的正义原则和规范在现实生活中难以具有约束力与可行性。要回应类似质疑, 我们必须回到原初状态以及正义原则背后更深层的逻辑中去。众所周知, 罗尔斯理论的最终目标是建立“公平之正义” (justice as fairness)^① 的社会, 决定社会基本结构的正义原则由每个自由平等的个体在相应的决策程序与情景中推导得出, 而原初状态就是对决策过程的高度提

① Justice as fairness 比较常见的有两种译法, 一种为“作为公平的正义”, 一种为“公平之正义”, 本文采用“公平之正义”的译法。

炼和形象模拟。简单来说，正义原则在原初状态中是这样被推导出来的：各派（parties）被置于“无知之幕”（veil of ignorance）之后且处于彼此对等的环境中，他们进行商讨并最终得出管理社会的基本原则。这就是说，社会基本正义原则既不源于抽象的道德法则或者外在权威，也不源于上帝意志或者宗教教条，而是源于“人”，源于原初状态中的各派。换言之，只有找到一个恰当的、合适的人的观念，并且在公平的、对等的讨论、商谈环境中所得出的原则才能称之为正义原则，其正当性才具有说服力。罗尔斯声称，“公平之正义建立在这种根深蒂固的人性观念之上，是一种基于观念的（conception-based）理论”^①，这一核心思路事实上贯穿着罗尔斯的整个正义理论。原初状态就是这样一个启发我们用恰当思考自身角色的代表装置；理解原初状态中“人”，是走进罗尔斯思想内核的关键所在。

一、“无知之幕”与抽象的人

关于原初状态中的“人”，我们需要区分罗尔斯前期和后期的思想。在《正义论》中，罗尔斯声称公平之正义应该表达我们作为自由平等理性人的本质。^②但到了《政治自由主义》中，他已不再将公平之正义建立在一种具有先天道德本质的个人观念之上了，对于人的道德本质的阐释变为了关于政治身份的、关于公民的观念。由此，原初状态思想实验的目标也转变为建构人的政治身份，或者说旨在模拟个体的公民身份。这些都体现了罗尔斯的正义原则证成方法的基础性转变——从道德建构主义转变为了政治建构主义。

罗尔斯“公平之正义”的理念起始于对一个封闭社会的构想，它由终生参与社会合作的自由平等公民所组成。公平之正义所指的主要对象是社会基本结构（basic structure of society），而社会基本结构则通过原初状态中的各派（parties）所规定的社会合作公平条款而确立，当这些理念被充分实践的时候就能够产生良序社会（well-ordered society）。可以看出，推

^① 罗尔斯：《道德理论中的康德式建构主义》，陈肖生等译，吉林出版集团有限责任公司 2013 年版，第 453 页。

^② 罗尔斯：《正义论》，何怀宏等译，中国社会科学出版社 2009 年版，第 201 页。

导社会合作公平条款的过程是罗尔斯正义理论得以落实的核心步骤，它确立了社会基本结构，并且直接决定着良序社会能否实现。社会合作条款由每个个体在相应的决策程序与情景中推导得出，而要使这些条款体现出公平之正义的理念，粗略地说就至少要确保：（1）各派为合理公民（rational citizens）；（2）他们需要受到适当约束；（3）他们决策的慎思（deliberation）过程是在对等的情况下所做出的。显然这是一个对决策个人与决策环境都要求严苛的推导过程，而这正是罗尔斯引入原初状态思想实验的理由。罗尔斯旨在用一个形象的契约场景将一系列复杂的理论综合起来，并对正义原则的产生情景进行概括，“为一个可以得出合乎理性（reasonable）的正义原则的讨论设计宪法”^①，即最根本的讨论规则。罗尔斯希望原初状态能成为一个大家都能理解与明白，并且指导自己参与正义原则讨论的思想实验和行动指南。

具体来说，罗尔斯这样设想原初状态中的各派：一群纯粹合理的（purely rational）、相互冷淡的（mutually disinterested）人聚集在一起，他们将慎思（deliberate）一套原则来管理所处社会。他们是由自身利益（own advantage）驱动的合理个体，每个人都致力于推进该社会中某一群公民的特定目标；此外，这些代表在原初状态中处于对称的位置：他们拥有相同的知识、相同的影响力、相同的发表意见的机会。出于以上设想，罗尔斯对原初状态的决策过程进行了必要的设计与限制：（1）“无知之幕”用于保障原初状态中各派的地位平等；（2）基于对等性（reciprocity）理念的协商框架则体现出每个人自由平等的基本权利以及程序的公平性，使得相互承认的原则成为可能。罗尔斯用“无知之幕”、对等性原则以及其他相应约束取代了一系列复杂公式，并通过推导程序向我们展示了这样一个结论：任何合理公民在相应的约束条件下都会选择两个正义原则作为他们所处社会的基本准则。也就是说，罗尔斯并没有给出任何独立于这个过程之外的正当性标准，他对于这一决策程序的假设是“形式的”而非“实质的”，没有依赖任何特定的善观念或出自任何完备性学说的论断。罗尔斯在这里的想法是，“正义协议得以达

① Rawls, John, “For the Record,” *The Harvard Review of Philosophy*, Spring, 1991, <http://www.hcs.harvard.edu/hrp/issues/1991/Rawls.pdf>.

成的那个环境的公平性，将会传递到被一致同意的正义原则上”^①。通过原初状态的设计，正义原则被表达为一种公平条件下自由平等人的建构，道德原则因此成为决策程序的产物。

“无知之幕”是原初状态的核心设计之一，是罗尔斯塑造自律平等个体的重要环节。对于个体，“无知之幕”既可以“移除”特征，又可以“赋予”知识。所谓“移除”的意思是，人们只拥有那些从政治角度来看具有重要性的特征，而且这些特征足够普遍，任何人都可以拥有；“赋予”则意味着，人们可以被认为具有某些知识，例如关于政治事务或经济理论的基础知识。“无知之幕”界定了人们在政治身份中所具有的特征以及他们所知道的信息种类，它要求人们以特定的方式思考自身，即基于每个人的一般性特征做出选择。举一个例子来说明：在“无知之幕”背后的人知道每个人是有性别的，但却不知道自己是男性还是女性，也就是说他们只知道定性的通用特征，而不知道自身的具体情况。很多批评者，例如桑德尔在他的《自由主义与正义的局限》中指出，如此的“无知之幕”是一个形而上学的屏障，为了确保决策过程的公平性，它抹去了每一个个体身上的独有特征，而这就意味着，罗尔斯认为每个人身上的独有特征对于个人身份无足轻重。进而，批评推进为，原初状态论证本身依赖于抽象的个人主义观念^②，原初状态中所体现的形而上个人身份与人们对自我身份的一般认识不符。

为了澄清误解，我们需要明白，首先，罗尔斯在论证中所使用的前提是为了营造适当的慎思环境，其正当性并不基于抽象的个人观，而是由公民的规范观念所决定。这种规范性主要体现在两个方面：（1）原初状态中各代表所选择的行动原则必须要能够充分表达他作为自由平等个体的本质，也就是说他处于自治的状态，在自律地行动；（2）原初状态中各代表的决策不受他律知识和动机的影响。由于这样的设置借鉴了康德的自律概念，批评者们就指出，罗尔斯原初状态中的代表体现了不切实际的、与日常不符的个人观。

① 罗尔斯：《道德理论中的康德式建构主义》，陈肖生等译，吉林出版集团有限责任公司 2013 年版，第 350 页。

② Sandel, Michael J., *Liberalism and the Limits of Justice*, Cambridge: Cambridge University Press, 1982, p. 94.

但事实上，原初状态旨在模拟理性公民在某些约束和限制下的慎思过程，也就是说，原初状态并不是让我们在现有的意义上去真的忘记和剥离我们每个人的独有特征和身份，而只是用明确直观、形象生动的方式让我们“进入”一种慎思的情景和状态，以做出相应的选择和安排。罗尔斯用了莎士比亚的戏剧来打比方，“当我们以这种方式假装处于原初状态时，我们的推理就不再使我们承诺一种关于自我本性的特殊形上学说，而是去推理我们在一处戏剧中的角色，比如莎翁悲剧中的麦克白和麦克白夫人，我们真的是一位介入一场绝望的政治权力斗争中的国王或王后”^①。可见，各种对个人身份的所谓约束与限制也只是进行公平慎思的前提，罗尔斯并非旨在塑造一个“抽象的”或“形而上”的主体。原初状态旨在强调，如果在如此的情景下能够得出公平正义的原则，那么其他特征就不能进入考虑选择原则的慎思，并且这是正当的。

哈贝马斯认为“无知之幕”有不合理的“信息强制”，诺奇克则指责罗尔斯的理论是“非历史主义”的。^②但罗尔斯认为，理性多元（reasonable pluralism）是现代社会的根本事实，现代社会中已经不存在任何一种学说能够被所有公民接受，并在其基础上解决政治正义的基本问题。换言之，现代公民不再可能将某一种统合性学说当作可行的政治协议的基础。“作为以现代民主社会为基本研究对象的自由主义政治哲学，其理论立场必须是公共理性的。”^③公共理性（public reason）是罗尔斯给予共享推理形式（shared form of reasoning）的名称，通过公共理性，基本政治问题有了被每个理性公民接受的理由，进而政府权力的基础也就得到了公共证成（public justification）。公共理性所考虑的不是一般性意义上的个人行为，而是作为社会公民的个人及其社会性行为，是公民们建立社会合作的组织化社会化行为，这里所需要的，“恰恰是要摆脱纯粹个人的原始事实，建立一种为全体（在实际情形下常常只能是大多数或绝大多数）公民共享的（公共理性）出发点”^④。从这个意义上

① 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第25页。

② 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第600—604页。

③ 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第606页。

④ 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第25页。

说，原初状态为我们找回的是一个在多元社会中具有公共理性的个体，它可以作为现代人公共反思与自我澄清的工具与手段，帮助当代社会每个公民再次认清了自己身上本已具备但却有些含混不清的理念，使每个个体能够以自律的状态在对等协商中达成政治共识。

二、原初状态中“人”的特征和功能

下面，我们将具体分析原初状态中各派的特征以及对他们的限制，以了解罗尔斯如何设计达成正义原则的决策情境。

具体来说，罗尔斯对原初状态中的各派赋予了以下特征：处于“无知之幕”之后以及具有合理性（rationality）。^① 我们也可将其称之为原初状态中的“公民前提”，公民前提包括两个类别：关于公民特征的前提和关于公民偏好的前提。这些前提不对应于任何特定人的特征或偏好，而是与所有公民的特征和偏好相匹配，因为前提必须与它们要支持的结论具有相同的普遍性，个人不允许使用他们的实际知识或特征作为考虑因素，这会导致决策情境不公平。

“特征前提”是从原初状态的“无知之幕”中推出的，如前所述，这些前提阻止了人们的个人兴趣或欲望进入争论，“无知之幕”背后的人不了解他们的种族、性别、阶级、天赋等关于自身的特殊信息，他们有的只是关于社会的一般性知识。换言之，所有的前提都不涉及任何特定的个人或公民，而是涉及所有公民的事实。“特征前提”可以被总结为以下四点：（1）所有公民（受适当限制）都知道他们有一个合理的生活计划；（2）所有公民都知道他们是男性或女性；（3）所有公民（受适当限制）都知道他们有阶级位置；（4）所有公民（受适当限制）都希望生活在一个自己被公平对待的社会中。

公民前提的第二类是“偏好前提”，它主要包括两个内容：首先，描述公民对首要善品（primary goods）的期望；其次，公民有合理地制订自身生活计划的能力。罗尔斯通过首要善品赋予了每个公民最基本的需求，首要善

^① Campbell, Catherine Galko, *Persons, Identity and Political Theory: A Defense of Rawlsian Political Identity*, London: Springer, 2014, p. 88.

品涉及每个合理公民在社会生活中都想要并且也必须有的东西，它不关乎每个人特有的欲望。除此之外，罗尔斯认为每个人都有能力合理地安排与计划自己的生活，他能够权衡自己的欲望，判断轻重缓急并采取最好的手段达到目的，等等。^①

从以上约束中可以看出，原初状态第一是体现了各派自由的特征，他们每个人都能拥有自己的生活计划与善观念，并且不受外力干涉，他们是自己善观念的“自身根源”（self-originating sources）与“自证根源”（self-authenticating sources）。第二是平等的特征，没有人被赋予更多的权利或能力，并且“无知之幕”阻止了任何人在其自然或社会馈赠中占上风。除此之外，罗尔斯在其中还融入了公民的第三个和第四个品质——合理的（rational）和理性的（reasonable）能力。“合理的”是表达每个参与者合理益处的观念（expresses a conception of each participant's rational advantage），是罗尔斯直截了当地分配给原初状态中各派的动机，它们直接表现为实现自身善观念的欲望以及发展和运用道德能力的需求。“理性的”被定义为在公平合作条款中行事的意愿以及在这些公平条款中隐含的对等性理念，它通过原初状态中各派进行慎思时的“约束框架”来表示。罗尔斯表示，理性和道德可以兼容，理性人和道德人合二为人，成为同一个体——“道德-理性人”或“理性-道德人”，他们不仅追求物质利益，而且追求正义价值。^② 通过以上的结构设计，罗尔斯将他眼中政治个人观的主要四个特性：自由的、平等的、“理性的”、“合理的”，都表现在了一个特定的决策程序和代表装置中。

一部分罗尔斯的反对者，如泰勒指出，罗尔斯在原初状态中所挑选出来的特征就是他认为的对于每个人而言最重要和最本质的特征，而其他特征则是可以被忽略的。泰勒认为，通过置所有的外部约束与影响于一旁而达成自由的自我，是没有特性与缺乏确定目的的，真正的自由必须是“处境中的”（situated）。^③ 事实上，设立原初状态的最主要目的是为了能够找到治理

① Campbell, Catherine Galko, *Persons, Identity and Political Theory: A Defense of Rawlsian Political Identity*, London: Springer, 2014, pp. 88-91.

② 张国清：《正义原则的证明问题》，《华中师范大学学报》（人文社会科学版）2017年第3期。

③ 金里卡：《当代政治哲学》，刘莘译，上海译文出版社2011年版，第235页。

社会的公平协议，罗尔斯选择了一系列他认为与慎思社会安排原则有关的特征，但我们不必将这套特征视为一个人最重要、最本质的特征。原初状态只是一种探讨政治原则的方式，而建立原初状态的目的，是为了确保任何个人在使用假设决策程序时都能够做出所有理性公民的选择。换言之，原初状态旨在模拟与慎思正义原则相关的内容，是用于确定政治原则的论证环节，原初状态中代理人具有共享的公民特征，但并不代表它们个人最本质的特征。扬·纳维森（Jan Narveso）也相信原初状态并不需要某种个人身份理论。他写道：“（接受罗尔斯观点）并不意味着为了公正的目的，我必须跳出我的皮囊，或者必须是一个既不是你也不是我的幽灵般的抽象自我。我们需要的是原则，原则不仅仅是为了我们一个人的利益，而应该是相互吸引的原则，这些原则构成了相互作用的合乎理性的基础。”^①可见纳维森也认为，原初状态只是选择了与政治相关的理由，其他理由并不在罗尔斯的假设之内，因而指责其包含一种形而上的或者抽象的个人观念的反驳意见是不得要领的。这些反对看起来很有道理，似乎得到了我们直觉上的支持，因为我们很容易地就会想到，处于原初状态的人与我们自己不是一个人，进而对罗尔斯提出抨击，但却忽略了，罗尔斯只是想通过原初状态的假设来找到制订社会基本原则的恰当代理人。

罗尔斯希望原初状态作为一种思想实验能够使“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”^②。原初状态在个人进行选择时采取的必要约束旨在模拟一个可以推出类似康德定言命令的情景，或者说，原初状态是“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”^③。原初状态的约束虽然是假设的，但其中所体现与反映的内容却是我们内心的确信，是沉淀于公共文化中的共识。例如，现代社会中公民的确信之一，是占有某一特殊社会地位的事实不能被人们当成提出或期待别人接受一

① Narveson, Jan, “Critical Notice of Michael Sandel Liberalism and the Limits of Justice,” *Canadian Journal of Philosophy*, Vol. 17, No. 1, 1987, p. 229.

② 博格：《罗尔斯：生平与正义理论》，顾肃等译，中国人民大学出版社2010年版，第194—195页。

③ 博格：《罗尔斯：生平与正义理论》，顾肃等译，中国人民大学出版社2010年版，第195页。

种有利于这种社会地位和正义观念的充分理由。同样地，认可一种特殊宗教的、哲学的、道德的完备性学说以及与之相联系的善观念这一事实，也不能被人们当成提出或期待别人接受一种有利于这些说教的正义观念的理由。因此，要在原初状态中铸造这种确信，就不能允许各派了解他们所代表的那些人的社会地位，或各派所代表的那些人的特殊完备性学说。^①也就是说，不接受特殊地位和他人完备性善观念的强迫，是罗尔斯意义上民主社会之公民的普遍确信，而原初状态的约束旨在体现与保障这些确信。

三、原初状态中“人”的实践价值

罗尔斯认为，在人们相互提出和施加自己合理主张的情形中，必须要达到一种公平或均衡的状态，这就是正义产生的典型情景。原初状态是这一典型情景的一种启发性代表装置。在他看来，要对公平之正义的社会进行一种行之有效的生成性解释，就必须超越近代社会契约理论依赖于某一种特殊人性设定的思维框架，以此获得无先验立场或无特殊偏颇的逻辑起点。原初状态为我们理解这种生成性解释提供了一种轻巧且精准的方式。

第一，原初状态从技术性与操作性的角度简化与优化了解释路径，成了每个人都能够理解与使用的有效工具。罗尔斯说，“原初状态……帮助我们厘定我们的思想，一旦我们能够采取一种清晰而有条理的有关正义要求的观点，我们便会把社会设想为自由平等公民之间世代代坚持的合作图式”^②。面对现实生活的复杂性与我们个人身份与认同的多样性（甚至是模糊性），原初状态用明晰的清单与顺序，帮助我们有的放矢、思路清晰地思考社会治理原则问题。除此之外，“原初状态还作为一种中介理念而发挥作用，通过这一中介理念，才能使我们所有人认可的确信产生相互沟通……这使我们能够在我们的各种判断中确立更高的一致性，而通过这种更深刻的自我理解，我们就能达到相互间更广泛的一致”^③。罗尔斯还认为，原初状态既保障了各派

① 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第22页。

② 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第23—24页。

③ 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第23—24页。

是平等自由和不受压迫的，也营造了一个各派能够产生有效沟通的场域，人们不至于因为难以达成某些一致而使得谈判无法进行。

第二，原初状态对于理解和塑造公民的“合理自律”与“充分自律”至关重要。“公民的合理自律是在原初状态下通过作为其代表的各派的慎思方式来塑造的。与之相对，公民的充分自律则是通过原初状态的结构方面来塑造的。”^① 罗尔斯认为，如若在现实社会中，公民能够按照正义原则来行动，并且在正义感的指导下，认肯正义原则就是原初状态中各派所采用的原则，那么公民就达到了充分自律。充分自律意味着公民在其政治生活中能够对正义原则有公共认识且能明智运用。也就是说，从合理自律到充分自律经历了这样一个过程：首先，原初状态的思想实验确立了其中各派的选择，即正义原则，这个时候原初状态中的各派做到了合理自律。进一步，由于原初状态的约束具有正当性，其表达了我们公共文化中被大家普遍接受的深思熟虑的确信，体现了每个人自由平等的地位，因此，当我们揭开“无知之幕”并走出原初状态时，其中确立的正义原则依然具有正当性与权威性，依然需要被自觉遵从。换言之，原初状态中的约束对应与转换成了现实社会中每个公民的“理性的”能力、正义感的能力。良序社会中的公民会在公共生活中按照正义原则，按照原初状态的约束来行动，以实现充分自律。在这个意义上，所谓的充分自律之“充分”主要就是指，良序社会中的公民能够认识到规范社会基本结构的正义原则是他们在原初状态中选择出来的，并愿意按照这个正义原则来行动。^② 可以说，充分自律归属于日常生活中的公民，他们以某种方式看待他们自己，并且认可和践行会被一致同意的首要正义原则。^③

只要理解了这一点，我们也就能更深刻地体会，为什么罗尔斯说，虽然原初状态仅仅是一个思想实验，但却体现了民主社会公共文化中最为根本的确信。原初状态启发与帮助着我们每个人实现自身的根本认同与全面需求，在其中所达成的正义原则不仅具有应然性，同时也具有实然性。原初状

① 罗尔斯：《政治自由主义》，万俊人译，译林出版社 2009 年版，第 71 页。

② 谭清华：《政治哲学与政治哲学中的人》，《学术探索》2015 年第 11 期。

③ 罗尔斯：《道德理论中的康德式建构主义》，陈肖生等译，吉林出版集团有限责任公司 2013 年版，第 348 页。

态通过对称的结构设计保障了其中每个人的平等地位，通过“无知之幕”确保了每个人不受特殊善观念的影响，这些特征与民主社会中自由平等的人性观是相符合的。与此同时，罗尔斯将首要善品作为动机直接植入了原初状态中，首要善品作为一个人实现其目标的一般通用手段（all-purpose means），保障了之于所有人都至关重要的权利、自由、财富和尊严等。由此可见，不管是个人的正义感还是善观念都能够能够在原初状态中得到体现，其中的各派既是“理性的”，也是“合理的”，各派的行动既由个体的一般欲望所驱动，也受其道德能力中的最高阶利益和终极目的所主导。我们再次看到，罗尔斯在原初状态中设立的人的观念，并非一种抽象的、不现实观念，这种个人观既缺乏动机要素，也不缺乏对日常合理性欲求的承认。在他看来，对原初状态的推理限制类似于我们在自然演绎系统中使用的推理限制。在自然演绎系统中，我们在进行推导时需要进行必要限制，目的是为了保障，如果所有前提为真，那么最终结果也为真。同理，当我们在演绎推理中讨论社会治理原则时，我们不可能考虑所有形式的利益或事实。如果我们把自然演绎系统的规则视为一个过滤器，那么原初状态也类似这样一个过滤器；通过该过滤器，与慎思社会治理原则无关的信息与不合理的要素被过滤掉，罗尔斯所设想的两个正义原则得以呈现。基于公平之正义的理念，罗尔斯更新了社会契约学说，“它将公平的社会合作条款设想为那些介入社会合作的人一致同意的条款。但是，他们的一致同意和其他的一致同意一样，必须在合适的条件下达成。特别是，这些条件必须使自由而平等的个人处于公平的境地，绝不允许某些个人占他人的便宜。进而言之，必须排除武力威胁、强制、欺骗和欺诈”^①。因此，虽然表面上原初状态只是一个假设性的启发装置，但其约束条件却是达成公平合作协议的必经之路，故具有实践价值与效力。

四、结语

政治哲学与道德哲学、伦理学等最重要的不同之处在于，其理论既要具

^① 罗尔斯：《政治自由主义》，万俊人译，译林出版社2009年版，第20—21页。

有逻辑上的应然性，还要在现实社会中具有可行性与可欲性。或者说，政治哲学与道德哲学、伦理学的区分在于个体外在与内在的区分，而罗尔斯的理论正是这两者的统一。

正义原则要想得到全面证明，既要实现其在理论上的逻辑自恰，又要它完美地实践于社会基本制度。面对当代自由民主的理性多元社会，如何在包容与平衡诸多完备性学说的基础上，为社会建构一套公平正义的规范性政治理论；如何让个体在多元身份中找到自己的公共身份，并以自律且理性的态度做出决策等都是当代政治哲学的棘手问题。罗尔斯在复杂混沌的现实环境中，以政治正义作为标尺，在政治自由主义的框架下，为我们梳理出了建构“公平之正义”社会的理论构架与核心要素。在此之后，为了让自己晦涩的理论得以清晰呈现，也为了使抽象的理论成为每个人都能理解且指导自身行动的工具，罗尔斯设计了原初状态的思想实验。通过原初状态这一启发性代表装置，罗尔斯澄清了深藏于每个现代公民心中的根本自我认同与价值信念。一方面，原初状态用对称的结构保障了其中各派的平等，用“无知之幕”屏蔽了特殊决策的可能，用对等性理念与首要善品赋予了各派动机，进而在公平的决策程序中将正义原则变为了自由平等公民间的建构。另一方面，原初状态将植根于每个个体心中的自由、平等、自治、善观念等价值均涵盖其中，通过限定各自主张的边界、通过对话理解以及谈判妥协等方式在政治基本原则问题上达成“重叠共识”，因此，这些个体在走出原初状态后依然会遵循此前推导出的各项原则。原初状态的决策过程是罗尔斯正义原则的正当性基石，为现实社会中的政治整合与公民间合作打下了坚实基础。因此，原初状态并非一个抽象空洞的逻辑游戏，而是一个直面当代政治哲学理论困境、具有正当性和实践价值的理念。